

МИНИСТЕРСТВО ПУТЕЙ СООБЩЕНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ПУТЕЙ СООБЩЕНИЯ (МИИТ)

Кафедра “Прикладная математика–2”

Л.Ф.Кочнева, А.С.Милевский

ЭКОНОМЕТРИКА.
ЧАСТЬ 2. МНОЖЕСТВЕННАЯ РЕГРЕССИЯ

Учебное пособие

для экономических специальностей

Москва – 2006

Классическая модель множественной линейной регрессии

4.1 Основные гипотезы

Модель множественной линейной регрессии имеет вид

$$y = \beta_0 + \beta_1 X_1 + \dots + \beta_m X_m + \varepsilon$$

где y – зависимая переменная, X_1, \dots, X_m – объясняющие переменные (также называемые регрессорами), $\beta_0, \beta_1, \dots, \beta_m$ – коэффициенты регрессии, ε – случайный член (или “ошибки измерений”). Для нахождения оценок коэффициентов проводятся n наблюдений ($n > m$), в результате чего получается n равенств:

$$y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_m X_{mi} + \varepsilon_i \quad (i=1, 2, \dots, n)$$

(индекс i показывает номер наблюдения). Это удобно записать в матричной форме:

$$Y = X \cdot \beta + \varepsilon, \quad (4.1)$$

где

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_{11} & \dots & x_{m1} \\ 1 & x_{12} & \dots & x_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \dots & x_{mn} \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Гипотезы, лежащие в основе классической модели множественной регрессии, обобщают соответствующие предположения модели парной регрессии:

- (A) “Истинная” зависимость Y от регрессоров X имеет вид (4.1).
- (B) Величины X – детерминированные (не случайные).
- (C) Столбцы матрицы X линейно независимы (другими словами, ранг матрицы X равен $m+1$).
- (D) $M\varepsilon_i = 0$, $D\varepsilon_i = \sigma^2$ не зависит от i

(E) $\text{COV}(\varepsilon_i, \varepsilon_j) = 0$ при $i \neq j$ (некоррелированность ошибок для разных наблюдений).

Замечание. Вместо (D) часто добавляют условие

(F) Ошибки ε_i имеют нормальное распределение $N(0, \sigma^2)$.

В этом случае модель называется *нормальной*.

Замечание. Условия (D) – (F) удобно записать в матричной форме:

(G) $M\varepsilon = 0$, $M(\varepsilon\varepsilon^T) = \sigma^2 E$ (E – единичная матрица)
 (F') $\varepsilon \sim N(0, \sigma^2 E)$

Пример.

В таблице представлены объём импорта (Y) и ВНП (X_1) и индекс потребительских цен (X_2) в США.

Годы	1964	1965	1966	1967	1968
X_1	636	689	753	796	868
X_2	93	95	97	100	104
Y	28	32	38	41	53

Построим уравнение линейной регрессии Y на X_1, X_2

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2.$$

В матричной записи

$$Y = \begin{pmatrix} 28 \\ 32 \\ 38 \\ 41 \\ 53 \end{pmatrix}, X = \begin{pmatrix} 1 & 636 & 93 \\ 1 & 689 & 95 \\ 1 & 753 & 97 \\ 1 & 796 & 100 \\ 1 & 868 & 104 \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

$$X^T X = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 636 & 689 & 753 & 796 & 868 \\ 93 & 95 & 97 & 100 & 104 \end{pmatrix} \cdot \begin{pmatrix} 1 & 636 & 93 \\ 1 & 689 & 95 \\ 1 & 753 & 97 \\ 1 & 796 & 100 \\ 1 & 868 & 104 \end{pmatrix} =$$

$$= \begin{pmatrix} 5 & 3742 & 489 \\ 3742 & 2833266 & 367516 \\ 489 & 367516 & 47899 \end{pmatrix}; (X^T X)^{-1} = \begin{pmatrix} 2452,8 & 1,822 & -39,02 \\ 1,822 & 0,0014 & -0,03 \\ -39,02 & -0,03 & 0,625 \end{pmatrix}$$

$$X^T Y = \begin{pmatrix} 192 \\ 147110 \\ 18942 \end{pmatrix}; \hat{\beta} = (X^T X)^{-1} X^T Y = \begin{pmatrix} -150,99 \\ 0,02 \\ 1,78 \end{pmatrix}$$

$$y = -150,99 + 0,02x_1 + 1,78x_2$$

4.2 Теорема Гаусса-Маркова

Будем искать коэффициенты β в задаче (4.1) *методом наименьших квадратов*. Другими словами, коэффициенты должны удовлетворять условию

$$\sum (y_i - \hat{y}_i)^2 \rightarrow \min; \quad \hat{y}_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots$$

или, в матричной форме,

$$(Y - X\beta)^T \cdot (Y - X\beta) \rightarrow \min$$

Если продифференцировать это равенство по β и приравнять производные нулю, получим

$$X^T X\beta = X^T Y$$

Отсюда

$$\boxed{\hat{\beta} = (X^T X)^{-1} X^T Y}, \quad (4.2)$$

если существует обратная матрица.

Теорема (Гаусса-Маркова). В предположениях (А) – (Е) оценки, полученные методом наименьших квадратов по формуле (4.2), являются несмещёнными и обладают наименьшей дисперсией среди всех линейных несмещённых оценок параметров β .

Доказательство. Докажем только несмещённость.

$$M\hat{\beta} = M(X^T X)^{-1} X^T Y = M(X^T X)^{-1} X^T (X\beta + \varepsilon) =$$

$$= (X^T X)^{-1} X^T M(X\beta + \varepsilon) = (X^T X)^{-1} X^T X\beta = \beta,$$

так как $M\varepsilon = 0$

4.3 Коэффициент детерминации

Дисперсия зависимой переменной (или *полная сумма квадратов*) обозначается TSS:

$$TSS = \sum (y_i - \bar{y})^2 = nD_y^* \quad (4.3)$$

Сумма

$$RSS = \sum (\hat{y}_i - \bar{y})^2 = nD_{\hat{y}}^* \quad (4.4)$$

определяет разброс оценки \hat{y} относительно среднего, или “объяснённую часть дисперсии”. И, наконец, сумма квадратов “остатков”

$$ESS = \sum (y_i - \hat{y}_i)^2 = \sum e_i^2 = nD_e^* \quad (4.5)$$

представляет собой “необъяснённую часть дисперсии” зависимой переменной. Нетрудно проверить, что

$$RSS + ESS = TSS, \quad (4.6)$$

Величина

$$R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS} \quad (4.7)$$

называется *коэффициентом детерминации* и представляет собой долю дисперсии зависимой переменной y , объяснённую при помощи оцененного уравнения регрессии. Эта величина принимает зна-

чения в пределах от 0 до 1. Крайний случай $R^2=1$ означает, что $ESS=0$, т.е. все точки наблюдений лежат на регрессионной плоскости.

Замечание. В модели *множественной линейной регрессии*, кроме β , есть ещё один параметр – σ^2 . Его оценку можно найти по формуле

$$\hat{\sigma}^2 = S^2 = \frac{ESS}{n - m - 1}, \quad (4.8)$$

Пример. В рассматриваемом примере

$$Y = \begin{pmatrix} 28 \\ 32 \\ 38 \\ 41 \\ 53 \end{pmatrix}, \hat{Y} = \begin{pmatrix} 27,59 \\ 32,22 \\ 37,07 \\ 43,28 \\ 51,85 \end{pmatrix},$$

$$ESS = 7,5906, RSS = 361,609, TSS = 369,24;$$

$$S^2 = \frac{ESS}{5 - 2 - 1} = 3,7953; R^2 = \frac{361,609}{369,24} = 0,9794$$

Замечание. Важное свойство коэффициента детерминации R^2 состоит в том, что при добавлении в модель новых объясняющих переменных X он не может уменьшиться. Поэтому при сравнении двух моделей с различным числом переменных не всегда ясно, за счёт чего возрос показатель детерминации: за счёт реального влияния дополнительного фактора на результат или просто из-за возрастания числа переменных. Для того чтобы можно было *сравнивать различные модели*, вводят так называемый *скорректированный коэффициент детерминации*:

$$R_{корр}^2 = 1 - \frac{ESS/(n - m - 1)}{TSS/(n - 1)} \quad (4.9)$$

Это число также не превосходит 1, но в некоторых случаях может оказаться и отрицательным. В рассматриваемом примере

$$R_{корр}^2 = 0,69589$$

4.4 Статистические свойства оценок коэффициентов

Теорема. В предположениях (A)–(E) и (F) теоремы Гаусса-Маркова оценки, полученные методом наименьших квадратов по формуле (4.2), обладают свойствами:

$$1. \frac{(n - m - 1) \cdot S^2}{\sigma^2} \sim \chi^2(n - m - 1).$$

$$2. S^2 \text{ и } \hat{\beta} \text{ независимы}$$

4.5 Доверительные интервалы

Будем считать, что условие нормальности (E) выполнено. Тогда можно построить *доверительные интервалы* для параметров регрессии, т.е. интервалы, *содержащие точные значения этих параметров с заданной вероятностью $1-\alpha$* . Формулы имеют следующий вид:

$$\hat{\beta}_i - t_{1-\alpha/2}(n - m - 1) \cdot S_{\beta_i} < \beta_i < \hat{\beta}_i + t_{1-\alpha/2}(n - m - 1) \cdot S_{\beta_i} \quad (4.10)$$

$$\frac{ESS}{\chi_{1-\alpha/2}^2(n - m - 1)} < \sigma^2 < \frac{ESS}{\chi_{\alpha/2}^2(n - m - 1)}. \quad (4.11)$$

Здесь α – заданный “уровень значимости”, т.е. *вероятность того, что построенный интервал не содержит истинного значения параметра*.

Далее, $t_{1-\alpha/2}(n-m-1)$, $\chi_{1-\alpha/2}^2(n-m-1)$, $\chi_{\alpha/2}^2(n-m-1)$ – *квантили, соответственно, закона Стьюдента и закона хи-квадрат*. Это – табличные значения.

Наконец, S_{β} –выборочные среднеквадратичные отклонения оценок параметров (или *стандартные ошибки параметров*), они находятся по формулам:

$$S_{\beta_i}^2 = S^2 \cdot (X^T \cdot X)^{-1}_{ii} \quad (4.12)$$

Как правило, стандартные ошибки параметров вычисляются в эконометрических программах вместе с оценками самих параметров.

Пример. В рассматриваемом примере

$$S_{\beta_0}^2 = S^2 \cdot 2452,8 = 9309,16; S_{\beta_0} = 96,48$$

$$S_{\beta_1}^2 = S^2 \cdot 0,0014 = 3,4339; S_{\beta_1} = 0,07$$

$$S_{\beta_2}^2 = S^2 \cdot 0,625 = 1533,01; S_{\beta_2} = 1,54$$

Таким образом,

$$y = -150,99 + 0,02x_1 + 1,78x_2$$

(96,48) (0,07) (1,54)

где в скобках подписаны стандартные ошибки параметров.

Далее, на уровне значимости $\alpha=0,1$:

$$t_{1-\alpha/2}(n-m-1) = t_{0,95}(2) = 2,92;$$

$$-15,99 - 2,92 \cdot 96,48 = -432,72; -15,99 + 2,92 \cdot 96,48 = 130,74$$

$$-432,72 < \beta_0 < 130,74;$$

$$-0,19 < \beta_1 < 0,23;$$

$$-2,71 < \beta_2 < 6,28;$$

$$\chi^2_{1-\alpha/2}(n-m-1) = \chi^2_{0,95}(2) = 5,99;$$

$$\chi^2_{\alpha/2}(n-m-1) = \chi^2_{0,05}(2) = 0,1;$$

$$\frac{7,5906}{5,99} = 1,27 < \sigma^2 < 75,9 = \frac{7,5906}{0,1}$$

4.6 Интервал для прогнозного значения

Рассмотрим модель(4.1) $Y = X \cdot \beta + \varepsilon$. Здесь X – матрица размера $(n \times m+1)$, каждая из n строк которой соответствует набору значений независимых переменных. Предположим, что у нас есть ещё один набор значений независимых переменных $x^{(n+1)} = (1, x_1^{(n+1)}, x_2^{(n+1)}, \dots, x_m^{(n+1)})$. Требуется найти доверительный интервал для соответствующего значения зависимой переменной $y^{(n+1)}$.

$$\hat{y}^{(n+1)} - t_{1-\alpha/2}(n-m-1) \cdot S_y < y^{(n+1)} < \hat{y}^{(n+1)} + t_{1-\alpha/2}(n-m-1) \cdot S_y$$

где

$$\hat{y}^{(n+1)} = \hat{\beta}_0 + \hat{\beta}_1 x_1^{(n+1)} + \dots + \hat{\beta}_m x_m^{(n+1)} = x^{(n+1)} \cdot \hat{\beta}$$

$$S_y = S \cdot \sqrt{1 + x^{(n+1)} \cdot (X^T X)^{-1} \cdot (x^{(n+1)})^T}$$

4.7 Проверка гипотез о параметрах регрессии

В этом параграфе также будем считать, что условие нормальности **(Е)** выполнено.

При исследовании модели приходится проверять гипотезы о равенстве истинных значений параметров регрессии заданным числам. Для этого достаточно построить доверительный интервал и посмотреть, попадает ли в него это заданное число.

Рассмотрим, например, гипотезу **H₀: $\beta_1 = A$** при альтернативной гипотезе **H₁: $\beta_1 \neq A$** . Учитывая формулы (4.4) получаем, что гипотеза **H₀** принимается в случае, если число **A** лежит в интервале

$$(\hat{\beta}_1 - t_{1-\alpha/2}(n-m-1) \cdot S_{\beta_1}; \hat{\beta}_1 + t_{1-\alpha/2}(n-m-1) \cdot S_{\beta_1}).$$

В противном случае принимается гипотеза **H₁**.

Особенно часто проверяют гипотезу **H₀: $\beta_1 = 0$** . Дело в том, что если принимается альтернативная гипотеза

$$\mathbf{H}_1: \beta_1 \neq 0,$$

то говорят, что коэффициент β_i *значим*. Таким образом, получаем следующий критерий значимости параметра β_i при помощи так называемой *t*-статистики:

$$\frac{\hat{\beta}_i}{S_{\beta_i}} \geq t_{1-\alpha/2}(n-m-1) \quad (4.13)$$

Пример. В рассматриваемом примере доверительный интервал для коэффициента β_1 содержит 0, так что коэффициент β_1 *не значим*. Коэффициент β_2 также не значим.

4.8 Проверка гипотезы о значимости уравнения регрессии

При построении эконометрической модели обязательно возникает вопрос: в какой мере можно ей доверять, насколько полученное уравнение *значимо*?

Для проверки значимости уравнения используется *F*-статистика Фишера. Такое распределение вероятности имеет случайная величина

$$F = \frac{(U_1^2 + U_2^2 + \dots + U_m^2)/m}{(W_1^2 + W_2^2 + \dots + W_l^2)/l},$$

где U_i, W_j – независимые случайные величины, распределенные по стандартному нормальному закону. Распределение Фишера обозначается $F(m, l)$, а числа m, l называются *степенями свободы*.

Оказывается, для множественной регрессии величина

$$F = \frac{RSS/m}{ESS/(n-m-1)} \quad (4.14)$$

где n – число наблюдений в выборке, распределена по закону $F(m, n-m-1)$.

Разделив в этих двух последних выражениях числитель и знаменатель на TSS, получим ещё одну формулу для F

$$F = \frac{R^2/m}{(1-R^2)/(n-m-1)}.$$

Для распределения Фишера имеются *таблицы квантилей* $F_{1-\alpha}(m, k)$, соответствующих уровню значимости α .

Пусть выбрано некоторое значение α .

- Если вычисленное значение *F*-статистики оказывается больше $F_{1-\alpha}(m, n-m+1)$, то оцененное уравнение регрессии считается значимым при выбранном α .

Пример. В рассматриваемом примере $F = \frac{361,6094 / 2}{7.5906 / 2} = 47,64$, а

квантиль $F_{0,9}(2; 2)$ по таблице равен 9. Таким образом, оценённое уравнение *значимо на уровне 0,1* (и с большим запасом).

4.9 Частная корреляция

Корреляция между двумя переменными, вычисленная после устранения влияния всех других переменных, называется *частной корреляцией*. Например, *длина волос* может коррелировать с *ростом* человека (чем выше человек, тем короче волосы), однако эта зависимость становится слабой или совсем исчезает, если устранить влияние *пола* наблюдаемых людей, поскольку женщины обычно ниже ростом и чаще имеют более длинные волосы, чем мужчины. *Частный коэффициент корреляции* между переменными X и Y при исключении влияния переменной Z вычисляется по формуле

$$r_{XY|Z} = \frac{r_{XY} - r_{XZ}r_{YZ}}{\sqrt{(1-r_{XZ})^2(1-r_{YZ})^2}}$$

Пример. В рассматриваемом примере *матрицы корреляций* и *частных корреляций* переменных Y, X_1, X_2 имеют вид

$$r = \begin{pmatrix} 1 & 0,98 & 0,99 \\ 0,98 & 1 & 0,99 \\ 0,99 & 0,99 & 1 \end{pmatrix}, \quad r_{частн} = \begin{pmatrix} 1 & 0,189 & 0,633 \\ 0,189 & 1 & 0,632 \\ 0,633 & 0,632 & 1 \end{pmatrix}$$

4.10 Спецификация модели

Спецификация модели включает пункты:

1. определение списка объясняющих и зависимых переменных;

2. выбор функциональной формы модели.
Рассмотрим подробнее первый пункт.

4.10.1 Отбор объясняющих переменных

При определении состава факторов, включаемых в уравнение регрессии, прежде всего руководствуются теоретическими представлениями о взаимосвязях факторов. Однако часто встречается ситуация, когда имеется большое число факторов, но нет априорной модели изучаемого явления, так что неясно, какие именно переменные включать в модель. В этом случае используются различные эвристические процедуры. Приведём простой пример такой процедуры.

Процедура пошагового отбора независимых переменных.

1. Из исходного набора переменных отбирается (включается в модель) переменная, имеющая наибольший по модулю коэффициент корреляции с зависимой переменной Y .
2. На каждом последующем шаге в модель добавляется та из переменных, добавление которой максимально увеличивает скорректированный коэффициент детерминации. Если добавление новых переменных не увеличивает этот коэффициент, процедура считается завершённой.

При отборе переменных также широко используется матрица частных корреляций.

4.10.2 Влияние ошибок спецификации

Неправильная спецификация переменных может произойти из-за

- отбрасывания переменной, которая должна быть в составе объясняющих переменных
- или
- включения переменной, которая не должна быть в составе объясняющих переменных.

Какие последствия имеют ошибки в спецификации переменных?

4.10.2.1 Отбрасывание значимой переменной

Пусть вместо “истинного” уравнения

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

из-за невключения переменной x_2 взято уравнение

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

В такой ситуации:

- МНК-оценка параметров β в общем случае оказывается смещённой;
- стандартные ошибки коэффициентов β имеют неотрицательное смещение;
- оценка дисперсии σ^2 имеет неотрицательное смещение

Отсюда видно, что последствия такой ошибки спецификации достаточно серьёзны.

4.10.2.2 Добавление незначимой переменной

Пусть вместо “истинного” уравнения

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

из-за невключения переменной x_2 взято уравнение

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

В такой ситуации:

- МНК-оценка параметров β оказывается несмещённой;
- стандартные ошибки коэффициентов β имеют неотрицательное смещение;
- оценка дисперсии σ^2 имеет неотрицательное смещение.

Последствия такой ошибки спецификации являются не столь серьёзными, как в предыдущем случае, однако увеличение стандартных ошибок может привести к неверным выводам при проверке гипотез и ухудшению качества интервальных оценок. В частности, может возникнуть статистическая незначимость коэффициентов.

4.11 Фиктивные переменные

Как правило, независимые переменные принимают непрерывный ряд значений (цена товара, уровень инфляции и т.д.). Однако нередко приходится принимать во внимание и какие-либо *качественные* признаки. Например, можно изучать влияние пола работника на уровень его зарплаты (переменная “пол” принимает всего два зна-

чения), или спросом на прохладительные напитки в зависимости от времени года (сколько различных значений принимает эта переменная?). В подобных ситуациях “качественному” значению условно присваивается числовая метка, как правило, это 0 или 1.

Если качественный признак принимает два значения, то это присвоение можно сделать произвольно (например, “наличие детей”=1, “отсутствие детей”=0). Если же качественный признак принимает K значений ($K > 2$), то вводят $K-1$ переменную, каждая из которых принимает значения 0 и 1.

Пример. Качественный признак – время года, значения: зима, весна, лето, осень. Фиктивные переменные: $X_1=1$, если зима, иначе $X_1=0$, $X_2=1$, если весна, иначе $X_2=0$, $X_3=1$, если лето, иначе $X_3=0$. Переменную X_4 аналогичным образом вводить бессмысленно (и нелзя!), так как всегда $X_4=1-X_1-X_2-X_3$.

4.12 Тест Чоу

Иногда выборка состоит из нескольких частей и требуется выяснить, следует ли их объединить в одну, или рассматривать каждую подвыборку отдельно. Отчасти этот случай близок к рассмотренному в предыдущем пункте. Например, следует ли рассматривать статистические данные для каждого времени года отдельно, или их можно объединить в общую модель?

Обозначим регрессию по первой подвыборке через A , по второй – через B , а объединённую – через U . Ясно, что

$$ESS_A + ESS_B \leq ESS_U.$$

Если левая часть неравенства близка к правой, то это говорит о том, что регрессия по объединённой выборке почти так же хороша, как “составная” их двух регрессий по двум подвыборкам. В этом случае объединение выборок допустимо. Проверка этой близости составляет суть **критерия Чоу**.

- *Составляется статистика*

$$F = \frac{(ESS_U - ESS_A - ESS_B)/(m+1)}{(ESS_A + ESS_B)/(n-2m-2)}$$

- *Если вычисленное значение F-статистики оказывается меньше $F_{1-\alpha}(m+1, n-2m-2)$, то объединение выборок возможно при данном α .*

Другой способ проверки состоит в использовании фиктивных переменных. Добавим в модель фиктивную переменную, равную 1 для первой подвыборки и 0 для второй. Оценим соответствующее уравнение регрессии. Если фиктивная переменная окажется *незначимой*, то объединение выборок возможно.

4.13 Контрольные вопросы

1. Запишите модель множественной регрессии в матричной форме.
2. Как рассчитываются МНК-оценки для коэффициентов регрессии?
3. Что такое доверительный интервал?
4. Что такое уровень значимости?
5. Какая оценка называется несмещённой?
6. Сформулируйте условия Гаусса – Маркова.
7. Что показывает коэффициент детерминации?
8. Чем скорректированный коэффициент детерминации отличается от обычного?
9. Какие значения могут принимать коэффициент детерминации и скорректированный коэффициент детерминации?
10. Как проверяется значимость коэффициентов регрессии?
11. Как вычисляются стандартные ошибки для коэффициентов регрессии?
12. Как проверяется значимость уравнения регрессии?
13. К каким последствиям может привести невключение объясняющей переменной в уравнении регрессии?
14. Как отразится на статистической значимости коэффициентов регрессии включение в уравнение регрессии лишних объясняющих переменных?
15. Дайте определение мультиколлинеарности.
16. Какие меры можно предпринять для устранения мультиколлинеарности.
17. Для каких целей используется матрица частных коэффициентов корреляции?
18. Коррелированность двух или нескольких объясняющих переменных в уравнении регрессии называется: а) мультиколлинеарностью; б) автоколлинеарностью; в) автокорреляцией.

19. Укажите верное утверждение: а) невключение нужной переменной в уравнение регрессии может привести к смещенности оценок его коэффициентов; б) если в уравнении регрессии включена ненужная переменная, то оценки становятся смещенными; в) включение в уравнение регрессии лишних объясняющих переменных не отражается на статистической значимости коэффициентов регрессии по t – статистике.
20. Включение в уравнение регрессии ненужной переменной приводит к: а) смещенности и эффективности оценок; б) смещенности и неэффективности оценок; в) несмещенности и неэффективности оценок.
21. Опишите процедуру включения фиктивных переменных.
22. В чём состоит тест Чоу?
23. Можно ли использовать фиктивные переменные вместо теста Чоу?

5. Нарушения допущений классической линейной модели

Кроме проверки значимости коэффициентов и качества уравнения регрессии в целом, необходима проверка выполнения условий Гаусса-Маркова, обеспечивающих несмещенность и эффективность оценок параметров регрессии.

5.1 Мультиколлинеарность

5.1.1 Определение

Одним из условий классической регрессионной модели является предположение (С) о линейной независимости столбцов матрицы X , т.е. о линейной независимости значений объясняющих переменных (включая константу). Нарушение этого условия – коррелированность двух или нескольких объясняющих переменных в уравнении регрессии. В этой ситуации нельзя построить оценку параметров регрессии, что формально видно из обращения в нуль определителя матрицы. Мультиколлинеарность может привести к незначимости коэффициентов регрессии.

Например, если в модели множественной линейной регрессии

$$y = \beta_0 + \beta_1 X_1 + \dots + \beta_m X_m + \varepsilon \quad (5.1)$$

есть линейная связь между двумя переменными X_1 и X_2 , пусть

$$X_1 = C \cdot X_2,$$

то уравнение (6.1) может быть записано в виде

$$y = \beta_0 + \beta_1' X_1 + \dots + \beta_m X_m + \varepsilon \quad (5.2)$$

где $\beta_1' = \beta_1 + C \cdot \beta_2$.

Можно ожидать получения значимых по t -статистике оценок для коэффициентов уравнения регрессии (6.2), т.е. для $\beta_0, \beta_1', \dots, \beta_m$.

В частности, это может быть для коэффициента β_1' . Но для коэффициентов β_1, β_2 по отдельности, конечно, получение значимых оценок невозможно. Это крайний случай строгой линейной зависимости между двумя объясняющими переменными ("полная мультиколлинеарность").

Полная мультиколлинеарность является скорее теоретическим примером. На практике более реальна ситуация, когда между объясняющими переменными существует *сильная корреляционная* зависимость, а не *строгая функциональная*. Например, если в состав объясняющих переменных входят располагаемый доход и потребление, то обе эти переменные будут сильно коррелированными.

Если связь между переменными X_1 и X_2 нестрогая, то тем труднее разделить влияние объясняющих переменных, чем сильнее они коррелированы между собой. t -статистики коэффициентов при таких переменных будут получаться низкими.

5.1.2 Способы обнаружения

Вот некоторые наиболее характерные признаки мультиколлинеарности:

- *Небольшое изменение исходных данных приводит к существенному изменению оценок параметров модели.*
- *Оценки коэффициентов имеют малую значимость по t -критерию, в то время как модель в целом является значимой по F -критерию.*
- *Оценки коэффициентов имеют неправильные с точки зрения теории знаки.*

Пример. В рассматриваемом примере коэффициенты модели незначимы, а само уравнение значимо. По признаку 2 можно сделать вывод, что в исходных данных присутствует мультиколлинеарность.

На практике о наличии мультиколлинеарности судят по *матрице парных корреляций* факторов модели.

$$\begin{pmatrix} 1 & r_{01} & r_{02} & \dots & r_{0m} \\ r_{10} & 1 & r_{12} & \dots & r_{1m} \\ r_{20} & r_{21} & 1 & \dots & r_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ r_{m0} & r_{m1} & r_{m2} & \dots & 1 \end{pmatrix} \quad (5.3)$$

В наибольшей степени за мультиколлинеарность ответствен тот признак, который имеет более высокие по модулю значения коэффициентов парной корреляции с остальными.

Можно вывести корреляционную матрицу и определить пары переменных, сильно коррелированных между собой. В то же время возможно взаимовлияние объясняющих переменных модели между собой. Поэтому более правильно использовать матрицу *частных коэффициентов корреляции*.

5.1.3 Методы устранения

Если задачей исследования является прогноз будущих значений зависимой переменной, то мультиколлинеарность не является серьезной проблемой, так как она не ухудшает значимость уравнения.

Мультиколлинеарность затрудняет определение степени влияния каждой из объясняющих переменных на зависимую.

- Простейшим методом устранения мультиколлинеарности является исключение отдельных переменных из модели. Но при этом могут возникнуть новые трудности. Например, не всегда ясно, какие переменные являются “лишними”. Удаление переменных может изменить содержательный смысл модели.
- Иногда можно увеличить объём выборки или качество исходных данных.

- Мультиколлинеарность может возникать в силу различных причин. Например, несколько объясняющих переменных могут иметь общий временной тренд. В этом случае типичное преобразование – переход к первым разностям, особенно, для временных рядов, например, от X_i к $\Delta X_i = X_{i+1} - X_i$, где i – момент времени.

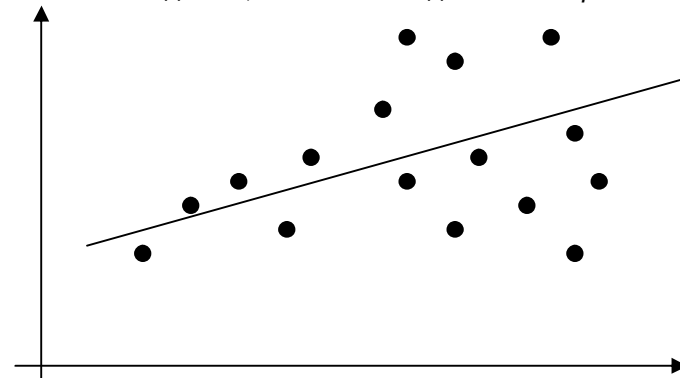
5.2 Гетероскедастичность

5.2.1 Определение

Условие **(D)** теоремы Гаусса-Маркова требует постоянства дисперсии случайного члена ε_i во всех наблюдениях:

$$D\varepsilon_i = \sigma^2 \text{ не зависит от } i$$

Это случай *гомоскедастичности*, при которой разброс точек наблюдений по вертикали явно не меняется с ростом номера наблюдений. Если же дисперсия случайного члена меняется от наблюдения к наблюдению, то мы имеем дело с *гетероскедастичностью*



При *гетероскедастичности* оценки коэффициентов регрессии будут несмещенными, но неэффективными. Вследствие этого окажутся завышенными t -статистики оценок коэффициентов, и будут сделаны ложные выводы о значимости коэффициентов регрессии.

5.2.2 Способы обнаружения

Существуют различные способы выявления гетероскедастичности.

5.2.2.1 Графический метод

При помощи метода наименьших квадратов оценивается регрессия $y=f(x_1, x_2, \dots, x_m)$. Вычисляются остатки e_i . Строится график зависимости остатков e_i от номера наблюдения. Далее анализируется внешний вид этого графика.

5.2.2.2 Тест Уайта

1. К исходной модели применяется обычный МНК, и находятся остатки e_i .
2. Осуществляется регрессия квадратов этих остатков на все регрессоры, их квадраты, попарные произведения и константу. Тогда величина nR^2 в случае отсутствия гетероскедастичности распределена по закону $\chi^2(N-1)$, где N – количество регрессоров.

5.2.2.3 Тест Голдфелда-Квандта

Предполагается, что стандартное отклонение σ_i пропорционально значению некоторой независимой переменной X_i .

1. Наблюдения упорядочиваются по возрастанию значений этой переменной X .
2. Берутся первые n' и последние n' наблюдений, и для них по отдельности строятся регрессии, и оцениваются суммы квадратов остатков ESS_1 и ESS_2 соответственно. Число n' можно взять примерно равным $n/3$.
3. Составляется статистика

$$\frac{ESS_2}{ESS_1} \sim F(n' - m - 1; n' - m - 1),$$

где m – число объясняющих переменных в исходном уравнении регрессии.

Большое значение этой статистики (больше соответствующего квантиля $F_{1-\alpha}(n'-m-1, n'-m-1)$) означает наличие гетероскедастичности.

5.2.2.4 Тест Глейзера

Оценивается регрессионная зависимость

$$|e_i| = a + bx_i^k + u_i$$

Параметр k изменяется с некоторым шагом, и для каждого его значения строится регрессия. Статистическая значимость коэффициента b при каком-либо значении означает наличие гетероскедастичности.

5.2.3 Методы устранения

Рассмотрим случай, когда стандартное отклонение σ_i пропорционально значению некоторой независимой переменной X_k . Разделив уравнение регрессии

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m + \varepsilon$$

на X_k , запишем его в виде

$$\frac{y}{x_k} = \frac{\beta_0}{x_k} + \frac{\beta_1 x_1}{x_k} + \dots + 1 + \dots + \frac{\beta_m x_m}{x_k} + \frac{\varepsilon}{x_k}$$

Если теперь перейти к новым переменным

$$\tilde{y} = \frac{y}{x_k}; \tilde{\varepsilon} = \frac{\varepsilon}{x_k}; \tilde{x}_i = \dots; \tilde{\beta}_i = \dots$$

то оно примет вид

$$\tilde{y} = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{x}_1 + \dots + \tilde{\beta}_m \tilde{x}_m + \tilde{\varepsilon}$$

причем $D(\tilde{\varepsilon}) = const$.

Аналогичные преобразования можно применять и при других видах зависимости σ_i от независимых переменных.

5.2.4 Контрольные вопросы

1. Какое условие Гаусса-Маркова должно выполняться, чтобы разброс точек наблюдений по вертикали явно не менялся с ростом номера наблюдений?
2. Когда наблюдается гетероскедастичность в модели регрессии?
3. Каким образом наличие гетероскедастичности повлияет на оценки коэффициентов регрессии?
4. Укажите способы выявления гетероскедастичности.
5. Явление гетероскедастичности возникает при нарушении условия Гаусса-Маркова: а) о постоянстве дисперсии случайного члена; б) об отсутствии систематической связи между значениями случайного члена в любых двух наблюдениях; в) о том, что математическое ожидание случайного члена в любом наблюдении должно быть равно нулю.
6. При гетероскедастичности оценки коэффициентов становятся: а) смещенными; б) неэффективными; в) эффективными, но несостоятельными.
7. В условиях гетероскедастичности t -статистики оценок коэффициентов окажутся: а) неизменными; б) завышенными; в) заниженными.

5.3 Автокорреляция

5.3.1 Определение

Условие (E) теоремы Гаусса-Маркова требует отсутствия корреляции между значениями случайного фактора ε_i во всех наблюдениях:

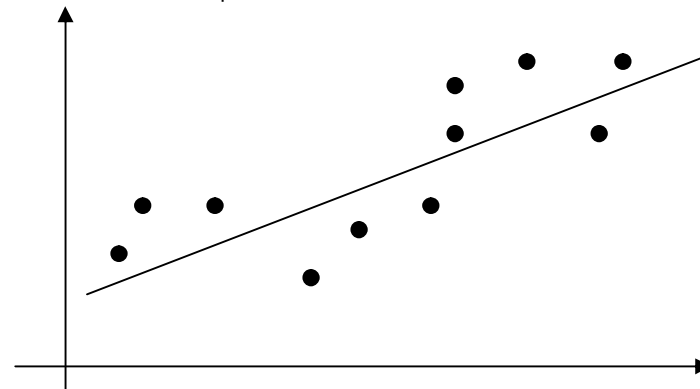
$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0 \text{ при } i \neq j$$

Фактически это условие соответствует независимости случайных членов в разных наблюдениях.

При нарушении этого условия, т.е. при наличии связи между случайными членами возникает явление *автокорреляции*.

Обычно *автокорреляция рассматривается при изучении временных рядов* (т.е. когда номер наблюдения соответствует моменту времени). В случае *положительной автокорреляции* реализации

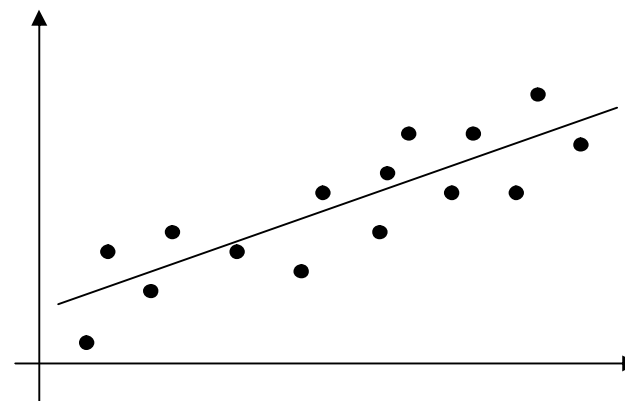
случайного члена ε_i для ряда последовательных наблюдений смещают значения зависимой переменной в одном направлении, затем для нескольких последовательных наблюдений – в противоположном направлении, потом снова в первоначальном направлении и т.д. (рис 9.1). В экономике положительная автокорреляция может быть связана с циклами



деловой активности, с сезонными колебаниями, влияние которых и отражается в случайном члене уравнения регрессии.

При *отрицательной автокорреляции* каждая реализация случайного члена ε_i , как правило, сменяется реализацией случайного члена

ε_{i+1} противоположного знака:



Для экономики обычно характерно возникновение положительной автокорреляции.

5.3.2 Способы обнаружения

5.3.2.1 Графический метод

При помощи метода наименьших квадратов оценивается регрессия $y=f(x_1, x_2, \dots, x_m)$. Вычисляются остатки e_i . Строится график зависимости остатков e_i от номера наблюдения. Далее анализируется внешний вид этого графика.

5.3.2.2 Метод рядов

В этом методе последовательно определяются знаки отклонений e_i . Ряд определяется как непрерывная последовательность одинаковых знаков.

При достаточно большом числе наблюдений для определения автокорреляции применяется следующая процедура.

- Рассчитываются статистики

$$M = \frac{2n_+n_-}{n_+ + n_-} + 1, \quad D = \frac{2n_+n_-(2n_+n_- - n_+ - n_-)}{(n_+ + n_-)^2(n_+ + n_- - 1)}$$

где n_+ – количество положительных отклонений, n_- – количество отрицательных.

- Если

$$M - u_{1-\alpha/2} \cdot D < k < M + u_{1-\alpha/2} \cdot D,$$

где k – количество рядов, то принимается гипотеза об отсутствии автокорреляции.

Для малых объёмов выборки есть таблицы критических значений k (см. приложение).

Пример. Рассмотрим ситуацию, когда знаки отклонений имеют вид:

(- - - - -)(+ + + + +)(- - - - -)(+ + + + +)

Здесь $k=4$, $n_+=10$, $n_-=10$. По таблице из приложения $k_1=6$, $k_2=16$. Число k не попадает в эти пределы, так что на уровне $\alpha=0,05$ гипотеза о наличии автокорреляции принимается.

5.3.2.3 Критерий Дарбина-Уотсона

Данный метод применяется для обнаружения автокорреляции вида

$$e_i = \rho \cdot e_{i-1} + u_i \quad (5.4)$$

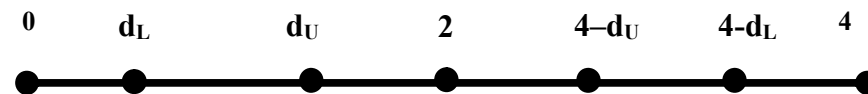
где u_i – случайные попарно независимые величины, ρ – некоторый параметр, называемый коэффициентом авторегрессии ($|\rho| < 1$). Такая ситуация называется *авторегрессионным процессом первого порядка*. Случайный член в данном наблюдении прямо связан лишь со случайным членом в предыдущем наблюдении. Если $\rho > 0$, то автокорреляция положительная, если $\rho < 0$, то отрицательная.

В критерии Дарбина-Уотсона проверяется гипотеза $\rho=0$ при альтернативной $\rho \neq 0$. Для обнаружения автокорреляции обычно используется статистика Дарбина-Уотсона:

$$DW = \frac{\sum_1^n (e_i - e_{i-1})^2}{\sum_1^n e_i^2}, \quad e_i = y_i - \hat{y}_i \quad (5.5)$$

Величина **DW** обычно рассчитывается в эконометрических пакетах. Для больших выборок $DW \approx 2 \cdot (1 - \rho)$. При значительной положительной автокорреляции $\rho \approx 1$ и $DW \approx 0$. При отрицательной автокорреляции $\rho \approx -1$ и $DW \approx 4$. А при отсутствии автокорреляции. При отрицательной автокорреляции $\rho \approx 0$ и $DW \approx 2$.

Оказывается, критические значения статистики Дарбина-Уотсона зависят не только от числа объясняющих переменных, но и от значений, которые они принимают в выборке. Поэтому в отличие от t -статистики и от F -статистики невозможно составить таблицы критических значений **DW** статистики Дарбина-Уотсона. Но можно указать верхнюю d_U и нижнюю d_L границы для d . Они определяются в зависимости от n и числа оцениваемых параметров h .



При $0 \leq DW \leq d_L$ принимается гипотеза $\rho > 0$ (положительная автокорреляция), при $4 - d_U \leq DW \leq 4$ принимается гипотеза $\rho < 0$ (отрицательная автокорреляция). Если **DW** попадает в интервал

$(d_U, 4-d_U)$, то принимается гипотеза об отсутствии авторегрессии первого порядка.

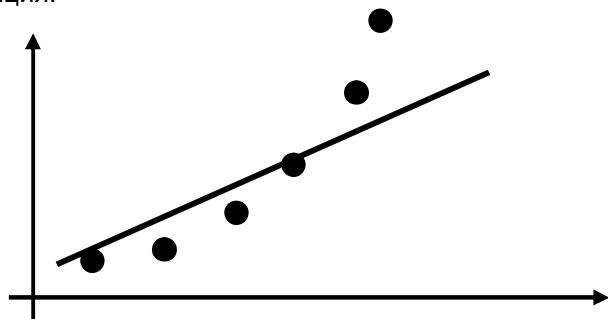
Обратите внимание, что на рисунке присутствует “зона неопределённости”, в которой неясно наличие или отсутствие автокорреляции.

5.3.3 Методы устранения

5.3.3.1 Исправление спецификации модели

Основной причиной наличия случайного фактора является несовершенство знаний о причинах и взаимосвязях, определяющих то или иное значение зависимой переменной. Поэтому свойства случайного члена зависят от выбора формы зависимости и состава объясняющих переменных. Другими словами, автокорреляция чаще всего вызывается ошибками в спецификации модели.

Например, при выборе линейной формы зависимости в ситуации, когда имеет место экспоненциальная, возникает положительная автокорреляция:



Если автокорреляция вызвана отсутствием в модели существенной переменной, можно попытаться добавить её в модель.

5.3.3.2 Авторегрессионное преобразование

Рассмотрим случай парной линейной регрессии. Если имеет место зависимость (5.4) и коэффициент ρ известен, для исходного уравнения регрессии

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (5.6)$$

можно выполнить *авторегрессионное* преобразование. Запишем это уравнение для предыдущего номера:

$$y_{i-1} = \beta_0 + \beta_1 x_{i-1} + \varepsilon_{i-1}$$

умножим его обе части на ρ и вычтем из уравнения (5.6).

$$y_i - \rho y_{i-1} = \beta_0(1 - \rho) + \beta_1(x_i - \rho x_{i-1}) + u_i$$

Если ввести новые переменные

$$y_i^* = y_i - \rho y_{i-1}, \quad x_i^* = x_i - \rho x_{i-1}, \quad \beta_0^* = \beta_0(1 - \rho),$$

то последнее уравнение можно записать в виде

$$y_i^* = \beta_0^* + \beta_1 x_i^* + u_i \quad (5.7)$$

В модели (5.7) автокорреляция будет отсутствовать. Тогда метод наименьших квадратов позволяет для уравнения (5.7) определить оценки коэффициентов β_0^*, β_1^* которые будут несмещенными и эффективными оценками, откуда находятся оценки коэффициентов и для исходного уравнения (5.6).

Однако проблема заключается в том, что величина ρ неизвестна. В качестве оценки для ρ часто используют $1 - DW/2$.

Другой метод оценки используется в **процедуре Кохрейна-Оркатта**. Она представляет собой итеративный процесс определения параметров β_0, β_1, ρ .

1. Методом наименьших квадратов находится оценка ρ_1 параметра ρ регрессии $e_i = \rho \cdot e_{i-1} + u_i$.
 2. Проводится авторегрессионное преобразование с этим значением параметра ρ и находятся МНК-оценки параметров β_i .
 3. Строится новый вектор остатков e_i и процедура повторяется.
- Процесс заканчивается, когда очередное приближение ρ мало отличается от предыдущего.

Метод Хилдрета-Лу использует перебор по значениям ρ из интервала $(-1, 1)$ с некоторыми шагом при оценивании β_0, β_1 по преобразованному уравнению. В процессе перебора определяется то зна-

чение ρ , при котором сумма квадратов отклонений для преобразованного уравнения минимальна. Вблизи найденного минимума процедура может повторяться с более мелкими шагами до достижения заданной точности.

5.3.4 Контрольные вопросы

1. При каких условиях возникает явление автокорреляции?
2. В чем заключается разница между положительной автокорреляцией и отрицательной автокорреляцией?
3. Какой критерий используется для обнаружения автокорреляции?
4. От каких факторов зависят критические значения статистики Дарбина-Уотсона?
5. Для каких моделей критерий Дарбина-Уотсона неприменим?
6. При отсутствии автокорреляции статистика Дарбина-Уотсона примет значение, близкое к: а) 2; б) 0; в) 4.
7. При каких условиях возникает автокорреляция первого порядка?
8. Сформулируйте условия существования положительной и отрицательной автокорреляции первого порядка.
9. Опишите метод Кохрейна-Оркатта.
10. В чем заключается сущность метода Хилдрета-Лу?
11. Когда случайные члены в уравнении регрессии связаны соотношением $e_i = \rho \cdot e_{i-1} + u_i$, положительная автокорреляция наблюдается при условии: а) $\rho > 0$, б) $\rho < 0$, в) $\rho = 0$.
12. При использовании метода Хилдрета-Лу используется перебор по значениям ρ из интервала: а) $(-1; 0)$, б) $(-1; 1)$, в) $(0; 4)$.

5.4 Стохастические объясняющие переменные

До этого раздела предполагалось, что объясняющие переменные X_1, X_2, \dots, X_m и матрица X результатов их наблюдений является неслучайными, детерминированными. Данное предположение позволяло упростить анализ модели. В этом разделе мы откажемся от этого допущения и выясним, к чему это может привести. Пусть, как и раньше

$$Y = X \cdot \beta + \varepsilon$$

Будем теперь считать элементы матрицы X случайными величинами. Предположим, что выполнены условия

(C1) Матрица X с вероятностью 1 имеет ранг $m+1$.

(D1) $M(\varepsilon | X) = 0$;

(D2) $M(\varepsilon \varepsilon^T | X) = \sigma^2 E$;

Здесь $M(\varepsilon | X)$ – условное математическое ожидание случайного вектора ε при фиксированной матрице X .

Можно показать, что в этих предположениях

- МНК-оценки параметров регрессии являются несмещёнными.
- Справедлив соответствующий вариант теоремы Гаусса-Маркова, т.е. среди всех линейных условно несмещённых оценок МНК-оценка обладает наименьшей условной дисперсией. Теорема.
- Если в каждом наблюдении значения объясняющих переменных (регрессоров) выбираются из одной и той же генеральной совокупности, а ошибки независимы, одинаково распределены и не зависят от регрессоров, то МНК-оценки состоятельны.
- Если регрессоры и ошибки коррелированы, то МНК-оценка будет в общем случае смещённой и несостоятельной. В этом случае можно попробовать применить *метод инструментальных переменных*.

5.5 Метод инструментальных переменных

Он состоит в замене объясняющей переменной, коррелированной с ε , на другую переменную, которая не коррелирована с ε . Рассмотрим пример

$$y = X\beta + \varepsilon$$

Допустим, что есть переменная Z , тесно связанная (коррелированная) с X , но не коррелированная с ε . Обозначим

$$\hat{\beta}_{IP} = (Z^T X)^{-1} Z^T Y$$

Можно показать, что с при “хорошей” корреляции между X и Z такая оценка является состоятельной, т.е. с ростом объема выборки эта оценка параметров β стремится к истинным значениям параметров.

5.6 Нелинейные формы зависимости

Использование лишь линейных зависимостей для описания экономических взаимосвязей часто оказывается недостаточным. Необходимо привлекать нелинейные соотношения. Например, производственная функция Кобба-Дугласа

$$Y = A \cdot K^\alpha \cdot L^\beta \quad (5.8)$$

где Y – выпуск, K – затраты капитала, L – затраты труда, A, α , β , – параметры, является степенной функцией.

В общем случае применяется производственная функция с постоянной эластичностью замещения CES:

$$Y = a \cdot [\delta K^{-\rho} + (1 - \delta)L^{-\rho}]^{-\frac{\eta}{\rho}} \quad (5.9)$$

с параметрами α , ρ , δ , η . Она более реалистична, чем функция Кобба-Дугласа, так как не позволяет сколь угодно заменять капитал трудом и наоборот.

Во многих случаях нелинейные зависимости путем замены переменных можно преобразовать к линейному виду. Так модель вида

$$y = a + \frac{b}{x}$$

в результате замены $z = \frac{b}{x}$ приводится к линейному виду

$$y = a + bz.$$

Другой пример с несколькими объясняющими переменными:

$$y = a + b_1 x_1 + b_2 x_1^2 + b_3 e^{x_1 + x_2}$$

Замена переменных

$$z_1 = x_1; z_2 = x_1^2; z_3 = e^{x_1 + x_2}$$

также приводит эту нелинейную зависимость к линейному виду. При этом исходные статистические данные легко пересчитываются для новых переменных.

Рассмотрим степенную зависимость:

$$Y = a \cdot x_1^{b_1} x_2^{b_2} \cdot \dots \cdot x_m^{b_m} \cdot \varepsilon,$$

где a, b_1 , ..., b_m – параметры, а ε – случайный множитель. Прологифмируем это соотношение

$$\ln y = \ln a + b_1 \ln x_1 + \dots + b_m \ln x_m + \ln \varepsilon$$

Если теперь обозначить

$$\tilde{y} = \ln y; \beta_0 = \ln a; \tilde{x}_i = \ln x_i; \tilde{\varepsilon} = \ln \varepsilon$$

то оно примет вид линейной регрессионной модели:

5.7 Нелинейные модели, неприводимые к линейному виду

Производственная функция с постоянной эластичностью замены CES зависит от параметров, которые надо оценить по статистическим данным. На первый взгляд кажется, что ее, как и функцию Кобба-Дугласа, можно преобразовать к линейному виду. Возведем обе части уравнение (12.2) в степень $-\rho/\eta$:

$$Y^{-\frac{\eta}{\rho}} = a^{-\frac{\eta}{\rho}} \delta K^{-\rho} + a^{-\frac{\eta}{\rho}} (1 - \delta) L^{-\rho}$$

Однако, здесь просто переобозначить, например, $Z = Y^{-\frac{\eta}{\rho}}$ нельзя, так как неизвестны параметры ρ , η и по выборочным данным Y_1, \dots, Y_n нельзя вычислить соответствующие значения новой переменной Z_1, \dots, Z_n .

В общем случае уравнение нелинейной регрессии с аддитивным случайным членом ε имеет вид

$$y = f(x_1, \dots, x_m, \beta_1, \dots, \beta_m) + \varepsilon$$

где y – зависимая переменная, X_1, \dots, X_m – объясняющие переменные, β_1, \dots, β_m – параметры модели.

Для нахождения оценок этих параметров можно использовать, как и в линейном случае, принцип наименьших квадратов:

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min$$

Сумма квадратов отклонений является функцией m переменных β_1, \dots, β_m . Задача минимизации решается при помощи итеративных методов оптимизации, например, методом Ньютона.

5.8 Контрольные вопросы

1. Какие условия накладываются в случае стохастических объясняющих переменных?
2. Как в этом случае изменяется формулировка теоремы Гаусса-Маркова?
3. Опишите метод инструментальных переменных.
4. Приведите примеры нелинейных зависимостей, используемых для описания экономических процессов.
5. Какие существуют способы приведения нелинейных зависимостей к линейному виду?
6. От чего зависит производственная функция с постоянной эластичностью замены CES и почему ее нельзя привести к линейному виду?

Литература

1. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учеб. пособие. – М.: Дело, 2004.
2. Бородич С.А. Эконометрика: Учеб. пособие. – Мн.: Новое знание, 2001.
3. Доугерти К. Введение в эконометрику – М.: ИНФРА-М, 2001.
4. Берндт Э.Р. Практика эконометрики: классика и современность. – М.: ЮНИТИ–ДАНА, 2005.

Приложение. Таблицы квантилей

Квантили u_p стандартного нормального закона распределения

$N(1,0)$.

p	0,9	0,95	0,975	0,99	0,995	0,999	0,9995
u_p	1,282	1,645	1,960	2,325	2,576	3,090	3,291

Квантили $\chi_p^2(k)$ закона распределения $\chi^2(k)$,

$k \backslash p$	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99
1	0,0002	0,001	0,004	0,0158	2,71	3,84	5,02	6,63
2	0,02	0,05	0,103	0,211	4,61	5,99	7,38	9,21
3	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,3
4	0,297	0,484	0,711	1,06	7,78	9,49	11,1	13,3
5	0,554	0,831	1,15	1,61	9,24	11,1	12,8	15,1
10	2,56	3,25	3,94	4,87	16	18,3	20,5	23,2
20	7,63	9,59	10,9	12,4	28,4	31,4	34,2	37,6
30	14,3	16,8	18,5	20,6	40,3	43,8	47,0	50,9
40	22,2	24,4	26,5	29,1	51,8	55,8	59,3	63,7
50	29,7	32,4	34,8	37,7	63,2	67,5	71,4	76,2
75	49,5	53	56,1	59,8	91,1	96,2	100,8	106,4
100	70,1	74,2	77,9	82,4	118,5	124,3	129,6	135,6

Квантили $t_p(k)$ закона распределения Стьюдента $T(k)$,

$k \backslash p$	0,9	0,95	0,975	0,99	0,995
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,92	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
10	1,372	1,812	2,228	2,764	3,169
20	1,325	1,725	2,086	2,528	2,845
30	1,312	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
120	1,289	1,658	2,98	2,358	2,617

Квантили $F_{0,9}(n_1, n_2)$ закона распределения Фишера

$n_1 \backslash n_2$	1	2	3	4	5	10	15	20	30	120
1	40	8,53	5,54	4,54	4,06	3,29	3,07	2,97	2,88	2,75
2	49,5	9	5,46	4,32	3,78	2,92	2,7	2,59	2,49	2,35
3	53,6	9,16	5,39	4,19	3,62	2,73	2,49	2,38	2,28	2,13
4	55,8	9,24	5,34	4,11	3,52	2,61	2,36	2,25	2,14	1,99
30	62,2	9,46	5,17	3,82	3,17	2,16	1,87	1,74	1,61	1,41

Квантили $F_{0,95}(n_1, n_2)$ закона распределения Фишера

$n_1 \backslash n_2$	1	2	3	4	5	10	15	20	30	120
1	161	18,5	10,13	7,71	6,61	4,96	4,54	4,35	4,24	3,92
2	199	19	9,55	6,94	5,79	4,1	3,68	3,49	3,39	3,07
3	216	19,16	9,28	6,59	5,41	3,71	3,29	3,1	2,99	2,68
4	225	19,25	9,12	6,39	5,19	3,48	3,05	2,87	2,76	2,45
30	250	19,46	8,62	5,75	4,5	2,7	2,25	2,04	1,84	1,55

Критические точки d_H, d_B распределения Дарбина-Уотсона при $\alpha=0,05$

$k \backslash n$	6	7	8	9	10	20	30	40	
1	d_H	0,61	0,7	0,763	0,824	0,879	1,2	1,352	1,44
	d_B	1,4	1,356	1,332	1,32	1,32	1,41	1,49	1,544
2	d_H		0,467	0,359	0,629	0,697	1,1	1,284	1,39
	d_B		1,896	1,777	1,7	1,641	1,537	1,567	1,6
3	d_H			0,368	0,435	0,525	0,998	1,214	1,338
	d_B			2,287	2,128	2,016	1,676	1,65	1,659
4	d_H				0,296	0,376	0,894	1,143	1,285
	d_B				2,39	2,414	1,828	1,74	1,721

Критические точки $K_H - K_B$ для проверки автокорреляции методом рядов при $\alpha=0,05$

$n_- \backslash n_+$	3	4	5	6	7	8	9	10	15	20
3	-	-	-	2-	2-	2-	2-	2-	3-	3-
4		-	2-9	2-9	2-	3-	3-	3-	3-	4-
5			2-10	3-10	3-11	3-11	3-	3-	4-	5-
6				3-11	3-12	3-12	4-13	4-13	5	6-
7					3-13	4-13	4-14	5-14	6-15	6-
8						4-14	5-14	5-15	6-16	7-17
9							5-15	5-16	7-18	8-18
10								6-16	7-18	9-20
15									10-22	12-25
20										14-28

Оглавление

4. Классическая модель множественной линейной регрессии	3
4.1. Основные гипотезы	3
4.2. Теорема Гаусса-Маркова	5
4.3. Коэффициент детерминации	6
4.4. Статистические свойства оценок коэффициентов	8
4.5. Доверительные интервалы	8
4.6. Интервал для прогнозного значения	10
4.7. Проверка гипотез о параметрах регрессии	10
4.8. Проверка гипотезы о значимости уравнения регрессии	11
4.9. Частная корреляция	12
4.10. Спецификация модели	13
4.11. Фиктивные переменные	15
4.12. Тест Чоу	15
4.13. Контрольные вопросы	16
5. Нарушения допущений классической линейной модели	17
5.1. Мультиколлинеарность	17
5.2. Гетероскедастичность	20
5.3. Автокорреляция	23
5.4. Стохастические объясняющие переменные	29
5.5. Метод инструментальных переменных	30
5.6. Нелинейные формы зависимости	31
5.7. Нелинейные модели, неприводимые к линейному виду	32
5.8. Контрольные вопросы	33
Литература	33
Приложение. Таблицы квантилей	34