4 Descomposició en Valors Singulars (SVD)

La descomposició QR d'una matriu és un mètode comú per a resoldre sistemes sobredeterminats i per trobar solucions a problemes de mínims quadrats, però no és pas l'únic. Un altre molt comú, i amb gran quantitat d'aplicacions, és l'anomenada *Descomposició en Valors Singulars* d'una matriu o, en anglès, *Singular Value Decomposition* (SVD).

Teorema 4.1 (SVD [2]). *Sigui $A \in \mathcal{M}_{m,n}(\mathbb{R})$, amb $m \geq n$. Aleshores, existeix una descomposició de la matriu A de la forma*

$$A = U\Sigma V^\top \tag{1}$$

satisfent les següents propietats:

- *$U \in \mathcal{M}_{m,n}(\mathbb{R})$ és matriu ortogonal, $U^\top U = I_n$. Els vectors-columna que formen U, és a dir $U = (u_1 \ u_2 \ \cdots \ u_n)$, reben el nom de* vector singulars per l'esquerra.
- *$V \in \mathcal{M}_{n,n}(\mathbb{R})$ és també ortogonal, $V^\top V = I_n$. Anàlogament al punt anterior, els vectors-columna que formen V, o sigui $V = (v_1 \ v_2 \cdots \ v_n)$, s'anomenen* vectors singulars per la dreta.
- *$\Sigma \in \mathcal{M}_{n,n}(\mathbb{R}) = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ amb $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$, anomenats* valors singulars *de la matriu A. El valor σ_1 es diu* valor singular principal.

Comentari 4.2.

1. *En el cas $A \in \mathcal{M}_{m,n}(\mathbb{R})$ amb $m < n$ el mateix teorema s'aplica. És trivial comprovar que la descomposició SVD d'$A^\top$ és $A^\top = V\Sigma^\top U^\top$.*
2. *Per a matrius complexes el teorema és també cert canviant "trasposada" per "Hermítica": $A = U\Sigma V^*$ amb $U^*U = V^*V = I_n$.*

Demostració del Teorema 4.1. Definim la matriu simètrica $S := A^\top A \in \mathcal{M}_{n,n}(\mathbb{R})$. Pel teorema espectral, existeix una base ortonormal de $\mathbb{R}^n$ que consisteix en els VEPs de S i, a més, els VAPs corresponents a cada vector són reals; i.e., existeixen n vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ t.q.:

$$Sv_i = \lambda_i v_i,$$

amb $\mathbb{R}^n = [v_1, v_2, \ldots, v_n]$ i $\lambda_i \in \mathbb{R}$, $\langle v_i, v_j \rangle = \delta_{i,j}$ per tot $i, j = 1 \div n$. En aquest cas, de fet, $\lambda_i \geq 0$ per tot $i = 1 \div n$, ja que S és semidefinida positiva. En efecte, sigui $\xi \in \mathbb{R}^n$, $\xi \neq 0$, llavors

$$\langle \xi, S\xi \rangle = \langle \xi, A^\top A\xi \rangle = \langle A\xi, A\xi \rangle = \|A\xi\|_2^2 \geq 0.$$

Sense pèrdua de generalitat, podem suposar que els vectors de la base ortonormal vénen ordenats de manera que els vaps respectius són decreixents, i.e.: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Sigui $1 \leq r \leq n$

[0]T. Lázaro, M. Ollé i J.R. Pacha. Departament Matemàtica Aplicada I

t.q.: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$, $\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$ (notem que pot ser $r = n$). A continuació definim

$$\sigma_1 := \sqrt{\lambda_1}, \sigma_2 := \sqrt{\lambda_2}, \ldots, \sigma_r := \sqrt{\lambda_r} \tag{2}$$

de manera que: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. Introduïm també els vectors

$$u_1 = \frac{1}{\sigma_1} A v_1, u_2 = \frac{1}{\sigma_2} A v_2, \ldots, u_r = \frac{1}{\sigma_r} A v_r.$$

Aquest vectors són ortogonals dos a dos i unitaris. Comprovem-ho:

$$\langle u_i, u_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A v_i, A v_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle v_i, A^\top A v_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle v_i, S v_j \rangle = \frac{\lambda_j}{\sigma_i \sigma_j} \langle v_i, v_j \rangle = \frac{\sigma_j}{\sigma_i} \langle v_i, v_j \rangle = \delta_{ij}$$

per $i, j = 1 \div r$, on hem fet servir que, d'acord amb (2) és $\lambda_j = \sigma_j^2$, per $j = 1 \div r$. A continuació, completem els vectors $u_1, u_2, \ldots, u_r \in \mathbb{R}^m$ amb $n - r$ vectors $u_{r+1}, u_{r+2}, \ldots, u_n \in \mathbb{R}^n$ de manera que $\{u_1, u_2, \ldots, u_n\}$ formin una família ortonormal a $\mathbb{R}^m$. Ara, en la base ortonormal $\{v_1, v_2, \ldots, v_n\}$ de $\mathbb{R}^n$ i per a la família ortonormal $\{u_1, u_2, \ldots, u_n\}$ de $\mathbb{R}^m$, tenim

$$\begin{aligned} A v_j &= \sigma_j \frac{1}{\sigma_j} A v_j = \sigma_j u_j, & &\text{per } j = 1 \div r, \\ A v_j &= 0, & &\text{per } j = r + 1 \div n \end{aligned} \tag{3}$$

d'acord amb la definició de u_j, $j = 1 \div r$. Ara, si escrivim $U = \begin{pmatrix} u_1 \ u_2 \cdots u_n \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R})$ com la matriu que té per columnes els vectors $u_1, u_2, \ldots, u_n$; $V = \begin{pmatrix} v_1 \ v_2 \cdots v_n \end{pmatrix} \in \mathcal{M}_{n,n}(\mathbb{R})$ com la matriu que té per columnes els vectors $v_1, v_2, \ldots, v_n$ i $\Sigma = \mathrm{diag}[\sigma_1, \sigma_2, \ldots, \sigma_r, 0, \ldots, 0] \in \mathcal{M}_{n,n}(\mathbb{R})$; podem expressar (3) en forma matricial com

$$AV = U\Sigma. \tag{4}$$

Finalment, com que les columnes de V són ortogonals, tenim que $V^\top V = V V^\top = I_n$, i multiplicant (4) per la dreta a totes dues bandes per $V^\top$ s'obté la descomposició buscada. □

La descomposició SVD d'una matriu satisfà moltes propietats importants. Al següent Teorema donem les més destacades.

Teorema 4.3 ([2]). *Sigui $A = U\Sigma V^\top$ la descomposició SVD de la matriu $A \in \mathcal{M}_{m,n}(\mathbb{R})$, amb $m \geq n$. Aleshores:*

1. *Suposeu que A és simètrica i que, per tant, admet una descomposició de la forma $A = UDU^\top$ amb $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, $U = (u_1 \ u_2 \ \cdots \ u_n)$ ortogonal, és a dir $UU^\top = I_n$. Aleshores una SVD d'A és $A = U\Sigma V^\top$ amb $\sigma_i = |\lambda_i|$ i $v_i = \mathrm{sign}(\lambda_i) u_i$, on sign és la funció "signe" i $\mathrm{sign}(0) = 1$.*

2. *Els vaps de la matriu simètrica $A^\top A \in \mathcal{M}_{n,n}(\mathbb{R})$ són σ_i^2. Els vectors singulars per la dreta $\{v_1, v_2, \ldots, v_n\}$ formen una base ortogonormal de veps.*

3. *Els vaps de la matriu simètrica* $AA^\top \in \mathcal{M}_{m,m}(\mathbb{R})$ *són* σ_i^2, $i = 1, 2, \ldots, n$, *i* $m - n$ *zeros. Els vectors singulars per l'esquerra* u_i *són els vep's ortonormals corresponents als vaps* σ_i^2.

4. *Siguin* $A \in \mathcal{M}_{n-n}(\mathbb{R})$ *una matriu quadrada i* $A = U\Sigma V^\top$ *la seva SVD amb* $\Sigma = \operatorname{diag}(\sigma_1, \cdots, \sigma_n)$, $U = (u_1\ u_2\ \cdots\ u_n)$ *i* $V = (v_1\ v_2\ \cdots\ v_n)$. *Considereu la matriu per blocs*

$$H = \begin{pmatrix} 0 & A^\top \\ A & 0 \end{pmatrix}.$$

Aleshores, els vaps d'H *són* $\{\pm\sigma_1, \pm\sigma_2, \ldots, \pm\sigma_n\}$ *i els seus veps unitaris corresponents són*

$$\frac{1}{\sqrt{2}} \begin{pmatrix} v_i \\ \pm u_i \end{pmatrix}.$$

5. *Si* A *té rang màxim, és a dir,* $\operatorname{rang}(A) = n$, *llavors*

$$\min_x \|Ax - b\| = \|Ax^* - b\|$$

amb $x^* = V\Sigma^{-1}U^\top b$. *Per tant, proporciona la solució del problema de mínims quadrats.*

6. *La norma euclídea d'*A *és* $\|A\|_2 = \sigma_1$. *Si* $A \in \mathcal{M}_{n,n}(\mathbb{R})$ *és no singular aleshores* $\|A^{-1}\|_2 = 1/\sigma_n$ *i el nombre de condició* $\mu_2(A) = \|A\|_2\|A^{-1}\|_2 = \sigma_1/\sigma_n$.

7. *Suposem* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$. *LLavors,* $\operatorname{rang}(A) = r$ *i*

$$\ker A = \operatorname{span}\{v_{r+1}, v_{r+2}, \ldots, v_n\}, \qquad \operatorname{Im} A = \operatorname{span}\{u_1, u_2, \ldots, u_r\}.$$

8. *Siguin* $S_n = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ *l'esfera unitat* n*-dimensional i* $A(S_n) = \{Ax : x \in S_n\}$ *la seva imatge per la matriu* $A \in \mathcal{M}_{m,n}(\mathbb{R})$. *Aleshores,* $A(S_n)$ *és un el·lipsoide de centre l'origen d*$\mathbb{R}^m$ *i eixos principals* $\sigma_i u_i$.

Demostració. Cadascun dels apartats de la proposició es proven a partir del Teorema 4.1. Deixem doncs la demostració d'aquest teorema com exercici (o bé podeu consultar la referència [2]). *Nota:* aquesta activitat és voluntària i *no* cal que la presenteu a l'informe de la pràctica. □

Exercici 4.4. *(Obligatori, cal que figuri a l'informe). Torneu a fer l'exercici 2 de la 2ª pràctica (extrapolació de la població a Catalunya) aplicant —en fer l'aproximaxió per mínims quadrats—, la propietat 5 del Teorema 4.3 (en comptes de la descomposició QR). Compareu els resultats.*

Una conseqüència molt interessant del Teorema 4.1 és que proporciona un mètode d'aproximació de la matriu A per matrius de menor rang. De fet,

Proposition 4.5 ([2]). *Considereu la descomposició SVD d'una matriu $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $A = U\Sigma V^\top$ i escriviu $U = [u_1\ u_2\ \dots u_n]$ i $V = [v_1\ v_2\ \dots v_n]$, on cada u_j, v_j és un vector columna d'm components. Observeu que podem escriure, equivalentment, que*

$$A = U\Sigma V^\top = \sum_{i=1}^{n} \sigma_i u_i v_i^\top$$

essent $\{\sigma_1, \sigma_2, \dots, \sigma_r, \sigma_{r+1} = 0, \dots, \sigma_n = 0\}$ els valors singulars d'A. Aleshores,

$$A_k := \sum_{i=1}^{k} \sigma_i u_i v_i^\top = U\Sigma_k V^\top,$$

amb $\Sigma_k = \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_k, 0, \dots, 0)$, és la matriu de rang k que millor aproxima A en norma euclídea. A més, $\|A - A_k\|_2 = \sigma_{k+1}$.

Demostració. Es comprova fàcilment que $\|A - A_k\|_2 = \sigma_{k+1}$. Sigui $B \in \mathcal{M}_{m,n}(\mathbb{R})$ de rang $k < n$ i, per tant, amb $\dim \operatorname{Nuc} B = n - k$. Considerem el subespai $F = [v_1, v_2, \dots, v_{k+1}]$. Aquests subespais, F i $\operatorname{Nuc} B$, es solapen, ja que la suma de les seves dimensions $(k+1) + (n-k) > n$. Sigui ξ un vector de la seva intersecció, que agafarem unitari; és a dir, t.q.: $\|\xi\|_2 = 1$. Llavors,

$$\|A - B\|_2^2 \geq \|(A - B)\xi\|_2^2 = \|A\xi\|_2^2 = \|U\Sigma V^\top \xi\|_2^2 = \|\Sigma V^\top \xi\|_2^2 \geq \sigma_{k+1}^2 \|V^\top \xi\|_2^2 = \sigma_{k+1}^2.$$

La qual cosa completa la prova. $\square$

Una aplicació divertida a la vegada que pràctica d'aquest darrer resultat és la compresió d'imatges. L'il·lustrarem amb un exemple complet. Suposeu que teniu una fotografia en format `jpg` com aquesta que anomenem `camell.jpg`. El primer que farem serà convertir-la a escala de grisos i

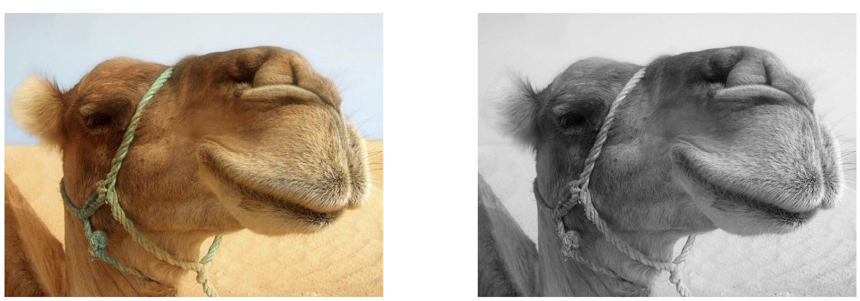

Figura 1: Retrat d'un camell (color i blanc i negre).

després transformar-la en una matriu en doble precisió segons la intensitat de cada píxel o grup de píxels. Malgrat es pot fer amb programes de tractament d'imatge de lliure distribució (per exemple, l'`opencv`) nosaltres usarem `Matlab`. Les comandes són aquestes:

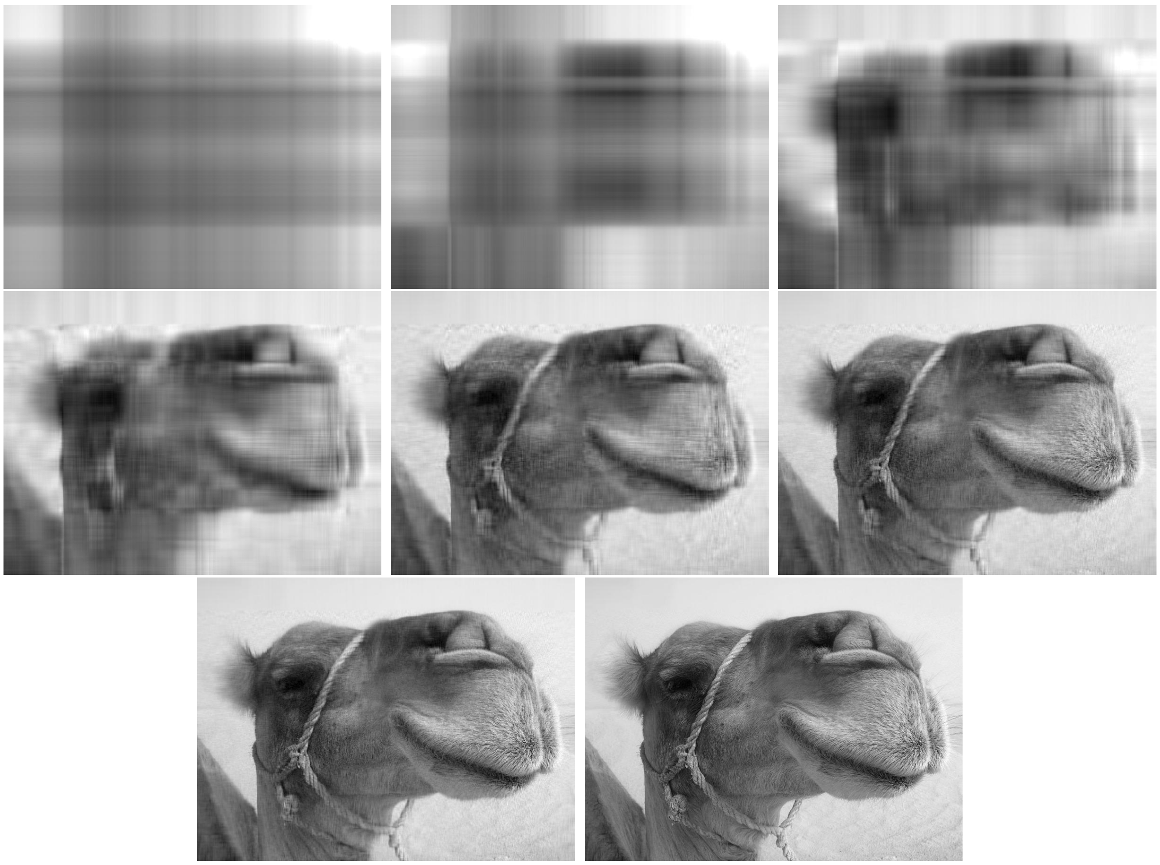

Figura 2: De dalt a baix i d'esquerra a dreta: fotografies corresponents a les diferents aproximacions usant els primers $1, 2, 5, 10, 25, 50, 100, 298$ valors singulars d'A, respectivament.

```
A = imread('camel.jpg');
A = rgb2gray(A);
A = im2double(A);
imshow(A)
dlmwrite('camel_matrix.dat', A, 'delimiter', '\t', 'precision', 4)
size(A)
```

La comanda `imshow(A)` mostra la imatge de nou mentre que `dlmwrite` escriu la matriu d'intensitats A en un fitxer ascii anomenat `camel_matrix.dat` de 600 files i 800 columnes separades per una tabulació. Si tornem a anomenar A la matriu (numèrica, d'intensitats) guardada al fitxer `camel_matrix.dat` i l'apliquem la SVD , obtenim les matrius que l'aproximen $A_1, A_2, A_5, A_{10}, A_{25}, A_{100}$

i A_{298}, obtingudes considerant els $1, 2, \ldots, 298$ primers valors no singulars d'A, respectivament. Si volem recuperar la imatge que correspon a cada una d'aquestes aproximacions, farem a MATLAB:

```
imwrite(Ak, sprintf('camel-%d.jpg', k));
```

on $k = 1, 2, \ldots, 298$. El resultat el podeu veure a la Figura 2.

Observeu que ja amb només 100 valors singulars (penúltima fotografia a Figura 2) la resolució és força raonable.

Exercici 4.6. *Feu el mateix amb una fotografia de vosaltres (és a dir, dels dos components de cada grup). És important que preneu la fotografia a una resolució baixa per a evitar que la matriu surti massa gran. L'exemple del camell está fet a una resolució de* 600×800.

5 Presentació d'informes

Els informes s'hauran de lliurar en format PDF a la Intranet (Campus Digital Atenea), a l'espai que, al llarg de la setmana del 17 al 22 de desembre i amb data límit de divendres 22 a les 24 hores estarà actiu per pujar fitxers.

Cada informe haurà de contenir la resolució comentada dels exercicis 4.4 i 4.6 plantejats en aquest guió, una descripció detallada de l'algorisme triat per programar la descomposició SVD, així com la bibliografia (llibres, articles, pàgines güeb,...) que heu consultat.

Referències

[1] G. Dahlquist, A. Björck. *Numerical methods in Scientific Computing, vol 1, 2.* SIAM, Philadelphia, 2008.

[2] J. Demmel *Applied Numerical Linear Algebra*. SIAM, Philadelphia, 1997. 1, 2, 3, 4

[3] Wikipedia: `http://en.wikipedia.org/wiki/Singular_value_decomposition`