# MAT 167: Advanced Linear Algebra
## Final Exam Solutions

**Problem 1** (15 pts)

**(a)** (5 pts) State the definition of a *unitary* matrix and explain the difference between an *orthogonal* matrix and an unitary matrix.

**Solution:** A unitary matrix is a square matrix of size $m \times m$ whose column vectors form an orthonormal basis for $\mathbb{C}^m$. In other words, a matrix $Q \in \mathbb{C}^{m\times m}$ is unitary if $Q^*Q = QQ^* = I_m$. In other words, $Q^{-1} = Q^*$. An orthogonal matrix is its counterpart for $\mathbb{R}^m$. In other words, an orthogonal matrix is a square matrix of size $m \times m$ whose column vectors form an orthonormal basis for $\mathbb{R}^m$ and satisfies $Q^{-1} = Q^T$.

**(b)** (5 pts) Prove if $Q \in \mathbb{C}^{m\times m}$ is unitary, then for any $\boldsymbol{x} \in \mathbb{C}^m$, we have $\|Q\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$.

**Solution:** We simply use the definitions of the 2-norm for vectors, the inner-product, and a unitary matrix.

$$\|Q\boldsymbol{x}\|_2^2 = \langle Q\boldsymbol{x}, Q\boldsymbol{x}\rangle = (Q\boldsymbol{x})^*(Q\boldsymbol{x}) = \boldsymbol{x}^*Q^*Q\boldsymbol{x} = \boldsymbol{x}^*\boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{x}\rangle = \|\boldsymbol{x}\|_2^2.$$

Since both $\|Q\boldsymbol{x}\|_2$ and $\|\boldsymbol{x}\|_2$ are nonnegative, we can take nonnegative square roots of both sizes of the above equality to get $\|Q\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$.

**(c)** (5 pts) Show that the column vectors of any unitary matrix $Q$ are linearly independent.

**Solution:** Let us consider a linear combination of columns of $Q$:

$$c_1\boldsymbol{q}_1 + \cdots + c_m\boldsymbol{q}_m = \boldsymbol{0},$$

where $\boldsymbol{q}_j$ is the $j$th column vector of $Q$. Using the notation, $\boldsymbol{c} = [c_1, \ldots, c_m]^T$, the above can be rewritten as:

$$Q\boldsymbol{c} = \boldsymbol{0}.$$

Multiplying $Q^*$ from left on the both sides yields:

$$Q^*Q\boldsymbol{c} = \boldsymbol{c} = \boldsymbol{0}.$$

Therefore, $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_m$ are linearly independent.

**Problem 2** (15 pts) Consider the following $2 \times 2$ matrix:

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad -\pi \le \theta < \pi.$$

**(a)** (3 pts) Show that $R_\theta$ is an *orthogonal matrix*.

**Solution:** This can be easily done by computing $R_\theta R_\theta^T = R_\theta^T R_\theta = I_2$.

**(b)** (3 pts) Determine the Fourier expansion of $\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with respect to the basis whose vectors are the column vectors of $R_\theta$.

**Solution:** Let the Fourier expansion of $\boldsymbol{x}$ be

$$\boldsymbol{x} = c_1 \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + c_2 \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = R_\theta \boldsymbol{c},$$

where $c_1$ and $c_2$ are the Fourier coefficients of $\boldsymbol{x}$, and $\boldsymbol{c} = \begin{bmatrix} c_1 & c_2 \end{bmatrix}^T$. Since these column vectors form an orthonormal basis, the Fourier coefficients can be computed by multiplying $R_\theta^T$ on both sides from left:

$$R_\theta^T \boldsymbol{x} = \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix} = R_\theta^T R_\theta \boldsymbol{c} = \boldsymbol{c}.$$

Thus, the Fourier expansion in this case is:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \cos\theta \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} - \sin\theta \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}.$$

**(c)** (3 pts) Show that $\|R_\theta\|_2 = 1$ regardless of the values of $\theta$.

**Solution:** The 2-norm of any matrix $A$ is simply the largest singular value $\sigma_1$ of $A$. Thus all we need to do is to compute the largest eigenvalue of $A^T A$. In this case, $R_\theta^T R_\theta = I_2$, thus

$$\det \begin{bmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 = 0.$$

Thus the eigenvalues are: $\lambda = 1, 1$. Thus the largest singular value $\sigma_1 = \sqrt{1} = 1$. Therefore we conclude: $\|R_\theta\|_2 = 1$ independent of the values of $\theta$.

**(d)** (3 pts) Compute $\|R_\theta\|_1$.

**Solution:** Recall the 1-norm of a matrix $A$ is the maximum of the 1-norm of the column vectors of $A$, i.e., $\|A\|_1 = \max_{1 \le j \le n} \|A(:, j)\|_1$. Thus, in this case, the 1-norm of the 1st column vector and that of the 2nd column vector are the same, i.e.,

$$\|R_\theta\|_1 = |\cos\theta| + |\sin\theta|.$$

**(e)** (3 pts) Show that $\|R_\theta\|_\infty = \|R_\theta\|_1$ for any value of $\theta$.

**Solution:** Recall the $\infty$-norm of a matrix $A$ is the maximum of the 1-norm of the row vectors of $A$, i.e., $\|A\|_\infty = \max_{1\le i\le m} \|A(i,:)\|_1$. Thus, in this case, the 1-norm of the 1st row vector and that of the 2nd row vector are the same, i.e.,

$$\|R_\theta\|_\infty = |\cos\theta| + |\sin\theta| = \|R_\theta\|_1.$$

**(f) Bonus problem** (5 pts) Show that $1 \le \|R_\theta\|_1 \le \sqrt{2}$. For what values of $\theta$ do the equalities hold?

[Hint: Check the periodicity of the 1-norm in this case, and then use the trig. identity: $\cos\theta + \sin\theta = \sqrt{2}\cos(\theta - \pi/4)$.]

**Solution:** Let

$$f(\theta) := \|R_\theta\|_1 = |\cos\theta| + |\sin\theta|.$$

First, notice that $f(\theta + \pi/2) = f(\theta)$ using the basic trig. identities, such as $\cos(\theta + \pi/2) = -\sin(\theta)$, $\sin(\theta + \pi/2) = \cos(\theta)$. This means that $f(\theta)$ is periodic with period $\pi/2$. Therefore, we only need to check this function over the interval $0 \le \theta \le \pi/2$. In this interval, clearly $\cos\theta \ge 0$ and $\sin\theta \ge 0$. Thus,

$$f(\theta) = \cos\theta + \sin\theta = \sqrt{2}\cos(\theta - \pi/4).$$

Thus it is clear that $f(\theta) \le \sqrt{2}$. The equality holds if $\theta = \pi/4$ in this interval. Now, it is also clear that $f(\theta) \ge 1$ in this interval, and the equality holds if $\theta = 0, \pi/2$. Using the periodicity of $f(\theta)$, we conclude that $1 \le f(\theta) \le \sqrt{2}$ and the first equality holds if $\theta = -\pi, -\pi/2, 0,$ and the second equality holds if $\theta = -3\pi/4, -\pi/4, \pi/4, 3\pi/4$.

**Problem 3** (10 pts) Compute the *QR factorization* of the following matrix:

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

**Solution:** Use the Gram-Schmidt procedure here. Let $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ be the column vectors of $A$.

**Step 1:**

$$\begin{cases} \boldsymbol{v}_1 = \boldsymbol{a}_1 = [1, 0, 1]^T \\ r_{11} = \|\boldsymbol{a}_1\| = \sqrt{2} \\ \boldsymbol{q}_1 = [1/\sqrt{2}, 0, 1/\sqrt{2}]^T. \end{cases}$$

**Step 2:**

$$\begin{cases} r_{12} = \langle \boldsymbol{a}_2, \boldsymbol{q}_1 \rangle = 1/\sqrt{2} \\ \boldsymbol{v}_2 = \boldsymbol{a}_2 - r_{12}\boldsymbol{q}_1 = [1/2, 0, 1/2]^T \\ r_{22} = \|\boldsymbol{v}_2\| = 1/\sqrt{2} \\ \boldsymbol{q}_2 = \boldsymbol{v}_2/r_{22} = [1/\sqrt{2}, 0, -1/\sqrt{2}]^T. \end{cases}$$

**Step 3:**

$$\begin{cases} r_{13} = \langle \boldsymbol{a}_3, \boldsymbol{q}_1 \rangle = 1/\sqrt{2} \\ r_{23} = \langle \boldsymbol{a}_3, \boldsymbol{q}_2 \rangle = 1/\sqrt{2} \\ \boldsymbol{v}_3 = \boldsymbol{a}_3 - r_{13}\boldsymbol{q}_1 - r_{23}\boldsymbol{q}_2 = [0, 1, 0]^T \\ r_{33} = \|\boldsymbol{v}_3\| = 1 \\ \boldsymbol{q}_3 = \boldsymbol{v}_3/r_{33} = [0, 1, 0]^T. \end{cases}$$

**Finally:** We get:

$$\underline{Q = [\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}}, \quad \underline{R = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1 \end{bmatrix}}.$$

**Problem 4** (10 pts) For every inner-product space $\mathcal{V}$ over the scalar field $\mathbb{C}$, show that if $\mathcal{M} \subseteq \mathcal{V}$, then $\mathcal{M}^{\perp}$ is a subspace of $\mathcal{V}$ even if $\mathcal{M}$ itself is not a subspace of $\mathcal{V}$.

**Solution:** Take any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{M}^{\perp}$ and any $\alpha \in \mathbb{C}$. Take any member $\boldsymbol{m} \in \mathcal{M}$. Then, by definition of $\perp$ space we must have $\langle \boldsymbol{x}, \boldsymbol{m} \rangle = \langle \boldsymbol{y}, \boldsymbol{m} \rangle = 0$. Now, the bilinearity of the inner-product, we have:

$$\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{m} \rangle = \langle \boldsymbol{x}, \boldsymbol{m} \rangle + \langle \boldsymbol{y}, \boldsymbol{m} \rangle = 0 + 0 = 0.$$

$$\langle \alpha \boldsymbol{x}, \boldsymbol{m} \rangle = \overline{\alpha} \langle \boldsymbol{x}, \boldsymbol{m} \rangle = \overline{\alpha} \cdot 0 = 0.$$

Thus $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{M}^{\perp}$ and $\alpha \boldsymbol{x} \in \mathcal{M}^{\perp}$. This implies that $\mathcal{M}^{\perp}$ is a subspace of $\mathcal{V}$, and this was derived independently from whether $\mathcal{M}$ is a subspace of $\mathcal{V}$ or not. This completes the proof.

**Problem 5** (10 pts) Consider a linear system $A\boldsymbol{x} = \boldsymbol{b}$, where $A \in \mathbb{C}^{m \times n}$, $\boldsymbol{x} \in \mathbb{C}^n$, and $\boldsymbol{b} \in \mathbb{C}^m$. Then show that this system of equations is *consistent* if and only if $\langle \boldsymbol{b}, \boldsymbol{y} \rangle = 0$ for any $\boldsymbol{y} \in \mathbb{C}^m$ that satisfies $A^*\boldsymbol{y} = \boldsymbol{0}$.

[Hint: Consider the orthogonal decomposition: $\mathbb{C}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$.]

**Solution:** ($\Longrightarrow$): Suppose $A\boldsymbol{x} = \boldsymbol{b}$ is a consistent system. This means that $\boldsymbol{b} \in \mathcal{R}(A)$. Now, any vector $\boldsymbol{y} \in \mathbb{C}^m$ with $A^*\boldsymbol{y} = \boldsymbol{0}$ means that $\boldsymbol{y} \in \mathcal{N}(A^*)$. But we know that $\mathcal{R}(A) \perp \mathcal{N}(A^*)$ in $\mathbb{C}^m$. Thus, clearly, $\boldsymbol{b} \perp \boldsymbol{y}$, i.e., $\langle \boldsymbol{b}, \boldsymbol{y} \rangle = 0$.

($\Longleftarrow$): Suppose $\boldsymbol{b} \perp \boldsymbol{y}$ for any $\boldsymbol{y} \in \mathcal{N}(A^*)$. Then, either $\boldsymbol{b} = \boldsymbol{0}$ or $\boldsymbol{b} \in \mathcal{R}(A)$. If $\boldsymbol{b} = \boldsymbol{0}$, then $A\boldsymbol{x} = \boldsymbol{0}$ always has a solution $\boldsymbol{x} = \boldsymbol{0}$. If $\boldsymbol{b} \in \mathcal{R}(A)$, then this means that $A\boldsymbol{x} = \boldsymbol{b}$ is consistent. Thus, either way, $A\boldsymbol{x} = \boldsymbol{b}$ is consistent.

**Problem 6** (10 pts) Consider the following matrix:

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

**(a)** (5 pts) Compute the *Singular Value Decomposition* (SVD) of $A$.

**Solution:** Follows the usual strategy.

$$\begin{aligned} \det(A^T A - \lambda I) &= \begin{bmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{bmatrix} \\ &= (1-\lambda)^2(2-\lambda) - (1-\lambda) - (1-\lambda) \\ &= (1-\lambda)((1-\lambda)(2-\lambda) - 2) \\ &= (1-\lambda)\lambda(\lambda - 3) = 0. \end{aligned}$$

Thus, we have $\lambda = 3, 1, 0$. Therefore, the singular values are: $\sqrt{3}, 1, 0$.

Now, to determine $V$, we solve the eigenvalue-eigenvector problems. For $\lambda = 3$:

$$\begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2x - y \\ -x - y - z \\ -y - 2z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Thus we can set $\begin{bmatrix} x & y & z \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}^T$, but we need to normalize it to have a unit length. Thus we have $\boldsymbol{v}_1 = \sqrt{6}^{-1} \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}^T$.

For $\lambda = 1$:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -y \\ -x + y - z \\ -y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Thus we can set $\begin{bmatrix} x & y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$, and after the normalization, to have a unit length, it becomes $\boldsymbol{v}_2 = \sqrt{2}^{-1} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$.

For $\lambda = 0$:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y \\ -x + 2y - z \\ -y + z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Thus, we can set $\begin{bmatrix} x & y & z \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$, and again after the normalization, it becomes $\boldsymbol{v}_3 = \sqrt{3}^{-1} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$.

Now, to compute $U$, we use the formula: $\boldsymbol{u}_j = \frac{1}{\sigma_j} A \boldsymbol{v}_j$ for $\sigma_j \neq= 0$. Thus,

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\boldsymbol{u}_2 = \frac{1}{1}\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 1\end{bmatrix}\begin{bmatrix}1/\sqrt{2}\\ 0\\ -1/\sqrt{2}\end{bmatrix} = \begin{bmatrix}-1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix}$$

Thus the SVD of $A$ is:

$$\underline{\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 1\end{bmatrix} = \begin{bmatrix}-1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}\begin{bmatrix}\sqrt{3} & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3}\\ -2/\sqrt{6} & 0 & 1/\sqrt{3}\\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3}\end{bmatrix}^T.}$$

**(b)** (5 pts) Compute the rank 1 approximation of $A$.

**Solution:** The rank 1 approximation of $A$ is of course

$$\begin{aligned}
\sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^T &= \sqrt{3}\begin{bmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}\begin{bmatrix}1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6}\end{bmatrix}\\
&= \frac{\sqrt{3}}{\sqrt{2}\cdot\sqrt{6}}\begin{bmatrix}-1\\ 1\end{bmatrix}\begin{bmatrix}1 & -2 & 1\end{bmatrix}\\
&= \frac{1}{2}\begin{bmatrix}-1 & 2 & -1\\ 1 & -2 & 1\end{bmatrix}\\
&= \underline{\begin{bmatrix}-1/2 & 1 & -1/2\\ 1/2 & -1 & 1/2\end{bmatrix}}.
\end{aligned}$$

**Problem 7** (15 pts) Consider the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

**(a)** (5 pts) What is the orthogonal projector $P$ onto $\mathcal{R}(A)$, and what is the projection of the vector $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$ under $P$?

**Solution:**

$$P_A = A(A^TA)^{-1}A^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \underline{\begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}}.$$

Therefore the image of $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$ under $P$ is:

$$P \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \underline{\begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}}.$$

**(b)** (10 pts) Same questions for $B$.

[Hint: Compute $B^\dagger$, i.e., the *pseudoinverse* of $B$, and construct an orthogonal projector using $B^\dagger$.]

**Solution:** In this case, $B$ is not of full rank and

$$B^TB = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

Thus, we cannot do $P_B = B(B^TB)^{-1}B^T$. Instead we need to compute the pseudoinverse $B^\dagger$ and construct the orthogonal projector $P_B = BB^\dagger$. To do so, we need to compute first the SVD of $B$.

$$\det \begin{bmatrix} 2-\lambda & 2 \\ 2 & 2-\lambda \end{bmatrix} = \lambda(4-\lambda) = 0.$$

So, $\lambda = 4, 0$ and $\sigma_1 = 2$, $\sigma_2 = 0$. As for the eigenvector for $\lambda = 4$,

$$\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2x + 2y \\ 2x - 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

So, $\boldsymbol{v}_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$, and after the normalization, we get: $\boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T$.

Now for the eigenvector for $\lambda = 0$,

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + 2y \\ 2x + 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus, with the normalization in mind, we get $\boldsymbol{v}_2 = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}^T$.
For $U$, we use the formula $\boldsymbol{u}_1 = \frac{1}{\sigma_1} A\boldsymbol{v}_1$ to get:

$$\boldsymbol{u}_1 = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}.$$

We need to find $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ orthogonal to $\boldsymbol{u}_1$ with $\|\boldsymbol{u}_2\| = \|\boldsymbol{u}_3\| = 1$. So, by looking the components of $\boldsymbol{u}_1$, we can easily find, for example, $\boldsymbol{u}_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$. Then, $\boldsymbol{u}_3$ can be $\boldsymbol{u}_3 = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}^T$. Thus, the SVD of $B$ is:

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}.$$

Therefore,

$$B^\dagger = V\Sigma^\dagger U^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Therefore,

$$P_B = BB^\dagger = \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \underline{\frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}}.$$

Therefore the image of $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$ under $P$ is:

$$P\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \underline{\begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}}.$$

**Problem 8** (15 pts) Suppose we are given three data points in $\mathbb{R}^2$, $(x, y) = (-\pi/2, 1), (0, 1), (\pi/2, -1)$. Now, we want to find the best function to fit these points in the form of $y = \alpha + \beta \sin x$ in the sense of the least squares.

**(a)** (5 pts) Write a system of equation in the form $A\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{x} = \begin{bmatrix} \alpha & \beta \end{bmatrix}^T$ as if this function passes through all these three points.

**Solution:** Since $\sin x = -1, 0, 1$ when $x = -\pi/2, 0, \pi/2$, respectively, the system of equation can be easily written as:

$$\underline{\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}}.$$

**(b)** (5 pts) Solve the least squares problem using the normal equation, and write the solution in the form of $y = \alpha + \beta \sin x$.

**Solution:** This is simply can be solved by

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (A^T A)^{-1} A^T \boldsymbol{b} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -1 \end{bmatrix}.$$

Thus the least square solution is:

$$\underline{y = \frac{1}{3} - \sin x.}$$

**(c)** (5 pts) Solve the least squares problem using the reduced QR factorization and confirm that your hand computed result agrees with that of (b).

**Solution:** Let the column vectors of $A$ be $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. Then the reduced QR factorization of $A$ in this case is:

$$r_{11} \leftarrow \|\boldsymbol{a}_1\| = \sqrt{3}. \boldsymbol{q}_1 \leftarrow \boldsymbol{a}_1 / r_{11} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T.$$

$$r_{12} \leftarrow \langle \boldsymbol{a}_2, \boldsymbol{q}_1 \rangle = 0;$$
$$\boldsymbol{q}_2 \leftarrow \boldsymbol{a}_2 - r_{12}\, \boldsymbol{q}_1 = \boldsymbol{a}_2;$$
$$r_{22} \leftarrow \|\boldsymbol{q}_2\| = \sqrt{2};$$
$$\boldsymbol{q}_2 \leftarrow \boldsymbol{q}_2 / r_{22} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^T.$$

Thus,

$$A = \hat{Q}\hat{R} = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} \\ 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix}.$$

Now, we transform the inconsistent system $A\boldsymbol{x} = \boldsymbol{b}$ using the orthogonal projector $\hat{Q}\hat{Q}^T$ to a consistent system: $A\boldsymbol{x} = \hat{Q}\hat{Q}^T\boldsymbol{b}$. Using the QR factorization of $A$, this simplifies to:

$$\hat{R}\boldsymbol{x} = \hat{Q}\boldsymbol{b}.$$

Therefore,

$$\boldsymbol{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \hat{R}^{-1}\hat{Q}\boldsymbol{b} = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \underline{\begin{bmatrix} 1/3 \\ -1 \end{bmatrix}}.$$

Therefore, the least squares solution based on the reduced QR factorization is the same as that by the normal equation.

**(d) Bonus Problem** (5 pts) After fitting the least squares solution you computed in (b) and (c) to the actual data, compute the residual (or error) vector and its length in 2-norm.

**Solution:** The error vector is simply,

$$\boldsymbol{b} - A\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1/3 \\ -1 \end{bmatrix} = \underline{\begin{bmatrix} -1/3 \\ 2/3 \\ -1/3 \end{bmatrix}}.$$

Thus its 2-norm is:

$$\left\| \begin{bmatrix} -1/3 \\ 2/3 \\ -1/3 \end{bmatrix} \right\| = \underline{\frac{\sqrt{6}}{3}}.$$