

Linear Algebra

Azure

2022

Contents

1	Day 1: Vectors & Matrices	3
1.1	Vectors	3
1.2	Matrices	4
1.3	System of Equations	6
1.4	Linear Transformations	10
1.5	Invertibility	12
2	Day 2: Vector Space	17
2.1	Span & Independence	17
2.2	Subspaces: Image & Kernel	19
2.3	Basis & Dimension	20
2.4	Injectivity & Surjectivity	23
2.5	Coordinates	25
3	Day 3: Orthogonality	28
3.1	Orthogonality	28
3.2	Orthogonal Basis	31
3.3	Orthogonal Transformations	36
4	Day 4: Determinants	38
4.1	Determinant Function	38
4.2	Properties of the Determinant	42
4.3	Volume	44
5	Day 5: Eigenvectors	46
5.1	Diagonalization	46
5.2	Characteristic Polynomial	48
5.3	Eigenspaces	49
5.4	Symmetry	51
5.5	Quadratic Forms	53

Vectors & Matrices

1.1 Vectors

Definition 1.1.1: Vector

A **vector** $\mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = (x_1, x_2, \dots, x_n) \quad \text{where each } x_i \in \mathbb{R}$$

Let $\mathbf{0} \in \mathbb{R}^n$ be $\mathbf{0} = \underbrace{(0, \dots, 0)}_n$. Then for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and scalar $c \in \mathbb{R}$, define:

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \quad c\mathbf{x} = (cx_1, cx_2, \dots, cx_n)$$

Also, define the **length** of \mathbf{x} :

$$\|\mathbf{x}\| = |\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}} \in \mathbb{R}$$

A **unit vector** is a vector with length 1.

Theorem 1.1.2: Properties of Vectors

For any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ and $c_1, c_2 \in \mathbb{R}$:

(a) **Commutativity**

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$$

Proof

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) = (y_1 + x_1, y_2 + x_2, \dots, y_n + x_n) = \mathbf{y} + \mathbf{x}$$

(b) **Additive Associativity**

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$$

Proof

$$\begin{aligned} (\mathbf{x} + \mathbf{y}) + \mathbf{z} &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) + \mathbf{z} \\ &= (x_1 + y_1 + z_1, x_2 + y_2 + z_2, \dots, x_n + y_n + z_n) \\ &= \mathbf{x} + (y_1 + z_1, y_2 + z_2, \dots, y_n + z_n) = \mathbf{x} + (\mathbf{y} + \mathbf{z}) \end{aligned}$$

(c) **Additive Identity**

There exists a unique $\mathbf{0}_v \in \mathbb{R}^n$ such that for all $\mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{0}_v + \mathbf{x} = \mathbf{x}$$

Moreover, $\mathbf{0}_v = \mathbf{0}$. This only holds for \mathbb{R}^n and not all vector spaces.

Proof

Suppose there are $\mathbf{0}_{v_1}, \mathbf{0}_{v_2}$ where $\mathbf{0}_{v_1} + \mathbf{x} = \mathbf{x}$ and $\mathbf{0}_{v_2} + \mathbf{x} = \mathbf{x}$ for all \mathbf{x} . Then:
 $\mathbf{0}_{v_1} = \mathbf{0}_{v_2} + \mathbf{0}_{v_1} = \mathbf{0}_{v_1} + \mathbf{0}_{v_2} = \mathbf{0}_{v_2}$
 Thus, $\mathbf{0}_v$ must be unique.
 Since $\mathbf{0} + \mathbf{x} = (0 + x_1, \dots, 0 + x_n) = (x_1, \dots, x_n) = \mathbf{x}$, then $\mathbf{0}_v = \mathbf{0}$.

(d) **Additive Inverse**

For any \mathbf{x} , there exists a unique $-\mathbf{x} \in \mathbb{R}^n$ such that:

$$\mathbf{x} + (-\mathbf{x}) = \mathbf{0}_v$$

Moreover, $-\mathbf{x} = (-1)\mathbf{x}$.

Proof

Suppose there are $(-\mathbf{x})_a, (-\mathbf{x})_b$ where $\mathbf{x} + (-\mathbf{x})_a = \mathbf{0}_v$ and $\mathbf{x} + (-\mathbf{x})_b = \mathbf{0}_v$. Then:
 $(-\mathbf{x})_a = (-\mathbf{x})_a + \mathbf{0}_v = (-\mathbf{x})_a + \mathbf{x} + (-\mathbf{x})_b = \mathbf{0}_v + (-\mathbf{x})_b = (-\mathbf{x})_b$
 Thus, $-\mathbf{x}$ must be unique.
 Since $(-1)\mathbf{x} + \mathbf{x} = (-x_1 + x_1, \dots, -x_n + x_n) = \mathbf{0}$, then $-\mathbf{x} = (-1)\mathbf{x}$.

(e) Distributivity

$$c_1(x + y) = c_1x + c_1y \quad (c_1 + c_2)x = c_1x + c_2x$$

Proof

$$\begin{aligned} c_1(x+y) &= (c_1x_1+c_1y_1, \dots, c_1x_n+c_1y_n) = (c_1x_1, \dots, c_1x_n) + (c_1y_1, \dots, c_1y_n) = c_1x + c_1y \\ (c_1 + c_2)x &= (c_1x_1 + c_2x_1, \dots, c_1x_n + c_2x_n) \\ &= (c_1x_1, \dots, c_1x_n) + (c_2x_1, \dots, c_2x_n) = c_1x + c_2x \end{aligned}$$

(f) Multiplicative Associativity

$$c_1(c_2x) = (c_1c_2)x$$

Proof

$$c_1(c_2)x = (c_1c_2x_1, \dots, c_1c_2x_n) = (c_1c_2)x$$

(g) Multiplicative Identity

$$1x = x$$

Proof

$$1x = (1x_1, \dots, 1x_n) = (x_1, \dots, x_n) = x$$

Theorem 1.1.3: Rescaling to a Unit Vector

Let $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then, $|cx| = |c||x|$.

Thus, $\frac{x}{|x|}$ is a unit vector.

Proof

$$\begin{aligned} |cx| &= \sqrt{(cx_1)^2 + \dots + (cx_n)^2} = \sqrt{c^2(x_1^2 + \dots + x_n^2)} = |c|\sqrt{x_1^2 + \dots + x_n^2} = |c||x| \\ \left|\frac{x}{|x|}\right| &= \frac{1}{|x|}|x| = 1 \end{aligned}$$

1.2 Matrices**Definition 1.2.1: Matrix**

A $m \times n$ matrix $A \in M_{m \times n}(\mathbb{R})$:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad \text{where each } a_{ij} \in \mathbb{R}$$

Let $0 \in M_{m \times n}(\mathbb{R})$: be a $m \times n$ matrix where the value of all entries are 0.

For $A, B \in M_{m \times n}(\mathbb{R})$ and scalar $c \in \mathbb{R}$, define:

$$A+B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix} \quad cA = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}$$

Theorem 1.2.2: Properties of Matrices

For any $A, B, C \in M_{m \times n}(\mathbb{R})$ and $c_1, c_2 \in \mathbb{R}$:

(a) Commutativity

$$A + B = B + A$$

Proof

$$A + B = \begin{bmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{bmatrix} = \begin{bmatrix} b_{11}+a_{11} & \dots & b_{1n}+a_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1}+a_{m1} & \dots & b_{mn}+a_{mn} \end{bmatrix} = B + A$$

(b) Additive Associativity

$$(A + B) + C = A + (B + C)$$

Proof

$$(A + B) + C = \begin{bmatrix} a_{11} + b_{11} + c_{11} & \dots & a_{1n} + b_{1n} + c_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} + c_{m1} & \dots & a_{mn} + b_{mn} + c_{mn} \end{bmatrix} = A + (B + C)$$

(c) Additive Identity

There exists a unique $0_M \in M_{m \times n}(\mathbb{R})$ such that for all $A \in M_{m \times n}(\mathbb{R})$:

$$0_M + A = A$$

Moreover, $0_M = 0$. This only holds for $M_{m \times n}(\mathbb{R})$ and not all vector spaces.

Proof

Suppose there are $0_{M_1}, 0_{M_2}$ where $0_{M_1} + A = A$ and $0_{M_2} + A = A$ for all A . Then:

$$0_{M_1} = 0_{M_2} + 0_{M_1} = 0_{M_1} + 0_{M_2} = 0_{M_2}$$

Thus, 0_M must be unique.

$$\text{Since } 0 + A = \begin{bmatrix} 0 + a_{11} & \dots & 0 + a_{1n} \\ \vdots & \ddots & \vdots \\ 0 + a_{m1} & \dots & 0 + a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = A, \text{ then } 0_M = 0.$$

(d) Additive Inverse

For any A , there exists a unique $-A \in M_{m \times n}(\mathbb{R})$ such that:

$$A + (-A) = 0_M$$

Moreover, $-A = (-1)A$.

Proof

Suppose there are $(-A)_a, (-A)_b$ where $A + (-A)_a = 0_M$ and $A + (-A)_b = 0_M$.

$$(-A)_a = (-A)_a + 0_M = (-A)_a + A + (-A)_b = 0_M + (-A)_b = (-A)_b$$

Thus, $-M$ must be unique.

$$\text{Since } (-1)A + A = \begin{bmatrix} -a_{11} + a_{11} & \dots & -a_{1n} + a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{m1} + a_{m1} & \dots & -a_{mn} + a_{mn} \end{bmatrix} = 0, \text{ then } -A = (-1)A.$$

(e) Distributivity

$$c_1(A + B) = c_1A + c_1B \quad (c_1 + c_2)A = c_1A + c_2A$$

Proof

$$\begin{aligned} c_1(A + B) &= \begin{bmatrix} c_1a_{11} + c_1b_{11} & \dots & c_1a_{1n} + c_1b_{1n} \\ \vdots & \ddots & \vdots \\ c_1a_{m1} + c_1b_{m1} & \dots & c_1a_{mn} + c_1b_{mn} \end{bmatrix} = c_1A + c_1B \\ (c_1 + c_2)A &= \begin{bmatrix} c_1a_{11} + c_2a_{11} & \dots & c_1a_{1n} + c_2a_{1n} \\ \vdots & \ddots & \vdots \\ c_1a_{m1} + c_2a_{m1} & \dots & c_1a_{mn} + c_2a_{mn} \end{bmatrix} = c_1A + c_2A \end{aligned}$$

(f) Multiplicative Associativity

$$c_1(c_2A) = (c_1c_2)A$$

Proof

$$c_1(c_2A) = \begin{bmatrix} c_1c_2a_{11} & \dots & c_1c_2a_{1n} \\ \vdots & \ddots & \vdots \\ c_1c_2a_{m1} & \dots & c_1c_2a_{mn} \end{bmatrix} = (c_1c_2)A$$

(g) **Multiplicative Identity**

$$1A = A$$

Proof

$$1A = \begin{bmatrix} 1a_{11} & \dots & 1a_{1n} \\ \vdots & \ddots & \vdots \\ 1a_{m1} & \dots & 1a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = A$$

1.3 System of Equations

Definition 1.3.1: Elementary Row Operations

There are three types of **elementary row operations**:

- (a) **Row Multiplication**: Multiplying a row by a nonzero scalar
- (b) **Row Addition**: Add a multiple of a row to another row
- (c) **Row Swapping**: Swapping two rows

Note that for any elementary row operation, an entry in the i -th column can be affected by another entry in the i -th column.

If for matrix $A, B \in M_{m \times n}(\mathbb{R})$, there is a sequence of elementary row operations that transforms A to B , then A and B are **row equivalent**.

Note if there is a sequence that transforms A to B , then performing the sequence in reverse will transform B to A so row equivalence between A and B is the same as row equivalence between B and A .

Definition 1.3.2: Reduced Row-Echelon Form: RREF

The **reduced row-echelon form (rref)** of matrix $A \in M_{m \times n}(\mathbb{R})$, $\text{rref}(A)$ satisfies:

- (a) If a row has nonzero entries, the first nonzero is 1
- (b) If a row has a leading 1, then each row before it has a leading 1
- (c) A column with a leading 1 has 0 for the other entries

For example:

$$\begin{bmatrix} \textcircled{1} & 2 & 0 & 1 \\ 0 & 0 & \textcircled{1} & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Definition 1.3.3: System of Equations: Augmented Matrix

A $m \times n$ system of equations written as a $m \times (n+1)$ **augmented matrix** $A \in M_{m \times (n+1)}(\mathbb{R})$:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \Leftrightarrow A = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right]$$

If the i -th column of the $\text{rref}(A)$ has a leading 1, then x_i is called a **pivot variable** else it is called a **free variable**. The **rank** of a matrix is equal to the number of pivot variables.

For example:

$$\left[\begin{array}{ccc|c} \textcircled{1} & 2 & 0 & 1 \\ 0 & 0 & \textcircled{1} & -1 \\ 0 & 0 & 0 & 0 \end{array} \right] \Leftrightarrow \begin{array}{l} x_1 + 2x_2 = 1 \\ x_3 = -1 \end{array} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_2 + 1 \\ x_2 \\ -1 \end{bmatrix}$$

Note a pivot variable, $\{x_1, x_3\}$, has a fixed value or its value depends on the free variables while the free variables, $\{x_2\}$, can be any value.

Theorem 1.3.4: Gauss-Jordan Elimination: Elementary row operations don't change solutions

Let $m \times n$ system of equations be the augmented matrix $A \in M_{m \times (n+1)}(\mathbb{R})$. By performing elementary row operations on A to get to $\text{rref}(A)$, the solutions are unchanged.

Proof

Suppose the i -th row is multiplied by scalar c .

$$\begin{bmatrix} a_{11} & \dots & a_{1n} & | & b_1 \\ \vdots & \ddots & \vdots & | & \vdots \\ ca_{i1} & \dots & ca_{in} & | & cb_i \\ \vdots & \ddots & \vdots & | & \vdots \\ a_{m1} & \dots & a_{mn} & | & b_m \end{bmatrix} \Leftrightarrow \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \dots \\ ca_{i1}x_1 + \dots + ca_{in}x_n = cb_i \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$

If (x_1^*, \dots, x_n^*) is a solution, then $a_{i1}x_1^* + \dots + a_{in}x_n^* = b_i$ for any $i \in \{1, \dots, m\}$:

$$ca_{i1}x_1^* + \dots + ca_{in}x_n^* = c(a_{i1}x_1^* + \dots + a_{in}x_n^*) = cb_i$$

If (x_1', \dots, x_n') is not a solution, then $a_{i1}x_1' + \dots + a_{in}x_n' \neq b_i$ for any $i \in \{1, \dots, m\}$:

$$ca_{i1}x_1' + \dots + ca_{in}x_n' = c(a_{i1}x_1' + \dots + a_{in}x_n') \neq cb_i$$

Thus, row multiplication does not change the solutions. Note c is nonzero since if $c = 0$, then any (x_1, \dots, x_n) satisfies $ca_{i1}x_1 + ca_{i2}x_2 + \dots + ca_{in}x_n = 0 = cb_i$ which includes non-solutions.

Suppose the i -th row multiplied by c is added to the j -th row.

$$\begin{bmatrix} a_{11} & \dots & a_{1n} & | & b_1 \\ \vdots & \ddots & \vdots & | & \vdots \\ ca_{i1} + a_{j1} & \dots & ca_{in} + a_{jn} & | & cb_i + b_j \\ \vdots & \ddots & \vdots & | & \vdots \\ a_{m1} & \dots & a_{mn} & | & b_m \end{bmatrix} \Leftrightarrow \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \dots \\ (ca_{i1} + a_{j1})x_1 + \dots + (ca_{in} + a_{jn})x_n = cb_i + b_j \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$

If (x_1^*, \dots, x_n^*) is a solution, then $a_{i1}x_1^* + \dots + a_{in}x_n^* = b_i$ for any $i \in \{1, \dots, m\}$:

$$(ca_{i1} + a_{j1})x_1^* + \dots + (ca_{in} + a_{jn})x_n^* = c(a_{i1}x_1^* + \dots + a_{in}x_n^*) + (a_{j1}x_1^* + \dots + a_{jn}x_n^*) = cb_i + b_j$$

If (x_1', \dots, x_n') is not a solution, then $a_{i1}x_1' + \dots + a_{in}x_n' \neq b_i$ for any $i \in \{1, \dots, m\}$:

$$(ca_{i1} + a_{j1})x_1' + \dots + (ca_{in} + a_{jn})x_n' = c(a_{i1}x_1' + \dots + a_{in}x_n') + (a_{j1}x_1' + \dots + a_{jn}x_n') \neq cb_i + b_j$$

Thus, row addition does not change the solutions.

Suppose the i -th row is swapped with the j -th row. Note row swapping is the same as:

$$\text{Add } i\text{-th row to } j\text{-th} \Rightarrow j'\text{-th} = i\text{-th} + j\text{-th}$$

$$\text{Subtract } i\text{-th row by } j'\text{-th} \Rightarrow i'\text{-th} = -j\text{-th}$$

$$\text{Add } i'\text{-th row to } j'\text{-th. Multiply the } i'\text{-th row by } -1 \Rightarrow j''\text{-th} = i\text{-th} \quad i''\text{-th} = j\text{-th}$$

Since each step does not change solutions, then row swapping does not change solutions.

Theorem 1.3.5: Row equivalent matrices have the same solutions

If $A, B \in M_{m \times n}(\mathbb{R})$ are row equivalent, then $Ax = 0$ and $Bx = 0$ have the same solutions

Proof

If A and B are row equivalent, then the augmented matrices $[A \mid 0], [B \mid 0] \in M_{m \times (n+1)}(\mathbb{R})$ are row equivalent. Then, there is a sequence of elementary row operations that transforms $[A \mid 0]$ to $[B \mid 0]$. By [theorem 1.3.4](#), the solutions to $[A \mid 0]$ don't change when transforming to $[B \mid 0]$ so the solutions to $[A \mid 0]$ and $[B \mid 0]$ are the same.

Note $Ax, Bx = 0$ since if $Ax = b$ where $b \neq 0$ and $Ax = c$ where $c \neq 0$, then performing elementary row operations to change A to B might not change b to c . But if $b = 0$, then any elementary row operation will keep b as 0 since the entries in b can only be affected by other entries in b which are all 0. Thus, if also $c = 0$, then $[A \mid 0]$ and $[B \mid 0]$ will be row equivalent.

Theorem 1.3.6: The $\text{rref}(A)$ is unique

Let matrix $A \in M_{m \times n}(\mathbb{R})$ be row equivalent to matrix $B, C \in M_{m \times n}(\mathbb{R})$ which are in reduced row-echelon form. Then, $B = C$.

Proof

Since A is row equivalent to B, C , then by [theorem 1.3.5](#), $Ax = 0$ and $Bx = 0$ have the same solutions and $Ax = 0$ and $Cx = 0$ have the same solutions. Thus, $Bx = 0$ and $Cx = 0$ have the same solutions. The following proof will be a proof by induction.

Suppose $A, B, C \in M_{m \times 1}(\mathbb{R})$. Since B, C are in reduced row-echelon form, then B, C are either:

$$M_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Since $M_1x = 0$ is only $x = 0$ and $M_2x = 0$ is any $x \in \mathbb{R}$ then either $B, C = M_1$ or $B, C = M_2$ since $Bx = 0$ and $Cx = 0$ has the same solutions. Thus, the base case holds true.

Suppose for some $n \in \mathbb{Z}_+$, any matrix $M \in M_{m \times n}(\mathbb{R})$ in reduced row echelon form is unique. Let $A, B, C \in M_{m \times (n+1)}(\mathbb{R})$ where A is row equivalent to B, C in reduced row echelon form. Let $A_n, B_n, C_n \in M_{m \times n}(\mathbb{R})$ be A, B, C without their $(n+1)$ -th column. Since A is row equivalent to B, C , then A_n is row equivalent to B_n, C_n which are also in reduced row-echelon form since removing the last column of any rref is still a rref. Since $A_n \in M_{m \times n}(\mathbb{R})$ which is row equivalent to reduced row echelon matrices $B_n, C_n \in M_{m \times n}(\mathbb{R})$, then $B_n = \text{rref}(A_n) = C_n$. Thus, the first n columns of B, C are the same. Suppose the $B \neq C$ so only the $(n+1)$ -th column can be different. Then there is a $i \in \{1, \dots, m\}$ where $b_{i(n+1)} \neq c_{i(n+1)}$. Let $(x_1^*, \dots, x_{n+1}^*)$ be a solution.

$$b_{i1}x_1 + \dots + b_{in}x_n + b_{i(n+1)}x_{n+1} = 0 \quad c_{i1}x_1 + \dots + c_{in}x_n + c_{i(n+1)}x_{n+1} = 0$$

Since $B_n = C_n$, then $b_{ij} = c_{ij}$ for $i = \{1, \dots, m\}$ and $j = \{1, \dots, n\}$. Thus, $b_{i(n+1)}x_{n+1} = c_{i(n+1)}x_{n+1}$. Since $b_{i(n+1)} \neq c_{i(n+1)}$, then $x_{n+1} = 0$. Thus, x_{n+1} is a pivot variable so the $(n+1)$ -th column of B, C have a leading 1. Thus, any other entry in the $(n+1)$ -th column is 0. Since B and C are in reduced row-echelon form, then this pivot is right after any other pivots before it since this 1 is in the final column, but since the entries of B_n, C_n are the same, then this 1 is in the same row in C as in B . Thus, the $(n+1)$ -th column of B and C are the same which contradicts the assumption that $(n+1)$ -th column are different. Thus, $B = C$ by induction.

Theorem 1.3.7: Number of solutions in a System of Equations

Let $m \times n$ system of equations be the augmented matrix $A \in M_{m \times (n+1)}(\mathbb{R})$. The system is called [consistent](#) if there is at least one solution and [inconsistent](#) if there are no solutions.

If the $\text{rref}(A)$ contains the row $[0 \dots 0 \mid 1]$, then the system has no solutions.

If there is at least one free variable, then there are infinitely many solutions and if all variables are pivots, then there is one solution.

Proof

Since a variable is either a pivot or free variable, then the $\text{rref}(A)$ either:

- contains the row $[0 \dots 0 \mid 1]$
- doesn't contains the row $[0 \dots 0 \mid 1]$ and have all pivot variables
- doesn't contains the row $[0 \dots 0 \mid 1]$, but have all pivot variables

Since $[0 \dots 0 \mid 1]$ implies $0 = 0x_1 + \dots + 0x_n = 1$, then if $\text{rref}(A)$ contains the row $[0 \dots 0 \mid 1]$, there cannot be any solution regardless of pivot and free variables since no $x = (x_1, \dots, x_n)$ will satisfy such a row. Now, suppose $\text{rref}(A)$ doesn't contains the row $[0 \dots 0 \mid 1]$.

Suppose the $\text{rref}(A)$ have all pivot variables. Since pivot variables are fixed or depend on free variables which don't exist, then the pivot variables are all fixed and thus, unique.

Suppose the $\text{rref}(A)$ has at least one free variable. Then at least one variable can be any real number and thus, there are infinitely many solutions.

Corollary 1.3.8: A unique solution must have as many equation as there are unknowns

Let $m \times n$ system of equations be the augmented matrix $A \in M_{m \times (n+1)}(\mathbb{R})$.

If there is a unique solution, then $m \geq n$.

Proof

By **theorem 1.3.7**, a unique solution must have all pivot variables. If $\text{rref}(A)$ has all pivots, then $m \geq n$ else there will be a column without a pivot.

Definition 1.3.9: Homogeneous & Inhomogeneous Equations

A $m \times n$ system of equations:

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

...

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

can also be written in **matrix form**:

$$Ax = b$$

$$\text{where } A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in M_{m \times n}(\mathbb{R}), \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$$

(a) **Closed under Addition**: For any $x_1, \dots, x_k \in \mathbb{R}^n$:

$$A(x_1 + \dots + x_k) = Ax_1 + \dots + Ax_k$$

(b) **Closed under Scalar Multiplication** For any $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$:

$$A(cx) = cAx$$

For $x_1, \dots, x_k \in \mathbb{R}^n$ and $c_1, \dots, c_k \in \mathbb{R}$:

$$A(c_1x_1 + \dots + c_kx_k)$$

$$\begin{aligned} & a_{11}(c_1x_{11} + \dots + c_kx_{k1}) + a_{12}(c_1x_{12} + \dots + c_kx_{k2}) + \dots + a_{1n}(c_1x_{1n} + \dots + c_kx_{kn}) \\ \Leftrightarrow & a_{21}(c_1x_{11} + \dots + c_kx_{k1}) + a_{22}(c_1x_{12} + \dots + c_kx_{k2}) + \dots + a_{2n}(c_1x_{1n} + \dots + c_kx_{kn}) \\ & \dots \\ & a_{m1}(c_1x_{11} + \dots + c_kx_{k1}) + a_{m2}(c_1x_{12} + \dots + c_kx_{k2}) + \dots + a_{mn}(c_1x_{1n} + \dots + c_kx_{kn}) \\ \hline & c_1(a_{11}x_{11} + \dots + a_{1n}x_{1n}) \quad c_k(a_{11}x_{k1} + \dots + a_{1n}x_{kn}) \\ = & c_1(a_{21}x_{11} + \dots + a_{2n}x_{1n}) \quad + \dots + \quad c_k(a_{21}x_{k1} + \dots + a_{2n}x_{kn}) \quad \Leftrightarrow c_1Ax_1 + \dots + c_kAx_k \\ & \dots \quad \dots \\ & c_1(a_{m1}x_{11} + \dots + a_{mn}x_{1n}) \quad c_k(a_{m1}x_{k1} + \dots + a_{mn}x_{kn}) \end{aligned}$$

A **Homogeneous equation** is in the form:

$$Ax = 0 \quad \text{where } A \in M_{m \times n}(\mathbb{R}), x \in \mathbb{R}^n, \text{ and } 0 \in \mathbb{R}^m$$

A **Inhomogeneous equation** is in the form:

$$Ax = b \quad \text{where } A \in M_{m \times n}(\mathbb{R}), x \in \mathbb{R}^n, \text{ and } b \neq 0 \in \mathbb{R}^m$$

Theorem 1.3.10: Relationship between Homogeneous and Inhomogeneous

Let x_0 be a solution to $Ax = b$. Then any solution x_s of $Ax = b$:

$$x_s = x_0 + x^*$$

where x^* is a solution to $Ax = 0$

Proof

Let x_0 be a solution to $Ax = b$. Suppose x_s be a solution to $Ax = b$.

$$b = Ax_s = A(x_0 + (x_s - x_0)) = Ax_0 + A(x_s - x_0) = b + A(x_s - x_0)$$

Thus, $A(x_s - x_0) = 0$ so $x^* = x_s - x_0$ is a solution to $Ax = 0$.

Example

Find the solution(s), (x_1, x_2, x_3, x_4) :

$$x_1 + 2x_3 + 4x_4 = -8$$

$$x_2 - 3x_3 - x_4 = 6$$

$$3x_1 + 4x_2 - 6x_3 + 8x_4 = 0$$

$$-x_2 + 3x_3 + 4x_4 = -12$$

What are the solutions if instead the equations are equal to 0?

$\begin{bmatrix} 1 & 0 & 2 & 4 & & -8 \\ 0 & 1 & -3 & -1 & & 6 \\ 3 & 4 & -6 & 8 & & 0 \\ 0 & -1 & 3 & 4 & & -12 \end{bmatrix}$	$\xRightarrow{\text{Add } -3(1\text{st}) \text{ to the (3rd)}}$	$\begin{bmatrix} 1 & 0 & 2 & 4 & & -8 \\ 0 & 1 & -3 & -1 & & 6 \\ 0 & 4 & -12 & -4 & & 24 \\ 0 & -1 & 3 & 4 & & -12 \end{bmatrix}$	
$\xRightarrow{\begin{array}{l} \text{Add } -4(2\text{nd}) \text{ to the (3rd)} \\ \text{Add (2nd) to the (4th)} \end{array}}$	$\begin{bmatrix} 1 & 0 & 2 & 4 & & -8 \\ 0 & 1 & -3 & -1 & & 6 \\ 0 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 3 & & -6 \end{bmatrix}$	$\xRightarrow{\begin{array}{l} \text{Multiply (4th) by } \frac{1}{3} \\ \text{Add } -4(4\text{th}) \text{ to (1st)} \\ \text{Add (4th) to (2nd)} \end{array}}$	$\begin{bmatrix} 1 & 0 & 2 & 0 & & 0 \\ 0 & 1 & -3 & 0 & & 4 \\ 0 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 1 & & -2 \end{bmatrix}$

Thus, the reduced row-echelon form of this matrix:

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

The pivot variables are x_1, x_2, x_4 and the free variable is x_3 so the rank is 3.

The solutions to the system of equations are:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2x_3 \\ 3x_3 + 4 \\ x_3 \\ -2 \end{bmatrix} = \begin{bmatrix} -2x_3 \\ 3x_3 \\ x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \\ 0 \\ -2 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ 3 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \\ 0 \\ -2 \end{bmatrix} = x_3(-2, 3, 1, 0) + (0, 4, 0, -2)$$

Thus, $x_3(-2, 3, 1, 0)$ are the solutions when the equations equal to 0.

1.4 Linear Transformations

Definition 1.4.1: Linear Transformation

Function $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a **linear transformation** if for any $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$:

(a) **Closed under Addition**: $T(x+y) = T(x) + T(y)$

(b) **Closed under Scalar Multiplication**: $T(cx) = cT(x)$

Note for any $x_1, \dots, x_k \in \mathbb{R}^n$ and $c_1, \dots, c_k \in \mathbb{R}$:

$$\begin{aligned} T(c_1x_1 + \dots + c_kx_k) &= T(c_1x_1) + T(c_2x_2 + \dots + c_kx_k) \\ &= c_1T(x_1) + T(c_2x_2) + T(c_3x_3 + \dots + c_kx_k) \\ &= c_1T(x_1) + c_2T(x_2) + T(c_3x_3) + T(c_4x_4 + \dots + c_kx_k) \\ &= \dots = c_1T(x_1) + \dots + c_kT(x_k) \end{aligned}$$

The **standard vectors** of \mathbb{R}^n are $e_1, \dots, e_n \in \mathbb{R}^n$ where:

$$e_i = (0, \dots, \underset{1}{0}, \dots, \underset{i-1}{0}, \underset{i}{1}, \underset{i+1}{0}, \dots, \underset{n}{0})$$

Theorem 1.4.2: $T(0) = 0$

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Then for $0 \in \mathbb{R}^n$:

$$T(0) = 0 \in \mathbb{R}^m$$

Proof

Let 0_n be the zero vector in \mathbb{R}^n and 0_m be the zero vector in \mathbb{R}^m . Since $0(0_n) = 0_n$, then:

$$T(0_n) = T(0(0_n)) = 0T(0_n) = 0_m$$

Theorem 1.4.3: Every Linear transformation is a Matrix transformation

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Then for any $x \in \mathbb{R}^n$:

$$T(x) = Ax$$

with $A \in M_{m \times n}(\mathbb{R})$ where $A = \begin{bmatrix} T(e_1) & T(e_2) & \dots & T(e_n) \end{bmatrix}$ is called the **standard matrix**

Proof

Since any $x \in \mathbb{R}^n$ is $x = (x_1, \dots, x_n) = x_1e_1 + \dots + x_ne_n$, then:

$$T(x) = T(x_1e_1 + \dots + x_ne_n) = x_1T(e_1) + \dots + x_nT(e_n)$$

$$\Leftrightarrow \begin{bmatrix} T(e_1) & T(e_2) & \dots & T(e_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax$$

Since $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and each $e_i \in \mathbb{R}^n$, then each $T(e_i) \in \mathbb{R}^m$. Thus, $A \in M_{m \times n}(\mathbb{R})$.

Corollary 1.4.4: Linear Transformation: Scaling

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation where for any $x \in \mathbb{R}^n$, the $T(x) = cx$ for some $c \in \mathbb{R}$. Then:

$$T(x) = \begin{bmatrix} c & 0 & \dots & 0 \\ 0 & c & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c \end{bmatrix} x$$

Proof

By **theorem 1.4.3**, $T(x) = \begin{bmatrix} T(e_1) & T(e_2) & \dots & T(e_n) \end{bmatrix} x$. Since $T(e_1) = ce_1$, then:

$$\begin{bmatrix} T(e_1) & T(e_2) & \dots & T(e_n) \end{bmatrix} = \begin{bmatrix} ce_1 & ce_2 & \dots & ce_n \end{bmatrix} = \begin{bmatrix} c & 0 & \dots & 0 \\ 0 & c & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c \end{bmatrix}$$

Example

Let $T: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ scales all $x \in \mathbb{R}^4$ by 2. Find T . Verify $T((-1, 2, 1, 3)) = (-2, 4, 2, 6)$

$$T(x) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} x \quad T\left(\begin{bmatrix} -1 \\ 2 \\ 1 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 2 * (-1) + 0 * 2 + 0 * 1 + 0 * 3 \\ 0 * (-1) + 2 * 2 + 0 * 1 + 0 * 3 \\ 0 * (-1) + 0 * 2 + 2 * 1 + 0 * 3 \\ 0 * (-1) + 0 * 2 + 0 * 1 + 2 * 3 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \\ 2 \\ 6 \end{bmatrix}$$

Corollary 1.4.5: Linear Transformation: 2D Rotation

Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear transformation. Suppose for any $x \in \mathbb{R}^n$, the T rotates x counterclockwise by an angle of θ . Then:

$$T(x) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} x$$

Proof

By **theorem 1.4.3**, $T(x) = [T(e_1) \ T(e_2)]x$. Since T rotates any $x \in \mathbb{R}^2$ counterclockwise by an angle of θ , then:

$$T(e_1) = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \quad T(e_2) = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$$

Example

Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ rotate all $x \in \mathbb{R}^2$ by $\frac{\pi}{6}$ radians = 30° degree counterclockwise. Find T . Find $\cos(75^\circ)$ and $\sin(75^\circ)$.

$$T(x) = \begin{bmatrix} \cos(\frac{\pi}{6}) & -\sin(\frac{\pi}{6}) \\ \sin(\frac{\pi}{6}) & \cos(\frac{\pi}{6}) \end{bmatrix} x = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} x$$

Note $75^\circ = 45^\circ + 30^\circ$ so apply T on the unit vector at 45° which is $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$

$$T\left(\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}\right) = \begin{bmatrix} \frac{\sqrt{3}}{2} \frac{\sqrt{2}}{2} + -\frac{1}{2} \frac{\sqrt{2}}{2} \\ \frac{1}{2} \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{2} \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}-\sqrt{2}}{4} \\ \frac{\sqrt{6}+\sqrt{2}}{4} \end{bmatrix} = \begin{bmatrix} \cos(75^\circ) \\ \sin(75^\circ) \end{bmatrix}$$

1.5 Invertibility**Definition 1.5.1: Product of Linear transformations**

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $S: \mathbb{R}^m \rightarrow \mathbb{R}^t$ be linear transformations. Then for any $x \in \mathbb{R}^n$:

$$(ST)x = S(Tx)$$

is a linear transformation where $ST: \mathbb{R}^n \rightarrow \mathbb{R}^t$

Let $x_1, x_2 \in \mathbb{R}^n$ and $c_1, c_2 \in \mathbb{R}$.

$$\begin{aligned} (ST)(c_1x_1 + c_2x_2) &= S(T(c_1x_1 + c_2x_2)) = S(c_1T(x_1) + c_2T(x_2)) \\ &= c_1S(T(x_1)) + c_2S(T(x_2)) = c_1(ST)(x_1) + c_2(ST)(x_2) \end{aligned}$$

Theorem 1.5.2: Product of Linear transformations are Matrix transformations

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $T(x) = Ax$ and $S: \mathbb{R}^m \rightarrow \mathbb{R}^t$ with $S(y) = By$ be linear transformations

where $A = \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix}$ for $A_i \in \mathbb{R}^m$. Then for any $x \in \mathbb{R}^n$:

$$(ST)x = (BA)x$$

$$\text{where } BA = \begin{bmatrix} BA_1 & BA_2 & \dots & BA_n \end{bmatrix}$$

Proof

$$(ST)x = B(Tx) = B(Ax) = B\left(\begin{bmatrix} A_1 & \dots & A_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)$$

$$= B(x_1A_1 + \dots + x_nA_n) = x_1BA_1 + \dots + x_nBA_n = \begin{bmatrix} BA_1 & \dots & BA_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Example

Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ rotate all $x \in \mathbb{R}^2$ by $\frac{\pi}{6}$ radians = 30° degree counterclockwise, then scale by 2. Find T .

The 30° degree counterclockwise rotation is $\begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$ and the scale by 2 transformation is $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ as noted by examples under [corollary 1.4.4](#) and [1.4.5](#).

$$T(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} x = \begin{bmatrix} 2\frac{\sqrt{3}}{2} + 0\frac{1}{2} & 2(-\frac{1}{2}) + 0\frac{\sqrt{3}}{2} \\ 0\frac{\sqrt{3}}{2} + 2\frac{1}{2} & 0(-\frac{1}{2}) + 2\frac{\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$$

Theorem 1.5.3: Properties of Matrix Products

- (a) **Associativity**: For $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{t \times m}(\mathbb{R})$, $C \in M_{s \times t}(\mathbb{R})$:
 $(CB)A = C(BA)$

Proof

Let $x \in \mathbb{R}^n$:

$$(CB)Ax = (CB)(A_1x_1 + \dots + A_nx_n) = CBA_1x_1 + \dots + CBA_nx_n = CBAx = C(BA)x$$

- (b) **Distributivity**: For $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{m \times n}(\mathbb{R})$, $C \in M_{t \times m}(\mathbb{R})$:
 $C(A+B) = CA + CB$

Proof

Let $x \in \mathbb{R}^n$:

$$\begin{aligned} C(A+B)x &= C((A_1 + B_1)x_1 + \dots + (A_n + B_n)x_n) \\ &= C((A_1x_1 + \dots + A_nx_n) + (B_1x_1 + \dots + B_nx_n)) \\ &= (CA_1x_1 + \dots + CA_nx_n) + (CB_1x_1 + \dots + CB_nx_n) \\ &= CAx + CBx = (CA+CB)x \end{aligned}$$

- (c) **Distributivity**: For $A \in M_{t \times m}(\mathbb{R})$, $B \in M_{t \times m}(\mathbb{R})$, $C \in M_{m \times n}(\mathbb{R})$:
 $(A+B)C = AC + BC$

Proof

Let $x \in \mathbb{R}^n$:

$$\begin{aligned} (A+B)Cx &= (A+B)(C_1x_1 + \dots + C_nx_n) \\ &= (A+B)C_1x_1 + \dots + (A+B)C_nx_n \\ &= (AC_1x_1 + \dots + AC_nx_n) + (BC_1x_1 + \dots + BC_nx_n) \\ &= ACx + BCx = (AC+BC)x \end{aligned}$$

- (d) **Scalar Multiplication**: For $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{t \times m}(\mathbb{R})$ and $c \in \mathbb{R}$:
 $(cB)A = c(BA) = B(cA)$

Proof

Let $x \in \mathbb{R}^n$:

$$\begin{aligned} (cB)Ax &= (cB)(A_1x_1 + \dots + A_nx_n) = cBA_1x_1 + \dots + cBA_nx_n = c(BA)x \\ &= BcA_1x_1 + \dots + BcA_nx_n = B(cA)x \end{aligned}$$

Definition 1.5.4: Identity Matrix

The $n \times n$ **identity matrix** $I_n \in M_{n \times n}(\mathbb{R})$:

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Definition 1.5.5: Elementary row operations are Matrix transformations

Elementary Row Operations as matrix are called **elementary matrices**.

Let $A \in M_{m \times n}(\mathbb{R})$. Then each elementary matrix $B \in M_{m \times m}(\mathbb{R})$ where:

Row Multiplication: Multiplying the i -th row by k

$$B = \begin{bmatrix} 1 & 1 & 2 & \dots & i & \dots & m \\ 1 & 1 & 0 & \dots & 0 & \dots & 0 \\ 2 & 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ i & 0 & 0 & \dots & k & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m & 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

Row Addition: Adding the k times the j -th row to the i -th row

$$B = \begin{bmatrix} 1 & 1 & 2 & \dots & i & \dots & j & \dots & m \\ 1 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 2 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ i & 0 & 0 & \dots & 1 & \dots & k & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ j & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 1 & 2 & \dots & j & \dots & i & \dots & m \\ 1 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 2 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ j & 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ i & 0 & 0 & \dots & k & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

Row Swapping: Swapping the i -th and j -th row

$$B = \begin{bmatrix} 1 & 1 & 2 & \dots & i & \dots & j & \dots & m \\ 1 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 2 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ i & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ j & 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 1 & 2 & \dots & j & \dots & i & \dots & m \\ 1 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 2 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ j & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ i & 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

Then, BA is the matrix after the elementary row operation is applied to A .

Definition 1.5.6: Invertibility

Linear Transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is **invertible** if there exist a linear transformation

$S: \mathbb{R}^m \rightarrow \mathbb{R}^n$ such that for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$:

$$(ST)x = x \quad (TS)y = y$$

Suppose $T(x) = Ax$ where $A \in M_{m \times n}(\mathbb{R})$ and $S(y) = By$ where $B \in M_{n \times m}(\mathbb{R})$.

Then if the property above holds true:

$$BAx = (ST)x = x = I_{n \times n}x \quad ABy = (TS)y = y = I_{m \times m}x$$

$$BA = I_{n \times n} \quad AB = I_{m \times m}$$

If A is invertible, then $B = A^{-1}$ is the **inverse transformation** of A .

Note if A is invertible, then A^{-1} is invertible since there is a A where $AA^{-1} = I_{m \times m}$ and $A^{-1}A = I_{n \times n}$ by the invertibility of A .

If $A \in M_{n \times n}(\mathbb{R})$, then A is called a **square matrix**.

Theorem 1.5.7: Only Square matrices can be Invertible

Let $A \in M_{m \times n}(\mathbb{R})$ and $B \in M_{n \times m}(\mathbb{R})$ such that $n \neq m$.

Then, either $AB \neq I_{m \times m}$ or $BA \neq I_{n \times n}$.

Proof

Suppose $n < m$. Then for any $x \in \mathbb{R}^n$, by [corollary 1.3.8](#), $By = x$ does not have a unique $y \in \mathbb{R}^m$ so either y does not exist or there are infinitely many y .

If y does not exist, then for $AB y = Ax$, the y does not exist so $AB \neq I_{m \times m}$ else $y = Ax$. If $By = x$ has infinitely many y , then there are y_1, y_2 where $y_1 \neq y_2$ such that $By_1 = x = By_2$ so $AB \neq I_{m \times m}$ else $y_1 = AB y_1 = AB y_2 = y_2$ contradicting $y_1 \neq y_2$. Thus, $AB \neq I_{m \times m}$.

Suppose $n > m$. Then for any $y \in \mathbb{R}^m$, by [corollary 1.3.8](#), $Ax = y$ does not have a unique $x \in \mathbb{R}^n$ so either x does not exist or there are infinitely many x .

If x does not exist, then for $BAx = By$, the x does not exist so $BA \neq I_{n \times n}$ else $x = By$. If $Ax = y$ has infinitely many x , then there are x_1, x_2 where $x_1 \neq x_2$ such that $Ax_1 = y = Ax_2$ so $BA \neq I_{n \times n}$ else $x_1 = BAx_1 = BAx_2 = x_2$ contradicting $x_1 \neq x_2$. Thus, $BA \neq I_{n \times n}$.

Theorem 1.5.8: Determining Invertibility

$A \in M_{n \times n}(\mathbb{R})$ is invertible if and only if $\text{rref}(A) = I_{n \times n}$

Proof

Let A be invertible. Suppose $\text{rref}(A) \neq I_{n \times n}$. Then there is at least one free variable.

Then for any $y \in \mathbb{R}^n$, by [theorem 1.3.7](#), $Ax = y$ has infinitely many $x \in \mathbb{R}^n$ since $x = I_{n \times n}x = A^{-1}Ax = A^{-1}y$ must exist due to the existence of A^{-1} by the invertibility of A . But then, there are x_1, x_2 where $x_1 \neq x_2$ such that $Ax_1 = y = Ax_2$ so $AA^{-1} \neq I_{n \times n}$ else $x_1 = A^{-1}Ax_1 = A^{-1}Ax_2 = x_2$ contradicting the invertibility of A . Thus, $\text{rref}(A) = I_{n \times n}$.

Let $\text{rref}(A) = I_{n \times n}$. Take the augmented matrix $[A \mid I_{n \times n}]$.

Since $\text{rref}(A) = I_{n \times n}$, then there is a sequence of elementary row operations that transforms A into $I_{n \times n}$ and thus, transform $[A \mid I_{n \times n}] = [I_{n \times n} \mid B]$ for some $B \in M_{n \times n}(\mathbb{R})$. Note

$$\begin{aligned} [A \mid I_{n \times n}] &\Leftrightarrow Ax = I_{n \times n}y & [I_{n \times n} \mid B] &\Leftrightarrow I_{n \times n}x = By \\ (BA)x = B(Ax) = By = x & & (AB)y = A(By) = Ax = y &\Rightarrow A^{-1} = B \end{aligned}$$

Corollary 1.5.9: Invertible $n \times n$ matrices have Rank n

$A \in M_{n \times n}(\mathbb{R})$ is invertible if and only if $\text{rank}(A) = n$

Proof

By [theorem 1.5.8](#), A is invertible $\Leftrightarrow \text{rref}(A) = I_{n \times n} \Leftrightarrow A$ has n pivots (i.e. $\text{rank}(A) = n$).

Corollary 1.5.10: Invertible Matrices have unique solutions

$A \in M_{n \times n}(\mathbb{R})$ is invertible if and only if for any $y \in \mathbb{R}^n$, there is a unique $x \in \mathbb{R}^n$ where:

$$Ax = y$$

Thus, A is invertible if and only if $Ax = 0$ has the trivial solution, $x = 0$.

Proof

Suppose A is invertible. Then by [corollary 1.5.9](#), $\text{rank}(A) = n$ so A has n pivots. Then for augmented matrix, $[A \mid y]$, by [theorem 1.3.7](#), there is one unique solution.

Suppose for any $y \in \mathbb{R}^n$, there is a unique $x \in \mathbb{R}^n$ where $Ax = y$. Then by [theorem 1.3.7](#), A has n pivots so $\text{rank}(A) = n$. Then by [corollary 1.5.9](#), A is invertible.

Suppose A is invertible. Since $A0 = 0$, then the only solution to $Ax = 0$ is $x = 0$.

Suppose $Ax = 0$ has only $x = 0$. Then, $\text{rref}(A)$ has n pivots so $\text{rank}(A) = n$. Thus, by [corollary 1.5.9](#), A is invertible.

Theorem 1.5.11: $AB = I_{n \times n}$ implies $BA = I_{n \times n}$

For $A, B \in M_{n \times n}(\mathbb{R})$, let $AB = I_{n \times n}$. Then A, B are invertible where:

$$A^{-1} = B \quad B^{-1} = A$$

Proof

Let $x \in \mathbb{R}^n$ be such that $Bx = 0$. Then, $x = I_{n \times n}x = ABx = B0 = 0$.

Then by **corollary 1.5.10**, B is invertible so B^{-1} exist where $B^{-1}B = I_{n \times n}$ and $BB^{-1} = I_{n \times n}$.

$$A = AI_{n \times n} = AI_{n \times n} = ABB^{-1} = I_{n \times n}B^{-1} = B^{-1}$$

Since B is invertible, then $A = B^{-1}$ is invertible so $A^{-1}A = I_{n \times n}$ and $AA^{-1} = I_{n \times n}$.

$$A^{-1} = A^{-1}I_{n \times n} = A^{-1}AB = I_{n \times n}B = B$$

Theorem 1.5.12: Invertibility Equivalences

Let $A \in M_{n \times n}(\mathbb{R})$. Then the following are equivalent:

- (a) A is invertible
- (b) $\text{rref}(A) = I_{n \times n}$
- (c) $\text{rank}(A) = n$
- (d) For any $y \in \mathbb{R}^n$, then $Ax = y$ has a unique solution x
- (e) $Ax = 0$ has only the trivial solution $x = 0$

Proof

$$(a) \Leftrightarrow \begin{cases} (b) & \text{theorem 1.5.8} \\ (c) & \text{corollary 1.5.9} \\ (d) & \text{corollary 1.5.10} \\ (e) & \text{corollary 1.5.10} \end{cases}$$

Theorem 1.5.13: Product of Invertible matrices is Invertible

Let $A, B \in M_{n \times n}(\mathbb{R})$ be invertible. Then, AB is invertible where:

$$(AB)^{-1} = B^{-1}A^{-1}$$

Proof

Since A and B are invertible, then there are $A^{-1}, B^{-1} \in M_{n \times n}(\mathbb{R})$ such that:

$$A^{-1}A = I_{n \times n} \quad AA^{-1} = I_{n \times n} \quad B^{-1}B = I_{n \times n} \quad BB^{-1} = I_{n \times n}$$

Then AB is invertible since $(AB)^{-1}$ exist as $(AB)^{-1} = B^{-1}A^{-1}$:

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_{n \times n}B = B^{-1}B = I_{n \times n}$$

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_{n \times n}A^{-1} = AA^{-1} = I_{n \times n}$$

Example

Let $T: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ be $T(x) = Ax$ where $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 7 & 11 \\ 3 & 7 & 14 & 25 \\ 4 & 11 & 25 & 50 \end{bmatrix}$. Find $T^{-1}(1, 1, -1, -6)$.

$$\begin{aligned} & \left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 2 & 4 & 7 & 11 & 0 & 1 & 0 & 0 \\ 3 & 7 & 14 & 25 & 0 & 0 & 1 & 0 \\ 4 & 11 & 25 & 50 & 0 & 0 & 0 & 1 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 5 & 13 & -3 & 0 & 1 & 0 \\ 0 & 3 & 13 & 34 & -4 & 0 & 0 & 1 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & -7 & -22 & 7 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 5 & 13 & -3 & 0 & 1 & 0 \\ 0 & 0 & -2 & -5 & 5 & 0 & -3 & 1 \end{array} \right] \\ & \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & -1 & -7 & 7 & -2 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 7 & -5 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 & -3 & 1 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -6 & 9 & -5 & 1 \\ 0 & 1 & 0 & 0 & 9 & -1 & -5 & 2 \\ 0 & 0 & 1 & 0 & -5 & -5 & 9 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 & -3 & 1 \end{array} \right] = A^{-1} \Rightarrow A \text{ is invertible} \\ & T^{-1}(1, 1, -1, -6) = A^{-1}(1, 1, -1, -6) = (2, 1, -1, 0) \end{aligned}$$

2 Vector Space

2.1 Span & Independence

Definition 2.1.1: Vector Space

V is a **vector space** over \mathbb{K} if for any $u, v, w \in V$ and $a, b \in \mathbb{K}$:

(a) **Commutativity**

$$u + v = v + u$$

(b) **Additive Associativity**

$$(u + v) + w = u + (v + w)$$

(c) **Additive Identity**

There exists a unique $0_v \in V$ such that for all $v \in V$:

$$0_v + v = v$$

(d) **Additive Inverse**

For any v , there exists a unique $-v \in V$ such that:

$$v + (-v) = 0_v$$

(e) **Distributivity**

$$a(u+v) = au + av \qquad (a+b)u = au + bu$$

(f) **Multiplicative Associativity**

$$a(bu) = (ab)u$$

(g) **Multiplicative Identity**

$$1u = u$$

Theorem 2.1.2: $0v = 0_v$, $a0_v = 0_v$, $(-1)v = -v$

Let V be a vector space where $v \in V$. Then:

$$0v = 0_v \qquad a0_v = 0_v \qquad (-1)v = -v$$

Proof

Since $0v = (0+0)v = 0v + 0v$, then:

$$0_v = 0v + (-0v) = 0v + 0v + (-0v) = 0v + 0_v = 0v$$

Since $a0_v = a(0_v+0_v) = a0_v + a0_v$, then:

$$0_v = a0_v + (-a0_v) = a0_v + a0_v + (-a0_v) = a0_v + 0_v = a0_v$$

Since $0_v = 0v = (-1+1)v = (-1)v + 1v = (-1)v + v$, then:

$$-v = 0_v + (-v) = (-1)v + v + (-v) = (-1)v + 0_v = (-1)v$$

Definition 2.1.3: Linear Combination, Span, and Independence

$x \in \mathbb{R}^n$ is a **linear combination** of $v_1, \dots, v_k \in \mathbb{R}^n$ if there are $c_1, \dots, c_k \in \mathbb{R}$ such that:

$$x = c_1v_1 + \dots + c_kv_k$$

The **span** of $v_1, \dots, v_k \in \mathbb{R}^n$ is the set of all linear combinations of v_1, \dots, v_k .

Also, $v_1, \dots, v_k \in \mathbb{R}^n$ are **linearly independent** if none of the v_i are linear combinations of the other v_i 's. Else, v_1, \dots, v_k are **linearly dependent**.

Theorem 2.1.4: Remove Linearly dependent vectors to get Linear independence

Let $u \in \mathbb{R}^n$ be a linear combination of $v_1, \dots, v_k \in \mathbb{R}^n$. Then:

$$\text{span}(v_1, \dots, v_k) = \text{span}(v_1, \dots, v_k, u)$$

Thus, by removing vectors that are linear combinations (i.e. linearly dependent vectors), then the resulting set of vectors will be linearly independent and the span is unaffected.

Proof

Since u is a linear combination of v_1, \dots, v_k , then there are $c_1, \dots, c_k \in \mathbb{R}$ such that:

$$u = c_1 v_1 + \dots + c_k v_k$$

Let u_1 be a linear combination of v_1, \dots, v_k, u . Then there are $a_1, \dots, a_k, a \in \mathbb{R}$ such that:

$$u_1 = a_1 v_1 + \dots + a_k v_k + a u = a_1 v_1 + \dots + a_k v_k + a(c_1 v_1 + \dots + c_k v_k) = (a_1 + a c_1) v_1 + \dots + (a_k + a c_k) v_k$$

Thus, $u_1 \in \text{span}(v_1, \dots, v_k)$ so $\text{span}(v_1, \dots, v_k, u) \subset \text{span}(v_1, \dots, v_k)$.

Let u_2 be a linear combination of v_1, \dots, v_k . Then there are $b_1, \dots, b_k \in \mathbb{R}$ such that:

$$\begin{aligned} u_2 &= b_1 v_1 + \dots + b_k v_k = [(b_1 - c_1) v_1 + \dots + (b_k - c_k) v_k] + [c_1 v_1 + \dots + c_k v_k] \\ &= (b_1 - c_1) v_1 + \dots + (b_k - c_k) v_k + u \end{aligned}$$

Thus, $u_2 \in \text{span}(v_1, \dots, v_k, u)$ so $\text{span}(v_1, \dots, v_k) \subset \text{span}(v_1, \dots, v_k, u)$.

Theorem 2.1.5: Condition for Linear independence

$v_1, \dots, v_k \in \mathbb{R}^n$ are linearly independent if and only if the only solution (c_1, \dots, c_k) :

$$c_1 v_1 + \dots + c_k v_k = 0$$

is $c_1 = \dots = c_k = 0$.

Note regardless of v_1, \dots, v_k , any $c_1 v_1 + \dots + c_k v_k = 0$ holds true when $(c_1, \dots, c_k) = 0 \in \mathbb{R}^k$.

$(c_1, \dots, c_k) = 0$ is called the **trivial solution**. Any $(c_1, \dots, c_k) \neq 0$ is a **nontrivial solution**.

Proof

Suppose $v_1, \dots, v_k \in \mathbb{R}^n$ are linearly independent.

Then for any v_i , there are no $c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_k \in \mathbb{R}$ such that:

$$v_i = c_1 v_1 + \dots + c_{i-1} v_{i-1} + c_{i+1} v_{i+1} + \dots + c_k v_k$$

Thus, there are no $(c_1, \dots, c_{i-1}, c_i = -1, c_{i+1}, \dots, c_k)$ such that:

$$0 = c_1 v_1 + \dots + c_{i-1} v_{i-1} + c_i v_i + c_{i+1} v_{i+1} + \dots + c_k v_k$$

The statement holds true if the equation was multiplied by any non-zero number. Thus, any (c_1, \dots, c_k) where at least one c_i is not 0 is not a solution. Since $(c_1, \dots, c_k) = 0$ is a solution for $c_1 v_1 + \dots + c_k v_k = 0$, then for linearly independent v_1, \dots, v_k , then $(c_1, \dots, c_k) = 0$.

Suppose the solution, (c_1, \dots, c_k) , to $c_1 v_1 + \dots + c_k v_k = 0$ is only $(c_1, \dots, c_k) = 0$.

Suppose there is a linearly dependent vector, v_i . Then there are $a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_k$ where:

$$v_i = a_1 v_1 + \dots + a_{i-1} v_{i-1} + a_{i+1} v_{i+1} + \dots + a_k v_k$$

$$0 = a_1 v_1 + \dots + a_{i-1} v_{i-1} + -v_i + a_{i+1} v_{i+1} + \dots + a_k v_k$$

Thus, $(a_1, \dots, a_{i-1}, -1, a_{i+1}, \dots, a_k)$ is a solution to $c_1 v_1 + \dots + c_k v_k = 0$ contradicting that $(c_1, \dots, c_k) = 0$. Thus, there are no linearly dependent vectors.

Theorem 2.1.6: Extending the Span of Linearly independent vectors

Let $v_1, \dots, v_k \in \mathbb{R}^n$ be linearly independent. If $v \in \mathbb{R}^n$ is not in the $\text{span}(v_1, \dots, v_k)$, then v_1, \dots, v_k, v are linearly independent.

Proof

Let $c_1, \dots, c_k, c \in \mathbb{R}$ be such that $c_1 v_1 + \dots + c_k v_k + c v = 0$. Suppose $c \neq 0$. Then:

$$c_1 v_1 + \dots + c_k v_k = -c v \quad \Rightarrow \quad v = \frac{-c_1}{c} v_1 + \dots + \frac{-c_k}{c} v_k$$

Then, v is a linear combination of v_1, \dots, v_k and thus, v is in the $\text{span}(v_1, \dots, v_k)$ which is a contradiction. Thus, $c = 0$. Then:

$$0 = c_1 v_1 + \dots + c_k v_k + c v = c_1 v_1 + \dots + c_k v_k$$

Since v_1, \dots, v_k are linearly independent, then each $c_1 = \dots = c_k = 0$. Thus, $(c_1, \dots, c_k, c) = 0$ so v_1, \dots, v_k, v are linearly independent.

2.2 Subspaces: Image & Kernel

Definition 2.2.1: Subspaces

$V \subset \mathbb{R}^n$ is a **subspace** of \mathbb{R}^n if:

(a) **Zero Vector Existence**

$$0 \in V$$

(b) **Closed under Addition**: If $v_1, v_2 \in V$, then:

$$v_1 + v_2 \in V$$

(c) **Closed under Scalar Multiplication**: If $v \in V$ and $c \in \mathbb{R}$, then:

$$cv \in V$$

Theorem 2.2.2: Union of Subspaces's condition for Subspace

Let $U, V \subset \mathbb{R}^n$ be subspaces. Then, $U \cup V$ is a subspace if and only if $U \subset V$ or $V \subset U$.

Proof

Suppose $U \cup V$ is a subspace. Suppose $U \not\subset V$ and $V \not\subset U$. Then there is a $u \in U$ where $u \notin V$ and a $v \in V$ where $v \notin U$. Thus, $u, v \in U \cup V$, but $u+v \notin U$ since $v \notin U$ and $u+v \notin V$ since $u \notin V$. Thus, $u+v \notin U \cup V$ contradicting $U \cup V$ is a subspace. Thus, $U \subset V$ or $V \subset U$.

If $U \subset V$, then $U \cup V = V$ is a subspace. If $V \subset U$, then $U \cup V = U$ is a subspace.

Theorem 2.2.3: Intersection of Subspaces is a Subspace

Let $U, V \subset \mathbb{R}^n$ be subspaces. Then, $U \cap V$ is a subspace.

Proof

Let $x, y \in U \cap V$ and $a, b \in \mathbb{R}$. Then $x, y \in U, V$. Since U and V are subspaces, then $ax+by \in U, V$. Thus, $ax+by \in U \cap V$ so $U \cap V$ is a subspace.

Definition 2.2.4: Image and Kernel

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the linear transformation $T(x) = Ax$.

The **image** of T is the set of all $Ax \in \mathbb{R}^m$ where $x \in \mathbb{R}^n$:

$$\text{im}(T) = \text{im}(A) = \{ Ax \mid x \in \mathbb{R}^n \}$$

The **kernel** of T is the set of all $x \in \mathbb{R}^n$ such that $Ax = 0$:

$$\text{ker}(T) = \text{ker}(A) = \{ x \mid Ax = 0 \}$$

Theorem 2.2.5: $\text{im}(A)$ is a Subspace that spans the columns of A

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the linear transformation $T(x) = Ax$.

Then, $\text{im}(A)$ is a subspace and $\text{im}(A) = \text{span}(A)$.

Proof

Since $Ax = A_1x_1 + \dots + A_nx_n$ for $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ where A_1, \dots, A_n are the columns of A :

$$\text{im}(A) = \{ Ax \mid x \in \mathbb{R}^n \} = \{ x_1A_1 + \dots + x_nA_n \mid x_1, \dots, x_n \in \mathbb{R} \} = \text{span}(A_1, \dots, A_n)$$

Since $A0 = 0$, then $0 \in \text{im}(A)$. Let $u, v \in \text{im}(A)$ and $a, b \in \mathbb{R}$.

Then there are $a_1, \dots, a_n \in \mathbb{R}$ and $b_1, \dots, b_n \in \mathbb{R}$ such that:

$$u = a_1A_1 + \dots + a_nA_n \quad v = b_1A_1 + \dots + b_nA_n$$

Thus, $au+bv = (aa_1 + bb_1)A_1 + \dots + (aa_n + bb_n)A_n \in \text{span}(A_1, \dots, A_n) = \text{im}(A)$.

Theorem 2.2.6: $\text{ker}(A)$ is a Subspace

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the linear transformation $T(x) = Ax$. Then, $\text{ker}(A)$ is a subspace.

Proof

Since $A0 = 0$, then $0 \in \text{ker}(A)$. Let $x_1, x_2 \in \text{ker}(A)$ and $a, b \in \mathbb{R}$ so $Ax_1 = Ax_2 = 0$.

Then, $A(ax_1 + bx_2) = aAx_1 + bAx_2 = a0 + b0 = 0$ so $ax_1 + bx_2 \in \text{ker}(A)$.

Theorem 2.2.7: Relationship between the Kernel and Linear independence

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the linear transformation $T(x) = Ax$ where $A = \begin{bmatrix} A_1 & \dots & A_m \end{bmatrix}$.

Then, A_1, \dots, A_m are linearly independent if and only if $\ker(A) = \{0\}$.

Proof

Note $Ax = \begin{bmatrix} A_1 & \dots & A_m \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x_1 A_1 + \dots + x_m A_m = 0$. Let A_1, \dots, A_m be linearly independent. Then, the only solution to $x_1 A_1 + \dots + x_m A_m = 0$ is $(x_1, \dots, x_m) = 0$ so $\ker(A) = \{0\}$. Suppose $\ker(A) = \{0\}$. Then, the only solution to $x_1 A_1 + \dots + x_m A_m = Ax = 0$ is $x = 0$. By **theorem 2.1.5**, A_1, \dots, A_m are linearly independent.

Example

Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ be $T(x) = Ax$ where $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 7 \\ 1 & 4 & 10 \end{bmatrix}$.

Find the $\text{im}(A)$ and $\ker(A)$ and determine if the columns of A are independent.

$\text{im}(A) = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 7 \\ 10 \end{bmatrix}\right) = c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 4 \\ 7 \\ 10 \end{bmatrix}$ for $c_1, c_2, c_3 \in \mathbb{R}$

$$\begin{bmatrix} 1 & 1 & 1 & | & 0 \\ 1 & 2 & 4 & | & 0 \\ 1 & 3 & 7 & | & 0 \\ 1 & 4 & 10 & | & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 & | & 0 \\ 0 & 1 & 3 & | & 0 \\ 0 & 2 & 6 & | & 0 \\ 0 & 3 & 9 & | & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -2 & | & 0 \\ 0 & 1 & 3 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \Leftrightarrow \begin{matrix} x_1 - 2x_3 = 0 \\ x_2 + 3x_3 = 0 \end{matrix} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_3 \\ -3x_3 \\ x_3 \end{bmatrix}$$

$\ker(A) = x_3 \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} \neq \{0\}$ for $x_3 \in \mathbb{R}$ so $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 7 \\ 10 \end{bmatrix}$ are not linearly independent.

2.3 Basis & Dimension**Definition 2.3.1: Basis & Dimension**

Let set $V \subset \mathbb{R}^n$ be subspace of \mathbb{R}^n . If $v_1, \dots, v_k \in \mathbb{R}^n$ are linearly independent and span V (i.e. $\text{span}(v_1, \dots, v_k) = V$), then v_1, \dots, v_k form a **basis** for V .

The **dimension** of V , $\dim(V)$, is the number of vectors in a basis of V .

Since e_1, \dots, e_n are linearly independent and any $x \in \mathbb{R}^n$, is $x = x_1 e_1 + \dots + x_n e_n$ so $\text{span}(e_1, \dots, e_n) = \mathbb{R}^n$, then $\dim(\mathbb{R}^n) = n$. e_1, \dots, e_n are called the **standard basis vectors**.

Theorem 2.3.2: # Linearly independent vectors in $V \leq$ # vectors that span V

Let $v_1, \dots, v_m \in V$ be linearly independent and $u_1, \dots, u_k \in V$ span V , then $m \leq k$

Proof

Let $A = \begin{bmatrix} v_1 & \dots & v_m \end{bmatrix} \in M_{n \times m}(\mathbb{R})$ and $B = \begin{bmatrix} u_1 & \dots & u_k \end{bmatrix} \in M_{n \times k}(\mathbb{R})$. Since $\text{im}(B) = \text{span}(B) = V$, then for $v_1, \dots, v_m \in V$, there are $c_1, \dots, c_m \in \mathbb{R}^k$ such that $v_i = Bc_i$:

$$A = \begin{bmatrix} v_1 & \dots & v_m \end{bmatrix} = \begin{bmatrix} Bc_1 & \dots & Bc_m \end{bmatrix} = B \begin{bmatrix} c_1 & \dots & c_m \end{bmatrix} = BC \quad \text{where } C \in M_{k \times m}(\mathbb{R})$$

Thus, if $Cx = 0$, then $Ax = BCx = B0 = 0$ so $\ker(C) \subset \ker(A)$. Since $v_1, \dots, v_m \in V$ be linearly independent, then by **theorem 2.2.7**, $\ker(A) = \{0\}$ so $\ker(C) = \{0\}$. Since $Cx = 0$ has a unique solution $x = 0$, then by **corollary 1.3.8**, $k \geq m$.

Corollary 2.3.3: The Dimension is unique

Suppose $v_1, \dots, v_m \in V$ and $u_1, \dots, u_k \in V$ are bases for V , then $\dim(V) = m = k$

Proof

Since v_1, \dots, v_m and u_1, \dots, u_k are bases, then v_1, \dots, v_m and u_1, \dots, u_k are linearly independent and span V . By [theorem 2.3.2](#), $m \leq k$ and $k \leq m$ so $m = k$.

Theorem 2.3.4: Linear combinations of a Basis are unique

Let $v_1, \dots, v_k \in \mathbb{R}^n$ form a basis for $V \subset \mathbb{R}^n$.

Then for any $v \in V$, there are unique (c_1, \dots, c_k) such that:

$$v = c_1 v_1 + \dots + c_k v_k$$

Proof

Since v_1, \dots, v_k form a basis for V , then $V = \text{span}(v_1, \dots, v_k)$. Then for any $v \in V$, then $v \in \text{span}(v_1, \dots, v_k)$. Thus, there are $c_1, \dots, c_k \in \mathbb{R}$ such that $v = c_1 v_1 + \dots + c_k v_k$.

Let $a_1, \dots, a_k \in \mathbb{R}$ such that $v = a_1 v_1 + \dots + a_k v_k$. Then:

$$0 = (c_1 - a_1)v_1 + \dots + (c_k - a_k)v_k$$

Since v_1, \dots, v_k form a basis for V , then v_1, \dots, v_k are linearly independent. Thus, by [theorem 2.1.5](#), $c_i - a_i = 0$ for $i = \{1, \dots, k\}$ so $c_i = a_i$ for $i = \{1, \dots, k\}$. Thus, (c_1, \dots, c_k) must be unique.

Theorem 2.3.5: Connection between Span, Linear independence, and Basis

Let $V \subset \mathbb{R}^n$ where $\dim(V) = m$. Then, $m \leq n$ and:

- (a) If $u_1, \dots, u_k \in V$ are linearly independent, then $k \leq m$
- (b) For $u_1, \dots, u_k \in V$, if $\text{span}(u_1, \dots, u_k) = V$, then $k \geq m$
- (c) $u_1, \dots, u_m \in V$ are linearly independent if and only if $\text{span}(u_1, \dots, u_m) = V$

Proof

Since $\dim(V) = m$, then there are $v_1, \dots, v_m \in \mathbb{R}^n$ that are linearly independent and span V . Then for any $v \in V$, there are c_1, \dots, c_m where:

$$v = c_1 v_1 + \dots + c_m v_m \Leftrightarrow \begin{bmatrix} v_1 & \dots & v_m \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = v \Leftrightarrow Ac = v \quad \text{where } A \in M_{n \times m}(\mathbb{R})$$

By [theorem 2.3.4](#), c is unique. Since $A \in M_{n \times m}(\mathbb{R})$, then by [corollary 1.3.8](#), $n \geq m$.

Let u_1, \dots, u_k be linearly independent. Since v_1, \dots, v_m span V , then by [theorem 2.3.2](#), $k \leq m$.

Let u_1, \dots, u_k span V . Since v_1, \dots, v_m are linearly independent, then by [theorem 2.3.2](#), $m \leq k$

Suppose $u_1, \dots, u_m \in V$ are linearly independent. Let $v \in V$. If $v \notin \text{span}(u_1, \dots, u_m)$, then by [theorem 2.1.6](#), u_1, \dots, u_m, v are linearly independent contradicting that $m+1 = \dim(V) = m$. Thus, any $v \in V$ is $v \in \text{span}(u_1, \dots, u_m)$ so $V \subset \text{span}(u_1, \dots, u_m)$. Since $\text{span}(u_1, \dots, u_m) \subset V$, then $\text{span}(u_1, \dots, u_m) = V$.

Suppose $\text{span}(u_1, \dots, u_m) = V$. Suppose there are c_1, \dots, c_m where not all $c_i = 0$ such that:

$$c_1 u_1 + \dots + c_m u_m = 0$$

If there is one $c_i \neq 0$, then $c_i u_i = c_1 u_1 + \dots + c_m u_m = 0$ implies $u_i = 0$ which is a dependent vector. Thus, there are at least two $c_i \neq 0$. Then, $u_i = \frac{-c_1}{c_i} u_1 + \dots + \frac{-c_{i-1}}{c_i} u_{i-1} + \frac{-c_{i+1}}{c_i} u_{i+1} + \dots + \frac{-c_m}{c_i} u_m$ where at least one $c_{j \neq i} \neq 0$ so u_i is a dependent vector. So there are less than m independent vectors contradicting that $m = \dim(V) < m$. Thus, all $c_i = 0$ so u_1, \dots, u_m are independent.

Theorem 2.3.6: Determining Basis for $\text{Im}(A)$ using rref

Let $A = [v_1 \ \dots \ v_k] \in M_{n \times k}(\mathbb{R})$ where each $v_i \in \mathbb{R}^n$. Then the v_i that are linearly independent are the columns in $\text{rref}(A)$ with pivots. Thus, such v_i form a basis for $\text{im}(A)$.

Proof

By [theorem 2.2.5](#), $\text{im}(A) = \text{span}(A) = \text{span}(v_1, \dots, v_k)$. Suppose the i -th column of $\text{rref}([A \mid 0])$ does not contain a pivot. Then, $Ax = 0$ where $x \in \mathbb{R}^k$ has a free variable at x_i so $\ker(A)$ has nonzero solutions. Thus, by [theorem 2.2.7](#), v_1, \dots, v_k are linearly dependent. But, if all columns without pivots are removed are removed from $\text{rref}(A)$ to make B , then $Bx = 0$ has only pivots and thus, by [theorem 1.3.7](#), there is a unique solution. Since $B0 = 0$, then $\ker(B) = \{0\}$ and thus, the columns with pivots are linearly independent and by [theorem 2.1.4](#), the $\text{span}(B) = \text{span}(A) = \text{im}(A)$ so the columns with pivots form a basis for $\text{im}(A)$.

Corollary 2.3.7: $\dim(\text{im}(A)) = \text{rank}(A)$

For any $A \in M_{m \times n}(\mathbb{R})$, then:

$$\dim(\text{im}(A)) = \text{rank}(A)$$

Proof

Let $A = [A_1 \ \dots \ A_n]$ where $A_i \in \mathbb{R}^m$. By [theorem 2.3.6](#), the $\dim(\text{im}(A))$, the number of vectors in a basis for $\text{im}(A)$ is the same as the number of pivots in $\text{rref}(A)$, $\text{rank}(A)$.

Theorem 2.3.8: Rank-Nullity Theorem

For any $A \in M_{m \times n}(\mathbb{R})$, the $\dim(\ker(A))$ is called the **nullity** of A where:

$$\dim(\text{im}(A)) + \dim(\ker(A)) = n$$

Proof

Note the number of free variables + the number of pivot variables = n .

By [corollary 2.3.7](#), the number of pivot variables, $\text{rank}(A) = \dim(\text{im}(A))$.

If the i -th column doesn't have a pivot, then the solutions to $Ax = 0$ are linear combinations of vectors v_i with 1 in the i -th row and 0 in any j -th row where the j -th column doesn't have a pivot so each v_i is linearly independent and $\text{span } \ker(A)$. Thus, the number of free variables is equal to the number of v_i , $\dim(\ker(A))$, so $\dim(\text{im}(A)) + \dim(\ker(A)) = n$.

Example

Let $T: \mathbb{R}^5 \rightarrow \mathbb{R}^2$ be $T(x) = Ax$ where $A = \begin{bmatrix} 1 & 2 & 3 & 3 & 0 \\ 4 & 5 & 6 & 9 & 6 \end{bmatrix}$.

Find a basis for the $\text{im}(A)$ and $\ker(A)$. Verify $\dim(\text{im}(A)) + \dim(\ker(A)) = 5$.

$$\begin{bmatrix} 1 & 2 & 3 & 3 & 0 & | & 0 \\ 4 & 5 & 6 & 9 & 6 & | & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 3 & 0 & | & 0 \\ 0 & -3 & -6 & -3 & 6 & | & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 3 & 0 & | & 0 \\ 0 & 1 & 2 & 1 & -2 & | & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 & 1 & 4 & | & 0 \\ 0 & 1 & 2 & 1 & -2 & | & 0 \end{bmatrix}$$

$$\text{im}(A) = \text{span}\left(\begin{bmatrix} 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}\right) = c_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 5 \end{bmatrix} \text{ for } c_1, c_2 \in \mathbb{R}$$

$$\begin{bmatrix} 1 & 0 & -1 & 1 & 4 & | & 0 \\ 0 & 1 & 2 & 1 & -2 & | & 0 \end{bmatrix} \Leftrightarrow \begin{matrix} x_1 - x_3 + x_4 + 4x_5 = 0 \\ x_2 + 2x_3 + x_4 - 2x_5 = 0 \end{matrix} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} x_3 - x_4 - 4x_5 \\ -2x_3 - x_4 + 2x_5 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -4 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\ker(A) = \text{span}\left(\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right) = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -4 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \text{ for } x_3, x_4 \in \mathbb{R}$$

Then, $\dim(\text{im}(A)) + \dim(\ker(A)) = 2 + 3 = 5$

2.4 Injectivity & Surjectivity

Definition 2.4.1: Injectivity and Surjectivity

Let $T: V \rightarrow W$ be a linear transformation.

T is **injective** if for any $w \in T(V) = \text{im}(T)$, there is a unique $v \in V$ such that $T(v) = w$

T is **surjective** if for any $w \in W$, there is a $v \in V$ such that $T(v) = w$

Theorem 2.4.2: Connection between Invertibility, Injectivity, and Surjectivity

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation where $T(x) = Ax$.

Then, T is invertible if and only if T is injective and surjective.

Proof

Suppose T is invertible. Let $y \in \mathbb{R}^n$. Then, by **theorem 1.5.12**, there is a unique $x \in \mathbb{R}^n$ such that $Ax = T(x) = y$. Thus, T is injective and surjective.

Suppose T is injective and surjective. Since T is surjective, then for any $y \in \mathbb{R}^n$, there is a $x \in \mathbb{R}^n$ such that $Ax = T(x) = y$. Since T is injective, then x is unique. Then, by **theorem 1.5.12**, T is invertible.

Theorem 2.4.3: Connection between Invertibility, Span, and Linear independence

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation $T(x) = Ax$ where $A = \begin{bmatrix} A_1 & \dots & A_n \end{bmatrix}$.

Then, T is invertible if and only if A_1, \dots, A_n are linearly independent and span \mathbb{R}^n .

Proof

Suppose T is invertible. By **theorem 1.5.12**, for any $y \in \mathbb{R}^n$, there is a unique $x \in \mathbb{R}^n$ where $Ax = y$. Thus, $\text{span}(A_1, \dots, A_n) = \text{im}(A)$. Since $\dim(\mathbb{R}^n) = n$, then A_1, \dots, A_n span \mathbb{R}^n , then by **theorem 2.3.5**, A_1, \dots, A_n are linearly independent.

Theorem 2.4.4: Injectivity \Leftrightarrow Surjectivity

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation where $T(x) = Ax$ where $A = \begin{bmatrix} A_1 & \dots & A_n \end{bmatrix}$.

Then, T is surjective if and only if T is injective.

Proof

Suppose T is surjective. Then for any $y \in \mathbb{R}^n$, there is a $x \in \mathbb{R}^n$ where $Ax = T(x) = y$. Thus, $\text{span}(A_1, \dots, A_n) = \text{im}(A) = \mathbb{R}^n$. Since $\dim(\mathbb{R}^n) = n$, then by **theorem 2.3.5**, A_1, \dots, A_n are linearly independent. By **theorem 2.4.3**, T is invertible, then by **theorem 2.4.2**, T is injective.

Suppose T is injective. Then for any $y \in \text{im}(A)$, there is a unique $x \in \mathbb{R}^n$ where $Ax = T(x) = y$. Suppose there is a linearly dependent A_i . Then there are $c_1, \dots, c_{i-1}, c_{i+1}, c_n$ where at least one $c_i \neq 0$ such that $A_i = c_1 A_1 + \dots + c_{i-1} A_{i-1} + c_{i+1} A_{i+1} + \dots + c_n A_n$. Then:

$$Ax = x_1 A_1 + \dots + x_n A_n$$

$$= (x_1 + x_i c_1) A_1 + \dots + (x_{i-1} + x_i c_{i-1}) A_{i-1} + (x_{i+1} + x_i c_{i+1}) A_{i+1} + \dots + (x_n + x_i c_n) A_n$$

If $x_i \neq 0$, then $x = (x_1, \dots, x_n)$ is not the only solution to $y = Ax$ contradicting that x is unique. Thus, $x_i = 0$. Similarly, if any other A_j is linearly dependent, then $x_j = 0$. Thus, $Ax = x_1 A_1 + \dots + x_n A_n = x_{i_1} A_{i_1} + \dots + x_{i_k} A_{i_k}$ where A_{i_1}, \dots, A_{i_k} are the A_i that are linearly independent. Thus, for $Ax = 0$, then each $x_{i_1} = 0$. Since then all $x_i = 0$, then no A_i is linearly dependent so A_1, \dots, A_n are linearly independent. By **theorem 2.4.3**, T is invertible, then by **theorem 2.4.2**, T is surjective.

Theorem 2.4.5: Invertibility Equivalences Extended

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation where $T(x) = Ax$ where $A = \begin{bmatrix} A_1 & \dots & A_n \end{bmatrix}$.

Then the following are equivalent:

- (a) A is invertible
- (b) $\text{rref}(A) = I_{n \times n}$
- (c) $\text{rank}(A) = n$
- (d) For any $y \in \mathbb{R}^n$, then $Ax = y$ has a unique solution x
- (e) $Ax = 0$ has only the trivial solution $x = 0$ (i.e $\ker(A) = \{0\}$)
- (f) A_1, \dots, A_n are linearly independent
- (g) $\text{im}(A) = \text{span}(A_1, \dots, A_n) = \mathbb{R}^n$
- (h) T is injective
- (i) T is surjective

Proof

$$(a) \Leftrightarrow \begin{cases} (b \rightarrow e) & \text{theorem 1.5.12} \\ (f) & \text{theorem 2.4.3 and 2.3.5.} \\ & (a) \Rightarrow (f) \text{ only needs 2.4.3, but } (a) \Leftarrow (f) \text{ needs 2.4.3, 2.3.5} \\ (g) & \text{theorem 2.4.3 and 2.3.5} \\ & (a) \Rightarrow (g) \text{ only needs 2.4.3, but } (a) \Leftarrow (g) \text{ needs 2.4.3, 2.3.5} \\ (h) & \text{theorem 2.4.2 and 2.4.4} \\ & (a) \Rightarrow (h) \text{ only needs 2.4.2, but } (a) \Leftarrow (h) \text{ needs 2.4.2, 2.4.4} \\ (i) & \text{theorem 2.4.2 and 2.4.4} \\ & (a) \Rightarrow (i) \text{ only needs 2.4.2, but } (a) \Leftarrow (i) \text{ needs 2.4.2, 2.4.4} \end{cases}$$

Example

Let $T: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ be $T(x) = Ax$ where $A = \begin{bmatrix} 1 & 0 & 2 & 4 \\ 0 & 1 & -3 & -1 \\ 3 & 4 & -6 & 8 \\ 0 & -1 & 3 & 1 \end{bmatrix}$.

Find if T is invertible, $\text{im}(T)$, $\dim(\text{im}(T))$, $\ker(T)$, and $\dim(\ker(T))$.

$$\begin{aligned} [A \mid I_{4 \times 4}] &= \left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 4 & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & -1 & 0 & 1 & 0 & 0 \\ 3 & 4 & -6 & 8 & 0 & 0 & 1 & 0 \\ 0 & -1 & 3 & 1 & 0 & 0 & 0 & 1 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 4 & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & -1 & 0 & 1 & 0 & 0 \\ 0 & 4 & -12 & -4 & -3 & 0 & 1 & 0 \\ 0 & -1 & 3 & 1 & 0 & 0 & 0 & 1 \end{array} \right] \\ &\Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 4 & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{array} \right] \Rightarrow \text{rref}(A) \neq I_{4 \times 4} \Rightarrow A \text{ is not invertible} \end{aligned}$$

Since the 1st and 2nd column have pivots, then $(1,0,3,0)$ and $(0,1,4,-1)$ form a basis for $\text{im}(T)$.

$$\text{im}(T) = c_1(1, 0, 3, 0) + c_2(0, 1, 4, -1) \quad \text{for } c_1, c_2 \in \mathbb{R}$$

where $(1,0,3,0)$, $(0,1,4,-1)$ form a basis for $\text{im}(T)$ so $\dim(\text{im}(T)) = 2$

To find $\ker(T)$, (i.e. solving $Ax = 0$), replace the right matrix by a column with all 0 entries:

$$\left[\begin{array}{cccc|c} 1 & 0 & 2 & 4 & 0 \\ 0 & 1 & -3 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \Leftrightarrow \begin{cases} x_1 + 2x_3 + 4x_4 = 0 \\ x_2 - 3x_3 - x_4 = 0 \end{cases} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2x_3 - 4x_4 \\ 3x_3 + x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ 3 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\ker(T) = x_3(-2, 3, 1, 0) + x_4(-4, 1, 0, 1) \quad \text{for } x_3, x_4 \in \mathbb{R}$$

where $(-2,3,1,0)$, $(-4,1,0,1)$ form a basis for $\ker(T)$ so $\dim(\ker(T)) = 2$

2.5 Coordinates

Definition 2.5.1: Coordinates

For $V \subset \mathbb{R}^n$, let $v_1, \dots, v_k \in V$ be a basis for V . Then the **coordinates** of $v \in V$ relative to basis $\mathcal{B} = v_1, \dots, v_k$ are $c_1, \dots, c_k \in \mathbb{R}$ such that:

$$v = c_1 v_1 + \dots + c_k v_k$$

Then the **coordinate vector** $[v]_{\mathcal{B}} = (c_1, \dots, c_k) \in \mathbb{R}^k$. Let $B = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \in M_{n \times k}(\mathbb{R})$.

$$v = c_1 v_1 + \dots + c_k v_k = B[v]_{\mathcal{B}}$$

Theorem 2.5.2: Properties of Coordinates

For $V \subset \mathbb{R}^n$, let $\mathcal{B} = v_1, \dots, v_k \in V$ be a basis for V . For $x, y \in V$ and $c \in \mathbb{R}$:

$$(a) [x + y]_{\mathcal{B}} = [x]_{\mathcal{B}} + [y]_{\mathcal{B}}$$

$$(b) [cx]_{\mathcal{B}} = c[x]_{\mathcal{B}}$$

Proof

Let $x = a_1 v_1 + \dots + a_k v_k$ and $y = b_1 v_1 + \dots + b_k v_k$ for $a_1, \dots, a_k, b_1, \dots, b_k \in \mathbb{R}$. Let $c_1, c_2 \in \mathbb{R}$.
 $[c_1 x + c_2 y]_{\mathcal{B}} = (c_1 a_1 + c_2 b_1, \dots, c_1 a_k + c_2 b_k) = c_1 (a_1, \dots, a_k) + c_2 (b_1, \dots, b_k) = c_1 [x]_{\mathcal{B}} + c_2 [y]_{\mathcal{B}}$

Theorem 2.5.3: Linear transformation by Coordinates

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation where $T(x) = Ax$. Let $\mathcal{B} = v_1, \dots, v_n$ be a basis for \mathbb{R}^n . Then for any $x \in \mathbb{R}^n$:

$$[T(x)]_{\mathcal{B}} = [Ax]_{\mathcal{B}} = A_{\mathcal{B}}[x]_{\mathcal{B}}$$

where the **\mathcal{B} -matrix of T** , $A_{\mathcal{B}} = \begin{bmatrix} [A(v_1)]_{\mathcal{B}} & \dots & [A(v_n)]_{\mathcal{B}} \end{bmatrix} \in M_{n \times n}(\mathbb{R})$

Proof

Let $x = c_1 v_1 + \dots + c_n v_n$ for $c_1, \dots, c_n \in \mathbb{R}$. Then $[x]_{\mathcal{B}} = (c_1, \dots, c_n)$:
 $[T(x)]_{\mathcal{B}} = [c_1 A(v_1) + \dots + c_n A(v_n)]_{\mathcal{B}} = c_1 [A(v_1)]_{\mathcal{B}} + \dots + c_n [A(v_n)]_{\mathcal{B}}$

$$= \begin{bmatrix} [A(v_1)]_{\mathcal{B}} & \dots & [A(v_n)]_{\mathcal{B}} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} [A(v_1)]_{\mathcal{B}} & \dots & [A(v_n)]_{\mathcal{B}} \end{bmatrix} [x]_{\mathcal{B}}$$

Theorem 2.5.4: Relationship between Ax and $A_{\mathcal{B}}x_{\mathcal{B}}$

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation where $T(x) = Ax$. Let $\mathcal{B} = v_1, \dots, v_n$ be a basis for \mathbb{R}^n where $[T(x)]_{\mathcal{B}} = A_{\mathcal{B}}[x]_{\mathcal{B}}$. Then:

$$AB = BA_{\mathcal{B}} \quad A_{\mathcal{B}} = B^{-1}AB$$

where $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} \in M_{n \times n}(\mathbb{R})$ is invertible and $[x]_{\mathcal{B}} = B^{-1}x$

Proof

Let $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$. Since $x = B[x]_{\mathcal{B}}$ and $T(x) = B[T(x)]_{\mathcal{B}}$, then:
 $T(x) = Ax = AB[x]_{\mathcal{B}} \quad T(x) = B[T(x)]_{\mathcal{B}} = BA_{\mathcal{B}}[x]_{\mathcal{B}}$
 Thus, $AB = BA_{\mathcal{B}}$. Since v_1, \dots, v_n is a basis, then by **theorem 2.4.5**, B is invertible.
 Since $x = B[x]_{\mathcal{B}}$, then $[x]_{\mathcal{B}} = B^{-1}x$

Corollary 2.5.5: Change of Bases

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation. Let $\mathcal{B} = v_1, \dots, v_n$ be a basis for \mathbb{R}^n where $[T(x)]_{\mathcal{B}} = A_{\mathcal{B}}[x]_{\mathcal{B}}$ and $\mathcal{B}' = v'_1, \dots, v'_n$ also be a basis for \mathbb{R}^n where $[T(x)]_{\mathcal{B}'} = A_{\mathcal{B}'}[x]_{\mathcal{B}'}$.
 $A_{\mathcal{B}'} = S^{-1}A_{\mathcal{B}}S$

where $S = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}^{-1} \begin{bmatrix} v'_1 & \dots & v'_n \end{bmatrix}$ is invertible and $[x]_{\mathcal{B}'} = S^{-1}[x]_{\mathcal{B}}$

Proof

Let $T(x) = Ax$. By **theorem 2.5.4**, for invertible $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$ and $B' = \begin{bmatrix} v'_1 & \dots & v'_n \end{bmatrix}$:

$$A = BA_{\mathcal{B}}B^{-1} \quad A = B'A_{\mathcal{B}'}(B')^{-1}$$

$$BA_{\mathcal{B}}B^{-1} = B'A_{\mathcal{B}'}(B')^{-1} \Rightarrow A_{\mathcal{B}'} = (B')^{-1}BA_{\mathcal{B}}B^{-1}B' = (B^{-1}B')^{-1}A_{\mathcal{B}}(B^{-1}B')$$

Since v_1, \dots, v_n and u_1, \dots, u_n are bases, then by **theorem 1.5.13**, $B^{-1}B'$ is invertible.

Since $x = B[x]_{\mathcal{B}}$ and $x = B'[x]_{\mathcal{B}'}$, then:

$$[x]_{\mathcal{B}'} = (B')^{-1}x = (B')^{-1}B[x]_{\mathcal{B}} = (B^{-1}B')^{-1}[x]_{\mathcal{B}}$$

Definition 2.5.6: Matrix Similarity

Let $A, B \in M_{n \times n}(\mathbb{R})$. Then A is **similar** to B if there is an invertible $X \in M_{n \times n}(\mathbb{R})$ where:
 $AX = XB$

Theorem 2.5.7: Properties of Matrix Similarity

Let $A, B, C \in M_{n \times n}(\mathbb{R})$.

(a) **Reflexivity**

A is similar to A

Proof

$$AI_{n \times n} = A = I_{n \times n}A$$

(b) **Symmetry**

If A is similar to B , then B is similar to A

Proof

$$\begin{aligned} \text{Since } A \text{ is similar to } B, \text{ there is an invertible } X \in M_{n \times n}(\mathbb{R}) \text{ where } AX = XB. \\ BX^{-1} = X^{-1}XBX^{-1} = X^{-1}AXX^{-1} = X^{-1}A \end{aligned}$$

(c) **Transitivity**

If A is similar to B and B is similar to C , then A is similar to C

Proof

$$\begin{aligned} \text{Since } A \text{ is similar to } B, \text{ there is an invertible } X_1 \in M_{n \times n}(\mathbb{R}) \text{ where } AX_1 = X_1B. \\ \text{Since } B \text{ is similar to } C, \text{ there is an invertible } X_2 \in M_{n \times n}(\mathbb{R}) \text{ where } BX_2 = X_2C. \\ AX_1X_2 = X_1BX_2 = X_1X_2C \\ \text{Since } X_1, X_2 \text{ are invertible, then by theorem 1.5.13, } X_1X_2 \text{ is invertible.} \end{aligned}$$

Example

Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be $T(x) = Ax = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 3 & -9 & 6 \end{bmatrix} x$ where $\mathcal{B} = (1,1,1)$, $(1,2,3)$, and $(1,3,6)$ is a basis for \mathbb{R}^3 . Find $A_{\mathcal{B}}$. Find and verify $[T(-1,0,1)]_{\mathcal{B}}$.

$$A_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 3 & -9 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$[(-1,0,1)]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

$$[T(-1,0,1)]_{\mathcal{B}} = A_{\mathcal{B}}[(-1,0,1)]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = (0,1,0)$$

Since $0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix} = (1,2,3)$ and $T(-1,0,1) = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 3 & -9 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = (1,2,3)$, then it is true that $[T(-1,0,1)]_{\mathcal{B}} = (0,1,0)$.

3 Orthogonality

3.1 Orthogonality

Definition 3.1.1: Dot Product

Let $x, y \in \mathbb{R}^n$. Then the **dot product** of x and y :

$$x \cdot y = x_1y_1 + \dots + x_ny_n = \sum_{i=1}^n x_iy_i$$

Theorem 3.1.2: Properties of the Dot Product

Let $x, y, z \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

(a) Positive Definite

$x \cdot x \geq 0$ where $x \cdot x = 0$ if and only if $x = 0$

Proof

$$x \cdot x = x_1x_1 + \dots + x_nx_n = x_1^2 + \dots + x_n^2 \geq 0$$

If $0 = x \cdot x = x_1^2 + \dots + x_n^2$, then each $x_i = 0$ so $x = 0$.

If $x = 0$, then $x \cdot x = x_1^2 + \dots + x_n^2 = 0 + \dots + 0 = 0$.

(b) Symmetry

$$x \cdot y = y \cdot x$$

Proof

$$x \cdot y = x_1y_1 + \dots + x_ny_n = y_1x_1 + \dots + y_nx_n = y \cdot x$$

(c) Scalar Multiplication

$$(cx) \cdot y = c(x \cdot y) = x \cdot (cy)$$

Proof

$$(cx) \cdot y = \sum_{i=1}^n cx_iy_i = c \sum_{i=1}^n x_iy_i = c(x \cdot y) = \sum_{i=1}^n x_i(cy_i) = x \cdot (cy)$$

(d) Distributivity

$$z \cdot (x + y) = (z \cdot x) + (z \cdot y) \quad (x + y) \cdot z = (x \cdot z) + (y \cdot z)$$

Proof

$$\begin{aligned} z \cdot (x + y) &= \sum_{i=1}^n z_i(x_i + y_i) = \sum_{i=1}^n z_ix_i + \sum_{i=1}^n z_iy_i = (z \cdot x) + (z \cdot y) \\ (x + y) \cdot z &= \sum_{i=1}^n (x_i + y_i)z_i = \sum_{i=1}^n x_iz_i + \sum_{i=1}^n y_iz_i = (x \cdot z) + (y \cdot z) \end{aligned}$$

Theorem 3.1.3: Dot Product: Length Property

Let $x \in \mathbb{R}^n$. Then, $x \cdot x = |x|^2$.

Proof

$$x \cdot x = x_1x_1 + \dots + x_nx_n = \sum_{i=1}^n x_i^2 = (\sqrt{\sum_{i=1}^n x_i^2})^2 = |x|^2.$$

Theorem 3.1.4: Dot Product: Cancellation Property

Let $x, y \in \mathbb{R}^n$. Then, $x = y$ if and only if $x \cdot z = y \cdot z$ for every $z \in \mathbb{R}^n$.

Proof

Suppose $x = y$. Then $x_i = y_i$ for $i = \{1, \dots, n\}$:

$$x \cdot z = x_1z_1 + \dots + x_nz_n = y_1z_1 + \dots + y_nz_n = y \cdot z$$

Suppose $x \cdot z = y \cdot z$ for every $z \in \mathbb{R}^n$. Then, $z \cdot (x - y) = 0$. Let $z = x - y$. Then:

$$0 = z \cdot (x - y) = (x - y) \cdot (x - y)$$

Thus, $x - y = 0$ so $x = y$.

Definition 3.1.5: Transpose

Let $x \in \mathbb{R}^n$ where $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. Then, the **transpose** of x :

$$x^T = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}$$

Let $A \in M_{m \times n}(\mathbb{R})$ where $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$. Then, the transpose of A :

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \in M_{n \times m}(\mathbb{R})$$

Theorem 3.1.6: Properties of Transpose

Let $A, B \in M_{m \times n}(\mathbb{R})$ and $c \in \mathbb{R}$. Then:

(a) **Addition**

$$(A + B)^T = A^T + B^T$$

Proof

$$(A + B)^T = \begin{bmatrix} a_{11} + b_{11} & a_{21} + b_{21} & \dots & a_{m1} + b_{m1} \\ a_{12} + b_{12} & a_{22} + b_{22} & \dots & a_{m2} + b_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} + b_{1n} & a_{2n} + b_{2n} & \dots & a_{mn} + b_{mn} \end{bmatrix} = A^T + B^T$$

(b) **Scalar Multiplication**

$$(cA)^T = cA^T$$

Proof

$$(cA)^T = \begin{bmatrix} ca_{11} & ca_{21} & \dots & ca_{m1} \\ ca_{12} & ca_{22} & \dots & ca_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{1n} & ca_{2n} & \dots & ca_{mn} \end{bmatrix} = cA^T$$

Theorem 3.1.7: Dot Product: Transpose Property

Let $x, y \in \mathbb{R}^n$. Then:

$$x \cdot y = x^T y = y^T x$$

Proof

$$x \cdot y = x_1 y_1 + \dots + x_n y_n = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x^T y$$

$$x \cdot y = y \cdot x = y^T x$$

Theorem 3.1.8: Transpose of Matrix Product

Let $A \in M_{m \times n}(\mathbb{R})$ and $B \in M_{t \times m}(\mathbb{R})$. Then, $(BA)^T \in M_{n \times t}(\mathbb{R})$ where:
 $(BA)^T = A^T B^T$

Proof

$$\begin{aligned}
 (BA)^T &= \left(\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{t1} & b_{t2} & \cdots & b_{tm} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \right)^T \\
 &= \left(\begin{bmatrix} \sum_{k=1}^m b_{1k} a_{k1} & \sum_{k=1}^m b_{1k} a_{k2} & \cdots & \sum_{k=1}^m b_{1k} a_{kn} \\ \sum_{k=1}^m b_{2k} a_{k1} & \sum_{k=1}^m b_{2k} a_{k2} & \cdots & \sum_{k=1}^m b_{2k} a_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^m b_{tk} a_{k1} & \sum_{k=1}^m b_{tk} a_{k2} & \cdots & \sum_{k=1}^m b_{tk} a_{kn} \end{bmatrix} \right)^T \\
 &= \begin{bmatrix} \sum_{k=1}^m b_{1k} a_{k1} & \sum_{k=1}^m b_{2k} a_{k1} & \cdots & \sum_{k=1}^m b_{tk} a_{k1} \\ \sum_{k=1}^m b_{1k} a_{k2} & \sum_{k=1}^m b_{2k} a_{k2} & \cdots & \sum_{k=1}^m b_{tk} a_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^m b_{1k} a_{kn} & \sum_{k=1}^m b_{2k} a_{kn} & \cdots & \sum_{k=1}^m b_{tk} a_{kn} \end{bmatrix} \in M_{n \times t}(\mathbb{R}) \\
 &= \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{t1} \\ b_{12} & b_{22} & \cdots & b_{t2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1m} & b_{2m} & \cdots & b_{tm} \end{bmatrix} = A^T B^T
 \end{aligned}$$

Corollary 3.1.9: $(A^T)^{-1} = (A^{-1})^T$

For invertible $A \in M_{n \times n}(\mathbb{R})$, then A^T is invertible where:

$$(A^T)^{-1} = (A^{-1})^T$$

Proof

Since A is invertible, there exists a A^{-1} such that $A^{-1}A = I_{n \times n}$ and $AA^{-1} = I_{n \times n}$
 $(A^{-1})^T A^T = (AA^{-1})^T = I_{n \times n}^T = I_{n \times n}$ $A^T (A^{-1})^T = (A^{-1}A)^T = I_{n \times n}^T = I_{n \times n}$
 Thus, $(A^T)^{-1} = (A^{-1})^T$ so A^T is invertible.

Corollary 3.1.10: $(Ax) \cdot y = x \cdot (A^T y)$

For $A \in M_{m \times n}(\mathbb{R})$ and $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, then $A^T \in M_{n \times m}(\mathbb{R})$ is unique such that:

$$(Ax) \cdot y = x \cdot (A^T y)$$

Proof

Since $Ax \in \mathbb{R}^m$, then:
 $(Ax) \cdot y = (Ax)^T y = (x^T A^T) y = x^T (A^T y) = x \cdot (A^T y)$
 Suppose there is a $B \in M_{n \times m}(\mathbb{R})$ such that $(Ax) \cdot y = x \cdot (By)$.
 Then for any $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$:
 $x \cdot (By) = (Ax) \cdot y = x \cdot (A^T y)$
 By **theorem 3.1.4**, $By = A^T y$ so $B = A^T$.

Definition 3.1.11: Orthogonality

Vectors $x, y \in \mathbb{R}^n$ are **orthogonal** (i.e. perpendicular) if $x \cdot y = 0$

Theorem 3.1.12: Pythagorean Theorem in higher dimensions

Let $x, y \in \mathbb{R}^n$. Then the following are equivalent:

- (a) $x \cdot y = 0$
- (b) $|x|^2 + |y|^2 = |x + y|^2$
- (c) If $n = 2$ or 3 and $x, y \neq 0$, then x and y are perpendicular

Proof

Note $|x + y|^2 = (x + y) \cdot (x + y) = (x \cdot x) + (x \cdot y) + (y \cdot x) + (y \cdot y) = |x|^2 + |y|^2 + 2(x \cdot y)$.

Suppose $x \cdot y = 0$. Then, $|x|^2 + |y|^2 = |x + y|^2$.

Suppose $|x|^2 + |y|^2 = |x + y|^2$. Then $2(x \cdot y) = 0$ so $x \cdot y = 0$.

Suppose $|x|^2 + |y|^2 = |x + y|^2$. By the regular definition of the Pythagorean theorem in $\mathbb{R}^2, \mathbb{R}^3$, x , y , and $x+y$ form a right triangle so x and y form a right angle and thus, perpendicular. Note if $x = 0$, then $|x|^2 + |y|^2 = |x + y|^2$ becomes $|y|^2 = |y|^2$ which is true, but has nothing to do with right triangles so exclude the case when $x = 0$. Similarly, exclude $y = 0$.

Theorem 3.1.13: $x \cdot y = |x||y| \cos(\theta)$

Let $x, y \in \mathbb{R}^n$. Then:

$$x \cdot y = |x||y| \cos(\theta)$$

where $\theta \in [0, \pi]$ is the angle between x and y

Proof

Since x , y , and $x-y$ form a triangle, by the Law of Cosine:

$$|x - y|^2 = |x|^2 + |y|^2 - 2|x||y| \cos(\theta)$$

where $\theta \in [0, \pi]$ is the angle between x and y . Since:

$$|x - y|^2 = (x - y) \cdot (x - y) = x \cdot x + y \cdot y - 2(x \cdot y) = |x|^2 + |y|^2 - 2(x \cdot y)$$

then $x \cdot y = |x||y| \cos(\theta)$.

3.2 Orthogonal Basis

Theorem 3.2.1: Orthogonal Projection

The **orthogonal projection** of $x \in \mathbb{R}^n$ onto $y \in \mathbb{R}^n$ is the component of x parallel to y :

$$\text{proj}_y x = \frac{x \cdot y}{|y|^2} y$$

Also, $\text{proj}_y x \in \text{span}(y)$ where $x - \text{proj}_y x$ and $\text{span}(y)$ are orthogonal.

Proof

Since $\text{proj}_y x$ is parallel to y , let $\text{proj}_y x = cy$ for some $c \in \mathbb{R}$.

Let x^\perp be the orthogonal component of x to y . Thus, $x = \text{proj}_y x + x^\perp = cy + x^\perp$.

Since x^\perp is orthogonal to y , then:

$$x \cdot y = (cy + x^\perp) \cdot y = cy \cdot y + x^\perp \cdot y = cy \cdot y = c|y|^2$$

Thus, $c = \frac{x \cdot y}{|y|^2}$ so $\text{proj}_y x = cy = \frac{x \cdot y}{|y|^2} y \in \text{span}(y)$. Let $a \in \mathbb{R}$.

$$\begin{aligned} (x - \text{proj}_y x) \cdot ay &= (x - \frac{x \cdot y}{|y|^2} y) \cdot ay = (x \cdot ay) - \frac{x \cdot y}{|y|^2} (y \cdot ay) \\ &= a(x \cdot y) - a \frac{x \cdot y}{|y|^2} |y|^2 = a(x \cdot y) - a(x \cdot y) = 0 \end{aligned}$$

Thus, $x - \text{proj}_y x$ and $\text{span}(y)$ are orthogonal so $x - \text{proj}_y x$ and $\text{proj}_y x$ are orthogonal.

Example

Let $x = (-4, 5, 7)$ and $y = (2, -4, 1)$. Find the vector decomposition of x onto y .

Parallel component: $\text{proj}_y x = \frac{(-4 \cdot 2 + 5 \cdot (-4) + 7 \cdot 1)}{2^2 + (-4)^2 + 1^2} (2, -4, 1) = \frac{-21}{21} (2, -4, 1) = (-2, 4, -1)$

Orthogonal component: $x - \text{proj}_y x = (-4, 5, 7) - (-2, 4, -1) = (-2, 1, 8)$

Theorem 3.2.2: Cauchy-Schwarz Inequality

For $x, y \in \mathbb{R}^n$, $|x \cdot y| \leq |x||y|$

Proof

Let $x = \text{proj}_y x + x^\perp = cy + x^\perp$ where x^\perp is the orthogonal component of x to y and $\text{proj}_y x = cy$ is the parallel component of x to y . By **theorem 3.2.1**, then $c = \frac{x \cdot y}{|y|^2}$.

$$|x|^2 = |cy + x^\perp|^2 = |cy|^2 + |x^\perp|^2 = \left(\frac{x \cdot y}{|y|^2}\right)^2 |y|^2 + |x^\perp|^2$$

$$|x|^2 |y|^2 = |y|^2 \left(\frac{x \cdot y}{|y|^2}\right)^2 |y|^2 + |y|^2 |x^\perp|^2 = (x \cdot y)^2 + |y|^2 |x^\perp|^2$$

Since $|y|^2 |x^\perp|^2 \geq 0$, then $(x \cdot y)^2 \leq |x|^2 |y|^2$ so $|x \cdot y| \leq |x||y|$.

Corollary 3.2.3: Triangle Inequality

For $x, y \in \mathbb{R}^n$, $\|x + y\| \leq \|x\| + \|y\|$

Proof

$$|x + y|^2 = (x + y) \cdot (x + y) = x \cdot x + y \cdot y + 2(x \cdot y) = |x|^2 + |y|^2 + 2(x \cdot y)$$

$$\leq |x|^2 + |y|^2 + 2|x \cdot y| \leq |x|^2 + |y|^2 + 2|x||y| = (|x| + |y|)^2$$
Definition 3.2.4: Orthonormal Vectors

Vectors $v_1, \dots, v_k \in \mathbb{R}^n$ is **orthogonal** if:

$$v_i \cdot v_j = 0 \text{ for } i \neq j$$

Then, $v_1, \dots, v_k \in \mathbb{R}^n$ is **orthonormal** if:

$$v_i \cdot v_j = 0 \text{ for } i \neq j \quad |v_i|^2 = v_i \cdot v_i = 1 \quad \Leftrightarrow \quad |v_i| = 1 \text{ for } i \in \{1, \dots, k\}$$

Theorem 3.2.5: Orthogonal sets are Linearly independent

Let $v_1, \dots, v_k \in \mathbb{R}^n$ be orthogonal. Then, v_1, \dots, v_k is linearly independent.

Proof

Let $c_1, \dots, c_k \in \mathbb{R}$ such that $0 = c_1 v_1 + \dots + c_k v_k$. Since $v_i \cdot v_j = 0$ for $i \neq j$, then:

$$0 = v_i \cdot 0 = v_i \cdot (c_1 v_1 + \dots + c_k v_k) = c_i (v_i \cdot v_i) = c_i |v_i|^2$$

Since $|v_i| > 0$, then $c_i = 0$. Since every $c_i = 0$, then v_1, \dots, v_k is linearly independent.

Theorem 3.2.6: Orthogonal Basis

For $V \subset \mathbb{R}^n$, let $v_1, \dots, v_k \in \mathbb{R}^n$ be an orthogonal basis for V . Then, for each $x \in V$:

$$x = \text{proj}_{v_1} x + \dots + \text{proj}_{v_k} x = \left(\frac{x \cdot v_1}{|v_1|^2}\right) v_1 + \dots + \left(\frac{x \cdot v_k}{|v_k|^2}\right) v_k$$

Then if v_1, \dots, v_k is an orthonormal basis, then:

$$x = (x \cdot v_1) v_1 + \dots + (x \cdot v_k) v_k$$

Proof

Since v_1, \dots, v_k is a basis for V , then for any $x \in V$, there are $c_1, \dots, c_k \in \mathbb{R}$ such that:

$$x = c_1 v_1 + \dots + c_k v_k$$

Since v_1, \dots, v_k is orthogonal, then $v_i \cdot v_j = 0$ for $i \neq j$. Thus:

$$v_i \cdot x = v_i \cdot (c_1 v_1 + \dots + c_k v_k) = c_i (v_i \cdot v_i) = c_i |v_i|^2 \quad \Rightarrow \quad c_i = \frac{x \cdot v_i}{|v_i|^2}$$

Since each $c_i = \frac{x \cdot v_i}{|v_i|^2}$, then:

$$x = c_1 v_1 + \dots + c_k v_k = \left(\frac{x \cdot v_1}{|v_1|^2}\right) v_1 + \dots + \left(\frac{x \cdot v_k}{|v_k|^2}\right) v_k = \text{proj}_{v_1} x + \dots + \text{proj}_{v_k} x$$

Additionally, if v_1, \dots, v_k is orthonormal, then each $|v_i| = 1$. Thus, $x = (x \cdot v_1) v_1 + \dots + (x \cdot v_k) v_k$.

Definition 3.2.7: Orthogonal Complement

For $V \subset \mathbb{R}^n$, the **orthogonal complement** of V :

$$V^\perp = \{ w \in \mathbb{R}^n : w \cdot v = 0 \text{ for } v \in V \}$$

Theorem 3.2.8: Orthogonal projection onto $V \subset \mathbb{R}^n$

Let $V \subset \mathbb{R}^n$. Then for any $x \in \mathbb{R}^n$:

$$x = x^{\parallel} + x^{\perp}$$

where $x^{\parallel} = \text{proj}_V x$, the **orthogonal projection** of x onto V , is the component of x in V and x^{\perp} is the component of x orthogonal to V are both unique

Proof

Since x^{\parallel} is in V and x^{\perp} is orthogonal to V , then $x^{\parallel} \cdot x^{\perp} = 0$.

Let $v_1, \dots, v_k \in \mathbb{R}^n$ be an orthogonal basis for V . Then by **theorem 3.2.6**, since $x^{\parallel} \in V$, then:

$$x^{\parallel} = \text{proj}_{v_1} x^{\parallel} + \dots + \text{proj}_{v_k} x^{\parallel}$$

Suppose $x = x_1^{\parallel} + x_1^{\perp} = x_2^{\parallel} + x_2^{\perp}$ for $x_1^{\parallel}, x_2^{\parallel}, x_1^{\perp}, x_2^{\perp} \in \mathbb{R}^n$. Then:

$$(x_1^{\parallel} - x_2^{\parallel}) + (x_1^{\perp} - x_2^{\perp}) = x_1^{\parallel} + x_1^{\perp} - x_2^{\parallel} - x_2^{\perp} = x - x = 0$$

Since $x_1^{\parallel}, x_2^{\parallel} \in V$ and $x_1^{\perp}, x_2^{\perp} \in V^{\perp}$, then:

$$0 = 0 \cdot (x_1^{\parallel} - x_2^{\parallel}) = [(x_1^{\parallel} - x_2^{\parallel}) + (x_1^{\perp} - x_2^{\perp})] \cdot (x_1^{\parallel} - x_2^{\parallel}) = (x_1^{\parallel} - x_2^{\parallel}) \cdot (x_1^{\parallel} - x_2^{\parallel}) = |x_1^{\parallel} - x_2^{\parallel}|^2$$

$$|x_1^{\parallel} - x_2^{\parallel}| = 0 \quad \Rightarrow \quad x_1^{\parallel} = x_2^{\parallel}$$

$$0 = 0 \cdot (x_1^{\perp} - x_2^{\perp}) = [(x_1^{\parallel} - x_2^{\parallel}) + (x_1^{\perp} - x_2^{\perp})] \cdot (x_1^{\perp} - x_2^{\perp}) = (x_1^{\perp} - x_2^{\perp}) \cdot (x_1^{\perp} - x_2^{\perp}) = |x_1^{\perp} - x_2^{\perp}|^2$$

$$|x_1^{\perp} - x_2^{\perp}| = 0 \quad \Rightarrow \quad x_1^{\perp} = x_2^{\perp}$$

Thus, x^{\parallel}, x^{\perp} are unique.

Theorem 3.2.9: Gram-Schmidt Process: Creating an Orthogonal Basis

For $V \subset \mathbb{R}^n$, let $u_1, \dots, u_k \in \mathbb{R}^n$ be a basis for V . Then let $v_1, \dots, v_k \in \mathbb{R}^n$:

$$v_1 = u_1$$

$$v_2 = u_2 - \text{proj}_{v_1}(u_2)$$

$$v_3 = u_3 - \text{proj}_{v_1}(u_3) - \text{proj}_{v_2}(u_3)$$

$$\vdots$$

$$v_k = u_k - \text{proj}_{v_1}(u_k) - \dots - \text{proj}_{v_{k-1}}(u_k)$$

Then, v_1, \dots, v_k is an orthogonal basis for V and $\frac{v_1}{|v_1|}, \dots, \frac{v_k}{|v_k|}$ is an orthonormal basis for V .

Proof

$$v_1 \cdot v_2 = u_1 \cdot (u_2 - \text{proj}_{v_1}(u_2)) = u_1 \cdot (u_2 - \frac{u_2 \cdot u_1}{|u_1|^2} u_1) = u_1 \cdot u_2 - u_1 \cdot u_2 = 0$$

Suppose for $m \leq k$, then v_1, \dots, v_{m-1} are orthogonal.

Since $v_m = u_m - \text{proj}_{v_1}(u_m) - \dots - \text{proj}_{v_{m-1}}(u_m)$, where by **theorem 3.2.1**, each $\text{proj}_{v_i}(u_m) = c_i v_i$ where $c_i = \frac{u_m \cdot v_i}{|v_i|^2}$ for $i = \{1, \dots, m-1\}$:

$$v_m = u_m - c_1 v_1 - \dots - c_{m-1} v_{m-1}$$

$$v_i \cdot v_m = v_i \cdot (u_m - c_1 v_1 - \dots - c_{m-1} v_{m-1}) = v_i \cdot u_m - c_i (v_i \cdot v_i) = v_i \cdot u_m - u_m \cdot v_i = 0$$

Thus, v_m is orthogonal to any v_i for $i = \{1, \dots, m-1\}$ so v_1, \dots, v_m is orthogonal. Thus, by proof by induction, v_1, \dots, v_k is orthogonal. Also, $v_1 = u_1 \in V$. Suppose for $m \leq k$, then $v_1, \dots, v_{m-1} \in V$. Then, $v_m = u_m - c_1 v_1 - \dots - c_{m-1} v_{m-1} \in V$. Thus, by proof by induction, $v_1, \dots, v_k \in V$. By **theorem 3.2.5**, v_1, \dots, v_k are linearly independent. Since u_1, \dots, u_k is a basis for V , then by **corollary 2.3.3**, $\dim(V) = k$. Then, by **theorem 2.3.5**, v_1, \dots, v_k span V and thus, form a basis for V . Thus, v_1, \dots, v_k is an orthogonal basis for V .

Since each $|\frac{v_i}{|v_i|}| = \frac{1}{|v_i|}|v_i| = 1$, then $\frac{v_1}{|v_1|}, \dots, \frac{v_k}{|v_k|}$ is an orthonormal basis for V .

Theorem 3.2.10: QR Factorization: Relationship between Basis and Orthonormal Basis

Let $u_1, \dots, u_k \in \mathbb{R}^n$ be linearly independent such that $A = \begin{bmatrix} u_1 & \dots & u_k \end{bmatrix} \in M_{n \times k}(\mathbb{R})$. Then, there is an orthogonal $v_1, \dots, v_k \in \mathbb{R}^n$ where $Q = \begin{bmatrix} \frac{v_1}{|v_1|} & \dots & \frac{v_k}{|v_k|} \end{bmatrix} \in M_{n \times k}(\mathbb{R})$ such that:
 $A = QR$

where $R \in M_{k \times k}(\mathbb{R})$ is invertible with $R = \begin{bmatrix} |v_1| & \frac{u_2 \cdot v_1}{|v_1|} & \frac{u_3 \cdot v_1}{|v_1|} & \dots & \frac{u_k \cdot v_1}{|v_1|} \\ 0 & |v_2| & \frac{u_3 \cdot v_2}{|v_2|} & \dots & \frac{u_k \cdot v_2}{|v_2|} \\ 0 & 0 & |v_3| & \dots & \frac{u_k \cdot v_3}{|v_3|} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & |v_k| \end{bmatrix}$

Proof

Let $V = \text{span}(u_1, \dots, u_k)$. Then by [theorem 2.3.5](#), u_1, \dots, u_k form a basis for V .

Thus, by [theorem 3.2.9](#), there is an orthonormal basis $\frac{v_1}{|v_1|}, \dots, \frac{v_k}{|v_k|}$ for V where $v_1 = u_1$ and $v_i = u_i - \text{proj}_{\text{span}(v_1, \dots, v_{i-1})} u_i$ for $i = \{2, \dots, k\}$.

By [theorem 3.2.6](#), each $\text{proj}_{\text{span}(v_1, \dots, v_{i-1})} u_i = \frac{u_i \cdot v_1}{|v_1|^2} v_1 + \dots + \frac{u_i \cdot v_{i-1}}{|v_{i-1}|^2} v_{i-1}$.

Thus, $u_i = \frac{u_i \cdot v_1}{|v_1|^2} v_1 + \dots + \frac{u_i \cdot v_{i-1}}{|v_{i-1}|^2} v_{i-1} + v_i$. Then:

$$\begin{bmatrix} u_1 & \dots & u_k \end{bmatrix} = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \begin{bmatrix} 1 & \frac{u_2 \cdot v_1}{|v_1|^2} & \frac{u_3 \cdot v_1}{|v_1|^2} & \dots & \frac{u_k \cdot v_1}{|v_1|^2} \\ 0 & 1 & \frac{u_3 \cdot v_2}{|v_2|^2} & \dots & \frac{u_k \cdot v_2}{|v_2|^2} \\ 0 & 0 & 1 & \dots & \frac{u_k \cdot v_3}{|v_3|^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = \begin{bmatrix} \frac{v_1}{|v_1|} & \dots & \frac{v_k}{|v_k|} \end{bmatrix} \begin{bmatrix} |v_1| & \frac{u_2 \cdot v_1}{|v_1|} & \frac{u_3 \cdot v_1}{|v_1|} & \dots & \frac{u_k \cdot v_1}{|v_1|} \\ 0 & |v_2| & \frac{u_3 \cdot v_2}{|v_2|} & \dots & \frac{u_k \cdot v_2}{|v_2|} \\ 0 & 0 & |v_3| & \dots & \frac{u_k \cdot v_3}{|v_3|} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & |v_k| \end{bmatrix}$$

$$A = QR$$

By [theorem 3.2.5](#), v_1, \dots, v_k are linearly independent so $|v_i| \neq 0$ else if $|v_i| = 0$, then $v_i = 0$ which is linearly dependent. Thus, $\text{rref}(R) = I_{n \times n}$ so R is invertible.

Theorem 3.2.11: Orthogonal projection matrix transformation

For $V \in \mathbb{R}^n$, let $v_1, \dots, v_k \in \mathbb{R}^n$ be an orthonormal basis for V . Then for any $x \in \mathbb{R}^n$:

$$\text{proj}_V x = AA^T x$$

$$\text{where } A = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \in M_{n \times k}(\mathbb{R})$$

Proof

By [theorem 3.2.8](#):

$$\begin{aligned} \text{proj}_V x &= (x \cdot v_1)v_1 + \dots + (x \cdot v_k)v_k = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \begin{bmatrix} x \cdot v_1 \\ \vdots \\ x \cdot v_k \end{bmatrix} = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \begin{bmatrix} v_1^T x \\ \vdots \\ v_k^T x \end{bmatrix} \\ &= \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix} x = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix}^T x \end{aligned}$$

Theorem 3.2.12: Reflection matrix transformation

For $V \in \mathbb{R}^n$, let $v_1, \dots, v_k \in \mathbb{R}^n$ be an orthonormal basis for V .

Then for any $x \in \mathbb{R}^n$, the [reflection](#) of x across V :

$$\text{reflect}_V x = (2AA^T - I_{n \times n})x$$

$$\text{where } A = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \in M_{n \times k}(\mathbb{R})$$

Proof

Note for $x = x^{\parallel} + x^{\perp}$ where x^{\parallel} is the component of x parallel to V and x^{\perp} is the component of x orthogonal to V , then $\text{reflect}_V x = x^{\parallel} - x^{\perp}$. Thus, by [theorem 3.2.11](#):

$$\text{reflect}_V x = x^{\parallel} - x^{\perp} = 2x^{\parallel} - x = 2AA^T x - x = (2AA^T - I_{n \times n})x$$

Example

Let $T: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ be $T(x) = Ax$ where $A = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 1 & 0 & 4 & 2 \\ 1 & 0 & 3 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$.

Find an orthonormal basis for $V = \text{im}(A)$, QR factorization of A , and proj_V .

First, find a basis for $\text{im}(A)$.

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

$$\text{im}(A) = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 3 \\ 1 \end{bmatrix}\right)$$

To find an orthonormal basis for the $\text{im}(A)$, apply the Gram-Schmidt Process.

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{1^2+1^2+1^2+1^2} \left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{2}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 4 \\ 3 \\ 1 \end{bmatrix} - \frac{1}{1^2+1^2+1^2+1^2} \left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 4 \\ 3 \\ 1 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{(\frac{1}{2})^2+(-\frac{1}{2})^2+(-\frac{1}{2})^2+(\frac{1}{2})^2} \left(\begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 4 \\ 3 \\ 1 \end{bmatrix} \right) \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 3 \\ 1 \end{bmatrix} - \frac{10}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{-2}{1} \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

$$\text{im}(A) = \text{span}\left(\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}\right)$$

Then, the QR factorization of A :

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 4 \\ 1 & 0 & 3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \sqrt{4} & \frac{2}{\sqrt{4}} & \frac{10}{\sqrt{4}} \\ 0 & \sqrt{1} & \frac{-2}{\sqrt{1}} \\ 0 & 0 & \sqrt{1} \end{bmatrix} \Rightarrow \begin{bmatrix} 2 \\ 4 \\ 3 \\ 1 \end{bmatrix} = \frac{10}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{1} \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 4 \\ 1 & 0 & 3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 & 5 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

And the proj_V :

$$\text{proj}_V = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \right)^T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

3.3 Orthogonal Transformations

Definition 3.3.1: Orthogonal Transformation

Linear Transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is **orthogonal** if for any $x \in \mathbb{R}^n$:

$$|T(x)| = |x|$$

Theorem 3.3.2: Orthogonal transformation Equivalences

Let linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be $T(x) = Ax$ for $A \in M_{n \times n}(\mathbb{R})$.

Then the following are equivalent:

- (a) For any $x \in \mathbb{R}^n$, then $|T(x)| = |x|$
- (b) For any $x, y \in \mathbb{R}^n$, then $T(x) \cdot T(y) = x \cdot y$
- (c) $T(e_1), \dots, T(e_n)$ is an orthonormal basis for \mathbb{R}^n
- (d) $A^T A = I_{n \times n}$
- (e) A is invertible where $A^{-1} = A^T$

Proof

Suppose for any $x \in \mathbb{R}$, then $|T(x)| = |x|$.

Thus, for $x, y \in \mathbb{R}^n$, then $|T(x) + T(y)| = |T(x + y)| = |x + y|$. Thus:

$$|T(x) + T(y)|^2 = |x + y|^2 = (x + y) \cdot (x + y) = |x|^2 + |y|^2 + 2(x \cdot y)$$

$$\begin{aligned} |T(x) + T(y)|^2 &= (T(x) + T(y)) \cdot (T(x) + T(y)) \\ &= |T(x)|^2 + |T(y)|^2 + 2(T(x) \cdot T(y)) = |x|^2 + |y|^2 + 2(T(x) \cdot T(y)) \end{aligned}$$

Thus, $T(x) \cdot T(y) = x \cdot y$.

Suppose for any $x, y \in \mathbb{R}^n$, then $T(x) \cdot T(y) = x \cdot y$.

$$T(e_i) \cdot T(e_j) = e_i \cdot e_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Thus, $T(e_1), \dots, T(e_n)$ are orthogonal. By **theorem 3.2.5**, $T(e_1), \dots, T(e_n)$ are linearly independent. Since $\dim(\mathbb{R}^n) = n$, then by **theorem 2.3.5**, $T(e_1), \dots, T(e_n)$ is a basis.

Suppose $T(e_1), \dots, T(e_n)$ is an orthonormal basis for \mathbb{R}^n .

Since $A = \begin{bmatrix} T(e_1) & \dots & T(e_n) \end{bmatrix}$ and $T(e_i) \cdot T(e_j) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$, then:

$$\begin{aligned} A^T A &= \begin{bmatrix} T(e_1)^T \\ \vdots \\ T(e_n)^T \end{bmatrix} \begin{bmatrix} T(e_1) & \dots & T(e_n) \end{bmatrix} \\ &= \begin{bmatrix} T(e_1)^T T(e_1) & T(e_1)^T T(e_2) & \dots & T(e_1)^T T(e_n) \\ T(e_2)^T T(e_1) & T(e_2)^T T(e_2) & \dots & T(e_2)^T T(e_n) \\ \vdots & \vdots & \ddots & \vdots \\ T(e_n)^T T(e_1) & T(e_n)^T T(e_2) & \dots & T(e_n)^T T(e_n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \end{aligned}$$

Suppose $A^T A = I_{n \times n}$.

Then by **theorem 1.5.11**, then A, A^T are invertible where $A^{-1} = A^T$.

Suppose A is invertible where $A^{-1} = A^T$.

Then for any $x \in \mathbb{R}^n$:

$$\begin{aligned} |T(x)|^2 &= T(x) \cdot T(x) = Ax \cdot Ax = (Ax)^T (Ax) = x^T A^T Ax \\ &= x^T A^{-1} Ax = x^T x = x \cdot x = |x|^2 \end{aligned}$$

Thus, $|T(x)| = |x|$.

Corollary 3.3.3: Inverse and Transpose of an Orthogonal matrix is Orthogonal

Let $A \in M_{n \times n}(\mathbb{R})$ be orthogonal. Then, A^{-1}, A^T are orthogonal.

Proof

By **theorem 3.3.2**, A is invertible so A^{-1} is invertible. Since A^{-1} is invertible, then $|A^{-1}(x)| = |x|$ so A^{-1} is orthogonal. Since $A^{-1} = A^T$, then A^T is orthogonal.

Corollary 3.3.4: Products of Orthogonal matrices are Orthogonal

Let $A, B \in M_{n \times n}(\mathbb{R})$ be orthogonal. Then, AB is orthogonal.

Proof

By **theorem 3.3.2**, A, B are invertible. Then by **theorem 1.5.13**, AB is invertible so AB is orthogonal.

4 Determinants

4.1 Determinant Function

Definition 4.1.1: Determinant Function

Function $D: M_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$ is a **determinant function** if for any $A \in M_{n \times n}(\mathbb{R})$ such that $A = [A_1 \ \dots \ A_n]$, then:

- (a) **Multilinearity**: For any $j \in \{1, \dots, n\}$, then $T: \mathbb{R}^n \rightarrow \mathbb{R}$:

$$T(x) = D(A_1, \dots, A_{j-1}, x, A_{j+1}, \dots, A_n)$$

is linear. So, for $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$:

- $T(x+y) = T(x) + T(y)$
- $T(cx) = cT(x)$

- (b) **Alternating**: If $B \in M_{n \times n}(\mathbb{R})$ is obtained by swapping two columns of A , then:

$$D(B) = -D(A)$$

Thus, if two columns are equal, then by swapping those columns, $B = A$.

$$D(A) = D(B) = -D(A) \Rightarrow D(A) = 0$$

- (c) **Identity**

$$D(I_{n \times n}) = 1$$

Since each $A_i = A_{i1}e_1 + \dots + A_{in}e_n$, then by multilinearity:

$$\begin{aligned} D(A) &= D(A_1, A_2, \dots, A_n) \\ &= \sum_{i_1=1}^n A_{1i_1} D(e_{i_1}, A_2, \dots, A_n) \\ &= \sum_{i_1=1}^n \sum_{i_2=1}^n A_{1i_1} A_{2i_2} D(e_{i_1}, e_{i_2}, \dots, A_n) \\ &\vdots \\ &= \sum_{i_1=1}^n \dots \sum_{i_n=1}^n A_{1i_1} \dots A_{ni_n} D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) \end{aligned}$$

By alternating, if $i_{k_1} = i_{k_2}$ for $k_1, k_2 \in \{1, \dots, n\}$, then $D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = 0$. Thus:

$$D(A) = \sum_{\{i_1, \dots, i_n\} = \{1, \dots, n\}} A_{1i_1} \dots A_{ni_n} D(e_{i_1}, e_{i_2}, \dots, e_{i_n})$$

where each i_k is unique from $\{1, \dots, n\}$. Since each A_{ki_k} for $k \in \{1, \dots, n\}$ are different columns and each i_k is unique, then each A_{ki_k} is from a unique row and column from A :

$$A = [A_1 \ \dots \ A_n] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \Rightarrow A_{ki_k} = a_{ij}$$

Since $(e_{i_1}, e_{i_2}, \dots, e_{i_n})$ is a rearrangement of (e_1, \dots, e_n) , then by alternating:

$$D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = (-1)^S D(e_1, \dots, e_n) = (-1)^S D(I_{n \times n}) = (-1)^S$$

where S is the number of swaps to turn $(e_{i_1}, e_{i_2}, \dots, e_{i_n})$ into (e_1, \dots, e_n) . So, $D(A)$ is unique:

$$D(A) = \sum_{\{i_1, \dots, i_n\} = \{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^S$$

Definition 4.1.2: Inversions

A $n \times n$ **pattern** $P = \{(i_1, j_1), \dots, (i_n, j_n)\}$ where $\{i_1, \dots, i_n\}, \{j_1, \dots, j_n\}$ are rearrangements of $\{1, \dots, n\}$. Then, (i_{k_1}, j_{k_1}) and (i_{k_2}, j_{k_2}) where $j_{k_1} < j_{k_2}$ is an **inversion** if $i_{k_1} > i_{k_2}$.

Then, the **signature** of P :

$$\text{sgn}(P) = (-1)^I \quad \text{where } I \text{ is the total number of inversions}$$

Theorem 4.1.3: $D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = \text{sgn}(e_{i_1}, e_{i_2}, \dots, e_{i_n})$

Let $e_{i_1}, e_{i_2}, \dots, e_{i_n} \in \mathbb{R}^n$ be a rearrangement of $e_1, \dots, e_n \in \mathbb{R}^n$. Let S be the number of swaps to turn $e_{i_1}, e_{i_2}, \dots, e_{i_n}$ into e_1, \dots, e_n and I be the number of inversions in $(1, i_1), \dots, (n, i_n)$. Then, $S = I$ so:

$$D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = (-1)^S = (-1)^I = \text{sgn}(e_{i_1}, e_{i_2}, \dots, e_{i_n})$$

Proof

Note $D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = (-1)^S$ and $\text{sgn}(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = (-1)^I$. Suppose $n = 2$. Then, either:

$$\begin{aligned} D\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right): & \quad 0 \text{ swaps, } 0 \text{ inversions} \\ D\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right): & \quad 1 \text{ swap, } 1 \text{ inversion} \end{aligned}$$

Thus, when $n = 2$, then $S = I$ so $D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = (-1)^S = (-1)^I = \text{sgn}(e_{i_1}, e_{i_2}, \dots, e_{i_n})$.

For $k \leq n$, let $e_{i_1}, e_{i_2}, \dots, e_{i_{k-1}}$ have the same number of swaps as for inversions.

Then, for $e_{i_1}, e_{i_2}, \dots, e_{i_{k-1}}, e_{i_k}$, the only possible new inversions compared to $e_{i_1}, e_{i_2}, \dots, e_{i_{k-1}}$ are e_{i_k} with each of $e_{i_1}, e_{i_2}, \dots, e_{i_{k-1}}$. Also, note the only new swaps are also e_{i_k} with each of $e_{i_1}, e_{i_2}, \dots, e_{i_{k-1}}$ since if $e_{i_1}, e_{i_2}, \dots, e_{i_{k-1}}$ is swapped into $e_{i_1}^*, e_{i_2}^*, \dots, e_{i_{k-1}}^*$ which has no inversions, then e_{i_k} can swap with $e_{i_{k-1}}^*$ first if needed, then $e_{i_{k-2}}^*$ second if needed, and etc. Thus, each $e_{i_1}, e_{i_2}, \dots, e_{i_{k-1}}$ does not have to swap with one another, the only swaps are e_{i_k} with $e_{i_1}, e_{i_2}, \dots, e_{i_{k-1}}$, and $e_{i_1}^*, e_{i_2}^*, \dots, e_{i_{k-1}}^*, e_{i_k}$ can be reached with no inversions.

But, since $D(e_{i_k}, e_{i_j}) = \text{sgn}(e_{i_k}, e_{i_j})$ for $j = \{1, \dots, k-1\}$, then the total new inversions is the same as the total new swaps. Thus, by proof by induction, $S = I$ so:

$$D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = (-1)^S = (-1)^I = \text{sgn}(e_{i_1}, e_{i_2}, \dots, e_{i_n})$$

Corollary 4.1.4: Determinant function redefined

For determinant function $D: M_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$, then for any $A \in M_{n \times n}(\mathbb{R})$:

$$D(A) = \sum_{\{i_1, \dots, i_n\} = \{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^I$$

where I is the number of inversions in $(1, i_1), \dots, (n, i_n)$

Proof

$D(A) = \sum_{\{i_1, \dots, i_n\} = \{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^S$ where S is the number of swaps to turn $(e_{i_1}, e_{i_2}, \dots, e_{i_n})$ into (e_1, \dots, e_n) . Since $D(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = (-1)^S$, then by **theorem 4.1.3**, $(-1)^S = \text{sgn}(e_{i_1}, e_{i_2}, \dots, e_{i_n}) = (-1)^I$ where I is the number of inversions in $(1, i_1), \dots, (n, i_n)$. Thus:

$$D(A) = \sum_{\{i_1, \dots, i_n\} = \{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^S = \sum_{\{i_1, \dots, i_n\} = \{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^I$$

Theorem 4.1.5: det() function

Function $\det: M_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$ is a determinant function

$$\det(A) = \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^I$$

where I is the number of inversions in $(1, i_1), \dots, (n, i_n)$

Proof

For $j \in \{1, \dots, n\}$, let $x, y \in \mathbb{R}^n$ and $c_1, c_2 \in \mathbb{R}$:

$$\begin{aligned} T(c_1x + c_2y) &= \det(A_1, \dots, A_{j-1}, c_1x + c_2y, A_{j+1}, \dots, A_n) \\ &= \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{(j-1)i_{j-1}} (c_1x + c_2y)_{ji_j} a_{(j+1)i_{j+1}} \dots a_{ni_n} (-1)^I \\ &= c_1 \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{(j-1)i_{j-1}} x_{ji_j} a_{(j+1)i_{j+1}} \dots a_{ni_n} (-1)^I \\ &\quad + c_2 \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{(j-1)i_{j-1}} y_{ji_j} a_{(j+1)i_{j+1}} \dots a_{ni_n} (-1)^I \\ &= c_1 \det(A_1, \dots, A_{j-1}, x, A_{j+1}, \dots, A_n) + c_2 \det(A_1, \dots, A_{j-1}, y, A_{j+1}, \dots, A_n) \\ &= c_1 T(x) + c_2 T(y) \end{aligned}$$

For $A = [A_1 \dots A_j \dots A_k \dots A_n]$, let $B = [A_1 \dots A_k \dots A_j \dots A_n]$.

By **theorem 4.1.3**, the I_B , number of inversions in B , is equal to S_B , the number of swaps in B to turn $e_{i_1}, \dots, e_{i_k}, \dots, e_{i_j}, \dots, e_{i_n}$ into e_1, \dots, e_n . Thus:

$$\begin{aligned} \det(B) &= \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{ki_k} \dots a_{ji_j} \dots a_{ni_n} (-1)^{I_B} \\ &= \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{ji_j} \dots a_{ki_k} \dots a_{ni_n} (-1)^{S_B} \end{aligned}$$

Note $e_{i_1}, \dots, e_{i_k}, \dots, e_{i_j}, \dots, e_{i_n}$ can first swap into $e_{i_1}, \dots, e_{i_j}, \dots, e_{i_k}, \dots, e_{i_n}$ and then perform swaps to turn into e_1, \dots, e_n . Since the number of swaps to turn $e_{i_1}, \dots, e_{i_j}, \dots, e_{i_k}, \dots, e_{i_n}$ into e_1, \dots, e_n is S_A , the number of swaps in A , then $S_B = S_A + 1$. Thus:

$$\begin{aligned} \det(B) &= \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{ji_j} \dots a_{ki_k} \dots a_{ni_n} (-1)^{S_A+1} \\ &= - \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{ji_j} \dots a_{ki_k} \dots a_{ni_n} (-1)^{S_A} = -\det(A) \end{aligned}$$

For $\det(I_{n \times n})$, the only nonzero $a_{1i_1} \dots a_{ni_n}$ is when each $a_{ki_k} = 1$ else $a_{ki_k} = 0$. Since there are no inversions from e_1, \dots, e_n to e_1, \dots, e_n , then $\det(I_{n \times n}) = a_{1i_1} \dots a_{ni_n} (-1)^I = 1 \dots 1 (-1)^0 = 1$.

Theorem 4.1.6: $\det(A^T) = \det(A)$

For $A \in M_{n \times n}(\mathbb{R})$, then $\det(A^T) = \det(A)$

Proof

Let $\det(A^T) = \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^{I_{A^T}}$.

For $a_{k_1 i_{k_1}}, a_{k_2 i_{k_2}}$ where $i_{k_1} < i_{k_2}$, suppose $k_1 < k_2$. Then, there is no inversion for $a_{k_1 i_{k_1}}$ and $a_{k_2 i_{k_2}}$. Then for $a_{i_{k_1} k_1}, a_{i_{k_2} k_2}$ where $k_1 < k_2$ and $i_{k_1} < i_{k_2}$, then there is no inversion.

For $a_{k_1 i_{k_1}}, a_{k_2 i_{k_2}}$ where $i_{k_1} < i_{k_2}$, suppose $k_1 > k_2$. Then, there is an inversion for $a_{k_1 i_{k_1}}$ and $a_{k_2 i_{k_2}}$. Then for $a_{i_{k_2} k_2}, a_{i_{k_1} k_1}$ where $k_2 < k_1$ and $i_{k_2} > i_{k_1}$, then there is an inversion.

Thus, transpose preserves inversions so $I_{A^T} = I_A$.

Since each $a_{1i_1} \dots a_{ni_n}$ in A^T can be arranged in order by i_1, \dots, i_n into $b_{1i_1} \dots b_{ni_n}$ which when transposed is $b_{i_1 1} \dots b_{i_n n}$ which is in A . Thus:

$$\det(A^T) = \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^{I_{A^T}} = \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} b_{i_1 1} \dots b_{i_n n} (-1)^{I_A} = \det(A)$$

Theorem 4.1.7: Cofactor Expansion: Determining $\det(A)$ by parts

For $A \in M_{n \times n}(\mathbb{R})$, let A_{ij} be A , but the i -th row and j -th column removed.

For any i -th row where $i \in \{1, \dots, n\}$, then by fixing i :

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

Or for any j -th column where $j \in \{1, \dots, n\}$, then by fixing j :

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

Proof

There are n possible a_{ij} choices in the first column and by choosing any such one, then that row is eliminated for choice in the following columns. Thus, there are $n-1$ possible a_{ij} choices in the second column and by choosing any such one, then that row is also eliminated for choice in the following columns. Repeating the pattern, then there are $n(n-1)(n-2)\dots 1 = n!$ total unique $a_{1i_1}, \dots, a_{ni_n}$ combinations. In the cofactor expansion, choose a fixed i . The case for a fixed j is analogous. For a fixed i , the cofactor expansion iterates through each of the n columns in row i so there are n unique a_{ij} . For each a_{ij} , the A_{ij} has the i -th row and j -th column removed so A_{ij} is a $(n-1)$ by $(n-1)$ matrix and thus, there are $(n-1)!$ unique $a_{1i_1}, \dots, a_{ni_n}$ combinations as proved earlier. Since each A_{ij} removes a different j -th column, then each $a_{1i_1}, \dots, a_{ni_n}$ from different columns are unique. Thus, the n unique a_{ij} has $(n-1)!$ unique $a_{1i_1}, \dots, a_{ni_n}$ so there are $n(n-1)! = n!$ unique $a_{1i_1}, \dots, a_{ni_n}$. Thus, the $a_{1i_1}, \dots, a_{ni_n}$ in the cofactor expansion must be equivalent to the $a_{1i_1}, \dots, a_{ni_n}$ in the original determinant.

For the fixed i , let's fix $j \in \{1, \dots, n\}$:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & a_{i-1,2} & a_{i-1,3} & \dots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i,1} & a_{i,2} & a_{i,3} & \dots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \dots & a_{i,n} \\ a_{i+1,1} & a_{i+1,2} & a_{i+1,3} & \dots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \dots & a_{n,n} \end{bmatrix}$$

In the original determinant, each $a_{1i_1}, \dots, a_{ni_n}$ associates $(-1)^{\# \text{inversions } a_{1i_1}, \dots, a_{ni_n}}$. As proven earlier, each $a_{1i_1}, \dots, a_{ni_n}$ is expressed in the cofactor expansion. So for any $a_{1i_1}, \dots, a_{ni_n}$ that contains a_{ij} with the fixed i, j , then from the $a_{ij} \det(A_{ij})$ in the cofactor expansion, the $\det(A_{ij})$ consists of the other a_{ij} in the $a_{1i_1}, \dots, a_{ni_n}$ since none of the other a_{ij} can exist in row i or column j by definition of the determinant and thus, $\det(A_{ij})$ must account for all the inversions exclusively between the other a_{ij} . To account for the inversions between the other a_{ij} and the fixed a_{ij} , refer to the matrix above. The only a_{ij} which contributes an inversion with the fixed a_{ij} must be in the lower left and upper right of the matrix by definition of the determinant. Let $A = \#a_{ij}$ in upper left, $B = \#a_{ij}$ in upper right, $C = \#a_{ij}$ in lower left, and $D = \#a_{ij}$ in lower right. Since each $a_{1i_1}, \dots, a_{ni_n}$ must have a a_{ij} in each row and column, then:

$$\begin{aligned} A+B &= i-1 & A+C &= j-1 & \Rightarrow & B+C &= i+j-2-2A \\ (-1)^{IA} &= (-1)^{B+C} &= (-1)^{i+j-2-2A} &= (-1)^{i+j}(-1)^{-2}(-1)^{-2A} &= & (-1)^{i+j} \end{aligned}$$

Thus:

$$\begin{aligned} \det(A) &= \sum_{\{i_1, \dots, i_n\}=\{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^{IA} \\ &= \sum_{j=1}^n a_{ij} \det(A_{ij}) (-1)^{IA} = \sum_{j=1}^n a_{ij} \det(A_{ij}) (-1)^{i+j} \end{aligned}$$

Example

Find $\det(A)$ where $A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 9 & 1 & 3 & 0 \\ 9 & 2 & 2 & 0 \\ 5 & 0 & 0 & 3 \end{bmatrix}$ using cofactor expansion

Using cofactor expansion along the 4th row:

$$\begin{aligned} \det(A) &= -5\det\begin{bmatrix} 0 & 1 & 2 \\ 1 & 3 & 0 \\ 2 & 2 & 0 \end{bmatrix} + 0\det\begin{bmatrix} 1 & 1 & 2 \\ 9 & 3 & 0 \\ 9 & 2 & 0 \end{bmatrix} - 0\det\begin{bmatrix} 1 & 0 & 2 \\ 9 & 1 & 0 \\ 9 & 2 & 0 \end{bmatrix} + 3\det\begin{bmatrix} 1 & 0 & 1 \\ 9 & 1 & 3 \\ 9 & 2 & 2 \end{bmatrix} \\ &= -5(4-12) + 3(2+18-9-6) = 40 + 15 = 55 \end{aligned}$$

4.2 Properties of the Determinant**Theorem 4.2.1: Relationship between Determinant and Elementary row operations**

Let $A, B \in M_{n \times n}(\mathbb{R})$.

(a) Row or Column Multiplication

Let B be obtained by multiplying the i -th row or j -th column of A by $c \in \mathbb{R}$. Then:

$$\det(B) = c \det(A)$$

Proof

Let B be A with the i -th column multiplied by c . Then:

$$\begin{aligned} \det(B) &= \det(B_1, \dots, B_i, \dots, B_n) = \det(A_1, \dots, cA_i, \dots, A_n) \\ &= c \det(A_1, \dots, A_i, \dots, A_n) = c \det(A) \end{aligned}$$

Let B be A with the j -th column multiplied by c . Then:

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

(b) Row or Column Addition

Let B be obtained by adding to the i -th row by $c \in \mathbb{R}$ times the j -th row of A or by adding to the i -th column by $c \in \mathbb{R}$ times the j -th column of A . Then:

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

Proof

Let B be A with the i -th column added by c times the j -th column. Then:

$$\begin{aligned} \det(B) &= \det(B_1, \dots, B_i, \dots, B_j, \dots, B_n) = \det(A_1, \dots, A_i + cA_j, \dots, A_j, \dots, A_n) \\ &= \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) + c \det(A_1, \dots, A_j, \dots, A_j, \dots, A_n) \\ &= \det(A) + c \cdot 0 = \det(A) \end{aligned}$$

Let B be A with the i -th row added by c times the j -th row. Then:

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

(c) Row or Column Swapping

Let B be obtained swapping the i -th and j -th rows of A or by swapping the i -th and j -th columns of A . Then:

$$\det(B) = -\det(A)$$

Proof

Let B be A with the i -th and j -th columns swapped. Then:

$$\begin{aligned} \det(B) &= \det(B_1, \dots, B_i, \dots, B_j, \dots, B_n) = \det(A_1, \dots, A_j, \dots, A_i, \dots, A_n) \\ &= -\det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) = -\det(A) \end{aligned}$$

Let B be A with the i -th and j -th rows swapped. Then:

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

Example

Find $\det(A)$ where $A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 9 & 1 & 3 & 0 \\ 9 & 2 & 2 & 0 \\ 5 & 0 & 0 & 3 \end{bmatrix}$ using elementary row operations

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 9 & 1 & 3 & 0 \\ 9 & 2 & 2 & 0 \\ 5 & 0 & 0 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -6 & -18 \\ 0 & 2 & -7 & -18 \\ 0 & 0 & -5 & -7 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -6 & -18 \\ 0 & 0 & 5 & 18 \\ 0 & 0 & -5 & -7 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -6 & -18 \\ 0 & 0 & 5 & 18 \\ 0 & 0 & 0 & 11 \end{bmatrix} = B$$

Since only row additions were performed, then $\det(A) = \det(B) = 1(1)(5)(11) = 55$.

Theorem 4.2.2: Invertible $A \Leftrightarrow \det(A) \neq 0$

Let $A \in M_{n \times n}(\mathbb{R})$. Then, A is invertible if and only if $\det(A) \neq 0$.

Proof

By [theorem 2.4.5](#), A is invertible if and only if $\text{rref}(A) = I_{n \times n}$. Then, there is a sequence of elementary row operations that transformation A into $I_{n \times n}$. By [theorem 4.2.1](#), then:

$$\det(A) = (-1)^S \det(I_{n \times n}) = (-1)^S \neq 0$$

where S is the number of row swaps in the sequence. Then if $\det(A) = 0$, then A cannot be invertible. Thus, A is invertible if and only if $\det(A) \neq 0$.

Theorem 4.2.3: $\det(AB) = \det(A)\det(B)$

For $A, B \in M_{n \times n}(\mathbb{R})$, then $\det(AB) = \det(A)\det(B)$

Proof

Suppose at least one of A, B is not invertible. Let A be not invertible.

Then by [theorem 4.2.2](#), $\det(A) = 0$. Since A is not invertible, then AB is not invertible.

$$\det(AB) = 0 = 0\det(B) = \det(A)\det(B)$$

Suppose A, B are invertible. By [theorem 2.4.5](#), $\text{rref}(A) = \text{rref}(B) = I_{n \times n}$. Thus, there is a sequence S_1 of elementary row operations that transforms A into $I_{n \times n}$. Thus, sequence S_1 will transform AB into $I_{n \times n}B = B$. And then there is another sequence S_2 of elementary row operations that transforms B into $I_{n \times n}$.

$$\det(A) = (-1)^{S_A} \quad \det(A) = (-1)^{S_B}$$

where S_A are the number of row swaps in S_1 and S_B are the number of row swaps in S_2 .

$$\det(AB) = (-1)^{S_A} \det(B) = (-1)^{S_A} (-1)^{S_B} = \det(A)\det(B)$$

Corollary 4.2.4: $\det(A^k) = (\det(A))^k$

For $A \in M_{n \times n}(\mathbb{R})$, then $\det(A^k) = (\det(A))^k$

Proof

$$\det(A^k) = \det(A)\det(A^{k-1}) = \det(A)\det(A)\det(A^{k-2}) = \dots = (\det(A))^k$$

Corollary 4.2.5: $\det(A^{-1}) = (\det(A))^{-1}$

For invertible $A \in M_{n \times n}(\mathbb{R})$, then $\det(A^{-1}) = (\det(A))^{-1}$

Proof

Since A is invertible, then $A^{-1}A = I_{n \times n}$. Then:

$$1 = \det(I_{n \times n}) = \det(A^{-1}A) = \det(A^{-1})\det(A) \quad \Rightarrow \quad \det(A^{-1}) = (\det(A))^{-1}.$$

Corollary 4.2.6: Similar matrices have the same Determinant

For similar $A, B \in M_{n \times n}(\mathbb{R})$, then $\det(A) = \det(B)$

Proof

Since A, B are similar, then there is an invertible $X \in M_{n \times n}(\mathbb{R})$ such that $AX = XB$. Then:
 $\det(A)\det(X) = \det(AX) = \det(XB) = \det(X)\det(B)$
 Since X is invertible, then $\det(X) \neq 0$ so $\det(A) = \det(B)$.

Corollary 4.2.7: Determinant of Orthogonal matrices

For orthogonal $A \in M_{n \times n}(\mathbb{R})$, then $\det(A) = \pm 1$.

Proof

Since A is orthogonal, then by **theorem 3.3.2**, $A^T A = I_{n \times n}$. Then:
 $1 = \det(I_{n \times n}) = \det(A^T A) = \det(A^T)\det(A) = (\det(A))^2 \Rightarrow \det(A) = \pm 1$

4.3 Volume

Definition 4.3.1: Volume of Parallelotope

Let $v_1, \dots, v_n \in \mathbb{R}^n$. Then a **parallelotope**, $P(v_1, \dots, v_n)$, is a n -th order parallelogram.

For $i = \{2, \dots, n\}$, let $v_i^\perp = v_i - \text{proj}_{\text{span}(v_1, \dots, v_{i-1})} v_i$. The volume of a parallelotope:

$$\text{Vol}_n(P(v_1, \dots, v_n)) = |v_1| |v_2^\perp| |v_3^\perp| \dots |v_n^\perp|$$

Theorem 4.3.2: $|\det(v_1, \dots, v_n)| = \text{Vol}_n(P(v_1, \dots, v_n))$

For $v_1, \dots, v_n \in \mathbb{R}^n$, let $A = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} \in M_{n \times n}(\mathbb{R})$. Then:

$$\text{Vol}_n(P(v_1, \dots, v_n)) = |\det(A)|$$

Proof

By **theorem 3.2.9**, $v_1, v_2^\perp, \dots, v_n^\perp \in \mathbb{R}^n$ is orthogonal. Let $B = \begin{bmatrix} v_1 & v_2^\perp & \dots & v_n^\perp \end{bmatrix}$. Thus:

$$B^T B = \begin{bmatrix} v_1^T \\ (v_2^\perp)^T \\ \vdots \\ (v_n^\perp)^T \end{bmatrix} \begin{bmatrix} v_1 & v_2^\perp & \dots & v_n^\perp \end{bmatrix} = \begin{bmatrix} v_1^T v_1 & v_1^T v_2^\perp & \dots & v_1^T v_n^\perp \\ (v_2^\perp)^T v_1 & (v_2^\perp)^T v_2^\perp & \dots & (v_2^\perp)^T v_n^\perp \\ \vdots & \vdots & \ddots & \vdots \\ (v_n^\perp)^T v_1 & (v_n^\perp)^T v_2^\perp & \dots & (v_n^\perp)^T v_n^\perp \end{bmatrix} = \begin{bmatrix} |v_1|^2 & 0 & \dots & 0 \\ 0 & |v_2^\perp|^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |v_n^\perp|^2 \end{bmatrix}$$

Since each $v_i^\perp = v_i - (c_{i1}v_1 + \dots + c_{i(i-1)}v_{i-1})$ for $c_{i1}, \dots, c_{i(i-1)} \in \mathbb{R}$, then by **theorem 4.2.1**:

$$\det(B) = \det(v_1, v_2^\perp, v_3^\perp, \dots, v_n^\perp) = \det(v_1, v_2, v_3^\perp, \dots, v_n^\perp) = \dots = \det(v_1, v_2, v_3, \dots, v_n) = \det(A)$$

Thus:

$$\begin{aligned} |\det(A)| &= \sqrt{\det(A)\det(A)} = \sqrt{\det(A^T)\det(A)} = \sqrt{\det(A^T A)} = \sqrt{\det(B^T B)} \\ &= \sqrt{|v_1|^2 |v_2^\perp|^2 \dots |v_n^\perp|^2} = |v_1| |v_2^\perp| \dots |v_n^\perp| = \text{Vol}_n(P(v_1, \dots, v_n)) \end{aligned}$$

Corollary 4.3.3: Expansion Factor: Linear transformation of a Parallelotope

Let linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be $T(x) = Ax$. Then for $v_1, \dots, v_n \in \mathbb{R}^n$,

$T(P(v_1, \dots, v_n))$ is a parallelotope where:

$$\text{Vol}_n(T(P(v_1, \dots, v_n))) = |\det(A)| \text{Vol}_n(P(v_1, \dots, v_n))$$

So, for $T(x) = Ax$, $|\det(A)|$ is called the **expansion factor** of T .

Proof

For any $v \in P(v_1, \dots, v_n)$, then $v = c_1 v_1 + \dots + c_n v_n$ for $c_1, \dots, c_n \in [0, 1]$. Thus:

$$T(P(v_1, \dots, v_n)) = \{ A(c_1 v_1 + \dots + c_n v_n) \} = \{ c_1 A v_1 + \dots + c_n A v_n \}$$

Thus, $T(P(v_1, \dots, v_n))$ is the parallelotope, $P(Av_1, \dots, Av_n)$. By **theorem 4.3.2**:

$$\text{Vol}_n(T(P(v_1, \dots, v_n))) = |\det([Av_1 \dots Av_n])| = |\det(A)\det([v_1 \dots v_n])| = |\det(A)| \text{Vol}_n(P(v_1, \dots, v_n))$$

Theorem 4.3.4: Cauchy-Binet Formula

For $k \leq n$, let $v_1, \dots, v_k \in \mathbb{R}^n$. Let $A = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \in M_{n \times k}(\mathbb{R})$. Then:

$$\text{Vol}_k(P(v_1, \dots, v_k)) = \sqrt{\det(A^T A)}$$

Proof

Since each $v_i^\perp = v_i - (c_{i1}v_1 + \dots + c_{i(i-1)}v_{i-1})$ for $c_{i1}, \dots, c_{i(i-1)} \in \mathbb{R}$, then:

$$B = \begin{bmatrix} v_1 & v_2^\perp & \dots & v_k^\perp \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \dots & v_k \end{bmatrix} \begin{bmatrix} 1 & -c_{21} & -c_{31} & \dots & -c_{k1} \\ 0 & 1 & -c_{32} & \dots & -c_{k2} \\ 0 & 0 & 1 & \dots & -c_{k3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = AX$$

Since $\det(X) = 1$, then by **theorem 4.2.2**, X is invertible so $BX^{-1} = AXX^{-1} = A$. By **theorem 3.2.9**, $v_1, v_2^\perp, \dots, v_k^\perp \in \mathbb{R}^n$ is orthogonal. Thus:

$$\begin{aligned} B^T B &= \begin{bmatrix} v_1^T \\ (v_2^\perp)^T \\ \vdots \\ (v_k^\perp)^T \end{bmatrix} \begin{bmatrix} v_1 & v_2^\perp & \dots & v_k^\perp \end{bmatrix} = \begin{bmatrix} v_1^T v_1 & v_1^T v_2^\perp & \dots & v_1^T v_k^\perp \\ (v_2^\perp)^T v_1 & (v_2^\perp)^T v_2^\perp & \dots & (v_2^\perp)^T v_k^\perp \\ \vdots & \vdots & \ddots & \vdots \\ (v_k^\perp)^T v_1 & (v_k^\perp)^T v_2^\perp & \dots & (v_k^\perp)^T v_k^\perp \end{bmatrix} = \begin{bmatrix} |v_1|^2 & 0 & \dots & 0 \\ 0 & |v_2^\perp|^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |v_k^\perp|^2 \end{bmatrix} \\ \sqrt{\det(A^T A)} &= \sqrt{\det((BX^{-1})^T (BX^{-1}))} = \sqrt{\det((X^{-1})^T) \det(B^T B) \det(X^{-1})} \\ &= \sqrt{\det((X^{-1})) \det(B^T B) \det(X)^{-1}} = \sqrt{\det(B^T B)} \\ &= \sqrt{|v_1|^2 |v_2^\perp|^2 \dots |v_k^\perp|^2} = |v_1| |v_2^\perp| \dots |v_k^\perp| = \text{Vol}_k(P(v_1, \dots, v_k)) \end{aligned}$$

Example

Find the volume of the parallelotope with sides $(1,0,0,0)$, $(1,1,1,1)$, and $(1,2,3,4)$

$$\begin{aligned} \text{Let } A &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 1 & 4 \end{bmatrix}. \text{ Then, } A^T A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 10 \\ 1 & 10 & 30 \end{bmatrix}. \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 10 \\ 1 & 10 & 30 \end{bmatrix} &\Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 9 \\ 0 & 9 & 29 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 9 \\ 0 & 0 & 2 \end{bmatrix} \\ \text{The volume is } &\sqrt{\det(A^T A)} = \sqrt{1(3)(2)} = \sqrt{6}. \end{aligned}$$

5 Eigenvectors

5.1 Diagonalization

Definition 5.1.1: Diagonalizable Matrices

Matrix $A \in M_{n \times n}(\mathbb{R})$ is **diagonal** if:

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \quad \text{where } a_{11}, \dots, a_{nn} \in \mathbb{R}$$

Then A is **diagonalizable** if A is similar to a diagonal matrix A_D

Theorem 5.1.2: Diagonalizability relationship between A and $A_{\mathcal{B}}$

Let $A \in M_{n \times n}(\mathbb{R})$. Then, A is diagonalizable if and only if for basis $\mathcal{B} = v_1, \dots, v_n \in \mathbb{R}^n$, then $A_{\mathcal{B}}$ is diagonalizable.

Proof

By **theorem 2.5.4**, there is an invertible matrix $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} \in M_{n \times n}(\mathbb{R})$ such that:

$$AB = BA_{\mathcal{B}}$$

Thus, A is similar to $A_{\mathcal{B}}$

If A is diagonalizable, then A is similar to diagonal $A_D \in M_{n \times n}(\mathbb{R})$.

Then by **theorem 2.5.7**, $A_{\mathcal{B}}$ is similar to A_D and thus, $A_{\mathcal{B}}$ is diagonalizable.

If $A_{\mathcal{B}}$ is diagonalizable, then $A_{\mathcal{B}}$ is similar to diagonal A_D .

Then by **theorem 2.5.7**, A is similar to A_D and thus, A is diagonalizable.

Theorem 5.1.3: Diagonalizable $A \Leftrightarrow Av = \lambda v$ for some v

$A \in M_{n \times n}(\mathbb{R})$ is diagonalizable if and only if there is a basis $\mathcal{B} = v_1, \dots, v_n \in \mathbb{R}^n$ where $A_{\mathcal{B}}$ is diagonal. Thus, for for some $\lambda_i \in \mathbb{R}$ where $i = \{1, \dots, n\}$:

$$[Av_i]_{\mathcal{B}} = [\lambda_i v_i]_{\mathcal{B}}$$

Proof

Suppose A is diagonalizable. Then there is an invertible $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$, diagonal $A_D \in M_{n \times n}(\mathbb{R})$ such that $AB = BA_D$. By **theorem 2.4.5**, the $\mathcal{B} = v_1, \dots, v_n$ form a basis for \mathbb{R}^n . By **theorem 2.5.4**, then $A_D = A_{\mathcal{B}}$ so $A_{\mathcal{B}}$ is a diagonal.

Suppose there is a basis $\mathcal{B} = v_1, \dots, v_n \in \mathbb{R}^n$ where $A_{\mathcal{B}}$ is diagonal.

By **theorem 2.5.4**, then $AB = BA_{\mathcal{B}}$ where B is invertible so A is diagonalizable.

By **theorem 2.5.3**, then $A_{\mathcal{B}} = \begin{bmatrix} [A(v_1)]_{\mathcal{B}} & \dots & [A(v_n)]_{\mathcal{B}} \end{bmatrix}$. Since $A_{\mathcal{B}}$ is a diagonal, then there are $\lambda_i \in \mathbb{R}$ for $i = \{1, \dots, n\}$ such that:

$$[A(v_i)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{bmatrix} = \lambda_i e_i = \lambda_i [v_i]_{\mathcal{B}} = [\lambda_i v_i]_{\mathcal{B}}$$

Definition 5.1.4: Eigenvalues and Eigenvectors

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation. Then, nonzero $v \in \mathbb{R}^n$ is an **eigenvector** of T if $T(v) = \lambda v$ for some **eigenvalue** $\lambda \in \mathbb{R}$ of T .

If eigenvectors $v_1, \dots, v_n \in \mathbb{R}^n$ with eigenvalues $\lambda_1, \dots, \lambda_n$ form a basis, then v_1, \dots, v_n is a **eigenbasis** for T .

Theorem 5.1.5: Diagonalizable \Leftrightarrow Existence of Eigenbasis

$A \in M_{n \times n}(\mathbb{R})$ is diagonalizable if and only if there is an eigenbasis $\mathcal{B} = v_1, \dots, v_n$ with eigenvalues $\lambda_1, \dots, \lambda_n$ for A . Then for $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$, $A_{\mathcal{B}} = \begin{bmatrix} \lambda_1 e_1 & \dots & \lambda_n e_n \end{bmatrix} \in M_{n \times n}(\mathbb{R})$:

$$AB = BA_{\mathcal{B}}$$

Proof

Suppose A is diagonalizable. Then, by **theorem 5.1.3**, there is a basis $\mathcal{B} = v_1, \dots, v_n$ such that $A_{\mathcal{B}}$ is diagonal where $[Av_i]_{\mathcal{B}} = [\lambda_i v_i]_{\mathcal{B}}$ for some λ_i . Since $Av_i = \lambda_i v_i$, then v_i is an eigenvector with eigenvalue λ_i . Thus, $\mathcal{B} = v_1, \dots, v_n$ is an eigenbasis with eigenvalues $\lambda_1, \dots, \lambda_n$ for A .

Suppose there is an eigenbasis $\mathcal{B} = v_1, \dots, v_n$ with eigenvalues $\lambda_1, \dots, \lambda_n$ for A .

Since $Av_i = \lambda_i v_i$, then $A_{\mathcal{B}} = \begin{bmatrix} [Av_1] & \dots & [Av_n] \end{bmatrix} = \begin{bmatrix} [\lambda_1 v_1] & \dots & [\lambda_n v_n] \end{bmatrix} = \begin{bmatrix} \lambda_1 e_1 & \dots & \lambda_n e_n \end{bmatrix}$, is diagonal. Thus, by **theorem 5.1.3**, then A is diagonalizable.

Since v_1, \dots, v_n is a basis, by **theorem 2.5.4**, then $AB = BA_{\mathcal{B}}$.

Theorem 5.1.6: Eigenvalues of an Orthogonal matrix

For orthogonal $A \in M_{n \times n}(\mathbb{R})$, the only possible eigenvalues are ± 1

Proof

Since A is orthogonal, then $|Ax| = |x|$. If $Ax = \lambda x$, then:

$$|x| = |Ax| = |\lambda x| = |\lambda||x|$$

Thus, $|\lambda| = 1$ so $\lambda = \pm 1$.

Theorem 5.1.7: Eigenvalues and Invertibility

$A \in M_{n \times n}(\mathbb{R})$ is invertible if and only if 0 is not an eigenvalue of A

Proof

Suppose A is invertible. Then by **theorem 2.4.5**, the only solution to $Ax = 0$ is $x = 0$. Thus, if $\lambda = 0$, then $Ax = \lambda x = 0$ has only $x = 0 \neq 0$ so $\lambda = 0$ is not an eigenvalue.

Suppose $\lambda = 0$ is not an eigenvalue of A . Then there are no nonzero $x \in \mathbb{R}^n$ such that $Ax = \lambda x = 0$ so only $x = 0$. Thus, by **theorem 2.4.5**, A is invertible.

5.2 Characteristic Polynomial

Theorem 5.2.1: Determining Eigenvalues

$\lambda \in \mathbb{R}$ is an eigenvalue of $A \in M_{n \times n}(\mathbb{R})$ if and only if $\det(A - \lambda I_{n \times n}) = 0$

Proof

Suppose $\lambda \in \mathbb{R}$ is an eigenvalue of A . Then there is a nonzero $v \in \mathbb{R}^n$ such that $Av = \lambda v$ so $0 = Av - \lambda v = (A - \lambda I_{n \times n})v$. Since nonzero $v \in \ker(A - \lambda I_{n \times n})$, then $A - \lambda I_{n \times n}$ is not invertible. By [theorem 4.2.3](#), then $\det(A - \lambda I_{n \times n}) = 0$.

Suppose $\det(A - \lambda I_{n \times n}) = 0$. By [theorem 4.2.3](#), then $A - \lambda I_{n \times n}$ is not invertible so there is a nonzero $v \in \mathbb{R}^n$ such that $(A - \lambda I_{n \times n})v = 0$. Thus, $Av = \lambda v$ so λ is an eigenvalue.

Theorem 5.2.2: $\det(A - \lambda I_{n \times n}) = \det(A_{\mathcal{B}} - \lambda I_{n \times n})$

Let $\mathcal{B} = v_1, \dots, v_n$ be a basis for \mathbb{R}^n . Then:

$$\det(A) = \det(A_{\mathcal{B}}) \quad \det(A - \lambda I_{n \times n}) = \det(A_{\mathcal{B}} - \lambda I_{n \times n})$$

Proof

By [theorem 2.5.4](#), $AB = BA_{\mathcal{B}}$ where $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$ is invertible. Thus:
 $\det(A) = \det(BA_{\mathcal{B}}B^{-1}) = \det(B)\det(A_{\mathcal{B}})\det(B^{-1}) = \det(B)\det(A_{\mathcal{B}})(\det(B))^{-1} = \det(A_{\mathcal{B}})$

For $i = \{1, \dots, n\}$:

$$[(A - \lambda I_{n \times n})]_{\mathcal{B}}[v_i]_{\mathcal{B}} = [(A - \lambda I_{n \times n})v_i]_{\mathcal{B}} = [Av_i]_{\mathcal{B}} - [\lambda v_i]_{\mathcal{B}} = A_{\mathcal{B}}[v_i]_{\mathcal{B}} - \lambda[v_i]_{\mathcal{B}} = (A_{\mathcal{B}} - \lambda I_{n \times n})[v_i]_{\mathcal{B}}$$

Thus, $\det(A - \lambda I_{n \times n}) = \det([(A - \lambda I_{n \times n})]_{\mathcal{B}}) = \det(A_{\mathcal{B}} - \lambda I_{n \times n})$.

Definition 5.2.3: Characteristic Polynomial

Since the entries, a_{ij}^* , of $A - \lambda I_{n \times n}$ are the same as the entries of A except on its diagonal entries, $a_{ii}^* = a_{ii} - \lambda$, then $\det(A - \lambda I_{n \times n}) = \sum_{\{i_1, \dots, i_n\} = \{1, \dots, n\}} a_{1i_1} \dots a_{ni_n} (-1)^I$ contains an arrangement $a_{1i_1} \dots a_{ni_n} (-1)^I = (a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda)$ which is a polynomial of degree n and the other arrangements contain less than n $(a_{ii} - \lambda)$ so $\det(A - \lambda I_{n \times n})$ is a polynomial of degree n with variable λ .

For $A \in M_{n \times n}(\mathbb{R})$ and $\lambda \in \mathbb{R}$, then $\det(A - \lambda I_{n \times n})$ is the [characteristic polynomial](#) of A .

Definition 5.2.4: Algebraic Multiplicity

Let $A \in M_{n \times n}(\mathbb{R})$ and $\lambda \in \mathbb{R}$.

Then, eigenvalue $\lambda^* \in \mathbb{R}$ of A has an [algebraic multiplicity](#) of k if:

$$\det(A - \lambda I_{n \times n}) = (\lambda - \lambda^*)^k f(\lambda)$$

where $f(\lambda)$ is a polynomial of degree $n-k$ with $f(\lambda^*) \neq 0$. Then, $\text{almu}(\lambda^*) = k \geq 1$.

Theorem 5.2.5: Relationship between Algebraic Multiplicity and Dimension

For $A \in M_{n \times n}(\mathbb{R})$, let $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ be distinct eigenvalues of A . Then $k \leq n$ where:

$$\text{almu}(\lambda_1) + \dots + \text{almu}(\lambda_k) \leq n$$

If $\lambda_1, \dots, \lambda_k \in \mathbb{C}$, then $k = n$ and $\text{almu}(\lambda_1) + \dots + \text{almu}(\lambda_k) = n$.

Proof

Since $\det(A - \lambda I_{n \times n})$ is a polynomial of degree n , then there are at most n real roots so $k \leq n$. If $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ are distinct eigenvalues of A , then:

$$\det(A - \lambda I_{n \times n}) = (\lambda - \lambda_1)^{m_1} \dots (\lambda - \lambda_k)^{m_k} f(\lambda)$$

$$\text{almu}(\lambda_1) + \dots + \text{almu}(\lambda_k) = m_1 + \dots + m_k \leq n$$

For $\lambda_1, \dots, \lambda_k \in \mathbb{C}$, then:

$$\det(A - \lambda I_{n \times n}) = (\lambda - \lambda_1)^{m_1} \dots (\lambda - \lambda_k)^{m_k}$$

$$\text{almu}(\lambda_1) + \dots + \text{almu}(\lambda_k) = m_1 + \dots + m_k = n$$

Corollary 5.2.6: Relationship between Algebraic Multiplicity and Determinant

Let $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ be distinct eigenvalues of $A \in M_{n \times n}(\mathbb{R})$ with $\text{almu}(\lambda_i) = m_i \in \mathbb{R}$. Then:
 $\det(A) = \lambda_1^{m_1} \dots \lambda_k^{m_k}$

Proof

By **theorem 5.2.5**, $\text{almu}(\lambda_1) + \dots + \text{almu}(\lambda_k) = n$. Thus:

$$\det(A - \lambda I_{n \times n}) = (\lambda - \lambda_1)^{m_1} \dots (\lambda - \lambda_k)^{m_k}$$

Then for $\lambda = 0$:

$$\det(A) = \lambda_1^{m_1} \dots \lambda_k^{m_k}$$

5.3 Eigenspaces**Definition 5.3.1: Eigenspaces**

For $A \in M_{n \times n}(\mathbb{R})$, let $\lambda \in \mathbb{R}$ be an eigenvalue. Then there is an eigenvector $v \neq 0 \in \mathbb{R}^n$ such that $Av = \lambda v$ so $0 = Av - \lambda v = (A - \lambda I_{n \times n})v$. Thus, $v \in \ker(A - \lambda I_{n \times n})$.

Then, the **eigenspace** of A for λ , the set of all eigenvectors v with eigenvalue λ :

$$E_\lambda = \ker(A - \lambda I_{n \times n})$$

Definition 5.3.2: Geometric Multiplicity

Let $A \in M_{n \times n}(\mathbb{R})$ and $\lambda \in \mathbb{R}$.

Then, the **geometric multiplicity** of eigenvalue $\lambda^* \in \mathbb{R}$ of A :

$$\text{gemu}(\lambda^*) = \dim(\ker(E_{\lambda^*})) = \dim(\ker(A - \lambda^* I_{n \times n})) = \text{nullity}(A - \lambda^* I_{n \times n})$$

Since λ^* is an eigenvalue, then there is a eigenvector $v \neq 0 \in \ker(A - \lambda^* I_{n \times n})$ so $\dim(\ker(A - \lambda^* I_{n \times n})) = \text{gemu}(\lambda) \geq 1$.

Theorem 5.3.3: Relationship between Algebraic Multiplicity and Geometric Multiplicity

Let $\lambda^* \in \mathbb{R}$ be an eigenvalue of $A \in M_{n \times n}(\mathbb{R})$. Then:

$$\text{gemu}(\lambda^*) \leq \text{almu}(\lambda^*)$$

Thus, if $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ are distinct eigenvalues of A , then:

$$\text{gemu}(\lambda_1) + \dots + \text{gemu}(\lambda_k) \leq n$$

Proof

Let $\text{gemu}(\lambda^*) = m$ so $\dim(\ker(A - \lambda^* I_{n \times n})) = m$. Thus, let v_1, \dots, v_m form a basis for E_{λ^*} . Then choose $u_1, \dots, u_{n-m} \in \mathbb{R}^n$ such that $\mathcal{B} = v_1, \dots, v_m, u_1, \dots, u_{n-m}$ form a basis for \mathbb{R}^n . Since $Av_i = \lambda^* v_i$ for $i = \{1, \dots, m\}$, then for $A_{\mathcal{B}} \in M_{n \times n}(\mathbb{R})$:

$$A_{\mathcal{B}} = \begin{bmatrix} \lambda^* & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1(n-m)} \\ 0 & \lambda^* & \dots & 0 & a_{21} & a_{22} & \dots & a_{2(n-m)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda^* & a_{m1} & a_{m2} & \dots & a_{m(n-m)} \\ 0 & 0 & \dots & 0 & a_{(m+1)1} & a_{(m+1)2} & \dots & a_{(m+1)(n-m)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{n1} & a_{n2} & \dots & a_{n(n-m)} \end{bmatrix} = \begin{bmatrix} \lambda^* I_{m \times m} & B_{m \times (n-m)} \\ 0_{(n-m) \times m} & D_{(n-m) \times (n-m)} \end{bmatrix}$$

where $[Au_j]_{\mathcal{B}} = (a_{1j}, \dots, a_{nj})$ for $j \in \{1, \dots, n-m\}$. Thus:

$$\det(A_{\mathcal{B}} - \lambda I_{n \times n}) = (\lambda^* - \lambda)^m \det(D - \lambda I_{(n-m) \times (n-m)})$$

By **theorem 5.2.2**, then $\det(A - \lambda I_{n \times n}) = \det(A_{\mathcal{B}} - \lambda I_{n \times n})$.

Thus, $\det(A - \lambda I_{n \times n}) = (\lambda^* - \lambda)^m \det(D - \lambda I_{(n-m) \times (n-m)})$ so $\text{gemu}(\lambda^*) = m \leq \text{almu}(\lambda^*)$.

Then for eigenvalues $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ for A , by **theorem 5.2.5**:

$$\text{gemu}(\lambda_1) + \dots + \text{gemu}(\lambda_k) \leq \text{almu}(\lambda_1) + \dots + \text{almu}(\lambda_k) \leq n$$

Theorem 5.3.4: Vandermonde Matrix

For $a_1, \dots, a_n \in \mathbb{R}$, the **Vandermonde Matrix**, $V \in M_{n \times n}(\mathbb{R})$:

$$V = \begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{bmatrix}$$

Then, $\det(V) = \prod_{1 \leq i < j \leq n} (a_j - a_i)$

Proof

Let $V = V_0$. Let V_i be V with the first $i-1$ rows and last $i-1$ columns removed. Let W_i be V_i , but each j -th column from the last to the second is subtracted by a_i times the $(j-1)$ -th column.

$$W_1 = \begin{bmatrix} 1 & a_1 - a_1 & a_1^2 - a_1 a_1 & \dots & a_1^{n-1} - a_1^{n-2} a_1 \\ 1 & a_2 - a_1 & a_2^2 - a_2 a_1 & \dots & a_2^{n-1} - a_2^{n-2} a_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n - a_1 & a_n^2 - a_n a_1 & \dots & a_n^{n-1} - a_n^{n-2} a_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & a_2 - a_1 & a_2(a_2 - a_1) & \dots & a_2^{n-2}(a_2 - a_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n - a_1 & a_n(a_n - a_1) & \dots & a_n^{n-2}(a_n - a_1) \end{bmatrix}$$

By **theorem 4.2.1**, $\det(W_1) = \det(V_1) = \det(V)$. Let A_i be W_i be the first row and column removed. By **theorem 4.1.7**, for $\det(W_1)$ across the first row, since only the first entry in the first row is nonzero, then $\det(W_1) = \det(A_1)$. But for each k -th row in A_i , $(a_{k+i} - a_i)$ is a factor for $k = \{1, \dots, n-i\}$. After factoring out each $(a_{k+1} - a_i)$, then A_i becomes V_{i+2} .

Thus, by **theorem 4.2.1**:

$$\det(V_1) = \det(W_1) = \det(A_1) = (a_2 - a_1)(a_3 - a_1) \dots (a_n - a_1) \det(V_2) = \left(\prod_{i=1}^n (a_{i+1} - a_1) \right) \det(V_2)$$

Repeating the process until V_n :

$$\begin{aligned} \det(V_1) &= \left(\prod_{i=1}^n (a_{i+1} - a_1) \right) \det(V_2) = \left(\prod_{i=1}^n (a_{i+1} - a_1) \right) \left(\prod_{i=2}^n (a_{i+2} - a_2) \right) \det(V_3) \\ &= \dots = \left(\prod_{i=1}^n (a_{i+1} - a_1) \right) \left(\prod_{i=2}^n (a_{i+2} - a_2) \right) \dots \left(\prod_{i=n-1}^n (a_{i+1} - a_{n-1}) \right) \det(V_n) \end{aligned}$$

Since $V_n = [1]$, then $\det(V_n) = 1$. Thus:

$$\det(V) = \det(V_1) = \left(\prod_{i=1}^n (a_{i+1} - a_1) \right) \dots \left(\prod_{i=n-1}^n (a_{i+1} - a_{n-1}) \right) = \prod_{1 \leq i < j \leq n} (a_j - a_i)$$

Theorem 5.3.5: Eigenvectors from different Eigenspaces are Linearly independent

Let $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ be distinct eigenvalues of $A \in M_{n \times n}(\mathbb{R})$ with eigenvector $v_i \in E_{\lambda_i}$.

Then, v_1, \dots, v_k are linearly independent.

Proof

Let $c_1, \dots, c_k \in \mathbb{R}$ be such that $c_1 v_1 + \dots + c_k v_k = 0$. Since $A v_i = \lambda_i v_i$, then for $m = \{1, \dots, k\}$:

$$A^m v_i = A^{m-1}(\lambda_i v_i) = \lambda_i A^{m-1}(v_i) = \lambda_i A^{m-2}(\lambda_i v_i) = \lambda_i^2 A^{m-2}(v_i) = \dots = \lambda_i^m v_i$$

Thus, the system of equations:

$$0 = A^0 0 = A(c_1 v_1 + \dots + c_k v_k) = c_1 \lambda_1 v_1 + \dots + c_k \lambda_k v_k$$

$$0 = A^2 0 = A^2(c_1 v_1 + \dots + c_k v_k) = c_1 \lambda_1^2 v_1 + \dots + c_k \lambda_k^2 v_k$$

\vdots

$$0 = A^k 0 = A^k(c_1 v_1 + \dots + c_k v_k) = c_1 \lambda_1^k v_1 + \dots + c_k \lambda_k^k v_k$$

as a matrix:

$$0_{n \times k} = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \begin{bmatrix} c_1 \lambda_1 & c_1 \lambda_1^2 & \dots & c_1 \lambda_1^k \\ c_2 \lambda_2 & c_2 \lambda_2^2 & \dots & c_2 \lambda_2^k \\ \vdots & \vdots & \ddots & \vdots \\ c_k \lambda_k & c_k \lambda_k^2 & \dots & c_k \lambda_k^k \end{bmatrix} = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} \begin{bmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_k \end{bmatrix} \begin{bmatrix} \lambda_1 & \lambda_1^2 & \dots & \lambda_1^k \\ \lambda_2 & \lambda_2^2 & \dots & \lambda_2^k \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_k & \lambda_k^2 & \dots & \lambda_k^k \end{bmatrix} = \text{VCM}$$

Since each i -th row has a factor of λ_i , by **theorem 4.2.1** and **5.3.4**, then:

$$\det(M) = \lambda_1 \dots \lambda_k \prod_{1 \leq i < j \leq k} (\lambda_j - \lambda_i)$$

Since each $\lambda_i \neq 0$ are distinct, then $\det(M) \neq 0$ so by **theorem 4.2.2**, then M is invertible.

$$0_{n \times k} = 0_{n \times k} M^{-1} = \text{VCM} M^{-1} = \text{VC} = \begin{bmatrix} c_1 v_1 & \dots & c_k v_k \end{bmatrix}$$

Since eigenvectors $v_i \neq 0$, then each $c_i = 0$. Thus, v_1, \dots, v_k are linearly independent.

Corollary 5.3.6: Diagonalizable $M_{n \times n}(\mathbb{R}) \Leftrightarrow \sum \text{gemu}(\lambda_i) = n$

Let $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ be distinct eigenvalues of $A \in M_{n \times n}(\mathbb{R})$.

Then, A is diagonalizable if and only if $\text{gemu}(\lambda_1) + \dots + \text{gemu}(\lambda_k) = n$.

Proof

Suppose A is diagonalizable. By **theorem 5.1.5**, there exists an eigenbasis v_1, \dots, v_n for A. Since each $v_i \in E_{\lambda_j}$ for some $j = \{1, \dots, k\}$, then $n \leq \text{gemu}(\lambda_1) + \dots + \text{gemu}(\lambda_k)$. By **theorem 5.3.3**, then $\text{gemu}(\lambda_1) + \dots + \text{gemu}(\lambda_k) = n$.

Suppose $\text{gemu}(\lambda_1) + \dots + \text{gemu}(\lambda_k) = n$. Let eigenvectors v_{i1}, \dots, v_{im_i} form a basis for E_{λ_i} so $m_1 + \dots + m_k = n$. By **theorem 5.3.5**, eigenvectors from different eigenspaces are linearly independent so $v_{11}, \dots, v_{1m_1}, v_{21}, \dots, v_{2m_2}, \dots, v_{k1}, \dots, v_{km_k}$ are linearly independent. By **theorem 2.3.5**, $v_{11}, \dots, v_{1m_1}, v_{21}, \dots, v_{2m_2}, \dots, v_{k1}, \dots, v_{km_k}$ form a basis of eigenvectors. By **theorem 5.1.5**, then A is diagonalizable.

Theorem 5.3.7: Diagonalizable $M_{n \times n}(\mathbb{R})$ with one Eigenvalue

Let diagonalizable $A \in M_{n \times n}(\mathbb{R})$ have only eigenvalue λ . Then:

$$A = \lambda I_{n \times n}$$

Proof

Since A is diagonalizable, by **corollary 5.3.6**, then $\text{gemu}(\lambda) = n$. Thus, there are n linearly independent eigenvectors v_1, \dots, v_n with eigenvalue λ . By **theorem 2.3.5**, v_1, \dots, v_n span \mathbb{R}^n . Then for any $v \in \mathbb{R}^n$, then $v = c_1 v_1 + \dots + c_n v_n$ for some $c_1, \dots, c_n \in \mathbb{R}$:

$$Av = A(c_1 v_1 + \dots + c_n v_n) = c_1 \lambda v_1 + \dots + c_n \lambda v_n = \lambda(c_1 v_1 + \dots + c_n v_n) = \lambda v$$

Thus, any $v \in \mathbb{R}^n$ is an eigenvector with eigenvalue λ .

$$Av = \lambda v \quad \Rightarrow \quad (A - \lambda I_{n \times n})v = 0$$

Then for $v \neq 0$, then $A - \lambda I_{n \times n} = 0_{n \times n}$ so $A = \lambda I_{n \times n}$.

5.4 Symmetry

Definition 5.4.1: Orthogonally Diagonalizable

Matrix $A \in M_{n \times n}(\mathbb{R})$ is **orthogonally diagonalizable** if there is an orthogonal matrix B such that $B^{-1}AB$ is diagonal

Since B is orthogonal, then by **theorem 3.3.2**, $B^{-1} = B^T$ so $B^{-1}AB = B^T AB$ is diagonal.

By **theorem 5.1.5**, then $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$ where v_1, \dots, v_n is an eigenbasis for A. Since B is orthogonal, then by **theorem 3.3.2**, v_1, \dots, v_n is orthonormal. Thus, if A is orthogonally diagonalizable, there is an orthonormal eigenbasis for A.

Theorem 5.4.2: Eigenvectors from different Eigenspaces are Orthogonal

Let λ_1, λ_2 be distinct eigenvalues of symmetric $A \in M_{n \times n}(\mathbb{R})$.

For $x_1 \in E_{\lambda_1}$ and $x_2 \in E_{\lambda_2}$, then $x_1 \cdot x_2 = 0$

Proof

$$\begin{aligned} \lambda_1(x_1 \cdot x_2) &= \lambda_1 x_1 \cdot x_2 = Ax_1 \cdot x_2 = (Ax_1)^T x_2 = x_1^T A^T x_2 \\ &= x_1 \cdot A^T x_2 = x_1 \cdot Ax_2 = x_1 \cdot (\lambda_2 x_2) = \lambda_2(x_1 \cdot x_2) \end{aligned}$$

Since $(\lambda_1 - \lambda_2)(x_1 \cdot x_2) = 0$, but $\lambda_1 \neq \lambda_2$, then $x_1 \cdot x_2 = 0$.

Theorem 5.4.3: Symmetric $M_{n \times n}(\mathbb{R})$ has n real Eigenvalues

Let $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ be distinct eigenvalues of symmetric $A \in M_{n \times n}(\mathbb{R})$ with $\text{almu}(\lambda_i) = m_i$. Then, $m_1 + \dots + m_k = n$.

Proof

Proof requires basic properties of complex vectors which isn't discussed in these notes.

Let $p_1(\lambda): \mathbb{R} \rightarrow \mathbb{R}$ be the characteristic polynomial of A and let $p_2(\lambda): \mathbb{C} \rightarrow \mathbb{C}$ be the characteristic polynomial of A . Thus, $p_2(\lambda) = p_1(\lambda) = \det(A - \lambda I_{n \times n})$ if $\lambda \in \mathbb{R}$.

Let $p_2(\lambda) = (\lambda - \lambda_1) \dots (\lambda - \lambda_n)$. Let $v = v_1 + v_2 i$ be an eigenvector with eigenvalue $\lambda_i = a + bi$.

$$Av_1 + iAv_2 = A(v_1 + v_2 i) = Av = \lambda_i v = (a + bi)(v_1 + v_2 i) = (av_1 - bv_2) + (av_2 + bv_1)i$$

Since:

$$A\bar{v} = A(v_1 - v_2 i) = Av_1 - iAv_2 \quad \bar{\lambda}_i \bar{v} = (a - bi)(v_1 - v_2 i) = (av_1 - bv_2) - (av_2 + bv_1)i$$

then $A\bar{v} = \bar{\lambda}_i \bar{v}$. Thus, $\bar{\lambda}_i$ is an eigenvalue of A with eigenvector \bar{v} . Then:

$$\bar{v}^T Av = \bar{v}^T (\lambda_i v) = \lambda_i (\bar{v}^T v) = \lambda_i |v|^2 \quad \bar{v}^T Av = \bar{v}^T A \bar{v} = \bar{A} \bar{v}^T v = \bar{\lambda}_i \bar{v}^T v = \bar{\lambda}_i |v|^2$$

Since $(\lambda_i - \bar{\lambda}_i)|v|^2 = 0$, but $|v| > 0$ since $v \neq 0$, then $a + bi = \lambda_i = \bar{\lambda}_i = a - bi$. Thus, $b = 0$ so all eigenvalues are real. Thus, any $p_1(\lambda) = p_2(\lambda) = (\lambda - \lambda_1) \dots (\lambda - \lambda_n)$ so A has n real eigenvalues when including their algebraic multiplicity.

Theorem 5.4.4: Spectral Theorem: Symmetric matrices have an orthonormal eigenbasis

Matrix $A \in M_{n \times n}(\mathbb{R})$ is orthogonally diagonalizable if and only if A is symmetric

Proof

Suppose A is orthogonally diagonalizable. Thus, A has an $\mathcal{B} = v_1, \dots, v_n \in \mathbb{R}^n$ is an orthonormal eigenbasis with eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Since A is diagonalizable, then by [theorem 5.1.5](#), for $B = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$ and $A_{\mathcal{B}} = \begin{bmatrix} \lambda_1 e_1 & \dots & \lambda_n e_n \end{bmatrix}$:

$$AB = BA_{\mathcal{B}}$$

By [theorem 3.3.2](#), B is invertible where $B^{-1} = B^T$. Thus, $A = BA_{\mathcal{B}}B^{-1} = BA_{\mathcal{B}}B^T$. Then:

$$A^T = (BA_{\mathcal{B}}B^T)^T = (B^T)^T A_{\mathcal{B}}^T B^T = BA_{\mathcal{B}}B^T = A$$

Thus, A is symmetric.

Suppose A is symmetric. If $n = 1$, then let $B = [1]$ which is orthogonal since $B^T B = I_{1 \times 1}$ by [theorem 3.3.2](#). Then, $B^{-1}AB = [1][a][1] = [a]$ is diagonal.

Suppose for $k \leq n$, then any symmetric $M_{(k-1) \times (k-1)}(\mathbb{R})$ is orthogonally diagonalizable. Then for symmetric $A \in M_{k \times k}(\mathbb{R})$, by [theorem 5.4.3](#), there is an eigenvalue λ with eigenvector v_1 such that $|v_1| = 1$. From v_1 , let v_1, u_2, \dots, u_k be a basis for \mathbb{R}^k . Apply [theorem 3.2.9](#) to get an orthonormal basis v_1, \dots, v_k . Let $B = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix}$ so B is orthogonal by [3.3.2](#) so $B^T = B^{-1}$.

Note $B^T A B e_1 = B^T A v_1 = B^T (\lambda v_1) = \lambda (B^T v_1) = \lambda e_1$. Also, $(B^T A B) = B^T A^T (B^T)^T =$

$$B^T A B \text{ so } B^T A B \text{ is symmetric. Thus, } B^T A B = \begin{bmatrix} \lambda & 0_{1 \times (k-1)} \\ 0_{(k-1) \times 1} & B^*_{(k-1) \times (k-1)} \end{bmatrix}.$$

Since $B^* \in M_{(k-1) \times (k-1)}(\mathbb{R})$, then B^* is orthogonally diagonalizable so there is an orthogonal

$C \in M_{(k-1) \times (k-1)}(\mathbb{R})$ such that $D = C^{-1}B^*C$ is diagonal. Let $X = \begin{bmatrix} 1 & 0_{1 \times (k-1)} \\ 0_{(k-1) \times 1} & C \end{bmatrix}$. Then:

$$\begin{aligned} X^T B^T A B X &= \begin{bmatrix} 1 & 0_{1 \times (k-1)} \\ 0_{(k-1) \times 1} & C^T \end{bmatrix} \begin{bmatrix} \lambda & 0_{1 \times (k-1)} \\ 0_{(k-1) \times 1} & B^* \end{bmatrix} \begin{bmatrix} 1 & 0_{1 \times (k-1)} \\ 0_{(k-1) \times 1} & C \end{bmatrix} \\ &= \begin{bmatrix} \lambda & 0_{1 \times (k-1)} \\ 0_{(k-1) \times 1} & C^T B^* C \end{bmatrix} = \begin{bmatrix} \lambda & 0_{1 \times (k-1)} \\ 0_{(k-1) \times 1} & D \end{bmatrix} \end{aligned}$$

Since $X^T B^T A B X = (BX)^T A (BX)$ is diagonal where BX is orthogonal since B, X are orthogonal by [corollary 3.3.4](#), then $A \in M_{k \times k}(\mathbb{R})$ is orthogonally diagonalizable. Thus, by proof by induction, $A \in M_{n \times n}(\mathbb{R})$ is orthogonally diagonalizable.

5.5 Quadratic Forms

Definition 5.5.1: Quadratic Form

f: $\mathbb{R}^n \rightarrow \mathbb{R}$ is a **quadratic form** if there are $a_{ij} \in \mathbb{R}$ such that:

$$f(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

Theorem 5.5.2: Matrix Form of the Quadratic Form

Let f: $\mathbb{R}^n \rightarrow \mathbb{R}$ be a quadratic form. Then, there is a unique symmetric $A \in M_{n \times n}(\mathbb{R})$:

$$f(x) = x \cdot Ax$$

If $B \in M_{n \times n}(\mathbb{R})$ satisfies $f(x) = x \cdot Bx$, then $A = \frac{1}{2}(B^T + B)$.

Proof

Let $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$. Since each $x = \sum_{k=1}^n x_k e_k$, then:

$$f(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \sum_{i=1}^n \sum_{j=1}^n (e_i \cdot A e_j) x_i x_j = \sum_{i=1}^n \sum_{j=1}^n (e_i \cdot A(x_j e_j)) x_i$$

$$= \sum_{i=1}^n (e_i \cdot Ax) x_i = \sum_{i=1}^n x_i e_i \cdot Ax = x \cdot Ax$$

Let $B \in M_{n \times n}(\mathbb{R})$ satisfy $f(x) = x \cdot Bx$.

$$f(x) = x \cdot Bx = \frac{1}{2}(Bx \cdot x + x \cdot Bx) = \frac{1}{2}(x \cdot B^T x + x \cdot Bx) = \frac{1}{2}(x \cdot (B^T + B)x) = x \cdot \frac{1}{2}(B^T + B)x$$

Thus, $A = \frac{1}{2}(B^T + B)$. Since $(\frac{1}{2}(B^T + B))^T = \frac{1}{2}((B^T)^T + B^T) = \frac{1}{2}(B + B^T) = \frac{1}{2}(B^T + B)$, then A is symmetric. Let symmetric $C \in M_{n \times n}(\mathbb{R})$ also satisfy $f(x) = x \cdot Cx$. Let symmetric $D = A - C$ where for any $x \in \mathbb{R}^n$:

$$x \cdot Dx = (x \cdot (A - C)x) = (x \cdot Ax) - (x \cdot Cx) = f(x) - f(x) = 0.$$

By **theorem 5.4.4**, then D has an orthonormal eigenbasis v_1, \dots, v_n with eigenvalues $\lambda_1, \dots, \lambda_n$. Then for $i = \{1, \dots, n\}$

$$0 = v_i \cdot Dv_i = v_i \cdot \lambda_i v_i = \lambda_i |v_i|^2$$

Since eigenvector $v_i \neq 0$ so $|v_i| > 0$, then all $\lambda_i = 0$. By **theorem 5.3.7**, $D = 0I_{n \times n} = 0_{n \times n}$ so $A = C$. Thus, A is unique.

Theorem 5.5.3: Eigenvalue Form of the Quadratic Form

For f: $\mathbb{R}^n \rightarrow \mathbb{R}$, let $f(x) = x \cdot Ax$. Then, there is an orthonormal eigenbasis v_1, \dots, v_n of A with eigenvalues $\lambda_1, \dots, \lambda_n$ such that:

$$f(x) = \lambda_1(x \cdot v_1)^2 + \dots + \lambda_n(x \cdot v_n)^2$$

Proof

By **theorem 5.4.4**, A has an orthonormal eigenbasis v_1, \dots, v_n with eigenvalues $\lambda_1, \dots, \lambda_n$. For $x \in \mathbb{R}^n$, by **theorem 3.2.6**, then $x = (x \cdot v_1)v_1 + \dots + (x \cdot v_n)v_n$. Since v_1, \dots, v_n is orthonormal, then $v_i \cdot v_i = 1$ and $v_i \cdot v_j = 0$ for $i \neq j$. Thus:

$$f(x) = x \cdot Ax = ((x \cdot v_1)v_1 + \dots + (x \cdot v_n)v_n) \cdot A((x \cdot v_1)v_1 + \dots + (x \cdot v_n)v_n)$$

$$= ((x \cdot v_1)v_1 + \dots + (x \cdot v_n)v_n) \cdot ((x \cdot v_1)\lambda_1 v_1 + \dots + (x \cdot v_n)\lambda_n v_n)$$

$$= \lambda_1(x \cdot v_1)^2 + \dots + \lambda_n(x \cdot v_n)^2$$

Definition 5.5.4: Definiteness

For f: $\mathbb{R}^n \rightarrow \mathbb{R}$, let $f(x) = x \cdot Ax$. Then:

- (a) f is **positive definiteness** if all eigenvalues of A are positive
- (b) f is **negative definiteness** if all eigenvalues of A are negative
- (c) f is **indefinite** if there is at least one positive and negative eigenvalue

References

- [1] Otto Bretscher, *Linear Algebra with Applications (4th Edition)*, ISBN-13: 978-0321796974