MA 1140: Lecture 6
Linear Transformation and Rank-Nullity Theorem

Dipankar Ghosh
(IIT Hyderabad)

February 20, 2019

Definition

A transformation (or 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 $v_1, v_2 \in V$.

Example

Let A be an $m \times n$ matrix over $\mathbb{R}$. Then the map $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(X) := AX$ for all $X \in \mathbb{R}^n$ is a linear transformation.

Proof. $T(X + Y) = A(X + Y) = AX + AY = T(X) + T(Y)$ and $T(cX) = A(cX) = c(AX) = cT(X)$.

An observation on matrix multiplication

- $$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} =$$
  $$x_1 \begin{pmatrix} 1 \\ 5 \\ 9 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 6 \\ 10 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 7 \\ 11 \end{pmatrix} + x_4 \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix}.$$
- $$\begin{bmatrix} C1 & C2 & \cdots & Cn \end{bmatrix}_{m \times n} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} =$$
  $$x_1(C1) + x_2(C2) + \cdots + x_n(Cn), \quad \text{where } C1, \ldots, Cn \in \mathbb{R}^m.$$

Theorem

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be represented by an $m \times n$ matrix.

Proof. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Want to construct an $m \times n$ matrix A such that $T(X) = AX$ for all $X \in \mathbb{R}^n$. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$. Set $A := \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix}$. Clearly A is an $m \times n$ matrix. We show that $T(X) = AX$ for every $X \in \mathbb{R}^n$.

- Consider $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$. Then
- $AX = \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix} X$
  $= x_1 T(e_1) + x_2 T(e_2) + \cdots + x_n T(e_n)$ (by the observation)
  $= T(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n) = T(X)$. $\square$

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. Use the last theorem and the example.

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 $T(0) = T(0+0) = T(0) + T(0)$.

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 $1 \leq i \leq n$.

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.

Another approach. Since $T \sim A$, $T(e_i) \sim$ the ith column of A.

Theorem

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

Proof. Every vector $v \in V$ has a unique expression:
$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$, because if
$v = d_1 v_1 + d_2 v_2 + \cdots + d_n v_n$ is another expression, then

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

Hence, by linearity, $T(v) = c_1 T(v_1) + \cdots + c_n T(v_n)$, which has a unique choice, once $T(v_i)$ is given for every i.

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 $1 \leq i \leq n$.

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: $v = c_1v_1 + \cdots + c_nv_n$ as before. Define $T(v) := c_1w_1 + \cdots + c_nw_n$. Then

- $T : V \to W$ is a linear map because:
- If $v = c_1v_1 + \cdots + c_nv_n$ and $u = d_1v_1 + \cdots + d_nv_n$, then $v + u = (c_1 + d_1)v_1 + \cdots + (c_n + d_n)v_n$. Hence $T(v + u) = T(v) + T(u)$.
- If $v = c_1v_1 + \cdots + c_nv_n$, then $cv = (cc_1)v_1 + \cdots + (cc_n)v_n$. Hence $T(cv) = cT(v)$.

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 $c_1 w_1 + c_2 w_2 = T(c_1 v_1 + c_2 v_2) \in \mathrm{Image}(T)$.

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.

Theorem

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

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 $T(v) = c_1T(u_1) + \cdots + c_nT(u_n) + d_1T(v_1) + \cdots + d_rT(v_r)$ $= d_1T(v_1) + \cdots + d_rT(v_r)$.

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

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

Lin. Independence: Let $b_1 T(v_1) + \cdots + b_r T(v_r) = 0$.
This implies that $b_1 v_1 + \cdots + b_r v_r \in \mathrm{Null}(T)$.
So $b_1 v_1 + \cdots + b_r v_r = a_1 u_1 + \cdots + a_n u_n$ for some $a_i \in \mathbb{R}$.
Thus $b_1 v_1 + \cdots + b_r v_r - a_1 u_1 - \cdots - a_n u_n = 0$.
Therefore $b_1 = \cdots = b_r = 0$.

Row and column spaces

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 **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 **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 **column rank** of A.
- The dimension of the row space of A is called **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).

Thank You!