Indian Institute of Science Banglore
Department of Computational and Data Sciences (CDS)
DS284: Numerical Linear Algebra
Mid-semester Exam 2024

Problem 1 [6×3.5=21]

Assert if the following statements are True or False. Give a detailed reasoning for your assertion. Marks will be awarded only for your general reasoning and not just for counter examples

(a) If $\hat{\mathbf{x}} \in \mathbb{R}^n$ satisfies the system of equations $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a full rank matrix with $m > n$ and $\mathbf{b} \in \mathbb{R}^m$, then $\hat{\mathbf{x}}$ should always satisfy the system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$

(b) Let $\mathbf{u} \in \mathbb{R}^m$ be such that $\|\mathbf{u}\|_2 = 1$ with all entries of $\mathbf{u}$ are positive. If we define $\mathbf{A} \in \mathbb{R}^{1\times m}$ as $\mathbf{A} = \mathbf{u}^T$, then induced norm of $\mathbf{A}$ in the 2-norm sense is the max element of $\mathbf{u}$ i.e., $\|\mathbf{A}\|_2 = \max\{u_i\}$.

(c) For the matrix $\mathbf{G} = \mathbf{I}_m - 3\mathbf{q}\mathbf{q}^T$ where $\mathbf{q} \in \mathbb{R}^m$ is a unit vector, the geometric multiplicity of the eigenvalue 1 is $m - 1$.

(d) Let $\mathbf{S} \in \mathbb{R}^{m\times m}$ be a symmetric positive definite matrix. There can exist a maximum element in $\mathbf{S}$ that does not lie on the diagonal. (**Hint:** If S_{ij} is the largest element for some $i \neq j$, then $S_{ii} + S_{jj} \leqslant 2S_{ij}$ should always be true.)

(e) Let $\mathbf{v}_1 \in \mathbb{R}^m$ be a unit-vector and the matrix $\mathbf{F} = \mathbf{I}_m - 2\mathbf{v}_1\mathbf{v}_1^T$ denote a Householder matrix. Furthermore, $(\epsilon_1,\mathbf{v}_1)$, $(\epsilon_2,\mathbf{v}_2)$, $(\epsilon_3,\mathbf{v}_3)$ $\cdots$ $(\epsilon_n,\mathbf{v}_n)$ represent the first $n(< m)$ largest eigenvalue/eigenvector pairs of a real symmetric matrix $\mathbf{S} \in \mathbb{R}^{m\times m}$. Now, assert the following statement: The n-dimensional space spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$ $\mathbf{v}_n$ is same as the eigenspace spanned by the eigenvectors of a matrix $\mathbf{T} = \mathbf{F}^{-1}\mathbf{S}\mathbf{F}$ corresponding to first n largest eigenvalues of $\mathbf{T}$.

(f) If $\mathbf{P} \in \mathbb{R}^{m\times m}$ is an orthogonal projector matrix, then one can choose the same set of orthogonal eigenvectors to diagonalize $\mathbf{P}$ and $\mathbf{I_m} - \mathbf{P}$.

a) False. This way of solving an equation is used when the system is over-determined, and hence, a least square approximation needs to be done.

In other words, we are basically projecting $\underset{\sim}{b}$ onto the column space of $\underset{\sim}{A}$ so that we can solve the equation. Furthermore, we use orthogonal projectors so that we can minimise the residual, i.e $\|\underset{\sim}{A}\hat{\underset{\sim}{x}} - \underset{\sim}{b}\|_2$.

b) False. We know that $\|\underset{\sim}{A}\|_2 = \|\underset{\sim}{u}^T\|_2 = \|\underset{\sim}{u}\|_2 = 1$

$\underset{\sim}{u}$ need not be the max. element in $\underset{\sim}{u}$. For e.g., if $\underset{\sim}{u} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$,

$\|\underset{\sim}{u}\|_2^2 = 1$, $\max(\underset{\sim}{u}_i) = \frac{\sqrt{2}}{2}$, $\|\underset{\sim}{A}\| = 1$

c) True. The only eigenvalue of $\underset{\sim}{I}$ will be 1. However, $\underset{\sim}{I}_m$ will have m linearly independent eigenvectors. All these eigenvectors shall correspond to the same eigenvalue (which is 1).

Now, for $\underset{\sim}{I} - 3\underset{\sim}{q}\underset{\sim}{q}^T$, we shall be getting two eigenvectors, i.e $\lambda_1 = 1$ & λ_2. The geometric multiplicity of λ_1 will be $(m-1)$ whereas for λ_2 it will be 1.

However, if $\underset{\sim}{u} = \underset{\sim}{0}$, then the statement will be false.

e) False. Because $\underset{\sim}{I} = \underset{\sim}{F}^{-1}\underset{\sim}{S}\underset{\sim}{F} \Rightarrow \underset{\sim}{S} = \underset{\sim}{F}\underset{\sim}{I}\underset{\sim}{F}^{-1}$, which implies the existence of a similarity transformation b/t $\underset{\sim}{S}$ & $\underset{\sim}{I}$.

The n-dimensional space occupied by $\underset{\sim}{v}_1 \cdots \underset{\sim}{v}_n$ is actually the eigenspace of $\underset{\sim}{S}$.
Because there is a similarity transformation b/t $\underset{\sim}{S}$ & $\underset{\sim}{I}$, they shall have the same eigenvalues. However, their eigenvectors shall differ.

Hence, the eigenspace of $\underset{\sim}{I}$ will not be the same as the eigenspace of $\underset{\sim}{S}$.

f) True. The eigenvalues of $\underset{\sim}{P}$ & $(\underset{\sim}{I}_m - \underset{\sim}{P})$ shall differ. However, the eigenvectors may remain the same.

Pf: Let $\underset{\sim}{P}\underset{\sim}{x} = \lambda_p \underset{\sim}{x}$

$$(\underset{\sim}{I}_m - \underset{\sim}{P})\underset{\sim}{x} = \underset{\sim}{I}_m \cdot \underset{\sim}{x} - \underset{\sim}{P} \cdot \underset{\sim}{x}$$
$$= \underset{\sim}{x} - \lambda_p \underset{\sim}{x}$$
$$= (1 - \lambda_p)\underset{\sim}{x}$$
$$= \lambda_{p'}\underset{\sim}{x}$$

Problem 2 [5+2+5+4=16]

Consider four vectors

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} \qquad \mathbf{a}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 0.4 \\ 0.7 \\ 0.4 \end{bmatrix}$$

(a) Construct the matrix $\mathbf{A}$ whose columns are the two vectors $\mathbf{a}_1, \mathbf{a}_2$. Compute the successive orthogonal transformations $\mathbf{Q}_1, \mathbf{Q}_2$ that transforms $\mathbf{A}$ into an upper triangle matrix $\mathbf{R}$ as shown below:

$$\mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \mathbf{R}$$

From this exercise, write down the orthonormal vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ that span the successive column subspaces of the matrix $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2]$

(b) Do the vectors $\mathbf{v}$ and $\mathbf{w}$ lie in the space spanned by $\mathbf{a}_1$ and $\mathbf{a}_2$? Explain.

(c) If the answer for (b) is "yes" for one of the vectors $\mathbf{v}$ and $\mathbf{w}$, write down the component(s) of this vector in the basis of $\mathbf{a}_1$ and $\mathbf{a}_2$ in a column vector format. If the answer for (b) is "no" for one of the vectors $\mathbf{v}$ and $\mathbf{w}$, then compute the solution $\mathbf{x}$ which minimizes the difference between $\mathbf{Ax}$ and this vector in the 2-norm sense.

(d) If the matrix $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2]$ denotes a data matrix related to 3 experiments and 2 features, construct a zero mean feature data matrix $\mathbf{A}_o$. Using this data matrix $\mathbf{A}_o$, now compute the unit vector in the direction of the second maximum variance.

a) $\underset{\sim}{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 4 & 0 \end{bmatrix}$ * Householder Triangularization

$\underset{\sim}{x_1} = A(1:3, 1) = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$

$\|\underset{\sim}{x_1}\|_2 = \sqrt{1^2+1^2+4^2} = \sqrt{18} = 3\sqrt{2}$

$\underset{\sim}{v_1} = +3\sqrt{2}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 1+3\sqrt{2} \\ 1 \\ 4 \end{pmatrix}$

$\|v_1\|_2 = \sqrt{(1+3\sqrt{2})^2+1+16} = \sqrt{36+6\sqrt{2}}$

$\underset{\sim}{u} = \dfrac{\underset{\sim}{v_1}}{\|v_1\|_2} = \begin{pmatrix} 0.786 \\ 0.150 \\ 0.600 \end{pmatrix}$ $(0.786\ \ 0.15\ \ 0.6)$

$$\underset{\sim}{F_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \underset{\sim}{u}\,\underset{\sim}{u}^T$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.618 & 0.118 & 0.472 \\ 0.118 & 0.023 & 0.09 \\ 0.472 & 0.09 & 0.36 \end{bmatrix}$$

b) $\alpha a_1 + \beta a_2 = v$?

$\alpha + \beta = 2$ ①
$\alpha + 2\beta = 2$ ②
$4\alpha = 1$ ③
} Overdetermined system of eqns. Hence, $\underset{\sim}{v}$ is not spanned by $\underset{\sim}{A}$.

$\alpha + \beta = 0.4$ ①
$\alpha + 2\beta = 0.7$ ②
$4\alpha = 0.4$ ③ $\Rightarrow \alpha = 0.1$, $\beta = 0.3 \Rightarrow$ Consistent soln.

Hence, $\underset{\sim}{v}$ is spanned by $\underset{\sim}{a_1}$ & $\underset{\sim}{a_2}$.

c) For $\underset{\sim}{v}$, $\underset{\sim}{A}\,\underset{\sim}{x} = \underset{\sim}{v}$ is overdetermined Then,

$\underset{\sim}{A}^T\underset{\sim}{A}\,\hat{\underset{\sim}{x}} = \underset{\sim}{A}^T\underset{\sim}{v}$

$$\left[\begin{array}{cc|c} 18 & 3 & k \\ 3 & 0 & 1 \end{array}\right]$$

$$A^TA = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 18 & 3 \\ 3 & 0 \end{bmatrix}$$

$$A^Tv = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

(2×3 · 3×1 = 2×1)

$3\alpha = 5 \Rightarrow \alpha = 5/3$, $\beta = \dfrac{5 - 3(5/3)}{10} = 0$

$\therefore \hat{x} = \begin{bmatrix} 5/3 \\ 0 \end{bmatrix}$

Verify:

$$\underset{\sim}{A}\underset{\sim}{\hat{x}} = \underset{3\times 2}{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 4 & 0 \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} 5/3 \\ 0 \end{bmatrix}} = \begin{bmatrix} 5/3 \\ 5/3 \\ 20/3 \end{bmatrix}$$

$$\underset{\sim}{r} = \underset{\sim}{v} - \underset{\sim}{A}\underset{\sim}{\hat{x}} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 5/3 \\ 5/3 \\ 20/3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \\ -17/3 \end{bmatrix}$$

For w_1 as calculated in b), $x = \begin{bmatrix} 0.1 \\ 0.3 \end{bmatrix}$

d) $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 4 & 0 \end{bmatrix}$ Mean of col. 1 = 2, " " " 2 = 1

$\therefore A_0 = \begin{bmatrix} -1 & 0 \\ -1 & 1 \\ 2 & -1 \end{bmatrix}$

$$A_0^T A_0 = \underset{2\times 3}{\begin{bmatrix} -1 & -1 & 2 \\ 0 & 1 & -1 \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} -1 & 0 \\ -1 & 1 \\ 2 & -1 \end{bmatrix}} = \begin{bmatrix} 6 & -3 \\ -3 & 2 \end{bmatrix}$$

Problem 3 [5+8=13]

Regression is one of the key aspects of machine learning. In most linear regression type problems, the aim is to find a linear map between input feature vector $\mathbf{x}$ and an output target scalar y based on a given dataset. One intrinsic assumption in linear regression is that all the data samples are given equal weightage. However, the data is usually collected from multiple sources with varying levels of fidelity involved in the measurement of target scalar y. This necessitates attributing different weights to different data samples. In such a scenario, we define the loss function (or the objective function that one has to minimize to find the optimal parameters in our regression) corresponding to our linear regression problem as follows:

$$\mathbf{L} = \sum_{i=1}^{n} w_i(y_i - \theta_0 - \theta_1 x_{i1} - \theta_2 x_{i2} - \theta_3 x_{i3} - ... - \theta_m x_{im})^2 \tag{1}$$

where n is the number of data samples in the given data set $(\mathbf{x}_i, y_i)_{i=1}^n$. Here $\mathbf{x}_i \in \mathbb{R}^m$, $y_i \in \mathbb{R}$, $w_i > 0$ are given and $n >> m$. The unknown quantities in linear regression are the parameter values $\theta_0, \theta_1, ..., \theta_m$, which need to be estimated through minimization of the loss function. Define a feature matrix $\mathbf{X} \in \mathbb{R}^{n\times(m+1)}$ such that the first entry of each row is 1, and the remaining m entries in each row correspond to the m entries of the feature vector $\mathbf{x}_i$. Based on the information above, answer the following questions:

(a) Show that if $\mathbf{X}$ is a full rank matrix, the optimal parameters obtained through minimization of $\mathbf{L}$ is unique. (*Hint:* Find the first derivatives of $\mathbf{L}$)

(b) To improve the generalizability of the regression model, it is typical to define a modified loss function $\mathbf{J} = \mathbf{L} + \eta \sum_{j=0}^{m} \theta_i^2$, where $\mathbf{L}$ is defined as in equation (1) and $\eta > 0$ is a user-controllable parameter. The set of optimal parameters can be obtained by minimizing this modified loss function $\mathbf{J}$. Show that the set of optimal parameters, i.e., $\theta_0, \theta_1, ..., \theta_m$ can always be made unique by varying the value of η irrespective of the rank of $\mathbf{X}$. Given a dataset $(\mathbf{x}_i, y_i)_{i=1}^n$, for what value of (or range of values of) η will the set of optimal parameters be unique?

Indian Institute of Science Banglore
Department of Computational and Data Sciences (CDS)
DS284: Numerical Linear Algebra
Mid-semester Exam 2023

Problem 1 [6x3=18 points]

Assert if the following statements are True or False. Give a detailed reasoning for your assertion. Marks will be awarded only for your reasoning.

(a) Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ with $m > n$ be a matrix with n orthogonal columns. Let $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ be its reduced $\mathbf{QR}$ factorization. Then $\hat{\mathbf{R}}$ is a diagonal matrix.

(b) There can exist a projector $\mathbf{P} \in \mathbb{R}^{m \times m}$ that is an orthogonal matrix but not a symmetric matrix.

(c) If $\mathbf{P} \in \mathbb{R}^{m \times m}$ is an orthogonal projector projecting a vector $\mathbf{v} \in \mathbb{R}^m$ to an n dimensional subspace of $\mathbb{R}^m$ $(n < m)$, then $\mathbf{I} - \mathbf{P}$ has n non-zero singular values.

(d) Two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ are related to each other such that $\mathbf{A} = \mathbf{U}\mathbf{B}\mathbf{V}^{\mathbf{T}}$ for some orthogonal matrices $\mathbf{U}$ and $\mathbf{V}$. Then $\mathbf{A}$ and $\mathbf{B}$ should have the same singular values.

(e) Suppose $\mathbf{A} \in \mathbb{R}^{102 \times 102}$ is such that $||\mathbf{A}||_2 = 50$ and $||\mathbf{A}||_F = 51$. Then the 2-norm condition number $\kappa(\mathbf{A})$ is always greater than or equal to 50.

(f) If a nonzero row is added to $\mathbf{A} \in \mathbb{R}^{m \times n}$ to obtain a new matrix $\mathbf{B} \in \mathbb{R}^{(m+1) \times n}$, then the largest singular value of $\mathbf{B}$ is always greater than or equal to the largest singular value of $\mathbf{A}$.

a) True. $\underset{\sim}{A} = \hat{\underset{\sim}{Q}}\hat{\underset{\sim}{R}}$ is basically orthogonalising the matrix $\underset{\sim}{A}$ such that $\underset{\sim}{Q}$ contains the orthonormal vectors that form a basis for the columns of $\underset{\sim}{A}$, and $\underset{\sim}{R}$ will be an UTM, which will contain the linear combinations of the orthonormal columns in $\underset{\sim}{Q}$ to get back the matrix $\underset{\sim}{A}$.

b) False. A Projector must be symmetric. The only projector which is an orthogonal matrix is the identity matrix. Proof: A projector must fulfill $\underset{\sim}{P}^2 = \underset{\sim}{P}$ & orthogonality implies $\underset{\sim}{P}^{-1} = \underset{\sim}{P}^T$. Because $\underset{\sim}{I}$ is symmetric, this statement is false

$\underset{\sim}{P}^2 = \underset{\sim}{P}$

$\underset{\sim}{P}^{-1} \cdot \underset{\sim}{P} \cdot \underset{\sim}{P} = \underset{\sim}{P}^{-1} \cdot \underset{\sim}{P}$

$\underset{\sim}{I} = \underset{\sim}{I}$ (shown)

c) No. of non-zero singular values = rank of the matrix. If $(\underset{\sim}{I}-\underset{\sim}{P})$ has n non-zero singular values, we know that $\dim(R(\underset{\sim}{I}-\underset{\sim}{P}))=n$. Hence, because $\underset{\sim}{P}$ & $(\underset{\sim}{I}-\underset{\sim}{P})$ project to orthogonal planes, $\dim(\text{Null}(\underset{\sim}{P}))=n$. Hence, $\dim(R(\underset{\sim}{P}))=m-n$. However, because $\underset{\sim}{P}$ maps to an n-dimensional subspace, we know that $\dim(R(\underset{\sim}{P}))=n$. Hence, because $m-n \neq n$, the statement is false.

d) If $\underset{\sim}{A}=\underset{\sim}{U}\underset{\sim}{B}\underset{\sim}{V}^T$, this proves the existence of a similarity transformation s/t $\underset{\sim}{A}$ & $\underset{\sim}{B}$. Hence $\underset{\sim}{A}$ & $\underset{\sim}{B}$ shall have the same eigenvalues (although they might have different eigenvectors). Because the singular values of the matrix are just sq. roots of their respective eigenvalues, $\underset{\sim}{A}$ & $\underset{\sim}{B}$ will have the same singular values.

e) We know that the largest singular value of $\underset{\sim}{A}=50$. We also know that $\sqrt{\sigma_1^2+\cdots+\sigma_n^2}=51$

$$\therefore \|\underset{\sim}{A}\|_2^2 = \|\underset{\sim}{A}\|_F^2 - \sum_{i=2}^{102}\sigma_i^2$$

$$50^2 = 51^2 - \sum_{i=2}^{102}\sigma_i^2$$

$$\sum_{i=2}^{102}\sigma_i^2 = (51-50)(51+101) = 101$$

Hence, the largest value of the smallest eigenvalue will be one. In other words, $\sigma_{102} \leq 1$

$$\therefore \hat{\kappa}^R(\underset{\sim}{A}) = \frac{\sigma_1}{\sigma_{102}} \geq \frac{50}{1}$$

$$\therefore \hat{\kappa}^R(\underset{\sim}{A}) \geq 50 \quad \text{(shown)}$$

f) The largest singular value of a matrix is basically its two-norm. To prove the statement, we have to prove that: $\|\underset{\sim}{A}\|_2 \leq \|B\|_2$

$\|\underset{\sim}{A}\|_2 \leq$

Problem 2 [1+6+5 = 12 points]

Consider the three vectors $\mathbf{a}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}$; $\mathbf{a}_2 = \begin{bmatrix}1\\1\\1\end{bmatrix}$; $\mathbf{a}_3 = \begin{bmatrix}0\\1\\1\end{bmatrix}$.

(a) Verify the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are linearly independent.

(b) Choosing $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ as the basis vectors that span $\mathbb{R}^3$, express the vector $\mathbf{v} = \begin{bmatrix}2\\2\\1\end{bmatrix}$ in the basis of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ by first setting up the required system of equations to be solved and then solve this system of equations by QR factorization.

(c) Construct the projector matrix $\mathbf{P}$ that projects any vector $\mathbf{v} \in \mathbb{R}^3$ orthogonally to the 2-dimensional subspace spanned by the vectors $\mathbf{a}_1, \mathbf{a}_2$. Subsequently find the orthogonal projection of $\mathbf{a}_3$ onto this subspace spanned by $\mathbf{a}_1, \mathbf{a}_2$.

a) To prove that $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$ & $\underset{\sim}{a}_3$ are L.I., we have to prove that the matrix $\underset{\sim}{A} = [\underset{\sim}{a}_1 \ \underset{\sim}{a}_2 \ \underset{\sim}{a}_3]$ is full rank:

Row echelon form by Gaussian elimination.

$$\underset{\sim}{A} = \begin{bmatrix}1&1&0\\0&1&1\\1&1&1\end{bmatrix} \xrightarrow{R_3 \to R_3 - R_1} \begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}$$

Because $\underset{\sim}{A}$ has three non-zero rows in the row echelon form, the three column vectors are L.I.

b) $x_1 \underset{\sim}{a}_1 + x_2 \underset{\sim}{a}_2 + x_3 \underset{\sim}{a}_3 = \underset{\sim}{v}$

$\therefore x_1 + x_2 = 2$ ①

$x_2 + x_3 = 2$ ②

$x_1 + x_2 + x_3 = 1$ ③

$$\begin{bmatrix}1&1&0\\0&1&1\\1&1&1\end{bmatrix} \underset{\sim}{x} = \begin{bmatrix}2\\2\\1\end{bmatrix}$$

From Schmidt orthogonalization:

$\underset{\sim}{v}_1 = \underset{\sim}{a}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix} \Rightarrow \underset{\sim}{q}_1 = \frac{\underset{\sim}{v}_1}{\sqrt{2}} = \sqrt{2}\begin{bmatrix}1\\0\\1\end{bmatrix}$

$\underset{\sim}{v}_2 = \underset{\sim}{a}_2 - (\underset{\sim}{q}_1^T \underset{\sim}{a}_2)\underset{\sim}{q}_1$

$= \begin{bmatrix}1\\1\\1\end{bmatrix} - \sqrt{2}\begin{bmatrix}1&0&1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} \cdot \begin{bmatrix}1\\0\\1\end{bmatrix}$

$= \begin{bmatrix}1\\1\\1\end{bmatrix} - \sqrt{2}\cdot 2 \cdot \begin{bmatrix}1\\0\\1\end{bmatrix}$

$= \begin{bmatrix}1-2\sqrt{2}\\1\\1-2\sqrt{2}\end{bmatrix}$

$\|\underset{\sim}{v}_2\| = \sqrt{2(1-2\sqrt{2})^2}$

$$\begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} & \frac{\sqrt{2}}{2} \\ 0 & 1 & 1 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} x = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

$$\underset{\sim}{Q} \qquad \underset{\sim}{R} \qquad \underset{\sim}{x} = \underset{\sim}{v}$$

$$\underset{\sim}{R}\underset{\sim}{x} = \underset{\sim}{Q}^{-1}\underset{\sim}{v}$$

$$= \underset{\sim}{Q}^{T}\underset{\sim}{v}$$

$$= \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} \sqrt{2} & \sqrt{2} & \frac{\sqrt{2}}{2} \\ 0 & 1 & 1 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} x = \begin{bmatrix} \frac{3\sqrt{2}}{2} \\ 2 \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$$

$$\frac{\sqrt{2}}{2}x_3 = -\frac{\sqrt{2}}{2} \Rightarrow x_3 = -1$$

$$x_2 + x_3 = 2 \Rightarrow x_2 = 3$$

$$\sqrt{2}x_1 + \sqrt{2}x_2 + \frac{\sqrt{2}}{2}x_3 = \frac{3\sqrt{2}}{2}$$

$$x_1 + x_2 + \frac{x_3}{2} = \frac{3}{2} \qquad (3,\ -\tfrac{1}{2})$$

$$x_1 = -1 \qquad \therefore x = \begin{bmatrix} -1 \\ 3/2 \\ -1 \end{bmatrix}$$

$P = A(A^TA)^{-1}A^T$, where $A = [u_1\ u_2] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \left(\underset{2\times3}{\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}} \underset{3\times2}{\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}} \right)^{-1} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix}_{2\times2} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}_{2\times3}$

$= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}_{3\times2} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & 0 \end{bmatrix}_{2\times3}$

$= \frac{1}{2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

Projection of $u_3 = u_3' = Pu_3$

$= \frac{1}{2} \underset{3\times3}{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}} \underset{3\times1}{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}}$

$= \frac{1}{2} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

Problem 3 [3+4+7+2+2+2+6+4 = 30 points]

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 $\frac{\|\delta \mathbf{A}\|}{\|\mathbf{A}\|} = O(\epsilon_{mach})$. 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).

a) Singular values of a matrix $\underset{\sim}{A}$ are basically sq. roots of eigenvalues of $(\underset{\sim}{A}\underset{\sim}{A}^T)$. In this case, because $A^T = A$ the singular values are sq. roots of the eigenvalues of A^2. Hence if A has $\lambda_1, \ldots \lambda_n$ as its eigenvalues, A^2 will have $\lambda_1^2 \ldots \lambda_n^2$ as its eigenvalues. Therefore the singular values should be $\sqrt{\lambda_1^2} \ldots \sqrt{\lambda_n^2} = |\lambda_1| \ldots |\lambda_n|$.

b) Let $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{u}$ From Cauchy-Schwarz inequality, $|\underset{\sim}{x}^T \underset{\sim}{A}\underset{\sim}{x}| = |\underset{\sim}{x}^T \underset{\sim}{u}|$

$\leq \|\underset{\sim}{x}^T\|_2 \|\underset{\sim}{u}\|_2$

$\therefore |\underset{\sim}{x}^T \underset{\sim}{A}\underset{\sim}{x}| \leq \|\underset{\sim}{A}\underset{\sim}{x}\|_2$, we know that $\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|$

Hence, $|x^T Ax| \leq \|Ax\|_2 \leq \|A\|_2 \Rightarrow |x^T Ax| \leq \|A\|_2$ (shown)

c) $A u = \lambda u$

$u^T A u = u^T \lambda u$

$\therefore u^T A u = \lambda u^T u$ (1)

$\lambda = u^T A u$

$\lambda + \delta\lambda = (u^T + \delta u^T)(A + \delta A)(u + \delta u)$

$= (u^T A + \delta u^T A + u^T \delta A + \delta u^T \delta A)(u + \delta u)$

$\lambda + \delta\lambda = (\lambda + \delta u^T A u + u^T \delta A u + \delta u^T \delta A u + u^T A \delta u + \delta u^T A \delta u + u^T \delta A \delta u + \delta u^T \delta A \delta u)$

$\lambda + \delta\lambda = \lambda + \delta u^T \lambda u + u^T \delta A u + u^T A \delta u$

$= \lambda + \lambda \delta u^T u + u^T \delta A u + u^T A \delta u$

$\lambda + \delta\lambda = \lambda + u^T \delta A u$

$\therefore \delta\lambda = u^T \delta A u$

$|\delta\lambda| = |u^T \delta A u| \leq \|\delta A\|_2$ (Cauchy)

d) relative condn. no $= \frac{|\delta\lambda|}{|\lambda|} \div \frac{\|\delta A\|_2}{\|A\|_2}$

$= \frac{|\delta\lambda| \|A\|_2}{|\lambda| \|\delta A\|} \leq \frac{\|A\|_2}{|\lambda|}$

$\leq \frac{\sigma_1}{|\lambda|}$

$\leq 1 \Rightarrow$ Very well condn. problem.

e) $\kappa(M) = \frac{\|M\|_2}{|\lambda|} = \frac{a}{a} = 1$

Indian Institute of Science Banglore
Department of Computational and Data Sciences (CDS)
DS284: Numerical Linear Algebra
Mid-semester Exam 2022

Problem 1 [5x3=15 points]

Assert if the following statements are True or False. Give detailed reasoning for your assertion. Marks will be awarded only for your reasoning.

(a) If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ $(n > 2)$ are two non-zero, non-parallel vectors, the determinant of the matrix $\mathbf{A} = \mathbf{u}\mathbf{v}^T + \mathbf{v}\mathbf{u}^T$ is always zero.

(b) Consider the system of equations $\mathbf{Px} = \mathbf{b}$ and $\mathbf{P}^T\mathbf{Px} = \mathbf{P}^T\mathbf{b}$, where $\mathbf{P} \in \mathbb{R}^{m\times n}$ with $m > n$ and $\mathbf{b} \in \mathbb{R}^m$. Assert the statements: (a) Both systems of equations do not always have a solution. (b) Both systems of equations have a unique solution if at all they have a solution. (c) Both systems of equations always have the same solution if they have a solution.

(c) Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ one can always find a matrix $\mathbf{B} \in \mathbb{R}^{m\times m}$ such that $\mathbf{AB} - \mathbf{BA} = \mathbf{I}$.

(d) For a matrix $\mathbf{F} = \mathbf{I}_m - 2\mathbf{qq}^T$, where $\mathbf{q} \in \mathbb{R}^m$ and m is an odd integer. Then the determinant of $\mathbf{F}$ is 1.

(e) Let $fl : \mathbb{R}^+ \to \mathbb{F}$ be a function that takes positive real numbers $x \in \mathbb{R}^+$ as input and returns the **nearest** floating point number approximation $x' \in \mathbb{F}$ (Assume that all inputs are representable as normalized floating point numbers and there is no risk of overflow/underflow in any of the operations). Then for $a, b, c \in \mathbb{R}^+$, $a + b = c \implies fl(fl(a) + fl(b)) = fl(c)$.

a) True. If two vectors are not parallel, the matrix $\underset{\sim}{u}\underset{\sim}{v}^T$ & $\underset{\sim}{v}\underset{\sim}{u}^T$ will be rank 2. Hence, the max. rank of $\underset{\sim}{A} = 2$. Because $\underset{\sim}{A} \in \mathbb{R}^{n\times n}$, where $n > 2$, $\underset{\sim}{A}$ will not be full-rank. Therefore it will have determinant = 0.

b)i) False. If $\underset{\sim}{P}$ is full rank, it is very possible that $\underset{\sim}{P}\underset{\sim}{x} = \underset{\sim}{b}$ does not have a soln. However $\underset{\sim}{P}^T\underset{\sim}{P}\underset{\sim}{x} = \underset{\sim}{P}^T\underset{\sim}{b}$ may have a solution. Basically, if $\underset{\sim}{P}\underset{\sim}{x} = \underset{\sim}{b}$ does not have a solution, it implies that $\underset{\sim}{b}$ does not lie in the column space of $\underset{\sim}{P}$. When we do $\underset{\sim}{P}^T\underset{\sim}{P}\underset{\sim}{x} = \underset{\sim}{P}^T\underset{\sim}{b}$, we shall be projecting $\underset{\sim}{b}$ onto the column space of $\underset{\sim}{P}$, thereby resulting in the presence of a solution.

b) ii) False, if the rank of R is less than n, the problem should have a 'family' of solutions, rather than a unique solution.

b) iii) True. If there is a unique solution to $Rx = b$, it means that b lies in the column space of R. Hence, when we project b onto R's column space, it shall maintain the same value. This is because we know that if we try to project a vector to a plane when it is already in the plane, it shall stay put.

c) False. Proof by counter example:

If $A = 0$, then $AB - BA = 0 \neq I$.

d) $\det(I - yy^T) = \det(I) - \det(yy^T)$

$= m - 0$

$= m$

and m may not be 1.

e) $fl(fl(a) + fl(b))$

$= fl(a(1+\varepsilon_1) + b(1+\varepsilon_2))$ Hence, the statement is false.

$= [a(1+\varepsilon_1) + b(1+\varepsilon_2)](1+\varepsilon_3)$

$= a(1+\varepsilon_1)(1+\varepsilon_3) + b(1+\varepsilon_2)(1+\varepsilon_3)$

$= a(1+\varepsilon_1+\varepsilon_3+\cancel{\varepsilon_1\varepsilon_3}^{0}) + b(1+\varepsilon_2+\varepsilon_3+\cancel{\varepsilon_2\varepsilon_3}^{0})$

$= a(1+\varepsilon_4) + b(1+\varepsilon_5)$

$= a + a\varepsilon_4 + b + b\varepsilon_5$

$= a + b + a\varepsilon_4 + b\varepsilon_5$

$= c + a\varepsilon_4 + b\varepsilon_5$

$\neq c(1+\varepsilon_6)$

Problem 2 [2+1+4 =7 points]

Consider three vectors

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{a}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

(a) Construct two orthonormal vectors $\mathbf{q}_1$ and $\mathbf{q}_2$ such that they span the successive column subspaces of the matrix $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2]$

(b) Does the vector $\mathbf{v}$ lie in the space spanned by $\mathbf{a}_1$ and $\mathbf{a}_2$? Explain.

(c) If the answer for (b) is yes, then solve for the linear combination coefficients of $\mathbf{a}_1$ and $\mathbf{a}_2$, when $\mathbf{v}$ is expressed as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$. If the answer for (b) is no, then compute the projection of $\mathbf{v}$ onto the space spanned by $\mathbf{a}_1$ and $\mathbf{a}_2$ and thereby compute the solution $\mathbf{x}$ which minimizes $||\mathbf{Ax} - \mathbf{v}||_2$

(a) $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$ $\quad v_1 = a_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$q_1 = \frac{v_1}{\|v_1\|_2} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$v_2 = a_2 - (q_1^T a_2) q_1$

$= \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$= \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - \frac{4}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$= \begin{bmatrix} -1/3 \\ 2/3 \\ -1/3 \end{bmatrix}$

$= \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$

$\|v_2\|_2 = \frac{1}{3} \sqrt{1 + 4 + 1} = \frac{\sqrt{6}}{3} =$

$$\therefore \underset{\sim}{q}_2 = \frac{\underset{\sim}{v}_2}{\|\underset{\sim}{v}_2\|_2} = \frac{3}{\sqrt{6}} \cdot \frac{1}{3} \cdot \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

b) $x_1 \underset{\sim}{u}_1 + x_2 \underset{\sim}{u}_2 = \underset{\sim}{v}$

$x_1 + x_2 = 2$ (1)
$x_1 + 2x_2 = 2$ (2)
$x_1 + x_2 = 1$ (3)

Inconsistent system.

Hence $\underset{\sim}{v}$ does not lie in the column space spanned by $\underset{\sim}{a}_1$ & $\underset{\sim}{a}_2$.

c) Least square soln.

$$\underset{\sim}{A}^T \underset{\sim}{A} \underset{\sim}{x} = \underset{\sim}{A}^T \underset{\sim}{v}$$

$$\underset{\sim}{A}^T \underset{\sim}{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix}$$

$$\underset{\sim}{A}^T \underset{\sim}{v} = \underset{2\times 3}{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix}} \underset{3\times 1}{\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix} \underset{\sim}{x} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 3 & 4 & 5 \\ 4 & 6 & 7 \end{array}\right]$$

$R_2 \to R_2 - \frac{4}{3} R_1$

$$\left[\begin{array}{cc|c} 3 & 4 & 5 \\ 0 & 2/3 & 1/3 \end{array}\right]$$

$\therefore x_2 = 1/3 \div 2/3 = 1/2$

$3x_1 + 4x_2 = 5$

$x_1 = \frac{5-2}{3} = 1$

$\therefore \underset{\sim}{x} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$

Problem 3 [5+3=8 points]

Suppose you are confronted with solving a linear system of equations $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is a symmetric non-singular matrix. Further, you would like to use an iterative solver to solve this system of equations as the matrix size is huge.

Remember that an iterative solver starts with some initial guess $\mathbf{x}_0$ and tries to seek sequence of approximations to $\mathbf{x}^*$ (the exact solution of $\mathbf{Ax} = \mathbf{b}$), in an iterative fashion so that the sequence of iterates $(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n, \ldots)$ converge to $\mathbf{x}^*$, within a given tolerance for large n. You usually specify the tolerance ε_{tol} on the norm of the relative residual $\varepsilon = \frac{||\mathbf{r}||}{||\mathbf{b}||} = \frac{||\mathbf{b}-\mathbf{Ax}||}{||\mathbf{b}||}$ for convergence hoping that $\mathbf{x}_i$ at i^{th} iteration will be close enough to $\mathbf{x}^*$ if ε_{tol} is small enough. For the problem at hand, you would like to achieve a norm of relative error in the solution $\eta = \frac{||\mathbf{x}_i-\mathbf{x}^*||}{||\mathbf{x}^*||}$ to be $\eta_{ach} = 10^{-6}$ and for this you specify ε_{tol} in your iterative solver to be 10^{-12}. Another piece of information you know about the matrix $\mathbf{A}$ is that its 2-norm is 10 and its smallest eigenvalue is 10^{-7}. Now answer the following

(a) Derive a relationship between ε and η in terms of property of the matrix $\mathbf{A}$.

(b) Based on the result derived in (a) and the other information you have about $\mathbf{A}$, will you achieve $\eta_{ach} = 10^{-6}$ if you prescribe $\varepsilon_{set} = 10^{-12}$

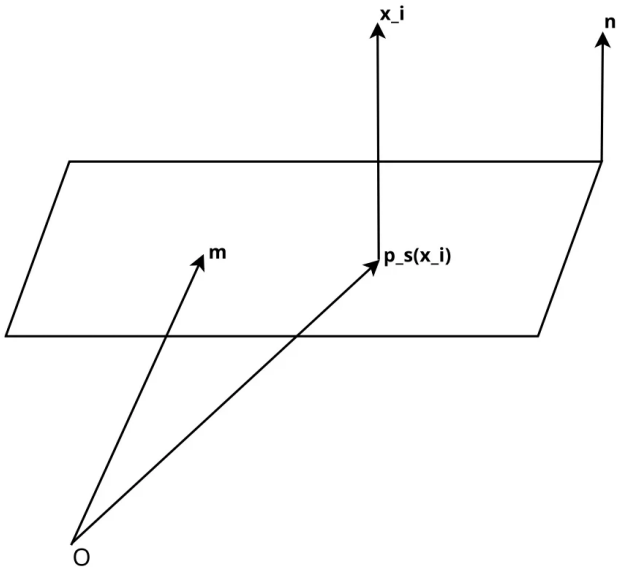
a) $\varepsilon = \dfrac{\|b - A\hat{x}\|}{\|b\|_2} = \dfrac{\|Ax^* - Ax_i\|_2}{\|Ax^*\|_2}$

$= \dfrac{\|A(x^* - x_i)\|_2}{\|Ax^*\|_2}$

$\leq \dfrac{\|A\|_2 \|x^* - x_i\|_2}{\|Ax^*\|_2}$

Problem 4 [4+3+3+2+3+5 =20 points]

You have n data points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^3$ corresponding to some measurements obtained in an experiment. The following exercise seeks you to derive an expression for the best fit 2-D plane $\mathbb{S}$, which minimizes the sum of squares of orthogonal distances from each of the n data points to the best-fit plane. In other words, we are seeking to minimize the orthogonal fitting errors.



(a) Assume $\mathbf{m} \in \mathbb{R}^3$ be a point on the plane $\mathbb{S}$. Let $\mathbf{n}$ be the unit vector in the direction orthogonal to the candidate plane $\mathbb{S}$. If $P_S(\mathbf{x}_i)$ is the point on this plane which is the orthogonal projection of $\mathbf{x}_i$ onto the plane $\mathbb{S}$, derive an expression for $P_S(\mathbf{x}_i)$ in terms of $\mathbf{m}$, $\mathbf{x}_i$ and $\mathbf{n}$. Write this expression for $P_S(\mathbf{x}_i)$ in terms of the orthogonal projector $\mathbf{I} - \mathbf{n}\mathbf{n}^T$ (Hint: use the fact that any vector on the plane $\mathbb{S}$ is orthogonal to $\mathbf{n}$)

(b) Let $\mathbf{B} \in \mathbb{R}^{3\times 2}$ be a matrix whose column vectors $\mathbf{b}_1, \mathbf{b}_2$ span the 2-dimensional vector space orthogonal to $\mathbf{n}$. Derive the expression for $\mathbf{n}\mathbf{n}^T$ in terms of the matrix $\mathbf{B}$ and hence write the final expression for $P_S(\mathbf{x}_i)$ in terms of $\mathbf{x}_i$, $\mathbf{m}$ and the matrix $\mathbf{B}$.

(c) Recall that our objective is to find the best-fit plane that minimizes the sum of squares of the Euclidean distances between each data point and its corresponding orthogonal projection onto the candidate plane $\mathbb{S}$. Pose this problem mathematically using the expression of $P_S(\mathbf{x}_i)$ derived in (b).

(d) Note that the minimization problem posed in (c) has to be minimized with respect to the point $\mathbf{m} \in \mathbb{R}^3$ and the orthogonal matrix $\mathbf{B} \in \mathbb{R}^{3\times 2}$. Using the fact that optimal $\mathbf{m}$ for any given $\mathbf{B}$ is of the form $\mathbf{m}^* = \frac{1}{n}\sum \mathbf{x}_i$, rewrite the minimization problem in (c) in terms of $\widetilde{\mathbf{x}}_i = \mathbf{x}_i - \mathbf{m}^*$

(e) Rewrite the minimization problem in terms of the matrix $\widetilde{\mathbf{X}} \in \mathbb{R}^{n\times 3}$ such that $\widetilde{\mathbf{X}} = [\widetilde{\mathbf{x}}_1\ \widetilde{\mathbf{x}}_2\ \ldots\ \widetilde{\mathbf{x}}_n]^T$ (Hint: Use the definition of Frobenius norm of a matrix in terms of the columns of the matrix)

(f) Comment about the minimization problem obtained in (e) by exploring the connections to the low-rank approximation of the given data matrix $\widetilde{\mathbf{X}}$ and thereby deduce that the minimizer $\mathbf{B}^*$ has to be related to the singular vectors of $\widetilde{\mathbf{X}}$.

Indian Institute of Science Banglore
Department of Computational and Data Sciences (CDS)
DS284: Numerical Linear Algebra
Mid-semester Exam 2021

Problem 1 [6x3=18 points]

Assert if the following statements are True or False. Give a detailed reasoning for your assertion. Marks will be awarded only for your reasoning.

(a) The accuracy of the solution $\mathbf{x}$ for the square system of equations $\mathbf{Ax} = \mathbf{b}$ only depends on (i) the condition number of $\mathbf{A}$ and (ii) the precision of arithmetic operations carried out.

(b) Let $\mathbf{P} \in \mathbb{R}^{n \times n}$ be an orthogonal matrix with $det(\mathbf{P}) = -1$. Then the matrix $\mathbf{P} + \mathbf{I}_n$ is not invertible (i.e a singular matrix).

(c) Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be of $rank(\mathbf{A}) = r (< \min(m, n))$. If $\mathbf{A} = \mathbf{U\Sigma V^T}$ be the full SVD of $\mathbf{A}$, then $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r\}$, the first r columns of $\mathbf{V}$ form an orthogonal basis for $null(\mathbf{A}^T)$

(d) Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be a matrix with orthogonal columns. Let $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ be its reduced $\mathbf{QR}$ factorization. Then $\hat{\mathbf{R}}$ is a diagonal matrix.

(e) If $\hat{\mathbf{x}} \in \mathbb{R}^n$ satisfies the system of equations $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a full rank matrix with $m > n$ and $\mathbf{b} \in \mathbb{R}^m$, then $\hat{\mathbf{x}}$ always satisfies the system of equations $\mathbf{Ax} = \mathbf{b}$

(f) Given a matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$ one can always find a matrix $\mathbf{B} \in \mathbb{R}^{m \times m}$ such that $\mathbf{AB} - \mathbf{BA} = \mathbf{I}$

(g) The following matrix $\mathbf{P}$ is a projector matrix

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

a) Assuming that system of equations $\tilde{A}\tilde{x} = \tilde{b}$ has a unique soln, the statement is true.

b)

c) False. The first r columns of V form a basis for range($\underset{\sim}{A}$), not null($\underset{\sim}{A}$). This is because the last (n-r) singular values are 0.

d) True. Because $\underset{\sim}{A}$ is an orthogonal matrix, we can just normalise the cols. of $\underset{\sim}{A}$ to get $\underset{\sim}{Q}$. We can put the 2-norm of each col. on the diagonal of $\underset{\sim}{R}$.

In math:

$$\text{let } \underset{\sim}{A} = [\underset{\sim}{a}_1 \cdots \underset{\sim}{a}_n] = \left[\frac{\underset{\sim}{a}_1}{\|\underset{\sim}{a}_1\|_2} \cdots \frac{\underset{\sim}{a}_n}{\|\underset{\sim}{a}_n\|_2}\right] \begin{bmatrix} \|\underset{\sim}{a}_1\|_2 & & 0 \\ & \ddots & \\ 0 & & \|\underset{\sim}{a}_n\|_2 \end{bmatrix}$$

e) False. If $\hat{\underset{\sim}{x}}$ satisfies $\underset{\sim}{A}^T\underset{\sim}{A}\hat{\underset{\sim}{x}} = \underset{\sim}{A}^T\underset{\sim}{b}$, this implies that $\hat{\underset{\sim}{x}}$ is the least square solution to the problem $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b}$. This soln. minimises the 2-norm of residual $\underset{\sim}{r}$, where $\underset{\sim}{r} = \underset{\sim}{b} - \underset{\sim}{A}\underset{\sim}{x}$.
We can also say that $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b}$ is definetly over determined, because $\underset{\sim}{A}$ is full rank & $m > n$. Hence, $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b}$ does not have a unique solution

f) False. Because range($\underset{\sim}{A}\underset{\sim}{B}$) = range($\underset{\sim}{B}\underset{\sim}{A}$), range($\underset{\sim}{A}\underset{\sim}{B} - \underset{\sim}{B}\underset{\sim}{A}$) $\neq$ range. Hence $\underset{\sim}{A}\underset{\sim}{B} - \underset{\sim}{B}\underset{\sim}{A} = \underset{\sim}{I}$

g) False. The only projector matrix which is full rank is $\underset{\sim}{I}$. Since $\underset{\sim}{P} \neq \underset{\sim}{I}$, and $\underset{\sim}{P}$ is full rank, $\underset{\sim}{P}$ is not a projector

$$P^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \leftarrow \text{(not equal to P)}$$

Problem 2 [8 points]

Consider the three vectors $\mathbf{a}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$; $\mathbf{a}_2 = \begin{bmatrix}1\\1\\1\end{bmatrix}$; $\mathbf{a}_3 = \begin{bmatrix}0\\1\\1\end{bmatrix}$

(a) Construct the matrix $\mathbf{A}$ whose columns are the three vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. Show that the *range*$(\mathbf{A})$ spans $\mathbb{R}^3$ [3 points]

(b) Construct the successive orthogonal transformations $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3$ which transforms $\mathbf{A}$ to an upper triangle matrix $\mathbf{R}$ as shown below:

$$\mathbf{Q}_3\mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \mathbf{R}$$

(Write down explicitly the matrices $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3$ in your answer sheet) [5 points]

a) $\underset{\sim}{A} = \begin{bmatrix}1 & 1 & 0\\ 1 & 1 & 1\\ 0 & 1 & 1\end{bmatrix}$ $\det(\underset{\sim}{A}) = \det\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} - \det\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} = 1$

$\therefore$ Because $\underset{\sim}{A}$ is non-singular, it has does not have a non-trivial null space. Hence, the cols. of $\underset{\sim}{A}$ span $\mathbb{R}^3$.

b) $\underset{\sim}{x_1} = \underset{\sim}{A}(1{:}m, 1) = \begin{bmatrix}1\\1\\0\end{bmatrix}$ $\underset{\sim}{v_1} = \operatorname{sgn}(x_{11}) \cdot \|\underset{\sim}{x_1}\|_2 \, e_1 + \underset{\sim}{x_1}$

$= (+1)(\sqrt{2})\begin{pmatrix}1\\0\\0\end{pmatrix} + \begin{pmatrix}1\\1\\0\end{pmatrix}$

$= \begin{pmatrix}1+\sqrt{2}\\1\\0\end{pmatrix}$

$\|\underset{\sim}{v_1}\| = \sqrt{1+2\sqrt{2}+2+1} = \sqrt{4+2\sqrt{2}}$

$\underset{\sim}{v_1} = \frac{\underset{\sim}{v_1}}{\|v_1\|_2} = \frac{1}{\sqrt{4+2\sqrt{2}}}\begin{pmatrix}1+\sqrt{2}\\1\\0\end{pmatrix}$

Problem 3 [24 points]

Consider a rank deficient matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with $m > n$ and a vector $\mathbf{b} \in \mathbb{R}^m$. Let $rank(\mathbf{A}) = r (< n)$. Now answer the following questions:

(a) Recall in the class we derived $\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$ minimizes the least square error $||\mathbf{b} - \mathbf{A}\mathbf{x}||_2$. Is this solution $\mathbf{x}$, a valid minimizer of $||\mathbf{b} - \mathbf{A}\mathbf{x}||_2$ for the problem at hand where $\mathbf{A}$ is rank deficient. Explain why/why not. [3 points]

(b) If $\mathbf{b}$ lies in the $range(\mathbf{A})$, explain why the system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ should have infinite solutions? [3 points]

(c) Let $\mathbf{U_1}\mathbf{\Sigma_1}\mathbf{V_1}^T$ be the reduced SVD of our rank deficient matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with $\mathbf{U_1} \in \mathbb{R}^{m \times r}$ $\mathbf{\Sigma_1} \in \mathbb{R}^{r \times r}$ and $\mathbf{V_1} \in \mathbb{R}^{n \times r}$, then show that $\mathbf{x} = \mathbf{V_1}\mathbf{\Sigma_1}^{-1}\mathbf{U_1}^T\mathbf{b}$ satisfies $\mathbf{A}\mathbf{x} = \mathbf{b}$ in the case where $\mathbf{b}$ lies in the $range(\mathbf{A})$. [5 points]

(d) Show that above solution $\mathbf{x} = \mathbf{V_1}\mathbf{\Sigma_1}^{-1}\mathbf{U_1}^T\mathbf{b}$ is the minimum L^2 norm solution out of all possible solutions that satisfy $\mathbf{A}\mathbf{x} = \mathbf{b}$ in the case where $\mathbf{b}$ lies in the $range(\mathbf{A})$ (Hint: Express $\mathbf{x} \in \mathbb{R}^n$ in the orthonormal basis of $\mathbf{V}$) [7 points]

(e) Let us consider the case where $\mathbf{b}$ does not lie in the $range(\mathbf{A})$. Now, to find the least square solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$, we need to solve $\mathbf{A}\mathbf{x} = \mathbf{P}\mathbf{b}$, where $\mathbf{P}$ is the orthogonal projector onto the $range(\mathbf{A})$. Comment on the solvability of the system $\mathbf{A}\mathbf{x} = \mathbf{P}\mathbf{b}$. Show that the solution $\mathbf{x} = \mathbf{V_1}\mathbf{\Sigma_1}^{-1}\mathbf{U_1}^T\mathbf{b}$ corresponds to the minimum L^2 norm least square solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$ [6 points]

Note for Problem 3: Answer all the above questions keeping in mind $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a rank deficient matrix with $m > n$. In data science, least squares problems of the above kind arise in the form of inverse problems relevant to various domain areas such as electromagnetic scattering, geophysics, image restoration, compressed sensing (reconstruction of sparse signal from small number of random measurements), inverse scattering, medical imaging and the study of atmospheres etc.,

a) If $\underset{\sim}{A}$ is not full rank, it means that the problem $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b}$ is under determined. Hence $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b}$ does not have a unique soln. Instead, it has a "family" of solutions. Hence, it is impossible to come up with a single solution that can minimise the value of $\|\underset{\sim}{b} - \underset{\sim}{A}\underset{\sim}{x}\|$

b) If $\underset{\sim}{A}$ is not full rank, i.e rank $(\underset{\sim}{A}) <$ min (m, n), the system is said to be under determined. In other words $\underset{\sim}{A}$ has a non-trivial null space. Hence, the system is bound to produce an infinite "family" of solutions.

c) $\underset{\sim}{b} = \underset{\sim}{A}\underset{\sim}{x}$

$= \underset{\sim}{U_1}\underset{\sim}{\Sigma_1}\underset{\sim}{V_1^T}(\underset{\sim}{V_1}\underset{\sim}{\Sigma_1^{-1}}\underset{\sim}{U_1^T})\underset{\sim}{b}$

$= \underset{\sim}{U_1}\underset{\sim}{\Sigma_1}\underset{\sim}{V_1^T}\underset{\sim}{V_1}\underset{\sim}{\Sigma_1^{-1}}\underset{\sim}{U_1^T}\underset{\sim}{b}$ (cols of $\underset{\sim}{V_1}$ are orthonormal

$= \underset{\sim}{U_1}\underset{\sim}{\Sigma_1}\underset{\sim}{\Sigma_1^{-1}}\underset{\sim}{U_1^T}\underset{\sim}{b}$ ($\underset{\sim}{\Sigma_1}$ is non-singular)

$= \underset{\sim}{U_1}\underset{\sim}{U_1^T}\underset{\sim}{b}$ ($\underset{\sim}{V_1}$ has orthogonal columns also)

$= \underset{\sim}{b}$ (shown)

d) $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{P}\underset{\sim}{b}$ will be solvable because $\underset{\sim}{P}$ ensures that $\underset{\sim}{x}$ lies in the space spanned by the columns of $\underset{\sim}{A}$.