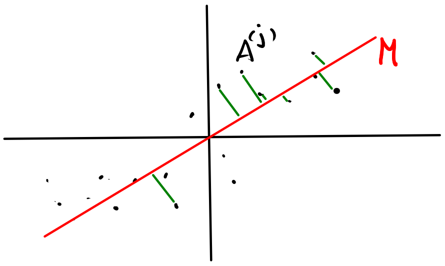
7.5 PROBLEMA DE MÍNIMOS CUADRADOS IV

• caso particular: sean $A^{(1)} \dots A^{(m)} \in \mathbb{R}^2$



¿cuál es la recta M que pasa por (0) que esté "más cerca" de esos puntos?

buscar la recta M que minimiza

$$\boxed{\mathcal{D} = \sum_{j=1}^{m} \| A^{(j)} - \mathbb{P}_M A^{(j)} \|^2}$$

• en general, si tenemos $\{A^{(j)}\}_{j=1}^{m}$ puntos de $\mathbb{R}^n$ y tenemos $\nu < \min\{n, m\}$: ¿cuál es el subespacio vectorial $M \subset \mathbb{R}^n$ de dimensión $\leq \nu$ que minimiza $\mathcal{D}$?

<u>teorema</u> (Eckhart-Young, Minsky, Schmidt)

sea $A \in \mathbb{C}^{n \times m}$, $A = U \Sigma V^*$ su SVD

sea $p = \min\{n, m\}$, sea $\nu < rg(A)$.

sea $A_\nu = \sum_{k=1}^{\nu} \sigma_k \, U^{(k)} \otimes V^{(k)}$

recordar que $A = \sum_{k=1}^{p} \sigma_k \, U^{(k)} \otimes V^{(k)}$

$\Rightarrow$ para toda $B \in \mathbb{C}^{n \times m}$, $rg(B) \leq \nu$

$$\| A - B \|_F \geq \| A - A_\nu \|_F = \left(\sum_{k=\nu+1}^{p} \sigma_k^2 \right)^{1/2} .$$

<u>observación</u>: este enunciado nos da una solución al problema

- $\|A-B\|_F^2 = \sum_{j=1}^{m} \|(A-B)^{(j)}\|^2 = \sum_{j=1}^{m} \|A^{(j)} - B^{(j)}\|^2$

- B tiene $rg(B) \leq \nu \Rightarrow$ las columnas de B generan un subespacio $M = \mathcal{L}(B^{(1)} \dots B^{(m)})$ de $\dim(M) \leq \nu$

$\hookrightarrow$ minimizar $\|A-B\|_F$ para toda B de $rg(B) \leq \nu$

es buscar un mínimo sobre un conjunto más grande respecto a la minimización de $\mathcal{D}$: para $\mathcal{D}$ se busca solo entre las matrices B cuyas columnas son $\mathbb{P}_M A^{(j)} \dots$ pero da lo mismo

- ¿ quienes son las columnas de A_ν, solución de la minimización sobre todas las B?

$$(A_\nu)_{ij} = \sum_{k=1}^{\nu} \sigma_k U_i^{(k)} \overline{V_j^{(k)}} \Rightarrow A_\nu^{(j)} = \sum_{k=1}^{\nu} \sigma_k V_{jk} U^{(k)}$$

donde $\sigma_k \overline{V_{jk}} = \langle A^{(j)}, U^{(k)} \rangle$ EJERCICIO

$$\hookrightarrow A_\nu^{(j)} = \sum_{k=1}^{\nu} \langle A^{(j)}, U^{(k)} \rangle U^{(k)}$$

esta es la proyección ortogonal de $A^{(j)}$ sobre $\mathcal{L}(U^{(1)} \dots U^{(\nu)}) \leftarrow$ llamemosla M

$$A^{(j)} = \underbrace{\sum_{k=1}^{\nu} \langle A^{(j)}, U^{(k)} \rangle U^{(k)}}_{\mathbb{P}_M A^{(j)}} + \underbrace{\sum_{k=\nu+1}^{m} \langle A^{(j)}, U^{(k)} \rangle U^{(k)}}_{(I-\mathbb{P}_M) A^{(j)}}$$

el teorema nos dice que la suma de las normas cuadradas de estos es mínima.

demostración:

- $$\|A - A_\nu\|_F = \left\| \sum_{k=1}^{p} \sigma_k\, U^{(k)} \otimes V^{(k)} - \sum_{k=1}^{\nu} \sigma_k\, U^{(k)} \otimes V^{(k)} \right\|_F$$

$$= \left\| \sum_{k=\nu+1}^{p} \sigma_k\, U^{(k)} \otimes V^{(k)} \right\|_F = \| U \tilde{\Sigma} V^* \|_F$$

$$0 \cdots 0 \;\; \sigma_{\nu+1} \cdots \sigma_p$$

$$= \left(\sum_{k=\nu+1}^{p} \sigma_k^2 \right)^{1/2}$$

- sea $B \in \mathbb{C}^{n \times m}$, $rg(B) = \ell \leq \nu$

y sea $E = A - B$, $E = X S Y^*$ su SVD

queremos demostrar que $\boxed{s_k \geq \sigma_{k+\nu}}$ $\forall k$

↳ porque, si esto es cierto, entonces

$$\|E\|_F^2 = \sum_{k=1}^{p} s_k^2 \geq \sum_{k=1}^{p} \sigma_{k+\nu}^2 = \sum_{k=\nu+1}^{p} \sigma_k^2 = \|A - A_\nu\|_F^2$$

la prueba termina

- $k=1$: sea $z_1 \in \mathcal{L}(V^{(1)} .. V^{(\nu+1)}) \cap Ker(B)$, $\|z_1\| = 1$

↳ este z_1 existe, porque

- $B : \mathbb{C}^m \to \mathbb{C}^n$, $\dim(Ker(B)) = m - rg(B)$

$\geq m - \nu$

- $\dim\left(\mathcal{L}(V^{(1)} .. V^{(\nu+1)})\right) = \nu + 1$

$[m - \nu + \nu + 1 = m + 1] \Rightarrow \mathcal{L}(V^{(1)} .. V^{(\nu+1)}) \cap Ker(B) \neq \{0\}$

$$S_1 = |||E|||_2 \geq \|E z_1\|_2 = \|(A-B) z_1\|_2$$

$z_1 \in Ker(B)$ $$= \|A z_1\|_2 = \left\| \sum_{k=1}^{p} \sigma_k < z_1, V^{(k)} > U^{(k)} \right\|_2$$

$z_1 \in \mathcal{L}(V^{(1)} .. V^{(\nu+1)})$ $$= \left\| \sum_{k=1}^{\nu+1} \sigma_k < z_1, V^{(k)} > U^{(k)} \right\|_2$$

$\{U^{(k)}\}$ ortonormales $$= \left(\sum_{k=1}^{\nu+1} \sigma_k^2 \, |< z_1, V^{(k)} >|^2 \right)^{1/2}$$

$\sigma_k \geq \sigma_{k+1}$ $$\geq \sigma_{\nu+1} \left(\sum_{k=1}^{\nu+1} |< z_1, V^{(k)} >|^2 \right)^{1/2}$$

$\|z_1\|_2 = 1$ $$= \sigma_{\nu+1}$$

• $k=2$: recordar que $S_2 = \max\limits_{\substack{\|z\|_2=1 \\ z \perp Y^{(1)}}} \|E z\|_2$.

sea $z_2 \in \underbrace{\mathcal{L}(V^{(1)} .. V^{(\nu+2)})}_{\dim\ \nu+2} \cap \underbrace{Ker(B) \cap (Y^{(1)})^{\perp}}_{\dim\ \geq\ m-\nu-1}$, $\|z_2\|_2 = 1$

$$S_2 = \max_{\substack{\|z\|=1 \\ z \perp Y^{(1)}}} \|E z\| \geq \|E z_2\| = \|A z_2\|_2$$

como antes... $$= \left\| \sum_{k=1}^{\nu+2} \sigma_k < z_2, V^{(k)} > U^{(k)} \right\|_2$$

$$= \left(\sum_{k=1}^{\nu+2} \sigma_k^2 \, |< z_2, V^{(k)} >|^2 \right)^{1/2}$$

$$\geq \sigma_{\nu+2} \left(\sum_{k=1}^{\nu+2} |< z_2, V^{(k)} >|^2 \right)^{1/2}$$

$$= \sigma_{\nu+2}$$

• y el mismo argumento se puede repetir para todos los S_k .

#