线性回归习题

宋歌 2015080086 数52

3/14/2018

考虑线性模型$Y_i = \beta_1 + \beta_1 X_i + e_i, 1 \leqslant i \leqslant n$，$\beta_1$和$\beta_2$是未知参数，$x_i$是固定设计点，随机误差$e_1, \ldots, e_n$ iid, 满足$E(e_i) = 0, var(e_i) = \sigma^2$. 定义如下四个函数：

$$L_1(\beta_1, \beta_2) = \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 X_i)^2;$$

$$L_2(\beta_1, \beta_2) = \sum_{i=1}^{n} (X_i - \frac{Y_i - \beta_1}{\beta_2})^2;$$

$$L_3(\beta_1, \beta_2) = \sum_{i=1}^{n} \frac{(Y_i - \beta_1 - \beta_2 X_i)^2}{1 + \beta_2^2};$$

$$L_4(\beta_1, \beta_2) = \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 X_i)^2 + (X_i - \frac{Y_i - \beta_1}{\beta_2})^2;$$

对函数$L_k(\beta_1, \beta_2)$关于(β_1, β_2)求极小值得到β_1和β_2的估计为$\hat{\beta}_{k,1}$和$\hat{\beta}_{k,2}$，$k = 1, 2, 3, 4$。请回答下面的问题：

1. 试解释函数$L_k(\beta_1, \beta_2)$ $(k = 1, 2, 3, 4)$的几何意义。
   L_1表示观测值y_i到拟合直线$y = \beta_1 + \beta_2 x$的竖直距离平方和；
   L_2表示观测值x_i到拟合直线$y = \beta_1 + \beta_2 x$的水平距离平方和；
   L_3表示观测点(x_i, y_i)到拟合直线$y = \beta_1 + \beta_2 x$的距离平方和；
   L_4表示观测点(x_i, y_i)到拟合直线$y = \beta_1 + \beta_2 x$的（竖直距离平方+水平距离平方）和。

2. 证明：$\hat{\beta}_1$和$\hat{\beta}_2$分别是β_1和β_2的无偏估计和相合估计。

由

$$\begin{cases} \frac{\partial L_1}{\partial \beta_1} = 0 \\ \frac{\partial L_1}{\partial \beta_2} = 0 \end{cases}$$

解得

$$\begin{cases} \hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x} \\ \hat{\beta}_2 = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) / \sum_{i=1}^n (x_i - \bar{x})^2 \end{cases}$$

令

$$b_i = \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

则有

$$\hat{\beta}_2 = \beta_2 + \sum_{i=1}^n b_i e_i$$

从而

$$E(\hat{\beta}_2) = \beta_2 + E(\sum_{i=1}^n b_i e_i) = \beta_2$$

又

$$E(\hat{\beta}_1) = E(\beta_1 + \beta_2 \bar{x} + \bar{e}) - \beta_2 \bar{x} = \beta_1 + E(\bar{e}) = \beta_1$$

从而$\hat{\beta}_1$和$\hat{\beta}_2$分别是β_1和β_2的无偏估计。

现看

$$\hat{\beta}_{2,n} = \hat{\beta}_2 = \beta_2 + \sum_{i=1}^n b_i e_i$$

$$\begin{aligned} P(|\hat{\beta}_{2,n} - \beta_2| > \varepsilon) = P(|\sum_{i=1}^n b_i e_i| > \varepsilon) \leqslant \frac{1}{\varepsilon^2} E(\sum_{i=1}^n b_i e_i)^2 \\ = \frac{\sigma^2}{\varepsilon^2} \sum_{i=1}^n b_i^2 = \frac{\sigma^2}{\varepsilon^2} \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{aligned}$$

假设数据x_i方差>0，此时

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=n(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2)\to\infty$$

从而

$$lim_{n\to\infty}P(|\hat{\beta}_{2,n}-\beta_2|>\varepsilon)=0$$

即$\hat{\beta}_{2,n}=\hat{\beta}_2$是$\beta_2$的相合估计。

又

$$\hat{\beta}_{1,n}=\hat{\beta}_1=\bar{y}-\hat{\beta}_2\bar{x}=\beta_1+\beta_2\bar{x}+\bar{e}-\hat{\beta}_2\bar{x}$$

从而

$$\hat{\beta}_{1,n}-\beta_1=(\beta_2-\hat{\beta}_2)\bar{x}+\bar{e}$$

$$\begin{aligned}P(|\hat{\beta}_{1,n}-\beta_1|>\varepsilon)=P(|\bar{e}-\bar{x}\sum_{i=1}^{n}b_ie_i|>\varepsilon)&\leqslant P(|\bar{e}|+|\bar{x}\sum_{i=1}^{n}b_ie_i|>\varepsilon)\\&\leqslant P(|\bar{e}|>\varepsilon/2)+P(|\bar{x}\sum_{i=1}^{n}b_ie_i|>\varepsilon/2)\end{aligned}$$

已知

$$P(|\bar{x}\sum_{i=1}^{n}b_ie_i|>\varepsilon/2)\to 0$$

由$E(e_i)=E(\bar{e})=0$ 知$\bar{e}\to 0$，从而

$$P(|\bar{e}|>\varepsilon/2)\to 0,\forall\varepsilon>0$$

故有

$$P(|\hat{\beta}_{1,n}-\beta_1|>\varepsilon)\to 0\ (n\to\infty)$$

即$\hat{\beta}_{1,n}=\hat{\beta}_1$是$\beta_1$的相合估计。

3. 试讨论：$\hat{\beta}_{k,1}$和$\hat{\beta}_{k,2}$ $(k=2,3,4)$是否分别是β_1和β_2的（渐近）无偏估计和相合估计。

- $k=2$

  由
$$\begin{cases} \frac{\partial L_2}{\partial \beta_1}=0 \\ \frac{\partial L_2}{\partial \beta_2}=0 \end{cases}$$
  解得
$$\begin{cases} \hat{\beta}_{2,1}=\bar{y}-\hat{\beta}_{2,2}\bar{x} \\ \hat{\beta}_{2,2}=(\frac{1}{n}\sum_{i=1}^n y_i^2-(\bar{y})^2)/(\frac{1}{n}\sum_{i=1}^n x_iy_i-\bar{x}\bar{y}) \end{cases}$$
  将$y_i=\beta_1+\beta_2x_i+e_i$代入得：
$$\hat{\beta}_{2,2}=\frac{\frac{1}{n}\sum_{i=1}^n[(\beta_2x_i+e_i)-(\beta_2\bar{x}+\bar{e})]^2}{\beta_2(\frac{1}{n}\sum_{i=1}^n x_i^2-\bar{x}^2)+(\frac{1}{n}\sum_{i=1}^n x_ie_i-\bar{e}\bar{x})}$$
  令
$$S^2=\frac{1}{n}\sum_{i=1}^n[(\beta_2x_i+e_i)-(\beta_2\bar{x}+\bar{e})]^2$$
  得到
$$\frac{1}{\hat{\beta}_{2,2}}=\frac{\beta_2(\frac{1}{n}\sum_{i=1}^n x_i^2-\bar{x}^2)}{S^2}+\frac{\frac{1}{n}\sum_{i=1}^n x_ie_i-\bar{e}\bar{x}}{S^2}$$
  已知$x_ie_i-\bar{x}e_i$ iid，故
$$\frac{1}{n}\sum_{i=1}^n x_ie_i-\bar{e}\bar{x}=\frac{1}{n}\sum_{i=1}^n(x_ie_i-\bar{x}e_i)\xrightarrow{P}E(x_ie_i-\bar{x}e_i)=0$$
  化简S^2得
$$S^2=\beta_2^2(\frac{1}{n}\sum_{i=1}^n x_i^2-\bar{x}^2)+2\beta_2(\frac{1}{n}\sum_{i=1}^n x_ie_i-\bar{e}\bar{x})+(\frac{1}{n}\sum_{i=1}^n e_i^2-\bar{e}^2)$$
  其中已知
$$\frac{1}{n}\sum_{i=1}^n x_ie_i-\bar{e}\bar{x}=\frac{1}{n}\sum_{i=1}^n(x_ie_i-\bar{x}e_i)\xrightarrow{P}E(x_ie_i-\bar{x}e_i)=0$$

$$\frac{1}{n}\sum_{i=1}^{n} e_i^2 - \bar{e}^2 = \frac{1}{n}\sum_{i=1}^{n}(e_i - \bar{e})^2 \xrightarrow{P} \sigma^2$$

故有

$$S^2 \xrightarrow{P} \beta_2^2(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2) + \sigma^2$$

从而

$$1/\hat{\beta}_{2,2} \xrightarrow{P} \frac{\beta_2(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2)}{\beta_2^2(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2)} + \sigma^2 = \frac{\beta_2}{(\beta_2^2 + \sigma^2/Var(x_i))}$$

假设$n \to \infty$时，$\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = Var(x) \to \infty$，则有$1/\hat{\beta}_{2,2} \xrightarrow{P} 1/\beta_2$
从而

$$\hat{\beta}_{2,2} \xrightarrow{P} \beta_2$$

即$\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = Var(x) \to \infty$ $(n \to \infty)$时，$\hat{\beta}_{2,2}$ 是β_2的相合估计。
又$E(\hat{\beta}_{2,1}) = E(\beta_1 + \beta_2\bar{x} + \bar{e}) - E(\hat{\beta}_{2,2})\bar{x}$，从而当$\hat{\beta}_{2,2}$是$\beta_2$的无偏估计时（不一定），$\hat{\beta}_{2,1}$ 是β_1的无偏估计。

- $k = 3$

  易知

  $$L_3(\beta_1, \beta_2) = \frac{1}{1+\beta_2^2}L_1(\beta_1, \beta_2) = \frac{\beta_2^2}{1+\beta_2^2}L_2(\beta_1, \beta_2)$$

  从而

  $$L_3(\hat{\beta}_{3,1}, \hat{\beta}_{3,2}) = \frac{1}{1+\hat{\beta}_{3,2}^2}L_1(\hat{\beta}_{3,1}, \hat{\beta}_{3,2}) = \frac{\hat{\beta}_{3,2}^2}{1+\hat{\beta}_{3,2}^2}L_2(\hat{\beta}_{3,1}, \hat{\beta}_{3,2})$$

  其中

  $$\begin{aligned} L_3(\hat{\beta}_{3,1}, \hat{\beta}_{3,2}) &= \frac{1}{1+\hat{\beta}_{3,2}^2}L_1(\hat{\beta}_{3,1}, \hat{\beta}_{3,2}) \geqslant \frac{1}{1+\hat{\beta}_{3,2}^2}L_1(\hat{\beta}_{1,1}, \hat{\beta}_{1,2}) \\ &= \frac{1+\hat{\beta}_{1,2}^2}{1+\hat{\beta}_{3,2}^2}L_3(\hat{\beta}_{1,1}, \hat{\beta}_{1,2}) \geqslant \frac{1+\hat{\beta}_{1,2}^2}{1+\hat{\beta}_{3,2}^2}L_3(\hat{\beta}_{3,1}, \hat{\beta}_{3,2}) \end{aligned}$$

$$L_3(\hat{\beta}_{3,1},\hat{\beta}_{3,2})=\frac{\hat{\beta}_{3,2}^2}{1+\hat{\beta}_{3,2}^2}L_2(\hat{\beta}_{3,1},\hat{\beta}_{3,2})\geqslant\frac{\hat{\beta}_{3,2}^2}{1+\hat{\beta}_{3,2}^2}L_2(\hat{\beta}_{2,1},\hat{\beta}_{2,2})$$
$$=\frac{\hat{\beta}_{3,2}^2}{1+\hat{\beta}_{3,2}^2}\frac{1+\hat{\beta}_{2,2}^2}{\hat{\beta}_{2,2}^2}L_3(\hat{\beta}_{2,1},\hat{\beta}_{2,2})\geqslant\frac{\hat{\beta}_{3,2}^2}{1+\hat{\beta}_{3,2}^2}\frac{1+\hat{\beta}_{2,2}^2}{\hat{\beta}_{2,2}^2}L_3(\hat{\beta}_{3,1},\hat{\beta}_{3,2})$$

从而有

$$\frac{1+\hat{\beta}_{1,2}^2}{1+\hat{\beta}_{3,2}^2}\leqslant 1,\ \frac{\hat{\beta}_{3,2}^2}{1+\hat{\beta}_{3,2}^2}\frac{1+\hat{\beta}_{2,2}^2}{\hat{\beta}_{2,2}^2}\leqslant 1$$

解得

$$\hat{\beta}_{1,2}^2\leqslant\hat{\beta}_{3,2}^2\leqslant\hat{\beta}_{2,2}^2$$

已证得

$$\hat{\beta}_{1,2}\xrightarrow{P}\beta_2,\ \hat{\beta}_{2,2}\xrightarrow{P}\beta_2$$

故有

$$\hat{\beta}_{3,2}\xrightarrow{P}\beta_2$$

即$\frac{1}{n}\sum_{i=1}^n(x_i-\bar{x})^2=Var(x)\to\infty\ (n\to\infty)$时，$\hat{\beta}_{3,2}$ 是β_2的相合估计。

又

$$\frac{\partial L_3}{\partial\beta_1}=0\Rightarrow\hat{\beta}_{3,1}=\bar{y}-\hat{\beta}_{3,2}\bar{x}$$

故$E(\hat{\beta}_{3,1})=E(\beta_1+\beta_2\bar{x}+\bar{e})-E(\hat{\beta}_{3,2})\bar{x}$，从而当$\hat{\beta}_{3,2}$ 是β_2的无偏估计时（不一定），$\hat{\beta}_{3,1}$ 是β_1的无偏估计。

- $k=4$

  易知

$$L_4(\beta_1,\beta_2)=(1+\frac{1}{\beta_2^2})L_1(\beta_1,\beta_2)=(1+\beta_2^2)L_2(\beta_1,\beta_2)$$

  从而

$$L_4(\hat{\beta}_{4,1},\hat{\beta}_{4,2})=(1+\frac{1}{\hat{\beta}_{4,2}^2})L_1(\hat{\beta}_{4,1},\hat{\beta}_{4,2})=(1+\hat{\beta}_{4,2}^2)L_2(\hat{\beta}_{4,1},\hat{\beta}_{4,2})$$

其中

$$L_4(\hat{\beta}_{4,1},\hat{\beta}_{4,2}) = (1+\frac{1}{\hat{\beta}_{4,2}^2})L_1(\hat{\beta}_{4,1},\hat{\beta}_{4,2}) \geqslant (1+\frac{1}{\hat{\beta}_{4,2}^2})L_1(\hat{\beta}_{1,1},\hat{\beta}_{1,2})$$

$$= (1+\frac{1}{\hat{\beta}_{4,2}^2})\frac{\hat{\beta}_{1,2}^2}{1+\hat{\beta}_{1,2}^2}L_4(\hat{\beta}_{1,1},\hat{\beta}_{1,2}) \geqslant (1+\frac{1}{\hat{\beta}_{4,2}^2})\frac{\hat{\beta}_{1,2}^2}{1+\hat{\beta}_{1,2}^2}L_4(\hat{\beta}_{4,1},\hat{\beta}_{4,2})$$

$$L_4(\hat{\beta}_{4,1},\hat{\beta}_{4,2}) = (1+\hat{\beta}_{4,2}^2)L_2(\hat{\beta}_{4,1},\hat{\beta}_{4,2}) \geqslant (1+\hat{\beta}_{4,2}^2)L_2(\hat{\beta}_{2,1},\hat{\beta}_{2,2})$$

$$= (1+\hat{\beta}_{4,2}^2)\frac{1}{1+\hat{\beta}_{2,2}^2}L_4(\hat{\beta}_{2,1},\hat{\beta}_{2,2}) \geqslant (1+\hat{\beta}_{4,2}^2)\frac{1}{1+\hat{\beta}_{2,2}^2}L_4(\hat{\beta}_{4,1},\hat{\beta}_{4,2})$$

从而有

$$(1+\frac{1}{\hat{\beta}_{4,2}^2})\frac{\hat{\beta}_{1,2}^2}{1+\hat{\beta}_{1,2}^2} \leqslant 1,\ (1+\hat{\beta}_{4,2}^2)\frac{1}{1+\hat{\beta}_{2,2}^2} \leqslant 1$$

解得

$$\hat{\beta}_{1,2}^2 \leqslant \hat{\beta}_{4,2}^2 \leqslant \hat{\beta}_{2,2}^2$$

已证得

$$\hat{\beta}_{1,2} \xrightarrow{P} \beta_2,\ \hat{\beta}_{2,2} \xrightarrow{P} \beta_2$$

故有

$$\hat{\beta}_{4,2} \xrightarrow{P} \beta_2$$

即$\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = Var(x) \to \infty\ (n \to \infty)$时，$\hat{\beta}_{4,2}$ 是β_2的相合估计。