Exercise 2 – Low-rank approximation techniques

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Prof. Dr. D. Kressner
Dr. L. Periša

Problem 1. Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$ and $k \leq n$. Use von Neumann's trace inequality to prove that

$$\max\{\|Q^T A\|_F : Q \in \mathbb{R}^{n\times k},\ Q^T Q = I_k\} = \sqrt{\sigma_1^2 + \cdots + \sigma_k^2},$$

where $\sigma_1, \ldots, \sigma_k$ are the singular values of A.

Problem 2.

1. Prove the basic properties of the angle between vectors stated on Slide 12 of the slides of Lecture 2.
2. Prove the basic properties of the angle between a vector and a subspace stated on Slide 13 of the slides of Lecture 2.

Problem 3. *The goal of this exercise is to prove the projector characterization* $\sin\theta_1(\mathcal{X}, \mathcal{Y}) = \|XX^T - YY^T\|_2$ *for orthnormal bases* X, Y *of* $\mathcal{X}, \mathcal{Y}$.

1. Explain why one can assume w.l.o.g that $X = \begin{bmatrix} I_k \\ 0 \end{bmatrix}$ (Hint: Utilize the QR decomposition.)
2. Assuming that $X = \begin{bmatrix} I_k \\ 0 \end{bmatrix}$, partition $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ with $Y_1 \in \mathbb{R}^{k\times k}$. Setting $P = XX^T - YY^T$, show that P^2 is block diagonal. What are the diagonal blocks?
3. Using Point 2, compute the largest singular value of P in terms of the singular values of $X^T Y$ and establish $\sin\theta_1(\mathcal{X}, \mathcal{Y}) = \|XX^T - YY^T\|_2$.

Problem 4.

A function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is called lower semi-continuous (upper semi-continuous) at $x_0 \in \mathbb{R}^{m\times n}$ if for every $\epsilon > 0$ there exists a neighborhood $\mathcal{U}$ of x_0 such that $f(x) \geq f(x_0) - \epsilon$ ($f(x) \leq f(x_0) + \epsilon$) for all $x \in \mathcal{U}$.

1. Prove that the rank function $A \mapsto \operatorname{rank}(A)$ is lower semi-continuous at every matrix $A_0 \in \mathbb{R}^{m\times n}$. (Hint: Use the stability of singular values.)
2. Construct an example to show that the rank function is in general *not* upper semi-continuous.

Problem 5. *The goal of this exercise is to recall the QR decomposition and illustrate its use in low-rank approximation.*

Let $X \in \mathbb{R}^{m\times n}$ with $m \geq n$. Then there is an orthogonal matrix $Q \in \mathbb{R}^{m\times m}$ such that

$$X = QR, \quad \text{with} \quad R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \diagdown\!\!\big| \\ 0 \end{pmatrix},$$

that is, $R_1 \in \mathbb{R}^{n\times n}$ is an upper triangular matrix. In practice, one often uses the economy-size QR decomposition instead: Letting $Q_1 \in \mathbb{R}^{m\times n}$ contain the first n columns of Q, one obtains

$$X = Q_1 R_1 = Q_1 \cdot \diagdown\!\!\big| \;.$$

The computation of such an economy-size QR decomposition requires $\mathcal{O}(mn^2)$ operations.

1. Given $A \in \mathbb{R}^{n\times n}$, partition $A = \begin{bmatrix} a_1, a_2, \ldots, a_n \end{bmatrix}$ with $a_i \in \mathbb{R}^n$. Using the QR decomposition, show *Hadamard's inequality*:

$$|\det(A)| \leq \|a_1\|_2 \cdot \|a_2\|_2 \cdots \|a_n\|_2.$$

   Characterize the set of all $n \times n$ matrices A for which equality holds.

2. Let $P \in \mathbb{R}^{m \times R}$, $Q \in \mathbb{R}^{n \times R}$, with $R \leq n \leq m$, be matrices with orthonormal columns. Let $A = PSQ^T$ and $r < R$. Prove that $P \cdot \mathcal{T}_r(S) \cdot Q^T$ is a best rank-r approximation of A in the sense that

$$\|A - P \cdot \mathcal{T}_r(S) \cdot Q^T\|_F = \|A - \mathcal{T}_r(A)\|_F.$$

3. Using the result from Point 2, develop an algorithm of complexity $\mathcal{O}(mR^2 + nR^2)$ for performing rank-r truncation of a matrix BC^T with $B \in \mathbb{R}^{m \times R}$, $C \in \mathbb{R}^{n \times R}$, $R \leq n \leq m$. Implement and test this algorithm in MATLAB.

4. Using the algorithm from Point 3, develop an algorithm of complexity $\mathcal{O}(mr^2 + nr^2)$ for recompressing the sum of two rank-r matrices back to rank r. Implement and test this algorithm in MATLAB.