2019CS50429

Declaration: Few results and derivations are referred from Trefethen & Bau, Numerical Linear Algebra; the primary reference for this course.

1. $A, B \in \mathbb{C}^{m\times m}$ $\quad C = AB$

$\left.\begin{array}{l} a_1 \geq \cdots \geq a_m \\ b_1 \geq \cdots \geq b_m \\ c_1 \geq \cdots \geq c_m \end{array}\right\}$ singular values

---

Before proving what is asked, we will prove some lemmas we'll use.

Lemma 1: $c_i \leq a_i b_1$ & $c_i \leq a_1 b_i$

Proof: By using min-max characterization of singular values, we have

$$c_i = \max_{S:\dim(S)=i} \min_{x\in S, \|x\|=1} \|ABx\|$$

$$\leq \|A\| \cdot \max_{S:\dim(S)=i} \min_{x\in S, \|x\|=1} \|Bx\|$$

$$= a_1 b_i$$

$$\Rightarrow c_i \leq a_1 b_i \quad \text{— (a)}$$

$\because \sigma_i(AB) = \sigma_i(BA)$ [where σ denotes singular value] [$\because$ they have identical spectra]

|||y $\Rightarrow c_i \leq a_i b_1$ — (b)

From (a) & (b),

$$\Rightarrow \boxed{c_i \leq \min(a_1 b_i, a_i b_1)}$$

2019CS50429

1. Lemma 2: $c_m \geq a_m b_m$

Proof: Note that the m^{th} singular value is the smallest/minimum value of singular value for those particular matrices.

$$c_m = c_{min} = \min_{x \neq 0} \frac{\|ABx\|_2}{\|x\|_2}$$

$$= \min_{x \neq 0} \frac{\|ABx\|_2}{\|Bx\|_2} \cdot \frac{\|Bx\|_2}{\|x\|_2}$$

$$\geq \min_{y \neq 0} \frac{\|Ay\|_2}{\|y\|_2} \min_{x \neq 0} \frac{\|Bx\|_2}{\|x\|_2}$$

$$= a_m b_m$$

$$\Rightarrow \boxed{c_m \geq a_m b_m}$$

Lemma 3: $c_1 \geq a_1 b_m$, $c_1 \geq a_m b_1$

Proof: Note that the 1st singular value is the the largest/maximum value of the singular value for those particular matrices.

$$c_1 = c_{max} = \max_{x \neq 0} \frac{\|ABx\|_2}{\|x\|_2}$$

$$= \max_{x \neq 0} \frac{\|ABx\|_2}{\|Bx\|_2} \cdot \frac{\|Bx\|_2}{\|x\|_2}$$

$$= \max_{y \neq 0} \frac{\|Ay\|_2}{\|y\|_2} \max_{x \neq 0} \frac{\|Bx\|_2}{\|x\|_2}$$

$$\geq \max_{y \neq 0} \frac{\|Ay\|_2}{\|y\|_2} \min_{y \neq 0} \frac{\|Bx\|_2}{\|x\|_2}$$

$$= a_1 b_m$$

2019CS50429

1. contd-3

$\therefore \; c_1 \geq a_1 b_m$

|||ly $c_1 \geq a_m b_1$

$\Rightarrow \; \boxed{c_1 \geq \max(a_1 b_m, a_m b_1)}$

Now,

Using lemma 1, $a_1 b_1 \geq c_1$

Using lemma 3, $c_1 \geq \max(a_1 b_m, a_m b_2)$

$\Rightarrow \; \boxed{a_1 b_1 \geq c_1 \geq \max(a_1 b_m, a_m b_1)}$

Again,

Using lemma 1, $\min(a_1 b_m, a_m b_1) \geq c_m$

Using lemma 2, $c_m \geq a_m b_m$

$\Rightarrow \; \boxed{\min(a_1 b_m, a_m b_1) \geq c_m \geq a_m b_m}$

Hence, proved.

2. $S \cap T = \{0\}$

$S + T = \mathbb{C}^m$

$S \perp T$

We know that $I-P$ is a complementary projector of P such that subspaces of $(I-P) \cap (P) = \{0\}$ ($\because$ range $(P) \cap$ null $(P) = \{0\}$).

Let $u = Px \in S$

and $v = (I-P)y \in T \in \mathbb{C}^m$

Since $S \perp T$,

From the orthogonal projectors' definition, we know that for S & T to be orthognal

$P^* = P$ (From trefethen & Bau) (Also below)

$u^* v = x^* P^* (I-P) y = x^* (P - P^2) y = 0$

($\because P = P^2$ for projectors)

$\therefore$ T is the set $\{v \in \mathbb{C}^m : u^* v = 0 \ \forall u \in S\}$

Proof of $P^* = P$ for orthogonal subspaces:

Let $\{q_1, q_2, \cdots q_m\}$ be orthonormal basis for $\mathbb{C}^m$ where $\{q_1, \ldots q_n\}$ is basis for S & $\{q_{n+1} \cdots q_m\}$ is basis for T. For $j \leq n$, we have $Pq_j = q$

& for $j > n$ we have $Pq_j = 0$.

2019CS50429

5

2. Now let Q be the unitary matrix whose jth column is q_j. We then have

$$PQ = \left[\, q_1 \,\middle|\, \cdots \,\middle|\, q_n \,\middle|\, 0 \,\middle|\, \cdots \right]$$

$$Q^* P Q = \underbrace{\begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix}}_{\text{first } n\text{-entries are } 1} = \Sigma$$

$$P = Q \Sigma Q^*$$

$$P^* = (Q \Sigma Q^*)^* = Q \Sigma^* Q^* = Q \Sigma Q^* = P$$

$$\Rightarrow \boxed{P^* = P}$$

2019CS50429

3. (a) Let us look at this problem geometrically.

- $\|x_i\|_2 = \|y_i\|_2$ denotes that the vectors have same length for each i.

$\|x_i - x_j\|_2 = \|y_i - y_j\|_2$ denotes that the vectors have same angle between them for each i & j.

For any full column rank matrix $A = [a_1, \dots a_n]$

and $P_u\, a = \dfrac{\langle u, a\rangle}{\langle u, u\rangle} u$ (P is projection)

Using Gram-Schmidt process, we have

$$\hat{R} = \begin{pmatrix} \langle e_1, a_1\rangle & \langle e_1, a_2\rangle & \langle e_1, a_3\rangle & \cdots \\ 0 & \langle e_2, a_2\rangle & \langle e_2, a_3\rangle & \cdots \\ 0 & 0 & \langle e_3, a_3\rangle & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where $e_k = \dfrac{u_k}{\|u_k\|}$ & $u_k = a_k - \sum_{j=1}^{k-1} \mathrm{proj}_{u_j} a_k$

& $a_k = \sum_{j=1}^{k} \langle e_j, a_k\rangle e_j$

e_k denote the orthonormal basis of A.

$\forall\, k = 1$ to n

2019CS50429

Now, for x & y,

$$\hat{R}_x = \begin{pmatrix} \langle e_1^x, x_1 \rangle & \langle e_1^x, x_2 \rangle & \langle e_1^x, x_3 \rangle & \cdots \\ 0 & \langle e_2^x, x_2 \rangle & \langle e_2^x, x_3 \rangle & \cdots \\ 0 & 0 & \langle e_3^x, x_3 \rangle & \vdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Illy for y.

inner product denotes dot product OR how much of the first vector points towards the second vector.

$\langle a, b \rangle = \|a\|_2 \|b\|_2 \cos\theta$

$\langle e_i, x_j \rangle = \|e_i\|_2 \|x_j\|_2 \cos\theta_{ij} = \|x_j\|_2 \cos\theta_{ij}$

where θ_{ij} is the angle b/w $\|e_i\|_2$ & $\|x_j\|_2$

$\langle e_i, x_j \rangle = \|x_j\|_2 \cos\theta_{ij} = \|y_j\|_2 \cos\theta_{ij} = \langle e_i, y_j \rangle$ —(1)

(since they have same lengths & angles between them)

$\therefore$ All the entries in $\hat{R}_x$ & $\hat{R}_y$ are equal by (1).

$\therefore$ x & y have same $\hat{R}$.

2. (b)

$Q x_i = y_i$

$\Rightarrow Q = y_i x_i^{-1}$

$\Rightarrow Q = \dfrac{y_i x_i^*}{x_i^* \cdot x_i}$

$$\left\{ \begin{array}{l} V V^{-1} = 1 \\ V^* V V^{-1} = V^* \\ V^{-1} = \dfrac{V^*}{V^* V} \end{array} \right\}$$

$\Rightarrow Q = \dfrac{y_i x_i^*}{\|x_i\|^2}$

$r = \|x_i\| * \|x_i\|$

For $m = 1$ to i

For $n = 1$ to i

$q_{mn} = y_{m1}^* x_{1n} / r$

Assuming that the vectors are non-zero (otherwise q could be anything trivially) vectors have full rank.

$\Rightarrow Q$ & R in the QR factorisation of a vector are both invertible.

Now, let $x_i = Q_x R$ & $y_i = Q_y R$ (calculating using Gram-Schmidt)

1. $Q Q_x R = Q_y R$
2. $Q Q_x = Q_y$ ($\because$ R is invertible)
3. $Q = Q_y Q_x^{-1}$ ($\because$ Q_x & Q_y are invertible)

Algorithm.

1. calculate Q_x & Q_y using modified Gram-Schmidt.
2. calculate Q_x^{-1}
3. Find $Q = Q_y Q_x^{-1}$.

Here, Q will be orthogonal since both Q_y & Q_x^{-1} are orthogonal & multiplication of orthogonal matrices is orthogonal.

4. $A \in \mathbb{C}^{m \times n}$

(a) $v \in \mathbb{C}^m$

$$F = I - \frac{2vv^*}{v^*v}$$

$$FA = \left[I - \frac{2vv^*}{v^*v} \right] A$$

$$= A - \frac{2vv^*A}{v^*v}$$

$$= A - v \left[\frac{2v^*A}{v^*v} \right]$$

comparing it with $FA = A + vw^*$, we see that

$$w^* = \frac{2v^*A}{v^*v} \quad \& \quad w \in \mathbb{C}^n$$

$\therefore$ w is a vector.

(i) computing F then performing matrix multiplication.

(a) $F = I - \frac{2vv^*}{v^*v}$ → m^2 multiplications (flops), m flops; v^*v → m multiplications (flops), $m-1$ additions (flops); m^2 subtractions

$\sim 2m^2 + 3m - 1$ operations (flops)

(b) FA $\quad F \in \mathbb{C}^{m \times m}$ $\quad A \in \mathbb{C}^{m \times n}$

$\sim (2m-1)mn$ operations (flops)

(a)+(b) $= \sim 2m^2n - mn + 2m^2 + 3m - 1$

$\boxed{\sim 2m^2n}$ operation count (flops)

Scanned with CamScanner

2019CS50429

(ii) Computing w, then vw^* then matrix addition

(a) $w^* = \dfrac{2 (v^* A)}{(v^* v)}$ → $n(2m-1)$ operations (flops)

m multiplications
m-1 additions

$\sim (2m-1) n + (2m-1)$

$\sim (2m-1)(n+1)$ operations

(b) vw^* $\quad v \in \mathbb{C}^m \quad w^* \in \mathbb{C}^{1 \times n}$

$\sim mn$ operations (flops)

(c) Add A & vw^* of size $m \times n$

$\sim mn$ operations (flops)

$(a) + (b) + (c) = (2m-1)(n+1) + mn + mn$

$= 2mn + 2m - n - 1 + mn + mn$

$= 4mn + 2m - n - 1$

$\boxed{\sim 4mn \text{ operations}}$ (flop)s

Method (ii) is more efficient in terms of time complexity.

Scanned with CamScanner

2019CS50429

5. $Ax = b$ $A \in \mathbb{C}^{m \times n}$

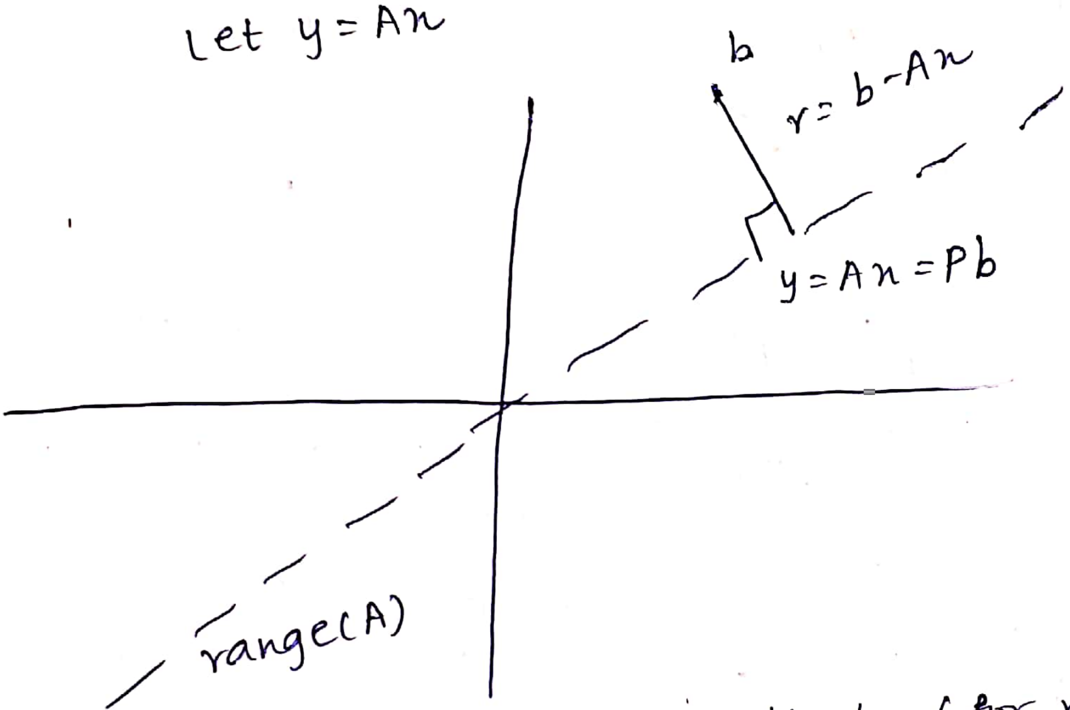
$m > n$

$\text{rank}(A) = r < n$

SVD of A $A = U \Sigma V^*$

(a) $y \in \text{range}(A)$

Let $y = Ax$



Geometrically, we can see that (for $r = b - Ax$) $\|r\|_2 = \|b - Ax\|_2$ is minimized only when $r \perp \text{range}(A)$

$\Rightarrow A^* r = 0$

From the properties of orthogonal projectors, we know that we can have P st $Pb = Ax$ & $P \in \mathbb{C}^{m \times m}$ is the orthogonal projector onto range(A).

$\boxed{\therefore y = Pb}$ → The explicit formula

5. (b) From the previous problem,

$\cancel{A^* r = 0}$

$\cancel{r = b - Ax}$

$\cancel{A^*(b - Ax) = 0}$

$\cancel{A^* b - A^* A x = 0}$

$\cancel{A^* A x = A^* b}$

$\cancel{\boxed{x = (A^* A)^{-1} A^* b}}$

$A^* r = 0$

$r = b - Ax$

$A^*(b - Ax) = 0$

$A^* b - A^* A x = 0$

$A^* A x = A^* b$

$x = (A^* A)^{-1} A^* b$

It is given by $\boxed{z = A^+ b + (I - A^+ A) w}$

where $A^+ = (A^* A)^{-1} A^*$ & w is an arbitrary n-dimensional vector.

From above, $A^+ b$ is already a solution. If we can show that $(I - A^+ A) w$ gives every vector in $K(A)$ & no other then, we have all solutions given by z.

Let $k = (I - A^+ A) w$, then:

$$Ak = A(I - A^+ A) w = (A - A A^+ A) w = (A - A) w = 0$$

$\therefore$ Every vector of the form $(I - A^+ A) w$ is a member of $K(A)$

Suppose k is a vector in the kernel of A, s.t $Ak = 0$. Then, as $(I - A^+ A) k = (k - A^+ A k)$ $= (k - A^+ 0) = k$, every vector in $K(A)$ may be written as $(I - A^+ A) w$

$\boxed{z = A^+ b}$ minimizes $\|x\|_2$.

6. (a)

$$p(x) = \sum_{k=0}^{n} p_k x^k$$

$$q(x) = \sum_{j=0}^{n} q_j x^j$$

$$(p(x), q(x)) = \sum_{k,j=0}^{n} \bar{p}_k q_j (x^k, x^j) \quad \text{--- (1)}$$

From

$$(p, q) = [p]^* G [q] \quad \text{--- (2)}$$

From (1) & (2),

we have G of the form

$$G = \begin{bmatrix} (x^1, x^1) & (x^2, x) & \cdots & (x^n, x) \\ \vdots & & \ddots & \\ (x^1, x^n) & & & (x^n, x^n) \end{bmatrix}$$

$$g_{mn} = (x^m, x^n)$$

$$g^*_{mn} = (x^n, x^m)$$

since $(x^m, x^n) = (x^n, x^m)$

$G = G^*$

Hence G is hermitian.

2019CS50429

(14)

6. (b) It means finding an orthogonal projection of q in the subspace <p>
If it is r, r is given by

$$r = \frac{(q,p)}{(p,p)} p$$

where $(p,q) = \int_{-1}^{1} \overline{p(x)}\, q(x)\, dx$

Since r is just a projector, an orthogonal projector,

(from Trefethen & Bau)

$$r = \left(p\,(p^* p)^{-1} p^*\right) q$$

∴ The matrix that acts on q is $(p(p^* p)^{-1} p^*)$.

Scanned with CamScanner

2019CS50429

4. (b)

$$\tilde{F} = I - \frac{2\tilde{v}\tilde{v}^{*}}{\tilde{v}^{*}\tilde{v}}$$

$$\frac{\|\tilde{v} - v\|_2}{\|v\|_2} = O(\epsilon_m)$$

$$\frac{\|\delta v\|_2}{\|v\|_2} = O(\epsilon_m)$$

$$\tilde{v} = v + \delta v$$