Limited Dependent Variables

Professor V.M. Auken, PhD

Maximum likelihood estimation

In the model $y = X\beta + \varepsilon$, it is assumed that the errors are normally and independently distributed with constant variance σ^2 i.e.,

$$\varepsilon \sim N(0, \sigma^2 I).$$

The normal density function for the errors is

$$f(\varepsilon_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}\varepsilon_i^2\right] \quad i = 1, 2, \ldots, n.$$

The likelihood function is the joint density of $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ given as

$$\begin{aligned} L(\beta, \sigma^2) &= \prod_{i=1}^{n} f(\varepsilon_i) \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^2\right] \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\varepsilon'\varepsilon\right] \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right]. \end{aligned}$$

Since the log transformation is monotonic, so we maximize $\ln L(\beta,\sigma^2)$ instead of $L(\beta,\sigma^2)$.

$$\ln L(\beta,\sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta).$$

The maximum likelihood estimators (m.l.e.) of β and σ^2 are obtained by equating the first order derivatives of $\ln L(\beta,\sigma^2)$ with respect to β and σ^2 to zero as follows:

$$\frac{\partial \ln L(\beta,\sigma^2)}{\partial\beta} = \frac{1}{2\sigma^2}2X'(y - X\beta) = 0$$

$$\frac{\partial \ln L(\beta,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(y - X\beta)'(y - X\beta).$$

The likelihood equations are given by

$$X'X\beta = X'y$$

$$\hat{\sigma}^2 = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta}).$$

Since rank$(X) = k$, so that the unique m.l.e. of β and σ^2 are obtained as

$$\tilde{\beta} = (X'X)^{-1}X'y$$

$$\tilde{\sigma}^2 = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta}).$$

Next we verify that these values maximize the likelihood function. First we find

$$\frac{\partial^2 \log L(\beta,\sigma^2)}{\partial \beta^2} = -\frac{1}{\sigma^2}X'X$$

$$\frac{\partial^2 \log L(\beta,\sigma^2)}{\partial^2(\sigma^2)^2} = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}(y - X\beta)'(y - X\beta)$$

$$\frac{\partial^2 \log L(\beta,\sigma^2)}{\partial \beta \partial \sigma^2} = -\frac{1}{\sigma^4}X'(y - X\beta).$$

Thus the Hessian matrix of second order partial derivatives of $\ln L(\beta,\sigma^2)$ with respect to β and σ^2 is

$$\begin{pmatrix} \dfrac{\partial^2 \ln L(\beta,\sigma^2)}{\partial\beta^2} & \dfrac{\partial^2 \ln L(\beta,\sigma^2)}{\partial\beta\partial\sigma^2} \\ \dfrac{\partial^2 \log L(\beta,\sigma^2)}{\partial\sigma^2\partial\beta} & \dfrac{\partial^2 \ln L(\beta,\sigma^2)}{\partial^2(\sigma^2)^2} \end{pmatrix}$$

which is negative definite at $\beta = \tilde{\beta}$ and $\sigma^2 = \tilde{\sigma}^2$.

This ensures that the likelihood function is maximized at these values.

Comparing with OLSEs, we find that

i. OLSE and m.l.e. of β are same. So m.l.e. of β is also an unbiased estimator of β.

ii. OLSE of σ^2 is s^2 which is related to m.l.e. of σ^2 as $\tilde{\sigma}^2 = \frac{n-k}{n}s^2$. So m.l.e. of σ^2 is a biased estimator of σ^2.

Cramer-Rao lower bound

Let $\theta = (\beta, \sigma^2)'$. Assume that both β and σ^2 are unknown. If $E(\hat{\theta}) = \theta$, then the Cramer-Rao lower bound for $\hat{\theta}$ is grater than or equal to the matrix inverse of

$$I(\theta) = -E\left[\frac{\partial^2 \ln L(\theta)}{\partial\theta\partial\theta'}\right]$$

$$= \begin{bmatrix} -E\left[\frac{\partial \ln L(\beta,\sigma^2)}{\partial\beta^2}\right] & -E\left[\frac{\partial \ln L(\beta,\sigma^2)}{\partial\beta\partial\sigma^2}\right] \\ -E\left[\frac{\partial \ln L(\beta,\sigma^2)}{\partial\sigma^2\partial\beta}\right] & -E\left[\frac{\partial \ln L(\beta,\sigma^2)}{\partial^2(\sigma^2)^2}\right] \end{bmatrix}$$

$$= \begin{bmatrix} -E\left[-\frac{X'X}{\sigma^2}\right] & -E\left[\frac{X'(y-X\beta)}{\sigma^4}\right] \\ -E\left[\frac{(y-X\beta)'X}{\sigma^4}\right] & -E\left[\frac{n}{2\sigma^4} - \frac{(y-X\beta)'(y-X\beta)}{\sigma^6}\right] \end{bmatrix}$$

$$= \begin{bmatrix} \frac{X'X}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{bmatrix}.$$

Then

$$[I(\theta)]^{-1} = \begin{bmatrix} \sigma^2(X'X)^{-1} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{bmatrix}$$

is the Cramer-Rao lower bound matrix of β and σ^2.

The covariance matrix of OLSEs of β and σ^2 is

$$\Sigma_{OLS} = \begin{bmatrix} \sigma^2(X'X)^{-1} & 0 \\ 0 & \frac{2\sigma^4}{n-k} \end{bmatrix}$$

which means that the Cramer-Rao

bound is attained for the covariance of b but not for s^2.

QUESTIONS!