Indian Institute of Science, Banglore
Department of Computational and Data Sciences (CDS)
DS284: Numerical Linear Algebra
Assignment 3 [Posted Sept 12, 2024]
Faculty Instructor: Dr. Phani Motamarri
TAs: Gourab Panigrahi, Srinibas Nandi, Nihar Shah, Rushikesh Pawar, Surya Neelakandan

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 Sep 29 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.)

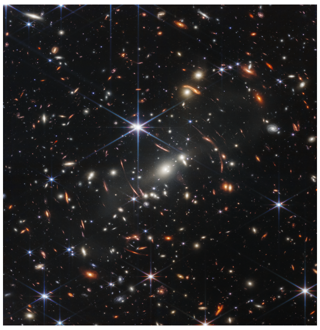
(b) Based on your observation in (a), how many entries need to be transmitted to earth to reconstruct the approximate image as opposed to sending the original image?

[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 $\sigma_1,\ \sigma_2,\ \sigma_3,\ \ldots\ \sigma_r$, $\mathbf{A}_v$ 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: $\|\mathbf{A} - \mathbf{A}_v\|_2 = \sigma_{v+1}$, $\|\mathbf{A} - \mathbf{A}_v\|_F = \sqrt{\sigma_{v+1}^2 + \sigma_{v+2}^2 + \ldots + \sigma_r^2}$

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 Sep 29 and will be graded

Consider a symmetric matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$. Answer the following 8 questions:

(a) Show that the singular values of $\mathbf{A}$ are absolute values of eigenvalues of $\mathbf{A}$. What can you say about the vector-induced matrix norm $\|\mathbf{A}\|_2$ in terms of eigenvalues of $\mathbf{A}$? Support your argument.

(b) Show that $|\mathbf{x}^T \mathbf{A} \mathbf{x}| \leqslant \|\mathbf{A}\|_2$ for any non-zero unit vector $\mathbf{x} \in \mathbb{R}^m$.

(c) Let the vector $\mathbf{u} \in \mathbb{R}^m$ be an eigenvector of the above symmetric matrix $\mathbf{A}$ corresponding to an eigenvalue λ i.e. $\mathbf{A}\mathbf{u} = \lambda \mathbf{u}$. Further, let the matrix $\mathbf{A}$ undergo a symmetric matrix perturbation by $\delta \mathbf{A}$ such that $\dfrac{\|\delta \mathbf{A}\|}{\|\mathbf{A}\|} = O(\epsilon_{mach})$. Let, $\tilde{\mathbf{u}} = \mathbf{u} + \delta \mathbf{u}$ and $\tilde{\lambda} = \lambda + \delta\lambda$ be the eigenvector-eigenvalue pair of the perturbed matrix $\tilde{\mathbf{A}} = \mathbf{A} + \delta \mathbf{A}$. Now, show that

$$|\delta\lambda| \leqslant \|\delta \mathbf{A}\|_2$$

(Hint:- Note that the perturbed matrix $\tilde{\mathbf{A}}$ is symmetric and start with the eigenvalue problem corresponding to $\tilde{\mathbf{A}}$ to first show that $|\delta\lambda| = |\mathbf{u}^T \delta \mathbf{A}\, \mathbf{u}|$. You may also assume that $\mathbf{A}$ is full rank and eigenvector-eigenvalue perturbations caused due to the symmetric perturbations in $\mathbf{A}$ are small and in the order of $||\delta \mathbf{A}||_2$.)

(d) Deduce the relative condition number for the problem of computing the eigenvalue λ of our symmetric matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ using the inequality derived in part (c).

(e) We now consider the problem of computing eigenvalues of the matrix $\mathbf{M} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, where a is a non-zero real number. As you can see, the two eigenvalues of this matrix $\mathbf{M}$ are a, a. Find the relative condition number for the mathematical problem of computing the eigenvalues for the above matrix $\mathbf{M}$ using the result obtained in part(d).

(f) An Algorithm S is designed to compute the eigenvalues of the above matrix $\mathbf{M} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ on a computer using floating point arithmetic. Algorithm S is designed to be a backward stable algorithm. To this end, comment on the relative forward error incurred in computing an eigenvalue of $\mathbf{M}$ by employing this backward stable Algorithm S.

(g) Furthermore, another Algorithm U is designed to compute eigenvalues of the above matrix $\mathbf{M}$ by solving the roots of the characteristic polynomial of $\mathbf{M}$ i.e. $p_{\mathbf{M}}(z) = \det(\mathbf{M} - z\mathbf{I}) = 0$. Assume that in $\mathbf{M} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, $a \in \mathbb{F}$, i.e. a does not have a floating point error when represented on a computer. List the steps of the Algorithm U required to compute eigenvalues of $\mathbf{M}$ on a computer. In doing so, deduce the floating-point approximation errors incurred both in evaluating the coefficients of $p_{\mathbf{M}}(z)$ and also, in the expressions employed to compute the roots of $p_{\mathbf{M}}(z)$ involving these coefficients. When doing this exercise, you may assume all the relative errors arising in floating point approximations to be the maximum possible relative error i.e. ϵ_M, the machine epsilon.

(h) Using the information in part (g), compute the forward relative error incurred in computing the eigenvalue a of $\mathbf{M}$ using the above Algorithm U on a computer and using this estimate, argue that the Algorithm U is unstable. (Hint: An Algorithm G is unstable if it is not both backward stable and stable. Also note that if the Algorithm G is backward stable or stable, then the relative forward error in the solution is $O(\kappa\epsilon_M)$ where κ is the condition number of the problem)

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.