Table of Contents

1 Vector Spaces and Subspaces

1.1 Definitions

A **vector space** is a set of vectors, along with two operations—vector addition and scalar multiplication—that satisfies the following properties:

- **Closure under addition:** If $\mathbf{u}$ and $\mathbf{v}$ are vectors in the vector space, then their sum $\mathbf{u} + \mathbf{v}$ is also in the vector space.
- **Closure under scalar multiplication:** If $\mathbf{u}$ is a vector in the vector space and c is a scalar, then $c\mathbf{u}$ is also in the vector space.
- **Other properties:** Vector spaces also satisfy properties like associativity, commutativity of addition, the existence of a zero vector, and additive inverses.

A **subspace** is a subset of a vector space that is itself a vector space. For a subset to be a subspace, it must satisfy three conditions:

1. **The zero vector is in the subset.**
2. **Closure under addition:** If $\mathbf{u}$ and $\mathbf{v}$ are in the subset, then $\mathbf{u} + \mathbf{v}$ is also in the subset.
3. **Closure under scalar multiplication:** If $\mathbf{u}$ is in the subset and c is a scalar, then $c\mathbf{u}$ is also in the subset.

1.2 Example: Verifying a Subspace

Problem: Let $V = \mathbb{R}^3$ and consider the set:

$$W = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0 \right\}.$$

Verify that W is a subspace of $\mathbb{R}^3$.

Step 1: Check the Zero Vector The zero vector in $\mathbb{R}^3$ is:

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Substituting into the condition $x_1 + x_2 + x_3 = 0$:

$$0 + 0 + 0 = 0.$$

Thus, $\mathbf{0} \in W$.

Step 2: Check Closure Under Addition Take two vectors $\mathbf{u}, \mathbf{v} \in W$:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}.$$

By definition of W, we know:

$$u_1 + u_2 + u_3 = 0, \quad v_1 + v_2 + v_3 = 0.$$

Adding $\mathbf{u}$ and $\mathbf{v}$:

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix}.$$

Checking the condition:

$$(u_1 + v_1) + (u_2 + v_2) + (u_3 + v_3) = (u_1 + u_2 + u_3) + (v_1 + v_2 + v_3) = 0 + 0 = 0.$$

Thus, $\mathbf{u} + \mathbf{v} \in W$.

Step 3: Check Closure Under Scalar Multiplication Take $\mathbf{u} \in W$ and a scalar $c \in \mathbb{R}$. Then:

$$c\mathbf{u} = c \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix}.$$

Checking the condition:

$$c(u_1 + u_2 + u_3) = c \cdot 0 = 0.$$

Thus, $c\mathbf{u} \in W$.

Conclusion: Since W satisfies all three conditions, it is a subspace of $\mathbb{R}^3$.

1.3 Numerical Example

Let $W = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0 \right\}$. Verify whether $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -3 \\ 3 \\ 0 \end{bmatrix}$ belong to W, and check if their sum is in W.

Step 1: Verify $\mathbf{u} \in W$ Check if $u_1 + u_2 + u_3 = 0$:

$$1 + (-2) + 1 = 0.$$

Thus, $\mathbf{u} \in W$.

Step 2: Verify $\mathbf{v} \in W$ Check if $v_1 + v_2 + v_3 = 0$:

$$-3 + 3 + 0 = 0.$$

Thus, $\mathbf{v} \in W$.

Step 3: Verify $\mathbf{u} + \mathbf{v} \in W$ Compute $\mathbf{u} + \mathbf{v}$:

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \begin{bmatrix} -3 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

Check if the sum satisfies the condition:

$$-2 + 1 + 1 = 0.$$

Thus, $\mathbf{u} + \mathbf{v} \in W$.

Conclusion: Both vectors $\mathbf{u}$ and $\mathbf{v}$ belong to W, and their sum is also in W, verifying closure under addition.

Linear Independence and Basis

Definitions

- **Linear Independence:** A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space is linearly independent if the only solution to the equation:

  $$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

  is $c_1 = c_2 = \cdots = c_n = 0$.

  - Here, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are vectors, and $c_1, c_2, \ldots, c_n$ are scalars.
  - If there exist non-zero scalars $c_1, c_2, \ldots, c_n$ such that the equation holds, the vectors are **linearly dependent**.

- **Basis:** A basis of a vector space is a set of linearly independent vectors that span the entire vector space.

  - **Span:** The span of a set of vectors is the set of all possible linear combinations of those vectors.
  - A basis provides the minimum number of vectors needed to represent every vector in the space.

Worked Example: Checking Linear Independence

Problem: Consider the vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ in $\mathbb{R}^3$. Determine if the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

Step 1: Set Up the Linear Independence Equation To check linear independence, solve the equation:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0},$$

where $\mathbf{0}$ is the zero vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Substituting the given vectors:

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Step 2: Combine the Vectors Adding the scaled vectors:

$$\begin{bmatrix} c_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c_2 \\ 0 \end{bmatrix} + \begin{bmatrix} c_3 \\ c_3 \\ 0 \end{bmatrix} = \begin{bmatrix} c_1 + c_3 \\ c_2 + c_3 \\ 0 \end{bmatrix}.$$

Equating components to the zero vector:

$$\begin{aligned} c_1 + c_3 &= 0, \\ c_2 + c_3 &= 0, \\ 0 &= 0. \end{aligned}$$

Step 3: Solve for the Scalars From the first equation:

$$c_3 = -c_1.$$

From the second equation:

$$c_3 = -c_2.$$

Equating $-c_1$ and $-c_2$:

$$c_1 = c_2.$$

Substituting $c_1 = c_2$ into $c_3 = -c_1$:

$$c_3 = -c_1.$$

The only solution is $c_1 = c_2 = c_3 = 0$.

Step 4: Conclusion Since the only solution is the trivial solution $c_1 = c_2 = c_3 = 0$, the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent.

Worked Example: Basis of a Subspace

Problem: Find a basis for the subspace $W \subseteq \mathbb{R}^3$ defined by:

$$W = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0 \right\}.$$

Step 1: Write the Condition as a Constraint The subspace is defined by the equation:

$$x_1 + x_2 + x_3 = 0.$$

Express x_3 in terms of x_1 and x_2:

$$x_3 = -(x_1 + x_2).$$

Thus, any vector in W can be written as:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ -(x_1 + x_2) \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Step 2: Identify the Basis The vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

are linearly independent and span W. Thus, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for W.

Step 3: Verify Linear Independence To verify, solve:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}.$$

Substituting $\mathbf{v}_1$ and $\mathbf{v}_2$:

$$c_1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Equating components:

$$\begin{aligned} c_1 &= 0, \\ c_2 &= 0, \\ -c_1 - c_2 &= 0. \end{aligned}$$

The only solution is $c_1 = c_2 = 0$, so the vectors are linearly independent.

Step 4: Conclusion The basis for W is $\{\mathbf{v}_1, \mathbf{v}_2\}$.

Matrix Operations

Definitions and Properties

- A **matrix** is a rectangular array of numbers arranged in rows and columns. A matrix with m rows and n columns is called an $m \times n$ matrix.
- Matrices can be added, subtracted, and multiplied if they conform to specific rules for their dimensions.
- The **identity matrix**, $\mathbf{I}$, is a square matrix with 1s on the diagonal and 0s elsewhere.
- The **transpose** of a matrix $\mathbf{A}$, denoted $\mathbf{A}^T$, swaps its rows and columns.

Matrix Multiplication

Matrix multiplication combines two matrices $\mathbf{A}$ and $\mathbf{B}$, provided that the number of columns of $\mathbf{A}$ matches the number of rows of $\mathbf{B}$.

For $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$, the product $\mathbf{C} = \mathbf{AB}$ is a matrix in $\mathbb{R}^{m\times p}$, where each entry c_{ij} is given by:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Worked Example: Matrix Multiplication

Problem: Compute the product of the following matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}.$$

Step 1: Verify Dimensions $\mathbf{A}$ is a 2×2 matrix, and $\mathbf{B}$ is also a 2×2 matrix. Since the number of columns in $\mathbf{A}$ matches the number of rows in $\mathbf{B}$, the product $\mathbf{C} = \mathbf{AB}$ is well-defined and will be a 2×2 matrix.

Step 2: Compute Each Entry of C Using the formula $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$:

$$\begin{aligned}
c_{11} &= a_{11}b_{11} + a_{12}b_{21} = (1)(2) + (2)(1) = 2 + 2 = 4, \\
c_{12} &= a_{11}b_{12} + a_{12}b_{22} = (1)(0) + (2)(3) = 0 + 6 = 6, \\
c_{21} &= a_{21}b_{11} + a_{22}b_{21} = (3)(2) + (4)(1) = 6 + 4 = 10, \\
c_{22} &= a_{21}b_{12} + a_{22}b_{22} = (3)(0) + (4)(3) = 0 + 12 = 12.
\end{aligned}$$

Step 3: Write the Result The resulting matrix $\mathbf{C}$ is:

$$\mathbf{C} = \begin{bmatrix} 4 & 6 \\ 10 & 12 \end{bmatrix}.$$

Determinants and Their Properties

Definition: The **determinant** of a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, denoted $\det(\mathbf{A})$, is a scalar that provides important information about the matrix, such as whether it is invertible.

Properties:

- $\det(\mathbf{A}) = 0$ implies that $\mathbf{A}$ is singular (not invertible).
- The determinant changes sign if two rows (or columns) of $\mathbf{A}$ are swapped.
- $\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$.

Worked Example: Computing a Determinant

Problem: Compute the determinant of the matrix:

$$\mathbf{A} = \begin{bmatrix} 4 & 3 \\ 6 & 3 \end{bmatrix}.$$

Step 1: Use the Formula for a 2×2 Matrix For $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant is:

$$\det(\mathbf{A}) = ad - bc.$$

Substitute the values:

$$\det(\mathbf{A}) = (4)(3) - (3)(6) = 12 - 18 = -6.$$

Step 2: Conclusion The determinant of $\mathbf{A}$ is -6, so $\mathbf{A}$ is invertible.

Inverse of a Matrix

Definition: The **inverse** of a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, denoted $\mathbf{A}^{-1}$, satisfies:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I},$$

where $\mathbf{I}$ is the identity matrix.

Formula for a 2×2 Matrix: If $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\det(\mathbf{A}) \neq 0$, then:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Worked Example: Computing an Inverse

Problem: Compute the inverse of the matrix:

$$\mathbf{A} = \begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix}.$$

Step 1: Compute the Determinant

$$\det(\mathbf{A}) = (4)(6) - (7)(2) = 24 - 14 = 10.$$

Step 2: Apply the Formula Substitute into the formula for the inverse:

$$\mathbf{A}^{-1} = \frac{1}{10} \begin{bmatrix} 6 & -7 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 0.6 & -0.7 \\ -0.2 & 0.4 \end{bmatrix}.$$

Step 3: Verify the Result (Optional) Multiply $\mathbf{A}$ and $\mathbf{A}^{-1}$ to confirm that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

Orthogonality and Projections

Definitions

- **Orthogonality:** Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are orthogonal if their dot product is zero:

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

- **Orthonormality:** A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is orthonormal if:
  - Each vector is a unit vector: $\|\mathbf{v}_i\| = 1$.
  - The vectors are mutually orthogonal: $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$.
- **Orthogonal Projection:** The projection of a vector $\mathbf{b}$ onto a vector $\mathbf{a}$ is given by:

$$\text{proj}_{\mathbf{a}}\mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a}.$$

Worked Example: Orthogonal Projection

Problem: Compute the projection of $\mathbf{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ onto $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Step 1: Compute the Dot Products

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= (1)(3) + (2)(4) = 3 + 8 = 11, \\ \mathbf{a} \cdot \mathbf{a} &= (1)(1) + (2)(2) = 1 + 4 = 5. \end{aligned}$$

Step 2: Compute the Projection Substitute into the formula:

$$\text{proj}_{\mathbf{a}}\mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a} = \frac{11}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{11}{5} \\ \frac{22}{5} \end{bmatrix}.$$

Step 3: Write the Result The projection is:

$$\text{proj}_{\mathbf{a}}\mathbf{b} = \begin{bmatrix} 2.2 \\ 4.4 \end{bmatrix}.$$

The Gram-Schmidt Process

Definition: The Gram-Schmidt process takes a set of linearly independent vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and produces an orthonormal set of vectors $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$.

Algorithm:

1. Start with $\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$.
2. For each subsequent vector $\mathbf{v}_k$, subtract its projections onto all previous $\mathbf{q}_i$:

$$\mathbf{u}_k = \mathbf{v}_k - \sum_{i=1}^{k-1} \text{proj}_{\mathbf{q}_i}\mathbf{v}_k.$$

3. Normalize $\mathbf{u}_k$ to obtain $\mathbf{q}_k$:

$$\mathbf{q}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.$$

Worked Example: Gram-Schmidt Process

Problem: Apply the Gram-Schmidt process to the set $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Step 1: Normalize $\mathbf{v}_1$

$$\|\mathbf{v}_1\| = \sqrt{(1)^2 + (1)^2} = \sqrt{2},$$
$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Step 2: Compute $\mathbf{u}_2$ Subtract the projection of $\mathbf{v}_2$ onto $\mathbf{q}_1$:

$$\text{proj}_{\mathbf{q}_1} \mathbf{v}_2 = \frac{\mathbf{q}_1 \cdot \mathbf{v}_2}{\mathbf{q}_1 \cdot \mathbf{q}_1} \mathbf{q}_1,$$
$$\mathbf{q}_1 \cdot \mathbf{v}_2 = \frac{1}{\sqrt{2}}(1) + \frac{1}{\sqrt{2}}(-1) = 0,$$
$$\text{proj}_{\mathbf{q}_1} \mathbf{v}_2 = 0 \cdot \mathbf{q}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$
$$\mathbf{u}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{q}_1} \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Step 3: Normalize $\mathbf{u}_2$

$$\|\mathbf{u}_2\| = \sqrt{(1)^2 + (-1)^2} = \sqrt{2},$$
$$\mathbf{q}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Step 4: Conclusion The orthonormal set is:

$$\mathbf{q}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Linear Transformations

Definitions

- A **linear transformation** is a function $T : \mathbb{R}^n \to \mathbb{R}^m$ that satisfies two properties for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and scalars c:
  1. **Additivity:** $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$.
  2. **Homogeneity:** $T(c\mathbf{u}) = cT(\mathbf{u})$.
- A linear transformation can be represented as a matrix multiplication: $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where $\mathbf{A} \in \mathbb{R}^{m\times n}$ is the matrix associated with the transformation.

Standard Matrix of a Linear Transformation

The standard matrix $\mathbf{A}$ of a linear transformation T is constructed by applying T to each column of the identity matrix $\mathbf{I}_n$:

$$\mathbf{A} = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \dots & T(\mathbf{e}_n) \end{bmatrix},$$

where $\mathbf{e}_i$ are the standard basis vectors in $\mathbb{R}^n$.

Worked Example: Constructing a Standard Matrix

Problem: Find the matrix $\mathbf{A}$ representing the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$, where:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + 3x_2 \\ -x_1 + 4x_2 \end{bmatrix}.$$

Step 1: Apply T to the Standard Basis Vectors The standard basis vectors in $\mathbb{R}^2$ are:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Compute $T(\mathbf{e}_1)$:

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2(1) + 3(0) \\ -1(1) + 4(0) \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Compute $T(\mathbf{e}_2)$:

$$T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2(0) + 3(1) \\ -1(0) + 4(1) \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Step 2: Form the Matrix A The columns of $\mathbf{A}$ are $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$:

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

Worked Example: Applying a Linear Transformation

Problem: Use the matrix $\mathbf{A} = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}$ to compute $T\left(\begin{bmatrix} 5 \\ 2 \end{bmatrix}\right)$.

Step 1: Perform the Matrix Multiplication Substitute $\mathbf{x} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$:

$$T\left(\begin{bmatrix} 5 \\ 2 \end{bmatrix}\right) = \mathbf{A}\mathbf{x} = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

Compute each entry:

$$\begin{aligned} \text{First row:} \quad & (2)(5) + (3)(2) = 10 + 6 = 16, \\ \text{Second row:} \quad & (-1)(5) + (4)(2) = -5 + 8 = 3. \end{aligned}$$

Step 2: Write the Result

$$T\left(\begin{bmatrix}5\\2\end{bmatrix}\right) = \begin{bmatrix}16\\3\end{bmatrix}.$$

Properties of Linear Transformations

- The transformation T is **invertible** if $\mathbf{A}$ is a square matrix and $\det(\mathbf{A}) \neq 0$.
- The image of T, also called the **column space** of $\mathbf{A}$, is the span of the columns of $\mathbf{A}$.
- The **null space** of $\mathbf{A}$ is the set of all vectors $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Worked Example: Finding the Null Space

Problem: Find the null space of $\mathbf{A} = \begin{bmatrix}2 & 3\\-1 & 4\end{bmatrix}$.

Step 1: Solve $\mathbf{A}\mathbf{x} = \mathbf{0}$ Substitute $\mathbf{x} = \begin{bmatrix}x_1\\x_2\end{bmatrix}$:

$$\begin{bmatrix}2 & 3\\-1 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}.$$

This gives the system of equations:

$$2x_1 + 3x_2 = 0,$$
$$-x_1 + 4x_2 = 0.$$

Step 2: Solve the System From the second equation:

$$x_1 = 4x_2.$$

Substitute into the first equation:

$$2(4x_2) + 3x_2 = 8x_2 + 3x_2 = 11x_2 = 0.$$

Thus, $x_2 = 0$ and $x_1 = 0$.

Step 3: Conclusion The null space of $\mathbf{A}$ is $\{\mathbf{0}\}$, meaning $\mathbf{A}$

Eigenvalues and Eigenvectors

Definitions

- An **eigenvector** of a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a non-zero vector $\mathbf{v}$ such that:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v},$$

  where λ is a scalar known as the **eigenvalue** associated with $\mathbf{v}$.

- Geometrically, $\mathbf{v}$ is a vector that does not change direction under the transformation $\mathbf{A}$, and λ describes the scaling factor.

Finding Eigenvalues and Eigenvectors

To find the eigenvalues and eigenvectors of a matrix $\mathbf{A}$:

1. Solve the **characteristic equation**: $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.
2. For each eigenvalue λ, solve $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$ to find the eigenvectors.

Worked Example: Eigenvalues and Eigenvectors

Problem: Find the eigenvalues and eigenvectors of:

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

Step 1: Set Up the Characteristic Equation The characteristic equation is:

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

Substitute $\mathbf{A}$ and $\mathbf{I}$:

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix}.$$

Compute the determinant:

$$\begin{aligned}
\det(\mathbf{A} - \lambda\mathbf{I}) &= \det\begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix} \\
&= (4-\lambda)(3-\lambda) - (1)(2) \\
&= (4-\lambda)(3-\lambda) - 2 \\
&= 12 - 4\lambda - 3\lambda + \lambda^2 - 2 \\
&= \lambda^2 - 7\lambda + 10.
\end{aligned}$$

Set the determinant to zero:

$$\lambda^2 - 7\lambda + 10 = 0.$$

Step 2: Solve for Eigenvalues Factorize the quadratic equation:

$$\lambda^2 - 7\lambda + 10 = (\lambda - 5)(\lambda - 2) = 0.$$

The eigenvalues are:

$$\lambda_1 = 5, \quad \lambda_2 = 2.$$

Step 3: Solve for Eigenvectors For each eigenvalue, solve $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$.

Eigenvector for $\lambda_1 = 5$: Substitute $\lambda_1 = 5$ into $\mathbf{A} - \lambda\mathbf{I}$:

$$\mathbf{A} - 5\mathbf{I} = \begin{bmatrix} 4-5 & 2 \\ 1 & 3-5 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}.$$

Solve $(\mathbf{A} - 5\mathbf{I})\mathbf{v} = \mathbf{0}$:

$$\begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This gives the system of equations:

$$-v_1 + 2v_2 = 0,$$
$$v_1 - 2v_2 = 0.$$

Both equations are equivalent. Let $v_2 = t$. Then:

$$v_1 = 2t.$$

The eigenvector is:

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} t = t \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Choose $t = 1$:

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Eigenvector for $\lambda_2 = 2$: Substitute $\lambda_2 = 2$ into $\mathbf{A} - \lambda\mathbf{I}$:

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 4-2 & 2 \\ 1 & 3-2 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}.$$

Solve $(\mathbf{A} - 2\mathbf{I})\mathbf{v} = \mathbf{0}$:

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This gives the system of equations:

$$2v_1 + 2v_2 = 0,$$
$$v_1 + v_2 = 0.$$

Both equations are equivalent. Let $v_2 = t$. Then:

$$v_1 = -t.$$

The eigenvector is:

$$\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} t = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Choose $t = 1$:

$$\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Step 4: Conclusion The eigenvalues and eigenvectors of $\mathbf{A}$ are:

- $\lambda_1 = 5$, $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.
- $\lambda_2 = 2$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Optimization Fundamentals

Optimization is the process of finding the best solution from a set of feasible solutions. In mathematical terms, it involves minimizing or maximizing an objective function subject to constraints.

Convex Sets and Functions

Convex Sets: A set $S \subseteq \mathbb{R}^n$ is convex if, for any two points $\mathbf{x}, \mathbf{y} \in S$ and any $\alpha \in [0, 1]$:

$$\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in S.$$

This means the line segment between any two points in S lies entirely within S.

Convex Functions: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if, for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha \in [0, 1]$:

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}).$$

This means the function lies below the line segment connecting any two points on its graph.

Gradient Descent

Definition: Gradient descent is an iterative optimization algorithm used to minimize a function $f(\mathbf{x})$. Starting from an initial guess $\mathbf{x}_0$, the algorithm updates $\mathbf{x}$ in the direction of the negative gradient:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \eta\nabla f(\mathbf{x}_k),$$

where:

- $\eta > 0$ is the learning rate (step size).
- $\nabla f(\mathbf{x}_k)$ is the gradient of f at $\mathbf{x}_k$.

Convergence: If f is convex and differentiable, gradient descent converges to a global minimum under appropriate choices of η.

Worked Example: Gradient Descent

Problem: Minimize the quadratic function:

$$f(x) = x^2 + 4x + 4.$$

Step 1: Compute the Gradient The gradient of f is the derivative:

$$\nabla f(x) = \frac{d}{dx}(x^2 + 4x + 4) = 2x + 4.$$

Step 2: Gradient Descent Update Rule Starting from an initial guess x_0, update x using:

$$x_{k+1} = x_k - \eta\nabla f(x_k).$$

Step 3: Perform Iterations Let $\eta = 0.1$ and $x_0 = 5$. Compute the first three iterations:

- $k = 0$:

$$\begin{aligned} x_1 &= x_0 - \eta \nabla f(x_0) \\ &= 5 - 0.1(2(5) + 4) \\ &= 5 - 0.1(14) \\ &= 5 - 1.4 = 3.6. \end{aligned}$$

- $k = 1$:

$$\begin{aligned} x_2 &= x_1 - \eta \nabla f(x_1) \\ &= 3.6 - 0.1(2(3.6) + 4) \\ &= 3.6 - 0.1(11.2) \\ &= 3.6 - 1.12 = 2.48. \end{aligned}$$

- $k = 2$:

$$\begin{aligned} x_3 &= x_2 - \eta \nabla f(x_2) \\ &= 2.48 - 0.1(2(2.48) + 4) \\ &= 2.48 - 0.1(8.96) \\ &= 2.48 - 0.896 = 1.584. \end{aligned}$$

Step 4: Interpret the Results After three iterations, the approximation to the minimum is $x_3 = 1.584$. Continuing further iterations will bring x_k closer to the true minimum at $x = -2$.

Properties of Convex Optimization

- If the objective function is convex, any local minimum is also a global minimum.
- Gradient descent is guaranteed to converge for appropriately chosen step sizes.
- Strongly convex functions have unique minima, and gradient descent converges faster.

Matrix Decompositions

Matrix decompositions are methods of breaking down a matrix into simpler components, making it easier to analyze or compute properties like determinants, eigenvalues, and solutions to linear systems.

LU Decomposition

Definition: LU decomposition factors a square matrix $\mathbf{A}$ into the product of a lower triangular matrix $\mathbf{L}$ and an upper triangular matrix $\mathbf{U}$:

$$\mathbf{A} = \mathbf{L}\mathbf{U},$$

where:

- $\mathbf{L}$ is a lower triangular matrix with ones on the diagonal.
- $\mathbf{U}$ is an upper triangular matrix.

Worked Example: LU Decomposition

Problem: Decompose the matrix:

$$\mathbf{A} = \begin{bmatrix} 4 & 3 \\ 6 & 3 \end{bmatrix}.$$

Step 1: Initialize L and U Start with:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}.$$

Step 2: Compute U Using $\mathbf{A} = \mathbf{L}\mathbf{U}$:

$$u_{11} = a_{11} = 4,$$
$$u_{12} = a_{12} = 3.$$

Step 3: Compute L

$$l_{21} = \frac{a_{21}}{u_{11}} = \frac{6}{4} = 1.5.$$

Step 4: Solve for u_{22}

$$u_{22} = a_{22} - l_{21} \cdot u_{12} = 3 - (1.5 \cdot 3) = -1.5.$$

Step 5: Write the Result

$$\mathbf{L} = \begin{bmatrix} 1 & 0 \\ 1.5 & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 4 & 3 \\ 0 & -1.5 \end{bmatrix}.$$

QR Decomposition

Definition: QR decomposition factors a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ into the product of an orthogonal matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$:

$$\mathbf{A} = \mathbf{Q}\mathbf{R}.$$

Worked Example: QR Decomposition

Problem: Compute the QR decomposition of:

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

Step 1: Apply the Gram-Schmidt Process

$$\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$
$$\mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Step 2: Form Q and R

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix},$$
$$\mathbf{R} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}.$$

Singular Value Decomposition (SVD)

Definition: SVD decomposes a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ into:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

where:

- $\mathbf{U} \in \mathbb{R}^{m \times m}$ is an orthogonal matrix (columns are left singular vectors).
- $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ is a diagonal matrix of singular values.
- $\mathbf{V} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix (columns are right singular vectors).

Worked Example: SVD

Problem: Compute the SVD of:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Step 1: Compute $\mathbf{\Sigma}$ The singular values are the square roots of the eigenvalues of $\mathbf{A}^T\mathbf{A}$:

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}.$$

The eigenvalues are 1 and 4, so the singular values are 1 and 2:

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Step 2: Compute U and V The eigenvectors of $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$ give $\mathbf{U}$ and $\mathbf{V}$, respectively:

$$\mathbf{U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$
$$\mathbf{V} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Step 3: Write the Result

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^T.$$