13. Sea $A \in \mathbb{R}^{m\times n}$ y $A = U\Sigma V^t$ una descomposición SVD. Demostrar:

a) $\|Ax\|_2/\|x\|_2$ se maximiza para $x = v_1$, con v_1 la primer columna de V.

b) $\|A\|_2 = \sigma_1$. Deducir que $\|A\|_2 = \sqrt{\rho(A^tA)}$[‡‡].

c) $\|A\|_F = \sqrt{\sum_{i=1}^n \sigma_i^2}$.

d) Si $m = n$ y A es inversible, entonces $\kappa_2(A) = \sigma_1/\sigma_n$.

e) $\max_i |a_{ii}| \leqslant \sigma_1$.

[‡‡]Dada una matriz $B \in \mathbb{R}^{n\times n}$ se define el *radio espectral* de B como $\rho(B) = \max\{|\lambda| : \lambda \text{ autovalor de } B\}$.

a) b)

$\|A\|_2 = \sigma_1 = \sqrt{\rho(A^TA)}$

$\downarrow$

Radio espectral: $\rho(A^TA) = \max\{|\lambda| \text{ con } \lambda \text{ autovalor } A^TA\}$

$\rho(A^TA) = |\lambda_1| = \lambda_1 = \sigma_1^2$ con λ_1 mayor autovalor en módulo

$\geqslant$ QVQ: $\|A\|_2 \geqslant \sigma_1$

$$\|A\|_2 = \max_{x:\|x\|_2=1} \|Ax\|_2 \geqslant \|A\tilde{x}\|_2 \text{ con } \|\tilde{x}\|_2 = 1$$

$$= \sqrt{(A\tilde{x})^TA\tilde{x}} = \sqrt{\tilde{x}^TA^TA\tilde{x}} = \sqrt{\lambda_1 z^Tz} = \sqrt{\lambda_1} = \sigma_1$$

$\uparrow$

Sea $z \neq 0$ tq $A^TAz = \lambda_1 z$ y además $\|z\|_2 = 1$.

Tomamos $\tilde{x} = z$.

$\boxed{\leq}$ QVQ: $\|A\|_2 \leq \sigma_1$

$$\|A\|_2 = \max_{x : \|x\|_2 = 1} \|Ax\|_2$$

$v_1 \cdots v_n$ base ortonormal $\quad x = \sum_{j=1}^{n} \beta_j v_j$

$$\|x\|_2 = 1 \quad \Rightarrow \quad x^T x = 1 \quad \Rightarrow \quad \left[\sum_{j=1}^{n} \beta_j v_j\right]^T \cdot \sum_{j=1}^{n} \beta_j v_j = 1$$

$$\Rightarrow \quad \sum_{j=1}^{n} \beta_j^2 = 1 \quad \text{pues } v_1 \cdots v_n \text{ ortonormales}$$

$$\|Ax\|_2^2 = (Ax)^T Ax = x^T A^T A x = \left[\sum_{j=1}^{n} \beta_j v_j^T\right] \cdot A^T A \cdot \left[\sum_{j=1}^{n} \beta_j v_j\right]$$

$$= \left[\sum_{j=1}^{n} \beta_j v_j^T\right] \cdot \left[\sum_{j=1}^{n} \beta_j A^T A v_j\right]$$

$$= \left[\sum_{j=1}^{n} \beta_j v_j^T\right] \cdot \left[\sum_{j=1}^{n} \beta_j \lambda_j v_j\right]$$

$$= \sum_{j=1}^{n} \beta_j^2 \lambda_j \leq \sum_{j=1}^{n} \beta_j^2 \lambda_1 = \lambda_1$$

$\lambda_j \leq \lambda_1 \ \forall j$

$$\therefore \|Ax\|_2^2 \leq \lambda_1 \quad \Rightarrow \quad \|Ax\|_2 \leq \sqrt{\lambda_1} = \sigma_1$$

$$\Rightarrow \quad \|A\|_2 \leq \sigma_1$$

Alternativa

b) a)

QVQ $\|A\|_2 \leqslant \sigma_1$

U ortogonal

$$\|A\|_2 = \max_{x:\|x\|_2=1} \|U\Sigma V^T x\|_2 \overset{\downarrow}{=} \max_{x:\|x\|_2=1} \|\Sigma V^T x\|_2 \underset{\uparrow}{=} \max_{y:\|y\|_2=1} \|\Sigma y\|_2$$

Cambio de variable

$y = V^T x \iff x = Vy$

V inversible $\Rightarrow$ hay una biyección

$\|y\|_2 = \|V^T x\|_2 = \|x\|_2$ misma norma

$$\|A\|_2 = \max_{y:\|y\|_2=1} \|\Sigma y\|_2 \iff \|A\|_2^2 = \max_{y:\|y\|_2=1} \|\Sigma y\|_2^2$$

$$\|A\|_2^2 = \max_{y:\|y\|_2=1} \sum_{i=1}^{n} \sigma_i^2 y_i^2 \underset{\downarrow}{\leqslant} \max_{y:\|y\|_2=1} \sigma_1^2 \overset{=1}{\sum y_i^2} = \sigma_1^2$$

$y_i \geqslant 0 \wedge \sigma_1 \geqslant \sigma_i \geqslant 0 \quad \forall i = 1 \dots n \Rightarrow \sigma_1^2 \geqslant \sigma_i^2 \quad \forall i = 1 \dots n$

$$\therefore \|A\|_2^2 \leqslant \sigma_1^2 \iff \|A\|_2 \leqslant \sigma_1$$

QVQ $\|A\|_2 = \sigma_1$

$$\|A\|_2 = \max_{x:\|x\|_2=1} \|Ax\|_2 \geqslant \|Av_1\|_2 \quad \text{pues } \|v_1\|_2 = 1$$

$$\|Av_1\|_2 = \|U\Sigma V^T v_1\|_2 \underset{\uparrow}{=} \|\Sigma e_1\|_2 = \|\sigma_1 e_1\|_2 = |\sigma_1| \cdot \overset{=1}{\|v_1\|_2} \underset{\uparrow}{=} \sigma_1$$

U ortogonal $\qquad\qquad$ $\sigma_1 \geqslant 0$

$$\therefore \|A\|_2 = \sigma_1$$

c)

QVQ $\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$

a) $\|A\|_F^2 = tr(A^TA)$ $\qquad A \in \mathbb{R}^{m\times n}$

$$tr(A^TA) = \sum_{i=1}^{n} (A^TA)_{ii} = \sum_{i=1}^{n} \sum_{k=1}^{m} (a_{ki})^2 = \|A\|_F^2$$

b) $\|QA\|_F = \|A\|_F$

$$\|QA\|_F^2 \underset{a)}{=} tr((QA)^T QA) = tr(A^T \underbrace{Q^T Q}_{I} A) = tr(A^TA) \underset{a)}{=} \|A\|_F^2$$

c) $\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$

$$\|A\|_F = \|U\Sigma V^T\|_F \underset{b)}{=} \|\Sigma V^T\|_F = \|V\Sigma^T\|_F \underset{b)}{=} \|\Sigma^T\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$$

d) $A \in \mathbb{R}^{n\times n}$ inversible

$K(A) = \|A\| \cdot \|A^{-1}\|$

$K_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \sigma_1 \cdot 1/\sigma_n = \sigma_1/\sigma_n$

$\downarrow$

$A^{-1} = (U\Sigma V^T)^{-1} = V\Sigma^{-1}U^T$ con $(\Sigma^{-1})_{ii} = 1/\sigma_i \quad \forall i = 1 \ldots n$

existe porque rango$(A) = n$

entonces $\sigma_i > 0 \quad \forall i = 1 \ldots n$

$\sigma_1 \geqslant \cdots \geqslant \sigma_n \Leftrightarrow 1/\sigma_1 \leqslant \cdots \leqslant 1/\sigma_n$