Indian Institute of Science, Banglore
Department of Computational and Data Sciences (CDS)

DS284: Numerical Linear Algebra

Assignment 3 [Posted Sept 13, 2025]

Faculty Instructor: Dr. Phani Motamarri
TAs: Naman Pesricha, Harshit Rawat, Nikhil Kodali, Sejal Maisheri, Ayush Kumar Pushp, Deepti Sahu

Submissions required: Problems 2 and 3. **Max Points: 50**

Notations: Vectors and matrices are denoted below by bold faced lower case and upper case alphabets respectively.

Problem 1

This exercise will walk you through the steps in proving the existence of SVD of any rectangular matrix $\mathbf{A}$ of size $m \times n$ with rank r.

(a) Matrices of the form $\mathbf{G} = \mathbf{A}^T\mathbf{A}$ are called Gram matrices where $\mathbf{A} \in \mathbb{R}^{m\times n}$.
Show that $\mathbf{x}^T\mathbf{G}\mathbf{x} \geqslant 0 \; \forall \; \mathbf{x} \in \mathbb{R}^n$ and hence show that all eigen-values of $\mathbf{G}$ are non-negative.

(b) Show that $\mathbf{A}$ and $\mathbf{A}^T\mathbf{A}$ have the same rank.

(c) Show that a vector $\mathbf{u}$ of the form $\mathbf{A}\mathbf{v}/\sigma$ $(\sigma > 0)$ is a unit eigen-vector of $\mathbf{A}\mathbf{A}^T$ where $\mathbf{v}$ and σ^2 form the eigen-vector, eigen-value pair of $\mathbf{A}^T\mathbf{A}$.

(d) Note that i^{th} eigen-vector, eigen-value pair of $\mathbf{A}^T\mathbf{A}$ can be written as $\mathbf{A}^T\mathbf{A}\mathbf{v}_i = (\sigma_i^2)\mathbf{v}_i$.
Consider the case of a full rank matrix $\mathbf{A}$ ie. $(\sigma_i > 0 \;\; \forall i)$, if we define a new vector $\mathbf{u}_i = \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i}$, show that $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{m\times m}$ is an orthogonal matrix, $\mathbf{\Sigma} \in \mathbb{R}^{m\times n}$ is diagonal matrix and $\mathbf{V} \in \mathbb{R}^{n\times n}$ is again an orthogonal matrix, $\mathbf{u}_i$ is i^{th} column of $\mathbf{U}$, and $\mathbf{v}_i$ is i^{th} column of $\mathbf{V}$
[Note: In low rank scenario some $\sigma_i = 0$ if other non-zero σ_i are sorted, we can compute $\mathbf{U}$ by adding additional column vectors that span $\mathbb{R}^m$ and add rows of 0-vector to $\mathbf{\Sigma}$.]

Problem 2

Solution to this problem needs to be submitted by 28 SEP and will be graded.

You are one of the scientists working at NASA's Goddard Space Flight Center in Greenbelt, Maryland and have been researching Wide-Field Slitless Spectroscopy to capture galaxy spectra of the distant universe. With the help of NASA's James Webb Space Telescope, you have successfully captured the deepest and sharpest infrared image of the distant universe to date. It is an image of the galaxy cluster SMACS 0723 and has been named Webb's First Deep Field.

Unfortunately, due to some technical difficulties, the space telescope has not been able to transmit full-resolution images to Earth. However, an onboard computer can be programmed remotely from Earth to transmit the image in a compressed format until the difficulties are resolved. The control station on Earth has decided to use SVD to compress the image. As a scientist tasked with programming the onboard computer, think about the following:

(a) How many singular values are required to approximate the image i.e., make it look indistinguishable from the original image? (Hint: Load the image in Python or Matlab or Octave or Julia and the matrix representation of the image will be accessible to you. For $r \times r$ pixel image, the image will have $r \times r \times 3$ matrix entries with the number 3 corresponding to color depth of the image representing Red, Blue, Green.)

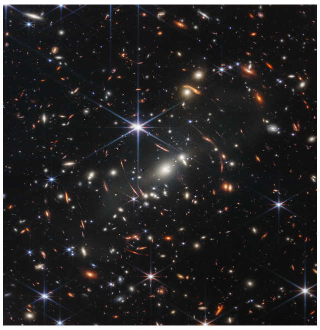
(b) Based on your observation in (a), how many entries need to be transmitted to earth to reconstruct the approximate image as opposed to sending the original image?

[Perform the tasks in a programming environment comfortable to you like Matlab/Octave/Python/Julia. You can use inbuilt functions for computing SVD.]

(c) What is the 2-Norm and Frobenius-Norm error between the matrix representation of the original image and the approximate image obtained for different number of singular values. Check if the following theorems hold for these errors:

For the matrix $\mathbf{A}$ of rank r, with singular values $\sigma_1,\ \sigma_2,\ \sigma_3,\ \dots\ \sigma_r,\ \mathbf{A}_v$ is the v-rank approximation of $\mathbf{A}$. ($\mathbf{A}_v = \sum_{i=1}^{v} \sigma_i \mathbf{u}_i \mathbf{v}_i$) such that $1 < v < r$, then: $\|\mathbf{A} - \mathbf{A}_v\|_2 = \sigma_{v+1}$, $\|\mathbf{A} - \mathbf{A}_v\|_F = \sqrt{\sigma_{v+1}^2 + \sigma_{v+2}^2 + \dots + \sigma_r^2}$

The image Webb's First Deep Field is as below and also downloadable from Teams assignment page as a PNG file.



Problem 3

Solution to this problem needs to be submitted by 28 SEP and will be graded.

Regression is one of the important aspects of data science and scientific computing. In most linear regression problems, you aim to find a linear map between input feature vector $\mathbf{x}$ and target scalar y based on a given dataset. The least squares approach we discussed in initial classes on vector norms is an approach to solve the regression problem where you try to minimise the squares of the residuals to find the model parameters. In the case of data in the 2D-plane, this least squares regression is like trying to minimize the errors vertically with respect to fitted line. This kind of least squares regression makes logical sense if you know *a priori* that the uncertainty is there only in your measurement variable y (target) but not in the input feature vector $\mathbf{x}$ at which you are measuring y. In certain applications, one would like to account for uncertainties in both $\mathbf{x}-$data and y-data. In these cases, we resort to orthogonal regression for building a robust model in contrast to the normal regression procedure described above. The following exercise seeks to derive an expression for the orthogonal regression where the best-fit line $\mathbb{L}$ is obtained by minimising the squares of orthogonal distances from each of the given data points to the best-fit line. You will also explore its connections to singular value decomposition. Say you have N data points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N \in \mathbb{R}^2$ corresponding to some measurements obtained in an experiment. We are now seeking to find the best-fit line $\mathbb{L}$ that minimizes the orthogonal fitting errors from the N points.

(a) Let $\bar{\mathbf{x}}_\mathbf{i}$ be the orthogonal projection of the given data point $\mathbf{x_i}$ on the line $\mathbb{L}$ you are seeking to find. Derive an expression for $\bar{\mathbf{x}}_\mathbf{i}$ in terms of unit-vector $\mathbf{n}$ which is normal to the best-fit line $\mathbb{L}$ and any point given $\mathbf{m} \in \mathbb{R}^2$ on $\mathbb{L}$.

(b) Recall our objective is to minimize the sum of squares of Euclidean distances between $\mathbf{x_i}$ and $\bar{\mathbf{x}}_\mathbf{i}$. Pose this minimization problem using the expression derived for $\bar{\mathbf{x}}_\mathbf{i}$ in (a). Note that this will be a minimization problem over both $\mathbf{m}$ and $\mathbf{n}$. One can easily show that the optimal $\mathbf{m}$ for any given $\mathbf{n}$ is of the form $\mathbf{m}^* = \frac{1}{N}\sum_i \mathbf{x_i}$ (you do not need to show now). Using this fact, rewrite the above minimization problem over $\mathbf{n}$ in terms of $\hat{\mathbf{x}}_\mathbf{i} = \mathbf{x_i} - \mathbf{m}^*$ and of course $\mathbf{n}$.

(c) Let $\mathbf{q} \in \mathbb{R}^2$ be a unit vector spanning the 1-dimensional vector space orthogonal to $\mathbf{n}$. Define a matrix $\mathbf{X} \in \mathbb{R}^{N\times 2}$ such that $\mathbf{X} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_N]^T$. Rewrite the minimization problem in (b) in terms of the matrix $\mathbf{X}$ and $\mathbf{q}$. [**Hint:** First get a relation between projectors corresponding to the vector spaces spanned by $\mathbf{q}$ and $\mathbf{n}$, then use the definition of Frobenius norm of matrix in terms of columns of a matrix. Note that the minimization problem is over $\mathbf{q}$ now.]

(d) Observe carefully the minimization problem you have obtained in part (c) above and explore connections to low-rank approximations of $\mathbf{X}$ to solve the minimization problem in (c). Finally comment on how $\mathbf{q}$ and hence $\mathbf{n}$ is related to singular vectors of $\mathbf{X}$.

(e) Once you know a point $\bar{\mathbf{x}}_\mathbf{i}$ on the best-fit line $\mathbb{L}$ and the unit normal vector $\mathbf{n}$ to it from the above exercise, what is the equation of the best-fit line $\mathbb{L}$ using the orthogonal regression procedure you carried out thus far?

Problem 4

(a) Geometrically, the orthogonal matrix is a matrix transformation that preserves 2-Norm of a matrix and causes rotation / reflection.

Can you justify $\mathbf{I} - 2\mathbf{P}$ is orthogonal matrix if $\mathbf{P}$ is orthogonal projector?

Prove the same algebraically as well.

(b) Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ and its projector $\mathbf{P}$ which projects all vectors orthogonally on to column space of $\mathbf{A}$, then answer the following questions:

- If $\mathbf{A}$ is full rank, what is $\mathbf{P}$?
- Given $\mathbf{P}$ is there any way to find out the null space of $\mathbf{A}$?
- What can you say about the eigen-values of $\mathbf{P}$?

(c) If $\mathbf{P} \in \mathbb{R}^{m\times m}$ be a non-zero projection. Show that $\|\mathbf{P}\|_2 \geqslant 1$ with equality, if and only if $\mathbf{P}$ is orthogonal projector.