

Def: Llamamos matrices de Stiefel a matrices ~~en~~ $\mathbb{C}^{m \times n}$ ($m \geq n$)

tal que sus vectores columnas son un conjunto ortonormal.
Denotaremos a este conjunto por $St_{m,n}$.

Ejemplo particular es que $St_{n,n} = \mathcal{U}(n)$ son las matrices unitarias.

(Se puede probar que $St_{m,n}$ tienen estructura de variedad diferenciable.)

La clase anterior vimos que toda proyección ortogonal podía escribirse como $\hat{Q} \cdot \hat{Q}^*$ con $\hat{Q} \in St_{m,n}$. : $\hat{Q} = (q_1 | \dots | q_n)$ $q_i \in \mathbb{C}^m$

(Observa que $\hat{Q} \hat{Q}^* q_i = q_i$ $\forall i=1, \dots, n$) y que $\hat{Q} \hat{Q}^* x = 0$
 $\forall x \in (\text{Im } \hat{Q})^\perp$ quien coincide con $\text{Ker } \hat{Q}^* \rightarrow \langle \hat{Q}^* v, w \rangle = \langle v, \hat{Q} w \rangle$

DESCOMPOSICIÓN QR

La descomposición QR dice que toda matriz $A \in \mathbb{C}^{m \times n}$ se puede escribir como $A = \hat{Q} \hat{R}$ con $\hat{Q} \in St_{m,n}$ y $\hat{R} \in \mathbb{C}^{n \times n}$ triangular superior.

Veamos esto. Si $A = (a_1 | \dots | a_n) \in \mathbb{C}^{m \times n}$ entonces ~~tenemos~~ ^{si consideramos} las

sucesivas subespacios $\langle a_1 \rangle \subseteq \langle a_1, a_2 \rangle \subseteq \dots \subseteq \langle a_1, \dots, a_j \rangle \subseteq \dots$

podemos construir una ~~matriz~~ $\hat{Q} = (q_1 | \dots | q_n)$ colección de vectores q_1, q_2, \dots

tales que $\langle a_1 \rangle = \langle q_1 \rangle$, $\langle a_1, a_2 \rangle = \langle q_1, q_2 \rangle$, \dots , $\langle a_1, \dots, a_j \rangle = \langle q_1, \dots, q_j \rangle$

(Siempre que $\{a_1, \dots, a_j\}$ sean l.i.)

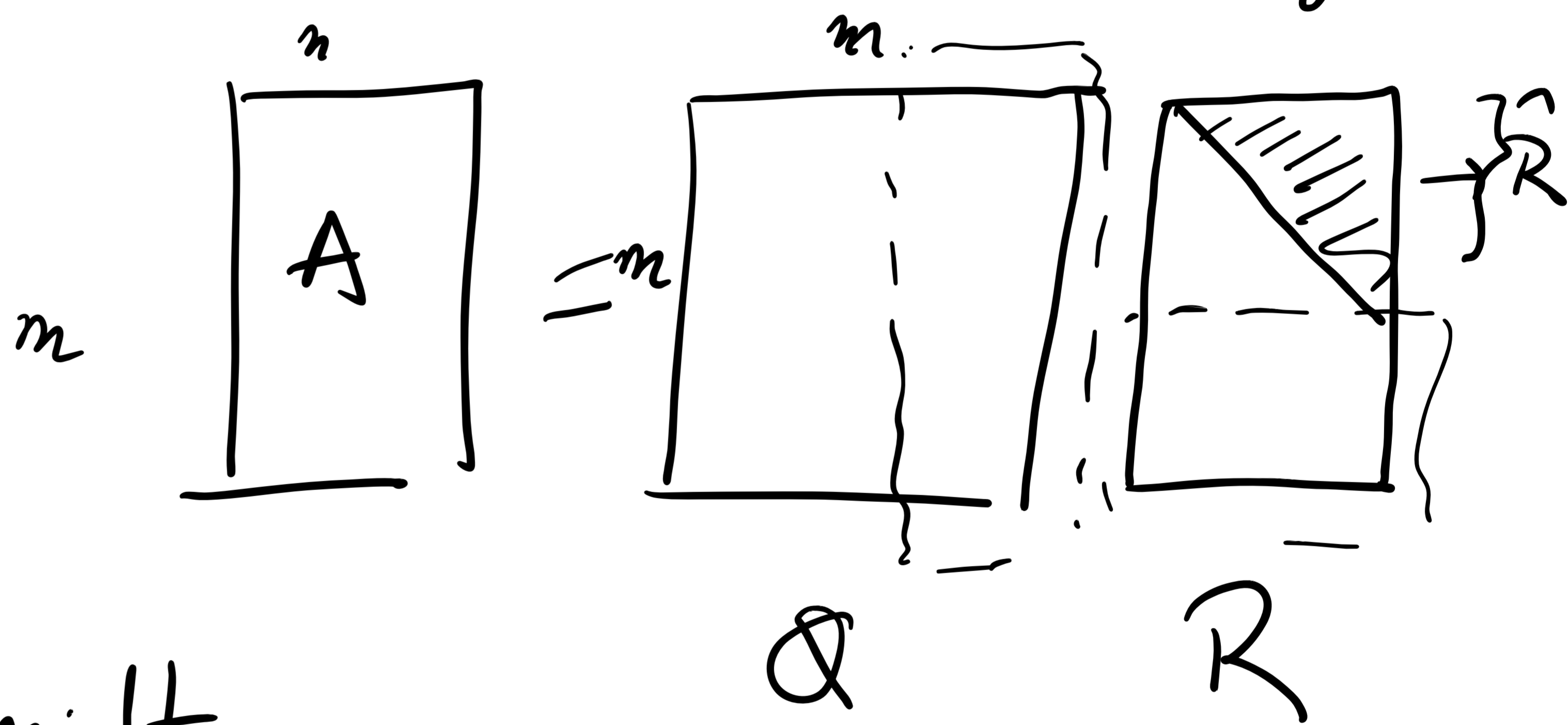
Luego tenemos que $\begin{cases} a_1 = r_{11} q_1 \\ a_2 = r_{12} q_1 + r_{22} q_2 \\ \vdots \\ a_m = r_{1m} q_1 + r_{2m} q_2 + \dots + r_{mm} q_m \end{cases}$ en el caso A de rango máximo.

y por lo tanto $A = \hat{Q} \cdot \hat{R} = \begin{pmatrix} | & & | \\ q_1 & & q_m \\ | & & | \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ 0 & r_{22} & & \\ & & \ddots & \\ & & & r_{mm} \end{pmatrix}$

(Si A no tiene rango máximo el proceso anterior vale igual eligiendo $q_i \perp \text{Im } A$ y tomando $r_{ij} = 0$ adecuadamente)

La Full descomposición QR se completa agregando $m-n$ vectores q_j ortogonales en \mathbb{C}^m y la matriz \hat{R} se extiende a una matriz $m \times m$ con ceros de la siguiente forma

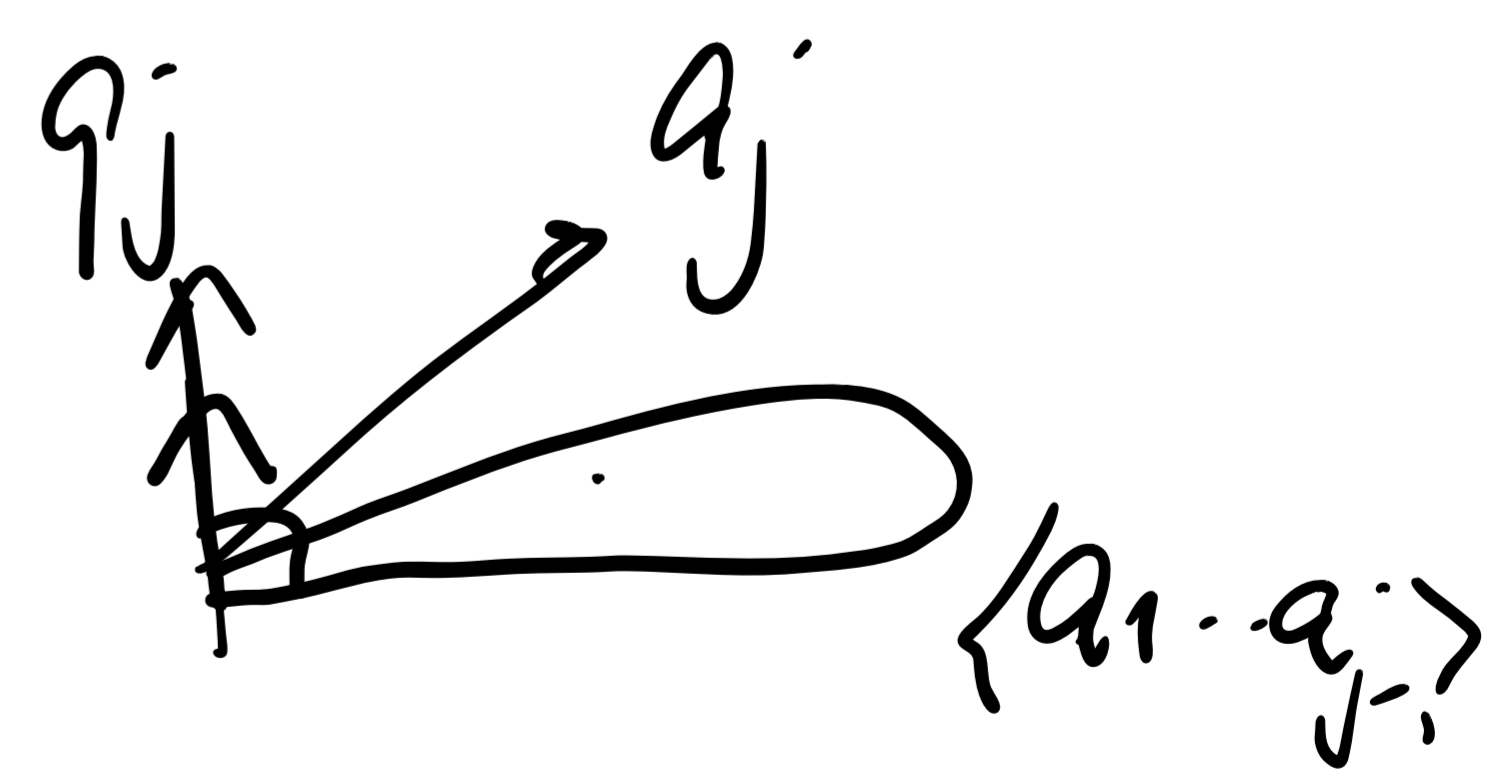
Observa que



Ortogonalización Gram-Schmidt

Un proceso constructivo para los q_i es utilizar el proceso de ortogonalización de Gram-Schmidt.

La idea es muy sencilla y conocida: (ver*)



Tomamos $q_1 = \frac{a_1}{\|a_1\|} = \frac{a_1}{r_{11}}$ donde $r_{11} = \|a_1\|$.

$p_2 = a_2 - q_1 q_1^* a_2$ $q_2 = p_2 / \|p_2\|$

$p_3 = a_3 - q_1 q_1^* a_3 - q_2 q_2^* a_3$ $q_3 = p_3 / \|p_3\|$

$\Rightarrow q_j = \frac{p_j a_j}{\|p_j a_j\|}$ donde $p_j = I - \hat{Q}_{j-1} \hat{Q}_{j-1}^*$ y $\hat{Q}_{j-1} = (q_1 | \dots | q_{j-1})$

donde $P_j = I_{\mathbb{R}^n} - \hat{Q}_{j-1} \hat{Q}_{j-1}^T$ es la proyección ortogonal en $\langle q_1, \dots, q_{j-1} \rangle^\perp$.
 Algoritmicamente G-S se escribe de la siguiente manera:

Algoritmo (G-S clásico)

for $j=1$ a n
 $v_j = a_j$
 for $i=1$ a $j-1$
 $r_{ij} = q_i^T v_j$
 $v_j = v_j - r_{ij} q_i$
 $r_{jj} = \|v_j\|$
 $q_j = v_j / r_{jj}$

No unicidad de descomposición QR

$$(q_1 | \dots | q_n) \begin{pmatrix} r_{11} & * & \\ & r_{22} & \\ & & \ddots & \\ & & & r_{nn} \end{pmatrix} =$$

"

$$(e^{i\theta} q_1 | q_2 | \dots | q_n) = \begin{pmatrix} e^{-i\theta} & & & & \\ & r_{11} & & & \\ & & r_{22} & & \\ & & & \ddots & \\ & & & & r_{nn} \end{pmatrix}$$

Unicidad de la descomposición

Sea $A \in \mathbb{C}^{m \times m}$ de rango máximo. posee una única descomposición QR si $r_{jj} > 0$ (i.e. la diagonal es positiva).

Dem: ~~versión repetida~~ $A = \hat{Q}_1 \hat{R}_1 = \hat{Q}_2 \hat{R}_2$ tenemos que $\hat{Q}_2^* \hat{Q}_1 = \hat{R}_2 \hat{R}_1^{-1}$
 y esta matriz luego esta matriz es unitaria y triangular superior y con valores reales positivos en la diagonal $\Rightarrow \hat{Q}_2^* \hat{Q}_1 = \hat{R}_2 \hat{R}_1^{-1} = Id$.

Queda como ejercicio ver la inversa de una matriz triangular superior (con elts de la diagonal positivos, lo cual implica que es invertible) es también triangular positiva con igual propiedad en la diagonal.

Solución de sistemas lineales

Si queremos resolver $Ax=b$ con $A \in \mathbb{C}^{n \times n}$ no singular podemos considerar la descomposición QR de $A=QR$ y escribir $Q \cdot Rx=b$ o $\underline{Rx=Q^*b}$.

• Calcular $y=Q^*b$ es muy fácil conociendo Q

• Luego se calcula la solución de $Rx=y$ lo cual también es muy fácil:

conociendo $\leftarrow \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & r_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix}$

luego $r_{nn}x_n = y_n$

$r_{(n-1)n}x_{n-1} + r_{(n-1)n-1}x_{n-1} = y_{n-1}$

\vdots

es fácil de calcular.

Polinomios Ortogonales

La descomp. QR ~~tiene~~ tiene en análogo en dim infinita, como por ejemplo en espacios de funciones.

Si cambiamos \mathbb{C}^m por $L^2[E, \mu]$ funciones de cuadrado integrable en $E=[a, b]$. Se puede definir un producto interno $\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx$

(Si $f: E \rightarrow \mathbb{R}$ podemos olvidarnos del conjugado) -

(Para más formalización de $L^2(E, \mu)$ ver al curso de Teoría de la Medida.)

Juego si consideramos la "matriz" $A = (1 | x | x^2 | \dots | x^{n-1})$

donde cada columna es una función de $L^2[-1,1]$ (de hecho la versión continua de la matriz de Vandermonde) tenemos que podemos escribir

$$A = QR = (q_0(x) | q_1(x) | \dots | q_{n-1}(x)) \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & & \\ & & 0 & \\ & & & r_{nn} \end{pmatrix}$$

donde $\langle q_i, q_j \rangle = \delta_{ij} \int_{-1}^1 q_i(x) q_j(x) dx$.

Del método de G-S podemos ver que $q_j(x)$ es un polinomio de grado j . ($q_j \in \langle 1, x, \dots, x^j \rangle$) $P_j(x)$

Los polinomios de Legendre son múltiplos de los $q_j(x)$ con

la convención $P_j(1) = 1$:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{3}{2}x^2 - \frac{1}{2} \\ &\vdots \end{aligned}$$

Resultado $\langle P_j, P_k \rangle = 0$ si $j \neq k$ y por eso se llaman polinomios ortogonales.

En otras variantes de productos interiores

$$\langle P, q \rangle = \int_{-1}^1 P(x) q(x) w(x) dx$$

que decantan en otros polinomios conocidos. $Q = (q_0(x) | \dots | q_{n-1}(x))$

El map $f \in L^2[-1,1] \xrightarrow{\Pi_{\langle q_i \rangle}} \sum_{i=0}^{n-1} \langle f, q_i \rangle q_i$

es la proyección ortogonal a $\langle q_0(x), \dots, q_{n-1}(x) \rangle$, i.e. $\Pi_{\langle q_i \rangle} = Q Q^T$

Gram-Schmidt revisitado

Vimos que el paso j del algoritmo lo podríamos identificar con la normalización del vector

$$v_j = P_j a_j = \text{Id}_m - \hat{Q}_{j-1} \hat{Q}_{j-1}^*$$

con $\hat{Q}_{j-1} = (q_1 \dots q_{j-1})$ ya construidos.

Una manera análoga (pero computacionalmente diferente como veremos más adelante) es observar que

$$P_j = P_{q_{j-1}^\perp} P_{q_{j-2}^\perp} \dots P_{q_2^\perp} P_{q_1^\perp}$$

i.e. producto de proyección en subespacios de codimensión 1 $\text{Id} - q_i q_i^*$.

Esta observación permite construir el proceso Modificado de G-S.

$$v_j^{(1)} = a_j$$

$$v_j^{(2)} = P_{q_1^\perp} v_j^{(1)} = v_j^{(1)} - q_1 q_1^* v_j^{(1)}$$

$$v_j^{(3)} = P_{q_2^\perp} v_j^{(2)} = v_j^{(2)} - q_2 q_2^* v_j^{(2)}$$

⋮

$$v_j := v_j^{(j)} = P_{q_{j-1}^\perp} v_j^{(j-1)} = v_j^{(j-1)} - q_{j-1} q_{j-1}^* v_j^{(j-1)}$$

Una vez construido el vector q_i podemos aplicar el proyector $P_{q_i^\perp}$ a $v_j^{(i)}$ con $j > i$:

Algoritmo: G-S Modificado

```

for i=1 a n
  v_i = a_i
for i=1 a n
  r_ii = ||v_i||
  q_i = v_i / r_ii
  for j=i+1 a n
    r_ij = q_i^* v_j
    v_j = v_j - r_ij q_i
  
```

Complejidad del Alg

Entendamos por complejidad al n^2 de operaciones aritméticas

Acá contaremos cada suma, producto, resta, división, y raíz cuadrada como una op. aritmética; un paso; un "flop".
(en \mathbb{C} o en \mathbb{R} contamos igual)

Por ejemplo, $z \in \mathbb{C}^m$
 $\langle z, w \rangle = z_1 \bar{w}_1 + \dots + z_m \bar{w}_m$
 contabiliza en m -productos y $(m-1)$ -sumas \Rightarrow # pasos $\approx 2m$

Es fácil ver que la complejidad está dominada por el loop interior

$$\left. \begin{array}{l} r_{ij} = q_i^* v_j \quad (\approx 2m) \\ v_j = v_j - r_{ij} q_i \quad (\approx 2m) \end{array} \right\} \approx 4m \quad \text{y luego los } \text{pasos recorridos de } i, j \text{ son } \frac{n(n+1)}{2}$$

de donde concluimos $4m \cdot \frac{n(n+1)}{2} \approx \boxed{2m \cdot n^2}$.

GS como iteración triangular

Observar que podemos ver que el proceso G-S puede ser visto como multiplicar por matrices triangulares superiores en cada paso:

$$\left(\sigma_1 | \dots | \sigma_n \right) \begin{pmatrix} 1 & & & & \\ & \frac{r_{12}}{r_{11}} & & & -\frac{r_{1n}}{r_{11}} \\ & 0 & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = \left(q_1 | \sqrt{\sigma_2^{(2)}} | \dots | \sqrt{\sigma_n^{(2)}} \right)$$

Después multiplicamos

$$\Rightarrow A = R_1 R_2 \dots R_n = Q$$

Por $R_2 = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ 0 & \frac{1}{r_{22}} & & & \\ & & \frac{-r_{23}}{r_{22}} & & \\ & & & \ddots & \\ & & & & \frac{-r_{2n}}{r_{22}} \\ & & & & & 1 \end{pmatrix}$

$$R_3 = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \frac{1}{r_{33}} & \\ & & & & \frac{-r_{34}}{r_{33}} \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$