Homework 3

This homework is due at 11 PM on February 2, 2024.

Submission Format: Your homework submission should consist of a single PDF file that contains all of your answers (any handwritten answers should be scanned).

1. **Miscellaneous**

   (a) Let m, n be positive integers where $m \geq n$, and let $U \in \mathbb{R}^{m \times n}$ be an orthonormal matrix (i.e., having orthonormal columns). Show that for any $\vec{x} \in \mathbb{R}^n$ we have

   $$\|U\vec{x}\|_2 = \|\vec{x}\|_2 . \tag{1}$$

   (b) Using the Gram-Schmidt process, find an orthonormal basis $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ of $\mathbb{R}^3$, where

   $$\vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} . \tag{2}$$

2. Diagonalization and Singular Value Decomposition

Let matrix $A = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

(a) Compute the eigenvalues and associated eigenvectors of A.

(b) Express A as $P \Lambda P^{-1}$, where Λ is a diagonal matrix and $PP^{-1} = I$. State P, Λ, and P^{-1} explicitly.

(c) Compute $\lim_{k \to \infty} A^k$.

(d) Give the singular values σ_1 and σ_2 of A.

3. Distinct Eigenvalues, Orthogonal Eigenspaces

Let $A \in \mathbb{S}^n$ (i.e., the set of $n \times n$ symmetric matrices) and $(\lambda_1, \vec{u}_1), (\lambda_2, \vec{u}_2), \lambda_1 \neq \lambda_2$ be distinct eigen-pairs of A. Show that $\vec{u}_1^\top \vec{u}_2 = 0$, i.e., eigenspaces corresponding to distinct eigenvalues are mutually orthogonal.

HINT: First try to prove that $\lambda_1 \vec{u}_1^\top \vec{u}_2 = \lambda_2 \vec{u}_1^\top \vec{u}_2$, then show that this implies $\vec{u}_1^\top \vec{u}_2 = 0$.

NOTE: This exercise is part of the proof of the spectral theorem.

4. Gram Schmidt

Any set of n linearly independent vectors in $\mathbb{R}^n$ could be used as a basis for $\mathbb{R}^n$. However, certain bases could be more suitable for certain operations than others. For example, an orthonormal basis could facilitate solving linear equations.

(a) Given a matrix $A \in \mathbb{R}^{n\times n}$, it could be represented as a multiplication of two matrices

$$A = QR, \tag{3}$$

where $Q \in \mathbb{R}^{n\times n}$ is an orthonormal matrix and $R \in \mathbb{R}^{n\times n}$ is an upper-triangular matrix. For the matrix A, describe how Gram-Schmidt process could be used to find the Q and R matrices, and apply this to

$$A = \begin{bmatrix} 3 & -3 & 1 \\ 4 & -4 & -7 \\ 0 & 3 & 3 \end{bmatrix} \tag{4}$$

to find an orthonormal matrix Q and an upper-triangular matrix R.

(b) Given an invertible matrix $A \in \mathbb{R}^{n\times n}$ and an observation vector $\vec{b} \in \mathbb{R}^n$, the solution to the equality

$$A\vec{x} = \vec{b} \tag{5}$$

is given as $\vec{x} = A^{-1}\vec{b}$. For the matrix $A = QR$ from part (a), assume that we want to solve

$$A\vec{x} = \begin{bmatrix} 8 \\ -6 \\ 3 \end{bmatrix}. \tag{6}$$

By using the fact that Q is an orthonormal matrix, find $\vec{v}$ such that

$$R\vec{x} = \vec{v}. \tag{7}$$

Then, given the upper-triangular matrix R in part (a) and $\vec{v}$, find the elements of $\vec{x}$ sequentially.

(c) Given an invertible matrix $B \in \mathbb{R}^{n\times n}$ and an observation vector $\vec{c} \in \mathbb{R}^n$, find the computational cost of finding the solution $\vec{z}$ to the equation $B\vec{z} = \vec{c}$ by using the QR decomposition of B. Assume that Q and R matrices are available, and adding, multiplying, and dividing scalars take one unit of "computation".

As an example, computing the inner product $\vec{a}^\top\vec{b}$ is said to be $\mathcal{O}(n)$, since we have n scalar multiplication for each $a_i b_i$. Similarly, matrix vector multiplication is $\mathcal{O}(n^2)$, since matrix vector multiplication can be viewed as computing n inner products. The computational cost for inverting a matrix in $\mathbb{R}^n$ is $\mathcal{O}(n^3)$, and consequently, the cost grows rapidly as the set of equations grows in size. This is why the expression $A^{-1}\vec{b}$ is usually not computed by directly inverting the matrix A. Instead, the QR decomposition of A is exploited to decrease the computational cost.

5. Eigenvectors of a Symmetric Matrix

Let $\vec{p}, \vec{q} \in \mathbb{R}^n$ be two linearly independent vectors, with unit norm ($\|\vec{p}\|_2 = \|\vec{q}\|_2 = 1$). Define the symmetric matrix $A := \vec{p}\vec{q}^\top + \vec{q}\vec{p}^\top$. In your derivations, it may be useful to use the notation $c := \vec{p}^\top \vec{q}$.

(a) Show that A is symmetric.

(b) Show that $\vec{p} + \vec{q}$ and $\vec{p} - \vec{q}$ are eigenvectors of A, and determine the corresponding eigenvalues.

(c) Determine the nullspace and rank of A.

(d) Find an eigenvalue decomposition of A, in terms of $\vec{p}, \vec{q}$. *HINT: Use the previous two parts.*

6. PSD Matrices

In this problem, we will analyze properties of positive semidefinite (PSD) matrices. A *symmetric* matrix $M \in \mathbb{R}^{n\times n}$ is a PSD matrix if $\vec{x}^\top M\vec{x} \geq 0$ for all $\vec{x} \in \mathbb{R}^n$, and we denote that as $M \succeq 0$ or $M \in \mathbb{S}^n_+$.

Assume $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix.

(a) Show that $A \succeq 0$ if and only if all eigenvalues of A are non-negative.

(b) Show that if $A \succeq 0$ then all diagonal entries of A are non-negative, $A_{ii} \geq 0$.

Now we will show that $A \in \mathbb{S}^n_+$ (i.e., $A \in \mathbb{R}^{n\times n}$ and $A \succeq 0$) if and only if there exists $P \in \mathbb{S}^n_+$ such that $A = P^\top P = P^2$. Such a matrix P is known as a *PSD square root* of A.

(c) First, show that if $A \in \mathbb{S}^n_+$, then there exists $P \in \mathbb{S}^n_+$ such that $A = P^2$.

(d) Now, show that for any matrix $Q \in \mathbb{R}^{m\times n}$, if $A = Q^\top Q$ then $A \in \mathbb{S}^n_+$.

NOTE: If we take $Q = P \in \mathbb{S}^n_+$, we have $A = P^2 \in \mathbb{S}^n_+$; this proves the other direction of the above "if-and-only-if."

We will use positive semidefinite matrices to construct the singular value decomposition (SVD). One can construct the SVD of a matrix $B \in \mathbb{R}^{m\times n}$ in two ways, using either the matrices $BB^\top \in \mathbb{R}^{m\times m}$ or $B^\top B \in \mathbb{R}^{n\times n}$, and usually one chooses the method depending on which matrix is smaller. The following result shows that the singular values will be the same no matter what construction you use.

(e) Let $B \in \mathbb{R}^{m\times n}$ be an arbitrary matrix (not necessarily PSD or even square). From the previous parts of this problem, $BB^\top$ and $B^\top B$ are PSD, thus having real non-negative eigenvalues. Prove that the non-zero eigenvalues of $BB^\top$ are the same as the non-zero eigenvalues of $B^\top B$.

7. Homework Process

With whom did you work on this homework? List the names and SIDs of your group members.

NOTE: If you didn't work with anyone, you can put "none" as your answer.