2-3

普通最小二乘（OLS）及其数学性质

主讲人：范国斌

普通最小二乘法（OLS）（Ordinary Least Squares)

OLS的基本思想（以一元为例）：

- 对于 $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ 不同的估计方法可以得到不同的样本回归参数 $\hat{\beta}_1$ 和 $\hat{\beta}_2$ ，所估计的 $\hat{Y}_i$ 也就不同。
- 理想的估计方法应使估计的 $\hat{Y}_i$ 与真实的 Y_i 的差(即剩余 e_i ）总的来说越小越好
- 因 e_i 可正可负，总有 $\sum e_i = 0$, 所以可以取 $\sum e_i^2$ 最小，即

$$\min \sum e_i^2 = \min \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

在观测值Y和X确定时，$\sum e_i^2$ 的大小决定于 $\hat{\beta}_1$ 和 $\hat{\beta}_2$

多元线性回归模型的OLS估计

多元情形下原则相同：

寻求剩余平方和最小的参数估计式 $\min:\sum e_i^2=\sum(Y_i-\hat{Y}_i)^2$

$$\min:\sum e_i^2=\sum[Y_i-(\hat{\beta}_1+\hat{\beta}_2X_{2i}+\hat{\beta}_3X_{3i}+\cdots+\hat{\beta}_kX_{ki})]^2$$

即 $\min:\sum e_i^2=\min:\mathbf{e}'\mathbf{e}=\min:(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})$

求偏导，并令其为0　　$\partial(\sum e_i^2)/\partial\hat{\beta}_j=0$

即

$$-2\sum\left[Y_i-(\hat{\beta}_1+\hat{\beta}_2X_{2i}+\hat{\beta}_3X_{3i}+\cdots+\hat{\beta}_{ki}X_{ki})\right]=0\longrightarrow\sum e_i=0$$

$$-2\sum X_{2i}\left[Y_i-(\hat{\beta}_1+\hat{\beta}_2X_{2i}+\hat{\beta}_3X_{3i}+\cdots+\hat{\beta}_{ki}X_{ki})\right]=0\longrightarrow\sum X_{2i}e_i=0$$

$$\vdots$$

$$-2\sum X_{ki}\left[Y_i-(\hat{\beta}_1+\hat{\beta}_2X_{2i}+\hat{\beta}_3X_{3i}+\cdots+\hat{\beta}_{ki}X_{ki})\right]=0\longrightarrow\sum X_{ki}e_i=0$$

用矩阵表示的正规方程

偏导数

$$\begin{bmatrix} \sum e_i \\ \sum X_{2i}e_i \\ \vdots \\ \sum X_{ki}e_i \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{bmatrix}}_{\boldsymbol{X'}} \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}}_{\boldsymbol{e}} = \mathbf{X'e} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{0}}$$

因为样本回归函数为 $Y = X\hat{\beta} + e$

两边左乘 X' $X'Y = X'X\hat{\beta} + X'e$

根据最小二乘原则 $X'e = 0$

则正规方程为 $X'X\hat{\beta} = X'Y$

OLS估计式

由正规方程 $X'X\hat{\beta} = X'Y$　　$(X'X)_{k\times k}$ 是满秩矩阵,其逆存在

多元回归中 $\hat{\beta} = (X'X)^{-1}X'Y$

只有两个解释变量时：

$$\hat{\beta}_1 = \overline{Y} - \hat{\beta}_2\overline{X}_2 - \hat{\beta}_3\overline{X}_3$$

$$\hat{\beta}_2 = \frac{(\sum y_i x_{2i})(\sum x_{3i}^2) - (\sum y_i x_{3i})(\sum x_{2i}x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i}x_{3i})^2}$$

$$\hat{\beta}_3 = \frac{(\sum y_i x_{3i})(\sum x_{2i}^2) - (\sum y_i x_{2i})(\sum x_{2i}x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i}x_{3i})^2}$$

注意：x, y 为 X、Y 的离差

对比

一元线性回归中

$$\hat{\beta}_1 = \overline{Y} - \hat{\beta}_2\overline{X}$$

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$$

OLS估计的数学性质

- 回归线通过样本均值 $\overline{Y} = \hat{\beta}_1 + \hat{\beta}_2 \overline{X}_2 + \hat{\beta}_3 \overline{X}_3 + \cdots + \hat{\beta}_k \overline{X}_k$
- 估计值 $\hat{Y}_i$ 的均值等于实际观测值 Y_i 的均值 $\sum \hat{Y}_i / n = \overline{Y}$
- 剩余项 e_i 的均值为零 $\overline{e_i} = \sum e_i / n = 0$
- 被解释变量估计值 $\hat{Y}_i$ 与剩余项 e_i 不相关

$$Cov(\hat{Y}_i, e_i) = 0 \quad 或 \quad \sum (e_i \hat{y}_i) = 0$$

- 解释变量 X_i 与剩余项 e_i 不相关

$$Cov(X_{ji}, e_i) = 0$$

（j=2,3,---k）