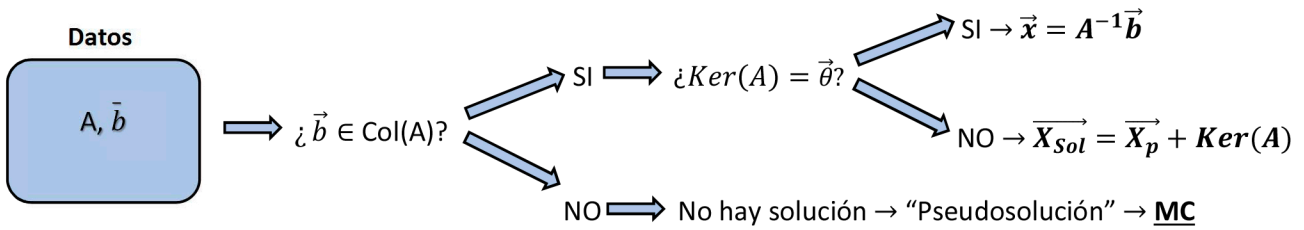


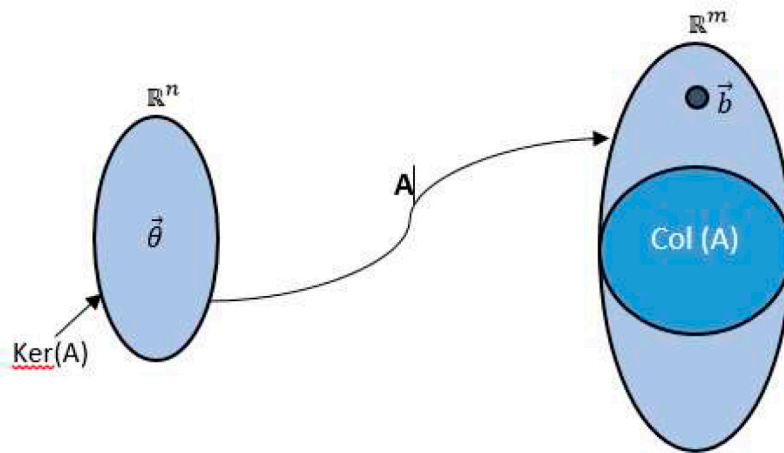
Por tanto, como esquema:



TIPO 1 → PROBLEMA “PURAMENTE” SOBREDETERMINADO

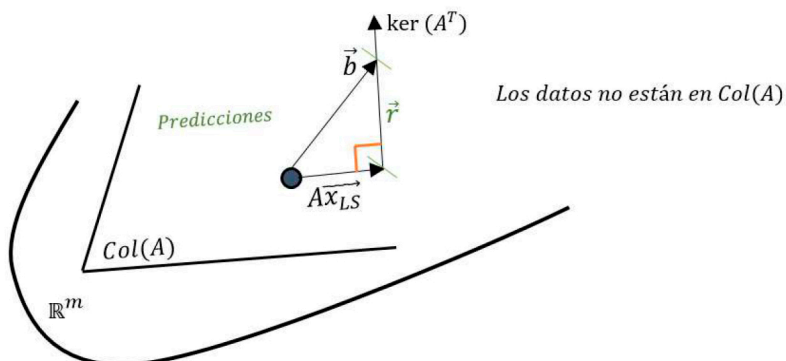
$m \gg n$ (más ecuaciones que incógnitas)

$\text{rang}(A) = n$



$$\dim(\text{Col}(A)) = n \neq m$$

De forma general $\vec{b}$ no se encuentra en $\text{Col}(A)$, por lo que este tipo de problemas **no tienen incertidumbre**, pero **tampoco solución** (No se puede hallar $A\vec{x} = \vec{b}$). Si se diese el caso de que $\vec{b}$ se encontrase en $\text{Col}(A)$, el problema tendría una solución única.



$\vec{r}$ (residuo) es la proyección ortogonal de $\vec{b}$ sobre $\text{Col}(A)$. La “pseudosolución” buscada es $\overrightarrow{x_{LS}}$.

La idea es hallar $\overrightarrow{x_{LS}}$, y predecir con ella $A\overrightarrow{x_{LS}}$, cometiendo un error de $\vec{r} = \vec{b} - A\overrightarrow{x_{LS}}$ (residuo).

Si $A\vec{x} = \vec{\theta} \leftrightarrow x \in \text{Ker}(A)$

Así, $\vec{x}$ es ortogonal a $\text{Col}(A^T)$, por lo que $\vec{x} \in [\text{Col}(A^T)]^\perp$:

$$A\vec{x} = \begin{pmatrix} \overrightarrow{f_1} \\ \overrightarrow{f_2} \\ \vdots \\ \overrightarrow{f_m} \end{pmatrix} \vec{x} = \begin{pmatrix} \overrightarrow{f_1}\vec{x} \\ \overrightarrow{f_2}\vec{x} \\ \vdots \\ \overrightarrow{f_m}\vec{x} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\vec{x}$ es ortogonal al espacio $\langle \overrightarrow{f_1}, \overrightarrow{f_2}, \dots, \overrightarrow{f_m} \rangle \equiv \text{Col}(A)$

$$(\text{Col}(A^T))^\perp = \text{Ker}(A) \qquad (\text{Col}(A))^\perp = \text{Ker}(A^T)$$

$$\vec{r} = \vec{b} - A\overrightarrow{x_{LS}} \in \text{Ker}(A^T) \qquad A^T(\vec{b} - A\overrightarrow{x_{LS}}) = \vec{\theta} \rightarrow \underbrace{A^T\vec{b} = A^T A\overrightarrow{x_{LS}}}_{\text{Sistema que da la solución por MC}}$$

Sistema que da la solución por MC

$$\vec{r} \perp \text{Col}(A)$$

$$\vec{b} - A\vec{x}_{LS} \perp \text{Col}(A)$$

Un vector es ortogonal a un subespacio vectorial si es ortogonal a cualquier base del subespacio:

$$\left. \begin{aligned} (\vec{b} - A\vec{x}_{LS})\vec{c}_1 &= 0 \\ (\vec{b} - A\vec{x}_{LS})\vec{c}_2 &= 0 \\ (\vec{b} - A\vec{x}_{LS})\vec{c}_n &= 0 \end{aligned} \right\} A^T A\vec{x}_{LS} = A^T \vec{b}$$

Cogiendo una ecuación cualquiera:

$$\left. \begin{aligned} \vec{b}\vec{c}_1 &= A\vec{x}_{LS}\vec{c}_1 \\ \vec{b}\vec{c}_2 &= A\vec{x}_{LS}\vec{c}_2 \\ \vec{b}\vec{c}_n &= A\vec{x}_{LS}\vec{c}_n \end{aligned} \right\} \begin{pmatrix} \vec{c}_1^T \\ \vec{c}_2^T \\ \vdots \\ \vec{c}_n^T \end{pmatrix} \vec{b} = A^T \vec{b}$$

$A^T A$ es una matriz de correlación de columnas:

$$\begin{pmatrix} \vec{c}_1\vec{c}_1 & \vec{c}_1\vec{c}_2 & \cdots & \vec{c}_1\vec{c}_n \\ \vec{c}_2\vec{c}_1 & \vec{c}_2\vec{c}_2 & \cdots & \vec{c}_2\vec{c}_n \\ \vdots & \vdots & \ddots & \vdots \\ \vec{c}_n\vec{c}_1 & \vec{c}_n\vec{c}_2 & \cdots & \vec{c}_n\vec{c}_n \end{pmatrix}$$

Propiedades de la matriz $A^T A$:

1. $A^T A \in \mathbf{M}_{m \times n}(\mathbb{R})$
2. Es simétrica.
3. $\text{rang}(A^T A) = \text{rang}(A)$
4. $\exists (A^T A)^{-1} \leftrightarrow \text{rang}(A) = n$
5. Es definida positiva (se puede usar **Cholesky o QR**):

Para cualquier $\vec{x} \in \mathbb{R}^n$ y $\vec{x} \neq 0$: $\vec{x}^T (A^T A) \vec{x} > 0$

QR:

$$\begin{aligned} A^T A &= QR \text{ (Q: ortogonal; R: triangular)} \\ QR\vec{x}_{LS} &= \vec{c} \rightarrow R\vec{x}_{LS} = Q^T \vec{c} = \vec{d} \text{ (se emplea sustitución regresiva)} \\ Q^T &= Q^{-1} \end{aligned}$$

Lo malo es que el condicionamiento de $A^T A = \text{cond}^2(A) \rightarrow$ mal condicionada.

REGRESIÓN LINEAL

Datos:

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$$

Recta de regresión: mejor recta de MC $\rightarrow y^*(x) = ax + b$

$$y_1 \approx y^*(x_1) = ax_1 + b \times 1$$

$$y_2 \approx y^*(x_2) = ax_2 + b \times 1$$

$$y_m \approx y^*(x_m) = ax_m + b \times 1$$

Expresándolo de forma matricial:

$$\overrightarrow{y_{obs}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m \approx a \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Se pretende que $\overrightarrow{y_{obs}}$ sea lo más próximo a:

$\langle \vec{x}, \vec{1} \rangle$, subespacio de dimensión 2, en $\mathbb{R}^m$:

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} = A = [\vec{x}, \vec{1}]$$

$$A^T A = \begin{pmatrix} \vec{x}\vec{x} & \vec{x}\vec{1} \\ \vec{1}\vec{x} & \vec{1}\vec{1} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m (\vec{x}_k)^2 & \sum \vec{x}_k \\ \sum \vec{x}_k & m \end{pmatrix}$$

$$A^T \overrightarrow{y_{obs}} = [\vec{x}, \vec{1}]^T \overrightarrow{y_{obs}} = \begin{bmatrix} \sum \vec{x}_k \vec{y}_k \\ \sum \vec{y}_k \end{bmatrix}$$

$$y^*(x) = ax + b$$

$$\begin{pmatrix} \sum_{k=1}^m (\vec{x}_k)^2 & \sum \vec{x}_k \\ \sum \vec{x}_k & m \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{bmatrix} \sum \vec{x}_k \vec{y}_k \\ \sum \vec{y}_k \end{bmatrix}$$

$$\left(\sum \vec{x}_k \right) a + mb = \sum y_k$$

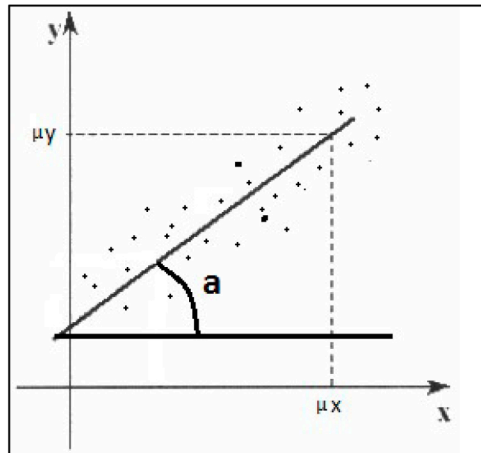
$$b = \mu y - a \mu x$$

μ es la media aritmética de los y $\left(\frac{\sum \vec{y}_k}{m}\right)$

Sustituyendo:

$$y^*(x) - \mu y = a(x - \mu x)$$

La recta de regresión pasa por el centro de gravedad de la nube:



a depende de la correlación (producto escalar) entre $\vec{y}$ y $\vec{x}$ (medible).

$$\left(\sum x_k^2\right)a + \left(\sum x_k\right)b = \sum x_k y_k$$

Dividiendo entre m :

$$a\mu x^2 + b\mu x = \mu xy$$

$$a\mu x^2 + (\mu y - a\mu x)\mu x = \mu xy$$

$$a = \frac{\mu xy - \mu x \mu y}{\mu x^2 - \mu^2 x}$$

Donde:

$(\mu xy - \mu x \mu y)$ es la covarianza centrada, $C_{xy} = C(x, y) = E[(x - \mu x)(y - \mu y)] \rightarrow$ se lleva al c.d.g.

Donde:

E es el operador "Esperanza"; $(x - \mu x) = x_c$ $(y - \mu y) = y_c$

$$(\mu x^2 - \mu^2 x) = C(x, x) = \mu_{xx} - \mu_x \mu_x$$

$C(x, x)$ es la Varianza de x , $\text{Var}(x)$; $\sqrt{\text{Var}(x)} = \sigma_x =$ desviación típica.

Con todo ello, la pendiente, a , será:

$$a = \frac{C(x, y)}{\text{Var}(x)} \quad y - \mu y = \frac{C(x, y)}{\text{Var}(x)} \times (x - \mu x)$$

$\vec{x}$	$\vec{y}$	$\vec{x}\vec{y}$	No centrado
$(\vec{x} - \mu\vec{x})$	$(\vec{y} - \mu\vec{y})$	$\vec{x}_c\vec{y}_c$	Centrado (depende de las dimensiones)
$\frac{(\vec{x} - \mu\vec{x})}{\sigma_x}$	$\frac{(\vec{y} - \mu\vec{y})}{\sigma_y}$	$\vec{x}_s\vec{y}_s$	Estandarización

$E(\vec{x}_s\vec{y}_s) = \varphi \rightarrow$ Pearson linear regression coefficient:

$$\varphi = \frac{C(x, y)}{\sigma_x \sigma_y}$$

Se cumple que $-1 \leq \varphi \leq 1$

$$y - \mu y = \frac{\varphi \sigma_y}{\sigma_x} (x - \mu x)$$

Donde:

$\frac{\varphi \sigma_y}{\sigma_x}$ es la pendiente, que depende de la correlación.

El mejor estimador de una variable cuando no hay correlación es la media:

$$y^*(x) - \mu y = 0$$

Si la pendiente es positiva ($a > 0$) $\rightarrow \varphi > 0$

Si la pendiente es negativa ($a < 0$) $\rightarrow \varphi < 0$

REGRESIÓN POTENCIAL/EXPONENCIAL

$y = ax + b \rightarrow$ relación afín entre x e y .

$y = Cx^\alpha$ donde C, α son dos constantes desconocidas.

Como se observa, se trata de una regresión no lineal

$$\left. \begin{aligned} y_1 &= Cx_1^\alpha \\ y_2 &= Cx_2^\alpha \\ y_m &= Cx_m^\alpha \end{aligned} \right\} \log(y) = \log(Cx^\alpha) = \log(C) + \alpha \log(x) \rightarrow \text{papel doble logarítmico.}$$

$$\begin{pmatrix} \log(y_1) \\ \log(y_2) \\ \vdots \\ \log(y_m) \end{pmatrix} = \alpha \begin{pmatrix} \log(x_1) \\ \log(x_2) \\ \vdots \\ \log(x_m) \end{pmatrix} + \log(C) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad \langle \log(\vec{x}), \vec{1} \rangle$$

$$\text{Proy} \perp \log(\vec{y}_{obs}) \rightarrow (\alpha, \log(C) \rightarrow C)$$

Se resuelve de forma análoga a la regresión lineal.

REGRESIÓN EXPONENCIAL

$$y(x) = Ce^{kx}$$

$$\log(y(x)) = \log(C) + kx \quad \langle \vec{1}, \vec{x} \rangle \rightarrow \text{papel semilogarítmico.}$$

$$\text{Proy} \rightarrow \log(\overrightarrow{y_{obs}}) \rightarrow (k, \log(C) \rightarrow C)$$

Se resuelve de forma análoga a la regresión lineal.

Ejercicio propuesto:

$$\frac{dy}{dx} = ky(B - y) \quad y(x = 0) = y_0 \quad \text{Conocido } k, B^T \in \mathbb{R}$$

Lo que se obtiene de esta relación diferencial es una curva sigmoide:

$$y(x) = S(x, k, B)$$

```

>> syms k
syms B
>> syms y
>> f=1/(k*y*B-k*(y^2))

f =

-1/(k*y^2 - B*k*y)

>> int(f,y)

ans =

(2*atanh((2*y)/B - 1))/(B*k)
  
```

$$x = \frac{2 \operatorname{arctanh}\left(\frac{2y}{B-1}\right)}{Bk} + C$$