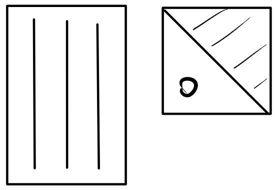
Ejemplo de factorización QR (reducida)

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 5 & 1 \\ 1 & 1 & 1 \\ 1 & 5 & 0 \end{pmatrix} =$$



$\hat{Q}$ $\hat{R}$

$$A^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad A^{(2)} = \begin{pmatrix} 1 \\ 5 \\ 1 \\ 5 \end{pmatrix} \qquad A^{(3)} = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

$$q_1 = \frac{A^{(1)}}{\underbrace{\|A^{(1)}\|}_{= r_{11} = 2}} = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$$

$$q_2 = \frac{V^{(2)}}{\underbrace{\|V^{(2)}\|}_{= r_{22} = 4}} , \quad V^{(2)} = A^{(2)} - \underbrace{\langle A^{(2)}, q_1 \rangle}_{r_{12}} q_1 = \begin{pmatrix} 1 \\ 5 \\ 1 \\ 5 \end{pmatrix} - \begin{pmatrix} 3 \\ 3 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ -2 \\ 2 \end{pmatrix}$$

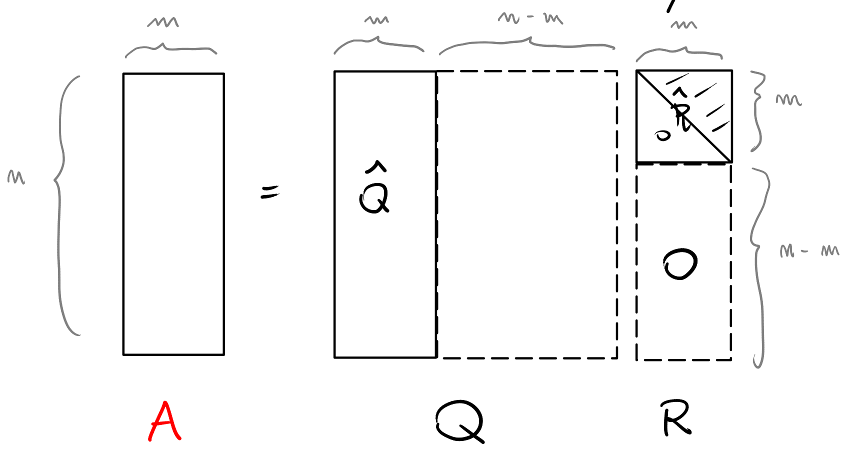
$$r_{12} = (1,5,1,5) \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} = 6$$

$$q_3 = \frac{V^{(3)}}{\underbrace{\|V^{(3)}\|}_{= r_{33} = 1}} , \quad V^{(3)} = A^{(3)} - \underbrace{\langle A^{(3)}, q_1 \rangle}_{r_{13} = 2} q_1 - \underbrace{\langle A^{(3)}, q_2 \rangle}_{r_{23} = -1} q_2 = \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix}$$

$$\Rightarrow \hat{Q} = \begin{pmatrix} 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 \end{pmatrix} , \quad \hat{R} = \begin{pmatrix} 2 & 6 & 2 \\ 0 & 4 & -1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \hat{Q}\hat{R} = A .$$

6.3 TRIANGULARIZACIÓN DE HOUSEHOLDER

1. en ciertas ocasiones es útil tener una factorización QR en la que Q es UNITARIA



$\mathbb{C}^{m \times n}$

$m \geq n$

factorización QR completa

$Q \in \mathbb{C}^{m \times m}$

$R \in \mathbb{C}^{m \times n}$

t.q. $Q^*Q = QQ^* = I$ ← unitaria (ortogonal en $\mathbb{R}$)

<u>observaciones</u> sobre QR completa:

- Q se puede obtener por Gram-Schmidt construyendo $\hat{Q}$ y completando la base → se escogen vectores aleatorios y se procede con Gram-Schmidt
- cualquier algoritmo que produzca la QR completa construye una Q cuyas primeras columnas generan el mismo subespacio generado por las columnas de A → la R tiene ceros debajo de la $\hat{R}$

2. problema: el método de Gram-Schmidt es intrínsecamente <u>inestable</u>:

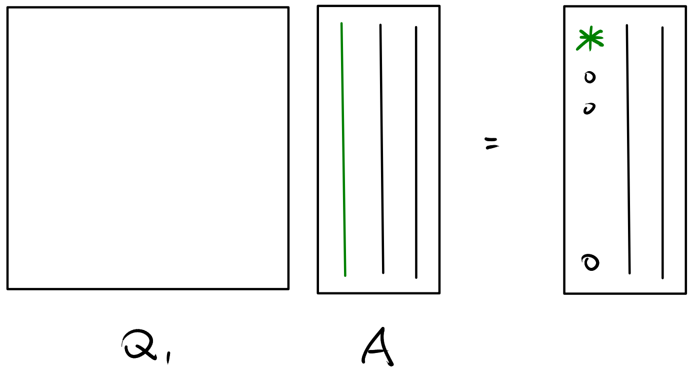
<u>ejemplo</u> ilustrativo: sean $A^{(1)}, A^{(2)} \in \mathbb{R}^n$ normalizados tales que $\langle A^{(1)}, A^{(2)} \rangle = 1 - \varepsilon$ por un ε pequeño

↘ casi paralelos

- $V^{(2)} = A^{(2)} - (1-\varepsilon) A^{(1)}$ ← diferencia de números cercanos: pérdida de precisión en la representación float
- $\| V^{(2)} \|^2 = \underbrace{\| A^{(2)} \|^2}_{=1} + (1-\varepsilon)^2 \underbrace{\| A^{(1)} \|^2}_{=1} - 2(1-\varepsilon) \underbrace{\langle A^{(2)}, A^{(1)} \rangle}_{(1-\varepsilon)} = 1 - (1-\varepsilon)^2$

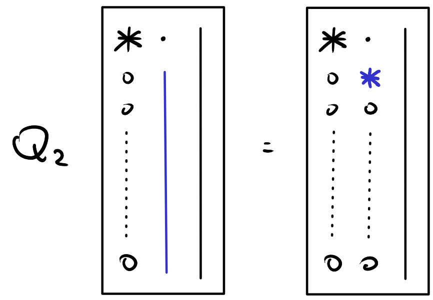
↑ número pequeño

al dividir las componentes de $V^{(2)}$ por este se amplifican los errores idea de Householder: buscar Q_1 unitaria t.q.



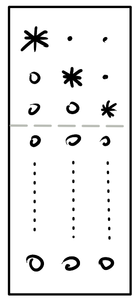
en la primera columna ponemos a cero todos los elementos excepto por el primero

↳ si sabemos encontrar una Q_1 así, entonces podemos hacer la misma operación para la segunda columna de $Q_1 A$, mirando solo a partir del segundo elemento:



⇒ si A tiene m columnas, podemos construir m matrices unitarias $\{Q_j\}_{j=1}^m$ tales que

$$\underbrace{Q_m Q_{m-1} \cdots Q_1}_{\text{producto de matrices unitarias: sigue siendo una matriz unitaria}} A =$$



triangular superior

construcción: reflexión de Householder

<u>lema</u>: sea $x \in \mathbb{C}^k$, $x \neq 0$, y sea $e_1 = (1, 0, \dots, 0)^t \in \mathbb{C}^k$

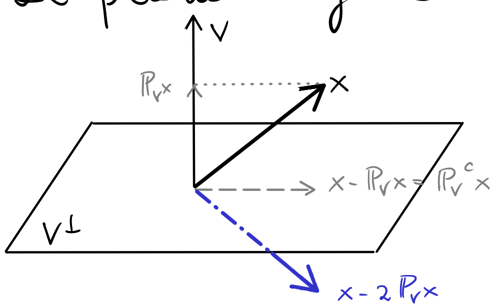
$\Rightarrow \exists\, v \in \mathbb{C}^k$, $\beta \in \mathbb{C}$, $\beta \neq 0$ t.q. $(I - 2\mathbb{P}_v)\, x = \beta e_1$.

<u>observaciones</u>:

- $I - 2\mathbb{P}_v$ es unitaria: $(I - 2\mathbb{P}_v)^*(I-2\mathbb{P}_v) \overset{\mathbb{P}_v^* = \mathbb{P}_v}{=} I - 4\mathbb{P}_v + 4\mathbb{P}_v^2 \overset{\mathbb{P}_v = \mathbb{P}_v^2}{=} I$

  ⇒ el lema nos dice que es una matriz del tipo deseado: es unitaria y manda cualquier x en un múltiplo de e_1

- geometricamente $I - 2\mathbb{P}_v$ es una reflexión respecto al plano ortogonal a v. Ejemplo: sea $v \in \mathbb{R}^3$, $\|v\| = 1$



$I - \mathbb{P}_v$ es una proyección ortogonal sobre $v^\perp$ (plano) que se obtiene restando $\mathbb{P}_v x$ a x. restando $\mathbb{P}_v x$ una vez más se manda x a su simétrico respecto al plano $v^\perp$

demostración (en $\mathbb{R}^k$): queremos hallar β, v t.q $(I - 2\mathbb{P}_v)x = \beta e_1$

recordar que $\mathbb{P}_v x = \langle x, \frac{v}{\|v\|} \rangle \frac{v}{\|v\|}$

- $\mathbb{P}_v x = \frac{1}{2}(x - \beta e_1) \iff \langle x, v \rangle v = \frac{\|v\|^2}{2}(x - \beta e_1)$
  $\iff v = \alpha(x - \beta e_1)$ para cualquier α

  $\Rightarrow \alpha \langle x, v \rangle = \frac{\|v\|^2}{2}$

- $$\begin{cases} \langle x, v \rangle = \alpha\left(\|x\|^2 - \beta \underbrace{\langle x, e_1 \rangle}_{x_1}\right) \\ \|v\|^2 = \alpha^2 \langle x - \beta e_1, x - \beta e_1 \rangle = \alpha^2\left(\|x\|^2 + \beta^2 - 2\beta \langle x, e_1 \rangle\right) \end{cases}$$

- desde la identidad $\alpha \langle x, v \rangle = \frac{\|v\|^2}{2}$ obtenemos

$$\|x\|^2 - \beta x_1 = \frac{\|x\|^2 + \beta^2 - 2\beta x_1}{2} \quad \Rightarrow \quad \beta = \pm \|x\| \qquad \#$$

$\cdot \sim \cdot$

Estabilidad de esta operación respecto al problema de manejar vectores casi paralelos:

- dado un x, hemos encontrado (en $\mathbb{R}^k$) dos vectores que pueden generar una reflexión que cumple con nuestros requisitos: $v = x \pm \|x\| e_1$
- si x es casi paralelo a e_1, eligiendo el signo − tenemos el mismo problema de antes: perdemos precisión en v y amplificamos este error al normalizarlo para construir $\mathbb{P}_v$. por la misma razón, si x es casi paralelo a $-e_1$ queremos usar el signo − y no el signo +
- $\Rightarrow$ $\boxed{v = x + \mathrm{signo}_+(x_1)\, \|x\| e_1}$, donde $\mathrm{signo}_+(0) = 1$

lo único que importa para ver si x es casi paralelo a e_1 o a $-e_1$ es el signo de $x_1 = \langle x, e_1 \rangle$. si $x_1 = 0$ es indiferente, y se elige +.

¿cómo cambia la demostración en $\mathbb{C}^k$?

- las diferencias estan en α y en el producto escalar $\langle x, e_1 \rangle = x_1$, que podrían ser complejos

- para $\beta \in \mathbb{R}$, tenemos $\begin{cases} \langle x, v \rangle = \bar{\alpha}\left(\|x\|^2 - \beta x_1\right) \\ \|v\|^2 = |\alpha|^2\left(\|x\|^2 + \beta^2 - 2\beta\, \mathrm{Re}(x_1)\right) \end{cases}$

  ↳ la identidad $\alpha \langle x, v \rangle = \frac{\|v\|^2}{2}$ sigue eliminando α de la ecuación ↳ solo hay que ocuparse de x_1

  la última identidad es entonces:

  $$\|x\|^2 - \beta x_1 = \frac{\|x\|^2 + \beta^2 - 2\beta\, \mathrm{Re}(x_1)}{2} \iff \beta^2 = \|x\|^2 - 2i\beta\, \mathrm{Im}(x_1)$$

  $\Rightarrow$ si $\mathrm{Im}(x_1) \neq 0$ no existe solución con β real

- para $\beta \in \mathbb{C}$, tenemos $\begin{cases} \langle x, v \rangle = \bar{\alpha}\left(\|x\|^2 - \bar{\beta} x_1\right) \\ \|v\|^2 = |\alpha|^2\left(\|x\|^2 + |\beta|^2 - 2\, \mathrm{Re}(\bar{\beta} x_1)\right) \end{cases}$

  desde aquí se obtiene

  $$\|x\|^2 - \bar{\beta} x_1 = \frac{\|x\|^2 + |\beta|^2 - 2\, \mathrm{Re}(\bar{\beta} x_1)}{2} \iff |\beta|^2 = \|x\|^2 - 2i\, \mathrm{Im}(\bar{\beta} x_1)$$

  $$\iff \begin{cases} |\beta|^2 = \|x\|^2 \\ \bar{\beta} x_1 \in \mathbb{R} \end{cases} : \quad \beta = \begin{cases} \pm \frac{x_1}{|x_1|} \|x\| & \text{si } x_1 \neq 0 \\ \|x\| & \text{si } x_1 = 0 \end{cases}$$

en $\mathbb{C}^k$ se elige $\boxed{v = x + \sigma_+(x_1)\, \|x\|\, e_1}$

donde $\sigma_+(z) = \begin{cases} \frac{z}{|z|} & \text{si } z \neq 0 \\ 1 & \text{si } z = 0 \end{cases}$ para $z \in \mathbb{C}$.

↳ notar que $\sigma_+(x_1) = \mathrm{signo}_+(x_1)$ si $x_1 \in \mathbb{R}$.