Linear Transformations

Jagannath
IIITDM Kancheepuram, Chennai

Linear Transformations

Definition:

Let V and W be vector spaces over the field F. A linear transformation from V into W is a function $T : V \longrightarrow W$ such that

$$T(c\alpha + \beta) = cT(\alpha) + T(\beta) \text{ for all } \alpha, \beta \in V, \ c \in F.$$

Examples of Linear Transformations

(1) Let V, W be vector spaces over the same field F. We define a function $0 : V \longrightarrow W$ as $0(v) = \mathbf{0}$ for all $v \in V$, where $\mathbf{0}$ is the zero vector of W. Then

$$0(c\alpha + \beta) = \mathbf{0} = c \cdot \mathbf{0} + \mathbf{0} = c0(\alpha) + 0(\beta) \text{ for all } \alpha, \beta \in V, c \in F.$$

Thus, 0 is a linear transformation, called the **zero linear transformation**.

(2) Let V be a vector space over the field F. We define a function $I : V \longrightarrow V$ as $I(v) = v$ for all $v \in V$. Then

$$I(c\alpha + \beta) = c\alpha + \beta = cI(\alpha) + I(\beta) \quad \text{for all } \alpha, \beta \in V, c \in F.$$

Thus, I is a linear transformation, called the **identity linear transformation**.

Examples

(3) Let $V = \left\{ f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n : n \in \mathbb{N}, c_i \in F \right\}$. That is, V is the vector space of all polynomials over the field F. We define a function $D : V \longrightarrow V$ as $(Df)(x) := c_1 + 2c_2 x + \cdots + nc_n x^{n-1}$. Then D is a linear transformation, called the **differentiation transformation**.

(4) Let $F = \mathbb{R}$. Let V be the vector space of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$. We define a function $T : V \longrightarrow \mathbb{R}$ as $(Tf)(x) := \int_0^x f(t)dt$. Then T is a linear transformation, called the **integral transformation**.

(5) Let $A \in F^{m \times n}$ be a fixed $m \times n$ matrix. Define a function $T : F^{n \times 1} \longrightarrow F^{m \times 1}$ as $T(X) = AX$. Then

$$T(cX + Y) = A(cX + Y) = cAX + AY = cT(X) + T(Y).$$

Hence, T is a linear transformation.

Properties of linear transformations

Property 1: $T(\mathbf{0}) = \mathbf{0}$.

Proof. As

$$\begin{aligned} T(\mathbf{0}) &= T(1 \cdot \mathbf{0} + \mathbf{0}) \\ &= 1 \cdot T(\mathbf{0}) + T(\mathbf{0}) \quad (\because T \text{ is a L.T. }) \\ &= T(\mathbf{0}) + T(\mathbf{0}). \end{aligned}$$

This implies

$$T(\mathbf{0}) = \mathbf{0}.$$

Properties of linear transformations

Property 2: The following two conditions are equivalent:

1. $T(c\alpha + \beta) = cT(\alpha) + T(\beta)$.
2. $T(\alpha + \beta) = T(\alpha) + T(\beta)$ and $T(c\alpha) = cT(\alpha)$.

Proof. $(1 \Rightarrow 2)$.

$$T(\alpha + \beta) = T(1 \cdot \alpha + \beta) = 1 \cdot T(\alpha) + T(\beta) = T(\alpha) + T(\beta)$$

and

$$T(c\alpha) = T(c\alpha + \mathbf{0}) = cT(\alpha) + T(\mathbf{0}) = cT(\alpha) + \mathbf{0} = cT(\alpha).$$

$(2 \Rightarrow 1)$.

$$T(c\alpha + \beta) = T(c\alpha) + T(\beta) = cT(\alpha) + T(\beta).$$

Properties of linear transformations

Property 3:

$$T(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n) = c_1 T(\alpha_1) + c_2 T(\alpha_2) + \cdots + c_n T(\alpha_n).$$

Proof.

$$\begin{aligned}
&T(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n) \\
&= c_1 T(\alpha_1) + T(c_2\alpha_2 + \cdots + c_n\alpha_n) \;\; (\because \text{ is a L.T.}) \\
&= c_1 T(\alpha_1) + c_2 T(\alpha_2) + T(c_3\alpha_3 + \cdots + c_n\alpha_n) \;\; (\because T \text{ is a L.T.}) \\
&\;\vdots \\
&= c_1 T(\alpha_1) + c_2 T(\alpha_2) + \cdots + c_n T(\alpha_n).
\end{aligned}$$

Problem 1:

Verify which of the following functions $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ are linear transformations?

(1) $T(x_1, x_2) = (1 + x_1, x_2)$.

Ans: It is not a linear transformation as
$T(0,0) = (1,0) \Longrightarrow T(\mathbf{0}) \neq \mathbf{0}$.

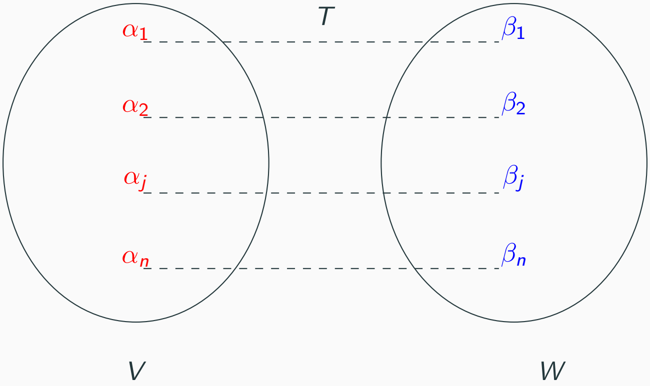
(2) $T(x_1, x_2) = (x_2, x_1)$.

Ans: It is a linear transformation as
$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Longrightarrow T(X) = AX.$$

(3) $T(x_1, x_2) = (x_1^2, x_2)$.

Ans: It is not a linear transformation (Verify!).

Linear transformations are special !!



Ordered basis, $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$

β_j *'s* need not be distinct

T is a unique linear transformation with $T(\alpha_j) = \beta_j$

Theorem 1.

Let V be a finite-dimensional vector space over the field F and let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an ordered basis for V. Let W be a vector space over the same field F and let $\beta_1, \beta_2, \ldots, \beta_n$ be any vectors in W. Then there is precisely one linear transformation $T : V \longrightarrow W$ such that $T(\alpha_j) = \beta_j$ for $j = 1, 2, \ldots, n$.

Proof: Since $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is an ordered basis for V, for a given vector $\alpha \in V$, there is a unique $n-$tuple $(x_1, x_2, \ldots, x_n)$ such that

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n.$$

Proof contd.

We define a function $T : V \longrightarrow W$ as

$$T(\alpha) = T(x_1\alpha_1 + \cdots + x_n\alpha_n) := x_1\beta_1 + \cdots + x_n\beta_n.$$

Claim 1: $T(\alpha_j) = \beta_j$.

$$\begin{aligned} T(\alpha_j) &= T(0\alpha_1 + \cdots + 0\alpha_{j-1} + 1\alpha_j + 0\alpha_{j+1} + \cdots + 0\alpha_n) \\ &= 0\beta_1 + \cdots + 0\beta_{j-1} + 1\beta_j + 0\beta_{j+1} + \cdots + 0\beta_n \\ &= \beta_j. \end{aligned}$$

Claim 2: T **is a linear transformation.**

We show that $T(c\alpha + \beta) = cT(\alpha) + T(\beta)$ for all $\alpha, \beta \in V$, $c \in F$.
Let $\beta = y_1\alpha_1 + y_2\alpha_2 + \cdots + y_n\alpha_n$. Then

$$c\alpha + \beta = (cx_1 + y_1)\alpha_1 + (cx_2 + y_2)\alpha_2 + \cdots + (cx_n + y_n)\alpha_n.$$

Proof contd.

Now, by the definition of T we have

$$\begin{aligned} T(c\alpha+\beta) &= (cx_1+y_1)\beta_1+(cx_2+y_2)\beta_2+\cdots+(cx_n+y_n)\beta_n \\ &= c(x_1\beta_1+x_2\beta_2+\cdots+x_n\beta_n)+(y_1\beta_1+y_2\beta_2+\cdots+y_n\beta_n) \\ &= cT(x_1\alpha_1+\cdots+x_n\alpha_n)+T(y_1\alpha_1+\cdots+y_n\alpha_n) \\ &= cT(\alpha)+T(\beta). \end{aligned}$$

Claim 3: T is unique.

It is enough to prove that if $U : V \longrightarrow W$ is a linear transformation with $U(\alpha_j) = \beta_j$ for $j = 1, 2, \ldots, n$, then $T(\alpha) = U(\alpha)$ for all $\alpha \in V$.

Consider

$$\begin{aligned} U(\alpha) &= U(x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n) \\ &= x_1U(\alpha_1)+x_2U(\alpha_2)+\cdots+x_nU(\alpha_n) \quad (\because \text{ U is a L.T.}) \\ &= x_1\beta_1+x_2\beta_2+\cdots+x_n\beta_n \quad (\because U(\alpha_j)=\beta_j) \\ &= T(\alpha). \end{aligned}$$

This completes the proof of the theorem.

Problem 2.

Let $B = \{\alpha_1 = (1,2), \alpha_2 = (3,4)\}$ be an ordered basis for $\mathbb{R}^2$. Let $\beta_1 = (3,2,1)$, $\beta_2 = (6,5,4) \in \mathbb{R}^3$. Find the unique linear transformation $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ such that $T(\alpha_j) = \beta_j$ for $j = 1, 2$.

Solution: $T(\alpha_1) = T(1,2) = (3,2,1) = \beta_1$ and $T(\alpha_2) = T(3,4) = (6,5,4) = \beta_2$.

Let $\alpha = (x,y) \in \mathbb{R}^2$. As $\{\alpha_1, \alpha_2\}$ is a basis, we have $\alpha = a\alpha_1 + b\alpha_2$ for some $a, b \in \mathbb{R}$. That is $(x,y) = a(1,2) + b(3,4) = (a+3b, 2a+4b)$. So,

$$a + 3b = x \text{ and } 2a + 4b = y.$$

On solving these two equations for a, b we obtain

$$a = \left(-2x + \frac{3}{2}y\right) \text{ and } b = \left(x - \frac{1}{2}y\right).$$

Problem 2 contd.

So,

$$(x, y) = \left(-2x + \frac{3}{2}y\right)\alpha_1 + \left(x - \frac{1}{2}y\right)\alpha_2.$$

Thus,

$$\begin{aligned} T(x, y) &= \left(-2x + \frac{3}{2}y\right)\beta_1 + \left(x - \frac{1}{2}y\right)\beta_2 \quad \text{(from the definition of } T\text{)} \\ &= \left(-2x + \frac{3}{2}y\right)(3, 2, 1) + \left(x - \frac{1}{2}y\right)(6, 5, 4) \\ &= \left(\frac{3}{2}y, x + \frac{1}{2}y, 2x - \frac{1}{2}y\right). \end{aligned}$$

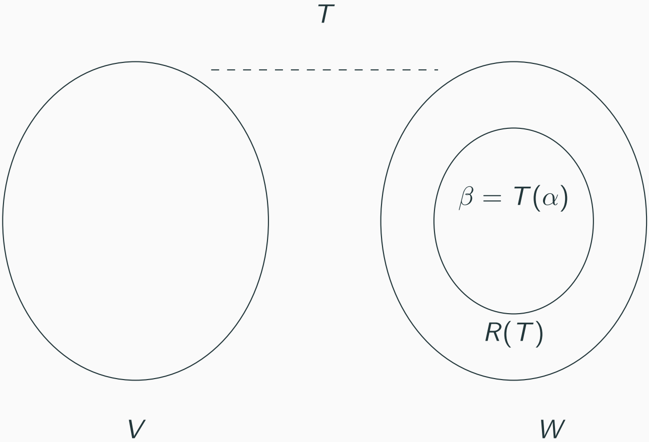
The transformation T is unique, thanks to Theorem 1.

Range of a linear transformation

Let V, W be vector spaces over a field F. Let $T : V \longrightarrow W$ be a linear transformation. The Range of T is defined by

$$R(T) := \{\beta \in W : T(\alpha) = \beta \text{ for some } \alpha \in V\}.$$

Range of T



Proposition-1: $R(T)$ is a subspace of W.

Proof: Let $T : V \longrightarrow W$ be a linear transformation. As $T(\mathbf{0}) = \mathbf{0}$, we have $\mathbf{0} \in R(T)$. Thus, $R(T) \neq \phi$.

Now, let $\beta_1, \beta_2 \in R(T)$ and $c \in F$. By the definition of $R(T)$ there exist $\alpha_1, \alpha_2 \in V$ such that $T(\alpha_1) = \beta_1$ and $T(\alpha_2) = \beta_2$.

As

$$c\beta_1 + \beta_2 = cT(\alpha_1) + T(\alpha_2) = T(c\alpha_1 + \alpha_2),$$

from the definition of $R(T)$ it follows that $c\beta_1 + \beta_2 \in R(T)$. This proves $R(T)$ is a subspace of W.

Definition: The rank of a linear transformation T is defined as the dimension of the range of T.
That is

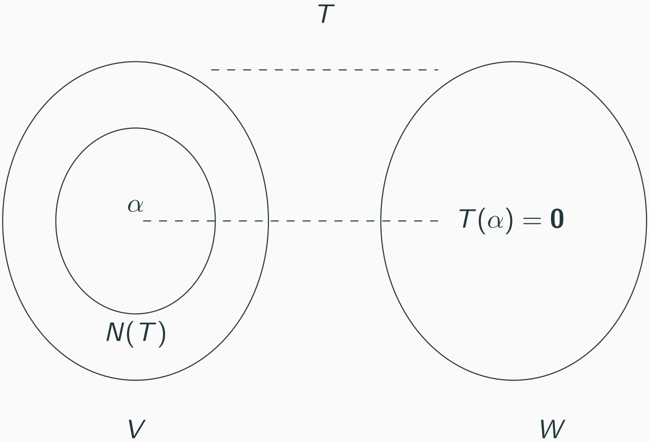
$$\textbf{Rank of } T := \textbf{dim } R(T).$$

The null space of a linear transformation

Let V, W be vector spaces over a field F. Let $T : V \longrightarrow W$ be a linear transformation. The null space of T is defined as

$$N(T) := \{\alpha \in V : T(\alpha) = \mathbf{0}\}.$$

Null space of T



Proposition-2: $N(T)$ is a subspace of V.

Proof: Let $T : V \longrightarrow W$ is a linear transformation. As $T(\mathbf{0}) = \mathbf{0}$, we have $\mathbf{0} \in N(T)$. This implies $N(T) \neq \phi$.

Let $\alpha_1, \alpha_2 \in N(T)$ and $c \in F$. Then $T(\alpha_1) = T(\alpha_2) = \mathbf{0}$. Since T is a linear transformation

$$T(c\alpha_1 + \alpha_2) = cT(\alpha_1) + T(\alpha_2) = c\mathbf{0} + \mathbf{0} = \mathbf{0}.$$

This implies $c\alpha_1 + \alpha_2 \in N(T)$. Hence $N(T)$ is a subspace of V.

Definition: The nullity of a linear transformation T is defined as the dimension of the null space of T. That is

$$\textbf{Nullity of } T = \textbf{dim } N(T).$$

Example 1. Find the rank and nullity of the zero linear transformation $O : V \longrightarrow W$ defined by $O(\alpha) = \mathbf{0}$ for all $\alpha \in V$.

Solution:

$$R(O) = \{\beta \in W : \beta = O(\alpha) \text{ for some } \alpha \in V\} = \{\mathbf{0}\}.$$

$$N(O) = \{\alpha \in V : \ O(\alpha) = \mathbf{0}\} = V.$$

Hence,

$$\text{Rank } (O) = \dim R(O) = 0$$

and

$$\text{Nullity } (O) = \dim N(O) = \dim V.$$

Example 2. Find the rank and nullity of the identity linear transformation $I : V \longrightarrow V$ defined by $I(\alpha) = \alpha$ for all $\alpha \in V$.

Solution:

$$\begin{aligned} R(I) &= \{\beta \in V : \beta = I(\alpha) \text{ for some } \alpha \in V\} \\ &= \{\beta \in V : \beta = I(\alpha) = \alpha \text{ for some } \alpha \in V\} \\ &= \{\alpha \in V\} \\ &= V. \end{aligned}$$

$$N(O) = \{\alpha \in V : \; I(\alpha) = \mathbf{0}\} = \{\mathbf{0}\}.$$

Hence,

$$\text{Rank } (I) = \dim R(I) = \dim V$$

and

$$\text{Nullity } (I) = \dim N(I) = 0.$$

Example 3: Find the rank and nullity of the linear transformation $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ defined as

$$T(x_1, x_2) = (x_1, 0, 0).$$

Solution:

$$\begin{aligned} R(T) &= \left\{Y \in \mathbb{R}^3 : \; Y = T(X) \text{ for some } X \in \mathbb{R}^2\right\} \\ &= \left\{Y \in \mathbb{R}^3 : \; Y = T(X) = (x_1, 0, 0) \text{ for some } X \in \mathbb{R}^2\right\} \\ &= \left\{Y = (x_1, 0, 0) : \; X = (x_1, x_2) \in \mathbb{R}^2\right\} \\ &= \{(x_1, 0, 0) : \; x_1 \in \mathbb{R}\} \\ &= \{x_1(1, 0, 0) : \; x_1 \in \mathbb{R}\} \\ &= \text{Span of } \{(1, 0, 0)\}. \end{aligned}$$

As $(1, 0, 0) \in \mathbb{R}^3$ is a non-zero vector, so $\{(1, 0, 0)\}$ is a linearly independent set in $\mathbb{R}^3$, thus forms a basis of $R(T)$. Hence,

$$Rank(T) = 1.$$

$$
\begin{aligned}
N(T) &= \left\{X \in \mathbb{R}^2 : \; T(X) = \mathbf{0}\right\} \\
&= \{X = (x_1, x_2) : \; T(x_1, x_2) = \mathbf{0}\} \\
&= \{(x_1, x_2) : \; (x_1, 0, 0) = (0, 0, 0)\} \\
&= \{(0, x_2) : \; x_2 \in \mathbb{R}\} \\
&= \{x_2(0, 1) : \; x_2 \in \mathbb{R}\} \\
&= \text{Span of } \{(0, 1)\}.
\end{aligned}
$$

As $(0, 1) \in \mathbb{R}^2$ is a non-zero vector, so $\{(0, 1)\}$ is a linearly independent set in $\mathbb{R}^2$, thus forms a basis of $N(T)$. Hence,

$$Nullity(T) = 1.$$

Example 4: Let $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be a function defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3).$$

Show that T is a linear transformation. Also find rank(T) and nullity(T).

Solution: Note that the function T defined above can also be written in the following form:

$$T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ -1 & -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

This is of the form

$$T(X) = AX.$$

Hence T is a linear transformation.

Now,

$$
\begin{aligned}
R(T) &= \left\{Y \in \mathbb{R}^3 : \; Y = T(X) \text{ for some } X \in \mathbb{R}^3\right\} \\
&= \left\{Y \in \mathbb{R}^3 : \; Y = AX, \; X \in \mathbb{R}^3\right\} \\
&= \left\{AX : \; X \in \mathbb{R}^3\right\} \\
&= \text{Set of linear combinations of columns of } A \\
&= \text{Column space of } A \\
&= \text{Row space of } A^t.
\end{aligned}
$$

To find the row space of A^t we first find the row-reduced echelon form of A^t.

$$A^t = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & -2 \\ 2 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -3 \\ 0 & -4 & 4 \end{bmatrix}$$

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

So,

$$\begin{aligned} R(T) &= \text{Row space of } A^t \\ &= \text{Span } \{(1,0,1),(0,1,-1)\} \\ &= \{a(1,0,1)+b(0,1,-1): \ a,b \in \mathbb{R}\} \\ &= \{(a,b,a-b) : a,b \in R\}. \end{aligned}$$

As the set $\{(1,0,1),(0,1,-1)\}$ is linearly independent and spans the range space $R(T)$, thus forms a basis for $R(T)$. Hence, $Rank(T) = \dim R(T) = 2$.

Now,

$$
\begin{aligned}
N(T) &= \left\{X \in \mathbb{R}^3 : \; T(X) = \mathbf{0}\right\} \\
&= \left\{X \in \mathbb{R}^3 : \; AX = \mathbf{0}\right\} \\
&= \text{Set of solutions of the system } AX = \mathbf{0} \\
&= \text{Solution space of } AX = \mathbf{0}.
\end{aligned}
$$

To find the solution space of $AX = \mathbf{0}$ we first find the row-reduced echelon form of A.

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ -1 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -4 \\ 0 & -3 & 4 \end{bmatrix}$$

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

$$AX = \mathbf{0} \implies x_1 + \frac{2}{3}x_3 = 0, x_2 - \frac{4}{3}x_3 = 0$$

Take $x_3 = a$. This implies $x_1 = -\dfrac{2}{3}a,\ x_2 = \dfrac{4}{3}a.$

Hence,

$$\begin{aligned} N(T) &= \text{Solution space of } AX = \mathbf{0} \\ &= \left\{ \left(-\frac{2}{3}a, \frac{4}{3}a, a \right) : a \in \mathbb{R} \right\} \\ &= \left\{ a \left(-\frac{2}{3}, \frac{4}{3}, 1 \right) : a \in \mathbb{R} \right\} \\ &= \text{Span } \left\{ \left(-\frac{2}{3}, \frac{4}{3}, 1 \right) \right\}. \end{aligned}$$

As the set $\left\{ \left(-\frac{2}{3}, \frac{4}{3}, 1 \right) \right\}$ is linearly independent and spans the null space $N(T)$, thus forms a basis for $N(T)$. Hence

$$nullity(T) = \dim N(T) = 1.$$

Some remarks

- The method that we used to find the rank and nullity of the linear transformation in Example-4 can also be used in Example-3.
- In all the four examples above, we have

$$rank(T) + nullity(T) = dim(V),$$

  where V represents the domain set of the linear transformation T.
- The above observation is indeed a fact, which holds in arbitrary finite dimensional vector space. In the next theorem we shall prove this fact.

Theorem 2 (Rank-Nullity-Dimension Theorem)

Note: This theorem is also known as the **Rank-Nullity Theorem**.

Statement: Let V and W be vector spaces over the field F and let $T : V \longrightarrow W$ be a linear transformation. Suppose that V is finite-dimensional. Then

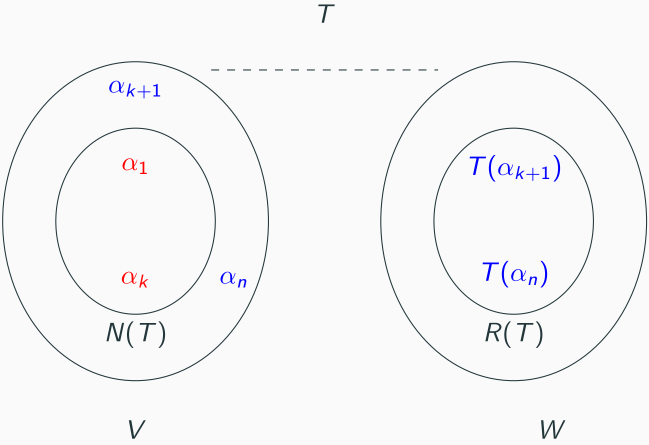
$$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

Proof: Let $\dim(V) = n$. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ be a basis for $N(T)$. This means nullity $(T) = k$. As $N(T)$ is a subspace of V, we have $k \leq n$.

Since $\{\alpha_1, \alpha_2, \ldots, \alpha_k\} \subseteq V$ and it is linearly independent in V, using Corollary 2 of Theorem 5 we can extend this linearly independent set to a basis of V. That is, there are non-zero vectors $\alpha_{k+1}, \ldots, \alpha_n \in V$ such that $\{\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n\}$ is a basis for V.

We prove that $B = \{T(\alpha_{k+1}), \ldots, T(\alpha_n)\}$ is a basis for $R(T)$.

Theorem 2 contd.



Theorem 2 contd.

Claim 1: The set $\{T(\alpha_{k+1}), \ldots, T(\alpha_n)\}$ spans $R(T)$.

Let $\beta \in R(T)$. Then by the definition of range of T there exists $\alpha \in V$ such that $\beta = T(\alpha)$. Since $\alpha \in V = \text{Span}\,\{\alpha_1, \ldots, \alpha_n\}$, there exist scalars $c_1, c_2, \ldots, c_n$ such that

$$\alpha = c_1\alpha_1 + \cdots + c_k\alpha_k + c_{k+1}\alpha_{k+1} + \cdots + c_n\alpha_n.$$

So,

$$\beta = T(\alpha) = T(c_1\alpha_1 + \cdots + c_k\alpha_k + c_{k+1}\alpha_{k+1} + \cdots + c_n\alpha_n).$$

As T is a linear transformation, we have

$$\beta = c_1 T(\alpha_1) + \cdots + c_k T(\alpha_k) + c_{k+1} T(\alpha_{k+1}) + \cdots + c_n T(\alpha_n).$$

But, $\alpha_1, \ldots, \alpha_k \in N(T)$, so we have $T(\alpha_1) = 0, \ldots, T(\alpha_k) = 0$.

Thus,

$$\beta = c_{k+1}T(\alpha_{k+1}) + \cdots + c_n T(\alpha_n).$$

Hence, the set $\{T(\alpha_{k+1}), \ldots, T(\alpha_n)\}$ spans $R(T)$.

Claim 2: $\{T(\alpha_{k+1}), \ldots, T(\alpha_n)\}$ **is an L.I. set.**

Consider

$$c_{k+1}T(\alpha_{k+1}) + \cdots + c_n T(\alpha_n) = \mathbf{0}.$$

As T is a linear transformation, we have

$$T(c_{k+1}\alpha_{k+1} + \cdots + c_n\alpha_n) = \mathbf{0}.$$

This implies

$$c_{k+1}\alpha_{k+1} + \cdots + c_n\alpha_n \in N(T).$$

But,

$$N(T) = \text{Span } \{\alpha_1, \ldots, \alpha_k\}.$$

So, there exist scalars $b_1, \ldots, b_k \in F$ such that

$$c_{k+1}\alpha_{k+1} + \cdots + c_n\alpha_n = b_1\alpha_1 + \cdots + b_k\alpha_k$$

$$\implies b_1\alpha_1 + \cdots + b_k\alpha_k - c_{k+1}\alpha_{k+1} - \cdots - c_n\alpha_n = \mathbf{0}.$$

Since $\{\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n\}$ is an L.I. set, we have

$$b_1 = \cdots = b_k = -c_{k+1} = \cdots = -c_n = 0.$$

This shows that

$$c_{k+1}T(\alpha_{k+1}) + \cdots + c_nT(\alpha_n) = \mathbf{0} \;\text{ implies }\; c_{k+1} = \cdots = c_n = 0.$$

This proves Claim 2. By Claims 1 and 2, B is a basis of $R(T)$ and dim $R(T) = |B| = n - k$. This implies $\dim R(T) = \dim(V) - \dim N(T)$. Hence,

$$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

Theorem 3.

If A is an $m \times n$ matrix, then row rank $(A) =$ column rank (A).

Proof: We define a linear transformation $T : F^{n\times 1} \longrightarrow F^{m\times 1}$ by $T(X) = AX$. By Rank-Nullity-Dimension Theorem,

$$\text{rank } (T) + \text{nullity } (T) = \dim V = \dim F^{n\times 1} = n - - - -(1).$$

Now,

$$\begin{aligned} R(T) &= \left\{Y \in F^{m\times 1} : T(X) = Y \text{ for some } X \in F^{n\times 1}\right\} \\ &= \left\{Y \in F^{m\times 1} : AX = Y \text{ for some } X \in F^{n\times 1}\right\} \\ &= \left\{AX : X \in F^{n\times 1}\right\} \\ &= \text{ Set of all linear combinations of columns of } A \\ &= \text{Column space } (A) \end{aligned}$$

$$\text{rank } (T) = \dim R(T) = \dim \text{column space } (A) = \text{column rank } (A) - (2)$$

Theorem 3 contd.

Now,

$$\begin{aligned} N(T) &= \left\{X \in F^{n\times 1} : T(X) = \mathbf{0}\right\} \\ &= \left\{X \in F^{n\times 1} : AX = \mathbf{0}\right\} \\ &= S \text{ (the solution space of } AX = \mathbf{0}.) \end{aligned}$$

Let R be the row-reduced echelon matrix row-equivalent to A. Let r be the number of non-zero rows of R.

$$r = \text{ row rank } (R) = \text{ row rank } (A) - - - -(3)$$

The system $RX = \mathbf{0}$ has $n - r$ free variables, thus

$$n - r = \dim\ S = \dim\ N(T) = \text{nullity}(T) - -(4)$$

Theorem 3 contd.

From (1), (2) and (4),

$$\text{column rank } (A) + n - r = n$$

$$\implies \text{column rank } (A) = r = \text{row rank } (A), \quad \text{by (3)}$$

This completes the proof.

Definition:

$$\text{rank } (A) = \text{column rank } (A) = \text{row rank } (A).$$

Problem 3

Find a linear transformation (if exists) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ such that $N(T) = \text{Span } \{(1,1,1)\}$ and $R(T) = \text{Span}\,\{(1,0,-1),(1,2,2)\}$.

Solution: It is given that $\{\alpha_1 = (1,1,1)\}$ is a basis for $N(T)$. By extending this basis we construct a basis for $V = \mathbb{R}^3$, say $\{\alpha_1 = (1,1,1), \alpha_2 = (0,1,1), \alpha_3 = (0,0,1)\}$ (We have solved similar problems in the past!).

Note that $\beta_1 = (0,0,0),\ \beta_2 = (1,0,-1),\ \beta_3 = (1,2,2) \in R(T)$. Let us construct T such that

$$T(\alpha_1) = T(1,1,1) = \beta_1 = (0,0,0),$$

$$T(\alpha_2) = T(0,1,1) = \beta_2 = (1,0,-1)$$

and

$$T(\alpha_3) = T(0,0,1) = \beta_3 = (1,2,2).$$

Problem 3 contd.

Let

$$(x, y, z) = a\alpha_1 + b\alpha_2 + c\alpha_3 = a(1,1,1) + b(0,1,1) + c(0,0,1)$$

$$\implies (x, y, z) = x\alpha_1 + (y - x)\alpha_2 + (z - y)\alpha_3$$

$$\implies \quad T(x, y, z) = x\beta_1 + (y - x)\beta_2 + (z - y)\beta_3$$

$$\implies T(x, y, z) = x(0,0,0) + (y - x)(1, 0, -1) + (z - y)(1, 2, 2)$$

$$\implies \quad T(x, y, z) = (-x + z, -2y + 2z, x - 3y + 2z)$$

$$\implies \quad T(x, y, z) = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 2 \\ 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

$L(V, W)$: Set of all linear transformations from V into W.

Let V, W be vector spaces over the field F. Define the set

$$L(V, W) = \{T : \ T : V \longrightarrow W \text{ is a L.T. }\}.$$

Observation 1: $L(V, W)$ is a vector space under the operations

$$(T + U)(\alpha) = T(\alpha) + U(\alpha)$$

and

$$(cT)(\alpha) = cT(\alpha)$$

for all $T, U \in L(V, W)$ and $c \in F$.

Observation 2: If V and W are finite dimensional vector spaces, then

$$\dim L(V, W) = \dim V \cdot \dim W.$$

Linear Operator

Definition:

If V is a vector space over the field F, then a linear operator T is a linear transformation from V into V.

One to one (1:1) function.

A function $f : X \longrightarrow Y$ is said to be an one to one function if each element in X has exactly one image in Y. In other words,

$$\text{if } f(x) = f(y), \text{ then } x = y.$$

Onto function.

A function $f : X \longrightarrow Y$ is said to be an onto function if the range of f is Y.

Invertible function.

A function $f : X \longrightarrow Y$ is said to be an invertible function if there exists a function $g : Y \longrightarrow X$ such that

(i) $gof : X \longrightarrow X$ and

(ii) $fog : Y \longrightarrow Y$ are identity functions.

Proposition 1. A function $f : X \longrightarrow Y$ is invertible if and only if f is one-to-one and onto.

If T is linear then T^{-1} is linear

Theorem 4. Let V and W be two vector spaces over the field F and let $T : V \longrightarrow W$ be a linear transformation. If T is invertible, then the inverse function $T^{-1} : W \longrightarrow V$ is a linear transformation.

Proof: Suppose that $T : V \longrightarrow W$ is an invertible linear transformation. Then there exists a function $T^{-1} : W \longrightarrow V$ such that $TT^{-1} : W \longrightarrow W$ and $T^{-1}T : V \longrightarrow V$ are identity functions.

We want to show that

$$T^{-1}(c\beta_1 + \beta_2) = cT^{-1}(\beta_1) + T^{-1}(\beta_2) \text{ for all } \beta_1, \beta_2 \in W, c \in F.$$

Let $\alpha_1 = T^{-1}(\beta_1)$ and $\alpha_2 = T^{-1}(\beta_2)$. Since T is invertible, α_1, α_2 are unique vectors in V such that $T(\alpha_1) = \beta_1$ and $T(\alpha_2) = \beta_2$. Since T is a linear transformation,

$$T(c\alpha_1 + \alpha_2) = cT(\alpha_1) + T(\alpha_2) = c\beta_1 + \beta_2.$$

$$\Longrightarrow T^{-1}(c\beta_1 + \beta_2) = T^{-1}T(c\alpha_1 + \alpha_2)$$

$$\Longrightarrow T^{-1}(c\beta_1 + \beta_2) = c\alpha_1 + \alpha_2 = cT^{-1}(\beta_1) + T^{-1}(\beta_2).$$

Hence T^{-1} is a linear transformation.

Problem 4. Let $T(x_1, x_2) = (x_1 + x_2, x_1)$ be a linear operator defined on F^2. Find T^{-1} if exists.

Solution:

$$T(x_1, x_2) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Longrightarrow T(X) = AX.$$

Now,

$$[A|I] = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 \end{array}\right] = \left[I|A^{-1}\right].$$

Thus,

$$T^{-1}(X) = A^{-1}X = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

This implies

$$T^{-1}(x_1, x_2) = (x_2, x_1 - x_2).$$

Problem 5. Find the inverse of a linear operator T on $\mathbb{R}^3$ defined as

$$T(x_1, x_2, x_3) = (3x_1, x_1 - x_2, 2x_1 + x_2 + x_3).$$

Solution:

$$T(x_1, x_2, x_3) = \begin{bmatrix} 3 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Now,

$$[A|I] = \left[\begin{array}{ccc|ccc} 3 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{3} & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right] = [I|A^{-1}].$$

Thus,

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

This implies

$$T^{-1}(x_1, x_2, x_3) = \left(\frac{1}{3}x_1, \frac{1}{3}x_1 - x_2, -x_1 + x_2 + x_3\right).$$

Definition. A linear transformation $T : V \longrightarrow W$ is **non-singular**

$$\text{if} \quad T(\alpha) = \mathbf{0} \quad \text{implies} \quad \alpha = \mathbf{0}.$$

$$\text{That is,} \quad N(T) = \{\mathbf{0}\}.$$

Lemma 1. Let $T : V \longrightarrow W$ be a linear transformation. Then the following statements are equivalent.

(1) T is one-to-one.

(2) T is non-singular.

Proof: (1) $\Longrightarrow$ (2). Suppose that T is one-to-one. Let $T(\alpha) = \mathbf{0}$. As T is a linear transformation we have $T(\mathbf{0}) = \mathbf{0}$. This implies $T(\alpha) = T(\mathbf{0})$. But T is one-to-one, so $\alpha = \mathbf{0}$. This shows T is non-singular.

(2) $\Longrightarrow$ (1). Suppose that T is non-singular. Then by definition $N(T) = \{\mathbf{0}\}$.

$$\begin{aligned}
\text{Let} \quad & T(\alpha) = T(\beta) \\
\Longrightarrow \quad & T(\alpha) - T(\beta) = \mathbf{0} \\
\Longrightarrow \quad & T(\alpha - \beta) = \mathbf{0} \quad (\because \text{ T is a L.T.}) \\
\Longrightarrow \quad & \alpha - \beta \in N(T) = \{\mathbf{0}\} \\
\Longrightarrow \quad & \alpha = \beta.
\end{aligned}$$

This proves T is one-to-one.

Non-singular linear transformations preserve linear independence

Theorem 5. Let $T : V \longrightarrow W$ be a linear transformation. Then T is non-singular if and only if T carries each linearly independent subset of V onto a linearly independent subset of W.

Proof:

Case 1: Suppose that T is non-singular. Then by definition $N(T) = \{\mathbf{0}\}$. Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ be a linearly independent set V. We show that $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_k)\}$ is linearly independent in W.

$$\begin{aligned}
\text{Let} \quad & c_1 T(\alpha_1) + c_2 T(\alpha_2) + \cdots + c_k T(\alpha_k) = \mathbf{0} \\
\implies \quad & T(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_k\alpha_k) = \mathbf{0} \quad (\because \text{ T is a L.T.}) \\
\implies \quad & c_1\alpha_1 + c_2\alpha_2 + \cdots + c_k\alpha_k \in N(T) = \{\mathbf{0}\} \\
\implies \quad & c_1\alpha_1 + c_2\alpha_2 + \cdots + c_k\alpha_k = \mathbf{0}.
\end{aligned}$$

As $S = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is linearly independent, we have

$$c_1 = c_2 = \cdots = c_k = 0.$$

This shows if

$$c_1 T(\alpha_1) + c_2 T(\alpha_2) + \cdots + c_k T(\alpha_k) = \mathbf{0},$$

then

$$c_1 = c_2 = \cdots = c_k = 0.$$

Hence, $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_k)\}$ is a linearly independent subset of W. This completes the proof of Case 1.

Case 2: Suppose that T carries linearly independent subset onto linearly independent subset.

Let $T(\alpha) = \mathbf{0}$. If $\alpha \neq \mathbf{0}$, then T carries a linearly independent set $\{\alpha\}$ onto a linearly dependent set $\{T(\alpha)\} = \{\mathbf{0}\}$, a contradiction. Thus, $\alpha = \mathbf{0}$. This implies T is non-singular.

Theorem 6. Let V and W be finite dimensional vector spaces over the field F such that $\dim\ V = \dim\ W$. If $T : V \longrightarrow W$ is a linear transformation, then the followings are equivalent.

(i) T is invertible.

(ii) T is non-singular.

(iii) T is onto.

(iv) T carries a basis of V to a basis of W. That is, if $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis for V, then $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is a basis for W.

Proof. Let $\dim\ V = \dim\ W = n$. By Rank-Nullity-Dimension Theorem,

$$\text{rank}\ (T) + \text{nullity}\ (T) = \dim\ V = n. \tag{1}$$

$(i) \Longrightarrow (ii)$. Assume that T is invertible. So, by proposition 1, T is one-to-one. Then Lemma 1 implies T is non-singular.

$(ii) \Longrightarrow (iii)$. Assume that T is non-singular. Then by definition $N(T) = \{\mathbf{0}\}$. So, nullity $(T) = \mathbf{0}$. From equation (1) we have rank $(T) = n$. Thus, $\dim R(T) = \dim W$. This implies

$$R(T) = W \quad (\because\ R(T) \subseteq W \text{ and } \dim R(T) = \dim W).$$

Hence, T is onto.

$(iii) \implies (iv)$. Assume that T is onto. That is $R(T) = W$. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis for V. Our aim is to show that $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ is a basis for W.

First, we prove that $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ spans W. Let $\beta \in W = R(T)$. Since T is onto, there exists $\alpha \in V$ such that $T(\alpha) = \beta$. Since $\alpha \in V$ and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis for V, there exists scalars $c_1, c_2, \ldots, c_n$ such that $\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$. So,

$$\beta = T(\alpha) = c_1 T(\alpha_1) + c_2 T(\alpha_2) + \cdots + c_n T(\alpha_n).$$

This implies $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ spans $R(T) = W$. Since dim $W = n$, $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ is a basis for $R(T) = W$.

$(iv) \implies (i)$. Let $\{\alpha_1, \dots, \alpha_n\}$ be a basis for V. By our assumption $\{T(\alpha_1), \dots, T(\alpha_n)\}$ forms a basis for $R(T)$. Since $\dim\ W = n = \dim\ R(T)$ and $R(T) \subseteq W$, we must have $R(T) = W$. Thus, T is onto.

Next, we show that T is one-to-one. That is $N(T) = \{\mathbf{0}\}$. Let $\alpha \in N(T)$. As we have a basis $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ for V, there exists scalars $c_1, c_2, \dots, c_n$ such that $\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$. So,

$$0 = T(\alpha) = c_1 T(\alpha_1) + c_2 T(\alpha_2) + \cdots + c_n T(\alpha_n).$$

Since, $\{T(\alpha_1), \dots, T(\alpha_n)\}$ is a basis, hence an L. I. set. So, $c_1 = c_2 = \cdots = c_n = 0$. This implies $\alpha = 0$. So, T is one-to-one. Hence T is invertible by Proposition 1.

If A is a given $m \times n$ matrix, then we can define a linear transformation from $\mathbb{R}^n$ into $\mathbb{R}^m$ by

$$T(x) = Ax.$$

What about the converse?

Theorem 7. Let V be an n-dimensional vector space over the field F and W an m-dimensional vector space over F. Let B be an ordered basis for V and B' an ordered basis for W. For each linear transformation $T : V \longrightarrow W$ there is an $m \times n$ matrix A with entries in F such that

$$[T(\alpha)]_{B'} = A[\alpha]_B,$$

for every vector $\alpha \in V$. Furthermore $T \longrightarrow A$ is a one-to-one correspondence between the set of all linear transformations from V into W and the set of all $m \times n$ matrices over the field F.

Proof.

Note that $T(\alpha_j) \in W$. Since $\{\beta_1, \ldots, \beta_m\}$ is a basis for W there exist unique scalars $A_{1j}, A_{2j}, \ldots, A_{mj}$ such that

$$T(\alpha_j) = A_{1j}\beta_1 + A_{2j}\beta_2 + \cdots + Amj\beta_m = \sum_{i=1}^{m} A_{ij}\beta_i$$

for $j = 1, 2, \ldots, n$. Therefore

$$[T(\alpha_j)]_{B'} = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ Amj \end{bmatrix}$$

for $j = 1, 2, \ldots, n$.

Define the matrix

$$A = \left[[T(\alpha_1)]_{B'}, [T(\alpha_2)]_{B'}, \ldots, [T(\alpha_n)]_{B'}\right] = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ A_{m1} & A_{m2} & \ldots & A_{mn} \end{bmatrix}.$$

This $m \times n$ matrix A is called the matrix of T relative to the ordered bases B, B' or the matrix of T relative to B, B'. The matrix A is denoted by

$$A = [T]_B^{B'}.$$

Our aim is to understand explicitly how the matrix A determines the linear transformation T.

We claim that

$$[T(\alpha)]_{B'} = A\,[\alpha]_B.$$

Proof. Let $\alpha \in V$. As $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis for V there exist unique scalars $x_1, x_2, \ldots, x_n$ such that

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \sum_{j=1}^{n} x_j\alpha_j.$$

Since T is a linear transformation, we have

$$\begin{aligned} T(\alpha) &= T\left(\sum_{j=1}^{n} x_j\alpha_j\right) = \sum_{j=1}^{n} x_j\,T(\alpha_j) \\ &= \sum_{j=1}^{n} x_j\left(\sum_{i=1}^{m} A_{ij}\beta_i\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x_j\right)\beta_i. \end{aligned}$$

So,

$$
\begin{aligned}
[T(\alpha)]_{B'} &= \begin{bmatrix} \sum_{j=1}^{n} A_{1j}x_j \\ \sum_{j=1}^{n} A_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} A_{mj}x_j \end{bmatrix} \\
&= \begin{bmatrix} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n \\ \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n \end{bmatrix} \\
&= \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ A_{m1} & A_{m2} & \ldots & A_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.
\end{aligned}
$$

This implies

$$[T(\alpha)]_{B'} = A\,[\alpha]_B,$$

$$\text{where } A = [[T(\alpha_1)]_{B'}, [T(\alpha_2)]_{B'}, \ldots, [T(\alpha_n)]_{B'}].$$

This completes the proof of the theorem.

Problem 6. Let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ be a linear transformation defined as

$$T(x_1, x_2) = (x_2, x_1 - x_2, x_1 + x_2)\,.$$

Let

$$B = \{\alpha_1 = (1, 0), \alpha_2 = (0, 1)\}$$

and

$$B' = \{\beta_1 = (1, 1, 1), \beta_2 = (1, 1, 0), \beta_3 = (1, 0, 0)\}$$

be respective ordered bases for $\mathbb{R}^2$ and $\mathbb{R}^3$. Find $[T]_B^{B'}$.

Solution.

$$\begin{aligned} T(\alpha_1) &= T(1, 0) \\ &= (0, 1, 1) \\ &= (1, 1, 1) + 0(1, 1, 0) - (1, 0, 0) \\ &= \beta_1 + 0\beta_2 - \beta_3. \end{aligned}$$

So, $[T(\alpha_1)]_{B'} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

Similarly,

$$\begin{aligned} T(\alpha_2) &= T(0,1) \\ &= (1,-1,1) \\ &= (1,1,1) - 2(1,1,0) + 2(1,0,0) \\ &= \beta_1 - 2\beta_2 + 2\beta_3. \end{aligned}$$

So, $[T(\alpha_2)]_{B'} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$.

Hence,

$$A = [T]_B^{B'} = \big[[T(\alpha_1)]_{B'}, \;\; [T(\alpha_2)]_{B'}\big] = \begin{bmatrix} 1 & 1 \\ 0 & -2 \\ -1 & 2 \end{bmatrix}.$$

Note: Let V be a finite dimensional vector space and B an ordered basis for V. If $T : V \longrightarrow V$ is a linear operator, then A is denoted as $[T]_B$. So by Theorem 7 we have

$$[T(\alpha)]_B = [T]_B \, [\alpha]_B \, .$$

The matrix $[T]_B$ is called the matrix of T relative to the ordered basis B.

Problem 7. Let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear transformation defined as $T(x_1, x_2) = (x_1, 0)$. Let $B = \{\alpha_1 = (1,1), \alpha_2 = (1,2)\}$ be an ordered basis for $\mathbb{R}^2$. Find $[T]_B$.

Solution.

$$T(\alpha_1) = T(1,1) = (1,0) = 2\alpha_1 - \alpha_2.$$

So,

$$[T(\alpha_1)]_B = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Similarly,

$$T(\alpha_2) = T(1,2) = (1,0) = 2\alpha_1 - \alpha_2.$$

So,

$$[T(\alpha_2)]_B = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Hence, $[T]_B = ([T(\alpha_1)]_B, [T(\alpha_2)]_B) = \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix}$.

Problem 8. Let $\mathbb{P}_3$ be the vector space of all real polynomials of degree at most three and $\mathbb{P}_2$ be the vector space of all real polynomials of degree at most two. Let D be the differentiation transformation from $\mathbb{P}_3$ into $\mathbb{P}_2$. Let $B = \{1, x, x^2, x^3\}$ and $B' = \{1, x, x^2\}$ be two ordered bases for $\mathbb{P}_3$ and $\mathbb{P}_2$, respectively. Find $[D]_B^{B'}$.

Solution.

$$D(\alpha_1) = D(1) = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 = 0 \cdot \beta_1 + 0 \cdot \beta_2 + 0 \cdot \beta_3$$

$$D(\alpha_2) = D(x) = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 = 1 \cdot \beta_1 + 0 \cdot \beta_2 + 0 \cdot \beta_3$$

$$D(\alpha_3) = D(x^2) = 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 = 0 \cdot \beta_1 + 2 \cdot \beta_2 + 0 \cdot \beta_3$$

$$D(\alpha_4) = D(x^3) = 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2 = 0 \cdot \beta_1 + 0 \cdot \beta_2 + 3 \cdot \beta_3$$

Hence,

$$[D]_B^{B'} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Theorem 8. Let V be a finite dimensional vector space over the field F and let $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ and $B' = \{\beta_1, \beta_2, \dots, \beta_n\}$ be two ordered bases for V. Suppose $T : V \longrightarrow V$ is a linear operator. If $P = [P_1, P_2, \dots, P_n]$ is the $n \times n$ matrix with columns $P_j = [\beta_j]_B$, then

$$[T]_{B'} = P^{-1} [T]_B P.$$

Proof: Reading assignment.

Similar matrices.

Let A and B be $n \times n$ matrices over the field F. We say B is similar to A over F if there exists an invertible $n \times n$ matrix P over F such that

$$B = P^{-1}AP.$$

Note. From Theorem 8, it follows that matrices $[T]_B$ and $T_{B'}$ are similar.

Problem 9. Let T be a linear operator on $\mathbb{R}^2$ defined as $T(x_1, x_2) = (x_1, 0)$. Let $B = \{\alpha_1 = (1,1), \alpha_2 = (1,2)\}$ be an ordered basis for $\mathbb{R}^2$. Let $B' = \{\beta_1 = (1,0), \beta_2 = (0,1)\}$ denotes the standard basis of $\mathbb{R}^2$. Find a matrix P such that $[T]_{B'} = P^{-1}[T]_B P$.

Solution. From Problem 7 we know that $[T]_B = \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix}$.

Similarly, we can show that $[T]_{B'} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Now we find the matrix P. Note that

$$\beta_1 = (1,0) = 2(1,1) - (1,2) = 2\alpha_1 - \alpha_2$$

and

$$\beta_2 = (0,1) = -(1,1) + (1,2) = -\alpha_1 + \alpha_2.$$

So,

$$P = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

and

$$P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Therefore

$$P^{-1}[T]_B P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = [T]_{B'}\,.$$

Problem 10. Let $\mathbb{P}_3$ be the vector space of all real polynomials of degree at most three. Let D be the differentiation operator on $\mathbb{P}_3$. Let $B = \{1, x, x^2, x^3\}$ and $B' = \{1, 2x, -3x^2, 2x^3\}$ be two ordered bases for $\mathbb{P}_3$. Find a matrix P such that $[D]_{B'} = P^{-1}[D]_B P$.

Eigenvalues / Characteristic values / Characteristic roots

Definitions.

- Let A be an $n \times n$ (square) matrix over the field F. A scalar $\lambda \in F$ is an eigenvalue of A if there exists a non-zero vector $X \in F^{n\times 1}$ such that
$$AX = \lambda X.$$

- Any non-zero vector X such that $AX = \lambda X$ is called an eigenvector of A corresponding to the eigenvalue λ.

- $E_A(\lambda) = \{X : AX = \lambda X\} = \{X : (\lambda I - A)X = \mathbf{0}\}$ is called the eigenspace of A associated to λ.

Proposition 2. The eigenspace $E_A(\lambda)$ associated with the eigenvalue λ is a subspace of $F^{n\times 1}$.

How to find the eigenvalues.

Let A be a given $n \times n$ matrix and λ be an eigenvalue of A. Then by definition there exists a non-zero vector X such that $AX = \lambda X$. This implies the system

$$(\lambda I - A)X = \mathbf{0}$$

has a non-trivial solution. This holds if and only if $(\lambda I - A)$ is not invertible. This holds if and only if

$$\det(\lambda I - A) = 0.$$

Characteristic Polynomial.

Let A be an $n \times n$ matrix over the field F. The polynomial $f(x) = \det(xI - A)$ is called the **characteristic polynomial of** A.

Note. Thus, the eigenvalues of a matrix A are the roots of the characteristic polynomial of matrix A.

How to find the eigenvectors.

Fix one eigenvalue λ of matrix A. Then solve the system $(\lambda I - A)X = 0$. The non-zero solutions are the eigenvectors of matrix A corresponding to the eigenvalue λ.

Problem 11. Find the eigenvalues and the corresponding eigenvectors of the matrix $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$.

Solution: Consider

$$\det\,(\lambda I - A) = 0$$

$$\implies \det \begin{pmatrix} \lambda - 1 & -2 \\ 0 & \lambda - 2 \end{pmatrix} = 0$$

$$\implies (\lambda - 1)(\lambda - 2) = 0$$

$$\implies \lambda = 1,\ 2.$$

So, $\lambda_1 = 1,\ \lambda_2 = 2$ are two eigenvalues of the given matrix.

The eigenspace corresponding to $\lambda = 1$.

$E_A(1) = \{X : (\lambda I - A)X = \mathbf{0}\} = \{X : (I - A)X = \mathbf{0}\} = \{X : (A - I)X = \mathbf{0}\}$.

Consider the system of equations

$$(A - I)X = \mathbf{0}$$

$$\implies \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$$\implies y = 0.$$

The solutions of this system are of the form $(a, 0)$, where $a \in \mathbb{R}$. Hence

$$E_A(1) = \{(a, 0) : a \in \mathbb{R}\} = \{a(1, 0) : a \in \mathbb{R}\} = \text{ Span } \{(1, 0)\}.$$

Each nonzero vector in $E_1(A)$ is an eigenvector of A corresponding to the eigenvalue $\lambda = 1$.

The eigenspace corresponding to $\lambda = 2$.

$$E_A(2) = \{X : (2I - A)X = \mathbf{0}\} = \{X : (A - 2I)X = \mathbf{0}\}.$$

Consider the system of equations

$$(A - 2I)X = \mathbf{0}$$

$$\implies \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$$\implies x = 2y.$$

Set $y = a$, then $x = 2a$. Thus, the solutions of this system are of the form $(2a, a)$, where $a \in \mathbb{R}$. Hence

$$E_A(2) = \{(2a, a) : a \in \mathbb{R}\} = \{a(2, 1) : a \in \mathbb{R}\} = \text{ Span } \{(2, 1)\}.$$

Each nonzero vector in $E_2(A)$ is an eigenvector of A corresponding to the eigenvalue $\lambda = 2$.

Problem 12. Find the eigenvalues and corresponding eigenspaces of the matrix

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}.$$

Solution. The characteristic polynomial of A

$$f_A(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - 5 & 6 & 6 \\ 1 & \lambda - 4 & -2 \\ -3 & 6 & \lambda + 4 \end{vmatrix} = (\lambda - 2)^2(\lambda - 1).$$

$$\Longrightarrow \lambda = 1, 2, 2.$$

Hence eigenvalues of A are $1, 2$.

The eigenspace corresponding to $\lambda = 1$.

$$E_A(1) = \{X : (I - A)X = \mathbf{0}\} = \{X : (A - I)X = \mathbf{0}\}.$$

$$A - I = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

$$(A - I)X = \mathbf{0} \Longrightarrow x_1 - x_3 = 0,\ x_2 + \frac{1}{3}x_3 = 0$$

Note that (i) pivot variables $= \{x_1,\ x_2\}$ and (ii) free variables $= \{x_3\}$. Let $x_3 = a$. This implies $x_1 = a$ and $x_2 = -\frac{a}{3}$. Thus

$$E_A(1) = \left\{\left(a, -\frac{a}{3}, a\right) : a \in R\right\} = \left\{\frac{a}{3}(3, -1, 3) : a \in R\right\} = \text{ Span}\left\{(3, -1, 3)\right\}.$$

The eigenspace corresponding to $\lambda = 2$.

$$E_A(2) = \{X : (2I - A)X = \mathbf{0}\} = \{X : (A - 2I)X = \mathbf{0}\}.$$

$$A - 2I = \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$(A - 2I)X = \mathbf{0} \Longrightarrow x_1 - 2x_2 - 2x_3 = 0.$$

Note that (i) pivot variables $= \{x_1\}$ and (ii) free variables $= \{x_2, x_3\}$. Let $x_2 = a$ and $x_3 = b$. This implies $x_1 = 2a + 2b$. Therefore

$$E_A(2) = \{(2a + 2b, a, b) : a, b \in R\} = \{a(2, 1, 0) + b(2, 0, 1) : a, b \in R\}.$$

Thus,

$$E_A(2) = \text{ Span } \{(2, 1, 0), (2, 0, 1)\}.$$

Notes. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $\lambda I - A = \begin{pmatrix} \lambda - a & -b \\ -c & \lambda - d \end{pmatrix}$

The characteristic polynomial of A is

$$f_A(\lambda) = det(\lambda I - A) = (\lambda - a)(\lambda - d) - bc = \lambda^2 - (a + d)\lambda + ad - bc.$$

That is

$$f_A(\lambda) = \lambda^2 - \text{trace }(A)\lambda + \text{det }(A).$$

But, if λ_1, λ_2 are the roots of the characteristic polynomial, then

$$f_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2$$

Hence,

$$\lambda_1 + \lambda_2 = \text{trace }(A) \text{ and } \lambda_1\lambda_2 = \text{det }(A).$$

Notes. Let A be an $n \times n$ matrix over the field $\mathbb{F}$.

- The characteristic polynomial of A is of the form

$$f_A(\lambda) = \lambda^n + (-1)^1 \text{trace } (A)\lambda^{n-1} + \cdots + (-1)^n \det(A).$$

- If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then

$$\text{trace } (A) = \lambda_1 + \cdots + \lambda_n$$

  and

$$\det(A) = \lambda_1 \cdots \lambda_n.$$

Theorem 9. (Cayley-Hamilton theorem)
Every square matrix satisfies its own characteristic polynomial.

That is, if A is an $n \times n$ matrix over the field $\mathbb{F}$ and $f(\lambda)$ is the characteristic polynomial of A, then $f(A) = 0$.

Example. Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$. Then from problem 11 we know that the characteristic polynomial of A is $f(\lambda) = \lambda^2 - 3\lambda + 2$. Then

$$f(A) = A^2 - 3A + 2I = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (\textit{verify}!)$$

Applications of Cayley-Hamilton theorem

Application 1. Let $f(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ be the characteristic polynomial of A. By Cayley-Hamilton theorem

$$\begin{aligned} & f(\lambda) = 0 \\ \implies \quad & A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I = 0 \\ \implies \quad & A^n = -c_{n-1}A^{n-1} - \cdots - c_1 A - c_0 I. \end{aligned}$$

Thus, all the higher powers of A starting from n can be calculated as a linear combination of lower powers $A^0, A^1, \ldots, A^{n-1}$.

Applications of Cayley-Hamilton theorem

Application 2. Let $f(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ be the characteristic polynomial of A. By Cayley-Hamilton theorem

$$\begin{aligned} & f(A) = 0 \\ \implies \quad & A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I = 0 \\ \implies \quad & A(A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I) = -c_0 I \\ \implies \quad & A\left(\frac{-1}{c_0}A^{n-1} - \frac{c_{n-1}}{c_0}A^{n-2} - \cdots - \frac{c_1}{c_0}I\right) = I \end{aligned}$$

Thus,

$$A^{-1} = \frac{-1}{c_0}A^{n-1} - \frac{c_{n-1}}{c_0}A^{n-2} - \cdots - \frac{c_1}{c_0}I.$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$. Then from problem 11 we know that the characteristic polynomial of A is $\lambda^2 - 3\lambda + 2$. By Cayley-Hamilton theorem

$$\begin{aligned} & f(A) = 0 \\ \implies \quad & A^2 - 3A + 2I = 0 \\ \implies \quad & A(A - 3I) = -2I \\ \implies \quad & A\left(\frac{-1}{2}A - \frac{3}{2}I\right) = I \end{aligned}$$

Thus,

$$A^{-1} = \frac{-1}{2}A + \frac{3}{2}I = \begin{pmatrix} 1 & -1 \\ 0 & \frac{1}{2} \end{pmatrix}.$$