Лекция 10

Системы одновременных уравнений

1. Общее понятие СОУ.

2. Косвенный метод наименьших квадратов.

3. Проблема идентифицируемости.

4. Двухшаговый метод наименьших квадратов.

1. Общее понятие СОУ.

При исследовании экономических процессов для их описания не всегда достаточно только одного взятого уравнения. Кроме того, некоторые переменные могут настолько взаимодействовать друг с другом, что трудно однозначно определить, какая из них является зависимой, а какая – независимой. Поэтому при построении эконометрической модели процесса прибегают к системам уравнений.

Выделяют следующие три вида эконометрических систем уравнений:

- *система независимых уравнений*, когда каждая зависимая переменная y_i рассматривается как функция одного и того же набора объясняющих факторов:

$$y_i = f_i(x_1, x_2, \ldots, x_p, \varepsilon_i), \quad i = \overline{1, m};$$

- *система рекурсивных уравнений*, когда в каждом последующем $i-$м уравнении системы зависимая переменная y_i представляет функцию всех зависимых и независимых переменных предшествующих уравнений:

$$
\begin{cases}
y_1 = f_1(x_1, x_2, \ldots, x_p, \varepsilon_1), \\
y_2 = f_2(x_1, x_2, \ldots, x_p, y_1, \varepsilon_2), \\
y_3 = f_3(x_1, x_2, \ldots, x_p, y_1, y_2, \varepsilon_3), \\
\qquad \vdots \\
y_m = f_m(x_1, x_2, \ldots, x_p, y_1, y_2, \ldots, y_{m-1}, \varepsilon_m);
\end{cases}
$$

● *система одновременных* (взаимозависимых) *уравнений* (СОУ), когда зависимые переменные y_i в одних уравнениях входят в левую часть системы, а в других уравнениях – в правую часть:

$$y_i = f_i(x_1, x_2, \ldots, x_p, y_1, \ldots, y_{m-1}, y_{m+1}, \ldots, y_m, \varepsilon_i), \quad i = \overline{1, m}.$$

В рассмотренных первых двух видах систем каждое уравнение этих систем может рассматриваться самостоятельно, отдельно, и для оценивания коэффициентов можно применять обычный МНК.

В рассмотренных первых двух видах систем каждое уравнение этих систем может рассматриваться *самостоятельно*, отдельно, и для оценивания коэффициентов можно применять обычный МНК.

В третьем виде систем каждое уравнение уже не может рассматриваться самостоятельно, и для нахождения его оценок традиционный МНК неприменим.

Например, если изучается модель спроса, как соотношение цены и количества потребляемых товаров, то одновременно для прогнозирования спроса необходима модель предложения, в которой рассматривается взаимосвязь меду количеством и ценой предлагаемых товаров:

$$Q_d = \beta_1 + \beta_2 P + \beta_3 I + \varepsilon_1, \quad (1)$$

$$Q_s = \beta_4 + \beta_5 P + \varepsilon_2, \quad (2)$$

$$Q_d = Q_s, \quad (3)$$

где Q_d – спрос, Q_s – предложение, P – цена товара, I – доход потребителя.

Уравнение (3) показывает, что рынок находится в состоянии равновесия и следует положить: $Q_d = Q_s = Q$. В этом случае наблюдаемое значение P – это *цена равновесия*, которая формируется вместе со спросом и предложением.

Переменные P и Q считаются *объясняемыми* (зависимыми) переменными, а доход I – *объясняющей* (независимой) переменной.

Переменные P и Q формируют свои значения, подчиняясь уравнениям (1) и (2), т.е. внутри модели. Такие переменные в эконометрике называют *эндогенными*.

Переменная же I считается заданной, её значения формируются вне модели (1-2). Переменные такого типа называют *экзогенными*. С математической точки зрения главное отличие между эндогенными и экзогенными переменными заключается в том, что экзогенные переменные *не коррелируют с ошибками* регрессии ε.

Эндогенные переменные в системе одновременных уравнений, записанной в общем виде, обозначаются символом y, а экзогенные - x.

В более общем виде СОУ включает множество эндогенных текущих значений переменных $y_{i,t}$ и множество *предопре-деленных* переменных, к которым относят *лаговые и текущие* значения экзогенных переменных x_j, а также *лаговые* значения эндогенных переменных, например, $y_{i,t-1}$.

Как уже отмечалось в СОУ одни и те же переменные одновременно рассматривают-ся как зависимые в одних уравнениях и как независимые – в других уравнениях. Такую запись СОУ называют *структурной формой модели*. Некоторые уравнения структурной формы могут быть тождествами, когда коэффициенты при переменных уравнения известны. Например, уравнение (3) является тождеством.

Простейшая структурная форма модели с двумя эндогенными и двумя экзогенными переменными имеет вид:

$$y_1 = \beta_1 x_1 + \gamma_1 y_2 + \varepsilon_1, \quad (4)$$

$$y_2 = \beta_2 x_2 + \gamma_2 y_1 + \varepsilon_2. \quad (5)$$

Использование МНК для оценивания *структурных* коэффициентов модели $\beta_1, \beta_2, \gamma_1, \gamma_2$ даёт смещенные и несостоятельные оценки вследствие нарушения предпосылок, лежащих в основе МНК.

Оценивание структурных коэффициентов СОУ требует специальных методов. Одним из них является *косвенный метод наименьших квадратов* (КМНК), в котором от структурной формы переходят к *приведенной форме* модели. Приведенная форма – это система независимых уравнений, в которой эндогенные переменные модели выражены только через предопределенные переменные, не зависящие от ошибок .

2. Косвенный метод наименьших квадратов.

Косвенный метод наименьших квадратов включает три шага:

1) получение приведенной формы модели и представление каждого коэффициента приведенной формы через структурные коэффициенты;

2) применение обычного МНК к каждому независимому уравнению приведенной формы и получение точечных оценок коэффициентов приведенной формы;

3) определение оценок структурных коэффициентов по оценкам коэффициентов приведенной формы с использованием полученных на первом шаге соотношений, связывающих эти коэффициенты.

Проиллюстрируем сказанное на примере модели (4-5).

Первый шаг. Выразим y_2 из уравнения (4):

$$y_2 = \frac{y_1 - \beta_1 x_1 - \varepsilon_1}{\gamma_1}.$$

Приравнивая правые части полученного соотношения и уравнения (5), после несложных преобразований получим

$$y_1 - \beta_1 x_1 - \varepsilon_1 = \gamma_1 \beta_2 x_2 + \gamma_1 \gamma_2 y_1 + \gamma_1 \varepsilon_2,$$

откуда находим

$$y_1 = \frac{\beta_1}{1 - \gamma_1 \gamma_2} x_1 + \frac{\beta_2 \gamma_1}{1 - \gamma_1 \gamma_2} x_2 + \frac{\varepsilon_1 + \gamma_1 \varepsilon_2}{1 - \gamma_1 \gamma_2}.$$

Если ввести обозначения

$$b_1 = \frac{\beta_1}{1-\gamma_1\gamma_2}, \quad c_1 = \frac{\beta_2\gamma_1}{1-\gamma_1\gamma_2}, \quad \nu_1 = \frac{\varepsilon_1 + \gamma_1\varepsilon_2}{1-\gamma_1\gamma_2}, \quad (6)$$

то получим уравнение

$$y_1 = b_1x_1 + c_1x_2 + \nu_1. \qquad (7)$$

Выполняя аналогичные преобразования с переменной из уравнения (5), можно получить:

$$y_2 = b_2x_1 + c_2x_2 + \nu_2, \qquad (8)$$

где

$$b_2 = \frac{\beta_1\gamma_2}{1-\gamma_1\gamma_2}, \quad c_2 = \frac{\beta_2}{1-\gamma_1\gamma_2}, \quad \nu_2 = \frac{\varepsilon_2 + \gamma_2\varepsilon_1}{1-\gamma_1\gamma_2}. \quad (9)$$

Система уравнений (7), (8) представляет приведенную форму модели, так как эндогенные переменные модели выражены только через экзогенные переменные .

Одновременно получены соотношения (6), (9), которые связывают коэффициенты приведенной формы со структурными коэффициентами. Отметим, что это *нелинейные* соотношения.

Второй шаг. Для получения оценок коэффициентов приведенной формы b_1, b_2, c_1, c_2 МНК требуются статистические данные в виде многомерной выборки:

t	y_{1t}	y_{2t}	x_{1t}	x_{2t}
1	y_{11}	y_{21}	x_{11}	x_{21}
2	y_{12}	y_{22}	x_{12}	x_{22}
…	…	…	…	…
n	y_{1n}	y_{2n}	x_{1n}	x_{2n}

На основании этих данных находятся оценки коэффициентов приведенной формы для каждого уравнения отдельно. Например, оценка b_1 найдется по формуле:

$$\hat{b}_1 = \frac{\sum_{t=1}^{n} x_{2t}^2 \cdot \sum_{t=1}^{n} x_{1t} y_{1t} - \sum_{t=1}^{n} x_{1t} x_{2t} \cdot \sum_{t=1}^{n} x_{2t} y_{1t}}{\sum_{t=1}^{n} x_{1t}^2 \cdot \sum_{t=1}^{n} x_{2t}^2 - \left(\sum_{t=1}^{n} x_{1t} x_{2t} \right)^2}. \quad (10)$$

Третий шаг. Заменяя в равенствах (6), (9) коэффициенты полученными на втором шаге их оценками , после несложных преобразований можно получить оценки структурных коэффициентов через оценки коэффициентов приведенной формы:

$$\hat{\beta}_1 = \frac{\hat{b}_1\hat{c}_2 - \hat{b}_2\hat{c}_1}{\hat{c}_2}, \quad \hat{\beta}_2 = \frac{\hat{b}_1\hat{c}_2 - \hat{b}_2\hat{c}_1}{\hat{b}_1}, \quad \hat{\gamma}_1 = \frac{\hat{c}_1}{\hat{c}_2}, \quad \hat{\gamma}_2 = \frac{\hat{b}_2}{\hat{b}_1}.$$

3. Проблема идентифицируемости.

Экономический смысл и интерес для анализа представляют только структурные коэффициенты модели. Однако переход от оценок коэффициентов приведенной формы к оценкам коэффициентов структурной формы не всегда возможен или возможен не единственным образом. Здесь возникает проблема идентифицируемости.

Идентифицируемость - это единственность соответствия между приведенной и структурной формами модели.

С позиции идентифицируемости структурные формы модели можно подразделить на три вида:

- идентифицируемые;
- неидентифицируемые;
- сверхидентифицируемые.

В первом случае все структурные коэффициенты однозначно, единственным образом определяются через коэффициенты приведенной формы. Число коэффициентов в обеих формах модели одинаково (как в рассмотренной простейшей структурной форме (4), (5)).

Во втором случае число коэффициентов *приведенной* формы меньше числа коэффициентов *структурной* формы и в результате последние не определяются через коэффициенты приведенной формы.

Наконец, в последнем случае число коэффициентов приведенной формы превышает число коэффициентов структурной формы, и они определяются не однозначно.

Каждое уравнение структурной формы модели требуется проверить на идентифицируемость. Модель считается идентифицируемой, если каждое уравнение системы идентифицируемо.

Если хотя бы одно уравнение системы неидентифицируемо, то вся модель считается неидентифицируемой.

Если хотя бы одно уравнение сверх-идентифицируемо, то и вся модель считается сверхидентифицируемой.

Для структурной формы модели введем следующие величины:

H – число эндогенных переменных системы;

h_j – число эндогенных переменных в j-м уравнении системы;

D – число предопределенных перемен-ных системы;

d_j – число предопределенных переменных в $j -$ м уравнении системы;

A_j – матрица коэффициентов при переменных (всех типов), не входящих в $j -$ е уравнение системы.

Необходимое условие идентифицируемост $j -$ го уравнения системы:

- если выполняется равенство $D - d_j = h_j - 1$, то уравнение идентифицируемо;
- если выполняется неравенство $D - d_j > h_j - 1$, то уравнение сверхидентифицируемо;
- если выполняется неравенство $D - d_j < h_j - 1$, то уравнение неидентифицируемо.

Достаточное условие идентифицируемости j –го уравнения системы:

Для того чтобы j –е уравнение системы было идентифицируемо достаточно, чтобы ранг матрицы A_j был равен ($H-1$).

Отсюда можно сформулировать необходимое и достаточное условие идентифицируемости j –го уравнения системы:

1. Если $D-d_j > h_j - 1$ и ранг матрицы A_j равен ($H-1$), то уравнение **сверхидентифицируемо.**

2. Если $D - d_j = h_j - 1$ и ранг матрицы A_j равен ($H - 1$), то уравнение **идентифицируемо.**

3. Если $D - d_j \geq h_j - 1$ и ранг матрицы A_j меньше ($H - 1$), то уравнение **неидентифицируемо.**

4. Если $D - d_j < h_j - 1$, то уравнение **неидентифицируемо.**

4. Двухшаговый метод наименьших квадратов.

Если СОУ сверхидентифицируема, то КМНК не используется, ибо он не даёт однозначных оценок структурных коэффициентов. Оценки сверхидентифицируемого уравнения осуществляются при помощи других методов, простейшим из которых является двухшаговый метод наименьших квадратов (ДМНК).

Алгоритм ДМНК включает следующие шаги:
1. Получение приведённой формы модели.
2. Применение обычного МНК к каждому независимому уравнению приведённой формы модели и получение точечных оценок коэффициентов приведённой формы.
3. Определение на основе полученных уравнений приведенной формы *расчётных* значений эндогенных переменных, которые фигурируют в качестве объясняющих переменных в структурной форме модели (в правой части).

4. Определение оценок структурных коэффициентов каждого уравнения в отдельности обычным МНК, используя в качестве факторов входящие в это уравнение предопределенные переменные и расчетные значения переменных, полученные на третьем шаге алгоритма. Как видно из алгоритма обычный МНК применяется дважды.

ДМНК можно применять и для точно идентифицируемой СОУ. В этом случае оценки, полученные КМНК и ДМНК, совпадут.

Рассмотрим применение ДМНК на примере системы (4)-(5):

$$y_1 = \beta_1 x_1 + \gamma_1 y_2 + \varepsilon_1,$$
$$y_2 = \beta_2 x_2 + \gamma_2 y_1 + \varepsilon_2.$$

Первый шаг. Получаем приведенную форму (см. пункт 2 лекции):

$$y_1 = b_1 x_1 + c_1 x_2 + v_1,$$

$$y_2 = b_2 x_1 + c_2 x_2 + v_2.$$

Второй шаг. Для оценки коэффициентов уравнений приведенной формы используем статистические данные таблицы 1. В результате получаем уравнения регрессии:

$$\tilde{y}_1 = \hat{b}_1 x_1 + \hat{c}_1 x_2,$$

$$\tilde{y}_2 = \hat{b}_2 x_1 + \hat{c}_2 x_2, \qquad (11)$$

где, например, оценка $\hat{b}_1$ вычисляется по формуле (10).

Третий шаг. В правые части исходной структурной формы входят в качестве объясняющих обе эндогенные переменные . Получаем расчётные значения этих переменных по уравнениям (11) с использованием статистических данных таблицы 1. В результате имеем:

Таблица 2

t	$\tilde{y}_{1t}$	$\tilde{y}_{2t}$
1	$\tilde{y}_{11}$	$\tilde{y}_{21}$
…	…	…
n	$\tilde{y}_{1n}$	$\tilde{y}_{2n}$

Четвёртый шаг. Для определения оценок первого уравнения структурной формы используем обычный МНК с данными, взятыми из таблицы 1 (x_{1t}, y_{1t}) и расчётными ($\tilde{y}_{2t}$) из таблицы 2:

Таблица 3

t	x_{1t}	y_{1t}	$\tilde{y}_{2t}$
1	x_{11}	y_{11}	$\tilde{y}_{21}$
…	…	…	…
n	x_{1n}	y_{1n}	$\tilde{y}_{2n}$

В результате получаем уравнение регрессии

$$\widetilde{y}_1 = \hat{\beta}_1 x_1 + \hat{\gamma}_1 y_2.$$

Аналогично по исходным данным

t	x_{2t}	$\widetilde{y}_{1t}$	y_{2t}
1	x_{21}	$\widetilde{y}_{11}$	y_{21}
…	…	…	…
n	x_{2n}	$\widetilde{y}_{1n}$	y_{2n}

оцениваются коэффициенты второго уравнения

$$\widetilde{y}_2 = \hat{\beta}_2 x_2 + \hat{\gamma}_2 y_1.$$

По своей сути использование расчётных значений представляет собой метод *инструментальных переменных*, и если коэффициент детерминации для оцениваемого приведенного уравнения низкий, то расчётные значения эндогенной переменной будут плохой аппроксимацией её фактических значений и применение двухшагового МНК может оказаться неэффективным.

В этом случае используют другие методы, например, *трёхшаговый* МНК.