Department of Mathematics, University of Utah

Analysis of Numerical Methods I

MATH 6610 – Section 001 – Fall 2025

Homework 5

SVD, II

Due Wednesday, September 24, 2025

Submission instructions:
Submit your assignment on gradescope.

Problem assignment:

1. ("Inf-sup" constants) Let $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ be arbitrary. Consider the non-negative constant,

$$\beta = \inf_{\boldsymbol{x}\in\mathbb{C}^n\setminus\{\boldsymbol{0}\}} \sup_{\boldsymbol{y}\in\mathbb{C}^n\setminus\{\boldsymbol{0}\}} \frac{|\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{y}|}{\|\boldsymbol{x}\|_2 \, \|\boldsymbol{y}\|_2}.$$

Show that $\beta > 0$ iff $\boldsymbol{A}$ is invertible. You may use inf / min and sup / max interchangeably for this problem.

2. (Interlacing of singular values) Let $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, and let $\boldsymbol{Q} \in \mathbb{C}^{(m-1)\times m}$ satisfy $\boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{I}$. Compute bounds for the singular values of $\boldsymbol{Q}\boldsymbol{A}$ in terms of the singular values of $\boldsymbol{A}$.

3. (Density of full-rank matrices) Let $\boldsymbol{A} \in \mathbb{C}^{m\times n}$. Show that in any matrix norm $\|\cdot\|$, $\boldsymbol{A}$ can be approximated to arbitrary tolerance (say $\epsilon > 0$) by a full-rank matrix $\boldsymbol{B} = \boldsymbol{B}(\epsilon)$.

4. (The Polar decomposition) The matrix exponential $e^{\boldsymbol{D}}$ of a diagonal matrix $\boldsymbol{D}$ is the diagonal matrix formed by elementwise exponentiating the diagonal entries. The matrix exponential of $i\boldsymbol{\Theta}$, where $\boldsymbol{\Theta} \in \mathbb{C}^{n\times n}$ is a Hermitian matrix is given by,

$$e^{i\boldsymbol{\Theta}} = \boldsymbol{V} e^{i\boldsymbol{\Lambda}} \boldsymbol{V}^*, \qquad \boldsymbol{\Theta} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^*,$$

where the latter equality is the eigenvalue decomposition of $\boldsymbol{\Theta}$.

(a) Show that if $\boldsymbol{U}$ is any $n \times n$ unitary matrix, then it can be written as $e^{i\boldsymbol{\Theta}}$ for some Hermitian matrix $\boldsymbol{\Theta}$.

(b) Let $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ be arbitrary. Show that $\boldsymbol{A}$ has the decompositions,

$$\boldsymbol{A} = \boldsymbol{R} e^{i\boldsymbol{\Theta}} = e^{i\boldsymbol{\Theta}} \boldsymbol{S},$$

where all matrices are square, $\boldsymbol{\Theta}$ is Hermitian, and $\boldsymbol{R}$ and $\boldsymbol{S}$ are Hermitian positive semi-definite matrices that are similar to each other.

5. (Moore-Penrose Pseudoinverse) Let $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ have rank $r \leq \min\{m, n\}$ and *reduced* SVD,

$$\boldsymbol{A} = \widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{\Sigma}}\widetilde{\boldsymbol{V}}^*.$$

The Moore-Penrose pseudoinverse of $\boldsymbol{A}$ is,

$$\boldsymbol{A}^+ := \widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{\Sigma}}^{-1}\widetilde{\boldsymbol{U}}^*.$$

(a) Prove that if $\boldsymbol{A}$ is invertible, then $\boldsymbol{A}^+ = \boldsymbol{A}^{-1}$.

(b) Prove that the matrices $\boldsymbol{A}\boldsymbol{A}^+$, $\boldsymbol{A}^+\boldsymbol{A}$, $\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^+$, and $\boldsymbol{I} - \boldsymbol{A}^+\boldsymbol{A}$ are all projection matrices, and characterize their ranges and kernels in terms of subspaces defined by $\boldsymbol{A}$. Are these matrices orthogonal projectors?

(c) Give conditions under which $\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A}$.

(d) Give conditions under which each of the more standard "inverse" equality $\boldsymbol{A}^+\boldsymbol{A} = \boldsymbol{I}_n$ holds. What about for $\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{I}_m$?

(e) Is the operation $\boldsymbol{A} \mapsto \boldsymbol{A}^+$ well-conditioned? I.e., for general $\boldsymbol{A}$ and a perturbation $\boldsymbol{B}$, can $\|(\boldsymbol{A}+\boldsymbol{B})^+ - \boldsymbol{A}^+\|/\|\boldsymbol{A}^+\|$ be controlled by $\|\boldsymbol{B}\|/\|\boldsymbol{A}\|$? If so, provide a condition number. If not, describe why not.

6. (Linear dimension reduction with PCA)
In a programming language of your choice, plot PCA-compressed representations and embeddings of the Yale Face Database:
`https://www.kaggle.com/datasets/olgabelitskaya/yale-face-database`
For this problem, you will consider three experiments:

Experiment A. For this experiment, grayscale images can be viewed as a matrix $\boldsymbol{A}$ of numbers, where the dimensions of the matrix correspond to the pixel dimensions of the image. Approximate an image $\boldsymbol{A}$ by a rank-k PCA approximation. (Pick a couple of your favorites.) When plotting the original images versus the approximations, how does the accuracy depend on k? Does the accuracy qualitatively depend on what kind of image (face) you use?

Experiment B. In order to treat an image as a single data point, vectorize its matrix representation (i.e., unwind the matrix into a vector of length m). For an experiment with n data points (images), let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ be the resulting matrix. Use the rank-k PCA compression of $\boldsymbol{A}$, each column of this rank-k matrix (appropriately unwound) is an approximation of one of the images. Explore the accuracy of rank-k compression for some values of k (experiment!). How does the qualitative accuracy of this experiment compare to the previous one?

Experiment C. In the context of Experiment B, an *embedding* of a data point is a size-k vector of projected coefficients. (This is the vector $\boldsymbol{c}$ from the previous assignment that explains PCA.) Plot embeddings of the data in 2 and 3 dimensions (so that they can be visualized). Are there any discernable patterns of the embedded data in light of classifications of each face in the original data?

In the above, design the details your own experiments: from all the images in a Yale database, select a couple of sets of them (maybe all of them, maybe only the ones with glasses, etc), and investigate the accuracy of rank-k approximation. Your assignment should sufficiently explain the details of your experiment(s) so that the experiment is reproducible, and should present some numerical results (plots) that are briefly discussed.

Note that PCA is a reduction scheme since in this context a rank-k matrix requires only $\mathcal{O}(km)$ storage instead of, e.g., $\mathcal{O}(m^2)$ storage, which is a substantial difference if $k \ll m$.