OLS 估计量的性质的推导证明（一些补充）

1、 线性：

（ 1 ）证明斜率系数估计量 $\hat{\beta}$ 是 Y 的线性函数。

$$\hat{\beta}=\frac{\sum x_i y_i}{\sum x_i^2}=\frac{\sum x_i(Y_i-\bar{Y})}{\sum x_i^2}=\frac{\sum x_i Y_i}{\sum x_i^2}-\frac{\bar{Y}\sum x_i}{\sum x_i^2}$$

$$=(\text{由于}\sum x_i=\sum(X_i-\bar{X})=\sum X_i-n\bar{X}=0)$$

$$=\frac{\sum x_i Y_i}{\sum x_i^2}=\sum k_i Y_i,\qquad \text{其中} k_i=\frac{x_i}{\sum x_i^2}$$

注意：$\sum k_i=\sum(\frac{x_i}{\sum x_i^2})$ （由于对确定量X_i而言$\sum x_i^2$是定值)

$$=\frac{1}{\sum x_i^2}\cdot\sum x_i(\text{前已证}\sum x_i=0)=0,\text{故}\sum k_i=0$$

$$\text{又}\sum k_i x_i=\sum(\frac{x_i}{\sum x_i^2}\bullet x_i)=\sum\frac{x_i^2}{\sum x_i^2}=1$$

$$\text{故}\sum k_i X_i=\sum k_i(x_i+\bar{X})=\sum k_i x_i+\bar{X}\bullet\sum k_i(\text{前已证}\sum k_i x_i=1,\sum k_i=0)$$

$$=1+0=1,\text{故}\sum k_i X_i=1$$

记得$\sum k_i=0$与$\sum k_i X_i=1$,对后面的证明会有用。

（2） 证明截距系数估计量$\hat{\alpha}$是Y的线性函数。

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=\sum(\frac{1}{n}-k_i\bar{X})Y_i=\sum w_i Y_i,\text{其中}w_i=\frac{1}{n}-k_i\bar{X}$$

注意： $\sum w_i = \sum(\frac{1}{n} - k_i\overline{X}) = n \bullet \frac{1}{n} - \sum k_i\overline{X} = 1 - \overline{X} \bullet \sum k_i$

(前已证$\sum k_i = 0$) $= 1$ ；

$\sum w_i X_i = \sum(\frac{1}{n} - k_i\overline{X})X_i = \frac{1}{n} \bullet \sum X_i - \overline{X} \bullet \sum k_i X_i$(前已证$\sum k_i X_i = 1$)

$= \frac{1}{n} \bullet \sum X_i - \overline{X} = 0;$

注意$\sum w_i = 1, \sum w_i X_i = 0$，对后面的证明有用。

2、无偏：

(1) $\hat{\beta}$是β的无偏估计量。

$\hat{\beta} = \sum k_i Y_i = \sum k_i(\alpha + \beta X_i + \varepsilon_i) = \alpha \sum k_i + \beta \sum k_i X_i + \sum k_i \varepsilon_i$

（前已证$\sum k_i X_i = 1, \sum k_i = 0$) $= \beta + \sum k_i \varepsilon_i$

由于$E(\sum k_i \varepsilon_i) = E(k_1\varepsilon_1 + k_2\varepsilon_2 + \ldots + k_n\varepsilon_n) = E(k_1\varepsilon_1) + E(k_2\varepsilon_2) + \ldots + E(k_n\varepsilon_n)$

$= k_1 E(\varepsilon_1) + k_2 E(\varepsilon_2) + \ldots + k_n E(\varepsilon_n)$(注意假设$E(\varepsilon_i) = 0$) $= 0$

所以对等式$\hat{\beta} = \beta + \sum k_i \varepsilon_i$两边取期望有，$E(\hat{\beta}) = \beta + E(\sum k_i \varepsilon_i) = \beta$

(2) $\hat{\alpha}$是α的无偏估计量, 即$E(\hat{\alpha}) = \alpha$

证明方法同上，参考课件。注意利用$\sum w_i = 1, \sum w_i X_i = 0$

课件上有错误:$\hat{\alpha} = \alpha + \sum k_i \varepsilon_i$应改为$\hat{\alpha} = \alpha + \sum w_i \varepsilon_i$

3、有效性：

证明思路：先计算$\hat{\beta}$的方差$Var(\hat{\beta})$,再证明对任一线性无偏估计量$\hat{\beta}^*$,(即$\hat{\beta}^*$ 满足 $\hat{\beta}^* = \sum c_i Y_i$ 且$E(\hat{\beta}^*)=\beta$ ），均满足$Var(\hat{\beta}^*)\geq Var(\hat{\beta})$。对$\hat{\alpha}$的有效性证明思路同$\hat{\beta}$。

对$\hat{\alpha},\hat{\beta}$的最小方差性证明上课件已经说的比较清楚，

也没有错误。这里仅仅对$\hat{\alpha},\hat{\beta}$的计算作一些说明。

（1） 计算$\hat{\beta}$与$\hat{\alpha}$的方差。

$$Var(\hat{\beta})(\text{注意前面证明无偏性的时候已证}\ \hat{\beta}=\beta+\sum k_i\varepsilon_i)$$

$$=Var(\beta+\sum k_i\varepsilon_i)(\text{注意到}\beta\text{为常数})=Var(\sum k_i\varepsilon_i)$$

$$=Var(k_1\varepsilon_1+k_2\varepsilon_2+...+k_n\varepsilon_n)(\text{注意到随机变量}\varepsilon_i\text{独立})$$

$$=Var(k_1\varepsilon_1)+Var(k_2\varepsilon_2)+...Var(k_n\varepsilon_n)=k_1^2Var(\varepsilon_1)+...k_n^2Var(\varepsilon_n)$$

$$(\text{注意到随机变量}\varepsilon_i\text{方差相同, 为}\sigma^2)=\sigma^2\sum k_i^2$$

$$\text{注意到}k_i=\frac{x_i}{\sum x_i^2},\text{故}k_i^2=\frac{x_i^2}{(\sum x_i^2)^2},\sum k_i^2=\frac{\sum x_i^2}{(\sum x_i^2)^2}=\frac{1}{\sum x_i^2}$$

$$\text{所以}Var(\hat{\beta})=\frac{\sigma^2}{\sum x_i^2}$$

$$Var(\hat{\alpha})=(\text{前几步思路同上，见课件})=\sigma^2\sum w_i^2=\sigma^2\sum(\frac{1}{n}-k_i\overline{X})^2$$

$$=\sigma^2\sum(\frac{1}{n^2}-\frac{2}{n}k_i\overline{X}+k_i^2\overline{X}^2)(\text{这里课件上有错误,请注意})$$

$$=\sigma^2(\frac{1}{n^2}\bullet n-\frac{2}{n}\overline{X}\sum k_i+\overline{X}^2\sum k_i^2)(\text{注意前已证}\sum k_i=0,\sum k_i^2=\frac{1}{\sum x_i^2})$$

$$=\sigma^2(\frac{1}{n}+\frac{\overline{X}^2}{\sum x_i^2})=\frac{\sigma^2 X_i^2}{n\sum x_i^2}$$

$$\text{最后一个等号处，用逆推比较清楚:}\frac{\sum X_i^2}{n\sum x_i^2}=\frac{\sum(x_i+\overline{X})^2}{n\sum x_i^2}=$$

$$\frac{\sum x_i^2+2X\sum x_i+n\overline{X}^2}{n\sum x_i^2}=\frac{1}{n}+\frac{\overline{X}^2}{\sum x_i^2}$$

4、关于$\hat{\alpha},\hat{\beta}$的协方差计算：

课本的证明方法略显复杂：

在证明前先注意两个公式：若X,Y,Z,W是随机变量，a,b,c,d是常数，则有

$\mathrm{cov}(X+a,Y+b)=\mathrm{cov}(X,Y)$,

$\mathrm{cov}(aX+bY,cW+dZ)=ac\,\mathrm{cov}(X,W)+ad\,\mathrm{cov}(X,Z)+bc\,\mathrm{cov}(Y,W)+bd\,\mathrm{cov}(Y,Z)$

并注意两个对随机变量ε_i的假设:

对i $\neq$ j，有$\mathrm{cov}(\varepsilon_i,\varepsilon_j)=0$；对i $=$ j，$\mathrm{cov}(\varepsilon_i,\varepsilon_j)=\mathrm{cov}(\varepsilon_i,\varepsilon_i)=Var(\varepsilon_i)=\sigma^2$

故 $\mathrm{cov}(\hat{\alpha},\hat{\beta})=\mathrm{cov}(\alpha+\sum w_i\varepsilon_i,\beta+\sum k_i\varepsilon_i)$(注意到$\alpha,\beta$为常数)

$$=\mathrm{cov}(\sum_{i=1}^{n}w_i\varepsilon_i,\sum_{j=1}^{n}k_j\varepsilon_j)=\mathrm{cov}[(w_1\varepsilon_1+w_2\varepsilon_2+...w_n\varepsilon_n),(k_1\varepsilon_1+k_2\varepsilon_2+...k_n\varepsilon_n)]$$

(由于对i $\neq$ j，有$\mathrm{cov}(\varepsilon_i,\varepsilon_j)=0$,所以只需考虑i =j 的情况)

$$=\mathrm{cov}(w_1\varepsilon_1,k_1\varepsilon_1)+\mathrm{cov}(w_2\varepsilon_2,k_2\varepsilon_2)+...\mathrm{cov}(w_n\varepsilon_n,k_n\varepsilon_n)$$

$$=w_1k_1\mathrm{cov}(\varepsilon_1,\varepsilon_1)+w_2k_2\mathrm{cov}(\varepsilon_2,\varepsilon_2)+....+w_nk_n\mathrm{cov}(\varepsilon_n,\varepsilon_n)$$

（注意到有同方差假设，$\mathrm{cov}(\varepsilon_i,\varepsilon_i)=Var(\varepsilon_i)=\sigma^2$）

$$=\sigma^2\sum_{i=1}^{n}w_ik_i=\sigma^2\sum_{i=1}^{n}(\frac{1}{n}-\overline{X}k_i)k_i=\frac{\sigma^2}{n}\sum_{i=1}^{n}k_i-\sigma^2\overline{X}\sum_{i=1}^{n}k_i^2$$

(注意到前面已证$\sum_{i=1}^{n}k_i=0,\sum_{i=1}^{n}k_i^2=\frac{1}{\sum_{i=1}^{n}x_i^2}$)

$$=-\frac{\sigma^2\overline{X}}{\sum_{i=1}^{n}x_i^2}$$

一种比较简单的算法如下：

由于$\hat{\alpha}=\overline{Y}-\hat{\beta}\overline{X}$,所以$E(\hat{\alpha})=\overline{Y}-E(\hat{\beta})\overline{X}$,$\hat{\alpha}-E(\hat{\alpha})=-\overline{X}[\hat{\beta}-E(\hat{\beta})]$

故 $\mathrm{cov}(\hat{\alpha},\hat{\beta})=E\{[\hat{\alpha}-E(\hat{\alpha})][\hat{\beta}-E(\hat{\beta})]\}=-\overline{X}E[\hat{\beta}-E(\hat{\beta})]^2$

$=-\overline{X}\cdot\mathrm{Var}(\hat{\beta})$(在证明$\hat{\beta}$的有效性时已求得$\mathrm{Var}(\hat{\beta})=-\frac{\sigma^2}{\sum x_i^2}$)

$$=-\frac{\overline{X}\sigma^2}{\sum x_i^2}$$

5、证明 $\hat{\sigma}^2 = s^2 = \dfrac{\sum e_i^2}{n-2}$（课件上误作 $\dfrac{\sum e_i}{n-2}$）

$$e_i = Y_i - \hat{Y}_i = (\alpha - \hat{\alpha}) + (\beta - \hat{\beta})X_i + \varepsilon_i$$

$$e_i^2 = (\alpha - \hat{\alpha})^2 + (\beta - \hat{\beta})^2 X_i^2 + \varepsilon_i^2 + 2X_i(\hat{\alpha} - \alpha)(\hat{\beta} - \beta)$$

$$-2\varepsilon_i(\hat{\alpha} - \alpha) - 2\varepsilon_i(\hat{\beta} - \beta)X_i \text{（课件上此处有误，请注意）}$$

由于前面已算得：$\hat{\alpha} - \alpha = \sum w_i\varepsilon_i,\ \hat{\beta} - \beta = \sum k_i\varepsilon_i$

又因为 $E(\alpha - \hat{\alpha})^2 = Var(\hat{\alpha}), E(\beta - \hat{\beta})^2 = Var(\hat{\beta}), E(\varepsilon_i^2) = \sigma^2$

$$E(\hat{\alpha} - \alpha)(\hat{\beta} - \beta) = \mathrm{cov}(\hat{\alpha}, \hat{\beta})$$

当 $i \neq j, \varepsilon_i, \varepsilon_j$ 独立 故 $E(\varepsilon_i\varepsilon_j) = E(\varepsilon_i) \bullet E(\varepsilon_j) = 0$

所以 $E[\varepsilon_i(\hat{\alpha} - \alpha)] = E[\varepsilon_i(w_1\varepsilon_1 + w_2\varepsilon_2 + \ldots + w_i\varepsilon_i + \ldots w_n\varepsilon_n)]$

$$= w_i E(\varepsilon_i^2) = w_i\sigma^2$$

同理可算得：$E[\varepsilon_i(\hat{\beta} - \beta)] = k_i\sigma^2$

故 $E(e_i^2) = Var(\hat{\alpha}) + X_i^2 Var(\hat{\beta}) + \sigma^2 + 2X_i\mathrm{cov}(\hat{\alpha}, \hat{\beta}) - 2E[\varepsilon_i(\hat{\alpha} - \alpha)] - 2X_i E[\varepsilon_i(\hat{\beta} - \beta)]$

$$= \frac{\sigma^2\sum X_i^2}{n\sum x_i^2} + X_i^2 \bullet \frac{\sigma^2}{\sum x_i^2} + \sigma^2 - 2X_i \bullet \frac{\sigma^2\overline{X}}{\sum x_i^2} - 2w_i\sigma^2 - 2X_i k_i\sigma^2$$

两边求和得

$$E(\sum e_i^2) = n \bullet \frac{\sigma^2\sum X_i^2}{n\sum x_i^2} + \frac{\sigma^2}{\sum x_i^2}\sum X_i^2 + n\sigma^2 - \frac{2\sigma^2\overline{X}}{\sum x_i^2} \bullet \sum X_i - 2\sigma^2\sum w_i - 2\sigma^2\sum k_i X_i$$

$$= \frac{\sigma^2\sum X_i^2}{\sum x_i^2} + \frac{\sigma^2\sum X_i^2}{\sum x_i^2} + n\sigma^2 - \frac{2\sigma^2\overline{X}}{\sum x_i^2} \bullet \sum X_i - 2\sigma^2 - 2\sigma^2$$

$$= 2\sigma^2\frac{\sum X_i^2 - \overline{X}\sum X_i}{\sum x_i^2} + (n-4)\sigma^2$$

$$= 2\sigma^2\frac{\sum(\overline{X} + x_i)^2 - \overline{X}\sum X_i}{\sum x_i^2} + (n-4)\sigma^2$$

$$= 2\sigma^2\frac{n\overline{X}^2 + 2\overline{X}\sum x_i + \sum x_i^2 - \overline{X}\sum X_i}{\sum x_i^2} + (n-4)\sigma^2$$

$$= 2\sigma^2\frac{n\overline{X}^2 + \sum x_i^2 - \overline{X} \bullet n\overline{X}}{\sum x_i^2} + (n-4)\sigma^2 = (n-2)\sigma^2$$

故 $E(\dfrac{\sum e_i^2}{n-2}) = \sigma^2$，即 $\hat{\sigma}^2 = s^2 = \dfrac{\sum e_i^2}{n-2}$ 为 σ^2 的无偏估计.