P.39. Finding the least-square estimator

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

$$e'e \Rightarrow (Y - X\beta)'(Y - X\beta) = f(\underset{\sim}{\beta})$$

$$f(\underset{\sim}{\beta}) = Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta$$

$$= \underbrace{Y'Y}_{①} - \underbrace{\beta'(X'X)(X'X)^{-1}X'Y}_{②} - \underbrace{Y'X(X'X)(X'X)^{-1'}\beta}_{③}$$

$$+ \underbrace{Y'X(X'X)^{-1}(X'X)(X'X)^{-1}X'Y}_{②} - \underbrace{Y'X(X'X)^{-1}(X'X)(X'X)^{-1}X'Y}_{①}$$

$$+ \underbrace{\beta'X'X\beta}_{③}$$

$$f(\underset{\sim}{\beta}) = \underset{\sim}{Y}'(I - X(X'X)^{-1}X')\underset{\sim}{Y} + (\underset{\sim}{\beta} - (X'X)^{-1}X'\underset{\sim}{Y})'\underset{\sim}{(X'X)}(\underset{\sim}{\beta} - (X'X)^{-1}X'\underset{\sim}{Y})$$

Now,

$$\frac{\partial f}{\partial \underset{\sim}{\beta}} = \frac{\partial}{\partial \underset{\sim}{\beta}}(\beta - (X'X)^{-1}(X'Y))'(X'X)(\beta - (X'X)^{-1'}X'Y)$$

$$= \frac{\partial}{\partial \underset{\sim}{\beta}}\left[(\beta'(X'X) - (X'Y)'\cancel{(X'X)^{-1}}\cancel{(X'X)})(\beta - (X'X)^{-1}X'Y)\right] =$$

$$= \frac{\partial}{\partial \underset{\sim}{\beta}}\left[\beta'X'X\beta - (Y'X)\beta - \beta'\cancel{(X'X)}\cancel{(X'X)^{-1}}X'Y + (Y'X)(X'X)^{-1}X'Y\right]$$

$$= \frac{\partial}{\partial \underset{\sim}{\beta}}\left[\beta'X'X\beta - Y'X\beta - \beta'(X'Y)\right] = \frac{\partial}{\partial \beta}\left[\beta'X'X\beta - Y'X\beta - (Y'X\beta)^T\right]$$

Using the Facts: $Y = AX \Rightarrow \frac{\partial}{\partial X}Y = A$ and $\alpha = Y^TAX \Rightarrow \frac{\partial \alpha}{\partial X} = Y^TA$

$$\frac{\partial \alpha}{\partial Y} = X^TA^T$$

$$\frac{\partial f}{\partial X} = \frac{\partial f}{\partial X^T}$$

$$\alpha = X^TAX \Rightarrow \frac{\partial \alpha}{\partial X} = X^T(A + A^T) \underset{\text{if A is symmetric}}{=} 2X^TA.$$

Proof: by the quadratic form definition:

$$\alpha = X^TAX = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_ix_j$$

differentiating with respect to the k-th element of $\underset{\sim}{X}$ we have:

$$\frac{\partial \alpha}{\partial x_k} = \sum_{j=1}^{n} a_{kj}x_j + \sum_{i=1}^{n} a_{ik}x_i, \quad \text{for all } k = 1 \ldots n.$$

$$\Rightarrow \frac{\partial \alpha}{\partial X} = X^TA^T + X^TA = X^T(A^T + A)$$

Hence we have:

$$\circledast 1 = \frac{\partial}{\partial \beta}\left[\beta'(x'x)\beta - y'x\beta - \beta'(x'y) \right] = 2\beta'(x'x) - y'x - y'x$$

$$= 2\beta'(x'x) - 2y'x$$

Now,

$$\frac{\partial}{\partial \beta} f(\beta) = 0 \Rightarrow 2\beta'(x'x) = +2y'x$$

$$(x'x)'\beta = x'y$$

$$\cancel{(x'x)^{-1}(x'x)}\beta = (x'x)^{-1}(x'y)$$

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

Calculo dos estimadores de min. Quadrados

Para Modelo de regressão linear na forma matricial:

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \dots + \alpha_k x_{ki} + \varepsilon_i$$

$$y = X\beta + \underset{\sim}{\varepsilon} \quad ; \quad \beta = (\alpha_0\ \alpha_1 \dots \alpha_k)$$

$\underset{\sim}{e} = (y - X\beta)$ é o vetor de erros.

queremos minimizar $e'e = (y - X\beta)'(y - X\beta)$

$$= y'y - \beta'X'y - y'X\beta + \beta'X'X\beta$$

$$= y'y - \beta'(X'X)(X'X)^{-1} \cdot X y - \underline{y'X(X'X)^{-1}(X'X)}\beta +$$

$$- y'X(X'X)^{-1}(X'X)(X'X)^{-1}X'y + \underline{y'X(X'X)^{-1}(X'X)}(X'X)^{-1}X'y$$

$$+ \beta'X'X\beta$$

$$= y'(I - X(X'X)^{-1}(X'X)(X'X)^{-1}X')y$$

$$+ \underline{y'X(X'X)^{-1}(X'X)}\left((X'X)^{-1}X'y - \beta\right)$$

$$+ \beta'(X'X)\left(\beta - (X'X)^{-1}X'y\right) \quad , \text{ mas } (X'X)^{-1}X'y - \beta = -(\beta - (X'X)^{-1}X'y),$$

$$= y'(I - X(X'X)^{-1}(X'X)(X'X)^{-1}X')y$$

$$+ \left(-y'X(X'X)^{-1}(X'X) + \beta'(X'X)\right)\left(\beta - (X'X)^{-1}X'y\right)$$

$$= y'(I - X(X'X)^{-1}(X'X)(X'X)^{-1}X')y$$

$$+ \left(-y'X(X'X)^{-1} + \beta'\right)(X'X)\left(\beta - (X'X)^{-1}X'y\right)$$

$$= y'(I - X(X'X)^{-1}(X'X)(X'X)^{-1}X')y$$

$$+ \underbrace{\left(\beta' - y'X(X'X)^{-1}\right)}_{(*1)}(X'X)\left(\beta - (X'X)^{-1}X'y\right) = (*2)$$

Note que: $(A^T)^{-1} = (A^{-1})^T$. Note também que $\underbrace{(X^TX)^T = (X^T)(X^T)^T = X^TX}_{\text{simétrica}}$.

Logo, de (*1) teremos: $\beta^T - y^TX(X^TX)^{-1}$

$$(\beta^T - y^TX(X'X)^{-1})^T = ((\beta^T)^T - (y^TX(X'X)^{-1})^T)^T = (\beta - ((X'X)^{-1})^T(y^TX)^T)^T =$$

$$= (\beta - (X^TX)^{-1}(X^Ty))^T = (\beta - (X^TX)^{-1}X^Ty)^T$$

Assim, (*2) fica:

$$= y'[I - (X(X'X)^{-1}(X'X)(X'X)^{-1}X']y +$$
$$(\beta - (X'X)^{-1}X'y)'(X'X)(\beta - (X'X)X'y)$$

Se derivarmos esta expressão em relação a β, relembrando que, se A é simétrica então $\frac{d}{d\underline{x}} x'AX = 2A\underline{x}$

então:

$$\frac{d}{d\beta}(\beta - (X'X)^{-1}X'y)'(X'X)(\beta - (X'X)^{-1}X'y) \overset{\text{regra da cadeia}}{=}$$

$$= 2(X'X)(\beta - (X'X)^{-1}X'y) \cdot \underbrace{\left(\frac{d}{d\beta}(\beta - (X'X)^{-1}X'y\right)}_{I}$$

agora, igualando a zero, teremos:

$$2(X'X)(\hat{\beta} - (X'X)^{-1}X'y) = 0$$

$$\cancel{2}(X'X)\hat{\beta} - \cancel{2}X'y = 0$$

$$(X'X)\hat{\beta} = X'y$$

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

Note que $\frac{d^2}{d\beta^2} e'e = 2X'X \geq 0$, pois $X'X$ é não negativa definida.

$$v'(X'X)v = (v'X')(Xv) =$$
$$= \langle v'X, Xv \rangle = \|vX\| \geq 0$$