# Homework 2

## This homework is due at 11 PM on February 6th, 2026.

**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. **Proof of the Fundamental Theorem of Linear Algebra**

   In this question, we will prove the fundamental theorem of linear algebra. For any $A \in \mathbb{R}^{m\times n}$, let $\mathcal{N}(A)$, $\mathcal{R}(A)$, and $\operatorname{rank}(A)$ denote the null space, range and rank of $A$ respectively. For any subspace, $\mathcal{S}$ with dimension, $\dim(\mathcal{S})$, let $\mathcal{S}^\perp$ denote its the subspace orthogonal to $\mathcal{S}$. The fundamental theorem of linear algebra states that,

   $$\mathcal{N}(A) \oplus \mathcal{R}\left(A^\top\right) = \mathbb{R}^n. \tag{1}$$

   The proof technique we employ will first show that,

   $$\mathcal{N}(A) = \mathcal{R}\left(A^\top\right)^\perp. \tag{2}$$

   Then we will prove that we can find orthonormal vectors $\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n$ such that $\mathcal{N}(A) = \operatorname{span}(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_\ell)$ and $\mathcal{R}\left(A^\top\right) = \operatorname{span}(\vec{e}_{\ell+1}, \vec{e}_{\ell+2}, \ldots, \vec{e}_n)$. As a corollary we get the rank-nullity theorem:

   $$\dim(\mathcal{N}(A)) + \operatorname{rank}(A) = n. \tag{3}$$

   (a) First, show that $\mathcal{N}(A) \subseteq \mathcal{R}\left(A^\top\right)^\perp$.
   *HINT: Consider $\vec{u}$ in $\mathcal{N}(A)$, $\vec{v} \in \mathcal{R}\left(A^\top\right)$ and show that $\vec{u}^\top \vec{v} = 0$.*

   (b) Now show that: $\mathcal{R}\left(A^\top\right)^\perp \subseteq \mathcal{N}(A)$.
   *HINT: Show that any vector $\vec{v}$ that is orthogonal to all vectors in the range of $A^\top$ satisfies $A\vec{v} = 0$. To do this, consider $\vec{v} \in R(A^\top)^\perp$ and what it implies for $\vec{v}^\top A^\top$.*

(c) Note that we could apply the orthogonal decomposition theorem (theorem 19 in the course reader) at this point to complete the proof. However, instead we'll work through how to re-derive that result directly. Let $\dim(\mathcal{N}(A)) = \ell$ and let $\vec{e}_1, \ldots, \vec{e}_\ell$ be an orthonormal basis for $\mathcal{N}(A)$. Consider an extension of the basis to an orthonormal basis, $\vec{e}_1, \ldots, \vec{e}_n$ for $\mathbb{R}^n$. We will prove that $\vec{e}_{\ell+1}, \ldots, \vec{e}_n$ form a basis for $\mathcal{R}\left(A^\top\right)$ and as a consequence, the dimension of $\mathcal{R}\left(A^\top\right)$ is $n - \ell$.

i. Show that $\mathcal{R}\left(A^\top\right)$ lies in the span of $\vec{e}_{\ell+1}, \ldots, \vec{e}_n$.
*HINT: Express any vector $\vec{u} \in \mathcal{R}\left(A^\top\right)$ as $\vec{u} = \sum_{i=1}^n \alpha_i \vec{e}_i$. What are the values of $\alpha_i$?*

ii. From part (i) we know that $\mathcal{R}\left(A^\top\right) \subseteq \operatorname{span}(\vec{e}_{\ell+1}, \ldots, \vec{e}_n)$, but we want something stronger. Show that in fact $\mathcal{R}\left(A^\top\right) = \operatorname{span}(\vec{e}_{\ell+1}, \ldots, \vec{e}_n)$.
*HINT: First, prove $\dim\left(\mathcal{R}\left(A^\top\right)\right) = \dim(\operatorname{span}(\vec{e}_{\ell+1}, \ldots, \vec{e}_n)) = n - \ell$ by contradiction. Assume $\dim\left(\mathcal{R}\left(A^\top\right)\right) = k < n - \ell$, show that a vector $\vec{u} \in \operatorname{span}(\vec{e}_{\ell+1}, \ldots, \vec{e}_n)$ and $\vec{u} \notin \mathcal{R}\left(A^\top\right)$ cannot exist.*
*Specifically, let $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ be an orthonormal basis for $\mathcal{R}\left(A^\top\right)$, we can find non-zero $\vec{u}_\perp = \vec{u} - \sum_{i=1}^k (\vec{f}_i^\top \vec{u})\vec{f}_i$ that is orthogonal to $\mathcal{R}\left(A^\top\right)$. Does $\vec{u}_\perp$ lie in $\mathcal{N}(A)$? Does $\vec{u}_\perp$ also lie in $\operatorname{span}(\vec{e}_{\ell+1}, \ldots, \vec{e}_n)$? Does this lead to a contradiction? Think of $n - \ell = 3$ and $k = 2$ for visualization.*
*HINT: Second, you can use without proof the fact that for two subspaces, $\mathcal{S}_1$ and $\mathcal{S}_2$, if $\mathcal{S}_1 \subseteq \mathcal{S}_2$ and $\dim(\mathcal{S}_1) = \dim(\mathcal{S}_2)$ then $\mathcal{S}_1 = \mathcal{S}_2$.*

 

(d) Using part (c) argue why $\mathcal{N}(A) \oplus \mathcal{R}\left(A^\top\right) = \mathbb{R}^n$ and why the rank nullity theorem holds.

**2. Eigenvalues of Symmetric Matrices**

Let $A \in \mathbb{S}^n$ (i.e., the set of $n \times n$ real symmetric matrices) with eigenvalues $\lambda_i$. Prove that all of the eigenvalues of $A$ are real (i.e. that $\lambda_i \in \mathbb{R}$ for each $i$).

*HINT: Consider the quantity $(Av)^* v$ for eigenvector $v$ where $*$ denotes the conjugate transpose. Note that this is the Hermitian inner product between $Av$ and $v$.*

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

**3. Distinct Eigenvalues, Orthogonal Eigenspaces**

Let $A \in \mathbb{S}^n$ (i.e. the set of $n \times n$ real 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

$$A = QR, \tag{4}$$

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

$$A = \begin{bmatrix} 3 & -3 & 1 \\ 4 & -4 & -7 \\ 0 & 3 & 3 \end{bmatrix} \tag{5}$$

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

$$A\vec{x} = \vec{b} \tag{6}$$

is given as $\vec{x} = A^{-1}\vec{b}$. For the matrix $A = QR$ from part 4(a), assume that we want to solve

$$A\vec{x} = \begin{bmatrix} 8 \\ -6 \\ 3 \end{bmatrix}. \tag{7}$$

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

$$R\vec{x} = \vec{v}. \tag{8}$$

Then, given the upper-triangular matrix $R$ in part 4(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. Determinants**

Consider a unit box $\mathcal{B}$ in $\mathbb{R}^2$ — i.e., the square with corners $\begin{bmatrix}0\\0\end{bmatrix}$, $\begin{bmatrix}0\\1\end{bmatrix}$, $\begin{bmatrix}1\\0\end{bmatrix}$, and $\begin{bmatrix}1\\1\end{bmatrix}$. Define $A(\mathcal{B})$ as the parallelogram generated by applying matrix $A$ to every point in $\mathcal{B}$.

(a) For $A = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$, calculate the location of each corner of $A(\mathcal{B})$.

(b) Write the area of $A(\mathcal{B})$ as a function of $\det(A)$.

*HINT: How are the basis vectors* $\begin{bmatrix}0\\1\end{bmatrix}$ *and* $\begin{bmatrix}1\\0\end{bmatrix}$ *transformed by the matrix multiplication?*

(c) Calculate the area of $A(\mathcal{B})$ for each of the following values of $A$.

i. $A = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$

ii. $A = \begin{bmatrix}2 & 1\\4 & 3\end{bmatrix}$

iii. $A = \begin{bmatrix}1 & 2\\2 & 4\end{bmatrix}$

iv. $A = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$

**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.