2.5 Parameterizations of a linear model

DATA: $(Y_i, Z_i^T)^T$, $i = 1, \dots, n$
(that we primarily observe)

remember: $X_i = t(Z_i)$ are regressors used in the model that allow to assume
$$E(Y_i \mid Z_i = z) = X_i^T \beta, \quad X_i = t(z_i)$$
for some β

- one set $z_1, \dots, z_n$ of observed covariates
  $\rightarrow$ ∞ many different sets of regressors $X_1, \dots, X_n$ can be proposed (based on different transformations t)
  $\rightarrow$ not necessarily "different" models

Def. 2.7 Equivalent linear models

Assume two linear models
$$M_1: \; Y \mid X_1 \sim (X_1 \beta, \sigma^2 I_n), \; X_1 \text{ matrix } n \times k, \; \text{rank}(X_1) = r,$$
$$M_2: \; Y \mid X_2 \sim (X_2 \beta, \sigma^2 I_n), \; X_2 \text{ matrix } n \times l, \; \text{rank}(X_2) = r.$$

We say that models M_1 and M_2 are equivalent if their regression spaces are the same. That is, if $\mathcal{M}(X_1) = \mathcal{M}(X_2)$.

---

equivalent models

$\Rightarrow$ • $H = X_1 (X_1^T X_1)^- X_1 = X_2 (X_2^T X_2)^- X_2$

• same $\hat{Y}$, U, SSe, MSe, ...

- two (or more) different parameterizations of the same situation

- different interpretation of regression coefficients

Example 1 Change of units of regressor

Y = consumption of a car [l/100km]

Z = weight of a car [kg] $\rightarrow Z_1, \dots, Z_n$

$X_1 = Z$ $\quad M_1: \; E(Y|Z=z) = \beta_0 + \beta_1 z$

$$X_1 = \begin{pmatrix} 1 & z_1 \\ \vdots & \vdots \\ 1 & z_n \end{pmatrix}$$

β_1 = change in $E(Y|\dots)$ related to 1kg change in a weight

---

$X_2 = \frac{Z}{1000}$ (= weight in tons)

$M_2: \; E(Y|Z=z) = \gamma_0 + \gamma_1 \left(\frac{z}{1000}\right)$

$$X_2 = \begin{pmatrix} 1 & \frac{z_1}{1000} \\ \vdots & \vdots \\ 1 & \frac{z_n}{1000} \end{pmatrix}$$

γ_1 = change in $E(Y|\dots)$ related to 1 ton change in a weight

Example 2 Two sample problem

Y = height [cm]

Z = gender $\in \{0,1\} \to z_1, \dots, z_n$

(WLG): upon reordering of data

$$\begin{pmatrix} z_1 \\ \vdots \\ z_{n_0} \\ \hline z_{n_0+1} \\ \vdots \\ z_{n_0+n_1} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \hline 1 \\ \vdots \\ 1 \end{pmatrix}$$

- we need to "model" $\mathbb{E}(Y \mid Z = z)$, $z \in \{0,1\}$

$\to$ $\mu_0 := \mathbb{E}(Y \mid Z = 0)$, $\mu_1 := \mathbb{E}(Y \mid Z = 1)$

- regression space

$$= \left\{ \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_0 \\ \hline \mu_1 \\ \vdots \\ \mu_1 \end{pmatrix} : \mu_0, \mu_1 \in \mathbb{R} \right\}$$

$\to$ any model matrix $\mathbb{X}$, whose $\mathcal{M}(\mathbb{X})$ is the above space must be of the form

$$\mathbb{X} = \begin{pmatrix} x_0^T \\ \vdots \\ x_0^T \\ \hline x_1^T \\ \vdots \\ x_1^T \end{pmatrix} \qquad \begin{array}{l} x_0 \neq 0 \\ x_1 \neq 0 \\ x_0 \neq x_1 \end{array}$$

with $\operatorname{rank}(\mathbb{X}) = 2$

$\to$ if used in a linear model $\mu_0 = x_0^T \beta$

$\mu_1 = x_1^T \beta$

POSSIBLE CHOICES OF x_0, x_1

(i) Group means, $\beta = (\beta_0, \beta_1)^T$

$$x_0 = (1, 0)^T \quad \rightarrow \quad \mu_0 = x_0^T \beta = \beta_0$$
$$x_1 = (0, 1)^T \qquad\qquad \mu_1 = x_1^T \beta = \beta_1$$

(ii) Group differences, $\beta = (\beta_0, \beta_1)^T$

$$x_0 = (1, 0)^T \quad \rightarrow \quad \mu_0 = x_0^T \beta = \beta_0$$
$$x_1 = (1, 1)^T \qquad\qquad \mu_1 = x_1^T \beta = \beta_0 + \beta_1$$

$$\beta_0 = \mu_0$$
$$\beta_1 = \mu_1 - \mu_0$$

(iii) Deviations from the means of the means, $\beta = (\beta_0, \beta_1)^T$

$$x_0 = (1, 1)^T \quad \rightarrow \quad \mu_0 = x_0^T \beta = \beta_0 + \beta_1$$
$$x_1 = (1, -1)^T \qquad\qquad \mu_1 = x_1^T \beta = \beta_0 - \beta_1$$

$$\beta_0 = \frac{\mu_0 + \mu_1}{2}$$

$$\beta_1 = \mu_0 - \frac{\mu_0 + \mu_1}{2}$$

$$= \frac{\mu_0 + \mu_1}{2} - \mu_1$$

(iv) Orthonormal basis, $\beta = (\beta_0, \beta_1)^T$

$$x_0 = \left(\frac{1}{\sqrt{n_0}}, 0\right)$$
$$x_1 = \left(0, \frac{1}{\sqrt{n_1}}\right)$$

$$X = \begin{pmatrix} \frac{1}{\sqrt{n_0}} & 0 \\ \vdots & \vdots \\ \frac{1}{\sqrt{n_0}} & 0 \\ 0 & \frac{1}{\sqrt{n_1}} \\ \vdots & \vdots \\ 0 & \frac{1}{\sqrt{n_1}} \end{pmatrix} \begin{array}{l} \\ n_0 \text{ rows} \\ \\ \\ n_1 \text{ rows} \\ \\ \end{array}$$

$(= Q)$

$$\rightarrow \mu_0 = x_0^T \beta = \frac{1}{\sqrt{n_0}} \beta_0 \qquad \beta_0 = \sqrt{n_0} \cdot \mu_0$$

$$\mu_1 = x_1^T \beta = \frac{1}{\sqrt{n_1}} \beta_1 \qquad \beta_1 = \sqrt{n_1} \cdot \mu_1$$

Easy homework:

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_{n_0} \\ \hline Y_{n_0+1} \\ \vdots \\ Y_{n_0+n_1} \end{pmatrix} \qquad \overline{Y}_0 := \frac{1}{n_0} \sum_{i=1}^{n_0} Y_i$$

$$\overline{Y}_1 := \frac{1}{n_1} \sum_{i=n_0+1}^{n_0+n_1} Y_i$$

$$\hat{Y} = \begin{pmatrix} \overline{Y}_0 \\ \vdots \\ \overline{Y}_0 \\ \hline \overline{Y}_1 \\ \vdots \\ \overline{Y}_1 \end{pmatrix} \qquad SS_e = \sum_{i=1}^{n_0} (Y_i - \overline{Y}_0)^2 + \sum_{i=n_0+1}^{n_0+n_1} (Y_i - \overline{Y}_1)^2$$

$$\nu_e = n - 2$$

2.6 Matrix algebra and a method of least squares

<u>Linear model</u> $\quad Y|X \sim (X\beta, \sigma^2 I_n)$

$\text{rank}(X_{n\times k}) = k$

<u>Quantities to calculate for the LSE:</u>

$H = X(X^TX)^{-1}X^T \quad , \quad M = I_n - H$

$\hat{Y} = HY = X(X^TX)^{-1}X^T ,$

$\text{var}(\hat{Y}|X) = \sigma^2 H = \sigma^2 \{X(X^TX)^{-1}X^T$

$U = MY = Y - \hat{Y}$

$\text{var}(U|X) = \sigma^2 M = \sigma^2 (I_n - H)$

$\hat{\beta} = (X^TX)^{-1}X^TY \quad , \quad \text{var}(\hat{\beta}|X) = \sigma^2 (X^TX)^{-1}$

$\rightarrow$ QR decomposition

$\rightarrow$ SVD decomposition

of X

<u>QR & LSE</u>

$\text{rank}(X_{n\times k}) = k \quad \rightarrow \quad X = Q \cdot R$

$Q_{n\times k} = (q_1, \dots, q_k)$

$=$ orthonormal basis of $\mathcal{M}(X)$

$R =$ upper Δ

$Q^TQ = I_k \quad , \quad \underline{Q \cdot Q^T = H}$

- $X^T X = R^T \underbrace{Q^T Q}_{I} R \;= R^T R$

  = Cholesky (square root) decomposition of $X^T X$ },
  special case of LU decomp.

- $(X^T X)^{-1} = (R^T R)^{-1} = \underline{\underline{R^{-1}(R^{-1})^T}}$
  easy

- $\det(X^T X) = \det(R^T R) = (\det R)^2 = \left(\prod_{j=0}^{k-1} r_{jj}\right)^2$

- $\det((X^T X)^{-1}) = (\det(X^T X))^{-1}$

- solution to normal eq.

$$X^T X \beta = X^T Y$$
$$R^T R \beta = R^T Q^T Y$$
$$R \beta = \underbrace{Q^T Y}_{RHS}$$

Δ

↳ easy to solve

- meaning of $c = Q^T Y = (c_1, \dots, c_k)^T$ (effects in R)

$$\hat{Y} = HY = Q \cdot \underbrace{Q^T Y}_{c} = \sum_{j=1}^{k} \underline{\underline{c_j}}\, q_j$$

↳ coefs. of linear combination for basis of $\mathcal{M}(X)$ to get $\hat{Y}$