Homework 3

Self grades are due at 11 PM on February 9, 2024.

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

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

Solution: Because U is orthonormal, we have $U^\top U = I$. Thus

$$\|U\vec{x}\|_2^2 = (U\vec{x})^\top (U\vec{x}) = \vec{x}^\top U^\top U\vec{x} = \vec{x}^\top \vec{x} = \|\vec{x}\|_2^2 . \tag{2}$$

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

$$\vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} . \tag{3}$$

Solution: We apply Gram-Schmidt process to the sequence $(\vec{v}_1, \vec{e}_1, \vec{e}_2, \vec{e}_3)$ where $\vec{e}_i$ is the i standard basis vector. Exactly one of them will be $\vec{0}$ after the process, and we can throw it out to obtain an orthonormal basis.

One can compute

$$\|\vec{v}_1\|_2^2 = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2} + \frac{1}{2} = 1, \tag{4}$$

so it is already normalized. To find $\vec{v}_2$ we compute

$$\vec{z}_2 = \vec{e}_1 - \underbrace{(\vec{v}_1^\top \vec{e}_1)}_{=1/\sqrt{2}} \vec{v}_1 \tag{5}$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \tag{6}$$

$$= \begin{bmatrix} 1/2 \\ -1/2 \\ 0 \end{bmatrix} \tag{7}$$

$$\vec{v}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|_2} \tag{8}$$

$$= \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} . \tag{9}$$

To find $\vec{v}_3$ we compute

$$\vec{z}_3 = \vec{e}_2 - \underbrace{(\vec{v}_1^\top \vec{e}_2)}_{=1/\sqrt{2}} \vec{v}_1 - \underbrace{(\vec{v}_2^\top \vec{e}_2)}_{=-1/\sqrt{2}} \vec{v}_2 \tag{10}$$

$$= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} + \frac{1}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} \tag{11}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0}. \tag{12}$$

We hit a $\vec{0}$ in the Gram-Schmidt process; we throw it away and move on to the next vector to compute $\vec{v}_3$.

$$\vec{z}_4 = \vec{e}_3 - \underbrace{(\vec{v}_1^\top \vec{e}_3)}_{=0} \vec{v}_1 - \underbrace{(\vec{v}_2^\top \vec{e}_3)}_{=0} \vec{v}_2 \tag{13}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{14}$$

$$\vec{v}_3 = \frac{\vec{z}_4}{\|\vec{z}_4\|_2} = \vec{z}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{15}$$

Thus we have that

$$(\vec{v}_1, \vec{v}_2, \vec{v}_3) = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}. \tag{16}$$

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.

Solution: Eigenvalues can be computed by first calculating A's characteristic polynomial:

$$\det(sI - A) = \det\left(\begin{bmatrix} s & -1 \\ -\frac{1}{2} & s - \frac{1}{2} \end{bmatrix}\right) \tag{17}$$

$$= s\left(s - \frac{1}{2}\right) - (-1)\left(-\frac{1}{2}\right) \tag{18}$$

$$= s^2 - \frac{1}{2}s - \frac{1}{2} \tag{19}$$

$$= \left(s - \frac{1}{4}\right)^2 - \frac{1}{16} - \frac{1}{2} \tag{20}$$

$$= \left(s - \frac{1}{4}\right)^2 - \frac{9}{16} \tag{21}$$

$$= \left(s - \frac{1}{4} - \frac{3}{4}\right)\left(s - \frac{1}{4} + \frac{3}{4}\right) \qquad a^2 - b^2 = (a-b)(a+b) \tag{22}$$

$$= (s-1)\left(s + \frac{1}{2}\right). \tag{23}$$

The eigenvalues of A are thus $\lambda_1 = 1$ and $\lambda_2 = -\frac{1}{2}$, the values of s at which $\det(sI - A) = 0$.

The eigenvectors associated with each eigenvalue λ can be calculated as values of $\vec{x} = \begin{bmatrix} x_a \\ x_b \end{bmatrix}$ for which $A\vec{x} = \lambda\vec{x}$, namely:

$$A\vec{x} = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_a \\ x_b \end{bmatrix} = \begin{bmatrix} x_b \\ \frac{x_a + x_b}{2} \end{bmatrix} \tag{24}$$

$$A\vec{x}_1 = \vec{x}_1 \iff x_b = x_a \iff \vec{x}_1 = \alpha_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \alpha_1 \neq 0 \in \mathbb{R}. \tag{25}$$

$$A\vec{x}_2 = -\frac{1}{2}\vec{x}_2 \iff x_b = -\frac{1}{2}x_a \iff \vec{x}_2 = \alpha_2 \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix}, \ \alpha_2 \neq 0 \in \mathbb{R}. \tag{26}$$

Note that the expressions above are valid eigenvectors for any nonzero values of α_1 and α_2.

(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.

Solution: Combining the calculations in part (a), we have that

$$A \begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \vec{x}_1 & \lambda_2 \vec{x}_2 \end{bmatrix} = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}. \tag{27}$$

For our calculations, we will use the eigenvalues and eigenvectors from part (a) with $\alpha_1 = \alpha_2 = 1$. (Your calculations may differ here; any nonzero values for α_1 and α_2 are permissible, and will result in scaled values of P and P^{-1}.) Filling in eigenvalue and eigenvector values, we have:

$$A \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}, \tag{28}$$

and rearranging,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}^{-1}. \tag{29}$$

Calculating the latter inverse explicitly, we have

$$\begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}^{-1} = -\frac{2}{3} \begin{bmatrix} -\frac{1}{2} & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix} \text{ because } \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{30}$$

so finally,

$$A = P\Lambda P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix}. \tag{31}$$

This is known as the *eigenvalue decomposition*, or *eigendecomposition*, of matrix A; for a more extensive description of this decomposition, see Calafiore & El Ghaoui section 3.5.

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

Solution: Using the diagonalization of A from part (b), we have:

$$A = P\Lambda P^{-1} \tag{32}$$
$$A^k = (P\Lambda P^{-1})^k \tag{33}$$
$$= (P\Lambda P^{-1})(P\Lambda P^{-1})\dots(P\Lambda P^{-1}) \qquad (k \text{ times}) \tag{34}$$
$$= P\Lambda \underbrace{P^{-1}P}_{I} \Lambda P^{-1} \dots P\Lambda P^{-1} \tag{35}$$
$$= P\Lambda^k P^{-1} \tag{36}$$
$$= \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}^k \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix} \tag{37}$$
$$= \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & \left(-\frac{1}{2}\right)^k \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix}. \tag{38}$$

Finally, because $\lim_{k\to\infty} \left(-\frac{1}{2}\right)^k = 0$, we have

$$\lim_{k\to\infty} A^k = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}. \tag{39}$$

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

Solution: Each singular value σ_i of A can be calculated as $\sigma_i = \sqrt{\lambda_i(AA^\top)} = \sqrt{\lambda_i(A^\top A)}$. (This is because A's singular value decomposition, canonically written $A = U\Sigma V^\top$, can be multiplied by a transposed version to give $AA^\top = U\Sigma^2 U^\top$, where Σ^2 is a diagonal matrix containing the squared singular values of A and $UU^\top = I$. For a thorough treatment of SVD, see Calafiore & El Ghaoui chapter 5.)

To find A's singular values, we thus perform the same calculation used in part (a) to find each $\lambda_i(AA^\top) = \sigma_i^2$:

$$AA^\top = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{bmatrix} \tag{40}$$

$$
\begin{aligned}
&= \frac{1}{2}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}. && (41) \\
\det\left((sI - AA^\top)\right) &= \det\left(\begin{bmatrix} s-1 & -\frac{1}{2} \\ -\frac{1}{2} & s-\frac{1}{2} \end{bmatrix}\right) && (42) \\
&= (s-1)\left(s-\frac{1}{2}\right) - \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) && (43) \\
&= s^2 - s - \frac{1}{2}s + \frac{1}{2} - \frac{1}{4} && (44) \\
&= s^2 - \frac{3}{2}s + \frac{1}{4} && (45) \\
&= \left(s-\frac{3}{4}\right)^2 - \frac{9}{16} + \frac{1}{4} && (46) \\
&= \left(s-\frac{3}{4}\right)^2 - \frac{5}{16} && (47) \\
&= \left(s-\frac{3}{4}-\frac{\sqrt{5}}{4}\right)\left(s-\frac{3}{4}+\frac{\sqrt{5}}{4}\right) \qquad a^2-b^2=(a-b)(a+b) && (48) \\
&= \left(s-\frac{3+\sqrt{5}}{4}\right)\left(s-\frac{3-\sqrt{5}}{4}\right) && (49) \\
&= (s-\sigma_1^2)(s-\sigma_2^2). && (50)
\end{aligned}
$$

Thus, the singular values of A are $\sigma_1 = \frac{\sqrt{3+\sqrt{5}}}{2}$ and $\sigma_2 = \frac{\sqrt{3-\sqrt{5}}}{2}$.

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.

Solution: We have

$$\begin{aligned}
\lambda_1 \vec{u}_1^\top \vec{u}_2 &= (\lambda_1 \vec{u}_1)^\top \vec{u}_2 && (51)\\
&= (A\vec{u}_1)^\top \vec{u}_2 && (52)\\
&= \vec{u}_1^\top A^\top \vec{u}_2 && (53)\\
&= \vec{u}_1^\top A \vec{u}_2 && (54)\\
&= \vec{u}_1^\top (\lambda_2 \vec{u}_2) && (55)\\
&= \lambda_2 \vec{u}_1^\top \vec{u}_2. && (56)
\end{aligned}$$

Thus we have

$$\lambda_1(\vec{u}_1^\top \vec{u}_2) = \lambda_2(\vec{u}_1^\top \vec{u}_2) \implies (\lambda_1 - \lambda_2)(\vec{u}_1^\top \vec{u}_2) = 0. \quad (57)$$

Since $\lambda_1 \neq \lambda_2$, we have $\lambda_1 - \lambda_2 \neq 0$, so we must have $\vec{u}_1^\top \vec{u}_2 = 0$.

Thus, $\vec{u}_1^\top \vec{u}_2 = 0$ for any $\vec{u}_1, \vec{u}_2$ corresponding to different eigenvalues. Stated differently, unique eigenvalues correspond to orthogonal eigenvectors.

This, in combination with the fact that the geometric multiplicity and algebraic multiplicity of a symmetric matrix are equal, allows us to construct an orthonormal set of eigenvectors. First, find all the distinct eigenvalues and their respective eigenvectors. Then, for all eigenvalues with algebraic multiplicity > 1, we know that the respective eigenspace is spanned by k linearly independent eigenvectors. Utilizing Gram-Schmidt, we can construct an orthonormal set of eigenvectors from this basis for this eigenspace. Putting the eigenvectors from these two cases together, we have constructed the U matrix of the decomposition.

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

$$A = QR, \tag{58}$$

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

$$A = \begin{bmatrix} 3 & -3 & 1 \\ 4 & -4 & -7 \\ 0 & 3 & 3 \end{bmatrix} \tag{59}$$

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

Solution: Let $\vec{a}_i$ and $\vec{q}_i$ denote the columns of A and Q, respectively. Using Gram-Schmidt, we obtain an orthonormal basis $\vec{q}_i$ for the column space of A.

$$\vec{p}_1 = \vec{a}_1, \vec{q}_1 = \frac{\vec{p}_1}{\|\vec{p}_1\|_2} \tag{60}$$

$$\vec{p}_2 = \vec{a}_2 - (\vec{a}_2^\top \vec{q}_1)\vec{q}_1, \quad \vec{q}_2 = \frac{\vec{p}_2}{\|\vec{p}_2\|_2} \tag{61}$$

$$\vec{p}_3 = \vec{a}_3 - (\vec{a}_3^\top \vec{q}_1)\vec{q}_1 - (\vec{a}_3^\top \vec{q}_2)\vec{q}_2, \vec{q}_3 = \frac{\vec{p}_3}{\|\vec{p}_3\|_2} \tag{62}$$

$$\vdots \tag{63}$$

Rearranging terms, we have

$$\vec{a}_1 = r_{11}\vec{q}_1 \tag{64a}$$

$$\vec{a}_i = r_{i1}\vec{q}_1 + \cdots + r_{ii}\vec{q}_i, \quad i = 2, ..., n, \tag{64b}$$

where each $\vec{q}_i$ has unit norm, and $r_{ij}\vec{q}_j$ denotes the projection of $\vec{a}_i$ onto the vector $\vec{q}_j$ for $j \neq i$. Stacking $\vec{a}_i$ horizontally into A and rewriting (64a-b) in matrix notation, we obtain $A = QR$. For the given matrix, we have

$$A = \begin{bmatrix} 0.6 & 0 & 0.8 \\ 0.8 & 0 & -0.6 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 5 & -5 & -5 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix}. \tag{65}$$

Note that an equivalent factorization is $A = (-Q)(-R)$.

(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

$$A\vec{x} = \vec{b} \tag{66}$$

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

$$A\vec{x} = \begin{bmatrix} 8 \\ -6 \\ 3 \end{bmatrix}. \tag{67}$$

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

$$R\vec{x} = \vec{v}. \tag{68}$$

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

Solution: We note that $Q^{-1} = Q^\top$.

$$A\vec{x} = \vec{b} \tag{69}$$

$$QR\vec{x} = \vec{b} \tag{70}$$

$$Q^\top QR\vec{x} = R\vec{x} = Q^\top\vec{b}. \tag{71}$$

Thus

$$\vec{v} = Q^\top\vec{b} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix}. \tag{72}$$

Given R and $\vec{v}$, we can find $\vec{x}$ by back-substitution:

$$\begin{bmatrix} 5 & -5 & -5 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix} \implies x_3 = 2 \implies x_2 = -1 \implies x_1 = 1 \implies \vec{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}. \tag{73}$$

(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.

Solution: We count the number of operations in back substitution. Solving the initial equation

$$r_{nn}x_n = \bar{b}_n \tag{74}$$

takes 1 multiplication. Solving each subsequent equation takes one more multiplication and one more addition than the previous. In total, we have $1 + 3 + 5 + \cdots$ of operations, which is on the order of $\mathcal{O}(n^2)$. Thus, matrix multiplication and back substitution are both $O(n^2)$. Given the QR decomposition of A, we can solve $A\vec{x} = \vec{b}$ in $O(n^2)$ times.

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.

Solution: We have

$$A^\top = (\vec{p}\vec{q}^\top + \vec{q}\vec{p}^\top)^\top = \vec{q}\vec{p}^\top + \vec{p}\vec{q}^\top = A. \tag{75}$$

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

Solution: We have

$$A\vec{p} = c\vec{p} + \vec{q}, \quad A\vec{q} = \vec{p} + c\vec{q}, \tag{76}$$

from which we obtain

$$A(\vec{p} - \vec{q}) = (c-1)(\vec{p} - \vec{q}), \quad A(\vec{p} + \vec{q}) = (c+1)(\vec{p} + \vec{q}). \tag{77}$$

Thus $\vec{u}_\pm := \vec{p} \pm \vec{q}$ is an (un-normalized) eigenvector of A, with eigenvalue $c \pm 1$.

(c) Determine the nullspace and rank of A.

Solution: If $\vec{x} \in \mathbb{R}^n$ is in the nullspace of A we must have: $A\vec{x} = 0$.

$$0 = A\vec{x} = \vec{p}(\vec{q}^\top \vec{x}) + \vec{q}(\vec{p}^\top \vec{x}). \tag{78}$$

Since $(\vec{q}^\top \vec{x})$ and $(\vec{p}^\top \vec{x})$ are scalars we can rewrite this as:

$$0 = A\vec{x} = (\vec{q}^\top \vec{x})\vec{p} + (\vec{p}^\top \vec{x})\vec{q} = 0. \tag{79}$$

However, since $\vec{p}, \vec{q}$ are linearly independent, the fact that a linear combination of $\vec{p}, \vec{q}$ is zero implies that $\vec{p}^\top \vec{x} = \vec{q}^\top \vec{x} = 0$. Hence, the nullspace of A is the set of vectors orthogonal to $\vec{p}$ and $\vec{q}$, i.e., $\mathcal{N}(A) = \operatorname{span}(\vec{p}, \vec{q})^\perp$. We have from the fundamental theorem of linear algebra and the fact that A is symmetric,

$$\mathcal{R}(A) = \mathcal{R}(A^\top) = \mathcal{N}(A)^\perp = (\operatorname{span}(\vec{p}, \vec{q})^\perp)^\perp = \operatorname{span}(\vec{p}, \vec{q}). \tag{80}$$

And since p and q are linearly independent, $\operatorname{rank}(A) = 2$.

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

Solution: Since the rank is 2, we need to find a total of two non-zero eigenvalues. First, we check that $\lambda = c \pm 1$ is not 0. We have $\vec{p} - \vec{q} \neq 0$ which implies $\|\vec{p} - \vec{q}\|_2^2 > 0$ which means $\|\vec{p}\|_2^2 + \|\vec{q}\|_2^2 - 2\vec{p}^\top \vec{q} > 0$. Therefore, we have $c < 1$ and through a similar proof with $\vec{p} + \vec{q}$, we have $-c < 1$. From these two facts, we get $|c| < 1$. Thus, we have found two linearly independent eigenvectors $\vec{u}_\pm = \vec{p} \pm \vec{q}$ that do not belong to the nullspace. Then, the eigenvalue decomposition is

$$A = (c-1)\vec{v}_-\vec{v}_-^\top + (c+1)\vec{v}_+\vec{v}_+^\top, \tag{81}$$

where $\vec{v}_\pm$ are the normalized vectors $\vec{v}_\pm = \vec{u}_\pm / \|\vec{u}_\pm\|_2$.

It is fine to have this kind of answer, but we can actually simplify it further. Since

$$\|\vec{p} \pm \vec{q}\|_2^2 = \vec{p}^\top \vec{p} \pm 2\vec{p}^\top \vec{q} + \vec{q}^\top \vec{q} = 2(1 \pm c), \tag{82}$$

we have

$$\vec{v}_{\pm} = \frac{1}{\sqrt{2(1 \pm c)}}(\vec{p} \pm \vec{q}), \tag{83}$$

so that the eigenvalue decomposition becomes

$$A = \frac{1}{2}\left((\vec{p} + \vec{q})(\vec{p} + \vec{q})^\top - (\vec{p} - \vec{q})(\vec{p} - \vec{q})^\top\right). \tag{84}$$

Note that this form removes explicit dependence on $c = \vec{p}^\top \vec{q}$.

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.

Solution: $\implies$:

i. Solution 1: We can plug in the Spectral Decomposition here:

$$\vec{x}^\top A\vec{x} = \vec{x}^\top U\Sigma U^\top \vec{x} = \vec{v}^\top \Sigma \vec{v} \geq 0, \tag{85}$$

where $\vec{v} := U^\top \vec{x}$ is a rotated version of $\vec{x}$ since U is orthonormal. Now, we just need to convert that final quadratic into any eigenvalue of A, and we can do that by choosing a $\vec{v}$ that pulls out whichever eigenvalue we want (e.g. if we want the first eigenvalue, we can choose the first unit vector). To be thorough, we can then realize that the set of $\vec{x}$'s such that $U^\top \vec{x} = \vec{e}_i$ for any unit vector, will pull out the ith eigenvalue, thus satisfying definition 2.

ii. Solution 2: We can just use the definition of an eigenvalue:

$$\vec{x}^\top A\vec{x} = \vec{x}\lambda\vec{x} = \lambda\vec{x}^\top \vec{x} = \lambda \|\vec{x}\|_2^2 \tag{86}$$

Since norms/anything squared is always non-negative, in order for $\lambda \|\vec{x}\|_2^2 \geq 0$, λ must be non-negative.

$\impliedby$: Using the Spectral Decomposition again, we arrive at the equation $\vec{v}^\top \Sigma\vec{v}$, which we can expand further:

$$\vec{v}^\top \Sigma \vec{v} = \sum_i \lambda_i v_i^2 \geq 0, \tag{87}$$

where the last inequality came from the fact that anything squared is non-negative and all eigenvalues are non-negative by assumption of the problem.

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

Solution: The quadratic form $\vec{x}^\top A\vec{x} \geq 0$ applies for all vectors $\vec{x}$. Therefore, let's choose a vector that will pull out A_{ii}: the ith unit vector. $A\vec{e}_i$ pulls out the ith column $\vec{a}_i$, followed by $\vec{e}_i^\top \vec{a}_i$, which will pull out the ith element of the ith column. Therefore, $\vec{e}_i^\top A\vec{e}_i = A_{ii} \geq 0$.

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 $A = P^2$.

Solution: Since A is symmetric positive semidefinite, A has non-negative eigenvalues. Thus, we can diagonalize A as $A = U\Sigma U^\top$, where the diagonal matrix of eigenvalues Σ has all non-negative entries on the diagonal. Then, we are able to define a matrix $A^{\frac{1}{2}} = U\Sigma^{\frac{1}{2}}U^\top$, where $\Sigma^{\frac{1}{2}}$ is a diagonal matrix with the square roots of A's eigenvalues. Note that $A^{\frac{1}{2}}$ is PSD since its eigenvalues are still non-negative. Thus, with $P = A^{\frac{1}{2}}$, we can show the following:

$$P^\top P = (A^{\frac{1}{2}})^\top A^{\frac{1}{2}} = (U\Sigma^{\frac{1}{2}}U^\top)^\top U\Sigma^{\frac{1}{2}}U^\top = U\Sigma^{\frac{1}{2}}U^\top U\Sigma^{\frac{1}{2}}U^\top = U\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}U^\top = U\Sigma U^\top = A. \tag{88}$$

(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."

Solution: We can plug in $A = Q^\top Q$ into the quadratic form as follows:

$$\vec{x}^\top A\vec{x} = \vec{x}^\top Q^\top Q\vec{x} = (Q\vec{x})^\top (Q\vec{x}) = \|Q\vec{x}\|_2^2 \geq 0. \tag{89}$$

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 $B^\top B$.

Solution: Say $\lambda \neq 0, \vec{v}$ is an eigenpair of $B^\top B$ which is a $n \times n$ matrix. Hence,

$$(B^\top B)\vec{v} = \lambda\vec{v} \tag{90}$$

Multiply both sides with B, to get,

$$B(B^\top B)\vec{v} = B(\lambda\vec{v}) \tag{91}$$

$$(BB^\top)B\vec{v} = \lambda B\vec{v} \tag{92}$$

As $\lambda \neq 0$ and $\vec{v} \neq \vec{0}$, we have $\lambda\vec{v} \neq \vec{0}$ and so, $(B^\top B)\vec{v} \neq \vec{0}$. Thus $B^\top(B\vec{v}) \neq \vec{0}$, which implies that $B\vec{v} \neq \vec{0}$. Therefore, $B\vec{v}$ is an eigenvector of $BB^\top$ corresponding to λ. Similarly, we can show that every non-zero eigenvalue of $BB^\top$ is an eigenvalue of $B^\top B$ and we are done.

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.