UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

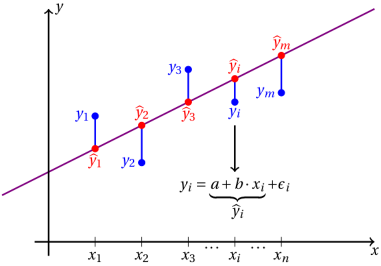
DEPARTAMENTO DE INFORMÁTICA

INF 285 - Computación Científica
Ingeniería Civil Informática

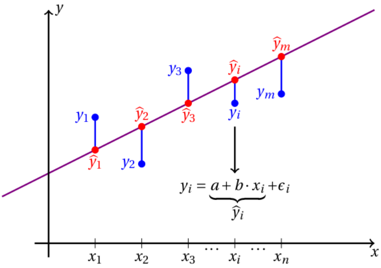
10: Mínimos Cuadrados + QR

Mínimos Cuadrados

2^{a} alternativa - Álgebra Lineal

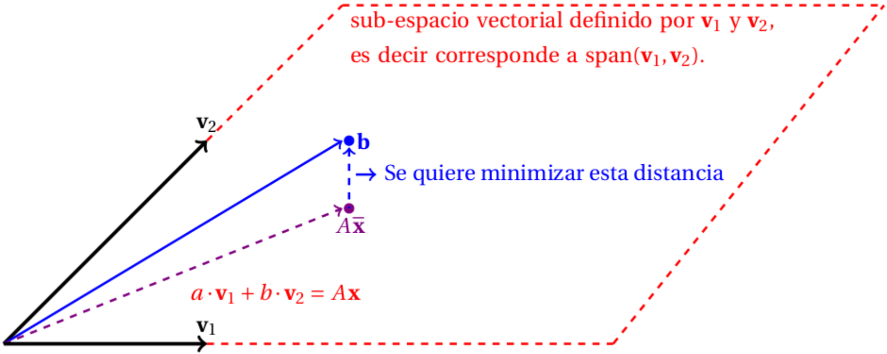


$$
\begin{array}{ccc}
a + b\,x_1 & = & y_1 \\
a + b\,x_2 & = & y_2 \\
a + b\,x_3 & = & y_3 \\
\vdots & \vdots & \vdots \\
a + b\,x_m & = & y_m
\end{array}
\Rightarrow
\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}
\begin{bmatrix} a \\ b \end{bmatrix}
=
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix}
\Rightarrow A\,\boldsymbol{x} = \boldsymbol{b}
$$



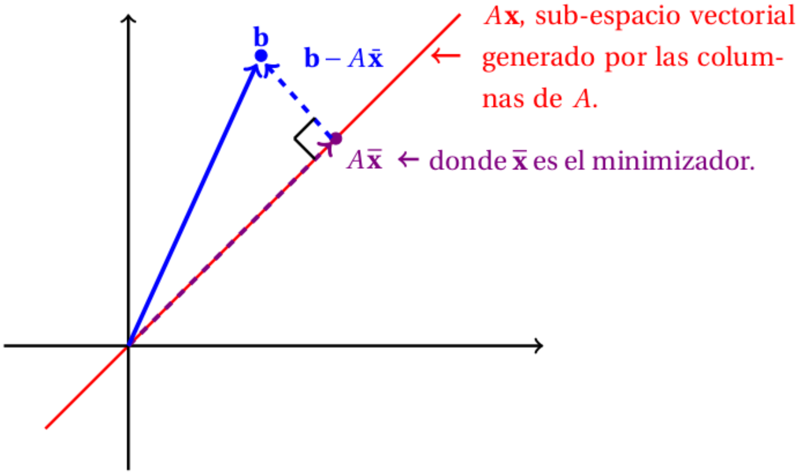
$$A\,\boldsymbol{x} = \boldsymbol{b} \Rightarrow \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \boldsymbol{b} \Rightarrow a\,\boldsymbol{v}_1 + b\,\boldsymbol{v}_2 = \boldsymbol{b}$$

$$A\,\boldsymbol{x} = \boldsymbol{b} \Rightarrow \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \boldsymbol{b} \Rightarrow a\,\boldsymbol{v}_1 + b\,\boldsymbol{v}_2 = \boldsymbol{b}$$

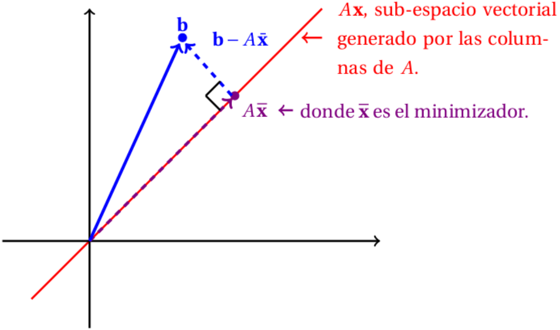


Mínimos Cuadrados

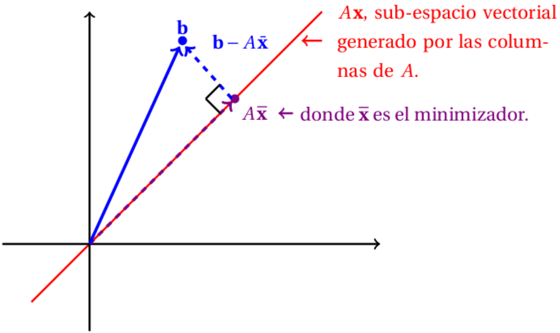
2ª alternativa - Álgebra Lineal



¿Qué relación podemos determinar para $\overline{\boldsymbol{x}}$?



$$\boldsymbol{r} = \boldsymbol{b} - A\,\overline{\boldsymbol{x}} \perp \operatorname{span}\left(\boldsymbol{a}_1, \boldsymbol{a}_2 \cdots, \boldsymbol{a}_n\right)$$



$$(A\,\boldsymbol{x})^*\;(\boldsymbol{b} - A\,\overline{\boldsymbol{x}}) = 0$$

$$\boldsymbol{x}^*\;(A^*\,\boldsymbol{b} - A^*A\,\overline{\boldsymbol{x}}) = 0$$

pero $x^* \neq 0$

Ecuaciones Normales $\Rightarrow A^*A\,\overline{\boldsymbol{x}} = A^*\,\boldsymbol{b}$

Ejemplo 1

Considere el siguiente problema de mínimos cuadrados,

$$\begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -3 \\ 15 \\ 9 \end{bmatrix}$$

Factorización QR

$$A\,\boldsymbol{x} = \boldsymbol{b} \Rightarrow \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boldsymbol{b}$$

Idea:

$\mathrm{span}(\boldsymbol{a}_1, \boldsymbol{a}_2) = \mathrm{span}(\boldsymbol{q}_1, \boldsymbol{q}_2)$

$\langle \boldsymbol{q}_1, \boldsymbol{q}_2 \rangle = 0$ y $||\boldsymbol{q}_i|| = 1$

$$\begin{aligned} \boldsymbol{a}_1 &= r_{11}\,\boldsymbol{q}_1 \\ \boldsymbol{a}_2 &= r_{12}\,\boldsymbol{q}_1 + r_{22}\,\boldsymbol{q}_2 \end{aligned}$$

¿Cómo obtenemos $\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_n$?

$$\underbrace{\begin{bmatrix} \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \\ \\ \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n \\ \\ \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}}_{R}$$

Ortogonalización de Gram-Schmidt

$$\boldsymbol{a}_1 = r_{11}\,\boldsymbol{q}_1 \Rightarrow \|\boldsymbol{a}_1\| = |r_{11}|\,\|\boldsymbol{q}_1\|$$

Pero $\|\boldsymbol{q}_1\| = 1$, luego $|r_{11}| = r_{11}$, entonces

$$\boldsymbol{q}_1 = \frac{\boldsymbol{a}_1}{r_{11}}$$

$$\boldsymbol{a}_1 = r_{11}\,\boldsymbol{q}_1 \Rightarrow \|\boldsymbol{a}_1\| = |r_{11}|\,\|\boldsymbol{q}_1\|$$

Pero $\|\boldsymbol{q}_1\| = 1$, luego $|r_{11}| = r_{11}$, entonces

$$\boldsymbol{q}_1 = \frac{\boldsymbol{a}_1}{r_{11}}$$

Ahora, tenemos que $\boldsymbol{a}_2 = r_{12}\,\boldsymbol{q}_1 + r_{22}\,\boldsymbol{q}_1$

¿Cómo obtenemos r_{22}?

Recordemos que $\boldsymbol{q}_1$ y $\boldsymbol{q}_2$ ortogonales!

$$\boldsymbol{a}_2 = r_{12}\,\boldsymbol{q}_1 + r_{22}\,\boldsymbol{q}_2$$

$$\begin{aligned}
\boldsymbol{q}_1^*\,\boldsymbol{a}_2 &= \boldsymbol{q}_1^*\,\left(r_{12}\,\boldsymbol{q}_1 + r_{22}\,\boldsymbol{q}_2\right)\\
&= r_{12}\,\underbrace{\boldsymbol{q}_1^*\,\boldsymbol{q}_1}_{1} + r_{22}\underbrace{\boldsymbol{q}_1^*\,\boldsymbol{q}_2}_{0}\\
&= r_{12}
\end{aligned}$$

Luego,

$$\begin{aligned}
\boldsymbol{a}_2 &= r_{12}\,\boldsymbol{q}_1 + r_{22}\,\boldsymbol{q}_2\\
\boldsymbol{a}_2 - r_{12}\,\boldsymbol{q}_1 &= r_{22}\,\boldsymbol{q}_2\\
\|\boldsymbol{a}_2 - r_{12}\,\boldsymbol{q}_1\| &= r_{22}
\end{aligned}$$

Entonces,

$$\boldsymbol{q}_2 = \frac{\boldsymbol{a}_2 - r_{12}\,\boldsymbol{q}_1}{r_{22}}$$

En forma general se obtiene que:

$$\boldsymbol{a}_k = r_{1k}\,\boldsymbol{q}_1 + r_{2k}\,\boldsymbol{q}_2 + \cdots + r_{k-1,k}\boldsymbol{q}_{k-1} + r_{kk}\,\boldsymbol{q}_k$$

$$\underbrace{\begin{bmatrix} \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \\ \\ \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n \\ \\ \end{bmatrix}}_{\tilde{Q}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}}_{\tilde{R}}$$

Factorización QR reducida

Factorización QR

Hasta ahora entonces tenemos un conjunto de n vectores ortogonales que permiten abarcar un subespacio de $\mathbb{R}^m$. Pero ¿podemos abarcar el espacio completo?

Nos faltan $m-n$ vectores, los cuales pueden ser agregados a la matriz Q:

$$\underbrace{\begin{bmatrix} \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \\ \\ \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \\ \boldsymbol{q}_1 & \cdots & \boldsymbol{q}_n & \color{blue}{\boldsymbol{q}_{n+1}} & \color{blue}{\cdots} & \color{blue}{\boldsymbol{q}_m} \\ \\ \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}}_{R}$$

Factorización QR completa

$$\underbrace{\begin{bmatrix} \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \\ \\ \end{bmatrix}}_{\substack{A \\ m\times n}} = \underbrace{\begin{bmatrix} \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_m \\ \\ \end{bmatrix}}_{\substack{Q \\ m\times m}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}}_{\substack{R \\ m\times n}}$$

Definición 1

Una matriz Q es **unitaria** *si* $Q^* = Q^{-1}$

La principal propiedad de las matrices unitarias, es que la norma Eucledіana de un vector se preserva:

$$\|Q\,\boldsymbol{x}\|_2^2 = (Q\,\boldsymbol{x})^*(Q\,\boldsymbol{x}) = \boldsymbol{x}^*\,Q^*\,Q\,\boldsymbol{x} = \boldsymbol{x}^*\,\boldsymbol{x} = \|\boldsymbol{x}\|_2^2 \qquad (1)$$

```
for k in range(1,n+1):
    $\mathbf{y} = \mathbf{a}_k$
    for i in range(1,k):
        $r_{ik} = \mathbf{q}_i^T \mathbf{a}_k$
        $\mathbf{y} = \mathbf{y} - r_{ik} \mathbf{q}_i$
    $r_{k,k} = \|\mathbf{y}\|_2$
    $\mathbf{q}_k = \dfrac{\mathbf{y}}{r_{k,k}}$
```

```
for k in range(1,n+1):
    $\mathbf{y} = \mathbf{a}_k$
    for i in range(1,k):
        $r_{ik} = \mathbf{q}_i^T \mathbf{y}$
        $\mathbf{y} = \mathbf{y} - r_{ik}\mathbf{q}_i$
    $r_{k,k} = \|\mathbf{y}\|_2$
    $\mathbf{q}_k = \frac{\mathbf{y}}{r_{k,k}}$
```