Problem 1:
1) Write $\| [A^T A \ B] \|_F^2$ as a function of the Frobenius norms of the blocks.
2) $2 \times 2$ 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 $|r_{44}|$ to $\|a_4 \|_2$ 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 $B C^T$ in the form "PSQ^T".
4) Can you write $A_1 A_2^T + A_3 A_4^T = B C^T$ for suitable matrices $B$ and $C$ (possibly "bigger" than the $A_i$)?

Problem 3:
1) If $A$ has rank $r$, consider the neighborhood $\mathcal{U} := \{ A + B : \| B \|_2 \leq \sigma_r(A)/2 \}$ and prove that $\text{rank}(A+B) \geq r$ (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 $A_n := \begin{bmatrix} 1 & 0 \\ 0 & -n \end{bmatrix}$...
Problem 1

(1) \[ \left\| \begin{bmatrix} A & B \\ C & D \end{bmatrix} \right\|_F^2 = \|A\|_F^2 + \|B\|_F^2 + \|C\|_F^2, \] therefore this is minimized if and only if \( \|A\|_F \) is minimized. A solution is then given by the truncated SVD \( T_r(A) \).

(2) Consider the matrices \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C = [0] \)

The best rank-1 approx. of \( A \) in the spectral norm is \( \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} = T_1(A) \)

which gives \[ \left\| \begin{bmatrix} A - T_1(A) & B \\ C & D \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \right\|_2 = \sqrt{1 + \left( \frac{1}{2} \right)^2} > 1 \]

On the other hand, choosing \( A_r = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \) gives \[ \left\| \begin{bmatrix} A - A_r & B \\ C & D \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \right\|_2 = 1. \] So this is a counterexample.

Problem 2

(1) Let \( A = QR \) be a QR decomposition of \( A \), so that \( R = Q^T A \).

By Binet theorem, \( \det(R) = \det(Q) \det(A) \). As \( R \) is triangular, we have \( \det(R) = |r_{11}| \cdot |r_{22}| \cdot \cdots |r_{nn}| \).

Moreover, by denoting \( Q = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} \) and \( A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \), we have \( r_{ii} = q_i^T a_i \Rightarrow |r_{ii}| \leq \|q_i\|_2 \cdot \|a_i\|_2 \) (Cauchy-Schwarz).

Therefore we can conclude that \( \det(A) \leq \prod_{i=1}^n \|a_i\|_2 \) \( \text{iff} \quad q_i \) and \( a_i \) are proportional.

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

Then, \( A = PSQ^T = (PU)Σ_r (VQ)^T \) where \( (PU)^T PU = U^TPPU = I \) and \( (VQ)^T VQ = Q^TVVQ = I \) is an SVD of \( A \). Its truncation to the first \( r \) singular values gives \( T_r(A) = (PU)_r Σ_r (VQ)_r^T \). Note that \( (PU)_r = P \cdot U_r \) and \( (VQ)_r = V_r \cdot Q_r \), so \( T_r(A) = P \cdot U_r Σ_r V_r^T Q^T = P \cdot T_r(S) \cdot Q^T \).
[Solutions]

3. \( B = Q_1 R_1 \quad C = Q_2 R_2 \) economy QR (cost \( O(mR^2 + nR^2) \))

\[ BC^T = Q_1 \left( R_1 \right) R_2 \]

- computing this costs \( O(R^3) \)

Compute SVD of \( R_1 R_2^T = \Sigma V^T \), (cost \( O(R^3) \))

Truncate the SVD \( \Sigma_r (R_1 R_2^T) = U_r \Sigma_r V_r^T \)

and return the decomposition \( \Sigma_r (BC^T) = (Q_1 U_r) \cdot \Sigma_r \cdot (Q_2 V_r)^T \). \( \square \)

4. \( A_1 A_2^T + A_3 A_4^T = \begin{bmatrix} A_1 & A_2 \\ \hline B \\ \hline C \end{bmatrix} \cdot \begin{bmatrix} A_2 & A_4 \end{bmatrix}^T = BC^T \)

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 \( O(mR^2 + nR^2) \). \( \square \)

Problem 4

\( \star \quad 1. \) If \( A = 0 \) any neighborhood is fine. Otherwise, let \( r = \text{rank}(A) \) and consider \( U := \{ A + B : \| B \|_2 < \sigma_r(A)/2 \} \). It is sufficient to prove that \( \forall C \in U \text{ rank}(C) \geq r \). Let \( C = A + B \in U \), 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 \( i = 1, \ldots, r \) we have \( \sigma_i(A) > \sigma_r(A)/2 \) and therefore \( \sigma_i(A + B) \neq 0 \) for \( i = 1, \ldots, r \), which means that \( \text{rank}(A + B) \geq r \). \( \square \)

\( 2. \) Consider the rank-1 matrix \( A = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \) and the rank-2 matrices \( A_n = \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \).

For \( \epsilon = \epsilon_2 \), for any neighborhood of \( A \) there exists \( n \in N \) s.t. \( A_n \in U \), as \( \text{rank}(A) = 2 > \text{rank}(A) + \epsilon = 3/2 \), the function "rank" is not upper semi-continuous. \( \square \)