Оценивание структурных эконометрических уравнений*

Стефен Поллок†

Колледж королевы Мэри Лондонского университета, Великобритания

Настоящая работа содержит отличающиеся от традиционных выводы ММПОИ- и 2ШМНК-оценок одиночных уравнений классической системы одновременных эконометрических уравнений. Принадлежность обеих к оценкам метода моментов подчеркивает их глубинные сходства. Оценка ММПОИ выводится из критерия наименьших квадратов, использующего интерпретацию структурного уравнения как модели с ошибками в переменных, а 2ШМНК-оценка получена при помощи асимптотически правомерного приближения. Оценку ММПОИ можно вычислить с помощью итерационного алгоритма, стартующего с оценки 2ШМНК. В работе также рассматриваются традиционные выводы 2ШМНК-оценки.

1 Введение

Задача оценивания одиночного уравнения классической эконометрической системы одновременных уравнений была решена исследователями из *Cowles Commission*, разработавшими метод максимального правдоподобия с ограниченной информацией (ММПОИ). Они представили два способа вывода этой оценки.

Первый вывод был произведен Anderson & Rubin (1949) на основе функции правдоподобия для модели одновременных уравнений в приведенной форме. Добавив в нее информацию, связанную с одиночным структурным уравнением, они нашли оценки параметров приведенной формы при ограничениях вместе с оценками параметров структурного уравнения.

В альтернативном выводе Koopmans, Rubin & Leipnick (1950) за отправную точку взяли функцию правдоподобия для структурных параметров полной системы. Они вывели целевую функцию для оценивания интересующего их уравнения путем удаления посторонних параметров через частичную максимизацию.

Эти два способа вывода ММПОИ должны были получить признание среди прикладных эконометристов. Однако обнаружились препятствия, помешавшие всеобщему принятию этой оценки. Первоначальные способы ее получения были слишком длинными и сложными, и немногие были способны овладеть ими. К тому же вычисление оценок подразумевало итерационную процедуру получения скрытых корней, с которой существовавшие на тот момент компьютеры едва справлялись.

Обе задачи были решены несколько лет спустя с помощью оценки двухшагового метода наименьших квадратов (2ШМНК), независимо полученной в Basmann (1957) и Theil (1958) простыми и понятными способами. Выводы обеих оценок производились в рамках всем знакомой классической модели линейной регрессии. Они строились в попытках найти простые пути преодоления той проблемы, из-за которой МНК в обычной линейной регрессии неверно оценивает структурные параметры.

Позже Theil (1961) сумел продемонстрировать сходство 2ШМНК- и ММПОИ-оценок, показав, что они обе являются элементами определенного им «k-класса оценок». Другие авторитетные ученые, включая Malinvaud (1966), также смогли показать это сходство. Несмотря

*Перевод А. Шамгунова и С. Анатольева. Цитировать как: Поллок, Стефен (2007) «Оценивание структурных эконометрических уравнений», Квантиль, №2, стр. 49–59. Citation: Pollock, Stephen (2007) “Estimation of Structural Econometric Equations,” Quantile, No.2, pp. 49–59.

†Адрес: Department of Economics, Queen Mary College, University of London, Mile End Road, London E1 4NS, United Kingdom. Электронная почта: *d.s.g.pollock@qmul.ac.uk*

на это, ММПОИ-оценка не была широко признана и часто представлялась без соответствующего вывода.

Недавно в работе, касающейся происхождения оценок 2ШМНК и ММПОИ, Anderson (2005) поведал, как Anderson & Rubin (1950) получили асимптотическое распределение ММПОИ-оценки путем нахождения асимптотического распределения статистики, по своей сути являвшейся оценкой 2ШМНК. Несмотря на то, что эта работа дает краткий отчет о первом выводе Anderson & Rubin (1949), она не проводит прямой параллели между обеими оценками.

Остается желание продемонстрировать близкую сущность этих оценок, что подразумевает прямой вывод обеих. Это и является целью данной работы.

2 Структурные уравнения

Классическая модель линейных одновременных эконометрических уравнений – это стохастическая система, связывающая M выходных эндогенных переменных с K входными экзогенными переменными. Особенностью этой модели является то, что каждая зависимая переменная вектора-строки $y_{t\bullet} = [y_{t1}, y_{t2}, \dots, y_{tM}]$ является функцией не только K экзогенных переменных вектора $x_{t\bullet} = [x_{t1}, x_{t2}, \dots, x_{tK}]$, но и некоторых других переменных из $y_{t\bullet}$.

Эту особенность можно интерпретировать как наличие постоянной обратной связи между выходом системы и входными данными. Поэтому j-е структурное уравнение, выражающее y_{tj} через элементы из $x_{t\bullet}$ и остальные переменные из $y_{t\bullet}$, можно записать как

$$y_{tj} = y_{t\bullet}c_{\bullet j} + x_{t\bullet}\beta_{\bullet j} + \varepsilon_{tj}, \tag{1}$$

где $c_{\bullet j}$ и $\beta_{\bullet j}$ являются векторами параметров этой системы. Также подразумевается, что $c_{jj} = 0$ с целью предотвратить появление y_{tj} и в левой, и в правой части. Уравнение также содержит случайное возмущение ε_{tj}.

Другим способом записи структурного уравнения, который ставит y_{tj} по соседству с другими экзогенными переменными расширенной системы, является выражение

$$y_{t\bullet}\gamma_{\bullet j} + x_{t\bullet}\beta_{\bullet j} + \varepsilon_{tj} = 0. \tag{2}$$

Таким образом, $\gamma_{\bullet j}$ и $c_{\bullet j}$ отличаются только своими j-ми элементами, равными $\gamma_{jj} = -1$ и $c_{jj} = 0$ соответственно. Условие $\gamma_{jj} = -1$, идентифицирующее зависимую переменную структурного уравнения, называется нормирующим правилом.

M структурных уравнений, собранные воедино, составляют следующую систему:

$$[y_{t1}, y_{t2}, \dots, y_{tM}] = y_{t\bullet}[c_{\bullet 1}, c_{\bullet 2}, \dots, c_{\bullet M}] + x_{t\bullet}[\beta_{\bullet 1}, \beta_{\bullet 2}, \dots, \beta_{\bullet M}] + [\varepsilon_{t1}, \varepsilon_{t2}, \dots, \varepsilon_{tM}], \tag{3}$$

которая в компактном виде выглядит следующим образом:

$$y_{t\bullet} = y_{t\bullet}C + x_{t\bullet}B + \varepsilon_{t\bullet} \tag{4}$$

Также ее можно записать как

$$y_{t\bullet}\Gamma + x_{t\bullet}B + \varepsilon_{t\bullet} = 0, \tag{5}$$

где $\Gamma = [\gamma_{\bullet 1}, \gamma_{\bullet 2}, \dots, \gamma_{\bullet j}]$.

3 Приведенная форма

Если мы готовы пренебречь деталями структуры эконометрической модели, то можно выразить каждую переменную выхода из $y_{t\bullet} = [y_{t1}, y_{t2}, \dots, y_{tM}]$, используя только экзогенные переменные. Такое представление называется приведенной формой модели. Приведенная форма получена из (5) умножением справа на матрицу, обратную к Γ, описывающую мгновенную обратную связь между переменными. Таким образом, получаем

$$y_{t\bullet} = x_{t\bullet}\Pi + \eta_{t\bullet}, \quad \text{где } \Pi = -B\Gamma^{-1} \text{ и } \eta_{t\bullet} = -\varepsilon_{t\bullet}\Gamma^{-1}. \tag{6}$$

Теперь следует сделать предположения о случайных элементах модели. Предположим, что элементы вектора $\varepsilon_{t\bullet} = [\varepsilon_{t1}, \varepsilon_{t2}, \ldots, \varepsilon_{tM}]$, являющиеся M структурными ошибками, распределены независимо по времени так, что для любого t выполняется

$$\mathbb{E}[\varepsilon_{t\bullet}] = 0 \quad \text{и} \quad \mathbb{V}[\varepsilon_{t\bullet}] = \mathbb{E}[\varepsilon'_{t\bullet}\varepsilon_{t\bullet}] = \Sigma_{\varepsilon\varepsilon}. \tag{7}$$

Также предполагается, что структурные ошибки распределены независимо от экзогенных переменных, так что $\mathbb{C}[\varepsilon_{t\bullet}, x_{s\bullet}] = 0$ для любых t и s.

Отсюда следует, что для вектора шоков приведенной формы $\eta_{t\bullet} = -\varepsilon_{t\bullet}\Gamma^{-1}$ выполнено

$$\mathbb{E}[\eta_{t\bullet}] = 0 \quad \text{и} \quad \mathbb{V}[\eta_{t\bullet}] = \Gamma'^{-1}\mathbb{V}[\varepsilon_{t\bullet}]\Gamma^{-1} = \Gamma'^{-1}\Sigma_{\varepsilon\varepsilon}\Gamma^{-1} = \Omega. \tag{8}$$

Преобразованные ошибки также независимы от $x_{t\bullet}$, откуда следует условие $\mathbb{C}[\eta_{t\bullet}, x_{s\bullet}] = 0$ для любых t и s.

4 Проблема идентификации и структурная модель

Структурная модель одновременных уравнений подвержена так называемой проблеме идентификации, ограничивающей возможности оценивания структурных параметров. Имея в наличии достаточный набор наблюдений, мы всегда можем оценить параметры статистической связи между эндогенными переменными из $y_{t\bullet}$ и экзогенными из $x_{t\bullet}$ в приведенной форме. Однако если мы собираемся оценить параметры структурных связей, то необходимо иметь априорную информацию о структуре модели.

Предположим, что статистические свойства данных можно полностью описать с помощью первых и вторых моментов. Обозначим дисперсионные матрицы $x_{t\bullet}$ и $y_{t\bullet}$ через $\mathbb{V}[x_{t\bullet}] = \Sigma_{xx}$ и $\mathbb{V}[y_{t\bullet}] = \Sigma_{yy}$, а матрицу их ковариаций через $\mathbb{C}[x_{t\bullet}, y_{t\bullet}] = \Sigma_{xy}$. Совмещая приведенную форму регрессионного отношения (6) с тривиальным равенством по $x_{t\bullet}$, получим следующую систему:

$$\begin{bmatrix} y_{t\bullet} & x_{t\bullet} \end{bmatrix} \begin{bmatrix} I & 0 \\ -\Pi & I \end{bmatrix} = \begin{bmatrix} \eta_{t\bullet} & x_{t\bullet} \end{bmatrix}. \tag{9}$$

Полагая $\mathbb{V}[\eta_{t\bullet}] = \Omega$ и $\mathbb{C}[\eta_{t\bullet}, x_{t\bullet}] = 0$, получим

$$\begin{bmatrix} I & -\Pi' \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \begin{bmatrix} I & 0 \\ -\Pi & I \end{bmatrix} = \begin{bmatrix} \Omega & 0 \\ 0 & \Sigma_{xx} \end{bmatrix}. \tag{10}$$

Умножая слева эту систему на матрицу, обратную к самой левой, получаем эквивалентное уравнение в виде

$$\begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \begin{bmatrix} I & 0 \\ -\Pi & I \end{bmatrix} = \begin{bmatrix} I & \Pi' \\ 0 & I \end{bmatrix} \begin{bmatrix} \Omega & 0 \\ 0 & \Sigma_{xx} \end{bmatrix} = \begin{bmatrix} \Omega & \Pi'\Sigma_{xx} \\ 0 & \Sigma_{xx} \end{bmatrix}. \tag{11}$$

Из этой системы можно выделить уравнения $\Sigma_{yy} - \Sigma_{yx}\Pi = \Omega$ и $\Sigma_{xy} - \Sigma_{xx}\Pi = 0$, из которых получаем параметры, характеризующие приведенную форму связей:

$$\Pi = \Sigma_{xx}^{-1}\Sigma_{xy} \quad \text{и} \quad \Omega = \Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}. \tag{12}$$

Эти параметры могут быть оценены с помощью эмпирических аналогов матриц моментов Σ_{xx}, Σ_{yy} и Σ_{xy}, доступных в виде $M_{xx} = T^{-1}\sum_t x'_{t\bullet}x_{t\bullet}$, $M_{yy} = T^{-1}\sum_t y'_{t\bullet}y_{t\bullet}$ и $M_{xy} = T^{-1}\sum_t x'_{t\bullet}y_{t\bullet}$.

Теперь скомбинируем структурное уравнение (5) с тривиальным равенством, получая при этом аналог соотношения (9):

$$\begin{bmatrix} y_{t\bullet} & x_{t\bullet} \end{bmatrix} \begin{bmatrix} \Gamma & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} \varepsilon_{t\bullet} & x_{t\bullet} \end{bmatrix}. \tag{13}$$

Используя $\mathbb{V}[\varepsilon]=\Sigma_{\varepsilon\varepsilon}$ и $\mathbb{C}[\varepsilon,x]=0$, получаем

$$\begin{bmatrix} \Gamma' & B' \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \begin{bmatrix} \Gamma & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} \Sigma_{\varepsilon\varepsilon} & 0 \\ 0 & \Sigma_{xx} \end{bmatrix}, \tag{14}$$

и, действуя по аналогии с (10)–(11), приходим к эквивалентному выражению

$$\begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \begin{bmatrix} \Gamma & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} \Gamma'^{-1} & \Pi' \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{\varepsilon\varepsilon} & 0 \\ 0 & \Sigma_{xx} \end{bmatrix} = \begin{bmatrix} \Omega\Gamma & \Pi'\Sigma_{xx} \\ 0 & \Sigma_{xx} \end{bmatrix}. \tag{15}$$

Из этого равенства получаем фундаментальные уравнения, связывающие структурные параметры Γ и B с матрицами моментов переменных. Это уравнения можно записать в двух альтернативных формах:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \Sigma_{yy}-\Omega & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \begin{bmatrix} \Gamma \\ B \end{bmatrix} = \begin{bmatrix} \Pi'\Sigma_{xy} & \Pi'\Sigma_{xx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \begin{bmatrix} \Gamma \\ B \end{bmatrix}. \tag{16}$$

Первое уравнение прямо следует из (15). Второе следует из равенств $\Sigma_{yy}=\Pi'\Sigma_{xx}\Pi+\Omega$ и $\Sigma_{xy}=\Sigma_{xx}\Pi$, полученных из (12). Действительно, заменяя Π' на $\Sigma_{yx}\Sigma_{xx}^{-1}$, мы можем выразить матрицу второго уравнения, используя только моменты переменных.

Примем уравнение (16) за основу, на базе которой мы можем оценить значения структурных параметров Γ и B. Очевидно, что в таком случае система содержит недостаточно информации для оценивания. В частности, составляющее ее уравнение $\Pi'\Sigma_{xy}\Gamma+\Pi'\Sigma_{xx}B=0$ получено преобразованием сопутствующего уравнения $\Sigma_{xy}\Gamma+\Sigma_{xx}B=0$, и поэтому не содержит дополнительной информации. На самом деле, если на матрицы моментов не наложены ограничения, кроме естественных условий симметричности и положительной определенности, то количество неизвестных параметров, которые можно вывести из уравнения (16), не может превышать MK, что равно количеству параметров матрицы Π в приведенной форме.

Теоретически, априорная информация о Γ и B может принимать разные формы. На практике обычно рассматриваются только линейные ограничения на параметры, часто являющиеся правилами нормализации, устанавливающими диагональные элементы Γ равными -1, и исключающими ограничениями, приравнивающими некоторые элементы Γ и B нулю. Если ни одно из ограничений не задействует более одного уравнения, то есть возможность рассматривать каждое уравнение по отдельности.

Если ограничения на параметры j-го уравнения принимают форму исключающих ограничений или правил нормализации, то их можно представить в виде уравнения

$$\begin{bmatrix} R'_\diamond & 0 \\ 0 & R'_* \end{bmatrix} \begin{bmatrix} \gamma_{\bullet j} \\ \beta_{\bullet j} \end{bmatrix} = \begin{bmatrix} r_j \\ 0 \end{bmatrix} \quad \text{или} \quad \begin{bmatrix} R'_\diamond & 0 \\ 0 & R'_* \end{bmatrix} \begin{bmatrix} \gamma_{\bullet j}+e_j \\ \beta_{\bullet j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{17}$$

где R_* содержит набор столбцов единичной матрицы I_K порядка K, $R_\diamond$, аналогично, состоит из набора столбцов единичной матрицы I_M порядка M, а r_j является вектором, состоящим из нулей и -1, согласно правилам нормализации. Вектор e_j есть j-й столбец I_M, единица из которого сокращает нормированный элемент $\gamma_{\bullet j}$.

Мы можем представить все эти ограничения в компактной форме:

$$\begin{bmatrix} \gamma_{\bullet j} \\ \beta_{\bullet j} \end{bmatrix} = \begin{bmatrix} S_\diamond & 0 \\ 0 & S_* \end{bmatrix} \begin{bmatrix} \gamma_{\diamond j} \\ \beta_{* j} \end{bmatrix} - \begin{bmatrix} e_j \\ 0 \end{bmatrix}, \tag{18}$$

где $\gamma_{\diamond j}$ и β_{*j} состоят из M_j и K_j не связанных ограничениями элементов $\gamma_{\bullet j}$ и $\beta_{\bullet j}$, и где $S_\diamond$ и S_* – аналоги $R_\diamond$ и R_*, состоящие из столбцов I_M и I_K соответственно.

При подстановке решения (18) в уравнение $\Sigma_{xy}\gamma_{\bullet j}+\Sigma_{xx}\beta_{\bullet j}=0$, являющееся j-м уравнением системы (16), получим

$$\Sigma_{xy}S_\diamond\gamma_{\diamond j}+\Sigma_{xx}S_*\beta_{*j}=\Sigma_{xy}e_j. \tag{19}$$

Это набор K уравнений с M_j+K_j неизвестными; и, при условии, что матрица $[\Sigma_{xy}, \Sigma_{xx}]$ имеет полный ранг, необходимым и достаточным условием для идентифицируемости параметров j-го уравнения является $K \geq M_j + K_j$.

Если это условие выполнено, то любого подмножества уравнений (19) размера $M_j + K_j$ будет достаточно, чтобы определить $\gamma_{\diamond j}$ и β_{*j}. Однако нам особо интересен набор из M_j+K_j независимых уравнений в форме

$$\begin{bmatrix} S'_\diamond \Pi' \Sigma_{xy} S_\diamond & S'_\diamond \Pi' \Sigma_{xx} S_* \\ S'_* \Sigma_{xy} S_\diamond & S'_* \Sigma_{xx} S_* \end{bmatrix} \begin{bmatrix} \gamma_{\diamond j} \\ \beta_{*j} \end{bmatrix} = \begin{bmatrix} S'_\diamond \Pi' \Sigma_{xy} e_j \\ S'_* \Sigma_{xy} e_j \end{bmatrix}, \tag{20}$$

полученной умножением слева уравнения (19) на матрицу $[\Pi S_\diamond, S_*]'$. Эти уравнения, полученные при использовании лишь связей между параметрами нашей модели и моментами векторов данных x и y, должны быть основой любой приемлемой оценки параметров индивидуальных структурных уравнений, независимо от предпосылок, из которых она получена.

5 Оценивание одиночного структурного уравнения методом наименьших квадратов

Рассмотрим тождество $\eta_{t\bullet}\Gamma = -\varepsilon_{t\bullet}$, описывающее связь между структурной и приведенной ошибками. Оно содержит равенство $\eta_{t\bullet}\gamma_{\bullet j} = -\varepsilon_{tj}$, которое можно использовать для придания j-му структурному уравнению вида

$$(y_{t\bullet} - \eta_{t\bullet})\gamma_{\bullet j} + x_{t\bullet}\beta_{\bullet j} = 0. \tag{21}$$

Это уравнение модели ошибок в переменных, в которой ошибки распространяются только на подмножество переменных.

Проигнорируем индекс, указывающий на расположение j-го структурного уравнения в системе M уравнений. Если бы дисперсионная матрица $\mathbb{V}[\eta_{t\bullet}] = \Omega$ была известна, то оценки параметров γ и β получались бы путем нахождения допустимых значений переменных, минимизирующих функцию

$$\sum_{t=1}^{T} \eta_{t\bullet}\Omega^{-1}\eta'_{t\bullet} = \sum_{t=1}^{T} (y_{t\bullet} - \mu_{t\bullet})\Omega^{-1}(y_{t\bullet} - \mu_{t\bullet})', \quad \text{где} \quad \mu_{t\bullet} = y_{t\bullet} - \eta_{t\bullet} = x_{t\bullet}\Pi, \tag{22}$$

при ограничении

$$\mu_{t\bullet}\gamma + x_{t\bullet}\beta = 0. \tag{23}$$

Последнее ограничение следует из соотношения $\Pi\Gamma + B = 0$, связывающего параметры приведенной и структурной моделей. Вместе уравнения (21)–(23) составляют постановку задачи, для которой Pollock (1979) получил оценки параметров структурной модели.

Минимизация выражения (22) происходит в два шага. Для начала можно заметить, что при фиксированных значениях γ и β уравнение (23) определяет гиперплоскость в пространстве размерности $K + M$, содержащую векторы $[y_{t\bullet}, x_{t\bullet}]'$, состоящие из наблюдений переменных системы. Точка $[\mu_{t\bullet}, x_{t\bullet}]'$ содержится в данной гиперплоскости, и в метрике, определенной матрицей Ω^{-1}, квадрат расстояния от нее до соответствующего вектора наблюдений равен

$$\|y_{t\bullet} - \mu_{t\bullet}\|^2_{\Omega^{-1}} = (y_{t\bullet} - \mu_{t\bullet})\Omega^{-1}(y_{t\bullet} - \mu_{t\bullet})'. \tag{24}$$

Сначала минимизируем это расстояние для любых заданных γ и β. Затем найдем те значения γ и β, которые минимизируют сумму из квадратов расстояний, отраженную в формуле (22). Поэтому рассмотрим следующую функцию Лагранжа:

$$L = (y_{t\bullet} - \mu_{t\bullet})\Omega^{-1}(y_{t\bullet} - \mu_{t\bullet})' + 2\lambda(\mu_{t\bullet}\gamma + x_{t\bullet}\beta). \tag{25}$$

Приравнивая производную этой функции по $\mu'_{t\bullet}$ к нулю, получим условие первого порядка на минимум расстояния

$$-(y_{t\bullet} - \mu_{t\bullet})\Omega^{-1} + \lambda\gamma' = 0, \tag{26}$$

из которого следует, что

$$(y_{t\bullet} - \mu_{t\bullet})\Omega^{-1} = \lambda\gamma', \tag{27}$$

или

$$(y_{t\bullet} - \mu_{t\bullet}) = \lambda\gamma'\Omega. \tag{28}$$

Вместе эти два уравнения дают

$$(y_{t\bullet} - \mu_{t\bullet})\Omega^{-1}(y_{t\bullet} - \mu_{t\bullet})' = \lambda^2\gamma'\Omega\gamma. \tag{29}$$

Но, умножая (28) справа на γ и пользуясь соотношением $-\mu_{t\bullet}\gamma = x_{t\bullet}\beta$, получим:

$$y_{t\bullet}\gamma - \mu_{t\bullet}\gamma = y_{t\bullet}\gamma + x_{t\bullet}\beta = \lambda\gamma'\Omega\gamma, \tag{30}$$

откуда следует, что

$$\lambda = \frac{y_{t\bullet}\gamma + x_{t\bullet}\beta}{\gamma'\Omega\gamma}.$$

Таким образом, (29) можно представить в следующем виде:

$$(y_{t\bullet} - \mu_{t\bullet})\Omega^{-1}(y_{t\bullet} - \mu_{t\bullet})' = \frac{(y_{t\bullet}\gamma + x_{t\bullet}\beta)^2}{\gamma'\Omega\gamma}. \tag{31}$$

Введем матрицы $Y' = [y'_{1\bullet}, y'_{2\bullet}, \ldots, y'_{t\bullet}]$ и $X' = [x'_{1\bullet}, x'_{2\bullet}, \ldots, x'_{t\bullet}]$, содержащие в себе весь набор наблюдений в системе за T периодов. Тогда выражение для суммы квадратов отклонений наблюдений от регрессионной гиперплоскости принимает вид

$$\sum_{t=1}^{T}(y_{t\bullet} - \mu_{t\bullet})\Omega^{-1}(y_{t\bullet} - \mu_{t\bullet})' = \frac{(Y\gamma + X\beta)'(Y\gamma + X\beta)}{\gamma'\Omega\gamma}. \tag{32}$$

Целевая функция (32) должна быть минимизирована при ограничениях, учитывающих всю априорную информацию о γ и β. В сочетании с информацией, содержащейся в выборке Y и X, этой априорной информации должно быть достаточно для идентификации параметров структурной модели. Априорная информация о γ и β обычно имеет форму *исключающих ограничений*, указывающих на то, что некоторые переменные, присутствующие в расширенной системе, отсутствуют в конкретном структурном уравнении. Также следует учесть *правило нормировки*, указывающее на то, что один из элементов $y_{t\bullet}$ является зависимой переменой в конкретном уравнении.

Общепринятым способом выражения априорной информации о параметрах является запись уравнения в виде

$$R'_1\gamma + R'_2\beta = r. \tag{33}$$

В случае, если все ограничения исключающие, то, пока элементы вектора r равны нулю, соответствующие элементы матриц R'_1 и R'_2 будут равны нулям и единицам, кроме соответствующего нормирующему правилу случая, когда элемент r равен -1. Возвращаясь к уравнению (17), получим, что $[R_\diamond, 0] = R_1$ и $[0, R_*] = R_2$.

Выпишем лагранжиан оптимизационной задачи оценивания параметров структурной модели при ограничениях:

$$L = \frac{(Y\gamma + X\beta)'(Y\gamma + X\beta)}{\gamma'\Omega\gamma} + 2\kappa'(R_1'\gamma + R_2'\beta - r). \tag{34}$$

Дифференцируя его по γ, используя формулу производной произведения и приравнивая результат к нулю, получим

$$\frac{(Y\gamma + X\beta)'Y}{\gamma'\Omega\gamma} - \frac{(Y\gamma + X\beta)'(Y\gamma + X\beta)\gamma'\Omega}{(\gamma'\Omega\gamma)^2} + \kappa' R_1' = 0. \tag{35}$$

Определив новый множитель $\mu = \kappa\gamma'\Omega\gamma$ и умножив обе части (35) на $\gamma'\Omega\gamma$, получим уравнение, транспонированное к которому имеет вид

$$Y'Y\gamma + Y'X\beta - \left\{ \frac{(Y\gamma + X\beta)'(Y\gamma + X\beta)}{\gamma'\Omega\gamma} \right\} \Omega\gamma + R_1\mu = 0. \tag{36}$$

Далее, дифференцируя функцию Лагранжа (34) по β и приравнивая результат к нулю, получаем уравнение

$$\frac{(Y\gamma + X\beta)'X}{\gamma'\Omega\gamma} + \kappa' R_2' = 0, \tag{37}$$

которое при умножении на $\gamma'\Omega\gamma$ и транспонировании дает

$$X'Y\gamma + X'Y\beta + R_2\mu = 0. \tag{38}$$

Комбинируя уравнения (36) и (38) вместе с уравнением ограничений (33), получим следующую систему:

$$\begin{bmatrix} Y'Y - \lambda\Omega & Y'X & R_1 \\ X'Y & X'X & R_2 \\ R_1' & R_2' & 0 \end{bmatrix} \begin{bmatrix} \gamma \\ \beta \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix}, \tag{39}$$

где

$$\lambda = \left\{ \frac{(Y\gamma + X\beta)'(Y\gamma + X\beta)}{\gamma'\Omega\gamma} \right\}. \tag{40}$$

Чтобы использовать эти уравнения для оценивания параметров структурного уравнения, следует оценить значение Ω, дисперсионной матрицы шоков приведенной формы модели

$$\Omega = \mathbb{V}[\eta_{t\bullet}] = \mathbb{E}\left[(y_{t\bullet} - x_{t\bullet}\Pi)'(y_{t\bullet} - x_{t\bullet}\Pi) \right]. \tag{41}$$

Прямое применение метода моментов дает подходящую оценку

$$\hat{\Omega} = \frac{1}{T}(Y - X\hat{\Pi})'(Y - X\hat{\Pi}) = \frac{1}{T} Y' \big\{ I - X(X'X)^{-1}X' \big\} Y, \tag{42}$$

где $\hat{\Pi} = (X'X)^{-1}X'Y$ – оценка коэффициентов приведенной модели из уравнения (6).

Вместе уравнения (39) и (40) составляют нелинейную систему, которую следует решать итерационными методами, чтобы оценить структурные параметры. Можно привести итерационный алгоритм решения системы. На первом шаге производится присвоение $\lambda = \lambda_{(0)} = T$ в уравнении (39). В результате решения уравнения получаем оценки первого приближения $\gamma_{(1)}$ и $\beta_{(1)}$. Подстановка этих оценок в (40) дает обновленный коэффициент $\lambda_{(1)}$, которым заменим $\lambda_{(0)} = T$. При вторичном решении уравнения (39) с $\lambda = \lambda_{(1)}$ получаем оценки второго порядка $\gamma_{(2)}$ и $\beta_{(2)}$.

Несложно понять, как этот алгоритм обобщается на любое количество итераций. Процедуру можно остановить, когда значения γ и β, полученные из последовательных итераций, приблизительно равны. На практике четырех-пяти циклов бывает достаточно.

Значения $\gamma_{(1)}$ и $\beta_{(1)}$, возникающие на первом шаге алгоритма, на самом деле являются оценками 2ШМНК. Значения, к которым сходится алгоритм, являются оценками ММПОИ. Еще две итерационные схемы, начинающиеся с 2ШМНК-оценок и сходящиеся к ММПОИ-оценкам, были представлены в Pollock (1983). Первая из них производит итерации с учетом последующих оценок Π, что позволяет учитывать информацию об ограничениях на структурные параметры, тогда как второй метод производит итерации с учетом оценки Ω при ограничениях.

6 Стандартные формы оценок

Полезно вывести стандартные формы уравнений для 2ШМНК- и ММПОИ-оценивания, используя систему из уравнений (39) и (40). Для начала примем обычное допущение, что, не считая правила нормализации, которое мы проигнорируем, априорные ограничения на γ и β принимают вид исключающих ограничений в форме отсутствия в интересующем нас структурном уравнении некоторых переменных, присутствующих в расширенной системе.

Примем форму записи $Y = [Y_*, Y_{**}]$ и $X = [X_*, X_{**}]$, где Y_{**} и X_{**} – матрицы исключенных переменных, и пусть первый столбец Y_* равен вектору наблюдений зависимой переменной структурного уравнения. В таком случае решаемые уравнения принимают вид

$$\begin{bmatrix} Y_*'Y_* - \lambda\hat{\Omega}_* & Y_*'X_* \\ X_*'Y_* & X_*'X_* \end{bmatrix} \begin{bmatrix} \gamma_* \\ \beta_* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{43}$$

$$\lambda = \left\{ \frac{(Y_*\gamma_* + X_*\beta_*)'(Y_*\gamma_* + X_*\beta_*)}{\gamma_*'\Omega_*\gamma_*} \right\}, \tag{44}$$

$$\hat{\Omega}_* = \frac{1}{T} Y_*'\{I - X(X'X)^{-1}X'\}Y_*. \tag{45}$$

При решении второго уравнения системы (43), представляющего из себя $X_*'Y_*\gamma_* + X_*'X_*\beta_* = 0$, получаем

$$\beta_* = -(X_*'X_*)^{-1}X_*'Y_*\gamma_*. \tag{46}$$

При подстановке обратно в первое уравнение системы (43) получаем

$$(Y_*'Y_* - \lambda\hat{\Omega}_*)\gamma_* + Y_*'X_*(X_*'X_*)^{-1}X_*'Y_*\gamma_* = 0, \tag{47}$$

что можно переписать как

$$\{Y_*'(I - P_*)Y_* - \lambda\hat{\Omega}_*\}\gamma_* = 0, \quad \text{где } \hat{\Omega}_* = \frac{1}{T}Y_*'(I - P)Y_*, \tag{48}$$

$P = X(X'X)^{-1}X'$ и $P_* = X_*(X_*'X_*)^{-1}X_*'$.

Заметим, что первое уравнение в (48) имеет форму уравнения для оценки модели ошибок в переменных. Стоит осознать этот факт, и альтернативные способы нахождения оценок γ_* и β_* напрашиваются сами собой. Сначала уравнение (48) решается путем нахождения характеристического корня λ и соответствующего характеристического вектора γ_*, используя любую стандартную технику, например, степенной метод.

Такая методология позволяет находить значение γ_* при произвольных ограничениях, например, при нормирующем ограничении, присваивающем первому элементу вектора γ_* значение -1. Путем подстановки правильно нормированной γ_* обратно в уравнение (46) можно найти оценку β_*.

Численные результаты работы этой схемы в точности те же, что получились бы при исполнении описанного ранее итерационного алгоритма до полной сходимости. На самом деле этот метод нахождения ММПОИ-оценок был общим результатом двух достаточно непохожих друг на друга выводов Anderson & Rubin (1949) и Koopmans, Rubin & Leipnik (1950).

Теперь рассмотрим случай, когда определяющие оценки уравнения нормируются с самого начала. Определим вектора $[-1, \gamma'_\diamond] = \gamma'_*$ и $[y_0, Y_\diamond] = Y_*$ для удобства умножения. В таком случае структурное уравнение, ранее записывавшееся как $Y_*\gamma_* + X_*\beta_* + \varepsilon = 0$, преобразуется в $y_0 = Y_\diamond\gamma_\diamond + X_*\beta_* + \varepsilon$, и соответствующая система уравнений, определяющая оценки параметров, выглядит как

$$\begin{bmatrix} y_0'y_0 - \lambda\hat{\omega}_{00} & y_0'Y_\diamond - \lambda\hat{\omega}_{0\diamond} & y_0'X_* \\ Y_\diamond'y_0 - \lambda\hat{\omega}_{\diamond 0} & Y_\diamond'Y_\diamond - \lambda\hat{\Omega}_{\diamond\diamond} & Y_\diamond'X_* \\ X_*'y_0 & X_*'Y_\diamond & X_*'X_* \end{bmatrix} \begin{bmatrix} -1 \\ \gamma_\diamond \\ \beta_* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{49}$$

где

$$\hat{\Omega}_* = \begin{bmatrix} \hat{\omega}_{00} & \hat{\omega}_{0\diamond} \\ \hat{\omega}_{\diamond 0} & \hat{\Omega}_{\diamond\diamond} \end{bmatrix}. \tag{50}$$

Игнорируя первое уравнение системы, служащее для выражения λ через $\gamma_\diamond$ и β_*, а также преобразуя оставшиеся уравнения, получаем систему

$$\begin{bmatrix} Y_\diamond'Y_\diamond - \lambda\hat{\Omega}_{\diamond\diamond} & Y_\diamond'X_* \\ X_*'Y_\diamond & X_*'X_* \end{bmatrix} \begin{bmatrix} \gamma_\diamond \\ \beta_* \end{bmatrix} = \begin{bmatrix} Y_\diamond'y_0 - \lambda\hat{\omega}_{\diamond 0} \\ X_*'y_0 \end{bmatrix}. \tag{51}$$

Положим $\lambda = T$. Это не только то значение, которое λ принимает на первом шаге итерационного алгоритма, описанного выше в разделе 5, но и значение, приводящее уравнения в соответствие с условиями на моменты (16) (по этой причине $\lambda = T$ можно было бы описать как асимптотическое значение, в свете анализа Anderson, 2005). Решением полученной системы будет 2ШМНК-оценка.

Перепишем уравнения, определяющие 2ШМНК-оценку, в более общей форме при помощи определения (42). Получаем

$$Y'Y - T\hat{\Omega} = Y'Y - (Y - X\hat{\Pi})'(Y - X\hat{\Pi}) = Y'Y - Y'(I - P)Y = Y'PY = \hat{\Pi}'X'X\hat{\Pi}, \tag{52}$$

где $P = X(X'X)^{-1}X'$. Также отметим, что $X'Y = X'\{X(X'X)^{-1}X'\}Y = X'X\hat{\Pi}$. Используя эти результаты, оценочные уравнения (51) можно преобразовать в

$$\begin{bmatrix} \hat{\Pi}_{X\diamond}'X'X\hat{\Pi}_{X\diamond} & \hat{\Pi}_{X\diamond}'X'X_* \\ X_*'X\hat{\Pi}_{X\diamond} & X_*'X_* \end{bmatrix} \begin{bmatrix} \gamma_\diamond \\ \beta_* \end{bmatrix} = \begin{bmatrix} \Pi_{X\diamond}'X'X\hat{\Pi}_{X0} \\ X_*'X\hat{\Pi}_{X0} \end{bmatrix}, \tag{53}$$

где $X\hat{\Pi}_{X\diamond} = X(X'X)^{-1}X'Y_\diamond$ и $X\hat{\Pi}_{X0} = X(X'X)^{-1}X'y_0$. Пользуясь последним равенством, перепишем уравнения в виде

$$\begin{bmatrix} \hat{\Pi}_{X\diamond}'X'Y_\diamond & \hat{\Pi}_{X\diamond}'X'X_* \\ X_*'XY_\diamond & X_*'X_* \end{bmatrix} \begin{bmatrix} \gamma_\diamond \\ \beta_* \end{bmatrix} = \begin{bmatrix} \Pi_{X\diamond}'X'y_0 \\ X_*'Xy_0 \end{bmatrix}, \tag{54}$$

являющимся прямым аналогом уравнений (20), описывающих связи между популяционными моментами и параметрами структурного уравнения.

7 Двухшаговый метод наименьших квадратов и метод инструментальных переменных

Уравнения на оценки 2ШМНК были независимо получены в Basmann (1957) и Theil (1958) с помощью подходов, сильно отличающихся от использованного в данной работе. Подход авторов заключался в подробном изучении причин того, почему невозможно получить состоятельные оценки параметров структурного уравнения методом наименьших квадратов.

Причина провала МНК-оценивания заключается в нарушении ключевого положения регрессионного анализа о некоррелированности ошибок с объясняющими переменными правой части уравнения. В уравнении $y_0 = Y_\diamond\gamma_\diamond + X_*\beta_* + \varepsilon$ присутствует прямая зависимость $Y_\diamond$ от структурных ошибок ε. Ошибки, однако, независимы от экзогенных переменных X_*.

Первоначальные выводы 2ШМНК-оценок основывались на идее, что можно очистить переменные $Y_\diamond$ от их зависимости от ε, после чего стандартная регрессия методом наименьших квадратов становится приемлемым способом оценивания. Таким образом, если бы $X\Pi_{X\diamond}$ были доступны, то ими можно было бы заменить $Y_\diamond$, и проблема зависимости была бы преодолена.

Несмотря на то, что значение $X\Pi_{X\diamond}$ неизвестно, его состоятельная оценка доступна в виде $\hat{Y}_\diamond = X\hat{\Pi}_{X\diamond}$. Нахождение $\hat{\Pi}_{X\diamond}$ представляет собой первый шаг процедуры 2ШМНК. Применение МНК к уравнению $y_o = \hat{Y}_\diamond\gamma_\diamond + X_*\beta_* + e$ является вторым шагом. Совместно эти этапы представляют собой подход Theil (1958), давшего имя оценке двухшагового метода наименьших квадратов.

Альтернативный подход, приводящий к той же оценке 2ШМНК, осуществляется через метод инструментальных переменных. Суть этого метода состоит в нахождении ряда переменных, коррелированных с регрессорами, но некоррелированных с ошибками.

В случае структурного уравнения подходящими инструментальными переменными будут переменные, экзогенные ко всей системе, содержащиеся в матрице X. Умножая слева структурное уравнение на X', получим

$$X'y_0 = X'Y_\diamond\gamma_\diamond + X'X_*\beta_* + X'\varepsilon. \tag{55}$$

В этой системе перекрестные члены соответствуют матрицам моментов, имеющим следующие пределы:

$$\operatorname{plim}(T^{-1}X'y_0) = \Sigma_{xy}e_0, \quad \operatorname{plim}(T^{-1}X'Y_\diamond) = \Sigma_{xy}S_\diamond, \quad \operatorname{plim}(T^{-1}X'X_*) = \Sigma_{xx}S_*, \tag{56}$$
$$\operatorname{plim}(T^{-1}X'\varepsilon) = 0.$$

При замене матриц моментов на их предельные значения получим уравнение

$$\Sigma_{xy}e_0 = \Sigma_{xy}S_\diamond\gamma_\diamond + \Sigma_{xx}S_*\beta_*, \tag{57}$$

ранее уже встречавшееся, см. (19). В этой системе K уравнений и $M_\diamond + K_*$ параметров. Можно предположить, что $[\Sigma_{xy}, \Sigma_{xy}]$ имеет полный ранг, в этом случае необходимым условием идентифицируемости параметров $\gamma_\diamond$ и β_* является $K \geq M_\diamond + K_*$, смысл которого в том, что число оцениваемых структурных параметров системы не должно превышать количества экзогенных переменных.

Эмпирическим аналогом (57) является уравнение

$$X'y_0 = X'Y_\diamond\gamma_\diamond + X'X_*\beta_*. \tag{58}$$

В случае, если $K = M_\diamond + K_*$, это уравнение можно явно разрешить и получить оценки. Однако если $K > M_\diamond + K_*$, то уравнения алгебраически несовместны. В этом случае говорят, что его параметры сверхидентифицированы. Для разрешения этой несовместности можно применить к (55) обобщенный метод наименьших квадратов. Ошибка в (55), $X'\varepsilon$, имеет дисперсионную матрицу $\mathbb{V}[X'\varepsilon] = \sigma^2 X'X$. При использовании ее в ОМНК-оценивании мы снова получим оценку 2ШМНК.

Basmann (1957) в своей работе получил оценку 2ШМНК, следуя описанному ОМНК-подходу. Эту оценку также можно рассматривать в рамках метода инструментальных переменных, как это было сделано в Sargan (1958).

Список литературы

Anderson, T.W. (2005). Origins of the limited information maximum likelihood and two-stage least squares estimators. *Journal of Econometrics* 127, 1–16.

Anderson, T.W. & H. Rubin (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* 20, 46–63.

Anderson, T.W. & H. Rubin (1950). The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* 21, 570–82.

Basmann, R.L. (1957). A generalized classical method of linear estimation of coefficients in a structural equation. *Econometrica* 25, 77–84.

Koopmans, T.C., H. Rubin & R.B. Leipnik (1950). Measuring the equation systems of dynamic economics. Глава 2 в *Statistical Inference in Dynamic Economic Models* под редакцией T.C. Koopmans. New York: John Wiley & Sons.

Malinvaud, E. (1966). Statistical Methods of Econometrics. Amsterdam: North-Holland.

Pollock, D.S.G. (1979). The Algebra of Econometrics. Chichester: John Wiley & Sons.

Pollock, D.S.G. (1983). Varieties of the LIML estimator. *Australian Economic Papers* 1983, 499–506.

Sargan, J.D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica* 26, 393–415.

Theil, H. (1958, 1961). Economic Forecasts and Economic Policy. Amsterdam: North-Holland.

Estimation of Structural Econometric Equations

Stephen Pollock

Queen Mary College, University of London, United Kingdom

Derivations are offered for the LIML and the 2SLS estimators of single equations of the classical econometric simultaneous-equation system that differ from the usual ones. By assimilating both estimators to the method of moments, their essential similarities are highlighted. The LIML estimator is derived from a least-squares criterion that exploits the interpretation of the structural equation as an error-in-variables model, and the 2SLS estimator is obtained by an approximation that is asymptotically valid. The LIML estimator may be calculated via an iterative procedure that begins with the 2SLS estimator. The conventional derivations of the 2SLS estimator are also reviewed.