# Linear Algebra

Azure

2022

# Contents

# 1 Vectors & Matrices

## 1.1 Vectors

**Definition 1.1.1: Vector**

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

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

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

$$x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \qquad cx = (cx_1, cx_2, ..., cx_n)$$

Also, define the length of x:

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

A unit vector is a vector with length 1.

**Theorem 1.1.2: Properties of Vectors**

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

(a) Commutativity

$$x + y = y + x$$

Proof

$$x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) = (y_1 + x_1, y_2 + x_2, ..., y_n + x_n) = y + x$$

(b) Addititive Associativity

$$(x + y) + z = x + (y + z)$$

Proof

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

(c) Additive Identity

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

$$0_v + x = x$$

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

Proof

Suppose there are $0_{v_1}, 0_{v_2}$ where $0_{v_1} + x = x$ and $0_{v_2} + x = x$ for all x. Then:

$$0_{v_1} = 0_{v_2} + 0_{v_1} = 0_{v_1} + 0_{v_2} = 0_{v_2}$$

Thus, $0_v$ must be unique.

Since $0 + x = (0 + x_1, ..., 0 + x_n) = (x_1, ..., x_n) = x$, then $0_v = 0$.

(d) Additive Inverse

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

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

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

Proof

Suppose there are $(-x)_a, (-x)_b$ where $x + (-x)_a = 0_v$ and $x + (-x)_b = 0_v$. Then:

$$(-x)_a = (-x)_a + 0_v = (-x)_a + x + (-x)_b = 0_v + (-x)_b = (-x)_b$$

Thus, $-x$ must be unique.

Since $(-1)x + x = (-x_1 + x_1, ..., -x_n + x_n) = 0$, then $-x = (-1)x$.

(e) Distributivity

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

Proof

$$c_1(x+y) = (c_1x_1+c_1y_1, ..., c_1x_n+c_1y_n) = (c_1x_1, ..., c_1x_n) + (c_1y_1, ..., c_1y_n) = c_1x+c_1y$$
$$(c_1+c_2)x = (c_1x_1 + c_2x_1, ..., c_1x_n + c_2x_n)$$
$$= (c_1x_1, ..., c_1x_n) + (c_2x_1, ..., c_2x_n) = c_1x + c_2x$$

(f) Multiplicative Associativity

$c_1(c_2x) = (c_1c_2)x$

Proof

$$c_1(c_2)x = (c_1c_2x_1, ..., c_1c_2x_n) = (c_1c_2)x$$

(g) Multiplicative Identity

1x = x

Proof

$$1x = (1x_1, ..., 1x_n) = (x_1, ..., x_n) = x$$

Theorem 1.1.3: Rescaling to a Unit Vector

Let x $\in \mathbb{R}^n$. Then, $\frac{x}{|x|}$ is a unit vector.

Proof

$$|\tfrac{x}{|x|}| = \sqrt{(\tfrac{x_1}{|x|})^2 + ... + (\tfrac{x_n}{|x|})^2} = \tfrac{1}{|x|}\sqrt{x_1^2 + ... + x_n^2} = \tfrac{1}{|x|}|x| = 1$$

## 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:

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

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

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

(b) Addititive 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.

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, 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.

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

(e) Distributivity

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

Proof

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

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

(f) Multiplicative Associativity

$c_1(c_2A) = (c_1c_2)A$

Proof

$$c_1(c_2)A = \begin{bmatrix} c_1c_2a_{11} & \dots & c_1c_2a_{1n} \\ \vdots & \ddots & \vdots \\ c_1c_2a_{m1} & \dots & c_1c_2a_{mn} \end{bmatrix} = (c_1c_2)A$$

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

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

Definition 1.3.3: System of Equations: Augmented Matrix

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

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

If the i-th column of the 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:

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

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 × 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 rref(A), the solutions are unchanged.

Proof

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

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

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

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

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

$ca_{i1}x_1^/ + ... + ca_{in}x_n^/ = c(a_{i1}x_1^/ + ... + a_{in}x_n^/) \neq cb_i$

Thus, row multiplication does not change the solutions. Note c is nonzero since if c = 0, then any $(x_1, ..., x_n)$ 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.

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

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

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

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

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

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:

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

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

Add i'-th row to j'-th. Multiply the i'-th row by -1 $\Rightarrow$ j''-th = i-th i''-th = j-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 ≠ 0 and Ax = c where c ≠ 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 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:

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

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 $B_n$ = 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,...,m} where $b_{i(n+1)} \neq c_{i(n+1)}$. Let $(x_1^*, ..., x_{n+1}^*)$ be a solution.

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

Since $B_n = C_n$, then $b_{ij} = c_{ij}$ for i = {1,...,m} and j = {1,...,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 rref(A) contains the row [0...0 | 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 rref(A) either:
- contains the row [0...0 | 1]
- doesn't contains the row [0...0 | 1] and have all pivot variables
- doesn't contains the row [0...0 | 1], but have all pivot variables

Since [0...0 | 1] implies $0 = 0x_1 + ... + 0x_n = 1$, then if rref(A) contains the row [0...0 | 1], there cannot be any solution regardless of pivot and free variables since no x = $(x_1, ..., x_n)$ will satisfy such a row. Now, suppose rref(A) doesn't contains the row [0...0 | 1].
Suppose the 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 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 × 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 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 × n system of equations:

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

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

$\ldots$

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

can also be written in matrix form:

Ax = b

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

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

$A(x_1 + ... + x_k) = Ax_1 + ... + Ax_k$

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

$A(cx) = cAx$

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

$$A(c_1x_1 + ... + c_kx_k)$$

$$\Leftrightarrow \begin{matrix} a_{11}(c_1x_{11} + ... + c_kx_{k1}) + a_{12}(c_1x_{12} + ... + c_kx_{k2}) + \ldots + a_{1n}(c_1x_{1n} + ... + c_kx_{kn}) \\ a_{21}(c_1x_{11} + ... + c_kx_{k1}) + a_{22}(c_1x_{12} + ... + c_kx_{k2}) + \ldots + a_{2n}(c_1x_{1n} + ... + c_kx_{kn}) \\ \ldots \\ a_{m1}(c_1x_{11} + ... + c_kx_{k1}) + a_{m2}(c_1x_{12} + ... + c_kx_{k2}) + \ldots + a_{mn}(c_1x_{1n} + ... + c_kx_{kn}) \end{matrix}$$

$$= \begin{matrix} c_1(a_{11}x_{11} + ... + a_{1n}x_{1n}) \\ c_1(a_{21}x_{11} + ... + a_{2n}x_{1n}) \\ \ldots \\ c_1(a_{m1}x_{11} + ... + a_{mn}x_{1n}) \end{matrix} + ... + \begin{matrix} c_k(a_{11}x_{k1} + ... + a_{1n}x_{kn}) \\ c_k(a_{21}x_{k1} + ... + a_{2n}x_{kn}) \\ \ldots \\ c_k(a_{m1}x_{k1} + ... + a_{mn}x_{kn}) \end{matrix} \Leftrightarrow c_1Ax_1 + ... + c_kAx_k$$

A Homogeneous equation is in the form:

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

A Inhomogeneous equation is in the form:

Ax = b    where A $\in M_{m\times n}(\mathbb{R})$, x $\in \mathbb{R}^n$, 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?

$$\begin{bmatrix} 1 & 0 & 2 & 4 & | & -8 \\ 0 & 1 & -3 & -1 & | & 6 \\ 3 & 4 & -6 & 8 & | & 0 \\ 0 & -1 & 3 & 4 & | & -12 \end{bmatrix} \underset{\Rightarrow}{\text{Add -3(1st) to the (3rd)}} \begin{bmatrix} 1 & 0 & 2 & 4 & | & -8 \\ 0 & 1 & -3 & -1 & | & 6 \\ 0 & 4 & -12 & -4 & | & 24 \\ 0 & -1 & 3 & 4 & | & -12 \end{bmatrix}$$

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

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: $\mathbb{R}^n \to \mathbb{R}^m$ 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 $x_1, ..., x_k \in \mathbb{R}^n$ and $c_1, ..., c_k \in \mathbb{R}$:

$$\begin{aligned} \mathrm{T}(c_1x_1 + ... + c_kx_k) &= \mathrm{T}(c_1x_1) + \mathrm{T}(c_2x_2 + ... + c_kx_k) \\ &= c_1\mathrm{T}(x_1) + \mathrm{T}(c_2x_2) + \mathrm{T}(c_3x_3 + ... + c_kx_k) \\ &= c_1\mathrm{T}(x_1) + c_2\mathrm{T}(x_2) + \mathrm{T}(c_3x_3) + \mathrm{T}(c_4x_4 + ... + c_kx_k) \\ &= ... = c_1\mathrm{T}(x_1) + ... + c_k\mathrm{T}(x_k) \end{aligned}$$

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

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

Theorem 1.4.2: T(0) = 0

Let T: $\mathbb{R}^n \to \mathbb{R}^m$ 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: $\mathbb{R}^n \to \mathbb{R}^m$ 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 $x = (x_1, ..., x_n) = x_1e_1 + ... + x_ne_n$, then:

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

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

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

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

Corollary 1.4.5: Linear Transformation: 2D Rotation

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

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

Proof

By theorem 1.4.3, T(x) $= \begin{bmatrix} T(e_1) & T(e_2) \end{bmatrix}$x. Since T rotates any x $\mathbb{R}^2$ counterclockwise by an angle of $\theta$, then:

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

Example

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

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

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

$$\mathrm{T}(\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}) = \begin{bmatrix} \frac{\sqrt{3}}{2}\frac{\sqrt{2}}{2} + -\frac{1}{2}\frac{\sqrt{2}}{2} \\ \frac{1}{2}\frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{2}\frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}-\sqrt{2}}{4} \\ \frac{\sqrt{6}+\sqrt{2}}{4} \end{bmatrix} = \begin{bmatrix} \cos(75°) \\ \sin(75°) \end{bmatrix}$$

## 1.5 Invertibility

Definition 1.5.1: Product of Linear Transformations

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

$$\mathrm{(ST)x} = \mathrm{S(Tx)}$$

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

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

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

Theorem 1.5.2: Product of Linear Transformations are Matrix Transformations

Let T: $\mathbb{R}^n \to \mathbb{R}^m$ with T(x) = Ax and S: $\mathbb{R}^m \to \mathbb{R}^t$ with S(y) = By be linear transformations where A $= \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix}$ for $A_i$ $\mathbb{R}^m$. Then for any x $\in \mathbb{R}^n$:

$$\mathrm{(ST)x} = \mathrm{(BA)x}$$

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

Proof

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

$$= \mathrm{B}(x_1A_1 + ... + x_nA_n) = x_1\mathrm{B}A_1 + ... + x_n\mathrm{B}A_n = \begin{bmatrix} BA_1 & \dots & BA_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Example

Let T: $\mathbb{R}^2 \to \mathbb{R}^2$ 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.

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

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

$(\mathrm{CB})\mathrm{Ax} = (\mathrm{CB})(A_1x_1+...+A_nx_n) = CBA_1x_1+...+CBA_nx_n = \mathrm{CBAx} = \mathrm{C(BA)x}$

(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 $\in \mathbb{R}^n$:

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

(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 $\in \mathbb{R}^n$:

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

(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 $\in \mathbb{R}^n$:

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

Definition 1.5.4: Identity Matrix

The n × 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

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

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

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

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

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

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

**Definition 1.5.6: Invertibility**

Linear Transformation T: $\mathbb{R}^n \to \mathbb{R}^m$ is invertible if there exist a linear transformation S: $\mathbb{R}^m \to \mathbb{R}^n$ such that for all x $\in \mathbb{R}^n$ and y $\in \mathbb{R}^m$:

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

BA = $I_{n\times n}$ 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 ABy = 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 B$y_1$ = x = B$y_2$ 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 A$x_1$ = y = A$x_2$ so BA $\neq I_{n\times n}$ else $x_1$ = BA$x_1$ = BA$x_2$ = $x_2$ 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 rref(A) = $I_{n\times n}$

Proof

Let A be invertible. Suppose 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 = $A^{-1}A$x = $A^{-1}y$ 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, rref(A) = $I_{n\times n}$.

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

$[A \mid I_{n\times n}]$ $\Leftrightarrow$ Ax = $I_{n\times n}$y $\quad$ $[I_{n\times n} \mid B]$ $\Leftrightarrow$ $I_{n\times n}$x = By

(BA)x = B(Ax) = By = x $\quad$ (AB)y = A(By) = Ax = y $\Rightarrow$ $A^{-1}$ = B

### Corollary 1.5.9: Invertible n × n matrices have Rank n

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

Proof

By theorem 1.5.8, A is invertible $\Leftrightarrow$ rref(A) = $I_{n\times n}$ $\Leftrightarrow$ A has n pivots (i.e. 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, 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 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, rref(A) has n pivots so 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 = $AI_{n\times n}$ = A$I_{n\times n}$ = A$BB^{-1}$ = $I_{n\times n}B^{-1}$ = $B^{-1}$

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) rref(A) = $I_{n\times n}$

(c) 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 soluton 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: $\mathbb{R}^4 \to \mathbb{R}^4$ 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 $T^{-1}(1,1,-1,-6)$.

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 2 & 4 & 7 & 11 & 0 & 1 & 0 & 0 \\ 3 & 7 & 14 & 25 & 0 & 0 & 1 & 0 \\ 4 & 11 & 25 & 50 & 0 & 0 & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 5 & 13 & -3 & 0 & 1 & 0 \\ 0 & 3 & 13 & 34 & -4 & 0 & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & -7 & -22 & 7 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 5 & 13 & -3 & 0 & 1 & 0 \\ 0 & 0 & -2 & -5 & 5 & 0 & -3 & 1 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & -1 & -7 & 7 & -2 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 7 & -5 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 & -3 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -6 & 9 & -5 & 1 \\ 0 & 1 & 0 & 0 & 9 & -1 & -5 & 2 \\ 0 & 0 & 1 & 0 & -5 & -5 & 9 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 & -3 & 1 \end{array}\right] = A^{-1} \Rightarrow \text{A is invertible}$$

$T^{-1}(1,1,-1,-6) = A^{-1}(1,1,-1,-6)$ = (2,1,-1,0)

# 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) Addititive 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    (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 = $0_v$    $a0_v = 0_v$    (-1)v = -v

Proof

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

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

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

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

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 $v_1, ..., v_k \in \mathbb{R}^n$ if there are $c_1, ..., c_k \in \mathbb{R}$ such that:

x = $c_1 v_1 + ... + c_k v_k$

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

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

Theorem 2.1.4: Remove Linearly dependent vectors to get Linearly independence

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

span$(v_1, \ldots, v_k)$ = span$(v_1, \ldots, v_k$,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 $v_1, \ldots, v_k$, then there are $c_1, \ldots, c_k \in \mathbb{R}$ such that:

u $= c_1v_1 + \ldots + c_kv_k$

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

$u_1 = a_1v_1 + \ldots + a_kv_k + au = a_1v_1 + \ldots + a_kv_k + a(c_1v_1 + \ldots + c_kv_k) = (a_1 + ac_1)v_1 + \ldots + (a_k + ac_k)v_k$

Thus, $u_1 \in$ span$(v_1, \ldots, v_k)$ so span$(v_1, \ldots, v_k$,u) $\subset$ span$(v_1, \ldots, v_k)$.

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

$u_2 = b_1v_1 + \ldots + b_kv_k = [(b_1 - c_1)v_1 + \ldots + (b_k - c_k)v_k] + [c_1v_1 + \ldots + c_kv_k]$

$= (b_1 - c_1)v_1 + \ldots + (b_k - c_k)v_k + u$

Thus, $u_2 \in$ span$(v_1, \ldots, v_k$,u) so span$(v_1, \ldots, v_k) \subset$ span$(v_1, \ldots, v_k$,u).

Theorem 2.1.5: Condition for Linear Independence

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

$c_1v_1 + \ldots + c_kv_k = 0$

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

Note regardless of $v_1, \ldots, v_k$, any $c_1v_1 + \ldots + c_kv_k = 0$ holds true when $(c_1, \ldots, c_k) = 0 \in \mathbb{R}^k$.

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

Proof

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

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

$v_i = c_1v_1 + \ldots + c_{i-1}v_{i-1} + c_{i+1}v_{i+1} + \ldots + c_kv_k$

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

$0 = c_1v_1 + \ldots + c_{i-1}v_{i-1} + c_iv_i + c_{i+1}v_{i+1} + \ldots + c_kv_k$

The statement holds true if the equation was multiplied by any non-zero number. Thus, any $(c_1, \ldots, c_k)$ where at least one $c_i$ is not 0 is not a solution. Since $(c_1, \ldots, c_k) = 0$ is a solution for $c_1v_1 + \ldots + c_kv_k = 0$, then for linearly independent $v_1, \ldots, v_k$, then $(c_1, \ldots, c_k) = 0$.

Suppose the solution, $(c_1, \ldots, c_k)$, to $c_1v_1 + \ldots + c_kv_k = 0$ is only $(c_1, \ldots, c_k) = 0$.

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

$v_i = a_1v_1 + \ldots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \ldots + a_kv_k$

$0 = a_1v_1 + \ldots + a_{i-1}v_{i-1} + -v_i + a_{i+1}v_{i+1} + \ldots + a_kv_k$

Thus, $(a_1, \ldots, a_{i-1}, -1, a_{i+1}, \ldots, a_k)$ is a solution to $c_1v_1 + \ldots + c_kv_k = 0$ contradicting that $(c_1, \ldots, c_k) = 0$. Thus, there are no linearly dependent vectors.

Theorem 2.1.6: Extending the Span of Linearly independent vectors

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

Proof

Let $c_1, \ldots, c_k, c \in \mathbb{R}$ be such that $c_1v_1 + \ldots + c_kv_k + cv = 0$. Suppose c $\neq 0$. Then:

$c_1v_1 + \ldots + c_kv_k =$ -cv $\quad \Rightarrow \quad$ v $= \frac{-c_1}{c}v_1 + \ldots + \frac{-c_k}{c}v_k$

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

$0 = c_1v_1 + \ldots + c_kv_k + cv = c_1v_1 + \ldots + c_kv_k$

Since $v_1, \ldots, v_k$ are linearly independent, then each $c_1 = \ldots = c_k = 0$. Thus, $(c_1, \ldots, c_k, c) = 0$ so $v_1, \ldots, v_k, v$ 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 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: $\mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation T(x) = Ax.

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

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

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

ker(T) = ker(A) = { x | Ax = 0 }

Theorem 2.2.5: Im(A) is a Subspace that spans the columns of A

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

Then, im(A) is a subspace and im(A) = span(A).

Proof

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

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

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

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

u = $a_1A_1 + ... + a_nA_n$ v = $b_1A_1 + ... + b_nA_n$

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

Theorem 2.2.6: ker(A) is a Subspace

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

Proof

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

Then, A($ax_1 + bx_2$) = aA($x_1$) + bA($x_2$) = a0 + b0 = 0 so $ax_1 + bx_2 \in$ ker(A).

**Theorem 2.2.7: Relationship between the Kernel and Linear Independence**

Let T: $\mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation T(x) = Ax where A = $\begin{bmatrix} A_1 & ... & A_m \end{bmatrix}$.
Then, $A_1, ..., A_m$ are linearly independent if and only if ker(A) = {0}.

Proof

Note Ax = $\begin{bmatrix} A_1 & ... & A_m \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x_1A_1 + ... + x_mA_m = 0$. Let $A_1, ..., A_m$ be linearly independent. Then, the only solution to $x_1A_1 + ... + x_mA_m = 0$ is $(x_1, ..., x_n) = 0$ so ker(A) = {0}.

Suppose ker(A) = {0}. Then, the only solution to $x_1A_1 + ... + x_mA_m$ = Ax = 0 is x = 0. By theorem 2.1.5, $A_1, ..., A_m$ are linearly independent.

## 2.3 Dimension

**Definition 2.3.1: Basis & Dimension**

Let set V $\subset \mathbb{R}^n$ be subspace of $\mathbb{R}^n$. If $v_1, ..., v_k \in \mathbb{R}^n$ are linearly independent and span V (i.e. span$(v_1, ..., v_k)$ = V), then $v_1, ..., v_k$ form a basis for V.
The dimension of V, dim(V), is the number of vectors in a basis of V.
Since $e_1, ..., e_n$, are linearly independent and any x $\in \mathbb{R}^n$, is x = $x_1e_1 + ... + x_ne_n$ so span$(e_1, ..., e_n) = \mathbb{R}^n$, then dim($\mathbb{R}^n$) = n. $e_1, ..., e_n$ are called the standard basis vectors.

**Theorem 2.3.2: # linearly independent vectors in V ≤ # vectors that span V**

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

Proof

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

A = $\begin{bmatrix} v_1 & ... & v_m \end{bmatrix} = \begin{bmatrix} Bc_1 & ... & Bc_m \end{bmatrix}$ = B$\begin{bmatrix} c_1 & ... & c_m \end{bmatrix}$ = BC 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 $v_1, ..., v_m \in$ 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 ≥ m.

**Corollary 2.3.3: The dimension is unique**

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

Proof

Since $v_1, ..., v_m$ and $u_1, ..., u_k$ are bases, then $v_1, ..., v_m$ and $u_1, ..., u_k$ are linearly independent and span V. By theorem 2.3.2, m ≤ k and k ≤ m so m = k.

**Theorem 2.3.4: Linear combinations of a Basis are unique**

Let $v_1, ..., v_k \in \mathbb{R}^n$ form a basis for V $\subset \mathbb{R}^n$.
Then for any v $\in$ V, there are unique $(c_1, ..., c_k)$ such that:

v = $c_1v_1 + ... + c_kv_k$

Proof

Since $v_1, ..., v_k$ form a basis for V, then V = span$(v_1, ..., v_k)$. Then for any v $\in$ V, then v $\in$ span$(v_1, ..., v_k)$. Thus, there are $c_1, ..., c_k \in \mathbb{R}$ such that v = $c_1v_1 + ... + c_kv_k$.
Let $a_1, ..., a_k \in \mathbb{R}$ such that v = $a_1v_1 + ... + a_kv_k$. Then:

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

Since $v_1, ..., v_k$ form a basis for V, then $v_1, ..., v_k$ are linearly independent. Thus, by theorem 2.1.5, $c_i - a_i = 0$ for i = {1,...,k} so $c_i = a_i$ for i = {1,...,k}. Thus, $(c_1, ..., c_k)$ 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 $u_1, ..., u_k \in$ V are linearly independent, then k $\leq$ m

(b) For $u_1, ..., u_k \in$ V, if span($u_1, ..., u_k$) = V, then k $\geq$ m

(c) $u_1, ..., u_m \in$ V are linearly independent if and only if span($u_1, ..., u_m$) = V

Proof

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

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

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 $u_1, ..., u_k$ be linearly independent. Since $v_1, ..., v_m$ span V, then by theorem 2.3.2, k $\leq$ m.

---

Let $u_1, ..., u_k$ span V. Since $v_1, ..., v_m$ are linearly independent, then by theorem 2.3.2, m $\leq$ k

---

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

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

$$c_1 u_1 + ... + c_m u_m = 0$$

If there is one $c_i \neq 0$, then $c_i u_i = c_1 u_1 + ... + c_m u_m = 0$ implies $u_i = 0$ which is a dependent vector. Thus, there are at least two $c_i \neq 0$. Then, $u_i = \frac{-c_1}{c_i} u_1 + ... + \frac{-c_{i-1}}{c_i} u_{i-1} + \frac{-c_{i+1}}{c_i} u_{i+1} + ... + \frac{-c_m}{c_i} u_m$ 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 $u_1, ..., u_m$ are independent.

**Theorem 2.3.6: Determining Basis for Im(A) using rref**

Let A = $\begin{bmatrix} v_1 & ... & v_k \end{bmatrix} \in M_{n\times k}(\mathbb{R})$ where each $v_i \in \mathbb{R}^n$.

Then the $v_i$ that are linearly independent are the columns in rref(A) which contain pivots. Thus, such $v_i$ form a basis for im(A).

Proof

By theorem 2.2.5, im(A) = span(A) = span($v_1, ..., v_k$). Suppose the i-th column of 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, $v_1, ..., v_k$ are linearly dependent. But, if all columns without pivots are removed are removed from rref(A) to make B, then Bx = 0 has only pivots and thus, by theorem 1.3.7, there is a unique solution and since B0 = 0, then ker(B) = {0} and thus, the columns with pivots are linearly independent.
If $v_i$ are linearly dependent, then the sequence of elementary row operations to transform A into rref(A) transform the entries of the i-th column into 0 except possibly the first entry and thus, the i-th column does not have a pivot. By theorem 2.1.4, removing linearly dependent vectors will not change the span(A) = span($v_1, ..., v_k$) = im(A) so the resulting vectors with pivots will be linearly independent and span im(A) and thus, form a basis for im(A).

Corollary 2.3.7: dim(im(A)) = rank(A)

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

dim(im(A)) = rank(A)

Proof

Let A = $\begin{bmatrix} A_1 & ... & A_n \end{bmatrix}$ where $A_i \in \mathbb{R}^m$. By theorem 2.3.6, the $A_i$ that form a basis for im(A) are columns that contain pivots so dim(im(A)), the number of vectors in a basis for im(A) is the same as the number of pivots in rref(A), 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(im(A)) + dim(im(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, rank(A) = dim(im(A)).
If the i-th column doesn't have a pivot, then the solutions to Ax = 0 are linear combinations of a vector $v_i$ 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 span ker(A). Thus, the number of free variables is equal to the number of $v_i$, dim(ker(A)), so dim(im(A)) + dim(im(A)) = n.

## 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) = 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 = $\begin{bmatrix} A_1 & ... & A_n \end{bmatrix}$.

Then, T is invertible if and only if $A_1, ..., A_n$ 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, span($A_1, ..., A_n$) = im(A). Since dim($\mathbb{R}^n$) = n, then $A_1, ..., A_n$ span $\mathbb{R}^n$, then by theorem 2.3.5, $A_1, ..., A_n$ are linearly independent.

### Theorem 2.4.4: Injectivity ⇔ Surjectivity

Let T: $\mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation where T(x) = Ax where A = $\begin{bmatrix} A_1 & \ldots & A_n \end{bmatrix}$.
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, span$(A_1, \ldots, A_n)$ = im(A) = $\mathbb{R}^n$. Since dim($\mathbb{R}^n$) = n, then by theorem 2.3.5, $A_1, \ldots, A_n$ 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 $\in$ 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 $c_1, \ldots, c_{i-1}, c_{i+1}, c_n$ where at least one $c_i \neq 0$ such that $A_i = c_1A_1 + \ldots + c_{i-1}A_{i-1} + c_{i+1}A_{i+1} + \ldots + c_nA_n$. Then:

$$\begin{aligned} \text{Ax} &= x_1A_1 + \ldots + x_nA_n \\ &= (x_1 + x_ic_1)A_1 + \ldots + (x_{i-1} + x_ic_{i-1})A_{i-1} + (x_{i+1} + x_ic_{i+1})A_{i+1} + \ldots + (x_n + x_ic_n)A_n \end{aligned}$$

If $x_i \neq 0$, then x = $(x_1, \ldots, x_n)$ 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 = $x_1A_1 + \ldots + x_nA_n = x_{i_1}A_{i_1} + \ldots + x_{i_k}A_{i_k}$ where $A_{i_1}, \ldots, A_{i_k}$ 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 $A_1, \ldots, A_n$ 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: $\mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation where T(x) = Ax where A = $\begin{bmatrix} A_1 & \ldots & A_n \end{bmatrix}$.
Then the following are equivalent:

(a) A is invertible
(b) rref(A) = $I_{n\times n}$
(c) 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 soluton x = 0 ( i.e ker(A) = {0} )
(f) $A_1, \ldots, A_n$ are linearly independent
(g) im(A) = span$(A_1, \ldots, A_n) = \mathbb{R}^n$
(h) T is injective
(i) T is surjective

Proof

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

Example

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

Find if T is invertible, im(T), dim(im(T)), ker(T), and dim(ker(T)).

$$[A \mid I_{4\times4}] = \begin{bmatrix} 1 & 0 & 2 & 4 & | & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & -1 & | & 0 & 1 & 0 & 0 \\ 3 & 4 & -6 & 8 & | & 0 & 0 & 1 & 0 \\ 0 & -1 & 3 & 1 & | & 0 & 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 2 & 4 & | & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & -1 & | & 0 & 1 & 0 & 0 \\ 0 & 4 & -12 & -4 & | & -3 & 0 & 1 & 0 \\ 0 & -1 & 3 & 1 & | & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 2 & 4 & | & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & -1 & | & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & | & -3 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & | & 0 & 1 & 0 & 1 \end{bmatrix} \Rightarrow \text{rref(A)} \neq I_{4\times4} \Rightarrow \text{A is not invertible}$$

Since the 1st and 2nd column have pivots, then (1,0,3,0) and (0,1,4,-1) form a basis for im(T).
im(T) = $c_1(1,0,3,0) + c_2(0,1,4,-1)$ for $c_1, c_2 \in \mathbb{R}$
where (1,0,3,0), (0,1,4,-1) form a basis for im(T) so dim(im(T)) = 2

To find ker(T), (aka solving Ax = 0 for x), replace the right matrix by a column with all 0 entries:

$$\begin{bmatrix} 1 & 0 & 2 & 4 & | & 0 \\ 0 & 1 & -3 & -1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \Leftrightarrow \begin{matrix} x_1 + 2x_3 + 4x_4 = 0 \\ x_2 - 3x_3 - x_4 = 0 \end{matrix} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2x_3 - 4x_4 \\ 3x_3 + x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ 3 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

ker(T) = = $x_3(-2,3,1,0) + x_4(-4,1,0,1)$ 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

# 3 Orthogonality

# 4 Eigenvectors

# References

[1] Otto Bretscher, *Linear Algebra with Applications (4th Edition)*, ISBN-13: 978-0321796974