Homework 1

Self grades are due at 11 PM on September 13, 2023.

1. 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:

$$\begin{aligned}
\det(sI - A) &= \det\left(\begin{bmatrix} s & -1 \\ -\frac{1}{2} & s - \frac{1}{2} \end{bmatrix}\right) && (1) \\
&= s\left(s - \frac{1}{2}\right) - (-1)\left(-\frac{1}{2}\right) && (2) \\
&= s^2 - \frac{1}{2}s - \frac{1}{2} && (3) \\
&= \left(s - \frac{1}{4}\right)^2 - \frac{1}{16} - \frac{1}{2} && (4) \\
&= \left(s - \frac{1}{4}\right)^2 - \frac{9}{16} && (5) \\
&= \left(s - \frac{1}{4} - \frac{3}{4}\right)\left(s - \frac{1}{4} + \frac{3}{4}\right) \qquad a^2 - b^2 = (a-b)(a+b) && (6) \\
&= (s-1)\left(s + \frac{1}{2}\right). && (7)
\end{aligned}$$

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:

$$A\vec{x} = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} x_a \\ x_b \end{bmatrix} = \begin{bmatrix} x_b \\ \frac{x_a + x_b}{2} \end{bmatrix} \tag{8}$$

$$A\vec{x}_1 = \vec{x}_1 \iff x_b = x_a \iff \vec{x}_1 = \alpha_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \alpha_1 \neq 0 \in \mathbb{R}. \tag{9}$$

$$A\vec{x}_2 = -\frac{1}{2}\vec{x}_2 \iff x_b = -\frac{1}{2}x_a \iff \vec{x}_2 = \alpha_2 \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix}, \ \alpha_2 \neq 0 \in \mathbb{R}. \tag{10}$$

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

$$A\begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1\vec{x}_1 & \lambda_2\vec{x}_2 \end{bmatrix} = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}. \tag{11}$$

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:

$$A \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}, \tag{12}$$

and rearranging,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}^{-1}. \tag{13}$$

Calculating the latter inverse explicitly, we have

$$\begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}^{-1} = -\frac{2}{3} \begin{bmatrix} -\frac{1}{2} & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix} \text{ because } \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{14}$$

so finally,

$$A = P\Lambda P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix}. \tag{15}$$

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:

$$A = P\Lambda P^{-1} \tag{16}$$

$$A^k = (P\Lambda P^{-1})^k \tag{17}$$

$$= (P\Lambda P^{-1})(P\Lambda P^{-1}) \dots (P\Lambda P^{-1}) \qquad (k \text{ times}) \tag{18}$$

$$= P\Lambda \underbrace{P^{-1}P}_{I} \Lambda P^{-1} \dots P\Lambda P^{-1} \tag{19}$$

$$= P\Lambda^k P^{-1} \tag{20}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}^k \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix} \tag{21}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & \left(-\frac{1}{2}\right)^k \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix}. \tag{22}$$

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

$$\lim_{k\to\infty} A^k = \begin{bmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}. \tag{23}$$

(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 $\lambda_i(AA^\top) = \sigma_i^2$:

$$\begin{aligned}
AA^\top &= \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{bmatrix} && (24) \\
&= \frac{1}{2} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}. && (25) \\
\det\left((sI - AA^\top)\right) &= \det\left(\begin{bmatrix} s-1 & -\frac{1}{2} \\ -\frac{1}{2} & s-\frac{1}{2} \end{bmatrix} \right) && (26) \\
&= (s-1)\left(s - \frac{1}{2}\right) - \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) && (27) \\
&= s^2 - s - \frac{1}{2}s + \frac{1}{2} - \frac{1}{4} && (28) \\
&= s^2 - \frac{3}{2}s + \frac{1}{4} && (29) \\
&= \left(s - \frac{3}{4}\right)^2 - \frac{9}{16} + \frac{1}{4} && (30) \\
&= \left(s - \frac{3}{4}\right)^2 - \frac{5}{16} && (31) \\
&= \left(s - \frac{3}{4} - \frac{\sqrt{5}}{4}\right)\left(s - \frac{3}{4} + \frac{\sqrt{5}}{4}\right) \qquad a^2 - b^2 = (a-b)(a+b) && (32) \\
&= \left(s - \frac{3+\sqrt{5}}{4}\right)\left(s - \frac{3-\sqrt{5}}{4}\right) && (33) \\
&= (s - \sigma_1^2)(s - \sigma_2^2). && (34)
\end{aligned}$$

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

2. Least Squares

(a) Let $A \in \mathbb{R}^{m\times n}$ and $\vec{y} \in \mathbb{R}^m$ be given. Suppose A has full column rank. Show that the minimizer $\vec{x}^\star$ of the least-squares problem:

$$\min_{\vec{x}\in\mathbb{R}^n} \|A\vec{x} - \vec{y}\|_2^2$$

is given by $\vec{x}^\star = (A^\top A)^{-1}A^\top\vec{y}$.

Solution: There are many possible ways to solve this problem; we present a solution below based on geometric intuition, which is detailed in Section 1.2 of the course reader. Intuitively speaking, we are searching over the set of all vectors of the form $\{A\vec{x} : \vec{x} \in \mathbb{R}^n\}$ in search for the one closest to $\vec{y}$, i.e., the vector in $R(A)$ that is closest to $\vec{y}$. To this end, we claim the following: Let $\vec{z} = A\vec{x}^\star \in R(A)$ be the unique nonzero vector in $R(A)$ such that $\vec{y} - \vec{z}$ is perpendicular to $R(A)$, i.e., $A^\top(\vec{y} - \vec{z}) = \vec{0}$. Then $\vec{z}$ is the closest vector to $\vec{y}$ in $R(A)$, and we will aim to find a vector $\vec{x}^\star$ such that $A\vec{x}^\star = \vec{z}$.

Let $\vec{u} \in R(A)$. Then:

$$\begin{aligned}\|\vec{y} - \vec{u}\|_2^2 &= \|(\vec{y} - \vec{z}) + (\vec{z} - \vec{u})\|_2^2 \\ &= \|\vec{y} - \vec{z}\|_2^2 + \|\vec{z} - \vec{u}\|_2^2 \\ &\geq \|\vec{y} - \vec{z}\|_2^2,\end{aligned}$$

with equality if and only if $\vec{z} = \vec{u}$. This proves the claim.

Now, let us study the equation:

$$A^\top(\vec{y} - A\vec{x}^\star) = A^\top(\vec{y} - \vec{z}) = \vec{0}.$$

Distributing $A^\top$ and rearranging terms, we obtain:

$$A^\top A\vec{x}^\star = A^\top\vec{y}^\star.$$

To obtain the final expression, it remains to recall that $A^\top A$ is invertible, given that A has full column rank (see Discussion 1, Problem 3). We can then apply $A^\top A$ to both sides of the above equation to obtain the desired expression, $\vec{x}^\star = (A^\top A)^{-1}A^\top\vec{y}$.

(b) The Michaelis-Menten model for enzyme kinetics relates the rate y of an enzymatic reaction to the concentration x of a substrate, as follows:

$$y = \frac{\beta_1 x}{\beta_2 + x}, \tag{35}$$

for constants $\beta_1, \beta_2 > 0$. This model will be used throughout the remaining sub-parts of this problem.

Show that the model can be expressed as a linear relation between the values $1/y = y^{-1}$ and $1/x = x^{-1}$. Specifically, give an equation of the form $y^{-1} = w_1 + w_2x^{-1}$, specifying the values of w_1 and w_2 in terms of β_1 and β_2.

Solution: Inverting each side of the equation, we have

$$\begin{aligned}y^{-1} &= \left(\frac{\beta_1 x}{\beta_2 + x}\right)^{-1} &(36)\\ &= \frac{\beta_2 + x}{\beta_1 x} &(37)\end{aligned}$$

$$= \frac{\beta_2}{\beta_1 x} + \frac{x}{\beta_1 x} \tag{38}$$

$$= \frac{\beta_2}{\beta_1} x^{-1} + \frac{1}{\beta_1} \tag{39}$$

$$= \frac{1}{\beta_1} + \frac{\beta_2}{\beta_1} x^{-1}. \tag{40}$$

$$\tag{41}$$

The above equation has exactly the desired form $y^{-1} = w_1 + w_2 x^{-1}$ for $w_1 = \frac{1}{\beta_1}$ and $w_2 = \frac{\beta_2}{\beta_1}$.

(c) In general, reaction parameters β_1 and β_2 (and, thus, w_1 and w_2) are not known a priori and must be fitted from data — for example, using least squares. Suppose you collect m measurements (x_i, y_i), $i = 1, \ldots, m$ over the course of a reaction. Formulate the least squares problem

$$\vec{w}^\star = \operatorname*{argmin}_{\vec{w}} \|X\vec{w} - \vec{y}\|_2^2, \tag{42}$$

where $\vec{w}^\star = \begin{bmatrix} w_1^\star & w_2^\star \end{bmatrix}^\top$, and you must specify $X \in \mathbb{R}^{m \times 2}$ and $\vec{y} \in \mathbb{R}^m$. Specifically, your solution should include explicit expressions for X and $\vec{y}$ as a function of (x_i, y_i) values and a final expression for $\vec{w}^\star$ in terms of X and $\vec{y}$, which should contain only matrix multiplications, transposes, and inverses.

Assume without loss of generality that $x_1 \neq x_2$.

Solution: To formulate the least squares problem as stated, X and $\vec{y}$ values should be set to

$$X = \begin{bmatrix} 1 & \ldots & 1 \\ x_1^{-1} & \ldots & x_m^{-1} \end{bmatrix}^\top, \quad \vec{y} = \begin{bmatrix} y_1^{-1} & \ldots & y_m^{-1} \end{bmatrix}^\top. \tag{43}$$

To solve this least squares problem, we note that the optimal residual vector $X\vec{w}^\star - \vec{y}$ must be orthogonal to $\mathcal{R}(X)$ by the orthogonality principle, and we have

$$X^\top (X\vec{w}^\star - \vec{y}) = 0. \tag{44}$$

Rearranging we get,

$$\vec{w}^\star = (X^\top X)^{-1} X^\top \vec{y}. \tag{45}$$

(d) Assume that we have used the above procedure to calculate values for $w_1^\star$ and $w_2^\star$, and we now want to estimate $\widehat{\vec{\beta}} = \begin{bmatrix} \widehat{\beta}_1 & \widehat{\beta}_2 \end{bmatrix}^\top$. Write an expression for $\widehat{\vec{\beta}}$ in terms of $w_1^\star$ and $w_2^\star$.

Solution: To calculate $\widehat{\vec{\beta}}$, we can simply reverse the calculations from part (b):

$$w_1 = \frac{1}{\beta_1} \implies \beta_1 = \frac{1}{w_1}, \tag{46}$$

$$w_2 = \frac{\beta_2}{\beta_1} \implies \beta_2 = \beta_1 w_2 = \frac{w_2}{w_1}. \tag{47}$$

Thus, $\widehat{\vec{\beta}} = \begin{bmatrix} \frac{1}{w_1^\star} & \frac{w_2^\star}{w_1^\star} \end{bmatrix}^\top$.

NOTE: This problem was taken (with some edits) from the textbook *Optimization Models* by Calafiore and El Ghaoui.

3. Vector Spaces and Rank

The *rank* of a $m \times n$ matrix A, $\operatorname{rank}(A)$, is the dimension of its *range*, also called span, and denoted $\mathcal{R}(A) := \{A\vec{x} \,:\, \vec{x} \in \mathbb{R}^n\}$.

(a) Assume that $A \in \mathbb{R}^{m \times n}$ takes the form $A = \vec{u}\vec{v}^\top$, with $\vec{u} \in \mathbb{R}^m$, $\vec{v} \in \mathbb{R}^n$, and $\vec{u}, \vec{v} \neq \vec{0}$. (Note that a matrix of this form is known as a *dyad*.) Find the rank of A.

HINT: Consider the quantity $A\vec{x}$ for arbitrary $\vec{x}$, i.e., what happens when you multiply any vector by matrix A.

Solution: For any $\vec{x} \in \mathbb{R}^n$, we have that $A\vec{x} = \vec{u}\vec{v}^\top\vec{x} = \vec{u}(\vec{v}^\top\vec{x}) = (\vec{v}^\top\vec{x})\vec{u}$. Note that $\vec{v}^\top\vec{x}$ is a scalar that can take on any value depending on choice of $\vec{x}$. Since the range of A is the subspace reachable through any choice of $\vec{x}$, $\mathcal{R}(A)$ is simply the 1-dimensional subspace spanned by $\vec{u}$ (i.e., the line pointing along $\vec{u}$). Since a single vector (namely, $\vec{u}$) spans $\mathcal{R}(A)$, the rank of A is 1.

(b) Show that for arbitrary $A, B \in \mathbb{R}^{m \times n}$,

$$\operatorname{rank}(A+B) \leq \operatorname{rank}(A) + \operatorname{rank}(B), \tag{48}$$

i.e., the rank of the sum of two matrices is less than or equal to the sum of their ranks.

HINT: First, show that $\mathcal{R}(A+B) \subseteq \mathcal{R}(A) + \mathcal{R}(B)$, meaning that any vector in the range of $A+B$ can be expressed as the sum of two vectors, each in the range of A and B, respectively. Remember that for any matrix A, $\mathcal{R}(A)$ is a subspace, and for any two subspaces S_1 and S_2, $\dim(S_1 + S_2) \leq \dim(S_1) + \dim(S_2)$.[1] *(Note that the sum of vector spaces $S_1 + S_2$ is not equivalent to $S_1 \cup S_2$, but is defined as $S_1 + S_2 := \{\vec{s_1} + \vec{s_2} | \vec{s_1} \in S_1, \vec{s_2} \in S_2\}$.)*

Solution: Given any vector $\vec{v} \in \mathcal{R}(A+B)$, there must by definition exist $\vec{x} \in \mathbb{R}^n$ such that $\vec{v} = (A+B)\vec{x}$. Thus, $\vec{v} = (A+B)\vec{x} = \underbrace{A\vec{x}}_{\in\mathcal{R}(A)} + \underbrace{B\vec{x}}_{\in\mathcal{R}(B)}$, so $\mathcal{R}(A+B) \subseteq \mathcal{R}(A) + \mathcal{R}(B)$, as hinted.

Computing the dimension of each side of the subset relationship, it follows that

$$\dim(\mathcal{R}(A+B)) \leq \dim(\mathcal{R}(A) + \mathcal{R}(B)). \tag{49}$$

Using the second part of the hint, we have that

$$\dim(\mathcal{R}(A) + \mathcal{R}(B)) \leq \dim(\mathcal{R}(A)) + \dim(\mathcal{R}(B)). \tag{50}$$

Combining the previous two equations,

$$\dim(\mathcal{R}(A+B)) \leq \dim(\mathcal{R}(A)) + \dim(\mathcal{R}(B)), \tag{51}$$

i.e., by definition,

$$\operatorname{rank}(A+B) \leq \operatorname{rank}(A) + \operatorname{rank}(B) \tag{52}$$

as desired.

(c) Consider an $m \times n$ matrix A that takes the form $A = UV^\top$, with $U \in \mathbb{R}^{m \times k}$, $V \in \mathbb{R}^{n \times k}$. Show that the rank of A is less or equal than k. *HINT: Use parts (a) and (b), and remember that this decomposition can*

[1] This fact can be proved by taking a basis of S_1 and extending it to a basis of S_2 (during which we can only add *at most* $\dim(S_2)$ basis vectors). This extended basis must now also be a basis of $S_1 + S_2$. Thus, $\dim(S_1 + S_2) \leq \dim(S_1) + \dim(S_2)$.

also be written as the dyadic expansion

$$A = UV^\top = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_k \end{bmatrix} \begin{bmatrix} \vec{v}_1^\top \\ \vdots \\ \vec{v}_k^\top \end{bmatrix} = \sum_{i=1}^{k} \vec{u}_i \vec{v}_i^\top, \tag{53}$$

for $U = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_k \end{bmatrix}$ *and* $V = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_k \end{bmatrix}$.

Solution: Starting with the dyadic expansion above, iteratively pulling out terms from this summation, and using the result from (a) that the rank of a dyadic matrix is 1 (or 0, if any $\vec{v}_i = \vec{0}$), we know by the rank relation from (b) that

$$\operatorname{rank}(A) = \operatorname{rank}\left(\sum_{i=1}^{k} \vec{u}_i \vec{v}_i^\top\right) \leq \operatorname{rank}\left(\sum_{i=1}^{k-1} \vec{u}_i \vec{v}_i^\top\right) + \underbrace{\operatorname{rank}\left(\vec{u}_k \vec{v}_k^\top\right)}_{\text{0 or 1}} \leq \dots \leq k, \tag{54}$$

as desired.

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

$$A = QR,$$

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.

Solution: Let a_i and q_i denote the columns of A and Q, respectively. Using Gram-Schmidt, we obtain an orthogonal basis q_i for the column space of A.

$$p_1 = a_1, q_1 = \frac{p_1}{\|p_1\|_2} \tag{55}$$

$$p_2 = a_2 - (a_2^\top q_1)q_1, \quad q_2 = \frac{p_2}{\|p_2\|_2} \tag{56}$$

$$p_3 = a_3 - (a_3^\top q_1)q_1 - (a_3^\top q_2)q_2, q_3 = \frac{p_3}{\|p_3\|_2} \tag{57}$$

$$\vdots \tag{58}$$

Rearranging terms, we have

$$a_1 = r_{11}q_1 \tag{59a}$$

$$a_i = r_{i1}q_1 + \cdots + r_{ii}q_i, \quad i = 2, ..., n, \tag{59b}$$

where each q_i has unit norm, and $r_{ij}q_j$ denotes the projection of a_i onto the vector q_j for $j \neq i$.

Stacking a_i horizontally into A and rewriting (59a-b) in matrix notation, we obtain $A = QR$. For the given matrix, we have

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

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 $b \in \mathbb{R}^n$, 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.

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

$$\begin{aligned} Ax &= b \\ QRx &= b \\ Q^\top QRx = Rx &= Q^\top b. \end{aligned}$$

Thus

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

Given R and v, we can find x by back-substitution:

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

(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.

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

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

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 $Ax = b$ in $O(n^2)$ times.

5. 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.