

Обобщенный МНК

лекция 16

ОБОБЩЕННЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

$$Y = X\beta + u \quad (Y = X\beta + \sigma^2 v)$$

Условия Гаусса-Маркова

0. модель правильно специфицирован и линейна по параметрам

1. $E(u_j) = 0$

2. Ошибки независимы и одинаковы распределены

$$\text{cov}(u_i; u_j) \equiv 0 \quad \forall i \neq j; \quad \text{Var}(u_i) = \sigma^2 \quad \forall i$$

$$\text{Var}(\mathbf{u}) = \sigma^2 \cdot \mathbf{I}$$

3. Регрессоры детерминированы

4. Регрессоры некоррелированы с ошибками $\text{COVAR}(\mathbf{X}, \mathbf{u})$

5. Ошибки имеют совместное нормальное распределение:
 $\mathbf{u} \sim N(0; \mathbf{I}\sigma^2)$

БЕЛЫЙ ШУМ

- Условия 1.-3., наложенные на распределения ошибок описываются только свойствами моментов 1 и 2 порядков (**слабый белый шум**):

$$E(\mathbf{u}) = \mathbf{0}; \quad \text{VAR}(\mathbf{u}) = \sigma^2 \mathbf{I}$$

- Условие 5. (требование нормальности)- **белый гауссовский шум**

$$\mathbf{u} \sim N(\mathbf{0}; \sigma^2 \mathbf{I})$$

- Условия 1.-4. могут быть заменены на более слабые:

$$E(\mathbf{u} | \mathbf{X}) = \mathbf{0}; \quad \text{COVAR}(\mathbf{u} | \mathbf{X}) = \sigma^2 \mathbf{I}; \quad \text{cov}(\mathbf{X}; \mathbf{u}) = \mathbf{0}$$

Классические условия регрессионного анализа (условия Гаусса-Маркова)

- I. Регрессионная модель линейна по параметрам (коэффициентам), корректно специфицирована, и содержит аддитивный случайный член.
- II. Случайный член имеет нулевое среднее.
- III. Объясняющая переменная не коррелирована со случайным членом.
- IV. Наблюдаемые значения случайного члена не коррелированы друг с другом.
- V. Случайный член имеет постоянную дисперсию
- VI. Случайный член распределен нормально (необязательное, но часто используемое условие).

Классические условия регрессионного анализа

- I. Регрессионная модель линейна по параметрам (коэффициентам) и включает константу, корректно специфицирована, и содержит аддитивный случайный член.

$$Y_i = \alpha + \beta X_i + u_i$$

II. Случайный член имеет нулевое среднее.

$$E(u_i) = 0$$

III. Объясняющая переменная не коррелирована со случайным членом.

$$\text{Cov}(X_i, u_i) = 0$$

IV. Наблюдаемые значения случайного члена не коррелированы друг с другом.

$$\text{Cov}(u_i, u_j) = 0, \quad i \neq j$$

V. Случайный член имеет постоянную дисперсию

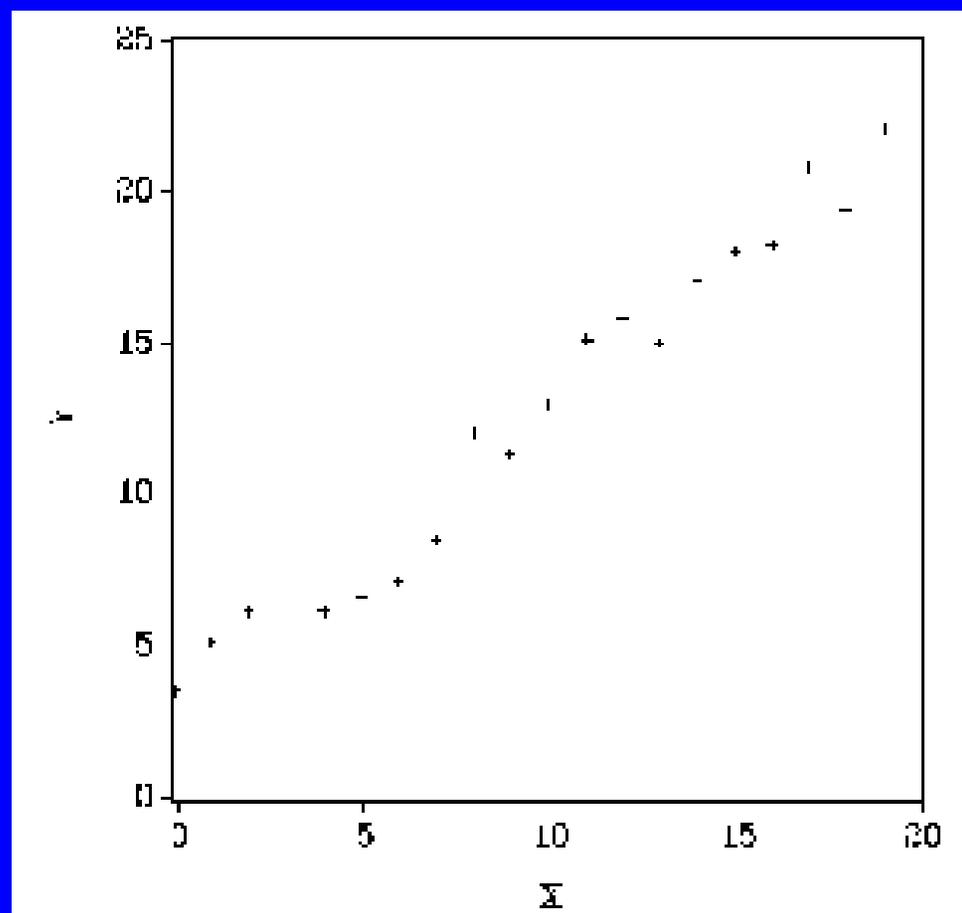
$$\mathit{Var}(u_i) = \mathit{Const.}$$

VI. Случайный член распределен нормально (необязательное, но часто используемое условие).

$$u_i \sim N(0, \sigma^2)$$

Классические условия регрессионного анализа

Типичная картина выполнения условий Гаусса-Маркова

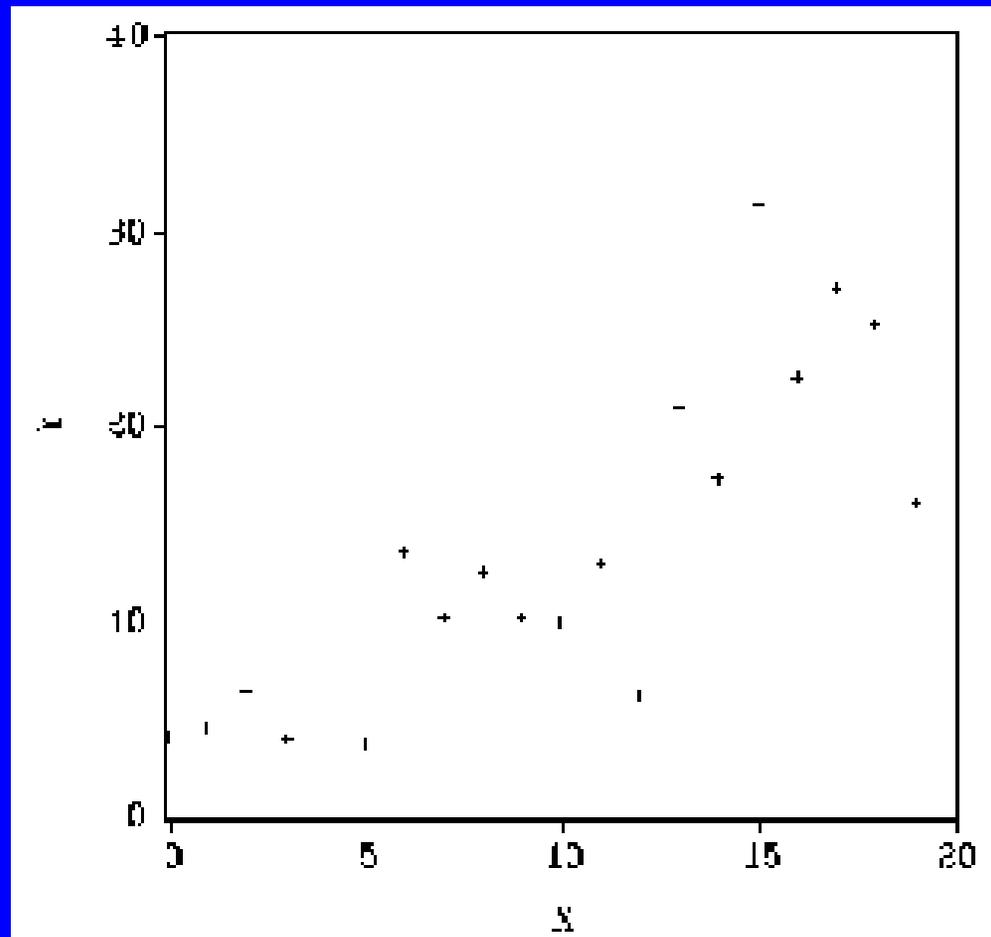


Классические условия регрессионного анализа

Типичная картина нарушения условий

$$\text{Var}(u_i) = \text{Const.}$$

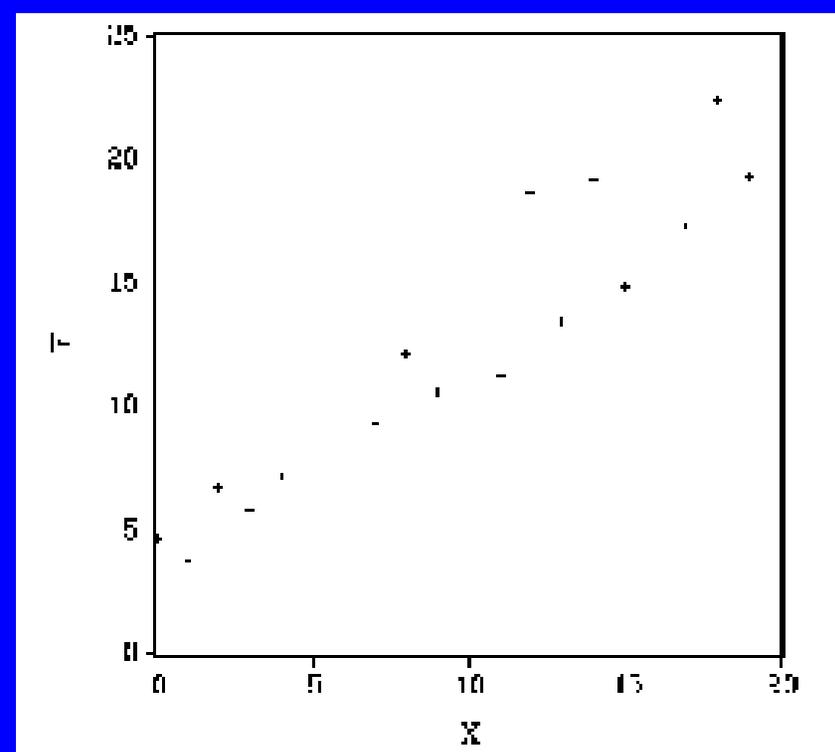
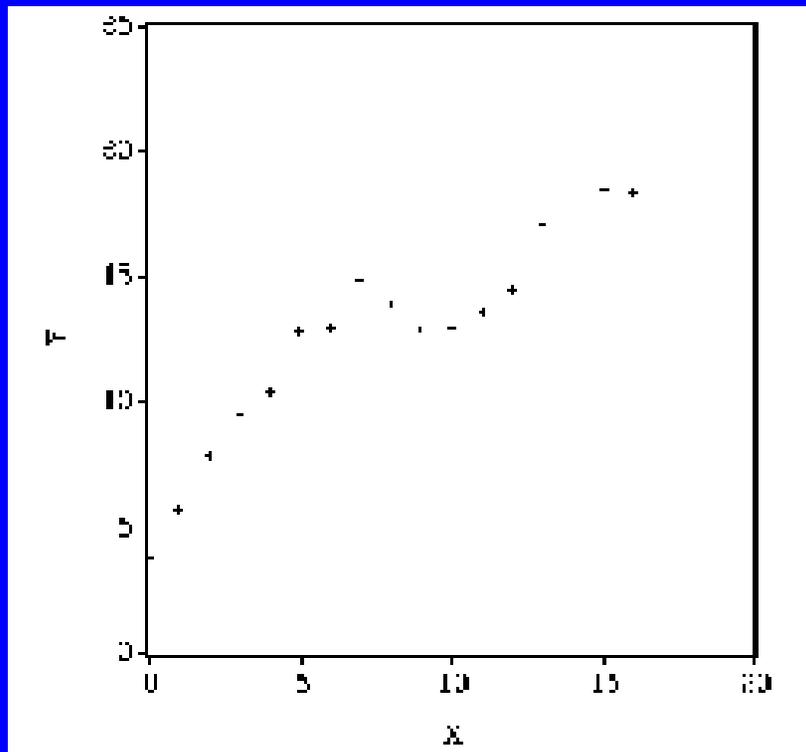
$$\text{Cov}(X_i, u_i) = 0$$



Классические условия регрессионного анализа

Типичные картины нарушения условия

$$\text{Cov}(u_i, u_j) = 0, \quad i \neq j$$



СВОЙСТВА ОЦЕНОК КОЭФФИЦИЕНТОВ

$$Y = X\beta + u$$

0. ОЦЕНКА МНК $\beta = (X'X)^{-1} X'y$

1. НЕСМЕЩЕННОСТЬ $E\hat{\beta} = \beta; Ee = 0$

2. МАТРИЦЫ КОВАРИАЦИЙ

$$\text{COVAR}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

$$E(ee') = \sigma^2 I; E(e'e) = (N - k)\sigma^2$$

3. ОЦЕНКА ДИСПЕРСИИ ОШИБОК

$$\widehat{\sigma^2} = s^2 = \frac{e'e}{N - k}$$

4. НЕЗАВИСИМОСТЬ $\text{COV}(\hat{\beta}; e) = 0$

5. ОЦЕНКА МАТРИЦ КОВАРИАЦИЙ

$$\widehat{\text{COV}}(e) = \frac{e'e}{N - k} I = s^2 I; \widehat{\text{COV}}(\hat{\beta}) = \frac{e'e(X'X)^{-1}}{N - k} = s^2 (X'X)^{-1}$$

6. НОРМАЛЬНОСТЬ

Если $u \sim N(0; \sigma^2 I)$, то

$$\hat{\beta} \in N(\beta; \sigma^2 (X'X)^{-1}) \sim N(\beta; s^2 (X'X)^{-1});$$

$$e \in N(0; \sigma^2 I) \sim N(0; s^2 I)$$

ТЕОРЕМА ГАУССА-МАРКОВА

В ЛИНЕЙНОЙ РЕГРЕССИОННОЙ МОДЕЛИ,
В КОТОРОЙ ОШИБКИ ЯВЛЯЮТСЯ БЕЛЫМ ШУМОМ,
А МНОЖЕСТВО РЕГРЕССОРОВ ЛНЗ:

1. ОЦЕНКИ ПАРАМЕТРОВ МОДЕЛИ
ЯВЛЯЮТСЯ ЛИНЕЙНЫМИ И НЕСМЕЩЕННЫМИ
2. В ЭТОМ КЛАССЕ ОНИ ЯВЛЯЮТСЯ ЭФФЕКТИВНЫМИ
(BLUE –best linear unbiased estimators)

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ

$$y = X\beta + u$$

1. X полного ранга, неслучайна
2. $E u = 0$
3. $\text{Var}(u) = \Omega$ положительно определена
4. $\text{cov}(X, u) = 0$
5. $u \in N(0; \Omega)$

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ и обычный МНК

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim (\mathbf{0}; \boldsymbol{\Omega}) \quad \left(\begin{array}{l} \exists i, j : \text{Var } u_i = \sigma_i^2 \neq \sigma_j^2 = \text{Var } u_j \\ \exists i, j : \text{cov}(u_i; u_j) \neq 0 \end{array} \right)$$

$$\text{Var}(\mathbf{y}) = \text{Var}(\mathbf{u}) = \boldsymbol{\Omega}$$

$$\hat{\boldsymbol{\beta}}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}; \quad \hat{\mathbf{y}} = \mathbf{P}\mathbf{y}; \quad \mathbf{e} = \mathbf{M}\mathbf{y}$$

$$\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'; \quad \mathbf{M} = \mathbf{I} - \mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'$$

Упражнение :

$$\mathbf{P}^2 = \mathbf{P}; \quad \mathbf{P}' = \mathbf{P}; \quad \mathbf{M}^2 = \mathbf{M}; \quad \mathbf{M}' = \mathbf{M}; \quad \mathbf{P}\mathbf{M} = \mathbf{M}\mathbf{P} = \mathbf{0}$$

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ и обычный МНК

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim (\mathbf{0}; \boldsymbol{\Omega}); \quad \text{Var}(\mathbf{y}) = \text{Var}(\mathbf{u}) = \boldsymbol{\Omega}$$

$$E \hat{\boldsymbol{\beta}}_{OLS} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' E(\mathbf{u}) = \boldsymbol{\beta}$$

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ и обычный МНК

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim (\mathbf{0}; \boldsymbol{\Omega}); \quad \text{Var}(\mathbf{y}) = \text{Var}(\mathbf{u}) = \boldsymbol{\Omega}$$

$$\text{Var}(\hat{\boldsymbol{\beta}}_{OLS}) = \text{Var}\left[(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}\right] = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \text{Var}(\mathbf{y}) \mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \Rightarrow$$

$$\Rightarrow \text{Var}(\hat{\boldsymbol{\beta}}_{OLS}) = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Omega}\mathbf{X} (\mathbf{X}'\mathbf{X})^{-1}$$

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ и обычный МНК

$$y = \mathbf{X}\beta + u, \quad u \sim (0; \Omega); \quad \text{Var}(y) = \text{Var}(u) = \Omega$$

$$\text{Var}(\hat{\beta}_{OLS}) = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\Omega\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$\text{if } \frac{\mathbf{X}'\mathbf{X}}{n} \xrightarrow[n \rightarrow \infty]{as} \mathbf{Q}; \quad \frac{\mathbf{X}'\Omega\mathbf{X}}{n} \xrightarrow[n \rightarrow \infty]{as} \tilde{\mathbf{Q}} \Rightarrow$$

$$\Rightarrow \text{Var}(\hat{\beta}_{OLS}) = \frac{1}{n} \left(\frac{\mathbf{X}'\mathbf{X}}{n} \right)^{-1} \frac{\mathbf{X}'\Omega\mathbf{X}}{n} \left(\frac{\mathbf{X}'\mathbf{X}}{n} \right)^{-1} \xrightarrow[n \rightarrow \infty]{as} \frac{1}{n} \cdot \mathbf{Q}^{-1} \cdot \tilde{\mathbf{Q}} \cdot \mathbf{Q}^{-1} \xrightarrow[n \rightarrow \infty]{as} \mathbf{0}$$

$\hat{\beta}_{OLS}$ состоятельна, но неэффективна

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ и обычный МНК

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim (\mathbf{0}; \boldsymbol{\Omega}); \quad \text{Var}(\mathbf{y}) = \text{Var}(\mathbf{u}) = \boldsymbol{\Omega}$$

$$\mathbf{E} \mathbf{e} = \mathbf{E}(\mathbf{M}\mathbf{y}) = \mathbf{M} \mathbf{E} \mathbf{y} = \mathbf{M} \mathbf{E} \mathbf{u} = \mathbf{0}$$

$$\text{Var}(\mathbf{e}) = \text{Var}(\mathbf{M}\mathbf{y}) = \mathbf{M} \text{Var}(\mathbf{y}) \mathbf{M}' = \mathbf{M}\boldsymbol{\Omega}\mathbf{M}$$

$$\begin{aligned} \mathbf{E}(\mathbf{e}'\mathbf{e}) &= \mathbf{E}\left(\sum_{t=1}^n e_t^2\right) = \sum_{t=1}^n \mathbf{E} e_t^2 = \sum_{t=1}^n \text{Var}(e_t^2) = \\ &= \text{tr}[\text{Var}(\mathbf{e})] = \text{tr}[\mathbf{M}\boldsymbol{\Omega}\mathbf{M}] = \text{tr}[\mathbf{M}^2\boldsymbol{\Omega}] = \text{tr}[\mathbf{M}\boldsymbol{\Omega}] \end{aligned}$$

Упражнение : $\text{tr}[\mathbf{A}\mathbf{B}] = \text{tr}[\mathbf{B}\mathbf{A}]$

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ И ОБЫЧНЫЙ МНК

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim (\mathbf{0}; \boldsymbol{\Omega}); \quad \text{Var}(\mathbf{y}) = \text{Var}(\mathbf{u}) = \boldsymbol{\Omega}$$

$$\text{Var}(\hat{\boldsymbol{\beta}}_{OLS}) = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$\hat{\sigma}^2 = \frac{\mathbf{e}'\mathbf{e}}{n-k} = \frac{\sum_{t=1}^n e_t^2}{n-k} \Rightarrow E\hat{\sigma}^2 = E\frac{\mathbf{e}'\mathbf{e}}{n-k} = \frac{\text{tr}[\mathbf{M}\boldsymbol{\Omega}]}{n-k}$$

$$\widehat{\text{Var}}\hat{\boldsymbol{\beta}}_{OLS} = \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1}$$

$$E\widehat{\text{Var}}\hat{\boldsymbol{\beta}}_{OLS} = \frac{\text{tr}[\mathbf{M}\boldsymbol{\Omega}]}{n-k} (\mathbf{X}'\mathbf{X})^{-1} \neq (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \text{Var}\hat{\boldsymbol{\beta}}$$

$\widehat{\text{Var}}\hat{\boldsymbol{\beta}}_{OLS}$ - смещена

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ

$$y = X\beta + u, \quad u \sim (0; \Omega)$$

Теорема Айткена :

Для обобщенной регрессионной модели оценка

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y \quad \text{является } \mathbf{BLUE} \text{ (best linear unbiased estimate)}$$

Лемма : $\exists \mathbf{R} : \det \mathbf{R} \neq 0, \mathbf{R}'\mathbf{R} = \Omega^{-1}$

Док - во : $\Omega^{-1} \geq 0, (\Omega^{-1})' = \Omega^{-1} \Rightarrow \exists \mathbf{S} : \mathbf{S}' = \mathbf{S}^{-1}$ (\mathbf{S} – ортогональная)

$\Omega^{-1} = \mathbf{S}'\mathbf{\Lambda}\mathbf{S}, \quad \Omega = \mathbf{S}'\mathbf{\Lambda}^{-1}\mathbf{S}, \quad \mathbf{\Lambda}$ – диагональная

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}, \quad \lambda_i - \text{с.ч.}\Omega^{-1} \Rightarrow \mathbf{R} = \mathbf{\Lambda}^{1/2}\mathbf{S}$$

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ

$$y = X\beta + u, \quad u \sim (0; \Omega)$$

Теорема Айткена :

Для обобщенной регрессионной модели оценка

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y \text{ является BLUE (best linear unbiased estimate)}$$

Док - во :

$$y^* = X^*\beta + u^*, \quad y^* = Ry; \quad X^* = RX; \quad u^* = Ru$$

Выполнены условия Теоремы Гаусса – Маркова, т.к. :

$$E u^* = 0; \quad \text{Var } u^* = R(\text{Var } u)R' = R\Omega R' = \Lambda^{1/2}SS'\Lambda^{-1}SS'\Lambda^{1/2} = \Lambda^{1/2}I\Lambda^{-1/2}\Lambda^{-1/2}I = I$$

$$\text{rank } X^* = \text{rank } X = k$$

$$\hat{\beta}_{OLS}^* = (X^{*'}X^*)^{-1} X^{*'}y^* = (X'R'RX)^{-1} X'R'Ry = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y = \hat{\beta}_{GLS}$$

Упражнение : доказать несмещенность $\hat{\beta}_{GLS}$

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim (\mathbf{0}; \boldsymbol{\Omega})$$

$$\hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y}$$

Замечание 1: $\hat{\boldsymbol{\beta}}_{GLS} = \arg \min (\mathbf{e}^{*'} \mathbf{e}^*) = \arg \min \left((\mathbf{y}^* - \mathbf{X}^* \boldsymbol{\beta})' (\mathbf{y}^* - \mathbf{X}^* \boldsymbol{\beta}) \right) =$
 $= \arg \min \left((\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{R}'\mathbf{R} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right) = \arg \min (\mathbf{e}' \boldsymbol{\Omega}^{-1} \mathbf{e})$

Упражнение: $\text{Var} \hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}$

Замечание 2: $R^2 = 1 - \frac{RSS}{TSS}$ не обладает свойствами обычного R^2 ,

в частности: $R^2 \notin [0;1]$, не монотонен по количеству регрессоров

ОБОБЩЕННАЯ РЕГРЕССИОННАЯ МОДЕЛЬ

$$y = X\beta + u, \quad u \sim (0; \Omega); \quad \hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y$$

$$y^* = X^*\beta + u^*, \quad u^* \sim (0; I); \quad y^* = Py; \quad X^* = PX; \quad u^* = P\varepsilon; \quad \hat{\beta}_{OLS} = (X^{*'}X^*)^{-1} X^{*'}y = \hat{\beta}_{GLS}$$

ВЫВОДЫ:

1. Для обобщенной регрессионной модели МНК-оценка остается несмещенной, состоятельной, но становится неэффективной
2. Оценка матрицы ковариаций МНК-оценки оказывается смещенной
3. Эффективной является оценка ОМНК
4. ОМНК-оценка может быть получена из МНК-оценки линейно-преобразованной исходной модели
5. Для нахождения ОМНК-оценки необходимо знать матрицу ковариаций ошибок
6. Естественным методом получения ОМНК-оценок является **ДОСТУПНЫЙ** ОБОБЩЕННЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ-оценка матрицы ковариаций ошибок и применение к ней ОМНК
7. Коэффициент детерминации R^2 не является адекватной мерой качества подгонки при использовании ОМНК

Конец лекции