

Least Squares Regression

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_{p-1} x_{ip-1} + \varepsilon_i$$

Y_i = i th response

x_1, \dots, x_{p-1} predictors/coordinates/explanatory variables

ε_i = i th error term

In matrix form

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon}$$

$$\underline{y} = (y_1, \dots, y_n)'$$

$$X = (1, x_1, \dots, x_{p-1})$$

$$\underline{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)'$$

$$\underline{\beta} = (\beta_0, \dots, \beta_{p-1})'$$

We assume $E(\underline{\varepsilon}) = \underline{0}$, $\text{var}(\underline{\varepsilon}) = \Sigma$

Typically $\underline{\varepsilon} \sim N(\underline{0}, \Sigma)$, $\Sigma = \sigma^2 I$

Since $y_i = (\beta_0 + \beta_1 \bar{x}_1 + \dots + \beta_{p-1} \bar{x}_{p-1})$

$$+ \beta_1 (x_{i1} - \bar{x}_1) + \dots + \beta_{p-1} (x_{ip-1} - \bar{x}_{p-1}) + \varepsilon_i$$

$$= \beta_0^* + \beta_1 x_{i1}^* + \dots + \beta_{p-1} x_{ip-1}^* + \varepsilon_i$$

$$Y = X^* \beta^* + E$$

where $X^* = [1, x_1 - 1 \bar{x}_1, \dots, x_{p-1} - 1 \bar{x}_{p-1}]$
 $= [1, Cx_1, \dots, C\bar{x}_{p-1}]$

Ordinary Least Squares (OLS)

$$\Sigma = \sigma^2 I$$

$$\begin{aligned} Q &= \sum_i^n (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_{p-1} x_{ip-1})^2 \\ &= (Y - X\beta)'(Y - X\beta) \\ &= Y'Y - 2Y'X\beta + \beta'X'X\beta \end{aligned}$$

$$\frac{\partial Q}{\partial \beta} = -2X'Y + 2X'X\beta$$

$$\frac{\partial Q}{\partial \beta} = 0 \quad X'X\beta = X'Y \quad \text{least squares equation}$$

Assume (for now) that $\text{rank}(X) = p < n$

Then $X'X$ $p \times p$ $\text{rank} = p$

Solving gives $\hat{\beta} = (X'X)^{-1}X'Y$

$$\begin{aligned} \text{(i)} \quad E(\hat{\beta}) &= (X'X)^{-1} X' E(Y) \\ &= (X'X)^{-1} X' \kappa \beta \\ &= \beta \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad \text{cov}(\hat{\beta}) &= (X'X)^{-1} X' \sigma^2 I X (X'X)^{-1} \\ &= \sigma^2 (X'X)^{-1} \end{aligned}$$

(iii) fitted values

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

$$E(\hat{\mu}) = X\beta = \mu$$

$$\text{cov}(\hat{\mu}) = \sigma^2 X(X'X)^{-1}X'$$

(iv) residual vector

$$Y - \hat{\mu} = [I - X(X'X)^{-1}X']Y$$

$$E(Y - \hat{\mu}) = 0$$

$$\text{var}(Y - \hat{\mu}) = \sigma^2 [I - X(X'X)^{-1}X']$$

• sum of squared residuals

$$\hat{Q} = (Y - \hat{\mu})'(Y - \hat{\mu}) = Y' [I - X(X'X)^{-1}X']Y$$

$$\begin{aligned} E(\hat{Q}) &= \text{tr}([I - X(X'X)^{-1}X'] \sigma^2 I) \\ &\quad + \beta' X' [] X \beta \\ &= \sigma^2 (n - p) + 0 \end{aligned}$$

$$\text{Hence } \hat{\sigma}^2 = \frac{1}{n-p} \mathbf{y}' [\mathbf{I} - \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'] \mathbf{y}$$

is an unbiased estimator of σ^2

- Distribution results if $\mathbf{y} \sim N_n(\mathbf{x}\beta, \sigma^2 \mathbf{I})$

$$(i) \hat{\beta} \sim N_p[\beta, \sigma^2(\mathbf{x}'\mathbf{x})^{-1}]$$

$$(ii) \hat{\mu} \sim N_n[\mathbf{x}\beta, \sigma^2 \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}']$$

$$(iii) \mathbf{y} - \hat{\mu} \sim N_n(0, \sigma^2 [\mathbf{I} - \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'])$$

Note that $\hat{\mu}$ and $\mathbf{y} - \hat{\mu}$ are independent

Also, $\hat{\beta}$ and $\mathbf{y} - \hat{\mu}$ are independent

$$(iv) \hat{\sigma}^2 \sim \sigma^2 \chi_{n-p}^2(0) / (n-p)$$

independent of $\hat{\beta}$ b/c

$$\hat{\sigma}^2 = \frac{1}{n-p} \mathbf{y}' \mathbf{A} \mathbf{y} \quad \text{where } \mathbf{A}'\mathbf{A} = \sigma^2 \mathbf{A}$$

Best Linear Unbiased Estimator (BLUE)

$$\mathbf{y} \sim (\mathbf{x}\beta, \sigma^2 \mathbf{I})$$

- $a_0 + a_1' \mathbf{y}$ is a linear unbiased estimator of $t'\beta$ if

$$E(a_0 + a_1' \mathbf{y}) = t'\beta \quad \text{for all } \beta$$

Gauss - Markou Theorem

The best linear unbiased estimator of $t'\beta$ (ie. the one with the smallest variance) is $t'\hat{\beta}$ where $\hat{\beta} = (X'X)^{-1}X'Y$ is the OLS estimator.

Proof Suppose $a_0 + a'Y$ is a LUE of $t'\beta$.

$$\begin{aligned} a' &= a' - t'(X'X)^{-1}X' + t'(X'X)^{-1}X' \\ &= b' + t'(X'X)^{-1}X' \end{aligned}$$

Then

$$\begin{aligned} t'\beta &= E(a_0 + a'Y) \\ &= a_0 + E(b'Y) + E(t'(X'X)^{-1}X'Y) \\ &= a_0 + b'X\beta + t'\beta \end{aligned}$$

for all β . It follows that $a_0 = 0$ and $b'X = 0$.

$$\begin{aligned} \text{var}(a_0 + a'Y) &= \sigma^2 a'a \\ &= \sigma^2 [b' - t'(X'X)^{-1}X'] [b - X(X'X)^{-1}t] \\ &= \sigma^2 b'b + t'(X'X)^{-1}t \\ &\geq t'(X'X)^{-1}t = \text{var}(t'\hat{\beta}) \end{aligned}$$

$$\text{ex. } \mathbf{Y} \sim \left(\frac{1}{n} \boldsymbol{\beta}_0, \sigma^2 \mathbf{I} \right)$$

OLS estimate of $\boldsymbol{\beta}_0$ is $\hat{\boldsymbol{\beta}}_0 = (\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'\mathbf{Y} = \bar{\mathbf{Y}}$

But $\bar{Y}_w = \sum_{i=1}^n w_i Y_i$ where $\sum_i w_i = 1$

is also a linear u/b estimator of $\boldsymbol{\beta}_0$

$$\text{var}(\bar{Y}_w) = \sigma^2 \sum w_i^2 = \sigma^2 \underline{w}' \underline{w}$$

$$\text{var}(\bar{Y}) = \sigma^2/n = \sigma^2 \frac{1}{n} \mathbf{1}' \mathbf{1} \frac{1}{n}$$

$$\text{But } 1 = (\underline{w}' \frac{1}{n})^2 \leq (\mathbf{1}' \mathbf{1}) (\underline{w}' \underline{w})$$

by Cauchy-Schwarz

$$\text{So } \frac{1}{n} \leq \underline{w}' \underline{w}$$

Generalized Least Squares

$$\mathbf{Y} \sim (\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{V}) \quad \mathbf{V} \text{ known and p.d.}$$

$$\mathbf{Y}^* = \mathbf{V}^{-1/2} \mathbf{Y} \sim (\mathbf{V}^{-1/2} \mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$$

The GLS estimator of $\boldsymbol{\beta}$ is the OLS with \mathbf{Y}^* in place of \mathbf{Y} and $\mathbf{X}^* = \mathbf{V}^{-1/2} \mathbf{X}$ in place of \mathbf{X} ; i.e.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{*'} \mathbf{X}^*)^{-1} \mathbf{X}^{*'} \mathbf{Y}^* = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1} \mathbf{Y}$$

Weighted Least Squares

Suppose $\mathbf{V} = \mathbf{W}^{-1}$ where $\mathbf{W} = \text{diag}(w_i)$ with $w_i > 0$ for all i

$$\text{i.e. } \text{var}(Y_i) = \sigma^2 / w_i$$

$$\text{cov}(Y_i, Y_j) = 0$$

$$\text{Then } \hat{\beta} = (X'W X)^{-1} X' W Y$$

An unbiased estimate of σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{n-p} Y^* [I - X^* (X'^* X)^{-1} X^*] Y^*$$

$$= \frac{1}{n-p} Y' [V^{-1} - V^{-1} X' (X' V^{-1} X)^{-1} X V^{-1}] Y$$

ANOVA Decomposition

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

$$I_n = \bar{J}_n + [X(X'X)^{-1}X' - \bar{J}_n] \\ + [\bar{J}_n - X(X'X)^{-1}X']$$

$$= A_1 + A_2 + A_3$$

$$\text{If } X = [1, X_c] \text{ where } X_c = [c_{x_1}, \dots, c_{x_{p-1}}]$$

then

$$X(X'X)^{-1}X' = \bar{J}_n + X_c (X_c' X_c)^{-1} X_c'$$

$$\text{so } A_2 = X_c (X_c' X_c)^{-1} X_c'$$