Parameter estimation in linear regression

TMA4267 L9

14.02.2017

$$\underset{n\times 1}{Y} = \underset{n\times p}{X}\underset{p\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \quad , \; E(\varepsilon) = 0, \; Cov(\varepsilon) = \sigma^2 I$$

Estimate β by minimizing $LS(\beta)$

(i) $LS(\beta) = (y - X\beta)^T(y - X\beta) \leftarrow \sum_{i=1}^{n}(y_i - x_i^T\beta)^2$

$$= y^Ty - 2y^TX\beta + \beta^TX^TX\beta$$

(ii) $\dfrac{\partial LS(\beta)}{\partial \beta} = 0$

$$\frac{\partial}{\partial\beta}\left(y^Ty - 2y^TX\beta + \beta^TX^TX\beta \right) = 0$$

Rules (see slides)

$$d^T = -2y^TX$$

$$D = X^TX$$

$$0 - 2X^Ty + 2X^TX\beta = 0$$

$$\boxed{X^Ty = X^TX\beta \qquad \text{normal equations}}$$

iii) Solving the normal equations (substitute β with $\hat{\beta}$)

$$\underline{\underline{\hat{\beta} = (X^T X)^{-1} X^T Y}}$$

iv) min or max

$$\frac{\partial^2}{\partial \beta^2} LS(\beta) = \frac{\partial}{\partial \beta}\left(-2X^T y - 2X^T X \beta\right)\Big|_{\beta=\hat{\beta}}$$

$$= 2X^T X$$

If this matrix has only positive eigenvalues this will be the minimum.

<u>Questions on slide:</u>

X rank p : $X^T X$ $p \times p$ matrix
symmetric
positive definite
inverse exists

If $v^T X^T X v > 0$ for all $v \neq 0$ then $X^T X$ positive definite.

$$v^T X X v = (Xv)^T (Xv) \geq 0$$

Assume that $v^T X^T X v = 0$ then $Xv = 0$.
If X has full rank then $Xv = 0$ only
has $v = 0$ as solution.
$\Rightarrow$ then $X^T X$ must be positive definite.

$\hat{\beta} = (X^T X)^{-1} X^T Y$ is the least squares
estimator of β. If we assume a
normal linear model then $\hat{\beta}$ is also
the maximum likelihood estimator.

Ex: Acid rain: fit in R using lm

$\hat{\beta}_0$ = "Estimate Intercept" = 5.67

= estimate of the pH in a lake when

$x_1 = 0, x_2 = 0, \ldots, \underset{\text{Telework}}{\underset{\shortparallel}{x_7}} = 0$.

$\underset{\text{Ca}}{\hat{\beta}_3}$ = 0.975

if Ca increase by one unit, and all the other covarietes are kept constant, then we predict that pH will increase by $\hat{\beta}_3 = 0.975$.

St. Error: $\hat{SD}(\hat{\beta})$

Properties of $\hat{\beta}$

$$\hat{\beta} = \underbrace{(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T}_{\substack{C \\ \parallel \\ \text{constants}}} \underset{\text{RV}}{\underset{\uparrow}{Y}} \qquad \text{and} \qquad E(Y) = \mathbf{X}\beta$$

$$Cov(Y) = \sigma^2 I$$

$$Y = \mathbf{X}\beta + \varepsilon \quad E(\varepsilon) = 0, \quad Cov(\varepsilon) = \sigma^2 I$$

Find $E(\hat{\beta})$ and $Cov(\hat{\beta})$:

$$E(\hat{\beta}) = E(CY) = C \underset{\mathbf{X}\beta}{\underset{\parallel}{E(Y)}} = \underbrace{(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}}_{I}\beta = \underline{\underline{\beta}}$$

unbiased

$$Cov(\hat{\beta}) = Cov(CY) = C \underbrace{Cov(Y)}_{\sigma^2 I} C^T$$

$$= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \sigma^2 I \left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\right]^T$$

$$= \sigma^2 (\mathbf{X}^T\mathbf{X})^{-1} \underbrace{\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}}_{I}$$

$\mathbf{X}^T\mathbf{X}$ is symmetric
$(\mathbf{X}^T\mathbf{X})^{-1}$ — " —

$$= \underline{\underline{\sigma^2 (\mathbf{X}^T\mathbf{X})^{-1}}}$$

In a normal model: $\hat{\beta} \sim N_p\left(\beta, \sigma^2 (\mathbf{X}^T\mathbf{X})^{-1}\right)$

Ex: Acid rain: What is $Var(\hat{\beta}_3)$?

$$\hat{SD}(\hat{\beta}_3) = 0.144 : \sqrt{\hat{\sigma}^2 \cdot \underbrace{(\mathbf{X}^T\mathbf{X})^{-1}[4,4]}_{.55}}$$

Need estimator for σ.

Last information on $\hat{\beta}$: From part 1:

$$(\hat{\beta} - E(\hat{\beta}))^T \, Cov(\hat{\beta})^{-1} (\hat{\beta} - E(\hat{\beta})) \sim \chi^2_p$$

$$\frac{1}{\sigma^2} (\hat{\beta} - \beta)^T (X^T X) (\hat{\beta} - \beta) \sim \chi^2_p$$

$$Cov(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$$
$$Cov(\hat{\beta})^{-1} = \frac{1}{\sigma^2} (X^T X)$$

Estimator for σ^2 [f 3.2.2]

$$Y = X\beta + \varepsilon \,, \quad \varepsilon \sim N_n(0, \sigma^2 I)$$

$$L(\beta, \sigma^2) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} \exp\left\{ -\frac{1}{2\sigma^2} (y - X\beta)^T (y - X\beta) \right\}$$

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

$$\frac{\partial l}{\partial \sigma^2} = 0 \iff \qquad \frac{d \ln x}{dx} = \frac{1}{x} \,, \quad \frac{d \frac{1}{x}}{dx} = -\frac{1}{x^2}$$

$$0 - \frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2} \frac{1}{\sigma^4} (y - X\hat{\beta})^T (y - X\hat{\beta}) = 0$$

$$\frac{n}{\sigma^2} = \frac{1}{\sigma^4}(y-X\hat{\beta})^T(y-X\hat{\beta})$$

$$\hat{\sigma}^2_{ML} = \frac{1}{n}\underbrace{(Y-X\hat{\beta})^T(Y-X\hat{\beta})}_{\hat{\varepsilon}} = \frac{1}{n}\hat{\varepsilon}^T\hat{\varepsilon}$$

$\hat{\varepsilon}^T\hat{\varepsilon}$ = sums of squares of errors SSE

But, this estimator is rarely used, because it is biased (use tr-formula Part 1 to find the mean)

However: is unbiased

$$\hat{\sigma}^2 = \frac{1}{n-p}\hat{\varepsilon}^T\hat{\varepsilon}$$

REML found by maximizing

restricted

$$L(\sigma^2) = \int_{\beta} L(\beta, \sigma^2)\,d\beta$$

TMA4295 Statistical Inference: more on this.

Ex: Acid rain $\hat{SD}(\hat{\beta}_3) = \sqrt{(X^TX)^{-1}[4,4]\cdot\hat{\sigma}^2}$

$\hat{\sigma}^2 = \frac{1}{n-p}\hat{\varepsilon}^T\hat{\varepsilon}$ $\hat{\sigma} = 0.1165$

$\hat{\sigma}$ is residual standard error in printout

Predicted values and residuals [F3.2.1]

$$E(Y) = X\beta \,, \quad \text{so} \quad \widehat{E(Y)} = X\hat{\beta} \equiv \hat{Y} \leftarrow \text{prediction}$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

$$\hat{Y} = X\hat{\beta} = \underbrace{X(X^T X)^{-1} X^T}_{H} Y = HY$$

$\underset{n\times n}{H} = X(X^TX)^{-1}X^T$ is called the "hat matrix" for putting the hat on Y.

Residuals: $\hat{\varepsilon} = Y - \hat{Y} = \underset{\underset{IY}{\parallel}}{Y} - HY = (I-H)Y$

Observe (see also RecEx3.P3a) that

$\underset{n\times n}{H}$ is symmetric
is idempotent $H^2 = H$
has rank p $\Rightarrow$ show this

$\underset{n\times n}{(I-H)}$ is also symmetric and idempotent,
and $\text{rank}(I-H) = n-p$. $\Rightarrow$ show this

$\Rightarrow$ Work $\underbrace{\text{with RecEx3.P3}}_{\text{supervision Thurs. 16.15 at Smia.}}$ - and be ready for L10!