Applied Linear Algebra

Gergo Nador

August 22, 2024

Contents

1 Introduction

1.1 Systems of Linear Equations

A **linear equation** is an equation of the form:

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b \tag{1}$$

where $a_1, a_2, \ldots, a_n$ and b are real numbers, and $x_1, x_2, \ldots, x_n$ are variables.

A **system of linear equations** is a collection of one or more linear equations that involve the same set of variables.

1.2 Solution Set

The **solution set** of a linear system refers to the collection of values $s_1, s_2, s_3, \ldots$ that satisfy all equations in the system, thereby rendering the system valid or "true."

1. **No solution**: The system is classified as *inconsistent*.
2. **Exactly one solution (Unique)**: The system is classified as *consistent*.
3. **Infinitely many solutions**: The system is classified as *consistent*.

Existence Problem: Is there at least one solution to the system? If so, is the solution unique?

1.3 Matrix Representation of a System of Linear Equations

Consider the system of linear equations:

$$\begin{align} 2x_1 + 3x_2 + x_3 &= 3 \tag{2}\\ 0x_1 + 7x_2 - 4x_3 &= 10 \tag{3}\\ 0x_1 + 0x_2 + x_3 &= 1 \tag{4}\end{align}$$

This system can be represented using matrices in two distinct forms:

1.3.1 Coefficient Matrix

The coefficient matrix, which contains only the coefficients of the variables, is given by:

$$\begin{bmatrix} 2 & 3 & 1 \\ 0 & 7 & -4 \\ 0 & 0 & 1 \end{bmatrix}$$

1.3.2 Augmented Matrix

The augmented matrix, which includes both the coefficients and the constants from the right-hand side of the equations, is represented as:

$$\begin{bmatrix} 2 & 3 & 1 & | & 3 \\ 0 & 7 & -4 & | & 10 \\ 0 & 0 & 1 & | & 1 \end{bmatrix}$$

Here, the augmented matrix is an $m \times (n+1)$ matrix, where m is the number of equations and n is the number of variables.

1.4 Elementary Row Operations

The following are the fundamental row operations that can be performed on matrices:

1. **Replacement:** Replace one row by itself and a multiple of another row.
2. **Swap:** Interchange two rows.
3. **Scaling:** Multiply all entries in a row by a non-zero constant.

1.5 Consistency and Matrices

1. If a system has no solution (i.e., it is inconsistent), the corresponding matrix will exhibit an inconsistency when reduced. For example:

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

   The reduced matrix shows the inconsistency, as the last row implies $0x_1 + 0x_2 + 0x_3 = 1$.

2. If a system has a unique solution, the reduced matrix will be in a simplified form, with ones on the diagonal of the coefficient matrix and real numbers in the last column, as seen in our example.

3. If a system has infinitely many solutions, the reduced matrix will contain a row of all zeros. For instance:

$$\begin{bmatrix} 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & b_2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

   This row of zeros indicates that the system has multiple solutions corresponding to the free variables.

1.6 Echelon Form

A matrix is said to be in echelon form if it satisfies the following conditions:

1. All non-zero rows are above any rows of all zeros.
2. The leading non-zero entry in each row (known as a pivot) is to the right of the leading non-zero entry in the row above it.
3. All entries in a column below a leading non-zero entry are zeros.

1.7 Reduced Echelon Form

A matrix is in reduced echelon form if it satisfies the following additional conditions beyond those for echelon form:

4. The leading non-zero entry in each row is 1.
5. Each leading 1 is the only non-zero entry in its column.

Theorem: Each matrix is row-equivalent to one and only one matrix in reduced echelon form.

1.8 Pivots

A leading non-zero entry in an echelon form matrix is referred to as a **pivot**, and the column that contains a pivot is called a **pivot column**.

Pivot columns correspond to **basic variables**, while non-pivot columns correspond to **free variables**.

1.8.1 Consistency of a System

If a system is consistent, then:

(a) It has a unique solution if there are no free variables.

(b) It has infinitely many solutions if there is at least one free variable.

1.9 Vector Equations

A matrix with only one column is referred to as a **column vector**.

For example:

$$\mathbf{u} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

where $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$.

The same rules apply to vectors as they do for numbers.

1.10 Linear Combinations

Given a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p \in \mathbb{R}^n$ and scalars $c_1, c_2, \dots, c_p \in \mathbb{R}$, the vector $\mathbf{y}$ given by:

$$\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

is called a **linear combination** of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ with **weights** $c_1, c_2, \dots, c_p$.

1.11 Vector Equations and Linear Systems

A vector equation of the form:

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

has the same solution set as the linear system whose augmented matrix is:

$$[\mathbf{a}_1\, \mathbf{a}_2\, \mathbf{a}_3\, \dots\, \mathbf{a}_n \,|\, \mathbf{b}]$$

where:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & | & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & | & b_m \end{bmatrix}$$

is an $m \times n$ matrix, where $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are the column vectors of the matrix.

If $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_p$ is denoted as $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and is called the **subset of $\mathbb{R}^n$ spanned by $\mathbf{v}_1, \dots, \mathbf{v}_p$**.

1.12 Matrix Equation

A column vector $\mathbf{x}$ and a matrix A can be combined as the product of a matrix and a vector:

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

which simplifies to:

$$A\mathbf{x} = \mathbf{b}$$

If A is an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$, the matrix equation $A\mathbf{x} = \mathbf{b}$ has the same solution as the corresponding vector equation.

1.12.1 Linear, Vector, and Matrix Equations

Consider the following system of **linear equations**:

$$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ 2x_1 + x_2 + 3x_3 &= 11 \\ x_1 + 2x_2 + x_3 &= 8 \end{aligned}$$

This system can be expressed as a **vector equation**:

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \\ 8 \end{bmatrix}$$

Alternatively, it can be written as a **matrix equation**:

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \\ 8 \end{bmatrix}$$

1.13 Equivalent Conditions for a Matrix Equation

If A is an $m \times n$ matrix, then the following statements are equivalent:

(a) For each $\mathbf{b} \in \mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

(b) Each $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns of A.

(c) The columns of A span $\mathbb{R}^m$.

(d) A has a pivot in every row.

1.14 Homogeneous Systems

A linear system is called **homogeneous** if it can be written in the form:

$$A\mathbf{x} = \mathbf{0}$$

where A is an $m \times n$ matrix, $\mathbf{x}$ is an $n \times 1$ column vector of variables, and $\mathbf{0}$ is the $m \times 1$ zero vector.

The homogeneous system always has at least one solution, known as the **trivial solution**, where $\mathbf{x} = \mathbf{0}$. This means that all variables are zero, and it satisfies the equation $A\mathbf{x} = \mathbf{0}$.

However, a homogeneous system may also have infinitely many **non-trivial solutions**, where $\mathbf{x} \neq \mathbf{0}$. Non-trivial solutions exist if and only if the system has at least one free variable. This occurs when the number of pivot positions (leading entries) in the matrix A is less than the number of variables.

1.15 Parametric Vector Form

When a homogeneous system has free variables, its solution set can be expressed in the **parametric vector form**.

To express the solutions in parametric vector form, follow these steps:

1. Solve the system using row reduction to bring the augmented matrix into row echelon form (or reduced row echelon form).

2. Express the basic variables in terms of the free variables.

3. Write the general solution as a linear combination of vectors, where each vector is multiplied by a free variable.

   For example, consider a homogeneous system with the solution:

   $$x_1 = 2x_3, \quad x_2 = -x_3, \quad x_3 \text{ is free}$$

   The solution can be written in parametric vector form as:

   $$\mathbf{x} = x_3 \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

where x_3 is a free variable, and the vector $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ is the direction vector corresponding to x_3. This form clearly shows that the solution set is a line through the origin in $\mathbb{R}^3$.

1.16 Linear Dependence

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ is said to be **linearly independent** if the vector equation:

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution, where $x_1 = x_2 = \cdots = x_p = 0$.

If the vector equation has non-trivial solutions (i.e., there are scalars $x_1, x_2, \dots, x_p$ not all zero such that the equation holds), then the vectors are said to be **linearly dependent**.

Additionally:

- Two vectors are linearly independent if neither vector is a scalar multiple of the other.
- More generally, a vector is independent of a set of vectors if it cannot be expressed as a linear combination of the other vectors in the set.

1.16.1 Important Theorems on Linear Dependence and Independence

(a) If a set $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

(b) A set of two vectors in $\mathbb{R}^n$ is linearly independent if and only if neither vector is a scalar multiple of the other.

(c) Any set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$, that is, if there are more vectors than the number of entries in each vector (more columns than rows in the matrix representation).

(d) A set $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is linearly dependent if and only if at least one vector in S is a linear combination of the others. Assuming $\mathbf{v}_r \neq \mathbf{0}$, the vector $\mathbf{v}_r$ $(1 \leq r \leq p)$ is a linear combination of the preceding vectors $\mathbf{v}_1, \dots, \mathbf{v}_{r-1}$.

1.17 The Invertible Matrix Theorem

Let A be an $n \times n$ matrix. The following statements are each equivalent to the assertion that A is an invertible matrix:

a. A is invertible.

b. $\det(A) \neq 0$

c. A is row equivalent to the identity matrix I_n.

d. A has n pivot positions (i.e., one for each row and column).

e. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

f. The columns of A are linearly independent.

g. The linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.

h. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

i. The columns of A span $\mathbb{R}^n$.

j. The linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = A\mathbf{x}$ is onto.

k. There is an $n \times n$ matrix C such that $CA = I$.

l. There is an $n \times n$ matrix D such that $AD = I$.

m. The transpose $A^\top$ is invertible.

n. The columns of A form a basis of $\mathbb{R}^n$.

o. Col $A = \mathbb{R}^n$.

p. $\dim(\text{Col } A) = n$.

q. rank $A = n$.

r. Nul $A = \{\vec{0}\}$.

s. $\dim(\text{Nul } A) = 0$.

t. The columns of A form a basis of $\mathbb{R}^n$.

u. Col $A = \mathbb{R}^n$.

v. $\dim(\text{Col } A) = n$.

w. rank $A = n$.

x. Nul $A = \{\vec{0}\}$.

y. $\dim(\text{Nul } A) = 0$.

2 Matrix Operations

2.1 Adding and Subtracting Matrices

When adding or subtracting matrices, the following conditions must be met:

- Matrices must be of the **same size** (i.e., they must have the same dimensions).
- Addition and subtraction are **performed entry by entry**.

2.1.1 Properties of Matrix Addition and Subtraction

Matrix addition and subtraction follow these fundamental properties:

$$
\begin{aligned}
A + B &= B + A && \text{(Commutative property)} \\
(A + B) + C &= A + (B + C) && \text{(Associative property)} \\
A + \mathbf{0} &= A && \text{(Identity element of addition)} \\
A + (-A) &= \mathbf{0} && \text{(Inverse element of addition)} \\
A - B &= A + (-B) && \text{(Subtraction as addition of the inverse)} \\
A + B &= A + C && \iff B = C\text{(Cancellation law)}
\end{aligned}
$$

2.2 Scaling of Matrices

Scaling a matrix involves multiplying each entry of the matrix by a scalar. This operation is defined as follows:

2.2.1 Properties of Scaling

Let A and B be matrices of the same size, and let s and r be scalars. The following properties hold:

$$
\begin{aligned}
s(A + B) &= sA + sB && s \in \mathbb{R} \\
(A + B)s &= sA + sB && s \in \mathbb{R} \\
(r + s)A &= rA + sA && r, s \in \mathbb{R} \\
r(sA) &= (rs)A && r, s \in \mathbb{R}
\end{aligned}
$$

2.3 Multiplying Matrices

Consider two matrices:

$$A : \text{an } m \times n \text{ matrix} \quad \text{and} \quad B : \text{a } k \times l \text{ matrix}$$

Matrix multiplication AB is defined only if the number of columns in matrix A equals the number of rows in matrix B, i.e., $n = k$.

2.3.1 Example

Given:

$$A = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \quad (3 \times 5)$$

$$B = \begin{bmatrix} \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \end{bmatrix} \quad (5 \times 2)$$

Here, the product AB is defined because the number of columns in A (5 columns) matches the number of rows in B (5 rows).

The resulting matrix AB will have dimensions equal to the number of rows of A by the number of columns of B, so AB will be a 3×2 matrix.

The product BA is **not defined** because the number of columns in B (2 columns) does not match the number of rows in A (3 rows).

2.3.2 Row-Column Rule for Matrix Multiplication

If the product AB is defined, then the (i, j) entry of AB is the dot product of the i-th row of A with the j-th column of B.

Given matrices A of size $m \times n$ and B of size $n \times l$, the (i, j) entry of the resulting matrix AB is calculated as:

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

2.3.3 Properties of Matrix Multiplication

Let A be an $m \times n$ matrix. The following properties hold:

$$\begin{aligned}
A(BC) &= (AB)C && \text{(Associative property)} \\
A(B + C) &= AB + AC && \text{(Distributive property)} \\
(B + C)A &= BA + CA && \text{(Distributive property)} \\
r(AB) &= (rA)B = A(rB) && \text{(Scalar multiplication)} \\
I_m A &= A = AI_n && \text{(Multiplication by the identity matrix)}
\end{aligned}$$

where I_m and I_n are the identity matrices of size $m \times m$ and $n \times n$, respectively.

2.3.4 Commutativity in Matrix Multiplication

In general, matrix multiplication is not commutative, meaning $AB \neq BA$ in most cases. However, if $AB = BA$, we say that matrices A and B **commute** with one another.

2.4 Matrix Powers

If A is a square matrix, i.e., an $n \times n$ matrix, then the k-th power of A is defined as:

$$A^k = \underbrace{A \cdot A \cdot A \cdots A}_{k \text{ times}}$$

where k is a positive integer.

2.5 Transpose of a Matrix

If A is an $m \times n$ matrix, then the transpose of A, denoted by $A^\top$, is the $n \times m$ matrix whose columns are formed from the rows of A. Specifically, the element at the i-th row and j-th column of $A^\top$ is equal to the element at the j-th row and i-th column of A.

Consider the matrix A of size $m \times n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The transpose of A, denoted by $A^\top$, is the $n \times m$ matrix given by:

$$A^\top = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}$$

2.5.1 Properties of the Transpose

The following properties of the transpose operation hold:

$$\begin{aligned} (A^\top)^\top &= A \\ (A+B)^\top &= A^\top + B^\top \\ (rA)^\top &= rA^\top && \text{where } r \in \mathbb{R} \\ (AB)^\top &= B^\top A^\top \end{aligned}$$

2.6 Inverse of a Matrix

We say that a square matrix A is **invertible** (or **nonsingular**) if and only if there exists a matrix A^{-1} such that:

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

where A^{-1} is called the **inverse** of A, and I is the identity matrix of the same dimension as A.

2.6.1 Inverse of a 2×2 Matrix

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible, and the inverse of A is given by:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

This can also be written as:

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

where $\det A = ad - bc$ is called the **determinant** of the matrix A.

The determinant of A is calculated as:

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

2.6.2 Invertible $n \times n$ Matrices

If A is an invertible $n \times n$ matrix, then the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution.

This solution can be found by multiplying both sides of the equation by A^{-1}, the inverse of A:

$$A\mathbf{x} = \mathbf{b}$$
$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$$
$$\mathbf{x} = A^{-1}\mathbf{b}$$

2.6.3 Properties of Inverses of Matrices

Let A and B be invertible $n \times n$ matrices. The following properties hold:

$$\begin{aligned} (A^{-1})^{-1} &= A \\ (AB)^{-1} &= B^{-1}A^{-1} \\ (A^\top)^{-1} &= (A^{-1})^\top \end{aligned}$$

2.6.4 Algorithm for Finding the Inverse of a Matrix

An $n \times n$ matrix A is invertible if and only if it is row equivalent to the identity matrix I, i.e., we obtain the identity matrix when we reduce A to its reduced echelon form (RREF).

1. Set up the augmented matrix $[A \mid I]$, where I is the identity matrix of the same size as A.
2. Perform row reduction operations on $[A \mid I]$ until the left side of the augmented matrix is reduced to the identity matrix. The right side of the matrix will then be A^{-1}.

   The process can be represented as:

$$[A \mid I] \xrightarrow{\text{RREF}} [I \mid A^{-1}]$$

2.7 The Invertible Matrix Theorem

Theorem (The Invertible Matrix Theorem):
Let A be an $n \times n$ matrix. The following statements are equivalent:

(a) A is invertible.

(b) A is row equivalent to the identity matrix I_n.

(c) A has n pivot positions (i.e., one for each row and column).

(d) The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(e) The columns of A are linearly independent.

(f) The linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.

(g) The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

(h) The columns of A span $\mathbb{R}^n$.

(i) The linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = A\mathbf{x}$ is onto.

(j) There is an $n \times n$ matrix C such that $CA = I$.

(k) There is an $n \times n$ matrix D such that $AD = I$.

(l) The transpose $A^\top$ is invertible.

3 Determinants

Submatrix: Let A_{ij} denote the submatrix obtained by deleting the ith row and jth column from matrix A.

3.1 Determinant of a Matrix for $n \geq 2$

For $n \geq 2$, the determinant of a matrix $A = [a_{ij}]$ is computed as the sum of n terms of the form $(-1)^{i+j} a_{ij} \cdot \det(A_{ij})$, where the signs alternate. This can be expressed as:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \cdot \det(A_{ij}) \tag{5}$$

$$\text{or} \quad \det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot \det(A_{ij}) \tag{6}$$

3.2 Invertibility of a Matrix

Consider the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. It is known that if the determinant of A is zero, i.e., $\det(A) = ad - bc = 0$, then the matrix A is not invertible. However, to understand why this is the case, we assume $a \neq 0$.

3.2.1 Row Reduction and Invertibility

Given the matrix A, assume $a \neq 0$. We perform the following row operations:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{R_2 \to R_2 - \frac{c}{a} R_1} \begin{bmatrix} a & b \\ 0 & ad - bc \end{bmatrix}$$

If $ad - bc \neq 0$, then the matrix has pivots in every row, indicating that A is invertible.

3.3 Cofactor

Given a matrix $A = [a_{ij}]$, the (i, j)-cofactor of A is defined as:

$$C_{ij} = (-1)^{i+j} \cdot \det(A_{ij}) \tag{7}$$

The determinant of the matrix A can therefore be expressed as:

$$\det(A) = \sum_{i=1}^{n} a_{ij} \cdot C_{ij} \tag{8}$$

$$\text{or} \quad \det(A) = \sum_{j=1}^{n} a_{ij} \cdot C_{ij} \tag{9}$$

3.3.1 Cofactor Expansion

The method described previously is known as cofactor expansion. This technique can be applied to any row or any column of a matrix. The results can be restated as follows:

Cofactor Expansion along any Row:

$$\begin{aligned} \det(A) &= (-1)^{i+1} a_{i1} \det(A_{i1}) + (-1)^{i+2} a_{i2} \det(A_{i2}) + \cdots + (-1)^{i+n} a_{in} \det(A_{in}) && (10) \\ &= a_{i1} C_{i1} + a_{i2} C_{i2} + \cdots + a_{in} C_{in} && (11) \\ &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) = \sum_{j=1}^{n} a_{ij} C_{ij} && (12) \end{aligned}$$

Cofactor Expansion along any Column:

$$
\begin{aligned}
\det(A) &= (-1)^{1+j}a_{1j}\det(A_{1j}) + (-1)^{2+j}a_{2j}\det(A_{2j}) + \cdots + (-1)^{n+j}a_{nj}\det(A_{nj}) && (13)\\
&= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} && (14)\\
&= \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}) = \sum_{i=1}^{n}a_{ij}C_{ij} && (15)
\end{aligned}
$$

3.3.2 Theorem: Determinant of a Triangular Matrix

If A is a triangular matrix, then $\det(A)$ is equal to the product of the entries on the main diagonal.

Note: The echelon form of a matrix is also a triangular matrix. If the echelon form of a matrix has a free variable, then the determinant is zero, which implies that the matrix is not invertible.

3.3.3 Theorem: Effect of Elementary Row Operations on the Determinant

Let A be a square matrix. If matrix B is obtained from A by one of the following elementary row operations, the determinant of B is related to the determinant of A as follows:

1. **Replacement:** $\det(B) = \det(A)$.
2. **Swap:** $\det(B) = -\det(A)$.
3. **Scaling:** $\det(B) = k \cdot \det(A)$.

3.3.4 Echelon Form and Determinant

It is always possible to obtain the echelon form of a matrix using only the operations of replacement and swapping rows. The determinant of the original matrix A can then be expressed as:

$$\det(A) = (-1)^r \cdot \det(U) \tag{16}$$

where U is the echelon form of A, and r is the number of row swaps performed.

3.3.5 Theorem: Properties of Determinants

1. A square matrix A is invertible if and only if $\det(A) \neq 0$.
2. The determinant of the transpose of a matrix is equal to the determinant of the original matrix, i.e., $\det(A^\top) = \det(A)$.
3. The determinant of the product of two matrices is the product of their determinants, i.e., $\det(AB) = \det(A) \cdot \det(B)$.

4 Vector Spaces

A vector space is a collection of vectors where linear combinations can be formed. This implies the following characteristics:

1. It is a set of vectors.
2. The set is non-empty.
3. Addition and scalar multiplication are well-defined operations.
4. The set always contains the zero vector.
5. The structure satisfies ten axioms that define the behavior of vector spaces.

4.1 Closure

A vector space is said to be *closed* under both scalar multiplication and vector addition.

4.2 Subspace

A subspace is a subset of a vector space that is itself a vector space under the same operations. Specifically, a subspace is a smaller space within a larger vector space that satisfies the closure properties.

> A subspace is a vector space in its own right.

4.2.1 Theorem

Let $\vec{v}_1, \ldots, \vec{v}_p$ be vectors in a vector space V. Then, the span of $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is a subspace of V.

4.3 The Null Space

The null space of a matrix A, denoted as $\text{Nul}(A)$, is the set of all vectors $\vec{x}$ such that $A\vec{x} = \vec{0}$. In other words, it is the set of solutions to the homogeneous linear equation $A\vec{x} = \vec{0}$. The null space is a fundamental concept in linear algebra because it provides insight into the properties of the matrix A, particularly in relation to the solutions of linear systems.

Formally, the null space is defined as:

$$\text{Nul}(A) = \{\vec{x} \mid A\vec{x} = \vec{0}\}$$

where A is an $m \times n$ matrix, and $\vec{x}$ is an n-dimensional vector.

Note: The null space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$.

4.3.1 Simplified Explanation

In simpler terms, the null space of a matrix A is the collection of all vectors that, when multiplied by A, result in the zero vector. Think of it as finding all possible inputs $\vec{x}$ that give an output of zero when plugged into the matrix A. This set of inputs is what we call the null space. The null space can tell us if the matrix A has non-trivial solutions to the equation $A\vec{x} = \vec{0}$, indicating that the matrix is not of full rank.

4.3.2 Finding the Null Space

The null space of a matrix is comprised of the vectors that form the solution set in the parametric vector form (PVF). These vectors span the null space.

Example: Consider the matrix A given by:

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

To find the null space, we first reduce A to its reduced row echelon form (RREF):

$$\text{RREF}(A) = \begin{bmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Let the solution vector $\vec{x}$ be:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

Expressing the solution in terms of free variables x_2 and x_5, we obtain:

$$\vec{x} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

These vectors form the parametric vector form, and the null space Nul(A) is spanned by:

$$\text{Nul}(A) = \text{span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Thus, every linear combination of $\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{v}_3 = \begin{bmatrix} -3 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ is an element in Nul(A).

4.4 The Column Space

The column space of a matrix A, denoted as Col(A), is the set of all linear combinations of the columns of A. If A is an $m \times n$ matrix with columns $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$, then the column space is given by:

$$\text{Col}(A) = \text{span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$$

The column space is a subspace of $\mathbb{R}^m$ and consists of all vectors $\vec{b}$ in $\mathbb{R}^m$ such that there exists a vector $\vec{x}$ in $\mathbb{R}^n$ satisfying the equation $A\vec{x} = \vec{b}$.

4.4.1 Contrast Between Nul(A) and Col(A) for an $m \times n$ Matrix A

Nul A	**Col** A
1. Nul(A) is a subspace of $\mathbb{R}^n$.	1. Col(A) is a subspace of $\mathbb{R}^m$.
2. Nul(A) is implicitly defined; that is, you are given only a condition ($A\vec{x} = 0$) that vectors in Nul(A) must satisfy.	2. Col(A) is explicitly defined; that is, you are told how to build vectors in Col(A).
3. It takes time to find vectors in Nul(A). Row operations on $[A\ \vec{0}]$ are required.	3. It is easy to find vectors in Col(A). The columns of A are displayed; others are formed from them.
4. There is no obvious relation between Nul(A) and the entries in A.	4. There is an obvious relation between Col(A) and the entries in A, since each column of A is in Col(A).
5. A typical vector $\vec{v}$ in Nul(A) has the property that $A\vec{v} = \vec{0}$.	5. A typical vector $\vec{v}$ in Col(A) has the property that the equation $A\vec{x} = \vec{v}$ is consistent.
6. Given a specific vector $\vec{v}$, it is easy to tell if $\vec{v}$ is in Nul(A). Just compute $A\vec{v}$.	6. Given a specific vector $\vec{v}$, it may take time to tell if $\vec{v}$ is in Col(A). Row operations on $[A\ \vec{v}]$ are required.
7. Nul(A) $= \{\vec{0}\}$ if and only if the equation $A\vec{x} = \vec{0}$ has only the trivial solution.	7. Col(A) $= \mathbb{R}^m$ if and only if the equation $A\vec{x} = \vec{b}$ has a solution for every $\vec{b}$ in $\mathbb{R}^m$.
8. Nul(A) $= \{\vec{0}\}$ if and only if the linear transformation $\vec{x} \mapsto A\vec{x}$ is one-to-one.	8. Col(A) $= \mathbb{R}^m$ if and only if the linear transformation $\vec{x} \mapsto A\vec{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$.

4.5 Bases

A basis of a vector space is a linearly independent set of vectors that spans the subspace. In essence, it is the minimal set of vectors needed to span the entire subspace.

4.5.1 Bases for the Null Space

The basis for the null space consists of the vectors in the parametric vector form (PVF) of the solution set to the homogeneous equation $A\vec{x} = \vec{0}$.

4.5.2 Bases for the Column Space

The basis for the column space is formed by the pivot columns of the matrix A.

4.6 Dimension of a Vector Space

The dimension of a vector space V is defined as the number of vectors in a basis for V.

4.6.1 Dimensions of the special spaces

- dim(Nul A) = the number of free variables in the system $A\vec{x} = \vec{0}$.
- dim(Col A) = the number of pivot columns in the matrix A.

4.7 Row Space

The row space of a matrix A is the set of all linear combinations of the row vectors of A. Formally, the row space of A is denoted as:

$$\text{Row}(A) = \text{Col}(A^T)$$

This means that the row space of A is equivalent to the column space of the transpose of A.

Properties of Row Space

- If matrices A and B are row equivalent, they share the same row space and therefore have the same basis for their row space.
- The basis for the row space of A consists of all nonzero rows in the echelon form of A.

5 Eigenvalues

5.1 Eigenvector

A nonzero vector $\vec{x}$ is called an **eigenvector** of a matrix A if it satisfies the equation:

$$A\vec{x} = \lambda\vec{x} \tag{17}$$

for some scalar λ.

5.2 Eigenvalue

A scalar λ is referred to as an **eigenvalue** of the matrix A if there exists a non-trivial solution $\vec{x}$ to the equation:

$$A\vec{x} = \lambda\vec{x} \tag{18}$$

In such cases, the vector $\vec{x}$ is called an eigenvector corresponding to the eigenvalue λ.

5.2.1 Checking if a value is an eigenvalue

A scalar λ is an eigenvalue of a matrix A if and only if the matrix equation

$$A - \lambda I = 0 \tag{19}$$

has a non-trivial solution (has a free variable).

5.2.2 Eigenvalues of a Triangular Matrix

The eigenvalues of a triangular matrix are the entries located on its main diagonal.

5.2.3 Eigenvalue: Zero ($\lambda = 0$)

The scalar 0 is an eigenvalue of a matrix A if and only if A is not invertible.

For $\lambda = 0$, the eigenvalue equation becomes:

$$A\vec{x} = \lambda\vec{x} = 0 \cdot \vec{x} = \vec{0} \tag{20}$$

Thus, 0 is an eigenvalue only if the equation $A\vec{x} = \vec{0}$ has a non-trivial solution.

5.3 Eigenspace

The **eigenspace** corresponding to an eigenvalue λ of a matrix A is defined as the null space of the matrix $A - \lambda I_n$, where I_n is the identity matrix of appropriate dimension. It is denoted as $\mathcal{E}_A(\lambda)$.

Note: If the dimension of the null space $\text{Null}(A - \lambda I)$ is greater than 1, we cannot guarantee that the vectors in this null space are linearly independent.

5.3.1 Eigenvectors to Distinct Eigenvalues

If $\vec{v}_1, \ldots, \vec{v}_r$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix, then the set $\{\vec{v}_1, \ldots, \vec{v}_r\}$ is linearly independent.

Note: The eigenspace corresponding to the eigenvalue $\lambda = 0$ is equivalent to the null space of the matrix A, i.e., the set of all solutions $\vec{x}$ satisfying:

$$A\vec{x} = 0\vec{x} = \vec{0} \tag{21}$$

5.4 Characteristic Equation

A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if it satisfies the equation:

$$\det(A - \lambda I) = 0 \tag{22}$$

This equation is referred to as the **characteristic equation** or **characteristic polynomial** of the matrix A.

5.5 Multiplicities

5.5.1 Geometric Multiplicity

The **geometric multiplicity** of an eigenvalue λ of a matrix A is defined as the dimension of the eigenspace corresponding to λ. Mathematically, it is given by:

$$\dim \text{Null}(A - \lambda I) \tag{23}$$

5.5.2 Algebraic Multiplicity

The **algebraic multiplicity** $\mu(\lambda)$ of an eigenvalue λ of a matrix A is the number of times λ appears as a root of the characteristic equation. For instance, in the diagonal matrix, if $\mu(2) = 2$, it indicates that the eigenvalue $\lambda = 2$ appears twice.

5.5.3 Inequality between the Multiplicities

The algebraic multiplicity of an eigenvalue is always greater than or equal to its geometric multiplicity:

$$\text{Algebraic multiplicity} \geq \text{Geometric multiplicity} \tag{24}$$

5.6 Diagonalisation

An $n \times n$ matrix A is said to be **diagonalisable** if A is similar to a diagonal matrix D. This means there exists an invertible matrix P such that:

$$A = PDP^{-1} \tag{25}$$

where the diagonal entries of D are the eigenvalues of A, and the columns of P are the corresponding eigenvectors.

5.6.1 When is a Matrix Diagonalisable?

A matrix A is diagonalisable if and only if A has n linearly independent eigenvectors.

1. Diagonalisation is guaranteed if, for every eigenvalue of A, the geometric multiplicity equals the algebraic multiplicity.
2. A matrix with n distinct eigenvalues is a special case of this scenario.

5.6.2 Method: Diagonalisation

1. Find the eigenvalues of A:
   (a) If A is triangular, the eigenvalues are the entries on the diagonal.
   (b) Otherwise, find the roots of the characteristic equation $\det(A - \lambda I_n) = 0$.
2. For each eigenvalue λ, find a basis for the corresponding eigenspace by solving $(A - \lambda I_n)\vec{x} = \vec{0}$ using Parametric Vector Form (P.V.F).
3. Construct the matrix D by placing the eigenvalues on the diagonal of an $n \times n$ zero matrix. It is good practice to place them in descending order: $d_1 \geq d_2 \geq d_3 \geq \ldots \geq d_n$.
4. Construct the matrix P using the eigenvectors found in step 2. Note that the eigenvector must be placed in the same column as its corresponding eigenvalue in D.
5. Find P^{-1}.
6. Verify that $A = PDP^{-1}$.

5.6.3 A Question of Power

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(DP^{-1}) = PD^2P^{-1} \tag{26}$$
$$A^3 = (PDP^{-1})(PDP^{-1})(PDP^{-1}) = PD(DDP^{-1}) = PD^3P^{-1} \tag{27}$$

In general:

$$A^k = PD^kP^{-1} \tag{28}$$

5.7 Trace

The sum of the main diagonal elements of a square matrix is called the **trace** of the matrix. Mathematically, for a matrix A, the trace is denoted as:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn} \tag{29}$$

Therefore, the characteristic equation for a 2×2 matrix can be expressed as:

$$\det(A - \lambda I) = \lambda^2 - \mathrm{tr}(A) \cdot \lambda + \det(A) \tag{30}$$

5.7.1 Properties of the Trace

1. The trace of a matrix A is equal to the sum of its eigenvalues:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i \tag{31}$$

2. The trace of the sum of two matrices equals the sum of their traces:

$$\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B) \tag{32}$$

3. The trace of the product of two matrices A and B is equal to the trace of the product B and A:

$$\mathrm{tr}(AB) = \mathrm{tr}(BA) \tag{33}$$

4. The trace of a scalar multiple of a matrix is equal to the scalar times the trace of the matrix:

$$\mathrm{tr}(cA) = c \cdot \mathrm{tr}(A) \tag{34}$$

5.8 Summary

Given a matrix equation:

$$A\vec{x} = \lambda\vec{x} \tag{35}$$

where λ is the eigenvalue and $\vec{x}$ is the corresponding eigenvector.

5.8.1 Checking if λ is an Eigenvalue

To verify whether a scalar λ is an eigenvalue of the matrix A, compute $[A - \lambda I]$ and reduce it to Row Echelon Form (REF). If the resulting system is consistent, then λ is an eigenvalue.

5.8.2 Checking if $\vec{v}$ is an Eigenvector

To check if a vector $\vec{v}$ is an eigenvector of the matrix A, multiply A by $\vec{v}$ and verify whether the result is a scalar multiple of $\vec{v}$.

6 Differential Equations

A **Differential Equation** is an equation that relates one or more functions and their derivatives. The "unknowns" in these equations are the functions themselves.

Given a differential equation of the form:

$$y'(t) = a \cdot y(t) \tag{36}$$

where a is a constant, solving the equation involves determining the function $y(t)$:

1. $y(t) = e^{at}$, so
$$y'(t) = \left(e^{at}\right)' = a \cdot e^{at} = a \cdot y(t) \tag{37}$$

2. $y(t) = C \cdot e^{at}$, so
$$y'(t) = \left(C \cdot e^{at}\right)' = C \cdot a \cdot e^{at} = a \cdot y(t) \tag{38}$$

6.1 General Solution

In summary, the **general solution** to the differential equation

$$y'(t) = a \cdot y(t) \tag{39}$$

is given by:

$$y(t) = C \cdot e^{at} \tag{40}$$

since

$$y'(t) = \frac{d}{dt}\left(C \cdot e^{at}\right) = C \cdot a \cdot e^{at} = a \cdot y(t) \tag{41}$$

The solutions are **uniquely determined** by an initial value $y(t_0) = r_0$:

$$y(t_0) = r_0 \tag{42}$$
$$C \cdot e^{at_0} = r_0 \tag{43}$$
$$C = r_0 \cdot e^{-at_0} \tag{44}$$

6.2 Systems of First-Order Differential Equations

Consider a system of first-order differential equations:

$$y_1'(t) = a_{11}y_1(t) + a_{12}y_2(t) + \cdots + a_{1n}y_n(t) \tag{45}$$
$$y_2'(t) = a_{21}y_1(t) + a_{22}y_2(t) + \cdots + a_{2n}y_n(t) \tag{46}$$
$$\vdots$$
$$y_n'(t) = a_{n1}y_1(t) + a_{n2}y_2(t) + \cdots + a_{nn}y_n(t) \tag{47}$$

In vector notation:

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix} \tag{48}$$

Or, equivalently:

$$\mathbf{y}'(t) = A\mathbf{y}(t) \tag{49}$$

where $A = [a_{ij}] \in \mathbb{R}^{n\times n}$.

Given $\mathbf{y}(t) = C \cdot \vec{v} \cdot e^{\lambda t}$, it is a solution to the differential equation $\mathbf{y}'(t) = A\mathbf{y}(t)$.
Since $A\vec{v} = \lambda\vec{v}$, we have:

$$\begin{align}
\mathbf{y}'(t) &= \left(C \cdot \vec{v} \cdot e^{\lambda t}\right)' \tag{50}\\
&= C \cdot \vec{v} \cdot \left(e^{\lambda t}\right)' \tag{51}\\
&= \lambda\vec{v} \cdot Ce^{\lambda t} \tag{52}\\
&= A\vec{v} \cdot Ce^{\lambda t} \tag{53}\\
&= A\mathbf{y}(t) \tag{54}
\end{align}$$

6.3 Decoupling

When the matrix A is diagonal, i.e.,

$$A = \mathrm{diag}(d_1, d_2, d_3, \dots, d_n), \tag{55}$$

the system of differential equations becomes decoupled, meaning each derivative depends only on its corresponding function:

$$\begin{align}
y_1'(t) &= d_1 y_1(t) \tag{56}\\
y_2'(t) &= d_2 y_2(t) \tag{57}\\
&\vdots \notag\\
y_n'(t) &= d_n y_n(t) \tag{58}
\end{align}$$

6.3.1 Example

Consider the system of differential equations:

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \tag{59}$$

This gives us the following equations:

$$\begin{align}
y_1'(t) &= 3y_1(t) \tag{60}\\
y_2'(t) &= -5y_2(t) \tag{61}
\end{align}$$

The solutions to these equations are:

$$\begin{align}
y_1(t) &= C_1 e^{3t} \tag{62}\\
y_2(t) &= C_2 e^{-5t} \tag{63}
\end{align}$$

Thus, the solution vector is:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} C_1 e^{3t} \\ C_2 e^{-5t} \end{bmatrix} = C_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{3t} + C_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-5t} \tag{64}$$

6.4 Theorem

If the matrix A is **diagonalizable** with eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ and corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, the solution to the differential equation $\mathbf{y}'(t) = A\mathbf{y}(t)$ is:

$$\mathbf{y}(t) = C_1\vec{v}_1 e^{\lambda_1 t} + C_2\vec{v}_2 e^{\lambda_2 t} + \cdots + C_n\vec{v}_n e^{\lambda_n t} \tag{65}$$

Given an **initial vector** $\mathbf{y}(t_0) = \vec{r} \in \mathbb{R}^n$, a **unique solution** can be determined by using:

$$C_1\vec{v}_1 e^{\lambda_1 t_0} + C_2\vec{v}_2 e^{\lambda_2 t_0} + \cdots + C_n\vec{v}_n e^{\lambda_n t_0} = \vec{r} \tag{66}$$

This will provide a unique solution for the constants $C_1, C_2, \ldots, C_n$ since $\{\vec{v}_1, \ldots, \vec{v}_n\}$ forms a **basis**.

6.4.1 Example

Consider the system of differential equations:

$$y_1'(t) = y_1(t) + y_2(t) \tag{67}$$
$$y_2'(t) = 4y_1(t) + y_2(t) \tag{68}$$

This can be written in matrix form as:

$$\mathbf{y}'(t) = A\mathbf{y}(t) = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \mathbf{y}(t) \tag{69}$$

We first need to find the eigenvalues λ_1 and λ_2 by solving the characteristic equation:

$$\det(A - \lambda I) = \lambda^2 - (\mathrm{tr}A)\lambda + \det A = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1) \tag{70}$$

This gives the eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -1$.
For $\lambda = 3$:

$$\begin{bmatrix} 1-3 & 1 \\ 4 & 1-3 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1/2 \\ 0 & 0 \end{bmatrix} \Rightarrow \vec{v}_1 = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \tag{71}$$

For $\lambda = -1$:

$$\begin{bmatrix} 1+1 & 1 \\ 4 & 1+1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1/2 \\ 0 & 0 \end{bmatrix} \Rightarrow \vec{v}_2 = \begin{bmatrix} \frac{1}{2} \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \tag{72}$$

The **general solution** is:

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-t} + c_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{3t} \tag{73}$$

Expanding this, we get:

$$y_1(t) = c_1 e^{-t} + c_2 e^{3t} \tag{74}$$
$$y_2(t) = -2c_1 e^{-t} + 2c_2 e^{3t} \tag{75}$$

Assume $\mathbf{y}(1) = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, and find the unique solution.
We start with:

$$c_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-1} + c_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{3} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \tag{76}$$

This gives us the system of equations:

$$c_1 e^{-1} + c_2 e^{3} = 3 \tag{77}$$
$$-2c_1 e^{-1} + 2c_2 e^{3} = 2 \tag{78}$$

This can be written in matrix form as:

$$\begin{bmatrix} e^{-1} & e^{3} \\ -2e^{-1} & 2e^{3} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \tag{79}$$

To solve for c_1 and c_2, we perform row reduction (REF):

$$\begin{bmatrix} e^{-1} & e^{3} & 3 \\ -2e^{-1} & 2e^{3} & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & e^{1} \\ 0 & 1 & 2e^{3} \end{bmatrix} \tag{80}$$

This gives the solutions:

$$c_1 = e^{1} \tag{81}$$
$$c_2 = 2e^{3} \tag{82}$$

Therefore, the solution is:

$$y_1(t) = e^{1} e^{-t} + 2e^{3} e^{3t} \tag{83}$$
$$= e^{1-t} + 2e^{3(1+t)} \tag{84}$$

$$y_2(t) = -2e^{1} e^{-t} + 2e^{3} e^{3t} \tag{85}$$
$$= -2e^{1-t} + 4e^{3(1+t)} \tag{86}$$

The solution can also be expressed as:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = C_1 \begin{bmatrix} v_{1_1} \\ v_{1_2} \end{bmatrix} e^{\lambda_1 t} + C_2 \begin{bmatrix} v_{2_1} \\ v_{2_2} \end{bmatrix} e^{\lambda_2 t} \tag{87}$$

Or, for this example:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = e^{1} \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-t} + 2e^{3} \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{3t} \tag{88}$$

7 Orthogonality

7.1 Vector Definitions

7.1.1 Dot Product (Inner Product)

The dot product of two vectors is defined as follows:

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

The dot product is the sum of the entry-by-entry products:

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \tag{89}$$

The dot product results in a scalar.

7.1.2 Length of a Vector (Norm)

The length or norm of a vector is defined as:

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

A vector whose length is 1 is referred to as a *unit vector*. It can be expressed as:

$$\vec{u} = \frac{1}{\|\vec{v}\|} \cdot \vec{v}$$

This process is also known as *normalizing* a vector.

7.1.3 Example: Length of a Vector

Given the vector:

$$\vec{v} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

Find a unit vector in the direction of $\vec{v}$:

$$\vec{u} = \frac{1}{\|\vec{v}\|}\vec{v} = \frac{1}{\sqrt{(-1)^2 + 2^2 + 0^2}} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{1+4+0}} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

This simplifies to:

$$\vec{u} = \begin{bmatrix} \frac{-1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{bmatrix} \approx \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \end{bmatrix}$$

7.1.4 Distance in $\mathbb{R}^n$

The distance between two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ is defined as:

$$\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \sqrt{(\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v})}$$

This can be expanded as:

$$\text{dist}(\vec{u}, \vec{v}) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$$

This distance is also referred to as the *Euclidean Distance* or the L_2-norm.

7.2 Orthogonal Vectors

Two vectors in $\mathbb{R}^n$ are said to be *orthogonal* to each other if:

$$\vec{u} \cdot \vec{v} = 0$$

This condition can also be expressed using the Pythagorean Theorem. Specifically, two vectors are orthogonal if and only if:

$$\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$$

Note the geometric interpretation of the sum of two vectors, which aligns with the properties of a right triangle.

7.3 Orthogonal Complement

If W is a subspace in $\mathbb{R}^n$, and $\vec{z}$ is orthogonal to all vectors in W, we say that $\vec{z}$ is orthogonal to W. The set of all vectors that are orthogonal to W is called the *Orthogonal Complement* of W and is denoted by $W^\perp$.

Note that $W^\perp$ is also a subspace of $\mathbb{R}^n$.

7.4 Null Space, Column Space, and Row Space

Let A be an $m \times n$ matrix. The relationships between the row space, column space, and null space of A are as follows:

$$(\text{Row } A)^\perp = \text{Null } A$$
$$(\text{Col } A)^\perp = \text{Null } A^T$$

These relationships indicate that the orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T.

7.5 Orthogonal Sets

If $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is also linearly independent. This implies that S is a basis for $\mathbb{R}^p$. If $p = n$, then S is a basis for $\mathbb{R}^n$.

1. Each distinct pair of vectors in S is orthogonal.
2. Independence does **not** entail orthogonality.
3. Orthogonality is stricter than independence.

7.6 Expressing a Vector as a Linear Combination of an Orthogonal Set

Given an orthogonal set of vectors $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ and a vector $\vec{y}$, we can express $\vec{y}$ as a linear combination of the orthogonal vectors in the set:

$$\vec{y} = c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_n\vec{u}_n$$

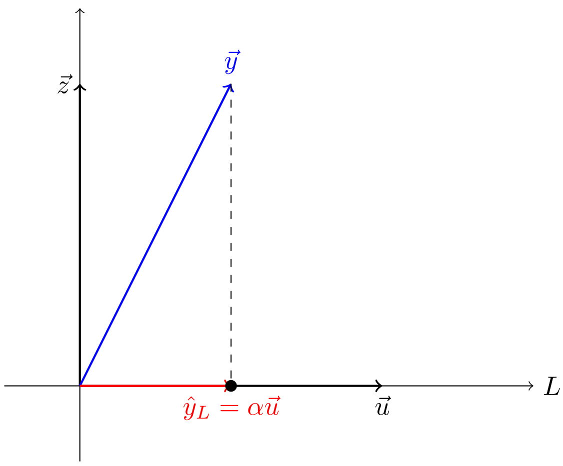
The coefficients c_i, often referred to as *weights*, can be computed using the formula:

$$c_i = \frac{\vec{y} \cdot \vec{u}_i}{\vec{u}_i \cdot \vec{u}_i}$$

7.7 Orthogonal Projection

Given a vector $\vec{y}$ in $\mathbb{R}^n$ and a subspace L spanned by a vector $\vec{u}$, the orthogonal projection of $\vec{y}$ onto L is the vector $\vec{y}_L$ that lies in L and is as close as possible to $\vec{y}$.

7.7.1 Process



1. **Decomposing $\vec{y}$:**
   The vector $\vec{y}$ can be decomposed into two components: one that lies in the subspace L (denoted by $\hat{y}_L$) and one that is orthogonal to L (denoted by $\vec{z}$):
   $$\vec{y} = \hat{y}_L + \vec{z}$$

2. **Orthogonal Condition:**
   Since $\vec{z}$ is orthogonal to $\vec{u}$, we have:
   $$\vec{z} \cdot \vec{u} = 0$$

3. **Expression for $\vec{z}$:**
   Substituting $\hat{y}_L = \alpha\vec{u}$ into the equation:
   $$\vec{z} = \vec{y} - \hat{y}_L = \vec{y} - \alpha\vec{u}$$
   The orthogonal condition becomes:
   $$(\vec{y} - \alpha\vec{u}) \cdot \vec{u} = 0$$

4. **Solving for α:**
   $$\alpha\vec{u} \cdot \vec{u} = \vec{y} \cdot \vec{u}$$
   $$\alpha = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}$$

5. **Orthogonal Projection:**
   Finally, the orthogonal projection of $\vec{y}$ onto L is given by:
   $$\hat{y}_L = \text{proj}_L\vec{y} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \cdot \vec{u}$$

7.8 The Orthogonal Decomposition Theorem

Let W be a subspace of $\mathbb{R}^n$. Then each vector $\vec{y}$ in $\mathbb{R}^n$ can be written uniquely as:

$$\vec{y} = \hat{y}_W + \vec{z}$$

where $\hat{y}_W$ is in W and $\vec{z}$ is in $W^{\perp}$.

Further Explanation:

If $\{\vec{u}_1, \ldots, \vec{u}_p\}$ is an orthogonal basis for W, then $\hat{y}_W$ can be expressed as:

$$\hat{y}_W = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 + \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2 + \cdots + \frac{\vec{y} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p}\vec{u}_p$$

And the orthogonal component $\vec{z}$ is given by:

$$\vec{z} = \vec{y} - \hat{y}_W$$

This decomposition allows us to break down any vector $\vec{y}$ in $\mathbb{R}^n$ into two components: one that lies entirely within the subspace W and one that lies in the orthogonal complement of W, denoted by $W^\perp$.

In a more concrete sense, if $\hat{y}_W$ can be written as:

$$\hat{y}_W = a\vec{u}_1 + b\vec{u}_2 + c\vec{u}_3 + d\vec{u}_4$$

where a, b, c, and d are scalar coefficients, then this represents the unique linear combination of the basis vectors in W.

7.9 Orthonormal Sets

- An **orthonormal set** is an orthogonal set of unit vectors.
- If a subspace W is spanned by an orthonormal set, then this set forms an orthonormal basis for W.

An $m \times n$ matrix U has orthonormal columns if and only if:

1. $U^T U = I$, where I is the identity matrix.
2. $\|U\vec{x}\| = \|\vec{x}\|$ for any vector $\vec{x}$.
3. $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$, meaning the dot product is preserved.
4. $(U\vec{x}) \cdot (U\vec{y}) = 0$ if and only if $\vec{x} \cdot \vec{y} = 0$, which implies that $\vec{x}$ and $\vec{y}$ are orthogonal.

7.10 Gram-Schmidt Process (GSP)

Given a basis, the Gram-Schmidt Process (GSP) provides a method for obtaining an orthogonal basis for any non-zero subspace W.

Assume a basis $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$, let:

$$\vec{v}_1 = \vec{x}_1$$

$$\vec{v}_2 = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1$$

$$\vec{v}_3 = \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2$$

$$\vdots$$

$$\vec{v}_p = \vec{x}_p - \frac{\vec{x}_p \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{x}_p \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 - \cdots - \frac{\vec{x}_p \cdot \vec{v}_{p-1}}{\vec{v}_{p-1} \cdot \vec{v}_{p-1}}\vec{v}_{p-1}$$

Then $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$ is an orthogonal set.

$$\vec{v}_p = \vec{x}_p - \left(\sum_{i=1}^{p-1} \frac{\vec{x}_p \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i}\right)\vec{v}_i = \vec{x}_p - \left(\sum_{i=1}^{p-1} \text{proj}_{\vec{v}_i}\vec{x}_p\right)$$

7.11 Least Squares Problem: Best Approximation

Problem: Not all systems of linear equations $A\vec{x} = \vec{b}$ have solutions. In particular, when $\vec{b}$ is not in the column space of A (denoted as Col A), there is no exact solution.

Solution: The Least Squares method finds the best approximation $\hat{\vec{b}} \in \text{Col } A$ such that there exists an $\hat{\vec{x}}$ satisfying:

$$A\hat{\vec{x}} = \hat{\vec{b}}$$

This $\hat{\vec{x}}$ is called the *least squares solution* to $A\vec{x} = \vec{b}$.

Geometric Interpretation: The vector $\vec{b}$ is decomposed into two components: one in Col A (which is $\hat{\vec{b}}$) and one orthogonal to Col A. The least squares solution $\hat{\vec{x}}$ is the point in Col A closest to $\vec{b}$, and it minimizes the distance between $A\hat{\vec{x}}$ and $\vec{b}$.

Mathematical Derivation: Let $A = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n]$ be the matrix of column vectors $\vec{a}_i$.

1. To find the best approximation $\hat{\vec{b}}$, we project $\vec{b}$ onto Col A. The condition for orthogonality implies:

$$\vec{a}_i \cdot (\vec{b} - \hat{\vec{b}}) = 0 \quad \text{for all } i$$

2. Express A^T as:

$$A^T = \begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vdots \\ \vec{a}_n^T \end{bmatrix}$$

3. The orthogonality condition becomes:

$$A^T(\vec{b} - A\hat{\vec{x}}) = 0$$

4. Expanding this, we get:

$$A^T\vec{b} - A^T A\hat{\vec{x}} = 0$$

5. Rearranging gives the normal equation:

$$A^T A\hat{\vec{x}} = A^T\vec{b}$$

Solving the Normal Equation: To solve for $\hat{\vec{x}}$, we solve the normal equation $A^T A\hat{\vec{x}} = A^T\vec{b}$. This is typically done by computing the matrix $A^T A$ and the vector $A^T\vec{b}$, then solving the resulting system using row reduction (RREF).

The augmented matrix is given by:

$$\left[A^T A \mid A^T\vec{b}\right] \sim \left[I \mid \hat{\vec{x}}\right]$$

where I is the identity matrix and $\hat{\vec{x}}$ is the least squares solution.

This approach ensures that the solution $\hat{\vec{x}}$ minimizes the sum of the squares of the residuals (the differences between the actual and predicted values).

7.11.1 Method: Finding the Least Squares Solution

To find the least squares solution $\hat{\vec{x}}$ to the equation $A\vec{x} = \vec{b}$, follow these steps:

1. Compute the matrix $A^T A$.
2. Compute the vector $A^T\vec{b}$.
3. Solve the normal equation $A^T A\hat{\vec{x}} = A^T\vec{b}$ by row reducing the augmented matrix $\left[A^T A \mid A^T\vec{b}\right]$.
4. State the solution $\hat{\vec{x}}$.

7.12 Linear Models

Problem: Real-world problems rarely adhere perfectly to mathematical functions.

Solution: The method of least squares is employed to find the function or model that yields the best approximation of the observed data by minimizing the error.

$$\hat{y} = \beta_0 + \beta_1 x \tag{90}$$

$$\mathbf{X}\boldsymbol{\beta} \approx \mathbf{y} \tag{91}$$

where $\hat{y}$ is the predicted value, β_0 and β_1 are the coefficients, and x is the independent variable.

The goal is to determine the parameters $\boldsymbol{\beta}$ that minimize the difference between the observed values and the values predicted by the linear model.

7.12.1 Linear Regression Model

To construct a simple linear regression model involves estimating the coefficients β_0 and β_1 that minimize the error. Specifically, the goal is to find β_0 and β_1 such that:

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\| \leq \|\mathbf{y} - \mathbf{X}\beta\| \tag{92}$$

where:

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

The *least squares* method is used to solve for the coefficients:

1. Compute $\mathbf{X}^t\mathbf{X}$
2. Compute $\mathbf{X}^t\mathbf{y}$
3. Solve the normal equation $[\mathbf{X}^t\mathbf{X} \quad \mathbf{X}^t\mathbf{y}]$ using row reduction to find $\hat{\beta}_0$ and $\hat{\beta}_1$
4. State the estimated coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$

8 Symmetric Matrices, SVD and PCA

8.1 Symmetric Matrices

A matrix A is called **symmetric** if it satisfies the condition $A = A^T$. In other words, a matrix is symmetric if it is equal to its transpose.

$$A = \begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{bmatrix} \tag{93}$$

Additionally, for a symmetric matrix A, the elements satisfy the relation $a_{ij} = a_{ji}$ for all i and j.

8.2 Diagonalizing Symmetric Matrices

Given a symmetric matrix A, the process of diagonalization involves the following steps:

1. **Find the eigenvalues:**
   To determine the eigenvalues λ of A, solve the characteristic equation:
   $$\det(A - \lambda I) = 0 \tag{94}$$
   This results in a polynomial equation in λ, whose roots are the eigenvalues.

2. **Find the eigenvectors:**
   For each eigenvalue λ_i, find the corresponding eigenvector $\mathbf{v}_i$ by solving the system:
   $$(A - \lambda_i I)\mathbf{v}_i = 0 \tag{95}$$
   Ensure that the eigenvectors are linearly independent and, if necessary, orthonormalize them.

3. **Construct the matrix P:**
   Form the matrix P by placing the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ as columns:
   $$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} \tag{96}$$

4. **Find the diagonal matrix D:**
   The diagonal matrix D is formed by placing the corresponding eigenvalues along its diagonal:
   $$D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{97}$$

5. **Verify orthogonality (if needed):**
   Ensure that the eigenvectors are orthogonal. For symmetric matrices, the eigenvectors corresponding to distinct eigenvalues are always orthogonal.

Thus, the original matrix A can be expressed as:

$$A = PDP^{-1} \tag{98}$$

where P is the orthogonal matrix of eigenvectors and D is the diagonal matrix of eigenvalues.

8.3 Theorems and Definitions Related to Symmetric Matrices

8.3.1 Orthogonality of Eigenvectors Theorem

If a matrix A is symmetric, then any two eigenvectors corresponding to distinct eigenvalues are orthogonal.

8.3.2 Orthogonal Matrix Inversion Theorem

If Q is an orthogonal matrix, then it satisfies the property:

$$Q^{-1} = Q^T \tag{99}$$

8.3.3 Definition of Orthogonal Diagonalization

An $n \times n$ matrix A is said to be **orthogonally diagonalizable** if there exists an orthogonal matrix P such that:

$$A = PDP^T \tag{100}$$

where D is a diagonal matrix.

8.3.4 Condition for Orthogonal Diagonalization

An $n \times n$ matrix A is **orthogonally diagonalizable** if and only if A is symmetric, i.e., $A = A^T$.

Key Difference from Regular Diagonalization The essential difference between orthogonal diagonalization and regular diagonalization is that the columns of the matrix P are orthonormal.

Method: Orthogonal Diagonalization The process of orthogonally diagonalizing a matrix involves the following steps:

1. Find the eigenvalues of A.
2. Find the basis vectors for the eigenspaces corresponding to each eigenvalue found in step 1.
3. Apply the Gram-Schmidt process, if necessary, to the eigenvectors obtained in step 2 to ensure they are orthogonal.
4. Normalize all vectors obtained from steps 2 and 3, and use them to construct the matrices P and P^T.
5. Construct the diagonal matrix D using the eigenvalues.

Note: Eigenvalues are also referred to as the **spectrum** of the matrix.

8.3.5 Spectral Theorem

If A is a symmetric matrix, i.e., $A = A^T$, then the following properties hold:

1. A has n real eigenvalues (counting algebraic multiplicities).
2. The dimension of the null space $\text{Nul}(A - \lambda_i I)$ is equal to the algebraic multiplicity of the eigenvalue λ_i.
3. The eigenspaces corresponding to distinct eigenvalues are mutually orthogonal.
4. A is orthogonally diagonalizable.

8.4 Singular Value Decomposition (SVD)

The Singular Value Decomposition (SVD) of a matrix $A_{m\times n}$ is expressed as:

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^T_{n\times n} \tag{101}$$

where:

- V consists of the eigenvectors of A^TA (right singular vectors). **Important (below)**
- U consists of the eigenvectors of AA^T (left singular vectors). **Important (below)**
- Σ is a diagonal matrix containing the singular values of A on its diagonal. The singular values are the square roots of the eigenvalues of A^TA or AA^T, and they are arranged in decreasing order of magnitude.

The corresponding vectors in U and V are placed in corresponding columns.

8.4.1 Important Consideration

It is crucial that you either:

1. Find V and then derive the vectors that constitute U, or
2. Find U and then derive the vectors that constitute V.

Do not attempt to use the eigenvectors of both $A^T A$ for V and AA^T for U independently without ensuring they correspond correctly, as this can lead to incorrect results.

8.4.2 Column-wise Interpretation of SVD

The Singular Value Decomposition (SVD) of a matrix A can be expressed as:

$$A = U\Sigma V^T = U\Sigma V^{-1} \tag{102}$$

This implies:

$$AV = U\Sigma \tag{103}$$

To better understand this, consider the decomposition by taking one column at a time. For each i-th column vector $\mathbf{v}_i$ of V, we have:

$$A\mathbf{v}_1 = \sigma_1 \mathbf{u}_1 \tag{104}$$
$$A\mathbf{v}_2 = \sigma_2 \mathbf{u}_2 \tag{105}$$
$$A\mathbf{v}_3 = \sigma_3 \mathbf{u}_3 \tag{106}$$
$$\vdots \tag{107}$$
$$A\mathbf{v}_n = \sigma_n \mathbf{u}_n \tag{108}$$

where σ_i are the singular values, and $\mathbf{u}_i$ are the corresponding left singular vectors.

8.4.3 Deriving Left Singular Vectors

To find the left singular vectors $\mathbf{u}_i$, the following relation is used:

$$\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i = \sigma_i^{-1} A\mathbf{v}_i \tag{109}$$

8.4.4 Guide: Choosing the Correct Matrix for SVD

When performing Singular Value Decomposition (SVD), it is important to choose the correct matrix to start with, depending on the shape of A:

- **If $m > n$ (more rows than columns):**
  - The number of rows in U will be greater than the number of rows in V.
  - If we start from $A^T A$, we will obtain too few eigenvectors. Instead, **depart from** AA^T.
- **If $n > m$ (more columns than rows):**
  - The number of rows in U will be fewer than the number of rows in V.
  - If we start from AA^T, we will obtain too few eigenvectors. Instead, **depart from** $A^T A$.

8.4.5 Method: Detailed Steps for SVD

To perform the Singular Value Decomposition (SVD) of a matrix A, follow these steps:

1. **Find the eigenvalues:**
   - (a) Compute the eigenvalues of AA^T if $m > n$.
   - (b) Compute the eigenvalues of $A^T A$ if $m < n$.

2. **Find the eigenvectors and normalize them:**
   - (a) Depending on the choice in step 1, if AA^T was used, find the eigenvectors $\mathbf{u}_i$.
   - (b) If A^TA was used, find the eigenvectors $\mathbf{v}_i$.
   - **Note:** Use the Gram-Schmidt process if $\lambda_i = \lambda_j$ for $i \neq j$.

3. **Compute the singular values σ_i:**

$$\sigma_i = \sqrt{\lambda_i} \tag{110}$$

   where λ_i are the eigenvalues found in step 1. Define the diagonal matrix Σ and use padding with zeros to adjust the size if necessary.

4. **Construct matrices U and V:**
   - (a) If deriving $\mathbf{v}_i$ from $\mathbf{u}_i$, use the relation $\mathbf{v}_i^T = \sigma_i^{-1}\mathbf{u}_i^T A$.
   - (b) If deriving $\mathbf{u}_i$ from $\mathbf{v}_i$, use the relation $\mathbf{u}_i = \sigma_i^{-1} A\mathbf{v}_i$.
   - **Note:** If $\lambda_i = 0$, you may need to use the Gram-Schmidt process again to ensure orthogonality.

5. **Verify the decomposition:**
   Test that $A = U\Sigma V^T$ holds true, confirming the accuracy of the SVD.

8.4.6 Constructing Orthogonal Vectors: Use Cases

Use Case:

1. You started with the "wrong" choice of AA^T or A^TA and are missing one or more eigenvectors.

2. One of the eigenvalues is zero, making it impossible to derive vectors from each other directly.

8.5 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is one of the most widely used techniques for dimensionality reduction in statistics and machine learning. It is essentially a method for compressing the data while preserving as much variance as possible.

8.5.1 Steps for Performing PCA

1. **Center the data:**
   Subtract the mean from all features (columns) to center the data around the origin. Let $\mathbf{X}$ be the data matrix after centering.

2. **Compute the covariance matrix:**
   Calculate the covariance matrix $\mathbf{C}$ of the centered data:

$$\mathbf{C} = \frac{1}{n}\mathbf{X}^T\mathbf{X} \quad \text{or} \quad \mathbf{C} = \frac{\mathbf{X}^T\mathbf{X}}{\sqrt{\mathrm{Var}(X_i)}} \tag{111}$$

   where n is the number of samples and $\mathrm{Var}(X_i)$ represents the variance of the features. The latter formulation normalizes the covariance matrix to obtain the correlation matrix Corr.

3. **Perform eigen decomposition:**
   Find the eigenvectors and eigenvalues of the covariance matrix $\mathbf{C}$ through eigen decomposition. The eigenvectors represent the directions of maximum variance (principal components), and the eigenvalues represent the variance along these directions.

4. **Sort the eigenvalues:**
   Sort the eigenvalues in descending order. The eigenvectors corresponding to the largest eigenvalues capture the most significant variance in the data.

5. **Form the matrix V:**
   Construct the matrix V by stacking the eigenvectors corresponding to the largest eigenvalues. This gives:
$$\mathbf{C}\mathbf{V} = \mathbf{V}\mathbf{D} \tag{112}$$
   where $\mathbf{D}$ is the diagonal matrix of eigenvalues.

6. **Calculate the principal components:**
   The principal components $\mathbf{T}$ are obtained by projecting the original data onto the eigenvector space:
$$\mathbf{T}_{n\times k} = \mathbf{X}_{n\times p}\mathbf{V}_{p\times k} \tag{113}$$
   where k is the number of principal components chosen.

PCA is often denoted as PCA($\mathbf{X}$) and is used to reduce the dimensionality of the data while retaining the components that contribute most to the variance.

8.5.2 Using SVD for PCA

In Principal Component Analysis (PCA), Singular Value Decomposition (SVD) can be used as an efficient method to decompose the centered data matrix $\mathbf{X}$ and extract the principal components.

- **Assume:** The centered data matrix $\mathbf{X}$ can be decomposed using SVD as:
$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T \tag{114}$$
  Here, $\mathbf{U}$ contains the left singular vectors, Σ is the diagonal matrix of singular values, and $\mathbf{V}$ contains the right singular vectors. The matrix $\mathbf{X}$ should be centered before applying SVD to ensure that the principal components correspond to the directions of maximum variance.

- **Principal Components:** The principal components $\mathbf{T}$ can be obtained by:
$$\mathbf{T} = \mathbf{U}\Sigma \tag{115}$$
  The matrix $\mathbf{T}$ represents the data projected onto the principal components.

- **Fraction of Variance Explained:** The proportion of total variance explained by the first r principal components is given by:
$$\frac{\sum_{i=1}^{r}\lambda_i}{\sum_{i=1}^{n}\lambda_i} \tag{116}$$
  where λ_i are the eigenvalues (or squared singular values) corresponding to the variance explained by each principal component. This ratio lies in the interval $[0, 1]$ and is used to determine how much of the original data's variance is captured by the first r components.

9 Extras

9.1 Rules for Diagonalizability of a Matrix

1. **Distinct Eigenvalues**:
   - **Rule**: If a matrix has n distinct eigenvalues (where n is the size of the matrix), then it is diagonalizable.
   - **Why**: Distinct eigenvalues ensure that the corresponding eigenvectors are linearly independent, which is sufficient for diagonalizability.

2. **Geometric Multiplicity Equals Algebraic Multiplicity**:

- **Rule**: A matrix is diagonalizable if, for each eigenvalue, the geometric multiplicity (the number of linearly independent eigenvectors associated with the eigenvalue) is equal to its algebraic multiplicity (the number of times the eigenvalue appears as a root of the characteristic polynomial).
- **Why**: If the geometric multiplicity of each eigenvalue matches its algebraic multiplicity, the matrix has enough linearly independent eigenvectors to form a basis, which is necessary for diagonalization.

3. **Symmetric Matrices**:
   - **Rule**: Any real symmetric matrix is diagonalizable.
   - **Why**: Symmetric matrices have real eigenvalues and a full set of orthonormal eigenvectors, making them diagonalizable by an orthogonal matrix.
4. **Normal Matrices**:
   - **Rule**: A matrix is diagonalizable if it is normal, meaning A satisfies $AA^* = A^*A$, where A^* is the conjugate transpose of A.
   - **Why**: Normal matrices have a complete set of orthonormal eigenvectors, ensuring diagonalizability.
5. **Diagonalizable by Similarity**:
   - **Rule**: A matrix A is diagonalizable if it is similar to a diagonal matrix, meaning there exists an invertible matrix P such that $P^{-1}AP$ is diagonal.
   - **Why**: By definition, diagonalization involves finding such a matrix P.
6. **Jordan Form and Nilpotent Matrices**:
   - **Rule**: A matrix is diagonalizable if its Jordan canonical form is already diagonal or if its Jordan blocks (corresponding to each eigenvalue) are 1×1.
   - **Why**: If the Jordan form has no blocks larger than 1×1, the matrix has a sufficient number of independent eigenvectors.
7. **Matrix Powers and Nilpotent Property**:
   - **Rule**: A matrix is diagonalizable if its minimal polynomial (the smallest polynomial that the matrix satisfies) has only distinct roots.
   - **Why**: If the minimal polynomial splits into distinct linear factors, the matrix is diagonalizable because it indicates the existence of a full set of eigenvectors.
8. **Defective Matrices**:
   - **Rule**: A matrix is **not** diagonalizable if it is defective, meaning it does not have a complete set of eigenvectors (i.e., geometric multiplicity is less than algebraic multiplicity for some eigenvalue).
   - **Why**: A matrix with a missing eigenvector cannot be diagonalized.
9. **Complex Matrices**:
   - **Rule**: A complex matrix is diagonalizable over the complex field if it has a full set of linearly independent eigenvectors.
   - **Why**: Diagonalizability is defined similarly in complex spaces, relying on the existence of a full set of eigenvectors.
10. **Projection Matrices**:
    - **Rule**: Projection matrices (idempotent matrices A such that $A^2 = A$) are diagonalizable.
    - **Why**: Projection matrices have eigenvalues 0 and 1, and their eigenspaces are sufficient for diagonalization.

9.2 Definition of a Singular Matrix

A **singular matrix** is a square matrix that does not have an inverse. In other words, if A is a singular matrix, then there does not exist a matrix B such that $AB = BA = I$, where I is the identity matrix of the same dimension.

9.2.1 Key Characteristics of a Singular Matrix:

1. **Determinant**:
   - The determinant of a singular matrix is 0. This is the primary condition for a matrix to be singular. If the determinant of a matrix A is zero ($\det(A) = 0$), then A is singular.
2. **Linear Dependence**:
   - The rows or columns of a singular matrix are linearly dependent. This means that at least one row or column can be expressed as a linear combination of the others.
3. **Rank**:
   - The rank of a singular matrix is less than its order (the number of rows or columns). The rank is the maximum number of linearly independent rows or columns in the matrix.
4. **Eigenvalues**:
   - A square matrix is singular if and only if it has at least one eigenvalue equal to zero. This is because having a zero eigenvalue implies that the matrix is not invertible.

9.2.2 Example

Consider the matrix A:

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

The determinant of A is:

$$\det(A) = (2 \times 2) - (4 \times 1) = 4 - 4 = 0$$

Since the determinant is 0, A is a singular matrix. Notice that the second row is a multiple of the first row, indicating linear dependence.

9.2.3 Summary

A singular matrix is one that cannot be inverted, has a determinant of zero, and exhibits linear dependence among its rows or columns. It is characterized by a rank that is less than its full order and the presence of at least one zero eigenvalue.

9.3 Rules of Determinant

1. **Determinant of a Product**:

   $$\det(AB) = \det(A) \times \det(B)$$

   The determinant of the product of two matrices is equal to the product of their determinants.
2. **Determinant of a Transpose**:

   $$\det(A^T) = \det(A)$$

   The determinant of a matrix is equal to the determinant of its transpose.
3. **Determinant of an Inverse**:

   $$\det(A^{-1}) = \frac{1}{\det(A)}$$

   if A is invertible. The determinant of the inverse of a matrix is the reciprocal of the determinant of the matrix.

4. **Determinant of an Identity Matrix**:

$$\det(I_n) = 1$$

The determinant of the identity matrix (of any size) is 1.

5. **Determinant and Row Operations**:

   - **Row Swap**: Swapping two rows (or columns) of a matrix multiplies the determinant by -1.
   - **Row Multiplication**: Multiplying a row (or column) by a scalar k multiplies the determinant by k.
   - **Row Addition**: Adding a multiple of one row (or column) to another row (or column) does not change the determinant.

6. **Determinant of a Triangular Matrix**:

$$\det(A) = \prod_{i=1}^{n} a_{ii}$$

The determinant of a triangular matrix (upper or lower) is the product of its diagonal elements.

7. **Determinant of Block Matrices**:

$$\det\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) = \det(A) \times \det(B)$$

For a block diagonal matrix, the determinant is the product of the determinants of the blocks.

8. **Determinant of a Scalar Multiple**:

$$\det(cA) = c^n \det(A)$$

for an $n \times n$ matrix A. The determinant of a matrix multiplied by a scalar c is c^n times the determinant of the matrix, where n is the size of the matrix.

9. **Zero Determinant**:

$$\det(A) = 0 \implies A \text{ is singular (non-invertible).}$$

If the determinant of a matrix is zero, the matrix is singular, meaning it does not have an inverse.

10. **Determinant and Eigenvalues**:

$$\det(A) = \prod_{i=1}^{n} \lambda_i$$

where λ_i are the eigenvalues of the matrix A. The determinant of a matrix is the product of its eigenvalues.

11. **Determinant of a 2x2 Matrix**:

$$\det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = ad - bc$$

For a 2×2 matrix, the determinant is calculated as the difference between the product of the diagonals.

12. **Cofactor Expansion (Laplace Expansion)**:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij})$$

The determinant of an $n \times n$ matrix can be calculated by expanding along any row or column, using cofactors. Here, a_{ij} is an element of the matrix and M_{ij} is the submatrix obtained by removing the ith row and jth column.

13. **Determinant of a Matrix with a Zero Row or Column**:

If a matrix has a row or a column of all zeros, its determinant is zero.

9.4 The Gram Matrix

Given a set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ in an inner product space (such as $\mathbb{R}^n$), the **Gram matrix** G is a matrix that encodes the inner products between each pair of vectors in the set. The Gram matrix is defined as follows:

$$G = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_n \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_n \cdot \mathbf{u}_1 & \mathbf{u}_n \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_n \cdot \mathbf{u}_n \end{pmatrix}$$

In other words, the entry G_{ij} of the Gram matrix G is given by the dot product (or inner product) of the vectors $\mathbf{u}_i$ and $\mathbf{u}_j$:

$$G_{ij} = \mathbf{u}_i \cdot \mathbf{u}_j$$

Properties of the Gram Matrix

- **Symmetry**: The Gram matrix G is symmetric, meaning $G_{ij} = G_{ji}$ for all i and j.
- **Positive Semi-Definiteness**: The Gram matrix is positive semi-definite, which implies that for any non-zero vector $\mathbf{x} \in \mathbb{R}^n$, the quadratic form $\mathbf{x}^T G \mathbf{x} \geq 0$.
- **Linear Independence**: The rank of the Gram matrix G is equal to the number of linearly independent vectors in the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$. If the set of vectors is linearly independent, the Gram matrix is positive definite, meaning $\mathbf{x}^T G \mathbf{x} > 0$ for any non-zero vector $\mathbf{x}$.

Example Consider a set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in $\mathbb{R}^3$, where:

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

The Gram matrix G for this set of vectors is:

$$G = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \mathbf{u}_1 \cdot \mathbf{u}_3 \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \mathbf{u}_2 \cdot \mathbf{u}_3 \\ \mathbf{u}_3 \cdot \mathbf{u}_1 & \mathbf{u}_3 \cdot \mathbf{u}_2 & \mathbf{u}_3 \cdot \mathbf{u}_3 \end{pmatrix} = \begin{pmatrix} 6 & 1 & -2 \\ 1 & 6 & -3 \\ -2 & -3 & 6 \end{pmatrix}$$

In this example, the Gram matrix G is symmetric and positive definite (assuming the vectors are linearly independent).

9.4.1 Conclusion

The Gram matrix provides valuable information about the geometric relationships between vectors, including their inner products, linear independence, and overall structure in the vector space.

9.5 Determining the Shortest Distance Between a Vector and a Subspace

To determine the shortest distance between a vector and a subspace, you can use the concept of orthogonal projection. Here's a step-by-step guide:

9.5.1 Step 1: Understand the Problem

Given a vector $\mathbf{v}$ and a subspace W of a vector space (typically $\mathbb{R}^n$), the shortest distance from $\mathbf{v}$ to W is the length of the vector that goes from $\mathbf{v}$ to its orthogonal projection onto W.

9.5.2 Step 2: Find an Orthogonal Basis for the Subspace

If W is spanned by a set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$, you may need to find an orthogonal (or orthonormal) basis for W using the **Gram-Schmidt process**.

However, if the vectors are already orthogonal, or if you already have an orthogonal basis, you can skip this step.

9.5.3 Step 3: Compute the Orthogonal Projection

The orthogonal projection of $\mathbf{v}$ onto the subspace W can be found as:

$$\text{Proj}_W(\mathbf{v}) = \sum_{i=1}^{k} \frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \mathbf{u}_i$$

where $\mathbf{u}_i$ are the basis vectors for W.

If the basis is orthonormal (i.e., $\mathbf{u}_i \cdot \mathbf{u}_i = 1$ for all i), the projection simplifies to:

$$\text{Proj}_W(\mathbf{v}) = \sum_{i=1}^{k} (\mathbf{v} \cdot \mathbf{u}_i)\mathbf{u}_i$$

9.5.4 Step 4: Calculate the Shortest Distance

The shortest distance d between the vector $\mathbf{v}$ and the subspace W is the length of the vector from $\mathbf{v}$ to its projection onto W. This is given by:

$$d = \|\mathbf{v} - \text{Proj}_W(\mathbf{v})\|$$

Here, $\|\mathbf{v} - \text{Proj}_W(\mathbf{v})\|$ is the Euclidean norm (or length) of the vector $\mathbf{v} - \text{Proj}_W(\mathbf{v})$.

9.5.5 Example

Suppose W is a subspace of $\mathbb{R}^3$ spanned by the vectors $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and we want to find the shortest distance from the vector $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ to W.

- **Orthonormal Basis:** The vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are already orthonormal.
- **Projection:**

$$\text{Proj}_W(\mathbf{v}) = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2$$

  Calculate each term:

$$\mathbf{v} \cdot \mathbf{u}_1 = 1 \quad \text{and} \quad \mathbf{v} \cdot \mathbf{u}_2 = 1$$

  So,

$$\text{Proj}_W(\mathbf{v}) = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

- **Shortest Distance:**

$$\mathbf{v} - \text{Proj}_W(\mathbf{v}) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$d = \|\mathbf{v} - \text{Proj}_W(\mathbf{v})\| = \sqrt{0^2 + 0^2 + 1^2} = \sqrt{1} = 1$$

So, the shortest distance from $\mathbf{v}$ to the subspace W is 1.

9.6 Identifying Eigenvalues and Their Multiplicities

The eigenvalues of a matrix are the solutions λ to the equation:

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

where A is a square matrix, and I is the identity matrix of the same size. The characteristic polynomial is derived from this determinant equation, and its roots (where the polynomial equals zero) are the eigenvalues.

9.6.1 Procedure

1. **Roots of the Polynomial:** The points where the characteristic polynomial graph intersects the x-axis represent the eigenvalues. Each crossing point indicates an eigenvalue of the matrix.
2. **Multiplicity of Each Eigenvalue:** The multiplicity of an eigenvalue is determined by the number of times the polynomial touches or crosses the x-axis at that eigenvalue's point. This can typically be inferred by observing:
   - If the polynomial *crosses* the x-axis, the eigenvalue has an *odd* multiplicity.
   - If the polynomial *touches* the x-axis (tangent) without crossing, the eigenvalue has an *even* multiplicity.
3. **Sign and Impact of Eigenvalues:**
   - Negative eigenvalues can suggest a system with certain stability properties or behaviors distinct from those with positive eigenvalues.
   - Positive eigenvalues are typically associated with growth or expansion in the context of dynamical systems.

9.6.2 Example

Given a plot of a characteristic polynomial, identify the eigenvalues and their multiplicities:

- The polynomial crosses the x-axis at $\lambda = -3$ with no rebound, indicating an eigenvalue of -3 with a multiplicity of 1.
- The polynomial crosses at $\lambda = 2$, turning back and re-crossing, indicating an eigenvalue of 2 with a multiplicity of 2.

9.7 Understanding Matrix Dimensions from SVD

Assume a matrix A of size $m \times n$ undergoes singular value decomposition (SVD) resulting in a specified number of left singular vectors, right singular vectors, and non-zero singular values. This provides specific insights into the dimensions of various subspaces associated with A.

9.7.1 Definitions

- **Left singular vectors** correspond to the column space of A^T (also called the row space of A), and their count gives the rank of A.
- **Right singular vectors** correspond to the column space of A, and their count often equals n but does not directly impact the rank.
- **Non-zero singular values** count directly indicates the rank of A.

9.7.2 Given

- **Three left singular vectors**
- **Five right singular vectors**
- **One non-zero singular value**

9.7.3 Implications

- The **rank of** A is 1 (indicated by one non-zero singular value).
- **Dimension of the column space of** A (dim Col A) is also 1, equal to the rank.
- **Dimension of the null space of** A (dim Nul A), for a matrix A of size $m \times n$, is $n -$ rank. Assuming $n = 5$ (suggested by the number of right singular vectors), it is $5 - 1 = 4$.
- **Dimension of the null space of** A^T (dim Nul A^T) is $m -$ rank. The assumption here requires knowledge of m, but given the correction, it is $m - 1$ where m is estimated based on the three left singular vectors or adjusted contextually.
- **Dimension of the column space of** A^T (dim Col A^T) is equal to the rank of A, which is 1.

9.7.4 Conclusion

The singular value decomposition helps in deducing the dimensions of the column and null spaces of both A and A^T, facilitating the understanding of matrix structure and its implications in vector space mappings.

9.8 Key Concepts in Linear Algebra

- **Vector spaces**: The fundamental structures in linear algebra where vectors are defined and operated upon.
- **Subspaces**: Subsets of vector spaces that themselves are vector spaces, closed under vector addition and scalar multiplication.
- **Column Space (Col A)**: The set of all linear combinations of the column vectors of a matrix A. Represents one type of subspace.
- **Null Space (Nul A)**: The set of all vectors $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$. Another important subspace that provides insight into the solutions of homogeneous linear systems.
- **Row Space**: The set of all linear combinations of the row vectors of a matrix A, equivalent in dimension to the column space of A^T.
- **Dimension of the Null Space (Nullity)**: Also known as the nullity of A, it is the dimension of the null space and provides insights into the number of parameters in the general solution to $A\mathbf{x} = \mathbf{0}$.
- **Rank (Rank = Dim Col A)**: The dimension of the column space of A, also equal to the maximum number of linearly independent columns of A. It is closely related to the dimensionality and invertibility of A.

9.9 Steps to Verify if a Vector is an Eigenvector

1. **Multiply the Matrix by the Vector**: Multiply the given matrix A by the vector $\mathbf{v}$.

$$\mathbf{w} = A\mathbf{v}$$

2. **Compare the Result with a Scalar Multiple of the Vector**: After performing the multiplication, check if the resulting vector $\mathbf{w}$ is a scalar multiple of the original vector $\mathbf{v}$. In other words, see if there exists a scalar λ such that:

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

   Here, λ is the eigenvalue associated with the eigenvector $\mathbf{v}$.

3. **Conclude**:
   - If such a scalar λ exists, then $\mathbf{v}$ is an eigenvector of A, and λ is the corresponding eigenvalue.
   - If no such scalar λ exists, then $\mathbf{v}$ is **not** an eigenvector of A.

9.10 Normalize a Vector

Let $\mathbf{v}$ be a vector in $\mathbb{R}^n$:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

The magnitude (or norm) of $\mathbf{v}$, denoted $\|\mathbf{v}\|$, is given by:

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

or equivalently,

$$\|\mathbf{v}\| = \sqrt{\sum_{i=1}^{n} v_i^2}$$

To normalize the vector $\mathbf{v}$, divide each component by the magnitude $\|\mathbf{v}\|$:

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{1}{\|\mathbf{v}\|} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \frac{v_1}{\|\mathbf{v}\|} \\ \frac{v_2}{\|\mathbf{v}\|} \\ \vdots \\ \frac{v_n}{\|\mathbf{v}\|} \end{pmatrix}$$

9.11 Inverse of a 3x3 Matrix

Given a 3×3 matrix A:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

Step 1: Calculate the determinant of A

The determinant $\det(A)$ is calculated as:

$$\det(A) = a(ei - fh) - b(di - fg) + c(dh - eg)$$

Step 2: Calculate the matrix of minors

The minor of an element is the determinant of the 2×2 matrix obtained by deleting the row and column of that element. For matrix A, the minors are:

$$\text{Minor}_{11} = \begin{vmatrix} e & f \\ h & i \end{vmatrix} = ei - fh$$

$$\text{Minor}_{12} = \begin{vmatrix} d & f \\ g & i \end{vmatrix} = di - fg$$

$$\text{Minor}_{13} = \begin{vmatrix} d & e \\ g & h \end{vmatrix} = dh - eg$$

And similarly for the other minors.

Step 3: Calculate the matrix of cofactors
To form the matrix of cofactors, apply the checkerboard pattern of signs (+, -, +, etc.) to the minors:

$$\text{Cofactor}(A) = \begin{pmatrix} +\text{Minor}_{11} & -\text{Minor}_{12} & +\text{Minor}_{13} \\ -\text{Minor}_{21} & +\text{Minor}_{22} & -\text{Minor}_{23} \\ +\text{Minor}_{31} & -\text{Minor}_{32} & +\text{Minor}_{33} \end{pmatrix}$$

This becomes:

$$\text{Cofactor}(A) = \begin{pmatrix} ei - fh & -(di - fg) & dh - eg \\ -(bi - ch) & ai - cg & -(ah - bg) \\ bf - ce & -(af - cd) & ae - bd \end{pmatrix}$$

Step 4: Transpose the cofactor matrix to get the adjugate matrix
Transpose the cofactor matrix:

$$\text{Adj}(A) = \begin{pmatrix} ei - fh & -(bi - ch) & bf - ce \\ -(di - fg) & ai - cg & -(af - cd) \\ dh - eg & -(ah - bg) & ae - bd \end{pmatrix}$$

Step 5: Divide the adjugate matrix by the determinant
Finally, the inverse of matrix A is given by:

$$A^{-1} = \frac{1}{\det(A)} \text{Adj}(A)$$

Full Formula for Inverse of 3×3 Matrix
The inverse of the matrix A can be compactly written as:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} ei - fh & -(bi - ch) & bf - ce \\ -(di - fg) & ai - cg & -(af - cd) \\ dh - eg & -(ah - bg) & ae - bd \end{pmatrix}$$

9.12 Orthogonal Complement: Trick

We use "the trick" where just find the nullspace of the transpose since this will be the orthogonal complement.

10 Exam Topics

10.1 Part 1

- **Trace and Eigenvalue:** The trace of a matrix is the sum of its diagonal elements. The eigenvalues of a matrix are the roots of its characteristic polynomial. Note that the trace of a matrix equals the sum of its eigenvalues.
- **Echelon Form:** The echelon form of a matrix does not provide information about orthogonality or subspace structure; it is primarily used for solving systems of linear equations and determining the rank of a matrix.
- **Dimensionality of a Subspace:** The dimension of a subspace is the number of vectors in a basis for that subspace. This is a measure of the "size" or "degree of freedom" within that subspace.
- **Eigenvalues and Invertibility:** Eigenvalues are crucial for understanding the invertibility of a matrix. If $\lambda = 0$ is an eigenvalue, then $\det(A) = 0$, meaning the matrix A is not invertible.
- **Linear Dependence and Independence:** A set of vectors is linearly dependent if one vector can be expressed as a linear combination of the others. Linear independence means no vector in the set can be written as a combination of the others.

- **Characteristic Polynomial and Rank:** The characteristic polynomial is derived from $\det(A - \lambda I) = 0$. The rank of a matrix A can be inferred from the number of non-zero eigenvalues, considering their multiplicity.
- **Multiplicities:** The algebraic multiplicity of an eigenvalue is its multiplicity as a root of the characteristic polynomial. The geometric multiplicity is the number of linearly independent eigenvectors corresponding to that eigenvalue. Algebraic multiplicity is always greater than or equal to geometric multiplicity.
- **Diagonalization:** A matrix A is diagonalizable if there exists a matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix. This happens when the sum of the geometric multiplicities equals the number of distinct eigenvalues.
- **Nullity and Free Variables:** The nullity of a matrix A is the dimension of the null space $\text{null}(A)$, which corresponds to the number of free variables in the solution to $A\mathbf{x} = 0$.
- **Orthonormal Sets:** A set of vectors is orthonormal if each vector is of unit length and orthogonal to each other. For vectors $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{u}^T\mathbf{v} = 0$ indicates orthogonality.
- **Linearly Independent Sets:** If a set of vectors is linearly independent, any subset of these vectors is also linearly independent.
- **Determinant Rules:** The determinant of a matrix can be used to check invertibility. If $\det(A) \neq 0$, the matrix is invertible; otherwise, it is singular (not invertible).
- **Solving $A\mathbf{x} = \lambda\mathbf{x}$:** To find eigenvalues and eigenvectors, solve the equation $A\mathbf{x} = \lambda\mathbf{x}$ or equivalently $(A - \lambda I)\mathbf{x} = 0$.
- **Eigenspace:** The eigenspace corresponding to an eigenvalue λ is the set of all vectors $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$. It is the null space of $(A - \lambda I)$.

10.2 Part 2

- **Linear Combination and Matrix Equations:** Understanding how much of vector A and vector B contribute to forming another vector is essential. This is often represented by a linear combination and can be set up as a matrix equation. Transitioning from a vector equation to a matrix equation is a key skill.
- **Systems of Matrix Equations:** Solving systems of linear equations can be generalized to matrix equations. This involves finding solutions to equations of the form $A\mathbf{x} = \mathbf{b}$, where A is a matrix and $\mathbf{x}$ and $\mathbf{b}$ are vectors.
- **Orthonormal Sets and Gram Matrix:** A set of vectors is orthonormal if each vector has unit length and all pairs of vectors are orthogonal. The Gram matrix is a matrix composed of dot products of vectors, and for an orthonormal set, it is the identity matrix.
- **Projection and Distance (Norm):** The projection of a vector $\mathbf{v}$ onto a vector $\mathbf{w}$ gives the component of $\mathbf{v}$ in the direction of $\mathbf{w}$. The distance between vectors can be measured using norms, particularly the Euclidean norm (also called the L_2 norm). The length of a vector and its projection onto another vector is often crucial in least squares problems and other applications.
- **Vector Intersections and Span:** The intersection of vector subspaces is the set of vectors that are in both subspaces. The span of vectors is the set of all possible linear combinations of those vectors. Understanding these concepts is essential for solving problems involving vector spaces.
- **Regression and Least Squares:** In the least squares method, we aim to minimize the error between the observed values and the values predicted by a linear model. The design matrix A is multiplied by the coefficient vector $\hat{\beta}$ to predict the output vector $\mathbf{y}$. The error is given by $\mathbf{y} - A\hat{\beta}$.
- **Mixing Problems and Limits:** Mixing problems often involve finding the concentration of substances in two or more tanks as they mix over time. The problem might involve solving differential equations and using limits (as $t \to \infty$ or $t \to 0$) to find steady-state solutions or initial conditions.

- **LU Decomposition:** LU decomposition is a method to factor a matrix A into a product of a lower triangular matrix L and an upper triangular matrix U. This method is preferred over using echelon form when solving for solutions to linear systems and for finding the determinant efficiently.
- **Key Points from Assignments:**
  - Be familiar with the Gram matrix, orthonormal sets, and how to calculate projections and distances.
  - Understand how to apply least squares and interpret the error in regression.
  - LU Decomposition is crucial for solving matrix equations more efficiently than using row reduction alone.