HINTS

Problem 1:

1) Write $\left\| \begin{bmatrix} A-A_r & B \\ C & 0 \end{bmatrix} \right\|_F^2$ as a function of the Frobenius norms of the blocks.

2) 2×2 matrices will not give you a counterexample; you can think of something with $A \in \mathbb{R}^{2\times 2}$ (diagonal), $B \in \mathbb{R}^{2\times 1}$, $C = 0$...

Problem 2:

1) If $A = QR$, $|\det(A)| = |\det(R)|$ (why?)
What is $\det(R)$ as a function of the entries of R?
How can you relate $|r_{ii}|$ to $\|a_{i}\|_2$ using $R = Q^T A$?

2) Can you write an SVD of A if you know the SVD of S?

3) Start by finding QR decompositions of B and C, and then express BC^T in the form "PSQ^T".

4) Can you write $A_1 A_2^T + A_3 A_4^T = BC^T$ for suitable matrices B and C (possibly "bigger" than the A_i)?

Problem 4:

1) If A has rank r, consider the neighborhood
$$\mathcal{U} := \{A + B : \|B\|_2 \le \sigma_r(A)/2\}$$
and prove that $\operatorname{rank}(A+B) \ge r$ (using the stability of the singular values).
[This means that $\forall \varepsilon > 0$ you can take the neighborhood $\mathcal{U}$]

2) Consider $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and the sequence $A_n := \begin{bmatrix} 1 & 0 \\ 0 & 1/n \end{bmatrix}$...

SOLUTIONS

Problem 1

① $\left\| \begin{bmatrix} A-A_r & B \\ C & 0 \end{bmatrix} \right\|_F^2 = \|A-A_r\|_F^2 + \|B\|_F^2 + \|C\|_F^2$, therefore this is minimized if and only if $\|A-A_r\|_F$ is minimized. A solution is then given by the truncated SVD $\mathcal{T}_r(A)$. □

② Consider the matrices $A = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $C = [0]$

The best rank-1 approx. of A in the spectral norm is $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \mathcal{T}_1(A)$ which gives

$$\left\| \begin{bmatrix} A-\mathcal{T}_1(A) & B \\ C & 0 \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} 1/2 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\|_2 = \sqrt{1+\left(\tfrac{1}{2}\right)^2} > 1$$

On the other hand, choosing $A_r = \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix}$ gives

$$\left\| \begin{bmatrix} A-A_r & B \\ C & 0 \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\|_2 = 1.$$ So this is a counterexample. □

Problem 2

① Let $A = QR$ be a QR decomposition of A, so that $R = Q^T \cdot A$

By Binet theorem, $|\det(R)| = \underbrace{|\det(Q^T)|}_{=1} \cdot |\det(A)| = |\det(A)|$.

As R is triangular, we have

$$|\det(R)| = |r_{11}| \cdot |r_{22}| \cdot \ldots \cdot |r_{nn}|.$$

Moreover, by denoting $Q = [q_1 | \cdots | q_n]$ and $A = [a_1 | \cdots | a_n]$, we have $r_{ii} = q_i^T a_i \Rightarrow |r_{ii}| \underset{\text{Cauchy-Schwarz}}{\leq} \underbrace{\|q_i\|_2}_{=1} \cdot \|a_i\|_2 = \|a_i\|_2$ (*) ← equality iff q_i and a_i are proportional $\forall i \Leftrightarrow R$ is diagonal

Therefore we can conclude that $|\det(A)| \leq \prod_{i=1}^{n} \|a_i\|_2$. (**) □

② Let $S = U\Sigma V^T$ be the SVD of S.

Then, $A = PSQ^T = (PU) \cdot \Sigma \cdot (VQ)^T$ where $(PU)^T PU = U^T P^T P U = I$ and $(VQ)^T VQ = Q^T V^T V Q = I$ is an SVD of A. Its truncation to the first r singular values gives

$\mathcal{T}_r(A) = (PU)_r \cdot \Sigma_r \cdot [(VQ)_r]^T$. Note that $(PU)_r = P \cdot U_r$ and $(VQ)_r = V \cdot Q_r$, so $\mathcal{T}_r(A) = P \cdot U_r \Sigma_r V_r^T Q^T = P \cdot \mathcal{T}_r(S) \cdot Q^T$. □

[SOLUTIONS]

③ $B = Q_1R_1 \qquad C = Q_2R_2 \qquad$ economy QR (cost $O(mR^2 + nR^2)$)

$$BC^T = Q_1 \cdot \underbrace{(R_1R_2^T)}_{\hookrightarrow \text{ computing this costs } O(R^3)} \cdot Q_2$$

Compute SVD of $R_1R_2^T = U\Sigma V^T$ (cost $O(R^3)$)

Truncate the SVD $\mathcal{T}_r(R_1R_2^T) = U_r\Sigma_rV_r^T$

and return the decomposition $\mathcal{T}_r(BC^T) = (Q_1U_r)\cdot\Sigma_r\cdot(Q_2V_r)^T$. □

④ $$A_1A_2^T + A_3A_4^T = \underbrace{\begin{bmatrix} A_1 & A_3 \end{bmatrix}}_{=:B} \cdot \left(\underbrace{\begin{bmatrix} A_2 & A_4 \end{bmatrix}}_{=:C}\right)^T = BC^T$$

We can apply the algorithm from point ③ to the $m\times 2r$ matrix B and $n \times 2r$ matrix C. As $R = 2r$, the cost is $O(mr^2 + nr^2)$. □

Problem 4

① If $A = 0$ any neighborhood is fine. Otherwise, let $r := \text{rank}(A)$ and consider $\mathcal{U} := \{A + B : \|B\|_2 < \sigma_r(A)/2\}$. It is sufficient to prove that $\forall C \in \mathcal{U}$ $\text{rank}(C) \geq r$. Let $C = A + B \in \mathcal{U}$, then by the lemma on the stability of the SVD we have that
$|\sigma_i(A+B) - \sigma_i(A)| \leq \|B\|_2 < \sigma_r(A)/2$. For $i = 1, \dots, r$ we have $\sigma_i(A) > \frac{\sigma_r(A)}{2}$
therefore $\sigma_i(A+B) \neq 0$ for $i = 1, \dots, r$, which means that $\text{rank}(A+B) \geq r$. □

② Consider the rank-1 matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and the rank-2 matrices $A_n = \begin{bmatrix} 1 & 0 \\ 0 & 1/n \end{bmatrix}$.
For $\varepsilon = 1/2$, for any neighborhood $\mathcal{U}$ of A there exists $n \in \mathbb{N}$ s.t. $A_n \in \mathcal{U}$, as $\text{rank}(A_n) = 2 > \text{rank}(A) + \varepsilon = \frac{3}{2}$, the function "rank" is <u>not</u> upper semi-continuous. □