Practice Problems: Mathematical Foundations for Machine Learning

Saurabh

June 22, 2025

1 Solution of Linear Systems and Vector Spaces

Question 1 (Easy). *Solve the following system of linear equations:*

$$\begin{aligned} x + 2y - z &= 3 \\ 2x + y + z &= 6 \\ x - y + 2z &= 3 \end{aligned}$$

Solution. *We write this system as an augmented matrix and use Gaussian elimination to reduce it to row echelon form.*

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 2 & 1 & 1 & 6 \\ 1 & -1 & 2 & 3 \end{array}\right] \xrightarrow[R_3 \to R_3 - R_1]{R_2 \to R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & -3 & 3 & 0 \\ 0 & -3 & 3 & 0 \end{array}\right] \xrightarrow{R_2 \to -\frac{1}{3}R_2} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & -3 & 3 & 0 \end{array}\right] \xrightarrow{R_3 \to R_3 + 3R_2} \left[\begin{array}{cc} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{array}\right.$$

The system is now in row echelon form. The matrix corresponds to the equations:

$$\begin{aligned} x + 2y - z &= 3 \\ y - z &= 0 \end{aligned}$$

From the second equation, we have $y = z$. Since the third row is all zeros, we have a free variable. Let $z = t$, where t is any real number. This implies $y = t$. Substitute $y = t$ and $z = t$ into the first equation: $x + 2(t) - t = 3 \implies x + t = 3 \implies x = 3 - t$. The general solution is a line in $\mathbb{R}^3$ described by the vector $(x, y, z) = (3 - t, t, t)$ for any $t \in \mathbb{R}$.

Question 2 (Easy). *Determine if the following set of vectors in $\mathbb{R}^3$ is linearly independent:*

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Solution. *To check for linear independence, we form a matrix A whose columns are the vectors in S and check its determinant. The vectors are linearly independent if and only if the determinant is non-zero.*

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

$$\det(A) = 1\begin{vmatrix}1 & 1\\0 & 1\end{vmatrix} - (-1)\begin{vmatrix}0 & 1\\1 & 1\end{vmatrix} + 0\begin{vmatrix}0 & 1\\1 & 0\end{vmatrix}$$
$$= 1(1\cdot 1 - 1\cdot 0) + 1(0\cdot 1 - 1\cdot 1) + 0 = 1(1) + 1(-1) = 1 - 1 = 0$$

Since the determinant is 0, the set of vectors is ***linearly dependent****.*

Alternatively, we can try to find a non-trivial solution to $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$*. This gives the system:*

$$\begin{aligned} c_1 - c_2 &= 0 \implies c_1 = c_2 \\ c_2 + c_3 &= 0 \implies c_3 = -c_2 \\ c_1 + c_3 &= 0 \end{aligned}$$

Substituting the first two relations into the third gives $(c_2) + (-c_2) = 0$*, which is* $0 = 0$*. This means there are infinitely many solutions. For a non-trivial solution, pick* $c_2 = 1$*. Then* $c_1 = 1$ *and* $c_3 = -1$*. Indeed,* $1\mathbf{v}_1 + 1\mathbf{v}_2 - 1\mathbf{v}_3 = \mathbf{0}$*. Since a non-trivial solution exists, the vectors are linearly dependent.*

Question 3 (Moderate)**.** *Let* $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ *be a basis for a subspace* $W \subset \mathbb{R}^3$*, where* $\mathbf{v}_1 = (1, 2, 0)$ *and* $\mathbf{v}_2 = (3, 1, -1)$*. Determine if the vector* $\mathbf{u} = (-1, 3, 1)$ *is in the subspace* W*. If it is, express* $\mathbf{u}$ *as a linear combination of the basis vectors.*

Solution. *The vector* $\mathbf{u}$ *is in the subspace* W *if and only if it can be written as a linear combination of the basis vectors* $\mathbf{v}_1$ *and* $\mathbf{v}_2$*. We are looking for scalars* c_1, c_2 *such that* $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{u}$*.*

$$c_1\begin{pmatrix}1\\2\\0\end{pmatrix} + c_2\begin{pmatrix}3\\1\\-1\end{pmatrix} = \begin{pmatrix}-1\\3\\1\end{pmatrix}$$

This vector equation corresponds to a system of three linear equations:

$$\begin{aligned} c_1 + 3c_2 &= -1 \\ 2c_1 + c_2 &= 3 \\ -c_2 &= 1 \end{aligned}$$

From the third equation, we immediately get $c_2 = -1$*. Substitute* $c_2 = -1$ *into the second equation:* $2c_1 + (-1) = 3 \implies 2c_1 = 4 \implies c_1 = 2$*. Now we have a potential solution* $(c_1, c_2) = (2, -1)$*. We must check if these values satisfy the remaining equation (the first one) to ensure the system is consistent. Check the first equation:* $c_1 + 3c_2 = (2) + 3(-1) = 2 - 3 = -1$*. The equation holds true. Since we found a consistent solution, the vector* $\mathbf{u}$ *is in the subspace* W*. The linear combination is* $\mathbf{u} = 2\mathbf{v}_1 - \mathbf{v}_2$*.*

Question 4 (Moderate)**.** *Find a basis for the column space and the rank of the matrix* A*:*

$$A = \begin{pmatrix}1 & 2 & 0 & 1\\2 & 4 & 1 & 3\\0 & 0 & 1 & 1\end{pmatrix}$$

Solution. *To find a basis for the column space Col(A), we reduce the matrix* A *to its row echelon form to identify the pivot columns.*

$$\begin{pmatrix}1 & 2 & 0 & 1\\2 & 4 & 1 & 3\\0 & 0 & 1 & 1\end{pmatrix} \xrightarrow{R_2 \to R_2 - 2R_1} \begin{pmatrix}1 & 2 & 0 & 1\\0 & 0 & 1 & 1\\0 & 0 & 1 & 1\end{pmatrix} \xrightarrow{R_3 \to R_3 - R_2} \begin{pmatrix}1 & 2 & 0 & 1\\0 & 0 & 1 & 1\\0 & 0 & 0 & 0\end{pmatrix}$$

The leading 1's (pivots) are in the first and third columns of the row echelon form. Therefore, the corresponding columns of the ***original matrix A*** *form a basis for its column space. The pivot columns are column 1 and column 3. A basis for Col(A) is* $\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$*. The rank of a matrix is the dimension of its column space (which is equal to the number of pivot columns). There are 2 pivots, so the* ***rank of A is 2****.*

Question 5 (Moderate)**.** *Let W be the set of all vectors in $\mathbb{R}^3$ of the form $(a, b, a+b)$. Show that W is a subspace of $\mathbb{R}^3$.*

Solution. *To show that W is a subspace of $\mathbb{R}^3$, we must verify three properties:*

1. ***The zero vector is in W:*** *The zero vector in $\mathbb{R}^3$ is $\mathbf{0} = (0,0,0)$. We must check if it can be written in the form $(a, b, a+b)$. If we choose $a = 0$ and $b = 0$, the form becomes $(0, 0, 0+0) = (0,0,0)$. So, $\mathbf{0} \in W$. The set is non-empty.*

2. ***Closure under vector addition:*** *Let $\mathbf{u}$ and $\mathbf{v}$ be two arbitrary vectors in W. Then they can be written as: $\mathbf{u} = (a_1, b_1, a_1 + b_1)$ for some scalars a_1, b_1. $\mathbf{v} = (a_2, b_2, a_2 + b_2)$ for some scalars a_2, b_2. Their sum is: $\mathbf{u}+\mathbf{v} = (a_1+a_2, b_1+b_2, (a_1+b_1)+(a_2+b_2))$ $\mathbf{u}+\mathbf{v} = (a_1+a_2, b_1+b_2, (a_1+a_2)+(b_1+b_2))$. Let $a' = a_1 + a_2$ and $b' = b_1 + b_2$. Then the sum is of the form $(a', b', a'+b')$, which matches the definition of a vector in W. Thus, W is closed under addition.*

3. ***Closure under scalar multiplication:*** *Let $\mathbf{u} = (a, b, a+b)$ be a vector in W and c be any scalar. $c\mathbf{u} = c(a, b, a+b) = (ca, cb, c(a+b)) = (ca, cb, ca+cb)$. Let $a'' = ca$ and $b'' = cb$. Then the resulting vector is $(a'', b'', a''+b'')$, which is in the form required for vectors in W. Thus, W is closed under scalar multiplication.*

Since all three properties hold, W is a subspace of $\mathbb{R}^3$.

Question 6 (Tough)**.** *Consider the system of equations:*

$$\begin{aligned} x + y + kz &= 1 \\ x + ky + z &= 1 \\ kx + y + z &= -2 \end{aligned}$$

Find the values of k for which the system has: (a) a unique solution, (b) no solution, (c) infinitely many solutions.

Solution. *The nature of the solution depends on the determinant of the coefficient matrix* $A = \begin{pmatrix} 1 & 1 & k \\ 1 & k & 1 \\ k & 1 & 1 \end{pmatrix}$.

$$\begin{aligned} \det(A) &= 1(k \cdot 1 - 1 \cdot 1) - 1(1 \cdot 1 - k \cdot 1) + k(1 \cdot 1 - k \cdot k) \\ &= (k-1) - (1-k) + k(1-k^2) \\ &= k - 1 - 1 + k + k - k^3 = -k^3 + 3k - 2 \end{aligned}$$

To find the roots of this polynomial, we can test integer factors of -2. If $k = 1$, $\det(A) = -1 + 3 - 2 = 0$. So $(k-1)$ is a factor. If $k = -2$, $\det(A) = -(-8) + 3(-2) - 2 = 8 - 6 - 2 = 0$. So $(k+2)$ is a factor. By polynomial division or synthetic division, we find the complete factorization: $\det(A) = -(k-1)^2(k+2)$.

(a) ***Unique solution:*** *A unique solution exists if and only if* $\det(A) \neq 0$. *This occurs when* $-(k-1)^2(k+2) \neq 0$, *which means* $\boldsymbol{k \neq 1}$ *and* $\boldsymbol{k \neq -2}$.

(b) ***No solution or infinitely many solutions:*** *This occurs when* $\det(A) = 0$, *i.e., for* $k = 1$ *or* $k = -2$. *We must analyze these cases using the augmented matrix.*

Case 1: k = 1

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -2 \end{array}\right] \xrightarrow{R_3 \to R_3 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -3 \end{array}\right]$$

The last row represents the equation $0x + 0y + 0z = -3$, *which is* $0 = -3$. *This is a contradiction. Therefore, for* $\boldsymbol{k = 1}$, *there is* ***no solution***.

Case 2: k = -2

$$\left[\begin{array}{ccc|c} 1 & 1 & -2 & 1 \\ 1 & -2 & 1 & 1 \\ -2 & 1 & 1 & -2 \end{array}\right] \xrightarrow[R_3 \to R_3 + 2R_1]{R_2 \to R_2 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & -2 & 1 \\ 0 & -3 & 3 & 0 \\ 0 & 3 & -3 & 0 \end{array}\right] \xrightarrow{R_3 \to R_3 + R_2} \left[\begin{array}{ccc|c} 1 & 1 & -2 & 1 \\ 0 & -3 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The last row is $0 = 0$, *which is consistent. Since there is a row of zeros, we have at least one free variable, and the system is consistent. Therefore, for* $\boldsymbol{k = -2}$, *there are* ***infinitely many solutions***.

Question 7 (Tough). *Let the matrix* A *be given by:*

$$A = \begin{pmatrix} 1 & -2 & 0 & 3 \\ 2 & -4 & 1 & 8 \\ -1 & 2 & -2 & -9 \end{pmatrix}$$

(a) *Find the Reduced Row Echelon Form (RREF) of A.*

(b) *Find a basis for the null space, Null(A).*

(c) *Find a basis for the column space, Col(A).*

(d) *State the dimensions of the null space and column space and verify the Rank-Nullity Theorem.*

Solution. *(a)* ***RREF of A****: We perform Gaussian elimination.*

$$\begin{bmatrix} 1 & -2 & 0 & 3 \\ 2 & -4 & 1 & 8 \\ -1 & 2 & -2 & -9 \end{bmatrix} \xrightarrow[R_3 \to R_3 + R_1]{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -2 & -6 \end{bmatrix} \xrightarrow{R_3 \to R_3 + 2R_2} \begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

To get RREF, continue reducing:

$$\xrightarrow{R_3 \to -1/2 R_3} \begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[R_1 \to R_1 - 3R_3]{R_2 \to R_2 - 2R_3} \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This is the ***RREF of A***.

(b) ***Basis for Null(A)***: *We solve* $A\mathbf{x} = \mathbf{0}$ *using the RREF. Let* $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$*. The RREF gives the system:* $x_1 - 2x_2 = 0 \implies x_1 = 2x_2$ $x_3 = 0$ $x_4 = 0$ *The variable* x_2 *corresponds to a non-pivot column, so it is a free variable. Let* $x_2 = t$*. The solution vector is* $\mathbf{x} = (2t, t, 0, 0)^T = t(2, 1, 0, 0)^T$*. A basis for Null(A) is the vector that spans this space:* $\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$.

(c) ***Basis for Col(A)***: *The pivot columns in the RREF are columns 1, 3, and 4. We take the corresponding columns from the* ***original matrix A****. A basis for Col(A) is* $\left\{ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ 8 \\ -9 \end{pmatrix} \right\}$.

(d) ***Rank-Nullity Theorem***: *The dimension of the null space, nullity(A), is the number of free variables, which is 1. The dimension of the column space, rank(A), is the number of pivot columns, which is 3. The Rank-Nullity Theorem states that rank(A)+ nullity(A)* $= n$*, where* n *is the number of columns of A. Here,* $n = 4$*. We check:* $3 + 1 = 4$*. The theorem is verified.*

2 Analytic Geometry

Question 8 (Easy). *Given the vectors* $\mathbf{u} = (2, -1, 3)$ *and* $\mathbf{v} = (1, 1, -2)$ *in* $\mathbb{R}^3$.

(a) Find the Euclidean norm (length) of $\mathbf{u}$.

(b) Find the distance between $\mathbf{u}$ *and* $\mathbf{v}$.

Solution. *(a) The Euclidean norm* $||\mathbf{u}||$ *is given by the square root of the sum of the squares of its components.*

$$||\mathbf{u}|| = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{4 + 1 + 9} = \sqrt{14}$$

(b) The distance between $\mathbf{u}$ *and* $\mathbf{v}$ *is the norm of their difference,* $||\mathbf{u} - \mathbf{v}||$. *First, find the difference vector:*

$$\mathbf{u} - \mathbf{v} = (2 - 1, -1 - 1, 3 - (-2)) = (1, -2, 5)$$

Now, find the norm of this vector:

$$||\mathbf{u} - \mathbf{v}|| = \sqrt{1^2 + (-2)^2 + 5^2} = \sqrt{1 + 4 + 25} = \sqrt{30}$$

Question 9 (Easy). *For the vectors* $\mathbf{a} = (1, 2, -3)$ *and* $\mathbf{b} = (3, 1, 2)$:

(a) Calculate the inner product (dot product) $\mathbf{a} \cdot \mathbf{b}$.

(b) Determine if the vectors are orthogonal.

Solution. *(a) The inner product is calculated as the sum of the products of corresponding components:*

$$\mathbf{a} \cdot \mathbf{b} = (1)(3) + (2)(1) + (-3)(2) = 3 + 2 - 6 = -1$$

(b) Two vectors are orthogonal if their inner product is zero. Since $\mathbf{a} \cdot \mathbf{b} = -1 \neq 0$, *the vectors are* ***not orthogonal***.

Question 10 (Moderate). *Find a non-zero vector* $\mathbf{w}$ *that is orthogonal to both* $\mathbf{u} = (1, 1, -2)$ *and* $\mathbf{v} = (2, -1, 1)$.

Solution. *A vector that is orthogonal to two other vectors in* $\mathbb{R}^3$ *can be found by computing their cross product. The cross product* $\mathbf{u} \times \mathbf{v}$ *produces a vector that is orthogonal to both* $\mathbf{u}$ *and* $\mathbf{v}$.

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & -2 \\ 2 & -1 & 1 \end{vmatrix}$$

where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ *are the standard basis vectors for* $\mathbb{R}^3$.

$$\mathbf{w} = \mathbf{i}(1 \cdot 1 - (-2) \cdot (-1)) - \mathbf{j}(1 \cdot 1 - (-2) \cdot 2) + \mathbf{k}(1 \cdot (-1) - 1 \cdot 2)$$

$$\mathbf{w} = \mathbf{i}(1 - 2) - \mathbf{j}(1 + 4) + \mathbf{k}(-1 - 2)$$

$$\mathbf{w} = -1\mathbf{i} - 5\mathbf{j} - 3\mathbf{k} = (-1, -5, -3)$$

To verify, we check the dot products:

$$\mathbf{w} \cdot \mathbf{u} = (-1)(1) + (-5)(1) + (-3)(-2) = -1 - 5 + 6 = 0$$

$$\mathbf{w} \cdot \mathbf{v} = (-1)(2) + (-5)(-1) + (-3)(1) = -2 + 5 - 3 = 0$$

Since both dot products are zero, the vector $\mathbf{w} = (-1, -5, -3)$ *is orthogonal to both* $\mathbf{u}$ *and* $\mathbf{v}$. *Any non-zero scalar multiple of* $\mathbf{w}$, *such as* $(1, 5, 3)$, *is also a valid answer.*

Question 11 (Moderate). *An inner product on $\mathbb{R}^2$ can be defined by a symmetric matrix. Consider the inner product defined by the matrix $A = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$, where $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$. Let $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$.*

(a) Calculate the inner product $\langle \mathbf{x}, \mathbf{y} \rangle$.

(b) Calculate the norm of $\mathbf{x}$ induced by this inner product, $||\mathbf{x}||_A = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

Solution. *(a)* ***Calculate the inner product:***

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

First, compute $A\mathbf{y}$:

$$A\mathbf{y} = \begin{pmatrix} 2(3) + (-1)(-1) \\ -1(3) + 3(-1) \end{pmatrix} = \begin{pmatrix} 7 \\ -6 \end{pmatrix}$$

Now compute the full product:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 7 \\ -6 \end{pmatrix} = 1(7) + 2(-6) = 7 - 12 = -5$$

(b) ***Calculate the norm of x:*** *The squared norm is $||\mathbf{x}||_A^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^T A \mathbf{x}$.*

$$||\mathbf{x}||_A^2 = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

First, compute $A\mathbf{x}$:

$$A\mathbf{x} = \begin{pmatrix} 2(1) + (-1)(2) \\ -1(1) + 3(2) \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$$

Now compute the full product:

$$||\mathbf{x}||_A^2 = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 5 \end{pmatrix} = 1(0) + 2(5) = 10$$

The norm is therefore $||\mathbf{x}||_A = \sqrt{10}$.

Question 12 (Moderate). *Given the vectors $\mathbf{v}_1 = (1, 1, 0)$, $\mathbf{v}_2 = (1, 0, 1)$, and $\mathbf{v}_3 = (0, 1, 1)$, which form a basis for $\mathbb{R}^3$. Use the Gram-Schmidt process to find an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$.*

Solution

The Gram-Schmidt process transforms a set of linearly independent vectors into an orthogonal basis, which can then be normalized to create an orthonormal basis. Let the given vectors be $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. We will first construct an orthogonal basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and then normalize each vector to get the orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$.

Step 1: Construct $\mathbf{q}_1$ from $\mathbf{v}_1$

First, we set the initial orthogonal vector $\mathbf{u}_1$ *to be equal to* $\mathbf{v}_1$*.*

$$\mathbf{u}_1 = \mathbf{v}_1 = (1, 1, 0)$$

Next, we find its magnitude (norm):

$$||\mathbf{u}_1|| = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}$$

The first orthonormal vector $\mathbf{q}_1$ *is found by normalizing* $\mathbf{u}_1$*:*

$$\mathbf{q}_1 = \frac{\mathbf{u}_1}{||\mathbf{u}_1||} = \frac{1}{\sqrt{2}}(1, 1, 0)$$

Step 2: Construct $\mathbf{q}_2$ from $\mathbf{v}_2$

The second orthogonal vector, $\mathbf{u}_2$*, is found by subtracting the projection of* $\mathbf{v}_2$ *onto* $\mathbf{u}_1$ *from* $\mathbf{v}_2$*.*

$$\mathbf{u}_2 = \mathbf{v}_2 - proj_{\mathbf{u}_1}(\mathbf{v}_2) = \mathbf{v}_2 - \frac{\mathbf{v}_2\mathbf{u}_1}{||\mathbf{u}_1||^2}\mathbf{u}_1$$

First, we calculate the dot product $\mathbf{v}_2\mathbf{u}_1$*:*

$$\mathbf{v}_2\mathbf{u}_1 = (1, 0, 1)(1, 1, 0) = 1(1) + 0(1) + 1(0) = 1$$

Now, substitute this into the equation for $\mathbf{u}_2$*:*

$$\mathbf{u}_2 = (1, 0, 1) - \frac{1}{2}(1, 1, 0) = \left(1 - \frac{1}{2}, 0 - \frac{1}{2}, 1 - 0\right) = \left(\frac{1}{2}, -\frac{1}{2}, 1\right)$$

To simplify the arithmetic in the next step, we can use a scaled version of $\mathbf{u}_2$*, as its direction is what matters. Let's use* $\mathbf{u}'_2 = 2\mathbf{u}_2$*:*

$$\mathbf{u}'_2 = 2\left(\frac{1}{2}, -\frac{1}{2}, 1\right) = (1, -1, 2)$$

Now, we normalize this cleaner vector to find $\mathbf{q}_2$*:*

$$||\mathbf{u}'_2|| = \sqrt{1^2 + (-1)^2 + 2^2} = \sqrt{1 + 1 + 4} = \sqrt{6}$$

$$\mathbf{q}_2 = \frac{\mathbf{u}'_2}{||\mathbf{u}'_2||} = \frac{1}{\sqrt{6}}(1, -1, 2)$$

Step 3: Construct $\mathbf{q}_3$ from $\mathbf{v}_3$

The third orthogonal vector, $\mathbf{u}_3$*, is found by subtracting the projections of* $\mathbf{v}_3$ *onto the previously found orthogonal vectors,* $\mathbf{u}_1$ *and* $\mathbf{u}'_2$*.*

$$\mathbf{u}_3 = \mathbf{v}_3 - proj_{\mathbf{u}_1}(\mathbf{v}_3) - proj_{\mathbf{u}'_2}(\mathbf{v}_3) = \mathbf{v}_3 - \frac{\mathbf{v}_3\mathbf{u}_1}{||\mathbf{u}_1||^2}\mathbf{u}_1 - \frac{\mathbf{v}_3\mathbf{u}'_2}{||\mathbf{u}'_2||^2}\mathbf{u}'_2$$

We calculate the required dot products first:

$$\mathbf{v}_3\mathbf{u}_1 = (0, 1, 1)(1, 1, 0) = 0(1) + 1(1) + 1(0) = 1$$
$$\mathbf{v}_3\mathbf{u}'_2 = (0, 1, 1)(1, -1, 2) = 0(1) + 1(-1) + 1(2) = 1$$

Now, we substitute these values into the equation for $\mathbf{u}_3$*:*

$$\begin{aligned}
\mathbf{u}_3 &= (0,1,1) - \frac{1}{2}(1,1,0) - \frac{1}{6}(1,-1,2) \\
&= (0,1,1) - \left(\frac{1}{2}, \frac{1}{2}, 0\right) - \left(\frac{1}{6}, -\frac{1}{6}, \frac{1}{3}\right) \\
&= \left(0 - \frac{1}{2} - \frac{1}{6}, \quad 1 - \frac{1}{2} + \frac{1}{6}, \quad 1 - 0 - \frac{1}{3}\right) \\
&= \left(-\frac{3}{6} - \frac{1}{6}, \quad \frac{6}{6} - \frac{3}{6} + \frac{1}{6}, \quad \frac{3}{3} - \frac{1}{3}\right) \\
&= \left(-\frac{4}{6}, \frac{4}{6}, \frac{2}{3}\right) = \left(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right)
\end{aligned}$$

Again, we can simplify by scaling. Let's use $\mathbf{u}_3' = \frac{3}{2}\mathbf{u}_3$ *to eliminate the fractions:*

$$\mathbf{u}_3' = \frac{3}{2}\left(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right) = (-1,1,1)$$

Finally, we normalize $\mathbf{u}_3'$ *to get* $\mathbf{q}_3$*:*

$$||\mathbf{u}_3'|| = \sqrt{(-1)^2 + 1^2 + 1^2} = \sqrt{3}$$

$$\mathbf{q}_3 = \frac{\mathbf{u}_3'}{||\mathbf{u}_3'||} = \frac{1}{\sqrt{3}}(-1,1,1)$$

Conclusion: The Orthonormal Basis

The resulting orthonormal basis for $\mathbb{R}^3$ *is* $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$*, where:*

$$\begin{aligned}
\mathbf{q}_1 &= \frac{1}{\sqrt{2}}(1,1,0) \\
\mathbf{q}_2 &= \frac{1}{\sqrt{6}}(1,-1,2) \\
\mathbf{q}_3 &= \frac{1}{\sqrt{3}}(-1,1,1)
\end{aligned}$$

Question 13 (Tough)**.** *Let* W *be the subspace of* $\mathbb{R}^4$ *spanned by the orthogonal vectors* $\mathbf{u}_1 = (1,1,0,1)$ *and* $\mathbf{u}_2 = (1,-1,1,0)$*. Let* $\mathbf{y} = (2,3,2,1)$*. Find the orthogonal projection of* $\mathbf{y}$ *onto the subspace* W*, and find the shortest distance from* $\mathbf{y}$ *to* W*.*

Solution. *The* ***orthogonal projection*** *of* $\mathbf{y}$ *onto the subspace* W*, denoted* $\hat{\mathbf{y}}$*, is the point in* W *closest to* $\mathbf{y}$*. Since* W *is spanned by an orthogonal set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2\}$*, the formula is:*

$$\hat{\mathbf{y}} = proj_W(\mathbf{y}) = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2$$

First, we calculate the required dot products:

$$\mathbf{y} \cdot \mathbf{u}_1 = (2)(1) + (3)(1) + (2)(0) + (1)(1) = 2 + 3 + 0 + 1 = 6$$

$$\mathbf{u}_1 \cdot \mathbf{u}_1 = ||\mathbf{u}_1||^2 = 1^2 + 1^2 + 0^2 + 1^2 = 3$$

$$\mathbf{y}\cdot\mathbf{u}_2 = (2)(1)+(3)(-1)+(2)(1)+(1)(0) = 2-3+2+0 = 1$$
$$\mathbf{u}_2\cdot\mathbf{u}_2 = ||\mathbf{u}_2||^2 = 1^2+(-1)^2+1^2+0^2 = 3$$

Now, substitute these values into the projection formula:

$$\hat{\mathbf{y}} = \frac{6}{3}\mathbf{u}_1 + \frac{1}{3}\mathbf{u}_2 = 2(1,1,0,1) + \frac{1}{3}(1,-1,1,0)$$
$$\hat{\mathbf{y}} = (2,2,0,2) + (\frac{1}{3}, -\frac{1}{3}, \frac{1}{3}, 0) = \left(\frac{7}{3}, \frac{5}{3}, \frac{1}{3}, 2\right)$$

This vector $\hat{\mathbf{y}}$ is the orthogonal projection of $\mathbf{y}$ onto W.

The ***shortest distance*** *from $\mathbf{y}$ to W is the norm of the component of $\mathbf{y}$ orthogonal to W, which is the vector $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.*

$$\mathbf{z} = (2,3,2,1) - \left(\frac{7}{3}, \frac{5}{3}, \frac{1}{3}, 2\right) = \left(2-\frac{7}{3}, 3-\frac{5}{3}, 2-\frac{1}{3}, 1-2\right) = \left(-\frac{1}{3}, \frac{4}{3}, \frac{5}{3}, -1\right)$$

The distance is the norm of $\mathbf{z}$:

$$||\mathbf{z}|| = \sqrt{\left(-\frac{1}{3}\right)^2 + \left(\frac{4}{3}\right)^2 + \left(\frac{5}{3}\right)^2 + (-1)^2} = \sqrt{\frac{1}{9}+\frac{16}{9}+\frac{25}{9}+\frac{9}{9}}$$
$$||\mathbf{z}|| = \sqrt{\frac{1+16+25+9}{9}} = \sqrt{\frac{51}{9}} = \frac{\sqrt{51}}{3}$$

The shortest distance is $\boldsymbol{\frac{\sqrt{51}}{3}}$.

Question 14 (Tough)**.** *Find the orthogonal projection of the point $P(1,0,0)$ onto the plane defined by the equation $x+2y+z=6$.*

Solution. *Let the plane be Π and the point be P. Let the projection of P onto the plane be Q.* ***Step 1: Identify the Normal Vector*** *The vector $\vec{PQ}$ must be orthogonal to the plane. The normal vector to the plane $x+2y+z=6$ is given by the coefficients of x, y, z. So, $\mathbf{n} = \langle 1,2,1\rangle$.*

Step 2: Define the line from P to Q *The vector $\vec{PQ}$ must be parallel to the normal vector $\mathbf{n}$. This means we can define a line from P to Q as $Q(t) = P + t\mathbf{n}$ for some scalar t. Let the coordinates of the projection point Q be (q_1, q_2, q_3). $Q(t) = (1,0,0) + t(1,2,1) = (1+t, 2t, t)$. So, $q_1 = 1+t$, $q_2 = 2t$, $q_3 = t$.*

Step 3: Solve for t *Since the point Q lies on the plane Π, its coordinates must satisfy the plane's equation. Substitute the expressions for q_1, q_2, q_3 into the plane equation:*

$$(1+t) + 2(2t) + (t) = 6$$
$$1+t+4t+t = 6$$
$$6t+1 = 6 \implies 6t = 5 \implies t = \frac{5}{6}$$

Step 4: Find the coordinates of Q *Now we find the coordinates of the projection point Q by substituting $t = 5/6$ back into our expressions from Step 2: $q_1 = 1+\frac{5}{6} = \frac{11}{6}$ $q_2 = 2\left(\frac{5}{6}\right) = \frac{10}{6} = \frac{5}{3}$ $q_3 = \frac{5}{6}$ The orthogonal projection of the point $P(1,0,0)$ onto the plane is the point $\boldsymbol{Q\left(\frac{11}{6}, \frac{5}{3}, \frac{5}{6}\right)}$.*

3 Determinant, Eigenvalues, SVD

Question 15 (Easy). *For the matrix* $A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 4 & -1 \\ 2 & 1 & 5 \end{pmatrix}$*:*

(a) Calculate the Trace of A*.*

(b) Calculate the Determinant of A*.*

Solution. *(a)* ***Trace of A****: The trace of a square matrix is the sum of the elements on its main diagonal.*

$$Tr(A) = 1 + 4 + 5 = 10$$

(b) ***Determinant of A****: The determinant can be calculated using cofactor expansion. Let's expand along the first row.*

$$\begin{aligned} \det(A) &= 1 \begin{vmatrix} 4 & -1 \\ 1 & 5 \end{vmatrix} - (-2) \begin{vmatrix} 0 & -1 \\ 2 & 5 \end{vmatrix} + 3 \begin{vmatrix} 0 & 4 \\ 2 & 1 \end{vmatrix} \\ &= 1(4(5) - (-1)(1)) + 2(0(5) - (-1)(2)) + 3(0(1) - 4(2)) \\ &= 1(20 + 1) + 2(0 + 2) + 3(0 - 8) \\ &= 1(21) + 2(2) + 3(-8) \\ &= 21 + 4 - 24 = 1 \end{aligned}$$

Question 16 (Easy). *Find the eigenvalues of the matrix* $B = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$.

Solution. *To find the eigenvalues, we solve the characteristic equation* $\det(B - \lambda I) = 0$*, where* I *is the identity matrix and* λ *represents the eigenvalues.*

$$B - \lambda I = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 - \lambda & -1 \\ 2 & 1 - \lambda \end{pmatrix}$$

Now, we compute the determinant of this matrix:

$$\begin{aligned} \det(B - \lambda I) &= (4 - \lambda)(1 - \lambda) - (-1)(2) \\ &= 4 - 4\lambda - \lambda + \lambda^2 + 2 \\ &= \lambda^2 - 5\lambda + 6 \end{aligned}$$

Set the characteristic polynomial to zero to find the eigenvalues:

$$\lambda^2 - 5\lambda + 6 = 0$$

Factoring the quadratic equation:

$$(\lambda - 2)(\lambda - 3) = 0$$

The eigenvalues are $\boldsymbol{\lambda_1 = 2}$ *and* $\boldsymbol{\lambda_2 = 3}$.

Question 17 (Moderate). *Consider the symmetric matrix* A*:*

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

(a) *Given that $\lambda_1 = 2$ is an eigenvalue of A, use the properties of trace and determinant to find the other two eigenvalues.*

(b) *Find a basis for the eigenspace corresponding to the eigenvalue $\lambda = 2$.*

Solution. *(a)* ***Find other eigenvalues using Trace and Determinant****: Let the eigenvalues be $\lambda_1, \lambda_2, \lambda_3$. We are given $\lambda_1 = 2$. The trace of a matrix is the sum of its eigenvalues: $Tr(A) = \lambda_1 + \lambda_2 + \lambda_3$. The trace is also the sum of the diagonal elements: $Tr(A) = 4 + 4 + 4 = 12$. So, $2 + \lambda_2 + \lambda_3 = 12 \implies \lambda_2 + \lambda_3 = 10$. (Eq. 1)*

The determinant of a matrix is the product of its eigenvalues: $\det(A) = \lambda_1\lambda_2\lambda_3$. $\det(A) = 4(16-4) - 2(8-4) + 2(4-8) = 4(12) - 2(4) + 2(-4) = 48 - 8 - 8 = 32$. *So, $2\lambda_2\lambda_3 = 32 \implies \lambda_2\lambda_3 = 16$. (Eq. 2)*

We solve the system from Eq. 1 and Eq. 2. From Eq. 1, $\lambda_3 = 10 - \lambda_2$. Substitute this into Eq. 2: $\lambda_2(10 - \lambda_2) = 16 \implies 10\lambda_2 - \lambda_2^2 = 16 \implies \lambda_2^2 - 10\lambda_2 + 16 = 0$. Factoring this quadratic: $(\lambda_2 - 2)(\lambda_2 - 8) = 0$. This gives $\lambda_2 = 2$ and $\lambda_3 = 8$. The three eigenvalues are $\{2, 2, 8\}$.

(b) ***Find basis for the eigenspace of*** $\lambda = 2$*: We need to find the basis for the null space of $(A - 2I)$.*

$$A - 2I = \begin{pmatrix} 4-2 & 2 & 2 \\ 2 & 4-2 & 2 \\ 2 & 2 & 4-2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

We solve $(A - 2I)\mathbf{x} = \mathbf{0}$ by row reducing the matrix:

$$\begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \xrightarrow{R_1 \to 1/2 R_1} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \xrightarrow[R_3 \to R_3 - 2R_1]{R_2 \to R_2 - 2R_1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

This gives the single equation $x_1 + x_2 + x_3 = 0$. There is one pivot and two free variables. Let $x_2 = s$ and $x_3 = t$. Then $x_1 = -s - t$. The general solution (the eigenvectors) is $\mathbf{x} = \begin{pmatrix} -s-t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$. A basis for the eigenspace is the set of vectors spanning this solution space:

$$\mathcal{B}_{\lambda=2} = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Question 18 (Moderate)**.** *Find the eigenvalues and corresponding eigenvectors of the matrix:*

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

Solution. *First, find the eigenvalues from* $\det(A - \lambda I) = 0$.

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{vmatrix} \\
&= (2-\lambda)\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ 0 & 2-\lambda \end{vmatrix} \\
&= (2-\lambda)((2-\lambda)^2 - 1) - 1(2-\lambda) \\
&= (2-\lambda)[(2-\lambda)^2 - 1 - 1] = (2-\lambda)[\lambda^2 - 4\lambda + 4 - 2] \\
&= (2-\lambda)(\lambda^2 - 4\lambda + 2) = 0
\end{aligned}$$

One eigenvalue is $\boldsymbol{\lambda_1 = 2}$. *For the others, we use the quadratic formula on* $\lambda^2 - 4\lambda + 2 = 0$*:* $\lambda = \frac{-(-4)\pm\sqrt{(-4)^2-4(1)(2)}}{2(1)} = \frac{4\pm\sqrt{16-8}}{2} = \frac{4\pm\sqrt{8}}{2} = \frac{4\pm 2\sqrt{2}}{2} = 2 \pm \sqrt{2}$. *The other eigenvalues are* $\boldsymbol{\lambda_2 = 2 + \sqrt{2}}$ *and* $\boldsymbol{\lambda_3 = 2 - \sqrt{2}}$.

Now, find the eigenvector for each eigenvalue by solving $(A - \lambda I)\mathbf{v} = \mathbf{0}$.

- *For* $\boldsymbol{\lambda_1 = 2}$*:* $(A - 2I)\mathbf{v} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0}$. *This gives the system* $x_2 = 0$ *and* $x_1 + x_3 = 0 \implies x_1 = -x_3$. *Let* $x_3 = t$. *Then* $x_1 = -t, x_2 = 0$. *An eigenvector is* $\mathbf{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

- *For* $\boldsymbol{\lambda_2 = 2 + \sqrt{2}}$*:* $(A - (2+\sqrt{2})I)\mathbf{v} = \begin{pmatrix} -\sqrt{2} & 1 & 0 \\ 1 & -\sqrt{2} & 1 \\ 0 & 1 & -\sqrt{2} \end{pmatrix}\mathbf{v} = \mathbf{0}$. *From the first row:* $-\sqrt{2}x_1 + x_2 = 0 \implies x_2 = \sqrt{2}x_1$. *From the third row:* $x_2 - \sqrt{2}x_3 = 0 \implies x_2 = \sqrt{2}x_3$. *These together imply* $x_1 = x_3$. *Let* $x_1 = t$. *Then* $x_3 = t$ *and* $x_2 = \sqrt{2}t$. *An eigenvector is* $\mathbf{v}_2 = \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$.

- *For* $\boldsymbol{\lambda_3 = 2 - \sqrt{2}}$*:* $(A - (2-\sqrt{2})I)\mathbf{v} = \begin{pmatrix} \sqrt{2} & 1 & 0 \\ 1 & \sqrt{2} & 1 \\ 0 & 1 & \sqrt{2} \end{pmatrix}\mathbf{v} = \mathbf{0}$. *From the first row:* $\sqrt{2}x_1 + x_2 = 0 \implies x_2 = -\sqrt{2}x_1$. *From the third row:* $x_2 + \sqrt{2}x_3 = 0 \implies x_2 = -\sqrt{2}x_3$. *These imply* $x_1 = x_3$. *Let* $x_1 = t$. *Then* $x_3 = t$ *and* $x_2 = -\sqrt{2}t$. *An eigenvector is* $\mathbf{v}_3 = \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$.

Question 19 (Moderate)**.** *Given the matrix* $A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$. *Find matrices* P *and* D *such that* $A = PDP^{-1}$, *where* D *is a diagonal matrix.*

Solution. *This is an eigendecomposition problem. We must find the eigenvalues and corresponding eigenvectors of* A. ***Step 1: Find Eigenvalues*** *Characteristic equation:*

$\det(A - \lambda I) = 0$.

$$\det\begin{pmatrix} 1-\lambda & 4 \\ 1 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 4 = 0$$

$$(1-\lambda)^2 = 4 \implies 1-\lambda = \pm 2$$

This gives two possibilities: $1-\lambda = 2 \implies \lambda_1 = -1$. $1-\lambda = -2 \implies \lambda_2 = 3$. ***Step 2: Find Eigenvectors***

- *For* $\lambda_1 = -1$*: Solve* $(A-(-1)I)\mathbf{v} = \mathbf{0}$. $\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{0}$. *This gives the equation* $x_1 + 2x_2 = 0 \implies x_1 = -2x_2$. *Let* $x_2 = 1$*, then* $x_1 = -2$. *An eigenvector is* $\mathbf{v}_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

- *For* $\lambda_2 = 3$*: Solve* $(A-3I)\mathbf{v} = \mathbf{0}$. $\begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{0}$. *This gives the equation* $x_1 - 2x_2 = 0 \implies x_1 = 2x_2$. *Let* $x_2 = 1$*, then* $x_1 = 2$. *An eigenvector is* $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Step 3: Construct P and D *The diagonal matrix* D *contains the eigenvalues on its diagonal, and the matrix* P *contains the corresponding eigenvectors as its columns.*

$$D = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}, \quad P = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix}$$

To complete the decomposition $A = PDP^{-1}$*, we need* P^{-1}*. For a 2x2 matrix* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$*, the inverse is* $\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

$$\det(P) = (-2)(1) - (2)(1) = -4$$

$$P^{-1} = -\frac{1}{4}\begin{pmatrix} 1 & -2 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} -1/4 & 1/2 \\ 1/4 & 1/2 \end{pmatrix}$$

The full diagonalization is $A = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} -1/4 & 1/2 \\ 1/4 & 1/2 \end{pmatrix}$.

Question 20 (Tough)**.** *Two* $n \times n$ *matrices* A *and* B *are simultaneously diagonalizable if they share a common basis of eigenvectors* ($A = PD_AP^{-1}$ *and* $B = PD_BP^{-1}$)*. This is possible if and only if* A *and* B *commute* ($AB = BA$)*. You are given:*

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \quad D_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Assuming A *and* B *are simultaneously diagonalizable, find the matrix* B.

Solution. ***Step 1: Find the eigenvector matrix P from A****: First, we find the eigenvalues of* A*. The characteristic equation* $\det(A-\lambda I) = 0$ *is:* $\det(A-\lambda I) = (1-\lambda)((2-\lambda)(1-\lambda)-1) + 1(-(1-\lambda)) = (1-\lambda)[\lambda^2 - 3\lambda + 2 - 1 - 1] = (1-\lambda)(\lambda^2 - 3\lambda) = \lambda(1-\lambda)(\lambda-3) = 0$. *The eigenvalues of* A *are* $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 3$.

Since A *is symmetric, its eigenvectors will be orthogonal. We find normalized eigenvectors to form an orthogonal matrix* P.

- *For* $\lambda_1 = 0$*:* $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \mathbf{x} = \mathbf{0} \implies \mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. *Normalized:* $\mathbf{p}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

- *For* $\lambda_2 = 1$*:* $\begin{pmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \mathbf{x} = \mathbf{0} \implies \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. *Normalized:* $\mathbf{p}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

- *For* $\lambda_3 = 3$*:* $\begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{pmatrix} \mathbf{x} = \mathbf{0} \implies \mathbf{v}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$. *Normalized:* $\mathbf{p}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$.

The matrix P has these eigenvectors as columns, ordered according to the eigenvalues of A (0, 1, 3).

$$P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix}$$

***Step 2: Compute B**: We use the diagonalization formula $B = PD_BP^{-1}$. Since P is an orthogonal matrix (its columns are orthonormal), we have $P^{-1} = P^T$.*

$$B = PD_BP^T = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{pmatrix}$$

First, compute PD_B:

$$PD_B = \begin{pmatrix} 1(1/\sqrt{3}) & 2(1/\sqrt{2}) & -1(1/\sqrt{6}) \\ 1(1/\sqrt{3}) & 2(0) & -1(-2/\sqrt{6}) \\ 1(1/\sqrt{3}) & 2(-1/\sqrt{2}) & -1(1/\sqrt{6}) \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & 2/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & -2/\sqrt{2} & -1/\sqrt{6} \end{pmatrix}$$

Now, multiply by P^T: $B = (PD_B)P^T$.

$$B_{11} = \frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}} + \frac{2}{\sqrt{2}}\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{6}}\frac{1}{\sqrt{6}} = \frac{1}{3} + \frac{2}{2} - \frac{1}{6} = \frac{2+6-1}{6} = \frac{7}{6}$$

$$B_{12} = \frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}} + \frac{2}{\sqrt{2}}(0) - \frac{1}{\sqrt{6}}\frac{-2}{\sqrt{6}} = \frac{1}{3} + \frac{2}{6} = \frac{4}{6}$$

$$B_{13} = \frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}} + \frac{2}{\sqrt{2}}\frac{-1}{\sqrt{2}} - \frac{1}{\sqrt{6}}\frac{1}{\sqrt{6}} = \frac{1}{3} - \frac{2}{2} - \frac{1}{6} = \frac{2-6-1}{6} = -\frac{5}{6}$$

By symmetry (B will be symmetric as D_B is diagonal and P is orthogonal), $B_{21} = B_{12}$, $B_{31} = B_{13}$, etc.

$$B_{22} = \frac{1}{3} + 0 + \frac{2}{\sqrt{6}}\frac{-2}{\sqrt{6}} = \frac{1}{3} - \frac{4}{6} = \frac{2-4}{6} = -\frac{2}{6}$$

$$B_{23} = \frac{1}{3} + 0 + \frac{2}{\sqrt{6}}\frac{1}{\sqrt{6}} = \frac{1}{3} + \frac{2}{6} = \frac{4}{6}$$

$$B_{33} = \frac{1}{3} + \frac{-2}{\sqrt{2}}\frac{-1}{\sqrt{2}} - \frac{1}{\sqrt{6}}\frac{1}{\sqrt{6}} = \frac{1}{3} + \frac{2}{2} - \frac{1}{6} = \frac{2+6-1}{6} = \frac{7}{6}$$

$$B = \frac{1}{6}\begin{pmatrix} 7 & 4 & -5 \\ 4 & -2 & 4 \\ -5 & 4 & 7 \end{pmatrix}$$

Question 21 (Tough). *Find the Singular Value Decomposition (SVD) of the matrix* $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$. *Then, find the best rank-1 approximation of A.*

Solution. *The SVD of A is* $A = U\Sigma V^T$.

1. ***Find V and Σ from** A^TA: $A^TA = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Find eigenvalues of A^TA: $\det(A^TA - \lambda I) = (2-\lambda)^2 - 1 = 0 \implies 2 - \lambda = \pm 1$. The eigenvalues are $\lambda_1 = 3, \lambda_2 = 1$. The singular values are their square roots: $\sigma_1 = \sqrt{3}, \sigma_2 = 1$. The matrix Σ is an $m \times n$ matrix (same size as A) with singular values on the diagonal: $\Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$.*

   Find normalized eigenvectors of A^TA for V: For $\lambda_1 = 3$: $(A^TA - 3I)\mathbf{v} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\mathbf{v} = \mathbf{0} \implies -x_1 + x_2 = 0$. Eigenvector is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Normalized: $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. For $\lambda_2 = 1$: $(A^TA - 1I)\mathbf{v} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\mathbf{v} = \mathbf{0} \implies x_1 + x_2 = 0$. Eigenvector is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Normalized: $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. So, $V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$.

2. ***Find U**: The columns of U are found by $\mathbf{u}_i = \frac{1}{\sigma_i}A\mathbf{v}_i$. $\mathbf{u}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$. $\mathbf{u}_2 = \frac{1}{1}\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$. Since U must be a 3×3 orthogonal matrix, we need a third vector $\mathbf{u}_3$ orthogonal to $\mathbf{u}_1, \mathbf{u}_2$. We can find it with the cross product of the non-normalized vectors $(2, 1, 1)$ and $(0, -1, 1)$. $(2, 1, 1) \times (0, -1, 1) = \det\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix} = \mathbf{i}(1 - (-1)) - \mathbf{j}(2 - 0) + \mathbf{k}(-2 - 0) = (2, -2, -2)$.*

   We can use the simpler parallel vector $(1, -1, -1)$. Normalized, $\mathbf{u}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$.

So, $U = \begin{pmatrix} 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \end{pmatrix}$.

3. ***Rank-1 Approximation***: *The best rank-k approximation is* $A_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$. *For rank-1, this is* $A_1 = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T$.

$$A_1 = \sqrt{3} \left(\frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right) \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \right)$$

$$A_1 = \frac{\sqrt{3}}{\sqrt{12}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} = \frac{\sqrt{3}}{2\sqrt{3}} \begin{pmatrix} 2 & 2 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 2 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

4 Differentiation and Gradients

Question 22 (Easy)**.** *Find the derivative of the univariate function $f(x) = x^3 \sin(x)$.*

Solution. *We use the product rule for differentiation, which states that $(u(x)v(x))' = u'(x)v(x) + u(x)v'(x)$. Let $u(x) = x^3$ and $v(x) = \sin(x)$. Then their derivatives are $u'(x) = 3x^2$ and $v'(x) = \cos(x)$. Applying the product rule:*

$$f'(x) = (3x^2)(\sin(x)) + (x^3)(\cos(x)) = 3x^2 \sin(x) + x^3 \cos(x)$$

Question 23 (Easy)**.** *Given the function $f(x, y) = 3x^2y^3 + e^{2x} - \ln(y)$, find the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.*

Solution. *To find $\frac{\partial f}{\partial x}$, we treat the variable y as a constant.*

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}(3x^2y^3) + \frac{\partial}{\partial x}(e^{2x}) - \frac{\partial}{\partial x}(\ln(y)) \\ &= (3y^3)(2x) + e^{2x}(2) - 0 = 6xy^3 + 2e^{2x}\end{aligned}$$

To find $\frac{\partial f}{\partial y}$, we treat the variable x as a constant.

$$\begin{aligned}\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}(3x^2y^3) + \frac{\partial}{\partial y}(e^{2x}) - \frac{\partial}{\partial y}(\ln(y)) \\ &= (3x^2)(3y^2) + 0 - \frac{1}{y} = 9x^2y^2 - \frac{1}{y}\end{aligned}$$

Question 24 (Moderate)**.** *Find the gradient of the function $f(x, y, z) = x^2 y \cos(z)$ at the point $P(2, 1, \pi)$.*

Solution. *The gradient of f is the vector of its partial derivatives: $\nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle$.*

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^2 y \cos(z)) = 2xy \cos(z)$$

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(x^2 y \cos(z)) = x^2 \cos(z)$$

$$\frac{\partial f}{\partial z} = \frac{\partial}{\partial z}(x^2 y \cos(z)) = -x^2 y \sin(z)$$

The gradient vector is $\nabla f(x, y, z) = \langle 2xy \cos(z), x^2 \cos(z), -x^2 y \sin(z) \rangle$. Now, evaluate the gradient at the point $P(2, 1, \pi)$. We use $\cos(\pi) = -1$ and $\sin(\pi) = 0$.

$$\left.\frac{\partial f}{\partial x}\right|_{(2,1,\pi)} = 2(2)(1)\cos(\pi) = 4(-1) = -4$$

$$\left.\frac{\partial f}{\partial y}\right|_{(2,1,\pi)} = (2)^2 \cos(\pi) = 4(-1) = -4$$

$$\left.\frac{\partial f}{\partial z}\right|_{(2,1,\pi)} = -(2)^2(1)\sin(\pi) = -4(0) = 0$$

The gradient at P is $\nabla f(2, 1, \pi) = \langle -4, -4, 0 \rangle$.

Question 25 (Moderate). *Let* $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ *and let* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ *be a vector-valued function defined by:*

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 x_2 \end{pmatrix}$$

Find the Jacobian matrix of $\mathbf{f}$*.*

Solution. *The Jacobian matrix* J *of* $\mathbf{f}$ *is an* $m \times n$ *matrix (here* 2×2*) where the entry* J_{ij} *is* $\frac{\partial f_i}{\partial x_j}$*.*

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix}$$

We calculate the four required partial derivatives:

$$\frac{\partial f_1}{\partial x_1} = \frac{\partial}{\partial x_1}(x_1^2 + x_2^2) = 2x_1$$

$$\frac{\partial f_1}{\partial x_2} = \frac{\partial}{\partial x_2}(x_1^2 + x_2^2) = 2x_2$$

$$\frac{\partial f_2}{\partial x_1} = \frac{\partial}{\partial x_1}(x_1 x_2) = x_2$$

$$\frac{\partial f_2}{\partial x_2} = \frac{\partial}{\partial x_2}(x_1 x_2) = x_1$$

Assembling these into the Jacobian matrix gives:

$$J = \begin{pmatrix} 2x_1 & 2x_2 \\ x_2 & x_1 \end{pmatrix}$$

Question 26 (Moderate). *In machine learning, L2 regularization is often added to a loss function to prevent overfitting. The loss function for Ridge Regression is given by* $f(\mathbf{w}) = ||\mathbf{y} - X\mathbf{w}||_2^2 + \lambda||\mathbf{w}||_2^2$*, where* X *is the* $m \times n$ *design matrix,* $\mathbf{y}$ *is the* $m \times 1$ *target vector,* $\mathbf{w}$ *is the* $n \times 1$ *vector of weights, and* λ *is the regularization parameter. Find the gradient of this function with respect to* $\mathbf{w}$*.*

Solution. *Let* $f(\mathbf{w}) = ||\mathbf{y} - X\mathbf{w}||_2^2 + \lambda||\mathbf{w}||_2^2$*. We can rewrite the squared Euclidean norms using transposes:* $f(\mathbf{w}) = (\mathbf{y} - X\mathbf{w})^T(\mathbf{y} - X\mathbf{w}) + \lambda\mathbf{w}^T\mathbf{w}$*.*

Step 1: Expand the expression. *Using properties of transposes, the first term expands to:*

$$(\mathbf{y}^T - \mathbf{w}^T X^T)(\mathbf{y} - X\mathbf{w}) = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T X\mathbf{w} - \mathbf{w}^T X^T\mathbf{y} + \mathbf{w}^T X^T X\mathbf{w}$$

Since $\mathbf{y}^T X\mathbf{w}$ *is a scalar, it equals its transpose:* $(\mathbf{y}^T X\mathbf{w})^T = \mathbf{w}^T X^T\mathbf{y}$*. We can combine the middle terms.*

$$f(\mathbf{w}) = \mathbf{y}^T\mathbf{y} - 2\mathbf{w}^T X^T\mathbf{y} + \mathbf{w}^T(X^T X)\mathbf{w} + \lambda\mathbf{w}^T\mathbf{w}$$

Step 2: Differentiate with respect to $\mathbf{w}$*. We use two standard results from matrix calculus:* $\nabla_{\mathbf{w}}(\mathbf{a}^T\mathbf{w}) = \mathbf{a}$ *and* $\nabla_{\mathbf{w}}(\mathbf{w}^T A\mathbf{w}) = (A + A^T)\mathbf{w}$*.*

- $\nabla_{\mathbf{w}}(\mathbf{y}^T\mathbf{y}) = \mathbf{0}$ *(no dependence on* $\mathbf{w}$*).*

- $\nabla_{\mathbf{w}}(-2\mathbf{w}^T X^T \mathbf{y}) = -2(X^T \mathbf{y})$.
- $\nabla_{\mathbf{w}}(\mathbf{w}^T (X^T X)\mathbf{w})$. *The matrix* $A = X^T X$ *is symmetric, so the derivative is* $2(X^T X)\mathbf{w}$.
- $\nabla_{\mathbf{w}}(\lambda \mathbf{w}^T \mathbf{w}) = \nabla_{\mathbf{w}}(\lambda \mathbf{w}^T I \mathbf{w})$. *The matrix is* λI*, which is symmetric. So the derivative is* $2(\lambda I)\mathbf{w} = 2\lambda \mathbf{w}$.

Step 3: Combine the results.

$$\nabla_{\mathbf{w}} f(\mathbf{w}) = \mathbf{0} - 2X^T \mathbf{y} + 2X^T X \mathbf{w} + 2\lambda \mathbf{w}$$

$$\nabla_{\mathbf{w}} f(\mathbf{w}) = 2X^T X \mathbf{w} - 2X^T \mathbf{y} + 2\lambda \mathbf{w}$$

Factoring out common terms gives the final gradient:

$$\nabla_{\mathbf{w}} f(\mathbf{w}) = 2(X^T(X\mathbf{w} - \mathbf{y}) + \lambda \mathbf{w})$$

Question 27 (Tough)**.** *Let* $f(x, y, z) = x^2 e^y + y \ln(z)$. *Find the directional derivative of* f *at the point* $P(1, 0, 2)$ *in the direction of the vector* $\mathbf{v} = \langle 2, -1, 2\rangle$. *What does this value represent?*

Solution. *The directional derivative of* f *in the direction of a unit vector* $\mathbf{u}$*, denoted* $D_{\mathbf{u}} f$*, is calculated as the dot product of the gradient of* f *and the unit vector* $\mathbf{u}$.

$$D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u}$$

Step 1: Find the gradient of f. *The gradient is the vector of partial derivatives:*

$$\nabla f(x, y, z) = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle$$

$$\frac{\partial f}{\partial x} = 2xe^y$$

$$\frac{\partial f}{\partial y} = x^2 e^y + \ln(z)$$

$$\frac{\partial f}{\partial z} = \frac{y}{z}$$

So, $\nabla f = \langle 2xe^y, x^2 e^y + \ln(z), \frac{y}{z}\rangle$.

Step 2: Evaluate the gradient at the point P(1, 0, 2).

$$\nabla f(1, 0, 2) = \langle 2(1)e^0, (1)^2 e^0 + \ln(2), \frac{0}{2}\rangle = \langle 2, 1 + \ln(2), 0\rangle$$

Step 3: Find the unit vector u in the direction of v. *First find the magnitude of* $\mathbf{v}$*:*

$$||\mathbf{v}|| = \sqrt{2^2 + (-1)^2 + 2^2} = \sqrt{4 + 1 + 4} = \sqrt{9} = 3$$

The unit vector $\mathbf{u}$ *is* $\mathbf{v}$ *divided by its magnitude:*

$$\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||} = \frac{1}{3}\langle 2, -1, 2\rangle = \left\langle \frac{2}{3}, -\frac{1}{3}, \frac{2}{3} \right\rangle$$

Step 4: Calculate the dot product $\nabla f(P) \cdot \mathbf{u}$.

$$D_{\mathbf{u}}f(1,0,2) = \langle 2, 1+\ln(2), 0\rangle \cdot \left\langle \frac{2}{3}, -\frac{1}{3}, \frac{2}{3} \right\rangle$$

$$= (2)\left(\frac{2}{3}\right) + (1+\ln(2))\left(-\frac{1}{3}\right) + (0)\left(\frac{2}{3}\right)$$

$$= \frac{4}{3} - \frac{1}{3} - \frac{\ln(2)}{3} = \frac{3-\ln(2)}{3} = 1 - \frac{\ln(2)}{3}$$

***Interpretation**: The value $1 - \frac{\ln(2)}{3}$ represents the instantaneous rate of change of the function f at the point $P(1,0,2)$ as we move from P in the direction of the vector $\mathbf{v}$. Since $\ln(2) \approx 0.693$, the value is positive, which means the function's value is increasing at that point along that specific direction.*

5 Higher-Order Derivatives and Taylor's Series

Question 28 (Easy). *Find the first and second derivatives of the function $f(x) = e^{x^2}$.*

Solution. ***First derivative***: *We use the chain rule, $\frac{d}{dx}e^{u(x)} = e^{u(x)}u'(x)$. Let $u(x) = x^2$, so $u'(x) = 2x$.*

$$f'(x) = \frac{d}{dx}(e^{x^2}) = e^{x^2} \cdot \frac{d}{dx}(x^2) = 2xe^{x^2}$$

Second derivative: *We differentiate $f'(x)$ using the product rule, $(uv)' = u'v + uv'$. Let $u(x) = 2x$ and $v(x) = e^{x^2}$. Then $u'(x) = 2$ and $v'(x) = 2xe^{x^2}$ (from our first derivative calculation).*

$$f''(x) = \frac{d}{dx}(2xe^{x^2}) = (2)(e^{x^2}) + (2x)(2xe^{x^2})$$

$$f''(x) = 2e^{x^2} + 4x^2e^{x^2} = 2e^{x^2}(1 + 2x^2)$$

Question 29 (Easy). *For the function $f(x, y) = x^3 - 3xy^2 + y^4$, compute all second-order partial derivatives: $f_{xx}, f_{yy}, f_{xy}, f_{yx}$.*

Solution. *First, find the first-order partial derivatives:*

$$f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^3 - 3xy^2 + y^4) = 3x^2 - 3y^2$$

$$f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(x^3 - 3xy^2 + y^4) = -6xy + 4y^3$$

Now, find the second-order partial derivatives by differentiating the first-order derivatives:

$$f_{xx} = \frac{\partial}{\partial x}(f_x) = \frac{\partial}{\partial x}(3x^2 - 3y^2) = 6x$$

$$f_{yy} = \frac{\partial}{\partial y}(f_y) = \frac{\partial}{\partial y}(-6xy + 4y^3) = -6x + 12y^2$$

$$f_{xy} = \frac{\partial}{\partial y}(f_x) = \frac{\partial}{\partial y}(3x^2 - 3y^2) = -6y$$

$$f_{yx} = \frac{\partial}{\partial x}(f_y) = \frac{\partial}{\partial x}(-6xy + 4y^3) = -6y$$

As expected from Clairaut's theorem on the equality of mixed partials (since these derivatives are continuous), we find that $f_{xy} = f_{yx}$.

Question 30 (Moderate). *Find the Hessian matrix of the function $f(x, y) = x\sin(y) + y\cos(x)$ at the point $(\frac{\pi}{2}, 0)$.*

Solution. *The Hessian matrix H is the matrix of second-order partial derivatives:*

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

First, find the first-order partial derivatives:

$$f_x = \frac{\partial}{\partial x}(x\sin(y) + y\cos(x)) = \sin(y) - y\sin(x)$$

$$f_y = \frac{\partial}{\partial y}(x\sin(y) + y\cos(x)) = x\cos(y) + \cos(x)$$

Now, find the second-order partial derivatives:

$$f_{xx} = \frac{\partial}{\partial x}(\sin(y) - y\sin(x)) = -y\cos(x)$$

$$f_{yy} = \frac{\partial}{\partial y}(x\cos(y) + \cos(x)) = -x\sin(y)$$

$$f_{xy} = \frac{\partial}{\partial y}(\sin(y) - y\sin(x)) = \cos(y) - \sin(x)$$

The Hessian matrix as a function of (x, y) *is:*

$$H(x,y) = \begin{pmatrix} -y\cos(x) & \cos(y) - \sin(x) \\ \cos(y) - \sin(x) & -x\sin(y) \end{pmatrix}$$

Now evaluate at the point $(\frac{\pi}{2}, 0)$*. We use the values:* $\cos(\frac{\pi}{2}) = 0, \quad \sin(\frac{\pi}{2}) = 1, \quad \cos(0) = 1, \quad \sin(0) = 0.$

$$H\left(\frac{\pi}{2}, 0\right) = \begin{pmatrix} -(0)\cos(\frac{\pi}{2}) & \cos(0) - \sin(\frac{\pi}{2}) \\ \cos(0) - \sin(\frac{\pi}{2}) & -\frac{\pi}{2}\sin(0) \end{pmatrix}$$

$$H\left(\frac{\pi}{2}, 0\right) = \begin{pmatrix} 0 & 1-1 \\ 1-1 & -\frac{\pi}{2}\cdot 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Question 31 (Moderate)**.** *Find the linearization (first-order Taylor expansion) of the function* $f(x, y) = \sqrt{x^2 + y^2}$ *around the point* $(3, 4)$*. Use it to approximate* $f(3.01, 3.98)$*.*

Solution. *The linearization* $L(x, y)$ *of a function* $f(x, y)$ *at a point* (a, b) *is the tangent plane to the function at that point. The formula is:*

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

Here, our function is $f(x, y) = (x^2 + y^2)^{1/2}$ *and the point is* $(a, b) = (3, 4)$*.* ***Step 1: Evaluate the function at the point.***

$$f(3, 4) = \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5$$

Step 2: Find the partial derivatives.

$$f_x = \frac{\partial}{\partial x}(x^2 + y^2)^{1/2} = \frac{1}{2}(x^2 + y^2)^{-1/2}(2x) = \frac{x}{\sqrt{x^2 + y^2}}$$

$$f_y = \frac{\partial}{\partial y}(x^2 + y^2)^{1/2} = \frac{1}{2}(x^2 + y^2)^{-1/2}(2y) = \frac{y}{\sqrt{x^2 + y^2}}$$

Step 3: Evaluate the partial derivatives at $(3, 4)$***.***

$$f_x(3, 4) = \frac{3}{\sqrt{3^2 + 4^2}} = \frac{3}{5}$$

$$f_y(3, 4) = \frac{4}{\sqrt{3^2 + 4^2}} = \frac{4}{5}$$

Step 4: Construct the linearization formula.

$$L(x, y) = 5 + \frac{3}{5}(x - 3) + \frac{4}{5}(y - 4)$$

This is the first-order Taylor expansion. To approximate $f(3.01, 3.98)$, we evaluate $L(3.01, 3.98)$. Here, $x - a = 3.01 - 3 = 0.01$ and $y - b = 3.98 - 4 = -0.02$.

$$f(3.01, 3.98) \approx L(3.01, 3.98) = 5 + \frac{3}{5}(0.01) + \frac{4}{5}(-0.02)$$

$$= 5 + \frac{0.03}{5} - \frac{0.08}{5} = 5 - \frac{0.05}{5} = 5 - 0.01 = 4.99$$

Question 32 (Tough)**.** *Find the second-order Taylor series expansion of the function $f(x, y) = e^{x-y}$ around the point $(a, b) = (1, 1)$.*

Solution. *The second-order Taylor expansion of $f(x, y)$ around (a, b) is given by:*

$$\begin{aligned} T_2(x, y) =& f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) \\ &+ \frac{1}{2!}\left(f_{xx}(a, b)(x - a)^2 + 2f_{xy}(a, b)(x - a)(y - b) + f_{yy}(a, b)(y - b)^2\right) \end{aligned}$$

Here, $(a, b) = (1, 1)$, so $x - a = x - 1$ and $y - b = y - 1$. ***Step 1: Evaluate the function and all required derivatives at (1, 1).***

- $f(x, y) = e^{x-y} \implies f(1, 1) = e^{1-1} = e^0 = 1.$
- $f_x = e^{x-y} \cdot (1) = e^{x-y} \implies f_x(1, 1) = e^0 = 1.$
- $f_y = e^{x-y} \cdot (-1) = -e^{x-y} \implies f_y(1, 1) = -e^0 = -1.$
- $f_{xx} = \frac{\partial}{\partial x}(e^{x-y}) = e^{x-y} \implies f_{xx}(1, 1) = 1.$
- $f_{yy} = \frac{\partial}{\partial y}(-e^{x-y}) = -e^{x-y} \cdot (-1) = e^{x-y} \implies f_{yy}(1, 1) = 1.$
- $f_{xy} = \frac{\partial}{\partial y}(e^{x-y}) = -e^{x-y} \implies f_{xy}(1, 1) = -1.$

Step 2: Substitute these values into the Taylor formula.

$$\begin{aligned} T_2(x, y) =& 1 + (1)(x - 1) + (-1)(y - 1) \\ &+ \frac{1}{2}\left((1)(x - 1)^2 + 2(-1)(x - 1)(y - 1) + (1)(y - 1)^2\right) \end{aligned}$$

$$T_2(x, y) = 1 + (x - 1) - (y - 1) + \frac{1}{2}\left((x - 1)^2 - 2(x - 1)(y - 1) + (y - 1)^2\right)$$

This expression is the final answer. It can be simplified by noticing the quadratic part is a perfect square:

$$T_2(x, y) = 1 + (x - y) + \frac{1}{2}\left([(x - 1) - (y - 1)]^2\right) = 1 + (x - y) + \frac{1}{2}(x - y)^2$$

This matches the standard Maclaurin series for $e^u \approx 1 + u + u^2/2$ where $u = x - y$.

Question 33 (Tough)**.** *Consider the function $f(x, y) = \cos(x) - y^2 - xy$.*

(a) Find the second-order Taylor expansion of $f(x, y)$ around the point $(0, 0)$.

(b) Use the Hessian matrix from the Taylor expansion to classify the critical point at $(0, 0)$.

Solution. *(a)* ***Second-order Taylor expansion around (0,0)****: The formula is:* $T_2(x, y) = f(0, 0) + f_x(0, 0)x + f_y(0, 0)y + \frac{1}{2}(f_{xx}(0, 0)x^2 + 2f_{xy}(0, 0)xy + f_{yy}(0, 0)y^2)$.

Step 1: Evaluate all derivatives at (0,0)*:*

- $f(x, y) = \cos(x) - y^2 - xy \implies f(0, 0) = \cos(0) - 0 - 0 = 1.$
- $f_x = -\sin(x) - y \implies f_x(0, 0) = -\sin(0) - 0 = 0.$
- $f_y = -2y - x \implies f_y(0, 0) = 0 - 0 = 0.$
- $f_{xx} = -\cos(x) \implies f_{xx}(0, 0) = -\cos(0) = -1.$
- $f_{yy} = -2 \implies f_{yy}(0, 0) = -2.$
- $f_{xy} = -1 \implies f_{xy}(0, 0) = -1.$

Since $f_x(0, 0) = 0$ and $f_y(0, 0) = 0$, the point $(0, 0)$ is indeed a critical point.

Step 2: Assemble the Taylor expansion*:*

$$T_2(x, y) = 1 + (0)x + (0)y + \frac{1}{2}((-1)x^2 + 2(-1)xy + (-2)y^2)$$

$$T_2(x, y) = 1 - \frac{1}{2}x^2 - xy - y^2$$

(b) ***Classify the critical point using the Hessian****: The Hessian matrix at a point (x, y) is $H(x, y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$. At the point $(0, 0)$, we have:*

$$H(0, 0) = \begin{pmatrix} f_{xx}(0, 0) & f_{xy}(0, 0) \\ f_{yx}(0, 0) & f_{yy}(0, 0) \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ -1 & -2 \end{pmatrix}$$

To classify the critical point, we use the Second Derivative Test. We compute the determinant of the Hessian, D.

$$D = \det(H) = (-1)(-2) - (-1)(-1) = 2 - 1 = 1$$

The test is as follows:

- *If $D > 0$ and $f_{xx} > 0$, it is a local minimum.*
- *If $D > 0$ and $f_{xx} < 0$, it is a local maximum.*
- *If $D < 0$, it is a saddle point.*
- *If $D = 0$, the test is inconclusive.*

In our case, $D = 1 > 0$ and $f_{xx}(0, 0) = -1 < 0$. Therefore, the critical point at $(0, 0)$ is a ***local maximum****.*

6 Unconstrained Optimization

Question 34 (Easy)**.** *Find the critical points of the function $f(x) = x^3 - 6x^2 + 9x + 2$.*

Solution. *To find the critical points of a single-variable function, we find the points where the first derivative is zero or undefined. First, find the derivative of $f(x)$:*

$$f'(x) = \frac{d}{dx}(x^3 - 6x^2 + 9x + 2) = 3x^2 - 12x + 9$$

The derivative is a polynomial, so it is defined for all real numbers x. We find the critical points by setting the derivative to zero:

$$3x^2 - 12x + 9 = 0$$

Divide the entire equation by 3 to simplify:

$$x^2 - 4x + 3 = 0$$

Factor the quadratic equation:

$$(x - 1)(x - 3) = 0$$

This gives two solutions. The critical points are $\boldsymbol{x = 1}$ and $\boldsymbol{x = 3}$.

Question 35 (Easy)**.** *Find the critical points of the function $f(x, y) = x^2 + y^2 - 2x - 6y + 14$.*

Solution. *To find critical points of a multivariable function, we set its gradient to the zero vector, $\nabla f = \mathbf{0}$. The gradient is the vector of partial derivatives. First, find the partial derivatives:*

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^2 + y^2 - 2x - 6y + 14) = 2x - 2$$

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(x^2 + y^2 - 2x - 6y + 14) = 2y - 6$$

Now, set both partial derivatives to zero to find the coordinates of the critical point(s):

$$2x - 2 = 0 \implies 2x = 2 \implies x = 1$$

$$2y - 6 = 0 \implies 2y = 6 \implies y = 3$$

The system of equations has only one solution. The only critical point is $\mathbf{(1, 3)}$. (This function describes a paraboloid, which has a single minimum).

Question 36 (Moderate)**.** *Find all critical points of the function $f(x, y) = x^3 + y^3 - 3xy$ and classify each as a local maximum, local minimum, or saddle point.*

Solution. ***Step 1: Find critical points**. Set the gradient $\nabla f = \langle f_x, f_y \rangle$ to $\mathbf{0}$.*

$$f_x = 3x^2 - 3y = 0 \implies y = x^2 \quad (1)$$

$$f_y = 3y^2 - 3x = 0 \implies x = y^2 \quad (2)$$

Substitute the first equation into the second: $x = (x^2)^2 \implies x = x^4 \implies x^4 - x = 0 \implies x(x^3 - 1) = 0$. This gives two possible values for x:

- $x = 0$. *Substitute back into eq(1):* $y = 0^2 = 0$. *This gives the critical point* $\mathbf{(0,0)}$.
- $x^3 - 1 = 0 \implies x = 1$. *Substitute back into eq(1):* $y = 1^2 = 1$. *This gives the critical point* $\mathbf{(1,1)}$.

Step 2: Classify points using the Second Derivative Test. *We need the Hessian matrix,* $H(x,y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$.

$$f_{xx} = 6x, \quad f_{yy} = 6y, \quad f_{xy} = -3$$

The determinant of the Hessian is $D(x,y) = f_{xx}f_{yy} - (f_{xy})^2 = (6x)(6y) - (-3)^2 = 36xy - 9$.

Case 1: Point (0,0) $D(0,0) = 36(0)(0) - 9 = -9$. *Since* $D < 0$*, the point* $(0,0)$ *is a* ***saddle point***.

Case 2: Point (1,1) $D(1,1) = 36(1)(1) - 9 = 27$. *Since* $D > 0$*, we check the sign of* $f_{xx}(1,1)$. $f_{xx}(1,1) = 6(1) = 6$. *Since* $D > 0$ *and* $f_{xx} > 0$*, the point* $(1,1)$ *is a* ***local minimum***.

Question 37 (Moderate)**.** *A company wants to build a rectangular storage container with an open top. It must have a volume of 10 cubic meters. The material for the base costs \$10 per square meter, and the material for the sides costs \$6 per square meter. Find the dimensions (length l, width w, height h) that will minimize the cost of the container.*

Solution. ***Step 1: Set up the objective and constraint equations.*** *Let the dimensions be length* l*, width* w*, and height* h*. The volume is the constraint equation:* $V = lwh = 10$*. The cost is the objective function we want to minimize. Cost of base = (Area of base)* $\times$ *(Cost per area) =* $(lw) \cdot 10$*. Cost of sides = (Area of sides)* $\times$ *(Cost per area) =* $(2lh + 2wh) \cdot 6$*. Total cost:* $C(l,w,h) = 10lw + 12lh + 12wh$.

Step 2: Reduce to a two-variable function. *From the volume constraint, we can express one variable in terms of the others:* $h = \frac{10}{lw}$*. Substitute this into the cost function:*

$$C(l,w) = 10lw + 12l\left(\frac{10}{lw}\right) + 12w\left(\frac{10}{lw}\right) = 10lw + \frac{120}{w} + \frac{120}{l}$$

Step 3: Find critical points. *Set the partial derivatives of* $C(l,w)$ *with respect to* l *and* w *to zero.*

$$C_l = \frac{\partial C}{\partial l} = 10w - \frac{120}{l^2} = 0 \implies 10wl^2 = 120 \implies w = \frac{12}{l^2} \quad (1)$$

$$C_w = \frac{\partial C}{\partial w} = 10l - \frac{120}{w^2} = 0 \implies 10lw^2 = 120 \implies l = \frac{12}{w^2} \quad (2)$$

Substitute the expression for w *from (1) into (2):*

$$l = \frac{12}{(12/l^2)^2} = \frac{12}{144/l^4} = \frac{12l^4}{144} = \frac{l^4}{12}$$

Since we are building a container, $l > 0$*, so we can divide by* l*:* $1 = \frac{l^3}{12} \implies l^3 = 12 \implies l = \sqrt[3]{12}$*. Now find* w *using equation (1):*

$$w = \frac{12}{l^2} = \frac{12}{(\sqrt[3]{12})^2} = \frac{12}{12^{2/3}} = 12^{1-2/3} = 12^{1/3} = \sqrt[3]{12}$$

So the optimal base is a square, $l = w$. Finally, find the height h:

$$h = \frac{10}{lw} = \frac{10}{\sqrt[3]{12} \cdot \sqrt[3]{12}} = \frac{10}{12^{2/3}}$$

The dimensions that minimize cost are $\boldsymbol{l = w = \sqrt[3]{12} \approx 2.29}$ ***m*** *and* $\boldsymbol{h = \frac{10}{12^{2/3}} \approx 1.91}$ ***m****. (A check with the second derivative test would confirm this is a minimum).*

Question 38 (Tough)**.** *Consider the function $f(x, y) = (x^2 + y^2)^2 - k(x^2 - y^2)$, where k is a constant parameter.*

(a) Show that the origin $(0, 0)$ is a critical point for any value of k.

(b) Find the Hessian matrix and use the second derivative test to determine how the classification of the critical point at $(0, 0)$ (local min, max, or saddle) depends on the value of k.

Solution. *(a)* ***Show (0,0) is a critical point****: We find the first partial derivatives of $f(x, y) = x^4 + 2x^2y^2 + y^4 - kx^2 + ky^2$. $f_x = 4x^3 + 4xy^2 - 2kx$. $f_y = 4x^2y + 4y^3 + 2ky$. Now evaluate them at the origin $(0, 0)$: $f_x(0, 0) = 4(0)^3 + 4(0)(0)^2 - 2k(0) = 0$. $f_y(0, 0) = 4(0)^2(0) + 4(0)^3 + 2k(0) = 0$. Since $\nabla f(0, 0) = \langle 0, 0\rangle$ for any value of k, the origin is always a critical point.*

(b) ***Classify the critical point based on k****: We find the second partial derivatives to construct the Hessian matrix. $f_{xx} = 12x^2+4y^2-2k$. $f_{yy} = 4x^2+12y^2+2k$. $f_{xy} = 8xy$. Now evaluate the Hessian at the origin $(0, 0)$: $H(0, 0) = \begin{pmatrix} f_{xx}(0,0) & f_{xy}(0,0) \\ f_{yx}(0,0) & f_{yy}(0,0) \end{pmatrix} = \begin{pmatrix} -2k & 0 \\ 0 & 2k \end{pmatrix}$. The determinant of the Hessian at the origin is $D = (-2k)(2k)-(0)^2 = -4k^2$.*

We analyze the classification based on the values of D and f_{xx}:

- ***Case 1:*** *$k \neq 0$. In this case, $k^2 > 0$, so the determinant $D = -4k^2$ will always be negative. If $D < 0$, the critical point is a* ***saddle point****. This is true for any non-zero value of k.*
- ***Case 2:*** *$k = 0$. If $k = 0$, the determinant $D = -4(0)^2 = 0$. When $D = 0$, the second derivative test is inconclusive. We must investigate the function itself. If $k = 0$, the function becomes $f(x, y) = (x^2 + y^2)^2$. At the critical point $(0, 0)$, $f(0, 0) = 0$. For any other point $(x, y) \neq (0, 0)$, the term $x^2 + y^2$ is a positive number, and so is $(x^2 + y^2)^2$. Thus, $f(x, y) > f(0, 0)$ for all points in the neighborhood of the origin. This means that for $k = 0$, the origin is a* ***local minimum*** *(in fact, it is a global minimum).*

Summary*:*

- *If $\boldsymbol{k \neq 0}$, the origin is a* ***saddle point****.*
- *If $\boldsymbol{k = 0}$, the origin is a* ***local minimum****.*