

А. В. ПРАСОЛОВ

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ЭКОНОМИЧЕСКОЙ ДИНАМИКИ

Издание второе, исправленное

РЕКОМЕНДОВАНО
УМО вузов РФ по образованию
в области прикладных математики и физики
в качестве учебного пособия для студентов,
обучающихся по направлению подготовки
«Прикладные математика и физика»,
а также для студентов,
обучающихся по другим направлениям и специальностям
в области естественных и математических наук,
техники и технологии



САНКТ-ПЕТЕРБУРГ · МОСКВА · КРАСНОДАР
2015

ББК 65.050

П 70

Прасолов А. В.

П 70 Математические методы экономической динамики: Учебное пособие. — 2-е изд., испр. — СПб.: Издательство «Лань», 2015. — 352 с.: ил. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-0797-2

Учебное пособие содержит методы и модели экономической динамики, т. е. той части экономической теории, которая устанавливает причины изменений в экономике, основываясь на количественных оценках. Изложенный в книге материал описывает математические методы исследования линейных и нелинейных уравнений, анализ влияния временных лагов, задачи идентификации и прогнозирования. Учебное пособие предназначено для студентов, обучающихся по направлению «Прикладная математика», а также для экономистов, биологов и социологов.

ББК 65.050

Обложка

А. Ю. ЛАПШИН

© Издательство «Лань», 2015

© А. В. Прасолов, 2015

© Издательство «Лань»,

художественное оформление, 2015

ПРЕДИСЛОВИЕ

В настоящее время под математической моделью динамического процесса понимается такая зависимость между количественными характеристиками системы, которая позволяет делать прогноз поведения системы на ближайшее или отдаленное будущее. В экономике как в науке о производстве, распределении и потреблении благ, прогнозирование и управление параметрами системы приводит к рациональному принятию решений на всех уровнях общества: правительство, регион, фирма, семья или человек. Поэтому развитие математического моделирования весьма актуально. Однако, в отличие от инженерных наук, в экономике еще слаба формальная логика, отсутствует четкая система предположений (как, например, аксиоматический подход в математике), принятая большинством специалистов или подразумеваемая по умолчанию. Это приводит к уменьшению эффективности экономических исследований и сужению применимости результатов. Тем не менее, предположив некоторую идеализацию, подобрав соответствующие абстракции, мы сможем приближенно описать динамику экономической системы математическими уравнениями, отражающими причинно-следственные связи.

В этой книге так или иначе затрагиваются следующие экономические проблемы:

- прогнозирование обменного курса валюты;
- прогнозирование цены товара на товарной бирже;
- управление инвестициями (распределение капиталовложений);

- макроэкономические динамические модели;
- модели конкуренции и взаимодействия нескольких фирм на общем рынке или ситуация использования общих ограниченных производственных ресурсов;
- циклические колебания деловой активности;
- модели международной торговли;
- управление уровнем таможенного импортного тарифа;
- динамические модели рекламной деятельности и некоторые другие.

Часть из описанных задач рассмотрена полностью, т. е. предложен алгоритм решения (в некоторых случаях приведены и результаты моделирования), другие — лишь подвергнуты анализу или намечены пути составления модели. Это не пренебрежение законченными формами, а следствие иных задач, стоящих перед автором. Все объясняется стремлением расширить инструментальные возможности теории математического моделирования динамических процессов. Условно можно выделить в теории линейную и нелинейную части. В первой, широко используемой в настоящее время, имеются некоторые слабые места, помимо самой линейности. К ним относятся неопределенность в выборе длины идентификационного интервала, разнообразии критериев качества модели и, как следствие, невозможность считать одну модель лучше другой. Получается, что исследователь, предлагая модель динамики экономисту, говорит, что для данного набора наблюдений за динамикой исследуемого параметра, для класса линейных моделей и для заданного (как правило, в том или ином смысле квадратичного) критерия качества модели, последняя является оптимальной. Экономисту же не хочется вдаваться в подробности, а почему собственно линейные и квадратичные, и какие нужны для исследователя данные, ему необходима надежная модель, с помощью которой он бы правильно принимал решения, т. е. критерий оптимальности должен являться конечной прикладной целью моделирования. В книге обсуждаются эти и другие вопросы теории линейных моделей.

Однако основная цель автора связана с желанием включить в систему инструментальных средств теории модели-

рования логистические уравнения с запаздыванием по времени. В биологии они известны как уравнения Лотки–Вольтерры. Они давно привлекают к себе внимание (в мире издаются статьи и монографии с анализом их применения), но без учета временного лага, а также без алгоритмов идентификации коэффициентов их использование нельзя признать целесообразным. В книге показано, что применение моделей Лотки–Вольтерры вполне конструктивно, а динамическое разнообразие превосходит формы поведения траекторий линейных систем. Это касается более сложной структуры равновесного множества, колебательности и асимптотического поведения.

Временные лаги (задержки в реакции системы на изменение ситуации) в экономике также давно применяются в описании динамических проблем, но сложность математического аппарата теории уравнений с последствием не позволяет широко применять указанные модели. Однако полностью игнорировать временные лаги было бы ошибкой. Поэтому в книге дан некоторый анализ зависимости экономических выводов по результатам математического моделирования от учета временных лагов. Приведены многочисленные примеры экономических ситуаций с оценкой времени запаздывания, рассмотрена конкурентная деятельность фирм с задержкой в реакции и, что самое интересное, дан алгоритм совместной идентификации параметров системы и запаздывания. В приложении к книге собраны многие (но не все) факты теории дифференциальных систем с последствием, но это, конечно, адресовано скорее математику, чем экономисту. Большая часть материала публикуется впервые.

Структурно книга содержит четыре главы и три приложения. Первая глава имеет вводный характер. В ней собраны первичные определения теории моделирования, изложена точка зрения автора на совместимость качественных понятий математики и экономической динамики, обозначена историческая палитра одного из разделов экономической динамики — теории роста.

Вторая глава посвящена анализу временных рядов в широком смысле. Обычно в таких задачах выделяют две

составляющих временного ряда — детерминированную и случайную. В этой книге акцент делается на детерминированную часть зависимости, в то время как случайная составляющая оценивается каким-либо известным способом. Построение тренда временного ряда, как модели процесса, осуществляется тремя путями: непосредственной аппроксимацией, т. е. выбором из заданного класса функций представителя, наилучшим образом приближающего временной ряд, построением авторегрессионных моделей (а также с запаздыванием и с наличием интегральных многообразий) и построением моделей управляемых динамических процессов, т. е. когда модель строится по известному отклику системы на данное входное воздействие.

Третья глава содержит описание, анализ и экономическую интерпретацию моделей типа Лотки–Вольтерры. Исследованию подвергнуты как простейшие (двумерные) системы, так и многомерные. Приведена зависимость модели от временной задержки в коэффициенте воспроизводства. Исследованы качественные свойства решений. В частности, доказано существование предельных циклов, определены условия асимптотической устойчивости всех положений равновесия. Предложены новые алгоритмы идентификации моделей Лотки–Вольтерры, как без запаздывания, так и с запаздыванием.

В четвертой главе исследуются несколько конкретных моделей с использованием реальных данных: прогнозирование курсов валют для заключения долгосрочных форвардных сделок; модель взаимодействия государства, импортера и контрабандиста в международной торговле с определением оптимального импортного тарифа; конкуренция на автомобильном рынке США и модель реакции продаж на изменение объема рекламы некоторого нового продукта. Если первая из перечисленных задач в полной мере использует результаты и рекомендации первой главы, то вторая и третья модели основываются в своем построении на экономических особенностях задач, что придает им самостоятельное значение.

Первое приложение, как уже отмечалось, демонстрирует различные аспекты теории дифференциальных урав-

нений с последствием. Приведены те теоремы и описаны те методы, на которых, так или иначе, строятся рассуждения основных глав книги. Материал дан с подробными доказательствами, поскольку автору хотелось создать как можно более полное произведение, а не отсылать читателя к дополнительной литературе.

Второе приложение содержит описание известного способа определения оптимального размера производства относительно затрат и существующих цен на производственные факторы.

В третьем приложении приведена теория производственных функций, выделены присущие им свойства и дан обзор различных форм представления производственной функции. Рассмотрена задача влияния научно-технического прогресса на процесс производства.

Список литературы содержит только фактические ссылки, показывающие происхождение того или иного результата.

ГЛАВА ПЕРВАЯ

ВВОДНАЯ

1.1. О ПРИМЕНЕНИИ МАТЕМАТИЧЕСКИХ ИНСТРУМЕНТОВ В ЗАДАЧАХ ЭКОНОМИЧЕСКОЙ ДИНАМИКИ

Прежде чем перейти к обсуждению задач экономической динамики и математических методов, которые используются при их анализе, сделаем несколько замечаний, обозначив гносеологический контур последующих утверждений. Мы не склонны «обожествлять» математику, понимая, что она представляет собой более или менее «счастливое» сочетание некоторого количества аксиом, определений и правил преобразования с последующими выводами. В математике выводы признаются правильными, если преобразования выполнены правильно. Все «правильные выводы» лежат во множестве различных выводов, заранее определенных аксиомами и правилами. Таким образом, математика по своей природе не добавляет ничего нового, она лишь помогает увидеть то, что на первый взгляд незаметно в явлении, и этим она полезна в приложениях.

Например, если нашей задачей является оценить будущие расходы жителей города Санкт-Петербурга на покупку легковых автомобилей, мы собираем информацию о поведении этой величины в прошлом, составляем приближенную функцию времени, которая в некотором смысле «похожа» на изучаемый параметр, и используем ее для прогноза. Экономика была задействована только при постановке задачи, сборе информации и интерпретации результата. В середине процесса оказалась математика, которая как аксиому восприняла модель. Математика не отвечает за то, что сделала экономика, и тем более она не отвечает за промежуточные проблемы, например, из ка-

кого класса функций выбиралась модель и т. п. Предсказанное значение расходов на машины вероятнее всего не реализуется, может быть, оно окажется близко к тому, что будет наблюдаться в будущем. Однако выделить вклад каждого фактора в общую ошибку практически невозможно. Могли допускаться ошибки при сборе информации, сама методика сбора данных является предметом специальных исследований. Так что экономисты, передавая данные математику, закладывают в окончательный прогноз большую долю неопределенности. Специалист по моделированию, предположительно, владеет и экономическим анализом, и математическим аппаратом, но он также вносит погрешность в прогноз, отдавая предпочтения некоторому узкому классу моделей. Далее, математик «строго логично и правильно» трансформирует модель так, чтобы получить прогноз в виде, удобном для интерпретации. В нашем примере он мог бы звучать следующим образом: «Расходы на автомобиль возрастут за год на 33%». Но неопределенность, внесенная на первых двух этапах работы над прогнозом, вряд ли будет нейтрализована математической строгостью. Из сказанного выше следует, что в неправильных прогнозах не всегда виновен применяемый математический аппарат.

В экономической теории широкое распространение получили термины «технический» и «фундаментальный» анализ. Согласно первому прогноз получается на основе предшествующего поведения экономической системы с сохранением правил, которые это поведение определяют. Все регрессионные, авторегрессионные и прочие способы продолжить набор данных на будущее, несомненно, относятся к техническому анализу. Экономические, фундаментальные оттенки в нем создаются либо выбором множества параметров, вовлекаемых в анализ, либо классом функций, составляющих модель. Именно так мы намерены поступать при решении задач в данной книге. В чистом виде фундаментальный анализ базируется на балансовых соотношениях и рассмотрении равновесных точек (типа равновесия по Вальрасу). К сожалению, экономическая теория не располагает такой системой законов, как, например,

физика или химия, что приводит ее в большой степени к констатации зависимостей и отдельных характеристик поведения. Самое трудное, на наш взгляд, место в развитии фундаментального анализа состоит в отсутствии общепринятых правил абстрагирования: для всякого исследователя, что-либо утверждающего, найдется другой, который будет настаивать на учете еще какого-либо факта; спор будет не по существу результата, а по системе предположений, предваряющих исследование.

Тем не менее, экономическая теория широко использует математический аппарат, развитый для других наук, понимая, что он не вполне соответствует существованию экономических задач, что надо бы разрабатывать математические инструменты, более подходящие для концепций и понятий в экономике. Ниже мы покажем, что одно из основных понятий в теории динамических систем — неустойчивость — в некотором, определяемом ниже, смысле является обязательным явлением в экономической динамике, что неустойчивости мы обязаны, образно говоря, прогрессом и экономическим ростом, что именно неустойчивость определяет структуру производства, распределения и потребления.

1.1.1. ОПРЕДЕЛЕНИЕ НЕУСТОЙЧИВОСТИ В МАТЕМАТИКЕ И ЕГО ТРАНСФОРМАЦИЯ НА ЭКОНОМИЧЕСКУЮ ДИНАМИКУ

Неустойчивость изначально определялась как отрицание соответствующей устойчивости по Ляпунову. Последняя, в свою очередь, предполагала, что динамический процесс определяется на бесконечном интервале времени с помощью дифференциальных уравнений с достаточно хорошими правыми частями. Кроме того, использование дифференциальных и разностных уравнений подразумевает обратимость времени, т. е. решение может быть продолжено как с возрастанием времени, так и с убыванием. В дальнейшем определение устойчивости было расширено для произвольного функционального пространства и сформулировано как некоторый вид непрерывности решений по начальным данным на бесконечном интервале вре-

мени (здесь уже можно обойтись и без обратимости времени). Понятие неустойчивости по начальным данным может быть дано следующим образом.

Рассмотрим n -мерное евклидово пространство R^n , точку $X_0 \subseteq R^n$ и однопараметрическое отображение $X = X(t, X_0)$, где параметр $t \subseteq [0, T(X_0)]$, т. е. каждое отображение имеет свое собственное время существования $T(X_0)$. Не останавливаясь на точном в математическом смысле определении, будем говорить, что отображение $X = X(t, X_0)$ неустойчиво в $X_0 = Y$, если существует такое положительное ε , что для всех X_0 , сколь угодно близких к Y , найдутся моменты времени $s(X_0) < T(X_0)$ такие, что $\|X(s(X_0), X_0) - X(s(X_0), Y)\| = \varepsilon$.

Рассмотрим несколько доводов, которые ограничивают или вовсе препятствуют применению понятия неустойчивости в экономической динамике. Сначала вспомним об обратимости времени в динамических системах [72]. Одна из самых распространенных моделей макроэкономической динамики формализует изменение основного капитала с учетом амортизации и инвестиций:

$$\dot{K}(t) = -\beta K + I.$$

При возрастании времени смысл этого равенства состоит в уменьшении основных фондов за счет износа и в увеличении последних со скоростью капитальных вложений. Разумность этого баланса не вызывает сомнений, когда исследователь интересуется прогнозом от текущего состояния экономики K_0 в будущее. Однако проверять с помощью этой модели состояние системы в прошлом не представляется адекватным: при отсутствии инвестиций убывание времени влечет возрастание $K(t)$ до бесконечности, что вызывает сомнения. Таким образом, модель в виде дифференциальных или разностных уравнений, где возможно обращение времени вспять, не является универсальной, и все математические результаты должны быть перепроверены. Хотя на эти выводы можно взглянуть и с другой, неожиданной точки зрения. Допустим, что часть основных фондов, списанная на амортизацию, имеет некоторую ценность. Будем считать, что в анализируемую

систему включены все капиталы, которые мы до этого считали выбывшими. Тогда сумма списанных и несписанных основных фондов будет постоянна при нулевых инвестициях. Это говорит о том, что первоначально система не являлась замкнутой.

Несколько иначе обстоит дело с дифференциальными включениями и дифференциальными уравнениями с запаздывающим аргументом. В этих случаях обратимость времени невозможна, и выводы, основанные на математических преобразованиях, допускают прозрачную экономическую интерпретацию. В. И. Зубов [31] определил такие системы как общие и создал теорию с применением функционалов Ляпунова.

Вторым важным обстоятельством, почему определение неустойчивости динамических процессов в экономике следует изменить, является иное соотношение времени жизни процесса и его компонент. В небесной механике время наблюдения, прогнозирования, управления (иначе говоря, использования) переменных системы гораздо меньше, чем характерные времена системы: периоды орбит, время релаксации и другие. В экономике, прогнозируя поведение отрасли на год, два или десятилетие вперед, мы абстрагируемся от времени жизни того или иного продукта, от технологического времени производства, от банкротств, слияний предприятий, т. е. скачкообразных изменений структуры экономической системы. Таким образом, происходит некоторое сглаживание динамического процесса, замена реальной траектории на модельную. При этом свойства устойчивости (и неустойчивости) реальной динамической системы и ее модели могут оказаться различными, и, следовательно, само понятие неустойчивости в его традиционном смысле становится бесполезным.

Кроме того, большое значение имеет сильная нестационарность реального процесса. При производстве, распределении и потреблении товаров и услуг внешняя среда играет огромную роль. Ее составляют природные и социальные условия экономического процесса, политические и юридические изменения, колебания в предпочтениях, случайные всплески инноваций и т. п. Учесть все эти и мно-

гие другие возмущающие факторы не способна ни одна математическая модель, и поэтому понятие устойчивости при постоянно действующих возмущениях актуально в экономической динамике. Оно подразумевает, что малые постоянно действующие возмущения изменяют динамический процесс мало, и если некоторая мера возмущений стремится к нулю, то и отклонения процесса уменьшаются до нуля. Однако, если качественные результаты о поведении системы получить удастся (см. например, [31]), то количественные оценки — дело весьма трудоемкое и часто невозможное. Отрицание этого вида устойчивости означает, что возмущенный процесс обязательно отклонится на некоторое пороговое значение от невозмущенного, сколь бы мало возмущение не было, причем за время, меньшее времени использования модели.

Обобщая все сказанное выше и учитывая применение этих понятий в экономике, *будем называть неустойчивым такой процесс, который за время, меньшее его собственного времени жизни, превзойдет некоторый пороговый уровень, за которым данный процесс существенно меняется, превращается в другой.*

Очевидно, что трансформация понятия неустойчивости не отрицает своей первоначальной сути, но теперь мы смогли бы сосредоточить внимание на пороговых уровнях и мерах отклонения одного процесса от другого.

1.1.2. ДИССИПАЦИЯ

Под диссипацией в физике и механике понимается необратимая потеря энергии из динамической системы (иногда говорят, что энергия упорядоченного движения превращается в энергию неупорядоченного). Математические решения диссипативных систем, продолжимых неограниченно вправо, обязательно погружаются в некоторый шар за конечное время. В экономике в принципе не должно быть недиссипативных систем. Можно, конечно, найти монотонно возрастающие переменные: например, количество воды, выпитой частью населения страны, все время растет, и такая переменная в составе динамической модели

делает модель недиссипативной, но скорость поглощения воды более удобна для установления причинно-следственных связей, и ей присуща диссипативность. Этот пример можно рассмотреть более подробно. Здесь важно то, что экономические системы производят продукт из чего-то ограниченного по своим количествам. Кроме того, полезность любого продукта или услуги ограничена, а потому и не надо много продуктов. В силу ограниченной производительности за единицу времени слишком много товаров и услуг изготовить невозможно. Значит, если по какой-то причине в обществе стало много какого-то товара (больше, чем надо), то через некоторое время его количество уменьшится, поскольку никому не выгодно его производить.

Диссипативные процессы проистекают еще и от диффузии, т. е. если продукт производится для некоторой замкнутой системы, то всегда существуют юридические, институциональные, таможенные «щели», через которые продукт уходит из замкнутой системы. И чем его больше в системе, тем сильнее отток. Как правило, неконтролируемый отток товаров, услуг, финансов, производственных факторов и прочего вреден субъекту экономики, так как он определил замкнутость экономической системы именно для сохранения внутри системы ресурсов и продуктов экономической деятельности. Диффузия приводит к уменьшению прямых и косвенных налогов, нарушает статистический контроль, ухудшает управляемость экономикой, а также служит провокационной составляющей в коррупции и прямом воровстве. Это характерно как в макроэкономическом масштабе, так и на уровне фирмы. Но для нашей концепции важно то, что диссипативность в экономике больше, если диффузия больше. Отсюда как будто бы следует, что из отрицательного воздействия на экономику диффузии вытекает вредность диссипации. Однако есть примеры положительного эффекта диффузии.

В России в годы перестройки сложилась благоприятная ситуация с притоком дешевой рабочей силы из-за рубежа (в основном из ближнего зарубежья, Китая, Вьетнама, Турции и др.), но, к сожалению, ни законодательство, ни социальные и гигиенические условия проживания не

подготовлены к постоянному притоку работников низкой культуры и квалификации. Возникают социальные проблемы на местном уровне. Правительство же неспособно решить эту задачу позитивно. Таким образом, диффузия фактора труда действует положительно на экономическое развитие страны в целом.

1.1.3. ПОЛОЖЕНИЯ РАВНОВЕСИЯ

В математике равновесные состояния системы, т. е. когда все переменные системы длительное время сохраняют постоянные значения, хорошо изучены. Можно переместить начало координат системы в это равновесное состояние и следить за развитием процесса в очень малой окрестности нуля, где действуют линейные законы в причинно-следственных связях. Линейные системы хорошо известны и потому широко применяются. Это касается как вопросов устойчивости или неустойчивости, так и построения периодических и почти периодических траекторий, а также выделения подпространств, где устанавливается то или иное поведение.

Выводы, полученные с помощью линейных приближений, распространяются на исходный процесс (часто без должного обоснования) и затем используются для прогнозирования или управления.

Здесь мы хотели бы обратить внимание на то, что равновесное состояние в экономике никогда не реализуется, можно сказать, что оно вообще не существует. Точнее говоря, оно всегда неустойчиво. Строго доказать это утверждение невозможно, так как математическое доказательство требует точных предположений (и, следовательно, моделей), а экономическое доказательство, как правило, состоит из перечисления явлений и их логических интерпретаций. Но так как и явления можно преподносить по-разному, и логика у разных исследователей своя, то новые концепции частью экономистов принимаются, а частью — нет. Приведем ряд доводов, подтверждающих (как нам кажется) утверждение о неустойчивости любого положения равновесия, наблюдаемого в экономике.

Начнем с микроэкономического уровня. Предположим, что фирма производит товар в установившемся режиме: источники сырья стабильны, рабочая сила набрана в нужном количестве и качестве, внутреннее инвестирование покрывает амортизационные расходы, спрос на товар устойчив и т. д. Долго ли так будет продолжаться? Что прежде всего ограничивает время этой стабильности? Время жизни товара. Наступит момент, когда упадет спрос, и надо будет снижать цену, чтобы сбыть товар, но стабильность при этом нарушается, так как какие-то издержки производства останутся неоплаченными. Предположим, фирма, изменяя технологию, снижает расходную часть бюджета, но если бы это было так просто, то оптимальное поведение фирмы состояло бы в своевременных отчислениях на технологические инновации и замену старого товара на новый. На самом деле фирмы запаздывают с инновационными инвестициями, последние не всегда эффективны, да и замена одного товара на другой часто совершается с ошибками. Частично временные задержки и риски сглаживает доступность кредитов, выпуск акций под модернизацию и т. д. Однако некоторая колебательность все равно присутствует.

Другой вид колебательности связан с ценами на производственные факторы, а также на транспортные и складские услуги. Предложение факторов производства подчинено тем же законам, которые мы обсуждали выше, и цены факторов производства имеют тенденцию к росту при ненулевой инфляции, и также демонстрируют колебательность, в которой есть как периодическая составляющая (например, сезонность), так и случайные всплески. Эти возмущающие явления препятствуют стабилизации динамических процессов фирмы.

Есть еще один существенный факт, мешающий стабильности фирмы, — это конкуренция. Когда все ниши экономической жизни заполнены, то расширение фирмы и начало новой жизни (а это постоянный процесс) приводит к столкновению интересов либо за производственные факторы, либо за потребителя. В первом случае возрастает спрос на некоторый фактор и, как следствие, растет его

цена. Во втором — вслед за ростом предложения товара падает цена. И то и другое нарушает баланс рассматриваемой фирмы. Она начинает сопротивляться, на что также нужны средства, и либо выигрывает, уничтожая конкурентов разорением или поглощением, либо сама терпит поражение, либо устанавливается паритет с равновесием в другой, отличной от прошлого, точке. Описанный процесс происходит обязательно.

Приведенные доводы должны, как нам кажется, объяснить, почему длительное стационарное состояние для фирмы невозможно. Но руководителю фирмы оно, положение равновесия, и не нужно в стратегическом смысле. Контролируемое падение некоторых показателей — это вполне обычная ситуация динамики. Чем меньше фирма, тем колебания, в относительном измерении, могут быть больше. Маленькая фирма может на время даже полностью прекратить производство товара или услуги и пережить трудные времена. Большие фирмы для устойчивости и уменьшения колебаний применяют диверсификацию производства, подыскивая такие виды деятельности, которые взаимно дополняют друг друга, уменьшая общий риск банкротства.

В макроэкономической динамике, т. е. таких показателей, как ВВП, инфляция, уровень безработицы, бюджет, индекс и продолжительность жизни и другие, положение равновесия также всегда является неустойчивым. Даже если считать, что постоянный рост ВВП в модели будет заменен ежегодным приростом, и тем самым в качестве желательного равновесного состояния экономической системы рассмотреть стационарное (или при определенной системе отсчета — нулевое) состояние, то все равно в социально-экономической жизни стран важнейшую роль играют факторы, выводящие систему из равновесия. Приведем ряд таковых.

Если политическая структура страны позволяет кристику и демократические инструменты, то любое стационарное состояние будет подвергаться атакам со стороны оппозиции и критическому анализу. Если бюджет дефицитный (ради текущих задач), то правительство будет

виновно в случаях непредвиденных расходов и неумения создать резервы для смягчения рисков. Если же бюджет профицитный, то правительство будет виновно в том, что оно скрывает средства от своего народа из-за каких-либо виртуальных причин в будущем. Таким образом, для критики всегда есть место. Управленческими рычагами в макроэкономическом смысле являются налоги, льготы, государственное управление в производстве и другие. Правящие и оппозиционные группы (или в современном мире — партии) неизменно используют стабильность и равновесность в своих целях: правящие группы — чтобы удержать власть, а оппозиционные — чтобы забрать. Но если бы партии, претендующие на власть, только критиковали (есть, конечно, и такие), то они не получили бы поддержки социального большинства. Мы не рассматриваем насильственные действия социальных групп, хотя в нашем контексте это также важная составляющая доказательства. Оппозиция, явно или неявно, ведет поиск экономически более эффективного хозяйствования на макроэкономическом уровне. В процессе борьбы за власть правящая группа, услышав новые веяния, может перехватить инициативу и внедрить новые идеи. Относительно социально-экономического прогресса безразлично, какая именно группа будет у власти после очередного раунда борьбы, лишь бы новая идея вывела страну из состояния равновесия (или регресса, о чем мы позже поговорим особо).

Таким образом, высшее политическое звено управления экономикой в современном обществе организовано так, что рано или поздно срабатывает некоторый механизм, выводящий экономику из равновесия. Но так же было и в прежние века, когда в странах доминировал тоталитаризм. Основным путем к благу общества были войны, т. е. стремление отнять готовый продукт в виде богатств, рабов, земель и ресурсов. Выбор наиболее эффективного поведения осуществлялся лидирующей группой, но ею учитывались и различные сценарии. Страны вступали в союзы против общего противника, затем могли воевать между собой. Поражение могло стоить лидирующей группе потери власти. Однако страна, которая долго не воевала, теряла спо-

способность к обороне и терпела поражение, когда подвергалась нападению. В данном случае роль оппозиции играли воинствующие соседи, которые воспринимали стационарное экономическое поведение как упадок и, следовательно, как знак слабости в военных действиях. Заметим, что поговорка «бей своих, чтобы чужие боялись» в нашем контексте имеет вполне конструктивный оттенок. А именно, внутри стран между различными кланами и социальными группами также шли войны, которые требовали средств, умения и агрессивности. Победившая социальная группа с большой вероятностью начинала внешние войны.

Неустойчивостью характеризуется также и управленческий аппарат страны. Он всегда начинает рассыпаться изнутри, если его постоянно не обновлять. Любой начальник, во-первых, смертен, т. е. является управленцем конечное число лет. Во-вторых, он обрастает связями, обещаниями, долгами, иждивенцами и привычками, которые, в конечном итоге, препятствуют управленческой деятельности. Можно предположить, что время деятельности любого начальника должно ограничиваться временем действующей социально-экономической технологии локально, на месте данного руководителя. Так, у военных летчиков есть правило посылать на переподготовку для работы с новой техникой только молодежь. Однако внутри управленческих структур есть еще один важный фактор дестабилизации: всякий начальник снизу испытывает давление других управленцев. Среди всех стратегий его поведения особое значение имеют три: уйти и уступить место претенденту, уволить последнего и послать конкурента на повышение. Первое встречается крайне редко. Второе ведет к застою (равновесию). Третье — не у всех есть силы для такого хода. Но, тем не менее, управленческому аппарату, будь то высший правительственный уровень или руководство фирмой, свойственно консервироваться и взрываться резкой сменой кадров с некоторой периодичностью, что также может рассматриваться как факт неустойчивости. Если считать, что талантливая молодежь появляется случайным образом, то и данный факт пульсации управленческих пирамид является случайным.

Часть причин, выводящих макроэкономическую систему из равновесия, носит непрямой характер. Представим, что экономика страны некоторое время находится в состоянии равновесия, а раньше — не находилась. Но есть взаимосвязь между удаленными в прошлое событиями и настоящим. Это, например, могут быть демографические всплески и провалы, войны, крупные природные аномалии и прочее. Волны колебаний, хоть и затухая, будут ощущаться и в настоящем. Такие волны выведут систему из состояния равновесия. Реакция правительства будет демпфирующей при правильном решении или наоборот, правительственные решения могут раскачать волну, создавая резонансные явления.

В этом параграфе мы привели причины, по которым стационарное положение равновесия не будет устойчивым. Однако фактически эти причины обуславливают только возмущенность динамики, но не неустойчивость. Последняя будет следовать из того, что любые приведенные факторы действуют, в принципе, по одной и той же схеме: фактор накапливает энергию преобразования некоторое время и до некоторого уровня, после которого и совершается преобразование. Так, предположим, что экономическая система имеет положение равновесия (или равновесную динамику), причем такое, что избранный возмущающий фактор (пусть это будет политическая борьба) на него не действует. Рассмотрим некоторую другую экономическую систему, с очень близкими макроэкономическими параметрами, но имеющую слабое влияние политической борьбы. Политические противники, воспользовавшись ненулевой разностью экономик, превратят ее в предмет разногласий, и со временем борьба достигнет уровня смены правительства. Таким образом, сколь угодно малое отличие экономик, побудив к поиску аргументов и средств борьбы, приведет возмущающий фактор к уровню преобразований, т. е. к неустойчивости.

В микроэкономике динамическое поведение отдельной фирмы может иметь устойчивые стационарные колебания длительное время за счет своевременного изменения технологии, а также диверсификации. Но это свойственно толь-

ко производствам с постоянным спросом на продукт: металлургия, энергетика, часть сельхозпроизводства, ювелирное дело и т. д. Здесь время жизни продукта (можно говорить и об услугах: медицинское, образовательное, банковское обеспечение) сравнительно велико, и изменение технологий не приводит к замене одного продукта другим. Однако и над перечисленными выше производствами товаров и услуг довлеет конкуренция, происходят случайные оплошности (или удачи) руководства фирмой, природные или социальные катаклизмы. Легко представить себе два банка с близкими активами, стратегиями, кадровым составом и действующие в одном правовом поле. Их динамические траектории близки некоторое время, но разразился дефолт в августе 1998 г., и один из банков получил кредит Центрального банка, а второй — нет. Это повлекло разорение второго. Или другой вариант: первый делал ставку на Иран, а второй — на Ирак. В 2000 г. обе страны имели примерно равные риски для инвестиций, но американская политика в этом районе принесла преимущества первому банку, точнее, погубила второй. Таким образом, и микроэкономическая динамика демонстрирует неустойчивость траекторий.

1.1.4. АТТРАКТОРЫ

В результате приведенных рассуждений рисуется картина экономической динамики, напоминающая странный аттрактор, т. е. минимальное инвариантное множество в фазовом пространстве, заполненное неустойчивыми траекториями. Соседние с ним траектории притягиваются к этому множеству. Возможно, что в фазовом пространстве существует не один аттрактор, и тогда переход из области притяжения одного в область притяжения другого может интерпретироваться как структурные изменения в экономике страны (или микроэкономические превращения производств одного типа с возникновением другого, например, банкротство и слияние). Такие изменения происходили в России в 1917–1927 и в 1988–1993 гг.

Всего 100 лет назад в России было широкое производство предметов для гужевого транспорта и валенок, чего

в настоящее время практически не сыскать. Если бы мы составили математическую модель производства валенок на данных по России за период 1900–1930 гг., то получили бы некоторый аттрактор в пространстве переменных: шерсть, труд, налоги, цены, госзаказ и т. п. По этой модели трудно было бы прогнозировать состояние с обувью в 2000 г., когда изменения моды и появление новых материалов и технологий практически уничтожили спрос на валенки. Новая модель обувного производства (аналогичного валенкам) должна включать переменные, отражающие использование продукции химических предприятий, таких, как клей, кожезаменители и пр., а также отражающие новейшие достижения дизайна. Следовательно, при таких преобразованиях меняется само фазовое пространство: в 1900 г. были одни переменные, описывающие состояние системы, а в 2000 г. стали другие. Здесь мы не можем говорить о переходе траектории из области притяжения одного аттрактора в область притяжения другого. Значит, появляется понятие «время жизни фазового пространства», и устойчивость или неустойчивость динамики экономического состояния необходимо увязывать с этим понятием, что существенно затрудняет математическое моделирование.

Заметим, что подобные рассуждения имеют место и в естественных науках: в физике нет абсолютно замкнутых систем, все внешние возмущения учесть не удастся практически никогда и т. п.; химические, биологические и инженерные науки также исходят из приближенного моделирования реальных процессов, которые, кстати, также необратимы и не вечны. Существенная разница самоорганизующихся систем заключается именно в том, что любое равновесие в экономике воспринимается как пустота, которую надо срочно заполнить. Приближаясь к равновесному состоянию, система позволяет за небольшую цену, затратив немного энергии, «перескочить» на другую траекторию, приносящую в будущем значительные выгоды. В неживой природе подобные стремления вообще отсутствуют, в живой природе — наблюдаются, например естественный отбор более приспособленных особей внутри вида,

но в этом случае изменения требуют длительного процесса, гораздо длинней промежутка моделирования. Социально-экономическая динамика обладает большой реакцией и возможностями в короткий срок получить значительный эффект. Таким образом, экономические процессы, обладающие меньшей инертностью, выходят из состояния равновесия за меньшее время.

В заключении размышления о понятии неустойчивости в экономической динамике отметим положительное значение последнего с общей исторической точки зрения.

1.2. О МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ В ЭКОНОМИЧЕСКОЙ ДИНАМИКЕ

Этот параграф объединяет понятия теории моделирования, которые при интуитивном восприятии кажутся простыми, однако для построения формальной модели лучше лишний раз уточнить их смысл.

1.2.1. НЕКОТОРЫЕ ПОНЯТИЯ ТЕОРИИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Первым таким понятием является «система». Очень часто понятие «система» используется в интуитивной форме, исследователи не определяют, что это такое. Здесь мы условимся считать, что *системой называется множество объектов произвольной природы, связанных между собой возможностью взаимного влияния.*

Может быть, действительно, само понятие системы не важно, однако, когда составляется модель экономических процессов, нужно понять, что является эндогенными и экзогенными (внутренними и внешними) параметрами и что такое замкнутая система. Например, в некоторых задачах производства экзогенными параметрами являются цены, по которым продается продукция, налоги, спрос и предложение сырья, капитала, труда и многое другое, что мы потребляем извне для производства; эндогенными являются количество энергии, требуемой для технологического процесса, количество труда и величина зарплаты,

использованные в производстве, и тому подобное, что мы устанавливаем сами. Условное разделение факторов, влияющих на объекты системы, зависит от постановки задачи и опыта исследователя, т. е. это довольно субъективное решение.

Когда составляется математическая модель, надо определить, что такое *объекты системы* и как они связаны внутри системы между собой. Рассмотрим, например, какие объекты могут объединяться в систему, если решается какая-либо макроэкономическая задача. Макропараметры — это валовой внутренний продукт (ВВП), налоги, уровень жизни и т. п. Что же в задачах макроэкономики является системой? Это народ, все источники сырья, технологии, основные фонды и т. д. Влияет ли динамика народонаселения на ВВП? Да, конечно. Чем больше народа участвует в производстве, тем больше выпуск продукции; чем выше покупательная способность, тем выше спрос; чем больше рождаемость, тем больше требуется рабочих мест и т. д. Динамика и распределение народонаселения в окрестности размещения основных фондов — это влияющие друг на друга объекты. Так, например, если возрастает продолжительность жизни людей, то работающему населению надо повысить производительность труда, чтобы прокормить неработающих. А если этого нет, то данную страну ожидает экономический кризис.

Каждому объекту в системе соответствуют некоторые функционалы, которые называются *переменными состояниями системы*. Если говорить о технических объектах, то все понятно (океан—корабль составляют систему: океан выталкивает корабль до какого-то уровня и создает волны, корабль движется по океану и имеет шесть степеней свободы). В экономике, если рассматривать микроэкономические проблемы одного предприятия, параметры, которыми описывается предприятие, будут такие: стоимость входа (все, что взяли снаружи, чтобы произвести одну единицу продукции), количество единиц продукции, произведенных за единицу времени. Если предприятие более эффективно, то за единицу времени выпускается больше единиц продукции или единица продукции стоит меньше.

Переменные, описывающие объект, — это функция той задачи, которую мы решаем. Главная цель человека, решающего данную задачу, заключается в том, что он должен правильно подобрать эти переменные.

Переменные, которые мы выбираем для описания объекта, должны удовлетворять таким свойствам:

- *изоморфизм* — если эти переменные не меняются, то считается, что и объект не меняется (взаимно-однозначное соответствие);
- *свойство полноты* — любое состояние системы представляется через минимальное семейство переменных; если из семейства можно изъять какую-нибудь переменную, не меняя при этом изоморфизма, то семейство уменьшится, оставаясь при этом полным. Таким образом, можно минимизировать количество данных, описывающих состояние объекта, сохраняя полноту описания.

Далее мы неоднократно будем пользоваться экономическими терминами, в самом определении которых просматривается зависимость от других. Например, рассмотрим некоторые определения [12].

«*Внутренний валовой продукт (ВВП)* — макроэкономический показатель статистики народного хозяйства. ВВП рассчитывается как совокупная стоимость конечной продукции отраслей материального производства и сферы услуг независимо от национальной принадлежности фирм, расположенных на территории данного государства. ВВП исчисляется тремя способами:

- по производству — как сумма условно чистой продукции (или добавленной стоимости) всех отраслей материального производства и сферы услуг, включая весь экспорт товаров и услуг;
- по распределению — как сумма государственного и частного потребления, государственных и частных инвестиций, прироста или уменьшения всех видов запаса или экспорта;
- по доходам — как сумма заработной платы наемных работников, всех видов прибыли, рентных доходов, амортизационных отчислений и косвенных налогов.

ВВП может быть рассчитан в текущих или неизменно сопоставимых ценах».

«*Индекс стоимости жизни (ИСЖ)* — это исчисляемый индекс цен и тарифов фиксированного набора товаров и услуг, входящих в потребление соответствующих категорий населения. Его функция — выразить динамику рыночной стоимости основных цен на товары и тарифов на услуги. При исчислении ИСЖ набор товаров и услуг постоянно увеличивается. Среди них есть цены на продукты питания, одежду, топливо и другие расходы (образование, медицинское обслуживание и т. д.)».

Очевидно, и ВВП, и ИСЖ связаны с инфляционными процессами (стоимостью денег), с демографическими проблемами, с безработицей и т. д. В определении ИСЖ количество услуг и товаров постоянно увеличивается, поэтому нельзя сравнивать стоимость жизни на больших интервалах времени, например в 1913 г. и в нынешнем. ИСЖ — характеристика экономики, но ее очень трудно применять, поэтому вводят понятие *изменение ИСЖ*, и оно должно быть положительным. Более активным является параметр продолжительности жизни.

Определим, что такое *структура системы*. Отдельные объекты системы могут по-разному влиять друг на друга. Это влияние может меняться с течением времени. Так, бюджетные инвестиционные потоки зависят от политических предпочтений, международной конъюнктуры и других нестационарных факторов. Поэтому назовем *структурой системы граф взаимного влияния объектов между собой*.

Рассмотрим пример структуры [71]. Простейшей замкнутой экономической моделью может служить однопродуктовая динамическая макроэкономическая модель. На рисунке 1 показано, как природные ресурсы и труд извне поступают в производственный технологический цикл и, объединяясь с основными производственными фондами, превращаются в валовой продукт. Он, в свою очередь, частью возвращается в производство (в виде возврата по кредитам или оборотного капитала и т. д.), а частью составляет конечный продукт, который делится на непродуц-

ЭКОНОМИКА

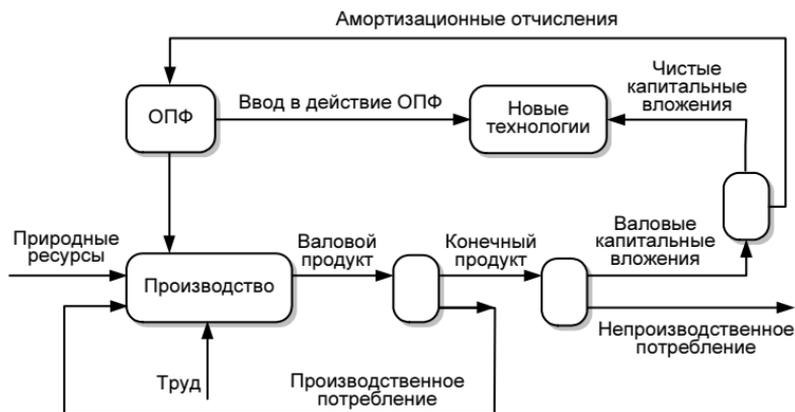


Рис. 1
Пример структуры экономической системы

венное потребление и валовые капитальные вложения. Последние используются на поддержание и развитие основных производственных фондов.

Математическая модель рассматривает зависимость между переменными состояния системы. Модели динамических процессов рассматриваются при постоянной структуре. Математические модели составляют из балансовых соотношений, то есть рассматривают систему как замкнутую и считают, что в нее попадает и как дальше распределяется. Так, например, если у заводского управления есть инвестиции, то их можно распределить на три части: дать цеху 1, цеху 2 и складу:

$$I = I_1 + I_2 + I_{st} + \delta,$$

где δ — сложно учитываемая добавка (могут быть управленческие затраты). Если будем считать в человеко-часах людей, занятых в производстве (получающих зарплату), то получим:

$$L = L_1 + L_2 + L_{st} + L_c.$$

Здесь индекс *st* обозначает склад, а *c* — бухгалтерию. Два последних соотношения являются элементами математической модели. Целая модель чаще всего бывает громоздкой, поэтому предусматривается возможность ее

агрегирования (группирование частей в единое целое) или декомпозиции (представления целого как независимых частей). И та и другая процедуры формально изменяют модель. Фактически мы от одной модели переходим к другой, более маленькой в случае декомпозиции. Будут ли применимы выводы, которые мы делаем для меньших систем, ко всей системе? Нет, не всегда.

Рассмотрим систему:

$$\dot{x} = Ax + Bu.$$

Предположим, что вектор переменных и вектор управлений состоят из трех частей: $x = (x_1, x_2, x_3)$, $u = (u_1, u_2, u_3)$. Декомпозиция в данном случае есть разделение на три подсистемы:

$$\dot{x} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \times \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

Математически заменяем одну систему на несколько других путем выбрасывания внедиагональных элементов. Получаем:

$$\begin{aligned} \dot{x}_1 &= A_{11}x_1 + B_{11}u_1; \\ \dot{x}_2 &= A_{22}x_2 + B_{22}u_2; \\ \dot{x}_3 &= A_{33}x_3 + B_{33}u_3. \end{aligned}$$

Если мы захотим перевести решение этой системы в нуль и для этого найдем u_1, u_2, u_3 , переводящие в нуль, то вся система в нуль не обратится из-за этих внедиагональных элементов, которые весьма существенны в рассмотрении.

В теории систем существуют примеры, когда управления, оптимальные для отдельных частей системы, в объединенном состоянии не выполняют своей функции, т. е. система в целом управляется не только неоптимально, но и недопустимо. Рассмотрим такую ситуацию.

Пусть, например, требуется решить задачу стабилизации для линейной стационарной системы

$$\dot{X} = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 2 \end{pmatrix} X + \begin{pmatrix} -1 & 0,5 \\ 0,5 & -3 \end{pmatrix} U.$$

Воспользуемся тем, что диагональные элементы матриц коэффициентов существенно больше по абсолютной величине внедиагональных, и продедем декомпозицию. Получим два независимых управляемых уравнения:

$$\dot{x}_1 = x_1 - u_1, \quad \dot{x}_2 = 2x_2 - 3u_2.$$

Для последних уравнений задача стабилизации решается управлениями $u_1 = 2x_1$, $u_2 = x_2$, а для функционалов

$$J_1(u_1) = \int_0^{\infty} (2x_1 - u_1)^2 dt, \quad J_2(u_2) = \int_0^{\infty} (x_2 - u_2)^2 dt,$$

задача стабилизации решается к тому же и оптимально. Однако, подставляя найденные оптимальные стабилизирующие управления в исходную общую систему, получим:

$$\dot{X} = \begin{pmatrix} -1 & 1 \\ 1,5 & -1 \end{pmatrix} X.$$

Характеристическое уравнение матрицы коэффициентов этой системы имеет вид $\lambda^2 + 2\lambda - 0,5 = 0$. Один из корней этого уравнения положителен и, значит, найденные управления не только неоптимальны, но и не решают задачу стабилизации.

1.2.2. ПОНЯТИЕ АДЕКВАТНОСТИ МАТЕМАТИЧЕСКОЙ МОДЕЛИ

Каким бы способом мы ни составляли модель, возникают проблемы ошибок моделирования. Ошибки бывают нескольких видов: ошибки данных, из которых строится модель, и ошибки неправильного подбора самой модели. В первом случае необходимо каким-либо статистическим приемом очертить область гарантированного решения задачи. Чтобы ее расширить, можно уточнять измерения или увеличивать их количество в зависимости от характера задачи. Во втором случае — варьировать подходы к моделированию. Но и на том и на другом пути необходимо уметь сравнивать результат моделирования с наблюдаемыми фактами. Если сравнение удовлетворительно, то говорят, что модель адекватна.

Пусть последовательность $\{x_i(t)\}$ описывает поведение системы. Пусть мы построили математическую модель и получили модельные переменные $\{y_i(t)\}$. Тогда можем ввести некую метрику расстояния x от y , т. е. $\|y_i(t) - x_i(t)\| \forall t, i$. Модель $\{y_i(t)\}$ адекватна реальному динамическому процессу $\{x_i(t)\}$, если для некоторого заданного ε выполняется неравенство: $\|y(t) - x(t)\| \leq \varepsilon$.

Рассмотрим пример. Пусть предсказывается курс доллара. Допустим, что мы придумали модель курса, нашли кредитора и брокера, который рискнул нам помочь опробовать модель в валютных спекуляциях. В общей сложности было сделано 100 экспериментов, и мы в 10 раз подняли наш исходный капитал. В этом случае брокер обязательно оценит наш алгоритм (модель) как весьма адекватный. Чтобы так удачно играть, надо каждый раз выигрывать чуть более 2%. Следовательно, относительная ошибка нашего алгоритма должна быть на 2% меньше средней относительной ошибки всех других игроков. Таким образом, можно составить указанную выше метрику и погрешность ε . Уже здесь видно, что без теории вероятности нам не обойтись, так как даже в простых примерах, где сравниваются измеряемые величины, присутствуют случайные погрешности. В экономических задачах присутствуют не только ошибки измерения, но и влияния неизвестных факторов, а значит, и понятие адекватности должно содержать вероятностные характеристики. В дальнейшем мы продолжим обсуждение понятия адекватности для конкретных прогнозных моделей. Отметим только, что критерий адекватности должен быть функциональным для каждой задачи, т. е. отражать конечный результат задачи.

1.2.3. СПОСОБЫ СОСТАВЛЕНИЯ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

В отличие от естественных наук в экономике не так много признанных учеными фактов, которые можно назвать законами. Такие фундаментальные утверждения естествознания, как закон сохранения энергии или закон сохранения масс, без сомнения, переносятся в экономику как балансовые соотношения. Но ведь, кроме того, в этих

законах большое значение имеет понятие замкнутой системы, да и само определение того, что сохраняется, в экономике весьма расплывчато. Примером сравнительно простого балансового равенства является статическая межотраслевая модель Леонтьева [36].

Распределение продукции каждой отрасли в межотраслевом балансе экономики страны описывается следующим соотношением:

$$x_i = \sum_{j=1}^n x_{ij} + y_i, \quad i = 1, \dots, n,$$

где x_i — валовая продукция в рублевом исчислении, произведенная i -й отраслью; x_{ij} — поставки i -й отрасли в j -ю отрасль в текущем году; y_i — конечное потребление i -й отрасли за текущий год.

Обычно межотраслевая модель строится на следующих предпосылках:

- в каждой отрасли имеется единственная технология производства;
- нормы производственных затрат не зависят от объема выпускаемой продукции;
- не допускаются замещения в производстве одних видов продукции другими.

При этих предположениях величина x_{ij} для всех i и j должна иметь следующий вид:

$$x_{ij} = a_{ij}x_j, \quad i = 1, \dots, n; \quad j = 1, \dots, n.$$

Величина a_{ij} называется коэффициентом прямых затрат. Этот коэффициент показывает, какое количество продукта i -й отрасли надо затратить на производство j -й отрасли. В матричной форме получаем соотношение:

$$X = AX + Y,$$

называемое балансом распределения продукции. Для межотраслевого баланса хорошо разработаны методики вычисления величин, входящих в модель, их количество, единицы измерения и т. д. Существуют, конечно, неучтенные продукты отраслей или контрабандный ввоз-вывоз продукции, но в первом приближении такой анализ считается вполне адекватным.

Примером динамического балансового соотношения может служить модель динамики основных фондов с учетом амортизации:

$$\dot{K} = -\beta K(t) + I(t),$$

где β — коэффициент старения, $K(t)$ — объем основных фондов в момент времени t и $I(t)$ — текущие инвестиции на поддержание и развитие основных фондов.

Последнее балансовое соотношение весьма приближенно, так как ни первое, ни второе слагаемые не соответствуют реальности. Например, первое слагаемое предполагает, что износ основных фондов пропорционален их количеству, но такое предположение требует огромного количества дополнительных посылок: всякая часть основных фондов стареет и выбывает из производства с той же скоростью, пополнение производится равномерно и элементами того же качества и т. д. Второе слагаемое, очевидно, должно быть нелинейно и содержать другой временной аргумент, так как не может некоторый объем денег мгновенно превратиться в работающие основные фонды. Кроме того, в разных задачах по-разному определяется замкнутая система и, как следствие, объем основных фондов: в макроэкономической проблематике не ясно, включать или не включать в объем предприятия иностранного капитала и предприятия, функционирующие за рубежом, но принадлежащие отечественным собственникам; в микромоделях производитель может арендовать фонды, выплачивая только арендную плату, которая вносится в издержки и, следовательно, относится к оборотному капиталу (конечно, арендодатель включит в арендную плату отчисления на покрытие износа, но это уже другое предприятие с другой моделью).

В качестве законов в экономике рассматривают стремление к равновесному состоянию между конкурирующими сторонами в процессе заключения тех или иных сделок. Например, вычисление цены товара происходит в точке пересечения кривых спроса и предложения или оптимальное по прибыли использование ресурсов определяется точкой касания изокванты (линии одинакового выпуска продук-

ции в стоимостном выражении при различных затратах ресурсов) и изокосты (линии одинаковых затрат).

Регрессионный метод. Описанные выше соотношения между переменными состояния системы основывались на некоторых теоретических соображениях. Представленные в аналитической форме, они, как правило, содержат некоторые неизвестные параметры, выбор которых нуждается в специальном рассмотрении. Иногда аналитическая зависимость, выступающая в качестве модели, просто постулируется. В обоих случаях для идентификации неизвестных параметров привлекаются данные наблюдения за процессом (или многократно повторяемого статического эксперимента). Формально математически регрессионный метод моделирования описывается следующим образом (в книге [28] этот метод назван «наивным»): пусть задан класс функций $\{f(X, A)\}$, где $f \in R^l$, $X \in R^m$, $A \in R^n$, зависящий от векторного параметра A . Пусть для некоторого множества аргументов $\Omega = \{X\}$ наблюдаются векторные значения модели $\{\bar{Y} = Y(\bar{X}), \bar{X} \in \Omega\}$. Требуется найти такой вектор A , чтобы достигался минимум некоторого наперед заданного функционала (нормы):

$$A = \arg \min_A \max_{\bar{X} \in \Omega} \|\bar{Y} - f(\bar{X}, A)\|.$$

В простейшем случае, когда в качестве функционала оказывается сумма квадратов отклонений наблюдаемых данных от модельных, а вектор параметров входит в f линейно, получаем задачу линейной регрессии с оценкой методом наименьших квадратов. Тогда модель динамического процесса выбирается из заданного класса функций при легко вычисляемом векторе A .

Система «вход–выход». Такими системами в теории управления называют модели, у которых имеется многомерный «вход», т. е. несколько таких переменных, которым могут быть присвоены произвольные (управляющие) значения, многомерный «выход», т. е. несколько переменных доступных наблюдениям (измерениям) в текущий момент времени, а также предполагается, что состояние системы описывается некоторыми внутренними переменными, недоступными управлению или наблюдению. Таким

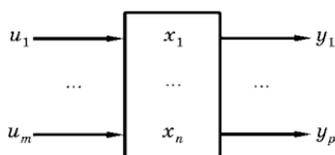


Рис. 2
Блок-схема системы
«вход-выход»

образом, «вход» влияет на внутренние переменные, а они, в свою очередь, определяют «выход» системы (рис. 2).

Зная вход $u(t)$ и выход $y(t)$, определить матрицы A , B и C возможно не всегда, а когда возможно, то не единственным

способом. Для анализа этого применяется теория минимальной реализации Калмана [39] (и ее частный случай — алгоритм Хосу).

Привлекательность таких моделей в экономике несомненна, однако специфический характер «входа» не позволяет применять инженерные методики восстановления модели по наблюдениям за «выходом». Необходима трансформация методов, гарантирующих идентификацию по любым известным «входам» и «выходам».

1.2.4. ЭКСТРАПОЛЯЦИЯ

Определим понятие экстраполяции в простейшем случае скалярной функции скалярного аргумента, полагая, что обобщение на произвольные операторы может быть получено без особого труда.

Определение. Пусть функция $\varphi(x)$ задана на промежутке $[a, b]$, а функция $f(x)$ — на промежутке $[\alpha, \beta] \supset [a, b]$. Если $f(x) = \varphi(x)$ на $[a, b]$ и на функцию $f(x)$ распространяется какое-либо качественное свойство функции φ , то $f(x)$ называется *экстраполирующей* функцией по отношению к $\varphi(x)$, а процесс построения $f(x)$ называется *экстраполяцией*.

Рассмотрим теперь некоторый процесс, т. е. вектор переменных состояния системы, изменяющийся во времени: $X(t)$, $t \in [t_1, t_2]$.

Определение. *Прогнозом* называется экстраполяция $X(t)$ при $t > t_2$ по информации на промежутке $[t_1, t_2]$.

Заметим, что экстраполяция осуществляется по данной функции $\varphi(x)$, где $x \in [a, b]$, и не требуется выяснения, почему здесь получилась такая функция $\varphi(x)$. В этом смыс-

ле экстраполяция является локальной математической моделью. Теперь перейдем к описанию некоторых способов прогнозирования.

Сначала остановимся на сохранении свойств гладкости. Пусть функция $\varphi(x)$ на промежутке $[a, b]$ n раз непрерывно дифференцируема. Тогда и $f(x)$ на $[a, b]$ n раз непрерывно дифференцируема. Потребуем сохранения этого свойства на промежутке $[b, \beta]$. Из высказанных выше предположений допустимо представление функции $f(x)$ в окрестности $x = b$:

$$f(x) = f(b) + f'(b)(x - b) + \dots + f^{(n-1)}(b) \frac{(x - b)^{n-1}}{(n-1)!} + f^{(n)}(\xi) \frac{(x - b)^n}{n!}, \quad (1)$$

где $\xi \in (b, x)$, $x \in [b, \beta]$.

В силу определения экстраполяции $\varphi^{(k)}(b) = f^{(k)}(b)$ при $k = 0, 1, \dots, n - 1$. Следовательно,

$$f(x) \approx \varphi(b) + \varphi'(b)(x - b) + \dots + \varphi^{(n-1)}(b) \frac{(x - b)^{n-1}}{(n-1)!}. \quad (2)$$

Оценить погрешность такого приближения невозможно, так как неизвестна $f^{(n)}(\xi)$. Однако ясно, что при увеличении n и уменьшении $x - b$ погрешность приближения уменьшается.

Последняя формула может оказаться полезной, только если известно аналитическое представление функции $\varphi(x)$, так как в этом случае можно выписать $\varphi^{(k)}(x)$. Если же $\varphi(x)$ задается на $[a, b]$ неточно или при помощи таблицы, то возникает независимая задача интерполяции на $[a, b]$.

В общем случае интерполяционный полином можно записать в форме Лагранжа [10]:

$$\hat{\varphi}(x) = \sum_{i=0}^m y_i \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}, \quad (3)$$

где y_i — значения функции в узлах x_i при $i = 0, 1, \dots, m$. Если экстраполяция осуществляется полиномом степени меньше m , то можно воспользоваться представлением (2), где производные вычислены по (3).

Будем рассматривать экстраполяцию как численный метод прогнозирования на один шаг дискретности h . Тогда метод будет устойчив [70], если все корни характеристического уравнения лежат в единичном круге комплексной плоскости, а корни, по модулю равные единице, обязательно простые. Из этого критерия следует теорема для полиномиальной экстраполяции по нескольким равноотстоящим узлам.

Утверждение. Полиномиальная экстраполяция с равноотстоящими узлами всегда неустойчива.

Действительно, пусть $Y(x)$ — вектор-функция ($Y = y_i$, $i = 1, \dots, m$). Интерполяционный полином в форме Лагранжа по $(m + 1)$ -узлу будет иметь вид:

$$Y(x) = \sum_{k=0}^m Y(x_k) \frac{(x - x_0) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_m)}{(x_k - x_0) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_m)}.$$

Положим теперь $Y_i = Y(x_i)$, $x_i = hi$, где $h > 0$ — шаг дискретности. Тогда

$$Y_{i+1} = \sum_{k=0}^m Y_{i-k} C_{m+1}^{k+1} (-1)^k, \quad i = 1, 2, \dots$$

Последнее равенство есть рекуррентное соотношение для определения очередного вектора Y_{i+1} по $m + 1$ предыдущему. Характеристическое уравнение для него имеет вид:

$$\lambda^{m+1} + \sum_{k=0}^m (-1)^{k+1} C_{m+1}^{k+1} \lambda^{m-k} = 0,$$

или иначе $(\lambda - 1)^{m+1} = 0$. Следовательно, корни характеристического уравнения равны единице и обязательно кратные. Значит, имеет место неустойчивость.

Неустойчивость экстраполяции с помощью полиномов не может быть устранена, если выбирать узлы интерполяции каким-либо специальным образом. Так, например, рассмотрим процесс экстраполяции по двум предыдущим узлам с произвольными шагами дискретности. Итак, пусть

$$y(x) = y(x_0) \frac{x - x_1}{x_0 - x_1} + y(x_1) \frac{x - x_0}{x_1 - x_0}.$$

Обозначим последовательность шагов дискретности h_k , $k = 1, 2, \dots$, так что $x_1 = x_0 + h_1$, $x_2 = x_1 + h_2 \dots$ Тогда рекур-

рентное соотношение численного метода записывается как линейное нестационарное разностное уравнение второго порядка:

$$y_{k+1} = \left(\frac{h_{k+1} + h_k}{h_k} \right) y_k - \left(\frac{h_{k+1}}{h_k} \right) y_{k-1}, \quad k = 1, 2, \dots$$

Общее решение этого уравнения может быть получено по индукции в форме

$$y_{k+1} = y_0 + \left(\sum_{i=1}^{k+1} h_i \right) \frac{y_1 - y_0}{h_1}.$$

Нетрудно видеть, что если решение ограничено, то оно и не продолжимо на всю ось (кроме постоянных решений, естественно). Значит, и численный метод неустойчив.

Экстраполяция может проводиться с сохранением некоторых качественных особенностей функции $\varphi(x)$, которые аналитически выражаются, например, условиями типа:

а) $G(x, \varphi(x)) = C_1$ при $x \in [a, b]$;

б) $\int_x^{x+h} G(\tau, \varphi(\tau)) d\tau = C_2$ при $x, x+h \in [a, b]$.

Если имеет место а), то естественно считать, что при $x \in [b, \beta]$ будет выполняться $G(x, f(x)) = C_1$. Последнее равенство является уравнением относительно $f(x)$ при $x \in [b, \beta]$. Оно разрешимо при известном условии $\partial G / \partial y \neq 0$ при $x \in [b, \beta]$. Однако необходимо потребовать существование непрерывных производных функций G и φ .

Если имеет место б), то для нахождения $f(x)$ можно записать интегральное уравнение: для всякого $x \in [b, \beta]$

$$\int_{\min\{(x-h), b\}}^b G(\tau, \varphi(\tau)) d\tau + \int_b^x G(\tau, f(\tau)) d\tau = C_2, \quad (4)$$

которое можно решать методами теории интегральных уравнений, например, если G линейны по f , то уравнение (4) будет интегральным уравнением Вольтерры I рода. Дифференцируя б) по x , можно получить более общее условие:

$$G(x, f(x)) = G(x-h, f(x-h)). \quad (5)$$

Если G разрешима относительно f , то для всякого фиксированного $x \in [b, \beta]$ можно получить $f(x)$. Дифференцируя уравнение (5) по x , получаем:

$$\frac{\partial G(x, f)}{\partial x} + \frac{\partial G(x, f)}{\partial f} f' = \frac{\partial G(x-h, f_h)}{\partial x} + \frac{\partial G}{\partial f} f'_h,$$

или иначе

$$\frac{df}{dx} = \frac{\frac{\partial G(x-h, f_h)}{\partial x} - \frac{\partial G(x, f(x))}{\partial x} + \frac{\partial G(x-h, f_h)}{\partial f} f'(x-h)}{\frac{\partial G(x, f(x))}{\partial f}}. \quad (6)$$

1.3.

МОДЕЛИ ЭКОНОМИЧЕСКОГО РОСТА

В качестве введения мы обратимся к мировому опыту в исследовании экономической динамики. Для этого используем обзор, который базируется на литературе по данному направлению, имеющейся в библиотеке Санкт-Петербургского государственного университета.

Основную работу по составлению обзора выполнил Д. И. Ткаченко. Конечно же, этот обзор получился далеко не полным по литературным источникам, по анализу полезности и применимости моделей, но для тех читателей, которые впервые сталкиваются с моделями в экономической динамике, прочесть эту главу будет несложно и полезно.

1.3.1.

ИСТОРИЧЕСКАЯ СПРАВКА

Первые модели экономического роста появились сравнительно недавно, во второй четверти XX в., да и вообще моделирование в экономике насчитывает немногим более двух веков. Историческая необходимость разработки экономических моделей возникла в связи с переходом от натурального хозяйства к товарному, бурным развитием мануфактурного производства и централизацией экономики. К XVIII в. в ряде европейских стран сложилось то, что принято называть национальной экономикой. Именно тогда и

именно в этих странах были предприняты первые попытки создания стройной, математически обоснованной экономической теории. Среди первопроходцев следует упомянуть Ф. Кенэ (1766) во Франции с его «Экономической таблицей» и представителей «классической» школы в Англии (А. Смит, Д. Рикардо и др.), работавших чуть позднее. Однако их рассуждения были все же достаточно умозрительными, «классики» использовали, как правило, лишь самый примитивный математический аппарат, а то и вовсе ограничивались словесными формулировками. Значительным шагом вперед стало включение динамики в экономическую теорию, предпринятое К. Марксом. Им было сформулировано условие расширенного воспроизводства и численно проанализированы схемы такого воспроизводства.

Однако в дальнейшем динамика надолго выпала из поля зрения экономистов. В последней четверти XIX в. ведущее место в экономической науке заняла «неоклассическая» школа, представители которой (А. Маршалл, Г. Кассель и др.) отказались от подхода к экономическому развитию как к целостному процессу и сконцентрировали внимание на проблемах краткосрочного порядка. Они заложили основы микроэкономики, сосредоточившейся на исследовании поведения отдельных производителей и потребителей. Ее составными частями стали теория потребительских предпочтений, теория фирмы, теория производственных функций. В ведении макроэкономики остались производство и распределение национального продукта, соотношения между макропараметрами, взаимосвязь отраслей экономики.

С начала XX в. происходит глобализация мировой экономики, наблюдается дальнейшая специализация отраслей и одновременно тесное взаимодействие между ними, благодаря достижениям науки радикально меняется способ производства. Все это вызывает необходимость составления долгосрочных прогнозов, анализа связей между разными экономическими явлениями, наилучшего использования ресурсов. Требуется усиление роли государства в экономике, особенно в свете периодически возникающих мировых экономических кризисов. Самый мощный из таких кризисов, «Великая депрессия» 1929–1933 гг.,

высветил несостоятельность прежних теорий. Требовалось создать теорию, которая объяснила бы причины этого кризиса, показала пути выхода из него и заложила основы новой экономической политики, позволяющей избежать подобных кризисов в дальнейшем. В результате на гребне «Великой депрессии» появились модели Дж. М. Кейнса и его последователей, перевернувшие представления о движущих силах и направлении развития экономики.

Основное методологическое отличие кейнсианцев от их предшественников состояло в том, что представители классического и неоклассического направлений лишь констатировали распределение доходов, складывающееся под действием стихийных рыночных сил. Кейнсианцы же указывали на необходимость более справедливого распределения доходов и предлагали механизм такого распределения с участием государства. Математической же новацией кейнсианцев стало включение инвестиций в модель производства валового продукта, что позволило рассматривать экономику в динамике.

На базе кейнсианства возникли (в конце 1930-х гг. и позже) тесно связанные между собой теории экономического роста (Р. Харрод) и цикла (Э. Хансен, П. Самуэльсон, Дж. Хикс). Примерно в это же время появились межотраслевые модели Дж. Неймана и В. Леонтьева. Разработка новых разделов математики, таких, как линейное программирование, дифференциально-разностные уравнения, теория игр, позволила получить содержательные результаты в рамках этих моделей. Во второй половине XX в. кейнсианская теория была дополнена изысканиями представителей неоклассического направления (включая теорию производственных функций), что привело к появлению неоклассической теории роста. К последним достижениям этой теории следует отнести включение в нее эндогенного научно-технического прогресса.

Множество всех динамических моделей макроэкономики можно классифицировать, используя различные признаки. В частности, модели различаются по характеру поведения экономики: равновесие, циклическое развитие, хаотические колебания, рост. Последние, в свою очередь,

различаются по характеру роста (монотонный, несбалансированный, циклический, магистральный). По отражению отраслевой структуры экономические модели делятся на односекторные (рассматривается вся экономика без деления ее на отрасли) и многосекторные. По количеству используемых факторов производства — однофакторные (основным фактором, как правило, выступает капитал или, реже, труд) и многофакторные. По сроку действия — краткосрочные и долгосрочные (в них исследуются асимптотические свойства, однако может не учитываться поведение на начальных этапах). Кроме того, к динамическим моделям макроэкономики применима общематематическая классификация. Они делятся на: аналитические (используются лишь для прогноза) и оптимизационные (для выбора одного из нескольких путей развития); стохастические (вероятностные) и детерминированные; открытые и замкнутые. По используемому математическому аппарату — с применением теории положительных матриц, линейного программирования, дифференциальных уравнений (обыкновенных и с отклоняющимся аргументом), выпуклых конусов и т. д. В последующих разделах мы коснемся большинства из перечисленных типов моделей экономического роста.

1.3.2. ОДНОСЕКТОРНЫЕ МОДЕЛИ РОСТА

1.3.2.1. МОДЕЛЬ МУЛЬТИПЛИКАТОРА КЕЙНСА

Наиболее простой, хотя и не первой хронологически, односекторной моделью является модель мультипликатора, предложенная Дж. М. Кейнсом (1936). В рамках этой модели устанавливается зависимость национального дохода от уровня инвестиционной активности. Национальный доход (валовой продукт) в этой модели, как и в большинстве других, является основным индикатором состояния экономики. Предполагается [2], что весь произведенный валовой продукт Y расходуется на потребительские нужды $C(Y)$ и идет на сбережения населения, которые, в свою очередь, целиком инвестируются в экономику, т. е.

$$Y = C(Y) + I. \quad (7)$$

Здесь I — автономные инвестиции, т. е. их величина фиксированна и не зависит от дохода. Более того, предполагается, что как раз доход зависит от этих инвестиций. Дифференцируя теперь данное соотношение по I , получаем зависимость между доходом и инвестициями, которую экономисты записывают обычно в форме конечных разностей:

$$\Delta Y = \frac{1}{1-c} \Delta I, \quad (8)$$

где $c = dC/dY$ — «предельная склонность к потреблению» (не обязательно, чтобы она была постоянной). Поскольку $c \in (0, 1)$, из (8) следует, что $\Delta Y > \Delta I$, т. е. единичный прирост инвестиций приводит к приросту выпуска, большему, чем единичный. Соответствующий эффект называется мультипликативным, а коэффициент $1/(1-c)$ — мультипликатором.

Однако серьезным недостатком модели мультипликатора является то, что изменение объемов производства зависит не от самих инвестиций I , а от их изменения ΔI , т. е. реальный объем текущих инвестиций не играет, по сути, никакой роли. Поскольку величина ΔI в какие-то периоды времени вполне может оказаться отрицательной при положительном общем уровне инвестиций I (это происходит во время снижения инвестиционной активности), выпуск продукции в такие периоды также будет снижаться. Таким образом, было бы некорректно относить модель Кейнса к моделям экономического роста, ибо рост в ней возможен, но не обязателен.

Кроме того, прирост инвестиций носит здесь совершенно произвольный, экзогенный характер, т. е. обусловлен решениями, принятыми вне модели. Модель Кейнса показывает влияние изменения инвестиций на состояние экономики, однако не объясняет причину этих изменений. Источник инвестиций, находящийся вне модели, оказывает решающее влияние на экономику страны. Таким источником могут быть иностранные кредиты, и, следовательно, модель справедлива для небольших стран с открытой экономикой, полностью зависимых от иностранного капитала.

Важно также отметить, что модель Кейнса вызывает вопросы, если проанализировать ее логически. Равенство (7) изначально выступает как распределительное соотношение для произведенного дохода. Однако далее это же равенство рассматривается фактически как производственная функция, т. е. как правило, по которому Y зависит от единственного фактора производства, роль которого играют инвестиции I . Допустимость дальнейших математических операций над соотношением (7) в таком качестве вызывает большие сомнения.

Исходная модель Кейнса может быть дополнена посредством включения в нее [85] таких макроэкономических параметров, как налоги T , государственные расходы G , чистый экспорт E (разность между экспортом и импортом) и т. п. Тогда равенство (7) уточняется:

$$Y = C(\bar{Y}) + I + G(Y) + E(Y),$$

где $\bar{Y} = Y - T(Y)$; дифференцируя его по I , получаем модель, аналогичную (8), но с мультипликатором:

$$\left[1 - \frac{dC}{d\bar{Y}} \times \left(1 - \frac{dT}{dY} \right) + \frac{dG}{dY} + \frac{dE}{dY} \right]^{-1}.$$

Можно также получить соотношения [89], отражающие зависимость ΔY от прироста государственных расходов, налогов или какого-либо иного фактора. Несмотря на упоминавшиеся вопросы (экзогенность переменных, необходимость роста, применимость не ко всякой экономике, логическая некорректность), модель Кейнса послужила серьезным толчком к созданию ряда более совершенных моделей экономического роста и цикла.

1.3.2.2.

МОДЕЛИ ХАРРОДА-ДОМАРА

Более содержательными достижениями кейнсианской экономической мысли представляются модели, разработанные Р. Харродом (1948) и Е. Домаром (1957). Их основным отличием от первоначальной модели Кейнса является индуцированный, эндогенный характер инвестиций. Вообще, Харродом был предложен критерий, отделяющий

динамическую модель экономики от статической [17]: в динамике сбережения населения не равны нулю. Положительность сбережений позволяет инвестировать средства в экономику. Поскольку инвестиции в основные фонды составляют некоторую долю национального дохода, это приводит к росту основных фондов, а он, в свою очередь, к росту производства продукции. Класс моделей, основанных на таких предположениях, получил название моделей экономического роста или моделей Харрода–Домара.

В простейшем из вариантов такая модель связывает национальный доход Y , инвестиции I и фонд основного капитала K (именуемый далее просто «капиталом»). Предполагается, что доля дохода, идущая на инвестиции, постоянна: $I = \sigma Y$. Инвестиции являются источником прироста основного капитала: $\dot{K} = I$. Доход пропорционален используемому капиталу: $Y = \mu K$. Последнее соотношение для однофакторной модели (единственный фактор производства — капитал) может рассматриваться как простейшая производственная функция, которой явно не хватало в модели мультипликатора. Объединение этих трех соотношений:

$$I = \sigma Y, \quad \dot{K} = I, \quad Y = \mu K, \quad (9)$$

в обыкновенное дифференциальное уравнение относительно Y позволяет найти решение в явном виде:

$$Y(t) = Y_0 e^{\sigma \mu t}. \quad (10)$$

Соотношение (10) показывает, что доход растет экспоненциально, а следовательно, и остальные переменные (K и I) растут экспоненциально с тем же показателем экспоненты. Эта модель имеет, конечно, краткосрочный характер, поскольку при больших значениях функция (10) будет быстро уходить в бесконечность, что противоречит экономической действительности. Подход Домара [106] методологически несколько отличается от подхода Харрода, однако полученные им результаты идентичны изложенным выше.

Базовая модель Харрода–Домара породила бесчисленное множество модификаций, усовершенствований и под-

ражений. Как правило, ревизии подвергается одно из соотношений в системе (9). Так, в модели Н. Кальдора [65] объем инвестиций предполагается зависящим не только от уровня национального дохода, но и от такой специфической величины, как прирост удельного дохода предпринимателей, а, к примеру, в модели А. Филлипса [9] инвестиции зависят от темпа роста выпуска продукции (т. е. от величины \dot{Y}) на некотором интервале, что, в частном случае, приводит к уравнению $\dot{K} = \alpha Y - \beta K$ вместо (9).

В. Леонтьевым [53] справедливо было замечено, что модель Харрода–Домара в виде (9) не является универсальной, поскольку применима лишь к странам, опирающимся на внутренние ресурсы и не импортирующим капитал. Для стран, зависимых от иностранного капитала, соотношение $I = \sigma Y$ следует заменить выражением:

$$I = \sigma Y + \omega Y_{imp}(t),$$

где $Y_{imp}(t)$ — выпуск продукции в развитых странах (одной или нескольких), уже найденный по формуле (10). Тогда производство в зависимой (развивающейся) стране определяется как

$$Y(t) = q_1 \exp(\sigma \mu t) + q_2 \exp(\sigma_{imp} \mu_{imp} t),$$

где q_1 и q_2 — некоторые константы, а σ_{imp} и μ_{imp} — параметры из формул (9) для страны, импортирующей капитал. В случае, если $\sigma \mu < \sigma_{imp} \mu_{imp}$, внешнее воздействие на экономику зависимой страны оказывается более значимым, чем ее внутренние параметры.

Исходная модель Харрода–Домара описывает лишь процесс производства, никак не учитывая при этом распределение полученного дохода. Очевидно, систему (9) можно дополнить распределительным соотношением, которое фигурировало еще у Кейнса:

$$Y = I + C. \quad (11)$$

Однако механическое включение в модель (9) потребления в виде $C = cY$ изменяет лишь коэффициенты в формуле (10), не меняя сути модели, т. е. экспоненциального роста дохода, поскольку инвестиции становятся равными

$I = (1 - c)Y$, что концептуально не отличается от соотношения (9).

Потребление может быть задано и экзогенно, в частности экспоненциально растущим [22] с неотрицательным коэффициентом роста r :

$$c(t) = c_0 e^{rt}.$$

Оказывается, возможность роста в такой модели зависит от величины r . При больших r рост имеет место лишь в начале временного интервала, а далее производство падает до отрицательной величины. В самом деле, если потребление растёт значительно быстрее выпуска, то в какой-то момент окажется $Y(t) < C(t)$, т. е. всего полученного дохода не хватает на то, чтобы удовлетворить потребительские запросы. Лишь при достаточно малых r имеет место рост выпуска, асимптотически стремящийся к экспоненциальному. Если в формуле (11) определять, согласно Р. Фришу, не само потребление, а его полезность [90], то функция дохода может быть найдена в виде:

$$Y(t) = e^{\mu t} \beta_1 \arctg(\beta_2 e^{-\mu t}) + \beta_3,$$

где β_j — положительные константы. В отличие от уравнения (10) доход здесь возрастает не экспоненциально, а по кривой, напоминающей логистическую, и асимптотически стремится к постоянной величине.

В одном из последующих разделов будет подробно рассказано о динамических межотраслевых моделях Леонтьева, односекторным аналогом которых можно считать слегка модифицированную модель Харрода–Домара, включающую наряду с равенствами $Y = \mu K$ и $\dot{K} = I$ соотношение

$$Y = \alpha Y + I + C, \quad (12)$$

характеризующее распределение произведенной продукции. Первый член в правой части уравнения (12) представляет собой воспроизводство ВВП (естественно, должно выполняться условие $\alpha < 1$), т. е. произведенный валовой продукт подразделяется в действительности не на две, а на три части. Соотношение (12) неявным образом заложено в (9) и (11), где валовой продукт рассматривается уже приведен-

ным, т. е. член αY в (12) перенесен налево и обе части поделены на $(1 - \alpha)$.

Важнейшим аспектом теории экономического роста является то, что вложенные в производство инвестиции не дают мгновенной отдачи в силу ненулевой длины производственного цикла, поэтому в соответствующие модели обычно вводятся запаздывания. Должно быть изменено соотношение, показывающее влияние инвестиций на рост основного капитала:

$$I(t) = \dot{K}(t+h);$$

член αY из тех же соображений необходимо также рассматривать в момент времени $(t+h)$. Получившееся в результате уравнение

$$Y(t) = \alpha Y(t+h) + \mu^{-1} \dot{Y}(t+h) + C(t),$$

является дифференциальным уравнением запаздывающего типа и решается обычно пошагово. При этом [91] если на каждом из таких шагов потребление не превышает некоторой величины (численное значение которой определяется на предыдущем шаге), то модель Харрода–Домара с запаздыванием является моделью экономического роста, т. е. во все моменты времени выполняется условие $\dot{Y}(t) > 0$.

Исходную модель Харрода–Домара принято интерпретировать как однофакторную с линейной зависимостью выпуска от этого единственного фактора (капитала). Между тем такую зависимость можно получить [86], отталкиваясь от двухфакторной производственной функции с нулевой эластичностью замены:

$$Y = \min \left\{ \frac{K}{v}, \frac{L}{a} \right\}, \quad (13)$$

рассматриваемой вместо однофакторной $Y = \mu K$ в (9). Здесь L — «труд», т. е. объем трудовых ресурсов, определяемый соотношением

$$\dot{L} = nL, \quad (14)$$

v и a — коэффициенты соответственно фондоемкости и трудоемкости.

Таким образом, характер и скорость роста выпуска определяются посредством немодельной процедуры выбора минимального из двух значений в уравнении (13). При $K/v > L/a$ наблюдается избыток капитала, и часть производственных мощностей остается незагруженной, зато обеспечивается полная занятость. При $K/v < L/a$, наоборот, обеспеченность капиталом невысока, и он весь используется в производстве, в то время как часть рабочей силы остается незанятой. Соотношение (13) преобразуется (с учетом (9)) в выражение

$$\frac{\dot{Y}}{Y} = \min \left\{ \frac{\sigma}{v}, n \right\},$$

т. е. выбирается одно из двух значений темпа роста. Кейнсианцы делают акцент на недостаточности инвестиционной активности, т. е. останавливаются на варианте $K/v \leq L/a$ и, следовательно, $\dot{Y}/Y = \sigma/v \leq n$. Темп σ/v называется равновесным (гарантированным) темпом роста, он допускает безработицу. Полная занятость возможна лишь при достижении «естественного» темпа роста $\sigma/v = n$.

В описанной выше модели выпуск продукции реально зависит от динамики лишь одного из двух факторов. Механическое добавление в уравнение (9) второго фактора (труда) $L = lY$, динамика которого определяется из (14), приводит к тому [36], что система становится переопределенной, т. е. число уравнений превышает число неизвестных. Такая система разрешима лишь при равенстве спроса на труд и предложения труда, т. е. при одновременном выполнении $\mu l K_0 = L_0$ и $\sigma \mu = n$. Изменить ситуацию могло бы использование нелинейной двухфакторной производственной функции, например в виде модели Кобба–Дугласа [29], однако и здесь осмысленные результаты получаются лишь при дополнительных предположениях. Помимо капитала и труда, возможно использование и других факторов производства, в частности научно-технического прогресса $a(t)$: $Y = f(K, L, a(t))$, или, к примеру, такого экзотического фактора, как размер налоговых отчислений [67], пропорциональных выпуску, что может рассматриваться как задача стабилизации.

1.3.2.3. МОДЕЛИ ЦИКЛИЧЕСКОГО РОСТА

Как уже отмечалось, кейнсианские теория роста и теория цикла имеют общие идеологические корни и представляют собой, по сути, двуединую теорию роста–цикла, в различных версиях которой одному из этих двух явлений уделяется большее внимание, чем другому. Так, базовая модель Харрода–Домара (9) при минимальных изменениях преобразуется в модель циклического роста. Основной причиной появления цикличности является фактор времени. Инвестиционный процесс протяжен во времени (это приводит к появлению запаздываний), причем на объем основных фондов влияют лишь «законченные» инвестиции I_+ , т. е. $\dot{K} = I_+$, а суммарные инвестиции за период длины θ определяются как

$$I(t) = \frac{1}{\theta} \int_0^{\theta} \eta(\tau) I_+(t - \tau) d\tau,$$

где η — весовая функция

$$\int_0^{\theta} \eta(\tau) d\tau = 1.$$

Частный случай определения $\eta(\tau)$ [27] приводит к модели М. Калецкого (1935), в которой переменная \dot{Y}/Y имеет периодические колебания вокруг положительного значения. Из этого следует, что в целом в модели наблюдается рост выпуска, однако вокруг тренда равновесного роста возникают колебания (а вокруг них, в свою очередь, вторичные циклы).

Несколько иной характер циклического роста и иные его предпосылки выявлены Р. Гудвином (1951). Он [107] вводит в исходную систему (9) и (11) инвестиционное запаздывание и преобразует ее затем следующим образом:

$$Y = C + I - \varepsilon \dot{Y}, \quad C = cY, \quad I = l(t) + \varphi(\dot{Y}).$$

Здесь $l(t)$ и $\varphi(\dot{Y})$ — соответственно автономные и индуцированные инвестиции (в главе 3 мы рассмотрим эту модель несколько подробнее и критичнее). При $l(t) = 0$ и некоторых ограничениях на функцию $\varphi(\dot{Y})$ экономика развивалась бы полностью циклически. Причиной же

появления тенденций к росту в этой модели являются автономные инвестиции. Положительность $l(t)$ приводит к тому, что цикл по переменной \dot{Y} сохраняется, однако с каждым шагом этот цикл сдвигается по Y в сторону роста Y . В то время как модели «чистого» роста Харрода–Домара описывают стабильное поступательное развитие экономики, модели Калецкого и Гудвина, не отрицая этого, указывают на циклический характер такого роста.

1.3.2.4.

НЕОКЛАССИЧЕСКИЕ МОДЕЛИ РОСТА

Следующей важной вехой в теории роста стали модели, разработанные в рамках теории «неоклассического синтеза». В 1950-е гг. кейнсианская модель Харрода–Домара была дополнена элементами неоклассической теории, в первую очередь теории производственных функций, а впоследствии к этой модели были приобщены также идеи Ф. Рамсея.

В отличие от кейнсианцев, у неоклассиков рост носит сбалансированный характер, не приводит к циклическим явлениям и обладает магистральными свойствами. Наибольшей известностью пользуются модели Солоу, Фелпса и модификации модели Рамсея.

В модели Р. Солоу (1956, одновременно была предложена Т. Свэнгом), в отличие от кейнсианских моделей, предполагается непрерывная взаимозаменяемость труда и капитала, а также полная занятость [89]. Это достигается при помощи двухфакторной производственной функции $Y = f(K, L)$, характеризующейся постоянной отдачей от масштаба (математики называют такую функцию однородной первой степени), т. е.

$$f(\alpha K, \alpha L) = \alpha f(K, L).$$

Фактически же производственная функция изящно превращается в однофакторную с единственным фактором $k = K/L$ — капиталом на одного рабочего:

$$\frac{1}{L} f(K, L) = f\left(\frac{K}{L}, 1\right) = f(k, 1) = \varphi(k).$$

Дополним функцию $\varphi(k)$ соотношениями $\dot{K} = I$ и $I = \sigma Y$ из (9) и функцией роста трудовых ресурсов (14). Тогда при

$$\sigma\varphi(k^*) = nk^* \quad (15)$$

выполняется равенство $\dot{k} = 0$, т. е. достигается положение равновесия в модели Солоу $k(t) \equiv k^*$. Совокупный выпуск при таком «равновесии» возрастает: $Y(t) = L_0\varphi(k^*)e^{nt}$, следовательно, модель Солоу является моделью равновесного роста. Если условие равновесия (15) не выполнено и $\dot{k}(0) > 0$, то $k(t)$ монотонно возрастает и асимптотически стремится к равновесному значению k^* (а при $\dot{k}(0) < 0$ — убывает до k^*). Значит, при любых начальных условиях имеет место:

$$Y(t) \xrightarrow{t \rightarrow \infty} L_0\varphi(k^*)e^{nt}.$$

Таким образом, модель Солоу обладает магистральным свойством, т. е. вне зависимости от начальных условий развитие экономики асимптотически стремится к сбалансированному росту, темп которого равен, в данном случае, темпу роста населения.

Эта модель, как, впрочем, и все предыдущие, является сугубо аналитической, т. е. лишь прогнозирует поведение экономики при заданных параметрах.

Следующим шагом стала разработка моделей, позволяющих выбрать наиболее выгодный для производства или потребления («оптимальный») путь развития экономики. Такой оптимизационной моделью является, в частности, модель Э. Фелпса (1961), которую можно считать творческим развитием модели Солоу. Индикатором оптимальности здесь [89] выступает, как и в большинстве подобных моделей, потребление (в данном случае удельное). Максимум потребления достигается при капитале, удовлетворяющем равенству:

$$\varphi'(k^*) = n. \quad (16)$$

Соотношение (16), именуемое «золотым правилом накопления», показывает, что условием оптимальности потребления является равенство предельной производительности капитала равновесному темпу роста экономики. Следует отметить, что модели Фелпса и его последователей имеют смысл лишь для долгосрочного периода, поскольку

оптимальность в них лишь асимптотическая и потери потребления в начальные периоды могут быть весьма велики.

Оптимизационная модель Ф. Рамсея (1928) первоначально была очень простой [89] и включала соотношения $\dot{K} = I$, $Y = f(K)$ и $Y = C + I$. Однако ее отличительной чертой было введение критерия полезности потребления. Целью, по Рамсею, является максимизация функционала суммарной полезности потребления:

$$U = \int_0^{\infty} u(C(t))dt,$$

где $U(C) > 0$, $U'(C) \geq 0$, $U''(C) \leq 0$, т. е. $U(C)$ растет с ростом потребления и стремится к уровню насыщения. Результатом является соотношение, связывающее I с $U(C)$, из которого определяется оптимальная инвестиционная политика.

Впоследствии модель Рамсея была дополнена уже упоминавшимися неоклассической производственной функцией $Y = f(K, L)$ с постоянной отдачей от масштаба и формулой (14) роста трудовых ресурсов. Кроме того, было учтено, что инвестиции идут не только на прирост новых мощностей \dot{K} , но и на возмещение выбывших:

$$I = \dot{K} + \mu K, \quad (17)$$

где μ — коэффициент амортизации.

Д. Кэсс и Т. Купманс (1965) применили [22] к модели Рамсея гипотезу «диахронной неравноценности» (гипотезу о существовании потребительских предпочтений во времени): полезность блага в настоящем выше полезности того же блага в будущем. Это выразилось во включении в критерий оптимальности дисконтирующего множителя $e^{-\delta t}$ (либо $(1 - \delta)^t$). Полученную модель:

$$\max \int_0^{\infty} e^{-\delta t} U(C(t))dt, \quad (18)$$

$$Y = C + I, \quad I = \dot{K} + \mu K, \quad Y = L\phi(k), \quad k = \frac{K}{L}, \quad \dot{L} = nl,$$

принято по-прежнему называть моделью Рамсея. Равновесное значение капитала в ней определяется [38] по «модифицированному золотому правилу»:

$$\varphi'(k^*) = \delta + \mu + n,$$

из которого затем находятся все переменные в (18). Как и в модели Солоу, оптимальный рост асимптотически стремится к сбалансированному, удовлетворяя тем самым магистральному свойству.

Существует масса дополнений и модификаций модели Рамсея, сводящихся, как правило, к уточнению или замене одного из соотношений в (18). Так, введение запаздывания инвестиций [99] не вносит качественных изменений в модель. Отличие возникает лишь в максимизирующем функционале в (18), к которому добавляется слагаемое, характеризующее прирост капитала за период запаздывания. Выясняется также, что с ростом запаздывания оптимальное значение душевого потребления падает.

Иная модификация модели Рамсея была предложена К. Шеллом [7]. Он включил в модель научно-технический прогресс:

$$Y(t) = e^{pt} F(K, L),$$

а также рассмотрел случай производственной функции, однородной степени $\gamma > 0$, т. е.

$$F(\alpha K, \alpha L) = \alpha^\gamma F(K, L)$$

(у Солоу было $\gamma = 1$). В каждой из таких моделей выводится свое «золотое» (или «слегка позолоченное») правило нахождения k^* .

Пожалуй, последним крупным вкладом в односекторную теорию роста (после моделей Харрода–Домара и неоклассических моделей роста) стали модели эндогенного научно-технического прогресса [87]. В таких моделях, в частности в модели П. Ромера (1986), НТП выступает в качестве одного из эндогенных (внутренних) факторов роста (наряду с трудом и капиталом), и вложения в НТП дают возрастающую общественную отдачу. Подобные попытки предпринимались и раньше [64]: предлагалось включить в распределение дохода затраты N на изменение уровня технических знаний (вложения в науку):

$$Y = C + I + N.$$

Производственная функция $Y = a(t)F(K, L)$ замыкалась соотношением $\dot{a} = \xi N - \tau a$, где $a(t)$ — НТП. Однако предположение о линейной связи вложений в науку N и дохода Y нивелирует различия между вложениями в науку, потреблением и инвестициями. Помимо Ромера, основоположником эндогенной теории роста считается Р. Лукас (1988), в модели которого [24] средства предлагается вкладывать не в НТП, а в образование — как инвестиции в человеческий капитал.

1.3.3.

МОДЕЛИ НЕЙМАНОВСКОГО ТИПА

Основным достоинством односекторных моделей была их простота, основным недостатком — игнорирование отраслевой структуры экономики и связей между отраслями. Устранить эти недостатки были призваны межотраслевые модели.

Попутно следует упомянуть о своеобразном промежуточном варианте между одно- и многосекторными моделями. Это двухсекторная модель расширенного воспроизводства, в которой вся экономика разбита на 2 «подразделения», а между ними установлена иерархическая зависимость. Теоретические основы этой модели были заложены К. Марксом, и ей, в силу известных идеологических причин, уделялось в нашей стране излишне пристальное внимание. Впрочем, и для неоклассических моделей роста был построен их двухсекторный аналог с двумя «подразделениями» [38].

Собственно же многосекторных (межотраслевых) моделей известно 2 основных типа — неймановские и леонтьевские модели, все остальное создано на их основе. Интересно, что модели Неймана и Леонтьева появились практически одновременно с кейнсианскими односекторными моделями роста и цикла и активно использовали их постулаты, в частности соотношения (9) и (17) были распространены на случай нескольких отраслей, а вот теория неоклассического роста в многосекторных моделях долгое время не использовалась, возможно, по причине ее нелинейности.

1.3.3.1.
СТАЦИОНАРНЫЕ ТРАЕКТОРИИ
В МОДЕЛИ НЕЙМАНА

В модели расширяющейся экономики Дж. фон Неймана (1932) всю экономику характеризует не один производственный процесс, как прежде, а конечное число таких процессов. В рамках таких процессов производится конечное число «продуктов», в разряд которых попадают также первичные факторы и конечное потребление. Каждый производственный процесс способен производить сразу несколько видов продукции, а каждый продукт может производиться несколькими процессами. Каждый i -й процесс, работая с интенсивностью $z_i \geq 0$, затрачивает j -й продукт в количестве $a_{ij} \geq 0$ и производит k -й продукт в количестве $b_{ik} \geq 0$, $i = \overline{1, n}$, $j, k = \overline{1, m}$.

Весь производственный механизм может быть описан технологическим множеством

$$G = \{(x, y) \mid \exists z: x = zA, y = zB\}, \quad (19)$$

где A и B — матрицы затрат и выпуска, составленные из описанных выше коэффициентов a_{ij} и b_{ij} .

При этом в модели Неймана делаются следующие допущения [65]. Модель предполагается линейной, а технология неизменной во времени, т. е. матрицы A и B постоянны. Процессы применяются в течение T периодов времени с траекторией интенсивностей

$$\{z(t)\}_{t=0}^T.$$

Система не подвергается воздействию внешних факторов и удовлетворяет условию замкнутости:

$$z(t+1)A \leq z(t)B, \quad t = \overline{0, T-1}, \quad (20)$$

означающему, что затраты продукции в текущем периоде не могут превышать затрат выпуска предыдущего периода. Каждый продукт имеет цену $p_j(t) \geq 0$, которая может изменяться во времени. Тем самым вводится в рассмотрение последнее из предположений — никакой процесс не приносит положительной прибыли:

$$Ap(t) \geq Bp(t+1), \quad t = \overline{0, T-1}. \quad (21)$$

Задача состоит в нахождении траектории интенсивностей, удовлетворяющей условию замкнутости (20), и траектории цен, удовлетворяющей условию неприбыльности (21).

Частным, но особенно интересным случаем допустимой траектории интенсивностей является «стационарная» траектория (траектория сбалансированного роста):

$$z(t) = v^t z.$$

Такая траектория существует для тех и только для тех v , которые удовлетворяют неравенству

$$vzA \leq zB.$$

При $v = 1$ состояние экономики со временем не меняется, при $v < 1$ экономика «свертывается», а при $v > 1$ — расширяется. Именно этот последний вариант наиболее важен и желателен, и модель Неймана часто называют моделью расширяющейся экономики.

Нейманом доказана следующая теорема [65]. Если при указанных выше предположениях выполнено

$$a_{ij} + b_{ij} > 0 \quad \forall i, j, \quad (22)$$

то существуют числа $v^* = \mu^*$ и вектора z^* и p^* , удовлетворяющие неравенствам $v^* z^* A \leq z^* B$, $\mu^* A p^* \geq B p^*$ и условиям дополняющей нежесткости (т. е. из строгого неравенства $v^* z^* a^j < z^* b^j$ сразу же следует $p_j^* = 0$, а из $\mu^* a_i p^* > b_i p^*$ следует $z_i^* = 0$). Такое v^* — максимальное из всех допустимых, оно определяется единственным образом и обеспечивает наискорейший рост экономики. Для квадратных матриц нахождение оптимальных значений v^* , z^* и p^* из теоремы Неймана сводится [65] к нахождению числа Фробениуса–Перрона и соответствующих ему левого и правого собственных векторов для матрицы $A(B - A)^{-1}$ (при дополнительных ограничениях на матрицы A и B). Выдвинутое Нейманом предположение (22) слишком жесткое, оно требует, чтобы в каждом процессе обязательно либо затрачивался, либо выпускался каждый из продуктов. Дж. Кемени, О. Моргенштерном и Дж. Томпсоном доказано, что это требование можно заменить более реалистичным: в матрице A нет нулевых строк, а в матрице B — нулевых столбцов.

1.3.3.2.
МАГИСТРАЛЬНАЯ ТЕОРИЯ

Логическим развитием модели расширяющейся экономики стала магистральная теория, идея которой предложена П. Самуэльсоном (1958). Если случилось так, что начальное значение z_0 вектора интенсивностей лежит на стационарной траектории z^* (называемой также лучом Неймана), то дальше экономика уже не свернет с этого луча. Однако такой случай исключителен, и, как правило, z_0 не принадлежит лучу Неймана. В этом случае решается задача линейного программирования на максимум целевой функции $z(T)Bq$ при ограничениях (20) и двойственная к ней задача на минимум для цен.

Тогда справедлива теорема о магистрали (в сильной форме), утверждающая, что при оптимальном развитии траектория интенсивностей $z(t)$ с произвольными начальными условиями z_0 большую часть своего пути, за исключением разве что нескольких начальных и конечных периодов, проходит в малой окрестности магистрального луча z^* . В зависимости от условий, налагаемых на матрицы и начальные данные, существует масса вариантов этой теоремы. Приведем один из таких вариантов [65].

Пусть число производственных процессов совпадает с числом производимых продуктов, т. е. $n = m$; пусть матрицы A и B неособые; пусть матрица $B - A$ неотрицательно обратима, т. е. существует $(B - A)^{-1} > 0$; пусть модули всех собственных чисел матрицы BA^{-1} отличны от v^* . Тогда для $\forall \varepsilon > 0 \exists T^* > 0$, что если $T > 2T^*$, то

$$d(z(t), z^*) < \varepsilon$$

для всех $t \in [T^*, T - T^*]$, где d — угловое расстояние:

$$d(x, y) = \sum_{i=1}^n \left| \frac{x_i}{|x|} - \frac{y_i}{|y|} \right|.$$

Если количество периодов невелико, то эта теорема теряет актуальность, поскольку нарушается неравенство $T > 2T^*$. При этом следует учитывать, что в реальности число периодов не может быть очень велико, так как коэффициенты неймановской, как и любой другой экономической

модели, остаются неизменными лишь в течение нескольких лет. Затем производится пересчет коэффициентов, который может изменить направление магистрального луча. Теорема, аналогичная данной, доказана и для магистралей в потреблении [69].

Как отмечалось выше, существуют различные варианты магистральных теорем. В частности, при наложении дополнительных ограничений на начальный вектор z_0 [30] периоды, в которые оптимальная траектория может выходить за пределы ε -окрестности, лежат только в конце пути, а в начале их нет. В некоторых случаях удается лишь доказать [65] близость оптимальных траекторий не к лучу Неймана, а к множествам, размерность которых больше единицы, называемым гранями Неймана. Наиболее же известны 3 версии теоремы о магистрали: в слабой, сильной и сильнейшей формах. Теорема в сильной форме сформулирована в начале этого раздела, она считается основной. Отличие теоремы в слабой форме состоит в том, что «исключительные» периоды могут быть распределены как угодно вдоль всего промежутка планирования, а не должны обязательно лежать в его начале и конце. Жесткие ограничения, налагаемые теоремой в сильнейшей форме [22], позволяют утверждать, что большинство точек оптимальной траектории лежат не просто в ε -окрестности, а непосредственно на магистрали.

1.3.3.3. МОДИФИКАЦИИ НЕЙМАНОВСКОЙ МОДЕЛИ

Как и для односекторных моделей, дальнейшее совершенствование модели Неймана может происходить по экстенсивному пути, посредством добавления в нее новых факторов. Так, существует модель Маркса–Неймана [66], к которой, впрочем, ни Маркс, ни Нейман отношения не имеют, она предложена М. Моришимой. В отличие от базовой неймановской модели она является двухфакторной и учитывает трудовые ресурсы. Фактически же все сводится к добавлению ставки зарплаты и нормы процента в условие бесприбыльности (21). Неймановская модель может быть также дополнена [29] соотношениями, раскры-

вающими структуру капитальных вложений: ввод новых мощностей, расширение имеющихся, сворачивание недостроенных производств и т. п.

Наиболее же существенной представляется модификация Д. Гейла (1956), подвергнувшего ревизии математические основы неймановской модели. В исходной неймановской модели технологическое множество (19) является выпуклым многогранным конусом, т. е. любой вектор из него можно представить в виде линейной комбинации конечного числа векторов. Гейл же [20] не требует, чтобы конус был многогранным, ему достаточно выпуклости и замкнутости конуса. Модель Неймана является, таким образом, частным случаем модели Гейла, однако, поскольку все равно отраслей в экономике конечное число, в действительности конус по-прежнему остается многогранным. Это означает, что полученные Гейлом результаты не меняют экономического содержания неймановской модели, но позволяют применить к ней аппарат теории выпуклых конусов, им же и разработанный.

В определенных кругах бытует мнение, что межотраслевая модель Леонтьева является частным случаем неймановской модели. Это справедливо лишь отчасти. Правильнее было бы сказать, что частный случай модели Леонтьева является частным случаем модели Неймана. В отсутствие инвестиций и потребления леонтьевская модель действительно может рассматриваться как частный случай неймановской при $n = m$ и $B = E$. Более того, можно доказать [52, 65] магистральную теорему (при более мягких предположениях) для случая неймановской модели, в которой каждый продукт может выпускаться несколькими производственными процессами, но при этом в каждом процессе выпускается лишь один продукт, т. е. матрица выпусков имеет вид:

$$B = \begin{pmatrix} 1 & 1 & \dots & 1 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & \dots & 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 1 & 1 & \dots & 1 \end{pmatrix}.$$

Эта промежуточная модель называется моделью Неймана–Леонтьева. П. Самуэльсон доказал [5] теорему о замещении, согласно которой из модели Неймана–Леонтьева можно выделить подмодель, являющуюся обычной моделью Леонтьева, но по своим производственным возможностям не уступающую всей модели. На эту подмодель переносятся и магистральные свойства.

1.3.4. МОДЕЛИ ЛЕОНТЬЕВСКОГО ТИПА

1.3.4.1. МЕТОД «ЗАТРАТЫ–ВЫПУСК»

Исторические корни метода «затраты–выпуск» можно обнаружить в XVIII–XIX вв. В частности, в упоминавшейся уже «Экономической таблице» Ф. Кенэ представлено движение потоков продукции между тремя общественными классами: крестьянами, земельными собственниками и промышленниками.

Элементы анализа «затраты–выпуск» есть и у Маркса. Все же в современном и развернутом виде модель «затраты–выпуск» впервые изложена В. Леонтьевым (1936), часто ее называют также моделью межотраслевого баланса или просто моделью Леонтьева.

Первоначально леонтьевская, как и неймановская, модель была замкнутой: все товары, производимые в одних секторах экономики, выступали как источник потенциальных затрат в других, т. е. все товары носили промежуточный характер. Однако уже вскоре потребление и инвестиции были выделены в самостоятельные структурные единицы, чего так и не было сделано в неймановской модели, которая подробно анализирует структуру производства, но не уделяет должного внимания проблемам распределения произведенного продукта и инвестирования. Кроме того, в отличие от модели Неймана, которая носит сугубо теоретический характер и совершенно оторвана от своих экономических корней (это в особенности относится к интерпретации Гейла), леонтьевская модель многократно и с успехом применялась на практике при составлении межотраслевых балансов.

Модель Леонтьева предполагала, что каждая отрасль производит один тип продукции и каждый продукт производится только одной отраслью, и была первоначально статической:

$$x = Ax + c. \quad (23)$$

Коэффициенты $a_{ij} \geq 0$ матрицы затрат A означают количество i -го продукта, необходимое для производства единицы j -го продукта, и удовлетворяют неравенству

$$\sum_{i=1}^n a_{ij} < 1 \quad j = \overline{1, n}. \quad (24)$$

Требуется найти неотрицательный вектор выпусков x при заданных матрице A и векторе конечного спроса c (куда включаются конечное потребление, инвестиции, экспорт и т. п.). Получены [69] условия на матрицу затрат, гарантирующие продуктивность системы (23), т. е. существование в ней неотрицательного вектора выпусков. Однако ввиду статического характера системы (23) найденный вектор x является постоянным, а потому рост производства в рамках такой модели невозможен.

Выпуск продукции перестает быть неизменным, если допустить [65] возможность изменения параметров системы (23), т. е.

$$x(t) = A(t)x(t) + c(t).$$

В этом случае изменение выпуска

$$x(t + \tau) - x(t) = [E - A(t + \tau)]^{-1} [c(t + \tau) - c(t)] + \{[E - A(t + \tau)]^{-1} - [E - A(t)]^{-1}\} c(t)$$

вызвано изменением конечного потребления (в 1-м слагаемом) и технологии производства (во 2-м). Однако и здесь нельзя гарантировать, что изменения коэффициентов матрицы затрат приведут именно к росту выпусков по всем отраслям. Заметим в очередной раз, что распределительное соотношение используется в качестве причинно-следственной модели производства.

Изменение коэффициентов $a_{ij}(t)$, задаваемое через некое экзогенное правило, можно рассматривать и как включение научно-технического прогресса в межотраслевую

модель. Иная трактовка роли НТП в леонтьевской модели предложена Э. Фелпсом: НТП рассматривается как отдельная «отрасль», так называемый технологический сектор экономики. Известны также попытки включения труда как второго фактора производства в модель межотраслевого баланса. Один из способов состоит в добавлении к исходной модели (23) ограничений на трудовые ресурсы в виде неравенств [58]. Другой способ — включение труда в производственную функцию (например, в виде модели Кобба–Дугласа) [89], что приводит к нелинейности и позволяет получить лишь численные решения. Изменяющиеся цены также оказывают влияние на леонтьевскую модель [90].

1.3.4.2.
МАГИСТРАЛЬНЫЙ РОСТ
В ЛЕОНТЬЕВСКИХ МОДЕЛЯХ

Более содержательной представляется полудинамическая модель Х. Никайдо [69]. Никайдо не вводит в систему (23) дополнительных факторов, зато учитывает, что между затратами на выпуск продукции и самим выпуском существует некоторое запаздывание, в простейшем виде представимое как

$$x(t) = Ax(t + 1) + c(t). \quad (25)$$

На матрицу затрат A накладываются дополнительные требования: она должна содержать только неотрицательные элементы, быть неразложимой (т. е. не приводимой к квазитреугольному виду), примитивной (не допускающей циклического разложения, подробнее см. [69]) и удовлетворять условию (24).

Если все это выполнено, то матрица A имеет число Фробениуса–Перрона $\lambda(A)$ (неотрицательное вещественное собственное число матрицы A , превышающее по модулю все остальные ее собственные числа) и соответствующий ему собственный вектор η такие, что верны строгие неравенства: $\lambda(A) > 0$, $\lambda(A) < 1$, $\eta > 0$. Это означает, что соответствующая (25) однородная система

$$x(t) = Ax(t + 1),$$

имеет решение

$$x(t) = \left(\frac{1}{\lambda(A)} \right)^t \eta, \quad (26)$$

которое при всех t остается положительным и является сбалансированно растущим.

Если начальный вектор x_0 удовлетворяет условию $x_0 = \omega \eta$, где ω — произвольная положительная константа, то выходящее из него решение $x(t)$ стремится к (26). Если же начальный вектор не удовлетворяет данному условию, то для модели Неймана можно было бы доказать магистральную теорему, однако в модели Никайдо этого не удастся. Объясняется это тем, что в неймановской модели существовало целое множество решений, удовлетворяющих неравенству (20), в случае же Никайдо решение $x(t)$ определено единственным образом (по начальному условию), поскольку обязано удовлетворять жесткому равенству $x(t) = Ax(t + 1)$. В результате вектор $x(t)$ не только не стремится к (26), но и, как правило, становится отрицательным, по крайней мере по одной из компонент.

Если доказана сходимость ряда

$$\sum_{v=0}^{\infty} A^v c(v)$$

(а она имеет место для всех ограниченных функций $c(t)$, а также для $c(t)$, являющихся полиномами), то решение однородной системы (25) может быть записано в виде:

$$x(t) = \left(\frac{1}{\lambda(A)} \right)^t \eta + \sum_{v=0}^{\infty} A^v c(v + t), \quad (27)$$

причем это решение при соответствующим образом выбранных начальных условиях асимптотически стремится к сбалансированно растущему решению (26) однородной системы. Однако если начальные условия произвольны, то, как и для однородного случая, стремление решения к (26) не имеет места. Кроме того, для нахождения решения $x(t)$ на любом конечном интервале по формуле (27) обязательно требуется знать значения вектора потребления на бесконечном интервале, что и практически вряд ли осуществимо, и экономически бессмысленно. Таким образом, хотя

на основе версии Никайдо леонтьевской модели можно проследить кое-какие тенденции развития экономики, относиться к такой модели следует с известной долей осторожности.

1.3.4.3. ДИНАМИЧЕСКИЕ ЛЕОНТЬЕВСКИЕ МОДЕЛИ

Автором наиболее плодотворной модификации модели Леонтьева стал сам В. Леонтьев (1949), который ввел динамику в межотраслевую модель не с черного хода, как его предшественники, а в явном виде. Впрочем, многие экономисты и после этого продолжали игнорировать дифференциальные зависимости в межотраслевых моделях. Выпуск в каждой из отраслей расходуется, по Леонтьеву [54], не только на воспроизводство и потребление, но также на увеличение запаса товаров:

$$x_i = \sum_{j=1}^n a_{ij} x_j + \sum_{j=1}^n \dot{S}_{ij} + c_i \quad i = \overline{1, n},$$

где S_{ij} — запас товаров, произведенных i -отраслью и используемых j -отраслью. Введя гипотезу линейности запасов $S_{ij} = b_{ij} x_j$, получим окончательный вид динамической леонтьевской модели:

$$x = Ax + B\dot{x} + c. \quad (28)$$

Иногда вектор приростов выпусков \dot{x} заменяют его разностным аналогом. Коэффициенты матрицы удельных капитальных вложений B неотрицательны, однако Леонтьев акцентировал внимание на необратимости процесса накопления, в связи с чем при $\dot{x}_j < 0$ должен обнуляться соответствующий коэффициент капиталоемкости: $b_{ij} = 0$, $i = \overline{1, n}$. На основе этого Леонтьевым была разработана теория многофазового процесса.

Сейчас динамическую модель (28) интерпретируют несколько иначе: к ней приходят на основе соотношений (9) и (17), распространенных на векторный случай, и $B\dot{x}$ рассматривается как вектор инвестиций, а коэффициенты матрицы A изменяются с учетом добавления к ним коэффициентов амортизации. Таким образом, леонтьевская

модель (28) является динамической, по Харроду (в ней ненулевые инвестиции), и, в отличие от неймановской модели, активно пользуется плодами кейнсианской экономической мысли. Нетрудно видеть, что при $n = 1$ получается одна из модификаций модели Харрода–Домара, так что динамическую модель Леонтьева можно рассматривать как распространение результатов Харрода–Домара на случай n отраслей.

Однако модель (28) оказалось проще сформулировать, чем найти и проинтерпретировать ее решение. Оно обычно ищется в предположении $\det B \neq 0$, однако при большом количестве отраслей это условие выполняется редко. Как правило, решение ищут для частных случаев системы (28): для «закрытой» модели (без потребления) [47], для потребления, растущего экспоненциально по всем отраслям с одинаковым темпом [22], и т. п.

Задача нахождения растущего решения динамической модели (28) может быть интерпретирована и как задача стабилизации [11, 21]. Потребление предполагается линейно зависящим от выпуска продукции: $c = Ux$, выступая таким образом в роли управления. Матрица U составляется так, чтобы любое решение замкнутой этим управлением системы стремилось к сбалансированно растущему. Интересной представляется также дифференциально-разностная модификация [37] модели (28), учитывающая запаздывания инвестиций и выпусков, а также различия в длине производственного цикла для разных отраслей:

$$x_i(t) = \sum_{j=1}^n [a_{ij}x_j(t + \tau_j) + b_{ij}\dot{V}_j(t + h_j)] + c_i(t) \quad i = \overline{1, n},$$

$$x(t) \leq V(t),$$

здесь учтена также неполная загруженность производственных мощностей V_j . Пример такой модели для двух отраслей рассмотрен еще Леонтьевым [54] при помощи преобразований Лапласа, однако он носит сугубо иллюстративный характер.

ГЛАВА ВТОРАЯ

ЛИНЕЙНЫЕ ДИНАМИЧЕСКИЕ МОДЕЛИ

2.1. ВВЕДЕНИЕ

Настоящая глава посвящена моделированию динамических процессов с целью получения численного результата, или, как принято говорить, количественному прогнозированию динамического процесса. В ней также рассматриваются управляемые экономические процессы. Исходным материалом является некоторый набор данных о процессе, наблюдаемых в дискретные моменты времени. Далее будут описаны различные подходы к изучению этих данных и к формированию прогноза на их основе.

Большой раздел математической статистики — анализ временных рядов — ставит перед собой сходные задачи, но он уделяет основное внимание статистической стороне. Здесь мы надеемся привлечь внимание к поиску наиболее подходящих функциональных зависимостей. Конечно, мы не сможем избежать использования вероятностных характеристик данных, но постараемся делать это в минимальной степени.

В качестве математических соотношений в этой главе основными являются линейные разностные и дифференциальные уравнения. Они являются наиболее изученными и широко используемыми.

2.2. АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

В этом параграфе рассматриваются некоторые способы непосредственного, прямого моделирования скалярного временного ряда. В этом направлении исследователями накоплен огромный опыт, и, тем не менее, к нему стоит

возвращаться, поскольку обычные предположения о поведении ряда являются следствием математических возможностей, а не причинно-следственных связей системы, породившей ряд данных. Ниже обсуждаются методы выделения класса функций, из которого находится тренд. Решается задача об оптимальном интервале идентификации.

2.2.1. НАБОР СКАЛЯРНЫХ ДАННЫХ

Пусть у нас имеется некоторый набор скалярных данных, характеризующих какую-либо экономическую переменную x . Они могут быть разного происхождения. Например, пусть данные — скалярный результат повторяющегося эксперимента (число забракованных изделий на производстве, число лесных пожаров на территории страны в течение года и т. д.), записанные в виде таблицы или графически, где на оси абсцисс откладывается номер наблюдения или эксперимента (рис. 3).

Тогда можно записать $x = x(i)$. Предположим, что наблюдения или эксперимент (теоретически) можно проводить бесконечное число раз. Величину x называют случайной, если не

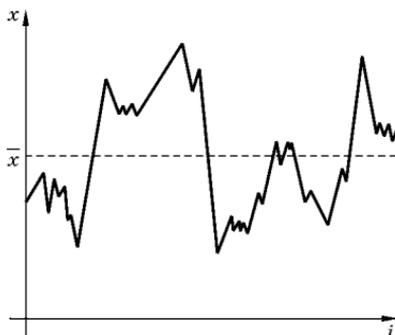


Рис. 3
Набор скалярных данных $x = x(i)$

существует правила, по которому в i -м наблюдении величина x приняла бы точно значение x_i . Пусть известно множество Ω , которому принадлежит значение величины x в i -м наблюдении. Если множество Ω конечно (или счетное), то рассматривают частоту, с которой случайная величина x принимает значение $x_i \in \Omega$ за N наблюдений. Предел этой частоты при $N \rightarrow \infty$ называется вероятностью, с которой x примет значение $x_i \in \Omega$. Обозначают определенную так вероятность $P_x(x = x_i)$. Если Ω более чем счетное, то для описания случайной величины применяют так называемую функцию распределения вероятностей для любых

вещественных ξ . Таким образом, знание функции распределения вероятностей случайной величины x не позволяет предсказать значение случайной величины при очередном наблюдении, но, в общем, для большого числа экспериментов о поведении случайной величины x можно сказать, что:

- ее среднее значение может быть вычислено как математическое ожидание

$$\bar{x} = \int_{\Omega} xf(x)dx = E(x),$$

где $f(x)$ — плотность вероятности ($f(x) = F'(x)$);

- разброс значений вокруг среднего характеризуется дисперсией

$$\sigma^2 = \int_{\Omega} (x - \bar{x})^2 f(x)dx;$$

а также получить, если это необходимо, множество других параметров, присущих нашей случайной величине.

К сожалению, и такая информация, как функция распределения вероятностей, чаще всего недоступна исследователю: он может опираться только на конечную реализацию $\{x_1, \dots, x_N\}$ случайной величины. Поэтому в дальнейшем, когда будет рассматриваться случайная величина, мы будем без специальных предположений об одинаковом распределении и независимости считать, что математическое ожидание равно среднему арифметическому, а дисперсия $D(x)$ равна разности среднего от квадратов и квадрата среднего:

$$D(x) = \sigma^2 = E(x^2) - (E(x))^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 - \left(\frac{1}{N} \sum_{i=1}^N x_i \right)^2.$$

2.2.2.

ХАРАКТЕРИСТИКИ МОДЕЛИ ВРЕМЕННОГО РЯДА

Теперь предположим относительно набора чисел $\{x_1, \dots, x_N\}$, что это наблюдения за некоторой экономической переменной в последовательные моменты времени, пусть для начала считается, что наблюдения делаются через равные промежутки времени. Тогда в каждый момент време-

ни $i = 1, \dots, N$, величина x_i должна рассматриваться как самостоятельная случайная величина с функцией распределения $F(i, x_i)$, зависящей от момента времени i . Более того, естественно считать, что наблюдаемая экономическая переменная только частично является случайной: она как-то зависит от времени, но мы не знаем как, а случайность появляется из-за погрешностей измерения или неучтенных случайных факторов.

Достаточно общей математической моделью служит модель вида:

$$x_i = u(i) + v_i, \quad (29)$$

где первое слагаемое правой части рассматривается как полностью детерминированная последовательность $\{u(i)\}$, которую можно назвать систематической составляющей; второе слагаемое — случайная последовательность $\{v_i\}$, подчиняющаяся некоторому вероятностному закону. Здесь различное обозначение аргумента у переменных вызвано тем, что мы хотим подчеркнуть происхождение переменной. Так, для реально существующей, но неизвестной нам функции $x(t)$, $t \in (0, N)$ наблюдение (измерение) ее в момент времени $t = i$ дает результат x_i , который совсем не обязательно равен $x(t)$. Иногда для упомянутых выше слагаемых используют соответственно термины сигнал и шум. Такое разделение носит теоретический характер, поскольку, измеряя, например, процентную ставку по кредитам $x(t)$, $t \in (0, N)$ в моменты времени $i = 1, \dots, N$, мы не можем повторить эти действия многократно для того, чтобы убедиться в систематичности $\{u(i)\}$ и случайности $\{v_i\}$. Точный смысл указанного разложения зависит не только от самих данных, но частично и от того, что понимается под повторением эксперимента, результатом которого являются эти данные. Если бы заранее знать, что $E(v_i) = 0$ для всех i , то можно было бы считать $\{v_i\}$ погрешностью измерения, но в экономических задачах и измерения (т. е. статистическая или бухгалтерская отчетность) могут быть предметом чьих-либо интересов и, следовательно, условие нулевого математического ожидания не выглядит достоверным.

Фактически математической моделью временного ряда $\{x_1, \dots, x_N\}$ является такая функция $u(t)$, при которой для соответствующих (29) $\{v_i\}$ будет выполнено $E(v_i) = 0$ и $D(v_i) \rightarrow \min$. Хорошо было бы при этом, чтобы функции распределения погрешностей $\{v_i\}$ не зависели от i и были бы нормальными. Практически это никогда не достижимо, но данное определение ценно тем, что теперь ясно, к чему надо стремиться. Все дальнейшее обсуждение моделирования временных рядов с целью составления прогноза исходит из этого определения, т. е. если даже специально не оговариваются необходимые предположения, все равно имеем их в виду.

Целью моделирования является предсказание значений $\{x_{N+1}, \dots, x_{N+L}\}$, где L — длительность прогноза. Очевидно, что, в лучшем случае, при большом количестве спорных предположений можно получить оценку функции распределения предсказания x_{N+1} (и следующих за ним), а не само значение. Тогда, в зависимости от задачи, интерес представляют:

- $E(x_{N+1})$, если прогноз делается из теоретических соображений, т. е. если исследователь просто хочет знать, что будет дальше;
- $E^+(x_{N+1})$ — верхний предел соответствующего доверительного интервала, если исследователь считает для себя более выгодным превысить прогноз (например, если речь идет о покупке товара по цене x);
- $E^-(x_{N+1})$ — нижний предел соответствующего доверительного интервала, когда исследователь опасается превысить в прогнозе будущее реальное значение (например, при продаже товара или при прогнозе налоговых сборов в бюджет страны).

Все три цели моделирования могут быть достигнуты, если нам известна функция распределения предсказания x_{N+1} .

Делая различные предположения о характере временного ряда, можно выделить, вообще говоря, два основных направления составления модели:

- влияние временного параметра проявляется только в систематической составляющей. Предполагается, что

математическое ожидание случайной составляющей тождественно равно нулю, дисперсия равна некоторой постоянной и что значения $\{v_i\}$ в различные моменты времени не коррелируют. В этом случае функцию $u(t)$, $t \in (0, N)$, выбирают из некоторого семейства базовых функций с каким-либо функционалом качества выбора;

- если влияние времени проявляется только в случайной составляющей, то используют математический аппарат стохастических процессов, в частности авторегрессии или скользящего среднего.

Рассмотрим эти два направления более предметно. В первом случае главным местом является выбор параметрического класса функций, из которого будет наилучшим образом выделен тренд (систематическая составляющая). Что такое «наилучшим образом», мы обсудим чуть позже, а пока остановимся на выборе класса функций. Нам необходимо, глядя на последовательность данных и выбирая класс функций, сопоставить причинно-следственные связи, которые могут быть наложены на данный набор чисел (фактически это установление того инварианта, который обсуждался в разделе об экстраполяции).

Если не ясна причинно-следственная связь, то хотя бы экономическая сущность величин должна определять искомый класс. Например, ставка процента по кредиту, очевидно, не может быть отрицательной, ее всплески должны иметь крутой подъем и сравнительно пологий спуск (если это не было ошибочным действием банка), ставка по проценту должна коррелировать с показателем инфляции и т. д.

Конкретный выбор определяется формой искомых функций, однако некоторые простые правила следует иметь в виду. Полиномы обычно хорошо ведут себя при интерполяции, но за пределами промежутка идентификации полином резко растет (или убывает) до бесконечности, причем тем быстрее, чем длиннее анализируемый ряд. Класс ограниченных функций может состоять из тригонометрических, причем монотонность данных может обеспечиваться монотонной частью синуса или косинуса, т. е. не надо избегать сверхнизких частот.

В качестве функционала, определяющего оптимальность выбора, остановимся на методе наименьших квадратов. Сначала опишем линейно-квадратичный случай, т. е. когда неизвестные параметры, задающие класс функций, входят в выражение функций линейно:

$$u(t) \in \left\{ \sum_{i=1}^p \beta_i u_i(t), \beta_i \in R^1, i = 1, \dots, p \right\}.$$

Рассмотрим некоррелированные случайные величины $\{x_1, \dots, x_N\}$, средние и дисперсии которых выражаются соотношениями:

$$E(x_t) = \sum_{i=1}^p \beta_i u_i(t), \quad t = 1, \dots, N, \quad (30)$$

$$E(x_t - E(x_t))^2 = \sigma^2, \quad t = 1, \dots, N.$$

Используя векторные обозначения:

$$B = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_p \end{pmatrix}, \quad U_t = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \dots \\ u_p(t) \end{pmatrix}, \quad t = 1, \dots, N,$$

математическое ожидание анализируемого ряда случайных величин можно записать в виде

$$E(x_t) = B^T U_t, \quad t = 1, \dots, N.$$

Будем обозначать через вектор \hat{B} оценку вектора B , представляющую собой решение системы нормальных уравнений,

$$A \hat{B} = C, \quad (31)$$

в которой

$$A = \sum_{t=1}^N U_t U_t^T, \quad C = \sum_{t=1}^N x_t U_t.$$

Решение системы (31) минимизирует сумму

$$\sum_{t=1}^N (x_t - B^T U_t)^2$$

на множестве всех p -мерных векторов B и называется оценкой метода наименьших квадратов вектора B . Если

$(p \times p)$ — матрица A невырожденная, то вектор $\hat{B} = A^{-1}C$. В противном случае необходимо рассматривать псевдообратную матрицу, что мы и сделаем в дальнейшем, когда это будет совершенно необходимо, а пока предположим, что A невырожденная.

Несмещенная оценка s^2 для σ^2 может быть получена (при $N > p$) из соотношения

$$(N - p)s^2 = \sum_{t=1}^N (x_t - \hat{B}^T U_t)^2 = \sum_{t=1}^N x_t^2 - \hat{B}^T A \hat{B}.$$

Оценка метода наименьших квадратов \hat{B} является несмещенной оценкой вектора B : $E(\hat{B}) = B$, и имеет ковариационную матрицу

$$E((\hat{B} - B)(\hat{B} - B)^T) = \sigma^2 A^{-1}.$$

Теорема Гаусса–Маркова утверждает, что компоненты вектора \hat{B} являются наилучшими линейными несмещенными оценками соответствующих компонент вектора B , т. е. каждый элемент вектора \hat{B} имеет дисперсию, наименьшую среди дисперсий всех несмещенных оценок соответствующего элемента вектора B , линейных по переменным $\{x_1, \dots, x_N\}$.

Если случайные величины $\{x_1, \dots, x_N\}$ независимы и нормально распределены, то \hat{B} является оценкой максимального правдоподобия вектора B . Оценкой максимального правдоподобия для σ^2 служит в этом случае величина $(N - p)s^2/N$. При этом оценка \hat{B} распределена по многомерному нормальному закону $N(B, \sigma^2 A^{-1})$ со средним B и ковариационной матрицей $\sigma^2 A^{-1}$.

Вторая ситуация характеризуется зависимостью от времени случайной составляющей в (29), в этом случае говорят о стохастическом процессе и о моделировании временного ряда в каких-либо предположениях об изменении функции распределения с течением времени. Процессы, вероятностная структура которых не изменяется со временем, называются **стационарными**.

Простейшими для изучения процессами являются или стационарные, или такие, в которых случайная составляющая может считаться стационарным процессом. Широко

распространенной моделью приближения временного ряда оказывается процесс **авторегрессии**, или **стохастическое разностное уравнение**,

$$x_t = \sum_{k=1}^p \alpha_k x_{t-k} + v_t, \quad t = p, p+1, p+2, \dots \quad (32)$$

где v_t — случайная величина с $E(v_t) = 0$, $E(v_t, v_s) = 0$ и $E(v_t^2) = \sigma^2$. Известное преобразование переводит разностное уравнение порядка p в систему того же порядка. Обозначая

$$y_t = \begin{pmatrix} x_t \\ x_{t-1} \\ \dots \\ x_{t-p+1} \end{pmatrix}, \quad w_t = \begin{pmatrix} v_t \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad G = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_{p-1} & \alpha_p \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

получим векторное уравнение

$$y_t = G y_{t-1} + w_t,$$

которое можно рассматривать теперь независимо от скалярного уравнения (32) с произвольной матрицей G . Записывая последнее уравнение для убывающих номеров t и подставляя одно в другое, составим соотношение

$$y_t = G^{s+1} y_{t-s-1} + \sum_{k=0}^s G^k w_{t-k}, \quad t, s > 0.$$

Из него непосредственно следует, что если собственные числа матрицы G (т. е. корни уравнения $\det(\lambda I - G) = 0$) по модулю меньше единицы, то влияние предшествующих состояний системы на y_t уменьшается по мере роста s , возмещаясь влиянием случайной составляющей. Так, в пределе получим:

$$y_t = \sum_{k=0}^{\infty} G^k w_{t-k}.$$

Таким образом, при сделанных относительно матрицы G предположениях процесс авторегрессии (32) сводится к бесконечному скользящему среднему. Рассмотрим более подробно проблему сглаживания временного ряда и скользящее среднее как один из способов сглаживания.

2.2.3. СГЛАЖИВАНИЕ ВРЕМЕННЫХ РЯДОВ

Выбор базисных функций для представления тренда $u(t)$ в (29) является своего рода искусством. Глядя на данные, мы не можем точно отделить тренд от случайных ошибок измерения или случайных неизвестных факторов. Поэтому бывает полезно пропустить реальные данные через некоторый «фильтр», чтобы часть случайных добавок к данным исчезла. Оставшаяся часть информации в этом случае будет использована в подборе класса базисных функций.

Геометрическое представление временного ряда на плоскости (t, x) изображается набором точек. Однако через каждую точку следовало бы провести вертикальный отрезок такой длины, которая соответствует области значений случайной добавки v_i к $u(t)$ при $i = t$. Таким образом, на плоскости (t, x) вместо набора точек будет изображена полоса, содержащая все точки (рис. 4).

Границы (края) полосы будут достаточно гладкими, если считать, что случайные добавки одинаково распределены, не коррелируют и имеют нулевое среднее. И, конечно же, предполагается, что тренд — достаточно гладкая функция. Ширина полосы тем больше, чем больше дисперсия $D(v_i)$, и равна нулю, если $\sigma^2 = 0$. Так как восстановить тренд по наблюдениям со случайными неизвестными ошибками невозможно, то примем, что нашей целью

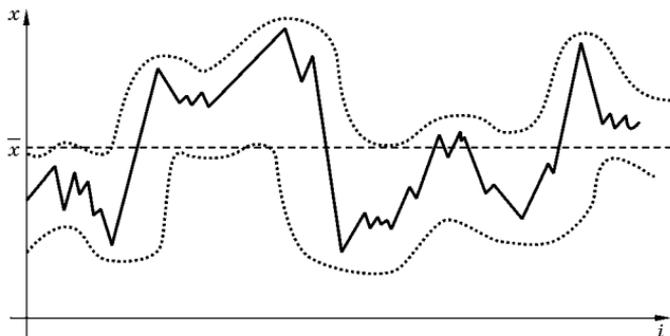


Рис. 4

Полоса, каждый вертикальный отрезок которой является множеством возможных значений $x(i)$

является построение такого гладкого тренда, чтобы дисперсия соответствующих отклонений была минимальна. Вообще говоря, этого мало. Надо еще нулевое среднее и некоррелированность ошибок, но в качестве первого приближения для оценки класса функций такое построение было бы весьма полезно. Отметим, что если в качестве тренда рассмотреть кусочно-линейную линию, соединяющую все точки временного ряда, то вычисленная дисперсия (оценка дисперсии) будет равна нулю, но в этом случае о гладкости тренда не может быть и речи. С другой стороны, если принять, что тренд равен постоянному среднему арифметическому значению по всем данным временного ряда, то будет соблюдено условие гладкости и нулевого среднего для отклонений v_i , но маловероятно, что оценка дисперсии будет наименьшей.

Построить описанную выше полосу не представляется возможным, так как ни тренд, ни распределения ошибок не известны, у нас есть только временной ряд $\{x_1, \dots, x_N\}$. Сглаженные данные (пропущенные через сглаживающий фильтр) позволяют сделать визуальный анализ данных как на предмет тренда, так и оценить функцию распределения ошибок.

Процедура сглаживания может основываться на различных принципах. Наиболее употребляемый алгоритм использует скользящее среднее, которое определяется следующим образом. Ряд данных $\{x_1, \dots, x_N\}$ заменяется другим, более коротким рядом:

$$x_i^* = \sum_{s=-m}^m c_s x_{i+s}, \quad i = m+1, \dots, N-m. \quad (33)$$

Положительные коэффициенты c_s являются весами усреднения и удовлетворяют соотношению

$$1 = \sum_{s=-m}^m c_s,$$

а m характеризует длину отрезка временного ряда, участвующего в формировании среднего. В случае равных коэффициентов c_s получим среднее арифметическое при $c_s = 1/(2m+1)$.

Для новой, «сглаженной» последовательности $\{x^*_{m+1}, \dots, x^*_{N-m}\}$ вследствие (29) и (33) будет справедливо выражение

$$x_i^* = \sum_{s=-m}^m c_s u(i+s) + v_i^*, \quad i = m+1, \dots, N-m,$$

где

$$v_i^* = \sum_{s=-m}^m c_s v_{i+s}.$$

Как и прежде предполагается, что $E(v_i) = 0$, $E(v_i^2) = \sigma^2$, $E(v_i, v_j) = 0$, $i \neq j$. В силу этого v_i^* имеет дисперсию

$$\sigma^2 \sum_{s=-m}^m c_s^2.$$

Веса c_s выбираются так, чтобы дисперсия величины v_i^* была значительно меньше, чем дисперсия величины v_i . Если значения $u(i+s)$, $s = 0, \pm 1, \dots, \pm m$ близки к значению $u(i)$, а гладкость функции $u(t)$ это предполагает, то и $E(x_i^*)$ близко к $u(t)$. Поэтому скользящее среднее, или сглаженный ряд $\{x^*_{m+1}, \dots, x^*_{N-m}\}$, имеет примерно ту же последовательность математических ожиданий, что и $\{x_1, \dots, x_N\}$, но зато меньшую дисперсию. Однако последовательные члены сглаженного ряда в общем случае являются коррелированными. Именно:

$$E(v_i^* v_{i+h}^*) = \sum_{s=-m}^m \sum_{r=-m}^m c_s c_r v_{i+s} v_{i+h+r} = \begin{cases} \sigma^2 \sum_{s=m+h}^m c_s c_{s-h}, & h = 0, 1, \dots, 2m. \\ 0, & h = 2m+1, \dots \end{cases}$$

Следовательно, использовать сглаженный таким образом временной ряд как модель для прогнозирования не рекомендуется, но некоторое представление о тренде и распределении ошибок скользящее среднее дает.

2.2.4.

ОБ ИНТЕРВАЛЕ ИДЕНТИФИКАЦИИ

Еще один вопрос, возникающий при построении модели временного ряда, связан с длиной интервала идентификации наилучших параметров модели для заданного класса функций. Конечно, это касается только тех задач, в которых нет недостатка в данных. Например, макроэкономические

параметры Великобритании могут быть получены за сотни лет: имеется в виду ВВП, денежная масса, индексы жизни и инфляции, процент по кредитам и т. д. Прогнозируя погодные явления, также можно собрать информацию за десятки лет, но для рабочего алгоритма, реально используемого на практике, исследователь вряд ли захочет создавать такой неудобный аппарат, когда всякий раз при прогнозе необходимо обращаться ко всей базе данных. К тому же большая база данных требует постоянного внимания и обслуживания, что влечет за собой увеличение стоимости прогноза. Поэтому стремление оптимизировать объем базы данных по издержкам в эксплуатации естественно при создании работающих алгоритмов прогнозирования. Обсудим этот вопрос подробнее.

В соответствии с введенными в (29) и (30) обозначениями модель временного ряда имеет систематическую и случайную составляющие. Забудем на некоторое время про случайное слагаемое и будем интерпретировать задачу следующим образом: пусть некоторая функция $x(t)$ определена при $t \in (0, N)$. Будем считать, что $x(t)$ — математическая модель данных. Она неизвестна. В нашем распоряжении имеется конечное число измерений $\{x(i)\}$, $i = 1, \dots, N$. Кроме того, исследователь ограничен некоторым классом функций $\{u(t, B)\}$, где $t \in (0, N)$ и B — векторный параметр, задающий конкретного представителя из класса $\{u(t, B)\}$. Так как целью моделирования является прогноз ряда данных, т. е. вычисление $\{x(i)\}$, $i = N + 1, \dots$, то задача сводится к выбору такого параметра B , чтобы соответствующая функция $u(t, B)$ была наилучшей аппроксимацией модели $x(t)$. Если бы мы хотели использовать все данные, то определение B можно было бы осуществить минимизируя функционал:

$$S(B) = \int_0^N (x(t) - u(t, B))^2 dt.$$

Однако в этом случае искомый параметр B зависел бы от всего объема имеющейся информации. Модифицируем задачу. Рассмотрим последнюю часть информационного массива (самую позднюю информацию) длины M , а в ка-

честве критерия близости выберем среднее из квадратов отклонений:

$$S(B, M) = \frac{1}{M} \int_{N-M}^N (x(t) - u(t, B))^2 dt.$$

Если $M = N$, то для прогноза используется весь временной ряд, если же $M = 1$, то прогноз строится по одной последней точке информации. Теперь можно, изменяя значение M , выбрать оптимальный интервал идентификации модели.

Утверждение. Для заданного временного ряда $\{x(i)\}$, $i = 1, \dots, N$, и заданного класса аппроксимации $\{u(t, B)\}$ оптимальное относительно $S(B, M)$ решение задачи дается решением системы уравнений:

$$\int_{N-M}^N (x(t) - u(t, B)) \frac{Du(t, B)}{DB} dt = 0,$$

$$\frac{1}{M} \int_{N-M}^N (x(t) - u(t, B))^2 dt = [x(N - M) - u(N - M, B)]^2. \quad (34)$$

Действительно, производная по B от $S(B, M)$ дает левую часть первого равенства. Выразим из него формально $B = B(M)$ и подставим в $S(B, M)$, а затем от получившегося выражения возьмем произвольную по M :

$$\frac{\partial S}{\partial M} = -\frac{1}{M^2} \int_{N-M}^N [...]^2 dt + \frac{1}{M} [x(N - M) - u(N - M, B(M))]^2 +$$

$$+ \frac{1}{M} \int_{N-M}^N [x(t) - u(t, B(M))] \left[-2 \frac{Du}{DB} \frac{dB}{dM} \right] dt.$$

Последнее слагаемое равно нулю в силу первого уравнения из (34) и независимости B от t . Приравнявая полученную производную к нулю и умножая на M , подтверждаем второе равенство в (34).

Отметим, что доказанное утверждение дает нам необходимые условия оптимальности. Кроме того, зависимость $S(B, M)$ существенно нелинейная по M , и поэтому следует ожидать не единственность решений системы (34).

Разностный аналог системы (34) достигается простой дискретизацией соответствующих функций и заменой интегралов на суммы:

$$\sum_{i=N-M}^{N-1} [x(i) - u(i, B)]^2 \frac{Du(i, B)}{DB} = 0,$$

$$\frac{1}{M} \sum_{i=N-M}^{N-1} [x(i) - u(i, B)]^2 = [x(N-M) - u(N-M, B)]^2.$$

Так, например, для функции $x(t) = t^2$ промежутка времени $t \in (0, 1)$ и класса аппроксимирующих функций $u(t, B) = bt$ получим после несложных вычислений на компьютере, что оптимальный интервал идентификации согласно формулам (34) равен 0,775, а оптимальное значение параметра равно 0,757. Кстати, в данном примере оптимальное значение оказалось единственным.

Полученные аналитические формулы и возможность построить вычислительный алгоритм не должны ввести нас в заблуждение, препятствуя ясному взгляду на сделанное: а так ли уж нас интересует минимизация функционала $S(B, M)$ по B и M ? На самом деле, мы бы хотели получить наилучший прогноз, т. е. аппроксимацию будущего, а не анализ прошлого. Трансформируем задачу так, чтобы оптимальный параметр выбирался по прошлому (иначе мы не сможем использовать модель в реальных условиях), но оптимальный интервал обеспечивал бы нам наилучший возможный прогноз.

Как и раньше, пусть N — длина интервала данных о динамическом процессе, h — заданная длина прогноза, M — искомая длина оптимального интервала идентификации модели (или, более точно, длина известного набора данных, по которым строится оптимальный вектор параметров B).

Для произвольного $t_0 \in [M, N - h]$ вычислим оптимальный вектор параметров из условия минимума следующего функционала:

$$S(B, M, t_0) = \frac{1}{M} \int_{t_0-M}^{t_0} [x(t) - u(t, B)]^2 dt.$$

Затем определим M , доставляющую наименьшее значение функционалу качества прогноза:

$$W(M) = \max_{t_0 \in [M, N-h]} \frac{1}{h} \int_{t_0}^{t_0+h} [x(t) - u(t, B(M, t_0))]^2 dt. \quad (35)$$

Заметим, что в данной постановке открывается путь к задаче оптимизации интервала прогноза по заданным наборам данных и класса аппроксимирующих функций.

2.3. МНОГОМЕРНЫЕ ВРЕМЕННЫЕ РЯДЫ

В этом параграфе мы в полной мере воспользуемся преимуществами метода наименьших квадратов, чтобы не обращать больше внимания на случайную составляющую временных рядов, которые будут являться исходным материалом для построения моделей. Так как в конечном итоге в большинстве случаев относительно случайных ошибок делаются такие предположения, что метод наименьших квадратов дает наилучшие несмещенные оценки, то будем впредь предполагать, что либо ошибок просто нет (и строится функция, ближайшая к данной), либо их случайный характер подавляется функционалом.

Далее будут рассмотрены способы моделирования непрерывных и дискретных процессов, идентификация коэффициентов моделей и содержательные экономические примеры. Изложенное в пунктах 2.3.3 и 2.3.6 в основном опирается на результаты А. В. Степанова [88].

2.3.1. ЗАДАЧА О ФОРВАРДНЫХ ОПЕРАЦИЯХ

Форвардными операциями называют [12] внебиржевые срочные валютные сделки, совершаемые банками и промышленно-торговыми корпорациями по телефону, интернету или телексу на договорной основе. С 1984 г. практикуются форвардные операции с кредитными инструментами — так называемые «соглашения о будущей ставке», которые представляют собой межбанковские срочные соглашения о взаимной компенсации убытков от изменения процентных ставок по депозитам до 1 года. Такие сделки

выступают альтернативой проводимым на бирже фьючерсным и опционным операциям с финансовыми инструментами. Поскольку такой контракт не является стандартным, то в момент его заключения стороны согласовывают все основные позиции: сроки, цены, объемы и т. д. Таким образом, для большинства крупных (с мировым размещением производства) компаний возникает задача о прогнозировании курса иностранной валюты. Крупные компании, такие, как «Самсунг», «Форд», «Шелл», традиционно имеют широкую сеть предприятий за границей. В большинстве случаев это компании высокоразвитых стран. Так, «Форд» еще с начала века имеет предприятия в Германии, Канаде, Великобритании, Италии и Франции. В настоящее время очень велика кооперация с Японией. Таким образом, проблема обмена деньгами стоит перед «Фордом» ежедневно. Кроме того, за последние годы усиливается тенденция привлечения среднего звена развитых стран к производству автомобилей и к созданию инфраструктуры обслуживания. В сферу интересов «Форда» вовлекаются Мексика, Индия, Бразилия, Беларусь, Китай, Россия и Румыния. По-видимому, этот список будет продолжен. Будем и далее условно считать, что обсуждается проблема некоторой американской компании А.

Каждый год, планируя финансовые операции, компания А предполагает истратить часть своего бюджета на покрытие расходов в указанных выше странах в валютах этих стран. Причем, так как планирование идет на год вперед, то необходимо оценить стоимость соответствующей валюты через год. Сама компания А международными финансовыми операциями не занимается, она делает это через один из мировых банков. Банк же за выполнение фьючерсной сделки с валютой берет предоплату, определяемую не только количеством обмененных денег, но и риском случайных колебаний в стоимости соответствующей валюты. Таким образом, улучшение прогнозирования стоимости валюты через год может дать значительную прибыль для компании А, а это, в свою очередь, может привести к снижению продажной цены на производимый товар или услугу и к повышению конкурентоспособности.

Рассмотрим сначала только одну страну, чью валюту компания А собирается купить через год. Пусть, например, это будет Италия. Проанализируем основные факторы, от которых зависит уровень обменного курса. При этом будем стремиться уменьшить количество переменных модели, захватывая большее количество объясняемых явлений.

Отметим одно важное свойство любого фьючерса: если нужно купить что-либо в будущем, то можно: 1) сделать это сейчас и сохранять товар до нужного времени; 2) отложить на это деньги сейчас и купить, когда придет время. В первом случае деньги, превращенные в товар сейчас, будут заморожены до будущего и компания понесет чистые потери, зато нет никакого риска. Второй случай предполагает возможность использовать необходимые деньги до будущего времени. Следовательно, цена покупки в будущем должна быть уменьшена (дисконтирована). Формально это записывается так: пусть $l(t)$ — цена лиры в долларах на момент t , $L(t)$ — цена фьючерса, $r(t)$ — ставка рефинансирования Центрального (федерального) банка для заемщиков, которыми являются коммерческие банки; v — объем покупаемой валюты в будущем. Текущее время примем $t = 0$ и временной момент будущего обозначим T . Тогда

$$L(0) = vl(T) \exp \left[- \int_0^T r(t) dt \right]. \quad (36)$$

Эта формула предполагает знание курса лиры $l(t)$ и ставки рефинансирования $r(t)$ в будущем: $t \in [0, T]$. Очевидно, необходимо уметь делать прогноз $l(t)$ и $r(t)$ по некоторой предыстории.

Теперь рассмотрим, какие факторы влияют на обменный курс лиры и доллара. Так как доллар — внутренняя денежная единица США, то на него влияют все макроэкономические параметры страны: спрос и предложение факторов производства, т. е. труда, капитала и природных ресурсов, а также производительность экономики, выражаемая ВВП. В данном случае рассматриваются страны с относительно стабильной политической и экономической жизнью, поэтому предполагается, что спрос и предложение

труда уравновешены достаточной зарплатой и небольшой безработицей. Давление капитала на экономику регулируется Центральным банком США с помощью ставки рефинансирования $r(t)$. Источники природных ресурсов не исчерпаны, иначе это привело бы к войнам, внешним конфликтам и т. п. Технологический уровень в экономике определяет производительность труда и, следовательно, ВВП. Таким образом, наиболее рациональным представляется избрать в качестве макроэкономических параметров ВВП США $G_A(t)$ и ставку рефинансирования $r_A(t)$. И то же самое для Италии: $G_I(t)$, $r_I(t)$.

В результате получаем вектор с пятью компонентами $\{G_A(t), r_A(t), G_I(t), r_I(t), l(t)\}$, для которого имеются в достаточном количестве данные за прошлые годы, и необходимо построить прогноз на один год для переменных $r_A(t)$, $l(t)$, чтобы воспользоваться формулой (36) и вычислить цену фьючерса. Так, возникает задача прогнозирования многомерного временного ряда. Один из способов ее решения — составление моделей двух выделенных переменных независимо от других (и друг от друга) по имеющимся данным, как скалярных временных рядов с использованием алгоритмов предыдущего параграфа. Будем считать этот способ крайним. Другой крайней формой моделирования будет попытка создать многомерную авторегрессионную модель некоторой степени:

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_{t-p},$$

где $X_t = \{G_A(t), r_A(t), G_I(t), r_I(t), l(t)\}$, p — степень авторегрессии. Возможны и промежуточные модели. Например, учитывая специальный характер задачи, модель можно формулировать как иерархическую структуру:

$$\begin{aligned} G_A &= G_A(t); & G_I &= G_I(t); \\ r_A(t) &= R_A(r_A(\cdot), G_A(\cdot)); & r_I(t) &= R_I(r_I(\cdot), G_I(\cdot)); \\ l(t) &= f(l(\cdot), r_A(\cdot), r_I(\cdot), G_A(\cdot), G_I(\cdot)), \end{aligned}$$

здесь все функции упомянутых выше переменных должны при построении модели конкретизироваться. В четвертой главе будут проанализированы несколько вариантов моделирования.

Таким образом, содержательной частью моделирования является построение линейной (как самой простой) системы дифференциальных или разностных уравнений. Выделим два принципиально различных подхода: первый основан на построении соответствующей линейной системы по минимальному объему информации, второй предполагает, что наблюдения за динамическим процессом производятся достаточно большое число раз, гораздо большее, чем размерность определяемой системы, и поэтому у исследователя имеется возможность не только создать модель, но как бы проверить ее адекватность на тех же данных.

2.3.2. АВТОРЕГРЕССИИ ПРИ МАЛЫХ ОБЪЕМАХ ИНФОРМАЦИИ

Предположим, что n -вектор состояния динамической системы (в предыдущем пункте это был пятимерный вектор $X_t = \{G_A(t), r_A(t), G_I(t), r_I(t), l(t)\}$) измеряется через временной интервал, принятый за единицу времени. В результате накапливается таблица наблюдений $\{X_0, X_1, \dots, X_N\}$, где $n < N$. Предположим, что причинно-следственный механизм динамического процесса состоит в линейной зависимости скорости изменения переменной от их текущего состояния, причем каждой от всех. В качестве модели динамики системы можно рассматривать линейные дифференциальные уравнения с решением $X(t)$:

$$\frac{d}{dt}X(t) = AX(t) + B, \quad (37)$$

где A — постоянная $n \times n$ -матрица, B — постоянный n -вектор. Наша задача состоит в идентификации A и B по наблюдениям $\{X_0, X_1, \dots, X_N\}$, $X_t = X(t)$. Предположим, что исходные наблюдения действительно удовлетворяют уравнению (37), и мы хотим восстановить A и B . Тогда решение поставленной задачи осуществляется следующим образом.

Из (37) следует, что

$$X_t = e^A \left(X_{t-1} + \int_0^1 e^{-As} B ds \right), \quad t = 1, 2, \dots, N,$$

или, вводя новые обозначения,

$$C = e^A, \quad (38)$$

$$D = e^A \int_0^1 e^{-As} B ds, \quad (39)$$

приходим к разностной системе

$$X_t = CX_{t-1} + D, \quad t = 1, 2, \dots, N. \quad (40)$$

Чтобы сначала избавиться от неизвестного вектора D , перейдем к новым векторам Y_t : $Y_t = X_t - X_{t-1}$, $t = 1, 2, \dots, N$. Новые векторы удовлетворяют разностной системе: $Y_t = CY_{t-1}$, $t = 1, 2, \dots, N$.

Используя таблицу наблюдений и последнее векторное равенство, получим матричное соотношение

$$\{Y_2, \dots\} = C\{Y_1, \dots\},$$

из которого при условии

$$\det\{Y_1, Y_2, \dots, Y_n\} \neq 0 \quad (41)$$

находим матрицу C :

$$C = \{Y_2, Y_3, \dots, Y_{n+1}\}\{Y_1, Y_2, \dots, Y_n\}^{-1}.$$

Вектор D соответственно выражается из уравнения (40):

$$D = X_t - CX_{t-1}, \quad t = 1, 2, \dots, N.$$

Теперь формулы (38)–(39) дадут нам искомые A и B :

$$A = \ln C, \quad B = \left(e^A \int_0^1 e^{-As} ds \right)^{-1} D. \quad (42)$$

Сделанные преобразования, хотя и имеют своим результатом формулы (42) в законченной форме, требуют ряда пояснений. При этом даже если считать, что наблюдения $\{X_0, X_1, \dots, X_N\}$ изначально удовлетворяют системе (37), т. е. $X(k) = X_k$, то и в этом случае с формулами (42) будут проблемы. Рассмотрим их подробнее. Очевидно, наименьшее значение N должно быть $(n + 1)$. Переход к переменным Y делает соответствующую разностную систему одно-родной:

$$\frac{d}{dt} Y_t = AY_t, \quad (43)$$

а для однородной системы справедливы следующие утверждения, необходимые для обоснования вводимых в алгоритм предположений. Так, условие (41) является в некотором смысле фундаментальным: из него следует, что ряд $\{Y_1, Y_2, \dots, Y_n\}$ для системы (43) находится в «общем положении», т. е. векторы Y_k не принадлежат никакому инвариантному подпространству матрицы A .

Утверждение 1. Пусть одна и та же траектория $Y(t)$ на промежутке $[a, b]$ описывается двумя системами линейных дифференциальных уравнений (43) и $\dot{Y} = \hat{A}Y$. Если существует набор чисел $\{t_k\}$, $a < t_1 < t_2 < \dots < t_n \leq b$, таких, что векторы $\{Y(t_1), Y(t_2), \dots, Y(t_n)\}$ линейно независимы, то $A = \hat{A}$.

Действительно, из условия совпадения траекторий следует, что для любого $t \in [a, b]$ имеет место равенство $(A - \hat{A})Y(t) \equiv 0$. Пусть произвольный вектор Z представлен в виде линейной комбинации

$$Z = \sum_{k=1}^n c_k Y(t_k),$$

что всегда возможно из-за предполагаемой линейной независимости $\{Y(t_k)\}$. Тогда $(A - \hat{A})Z = 0$. Предположим, что $A \neq \hat{A}$, т. е. $a_{ij} \neq \hat{a}_{ij}$ для некоторых i, j . Выбирая в качестве Z любой вектор с отличным от нуля элементом z_j на j -м месте, получим, что $(a_{ij} - \hat{a}_{ij})z_j \neq 0$, значит, $(A - \hat{A})Z \neq 0$. Это противоречие и доказывает утверждение.

Утверждение 2. Если выполнено условие утверждения 1, то для некоторого достаточно малого положительного числа Δ среди точек разбиения $[a, b]$ на Δ -интервалы $\tau_m = a + m\Delta$, $m = 1, \dots, N$, существует набор $\{Y(\tau_{m_k})\}$ линейно независимых векторов.

Действительно, условие линейной зависимости — это условие типа

$$\det\{Y(t_1), \dots, Y(t_n)\} = 0. \tag{44}$$

Следовательно, множество точек t , содержащее все t_k , удовлетворяющие формуле (44), будет замкнуто, так как определитель является непрерывной функцией компонент $Y(t_k)$, а $Y(t)$ — непрерывная вектор-функция на $[a, b]$.

Отсюда вытекает, что множество моментов t_k , для которых (44) не выполняется, обязательно открыто. Тогда, если (44) не выполнено для какого-то набора $\{t_k\}$, то оно не выполнено для любого другого набора $\{t_k^*\}$, связанного с t_k неравенством

$$\max_{k=1, \dots, n} |t_k - t_k^*| < \varepsilon,$$

где ε — достаточно малое число. Теперь осталось выбрать $\Delta < 2\varepsilon$, так как в этом случае в каждом интервале $t_k - \varepsilon < t < t_k + \varepsilon$ обязательно окажется точка τ_{m_k} .

Итак, считаем, что предположения о существовании n линейно независимых векторов в таблице и о равенстве интервалов наблюдения выполнены. Ссылаясь на соответствующую теорему в [19], можем утверждать, что вещественная матрица A будет решением уравнения (38) тогда и только тогда, когда C невырожденная и элементарные делители, соответствующие вещественным отрицательным собственным числам C , повторяются четное число раз. Про матрицу C с такими свойствами будем говорить, что она логарифмируема.

Матричный логарифм в (41) имеет неоднозначность двух типов [19]: счетную неоднозначность, проистекающую из периодичности функции $\exp z$, и континуальную неоднозначность, связанную со структурой оператора $\exp A$. Если последний оператор имеет простую структуру, то континуальной неоднозначности нет. В общем случае может быть приведен корректный алгоритм, учитывающий только счетную неоднозначность.

Рассмотрим одно общее утверждение о многозначности решения уравнения (38), показывающее, что для одной и той же матрицы C будет существовать счетное множество вещественных логарифмов. Будем предполагать далее, что C имеет простую структуру. Определим матрицу S , составленную из собственных векторов вещественной матрицы A , следующим образом: первыми в S стоят собственные векторы, соответствующие вещественным собственным числам матрицы A , а затем стоят парами комплексно-сопряженные векторы, соответствующие комплексно-сопряженным собственным числам (собственны-

ми векторами мы называем в данном случае соответствующие векторы жорданова базиса). Пусть диагональная матрица V определена так: первыми элементами диагонали пусть будут нули в таком количестве, сколько вещественных собственных чисел у матрицы A с учетом кратности; далее следуют парами по диагонали k_1 и $-k_1$, k_2 и $-k_2$ и т. д. по числу пар комплексно-сопряженных собственных чисел A (k_i — целые числа).

Утверждение 3. Если $\exp A_1 = \exp A$, то

$$A_1 = A + S \times 2\pi i \times VS^{-1}, \quad (i = \sqrt{-1}).$$

Здесь A и A_1 — вещественные матрицы, а S и V — определенные выше по A .

Действительно, по условию $A = S\Lambda S^{-1}$, где Λ — жорданова форма A . По определению функции от матрицы [19]

$$\exp A = S \exp \Lambda S^{-1}$$

и

$$Ln(\exp A) = S Ln(\exp \Lambda) S^{-1}.$$

Так как $\exp \lambda = \exp(\lambda + 2k\pi i)$, то на диагонали матрицы $Ln(\exp \Lambda)$, кроме главного значения Ln , должны быть слагаемые $2k\pi i$, где k — целое число, причем для каждого собственного числа — свое. Следовательно,

$$Ln(\exp \Lambda) = \Lambda_0 + 2\pi i \times \text{diag}(l_1, \dots, l_n).$$

Жорданова форма Λ_0 соответствует главному значению Ln и содержит все особенности структуры оператора $\exp A$.

Таким образом,

$$A_1 = S[\Lambda_0 + 2\pi i \times \text{diag}(l_1, \dots, l_n)]S^{-1},$$

$$A = S[\Lambda_0 + 2\pi i \times \text{diag}(\bar{l}_1, \dots, \bar{l}_n)]S^{-1}.$$

Значит,

$$A_1 = A + S[2\pi i \times \text{diag}(\bar{l}_1, \dots, \bar{l}_n)]S^{-1},$$

где \bar{l}_k — целые числа, но A и A_1 вещественные, следовательно, $\text{diag}(\bar{l}_1, \dots, \bar{l}_n)$ должна иметь вид V .

Действительно, пусть S_m — собственный вектор, соответствующий вещественному собственному числу. Тогда S_m — вещественный. Матрица

$$V_1 = S[2\pi i \times \text{diag}(\overline{l_1}, \dots, \overline{l_n})]S^{-1}$$

должна быть вещественной. Из соотношения $V_1 S_m = i \overline{l_m} S_m$ получаем, что $\overline{l_m}$ обязательно должно иметь нулевое значение.

Утверждение 4. Пусть таблица наблюдений Y_j «порождена» системой (43) с неизвестной матрицей A , наблюдения велись через равные интервалы времени $\Delta > 0$ и среди Y_j ($j = 0, 1, \dots, N > n$) имеется набор $\{Y_{m_k}\}$ линейно независимых векторов ($k = 1, \dots, n$). Тогда, если матрица $C = \{Y_{m_{k+1}}\}\{Y_{m_k}\}^{-1}$ логарифмируема, то матрица A может быть восстановлена с точностью до слагаемого $S \times 2\pi i V S^{-1} / \Delta$, где S — матрица, построенная из собственных векторов C способом, описанным выше; V — диагональная матрица, также описанная выше. При этом указанное слагаемое зависит от целочисленных параметров, число которых совпадает с числом пар комплексно-сопряженных собственных чисел матрицы C с учетом их кратности.

Действительно, пусть матрица A_1 — какое-либо решение уравнения (38). Тогда имеем:

$$Y_{m_{k+1}} = \exp(A_1 \Delta) Y_{m_k}, \quad k = 1, \dots, n,$$

но в силу (43) будет:

$$Y_{m_{k+1}} = \exp(A \Delta) Y_{m_k},$$

значит,

$$[\exp(A_1 \Delta) - \exp(A \Delta)] Y_{m_k} = 0.$$

Аналогично доказательству утверждения 1 получаем, что $\exp(A_1 \Delta) = \exp(A \Delta)$, так как $\{Y_{m_k}\}$ линейно независимы. Теперь из утверждения 3 следует требуемое.

Отметим, что наборы собственных векторов A_1 и C совпадают, и C всегда не вырождена, если $\{Y_{m_k}\}$ — набор линейно независимых векторов, так как

$$\det[\exp(A \Delta)] = \exp[sp(A \Delta)],$$

по теореме Лиувилля.

Из утверждения 3 следует, что для нахождения матрицы A достаточно получить главное значение логарифма, а в случае наличия некоторой дополнительной информации, варьируя элементы V , можно подобрать соответствующую ветвь. Если собственные числа C лежат в правой полуплоскости, то скалярным сжатием можно привести их в единичный круг комплексной плоскости $R \equiv \{z : |z - 1| < 1\}$ и тогда главное значение логарифма может быть получено простым разложением в ряд:

$$A\Delta = \sum_{l=1}^{\infty} \frac{(-1)^{l-1}}{l} (C - E)^l.$$

Следующее утверждение дает принципиальные условия разрешимости (38) для более широкого класса динамических процессов.

Утверждение 5. Если система дифференциальных уравнений

$$\dot{Y} = F(t, Y),$$

«породившая» таблицу наблюдений Y_i , имеет непрерывную по совокупности аргументов функцию F при $t \in [t_0, t_N]$, $\|Y\| \leq H$, то матрица C при надлежащем выборе Δ будет иметь собственные числа в круге R .

Действительно, для произвольного t_i имеем:

$$Y_{i+1} = Y_i + \int_{t_i}^{t_i+\Delta} F(s, Y(s)) ds = Y_i + \Delta F(\xi_i, Y(\xi_i)),$$

где $\xi_i \in (t_i, t_i + \Delta)$. Тогда

$$C = E + \Delta \{F(\xi_{m_k}, Y(\xi_{m_k}))\} \{Y_{m_k}\}^{-1}.$$

Оценим элементы $\{Y_{m_k}\}^{-1}$ при $\|Y_{m_k}\| \leq H$. По определению обратной матрицы:

$$\{Y_{m_k}\}^{-1} = (\det\{Y_{m_k}\})^{-1} \overline{\{Y_{m_k}\}},$$

где $\overline{\{Y_{m_k}\}}$ — матрица алгебраических дополнений. Из линейной независимости Y_{m_k} следует существование ε такого, что

$$\det\{Y_{m_k}\} \geq \varepsilon > 0.$$

Каждый элемент матрицы $\overline{\{Y_{m_k}\}}$ оценивается по модулю сверху числом $(n - 1)!H^{n-1}$, где n — порядок системы.

Тогда элементы матрицы $\{Y_{m_k}\}^{-1}$ допускают оценку $(n-1)!H^{n-1}/\varepsilon$. Так как каждый элемент матрицы

$$\{F(\xi_{m_k}, Y(\xi_{m_k}))\}\{Y_{m_k}\}^{-1}$$

не превосходит по модулю некоторого числа \hat{F} в силу непрерывности на замкнутом множестве вектор-функции $F(t, x)$, то любой элемент матрицы $\{F(\xi_{m_k}, Y(\xi_{m_k}))\}\{Y_{m_k}\}^{-1}$ по модулю меньше

$$\frac{n!\hat{F}H^{n-1}}{\varepsilon}.$$

Таким образом, при

$$\Delta < \frac{\Delta}{n \times n! \hat{F} H^{n-1}}$$

собственные числа S расположены в круге R .

Из последнего утверждения следует, что шаг наблюдений должен быть достаточно мал, но это, в свою очередь, влечет уменьшение ε . Для линейных систем можно сформулировать следующую полезную практически рекомендацию: шаг наблюдений Δ удобно выбирать равным $1/8$ периода самых коротко периодических колебаний наблюдаемого процесса.

Таким образом, показано, что формулы (42) дают решение задачи восстановления при минимально возможной информации. Отметим, что для дифференциальной системы (37) решением являются формулы (42), а для разностной системы (40) решением будут матрица S и вектор D .

2.3.3. АВТОРЕГРЕССИИ ПРИ БОЛЬШИХ ОБЪЕМАХ ДАННЫХ

Большие объемы информации, т. е. большие N , позволяют применить алгоритм предыдущего пункта, но он содержит столь многочисленные предположения и ограничения, что лучше решать приближенную задачу, опираясь на метод наименьших квадратов, и строить авторегрессии не очень высокого порядка. Итак, пусть $n \ll N$. Теперь мы не предполагаем, что ряд векторов $\{X_0, X_1, \dots, X_N\}$ — это наблюдения за решением системы (37) или (40). Пусть динамический процесс описывается нелинейной системой или

линейной, но измерения делаются с ошибками. Постараемся подобрать такую систему (40) (т. е. матрицу C и вектор D), чтобы функционал

$$F(C, D) = \sum_{k=0}^{N-1} (X_{k+1} - CX_k - D)^2, \quad (45)$$

достигал своего наименьшего значения. Если мы в дальнейшем предполагаем пользоваться построенной авторегрессией для прогноза в дискретные моменты времени (с тем же шагом дискретности), то считаем задачу решенной. Если же необходимо прогнозировать $X(t)$ для любых $t > N$, то переходим к системе (37) уже обсуждавшимся способом по формулам (42).

Прежде чем приступить к решению этой задачи, введем определение псевдообратной матрицы [19].

Определение. Матрица P^+ размером $(m \times n)$ называется псевдообратной для $(m \times n)$ матрицы P , если выполняются равенства:

$$PP^+P = P, \quad (46)$$

$$P^+ = RP^T = P^TS, \quad (47)$$

где R и S — некоторые матрицы, P^T — транспонированная к P матрица.

Если P — квадратная и неособенная матрица, то для нее существует обратная матрица P^{-1} . Если P — квадратная и особенная матрица, или не квадратная, а прямоугольная, то матрица P не имеет обратной. Однако в этом случае для произвольной прямоугольной матрицы P существует псевдообратная матрица P^+ , которая обладает некоторыми свойствами обратной матрицы. В случае, когда P — квадратная неособенная матрица, псевдообратная матрица P^+ совпадает с обратной P^{-1} .

Если P — квадратная неособенная матрица, то уравнение (46) имеет единственное решение $P^+ = P^{-1}$. Если же P — произвольная прямоугольная матрица размерности $(m \times n)$, то решение уравнения (46), во-первых, имеет размерность $(m \times n)$, во-вторых, определяется неоднозначно. В общем случае оно имеет бесконечное число решений, но среди всех этих решений имеется только одно, обла-

дающее тем свойством, что его строки и столбцы являются линейными комбинациями соответственно строк и столбцов транспонированной матрицы P^T [19], т. е. выполняется свойство (47).

Таким образом, имеет место следующее утверждение.

Утверждение 6. Для произвольной прямоугольной матрицы P существует и единственна ее псевдообратная матрица.

Для определения псевдообратной матрицы можно дать следующий алгоритм, который называется алгоритмом Гревилля.

Пусть P — прямоугольная матрица размерности $(m \times n)$, в которой число строк больше или равно числу столбцов.

Пусть p_k — k -й столбец в $(m \times n)$ -матрице P , b_k — последняя строка в матрице P_k^+ ($k = 1, \dots, m, P_1 = p_1, P_m = P$). Тогда

$$P_1^+ = p_1^+ \frac{p_1^T}{p_1^T p_1}, \quad (48)$$

и для $k > 1$ имеют место рекуррентные формулы:

$$P_k^+ = \begin{pmatrix} B_k \\ b_k \end{pmatrix}, \quad B_k = P_{k-1}^+ - d_k b_k, \quad d_k = P_{k-1}^+ p_k; \quad (49)$$

если при этом $c_k = p_k - P_{k-1} d_k \neq 0$, то

$$b_k = c_k^+ = (p_k - P_{k-1} d_k)^+; \quad (50)$$

если же $c_k = 0$, т. е. $p_k = P_{k-1} d_k$, то

$$b_k = (1 + d_k^T d_k)^{-1} d_k^T P_{k-1}^+. \quad (51)$$

Матрица

$$\begin{pmatrix} B_k \\ b_k \end{pmatrix}$$

является псевдообратной для матрицы P_k , если матрица B_k и строка b_k определяются формулами (48)–(51). Этот метод не требует вычисления детерминантов и может быть использован для вычисления обратной матрицы.

Вернемся к задаче определения минимума функционала (45). Введем в рассмотрение матрицу $H = \{C, D\}$, тогда систему (40) можно переписать в виде:

$$\{X_1, X_2, \dots, X_N\} = H \begin{Bmatrix} X_0 & X_1 & \dots & X_{N-1} \\ 1 & 1 & \dots & 1 \end{Bmatrix}. \quad (40')$$

Определим норму $\|H\|$ ($n \times (n + 1)$)-матрицы $H = \{h_{ij}\}$, ($i = 1, \dots, n, j = 1, \dots, n + 1$) как неотрицательное число, задаваемое формулой

$$\|H\|^2 = \sum_{i,j} |h_{i,j}|^2 \equiv Sp(H^T H).$$

Тогда соответствующая норма вектора h размерности $(n + 1)$ будет задаваться формулой

$$\|h\|^2 = h^T h = \sum_{i=1}^{n+1} |h_i|^2.$$

При этом очевидно, что

$$\|H\|^2 = \sum_{j=1}^{n+1} |H_{.j}|^2 = \sum_{i=1}^n |H_{i.}|^2,$$

где $H_{.j}$ ($j = 1, \dots, n + 1$) есть j -й столбец матрицы H . Соответственно, $H_{i.}$ ($i = 1, \dots, n$) — i -я строка матрицы H .

Определение. ($n \times (n + 1)$)-матрицу $H^0 = \{C^0, D^0\}$ назовем наилучшим приближенным в смысле квадратичного отклонения решением системы (40), если при значениях $C = C^0, D = D^0$ «квадратичное отклонение»

$$F = \sum_{k=0}^{N-1} (X_{k+1} - CX_k - D)^2,$$

достигает своего наименьшего значения и среди всех матриц C и столбцов D , для которых это отклонение имеет минимальное значение, матрица $H^0 = \{C^0, D^0\}$ имеет наименьшую норму.

Тогда имеет место следующее утверждение.

Утверждение 7. Линейная алгебраическая система (40') относительно H имеет одно и только одно наилучшее в смысле квадратичного отклонения приближенное решение, которое определяется по формуле

$$H = \{X_1, X_2, \dots, X_N\} \begin{Bmatrix} X_0 & X_1 & \dots & X_{N-1} \\ 1 & 1 & \dots & 1 \end{Bmatrix}^+,$$

где

$$\begin{Bmatrix} X_0 & X_1 & \dots & X_{N-1} \\ 1 & 1 & \dots & 1 \end{Bmatrix}^+$$

— псевдообратная матрица к матрице

$$\begin{Bmatrix} X_0 & X_1 & \dots & X_{N-1} \\ 1 & 1 & \dots & 1 \end{Bmatrix}.$$

Действительно, введем следующие обозначения:

$$G = \{X_1, X_2, \dots, X_N\}, \quad P = \begin{Bmatrix} X_0 & X_1 & \dots & X_{N-1} \\ 1 & 1 & \dots & 1 \end{Bmatrix}.$$

Очевидно, имеем:

$$\begin{aligned} F &= \sum_{k=0}^{N-1} (X_{k+1} - CX_k - D)^2 = \\ &= \left\| \{X_1, X_2, \dots, X_N\} - H \begin{Bmatrix} X_0 & X_1 & \dots & X_{N-1} \\ 1 & 1 & \dots & 1 \end{Bmatrix} \right\|^2 = \\ &= \sum_{i=1}^n (G_i - H_i P)^2. \end{aligned}$$

Отсюда следует, что i -я строка искомой матрицы H должна быть наилучшим приближенным в смысле квадратичного отклонения решением системы линейных уравнений:

$$H_i P = G_i, \quad (52)$$

где наилучшим приближенным в смысле квадратичного отклонения решением понимается строка H_i , минимальная по норме

$$|H_i| = H_i H_i^T = \sum_{i=1}^{n+1} |h_{il}|^2,$$

и доставляющая минимум функционалу

$$(G_i - H_i P)^2 = \sum_{l=1}^N \left| g_{il} - \sum_{k=1}^{n+1} h_{ik} p_{kl} \right|^2. \quad (53)$$

Покажем, что система (52) имеет единственное наилучшее в смысле квадратичного отклонения приближенное решение, которое определяется по формуле

$$H_i^0 = G_i P^+,$$

где P^+ — псевдообратная для P матрица.

Положим $S_1 = G_i - H_i^0 P = G_i - G_i P^+ P$, $S_2 = (H_i^0 - H_i) P$.

Рассмотрим произвольную строку H_i , для которой

$$\begin{aligned} G_i - H_i P &= G_i - H_i P - H_i^0 P + H_i^0 P = \\ &= G_i - H_i^0 P + (H_i^0 - H_i) P = S_1 + S_2. \end{aligned}$$

Тогда минимизируемый функционал (53) запишется:

$$\begin{aligned} (G_i - H_i P)^2 &= (G_i - H_i P)(G_i - H_i P)^T = \\ &= (S_1 + S_2)(S_1 + S_2)^T = S_1^2 + S_2^2 + S_1 S_2^T + S_2 S_1^T, \end{aligned}$$

но

$$\begin{aligned} S_1 S_2^T &= G_i (E - P^+ P) P^T (H_i^0 - H_i)^T = \\ &= G_i (P^T - P^+ P P^T) (H_i^0 - H_i)^T. \end{aligned} \quad (54)$$

Произвольную прямоугольную $((n + 1) \times N)$ -матрицу P ранга k можно разложить в произведение двух матриц Q и R , имеющих соответственно размеры $((n + 1) \times k)$ и $(k \times N)$:

$$P = QR = \begin{pmatrix} q_{11} & \dots & q_{1k} \\ q_{21} & \dots & q_{2k} \\ \dots & \dots & \dots \\ q_{n+11} & \dots & q_{n+1k} \end{pmatrix} \begin{pmatrix} r_{11} & \dots & r_{1N} \\ r_{21} & \dots & r_{2N} \\ \dots & \dots & \dots \\ r_{k1} & \dots & r_{kN} \end{pmatrix}. \quad (55)$$

Это разложение можно получить, если в качестве столбцов матрицы Q взять любые k линейно независимых столбцов матрицы P . Тогда j -й столбец матрицы R будет составлен из коэффициентов $r_{1j}, r_{2j}, \dots, r_{kj}$, на которые нужно домножить столбцы матрицы Q , чтобы получить j -й столбец матрицы P .

Разложение вида (55) называется скелетным разложением.

Найдем выражения для псевдообратных матриц к матрицам Q и R . По определению псевдообратной матрицы

$$QQ^+Q = Q, \quad Q^+ = \bar{T}Q^T,$$

где \bar{T} — некоторая матрица. Тогда $Q\bar{T}Q^TQ = Q$. Домножим слева на Q^T , тогда $\bar{T} = (Q^TQ)^{-1}$, так как Q^TQ — неособенная квадратная матрица. Следовательно, $Q^+ = (Q^TQ)^{-1}Q^T$. Аналогично, $R^+ = R^T(RR^T)^{-1}$.

Теперь покажем, что

$$P^+ = R^+Q^+ = R^T(RR^T)^{-1}(Q^TQ)^{-1}Q^T \quad (56)$$

является псевдообратной к матрице P , т. е. имеют место соотношения (46), (47). Действительно,

$$\begin{aligned} PP^+P &= QRR^T(RR^T)^{-1}(Q^TQ)^{-1}Q^TQR = QR = P, \\ P^+ &= R^TWQ^T = R^TW(RR^T)^{-1}RR^TQ^T = \bar{T}R^TQ^T = \bar{T}P^T, \\ P^+ &= R^TWQ^T = R^TQ^TQ(Q^TQ)^{-1}WQ^T = R^TQ^TS, \end{aligned}$$

где

$$\begin{aligned} W &= (RR^T)^{-1}(Q^TQ)^{-1}, \\ \bar{T} &= R^TW(RR^T)^{-1}R, \quad S = Q(Q^TQ)^{-1}WQ^T. \end{aligned}$$

Таким образом, для псевдообратной к P матрицы P^+ имеет место равенство (56).

Вернемся к выражению (54) и используем свойство (56), тогда

$$P^+PP^T = R^T(RR^T)^{-1}(Q^TQ)^{-1}Q^TQRR^TQ^T = R^TQ^T = P^T.$$

Следовательно, $S_1S_2^T = 0$, но тогда и $S_2S_1^T = (S_1S_2^T)^T = 0$. Поэтому

$$(G_i - H_iP)^2 = |S_1|^2 + |S_2|^2 = |G_i - H_i^0P|^2 + |(H_i^0 - H_i)P|^2, \quad (57)$$

откуда следует, что для любой строки H_i

$$|G_i - H_iP| \geq |G_i - H_i^0P|.$$

Предположим, что

$$|G_i - H_iP| = |G_i - H_i^0P|,$$

тогда из (57) следует, что $KP = 0$, где $K = H_i - H_i^0$. С другой стороны,

$$|H_i|^2 = (H_i^0 + K)(H_i^0 + K)^T = |H_i^0|^2 + |K|^2 + KH_i^0{}^T + H_i^0K^T, \quad (58)$$

но

$$KH_i^0{}^T = K(G_iP^+)^T = KP^+{}^TG_i^T = KP^T\bar{T}G_i^T = 0,$$

так как $KP = 0$. Следовательно,

$$H_i^0 K^T = (KH_i^{0T})^T = 0.$$

Тогда (58) примет вид:

$$|H_i|^2 = |H_i^0|^2 + |K|^2,$$

т. е. $|H_i|^2 \geq |H_i^0|^2$. Равенство имеет место только при $K = 0$, т. е. при $H_i = H_i^0$, где $H_i^0 = G_i P^+$.

Таким образом, мы показали, что i -я строка искомой матрицы H является наилучшим приближенным решением системы линейных уравнений

$$H_i P = G_i,$$

и это наилучшее решение имеет вид:

$$H_i = G_i P^+.$$

Так как это равенство справедливо при любом $i = 1, \dots, n$, то

$$H^0 = GP^+.$$

Это и требовалось установить.

2.3.4. ИДЕНТИФИКАЦИЯ АВТОРЕГРЕССИОННОЙ МОДЕЛИ С ЗАПАЗДЫВАНИЕМ

Проблема идентификации величины временного лага в различных постановках в эконометрической литературе рассматривается уже около 50 лет. По-видимому, первые содержательные определения и результаты принадлежат Акаике. Им предложен метод выделения главных, в некотором смысле, переменных регрессионной зависимости. Рассмотрены случаи линейной и нелинейной регрессии, а также как частный случай — авторегрессионная зависимость. В качестве критерия оптимальности выбора использовались AIC — Акаике информационный критерий, FPE — окончательная ошибка предсказания и BIC — Байесов информационный критерий. Последние десятилетия эти критерии и методы их использования активно анализировались статистически, при разных предположениях

доказывались теоремы о единственности экстремума, о состоятельности оценок и другие (см. например, [108, 109, 115, 120]).

Использование моделей с временным лагом (или, как принято говорить в математической литературе, с запаздывающим аргументом) в экономике вполне естественно, так как время реакции системы всегда значительно. Например, от инвестиций до выхода новой продукции иногда проходят годы, в особенности когда инвестиции делаются в науку, образование или медицину. Промышленность и сельское хозяйство имеют запаздывания, определяемые технологическими циклами; в торговле, и вообще в распределительной сфере, оборот капитала происходит быстрее и, следовательно, запаздывание реакции меньше. Спекулятивные операции могут носить мгновенный характер, т. е. исключают временной лаг. Таким образом, в макроэкономических исследованиях естественно учитывать лаги, но возникает проблема их определения. Обсудим этот вопрос более подробно.

Поскольку практически во всякой экономической деятельности присутствует запаздывание, то наиболее правильно и точно было бы вычислить его величину, основываясь на технологических картах рассматриваемой деятельности, включая в анализ все ее этапы. Например, для сельскохозяйственного предприятия с земледельческой ориентацией время реакции разумно связать с годовым циклом получения урожая, хотя предприятие может рассматривать долгосрочные программы мелиорации или крупного строительства, и тогда реакция производства на инвестиции очевидно затянется. Если же сельскохозяйственное предприятие производит продукцию из мяса крупного рогатого скота, то запаздывание между инвестициями и их влиянием на выход скорее всего будет два года, потому что сдача двухгодичных бычков более выгодна. Свиноводство определяется реакцией в один год, поэтому и модель должна быть с временным лагом в один год. В строительстве жилых домов временной лаг может быть год или полтора года, а в промышленном строительстве пуск объекта может отстоять от начала на пять и более лет.

Все эти примеры показывают, как трудно вычислить запаздывание в конкретных производствах. Еще труднее это делать в распределительной и потребительной сферах, поскольку кроме технологических необходимо учитывать социальные и психологические факторы.

Однако можно попытаться дать интегральную оценку времени реакции экономической системы на единицу инвестиций.

Критерий. Мы считаем, что такой оценкой может служить срок, на который дается кредит коммерческими банками для расширения или модернизации данного вида экономической деятельности, точнее, не срок полного возврата кредита, а срок погашения первой его части.

Усредняя названную величину по различным коммерческим банкам и по близким экономическим субъектам, получим оценку временного лага для динамической модели, но это, к сожалению, умозрительное определение является не вполне конструктивным: вряд ли банки раскроют свои базы данных для такого исследования, поскольку задача крупных кредитов — дело всегда тонкое и конфиденциальное. Поэтому со всей остротой встает проблема идентификации запаздывания по наблюдениям за динамикой экономической деятельности.

Рассмотрим сначала простейшее уравнение с запаздыванием

$$\dot{x} = -x(t - h),$$

где $h > 0$ — запаздывание. Пусть наблюдается его решение $x(t)$ при $t = k\delta$, $\delta > 0$ — шаг дискретности наблюдений, $k = 0, 1, 2, \dots, N$. Предположим также, что $h \gg \delta$. Выпишем уравнение в интегральной форме:

$$x(t) = x(h) - \int_0^{t-h} x(s) ds, \quad t \in [h, N\delta]. \quad (59)$$

Заменяя приближенно интеграл суммой, а запаздывание h целым числом l шагов дискретности δ , получим из (59):

$$x(k\delta) = x(l\delta) - \delta \sum_{i=0}^{k-l-1} x(i\delta). \quad (60)$$

Введем теперь в рассмотрение функционал квадратичных отклонений

$$F(l) = \sum_{k=l+1}^N \left(x_k - x_l + \delta \sum_{i=0}^{k-l-1} x_i \right)^2 \frac{1}{N-l}. \quad (61)$$

Очевидно, если $x_k = x(k\delta)$ и $h \approx l\delta$, то $F(l) = 0$. В остальных случаях $F(l) > 0$. Таким образом, для нахождения h надо вычислить l , так как δ предполагается известной. По известным наблюдениям $\{x_k\}$ требуемое значение l вычисляется как доставляющее наименьшее значение функционалу (61). К сожалению, алгоритм получился нелинейный, т. е. возможно, что будет много локальных экстремумов, но так как $l = 1, \dots, N$, то требуемое значение можно найти простым перебором всех возможных l .

Обобщим теперь полученный алгоритм на произвольную линейную систему дифференциальных уравнений с запаздыванием, т. е. предположим, что наблюдается временной векторный ряд $\{X_0, X_1, \dots, X_N\}$ и мы хотели бы по нему построить линейную систему авторегрессии с запаздыванием:

$$\dot{X}(t) = AX(t) + BX(t-h). \quad (62)$$

Здесь $n \times n$ -матрицы A, B и запаздывание h неизвестны.

Переходя к формуле Коши для системы (62), получим многомерный аналог формулы (60):

$$X(k\delta) = e^{A(k-l)\delta} \left[X(l\delta) + \int_{l\delta}^{k\delta} e^{A(l\delta-t)} BX(t-l\delta) dt \right],$$

или в иных обозначениях для соседних точек решения системы (62):

$$X_{k+1} = e^{A\delta} X_k + \int_{(k-l)\delta}^{(k+1-l)\delta} \exp[A(k+1-l)\delta - As] BX(s) ds.$$

Приближенное значение интеграла в последнем выражении выберем из условия линейности по матричной экспоненте. Тогда в соответствии с этим определим аналог функционала (61):

$$F_l(D, B) = \sum_{k=l}^{N-1} (X_{k+1} - DX_k - B\delta X_{k+1-l})^2 \frac{1}{N-l}. \quad (63)$$

Заметим, что для каждого l функционал (63) квадратичный по $D = \exp(\Delta\delta)$, B , и можно использовать следующий понятный алгоритм вычисления наилучших матриц, доставляющих наименьшее значение (63).

Итак, пусть l фиксированно и N , δ , X_k известны. Преобразуем функционал (63) так, чтобы вместо двух матриц D , B надо было бы искать одну прямоугольную матрицу $G = \{D, B\}$:

$$F_l(G) = \sum_{k=l}^{N-1} \left(X_{k+1} - G \begin{Bmatrix} X_k \\ \delta X_{k+1-l} \end{Bmatrix} \right)^2 \frac{1}{N-l}.$$

В этих обозначениях производная функционала F по элементам матрицы G будет прямоугольной матрицей вида:

$$\frac{\partial F_l}{\partial G} = \sum_{k=l}^{N-1} (-2) \left(X_{k+1} \begin{Bmatrix} X_k \\ \delta X_{k+1-l} \end{Bmatrix}^T \right) + 2 \left(G \begin{Bmatrix} X_k \\ \delta X_{k+1-l} \end{Bmatrix} \begin{Bmatrix} X_k \\ \delta X_{k+1-l} \end{Bmatrix}^T \right).$$

Если приравнять ее к нулю, то можно выразить матрицу G :

$$G = \left(\sum_{k=l}^{N-1} X_{k+1} \begin{Bmatrix} X_k \\ \delta X_{k+1-l} \end{Bmatrix}^T \right) \left(\sum_{k=l}^{N-1} \begin{Bmatrix} X_k \\ \delta X_{k+1-l} \end{Bmatrix} \begin{Bmatrix} X_k \\ \delta X_{k+1-l} \end{Bmatrix}^T \right)^{-1},$$

где вторая матрица, естественно, должна быть невырождена. Это будет иметь место, если среди $N-l$ векторов $\{X_k, \delta X_{k-l+1}\}^T$ есть $2n$ линейно независимых. Последнее утверждение требует некоторых пояснений. В более простых обозначениях оно звучит так: пусть среди j -векторов $\{V_1, \dots, V_N\}$ имеется ровно j линейно независимых, тогда ранг суммы матриц первого ранга

$$\sum_{k=1}^N V_k V_k^T$$

будет равен j . Отметим, что каждое слагаемое последней суммы аффинно эквивалентно $j \times j$ -матрице, состоящей из всех нулей и одной единицы где-нибудь на главной диагонали. Тогда линейное преобразование $S: V \rightarrow \hat{V}$, где матрица преобразования составлена из упомянутых выше линейно независимых j -векторов $\{V_1, \dots, V_N\}$, переводит их в единичные векторы. Значит, если бы $N = j$, т. е. вся

система векторов была бы линейно независима, то указанная сумма была бы равна SS^T и таким образом была бы невырожденная. В общем случае, применяя ортогонализацию Грама–Шмидта, выражаем исходную сумму через j аналогичных слагаемых, но использующих ортонормированные векторы, и затем составляем преобразование S .

Зная матрицу G , вычислим матричный логарифм от ее части D . Это можно сделать, пользуясь формулами раздела 2.3.2. этой главы. Таким образом, определены матрицы коэффициентов системы (62) при фиксированном l . Обозначим их A_l, B_l . Перебирая все возможные значения $l = 1, \dots, N - 1$, рассмотрим изменение $F_l(D_l, B_l)$ и остановимся на том значении, для которого функционал (63) имеет наименьшую величину.

2.3.5.

ОБСУЖДЕНИЕ КРИТЕРИЯ ОПТИМАЛЬНОСТИ

Обратим внимание, что функционал (45) вынуждает нас выбрать такие коэффициенты системы (40), чтобы из любого положения X_k динамическая система в следующий момент времени ($k + 1$) как можно ближе переместилась к X_{k+1} . Это задает тенденцию, фиксирует причинно-следственную базу модели, но не исчерпывает многообразия решений линейных систем. Так, общее решение разностной системы (40) имеет вид:

$$X_{k+1} = C^{k+1} X_0 + \sum_{i=0}^k C^i D, \quad (64)$$

и следовало бы вместо функционала (45) минимизировать другое квадратичное отклонение:

$$\tilde{F}(C, D) = \sum_{k=0}^{N-1} \left(X_{k+1} - C^{k+1} X_0 - \sum_{i=0}^k C^i D \right)^2. \quad (65)$$

Однако здесь проблема в нелинейной зависимости функционала от искомых параметров (и от X_0 дополнительно, демонстрируя особый характер начального вектора). В настоящее время существует большое количество компьютерных программ для поиска минимумов, но нелинейные задачи на экстремум сопряжены с трудностями опре-

деления глобального решения (в отличие от локальных), к тому же объем вычислений для длинных временных рядов будет слишком велик. Отметим лишь пока, что решение задачи с функционалом (65) отличается от оптимальных C, D , найденных по функционалу (45).

Привлекательность последнего функционала, более простого в вычислении, а также базовое предположение о линейности поведения в каждый момент времени все равно не избавляют нас от выбора начального вектора для прогноза. Предположим, что модель (40) каким-либо способом построена и нам необходимо ею воспользоваться, чтобы определить будущее поведение динамической системы, т. е. при $t > N$. Можно, кстати, для этой цели привлечь формулу (64)

$$X_{k+1+N} = C^{k+1} X_N + \sum_{i=0}^k C^i D, \quad k = 0, 1, \dots, L,$$

где L — длительность прогноза. Вот только правильно ли в последней формуле в качестве начальной точки прогноза рассматривать конечную точку предыстории, ведь мы знаем (предполагаем), что всякий вектор X_k является реализацией случайного вектора и систематической составляющей, поэтому необходимо обосновать выбор начального вектора. Для этого вернемся в начало временного ряда $\{X(0), \dots, X(N)\}$ и рассмотрим X_0 . Почему в формуле (64) стоит первый вектор ряда, он ведь так же, как другие, содержит ошибки в компонентах? Очевидно, этим вектором в (64) можно распорядиться для улучшения приближения. Используем полученную свободу для минимизации функционала (65), считая, что C, D найдены по функционалу (45). Приведем соответствующие формулы.

При фиксированных C и D будем иметь:

$$\tilde{F}(X) = \sum_{k=0}^{N-1} \left(\left(X_{k+1} - \sum_{i=0}^k C^i D \right) - C^{k+1} X \right)^2.$$

Далее берем производную по X и приравняем ее к нулю:

$$\frac{D}{DX} \tilde{F}(X) = -2 \sum_{k=0}^{N-1} (C^T)^{k+1} \left(\left(X_{k+1} - \sum_{i=0}^k C^i D \right) - C^{k+1} X \right) = 0.$$

Уравнение для оптимального X принимает вид:

$$\sum_{k=0}^{N-1} (C^T)^{k+1} \left(X_{k+1} - \sum_{i=0}^k C^i D \right) = \sum_{k=0}^{N-1} (C^T)^{k+1} (C^{k+1}) X. \quad (66)$$

Использование такой схемы (вычисления модели по линейно-квадратичному вычислительному пути и оптимизация начального вектора так же простым и практически однозначным путем) может быть полезно в двух направлениях. Если задача настолько громоздка (большие n и N), что невозможно искать и анализировать минимумы функционала (65), то предлагается ограничиться вычислениями оптимальных точек функционала (45) и решением системы (66). Если же вычислительные возможности исследователя достаточно велики, то надлежащий экстремум функционала (65) гораздо быстрее получить, пользуясь приближенной оптимизацией (45) и (66).

Рассмотрим на простом примере описанную ситуацию. Пусть скалярный временной ряд состоит из чисел $\{1, 2, 4, 5\}$. Требуется построить линейную однородную модель $x_{k+1} = cx_k$ так, чтобы она наилучшим образом, в смысле квадратичного отклонения, описывала предлагаемый ряд, т. е. величина функционала

$$\tilde{F}(c, x) = (2 - cx)^2 + (4 - c^2x)^2 + (5 - c^3x)^2,$$

достигала наименьшего значения для любых c, x . Вычислим соответствующие частные производные и приравняем их к нулю:

$$\begin{aligned} (2 - cx) + 2c(4 - c^2x) + 3c^2(5 - c^3x) &= 0, \\ (2 - cx) + c(4 - c^2x) + c^2(5 - c^3x) &= 0. \end{aligned}$$

Из второго уравнения выразим cx :

$$cx = \frac{2 + 4c + 5c^2}{1 + c^2 + c^4}.$$

Теперь cx подставим в первое уравнение и получим соотношение для c :

$$4c^4 - c^3 - 8c - 4 = 0.$$

Его корни соответствуют точкам, подозрительным на экстремум. С помощью компьютера определим, что такими точками являются $c_1 = -0,465$; $c_2 = 1,477$. Из форму-

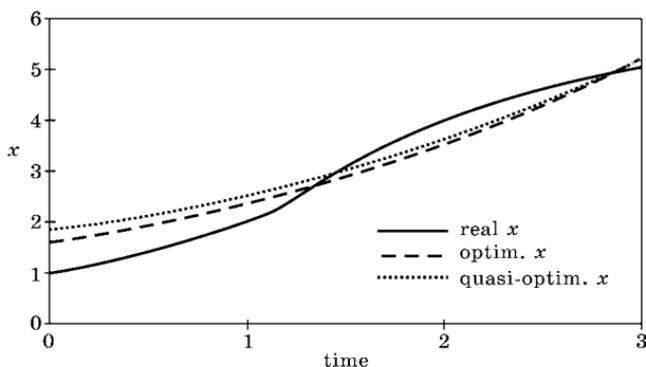


Рис. 5

Графики сравнения оптимальных и почти оптимальных кривых с действительными данными

лы для cx получим $x_1 = -2,08$; $x_2 = 1,604$. Значения функционала будут следующие: $\tilde{F}(c_1, x_1) = 53,3$; $\tilde{F}(c_2, x_2) = 0,78$. Очевидно, что первая точка нас не устраивает, но необходимо дополнительное исследование, чтобы определить ее характер.

Вычислим теперь приближенную оптимальную модель по функционалу (45) и начальному вектору из (66). Наименьшее значение функции (функционала квадратичных отклонений)

$$F(c) = \sum_{k=0}^2 (x_{k+1} - cx_k)^2,$$

достигается при $c^* = 10/7 = 1,428$. Наилучшее начальное значение для решения в форме (64) найдем по формуле (66): $x^* = 1,741$. Окончательно, $\tilde{F}(c^*, x^*) = 0,991$, что конечно чуть хуже, чем в оптимальной точке. На рисунке 5 приведены графики реального, оптимального и примерно оптимального динамического процессов.

2.3.6.

МОДЕЛЬ РАВНОВЕСНОЙ ЦЕНЫ ПО МЕНДЕЛЕЕВУ

В данном разделе рассматривается формирование цены через воздействие рыночного механизма спроса и предложения. На основе данных — наблюдений за ценой, спросом, предложением на определенном промежутке времени — осуществляется процесс идентификации, построение

локальной модели изменения предложения, спроса, цены. Посредством этой модели возможно прогнозирование предложения, спроса и цены на некоторый промежуток времени, следующий за тем временем, когда производились наблюдения.

Д. И. Менделеев 100 лет назад считал, что знания, полученные в естественных науках, обязательно должны найти адекватное приложение в экономике. В статье «Приемы естествознания в изучении цен» [59] им была предложена интерпретация, согласно которой изменение цен на рынке подобно изменению удельных объемов веществ.

Удельным объемом веществ выражается пространство, занимаемое единицей массы или веса вещества. Ценой выражается плата за единицу веса товара или вещества. Понятие о массах, входящее в оба определения, связывает их между собой и заставляет вспомнить, что это понятие массы проникает во все важнейшие отрасли естествознания. Участвуя в определении цен, оно дает им некоторое естественное право быть предметом точного изучения. Очевидно, что применительно к цене не следует касаться того, что хотя и продается, но ценится не по массе. Таковы, например, места в театре, картины, дома, земля, акции и векселя, услуги.

Можно продолжить данную интерпретацию в направлении конструктивного построения математической модели цены заданного продукта. Используем метод локального моделирования, основанный на построении прогноза по наблюдениям в предшествующие моменты времени.

В естественных науках хорошо известен и действует закон Менделеева–Клапейрона:

$$P\bar{V} = \frac{m}{M}\bar{R}T, \quad (67)$$

где \bar{V} — объем вещества, m — масса, M — молярная масса, T — температура вещества, P — давление, \bar{R} — универсальная постоянная.

Если ввести $V = \frac{\bar{V}}{m}, R = \frac{\bar{R}}{M},$

то V будет играть роль удельного объема, а R будет постоянной для каждого вещества и будет служить характеристикой данного вещества.

Тогда уравнение Менделеева–Клапейрона (67) можно записать следующим образом:

$$PV = RT. \quad (68)$$

Выделим три переменные состояния системы применительно к изучению поведения цены данного продукта на рынке товаров, как это следует из рассуждений Д. И. Менделеева: $P(t)$ — предложение, $T(t)$ — спрос, $V(t)$ — цена. Будем считать, что для любого момента времени верно соотношение вида (68), но применительно к экономической ситуации. Тогда константа R будет характеризовать заданный товар на рынке и будет подлежать определению для каждого товара. Уравнение (68) будет определять равновесные состояния системы.

Аналогичное соотношению (68) выражение для цены продукта можно найти в работе [63]. Там для неравновесной цены выводится соотношение

$$V = \frac{T}{(P - \tilde{y} / \tau_y)},$$

где \tilde{y} — запасы товара вида y в натуральном выражении, τ_y — долговечность товара вида y , связанная одновременно с физическим износом и моральным старением.

Часто для определения цены пользуются выражением [42]

$$dV/dt = T(V) - P(V).$$

В работе [18] для цены дано соотношение

$$dV(t)/dt = HV(t)(T(t) - P(t)),$$

где H — положительная константа.

В термодинамике при изотермическом, изобарном, изохорном процессах из уравнения (68) вытекают законы Бойля–Мариотта, Гей-Люссака, Шарля [93, 49]. В экономике из соотношения (68) вытекают законы спроса и предложения. Причем уравнение (68) не определяет кривых спроса и предложения, а дает характер зависимости равновесной цены от изменения в спросе и предложении.

Из выделенных применительно к экономике переменных наибольшие нестационарности наблюдаются у спроса $T(t)$

и предложения $P(t)$. Это и понятно, так как на спрос влияют неценовые детерминанты (вкусы или предпочтения потребителей, цены на сопряженные товары, потребительские ожидания относительно будущих цен и доходов, влияние рекламы и др.). На предложение сильное влияние оказывают технология производства, налоги и дотации, цены на ресурсы, ожидания изменения цен, цены на другие товары, число продавцов на рынке, наличие свободных капиталов, свободной квалифицированной рабочей силы, возможность быстро переориентировать производство, что требует достаточной децентрализации управления и разгосударствления. Очевидно, все эти факторы являются внешними по отношению к рассматриваемой системе. Цена же при либерализации в первом приближении может считаться функцией спроса и предложения. Все перечисленные выше факторы влияют на динамику изменения спроса и предложения, а значит, отражаются на числовых данных — наблюдениях за ними. Эти данные в предлагаемой модели используются для построения уравнений измерения спроса и предложения, а через уравнение (68) указанные данные будут оказывать влияние и на изменение цены. Таким образом, считаем, что

$$V = RT/P. \quad (69)$$

Для получения предложения $P(t)$ и спроса $T(t)$ и возможности их прогнозирования можно воспользоваться локальным моделированием. Будем искать аппроксимацию $P(t)$ и $T(t)$ в виде:

$$\begin{cases} \dot{P} = a_{11}P + a_{12}T + b_1 \\ \dot{T} = a_{21}P + a_{22}T + b_2 \end{cases}. \quad (70)$$

Коэффициенты данной системы $\{a_{ij}\}$ и $\{b_{ij}\}$ определяются по векторам наблюдений в дискретные моменты времени, предшествующие текущему:

$$\begin{aligned} X_0 \begin{pmatrix} P(t_0) \\ T(t_0) \end{pmatrix}, \quad X_1 &= \begin{pmatrix} P(t_0 + h) \\ T(t_0 + h) \end{pmatrix}, \\ X_2 &= \begin{pmatrix} P(t_0 + 2h) \\ T(t_0 + 2h) \end{pmatrix}, \quad X_3 = \begin{pmatrix} P(t_0 + 3h) \\ T(t_0 + 3h) \end{pmatrix}, \end{aligned}$$

где h — шаг, с которым проводятся наблюдения.

Введем векторы $Y_k = X_k - X_{k-1}$, $k = 1, 2, 3$. Составим из них матрицы $D = \{Y_1, Y_2\}$, $G = \{Y_2, Y_3\}$. Тогда существует шаг h , при котором мы можем восстановить матрицу $A = \{a_{ij}\}$ и вектор $b = \{b_i\}$. При этом выражения для них будут иметь вид

$$A = \frac{1}{h} Ln[\{Y_2, Y_3\}\{Y_1, Y_2\}^{-1}], \quad (71)$$

$$b = (e^{Ah} - E)^{-1}A(X_3 - e^{Ah}X_2). \quad (72)$$

Окончательно решение уравнения (70) будет иметь вид:

$$X(t) = \begin{pmatrix} P(t) \\ T(t) \end{pmatrix} = e^{A(t-t_0)} \left[\begin{pmatrix} P(t_0) \\ T(t_0) \end{pmatrix} + \int_{t_0}^t e^{-A(t-\tau)} b d\tau \right]. \quad (73)$$

Формулу (73) можно упростить, если воспользоваться соотношением

$$e^{A(t-t_0)} - E = A \int_{t_0}^t e^{A(t-\tau)} d\tau.$$

Тогда для получения $P(t)$ и $T(t)$ и возможности их прогнозирования для $t > t_0 + 3h$ можно записать:

$$X(t) = \begin{pmatrix} P(t) \\ T(t) \end{pmatrix} = e^{A(t-t_0)} X_0 + A^{-1}[e^{A(t-t_0)} - E] b, \quad (74)$$

где A и b вычисляются по формулам (71), (72).

Неизвестную величину R можно определить как аргумент среднеквадратичной минимизации функционала:

$$I(R) = \sum_{k=0}^3 \left[V(t_0 + kh) - \frac{RT(t_0 + kh)}{P(t_0 + kh)} \right]^2.$$

В результате выражение для R будет иметь вид:

$$R = \frac{\sum_{k=0}^3 V(t_0 + kh) \frac{T(t_0 + kh)}{P(t_0 + kh)}}{\sum_{k=0}^3 \frac{T^2(t_0 + kh)}{P^2(t_0 + kh)}}.$$

Таким образом, располагая четырьмя наблюдениями за ценой $V(t)$, спросом $T(t)$ и предложением $P(t)$ определенного товара на промежутке времени $t_0 \leq t \leq t_0 + 3h$, мы

Таблица 1

Данные биржи для листа 12X18H10T

9.05 15.05	16.05 22.05	23.05 29.05	30.05 5.06	6.06 12.06	13.06 19.06	20.06 26.06	27.06 3.07
Предложение, кг							
11 500	14 800	10 600	11 000	14 200	12 200	11 500	12 500
Спрос, кг							
6400	8640	6140	6175	8200	7000	6100	7400
Цена, руб.							
2400	2500	2500	2400	2500	2500	2400	2500

Таблица 2

Прогноз для листа 12X18H10T

9.05 15.05	16.05 22.05	23.05 29.05	30.05 5.06	6.06 12.06	13.06 19.06	20.06 26.06	27.06 3.07
Предложение, кг							
11 500	15 018	10 738	11 053	13 982	12 009	12 300	13 162
Спрос, кг							
6400	8735	6200	6198	8105	6917	6999	7605
Цена, руб.							
2400	2508	2505	2402	2492	2493	2445	2490

можем восстановить матрицу A , вектор b , константу R , а затем использовать их для определения прогноза для $P(t)$, $T(t)$ согласно (74), $V(t)$ согласно (69) при $t > t_0 + 3h$.

По описанному алгоритму составлена программа на ЭВМ. В качестве наблюдений использовались данные товарно-фондовой биржи «Санкт-Петербург». Эти данные отражают торги на бирже по секции «Черная металлургия» таких товаров, как лист 12X18H10T, за май–июнь 1994 г. Данные, взятые с биржи по этим товарам, приведены в таблице 1. Шаг наблюдений составляет одну неделю. Первые четыре столбца из таблицы 1 используются как основа для построения матрицы A , столбца b , константы R .

Результаты подсчетов приведены в таблице 2, причем первые четыре столбца в этих таблицах соответствуют тому же промежутку времени, за который брались исход-

ные наблюдения с биржи, используемые для построения модели, а следующие четыре столбца отражают месячный прогноз изменения спроса, предложения, цены указанных товаров. Этот прогноз можно сравнить с действительными показателями биржи за тот же промежуток

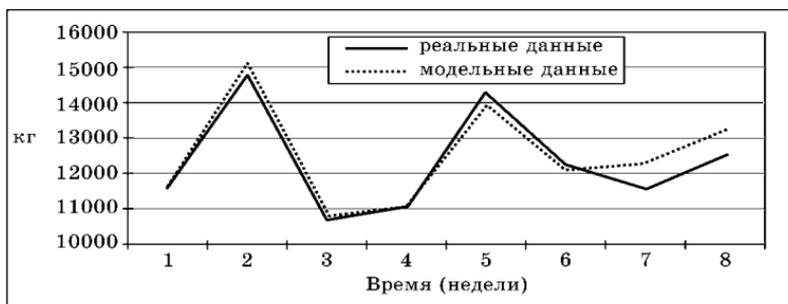


Рис. 6
 Диаграмма предложения на лист 12X18H10T:
 реальные и модельные данные

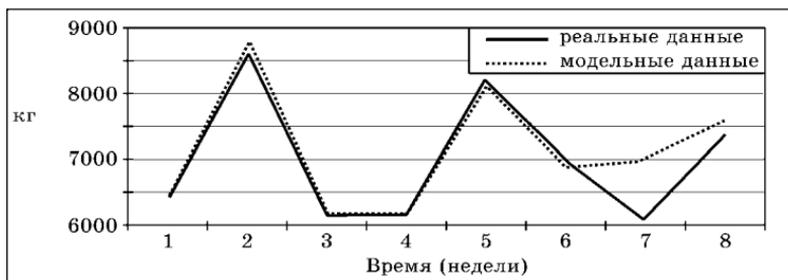


Рис. 7
 Диаграмма спроса на лист 12X18H10T: реальные и модельные данные

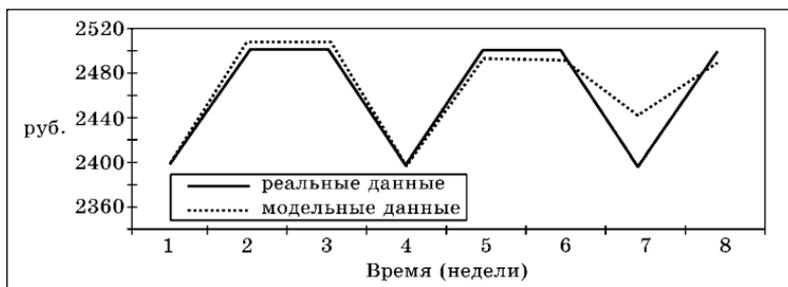


Рис. 8
 Диаграмма цены листа 12X18H10T: реальные и модельные данные

времени. Данные показатели содержатся в последних четырех столбцах таблицы 1.

Отличие расчетных и действительных цифр можно проследить по диаграммам (ниже приведены соответствующие графики, см. рис. 6–8). Сплошные линии отражают действительные показатели, а пунктирные — расчетные.

Представленная модель формирования цены, спроса и предложения является локальной. Она дает прогноз на ближайшее будущее и не служит для вывода всеобъемлющих, глобальных уравнений изменения предложения, спроса, цены на продолжительные промежутки времени.

2.4.

МОДЕЛИ УПРАВЛЯЕМЫХ ПРОЦЕССОВ

В экономике управляемые процессы делятся на две группы: выбор специальных значений параметров модели такой, чтобы процесс развивался в нужном направлении, или организация какого-либо экзогенного фактора специальным образом, чтобы динамика процесса удовлетворяла требованиям управляющего. Примеры первого вида управления (параметрическое управление): налоговая политика государства, управление процентной ставкой Центрального банка, установление пенсионного возраста и другие макроэкономические рычаги правительства и парламента страны. Второй прием управления более характерен для микроэкономических задач предприятия или планирования муниципальной деятельности. Например, банк, страховая компания или пенсионный фонд могут ставить задачу разумного (или даже оптимального) инвестиционного распределения капитала с течением времени по различным государственным или частным проектам, предусматривая наибольший доход и наименьший риск. В данном параграфе рассматривается теория построения моделей динамики для многомерного экономического процесса с экзогенными управляющими переменными. Сначала будет изложена постановка задачи динамического распределения инвестиций, затем рассказаны методы формирования моделей и управления ими для дискретного и непрерывного времени.

2.4.1. ЗАДАЧА РАСПРЕДЕЛЕНИЯ ИНВЕСТИЦИЙ

Примерно 40 лет назад факультет прикладной математики — процессов управления — по заказу ленинградских городских исполнительных властей решал задачу оптимального распределения сил и средств. Задача решалась под руководством профессора В. И. Зубова и формулировалась следующим образом: пусть имеется ограниченный фиксированный бюджет (города, региона или другого административного образования), требуется его распределить по отраслям народного хозяйства так, чтобы либо отрасли достигли в своем развитии заданного уровня, либо определенные отрасли развивались быстрее и т. д. Основным в математической модели было условие — прирост отрасли пропорционален инвестициям.

Это известная задача плановой экономики в настоящее время, когда государство влияет на развитие отраслей косвенно (через дотации и налоги), нуждается в изменении. А именно, на первый план выходит зависимость прироста отрасли от его основных фондов (капиталов) и от развития смежных отраслей, а также, естественно, от дотаций и налогов. Поэтому резко усложняется составление математической модели процесса. Заметим, что в период существенных изменений в экономике (инфляция, структурная перестройка, приватизация и т. п.) любые математические модели являются нестационарными и нелинейными, а это приводит к необходимости постоянной проверки адекватности модели.

Мы предлагаем математический подход к описанной выше задаче, основанный на восстановлении линейной управляемой системы по дискретным наблюдениям за ее состоянием. Метод опробован на технологических процессах невысоких порядков. В задаче распределения при наличии достаточной информации о состоянии отраслей может быть составлен алгоритм (и соответствующее программное обеспечение), позволяющий давать дотации (на городском или республиканском уровнях) целенаправленно. Говорить об оптимальном управлении здесь нельзя, так как модель строится приближенно, но алгоритм будет

предусматривать постоянную коррекцию в зависимости от оценки правильности распределения.

Рассматривается задача распределения ресурсов в различные взаимосвязанные сферы при ограничении на общий объем ресурсов и отсутствии математической модели системы. Частным случаем этой задачи является распределение капиталовложений по отраслям [33]. Решение предлагается строить с помощью локального математического моделирования.

Пусть имеется n отраслей, количественные характеристики которых определяются на текущий момент t доступными измерениями величинами $x_1(t), x_2(t), \dots, x_n(t)$. Обозначим через $u_s(t)$ капиталовложения в отрасль s . Будем считать, что выделение капиталовложений в каждую из отраслей в текущий момент t осуществляется из общего количества:

$$\sum_{s=1}^n u_s(t) = u(t). \quad (75)$$

Пусть отрасли взаимосвязаны, и в момент $t = 0$ их состояние описывается величинами x_{s0} , $s = 1, 2, \dots, n$. Требуется построить функции $u_s(t)$, $t \in [0, T]$ так, чтобы выполнялось (75), $u_s(t) \geq 0$ и $x_s(T) = \bar{x}_s$, $s = 1, 2, \dots, n$, где \bar{x}_s — заданный уровень развития отрасли.

В работе [33] математическая модель либо не учитывает взаимосвязи отраслей и выражается в виде системы

$$\dot{x}_s = b_s u_s,$$

либо добавляет уравнение баланса. Модель хороша своей простотой, но оставляет в стороне вопрос о происхождении коэффициентов системы и коэффициентов балансовых соотношений.

Применим идеи локального моделирования в сформулированной выше задаче. Будем строить математическую модель развития отраслей в виде линейной стационарной системы

$$\dot{X}(t) = AX(t) + BU(t), \quad (76)$$

где $A = \{a_{ij}\}$, $B = \{b_{ij}\}$ — постоянные $(n \times n)$ -матрицы, $X = \{x_s\}$, $U = \{u_s\}$, $s = 1, 2, \dots, n$.

Как следует из предположений, измерению доступен весь вектор X в любые моменты времени. Пусть для некоторого шага измерений $h > 0$ известны векторы $X(0)$, $X(h)$, ..., $X(Nh)$, причем величина h настолько мала, что $Nh \ll T$ и $N > 2n + 1$. Тогда при условии, что вектор U не меняется на интервале $[0, (n + 1)h]$, введя новые переменные $Y_k = X(kh + h) - X(kh)$, $k = 0, 1, 2, \dots, n$, получим:

$$A = \frac{1}{h} Ln(\{Y_1, Y_2, \dots, Y_n\} \{Y_0, \dots, Y_{n-1}\}^{-1}), \quad (77)$$

$$BU = \left[\int_0^h e^{-As} ds \right]^{-1} [e^{-Ah} X(kh + h) - X(kh)]$$

для любого $k = 0, 1, \dots, n - 1$. (78)

Введем условие. Пусть вектор U меняется на интервале $[(n + 2)h, (2n + 1)h]$ на каждом шаге h так, чтобы матрица

$$\Omega = \{U((n + 2)h), \dots, U((2n + 1)h)\}$$

была невырожденная.

Если это условие выполнено, то из (78) следует:

$$B = \left[\int_0^h e^{-As} ds \right]^{-1} \{e^{-Ah} X((n + 2)h) - X((n + 1)h), \dots, \dots, e^{-Ah} X((2n + 2)h) - X((2n + 1)h)\} \Omega^{-1}. \quad (79)$$

Заметим, что в (78) обратная матрица обязательно существует для достаточно малых h . Покажем, что если шаг наблюдений в таблице $\{Y_1, Y_2, \dots, Y_{n+1}\}$ достаточно мал, то указанная матрица обязательно обратима.

Это следует из простого утверждения: существует $h_0 > 0$ такое, что при $h \in (0, h_0)$

$$\det \left[\int_0^h \exp(-A\tau) d\tau \right] \neq 0.$$

Действительно, известно, что существует невырожденная матрица S такая, что

$$\int_0^h \exp(-A\tau) d\tau = S^{-1} \int_0^h \exp(\Lambda\tau) d\tau S,$$

где Λ — жорданова форма матрицы $(-A)$. Следовательно, утверждение будет доказанным, если установим, что

$$\int_0^h \exp(\lambda_j \tau) d\tau \neq 0,$$

где λ_j — собственное число матрицы $(-A)$.

При $\lambda_j = 0$

$$\int_0^h \exp(\lambda_j \tau) d\tau = h \neq 0.$$

При $\lambda_j \neq 0$ интеграл отличен от нуля для $\lambda_j \neq 2\pi li/h$, где l — целое число. Последнее неравенство будет верно, если для всех j имеет место $\Re \lambda_j \neq 0$. Если же существует чисто мнимое λ_j , то h_0 найдется ввиду конечности спектра матрицы.

Получение B в виде (79) относится к более общему случаю, чем задача распределения ресурсов. Априори известно, что вложения капитала делаются в конкретную отрасль и причем обязательно. Значит, B должна быть диагональной и $u_s \neq 0$. Эти рассуждения позволяют из (78) получить

$$B = \text{diag}\{b_1, \dots, b_n\}, \quad b_s = \frac{1}{u_s} \left\{ \left[\int_0^h e^{-As} ds \right]^{-1} [e^{-Ah} X(nh + h) - X(nh)] \right\}_s.$$

Таким образом, введенное условие $\det \Omega \neq 0$ не является обязательным для функционирования модели.

В результате модель (76) построена. Для того чтобы равенство (75) выполнялось для всех $t \in [0, T]$, исключим u_n из рассмотрения. Тогда окончательно получим математическую модель динамики инвестиций:

$$\dot{X} = AX + \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{n-1} \\ -b_n & -b_n & \dots & -b_n \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_{n-1} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ b_n u(t) \end{pmatrix}. \quad (80)$$

Управление $\{u_1(t), \dots, u_{n-1}(t)\}$, переводящее $X(t)$ из состояния $X(Nh)$ в \bar{X} , строится известным образом. Обозначим:

$$Z(t) = e^{-At} \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{n-1} \\ -b_n & -b_n & \dots & -b_n \end{pmatrix};$$

$$P(T) = \int_{Nh}^T Z(t)Z^T(t)dt.$$

Если $\det P(T) \neq 0$, то:

$$\begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_{n-1} \end{pmatrix} = Z^T(t)C + \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_{n-1} \end{pmatrix}, \quad (81)$$

где

$$C = P^{-1}(T) \left[e^{-AT} \bar{X} - \int_{Nh}^T e^{-A\tau} \begin{pmatrix} 0 \\ \dots \\ b_n u(\tau) \end{pmatrix} d\tau - e^{-ANh} X(Nh) \right],$$

$$\int_{Nh}^T Z(\tau) \begin{pmatrix} v_1(\tau) \\ \dots \\ v_{n-1}(\tau) \end{pmatrix} d\tau = \begin{pmatrix} 0 \\ \dots \\ 0 \end{pmatrix}. \quad (82)$$

При этом ортогональные добавки $v_s(t)$ надо подбирать так, чтобы $u_1(t), \dots, u_n(t) > 0$.

Последние условия на ортогональные добавки $v_s(t)$ не исчерпывают возможностей уравнения (81). Покажем это на модельном примере.

Пусть x_i — выпуск i -го продукта за единицу времени, p_i — цена единицы i -го продукта, $\alpha \in (0, 1)$ — уровень налога (например, НДС). Тогда правительство собирает налог в виде:

$$I = \alpha \int_0^T \sum_{i=1}^n p_i x_i(t) dt.$$

Учитывая, что $p_i > 0$, $x_i > 0$, сумма налога будет положительна. Из (80) и (81) получим;

$$X(t) = e^{At} X_0 + e^{At} \int_0^t e^{-A\tau} \left[BU + \begin{pmatrix} 0 \\ \dots \\ b_n u(\tau) \end{pmatrix} \right] d\tau.$$

Обозначая $L^T = (p_1, \dots, p_n)$, $\bar{V}^T = (v_1, \dots, v_{n-1})$, можем вычислять сумму налога как функцию вектора \bar{V} :

$$I = \alpha L^T \int_0^T e^{At} dt X_0 + \alpha L^T \int_0^T e^{At} \int_0^t e^{-A\tau} B \left[B^T e^{-A^T \tau} + \begin{pmatrix} 0 \\ \dots \\ b_n u(\tau) \end{pmatrix} \right] d\tau dt + \\ + \alpha L^T \int_0^T e^{At} \int_0^t e^{-A\tau} BV(\tau) d\tau dt.$$

Последний интеграл упрощается интегрированием по частям, если матрица A невырожденная. Тогда наибольшее значение I достигается при $V = \bar{V}$:

$$\bar{V} = \arg \max \left[-\alpha L^T A^{-1} B \int_0^T V(\tau) d\tau \right].$$

Здесь максимум достигается при выполнении ограничений ортогональности (82) и содержательных неравенств $u_i > 0$. Вопрос о существовании допустимых управлений остается открытым.

В задачах численного моделирования очень важным является вопрос об адекватности математической модели. Будем считать решение уравнения (80) с управлением (81) модельным решением в отличие от истинного, доступного измерениям. Предположим, что начиная с момента $t^* > Nh$ модельное решение удаляется от истинного. Тогда прекращаем управление по правилу (81), фиксируем распределение $u_1(t^*), \dots, u_n(t^*)$ и вновь вычисляем модель. Такой способ позволяет проверять адекватность не модели, а закона управления, гибко реагировать на нестационарность реального процесса.

2.4.2.
ВОССТАНОВЛЕНИЕ
РАЗНОСТНЫХ ЛИНЕЙНЫХ СИСТЕМ.
СПЕЦИАЛЬНЫЙ ВХОД

Алгоритм Хоу. Пусть имеется система линейных разностных уравнений с m «входами» и p «выходами»:

$$X_{k+1} = FX_k + GU_k, \quad Y_k = HX_k, \quad k = 0, 1, 2, \dots, \quad (83)$$

где $X \in R^n, U \in R^m, Y \in R^p; F, G, H$ — постоянные матрицы размерностей $(n \times n), (n \times m), (p \times n)$, соответственно. Поставим задачу определения коэффициентов системы (83) по «выходам» системы, если «входы» также известны, а точнее «входы» являются единичным импульсным воздействием, т. е. при $k = 0$ на «входы» системы подается единичное воздействие, а при $k = 1, 2, 3, \dots U_k = (0, \dots, 0)^T$. Начальное состояние X_0 считаем нулевым. Так как единичное воздействие может подаваться на любом из «входов», то будем рассматривать сразу все единичные векторы:

$$\bar{U}_0 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

— $(m \times m)$ -единичная матрица «входов».

Соответственно \bar{X}_k — $(n \times m)$ -матрица состояний, \bar{Y}_k — $(p \times m)$ -матрица «выходов». В этих обозначениях имеем:

$$\bar{Y}_k = HF^{k-1}G, \quad k = 1, 2, 3, \dots \quad (84)$$

Сформулируем математически задачу: определить матрицы H, F, G по последовательности матриц $\{\bar{Y}_k\}$ с минимальным числом n . Так ставится задача в теории реализации Калмана [39].

Алгоритм Хоу дает одно из возможных решений задачи в предположении, что последовательность $\{\bar{Y}_k\}$ удовлетворяет (84), т. е. система линейна, стационарна и конечномерна (исходная модель (83) имеет постоянные коэффициенты и конечную размерность).

Построим блочную матрицу из последовательности $\{\bar{Y}_k\}$ следующим образом:

$$\bar{H}_{N'N} = \begin{pmatrix} \bar{Y}_1 & \bar{Y}_2 & \dots & \bar{Y}_N \\ \bar{Y}_2 & \bar{Y}_3 & \dots & \bar{Y}_{N+1} \\ \dots & \dots & \dots & \dots \\ \bar{Y}_{N'} & \bar{Y}_{N'+1} & \dots & \bar{Y}_{N+N'-1} \end{pmatrix}.$$

Утверждение 1. Если $\{\bar{Y}_k\}$ удовлетворяет (84), то существует такое число r , что для некоторых β_1, \dots, β_r выполняются матричные равенства:

$$\bar{Y}_{r+k+1} = -\sum_{i=1}^r \beta_i \bar{Y}_{i+k}, \quad k = 0, 1, 2, \dots \quad (85)$$

Действительно, матрица F неизвестна, однако для нее существуют числа r и β_1, \dots, β_r , такие, что

$$F^r + \beta_r F^{r-1} + \beta_{r-1} F^{r-2} + \dots + \beta_1 E = 0. \quad (86)$$

Это следует из теоремы Гамильтона–Кели для $r = n$. Может быть и $r < n$, если минимальный аннулирующий полином не совпадает с характеристическим полиномом матрицы F . Здесь n — размерность квадратной матрицы F , которая неизвестна нам, но обязательно конечна в силу предположения утверждения. Умножим (86) справа на $F^k G$, а слева на H . Получим (85) при $k = 0, 1, 2, \dots$, что и требовалось показать.

Определим обобщенную единичную матрицу

$$E_n^m = \begin{cases} \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{pmatrix} & \text{при } m < n, \\ \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix} & \text{при } m > n, \\ \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} & \text{при } n = m. \end{cases}$$

Утверждение 2. Если известны числа r и β_1, \dots, β_r из утверждения 1, то можно считать

$$H = E_{pr}^p, \quad G = \bar{H}_{rr} E_m^{mr},$$

$$F = C = \begin{pmatrix} O_p & E_p & O_p & \dots & O_p \\ O_p & O_p & E_p & \dots & O_p \\ \dots & \dots & \dots & \dots & \dots \\ O_p & O_p & O_p & \dots & E_p \\ -\beta_1 E_p & -\beta_2 E_p & -\beta_3 E_p & \dots & -\beta_r E_p \end{pmatrix}, \quad (87)$$

где O_p — нулевая ($p \times p$)-матрица, $E_p = E_p^p$.

Действительно, утверждение будет доказано, если из (87) будет следовать (84) для $k = 1, 2, \dots$. Определим операцию сдвига для матрицы \bar{H}_{rr} :

$$\sigma \bar{H}_{rr} = \begin{pmatrix} \bar{Y}_2 & \bar{Y}_3 & \dots & \bar{Y}_{r+1} \\ \bar{Y}_3 & \bar{Y}_4 & \dots & \bar{Y}_{r+2} \\ \dots & \dots & \dots & \dots \\ \bar{Y}_{r+1} & \bar{Y}_{r+2} & \dots & \bar{Y}_{2r} \end{pmatrix}.$$

В силу (85) имеем:

$$\sigma \bar{H}_{rr} = C \bar{H}_{rr},$$

и далее по индукции можно установить, что

$$\sigma^k \bar{H}_{rr} = C^k \bar{H}_{rr} \quad \text{при } k = 0, 1, 2, \dots \quad (88)$$

Легко можно убедиться, что

$$\bar{Y}_k = E_{pr}^p (\sigma^{k-1} \bar{H}_{rr}) E_m^{mr}, \quad k = 1, 2, \dots$$

Используя (87) и (88), получаем:

$$\bar{Y}_k = E_{pr}^p C^{k-1} \bar{H}_{rr} E_m^{mr} = HF^{k-1}G.$$

Что и требовалось показать.

Отметим, что в данном случае $n = pr$, т. е. эта реализация не обязательно минимальная.

Теорема Хоу. Если число r удовлетворяет условию (85), то последовательность $\{\bar{Y}_k\}$ реализуется с помощью матриц:

$$\begin{aligned}
 F &= E_{pr}^n P(\sigma \bar{H}_{rr}) M E_n^{mr}, \\
 G &= E_{pr}^n P \bar{H}_{rr} E_m^{mr}, \\
 H &= E_{pr}^p \bar{H}_{rr} M E_n^{mr},
 \end{aligned} \tag{89}$$

где матрицы P и M квадратные, невырожденные и определяются равенством

$$P \bar{H}_{rr} M = E_n^{pr} E_m^n. \tag{90}$$

$n = \text{rang } \bar{H}_{rr}$.

Доказательство. Для простоты будем писать в доказательстве обозначения матриц без указания несущественных размерностей. Так, вместо \bar{H}_{rr} будем писать без индексов \bar{H} . Вспомним, что верхний индекс «+» мы вводили для обозначения псевдообратной матрицы:

$$\bar{H} \bar{H}^+ \bar{H} = \bar{H}. \tag{91}$$

Убедимся, что

$$\bar{H}^+ = M E_n E^n P. \tag{92}$$

Для этого умножим (92) справа и слева на \bar{H} , воспользуемся слева определением псевдообратной матрицы (91): $\bar{H} = \bar{H} M E_n E^n P \bar{H}$. Теперь умножим последнее равенство слева на P , а справа на M . Применяя три раза определение (90) к полученному равенству, убеждаемся, что оно верное.

Далее проверим формулы (89), многократно используя тождества (90)–(92), а также операцию сдвига. Нам необходимо из (89) получить соотношения (84): $\bar{Y}_k = H F^{k-1} G$.

Согласно определению операции сдвига:

$$\bar{Y}_k = E^p [\sigma^{k-1} \bar{H}] E_m =$$

(по формуле (88) получаем)

$$= E^p C^{k-1} \bar{H} E_m =$$

(по определению псевдообратной матрицы)

$$= E^p C^{k-1} \bar{H} \bar{H}^+ \bar{H} E_m =$$

(из формулы (92))

$$= E^p C^{k-1} \bar{H} M E_n E^n P \bar{H} E_m =$$

(из формул (89) для G)

$$= E^p C^{k-1} \bar{H} M E_n G =$$

(далее используем матрицу

$$D = \begin{pmatrix} O_m & O_m & \dots & O_m & -\beta_1 E_m \\ E_m & O_m & \dots & O_m & -\beta_2 E_m \\ O_m & E_m & \dots & O_m & -\beta_3 E_m \\ \dots & \dots & \dots & \dots & \dots \\ O_m & O_m & \dots & E_m & -\beta_r E_m \end{pmatrix},$$

так же, как ранее использовали матрицу C : $C\bar{H} = \bar{H}D$)

$$= E^p \bar{H} D^{k-1} M E_n G =$$

(по определению псевдообратной матрицы)

$$= E^p \bar{H} \bar{H}^+ \bar{H} D^{k-1} M E_n G =$$

(из формулы (92))

$$= E^p \bar{H} M E_n E^n P \bar{H} D^{k-1} M E_n G =$$

(по формуле (89) для матрицы H)

$$= \bar{H} E^n P \bar{H} D^{k-1} M E_n G =$$

(в силу равенства $C\bar{H} = \bar{H}D$)

$$= \bar{H} E^n P C^{k-1} \bar{H} M E_n G =$$

(остается только показать, что)

$$= E^n P C^{k-1} \bar{H} M E_n = (E^n P C \bar{H} M E_n)^{k-1} = F^{k-1}.$$

Для случая $k = 3$ имеем:

$$\begin{aligned} F^2 &= E^n P C \bar{H} M E_n E^n P C \bar{H} M E_n = E^n P C \bar{H} \bar{H}^+ C \bar{H} M E_n = \\ &= E^n P C \bar{H} \bar{H}^+ \bar{H} D M E_n = E^n P C \bar{H} D M E_n = E^n P C^2 \bar{H} M E_n. \end{aligned}$$

В общем случае последнее соотношение доказывается по индукции. Теорема доказана.

Рассмотрим пример с одним «входом» и одним «выходом», т. е. когда $p = m = 1$.

Пример. $\{\bar{Y}_k\} = 1, 1, 1, 2, 1, 3, \dots$

Определить $\{\bar{Y}(7)\}, \{\bar{Y}(8)\}$ и т. д.

Составим матрицу

$$\bar{H} = \begin{pmatrix} 1 & 1 & 1 & 2 & 1 & 3 & \dots \\ 1 & 1 & 2 & 1 & 3 & \dots & \dots \\ 1 & 2 & 1 & 3 & \dots & \dots & \dots \\ 2 & 1 & 3 & \dots & \dots & \dots & \dots \\ 1 & 3 & \dots & \dots & \dots & \dots & \dots \\ 3 & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Попытаемся найти число r .

Если $r = 1$, то должно быть $\bar{Y}(k+1) = -\beta\bar{Y}(k)$. По первым числам получим: $\beta = -1$, но $\bar{Y}(4) = 2$. Значит $r \neq 1$.

Если $r = 2$, то должно быть

$$\bar{Y}(k+2) = -\beta_1\bar{Y}(k) - \beta_2\bar{Y}(k+1).$$

Составим систему для нахождения β_1 и β_2 :

$$\begin{pmatrix} -\bar{Y}(1) & -\bar{Y}(2) \\ -\bar{Y}(2) & -\bar{Y}(3) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \bar{Y}(3) \\ \bar{Y}(4) \end{pmatrix}.$$

Система не имеет решения, так как матрица коэффициентов вырождена. Значит, $r \neq 2$.

Пусть $r = 3$. Составим систему для нахождения $\beta_1, \beta_2, \beta_3$:

$$\begin{pmatrix} -\bar{Y}(1) & -\bar{Y}(2) & -\bar{Y}(3) \\ -\bar{Y}(2) & -\bar{Y}(3) & -\bar{Y}(4) \\ -\bar{Y}(3) & -\bar{Y}(4) & -\bar{Y}(5) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} \bar{Y}(4) \\ \bar{Y}(5) \\ \bar{Y}(6) \end{pmatrix}.$$

Система имеет единственное решение. Отсюда следует, что $n = r = 3$. Согласно теореме Хоу,

$$F = P \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix} M, \quad G = P \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad H = (1 \ 1 \ 1)M,$$

где квадратные неособые P и M должны быть получены из

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} M.$$

Положим $P = E$. Тогда

$$M = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}.$$

Тогда

$$F = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & -1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad H = (1 \ 0 \ 0),$$

$$X(k+1) = FX(k) + G\bar{U}.$$

И, следовательно,

$$X(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad X(1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad X(2) = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad X(3) = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad X(4) = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix},$$

$$X(5) = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \quad X(6) = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}, \quad X(7) = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}, \quad X(8) = \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}.$$

Теперь, пользуясь равенством $\bar{Y}(k) = HX(k)$, получаем ответ: $\bar{Y}(7) = 2, \bar{Y}(8) = 3$.

Отметим, что выбор матрицы P был произволен: мы могли бы выбрать любую невырожденную матрицу, а по ней однозначно определить M . Покажем, что в одномерном случае указанная неоднозначность равносильна неоднозначному выбору базиса для внутренней переменной $X(k)$ (при $n = r$).

Введение нового базиса означает обратимую замену переменных $X = D\tilde{X}$ ($\det D \neq 0$). Тогда система (83) принимает вид:

$$\tilde{X}_{k+1} = D^{-1}FD\tilde{X}_k + D^{-1}G\bar{U}_k, \quad \bar{Y}_k = HD\tilde{X}_k,$$

но при $\tilde{X}(0) = 0, \bar{U}_0 = 1$ и $\bar{U}_k = 0$ для $k = 1, 2, \dots$ имеем

$$\bar{Y}_k = (HD)(D^{-1}FD)^{k-1}(D^{-1}G) = HF^{k-1}G,$$

т. е. «выход» системы от выбора базиса не зависит.

С другой стороны, при $n = r$ и $p = m = 1$ из (67) имеем $F = P(\sigma \bar{H}_{rr})M$ для одного базиса и

$$D^{-1}FD = D^{-1}P(\sigma \bar{H}_{rr})MD = \hat{P}(\sigma \bar{H}_{rr})\hat{M},$$

причем для \hat{P} и \hat{M} удовлетворяется неравенство (90):

$$D^{-1}P\bar{H}_{rr}MD = E_n^n.$$

Рассмотрим теперь модельный пример применения теории реализации Калмана. Особое внимание обратим на то, как сложно в экономических моделях организовать единичное импульсное воздействие, откликом на которое и является выход системы. Если в инженерных лабораториях сравнительно легко построить эксперимент и повторить его много раз, то отдельные люди и социальные группы неохотно идут на такие манипуляции. Тем не менее, составленный ниже пример может служить образцом организации маркетинга нового товара только за счет налаженной информационной системы.

Пусть требуется определить количество лекарства определенного вида для города. Специалист знает, что это лекарство может использоваться для лечения 2 болезней. Он выбирает время, когда количество больных этими болезнями стабилизировалось, т. е. не меняется в течение, например, 10 дней (в данном примере все величины условные). Затем он равномерно по всем аптекам города мгновенно распределяет порцию, которую принимает за единицу лекарства (возможно, есть необходимость коррекции величины порции ввиду незамкнутости системы, т. е. аптеки доступны не только жителям города). Сразу после этого специалист собирает из поликлиник города информацию о количестве больных указанными болезнями с частотой 1 раз в 2 дня. При этом порция лекарств должна быть небольшой, чтобы население разобрало его за короткий промежуток, в идеальном случае за те же 2 дня. Информация о количестве больных, точнее об отклонении количества больных от принятого за нулевой уровень, и есть отклик системы на единичное импульсное воздействие (табл. 3). Собрав достаточное количество информации, можно приступить к восстановлению системы (83). В данном случае $m = 1, p = 2$.

Таблица 3

**Данные для построения модели лекарственного воздействия
(маркетинга нового лекарства)**

k	День наблюдения	$\bar{Y}_1(k)$	$\bar{Y}_2(k)$	k	День наблюдения	$\bar{Y}_1(k)$	$\bar{Y}_2(k)$
1	2	-1	-10	5	10	-6	-15
2	4	0	-20	6	12	-3	-10
3	6	-4	-30	7	14	-1	0
4	8	-8	-25				

Теперь может быть решена задача восстановления разностной системы по приведенным наблюдениям и определено, какое количество лекарства необходимо продавать в аптеках, чтобы уровень болезни не превышал заданных пределов, например, $\bar{Y}_1 < -10$, $\bar{Y}_2 < -50$.

**2.4.3.
ВОССТАНОВЛЕНИЕ
РАЗНОСТНЫХ ЛИНЕЙНЫХ СИСТЕМ.
ПРОИЗВОЛЬНЫЙ ВХОД**

Мы начали этот параграф с задачи распределения капиталовложений. Рассмотрели метод локального моделирования, основанный на минимально возможной информации. Затем проанализировали один из алгоритмов теории реализации для нахождения минимальной по размерности модели, соответствующей данному выходу. И в том и в другом случае для построения модели требовался некоторый специальный вход, т. е., другими словами, экзогенные переменные должны быть задаваемы заранее. В экономике это практически невозможно. Инженерные науки (да и практика) допускают проведение специальных экспериментов по идентификации модели, скажем, в лабораторных условиях создаются подобию реальным объектам и на них уточняются параметры, совершенствуются предположения модели и т. д. Экономические задачи характерны тем, что входные воздействия не являются произвольными. Частично они доступны измерению, следовательно, можно считать их известными функциями времени.

Возвращаясь к задаче о распределении капитала, заметим, что более реальна следующая организация входа в систему. Пусть инвестор имеет план вкладывать $I(t)$ капитала в момент времени t (здесь считаем t дискретным временем). Распределение по предприятиям (или проектам) предполагаем постоянным (собственно, это и есть «портфель»), $g^T = \{g_1, \dots, g_n\}$, так что

$$g_i \geq 0, \quad \sum g_i = 1.$$

Таким образом, в момент времени t согласно плану инвестиций i -ое предприятие получит $g_i I(t)$ капитала. Такая же ситуация наблюдается, когда государство из бюджета направляет трансфертные платежи в дотационные регионы. План платежей утверждается вместе с бюджетом, причем распределение g и график выплат $I(t)$ обсуждаются отдельно (имеют самостоятельное значение).

Предположим, что каждое предприятие, рассматриваемое инвестором в плане капиталовложений, описывается скалярной характеристикой x_i , смысл которой может быть разным в зависимости от доступности (например, текущая стоимость акций на рынке ценных бумаг) или от прямой значимости в модели (например, выход продукции предприятия на единицу капиталовложений). Ясно, что в первом случае модель нужна экономическому субъекту, не допущенному к фактическим данным предприятия. Во втором случае имеет место полная информированность инвестора. Так должно быть при анализе работы муниципальной собственности (или с частичным муниципальным капиталом), так как сравнительно небольшое количество предприятий может быть опрошено с достаточной частотой и строгостью. Составим из переменных x_i вектор:

$$X(t) = \{x_i(t)\}, \quad i = 1, \dots, n.$$

Рассмотрим самую простую ситуацию: нам известно поведение предприятий и объемы инвестиций при $t = 1, 2, \dots, n + 1$, вектор g также известен. Предположим, что предприятия в своей деятельности связаны между собой и выход их продукции (или популярность их акций) меняется в зависимости от их текущего состояния и инвести-

ционного потока. Тогда в первом приближении мы можем считать, что модель динамики этих предприятий линейна:

$$X(t+1) = AX(t) + gI(t), \quad t = 1, 2, \dots, \quad (93)$$

где A — постоянная матрица модели, которую необходимо получить из наблюдений за работой предприятий при $t < n + 2$. Пользуясь тем же приемом, что и в начале этого параграфа, получим решение задачи восстановления матрицы коэффициентов:

$$A = \{X(2) - gI(1), X(3) - gI(2), \dots, X(n+1) - gI(n)\} \{X(1), \dots, X(n)\}^{-1}. \quad (94)$$

Если $\det\{X(1), \dots, X(n)\} \neq 0$, то модель (93) с матрицей (94) позволяет прогнозировать поведение предприятий при $t < n + 2$. Это, в свою очередь, позволит оценить правильность планирования или даже построить оптимальный в том или ином смысле план. Если же векторы наблюдений оказываются линейно зависимыми, то необходимо пересмотреть правильность представления производств в системе. Например, если в небольшом городке имеется так называемое «градообразующее» предприятие, то очевидно многие мелкие муниципальные производства или сервисные центры будут зависимы от успехов или неудач главного производства, и тогда показатели этих центров можно считать зависимыми. Таким образом, часть переменных системы x_i выражается линейно через оставшиеся так, чтобы модель содержала только динамически линейно независимые переменные. Технически это осуществляется так. Пусть ранг матрицы $\{X(1), \dots, X(n)\}$ равен $k < n$. Тогда выделим из последней матрицы k линейно независимых строк и будем считать соответствующие переменные главными (и соответствующие предприятия), а оставшиеся — зависимыми.

2.5. ЗАКЛЮЧЕНИЕ

В этой главе представлены некоторые модели прогнозирования динамики. Все они объединены тем, что основаны на линейных системах дифференциальных, разностных или функционально-дифференциальных (с временным

лагом) уравнений. Рассмотрены характерные свойства таких моделей, приведены конструктивные алгоритмы, дан анализ применимости к конкретным задачам экономики. В качестве содержательных примеров обсуждены задачи о форвардных контрактах, о рыночном равновесии цены, о распределении инвестиций (или бюджетных трансфертах) и др. Важно отметить, что линейные системы позволяют управлять процессами в экономике, причем, не только выбирая оптимальные параметры, но и формируя соответствующие обратные связи.

Однако самым большим достоинством линейных моделей является их широкая распространенность в среде математиков, инженеров, экономистов, т. е. тех, кто применяет методы моделирования. Нелинейные модели являются более богатыми в функциональном смысле, они не ограничены только полиномами, экспонентами и простейшими периодическими колебаниями. Поэтому, отдавая дань линейным моделям, мы призываем изучать и применять менее известные, нелинейные.

ГЛАВА ТРЕТЬЯ

МОДЕЛИ ЛОТКИ–ВОЛЬТЕРРЫ В ЭКОНОМИКЕ

Данная глава содержит описание, анализ и способы применения замечательной модели, известной в современной литературе как модель Лотки–Вольтерры [16, 83]. Чаще всего ее используют для описания процессов в биологии, медицине и экологии [72]. В последнее время ею заинтересовались экономисты [62]. Ниже мы развиваем модель Лотки–Вольтерры, дополняя ее новыми результатами и приложениями.

Структура главы организована так, чтобы побудить экономистов, не имеющих неприязни к математике, шире привлекать методы нелинейной динамики к экономическим процессам. В начале главы описываются формально-экономические постановки динамических ситуаций, где линейными моделями не обойтись.

Особую роль в моделировании играет временной лаг. Уравнения с временным лагом (или, иначе говоря, с последствием) имеют адекватную интерпретацию в экономических терминах, но этого мало. Оказывается, что попытки игнорировать существование временной задержки в причинно-следственных связях часто ведут к ошибочным заключениям.

В этой главе исследована нелинейная модель Лотки–Вольтерры с запаздывающим аргументом. Показаны условия устойчивости и неустойчивости положений равновесия, наличия колебаний и т. п. Важное место в ней занимает идентификация коэффициентов модели и величины запаздывания. Все эти результаты, объединенные в одной

книге, дают разносторонний анализ и конструктивные алгоритмы, которые позволяют постепенно сокращать использование в приложениях линейных моделей и заменять их более богатыми, нелинейными с запаздыванием.

3.1. ОСНОВНЫЕ ИДЕИ, МЕТОДЫ И РЕЗУЛЬТАТЫ

3.1.1. ОПИСАНИЕ МОДЕЛИ

Считается, что первая формализация динамики биологических популяций восходит к Мальтусу, предположившему, что скорость изменения числа особей N в популяции пропорциональна объему популяции:

$$dN(t) = \varepsilon N(t)dt.$$

Естественно, это уравнение приводит к экспоненциальному росту объема популяции, чего в природе не наблюдается (в течение продолжительного времени).

В следующей модели был исключен бесконечный рост и был принят во внимание эффект насыщения. Она появилась в 1838 г. в работе Ферхюльста под названием «логистическая» модель и использовалась для описания динамики биологических популяций:

$$dN(t) = [\varepsilon_0 - \gamma N(t)]N(t)dt.$$

Здесь коэффициент прироста ε заменен линейной функцией от объема популяции, что интерпретируется следующим образом: если жизненные ресурсы популяции ограничены, т. е. пища, жизненное пространство, энергетические источники и т. д. ограничены, то объем популяции должен с течением времени стабилизироваться, а не расти бесконечно. В последнем уравнении такой уровень объема популяции равен ε_0/γ .

В наиболее общей форме (без последствия) логистическая модель рассматривалась А. Н. Колмогоровым [45] для сообщества из n видов:

$$\dot{N}_i(t) = N_i(t)f_i[N_1(t), \dots, N_n(t)], \quad i = 1, \dots, n.$$

Функции f_i удовлетворяют некоторым специальным свойствам, которые допускают простую биологическую интерпретацию. На этом обобщении мы не будем останавливаться, так как слишком большой произвол в выборе указанных функций препятствует созданию конструктивных методов идентификации и модели.

Сам В. Вольтерра использовал линейные функции:

$$\dot{N}_i(t) = N_i(t) \left[\varepsilon_i - \sum_{k=1}^n \gamma_{ik} N_k(t) \right], \quad i = 1, \dots, n. \quad (95)$$

Он исследовал случаи $n = 2, 3$, а также сделал несколько обобщений по системе с произвольным n .

Особую ценность работе В. Вольтерры придает то, что он, специалист в области интегральных и интегро-дифференциальных уравнений, ввел в рассмотрение последствие, т. е. он предположил, что взаимодействие видов на общих ресурсах изменяет коэффициенты прироста через некоторое время. Так, в задаче «хищник–жертва» предполагается, что объем популяции «жертв», являясь пищей для «хищников», определяет объем популяции последних через вегетативный период. В результате некоторых логических шагов В. Вольтерра получает систему с распределенным запаздыванием:

$$\begin{cases} \dot{N}_1(t) = \left[\varepsilon_1 - \gamma_1 N_2(t) - \int_0^{+\infty} F_1(s) N_2(t-s) ds \right] N_1(t), \\ \dot{N}_2(t) = \left[-\varepsilon_2 + \gamma_2 N_1(t) + \int_0^{+\infty} F_2(s) N_1(t-s) ds \right] N_2(t). \end{cases}$$

Здесь N_1, N_2 — объемы популяций «жертвы» и «хищника» соответственно, а F_1, F_2 — функции, характеризующие распределение «хищников» и «жертв» по возрастам.

Если считать, что запаздывание единственное, то без ограничений на коэффициенты модель Лотки–Вольтерры с последствием имеет вид:

$$\dot{N}_i(t) = N_i(t) \left[\varepsilon_i - \sum_{k=1}^n \gamma_{ik} N_k(t - \tau) \right], \quad i = 1, \dots, n. \quad (96)$$

Именно этой системе дифференциальных уравнений с последствием мы будем далее уделять наибольшее внимание.

3.1.2. ЦЕЛИ ИССЛЕДОВАНИЯ И ТЕХНИКА АНАЛИЗА

Поскольку предполагается, что эту книгу будут читать как математики, так и экономисты, есть смысл остановиться на тех задачах, которые могут быть решены предлагаемыми методами.

С точки зрения математика, система нелинейных дифференциальных уравнений с последствием является самостоятельным объектом исследования. Обычно в таком случае математики рассматривают следующие вопросы:

- существование и единственность решений задачи Коши на конечном и бесконечном интервалах времени;
- существование стационарных точек (положений равновесия системы);
- в более общем смысле, существование инвариантных множеств;
- исследование свойств устойчивости и неустойчивости инвариантных множеств;
- колебательность решений, т. е. существование периодических, почти периодических или рекуррентных решений.

Так как теория дифференциальных уравнений с последствием еще не является неотъемлемой частью образования математиков, то в этой книге дано большое приложение, являющееся фактически современным учебником по упомянутым выше вопросам. В нем все необходимые факты теории приведены с подробными доказательствами, что, с нашей точки зрения, важно для тех математиков, которые захотят приложить свои усилия к развитию теории. Здесь же, в тексте, мы будем ограничиваться бездоказательными утверждениями, отсылая читателя к соответствующему месту приложения.

Для экономиста применение математических моделей с целью прогнозирования поведения системы сопряжено с рядом условий, которые требуют ответов на очень слож-

ные, с формально математической позиции, вопросы, а именно:

- какие экономические предположения легли в основу модели;
- как идентифицировать все параметры модели и по какой информации;
- не являются ли прогнозируемые качественные картины поведения экономической системы тривиальным следствием предположений?

В этом смысле условия адекватности и целесообразности моделирования часто приводят экономистов к двум крайностям: либо безраздельно верить модели и ее формальным следствиям, либо избегать математического прогнозирования поведения модели.

Мы бы хотели видеть содержательное применение модели к качественным и количественным прогнозам динамики, но в то же время с честной интерпретацией экономических предположений.

Один биолог, специалист по эволюции видов, утверждал как-то, что «снежный человек» исчез как биологический вид из-за конкуренции с человеком за общие жизненные ресурсы и что это в точности соответствует выводам В. Вольтерры, полученным на основании модели конкуренции двух видов за один ресурс. И, действительно, в работе [16] рассматривается модель

$$\begin{cases} \dot{N}_1(t) = [\varepsilon_1 - \gamma_1 F(N_1, N_2)]N_1(t), \\ \dot{N}_2(t) = [\varepsilon_2 - \gamma_2 F(N_1, N_2)]N_2(t), \end{cases}$$

где $F(N_1, N_2)$ — монотонно возрастающая по каждому аргументу положительная функция, такая, что

$$F(N_1, N_2)_{N_1+N_2 \rightarrow +\infty} \rightarrow +\infty.$$

Биологически $F(N_1, N_2)$ интерпретируется как количество пищи, поедаемой за единицу времени. По Вольтерре, объем популяции одного из видов стремится к нулю с течением времени, тогда как объем другого достигает некоторого конечного предела.

Однако человек и «снежный человек» едят разную пищу, по-разному используют природные ресурсы и энергию.

И если человек постоянно совершенствует технологии жизнеобеспечения, то «снежный человек» приспособляется чисто биологически. Поэтому к этим двум видам нельзя применить последнюю модель без более подробного анализа коэффициентов воспроизводства, стоящих в уравнениях в квадратных скобках.

Возможно, что в силу тех же моделей Лотки–Вольтерры с последствием, которые будут изучаться ниже, динамика системы человек–«снежный человек» имеет колебания с периодом в несколько тысячелетий, согласно которым цивилизация человека рухнет, а биологически адаптивный «снежный человек» увеличит на Земле свою популяцию за следующие тысячелетия.

В любом случае, прежде чем формулировать прогноз по математической модели, необходимо как-то подтвердить ее адекватность. К сожалению, у экономиста очень мало информации для этого. Биолог может опробовать модель даже в лабораторных условиях, т. е. повторить всю динамику с теми же внешними воздействиями. Экономист лишен этой возможности. Он может только, в лучшем случае, лишь собрать все имеющиеся в истории факты и выделить из них близкие к рассматриваемой ситуации. Однако на этом пути часто субъективизм исследователя играет главную роль.

3.2.

ЭКОНОМИЧЕСКИЕ ИНТЕРПРЕТАЦИИ

3.2.1.

ОБЩИЕ РАССУЖДЕНИЯ

В этом разделе мы обсудим математическую формализацию динамики взаимодействия нескольких фирм, отраслей или государств. Начнем с самого простого — существуют две конкурирующие фирмы в одной экономической нише, т. е. с общими ресурсами, потребителями и одинаковыми товарами или услугами. Предположим, что цена не меняется при изменении объема товара на рынке (т. е. рассматриваемые фирмы не являются монополистами) и что товар незамещаем. Примерами таких конкурирующих фирм могут быть две транспортные компании по перевоз-

ке грузов или людей в данном городе. Они могут менять организацию труда, технологию обслуживания, вводить скидки, тратиться на рекламу, покупать более совершенные транспортные средства, сокращать стоимость ремонта и хранения, выбирать новые оптимальные маршруты, чтобы победить конкурентов. Можно рассматривать их взаимодействие на фоне существующего мощного и не подверженного частым изменениям городского транспортного хозяйства, т. е. предположить, что потребитель откажется от услуг рассматриваемых компаний, когда городской транспорт станет предпочтительней по цене или по качеству.

Другие примеры — это производство хлебо-булочных изделий, услуги по ремонту квартир или автомобилей, туристический бизнес и т. д.

Ясно, что такие конкурентные модели могут иметь большие размерности по числу фирм, но эту сложность преодолевают, объединяя фирмы с близкими технологиями производства товаров (или услуг) простым суммированием объемов товара на рынке.

Более сложным примером экономического взаимодействия является система мирового рынка. Так как мировые ресурсы труда, капитала, энергоносителей, земли и т. д. ограничены, то рассматриваемая динамическая модель конкуренции между странами вполне может оказаться адекватной. Конечно, невозможно составить систему по всем государствам и по всем товарам, но представляется разумным ограничиться объединениями стран с примерно равными ВВП: США + Канада + Мексика, Европа, Япония + Ю. Корея + Тайвань, Россия + Бразилия + арабские страны и т. д.; или иначе: Северная Америка, Южная Америка, Европа, Азия, Африка.

Для такой модели экономическая интерпретация требует существенно более сложных предположений: многие компании, являющиеся международными по производству и сбыту, трудно отнести к какой-либо стране или континенту. Так, «Форд» имеет отделения в Европе, Южной Америке и Азии. Японские и южнокорейские фирмы по производству бытовой электроники создают свои отделения

в Америке и Европе. Немецкие химические предприятия имеют филиалы в Южной Америке и Азии. Перемешивание капитала и производства затрудняет создание понятной и абсолютно «чистой», в смысле логики, модели. Однако опора на такие интегральные показатели, как ВВП и индекс жизни, могла бы дать достаточно адекватную модель.

Заметим, что субъекты в экономике не всегда только борются за общие ресурсы или потребителей. В каждом государстве существуют отрасли, взаимодействие которых является более сложным. Рассмотрим, например, сельское хозяйство, промышленность, добывающий сектор экономики и бюджетные средства государства. Они одновременно и конкурируют, и развивают друг друга. Так как в мире нет совершенно изолированных стран, то могут существовать государства с одним доминирующим сектором из вышеназванных. Скажем условно, что Бирма, Таиланд и Никарагуа — сельскохозяйственные страны, Кувейт и Венесуэла — добывающие, Англия и Япония — промышленные, Германия, Швеция и Люксембург — бюджетные. Тем не менее, в большинстве стран функционируют все из названных экономических секторов, и они, естественно, взаимодействуют. Правительства и парламенты стараются управлять указанным взаимодействием, например вводя специальные экспортно-импортные налоги или поддерживая сельское хозяйство специальными закупочными ценами. Далее в этой главе мы предложим соответствующую модель Лотки–Вольтерры.

Многие экономисты пытались создать математические модели циклов деловой активности. Некоторые из циклов имеют свои названия: длинные и короткие кондратьевские волны, сезонные циклы и т. д. Про другие говорят: «...имеется явно выраженная тенденция к периодическим колебаниям». Однако это явление, колебательность, до сих пор остается неформализованным. Отдельные попытки (как, например, обсуждаемая ниже модель Гудвина) следует признать неудачными. Типичное объяснение колебаний, основанное на внедрениях новых технологий, нам не кажется убедительным, хотя бы потому что в широком спек-

тре наук и с учетом случайности появления содержательного открытия или изобретения вряд ли может определить-ся периодичность или даже общая закономерность, кроме нарастающего суммирования всех достижений.

3.2.2. КОНКУРЕНЦИЯ ПРЕДПРИЯТИЙ НА ОБЩЕМ РЫНКЕ

Сначала рассмотрим динамическую модель одного предприятия. Пусть некоторая фирма, обладая основными фондами $K(t)$, привлекая рабочую силу $L(t)$, используя природные ресурсы (сырье, вода, энергия, земля и т. д.) $R(t)$, получает объем товара $x(t)$. Для задач со сложной структурой использования природных ресурсов можно условно считать, что $R(t)$ — это оборотный капитал. Аналогично, $K(t)$ — основной капитал. Объем товара $x(t)$ выражен в текущих ценах. Чтобы упростить рассмотрение экономических предположений на первом этапе, фиксируем текущие цены. Это значит, что рассматриваемая фирма своим объемом товара не влияет на рыночную равновесную цену, т. е. фирма не является монополистом. Обозначая производственную функцию фирмы φ , получим первое математическое выражение: $x(t) = \varphi(K(t), L(t), R(t))$. Мы не будем рассматривать подробно теорию производственных функций, скажем лишь, что φ и все ее первые производные являются строго положительными (дифференцируемость может быть и кусочной) [78]. Здесь и далее t обозначает непрерывное или дискретное время. Если время непрерывно, то скорость изменения любой переменной обозначаем \dot{x} или dx/dt ; если же время понимается как дискретное, то скорость изменения любой переменной по времени записываем как первую разность Δx .

Чтобы расширить производство, руководство фирмы должно расширить основные фонды и использовать больше трудовых и природных ресурсов. Примем как простейший вариант исследования, что основные фонды на свое поддержание требуют амортизационных отчислений, пропорциональных объему основных фондов, а на развитие инвестиций — $I(t)$. Далее, пусть используемые ресурсы

пропорциональны основным фондам, включенным в производство: $L(t) = lK(t)$, $R(t) = rK(t)$, и фирма не тратит прибыль ни на что другое, кроме развития (налоговую составляющую примем пропорциональной выводу и включенной в производственную функцию): $I(t) = x - lK - rK$. Добавляя к этим предположениям требование линейности функции φ , получим уравнение изменения основных фондов:

$$\dot{K}(t) = [-\beta + \varphi_K + (\varphi_L - 1)l + (\varphi_R - 1)r]K(t),$$

где β — коэффициент износа основных фондов, $\varphi_K = \partial\varphi/\partial K$, $\varphi_L = \partial\varphi/\partial L$, $\varphi_R = \partial\varphi/\partial R$. Последнее уравнение дает экспоненциальный рост основных фондов, если $\varepsilon = -\beta + \varphi_K + (\varphi_L - 1)l + (\varphi_R - 1)r > 0$. Соответственно, и выход продукции будет расти или убывать экспоненциально:

$$x(t) = x(0)\exp(\varepsilon t).$$

Таким образом, модель

$$\dot{x}(t) = \varepsilon x(t),$$

с постоянным коэффициентом роста ε , характеризует динамику одного предприятия при отсутствии каких-либо экономических ограничений. Однако ограничения всегда существуют, например:

- рост производства приводит к насыщению рынка и снижению спроса;
- со временем происходит моральное старение товара или услуги;
- расширение производства требует привлечения рабочей силы, но если рынок труда ограничен, то расширение достигает верхнего предела;
- природные ресурсы всегда ограничены либо по объему, либо по цене, когда спрос на ресурсы растет (это же касается и труда).

Чтобы перейти к модели Лотки–Вольтерры, достаточно в последнем уравнении коэффициент ε представить как убывающую линейную функцию растущего значения $x(t)$ (линейную — поскольку это простейшая убывающая функция): $\varepsilon = \hat{\varepsilon} - \gamma x(t)$, с постоянными $\hat{\varepsilon}$, γ . Так как логика этих действий понятна, проделаем это сразу в двухмерном случае.

Итак, первая фирма имеет производственную функцию $x = \varphi(K_x, L_x, R_x)$, а вторая — $y = \psi(K_y, L_y, R_y)$, где φ и ψ — однородные линейные функции своих аргументов. Предположим, что каждая фирма тратит прибыли только на инвестиции, а затраты на трудовые ресурсы и оборотный капитал пропорциональны соответствующим привлекаемым основным фондам. Коэффициенты пропорциональности вследствие предположения об ограниченности ресурсов (на общем рынке труда и сырья) будем считать линейными убывающими функциями K_x, K_y . Тогда изменения основных фондов обеих фирм складываются из износа и инвестиций:

$$\begin{cases} \dot{K}_x(t) = -\beta_x K_x + \varphi(K_x, L_x, R_x) - L_x - R_x, \\ \dot{K}_y(t) = -\beta_y K_y + \psi(K_y, L_y, R_y) - L_y - R_y. \end{cases}$$

Подставляя линейные выражения в качестве коэффициентов для труда и сырья

$L_x = (l_{0x} - l_{1x}K_x - l_{2x}K_y)K_x, \dots, R_y = (r_{0y} - r_{1y}K_x - r_{2y}K_y)K_y$,
получим систему относительно K_x, K_y

$$\begin{cases} \dot{K}_x(t) = [-\beta_x + \varphi_K + (\varphi_L - 1)l_{0x} + (\varphi_R - 1)r_{0x} - \\ - (\varphi_L l_{1x} + \varphi_R r_{1x})K_x - (\varphi_L l_{2x} + \varphi_R r_{2x})K_y]K_x, \\ \dot{K}_y(t) = [-\beta_y + \psi_K + (\psi_L - 1)l_{0y} + (\psi_R - 1)r_{0y} - \\ - (\psi_L l_{1y} + \psi_R r_{1y})K_x - (\psi_L l_{2y} + \psi_R r_{2y})K_y]K_y. \end{cases} \quad (97)$$

Чтобы увидеть систему (95), переобозначим коэффициенты:

$$\begin{aligned} \varepsilon_1 &= -\beta_x + \varphi_K + (\varphi_L - 1)l_{0x} + (\varphi_R - 1)r_{0x}, \\ \gamma_{11} &= \varphi_L l_{1x} + \varphi_R r_{1x}, \quad \gamma_{12} = \varphi_L l_{2x} + \varphi_R r_{2x}, \\ \varepsilon_2 &= -\beta_y + \psi_K + (\psi_L - 1)l_{0y} + (\psi_R - 1)r_{0y}, \\ \gamma_{21} &= \psi_L l_{1y} + \psi_R r_{1y}, \quad \gamma_{22} = \psi_L l_{2y} + \psi_R r_{2y}. \end{aligned}$$

Экономический смысл коэффициентов β понятен. Остановимся на интерпретации остальных параметров. Величины $\varphi_K, \varphi_L, \varphi_R, \psi_L, \psi_K, \psi_R$ характеризуют производственные функции обеих фирм, следовательно, мы можем считать их заданными. По предположению о производственных функциях эти величины положительны. Коэффициенты ε и γ должны быть идентифицированы по информации об объемах основных фондов K_x, K_y . Остались только коэф-

фициенты l и r , но их 12 на 6 приведенных выше уравнений. Поэтому при идентификации необходимо уделить этому внимание.

Ограниченность общих ресурсов приводит к тому, что с ростом основных фондов K_x , K_y коэффициенты возобновления капитала становятся равными нулю, а затем становятся отрицательными, т. е. фонды начинают убывать. Область в плоскости переменных K_x , K_y , когда основные фонды еще не выбывают из производства, аналитически выражается неравенствами:

$$\begin{aligned} & -\beta_x + \varphi_K + (\varphi_L - 1)l_{0x} + (\varphi_R - 1)r_{0x} - \\ & - (\varphi_L l_{1x} + \varphi_R r_{1x})K_x - (\varphi_L l_{2x} + \varphi_R r_{2x})K_y > 0, \\ & -\beta_y + \psi_K + (\psi_L - 1)l_{0y} + (\psi_R - 1)r_{0y} - \\ & - (\psi_L l_{1y} + \psi_R r_{1y})K_x - (\psi_L l_{2y} + \psi_R r_{2y})K_y > 0. \end{aligned}$$

При достаточно малых K_x , K_y эти неравенства выполняются, увеличение какого-нибудь из них приводит к уменьшению коэффициента пропорциональности между стоимостью ресурсов и привлекаемых основных фондов на единицу продукции. Так, например, если первое из указанных выше неравенств перестает выполняться при некоторых значениях K_x , K_y , то это значит, что первая из фирм получает продукцию вообще без привлечения рабочей силы, это мы должны исключить. Более подробно такие особенности модели необходимо обсуждать в каждом конкретном случае.

Теперь ясно, что взаимная конкуренция n фирм за общие ресурсы может быть описана системой (95), однако параметры модели ε и γ теперь уже не носят абстрактный характер, т. е. они представляют собой некоторые вполне определенные зависимости от экономически интерпретируемых параметров.

3.2.3. ДИНАМИКА ИЕРАРХИЧЕСКИ СВЯЗАННЫХ ОТРАСЛЕЙ

Взаимодействие отраслей очень похоже на взаимодействие предприятий: отрасли также обладают основными фондами, привлекают трудовые и природные ресурсы и, следовательно, конкурируют между собой. Однако здесь возможна и зависимость отраслей.

До сих пор мы предполагали, что коэффициенты ε в системе (95) положительны, т. е. при отсутствии конкуренции и ограниченных ресурсов все субъекты развиваются экспоненциально (неограниченно). Но существуют отрасли, чья зависимость от других субъектов экономики является гораздо более сложной. Так, например, бюджетные отрасли государства (медицина, вооруженные силы, образование, наука и т. д.) в большой степени не предполагают самокупаемость и страдают, когда бюджет страны уменьшается. А так как бюджет в основном состоит из налогов, и последние пропорциональны выходам других отраслей, то следует ожидать появления отрицательных коэффициентов ε .

Рассмотрим проблему дотирования сельского хозяйства. В экономической истории нашей страны были различные взаимоотношения между сельхозпроизводством и промышленностью: крепостное право сменилось патриархальным земледелием, основанном на коллективном владении землей, столыпинская попытка активизировать сельское хозяйство инициативой единоличников с частным землевладением закончилась революцией. Социалистическое правление также не привело к стабильности в сельхозпроизводстве: НЭП, раскулачивание, коллективизация, ценовые «ножницы», ограничительные меры по производству и основным фондам, отток рабочей силы из села, неподготовленное и не поддержанное экономически (и юридически) фермерство — вот неполный перечень ударов, нанесенных по сельскому хозяйству нашей страны в XX столетии. Интересно, что в 80-х гг. XIX века Д. И. Менделеев, рассматривая рентабельность производства хлеба в Российской Империи, подробно, с анализом статистических данных, показал, что в России с ее неустойчивым климатом невозможно иметь надежного производства зерна, что огромный экспорт ржи в Европу в конце XIX века был следствием жестокой эксплуатации крестьянства и недоедания российских граждан. Его точку зрения разделял и С. Ю. Витте, известный политический и экономический деятель тех лет.

Однако и в индустриально развитых странах, таких как США, Франция или Германия, производство сельскохозяйственной продукции является постоянным предметом

беспокойства правительства, которые вынуждены делать специальные государственные инвестиции, или проводить через парламенты налоговые льготы, или осуществлять бюджетные закупки по завышенным ценам. Таких способов помощи своим домашним аграриям история экономики знает множество — от создания протекционистских барьеров на пути дешевого импорта до государственных дотаций на покупку комбайнов и тракторов. Все эти способы в конечном итоге приводят к математической модели «хищник–жертва» в экономическом контексте того, что сельское производство «паразитирует» на промышленности. Модель (95), а точнее модель (97), будет иметь следующий вид:

$$\begin{cases} \dot{K}_x(t) = [-\varepsilon_x + \gamma_x K_y(t)]K_x(t), \\ \dot{K}_y(t) = [\varepsilon_y - \gamma_{yx} K_x(t) - \gamma_{yy} K_y(t)]K_y(t). \end{cases} \quad (98)$$

Здесь все коэффициенты положительны. Далее мы исследуем эту систему и обобщим ее на многомерный случай.

3.2.4. МЕЖДУНАРОДНАЯ КОНКУРЕНЦИЯ

Экономическое взаимодействие государств определяется множеством различных факторов экономического, политического, географического, исторического содержания. Не существует математических моделей, охватывающих все проявления взаимодействия. Как правило, анализу подвергаются отдельные его виды, например, политические групповые интересы государств исследуются как решение игровой задачи или вопрос о торговых потоках сводится к поиску точки экстремума функции многих переменных (функции дохода). Ввиду сложности и многозначности межгосударственных отношений модели в основном носят статический характер. Из всех видов взаимодействий между государствами остановимся только на торговых, игнорируя военные, политические и другие неторговые отношения. Они более изучены в различных работах, где делаются попытки, опираясь на закон сравнительных преимуществ Д. Рикардо, осуществить всесторонний анализ внешней торговли.

Почти за двести лет этот закон улучшался, изменялся и отвергался, но остался во всех учебниках по внешней торговле. Кроме того, к классической теории сравнительных преимуществ в международной торговле в настоящее время добавилась монетаристская теория, основанная на платежном балансе в экспортно-импортных операциях стран. Так что этот раздел экономической теории можно рассматривать как достаточно подготовленный для математического моделирования.

Здесь мы сформулируем несколько известных утверждений теории международной торговли по книге П. Х. Линдберта «Экономика мирохозяйственных связей» [55].

Закон Д. Рикардо гласит: каждая страна располагает сравнительным преимуществом в производстве какого-то товара и получает выигрыш, торгуя им в обмен на остальные. Следующее основополагающее утверждение носит название «теории Хекшера–Олина» (оба шведские экономисты, первый ученик второго): товары, требующие для своего производства значительных затрат (избыточных факторов производства) и небольших затрат (дефицитных факторов), экспортируются в обмен на товары, производимые с использованием факторов в обратной пропорции. Так, в скрытом виде экспортируются избыточные факторы и импортируются дефицитные факторы производства.

Этот вывод не всегда соответствует экономической статистике, поэтому американский экономист П. Самуэльсон ограничил применение последнего утверждения несколькими условиями:

- речь идет о двух странах, двух товарах и двух факторах;
- предложение факторов в каждой стране фиксировано и их перемещение допускается между секторами внутри страны, но не между странами;
- страны отличаются одна от другой лишь обеспеченностью факторами производства;
- в обеих странах технология такова, что обеспечивает неизменный эффект масштаба.

Кроме того, в распределительной части теории П. Самуэльсон вместе с В. Столпером установили, что при тех

же предпосылках установление торговых отношений и свободная торговля неизбежно ведут к росту вознаграждения фактора, интенсивно используемого в производстве товара, цена на который растет, и снижению вознаграждения фактора, интенсивно используемого в производстве товара, цена на который падает, вне зависимости от того, какова структура потребления этих товаров владельцами факторов производства. Из этого, в частности, следует, что с открытием внешней торговли должна возрастать классовая (социальная) напряженность в обществе.

Далее П. Самуэльсон развивает эту теорию в так называемой «теореме о выравнивании цен на факторы производства»: если к приведенным выше предпосылкам добавить следующие:

- все рынки являются конкурентными;
- в обеих странах каждый фактор используется с полной загрузкой как в условиях международной торговли, так и в ее отсутствие;
- транспортные и информационные издержки отсутствуют;
- торговля является свободной (бестарифной);
- производственные функции не обладают свойством «обратимости факторной интенсивности»;
- обе страны после установления торговых отношений продолжают производить оба товара;

тогда внешняя торговля ведет к выравниванию цен не только на товары, но и на факторы производства вне зависимости от структуры спроса или обеспеченности факторами в каждой стране.

Теория международной торговли высветила проблему мобильности факторов производства как с ее положительных сторон, так и с отрицательных. Например, перемещение фактора труда, т. е. миграция людей в поисках удовлетворяющей их работы совсем необязательна, так как экспорт трудоемкой продукции равносителен перемещению фактора труда. Но, с другой стороны, имеет место теорема Рыбчинского (в некотором смысле двойственная теореме Столпера–Самуэльсона): увеличивающееся предложение одного из факторов приводит к еще большему

процентному увеличению выпуска в том секторе, где этот фактор используется интенсивнее, в то время как в другом секторе наблюдается абсолютное сокращение темпов выпуска. Отсюда следует, что остальные факторы как бы притягиваются к производству с гипертрофированным (и удешевленным) фактором, оставляя те производства, где они доминировали.

Таким образом, открытая внешняя торговля, решая некоторые проблемы государства расширением рынков и базы производственных факторов, приобретает новые трудно формализуемые изменения в структуре производства и потребления. Однако и протекционизм обладает существенными недостатками общеэкономического плана. Хотя в той или иной мере правительства постоянно прибегают к различным формам протекционизма, в исторической перспективе защита отечественного производителя часто выливается в консервацию недостатков развития, созданию монополистических групп, установлению необоснованно высоких цен на товары и т. д.

3.2.5. КОНДРАТЬЕВСКИЕ ВОЛНЫ

В настоящее время в России происходит структурное преобразование экономики от плановых методов управления в условиях полной государственной собственности к рыночным способам производства, распределения и потребления основных благ. Естественно, что это изменение в системе производства, пропорциях потребления приводит к тому, что в этих условиях макроэкономические параметры не ведут себя стабильно, они резко изменяются, подчиняясь не только экономическим закономерностям, но и нестационарным политическим влияниям. Поэтому построение математических моделей для неустановившихся, переходных процессов и их использование в прогнозировании нуждается в специальных обоснованиях. Если же взглянуть на изменение параметров данной страны за длительный промежуток времени (сотни лет), то можно считать некоторые высокочастотные колебания интегрально незначимыми, и тогда выделяются три главные компоненты изменения:

- более или менее монотонный тренд;
- низкочастотные колебания с периодом, примерно равным 50 годам (длинные кондратьевские волны);
- высокочастотные колебания с периодом, примерно равным 12 годам (короткие кондратьевские волны).

Наш замечательный соотечественник Николай Дмитриевич Кондратьев своей работой 1928 г. «Динамика цен промышленных и сельскохозяйственных товаров» [46] дал мощный толчок изучению колебательных процессов макроэкономических параметров. Вслед за открытием длинных и коротких волн в колебаниях цен появились многочисленные работы, описывающие аналогичные явления в деловой активности, параметрах спроса и т. п. Эта тема является очень актуальной как для исследования причин динамики, так и для прогнозов экономического развития. Если в настоящее время в России трудно выделить определенно одну или несколько причин, оказывающих основное влияние на протекающие процессы, то в масштабе столетий вполне можно составить некоторые концепции и даже модели причинно-следственных связей в макроэкономике страны. Большинство исследователей склонно считать причиной больших кондратьевских волн технологические инновации. Остается только научиться выделять те из появившихся изобретений, которые более значимы для рассматриваемых макропараметров.

Поскольку эта проблема вряд ли имеет однозначное решение, то для описания колебаний привлекается аппарат диффузии или психологического конфликта. Поиск причин колебаний продолжается с все возрастающей силой: функционируют научные фонды, ассоциации, конференции, выпускается ряд журналов. По-видимому, в скором времени законы колебаний в экономике будут установлены. В этой книге мы хотели бы предложить один из способов описания колебаний, основанный на модели Лотки–Вольтерры с запаздыванием. Основной посылкой, приведшей к этой модели, является ограниченность экономического пространства (точнее, ограниченность факторов производства, которые используются данной экономикой), а также длительное время реакции экономики на измене-

ние среды. Сам Н. Д. Кондратьев тоже пишет об этом, хотя и не делает акцента на ограниченности ресурсов: «Можно говорить о равновесии (и о ценах равновесия) применительно к более длительному периоду, в течение которого меняется не только спрос, но и предложение: однако последнее меняется на основе в общем того же фонда основных капитальных благ (факторов производства). Можно, наконец, говорить о равновесии (и о ценах равновесия) применительно к еще более длительному периоду, в течение которого меняется не только спрос, не только предложение товаров, но и количество упомянутых основных капитальных благ. Эти блага (крупнейшие строительные сооружения, мелиорация, кадры квалифицированного труда и т. д.) обладают способностью длительного служения. Однако и создание их равным образом требует длительного времени, которое не укладывается в рамки обычного торгово-промышленного цикла» [46]. Последний абзац цитируемого фрагмента подталкивает к введению временного лага в математическую модель колебаний, так как без запаздывания модель Лотки–Вольтерры не допускает колебаний вокруг стационарной точки.

3.3. ПОСЛЕДЕЙСТВИЕ В ЭКОНОМИЧЕСКИХ МОДЕЛЯХ

В данном разделе приводятся некоторые замечания методического характера. Они возникли после ознакомления автора с большим количеством математических динамических моделей в макро- и микроэкономике, точнее говоря, в динамических моделях, использующих дифференциальные (и разностные) уравнения с временным лагом (или запаздыванием аргумента). Оказалось, что многие экономисты придают большое значение временному лагу в причинно-следственных связях динамических процессов и стараются отразить его в соответствующих моделях, однако сравнительно сложный математический аппарат дифференциальных уравнений с последствием не позволяет раскрыть все его богатство до конца или хотя бы до числового результата. В некоторых случаях игнорирование

запаздывания приводит к неадекватным результатам (далее этому посвящен первый пример). В других ситуациях делаются попытки добиться от модели каких-либо специальных свойств, вводя запаздывание (или даже опережение) аргумента, но при этом допускаются неадекватные преобразования. Иллюстрацией такого подхода автор считает широко известную модель Гудвина.

Завершит раздел описание зависимости решений от величины запаздывания, будет приведена форма периодического решения в скалярном случае.

3.3.1. МОДЕЛЬ ДИНАМИКИ ЦЕНЫ В ОКРЕСТНОСТИ РАВНОВЕСНОЙ ТОЧКИ

Хорошо известно [82], что в идеальном случае цена на некоторый товар p определяется равновесием спроса D и предложения S . На рисунке 9 это равновесие соответствует точке E на плоскости (q, p) , т. е. количеству товара и его цене.

Известно также, что положение равновесия будет притягивающим (асимптотически устойчивым по Ляпунову), если угол α будет больше угла β (имеются в виду соответствующие углы наклона касательных к кривым спроса и предложения). Это доказывается построением графика траектории изменения цены на рисунке 9 вокруг положения равновесия [62].

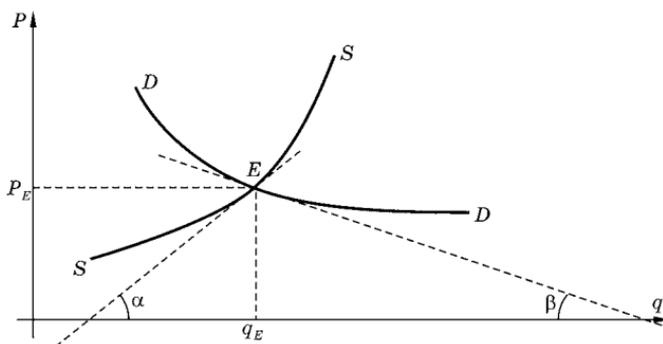


Рис. 9
Равновесная точка E в пересечении кривых спроса
и предложения

Если необходимо получить аналитическое выражение цены как функции времени, то обычно пользуются одним из уравнений:

$$\dot{p}(t) = D(p) - S(p), \quad \dot{p}(t) = p(t)[D(p) - S(p)],$$

где $D(p)$, $S(p)$ — функции, обратные соответствующим функциям спроса и предложения (на рис. 9 обращение достигается заменой оси цены на ось количеств, и наоборот). В данном случае безразлично, какое из уравнений используется, поэтому остановимся на первом. Для простоты будем считать, что спрос и предложение являются линейными функциями цены:

$$S(p) = \text{ctg}\alpha(p - p_E) + q_E, \quad D(p) = -\text{ctg}\beta(p - p_E) + q_E.$$

Здесь $\alpha, \beta \in [0, \pi/2]$, поэтому соответствующие котангенсы будут обязательно положительны. Уравнение для изменения цены примет вид:

$$\dot{p} = -(\text{ctg}\alpha + \text{ctg}\beta)p + (\text{ctg}\alpha + \text{ctg}\beta)p_E.$$

Очевидно, что положение равновесия p_E в последнем уравнении является асимптотически устойчивым по Ляпунову, для любых соотношений α и β , что противоречит упомянутому выше правилу. Если же учесть, что предложение приходит на рынок с некоторым опозданием h , связанным с технологическим циклом производства данного товара (этот факт отмечается в [62]), то получим адекватную модель

$$\dot{p}(t) = -\text{ctg}\beta p(t) - \text{ctg}\alpha p(t - h) + (\text{ctg}\alpha + \text{ctg}\beta)p_E.$$

Теперь то же самое положение равновесия p_E будет асимптотически устойчиво при произвольном запаздывании, когда $\text{ctg}\alpha < \text{ctg}\beta$, то есть когда $\alpha > \beta$.

Для того чтобы был понятен последний вывод и последующие рассуждения, уместно привести анализ простейшего уравнения с одним сосредоточенным запаздыванием [74] (также Приложение 1):

$$\dot{x}(t) = ax(t) + bx(t - h).$$

В плоскости параметров a и b область асимптотической устойчивости имеет вид, изображенный на рисунке 10.

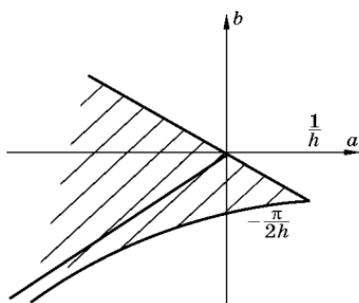


Рис. 10
Область устойчивости в пространстве параметров (a, b)

Так введение запаздывания в модель позволило устранить очевидное противоречие. Заметим, что иногда технологическое запаздывание отсутствует или очень мало. Такому случаю соответствует ситуация, когда данный товар в достаточных количествах имеется на складах оптовиков.

3.3.2. МАКРОЭКОНОМИЧЕСКИЕ МОДЕЛИ С ЗАПАЗДЫВАНИЕМ

Макроэкономические модели, связывающие валовой внутренний продукт, инвестиции, основные фонды и потребление, рассматривались П. Самуэльсоном, Дж. Хиксом М. Калецким, Р. Гудвином и другими авторами. Целью макроэкономического моделирования, как всегда, является некоторое прогнозирование, определение тенденций и попытки государственного управления экономикой страны. В данной работе рассматривается только одна из сторон всей динамики: циклические колебания, причем не возбужденные некоторой экзогенной составляющей для выбранной системы, а колебания, порождаемые внутренними причинами. Р. Гудвин составил математическую модель [107], делая упор на нелинейностях и введении запаздывания, чтобы полученные уравнения имели устойчивый предельный цикл. В работе [101] почти через 40 лет Д. Эрроусмит и К. Плейс практически процитировали процесс моделирования самого Р. Гудвина, сохранив всю логику и результат. В нашей литературе подобные модели также неоднократно появлялись, например [36], и с более понятной логикой предположений. Однако в большом обзоре [73] А. Поманского и Г. Трофимова опять приводится логика Р. Гудвина.

Вообще говоря, экономические предположения Р. Гудвина не могут считаться адекватными, и само построение

модели вызывает критические замечания, но здесь будет обращено внимание только на математическую часть, то есть на ошибочность математических преобразований.

Итак, пусть в любой момент времени t экономика страны располагает основным капиталом K , объемом производства (ВВП) Y с общим потреблением C . Эти величины связаны между собой соотношениями:

$$C_t = aY_t + b, \quad Y_t = C_t + \dot{K}_t, \quad (99)$$

где первое равенство означает, что потребление линейно зависит от объема производства, а второе — что все произведенное тратится либо на потребление, либо на поддержание основных фондов. Далее Р. Гудвин предлагает управлять инвестициями, чтобы основные фонды с течением времени колебались вокруг величины γY . В качестве управляющего сигнала он рассматривает импульсно-релейную функцию от Y . Естественно, объем производства и капитал получают как кусочно-линейные или кусочно-постоянные функции времени. Чтобы как-то сгладить скачки ВВП, Р. Гудвин не учитывает, что второе равенство в уравнении (99) означает распределение ВВП, и говорит, что это закон изменения Y по управлению \dot{K} , и вводит первое запаздывание ε :

$$Y(t + \varepsilon) = \frac{b + \dot{K}(t)}{1 - a}.$$

Так как это фактически нелинейное функциональное уравнение и решение его проблематично, то далее Y разлагается в ряд по ε и в уравнении сохраняется только член первого порядка

$$Y(t) + \varepsilon \dot{Y}(t) = (b + \dot{K}) / (1 - a).$$

Второй раз Р. Гудвин, стремясь сохранить колебательность, но в то же самое время добавляя постепенный рост экономики, полагает, что инвестиции (управление) зависят от $\dot{Y}(t - \theta)$, причем нелинейно и гладко. Если быть последовательным, то в математической форме получится уравнение опережающего типа, а их теория еще далека от совершенства. Но снова разлагая соответствующие функции в ряд Тейлора по степеням второго запаздывания,

можно получить дифференциальное уравнение второго порядка с гладкой нелинейностью относительно скорости

$$\varepsilon \theta \ddot{Y}(t) + [\varepsilon + (1-a)\theta] \dot{Y}(t) + (1-a)Y(t) = b + \dot{K}(\dot{Y}(t)).$$

Вот оно-то и имеет устойчивый предельный цикл.

Таким образом, если бы преобразования были верны, то из экономических предпосылок следовало бы существование устойчивых колебаний вокруг магистрального изменения макропараметров экономики. Но так ли верны преобразования?

Ниже приводится пример того, как дважды проделанная Р. Гудвином операция изменяет качественные свойства модели.

Из рисунка 10 видно, что уравнение

$$\dot{x}(t) = -ax(t-h)$$

имеет асимптотически устойчивое нулевое решение при $0 < ah < \pi/2$; допускает периодическое решение $x(t) = A \sin(\pi t/(2h)) + B \cos(\pi t/(2h))$ при $ah = \pi/2$ (с любыми A, B) и неустойчиво во всех остальных случаях.

Разложение правой части последнего уравнения в ряд по степеням h приводит к обыкновенным дифференциальным уравнениям такого порядка, сколько членов разложения оставлено в уравнении. В первом приближении уравнение имеет вид:

$$(1-ah)\dot{x}(t) = -ax(t).$$

При $0 < ah < 1$ имеет место асимптотическая устойчивость, при $ah > 1$ нулевое решение будет неустойчиво, периодических решений вообще нет. Во втором приближении исходное уравнение превращается в

$$ah^2 \ddot{x}(t) + 2(1-ah)\dot{x}(t) + 2ax(t) = 0,$$

которое ведет себя так же, как уравнение первого приближения, но при $ah = 1$ допускает периодическое решение с частотой $\sqrt{2}/h$. Приближение третьего порядка всегда дает неустойчивость. Таким образом, качественная картина в поведении решений точной модели и приближенной существенно зависит от величины запаздывания. Если же в мо-

делировании используется уравнение опережающего типа, то там стационарное решение всегда неустойчиво.

Таким образом, игнорирование временных лагов в прогнозных моделях может приводить к неадекватным выводам, но также и поверхностное манипулирование аппаратом уравнений с запаздыванием аргумента снижает ценность математического моделирования.

Выводы о существовании единственного устойчивого предельного цикла нельзя считать исходящими из экономических предпосылок, так как последние были сильно искажены последующими грубыми преобразованиями. Использование математического аппарата с недостаточно развитой теорией может привести к неадекватным предсказаниям, хотя это является очень хорошим стимулом для математиков в поисках новых инструментальных средств.

3.3.3. ИЗМЕНЕНИЕ КАЧЕСТВЕННОГО ПОВЕДЕНИЯ МОДЕЛИ ПРИ УВЕЛИЧЕНИИ ЗАПАЗДЫВАНИЯ

Как было только что показано, для линейных систем характерно ухудшение устойчивости при увеличении запаздывания (Приложение 1), и так как для систем с последствием в общем случае имеют место теоремы об устойчивости по первому приближению (Приложение 1), то локально это всегда будет верно. Но нелинейные системы в удалении от стационарных точек могут проявить другие свойства. Рассмотрим на скалярном примере модели Лотки–Вольтерры, как изменяется качественное поведение решения при увеличении запаздывания.

Сначала преобразуем скалярное логистическое уравнение, которое явилось прототипом рассматриваемой модели, к самому простому виду. А именно, уравнение

$$\dot{x}(t) = x(t)[a - bx(t - h)],$$

где $a, b > 0$, заменой времени $t = hs$ и заменой масштаба измерения $x = y/b$ приведем к виду:

$$\dot{y}(s) = hy(s)[a - y(s - 1)].$$

Или в интегральной форме, которая более удобна для вычислительных процедур, последнее уравнение принимает вид:

$$y(s) = y(0) \exp \left[h \int_0^s (a - y(r-1)) dr \right]. \quad (100)$$

Мы можем положить $a = 1$ без потери общности, и тогда справедливы следующие заключения [8]:

- $y = 1$ — положение равновесия уравнения (100);
- если решение дольше чем единицу времени находится ниже положения равновесия, то y возрастает;
- если решение дольше чем единицу времени находится выше 1, то y убывает;
- так как за единицу времени решение не может вырасти больше чем в e^h раз, то величина $y(0)e^h$ и ограничивает решение сверху;
- локально, в некоторой окрестности 1, решение, как уже упоминалось ранее, асимптотически стремится к положению равновесия при $h < \pi/2$, если же $h > \pi/2$, то $y = 1$ будет неустойчиво по Ляпунову.

Вычислительные процедуры показывают, что при малых h асимптотическая устойчивость будет глобальной в положительной области начальных данных, а при $h > \pi/2$ будут возникать периодические колебания вокруг нетривиального положения равновесия, которые будут к тому же глобально притягивающими.

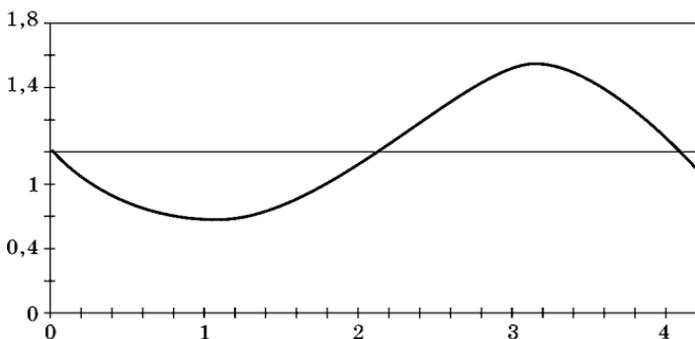


Рис. 11

Решение скалярного логистического уравнения с запаздыванием, немного большим, чем критическое

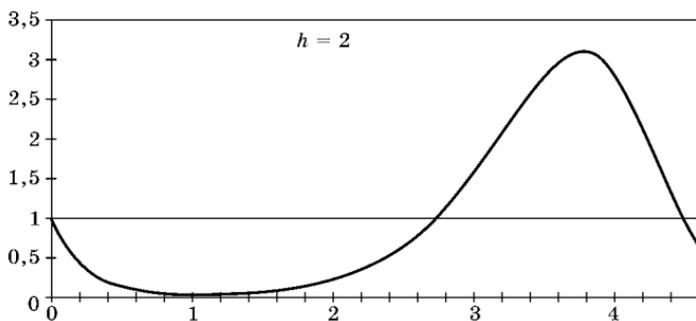


Рис. 12
Запаздывание существенно больше критического

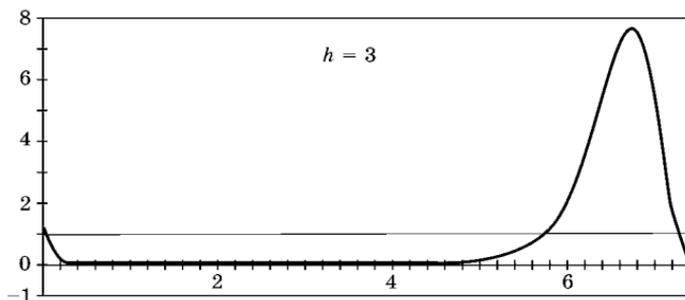


Рис. 13
Запаздывание вдвое превосходит критическое

Вид этих колебаний при возрастании запаздывания h изображен на рисунках ниже. Рисунок 11 соответствует $h = 1,6$ и похож на синусоидальное колебание, но в нем уже заметна тенденция, которая на рисунках 12 и 13 проявляется со всей очевидностью.

Нижняя часть колебания вытягивается горизонтально, а верхняя — вертикально.

При этом площади отрицательной и положительной частей по отношению к положению равновесия остаются равными. Действительно, раз мы говорим о периодических решениях, то к равенству (100) добавим равенство $y(s + T) = y(s)$ для любого s и периода T . Тогда из (100) следует:

$$y(T) = y(0) \exp \left[h \int_{-1}^{T-1} (1 - y(r)) dr \right]$$

или после преобразований

$$1 = \frac{1}{T} \int_0^T y(r) dr,$$

что и означает равенство соответствующих частей. Также с помощью вычислений на компьютере можно получить зависимость максимальной амплитуды и периода от величины запаздывания h . На рисунке 14 приведена такая зависимость.

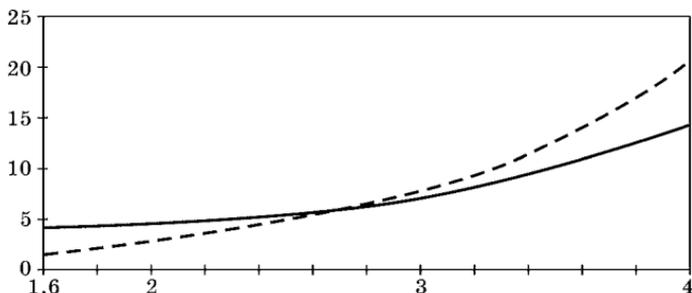


Рис. 14

Зависимость периода и амплитуды стабильного колебания при возрастании запаздывания

Здесь ось абсцисс соответствует величине запаздывания, сплошная линия — график периода и пунктирная линия — график амплитуды.

3.4. ПОЛНЫЙ АНАЛИЗ ВЗАИМОДЕЙСТВИЯ ЭКОНОМИЧЕСКИХ АГЕНТОВ

В данном разделе мы рассмотрим ситуации без запаздывания и с запаздыванием, двухмерный и многомерный случаи. Начнем с обсуждения возможных положений равновесия и вариантов поведения экономических агентов. Выяснив разные, с математической точки зрения, ситуации, приступим к их анализу, двигаясь от простого к сложному. Теоретические (т. е. абсолютные) факты будем чередовать с частными (но наглядными), полученными вычислительным путем, иллюстрациями. Экономическую интерпретацию результатов сделаем в заключении каждого подраздела.

**3.4.1.
СТАЦИОНАРНЫЕ ТОЧКИ МОДЕЛИ**

Итак, мы рассматриваем систему дифференциальных уравнений

$$\dot{x}_i(t) = x_i(t) \left[\varepsilon_i - \sum_{k=1}^n \gamma_{ik} x_k(t - \tau) \right], \quad i = 1, \dots, n, \quad (101)$$

где x_i характеризует одну из n конкурирующих фирм. Это может быть основной капитал или выход продукта. Величину последействия будем специально оговаривать, когда в этом будет необходимость.

По определению стационарные точки системы (101) могут быть вычислены как решение нелинейных уравнений

$$0 = x_i \left[\varepsilon_i - \sum_{k=1}^n \gamma_{ik} x_k \right], \quad i = 1, \dots, n,$$

где коэффициенты ε, γ являются положительными или отрицательными, в зависимости от постановки задачи. Все множество стационарных точек, как известно из алгебры [19], состоит из нескольких нулевых значений и соответствующих решений оставшейся линейной системы:

$$0 = \left[\varepsilon_i - \sum_{k=1}^n \gamma_{ik} x_k \right], \quad i = 1, \dots, n,$$

здесь мы оставили квадратные скобки, чтобы подчеркнуть, что это не вся система, а только ее «нетривиальная» часть. Например, для системы (101) третьего порядка множество стационарных точек запишем следующим образом:

$$\left\{ \begin{array}{l} (0, 0, 0); \left(0, 0, \frac{\varepsilon_3}{\gamma_{33}} \right); \left(0, \frac{\varepsilon_2}{\gamma_{22}}, 0 \right); \left(\frac{\varepsilon_1}{\gamma_{11}}, 0, 0 \right); \\ \left(0, \left[0 = \varepsilon_i - \sum_{j=2,3} \gamma_{ij} x_j \right]_{i=2,3} \right); \left(x_2 = 0, \left[0 = \varepsilon_i - \sum_{j=1,3} \gamma_{ij} x_j \right]_{i=1,3} \right); \\ \left(x_3 = 0, \left[0 = \varepsilon_i - \sum_{j=1,2} \gamma_{ij} x_j \right]_{i=1,2} \right); \left(\left[0 = \varepsilon_i - \sum_{j=1,2,3} \gamma_{ij} x_j \right]_{i=1,2,3} \right) \end{array} \right\}.$$

Если соответствующая линейная система не имеет решений, то рассматриваемое положение равновесия не существует; если же у системы имеется бесконечное множество решений, то стационарные точки составят линейное многообразие. Пусть, например, пятая точка указанного множества задается системой

$$\begin{cases} x_1 = 0, \\ 0 = 1 - 0,5x_2 - 0,3x_3, \\ 0 = 2 - 0,9x_2 - 0,6x_3. \end{cases}$$

Очевидно, у нее есть единственное решение $(0, 0, 3,33)$. Несколько измененная система

$$\begin{cases} x_1 = 0, \\ 0 = 1 - 0,5x_2 - 0,3x_3, \\ 0 = 2 - x_2 - 0,6x_3 \end{cases}$$

имеет целую прямую решений: $\{0, 2 - 0,6x_3, x_3\}$.

И также немного измененная система

$$\begin{cases} x_1 = 0, \\ 0 = 1 - 0,5x_2 - 0,3x_3, \\ 0 = 1,9 - x_2 - 0,6x_3 \end{cases}$$

не имеет решений.

Отметим очень важный факт о системе (101): *при любом запаздывании τ любая координатная плоскость и любое пересечение координатных плоскостей инвариантны относительно системы (101)*. Это легко увидеть, если переписать систему (101) в интегральной форме:

$$x_i(s) = x_i(0) \exp \int_0^s \left[\varepsilon_i - \sum_{k=1}^n \gamma_{ik} x_k(t-\tau) \right] dt, \quad i=1, \dots, n. \quad (102)$$

Очевидно, что для любых начальных функций (Приложение 1) знак решения определяется значением $x_i(0)$. Таким образом, если решение начинается (в момент времени $s=0$) в положительном ортанте пространства переменных, то оно там и остается дальше. При бесконечном времени существования решения оно может примыкать к той или иной грани положительного ортанта. Это свойство модели очень хорошо подходит для экономических ин-

терпретаций, так как выход фирмы или основные фонды не должны принимать отрицательных значений. Если какое-либо x_i стремится к нулю с течением времени, то это значит, что соответствующая фирма (или отрасль) прекращает свою деятельность.

Из вышеизложенного следует, что нас интересуют только стационарные точки, лежащие внутри и на границах положительного ортанта пространства переменных системы (101). Более того, если стационарная точка принадлежит координатной плоскости, то важны для заключительных выводов только асимптотически устойчивые точки. Исследование поведения решений на такой координатной плоскости, как на инвариантном множестве, сводится к тем же рассуждениям, что и во всем пространстве.

Таким образом, в дальнейшем нас интересуют только нетривиальные положительные стационарные точки (и существование притягивающих точек на координатных плоскостях).

Обозначим $E = \{\varepsilon_i\}_{i=1, \dots, n}$, $\Gamma = \{\gamma_{ij}\}_{i, j=1, \dots, n}$, тогда линейная алгебраическая система, дающая нетривиальное положительное решение, записывается в виде:

$$\Gamma X = E, \quad (103)$$

где X — вектор искомой координатной точки. О существовании положительных решений системы (103) может быть сформулирована следующая теорема [50]: справедлива альтернатива: либо система (103) имеет положительное решение, либо имеет решение система неравенств:

$$\Gamma^T Y \geq 0, \quad (E, Y) < 0, \quad (104)$$

где Y — некоторый n -мерный вектор, $(*, *)$ значит скалярное произведение.

Геометрически свойство положительности решений равносильно тому, что вектор E принадлежит выпуклой конической оболочке столбцов матрицы Γ .

Этой теоремой трудно пользоваться, поэтому для всякого частного примера мы будем формулировать условия существования положительного решения системы (103) специально.

3.4.2. КОНКУРЕНТНАЯ ДЕЯТЕЛЬНОСТЬ БЕЗ ЗАПАЗДЫВАНИЯ

Рассмотрим сначала случай деятельности двух фирм, конкурирующих за общие ресурсы. Как следует из вышеизложенного, динамика деятельности фирм без запаздывания ($\tau = 0$) описывается уравнениями:

$$\begin{aligned}\dot{x}_1(t) &= x_1(t)[\varepsilon_1 - \gamma_{11}x_1(t) - \gamma_{12}x_2(t)], \\ \dot{x}_2(t) &= x_2(t)[\varepsilon_2 - \gamma_{21}x_1(t) - \gamma_{22}x_2(t)],\end{aligned}$$

где шесть коэффициентов ε, γ обязательно положительны. Заменой масштаба измерения величин

$$y_1 = (\gamma_{11}/\varepsilon_1)x_1, \quad y_2 = (\gamma_{12}/\varepsilon_1)x_2,$$

приводим систему к виду, содержащему только два параметра, существенно влияющих на поведение решений:

$$\begin{aligned}\dot{y}_1(t) &= \varepsilon_1 y_1(t)[1 - y_1(t) - y_2(t)], \\ \dot{y}_2(t) &= \varepsilon_2 y_2(t)[1 - \alpha y_1(t) - \beta y_2(t)],\end{aligned}\tag{105}$$

здесь параметры α, β положительны и выражаются через исходные коэффициенты по формулам

$$\alpha = (\gamma_{21}\varepsilon_1)/(\gamma_{11}\varepsilon_2), \quad \beta = (\gamma_{22}\varepsilon_1)/(\gamma_{12}\varepsilon_2).$$

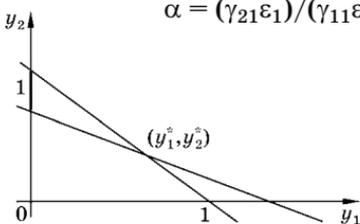


Рис. 15
Нетривиальное положение
равновесия системы (105)

Непосредственное вычисление нетривиального положения равновесия системы (105) дает результат:

$$\begin{aligned}y_1^* &= (\beta - 1)/(\beta - \alpha), \\ y_2^* &= (1 - \alpha)/(\beta - \alpha).\end{aligned}$$

На рисунке 15 эта точка получена при пересечении прямых линий, уравнения которых могут быть взяты в квадратных скобках системы (105).

Из последних вычислений следует вывод 1.

Вывод 1. Нетривиальная положительная стационарная точка системы (105) существует и единственна, если

$$\alpha < 1, \quad \beta > 1\tag{106}$$

или

$$\alpha > 1, \beta < 1. \tag{107}$$

Если $\alpha = \beta = 1$, то стационарные решения (y_1^*, y_2^*) составляют прямую линию с уравнением $y_1^* + y_2^* = 1$; и условие $\alpha = \beta \neq 1$ приводит к отсутствию нетривиальных стационарных точек вообще.

Обсуждаемое положение равновесия (y_1^*, y_2^*) может быть притягивающим (асимптотически устойчивым) или отталкивающим (неустойчивым). Для определения свойства устойчивости используем функцию Ляпунова:

$$V(y_1, y_2) = \frac{1}{\varepsilon_1} \left(y_1 - y_1^* - y_1^* \ln \frac{y_1}{y_1^*} \right) + \frac{1}{\varepsilon_2} \left(y_2 - y_2^* - y_2^* \ln \frac{y_2}{y_2^*} \right). \tag{108}$$

Так как мы рассматриваем только положительные y_1, y_2 , то V определена, непрерывна, $V(y_1^*, y_2^*) = 0$. При всех остальных положительных y_1, y_2 функция V строго положительна. Значит, она является определенно положительной для положения равновесия (y_1^*, y_2^*) .

Поскольку функции такого типа редко используются исследователями, покажем на рисунке 16 примерный график в одномерном случае.

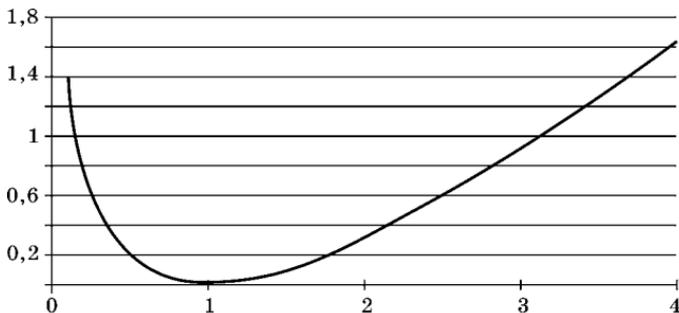


Рис. 16

Примерное изображение функции Ляпунова (108) в одномерном случае $V(y_1) = y_1 - 1 - \ln y_1$

В двумерном пространстве изобразим поверхности уровня функции (см. рис. 17).

Ее производная по времени в силу системы (108) имеет вид:

$$\dot{V}|_{(105)} = -(y_1 - y_1^*)^2 - (1 + \alpha)(y_1 - y_1^*)(y_2 - y_2^*) - \beta(y_2 - y_2^*)^2.$$

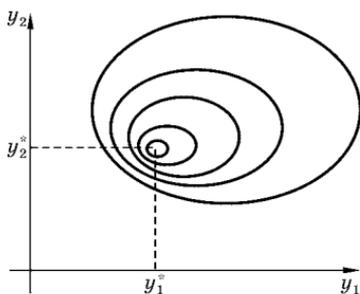


Рис. 17
Примерное изображение
поверхностей уровня
функции Ляпунова (108)
в двухмерном случае

Она определенно отрицательна, когда $4\beta > (1 + \alpha)^2$. Последнее неравенство выполняется, когда верно (106), и не выполняется при (107). Таким образом, справедлив следующий вывод.

Вывод 2. Если верно условие (106), то для любых положительных начальных данных решение с течением времени асимптотически стремится к (y_1^*, y_2^*) . Если же

имеет место условие (107), то положение равновесия неустойчиво (точнее говоря, представляет собой особую точку типа «седло»). В этом случае на плоскости переменных y_1, y_2 притягивающими будут точки координатных осей $(0, 1/\beta), (1, 0)$. Последнее утверждение вывода 2 легко установить, выписав систему в отклонениях относительно указанных точек. Линейные части этих систем имеют вид:

$$\begin{cases} \dot{y}'_1 \approx \varepsilon_1(1 - 1/\beta)y'_1, \\ \dot{y}'_2 \approx \varepsilon_2 1/\beta(-\alpha y'_1 - \beta y'_2); \end{cases} \quad \begin{cases} \dot{y}'_1 \approx -\varepsilon_1(y'_1 + y'_2), \\ \dot{y}'_2 \approx \varepsilon_2(1 - \alpha)y'_2; \end{cases}$$

где штрих показывает отклонение переменной от соответствующего равновесного значения. Как известно, через точку (y_1^*, y_2^*) в этом случае проходит сепаратриса, делящая положительный квадрант плоскости на две части: в одной части собраны все траектории, стремящиеся к одной координатной оси, а в другой — ко второй. Движение по самой сепаратрисе осуществляется к точке (y_1^*, y_2^*) . Таким образом приходим к выводу 3.

Вывод 3. Траектория системы (105) с любыми неотрицательными начальными данными и любыми коэффициентами с течением времени асимптотически приблизится либо к нетривиальному положению равновесия, либо к одной из равновесных точек на координатных осях.

На следующих графиках (рис. 18, 19) изображено поле направлений в случае устойчивости и неустойчивости нетривиального положения равновесия системы (105).

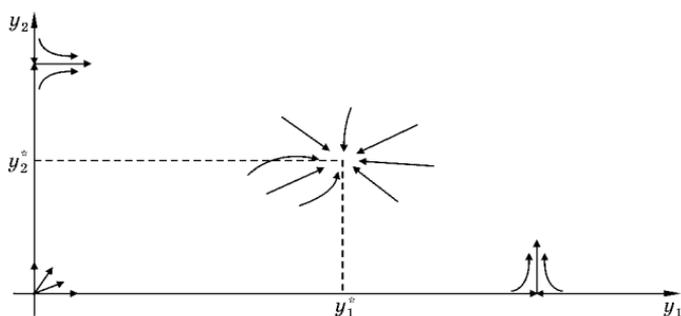


Рис. 18

Нетривиальное положение равновесия является притягивающей точкой

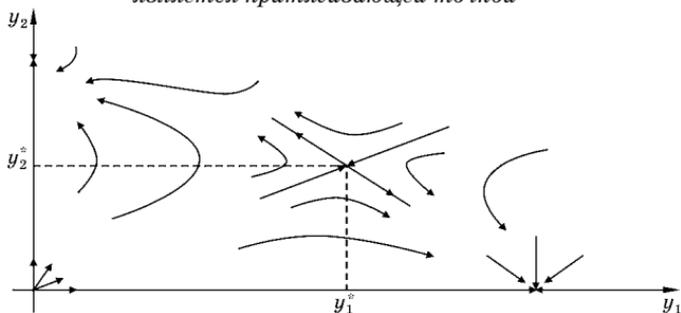


Рис. 19

Нетривиальное положение равновесия является неустойчивым, типа «седло»

Экономическая интерпретация. При конкуренции двух экономических агентов за общие ресурсы с течением времени возможны следующие результаты:

- при $\gamma_{12}/\gamma_{22} < \varepsilon_1/\varepsilon_2 < \gamma_{11}/\gamma_{21}$ объемы производства обеих фирм стремятся к величинам $(\gamma_{22}\varepsilon_1 - \gamma_{12}\varepsilon_2)/(\gamma_{22}\gamma_{11} - \gamma_{12}\gamma_{21})$ и $(\gamma_{11}\varepsilon_2 - \gamma_{21}\varepsilon_1)/(\gamma_{22}\gamma_{11} - \gamma_{12}\gamma_{21})$ соответственно. Их начальное состояние не играет никакой роли, объемы производства могут сокращаться или возрастать, изменяется также их суммарный объем;
- при $\gamma_{11}/\gamma_{21} < \varepsilon_1/\varepsilon_2 < \gamma_{12}/\gamma_{22}$ одна из фирм (за исключением крайне редкого случая сепаратрисы) с течением времени прекратит производство, другая же либо сократит, либо увеличит выпуск в соответствии с наличным ресурсом. Окончательный выпуск для первой фирмы численно равен $\varepsilon_1/\gamma_{11}$, а для второй — $\varepsilon_2/\gamma_{22}$. Этот случай

открывает возможность для руководителей фирм управлять процессом конкуренции. Если изменить начальное состояние производства так, чтобы оказаться в предпочтительной части плоскости, то появится шанс выжить. Достичь этого можно, например, взяв большой кредит и резко расширив производство. Однако расширение не должно превосходить равновесного значения;

- при иных соотношениях коэффициентов модели следует отказаться от ее использования, так как система не будет иметь положительного равновесия.

Рассмотрим случай произвольного числа экономических агентов, конкурирующих за общие ресурсы. Нетривиальное положительное решение системы (103), как уже отмечалось, будет существовать, если положительный вектор E принадлежит выпуклой конической оболочке, натянутой на столбцы матрицы Γ . Это не очень конструктивный критерий. Как правило, в таких случаях легче решить систему (103), чем проверять общие критерии. Предположим, что по имеющимся коэффициентам ε, γ решение системы (103) X^* получено. Вопрос об устойчивости положения равновесия в точке X^* решается с помощью определенной положительной функции Ляпунова

$$V(X - X^*) = \sum_{i=1}^n \left(x_i - x_i^* - x_i^* \ln \frac{x_i}{x_i^*} \right). \quad (109)$$

Заметим, что эта функция определена только для строго положительных x_i и обращается в нуль только в точке равновесия X^* . Ее производная по времени в силу системы (101) записывается проще, если преобразовать систему (101), используя решение алгебраической системы (103)

$$\dot{x}_i(t) = -x_i(t) \left[\sum_{k=1}^n \gamma_{ik} (x_k(t) - x_k^*) \right], \quad i = 1, \dots, n. \quad (101^*)$$

Итак,

$$\dot{V}|_{(X^*)} = - \sum_{i=1}^n \sum_{k=1}^n (x_i - x_i^*) \gamma_{ik} (x_k - x_k^*).$$

Очевидно, это определено отрицательная квадратичная форма, если матрица $^{1/2}(\Gamma^T + \Gamma)$ является определено положительной.

Вывод 4. Нетривиальное положение равновесия системы (101) X^* будет асимптотически устойчивым, если матрица с положительными элементами $\frac{1}{2}(\Gamma^T + \Gamma)$ будет определено положительной, при этом траектории системы (101) с любыми строго положительными начальными данными неограниченно с течением времени приближаются к X^* .

Если среди начальных данных имеется нулевое, то есть траектория начинается с какой-либо координатной плоскости, то в силу инвариантности она будет всегда оставаться на этой координатной плоскости и соответствующее положение равновесия будет оставаться условно асимптотически устойчивым. Относительно же всего пространства координатные положения равновесия (т. е. имеющие нулевые координаты) будут неустойчивыми.

Ситуация, когда матрица $\frac{1}{2}(\Gamma^T + \Gamma)$ не является определено положительной, не описывается так просто в общем случае, так как условия существования нетривиального положительного равновесного состояния на координатных плоскостях являются более узкими, чем в общем случае. Поэтому выводы о притяжении координатных стационарных точек необходимо всякий раз перепроверять.

Так, например, проанализируем 3-мерную модель с двумя свободными параметрами:

$$\begin{cases} \dot{x}_1 = x_1(1 - x_1 - x_3), \\ \dot{x}_2 = x_2(\varepsilon_2 - x_2 - 0,5x_3), \\ \dot{x}_3 = x_3(\varepsilon_3 - x_1 - 0,5x_2 - x_3), \end{cases}$$

где $\varepsilon_2, \varepsilon_3 > 0$. Система имеет положительную нетривиальную стационарную точку, если $\varepsilon_2 + 1,5 \leq 2\varepsilon_3 \leq \varepsilon_2 + 2$, которая неустойчива, так как матрица

$$\frac{1}{2}(\Gamma^T + \Gamma) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0,5 \\ 1 & 0,5 & 1 \end{pmatrix}$$

не является определено положительной. Таким образом, общее нетривиальное равновесие неустойчиво. Рассмотрим координатную грань $x_3 = 0$: на ней нетривиальное положение равновесия при уже принятых ограничениях всегда

существует и положительно: $(1, \varepsilon_2, 0)$. Составив систему в отклонениях относительно последнего положения равновесия, легко видеть, что ее линейная часть имеет отрицательные характеристические числа, если $2\varepsilon_3 < \varepsilon_2 + 2$, т. е. эта стационарная точка всегда будет притягивающей.

На координатной плоскости $x_2 = 0$ стационарные точки существуют только при $\varepsilon_3 = 1$. Это влечет строгую отрицательность ε_2 , что недопустимо. На координатной плоскости $x_1 = 0$ положительные стационарные точки могут существовать (при $\varepsilon_2 \leq 2\varepsilon_3 \leq 4\varepsilon_2$), и тогда они притягивающие, а могут и не существовать.

Из асимптотической устойчивости положения равновесия следует ограниченность и продолжимость на бесконечный интервал времени решений системы (101) при $\tau = 0$. В случае неустойчивости такой вывод не очевиден. Поэтому рассмотрим систему в интегральной форме (105). При условии положительности ε, γ, x_i будут верны неравенства:

$$x_i(s) = x_i(0) \exp \int_0^s \left[\varepsilon_i - \sum_{k=1}^n \gamma_{ik} x_k(t) \right] dt < x_i(0) e^{\varepsilon_i s}, \quad i = 1, \dots, n,$$

а это влечет продолжимость решений. Но можно показать и ограниченность всех решений при всех положительных коэффициентах и начальных данных (отсюда будет следовать и продолжимость).

Предположим, что какая-либо переменная $x_i(s)$ от положительного начального значения неограниченно возрастает. Тогда найдутся два достаточно близких момента времени $\hat{s} < \tilde{s}$, что $x_i(\hat{s}) < x_i(\tilde{s})$ и

$$\varepsilon_i - \gamma_{ii} x_i(\hat{s}) - \sum_{k=1, k \neq i}^n \gamma_{ik} x_k(\hat{s}) < 0,$$

$$\varepsilon_i - \gamma_{ii} x_i(\tilde{s}) - \sum_{k=1, k \neq i}^n \gamma_{ik} x_k(\tilde{s}) < 0,$$

следовательно, учитывая гладкость решений, получим противоречие: с одной стороны, монотонное возрастание, т. е. $\dot{x}_i(\hat{s}) > 0$, с другой — правые части системы (101) при этих значениях переменных отрицательны.

Следует отметить, что за пределами положительного ортанта пространства переменных система (101) имеет непродолжимые решения, т. е. непрерывные и становящиеся бесконечно большими при конечных значениях времени. В этом контексте очень важным является анализ правых частей при различных обобщениях систем Лотки–Вольтерры. Так, например, обобщенная логистическая модель А. Н. Колмогорова для всех вариантов f_i требует проверки продолжимости.

В случае неустойчивости нетривиального положения равновесия (и в то же время ограниченности решений в положительном ортанте) интересно было бы исследовать вопрос о колебательности решений. В случае $n = 2$ сепаратриса неустойчивого равновесного состояния проходит через начало координат и делит положительный квадрант на две части, в каждой из которых есть притягивающие точки на координатных осях и, следовательно, колебательности нет. В пространстве размерности более двух аналогичное утверждение неверно.

3.4.3. КОНКУРЕНТНАЯ ДЕЯТЕЛЬНОСТЬ С ВРЕМЕННЫМ ЛАГОМ

Вернемся к уравнениям (105), предполагая, что запаздывание τ строго положительно:

$$\begin{aligned} \dot{y}_1(t) &= \varepsilon_1 y_1(t)[1 - y_1(t - \tau) - y_2(t - \tau)], \\ \dot{y}_2(t) &= \varepsilon_2 y_2(t)[1 - \alpha y_1(t - \tau) - \beta y_2(t - \tau)]. \end{aligned} \quad (110)$$

И нетривиальное положение равновесия также останется прежним (y_1^*, y_2^*) . При малом положительном запаздывании все выводы о качественном поведении решений системы (110) продолжают иметь место, но с увеличением запаздывания картина меняется: неустойчивые стационарные точки сохраняют неустойчивость, а устойчивые — теряют устойчивость. В последнем случае вокруг нетривиального равновесного центра (y_1^*, y_2^*) возникает притягивающий предельный цикл (или иначе — стабильное колебание). Примерное поведение траекторий изображено на рисунке 20.

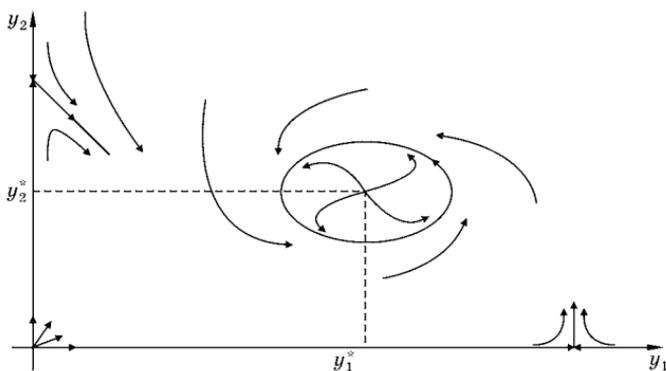


Рис. 20

Возникновение притягивающего предельного цикла при возрастании запаздывания

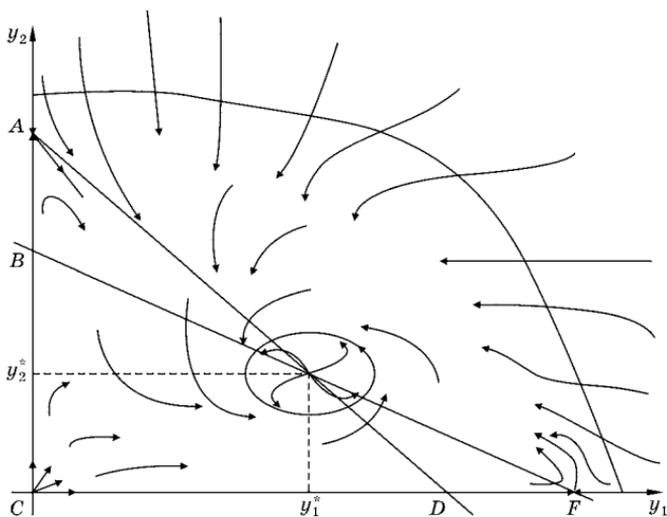


Рис. 21

Примерный портрет фазовой плоскости (y_1, y_2) для системы (110) при условии (106)

Обсудим эти изменения свойств траекторий с увеличением запаздывания более подробно.

Утверждение. Если выполнено условие (106), то для любого фиксированного запаздывания τ существует область в положительном квадранте фазовой плоскости, содержащая треугольники ABE и DEF и четырехугольник

$CBED$ (рис. 21), и такая, что все траектории изнутри этой области в ней и остаются, и любая траектория снаружи ее входит внутрь.

Для доказательства заметим, что из (110) следуют неравенства:

- если точка $(y_1(t - \tau), y_2(t - \tau)) \in \triangle ABE$, то $\dot{y}_1(t) > 0, \dot{y}_2(t) < 0$;
- если точка $(y_1(t - \tau), y_2(t - \tau)) \in \triangle DEF$, то $\dot{y}_1(t) < 0, \dot{y}_2(t) > 0$;
- если точка $(y_1(t - \tau), y_2(t - \tau)) \in BEDC$, то $\dot{y}_1(t) > 0, \dot{y}_2(t) > 0$;
- если же точка $(y_1(t - \tau), y_2(t - \tau))$ не принадлежит указанным областям, то обе производные отрицательны.

Будем рассматривать только непрерывную часть траектории (начальная функция может иметь разрывы). Пусть в начальный момент времени t^* первая переменная имеет любое, сколь угодно большое значение, а запаздывающая точка, определяющая знак производной, расположена сколь угодно близко к началу координат (точка C), что соответствует наибольшему значению производной. Тогда за время τ переменная $y_1(t^*)$ возрастет не более чем в $e^{\varepsilon_1 \tau}$ раз, а запаздывающая точка $(y_1(t^* - \tau), y_2(t^* - \tau))$ перейдет непременно в $(y_1(t^*), y_2(t^*))$, в силу непрерывности траектории. И так как последняя, по предположению, имеет большое значение первой переменной, то и производная первой переменной станет отрицательной. Знак указанной производной сохранится до тех пор, пока запаздывающая точка не вернется к началу координат. Вторая переменная имеет аналогичное поведение и поэтому все удаленные точки движутся к началу системы координат, а затем на некоторое расстояние отходят в сторону. Если считать вторую переменную y_2 близкой нулю (это равносильно рассмотрению траектории, проходящей почти по оси абсцисс), то первое уравнение системы (110) можно заменить на $\dot{y}_1(t) = \varepsilon_1 y_1(t)[1 - y_1(t - \tau)]$. И аналогично, полагая малой первую переменную, получим приближенное второе уравнение из системы (110): $\dot{y}_2(t) = \varepsilon_2 y_2(t)[1 - \beta y_2(t - \tau)]$. В разделе 3.3.3. мы подробно останавливались на поведении решений таких уравнений. Таким образом, упомянутая в утверждении область лежит внутри прямоугольника:

$$\left\{ 0 < y_1 < e^{\varepsilon_1 \tau}, 0 < y_2 < \frac{1}{\beta} e^{\varepsilon_2 \tau} \right\}.$$

Появление временного лага способно изменить всю качественную картину траекторий на фазовой плоскости: так с возрастанием запаздывания притягивающие точки становятся отталкивающими. Сформулируем все возможные ситуации в одном утверждении.

Утверждение. Положение равновесия $(1, 0)$ системы (110) будет асимптотически устойчиво, если $\alpha > 1, \varepsilon_1\tau < \pi/2$; и неустойчиво, если либо $\alpha < 1$, либо $\alpha > 1, \varepsilon_1\tau > \pi/2$.

Положение равновесия $(0, 1/\beta)$ будет асимптотически устойчиво, если $\beta < 1, \varepsilon_2\tau < \pi/2$; и неустойчиво, если либо $\beta > 1$, либо $\beta < 1, \varepsilon_2\tau > \pi/2$.

Положение равновесия (y_1^*, y_2^*) будет асимптотически устойчиво, если $\alpha < 1, \beta > 1, -z_1\tau < \pi/2$; и неустойчиво, если либо $\alpha > 1$, либо $\beta < 1$, либо $\alpha < 1, \beta > 1, -z_1\tau > \pi/2$. Здесь

$$z_1 = \frac{1}{2} \left(-\sqrt{(y_1^* \varepsilon_1 - y_2^* \varepsilon_2 \beta)^2 + 4\alpha y_1^* \varepsilon_1 y_2^* \varepsilon_2} - y_1^* \varepsilon_1 - y_2^* \varepsilon_2 \beta \right) < 0.$$

Обсудим, как получаются последние выводы. Для каждой стационарной точки составляется система уравнений в отклонениях и рассматривается ее линейная часть. Например, первая стационарная точка порождает линейное приближение системы в отклонениях вида:

$$\dot{y}'_1 = -\varepsilon_1(y'_{1\tau} + y'_{2\tau}); \quad \dot{y}'_2 = \varepsilon_2(1 - \alpha)y'_2.$$

Очевидно, что второе уравнение дает затухающие до нуля решения при $\alpha > 1$, а первое уравнение в этом случае становится неоднородным линейным с экспоненциально затухающей неоднородностью. Его однородная часть будет асимптотически устойчивой, если запаздывание не очень велико, а именно, если $\varepsilon_1\tau < \pi/2$. Последнее можно получить из приложения 1, где дано точное решение области асимптотической устойчивости для скалярного линейного уравнения с одним запаздыванием. Таким образом устанавливается вывод утверждения.

Несколько более сложно решается вопрос с нетривиальной стационарной точкой (y_1^*, y_2^*) . В этом случае система в отклонениях (ее линейная часть) имеет вид:

$$\dot{y}'_1 = -\varepsilon_1 y_1^* (y'_{1\tau} + y'_{2\tau}); \quad \dot{y}'_2 = -\varepsilon_2 y_2^* (\alpha y'_{1\tau} + \beta y'_{2\tau}). \quad (111)$$

Ее характеристический квазиполином определяется следующим равенством:

$$\det \left[\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + e^{-\lambda\tau} \begin{pmatrix} \varepsilon_1 y_1^* & \varepsilon_1 y_2^* \\ \alpha \varepsilon_2 y_1^* & \beta \varepsilon_2 y_2^* \end{pmatrix} \right] = 0.$$

С помощью замены переменных $z = \lambda e^{\lambda\tau}$ последнее уравнение превращается в квадратное

$$z^2 + (y_1^* \varepsilon_1 + y_2^* \varepsilon_2 \beta)z + y_1^* \varepsilon_1 y_2^* \varepsilon_2 (\beta - \alpha) = 0.$$

Если $\alpha < 1$, $\beta > 1$, оно имеет всегда два вещественных отрицательных корня:

$$z_{1,2} = \frac{1}{2} \left(\mp \sqrt{(y_1^* \varepsilon_1 - y_2^* \varepsilon_2 \beta)^2 + 4\alpha y_1^* \varepsilon_1 y_2^* \varepsilon_2} - y_1^* \varepsilon_1 - y_2^* \varepsilon_2 \beta \right). \quad (112)$$

Таким образом, приведенное выше утверждение полностью описывает асимптотическое поведение траекторий во всей положительной четверти фазовой плоскости.

Рассмотрим теперь вопрос о наличии периодических решений системы (110). Как мы уже отмечали, при нулевом временном лаге колебаний быть не может (имеется в виду колебательное движение по замкнутой траектории). Введение запаздывания изменяет картину. Теперь не точка, а отрезок траектории определяет решение задачи Коши для системы (110), и вследствие этого траектория может пересекать сепаратрису, проходящую от начала системы координат к нетривиальной равновесной точке (y_1^*, y_2^*) . Покажем, что если и существует замкнутая траектория, по которой происходит периодическое движение, то оно описывается равенствами:

$$\frac{1}{T} \int_0^T y_1(t) dt = y_1^*, \quad \frac{1}{T} \int_0^T y_2(t) dt = y_2^*,$$

где T — соответствующий период. Действительно, из (110) следуют интегральные равенства:

$$\begin{aligned} y_1(t) &= y_1(0) \exp \varepsilon_1 \int_0^t [1 - y_1(s - \tau) - y_2(s - \tau)] ds, \\ y_2(t) &= y_2(0) \exp \varepsilon_2 \int_0^t [1 - \alpha y_1(s - \tau) - \beta y_2(s - \tau)] ds. \end{aligned} \quad (113)$$

Предполагаемая периодичность означает, что для любых t выполняются тождества: $y_1(t) = y_1(t + T)$, $y_2(t) = y_2(t + T)$ и, в частности, $y_1(0) = y_1(T)$, $y_2(0) = y_2(T)$. Тогда из (110) имеем соотношения:

$$\int_0^T [1 - y_1(s - \tau) - y_2(s - \tau)] ds = 0,$$

$$\int_0^T [1 - \alpha y_1(s - \tau) - \beta y_2(s - \tau)] ds = 0,$$

которые представляют собой линейную алгебраическую систему. Ее решение записывается в виде:

$$\frac{1}{T} \int_0^T y_1(s - \tau) ds = y_1^*, \quad \frac{1}{T} \int_0^T y_2(s - \tau) ds = y_2^*.$$

В силу аддитивности интеграла и периодичности будут справедливы следующие равенства:

$$\int_0^T y_1(s - \tau) ds = \int_{-\tau}^{T-\tau} y_1(s) ds = \int_{-\tau}^0 + \int_0^T - \int_{T-\tau}^T = \int_0^T y_1(s) ds,$$

$$\int_0^T y_2(s - \tau) ds = \int_0^T y_2(s) ds.$$

А это и означает, что среднее интегральное по периодическому движению совпадает с положением равновесия (y_1^*, y_2^*) . Назовем такие движения (y_1^*, y_2^*) центральными. Заметим, что такая покоординатная центральность не гарантирует нам того, что периодическая траектория будет охватывать со всех сторон центр (y_1^*, y_2^*) , т. е. если провести луч из этого центра в любую точку периодической траектории, то совсем необязательно при полном обходе точкой контура луч повернется на полный оборот вокруг центра. Вычисления показывают существование предельных циклов, которые нельзя деформировать в единичную окружность с центром в (y_1^*, y_2^*) непрерывным и однородным относительно этого центра преобразованием (гомеоморфизмом).

В его сколь угодно малой окрестности, где главную роль играет линейное приближение системы в отклоне-

ниях (111), при условии выполнения (106) будет справедливо следующее свойство: линейная система (111) имеет вокруг нуля (что соответствует точке (y_1^*, y_2^*) для системы (110)) два семейства периодических решений — одно соответствует запаздыванию $\tau_1 = -\pi/(2z_1)$ с периодом $T_1 = -2\pi/z_1$, а второе запаздыванию $\tau_2 = -\pi/(2z_2)$ с периодом $T_2 = -2\pi/z_2$. Амплитуда этих периодических решений зависит от начальных функций. Отрицательные величины z_1 и z_2 определены в (112).

Вычисление периода здесь осуществляется по уже пройденной схеме: для указанного фиксированного запаздывания τ характеристический квазиполином $z_1 = \lambda e^{\lambda\tau}$ соответствует дифференциальному уравнению с последствием $\dot{\xi}(t) = z_1 \xi(t - \tau)$, которое заменой шкалы времени $t = 2\tau s/\pi$ сводится к уравнению

$$\dot{\eta}(s) = \frac{2\tau z_1}{\pi} \eta\left(s - \frac{\pi}{2}\right) = -\eta\left(s - \frac{\pi}{2}\right).$$

Последнее имеет решение вида $\eta(s) = C_1 \cos(s) + C_2 \sin(s)$ с периодом 2π . Возвращаясь к времени t , получим период, равный четырем запаздываниям.

Оказывается, что система (113) всегда, при любом запаздывании, имеет (y_1^*, y_2^*) -центральное периодическое решение. Ниже мы обоснуем это утверждение. Для этого из выражений (113) строятся итеративные процедуры так, чтобы каждая итерация была периодическая и (y_1^*, y_2^*) -центральная. Предел полученных итераций и будет искомым периодическим решением.

Сначала отметим, что упомянутое выше произвольное запаздывание можно заменить на $\tau_1 \geq -\pi/(2z_1)$, поскольку в противном случае все решения из положительного квадранта притягиваются к стационарной точке (y_1^*, y_2^*) , т. е. она в наших условиях является глобальным аттрактором. Естественно, стационарное решение удовлетворяет условию периодичности и может служить пределом, упомянутым выше. Но так как практической пользы от такого расширенного толкования периодических решений нет, то сразу будем считать, что временной лаг больше указанной величины.

Итак, для произвольного запаздывания τ и произвольного T выберем пару непрерывных, (y_1^*, y_2^*) -центральных, T -периодических функций $(y_1^{(0)}(t), y_2^{(0)}(t))$. Подставим их в интегральное представление системы (110) и получим другую пару функций:

$$\hat{y}_1^{(1)}(t) = y_1^{(0)}(0) \exp \varepsilon_1 \int_0^t [1 - y_1^{(0)}(s - \tau) - y_2^{(0)}(s - \tau)] ds,$$

$$\hat{y}_2^{(1)}(t) = y_2^{(0)}(0) \exp \varepsilon_2 \int_0^t [1 - \alpha y_1^{(0)}(s - \tau) - \beta y_2^{(0)}(s - \tau)] ds,$$

но они не являются (y_1^*, y_2^*) -центральными. Для реализации этого свойства нормируем каждую из полученной пары функций с помощью интеграла по периоду:

$$y_1^{(1)}(t) = Ty_1^* \left[y_1^{(0)}(0) \int_0^T \exp \varepsilon_1 \int_0^t [1 - y_1^{(0)}(s - \tau) - y_2^{(0)}(s - \tau)] ds dt \right]^{-1} \hat{y}_1^{(1)}(t),$$

$$y_2^{(1)}(t) = Ty_2^* \left[y_2^{(0)}(0) \int_0^T \exp \varepsilon_2 \int_0^t [1 - \alpha y_1^{(0)}(s - \tau) - \beta y_2^{(0)}(s - \tau)] ds dt \right]^{-1} \hat{y}_2^{(1)}(t).$$

Теперь функции $(y_1^{(1)}(t), y_2^{(1)}(t))$ и непрерывны, и (y_1^*, y_2^*) -центральны, и T -периодичны. Таким же образом можно получить любые итерации $(y_1^{(k)}(t), y_2^{(k)}(t))$, заменяя в последних формулах верхний индекс ноль на k , а единицу на $(k + 1)$. Для них всех останутся справедливы последние свойства. Отметим, что для указанных итераций образовалась последовательность начальных точек

$$\left(Ty_1^* \left[\int_0^T \exp \varepsilon_1 \int_0^t [1 - \dots]_{(k)} ds dt \right]^{-1}, Ty_2^* \left[\int_0^T \exp \varepsilon_2 \int_0^t [1 - \dots]_{(k)} ds dt \right]^{-1} \right).$$

Поскольку каждая компонента пары $(y_1^{(k)}(t), y_2^{(k)}(t))$ непрерывна на периоде, положительна и имеет фиксированное среднее, то существует общая для всех функций мажоранта. Если бы это было не так, т. е. для любого сколь угодно большого числа M нашлись бы номер k_M и момент времени $t_M \in (0, T)$ такие, что $y_{1,2}^{(k_M)}(t_M) > M$, то обязательно можно было бы (из центральности) определить положи-

тельное число $w < \tau$ такое, что $Mw > Ty_{1,2}^*$, а это повлекло бы неограниченность правых частей системы (110).

Из существования общей для функциональной последовательности мажоранты следует существование предельной пары функций. Она так же должна быть (y_1^*, y_2^*) -центральной и T -периодична. Этот предел и является искомым периодическим решением системы (110).

Замечание. Процедуру построения итераций вряд ли удастся превратить в работающий численный метод получения периодических решений. Одна из проблем здесь связана с точным знанием периода T .

Рассмотрим две экономических модели для того, чтобы представлять себе значения параметров, входящих в модель. Мы будем строить систему (101) для двух фирм. Сначала предположим, что производственные фирмы работают на общем рынке и используют одни и те же источники производственных факторов: имеются в виду трудовые, сырьевые и капитальные ресурсы. Это могут быть домостроительные комбинаты, предприятия легкой промышленности и т. п.

Напомним, что речь идет о системе уравнений:

$$\begin{aligned}\dot{x}_1(t) &= x_1(t)[\varepsilon_1 - \gamma_{11}x_1(t - \tau) - \gamma_{12}x_2(t - \tau)], \\ \dot{x}_2(t) &= x_2(t)[\varepsilon_2 - \gamma_{21}x_1(t - \tau) - \gamma_{22}x_2(t - \tau)].\end{aligned}$$

Оценим параметры модели, которые отвечают за собственные возможности предприятий, без учета конкурирующей фирмы. Пусть технологические и управленческие возможности первого предприятия таковы, что оно способно удвоить выпуск продукции за 5 лет. Тогда из первого уравнения модели получим при малых значениях x_1 :

$$x_1(5) = x_1(0)e^{\varepsilon_1 5}.$$

Отсюда имеем приближенное равенство $\varepsilon_1 \approx 0,14$. Аналогично предположим, что второе предприятие при малых собственных объемах выпуска способно утроить за 5 лет объем производства. Тогда из второго уравнения модели получим $\varepsilon_2 \approx 0,22$.

Каждый из управляющих может оценить оптимальный размер предприятия, пользуясь известным критерием

через цены на факторы производства и вид производственной функции (Приложение 3). Пусть из этих оценок следует, что без учета конкуренции первое предприятие способно выпускать 20 единиц продукции, а второе — 10. Но мы знаем, что при отсутствии конкуренции предельный объем выпуска на первом предприятии будет равен $\varepsilon_1/\gamma_{11}$, а на втором — $\varepsilon_2/\gamma_{22}$. Тогда вычисляются соответствующие γ : $\gamma_{11} = 0,007$; $\gamma_{22} = 0,022$. Самое трудное оценить взаимное влияние двух фирм, т. е. γ_{12} , γ_{21} . Либо следует предположить, что управляющие фирм имеют возможность договориться об объемах поставок товара на общий рынок, либо делаются регулярные маркетинговые исследования, из которых можно судить о равновесии в конкурентных отношениях. Предположим, что равновесие достигается при $x_1^* = 15$; $x_2^* = 8$; тогда вычисляем оставшиеся параметры: $\gamma_{12} = 0,0044$; $\gamma_{21} = 0,003$.

Итак, модель взаимодействия двух производственных фирм построена. Чтобы применить полученные выше выводы, нам необходимо знать α , β , z_1 . Согласно приведенным ранее формулам $\alpha = 0,27$; $\beta = 3,18$; и $z_1 = -0,19$. Таким образом, нетривиальное положение равновесия на рынке двух производственных фирм (15, 8) будет асимптотически устойчивым, т. е. какими бы ни были начальные условия, с течением времени оба производства достигнут равновесной точки, при этом они могут не волноваться относительно временных лагов в их технологических процессах или в финансовых стратегиях, так как устойчивость стационарной точки будет гарантирована до запаздывания, равного 8 годам. Никаких колебательных процессов в динамике наблюдаться не должно.

Следующий пример конкурентной борьбы на общем рынке ресурсов относится к торговле. Давайте рассмотрим взаимное влияние мелкой уличной торговли (первая фирма) и торговли, располагающей стационарными помещениями, складами, демонстрационными залами и т. д. (вторая фирма). Так как каждая из фирм не способна закрыть весь рынок — уличная ближе к покупателю и мобильней, но не может охватить всех товаров и предоставить удобные и современные способы торговли, примем условно, что

уличная торговля в принципе может удовлетворить только 30% рынка, а стационарная — 80% (далее в этих единицах и будем измерять деятельность торговли). Кроме того, положим, что первая фирма может увеличить свои торговые точки в десять раз за один год, если стационарная торговля вдруг перестанет существовать. Для последней возможно лишь удвоение за год, поскольку строительство, аренда, регистрация, подготовка продавцов, оптовые закупки — все требует времени. Таким образом, как и в первом примере, получим для коэффициентов модели значения: $\varepsilon_1 \approx 2,3$; $\varepsilon_2 \approx 0,69$; $\gamma_{11} = 0,077$; $\gamma_{22} = 0,0086$. В результате различных видов конкурентной борьбы (рекламы, демпинговых цен, влияния на городские власти и т. п.) сформировалось равновесие: $x_1^* = 25\%$, $x_2^* = 75\%$, из которого можно получить недостающие параметры $\gamma_{12} = 0,005$; $\gamma_{21} = 0,0018$. Модель готова — можно ознакомиться, что будет из нее следовать. Так как $\alpha = 0,078$; $\beta = 5,73$ и $z_1 = -1,94$, то стационарная точка (25%, 75%) будет асимптотически устойчивой только в том случае, если реакция рынка на действие каждой фирмы будет происходить без большого запаздывания, а именно, временной лаг не должен превышать 0,81 года. В противном случае вокруг стационарной точки начнутся колебания: то уличная торговля захватит больше 25% рынка, то стационарная торговля выйдет за свои 75%. Чем больше запаздывание, тем больше период колебаний и их амплитуда. Когда запаздывание чуть превысит критическую величину 0,81, период колебаний будет несколько больше трех лет (четыре временных лага). Если воспользоваться компьютером и провести численное моделирование, то можно увидеть следующую картину. При временном лаге 0,82 (т. е. чуть больше критического) уличная торговля начинает испытывать колебания с амплитудой 7% от всего рынка, т. е. примерно 25% от своего объема. С увеличением лага амплитуда быстро растет, и при $\tau = 0,9$ в течение каждых четырех лет уличная торговля изменяется от нуля до 45% всего рынка. Но колебания охватывают и стационарную торговлю: так, при запаздывании 0,82 амплитуда составляет 2% всего рынка или 3,2% от своего стационарного объема. Такие явления становятся

заметными на сборе налогов и на городском бюджете. С ростом запаздывания амплитуда колебаний и период растут почти экспоненциально. Следует отметить, что на продолжительность временного лага оказывают влияние сроки кредитов в банковской системе, политическое устройство городских властных структур, удаленность от мировых и отечественных торговых потоков и т. д.

3.4.4. ВЗАИМОЗАВИСИМАЯ ДЕЯТЕЛЬНОСТЬ С ПОСЛЕДЕЙСТВИЕМ

В этом разделе мы рассмотрим такое взаимодействие экономических агентов, когда одни из них существенно зависят от других. Простейшими примерами таких отношений служат модели бюджетного или аграрного секторов в экономике страны. Биологическим аналогом подобных моделей служит система «хищник–жертва». Без временного лага для двух агентов она записывается следующим образом:

$$\begin{aligned}\dot{x}_1(t) &= x_1(t)[- \varepsilon_1 + \gamma_{12}x_2(t)], \\ \dot{x}_2(t) &= x_2(t)[\varepsilon_2 - \gamma_{21}x_1(t) - \gamma_{22}x_2(t)],\end{aligned}\quad (114)$$

где все коэффициенты положительны и $x_1(t)$ характеризует зависимую отрасль, например сельское хозяйство, а $x_2(t)$ — отрасль донора, например всю промышленность или добывающий сектор.

Как и в предыдущем разделе введем более удобные координаты (при которых модель имеет меньше параметров):

$$y_1 = (\gamma_{21}/\varepsilon_2)x_1, \quad y_2 = (\gamma_{12}/\varepsilon_1)x_2.$$

Тогда система (114) будет иметь вид:

$$\begin{aligned}\dot{y}_1(t) &= \varepsilon_1 y_1(t)[-1 + y_2(t)], \\ \dot{y}_2(t) &= \varepsilon_2 y_2(t)[1 - y_1(t) - \alpha y_2(t)],\end{aligned}\quad (115)$$

где $\alpha = \gamma_{22}\varepsilon_1/(\gamma_{12}\varepsilon_2)$. Модель имеет три стационарные точки $(0, 0)$, $(0, 1/\alpha)$, $(1 - \alpha, 1)$. Так же, как и раньше, составляя систему в отклонениях, получим, что первая точка всегда является неустойчивой, типа «седло», вторая будет асимптотически устойчивой при $\alpha > 1$ и неустойчивой при

$0 < \alpha < 1$. Третья точка, соответствующая нормальному функционированию экономики (в том смысле, что ни одна из отраслей не прекращает своей деятельности), является асимптотически устойчивой при $0 < \alpha < 1$. Последнее утверждение легко обосновывается с помощью второго метода Ляпунова. Рассмотрим определенно положительную функцию относительно третьей равновесной точки

$$V(y_1, y_2) = \frac{1}{\varepsilon_1} \left(y_1 - (1 - \alpha) - (1 - \alpha) \ln \frac{y_1}{1 - \alpha} \right) + \frac{1}{\varepsilon_2} (y_2 - 1 - \ln y_2).$$

Ее полная производная по времени в силу системы (115) является неположительной функцией

$$\dot{V}|_{(21)} = -\alpha(1 - y_2)^2.$$

Согласно теореме Барбашина–Красовского [6], если производная не положительна, то для асимптотической устойчивости достаточно, чтобы множество (y_1, y_2) , где она равна нулю, не содержало целых траекторий, кроме стационарной точки. Это в нашем случае выполнено.

Таким образом, при $0 < \alpha < 1$ нетривиальное положительное равновесие будет глобальным аттрактором, т. е. из любых начальных состояний выходы соответствующих отраслей стремятся к установившемуся значению. При этом сам характер стремления может быть различным. Если

$$0 < \alpha < \frac{2}{\varepsilon_2} \left(\sqrt{\varepsilon_1^2 + \varepsilon_1 \varepsilon_2} - \varepsilon_1 \right) < 1,$$

то стационарная точка будет особой точкой, типа «фокус», иначе — типа «узел». Это значит, что в первом случае траектории достигают равновесной точки, вращаясь вокруг нее по спирали.

В предельной ситуации, когда $\alpha = 0$, вокруг положения равновесия образуются замкнутые траектории с периодическими движениями по ним. Уравнения этих замкнутых кривых можно получить как уравнения поверхностей уровней определенно положительной функции Ляпунова, которая уже использовалась нами:

$$V(y_1, y_2) = 1/\varepsilon_1 (y_1 - 1 - \ln y_1) + 1/\varepsilon_2 (y_2 - 1 - \ln y_2) = \text{const.}$$

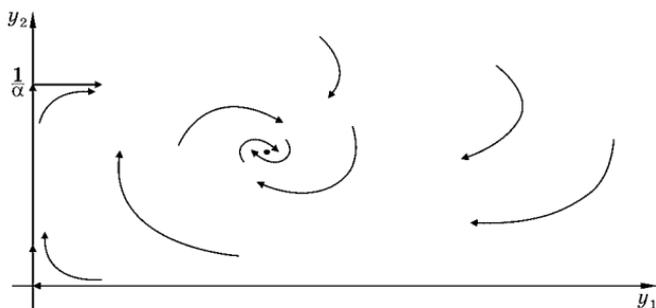


Рис. 22

Нетривиальное положение равновесия является притягивающим «фокусом»

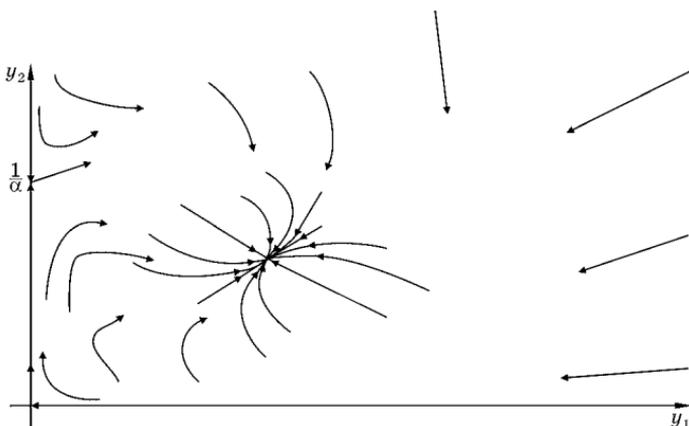


Рис. 23

Нетривиальное положение равновесия является притягивающим «узлом»

Ниже на рисунках 22 и 23 изображены примерные траектории в положительном квадранте координатной плоскости описанных выше случаев.

Вывод. При отсутствии временного лага поведение взаимозависимых отраслей, допускающих их совместное существование, определяется параметрами модели: если верно неравенство

$$\frac{\gamma_{22}}{\gamma_{12}} < \frac{2}{\varepsilon_1} \left(\sqrt{\varepsilon_1^2 + \varepsilon_1 \varepsilon_2} - \varepsilon_1 \right),$$

то «выходы» отраслей будут колебаться вокруг стационарной точки $(1/\gamma_{21}(\varepsilon_2 - \gamma_{22}\varepsilon_1/\gamma_{12}), \varepsilon_1/\gamma_{12})$, втягиваясь в нее по спирали. Если же указанное выше неравенство не выпол-

няется, то динамике характерно монотонное стремление к равновесной точке.

Если $\gamma_{22} = 0$, что соответствует ситуации неограниченного развития промышленности, взаимодействие промышленности и сельского хозяйства (как пример) характеризуется периодическими колебаниями вокруг нетривиальной стационарной точки с периодом

$$T = \frac{2\pi}{\sqrt{\varepsilon_1 \varepsilon_2}}$$

(отметим, что этот факт был установлен еще В. Вольтеррой).

Учет временного лага в описанных выше моделях возможен так же, как это мы делали раньше: в системах (114) и (115) в квадратных скобках переменные системы берутся не в текущий момент времени, а в предшествующий — $(t - \tau)$. Будем рассматривать только нетривиальное положение равновесия $(1 - \alpha, 1)$, да и то только при условии $0 < \alpha < 1$, то есть когда без запаздывания оно асимптотически устойчиво по Ляпунову. Из системы

$$\begin{aligned} \dot{y}_1(t) &= \varepsilon_1 y_1(t)[-1 + y_2(t - \tau)], \\ \dot{y}_2(t) &= \varepsilon_2 y_2(t)[1 - y_1(t - \tau) - \alpha y_2(t - \tau)] \end{aligned}$$

составим систему в отклонениях относительно рассматриваемого равновесия, выпишем ее линейную часть (так же, как это делалось в предыдущем разделе):

$$\dot{y}'_1 = \varepsilon_1(1 - \alpha)y'_{2\tau}; \quad \dot{y}'_2 = -\varepsilon_2(y'_{1\tau} + \alpha y'_{2\tau}).$$

Ее характеристический квазиполином определяется следующим равенством:

$$\det \left[\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + e^{-\lambda\tau} \begin{pmatrix} 0 & -\varepsilon_1(1 - \alpha) \\ \varepsilon_2 & \varepsilon_2\alpha \end{pmatrix} \right] = 0.$$

Относительно $z = \lambda e^{\lambda\tau}$ последнее уравнение превращается в квадратное

$$z^2 + \varepsilon_2\alpha z + \varepsilon_1\varepsilon_2(1 - \alpha) = 0.$$

Его решения при обозначенных условиях будут всегда лежать в левой полуплоскости комплексной плоскости:

$$z_{\pm} = -\frac{\varepsilon_2\alpha}{2} \pm \sqrt{\frac{\varepsilon_2^2\alpha^2}{4} - \varepsilon_1\varepsilon_2(1 - \alpha)}.$$

И если решение получилось вещественное — это будет, как уже отмечалось в предыдущем разделе, при

$$\frac{2}{\varepsilon_2} \left(\sqrt{\varepsilon_1^2 + \varepsilon_1 \varepsilon_2} - \varepsilon_1 \right) < \alpha < 1$$

— то, пользуясь старым приемом, устанавливаем, при каких временных лагах стационарная точка останется притягивающей: $\tau < -\pi/(2z_-)$. При комплексных решениях последнего уравнения необходимо воспользоваться замечанием в конце второго параграфа в Приложении 1. Из него следует, что положение равновесия останется асимптотически устойчивым, если точка с координатами

$$\left(\frac{\varepsilon_2 \alpha \tau}{2}, \tau \sqrt{\varepsilon_1 \varepsilon_2 (1 - \alpha) - \frac{\varepsilon_2^2 \alpha^2}{4}} \right)$$

попадет в часть плоскости, расположенную выше оси абсцисс, но ниже кривой, полярные координаты которой связаны уравнением

$$\rho = \pi/2 - \varphi, \varphi \in [0, \pi/2].$$

На картинке область устойчивости изображена заштрихованной фигурой. Условия выглядят громоздко, но они легко проверяемы. Очевидно, что при малых запаздываниях упомянутая точка обязательно принадлежит фигуре и, следовательно, в положение равновесия втягиваются все траектории по спиралям. С ростом запаздывания после некоторого критического значения в динамике поведения отраслей появляются устойчивые колебания с возрастающими амплитудой и периодом. Рассмотрим числовой пример.

Пусть x_1 — валовой продукт сельского хозяйства за год, а x_2 — валовой продукт промышленности страны. Сделаем ряд правдоподобных предположений, чтобы идентифицировать коэффициенты модели (114). Если считать, что промышленность столь динамична, что способна за пять лет, начиная с малых объемов, удвоить свой выпуск, то из интегрального представления второго из уравнений (114) получим $\varepsilon_2 = 0,14$. Столкнувшись с перепроизводством или

нехваткой производственных факторов, промышленность остановится в развитии. Этот предел обозначим как 100%. Тогда без влияния сельского хозяйства второе уравнение позволяет вычислить следующий параметр: $\gamma_{22} = 0,0014$. Предположим теперь, что правительство намерено помогать сельскому хозяйству, передавая часть выхода промышленности или перераспределяя инвестиции, или влияя на цены, или выплачивая дотации, а само сельское хозяйство таково, что без помощи извне оно за десять лет сократит производство в десять раз, т. е. окончательно придет в упадок. Последнее предположение дает $\varepsilon_1 = 0,23$. Предположительные расчеты в правительстве показали, что планируемая помощь приведет к устойчивому функционированию экономики около равновесной точки (10%, 80%) (единицы измерения соответствуют предельному выпуску промышленности). Отсюда вычисляем: $\gamma_{12} = 0,0029$, $\gamma_{21} = 0,0028$. В этом случае оставшийся параметр модели α принимает значение 0,8. Таким образом, при малых запаздываниях положение равновесия будет асимптотически устойчиво. Для оценки допустимых запаздываний необходимо воспользоваться критерием, описанным выше. Поскольку это происходит первый раз, сделаем все операции подробно. Поскольку

$$\frac{2}{\varepsilon_2} \left(\sqrt{\varepsilon_1^2 + \varepsilon_1 \varepsilon_2} - \varepsilon_1 \right) = 0,88 > \alpha,$$

мы не можем применить неравенство $\tau < -\pi/(2z_-)$. Найденные значения параметров модели позволяют утверждать, что если точка $(0,056\tau, 0,0575\tau)$ расположена ниже линии с уравнением $\rho = \pi/2 - \varphi$, $\varphi \in [0, \pi/2]$ в полярных координатах, то соответствующие τ допустимы. Для произвольных τ упомянутая точка лежит на центральной прямой с углом наклона $\varphi = \arctg(0,0575/0,056) \approx 0,8$. Тогда длина отрезка этой прямой, лежащего внутри фигуры, меньше 0,77. Отсюда $\tau < 0,77 \cos 0,8/0,056 \approx 9,58$. Окончательно утверждается, что равновесие останется асимптотически устойчивым при запаздываниях почти до 10 лет. Хотя этот срок и кажется большим, но экономические отношения промышленности и сельского хозяйства могут строиться на

долгосрочной основе. Например, инвестиции в строительство заводов сельхозтехники, долгосрочные программы мелиорации, электрификации и газификации села согласно этой модели не должны длиться более 10 лет.

3.4.5. ПРОИЗВОЛЬНОЕ ЧИСЛО УЧАСТНИКОВ РЫНКА

В задачах со многими участниками экономической деятельности взаимное влияние и конкуренция так переплетаются, что не представляется возможным использовать какую-то определенную модель. Поэтому в данном разделе сформулированы общие принципы и способы моделирования динамики сложных экономических систем. Обозначим

$$\text{diag}\{X\} = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x_n \end{pmatrix}.$$

Тогда многомерную модель Лотки–Вольтерры можно записать в форме

$$\dot{X}(t) = \text{diag}\{X(t)\}(E - \Gamma X(t - \tau)).$$

Как уже отмечалось в начале этого параграфа, нетривиальное положение равновесия всей системы дается решением уравнения (103). Для экономической осмысленности результата необходимо, чтобы это решение было положительным. Обозначим его звездочкой сверху: $X^* = \Gamma^+ E$. Здесь верхний индекс «+», как и раньше, показывает на псевдообратную матрицу, так как обратная матрица может и не существовать в общем случае. Система в отклонениях относительно рассматриваемого положительного равновесия записывается в форме

$$\dot{X}' = \text{diag}\{X' + X^*\}(-\Gamma X'_t),$$

а характеристическое уравнение ее линейного приближения имеет вид:

$$\det\{\lambda I + \text{diag}\{\Gamma^+ E\}\Gamma e^{-\lambda\tau}\} = 0.$$

Обозначая, как и раньше, $z = \lambda e^{\lambda t}$, можем вычислить $\{z_i\}$ — спектр матрицы $-\text{diag}\{\Gamma^+ E\}\Gamma$.

Утверждение. Если среди комплексных чисел z_i есть хотя бы одно с положительной вещественной частью, то положение равновесия неустойчиво по Ляпунову. В этом случае на гранях положительного ортанта есть притягивающее положение равновесия. Оно соответствует уничтожению одного из экономических агентов с течением времени.

Если все числа z_i имеют строго отрицательные вещественные части, то при малых запаздываниях положение равновесия асимптотически устойчиво по Ляпунову, т. е. все траектории из любых частей положительного ортанта с течением времени притягиваются к равновесной точке. Установившиеся колебания отсутствуют. Все экономические агенты развиваются предсказуемым плавным образом, стремясь к соответствующим предельным величинам.

Увеличение запаздывания до некоторого критического уровня приводит к нарушению асимптотической устойчивости положения равновесия, но в его окрестности возникают колебательные движения. Сначала они являются периодическими, а затем при дальнейшем увеличении запаздывания становятся почти периодическими и рекуррентными.

Критическое значение запаздывания τ^* определяется следующим образом: для всякого комплексного собственного числа $z_i = -a + ib$ ($a, b > 0$) матрицы $-\text{diag}\{\Gamma^+ E\}\Gamma$ строится центральная прямая линия точек $(a\tau, b\tau)$, которая при некотором τ^* пересекает кривую $\rho = \pi/2 - \varphi$, $\varphi \in [0, \pi/2]$, или в более громоздком аналитическом виде:

$$\tau^* = \frac{\pi - 2\arctg \frac{b}{a}}{2\sqrt{a^2 + b^2}}.$$

Оставшийся случай характеризуется неположительными вещественными частями комплексных чисел z_i и существованием пар чисто мнимых собственных чисел матрицы $-\text{diag}\{\Gamma^+ E\}\Gamma$. При нулевых временных лагах для двух уравнений в окрестности равновесия формируются периодические движения. Запаздывание делает стационарную

точку в этом случае неустойчивой, но будет ли существовать стабильное колебание, или будет ли существовать многообразие (не обязательно линейное) со стабильными колебаниями в многомерном случае, в настоящий момент неизвестно.

3.5. ПРОБЛЕМА ИДЕНТИФИКАЦИИ СИСТЕМЫ

Реальные динамические процессы проходят под воздействием множества причин и возмущающих факторов. Никакая математическая формализация процессов не может их учесть абсолютно полностью. Поэтому этап идентификации системы, т. е. установления параметров уравнений, отвечающих за данный конкретный процесс, является оптимизационным этапом моделирования. Все предшествующие рассуждения относились к выбору класса функций, заданного с точностью до некоторых параметров. Теперь же настала очередь подобрать параметры так, чтобы некоторый функционал качества модели достиг своего минимума. В данном случае классом функций модели являются решения системы логистических уравнений, или иначе модели Лотки–Вольтерры. Ниже будут обсуждены и даны соответствующие формулы и рекомендации для модели без запаздывания, с запаздыванием и с ошибками в данных наблюдения.

3.5.1. ИДЕНТИФИКАЦИЯ МОДЕЛИ БЕЗ ВРЕМЕННОГО ЛАГА И БЕЗ ИСКАЖЕНИЯ ИНФОРМАЦИИ

Рассмотрим сначала скалярное логистическое уравнение

$$\dot{x}(t) = x(t)[a - bx(t)], \quad (116)$$

где a, b — положительные параметры. Предположим, что мы знаем решение этого уравнения в дискретные моменты времени $t_i = i\Delta, i = 0, \dots, N: x_i$. Идентификация уравнения (116) состоит в нахождении чисел a, b по наблюдениям $\{x_i\}$.

Данное уравнение решается просто:

$$x(t)/(a - bx(t)) = x(0)/(a - bx(0))e^{at},$$

и, следовательно, имея всего три наблюдения x_0, x_1, x_2 , можно выписать алгебраическую систему двух уравнений для нахождения чисел a, b . Однако этот путь не приведет к общему конструктивному алгоритму.

Проведем дискретизацию уравнения (116). Так как

$$\frac{d}{dt} \ln x(t) = a - bx(t),$$

то для соседних моментов времени t_k, t_{k+1} , интегрируя последнее равенство, получим:

$$\ln x_{k+1} = \ln x_k + \int_{t_k}^{t_{k+1}} [a - bx(t)] dt.$$

Вычисление интеграла невозможно провести точно, поскольку значение $x(t)$ между наблюдениями неизвестно. Поэтому интеграл придется аппроксимировать. Это можно сделать различными способами с различной точностью окончательного результата. Имея в виду общий алгоритм, остановимся на приближениях первого порядка:

$$\ln x_{k+1} \approx \ln x_k + \Delta(a - bx_k).$$

Из этого равенства выводим функционал, минимум которого следует находить, чтобы восстановить a, b :

$$F(a, b) = \sum_{k=0}^{N-1} [\ln x_{k+1} - \ln x_k - \Delta(a - bx_k)]^2. \quad (117)$$

К сожалению, нельзя утверждать, что при точном значении чисел a, b функционал примет нулевое значение, но, учитывая ошибки измерения и моделирования, реальные процессы все равно не давали бы точного нуля. Поэтому примем как определение, что числа a, b , полученные при минимизации функционала $F(a, b)$, определяют наилучшую модель из всех уравнений (116) для данного набора наблюдений за процессом $\{x_i\}$.

Обобщим этот результат на векторный случай. Пусть X — это n -мерный вектор с компонентами x_i , дискретное время же обозначим целочисленным аргументом $x_i(k)$,

$i = 1, \dots, n; k = 0, 1, \dots, N$. Из формы системы уравнений (101) следует (при нулевом запаздывании), что

$$\ln x_i(k+1) \approx \ln x_i(k) + \Delta \left[\varepsilon_i - \sum_{j=1}^n \gamma_{ij} x_j(k) \right].$$

Здесь знак приближенного равенства подчеркивает наличие такой же процедуры аппроксимации интеграла, какая имела место в скалярном случае.

Выберем функционал, характеризующий качество настройки параметров системы, в виде суммы квадратов отклонений по всем векторам наблюдений:

$$F(E, \Gamma) = \sum_{k=0}^{N-1} \sum_{i=1}^n \left\{ \ln x_i(k+1) - \ln x_i(k) - \Delta \left[\varepsilon_i - \sum_{j=1}^n \gamma_{ij} x_j(k) \right] \right\}^2.$$

Нетрудно видеть, что функционал является квадратичным по искомым параметрам $\varepsilon_i, \gamma_{ij}$. Так же, как это делалось уже неоднократно (разделы 2.3.3 и 2.3.4), заменой переменных:

$$D = E\Delta, \quad G = -\Gamma\Delta, \quad Y_k = \{\ln x_i(k+1) - \ln x_i(k)\}_{i=1}^n, \quad Z_k = \begin{Bmatrix} X(k) \\ 1 \end{Bmatrix},$$

переводим системы к известной задаче минимизации квадратичного функционала:

$$F'(D, G) = \sum_{k=0}^{N-1} \|Y_k - \{G, D\}Z_k\|^2.$$

Тогда, как следует из рассуждений в разделе 1.3.4, если среди векторов $\{Z_k\}$, $k = 0, \dots, N-1$, есть $(n+1)$ линейно независимый вектор, то минимум последнего функционала достигается в единственной точке

$$\{G, D\} = \left(\sum_{k=0}^{N-1} Y_k Z_k^T \right) \left(\sum_{k=0}^{N-1} Z_k Z_k^T \right)^{-1}.$$

Замечание. Здесь очень важно отметить, что наблюдения $x_i(k)$, $i = 1, \dots, n; k = 0, 1, \dots, N$ должны быть строго положительными, так как иначе невозможно использовать

функцию \ln . Это обстоятельство ограничивает применение рассматриваемых моделей. Например, ВВП имеет колебания около кривой роста экономики, поэтому считается обычным моделировать не сам ВВП, а его изменение, но изменение ВВП может быть отрицательным в период политических и экономических потрясений в стране. Таким образом, в задачах макроуровня модели Лотки–Вольтерры наиболее естественны для стран со стабильной экономикой.

3.5.2. ИДЕНТИФИКАЦИЯ МОДЕЛИ БЕЗ ЗАПАЗДЫВАНИЯ И С ОШИБКАМИ В ИНФОРМАЦИИ

В параграфах первой главы мы рассмотрели временные ряды с ошибками измерения, на которые накладывались ограничения независимости распределения от времени и наличия нулевого среднего. Предположим и теперь, что

$$x_i(k) = \hat{x}_i(k) + \mu_i, \quad i = 1, \dots, n; \quad k = 0, 1, \dots, N, \quad (118)$$

где знаком « \wedge » обозначены детерминированные составляющие наблюдений, а μ_i — случайная ошибка измеряемых величин для каждой компоненты вектора $X(k)$, имеющая нулевое математическое ожидание и достаточно малую дисперсию. Тогда, раскладывая функцию \ln в ряд Тейлора, получим приближенное равенство:

$$\ln x_i(k) = \ln(\hat{x}_i(k) + \mu_i) \approx \ln \hat{x}_i(k) + \frac{\mu_i}{\hat{x}_i(k)}.$$

Таким образом, $\ln x_i(k)$ определен и непрерывен по m , если $x_i(k) > 0$, $\hat{x}_i(k) > 0$.

Рассмотрим скалярный случай как более простой, т. е. проверим, какие условия на последовательности $\{x_i(k)\}$, $\{\hat{x}_i(k)\}$, $\{\mu_i(k)\}$ надо наложить, чтобы оптимальные значения a , b , полученные при минимизации функционала (117) с возмущенными данными (118), были непрерывны по возмущениям μ . Здесь $\{\mu_i(k)\}$ — последовательность реализаций случайного возмущения μ .

Вычисляя частную производную от (117) по a , получим:

$$N\Delta a = \sum_{k=0}^{N-1} \ln \frac{x(k+1)}{x(k)} - \Delta b \sum_{k=0}^{N-1} x(k).$$

Затем, вычисляя частную производную от (117) по b и подставляя найденное значение a , определим:

$$b = \frac{1}{\Delta} \frac{\sum x(k) \ln \frac{x(k+1)}{x(k)} - \frac{1}{N} \sum x(k) \sum \ln \frac{x(k+1)}{x(k)}}{\sum x^2(k) - \frac{1}{N} (\sum x(k))^2}.$$

Здесь все суммы по k от 0 до $N - 1$. В знаменателе последней формулы находится оценка среднеквадратичного отклонения $\{x(k)\}$ от его среднего значения, умноженного на длину выборки. Поэтому условия корректности последних формул состоят из неравенств:

$$x(k) > 0, \quad \sum (x(k) - \bar{x})^2 > 0, \quad (119)$$

где \bar{x} — среднее выборки $\{x(k)\}$.

Первое из неравенств (119) легко проверить перед обработкой данных, поэтому его можно не интерпретировать. Второе должно выполняться всегда в силу предположения (118) и того, что данные $\{\hat{x}(k)\}$ удовлетворяют уравнению Лотки–Вольтерры (и не соответствуют положению равновесия), а при достаточно малой дисперсии случайной ошибки m последовательность $\{x(k)\}$ также будет удовлетворять второму неравенству в (119).

Теперь осталось заметить, что, подставляя (118) в выражения для a и b , получим непрерывные функции μ (при достаточно малых μ).

Следовательно, если $\mu \rightarrow 0$, то везде $x(k)$ заменится на $\hat{x}(k)$, а для последних неравенства (119) обязательно имеют место.

В многомерном случае к условиям (119) добавляется условие невырожденности матрицы

$$\sum_{k=0}^{N-1} Z_k Z_k^T.$$

А это, в свою очередь, сводится к существованию среди N векторов Z_k ровно $(n + 1)$ линейно независимых, что и необходимо потребовать от многомерных данных для идентификации модели.

3.5.3. ИДЕНТИФИКАЦИЯ МОДЕЛИ С ЗАПАЗДЫВАНИЕМ

Так как на протяжении всей книги в сферу моделирования динамических процессов включаются уравнения с запаздывающим аргументом, то естественно расширить алгоритм раздела 2.3.4, в котором предлагается способ идентификации линейной системы с запаздыванием, на нелинейные модели Лотки–Вольтерры. Подход к проблеме одновременной идентификации величины запаздывания и коэффициентов системы сохраним прежний: для каждого дискретного запаздывания из определенного интервала получим оптимальные коэффициенты, а затем из конечного множества систем выберем ту, для которой значение функционала качества будет наименьшим.

Пусть $x_i(k), i = 1, \dots, n; k = 0, 1, \dots, N$ — наблюдения за состоянием системы (101) в дискретные моменты времени $t = k\Delta$. Будем считать временной лаг кратным интервалу наблюдения: $\tau = l\Delta$. Тогда из системы (101) следует, что

$$x_i(k+1) = x_i(k) \exp \int_{k\Delta}^{(k+1)\Delta} \left[\varepsilon_i - \sum_{j=1}^n \gamma_{ij} x_j(s-\tau) ds \right].$$

Если прологарифмировать последнее равенство, заменить переменную интегрирования и аппроксимировать интеграл по формуле трапеции, то получим приближенное (с точностью, пропорциональной Δ) равенство

$$\ln x_i(k+1) \approx \ln x_i(k) + \Delta \left[\varepsilon_i - \frac{1}{2} \sum_{j=1}^n \gamma_{ij} (x_j(k-l) + x_j(k+1-l)) \right]$$

при $i = 1, \dots, n; k = l, l+1, \dots, N-1$.

Теперь можно выписать функционал, минимизация которого будет приближать нас к решению задачи идентификации:

$$F_l(E, \Gamma) = \sum_{k=l}^{N-1} \sum_{i=1}^n \left\{ \ln x_i(k+1) - \ln x_i(k) - \Delta \left[\varepsilon_i - \frac{1}{2} \sum_{j=1}^n \gamma_{ij} (x_j(k-l) + x_j(k+1-l)) \right] \right\}^2,$$

где $l = 1, \dots, N-2$.

Алгоритм подбора тренда векторного временного ряда в виде решения соответствующей системы уравнений Лотки–Вольтерры заключается в следующем: для каждого $l = 1, \dots, N - 2$ делаем замену переменных:

$$D = E\Delta, \quad G = -\Gamma\Delta, \quad Y_k = \{\ln x_i(k+1) - \ln x_i(k)\}_{i=1}^n,$$

$$Z_k = \left\{ \begin{array}{c} \frac{1}{2}(X(k-l) + X(k+1-l)) \\ 1 \end{array} \right\},$$

и решаем задачу минимизации функционала:

$$F_l(D, G) = \frac{1}{N-l} \sum_{k=l}^{N-1} (Y_k - \{G, D\}Z_k)^2.$$

Получаем в результате $E_l = (1/\Delta)D_l$, $\Gamma_l = (1/\Delta)G_l$, и из этого конечного множества параметров выбираем номер l , для которого функционал $F_l(E_l, \Gamma_l)$ принимает наименьшее значение.

3.6. ЗАКЛЮЧЕНИЕ

В этой главе рассмотрена многомерная модель Лотки–Вольтерры с временным лагом. Выделен класс экономических задач, динамика в которых может быть описана такой моделью. В данный круг проблем попадают все виды взаимодействия предприятий на общем экономическом поле: в производстве и распределении товаров, причем как в конкурентной борьбе, так и в случае зависимости одного предприятия от функционирования другого. Аналогично обстоят дела с взаимодействием отраслей, регионов и стран. Поэтому рассматриваемая модель может оказаться полезной при анализе динамики внешней торговли или цикличности экономического развития стран.

Использование нелинейных уравнений открывает путь для формирования в пространстве переменных нескольких областей притяжения (или отталкивания), а введение запаздывания является не только естественным в экономических задачах, но и существенно обогащает динамические свойства модели, делает предметом исследования как простейшие периодические колебания, так и почти периоди-

ческие, и рекуррентные движения. В главе дано подробное изучение таких динамических свойств в многомерном случае. Частью аналитически, частью с привлечением численных методов показано, как выглядят такие сложные движения и как изменяются их характеристики при возрастании величины временного лага.

Предложение шире использовать столь сложные модели повисло бы в воздухе, если бы не конкретные алгоритмы идентификации коэффициентов модели. Предложенный численный метод вполне обозримый и дает достаточно адекватную математическую модель для динамических процессов, чьи движущие причины могут быть отнесены к конкуренции за общие жизненные ресурсы.

Внедрение в экономическую практику описанных выше моделей помогло бы сделать более точным прогноз, разобраться в причинах тех или иных явлений, а также предпринять соответствующие управленческие шаги для изменения ситуации.

ГЛАВА ЧЕТВЕРТАЯ

РАЗЛИЧНЫЕ ДИНАМИЧЕСКИЕ МОДЕЛИ

В этой главе собраны несколько не связанных между собой задач экономики. Все они, однако, дополняют примеры первых двух глав, еще раз демонстрируя технику линейного и нелинейного анализа, а также иллюстрируют разнообразие возможных подходов к моделированию динамических процессов. Открывает изложение задача прогнозирования цены фьючерса. На примере данных об экономическом состоянии нескольких стран выбирается наилучший в некотором смысле алгоритм прогнозирования на длительный интервал времени. Используются методы, приведенные в первой главе. Второй параграф посвящен динамическим и статическим задачам в международной торговле. Рассматривается одна модель оптимального протекционистского тарифа. Далее к задачам оценки реакции рекламы на продажу нового товара применяются необычные модели. Завершает главу применение моделей Лотки–Вольтерры к динамике продаж на внутреннем автомобильном рынке США. На базе реальных данных о продажах определяются временные лаги и остальные параметры конкурентной модели.

4.1.

ПРОГНОЗИРОВАНИЕ КУРСОВ ВАЛЮТ

Ниже будут проанализированы обменные курсы трех стран (Италии, Англии и Германии) по отношению к США. Рассмотрены различные вычислительные процедуры, различные критерии выбора метода и исследована длина ин-

тервала идентификации модели. Цель прогноза — помощь при заключении форвардных сделок на покупку иностранной валюты, поэтому интервал прогнозирования — один год. Данные для расчетов получены автором в совместной работе с К. Вэйем над докладом [116].

4.1.1. ФОРМУЛИРОВКА ПРОБЛЕМЫ

Предположим, что некоторая транснациональная промышленная компания имеет свои производственные филиалы в других странах, например, это может быть шведская компания «Электролюкс» или немецкая «Хенкель». Для большей определенности (а точнее, из-за возможности получить статистические данные) мы остановим свой выбор на некоторой крупной американской компании с производственными филиалами в Италии, Англии и Германии. В процессе анализа алгоритмов будем рассматривать покупку итальянской лиры на доллары, а результирующие утверждения отнесем и к английскому фунту, и к немецкой марке.

Итак, предположим, что компания, планируя свои финансовые операции на год вперед, хочет купить один миллиард итальянских лир в конце года. В начале года еще неизвестно, каким будет обменный курс в конце года. Неизвестно также годовое значение ставки по безрисковым государственным ценным бумагам (*treasury bonds*), которое определяет так называемый дискант-фактор, используемый при назначении форвардной цены. Компания обращается в коммерческий банк или другой соответствующий финансовый институт для заключения форвардной сделки на год вперед на покупку миллиарда лир. Банк каким-либо образом прогнозирует обменный курс на конец года и сообщает свой прогноз компании. Так как она может полученную в результате прогноза сумму разместить на год в различные ценные бумаги, спекулятивные операции, дать в кредит или инвестировать в какой-либо проект, что принесет прибыль, но и подвержено риску, то компания вычисляет возможную прибыль за год от прогнозируемой суммы с самым малым риском. Очевидно,

существует еще и плата банку за выполнение операции, но это фиксированная величина и она не нуждается в анализе в рамках этой задачи.

Принимая во внимание дисконт-фактор, выраженный через процентную ставку по безрисковым государственным облигациям США, получаем формулу форвардной цены:

$$f(t) = l(t+T) \exp \left[- \int_t^{t+T} r(s) ds \right], \quad (120)$$

где $l(t)$, $r(t)$ — цена лиры в американских долларах в момент времени t и процентная ставка соответственно; T — срок форвардного контракта (в нашем случае это один год). Для того чтобы воспользоваться формулой (120), необходимо знать прогноз цены лиры $l(t+T)$ и процентную ставку $r(t)$ на всем интервале $[t, t+T]$. Таким образом, перед нами задача экстраполяции: по данным на интервале идентификации $[t-h, t]$ построить прогноз на интервал $[t, t+T]$ для функций $l(t)$, $r(t)$. Здесь h — длина интервала идентификации.

Предположим, что банк имеет всю необходимую для нашего анализа информацию. Сначала для определенности в формулах будем считать длину интервала идентификации равной двум годам. Позже мы обсудим оптимизацию по этой длине.

Выражение (120) не предусматривает никаких гарантий или компенсаций банку за риск ошибочного прогноза. Более подробно это выглядит так. Функция $f(t)$ для любых t является случайной величиной, так как l , r — случайные. Мы можем остановиться сначала на модели стационарного случайного процесса, поскольку идентифицировать более общую функцию распределения все равно не удастся. Разброс значений вокруг соответствующего математического ожидания дает банку основание увеличить цену форвардной сделки на величину риска, которая также вычисляется в виде прогноза. Значит, целью прогнозной модели являются две случайные величины: тренд как следствие выражения (120) и ширина полосы 90%-ного разброса, соответствующего 90%-ному доверительному интервалу. Можно брать и другие условия гарантии, например, если взять 50%-ный доверительный интервал, то,

конечно, плата за риск будет меньше, но вряд ли банк согласится на сделки, положительный исход которых может равным образом чередоваться с отрицательным.

В нашем распоряжении имеются следующие данные:

- валовой внутренний продукт (ВВП), номинальный и реальный (поквартально);
- процентная ставка безрисковых государственных облигаций с учетом инфляции и без таковой (помесячно);
- индекс цен на потребительские товары (для США) и индекс стоимости жизни (для Италии, Англии и Германии) (помесячно);
- курсовая стоимость соответствующих валют в долларах США (помесячно)

для упомянутых четырех стран с 01.01.76 по 03.01.98.

Эти экономические параметры были выбраны, так как изначально предполагалось, что обменный курс валюты является следствием макроэкономических процессов в странах, чьи валюты участвуют в обмене. Так, ВВП характеризует значимость страны в мировом экономическом порядке и внутреннее текущее состояние производства; процентная ставка государственных ценных бумаг коррелирует со ставкой рефинансирования Центрального банка, а последняя является одним из основных рычагов государственного управления (как в кейнсеанской теории, так и в монетаризме) деловой активностью (спросом на деньги). Индексы жизни определяют потребительские свойства общества и тесно связаны с инфляцией. Предполагается, что выбранные переменные экономики страны достаточно полно характеризуют ее динамику и, следовательно, определяют обменный курс валют. Мы не рассматриваем платежный баланс экспортно-импортного взаимодействия стран, поскольку сами страны выбраны произвольно: они все принадлежат группе промышленно развитых стран с высоким жизненным уровнем и большим влиянием на мировую политику. В этих странах уже много лет ведется открытая (до известных пределов) внешнеторговая политика, происходит перемешивание капиталов и трудовых ресурсов. В такой ситуации платежный баланс играет меньшую роль, чем мобильность факторов производства.

4.1.2. КРИТЕРИИ КАЧЕСТВА МЕТОДОВ ПРОГНОЗА

Прежде чем приступить к обсуждению различных методов прогнозирования, важно сформулировать критерий, с помощью которого можно оценивать качество модели. В первой главе объяснялось, что такое адекватность модели. Адекватных моделей может быть много, и в этой ситуации необходима дополнительная оценка качества. Если, например, прогноз требует большого набора данных, а они стоят дорого, то предпочтение будет отдано более дешевому пути. Или при необходимости очень быстро принимать решение (в реальном масштабе времени на энергетических установках) предпочтение будет отдано более быстрому из двух одинаково точных. Также скорость ввода и обработки данных важны на биржевых торгах. В нашем случае упомянутые ограничения не играют никакой роли, и мы можем все внимание уделить точности прогнозирования. Но и на этом пути возможны различные формы функционала качества. Будем исходить из того, что нас интересует наилучшее приближение конечного результата — цены форвардного контракта, поэтому рассмотрим сумму ошибок расчетной и истинной цены по всем возможным данным:

$$FPC = \frac{1}{N-T-h} \sum_{t_0=h}^{N-T} \left[l(t_0+T) \exp \left[- \int_{t_0}^{t_0+T} r(t) dt \right] - \bar{l}(t_0+T) \exp \left[- \int_{t_0}^{t_0+T} \bar{r}(t) dt \right] \right]. \quad (121)$$

Здесь $\bar{l}(t)$, $\bar{r}(t)$ обозначены расчетные или модельные значения стоимости лиры и процентной ставки в отличие от не помеченных l , r . Так как компания, о которой идет речь, американская, то $r(t)$, $\bar{r}(t)$ в формуле (121) касаются США. N — это длина анализируемых данных, в нашем случае это 21 год. Если, переходя к дискретному времени (и заменяя интегралы соответствующими суммами), в качестве единицы выберем один месяц, то $N = 252$. Если алгоритм потребует трехлетних данных для одного вы-

числительного цикла ($T = 12, h = 24$), то есть возможность получить среднюю ошибку по 216 экспериментам прогноза, это достаточное число для того, чтобы поверить в результат. Функционал (121) дает нам величину ошибки (по всем анализируемым данным) цены фьючерса. Теоретически *FPC* (Futures Price Criterion) является оценкой случайной величины с некоторой функцией распределения, но как первый шаг исследования достаточно принимать во внимание только среднее значение модуля ошибки для всех методов прогнозирования. Практически величина *FPC* означает, что если для данного метода прогнозирования лиры в долларах $FPC = 0,0001$, то компания в среднем (за много раз повторенный контракт) теряет 0,01 цента за одну лиру, или на каждый миллиард покупаемых лир потери составят 100 000 \$. Если другой метод имеет значение $FPC = 0,00001$, то потери на таком же форвардном контракте составят 10 000 \$.

Критерий (121) хорош для оценки анализируемого метода как для выбора одного из нескольких методов, так и для оптимизации по длине интервала идентификации, хотя последняя задача приведет к рассмотрению и других особенностей методов. Но если мы хотим учесть оплату риска, то необходимо получить функцию распределения (точнее, оценку функции распределения) разницы между действительной и модельной ценой:

$$\delta(t) = l(t+T) \exp \left[- \int_t^{t+T} r(t) dt \right] - \bar{l}(t+T) \exp \left[- \int_t^{t+T} \bar{r}(t) dt \right],$$

где $t \in [h, N - T]$. Предполагается, что стохастический процесс $\delta(t)$ является стационарным, т. е. функция распределения $\delta(t)$ при любом t постоянна. В этом случае переменную t можно рассматривать как номер независимого эксперимента (это, конечно, очень сильное предположение для экономических взаимоотношений), который производится $(N - T - h)$ раз.

Таким образом, окончательный выбор метода и длины интервала идентификации делается из анализа 90%-ного доверительного интервала случайной величины $\delta(t)$.

4.1.3. ОБСУЖДЕНИЕ МЕТОДОВ ПРОГНОЗИРОВАНИЯ

Для всех методов в этом разделе идентификационный интервал предполагается равным 24 месяцам, т. е. $h = 24$. Мы также постараемся для каждого метода сформулировать ясные ограничения и предположения, на которых метод основан.

4.1.3.1. НЕПОСРЕДСТВЕННАЯ АППРОКСИМАЦИЯ ВРЕМЕННЫХ РЯДОВ

Простейший алгоритм прогнозирования временных рядов, как подробно излагалось в первой главе, состоит в подборе линейной комбинации функций из определенного набора. Предположим, что чуть-чуть сглаженный (например, с помощью скользящего среднего) временной ряд похож на комбинацию колебаний с высокими и низкими частотами, наложенных на некоторый полином. Тогда следующие функции могут быть использованы как базовые:

$$\{1, t, t^2, \dots, \sin \omega_1 t, \cos \omega_1 t, \sin \omega_2 t, \cos \omega_2 t, \dots\}. \quad (122)$$

Здесь многоточия означают, что последовательности функций могут быть продолжены. Далее мы будем приводить формулы только с частью возможных функций. Итак, приближение, например процентной ставки $r(t)$, записывается в виде:

$$\bar{r}(t) = a_0 + a_1 t + a_2 t^2 + a_3 \sin \omega t + a_4 \cos \omega t. \quad (123)$$

Согласно объяснению в параграфе 1.2, если предположить, что

$$r(t) = \bar{r}(t) + \varepsilon,$$

где ε — случайная величина с постоянной функцией распределения и нулевым математическим ожиданием, тогда оценка, получаемая методом наименьших квадратов, будет несмещенной и с минимальной дисперсией.

Так как коэффициенты a_0, a_1, a_2, a_3, a_4 входят в (123) линейно, то легко для них построить систему уравнений: составим сумму квадратов отклонений:

$$S(a_0, \dots, a_4) = \sum_{k=0}^{23} (r_{k+1} - a_0 - a_1 k - a_2 k^2 - a_3 \sin \omega k - a_4 \cos \omega k)^2 \quad (124)$$

и решим относительно a_0, a_1, a_2, a_3, a_4 алгебраическую систему:

$$\partial S / (\partial a_0) = 0, \quad \partial S / (\partial a_1) = 0, \quad \dots, \quad \partial S / (\partial a_4) = 0.$$

Замечание 1. Частота колебаний ω не входит в выражение для S так же, как коэффициенты a во второй степени. Поэтому ее неудобно рассматривать как одну из переменных последней системы. Чтобы избежать этого, в вычислительной процедуре предусматривается простой перебор некоторого очевидного интервала, в котором ω может находиться, и для каждого значения частоты вычисляются оптимальные относительно функционала (124) коэффициенты a . Из всех полученных таким образом точек выбирается та, в которой S принимает наименьшее значение. Это не дает точного решения задачи нахождения экстремума, но зато сама собой исчезает проблема большого числа локальных экстремумов, которая имела бы место при решении нелинейной задачи.

Замечание 2. Метод наименьших квадратов имеет дополнительные возможности, которые мы здесь не использовали. Например, введение весовых коэффициентов перед квадратами в сумме дает некоторое ранжирование значимости данных. Так это позволяет делать более удаленные данные временного ряда от начальной точки прогноза менее значимыми. В данной работе оптимизация по весовым коэффициентам не производилась.

Основная проблема в такого рода приближениях заключается в противоборстве желания применить полиномиальные функции, поскольку они хорошо изучены и дают сколь угодно точный результат в задаче интерполирования (например, ортогональное семейство полиномов Чебышева), с неустойчивым поведением полиномов вне интервала идентификации. Этот факт установлен в общем виде в параграфе 1.2.4.

На рисунке 24 показан результат приближения стоимости доллара в лирах на протяжении 21 года графиком функции вида (123). Весь промежуток наблюдения разбит специальными разделителями (селекторами) на отрезки: двухлетний — идентификации и однолетний — прогноза.

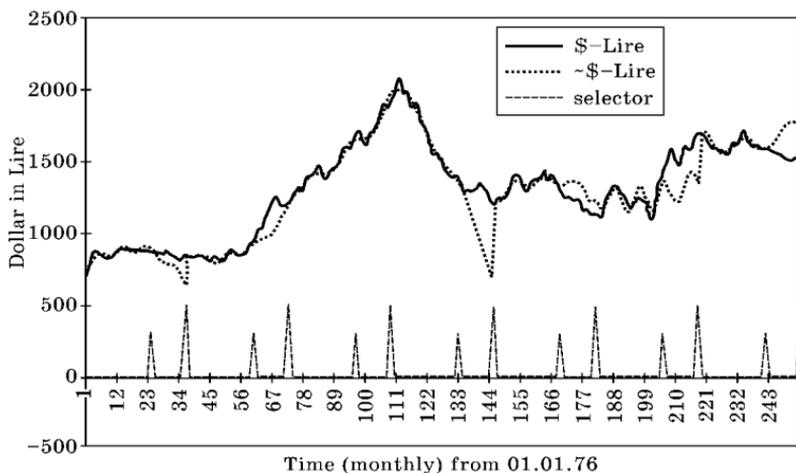


Рис. 24

Приближение цены 1 доллара США в лирах с помощью квадратичного полинома и одного колебания

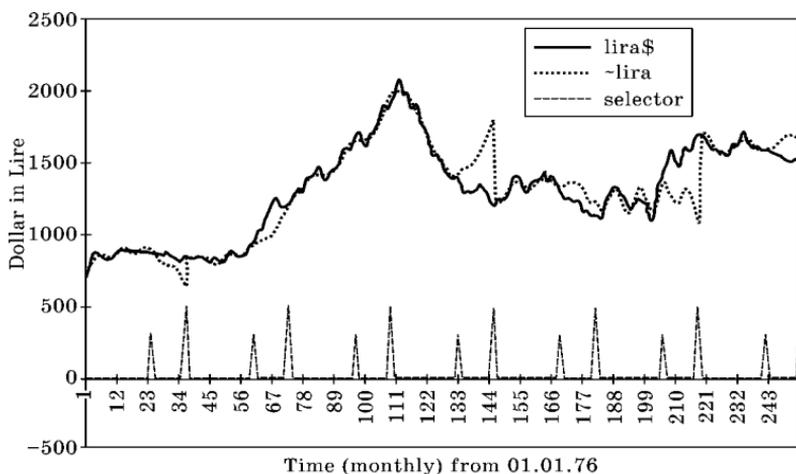


Рис. 25

Приближение цены 1 доллара США в лирах с помощью комбинации двух колебаний

Так как полином используется на монотонных участках, то чтобы устранить резкий рост полинома, предлагается добавить в функцию медленно меняющийся синус (и косинус), т. е. низкочастотное колебание. Результат показан на рисунке 25.

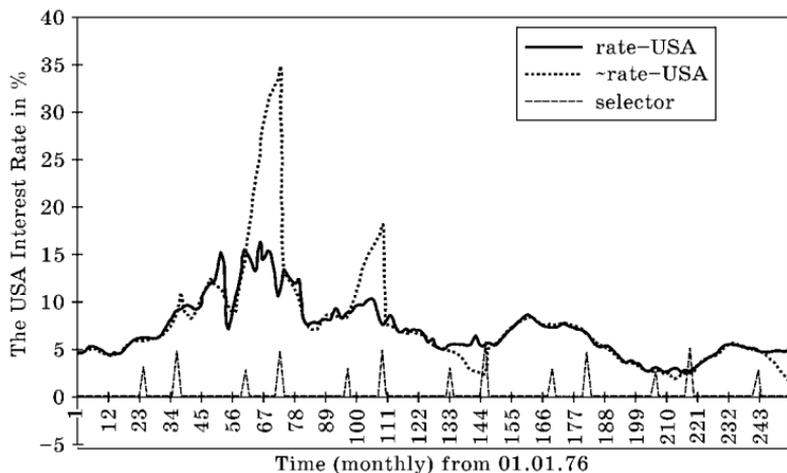


Рис. 26

Приближение процентной ставки США с помощью квадратичного полинома и одного колебания

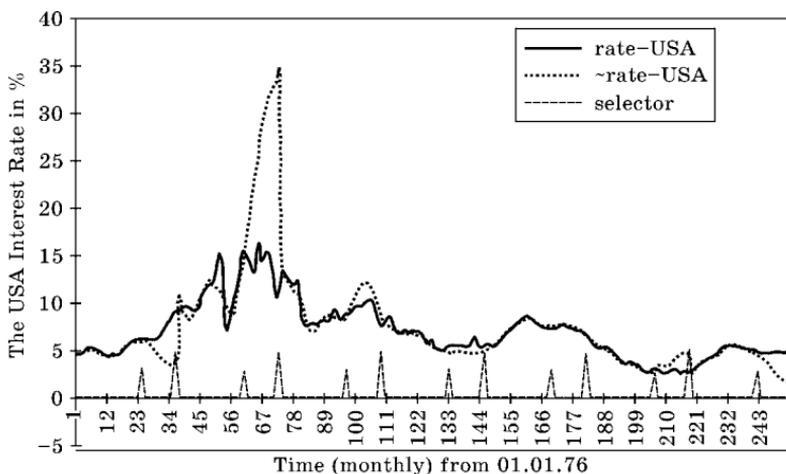


Рис. 27

Приближение процентной ставки США с помощью двух колебаний

Улучшение приближения более наглядно представлено на других данных, а именно, на процентной ставке в США, как это показано на рисунке 26 (хотя большие отклонения прогноза все равно имеют место). Быстрый рост отклонений может быть уменьшен применением двух колеблющихся

функций, за исключением периода между 1979 и 1982 годами. Из графика (см. рис. 27) видно, что только информация рассматриваемых данных не способна стать основанием более точного предсказания: на промежутке идентификации данные колеблются вокруг возрастающего среднего с возрастающей амплитудой (что и демонстрирует модель как прогноз), но далее среднее начинает убывать, а колебания — уменьшать амплитуду, что никак не следует из данных интервала идентификации. Применение нескольких колебаний лишь отчасти решает проблему: просто рост модельной функции на интервале прогноза меньше.

Приведенные выше примеры показывают необходимость использования интегрального критерия FPC (или аналогичного ему), чтобы избежать субъективных визуальных оценок графиков данных, особенно потому что FPC содержит комбинацию сразу двух случайных величин. Вычисления дают следующий результат. Класс регрессионных функций, состоящий из квадратичного полинома и одной периодической функции соответствует $FPC = 0,0002$, а в случае двух периодических функций — $FPC = 0,00013$, т. е. второй способ приближения является более точным. Таким образом, покупая один миллиард лир, компания на годовом прогнозе имела бы средние потери 200 000 \$ при первом методе (с полиномом) и 130 000 \$ — используя второй путь (два колебания). Очевидно, годовая разница в 70 000 \$ стоит того, чтобы отказаться от полиномов.

4.1.3.2.

МЕТОД ОБЩЕЙ КОРРЕЛЯЦИИ

Под названием метод общей корреляции мы имеем в виду построение векторной линейной авторегрессии по всем данным, имеющим отношение к обменному курсу. Оно более сложно в алгоритмическом плане, чем прямой непосредственный анализ скалярных временных рядов, но зато объединение макроэкономических параметров в одной системе создает возможность проявления более тонких экономических влияний одних данных на другие.

Рассмотрим вектор, состоящий из семи макроэкономических параметров: ВВП США, процентная ставка США,

индекс цен на потребительские товары США, ВВП Италии, процентная ставка в Италии, индекс стоимости жизни в Италии, курсовая стоимость лиры в долларах:

$$X_k \equiv \{GDP-US, Rate-US, CPI-US, GDP-It, Rate-It, CPI-It, Lira\}_k,$$

где $k = 1, \dots, 24$. Наша задача состоит в получении последнего элемента вектора X_{36} и среднего интегрального от второго элемента вектора X_k , $k = 25, \dots, 36$ по значениям X_k , $k = 1, \dots, 24$. Этих приближений достаточно, чтобы вычислить цену форвардного контракта по формуле (120). Модель строится в виде линейной разностной системы как простейшей формы зависимости:

$$X_{k+1} = AX_k + B. \tag{125}$$

Матрица A и вектор B могут быть найдены с помощью метода наименьших квадратов и функционала:

$$S(A, B) = \sum_{k=1}^{23} (X_{k+1} - AX_k - B)^2. \tag{126}$$

Как было описано в параграфе 2.3, функционал (126) достигает своего минимума при следующих значениях параметров:

$$\{A, B\} = \{X_2, X_3, \dots, X_{24}\} \begin{Bmatrix} X_1 & X_2 & \dots & X_{22} & X_{23} \\ 1 & 1 & \dots & 1 & 1 \end{Bmatrix}^+,$$

где верхний индекс «+» означает псевдообратную матрицу.

Ниже мы повторим некоторые формулы, полученные во второй главе. Обозначая (8×23) -матрицу:

$$D = \begin{Bmatrix} X_1 & X_2 & \dots & X_{22} & X_{23} \\ 1 & 1 & \dots & 1 & 1 \end{Bmatrix},$$

получим ее скелетное разложение на произведение двух других матриц, размерами $8 \times r$ и $r \times 23$:

$$D = GH = \begin{pmatrix} g_{11} & \dots & g_{1r} \\ g_{21} & \dots & g_{2r} \\ \dots & \dots & \dots \\ g_{81} & \dots & g_{8r} \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1,23} \\ \dots & \dots & \dots & \dots \\ h_{r1} & h_{r2} & \dots & h_{r,23} \end{pmatrix},$$

где $r = \text{rank } D$. Чтобы получить скелетное разложение, достаточно в качестве столбцов матрицы G взять любые r линейно не зависимые столбцы матрицы D . В этом случае любой столбец матрицы D является линейной комбинацией столбцов матрицы G с коэффициентами $h_{1j}, h_{2j}, \dots, h_{rj}$. Эти коэффициенты составляют соответствующий столбец матрицы H .

Следующая формула в указанных обозначениях является основной для алгоритма вычисления неизвестных параметров системы (125):

$$D^+ = H^+G^+ = H^T(HH^T)^{-1}(G^TG)^{-1}G^T.$$

Однако решение системы (125) определяется некоторым вектором начальных данных и сейчас возникает проблема, какой из векторов мог бы претендовать на роль начального. Это мог бы быть последний вектор наблюдений, или первый, или средний в каком-либо смысле. Но в нашей задаче все векторы X_k имеют случайные элементы, поэтому не остается никаких оснований выбирать начальный вектор из множества идентификации. Это наталкивает на мысль продолжить оптимизацию среднеквадратичных отклонений при фиксированных A, B и неизвестном начальном векторе Z .

Частное решение системы (125) записывается в форме

$$X_{k+1} = A^k Z + A^{k-1} B + \dots + AB + B,$$

где Z — начальный вектор, и функционал отклонений на этом решении примет вид:

$$S(Z) = \sum_{k=1}^{24} (X_k - A^{k-1} Z - A^{k-2} B - \dots - AB - B)^2.$$

Все члены внутри скобок, не зависящие от Z , обозначим как Y . Тогда наименьшее значение функционала достигается при

$$Z = \left(\sum_{k=0}^{23} (A^T)^k A^k \right)^{-1} \left(\sum_{k=0}^{23} A^k Y_{k+1} \right).$$

Теперь алгоритм готов к использованию.

Прокомментируем результат предлагаемого моделирования на примере трехлетнего набора данных, начиная

с 01.01.76. Все данные были трансформированы так, чтобы в вычислениях участвовали величины одного с единицей порядка:

$$G_A = \frac{GDP_A}{10000}; \quad R_A, CPI_A, R_I, CPI_I = \frac{R_A, CPI_A, R_I, CPI_A}{100};$$

$$G_I = \frac{GDP_I}{1000}; \quad L = \frac{1000}{Lira}.$$

Здесь обозначения соответствуют аббревиатуре английских названий рассматриваемых переменных, индекс обозначает страну.

После идентификации модели (125) была получена следующая разностная система (значения коэффициентов округлялись):

$$\begin{pmatrix} G_A \\ R_A \\ C_A \\ G_I \\ R_I \\ C_I \\ L \end{pmatrix}_{k+1} = \begin{pmatrix} 0,92 & -0,16 & 0,09 & 0,3 & -0,01 & -0,01 & 0,0 \\ 0,07 & 0,74 & 0,04 & -0,4 & -0,05 & 0,005 & -0,02 \\ -0,05 & -0,27 & 0,78 & -0,1 & -0,1 & 0,03 & -0,02 \\ -0,04 & 0,09 & -0,02 & 0,94 & -0,004 & -0,007 & -0,002 \\ -2,9 & -0,9 & 1,2 & 2,1 & 0,03 & 0,08 & -0,1 \\ 2,0 & -2,4 & -2,2 & 0,1 & 0,41 & 0,24 & -0,3 \\ -0,7 & 4,0 & 1,0 & 0,6 & 0,7 & 0,25 & 0,63 \end{pmatrix} \times$$

$$\begin{pmatrix} G_A \\ R_A \\ C_A \\ G_I \\ R_I \\ C_I \\ L \end{pmatrix}_k + \begin{pmatrix} -0,02 \\ 0,1 \\ 0,1 \\ 0,03 \\ 1,0 \\ -0,5 \\ 0,17 \end{pmatrix}.$$

Оптимальное значение начального вектора для этого примера имеет вид:

$$\{0,41 \ 0,045 \ 0,07 \ 0,22 \ 0,16 \ 0,11 \ 1,43\}.$$

Подставляя элементы начального вектора в полученную разностную систему, можно оценить вклад каждого фактора в изменение соответствующих переменных на каждом шаге.

G_A — реальный ВВП США имеет сильное влияние на все другие переменные, но в то же самое время его собственное изменение в наибольшей степени зависит только от него самого. Это вполне вероятно, так как ВВП США в двадцать раз превосходит ВВП Италии и к тому же экономика США разнообразна и самодостаточна настолько, чтобы небольшие изменения деловой активности и потребления не влияли существенно на изменение ВВП. Зависимость других промышленно развитых стран от ВВП США объясняется большой политической и экономической зависимостью, что можно условно назвать зоной влияния доллара. Если, скажем, правительство США решит изъять из обращения часть долларовой массы с помощью увеличения процентной ставки, то это затронет не только американцев, но и всех, кто использует доллары в экономической деятельности, т. е. все рассматриваемые страны.

R_A — процент выплат по государственным казначейским облигациям определяет спрос на деньги внутри США. Так как доллар в известном смысле является сильной мировой валютой, то большое число долларов участвует в мировой экономике. Поэтому R_A , G_I , G_A в большой степени зависят от R_A .

C_A — эта переменная, связанная инфляционными процессами в Америке, зависит в большей степени от себя самой и ВВП США.

G_I — оказывается (согласно модели), что на изменение ВВП Италии наибольшее влияние создается самим ВВП Италии, а также ВВП США и процентной ставкой США (т. е. спросом на деньги с США).

R_I — эта переменная очень сильно зависит от ВВП США, ВВП Италии и обменного курса лиры и доллара.

C_I — здесь удивительный факт: индекс жизненного уровня в Италии в большей степени зависит от экономики Америки, т. е. от L , G_A , R_A , и в меньшей степени от ВВП Италии. Возможно, что инфляционные ожидания итальянских потребителей чутко реагируют на изменения мировой (и, в особенности, американской) экономики, поскольку собственные макропараметры имеют поведение

с большими отклонениями от среднего уровня, т. е. неустойчивое в широком понимании этого слова.

L — обменный курс (точнее, после преобразований, сделанных выше, стоимость 1000 лир в долларах США) также сильно связан с G_A и R_A , но в меньшей степени с G_I и R_I .

Результаты, полученные с использованием метода общей корреляции, представлены на рисунках 28 и 29.

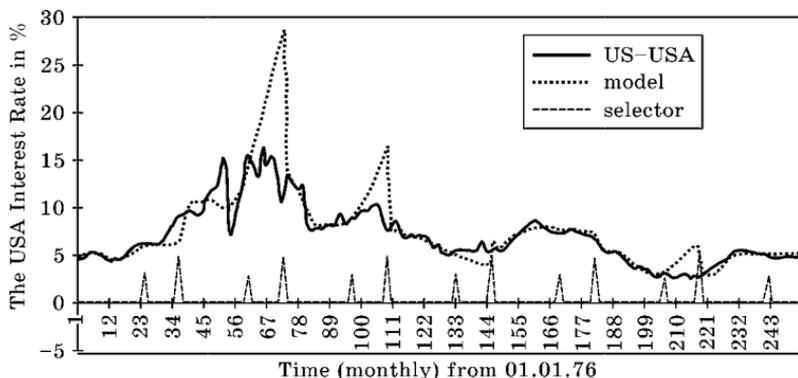


Рис. 28
 Модель общей корреляции переменных.
 Результат для процентной ставки США

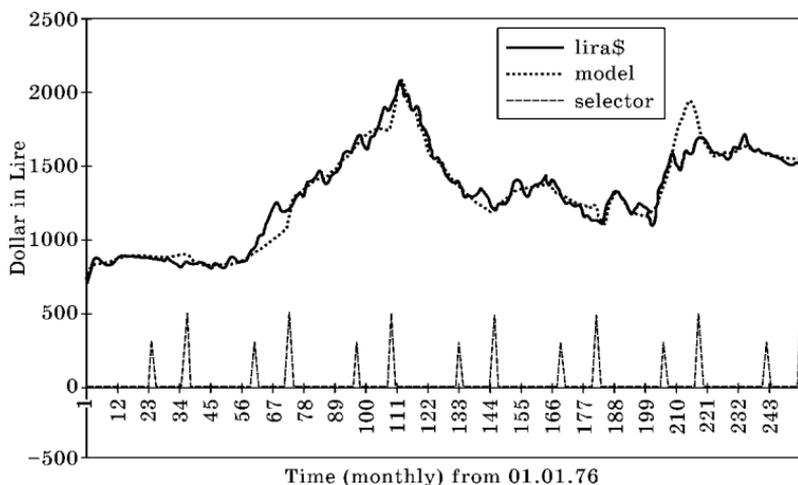


Рис. 29
 Модель общей корреляции переменных.
 Результат для стоимости 1 доллара США в лирах

Эти графики отражают способность метода общей корреляции автоматически выявить иерархические зависимости между макропараметрами и их общее влияние на обменный курс.

Давайте проверим по критерию FPC , как этот метод соотносится с прямым моделированием временных рядов, которое рассмотрено в предыдущем разделе. Используя весь набор данных, получим среднюю ошибку прогноза $FPC = 0,00015$. Это почти совпадает с двухчастотным экстраполизированием. Результат в данном случае был предсказуем, так как разностная система седьмого порядка имеет характеристическое уравнение седьмого порядка, и, следовательно, динамика решения может описываться, в лучшем случае, трехчастотными колебаниями (т. е. более разнообразный класс функций, чем при прямой экстраполяции). Но объединение семи экономических переменных в одну систему потребовало потратить появившиеся свободы выбора на удовлетворение особенностей всех переменных. Так, оба ВВП имеют монотонно возрастающие составляющие, в то время как остальные переменные регулируются правительствами (и центральными банками) с целью удержания в определенных пределах. Появление двух монотонных составляющих у решения системы (125) влечет существование только двух пар комплексно-сопряженных корней соответствующего характеристического уравнения седьмой степени, т. е. существование только двухчастотных колебаний.

Возможно некоторое улучшение процедуры моделирования в смысле оптимизации преобразований переменных, сделанных выше. Это обуславливается тем, что задачи большой размерности могут иметь чисто вычислительные проблемы. В качестве функционала такой оптимизации можно предложить сумму отклонений модулей корней матрицы A системы (125) от единицы. Фактически данные в рассматриваемых нами задачах таковы, что линейные вычислительные процедуры не являются жесткими. Можно предположить, что распространение соответствующих оптимальных преобразований на другие валюты (фунты и марки) невозможно и, следовательно, необходимо подстраивать алгоритм под разные валюты.

4.1.4. ОПТИМИЗАЦИЯ ДЛИНЫ ИНТЕРВАЛА

Рассмотрим далее длину интервала идентификации как параметр методов, т. е. при таком взгляде FPC будет функцией длины интервала. Вычисления по всем данным с различными интервалами идентификации дают следующие результаты:

- первый прямой метод, использующий полином и одно колебание, имеет оптимальный интервал в смысле наименьшего FPC , равный 30 месяцам с $FPC = 0,00016$;
- второй прямой метод, использующий два колебания, не имеет четко выраженного экстремума FPC по длине интервала. Значения FPC слабо меняются (около 0,000134), в то время как длина интервала изменяется от 22 месяцев до 28. Так что двухлетний интервал идентификации можно считать вполне обоснованным;
- метод общей корреляции имеет четкий минимум $FPC = 0,000149$ при 30-месячном интервале.

К сожалению, этот рассмотренный нами критерий не позволяет абсолютно уверенно выделить лучший метод. Следовательно, мы должны ввести нечто более общее, позволяющее распознать детали второго порядка моделирования. Так как нас все равно будет интересовать функция распределения ошибки, давайте в дополнение к FPC примем во внимание 90%-ные доверительные границы ошибки для каждого фиксированного интервала идентификации. Следующие три графика (см. рис. 30) показывают зависимости 90%-ного доверительного интервала от продолжительности времени идентификации по трем обсуждаемым методам для стоимости 1000 лир в долларах США (точнее, не для самой цены, а для разности реальной и модельной цен).

Область между верхним и нижним пределами и есть 90%-ная доверительная полоса для ошибки в прогнозе цены за 1000 лир в долларах США как функция длины интервала идентификации. По графику можно заметить, что наилучшее поведение полосы проявилось, когда длина была в пределах от 72 до 108 месяцев: среднее значение

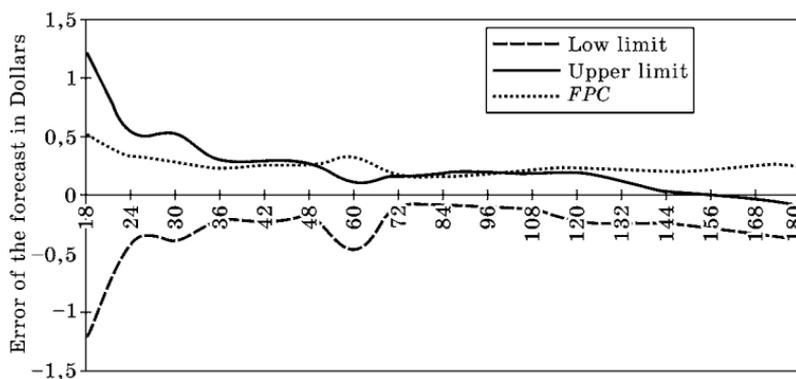


Рис. 30

Модель с полиномом и одним колебанием

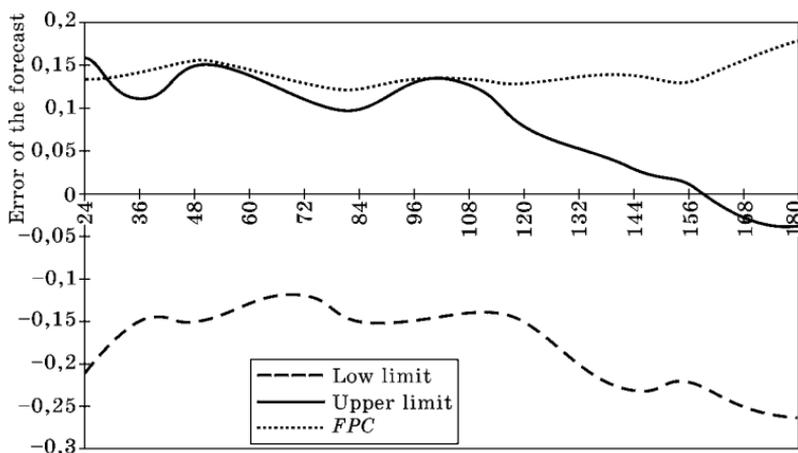


Рис. 31

Модель с двумя колебаниями

ошибки равно нулю и ширина минимальна. Этот вывод мы не могли бы получить только по *FPC*, так как последний имеет несколько экстремумов и к тому же неясно, будет ошибка смещенной или нет. Но самое главное, что из последнего графика видна средняя цена за риск: 19 центов при покупке каждой тысячи лир.

Модель, построенная при другой длине данных, дает либо большую цену риска, либо (смещаясь относительно нуля) вносит асимметрию между покупкой и продажей валюты.

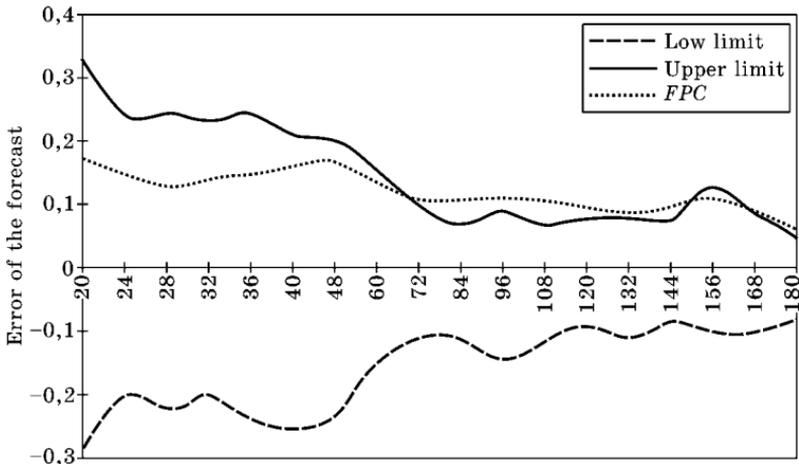


Рис. 32
Метод общей корреляции

Все вышеизложенное позволяет выделить построение доверительной полосы, как особую графическую форму функционала качества (мы следим за шириной полосы и за симметрией относительно нуля). Назовем его *RCC* («risk charge criterion»).

На рисунке 31 можно видеть, что оптимальная длина интервала идентификации лежит в отрезке [72, 96] с 13 центами в качестве цены риска.

Результат применения метода общей корреляции показан на рисунке 32.

Очевидно, 90%-ная полоса между доверительными границами сейчас более узкая, чем в двух предыдущих случаях. Она почти всегда симметрична относительно оси абсцисс, и ее ширина постоянно убывает. И опять оптимальный интервал идентификации равен 72 месяцам (и более). В этом случае (т. е. в случае выбора длины интервала 72 месяца и выбора модели методом общей корреляции) рассматриваемая компания будет платить за риск 11 центов за каждые 1000 лир, т. е. $RCC = 0,11\$$.

Для сравнения с предыдущими рисунками продемонстрируем на двух примерах шестилетний интервал идентификации (см. рис. 33, 34).

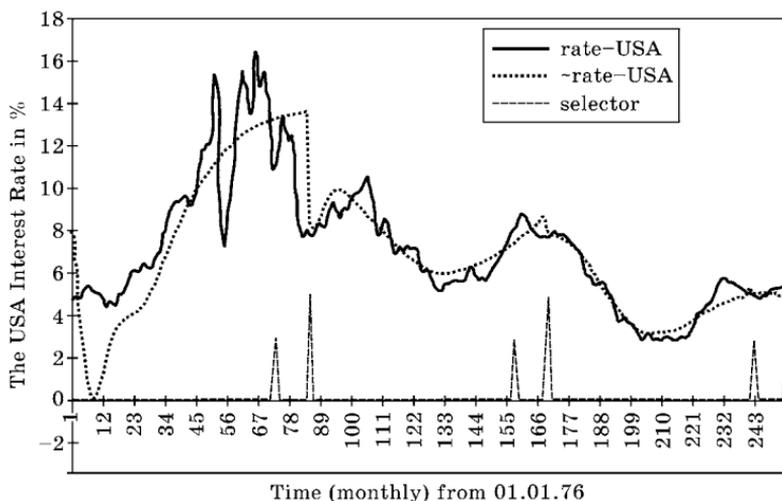


Рис. 33

Метод общей корреляции с оптимальным интервалом идентификации для процентной ставки США

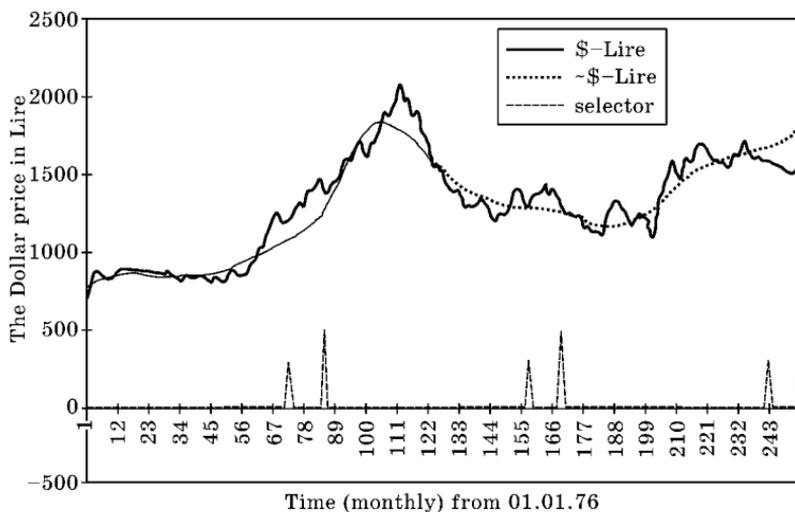


Рис. 34

Метод общей корреляции с оптимальным интервалом идентификации для цены 1 доллара США в лирах

Здесь наблюдается более плавное прогнозирование, лучше отражающее основные изменения переменных.

4.1.5. ОСНОВНЫЕ ВЫВОДЫ ИЗ РЕШЕНИЯ ЗАДАЧИ

Были предложены несколько методов прогнозирования цены форвардной покупки валюты, проведен анализ на реальных данных о развитии экономик четырех стран. Использованы два критерия качества модели: *FPC*, характеризующий средний модуль ошибки, и *RCC*, определяющий стоимость риска 90%-ного доверительного интервала в оптимальной точке модели.

Хотя второй критерий является более сложным и требующим глубокого понимания задачи, что затрудняет его широкое использование, мы сочли важным привести результаты вычислений цены риска для разных методов и разных стран (табл. 4, 5).

Таблица 4

Значения рисковых цен для различных методов и разных валют
(в долларах США)

Методы	RCC для 1000 лир	Для фунта	Для марки
Полином и одно колебание	0,14	0,43	0,12
Два колебания	0,12	0,2	0,09
Общая корреляция	0,11	0,16	0,06

Таблица 5

Оптимальные интервалы идентификации (в месяцах)

Методы	Для лиры	Для фунта	Для марки
Полином и одно колебание	72	60	60
Два колебания	72	36	36
Метод общей корреляции	72	72	72

Окончательно, в качестве лучшего, выбран метод общей корреляции с интервалом идентификации 72 месяца.

Заметим, что приведенные приращения стоимости валюты, как следует из графиков, действительно 3–4 года, следовательно, их можно использовать до 2001 г.

4.2. ЗАДАЧИ МЕЖДУНАРОДНОЙ ТОРГОВЛИ

4.2.1. НЕКОТОРЫЕ ФОРМАЛИЗАЦИИ МЕЖДУНАРОДНОЙ ТОРГОВЛИ

Анализ международных отношений, выполненный на основе теории сравнительных преимуществ, в параграфе 3.2.4 может быть формализован математически различными путями. Используя аппарат теории общего равновесия, авторы книги [105] получили ответы на многочисленные вопросы о структуре торговли между несколькими странами, об объемах товарных потоков, распределении прибыли от введения тарифов между участниками процесса и др. В основном обсуждается статическая картина, т. е. когда равновесие в торговле уже достигнуто. Ниже мы составим некоторые количественные соотношения исходя из оптимальности поведения стран по отношению к собственным прибылям и ресурсам. И далее предложим нелинейную динамическую модель развития торговли в направлении равновесного состояния.

Итак, пусть в простейшем случае имеются две страны, производящие один товар с использованием двух производственных факторов — труда L и капитала K . В первой стране производится x_1 единиц товара из L_1 единиц труда и K_1 единиц капитала. Аналогично во второй стране

$$x_1 = G(L_1, K_1), \quad x_2 = G(L_2, K_2).$$

Производственная функция G , общая для обеих стран, так как мы предполагаем, что технологически производство товара одно и то же, является однородной по обоим аргументам. В каждой стране без взаимной торговли спрос на товар удовлетворяется по одной цене p . Производство осуществляется оптимальными по прибыли затратами производственных факторов. Условия оптимальности дают две системы уравнений относительно затрат L_1, K_1 и L_2, K_2 (Приложение 2):

$$\begin{cases} p \frac{\partial G}{\partial L}(L_1, K_1) = w_1, & p \frac{\partial G}{\partial L}(L_2, K_2) = w_2, \\ p \frac{\partial G}{\partial K}(L_1, K_1) = r_1, & p \frac{\partial G}{\partial K}(L_2, K_2) = r_2. \end{cases}$$

Здесь w_1, w_2 — цена фактора труда (почасовая заработная плата) в первой и второй странах, r_1, r_2 — цена капитала (процент кредита в коммерческих банках).

Предположим теперь, что одна из стран намерена обеспечить выпуск товара для потребителей в обеих странах: $x_1 + x_2 = G(L, K)$. Тогда прибыль первой страны выражалась следующим образом:

$$I_1 = pG(L, K) - w_1L - r_1K,$$

а прибыль второй страны соответственно:

$$I_2 = pG(L, K) - w_2L - r_2K.$$

Разница этих двух величин и определит, какой из стран предпочтительнее взять на себя производство товара:

$$M = I_1 - I_2 = (w_2 - w_1)L + (r_2 - r_1)K.$$

Точнее говоря, важен только знак величины M . В силу однородности функции G в последней формуле можно L, K заменить на L_1, K_1 . Действительно, так как для любого положительного числа t справедливо равенство $tx_1 = G(tL_1, tK_1)$, то замечая, что $x_1 + x_2 = (1 + x_2/x_1)x_1$, получим:

$$M = (w_2 - w_1)(1 + x_2/x_1)L_1 + (r_2 - r_1)(1 + x_2/x_1)K_1,$$

откуда и следует вывод о знаке M . Если в производстве товара интенсивно используется труд (т. е. $L_1 > K_1$) и фактор труда в первой стране находится в относительном избытке (т. е. $w_2 > w_1$), то $M > 0$, т. е. первой стране более выгодно производить товар для потребителей обеих стран.

Естественно, что вторая страна в этом случае должна предложить взамен другой товар, с интенсивным использованием капитала. Таким образом, формализована теорема Хекшера–Олина.

4.2.2. ОБ ОДНОМ ИЗ ВОЗМОЖНЫХ ПОДХОДОВ К АНАЛИЗУ ПРОТЕКЦИОНИЗМА

Одна из наиболее ярких работ по истории протекционизма, «Толковый тариф или исследование о развитии промышленности России в связи с ее общим таможенным тарифом 1891 года» [60], принадлежит нашему великому соотечественнику Дмитрию Ивановичу Менделееву. В ней

анализируются причины, по которым нельзя вслепую следовать за теоретиками свободной торговли. Хотя идеи Д. И. Менделеева мало обоснованы теоретически и слабо формализованы (мне, например, не удалось понять, почему он предлагает тот или иной уровень тарифа), им сформулированы задачи протекционизма, т. е. определено, кто и сколько должен получать и куда эти средства направлять. С этих же позиций ведется анализ протекционизма в настоящей заметке, но здесь приводятся также формализованные количественные соотношения. В результате могут быть получены как экономические критерии введения импортного тарифа, так и выбраны оптимальные, в том или ином смысле, размеры тарифа. Предпринята попытка рассмотреть динамическую модель протекционизма, учитывающая временные факторы в международной торговле. Включение в анализ такого параметра, как контрабанда, обуславливается простым наблюдением, что если барьер будет установлен слишком высоко, то импортеры найдут на границе проход за меньшие деньги. Эту мысль также высказывал Д. И. Менделеев.

Рассмотрим взаимосвязи между величиной импорта данного товара, объемом отечественного производства, различием таможенной пошлины и налогом на добавленную стоимость для отечественного производителя, объемом контрабанды данного товара и эффективностью таможенного контроля. Математическая модель строится на простейших предположениях, допускающих ясную экономическую интерпретацию.

Предположим, что на внутреннем рынке некоторый товар (или услуга) продается по цене p , причем весь объем продаж состоит из трех частей: x — отечественного производства, y — импорта и z — контрабанды. Величины измеряются в единицах продукта; цена единицы продукта в стране, откуда производится импорт и контрабанда, измеряются в условных единицах (например, доллар). Обозначим таможенную пошлину τ , а налог на добавленную стоимость s .

Согласно введенным обозначениям, в бюджет страны из таможенного и налогового комитетов поступает:

$$S = \tau qy + sxp + \alpha zp, \quad (127)$$

где последнее слагаемое отражает реализацию по цене p внутри страны αz единиц контрабанды ($0 \leq \alpha \leq 1$ — коэффициент эффективности таможни).

Основное предположение. Будем считать, что население платит за данный продукт постоянную сумму M , а ввозимый продукт незамещаем.

Таким образом, кривая спроса имеет вид:

$$p(x + y + z) = M. \quad (128)$$

Будем считать также, что импорт определяется выгодой импортера, способного менять величину импорта в зависимости от условий на внутреннем рынке. Прибыль импортера будет вычисляться по формуле

$$D = y(p - q - \tau q). \quad (129)$$

Прибыль — это движущая сила контрабандистов:

$$G = z[(1 - \alpha)p - q]. \quad (130)$$

Мотивы изменений объема внутреннего производства могут быть самые различные: создание рабочих мест, формирование экспортного потенциала страны, укрепление производства, связанного с национальной безопасностью, и т. д. В каждом случае будет действовать свой закон изменения величины x в зависимости от условий рынка.

Рассмотрим теперь простейшие задачи, с помощью которых можно провести небольшое аналитическое исследование.

Задача 1. Предположим, что $x = \text{const}$, $z = 0$, т. е. контрабанда отсутствует, а внутреннее производство — величина постоянная. Тогда при фиксированной $\tau \geq 0$ будут иметь место соотношения:

$$p_\tau(x + y_\tau) = M, \quad D_\tau = y_\tau[M/(x + y_\tau) - q(1 + \tau)], \\ S_\tau = \tau q y_\tau + s x p_\tau.$$

Предположим, что импортер реализует наилучшую для себя стратегию, т. е. $y_\tau = \arg \max_{y_\tau} D_\tau(y_\tau)$. Оптимальную стратегию получим из условия $\frac{d}{d y_\tau} D_\tau = 0$. Это уравнение имеет решение:

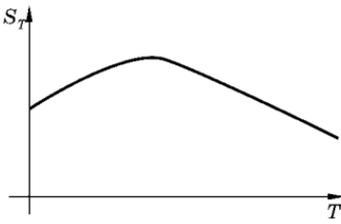


Рис. 35

Примерный вид зависимости бюджетных прибылей от таможенного тарифа

$$y_{\tau} = \sqrt{\frac{Mx}{q(1+\tau)}} - x.$$

Очевидно, что $y_{\tau} > 0$, пока $\tau < M/(qx) - 1$. Цена на внутреннем рынке при этом условии вычисляется по формуле

$$p_{\tau} = \sqrt{\frac{Mq(1+\tau)}{x}}.$$

Подставим теперь оптимальные y_{τ} , p_{τ} в формулу расчета бюджетных прибылей:

$$S_{\tau} = sx\sqrt{\frac{Mq(1+\tau)}{x}} + \tau q \left(\sqrt{\frac{Mx}{q(1+\tau)}} - x \right).$$

График динамики бюджетных прибылей S_{τ} имеет примерно следующий вид (рис. 35).

Предположим, что государство устанавливает уровень таможенной пошлины τ из условия наибольших отчислений в бюджет из налоговой и таможенной служб, т. е. из наибольшего значения функции S_{τ} . Упростим последнюю функцию, сделав замену $\mu = \sqrt{1+\tau}$:

$$S_{\mu} = q(\mu^2 - 1) \left(\sqrt{\frac{xM}{q}} \frac{1}{\mu} - x \right) + \sqrt{xMq}\mu. \quad (131)$$

Очевидно, что нам интересно, учитывая ограничения на τ , получить наибольшее значение S_{μ} при $1 < \mu < \sqrt{\frac{M}{xq}}$. Пусть $m = \sqrt{\frac{M}{xq}}$, тогда и наибольшее значение S_{μ} определяется корнем уравнения

$$\mu^3 - \frac{m(1+s)}{2}\mu^2 - \frac{m}{2} = 0.$$

В заданных пределах корень существует и является единственным корнем уравнения, в чем можно убедиться графически. Его можно вычислить по громоздкой формуле Кардано:

$$\mu = \frac{m(s+1)}{6} + \sqrt[3]{\frac{m^3(1+s)^3}{216}} + \frac{m}{4} + \frac{m}{4} \sqrt{1 + \frac{m^2(1+s)^3}{27}} + \\ + \sqrt[3]{\frac{m^3(1+s)^3}{216}} + \frac{m}{4} - \frac{m}{4} \sqrt{1 + \frac{m^2(1+s)^3}{27}}.$$

Рассмотрим сравнительно простой пример — производство и торговля телевизорами в России, который, однако, при детальном анализе может превратиться в громоздкое исследование. Все приведенные ниже предположения носят упрощенный характер.

Итак, в России в настоящее время (пример составлен для 1999 г.) проживает около 40 млн семей, из которых почти половина находится за чертой бедности, т. е. даже не мечтает о покупке нового телевизора. Оставшиеся потенциальные покупатели могли бы купить 20 млн штук телевизоров. Предположим, что по различным причинам телевизор меняют один раз в 10 лет, тогда получим ориентировочный годовой объем покупки. Конечно, существует много различных марок, имеющих несовпадающие характеристики, но для простоты примем, что в России половина продаж — отечественные (включая белорусские), а половина — импортные модели, и до введения импортных пошлин цена одного усредненного телевизора составляет 220 долларов США. Оценка оптовой цены импортируемых телевизоров в странах производителя затруднена, поэтому будем считать почти реальной (с учетом транспортировки) цену в 200 долларов США. Тогда $M = 440$ [млн долл. США]. Рассчитаем размер тарифа по указанному выше алгоритму. Он определяется экстремумом функции

$$S_{\mu} = 200 \left(1 - \mu^2 - \sqrt{2,2} \frac{1}{\mu} - 1,2 \sqrt{2,2\mu} \right),$$

который достигается при $\mu \approx 1,32$, что соответствует тарифу $\tau \approx 0,73$. И розничная цена, согласно модели, поднимется до 391 доллара США. Импорт упадет до 0,126 млн шт. При этом налоговая прибыль государства возрастет с 44 до 96,6 млн долларов США. Российские производители телевизоров получают дополнительно 171 млн долларов

США чистой прибыли в год, на которые можно построить новые заводы или выпустить современные модели телевизоров. Огромные потери понесет потребитель: почти 750 тыс. российских семей откажутся в течение года от покупки нового телевизора. Но это — отрицательные последствия любой формы протекционизма.

Задача 2. Предположим, что промышленность страны отвечает на рост цены расширением производства. При этом расширение производства, т. е. увеличение x , может быть осуществлено только за счет средств, полученных от продажи дополнительно произведенного продукта по возросшим из-за пошлины ценам. Уравнение изменения будет иметь вид:

$$\dot{x}(t) = cx(t)[p_\tau(t) - p^0], \quad (132)$$

где p_τ — цена, установившаяся после введения пошлины, p^0 — цена до введения пошлины (она отражает себестоимость, так как если бы по этой цене продавалось x единиц продукта, значит p^0 — было бы выше себестоимости только на величину торговых, транспортных, налоговых и прочих добавок). Расширение производства (увеличение x) будет возможно, если $p_\tau > p^0$ в соответствии с уравнением (132). Здесь c означает долю дохода, идущую на расширение производства. Из (132) видно, что его рост прекратится, когда $p_\tau = p^0$, т. е. с увеличением x цена, определяемая функцией спроса (128), будет падать до p^0 , если только не изменит своего поведения импортер.

Предположим, что импортер продолжает ввозить y_τ единиц продукта, а производитель, получив кредит, сумел увеличить производство до \hat{x} так, что

$$p^0(y_\tau + \hat{x}) = M. \quad (133)$$

Более строгий подход требует учитывать время выплаты кредита и процент, либо следует предположить, что необходимая сумма получена в виде государственной дотации.

Указанные предположения позволяют определить прибыль государства из следующего выражения:

$$S = \tau q y_\tau + s p^0 (M/p^0 - y_\tau).$$

Покажем, что прибыль государства больше, если $y_\tau < y$, т. е. $\hat{x} > x$, так что

$$\begin{aligned} \Delta S &= (\tau q y_\tau - s x p_\tau) - \left(\tau q y_\tau + s p^0 \left(\frac{M}{p^0} - y_\tau \right) \right) = \\ &= s \left(\frac{x M}{x + y_\tau} - M + p^0 y_\tau \right) = \frac{s y_\tau p^0}{x + y_\tau} (y_\tau - y) < 0. \end{aligned}$$

Оптимальное по τ, s приращение прибыли может быть получено аналогично, однако более целесообразно рассматривать оптимизацию прибыли государства одновременно с влиянием на коэффициент c (расширением производства), так как получаемая прибыль сокращает инвестиции в производство. Например, уравнение (126) может иметь вид:

$$\dot{x}(t) = c(s)x(t)[p_\tau(t) - p^0],$$

где $c(s)$ — убывающая функция s .

Задача 3. Предположим, что существует ненулевая контрабанда. Это, очевидно, может быть только тогда, когда существует таможенный барьер. В противном случае нет смысла рисковать. Для упрощения предположим, что $x = \text{const}$. Если α — показатель эффективности таможни, то из z единиц данного продукта, предназначенных к контрабанде, αz задерживаются таможней, а $(1 - \alpha)z$ попадают на внутренний рынок. Будем считать, что выполняют-ся соотношения:

$$p^0(y^0 + x^0) = M, \quad p_\tau(y_\tau + z_\tau + x^0) = M,$$

где y_τ — оптимальная стратегия импортера. Тогда прибыль импортера равна

$$D = y_\tau [M / (y_\tau + z_\tau + x^0) - q(1 + \tau)],$$

а контрабандиста —

$$G = z_\tau [(1 - \alpha)M / (y_\tau + z_\tau + x^0) - q].$$

Из вышесказанного следуют три вывода:

- если $(1 - \alpha)p_\tau < q$, то $z_\tau = 0$ — контрабанда вообще невыгодна, независимо от существования импорта;

- если $\alpha\tau > q\tau$, то $z\tau = 0$ — т. е. контрабандой, в сравнении с легальным импортом, заниматься невыгодно;
- если $\alpha\tau < q\tau$, то контрабанда будет значительной (хотя, по-видимому, будут и честные импортеры).

Последние формулы позволяют решать задачи оптимизации уровня таможенных пошлин в различных постановках: с одной, с двумя или даже с тремя целевыми функциями. Например, оптимальное значение импорта

$$y_\tau = \sqrt{\frac{M(x^0 + z_\tau)}{q(1 + \tau)}} - x^0 - z_\tau,$$

как и ранее, получается из условия $(d/(dy))D = 0$. Тогда государство получит

$$S = \tau q \left[\sqrt{\frac{M(x^0 + z_\tau)}{q(1 + \tau)}} - x^0 - z_\tau \right] + \sqrt{\frac{Mq(1 + \tau)}{x^0 + z_\tau}} (sx^0 + \alpha z_i).$$

Этой формулой можно воспользоваться для вычисления z_τ , α по известному S . Если z_τ , α известны, то можно вычислить оптимальные τ , s .

Приведем еще один пример из туристического бизнеса (он выполнен в дипломной работе И. Н. Соковой (1998 г.)). Товаром является туристическая путевка, называемая условно «Отдых у моря в течение двух недель». Будем рассматривать курорты, отдых на которых является равноценным, например российские курорты и курорты Болгарии. Согласно основному предположению, население платит за данный продукт сумму M , которая в данном случае вычисляется по формуле

$$M = p_1x + p_2y,$$

где p_1 — цена путевки в Болгарию; x — годовое количество туристов, посещающих Болгарию; p_2 — цена путевки на российские курорты; y — годовое количество туристов, посещающих российские курорты.

Частичное обследование туристических фирм Санкт-Петербурга в 1997–1998 гг. показало, что $p_1 = 450$ \$, $x = 3\,370\,000$, $p_2 = 710$ \$, $y = 678\,000$. Таким образом $M = 1\,997\,880\,000$ \$.

Вычислим тариф по указанному выше алгоритму. А именно его дает экстремум функции

$$S_{\mu} = q(\mu^2 - 1) \left(\sqrt{\frac{xM}{q} \frac{1}{\mu}} - x \right) + \sqrt{xMq\mu}.$$

Причем $s = 0$, так как туристические путевки не облагаются налогом на добавленную стоимость.

$$\begin{aligned} S_{\mu} &= 450(\mu^2 - 1) \left(\frac{678\,000 \times 1\,997\,880\,000}{450} \frac{1}{\mu} - 678\,000 \right) = \\ &= 45\,000 \left(\sqrt{301\,013\,920\mu} - 6780\mu^2 - \frac{\sqrt{301\,013\,920}}{\mu} + 6780 \right). \end{aligned}$$

Наибольшее значение S_{μ} определяется корнем уравнения

$$\mu^2 - (m/2)\mu - m/2 = 0, \quad (134)$$

где

$$m = \sqrt{\frac{M}{xq}} = \sqrt{\frac{1\,997\,880\,000}{450 \cdot 678\,000}} \approx 2,56.$$

Таким образом, (128) $\Leftrightarrow \mu^3 - 1,28\mu^2 - 1,28 = 0$. Корень этого уравнения вычислим по формуле Кардано, приведенной выше:

$$\mu = \frac{m}{6} + \sqrt[3]{\frac{m^3}{216} + \frac{m}{4}} + \frac{m}{4} \sqrt{1 + \frac{m^2}{27}} + \sqrt[3]{\frac{m^3}{216} + \frac{m}{4} - \frac{m}{4} \sqrt{1 + \frac{m^2}{27}}} \approx 1,7;$$

$$\mu = \sqrt{1 + \tau} \Rightarrow \tau \approx 1,89.$$

Вычисления показали, что для государства из чисто фискальных соображений надо ввести таможенный тариф на путевки в Болгарию, сделав их цену порядка 2000 \$. Очевидно, за эти деньги российские граждане не поедут ни в Болгарию, ни на свои курорты, а предпочтут Турцию, Испанию, Египет и другие курортные зоны мира. Т. е. при увеличении таможенного тарифа произойдут замещения товара другими, более дорогими до введения пошлины. Из этого отрицательного опыта использования алгоритма, очевидно, что следует дополнить его предположениями о замещаемости товаров другими.

4.3. МОДЕЛЬ ВЛИЯНИЯ РЕКЛАМЫ НА ПРОДАЖУ НОВОГО ТОВАРА

В данном разделе обсуждается математическая формализация такой трудной экономической деятельности, как рекламирование товара. Особенность здесь обусловлена следующими чертами.

1. Многообразие форм рекламной деятельности. Это могут быть телевизионные или радиоролики (иногда большие программы), газетные материалы, рекламные щиты, презентации, индивидуальные беседы у специальных стендов и т. д. Информация может содержать только упоминание о товаре, глубокое описание потребительских свойств или сравнение с другими товарами, предшествующими данному на рынке или конкурирующими с ним. Все формы, рассчитанные на возможно большую осведомленность потребителей, сильно отличаются друг от друга и, естественно, требуют для своего моделирования привлечения различных математических инструментов. Поэтому наиболее перспективный вид модели должен, по-видимому, содержать интегрированные критерии и параметры процесса рекламирования.

2. Другой отличительной чертой рассматриваемого процесса является изменение товара (точнее, изменение отношения к нему потребителей) в течение времени существования его на рынке: новый товар требует ознакомительной информации о себе, т. е. пространной и содержательной рекламы, а также информации о производителе, чтобы снять страх новизны у возможного потребителя. Традиционный товар нуждается в простом напоминании о нем, о его обычных постоянных и хороших качествах. Товар, доживающий свой век, ориентирован на консервативную (по разным причинам) часть населения, и это накладывает особый отпечаток на рекламу. Очевидно, математические модели во всех перечисленных ситуациях должны быть различны.

3. К отрицательным особенностям рекламной деятельности следует отнести недостаток информации о затратах на рекламу и о результатах ее. В российском законодательстве расходы на рекламу отнесены к прибылям, и со-

ответственно эти расходы облагаются более высоким налогом. Поэтому бухгалтерии стараются скрыть рекламные затраты и перенести их в издержки производств, чтобы впоследствии ввести в цену. По этой причине большинство предприятий не дает точных сведений о расходах. Таким образом, модели, описывающие влияние рекламы на продажи, как правило, не имеют апробации с использованием реальных данных.

И тем не менее, экономическая наука все в большей степени привлекает математический аппарат для создания алгоритмов принятия решений в части рекламной деятельности. Так, например, в статье [119] Морис Сэзини рассматривает оптимальную стратегию рекламирования, которое осуществляется посредством почтовой рассылки рекламного издания некоторой фирмы в различных районах и с вариацией графика рассылки. Обсуждается ряд научных публикаций, предшествующих данной, за период более чем 30 лет. Основным в работе является уравнение зависимости осведомленности (объема продаж) от уровня рекламы:

$$\dot{A} = g(u, A), \quad (135)$$

где A — осведомленность о товаре, u — объем рекламы, и на функцию g (реакция на рекламу) накладываются следующие ограничения:

- $g_u > 0$, т. е. реакция на рекламу возрастает, если возрастают рекламные затраты (индекс указывает на производную по соответствующему аргументу);
- $g_A < 0$, т. е. реакция убывает при большей осведомленности потребителей о товаре;
- $g_{AA} \leq 0$, т. е. скорость убывания реакции от увеличивающейся осведомленности также не возрастает;
- для любого уровня рекламы существует устойчивое равновесное состояние осведомленности;
- скорость возрастания реакции при увеличении объема рекламы убывает с увеличением осведомленности.

Эти свойства основного уравнения носят феноменологический характер, что, с одной стороны, делает модель привлекательной, но с другой — ставит ее в большую зависимость от реального процесса рекламирования. Далее

большинство авторов максимально упрощает функцию реакции g с сохранением перечисленных свойств и решает задачу оптимального управления на различных классах функций $u(t)$.

Рассмотрим эту же проблему с другой позиции, попытаемся установить причинно-следственные связи (пусть даже обладающие теми же феноменологическими свойствами) между экономическими характеристиками потребительского сообщества, где производится процесс рекламирования. Для определенности остановимся на проблеме рекламирования только нового на рынке товара (или услуги).

Итак, пусть N — общее число возможных покупателей товара (например, население страны). Обозначим αN ту часть населения, которая знает о существовании рассматриваемого товара; очевидно, $\alpha \in [0, 1]$. Далее, пусть $\beta(\alpha N)$ — та часть населения, которая хочет купить данный продукт ($0 < \beta < 1$). И часть населения, в конце концов сделавшая покупку, будет долей предыдущего количества: $\gamma\beta\alpha N$, где также $0 < \gamma < 1$. Будем считать, что в дискретный момент времени t общее число единиц продукта, купленных за все время его существования на рынке, $M(t)$. Тогда имеет место балансовое соотношение

$$M(t+1) = M(t) + \alpha(t)\beta(t)\gamma(t)N - \alpha(t-h)\beta(t-h)\gamma(t-h)N, \quad (136)$$

где h — время жизни товара. В последней формуле удобнее перейти к относительным величинам: $m(t) = M(t)/N$, и тогда уравнение, определяющее количество используемых единиц товара, примет вид:

$$m(t+1) = m(t) + \alpha(t)\beta(t)\gamma(t) - \alpha(t-h)\beta(t-h)\gamma(t-h). \quad (137)$$

Рассмотрим теперь подробно, от чего зависят параметры α , β , γ . Они, очевидно, должны принимать значения в интервале $[0, 1]$, так как характеризуют долю общего количества (мы не рассматриваем здесь случай, когда один потребитель покупает несколько единиц товара. В этом случае требуется некоторая модификация модели.).

Поскольку α характеризует только осведомленность населения о новом товаре, то величина этого параметра зависит от уровня затрат на рекламу и от количества товара, уже используемого населением,

$$\alpha(t) = \alpha(u, m(t), m(t - 1), \dots, m(t - h)).$$

Так как самый простой вид зависимости — это линейный закон, то примем следующую формулу для α :

$$\alpha(t) = \alpha_0 + \alpha_1 u + \sum_{i=0}^h \alpha_{i+2} m(t - i), \quad (138)$$

где все коэффициенты α положительны. Второй параметр β отражает интерес к покупке данного товара и зависит от уровня рекламы и благосостояния населения:

$$\beta(t) = \beta_0 + \beta_1 u + \beta_2 W,$$

в качестве характеристики благосостояния можно рассматривать, например, ВВП на душу населения. Здесь мы также ограничились линейной формой зависимости. Параметр γ показывает только факт покупки товара при том, что желание купить у потребителя уже есть, и поэтому в простейшем случае он зависит от цены и уровня благосостояния:

$$\gamma(t) = \gamma_0 - \gamma_1 p + \gamma_2 W.$$

Цена на товар, вообще говоря, не является независимой внешней переменной. Если абстрагироваться от замещаемости товара, от особенностей монопольного назначения цены и других усложняющих моделирование тонкостей, то кривую спроса выразим известным соотношением

$$m(t)p(t) = \text{const.}$$

Объединяя все это в одно выражение, получим модель динамики покупок нового товара с дискретным временем t :

$$\begin{aligned} m(t+1) = m(t) + (\alpha_0 + \alpha_1 u + \\ + \sum_{i=0}^h \alpha_{i+2} m(t-i))(\beta_0 + \beta_1 u + \beta_2 W) \left(\gamma_0 - \frac{\gamma_1}{m(t)} + \gamma_2 W \right) - \\ - (\alpha_0 + \alpha_1 u + \sum_{i=0}^h \alpha_{i+2} m(t-i-h))(\beta_0 + \beta_1 u + \beta_2 W) \left(\gamma_0 - \frac{\gamma_1}{m(t-h)} + \gamma_2 W \right). \end{aligned} \quad (139)$$

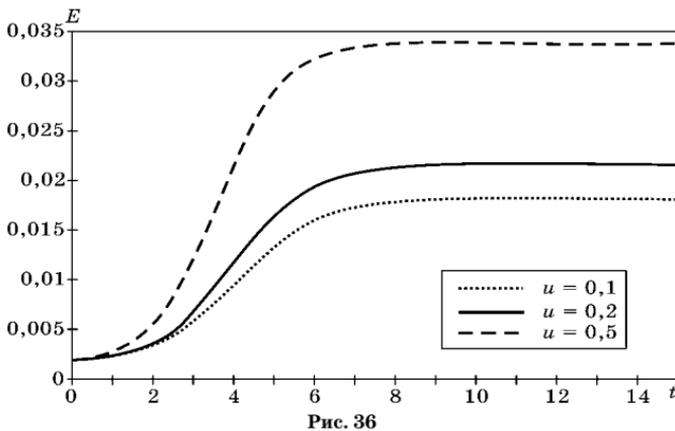


Рис. 36
Зависимость относительного объема продаж от рекламы

Модель с непрерывным временем легко трансформируется из разностного уравнения (139):

$$\begin{aligned} \dot{m}(t) = & \left(\alpha_0 + \alpha_1 u + \int_{t-h}^t \alpha(s) m(s) ds \right) (\beta_0 + \beta_1 u + \beta_2 W) \times \\ & \times \left(\gamma_0 - \frac{\gamma_1}{m(t)} + \gamma_2 W \right) - \left(\alpha_0 + \alpha_1 u + \int_{t-2h}^{t-h} \alpha(s) m(s) ds \right) \times \\ & \times (\beta_0 + \beta_1 u + \beta_2 W) \left(\gamma_0 - \frac{\gamma_1}{m(t-h)} + \gamma_2 W \right). \end{aligned} \quad (140)$$

Очевидно, это существенно нелинейное дифференциальное уравнение с распределенным запаздыванием. Выбирая различные комбинации положительных параметров α , β , γ , можно получить определенные феноменологические свойства решений. Существенное ограничение модели состоит в том, что, взяв простейшие из функций — линейные, мы вынуждены использовать модель только локально, на короткие временные горизонты, так как в уравнении (140) должны соблюдаться следующие неравенства:

$$\begin{aligned} 0 < \left(\alpha_0 + \alpha_1 u + \int_{t-h}^t \alpha(s) m(s) ds \right) < 1, \\ 0 < (\beta_0 + \beta_1 u + \beta_2 W) < 1, \quad 0 < \left(\gamma_0 - \frac{\gamma_1}{m(t)} + \gamma_2 W \right) < 1. \end{aligned}$$

Так как трудно представить себе поведение решения уравнения (140) или (139) при дискретном времени, приведем график решения с фиксированными параметрами при нескольких различных постоянных u (рис. 36):

$$m(t+1) = m(t) + (0,1 + 0,01u)(0,1 + 0,01u) \times \\ \times (0,5 - 0,001/m(t)) - (0,1 + 0,01u)(0,1 + 0,01u) \times \\ \times (0,5 - 0,001/m(t-2)).$$

Ясно, что для постоянных затрат на рекламу при малой продолжительности жизни товара при получении информации о новом товаре только через рекламу имеют место те же феноменологические свойства, что и для уравнения (135). Однако модель (140) является гораздо более богатой в смысле поведения решений.

4.4. ЗАДАЧА КОНКУРЕНЦИИ НА АВТОМОБИЛЬНОМ РЫНКЕ США

В данном параграфе рассматривается задача моделирования конкурентной динамики на внутреннем рынке США легковых автомашин в течение нескольких последних десятилетий. Американские автомобили отечественного производства конкурируют с автомашинами иностранного производства. Предложена математическая динамическая модель в форме двух дифференциальных уравнений типа Лотки–Вольтерры. Приведены основные свойства модели: равновесные траектории, асимптотическое поведение, диссипативность и колебательность. Особое внимание уделено проблеме идентификации с одновременным получением параметров модели, как коэффициентов, так и временного лага. Анализ выбранной модели показал, что введение временного лага в данном случае приносит новую динамику по сравнению с моделью без лагов.

4.4.1. ФОРМУЛИРОВКА ЗАДАЧИ

Мы ставим цель построить и проанализировать динамическую модель взаимодействия двух экономических агентов, принимая во внимание временные лаги и нелинейности. По экономическим параметрам эта задача относится

к микроэкономике, поскольку мы не рассматриваем изменение цены на рынке как следствие изменения объемов производства. В качестве примера рассматривается американский внутренний рынок автомашин. Это сравнительно замкнутая система с хорошо информированными потребителями. На этом рынке действуют отечественные производители (GM, Chrysler, Ford) и много иностранных компаний (для нас неважно, связаны ли они между собой общими капиталами). Для анализа мы будем использовать данные с американской Интернет-страницы правительственного Бюро экономического анализа о месячных продажах отечественных и иностранных машин с 1967 по 2005 гг. Конечно, данные о продажах отличаются от данных производства автомобилей на величину экспортируемых машин. Однако экспорт американских машин не очень большой, и к тому же, с конца 1970-х гг. общий экспорт Америки существенно снизился по сравнению с импортом. Поэтому мы предполагаем, что данные о продажах отражают данные производства с точностью до постоянного множителя (или, что то же самое, с точностью до единиц измерения). Таким образом, первый экономический агент — это американская автомобильная промышленность с основными фондами $K_1(t)$ и общими продажами на своем внутреннем рынке $Y_1(t)$ (что соответствует выходу отечественной промышленности по предположению).

Второй экономический агент — это все иностранные компании, участвующие в продажах на американском внутреннем рынке автомашин. Так как у нас имеются только данные по продажам импортных машин, необходимо предположить, что иностранные компании поставляют на этот рынок такое количество машин, которое допускается конкуренцией без уменьшения цены. Чтобы исключить из рассмотрения изменение цены в результате давления конкурентов, мы предположим, что установилась некоторая постоянная цена p , определенная спросом, и в то же самое время конкурентное движение цены вниз будем рассматривать как рекламные действия, т. е. как некоторый вид издержек производства. Мы не будем принимать во внимание возможные замещения в спросе одного вида машин

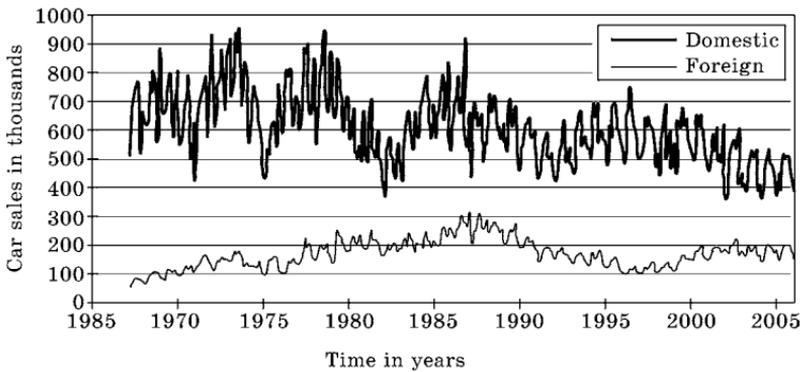


Рис. 37

Данные о продажах иностранных и отечественных машин на американском рынке (в тыс. шт., ежемесячно)

на другой, что, конечно же, происходит в реальности из-за моды, изменения цены на топливо или других причин. Также предположим, что все импортные машины подчинены одному и тому же закону динамики, независимо от страны или поставщика, и в соответствии с этим мы можем рассмотреть второго экономического агента с некоторыми условными параметрами: $Y_2(t)$ — условный выход (равный продажам импортированных машин) и $K_2(t)$ — соответствующий основной капитал.

Реальные данные о продажах машин на американском рынке, взятые на сайте БЭА, демонстрируют нам (рис. 37) сложное взаимное поведение двух описанных выше экономических агентов. Кроме обычных ошибок измерения здесь ясно наблюдаются колебания сложной природы. Это наиболее важный факт для выбора математической модели. Очевидно, мы должны учесть нелинейность и существование временного лага во взаимодействии агентов. Такими свойствами обладают логистические уравнения с запаздыванием во времени, которые по сложившейся традиции носят имя моделей Лотки–Вольтерры. Они хорошо изучены и применялись для анализа динамики биологических популяций, а также в последнее время в экономике [16, 75, 103, 112, 113, 121]. В нашей задаче основным нелинейным свойством является эффект насыщения рынка: покупатели не хотят покупать больше, чем определенное количество

машин, даже если цена несколько снижена. В процессе моделирования этот факт превращается в требование диссипативности модели. Предположим, что управляющий первого экономического агента (отечественное производство машин) принял решение резко увеличить выпуск автомобилей с помощью импульсного инвестирования. В этом случае даже с вытеснением импортных машин существует ограничение продаж из-за ограничения в спросе. Таким образом, рынок демпфирует неожиданные изменения в предложении, а следовательно, в данных. Модели Лотки–Вольтерры достаточно исследованы, и они обладают свойством диссипации, чтобы использовать их для описания конкурентной динамики.

Второе важное свойство данных — это присутствие временного лага в причинно-следственных связях, которые лежат в основе моделирования. Для экономики временные лаги вполне естественны: в производстве время реакции на изменение спроса или стратегии конкурента не может быть очень маленьким, так как требуется время и капитал для расширения производства (то же самое для сокращения). В математическом плане введение временных лагов приводит к дифференциальным уравнениям с запаздыванием, которые имеют более сложный характер, чем без запаздывания. Это позволяет модели быть более адекватной реальным данным.

Все это подталкивает нас к использованию моделей Лотки–Вольтерры для описания динамики автомобильных продаж. Модель состоит из двух нелинейных однородных уравнений с полиномиальной правой частью второго порядка и запаздывающим аргументом.

4.4.2. ВЫВОД УРАВНЕНИЙ

Ниже мы повторим некоторые рассуждения и выводы из главы 3, поскольку они относятся теперь к конкретным отраслям промышленности.

Рассмотрим два экономических агента: американская автомобильная промышленность и промышленность остального мира. Мы выводим динамические уравнения в

обычном виде. Сначала объясним уравнение в скалярной форме и затем распространим его на систему второго порядка. Предположим, что выход промышленности Y является функцией основного капитала K и фактора труда L , т. е. $Y = F(K, L)$. Для простоты предположим, что все функции линейны и L пропорциональна K , таким образом, $Y = aK$. Затем предположим, что весь выход Y делится на C — общие производственные затраты и I — общие инвестиции: $Y = C + I$. Обычно при этом предполагается $C = sY$. Таким образом, $I = (1 - s)Y$. И последний шаг состоит во введении динамики: $\dot{K} = I$. Заметим, что здесь всегда I является внутренней переменной, т. е. мы не предполагаем каких-либо внешних по отношению к данной промышленности. Окончательно имеем дифференциальное уравнение промышленной динамики

$$\dot{K}(t) = a(1 - s)K(t). \quad (141)$$

Если мы принимаем во внимание эффект насыщения на рынке, где продукт промышленности продается, то тогда уравнение (141) станет более сложным: мы должны включить в уравнение (141) эффект насыщения. Это может быть сделано разными путями. Здесь хотелось бы рассмотреть s (норму общих издержек) как функцию состояния рынка. Если говорить более точно, то величина $sY(t)$ является частью выхода $Y(t)$, которая возвращается в производство и процесс распределения. Если производство не испытывает давления со стороны конкурента, то оно может установить монопольную цену, платить зарплату в соответствии с соглашениями и высокие дивиденды акционерам и т. д. В этом случае $sY > 0$ и $I > 0$. В противном случае производство нуждается в некоторой внешней помощи или продаже основных фондов (имущества), может быть, выпуске облигаций, выводе из производства части основных фондов и сдаче их в аренду и т. д. Когда конкуренты приходят на рынок, то в значительной степени промышленность, о которой идет речь, должна тратить деньги на рекламу или уменьшать цену, или уменьшать дивиденды, или как-то иначе удерживать свое положение на рынке. Все эти действия мы можем суммировать одной функцией:

$s = s(K)$, или в простейшей линейной форме: $s = s_0 + s_1K$. Возвращаясь к уравнению (141), подставим вновь полученную функцию и получим следующее нелинейное уравнение с некоторыми новыми параметрами:

$$\dot{K}(t) = aK(t)[b + cK(t)]. \quad (142)$$

Обычно эффект насыщения имеет место, если $a, b > 0$; $c < 0$. Предположим, что на автомобильном рынке два разных экономических агента действуют с подобными динамическими уравнениями, как (142). Следовательно, их общая динамика может быть выражена следующими уравнениями:

$$\begin{cases} \dot{K}_1(t) = K_1(t)[a_1 + a_2K_1(t) + a_3K_2(t)], \\ \dot{K}_2(t) = K_2(t)[b_1 + b_2K_1(t) + b_3K_2(t)], \end{cases} \quad (143)$$

где a_i, b_i — фиксированные параметры, K_1 — капитал американской автомобильной промышленности, K_2 — капитал заморских производителей, которые импортируют машины на американский рынок.

К сожалению, система (143) не является реалистической потому, что динамические переменные K_1 и K_2 везде входят в один и тот же момент времени. В действительности есть несколько лагов в нашем предыдущем выводе. Первый лаг появляется между временем инвестирования и соответствующим временем выхода, обозначим его $h > 0$. Второй лаг, который должен быть учтен, является разницей во времени между изменениями в рыночной ситуации и принятием решения с целью улучшения этой ситуации посредством изменения s . Этот лаг обозначим $r > 0$. Таким образом, система (143) трансформируется следующим образом:

$$\begin{cases} \dot{K}_1(t) = K_1(t-h)[a_1 + a_2K_1(t-h-r) + a_3K_2(t-h-r)], \\ \dot{K}_2(t) = K_2(t-h)[b_1 + b_2K_1(t-h-r) + b_3K_2(t-h-r)]. \end{cases} \quad (144)$$

Конечно, различные экономические системы имеют разные лаги, и, следовательно, каждое уравнение должно иметь собственные лаги, но это сделало бы задачу недоступной анализу. Далее мы будем полагать $h = 0$.

Система (144) имеет неопределенные параметры, которые мы должны идентифицировать, используя реальные данные. Вспоминая, что Y пропорциональна K , преобразуем систему (144) в окончательной форме:

$$\begin{cases} \dot{Y}_1(t) = Y_1(t)[a_1 + a_2 Y_1(t-r) + a_3 Y_2(t-r)], \\ \dot{Y}_2(t) = Y_2(t)[b_1 + b_2 Y_1(t-r) + b_3 Y_2(t-r)]. \end{cases} \quad (145)$$

Это более удобно для идентификации коэффициентов, так как мы имеем только данные о продажах машин.

Из содержательного смысла задачи очевидно, что a_1 и b_1 должны быть положительны, потому что, когда Y_1 и Y_2 малы, ничего не препятствует росту продаж. Другие параметры могут иметь после идентификации любой знак. Намереваясь получить общий алгоритм для всевозможных случаев, мы будем при необходимости давать соответствующую экономическую интерпретацию.

4.4.3. ИДЕНТИФИКАЦИЯ И ЧИСЛИТЕЛЬНОЕ ПРИЛОЖЕНИЕ

В главе 3 подробно обсуждались качественные свойства систем типа (145). Для справедливого отнесения реальных процессов к тем или иным группам поведений, нам необходимо знать реальные коэффициенты системы (145). Кроме того, чтобы использовать эти модели для прогнозирования, также подразумевается процедура идентификации параметров, в частности величины запаздывания. Однако мы оставим, по возможности, в стороне статистические проблемы, потому что они требуют значительного места для подробного анализа, но в то же время не принесут нам более конструктивных алгоритмов, чем это получается без глубокого статистического анализа.

Наша конкретная задача идентификации заключается в получении вполне определенной системы дифференциальных уравнений с одним сосредоточенным запаздыванием (лагом). Эта определенность основывается на нашем субъективном выборе модели в форме уравнений Лотки–Вольтерры. Таким образом, после процедуры дискретизации мы приходим к задаче идентификации нелинейной разностной

системы со специальным видом лага, т. е. мы не можем использовать известные алгоритмы выбора оптимальной структуры распределенного лага. Таким алгоритмам посвящено много работ, например, назовем несколько последних публикаций: [109, 115, 120]. Эвристические посылки авторов проявляются при формализации задачи идентификации в следующих направлениях:

- способ идентификации — параметрический или непараметрический. Первый способ более распространен и изучен. В научных журналах имеется немного работ по параметрическому подходу, тем не менее, наблюдается некоторое развитие, например применение сплайнов в работе [108];
- критерий качества модели — начиная с широко распространенного критерия «среднего квадрата ошибки» до различных «информационных» критериев есть много разных способов измерения качества прогноза. В последние годы исследователи сравнивали различные критерии, обобщали, развивали и находили лучший (см., например, работы [104, 109]), но окончательный выбор делался с помощью субъективного мнения или некоторого другого критерия;
- следующая существенная разница между моделями состоит в предпочтениях исследователя к системе функций, среди которых выбирается наилучший представитель. Можно встретить очень простые классы, как линейные или полиномиальные функции, или очень сложные системы, подобно модели Хотеллинга (см. [117]), где обыкновенное дифференциальное уравнение комбинируется с уравнением в частных производных диффузии. Последняя модель отражает специфические свойства процесса, но при этом является настолько сложной, что никто не может применять ее к серьезным многомерным задачам. С нашей точки зрения, было бы гораздо более плодотворно постепенное усложнение теории моделирования новыми инструментами и расширение класса привлекаемых функций хорошо изученными объектами, такими как уравнения Лотки–Вольтерры, а также уравнениями с запаздывающим аргументом.

Примерно те же позиции можно встретить в работе [121] и у других авторов.

Таким образом, наша задача состоит в идентификации коэффициентов модели. Для этой цели мы преобразуем систему (145) в интегральную форму:

$$\begin{aligned}
 x_1(t) &= x_1(0) \exp \left[\int_0^t (a_1 + a_2 x_1(s-r) + a_3 x_2(s-r)) ds \right], \\
 x_2(t) &= x_2(0) \exp \left[\int_0^t (b_1 + b_2 x_1(s-r) + b_3 x_2(s-r)) ds \right]
 \end{aligned}$$

(мы ввели другие переменные, чтобы не путать впоследствии разностный аналог с непрерывной исходной моделью). Интегралы в последних выражениях могут быть приближены линейными функциями. В нашем случае единицей времени является месяц. По этой причине предполагаем, что r — целое число (количество месяцев), пока что неизвестное. В результате у нас имеется нелинейная разностная система с неизвестными параметрами:

$$\begin{aligned}
 x_1(i+1) &= x_1(i) \exp \left[a_1 + a_2 \frac{x_1(i+1-r) + x_1(i-r)}{2} + \right. \\
 &\quad \left. + a_3 \frac{x_2(i+1-r) + x_2(i-r)}{2} \right], \\
 x_2(i+1) &= x_2(i) \exp \left[b_1 + b_2 \frac{x_1(i+1-r) + x_1(i-r)}{2} + \right. \\
 &\quad \left. + b_3 \frac{x_2(i+1-r) + x_2(i-r)}{2} \right].
 \end{aligned} \tag{146}$$

Неизвестные коэффициенты включены в систему (146) линейно, следовательно, мы можем использовать для их вычисления метод наименьших квадратов для различных значений r . После применения МНК оптимальные коэффициенты становятся функциями r : $a_i = a_i(r)$, $b_i = b_i(r)$. Чтобы оценить величину оптимального лага, можно использовать часть данных (здесь используется двухлетний прогноз), сравнивая модельные и реальные данные.

Качество (адекватность) выбираемой модели проверяется в два этапа: для имеющихся данных (467 пар чисел) строим два функционала, зависящих от величины лага r :

$$S_r^1(a_1, a_2, a_3) = \sum_{k=r+1}^{443} \left[\ln \frac{x_1(k+1)}{x_1(k)} - a_1 - \frac{a_2}{2}(x_1(k-r) + x_1(k-r+1)) - \frac{a_3}{2}(x_2(k-r) + x_2(k-r+1)) \right]^2,$$

$$S_r^2(b_1, b_2, b_3) = \sum_{k=r+1}^{443} \left[\ln \frac{x_2(k+1)}{x_2(k)} - b_1 - \frac{b_2}{2}(x_1(k-r) + x_1(k-r+1)) - \frac{b_3}{2}(x_2(k-r) + x_2(k-r+1)) \right]^2.$$

Так как S_r^1 и S_r^2 квадратичные по неизвестным параметрам, то найти наименьшее значение не представляет труда. В результате получим оптимальные коэффициенты модели при произвольном r . На втором этапе, чтобы выбрать оптимальное r , воспользуемся остатком данных (24 пары чисел) и построим функционал, являющийся усредненным квадратом ошибки прогноза на двухгодичном интервале:

$$D_r = \frac{1}{24} \sum_{k=443}^{466} \left[x_1(k+1) - x_1(k) \exp(a_1 + \frac{a_2}{2}(x_1(k-r) + x_1(k-r+1)) + \frac{a_3}{2}(x_2(k-r) + x_2(k-r+1))) \right]^2 + \\ + \left[x_2(k+1) - x_2(k) \exp(b_1 + \frac{b_2}{2}(x_1(k-r) + x_1(k-r+1)) + \frac{b_3}{2}(x_2(k-r) + x_2(k-r+1))) \right]^2.$$

Последний функционал зависит от r более сложным образом, чем это было в предыдущих случаях. Он может иметь один и более локальных минимумов. Поэтому простых формул вычисления оптимального лага мы предложить не можем. Однако вычислительный алгоритм в каждом конкретном случае осуществляется проверкой нескольких условий: оптимальная точка $r = r_{opt}$ должна доставлять

локальный минимум функционалу D_r , соответствующие параметры $a_1(r_{opt})$ и $b_1(r_{opt})$ должны быть положительны и исследуемая стационарная точка модели должна иметь положительные координаты.

Самое неприятное в нахождении r_{opt} состоит в том, что данные, на которых основано построение модели, имеют случайную составляющую, распределение которой заранее неизвестно. Таким образом, все получаемые значения оптимальных коэффициентов, функционалов и лагов становятся функциями случайных величин. Определить соответствующие характеристики (математические ожидания и дисперсии) напрямую не представляется возможным. Поэтому мы примем следующий алгоритм выбора:

- визуально по зависимости D_r выделим несколько лагов, подозрительных на оптимальность с соблюдением указанных выше условий;
- проверим избранные точки на чувствительность, сглаживая исходные данные с помощью скользящих средних;
- проверим избранные точки на чувствительность с помощью случайных возмущений;
- примем окончательное решение о величине лага.

Все последующие вычисления выполнены с применением пакета «Математика-5» (рис. 38).

Для прогнозирования мы применим способ оптимизации начальных данных прогноза, описанный в параграфе 2.3.5 для задачи авторегрессии без временных лагов. Здесь вычисление прогнозируемых значений производится не по реальным данным, а по усредненным с прогнозными по оптимальной модели из ближайшего прошлого. Таблица 6 демонстрирует значительное повышение качества прогноза для большинства учитываемых лагов.

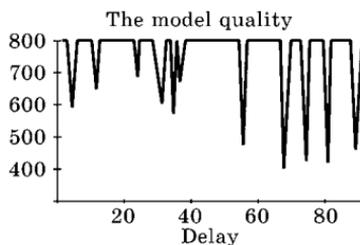


Рис. 38
На графике представлен функционал качества двухгодичного прогноза D_r , средний квадрат ошибки прогноза

Таблица 6

Сравнение прогнозов по обычным и смешанным начальным данным

Величина лага									
11	31	55	67	68	73	74	80	88	89
Средний квадрат ошибки прогноза (обычные начальные данные)									
653	747	708	600	606	538	643	603	773	798
Средний квадрат ошибки прогноза (смешанные начальные данные)									
654	617	476	401	433	486	427	421	462	477

Кроме этого из таблицы и графика функционала качества видно, что наилучшим можно считать лаг, равный 67 месяцам (или примерно 5,5 лет), и модель с оптимальными, по классическому методу наименьших квадратов, коэффициентами:

$$\begin{cases} \dot{x}_1(t) = x_1(t)[0,0121 - 0,00003x_1(t-67) + 0,000033x_2(t-67)], \\ \dot{x}_2(t) = x_2(t)[0,0194 - 0,000013x_1(t-67) - 0,000064x_2(t-67)], \end{cases}$$

где шкала времени имеет в качестве единицы один месяц. Далее, по порядку ухудшения качества двухлетнего прогноза следуют модели с лагами 80, 74, 68 и т. д.

Проверим чувствительность результата с помощью сглаженных данных, возмущенных данных, а также по количеству отвергнутых коэффициентов моделей. Сглаженные по два временные ряды сохранили локальные минимумы ошибок при лагах 11, 31, 67, 68, 74, 88, 89; сглаживание по три дает наилучшие модели с лагами 31 и 89, а сглаживание по четыре — только лаг 31.

Возмущение временных рядов проводится с помощью датчика равномерно распределенных случайных чисел с разбросом около нуля по 5, 15 и 25 единиц. Расчеты показали, что модель с лагом 67 сохраняет локальный минимум ошибки прогноза с высокой вероятностью. Остальные лаги менее устойчивы к возмущениям временных рядов.

В результате всех численных экспериментов окончательно считаем модель с лагом 67 наилучшей. Она имеет в качестве равновесной точки положение $\{603,284; 179,548\}$. Модельная траектория колеблется вокруг равновесной точки с периодом примерно 30 лет. Приведенные графики де-

монстрируют расположение модельной кривой среди реальных данных, а также прогноз по избранной модели на 4 года (рис. 39–42).

Замечание. Вопрос об устойчивости предлагаемой методики изучен численными методами и довольно поверхностно. Автора успокаивает только то, что субъективизм в выборе результата моделирования присутствует всегда в прогнозных задачах. Иногда, если это возможно, исследователи вводят предположения, которые позволяют им получить однозначный результат и доказать его свойства

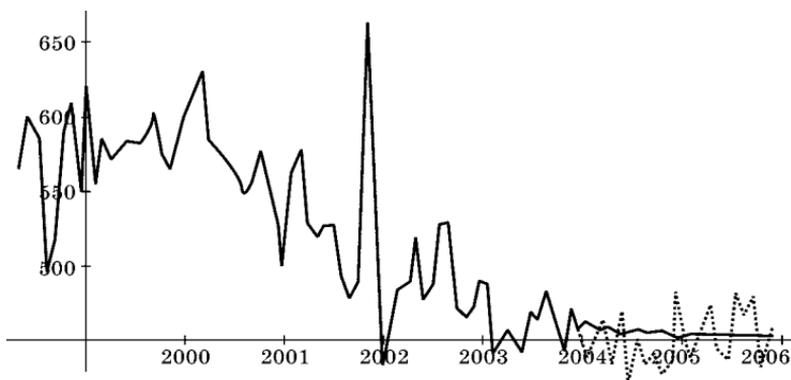


Рис. 39

Результат моделирования для месячного объема продаж американских автомобилей: сравнение реальных и модельных данных в интервале от 2004 до 2006 г.

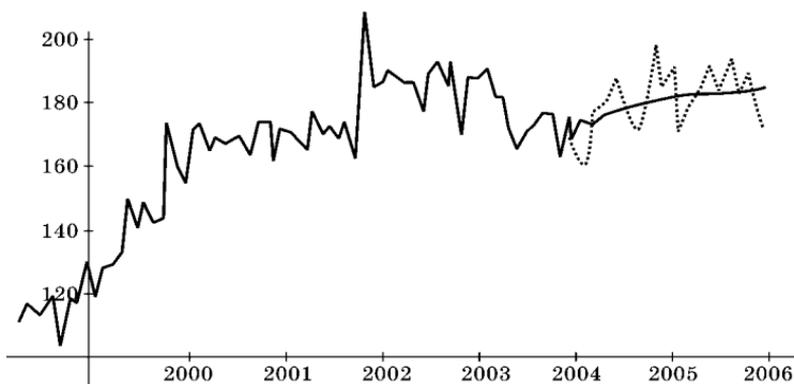


Рис. 40

Результат моделирования для месячного объема продаж иностранных автомобилей: сравнение реальных и модельных данных в интервале от 2004 до 2006 г.

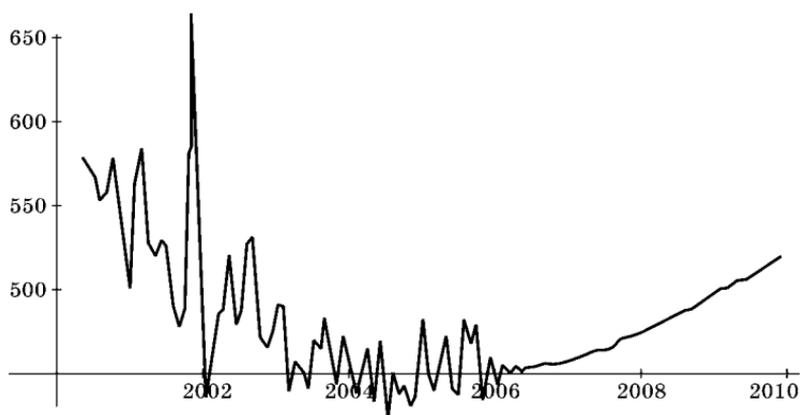


Рис. 41

Прогнозирование месячного объема продаж американских автомобилей с 2006 до 2010 г.

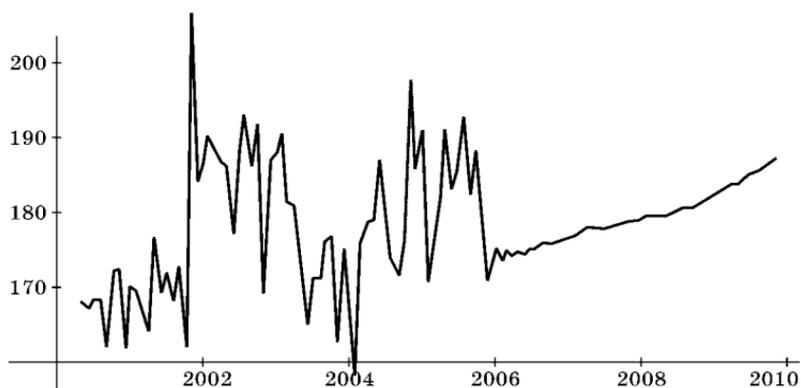


Рис. 42

Прогнозирование месячного объема продаж иностранных автомобилей с 2006 до 2010 г.

строго математически, но в таком подходе субъективизм вносится в процесс моделирования изначально, что, равным счетом, представляет собой то же самое.

Экономическая интерпретация полученного результата:

- коэффициент воспроизводства американских автомобилей без учета эффекта насыщения a_1 меньше соответствующего коэффициента b_1 : это значит, что промышленности Европы и Азии могут наращивать объемы продаж на американском рынке быстрее, чем промышленность

- Америки; наибольшее возможное увеличение поступлений автомобилей в год Америка может осуществить до 14%, а иностранные поставщики до 24%;
- коэффициенты a_2 , b_2 , b_3 отрицательны, а коэффициент a_3 положителен: все эти параметры отвечают в модели за уменьшение спроса при увеличении объемов продаж всеми поставщиками. Положительное влияние объема продаж иностранных машин на воспроизводство американских, по-видимому, объясняется национальными приоритетами правительства, т. е. на увеличение продаж импортного происхождения американское правительство отвечает стимуляцией собственного автомобильного сектора;
 - временной лаг составляет примерно 5,5 лет: такая задержка в реакции высокотехнической промышленности соответствует времени перехода на новую продукцию, способную изменить конкурентную ситуацию на рынке. Очевидно, что автомобильная индустрия разных стран за разное время обновляет ассортимент, но в данном случае мы имеем дело со средним статистическим сроком. При этом речь не идет о принципиально новых решениях, таких как переход на новое топливо, введение новых материалов или обширная компьютеризация автомобиля. Достаточно найти путь к сердцу потребителя;
 - равновесное состояние значительно выше текущих продаж американских автомобилей, в то время как продажи импортных авто колеблются около равновесной точки: в последние 20 лет на американском рынке продажи отечественная индустрия постепенно сдает позиции импорту (за указанное время в американской внешней торговле импорт на 20% превышает экспорт в целом). Поскольку модель строилась по продолжительному временному интервалу, с 1967 по 2005 г., на динамику продаж, очевидно, влияют другие, экзогенные по отношению к рассматриваемым экономическим агентам, факторы. Таким образом, прогноз будущего роста продаж автомобилей американского производства носит гипотетический характер: сначала необходимо устранить общеэкономические причины доминирования импорта.

ТЕОРИЯ ДИФФЕРЕНЦИАЛЬНЫХ СИСТЕМ С ПОСЛЕДЕЙСТВИЕМ

КЛАССИФИКАЦИЯ, ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ, ТЕОРЕМЫ СУЩЕСТВОВАНИЯ И ЕДИНСТВЕННОСТИ ДЛЯ УРАВНЕНИЙ С ПОСЛЕДЕЙСТВИЕМ

Рассмотрим несколько дифференциальных уравнений:

$$\dot{x} = x(t-1), \quad \dot{x} = f(t, x(t), x(t-h)), \quad \ddot{x} = f(t, x(t/2), \dot{x}(t-e^{-t})).$$

Это примеры уравнений с запаздывающим аргументом, то есть старшая производная входит в уравнение только для самого большого значения времени. Если старшая производная, входящая в уравнение, вычисляется в момент времени, предшествующий текущему, то такое уравнение с отклоняющимся аргументом называется уравнением опережающего типа. Например:

$$\dot{x} = f(t, x(t-2), \dot{x}(t-1), \ddot{x}(t-1/2)).$$

Все остальные уравнения называются уравнениями нейтрального типа. Например:

$$\dot{x} = -x(t) + \dot{x}(t-1).$$

Вообще говоря, дифференциальными уравнениями с отклоняющимся аргументом называются такие уравнения, в которых неизвестная функция появляется при различных значениях аргумента. Иногда встречаются названия «дифференциально-разностные уравнения», а для уравнений с запаздыванием — «уравнения с временной задержкой» или «уравнения с последствием».

Отклоняющийся аргумент могут иметь и интегральные уравнения, и уравнения в частных производных, и разностные уравнения.

Мы будем рассматривать только дифференциальные уравнения с последствием, как наиболее известный и широко встречающийся тип.

В общем виде дифференциальные уравнения с последствием записываются так:

$$\dot{x}_i(t) = F_i(t, x_i(\cdot)), \quad i = 1, 2, \dots, n; \quad t \geq t_0, \quad (147)$$

где $x_t(\cdot)$ обозначает предысторию решения $x(s)$ при $s \in [t - h, t]$ — если предыстория ограничена ($h = \text{const} > 0$); $x(s)$ при $s \in [\alpha, t]$, $\alpha \leq t_0 \leq t$ — если вся предыстория движения влияет на дальнейшее протекание процессов (α может равняться $-\infty$). Другое обозначение $x(t + \cdot) \equiv x_t(\cdot)$.

В уравнении (147) $F_i(t, x_t(\cdot))$ — функционал, заданный на некотором классе кривых $x(s)$ при $s \in [t - h, t]$, $h > 0$, это может быть класс непрерывных или кусочно-непрерывных функций.

Для того чтобы знать значение производной $\dot{x}(t)$ в уравнении (147), необходимо, чтобы была задана некоторая кривая $\varphi(s)$, $s \in [t - h, t]$, и тогда функционалы $F_i(t, \varphi_t(\cdot))$ будут определены. При этом если $t = t_0$, то кривая $\varphi(s) = \varphi_{t_0}(\cdot)$ никак не связана с уравнением (147), а если $t > t_0$, то часть кривой $\varphi(s)$ при $s \in [t_0, t]$ (или вся кривая, когда $t_0 \leq t - h$) должна быть решением уравнения (147), т. е. при подстановке в него давать тождество.

Для дифференциального уравнения с последействием задача Коши формулируется так: пусть задана некоторая кривая $\varphi_{t_0}(\cdot) = \varphi(s)$ при $s \in [t_0 - h, t_0]$ (например, кусочно-непрерывная), назовем ее начальной вектор-функцией (или просто начальной функцией). Требуется построить непрерывную кривую $x(t, \varphi_{t_0}(\cdot), t_0)$ при $t > t_0$ так, чтобы $x(t, \varphi_{t_0}(\cdot), t_0)$ удовлетворяла уравнению (147). Получим решение задачи Коши в виде «склейки» двух кривых:

$$x(s) = \begin{cases} \varphi(s) & \text{при } s \in [t_0 - h, t_0], \\ x(s, \varphi_{t_0}(\cdot), t_0) & \text{при } s > t_0. \end{cases}$$

Естественно, сразу возникает вопрос о существовании решения задачи Коши. Есть простой способ убедиться в том, что это решение существует, если функционалы $F_i(t, \varphi(\cdot))$ удовлетворяют условию Липшица по φ и по t .

Определение 1. Функционал $G(t, \varphi(\cdot))$ удовлетворяет условию Липшица по φ , если существует положительная константа L такая, что

$$|G(t, \varphi(\cdot)) - G(t, \phi(\cdot))| < L \|\varphi(\cdot) - \phi(\cdot)\| \quad (148)$$

при $t > t_0$.

Пока оставим открытым вопрос о норме $\|\cdot\|$ в уравнении (148) и области, в которой уравнение (148) имеет место.

Зададим некоторую начальную функцию $\varphi(s)$. Подставляя ее в (147), получаем $F_i(t_0, \varphi_{t_0}(\cdot)) = \dot{x}(t_0)$. Имея производную слева решения $x(t)$ при $t = t_0$, сдвинемся вправо по этому направлению, оставляя слева вполне определенную кривую — сначала $\varphi(s)$

и далее $x(t, \varphi_{t_0}(\cdot), t_0)$. Этот процесс возможен, так как аналогично обыкновенным дифференциальным уравнениям при принятых предположениях производная в (147) всякий раз определена, и решение существует и продолжимо вправо на некоторое расстояние, а именно до тех пор, пока выполняется условие Липшица. Такой метод называется «методом шагов».

Рассмотрим простой пример:

$$\dot{x}(t) = x(t-1), \quad t_0 = 0, \quad \varphi(s) = s \quad \text{при } s \in [-1, 0].$$

Тогда для $t \in [0, 1]$ будем иметь:

$$x(t) = x(0) + \int_0^t x(t-1) dt = \int_{-1}^{t-1} x(s) ds = \frac{s^2}{2} \Big|_{-1}^{t-1} = \frac{1}{2}(t-1)^2 - \frac{1}{2} = \frac{t}{2}(t-2).$$

Далее при $t \in [1, 2]$:

$$x(t) = x(1) + \frac{1}{2} \int_0^{t-1} (s^2 - 2s) ds = \left(\frac{s^3}{6} - \frac{s^2}{2} \right) \Big|_0^{t-1} - \frac{1}{2} = \frac{1}{6}(t-1)^3 - \frac{1}{2}(t-1)^2 - \frac{1}{2}.$$

Обычным интегрированием получим решение последовательно на промежутках $(0, 1]$, $(1, 2]$, $(2, 3]$...

Казалось бы, что доказательство существования решения задачи Коши этими рассуждениями получено и «метод шагов» можно считать пригодным для всех уравнений общего вида, но рассмотрим уравнение

$$\dot{x}(t) = -x(t + \sin t - 1)$$

при $t_0 = 0$.

Здесь аргумент $g(t) = t + \sin t - 1$ при $t = \pi/2 + 2k\pi$ ($k = 0, \pm 1, \pm 2, \dots$) обладает свойством $g(t) = t$ и $h = 2$, поскольку наибольшее отклонение аргумента равно 2. Применим «метод шагов». При $t = 0$ имеем $g(t) = 1$. Пусть, например, начальная функция будет $\varphi(s) = 1$ при $s \in [-1, 0]$. Тогда для $t \in [0, t_1]$ имеем $x(t) = 1 - t$ и t_1 является корнем уравнения $t_1 + \sin t_1 = 1$. Далее при $t \in [t_1, t_2]$, где t_2 является положительным корнем уравнения $t_1 = t_2 + \sin t_2 - 1$, и т. д., получаем последовательность моментов времени $\{t_n\} = 0, t_1, t_2, \dots$. Общий член этой последовательности определяется неявным рекуррентным соотношением

$$t_n = t_{n-1} + 1 - \sin t_{n-1}.$$

Последовательность является монотонно возрастающей и имеет предел $t_n \rightarrow \pi/2$. Установим это. Последовательность монотонно возрастает при $t_n \in (0, \pi/2)$, так как $1 - \sin t_n > 0$, а при $t_n \rightarrow \pi/2$ имеем $\Delta_n = 1 - \sin t_n \rightarrow 0$. Осталось показать, что если $t_n \in (0, \pi/2)$, то и $t_{n+1} \in (0, \pi/2)$.

Для этого воспользуемся неравенством $\sin(x) > 1 - (x - \pi/2)^2$, которое верно при всех x . Из него следует, что $t_{k+1} < tk + (t_k - \pi/2)^2$. Обозначая $s_k = t_k - \pi/2$, получаем неравенство

$$s_{k+1}^2 - s_{k+1} + s_k > 0.$$

Так как по предположению $s_k < 0$, то либо

$$s_{k+1} < \frac{1}{2} - \sqrt{\frac{1}{4} - s_k}, \text{ либо } s_{k+1} < \frac{1}{2} + \sqrt{\frac{1}{4} - s_k}.$$

Но второе невозможно, и значит,

$$s_{k+1} < \frac{1}{2} - \sqrt{\frac{1}{4} - s_k} < 0.$$

Таким образом, «метод шагов» не позволяет построить решение этого уравнения более чем на отрезке $[0, \pi/2)$, да и на этом отрезке времени построение требует бесконечного числа шагов.

Более надежным способом существования решения системы (147) служит метод последовательных приближений или другие общие способы.

Теорема 1. Предположим, что функционалы $F_i(t, \varphi(\cdot))$ определены при $t \in [t_0, t_0 + T]$, непрерывны по t здесь и удовлетворяют условию Липшица по φ в области

$$\|\varphi(\cdot)\|^{(h)} < H. \tag{149}$$

Тогда для любой непрерывной начальной функции $\varphi(\cdot)$, удовлетворяющей (149), существует и единственно решение задачи Коши для уравнения (147) при $t \in [t_0 - h, t_0 + \alpha]$, $\alpha > 0$.

Здесь норма в (149) определяется следующим образом:

$$\|\varphi(\cdot)\|^{(h)} = \sup\{|\varphi_i(s)|, i = \overline{1, n}, s \in [t_0 - h, t_0]\}. \tag{150}$$

Норма в определении условия Липшица пусть также имеет вид (150) в области (149).

Теорема 1 — это локальная теорема существования и единственности. Для ее доказательства изберем метод, основанный на «принципе сжатых отображений». Без подробных пояснений и доказательств, которые можно найти в учебниках по функциональному анализу или по дифференциальным уравнениям, это принцип может быть сформулирован так.

Принцип сжатых отображений. Пусть в полном метрическом пространстве M дан оператор A , переводящий элементы пространства M снова в элементы этого пространства. Пусть, кроме того, для любых x и y из M

$$\rho(A(x), A(y)) \leq \mu\rho(x, y),$$

где $\mu < 1$ и не зависит от x и y . Тогда существует одна и только одна точка x_0 , такая, что $A(x_0) = x_0$.

Пространство M будет метрическим, если в нем задано расстояние $\rho(x, y)$ между элементами, и оно будет полным, если любая сходящаяся в себе последовательность элементов M имеет предел в M .

Доказательство. Хорошо известно, что пространство непрерывных функций, заданных на замкнутом интервале, является полным метрическим с метрикой (150), которая называется равномерной или чебышевской. Определим по уравнению (147) оператор $A(x(t))$ следующим образом:

$$A(x(t)) = \begin{cases} \varphi(t_0) + \int_{t_0}^t F(s, x_s(\cdot)) ds, & t > t_0 \\ x(t) = \varphi(t), & t \in [t_0 - h, t_0]. \end{cases}$$

Как и прежде, $x = \{x_i\}$, $F = \{F_i\}$, где $i = \overline{1, n}$. Оператор $A(x(t))$ определен при $t \in [t_0 - h, t_0 + T]$ и $\|\varphi(\cdot)\|^{(h)} < H$. Покажем, что функция $A(x(t))$ при достаточно малом $\alpha > 0$ непрерывна, т. е. что

$$\|A(x(t + \Delta t)) - A(x(t))\| \xrightarrow{\Delta t \rightarrow 0} 0 \quad (151)$$

при $t \in [t_0, t_0 + \alpha]$.

Здесь в (151) норма обычная, евклидова: $\sqrt{\sum (\cdot)_i^2}$.

По определению A имеем:

$$\|A(x(t + \Delta t)) - A(x(t))\| = \left\| \int_t^{t + \Delta t} F(s, x_s(\cdot)) ds \right\| \leq \Delta t K,$$

где $K = \max_{s \in [t_0 - h, t_0 + \alpha]} \|F(s, x_s(\cdot))\|$.

Теперь убедимся, что оператор A для достаточно малого α — сжимающий:

$$\begin{aligned} & \|A(x(t + \Delta t)) - A(x(t))\|^{(h)} = \\ & = \sup \{ |A_i(x_i(t + s)) - A_i(y_i(t + s))| \text{ при } i = \overline{1, n}, s \in [-h, 0] \} = \\ & = \sup \left\{ \left| \int_{t_0}^{t+s} [F_i(\tau, x_\tau(\cdot)) - F_i(\tau, y_\tau(\cdot))] d\tau \right| \text{ при } i = \overline{1, n}, s \in [-h, 0] \right\} \leq \\ & \leq \sup_{i, s} \int_{t_0}^{t+s} |F_i(\tau, x_\tau(\cdot)) - F_i(\tau, y_\tau(\cdot))| d\tau \leq \\ & \leq \sup_{i, s} \int_{t_0}^{t+s} L_i \|x_\tau(\cdot) - y_\tau(\cdot)\|^{(h)} d\tau \leq L \sup_s \int_{t_0}^{t+s} \|x_\tau(\cdot) - y_\tau(\cdot)\|^{(h)} d\tau \leq \\ & \leq L \int_{t_0}^t \|x_\tau(\cdot) - y_\tau(\cdot)\|^{(h)} d\tau \leq L\alpha \sup_\tau \|x_\tau(\cdot) - y_\tau(\cdot)\|^{(h)} = L\alpha \|x_t(\cdot) - y_t(\cdot)\|^{(h)}. \end{aligned}$$

Здесь мы воспользуемся тем, что

$$L = \sqrt{\sum L_i^2} > L_i;$$

$t \in [t_0, t_0 + \alpha]$; $x(s) = y(s)$ при $s \leq t_0$. При этом последнее неравенство в цепочке соотношений будет верно при $0 < \alpha < h$. Таким образом, оператор A будет сжимающим при $\alpha < \min\{h, 1/L, T\}$. Тем самым теорема доказана.

Существует еще несколько способов доказательства существования и единственности решений уравнения (147). Это теоремы типа теоремы Пикара, основанные на методе последовательных приближений; типа теоремы Пеано, содержащие построение ломаных Эйлера. Все эти теоремы используют понятие нормы. В теореме 1 мы в качестве нормы приняли равенство (150). Изменяя определение нормы, можно получить другие доказательства этих теорем. Например, можно рассмотреть нормы вида:

$$\|x(t)\| = \sup\{|x_i(s)| \text{ при } i = \overline{1, n}; s \in [t_0, t]\}$$

или

$$\|x_t(\cdot)\|_2^{(h)} = \int_{t-h}^t \|x(s)\|^2 ds.$$

Предположим, что решение существует и единственно на $[t_0, t_0 + \alpha]$. Величина α зависит от H, h, L .

Зададим начальную кривую $\varphi(\cdot)$ следующим образом: на отрезке $[t_0, t_0 + \alpha]$ пусть $\psi(s) = x(s)$, а на отрезке $[t_0 + \alpha - h, t_0]$ пусть $\psi(s) = \varphi(s)$. Если решение $x(s)$ не нарушило неравенства $\|x_t(\cdot)\|^{(h)} < H$, то оно может быть продолжено на отрезок $[t_0 + \alpha, t_0 + 2\alpha]$ и т. д. до тех пор, пока не нарушится указанное выше неравенство или $K\alpha < T$. Можно считать, что доказано следующее утверждение.

Теорема 2. Единственное решение уравнения (147) с ограничениями теоремы 1 продолжимо до тех пор, пока $x(t)$ удовлетворяет ограничениям

$$\|x_t(\cdot)\|^{(h)} < H \text{ и } t_0 \leq t \leq T + t_0.$$

Теорема 3. Решение уравнения (147), удовлетворяющее условиям теоремы 1, непрерывно в пространстве C_0 зависит от начальной функции. Более того, если

$$\|\varphi(\cdot) - \psi(\cdot)\|^{(h)} < \delta,$$

то $\|x(t, \varphi(\cdot), t_0) - x(t, \psi(\cdot), t_0)\| < \delta e^{L(t-t_0)}$ при $t > t_0$.

До к а з а т е л ь с т в о. Если начальные функции $\varphi(\cdot)$ и $\psi(\cdot)$ выбраны таким образом, то

$$\begin{aligned} & \|x(t + \cdot, \varphi(\cdot), t_0) - x(t + \cdot, \psi(\cdot), t_0)\|^{(h)} < \\ & < \|\varphi(t_0) - \psi(t_0)\| + \int_{t_0}^t L \|x(s + \cdot, \varphi(\cdot), t_0) - x(s + \cdot, \psi(\cdot), t_0)\|^{(h)} ds. \end{aligned}$$

Обозначим $u(t) = \|x(t + \cdot, \varphi(\cdot), t_0) - x(t + \cdot, \psi(\cdot), t_0)\|^{(h)}$, тогда

$$u(t) \leq \delta + L \int_{t_0}^t u(s) ds,$$

и по лемме Гронуолла имеем

$$u(t) \leq \delta \exp[L(t - t_0)].$$

Это утверждение, которое иначе называют **теоремой об интегральной непрерывности**, и гарантирует близость решений на конечном промежутке времени, если начальные функции близки.

ЛИНЕЙНАЯ СТАЦИОНАРНАЯ СИСТЕМА С ПОСЛЕДСТВОМ

Большая часть материала этого параграфа основывается на работах В. И. Зубова по теории уравнений с запаздыванием [32]. В общем виде линейная стационарная система может быть записана так:

$$\dot{X}(t) = \int_{-h}^0 dG(\tau)X(t + \tau), \quad (152)$$

где $X = \{x_i\}_{i=1, n}$; $G(\tau) = \{g_{i, j}(\tau)\}_{i, j=1, n}$; $\tau \in [-h, 0]$, $h > 0$, $g_{ij}(\tau)$ — функции ограниченной вариации. Например, в частном случае:

$$\dot{X}(t) = AX(t) + BX(t - h),$$

матрица $G(\tau)$ связана с матрицей A и B следующим образом:

$$G(\tau) = \begin{cases} 0, & \tau = -h; \\ B, & \tau \in (-h, 0); \\ A + B, & \tau = 0. \end{cases}$$

Аналогично могут быть описаны системы с любым количеством сосредоточенных запаздываний:

$$\dot{X} = \sum_{k=1}^n A_k X(t - h_k).$$

В линейности правой части (152) легко убедиться, вспоминая определение интеграла Стильтеса. Нестационарная линейная система в отличие от (152) имеет вид:

$$\dot{X}(t) = \int_{-h}^0 dG(t, \tau)X(t + \tau).$$

Правая часть (152) удовлетворяет условию Липшица по X :

$$\left\| \int_{-h}^0 dG(\tau)X(t + \tau) - \int_{-h}^0 dG(\tau)Y(t + \tau) \right\| \leq \left\| \int_{-h}^0 dG(\tau) \right\| \cdot \|X_t(\cdot) - Y_t(\cdot)\|^{(h)}.$$

Здесь

$$\left\| \int_{-h}^0 dG(\tau) \right\| = L$$

— постоянная Липшица.

Применяя теорему об оценке решений дифференциальных уравнений, получаем:

$$\|X(t)\| < Ce^{L(t-t_0)}. \tag{153}$$

Так же, как для линейных систем обыкновенных дифференциальных уравнений, будем искать решение (152) в виде экспоненты, то есть $X(T) = e^{\lambda t}C$, где C — постоянный вектор:

$$\lambda e^{\lambda t}C = \int_{-h}^0 dG(\tau)e^{\lambda(t+\tau)}C.$$

Обозначим

$$A(\lambda) = \lambda E - \int_{-h}^0 dG(\tau)e^{\lambda\tau},$$

$\Delta A(\lambda) = 0$ — характеристическое уравнение для (152).

Следующая теорема показывает связь между линейными стационарными системами обыкновенных дифференциальных уравнений и системой (152).

Теорема 4. Пусть λ_k — корень уравнения $\Delta A(\lambda) = 0$. Тогда комплексному числу λ_k отвечает решение системы (152) вида:

$$X_k(t) = P_k(t)e^{\lambda_k t},$$

где $P_k(t)$ — полином относительно t с векторными коэффициентами и степени ниже кратности корня λ_k .

Д о к а з а т е л ь с т в о. Непосредственной подстановкой в систему (152) убедимся, что

$$X_k(t) = \frac{1}{2\pi i} \int_{C_k^+} e^{\lambda t} A^{-1}(\lambda)F(\lambda)d\lambda \tag{154}$$

будет решением. Здесь $\int_{C_k^+}$ — интеграл по замкнутой окружности C_k с центром λ_k и радиусом настолько маленьким, что в круге C_k больше корней нет. Знак «+» означает, что обход совершается

против часовой стрелки. Вектор-функция $F(\lambda)$ произвольная однозначная аналитическая в круге C_k .

Выполним подстановку:

$$\begin{aligned} \frac{d}{dt} \int_{C_k^+} e^{\lambda t} A^{-1}(\lambda) F(\lambda) d\lambda - \int_{-h}^0 dG(\vartheta) \int_{C_k^+} e^{\lambda(t+\vartheta)} A^{-1}(\lambda) F(\lambda) d\lambda = \\ = \int_{C_k^+} e^{\lambda t} (\lambda E - \int_{-4}^0 e^{\lambda \vartheta} dG(\vartheta)) A^{-1}(\lambda) F(\lambda) d\lambda = \int_{C_k^+} e^{\lambda t} F(\lambda) d\lambda. \end{aligned}$$

Последний интеграл по теореме Коши равен нулю. Таким образом, (154) есть решение (152).

Далее функцию (154) по теореме о вычетах можно представить в следующей форме:

$$X_k(t) = \operatorname{Res}_{\lambda=\lambda_k} [e^{\lambda t} A^{-1}(\lambda) F(\lambda)].$$

Положим

$$A^{-1}(\lambda) = \frac{\bar{A}(\lambda)}{\Delta A(\lambda)},$$

где $\bar{A}(\lambda)$ — матрица алгебраических дополнений для элементов $A(\lambda)$. Тогда X_k можно представить в форме:

$$X_k(t) = \frac{1}{(s_k - 1)!} \left. \frac{d^{s_k-1}}{d\lambda^{s_k-1}} \left[\frac{e^{\lambda t} \bar{A}(\lambda) F(\lambda) (\lambda - \lambda_k)^{s_k}}{\Delta A(\lambda)} \right] \right|_{\lambda=\lambda_k} = e^{\lambda_k t} \sum_{l=1}^{s_k-1} t^l C_l.$$

При этом ввиду произвольности однозначной аналитической функции F постоянные векторы C_l выражаются через s_k произвольных векторов

$$\left. \frac{d^l F}{d\lambda^l} \right|_{\lambda=\lambda_k} \quad (l = 0, \dots, s_k - 1)$$

при помощи линейных комбинаций с коэффициентами, зависящими лишь от матрицы $A(\lambda)$.

Из линейности системы (152) следует, что решением будет также любая линейная комбинация вектор-функций $X_k(t)$, а именно:

$$X(t) = \sum_{k=0}^{\infty} \alpha_k X_k = \sum_{k=0}^{\infty} \frac{\alpha_k}{2\pi i} \int_{C_k^+} e^{\lambda t} A^{-1}(\lambda) F_k(\lambda) d\lambda,$$

где $F_k(\lambda)$ — векторные функции для каждого λ_k , обладающие свойствами $F(\lambda)$.

Таким образом, получен вид частного решения. Если задана задача Коши, то есть если необходимо построить решение системы (152) по начальной функции $\varphi(\cdot) \in C_{[-h, 0]}$, то для представления решения пользуются преобразованием Лапласа.

Теорема 5. Векторная функция

$$X(t, \varphi) = \frac{V.P.}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{\lambda t} A^{-1}(\lambda) Q(\lambda, \varphi) d\lambda,$$

где $\sigma > L$ из неравенства (153), является непрерывным решением системы (152). Здесь

$$Q(\lambda, \varphi) = \int_{-h}^0 e^{\lambda \vartheta} dG(\vartheta) \int_{\vartheta}^0 e^{-\lambda \tau} \varphi(\tau) d\tau + \varphi(0).$$

Для доказательства сделаем преобразование Лапласа над решением системы (152) (при $\text{Re}(\lambda) > L$):

$$\int_0^{\infty} e^{-\lambda t} \dot{X}(t) dt = \int_0^{\infty} e^{-\lambda t} \int_{-h}^0 dG(\vartheta) X(t + \vartheta) dt$$

и далее

$$-\varphi(0) + \lambda \int_0^{\infty} e^{-\lambda t} X(t) dt = \int_{-h}^0 e^{\lambda \vartheta} dG(\vartheta) \left[\int_0^{\infty} e^{-\lambda \tau} X(\tau) d\tau + \int_{\vartheta}^0 e^{-\lambda \tau} X(\tau) d\tau \right].$$

Отсюда следует, что

$$\int_0^{\infty} e^{-\lambda t} X(t) dt = A^{-1}(\lambda) \left[\varphi(0) + \int_{-h}^0 e^{\lambda \vartheta} dG \int_{\vartheta}^0 e^{-\lambda \tau} \varphi(\tau) d\tau \right].$$

Получили образ решения задачи Коши $X(t, \varphi)$. Обратное преобразование завершает теорему.

Оказывается, что решение задачи Коши может быть разложено в асимптотический ряд

$$\sum_{k=1}^{\infty} Y_k(t, \varphi),$$

где $Y_k = \frac{1}{2\pi i} \int_{C_k^+} e^{\lambda t} A^{-1}(\lambda) Q(\lambda, \varphi) d\lambda.$

Более подробно с подобными представлениями решений системы (152) можно ознакомиться в работе [32]. Из этих представлений, в частности, следуют утверждения об устойчивости нулевого решения системы (152), а также о наличии периодических движений.

Аналогично случаю обыкновенных дифференциальных уравнений асимптотическая устойчивость будет иметь место, когда корни λ_k имеют отрицательную вещественную часть, а неустойчивость следует из наличия хотя бы одного λ_k с положительной вещественной частью. В связи с этим важное место в качественной теории занимает вопрос о расположении корней характеристического уравнения на комплексной плоскости.

СВОЙСТВА КОРНЕЙ ХАРАКТЕРИСТИЧЕСКОГО УРАВНЕНИЯ

Для линейной стационарной системы с последействием характеристическое уравнение имеет вид:

$$\det \left(\lambda E - \int_{-h}^0 dG(\tau) e^{\lambda \tau} \right) = 0. \quad (155)$$

Отметим некоторые свойства $\det A(\lambda)$ как функции комплексной переменной λ .

Свойство 1.

$$\int_{-h}^0 e^{\lambda \tau} dg_{ij}(\tau)$$

— целая функция λ , так как $e^{\lambda \tau}$ — целая функция. Целой функцией называется аналитическая функция, имеющая особенности типа полюс только на бесконечности.

Свойство 2. Характеристическое уравнение (155) имеет, вообще говоря, счетное число корней. Например, для уравнения $\dot{x}(t) = x(t-1)$ характеристическое уравнение будет иметь вид $\lambda = e^{-\lambda}$. Пусть $\lambda = x + iy$. Тогда

$$\begin{cases} x = e^{-x} \cos y, \\ y = -e^{-x} \sin y; \end{cases}$$

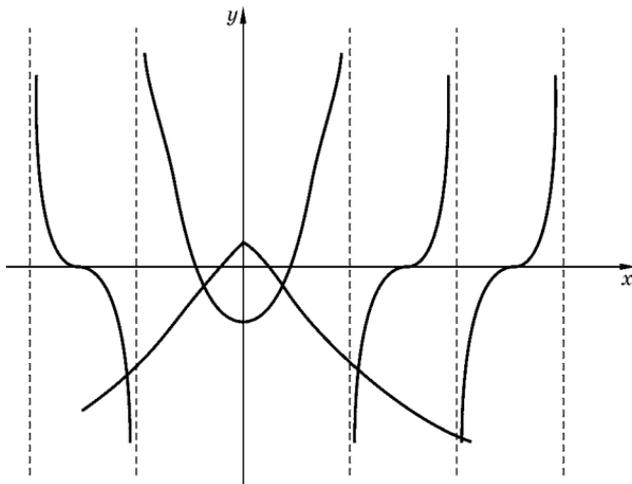


Рис. 43

Пересечения сплошных линий соответствуют корням нелинейной системы уравнений. Пунктиром нарисованы асимптоты арктангенса

или иначе

$$\begin{cases} y = \pm \sqrt{e^{-2x} - x^2}, \\ x = -y \operatorname{ctg} y, \end{cases}$$

что графически можно приблизительно изобразить следующим образом (рис. 43).

Свойство 3. В каждой вертикальной полосе конечной ширины характеристическое уравнение (155) имеет конечное число корней. Подтвердим этот факт доказательством.

Рассмотрим полосу комплексной плоскости λ : $a < \operatorname{Re} \lambda < b$. Уравнение $\det A(\lambda) = 0$ можно представить как соотношение

$$\lambda^n + \sum_{k=0}^{n-1} b_k(\lambda) \lambda^k = 0,$$

где $b_k(\lambda)$ — целые функции, так как составлены из

$$\int_{-h}^0 dg_{ij}(\tau) e^{\lambda \tau}$$

с помощью операций сложения и умножения. Функции ограничены по абсолютной величине в рассматриваемой полосе. Следовательно, существует такое положительное число B , что для произвольных комплексных чисел λ таких, что $a \leq \operatorname{Re} \lambda \leq b$ и $|\operatorname{Im} \lambda| > B$, будет верно $|\det A(\lambda)| > 0$. Действительно, при больших $|\lambda|$ имеем:

$$|\lambda^n| > \left| \sum_{k=0}^{n-1} b_k \lambda^k \right|.$$

Тогда все корни характеристического уравнения из рассматриваемой полосы оказываются сосредоточенными в прямоугольнике конечных размеров. Из этого следует, что число корней здесь конечно, так как в противном случае указанный прямоугольник содержал бы точку сгущения корней, то есть существенно особую точку функции $\det A(\lambda)$. Но это исключено, так как $\det A(\lambda)$ — целая функция.

Свойство 4. Все корни характеристического уравнения (155) лежат на комплексной плоскости левее некоторой вертикальной прямой.

Предположим, что это не так, то есть для любого числа L найдется корень уравнения (155) λ_k такой, что $\operatorname{Re} \lambda_k > L$. Но если в качестве L рассмотреть полную вариацию матрицы G и вспомнить неравенство (153), то получаем противоречие: частное решение $X_k(t) = P(t) e^{\lambda_k t}$ неравенству (153) не удовлетворяет.

Если система дифференциальных уравнений имеет одно сосредоточенное запаздывание, то характеристическое уравнение этой системы будет квазиполиномом

$$h(z, e^z) = 0,$$

где $h(z, \xi)$ — полином относительно двух переменных степени n по z и степени m по ξ . Например, для системы

$$\dot{X} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} X + \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix} X(t-h),$$

получим:

$$\det \left[\lambda E - \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix} e^{-h\lambda} \right] = 0.$$

Или, заменяя,

$$\det \begin{pmatrix} \frac{z}{h} + e^{-z} & -1 \\ -1 + 2e^{-z} & \frac{z}{h} - 2 - e^{-z} \end{pmatrix} = 0.$$

Раскрыв определитель и домножив на e^{2z} , окончательно имеем:

$$z^2 e^{2z} - 2hze^{2z} - h^2 - h^2 e^{2z} = 0.$$

Получили квазиполином второй степени по обоим переменным.

Согласно сформулированным в конце прошлого параграфа утверждениям, свойству устойчивости нулевого решения соответствует расположение корней характеристического уравнения в левой полуплоскости комплексной плоскости. Для полинома условия такого расположения дают критерии типа Рауса–Гурвица. Для квазиполиномов ряд необходимых и достаточных условий получен Л. С. Понтрягиным. Из-за недостатка места мы не будем на них останавливаться.

Для более близкого знакомства с квазиполиномами рассмотрим простой скалярный случай, который неоднократно встречается в основных главах книги:

$$\dot{x}(t) = ax(t) + bx(t-h).$$

Характеристическое уравнение имеет вид:

$$\lambda = a + be^{-\lambda h}.$$

Производя замену $z = \lambda h$, $a_1 = ah$, $b_1 = bh$, получим квазиполином

$$ze^z - a_1 e^z - b_1. \quad (156)$$

Необходимо выяснить условия, налагаемые на a_1 и b_1 , чтобы нули этого квазиполинома располагались слева от мнимой оси.

Отметим сначала, что при $a_1 + b_1 = 0$ функция (156) имеет нуль при $z = 0$. Следовательно, на плоскости (a_1, b_1) линия $a_1 + b_1 = 0$ принадлежит границе области асимптотической устойчивости в пространстве параметров. Так как нули функции (156), переходя слева от мнимой оси направо, пересекают мнимую ось, и движение это непрерывно по параметрам a_1 и b_1 , то указанной границе принадлежат также кривые, получаемые подстановкой ix вместо z в (156) при $0 \leq x \leq \infty$:

$$ixe^{ix} - a_1e^{ix} - b_1 = 0.$$

Или иначе

$$b_1 = -x/\sin x, \quad a_1 = x(\cos x/\sin x), \quad 0 \leq x \leq \infty. \quad (157)$$

Это параметрическое задание кривых на плоскости (a_1, b_1) , соответствующих нулям на мнимой оси. Так как при $x = k\pi$ знаменатели в (157) обращаются в нуль, то область определения параметра x разобьем на интервалы $(0, \pi)$, $(\pi, 2\pi)$, $(2\pi, 3\pi)$, ... Для каждого интервала рисунка 44 изображена своя кривая.

При движении точки на плоскости (a_1, b_1) изменяется расположение корней характеристического уравнения (156) на комплексной плоскости, причем, если точка плоскости параметров пересекает кривую, изображенную на рисунке 44, то какой-то корень уравнения (156) пересекает мнимую ось.

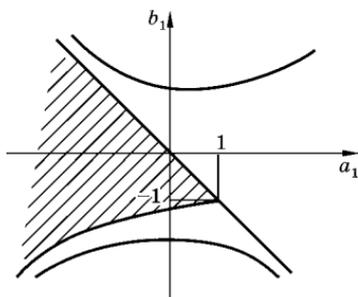


Рис. 44
Заштрихованная часть плоскости соответствует области устойчивости в плоскости параметров

Покажем, что заштрихованная область рисунка 44 является областью асимптотической устойчивости в пространстве параметров, то есть множеством точек, каждая из которых определяет устойчивый квазиполином (156). Сначала убедимся, что все точки слева от нуля на оси абсцисс соответствуют устойчивым квазиполиномам. Действительно, здесь $b_1 = 0$ и квазиполином (2.) вырождается в полином $z - a_1$, у которого один отрицательный корень. Далее, точка из заштрихованной области может выйти только через две границы: через верхнюю, но тогда квазиполином (156) приобретает нулевой корень, или через нижнюю, но тогда квазиполином приобретает два чисто мнимых корня. Все это приводит к нарушению устойчивости.

Такой способ исследования расположения корней носит название **метод D -разбиения**. Он может с успехом применяться для квазиполиномов низких порядков с малым числом свободных параметров. Несколько простейших квазиполиномов исследовано в работе [8] с помощью упоминавшихся теорем Понтрягина.

Замечание. При исследовании одной из моделей в третьей главе нам потребовалось определить критическое значение запаздывания τ , для которого решение скалярного уравнения

$$\lambda e^{\lambda\tau} = -m + in, \quad (158)$$

где $m, n > 0$, i -мнимая единица, остается в левой полуплоскости плоскости комплексных чисел. Применим метод D -разбиений для этой задачи.

Заменим неизвестную переменную λ на другую так, чтобы величина запаздывания при решении не играла роли: $\lambda\tau = z$. Тогда уравнение (158) примет вид:

$$ze^z = -m_1 + in_1, \quad (159)$$

где $m_1 = m\tau$, $n_1 = n\tau$. Определим теперь линию на плоскости коэффициентов (m_1, n_1) , которая соответствует чисто мнимым корням уравнения (159). Для этого сделаем в (159) подстановку $z = ix$ при всех вещественных x :

$$ix(\cos x + i \sin x) = -m_1 + in_1.$$

Выделение вещественной и мнимой частей приводит нас к параметрическому уравнению искомой кривой

$$m_1 = x \sin x, n_1 = x \cos x. \quad (160)$$

Возводя оба равенства во вторую степень и складывая, получим для произвольного параметра x :

$$m_1^2 + n_1^2 = x^2.$$

Из этого равенства следует симметрия искомой линии относительно осей координат и центра.

Для изображения линии воспользуемся полярными координатами:

$$m_1 = \rho \cos \varphi, n_1 = \rho \sin \varphi.$$

Откуда

$$\varphi = \arctg(n_1/m_1) = \arctg(\ctg x) = \pi/2 + k\pi - x, \quad k = 0, \pm 1, \pm 2 \dots$$

Здесь мы воспользовались равенствами (160) и известным тождеством

$$\arctg x + \arctg x = \pi/2 + k\pi, \quad k = 0, \pm 1, \pm 2 \dots$$

Из (160) следует, что $\rho = x \sin(x + \varphi)$ и, следовательно, наша кривая состоит из счетного числа ветвей:

$$\rho_k = (-1)^k(\pi/2 + k\pi - \varphi), \quad k = 0, \pm 1 \dots$$

Напомним, что мы стремимся определить область параметров (m_1, n_1) , в которой имеет место асимптотическая устойчивость дифференциального уравнения, соответствующего характеристическому уравнению (159). Исследование вещественного квазиполинома показало (см. выше), что отрезок оси m_1 от нуля до $\pi/2$ принадлежит области асимптотической устойчивости, точнее говоря, только он и принадлежит. Поэтому для построения области достаточно выделить из всех ветвей описанной выше линии одну замкнутую кривую, содержащую интервал $(0, \pi/2)$ внутри себя. Такая замкнутая кривая получается сопряжением двух кусков:

$$\begin{aligned} \rho_0 &= \pi/2 - \varphi, \quad \varphi \in (0, \pi/2) \\ \text{и } \rho_{-1} &= \pi/2 + \varphi, \quad \varphi \in (-\pi/2, 0). \end{aligned}$$

На рисунке 45 дано графическое изображение зоны устойчивости.

Возвращаясь к параметрам уравнения (158), получаем следующий критерий того, что корни (158) имеют отрицательные вещественные части.

Теорема 6. Для любых положительных чисел m, n точка с координатами (m, n) должна лежать ниже кривой $\rho_0 = \pi/2 - \varphi, \varphi \in (0, \pi/2)$, где

$$\rho_0 = \tau \sqrt{m^2 + n^2}, \quad \varphi = \arctg \frac{n}{m}.$$

Иными словами, величина запаздывания должна удовлетворять неравенству

$$\tau < \frac{\pi - \arctg \frac{n}{m}}{\sqrt{m^2 + n^2}}. \tag{161}$$

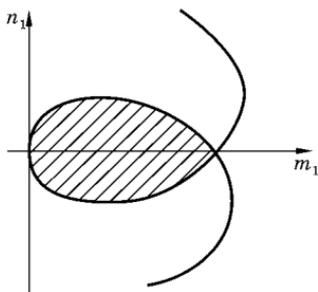


Рис. 45
Линия, задаваемая параметрически уравнениями (160) при $-\pi < x < \pi/2$

ПРЯМОЙ МЕТОД ЛЯПУНОВА ДЛЯ СИСТЕМ С ПОСЛЕДЕЙСТВИЕМ

Предположим, что основные понятия и результаты второго метода Ляпунова для систем обыкновенных дифференциальных уравнений хорошо известны. Использование функций Ляпунова для исследования устойчивости решений систем с последствием общего вида

$$\dot{x}_i = F_i(t, X_i(\cdot)); \quad i = \overline{1, n} \tag{162}$$

начиналось не очень успешно: в 1954 г. Л. Э. Эльсгольцем были сформулированы теоремы Ляпунова для системы (162) без каких-либо изменений [100]. Метод функций был признан неперспективным, поскольку теоремы не допускали обращения.

Приведем одну из таких теорем.

Теорема 7. Если существует определенно положительная функция $V(t, X)$, чья производная в силу системы (162) неположительна, то нулевое решение системы (162) устойчиво по Ляпунову.

Естественно, что предполагается выполнение условий существования нулевого решения и продолжимость вправо до бесконечности решений задачи Коши системы (162) из некоторой окрестности нуля. Теорема 1 не является практически важной, поскольку до сих пор остается открытым вопрос о существовании таких функций. Кроме того, производная функция $V(t, X)$ в силу системы (162) будет уже функционалом, определенным на отрезках траектории, что усложняет проверку знакоопределенности производной.

В 1956 г. Н. Н. Красовским был предложен метод исследования устойчивости решений системы (162) при помощи функционалов [48]. Вместо того чтобы следить за отклонением от нуля вектора $X(t)$, было предложено выделять отрезок траектории $X_t(\theta)$, $\theta \in [-h, 0]$ и в качестве обобщенного расстояния рассматривать некоторые функционалы $V(t, X_t(\cdot))$, то есть перейти из пространства E^n в пространство кусочно-непрерывных вектор-функций. Все определения второго метода Ляпунова в этом пространстве переносятся без особых изменений. Н. Н. Красовский доказал ряд теорем об устойчивости и асимптотической устойчивости, причем эти теоремы допускают обращение. Приведем без доказательства одну из них.

Теорема 8. Для того чтобы тривиальное решение системы (162) было равномерно асимптотически устойчиво по t_0 и $\varphi_{t_0}(\cdot)$, необходимо и достаточно, чтобы существовал функционал $V(\psi(\cdot), t)$, обладающий следующими свойствами:

1) $V(\psi(\cdot), t)$ определен в области

$$\|\psi(\cdot)\|^{(h)} < H, \quad t > t_0; \quad (163)$$

2) $V(\psi(\cdot), t)$ непрерывен по t и удовлетворяет условию Липшица по $\psi(\cdot)$ в (163);

3) $V(\psi(\cdot), t)$ определенно положителен и допускает бесконечно малый высший предел в (163);

4) вдоль решений системы (162) правая верхняя производная

$$\limsup_{\Delta t \rightarrow +0} \frac{\Delta V}{\Delta t} \Big|_{(1)}$$

является определенно отрицательным функционалом в (163).

Здесь следует отметить, что примерно в то же время В. И. Зубов получил результаты, хотя и не направленные на исследование систем с последствием, но имеющие более общий характер. А именно, было введено понятие «общих систем» в произвольном метрическом пространстве и доказаны необходимые и достаточные условия устойчивости инвариантных множеств. Системы с последствием, если их представить как операторные уравнения в соответствующем метрическом пространстве, будут частным случаем обеих систем. И даже более того: они близки к динамическим системам, но заданным при $t > t_0$.

Б. С. Разумихин в 1956 г. [79] продолжил изучение возможности использовать функции Ляпунова, так как проблема построения функционалов еще, по-видимому, далека от конструктивного решения, а набор функций Ляпунова в конкретных областях исследования уже довольно большой. Остановимся более подробно на результатах Б. С. Разумихина.

Наложим на функционалы $F_i(t, X_t(\cdot))$ ограничения, гарантирующие существование, единственность и продолжимость решений задачи Коши для системы (162) в области (163), а именно $F_i(t, \varphi(\cdot)) \equiv 0$ при $\varphi(\cdot) \equiv 0$:

$$\|F_i(t, \varphi(\cdot)) - F(t, \psi(\cdot))\| \leq L \|\varphi(\cdot) - \psi(\cdot)\|^{(h)}, \quad L > 0.$$

Рассмотрим сначала вопрос об устойчивости нулевого решения системы (162).

Теорема 9. Нулевое решение системы (162) будет устойчиво по Ляпунову, если существует функция $V(t, X)$ такая, что

1) $V(t, X)$ определена и непрерывно дифференцируема в области

$$\|X\| < H, \quad t \geq t_0; \tag{164}$$

2) $V(t, X)$ положительно определена в (164);

3) производная $V(t, X)$ в силу системы (162)

$$\frac{dV}{dt} = U(t, X_t(\cdot)) = \frac{\partial V(t, X(t))}{\partial t} + \sum_{i=1}^n \frac{\partial V(t, X(t))}{\partial x_i} F_i(t, X_t(\cdot))$$

является функционалом, который определен, вообще говоря, на любых непрерывных кривых $Y_t(\cdot)$ из области (163) и который обладает свойством

$$U(t, Y_t(\cdot)) \leq 0, \tag{165}$$

вдоль всяких непрерывных кривых, удовлетворяющих условию

$$Y(t) = X(t); V(s, Y(s)) \leq V(t, X(t)) \text{ при } s \in [t - h, t]. \quad (166)$$

Доказательство. Рассмотрим произвольное положительное число ε такое, что $0 < \varepsilon < H$. Обозначим $c = \inf V(t, X)$ по $t \geq t_0$, $\|X\| = \varepsilon$. Из теоремы о непрерывной зависимости решений от начальных данных и из гладкости $V(t, X)$ следует, что существует положительное число $\delta = \delta(h, t_0, C)$ такое, что для любой непрерывной начальной вектор-функции $\varphi_{t_0}(\cdot)$, подчиненной неравенству $\|\varphi_{t_0}(\cdot)\|^{(h)} < \delta$, будет иметь при $t \in [t_0, t_0 + 2h]$ оценку

$$V(t, X(t, \varphi_{t_0}(\cdot))) < C. \quad (167)$$

При этом $X(t, \varphi_{t_0}(\cdot)) < C$ для $t \in [t_0 + h, t_0 + 2h]$ будет непрерывно дифференцируемая вектор-функция. Неравенство (167) будет выполняться и при всех $t \geq t_0 + 2h$. Действительно, предположим противное, то есть пусть существует момент $T > t_0 + 2h$, когда $V(T, X(T, \varphi_{t_0}(\cdot))) = C$, а при $t < T$ имело место (167). Тогда в момент T было бы верно неравенство $U(T, X_T(\cdot)) > 0$, что противоречит условию (165), поскольку $X(t, \varphi_{t_0}(\cdot))$ оказалось такой кривой, что выполнено (166):

$$V(t, X(t, \varphi_0(\cdot))) < V(T, X(T, \varphi_0(\cdot))) \text{ при } t_0 < t < T.$$

Эти рассуждения необходимо дополнить рассмотрением случая, когда $U(T, X_T(\cdot)) = 0$. Обозначим для краткости изложения $v(t) = V(t, X(t, \varphi_{t_0}(\cdot)))$ — непрерывно дифференцируемая функция при $t \in [t_0 + 2h, T]$; ясно, что $v(t)$ имеет ограниченную вариацию на $[t_0 + 2h, T]$ и конечное число промежутков монотонности. Значит, существует число T_1 такое, что $v(t)$ при $t \in [T_1, T]$ монотонно возрастает, где T_1 достаточно близко к T . Обозначим далее $c_1 = \max(v(T_1), v(t_0 + 2h))$. Ясно, что $c_1 < c$. Из непрерывности $v(t)$ следует, что уравнение $v(t) = c$ имеет решение в промежутке $[T_1, T]$. Обозначим его T_2 . На замкнутом интервале $[T_2 - h, T_2]$ функция $v(t)$ достигает своего наибольшего значения c_2 . При этом $c_2 < c$ и, если $v(T_2) = c_2$, то процесс остановим, если же наибольшее значение достигается в полуинтервале $[T_2 - h, T_2)$, то обозначим T_3 решение уравнения $v(t) = c_2$ на интервале $[T_2, T)$, которое обязательно существует. Продолжая процесс далее, получаем $\{c_i\}$ и $\{T_i\}$. Из построения $c_i < c$, $T_i < T$. Очевидно, $c_i \rightarrow c_0$, $T_i \rightarrow T_0$, где i либо стремится к ∞ , либо прекратится в конечное число шагов. Предположим, что $T_0 = T$, тогда $c_0 = c$, и, следовательно, момент времени T не будет первым моментом выхода $v(t)$ на уровень c . Значит, $T_0 < T$ и $c_0 < c$. Но тогда существует $T^* \in (T_0, T)$, где $\dot{v}(T^*) > 0$, и выполнено условие $v(t) \leq v(T^*)$ при $t \in [T^* - h, T^*]$, то есть $U(T^*) \leq 0$. Получили противоречие.

Предположим теперь, что некоторое решение $X(t)$ по норме превзошло ε , тогда $V(t, X)$ стало бы больше c , что противоречит выбору δ . Теорема доказана.

Теорема 10. Если выполнены условия теоремы 3 и $V(t, X)$ допускает бесконечно малый высший предел, тогда нулевое решение системы (162) будет равномерно по t_0 устойчиво по Ляпунову.

Доказательство. Для того чтобы убедиться в справедливости теоремы, достаточно заметить, что δ , выбираемое в доказательстве теоремы 3, не будет зависеть от t_0 . Действительно, из того что L в условии Липшица не зависит от t_0 , по теореме о непрерывной зависимости решений от начальных данных существует $\delta_1 = \delta_1(h, \Delta) > 0$ такое, что для $\|\varphi_{t_0}(\cdot)\|^{(h)} < \delta_1$ будет $\|x(s, \varphi_{t_0}(\cdot))\| < \Delta$ при $s \in [t_0, t_0 + 2h]$ для любого положительного значения Δ . Так как $V(t, X)$ допускает бесконечно малый высший предел, существует $W(X)$ — положительно определенная функция и справедливо неравенство $V(t, X) \leq W(X)$. Подбирая $\Delta = \Delta(c)$ так, чтобы $V(t, X) < c$, получим $\delta_1 = \delta_1(h, c)$. Например, Δ таково, что

$$\sup_{\|X\| < \Delta} W(X) = c.$$

Теорема доказана.

Приведем теперь теорему об асимптотической устойчивости.

Теорема 11. Нулевое решение системы (162) будет асимптотически устойчиво, если существует функция $V(t, X)$, определенная в (164) и обладающая свойствами:

- 1) $V(t, X)$ имеет в (164) ограниченные и непрерывные частные производные;
- 2) $V(t, X)$ определенно положительна и допускает бесконечно малый высший предел;
- 3) производная $V(t, X)$ в силу системы (162)

$$dV/(dt) = U(t, X_t(\cdot))$$

является функционалом, который определен, вообще говоря, на любых непрерывных кривых $y_t(\cdot)$ из области (163) и который обладает свойством

$$U(t, y_t(\cdot)) \leq -W(\|y(t)\|)$$

вдоль всяких непрерывных кривых, удовлетворяющих условию (166). Здесь $W(r) > 0$ при $r \neq 0$, $W(0) = 0$.

Доказательство. Покажем сначала, что при условии 1 теоремы функционал $U(t, y_t(\cdot))$ удовлетворяет условию Липшица для любых непрерывных кривых $\varphi(\theta)$ и $\psi(\theta)$ ($\theta \in [-h, 0]$) из (163), совпадающих своими правыми концами, то есть $\varphi(0) = \psi(0)$:

$$|U(t, \varphi(\cdot)) - U(t, \psi(\cdot))| \leq D\|\varphi(\cdot) - \psi(\cdot)\|^{(h)}.$$

Из определения $U(t, y_i(\cdot))$ имеем:

$$\begin{aligned} & \left| \frac{\partial V}{\partial t} \Big|_{t, \varphi(0)} + \sum_{i=1}^n \frac{\partial V}{\partial x_i} \Big|_{t, \varphi(0)} \times F_i(t, \varphi(\cdot)) - \frac{\partial V}{\partial t} \Big|_{t, \psi(0)} - \sum_{i=1}^n \frac{\partial V}{\partial x_i} \Big|_{t, \psi(0)} \times F_i(t, \psi(\cdot)) \right| = \\ & = \left| \sum_{i=1}^n \frac{\partial V}{\partial t} \Big|_{t, \varphi(0)} \times [F_i(t, \varphi(\cdot)) - F_i(t, \psi(\cdot))] \right| \leq M \sum_{i=1}^n |F_i(t, \varphi(\cdot)) - F_i(t, \psi(\cdot))| \leq \\ & \leq ML\sqrt{n} \|\varphi(\cdot) - \psi(\cdot)\|^{(h)}. \end{aligned}$$

Итак, $D = ML\sqrt{n}$. Здесь L — константа Липшица для функционалов F_i , а число M ограничивает в (164) частные производные.

Далее, как в доказательстве теоремы 3, по числу $C > 0$ определим положительное число δ . Из устойчивости следует, что при $t > t_0$

$$v(t) = V(t, X(t, \varphi_{t_0}(\cdot))) < C,$$

при $\|\varphi_{t_0}(\cdot)\|^{(h)} < \delta$. Для асимптотической устойчивости достаточно показать, что $V(t, X(t, \varphi_{t_0}(\cdot))) \rightarrow 0$ при $t \rightarrow \infty$. Аналогично тому, как это было сделано в теореме 3, легко установить следующее неравенство, используя свойство 3: для произвольного $t > t_0 + 2h$ имеем:

$$\sup_{s \in [t-h, t]} v(s) > v(t). \quad (168)$$

Отсюда следует, что

$$\max_{s \in [t-h, t]} v(s) > \max_{s \in [t, t+h]} v(s). \quad (169)$$

При этом мы исходим из дифференцируемости $v(s)$, а также из следующих соображений. Имеем $\max_{s \in [t-h, t]} v(s) = v(\xi_1)$, а

$$\max_{s \in [t, t+h]} v(s) = V(\xi_2).$$

Допустим противное:

$$v(\xi_1) \leq v(\xi_2). \quad (170)$$

Из (168) следует, что $\xi_1 \in [t-h, t)$, $\xi_2 \in (t, t+h)$. Поскольку $v(\xi_2)$ — наибольшее значение на $(t, t+h)$, то $v(s) \leq v(\xi_2)$ при $s \in [t, \xi_2]$. Аналогично, $v(s) \leq v(\xi_1)$ при $s \in [t-h, t]$. Следовательно, из (170) получаем $v(s) \leq v(\xi_2)$ при $s \in [t-h, \xi_2]$, но тогда из свойства 3 имеем $U(\xi_2, X(\xi_2, \varphi_{t_0}(\cdot))) < 0$, хотя в ξ_2 функция $U(s)$ имеет нулевую производную. Значит, неравенство (170) не верно.

Определим функцию

$$l(t) = \sup_{s \geq t} v(s).$$

Очевидно, что $l(t) > 0$ и монотонно убывает. Более того, из (169) следует, что $l(t) > l(t+h)$. По определению имеем $l(t) \geq v(t)$. Для доказательства теоремы достаточно показать, что $l(t) \rightarrow 0$ при

$t \rightarrow \infty$. Пусть это не так, то есть существует $l_0 > 0$ такое, что $l(t) \rightarrow l_0$ при $t \rightarrow \infty$. Тогда, очевидно, $l(t) - l(t+h) \rightarrow 0$ при $t \rightarrow \infty$, то есть

$$\max_{s \in [t, t+h]} v(s) - \max_{s \in [t+h, t+2h]} v(s) \xrightarrow{t \rightarrow \infty} 0. \quad (171)$$

Как и выше, оба максимума обозначим соответственно $v(\xi)$ и $v(\zeta)$. Построим непрерывную кривую $y(\zeta + \cdot)$ так, чтобы:

$$y(\zeta) = x(\zeta, \varphi(t_0 + \cdot)), \quad \|y(\zeta + \cdot) - x(\zeta + \cdot)\|^{(h)} < \alpha, \quad V(s, y(s)) \leq v(\zeta)$$

при $s \in [\zeta - h, \zeta]$. Рассмотрим два случая: а) $\zeta \in (t+h, t+2h)$; б) $\zeta = t+h$. В случае а) $v(s)$ во внутренней точке ζ достигает экстремума и, значит, $\dot{v}(\zeta) = 0$. В то же время вдоль построенной кривой $y_\zeta(\cdot)$ имеем из условия $l_0 > 0$ неравенство $v(\xi) \geq l_0 > 0$. Из свойства 2 получим $\|x(\xi)\| > \mu = \mu(l_0) > 0$ и далее:

$$U(\zeta, y(\zeta + \cdot)) \leq -w(\mu) < 0. \quad (172)$$

Выберем теперь $\alpha > 0$ таким, чтобы из условия Липшица для U имело место неравенство

$$|U(\zeta, x(\zeta + \cdot)) - U(\zeta, y(\zeta + \cdot))| < -\frac{1}{2}w(\mu) \quad (173)$$

при $\|x(\zeta + \cdot) - y(\zeta + \cdot)\|^{(h)} < \alpha$. По выбранному $\alpha > 0$ найдем положительное число β так, чтобы кривую $y_\zeta(\cdot)$ можно было построить при

$$v(\xi) - v(\zeta) < \beta. \quad (174)$$

И по числу β из предельного соглашения (171) найдем T такое, чтобы (174) было верно при $t > T$.

Таким образом, случай а) приводит к противоречию: с одной стороны, $\dot{v}(\zeta) = U(\zeta, x(\zeta + \cdot)) = 0$, а с другой стороны, имеют место (172) и (173).

В случае б) имеем $\dot{v}(\zeta) < 0$. Выделим две ситуации: либо при увеличении t б) сводится к а), либо $v(t)$ при достаточно больших $t > T$ строго монотонно убывает и при всяком $t > T$ имеем ситуацию б). Но тогда $l(t) \equiv v(t)$ и $l(t)$ становится непрерывно дифференцируемой функцией. По теореме Лагранжа монотонная стремящаяся к пределу функция будет иметь производную, стремящуюся к нулю при $t \rightarrow \infty$. И применяя уже проделанное построение $y(t)$, как и в случае а), получим противоречие, которое и доказывает теорему.

Б. С. Разумихиным показано [79], что множество удовлетворяющих неравенству (166) непрерывных кривых $y(t)$, на которых проверяется отрицательность производной $V(t, x)$, можно уменьшить. А именно, обозначим $S_i(t, x)$ наибольшее значение

$|F_i(t, y(t + \cdot))|$ по всем непрерывным кривым $y(t)$, удовлетворяющим (166).

Тогда можно утверждать, что теоремы 3 и 4 останутся справедливыми, если вместо множества кривых $y_i(\cdot)$ рассматривать множество непрерывно дифференцируемых кривых $z_i(\cdot)$, удовлетворяющих неравенствам

$$V(s, z(s)) \leq V(t, z(t)), \quad s \in [t - h, t], \quad z(t) = X(t)$$

и

$$\left| \frac{dz_i}{ds} \right| \leq S_i(t, X). \quad (175)$$

Этот процесс уменьшения множества кривых можно продолжить аналогичным образом. Пусть $K_i(t, X) = \sup |F_i(t, z_i(\cdot))|$ при всех $z_i(\cdot)$, удовлетворяющих (166) и (175). Очевидно, что $K_i(t, X) \leq S_i(t, X)$. Отрицательность $U(t, X_i(\cdot))$ будем проверять теперь лишь на непрерывно дифференцируемых кривых $\xi_i(\cdot)$, таких, что

$$\xi_i(t) = X(t), \quad V(s, \xi(s)) \leq V(t, X(t)), \quad s \in [t - h, t], \quad \left| \frac{d\xi_i(t)}{dt} \right| \leq K_i(t, X).$$

На простых примерах показано, что этот процесс уточняет оценку области устойчивости в пространстве параметров системы, но сходится ли он к точному решению задачи об устойчивости, пока установить не удалось.

Например, пусть требуется найти область в пространстве параметров a и b , для которой нулевое решение уравнения

$$\dot{x} = ax + bx(t - h) \quad (176)$$

будет асимптотически устойчиво. Применяя теорему 5 с функцией $V(t, X) = x^2/2$, имеем $\dot{V} = ax^2 + bxx_h$, где $x_h = x(t - h)$. Из неравенства (166) получим при $t > t_0$

$$|x_h| \leq |x|, \quad (177)$$

и тогда $\dot{V} \leq (a + |b|)x^2$, т. е. при $a < -|b|$ нулевое решение асимптотически устойчиво (рис. 46).

На рисунке 47 изображена точная граница области устойчивости в плоскости a, b . На рисунке 48 изображена область, уточненная методом, описанным выше. Из уравнения (176) получим оценку с помощью (177):

$$|x(t) - x(t - h)| = |\dot{x}(\xi)| h \leq h \sup_{\xi \in [t-h, t]} |\dot{x}(\xi)|$$

или

$$|x - x_h| \leq (|a| + |b|)|x|,$$

т. е.

$$-2xx_h \leq h^2(|a| + |b|)^2 x^2 - x^2 - x_h^2.$$

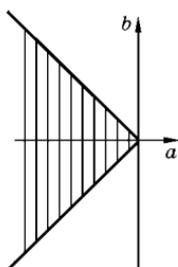


Рис. 46
Оценка области устойчивости без ограничения на производные

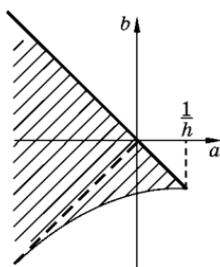


Рис. 47
Точная область устойчивости

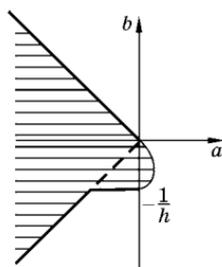


Рис. 48
Оценка области устойчивости с ограничениями на производные

Используем это неравенство для оценки \dot{V} при отрицательных значениях b :

$$\dot{V} \leq ax^2 + \left(-\frac{b}{2}\right)[x^2(h^2(|a|-b)^2 - 1) - x_h^2].$$

Очевидно, что правая часть неравенства будет отрицательна при

$$2a - 2h^2(a^2 - 2|a|b + b^2) + b < 0.$$

Избавляясь от параметра h введением новых переменных $\xi = bh$ и $\eta = ah$, получим область, изображенную на рисунке 48.

В некоторых частных задачах не все F_i зависят от предшествующего состояния переменных. В этих случаях бывает целесообразно рассматривать функции Ляпунова вида $V(t, x(t), \hat{x}(t-h))$, где под \hat{x} понимается вектор с теми компонентами вектора x , для которых $F_i = F_i(t, x(t))$. Вообще говоря, увеличение числа переменных функции V таким образом неперспективно, поскольку при дифференцировании V в силу системы (162) получим функционал, заданный на решениях $x(t+\theta)$ при $\theta \in [-2h, 0]$, а это внесет только дополнительные трудности. Польза от такого приема может быть в задачах управления, когда управляющие воздействия входят в часть уравнений системы. Например, если математическая модель описывается уравнением

$$\ddot{x} + a\dot{x} + bx = u(x(t-h), \dot{x}(t-h)),$$

то задача сводится к системе

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -ax_2 - bx_1 + u(x_1(t-h), x_2(t-h)), \end{cases}$$

и в качестве V можно выбирать функцию трех переменных

$$V = V(x_1, x_2, x_{1h}).$$

Перенос результата теоремы 5 на этот случай осуществляется простыми преобразованиями. Обозначим те номера i , для которых F_i зависит лишь от текущего значения $x(t)$ символом i_j , где $j = 1, 2, \dots, m$ ($m < n$); $\chi_{i_j}(t) = x_{i_j}(t - h)$. Тогда

$$\dot{\chi}_{i_j} = F_{i_j}(t - h, x(t - h)) = G_{i_j}(t, x(t - h)). \quad (178)$$

В функции G_{i_j} m компонент вектора $x(t - h)$ можно записать как χ_{i_j} , а остальные — как $x_i(t - h)$.

Рассмотрим вектор

$$z(t) = \{x_i(t), \chi_{i_j}(t)\}, \quad i = \overline{1, n}, \quad j = \overline{1, m},$$

очевидно, объединяя системы (162) и (178), получаем:

$$\dot{z}_k = Z_k(t, z_t(\cdot)), \quad k = 1, 2, \dots, m + n, \quad (179)$$

с начальными функциями

$$z_0(t_0 + \theta) = \{\varphi_i(t_0 + \theta), \varphi_{i_j}(t - h)\}, \quad \theta \in [-h, 0].$$

Таким образом, может быть сформулирована теорема.

Теорема 12. Если нулевое решение системы (179) асимптотически устойчиво, то асимптотически устойчиво также и нулевое решение системы (162).

В качестве функции Ляпунова для системы (179) можно использовать $V(t, z)$, что, собственно, и явилось нашей задачей. Отметим здесь одно обстоятельство: используя для системы (179) теорему 11, мы столкнемся с необходимостью проверить, допускает ли $V(t, z)$ бесконечно малый высший предел, то есть существует ли $\Phi(\|z\|)$ положительно определенная, непрерывная и ограничивающая V сверху. А также надо проверить отрицательность U , т. е.

$$U(t, z_t(\cdot)) \leq -W(\|z(t)\|).$$

Здесь важно, что в обоих случаях стоят нормы $z(t)$, а не $x(t)$.

Замечание. Одновременно с теоремой 9 появилась теорема Н. Н. Красовского, в которой роль условия (166) отводилась неравенству

$$V(s, y(s)) \leq f(V(t, y(t))), \quad (180)$$

где $f(r) > r$, f строго монотонна. Теоремы с условиями (166) и (180) применялись и обобщались различными отечественными и зарубежными математиками, однако предпочтение отдавалось условию (180). Не последнюю роль в этом сыграла заметка польского математика З. Миколайской [114], в которой приводится пример нестационарной системы с всего лишь устойчивым нуле-

вым решением при выполнении условий Разумихина об асимптотической устойчивости. Из этого делается вывод о невозможности применения условия (166) для выяснения свойства асимптотической устойчивости. Однако рассматриваемая теорема Разумихина (теорема 11) существенным образом использует наличие не зависящей от t постоянной Липшица для правых частей уравнения (162). Пример Миколайской не удовлетворяет этому требованию.

Рассмотрим признаки неустойчивости решений. Будем говорить, что имеет место неустойчивость нулевого решения системы (162), если выполнено отрицание определения устойчивости по Ляпунову. Если же выполнено отрицание определения равномерной устойчивости по Ляпунову, то будем говорить, что имеет место неустойчивость в широком смысле.

В 1960 г. [98] С. Н. Шиманов применил метод функционалов Ляпунова–Красовского для исследования нулевого решения системы (162), им доказаны аналоги теорем Ляпунова и Четаева в следующих формулировках.

Теорема 13. Если существует функционал $V(t, X_t(\cdot))$, определенный и ограниченный в области

$$\|X(t + \cdot)\|^{(h)} < H, \quad V(t, X(t + \cdot)) > 0, \quad t \geq t_0,$$

допускающий бесконечно малый высший предел и такой, что его производная в силу системы (162)

$$\limsup_{\Delta t \rightarrow 0} \left(\frac{\Delta V}{\Delta t} \right)$$

является положительно определенным функционалом, то нулевое решение неустойчиво.

Теорема 14. Если существует ограниченный в (163) функционал $V(t, X(t + \cdot))$, удовлетворяющий свойствам:

1) для любого $t \geq t_0$ существует кривая $X(\cdot)$ со сколь угодно малой нормой и $V(t, X(\cdot)) > 0$;

2) производная $V(t, X(t + \cdot))$ в силу системы (162) представима в виде $\lambda V(t, X(t + \cdot)) + W(t, X(t + \cdot))$, где $\lambda = \text{const} > 0$, $W \geq 0$, то нулевое решение неустойчиво.

Аналогично формулируется и аналог первой теоремы Ляпунова о неустойчивости. Для исследования неустойчивости в широком смысле приведем теорему [76], которая допускает обращение.

Теорема 15. Нулевое решение системы (162) неустойчиво в широком смысле тогда и только тогда, когда существует функционал $V(t, X(t + \cdot))$, обладающий следующими свойствами:

1) он ограничен в области (163);

2) для любого $\varepsilon > 0$ существует момент t^* и кривая $X^*(\cdot)$ такие, что $t^* > 0$, $\|X^*(\cdot)\|^{(h)} < \varepsilon$, $V(t^*, X^*(\cdot)) > 0$;

3) производная $V(t, X(t + \cdot))$ в силу системы (162) представима в виде:

$$\dot{V} = \lambda V(t, X_t(\cdot)) + W(t, X_t(\cdot)), \quad (181)$$

где $\lambda = \text{const} > 0$, $W \geq 0$.

Доказательство. Достаточность. Допустим противное, имеет место равномерная устойчивость, т. е. по любому положительному числу ε и любому t_0 можно подобрать положительное число $\delta = \delta(\varepsilon)$ такое, что для любых начальных кривых $X(t_0 + \cdot)$, удовлетворяющих условию $\|X(t_0 + \cdot)\|^{(h)} < \delta$, при любых $t > t_0$ будем иметь оценку $\|X(t, X(t_0 + \cdot), t_0)\| < \varepsilon$. Рассмотрим произвольное $\varepsilon < H$, по соответственному числу $\delta = \delta(\varepsilon)$ найдем такую кривую $X^*(\cdot)$ и момент времени t^* , что $\|X^*(\cdot)\|^{(h)} < \delta$, $V(t^*, X^*(\cdot)) > 0$, что возможно в силу свойства 2. Ввиду независимости δ от t_0 решение $X(t, X^*(t^* + \cdot))$ будет по норме меньше ε при $t > t^*$. Тогда из уравнения (181) имеем:

$$V(t, X(t + \cdot)) \geq V(t^*, X^*(\cdot))e^{\lambda(t-t^*)}, \quad (182)$$

то есть $V(t, X(t + \cdot)) \rightarrow \infty$ при $t \rightarrow \infty$, и это справедливо пока $\|X_t(\cdot)\|^{(h)} < H$. Но, по предположению, $\|X_t(\cdot)\|^{(h)} < \varepsilon < H$, следовательно, (182) верно при всех $t > t^*$ и $V(t, X_t(\cdot))$ принимает сколь угодно большие значения, что противоречит свойству 1.

Необходимость. Система (162) определяет двухпараметрическое преобразование начальной кривой $X_0(\cdot)$ в отрезок траектории $X(t + \cdot, X_0(\cdot), t_0)$ с параметрами $t_0 \geq 0$ и $t \geq t_0$. Это преобразование обладает групповым свойством: при $t_0 \leq s \leq t$ и $y(s + \cdot) = X(t + \cdot, X_0(\cdot), t_0)$ имеем неравенство

$$X(t + \cdot, X_0(\cdot), t_0) = X(t + \cdot, y(s + \cdot), s).$$

Предположим, что нулевое решение системы (162) неустойчиво в широком смысле, то есть по некоторому $\varepsilon > 0$ можно указать три последовательности $\{t_0^{(k)}\}$, $\{X_{0k}(\cdot)\}$, $\{T_k\}$ такие, что $t_0^{(k)} \leq t_0^{(k+1)}$, $\|X_{0k}(\cdot)\|^{(h)} < \varepsilon$, $\|X_{0k}(\cdot)\|^{(h)} \xrightarrow[k \rightarrow \infty]{} 0$, $T_k = T_k(\varepsilon, t_0^{(k)}, X_{0k}(\cdot)) > t_0^{(k)}$,

$$\limsup_{k \rightarrow \infty} \|X(T_k, X_{0k}(\cdot), t_0^{(k)})\| \geq \varepsilon. \quad (183)$$

В полуцилиндре пространства $R^1 \times PC_0[-h, 0]$, обозначенном

$$P = \left\{ t \geq 0, \varphi(\cdot) : \|\varphi(\cdot)\|^{(h)} \leq \frac{\varepsilon}{2} \right\},$$

определим функционал $\tau(t, \varphi(\cdot))$ следующим образом: τ — первый момент времени, когда имеет место соотношение $\|X(\tau, \varphi(\cdot), t)\| = \varepsilon/2$, т. е. точка X траектории выходит на боковую поверхность P в момент τ . Если точка X остается в P , то $\tau = \infty$. Теперь определим функционал $V(t, \varphi(\cdot))$ соотношением $V(t, \varphi(\cdot)) = \exp[t - \tau(t, \varphi(\cdot))]$. Из неравенства (183) следует, что при достаточно больших k и сколь угодно малых $\Delta > 0$ существует пара $\{t^*, \varphi^*(\cdot)\} \in P$ и $\|\varphi^*(\cdot)\|^{(h)} < \Delta$, такие, что $t^* < \tau(t^*, \varphi^*(\cdot)) < +\infty$, следовательно, свойство 2 имеет место. Доопределяя $V(t, \varphi(\cdot))$ в (163) значением 1 вне P , получаем свойство 1. Для проверки свойства 3 достаточно заметить, что вдоль решения системы (162) $\tau(t, X(t + \cdot))$ постоянно, так как имеет место групповое свойство преобразования X . Значит, $\dot{V} = V$, т. е. $\lambda = 1$, $W \equiv 0$. Теорема доказана.

На возможность использовать функции Ляпунова для исследования неустойчивости указывают следующие теоремы [76].

Теорема 16. Нулевое решение системы (162) будет неустойчиво в широком смысле, если существует допускающая бесконечно малый предел функция $V(t, X)$ в области (164), обладающая свойствами:

1) в сколь угодно малой окрестности нуля найдутся такая точка X^* и такой момент $t^* > h$, что $V(t^*, X^*) > 0$;

2) правое нижнее производное число $V(t, X)$ в силу системы (162) является функционалом $U(t, X(t + \cdot))$, который задан, вообще говоря, на любых кусочно-непрерывных кривых $y(t + \cdot)$ из области (163) и к тому же положительно определен на кривых, удовлетворяющих условию (166).

Под положительной определенностью функционала здесь имеется в виду существование положительно определенной функции ω такой, что

$$U(t, X(t + \cdot)) \geq \omega(\|X(t)\|). \quad (184)$$

Доказательство. Предположим противное, т. е. пусть нулевое решение будет равномерно относительно t_0 устойчиво по Ляпунову или для любого ε ($0 < \varepsilon < H$) найдется $\delta(\varepsilon)$ ($0 < \delta < \varepsilon$) такое, что всякие начальные кривые $\varphi_{t_0}(\cdot)$ из области $\|\varphi(t_0 + \cdot)\|^{(h)} < \delta(\varepsilon)$ порождают решения, удовлетворяющие неравенству $\|X(t, \varphi(t_0 + \cdot))\| < \varepsilon$ при любом $t > t_0$. Фиксируем эти числа ε и δ , и по δ определим t^* и X^* , которые по свойству 2 функции V доставляют ей положительное значение. Так как δ не зависит от t_0 , то будем считать, что $t_0 = t^*$. Таким образом, получим $V(t^*, X^*) > 0$, $\|X^*\| < \delta(\varepsilon)$, $\|X(t, \varphi(t^* + \cdot))\| < \varepsilon$, а в качестве начальной функции выберем такую $\varphi(t^* + \cdot)$, что $\varphi(t^*) = X^*$, $\|\varphi(t^* + \cdot)\|^{(h)} < \delta$ и $V(s, \varphi(s)) \leq V(t^*, X^*)$ при $s \in [t^* - h, t^*]$. Это всегда можно сделать, так как $V(t, X)$ допускает

бесконечно малый высший предел и имеет место строгое неравенство $V(t^*, X^*) > 0$. Например:

$$\varphi_{t^*}(\vartheta) = \psi_k(\vartheta) = \begin{cases} 0 & \text{при } \vartheta \in \left[t^* - h, t^* - \frac{1}{k} \right], \\ k\vartheta + 1 - kt^* & \text{при } \vartheta \in \left[t^* - \frac{1}{k}, t^* \right]. \end{cases}$$

Определенная таким образом $\psi_k(\vartheta)$ непрерывна и при $k \rightarrow \infty$ приближается к функции

$$\psi_\infty(\vartheta) = \begin{cases} 0 & \text{при } \vartheta \in [t^* - h, t^*), \\ X^* & \text{при } \vartheta = t^*. \end{cases}$$

Кроме того, для любого $\vartheta \in [t^* - h, t^*]$ и любых $k > 1$ имеем $\psi_k(\vartheta) \geq \psi_{k+1}(\vartheta)$. Следовательно,

$$U(t^*, \varphi(t^* + \cdot)) \geq \omega(\|X^*\|) > 0 \quad (185)$$

по свойству 2 функции V и по определению положительной определенности (184). Последнее неравенство непременно строгое, поскольку функция V допускает бесконечно малый высший предел, и отсюда $\|X^*\| \neq 0$, если $V(t^*, X^*) > 0$ для любого $t > t_0$. Но тогда существует $\Delta > 0$ такое, что из (185) следует

$$V(t) \geq V(t^*, X^*) + \frac{1}{2}\omega(\|X^*\|)(t - t^*)$$

при $t > t^*$, и $V(t, X(t))$ будет монотонно возрастать, пока выполняется неравенство (166) и $\|X(t)\| < H$. Однако ввиду монотонности V неравенство (166) будет выполняться всегда, а $\|X(t)\| < \varepsilon < H$ по предположению. Все это противоречит ограниченности V , которая имеет место в области (164). Тем самым теорема доказана.

Следующая теорема для системы (162) доказывается в точности так же, как и теорема 16, с теми естественными изменениями, которые отличают первую теорему Ляпунова о неустойчивости от второй.

Теорема 17. Нулевое решение системы (162) будет неустойчиво в широком смысле, если существует функция $V(t, X)$, определенная и допускающая бесконечно малый высший предел в области (164) и обладающая свойствами:

1) для любого $\varepsilon > 0$ существует точка X^* и момент $t^* > t_0$ такие, что $\|X^*\| < \varepsilon < H$ и $V(t^*, X^*) > 0$;

2) правое нижнее производное число V в силу системы (162) представимо в виде:

$$D_+V = \lambda V + W(t, X(t + \cdot)),$$

где $\lambda > 0$, а функционал W , определенный на любых кусочно-непрерывных кривых $y(t + \cdot)$ из области (163), удовлетворяет неравенству $W(t, y(t + \cdot)) \geq 0$ для всех кривых $y(t + \cdot)$, подчиненных условию (166).

Например, для уравнения $\dot{x} = ax + bx_n$, применяя теорему 10, получим следующее решение. Пусть $V = x^2/2$. Тогда $\dot{V} = ax^2 + bx_n x_n$. Условие (166) даст неравенство $|x(t - h)| \leq |x(t)|$. Оценим \dot{V} снизу, используя последнее неравенство: $\dot{V} \geq (a - |b|)x^2$. Значит, \dot{V} будет положительно определена, если $a > |b|$.

Японским математиком Дж. Като [111] сформулирована теорема, аналогичная теореме 16, и, как может показаться на первый взгляд, даже более общая. Но введенные Дж. Като абстрактные функции Ляпунова фактически ничего не обобщают, а лишь затрудняют понимание вопроса.

В отличие от теоремы 16 у Дж. Като неустойчивость понимается в смысле отрицания устойчивости, от функции Ляпунова требуется непрерывность, и к тому же отсутствует доказательство.

Теоремы 16 и 17 также можно сформулировать для неустойчивости, но тогда необходимо потребовать, чтобы при произвольном $t > 0$ и сколь угодно малом Δ существовала точка, такая, что $\|X\| < \Delta$ и $V(t, X) > 0$.

Определить свойство неустойчивости нулевого решения системы (162) с помощью функции удобнее, чем с помощью функционалов. О существовании таких функций говорит следующая теорема (которая в части достаточности является следствием теоремы 17, поэтому приводится только доказательство необходимости).

Теорема 18. Чтобы нулевое решение системы (162) было неустойчивым в широком смысле, необходимо и достаточно существования функции $V(t, X)$ со следующими свойствами:

- 1) $V(t, X)$ определена в полуцилиндре (164);
- 2) она ограничена в (164) и при любом фиксированном $t > 0$

$$\lim_{\|X\| \rightarrow 0} |V(t, X)| = 0;$$

- 3) для любого $\varepsilon > 0$ существует пара $\{t^*, X^*\}$ такая, что

$$t^* \geq 0, \quad \|X^*\| < \varepsilon, \quad V(t^*, X^*) > 0;$$

- 4) производная $V(t, X)$ в силу системы (162) имеет вид:

$$\frac{d}{dt} V(t, X) = \lambda(t, X(t + \cdot))V(t, X(t)) + W(t, X(t + \cdot)),$$

где функционалы λ и W определены вдоль всяких кусочно-непрерывных кривых $y(t + \cdot)$ из области (163) и λ непременно

ограничен снизу положительным числом α , $W \geq 0$, когда кривые $y(t + \cdot)$ удовлетворяют неравенству:

$$e^{\alpha(t-s)}V(s, y(s)) \leq V(t, y(t)) \text{ при } s \in [t-h, t]. \quad (186)$$

Доказательство. Необходимость. Пусть нулевое решение системы (163) неустойчиво в широком смысле. Определим (аналогично тому, как это было сделано В. И. Зубовым в доказательстве теоремы обращения без запаздывания) функционал $T(t_0, X_0(t_0 + \cdot))$ как первый момент выхода траектории $X(t, X_0(t_0 + \cdot))$ через боковую поверхность ε -полуцилиндра. Если траектория не выходит из полуцилиндра, то $T(t_0, X_0(t_0 + \cdot)) = \infty$. Теперь рассмотрим функцию $\tau(t_0, X_0)$, определенную в ε -полуцилиндре следующим образом:

$$\tau(t_0, X_0) = \inf\{T(t_0, Z_{t_0}(\cdot)) \text{ по } \|Z_{t_0}(\cdot)\|^{(h)} \leq \|X_0\|, Z(t_0) = X_0\} \quad (187)$$

(здесь $Z(t_0 + \cdot)$ — непрерывные кривые). Далее определим $R(t, X) = \exp\{t - \tau(t, X)\}$. Теперь доопределим при $\|X\| < H$ вне ε -полуцилиндра R по непрерывности постоянным значением — единицей. Выполнено свойство 1. Чтобы «увидеть» свойство 2, достаточно сопоставить три факта: $X(t) \equiv 0$ — решение системы (162), не выходящее из ε -полуцилиндра, определение (187) и непрерывную зависимость решений от начальных функций. Свойство 3 для функции $R(t, X)$ выполняется потому, что из неустойчивости для сколь угодно малой по норме некоторой начальной кривой $\varphi(t_0 + \cdot)$ и некоторого момента решение за конечное время $T(t_0, \varphi(t_0 + \cdot))$ покинет ε -полуцилиндр. Тогда на решении $X(t, \varphi(t_0 + \cdot))$ можно взять произвольную точку X^* и соответствующий момент времени обозначить t^* , лишь бы для X^* и t^* выполнялось соотношение $\|X(t, \varphi(t_0 + \cdot))\| \leq \|X^*\|$ при $t \in [t^* - h, t^*]$. Это всегда можно сделать, так как решение — непрерывная вектор-функция, которая по норме за время $T(t_0, \varphi) - t_0$ принимает все значения между $\|\varphi(t_0 + \cdot)\|^{(h)}$ и ε .

Теперь покажем непрерывность функции $R(t, X)$. Доказательство будет основываться на определении \inf в (187), на непрерывной зависимости решений от начальных данных.

Обозначим $Q(t, X)$ — множество непрерывных кривых $Z(t + \cdot)$ таких, что $\|Z(t + \cdot)\|^{(h)} \leq \|X\|$ и $Z(t) = X$. Так как $\tau(t, X) = \inf\{T(t, Z_t(\cdot)) \text{ при } Z(t + \cdot) \in Q(t, X)\}$, то непрерывность $\tau(t, X)$ будет установлена, если по любому положительному числу v и для произвольной кривой $Z_1(t_1 + \cdot)$ такой, что $Z_1 \in Q(t_1, X_1)$ и $T(t_1, Z_1(t_1 + \cdot)) < \infty$, найдется положительное число, для которого при $|t_1 - t_2| + \|X_1 - X_2\| < \delta$ существует кривая $Z_2(t_2 + \cdot) \in Q(t_2, X_2)$, удовлетворяющая неравенству $|T(t_1, Z_1(t_1 + \cdot)) - T(t_2, Z_2(t_2 + \cdot))| < v$. Фиксируем некоторую описанную выше непрерывную $Z_1(t_1 + \cdot)$ и соответствующее

число $T(v, Z_1)$. Для любого $\eta > 0$ можно подобрать положительное число μ , которое зависит от η и $T(t_1, Z_1(t_1 + \cdot))$ и для которого верна импликация

$$\|Z_1(t_1 + \cdot) - Z(t_1 + \cdot)\|^{(h)} < \mu \rightarrow \|X(T, Z_1(t_1 + \cdot)) - X(T, Z(t_1 + \cdot))\| < \eta.$$

Здесь $Z(t_1 + \cdot)$ — произвольная непрерывная кривая. Далее предположим, что траектория $X(t, Z_1(t_1 + \cdot))$ проходит боковую поверхность ε -полуцилиндра с нулевым углом наклона, т. е. скалярное произведение

$$X'(T, Z_1(t_1 + \cdot))\dot{X}(T, Z_1(t_1 + \cdot)) = \frac{d}{dt} \left[\frac{1}{2} \|X(T, Z_1(t_1 + \cdot))\|^2 \right] \geq \beta > 0$$

(штрих означает транспонирование). Это предположение не ограничивает общности рассмотрения, так как в определении неустойчивости можно считать его выполненным, полагая при необходимости указанное в нем ε чуть меньшим. Тогда по непрерывности $F(t, X(t + \cdot))$ будем иметь неравенство $X'(T, Z(t_1 + \cdot))F(T, X(T + \cdot, Z)) \geq \beta/2 > 0$, если только η достаточно мало. И, следовательно, траектория $X(t, Z(t_1 + \cdot))$ выйдет из ε -полуцилиндра не позже момента $T(t_1, Z_1(t_1 + \cdot)) + 2\eta/\beta$. Получим неравенство $|T(t_1, Z(t_1 + \cdot)) - T(t_1, Z_1(t_1 + \cdot))| < 2\eta/\beta$. По числу μ , очевидно, можно подобрать положительное число δ такое, что при $|t_1 - t_2| + \|X_1 - X_2\| < \delta$ существует кривая

$$Z_2(t_2 + \cdot) \in (Q(t_2, X_2)) / (\|Z_2(t_2 + \cdot) - Z_1(t_1 + \cdot)\|^{(h)} < \mu).$$

Таким образом, по произвольному v определили $\eta = \beta v/2$, и по η получили μ , а по μ нашли δ . Так как $\mu = \mu(\eta, T)$, то $\delta = \delta(v, Z_1(t_1 + \cdot))$.

Предположим теперь, что $\tau(t, X)$ разрывна в (t_1, X_1) , т. е. при $t \rightarrow t_1$ и $X \rightarrow X_1$ имеем для некоторого положительного v_1 неравенство $|\tau(t_1, X_1) - \tau(t, X)| > v_1$. Пусть, например, $\tau_1 < \tau$. Тогда по определению τ как $\inf(T)$ найдется T^* , разделяющее числа $\tau_1 = \tau(t_1, X_1)$ и $\tau = \tau(t, X)$. Например, $T^* = (\tau(t_1, X_1) - \tau(t, X))/2$. Тогда подберем $v < v_1/2$ и по только что доказанному выше свойству функционала T найдем $T_2(t_2, X_2)$, меньшее $\tau(t_2, X_2)$ при $|t_1 - t_2| + \|X_1 - X_2\| < \delta(v_1, T^*)$. Но это противоречит определению $\tau(t_2, X_2)$ как $\inf(T_2)$. Следовательно, $\tau(t, X)$ непрерывна.

Далее по построенной непрерывной функции $R(t, X)$ определим среднюю функцию $V(t, X)$, удовлетворяющую свойствам 1–4. Для этого обозначим

$$\omega(t, X; \xi, \zeta) = \begin{cases} C \exp\left(\frac{r^2}{r^2 - 1}\right) & \text{при } r = \sqrt{(t - \xi)^2 + (X - \zeta)^2} < 1, \\ 0 & \text{при } r \geq 1, \end{cases}$$

где C выбирается так, чтобы $\int \omega(t, X; \xi, \zeta) dt dX = 1$ (интегрирование ведется по области (164)). Для произвольного положительного ρ введем

$$\omega_\rho(t, X, \xi, \zeta) = \omega\left(\frac{t}{\rho}, \frac{1}{\rho} X; \frac{1}{\rho} \xi, \frac{1}{\rho} \zeta\right).$$

Очевидно, что $\omega_\rho = 0$ при $r \geq \rho$:

$$\int \omega_\rho(t, X; \xi, \zeta) dt dX = \rho^{n+1} = \rho_1$$

и $\omega_\rho(t, X; \xi, \zeta)$ имеет непрерывные производные всех порядков по всем аргументам.

Рассмотрим среднюю функцию

$$V_\rho(t, X) = \frac{1}{\rho_1} \int R(\xi, \zeta) \omega_\rho(t, X; \xi, \zeta) d\xi d\zeta. \quad (188)$$

Эта функция определена в полуцилиндре (164), как и функция $R(t, X)$. При $\rho \rightarrow 0$ функция $V_\rho(t, X)$ равномерно сходится к $R(t, X)$ и, кроме того, $V_\rho(t, X)$ имеет непрерывные производные любых порядков при $\rho > 0$. Из свойств функции $R(t, X)$ и определения (188) следует, что $V_\rho(t, X)$ удовлетворяет свойствам 2 и 3 теоремы.

Рассмотрим свойство 4. Из определения функции $\omega_\rho(t, X; \xi, \zeta)$ имеем, что производную в силу системы (162) можно представить в виде:

$$\frac{d}{dt} \omega_\rho(t, X(t); \xi, \zeta) = \sigma(t, X(t+\cdot); \xi, \zeta) \omega_\rho(t, X(t); \xi, \zeta).$$

Здесь функционал σ является непрерывной функцией параметров ξ, ζ . Тогда производная V_ρ в силу системы (162) может быть записана в виде:

$$\begin{aligned} \frac{d}{dt} V_\rho(t, X(t)) &= \frac{1}{\rho_1} \int R(\xi, \zeta) \frac{d}{dt} \omega_\rho(t, X(t); \xi, \zeta) d\xi d\zeta = \\ &= \sigma(t, X(t+\cdot); \xi_1, \zeta_1) V_\rho(t, X(t)), \end{aligned}$$

так как справедлива теорема о среднем значении подынтегральной функции. Так как при $\rho \rightarrow 0$ имеют место пределы $\xi_1 \rightarrow t$, $\zeta_1 \rightarrow X(t)$, то при достаточно малом фиксированном $\rho > 0$ можно обозначить $\lambda_\rho(t, X(t+\cdot)) = \sigma(t, X(t+\cdot)); \xi_1, \zeta_1$. Теперь для доказательства свойства 4 осталось положить $W \equiv 0$ и из неравенства (186) получить ограничение $\lambda_\rho \geq \alpha > 0$. Но это очевидно, так как из (186) следует

$$\frac{V_\rho(t, y(t)) - V_\rho(s, y(s))}{t - s} \geq \alpha V_\rho(t, y(t)) e^{\alpha(s-t)}.$$

Если положить $y(t) = X(t)$, то при $s \rightarrow t$ для левой нижней производной имеем

$$D_-V_\rho(t, X(t)) = \lambda_\rho(t, y(t + \cdot))V_\rho(t, X(t)) \geq \alpha V_\rho(t, X(t))$$

для тех $y(t)$, которые удовлетворяют (186). Теорема доказана.

В качественной теории дифференциальных и дифференциально-разностных уравнений большое значение имеют теоремы об устойчивости по первому приближению. Ниже мы приведем практически без изменений соответствующий материал Н. Н. Красовского [48]. При этом мы несколько повторим вводные предположения, чтобы оставить оригинальные обозначения Н. Н. Красовского (часто приходится слышать, что последние удобнее используемых нами выше).

Рассмотрим систему уравнений с последствием

$$dx_i/(dt) = F_i(x_1(t + \vartheta), \dots, x_n(t + \vartheta), t), \tag{189}$$

где функционалы F_i удовлетворяют следующим условиям:

1) функционалы F_i определены и кусочно-непрерывны в области $\|x\|^{(h)} < H, t \geq 0$ (где H — фиксированная постоянная) в следующем смысле: существует последовательность чисел t_k ($k = 0, 1, 2, \dots$) таких, что в каждой области $\|x(\vartheta)\|^{(h)} < H, t_k \leq t < t_{k+1}$ функционалы F_i непрерывны по t и могут быть доопределены (с сохранением непрерывности) во всей $\|x\|^{(h)} < H, t_k \leq t < t_{k+1}$ области так, что для каждого $t^* \in [t_k, t_{k+1}]$ и для каждой непрерывной функции $x_i^*(\vartheta)$ ($i = 1, \dots, n$) для любого наперед заданного числа $\varepsilon > 0$ можно указать $\delta > 0$ такое, что

$$\begin{aligned} &|F_i(x_1^*(\vartheta), \dots, x_n^*(\vartheta), t^*) - F_i(x_1^*(\vartheta), \dots, \\ & \quad x_n^*(\vartheta), t)| < \varepsilon \quad (i = 1, \dots, n), \end{aligned}$$

если только $|t^* - t| < \delta$ и $t \in [t_k, t_{k+1}]$; так как функционалы F_i могут претерпевать разрывы в отдельные моменты времени t_k , и по ряду других причин, будем в уравнениях везде под $dx_i/(dt)$ понимать лишь правую производную от x по времени t ;

2) функционалы F_i удовлетворяют условиям Липшица по x , т. е.

$$\begin{aligned} &|F_i(x_1''(\vartheta), \dots, x_n''(\vartheta), t) - F_i(x_1'(\vartheta), \dots, x_n'(\vartheta), t)| < L \|x'' - x'\|^{(h)}, \\ &(L = \text{const}, \quad i = 1, \dots, n). \end{aligned}$$

Норму введем следующим образом:

$$\|x\|^{(h)} = \sup |x_i(\vartheta)| \quad \text{при } -h \leq \vartheta \leq 0, \quad i = 1, \dots, n).$$

Как обычно, предполагаем, что $F_i(x(\vartheta), t) = 0$ при $x_j(\vartheta) \equiv 0$ $i = 1, \dots, n, j = 1, \dots, n, -h \leq \vartheta \leq 0$.

Лемма 1. Если решения $x(x_0(\vartheta_0), t_0, t)$ системы (189) удовлетворяют условию:

$$\|x(x_0(\vartheta_0), t_0, t)\|^{(h)} \leq B \|x_0(\vartheta_0)\|^{(h)} \exp(-\alpha(t - t_0)) \quad \text{при } t \geq t_0, \quad (190)$$

$$\|x_0(\vartheta_0)\|^{(h)} < H_0 = H/B, \quad (191)$$

то в области (191) можно построить функционал $V(x(\vartheta), t)$, удовлетворяющий следующим условиям:

$$c_1 \|x(\vartheta)\|^{(h)} \leq V(x(\vartheta), t) \leq c_2 \|x(\vartheta)\|^{(h)}, \quad (192)$$

$$\limsup_{\Delta t \rightarrow +0} \left(\frac{\Delta V}{\Delta t} \right)_{(189)} \leq -c_3 \|x(\vartheta)\|^{(h)}, \quad (193)$$

$$|V(x''(\vartheta), t) - V(x'(\vartheta), t)| \leq c_4 \|x''(\vartheta) - x'(\vartheta)\|^{(h)} \quad (194)$$

(c_1, \dots, c_4 — положительные постоянные).

Доказательство. Покажем, что функционал

$$V(x_0(\vartheta_0), t_0) = \int_{t_0}^{t_0+T} \|x(x_0(\vartheta_0), t_0, \xi + \vartheta)\|^{(h)} d\xi + \\ + \sup(\|x(x_0(\vartheta_0), t_0, \xi + \vartheta)\|^{(h)} \quad \text{при } t_0 \leq \xi \leq t_0 + T) \left(T = \frac{1}{\alpha} \ln B \right) \quad (195)$$

удовлетворяет всем условиям теоремы. В самом деле, при начальных кривых $x_0(\vartheta_0)$ из области (191) вследствие условия (190) решения $x(x_0(\vartheta_0), t_0, t)$ при $t \geq t_0$ лежат в области

$$\|x(\vartheta)\|^{(h)} < H,$$

где определены правые части системы (189), и, следовательно, эти решения $x(x_0(\vartheta_0), t_0, t)$ продолжимы при всех $t \geq t_0$, т. е. функционал (195) определен в области (191). Левое неравенство (192) (при $c_1 = 1$) следует сразу из очевидного соотношения

$$\sup(\|x(x_0(\vartheta_0), t_0, \xi + \vartheta)\|^{(h)} \quad \text{при } t_0 \leq \xi \leq t_0 + T) \geq \|x_0(\vartheta_0)\|^{(h)}.$$

Первое неравенство (192) следует по условиям (190) из соотношений

$$V(x_0(\vartheta_0), t_0) \leq \int_{t_0}^{t_0+T} B \|x_0(\vartheta_0)\|^{(h)} \exp[-\alpha(t - t_0)] dt + \|x_0(\vartheta_0)\|^{(h)} B = \\ = B \left(1 + \frac{1}{\alpha} \left(1 - \frac{1}{2B} \right) \right) \|x_0(\vartheta_0)\|^{(h)}.$$

Вследствие условия (190) второе слагаемое в правой части равенства (195) не возрастает с ростом t_0 при движении вдоль траектории системы (189), поэтому

$$\limsup_{\Delta t \rightarrow +0} \left(\frac{\Delta V}{\Delta t} \right)_{\substack{\text{при } t=t_0 \\ \text{в силу (189)}}} \leq \\ \leq \lim_{\Delta t \rightarrow +0} \left(\int_{t_0+\Delta t}^{t_0+T+\Delta t} \|x(x_0(\vartheta_0), t_0, \xi + \vartheta)\|^{(h)} d\xi - \int_{t_0}^{t_0+T} \|x(x_0(\vartheta_0), t_0, \xi + \vartheta)\|^{(h)} d\xi \right) = \\ = -\|x_0(\vartheta_0)\|^{(h)} + \|x(x_0(\vartheta_0), t_0, t_0+T+\vartheta)\|^{(h)},$$

и теперь по условиям (190) и по выбору числа T получим

$$\limsup_{\Delta t \rightarrow +0} \left(\frac{\Delta V}{\Delta t} \right)_{\substack{\text{при } t=t_0 \\ \text{в силу (189)}}} \leq -\frac{1}{2} \|x_0(\vartheta_0)\|^{(h)},$$

что и доказывает неравенство (193) (при $c_3 = 1/2$).

Проверим выполнение неравенства (194). Оценивая интеграл и используя неравенства (190), получим:

$$\begin{aligned} & |V(x_0''(\vartheta_0), t_0) - V(x_0'(\vartheta_0), t_0)| \leq \\ & \leq \int_{t_0}^{t_0+T} \|x(x_0''(\vartheta_0), t_0, \xi + \vartheta) - x(x_0'(\vartheta_0), t_0, \xi + \vartheta)\|^{(h)} d\xi + \\ & + \sup(\|x(x_0''(\vartheta_0), t_0, \xi + \vartheta) - x(x_0'(\vartheta_0), t_0, \xi + \vartheta)\|^{(h)} \text{ при } t_0 \leq \xi \leq t_0+T) \leq, \end{aligned}$$

что и доказывает неравенство (194). Лемма доказана.

Следствием леммы 1 является следующее утверждение.

Теорема 19. Пусть $F_i(x_1(\vartheta), \dots, x_n(\vartheta))$ — линейные функционалы. Если корни «характеристического» уравнения

$$\begin{vmatrix} F_{11} - \lambda & \dots & F_{1n} \\ \dots & \dots & \dots \\ F_{n1} & \dots & F_{nn} - \lambda \end{vmatrix} = 0 \tag{196}$$

$$(F_{ij} = F_i(0, \dots, e^{\lambda \vartheta}, \dots, 0))$$

удовлетворяют неравенству

$$\operatorname{Re} \lambda < -\gamma \quad (\gamma > 0 - \text{const}), \tag{197}$$

то существует функционал $V(x(\vartheta), t)$, удовлетворяющий условиям (192), (193), (194).

П р и м е ч а н и е. Как мы видели при доказательстве леммы 1, существование функционала, удовлетворяющего оценкам (192), (193), (194), связано не столько с фактом линейности функционалов F_i , сколько определяется характером (190) приближения траекторий к точке с возрастанием времени. Следует, впрочем, отметить, что неравенство (190) характерно как раз для линейных систем, правые части которых не зависят явно от времени.

Теорема 19 позволяет доказать для уравнений с последствием теоремы об устойчивости по первому приближению. Докажем здесь одну теорему такого рода [48].

Рассмотрим систему уравнений

$$dx_i/(dt) = F_i(x_1(\vartheta), \dots, x_n(\vartheta)) + R_i(x_1(\vartheta), \dots, x_n(\vartheta), t), \quad (198)$$

где F_i — линейные функционалы, R_i — некоторые непрерывные функционалы. Предположим, что функционалы R_i удовлетворяют неравенству

$$|R_i(x_1(\vartheta), \dots, x_n(\vartheta), t)| \leq \beta \|x(\vartheta)\|^{(h)}. \quad (199)$$

Теорема 20. Если корни λ уравнения (196) удовлетворяют неравенству (197), то можно указать постоянную $\beta > 0$ такую, что решение $x = 0$ системы (189) будет асимптотически устойчивым при любом выборе непрерывных функционалов R_i , удовлетворяющих условиям (199).

Доказательство. При условиях теоремы 20 согласно теореме 19 существует функционал $V(x(\vartheta), t)$, удовлетворяющий условиям (192), (193), (194). Вычислим

$$\limsup_{\Delta t \rightarrow +0} \left(\frac{\Delta V}{\Delta t} \right)_{(57)}$$

вдоль траекторий системы (198):

$$\begin{aligned} & \limsup_{\Delta t \rightarrow +0} \left(\frac{\Delta V}{\Delta t} \right)_{(198)} = \\ & = \limsup_{\Delta t \rightarrow +0} \left(\frac{1}{\Delta t} [V(x^{(2)}(t + \Delta t + \vartheta), t + \Delta t) - V(x^{(2)}(t + \vartheta), t)] \right) + \\ & + \limsup_{\Delta t \rightarrow +0} \left(\frac{1}{\Delta t} [V(x^{(2)}(t + \Delta t + \vartheta), t + \Delta t) - V(x^{(1)}(t + \Delta t + \vartheta), t + \Delta t)] \right) \leq \\ & \leq -c_3 \|x(t + \vartheta)\|^{(h)} + c_4 \limsup_{\Delta t \rightarrow +0} \left(\frac{1}{\Delta t} [\|x^{(2)}(t + \Delta t + \vartheta) - x^{(1)}(t + \Delta t + \vartheta)\|^{(h)}] \right). \end{aligned}$$

Здесь для краткости приняты обозначения: $x^{(2)}(t)$ — траектория системы (198), и $x^{(1)}(t)$ — траектория системы (189). Из уравнений (198), учитывая неравенство (199), получим теперь окончательную оценку:

$$\limsup_{\Delta t \rightarrow +0} \left(\frac{\Delta V}{\Delta t} \right)_{(198)} \leq (-c_3 + c_4 \beta) \|x(t + \vartheta)\|^{(h)}.$$

Если число $\beta > 0$ удовлетворяет условию

$$\beta < c_3/c_4,$$

то величина

$$\lim_{\Delta t \rightarrow +0} \sup \left(\frac{\Delta V}{\Delta t} \right)_{(198)}$$

будет определенно-отрицательным функционалом, т. е. функционал $V(x(\theta), t)$ будет удовлетворять всем условиям Ляпунова об асимптотической устойчивости вдоль траекторий системы (198), откуда и следует асимптотическая устойчивость решения $x = 0$. Теорема доказана.

В частном случае уравнений с запаздываниями

$$\frac{dx_i}{dt} = \sum_{j=1}^n a_{ij}x_j(t) + b_{ij}x_j(t - h_{ij}) + R_i(x_1, \dots, x_n, x_1(t - h_{ij}^*(t)), \dots, \dots, x_n(t - h_{ij}^*(t)), t) \tag{200}$$

($h_{ij} = \text{const}$, $0 \leq h_{ij} \leq h$, $0 \leq h_{ij}^*(t) \leq h$), теорема 20 читается следующим образом.

Если корни λ «характеристического» уравнения

$$\begin{vmatrix} a_{11} + b_{11}e^{-\lambda h_{11}} - \lambda & \dots & a_{1n} + b_{1n}e^{-\lambda h_{1n}} \\ \dots & \dots & \dots \\ a_{n1} + b_{n1}e^{-\lambda h_{n1}} & \dots & a_{nn} + b_{nn}e^{-\lambda h_{nn}} - \lambda \end{vmatrix} = 0$$

удовлетворяют неравенству

$$\text{Re} \lambda < -\gamma \quad (\gamma > 0 - \text{const}),$$

то можно указать постоянную $\beta > 0$ такую, что решение $x = 0$ системы (200) будет асимптотически устойчивым, каковы бы ни были непрерывные функции $R_i(x_1, \dots, x_n, y_1, \dots, y_n, t)$ и запаздывания $h_{ij}^*(t)$, удовлетворяющие неравенству

$$|R_i(x_1, \dots, x_n, y_1, \dots, y_n, t)| \leq \beta(\|x\| + \|y\|).$$

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ В СЛУЧАЕ ДИНАМИЧЕСКОГО ПРОЦЕССА С ДИСКРЕТНЫМ ВРЕМЕНЕМ

В данном параграфе метод функций Ляпунова с условием отрицательной определенности Разумихина применяется к разностным системам.

Дискретное время часто используется в экономических моделях, поскольку статистические данные собираются по известным правилам через фиксированные интервалы времени. Тогда возникает ситуация, в которой математическую модель системы удобно записывать в форме разностного уравнения

$$X_{k+1} = F(X_k, X_{k-1}, X_{k-2}, \dots, X_{k-m}), \quad (201)$$

где $X_k = \{x_{ki}\}$, $i = 1, \dots, n$; $k = 0, 1, \dots, m \geq 1$.

Такой же вид получит непрерывная система после применения какого-то численного метода интегрирования или другой дискретизации.

Векторное уравнение (201) имеет решение $\{X_k\}$, если заданы начальные векторы $X_0, X_{-1}, \dots, X_{-m}$ и при этом F — однозначная векторная функция, определенная в соответствующей области пространства $E^{n(m+1)}$. Решение продолжимо при $k > 0$, если X_k не выходят за границу области определения F . Если $F(0, \dots, 0) = 0$, то система (201) имеет тривиальное решение. Предположим, что это так и, кроме того, F удовлетворяет условию Липшица по всем аргументам.

Определение 2. Тривиальное решение системы (201) называется устойчивым по Ляпунову, если для любого $\varepsilon > 0$ существует $\delta > 0$, такое, что для любых $X_{-m}, X_{1-m}, \dots, X_0$, отвечающих неравенству

$$\sum_{i=-m}^0 \|X_i\| < \delta,$$

будет верно неравенство $\|X_k\| < \varepsilon$ для всех $k > 0$. Если к тому же существует $\Delta > 0$, такое, что для любых X_{-m}, \dots, X_0 , удовлетворяющих условию

$$\sum_{i=-m}^0 \|X_i\| < \Delta,$$

будет $\lim_{k \rightarrow \infty} \|X_k\| = 0$, то говорят, что имеется асимптотическая устойчивость.

Известно, что для анализа системы (201) на устойчивость применяются функции Ляпунова, а именно справедлива, например, следующая теорема.

Теорема 21. Если существует определенно-положительная функция $V_k = V(X_k, X_{k-1}, \dots, X_{k-m})$, $k = 0, 1, \dots$, такая, что $U_k = V_{k+1} - V_k$ (разность) неположительна, то нулевое решение системы (201) устойчиво по Ляпунову.

Обратим здесь внимание на то, что функции V_k и U_k имеют $n(m+1)$ аргументов.

В предыдущем параграфе для систем с последствием приведен тип теорем, в котором знак производной функции Ляпунова проверяется не на всей области определения, а лишь на некотором классе функций. Применим эту идею для системы (201). Будем считать, что (201) — это разностная система первого по-

рядка с m запаздываниями, k — текущее значение независимой переменной, а $k - 1, \dots, k - m$ — запаздывающие аргументы. Тогда оказывается справедливым утверждение [74].

Теорема 22. Если существует такая определенно-положительная функция $V = V(X)$, что разность

$$U(X_k, \dots, X_{k-m}) = V(X_{k+1}) - V(X_k)$$

есть неположительная функция произвольных X_k и всяких X_j , удовлетворяющих неравенству

$$V(X_j) \leq V(X_{k+1}), \quad j \in [k - m, k], \quad (202)$$

то нулевое решение (201) устойчиво по Ляпунову.

Доказательство. Рассмотрим произвольное малое положительное число ε и по нему построим число

$$c = \inf\{V(X), \varepsilon \leq \|X\| \leq \varepsilon + D\}.$$

Здесь $D > 0$ зависит от ε и F . Так как $V(X)$ определенно-положительна, то такое c обязательно существует. Рассмотрим произвольное $0 < c_1 < c$. Из непрерывности V и соответствия F условию Липшица следует, что существует такое число $\delta = \delta(c_1)$, что для набора векторов X_{-m}, \dots, X_0 , удовлетворяющих

$$\sum_{i=-m}^0 \|X_i\| < \delta,$$

будет $V(X_i) < c_1$ при $i = 0, 1, \dots, m$. Тогда и далее при $i > m$ будет $V(X_i) < c_1$. Пусть это не так. Тогда существует номер $l > m$, такой, что $V(X_l) \geq c_1$. Не умаляя общности, можно считать, что $V(X_l) = c_1$, так как c_1 достаточно произвольно. Далее, очевидно, что либо $V(X_{l+1}) > c_1$, либо $V(X_{l+1}) \leq c_1$. Второй случай считаем неинтересным. В первом случае имеем $V(X_{l+1}) - V(X_l) > 0$. Но по условию теоремы $U(X_l, X_{l-1}, \dots, X_{l-m}) \leq 0$, так как $V(X_i) \leq V(X_{l+1})$ при $i \leq l$. Это противоречие и доказывает, что $V(X_k) \leq c_1$ для всех k . Предположим теперь, что по найденным ε и δ решение X_k не удовлетворяет неравенству $\|X_k\| < \varepsilon$ при всех k , т. е. нулевое решение неустойчиво. Тогда существует номер \bar{k} такой, что $\|X_{\bar{k}}\| \geq \varepsilon$. Будем считать D настолько большим, чтобы $\|X_{\bar{k}}\| \leq \varepsilon + D$. Это возможно, так как F удовлетворяет условию Липшица и $\|X_i\| < \varepsilon$ при $i < \bar{k}$. Отсюда следует, что $\varepsilon \leq \|X_{\bar{k}}\| \leq \varepsilon + D$ и $V(X_{\bar{k}}) \geq c$. Но уже доказано, что этого не может быть. Значит, нулевое решение устойчиво.

Пример 1. Рассмотрим скалярное разностное уравнение третьего порядка

$$x_{k+1} = x_k^2 x_{k-1} + \alpha x_{k-2}, \quad k = 2, 3, \dots \quad (203)$$

Исследуем его нулевое решение на устойчивость с помощью теоремы. Пусть $V(x) = x^2$. Тогда из неравенства Разумихина (202) получим:

$$|x_{k-1}| \leq |x_k|, \quad |x_{k-2}| \leq |x_k|.$$

Оценим разность $V(x_{k+1}) - V(x_k)$, учитывая последние неравенства:

$$\begin{aligned} U(x_k, x_{k-1}, x_{k-2}) &= x_k^4 x_{k-1}^2 + 2\alpha x_k^2 x_{k-1} x_{k-2} - x_k^2 \leq \\ &\leq x_k^2 [x_k^2 + 2|\alpha| x_k^2 + (\alpha^2 - 1)]. \end{aligned}$$

Очевидно, что при $|\alpha| < 1$ существует окрестность нуля

$$\left\{ x : |x| < \sqrt{1 - |\alpha|} \right\},$$

в которой наблюдается устойчивость.

Рассмотрим теперь свойство асимптотической устойчивости нулевого решения системы (201).

Теорема 23. Если существует такая определенно-положительная функция $V = V(X)$, что разность

$$U(X_k, X_{k-1}, \dots, X_{k-m}) = V(X_{k+1}) - V(X_k)$$

удовлетворяет неравенству $U \leq -\omega(\|X_k\|)$ для всяких X_i , таких, что

$$V(X_i) \leq f(V(X_{k+1})), \quad k - m \leq i \leq k, \quad (204)$$

то нулевое решение системы (201) асимптотически устойчиво. Здесь $\omega = \omega(r)$ — определенно-положительная функция скалярного аргумента, f — непрерывная монотонно возрастающая функция, удовлетворяющая неравенству $f(r) > r$.

Доказательство. Ясно, что если нашлась такая функция $V = V(X)$, то она удовлетворяет условиям теоремы 22, и, следовательно, имеет место устойчивость по Ляпунову. Осталось доказать, что решения из некоторой окрестности нуля асимптотически приближаются к нулю.

Из доказательства теоремы 22 видно, что последовательность $V_k = V(X_k)$ ограничена сверху числом c и имеет положительные члены. Построим убывающую мажоранту W_k этой последовательности. Из условия теоремы последовательность V_k не может иметь промежутков возрастания длиной более m , так как всякий интервал монотонного возрастания (и даже неубывания) согласно условию теоремы должен закончиться строгим убыванием. Значит, последовательность V_k либо глобально убывает, т. е. так, что $V_k > V_j$ для всех $j > k$, либо колеблется, но обязательно так, что для всякого локального максимума с номером \bar{k} имеет место

неравенство $V_k > V_{\bar{k}}$, где $k \in (\bar{k} - m, \bar{k})$. Отсюда следует существование убывающей мажоранты $W_k: W_k \geq V(X_k)$.

Если мажоранта убывает до нуля, то и $V_k \rightarrow 0$ при $k \rightarrow \infty$. Предположим, что мажоранта стремится к некоторому ненулевому значению $W_k \rightarrow W_0$. Тогда существует последовательность $\{V_{k_i}\}$, такая, что $V_{k_i} \rightarrow V_0$ при $i \rightarrow \infty$ и $V_{k_i} > V_0$. Из $\{k_i\}$ выберем подпоследовательность $\{k_{i_j}\}$, такую, что

$$\lim_{j \rightarrow \infty} X_{k_{i_j}} = \bar{X} \neq 0$$

и $V(\bar{X}) = V_0$. Возможность этого следует из неравенства $\|X_k\| < \varepsilon$. Без ограничения общности считаем, что $\{k_{i_j}\} = \{k_i\}$. Построенная последовательность обладает свойством: для малого δ существует такой номер \bar{k} , что а) $W_{\bar{k}-m} - W_{\bar{k}} < \delta$; б) $V(X_{\bar{k}}) > V(X_{\bar{k}-1})$; такой номер \bar{k} обязательно существует, потому что иначе V_k всегда при больших номерах не возрастает.

Далее, из а) следует при $i \in [\bar{k} - m, \bar{k} - 1]$

$$V(X_i) \leq V(X_{\bar{k}}) + \delta < f(V(X_{\bar{k}})).$$

Второе неравенство основывается на свойстве $f(r) > r$. Из неравенства (204) получаем $V(X_{\bar{k}}) < V(X_{\bar{k}-1})$, что противоречит б). Последнее неравенство будет непременно строгим, так как $X_{\bar{k}}$ при больших \bar{k} близко к $X_0 \neq 0$, а тогда $\omega(\|X_{\bar{k}}\|) \neq 0$.

Тем самым установлено, что мажоранта W_k монотонно убывает до нуля и, следовательно, $X_k \rightarrow 0$ при $k \rightarrow \infty$, т. е. имеет место асимптотическая устойчивость.

Профессора А. Ю. Александров и А. П. Жабко, обсуждая теоремы 22 и 23, пришли к выводу, что в следующей формулировке предлагаемый метод будет более конструктивен и широко применим.

Теорема 24. Нулевое решение системы (201) асимптотически устойчиво по Ляпунову, если:

- 1) существует определенно-положительная функция $V = V(X)$;
- 2) существует непрерывная, монотонно возрастающая функция f , удовлетворяющая неравенству $f(r) > r$ при $r > 0$ и $f(0) = 0$;
- 3) существует малое число $\delta > 0$ и неотрицательное целое l , и выполнено условие

$$f(V(X_{k+1})) \leq \max\{V(X_{k-j}) \mid j = 0, \dots, m + l\}$$

на множестве

$$\{X_s: \|X_s\| \leq \delta \cap \{V(X_s) \leq \min_{\|Z\|=\delta} V(Z)\}, \\ s = k - m - l, \dots, k - l\},$$

а при $s = k - l + 1, \dots, k + 1$ значения X_s определены системой (201).

Пример 2. Рассмотрим асимптотическую устойчивость тривиального решения уравнения

$$x_{k+1} = x_k x_{k-1}.$$

При $|x_k| < 1$ решения асимптотически стремятся к нулю, $x_k = 1$ — второе положение равновесия. Для иллюстрации теоремы 23 в качестве $V(x)$ рассмотрим x^2 , и $f(r) = \gamma r$, $\gamma > 1$. Тогда

$$V(x_{k+1}) - V(x_k) = x_k^2 x_{k-1}^2 - x_k^2 = x_k^2 (x_{k-1}^2 - 1).$$

Из равенства (204) следует, что $|x_{k-1}| < |x_k|$, и тогда при $|x_k| < 1$ получим определенную отрицательность разности.

Для линейного разностного уравнения имеется критерий Шура–Кона, с помощью которого можно построить область устойчивости в пространстве параметров.

Приведем сравнение точной границы области с полученной по теореме 22.

Пусть уравнение имеет вид:

$$x_{k+1} = -ax_k - bx_{k-1}.$$

Рассмотрим функцию Ляпунова $V(x) = x^2$. Тогда неравенство Разумихина определит множество x_{k-1} , на котором проверяется знак первой разности: $|x_{k-1}| \leq |x_k|$. Составим разность:

$$V_{k+1} - V_k = a^2 x_k^2 + 2abx_k x_{k-1} + b^2 x_{k-1}^2 - x_k^2.$$

Оценивая x_{k-1} по модулю, получаем:

$$V_{k+1} - V_k \leq (a^2 + b^2 + 2|ab| - 1)x_k^2.$$

Очевидно, что разность отрицательна, когда $|a| + |b| \leq 1$. На рисунке 49 плоскости параметров эта область заштрихована дважды. Один раз заштрихована точная область устойчивости.

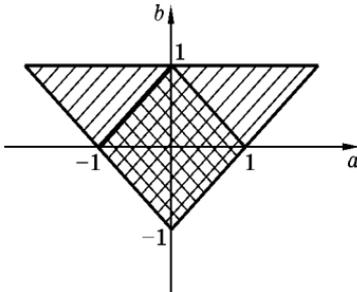


Рис. 49

Оценка области устойчивости на примере простейшего линейного уравнения второго порядка

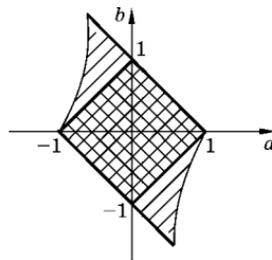


Рис. 50

Оценка области устойчивости на примере линейного уравнения третьего порядка

Рассмотрим уравнение третьего порядка, но оставим два параметра, чтобы можно было изобразить область на рисунке:

$$\begin{aligned} x_{k+1} &= -ax_k - bx_{k-2}, \\ k &= 2, 3, \dots \end{aligned}$$

Повторяя рассуждения, и здесь получим ту же область $|a| + |b| \leq 1$.

Точная область, заштрихованная на рисунке 50 один раз, согласно критерию Шура–Кона вычисляется по формулам:

$$|a + b| < 1, \quad |ab| < 1 - b^2.$$

Таким образом, теоремы 22 и 23 для линейных систем достаточно эффективны.

ЗАДАЧА СТАБИЛИЗАЦИИ ЛИНЕЙНЫХ СИСТЕМ С ПОСЛЕДЕЙСТВИЕМ

Рассмотрим теперь экономическую систему, у которой есть эндогенные и экзогенные переменные. Будем последние считать переменными регулирования, а про закон управления предположим, что он формируется на основе сигнала, не совпадающего по времени происхождения с временем исполнения. Пусть

$$\dot{Y} = AY + Bu, \quad (205)$$

где $A = \{a_{ij}\}$, $B = \{b_i\}$ — постоянные матрица и вектор. Будем искать стабилизирующее управление в виде:

$$u = \sum_{i=1}^n c_i y_i(t - \tau), \quad (206)$$

при условии, что c_i и $\tau > 0$ — постоянные числа.

Возникает вопрос — при какой матрице A и при каком векторе B возможно подобрать коэффициенты усиления c_1, \dots, c_n прямого регулятора (206) так, чтобы программное движение было асимптотически устойчивым? Ответ на этот вопрос дает теорема В. И. Зубова [32].

Теорема 25. Если векторы $B, AB, \dots, A^{n-1}B$ линейно независимы, то при любом достаточно малом запаздывании $\tau > 0$ существуют такие коэффициенты усиления c_1, \dots, c_n , при которых программное движение будет асимптотически устойчивым.

Доказательство. В системе (205) сделаем замену переменных $Y = SY$, где $S = \{B, AB, \dots, A^{n-1}B\}$.

Тогда вместо (205) запишем:

$$\begin{cases} \dot{\bar{y}}_1 = -p_n \bar{y}_n + u, \\ \dot{\bar{y}}_s = \bar{y}_{s-1} - p_{n-s+1} \bar{y}_n, \quad s = 2, \dots, n, \end{cases} \quad (207)$$

где числа p_s являются коэффициентами характеристического полинома матрицы A :

$$\lambda^n + p_1 \lambda^{n-1} + \dots + p_n = 0.$$

Будем искать управление в виде линейной комбинации переменных $\bar{y}_s(t - \tau)$, т. е.

$$u = \sum_{i=1}^n \gamma_i \bar{y}_i(t - \tau). \quad (208)$$

Характеристический квазиполином системы (207) с управлением (208) может быть записан в виде:

$$\Delta(\lambda, \tau) = (-1)^n \left[P_n(\lambda) - \sum_{k=1}^n \gamma_k e^{-\lambda \tau} P_{n-k}(\lambda) \right].$$

Пусть $\mu_1(\tau), \mu_2(\tau), \dots$ — корни уравнения

$$\Delta(\lambda, \tau) = 0. \quad (209)$$

Программное (в данном случае нулевое) движение системы (207) и (208) будет асимптотически устойчивым тогда и только тогда, когда все корни уравнения (209) имеют отрицательные вещественные части. Таким образом, для доказательства теоремы следует установить, что при достаточно малом $\tau > 0$ числа $\gamma_1, \dots, \gamma_n$ всегда можно выбрать так, что $\operatorname{Re} \mu_i(\tau) < 0$.

Обозначим через $\mu_1^{(0)}, \dots, \mu_n^{(0)}$ корни уравнения (209) при $\tau = 0$. Как известно, величины $\gamma_1, \dots, \gamma_n$ можно выбрать таким образом, чтобы эти корни принимали наперед заданные значения. Будем считать, что $\gamma_1, \dots, \gamma_n$ выбраны так, что $\mu_i^{(0)} < 0$, $i = 1, \dots, n$, и $\mu_i^{(0)} \neq \mu_j^{(0)}$ ($i \neq j$). Пусть $\alpha > 0$, причем $\mu_j^{(0)} < -\alpha$, $j = 1, \dots, n$. Покажем, что существует такое число $\tau_1 > 0$, что при любом $\tau > 0$ ($\tau \leq \tau_1$) все корни уравнения (209) удовлетворяют условию

$$\operatorname{Re} \mu_j(\tau) < -\alpha.$$

Предположим, что это не так, т. е. для любого $\tau_1 > 0$ есть корни такие, что $\operatorname{Re} \mu_j(\tau) \geq -\alpha$.

Из вида функции (209) вытекает, что все корни уравнения (209), обладающие свойством $\operatorname{Re} \mu_j(\tau) \geq -\alpha$, обязательно принадлежат прямоугольнику

$$\{-\alpha \leq \operatorname{Re} \lambda \leq \beta, \quad -\omega \leq \operatorname{Im} \lambda \leq \omega\},$$

где $\beta > 0$ и $\omega > 0$. Если при любом достаточно малом $\tau_1 > 0$ будут существовать корни уравнения (209), попадающие в этот прямоугольник, то из непрерывности функции $\Delta(\lambda, \tau)$ по обоим аргументам и замкнутости множества корней будет следовать, что при $\tau = 0$ уравнение (209) также имеет корни, расположенные в этом прямоугольнике. Это противоречит принятому предположению, что $\mu_j^{(0)} < -\alpha, j = 1, \dots, n$. Следовательно, существование числа τ_1 , установлено.

Таким образом, при выбранных величинах $\gamma_1, \dots, \gamma_n$ при всех запаздываниях $\tau < \tau_1$, нулевое решение системы (207), а следовательно, и программное движение исходной системы будут асимптотически устойчивыми.

Выполним замену переменных в системе (207) еще раз:

$$\bar{Y} = \begin{pmatrix} 1 & p_1 & p_2 & \dots & p_{n-1} \\ 0 & 1 & p_1 & \dots & p_{n-2} \\ 0 & 0 & 1 & \dots & p_{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} z^{(n-1)} \\ z^{(n-2)} \\ z^{(n-3)} \\ \dots \\ z \end{pmatrix}.$$

Тогда имеем одно дифференциальное уравнение

$$z(n) + p_1 z^{(n-2)} + \dots + p_n z = u. \tag{210}$$

Предположим теперь, что управление задается в форме:

$$u = Q_{n-1} \left(\frac{d}{dt} \right) z(t-h) + Q_{n-2} \left(\frac{d}{dt} \right) z(t-2h) + \dots + Q_0 z(t-nh), \tag{211}$$

где $h > 0$ — запаздывание, $Q_i(d/(dt))$ — полином с постоянными коэффициентами от оператора дифференцирования $d/(dt)$ степени i ($i = 0, 1, \dots, n - 1$). Требуется указать условия, при выполнении которых можно выбрать коэффициенты полиномов Q_i так, чтобы нулевое решение уравнения (210) при управлении (211) было асимптотически устойчивым. Тогда, в силу отсутствия особенности указанных выше преобразований, асимптотически устойчивым будет и нулевое решение системы (207), причем управление через переменные Y выписывается следующим образом:

$$u = C'_1 Y(t-h) + C'_2 Y(t-2h) + \dots + C'_n Y(t-nh),$$

где постоянные векторы C_i могут быть получены из коэффициентов полиномов Q_i , а также из коэффициентов указанных выше неособых преобразований.

Сформулированную задачу решает следующая теорема [74].

Теорема 26. Если корни уравнения

$$\lambda^n + \sum_{i=1}^n p_i \lambda^{n-i} = 0 \quad (212)$$

удовлетворяют условию $\operatorname{Re} \lambda < 1/h$, то всегда можно подобрать коэффициенты полиномов Q_i так, чтобы нулевое решение уравнения (210) с управлением (211) было асимптотически устойчивым.

Доказательство. Асимптотическая устойчивость будет иметь место тогда и только тогда, когда корни квазиполинома

$$\lambda^n + \sum_{i=1}^n p_i \lambda^{n-i} = \sum_{j=1}^n Q_{n-j}(\lambda) e^{-jh\lambda} \quad (213)$$

лежат в левой полуплоскости комплексной плоскости. Уравнение (212) можно представить в виде:

$$(\lambda - a_1) \dots (\lambda - a_n) = 0,$$

где a_i — вещественные или комплексные корни (212).

Рассмотрим уравнение

$$(\lambda - a_i + q_i e^{-\lambda h}) \dots (\lambda - a_n + q_n e^{-\lambda h}) = 0, \quad (214)$$

причем будем считать q_i вещественным, если a_i вещественное, и комплексным, если a_i комплексно. Если для каждого сомножителя в (214) корни лежат в левой полуплоскости при некотором наборе q_i ($i = 1, 2, \dots, n$), то раскрывая скобки и перенося в правую часть равенства слагаемые с экспоненциальными членами, получим, что корни уравнения (213) также лежат в левой полуплоскости, причем $Q_i(\lambda)$ определяются через a_i и q_i однозначно. Рассмотрим теперь отдельно каждый сомножитель. Сначала предположим, что a_i и q_i вещественны. Тогда при выполнении условий теоремы всегда можно выбрать вещественное q_i , чтобы у любого корня квазиполинома

$$\lambda - a_i + q_i e^{-h\lambda} = 0 \quad (215)$$

вещественная часть была отрицательна. Этот факт следует из рассмотрения примера (156) в данном приложении.

Пусть $a_i = \alpha + i\beta$ — комплексное число (здесь в правой части последнего неравенства i — мнимая единица в отличие от индекса $i = 1, 2, \dots, n$).

«Управляющий» параметр q_i также будем предполагать комплексным и будем искать его в виде $\rho e^{i\varphi}$. Тогда приходим к уравнению

$$\lambda - \alpha - i\beta = \rho e^{i\varphi - h\lambda}, \quad (216)$$

где $\rho > 0$, φ — вещественное число. Заменой $\mu = \lambda - i\beta$ можно (216) свести к (215), если положить, например,

$$\varphi = \pi(2k + 1) + h\beta, \quad k = 0, \pm 1, \pm 2, \dots,$$

причем $\operatorname{Re} \mu = \operatorname{Re} \lambda$, и, значит, если $\alpha < 1/h$, то комплексное q_i при выполнении условий теоремы всегда существует.

Для завершения доказательства осталось показать, что коэффициенты полиномов Q_i будут вещественны при таком выборе q_i . Рассмотрим произведение двух сомножителей из (214) с комплексно-сопряженным a_i , т. е.

$$(\lambda - \xi + \chi_1 e^{-h\lambda})(\lambda - \bar{\xi} - \chi_2 e^{-h\lambda}) = 0,$$

где $\bar{\xi}$ — комплексно-сопряженное ξ число. Раскрывая скобки в последнем уравнении, получаем:

$$(\lambda - \xi)(\lambda - \bar{\xi}) + \chi_1 \chi_2 e^{-2h\lambda} + (\lambda - \xi)\chi_2 e^{-h\lambda} + (\lambda - \bar{\xi})\chi_1 e^{-h\lambda} = 0.$$

Очевидно, что коэффициенты получаются вещественные, если $\chi_2 = \bar{\chi}_1$. Легко проследить, что именно так и выбираются соответствующие q_i . Таким образом, собирая сомножители с комплексно-сопряженными a_i парами, получим коэффициенты Q_i вещественными.

Доказательство теоремы закончено.

В. Л. Харитоновым [94] предложен алгоритм вычисления наибольшего значения запаздывания h_0 такого, что при всех $h \in [0, h_0]$ тривиальное решение системы

$$\frac{d}{dt} X(t) = \sum_{k=0}^m A_k X(t - hk), \tag{217}$$

асимптотически устойчиво. Характеристическое уравнение системы (217) имеет вид:

$$\det \left(\lambda E - \sum_{k=0}^m e^{-kh\lambda} A_k \right) = \sum_{s=0}^N P_s(\lambda) e^{-sh\lambda} = 0, \tag{218}$$

где $P_s(\lambda)$ — полиномы по λ . При этом степень полинома $P_0(\lambda)$ равна n и превосходит степени всех оставшихся полиномов. Умножив среднюю часть (218) на $e^{\lambda h N}$, получаем квазиполином с главным членом

$$f(\lambda) = \sum_{s=0}^N P_{N-s}(\lambda) e^{sh\lambda}.$$

Пусть все собственные числа матрицы

$$A = \sum_{k=0}^m A_k$$

лежат в левой полуплоскости. Из теоремы 1 тогда следует, что при достаточно малых $h > 0$ нули $f(\lambda)$ имеют отрицательные вещественные части. При увеличении h нарушение асимптотической устойчивости произойдет тогда, когда у $f(\lambda)$ появятся чисто мнимые нули. Рассмотрим полином

$$F_{\omega}(u) = \sum_{s=0}^N P_{N-s}(\omega) u^s.$$

Если ω_0 — вещественный нуль $f(i\omega)$, то полином $F_{\omega_0}(u)$ имеет нуль $u = e^{ih\omega_0}$, расположенный на единичной окружности комплексной плоскости. С другой стороны, если при некотором вещественном ω_0 полином $F_{\omega_0}(u)$ имеет нуль на единичной окружности, то существует h , при котором $f(i\omega_0) = 0$.

Введем в рассмотрение полином $F_{\omega}^*(u) = u^N \bar{F}_{\omega}(u^{-1})$, где черта означает, что все коэффициенты полиномов $P_s(i\omega)$ заменены на комплексно-сопряженные. Известно, что если полином имеет нуль, расположенный на единичной окружности, то этот нуль является общим для полиномов $F_{\omega}(u)$ и $F_{\omega}^*(u)$.

Для нахождения значений ω , при которых $F_{\omega}(u)$ имеет нуль, по модулю равный единице, составим результат полиномов $F_{\omega}(u)$ и $F_{\omega}^*(u)$ и обозначим его $R(F_{\omega}(u), F_{\omega}^*(u)) = P(\omega)$. Этот результат является полиномом по ω . Найдем его вещественные нули $\omega_1, \omega_2, \dots, \omega_M$ и рассмотрим полиномы $F_k(u) = F_{\omega_k}(u)$ ($k = 1, 2, \dots, M$). Нули полинома $F_k(u)$, модуль которых равен единице, обозначим через $u_1^{(k)}, u_2^{(k)}, \dots, u_{r_k}^{(k)}$ (может случиться, что полином $F_k(u)$ не имеет таких нулей, тогда $r_k = 0$): $u_j^{(k)} = \exp(i\varphi_{kj})$ ($\varphi_{kj} \in (0, 2\pi)$), $k = 1, 2, \dots, M$; $j = 1, 2, \dots, r_k$.

Введем числа:

$$h_{kj} = \begin{cases} \varphi_{kj} \omega_k^{-1} & \text{при } \omega_k > 0, \\ (\varphi_{kj} - 2\pi) \omega_k^{-1} & \text{при } \omega_k < 0, \quad k = 1, \dots, M; \quad j = 1, \dots, r_k. \end{cases}$$

Теорема 27. Для того чтобы тривиальное решение системы (217) было асимптотически устойчиво при всех $h \in [0, h]$, необходимо и достаточно, чтобы

$$h_1 < \min_{k,j} \{h_{kj}\}.$$

Из теоремы 27 следует, что искомое предельное запаздывание

$$h_0 = \min_{k,j} \{h_{kj}\}.$$

З а м е ч а н и е. Условие обращения в нуль результата можно заменить любым условием, позволяющим выделить вещественные значения ω , при которых полином $F_{\omega}(u)$ имеет корни на единичной окружности.

Пример [95]. Рассмотрим уравнение

$$\sum_{k=0}^n p_{n-k} x^{(k)}(t) + \sum_{k=0}^{n-1} q_{n-k} x^{(k)}(t-h) = 0.$$

Соответствующий квазиполином $f(\lambda) = e^{\lambda h} p(\lambda) + q(\lambda)$. Пусть корни полинома $p(\lambda) + q(\lambda)$ лежат в левой полуплоскости. Полином $F_{\omega}(u) = p(i\omega)u + q(i\omega)$. Корень этого полинома лежит в единичной окружности тогда и только тогда, когда $|p(i\omega)| = |q(i\omega)|$. Пусть $\omega_1, \omega_2, \dots, \omega_M$ — вещественные корни последнего уравнения. Полином $F_k(u)$ имеет корень

$$u_k = e^{i\varphi_k} = -q(i\omega_k) p^{-1}(i\omega_k), \varphi_k \in (0, 2\pi) (k = 1, 2, \dots, M).$$

Числа:

$$h_k = \begin{cases} \varphi_k \omega_k^{-1} & \text{при } \omega_k > 0, \\ (\varphi_k - 2\pi) \omega_k^{-1} & \text{при } \omega_k < 0, \end{cases}$$

а $h_0 = \min_k \{h_k\}$.

Аналогично к применению теоремы 27 сводится и задача стабилизации движения. А именно, пусть задана управляемая система

$$\frac{d}{dt} X(t) = \sum_{k=0}^m A_k X(t - kh) + \sum_{k=0}^l B_k U_k, \tag{219}$$

где B_k — постоянные $(n \times r_k)$ -матрицы, U_k — r_k -мерные векторы управления ($k = 1, 2, \dots, l$). Выберем управление:

$$U_k = C_k^* X(t - kh), \quad k = 0, 1, 2, \dots, l, \tag{220}$$

так, чтобы система (219), замкнутая управлением (220) была асимптотически устойчива при $h = 0$ (это возможно, когда пара (A, B) стабилизируема, здесь $B = (B_0, B_1, \dots, B_l)$). Далее задача сводится к определению наибольшего допустимого запаздывания.

ОБ ОПТИМАЛЬНОМ РАЗМЕРЕ ПРОИЗВОДСТВА

В этом приложении приведен параграф (глава 3, параграф 6) монографии Ю. П. Иванилова и А. В. Лотова «Математические модели в экономике» (1979) [36], так как дважды в тексте (2-й и 3-й глав) мы опираемся на выводы этого параграфа. Изложение ведется практически без изменений содержательной части с заменой терминов «экономическое стимулирование» на «прибыль предприятия».

Ниже описывается математическое исследование одной очень простой модели максимизации прибыли предприятия. Производство моделируется с помощью производственных функций, свойства которых предполагаются известными. Сначала строится производственная программа, наиболее выгодная для предприятия, обеспечивающая максимальную прибыль. Затем демонстрируется возможность с помощью назначения цен на сырье сделать выгодным для предприятия рациональное использование дефицитных ресурсов.

Перейдем к описанию модели предприятия. Предположим, что цена продукции предприятия p фиксированна. Будем считать, что выпуск измеряется в натуральных показателях (если же выпуск рассматривается в стоимостном выражении, то p — индекс цен). Тогда стоимость произведенной продукции будет равна

$$pf(x),$$

где $f(x)$ — производственная функция, которая определяет зависимость между выпуском и количеством занятых в производстве ресурсов: x^i , $i = \overline{1, n}$. Здесь $x = \{x^1, \dots, x^n\}$. Мы подробно исследуем случай одного продукта. Предположим, что предприятие может использовать произвольное количество производственных ресурсов, выплачивая за их использование некоторые суммы

(в виде арендной платы, зарплаты и т. д.). Пусть цены ресурсов задаются вектором q . Тогда издержки производства составят величину

$$qx = q_1x^1 + \dots + q_nx^n,$$

а прибыль $J(x)$ будет равна разности между стоимостью произведенной продукции и издержками производства:

$$J(x) = pf(x) - qx.$$

Как мы уже говорили, предполагается, что предприятие планирует свою деятельность так, чтобы максимизировать прибыль:

$$\max_{x \geq 0} (pf(x) - qx) = \max_{x \geq 0} J(x),$$

или, что то же самое, минимизировать убытки:

$$\min_{x \geq 0} (qx - pf(x)).$$

Предположим, что предприятие невелико и имеется достаточное количество ресурсов производства, чтобы обеспечить его деятельность, так что ограничения сверху на переменные $x = (x^1, \dots, x^n)$ отсутствуют. Для оптимального решения должны выполняться следующие условия:

$$\begin{aligned} \frac{\partial J}{\partial x^i} = p \frac{\partial f}{\partial x^i} - q_i \leq 0, \quad i = 1, \dots, n, \\ \sum_{i=1}^n \frac{\partial J}{\partial x^i} x^i = \sum_{i=1}^n \left(p \frac{\partial f}{\partial x^i} - q_i \right) x^i = 0. \end{aligned} \tag{221}$$

Второе условие носит название условия дополняющей нежесткости и означает, что если $p \frac{\partial f}{\partial x^i} - q_i < 0$, то $x^i = 0$ и соответствующий ресурс в производстве не используется. В самом деле, если мы введем некоторое дополнительное количество этого ресурса dx^i в производство, то изменение прибыли будет отрицательно, т. е.

$$(p \frac{\partial f}{\partial x^i} - q_i) dx^i < 0,$$

что нецелесообразно.

Поэтому для всех ресурсов, затрачиваемых в производстве, мы получим условие максимума прибыли в форме:

$$p \frac{\partial f}{\partial x^i} - q_i = 0. \tag{222}$$

Оно означает, что в случае, когда предприятие работает оптимально (с максимальной прибылью), то величины $p \frac{\partial f}{\partial x^i}$ для всех используемых в производстве ресурсов равны их ценам.

Экономический смысл этого соотношения понятен: если бы вместо соотношения (222) имело место, например, соотношение

$$p\partial f/\partial x^i - q_i > 0,$$

или, что то же самое,

$$pdf > q_i dx^i, \quad (223)$$

то из условия (223) следовали бы такие выводы: при небольшом увеличении количества i -го ресурса на величину dx^i стоимость дополнительно произведенной продукции pdf превосходила бы затраты на приобретение этого ресурса. Поэтому имело бы смысл увеличить использование этого ресурса и, таким образом, прибыль предприятия. Аналогично, обратное соотношение:

$$pdf < q_i dx^i,$$

при стремлении увеличить прибыль количество используемого в производстве ресурса уменьшится, и так до тех пор, пока для всех используемых ресурсов не будет выполнено соотношение (221) (или соответствующий ресурс не будет исключен из производства: $x^i = 0$). Мы пришли к условиям оптимальности (221) вместе с условиями дополняющей нежесткости.

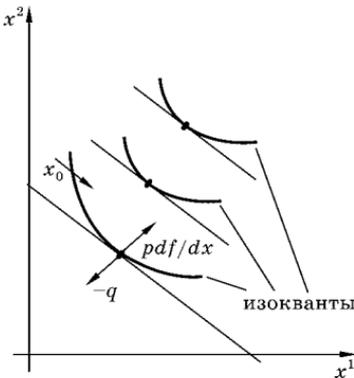


Рис. 51

Изокванты касаются изокост в точках оптимального объема производства

Вычертим в плоскости $\{x^1, x^2\}$ изокванты (линии одинакового выпуска продукции в стоимостном выражении при различных затратах ресурсов) и изокосты (линии одинаковых затрат). Уравнения изоквант имеют следующий вид (рис. 51):

$$pf(x^1, x^2) = \text{const.}$$

Уравнения изокост определяют прямые:

$$q_1 x^1 + q_2 x^2 = \text{const.}$$

Согласно условию максимума прибыли в точках, соответствующих оптимальным затратам

и выпускам, изокоста касается изокванты. В самом деле, пусть это не так. Если, например, режим функционирования предприятия изображается точкой x^0 , то, смещаясь из этой точки вдоль изокосты в направлении, указанном на рисунке стрелкой, мы с теми же затратами увеличиваем выпуск.

Естественно предположить, что все входящие в производственную функцию ресурсы используются. Тогда мы приходим к следующей системе условий:

$$\begin{aligned}
 p \frac{\partial f(x^1, \dots, x^n)}{\partial x^1} - q_1 &= 0 \\
 \dots & \\
 p \frac{\partial f(x^1, \dots, x^n)}{\partial x^n} - q_n &= 0.
 \end{aligned}
 \tag{224}$$

Число этих условий совпадает с числом используемых ресурсов (если некоторый ресурс не используется, то его можно не рассматривать, сократив число ресурсов).

Таким образом, при фиксированных ценах мы имеем n уравнений, из которых можно определять значения количеств ресурсов $\bar{x}^1, \dots, \bar{x}^n$, при которых прибыль предприятия максимальна. Якобиан системы (224) совпадает с матрицей Гессе

$$H = \left[\frac{\partial^2 f}{\partial x^i \partial x^j} \right]$$

производственной функции. Если определитель этой матрицы отличен от нуля, то по теореме о неявных функциях мы можем найти вектор-функцию:

$$\begin{aligned}
 \bar{x}^1 &= \bar{x}^1(p, q_1, \dots, q_n) \\
 \dots & \\
 \bar{x}^n &= \bar{x}^n(p, q_1, \dots, q_n).
 \end{aligned}
 \tag{225}$$

Эта функция имеет смысл функции спроса на ресурсы, а условия (224) — уравнений спроса на ресурсы. Вектор-функция (225) определяет количества ресурсов, при которых прибыль предприятия максимальна в заданных ценах на продукцию и на ресурсы.

При пропорциональном изменении цен ($p \rightarrow tp, q_1 \rightarrow tq_1, \dots, q_n \rightarrow tq_n$) система (224) не меняется. Не меняется и ее решение. Таким образом, функция (225) — однородная функция нулевой степени. Иными словами,

$$\bar{x}(tp, tq) = \bar{x}(p, q).$$

Подставляя функцию спроса на ресурсы в производственную функцию, мы получим функцию оптимального выпуска:

$$\bar{y} = f(\bar{x}(p, q)) = \bar{y}(p, q).$$

Она определяет желательный для предприятия выпуск продукции при заданных ценах.

Рассмотрим теперь поведение предприятия при изменении цен с течением времени. При этом по-прежнему будем считать, что действия предприятия не влияют на цены.

Пусть в исходный момент предприятие функционирует оптимально, находясь в точке \bar{x}_0 , определяемой текущими ценами из соотношения (224). Пусть изменение цен во времени известно предприятию, и это изменение происходит «плавно» (без скачков). Предприятию выгодно планировать производство в соответствии с меняющимися ценами так, чтобы траектория затрат лежала на линии, соединяющей точки касания изоквант и изокост.

Все приведенные соображения справедливы и для случая любого числа продуктов, когда $p = (p_1, \dots, p_n)$ и $f = (f^1, \dots, f^m)$ — векторы.

Для случая одного продукта из условия (222) имеем

$$p = \frac{q_1}{(\partial f / \partial x^1)} = \dots = \frac{q_n}{(\partial f / \partial x^n)}, \quad (226)$$

т. е. в оптимальном для предприятия плане предельные производительности ресурсов пропорциональны их ценам.

Теперь рассмотрим вопрос о работе предприятия в тех случаях, когда ресурсы ограничены. Покажем, как с помощью регулирования цен можно добиться того, чтобы предприятию было выгодно использовать ресурсы в количествах, не превышающих выделенные лимиты.

Пусть в производстве используются несколько замещаемых ресурсов, количества которых ограничены:

$$0 \leq x^i \leq \hat{x}^i, \quad i = 1, \dots, n. \quad (227)$$

Стоимость продукции, производимой предприятием, равна

$$pf(x^1, \dots, x^n).$$

Как и раньше, использование ресурсов в производстве связано с издержками:

$$g(x^1, \dots, x^n) = \sum_{i=1}^n q_i x^i.$$

Предприятие будет планировать свою деятельность так, чтобы максимизировать разность стоимости произведенной продукции и издержек:

$$pf(x^1, \dots, x^n) - g(x^1, \dots, x^n) \rightarrow \max \quad (228)$$

при выполнении ограничений (227).

Можно не ограничивать ресурсы предприятия, а назначить цены μ_i ($i = 1, \dots, n$), по которым оплачиваются используемые в производстве ресурсы, получаемые сверх лимитов (227). Пусть предприятие может продавать излишки ресурсов также по цене μ_i , $i = 1, \dots, n$. В этом случае прибыль предприятия, которую он будет максимизировать, равна

$$L(x, \mu) = pf(x^1, \dots, x^n) - g(x^1, \dots, x^n) - \sum_{i=1}^n \mu_i(x^i - \hat{x}^i). \quad (229)$$

В математике функция (229) носит название *функции Лагранжа*. Известно, что при некоторых предположениях о функции f существуют такие множители Лагранжа μ_i , что точка максимума функции (228) при ограничениях (227) является стационарной точкой функции Лагранжа, т. е.

$$\partial L / \partial x^i = p \partial f / \partial x^i - \partial g / \partial x^i - \mu_i = 0 \quad (230)$$

при оптимальном использовании ресурсов.

То, что соотношение (230) должно выполняться, достаточно очевидно, ибо как только цены зафиксированы, в случае

$$\partial L / \partial x^i > 0$$

следует увеличивать использование ресурса x^i , увеличивая тем самым прибыль предприятия (229), а в случае

$$\partial L / \partial x^i < 0$$

количество используемого ресурса следует соответственно уменьшать.

Итак, должны существовать такие величины μ_i , для которых условиям (230) удовлетворяют оптимальные количества ресурсов. В самом деле, если ресурс используется сверх установленных пределов, то цена на него поднимается. Тогда использование в производстве этого ресурса уменьшится. Если же у предприятия остается излишек ресурса, то цену на него можно снизить и использование этого ресурса в производстве должно увеличиться. В обоих случаях такой механизм изменения цен уменьшает прибыль предприятия. В конце концов, между ценами $\bar{\mu}$ и количествами используемых ресурсов \bar{x} должно установиться равновесие. Вектор \bar{x} максимизирует прибыль при заданных ценах $\bar{\mu}$:

$$L(x, \bar{\mu}) \leq L(\bar{x}, \bar{\mu}). \quad (231)$$

В то же время, как уже говорилось, цены $\bar{\mu}$ таковы, что они уменьшают прибыль предприятия по сравнению с произвольными ценами:

$$L(\bar{x}, \bar{\mu}) \leq L(\bar{x}, \mu). \quad (232)$$

Объединяя (231) и (232), получим

$$L(x, \bar{\mu}) \leq L(\bar{x}, \bar{\mu}) \leq L(\bar{x}, \mu). \quad (233)$$

Точка $(\bar{x}, \bar{\mu})$, удовлетворяющая условиям (233), называется *седловой точкой* функции Лагранжа $L(x, \mu)$.

В математике доказывается, что при условии существования седловой точки выполняется условие

$$\max_{x \geq 0} \min_{\mu \geq 0} L(x, \mu) = \min_{\mu \geq 0} \max_{x \geq 0} L(x, \mu).$$

Иными словами, при фиксированном режиме x работы предприятия цены выбираются так, чтобы уменьшить прибыль предприятия $L(x, \mu)$, доведя ее до минимума

$$\min_{\mu \geq 0} L(x, \mu),$$

а предприятие должно выбрать режим работы x так, чтобы максимизировать прибыль с учетом реакции механизма изменения цен на использование ресурсов:

$$\max_{x \geq 0} \min_{\mu \geq 0} L(x, \mu). \quad (234)$$

Наоборот, при фиксированных ценах μ предприятие максимизирует прибыль:

$$\max_{x \geq 0} L(x, \mu).$$

Назначение цен уменьшает эту прибыль и приводит ее к минимуму

$$\min_{\mu \geq 0} \max_{x \geq 0} L(x, \mu).$$

Покажем, что задача (234) эквивалентна исходной задаче предприятия (227), (228).

Имеем:

$$\min_{\mu \geq 0} L(x, \mu) = \begin{cases} pf(x) - g(x), & \text{если } x^i \leq \hat{x}^i \text{ для всех } i = 1, \dots, n, \\ -\infty, & \text{если } x^i > \hat{x}^i \text{ хотя бы для одного } i. \end{cases} \quad (235)$$

В самом деле, если $x^i < \hat{x}^i$, то $\hat{x}^i - x^i > 0$ и минимум $\mu_i(\hat{x}^i - x^i)$ по неотрицательным μ_i достигается при $\mu_i = 0$. Если же $\hat{x}^i - x^i < 0$, то, неограниченно увеличивая μ_i , можно сделать убытки предприятия сколь угодно большими ($L \rightarrow -\infty$).

Максимизируя (235), мы, во-первых, должны удовлетворить всем условиям (228), чтобы не получить $L = -\infty$, а во-вторых, добившись выполнения этих условий, максимизировать $pf(x) - g(x)$, т. е. сделать то же, что и в задаче (227), (228).

Заметим, что из (230) имеем условие

$$p = \frac{\frac{\partial g}{\partial x^1} + \mu_1}{\frac{\partial f}{\partial x^1}} = \dots = \frac{\frac{\partial g}{\partial x^n} + \mu_n}{\frac{\partial f}{\partial x^n}},$$

аналогичное условию (226). Экономически целесообразные предельные издержки $\partial g / \partial x^i + \mu_i$ равны предельным производительностям $p \partial f / \partial x^i$ ресурсов в стоимостном выражении.

Введение дополнительной платы за ресурс по цене μ_i «уравнивает» ценность взаимозамещаемых ресурсов, делая их одинаково «эффективными» для производства, так что при этом нет дополнительного стимула к использованию в первую очередь дефицитных, особо эффективных ресурсов.

ПРОИЗВОДСТВЕННАЯ ФУНКЦИЯ

ВВЕДЕНИЕ

Ряд характерных особенностей, задач и целей выделяют экономисты из множества объектов управления. В настоящее время экономическая жизнь предприятия во многом определяется способностью с необходимой точностью описать те или иные явления экономики, умением планировать ведение хозяйства, владением искусством управления.

В данном приложении производственные функции представляются как инструмент для моделирования деятельности предприятия, для качественного исследования процесса производства и реализации управляющих воздействий. Описывается несколько существующих подходов к определению и построению производственных функций. Измерение текущего состояния объекта лежит в основе любой системы управления, благодаря которому возможна организация таких процессов, как планирование и регулирование, поэтому здесь, в приложении 3, приводится качественный анализ математических характеристик производственных функций.

В тексте описываются конкретные типы производственных функций с указанием их специфических особенностей. При моделировании процесса производства, кроме отражения его основных тенденций, весьма важно учитывать влияние достижений науки и техники. Рассматриваются формы научно-технического прогресса, нейтральные по отношению к тому или иному ресурсу. Кроме того, описывается прогресс, определяемый внутри самой отрасли. Одна из основных задач исследования деятельности предприятия заключается в том, чтобы в каждом отдельном случае построить такую модель, которая как можно точнее воспроизводила бы изучаемый производственный процесс, поэтому в этом приложении излагаются этапы экономико-статистического моделирования и перечисляются возможные методы идентификации параметров производственных функций.

ПОНЯТИЕ ПРОИЗВОДСТВЕННОЙ ФУНКЦИИ

ОПРЕДЕЛЕНИЯ ПРОИЗВОДСТВЕННОЙ ФУНКЦИИ

Можно выделить три основных подхода к определению понятия производственной функции (ПФ) и к ее построению [23]:

- эмпирический подход — изучаемая экономическая система рассматривается как черный ящик, вид функции задается заранее, а ее параметры определяются на основе статистической обработки наблюдавшихся затрат ресурсов и выпусков продукции;
- теоретический подход — экономическая система разбивается на элементарные подсистемы, для каждой из которых тем или иным способом задаются «элементарные» ПФ, а ПФ всей системы в целом определяется как эффективная граница множества ее производственных возможностей, описываемого в терминах затрат ресурсов и выпусков продукции;
- композиционный подход базируется на микроописании системы в виде совокупности «элементарных» подсистем, но в отличие от теоретического подхода в модель включается описание экономического механизма, регулирующего взаимодействие подсистем.

ПФ, построенная по эмпирическому подходу, может быть использована для прогнозирования развития системы тогда и только тогда, когда социально-экономические и организационные условия производства не меняются. В теоретическом подходе описывается не реальное поведение, а лишь потенциальные возможности производства. Благодаря же композиционному подходу, ПФ отражает как материальные, так и организационные аспекты производства.

ЭМПИРИЧЕСКИЙ ПОДХОД К ПОНЯТИЮ ПФ

Пусть R_+^n — неотрицательный ортант n -мерного евклидова пространства, каждый вектор $x = (x_1, \dots, x_n)$ которого интерпретируется как вектор затрат производственных ресурсов. Пусть R_+^m — неотрицательный ортант m -мерного евклидова пространства, каждый вектор $y = (y_1, \dots, y_m)$ которого интерпретируется как набор количественных оценок результатов производства при определенных затратах ресурсов. Если P — некоторый производственный процесс, то производственной функцией f этого процесса будем называть отображение

$$f: D \rightarrow V, \text{ где } D \subseteq R_+^n, V \subseteq R_+^m, \quad (236)$$

моделирующее выпуск продукции в процессе P , такое, что каждой точке пространства затрат соответствует единственный максимальный вектор количественных оценок, получаемый при использовании этих затрат. В дальнейшем будем рассматривать случай $m = 1$, т. е. когда имеется единственная количественная оценка результатов производства, и в этом случае производственную функцию естественно записывать как обычную функцию нескольких переменных:

$$y = f(x_1, \dots, x_n). \quad (237)$$

ТЕОРЕТИЧЕСКИЙ ПОДХОД К ПОНЯТИЮ ПФ

Производственная функция моделируемой экономической системы, как уже отмечалось, определяется как эффективная граница множества производственных возможностей. Далее в работе для анализа границ производственных возможностей и построения соответствующих ПФ описывается метод множеств достижимости.

ОБ ОГРАНИЧЕНИЯХ ПРОИЗВОДСТВЕННЫХ ФУНКЦИЙ

Предполагается, что производственная функция удовлетворяет ряду аксиом, отражающих определенные экономические факты процесса производства. Понятно, что ограничения на производственную функцию устанавливают сообразно условиям конкретной действительности.

Рассмотрим предположения и дадим их экономико-математическое обоснование.

1. Неотрицательность. В силу определения производственной функции для любой точки $X = (x_1, \dots, x_n)$ выполняется

$$y = f(x_1, \dots, x_n) \geq 0.$$

2. Непрерывность. Соответствует условию малых изменений выпуска при малых вариациях ресурсов. Однако это упрощение условно, поскольку, в принципе, можно представить себе случай, когда малое изменение какого-то ресурса создаст качественно новую ситуацию, в которой возможно осуществить скачок в объеме выпуска. Это замечание особенно существенно для объектов микроуровня.

3. Гладкость. В случае, когда речь идет всего лишь об аппроксимации, непрерывную функцию всегда можно с любой точностью приблизить к гладкой (обычно считают ПФ дважды дифференцируемой).

4. Однородность. Свойство расширения производственных возможностей, которое означает отсутствие структурных и качественных сдвигов для различных точек наблюдения:

$$f(\lambda x_1, \dots, \lambda x_n) = \lambda^k f(x_1, \dots, x_n) > f(x_1, \dots, x_n)$$

для $\forall (x_1, \dots, x_n) \in R_+^n$ при $\lambda > 1, k > 0$. Отметим, что данное предположение относительно производства сильно упрощает, а иногда и искажает реальное наблюдение.

5. Неубывание по некоторому ресурсу. В области взаимозаменяемости ресурсов существует подмножество пространства затрат, называемое экономической областью, в которой увеличение любого вида затрат при неизменном объеме остальных не приводит к уменьшению выпуска, т. е. $(\partial y / \partial x_i) = 0$, и, кроме того, существуют процессы, в которых чрезмерное увеличение ресурса оказывает отрицательное воздействие на выпуск.

6. Строгая вогнутость по координатам. Эффект убывающей доходности, т. е. в области определения производственной функции существует непустое множество M , в котором выполняется $(\partial^2 y / \partial x_i^2) < 0, i = 1, \dots, n$. Этот факт имеет следующее экономическое объяснение: увеличение затрат лишь одного производственного фактора приводит к снижению его эффективности, предельной производительности. Поскольку каждая последующая единица возрастающего производственного ресурса обеспечивается все меньшим приходящимся на нее количеством других ресурсов, эффективность растущего ресурса уменьшается.

7. Строгая выпуклость изоквант. Эффект возрастания издержек при замещении факторов производства.

8. Супераддитивность. Означает, что резервирование ресурсов невыгодно для системы, иначе бы возник дополнительный эффект от дробления и резервирования ресурсов, т. е. $\forall x', x''$ из области определения функции выполняется

$$f(x' + x'') \geq f(x') + f(x'').$$

9. Делимость. Соответствует тому, что для целого $n > 0$ справедливо $f(x/n) \geq f(x)/n$.

МАТЕМАТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ПРОИЗВОДСТВЕННОЙ ФУНКЦИИ И ИХ ЭКОНОМИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ

Каждый раз, когда моделируется поведение какой-либо реально существующей системы, следует исходить из существа экономических отношений, из качественного анализа экономических связей. В связи с этим отметим следующие характеристики производственных функций.

1. Предельные значения производства. Пусть x_1, \dots, x_n — затраты производства, а $f(x_1, \dots, x_n)$ — производственная функция. Тогда для любого набора: $(i_1, \dots, i_k), k \geq 1, i_s \neq i_r, r \neq s, i_s \in \{1, \dots, n\}$, рассматриваются пределы:

$$\lim_{x_{i_1} \rightarrow \infty, \dots, x_{i_k} \rightarrow \infty} f(x_1, \dots, x_n),$$

показывающие объем выпуска при $x_{i_1} \rightarrow \infty, \dots, x_{i_k} \rightarrow \infty$.

2. Предельная норма замещения фактора j фактором i :

$$\frac{\partial f(x_1, \dots, x_n)}{\partial x_i} / \frac{\partial f(x_1, \dots, x_n)}{\partial x_j}, \quad i, j = 1, \dots, n, \quad i \neq j.$$

Эта характеристика выражает соотношение между изменением объема выпуска, вызванным изменением на единицу i -го ресурса, и таким же изменением, связанным с вариацией на единицу j -го ресурса (исходным значением ресурсов в обоих случаях является точка (x_1, \dots, x_n)).

Другая интерпретация функции предельной нормы замены состоит в следующем: ее значение в точке (x_1, \dots, x_n) приближенно определяет, насколько нужно увеличить ресурс j , чтобы скомпенсировать уменьшение ресурса i на единицу, обеспечивая тот же выход продукции, при условии, что остальные ресурсы не изменились, а исходные значения ресурсов составляли $(\bar{x}_1, \dots, \bar{x}_n)$.

Опытные данные свидетельствуют: в большинстве ПФ при увеличении затрат фактора i и уменьшении затрат фактора j и постоянстве затрат всех остальных факторов и выпуска продукции предельная норма замещения фактора j фактором i уменьшается. Это свойство получило название *закона убывания предельной нормы замещения*. Более точно этот закон выражается неравенством [9]

$$\frac{d \left\{ \frac{\partial f}{\partial x_i} / \frac{\partial f}{\partial x_j} \right\}}{d \left\{ \begin{matrix} x_i \\ x_j \end{matrix} \right\}} < 0,$$

в котором величины f и $x_k, k \neq i, j$ принимаются постоянными.

3. Темп прироста α_u величины $u(\lambda)$. Это относительная скорость роста величины $u(\lambda)$ при изменении λ . В случае дискретного процесса: $\alpha_u = \Delta u(\lambda) / u(\lambda) = (u(\lambda + 1)) / u(\lambda)$. В непрерывном случае: $\alpha_u = du(\lambda) / d\lambda u(\lambda)$.

4. Эластичность выпуска по i -му ресурсу

$$\frac{\partial f(x_1, \dots, x_n)}{\partial x_i} / \frac{f(x_1, \dots, x_n)}{x_i}, \quad i = 1, \dots, n,$$

показывает в точке $(\bar{x}_1, \dots, \bar{x}_n)$, на сколько процентов изменится выпуск при увеличении затрат данного ресурса на один процент при фиксированных остальных ресурсах.

5. Эластичность производства

$$\varepsilon(x) = \lim_{t \rightarrow 1} \frac{t}{f(tx)} \frac{\partial f(tx)}{\partial t} = \lim_{t \rightarrow 1} \frac{\partial \ln f(tx)}{\partial \ln t}, \text{ где } x = (x_1, \dots, x_n), t > 1,$$

является локальным показателем изменения выпуска при расширении масштаба производства в точке \bar{x} . Очевидно, что

$$\begin{aligned} \varepsilon(x) &= \lim_{t \rightarrow \infty} \frac{t}{f(tx)} \frac{\partial f(tx)}{\partial t} = \lim_{t \rightarrow \infty} \frac{t}{f(tx)} \left[\frac{\partial f(tx)}{\partial \tilde{x}_1} \frac{\partial \tilde{x}_1}{\partial t} + \dots + \frac{\partial f(tx)}{\partial \tilde{x}_n} \frac{\partial \tilde{x}_n}{\partial t} \right] = \\ &= \lim_{t \rightarrow \infty} \frac{t}{f(tx)} \frac{1}{t} \sum_{i=1}^n \frac{\partial f}{\partial \tilde{x}_i} \tilde{x}_i = \sum_{i=1}^n \frac{\partial f}{\partial \tilde{x}_i} \frac{\tilde{x}_i}{f} = \sum_{i=1}^n \varepsilon_i(x). \end{aligned}$$

Таким образом, эластичность производства в некоторой точке x равна сумме эластичностей выпуска по отношению к затратам различных ресурсов в этой точке.

6. Показатели эластичности замены факторов для двухатомной производственной функции. Эластичность замены факторов по Аллену есть количественная оценка скорости изменения предельной нормы замещения при движении вдоль изокванты и равна

$$\sigma^A(x_1, \dots, x_n) = \left(\frac{d(f_1/f_2)}{d(x_2/x_1)} \bigg/ \frac{(f_1/f_2)}{(x_2/x_2)} \right)^{-1} \text{ при } f(x_1, x_2) = \text{const.}$$

Эластичность замены факторов по Михайлевскому — характеристика ПФ, имеющая вид:

$$\sigma^M(x_1, x_2, dx_1, dx_2) = \frac{d(x_2/x_1)}{d\left(\frac{f_1(x_1, x_2)}{f_2(x_1, x_2)}\right)} \bigg/ \frac{x_2/x_1}{f_1(x_1, x_2)/f_2(x_1, x_2)}.$$

Эластичность замены факторов соотносит темпы изменения пропорции между замещаемыми факторами с темпами изменения их предельных продуктов; если в определении эластичности замены по Аллену факторы изменялись таким образом, чтобы обеспечивалось постоянство выпуска (т. е. в пределах одной изокванты), то здесь допускается их произвольное изменение.

Приведем рабочие формулы, непосредственно выражающие эластичности замены факторов через аргументы и значения производственной функции и ее производных:

$$\sigma^A(x_1, x_2) = \frac{f_1 f_2 (x_1 f_1' + x_2 f_2')}{x_1 x_2 (-f_{11} f_2^2 + 2 f_{12} f_1' f_2' - f_{22} f_1'^2)},$$

$$\sigma^M(x_1, x_2, dx_1, dx_2) = f_1 f_2 \frac{-\frac{1}{x_1} dx_1 + \frac{1}{x_2} dx_2}{(f_{11} f_2 - f_{12} f_1) dx_1 + (f_{12} f_2 - f_{22} f_1) dx_2}.$$

Свойства:

- а) $\sigma_{12}^A = \sigma_{21}^A$ (симметричность);
 б) инвариантность эластичности замены факторов по Аллену и Михайлевскому относительно преобразования h , т. е.

$$\sigma_{h(f)}^A = \sigma_f^A, \sigma_{h(f)}^M = \sigma_f^M,$$

где h — любая дважды дифференцируемая функция от одного аргумента;

- в) $\sigma_f^M = \sigma_f^A$ для однородных функций любой степени однородности.

Для линейно-однородной ПФ известно определение эластичности замены факторов, введенное Дж. Хиксом:

$$\sigma_f^H(x_1, x_2) = \frac{f_1(x_1, x_2) f_2(x_1, x_2)}{f(x_1, x_2) f_{12}(x_1, x_2)}.$$

Свойства:

- а) $\sigma_f^H(x_1, x_2) = \sigma_f^H(x_2, x_1)$ (симметричность);
 б) $\sigma_f^H = \sigma_f^A = \sigma_f^M$ для однородных двухфакторных функций первой степени однородности.

Для функции произвольной степени однородности γ имеет место равенство [41]

$$\sigma_f^A(x_1, x_2) = \frac{f_1 f_2}{(1-\gamma) f_1 f_2 + \gamma f f_{12}}.$$

7. Показатели эластичности замены факторов для многофакторной производственной функции. Определение эластичности замены факторов по Мак-Фаддену, по сути дела, есть не что иное, как сведение многофакторной функции к двухфакторной и последующего применения определения эластичности замены по Аллену:

$$\sigma_{ij}^F(x_1, \dots, x_n) = \frac{(f_j / f_i) \partial(f_j / f_i)}{(x_i / x_j) \partial(x_i / x_j)},$$

где $x_k = \text{const}$ при $k \neq j$, $i, j = \text{const}$.

Эластичность замены i -го фактора j -м по Аллену определяется следующим образом:

$$\sigma_{ij}^A(x_1, \dots, x_n) = \frac{x_1 f_1 + \dots + x_n f_n}{x_i x_j} \frac{H_{ij}}{H}, \text{ где } H = \begin{vmatrix} 0 & f_1 & \dots & f_n \\ f_1 & f_{11} & \dots & f_{1n} \\ \dots & \dots & \dots & \dots \\ f_n & f_{n1} & \dots & f_{nn} \end{vmatrix},$$

H_{ij} — алгебраическое дополнение к элементу f_{ij} матрицы H , $f_i = \partial f / \partial x_i$, $f_{ij} = \partial^2 f / \partial x_i \partial x_j$.

Величина σ_{ij}^A с точностью до множителя

$$\left(x_j f_j / \sum_{j=1}^n x_j f_j \right)$$

показывает, на сколько процентов придется увеличить количество ресурса x_i , чтобы сохранить неизменными выпуск и предельные нормы замены остальных i -х факторов в условиях, когда предельная норма замены j -го фактора i -м уменьшилась на 1%.

Ввиду симметричности матрицы H имеет место равенство $\sigma_{ij}^A = \sigma_{ji}^A$ (предполагаем, что ПФ дважды дифференцируема).

Эластичность замены факторов по Михайлевскому:

$$\sigma_{ij}^M(x_1, \dots, x_n, dx_1, \dots, dx_n) = s_{ij} d \left(\frac{x_i}{x_j} \right) / \left(\frac{x_i}{x_j} \right) ds_{ij},$$

где $s_{ij} = (\partial y / \partial x_j) / (\partial y / \partial x_i)$, $ds_{ij} = \sum_{k=1}^n \frac{\partial s_{ij}}{\partial x_k} dx_k$.

Величина σ_{ij}^M показывает соотношение между процентным изменением отношения i -го и j -го факторов и процентным изменением предельной нормы замены при переходе от точки (x_1, \dots, x_n) к близкой точке $(x_1 + dx_1, \dots, x_n + dx_n)$.

8. Эластичность выпуска по масштабу производства. Характеристика

$$\left(\frac{\partial f(\lambda x_1, \dots, \lambda x_n)}{\partial \lambda} / \frac{f(\lambda x_1, \dots, \lambda x_n)}{\lambda} \right),$$

показывающая, насколько ПФ $f(x_1, \dots, x_n)$ близка к однородной, причем эта близость рассматривается локально — в окрестности точки $(\bar{x}_1, \dots, \bar{x}_n)$, или, другими словами, эластичность выпуска по масштабу производства позволяет установить, на сколько процентов изменится индекс роста с изменением λ на 1%, при условии, что исходной точкой для изменения ресурсов являлись значения $\bar{x}_1, \dots, \bar{x}_n$.

9. Индекс роста. Индексом роста объема выпуска при увеличении размеров всех ресурсов в λ раз называется коэффициент

$$\text{ind}_j f(x) = f(\lambda x_1, \dots, \lambda x_n) / f(x_1, \dots, x_n).$$

ПОДХОДЫ К ПОСТРОЕНИЮ ПРОИЗВОДСТВЕННЫХ ФУНКЦИЙ

МЕТОД МНОЖЕСТВ ДОСТИЖИМОСТИ

В работе [23] для анализа математической модели с экзогенными переменными, которые в данном случае будем считать управлениями, рассматриваются некоторые агрегированные переменные, являющиеся показателями функционирования

изучаемой системы. Поскольку связь показателей с управлениями задается соотношениями модели, множеству допустимых управлений соответствует множество допустимых значений показателей, которое впредь будем называть обобщенным множеством достижимости (ОМД).

Рассмотрим случай линейной модели

$$Az \leq b, \quad (238)$$

где $z = (z_1, \dots, z_n)$ — вектор решений, $z \in R^n$.

Напомним, что по определению соотношение:

- 1) $z' \leq z''$, $\forall z', z'' \in R^n$ означает $z'_k \leq z''_k$, $\forall k = 1, n$;
- 2) $z' \leq z''$, $\forall z', z'' \in R^n$ означает $z'_k \leq z''_k$, $\forall k = 1, n$, но, по крайней мере, существует один k такой, что $z'_k < z''_k$;
- 3) $z' < z''$, $\forall z', z'' \in R^n$ означает $z'_k < z''_k$, $\forall k = 1, n$.

Если ввести понятие множества допустимых решений $G \in R^n$, то для модели (238) оно имеет вид

$$G = \{z \in R^n: Az \leq b\}. \quad (239)$$

Пусть G не пусто, и пусть задано m показателей качества решения модели в агрегированном виде и связанных с ним соотношением

$$v = Cz, \quad (240)$$

где C — матрица, $v = (v_1, \dots, v_m)$, $v \in R^m$. Тогда множество достижимых значений агрегированных показателей (ОМД) определяется так:

$$G_v = \{v \in R^m: v = Cz, Az \leq b\}. \quad (241)$$

Метод множеств достижимости состоит в построении ОМД в явном виде:

$$G_v = \{v \in R^m: Dv \leq d\}, \quad (242)$$

т. е. в нахождении таких матрицы D и вектора d , чтобы множество решений системы линейных неравенств (242) точно или приближенно совпадало с G_v , и в последующем изучении этого множества.

В дальнейшем уделяется внимание использованию этого подхода для оценки границ производственных возможностей и построения ПФ, описывающих потенциальные возможности системы.

ПОСТРОЕНИЕ ТЕОРЕТИЧЕСКИХ ПРОИЗВОДСТВЕННЫХ ФУНКЦИЙ НА ОСНОВЕ ММД

Пусть рассматриваемая экономическая система описывается совокупностью p производственных процессов, каждый из которых характеризуется столбцом из q чисел, показывающих объем выпуска (или потребления, если число меньше нуля) того или

иного продукта (ресурса) при единичной интенсивности производства. Тогда математическая модель системы будет иметь вид:

$$w = Bu, \quad u \geq 0, \quad (243)$$

где $u \in R^p$ — вектор интенсивностей использования технологических процессов, B — матрица технологических процессов, $w \in R^q$ — конечный выпуск продукции (или затраты ресурсов при отрицательном значении величины выпуска).

Множество производственных возможностей в данном случае — это конус K в пространстве продуктов R^q :

$$K = \{w \in R^q: w = Bu, u \geq 0\}. \quad (244)$$

Пусть на модель (243) наложены ограничения на интенсивности технологических процессов и на затраты некоторых конечных ресурсов так, что модель системы приобретает вид:

$$w = Bu, \quad u \geq 0, \quad Fw \leq f, \quad Hu \leq h, \quad (245)$$

где F, H, f, h — заданные матрицы и векторы.

В качестве показателей, представляющих интерес в исследовании, естественно рассматривать наиболее важные из продуктов или их некоторые комбинации (агрегированные продукты):

$$v = Cw. \quad (246)$$

Заметим, что в этом случае множество производственных возможностей, характеризуемых вектором $v \in R^n$, уже не является конусом.

Введем в рассмотрение эффективную (паретовскую) границу $P(G_v)$:

$$P(G_v) = \{v \in G_v: \{v' \in G_v, v' \geq v\} \Rightarrow v' = v\},$$

и слабо эффективную (слейторовскую) границу $S(G_v)$:

$$S(G_v) = \{v \in G_v: v' > v \Rightarrow v' \notin G_{vu}\}.$$

Для точек $P(G_v)$ не существует достижимых сочетаний агрегированных продуктов, которые лучше их хотя бы по одному продукту и не хуже по всем остальным. Для точек $S(G_v)$ не существует достижимых сочетаний продуктов, которые лучше их сразу по всем показателям. Ясно, что $P(G_v) \subset S(G_v)$.

Множество $P(G_v)$ в математической экономике принято называть *поверхностью эффективного выпуска*. Она характеризует предельные возможности изучаемой производственной системы. Эффективную границу (или несколько более широкое множество — слабо эффективную границу $S(G_v)$) принято считать теоретической производственной функцией изучаемой системы.

Часто при построении ПФ заранее известно, какие из составляющих вектора v производятся, а какие — потребляются (т. е. заранее известен знак координат вектора v). В этом случае можно ввести понятие вектора ресурсов x и вектора продуктов y . Пусть первые m_1 координат соответствуют ресурсам, а остальные m_2 — продуктам: $v = \{-x, y\}$. Тогда величины x и y будут неотрицательными, а ПФ приобретает вид $y(x)$.

Метод множеств достижимости по отношению к рассматриваемым производственным системам состоит в построении множества в явном виде и в изучении его двумерных сечений. При выборе координат вектора v , в которых будут рассматриваться двумерные сечения множества G_v , могут встретиться следующие 3 ситуации:

- обе координаты — продукты y_i и y_j ;
- обе координаты — ресурсы x_k и x_l ;
- одна координата — ресурс x_k , другая — продукт y_l .

В первом случае граница сечения является множеством наилучших сочетаний этих двух продуктов при заданных значениях ресурсов и других продуктов. При соответствующем выборе этих заданных значений такие границы будут соответствовать сечениям поверхности эффективного выпуска.

Во втором случае граница сечения является множеством наилучших сочетаний этих двух ресурсов при заданных значениях продуктов и других ресурсов. Для части таких двумерных сечений эти границы будут соответствовать сечениям поверхности эффективных затрат, т. е. изоквантам ПФ.

Однако, глядя на отдельное сечение, невозможно утверждать, является ли его «эффективная» граница сечением эффективной границы множества G_v . Для того чтобы это установить, требуется рассматривать значительное число сечений в этих координатах. Известен также подход с использованием Парето оболочки ОМД [23].

ПОСТРОЕНИЕ ПРОИЗВОДСТВЕННЫХ ФУНКЦИЙ С ПОСТОЯННОЙ ЭЛАСТИЧНОСТЬЮ ЗАМЕЩЕНИЯ ФАКТОРОВ

В основе сопоставления различных типов функций и оценки их возможности практического применения лежит главная характеристика функций — показатель эластичности замещения ресурсов, исходя из которой часто строятся варианты ПФ.

Из определения эластичности замены факторов по Михайлевскому с учетом того, что предельная норма замещения фактора i фактором j определяется как

$$s_{ij} = (\partial y / \partial x_j) / (\partial y / \partial x_i) = \partial x_i / \partial x_j,$$

получаем

$$\sigma_{ij}(x_1, \dots, x_n, dx_1, \dots, dx_n) = \frac{s_{ij} d\left(\frac{x_i}{x_j}\right)}{\left(\frac{x_i}{x_j}\right) ds_{ij}} = \frac{d\left(\ln \frac{x_i}{x_j}\right)}{d(\ln s_{ij})} = \sigma = \text{const},$$

или

$$d\left(\ln \frac{dx_i}{dx_j}\right) = \frac{1}{\sigma} d\left(\ln \frac{x_i}{x_j}\right), \quad d\left(\ln \left(\frac{dx_i}{dx_j}\right) - \ln \left(\frac{x_i}{x_j}\right)^{1/\sigma}\right) = 0.$$

Тогда

$$d\left(\ln \frac{dx_i / dx_j}{(x_i / x_j)^{1/\sigma}}\right) = 0 \Rightarrow \ln \frac{dx_i / dx_j}{(x_i / x_j)^{1/\sigma}} = \ln \left(-\frac{\alpha_j}{\alpha_i}\right),$$

где (α_j / α_i) — постоянная или переменная, не зависящая от x_i и x_j .

Таким образом, получили следующее дифференциальное уравнение:

$$\frac{dx_i}{dx_j} = -\frac{\alpha_j}{\alpha_i} \left(\frac{x_i}{x_j}\right)^{1/\sigma}, \quad i, j = 1, \dots, n, \quad i \neq j. \quad (247)$$

Решением уравнения (247) является изокванта производственной функции $y = f(x) = f(x_1, \dots, x_n)$. Геометрически для каждого значения σ на изокванте:

$$y(x_1^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_j, \dots, x_n^0) = y_0,$$

где y_0, x_k^0 ($k \neq i, j$) — постоянные, можно определить вид функции:

$$y(x_i, x_j, \alpha_i, \alpha_j) = y_0, \quad (248)$$

где постоянная y_0 (и поэтому α_i, α_j) на самом деле зависит от каждого набора определенных значений переменных x_k ($k \neq i, j$), которые считаются параметрами функции (248). Определенному значению y_0 соответствует определенная изокванта. Совокупность всех таких изоквант определяет вид функции y в подпространстве (x_i, x_j, y) . Заменяя (x_i, x_j) парой (x_i, x_k) , получаем функцию y в подпространстве (x_i, x_j, x_k, y) .

Продолжая процесс замены до последней пары, получим функцию $y = f(x) = f(x_1, \dots, x_n)$.

Рассмотрим случай, когда $(\alpha_i / \alpha_j) = \text{const}$. Тогда имеем:

$$\frac{\alpha_i dx_i}{(x_i)^{1/\sigma}} = -\frac{\alpha_j dx_j}{(x_j)^{1/\sigma}}.$$

Решая уравнение, получим:

$$\frac{\sigma}{\sigma-1} \alpha_i x_i^{\frac{\sigma-1}{\sigma}} + \frac{\sigma}{\sigma-1} \alpha_j x_j^{\frac{\sigma-1}{\sigma}} = y_0^{(1)}, \quad (249)$$

где $y_0^{(1)}$ зависит от $x_k \neq x_i, x_j$. Фиксируя x_i , будем дифференцировать (249) по x_k , считая $x_j = x_j(x_k)$, а все другие x_t постоянными.

Тогда

$$\alpha_j x_j^{-1/\sigma} \frac{dx_j}{dx_k} = \frac{dy_0^{(1)}}{dx_k}.$$

С учетом уравнения (247) имеем:

$$-\alpha_k x_k^{-1/\alpha} = \frac{dy_0^{(1)}}{dx_k}, \Rightarrow y_0^{(1)} = -\frac{\sigma}{\sigma-1} \alpha_k x_k^{\frac{\sigma-1}{\sigma}} y_0^{(2)},$$

где $y_0^{(2)}$ зависит от $x_i \neq x_j, x_k$. Процесс решения продолжается до получения следующего результата:

$$y_0^{(n-1)} = \frac{\sigma}{\sigma-1} \sum_{j=1}^n \alpha_j x_j^{\frac{\sigma-1}{\sigma}}.$$

Постоянная в левой части последнего уравнения в действительности является значением функции y на какой-либо изокванте. Для каждого набора значений (x_1, \dots, x_n) имеем определенное значение $y_0^{(n-1)}$. Отсюда получаем функцию

$$y = \frac{\sigma}{\sigma-1} \sum_{j=1}^n \alpha_j x_j^{\frac{\sigma-1}{\sigma}}. \quad (250)$$

Функция (250) обладает всеми свойствами производственной функции при $1 < \sigma < \infty$, $\alpha_j > 0$, $j = 1, \dots, n$, поэтому сама является производственной функцией. В самом деле,

$$y \geq 0, \quad \frac{\partial y}{\partial x_j} = \alpha_j x_j^{-1/\sigma} \geq 0, \quad \frac{\partial^2 y}{\partial x_j^2} = -\frac{\alpha_j}{\sigma} x_j^{-1/\sigma-1} \leq 0,$$

$$f(\lambda x) = \frac{\sigma}{\sigma-1} \sum_{j=1}^n \alpha_j (\lambda x_j)^{\frac{\sigma-1}{\sigma}} = \lambda^{\frac{\sigma-1}{\sigma}} f(x),$$

значит, $y = f(x)$ — однородная функция, однако отдача снижается от расширения производства, $0 < (\sigma - 1)/\sigma < 1$. После замены $y_0^{(1)} = (y_0^{(1)})^{1/a}$ получим

$$y_0^{(1)} = \left\{ \frac{\sigma}{\sigma-1} \left[\alpha_j x_i^{\frac{\sigma-1}{\sigma}} + \alpha_j x_j^{\frac{\sigma-1}{\sigma}} \right] \right\}^a.$$

Описанным выше способом получим функцию с постоянной эластичностью замены факторов в следующем виде:

$$y = \left\{ \frac{\sigma}{\sigma-1} \sum_{j=1}^n \alpha_j x_j^{\frac{\sigma-1}{\sigma}} \right\}^a, \quad (251)$$

где $0 < \sigma < \infty$ и $\sigma \neq 1$.

ПРОИЗВОДСТВЕННЫЕ ФУНКЦИИ
С ЕДИНИЧНОЙ ЭЛАСТИЧНОСТЬЮ ЗАМЕЩЕНИЯ ФАКТОРОВ

При $\sigma = 1$ уравнение (247) принимает вид:

$$dx_i/dx_j = -\alpha_j/\alpha_i(x_i/x_j). \quad (252)$$

Решая это уравнение, получим:

$$\alpha_i \ln x_i + \alpha_j \ln x_j = y_0^{(1)},$$

где $y_0^{(1)}$ зависит от $x_k \neq x_i, x_j$. Фиксируя x_i , будем дифференцировать (252) по x_k , считая $x_j = x_j(x_k)$, а все другие x_t постоянными.

Получаем:

$$\frac{\alpha_j}{x_j} \frac{dx_j}{dx_k} = \frac{dy_0^{(1)}}{dx_k}. \quad (253)$$

Поскольку

$$\frac{dx_j}{dx_k} = -\frac{\alpha_k}{\alpha_j} \left(\frac{x_j}{x_k} \right),$$

из соотношения (253) получаем:

$$dy_0^{(1)} = -\frac{\alpha_k}{x_k} dx_k, \text{ или } y_0^{(1)} = -\alpha_k \ln x_k + y_0^{(2)},$$

где $y_0^{(2)}$ зависит от $x_t \neq x_i, x_j, x_k$. Процесс решения продолжается до последней переменной. В результате имеем:

$$y = \sum_{j=1}^n \alpha_j \ln x_j. \quad (254)$$

Если считать, что $x_j = 1$ — минимальная единица j -го ресурса, необходимая для производства, то (254) является производственной функцией. В самом деле,

$$y \geq 0 \text{ при } x_j \geq 1 \text{ и } \alpha_j \geq 0, \partial y / \partial x_j = \alpha_j / x_j \geq 0, \partial^2 y / \partial x_j^2 = -\alpha_j / x_j^2 \leq 0.$$

Кроме того,

$$f(\lambda x) = \sum_{j=1}^n \alpha_j \ln(\lambda x_j) = \ln \lambda \sum_{j=1}^n \alpha_j + f(x) > f(x),$$

так как $\lambda > 1$. Если вместо $y_0^{(1)}$ подставить $\ln y_0^{(1)}$, то получим $y_0^{(1)} = x_i^{\alpha_i} x_j^{\alpha_j}$. Аналогично описанным выше способом получим функцию Кобба–Дугласа:

$$y = \prod_{j=1}^n x_j^{\alpha_j}. \quad (255)$$

Эта функция удовлетворяет всем свойствам производственной функции при $x_j \geq 0, 1 > \alpha_j > 0, j = 1, \dots, n$.

КОНКРЕТНЫЕ ВИДЫ ПРОИЗВОДСТВЕННЫХ ФУНКЦИЙ, ИХ ОСОБЕННОСТИ

Выделим из существующего многообразия некоторые производственные функции и отметим их наиболее важные особенности.

1а. Однородная ПФ:

$$f(\lambda x_1, \dots, \lambda x_n) = r(\lambda)f(x_1, \dots, x_n), r(1) = 1.$$

Одновременный рост всех ресурсов в одно и то же число λ раз вызывает рост выпуска, зависящий только от λ .

1б. Однородная ПФ порядка k

$$r(\lambda) = \lambda^k:$$

- основное значение для анализа динамики выпуска представляют не объемы применяемых ресурсов, а соотношения между ними;
- эластичности выпуска

$$\alpha_k = \frac{\partial f}{\partial x_i} \frac{x_i}{f}, i = \overline{1, n},$$

по факторам x_i (факторные эластичности) такие, что $0 < \alpha_i < 1$ (учитывается строгая вогнутость ПФ);

- произвольная степень однородности.
2. Однородная линейная ПФ порядка k

$$y = \left(\sum_{i=1}^n \alpha_i x_i \right)^k$$

- определенное количество i -го фактора всегда, независимо от степени насыщения им производства, может заменить единицу j -го фактора (в случае, когда норма замены постоянна), т. е. «вклад» ресурсов в выпуск пропорционален и не зависит от того, каковы были первоначальные размеры ресурсов;
- применяется для моделирования систем, в которых выпуск продукции является результатом одновременного функционирования множества различных технологий;
- отражает монотонно возрастающее производство.

3. ПФ Леонтьева

$$y = \min_i (f_i(x_1), \dots, f_n(x_n)),$$

где $f_i(x_i)$ — частная ПФ, описывающая i -й фактор:

- ресурсы жестко дополняют друг друга (замена невозможна);
- выпуск определяется одним ресурсом, образующим «узкое место» системы;

- применяется в практике плановых расчетов, в частности, по этой ПФ практически определяется потребность в ресурсах на определенную производственную программу;
- является решением задачи математического программирования (МП)

$$\begin{cases} y \rightarrow \max; \\ y \leq f_i(x_i), \quad i = 1, \dots, n. \end{cases}$$

4. Ординарная ПФ

$$y = f(x_1, x_2).$$

Назовем ординарной ПФ такую, у которой эластичность замещения факторов по Аллену определена и непрерывна по x , $\forall x > 0$, где $x = (x_1/x_2)$, ограничена и существенно отлична от единицы:

- такие функции не могут обладать степенью однородности γ больше единицы;
- ПФ вогнута (строго) $\Leftrightarrow 1/\sigma \geq 0$ и $\gamma \leq 1$ ($1/\sigma > 0$ и $\gamma < 1$).

5. Неординарная ПФ

$$y = f(x_1, x_2).$$

Функции, эластичность замещения которых близка к 1, интересны главным образом тем, что их факторные эластичности ограничены более узкими границами, чем интервал $(0, \gamma)$ (это делает ПФ более гибкой: свойства вогнутости $(\partial^2 y / \partial x_k^2) < 0, k = 1, 2$, не нарушаются при приближении к осям координат x_i и x_j . Такие функции, по-видимому, сохраняют адекватность реальности в более широкой области (x_1, x_2)):

- приближение $\sigma_{ij}^A \rightarrow 1 - 0$ при $x \rightarrow \infty$, где $x = (x_1/x_2)$ может интерпретироваться в макроэкономике как приближение свойств развитой технологии в интенсивном процессе капиталообразования;
- необходимым условием вогнутости является $0 < \gamma < \min[2; 1 + 1/\sigma]$.

6. ПФ Кобба–Дугласа

$$y = \alpha_0 x_1^{\alpha_1}, \dots, x_n^{\alpha_n} :$$

- если, по крайней мере, хотя бы один ресурс отсутствует, выпуск обращается в ноль;
- в результате роста любого ресурса выпуск также растет;
- неординарная ПФ;
- параметры $\alpha_i, i = 1, n$ представляют собой соответственно эластичность выпуска по факторам $x_i, i = 1, \dots, n$, т. е. увеличение

на 1% фактора x_i приводит к возрастанию на $\alpha_i\%$ результата производства;

- если $\alpha_1 + \dots + \alpha_n = 1$, то прирост производства пропорционален увеличению факторов x_i , $i = 1, \dots, n$;
- если $\alpha_1 + \dots + \alpha_n < 1$, то при одновременном возрастании на 1% используемого объема всех факторов результат производства возрастет менее чем на 1%;
- если $\alpha_1 + \dots + \alpha_n > 1$, то наблюдается возрастающая эффективность использования факторов x_i , $i = 1, \dots, n$;
- строго квазивогнута при $\forall \gamma$;
- прирост выпуска, вызванный изменением i -го ресурса на 1%, не зависит от того, каковы были первоначальные размеры ресурсов;
- отсутствие экстремальных точек, т. е. с ростом факторов выпуск не может сокращаться;
- отсутствие точек перегиба обуславливает отражение функций эффективности расширения производства только одного типа — постоянной, возрастающей или убывающей. Функция не может охватывать все три типа одновременно.

7. ПФ Кобба–Дугласа с переменной эффективностью масштаба:

$$y = (\alpha_0 x_1^{\alpha_1}, \dots, x_n^{\alpha_n})^{\varphi(y)}.$$

При $\varphi(y) > 0$, $y > 0$ ПФ приводится к виду:

$$\frac{1}{y^{\varphi(y)}} = \alpha_0 x_1^{\alpha_1}, \dots, x_n^{\alpha_n}.$$

8. ПФ постоянной эластичности замены факторов по Михайлевскому (функция CES):

$$y = (\alpha_1 x_1^{\alpha_{n+1}} + \dots + \alpha_n x_n^{\alpha_{n+1}})^{\alpha_{n+2}}.$$

Применяется в случаях, когда отсутствует точная информация об уровне взаимозаменяемости производственных факторов и вместе с тем есть основания предполагать, что этот уровень существенно не изменяется при изменении объемов вовлекаемых ресурсов:

- ординарная ПФ;
- при $k = 1$, $\alpha_3 > 0$, (исследуем вторую запись ПФ), рассмотрим уравнения изоквант; уравнения асимптот следующие:

$$x_1 = \frac{\alpha_1^{1/\alpha_3}}{\alpha_0} y_0 \quad \text{и} \quad x_2 = \frac{(1-\alpha)_1^{1/\alpha_3}}{\alpha_0} y_0.$$

Итак, если данного выпуска y_0 нельзя достичь при затратах x_2 , меньших

$$x_2^0 = \frac{y_0}{\alpha_0} (1-\alpha)_1^{1/\alpha_3},$$

то и выпуска, большего $\alpha_0(1-\alpha_1)^{-1/\alpha_3} x_2^0$, нельзя достичь при затратах x_2^0 , причем и сам этот выпуск недостижим при конечном x_1 ;

- при эластичности замены между двумя любыми факторами $\sigma \rightarrow \infty$ функция CES превращается в однородную линейную ПФ порядка k ;
- при $\sigma \rightarrow 0+$ функция CES переходит в ПФ

$$y = \min_i [(f_i(x_1), \dots, f_n(x_n))]^k,$$

называемую функцией Леонтьева–Харрода–Домара;

- при $\sigma \rightarrow 0$ функция CES переходит в ПФ

$$y = \max_i [(f_i(x_1), \dots, f_n(x_n))]^k,$$

где $f_i(x_i)$ — частная ПФ, описывающая i -й фактор;

- при $\sigma \rightarrow 1$ функция CES переходит в ПФ Кобба–Дугласа;
 - предпочтительна в работах, основной задачей которых является изучение проблем замещения факторов в производстве.
9. ПФ Аллена:

$$y = \alpha_0 x_1 x_2 - \alpha_1 x_1^2 - \alpha_2 x_2^2.$$

Предназначена для описания производственных процессов, в которых чрезмерный рост любого из факторов оказывает отрицательное воздействие на объем выпуска, т. е. при $\alpha_1, \alpha_2 > 0$

$$\lim_{x_1 \rightarrow \infty} y = -\infty, \quad \lim_{x_2 \rightarrow \infty} y = -\infty.$$

10. Однородная ПФ с линейной эластичностью замены факторов по Аллену (функция LES)

$$y = x_1^{\alpha_0} (\alpha_1 x_1 + \alpha_2 x_2)^{\alpha_3}.$$

Рекомендуется для описания производственных процессов, у которых возможность замещения вовлекаемых факторов существенно зависит от их пропорций.

$$\sigma^A(x_1, x_2) = 1 + c(x_1/x_2),$$

где $c = \alpha_1/\alpha_2 + \alpha_1\alpha_3/(\alpha_0\alpha_2)$.

11. «Параллельная» ПФ Леонтьева

$$y = \min(f_{11}(x_1), \dots, f_{1n}(x_n)) + \dots + \min(f_{k1}(x_1), \dots, f_{kn}(x_n)),$$

где $f_{ij}(x_j)$ — ПФ, описывающая j -й фактор в i -м процессе.

Отражает процесс, в котором объем выпуска складывается из выпусков параллельных производственных процессов, использующих общие ресурсы.

12. ПФ Солоу

$$y = (\alpha_1 x_1^{\alpha_{n+1}} + \dots + \alpha_n x_n^{\alpha_{2n}})^{\alpha_{2n+1}}.$$

Функцию можно рассматривать как линейную относительно новых переменных

$$z = y^{1/\alpha_{n+1}}, \quad \tilde{x}_1 = x_1^{\alpha_{n+1}}, \dots, \tilde{x}_2 = x_2^{\alpha_{2n}}.$$

Рекомендуется использовать в тех случаях, когда предположение об однородности представляется неоправданным.

13. Ограниченная ПФ с постоянной эластичностью замены

$$y = \min\{f_1(x_1), \dots, f_k(x_k), (\alpha_1 x_{k+1}^{\alpha_{n+1}}, \dots, \alpha_n x_{k+n}^{\alpha_{n+1}})^{\alpha_{n+2}}\},$$

где $f_1(x_i)$ — частная ПФ, описывающая i -й фактор:

- между переменными x_1, \dots, x_k и группой остальных переменных отсутствует взаимозаменяемость;
- аналогична функции CES, если ее значения не превосходят пороговой величины $\min\{f_i(x_i), i = 1, k\}$; в области превышения этого порога «включается» ПФ Леонтьева;
- решение задачи МП:

$$\begin{cases} y \rightarrow \max; \\ y_i \leq f_i(x_i), \quad i = 1, k; \\ y \leq (\alpha_1 x_{k+1}^{\alpha_{n+1}} + \dots + \alpha_n x_{k+n}^{\alpha_{n+1}})^{\alpha_{n+2}}. \end{cases}$$

14. Многорежимная ПФ

$$y = (\alpha_{11} x_1^{\beta_1}, \dots, \alpha_{1n} x_n^{\beta_1})^{\gamma_1} \times \dots \times (\alpha_{k1} x_1^{\beta_k}, \dots, \alpha_{kn} x_n^{\beta_k})^{\gamma_k}.$$

Используется при описании процессов, в которых уровень отдачи каждой новой единицы ресурса скачкообразно меняется в зависимости от соотношения факторов.

Пример. Пусть

$$y = \alpha_0 (x_2^{\beta_2} - \alpha_2 x_1^{\beta_2})^{-1} \dots \alpha_0 (x_n^{\beta_n} - \alpha_n x_1^{\beta_n})^{-1},$$

где $\beta_2, \dots, \beta_n < 0, \alpha_2, \dots, \alpha_n > 0$.

Если $x_2/x_1 \gg 1/\alpha_2, \dots, x_n/x_1 \gg 1/\alpha_n$, то функция близка к степенной функции $y = \alpha_0 x_1^{\beta_2 + \dots + \beta_n}$ и слабо зависит от x_2, \dots, x_n .

Если $x_i/x_1 \gg 1/\alpha_i, i = 2, n$, то функция близка к ПФ Кобба–Дугласа $y = \alpha_0 x_1^{\beta_2} \dots x_n^{\beta_n}$ и слабо зависит от x_1 .

15. ПФ Сато

$$y = \alpha_0 x_1^{\alpha_1} x_2^{\alpha_2} (\alpha_3 x_1^{-\alpha_4} + (1 - \alpha_3) x_2^{-\alpha_4})^{\frac{-\alpha_5}{\alpha_4}} = y_1(x_1, x_2) y_2(x_1, x_2):$$

- предполагается, что $y_1(x_1, x_2)$ и $y_2(x_1, x_2)$ строго положительны внутри первого ортанта, монотонны по аргументам;
- степень однородности ПФ Сато $\gamma = \alpha_1 + \alpha_2 + \alpha_5 > 0$;
- для строгой вогнутости по координатам необходимо выполнение $0 \leq \alpha_1 \leq 1, 0 \leq \alpha_2 \leq 1$ (если $0 \leq \sigma^A \leq 1, \forall x_1 > 0, x_2 > 0$, то условие является и достаточным);

- каковы бы ни были $\bar{\gamma}$ и $\bar{\sigma}^A$, удовлетворяющие неравенству $0 < \gamma < \min[2, 1 + (1/\sigma^A)]$, может быть построена ПФ Сато, строго вогнутая по координатам, такая, что $\gamma = \bar{\gamma}$ и $\sigma^A(\bar{x}_1/\bar{x}_2) = \bar{\sigma}^A$ в некоторой точке $(\bar{x}_1/\bar{x}_2) \in (0, \infty)$. При этом если $\bar{\sigma}^A > 0$, то $\sigma^A(x_1/x_2)$ определена и положительна при всех $(x_1/x_2) > 0$.

16. Гомотетичная ПФ

$$y = f(g(x)),$$

где $g(x)$ — линейно-однородна, $f(g)$ — непрерывна, неотрицательна и монотонно возрастает:

- рассматривается гипотеза о существовании определенного эффекта при переходе к новым масштабам, который, однако, не зависит от самих масштабов затрат;
- норма замены ресурсов инвариантна относительно гомотетичного преобразования.

Пример

$$f(x) = a(g(x))^m, \text{ где } a, m = \text{const};$$

- имеет место спецификация однородной степени m функции.

17. Негомотетичная ПФ:

- нормы взаимозаменяемости ресурсов зависят не только от пропорций в затратах, но и от уровня выпуска;
- пусть норма замены $w = ay^\gamma(x_2/x_1)^{e+1}$, где a, e, γ — const; если $\gamma \neq 0$, имеем еще один показатель эластичности замещения; его можно называть масштабным: $\sigma y = \partial \ln(x_2/x_1) / \partial \ln y = -\gamma / (e + 1)$;
- рост масштабов производства может как увеличивать трудности взаимозамещения ресурсов, так и способствовать их снижению, если σy соответственно больше или меньше нуля.

18. Однородная ПФ с произвольной эластичностью замены по Алену (функция VES)

$$y = \alpha_0 x_1^{\alpha_1} \left(- \int \frac{dr}{r + \alpha_2 r^2 \exp\left(-\int \frac{dr}{r\sigma(r)}\right)} \right):$$

- функция VES однородна степени α_1 ;
- $\sigma^A(x_1, x_2) = \varepsilon\sigma(r)$, где $r = x_1/x_2$, $\sigma(r)$ — заданная непрерывная функция от одного переменного;
- при $\sigma(r) = \text{const}$ получаем однородную функцию CES.

19. Краткосрочная и долгосрочная ПФ.

Различия между кратко- и долгосрочными аспектами производства связаны прежде всего с различными возможностями замещения факторов — аргументов ПФ. В течение долгосрочного

периода осуществимость замещения определяется только технологией, т. е. накопленным фондом реализуемых технических знаний. Реально изменение соотношения между трудом и капитальными фондами не может быть произведено быстро не только в силу технологических ограничений, но и из-за неэластичности уже занятого в производстве парка фондов. Коль скоро оборудование установлено и пущено, а рабочие заняты, замещение факторов существенно затрудняется. Продолжительность краткосрочного периода зависит от ряда причин и, прежде всего, от продолжительности срока службы применяемого оборудования.

Можно ожидать, что эластичность замещения ПФ краткосрочного периода будет меньше эластичности долгосрочной ПФ.

Долгосрочная ПФ вряд ли может быть нейтрально зависящей от времени, ибо технологические эпохи отделяются друг от друга крупными структурными изменениями, научно-техническими революциями, изменениями в социальной организации.

НТП В ПРОИЗВОДСТВЕННЫХ ФУНКЦИЯХ

ОТРАЖЕНИЕ ТЕХНИЧЕСКОГО ПРОГРЕССА В ПФ

Понятие НТП трактуется в теории ПФ расширительно.

Основным направлением технического процесса (ТП) является совершенствование средств труда — «материализация» в них достижений науки. С другой стороны, имеет место ТП, который можно назвать нематериализованным. Не воплощаясь в конкретных материальных носителях, он способствует росту эффективности производства, проявляется в совершенствовании организации и управления производством и часто связан с использованием накопленного опыта и эмпирических знаний. Кроме того, что действительно отражает НТП, сюда иногда включаются всевозможные изменения свойств производственного объекта, возникающие во времени, в том числе, например, и просто влияние неучтенных в основном списке факторов. В это понятие войдут также изменения, связанные с изменениями внешней для данной системы среды и т. п.

Другое различие форм ТП связано с влиянием его на основные соотношения и показатели, характеризующие экономический рост. Так, например, нововведения могут менять соотношения факторов производства и влиять на их замещение: они могут быть направлены на относительную экономию труда или

относительную экономию основных фондов. С другой стороны, ТП может быть нейтральным по фондовооруженности труда, т. е. не изменять соотношения факторов. С введением новшеств выпуск продукции можно увеличивать при том же количестве затрат живого и овеществленного труда.

В наиболее обобщенной форме технический прогресс в ПФ представляется следующим образом:

$$y = f(x_1, \dots, x_n, T), \quad (256)$$

где T — вектор параметров, каждое значение которого выражает определенный уровень технического развития. В качестве этого параметра часто используется показатель времени t . Под ТП понимается монотонное возрастание функции $f(x_1, \dots, x_n, T)$ по T .

ЭКЗОГЕННЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ ПРОГРЕСС

Нейтральный технический прогресс. Возможным путем решения проблемы выбора конкретной формы ПФ (256) является принятие допущения, что ТП определенным образом влияет на связи между конкретными макроэкономическими показателями. В частном случае эти связи могут быть инвариантными, «нейтральными» относительно технических изменений. При этом выясняется, что заданный таким образом тип технического прогресса и форма ПФ часто не являются независимыми. Рассмотрим некоторые формы НТП, отметим их особенности.

1. Автономный технический прогресс (нейтральность НТП по Хиксу)

$$y = A(t)f(K, L): \quad (257)$$

- такой способ учета НТП основан на гипотезе о том, что технические изменения проявляются не через отдельные ресурсы или способ их соединения, а лишь через общую интенсификацию производства;
- нейтральность НТП по отношению ко всем ресурсам одновременно.

Примеры:

а) ПФ Кобба–Дугласа–Тинбергена

$$y = ae^{\gamma t} K^\lambda L^{1-\lambda}.$$

Действие НТП стационарное, т. е. ежегодно увеличивающее интенсивность выпуска в одно и то же число раз;

б) пусть

$$y = A(t)f(K, L),$$

где f — функция, однородная степени k .

- ТП проявляется, даже если другие факторы не меняются;
- ТП в равной мере действует на старые и вновь введенные ресурсы;
- такой подход приемлем для коротких интервалов времени;
- ТП — следствие улучшения организации труда, управления производством и т. д.

2. Нейтральность ТП по фондовооруженности труда.

Предположим, что зависимость между предельной нормой технической замены и фондовооруженностью труда инвариантна относительно технического прогресса:

$$\frac{\partial Y}{\partial K} / \frac{\partial Y}{\partial L} = \Phi\left(\frac{K}{L}\right) \quad (258)$$

(влияние ТП на это соотношение можно было бы учесть, записав зависимость (258) в форме:

$$\frac{\partial Y}{\partial K} / \frac{\partial Y}{\partial L} = \Phi\left(\frac{K}{L}, t\right),$$

где t — время). Разделив переменные и интегрируя, получим:

$$Y = A(Y, t)G(t, H(K, L)),$$

где $A = A(Y, t)$ — константа интегрирования, G — гомотетичная ПФ, H — однородная первой степени ПФ.

Здесь t как фактор взаимодействует с функцией факторов производства $H(K, L)$.

В данном случае можно говорить о техническом прогрессе, увеличивающем выпуск.

3. Нейтральность ТП по фондоотдаче (нейтральность по Харроду).

Пусть зависимость между средней и предельной производительностью производственных фондов инвариантна относительно технического прогресса, т. е.

$$\partial Y / \partial K = \Phi(Y/L). \quad (259)$$

Разделив переменные и интегрируя, получим:

$$Y = H(K, A(L, t)),$$

где H — линейно-однородна по $A(L, t)$ и K , $A(L, t)$ является константой интегрирования (отметим, что однородность функции H не предполагалась с самого начала, но была выведена как свойство конкретной инвариантной формы).

Этот случай отражает увеличение эффективности труда, которое может быть связано не только с ростом навыков занятых

в производстве, но и с вводом в действие более прогрессивного оборудования.

Данный тип ТП аналогичен увеличению числа занятых в материальном производстве, т. е. является трудодобавляющим. Он позволяет работнику сделать больше, чем тот делал раньше, и связан с изменениями в производстве, направленными на экономию труда.

4. Нейтральность ТП по производительности труда (нейтральность по Солоу).

Соотношение

$$\partial Y / \partial L = \Phi(Y/L), \quad (260)$$

инвариантное относительно технических изменений, дает линейно-однородную ПФ вида

$$Y = H[A(K, t), L],$$

где H — линейно-однородная функция относительно $A(K, t)$ и L , $A(K, t)$ — константа интегрирования.

Данный тип ТП является фондодобавляющим, т. е. экономиями производственные фонды, увеличивающим их эффективность.

5. Пример ПФ Кобба–Дугласа:

$$y = A(t)K^{\alpha_1}L^{\alpha_2}.$$

ПФ нейтральна по фондовооруженности, фондоотдаче и производительности труда одновременно.

6. Общий случай фактородобавляющего ТП:

$$Y = F[H(A(t)K, B(t)L)]: \quad (261)$$

- ПФ (261) есть монотонное преобразование линейно-однородной функции с фактороувеличивающимся техническим прогрессом;
- фактородобавляющий ТП является приближением к реальности, так как он отражает тот факт, что технический прогресс, как правило, меняет эффективность живого и овеществленного труда.

7. Частный случай ПФ (261):

$$Y = H(A(t)K, B(t)L): \quad (262)$$

- ПФ (27) является обобщением трех предыдущих случаев технического прогресса;
- если $A(t)/B(t) \equiv \text{const}$, то имеем увеличивающий выпуск ТП;
- если $A(t) \equiv \text{const}$, то — трудодобавляющий ТП;
- если $B(t) \equiv \text{const}$, то — фондодобавляющий ТП.

ЭКЗОГЕННЫЙ И ЭНДОГЕННЫЙ НТП

Перечисленные типы НТП отражают так называемый экзогенный прогресс, т. е. технологические изменения, происхождение которых не обуславливается внутри самой модели (производственной функции), а лишь учитывается ею. Существуют, однако, модели, в которых происхождение НТП определяется внутри данной модели (например, политикой капиталовложений в отраслевые НИИ), зависит от действий участников экономической системы (опытное производство и т. п.), моделируемой ПФ. НТП такого типа называется эндогенным.

ПРИМЕРЫ МАКРОЭКОНОМИЧЕСКИХ И МИКРОЭКОНОМИЧЕСКИХ МОДЕЛЕЙ

ЭКЗОГЕННЫЙ НТП В ПРОИЗВОДСТВЕННОЙ ФУНКЦИИ

В реальной экономике НТП действует не аналогично отдельному автономному ресурсу, а так, что технические изменения происходят в некоторой взаимосвязи с динамикой основных фондов. Другими словами, НТП проявляется и в повышении качества отдельных ресурсов, в основном за счет капиталовложений, и в совершенствовании их взаимосвязей — прогресс организации, управления, прогресс технологии на имеющихся ресурсах и т. п. Представляют интерес попытки представления НТП как следствия капиталовложений, главным образом капиталовложений в основные производственные фонды [97].

Американский экономист Р. Солоу в 1953 г. предложил ПФ

$$y = A\Phi^\lambda(t)L^{1-\lambda}(t), \quad (263)$$

где

$$\Phi = e^{-\mu t} \int_{-\infty}^t e^{(\mu+\beta/2)\tau} k(\tau) d\tau, \quad (264)$$

Φ — расчетный объем основных фондов; μ — норма выбытия фондов; β — темп НТП, проявляющегося в качестве вновь вводимых основных фондов; $k(\tau)$ — капиталовложения в основные производственные фонды в году τ . В рамках этой ПФ НТП не влечет никаких изменений в уровне производства, пока не будет овеществлен в основных фондах.

Функция Солоу является упрощенным вариантом функции

$$y(t) = \int_{-\infty}^t e^{\beta t} \Phi^\lambda(\tau, t) L^{1-\lambda}(\tau, t) d\tau, \quad (265)$$

где $\Phi(\tau, t)$ — объем основных производственных фондов, введенных в году τ и действующих в году t ; $L(\tau, t)$ — объем занятости, связанной с этими фондами; β — темп НТП.

Здесь предполагается, что общий объем продукции складывается из продукции, созданной на фондах различных возрастов. В этом случае в качестве показателя выпуска может использоваться показатель чистой продукции.

УРАВНЕНИЕ КАПИТАЛОВЛОЖЕНИЙ В ОСНОВНЫЕ ФОНДЫ

Вызывает интерес описанное в работе [9] уравнение решения фирм относительно капиталовложений в основные фонды. Рассмотрим это уравнение. Предположим, что

$$\frac{(dK/dt)}{K} = b \log \left(\frac{P}{(1+c)r p_k K} \right), \quad (266)$$

где r — норма процента на заемные средства; p_k — цена капитальных ресурсов; P — реальная прибыль; $b > 0$, $c > 0$ — константы (c можно рассматривать как премию за риск); $P/K p_k$ — норма прибыли на существующий капитал.

Обоснование уравнения капиталовложений в основные фонды:

- геометрический темп роста объема основного капитала в отрасли увеличивается при повышении отношения $(P/K p_k)$ и при уменьшении r ;
- член $(1+c)r$ выражает норму прибыли, требующуюся фирмам для поддержания имеющегося объема основного капитала;
- если норма прибыли падает ниже $(1+c)r$, то часть основных фондов не будет заменяться по мере износа и (dK/dt) будет меньше нуля;
- если же норма прибыли превысит величину $(1+c)r$, то закупки новых капитальных ресурсов будут превышать потребности в замене и (dK/dt) будет больше нуля;
- член $(1+c)r p_k$ можно рассматривать как цену (включающую премию за риск) использования единицы основного капитала.

Будем называть функцию в правой части уравнения (266) функцией капиталовложений.

МАКРОЭКОНОМИЧЕСКАЯ МОДЕЛЬ С ЭКЗОГЕННЫМ ТП

Рассмотрим односекторную замкнутую модель экономики, динамика которой определяется тем, в какой пропорции делится национальный доход между капиталовложениями на расширение основных фондов и на улучшение производства («на науку») [4].

Для простоты считаем, что потребление отсутствует. Можно было бы, например, считать, что норма потребления фиксированна, т. е. составляет заданную долю национального дохода, но это то же самое, что рассматривать другую производственную функцию.

Будем считать, что вложения в «науку» действуют таким образом, что мы имеем дело с прогрессом, нейтральным по Хиксу, т. е.

$$Y = A(Q)F(K, L), \quad (267)$$

где $F(K, L)$ — ПФ, Q — суммарный объем капиталовложений в НТП. Пусть

$$F(K, L) = g(K)h(L) \quad A'(Q) > 0, \quad A''(Q) < 0, \quad A(0) = 1, \\ \lim_{Q \rightarrow \infty} A(Q) = \infty, \quad \lim_{Q \rightarrow \infty} A'(Q) = 0, \quad \lim_{Q \rightarrow 0} A(Q) = \infty.$$

Пусть $u(0 \leq u \leq 1)$ — норма накопления, т. е. доля национального дохода Y , направляемая на увеличение основных фондов. Тогда

$$\dot{K} = uY, \quad (268)$$

величина $(1 - u)$ представляет собой долю национального дохода Y , направляемую в НТП, на «науку». Значит,

$$\dot{Q} = (1 - u)Y. \quad (269)$$

Пусть

$$\dot{L} = p(L), \quad (270)$$

где $p(L)$ — такая функция, что величина L не может стать бесконечной за конечное время. Пусть (K_0, Q_0, L_0) — начальное состояние. Задание управления $u(t)$, $t \geq 0$, $0 \leq u(t) \leq 1$ позволяет из соотношений (268)–(270) определить траекторию $(K(t), Q(t), L(t))$.

УРАВНЕНИЕ ЭКОНОМИЧЕСКОГО РОСТА

Обозначим через Y — национальный доход; K — капитал системы; t — время; L — численность занятых в сфере производства национального дохода; P — потребление занятых вне сферы производства национального дохода.

Пусть

$$Y = F(K, L, t); \quad (271)$$

$$Y = \frac{\partial F}{\partial L} L + \frac{\partial F}{\partial K} K, \quad (272)$$

т. е. выпуск — линейно-однородная функция по K и L ;

$$\frac{\partial F}{\partial L} = \frac{Y - \dot{K} - P}{L}, \quad (273)$$

т. е. предполагается, что занятость на каждом элементе экономики регулируется условием максимизации прибыли;

$$\partial F / \partial t = vY, \quad (274)$$

где v — темп роста эффективности экономики (т. е. предполагается, что эффект от НТП и организации процесса пропорционален масштабу экономики).

Продифференцировав связь (263) и используя соотношения (272)–(274), получим:

$$\frac{\dot{Y}}{Y} = \frac{\dot{K} + P}{Y} \frac{\dot{K}}{K} + \left(1 - \frac{\dot{K} + P}{Y}\right) \frac{\dot{L}}{L} + v. \quad (275)$$

Положим

$$K_1 = K \left(1 + \frac{P}{K}\right), \quad \frac{\dot{K}_1}{K_1} = \frac{\dot{K}}{K}.$$

Тогда

$$\frac{\dot{Y}}{Y} = \frac{\dot{K}_1^2}{YK_1} + \left(1 - \frac{\dot{K}_1}{Y}\right) \frac{\dot{L}}{L} + v. \quad (276)$$

Далее параметры v и $v_L = \dot{L}/L$ считаются постоянными и заданными. Таким образом, в уравнение (276) входят две неизвестные переменные Y и K_1 . Чтобы замкнуть описание модели, необходимо или задать цель развития в виде функционала, или задать каким-то образом еще одну связь между Y и K_1 .

Итак, при моделировании развития экономических систем вместо непосредственного задания ПФ с последующей идентификацией ее параметров можно использовать уравнение экономического роста (275), в котором необходимо задавать только один параметр v и которое позволяет рассматривать проблему развития как проблему управления [26].

ПОСТРОЕНИЕ ПРОИЗВОДСТВЕННЫХ ФУНКЦИЙ В УСЛОВИЯХ КОНКРЕТНОЙ ДЕЙСТВИТЕЛЬНОСТИ

Каждая производственная функция представляется конкретным алгебраическим уравнением, поэтому основная задача исследования состоит в том, чтобы в каждом отдельном случае подобрать такое уравнение, которое как можно ближе воспроизводило бы исследуемый производственный процесс или его отдельные стороны.

Построению математической модели производственной функции должен предшествовать качественный анализ моделируемого экономического процесса. Экономико-статистическое моделирование включает:

- выбор зависимых переменных и отбор независимых переменных факторов;
- получение статистических данных и их обработка;
- установление математической формы связи;
- решение модели;
- корректировка и экономическая интерпретация модели.

Выбор зависимой переменной производится на основе логического анализа возможных взаимосвязей между признаками, формирующими данный производственный процесс. За зависимую переменную принимают такой показатель, который, исходя из поставленной задачи исследования, наиболее полно характеризует изучаемые производственные взаимосвязи, экономический процесс. В сложном переплетении явлений экономической жизни практически невозможно учесть влияние всех факторов. В зависимости от цели исследования необходимо из всего множества отбирать только те количественно измеримые факторы, которые оказывают наиболее существенное влияние на изменение производственного результата. Включаемые в модель факторы не должны находиться между собой в функциональной связи, так как тогда они будут характеризовать одну и ту же сторону изучаемого явления и косвенно дублировать друг друга.

Под подготовкой исходной информации понимают получение данных и их обработку. Статистические данные должны отвечать ряду требований. Во-первых, объем выборки должен быть достаточно большим, чтобы полученные результаты с заданной вероятностью достоверно характеризовали бы генеральную совокупность; во-вторых, выборка должна быть однородной, и в-третьих, исходные данные должны быть выражены количественно.

Один из самых сложных этапов экономико-статистического моделирования — определение математической формы связи между переменными изучаемого экономического процесса в виде алгебраического уравнения. Уравнение должно, с одной стороны, правильно отражать основную тенденцию связи между переменными, а с другой — поддаваться дальнейшей математической обработке и не быть слишком сложным.

Построенную модель производственной функции необходимо решить, т. е. функция $y = f(x_1, \dots, x_n, a_1, \dots, a_m)$ рассматрива-

ется как регрессионное уравнение, для которого необходимо найти оценки параметров a_1, \dots, a_m .

Составная часть моделирования производственных функций — корректировка результатов решения полученной модели. В основе корректировки модели должно быть тесное сочетание количественных и главным образом качественных методов оценки. Количественную оценку выполняют с помощью различных статистических коэффициентов и критериев. Суть качественного анализа состоит, прежде всего, в истолковании смысла вычисленных параметров. Здесь в первую очередь необходимо проверить, соответствуют ли знаки числовых значений параметров принятым логическим представлениям о направлении влияния отобранных факторов на исследуемый производственный результат. Наличие этого соответствия будет свидетельствовать о том, что с логической стороны полученную модель можно считать пригодной для дальнейшего использования.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Абрагимов А. П. и др.* Исследование динамики факторных эластичностей производственных функций. М., 1988.
2. *Аллен Р.* Математическая экономия. М.: ИЛ, 1963.
3. *Андерсон Т.* Статистический анализ временных рядов. М.: Мир, 1976.
4. *Ашманов С. А.* Математические модели в экономике. Ч. 2. М., 1987.
5. *Ашманов С. А.* Математические модели и методы в экономике. М.: МГУ, 1980.
6. *Барбашин Е. А.* Введение в теорию устойчивости. М.: Наука, 1967.
7. *Баркалов Н. Б.* Производственные функции в моделях экономического роста. М.: МГУ, 1981.
8. *Беллман Р., Кук К.* Дифференциально-разностные уравнения. М.: Мир, 1967.
9. *Бергстром А.* Построение и применение экономических моделей. М.: Прогресс, 1970.
10. *Березин И. С., Жидков Н. П.* Методы вычислений: В 2 т. М.: Физматгиз, 1959.
11. *Булавский В. А.* Управление динамикой экономической системы через внешнее потребление // Экономика и математические методы. 1995. № 2.
12. *Вечканов Г. С., Вечканова Г. Р., Пуляев В. Т.* Краткая экономическая энциклопедия. СПб.: Петрополис, 1998.
13. *Винокуров В. Р.* Асимптотическое поведение решений одного класса интегродифференциальных уравнений Вольтерра // Дифференциальные уравнения. 1967. Т. 3. № 7. С. 1095–1099.
14. *Витте С. Ю.* Конспект лекций о народном и государственном хозяйстве. СПб., 1912.
15. *Вишнев С. М.* Экономические параметры. Введение в теорию показателей экономических систем и моделей. М., 1968.
16. *Вольтерра В.* Математическая теория борьбы за существование. М.: Наука, 1976.
17. *Воркуев Б. Л.* Модели макро- и микроэкономики. М.: Теис, 1999.

18. *Вороновицкий М. М.* Модель инфляции в экономике переходного периода // Экономика и математические методы. 1994. Т. 30. Вып. 3. С. 117–128.
19. *Гантмахер Ф. Р.* Теория матриц. М.: Наука, 1967.
20. *Гейл Д.* Замкнутая линейная модель производства // Линейные неравенства и смежные вопросы. М.: ИЛ, 1959.
21. *Гершензон М. А.* Анализ упрощенных динамических моделей межотраслевого баланса. Новосибирск: Наука, 1975.
22. *Гранберг А. Г.* Динамические модели народного хозяйства. М.: Экономика, 1985.
23. *Джолдыбаева С. М., Лотов А. В.* Агрегированные производственные функции линейных моделей. М., 1989.
24. *Д'Отюм А., Шараев Ю.* Образование и эндогенный экономический рост: Модель Лукаса. М., 1998.
25. *Друнченко В. П.* Теорема о существовании и единственности решения задачи Коши, дифференциального уравнения с запаздывающим аргументом, имеющим кусочно-непрерывную начальную функцию // Вестн. СПбГУ. 1996. Вып. 1. С. 25–30.
26. *Дубовский С. В.* Математические модели экономических процессов: Обзор. М., 1977.
27. *Дубовский С. В., Осипов С. Н.* Вопросы структурной устойчивости в модели экономического роста и циклов // Автоматика и телемеханика. 1994. № 9.
28. *Ермаков С. М., Жиглявский А. А.* Математическая теория оптимального эксперимента. М.: Наука, 1987.
29. *Ефимов М. Н., Мовшович С. М.* Анализ сбалансированного роста в динамической модели народного хозяйства // Экономика и математические методы. 1973. № 1.
30. *Заславский А. Я.* Теоремы о магистрали в моделях с переменной технологией // Экономика и математические методы. 1989. № 1.
31. *Зубов В. И.* Колебания и волны. Л.: Изд-во ЛГУ, 1989.
32. *Зубов В. И.* Лекции по теории управления. М.: Наука, 1975.
33. *Зубов В. И., Петросян Л. А.* Математические методы в планировании. Л.: Изд-во ЛГУ, 1982.
34. *Иванов Ю. П. и др.* Производственная народно-хозяйственная функция. М., 1983.
35. *Иванов Ю. П., Ланец С. А.* Анализ и построение производственных функций с переменной эластичностью замещения по ресурсам. М., 1984.
36. *Иванов Ю. П., Лотов А. В.* Математические модели в экономике. М.: Наука, 1979.
37. *Иванов Ю. Н., Токарев В. В., Уздемир А. П.* Математическое описание элементов экономики. М.: Физматлит, 1994.
38. *Интрилигатор М.* Математические методы оптимизации и экономическая теория. М.: Прогресс, 1975.
39. *Калман Р., Фалб П., Арbib М.* Очерки по математической теории систем. М.: Мир, 1971.
40. *Канторович Л. В.* Математические методы организации и планирования производства. Л., 1939.

41. *Канторович Л. В. и др.* Современный математический аппарат управления экономикой. Свердловск, 1974.
42. *Карлин С.* Математические методы в теории игр, программировании и экономике. М.: Мир, 1964.
43. *Клейнер Г. Б.* Методы анализа производственных функций. М., 1980.
44. *Клейнер Г. Б.* Производственные функции: теория, методы, применение. М., 1986.
45. *Колмогоров А. Н.* Качественное изучение математических моделей динамики популяций // Проблемы кибернетики. М.: Наука, 1972. Вып. 25.
46. *Кондратьев Н. Д.* Особое мнение. Избранные произведения: В 2 кн. М.: Наука, 1993. Кн 2.
47. *Коссов В. В.* Межотраслевые модели. М.: Экономика, 1973.
48. *Красовский Н. Н.* Некоторые задачи теории устойчивости движения. М.: Физматгиз, 1959.
49. *Кричевский И. Р.* Понятия и основы термодинамики. М.: Химия, 1970.
50. *Кузютин В. Ф., Зенкевич Н. А., Еремеев В. В.* Геометрия. СПб.: Изд-во «Лань», 2003.
51. *Кэмпбелл Р., Макконелл К., Брю С.* Экономика: Принципы, проблемы и политика. Т. 1. Таллин, 1993.
52. *Ланкастер К.* Математическая экономика. М.: Советское радио, 1972.
53. *Леонтьев В.* Экономические эссе. М.: Политиздат, 1990.
54. *Леонтьев В. и др.* Исследования структуры американской экономики. М.: Госстатиздат, 1958.
55. *Линдерт П. Х.* Экономика мирохозяйственных связей. М.: Прогресс, 1992.
56. *Маленко Э.* Лекции по микроэкономическому анализу. М., 1985.
57. *Матросов В. М.* Метод векторных функций Ляпунова в анализе сложных систем с распределенными параметрами // Автоматика и телемеханика. 1973. № 1. С. 5–22.
58. *Медницкий В. Г., Медницкий Ю. В., Колбанов В. М., Королев В. Г.* Формы динамического равновесия замкнутой экономики // Экономика и математические методы. 1998. № 2.
59. *Менделеев Д. И.* Приемы естествознания в изучении цен // Полн. собр. соч.: В 25 т. Л.; М.: Изд-во АН СССР, 1937–1954. Т. 21. С. 33–42.
60. *Менделеев Д. И.* Толковый тариф или исследование о развитии промышленности России в связи с ее общим таможенным тарифом // Полн. собр. соч.: В 25 т. М.–Л.: Изд-во АН СССР, 1937–1954. Т. 19. С. 23–937.
61. Методы построения и использования макроэкономических и отраслевых производственных функций. М., 1979.
62. Микроэкономика / Под ред. Е. Б. Яковлевой. М.–СПб.: Поиск, 1998.
63. *Милованов В. П.* Об одном подходе к моделированию механизмов ценообразования // Экономика и математические методы. 1994. Т. 30. Вып. 1. С. 137–147.

64. Митягин Б. С. Заметки по математической экономике // Успехи математических наук. 1972. № 3.
65. Моделирование народнохозяйственных процессов / Под ред. В. С. Дадаева. М.: Экономика, 1973.
66. Моришима М. Равновесие, устойчивость, рост. М.: Наука, 1972.
67. Москаленко А. И. Оптимальное управление моделями экономической динамики. Новосибирск: Наука, 1999.
68. Немчинов В. С. Экономика и математика. М., 1965.
69. Никайдо Х. Выпуклые структуры и математическая экономика. М.: Мир, 1972.
70. Ортега Дж., Пул У. Введение в численные методы решения дифференциальных уравнений. М.: Наука, 1986.
71. Основы теории оптимального управления / Под редакцией В. Ф. Кротова. М.: Высшая школа, 1990.
72. Петросян Л. А., Захаров В. В. Математические модели в экологии. СПб.: Изд-во СПбГУ, 1997.
73. Поманский А. Б., Трофимов Г. Ю. Математические модели в теории экономического цикла // Экономика и математические методы. 1989. Т. 25. Вып. 5. С. 825–840.
74. Прасолов А. В. Аналитические и численные методы исследования динамических процессов. СПб.: Изд-во СПбГУ, 1995.
75. Прасолов А. В. Математические модели динамики в экономике. СПб.: СПбГУЭФ, 2000.
76. Прасолов А. В. Математические модели управления. СПб.: Изд-во ЛГУ, 1991.
77. Прасолов А. В., Степанов А. В. Применение метода локального моделирования к экономическим задачам. СПб.: Изд-во СПбГУ, 1995.
78. Прасолов А. В., Чуркин А. В. Производственные функции и их использование в теории фирмы. СПб.: Изд-во СПбГУ, 1998.
79. Разумихин Б. С. Об устойчивости систем с запаздыванием // Прикл. мат. и мех. 1956. Т. 20. Вып. 4. С. 500–512.
80. Рикардо Д. Начала политической экономии и налогового обложения // Собр. соч.: В 5 т. Т. 1. М.: Госполитиздат, Соцэкгиз. 1955–1961.
81. Роговский Е. А. О макроэкономических производственных функциях. М., 1984.
82. Самуэльсон П. Экономика: В 2 т. М.: ВНИИСИ, 1994.
83. Свиричев Ю. М., Логофет Д. О. Устойчивость биологических сообществ. М.: Наука, 1978. 352 с.
84. Свириденко К. С., Фадеев В. А. Вычисление дифференциальных характеристик двухфакторных производственных зависимостей. М., 1988.
85. Современная экономическая мысль / Под ред. С. Вайнтрауба. М.: Прогресс, 1981.
86. Современные буржуазные теории экономического роста и цикла / Под ред. Р. М. Энтова и др. М.: Наука, 1979.
87. Солоу Р. Перспективы теории роста // Мировая экономика и международные отношения. 1996. № 8.

88. *Степанов А. В.* Математические методы и алгоритмы численного моделирования динамических процессов: Дис. канд. физ.-мат. наук. СПбГУ, 1996.
89. *Столерю Л.* Равновесие и экономический рост. М.: Статистика, 1974.
90. *Гинбэрхэн Я., Бос Х.* Математические модели экономического роста. М.: Прогресс, 1967.
91. *Ткаченко Д. И.* Об условиях роста производства в одной модели с запаздыванием // Процессы управления и устойчивость: Труды XXXI научной конференции факультета ПМ-ПУ СПбГУ. СПб.: НИИ Химии СПбГУ, 2000.
92. *То, Кам Ту.* О некоторых методах построения производственных функций. М., 1988.
93. *Ферми Э.* Термодинамика. Харьков: Изд-во ХаГУ, 1969. 139 с.
94. *Харитонов В. Л.* К определению максимально допустимого запаздывания в задачах стабилизации // Дифференциальные уравнения. 1982. Т. 18. № 4. С. 723–724.
95. *Цыпкин Я. З.* Устойчивость систем с запаздывающей обратной связью // Автоматика и телемеханика. 1946. Т. 7. № 2, 3. С. 107–128.
96. *Чижова О. Н.* О нелокальной продолжимости систем с неограниченным запаздыванием // Вопросы механики и процессов управления. СПб.: СПбГУ. 2000. Вып. 18. С. 5.
97. *Шапиро Л. Д. и др.* Экономико-математическое моделирование. Томск, 1987.
98. *Шиманов С. Н.* О неустойчивости движения систем с запаздыванием по времени // Прикл. мат. и мех. 1960. Т. 24. № 1. С. 55–63.
99. *Щеголев И. Р.* Исследование некоторых моделей экономического роста // Кибернетику — на службу коммунизму. М.: Энергия, 1971. Т. 6.
100. *Эльсгольц Л. Э.* Устойчивость решений дифференциально-разностных уравнений // Успехи мат. наук. 1954. Т. 9. Вып. 4. С. 95–112.
101. *Эрроусмит Д., Плейс К.* Обыкновенные дифференциальные уравнения: Качественная теория с приложениями. М.: Мир, 1986, 243 с.
102. *Bellman R., Cooke K.* Differential-Difference Equations. N. Y.: Academic Press, 1963.
103. *Beltrami E.* Mathematical models for Society and Biology. N. Y.: Academic Press, 2002.
104. *Carter R. A. L., Zellner A.* The ARAR Error Model for Univariate Time Series and Distributed Lag // Studies in Nonlinear Dynamics & Econometrics. 2004. 8(1), article 2.
105. *Dixit A., Norman V.* International Trade Theory. Cambridge: Cambridge University Press, 1980. 339 p.
106. *Domar E. D.* Essays in the theory of economic growth. N. Y., 1957.
107. *Goodwin R. M.* The nonlinear accelerator and the persistence of business cycles // Econometrica 19. 1951.

108. *Huang J. Z., Yang L.* Identification of Non-Linear Additive Autoregressive Models // *J. Royal Statistical Society, (Series B)*, 66. 2004. Part 2. Pp. 463–477.
109. *Ivanov V., Kilian L.* A Practitioner's Guide to Lag Order Selection for VAR Impulse Response Analysis // *Studies in Nonlinear Dynamics & Econometrics*. 2005. 9(1), article 2.
110. *Junius T.* Remarks on models and methods in growth theory // *V symposium of operations research. Königstein. Ts*, 1981. Part II: sections 3–7.
111. *Kato J.* On Liapunov-Razumikhin type theorems. *Lect. Notes Math.* 1971. 243. Pp. 54–65.
112. *McDonalds N.* *Biological Delay Systems: Linear Stability Theory.* Cambridge: Cambridge University Press, 1988.
113. *McDonalds N.* *Lecture Notes in Biomathematics. Time lags in Biological Models.* Berlin: Springer-Verlag, 1978.
114. *Mikolajska Z.* Remarque sur la stabilite d'une solution du systeme d'eqs. diff. a parametre. *Collog. math.*, 18. 1967. Pp. 59–66.
115. *Ng S., Perron P.* Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power // *Econometrica*, 69. N. 6. 2001. P. 1519.
116. *Prasolov A. V., Wei K. C.* On Forecast of Exchange Rate of a Foreign Currency // *IEEE International Conference on Control Applications & IEEE International Symposium on Computer-Aided Control Systems Design*, September 25–27, 2000, Anchorage Hilton, Anchorage, Alaska, USA.
117. *Puu T.* *Nonlinear Economic Dynamics.* Berlin; Heidelberg: Springer-Verlag, 1989.
118. *Reichlin P.* Endogenous Cycles in Competitive Models: An Overview // *Studies in Nonlinear Dynamics & Econometrics*. 1996. 1(4), article 1.
119. *Sasiemi M. W.* Optimal advertising strategies // *Marketing Science*. 1989. Vol. 8. № 4. Pp. 358–372.
120. *Winker P.* Optimized Multivariate Lag Structure Selection: Using the Global Optimization Heuristic Threshold Accepting // *Computational Economics*, 16. 2000. Pp. 87–103.
121. *Zellner A.* My Experiences with Nonlinear Dynamic Model in Economics // *Studies in Nonlinear Dynamics & Econometrics*. 2002. 6(2), article 1.

УКАЗАТЕЛЬ ИМЕН

- А**
Акаике Х. 99
Александров А. Ю. 291
Аллен Р. 313, 314, 323, 325
- Б**
Байес Т. 99
Барбашин Е. А. 183
Бойль Р. 109
- В**
Вальрас Л. 9
Витте С. Ю. 145
Вольтерра В. 5, 6, 37, 133,
135–138, 140, 142, 150, 151,
157, 171, 185, 188, 190, 193,
194, 195, 196, 198, 235, 237,
238, 241, 242
Вэй К. 199
- Г**
Гамильтон У. 122
Гаусс К. 73
Гейл Д. 59, 60
Гей-Люссак Ж. 109
Гессе Л. 303
Грам И. 104
Гревилль 94
Гронуолл Т. 256
Гудвин Р. 49, 50, 140, 152,
154, 155, 156
Гурвиц А. 262
- Д**
Домар Е. 43–47, 49, 50, 53, 65,
325
Дуглас П. 48, 62, 321,
323–326, 329, 331
- Ж**
Жабко А. П. 291
- З**
Зубов В. И. 12, 115, 256, 267,
280, 293
- И**
Иванилов Ю. П. 300
- К**
Калецкий М. 49, 50, 154
Калман Р. 34, 121, 128
Кальдор Н. 45
Кардано Дж. 224, 229
Кассель Г. 39
Като Дж. 279
Кейнса Дж. М. 40–43, 45
Кели А. 122
Кемени Дж. 56
Кенэ Ф. 39, 60
Клапейрон Б. 108, 109
Кобб К. 48, 62, 321, 323,
324–326, 329, 331
Колмогоров А. Н. 134, 171
Кон Т. 292, 293
Кондратьев Н. Д. 150, 151
Коши О. 102, 136, 175,
251–253, 258, 259, 266, 267
Красовский Н. Н. 183, 266,
274, 275, 283
Купманс Т. 52
Кэсс Д. 52
- Л**
Лагранж Ж. 35, 36, 271, 305,
306
Лаплас П. 65, 258, 259

- Леонтьев В. В. 31, 40, 45, 46,
54, 59, 60, 61, 64, 65, 322,
325, 326
- Линдерт П. Х. 147
- Липшиц Р. 251, 252, 253, 257,
266, 269, 270, 271, 275, 283,
288, 289
- Лиувиль Ж. 90
- Лотка А. 5, 6, 133, 135, 138,
140, 142, 150, 151, 157, 171,
188, 190, 193–196, 198, 235,
237, 238, 241, 242
- Лотов А. В. 300
- Лукас Р. 54
- Ляпунов А. М. 10, 12, 152,
153, 158, 165, 166, 168, 183,
185, 189, 265–267, 269, 273–
275, 277–279, 287–292
- Мак-Фадден Д.** 314
- Мальтус Т. 134
- Мариотт Э. 109
- Марков А. А. 73
- Маркс К. 39, 54, 58, 60
- Маршалл А. 39
- Менделеев Д. И. 107–109, 145,
221, 222
- Миколайска З. 274, 275
- Михайлевский 313–315, 318,
324
- Моргенштерн О. 56
- Моришима М. 58
- Неймана Дж. фон** 40, 54–60, 63
- Никайдо Х. 62, 63, 64
- Олин Б.** 147, 221
- Парето В.** 318
- Пеано Дж. 255
- Перрон О. 56, 62
- Пикар Ш. 255
- Плейс К. 154
- по производству 25
- Поманский А. 154
- Понтрягин Л. С. 262, 264
- Разумихин Б. С.** 267, 271, 275,
287, 290, 292
- Рамсей Ф. 50, 52, 53
- Раус Е. 262
- Рикардо Д. 39, 147
- Ромера П. 53, 54
- Рыбчинский Т. 148
- Самуэльсон П.** 40, 57, 60, 147,
148, 154
- Сато М. 326, 327
- Свэн Т. 50
- Смит А. 39
- Сокова И. Н. 228
- Солоу Р. 50, 51, 53, 325, 331,
332
- Степанов А. В. 81
- Стилтьес Т. 256
- Столпер В. 147, 148
- Сэзини М. 231
- Тейлор В.** 155, 193
- Тинбергген Я. 329
- Ткаченко Д. И. 38
- Томпсон Дж. 56
- Трофимов Г. Ю. 154
- Фелпс Э.** 50, 51, 62
- Ферхюльст П. 134
- Фриш Р. 46
- Фробениус Ф. 56, 62
- Хансен Э.** 40
- Харитонов В. Л. 297
- Харрод Р. 40, 43–47, 49, 50,
53, 65, 325, 330
- Хекшер 147, 221
- Хикс Дж. 40, 154, 314, 329,
334
- Хотеллинг Х. 242
- Хоу Б. Л. 34, 121, 123, 126
- Четаев Н. Г.** 275
- Шарль Ж.** 109
- Шелл К. 53
- Шиманов С. Н. 275
- Шмидт Э. 104
- Шур 292, 293
- Эйлер Л.** 255
- Эльсгольд Л. Э. 266
- Эрроусмит Д. 154

ОГЛАВЛЕНИЕ

Предисловие	3
<i>Глава первая</i>	
Вводная	8
1.1. О применении математических инструментов в задачах экономической динамики	8
1.1.1. Определение неустойчивости в математике и его трансформация на экономическую динамику ...	10
1.1.2. Диссипация	13
1.1.3. Положения равновесия	15
1.1.4. Аттракторы	21
1.2. О математическом моделировании в экономической динамике	23
1.2.1. Некоторые понятия теории математического моделирования	23
1.2.2. Понятие адекватности математической модели ...	29
1.2.3. Способы составления математических моделей ...	30
1.2.4. Экстраполяция	34
1.3. Модели экономического роста	38
1.3.1. Историческая справка	38
1.3.2. Односекторные модели роста	41
1.3.2.1. Модель мультипликатора Кейнса	41
1.3.2.2. Модели Харрода–Домара	43
1.3.2.3. Модели циклического роста	49
1.3.2.4. Неоклассические модели роста	50
1.3.3. Модели неймановского типа	54
1.3.3.1. Стационарные траектории в модели Неймана	55
1.3.3.2. Магистральная теория	57
1.3.3.3. Модификации неймановской модели	58
1.3.4. Модели леонтьевского типа	60
1.3.4.1. Метод «затраты–выпуск»	60
1.3.4.2. Магистральный рост в леонтьевских моделях	62
1.3.4.3. Динамические леонтьевские модели	64

Глава вторая

Линейные динамические модели	66
2.1. Введение	66
2.2. Анализ временных рядов	66
2.2.1. Набор скалярных данных	67
2.2.2. Характеристики модели временного ряда	68
2.2.3. Сглаживание временных рядов	75
2.2.4. Об интервале идентификации	77
2.3. Многомерные временные ряды	81
2.3.1. Задача о форвардных операциях	81
2.3.2. Авторегрессии при малых объемах информации ...	85
2.3.3. Авторегрессии при больших объемах данных ...	92
2.3.4. Идентификация авторегрессионной модели с запаздыванием	99
2.3.5. Обсуждение критерия оптимальности	104
2.3.6. Модель равновесной цены по Менделееву	107
2.4. Модели управляемых процессов	114
2.4.1. Задача распределения инвестиций	115
2.4.2. Восстановление разностных линейных систем. Специальный вход	121
2.4.3. Восстановление разностных линейных систем. Произвольный вход	129
2.5. Заключение	131

Глава третья

Модели Лотки–Вольтерры в экономике	133
3.1. Основные идеи, методы и результаты	134
3.1.1. Описание модели	134
3.1.2. Цели исследования и техника анализа	136
3.2. Экономические интерпретации	138
3.2.1. Общие рассуждения	138
3.2.2. Конкуренция предприятий на общем рынке ...	141
3.2.3. Динамика иерархически связанных отраслей ...	144
3.2.4. Международная конкуренция	146
3.2.5. Кондратьевские волны	149
3.3. Последствие в экономических моделях	151
3.3.1. Модель динамики цены в окрестности равновесной точки	152
3.3.2. Макроэкономические модели с запаздыванием	154
3.3.3. Изменение качественного поведения модели при увеличении запаздывания	157
3.4. Полный анализ взаимодействия экономических агентов	160
3.4.1. Стационарные точки модели	161
3.4.2. Конкурентная деятельность без запаздывания ...	164
3.4.3. Конкурентная деятельность с временным лагом	171
3.4.4. Взаимозависимая деятельность с последствием	182
3.4.5. Произвольное число участников рынка	188

3.5. Проблема идентификации системы	190
3.5.1. Идентификация модели без временного лага и без искажения информации	190
3.5.2. Идентификация модели без запаздывания и с ошибками в информации	193
3.5.3. Идентификация модели с запаздыванием	195
3.6. Заключение	196

Глава четвертая

Различные динамические модели	198
4.1. Прогнозирование курсов валют	198
4.1.1. Формулировка проблемы	199
4.1.2. Критерии качества методов прогноза	202
4.1.3. Обсуждение методов прогнозирования	204
4.1.3.1. Непосредственная аппроксимация временных рядов	204
4.1.3.2. Метод общей корреляции	208
4.1.4. Оптимизация длины интервала	215
4.1.5. Основные выводы из решения задачи	219
4.2. Задачи международной торговли	220
4.2.1. Некоторые формализации международной торговли	220
4.2.2. Об одном из возможных подходов к анализу протекционизма	221
4.3. Модель влияния рекламы на продажу нового товара	230
4.4. Задача конкуренции на автомобильном рынке США	235
4.4.1. Формулировка задачи	235
4.4.2. Вывод уравнений	238
4.4.3. Идентификация и числительное приложение	241

Приложение 1

Теория дифференциальных систем с последствием	250
Классификация, основные определения, теоремы существования и единственности для уравнений с последствием	250
Линейная стационарная система с последствием	256
Свойства корней характеристического уравнения	260
Прямой метод Ляпунова для систем с последствием	265
Исследование устойчивости в случае динамического процесса с дискретным временем	287
Задача стабилизации линейных систем с последствием	293

Приложение 2

Об оптимальном размере производства	300
--	------------

Приложение 3

Производственная функция	308
Введение	308
Понятие производственной функции	309

Определения производственной функции	309
Об ограничениях производственных функций	310
Математические характеристики производственной функции и их экономическая интерпретация	311
Подходы к построению производственных функций	315
Метод множеств достижимости	315
Построение теоретических производственных функций на основе ММД	316
Построение производственных функций с постоянной эластичностью замещения факторов	318
Конкретные виды производственных функций, их особенности	322
НТП в производственных функциях	328
Отражение технического прогресса в ПФ	328
Экзогенный научно-технический прогресс	329
Экзогенный и эндогенный НТП	332
Примеры макроэкономических и микроэкономических моделей	332
Экзогенный ТП в производственной функции	332
Уравнение капиталовложений в основные фонды	333
Макроэкономическая модель с экзогенным ТП	333
Уравнение экономического роста	334
Построение производственных функций в условиях конкретной действительности	335
Библиографический список	338
Указатель имен	344

Александр Витальевич ПРАСОЛОВ

**МАТЕМАТИЧЕСКИЕ
МЕТОДЫ
ЭКОНОМИЧЕСКОЙ
ДИНАМИКИ**

Учебное пособие

Издание второе, исправленное

Заведующая редакцией
физико-математической литературы *Н. Р. Нигмадзянова*
Выпускающие *Н. А. Крылова, Е. П. Королькова*

ЛР № 065466 от 21.10.97

Гигиенический сертификат 78.01.07.953.П.007216.04.07
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»

lan@lanbook.ru; www.lanbook.com
192029, Санкт-Петербург, Общественный пер., 5.
Тел./факс: (812) 412-29-35, 412-05-97, 412-92-72.
Бесплатный звонок по России: 8-800-700-40-71

Подписано в печать 27.04.15.
Бумага офсетная. Гарнитура Школьная. Формат 84×108^{1/32}.
Печать офсетная. Усл. п. л. 18,48. Тираж 500 экз.

Заказ № .

Отпечатано в полном соответствии с качеством
предоставленного оригинал-макета.
в ПАО «Т8 Издательские Технологии».
109316, г. Москва, Волгоградский пр., д. 42, к. 5.