

Computación Científica I,¹ 2007
Guía N^o 2
Valparaíso–Santiago, 13 de mayo de 2007²

1. DESCOMPOSICIÓN DE VALOR SINGULAR

1. Considere la matriz $A = \begin{bmatrix} 3 & 2 \\ -1 & 2 \\ 2 & 0 \end{bmatrix}$.

- (a) Determine los valores singulares de A .
- (b) Determine los vectores singulares *derechos* de A , i.e., los vectores-columna de la matriz \widehat{V} de la SVD *reducida* $\widehat{U}\widehat{\Sigma}\widehat{V}^*$ de A . Note bien que lo pedido son los vectores-columna de \widehat{V} y no los de \widehat{V}^* .
- (c) Determine los vectores singulares *izquierdos* de A , i.e., los vectores-columna de la matriz \widehat{U} de la SVD *reducida* $\widehat{U}\widehat{\Sigma}\widehat{V}^*$ de A .
- (d) Verifique que la SVD de A que Ud. obtuvo es correcta.

DESARROLLO.

- (a) *Preámbulo.* Consideremos la SVD *reducida* de $A \in M(m \times n, \mathbb{C})$:

$$A = U\Sigma V^* \quad \text{y, equivalentemente,} \quad A^* = V\Sigma U^*,$$

donde $U \in M(m \times n, \mathbb{C})$ y $V \in M(n \times n, \mathbb{C})$ tienen columnas respectivamente ortonormales (i.e., $(U^j)^*U^k = \delta_{jk}$ y $(V^j)^*V^k = \delta_{jk}$)³, y

$$\Sigma = \text{diag} [\sigma_1, \sigma_2, \dots, \sigma_n] \in M(n \times n, \mathbb{C}), \quad \text{con} \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0,$$

es una matriz diagonal. De aquí se desprende que:

$$A^*A = V\Sigma U^*U\Sigma V^* = V\Sigma^2 V^*,$$

donde $A^*A \in M(n \times n, \mathbb{C})$. Luego, $(A^*A).V = V.\Sigma^2$, i.e.:

$$\begin{aligned} (A^*A).[V^1|V^2|\dots|V^n] &= [V^1|V^2|\dots|V^n].\text{diag} [\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2] \\ &= [\sigma_1^2 V^1, \sigma_2^2 V^2, \dots, \sigma_n^2 V^n], \end{aligned}$$

donde V^k es la k -ésima columna de V . Por consiguiente, los vectores-columna de la matriz V de la SVD de A deben satisfacer:

$$(A^*A).V^k = \sigma_k^2 V^k, \quad k = 1 : n,$$

i.e., el k -ésimo vector-columna de la matriz V debe ser un vector propio de la matriz A^*A con valor propio σ_k^2 . Esta observación permite obtener la matriz V con cierta facilidad. En el caso real habrá que considerar, evidentemente, la matriz $A^T A$.

¹Proyecto UTFSM-DGIP 2006-2007 Desarrollo de Textos Docentes.

²© Luis Salinas Carrasco, Valparaíso, 13 de mayo de 2007. De antemano se agradece toda corrección, crítica o comentario que el amable lector tenga a bien hacer llegar a luis.salinas@usm.cl.

³ δ_{jk} es el llamado *símbolo de Kronecker* definido mediante $\delta_{jk} = \begin{cases} 1, & \text{si } j = k, \\ 0, & \text{si } j \neq k. \end{cases}$

- (b) Cálculo de $A^T A$. Se obtiene directamente $A^T A = \begin{bmatrix} 14 & 4 \\ 4 & 8 \end{bmatrix}$.

Cálculo de los valores propios de $A^T A$ mediante la solución de la ecuación $\det(A^T A - \lambda I) = 0$. Se obtiene $\lambda_1 = \sigma_1^2 = 16$ y $\lambda_2 = \sigma_2^2 = 6$.

Cálculo de los valores singulares de $A^T A$ tomando raíz de los valores propios ya determinados. Se obtiene $\sigma_1 = 4$ y $\sigma_2 = \sqrt{6}$.

- (c) Como se sabe, los vectores singulares *derechos* v_i de A corresponden a los vectores propios de la matriz $A^T A$, normalizados a norma 1. Luego, hay que resolver el sistema $A^T A.v = \lambda v$.

Caso $\lambda = \lambda_1 = 16$: se obtiene $v_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$.

Caso $\lambda = \lambda_2 = 6$: se obtiene $v_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$.

Nota: Los *signos* de v_1 y v_2 no están unívocamente determinados, pero ello no influye en el desarrollo ulterior.

- (d) Análogamente, los vectores singulares *izquierdos* de A se obtienen resolviendo el sistema $A\hat{V} = \hat{U}\hat{\Sigma}$, donde $\hat{V} = [v_1 | v_2]$, $\hat{\Sigma} = \text{diag}[\sigma_1, \sigma_2]$ y los σ_i son los valores singulares obtenidos en (a). Luego, basta resolver el sistema

$$\underbrace{\begin{bmatrix} 8/\sqrt{5} & 1/\sqrt{5} \\ 0 & \sqrt{5} \\ 4/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}}_{=AV} = \begin{bmatrix} 3 & 2 \\ -1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & \sqrt{6} \end{bmatrix} = \underbrace{\begin{bmatrix} 4 u_{11} & \sqrt{6} u_{12} \\ 4 u_{21} & \sqrt{6} u_{22} \\ 4 u_{31} & \sqrt{6} u_{32} \end{bmatrix}}_{=U\Sigma}$$

para $u_1 = [u_{11}, u_{21}, u_{31}]^T$ y $u_2 = [u_{12}, u_{22}, u_{32}]^T$. De consiguiente:

$$u_1 = \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \end{bmatrix} = \frac{1}{\sigma_1} \text{columna}_1(AV) = \frac{1}{4} \text{columna}_1 \left(\begin{bmatrix} 8/\sqrt{5} & 1/\sqrt{5} \\ 0 & \sqrt{5} \\ 4/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \right) = \begin{bmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{bmatrix}$$

$$u_2 = \begin{bmatrix} u_{12} \\ u_{22} \\ u_{32} \end{bmatrix} = \frac{1}{\sigma_2} \text{columna}_2(AV) = \frac{1}{\sqrt{6}} \text{columna}_2 \left(\begin{bmatrix} 8/\sqrt{5} & 1/\sqrt{5} \\ 0 & \sqrt{5} \\ 4/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \right) = \begin{bmatrix} 1/\sqrt{30} \\ \sqrt{5/6} \\ -\sqrt{2/15} \end{bmatrix}$$

- (e) Verificación:

$$\hat{U} \hat{\Sigma} \hat{V}^T = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{30} \\ 0 & \sqrt{5/6} \\ 1/\sqrt{5} & -\sqrt{2/15} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & \sqrt{6} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -1 & 2 \\ 2 & 0 \end{bmatrix} = A$$

2. (a) Construya su propia demostración para el siguiente teorema o estudie a fondo las demostraciones que proponen algunos de los textos de referencia para este importante resultado del Algebra Lineal Computacional. Asegúrese de haber comprendido correctamente las demostraciones que estudie.

Teorema 1. [8, Theor. 4.1, p. 29] *Sea $A \in M(m \times n, \mathbb{C})$. Entonces existen matrices unitarias $U \in M(m \times m, \mathbb{C})$ y $V \in M(n \times n, \mathbb{C})$, y una matriz diagonal $\Sigma = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_n] \in M(m \times n, \mathbb{C})$ con $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$,⁴ tales que $A = U\Sigma V^*$.*

⁴Obsérvese que la matriz Σ no es, en general, cuadrada. No obstante, aún en ese caso se acostumbra a decir que Σ es una matriz diagonal. Cuando Σ no es cuadrada, sus filas de por debajo de la fila n , si $m \geq n$, o sus columnas a la derecha de la columna m , si $m \leq n$, son todas nulas.

Los valores singulares están siempre unívocamente determinados.

Si A es una matriz cuadrada y los σ_k son todos distintos, entonces los los vectores singulares izquierdos:

$$U^j = j\text{-ésima columna de la matriz } U, \quad j = 1 : m = n,$$

los vectores singulares derechos:

$$V^k = k\text{-ésima columna de la matriz } V, \quad j = 1 : n = m,$$

están unívocamente determinados excepto por signos complejos (i.e., factores escalares complejos de módulo 1).

(b) Las demostraciones más populares del resultado precedente, son inductivas. Estudie la posibilidad de traducir algunas de esas demostraciones inductivas en algoritmos computacionales que permita obtener la SVD de una matriz dada A .

3. (Cf. [8, Lect. 5, p. 32]) Sea $A \in M(m \times n, \mathbb{C})$ una matriz dada y $A = U\Sigma V^*$ su SVD reducida. Sea $r = |\{\sigma_j : \sigma_j > 0\}|$. Demuestre o refute, y en todo caso, discuta:

(a) $\text{rank } A = r$.

(b) $\text{Range}(A) = \text{Col}(A) = \langle U^1, \dots, U^r \rangle$, donde $U^j = j$ -ésimo vector-columna de U .

(c) $\text{Null}(A) = \langle V^{r+1}, \dots, V^n \rangle$, donde $V^k = k$ -ésimo vector-columna de V .

(d) $\|A\|_2 = \sigma_1$; $\|A\|_{\text{Frobenius}} = \sqrt{\sigma_1^2 + \dots + \sigma_r^2}$; $\min_{x \neq 0} \frac{\|A \cdot x\|_2}{\|x\|_2} = \sigma_n$ ($m \geq n$).

(e) Los valores singulares σ_j no nulos vienen dados por $\sigma_j = \sqrt{\lambda_j}$, donde λ_j es el j -ésimo valor propio no nulo de la matriz $A^*A \in M(n \times n, \mathbb{C})$ o de la matriz $AA^* \in M(m \times m, \mathbb{C})$. Ambas matrices tiene los mismos valores propios no nulos. Los valores propios de A^*A y AA^* son reales y no negativos.

(f) Si $A^* = A$, los valores singulares de A coinciden con los valores absolutos de los valores propios de A .

(g) Si $A \in M(m \times m, \mathbb{C})$, $|\det A| = \prod_{j=1}^m \sigma_j$.

4. [8, Lect. 5, p. 35] APROXIMACIÓN DE BAJO RANGO. Sea $A \in M(m \times n, \mathbb{C})$ una matriz dada y $A = U\Sigma V^*$ su SVD reducida. Sea $p = \min\{m, n\}$ y $r = |\{\sigma_j : \sigma_j > 0\}|$. Discuta, verifique, corrija o eventualmente refute los siguientes argumentos y resultados:

(a) Dado que $\Sigma = \text{diag} [\sigma_1, \dots, \sigma_r, \underbrace{0, \dots, 0}_p] = \sum_{j=1}^r \text{diag} [0, \dots, 0, \underbrace{\sigma_j}_j, 0, \dots, \underbrace{0}_p]$, se tiene:

$$A = U\Sigma V^* = \sum_{j=1}^r U \cdot \text{diag} [0, \dots, 0, \underbrace{\sigma_j}_j, 0, \dots, \underbrace{0}_p] \cdot V^* = \sum_{j=1}^r \sigma_j U^j \cdot (V^j)^*,$$

(b) Sea $A_\nu = \sum_{j=1}^\nu \sigma_j U^j \cdot (V^j)^*$, $\nu = 0 : r$. Entonces:

$$A - A_\nu = \sum_{j=\nu+1}^r \sigma_j U^j \cdot (V^j)^* = \underbrace{[U^{\nu+1} | \dots | U^r]}_{U_{\nu+1}} \cdot \underbrace{\text{diag} [\sigma_{\nu+1}, \dots, \sigma_r]}_{\Sigma_{\nu+1}} \cdot \underbrace{\begin{bmatrix} (V^{\nu+1})^* \\ \vdots \\ (V^r)^* \end{bmatrix}}_{V_{\nu+1}^*}$$

constituye una SVD para $A - A_\nu$: $A - A_\nu = U_{\nu+1} \cdot \Sigma_{\nu+1} \cdot V_{\nu+1}^*$. Por definición de la SVD se tiene entonces:

$$\|A - A_\nu\|_2 = \text{primer (mayor) valor singular} = \sigma_{\nu+1}. \tag{1}$$

(c) Sea $E = \langle V^1, \dots, V^{\nu+1} \rangle$ el espacio de \mathbb{C}^n engendrado por los primeros $\nu + 1$ vectores singulares derechos. Para todo $w \in E$ se tiene entonces:

$$\begin{aligned} (A.w)^* \cdot (A.w) &= w^* A^* A w = w^* V \Sigma U^* U \Sigma V^* w = \underbrace{w^* V}_{z^*} \Sigma^2 \underbrace{V^* w}_z = z^* \Sigma^2 z \\ &= \sigma_1^2 |z_1|^2 + \dots + \sigma_{\nu+1}^2 |z_{\nu+1}|^2 \geq \underbrace{(|z_1|^2 + \dots + |z_{\nu+1}|^2)}_{=\|z\|^2=\|w\|^2} \sigma_{\nu+1}^2. \end{aligned}$$

Nótese que $\|z\|^2 = \|V^* \cdot w\|^2 = \|w\|^2$, pues la matriz $V \in M(n \times n, \mathbb{C})$ es unitaria y, por lo tanto, preserva la norma Euclideana. Por consiguiente se tiene:

$$\|A.w\|^2 \geq \sigma_{\nu+1}^2 \|z\|^2 = \sigma_{\nu+1}^2 \|w\|^2 \quad \forall w \in E. \tag{2}$$

(d) Sobre la base de los resultados precedentes se puede demostrar el siguiente

Teorema 2. Para todo $\nu = 0 : r$ se tiene:

$$\|A - A_\nu\|_2 \stackrel{(1)}{=} \sigma_{\nu+1} \stackrel{\perp}{=} \inf_{\substack{B \in M(m \times n, \mathbb{C}) \\ \text{rank } B \leq \nu}} \|A - B\|_2. \tag{3}$$

Si $\nu = \min\{m, n\}$, hacemos $\sigma_{\nu+1} = 0$.

Demostración. Sólo hay que probar la segunda igualdad en (3). Supongamos que exista una matriz $B \in M(m \times n, \mathbb{C})$ con $\text{rank } B \leq \nu$ tal que:

$$\|A - B\|_2 < \|A - A_\nu\|_2 = \sigma_{\nu+1}.$$

Consideremos entonces el subespacio:

$$W = \text{Null}(B) \subseteq \mathbb{C}^n, \quad \text{con} \quad \dim W = n - \text{rank } B \geq n - \nu.$$

Para $w \in W$ se tiene entonces $B.w = 0$, $A.w = (A - B).w$, y:

$$\|A.w\|_2 = \|(A - B).w\|_2 \leq \|A - B\|_2 \|w\|_2 < \sigma_{\nu+1} \|w\|_2.$$

Luego, W es un subespacio de \mathbb{C}^n de dimensión a lo menos $n - \nu$ donde se cumple $\|A.w\|_2 < \sigma_{\nu+1} \|w\|_2$.

Por otro lado, de (c) sabemos que el subespacio E de \mathbb{C}^n generado por los $\nu + 1$ vectores singulares derechos V^j , $j = 1 : \nu + 1$, de A , posee las siguientes propiedades:

$$\dim E = \nu + 1 \quad \text{y} \quad \|A.w\|_2^2 \stackrel{(2)}{\geq} \sigma_{\nu+1}^2 \|w\|^2 \quad \forall w \in E.$$

En vista de $\dim W \geq n - \nu$ y $\dim E = \nu + 1$, los subespacios W y E de \mathbb{C}^n deben tener una intersección que no se reduce al subespacio trivial $\{0\}$. En efecto, si $W \cap E \neq \{0\}$, entonces se tendría:

$$n + 1 = (n - \nu) + (\nu + 1) \leq \dim W + \dim E \leq \dim \mathbb{C}^n = n,$$

que obviamente es una contradicción. Existe, por lo tanto, un vector $0 \neq w \in W \cap E$ que debe satisfacer simultáneamente $\|A.w\|_2 < \sigma_{\nu+1} \|w\|_2$, por estar en W , y $\|A.w\|_2 \geq \sigma_{\nu+1} \|w\|_2$, por estar en E . Ello representa, evidentemente, otra contradicción. ■

Nota. El lector diligente seguramente habrá observado que la demostración precedente sólo prueba la desigualdad $\sigma_{\nu+1} \leq \inf_{\substack{B \in M(m \times n, \mathbb{C}) \\ \text{rank } B \leq \nu}} \|A - B\|_2$ en (3). ¿Cómo se obtendría la

otra desigualdad?

5. Sea $A \in M(n \times n, R)$ una matriz positiva definida. Para $i_k \in \mathbb{N}$ con $1 \leq i_1 < i_2 < \dots < i_p \leq n$ sea $A[i_1, \dots, i_p]$ la submatriz de A formada por las filas i_1, \dots, i_p y las columnas i_1, \dots, i_p de la matriz A . Note, entonces, que la diagonal de $A[i_1, \dots, i_p]$ está contenida en la diagonal de A . Se dice que $A[i_1, \dots, i_p]$ es una *submatriz principal* de A .

(a) Demuestre o refute, y en cualquier caso, discuta: $\det A[i_1, \dots, i_p] > 0$ para toda selección de los índices $1 \leq i_1 < i_2 < \dots < i_p \leq n$.

(b) ¿Es equivalente (a) con el hecho que A es positiva definida?

(c) ¿Qué relación hay entre los valores singulares de las matrices $A[i_1, \dots, i_p]$ y los valores singulares de la matriz A ?

(d) Sea $V(A, \gamma)$ el volumen del elipsoide $\{x \in \mathbb{R}^n : x^T A x \leq \gamma\}$. Calcule el cociente $\frac{V(A, \gamma)}{\det(A)^{1/2}}$.

6. [4, p. 415] Sea $x \in \mathbb{C}^n$ un vector no nulo dado. Considere el vector x como una matriz $X \in M(n \times 1, \mathbb{K})$. Sean $W = [1] \in M(1 \times 1, K)$, $\Sigma = [\|x\|_2, 0, \dots, 0]^T \in M(n \times 1, \mathbb{K})$, $V = \left[\frac{x}{\|x\|_2}, v_2, \dots, v_n \right] \in M(n \times 1, \mathbb{K})$, donde v_2, \dots, v_n es un sistema de $n - 1$ vectores ortonormales cualesquiera, todos ellos ortogonales al vector x . Demuestre que una descomposición de valor singular de la matriz X viene dada por $X = V\Sigma W^*$.

7. [4, p. 416] Si $A \in M(n \times n, \mathbb{K})$ es no singular, demuestre que el siguiente procedimiento produce una descomposición de valor singular $A = V\Sigma W^*$:

(i) Forme la matriz Hermitiana positiva definida AA^* y determine una diagonalización unitaria $AA^* = U\Lambda U^*$ hallando los valores propios positivos $\lambda_i, i = 1, \dots, n$, y el correspondiente sistema de $u_i, i = 1, \dots, n$, de vectores propios ortonormales.

(ii) Defina $\Sigma := \Lambda^{1/2}$ y $V = U = [u_1, \dots, u_n]$.

(iii) Defina $W = A^*V\Sigma^{-1}$.

a) Demuestre todo. En particular, demuestre que W es unitaria con $A = V\Sigma W^*$.

b) Escriba un programa que permita obtener la descomposición de valor singular.

c) Usando su programa, obtenga descomposiciones de valor singular para tres o cuatro matrices de dimensiones 3, 5, y 8, por ejemplo. Para resolver esta parte del ejercicio el estudiante probablemente querrá aplicar el algoritmo descrito en el ejercicio precedente. Converse con sus colegas que abordaron ese ejercicio para que le faciliten os códigos correspondientes.

d) Usando un programa *ad hoc* verifique que el producto de las matrices de las factorizaciones que obtenga, efectivamente dan las matrices originales.

8. [4, p. 416] Sea $A \in M(n \times n, K)$ una matriz no necesariamente no-singular. Demuestre que el siguiente procedimiento produce una descomposición de valor singular $A = V\Sigma W^*$:

(i) Sea $I \in M(n \times n, K)$ la matriz identidad. Entonces existe una constante $c = c(A)$ tal que la matriz $A_\varepsilon = A + \varepsilon I$ es no-singular para todo $\varepsilon \in]0, c[\subset \mathbb{R}$. *Sugerencia:* Observe las raíces de la ecuación característica $\det(A - \lambda I) = 0$ en el plano complejo.

(ii) Elija ahora $\varepsilon \in]0, c[\subset \mathbb{R}$. Aplique el procedimiento del ejercicio anterior para obtener una descomposición de valor singular $A_\varepsilon = V_\varepsilon \Sigma_\varepsilon W_\varepsilon^*$

(iii) Aplique ahora el siguiente resultado, *que no se pide demostrar*:

Lema Auxiliar. Sea $\{U_p\}_{p \in \mathbb{N}} = [u_{ij}^p]_{p \in \mathbb{N}}$ una sucesión de matrices unitarias en $M(n \times n, \mathbb{K})$. Entonces existe una subsucesión $\{U_{p_k}\}_{k \in \mathbb{N}} = [u_{ij}^{p_k}]_{k \in \mathbb{N}}$ de la sucesión original, tal que

$$\lim_{k \rightarrow \infty} u_{ij}^{p_k} = u_{ij}^0, \quad \forall i, j = 1, \dots, n,$$

donde la matriz límite $U_0 = [u_{ij}^0] \in M(n \times n, \mathbb{K})$ también es unitaria.

Sobre la base de este lema auxiliar se obtiene una sucesión $\{\varepsilon_k\}_{k \in \mathbb{N}}$ en $]0, c[$, una matriz diagonal $\Sigma \in M(n \times n, \mathbb{K})$ con coeficientes no negativos, y matrices unitarias $V, W \in M(n \times n, \mathbb{K})$ tales que:

$$\lim_{k \rightarrow \infty} \varepsilon_k = 0, \quad \lim_{k \rightarrow \infty} V_{\varepsilon_k} = V, \quad \lim_{k \rightarrow \infty} \Sigma_{\varepsilon_k} = \Sigma, \quad \lim_{k \rightarrow \infty} W_{\varepsilon_k} = W.$$

(iv) Demuestre que $A = V\Sigma W^*$.

9. [4, p. 421, pr. 7] Sea $A \in M(m \times n, \mathbb{K})$ una matriz dada con descomposición de valor singular $A = V\Sigma W^*$. Defina $A^\dagger = W\Sigma^\dagger V^*$, donde Σ^\dagger es la transpuesta de Σ en la cual los valores singulares *positivos* de A se reemplazan por sus valores recíprocos. Demuestre que:

(a) AA^\dagger y $A^\dagger A$ son hermitianas, (b) $AA^\dagger A = A$, (c) $A^\dagger AA^\dagger = A^\dagger$.

Demuestre que $A^\dagger = A^{-1}$ si A es cuadrada y no-singular. La matriz A^\dagger se llama *inversa generalizada* o *inversa de Moore-Penrose* de la matriz A . Se puede demostrar que la inversa generalizada existe para toda matriz A , incluso para matrices cuadradas singulares o para matrices no cuadradas.

Demuestre, además, que A^\dagger queda unívocamente determinada por las propiedades (a), (b), (c).

10. [4, p. 422, pr. 10] La descomposición de valor singular se puede obtener sin utilizar explícitamente vectores ni valores propios. Los vectores singulares (de izquierda y de derecha) y los valores singulares pueden construirse directamente a partir de la llamada *caracterización variacional de la norma espectral* (que es, simplemente, la definición usual de la norma de un operador lineal): si $A \in M(n \times n, \mathbb{K})$ es una matriz dada, entonces se define:

$$\|A\|_2 = \sup_{\|x\|_2=1} \|Ax\|_2. \quad (\text{Norma Espectral})$$

(a) Sea $n \geq 2$ y suponga que $B \in M(n \times n, \mathbb{K})$ tiene la forma:

$$B = \begin{bmatrix} \sigma_1 & w^* \\ 0 & X \end{bmatrix}, \quad \text{con } \sigma_1 = \|B\|_2, \quad w \in \mathbb{C}^{n-1}, \quad X \in M((n-1) \times (n-1), \mathbb{K}).$$

Demuestre que $w = 0$. *Sugerencia:* Si $\sigma > 0$, considere $\zeta = \frac{1}{(\sigma_1^2 + w^*w)^{1/2}} \begin{bmatrix} \sigma_1 \\ w \end{bmatrix}$, demuestre que $\|B\zeta\|_2^2 \geq \sigma_1^2 + w^*w$, y use la fórmula (Norma Espectral). Además, ¡consulte [8]!

(b) Sea $A \in M(n \times n, \mathbb{K})$, defina $\sigma_1 = \|A\|_2$, y use la fórmula (Norma Espectral) para probar que existe un vector unitario $\|x\|_1$ tal que $\|Ax_1\|_2 = \sigma_1$. Defina $y_1 = \sigma_1^{-1} Ax_1$ para la parte (c). *Nota:* Posiblemente el estudiante quiera usar aquí el teorema de Weierstrass para funciones continuas definidas sobre conjuntos compactos.

(c) Sean $V_1, W_1 \in M(n \times n, \mathbb{K})$ matrices unitarias cuyas primeras columnas son x_1 e y_1 respectivamente. Demuestre que $V_1^* AW_1$ tiene norma espectral σ_1 y que tiene la forma de la matriz B de la parte (a). Concluya que $V_1^* AW_1 = \begin{bmatrix} \sigma_1 & 0 \\ 0 & X \end{bmatrix}$.

(d) Formule un proceso de “deflación inductiva” de A introduciendo columnas y filas con “ceros fuera de la diagonal” (cuando se las observa en la matriz A) mediante pre- y post-multiplicación por matrices unitarias apropiadas. Verifique que el proceso de inducción termina tras un número finito de pasos y que al final se obtiene la descomposición de valor singular de A .

(e) ¿Qué ocurre si la matriz original A no es cuadrada?

11. (Cf. [4, Theor. 7.3.2, p. 412]) Considere una descomposición de valor singular $A = U\Sigma V^*$ de una matriz cuadrada $A \in M(n \times n, \mathbb{K})$. Escriba:

$$A = U\Sigma V^* = (U\Sigma U^*) (UV^*) = PW, \quad \text{con } P = U\Sigma U^* \text{ y } W = UV^*. \quad (4)$$

Demuestre o refute, y en todo caso, discuta:

- (a) P es positiva semidefinida, i.e., $x^*Px \geq 0$ (note que dice “ \geq ” y no “ $>$ ”) para todo $x \in \mathbb{K}^n$, con $x \neq 0$.
- (b) $\text{rank } P = \text{rank } A$.
- (c) W tiene filas ortonormales, i.e. $WW^* = I$.
- (d) P está unívocamente determinada y satisface $P = (AA^*)^{1/2}$. Dé una interpretación razonable para esta última fórmula.

Suponga ahora que la matriz original A no es cuadrada, digamos $A \in M(m \times n, \mathbb{K})$ con $m \geq n$. A “*bruta forza*”, haga cuadradas (de $m \times m$) todas las matrices (exceptuando V^* que conviene dejarla como una matriz de $m \times n$) de la descomposición de valor singular, agregando 0’s a diestra y siniestra.

- (e) Examine si la factorización (4) y las propiedades (a)-(d) siguen siendo válidas o no.
- (f) La nueva matriz W ¿está unívocamente determinado cuando $\text{rank } A = n$?
- (g) Si A es real, ¿es posible tomar P y W como matrices reales?
- (h) ¿Qué pasa si $m \leq n$?

Como el lector sospechará, (casi) todas las respuestas deberían ser afirmativas. La factorización obtenida se conoce como “*representación polar*” de una matriz dada.

- (i) Escriba un programa para obtener la representación polar de cualquier matriz.

12. [4, p. 415][8, p. 37, exc. 5.2] Usando SVD⁵ demuestre que toda matriz $A \in M(m \times n, \mathbb{C})$ es el límite (¿cuál norma sería apropiada para calcular el límite?) de una sucesión de matrices $A_k = \left(a_{ij}^{(k)} \right)_{m \times n} \in M(m \times n, \mathbb{C})$, $k \in \mathbb{N}$, de rango completo. En otras palabras, demuestre que el conjunto de las matrices de rango completo *es denso* en $M(m \times n, \mathbb{C})$.

13. [8, p. 37, exc. 5.3] [For all students!!!] Considere la matriz $\begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix}$.

- (a) Determine s.c.P.L.P.⁶ una SVD real de A de la forma $A = U\Sigma V^T$. La SVD no es única (compruebe este hecho). Así, encuentre la SVD que tiene el menor número de signos negativos en U y V .
- (b) Haga una tabla con los valores singulares, los vectores singulares izquierdos, y los vectores singulares derechos de A . Dibuje cuidadosamente la bola unitaria de \mathbb{R}^2 y su imagen bajo A , junto con los vectores singulares, indicando claramente las coordenadas de sus vértices. Use sistemas de coordenadas claramente rotulados con los nombres de las variables involucradas y todos los signos convencionales del Dibujo Técnico que aseguren una buena comprensión de sus diagramas.
- (c) Determine las normas 1, 2, ∞ , y de Frobenius-Hilbert-Schmidt de la matriz A .
- (d) Determine A^{-1} indirectamente vía SVD.
- (e) Verifique que: $\det A = \lambda_1\lambda_2$ y $|\det A| = \sigma_1\sigma_2$.
- (f) Determine el área del elipsoide que resulta como imagen de la bola unitaria de \mathbb{R}^2 bajo A .

⁵“singular value decomposition”

⁶s.c.P.L.P. = sólo con papel, lápiz y pensamiento

14. [8, p. 37, exc. 5.4] Suponga que $A \in M(m \times m, \mathbb{C})$ tiene una SVD $A = U\Sigma V^*$. Encuentre una EVD ⁷ $A = X\Lambda X^{-1}$ de la matriz hermitiana $\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix} \in M(2m \times 2m, \mathbb{C})$.

15. Cf.: <http://www.computer.org/cise/homework/v5n3/index.htm>

James G. Nagy and Dianne P. O’Leary, [6]. Los llamados *problemas inversos* constituyen uno de los desafíos más exigentes en la computación aplicada a la ciencia y a la ingeniería, debido a que involucran la determinación de los parámetros de un sistema observado sólo indirectamente. Un ejemplo de tales problemas es, por ejemplo, la determinación de un sistema a partir solamente de su *espectro*. Otro ejemplo típico es la determinación de posibles fracturas en un (es)tanque exclusivamente a partir de mediciones de tipo sonar.

El problema que se quiere discutir aquí es el siguiente: dada una imagen borrosa y un modelo lineal para la borrosidad, reconstruir la imagen original. Este problema inverso lineal ilustra el impacto del *mal condicionamiento* en la selección de algoritmos. La herramienta principal para este proyecto es la SVD.

(1) MAL CONDICIONAMIENTO. Consideremos un sistema de ecuaciones lineales:

$$K.f = g, \quad K \in M(n \times n, \mathbb{R}), \quad f, g \in \mathbb{R}^n.$$

Supondremos que la matriz K está escalada de modo que su mayor valor singular es $\sigma_1 = 1$. Si el menor valor singular σ_n es aproximadamente 0, entonces K está mal condicionada. Distinguiremos dos tipos de mal condicionamiento:

- La matriz K se considera *numéricamente de rango deficiente* si existe un j tal que $\sigma_j \gg \sigma_{j+1} \approx \dots \approx \sigma_n \approx 0$, i.e., si hay una brecha ostensible entre los valores singulares grandes y los pequeños.
- Si los valores singulares decaen a 0 sin ninguna brecha especial en el espectro, se dice que el sistema lineal $K.f = g$ es un *problema mal condicionado discreto*

El cálculo de soluciones aproximadas muy cercanas a las soluciones exactas en un problema mal condicionado discreto del tipo $K.f = g$ es, por lo general, un problema muy difícil, especialmente debido a que en muchas aplicaciones reales, g no se conoce exactamente. En realidad la data típicamente tiene la forma:

$$g = K.f + \eta,$$

donde η representa un ruido (desconocido) o errores de medición. El objetivo es, entonces, dada una matriz mal condicionada K y un vector g , calcular una (buena) aproximación del vector desconocido f .

La solución ingenua de $K.f = g$, usualmente, no es apropiada cuando K es muy mal condicionada, como en el caso que nos interesa. En estos casos se usa un procedimiento de *regularización* para hacer el problema menos sensible al ruido. En algunos círculos se habla de *robustificación* del proceso.

(2) REGULARIZACIÓN DE TIKHONOV. El método de regularización más conocido es la llamada *regularización de Tikhonov* que calcula una solución del problema de mínimos cuadrados amortiguado:

$$\min_f \{ \|g - Kf\|_2^2 + \alpha^2 \|f\|_2^2 \}. \tag{5}$$

El término $\alpha^2 \|f\|_2^2$ impone una penalización al tamaño de la solución. Dicha penalización reduce el efecto de los valores singulares pequeños.

⁷“eigenvalue decomposition”

- (3) PROBLEMA 1. Demuestre que la ecuación (5) es equivalente al problema de mínimos cuadrados lineal:

$$\min_f \left\| \begin{bmatrix} g \\ 0 \end{bmatrix} - \begin{bmatrix} K \\ \alpha I \end{bmatrix} f \right\|_2^2. \quad (6)$$

El escalar α (llamado *parámetro de regularización*) controla el grado de *suavidad* de la solución. Evidentemente, $\alpha = 0$ implica ausencia de regularización y en este caso la solución de (6) estará muy probablemente corrupta con mucho ruido. Por otro lado, si α es grande, la solución calculada no podrá ser una buena aproximación de la solución exacta f . La elección de un valor apropiado para α no es algo trivial. Varios algoritmos han sido propuestos en la literatura pero aquí se usará un método manual.

- (4) PROBLEMA 2. Demuestre que si K tiene una SVD $K = U\Sigma V^T$, entonces la ecuación (6) puede transformarse en el problema de cuadrados mínimos equivalente:

$$\min_f \left\| \begin{bmatrix} \hat{g} \\ 0 \end{bmatrix} - \begin{bmatrix} \Sigma \\ \alpha I \end{bmatrix} \hat{f} \right\|_2^2, \quad \text{con } \hat{f} = V^T f, \quad \hat{g} = U^T g. \quad (7)$$

- (5) PROBLEMA 3. Obtenga una fórmula para la solución de la ecuación (7). *Hint*: haga cero la derivada de la función de minimización y resuelva para f .

Todo esto constituye un algoritmo para determinar la solución de Tikhonov de un problema mal condicionado discreto.

- (6) SVD TRUNCADA. Otro método para regularizar el problema consiste en truncar la SVD. El siguiente Problema 4 ilustra cómo expresar la solución al problema de mínimos cuadrados en términos de la SVD.

- (7) PROBLEMA 4. Demuestre que la solución del problema:

$$\min_f \{ \|g - Kf\|_2^2 \}. \quad (8)$$

viene dada por:

$$f_{ls} = V\Sigma^\dagger U^T \equiv \sum_{i=1}^n \frac{u_i^T g}{\sigma_i} v_i, \quad (9)$$

donde u_i es la i -ésima columna de U , y v_i es la i -ésima columna de V .

Se puede ver que hay problemas en f_{ls} si un pequeño valor de σ_i divide un término $u_i^T g$ que contiene errores. En tales casos, f_{ls} quedará dominado por errores.

Para eludir esta dificultad, Richard Hanson y James Varah han sugerido truncar la expansión (9):

$$f_{ls} = \sum_{i=1}^p \frac{u_i^T g}{\sigma_i} v_i, \quad (10)$$

para algún valor $p < n$.

Ahora tenemos todos los ingredientes para resolver el problema de eliminación de la borrosidad en procesamiento de imágenes. Supongamos que tenemos una imagen borrosa o ruidosa G junto con algún conocimiento del operador de borrosidad, y queremos construir la imagen verdadera original F . Este es un ejemplo de un problema mal condicionado discreto en el cual los vectores en el sistema lineal $g = K.f + \eta$ representan los arreglos bidimensionales (F y G) de las imágenes apilados por columnas para formar vectores (por ejemplo, debajo de la primera columna de F viene la segunda columna de F y así sucesivamente hasta concluir con la n -ésima columna, obteniéndose de esta manera un vector de n^2 componentes). En la notación de Matlab se tendría:

$$f = \text{reshape}(F, n, 1), \quad g = \text{reshape}(G, n, 1). \quad (11)$$

El objetivo de este problema es, dados K y G , reconstruir una aproximación de la imagen desconocida F .

Si suponemos que F y G contienen $\sqrt{n} \times \sqrt{n}$ pixeles, entonces f y g son vectores de largo n y K es una matriz de $n \times n$ que representa la borrosidad en la imagen. Por lo general esta matriz es muy grande para usar SVD. Sin embargo, en algunos casos se puede escribir K como un producto de Kronecker $K = A \otimes B$ y entonces se puede usar SVD.

- (8) DOS PALABRITAS ACERCA DE LOS PRODUCTOS DE KRONECKER. El producto de Kronecker $A \otimes B$, donde A y B son matrices de $m \times m$ se define mediante:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mm}B \end{bmatrix}. \tag{12}$$

Teorema 3. Si $A = U_A \Sigma_A V_A^T$ y $B = U_B \Sigma_B V_B^T$, entonces $K = U \Sigma V^T$ donde $U = U_A \otimes U_B$, $\Sigma = \Sigma_A \otimes \Sigma_B$ y $V = V_A \otimes V_B$.

Tarea: Demuestre el teorema precedente.

Por consiguiente, si una matriz grande es el producto de Kronecker de dos matrices más pequeñas, entonces se puede calcular la SVD de la matriz grande. En la página “web” de J.G. Nagy y D.P. O’Leary mencionada al comienzo de este ejercicio se puede hallar un programa Matlab de muestra que ilustra la propiedad discutida.

Para resolver el problema de eliminación de la borrosidad de una imagen hay que operar cuidadosamente con matrices más pequeñas pues, de otro modo, rápidamente aparecen problemas de memoria o almacenamiento de data.

- (9) PROBLEMA 5. Escriba un programa que toma matrices A , B y una imagen G , y calcule aproximaciones de la imagen F mediante la regularización de Tikhonov y la SVD truncada. Para cada uno de estos algoritmos, realice algunos experimentos para hallar el valor del parámetro de regularización (α para Tikhonov y p para la SVD truncada) que da la imagen más nítida. La referida página “web” de J.G. Nagy y D.P. O’Leary contiene datos de muestra: una imagen borrosa G y matrices A , B . La tarea consiste en restaurar la imagen hasta el punto que se pueda leer el texto que contiene. Compare la efectividad de los dos algoritmos.

2. PROYECTORES

1. PROYECTORES EN GENERAL. Examine las proposiciones siguientes. Demuestre o refute lo que corresponda y, en todo caso, discuta.

- (a) Un *proyector* en \mathbb{C}^m es una matriz cuadrada $P \in M(m \times m, \mathbb{C})$ tal que $P^2 = P$. Evidentemente, el proyector P proyecta el espacio \mathbb{C}^m sobre el subespacio $\text{Range}(P) = \text{Col}(P) \subseteq \mathbb{C}^m$.

EJEMPLO. Considere el espacio \mathbb{R}^2 equipado con la base $\mathcal{B} = \{b_1, b_2\}$ con $b_1 = [2, 1]^T$ y $b_2 = [3, 7]^T$. Claramente \mathcal{B} no es una base ortonormal. Entonces todo vector $x = [x_1, x_2]^T \in \mathbb{R}^2$ puede escribirse en términos de la base \mathcal{B} :

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x = \xi_1 b_1 + \xi_2 b_2 = \xi_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \xi_2 \begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 2\xi_1 + 3\xi_2 \\ \xi_1 + 7\xi_2 \end{bmatrix},$$

que da origen a un sistema de ecuaciones lineales de la forma $A\xi = x$:

$$\begin{bmatrix} 2 & 3 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Se observa que el primer vector-columna de A es el vector basal b_1 y el segundo, b_2 . Como \mathcal{B} es una base de \mathbb{R}^2 , la matriz A es invertible. ξ_1 y ξ_2 son las *coordenadas* de x con respecto a la base \mathcal{B} . Las coordenadas del vector x con respecto a la base \mathcal{B} vienen dadas entonces por:

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 7 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{7}{11} & \frac{-3}{11} \\ \frac{-1}{11} & \frac{2}{11} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

La *proyección del vector x sobre el subespacio $\mathbb{R}b_1$ a lo largo del subespacio $\mathbb{R}b_2$* viene dada por la aplicación $\pi_1 : x \mapsto \xi_1 b_1$, i.e.:

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &\mapsto \xi_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \xi_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \left(\frac{7}{11} x_1 - \frac{3}{11} x_2 \right) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \left(\begin{bmatrix} \frac{7}{11} & -\frac{3}{11} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) \\ &= \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{7}{11} & -\frac{3}{11} \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{14}{11} & \frac{-6}{11} \\ \frac{7}{11} & \frac{-3}{11} \end{bmatrix}}_{=:P} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned}$$

Se observa que la matriz cuadrada P satisface:

$$P^2 = \begin{bmatrix} \frac{14}{11} & \frac{-6}{11} \\ \frac{7}{11} & \frac{-3}{11} \end{bmatrix} \begin{bmatrix} \frac{14}{11} & \frac{-6}{11} \\ \frac{7}{11} & \frac{-3}{11} \end{bmatrix} = \begin{bmatrix} \frac{14}{11} & \frac{-6}{11} \\ \frac{7}{11} & \frac{-3}{11} \end{bmatrix} = P,$$

de modo que P es efectivamente un proyector de \mathbb{R}^2 . Diremos que el proyector es *oblicuo* pues la base \mathcal{B} no es ortogonal.

- (b) *Tarea.* Determine la matriz Q del *proyector* $\pi_2 : x \mapsto \xi_2 b_2$ de \mathbb{R}^2 sobre el subespacio $\mathbb{R}b_2$ a lo largo del subespacio $\mathbb{R}b_1$.
- (c) *Tarea.* Determine todos los proyectores $P \in M(2 \times 2, \mathbb{C})$ y $P \in M(3 \times 3, \mathbb{C})$

2. PROYECTORES EN GENERAL (NO NECESARIAMENTE ORTOGONALES).

- (a) Sean $I, P \in M(m \times m, \mathbb{C})$, donde I es la matriz identidad. Entonces se tiene:
 - (i) P es un proyector si y sólo si $I - P$ es un proyector.
 $I - P$ es el llamado *proyector complementario* de P . Evidentemente, P es el proyector complementario de $I - P$.
 Ahora supóngase que P es un proyector. Entonces:
 - (ii) $\text{Range}(I - P) = \text{Null}(P)$; $\text{Range}(P) = \text{Null}(I - P)$.
 - (iii) $\text{Range}(P) \cap \text{Null}(P) = \{0\}$, $\text{Range}(P) \oplus \text{Null}(P) = \mathbb{C}^m$.
 - (iv) $\mathbb{C}^m = \text{Range}(P) \oplus \text{Null}(P)$ (basta observar que $v = P.v + (I - P).v$)
 - (v) ¿Serían válidas (ii), (iii), (iv) si P no fuera un proyector? Dé contra-ejemplos.
- (b) Recíprocamente, si S_1 y S_2 son dos subespacios de \mathbb{C}^m tales que $S_1 \oplus S_2 = \mathbb{C}^m$ (i.e., en particular $S_1 \cap S_2 = \{0\}$), entonces existe un único proyector P tal que $S_1 = \text{Range}(P) = \text{Null}(I - P)$ y $S_2 = \text{Null}(P) = \text{Range}(I - P)$.
 P se denomina *proyector de \mathbb{C}^m sobre S_1 a lo largo de S_2* e $I - P$ es el *proyector de \mathbb{C}^m sobre S_2 a lo largo de S_1* y satisface $S_2 = \text{Range}(I - P)$ y $\text{Null}(I - P) = S_1$.
 - (i) Obtenga una expresión algebraica para P , por ejemplo, en función de una base de S_1 .
 - (ii) Determine $\text{rank}(P)$ y $\text{rank}(I - P)$ en función de m y de las dimensiones de los espacios S_1 y S_2 .

3. PROYECTORES ORTOGONALES. Examine las proposiciones siguientes. Demuestre o refute lo que corresponda y, en todo caso, discuta.

- (a) Se dice que un proyector $P \in M(m \times m, \mathbb{C})$ es un *proyector ortogonal* ssi los subespacios $S_1 = \text{Range}(P)$ y $S_2 = \text{Null}(P)$ son ortogonales, i.e.:

$$u^*v = 0 \quad \forall u \in \text{Range}(P) \quad \forall v \in \text{Null}(P).$$

El proyector ortogonal complementario de P viene dado por $I - P$.

- (b) ¡Los proyectores ortogonales no son matrices ortogonales! Caracterice todos los proyectores ortogonales de 2×2 y 3×3 que son matrices ortogonales, resp., unitarias.
- (c) Sea $P \in M(m \times m, \mathbb{C})$ un proyector. Entonces: P es un proyector ortogonal si y sólo si $P = P^*$, i.e., si y sólo si P es una matriz Hermitiana. (Cf. [8, Theor. 6.1, p. 44])
- (d) Sea $\{q_1, q_2, \dots, q_n\}$ un conjunto cualquiera de vectores ortonormales en \mathbb{C}^m , $n \leq m$. Sea $S_1 = \langle q_1, q_2, \dots, q_n \rangle$ el subespacio de \mathbb{C}^m generado por los vectores ortonormales q_k . Sea $\widehat{Q} = [q_1 | q_2 | \dots | q_n] \in M(m \times n, \mathbb{C})$. Entonces el proyector *ortogonal* P sobre el subespacio S_1 viene dado por $P = \widehat{Q}\widehat{Q}^*$.
- (e) *La fórmula maravillosa.* En el caso $n = 1$ el proyector ortogonal P sobre el subespacio $\mathbb{C}q$, $q = q_1$ con $\|q\|_2 = 1$, viene dado por $P = qq^*$.

El proyector ortogonal P sobre el subespacio $\mathbb{C}a$ generado por un vector no nulo *cualquiera* $a \in \mathbb{C}^m$ (i.e., $\|a\|_2$ no necesariamente es 1) viene dado por $P = \frac{aa^*}{a^*a}$, que ciertamente es una fórmula maravillosa.

- (f) Sea $\{a_1, a_2, \dots, a_n\}$ un conjunto cualquiera de vectores *linealmente independientes* de \mathbb{C}^m , $n \leq m$. Sea $S_1 = \langle a_1, a_2, \dots, a_n \rangle$ el subespacio de \mathbb{C}^m generado por los vectores l.i. a_k . Sea $A = [a_1 | a_2 | \dots | a_n] \in M(m \times n, \mathbb{C})$. Entonces el proyector *ortogonal* P sobre el subespacio S_1 viene dado por $P = A(A^*A)^{-1}A^*$.

4. [1, p. 168, Theor. 4.25] Sean P_1, P_2 proyectores. Demuestre:

$$P := P_1 + P_2 \text{ es proyector} \Leftrightarrow P_1P_2 = P_2P_1 = 0. \quad (13)$$

Entonces P es el proyector sobre $\text{Range}(P) = \text{Range}(P_1) \oplus \text{Range}(P_2)$ a lo largo de $\text{Null}(P) = \text{Null}(P_1) \cap \text{Null}(P_2)$.

5. [1, p. 169, Theor. 4.26] Sean P_1, P_2 proyectores. Demuestre:

$$P := P_1P_2 \text{ es proyector} \Leftrightarrow P_1P_2 = P_2P_1. \quad (14)$$

Entonces P es el proyector sobre $\text{Range}(P) = \text{Range}(P_1) \cap \text{Range}(P_2)$ a lo largo de $\text{Null}(P) = \text{Null}(P_1) \oplus \text{Null}(P_2)$.

6. [1, p. 170, Theor. 4.27] Sean P_1, P_2 proyectores. Demuestre:

$$P := P_1 - P_2 \text{ es proyector} \Leftrightarrow P_1P_2 = P_2P_1 = P_2. \quad (15)$$

Entonces P es el proyector sobre $\text{Range}(P) = \text{Range}(P_1) \cap \text{Range}(P_2)$ a lo largo de $\text{Null}(P) = \text{Null}(P_1) \oplus \text{Null}(P_2)$.

7. [8, p. 47, exc. 6.2] Sea $F \in M(m \times m, \mathbb{C})$ la matriz que “invierte” las componentes de los vectores $x \in \mathbb{C}^m$, i.e., transforma el vector $(x_1, \dots, x_m)^*$ en el vector $(x_m, \dots, x_1)^*$. Sea $E = \frac{1}{2}(I_m + F)$. Trefethen dice que E es la matriz que “extrae la parte par” de los vectores $x \in \mathbb{C}^m$. Why? ¿Es E un proyector?, ¿oblícuo?, ¿ortogonal?, ¿cuáles son sus coeficientes?

8. [8, p. 47, exc. 6.3] Sea $A \in M(m \times n, \mathbb{C})$ con $m \geq n$. Demuestre que A^*A es no singular si y sólo si A tiene rango completo.

9. [8, p. 47, exc. 6.4, and p. 55, exc. 7.1] Considere las matrices $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Responda, s.c.P.L.P.:

- (a) ¿Cuáles son los proyectores ortogonales P y Q sobre $\text{Range}(A)$ y $\text{Range}(B)$, respectivamente, y cuáles son las imágenes del vector $(1, 2, 3)^*$ bajo P y Q ?
 (b) Determine s.c.P.L.P. factorizaciones QR reducidas y completas para A y B .

10. [8, p. 47, exc. 6.5] Sea $P \in M(m \times m, \mathbb{C})$ un proyector. Demuestre que $\|P\|_2 \geq 1$, con igualdad si y sólo si P es un proyector ortogonal.

11. [2, §2.6.1, p. 75] PROYECCIONES RELACIONADAS CON LA SVD. Hay varias proyecciones ortogonales asociadas con la descomposición de valor singular. Supóngase que $A = U\Sigma V^T \in M(m \times n, \mathbb{R})$ es la SVD de A y que $r = \text{rank}(A)$. Considérese las siguientes particiones de U y V :

$$U = \begin{bmatrix} U_r & \tilde{U}_r \\ r & m-r \end{bmatrix}, \quad V = \begin{bmatrix} V_r & \tilde{V}_r \\ r & n-r \end{bmatrix}. \quad (16)$$

Entonces se tiene:

$$\begin{aligned} V_r V_r^T &= \text{proyección ortogonal sobre } \text{Null}(A)^\perp = \text{Range}(A^T), \\ \tilde{V}_r \tilde{V}_r^T &= \text{proyección ortogonal sobre } \text{Null}(A), \\ U_r U_r^T &= \text{proyección ortogonal sobre } \text{Range}(A), \\ \tilde{U}_r \tilde{U}_r^T &= \text{proyección ortogonal sobre } \text{Range}(A)^\perp = \text{Null}(A^T). \end{aligned}$$

Demuestre estas proposiciones.

12. [2, Probl. 2.6.1, p. 79] Demuestre que si P es un proyector ortogonal, entonces $Q = I - 2P$ es (una matriz) ortogonal.

13. [2, Probl. 2.6.2, p. 79] ¿Qué puede Ud. decir acerca de los valores singulares de una proyección ortogonal?

3. FACTORIZACIÓN QR

1. [8, p. 55, exc. 7.2] Sea A una matriz tal que sus columnas impares son ortogonales a sus columnas pares. En una factorización reducida QR de A , $A = \hat{Q}\hat{R}$, ¿qué estructura especial tiene \hat{R} ?

2. [8, p. 55, exc. 7.3] Sea $A \in M(m \times m, \mathbb{C})$ y a_j su j -ésima columna. Demuestre algebraicamente (hay otras demostraciones analíticas) la *desigualdad de Hadamard* $|\det A| \leq \prod_{j=1}^m \|a_j\|_2$.

Dé una interpretación geométrica de este resultado con base en el hecho que el determinante es igual al volumen de un cierto paralelepípedo (ya no hay misterio, le dije todo, ¿pero cuál paralelepípedo?).

3. [8, p. 55, exc. 7.4] Sean $a, b \in \mathbb{R}^3$ y $c, d \in \mathbb{R}^3$ dos pares de vectores no nulos l.i. Considere los planos $E = \langle a, b \rangle \subset \mathbb{R}^3$ $F = \langle c, d \rangle \subset \mathbb{R}^3$ generados por los pares de vectores a, b y c, d , respectivamente. Desarrolle un método que permita hallar un vector no nulo $v \in E \cap F \subset \mathbb{R}^3$ mediante el cálculo de las factorizaciones QR de tres matrices de 3×2 apropiadas.

donde el $\cos \phi$ superior izquierdo ocupa el lugar (i, i) , el $\cos \phi$ inferior derecho el lugar $(i + 1, i + 1)$, y el ángulo ϕ se elige de modo que $\operatorname{tg} \phi = h_{i+1,i}/h_{i,i}$ o, lo que es lo mismo, se hace $\sin \phi = ch_{i+1,i}$, $\cos \phi = ch_{i,i}$, y se elige c de modo que la suma de los cuadrados sea 1. Demuestre esta afirmación (c).

El producto de las matrices $S_i^{(k)T}$ constituye la matriz $Q^{(k)T}$ buscada:

$$Q^{(k)T} = S_{n-1}^{(k)T} \dots S_2^{(k)T} S_1^{(k)T}.$$

Evidentemente, el mismo argumento se aplica a cualquier etapa k .

(d) Programe el algoritmo QR en MATLAB y aplique su programa para obtener los valores propios de las matrices

$$H_U = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 0 & 1/4 & 1/5 & 1/6 \\ 0 & 0 & 1/6 & 1/7 \end{bmatrix}, \quad H_L = \begin{bmatrix} 4 & 1 & 1 & 1 \\ 1 & 4 & 1 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}.$$

5. Sea $w = e^{2\pi i/n}$ con $i^2 = -1$ y $n \in \mathbb{N}$ fijo. Considere la matriz de Vandermonde $F = n^{-1/2} [w^{k\ell}]_{k,l=0:(n-1)} \in M(n \times n, \mathbb{C})$. Verifique que la matriz F es simétrica (¿es Hermitiana?) y que satisface $F^{-1} = \overline{F} \equiv F^*$ y $F^4 = I$. Encuentre la factorización QR de la matriz F .

6. Sea $A \in M(n \times n, \mathbb{C})$ y QR la factorización QR de A . Obtenga las factorizaciones QR de los factores Q y R de A .

4. ORTOGONALIZACIÓN DE GRAM-SCHMIDT

1. [8, p. 62, exc. 8.1] Sea $A \in M(m \times n, \mathbb{C})$. Determine el número exacto de sumas, restas multiplicaciones y divisiones de punto flotante involucradas en el cálculo de la factorización $A = \hat{Q}\hat{R}$ mediante el algoritmo 8.1 de [8].

2. [8, p. 62, exc. 8.2] Desarrolle una función $[Q,R]=\text{mgs}(A)$ en MATLAB que compute una factorización QR reducida $A = QR$ de una matriz $A \in M(m \times n, \mathbb{C})$ con $m \geq n$ mediante el método de ortogonalización modificado de Gram-Schmidt. Las variables de salida deben ser, por supuesto, una matriz $Q \in M(m \times n, \mathbb{C})$ con columnas ortonormales, y una matriz triangular superior $R \in M(n \times n, \mathbb{C})$. No olvide considerar el caso en que el algoritmo de Gram-Schmidt pudiera interrumpirse (momentáneamente) antes de procesar todas las columnas de la matriz A .

3. *El algoritmo de ortogonalización de Gram-Schmidt y la factorización QR.* Considere una matriz invertible M de 3×3 (para simplificar) cuyas columnas denotaremos por \mathbf{a} , \mathbf{b} , y \mathbf{c} , respectivamente. Evidentemente, los vectores \mathbf{a} , \mathbf{b} , y \mathbf{c} , son linealmente independientes (¿Por qué?).

A partir de los vectores \mathbf{a} , \mathbf{b} , y \mathbf{c} , queremos construir tres vectores *ortogonales* entre sí \mathbf{A} , \mathbf{B} , y \mathbf{C} , que expandan o desplieguen el mismo espacio vectorial que los vectores originales \mathbf{a} , \mathbf{b} , y \mathbf{c} .

Partimos definiendo (de manera arbitraria) $\mathbf{A} := \mathbf{a}$. El vector \mathbf{B} se define como \mathbf{b} menos su proyección a lo largo del vector \mathbf{A} , i.e.:

$$\mathbf{B} = \mathbf{b} - \frac{\mathbf{A}^T \mathbf{b}}{\mathbf{A}^T \mathbf{A}} \mathbf{A}.$$

(a) Verifique que \mathbf{A} y \mathbf{B} son ortogonales.

El vector \mathbf{C} se define como \mathbf{c} menos su proyección en el subespacio vectorial desplegado por los vectores \mathbf{A} y \mathbf{B} , i.e.:

$$\mathbf{C} = \mathbf{c} - \frac{\mathbf{A}^T \mathbf{c}}{\mathbf{A}^T \mathbf{A}} \mathbf{A} - \frac{\mathbf{B}^T \mathbf{c}}{\mathbf{B}^T \mathbf{B}} \mathbf{B}.$$

(b) Verifique que los vectores \mathbf{A} , \mathbf{B} , y \mathbf{C} son mutuamente ortogonales.

Es claro que este procedimiento se puede extender a cualquier conjunto finito de vectores linealmente independientes o, equivalentemente, a matrices invertibles de cualquier dimensión finita. Este algoritmo se conoce con el nombre de *método de ortogonalización de Gram-Schmidt*.

Normalizando los vectores \mathbf{A} , \mathbf{B} , \mathbf{C} obtenemos los vectores:

$$\mathbf{q}_1 = \mathbf{A}/\|\mathbf{A}\|, \quad \mathbf{q}_2 = \mathbf{B}/\|\mathbf{B}\|, \quad \mathbf{q}_3 = \mathbf{C}/\|\mathbf{C}\|,$$

que podemos considerar como los vectores-columna de una matriz Q .

(c) ¿Es ortogonal la matriz Q ? (¿Cómo se definen las matrices ortogonales?).

Puesto que los vectores \mathbf{a} , \mathbf{b} , y \mathbf{c} , son combinaciones lineales de los vectores \mathbf{q}_1 , \mathbf{q}_2 , y \mathbf{q}_3 (y *vice versa*), existe una matriz R tal que $A = QR$.

(d) Verifique que R es una matriz triangular superior dada por

$$\begin{bmatrix} \mathbf{q}_1^T \mathbf{a} & \mathbf{q}_1^T \mathbf{b} & \mathbf{q}_1^T \mathbf{c} \\ 0 & \mathbf{q}_2^T \mathbf{b} & \mathbf{q}_2^T \mathbf{c} \\ 0 & 0 & \mathbf{q}_3^T \mathbf{c} \end{bmatrix}$$

(e) Escriba un programa para MATLAB que permita calcular la factorización QR de una matriz invertible A de cualquier dimensión finita.

(f) Estudie la generalización de la factorización QR de cualquier matriz A de $m \times n$ con n columnas linealmente independientes. Escriba el programa MATLAB correspondiente.

4. EL ALGORITMO DE GRAM-SCHMIDT. *Objetivo:* Factorizar una matriz $A = [a_1|a_2|\dots|a_n] \in M(m \times n, \mathbb{C})$, $n \leq m$ (factorización QR reducida), en la forma $A = \widehat{Q}\widehat{R}$, donde $\widehat{Q} = [q_1|q_2|\dots|q_n] \in M(m \times n, \mathbb{C})$ es una matriz con columnas ortonormales y $\widehat{R} = [r_{ij}] \in M(n \times n, \mathbb{C})$ es una matriz triangular superior:

$$\begin{aligned} [a_1|a_2|a_3|\dots|a_n] &= [q_1|q_2|q_3|\dots|q_n] \begin{bmatrix} r_{11} & r_{12} & r_{13} & \dots & r_{1n} \\ & r_{22} & r_{23} & \dots & r_{2n} \\ & & r_{33} & \dots & r_{3n} \\ & & & \ddots & \vdots \\ & & & & r_{nn} \end{bmatrix} \\ &= \left[r_{11}q_1 \mid r_{12}q_1 + r_{22}q_2 \mid r_{13}q_1 + r_{23}q_2 + r_{33}q_3 \mid \dots \mid r_{1n}q_1 + \dots + r_{nn}q_n \right] \end{aligned}$$

De aquí resulta:

$$a_j = r_{1j}q_1 + r_{2j}q_2 + \dots + r_{j-1,j}q_{j-1} + r_{j,j}q_j, \quad j = 1 : n, \tag{17}$$

de donde se obtiene inmediatamente:

$$r_{11} = \|a_1\|, \quad q_1 = a_1/r_{11},$$

y, en seguida, usando la ortonormalidad de las columnas q_j , resulta para $j = 2 : n$:

$$r_{ij} = q_i^* a_j \quad (i = 1 : j - 1), \quad v_j = a_j - \sum_{k=0}^{j-1} r_{kj}q_k, \quad r_{jj} = \|v_j\|, \quad q_j = v_j/r_{jj}.$$

Estos resultados se pueden resumir en el siguiente algoritmo:

Algorithm 1 Gram-Schmidt Clásico

```

for  $j = 1$  to  $n$  do
     $v_j = a_j = j$ -ésima columna de la matriz  $A$ 
    for  $i = 1$  to  $j - 1$  do
         $r_{ij} = q_i^* a_j$ 
         $v_j = v_j - r_{ij} q_i$ 
    end for
     $r_{jj} = (v_j^* v_j)^{1/2}$ 
     $q_j = v_j / r_{jj}$ 
end for

```

Tarea: Determine la complejidad del algoritmo clásico de Gram-Schmidt.

De (17) se observa que:

$$\langle a_1, a_2, \dots, a_{j-1}, a_j \rangle \subseteq \langle q_1, q_2, \dots, q_{j-1}, q_j \rangle, \quad j = 1 : n.$$

Si los a_k , $k = 1 : j$, son l.i., entonces evidentemente se tiene la igualdad. Si, en cambio, se tiene que los a_k son l.i. solamente para $k = 1 : j - 1$, y que los a_k , $k = 1 : j$ son l.d., entonces:

$$\langle a_1, a_2, \dots, a_{j-1}, a_j \rangle \subseteq \langle q_1, q_2, \dots, q_{j-1} \rangle. \quad (18)$$

En este caso no habría problemas para calcular los coeficientes r_{ij} , $i = 1 : j - 1$, pero se obtendría $r_{jj} = 0$, lo que impediría calcular q_j . En tal caso se podría definir, en principio, q_j como *cualquier* vector de \mathbb{C}^m ortonormal a $\langle q_1, q_2, \dots, q_{j-1} \rangle$. Pero entonces podría ocurrir que $q_j \in \langle a_1, a_2, \dots, a_{n-1}, a_n \rangle$, lo que en iteraciones subsecuentes conduciría a nuevos coeficientes $r_{kk} = 0$, lo que pudiere ser indeseable. Por esta razón, cuando ocurre (18) es preferible definir q_j como *cualquier vector unitario ortogonal a* $\langle a_1, a_2, \dots, a_{n-1}, a_n \rangle$. Es claro que esta situación ocurre precisamente cuando la matriz inicial A no es de rango completo.

Tarea: Construya ejemplos donde ocurra el fenómeno discutido en las líneas precedentes. Modifique convenientemente el algoritmo clásico de Gram-Schmidt de modo que pueda operar adecuadamente en la situación descrita.

El algoritmo de Gram-Schmidt clásico es inestable numéricamente en el sentido que es muy sensible a los efectos de los errores de redondeo. De consiguiente es necesario modificarlo; para hacerlo, expresaremos el algoritmo en términos de proyectores. Primeramente constatemos que:

$$\begin{aligned}
 r_{11}q_1 &= a_1 =: P_1 a_1, \quad \text{donde } P_1 = I, \\
 r_{22}q_2 &= a_2 - r_{12}q_1 = \text{proyección ortogonal de } a_2 \text{ sobre } \langle q_1 \rangle^\perp =: P_2 a_2, \\
 r_{33}q_3 &= a_3 - r_{13}q_1 - r_{23}q_2 = \text{proyección ortogonal de } a_3 \text{ sobre } \langle q_1, q_2 \rangle^\perp =: P_3 a_3, \\
 &\dots \quad \dots \quad \dots \\
 r_{nn}q_n &= a_n - r_{1n}q_1 - \dots - r_{n-1,n}q_{n-1} \\
 &= \text{proyección ortogonal de } a_n \text{ sobre } \langle q_1, \dots, q_{n-1} \rangle^\perp =: P_n a_n.
 \end{aligned}$$

Como se sabe (del capítulo relativo a proyectores), el proyector ortogonal sobre $\langle q_1, \dots, q_{j-1} \rangle$ viene dado por:

$$\widehat{Q}_{j-1} \widehat{Q}_{j-1}^* \in M(m \times m, \mathbb{C}), \quad \text{donde } \widehat{Q}_{j-1} = [q_1 | \dots | q_{j-1}] \in M(m \times n, \mathbb{C}),$$

de modo que el proyector ortogonal P_j sobre $\langle q_1, \dots, q_{n-1} \rangle^\perp$ viene dado por:

$$P_j = I - \widehat{Q}_{j-1} \widehat{Q}_{j-1}^* \in M(m \times m, \mathbb{C}).$$

Tarea: Verifique estas afirmaciones. Demuestre:

$$\begin{aligned} \text{rank}(\widehat{Q}_{j-1} \widehat{Q}_{j-1}^*) &= \dim \langle q_1, \dots, q_{j-1} \rangle = j - 1, \\ \text{rank}(P_j) &= \text{rank}(I - \widehat{Q}_{j-1} \widehat{Q}_{j-1}^*) = \dim \langle q_1, \dots, q_{j-1} \rangle^\perp = m - j + 1. \end{aligned}$$

Tarea: Demuestre que el proyector ortogonal P_j se puede escribir en la forma:

$$P_j \equiv P_{\langle q_1, \dots, q_{n-1} \rangle^\perp} = P_{q_{j-1}^\perp} P_{q_{j-2}^\perp} \dots P_{q_2^\perp} P_{q_1^\perp}.$$

Usando estos proyectores el algoritmo clásico de Gram-Schmidt genera la siguiente sucesión de vectores v_j :

$$\begin{aligned} v_1 &= P_1 a_1 = I a_1 = a_1, \\ v_2 &= P_{q_1^\perp} a_2, \\ v_3 &= P_{q_2^\perp} P_{q_1^\perp} a_3, \\ v_4 &= P_{q_3^\perp} P_{q_2^\perp} P_{q_1^\perp} a_4, \\ &\dots \quad \dots \quad \dots \\ v_n &= P_{q_{n-1}^\perp} \dots P_{q_2^\perp} P_{q_1^\perp} a_n, \end{aligned}$$

Examinando la sucesión de cálculos precedentes se observa que las sucesivas operaciones se pueden organizar del siguiente modo:

$$\begin{array}{ccccccc} v_1 & \leftarrow & a_1 & & & & \\ v_2 & \leftarrow & a_2 & \quad v_2 & \leftarrow & P_{q_1^\perp} v_2 & \\ v_3 & \leftarrow & a_3 & \quad v_3 & \leftarrow & P_{q_1^\perp} v_3 & \quad v_3 & \leftarrow & P_{q_2^\perp} v_3 \\ v_4 & \leftarrow & a_4 & \quad v_4 & \leftarrow & P_{q_1^\perp} v_4 & \quad v_4 & \leftarrow & P_{q_2^\perp} v_4 & \quad v_4 & \leftarrow & P_{q_3^\perp} v_4 \\ \dots & & \dots & & \dots & & \dots & & \dots & & \dots & & \\ v_n & \leftarrow & a_n & \quad v_n & \leftarrow & P_{q_1^\perp} v_n & \quad v_n & \leftarrow & P_{q_2^\perp} v_n & \quad v_n & \leftarrow & P_{q_3^\perp} v_n & \quad \dots & v_n & \leftarrow & P_{q_{n-1}^\perp} v_n \end{array}$$

donde el orden de la operaciones es *top-down* y *left-right*. Se constata en la práctica que este algoritmo es estable. Recurriendo nuevamente a un pseudo-lenguaje de programación, el algoritmo de Gram-Schmidt modificado queda:

Algorithm 2 Gram-Schmidt Modificado

```

for  $i = 1$  to  $n$  do
     $v_i = a_i = i$ -ésima columna de la matriz  $A$ 
end for
for  $i = 1$  to  $n$  do
     $r_{ii} = (v_i^* v_i)^{1/2}$ 
     $q_i = v_i / r_{ii}$ 
    for  $j = i + 1$  to  $n$  do
         $r_{ij} = q_i^* v_j$ 
         $v_j = v_j - r_{ij} q_i$ 
    end for
end for

```

Tarea: Determine la complejidad del algoritmo modificado de Gram-Schmidt.

Tarea: Modifique convenientemente el algoritmo *modificado* de Gram-Schmidt de modo que pueda manejar los casos en que hay coeficientes diagonales nulos.

5. EL ALGORITMO DE GRAM-SCHMIDT (MODIFICADO) COMO UNA ORTOGONALIZACIÓN TRIANGULAR. Sea $A = [a_1 | a_2 | \dots | a_n] \in M(m \times n, \mathbb{C})$, $m > n$, una matriz de rango completo. Es fácil ver que el algoritmo de Gram-Schmidt modificado puede describirse mediante una sucesión de transformaciones matriciales:

$$\begin{aligned} [v_1^{(0)} | v_2^{(0)} | v_3^{(0)} | \dots | v_n^{(0)}] &\leftarrow [a_1 | a_2 | a_3 | \dots | a_n], \\ [q_1 | v_2^{(1)} | v_3^{(1)} | \dots | v_n^{(1)}] &\leftarrow [v_1^{(0)} | v_2^{(0)} | v_3^{(0)} | \dots | v_n^{(0)}] \cdot R_1, \\ [q_1 | q_2 | v_3^{(2)} | \dots | v_n^{(2)}] &\leftarrow [q_1 | v_2^{(1)} | v_3^{(1)} | \dots | v_n^{(1)}] \cdot R_2, \\ &\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ [q_1 | \dots | q_{j-1} | q_j | v_{j+1}^{(j)} | \dots | v_n^{(j)}] &\leftarrow [q_1 | \dots | q_{j-1} | v_j^{(j-1)} | v_{j+1}^{(j-1)} | \dots | v_n^{(j-1)}] \cdot R_j, \\ &\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ [q_1 | \dots | q_n] &\leftarrow [q_1 | \dots | q_{n-1} | v_n^{(n-1)}] \cdot R_n. \end{aligned}$$

Las matrices $R_j \in M(n \times n, \mathbb{C})$ son matrices de la forma:

$$R_j = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & -\frac{r_{j,j+1}}{r_{jj}} & \dots & -\frac{r_{jn}}{r_{jj}} \\ & & \frac{1}{r_{jj}} & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} = I + Z_j,$$

donde $I \in M(n \times n, \mathbb{C})$ es la matriz identidad y $Z_j \in M(n \times n, \mathbb{C})$ es una matriz cuyos coeficientes $z_{ik}^{(j)}$ son todos nulos excepto los de la fila j a partir de la columna j :

$$z_{ik}^{(j)} = \begin{cases} 0, & \text{si } (i \neq j) \vee (i = j \wedge k < j), \\ -1 + 1/r_{jj}, & \text{si } i = j = k, \\ -r_{j,k}/r_{jj}, & \text{si } i = j \wedge k > j. \end{cases}$$

Notar que $r_{jj} \neq 0$ para todo $j = 1 : n$ pues A es de rango maximal. Evidentemente, $\det(R_j) = 1/r_{jj} \neq 0$, de modo que las matrices R_j son invertibles. Se les denomina matrices de Frobenius.

(a) Calcule R_j^{-1} .

Al concluir el algoritmo de Gram-Schmidt modificado se tiene:

$$[q_1 | q_2 | q_3 | \dots | q_n] = [a_1 | a_2 | a_3 | \dots | a_n] R_1 R_2 \dots R_n. \tag{19}$$

De (19) se deduce:

$$\underbrace{[a_1 | a_2 | a_3 | \dots | a_n]}_{=A} = \underbrace{[q_1 | q_2 | q_3 | \dots | q_n]}_{=\hat{Q}} \underbrace{R_n^{-1} \dots R_2^{-1} R_1^{-1}}_{=\hat{R}} \tag{20}$$

(b) Determine R en términos de los coeficientes $r_{ik}^{(j)}$ de las matrices R_j . Demuestre que R es una matriz triangular superior de tal modo que (20) representa la factorización QR reducida $\widehat{Q}\widehat{R}$ de la matriz A .

6. Extienda los algoritmos clásico y modificado de Gram-Schmidt a las siguientes situaciones:

(a) Matrices $A \in M(m \times n, \mathbb{C})$ con $m < n$.

(b) Matrices $A \in M(\infty \times n, \mathbb{C})$.

7. [7, §2.4, p. 28] *Polinomios de Jacobi*. Obtenga (unos cuantos de) los polinomios de Jacobi $P_n^{(\alpha, \beta)}(x)$, $\alpha, \beta > -1$, $n \in \mathbb{N}_0$, $-1 \leq x \leq 1$, aplicando el método de ortonormalización de Gram-Schmidt a los polinomios $a_k(x) := x^k$, $k \in \mathbb{N}_0$, sobre el intervalo $[-1, 1]$, con respecto al producto interno:

$$(f, g) := \int_{-1}^1 \overline{f(x)} g(x) (1-x)^\alpha (1+x)^\beta dx, \quad \alpha, \beta > -1.$$

(a) Analíticamente (i.e., s.c.P.L.P.).

(b) Computacionalmente. Para el cálculo de las integrales use una regla de Simpson apropiada.

Casos particulares de los polinomios de Jacobi $P_n^{(\alpha, \beta)}(x)$ son:

(i) los *polinomios de Legendre* $P_n(x) = P_n^{(0,0)}(x)$,

(ii) los *polinomios ultrasféricos* $S_n^{(\alpha)}(x) = P_n^{(\alpha, \alpha)}(x)$,

(iii) los *polinomios de Chebyshev de primera especie* $T_n(x) = P_n^{(-\frac{1}{2}, -\frac{1}{2})}(x) = \cos(n \arccos x)$,

(iv) los *polinomios de Chebyshev de segunda especie* $U_n(x) = P_n^{(\frac{1}{2}, \frac{1}{2})}(x) = \frac{\sin((n+1) \arccos x)}{\sin(\arccos x)}$.

También obtenga (unos cuantos de) estos polinomios y demuestre las fórmulas indicadas para los polinomios de Chebyshev.

8. [7, §2.4, p. 28] *Polinomios de Laguerre*. Obtenga los polinomios de Laguerre $L_n^{(\alpha)}(x)$, $\alpha > -1$, $n \in \mathbb{N}_0$, $x \geq 0$, aplicando el método de ortonormalización de Gram-Schmidt a los polinomios $a_k(x) := x^k$, $k \in \mathbb{N}_0$, sobre el intervalo $[0, \infty]$, con respecto al producto interno:

$$(f, g) := \int_0^\infty \overline{f(x)} g(x) e^{-x} x^\alpha dx, \quad \alpha > -1.$$

(a) Analíticamente (i.e., s.c.P.L.P.).

(b) Computacionalmente. Para el cálculo de las integrales use una regla de Simpson apropiada.

9. [7, §2.4, p. 28] *Polinomios de l'Hermite*. Obtenga los polinomios de l'Hermite $H_n(x)$, $n \in \mathbb{N}_0$, $x \in \mathbb{R}$, aplicando el método de ortonormalización de Gram-Schmidt a los polinomios $a_k(x) := x^k$, $k \in \mathbb{N}_0$, sobre \mathbb{R} , con respecto al producto interno:

$$(f, g) := \int_{-\infty}^\infty \overline{f(x)} g(x) e^{-x^2} dx.$$

(a) Analíticamente (i.e., s.c.P.L.P.).

(b) Computacionalmente. Para el cálculo de las integrales use una regla de Simpson apropiada.

5. TRIANGULARIZACIÓN DE HOUSEHOLDER

1. [3, p. 271, exc. 6.3.6] Sea $P = I - 2ww^T \in M(n \times n, \mathbb{R})$, $w \in \mathbb{R}^n$. Demuestre:

- (i) P es simétrica y ortogonal si y sólo si $ww^T = 1$.
- (ii) Si $ww^T = 1$, entonces $Pw = -w$.
- (iii) Si $u^T w = 0$, entonces $Pu = 0$.

Interprete geoméricamente los resultados precedentes en dos y tres dimensiones, como reflexiones con respecto a planos apropiados.

2. [8, p. 76, exc. 10.1], [3, p. 271, exc. 6.3.7] Obtenga los valores y vectores propios, determinantes y valores singulares de los reflector de Householder.

Más generalmente, ¿qué puede Ud. decir acerca de los valores y vectores propios, determinantes y valores singulares de las matrices de la forma $I + uv^T$ para vectores dados $u, v \in \mathbb{R}^n$.

- 3.** [8, p. 76, exc. 10.2] (a) Escriba una función de MATLAB `[W,R]=house(A)` que compute una representación implícita de una factorización QR total $A = QR$ de una matriz $A \in M(m \times n, \mathbb{C})$, con $m \geq n$, usando reflexiones de Householder. Las variables de salida deben ser: i) una matriz triangular inferior $W \in M(m \times n, \mathbb{C})$ cuyas columnas son los vectores v_k que definen las sucesivas reflexiones de Householder, y ii) una matriz triangular $R \in M(n \times n, \mathbb{C})$.
 (b) Escriba una función de MATLAB `Q=formQ(W)` que toma como entrada la matriz W producida por `house` y genera la correspondiente matriz ortogonal $Q \in M(m \times m, \mathbb{C})$.

- 4.** [8, p. 76, exc. 10.3] Sea $Z = \begin{bmatrix} 1 & 4 & 7 & 4 & 4 \\ 2 & 5 & 8 & 2 & 2 \\ 3 & 6 & 7 & 3 & 2 \end{bmatrix}^T$. Compute la factorización QR reducida

mediante: (a) s.c.P.L.P.,

(b) la rutina de Gram-Schmidt `mgs` que Ud. desarrolló en el capítulo dedicado a la ortogonalización de Gram-Schmidt.

(c) las rutinas de Householder `[W,R]=house(A)` y `Q=formQ(W)` del ejercicio (3),

(d) el comando `[Q,R]=qr(Z,0)` de MATLAB.

Compare los tres últimos procedimientos y comente las diferencias que observe.

- 5.** [8, p. 76, exc. 10.4] Considere las matrices ortogonales de 2×2 :

$$F = \begin{bmatrix} -\cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad J = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}.$$

(a) Observe (i.e., ¡demuestre!) que F es un reflector (¿con respecto a qué plano?) con $\det F = -1$, i.e., corresponde al caso especial de un reflector de Householder en dimensión 2.

(b) Observe (i.e., ¡demuestre!) que J es una rotación (¿con respecto a qué eje?) con $\det J = 1$. Una tal matriz se denomina *rotación de Givens*.

(c) Describa exactamente qué efectos geométricos tiene la multiplicación por F y J por la izquierda en el plano \mathbb{R}^2 . Ciertamente J rota el plano en un ángulo de θ radianes, pero la rotación es “contra-reloj” o a “favor-de-reloj” (antes de la era digital, por supuesto).

(d) Como el lector sabe, el algoritmo de factorización QR mediante reflectores de Householder es el siguiente:

for $k = 1$ to n

$$x = A_{k:m,k}$$

$$v_k = \text{sgn}(x_1) \|x\|_2 e_1 + x, \quad \text{donde } e_1 = [1, 0, \dots, 0]^T \in \mathbb{C}^{m-k+1}$$

$$v_k = v_k / \|v_k\|_2$$

$$A_{k:m,k:n} = A_{k:m,k:n} - 2v_k (v_k^* A_{k:m,k:n})$$

Demuestre que la complejidad de este algoritmos es del orden de $2mn^2 - \frac{2}{3}n^3$ “flops”.

- (e) Diseñe un algoritmo de factorización QR análogo al precedente pero con base en unas rotaciones de Givens en lugar de reflexiones de Householder.
- (f) Demuestre que su algoritmo requiere seis “flops” por celda en lugar de cuatro, de modo que la cantidad asintótica de operaciones es 50 % mayor que la determinada en (d).
- (g) Programe en MATLAB todos los algoritmos mencionados en este ejercicio, Verifique su correcta operación.

6. [4, p. 77, exc. 4.] Demuestre que todo vector $x \in \mathbb{R}^n$ de largo Euclideo $r = (x^T x)^{1/2}$ puede ser transformado en cualquier otro vector $y \in \mathbb{R}^n$ de largo r , $y \neq x$, mediante una transformación de Householder. Estudie también el caso en que $x, y \in \mathbb{C}^n$.

7. [4, p. 78, exc. 5.] *Obtención a la forma de Hessenberg mediante la transformación de Householder.* El método de Householder para calcular valores propios de $A \in M(n \times n, \mathbb{R})$, como el método de Givens, primeramente reduce A a la forma de Hessenberg superior (o a la forma tridiagonal en el caso simétrico). Demuestre constructivamente que una matriz de la forma:

$$A = [a_{ij}]_{n \times n} = \begin{bmatrix} * & * & * & \dots & * & * & * & \dots & * \\ * & * & * & \dots & * & * & * & \dots & * \\ 0 & * & * & \dots & * & * & * & \dots & * \\ 0 & 0 & * & \dots & * & * & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & * & * & * & \dots & * \\ 0 & 0 & 0 & \dots & 0 & * & * & \dots & * \\ \hline 0 & 0 & 0 & \dots & 0 & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & * & \dots & * \end{bmatrix}, \quad \text{con } a_{ij} = 0 \begin{cases} \forall i = (j + 2) : n, \\ \forall j = 1 : k \leq n, \end{cases}$$

con 0's debajo de la celda $(j + 1)$ de la columna j , $j = 1 : k$, puede ser transformada en una matriz de la misma forma, ahora con $j = 1 : (k + 1)$, por medio de una sola *transformación de similitud* real ortogonal usando una transformación de Householder. Concluya, entonces, que toda matriz $A \in M(n \times n, \mathbb{R})$ se puede reducir a la forma de Hessenberg superior mediante una sucesión de $(n - 2)$ similitudes de Householder y que, del mismo modo, una matriz $A \in M(n \times n, \mathbb{R})$ simétrica se puede reducir a la forma tridiagonal.

Hint: Para la $(k + 1)$ -ésima columna elija una transformación de Householder $U \in M((n - k - 1) \times (n - k - 1), \mathbb{R})$ que transforma las $(n - k - 1)$ celdas por debajo de la diagonal en un múltiplo apropiado de $[1, 0, \dots, 0]^T \in \mathbb{R}^{n-k-1}$. Después, aplique a la matriz completa la transformación de similitud determinada por la matriz ortogonal $\begin{bmatrix} I & 0 \\ 0 & U \end{bmatrix} \in M(n \times n, \mathbb{R})$ y verifique que el patrón de ceros deseado se mantiene.

8. [5, p. 175, exc. 12] Demuestre que el método de Householder reduce toda matriz real antisimétrica A (i.e., $A = -A^T$) a una matriz real antisimétrica *tridiagonal*. Examine la evolución numérica de este algoritmo al aplicarlo a la matriz *circulante*:

$$C = \begin{bmatrix} 0 & 1 & 1 & 0 & -1 & -1 \\ -1 & 0 & 1 & 1 & 0 & -1 \\ -1 & -1 & 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 & 1 \\ 1 & 0 & -1 & -1 & 0 & 1 \\ 1 & 1 & 0 & -1 & -1 & 0 \end{bmatrix}$$

REFERENCIAS

- [1] A. Basilevsky. *Applied Matrix Algebra in the Statistical Sciences*. North-Holland, New York-Amsterdam--Oxford, 1983.
- [2] G.H. Golub and C. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore-London, 1996 (Third Edition). This encyclopedic work is the must in the field!
- [3] G.H. Golub and J.M. Ortega. *Scientific Computing: An introduction with Parallel Computing*. Academic Press, Boston, 1993.
- [4] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985 (reprinted 1994).
- [5] E. Isaacson and H.B. Keller. *Analysis of Numerical Methods*. John Wiley and Sons, New York, 1966. There is a more recent edition by Dover, New York, 1994.
- [6] J.G. Nagy and D.P. O'Leary. Image deblurring: I can see clearly now. *Computing in Science and Engineering*, 5(3):82–84, May-June 2003. Copublished by the IEEE Computer Society and the American Institute of Physics.
- [7] G. Szegő. *Orthogonal Polynomials*, volume XXIII. American Mathematical Society, Colloquium Publications, New York, 1939. One of the best of all classics!
- [8] L.N. Trefethen and D. Bau III. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997. A very good book indeed! More advanced than Strang's book.

LSC/lsc, Valparaíso, 13 de mayo de 2007