Homework 2

This homework is due at 11 PM on February 2nd, 2023.

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. **Norms**

   (a) Show that the following inequalities hold for any vector $\vec{x} \in \mathbb{R}^n$:

   $$\frac{1}{\sqrt{n}} \|\vec{x}\|_2 \leq \|\vec{x}\|_\infty \leq \|\vec{x}\|_2 \leq \|\vec{x}\|_1 \leq \sqrt{n} \|\vec{x}\|_2 \leq n \|\vec{x}\|_\infty . \tag{1}$$

   As an aside: note that we can interpret different norms as different ways of computing distance between two points $\vec{x}, \vec{y} \in \mathbb{R}^2$. The ℓ_2 norm is the distance as the crow flies (i.e. point-to-point distance), the ℓ_1 norm, also known as the Manhattan distance is the distance you would have to cover if you were to navigate from $\vec{x}$ to $\vec{y}$ via a rectangular street grid, and the ℓ_∞ norm is the maximum distance that you have to travel in either the north-south or the east-west direction.

   (b) We define the *cardinality* of the vector $\vec{x}$ as the number of non-zero elements in $\vec{x}$. This is also commonly known as the ℓ_0 norm of the vector $\vec{x}$, denoted by $\|\vec{x}\|_0$. Show that for any non-zero vector x,

   $$\|\vec{x}\|_0 \geq \frac{\|\vec{x}\|_1^2}{\|\vec{x}\|_2^2}. \tag{2}$$

   Find all vectors $\vec{x}$ for which the lower bound is attained.

2. 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 $\langle \vec{u}_1,\ \vec{u}_2 \rangle = 0$, i.e eigenspaces corresponding to distinct eigenvalues are mutually orthogonal.

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

3. 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

$$A = QR,$$

where Q is an orthonormal in $\mathbb{R}^n$ and R 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

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

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

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

$$Ax = b$$

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

$$Ax = \begin{bmatrix} 8 \\ -6 \\ 3 \end{bmatrix}.$$

By using the fact that Q is an orthonormal matrix, find v such that

$$Rx = v.$$

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

(c) Given an invertible matrix $B \in \mathbb{R}^{n\times n}$ and an observation vector $c \in \mathbb{R}^n$, find the computational cost of finding the solution z to the equation $Bz = c$ 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 $a^\top b$ 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 $A^{-1}b$ is usually not computed by directly inverting the matrix A. Instead, the QR decomposition of A is exploited to decrease the computational cost.

4. 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 $\vec{p} + \vec{q}$ and $\vec{p} - \vec{q}$ are eigenvectors of A, and determine the corresponding eigenvalues.

(b) Determine the nullspace and rank of A.

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

(d) **(OPTIONAL)** Now consider general vectors $\vec{p}_{\text{new}}, \vec{q}_{\text{new}}$ that are scaled versions of $\vec{p}, \vec{q}$. Note that $\vec{p}_{\text{new}}, \vec{q}_{\text{new}}$ are not necessarily norm 1. Define the matrix $A_{\text{new}} := \vec{p}_{\text{new}}\vec{q}_{\text{new}}^\top + \vec{q}_{\text{new}}\vec{p}_{\text{new}}^\top$.

Write A_{new} as a function of $\vec{p}, \vec{q}$ and the norms of $\vec{p}_{\text{new}}, \vec{q}_{\text{new}}$, and the eigenvalues of matrix A_{new} as a function of $\vec{p}, \vec{q}$ and the norms of $\vec{p}_{\text{new}}, \vec{q}_{\text{new}}$.

5. PSD Matrices

In this problem, we will analyze properties of positive semidefinite (PSD) matrices. A matrix M is a PSD matrix if all its eigenvalues are non-negative, and we denote that as $M \succeq 0$.
Assume $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix.

(a) Show that $\forall \vec{x} \in \mathbb{R}^n, \vec{x}^\top A \vec{x} \geq 0 \iff$ all eigenvalues of A are non-negative.

Now we will show that a symmetric matrix A is positive semidefinite if and only if there exists a symmetric matrix $P \in \mathbb{R}^{n \times n}$ such that $A = P^\top P$.

(b) First, show that A having non-negative eigenvalues allows us to decompose $A = P^\top P$ where $P \succeq 0$.

(c) Now, show that any matrix of the form $A = P^\top P$ is positive semidefinite, i.e. $A \succeq 0$.

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

6. 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.