Lectures 6 and 7
Linear Transformation and Rank-Nullity Theorem

Dipankar Ghosh

Department of Mathematics
Indian Institute of Technology Hyderabad

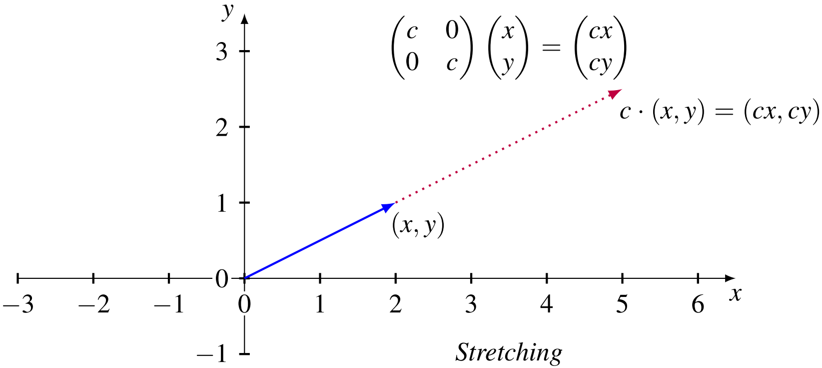
January 21, 2020

A ‘Linear Transformation’ is nothing but a map between vector spaces. Let us start with some well known maps:

$$T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto c \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{where } c \in \mathbb{R}$$

$$\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} cx \\ cy \end{pmatrix}$$

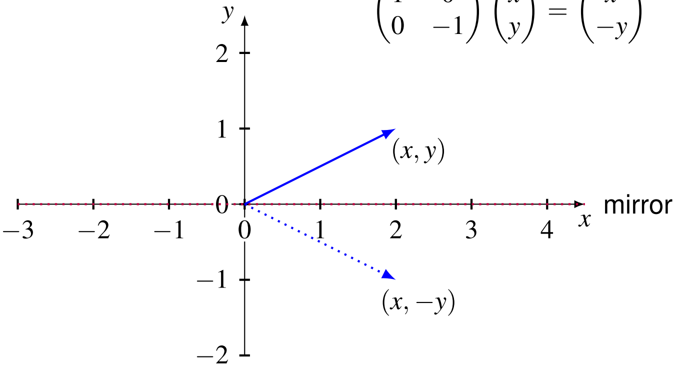


Reflection with x-axis as mirror

$$T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ -y \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$$



Projection on the x-axis

$$S : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

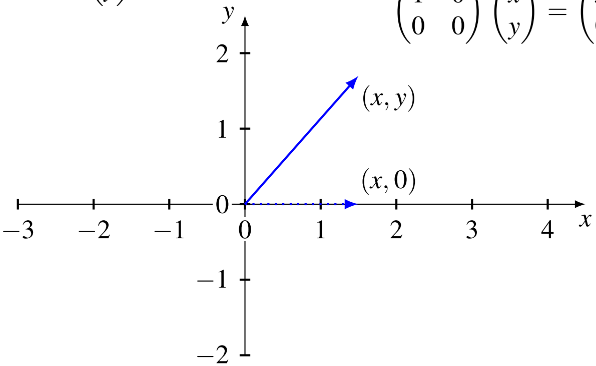
$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto x$$

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x$$

$$T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$$

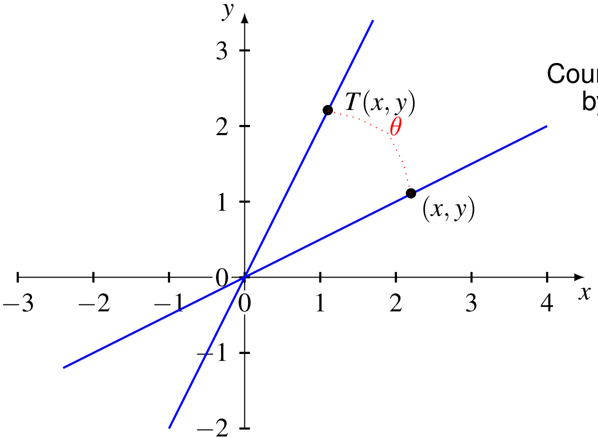
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$$



projection on x-axis

Consider the map $T : \mathbb{R}^2 \to \mathbb{R}^2$ which performs a rotation in the xy-plane counterclockwise by an angle θ about the origin.

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$



Counterclockwise rotation by an angle θ

Linear transformation, or linear map

Definition

A map between vector spaces which satisfies the rule of linearity is called linear transformation (or linear map).

More precisely, let V and W be vector spaces over $\mathbb{R}$. A linear transformation $T : V \to W$ is a function such that

$$T(c_1 v_1 + c_2 v_2) = c_1 T(v_1) + c_2 T(v_2)$$

for all $c_1, c_2 \in \mathbb{R}$ and $v_1, v_2 \in V$. Equivalently, a linear map is a map which respects both vector addition and scalar multiplication.

Example (A matrix can be thought of as a linear map)

Starting with an $m \times n$ matrix A over $\mathbb{R}$, one can construct a linear map $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T_A(X) := AX$ for all $X \in \mathbb{R}^n$.

Proof. $T_A(X+Y) = A(X+Y) = AX + AY = T_A(X) + T_A(Y)$ and $T_A(cX)$ $= A(cX) = c(AX) = cT_A(X)$.

Differentiation and integration transformation

Example (Differentiation transformation)

Let $V = \mathbb{R}[x]$, the set of all polynomials in x over $\mathbb{R}$. Define a map $D : V \to V$ as follows: If $f = a_0 + a_1x + a_2x^2 + \cdots + a_rx^r$, then

$$D(f) := a_1 + 2a_2x + \cdots + ra_rx^{r-1}.$$

Then D is a linear transformation.

Example (Integration transformation)

Let V be the set of all continuous functions from $\mathbb{R}$ into $\mathbb{R}$. Define a map $T : V \to V$ as follows: If $f \in V$, then $T(f)$ is given by

$$T(f)(x) = \int_0^x f(t)dt \quad \text{for all } x \in \mathbb{R}.$$

Then T is a linear transformation.

What is $T(0)$?

- Let $T : V \to W$ be a linear transformation.
- What is $T(0)$?
- Answer: $T(0) = 0$, because $T(0) = T(0+0) = T(0) + T(0)$.

An observation on matrix multiplication

- Multiplying a matrix A with a column vector b yields a linear combination of the columns of A.
- $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} =$

  $x_1 \begin{pmatrix} 1 \\ 5 \\ 9 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 6 \\ 10 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 7 \\ 11 \end{pmatrix} + x_4 \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}.$
- $\begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix}_{m \times n} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1(C_1) + x_2(C_2) + \cdots + x_n(C_n),$

  where $C_1, \ldots, C_n \in \mathbb{R}^m$.

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists an $m \times n$ matrix A such that T can be represented by A, i.e., $T(X) = AX$ for every $X \in \mathbb{R}^n$.

Proof. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$. Consider $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$. Then $X = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{i=1}^{n} x_i e_i$.

Applying T on the above equalities, we have that

$$\begin{aligned} T(X) &= x_1 T(e_1) + x_2 T(e_2) + \cdots + x_n T(e_n) \\ &= \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix} X \quad \text{[by the observation].} \end{aligned}$$

The theorem follows by setting $A := \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix}$.

Remark. A linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ is uniquely determined by its action on $\{e_1, \ldots, e_n\}$, i.e., by $T(e_i)$ for all $1 \leqslant i \leqslant n$.

Correspondence between linear maps and matrices

Corollary

There is a one to one correspondence between the set of all linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ and the collection of all $m \times n$ matrices over $\mathbb{R}$.

Proof. The correspondences are given by

$$\varphi : T \mapsto \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix} \quad \text{and} \quad \psi : A \mapsto T_A.$$

It can be verified that the compositions $\varphi \circ \psi$ and $\psi \circ \varphi$ are the identity maps on the respective domains.

Theorem

Let $T : V \to W$ be a linear transformation. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ and $\mathcal{B}' = \{w_1, \ldots, w_m\}$ be two ordered bases of V and W respectively. Then there exists an $m \times n$ matrix A such that T can be represented by A, i.e., $A[v]_{\mathcal{B}} = [T(v)]_{\mathcal{B}'}$ for every $v \in V$. The ith column of A, which is same as Ae_i, will be obtained by $[T(v_i)]_{\mathcal{B}'}$.

Sketch of the Proof.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, leftrightarrow] & W \arrow[d, leftrightarrow] \\ \mathbb{R}^n & \mathbb{R}^m \end{tikzcd} \qquad \begin{tikzcd} v \arrow[r, "T"] \arrow[d, leftrightarrow] & T(v) \arrow[d, leftrightarrow] \\ {[v]_{\mathcal{B}}} & {[T(v)]_{\mathcal{B}'}} \end{tikzcd} \qquad \begin{tikzcd} V \arrow[r, "T"] \arrow[d, leftrightarrow] & W \arrow[d, leftrightarrow] \\ \mathbb{R}^n \arrow[r, dashed, "???"] & \mathbb{R}^m \end{tikzcd}.$$

By the last theorem, there exists A such that the following diagram is commutative.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, leftrightarrow] & W \arrow[d, leftrightarrow] \\ \mathbb{R}^n \arrow[r, dashed, "A"] & \mathbb{R}^m \end{tikzcd}.$$

Observe that the ith column of A is $[T(v_i)]_{\mathcal{B}'}$.

A linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ is determined by $T(e_i)$

Theorem

Consider the standard basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$.
Then any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is uniquely determined by $T(e_i)$ for all $1 \leqslant i \leqslant n$.

Proof. Every vector $v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ has a unique expression:

$$v = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

Hence, by linearity, $T(v) = x_1 T(e_1) + \cdots + x_n T(e_n)$, which has a unique choice once $T(e_i)$ is given for every i.

A linear map is uniquely determined by its action on a basis

Theorem

Let V be finite dimensional, and $\{v_1, \ldots, v_n\}$ be a basis of V. Then any linear transformation $T : V \to W$ is uniquely determined by $T(v_i)$ for all $1 \leqslant i \leqslant n$.

Proof. Every vector $v \in V$ has a unique expression: $v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$, because if $v = d_1v_1 + d_2v_2 + \cdots + d_nv_n$ is another expression, then

$$(c_1 - d_1)v_1 + \cdots + (c_n - d_n)v_n = 0 \quad \implies \; c_i = d_i \text{ for all } i.$$

Hence, by linearity, $T(v) = c_1T(v_1) + \cdots + c_nT(v_n)$, which has a unique choice, once $T(v_i)$ is given for every i.

A linear map is determined by its action on a basis

Theorem

Let V be finite dimensional, and $\{v_1, \ldots, v_n\}$ be a basis of V. Let $\{w_1, \ldots, w_n\}$ be any collection of n vectors in W. Then there is EXACTLY one linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for all $1 \leqslant i \leqslant n$.

Proof. Once we show the existence, uniqueness follows from the last theorem. We define a map as follows: Every vector $v \in V$ has a UNIQUE expression: $v = c_1 v_1 + \cdots + c_n v_n$ as before.
Define $T(v) := c_1 w_1 + \cdots + c_n w_n$. Then

- $T : V \to W$ is a linear map because:
- If $v = c_1 v_1 + \cdots + c_n v_n$ and $u = d_1 v_1 + \cdots + d_n v_n$, then $v + u = (c_1 + d_1)v_1 + \cdots + (c_n + d_n)v_n$. Hence $T(v + u)$ $= (c_1 + d_1)w_1 + \cdots + (c_n + d_n)w_n = T(v) + T(u)$.
- If $v = c_1 v_1 + \cdots + c_n v_n$, then $cv = (cc_1)v_1 + \cdots + (cc_n)v_n$. Hence $T(cv) = (cc_1)w_1 + \cdots + (cc_n)w_n = cT(v)$.

Null space and nullity of a linear transformation

- Let $T : V \to W$ be a linear transformation. Then
- $\mathrm{Null}(T) := \{v \in V : T(v) = 0\}$ is a subspace of V, because:
- It is non-empty as $0 \in \mathrm{Null}(T)$.
- If $u, v \in \mathrm{Null}(T)$ and $c, d \in \mathbb{R}$,
  then $T(cu + dv) = cT(u) + dT(v) = 0$,
  hence $cu + dv \in \mathrm{Null}(T)$.

Definition (Null space and nullity)

- $\mathrm{Null}(T) := \{v \in V : T(v) = 0\}$ is called the **null space** of T.
- The **nullity** of T is the dimension of the null space of T.

Range (or Image) of a linear transformation, and rank

- Let $T : V \to W$ be a linear transformation. Then
- $\mathrm{Image}(T) := \{w \in W : w = T(v) \text{ for some } v \in V\}$ is a subspace of W, because:
- It is non-empty as $0 \in \mathrm{Image}(T)$.
- If $w_1, w_2 \in \mathrm{Image}(T)$ and $c_1, c_2 \in \mathbb{R}$, then $w_1 = T(v_1)$ and $w_2 = T(v_2)$ for some $v_1, v_2 \in V$, hence $c_1w_1 + c_2w_2 = T(c_1v_1 + c_2v_2) \in \mathrm{Image}(T)$.

Definition (Range space and rank)

- $\mathrm{Image}(T) := \{w \in W : w = T(v) \text{ for some } v \in V\}$ is called the **range space** of T.
- The **rank** of T is the dimension of the range space of T.

Rank-Nullity Theorem

Theorem

Let $T : V \to W$ *be a linear transformation, where* $\dim(V)$ *is finite. Then* $\operatorname{rank}(T) + \operatorname{nullity}(T) = \dim(V)$.

Proof. Start with a basis $\{u_1, \ldots, u_n\}$ of $\mathrm{Null}(T)$. Extend this to a basis $\{u_1, \ldots, u_n, v_1, \ldots, v_r\}$ of V. It is enough to prove that

$$\{T(v_1), \ldots, T(v_r)\} \text{ is a basis of } \mathrm{Image}(T).$$

Spanning: Any vector of $\mathrm{Image}(T)$ looks like $T(v)$ for some $v \in V$.
Write $v = c_1u_1 + \cdots + c_nu_n + d_1v_1 + \cdots + d_rv_r$. Then
$T(v) = c_1T(u_1) + \cdots + c_nT(u_n) + d_1T(v_1) + \cdots + d_rT(v_r)$
$= d_1T(v_1) + \cdots + d_rT(v_r)$.
Lin. Independence: Let $b_1T(v_1) + \cdots + b_rT(v_r) = 0$.
This implies that $b_1v_1 + \cdots + b_rv_r \in \mathrm{Null}(T)$.
So $b_1v_1 + \cdots + b_rv_r = a_1u_1 + \cdots + a_nu_n$ for some $a_i \in \mathbb{R}$.
Thus $b_1v_1 + \cdots + b_rv_r - a_1u_1 - \cdots - a_nu_n = 0$.
Therefore $b_1 = \cdots = b_r = 0$.

Definition

- Let A be an $m \times n$ matrix over $\mathbb{R}$.
- The subspace of $\mathbb{R}^m$ generated by all columns (column vectors) of A is called the **column space** of A.
- The subspace of $\mathbb{R}^n$ generated by all rows (row vectors) of A is called the **row space** of A.

Example

- Let $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$.
- Column space of A is $\textbf{Span}\left\{ \begin{pmatrix} 1 \\ 5 \\ 9 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 10 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \\ 11 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix} \right\}$.
- Column space of A is a subspace of $\mathbb{R}^3$.

Examples: Row and column spaces

Example

- Let $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$.
- Row space of A is $\textbf{Span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 6 \\ 7 \\ 8 \end{pmatrix}, \begin{pmatrix} 9 \\ 10 \\ 11 \\ 12 \end{pmatrix} \right\}$.
- Row space of A is a subspace of $\mathbb{R}^4$.

Example

If $A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, Column Sp. is $\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x_3 = 0 \right\}$.

Row rank and column rank

Definition

- Let A be an $m \times n$ matrix over $\mathbb{R}$.
- The dimension of the column space of A is called the **column rank** of A.
- The dimension of the row space of A is called the **row rank** of A.

Example

- Let $A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & 1 \end{pmatrix}$.
- Column rank of A is 2. Row rank of A is 2.
- Column rank of B is 3. Row rank of B is 3.

As a consequence of Rank-Nullity Theorem, we will prove that for an arbitrary matrix D, row rank(D) = column rank(D).

For every matrix, row rank = column rank (an application of the Rank-Nullity Theorem)

Theorem

For an $m \times n$ matrix A over $\mathbb{R}$, $\text{row rank}(A) = \text{column rank}(A)$.

1. Consider A as a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$.
2. An element of $\mathrm{Image}(A)$ looks like AX for some $X \in \mathbb{R}^n$.
3. But $AX = x_1A_1 + \cdots + x_nA_n$, where A_i is the ith column of A.
4. Therefore $\mathrm{Image}(A) = \mathrm{Span}\{A_1, \ldots, A_n\} = \text{Column Sp.}(A)$.
5. Hence $\mathrm{rank}(A) := \dim(\mathrm{Image}(A)) = \text{column rank}(A)$.
6. **Rank-nullity theorem:** $\mathrm{rank}(A) + \mathrm{nullity}(A) = \dim(\mathbb{R}^n)$.
7. Therefore $\quad \text{column rank}(A) = n - \mathrm{nullity}(A)$.

So it is enough to show that

$$\text{row rank}(A) = n - \mathrm{nullity}(A).$$

Elementary row operations preserve row space, hence rank

Theorem

Let A and B be row equivalent. Then A and B have the same row space. In particular, $\text{row rank}(A) = \text{row rank}(B)$.

Proof. Note that A and B have the same order (say, $m \times n$). Let $R_1, \ldots, R_m \in \mathbb{R}^n$ be the row vectors of A. We observe that the elementary row operations preserve the row space:

1. Effect of the **1st type** elementary row operation, e.g., $\text{Span}\{R_1, R_2, R_3, \ldots, R_m\} = \text{Span}\{R_2, R_1, R_3, \ldots, R_m\}$.
2. Effect of the **2nd type** elementary row operation, e.g., $\text{Span}\{R_1, R_2, R_3, \ldots, R_m\} = \text{Span}\{R_1, c \cdot R_2, R_3, \ldots, R_m\}$, where $c \neq 0$ (important!).
3. Effect of the **3rd type** elementary row operation, e.g., $\text{Span}\{R_1, R_2, R_3, \ldots, R_m\} = \text{Span}\{R_1, R_2 - c \cdot R_1, R_3, \ldots, R_m\}$, where $c \in \mathbb{R}$.

Elementary row operations preserve the nullity of a matrix

1. Let A and B be row equivalent matrices over $\mathbb{R}$.
2. Then $AX = 0$ and $BX = 0$ have the same solution set, i.e., $\mathrm{Null}(A) = \mathrm{Null}(B)$.
3. Therefore $\mathrm{nullity}(A) = \mathrm{nullity}(B)$.

Proof of "row rank$(A) = n -$ nullity(A)"

1. Let A be an $m \times n$ matrix over $\mathbb{R}$.
2. A is row-equivalent to a **row-reduced echelon** matrix B.
3. Since $\text{row rank}(A) = \text{row rank}(B)$ and $\text{nullity}(A) = \text{nullity}(B)$, it is enough to prove that

$$\text{row rank}(B) + \text{nullity}(B) = n.$$

4. We will study some examples to observe this inequality. But I will leave it as an exercise to verify this inequality in the general situation.

For a row-reduced echelon B, row rank$(B) = n -$ nullity(B)

Consider a row-reduced echelon matrix $B = \begin{pmatrix} 1 & 0 & \frac{3}{5} & \frac{7}{5} \\ 0 & 1 & -\frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

- We shall show that $\text{row rank}(B) + \text{nullity}(B) = 4$.
- It can be observed that $\text{row rank}(B)$ is the number of non-zero rows of B, i.e., the number of pivots of B.
  So $\text{row rank}(B) = 2$.
- Consider the system $BX = 0$. The pivot variables are x_1, x_2. The free variables are x_3 and x_4. The system $BX = 0$ is

$$\begin{aligned} x_1 \qquad\quad + \frac{3}{5}x_3 + \frac{7}{5}x_4 &= 0 \\ x_2 - \frac{1}{5}x_3 + \frac{1}{5}x_4 &= 0 \end{aligned}$$

- We claim that $\text{nullity}(B)$ is the number of free variables, because ...

How to solve $BX = 0$ when B is row-reduced echelon?

- Consider $B = \begin{pmatrix} 1 & 0 & \frac{3}{5} & \frac{7}{5} \\ 0 & 1 & -\frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$, a row-reduced echelon matrix.
- The corresponding homogeneous system can be written as

$$\begin{aligned} x_1 \qquad + \frac{3}{5}x_3 + \frac{7}{5}x_4 &= 0 \\ x_2 - \frac{1}{5}x_3 + \frac{1}{5}x_4 &= 0 \end{aligned}$$

- The solutions of the system are given by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -\frac{3}{5}x_3 - \frac{7}{5}x_4 \\ \frac{1}{5}x_3 - \frac{1}{5}x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} -3/5 \\ 1/5 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -7/5 \\ -1/5 \\ 0 \\ 1 \end{pmatrix}$$

- $\operatorname{nullity}(B) = \dim(\operatorname{Null}(B))$ = the number of free variables.

Isomorphism of vector spaces

Definition

A linear map $T : V \to W$ is said to be an **isomorphism** if there is a linear map $S : W \to V$ such that
$S \circ T = 1_V : V \to V$ (identity map) and $T \circ S = 1_W : W \to W$.
If $T : V \to W$ is an isomorphism, we say that V and W are isomorphic, and we write $V \cong W$.

Example

Let A be an $n \times n$ matrix over $\mathbb{R}$. Consider $A : \mathbb{R}^n \to \mathbb{R}^n$ as a linear map. When is it an isomorphism?
Answer: When there is an inverse linear map $B : \mathbb{R}^n \to \mathbb{R}^n$ such that $A \circ B = 1_{\mathbb{R}^n}$ and $B \circ A = 1_{\mathbb{R}^n}$, i.e., when there is an $n \times n$ matrix B over $\mathbb{R}$ such that $AB = I_n$ and $BA = I_n$, i.e., when A is an invertible matrix.

Isomorphism of vector spaces

Theorem

Let $T : V \to W$ be a linear map. The following are equivalent:

1. *T is an isomorphism.*
2. *T is bijective (i.e., as a set map, it is injective and surjective).*

Proof.
(1) $\Rightarrow$ (2): Since T is an isomorphism, there is a linear map $S : W \to V$ such that $S \circ T = 1_V$ and $T \circ S = 1_W$.
Let $T(u) = T(v)$ for some $u, v \in V$.
Apply S on this equality, to get $u = v$. So T is injective.
For surjectivity, note that $T(S(w)) = w$ for every $w \in W$.

(2) $\Rightarrow$ (1): Since T is bijective, there is an inverse SET map $S : W \to V$ such that $S \circ T = 1_V$ and $T \circ S = 1_W$.
All we need to show that $S : W \to V$ is a linear map.
You may try on your own. Otherwise, see the next slide.

(2) $\Rightarrow$ (1):

...

Let $w_1, w_2 \in W$. Want to show $S(w_1 + w_2) = S(w_1) + S(w_2)$.
Set $v_1 := S(w_1)$ and $v_2 := S(w_2)$.
Hence, since T is inverse of S (as a set map), it follows that $T(v_1) = w_1$ and $T(v_2) = w_2$. So $T(v_1 + v_2) = w_1 + w_2$ because T is linear. Therefore $S(w_1 + w_2) = v_1 + v_2 = S(w_1) + S(w_2)$.
Similarly, one can prove that $S(cw) = cS(w)$ for every scalar $c \in \mathbb{R}$ and every vector $w \in W$.

Conditions for a linear transformation to be isomorphism

Theorem

Let $T : V \to V$ be a linear map (or **linear operator***), where $\dim(V) = n < \infty$. Then the following statements are equivalent:*

1. *T is an isomorphism (see the definition in the 1st slide).*
2. *T is bijective (as a set map).*
3. *T is injective.*
4. $\mathrm{Ker}(T) = 0$, *i.e.*, $\{T(v) = 0 \;\Rightarrow\; v = 0\}$, *i.e.*, $\mathrm{Null}(T) = 0$.
5. *T is surjective.*

Proof. We already proved (1) $\Leftrightarrow$ (2). The following implications ar trivial: (2) $\Rightarrow$ (3) $\Rightarrow$ (4).
(4) $\Rightarrow$ (5): Since $\mathrm{Null}(T) = 0$, $\mathrm{nullity}(T) = \dim(\mathrm{Null}(T)) = 0$.
Hence, by Rank-Nullity Theorem, $\mathrm{rank}(T) = \dim(V)$.
So $\mathrm{Image}(T) = V$, i.e., T is surjective.
(5) $\Rightarrow$ (2): Since T is surjective, $\mathrm{rank}(T) = \dim(V)$, hence $\mathrm{nullity}(T) = 0$, i.e., $\mathrm{Ker}(T) = 0$. Then, by linearity, T is injective.

Theorem

Let A be an $n \times n$ matrix over $\mathbb{R}$. The following are equivalent:

1. *A is invertible.*
2. *The homogeneous system $AX = 0$ has only the trivial solution.*
3. *For every $b \in \mathbb{R}^n$, the system $AX = b$ has a solution.*
4. *$\mathbb{R}^n$ is spanned by the column vectors of A.*
5. *The column vectors of A are linearly independent.*
6. *$\mathbb{R}^n$ is spanned by the row vectors of A.*
7. *The row vectors of A are linearly independent.*

Proof. Consider A as a linear map $A : \mathbb{R}^n \to \mathbb{R}^n$. Then (2) is same as $\mathrm{Null}(A) = 0$. Moreover (3) is same as A is surjective. Thus, by the previous theorem, we have (1), (2) and (3) are equivalent.
Since AX is nothing but a linear combination of column vectors of A, it follows that (3) and (4) are equivalent.

Conditions for a square matrix to be invertible contd...

Theorem

Let A be an $n \times n$ matrix over $\mathbb{R}$. The following are equivalent:

1. *A is invertible.*
2. *The homogeneous system $AX = 0$ has only the trivial solution.*
3. *For every $b \in \mathbb{R}^n$, the system $AX = b$ has a solution.*
4. *$\mathbb{R}^n$ is spanned by the column vectors of A.*
5. *The column vectors of A are linearly independent.*
6. *$\mathbb{R}^n$ is spanned by the row vectors of A.*
7. *The row vectors of A are linearly independent.*

Proof. We already proved (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4).
(4) $\Leftrightarrow$ (5) and (6) $\Leftrightarrow$ (7): Since $\dim(\mathbb{R}^n) = n$, any collection of n vectors in $\mathbb{R}^n$ is linearly independent if and only if it spans $\mathbb{R}^n$.
(4) $\Leftrightarrow$ (6): It follows from the above equivalences "(4) $\Leftrightarrow$ (5) and (6) $\Leftrightarrow$ (7)" and the fact that $\text{column rank}(A) = \text{row rank}(A)$.

Thank You!