

MA1522 Notes (AY24/25 Sem1)

Michael Yang

September 3, 2024

Chapter 1: Linear Systems

1.1 Introduction to Linear Systems

Definition

A **linear equation** with n variables in **standard form** has the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

Here a_1, a_2, \dots, a_n are known constants, called the **coefficients**, b is called the **constant** and x_1, x_2, x_n are variables.

The linear equation is homogeneous if $b = 0$, i.e.

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0.$$

A **system of linear equations**, or a **linear system** consists of a finite number of linear equations. In general, a linear system with m variables and n equations in standard form is written as

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

The linear system is **homogeneous** if $b_1 = b_2 = \cdots = b_m = 0$, that is, all the linear equations are homogeneous.

Given a linear system, we say that

$$x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$$

is a **solution** to the **linear system** if the equations are simultaneously satisfied after making the substitution.

The **general solution** to a linear system captures all possible solutions to the linear system.

A linear system is said to be **inconsistent** if it does not have any solutions. It is **consistent** otherwise, that is, a linear system is consistent if it has at least one solution.

1.2 Solving a Linear System and Row-Echelon Form

Definition

(Augmented Matrix)

A linear system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

can be expressed uniquely as an **augmented matrix**

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right)$$

Definition

In the augmented matrix, a **zero row** is a row with all entries 0. A **leading entry** of a row is the first nonzero entry of the row counting from the left.

An augmented matrix is in **row-echelon form (REF)** if

1. If **zero rows** exists, they are at the bottom of the matrix.
2. The **leading entries** are **further to the right** as we move down the rows.

An augmented matrix in **REF** has the form

$$\left(\begin{array}{cccccc|c} * & & & & & & * \\ 0 & \dots & 0 & * & & & * \\ 0 & \dots & 0 & 0 & \dots & 0 & * \\ \vdots & & & & & & \vdots \\ 0 & \dots & & & \dots & 0 & 0 \end{array} \right)$$

In the **row-echelon form**, a **pivot column** is a column containing a **leading entry**. Otherwise, it is called a **non-pivot column**.

The augmented matrix is in **reduced row-echelon form (RREF)** if further

1. The leading entries are 1.
2. In each pivot column, all entries except the leading entry is 0.

A matrix in **RREF** has the form

$$\left(\begin{array}{cccccc|cc} 0 & \dots & 1 & & * & 0 & * & 0 & * & * \\ 0 & \dots & 0 & \dots & 0 & 1 & * & 0 & * & * \\ 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 & 1 & * \\ 0 & \dots & 0 & \dots & & & & 0 & 0 & 0 \\ \vdots & & & & \vdots & & & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 0 & & \dots & 0 & 0 & 0 \end{array} \right)$$

Solutions from REF and RREF

- If the augmented matrix is in row-echelon form, we perform **back substitution** to obtain the solutions.
- If the augmented matrix is in reduced row-echelon form, we will read off the solutions directly.

Number of solutions from row-echelon form.

- **No solution:** a row of zero before the bar (coefficient matrix) and a non zero number after the bar.
- **Unique solution:** all columns of coefficient matrix are pivot columns (not possible if there is more variables than equations)
- **Infinitely many solutions:** when there is a non-pivot column in the augmented matrix before the bar

1.3 Elementary Row Operations

Definition

There are 3 types of **elementary row operations**.

1. Exchanging 2 rows, $R_i \leftrightarrow R_j$;
2. Adding a multiple of a row to another, $R_i + cR_j, c \in \mathbb{R}$ and $i \neq j$;
3. Multiplying a row by a nonzero constant, $aR_j, a \neq 0$.

Two **(augmented) matrices** are **row equivalent** if one can be obtained from the other by **elementary row operations**.

Theorem

Two linear systems have the **same solutions** if their **augmented matrices** are **row equivalent**.

Definition

(Reverse of elementary row operations)

Every elementary row operation has a reverse elementary row operation. The reverse of the row operations are given as such.

1. The reverse of exchanging 2 rows, $R_i \leftrightarrow R_j$, is itself.
2. The reverse of adding a multiple of a row to another, $R_i + cR_j$ is subtracting the multiple of that row, $R_i - cR_j$.
3. The reverse of multiplying a row by a nonzero constant, aR_j is the multiplication of the reciprocal of the constant, $\frac{1}{a}R_j$.

1.4 Row Reduction, Gaussian and Gauss-Jordan Elimination

Definition

Gaussian elimination

Step 1: Locate the leftmost column that does not consist entirely of zeros.

Step 2: Interchange the top row with another row, if necessary, to bring a nonzero entry to the top of the column found in Step 1.

Step 3: For each row below the top row, add a suitable multiple of the top row to it so that the entry below the leading entry of the top row becomes zero.

Step 4: Now cover the top row in the augmented matrix and begin again with Step 1 applied to the submatrix that remains. Continue this way until the entire matrix is in row-echelon form.

The (augmented) matrix is now in row-echelon form. The result of step 1 to 4 reduces to (augmented) matrix to a row-echelon form. The process up to step 4 is called **Gaussian Elimination**.

Gauss-Jordan elimination

If we continue to perform the next 2 steps, the entire process is called **Gauss-Jordan Elimination**.

Step 5: Multiply a suitable constant to each row so that all the leading entries become 1.

Step 6: Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading entries.

The (augmented) matrix is now in reduced row-echelon form.

1.5 More on Linear Systems

Summary to solving linear systems

1. Write the linear system in its standard form.
2. Form the augmented matrix of the linear system.
3. Reduce the augmented matrix to either a row-echelon form or reduced row echelon form. May use Gaussian/Gauss-Jordan elimination.
4. Check if the system is consistent
 - If the last column is a pivot column, the system is inconsistent.
 - Otherwise, the system is consistent. If there are any non-pivot columns in the left hand side of the augmented matrix, assign the corresponding variables as parameters, s, t , or s_1, s_2, \dots, s_k .
5. If the system is in reduced row-echelon form, read off the solutions directly.
6. If the system is in row-echelon form only, do back substitution, starting from the lowest nonzero row.
7. Write down the (general) solution to the system.

Chapter 2: Matrix Algebra

2.1 Definition and Special types of Matrices

Definition

A (real-valued) **matrix** is a rectangular array of (real) numbers

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = (a_{ij})_{m \times n} = (a_{ij})_{i=1}^m{}_{j=1}^n$$

where $a_{ij} \in \mathbb{R}$ are real numbers. The size of the matrix is said to be $m \times n$ (read as m by n), where m is the number of rows and n is the number of columns.

The numbers in the array are called **entries**. The (i, j) -entry, a_{ij} , $i = 1, \dots, m$, $j = 1, \dots, n$, is the number in the i -th row j -th column.

Definition

(Special Types of Matrices)

Vectors: A $n \times 1$ matrix is called a (column) **vector**, and a $1 \times n$ matrix is called a (row) **vector**.

Zero matrices: All entries equal 0, denoted as $\mathbf{0}_{m \times n}$. **Not necessarily a square matrix.**

Square matrices: Number of rows = number of columns.

$$\mathbf{A} = (a_{ij})_n = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

A size $n \times n$ matrix is a square matrix of **order** n .

The entries a_{ii} , $i = 1, 2, \dots, n$, (explicitly, a_{11} , a_{22} , \dots , a_{nn}) are called the **diagonal entries** of the (square) **matrix**.

Diagonal matrix: $\mathbf{D} = (a_{ij})_n$, $a_{ij} = 0$ for $i \neq j$. Denote as $\mathbf{D} =$

$$\text{diag}(d_1, d_2, \dots, d_n) = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

Scalar matrix: $\mathbf{C} = (a_{ij})$, $a_{ij} = \begin{cases} c & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$, $\mathbf{C} =$

$$\text{diag}(c, c, \dots, c) = \begin{pmatrix} c & 0 & \dots & 0 \\ 0 & c & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c \end{pmatrix}$$

Identity matrix: $\mathbf{I} = (a_{ij})$, $a_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$, $\mathbf{I}_n =$

$$\text{diag}(1, 1, \dots, 1) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

A scalar matrix can also be denoted as $\mathbf{C} = c\mathbf{I}$, where \mathbf{I} is the identity matrix.

Upper triangular: $\mathbf{A} = (a_{ij})$, $a_{ij} = 0$ for $i > j$:

$$\begin{pmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & * \end{pmatrix}$$

Strictly upper triangular: $\mathbf{A} = (a_{ij})$, $a_{ij} = 0$ for $i \geq j$:

$$\begin{pmatrix} 0 & * & \dots & * \\ 0 & 0 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Lower triangular: $\mathbf{A} = (a_{ij})$, $a_{ij} = 0$ for $i < j$:

$$\begin{pmatrix} * & 0 & \dots & 0 \\ * & * & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{pmatrix}$$

Strictly lower triangular: $\mathbf{A} = (a_{ij})$, $a_{ij} = 0$ for $i \leq j$:

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ * & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & 0 \end{pmatrix}$$

Symmetric matrices: $\mathbf{A} = (a_{ij})$, $a_{ij} = a_{ji}$:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

2.2 Matrix Algebra

Definition

Scalar multiplication: $c \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{pmatrix}$

Matrix addition:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

Theorem

(Properties of Matrix Addition and Scalar Multiplication)

For matrices $\mathbf{A} = (a_{ij})_{m \times n}$, $\mathbf{B} = (b_{ij})_{m \times n}$, $\mathbf{C} = (c_{ij})_{m \times n}$ and real numbers $a, b \in \mathbb{R}$,

1. (Commutative) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. (Associative) $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
3. (Additive identity) $\mathbf{0}_{m \times n} + \mathbf{A} = \mathbf{A}$
4. (Additive inverse) $(\mathbf{A}) + (-\mathbf{A}) = \mathbf{0}_{m \times n}$
5. (Distributive law) $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$
6. (Scalar addition) $(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$
7. (Associative) $(ab)\mathbf{A} = a(b\mathbf{A})$
8. If $a\mathbf{A} = \mathbf{0}_{m \times n}$, then either $a = 0$ or $\mathbf{A} = \mathbf{0}$.

Definition

Matrix multiplication

$$\mathbf{AB} = \mathbf{A} = (a_{ij})_{m \times p} \mathbf{B} = (b_{ij})_{p \times n} = (\sum_{k=1}^p a_{ik} b_{kj})_{m \times n}$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mp} \end{pmatrix} =$$

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1p}b_{p1} & a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1p}b_{p2} & \cdots & a_{11}b_{1n} + a_{12}b_{2n} + \cdots + a_{1p}b_{pn} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2p}b_{p1} & a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2p}b_{p2} & \cdots & a_{21}b_{1n} + a_{22}b_{2n} + \cdots + a_{2p}b_{pn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mp}b_{p1} & a_{m1}b_{12} + a_{m2}b_{22} + \cdots + a_{mp}b_{p2} & \cdots & a_{m1}b_{1n} + a_{m2}b_{2n} + \cdots + a_{mp}b_{pn} \end{pmatrix}$$

Caution: Matrix multiplication is not commutative $\mathbf{AB} \neq \mathbf{BA}$ in general.

Definition

If we multiply \mathbf{A} to the left of \mathbf{B} , we are **pre-multiplying** \mathbf{A} to \mathbf{B} .

If we multiply \mathbf{A} to the right of \mathbf{B} , we are **post-multiplying** \mathbf{A} to \mathbf{B} .

Theorem

(Properties of Matrix Multiplication)

1. (Associative) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
2. (Left distributive law) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
3. (Right distributive law) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
4. (Commute with scalar multiplication) $c(\mathbf{AB}) = (c\mathbf{A})\mathbf{B} = \mathbf{A}(c\mathbf{B})$
5. (Multiplicative identity) For any $m \times n$ matrix \mathbf{A} , $\mathbf{I}_n\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_m$
6. (Nonzero Zero divisor) There exists $\mathbf{A} \neq \mathbf{0}_{m \times p}$ and $\mathbf{B} \neq \mathbf{0}_{p \times n}$ such that $\mathbf{AB} = \mathbf{0}_{m \times n}$
7. (Zero matrix) For any $m \times n$ matrix \mathbf{A} , $\mathbf{A}\mathbf{0}_{n \times p} = \mathbf{0}_{m \times p}$ and $\mathbf{0}_{p \times m}\mathbf{A} = \mathbf{0}_{p \times n}$

Definition

Define the power of square matrices inductively as such.

1. $\mathbf{A}^0 = \mathbf{I}$
2. $\mathbf{A}^n = \mathbf{A}\mathbf{A}^{n-1}$, for $n \geq 1$.

2.3 Linear System and Matrix Equation

Definition

(Matrix Equation)

A linear system in standard form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

can be expressed as a **matrix equation**

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \mathbf{Ax} = \mathbf{b}$$

Here $\mathbf{A} = (a_{ij})_{m \times n}$ is called the **coefficient matrix**, $\mathbf{x} = (x_i)_{n \times 1}$ the **variable vector**, and $\mathbf{b} = (b_i)_{m \times 1}$ the **constant vector**.

It can also be expressed as a **vector equation**:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n = \mathbf{b}$$

Here \mathbf{a}_i is called the **coefficient vector** for variable x_i , for $i = 1, \dots, n$.

Properties of Homogeneous Linear System

A homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ is **always consistent**, since the **zero vector** is a solution, $\mathbf{A}\mathbf{0} = \mathbf{0}$.

Definition

The zero vector is called the **trivial solution**. If $x \neq 0$ is a nonzero solution to the homogeneous system, it is called a **nontrivial solution**.

Theorem

A homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ has **infinitely many solutions** if and only if it has a **nontrivial solution**.

Lemma

Let \mathbf{v} be a particular solution $\mathbf{Ax} = \mathbf{b}$, and \mathbf{u} be a particular solution to the homogeneous system $\mathbf{Ax} = \mathbf{0}$ with the same coefficient matrix \mathbf{A} . Then $\mathbf{v} + \mathbf{u}$ is also a solution to $\mathbf{Ax} = \mathbf{b}$.

Lemma

Suppose \mathbf{v}_1 and \mathbf{v}_2 are solutions to the linear system $\mathbf{Ax} = \mathbf{b}$. Then $\mathbf{v}_1 - \mathbf{v}_2$ is a solution to the homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ with the same coefficient matrix.

Definition

A $p \times q$ **submatrix** of an $m \times n$ matrix \mathbf{A} , $p \leq m$, $q \leq n$, is formed by taking a $p \times q$ block of the entries of the matrix \mathbf{A} .

Theorem

(Block Multiplication)

Let \mathbf{A} be an $m \times p$ matrix and \mathbf{B} a $p \times n$ matrix. Let \mathbf{A}_1 be a $(m_2 - m_1 + 1) \times p$ submatrix of \mathbf{A} obtained by taking rows m_1 to m_2 , and \mathbf{b}_1 a $p \times (n_2 - n_1 + 1)$ submatrix of \mathbf{B} obtained by taking columns n_1 to n_2 . Then the product $\mathbf{A}_1 \mathbf{B}_1$ is a $(m_2 - m_1 + 1) \times (n_2 - n_1 + 1)$ submatrix of \mathbf{AB} obtained by taking rows m_1 to m_2 and columns n_1 to n_2 .

In particular, let \mathbf{b}_i be the i -th column of \mathbf{B} . Then

$$\mathbf{AB} = \mathbf{A}(\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_n) = (\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \dots \quad \mathbf{Ab}_n).$$

Also, if \mathbf{a}_i is the i -th row of \mathbf{B} , then

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{pmatrix} \mathbf{B} = \begin{pmatrix} \mathbf{a}_1 \mathbf{B} \\ \mathbf{a}_2 \mathbf{B} \\ \vdots \\ \mathbf{a}_m \mathbf{B} \end{pmatrix}$$

2.4 Inverse of Matrices

Definition

A $n \times n$ square matrix \mathbf{A} is **invertible** if there exists a **square** matrix \mathbf{B} of the same size such that $\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}$.

A matrix is said to be **non-invertible** otherwise.

A non-invertible square matrix is called a **singular matrix**.

Theorem

(Uniqueness of inverse)

If \mathbf{B} and \mathbf{C} are both inverses of a square matrix \mathbf{A} , then $\mathbf{B} = \mathbf{C}$.

Definition

Since the inverse is **unique**, we can denote the **inverse** of an **invertible** matrix \mathbf{A} by \mathbf{A}^{-1} and call it the **inverse** of \mathbf{A} . That is, \mathbf{A} is invertible and \mathbf{A}^{-1} is its (unique) inverse if

$$\mathbf{AA}^{-1} = \mathbf{I}_n = \mathbf{A}^{-1}\mathbf{A}.$$

Theorem

(Inverse of 2 by 2 square matrices)

A 2×2 square matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible if and only if $ad - bc \neq 0$. In this case, the inverse is given by

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

Theorem

(Cancellation Law for Matrices)

Let \mathbf{A} be an **invertible** matrix of order n .

1. (Left cancellation) If \mathbf{B} and \mathbf{C} are $n \times m$ matrices with $\mathbf{AB} = \mathbf{AC}$, then $\mathbf{B} = \mathbf{C}$.
2. (Right cancellation) If \mathbf{B} and \mathbf{C} are $m \times n$ matrices with $\mathbf{BA} = \mathbf{CA}$, then $\mathbf{B} = \mathbf{C}$.

Caution: If $\mathbf{AB} = \mathbf{CA}$, we cannot conclude that $\mathbf{B} = \mathbf{C}$.

Theorem

Suppose \mathbf{A} is an $n \times n$ invertible square matrix. Then for any $n \times 1$ vector \mathbf{b} , $\mathbf{Ax} = \mathbf{b}$ has a unique solution.

Corollary

Suppose \mathbf{A} is **invertible**. Then the **trivial solution** is the **only solution** to the homogeneous system $\mathbf{Ax} = \mathbf{0}$.

Algorithm to compute inverse

Suppose \mathbf{A} is an invertible $n \times n$ matrix. By uniqueness of the inverse, there must be a unique solution to

$$\mathbf{AX} = \mathbf{I}.$$

By block multiplication, we are solving the augmented matrix

$$(\mathbf{A} \mid \mathbf{I}) \xrightarrow{RREF} (\mathbf{I} \mid \mathbf{A}^{-1}).$$

Theorem

(Properties of inverses)

Let \mathbf{A} be an **invertible matrix** of order n .

1. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
2. For any **nonzero** real number $a \in \mathbb{R}$, $(a\mathbf{A})$ is **invertible** with **inverse** $(a\mathbf{A})^{-1} = \frac{1}{a}\mathbf{A}^{-1}$.
3. \mathbf{A}^T is **invertible** with **inverse** $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$. (that is, the inverse of the transpose is the transpose of the inverse).
4. If \mathbf{B} is an **invertible** matrix of order n , then (\mathbf{AB}) is **invertible** with **inverse** $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

By (4): Product of invertible matrices is invertible.

If $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k$ are **invertible** matrices of the same size, then the product $\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k$ is **invertible** with $(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1}\cdots\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}$.

Definition

The **negative power** of an **invertible** matrix is defined to be

$$\mathbf{A}^{-n} = (\mathbf{A}^{-1})^n$$

for any $n > 0$.

2.5 Elementary Matrices

Definition

A square matrix of order n \mathbf{E} is called an **elementary matrix** if it can be obtained from the identity matrix \mathbf{I}_n by performing a **single elementary row operation** where

$$\mathbf{I}_n \xrightarrow{r} \mathbf{E}$$

is an elementary row operation.

The **elementary row operation** performed to obtain \mathbf{E} is said to be the **row operation** corresponding to the elementary matrix.

Theorem

(Elementary matrix and elementary row operation)

Let \mathbf{A} be an $n \times m$ matrix and \mathbf{E} be the **elementary matrix corresponding** to the **elementary row operation** r .

Then the product \mathbf{EA} is the **resultant** of performing the row operation r on \mathbf{A} ,

$$\mathbf{A} \xrightarrow{r} \mathbf{EA}.$$

Suppose now \mathbf{B} is row equivalent to \mathbf{A} ,

$$\mathbf{A} \xrightarrow{r_1} \xrightarrow{r_2} \cdots \xrightarrow{r_k} \mathbf{B}.$$

Let \mathbf{E}_i be the elementary matrix corresponding to the row operation r_i , for $i = 1, 2, \dots, k$. Then

$$\mathbf{B} = \mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A}.$$

Theorem

Two $n \times m$ matrices \mathbf{A} and \mathbf{B} are **row equivalent** if and only if there exists **elementary matrices** $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$ such that $\mathbf{B} = \mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A}$.

Theorem

(Inverse of elementary matrices)

Every elementary matrix \mathbf{E} is **invertible**. The inverse \mathbf{E}^{-1} is the elementary row operation corresponding to the **reverse** of the original corresponding row operation.

1. $\mathbf{I}_n \xrightarrow{R_i + cR_j} \mathbf{E} \xrightarrow{R_i - cR_j} \mathbf{I}_n \implies \mathbf{E} : R_i + cR_j, \mathbf{E}^{-1} : R_i - cR_j.$
2. $\mathbf{I}_n \xrightarrow{R_i \leftrightarrow R_j} \mathbf{E} \xrightarrow{R_i \leftrightarrow R_j} \mathbf{I}_n \implies \mathbf{E} : R_i \leftrightarrow R_j, \mathbf{E}^{-1} : R_i \leftrightarrow R_j.$
3. $\mathbf{I}_n \xrightarrow{cR_i} \mathbf{E} \xrightarrow{\frac{1}{c}R_i} \mathbf{I}_n \implies \mathbf{E} : cR_i, \mathbf{E}^{-1} : \frac{1}{c}R_i.$

2.6 Equivalent Statements for Invertibility

Theorem

If $\mathbf{A} = \mathbf{E}_k \dots \mathbf{E}_2 \mathbf{E}_1$ is a product of elementary matrices, then \mathbf{A} is invertible.

Corollary

If the reduce row-echelon form of \mathbf{A} is the identity matrix, then \mathbf{A} is invertible.

Theorem

A square matrix \mathbf{A} is invertible if and only if the homogeneous system $\mathbf{A}\mathbf{x} = \mathbf{0}$ has only the trivial solution

Theorem

A square matrix \mathbf{A} is invertible if and only if it's reduced row-echelon form is the identity matrix.

Theorem

A square matrix \mathbf{A} is invertible if and only if it is a product of elementary matrices.

Definition

Let \mathbf{A} be a $n \times m$ matrix.

1. A $m \times n$ matrix \mathbf{B} is said to be a left inverse of \mathbf{A} if $\mathbf{BA} = \mathbf{I}_m$, where \mathbf{I}_m is the $m \times m$ identity matrix.
2. A $m \times n$ matrix \mathbf{B} is said to be a right inverse of \mathbf{A} if $\mathbf{AB} = \mathbf{I}_n$, where \mathbf{I}_n is the $n \times n$ identity matrix.

\mathbf{B} is a left inverse of \mathbf{A} if and only if \mathbf{A} is a right inverse of \mathbf{B} .

Theorem

A square matrix \mathbf{A} is invertible if and only if it has a left inverse.

Theorem

A square matrix \mathbf{A} is invertible if and only if it has a right inverse.

Theorem

A square matrix \mathbf{A} is invertible if and only if $\mathbf{Ax} = \mathbf{b}$ has a unique solution for all \mathbf{b} .

Algorithm for Finding Inverse

Let \mathbf{A} be a $n \times n$ matrix.

Step 1: Form the $n \times 2n$ (augmented matrix) $(\mathbf{A} \mid \mathbf{I}_n)$.

Step 2: Reduce the matrix $(\mathbf{A} \mid \mathbf{I}_n) \rightarrow (\mathbf{R} \mid \mathbf{B})$ to its REF or RREF.

Step 3: If RREF $\mathbf{R} \neq \mathbf{I}$ or REF has a zero row, then \mathbf{A} is not invertible. If RREF $\mathbf{R} = \mathbf{I}$ or REF has no zero row, \mathbf{A} is invertible with inverse $\mathbf{A}^{-1} = \mathbf{B}$.

2.7 LU Factorization

Definition

A square matrix \mathbf{L} is a **unit lower triangular** matrix if \mathbf{L} is a lower triangular matrix with 1 in the diagonal entries.

An **LU factorization** of a $m \times n$ matrix \mathbf{A} is the decomposition

$$\mathbf{A} = \mathbf{LU},$$

where \mathbf{L} is a unit lower triangular matrix, and \mathbf{U} is a row-echelon form of \mathbf{A} .

If such LU factorization exists for \mathbf{A} , we say that \mathbf{A} is LU factorizable.

Lemma

(Product of unit lower triangular matrix is unit lower triangular)

Let \mathbf{A} and \mathbf{B} be unit lower triangular matrices of the same size. Then \mathbf{AB} is a unit lower triangular matrix too.

Lemma

If \mathbf{E} is an elementary matrix corresponding to the operation $R_i + cR_j$ for $i > j$ for some real number c , then \mathbf{E} is a lower triangular matrix.

Algorithm to LU factorization

Suppose $\mathbf{A} \xrightarrow{r_1, r_2, \dots, r_k} \mathbf{U}$, where each row operation r_l is of the form $R_i + cR_j$ for some $i > j$ and real number c , and \mathbf{U} is a row-echelon form of \mathbf{A} . Let \mathbf{E}_i be the elementary matrix corresponding for r_i , for $i = 1, 2, \dots, k$. Then

$$\mathbf{E}_k \dots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{U} \implies \mathbf{A} = \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \dots \mathbf{E}_k^{-1} \mathbf{U} = \mathbf{LU}$$

where $\mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\dots\mathbf{E}_k^{-1}$. Then

$$\mathbf{A} = \mathbf{LU} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ * & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & 1 \end{pmatrix} \begin{pmatrix} * & \dots & 0 & * & \dots & * \\ 0 & \dots & 0 & * & \dots & * \\ \vdots & & & & & \vdots \\ 0 & \dots & & & \dots & * \end{pmatrix}$$

Solving Linear System using LU Factorization

Let $\mathbf{A} = \mathbf{LU}$ be a LU-factorization. Consider the linear system $\mathbf{Ax} = \mathbf{b}$.

- Since \mathbf{L} is a unit lower, can solve $\mathbf{Ly} = \mathbf{b}$ by substitution starting from top row.
- Since \mathbf{U} is in row-echelon form, can solve $\mathbf{Ux} = \mathbf{y}$ by back substitution.

Definition

A $n \times n$ matrix \mathbf{P} is a **permutation matrix** if every rows and columns has a 1 in only one entry, and 0 everywhere else. Equivalently, \mathbf{P} is a permutation matrix if and only if \mathbf{P} is the product of elementary matrices corresponding to row swaps.

2.8 Determinant by Cofactor Expansion

Definition

The **determinant** of \mathbf{A} is defined to be

$$\det(\mathbf{A}) = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} = \sum_{k=1}^n a_{ik}A_{ik} \quad (1)$$

$$= a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj} = \sum_{k=1}^n a_{kj}A_{kj} \quad (2)$$

where

$$A_{ij} = (-1)^{i+j}\det(\mathbf{M}_{ij})$$

is the (i, j) -**cofactor** of \mathbf{A} , and \mathbf{M}_{ij} , the (i, j) **matrix minor** of \mathbf{A} , obtained from \mathbf{A} by deleting the i -th row and j -th column.

This is called the **cofactor expansion** along $\begin{cases} \text{row } i & (1) \\ \text{column } j & (2) \end{cases}$.

The determinant of \mathbf{A} is also denoted as $\det(\mathbf{A}) = |\mathbf{A}|$.

Theorem

(Determinant is invariant under transpose)

The determinant of a square matrix \mathbf{A} is equal to the determinant of its transpose,

$$\det(\mathbf{A}) = \det(\mathbf{A}^T).$$

Corollary

The **determinant** of a **triangular matrix** is the **multiplication of the diagonal entries**. That is, if $\mathbf{A} = (a_{ij})_n$ is a **triangular matrix**, then

$$\det(\mathbf{A}) = a_{11}a_{22} = \dots a_{nn} = \prod_{k=1}^n a_{kk}.$$

2.9 Determinant by Reduction

Theorem

Suppose \mathbf{B} is obtained from \mathbf{A} by a **single elementary row operation**, $\mathbf{A} \xrightarrow{r} \mathbf{B}$. Then the **determinant** of \mathbf{B} is obtained from the **determinant** of \mathbf{A} as such.

- If $r = R_i + aR_j$, then $\det(\mathbf{B}) = \det(\mathbf{A})$;
- If $r = cR_i$, then $\det(\mathbf{B}) = c \det(\mathbf{A})$;
- If $r = R_i \leftrightarrow R_j$, then $\det(\mathbf{B}) = -\det(\mathbf{A})$.

Corollary

The determinant of an elementary matrix \mathbf{E} is given as such.

- If \mathbf{E} corresponds to $R_i + aR_j$, then $\det(\mathbf{E}) = 1$.
- If \mathbf{E} corresponds to cR_i , then $\det(\mathbf{E}) = c$.
- If \mathbf{E} corresponds to $R_i \leftrightarrow R_j$, then $\det(\mathbf{E}) = -1$.

Theorem

Let \mathbf{A} and \mathbf{R} be square matrices such that

$$\mathbf{R} = \mathbf{E}_k \dots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A}$$

for some elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$. Then

$$\det(\mathbf{R}) = \det(\mathbf{E}_k) \dots \det(\mathbf{E}_2) \det(\mathbf{E}_1) \det(\mathbf{A}).$$

Corollary

Let \mathbf{A} be a $n \times n$ square matrix.

Suppose $\mathbf{A} \xrightarrow{r_1} \xrightarrow{r_2} \dots \xrightarrow{r_k} \mathbf{R} = \begin{pmatrix} d_1 & * & \dots & * \\ 0 & d_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$, where \mathbf{R} is the reduced row-echelon form of \mathbf{A} . Let \mathbf{E}_1 be the elementary matrix corresponding to the elementary row operation r_i , for $i = 1, \dots, k$. Then

$$\det(\mathbf{A}) = \frac{d_1 d_2 \dots d_n}{\det(\mathbf{E}_k) \dots \det(\mathbf{E}_2) \det(\mathbf{E}_1)}.$$

2.10 Properties of Determinant

Theorem

A square matrix \mathbf{A} is invertible if and only if $\det(\mathbf{A}) \neq 0$.

Theorem

(Determinant of product is the product of determinant)

Let \mathbf{A} and \mathbf{B} be square matrices of the same size. Then

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}).$$

By induction, for square matrices $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k$ of the same size,

$$\det(\mathbf{A}_1\mathbf{A}_2 \dots \mathbf{A}_k) = \det(\mathbf{A}_1)\det(\mathbf{A}_2) \dots \det(\mathbf{A}_k).$$

Theorem

(Determinant of inverse is the inverse of determinant)

If \mathbf{A} is invertible, then

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

Theorem

(Determinant of scalar multiplication)

For any square matrix \mathbf{A} of order n and scalar c ,

$$\det(c\mathbf{A}) = c^n \det(\mathbf{A}).$$

Definition

Let \mathbf{A} be a $n \times n$ square matrix. The **adjoint** of \mathbf{A} , denoted as $\mathbf{adj}(\mathbf{A})$, is the $n \times n$ square matrix whose (i, j) entry is the (j, i) -cofactor of \mathbf{A} ,

$$\mathbf{adj}(\mathbf{A}) = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

Theorem

(Adjoint formula)

Let \mathbf{A} be a square matrix and $\mathbf{adj}(\mathbf{A})$ be its adjoint. Then

$$\mathbf{A}(\mathbf{adj}(\mathbf{A})) = \det(\mathbf{A})\mathbf{I},$$

where \mathbf{I} is the identity matrix.

Corollary

(Adjoint formula for inverse)

Let \mathbf{A} be an **invertible** matrix. Then the **inverse** of \mathbf{A} is given by

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

Chapter 3: Euclidean Vector Spaces

3.1 Euclidian Vector Spaces

Definition

A (real) n -**vector** is a collection of n ordered real numbers,

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \text{ where } v_i \in \mathbb{R} \text{ for } i = 1, \dots, n.$$

The real number v_i is called the i -th coordinate of the vector \mathbf{v} . The **Euclidean n -space**, denoted \mathbb{R}^n , is the collection of all n -vectors

$$\mathbb{R}^n = \left\{ v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \mid v_i \in \mathbb{R} \text{ for } i = 1, \dots, n. \right\}$$

Properties of Vector Addition and Scalar Multiplication

Since vectors are matrices (column vectors are $n \times 1$ matrices and row vectors are $1 \times n$ matrices), the properties of matrix addition and scalar multiplication holds for vectors. For any vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ and scalars $a, b \in \mathbb{R}$,

1. The sum $\mathbf{u} + \mathbf{v}$ is a vector in \mathbb{R}^n
2. (Commutative) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. (Associative) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.
4. (Zero vector) $\mathbf{0} + \mathbf{v} = \mathbf{v}$.
5. The negative $-\mathbf{v}$ is a vector in \mathbb{R}^n such that $\mathbf{v} - \mathbf{v} = \mathbf{0}$.
6. (Scalar multiple) $a\mathbf{v}$ is a vector in \mathbb{R}^n .
7. (Distribution) $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$.
8. (Distribution) $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$.
9. (Associativity of scalar multiplication) $(ab)\mathbf{u} = a(b\mathbf{u})$.
10. If $a\mathbf{u} = \mathbf{0}$, then either $a = 0$ or $\mathbf{u} = \mathbf{0}$.

Definition

A **linear combination** of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k \in \mathbb{R}^n$ is $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_k\mathbf{u}_k$ for some $c_1, c_2, \dots, c_k \in \mathbb{R}$.

Definition

A set V equipped with **addition** and **scalar multiplication** is said to be a **vector space** over \mathbb{R} if it satisfies the following axioms.

1. For any vectors \mathbf{u}, \mathbf{v} in V , the sum $\mathbf{u} + \mathbf{v}$ is in V .

2. (Commutative) For any vectors \mathbf{u}, \mathbf{v} in V , $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. (Associative) For any vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V , $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
4. (Zero vector) There is a vector $\mathbf{0}$ in V such that $\mathbf{0} + \mathbf{v} = \mathbf{v}$ for all vectors \mathbf{v} in V .
5. (Negative) For any vector \mathbf{u} in V , there exists a vector $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. For any scalar a in \mathbb{R} and vector \mathbf{v} in V , $a\mathbf{v}$ is a vector in V .
7. (Distribution) For any scalar a in \mathbb{R} and vector \mathbf{u}, \mathbf{v} in V , $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$.
8. (Distribution) For any scalars a, b in \mathbb{R} and vector \mathbf{u} in V , $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$.
9. (Associativity of scalar multiplication) For any scalars a, b in \mathbb{R} and vector \mathbf{u} in V , $a(b\mathbf{u}) = (ab)\mathbf{u}$.
10. For any vector \mathbf{u} in V , $1\mathbf{u} = \mathbf{u}$.

3.2 Dot Product, Norm, Distance

Definition

The **inner product** (or **dot product**) of vectors $\mathbf{u} = (u_i)$ and $\mathbf{v} = (v_i)$ in \mathbb{R}^n is defined to be

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Define the **norm** of a vector $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} = (u_i)$, to be the square root of the inner product of \mathbf{u} with itself, and is denoted as $\|\mathbf{u}\|$,

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}.$$

This is also known as the **length** or **magnitude** of the vector.

Theorem

(Properties of inner product and norm)

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ be vectors and $a, b, c \in \mathbb{R}$ be real numbers.

1. Inner product is **symmetric**,

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

2. Inner product **commutes** with scalar multiple,

$$c\mathbf{u} \cdot \mathbf{v} = (c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v}).$$

3. Inner product is **distributive**,

$$\mathbf{u} \cdot (a\mathbf{v} + b\mathbf{w}) = a\mathbf{u} \cdot \mathbf{v} + b\mathbf{u} \cdot \mathbf{w}$$

4. Inner product is **positive definite**, $\mathbf{u} \cdot \mathbf{u} \geq 0$ with equality if and only if $\mathbf{u} = \mathbf{0}$.

5. $\|c\mathbf{u}\| = |c| \|\mathbf{u}\|$.

Definition

A vector $\mathbf{u} \in \mathbb{R}^n$ is a **unit vector** if its norm is 1,

$$\|\mathbf{u}\| = 1$$

Normalizing a vector

Let \mathbf{u} be a nonzero vector $\mathbf{u} \neq \mathbf{0}$. By multiplying by the reciprocal of the norm, we get a unit vector,

$$\mathbf{u} \longrightarrow \frac{\mathbf{u}}{\|\mathbf{u}\|}$$

This is called **normalizing** \mathbf{u} .

Definition

The **distance** between two vectors \mathbf{u} and \mathbf{v} , denoted as $d(\mathbf{u}, \mathbf{v})$, is defined to be

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Define the **angle** θ between two **nonzero** vectors, $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$ to be such that

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

3.3 Linear Combinations and Linear Spans

Definition

A linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k \in \mathbb{R}^n$ is

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_k\mathbf{u}_k, \text{ for some } c_1, c_2, \dots, c_k \in \mathbb{R}.$$

The scalars c_1, c_2, \dots, c_k are called **coefficients**.

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ be vectors in \mathbb{R}^n . The **span** (or **linear span**) of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ is the subset of \mathbb{R}^n containing all the linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$,

$$\text{span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\} = \{c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_k\mathbf{u}_k \mid c_1, c_2, \dots, c_k \in \mathbb{R}\}.$$

Algorithm to Check for Linear Combination

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a set of vectors in \mathbb{R}^n .

1. Form the $n \times k$ matrix $\mathbf{A} = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_k)$ whose columns are the vectors in S .
2. Then a vector \mathbf{v} in \mathbb{R}^n is in $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ if and only if the system $\mathbf{A}\mathbf{x} = \mathbf{v}$ is consistent.
3. If the system is consistent, then the solutions to the system are the possible coefficients of the linear

combination. That is, if $\mathbf{u} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$ is a solution to $\mathbf{A}\mathbf{x} = \mathbf{v}$, then

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_k\mathbf{u}_k.$$

Algorithm to Check if $\text{span}(S) = \mathbb{R}^n$

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a set of vectors in \mathbb{R}^n .

1. Form the $n \times k$ matrix $\mathbf{A} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ whose columns are the vectors in S .
2. Then $\text{span}(S) = \mathbb{R}^n$ if and only if the system $\mathbf{A}\mathbf{x} = \mathbf{v}$ is consistent for all \mathbf{v} .
3. This is equivalent to the reduced row-echelon form of \mathbf{A} having no zero rows.

Properties of linear span

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a finite set of vector. The span of S , $\text{span}(S)$ has the following properties.

1. The span of S **contains the origin**,

$$\mathbf{0} \in \text{span}(S).$$

2. The span of S is **closed under vector addition**, for any $\mathbf{u}, \mathbf{v} \in \text{span}(S)$,

$$\mathbf{u} + \mathbf{v} \in \text{span}(S)$$

3. The span S is **closed under scalar multiplication**, for any $\mathbf{u} \in \text{span}(S)$ and real number $\alpha \in \mathbb{R}$,

$$\alpha \mathbf{u} \in \text{span}(S).$$

Properties (ii) and (iii) can be combined together into one property (ii'): The span is **closed under linear combinations**, that is, if \mathbf{u}, \mathbf{v} are vectors in $\text{span}(S)$ and α, β are any scalars, then the linear combination $\alpha \mathbf{u} + \beta \mathbf{v}$ is a vector in $\text{span}(S)$.

Theorem

(Linear span is closed under linear combinations)

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a set of vectors in \mathbb{R}^n . For any vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ in $\text{span}(S)$, the span of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ is a subset of $\text{span}(S)$,

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\} \subseteq \text{span}(S).$$

Algorithm to check for Set Relations between Spans

Suppose we are given 2 sets of vectors $T = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$.

1. By the corollary, if $\mathbf{v}_i \in \text{span}(S)$ for $i = 1, \dots, m$, we can conclude that $\text{span}(T) \subseteq \text{span}(S)$.
2. Recall that to check if $\mathbf{v}_i \in \text{span}(S)$, we check that the system $(\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k \mid \mathbf{v}_i)$ is consistent for all $i = 1, \dots, m$.
3. There are in total m such linear systems to check. However, since they have the same coefficient matrix, we may combine and check them together, that is, check that

$$(\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k \mid \mathbf{v}_1 \mid \mathbf{v}_2 \mid \dots \mid \mathbf{v}_m)$$

is consistent.

Theorem

(Algorithm to check for set relations between spans)

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ and $T = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be sets of vectors in \mathbb{R}^n . Then $\text{span}(T) \subseteq \text{span}(S)$ if and only if $(\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k \mid \mathbf{v}_1 \mid \mathbf{v}_2 \mid \dots \mid \mathbf{v}_m)$ is consistent.

3.4 Subspaces

Definition

The set of solutions to a linear system $\mathbf{Ax} = \mathbf{b}$ can be expressed **implicitly** as

$$V = \{\mathbf{u} \in \mathbb{R}^n \mid \mathbf{Au} = \mathbf{b}\}$$

or **explicitly** as

$$V = \{\mathbf{u} + s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \dots + s_k\mathbf{v}_k \mid s_1, s_2, \dots, s_k \in \mathbb{R}\},$$

where $\mathbf{u} + s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \dots + s_k\mathbf{v}_k$, $s_1, s_2, \dots, s_k \in \mathbb{R}$ is the general solution.

Definition

A subset V of \mathbb{R}^n is a **subspace** if it satisfies the following properties.

1. V contains the zero vector, $\mathbf{0} \in V$.
2. V is **closed under scalar multiplication**. For any vector \mathbf{v} in V and scalar α , the vector $\alpha\mathbf{v}$ is in V .
3. V is **closed under addition**. For any vectors \mathbf{u}, \mathbf{v} in V , the sum $\mathbf{u} + \mathbf{v}$ is in V .

Property (i) can be replaced with property (i'): V is **nonempty**.

Properties (ii) and (iii) is equivalent to property (ii'): V is **closed under linear combination**. For any \mathbf{u}, \mathbf{v} in V , and scalars α, β , the linear combination $\alpha\mathbf{u} + \beta\mathbf{v}$ is in V .

Theorem

(Solution set of a homogeneous system is a subspace)

The solution set $V = \{\mathbf{u} \mid \mathbf{Au} = \mathbf{b}\}$ to a linear system $\mathbf{Ax} = \mathbf{b}$ is a **subspace** if and only if $\mathbf{b} = \mathbf{0}$, that is, the system is **homogeneous**.

Definition

The **solution set** to a **homogeneous system** is call a **solution space**.

Theorem

(Subspaces are equivalent to linear spans)

A subset $V \subseteq \mathbb{R}^n$ is a subspace if and only if it is a linear span, $V = \text{span}(S)$, for some finite set $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$.

Check if a set is a subspace

To show that a set V is a subspace, we can either

- find a spanning set, that is, find a set S such that $V = \text{span}(S)$, or
- show that V satisfies the 3 conditions of being a subspace.

To show that a subset V is not a subspace, we can either

- show that it does not contain the zero vector, $\mathbf{0} \notin V$,
- find a vector $\mathbf{v} \in V$ and a scalar $\alpha \in \mathbb{R}$ such that $\alpha\mathbf{v} \notin V$, or
- find vectors $\mathbf{u}, \mathbf{v} \in V$ such that the sum is not in V , $\mathbf{u} + \mathbf{v} \notin V$.

Theorem

(Affine spaces)

The solution set $W = \{ \mathbf{w} \mid \mathbf{A}\mathbf{w} = \mathbf{b} \}$ of a non-homogeneous linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{b} \neq \mathbf{0}$, is given by

$$\mathbf{u} + V := \{ \mathbf{u} + \mathbf{v} \mid \mathbf{v} \in V \}$$

where $V = \{ \mathbf{v} \mid \mathbf{A}\mathbf{v} = \mathbf{0} \}$ is the solution space to the associated homogeneous system and \mathbf{u} is a particular solution, $\mathbf{A}\mathbf{u} = \mathbf{b}$.

That is, vectors in $\mathbf{u} + V$ are of the form $\mathbf{u} + \mathbf{v}$ for some \mathbf{v} in V .

3.5 Linear Independence

Definition

A set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ is **linearly independent** if the **only coefficients** $c_1, c_2, \dots, c_k \in \mathbb{R}$ satisfying the equation

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_k\mathbf{u}_k = \mathbf{0}$$

are $c_1 = c_2 = \dots = c_k = 0$. Otherwise, we say that the set is **linearly dependent**.

Algorithm to Check for Linear Independence

Let $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a set of vectors in \mathbb{R}^n .

- $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ is **linearly independent** if and only if the **homogeneous system** $(\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k)\mathbf{x} = \mathbf{0}$ has only the **trivial solution**.
- The homogeneous system has only the **trivial solution** if and only if the **reduced row-echelon form** of $(\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k)$ has **no non-pivot column**.

Theorem

(Solution set of a homogeneous system is a subspace)

A subset $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ of \mathbb{R}^n is **linearly independent** if and only if the **reduced row-echelon form** of $\mathbf{A} = (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k)$ has **no non-pivot columns**.

3.6 Basis and Coordinates

Definition

Let V be a subspace of \mathbb{R}^n . A set $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ is a **basis** for V if

- $\text{span}(S) = V$ and
- S is linearly independent.

Theorem

Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ is a basis for V . Then every vector \mathbf{v} in the subspace V **can be written** as a linear combination of vectors in S **uniquely**.

Theorem

(Basis for Solution Set of Homogeneous System)

Let $V = \{\mathbf{u} | \mathbf{A}\mathbf{u} = \mathbf{0}\}$ be the solution space to some homogeneous system. Suppose

$$s_1\mathbf{u}_1 + s_2\mathbf{u}_2 + \dots + s_k\mathbf{u}_k, s_1, s_2, \dots, s_k \in \mathbb{R}$$

is a general solution to the homogeneous system $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Then $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ is a basis for the subspace $V = \{\mathbf{u} | \mathbf{A}\mathbf{u} = \mathbf{0}\}$.

Theorem

Basis for the zero space $\{\mathbf{0}\}$ of \mathbb{R}^n is the **empty set** $\{\}$ or \emptyset .

Theorem

A $n \times n$ square matrix \mathbf{A} is invertible if and only if the columns are linearly independent.

Theorem

A $n \times n$ square matrix \mathbf{A} is invertible if and only if the columns spans \mathbb{R}^n .

Corollary

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ be a subset of \mathbb{R}^n containing n vectors. Then S is linearly independent if and only if S spans \mathbb{R}^n .

Corollary

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a subset of \mathbb{R}^n and $\mathbf{A} = (\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_k)$ be the matrix whose columns are vectors in S . Then S is a **basis** for \mathbb{R}^n if and only if $k = n$ and \mathbf{A} is an **invertible matrix**.

Theorem

A $n \times n$ **square** matrix \mathbf{A} is invertible if and only if the **columns** of \mathbf{A} form a **basis** for \mathbb{R}^n .

Theorem

A $n \times n$ **square** matrix **A** is invertible if and only if the **rows** of **A** form a **basis** for \mathbb{R}^n .

Theorem

(Equivalent Statements for Invertibility)

Let **A** be a square matrix of order n . The following statements are equivalent.

1. **A** is **invertible**.
2. \mathbf{A}^T is **invertible**.
3. **A** has a **left-inverse**, that is, there is a matrix **B** such that $\mathbf{BA} = \mathbf{I}$.
4. **A** has a **right-inverse**, that is, there is a matrix **B** such that $\mathbf{AB} = \mathbf{I}$.
5. The **reduced row-echelon form** of **A** is the **identity matrix**.
6. **A** can be expressed as a **product** of **elementary matrices**.
7. The **homogeneous system** $\mathbf{Ax} = \mathbf{0}$ has **only the trivial solution**.
8. For **any** **b**, the system $\mathbf{Ax} = \mathbf{b}$ is **consistent**.
9. The **determinant** of **A** is **nonzero**, $\det(\mathbf{A}) \neq 0$.
10. The **columns/rows** of **A** are **linearly independent** for \mathbb{R}^n .
11. The **columns/rows** of **A** **spans** \mathbb{R}^n .

Definition

(Coordinates Relative to a Basis)

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a basis for a subspace V of \mathbb{R}^n .
Then given any vector $\mathbf{v} \in V$, we can write \mathbf{v} unique as

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k.$$

The coordinates of \mathbf{v} relative to the basis S is defined to be the vector

$$[\mathbf{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}.$$

Algorithm for Computing Relative Coordinate

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a basis for a subspace V of \mathbb{R}^n .

For $\mathbf{v} \in V$, find real numbers $c_1, c_2, \dots, c_k \in \mathbb{R}$ such that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k = \mathbf{v}.$$

That is, we are solving for

$$(\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k \mid \mathbf{v}).$$

Theorem

Let V be a subspace of \mathbb{R}^n and B a basis for V .

1. For any vectors $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} = \mathbf{v}$ if and only if $[\mathbf{u}]_B = [\mathbf{v}]_B$.
2. For any $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m \in V$,

$$[c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_m\mathbf{v}_m]_B = c_1[\mathbf{v}_1]_B + c_2[\mathbf{v}_2]_B + \dots + c_m[\mathbf{v}_m]_B.$$

3.7 Dimensions

Theorem

Let V be a subspace of \mathbb{R}^n and B a basis for V . Suppose B contains k vectors, $|B| = k$. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ be vectors in V . Then

1. $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ is linearly independent (respectively, dependent) if and only if $[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, \dots, [\mathbf{v}_m]_B$ is linearly independent (respectively, dependent) in \mathbb{R}^k ; and
2. $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ spans V if and only if $[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, \dots, [\mathbf{v}_m]_B$ spans \mathbb{R}^k .

Corollary

Let V be a subspace of \mathbb{R}^n and B a basis for V . Suppose B contains k vectors, $|B| = k$.

1. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is a subset of V with $m > k$, then S is **linearly dependent**.
2. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is a subset of V with $m < k$, then S **cannot span** V .

Corollary

Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ and $T = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ are bases for a subspace $V \subseteq \mathbb{R}^n$. Then $k = m$.

Definition

Let V be a subspace of \mathbb{R}^n . The **dimension** of V , denoted by $\dim(V)$, is defined to be the **number of vectors** in any **basis** of V .

Theorem

(Dimension of solution space)

Let \mathbf{A} be a $m \times n$ matrix. The **number of non-pivot columns** in the reduced row-echelon form of \mathbf{A} is the **dimension** of the solution space

$$V = \{\mathbf{u} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{u} = \mathbf{0}\}.$$

Theorem

(Spanning Set Theorem)

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a subset of vectors in \mathbb{R}^n , and let $V = \text{span}(S)$. Suppose V is not the zero space, $V \neq \{\mathbf{0}\}$. Then there must be a subset of S that is a basis for V .

Theorem

(Linear Independence Theorem)

Let V be a subspace of \mathbb{R}^n and $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ a linearly independent subset of V , $S \subseteq V$. Then there must be a set T containing S , $S \subseteq T$ such that T is a basis for V .

Theorem

Let U and V be **subspaces** of \mathbb{R}^n .

1. If $U \subseteq V$, then $\dim(U) \leq \dim(V)$.
2. If $U \subseteq V$, and $U \neq V$, then $\dim(U) < \dim(V)$. That is, $U \subseteq V$, then $\dim(U) \leq \dim(V)$ with **equality** if and only if $U = V$.

Theorem

(B1)

Let V be a k -dimensional subspace of \mathbb{R}^n , $\dim(V) = k$. Suppose $S \subseteq V$ is a **linearly independent** subset containing k vectors, $|S| = k$. Then S is a **basis** for V .

In summary,

1. $|S| = \dim(V)$
2. $S \subseteq V$
3. S is linearly independent

Theorem

(B2)

Let V be a k dimensional subspace of \mathbb{R}^n , $\dim(V) = k$. Suppose S is a set containing k vectors, $|S| = k$, such that $V \subseteq \text{span}(S)$. Then S is a **basis** for V .

In summary,

1. $|S| = \dim(V)$
2. $V \subseteq \text{span}(S)$

3.8 Transition Matrices

Definition

Let V be a subspace of \mathbb{R}^n . Suppose $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and $T = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ are **basis** for the subspace V . Define the **transition matrix** from T to S to be

$$\mathbf{P} = ([\mathbf{v}_1]_S \quad [\mathbf{v}_2]_S \quad \dots \quad [\mathbf{v}_k]_S),$$

the matrix whose columns are the coordinates of the vectors in T relative to the basis S .

Theorem

(Transition Matrix)

Let V be a subspace of \mathbb{R}^n . Suppose $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and $T = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ are **bases** for the subspace V . Let \mathbf{P} be the transition matrix from T to S . Then for any vector w in V ,

$$[\mathbf{w}]_S = \mathbf{P}[\mathbf{w}]_T.$$

Algorithm to find Transition Matrix

Let $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and $T = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be basis for a subspace V in \mathbb{R}^n . To find \mathbf{P} , the transition matrix from T to S ,

$$("S" | "T") = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_k \quad | \quad \mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_k) \xrightarrow{\text{rref}} \left(\begin{array}{c|c} \mathbf{I}_k & \mathbf{P} \\ \hline \mathbf{0}_{(n-k) \times k} & \mathbf{0}_{(n-k) \times k} \end{array} \right)$$

Theorem

(Inverse of Transition Matrix)

Suppose $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and $T = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ are **bases** for a subspace V of \mathbb{R}^n . Let \mathbf{P} be the **transition matrix from T to S** . Then \mathbf{P}^{-1} is the **transition matrix from S to T** .