Exercise 3 – Low-rank approximation techniques

Prof. Dr. D. Kressner
A. Cortinovis

Problem 1

Given a matrix of the form

$$\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \quad A \in \mathbb{R}^{m \times n}, \quad B \in \mathbb{R}^{m \times k}$$

and $C \in \mathbb{R}^{l \times n}$, consider the problem of finding a matrix A_r, with $\operatorname{rank}(A_r) \le r$, such that $\left\| \begin{bmatrix} A - A_r & B \\ C & 0 \end{bmatrix} \right\|$ is small as possible.

1. Prove that for the Frobenius norm ($\|\cdot\| \equiv \|\cdot\|_F$) the solution is given by $\mathcal{T}_r(A)$.
2. Provide an example which shows that the same statement does not hold in general for the spectral norm ($\|\cdot\| \equiv \|\cdot\|_2$).

Problem 2: Using the QR decomposition for low-rank approximation.

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

$$X = QR, \quad \text{with} \quad R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \triangledown \\ 0 \end{pmatrix},$$

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

$$X = Q_1 R_1 = Q_1 \cdot \triangledown .$$

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

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

   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 $R \le n \le m$, 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 $B \in \mathbb{R}^{m \times R}$, $C \in \mathbb{R}^{n \times R}$, $R \le n \le m$. 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.

Problem 3: Principal Component Analysis (PCA)

The iris dataset at `https://archive.ics.uci.edu/ml/datasets/Iris` contains 3 classes of 50 instances each, where each class refers to a type of iris plant.

1. Import the data in Matlab / Python / Julia. If you use Matlab, transform the file `iris.data` in a `iris.csv` file, then use the commands

   ```
   T = readtable('iris.csv'); A = table2array(T(:,[1 2 3 4]));
   ```

2. Subtract the mean from each column.

3. Compute the first principal component and plot the projection of the data on this component (use three different colors for the three types of iris plants). Is one component enough to "explain" the data?

4. Repeat the previous point using the first and second principal components.

Problem 4: Semi-continuity of the rank

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 $\varepsilon > 0$ there exists a neighborhood $\mathcal{U}$ of x_0 such that $f(x) \geq f(x_0) - \varepsilon$ ($f(x) \leq f(x_0) + \varepsilon$) 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.