$$\min_{\beta} \|y - x\beta\|^2$$

i) $\nabla_\beta\, y^T\beta = y$

ii) $\nabla^2_\beta\, y^T\beta = 0$

$\boxed{iii)}$ $\nabla_\beta\, \beta^T x \beta = \underbrace{2x\beta}$ (si x es simétrica)

iv) $\nabla^2_\beta\, \beta^T x\beta = 2x$ "

$$\|y - x\beta\|^2$$
$$= \left(\sqrt{(y-x\beta)^T(y-x\beta)}\right)^2$$
$$= (y-x\beta)^T(y-x\beta)$$
$$= (y^T - (x\beta)^T)(y - x\beta)$$
$$= (y^T - \beta^T x^T)(y - x\beta)$$
$$= \underbrace{\overset{0}{\cancel{y^T y}}} - y^T x\underbrace{\beta} - \underbrace{\beta^T x^T y} + \beta^T x^T x \beta = L(\beta)$$

gram matrix

$\frac{d}{d\beta}\, \beta^T z = z^T$ → agregar

i) simétrica
ii) $\det(x^T x) = 0$
si las columnas son L.I

$$\boxed{\frac{\partial L(\beta)}{\partial \beta} = -y^T x - y^T x + 2x^T x\beta}$$

$$= -\cancel{2}y^T x + \cancel{2}x^T x\beta = 0$$

$n\,(\quad |_{p>>1}$

Condiciones de opt.

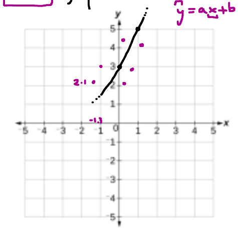
$$x^T x\,\beta = y^T x$$
$$\rightarrow (x^T x)^{-1}(x^T x)\beta = (x^T x)^{-1} y^T x$$
$$\boxed{\beta^* = \boxed{(x^T x)^{-1}} x^T y}$$

$\underbrace{x^T x} = \mathbb{I}$

$\Rightarrow \boxed{x^T y}$

$y = a\underbrace{x} + b$

β



$y = 2x + 3$
2 → pendiente
3 → punto d. corte
0

$\| \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$$\min_{\beta} \sum_{i=1}^{n} (y_i - x_i\beta)^2 = (y_1 - \beta_0 - \beta_1 x_1)^2 + (y_2 - \beta_0 - \beta x_2)^2 + \vdots + (y_n - \beta_0 - \beta_1 x_n)^2$$

$$X = \| \begin{bmatrix} [1 & - & x_1] \\ 1 & & x_2 \\ \vdots & & \vdots \\ 1 & & x_n \end{bmatrix}_{n\cdot 1} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}_{2\cdot 1} \|^2$$

$$\|x\|^2 = x^T x = \sum_i x_i^2$$

$$(\ )^{O}\ OO \rightarrow \beta_0 \left\| \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} \beta_0 + x_1\beta_1 \\ \beta_0 + x_2\beta_1 \\ \vdots \\ \beta_0 + x_n\beta_1 \end{pmatrix} \right\|^2 = \begin{pmatrix} y_1 - \beta_0 - x_1\beta_1 \\ \vdots \end{pmatrix}^T \begin{pmatrix} y_1 - \beta_0 - x_1\beta_1 \\ \end{pmatrix} = (y_1 - \beta_0 - \beta_1 x_1)^2 + \cdots$$

β_1

$$\hat{\alpha} = \bar{y} - (\hat{\beta}\,\bar{x}),$$
$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$
$$= \frac{s_{x,y}}{s_x^2}$$
$$= r_{xy}\frac{s_y}{s_x}.$$

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

$$L(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

La primera prueba es que $\beta_0 = \bar{y} - \bar{x}\beta_1$

$$\frac{\partial L(\beta_0,\beta_1)}{d\beta_0} = \frac{dL}{d\beta_0}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)$$
$$\overset{(*)}{=} \sum_{i=1}^{n} \frac{d}{d\beta_0}(\overset{0}{\cancel{y_i}} - \beta_0 - \overset{0}{\cancel{\beta_1 x_i}})^2$$

(*) smooth

$$= \sum_{i=1}^{n} \underbrace{2}(y_i - \beta_0 - \beta_1 x_i)(-1)$$
$$= -2\sum_{i=1}^{n} y_i - \beta_0 - \beta_1 x_i$$
$$= -2\Big(\underbrace{\sum_{i=1}^{n} y_i}_{Ⓐ} - \underbrace{\sum_{i=1}^{n} \beta_0}_{Ⓑ} - \sum_{i=1}^{n} \beta_1 x_i\Big)$$
$$= -2(n\cdot\bar{y} - n\cdot\beta_0 - n\cdot\beta_1\cdot\bar{x})$$
$$= \cancel{-2}(\cancel{n}\bar{y} - \cancel{n}\beta_0 - \cancel{n}\cdot\beta_1\bar{x})$$
$$\Rightarrow 0 = \bar{y} - \boxed{\beta_0} - \beta_1\bar{x}$$

→ El objetivo es despejar β_0

$$\Rightarrow \boxed{\beta_0 = \bar{y} - \beta_1\bar{x}}$$

Atentos con esta expresión: β_0 está expresado en función de β_1. Cuando esto sucede lo más probable sustituimos el valor en la prueba de β_1

Ⓐ $\frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y} \Rightarrow \bar{y}\cdot n = \sum_{i=1}^{n} y_i$

Ⓑ $\sum_{i=1}^{n} a = a\sum_{i=1}^{n} 1 = n\cdot a$

Segunda parte: Demostramos ahora para β_1

$$\frac{dL}{d\beta_1}\sum_{i=1}^{n}(y_i - \beta_0 \underbrace{- \beta_1 x_i})^2$$

→ Este es el término importante ahora

$$= 2(y_i - \beta_0 - \beta_1 x_i)\cdot(-x_i)$$
$$= 2(y_i - (\bar{y} - \bar{x}\beta_1) - \beta_1 x_i)(-x_i)$$
$$= -2(y_i - \bar{y} + \bar{x}\beta_1 - \beta_1 x_i)x_i$$

$$-\sum_{i=1}^{n} y_i x_i + \bar{y}\sum_{i=1}^{n} x_i + \beta_1\sum_{i=1}^{n} x_i^2 \,|\, \bar{x}\beta_1\sum_{i=1}^{n} x_i$$

$$= -\sum_{i=1}^{n} y_i x_i + \bar{y}\cdot n\bar{x} + \beta_1\Big(\sum_{i=1}^{n} x_i^2 - \bar{x}\cdot n\cdot\bar{x}\Big)$$

Como en el caso anterior buscamos despejar para β_1

Es fácil probar usando Ⓐ y Ⓑ que la expresión es equivalente a

$$\beta_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{y}\bar{x}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2(y_i - \bar{y})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Demostración Ⓐ

Cuando demostramos β_1 necesitamos desarrollar dos estadísticos importantes.

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$
$$= \sum_{i=1}^{n} x_i y_i \boxed{- \sum_{i=1}^{n} x_i\bar{y} - \sum_{i=1}^{n}\bar{x} y_i} + \bar{x}\bar{y}n \quad ✓$$
$$= \sum_{i=1}^{n} x_i y_i - n\sum_{i=1}^{n} x_i\sum_{i=1}^{n} y_i - n\cancel{\sum_{i=1}^{n} x_i}\sum_{i=1}^{n} y_i + n\cancel{\sum_{i=1}^{n} x_i}\sum_{i=1}^{n} y_i$$
$$= \sum_{i=1}^{n} y_i x_i - n\sum_{i=1}^{n} x_i\sum_{i=1}^{n} y_i$$

Con esto demuestro $\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} y_i x_i - n\sum_{i=1}^{n} x_i\sum_{i=1}^{n} y_i$

Demostración Ⓑ

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - 2x_i\bar{x} + \bar{x}^2$$
$$= \sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\bar{x}^2$$
$$= \sum_{i=1}^{n} x_i^2 - 2\bar{x}.n.\bar{x} + n\bar{x}^2$$
$$= \sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2$$
$$= \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 = \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i x_i$$

En realidad cualquiera de las versiones está ok! cada quien lo escribe como parezca apropiado