HINTS:

PROBLEM 1:

- if A is column vector: how many choices do you have for x?
- if A is a row vector: $\|Ax\|_2$ is a scalar product and you can use Cauchy-Schwarz inequality
- if A is diagonal, write $\|Ax\|_2$ explicitely and use that $\lambda_i^2 \leq \lambda_{max}^2$ for all $i=1,\dots,n$, where $\lambda_{max} = \max_{i=1\dots n} |\lambda_i|$

PROBLEM 2: Remember the relation between the singular values of a matrix A and the eigenvalues of AA^T and A^TA (you need to be a bit careful with the "shapes" of these matrices).

PROBLEM 3: Write down everything explicitly as a function of the entries of the matrices A and B.

PROBLEM 4:

- Consider the elements on the diagonal of A: their sum must be n, but what happens to $\|A\|_2$ if one of them is larger than 1?
- If $a_{ij} \neq 0$ for some $i \neq j$, can you find a vector x such that $\|x\|_2 = 1$ and $\|Ax\|_2 > 1$, in order to contradict the assumption that $\|A\|_2 \leq 1$?

PROBLEM 5: Use the given relation (3 times!) and the property that

$$\max\{f_1(X_1,Y) + f_2(X_2,Y) \mid Y \text{ satisfies sth}\} \leq$$
$$\leq \max\{f_1(X_1,Y) \mid Y \text{ satisfies sth}\} + \max\{f_2(X_2,Y) \mid Y \text{ satisfies sth}\}$$

PROBLEM 6: You should note (by trying the examples in Matlab) that

$$\text{rank}(A * B) \leq \min\{\text{rank}(A) \cdot \text{rank}(B),\ n\}$$

[why isn't it $=$?]

To prove it, recall that you can write a matrix C of rank r_c as the sum of r_c rank-1 matrices $C = \sum_{i=1}^{r_c} x_i y_i^T$ for some vectors x_i, y_i. Given such decompositions for A and B, try to find a decomposition for $A * B$ using $r_A \cdot r_B$ terms.

SOLUTIONS:

PROBLEM 1

ⓐ If A is a column vector $(n=1)$ then $A = a \in \mathbb{R}^m$

By definition, $\|A\|_2 = \max\limits_{\|x\|=1} \|Ax\|_2 = \max\{\|a\|_2, \|-a\|_2\} = \|a\|_2$

$\hookrightarrow x \in \mathbb{R}$ so it is either 1 or -1

If A is a row vector $(m=1)$ then $A = a^T$ for $a \in \mathbb{R}^m$

By definition $\|A\|_2 = \max\limits_{\substack{\|x\|_2=1 \\ x\in\mathbb{R}^n}} \|Ax\|_2 = \max\limits_{\substack{\|x\|_2=1 \\ x\in\mathbb{R}^n}} \|\underbrace{a^T x}_{\in\mathbb{R}}\|_2 =$

$$= \max_{\substack{\|x\|_2=1 \\ x\in\mathbb{R}^m}} |\langle a, x\rangle| \underset{\text{Cauchy-Schwarz}}{\leq} \max_{\|x\|_2=1,\, x\in\mathbb{R}^n} \|a\|_2 \cdot \|x\|_2 = \|a\|_2 \qquad \square$$

ⓑ $\|\mathrm{diag}(\lambda_1, \dots, \lambda_n)\|_2 \overset{\text{def.}}{=} \max\limits_{\substack{x\in\mathbb{R}^n \\ \|x\|_2=1}} \|\mathrm{diag}(\lambda_1, \dots, \lambda_n) \cdot x\|_2 =$

$$= \max_{\substack{x\in\mathbb{R}^n \\ \|x\|_2=1}} \|(\lambda_1 x_1, \dots, \lambda_n x_n)^T\|_2 \underset{\text{by part ⓐ}}{=} \max_{\substack{x\in\mathbb{R}^n \\ \|x\|_2=1}} \sqrt{\lambda_1^2 x_1^2 + \dots + \lambda_n^2 x_n^2} \leq$$

$$\underset{\text{let } \lambda_{max} = \max\{|\lambda_1|, \dots, |\lambda_n|\}}{\leq} \max_{\substack{x\in\mathbb{R}^n \\ \|x\|_2=1}} \sqrt{\lambda_{max}^2 x_1^2 + \dots + \lambda_{max}^2 x_n^2} = |\lambda_{max}| \cdot \max_{\substack{x\in\mathbb{R}^n \\ \|x\|_2=1}} \underbrace{\sqrt{x_1^2 + \dots + x_n^2}}_{=\|x\|_2} = |\lambda_{max}|$$

This proves that $\|\mathrm{diag}(\lambda_1, \dots, \lambda_n)\|_2 \leq |\lambda_{max}|$. To prove the opposite inequality, let i be an index such that $|\lambda_i| = |\lambda_{max}|$ and take the canonical vector $e_i = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ ← i-th row

We have that $\|e_i\|_2 = 1$ and $\|\mathrm{diag}(\lambda_1, \dots, \lambda_n) e_i\|_2 = \|\lambda_{max} e_i\|_2 = |\lambda_{max}|$ so $\|\mathrm{diag}(\lambda_1, \dots, \lambda_n)\|_2 \geq |\lambda_{max}|$. This concludes the proof. $\square$

[SOLUTIONS]

PROBLEM 2

$A \in \mathbb{R}^{m\times n}$ $\quad m \geq n \quad$ $A = \square$ $\quad$ singular values $\quad \sigma_1, \dots, \sigma_n$

- $B = [A \quad I_m]$

The singular values of B are the square root of the eigenvalues of $BB^T = AA^T + I_m$. The eigenvalues of AA^T are the squares of the sing. values of A with $m-n$ additional zero eigenvalues. Therefore

$$\lambda_i(BB^T) = \begin{cases} \sigma_i(A)^2 + 1 & i = 1,\dots,n \\ 1 & i = n+1,\dots,m \end{cases} \Rightarrow \sigma_i(B) = \begin{cases} \sqrt{\sigma_i^2(A)+1} & i = 1,\dots,n \\ 1 & i = n+1,\dots,m \end{cases}$$

- $B = \begin{bmatrix} A \\ I_n \end{bmatrix}$ The sing. val. of B are the square root of the eigenvalues of $B^TB = A^TA + I_n$. As the eigs of A^TA are the squares of the sing. val. of A we have

$$\sigma_i(B) = \sqrt{\sigma_i(A)^2 + 1} \quad \text{for } i = 1, \dots, n.$$

- $B = \begin{bmatrix} A & I_m \\ I_n & 0 \end{bmatrix}$ The singular values of B are the square roots of the eigenvalues of $C := B^TB = \begin{bmatrix} A^TA + I_n & A^T \\ A & I_m \end{bmatrix}$

Let $\begin{bmatrix} x \\ y \end{bmatrix}$ be an eigenvector of C. Then we must have

$$\begin{cases} A^TAx + x + A^Ty = \lambda x \\ Ax + y = \lambda y \end{cases} \longrightarrow (\lambda - 1)A^Ty = A^TAx$$

$\rightarrow$ if $\lambda \neq 1$, $\quad A^TAx + x + \frac{1}{\lambda - 1} A^TAx = \lambda x$

$$(\lambda-1)A^TAx + (\lambda-1)x + A^TAx = \lambda(\lambda-1)x$$

$$\lambda A^TAx = (\lambda-1)^2 x \longrightarrow \text{if } \lambda \neq 0, \boxed{A^TAx = \frac{(\lambda-1)^2}{\lambda}x} \ (*)$$

Note that there are (at least) $m-n$ eigenvalues equal to 1, corresponding to $x = 0$ and $A^Ty = 0$. If $\lambda = 0$ then $x = 0 \Rightarrow y = 0$ $\Rightarrow$ there are no eigenvalues 0.

Therefore, thanks to $(*)$ the other eigenvalues are of the form $\frac{(\lambda-1)^2}{\lambda} = \sigma_i(A)^2$

$$\Leftrightarrow \lambda^2 - 2\lambda + 1 = \lambda\sigma_i(A)^2 \Leftrightarrow \lambda^2 - \lambda(2 + \sigma_i(A)^2) + 1 = 0$$

$$\Leftrightarrow \lambda = \frac{1}{2}\left(2 + \sigma_i(A)^2 \pm \sqrt{\sigma_i(A)^4 + 4\sigma_i(A)^2}\right) = \frac{1}{2}\left(2 + \sigma_i(A)^2 \pm \sigma_i(A)\sqrt{4 + \sigma_i(A)^2}\right)$$

The singular values of B are then

$$\begin{cases} \sqrt{\frac{1}{2}\left(2 + \sigma_i(A)^2 \pm \sigma_i(A)\sqrt{4 + \sigma_i(A)^2}\right)} & (2n \text{ sing. val. like this}) \\ 1 & (m-n \text{ times}) \end{cases}$$

[SOLUTIONS]

PROBLEM 3

Prove that $\mathrm{trace}(B^TA) = \mathrm{trace}(AB^T) = \mathrm{vec}(B)^T \cdot \mathrm{vec}(A)$ $\quad A, B \in \mathbb{R}^{m\times n}$

Recall that $\mathrm{vec}(B) = (b_{11}\ b_{21}\ b_{31} \cdots b_{m1}\ b_{12}\ b_{22} \cdots b_{m2} \cdots b_{1n}\ b_{2n} \cdots b_{mn})^T$

therefore $\mathrm{vec}(B)^T \cdot \mathrm{vec}(A) = \sum_{i=1}^{m} \sum_{j=1}^{n} b_{ij} a_{ij}$.

By definition, $\mathrm{trace}(B^TA) = \sum_{i=1}^{n} (B^TA)_{ii} = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} (B^T)_{ij} a_{ji} \right) =$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ji} b_{ji} = \sum_{i=1}^{m} \sum_{j=1}^{n} b_{ij} a_{ij}$$

and, analogously, $\mathrm{trace}(AB^T) = \sum_{i=1}^{m} (AB^T)_{ii} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} (B^T)_{ji} \right) =$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij} \qquad \square$$

PROBLEM 4 Prove that the only matrix $A \in \mathbb{R}^{n\times n}$ s.t. $\mathrm{trace}(A) = n$ and $\|A\|_2 \leq 1$ is $A = I$.

Proof: $\mathrm{trace}(A) = n$ means that $a_{11} + a_{22} + \dots + a_{nn} = n$

As $a_{ii} = \langle e_i, Ae_i \rangle \underset{\uparrow \text{Cauchy-Schwarz}}{\leq} \|e_i\|_2 \cdot \|Ae_i\|_2 \leq 1 \cdot \|A\|_2 \leq 1$, we have $a_{ii} = 1$ for $i = 1, \dots, n$.

By contradiction, assume that $\exists\, i \neq j$ s.t. $a_{ij} \neq 0$. Then $\|A \cdot e_j\|_2 \geq \sqrt{a_{ij}^2 + a_{jj}^2} = \sqrt{1 + a_{ij}^2} > 1$, so $\|A\|_2 > 1$.

Therefore, $a_{ij} = 0 \quad \forall\, i \neq j \Rightarrow A = I$ $\qquad \square$

PROBLEM 5 $\max\{\mathrm{trace}(P^TMQ) : P \in \mathbb{R}^{m\times k}, Q \in \mathbb{R}^{n\times k}, P^TP = Q^TQ = I_k\} = \sigma_1 + \dots + \sigma_k$

$$\sigma_1(A+B) + \dots + \sigma_k(A+B) \overset{\text{property } (*)}{=} \max\{\mathrm{trace}(P^T(A+B)Q) : P \in \mathbb{R}^{m\times k}, Q \in \mathbb{R}^{n\times k}, P^TP = Q^TQ = I_k\}$$

$$\overset{\text{linearity of trace}}{=} \max\{\mathrm{trace}(P^TAQ) + \mathrm{trace}(P^TBQ) : P \in \mathbb{R}^{m\times k}, Q \in \mathbb{R}^{n\times k}, P^TP = Q^TQ = I_k\}$$

$$\leq \max\{\mathrm{trace}(P^TAQ) : P \in \mathbb{R}^{m\times k}, Q \in \mathbb{R}^{n\times k}, P^TP = Q^TQ = I_k\} + \max\{\mathrm{trace}(P^TBQ) \mid \dots\}$$

$$\overset{\text{property } (*)}{=} \sigma_1(A) + \dots + \sigma_k(A) + \sigma_1(B) + \dots + \sigma_k(B). \qquad \square$$

PROBLEM 6 We prove that $\mathrm{rank}(A * B) \leq \min\{\mathrm{rank}(A) \cdot \mathrm{rank}(B), n\}$

We have $A = \sum_{i=1}^{r_A} x_i y_i^T$, $B = \sum_{j=1}^{r_B} z_j w_j^T$, for col. vectors x_i, y_i, z_j, w_j

Therefore,

$$A * B = \left(\sum_{i=1}^{r_A} x_i y_i^T \right) * \left(\sum_{j=1}^{r_B} z_j w_j^T \right) = \sum_{i=1}^{r_A} \sum_{j=1}^{r_B} (x_i y_i^T) * (z_j w_j^T) =$$

$$= \sum_{i=1}^{r_A} \sum_{j=1}^{r_B} (x_i * z_j) \cdot (y_i * w_j)^T \rightarrow A * B \text{ has rank } \leq r_A r_B.$$

Of course, $\mathrm{rank}(A * B) \leq n$ because $A * B \in \mathbb{R}^{n\times n}$. The example from part c also shows that in general $\mathrm{rank}(A * B) \neq \min\{r_A r_B, n\}$. $\square$