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

DS284: Numerical Linear Algebra

Final Exam – Aug 2023 Term

Faculty Instructor: Dr. Phani Motamarri
TAs: Kartick Ramakrishnan, Sayan Dutta, Sundaresan G

Duration: 9:00 AM to 12:00 noon | Max Points: 100

Notations: (i) Vectors and matrices are denoted by bold faced lower case and upper case alphabets respectively. (ii) Set of all real numbers is denoted by $\mathbb{R}$ (iii) Set of all n dimensional column vectors is denoted by $\mathbb{R}^n$ and set of all $m \times n$ matrices is denoted by $\mathbb{R}^{m\times n}$. (iv) ϵ_M denotes machine epsilon.

- **There are 5 questions and 3 pages. Start each problem on a new page.**

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) 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 $n << m$ and $\mathbf{b} \in \mathbb{R}^m$, then $\hat{\mathbf{x}}$ always satisfies the system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$.

(b) Let $\mathbf{P},\ \mathbf{Q} \in \mathbb{R}^{1\times(m+1)}$ be two non-zero matrices. Further, let $\mathbf{A} \in \mathbb{R}^{(m+1)\times m}$ be a full rank matrix such that $\mathbf{PA} = \mathbf{QA} = \mathbf{0}_\mathbf{m}^\mathbf{T}$. Then $\mathbf{P}\mathbf{Q}^T \neq 0$. (Note: the notation $\mathbf{0}_\mathbf{m}$ denotes the m-dimensional column vector with each entry zero)

(c) Let $\mathbf{W}$ be a subspace of $\mathbb{R}^n$. There exists some $\mathbf{A} \in \mathbb{R}^{n\times n}$ such that $\mathbf{W} = null(\mathbf{A})$.

(d) Let $\mathbf{E} \in \mathbb{R}^{m\times m}$ extract the "even part" of a vector $\mathbf{x} \in \mathbb{R}^m$, i.e. $\mathbf{Ex} = (\mathbf{x} + \mathbf{Fx})/2$ where $\mathbf{F} \in \mathbb{R}^{m\times m}$ that flips $(x_1, \ldots, x_m)^T$ to $(x_m, \ldots, x_1)^T$. Then $\mathbf{E}$ is an orthogonal projector.

(e) Consider a non-zero symmetric matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ and a non-zero vector $\mathbf{b} \in \mathbb{R}^m$. Let $\mathbf{K}_n = \mathbf{Q}_n\mathbf{R}_n$ be the QR factorization of the n-dimensional Krylov subspace $\mathbf{K}_n = [\mathbf{b} \quad \mathbf{Ab} \quad \mathbf{A}^2\mathbf{b} \ \ldots \ \mathbf{A}^{n-1}\mathbf{b}]$ for $n << m$. As $n \to m$, columns of $\mathbf{Q}_n$ approach the eigenvectors of $\mathbf{A}$.

(f) Let $\mathbf{A} \in \mathbb{R}^{m\times m}$, not necessarily symmetric, and $\mathbf{A_s} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$ be positive definite. $\mathbf{K_n} = \{\mathbf{b}, \mathbf{Ab}, \mathbf{A}^2\mathbf{b}, \ldots, \mathbf{A}^{n-1}\mathbf{b}\}$ denotes n-dimensional Krylov subspace constructed with $\mathbf{b} \in \mathbb{R}^m$ as a starting vector. For solving the linear system $\mathbf{Ax} = \mathbf{b}$, if one seeks a vector $\mathbf{x_1} \in \mathbf{K_1}$ to minimize the residual vector $||r||_2 = ||\mathbf{b} - \mathbf{Ax}||_2$, then $\mathbf{x_1}$ must be in the direction of $\mathbf{b}$.

Problem 2 [6+5+4+5=20 points]

Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be a full rank matrix, and $\mathbf{A} = \mathbf{QR}$ denote the full QR decomposition of $\mathbf{A}$, where $\mathbf{Q} \in \mathbb{R}^{m\times m}$ is an orthogonal matrix and $\mathbf{R} \in \mathbb{R}^{m\times m}$ is an upper triangular matrix. Let $\mathbf{q}_i$ and $\mathbf{a}_i$ denote the i^{th} column of the matrices $\mathbf{Q}$ and $\mathbf{A}$ respectively for $1 \leqslant i \leqslant m$. Now, answer the following questions with clear arguments:

(a) Construct the projector matrix $\mathbf{P}$ that orthogonally projects a vector onto a subspace spanned by the set of vectors $< \mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_{j-1} >$ for some j ($\leqslant m$).

(b) Write down the expression for the complimentary projector to $\mathbf{P}$ obtained above in (a). Show that this complimentary projector is symmetric. Let $\tilde{\mathbf{a}}_j$ be the projection of $\mathbf{a}_j$ obtained using this complementary projector. To this end, write the expression for $\tilde{\mathbf{a}}_j$ in terms of $< \mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_{j-1} >$ and $\mathbf{a}_j$. What subspace does $\tilde{\mathbf{a}}_j$ belong?

(c) Show that the absolute value of diagonal entry of the $\mathbf{R}$ matrix i.e $|\mathbf{R}_{jj}|$ is related to the 2-norm of the vector $\tilde{\mathbf{a}}_j$ obtained in (b).

(d) Using the above results and $\mathbf{QR}$ decomposition of $\mathbf{A}$. Show that: $|det(\mathbf{A})| \leqslant \prod_{j=1}^{m} ||\mathbf{a}_j||_2$

Problem 3 [4+5+8+4=21 points]

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$ be two distinct unit vectors. We have a matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ of the form $\mathbf{A} = \mathbf{I} + \mathbf{u}\mathbf{v}^T$ and this form is known as the rank one perturbation of the identity matrix $\mathbf{I} \in \mathbb{R}^{m\times m}$ and arises in many scientific applications.

(a) Show that $\mathbf{A}$ has an eigenvalue 1 with geometric multiplicity $m - 1$.

(b) If $\mathbf{A}$ is non-singular, show that an eigenvector of $\mathbf{A}$ for a non-zero eigenvalue $\lambda \neq 1$ must be of the form $c\mathbf{u}$ for some scalar c. Find this eigenvalue λ in terms of the vectors $\mathbf{u}$ and $\mathbf{v}$.

(c) If $\mathbf{A}$ is non-singular, derive an expression for $\mathbf{A}^{-1}$ in terms of $\mathbf{u}$ and $\mathbf{v}$. (*Hint:* Use the eigenvector of $\mathbf{A}$ corresponding to non-unit eigenvalue.)

(d) If $\mathbf{v}^T\mathbf{u} = -1$, show that $\mathbf{A}$ is singular. What can you say about the null space of $\mathbf{A}$?

Problem 4 [6+3+6+6+4=25 points]

Regression is one of the important aspects of data science. 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. Least squares approach we discussed in the class is an approach to solve the regression problem where you try to minimize 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 you to derive an expression for the orthogonal regression where the best-fit line $\mathbb{L}$ is obtained by minimizing the squares of orthogonal distances from each of the given data points to the best-fit line. You will also explore its connections to SVD. Say you have N data points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N \in \mathbb{R}^2$ 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 $\mathbf{x}_\mathbf{i}$ 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 $\mathbf{x}_\mathbf{i}$ 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 $\mathbf{m}^* = \frac{1}{N}\sum_i \mathbf{x}_\mathbf{i}$ (you do not need to show now). Using this fact, rewrite the above minimization problem over $\mathbf{n}$ in terms of $\hat{\mathbf{x}}_\mathbf{i} = \mathbf{x}_\mathbf{i} - \mathbf{m}^*$ 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 5 [5+5+6=16 points]

If $\mathbf{A} \in \mathbb{R}^{m\times m}$ is a symmetric invertible matrix with an exceedingly large condition number. To this end, let one eigenvalue of $\mathbf{A}$ be much smaller than others in absolute value, i.e. $|\lambda_1| <<< |\lambda_2| \leqslant |\lambda_3| ... \leqslant |\lambda_m|$ (i.e. assume $\mathbf{A}$ is an ill-conditioned matrix with $\kappa(\mathbf{A})$ on the order of ϵ_M^{-1}). Let $\mathbf{q}_1, \mathbf{q}_2, ...\mathbf{q}_m$ be the corresponding eigenvectors. Now answer the following questions:

(a) If the system of equations $\mathbf{A}\mathbf{w} = \mathbf{v}$ is solved using some backward stable algorithm for a given $\mathbf{v} \in \mathbb{R}^m$ yielding a computed vector $\tilde{\mathbf{w}} = \mathbf{w} + \delta\mathbf{w}$, show that $\delta\mathbf{w} = -(\mathbf{A} + \delta\mathbf{A})^{-1}(\delta\mathbf{A})\mathbf{w}$ for some $\delta\mathbf{A}$ such that $\|\delta\mathbf{A}\| = O(\epsilon_M)\|\mathbf{A}\|$.

(b) Assuming that $\mathbf{v}$ is a vector with components in the directions of all eigenvectors $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_m$ of $\mathbf{A}$, show that $\frac{\mathbf{w}}{\|\mathbf{w}\|_2} \approx \mathbf{q}_1$, where $\mathbf{w} = \mathbf{A}^{-1}\mathbf{v}$ is the exact solution of the system of equations given in (a). [**Hint**: First show that $\mathbf{w}$ is approximately in the direction of $\mathbf{q}_1$]

(c) Using the Taylor series expansion $(\mathbf{A} + \delta\mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\delta\mathbf{A})\mathbf{A}^{-1} + O(\epsilon_M^2)$, show that $\frac{\tilde{\mathbf{w}}}{\|\tilde{\mathbf{w}}\|_2} \approx \mathbf{q}_1$. [**Hint**: Using the expression $\delta\mathbf{w}$ derived in (a) and the fact $\delta\mathbf{A}$ is random roundoff perturbation, show that $\delta\mathbf{w}$ is in the direction of $\mathbf{q}_1$]

[This sequence of steps in (a), (b) and (c) show that, though the computed solution $\tilde{\mathbf{w}}$ is far away from $\mathbf{w}$ for the ill-conditioned system $\mathbf{A}\mathbf{w} = \mathbf{v}$, but $\frac{\tilde{\mathbf{w}}}{\|\tilde{\mathbf{w}}\|_2}$ need not be far from $\frac{\mathbf{w}}{\|\mathbf{w}\|_2}$.]