МАТЕМАТИКО-СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

ВЕРИФИКАЦИЯ ЭКОНОМЕТРИЧЕСКОЙ МОДЕЛИ С УЧЕТОМ АПРИОРНЫХ ОГРАНИЧЕНИЙ НА СТРУКТУРНЫЕ ПАРАМЕТРЫ*

Н.В. Суворов

В статье описан метод верификации статистической модели, которая, во-первых, представлена временными рядами исходных данных и, во-вторых, является линейной по оцениваемым параметрам. Опыт статистических расчетов на основе реальных эмпирических данных свидетельствует о том, что наиболее известные и традиционно применяемые в практике эконометрического моделирования математико-статистические методы (метод наименьших квадратов, метод максимального правдоподобия и близкие к ним методы) очень часто не позволяют обеспечить успешную верификацию теоретически требуемых форм эконометрических моделей. Разработанный метод, называемый альтернативным методом линейной регрессии (АМЛР), обеспечивает учет априорных ограничений абсолютных значений и знаков параметров идентифицируемой модели. В основе АМЛР лежит концепция наилучшего линейного индекса, известная в теории статистики с конца 1950-х годов. В математическом отношении АМЛР основывается на методе главных компонент. Проанализированы условия применения АМЛР в эконометрическом моделировании и методы преобразования исходной статистической информации, обеспечивающие корректность применения разработанной процедуры оценивания.

Специальной проблемой предложенного метода является определение уровня точности аппроксимации зависимой переменной модели. В связи с этим для оценки уровня точности статистической модели, верифицируемой при помощи АМЛР, разработан оригинальный метод декомпозиции временного ряда на регулярную и стохастическую компоненты. Проанализированы свойства предлагаемого метода декомпозиции и дана числовая иллюстрация его применения в эконометрических расчетах.

Ключевые слова: статистическая модель, временные ряды, наилучший линейный индекс, декомпозиция временного ряда.
JEL: C01, C51.

Применение эконометрических методов базируется, как правило, на представлении изучаемого процесса в виде линейной регрессионной модели, которая в стандартном виде задается соотношениями

$$y = Xa + \varepsilon,$$

$$X = \begin{bmatrix} x_{11} \ldots x_{1m} \\ \ldots\ldots\ldots\ldots \\ x_{T1} \ldots x_{Tm} \end{bmatrix},\ y' = (y_1,...,y_T),\ a' = (a_1,...,a_m),\ \varepsilon' = (\varepsilon_1,...,\varepsilon_T), \quad (1)$$

где m - число оцениваемых структурных параметров; T - число наблюдений (длина временных рядов переменных, если речь идет о динамических процессах); y - вектор значений зависимой переменной; X - матрица наблюдений объясняющих переменных; α - вектор структурных параметров; ε - вектор случайных отклонений, которые, по предположению, обладают нулевым математическим ожиданием [$M(\varepsilon) = 0$] и фиксированной дисперсией [$M(\varepsilon\varepsilon') = \sigma_\varepsilon^2 E$, где E - единичная матрица, σ_ε^2=const]; «′» - символ транспонирования. Коэффициенты модели (1) предполагаются постоянными и отражают степень влияния каждого из факторов, включенных в модель, в среднем за период $1,...,T$. Условия существования оценок $\{a_i\}$ в модели (1): 1) число наблюдений не меньше числа оцениваемых параметров и 2) ранг матрицы X равен числу оцениваемых параметров, то есть столбцы матрицы X линейно независимы. В регрессионном анализе рассматриваются задачи, в которых $T \geqslant m$ и, как правило, T существенно больше m.

Опыт статистических расчетов на основе реальных эмпирических данных свидетельствует о том, что наиболее известные и традицион-

Суворов Николай Владимирович (suvor_n@ecfor.ru) - д-р экон. наук, профессор, руководитель лаборатории прогнозирования динамики и структуры народного хозяйства Института народнохозяйственного прогнозирования Российской академии наук (г. Москва, Россия).

*Статья подготовлена с использованием материалов, разрабатывавшихся при финансовой поддержке Российского фонда фундаментальных исследований (проект № 16-06-00319).

но применяемые в практике эконометрического моделирования математико-статистические методы (метод наименьших квадратов, метод максимального правдоподобия и близкие к ним методы) очень часто не позволяют обеспечить успешную верификацию теоретически требуемых форм эконометрических моделей.

Типовыми случаями при этом являются: 1) несоответствие знака или порядка какого-либо из структурных параметров модели (1) априорным представлениям исследователя о моделируемом процессе; 2) большие величины стандартных ошибок оценок отдельных структурных параметров моделей типа (1) либо неудовлетворительные в целом результаты оценивания регрессионной модели.

В связи с этим настоятельной оказывается разработка метода идентификации структурных параметров эконометрической модели типа (1), альтернативного методам наименьших квадратов, максимального правдоподобия или их модификациям. Принципиально важным свойством этого гипотетического метода верификации статистической модели должна являться возможность обеспечения теоретических (или априорных) требований, предъявляемых к спецификации[1] оцениваемой статистической конструкции.

Формулировка метода

Упомянутые выше априорные требования к знакам и уровням оцениваемых параметров могут быть специфицированы различным образом.

Так, требование неотрицательности некоторого параметра a_i из (1) традиционно формализуется в виде ограничения

$$a_i \geq 0, \tag{2}$$

добавляемого к соотношениям модели (1). Однако, как хорошо известно, нахождение решения системы (1) с учетом (2) по методу наименьших квадратов или по методу максимума правдоподобия дает следующий результат: либо $a_i = 0$ [это означает, что ограничение (2) является существенным], либо использование (2) излишне [в том случае, если оценивание модели (1) без учета (2) и так приводит к неотрицательному значению a_i].

Другой способ задания ограничений на параметры модели (1), разработанный в эконометрической теории, - представление области возможных значений всех или части параметров в виде

$$r = Ra + \delta, \tag{3}$$

где r - заданный вектор размерности k ($k \leq m$); R - заданная матрица размерности $k*m$; δ - k-мерный вектор погрешностей с нулевым математическим ожиданием и известной ковариационной матрицей Ψ размерности $k*k$.

Дополнение исходной модели (1) соотношением (3) приводит к определению искомого вектора a по формуле:

$$a = (X'X + \sigma_\varepsilon^2 R'\Psi^{-1}R)^{-1}(X'y + \sigma_\varepsilon^2 R'\Psi^{-1}r). \tag{4}$$

Поскольку значение σ_ε^2 априори неизвестно, вычисление a в соответствии с (4) требует использования итерационной процедуры. На первом шаге производится вычисление оценки a модели (1) «в чистом виде» [то есть без использования ограничений (3)] и находится первоначальная оценка дисперсии σ_ε^2; далее производится вычисление a в соответствии с (4), что позволяет сформировать новую оценку σ_ε^2. Вычисления продолжаются до достижения сходимости. Данный метод известен в эконометрической литературе как метод Дарбина или метод Тейла-Гольдбергера [1].

В частном случае, когда R - единичная матрица порядка m, Ψ - диагональная матрица также порядка m, соотношение (3) означает, что предварительная (априорная) информация о параметрах модели (1) представлена элементами m-мерного вектора r, а априорные дисперсии оцениваемых параметров - диагональными эле-

[1] Под спецификацией понимается перечень объясняющих переменных оцениваемой модели.

ментами матрицы Ψ. При названных условиях объединение выборочной информации [представленной данными модели (1)] и априорной информации, представленной в виде (3), приводит к тому, что оценка вектора структурных параметров в соответствии с (4) оказывается средневзвешенной из оценок параметров, получаемых исходя из выборочной информации и их априорных оценок.

Формализация ограничений на параметры в форме (3) в принципе позволяет обеспечить их (параметров) нахождение в заранее предполагаемых областях. При этом использование (3) предполагает, что априорные значения искомых параметров являются центром области их возможного нахождения, а априорные значения дисперсий определяют величину этой области.

Вместе с тем, как показывает наш опыт статистических расчетов, объединение выборочной информации в виде (1) и априорных ограничений в виде (3) приводит к тому, что, как правило, именно априорная информация оказывает преимущественное влияние на результаты оценивания. Очевидно, что потребность в использовании метода смешанного оценивания возникает в тех случаях, когда оценивание модели (1) «в чистом виде» дает неудовлетворительные результаты. Например, пусть по результатам расчетов i-й структурный параметр модели (1) принимает отрицательное значение и имеет большое стандартное отклонение. Предположим также, что для указанного структурного параметра имеется априорная информация в виде $0 \le a_i \le 1$, а интервал между нулевым и единичным значениями исчерпывает возможную область нахождения данного параметра. Тогда в качестве априорного значения a_i кажется естественным принять 1/2, а его стандартное отклонение не должно превышать 1/6, если исходить из гипотезы, что δ_i - нормально распределенная случайная величина. При сделанных предположениях результат оценивания a_i [каково бы ни было происхождение статистических данных, используемых в модели (1)], как правило, заведомо будет близок к 1/2. Однако в случае, когда предположение, что $0 \le a_i \le 1$ неверно, результат применения метода смешанного оценивания следует признать неудовлетворительным.

В более общем плане можно говорить о том, что при использовании метода смешанного оценивания ошибочность априорных представлений относительно значений структурных параметров модели (1) далеко не всегда может быть компенсирована выборочной информацией.

Необходимо отметить, что исследователь, занимающийся оцениванием статистических моделей, как правило, заранее формулирует гипотезы о направлении воздействия тех или иных факторов (объясняющих переменных) на объясняемую переменную. Однако формализовать это знание в виде ограничений типа (3) далеко не всегда возможно.

В наших работах [2, 3] на примере оценивания макроэкономической и отраслевых производственных функций для отечественной экономики изложен метод (называемый АМЛР), который, с одной стороны, решает проблему задания ограничений на допустимую область возможных значений структурных параметров эконометрической модели типа (1), а с другой - не требует задания в явном виде предварительных значений оцениваемых параметров.

В математическом отношении АМЛР основывается на методе главных компонент. С точки зрения математической теории индексов (как направления теории экономической статистики), в основе АМЛР лежит конструкция наилучшего линейного индекса (см., например, [4]).

Понятие наилучшего линейного индекса было введено в научный оборот в теории статистики в конце 1950-х - начале 1960-х годов. Необходимость его использования возникает в случае, когда некоторый экономический процесс может быть описан несколькими альтернативными временными рядами индексных чисел.

Рассмотрим следующий пример. Пусть динамика физического объема промышленного производства за период времени $(1, 2,..., T)$ представлена временными рядами погодовых базисных индексов, исчисление которых осуществляется 1) исходя из постоянных цен на начало периода и 2) исходя из постоянных цен на конец периода. Соответствующие временные ряды обозначим как $I(1) = (i_{11,}\, i_{21},..., i_{T1})$ и $I(2) = (i_{12,}\, i_{22},..., i_{T2})$. Полное совпадение $I(1)$ и $I(2)$ практически исключено, особенно если T достаточно велико. В указанных условиях для оценки динамики производства как раз и может быть использован наилучший линейный индекс, представляющий собой линейную комбинацию $I(1)$ и $I(2)$. Метод его исчисления заключается в следующем.

Образуем из рядов $I(1)$ и $I(2)$ матрицу размерности T*2 вида

$$J = \begin{bmatrix} i_{11} & i_{12} \\ \multicolumn{2}{c}{\cdots\cdots\cdots} \\ i_{t1} & i_{t2} \\ \multicolumn{2}{c}{\cdots\cdots\cdots} \\ i_{T1} & i_{T2} \end{bmatrix}.$$

Для построения наилучшего линейного индекса необходимо определить матрицу W вида

$$W = \begin{bmatrix} u_1 v_1 & u_1 v_2 \\ \multicolumn{2}{c}{\cdots\cdots\cdots} \\ u_t v_1 & u_t v_2 \\ \multicolumn{2}{c}{\cdots\cdots\cdots} \\ u_T v_1 & u_T v_2 \end{bmatrix} = uv', \qquad (5)$$

где u, v - векторы соответствующей размерности. Матрица W рассматривается как аппроксимация исходной матрицы J.

Близость матриц J и W определяется на основе требования минимизации выражения:

$$\sum_{k,t} (i_{tk} - u_t v_k)^2 . \qquad (6)$$

Минимизация выражения (6) по переменным $\{u_t,\, v_k\}$ эквивалентна нахождению следа (суммы диагональных элементов) матрицы $\Delta\Delta'$, где Δ - матрица остатков: $\Delta = J\text{-}W$. По определению, след матрицы $tr(\Delta\Delta')$ есть:

$$tr(\Delta\Delta') = tr(JJ') - 2tr(Jvu') + tr[v'(u'u)v] =$$
$$= tr(JJ') - 2tr(Jvu') + (v'v)(u'u).$$

Непосредственно проверяется, что $tr(Jvu') = u'Jv$; соответственно, минимизируемый функционал (6) есть:

$$tr(JJ') - 2u'Jv + (v'v)(u'u). \qquad (7)$$

Приравнивание к нулю производных выражения (7) дает следующие необходимые условия экстремума:

$$Jv = (v'v)\, u;\ J'u = (u'u)v.$$

Далее приведенные выражения преобразуются к виду:

$$J'Jv = [(v'v)(u'u)]v;\ JJ'u = [(v'v)(u'u)]u. \qquad (8)$$

Условия (8) означают, что вектор v является собственным вектором матрицы $J'J$, а вектор u - собственным вектором матрицы JJ'. Скалярная величина $[(v'v)\ (u'u)]$ имеет смысл собственного числа матриц JJ' или $J'J$.

В силу приведенных выше соотношений член $u'Jv$ выражения (7) оказывается равным $(v'v)(u'u)$, поэтому (7) приводится к виду:

$$tr(JJ') - (v'v)(u'u).$$

Минимум этого последнего выражения достигается, очевидно, при максимальном значении $(v'v)(u'u)$. Из теории матриц известно, что собственные значения симметрических действительных матриц неотрицательны. Отсюда следует, что v и u в (8) должны выбираться как первые собственные векторы (то есть соответствующие максимальному собственному числу).

Вектор $u = (u_1,..., u_T)$ и является наилучшим линейным индексом, поскольку наиболее точно (в квадратичной метрике) воспроизводит индексы $I(1)$ и $I(2)$, являясь при этом линейной комбинацией указанных индексов:

$$u = v_1 I(1) + v_2 I(2). \qquad (9)$$

В силу того что собственные векторы определены с точностью до постоянного множителя, без ограничения общности можно положить, что $v_1 + v_2 = 1$. При этом веса v_1 и v_2, с которыми исходные индексы агрегируются в наилучший линейный индекс, всегда являются положительными вследствие того, что все элементы матрицы J - неотрицательные величины.

Кроме того, вектор u является первой главной компонентой матрицы исходных данных J.

Проведем замену переменных в (9), положив, что $I(1) = y\text{-}x_1$, $I(2) = y\text{-}x_2$, где y, x_1, x_2 - некоторые векторы размерности T. Тогда выражение (9) преобразуется к виду

$$y = v_1 x_1 + v_2 x_2 + u \qquad (10)$$

или к виду

$$y\text{-}x_2 = v_1(x_1\text{-}x_2)+u = I(2) = v_1(I(2) - I(1))+u, \quad (11)$$

поскольку $v_1 + v_2 = 1$.

Выражение (10) правомерно трактовать как функцию, зависимая переменная которой является линейно однородной относительно переменных x_1, x_2, а также включает некоторую дополнительную переменную u. При этом если переменные y, x_1, x_2 не могут быть однозначно определены исходя из известных векторов $I(1)$ и $I(2)$, то для (11) зависимая и объясняющая переменные определяются однозначно.

Таким образом, построение наилучшего линейного индекса в рассматриваемом примере может быть интерпретировано и как процедура идентификации модели (11), в которой оцениваемый структурный параметр v_1 априори удовлетворяет условию $0 \le v_1 \le 1$.

Описанная выше процедура построения наилучшего линейного индекса может быть реализована для различного количества альтернативных временных рядов индексных чисел. Следовательно, правомерно утверждать, что процедура агрегирования m временных рядов длины T по правилу исчисления наилучшего линейного индекса при условии $m < T$ эквивалентна верификации некоторой статистической модели типа (1):

$$y_t = \sum_{i=1}^{m} a_i x_{ti} + f(t), \qquad (12)$$

включающей m объясняющих переменных, на структурные параметры которой наложены априорные ограничения:

$$0 \le a_i \le 1; \ \sum_{i=1}^{m} a_i = 1, \qquad (13)$$

а функция $f(t)$ по смыслу аналогична u из выражений (10) - (11).

Применительно к проблеме верификации модели (12) необходимо отметить два момента.

Во-первых, требование неотрицательности оцениваемых параметров не может считаться ограничительным: если по априорным соображениям i-й параметр должен лежать в отрицательной области, то можно переопределить i-ю объясняющую переменную, заменив x_{ti} на $(-x_{ti})$[2].

Во-вторых, использование конструкции наилучшего линейного индекса для верификации модели (12) возможно лишь в случае, если для всех значений t выполнятся требование

$$z_{ti} = (y_t\text{-}x_{ti}) \ge 0, \ i = 1,\dots,m.$$

Данное требование вытекает из того, что все элементы матрицы J, как уже было отмечено выше, должны быть неотрицательны. Это может быть обеспечено преобразованием временных рядов $(z_{1i},\dots,z_{ti},\dots, z_{Ti})$ по правилу так называемого линейного нормирования:

$$z^s_{ti} = (z_{ti} - z^{min}_i) \,/\, (z^{max}_i - z^{min}_i), \qquad (14)$$

где z^{max}_i, z^{min}_i - соответственно максимальное и минимальное значения i-го временного ряда; $\{z^s_{ti}\}$ - стандартизованные значения элементов соответствующих временных рядов. В результате указанного преобразования значения временных рядов переменных $\{z^s_{ti}\}$, во-первых, заведомо будут неотрицательными и, во-вторых, будут находиться в интервале от 0 до 1. Тем самым также элиминируется воздействие различий в размерности исходных данных на результаты оценивания параметров модели.

Таким образом, применение АМЛР для оценивания параметров (12) с учетом ограничений (13) включает: 1) расчет разностей $z_{ti} = (y_t\text{-}x_{ti})$ и последующую стандартизацию полученных временных рядов по правилу линейного нор-

[2] Исследователь, занимающийся оцениванием статистических моделей, как правило, заранее формулирует гипотезы о направлении воздействия тех или иных факторов (объясняющих переменных) на объясняемую переменную. При отсутствии такого априорного знания возможна оценка альтернативных спецификаций модели с тем, чтобы выявить наиболее предпочтительную.

мирования, что дает векторы $Z^s_i = (z^s_{1i}, \ldots, z^s_{Ti})$; 2) формирование из векторов $Z^s_1,\ldots,Z^s_m$ матрицы Z^s, аналогичной матрице J из рассмотренного выше примера [см. (4)]; 3) расчет первого собственного вектора $V^s = (v^s_1,\ldots,v^s_m)$ матрицы $Z^{s\prime} Z^s$; 4) расчет параметров модели (12) в виде[3]

$$a_i=[v^s_i/(z^{max}_i - z^{min}_i)]:[\sum_{i=1}^{m} v^s_i/(z^{max}_i - z^{min}_i)].$$

Насколько существенным для оценивания с помощью АМЛР модели (12) является требование $\sum_{i=1}^{m} a_i=1$? Для иллюстрации данной проблемы рассмотрим модель:

$$y_t = ax_t + f(t), \tag{15}$$

в которой y_t, x_t - соответственно объясняемая и объясняющая переменные; a - подлежащий оцениванию структурный параметр при объясняющей переменной; $f(t)$ - неизвестная функция времени, также подлежащая определению; при этом из априорных соображений искомый параметр заключен в интервале: $0 < a < 1$. Оценивание модели (15) на основе АМЛР с учетом этого ограничения предполагает преобразование (15) к виду

$$f(t) = a(y_t - x_t) + (1-a)y_t, \tag{16}$$

где временные ряды $\{(y_t - x_t)\}$ и $\{y_t\}$ по смыслу аналогичны $I(1)$ и $I(2)$, а $f(t)$ - вектору u из выражения (9). Стандартизация временных рядов, входящих в (16), в соответствии с правилом (14) дает

$$z^s_{t1} = [(y_t - x_t) - (y - x)^{min}]/[(y-x)^{max} - (y-x)^{min}],$$
$$z^s_{t2} = (y_t - y^{min})/(y^{max} - y^{min}).$$

Элементы v^s_1, v^s_2 первого собственного вектора матрицы $Z^{s\prime} Z^s$, образованной из этих стандартизованных данных, суть

$$v^s_1 = ka[(y-x)^{max} - (y-x)^{min}],$$
$$v^s_2 = k(1-a)(y^{max} - y^{min}),$$

где k - некоторый коэффициент, поскольку, как уже отмечалось, элементы собственных векторов определяются с точностью до постоянного множителя. Тогда

$$v_1 = v^s_1/[(y-x)^{max} - (y-x)^{min}] = ka,$$
$$v_2 = v^s_2/(y^{max} - y^{min}) = k(1-a),$$

откуда следует, что

$$k = v_1 + v_2$$

и, соответственно, $a = v_1/(v_1 + v_2)$, то есть искомый коэффициент модели (15), во-первых, определяется однозначно и, во-вторых, $0 < a < 1$.

Оценивание же на основе АМЛР модели (15) при условии $a > 1$ приводит к необходимости рассмотрения выражения

$$f(t) = a(y_t - x_t) + (a-1)(-y_t). \tag{17}$$

Преобразование данных модели (17) и последующее определение v^s_1, v^s_2 и v_1, v_2 аналогично тому, как это было сделано для (16), дают, в конечном счете:

$$v_1 = ka,\ v_2 = k(a-1),$$

откуда следует, что $k = v_1 - v_2$ и $a = v_1/(v_1 - v_2)$. То есть несмотря на то, что v_1, v_2 - неотрицательные величины, искомый коэффициент a может оказаться как положительной, так и отрицательной величиной в зависимости от знака разности $(v_1 - v_2)$. Очевидно, что при $(v_1 - v_2) \le 0$ правомерно сделать вывод о неприемлемости гипотезы $a > 1$ для модели (17).

Из приведенного примера видно, что использование АМЛР предполагает приведение спецификации оцениваемой модели к такому виду, для которого требование

$$0 \le a_i \le 1;\ \sum_{i=1}^{m} a_i = 1 \tag{18}$$

выполнялось бы по определению или, по крайней мере, представляло бы один из возможных вариантов модели. В силу того что исходные данные статистической модели могут иметь самую различную размерность, обеспечение указанного требования представляется трудновыполнимым. Тем не менее можно указать несколь-

[3] Напомним, что $(v^s_1,\ldots,v^s_m)$ определяются с точностью до постоянного множителя.

ко достаточно общих приемов, способствующих решению данной проблемы.

Рассмотрим модель, включающую две объясняющие переменные (все обозначения аналогичны принятым ранее):

$$y_t = a_1 x_{t1} + a_2 x_{t2} + f(t). \tag{19}$$

Предположим, что средние значения зависимой и объясняющих переменных $\bar{y}, \bar{x}_1, \bar{x}_2$ отличны от нуля. Тогда выражение (19) представимо в виде

$$\begin{gathered} y_t = [a_1/(\bar{y}/\bar{x}_1)]\,(\bar{y}/\bar{x}_1)\,x_{t1} + \\ +[a_2/(\bar{y}/\bar{x}_2)]\,(\bar{y}/\bar{x}_2)x_{t2} + f(t) = \\ = b_1\{(\bar{y}/\bar{x}_1)x_{t1}\} + b_2\{(\bar{y}/\bar{x}_2)x_{t2}\} + f(t); \\ b_1 = a_1/(\bar{y}/\bar{x}_1),\ b_2 = a_2/(\bar{y}/\bar{x}_2), \end{gathered} \tag{20}$$

где в фигурных скобках стоят преобразованные значения объясняющих переменных, а b_1, b_2 - параметры, аналогичные коэффициентам эластичности[4]. Эти параметры являются безразмерными величинами.

Рассмотрим случай $b_1 \geq 0$, $b_2 \geq 0$ (как это имеет место для модели производственной функции). Если среднее значение $\bar{f}$ функции $f(t)$ равно нулю, то $b_1+b_2=1$; при $\bar{f} > 0$ $b_1+b_2 < 1$; при $\bar{f} < 0$ $b_1+b_2 > 1$. Тогда применение АМЛР гарантирует получение оценок b_1, b_2 (а значит, и оценок a_1, a_2), соответствующих априорным представлениям при $b_1+b_2 \leq 1$, а при $b_1+b_2 > 1$ оказывается возможной проверка корректности этой гипотезы.

В случае, если первоначально в (20) $b_1 \geq 0$ и $b_2 \leq 0$, то замена знака у второй объясняющей переменной, очевидно, не может гарантировать соблюдения равенства суммы искомых параметров эластичности единичному значению. Однако как показывает практика реализованных нами статистических расчетов, параметры эластичности, определенные в соответствии с (20), для линейных моделей, как правило, существенно не превышают единичного значения по абсолютной величине. Это позволяет осуществить оценивание параметров модели типа (20), произведя сначала оценку модели

$$y_t = b_1\{(\bar{y}/\bar{x}_1)x_{t1}\} + f(t), \tag{21}$$

а затем (исходя из полученного значения b_1) - оценку модели

$$y_t - b_1\{(\bar{y}/\bar{x}_1)x_{t1}\} = b_2\{(\bar{y}/\bar{x}_2)x_{t2}\} + f(t), \tag{22}$$

где все обозначения аналогичны (20) [отметим, что функции $f(t)$ в (21) и (22) не совпадают]. То есть искомые параметры модели (20) могут быть оценены при помощи АМЛР по отдельности[5].

Обобщая вышесказанное, правомерно заключить, что, во-первых, использование АМЛР предполагает предварительную трансформацию исходных данных статистической модели, обеспечивающую приведение модели к форме (12) с ограничениями (13), которую (форму) можно считать канонической. Приведение же к канонической форме в принципе возможно для самых разнообразных наборов статистических данных. Во-вторых, не вызывает сомнения, что способ преобразования переменных статистической модели, позволяющий использовать АМЛР, может определяться и спецификой предметной области, применительно к которой формулируется модель.

Оценка точности эконометрической модели, верифицированной при помощи АМЛР

Как было отмечено ранее, АМЛР применяется для оценивания эконометрической модели (12) с ограничениями (13); при этом $f(t)$ - некоторая временная функция, рассчитываемая остатком после того, как получены значения искомых параметров $\{a_i\}$. Однако в отличие, например, от метода наименьших квадратов, АМЛР

[4] Отметим, что модели, представленные данными в логарифмической шкале, не требуют предварительного преобразования переменных, поскольку оцениваемые коэффициенты изначально являются коэффициентами эластичности.

[5] Очевидно, что на основе описанной схемы сначала может быть рассчитан параметр b_2, а затем параметр b_1. По нашему представлению, последовательность оценки параметров рассматриваемой модели может определяться лишь спецификой каждой конкретной задачи, требующей использования АМЛР.

не позволяет произвести оценку точности модели, а также рассчитать стандартные отклонения оцененных структурных параметров. Априори ясно, что временной ряд $f(t)$ может гипотетически содержать как случайную компоненту, так и компоненту, обусловленную воздействием некоторых неизвестных факторов, не включенных в явном виде в модель. Из сказанного следует необходимость разработки математического метода, обеспечивающего представление временного ряда $f(t)$ в виде

$$f(t) = q(t) + \varepsilon(t), \qquad (23)$$

где $q(t)$ - регулярная составляющая; $\varepsilon(t)$ - нерегулярная (случайная) составляющая. При этом естественными требованиями для нерегулярной компоненты исследуемого временного ряда являются: 1) математическое ожидание $\varepsilon(t)$ равно нулю; 2) $\varepsilon(t)$ не коррелирована с $q(t)$. Требования к свойствам $q(t)$ не могут быть сформулированы сколько-нибудь конкретно, если исходить лишь из выражения (23), не принимая никаких дополнительных гипотез. Ясно лишь, что данная функция должна обладать свойством большей гладкости (в математическом смысле) в сравнении с $f(t)$[6].

Разделение (декомпозиция) временного ряда на две указанные выше составляющие - традиционная задача анализа временных рядов. Анализ временных рядов - направление теории математической статистики, сформировавшееся на рубеже XIX и XX веков. Соответственно и разработка методов решения данной задачи имеет весьма давнюю историю. Ни в коей мере не претендуя на исчерпывающую классификацию и подробное описание специфики предлагавшихся до настоящего времени подходов, отметим следующее.

1. Наиболее элементарный метод разделения исследуемого временного ряда в соответствии с (23) связан с оценкой регрессионной модели, в которой функция $q(t)$ задается в виде некоторой аналитической функции времени (например, полиномом определенной степени). Верификация такой модели методом наименьших квадратов (или какими-либо аналогичными методами) позволяет оценить $\varepsilon(t)$. Однако очевидно, что произвол в выборе аналитического вида функции $q(t)$ влечет неоднозначность оценки $\varepsilon(t)$.

2. Подавляющее большинство известных на сегодняшний день методов декомпозиции временных рядов (безразлично, какой природы данные представляют эти ряды) базируется либо на методе выделения регулярной компоненты (тренда) посредством сглаживания исходного временного ряда по методу скользящих средних, либо на обработке временного ряда при помощи аппарата авторегрессионных моделей, либо на определенной комбинации указанных методов.

Использование этих методов сопряжено с априорным заданием некоторых параметров (например, порядка авторегрессионного уравнения, порядка сглаженной средней и т. д.), что порождает неоднозначность результатов выделения функции ε_t.

В качестве недостатка методов выделения тренда, основывающихся на осреднении соседних значений временного ряда, следует отметить также, что процедура сглаживания может быть источником корреляционных связей между членами временного ряда (составленного из сглаженных данных), отсутствующих в первичных данных.

3. Можно констатировать многочисленность известных моделей декомпозиции временных рядов при отсутствии явно предпочтительной схемы.

Рассматриваемый далее подход к решению задачи декомпозиции временного ряда базируется на модели регрессии с переменными структурными параметрами [5]. А именно, рассмотрим регрессионную модель следующего вида:

$$f_t = q_t + \varepsilon_t; \qquad (24)$$
$$q_t - q_{t-1} = \delta_t \qquad (25)$$
$$t = 1,\ldots,T,$$

где $\{f_t\}$, $\{q_t\}$ - соответственно значения функций $f(t)$ и $q(t)$ в момент времени t; T - длина рассматриваемых временных рядов; $\{\varepsilon_t\}$ и $\{\varepsilon_t\}$ - две группы случайных отклонений, имеющих в общем случае неизвестные и различные дисперсии; при этом имеется T соотношений (24) и T-1 соотношений (25).

[6] В связи с этим в дальнейшем изложении термины «сглаженный ряд» и «тренд» употребляются как синонимы регулярной компоненты, то есть функции $q(t)$.

Совокупность соотношений (24) - (25) правомерно трактовать как регрессионную модель, в которой единственная объясняющая переменная тождественно равна единичному значению, на искомые структурные параметры $\{q_t\}$ которой наложены ограничения вида (25).

Использование метода наименьших квадратов для оценки параметров $\{q_t\}$ модели (24) - (25) приводит к задаче минимизации по параметрам $\{q_t\}$ выражения

$$\sum_{t=1}^{T}(f_t - q_t)^2 + \mu\sum_{t=1}^{T-1}(q_t - q_{t-1})^2, \qquad (26)$$

где $\mu = \sigma^2_\varepsilon / \sigma^2_\delta$, то есть соотношение дисперсий случайных отклонений $\{\varepsilon_t\}$ и $\{\delta_t\}$.

Перейдем к векторно-матричным обозначениям. Положим, что f, q - векторы-столбцы размерности T значений $\{f_t\}$ и $\{q_t\}$ соответственно; E - единичная матрица размерности $(T*T)$; D - матрица $[(T\text{-}1)*T]$, элементы $\{d_{ij}\}$ которой суть: $d_{ii} = -1$ и $d_{ii+1} = 1$ $(i = 1,\ldots,T\text{-}1)$, а остальные элементы являются нулевыми. Тогда решение задачи минимизации функционала (26) будет иметь следующий вид:

$$q = (E + \mu D'D)^{-1}f, \qquad (27)$$

где «$'$» - символ транспонирования. Величина μ должна быть задана экзогенно и является параметром метода.

Введем в рассмотрение также вектор-столбец погрешностей ε размерности T, элементы которого суть $\{\varepsilon_t\}$ из (24), так что $\varepsilon = f\text{-}q$.

Оцениватель (27) обладает следующими свойствами:

1. При $\mu=0$ $q=f$, то есть фактические и сглаженные (в традиционном определении математической статистики) значения анализируемого временного ряда совпадают; при $\mu\to\infty$ элементы вектора q - среднее значение временного ряда f.

2. При заданном значении μ минимизация (26) означает, что полученный в соответствии с (27) сглаженный временной ряд q аппроксимирует фактические значения исследуемого временного ряда с некоторой погрешностью ζ, зависящей от μ, то есть

$$\sum_{t=1}^{T}(f_t - q_t)^2 = \zeta.$$

Следовательно, сглаженный временной ряд q является наиболее гладкой (в евклидовой метрике) кривой (или кривой наименьшей длины), аппроксимирующей фактические значения исследуемого временного ряда. Действительно, задача минимизации (26) может быть интерпретирована как задача Лагранжа нахождения условного минимума функционала

$$\sum_{t=1}^{T-1}(q_t - q_{t-1})^2$$

при ограничении вида

$$\sum_{t=1}^{T}(f_t - q_t)^2 = \zeta, \qquad (28)$$

где ζ - заранее известный уровень погрешности аппроксимации исследуемого временного ряда, а величина $(1/\mu)$ - множитель Лагранжа для ограничения (28).

3. При любом заданном значении μ среднее значение элементов вектора погрешностей ε - нулевое значение. Из определения вектора сглаженных значений исследуемого временного ряда по формуле (27) следует, что

$$I'(E + \mu D'D)\, q = I'q + \mu I'D'Dq = I'f,$$

где I - единичный вектор-столбец размерности T. В свою очередь, каждый столбец матрицы D' содержит лишь два ненулевых элемента: +1 и −1. Это означает, что произведение $I'D'$ - нулевой вектор-строка, откуда

$$I'q = I'f.$$

Из данного выражения следует, что средние значения исходного и сглаженного временных рядов совпадают, то есть среднее значение элементов вектора ε равно нулю.

4. Выбор того или иного возможного значения μ не связан с характером и масштабом данных, составляющих анализируемый временной ряд, поскольку значения элементов матриц, входящих в выражение (27), представлены лишь значениями +1, -1 и 0.

5. Значения $\{q_t\}$ и $\{\varepsilon_t\}$ в общем случае являются коррелированными. Для доказательства данного утверждения рассмотрим скалярное произведение $\varepsilon' q$, которое с учетом вышеприведенных соотношений есть

$$\varepsilon' q = (f\text{-}q)' q = f' q - q' q =$$
$$= f' (E+\mu D' D)^{-1} f - f' (E+\mu D' D)^{-1} (E+\mu D' D)^{-1} f =$$
$$= f' \{(E+\mu D' D)^{-1} [E - (E+\mu D' D)^{-1}]\} f, \quad (29)$$

то есть представляет собой квадратичную форму относительно вектора f. Аналитическое исследование матрицы из выражения (29) затруднительно, однако результаты численных экспериментов показывают, что для фиксированного значения T при любом конечном значении μ симметрическая матрица $(E+\mu D' D)^{-1} [E - (E + +\mu D' D)^{-1}]$ является положительно определенной. Таким образом, векторы ε и q не являются ортогональными.

Суммируя сказанное выше, правомерно заключить, что, с одной стороны, метод расчета теоретического (сглаженного) ряда q в соответствии с (27) обладает очевидными достоинствами: во-первых, данный метод не требует задания аналитического вида временной функции для оценки регулярной составляющей исследуемого временного ряда; во-вторых, оценка регулярной компоненты является в математическом смысле наиболее гладкой кривой из всех возможных при заданном уровне погрешности сглаживания (свойство 2). С другой стороны, рассматриваемый метод декомпозиции временного ряда не обеспечивает независимости регулярной и стохастической компонент; также остается открытым вопрос о правиле выбора наиболее подходящего значения параметра μ.

Проведенный нами анализ перечисленных специфических свойств рассматриваемого здесь метода оценивания в совокупности с результатами численных экспериментов позволил разработать следующую его (метода) модификацию, дающую возможность как решить проблему обеспечения некоррелированности ε и q, так и обосновать метод выбора значения параметра μ.

Более конкретно результаты проведенных численных экспериментов показывают, что степень корреляции ε и q изменяется неравномерно по мере уменьшения значения параметра μ от больших значений к меньшим; при $\mu \to 0$ уровень коррелированности ε и q постепенно нарастает, стабилизируясь начиная с некоторого значения данного параметра. Последнее свойство позволяет следующим образом скорректировать изложенную ранее процедуру получения оценок регулярной и случайной компонент временного ряда:

1) для каждого заранее заданного значения μ производится расчет $q(\mu)$ в соответствии с (6) и далее рассчитывается $\varepsilon(\mu) = f - q(\mu)$;

2) производится расчет линейной регрессии вида

$$q(\mu) = b_\mu \, \varepsilon(\mu) + b_0 + u'(\mu), \quad (30)$$

где b_0, b_μ - параметры регрессионного уравнения, подлежащие оцениванию; $u'(\mu)$ - составляющая временного ряда $q(\mu)$, некоррелированная с $\varepsilon(\mu)$;

3) производится вычисление скорректированного вектора регулярной компоненты q^* исследуемого временного ряда в виде

$$q^* = q(\mu) - b_\mu \, \varepsilon(\mu). \quad (31)$$

Вектор q^*, начиная с некоторого значения μ, принимается в качестве окончательной оценки регулярной компоненты анализируемого исходного временного ряда.

Иными словами, специфическая особенность вектора $\varepsilon(\mu)$, получаемого в результате использования формулы (27), заключается в том, что при $\mu \to 0$ структура этого вектора остается стабильной (при абсолютном уменьшении составляющих его элементов $\{\varepsilon_t(\mu)\}$). В результате оцениваемый в соответствии с (30) регрессионный параметр b_μ увеличивается во столько же раз, во сколько раз снижаются значения $\{\varepsilon_t(\mu)\}$, так что значение элементов вектора q^*, рассчитываемого в соответствии с (31), остается неизменным. Последнее, в свою очередь, означает, что существует некоторая область значений параметра μ, в пределах которой вектор q^* оказывается стабильным. Тем самым решается проблема выбора «оптимального» значения μ.

Использование указаанного метода для декомпозиции временных рядов проиллюстрируем на данных отечественной статистики.

В [3] приведены, в частности, результаты оценивания при помощи АМЛР линейно однородной производственной функции для промышленности России:

$$y_t - l_t = \alpha(k_t - l_t) + f(t),$$

где y_t, k_t, l_t - темпы изменения (разности натуральных логарифмов) выпуска и производственных ресурсов основного капитала и труда (численности занятых) в году t; α - коэффициент эластичности производительности труда по капиталовооруженности. По результатам проведенных расчетов среднее значение α для периода 1993-2012 гг. составило 0,31. Погодовые значения $\{f_t\}$ функции-остатка $f(t) = y_t - l_t - 0{,}31(k_t - l_t)$, а также оценки $\{q^*_t\}$ и $\{\varepsilon^*_t\}$, исчисленные в соответствии с (30) - (31), приведены в таблице 1.

Проведенные расчеты показали, что динамика регулярной компоненты q^* объясняет 69,5% дисперсии исходного временного ряда f, а динамика случайной компоненты ε^* - соответственно 30,5%.

Таблица 1

Результаты декомпозиции исследуемого временного ряда

Год	f_t	q^*_t	ε^*_t	Год	f_t	q^*_t	ε^*_t
1993	-0,1558	-0,1942	0,0384	2003	0,0580	0,0454	0,0125
1994	-0,2601	-0,1440	-0,1160	2004	0,0514	0,0532	-0,0018
1995	-0,0494	-0,1288	0,0794	2005	0,0497	0,0545	-0,0048
1996	-0,0543	-0,0291	-0,0252	2006	0,0611	0,0586	0,0025
1997	0,0092	-0,0353	0,0445	2007	0,0658	0,0420	0,0238
1998	-0,0481	0,0258	-0,0739	2008	0,0060	-0,0101	0,0161
1999	0,0951	0,0512	0,0440	2009	-0,0976	0,0025	-0,1001
2000	0,1190	0,0853	0,0337	2010	0,0705	0,0006	0,0699
2001	0,0514	0,0686	-0,0172	2011	0,0488	0,0511	-0,0023
2002	0,0305	0,0483	-0,0178	2012	0,0334	0,0391	-0,0057

Таблица 2

Статистические характеристики уравнения (30) при различных значениях параметра μ

$\mu^{1/2}$	$\sigma_\varepsilon(\mu)$	$b_\mu(\mu)$	$R^2(\mu)$
100	0,08952	0,001304	0,430
10	0,08106	0,098382	0,371
5	0,06913	0,243789	0,307
2,5	0,05380	0,443746	0,222
1	0,03491	0,761214	0,142
0,8	0,02990	0,955874	0,150
0,5	0,01902	1,841713	0,194
0,25	0,00701	6,265127	0,261
0,1	0,00130	37,2170	0,297
0,01	1,3426E-05	3684,68	0,305
0,001	1,3430E-07	368430,9	0,305
0,0001	1,3430E-09	36843052	0,305
0,00001	1,3430E-11	3684306180	0,305

В таблице 2 представлены значения среднеквадратических отклонений $\sigma_\varepsilon(\mu)$, значения коэффициента регрессии $b_\mu(\mu)$ и коэффициента детерминации $R^2(\mu)$ для уравнения (30), соответствующие различным значениям μ. Как можно видеть, $R^2(\mu)$ сначала снижается по мере уменьшения μ от очень больших величин (соответствующих практически постоянным величинам $\{b_t\}$) до μ = 1. При дальнейшем уменьшении μ $R^2(\mu)$ возрастает, и начиная с μ=0,01, он остается практически стабильным при том, что значения $\sigma_\varepsilon(\mu)$ стремятся к нулю, а значения $b_\mu(\mu)$ растут.

Из приведенных данных следует, что на интервале возможных значений параметра метода μ степень коррелированности ε(μ) и q(μ) изменяется неравномерно. Имеется область значений μ, для которой корреляция ε(μ) и q(μ) может считаться незначимой. Однако в области минимальной корреляции ε(μ) и q(μ) динамика регулярной компоненты объясняет лишь незначительную часть вариации исследуемых временных рядов. В области μ ≤ 0,01 свойства регулярной компоненты q^* для рассматриваемых временных рядов аналогичны; при этом регулярная компонента q^* вносит существенно больший вклад в вариацию исходных временных рядов в сравнении с q(μ), когда значение μ таково, что минимизирует корреляцию ε(μ) и q(μ).

Следовательно, использование специфических свойств оценивателя (27) при $\mu \to 0$ в совокупности с последующей корректировкой сглаженного ряда $q(\mu)$, приводящей к вычислению регулярной компоненты q^* по формуле (31), обеспечивает решение задачи декомпозиции исследуемого временного ряда f таким образом, что, с одной стороны, q^* обладает наиболее высоким уровнем корреляции с f, а с другой - компоненты временных рядов ε^* и q^* являются некоррелированными.

* *
*

В *заключение* отметим следующее.

1. В плане используемого математического аппарата АМЛР, безусловно, имеет значительную аналогию с методами обработки данных, развитыми ранее как в математической, так и экономической статистике. При этом следует прежде всего упомянуть различные модификации методов факторного анализа, и прежде всего метод главных компонент как средство анализа статистических данных самой разнообразной природы (см., в частности, [4, 6, 7]).

Преобразование исходных статистических данных, использованное нами, применяется в многомерном статистическом анализе при агрегировании частных индексов (или шкал) для получения сводного (обобщающего) индекса.

В то же время разработанный метод принципиально отличается от метода ортогональной регрессии, давно известного как в трудах по математической статистике, так и в эконометрии и также основывающегося на методе главных компонент (см., например, [8]).

Как известно, в традиционной регрессионной модели типа (1) нахождение искомых структурных параметров модели основывается на минимизации суммы квадратов отклонений фактических и теоретических значений объясняемой переменной. В модели ортогональной регрессии используется иной критерий определения параметров. Например, процедура определения структурных параметров модели

$$y_t = ax_t + b$$

в соответствии с методом ортогональной регрессии сводится, во-первых, к стандартизации входящих в модель переменных [в самом простом случае - центрированию, то есть переходу к величинам $\{(y_t - \bar{y}), (x_t - \bar{x})\}$, где $\bar{y}$, $\bar{x}$ - средние соответствующих рядов исходных данных, с тем чтобы исключить из оцениваемой модели свободный член], и во-вторых, к минимизации суммы квадратов отклонений $\{\varepsilon_t\}$

$$\varepsilon_t = a_1(y_t - \bar{y}) + a_2(x_t - \bar{x}),$$

при условии $a^2_1 + a^2_2 = 1$. В этом случае в качестве искомых коэффициентов a_1, a_2 должны быть приняты элементы собственного вектора матрицы, аналогичной матрице $J'J$ из (4), соответствующего второму, то есть минимальному собственному значению. В АМЛР, напротив, определение искомых параметров статистической модели основывается на определении первого собственного вектора матрицы $J'J$, что и определяет специфику разработанного метода обработки данных.

2. Метод разделения функции-остатка $f(t)$ для модели, оцениваемой на основе АМЛР, на регулярную и случайную компоненты, описанный в данной статье, позволяет оценить точность модели, а также выявить вклад в динамику объясняемой переменной факторов, в явном виде не идентифицированных в рамках исследуемой модели. Данный результат, как нам представляется, имеет чрезвычайно важное значение для практики эконометрического моделирования.

Предлагаемый метод декомпозиции временного ряда на регулярную и нерегулярную (стохастическую) составляющие также позволяет подойти к решению следующей задачи анализа математических моделей, представленных временными рядами включаемых в эти модели переменных. Пусть ставится задача построения статистической модели

$$y_t = g(x_{1t},..,x_{mt}) + \varepsilon_t, \tag{32}$$

где y_t - объясняемая переменная; $\{x_{it}\}$ - объясняющие переменные модели (32); ε_t - случайная компонента, отражающая точность аппроксимации функцией $g(x_{1t},..,x_{mt})$ объясняемой переменной. Набор объясняющих переменных может быть различен для различных вариантов спецификации модели (32), что, вообще говоря, порождает различные представления о характере и уровне $\{\varepsilon_t\}$.

Предварительное знание об уровне погрешности и виде регулярной компоненты, а также их вкладе в совокупную вариацию исследуемого временного ряда, обеспечиваемое расчетом ε^* и q^* в соответствии с изложенным ранее методом, позволяет составить представление о потенциальном пределе точности модели (32) еще до задания конкретного варианта спецификации уравнения (32).

В частности, результаты декомпозиции временного ряда темпов изменения сельскохозяйственного производства за период 1993-2012 гг. при помощи описанного здесь метода показывают, что динамика регулярной компоненты q^* объясняет 37% дисперсии исходного временного ряда, а динамика случайной компоненты ε^* - соответственно 63%. Таким образом, не более чем 40% вариации анализируемого временного ряда могут быть объяснены динамикой некоторых (не идентифицированных заранее) систематических факторов (это может быть прежде всего динамика производственного потенциала); большая же часть вариации должна быть приписана воздействию случайных (например, погодных) факторов.

В прикладном плане следует отметить простоту численной реализации описанного выше вычислительного метода декомпозиции. Его практическое использование доступно, например, в среде табличного процессора Excel, то есть не требует разработки специализированных программных продуктов.

Литература

1. **Джонстон Дж.** Эконометрические методы. М.: Наука, 1980.

2. **Суворов Н.В.** Актуальные направления и проблемы совершенствования модельного инструментария макроэкономического анализа // Проблемы прогнозирования. 2015. № 5. С. 25-39.

3. **Суворов А.В., Суворов Н.В., Балашова Е.Е.** и др. Человеческий капитал как фактор социально-экономического развития России. СПб: Нестор-История, 2016.

4. **Аллен Р.** Экономические индексы. М.: Финансы и статистика, 1980.

5. **Суворов Н.В.** Метод построения регрессионных моделей с динамическими структурными параметрами // Проблемы прогнозирования. 2005. № 4. С. 143-153.

6. **Лоули Д., Максвелл А.** Факторный анализ как статистический метод. М.: Мир, 1967.

7. **Айвазян С.А., Енюков И.С., Мешалкин Л.Д.** Прикладная статистика. Т. 2. Исследование зависимостей. М.: Финансы и статистика, 1985.

8. **Маленво Э.** Статистические методы эконометрии. М.: Статистика, 1976.

VERIFICATION OF AN ECONOMETRIC MODEL BASED ON A PRIORI CONSTRAINTS ON THE STRUCTURAL PARAMETERS*

Nikolai V. Suvorov

Author affiliation: Russian Academy of Sciences (Moscow, Russia). E-mail: suvor_n@ecfor.ru.

The article describes a method for verification of a statistical model, which, firstly, is represented by the time series of original data and, secondly, is linear in the estimated parameters. Experience in statistical calculations on real empirical data shows that the most well-known and conventionally used in the econometric modeling of mathematical-statistical methods (least squares, maximum likelihood method, and similar methods) often do not ensure successful verification of theoretically required forms of econometric models. The developed method which is called an alternative method of linear regression (AMLR) provides an account of a priori restrictions on the absolute values and signs of the parameters identified by the model. The AMLR based on the concept of best linear index, is known in the theory of statistics from the end of the 1950s. Mathematically AMLR it based on the method of principal components. The article analyzes conditions for applying the AMLR in econometric modeling and methods of transformation of the initial statistical information to ensure correct application of the developed evaluation procedures.

Special problems of the proposed method are to determine the level of accuracy of approximation of the dependent variable of the model. In this regard, to assess the level of precision of the statistical model verifiable by using the AMLR, was developed an original method of decomposition of the time series on the regular and stochastic components. The author analyzes the properties of the proposed method of decomposition and gave a numerical illustration of its use in econometric calculations.

Keywords: statistical model, time series, best linear index, decomposition of the time series.
JEL: C01, C51.

* This article was prepared using the materials developed with the financial support of the Russian Foundation for Basic Research (Project № 16-06-00319).

References

1. **Johnston J.** *Econometric methods.* 2-nd ed. New York, McGraw-Hill Book Company, 1972 (Russ. ed.: Dzhonston Dzh. *Ekonometricheskie metody*. Moscow, Nauka Publ., 1980).

2. **Suvorov N.V.** Aktual'nye napravleniya i problemy sovershenstvovaniya model'nogo instrumentariya makroekonomicheskogo analiza [Urgent directions and problems of improvement of model tools of the macroeconomic analysis]. *Problemy prognozirovaniya*, 2015, no. 5, pp. 25-39. (In Russ.).

3. **Suvorov A.V., Suvorov N.V., Balashova E.E.** et al. *Chelovecheskii kapital kak faktor sotsial'no-ekonomicheskogo razvitiya Rossii* [Human capital as the factor of the social and economic development of Russia]. St. Petersburg, Nestor-history Publ., 2016. (In Russ.).

4. **Allen R.** *Index numbers in theory and practice*. Chicago: Aldine Publ., 1975 (Russ. ed.: Allen R. *Ekonomicheskie indeksy*. Moscow, Finansy i statistika Publ., 1980).

5. **Suvorov N.V.** Metod postroeniya regressionnykh modelei s dinamicheskimi strukturnymi parametrami [The method of constructing regression models with dynamic structural parameters]. *Problemy prognozirovaniya*, 2005, no. 4, pp. 143-153. (In Russ.).

6. **Lawley D.N., Maxwell A.E.** *Factor analysis as a statistical method*. London, Butterworths, 1963 (Russ. ed.: Louli D., Maksvell A. *Faktornyi analiz kak statisticheskii metod*. Moscow, Mir Publ., 1967).

7. **Aivazyan S.A., Enyukov I.S., Meshalkin L.D.** *Prikladnaya statistika. T. 2. Issledovanie zavisimostei* [Applied statistics. Vol. 2. Study of dependences]. Moscow, Finansy i statistika Publ., 1985. (In Russ.).

8. **Malinvaud E.** *Méthodes statistiques de l'économétrie*. Deuxième ed. Paris, 1969 (Russ. ed.: Malenvo E. *Statisticheskie metody ekonometrii*. Moscow, Statistika Publ., 1976).