РАСПРЕДЕЛЕННЫЕ ЛАГИ

(Обзор)

Ф. Дж. ДРАЙМС

(США)

1. ВВЕДЕНИЕ И ПРЕДВАРИТЕЛЬНЫЕ ПОНЯТИЯ

Проблема распределенных лагов в экономике возникает из следующих соображений. Пусть рассматривается некоторый экономический процесс y_t , который в качестве аргумента (систематического или нестохастического) имеет некоторую экономическую переменную x_t , влияние которой на y_t , однако, не проявляется немедленно, а растянуто или распределено на много (возможно бесконечно много) промежутков времени. Формально это можно записать в виде модели

$$y_t = \sum_{i=0}^{\infty} w_i x_{t-i} + u_t, \quad (1)$$

где y_t — значение интересующего нас экономического процесса в момент t , а x_t — экономическая переменная, являющаяся аргументом этого процесса. Величины w_i , $i = 0, 1, 2, \dots$, — фиксированные, но неизвестные постоянные, а u_t — определенная подходящим образом случайная переменная. Теперь легко видеть, что изменение x_t не вызывает немедленного и полного изменения соответствующего значения y_t . Рассмотрим, например, последовательность

$$x_t = \begin{cases} 0, & t = 0, -1, -2, \dots, \\ 1, & t = 1, 2, 3, \dots \end{cases}$$

Пусть условные средние y заданы

$$E(y_t | x_{t-i}, i = 0, 1, 2, \dots) = \begin{cases} 0 & \text{при } t = 0, -1, -2, \dots, \\ \sum_{i=0}^{t-1} w_i & \text{при } t = 1, 2, \dots \end{cases} \quad (2)$$

Таким образом, несмотря на непрерывное изменение аргумента, воздействие этих изменений на зависимую переменную распределяется на бесконечный период, причем первоначальное воздействие этого изменения будет равно w_0 , кумулятивное воздействие после двух периодов $w_0 + w_1$, после трех $w_0 + w_1 + w_2$ и т. д. Теперь полезно ввести

Определение 1. Модель (1) называется *общей моделью бесконечно распределенного лага*. Последовательность $\{w_i : i = 0, 1, 2, \dots\}$ называется *общей структурой бесконечного лага*. Элементы w_i этой последовательности называются *коэффициентами лага*.

Различные специальные разновидности модели (1) применяются на практике. Возможно первое из таких приложений было сделано И. Фишером [1], который использовал довольно грубую схему

$$\begin{aligned} w_i &= 0, & i > n, \\ w_i &= a(n+1-i), & i = 0, 1, 2, \dots, n. \end{aligned}$$

Независимо от Фишера, Элмон в [2] предложил весьма похожую схему

$$\begin{aligned} w_\tau &= 0, & \tau > n, \\ w_\tau &= P^*(t), & \tau = 0, 1, 2, \dots, n, \end{aligned}$$

где

$$P^*(t) = \sum_{j=0}^k b_j s_j(t), \quad (3)$$

причем $s_j(t)$ представляют собой подходящим образом определенные полиномы Лагранжа. Элмон использовал эту модель для описания процесса превращения ассигнований в затраты. Благодаря этому фирмы или правительственные учреждения могли принимать решения относительно определенного уровня бюджетных ассигнований для каждого вида затрат. Обычно власти осуществляли бюджетные ассигнования лишь по прошествии нескольких периодов времени (лет). В этом случае имеются ясные причины институционального характера брать коэффициенты лага равными нулю при $\tau > n$. Можно показать, что введение полиномов Лагранжа в модели Элмона совершенно излишне, и мы можем дать лучшее определение этого класса модели распределенного лага с помощью (1) при условиях

$$w_\tau = 0, \tau > n, \quad w_\tau = P(\tau), \quad P(t) = \sum_{i=0}^k \beta_i t^i. \quad (4)$$

Следовательно, мы имеем

Определение 2. Модель (1) с ограничениями (4) называется *моделью полиномиального лага* порядка n и степени k .

Эта модель нашла обширную область использования в недавних прикладных работах. В частности, она была применена в уортонской модели [3] и модели MPS [4]*. Обе они — модели экономики США. Подобным же образом предполагается использовать модель полиномиального лага в проекте LINK, который по существу является глобальной моделью торговых потоков.

Несомненно, наиболее широкое практическое применение моделей распределенного лага началось с предположений, впервые сформулированных голландским эконометристом Л. М. Койком [5]. Его задачей было «объяснить» инвестиции в основной капитал. Он сделал предположение, что после некоторого промежутка времени, скажем τ_0 , можно определить поведение коэффициентов лага так

$$w_\tau = a\lambda^\tau, \quad \tau = \tau_0, \tau_0 + 1, \dots \quad (5)$$

Если для простоты записи взять $\tau_0 = 0$, то модель (1) может быть переписана в виде

$$y_t = a \sum_{i=0}^{\infty} \lambda^i x_{t-i} + u_t. \quad (6)$$

Это приводит к следующему.

* Прежде была известна как MIT—FRB эконометрическая модель.

Определение 3. Модель (6) называется *моделью геометрически распределенного лага*.

После первых работ Койка было предложено множество остроумных усовершенствований модели геометрического лага. Среди них выделяются работы Кагана [6] и Нерлова [7].

Обсуждение взаимосвязей этих усовершенствований и общее описание таких моделей будет значительно проще, если ввести понятие оператора лага. Эти вопросы будут рассмотрены здесь лишь несколько поверхностно. Подробно с ними можно познакомиться в работе автора [8, гл. 2] или Грилихеса [9].

Определение 4. Пусть X — множество всех функций $\{x: N \rightarrow R\}$, где N — множество $\{0, \pm 1, \pm 2, \dots\}$, а R — одномерное евклидово пространство. *Оператор лага* L определяется выражением

$$Lx(t) = x(t - 1) \text{ для } \forall x \in X.$$

Степени оператора лага определяются итерациями

$$L^2x(t) = L[Lx(t)] = Lx(t - 1) = x(t - 2). \tag{7}$$

В общей форме

$$L^kx(t) = x(t - k), \quad k = 1, 2, \dots, \tag{8}$$

и, кроме того, по соглашению

$$L^0 = I, \quad Ix(t) = x(t), \quad \forall x \in X. \tag{9}$$

Оператор I называется *единичным оператором*. Из (7) и (8) легко видеть, что

$$L^{k+s} = L^{s+k}, \quad (L^k)^s = L^{ks}. \tag{10}$$

Кроме того, можно показать [8, гл. 2], что множество всех конечных линейных комбинаций элементов множества $\{I, L, L^2, \dots\}$ над полем комплексных чисел образуют алгебру, которая *изоморфна алгебре полиномов* от комплексных чисел.

Важность такой трактовки в том, что при операциях с полиномами (или степенными рядами) от L мы можем обращаться с последними, как с комплексными величинами. Удобство такого подхода станет очевидным из дальнейшего обсуждения.

В только что введенных обозначениях можно переписать общую модель бесконечно распределенного лага (1) в виде

$$y_t = W(L)x_t + u_t, \quad W(L) = \sum_{i=0}^{\infty} w_i L^i. \tag{11}$$

Определение 5. Величину $W(L)$ из (11) называют *оператором общего бесконечного лага*.

Введенное Койком и записанное в форме (6) определение можно представить в виде

$$y_t = \frac{\alpha I}{I - \lambda L} x_t + u_t \tag{12}$$

при условии, что $|\lambda| < 1$. Последнее условие неявно содержится в (6).

Вклад работ Кагана (см. [6]) связан с использованием адаптивных ожиданий, которое существенно в (12). Предположим, что поведение экономических агентов определяется выражением

$$y_t = \alpha r_t^* + u_t. \tag{13}$$

В (13) y — некоторая наблюдаемая экономическая переменная, а p^* — ожидаемая величина. В модели адаптивного ожидания предполагается, что экономические агенты корректируют свои ожидания линейно, в соответствии с самыми последними наблюдениями отклонений ожиданий от их реализации. В частности, если p — наблюдаемая величина переменной, то предполагаемое состояние будет

$$p_t^* - p_{t-1}^* = \beta(p_{t-1} - p_{t-1}^*), \quad \beta \in [0, 1). \quad (14)$$

Решая (14) путем использования методов оператора лага, имеем

$$p_t^* = \frac{(1 - \gamma)I}{I - \gamma L} p_{t-1}, \quad \gamma = 1 - \beta. \quad (15)$$

Подставляя это выражение в (13), найдем

$$y_t = \frac{\alpha(1 - \gamma)I}{I - \gamma L} p_{t-1} + u_t, \quad (16)$$

а это выражение представляет собой модель распределенного геометрического лага.

Формулировка Нерлова столь же остроумна. В его схеме рассуждений принимается, что экономические агенты, получая информацию из внешней среды, определяют оптимальное значение некоторой экономической переменной, скажем y_t^* . В то же время, они поддерживают определенное значение этой величины, скажем y_{t-1} . Изменения, которые они вносят в него, пропорциональны разности между действительными и желаемыми значениями. Таким образом, предполагается, что

$$y_t - y_{t-1} = \alpha(y_t^* - y_{t-1}) + u_t, \quad \alpha \in [0, 1). \quad (17)$$

Причем, это соотношение справедливо лишь для $y_t^* - y_{t-1} \geq 0$. Например, если y_t^* желательный объем капитала и если $y_t^* - y_{t-1} < 0$, то его снижение может происходить лишь по мере амортизации в предположении, что нет рынков на уже бывшие в употреблении элементы основного капитала. В действительности это предположение может быть не слишком реалистично для отдельных фирм, хотя совершенно резонно для экономики в целом. Во всяком случае, решая (17), получаем

$$y_t = \frac{\delta I}{I - \delta L} y_t^* + \frac{I}{I - \delta L} u_t, \quad \delta = 1 - \alpha. \quad (18)$$

Но y_t^* не есть величина наблюдаемая. Мы можем, однако, постулировать, что оптимизация происходит соответственно выражению

$$y_t^* = \alpha x_t, \quad (19)$$

где x_t — наблюдаемая экономическая величина. Следовательно,

$$y_t = \frac{\alpha \alpha I}{I - \delta L} x_t + \frac{I}{I - \delta L} u_t. \quad (20)$$

А это (за исключением случайной компоненты) совпадает с моделью геометрически распределенного лага, представленной выражением (12). Достоинства формулировки Нерлова станут вполне очевидны, когда мы рассмотрим проблему оценки в моделях такого рода.

Определенным недостатком структуры геометрического лага является то, что последовательность коэффициентов лага монотонно убывает, т. е. изменение аргумента в большей части сказывается немедленно. В после-

дующие моменты коэффициенты лага убывают по величине. Во многих экономических задачах, однако, есть основания полагать, что коэффициенты лага сначала растут, достигают пика и лишь затем начинают убывать. Поведение такого рода не может быть описано с помощью структуры геометрического лага. Это отметил Солоу [10], предложивший *поскальку структуру лага*, которая определяется так

$$w_i = (1 - \lambda)^r \binom{r+i-1}{i} \lambda^i, \quad i = 0, 1, 2, \dots, \quad (21)$$

$$\lambda \in [0, 1), \quad r \in (0, \infty).$$

Этому описанию структуры лага соответствует оператор

$$W(L) = \frac{(1 - \lambda)^r I}{(I - \lambda L)^r}. \quad (22)$$

В немногих практических приложениях этой структуры лага *величина r обычно берется целой*. Нетрудно показать, что выписанный выше оператор может быть получен последовательным применением r операторов геометрического лага, причем все они имеют один и тот же параметр λ . Более удобно структуру лага (21) называть *отрицательной биномиальной структурой лага*.

Последнее достижение в развитии идей распределенного лага (из опубликованных в эконометрической литературе) — работа Юргенсона [11]. Он предложил следующий рациональный оператор лага

$$W(L) = \frac{A(L)}{B(L)}, \quad (23)$$

где

$$A(L) = \sum_{i=0}^n a_i L^i \quad \text{и} \quad B(L) = \sum_{j=0}^n b_j L^j, \quad m \leq n, \quad (24)$$

причем полиномы A и B не имеют общих членов. Хотя Юргенсон пытался показать, что рациональный лаг есть приближение структуры бесконечного лага (см. формулу (1)), привлекая для этого теорему о приближении Вейерштрасса, доказательство это нельзя считать удовлетворительным. Этот вопрос обсуждался в работах автора [8, гл. 3] и Симса [12]. Спецификацию лага (23) — (24) лучше рассматривать применительно к изучению какого-либо определенного явления в экономике, а не как аппроксимацию совершенно общей, основной структуры лага.

В заключение следует отметить, что подобное «ненадежное» развитие теории распределенного лага наблюдается и в инженерной литературе. В частности, рационально распределенный лаг привлекает внимание многих, см. [13] и недавнюю обзорную статью Эстрома и Эйхоффа [14]).

2. ХАРАКТЕРИЗАЦИЯ СТРУКТУРЫ ЛАГА

Часто бывает полезно дать общую характеристику распределения лага. В частности, мы можем интересоваться выяснением того, влияет ли аргумент на зависимую переменную «концентрированно» или «рассеянно», или определением среднего периода лага, на протяжении которого такое влияние заметно. Для определенного типа лага подобную характеристику вполне можно дать. Для этого заметим, что из (21) имеем

$$\lim_{i \rightarrow \infty} E(y_t | x_{t-i}, i = 0, 1, 2, \dots) = \sum_{i=0}^{\infty} w_i = \omega. \quad (25)$$

Ясно, что в экономических задачах естественным является условие $|\omega| < \infty$. Часто требуется также, чтобы коэффициенты лага были знакопостоянны. Когда выполнены оба эти условия, можно определить

$$w_i^* = \frac{w_i}{\omega}, \quad i = 0, 1, 2, \dots \quad (26)$$

Заметим, кроме того, что

$$w_i^* \in [0, 1), \quad \sum_{i=0}^{\infty} w_i^* = 1. \quad (27)$$

И мы имеем

Определение 6. Если структура лага $\{w_i: i = 0, 1, 2, \dots\}$ обладает следующими свойствами $w_i w_j \geq 0$ для всех i и j и $\sum_{i=0}^{\infty} w_i = \omega < \infty$, то она называется нормализуемой структурой лага.

Замечание 1. Нормализуемая структура лага изоморфна функциям вероятностной меры. Многие из ранних работ, посвященных распределенному лагу, основаны на этом изоморфизме. Напомним, что если y — геометрическая (дискретная) случайная переменная, то функция $x = y - 1$ задается выражением

$$Pr\{x = i\} = (1 - \lambda)\lambda^i, \quad i = 0, 1, 2, \dots,$$

причем ординаты этого выражения дают структуру геометрического лага, введенного Койком [5].

Пусть $\{x_i: i = 1, 2, \dots, r\}$ — независимые случайные переменные с одинаковыми независимыми функциями распределения (о.н.р.) с приведенной выше функцией меры. Тогда

$$x = \sum_{i=1}^r x_i$$

имеет функцию распределения меры

$$Pr\{x = i\} = (1 - \lambda)^r \binom{r+i-1}{i} \lambda^i, \quad i = 0, 1, 2, \dots$$

но ординаты этой функции даются паскалевой структурой лага Солоу [10]. Используя отмеченный выше изоморфизм, мы можем определить различные характеристики центральной тенденции или дисперсии точно так же, как для случайных переменных.

Определение 7. Пусть $\{w_i: i = 0, 1, 2, \dots\}$ — структура лага и t — некоторое действительное число. Тогда

$$W(t) = \sum_{i=0}^{\infty} w_i t^i$$

называется производящей функцией лага (точнее — структуры лага).

Если лаг нормализуем, то путем соответствующего переопределения коэффициентов лага (если это необходимо) замечаем, что

$$W'(1) = \sum_{i=0}^{\infty} i w_i = m(w). \quad (28)$$

Определение 8. Если $\{w_i: i = 0, 1, 2, \dots\}$ — нормализованная структура лага, т. е. если $w_i \in [0, 1]$, $\sum_{i=0}^{\infty} w_i = 1$, то величина (28) будет средним лагом, а величина

$$W''(1) - [W'(1)]^2 = V(w) \tag{29}$$

называется дисперсией лага.

Определенные выше величины аналогичны среднему и дисперсии дискретных случайных величин и имеют ту же интерпретацию. Заметим, однако, что такие величины (хотя и могут быть определены) бессмысленны, когда структура лага ненормализуема.

Пример 1. Для геометрической структуры лага имеем

$$W(t) = (1 - \lambda) \sum_{i=0}^{\infty} (\lambda t)^i = \frac{1 - \lambda}{1 - \lambda t}, \quad m(w) = \frac{\lambda}{1 - \lambda}, \quad V(w) = \frac{\lambda^2}{(1 - \lambda)^2}. \tag{30}$$

Мы видим, что и средний лаг $m(w)$ и дисперсия лага $V(w)$ — возрастающие функции λ .

Пример 2. Для отрицательного биномиального (Паскалева) распределения лага имеем

$$W(t) = \frac{(1 - \lambda)^r}{(1 - \lambda t)^r}, \quad m(w) = \frac{\lambda r}{1 - \lambda}, \quad V(w) = \frac{\lambda r^2}{(1 - \lambda)^2}. \tag{31}$$

Здесь средний лаг и вариабельность лага — возрастающие функции λ и r .

Пример 3. Положим, что общая рациональная структура лага нормализуема и уже нормализована. Тогда

$$W(t) = \frac{A(t)}{B(t)}, \quad m(w) = \frac{A'(1)}{A(1)} - \frac{B'(1)}{B(1)},$$

$$V(w) = \frac{A''(1)}{A(1)} - \left(\frac{A'(1)}{A(1)}\right)^2 - \frac{B''(1)}{B(1)} + \left(\frac{B'(1)}{B(1)}\right)^2. \tag{32}$$

3. ОЦЕНКА РАСПРЕДЕЛЕНИЯ ЛАГА

а) **Полиномиальный лаг.** Начнем рассмотрение класса структур полиномиального лага, так как проблемы оценки в этом случае особенно просты. Задача формулируется следующим образом. Дана модель

$$y_t = \sum_{\tau=0}^n w_{\tau} x_{t-\tau} + u_t, \quad t = 1, 2, \dots, T, \tag{33}$$

причем n известно и выполняются условия

$$w_{\tau} = P(\tau), \quad P(t) = \sum_{i=0}^k \beta_i t^i, \quad k < n. \tag{34}$$

Найти эффективные способы оценки коэффициентов лага.

Решение этой проблемы известно при обычных предположениях относительно последовательности ошибок $\{u_t: t = 1, 2, \dots\}$, а именно, если последняя представляет собой последовательность независимых и одинаково распределенных случайных переменных с пулевой средней и дисперсией σ^2 , а последовательность u независима от последовательности x , то задача решается следующим образом. Используя (34) и (33) замечаем,

что

$$y_t = \sum_{i=0}^k \beta_i z_{ti} + u_t, \quad z_{ti} = \sum_{i=0}^n \tau^i x_{t-i}, \quad t = n+1, n+2, \dots, T, \quad (35)$$

тогда теорема Гаусса — Маркова показывает, что

$$\hat{\beta} = (Z'Z)^{-1}Z'y, \quad Z = (z_{ti}), \quad t = n+1, n+2, \dots, T, \\ i = 0, 1, 2, \dots, k, \quad y = (y_{n+1}, y_{n+2}, \dots, y_T)', \quad \beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)' \quad (36)$$

— наилучшая линейная несмещенная оценка (н.л.н.о.) вектора коэффициентов β . Вывести н.л.н.о. для коэффициентов лага сравнительно легко, так как мы имеем соотношение

$$w = Q\beta, \quad Q = (\tau^i), \quad \tau = 0, 1, 2, \dots, n; \quad i = 0, 1, 2, \dots, k, \quad (37)$$

принимая, что $\tau^i = 1$ при $\tau = 0$ и $i = 0$. Из (37) видно, что

$$\hat{\omega} = Q\hat{\beta} \quad (38)$$

есть н.л.н.о. величины ω .

Теория оценок распределенного полиномиального лага довольно проста, она не намного труднее, чем теория оценок в обычных регрессионных моделях. В практических приложениях первостепенными являются два вопроса: во-первых, какова длина лага (т. е. каково n) и порядок многочлена (т. е. k) и, во-вторых, насколько спецификация (34) совместна с реальными данными?

По первому вопросу заметим, что во всех проведенных выше рассуждениях, мы исходили из того, что и n и k известны. При такой априорной информации имеется эффективная процедура оценки неизвестных параметров. Для приложений, однако, типично, что ни n , ни k неизвестны и должны быть оценены одновременно с коэффициентами лага. Мало исследований было проведено для решения этой проблемы. Стандартная процедура заключается в подборе достаточно узкого диапазона изменения n и k и выборе такого набора оценок этих параметров, который минимизирует, скажем, остаточную сумму квадратов. Можно показать, что если такие априорные ограничения на величины n и k сделаны верно, то описанная выше процедура даст минимум x — квадрат оценки. Второй из поставленных вопросов обсуждается в работе автора [8, гл. 8].

б) Геометрический лаг. Рассмотрим некоторые проблемы оценок, возникающие при геометрической структуре лага; обсуждение общего рационального лага будет значительно менее полным. Возьмем сначала модель

$$y_t = \frac{\alpha I}{I - \lambda L} x_t + u_t, \quad t = 1, 2, \dots, T, \quad (39)$$

где $\{u_t: t = 1, 2, \dots\}$ — последовательность о.н.р. случайных переменных с нулевым средним и конечными (абсолютными) моментами — по крайней мере до восьмого порядка. Допустим сначала, что экзогенная переменная представляет собой равномерно ограниченную последовательность фиксированных констант с определенными свойствами, которые будут описаны ниже.

Из (39) видно, что основная проблема здесь — как и почти во всей теории оценки распределенного лага — нелинейности. Проблемы, возникающие при рассмотрении модели

$$y_t = g_t(\theta) + u_t, \quad (40)$$

охватывают почти все трудности, с которыми мы столкнемся на протяжении этого раздела. В (40) $g_t(\theta)$ — вектор неизвестных параметров, а индекс символа функции — просто «дата» наблюдения. Обращаясь к (39), имеем

$$\theta = (\alpha, \lambda)', \quad g_t(\theta) = \alpha \sum_{i=0}^{\infty} \lambda^i x_{t-i}. \tag{41}$$

Оценки неизвестных параметров модели (40) могут быть получены как

$$\sup_{\theta \in \Theta} \sum_{t=1}^T [y_t - g_t(\theta)]^2,$$

где Θ — соответствующее пространство параметров. Общие условия, при которых полученные таким образом оценки соответствуют действительности, даны в работах Йенриха [15] и Маленво [16]. Йенрих получил также условия, при которых их предельное распределение является нормальным. Ранее эти проблемы рассматривались Вальдом [17] и Волфовицем [18]. Вычислительные алгоритмы для получения таких оценок были проверены Стиглицем и Мак-Брайдом [13] и автором [19, 8]. Моррисон в [20] на примере небольшой размерности сравнил их с процедурой Монте-Карло.

Минимум оценки χ — квадрат (MCS) или максимум правдоподобия (ML) — в случае, когда u — процесс предполагается нормальным — получается следующим образом. Определим

$$a_0 = \sum_{i=0}^{\infty} \lambda^i x_{-i}, \quad x_i^* = \sum_{i=0}^{t-1} \lambda^i x_{t-i} \tag{42}$$

и заметим, что можно записать

$$y_t = a_0 \lambda^t + \alpha x_t^* + u_t. \tag{43}$$

При заданном λ можно вычислить

$$\begin{pmatrix} \tilde{a}_0 \\ \tilde{\alpha} \end{pmatrix} = \begin{bmatrix} \lambda^2(1 - \lambda^{2T}) / (1 - \lambda^2) \sum \lambda^i x_i^* \\ \sum \lambda^i x_i^* \\ \sum x_i^{*2} \end{bmatrix}^{-1} \begin{bmatrix} \sum \lambda^i y_i \\ \sum x_i^* y_i \end{bmatrix}, \tag{44}$$

рассматривая a_0 как подлежащий оценке параметр (a_0 называется *параметром усечения*). Впервые a_0 было рассмотрено как параметр требующий оценки. Клейном [21]. Часто в экономических задачах накладывается ограничение $\lambda \in [0, 1)$. С незначительной потерей точности можно записать это ограничение как $\lambda \in [0, 1 - \delta]$, $\delta > 0$, а затем можно разделить этот интервал точками λ_i , $i = 0, 1, 2, \dots, r$, такими, что $\lambda_0 = 0$, а $\lambda_r = 1 - \delta$ и для каждого λ_i вычислить величины из (44). Набор трех чисел $\tilde{\lambda}$, $\tilde{\alpha}$, $\tilde{a}_0$, который дает наименьшую остаточную сумму квадратов и есть в зависимости от обстоятельств MCS- или ML-оценка. В этой связи можно заметить, что

a_0 не является состоятельной оценкой. Действительно, если $\lim_{T \rightarrow \infty} \left(\sum x_i^{*2} / T \right)$

— величина положительная, $\text{alim}_{T \rightarrow \infty} \left(\sum x_i^* / T \right)$ — ограничена, то матрица,

обратная к (44), асимптотически стремится к единичной. Таким образом, *параметр усечения может практически приниматься во внимание лишь для*

небольших задач. Результаты Моррисона [20] показывают, что для небольших задач учет a_0 дает заметное улучшение оценок для α и λ .

Алгоритм Стиглица — Мак-Брайда (SM) работает так. Для оценки MCS условия первого порядка имеют вид

$$\alpha \sum (y_i - \alpha x_i^*) \frac{\partial x_i^*}{\partial \lambda} = 0, \quad \sum (y_i - \alpha x_i^*) x_i^* = 0, \quad (45)$$

где

$$x_i^* = \frac{I}{I - \lambda L} x_i = \sum_{i=0}^{\infty} \lambda^i x_{i-i}, \quad \frac{\partial x_i^*}{\partial \lambda} = \frac{I}{I - \lambda L} x_{i-1}^* = x_{i-1}^{**}. \quad (46)$$

Из (46) получаем рекурсивные соотношения

$$x_i^* = \lambda x_{i-1}^* + x_i, \quad x_i^{**} = \lambda x_{i-1}^{**} + x_i^*. \quad (47)$$

Тогда при известной величине λ значения x_i^* и x_{i-1}^{**} можно легко вычислить, исходя из информации, которую дают нам заданные начальные условия x_0^* и x_0^{**} . В SM -алгоритме эти начальные значения берутся нулевыми. Это удобно. Можно показать, что такой выбор не оказывает никакого влияния на свойства оценки. Фактически, использование этой пары начальных условий эквивалентно использованию x_i^* способом, определенным в (42), и тому, что мы пренебрегаем параметром усечения a_0 .

Существенным аспектом SM -алгоритма является замена

$$y_i = \frac{I - \lambda L}{I - \lambda L} \lambda y_i = y_i^* - \lambda y_{i-1}^*, \quad y_i^* = \frac{I}{I - \lambda L} y_i.$$

Уравнения (45) записываются в виде

$$\begin{aligned} \alpha \sum x_i^{*2} + \lambda \sum y_{i-1}^* x_i^* &= \sum y_i^* x_i^*, \\ \alpha \sum x_i^* x_{i-1}^{**} + \lambda \sum y_{i-1}^* x_{i-1}^{**} &= \sum y_i^* x_{i-1}^{**}. \end{aligned} \quad (48)$$

Исходя из начальной оценки λ , могут быть вычислены величины, помеченные звездочками, и, таким образом, получено первое приближение $\left(\begin{smallmatrix} \bar{\alpha} \\ \bar{\lambda} \end{smallmatrix} \right)_1$. Пересчитывая значения помеченных звездочками величин и используя $\left(\begin{smallmatrix} \bar{\alpha} \\ \bar{\lambda} \end{smallmatrix} \right)_1$, можно получить второе приближение и т. д. пока процесс

не сойдется. К сожалению нет общей теоремы, определяющей условия сходимости x -процесса. Ясно, конечно, что если процесс сошелся, то последняя итерация соответствует решению уравнений (45). Надо отметить два факта. Первый — процедура получения оценки не гарантирует, что ограничения на величину λ будут удовлетворяться; второй — нет гарантии, что получаемое решение есть глобальный минимум минимизируемой функции. Оно вполне может оказаться просто одной из стационарных точек.

Асимптотические свойства ML - (или MCS)-оценок (в зависимости от обстоятельств) определены. При подходящих условиях на последовательности x можно показать, что асимптотически такие оценки подчиняются условию

$$\sqrt{T} \left[\begin{pmatrix} \hat{\alpha} \\ \hat{\lambda} \end{pmatrix} - \begin{pmatrix} \alpha \\ \lambda \end{pmatrix} \right] \sim N(0, \Phi), \quad (49)$$

где

$$\Phi = \sigma^2 \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \begin{bmatrix} \sum x_t^{**2} & \alpha \sum x_t^* x_{t-1}^{**} \\ \alpha \sum x_{t-1}^{**} x_t^* & \alpha^2 \sum x_{t-1}^{**2} \end{bmatrix} \right\}^{-1} \quad (50)$$

и $\begin{pmatrix} \hat{\alpha} \\ \hat{\lambda} \end{pmatrix}$ — оценки заключительной итерации.

Таким образом, если мы готовы положиться на асимптотическую теорию, проблема оценки полностью решена. Есть много вариантов этой основной модели. Наиболее очевидная возможность развития модели заключается во введении дополнительных независимых переменных, ни одна из которых не связана с параметрами, вызывающими нелинейность в схеме распределенного лага. В этом случае модель имеет вид

$$y_t = \frac{\alpha_0 I}{I - \lambda L} x_{t0} + \sum_{i=1}^n \alpha_i x_{ti} + u_t. \quad (51)$$

Проблемы оценки для этой модели вполне совпадают с только что рассмотренными. Несколько более запутанным представляется набор проблем, возникающих при рассмотрении модели

$$y_t = \frac{\alpha L}{I - \lambda L} x_t + \frac{I}{I - \rho L} \varepsilon_t, \quad (52)$$

где $\{\varepsilon_t: t = 0, \pm 1, \pm 2, \dots\}$ — последовательность о.н.р. случайных переменных с нулевым средним и дисперсией σ^2 . Получаем *MCS*-оценки путем минимизации

$$S = (y - \alpha x^*)' V^{-1} (y - \alpha x^*). \quad (53)$$

Здесь символы имеют тот же смысл, что и выше, а

$$\text{Cov}(u) = \sigma^2 V, \quad u = (u_1, u_2, \dots, u_T)', \quad u_t = \frac{I}{I - \rho L} \varepsilon_t, \quad (54)$$

$$V = (v_{ts}), \quad v_{ts} = \frac{\rho^{|t-s|}}{1 - \rho^2}, \quad t, s = 1, 2, \dots, T.$$

Очевидно, оценки здесь могут быть получены путем двойной поисковой процедуры в области допустимых изменений параметров λ и ρ . Итак, если мы положим $\theta = (\alpha, \lambda, \rho)'$ и $S_T(\theta)$ для минимизируемой в (53) величины, деленной на T , то получим оценку в виде оператора

$$\sup_{\theta \in \Theta} [S_T(\theta)].$$

Здесь оценки *ML* и *MCS* различаются на величину, которая прежде максимизировалась

$$L_T(\theta, y, x) = -\frac{T}{2} \ln 2\pi - \frac{T}{2} \ln \sigma^2 - \frac{T}{2\sigma^2} S_T(\theta) + \frac{1}{2} \ln(1 - \rho^2). \quad (55)$$

Частичной максимизацией по σ^2 и подстановкой в (55) находим

$$L_T(\theta, y, x) = -\frac{T}{2} [\ln(2\pi) + 1] - \frac{T}{2} \ln \left[\frac{S_T(\theta)}{(1 - \rho^2)^{1/T}} \right] \quad (56)$$

и видим, что ML -оценка получается с помощью оператора

$$\sup_{\theta \in \Theta} \frac{[S_T(\theta)]}{(1 - \rho^2)^{1/T}}.$$

Если допустить, что $\lambda \in [0, 1 - \delta_1]$, $\rho \in [-1 + \delta_2, 1 - \delta_2]$, $\delta_1, \delta_2 > 0$, то обычно это приводит к незначительному уменьшению практической приложимости, но значительно упрощает математическую аргументацию. Заметив это, мы видим, что для любого допустимого $\theta \lim_{T \rightarrow \infty} (1 - \rho^2)^{1/T} = 1$.

Следовательно, можно ожидать, что эти две оценки асимптотически эквивалентны, что фактически и имеет место.

Итак, асимптотические свойства этих оценок установлены. Мы не приводим здесь детальной аргументации, однако кратко упомянем о несколько ином доказательстве, которое было дано у автора [22] и [8, гл. 6]. Там показано, что рассматриваемые выше оценки сходятся к соответствующим параметрам *не только по вероятности, но и с вероятностью единица*. Итак, рассмотрим MCS -оценку, определяемую выражением

$$S_T(\hat{\theta}) \leq S_T(\theta), \quad \forall \theta \in \Theta.$$

Ясно, что α может быть исключено частичной минимизацией

$$\begin{aligned} S_T(\omega) &= \frac{1}{T} y' \left[V^{-1} - \frac{V^{-1} x^* x'^* V^{-1}}{x'^* V^{-1} x^*} \right] y = \\ &= \frac{1}{T} y' V^{-1} y \left[1 - \frac{(y' V^{-1} x^*)^2}{(x'^* V^{-1} x^*) (y' V^{-1} y)} \right], \end{aligned} \quad (57)$$

где $\omega = (\lambda, \rho)'$.

Значит, принимая $x_i^* \neq 0$ для всех t , мы видим, что $S_T(\omega)$ не равно нулю и ограничено сверху величиной $\frac{1}{T} y' V^{-1} y$; таким образом, любые асимптотические свойства, которые мы определим для оценки MCS , будут приложимы и для оценки ML . Для доказательства сходимости оценки MCS с вероятностью единица, поступим следующим образом. Сначала покажем, что для любого допустимого значения ω , $S_T(\omega)$ сходится к ее пределу, скажем $S(\omega)$, с вероятностью единица *равномерно по ω* . Сходимость с вероятностью единица означает следующее. Пусть ξ — пространство выборок, над которым определен процесс ошибок (u — процесс), и пусть s — подмножество ξ , над которым сходимости нет. Сходимость с вероятностью единица означает, что *мера подмножества s равна нулю*. Равномерность сходимости по ω означает, что предельное отклонение не зависит от того, какое ω (из множества допустимых) мы рассматриваем. Таким образом, можно сказать, что почти для всех последовательностей $\{\omega_t: t = 1, 2, \dots\}$ $S_T(\omega)$ сходится к $S(\omega)$ равномерно по $\omega \in \Omega$. Здесь Ω множество допустимых значений параметра ω .

Затем мы замечаем, что последовательность оценок $\{\hat{\omega}_T\}$ такая, что $S_T(\hat{\omega}_T) \leq S_T(\omega) \forall$ для всех $\omega \in \Omega$ — бесконечная ограниченная последовательность, а значит она имеет, по крайней мере, одну предельную точку, скажем ω_* . Наконец, мы показываем, что $\omega_* = \omega_0$, т. е. истинной величине параметра. Так как наше рассуждение справедливо почти для всех последовательностей $\{\omega_t: t = 1, 2, \dots\}$ и ω_* — любая предельная точка, мы заключаем, что $\hat{\omega}_T$ сходится к ω_0 с вероятностью единица при выполнении некоторых условий идентификации. Так как оценка α определяется выражением

$$\hat{\alpha} = (x'^* V^{-1} x^*)^{-1} x'^* V^{-1} y, \quad (58)$$

мы заключаем также, что $\hat{a}$ сходится к a_0 с вероятностью единица. Это доказательство можно легко распространить на случай, когда процесс ошибок является авторегрессивным процессом n -го порядка или когда модель содержит дополнительные независимые переменные, ни одна из которых не имеет бесконечно распределенного лага.

Теперь мы замечаем, что

$$\frac{x^*{}' V^{-1} x^*}{T} = \frac{1}{T} (1 - \rho^2) x_1^{*2} + \frac{1}{T} \sum_{t=2}^T (x_t^* - \rho x_{t-1}^*)^2 \quad (59)$$

и, вводя

$$h_t(\omega) = x_t^* - \rho x_{t-1}^* = \sum_{i=0}^{\infty} \lambda^i (x_{t-i} - \rho x_{t-1-i}), \quad (60)$$

имеем при условии, если последовательность x ограничена,

$$|h_t(\omega)| \leq \frac{2K}{\delta_1}, \quad (61)$$

где K — граница последовательности x . Кроме того, если $\omega_1, \omega_2 \in \Omega$, то

$$|h_t(\omega_2) - h_t(\omega_1)| \leq |\lambda_2 - \lambda_1| \frac{2K}{\delta_1^2} + |\rho_2 - \rho_1| \frac{K}{\delta_1}, \quad (62)$$

откуда видно, что $\{h_t\}$ — последовательность равномерно ограниченных непрерывных функций. Так как

$$|h_t^2(\omega_2) - h_t^2(\omega_1)| \leq |h_t(\omega_2) - h_t(\omega_1)| |h_t(\omega_2) + h_t(\omega_1)|, \quad (63)$$

то и $\{h_t^2\}$ — последовательность равномерно ограниченных непрерывных функций. Также, очевидно, справедливо

$$H_T(\omega) = \frac{1}{T} \sum_{t=2}^T h_t^2(\omega). \quad (64)$$

Но если последовательность x такова, что

$$c(i, j) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T x_{t-i} x_{t-j}$$

существуют и имеют определенные конечные величины, такие, что последовательность $\{H_T(\omega)\}$ сходится точно по ω , тогда по теореме Арзела — Асколи [23] она сходится по ω и равномерно. Далее замечаем, что

$$\frac{1}{T} u' V^{-1} x^* = \frac{1}{T} \left[\sum_{t=1}^T u_t x_t^* - \rho \sum_{t=1}^{T-1} u_t x_{t+1} - \rho \sum_{t=1}^{T-1} u_{t+1} x_t - \rho^2 \sum_{t=2}^{T-1} u_t x_t^* \right] \quad (65)$$

и, таким образом, если каждая из этих сумм равномерно сходится по λ , то равномерная сходимость по ρ очевидна. Но достаточно будет показать, что $\frac{1}{T} \sum u_t x_t^*$ сходится к пределу с вероятностью единица равномерно по λ .

С другими членами мы можем обращаться точно так же. Но

$$u_t = \rho_0^t u_0 + \sum_{j=0}^{t-1} \rho_0^j \varepsilon_{t-j}, \quad x_t^* = \lambda^t x_0^* + \sum_{i=0}^{t-1} \lambda^i x_{t-i}. \quad (66)$$

Следовательно,

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T x_t u_t &= \frac{1}{T} \left[u_0 x_0 \sum_{t=1}^T (\lambda \rho_0)^t + u_0 \sum_{t=1}^T \rho_0^t \sum_{i=0}^{t-1} \lambda^i x_{t-i} + \right. \\ &\left. + x_0 \sum_{t=1}^T \lambda^t \sum_{j=0}^{t-1} \rho_0^j \varepsilon_{t-j} + \sum_{t=1}^T \sum_{i=0}^{t-1} \sum_{j=0}^{t-1} \lambda^i \rho_0^j x_{t-i} \varepsilon_{t-j} \right], \end{aligned} \quad (67)$$

и будет достаточно показать, что последний член в правой части сходится к нулю равномерно по λ , а с другими членами поступить аналогично. Для этой цели заметим, что при изменении порядка суммирования получится

$$\frac{1}{T} \sum_{t=1}^T \sum_{i=0}^{t-1} \sum_{j=0}^{t-1} \lambda^i \rho_0^j x_{t-i} \varepsilon_{t-j} = \sum_{i=0}^{T-1} \sum_{j=0}^{T-1} \lambda^i \rho_0^j W_{T, i, j} \quad (68)$$

где

$$W_{T, i, j} = \frac{1}{T} \sum_{t=\max(i, j)+1}^T x_{t-i} \varepsilon_{t-j} \quad (69)$$

и, очевидно, достаточно будет показать, что

$$\limsup_{T \rightarrow \infty} \sup_{i, j \leq T} |W_{T, i, j}|$$

сходится к нулю. Используя неравенство Чебышева для моментов шестого порядка, имеем

$$Pr\{|W_{T, i, j}| > r\} < \frac{E(W_{T, i, j})^6}{r^6}, \quad (70)$$

но

$$E(W_{T, i, j})^6 \leq \frac{K_1}{T^4}, \quad (71)$$

где K_1 — некоторая граница, не зависящая от T , i или j . Теперь для любого $r > 0$

$$\sum_{T=1}^{\infty} \sum_{i=0}^{T-1} \sum_{j=0}^{T-1} Pr\{|W_{T, i, j}| > r\} \leq K_2 \sum_{T=1}^{\infty} \frac{1}{T^2} < \infty \quad (72)$$

и по лемме Бореля — Кантелли [24] заключаем, что

$$Pr\{W_{T, i, j} > r, \text{ i.o.}\} = 0. \quad (73)$$

Это означает, что $\sup_{i, j \geq T} |W_{T, i, j}|$ стремится к нулю с вероятностью единица. Таким образом, мы установили, что $\frac{1}{T} \sum x_t u_t$ сходится к нулю с вероятностью единица равномерно по λ и ρ_0 .

Наконец, рассмотрим член

$$\frac{1}{T} u' V^{-1} u = \frac{1}{T} (1 - \rho^2) u_1^2 + \frac{1}{T} \left[\sum_{t=2}^T u_t^2 - 2\rho \sum_{t=2}^T u_t u_{t-1} + \rho^2 \sum_{t=1}^T u_t^2 \right]. \quad (74)$$

Ясно, что если мы сможем показать, что $\frac{1}{T} \sum u_t^2$ и $\frac{1}{T} \sum u_t u_{t-1}$ приближаются к их пределам с вероятностью единица, то равномерность сходимости по ρ будет очевидна из (74). Заметим, что

$$E(u_t^2) = \frac{\sigma_0^2}{1 - \rho_0^2}, \quad E|u_t u_{t-1}| \leq \frac{\sigma_0^2}{1 - \rho_0^2}, \quad \forall t \tag{75}$$

и, кроме того, что $\{u_t^2 : t = 1, 2, \dots\}$ и $\{u_t u_{t-1} : t = 2, 3, \dots\}$ — строго стационарный процесс. Следовательно, по теореме Биркгоффа — Хинчина [25] величины $\frac{1}{T} \sum u_t^2$ и $\frac{1}{T} \sum u_t u_{t-1}$ сходятся к их пределам с вероятностью единица. Но легко показать, что

$$p \lim_{T \rightarrow \infty} \frac{1}{T} \sum u_t^2 = \frac{\sigma_0^2}{1 - \rho_0^2}, \quad p \lim_{T \rightarrow \infty} \frac{1}{T} \sum u_t u_{t-1} = \frac{\sigma_0^2 \rho_0}{1 - \rho_0^2}, \tag{76}$$

а значит пределы идентифицированы. Предшествующее доказательство показало, что $S_T(\omega)$ равномерно сходится к ее пределу, скажем $S(\omega)$ с вероятностью единица. Кроме того мы выяснили, что

$$S(\omega) = \sigma_0^2 + \frac{\sigma_0^2(\rho - \rho_0^2)^2}{1 - \rho_0^2} + \varphi(\omega), \tag{77}$$

где

$$\varphi(\omega) = \alpha_0^2 \lim_{T \rightarrow \infty} \left[\frac{x_0^* V^{-1} x_0^*}{T} - \frac{x_0^{*'} V^{-1} x^* (x^{*'} V^{-1} x_0^{*'})'}{x^{*'} V^{-1} x^*} \right] \tag{78}$$

и вектор x_0^* вычисляется при истинном значении параметра λ_0 . Ясно, что для всех допустимых ω , $\varphi(\omega) \geq 0$ и, в частности,

$$\varphi(\omega_0) = 0. \tag{79}$$

Таким образом, имеем

$$S(\omega_0) \leq S(\omega) \text{ для всех } \omega \in \Omega. \tag{80}$$

Теперь рассмотрим подпоследовательность $\{\hat{\omega}_{T_i}\}$ последовательности $\{\hat{\omega}_T\}$, причем первая сходится к ω_* , которая является предельной точкой последней. По построению подпоследовательности мы имеем

$$S_{T_i}(\hat{\omega}_{T_i}) \leq S_T(\omega) \quad \forall \omega \in \Omega, \tag{81}$$

следовательно,

$$S(\omega_*) \leq S(\omega) \quad \forall \omega \in \Omega, \tag{82}$$

а вследствие (80)

$$S(\omega_*) = S(\omega_0) = \sigma_0^2, \tag{83}$$

ω_0 — истинный вектор параметров $(\lambda_0, \rho_0)'$. Но (83) немедленно влечет

$$\rho_* = \rho_0, \quad \varphi(\omega_*) = 0. \tag{84}$$

Для идентификации модели наложим условие

$$\varphi(\omega_*) = 0, \tag{85}$$

что предполагает

$$\lambda_* = \lambda_0. \tag{86}$$

Если (85) и (86) удовлетворяются, мы заключаем, что $\omega_* = \omega_0$ и так как ω_* есть любая предельная точка последовательности $\{\hat{\omega}_T\}$, то $\{\hat{\omega}_T\}$ сходится к ω_0 с вероятностью единица.

Замечание 2. Важно отметить, что условия идентифицируемости (85) и (86) совсем не бессмысленны. Рассмотрим, например, x -последовательность вида: $x_i = (-1)^i$. Она отвечает всем требованиям к таким последовательностям, которые были сформулированы выше. С другой стороны, находим

$$x_i^* = \sum_{i=0}^{\infty} \lambda^i x_{i-i} = (-1)^i \sum_{i=0}^{\infty} (-\lambda)^i = (-1)^i \frac{1}{1+\lambda}, \quad \lim_{T \rightarrow \infty} \frac{1}{T} x^{*'} V^{-1} x = \frac{1+\rho}{1+\lambda},$$

Следовательно,

$$\varphi(\omega) = \alpha_0^2 \left[\frac{(1+\rho)^2}{(1+\lambda_0)^2} - \frac{(1+\rho)^2}{(1+\lambda_0)^2} \right] = 0,$$

откуда видно, что $S(\omega)$ не зависит от λ . Таким образом, доказательство, приведенное выше, ничего не говорит о пределе последовательности $\{\hat{\lambda}_T\}$. Для той конкретной x -последовательности, которую мы рассматриваем сейчас, показано также, что $\{\hat{\alpha}_T\}$, определенная выражением (58), не сходится к α_0 , если λ_0 неизвестно и не используется при получении $\hat{\alpha}_T$. Если истинное значение параметра ω_0 — внутренняя точка множества Ω , то асимптотическое распределение оценки MCS устанавливается разложением

$$\frac{\partial S}{\partial \theta}(\hat{\theta}) = \frac{\partial S}{\partial \theta}(\theta_0) + \frac{\partial^2 S}{\partial \theta \partial \theta}(\bar{\theta})(\hat{\theta} - \theta_0), \quad (87)$$

где $\hat{\theta}$ — MCS -оценка θ_0 . Последняя — истинное значение параметра и $\bar{\theta}$ удовлетворяет неравенству $|\bar{\theta} - \theta_0| \leq |\theta - \theta_0|$. Доказательство требует применения центральной предельной теоремы для m зависимых переменных — см., например, [8, гл. 4]. Приближенное доказательство впервые было дано Манном и Вальдом [26], особенно лемма 2. Затем определим, что асимптотически

$$\sqrt{T}(\hat{\theta} - \theta_0) \sim N(0, C_{MCS}), \quad (88)$$

где

$$C_{MCS} = \sigma_0^2 \lim_{T \rightarrow \infty} \begin{bmatrix} \frac{x^{*'} V^{-1} x^*}{T} & \frac{1}{T} \alpha_0 x^{*'} V^{-1} \frac{\partial x^*}{\partial \lambda} & 0 \\ \frac{1}{T} \alpha_0 \left(\frac{\partial x^*}{\partial \lambda} \right)' V^{-1} x^* & \alpha_0^2 \left(\frac{\partial x^*}{\partial \lambda} \right)' V^{-1} \left(\frac{\partial x^*}{\partial \lambda} \right) & 0 \\ 0 & 0 & \frac{\sigma_0^2}{1 - \rho_0^2} \end{bmatrix}^{-1} \quad (89)$$

Это следует понимать так, что x^* и $\partial x^* / \partial \lambda$ вычисляются при истинном значении параметра λ_0 .

Замечание 3. Интересной особенностью только что упомянутого результата является то, что асимптотически оценка $(\alpha, \lambda)'$ не зависит от оценки ρ . Вследствие этого можно определить двухшаговую оценку инструментальных переменных (IV), с тем же асимптотическим распределением, что и оценка MCS (или ML). Этот факт был установлен Амеция и Фуллером [27]. Чтобы показать, как это было обнаружено, запишем модель

в виде

$$y_t = \lambda y_{t-1} + \alpha x_t + v_t, \quad v_t = u_t - \lambda u_{t-1}. \quad (90)$$

Примем

$$\text{Cov}(v) = \sigma^2 \Phi, \quad v = (v_2, v_3, \dots, v_T)'. \quad (91)$$

Пусть имеются начальные состоятельные оценки для α и λ , например, $\hat{\alpha}$ и $\hat{\lambda}$. Тогда можно вычислить

$$\tilde{u}_t = y_t - \tilde{y}_t', \quad \tilde{y}_t = \hat{\alpha} \sum_{i=0}^{t-1} \hat{\lambda}^i x_{t-i}, \quad \hat{\rho} = \frac{\sum \tilde{u}_t u_{t-1}}{\tilde{u}_{t-1}} \quad (92)$$

и, таким образом, получить состоятельную оценку Φ , так как последняя зависит только от λ и ρ . Затем можно показать, что оценка

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\lambda} \end{pmatrix}_{IV} = [(x, \tilde{y}_{-1})' \tilde{\Phi}^{-1}(x, y_{-1})]^{-1} (x, \tilde{y}_{-1})' \tilde{\Phi}^{-1} y$$

имеет то же асимптотическое распределение, что и предельное (асимптотическое) распределение первых двух компонент $\hat{\theta}$, определенное (88) и (89). Исходные состоятельные оценки для α и λ были ранее получены из (90) путем использования x_t и x_{t-1} как инструментальных переменных.

Замечание 4. Очень соблазнительно использовать описанную выше упрощенную процедуру для анализа модели

$$y_t = \lambda y_{t-1} + \alpha x_t + \frac{I}{I - \rho L} \varepsilon_t, \quad (93)$$

так как при решении, и если $|\lambda| < 1$, находим

$$y_t = \frac{\alpha I}{I - \lambda L} x_t + \frac{I}{(I - \lambda L)(I - \rho L)} \varepsilon_t. \quad (94)$$

Однако получившиеся оценки не имеют того же асимптотического распределения, что и *MCS*-оценка. Причина этого впервые была рассмотрена Амеция и Фуллером [27], а также автором [8]. Она заключается в том, что в модели (93) λ является и средним и дисперсией y -процесса, как это видно из (94). В модели, рассмотренной ранее, λ была только средним, но не дисперсией (для y -процесса). Это объяснение становится совершенно ясным, если вспомнить, что в случае нормальных распределений (которые имеют место асимптотически) оценки среднего и дисперсии взаимно независимы. В этой связи Маддала [28] также отметил, что в случае, когда процесс ошибок является нормально распределенным, информационная матрица — блочно-диагональная в первом случае, но не в случае, представленном моделью (93).

Заканчивая изучение структуры геометрического лага, рассмотрим проблемы, возникающие при анализе модели (93), и, в частности, в какой степени двухшаговая оценка инструментальных переменных, рассмотренная выше, и другие двухшаговые процедуры отличаются от оценки *MCS*. Оценка *MCS* параметров α , λ и ρ имеет асимптотическое распределение

$$\sqrt{T}(\hat{\theta} - \theta) \sim N(0, \sigma^2 \Omega^*), \quad (95)$$

где $\theta = (\alpha, \lambda, \rho)'$, а θ — *MCS*-оценка θ и

$$\Omega^* = \lim_{T \rightarrow \infty} \left\{ \frac{1}{T} \begin{bmatrix} x'V^{-1}x, x'V^{-1}\bar{y}_{-1} & 0 \\ \bar{y}'_{-1}V^{-1}x, \bar{y}'_{-1}V^{-1}\bar{y}_{-1} + \frac{T\sigma^2}{1-\lambda^2} \frac{T\sigma^2}{1-\lambda^2} & \\ 0 & \frac{T\sigma^2}{1-\lambda\rho} & \frac{T\sigma^2}{1-\rho^2} \end{bmatrix} \right\}^{-1}. \quad (96)$$

Важное различие между моделями (93) и (52) заключается в том, что в первой из них оценки параметров ρ и (α, λ) зависимы (даже асимптотически). Следовательно, качество оценки ρ даже асимптотически будет влиять на качество оценки (α, λ) . Как видим, здесь положение противоположное тому, что было в модели (52), где свойства оценки параметра ρ — помимо состоятельности — не вытекали из определения (асимптотических) свойств *MCS*-оценки параметров (α, λ) . Таким образом, можно ожидать, что двухшаговая оценка инструментальных переменных, рассмотренная ранее, не будет совпадать асимптотически с *MCS*-оценкой (α, λ) . Этот результат будет достаточно очевиден, если мы заметим, что исходная оценка инструментальных переменных (α, λ)

$$\begin{pmatrix} \bar{\alpha} \\ \bar{\lambda} \end{pmatrix} = \left[\begin{pmatrix} x' \\ x_{-1}' \end{pmatrix} (x, y_{-1}) \right]^{-1} \begin{pmatrix} x'y \\ x_{-1}'y \end{pmatrix} \quad (97)$$

и, после итерации

$$\begin{pmatrix} \bar{\alpha} \\ \bar{\lambda} \end{pmatrix}_{IV} = \left[\begin{pmatrix} x' \\ \tilde{y}'_{-1} \end{pmatrix} \tilde{V}^{-1}(x, y_{-1}) \right]^{-1} \begin{pmatrix} x'\tilde{V}^{-1}y \\ \tilde{y}'_{-1}\tilde{V}^{-1}y \end{pmatrix}. \quad (98)$$

Оценка V^{-1} получается из оценки ρ , заданной выражением

$$\bar{\rho} = \frac{\sum_{t=3}^T \tilde{u}_t \tilde{u}_{t-1}}{\sum_{t=3}^T \tilde{u}_{t-1}^2}, \quad \tilde{u}_t = y_t - (x_t, y_{t-1}) \begin{pmatrix} \bar{\alpha} \\ \bar{\lambda} \end{pmatrix}, \quad t = 2, 3, \dots, T. \quad (99)$$

Асимптотическое распределение оценки (98) совершенно очевидно

$$\sqrt{T} \left[\begin{pmatrix} \bar{\alpha} \\ \bar{\lambda} \end{pmatrix}_{IV} - \begin{pmatrix} \alpha \\ \lambda \end{pmatrix} \right] \sim N(0, \sigma^2 C_{IV}), \quad (100)$$

где

$$C_{IV} = \lim_{T \rightarrow \infty} \left\{ \frac{1}{T} \begin{bmatrix} x'V^{-1}x & x'V^{-1}\bar{y}_{-1} \\ \bar{y}'_{-1}V^{-1}x & \bar{y}'_{-1}V^{-1}\bar{y}_{-1} \end{bmatrix} \right\}^{-1}, \quad (101)$$

где

$$\bar{y}_{-1} = (\bar{y}_1, \bar{y}_2, \dots, \bar{y}_{T-1}) \bar{y}'_t = \frac{\alpha I}{I - \lambda L} x_t. \quad (102)$$

Можно показать, что предельное асимптотическое распределение *MCS*-оценки (α, λ) дается выражением

$$\sqrt{T} \left[\begin{pmatrix} \bar{\alpha} \\ \bar{\lambda} \end{pmatrix}_{MCS} - \begin{pmatrix} \alpha \\ \lambda \end{pmatrix} \right] \sim N(0, C_{MCS}), \quad (103)$$

где

$$C_{MCS} = \lim_{T \rightarrow \infty} \left\{ \frac{1}{T} \left[x'V^{-1}x, x'V^{-1}\bar{y}_{-1} \right. \right. \\ \left. \left. \bar{y}'_{-1}V^{-1}x, \bar{y}'_{-1}V^{-1}\bar{y}_{-1} + \frac{T\sigma^2(\lambda - \rho)^2}{(1 - \lambda^2)(1 - \lambda\rho)^2} \right] \right\}^{-1}. \quad (104)$$

Таким образом, МС-оценка параметров (α, λ, ρ) есть оценка, которая (асимптотически) эффективна по отношению к двухшаговой оценке инструментальных переменных параметров (α, λ) при $\lambda \neq \rho$. Когда $\lambda = \rho$, две эти оценки равно эффективны. Следует подчеркнуть, однако, что при получении оценок мы в обоих случаях не пользуемся условием $\lambda = \rho$. Предложено множество других двухшаговых процедур для оценки параметров модели (93) — см., например, Гупта [29], автор [19, 30]. В общем такие процедуры хуже МС-оценок (в отношении асимптотической эффективности). Более широкое обсуждение таких вопросов можно найти у автора [8] и Гретера и Маддала [31].

в) Общий рациональный лаг. Как отмечалось ранее, эта формулировка обобщает полиномиальный, геометрический и отрицательный биномиальный варианты модели распределенного лага. Описание рационального лага имеет вид

$$y_t = \frac{A(L)}{B(L)} x_t + u_t, \quad t = 1, 2, \dots, T, \quad (105)$$

где

$$A(L) = \sum_{i=0}^m a_i L^i, \quad \sum_{j=0}^n b_j L^j, \quad b_0 = 1, \quad m \leq n. \quad (106)$$

Подход к оценке параметров этой модели был предложен Юргенсоном [32]. Подход этот заключается в следующем. Модель (105) после сокращения приводится к виду

$$y_t = B^*(L) y_t + A(L) x_t + B(L) u_t, \quad (107)$$

где

$$B^*(L) = - \sum_{j=1}^n b_j L^j. \quad (108)$$

Это — авторегрессивная модель со скользящей средней, и предложение Юргенсона заключалось в том, чтобы положить

$$\varepsilon_t = B(L) u_t, \quad (109)$$

принять, что $\{\varepsilon_t : t = 1, 2, \dots\}$ — последовательность о.н.р. случайных переменных с нулевой средней и дисперсией σ^2 , и оценивать параметры с помощью метода наименьших квадратов. Этот подход, однако, пригоден лишь, если в описании исходной модели (105) процесс ошибок имеет форму

$$u_t = \frac{I}{B(L)} \varepsilon_t. \quad (110)$$

Напомним, что подобное условие соблюдалось в формулировке модели геометрического лага Нерлова (18) и (20). Мы не можем, однако, доверчиво принять (110) как окончательную гипотезу. Если, как это делалось в (105), принять традиционные допущения относительно процесса ошибок, то использование обычных методов наименьших квадратов (применительно к (107)) даст несостоятельные оценки, так как (независимые) пе-

ременные в правой части не являются более независимыми или даже некоррелированными с процессом ошибок в (107). Таким образом, так же, как и было в случае рассмотренной выше приведенной модели геометрического лага, требуются различные процедуры оценки. Мы сначала отметили, что если процесс ошибок в (105) представляет собой последовательность о.н.р. случайных переменных, мы можем получить оценки методами минимума χ -квадрат (здесь — нелинейным методом наименьших квадратов). Итак, рассматриваем задачу минимизации

$$S(\theta) = \sum_{i=1}^T \left(y_i - \frac{A(L)}{B(L)} x_i \right)^2, \quad \theta = (a_0, a_1, \dots, a_m, b_1, b_2, \dots, b_n) \quad (111)$$

относительно неизвестных элементов A и B . В действительности, это задача выявления состоятельности и асимптотической нормальности оценки, скажем $\hat{\theta}$, полученной с помощью операции $\sup_{\theta} S(\theta)$. Она может быть решена с помощью методов, изложенных в этом разд., п. б). Проблема здесь в том, что если m велико, скажем, больше двух, применение поисковой техники — дело нереальное. Таким образом, мы приходим к итеративной процедуре следующего вида. Пусть

$$a = (a_0, a_1, a_2, \dots, a_m)', \quad b = (b_1, b_2, \dots, b_n)'. \quad (112)$$

Продифференцируем (111) для получения * нормальных уравнений

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^T \left(y_i - \frac{A(L)}{B(L)} x_i \right) \frac{L^i}{B(L)} x_i = 0, \quad i = 0, 1, \dots, m, \quad (113)$$

$$\frac{\partial S}{\partial b} = 2 \sum_{i=1}^T \left(y_i - \frac{A(L)}{B(L)} x_i \right) \frac{A(L)}{[B(L)]^2} L^j x_i = 0, \quad j = 1, 2, \dots, n. \quad (114)$$

Положим

$$y_i^* = \frac{I}{I - B^*(L)} y_i, \quad x_i^* = \frac{I}{I - B^*(L)} x_i, \quad x_i^{**} = \frac{A(L)}{I - B^*(L)} x_i^*. \quad (115)$$

и перепишем (113) и (114) в виде

$$\sum_{i=1}^T \left(\sum_{i=0}^m a_i x_{i-i}^* - \sum_{j=1}^n b_j y_{i-j}^* \right) x_{i-i}^* = \sum_{i=1}^T y_i^* x_{i-i}^*, \quad i = 0, 1, 2, \dots, m, \quad (116)$$

$$\sum_{i=1}^T \left(\sum_{i=0}^m a_i x_{i-i}^* - \sum_{j=1}^n b_j y_{i-j}^* \right) x_{i-j}^{**} = \sum_{i=1}^T y_i^* x_{i-j}^{**}, \quad j = 1, 2, \dots, n. \quad (117)$$

Итеративная процедура, предложенная Стиглицем и Мак-Брайдом (SM) [13] есть прямое обобщение процедуры, описанной в предыдущем разделе при рассмотрении структуры геометрического лага. Итак, если имеется начальная *состоятельная оценка* векторов a и b (скажем, $\tilde{a}$ и $\tilde{b}$), то величины y_i^* , x_i^* , x_i^{**} могут быть вычислены рекурсиями и при исполь-

* Нижний предел индекса суммирования, разумеется, неверен. Вообще, так как наблюдения поступают в форме пар $\{(y_t, x_t): t = 1, 2, \dots, T\}$ проще вести суммирование при t , меняющемся от $n+1$ до T .

зовании соотношений (115) и значений $\tilde{a}$ и $\tilde{b}$. Затем решаем уравнения (116) и (117) и получаем первое приближение $\tilde{a}^{(1)}$ и $\tilde{b}^{(1)}$. Используя полученные значения, мы можем пересчитать y_i^* , x_i^* , x_i^{**} и получить второе приближение $\tilde{a}^{(2)}$ и $\tilde{b}^{(2)}$ и т. д. — пока процесс не сойдется. Результаты последней итерации, например, $(\hat{a}, \hat{b})$, и есть решение уравнений (113) и (114). Так как итеративный процесс начат с состоятельной оценки, то можно показать, что каждая из промежуточных итераций состоятельна и, таким образом, и последняя итерация, если процесс сходится, дает состоятельные корни нормальных уравнений. Его асимптотическое распределение можно установить применением теоремы о среднем значении к функции градиента

$\frac{\partial S}{\partial \begin{pmatrix} a \\ b \end{pmatrix}}$ К сожалению нет общих теорем, точно описывающих условия при которых итеративный процесс сходится.

4. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Значительная часть этого обзора посвящена формулировке, интерпретации и, в особенности, проблемам оценки, возникающим при рассмотрении моделей распределенного лага. Мы рассмотрели модель рационально распределенного лага и две важных ее разновидности — модели лага с геометрическим и полиномиальным распределением. При изучении этих моделей предполагалось, что процесс ошибок представляет собой либо последовательность независимых случайных переменных с одинаковым распределением, либо является авторегрессионным процессом первого порядка. Нет, однако, оснований ограничиваться рассмотрением лишь таких процессов ошибок. Действительно, Хэннэн [33] рассмотрел случай геометрического лага с линейным процессом ошибок (при постоянной ковариации), а автор [8, гл. 10] описывает модель общего рационального лага, где процесс ошибок также является линейным.

Много вопросов, касающихся проверки устойчивости, согласия и выбора модели (применительно к моделям распределенного лага), затронуто в работе автора [8, гл. 11]. Правда, там не предложено каких-либо определенных подходов к решению этих проблем. Так, в случае невложенных моделей нет процедуры, позволяющей отдать предпочтение одной модели перед другой. Так, например, если модель 1 описана в пространстве параметров ξ_1 , а модель 2 в пространстве параметров ξ_2 , причем $\xi_1 \not\subset \xi_2$ и $\xi_2 \not\subset \xi_1$ (иначе говоря — обе модели принадлежат одной и той же спецификации), то не существует удовлетворительного метода для достаточно ясного выбора наиболее подходящей модели на основе выборки. Совершенно недостаточно исследована также проблема оценок распределения распределенного лага на основе малых выборок. Центральной проблемой является проблема оценок, которая, в свою очередь, есть следствие сильно нелинейного характера большинства формулировок моделей распределенного лага. Наконец, совершенно неясными остаются проблемы, возникающие из-за неправильной спецификации. Хорошее начало исследованиям в этом направлении было положено Симсом [12].

Таким образом, хотя литература по моделям распределенного лага довольно обширна, нельзя утверждать, что проблемы, возникающие при рассмотрении таких моделей и их приложений, удовлетворительно разрешены.

ЛИТЕРАТУРА

1. I. Fisher. A Note on a Short-cut Method for Calculating Distributed Lags, Bulletin de l'Institut Internationale de Statistique, 1937, v. 29, p. 323—328.
2. S. Almon. The Distributed Lag between Capital Appropriations and Expenditures. *Econometrica*, 1965, v. 33, p. 178—196.
3. M. D. McCarthy. The Wharton Quarterly Econometric Forecasting Model, Mark III. Studies in Quantitative Econometrics, 1972, N 6, Economics Research Unit Department of Economics, University of Pennsylvania.
4. R. H. Raschem, H. T. Shapiro. The MIT-FRB Model: Its Special Features. *Amer. Econ. Rev.*, 1968, b. 58, p. 123—149.
5. L. M. Koyck. Distributed Lags and Investment Analysis. Amsterdam. North Holland Publishing Co., 1954.
6. P. Cagan. The Monetary Dynamics of Hyperinflation, in Studies in the Quantity Theory of Money, M. Friedman (Ed.), Chicago, University of Chicago Press, 1946.
7. M. Nerlove. The Dynamics of Supply: Estimation on Farmers' Response to Price. Baltimore. The Johns Hopkins Press, 1958.
8. P. J. Dhrymes. Distributed Lags: Problems of Estimation and Formulation. Holden—Day, San Francisco, 1971.
9. Z. Griliches. Distributed Lags: A Survey. *Econometrica*, 1967, v. 35, p. 16—49.
10. R. M. Solow. On a Family of Lag Distributions, *Econometrica*, 1960, v. 28, p. 393—406.
11. D. W. Jorgenson. Anticipations and Investment Behavior. Brookings Quarterly Econometric Model of the United States, J. S. Duesenberry, G. Fromm, L. R. Klein, and E. Kuh (Eds.), Chicago, Rand McNally and Co., 1965.
12. Lags in Econometrics, *Econometrica*, 1971, v. 39, p. 545—563.
13. K. Steiglitz, L. E. McBride. A Technique for the Identification of Linear System. *IEEE Transactions on Automatic Control*, AC-10, 1965, p. 461—464.
14. K. J. Astrom and P. Eykhoff. Identification and Process Parameter Estimation, presented at the 1970 IFAC Symp. Identification, Prague.
15. R. I. Jennrich. Asymptotic Properties of Non-Linear Least Squares. *Annals of Mathematical Statistics*, 1969, v. 40, p. 633—643.
16. E. Malinvaud. The Consistency of Non-Linear Regression. *Annals of Mathematical Statistics*, 1970, v. 41, p. 956—969.
17. A. Wald. A Note on the Consistency of the Maximum Likelihood Estimate. *Annals of Mathematical Statistics*, 1949, v. 20, p. 595—601.
18. J. Wolfowitz. On Wald's Proof of the Consistency of the Maximum Likelihood Estimate. *Annals of Mathematical Statistics*, 1949, v. 20, p. 601—602.
19. P. J. Dhrymes. Efficient Estimation of Distributed Lags with Autocorrelated Errors. *International Econ. Rev.*, 1969, v. 10, p. 47—67.
20. J. L. Morrison. Small Sample Properties of Selected Distributed Lag Estimators: A Monte Carlo Experiment. *International Econ. Rev.*, 1970, v. 11, p. 13—23.
21. L. R. Klein. The Estimation of Distributed Lags, *Econometrica*, 1958, v. 26, p. 553—556.
22. P. J. Dhrymes. On the Strong Consistency of Estimators for Certain Distributed Lag Models with Autocorrelated Errors. 1972, v. 12, p. 329—343.
23. H. L. Royden. *Real Analysis*. N. Y. McMillan, 1963.
24. L. Breiman. *Probability*, Readings, Mass., Addison Wesley, 1968.
25. I. I. Gikhman, A. V. Skorokhod. *Introduction to the Theory of Random Processes*. Philadelphia, 1969.
26. H. B. Mann, A. Wald. On the Statistical Treatment of Stochastic Difference Equations. *Econometrica*, 1943, v. 11, p. 173—220.
27. T. Amemiya, W. Fuller. A Comparative Study of Alternative Estimators in a Distributed Lag Model. *Econometrica*, 1967, v. 35, p. 509—529.
28. G. S. Maddala. Generalized Least Squares with an Estimated Covariance Matrix. *Econometrica*, 1971, v. 39, p. 23—34.
29. Y. P. Gupta. Least Squares Variant of the Dhrymes Two-Step Estimation Procedure of the Distributed Lag Model, *International Econ. Rev.*, 1969, v. 10, p. 112—113.
30. P. J. Dhrymes, L. R. Klein, K. Steiglitz. Estimation of Distributed Lags. *International Econ. Rev.*, 1970, v. 11, p. 235—250.
31. D. Grether, G. S. Maddala. A Comparison of Several Two-Step Distributed Lag Estimators. *International Econ. Rev.*
32. D. W. Jorgenson. Rational Distributed Lag Functions. *Econometrica*, 1960, v. 34, p. 135—149.
33. E. J. Hannan. The Estimation of Relationships Involving Distributed Lags. *Econometrica*, 1965, v. 33, p. 409—418.

Поступила в редакцию
10 IX 1972