# СИСТЕМЫ ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ

Объектом статистического изучения в социальных науках являются сложные системы. Построение изолированных уравнений регрессии недостаточно для описания таких систем и объяснения механизма их функционирования. Изменение одной переменной, как правило, не может происходить без изменения других. Поэтому важное место занимает проблема описания структуры связей между переменными системой так называемых **одновременных уравнений.** Так, если изучается модель спроса как отношение цен и количества потребляемых товаров, то одновременно для прогнозирования спроса необходима модель предложения товаров, в которой рассматривается также взаимосвязь между количеством и ценой предлагаемых благ. Это позволяет достичь равновесия между спросом и предложением.

Системы уравнений здесь могут быть построены по – разному.

Возможна **система независимых уравнений,** когда каждая зависимая переменная $y$ рассматривается как функция одного и того же набора факторов $x$:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m + \varepsilon_{1,} \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m + \varepsilon_{2,} \\ \ldots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m + \varepsilon_{n.} \end{cases} \quad (1)$$

Набор факторов $x_j$ в каждом уравнении может варьироваться. Каждое уравнение может рассматриваться самостоятельно. Для нахождения его параметров используется МНК. По существу, каждое уравнение этой системы является уравнением регрессии.

Наибольшее распространение в эконометрических исследованиях получила **система одновременных (совместных, взаимозависимых) уравнений.** В ней одни и те же зависимые переменные в одних уравнениях входят в левую часть, а в других уравнениях – в правую часть:

$$\begin{cases} y_1 = b_{12}y_2 + b_{13}y_3 + \ldots + b_{1n}y_n + a_{11}x_1 + \ldots + a_{1m}x_m + \varepsilon_{1,} \\ y_2 = b_{21}y_1 + b_{23}y_3 + \ldots + b_{2n}y_n + a_{21}x_1 + \ldots + a_{2m}x_m + \varepsilon_{2,} \\ \ldots \\ y_n = b_{n1}y_1 + b_{n2}y_2 + \ldots + b_{n,n-1}y_{n-1} + a_{n1}x_1 + \ldots + a_{nm}x_m + \varepsilon_{n.} \end{cases} \quad (2)$$

В эконометрике эта система уравнений называется также **структурной формой модели.** Для нахождения параметров каждого уравнения традиционный МНК неприменим, здесь используются специальные методы оценивания. В этом случае каждое из уравнений не может рассматриваться самостоятельно.

### Структурная и приведенная формы модели.

Система одновременных уравнений (т.е. структурная форма модели) обычно содержит эндогенные и экзогенные переменные.

**Эндогенные переменные** – это зависимые переменные, число которых равно числу уравнений в системе. Они обозначаются через $y$

**Экзогенные переменные** – это предопределенные переменные, влияющие на эндогенные переменные, но не зависящие от них. Они обозначаются через $x$.

Простейшая структурная форма модели имеет вид:

$$\begin{cases} y_1 = b_{12}y_2 + a_{11}x_1 + \varepsilon_{1,} \\ y_2 = b_{21}y_1 + a_{22}x_2 + \varepsilon_{2,} \end{cases}$$

где $y_1$,$y_2$ - эндогенные переменные, $x_1$,$x_2$ - экзогенные.

Классификация переменных на эндогенные и экзогенные зависит от теоретической концепции принятой модели. Экономические переменные могут выступать в одних моделях как эндогенные, а в других - как экзогенные переменные. Внеэкономические переменные (например, климатические условия) входят в систему как экзогенные переменные. В качестве экзогенных переменных можно рассматривать значения эндогенных переменных за предшествующий период времени (лаговые переменные). Например, потребление текущего года $y_t$ может зависеть также и от уровня потребления в предыдущем году $y_{t-1}$.

Структурная форма модели позволяет увидеть влияние изменений любой экзогенной переменной на значения эндогенной переменной. Целесообразно в качестве экзогенных переменных выбирать такие переменные, которые могут быть объектом регулирования. Меняя их и управляя ими, можно заранее иметь целевые значения эндогенных переменных.

Коэффициенты $b_i$ при эндогенных и $a_j$ - при экзогенных переменных называются **структурными коэффициентами модели.** Все переменные в модели могут быть выражены в отклонениях

$(x-\bar{x})$ и $(y-\bar{y})$ от среднего уровня, и тогда свободный член в каждом уравнении отсутствует.

Использование МНК для оценивания структурных коэффициентов модели дает смещенные и несостоятельные оценки. Поэтому обычно для определения структурных коэффициентов модели структурная форма преобразуется в приведенную.

**Приведенная форма модели** представляет собой систему линейных функций эндогенных переменных от экзогенных:

$$\begin{cases} \hat{y}_1 = \delta_{11}x_1 + \ldots + \delta_{1m}x_{m,} \\ \hat{y}_2 = \delta_{21}x_1 + \ldots + \delta_{2m}x_{m,} \\ \ldots \\ \hat{y}_n = \delta_{n1}x_1 + \ldots + \delta_{nm}x_{m.} \end{cases} \quad (3)$$

$\delta_{ij}$ – коэффициенты приведенной формы модели.

По своему виду приведенная форма модели ничем не отличается от системы независимых уравнений, Применяя МНК, можно оценить $\delta_{ij}$, а затем оценить значения эндогенных переменных через экзогенные.

Приведенная форма позволяет выразить значения эндогенных переменных через экзогенные, однако аналитически уступает структурной форме модели, т.к. в ней отсутствуют оценки взаимосвязи между эндогенными переменными.

## Проблема идентификации

При переходе от приведенной формы модели к структурной исследователь сталкивается с проблемой идентификации. **Идентификация** – это единственность соответствия между приведенной и структурной формами модели.

Структурная модель (2) в полном виде, состоящая в каждом уравнении системы из $n$ эндогенных и $m$ экзогенных переменных, содержит $n(n\text{-}1+m)$ параметров. Приведенная модель (3) в полном виде содержит $nm$ параметров. Таким образом, в полном виде структурная модель содержит большее число параметров, чем приведенная форма модели. Поэтому $n(n\text{-}1+m)$ параметров структурной модели не могут быть однозначно определены через $nm$ параметров приведенной формы модели.

Чтобы получить единственно возможное решение для структурной модели, необходимо предположить, что некоторые из структурных коэффициентов модели равны нулю. Тем самым уменьшится число структурных коэффициентов.

С позиции идентифицируемости структурные модели можно подразделить на три вида:

- идентифицируемые;
- неидентифицируемые;
- сверхидентифицируемые.

Модель **идентифицируема,** если все структурные ее коэффициенты определяются однозначно, единственным образом по коэффициентам приведенной формы модели, т.е. число параметров структурной модели равно числу параметров приведенной формы модели.

Модель **неидентифицируема,** если число приведенных коэффициентов меньше числа структурных коэффициентов, и в результате структурные коэффициенты не могут быть оценены через коэффициенты приведенной формы модели. Модель (2) в полном виде **всегда неидентифицируема.**

Модель **сверхидентифицируема,** если число приведенных коэффициентов больше числа структурных коэффициентов. В этом случае на основе приведенных коэффициентов можно получить два или более значений одного структурного коэффициента. Сверхидентифицируемая модель, в отличие от неидентифицируемой, практически решаема, но требует для этого специальных методов исчисления параметров.

Структурная модель всегда представляет собой систему совместных уравнений, каждое из которых требуется проверять на идентификацию. **Модель считается идентифицируемой, если каждое уравнение системы идентифицируемо.** Если хотя бы одно из уравнений системы неидентифицируемо, то и вся модель считается неидентифицируемой. Сверхидентифицируемая модель содержит хотя бы одно сверхидентифицируемое уравнение.

Обозначим $H$ – число эндогенных переменных в $i$- ом уравнении системы, $D$ – число экзогенных переменных, которые содержатся в системе, но не входят в данное уравнение. Тогда условие идентифицируемости уравнения может быть записано в виде следующего счетного правила:

$D+1 = H$ – уравнение идентифицируемо;

$D+1 < H$ – уравнение неидентифицируемо;

$D+1 > H$ – уравнение сверхидентифицируемо.

Это счетное правило отражает необходимое, но не достаточное условие идентификации. Более точно условия идентификации определяются, если накладывать ограничения на коэффициенты матриц параметров структурной модели. Уравнение идентифицируемо, если по отсутствующим в нем переменным (эндогенным и экзогенным) можно из коэффициентов при них в других уравнениях системы получить матрицу, определитель которой не равен нулю, а ранг матрицы не меньше, чем число эндогенных переменных в системе без одного.

**Пример**. Рассмотрим следующую макроэкономическую модель:

$$\begin{cases} M_t = a_1 + b_{12}N_t + b_{13}S_t + b_{14}E_{t-1} + b_{15}M_{t-1} + \varepsilon_1, \\ N_t = a_2 + b_{21}M_t + b_{23}S_t + b_{26}Y_t + \varepsilon_2, \\ S_t = a_3 + b_{31}M_t + b_{32}N_t + b_{37}X_t + \varepsilon_3, \end{cases}$$

где $M$ – доля импорта в ВВП;

$N$ – общее число прошений об освобождении от таможенных пошлин;

$S$ – число удовлетворенных прошений;

$E$ – фиктивная переменная, означающая, является ли курс доллара искусственно завышенным или нет;

$Y$ – реальный ВВП;

$X$ – реальный объём чистого экспорта;

$t$ – текущий период;

$t\text{-}1$ – предыдущий период.

Проверим данную модель на идентификацию и определим, каким методом могут быть рассчитаны её коэффициенты (в случае, если модель сверх – или точно идентифицируема).

Сначала рассмотрим общие характеристики структурной формы. Здесь три эндогенные переменные – $M_t$, $N_t$ и $S_t$, они стоят в левых частях уравнений. Кроме того, в правых частях находятся четыре предопределенные переменные – одна лаговая ($M_{t-1}$) и три экзогенные – $E_{t-1}$, $Y_t$ и $X_t$. Теперь проверим каждое уравнение.

**Уравнение I**. В этом уравнении присутствуют три эндогенные переменные ($M_t$, $N_t$ и $S_t$), но отсутствуют две предопределенные переменные - $Y_t$ и $X_t$. Поэтому $H=3$, $D=2$, и необходимое условие идентификации выполняется, поскольку $D+1=H$. Это означает, что первое уравнение точно идентифицируемо.

**Уравнение II**. В этом уравнении присутствуют три эндогенные переменные ($M_t$, $N_t$ и $S_t$), но отсутствуют три экзогенные - $E_{t-1}$, $M_{t-1}$ и $X_t$. Поэтому $H=3$, $D=3$, $D+1>H$ и второе уравнение по необходимому условию является сверхидентифицируемым.

**Уравнение III**. В этом уравнении, как и в других уравнениях, присутствуют все три эндогенные переменные, но отсутствуют три экзогенные - $E_{t-1}$, $M_{t-1}$ и $Y_t$ . Поэтому $H=3$, $D=3$, $D+1>H$, и третье уравнение системы является сверхидентифицируемым.

Проверим каждое уравнение на выполнение достаточного условия идентификации. Для этого сначала запишем расширенную матрицу системы в виде следующей таблицы:

| Уравнение | $M_t$ | $N_t$ | $S_t$ | $E_{t-1}$ | $M_{t-1}$ | $Y_t$ | $X_t$ |
|---|---|---|---|---|---|---|---|
| I | -1 | $b_{12}$ | $b_{13}$ | $b_{14}$ | $b_{15}$ | 0 | 0 |
| II | $b_{21}$ | -1 | $b_{23}$ | 0 | 0 | $b_{26}$ | 0 |
| III | $b_{31}$ | $b_{32}$ | -1 | 0 | 0 | 0 | $b_{37}$ |

Как видим, в эту матрицу включены коэффициенты при всех переменных и не включены свободные члены, поскольку они могут быть исключены из системы, если задавать все переменные в отклонениях от среднего значения. Кроме того, здесь все переменные перенесены в правые части уравнений.

Достаточное условие идентификации для соответствующего уравнения будет выполнено, если ранг подматрицы, построенной только из коэффициентов при переменных, отсутствующих в этом уравнении, равен количеству эндогенных переменных в системе минус единица.

Рассмотрим подробно этот процесс для первого уравнения системы. Первому уравнению соответствует первая строка расширенной матрицы, поэтому первую строку не следует включать в подматрицу. Из остальной части расширенной матрицы оставим только столбцы, которые имеют нули в первой строке. Получаем подматрицу:

$$\begin{bmatrix} b_{26} & 0 \\ 0 & b_{37} \end{bmatrix},$$

определитель которой не равен нулю, поскольку $b_{26} \cdot b_{37} \neq 0$. Таким образом, ранг подматрицы равен двум, т.е. числу эндогенных переменных в системе минус единица. Достаточное условие идентификации для первого уравнения выполнено.

Аналогично рассмотрим другие уравнения. Подматрица для второго уравнения имеет вид:

$$\begin{bmatrix} b_{14} & b_{15} & 0 \\ 0 & 0 & b_{37} \end{bmatrix}.$$

Её ранг также равен двум, поскольку определитель, составленный, например, из первого и третьего столбцов, очевидно, не равен нулю.

Подматрица для третьего уравнения имеет вид:

$$\begin{bmatrix} b_{14} & b_{15} & 0 \\ 0 & 0 & b_{26} \end{bmatrix}.$$

Она также имеет ранг, равный двум.

Таким образом, достаточное условие идентификации выполнено для каждого уравнения системы. Поскольку среди уравнений системы нет неидентифицируемых, а второе и третье уравнения являются сверхидентифицированными, то и модель в целом сверхидентифицирована. Для определения параметров первого уравнения должен быть применен косвенный МНК (поскольку оно точно идентифицировано), а для других уравнений – двухшаговый МНК.

Приведенная форма модели имеет вид:

$$\begin{cases} M_t = A_1 + A_2 E_{t-1} + A_3 M_{t-1} + A_4 X_t + A_5 Y_t + V_1, \\ N_t = B_1 + B_2 E_{t-1} + B_3 M_{t-1} + B_4 X_t + B_5 Y_t, + V_2 \\ S_t = C_1 + C_2 E_{t-1} + C_3 M_{t-1} + C_4 X_t + C_5 Y_t, + V_3. \end{cases}$$

Здесь $V_1, V_2, V_3$ - случайные члены. Как обычно, в правой части приведенной формы стоят только предопределенные переменные. Для определения параметров ПФМ применяется обычный МНК.

## Оценивание параметров структурной модели

Коэффициенты структурной модели могут быть оценены разными способами в зависимости от вида системы одновременных уравнений. Наибольшее распространение получили два метода оценивания коэффициентов структурной модели: **косвенный МНК** и **двухшаговый МНК.**

**Косвенный МНК (**КМНК) применим в случае точно идентифицируемой структурной модели. Процедура следующая:

1. Структурная модель преобразуется в приведенную форму.
2. Для каждого уравнения приведенной формы обычным МНК оцениваются коэффициенты $\delta_{ij}$

3. Коэффициенты приведенной модели трансформируются в параметры структурной модели.

Рассмотрим применение КМНК для модели:

$$\begin{cases} y_1 = b_{12}y_2 + a_{11}x_1 + \varepsilon_{1,} \\ y_2 = b_{21}y_1 + a_{22}x_2 + \varepsilon_2 \end{cases}$$

Для построения модели имеем таблицу:

| № п/п | $y_1$ | $y_2$ | $x_1$ | $x_2$ |
|---|---|---|---|---|
| 1 | 2 | 5 | 1 | 3 |
| 2 | 3 | 6 | 2 | 1 |
| 3 | 4 | 7 | 3 | 2 |
| 4 | 5 | 8 | 2 | 5 |
| 5 | 6 | 5 | 4 | 6 |
| Средние | 4 | 6,2 | 2,4 | 3,4 |

Приведенная форма модели имеет вид:

$$\begin{cases} y_1 = \delta_{11}x_1 + \delta_{12}x_2 + u_{1,} \\ y_2 = \delta_{21}x_1 + \delta_{22}x_2 + u_{2,} \end{cases}$$

где $u_{1,}u_2$ – случайные ошибки приведенной формы модели.

Для каждого уравнения приведенной формы применим традиционный МНК и определим $\delta$- коэффициенты. Для простоты работаем в отклонениях, т.е. $y = y - \bar{y}, x = x - \bar{x}$. Тогда система нормальных уравнений для первого уравнения системы составит:

$$\begin{cases} \delta_{11}\sum x_1^2 + \delta_{12}\sum x_1x_2 = \sum y_1x_1 \\ \delta_{11}\sum x_1x_2 + \delta_{12}\sum x_2^2 = \sum y_1x_2 \end{cases}$$

Для приведенных данных система составит:

$$\begin{cases} 5{,}2\delta_{11} + 4{,}2\delta_{12} = 6, \\ 4{,}2\delta_{11} + 17{,}2\delta_{12} = 10. \end{cases}$$

Отсюда получаем первое уравнение ( и аналогично второе):

$$\begin{cases} y_1 = 0{,}852x_1 + 0{,}373x_2 + u_1 \\ y_2 = -0{,}072x_1 - 0{,}00557x_2 + u_2 \end{cases}$$

Перейдем к структурной форме следующим образом: исключим из первого уравнения приведенной формы $x_2$ , выразив его из второго уравнения приведенной формы и подставив в первое уравнение:

$$x_2 = \frac{-0{,}072x_1 - y_2}{0{,}00557}$$

Первое уравнение структурной формы:

$$\hat{y}_1 = 0{,}852x_1 + 0{,}373\left(\frac{-0{,}072x_1 - y_2}{0{,}00557}\right) = -66{,}966y_2 - 3{,}97x_1$$

Аналогично исключим из второго уравнения $x_1$ выразив его через первое уравнение и подставив во второе:

$$x_1 = \frac{y_1 - 0{,}373x_2}{0{,}852}; \Rightarrow \hat{y}_2 = -0{,}072\left(\frac{y_1 - 0{,}373x_2}{0{,}852}\right) - 0{,}00557x_2$$

$\hat{y}_2 = -0{,}085y_1 + 0{,}026x_2 -$ второе уравнение структурной формы.

Структурная форма модели имеет вид:

$$\begin{cases} y_1 = -66{,}966y_2 - 3{,}97x_1 + \varepsilon_{1,} \\ y_2 = -0{,}085y_1 + 0{,}026x_2 + \varepsilon_2 \end{cases}$$

Эту же систему можно записать, включив в нее свободный член уравнения, т.е. перейти от переменных в виде отклонений от среднего к исходным переменным $y$ и $x$:

$$A_{01} = \bar{y}_1 - b_{12}\bar{y}_2 - a_{11}\bar{x}_1 = 428{,}717$$

$$A_{02} = \bar{y}_2 - b_{21}\bar{y}_1 - a_{22}\bar{x}_2 = 6{,}451$$

Тогда структурная модель имеет вид:

$$\begin{cases} y_1 = 428{,}717 - 66{,}966y_2 - 3{,}97x_1 + \varepsilon_{1,} \\ y_2 = 6{,}451 - 0{,}085y_1 + 0{,}026x_2 + \varepsilon_2 \end{cases}$$

Если к каждому уравнению структурной формы применить традиционный МНК, то результаты могут сильно отличаться. В данном примере будет:

$$\begin{cases} y_1 = -1{,}09 + 0{,}364y_2 + 1{,}192x_1 + \varepsilon_{1,} \\ y_2 = 5{,}2 + 0{,}533y_1 - 0{,}333x_2 + \varepsilon_2 \end{cases}$$

**Двухшаговый МНК.** ДМНК используется для сверхидентифицируемых систем. Основная идея ДМНК**:** на основе приведенной формы модели получить для сверхидентифицируемого уравнения теоретические значения эндогенных переменных, содержащихся в правой части уравнения. Далее, подставив их вместо фактических значений, можно применить обычный МНК к структурной форме сверхидентифицируемого уравнения. Здесь дважды используется МНК: на первом шаге при определении приведенной формы модели и нахождении на ее основе оценок теоретических значений эндогенной переменной $\hat{y}_i = \delta_{i1}x_1 + \delta_{i2}x_2 + ... + \delta_{im}x_{m,}$ и на втором шаге применительно к структурному сверхидентифицируемому уравнению при определении структурных коэффициентов модели по данным теоретических (расчетных) значений эндогенных переменных.

Сверхидентифицируемая структурная модель может быть двух типов:
- все уравнения системы сверхидентифицируемые;
- система содержит также точно идентифицируемые уравнения.

В первом случае для оценки структурных коэффициентов каждого уравнения используется ДМНК. Во втором случае структурные коэффициенты для точно идентифицируемых уравнений находятся из системы приведенных уравнений.

Рассмотрим модель:

$$\begin{cases} y_1 = b_{12}(y_2 + x_1) + \varepsilon_{1,} \\ y_2 = b_{21}y_1 + a_{22}x_2 + \varepsilon_2 \end{cases}$$

Она получена из предыдущего примера наложением ограничения $b_{12} = a_{11.}$ Поэтому первое уравнение стало сверхидентифицируемым.

На первом шаге найдем приведенную форму модели. С использованием тех же исходных данных получим систему:

$$\begin{cases} y_1 = 0{,}852x_1 + 0{,}373x_2 + u_{1,} \\ y_2 = -0{,}072x_1 - 0{,}00557x_2 + u_{2.} \end{cases}$$

На основе второго уравнения этой системы можно найти теоретические значения для эндогенной переменной $y_{2,}$ т.е. $\hat{y}_{2.}$ Подставим в это уравнение значения $x_1$ и $x_2$ в форме отклонений от средних значений, запишем в виде таблицы:

| $x_1$ | $x_2$ | $\hat{y}_2$ | $\hat{y}_2 + x_1 = z$ | $y_1$ | $y_1 z$ | $z^2$ |
|---|---|---|---|---|---|---|
| -1,4 | -0,4 | 0,103 | -1,297 | -2 | 2,594 | 1,682 |
| -0,4 | -2,4 | 0,042 | -0,358 | -1 | 0,358 | 0,128 |
| 0,6 | -1,4 | -0,035 | 0,565 | 0 | 0 | 0,319 |
| -0,4 | 1,6 | 0,02 | -0,38 | 1 | -0,38 | 0,144 |
| 1,6 | 2,6 | -0,13 | 1,47 | 2 | 2,94 | 2,161 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 5,512 | 4,434 |

После того, как найдены оценки $\hat{y}_{2,}$ заменим в уравнении $y_1 = b_{12}(y_2 + x_1)$ фактические значения $y_2$ их оценками $\hat{y}_{2,}$ найдем значения новой переменной $z = \hat{y}_2 + x_{1.}$ Применим МНК к уравнению:

$$y_1 = b_{12} z\,.$$

Получим:

$$b_{12} = \frac{\sum y_1 z}{\sum z^2} = \frac{5{,}512}{4{,}434} = 1{,}243.$$

В целом рассматриваемая система будет иметь вид:

$$\begin{cases} y_1 = 1{,}243(y_2 + x_2) \\ y_2 = -0{,}085 y_1 + 0{,}026 x_2 \end{cases}$$

Второе уравнение не изменилось по сравнению с предыдущим примером.

ДМНК является наиболее общим и широко распространенным методом решения системы одновременных уравнений. Для точно идентифицируемых уравнений ДМНК дает тот же результат, что и КМНК.

## Применение систем эконометрических уравнений.

Наиболее широко системы одновременных уравнений используются при построении макроэкономических моделей экономики страны. В большинстве случаев это мультипликаторные модели кейнсианского типа. Статическая модель Кейнса народного хозяйства в самом простом виде следующая:

$$\begin{cases} C = a + by + \varepsilon, \\ y = C + I, \end{cases}$$

где $C$ - личное потребление;
$y$ - национальный доход в постоянных ценах;
$I$ - инвестиции в постоянных ценах.

В силу наличия тождества в модели (второе уравнение системы) $b \le 1.$ Он характеризует предельную склонность к потреблению. Если $b = 0{,}65,$ из каждой дополнительной тысячи рублей дохода на потребление расходуется в среднем 650 рублей и 350 рублей инвестируется. Если $b>1$ то $y<C+I$, и на потребление расходуются не только

доходы, но и сбережения. Параметр $a$ Кейнс истолковывал как прирост потребления за счет других факторов.

Структурный коэффициент $b$ используется для расчета мультипликаторов. По данной функции потребления можно определить два мультипликатора – инвестиционный мультипликатор потребления $M_c$ и национального дохода $M_y$:

$$M_c = \frac{b}{1-b}, \text{ т.е. при } b = 0{,}65 \quad M_c = \frac{0{,}65}{1-0{,}65} = 1{,}857$$

Это означает, что дополнительные вложения 1 тыс. руб. приведут при прочих равных условиях к дополнительному увеличению потребления на 1,857 тыс. руб.

$$M_y = \frac{1}{1-b}, \text{ т.е. при } b = 0{,}65 \; M_y = \frac{1}{1-0{,}65} = 2{,}857,$$

т.е. дополнительные вложения 1 тыс. руб. на длительный срок приведут при прочих равных условиях к дополнительному доходу 2,857 тыс. руб.

Эта модель точно идентифицируема, и для получения $b$ применяется КМНК. Строится система приведенных уравнений:

$$\begin{cases} C = A + B \cdot I + U_1 \\ y = A' + B' \cdot I + U_{2,} \end{cases}$$

в которой $A = A'$, а параметры $B$ и $B'$ являются мультипликаторами, т.е. $B = M_c$ и $B' = M_y$. Для проверки подставим балансовое равенство в первое уравнение структурной модели:

$$C = a + by + \varepsilon = a + b(C + I) + \varepsilon = a + bc + bI + \varepsilon \Rightarrow$$

$$\Rightarrow C(1-b) = a + bI + \varepsilon \Rightarrow C = \underbrace{\frac{a}{1-b}}_{A} + \underbrace{\frac{b}{1-b}}_{M_c} I + \varepsilon \underbrace{\frac{1}{1-b}}_{U_1}$$

Аналогично поступим и со вторым уравнением структурной модели:

$$y = C + I \Rightarrow y = a + by + \varepsilon + I \Rightarrow y(1-b) = a + I + \varepsilon \Rightarrow$$

$$\Rightarrow y = \underbrace{\frac{a}{1-b}}_{A' = A} + \underbrace{\frac{1}{1-b}}_{M_y} I + \underbrace{\frac{1}{1-b}}_{U_2} \varepsilon$$

Таким образом, приведенная форма содержит мультипликаторы, интерпретируемые как коэффициенты множественной регрессии, отвечающие на вопрос, на сколько единиц изменится значение эндогенной переменной, если экзогенная изменится на 1 единицу. Это делает модель удобной для прогнозирования.

В более поздних исследованиях статическая модель Кейнса включала уже не только функцию потребления, но и функцию сбережений:

$$\begin{cases} C = a + by + \varepsilon_{1,} \\ r = T + K(C + I) + \varepsilon_{2,} \\ y = C + I - r, \end{cases}$$

где $r$ — сбережения.

Здесь три эндогенные переменные - $C, r$ и $y$ — и одна экзогенная - $I$. Система идентифицируема: в первом уравнении $H$=2 и $D$=2, во втором $H$=1, $D$=0; $C + I$ рассматривается как предопределенная переменная.

Наряду со статическими широкое распространение получили динамические модели экономики. Они содержат в правой части лаговые переменные, а также учитывают тенденцию. Например, модель Кейнса экономики США 1950-1960 гг. в упрощенном варианте:

$$\begin{cases} C_t = b_1 S_t + b_2 P_t + b_3 + \varepsilon_{1,} \\ I_t = b_4 P_t + b_5 P_{t-1} + b_6 + \varepsilon_{2,} \\ S_t = b_7 R_t + b_8 R_{t-1} + b_9 t + b_{10} + \varepsilon_{3,} \\ R_t = S_t + P_t + T_{t,} \\ R_t = C_t + I_t + G_t \end{cases}$$

$T_t$ — чистые трансферты в пользу администрации;

$I_t$ — кап. вложения;

$G_t$ — правительственные расходы;

$S_t$ — заработная плата в период $t$;

$P_t$ — прибыль;

$P_{t-1}$ — прибыль в период $t-1$;

$R_t$ — общий доход.

Модель содержит 5 эндогенных переменных - $C_t, I_t, S_t, R_t$ ( в левой части системы) и $P_t$ (зависимая переменная, определяемая по первому тождеству), три экзогенные переменные - $T_t, G_t, t$ и две лаговые предопределенные переменные $P_{t-1}$ и $R_{t-1}$. Данная модель сверхидентифицируема и решается ДМНК. Для прогнозных целей используется приведенная форма модели:

$$\begin{cases} C_t = d_1T + d_2G + d_3t + d_4P_{t-1} + d_5R_{t-1} + U_1, \\ I_t = d_6T + d_7G + d_8t + d_9P_{t-1} + d_{10}R_{t-1} + U_2, \\ S_t = d_{11}T + d_{12}G + d_{13}t + d_{14}P_{t-1} + d_{15}R_{t-1} + U_3, \\ R_t = d_{16}T + d_{17}G + d_{18}t + d_{19}P_{t-1} + d_{20}R_{t-1} + U_4, \\ P_t = d_{21}T + d_{22}G + d_{23}t + d_{24}P_{t-1} + d_{25}R_{t-1} + U_5 \end{cases}$$

Здесь мультипликаторами являются коэффициенты при экзогенных переменных. Они отражают влияние экзогенной переменной на эндогенную переменную.

Система одновременных уравнений нашла применение в исследованиях спроса и предложения. Линейная модель спроса и предложения имеет вид:

$$\begin{cases} Q^d = a_0 + a_1P + \varepsilon_1, \text{ - объём спроса,} \\ Q^s = b_0 + b_1P + \varepsilon_2, \text{ - объём предложения,} \\ Q^d = Q^s \end{cases}$$

Здесь 3 эндогенные переменные: $Q^d, Q^s$ и $P$. При этом, если $Q^d$ и $Q^s$ представляют собой эндогенные переменные, исходя из структуры самой системы, то $P$ является эндогенной по экономическому содержанию (цена зависит от спроса и предложения), а также в результате наличия тождества $Q^d = Q^s$. Приравняем уравнения, получим:

$$a_0 + a_1P + \varepsilon_1 = b_0 + b_1P + \varepsilon_2$$

$$P = \frac{b_0 - a_0}{a_1 - b_1} + \frac{\varepsilon_2 - \varepsilon_1}{a_1 - b_1}$$

Модель не содержит экзогенной переменной. Однако, чтобы модель имела статистическое решение и можно было убедиться в ее справедливости, в модель вводятся экзогенные переменные.

Например, модель вида:

$$\begin{cases} Q^d = a_0 + a_1 P + a_2 R + \varepsilon_{1,} \\ Q^s = b_0 + b_1 P + b_2 W + \varepsilon_{2,} \\ Q^d = Q^s, \end{cases}$$

где $R -$ доход на душу населения; $W -$ климатические условия (при спросе и предложении зерна).

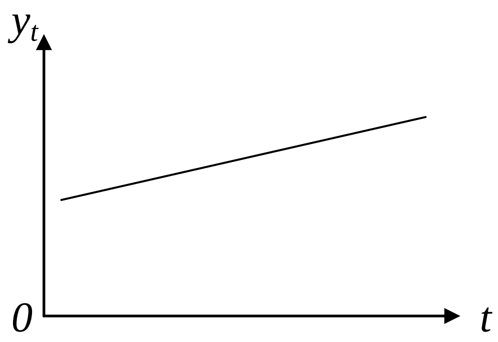
Переменные $R$ и $W$ экзогенные. Введя их в модель получаем идентифицированную структурную модель, где можно применить КМНК.

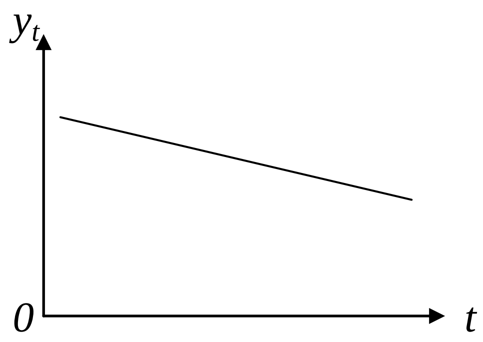
# ВРЕМЕННЫЕ РЯДЫ В ЭКОНОМЕТРИЧЕСКИХ ИССЛЕДОВАНИЯХ.

**Временно́й ряд –** это совокупность значений какого – либо показателя за несколько последовательных моментов или периодов времени. Каждое значение (уровень) временного ряда формируется под воздействием большого числа факторов, которые можно условно разделить на три группы:

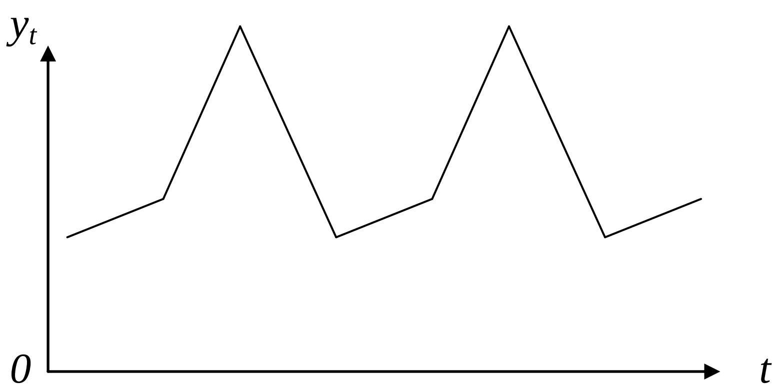
- факторы, формирующие тенденцию ряда;
- факторы, формирующие циклические колебания ряда;
- случайные факторы.

Тенденция характеризует долговременное воздействие факторов на динамику показателя. Тенденция может быть возрастающей или убывающей.




Циклические колебания могут носить сезонный характер или отражать динамику конъюнктуры рынка, а также фазу бизнес – цикла, в которой находится экономика страны.



Реальные данные часто содержат все три компоненты. В большинстве случаев временной ряд можно представить как сумму или произведение трендовой $(T)$, циклической $(S)$ и случайной $(E)$ компонент. В случае суммы имеет место **аддитивная** модель временного ряда:

$$y = T + S + E, \tag{1}$$

в случае произведения – **мультипликативная** модель:

$$y = T \cdot S \cdot E \tag{2}$$

Основная задача эконометрического исследования отдельного временного ряда – выявление количественного выражения каждой из компонент и использование полученной информации для прогноза будущих значений ряда или построение модели взаимосвязи двух или более временных рядов.

Сначала рассмотрим основные подходы к анализу отдельного временного ряда. Такой ряд может содержать, помимо случайной составляющей, либо только тенденцию, либо только сезонную (циклическую) компоненту, либо все компоненты вместе. Для того, чтобы выявить наличие той или иной неслучайной компоненты, исследуется корреляционная зависимость между последовательными уровнями временного ряда, или **автокорреляция уровней ряда.** Основная идея такого анализа заключается в том, что при наличии во временном ряде тенденции и циклических колебаний значения каждого последующего уровня ряда зависят от предыдущих.

Количественно автокорреляцию можно измерить с помощью линейного коэффициента корреляции между уровнями исходного временного ряда и уровнями этого ряда, сдвинутыми на несколько шагов во времени.

Коэффициент автокорреляции уровней ряда первого порядка измеряет зависимость между соседними уровнями ряда $t$ и $t-1,$ т.е. при лаге 1.

Он вычисляется по следующей формуле:

$$r_1 = \frac{\sum_{t=2}^{n}(y_t - \bar{y}_1)(y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^{n}(y_t - \bar{y}_1)^2 \sum_{t=2}^{n}(y_{t-1} - \bar{y}_2)^2}}, \tag{3}$$

где в качестве средних величин берутся значения:

$$\bar{y}_1 = \frac{\sum_{t=2}^{n} y_t}{n-1}; \qquad \bar{y}_2 = \frac{\sum_{t=2}^{n} y_{t-1}}{n-1}; \tag{4}$$

В первом случае усредняются значения ряда, начиная со второго до последнего, во втором случае - значения ряда с первого до предпоследнего.

Формулу (3) можно представить как формулу выборочного коэффициента корреляции:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}, \tag{5}$$

где в качестве переменной $x$ берется ряд $y_2, y_3, ..., y_n$, а в качестве переменной $y$ – ряд $y_1, y_2, ..., y_{n-1}$.

Если значение коэффициента (3) близко к единице, это указывает на очень тесную зависимость между соседними уровнями временного ряда и о наличии во временном ряде сильной линейной тенденции.

Аналогично определяются коэффициенты автокорреляции более высоких порядков. Так, коэффициент автокорреляции второго порядка характеризует тесноту связи между уровнями $y_t$ и $y_{t-2}$ и определяется по формуле:

$$r_2 = \frac{\sum_{t=3}^{n} (y_t - \bar{y}_3)(y_{t-2} - \bar{y}_4)}{\sqrt{\sum_{t=3}^{n} (y_t - \bar{y}_3)^2 \sum_{t=2}^{n} (y_{t-2} - \bar{y}_4)^2}}, \tag{6}$$

где в качестве одной средней величины берут среднюю уровней ряда с третьего до последнего, а в качестве другой - среднюю с первого уровня до $y_{n-2}$:

$$\bar{y}_3 = \frac{\sum_{t=3}^{n} y_t}{n-2}; \qquad \bar{y}_4 = \frac{\sum_{t=3}^{n} y_{t-2}}{n-2}; \tag{7}$$

Число периодов, по которым рассчитывается коэффициент автокорреляции, называют **лагом.** С увеличением лага число пар значений, по которым рассчитывается коэффициент автокорреляции, уменьшается. Для обеспечения статистической достоверности максимальный лаг, как считают некоторые известные эконометристы, не должен превышать четверти общего объема выборки.

Коэффициент автокорреляции строится по аналогии с линейным коэффициентом корреляции, и поэтому он характеризует тесноту только линейной связи текущего и предыдущего уровней ряда. По нему можно судить о наличии линейной или близкой к линейной тенденции. Однако для некоторых временных рядов с сильной нелинейной тенденцией (например, параболической или экспоненциальной), коэффициент автокорреляции уровней ряда может приближаться к нулю.

Кроме того, по знаку коэффициента автокорреляции нельзя делать вывод о возрастающей или убывающей тенденции в уровнях ряда. Большинство временных рядов экономических данных имеют положительную автокорреляцию уровней, однако при этом не исключается убывающая тенденция.

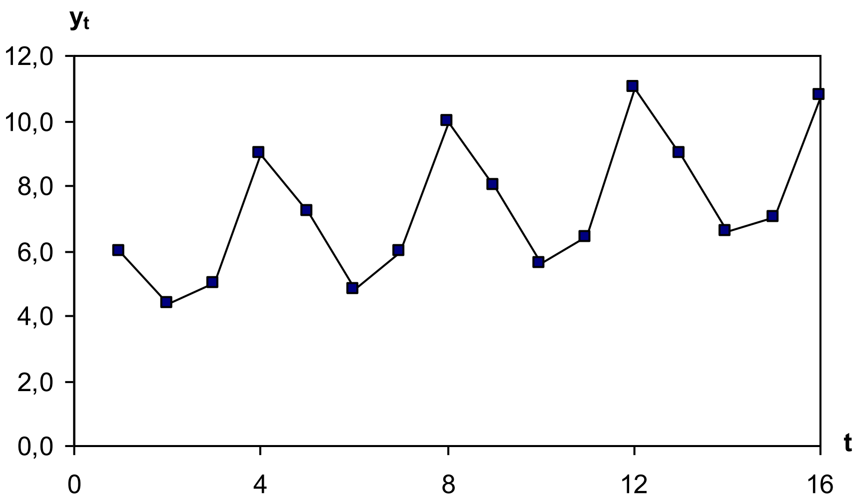
Последовательность коэффициентов автокорреляции уровней различных порядков, начиная с первого, называется **автокорреляционной функцией временного ряда.** График зависимости ее значений от величины лага называется **коррелограммой.** Анализ автокорреляционной функции и коррелограммы помогает выявить структуру ряда. Здесь уместно привести следующие качественные рассуждения.

Если наиболее высоким является коэффициент автокорреляции первого порядка, очевидно, исследуемый ряд содержит только тенденцию. Если наиболее высоким оказался коэффициент автокорреляции порядка $\tau$, ряд содержит циклические колебания с периодичностью в $\tau$ моментов времени. Если ни один из коэффициентов автокорреляции не является значимым, то либо ряд не содержит тенденции и циклических колебаний и имеет только случайную составляющую, либо ряд содержит сильную нелинейную тенденцию, для исследования которой нужно провести дополнительный анализ.

**Пример**. Пусть имеются данные об объёмах потребления электроэнергии жителями района за 16 кварталов, млн. квт.-ч:

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $y_t$ | 6,0 | 4,4 | 5,0 | 9,0 | 7,2 | 4,8 | 6,0 | 10,0 | 8,0 | 5,6 | 6,4 | 11,0 | 9,0 | 6,6 | 7,0 | 10,8 |

Нанесем эти значения на график:




Определим автокорреляционную функцию данного временного ряда. Рассчитаем коэффициент автокорреляции первого порядка. Для этого определим средние значения:

$$\bar{y}_1 = \frac{4{,}4 + 5{,}0 + 9{,}0 + \ldots + 10{,}8}{15} = 7{,}3867;$$

$$\bar{y}_2 = \frac{6{,}0 + 4{,}4 + 5{,}0 + \ldots + 7{,}0}{15} = 7{,}0667.$$

С учетом этих значений можно построить вспомогательную таблицу:

| t | $y_t$ | $y_t - \bar{y}_1$ | $y_t - \bar{y}_2$ | $(y_t - \bar{y}_1)(y_{t-1} - \bar{y}_2)$ | $(y_t - \bar{y}_1)^2$ | $(y_t - \bar{y}_1)^2$ |
|---|---|---|---|---|---|---|
| 1 | 6,0 | | -1,0667 | | | 1,137778 |
| 2 | 4,4 | -2,9867 | -2,6667 | 3,185778 | 8,920178 | 7,111111 |
| 3 | 5,0 | -2,3867 | -2,0667 | 6,364444 | 5,696178 | 4,271111 |
| 4 | 9,0 | 1,6133 | 1,9333 | -3,33422 | 2,602844 | 3,737778 |
| 5 | 7,2 | -0,1867 | 0,1333 | -0,36089 | 0,034844 | 0,017778 |
| 6 | 4,8 | -2,5867 | -2,2667 | -0,34489 | 6,690844 | 5,137778 |
| 7 | 6,0 | -1,3867 | -1,0667 | 3,143111 | 1,922844 | 1,137778 |
| 8 | 10,0 | 2,6133 | 2,9333 | -2,78756 | 6,829511 | 8,604444 |
| 9 | 8,0 | 0,6133 | 0,9333 | 1,799111 | 0,376178 | 0,871111 |
| 10 | 5,6 | -1,7867 | -1,4667 | -1,66756 | 3,192178 | 2,151111 |
| 11 | 6,4 | -0,9867 | -0,6667 | 1,447111 | 0,973511 | 0,444444 |
| 12 | 11,0 | 3,6133 | 3,9333 | -2,40889 | 13,05618 | 15,47111 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 13 | 9,0 | 1,6133 | 1,9333 | 6,345778 | 2,602844 | 3,737778 |
| 14 | 6,6 | -0,7867 | -0,4667 | -1,52089 | 0,618844 | 0,217778 |
| 15 | 7,0 | -0,3867 | -0,0667 | 0,180444 | 0,149511 | 0,004444 |
| 16 | 10,8 | 3,4133 | | -0,22756 | 11,65084 | |
| Итог | | | | 9,813333 | 65,3173 | 54,0533 |

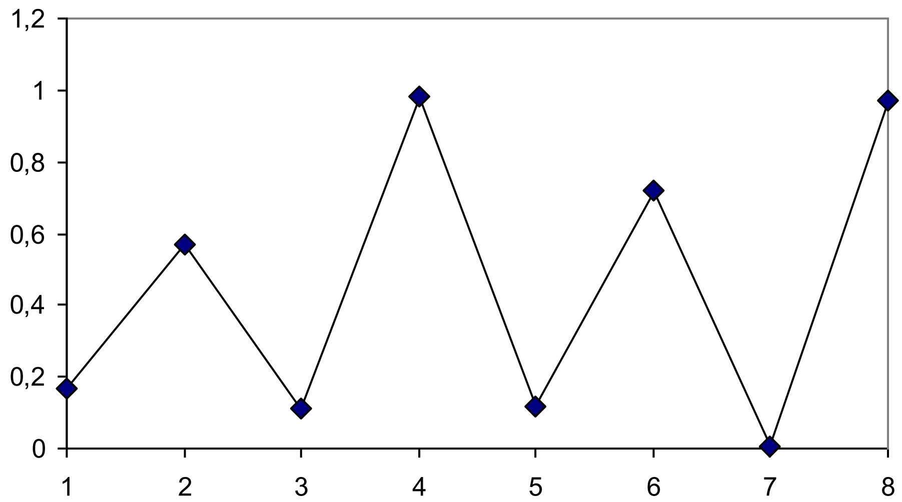
С помощью итоговых сумм подсчитаем величину коэффициента автокорреляции первого порядка:

$$r_1 = 0{,}165154\,.$$

Это значение свидетельствует о слабой зависимости текущих уровней ряда от непосредственно им предшествующих. Однако из графика очевидно наличие возрастающей тенденции уровней ряда, на которую накладываются циклические колебания.

Продолжая аналогичные расчеты для второго, третьего и т.д. порядков, получим автокорреляционную функцию, значения которой сведем в таблицу:

| Лаг | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $r_t$ | 0,16515 | 0,56687 | 0,11355 | 0,98302 | 0,11871 | 0,72204 | 0,00336 | 0,97384 |




Из коррелограммы видно, что наиболее высокий коэффициент корреляции наблюдается при значении лага, равном четырем, следовательно, ряд имеет циклические колебания периодичностью в четыре квартала. Это подтверждается и графическим анализом структуры ряда.

В случае, если при анализе структуры временного ряда обнаружена только тенденция и отсутствуют циклические колебания (случайная составляющая присутствует всегда), следует приступать к моделированию тенденции. Если же во временном ряде имеют место и циклические колебания, прежде всего следует исключить именно циклическую составляющую, и лишь затем приступать к моделированию тенденции. Выявление тенденции состоит в построении аналитической функции, характеризующей зависимость уровней ряда от времени, или **тренда**. Этот способ называют **аналитическим выравниванием временного ряда**.

Зависимость от времени может принимать разные формы, поэтому для её формализации используют различные виды функций:

- линейный тренд: $\hat{y}_t = a + b \cdot t$;
- гипербола: $\hat{y}_t = a + b / t$;
- экспоненциальный тренд: $\hat{y}_t = e^{a+b\cdot t}$ (или $\hat{y}_t = a \cdot b^t$);
- степенной тренд: $\hat{y}_t = a \cdot t^b$;
- параболический тренд второго и более высоких порядков:

$$\hat{y}_t = a + b_1 \cdot t + b_2 \cdot t^2 + ... + b_k \cdot t^k .$$

Параметры каждого из трендов можно определить обычным МНК, используя в качестве независимой переменной время $t = 1, 2, ..., n$, а в качестве зависимой переменной – фактические уровни временного ряда $y_t$ (или уровни за вычетом циклической составляющей, если таковая была обнаружена). Для нелинейных трендов предварительно проводят стандартную процедуру их линеаризации.

Существует несколько способов определения типа тенденции. Чаще всего используют качественный анализ изучаемого процесса, построение и визуальный анализ графика зависимости уровней ряда от времени, расчет некоторых основных показателей динамики. В этих же целях можно использовать и коэффициенты автокорреляции уровней ряда. Тип тенденции можно определить путем сравнения коэффициентов автокорреляции первого порядка, рассчитанных по исходным и преобразованным уровням ряда. Если временной ряд имеет линейную тенденцию, то его соседние уровни $y_t$ и $y_{t-1}$ тесно коррелируют. В этом случае коэффициент автокорреляции первого порядка уровней исходного ряда должен быть высоким. Если временной ряд содержит нелинейную тенденцию, например, в форме экспоненты, то коэффициент автокорреляции первого порядка по логарифмам уровней исходного ряда будет выше, чем соответствующий коэффициент, рассчитанный по уровням ряда. Чем сильнее выражена нелинейная тенденция в изучаемом временном ряде, тем в большей степени будут различаться значения указанных коэффициентов.

Выбор наилучшего уравнения в случае, если ряд содержит нелинейную тенденцию, можно осуществить путем перебора основных форм тренда, расчета по каждому уравнению скорректированного коэффициента детерминации $\overline{R}^2$ и выбора уравнения тренда с максимальным значением этого коэффициента. Реализация этого метода относительно проста при компьютерной обработке данных.

При анализе временных рядов, содержащих сезонные или циклические колебания, наиболее простым подходом является расчет значений сезонной компоненты методом скользящей средней и построение аддитивной или мультипликативной модели временно́го ряда в форме (1) или (2).

Если амплитуда колебаний приблизительно постоянна, строят аддитивную модель (1), в которой значения сезонной компоненты предполагаются постоянными для различных циклов. Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель (2), которая ставит уровни ряда в зависимость от значений сезонной компоненты.

Построение модели (1) или (2) сводится к расчету значений $T$, $S$ или $E$ для каждого уровня ряда. Процесс построения модели включает в себя следующие шаги:

1. Выравнивание исходного ряда методом скользящей средней.
2. Расчет значений сезонной компоненты $S$.
3. Устранение сезонной компоненты из исходных уровней ряда и получение выровненных данных ($T+E$) в аддитивной или ($T \cdot E$) в мультипликативной модели.
4. Аналитическое выравнивание уровней ($T+E$) или ($T \cdot E$) и расчет значений $T$ с использованием полученного уравнения тренда.
5. Расчет полученных по модели значений ($T+S$) или ($T \cdot S$)
6. Расчет абсолютных и относительных ошибок.

**Пример**. ***Построение аддитивной модели временного ряда***. Рассмотрим данные об объёме потребления электроэнергии жителями района из ранее приведенного примера. Из анализа автокорреляционной функции было показано, что данный временно́й ряд содержит сезонные колебания периодичностью в 4 квартала. Объёмы потребления электроэнергии в осенне – зимний период (I и IV кварталы) выше, чем весной и летом (II и III кварталы). По графику этого ряда можно установить наличие приблизительно равной амплитуды колебаний. Это говорит о возможном наличии аддитивной модели. Рассчитаем её компоненты.

***Шаг 1***. Проведем выравнивание исходных уровней ряда методом скользящей средней.

Поскольку циклические колебания имеют периодичность в 4 квартала, просуммируем уровни ряда последовательно за каждые 4 квартала со сдвигом на один момент времени и определим условные годовые объёмы потребления электроэнергии (колонка 3 в таблице 1).

Разделив полученные суммы на 4, найдем скользящие средние (колонка 4 таблицы 1). Полученные таким образом выровненные значения уже не содержат сезонной компоненты.

Поскольку скользящие средние получены осреднением четырех соседних уровней ряда, т.е. четного числа значений, они соответствуют серединам подынтервалов, состоящих из четверок чисел, т.е. должны располагаться между третьим и четвертым значениями четверок исходного ряда. Для того, чтобы скользящие средние располагались на одних временных отметках с исходным рядом, пары соседних скользящих средних ещё раз усредняются и получаются центрированные скользящие средние (колонка 5 таблицы 1). При этом теряются первые две и последние две отметки временного ряда, что связано с осреднением по четырем точкам.

Таблица 1.

| № квартала | Потребление электроэнергии $y_t$ | Итого за четыре квартала | Скользящая средняя за четыре квартала | Центрированная скользящая средняя | Оценка сезонной компоненты |
|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** |
| 1 | 6,0 | | | | |
| 2 | 4,4 | | | | |
| 3 | 5,0 | 24,4 | 6,10 | 6,25 | -1,250 |
| 4 | 9,0 | 25,6 | 6,40 | 6,45 | 2,550 |
| 5 | 7,2 | 26,0 | 6,50 | 6,625 | 0,575 |
| 6 | 4,8 | 27,0 | 6,75 | 6,875 | -2,075 |
| 7 | 6,0 | 28,0 | 7,00 | 7,1 | -1,100 |
| 8 | 10,0 | 28,8 | 7,20 | 7,3 | 2,700 |
| 9 | 8,0 | 29,6 | 7,40 | 7,45 | 0,550 |
| 10 | 5,6 | 30,0 | 7,50 | 7,625 | -2,025 |
| 11 | 6,4 | 31,0 | 7,75 | 7,875 | -1,475 |
| 12 | 11,0 | 32,0 | 8,00 | 8,125 | 2,875 |
| 13 | 9,0 | 33,0 | 8,25 | 8,325 | 0,675 |
| 14 | 6,6 | 33,6 | 8,40 | 8,375 | -1,775 |
| 15 | 7,0 | 33,4 | 8,35 | | |
| 16 | 10,8 | | | | |

***Шаг 2***. Найдем оценки сезонной компоненты как разность между фактическими уровнями ряда (колонка 2 таблицы 1) и центрирован-

ными скользящими средними (колонка 5). Эти значения помещаем в колонку 6 таблицы 1 и используем для расчета значений сезонной компоненты (таблица 2), которые представляют собой средние за каждый квартал (по всем годам) оценки сезонной компоненты $S_i$. В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период (в данном случае – за год) взаимопогашаются. В аддитивной модели это выражается в том, что сумма значений сезонной компоненты по всем точкам (здесь – по четырем кварталам) должна быть равна нулю.

Таблица 2.

| Показатели | Год | № квартала, $i$ | | | |
|---|---|---|---|---|---|
| | | I | II | III | IV |
| | 1 | - | - | -1,250 | 2,550 |
| | 2 | 0,575 | -2,075 | -1,100 | 2,700 |
| | 3 | 0,550 | -2,025 | -1,475 | 2,875 |
| | 4 | 0,675 | -1,775 | - | - |
| Итого за $i$ – й квартал (за все годы) | | 1,800 | -5,875 | -3,825 | 8,125 |
| Средняя оценка сезонной компоненты для $i$ – го квартала, $\overline{S}_i$ | | 0,600 | -1,958 | -1,275 | 2,708 |
| Скорректированная сезонная компонента, $S_i$ | | 0,581 | -1,977 | -1,294 | 2,690 |

Для данной модели сумма средних оценок сезонной компоненты равна:

$$0,6-1,958-1,275+2,708=0,075.$$

Эта сумма оказалась не равной нулю, поэтому каждую оценку уменьшим на величину поправки, равной одной четверти полученного значения:

$$\Delta=0,075/4=0,01875.$$

Рассчитаем скорректированные значения сезонной компоненты (они записаны в последней строке таблицы 2):

$$S_i = \overline{S}_i - \Delta, \quad i = \overline{1,4} \tag{8}$$

Эти значения при суммировании уже равны нулю:

$$0,581-1,977-1,294+2,69=0.$$

***Шаг 3***. Исключаем влияние сезонной компоненты, вычитая её значения из каждого уровня исходного временного ряда. Получаем величины:

$$T+E=Y-S \tag{9}$$

Эти значения рассчитываются в каждый момент времени и содержат только тенденцию и случайную компоненту (колонка 4 следующей таблицы):

Таблица 3.

| $t$ | $y_t$ | $S_t$ | $T+E=y_t-S_i$ | $T$ | $T+S$ | $E=y_t-(T+S)$ | $E^2$ |
|---|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| 1 | 6,0 | 0,581 | 5,419 | 5,902 | 6,483 | -0,483 | 0,2332 |
| 2 | 4,4 | -1,977 | 6,377 | 6,088 | 4,111 | 0,289 | 0,0833 |
| 3 | 5,0 | -1,294 | 6,294 | 6,275 | 4,981 | 0,019 | 0,0004 |
| 4 | 9,0 | 2,69 | 6,310 | 6,461 | 9,151 | -0,151 | 0,0228 |
| 5 | 7,2 | 0,581 | 6,619 | 6,648 | 7,229 | -0,029 | 0,0008 |
| 6 | 4,8 | -1,977 | 6,777 | 6,834 | 4,857 | -0,057 | 0,0032 |
| 7 | 6,0 | -1,294 | 7,294 | 7,020 | 5,726 | 0,274 | 0,0749 |
| 8 | 10,0 | 2,69 | 7,310 | 7,207 | 9,897 | 0,103 | 0,0107 |
| 9 | 8,0 | 0,581 | 7,419 | 7,393 | 7,974 | 0,026 | 0,0007 |
| 10 | 5,6 | -1,977 | 7,577 | 7,580 | 5,603 | -0,003 | 0,0000 |
| 11 | 6,4 | -1,294 | 7,694 | 7,766 | 6,472 | -0,072 | 0,0052 |
| 12 | 11,0 | 2,69 | 8,310 | 7,952 | 10,642 | 0,358 | 0,1278 |
| 13 | 9,0 | 0,581 | 8,419 | 8,139 | 8,720 | 0,280 | 0,0785 |
| 14 | 6,6 | -1,977 | 8,577 | 8,325 | 6,348 | 0,252 | 0,0634 |
| 15 | 7,0 | -1,294 | 8,294 | 8,512 | 7,218 | -0,218 | 0,0474 |
| 16 | 10,8 | 2,69 | 8,110 | 8,698 | 11,388 | -0,588 | 0,3458 |

***Шаг 4***. Определим трендовую компоненту данной модели. Для этого проведем выравнивание ряда ($T$+$E$) с помощью линейного тренда:

$$T=5{,}715+0{,}186\cdot t; \quad R^2=0{,}91497.$$

Подставляя в это уравнение значения $t=1,2,\ldots,16$, найдем уровни $T$ для каждого момента времени (колонка 5 таблицы 3).

***Шаг 5***. Найдем значения уровней ряда, полученные по аддитивной модели. Для этого прибавим к уровням $T$ значения сезонной компоненты для соответствующих кварталов, т.е. к значениям в колонке 5 таблицы 3 прибавим значения в колонке 3. Результаты операции представлены в колонке 6 таблицы 3.

***Шаг 6***. В соответствии с методикой построения аддитивной модели расчет ошибки производим по формуле:

$$E=Y-(T+S) \tag{10}$$

Это абсолютная ошибка. Численные значения абсолютных ошибок приведены в колонке 7 таблицы 3.

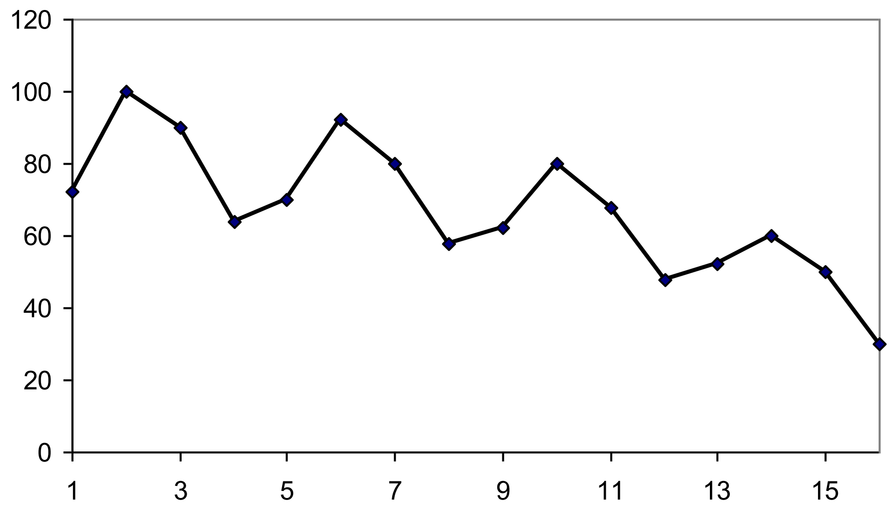
По аналогии с моделью регрессии для оценки качества построения модели или для выбора наилучшей модели можно применять сумму квадратов полученных абсолютных ошибок. Для данной аддитивной модели сумма квадратов абсолютных ошибок равна 1,10. По отношению к общей сумме квадратов отклонений уровней ряда от его среднего уровня, равной 71,59, эта величина составляет чуть более 1,5%. Следовательно, можно сказать, что аддитивная модель объясняет 98,5% общей вариации уровней временного ряда потребления электроэнергии за последние 16 кварталов.

**Пример**. ***Построение мультипликативной модели временного ряда***. Пусть имеются поквартальные данные о прибыли компании за последние четыре года:

Таблица 4.

| Квартал<br>Год | I | II | II | IV |
|---|---|---|---|---|
| 1 | 72 | 100 | 90 | 64 |
| 2 | 70 | 92 | 80 | 58 |
| 3 | 62 | 80 | 68 | 48 |
| 4 | 52 | 60 | 50 | 30 |

График временного ряда свидетельствует о наличии сезонных колебаний периодичностью 4 квартала и общей убывающей тенденции уровней ряда:




Прибыль компании в весенне – летний период выше, чем в осенне – зимний период. Поскольку амплитуда сезонных колебаний уменьшается, можно предположить существование мультипликативной модели. Определим её компоненты.

**Шаг 1**. Проведем выравнивание исходных уровней ряда методом скользящей средней. Методика, применяемая на этом шаге, полностью совпадает с методикой аддитивной модели. Результаты расчетов оценок сезонной компоненты представлены в таблице:

Таблица 5.

| № квартала | Прибыль компании | Итого за четыре квартала | Скользящая средняя за четыре квартала | Центрированная скользящая средняя | Оценка сезонной компоненты |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 72 | | | | |
| 2 | 100 | | | | |
| 3 | 90 | 326 | 81,500 | 81,250 | 1,108 |
| 4 | 64 | 324 | 81,000 | 80,000 | 0,800 |
| 5 | 70 | 316 | 79,000 | 77,750 | 0,900 |
| 6 | 92 | 306 | 76,500 | 75,750 | 1,215 |
| 7 | 80 | 300 | 75,000 | 74,000 | 1,081 |
| 8 | 58 | 292 | 73,000 | 71,500 | 0,811 |
| 9 | 62 | 280 | 70,000 | 68,500 | 0,905 |
| 10 | 80 | 268 | 67,000 | 65,750 | 1,217 |
| 11 | 68 | 258 | 64,500 | 63,250 | 1,075 |
| 12 | 48 | 248 | 62,000 | 59,500 | 0,807 |
| 13 | 52 | 228 | 57,000 | 54,750 | 0,950 |
| 14 | 60 | 210 | 52,500 | 50,250 | 1,194 |
| 15 | 50 | 192 | 48,000 | | |
| 16 | 30 | | | | |

**Шаг 2**. Найдем оценки сезонной компоненты как частное от деления фактических уровней ряда на центрированные скользящие средние (колонка 6 таблицы). Используем эти оценки для расчета значений сезонной компоненты $S$. Для этого найдем средние за каждый квартал оценки сезонной компоненты $S_i$. Взаимопогашаемость сезонных воздействий в мультипликативной модели выражается в том, что сумма значений сезонной компоненты по всем кварталам должна равняться числу периодов в цикле. В нашем случае число периодов одного цикла (год) равно четырем кварталам. Результаты расчетов сведем в таблицу:

Таблица 6.

| Показатели | Год | № квартала, $i$ | | | |
|---|---|---|---|---|---|
| | | I | II | III | IV |
| | 1 | - | - | 1,108 | 0,800 |
| | 2 | 0,900 | 1,215 | 1,081 | 0,817 |
| | 3 | 0,905 | 1,217 | 1,075 | 0,807 |
| | 4 | 0,950 | 1,194 | - | - |
| Итого за $i$ – й квартал (за все годы) | | 2,755 | 3,626 | 3,264 | 2,424 |
| Средняя оценка сезонной компоненты для $i$ – го квартала, $\bar{S}_i$ | | 0,918 | 1,209 | 1,088 | 0,808 |
| Скорректированная сезонная компонента, $S_i$ | | 0,913 | 1,202 | 1,082 | 0,803 |

Здесь сумма средних оценок сезонных компонент по всем четырем кварталам

$$0{,}918 + 1{,}209 + 1{,}088 + 0{,}808 = 4{,}023.$$

не равна четырем. Чтобы эта сумма равнялась четырем, умножим каждое слагаемое на поправочный коэффициент

$$k = 4/4{,}023 = 0{,}9943,$$

т.е.

$$S_i = \bar{S}_i \cdot k, \quad i = \overline{1{,}4} \qquad (11)$$

Значения скорректированных сезонных компонент записаны в последней строке таблицы 6. Теперь их сумма равна четырем. Занесем эти значения в новую таблицу (колонка 3 таблицы 7):

Таблица 7.

| $t$ | $y_t$ | $S_i$ | $T \cdot E = = Y_t / S_i$ | $T$ | $T \cdot S$ | $E = y_t : :(T \cdot S)$ | $E = y_t - -(T \cdot S)$ | $E^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 72 | 0,913 | 78,86 | 87,80 | 80,16 | 0,898 | -8,165 | 66,66 |
| 2 | 100 | 1,202 | 83,19 | 85,03 | 102,20 | 0,978 | -2,204 | 4,86 |
| 3 | 90 | 1,082 | 83,18 | 82,25 | 89,00 | 1,011 | 1,002 | 1,00 |
| 4 | 64 | 0,803 | 79,70 | 79,48 | 63,82 | 1,003 | 0,179 | 0,03 |
| 5 | 70 | 0,913 | 76,67 | 76,70 | 70,03 | 1,000 | -0,030 | 0,00 |
| 6 | 92 | 1,202 | 76,54 | 73,93 | 88,86 | 1,035 | 3,139 | 9,85 |
| 7 | 80 | 1,082 | 73,94 | 71,15 | 76,99 | 1,039 | 3,013 | 9,08 |
| 8 | 58 | 0,803 | 72,23 | 68,38 | 54,91 | 1,056 | 3,093 | 9,57 |
| 9 | 62 | 0,913 | 67,91 | 65,60 | 59,90 | 1,035 | 2,105 | 4,43 |
| 10 | 80 | 1,202 | 66,56 | 62,83 | 75,52 | 1,059 | 4,482 | 20,08 |
| 11 | 68 | 1,082 | 62,85 | 60,05 | 64,98 | 1,047 | 3,024 | 9,14 |
| 12 | 48 | 0,803 | 59,78 | 57,28 | 45,99 | 1,044 | 2,007 | 4,03 |
| 13 | 52 | 0,913 | 56,96 | 54,50 | 49,76 | 1,045 | 2,240 | 5,02 |
| 14 | 60 | 1,202 | 49,92 | 51,73 | 62,18 | 0,965 | -2,176 | 4,73 |
| 15 | 50 | 1,082 | 46,21 | 48,95 | 52,97 | 0,944 | -2,966 | 8,79 |
| 16 | 30 | 0,803 | 37,36 | 46,18 | 37,08 | 0,809 | -7,080 | 50,12 |

***Шаг 3***. Разделим каждый уровень исходного ряда на соответствующие значения сезонной компоненты. Тем самым мы получим величины

$$T \cdot E = Y / S, \qquad (12)$$

которые содержат только тенденцию и случайную компоненту (колонка 4).

***Шаг 4***. Определим трендовую компоненту в мультипликативной модели. Для этого рассчитаем параметры линейного тренда, используя уровни ($T$+$E$). Уравнение тренда имеет вид:

$$T = 90{,}59 - 2{,}773 \cdot t, \quad R^2 = 0{,}9152.$$

Подставляя в это уравнение значения $t = 1,2,\ldots,16$, найдем уровни $T$ для каждого момента времени (колонка 5 таблицы).

***Шаг 5***. Найдем уровни ряда по мультипликативной модели, умножив уровни Т на значения сезонной компоненты для соответствующих кварталов (колонка 6 таблицы).

***Шаг 6***. Расчет ошибок в мультипликативной модели произведем по формуле:

$$E = Y / (T \cdot S). \qquad (13)$$

Численные значения ошибок приведены в колонке 7 таблицы. Для того, чтобы сравнить мультипликативную модель и другие модели временного ряда, можно по аналогии с аддитивной моделью использовать сумму квадратов абсолютных ошибок. Абсолютные ошибки в мультипликативной модели определяются как:

$$E = y_t - (T \cdot S) \tag{14}$$

В данной модели сумма квадратов абсолютных ошибок составляет 207,4. Общая сумма квадратов отклонений фактических уровней этого ряда от среднего значения равна 5023. Таким образом, доля объясненной дисперсии уровней ряда составляет 95,9%.

Прогнозирование по аддитивной или мультипликативной модели временного ряда сводится к расчету будущего значения временного ряда по уравнению модели в виде

$$\hat{y}_t = T + S$$

для аддитивной или

$$\hat{y}_t = T \cdot S$$

для мультипликативной модели.

## Динамические эконометрические модели

Теперь рассмотрим модели временных рядов, где в качестве исходных статистических данных мы располагаем наблюдениями **двух временных рядов** $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$. Целью регрессионного анализа в данном случае является построение линейной регрессионной модели, позволяющей с наименьшими ошибками прогнозировать значения $y_t$ по значениям $x_t$ для $t>n$.

Подобные модели естественны в ситуациях, когда две переменные $x$ и $y$ связаны так, что воздействия единовременного изменения одной из них ($x$) на другую ($y$) сказывается в течение достаточно продолжительного времени, т.е. наблюдается распределенный во времени эффект воздействия. В частности, такие связи возникают между регистрируемыми во времени входными и выходными характеристиками процессов накопления и распределения ресурсов (например, процессов преобразования доходов населения в его расходы) или процессов трансформации затрат в результаты (например, процессов воспроизводства основных доходов).

Эконометрическая модель является динамической, если в данный момент $t$ она учитывает значения входящих в неё переменных, относящихся как к текущему, так и к предыдущим моментам времени, т.е. модель учитывает, отражает динамику исследуемых переменных в каждый момент времени.

Переменные, влияние которых характеризуется определенным запаздыванием, называются **лаговыми переменными**.

Классифицируются динамические модели по – разному. Один из вариантов классификации следующий:

1. ***Модели с распределенными лагами***. Они содержат в качестве лаговых переменных лишь независимые (объясняющие) переменные, например:
$$y = a + b_0 x_t + b_1 x_{t-1} + \ldots + b_p x_{t-p} + \varepsilon_t \tag{15}$$
2. ***Авторегрессионные модели***, уравнения которых включают в качестве объясняющих переменных лаговые значения зависимых переменных, например:
$$y = a + b x_t + c_1 y_{t-1} + c_2 y_{t-2} + \varepsilon_t \tag{16}$$

Рассмотрим модель (15), приняв, что $p$ – конечное число. Модель говорит о том, что если в некоторый момент времени $t$ происходит изменение $x$, это изменение будет влиять на значение $y$ в течение $p$ последующих моментов времени. Коэффициент $b_0$ называется **краткосрочным мультипликатором**, т.к. он характеризует изменение среднего значения $y$ при единичном изменении $x$ в тот же самый момент времени. Сумма $\sum_{j=0}^{p} b_j$ называется **долгосрочным мультипликатором**; он характеризует изменение $y$ под воздействием единичного изменения $x$ в каждом из моментов времени. Любая сумма $\sum_{j=0}^{k} b_j \ (k < p)$ называется **промежуточным мультипликатором**.

**Относительные коэффициенты** модели (15) с распределенным лагом определяются выражениями:

$$\beta_j = \frac{b_j}{\sum_{j=0}^{p} b_j}; \quad \sum_{j=0}^{p} \beta_j = 1 \tag{17}$$

(условие нормировки имеет место, только если все $b_j$ имеют одинаковые знаки). Значения $\beta_j$ являются **весами** для соответствующих коэффициентов $b_j$. Каждый из них измеряет долю общего изменения $y$, приходящегося на момент $(t+j)$.

**Средний лаг** определяется по формуле средней арифметической взвешенной:

$$\bar{l} = \sum_{j=0}^{p} j \cdot \beta_j \qquad (18)$$

Он означает период, в течение которого происходит изменение результата от изменения $x$ в момент $t$. Небольшая величина (18) означает быструю реакцию $y$ на изменение $x$, высокое значение говорит о том, что воздействие фактора $y$ будет сказываться в течение длительного времени.

**Медианный лаг** – это величина лага, для которого

$$\sum_{j=0}^{l_{Me}} \beta_j \approx 0{,}5 \qquad (19)$$

Это время, в течение которого с момента $t$ будет реализована половина общего воздействия фактора на результат.

Рассмотрим условный пример. Предположим, модель зависимости объёмов продаж компании от расходов на рекламу имеет вид:

$$\hat{y}_t = -0{,}67 + 4{,}5x_t + 3x_{t-1} + 1{,}5x_{t-2} + 0{,}5x_{t-3}.$$

Краткосрочный мультипликатор равен 4,5: увеличение расходов на рекламу на 1 млн. руб. приводит к среднему росту продаж компании на 4,5 млн. руб. в том же периоде.

В момент $(t+1)$ такой рост составит 4,5+3,0=7,5 млн. руб., в момент $(t+2)$ - 7,5+1,5=9 млн. руб. и т.д. долгосрочный мультипликатор равен 9,5. В долгосрочной перспективе (в течение 3 месяцев) увеличение расходов на 1 млн. руб. приведет к общему росту продаж на 9,5 млн. руб.

Относительные коэффициенты:

$$\beta_0 = \frac{4{,}5}{9{,}5} = 0{,}474; \quad \beta_1 = \frac{3}{9{,}5} = 0{,}316; \quad \beta_2 = \frac{1{,}5}{9{,}5} = 0{,}158; \quad \beta_3 = \frac{0{,}5}{9{,}5} = 0{,}053.$$

47,4% общего увеличения объёма продаж от роста затрат на рекламу происходит в текущем месяце, 31,6% - в следующем месяце и т.д.

Средний лаг равен:

$$\bar{l} = 0 \cdot 0{,}474 + 1 \cdot 0{,}316 + 2 \cdot 0{,}158 + 3 \cdot 0{,}053 = 0{,}791 \text{ (мес.)}$$

- небольшая величина, поскольку большая часть эффекта роста затрат на рекламу проявляется сразу же. Медианный лаг в данном примере составляет чуть более 1 месяца.

Модель (15) можно свести к уравнению множественной регрессии через замены переменных:

$$x_0^* = x_t, \quad x_1^* = x_{t-1}, \ldots, x_p^* = x_{t-p}, \tag{20}$$

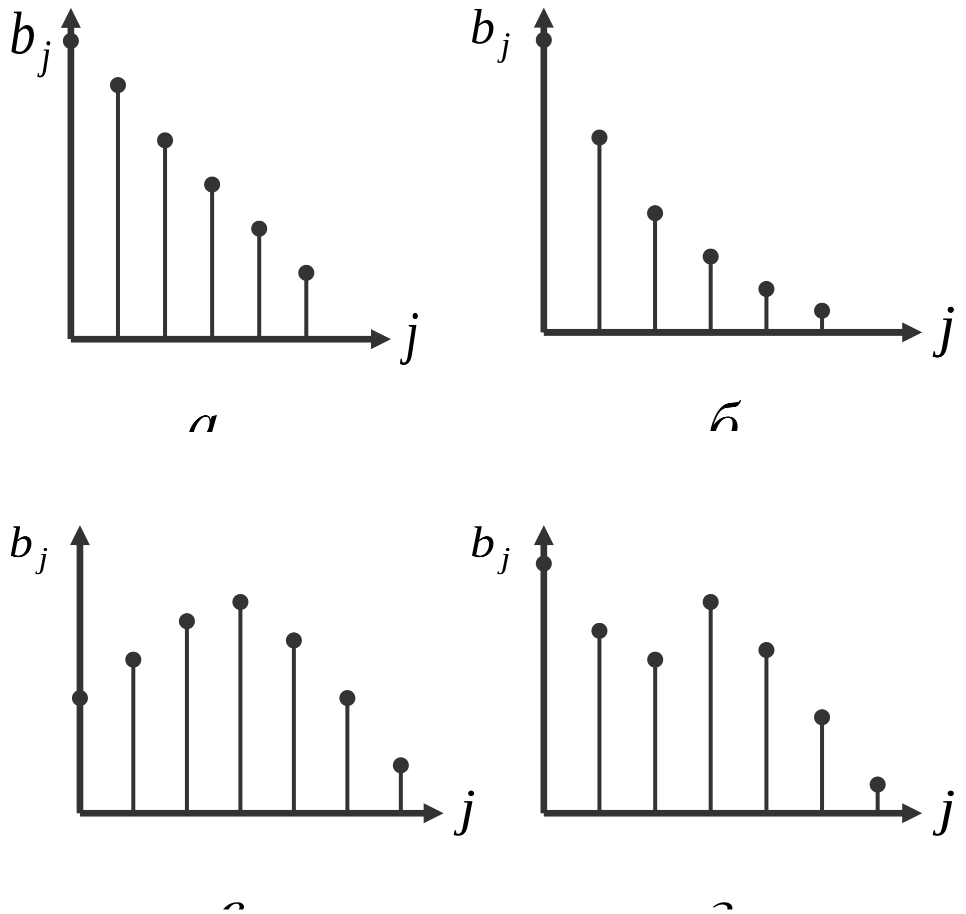
в результате получаем:

$$y = a + b_0 x_1^* + b_1 x_2^* + \ldots + b_p x_p^* + \varepsilon_t \tag{21}$$

Однако применение обычного МНК затруднительно по следующим причинам:

1. Текущие и лаговые значения $x$ тесно связаны между собой, что приводит к высокой мультиколлинеарности факторов.
2. При большой величине лага велико число параметров, что приводит к уменьшению числа степеней свободы.
3. Часто возникает проблема автокорреляции остатков.

Поэтому оценки параметров становятся неточными и неэффективными. Для получения более обоснованных оценок нужна информация о структуре лага. Эта структура может быть различной. На рисунке представлены некоторые её формы:



Если с ростом величины лага коэффициенты при лаговых переменных убывают, то имеет место линейная (или треугольная) структура лага (*а*), а также геометрическая структура (*б*). Возможны и другие структуры лага (*в* или *г*). Рассмотрим некоторые подходы к расчету лагов.

**Лаги Алмон**. Предполагается, что в модели (15) с конечной максимальной величиной лага $p$ значения коэффициентов $b_j$ описываются полиномом $k$ – й степени:

$$b_j = c_0 + c_1 j + c_2 j^2 + \ldots + c_k j^k \qquad (22)$$

Каждый коэффициент, таким образом, запишется так:

$$\begin{aligned} & b_0 = c_0, \\ & b_1 = c_0 + c_1 + \ldots + c_k, \\ & b_2 = c_0 + 2c_1 + \ldots + 2^k c_k, \\ & \ldots \\ & b_p = c_0 + p \cdot c_1 + \ldots + p^k c_k. \end{aligned} \qquad (23)$$

Подставим эти соотношения в (15) и перегруппируем слагаемые, получим:

$$y_t = a + c_0 \sum_{j=0}^{p} x_{t-j} + c_1 \sum_{j=0}^{p} j \cdot x_{t-j} + c_2 \sum_{j=0}^{p} j^2 \cdot x_{t-j} + \ldots + c_k \sum_{j=0}^{p} j^k \cdot x_{t-j} + \varepsilon_t \quad 24)$$

Обозначим суммы соответственно как новые переменные $z_0, z_1, \ldots, z_k$, перепишем (24) в виде:

$$y_t = a + c_0 z_0 + c_1 z_1 + c_2 z_2 + \ldots + c_k z_k + \varepsilon_t \qquad (25)$$

Параметры $c_j$ определяются по МНК.

Достоинства метода:

1. Универсальность, применимость для моделирования процессов с разнообразными структурами лагов.
2. При малых $k$ (2 или 3) можно построить модели с распределенным лагом любой длины.

Ограничения метода:

1. Величина $p$ должна быть известна заранее. При этом приходится задавать максимально возможную величину лага. Выбор меньшего лага, чем его реальное значение, приведет к неверной спецификации модели, невозможности обеспечить случайность остатков, поскольку влияние значимых факторов будет выражено в остатках. Оценки параметров при этом окажутся неэффективными и сме-

щенными. Включение в модель большей величины лага, чем его реальное значение, снижает эффективность оценок из – за наличия статистически незначимых факторов.

2. Необходимость установить степень полинома. Обычно принимают $k$=2 или 3 по правилу: степень полинома $k$ должна быть на единицу больше числа экстремумов в структуре лага. В крайнем случае $k$ определяется из сравнения моделей для различных $k$.
3. Возможна мультиколлинеарность факторов $z_j$, однако она сказывается здесь в меньшей степени, чем в модели (15).

**Метод Койка**. Этот метод применяется в модели с бесконечным лагом:

$$y_t = a + b_0 x_t + b_1 x_{t-1} + b_2 x_{t-2} + \ldots + \varepsilon_t \tag{26}$$

Здесь обычный МНК применить нельзя. Для идентификации модели (26) предполагается, что параметры с увеличением лага убывают в геометрической прогрессии, т.е. с постоянным темпом $\lambda \in (0,1)$:

$$y_t = a + b_0 x_t + b_0 \lambda x_{t-1} + b_0 \lambda^2 x_{t-2} + \ldots + \beta_0 \lambda^m x_{t-m} + \ldots + \varepsilon_t \tag{27}$$

Запишем выражение (27) для момента $(t\text{-}1)$:

$$y_{t-1} = a + b_0 x_{t-1} + b_0 \lambda x_{t-2} + b_0 \lambda^2 x_{t-3} + \ldots + \varepsilon_t \tag{28}$$

Умножим (28) на $\lambda$ и вычтем из (27):

$$y_t = a(1-\lambda) + b_0 x_t + \lambda y_{t-1} + \varepsilon_t - \lambda \varepsilon_{t-1}$$

или

$$y_t = a(1-\lambda) + b_0 x_t + \lambda y_{t-1} + u_t \tag{29}$$

Это модель авторегрессии. Определив её параметры, находим $\lambda$, $a$, $b_0$ исходной модели, а затем и параметры $b_j = \lambda^j b_0,\ j = 1,2,3,\ldots$. Данная модель позволяет определить долгосрочный мультипликатор $\sum_{j=0}^{\infty} b_j = b_0 \dfrac{1}{1-\lambda}$ и средний лаг $\bar{l} = \dfrac{\lambda}{1-\lambda}$.

Теперь перейдем к рассмотрению авторегрессионных моделей.

**Модель адаптивных ожиданий**. Моделирование ожиданий является сложной задачей, поскольку фактор ожидания имеет качественную специфику. Например, инвестиции связаны не только с нормой процента, но и с ожиданиями инвесторов. Если в стране существенная безработица, то действия правительства в направлении стимулирования могут рассматриваться как позитивные, и это способствует инвестициям. Если экономика близка к полной занятости, то та же самая политика будет рассматриваться как ведущая к росту инфляции и приведет к падению инвестиционной активности.

Модель адаптивных ожиданий заключается в простой процедуре корректировки ожиданий, когда в каждый момент времени реальное значение переменной сравнивается с её ожидаемым значением. Если реальное значение оказывается больше, то ожидаемое в следующий момент значение корректируется в сторону его увеличения, если меньше – то в сторону уменьшения. Предполагается, что размер корректировки пропорционален разности между реальным и ожидаемым значениями переменной. Таким образом, основную идею можно записать формулой:

$$x_{t+1}^e - x_t^e = \lambda \left(x_t - x_t^e\right) \quad \left(0 \le \lambda \le 1\right) \tag{30}$$

где $x_t^e$ - значение $x$, ожидаемое в момент $t$ (expected). Это выражение можно переписать в форме взвешенного среднего:

$$x_{t+1}^e = \lambda x_t + \left(1-\lambda\right) x_t^e \tag{31}$$

Модель (30) и является моделью адаптивных ожиданий. Это выражение иногда называют моделью обучения на ошибках, т.к. ожидания экономических объектов складываются из прошлых ожиданий, поправленных на величину ошибки в ожиданиях, допущенных ранее.

При $\lambda$=0 ожидания являются статичными, неизменными, т.е. $x_{t+1}^e = x_t^e$.

При $\lambda$=1 ожидания реализуются мгновенно, т.е. $x_{t+1}^e = x_t$.

Чем больше $\lambda$, тем быстрее ожидаемое значение адаптируется к предыдущим реальным значениям переменной.

Долгосрочная функция модели адаптивных ожиданий записывается в виде:

$$y_t = a + b x_{t+1}^e + \varepsilon_t \tag{32}$$

Подставим сюда выражение (31), получим:

$$y_t = a + \lambda b x_t + \left(1-\lambda\right) b x_t^e + \varepsilon_t \tag{33}$$

Запишем его для ($t$-$1$):

$$y_{t-1} = a + \lambda b x_{t-1} + \left(1-\lambda\right) b x_{t-1}^e + \varepsilon_{t-1} \tag{34}$$

Умножим (34) на ($1$-$\lambda$) и вычтем почленно из (32):

$$y_t = \lambda a + \lambda b x_t + \left(1-\lambda\right) b y_{t-1} + u_t, \tag{35}$$

где $u_t = \varepsilon_t - \left(1-\lambda\right)\varepsilon_{t-1}$.

Это модель авторегрессии, в которой все переменные имеют фактические, а не ожидаемые значения. Модель в форме (35) называется краткосрочной функцией модели адаптивных ожиданий.

**Модель неполной (частичной) корректировки**. Здесь поведенческое уравнение определяет не фактическое значение $y_t$, а её желаемый (целевой) уровень $y_t^*$:

$$y_t^* = a + bx_t + u_t \tag{36}$$

Примером такой модели служит политика компаний относительно распределения дивидендов: прибыль расходуется частично на уплату дивидендов, частью на инвестиции. Когда прибыль увеличивается, дивиденды тоже растут, но не в той же пропорции (это объясняется желанием руководства фирмы в любом случае не уменьшать дивиденды, т.к. это ударяет по репутации фирмы).

В модели предполагается, что фактическое приращение зависимой переменной пропорционально разнице между её желаемым уровнем и значением в предыдущий период:

$$y_t - y_{t-1} = \lambda\left(y_t^* - y_{t-1}\right) + v_t \tag{37}$$

($v_t$ – случайный член). Это выражение можно переписать так:

$$y_t = \lambda y_t^* + \left(1 - \lambda\right) y_{t-1} + v_t, \tag{38}$$

т.е. в форме взвешенного среднего. Чем больше $\lambda$, тем быстрее идет корректировка. При $\lambda$ =1 полная корректировка происходит за один период. При $\lambda$ =0 корректировка не происходит совсем.

Подставим (36) в (38), получим:

$$y_t = a\lambda + b\lambda x_t + \left(1 - \lambda\right) y_{t-1} + v_t + \lambda u_t \tag{39}$$

Это и есть модель частичной корректировки, которая также является моделью авторегрессии.

Несколько слов об оценке параметров уравнений авторегрессии. Рассмотрим уравнение:

$$y_t = a + bx_t + cy_{t-1} + \varepsilon_t \tag{40}$$

Во всех рассмотренных выше моделях стоит проблема оценивания параметров. Обычный МНК чаще всего даёт смещенные и несостоятельные оценки, вследствие автокорреляции между случайными отклонениями $\varepsilon_t$ и $\varepsilon_{t-1}$ и корреляции между $y_{t-1}$ и $\varepsilon_t$.

Один из возможных методов расчета параметров – метод инструментальных переменных, состоящий в замене $y_{t-1}$ на новую переменную, которая тесно коррелирует с $y_{t-1}$, но не коррелирует с остатками. Это можно сделать двумя способами.

1. Провести регрессию

$$y_{t-1} = \delta_0 + \delta_1 x_{t-1} + u_t, \tag{41}$$

или

$$y_{t-1} = \hat{y}_{t-1} + u_t$$

и подставить $\hat{y}_{t-1}$ в уравнение авторегрессии, получаем:

$$y_t = a + bx_t + c\hat{y}_{t-1} + \nu_t, \tag{42}$$

и далее применяем обычный МНК.

2. Подставим (41) в (40), получим модель с распределенным лагом:

$$y_t = (a + c\delta_0) + bx_t + c\delta_1 x_{t-1} + (cu_t + \varepsilon_t), \tag{43}$$

для которой не нарушаются предпосылки обычного МНК.