Big Data Computing **A.Y. 2023/2024**
Homework 3 **Deadline: December 3^{rd}**

Instructions. You should hand in your homeworks within the due date and time. Late deliveries will be penalized, please check homework delivery policies in the exam info.

Handing in. You should submit your work as *one zip file* containing i) a Colab notebook corresponding to your first assignment, with text explaining your implementation choices and an explanation of your results and ii) *one pdf* containing your answers and solutions to the theoretical exercises (i.e., Assignments 2, 3 and 4). Answers to theoretical exercises should be either typed up or hand written **very clearly** and scanned. The first option is strongly preferred.

Please deliver your homework by attaching the above zip file to an email addressed to me as follows:
To: `becchetti@diag.uniroma1.it`
Subject: `HW3 <Your last name> <Your first name> <Your Sapienza ID number>`

Typesetting in Latex. Latex is very handy for typesetting (especially math), but you need to install it. If you do not want to install Latex, you can go for Overleaf, providing an integrated, Web interface, accessible for free in its basic version (which is enough for your needs). It allows you to both type in Latex using a Web interface, and compiling your code to produce a pdf document. Overleaf's documentation also contains a tutorial on Latex essentials.

Important note. Grading of your answers to the theoretical assignments will depend on i) correctness and solidity of the arguments, ii) clarity of the presentation and iii) mathematical rigour. E.g., ill-defined quantities, missing assumptions, undefined symbols etc. are going to penalize you. Rather than writing a lot, try to write what is needed to answer and write it well.

Assignment 1.

The goal of the first assignment is to appreciate the practical meaning of low-rank approximation. To this purpose, we are going to take a look at image reconstruction. At a high level, your task is the following: i) read in a black/white photo and convert it into a matrix; ii) perform a singular value decomposition of the matrix thus obtained; iii) reconstruct the photo corresponding to using only 5%, 10%, 25%, 50% of the singular values. Things you should look at:

- Print the reconstructed photo. How good is the quality of the reconstructed photo?
- What percent of the Frobenius norm of the matrix corresponding to the original picture is retained in each case?

In practice, to perform a first test, you can:

- Use `skimage`, the Python image preprocessing package, to convert pictures into numpy arrays. To begin, use `from skimage import data` and then something like `image = data.coins()` to have the matrix corresponding to the b/w picture of a set of coins arranged on a dark surface (this is a small dataset that is made available by `skimage`).
- Use `matplotlib` to print the picture corresponding to a matrix. For example, the matrix might be the low-rank approximation of the original matrix obtained by retaining 25% of the singular values.

This is a really simple assignment, but one that allows you to fully appreciate the reconstruction properties of the Singular Value Decomposition. Once you have the pipeline in place (my estimate is that this is going to take less than 1 hour) you are encouraged to test it on different pictures, just for your own curiosity.

Assignment 2.
Consider a square *symmetric* matrix $A \in \mathbb{R}^{n \times n}$ with spectral decomposition

$$A = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T,$$

where of course some of the λ_i's might be zero. Answer the following questions:

1. For an interger $k \geq 1$, write A^k in terms of the $\mathbf{u}_i$'s and the λ_i's.
2. Assume that A is invertible, i.e., $\lambda_i \neq 0$ for $i = 1, \ldots, n$. Using your answer to the previous question: i) guess the expression of the inverse A^{-1} of A in terms of the $\mathbf{u}_i$'s and the λ_i's and ii) *prove* that your guess is correct.

Note: you should provide rigorous arguments for your answers.

Assignment 3.
Consider any real matrix $A \in \mathbb{R}^{n \times m}$. Let $A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ be the SVD of A. *Prove* the following:

1. Write the expression of $(AA^T)^k$ in terms of the σ_i's, the $\mathbf{u}_i$'s and the $\mathbf{v}_i$'s, where $k \geq 1$ is an integer;
2. Consider a *square* matrix Q, whose columns form an orthonormal basis. *Prove* that its rows are also orthonormal;
3. Assume next that A is square (but not necessarily symmetric), i.e., $n = m$. Assume further that A is invertible. Prove that $B = \sum_{i=1}^{n} \frac{1}{\sigma_i} \mathbf{v}_i \mathbf{u}_i^T$ is the inverse of A.

Hints: To answer point 1 above, you might reuse part of the work you did for Assignment 2 (in this case, please state clearly what you are reusing). To answer point 3 above, at some point you may need to use the claim you proved for point 2.

Note: you should provide rigorous arguments for your answers.

Assignment 4.

A square matrix $B \in \mathbb{R}^{n \times n}$ is positive semi-definite (PSD) if $\mathbf{x}^T B \mathbf{x} \geq 0$, for every $\mathbf{x} \in \mathbb{R}^n$. Answer the following:

1. Consider a real matrix $A \in \mathbb{R}^{n \times m}$. Prove that AA^T is PSD.
2. **Bonus question.** Consider a real, symmetric matrix $A \in \mathbb{R}^{n \times n}$. Prove that A is PSD if and only if *all* its eigenvalues are non-negative.

Note 1: you should provide rigorous arguments for your answers.

Note 2: you do not have to answer the bonus question, unless you answered all other questions of this homework and you are reasonably confident in the quality of your answers. In this case, the bonus question might give you extra points.

Important note. Please check the collaboration policy on the course web page.