MATH 311 - Vector Spaces in $\mathbb{R}^n$

Jasraj Sandhu

May 2023

Proofs

Determine if the following statements are true or false. It the statement is true, provide a proof. If the statement is false, prove its negation is true by giving a counterexample.

1. If A is skew-symmetric, then A^{-1} is skew-symmetric.

   It is true.

   Proof. Suppose A is skew-symmetric. This means that $A^\top = -A$. We want to prove that $(A^{-1})^\top = -A^{-1}$ Then

$$\begin{aligned}(A^{-1})^\top &= (A^\top)^{-1} \\ &= (-A)^{-1} \\ &= -A^{-1} \ .\end{aligned}$$

   Hence, A^{-1} is skew-symmetric. □

2. If A is skew-symmetric, then A^{-1} is symmetric.

   It is false. We prove the negation: There exists a skew-symmetric matrix A wuch that A^{-1} is not symmetric.

   Proof. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then A is skew-symmetric since

$$A^\top = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -A \ .$$

   However,

$$\begin{aligned}A^{-1} &= \frac{1}{0-(-1)} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &\neq \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= (A^{-1})^\top \ .\end{aligned}$$

   Thus, A^{-1} is not symmetric. □

3. Suppose A and B are skew-symmetric.

   (a) If $AB = BA$, then AB is skew-symmetric.

   It is false. We prove the negation: There exist skew-symmetric matrices A and B such that $AB = BA$, but AB is not skew-symmetric.

   Proof. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$. Then A and B are skew-symmetric since

$$\begin{aligned} A^\top &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^\top \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= -\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= -A \end{aligned}$$

   and

$$\begin{aligned} B^\top &= \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}^\top \\ &= -\begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \\ &= -\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} \\ &= -B\ , \end{aligned}$$

   and $AB = BA$ since

$$\begin{aligned} AB &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= BA\ . \end{aligned}$$

However,

$$
\begin{aligned}
(AB)^\top &= \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}^\top \\
&= \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \\
&\neq \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \\
&= -(AB) \ .
\end{aligned}
$$

Thus, AB is not skew-symmetric. $\square$

(b) If $AB = BA$, then AB is symmetric.

It is true.

Proof. Suppose A and B are skew-symmetric and $AB = BA$. Then $A^\top = -A$ and $B^\top = -B$. We want to prove that $(AB)^\top = AB$. So, we have that

$$\begin{aligned} (AB)^\top &= (BA)^\top \\ &= A^\top B^\top \\ &= (-A)(-B) \\ &= AB\ . \end{aligned}$$

Thus, AB is symmetric. $\square$

Spans and Subspaces

Definition (Linear Combination)

A linear combination of vectors $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$ is a vector

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_k\vec{v}_k \in \mathbb{R}^n$$

where $a_1, \ldots, a_k \in \mathbb{R}$.

Definition (Span)

The span of vectors $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$ is the set of **all** linear combinations of $\vec{v}_1, \ldots, \vec{v}_k$. That is,

$$\text{span}\{\vec{v}_1, \ldots, \vec{v}_k\} = \{a_1\vec{v}_1 + \ldots + a_k\vec{v}_k \mid a_1, \ldots, a_k \in \mathbb{R}\} .$$

Definition (Subspace)

A set of vectors U in $\mathbb{R}^n$ is called a **subspace** of $\mathbb{R}^n$ if it satisfies the following properties:

1. The zero vector $\vec{0} \in U$.
2. If $\vec{v}, \vec{w} \in U$, then $\vec{u} + \vec{w} \in U$.
3. If $\vec{v} \in U$ and $c \in \mathbb{R}$, then $c\vec{v} \in U$.

(The Subspace Test is used to determine whether or not a set S is a subspace of $\mathbb{R}^n$. That is, whether $S \in U$.)

Theorem 1:

If $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$, then $V = \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$ is a subspace of $\mathbb{R}^n$.

Proof. Suppose $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$. We perform the subspace test to prove that V is a subspace of R^n.

1. $\vec{0} = 0 \cdot \vec{v}_1 + \ldots + 0 \cdot \vec{v}_k \in V$. In other words, the zero vector which can be expressed as a linear combination of the vectors $\vec{v}_1, \ldots, \vec{v}_k$ is in the span.

2. Suppose $\vec{v}, \vec{w} \in V$. Then
$$\vec{v} = a_1\vec{v}_1 + \ldots + a_k\vec{v}_k$$
and
$$\vec{w} = b_1\vec{v}_1 + \ldots + b_k\vec{v}_k$$
for some $a_1, \ldots, a_k \in \mathbb{R}$ and $b_1, \ldots, b_k \in \mathbb{R}$. Then
$$\begin{aligned} \vec{v} + \vec{w} &= (a_1\vec{v}_1 + \ldots + a_k\vec{v}_k) + (b_1\vec{v}_1 + \ldots + b_k\vec{v}_k) \\ &= (a_1 + b_1)\vec{v}_1 + \ldots + (a_k + b_k)\vec{v}_k \end{aligned}$$
where $a_1 + b_1, \ldots, a_k + b_k \in \mathbb{R}$. Hence,
$$\vec{v} + \vec{w} \in V = \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\} \ .$$

3. Suppose $\vec{v} \in V$ and $c \in \mathbb{R}$. Then
$$\vec{v} = a_1\vec{v}_1 + \ldots + a_k\vec{v}_k$$
for some $a_1, \ldots, a_k \in \mathbb{R}$. So, we have that
$$\begin{aligned} c\vec{v} &= c(a_1\vec{v}_1 + \ldots + a_k\vec{v}_k) \\ &= (ca_1)\vec{v}_1 + \ldots + (ca_k)\vec{v}_k \end{aligned}$$
where $ca_1, \ldots, ca_k \in \mathbb{R}$. Hence,
$$c\vec{v} \in V = \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\} \ .$$

Thus, by the Subspace Test, $V = \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$ is a subspace of $\mathbb{R}^n$. $\square$

Subspace Examples

1. Check if $S = \{\vec{0}\} \in U$.

Image Spaces

Definition (Image Space)

The image space of an $m \times n$ matrix A is defined as

$$\operatorname{Im}(A) = \{A\vec{x} \in \mathbb{R}^m \mid \vec{x} \in \mathbb{R}^n\} \ .$$

Theorem 2

Let $\vec{c}_1, \ldots, \vec{c}_n$ be the columns of an $m \times n$ matrix A. Then

$$\operatorname{Im}(A) = \operatorname{span}\{\vec{c}_1, \ldots, \vec{c}_n\} \ .$$

Proof. Suppose A is an $m \times n$ matrix with columns $\vec{c}_1, \ldots, \vec{c}_n$. We want to show that $\operatorname{Im}(A) = \operatorname{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$.

First, we prove that $\operatorname{Im}(A) \subseteq \operatorname{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$. Suppose $\vec{u} \in \operatorname{Im}(A)$. We show that $\vec{u} \in \operatorname{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$. Since $\vec{u} \in \operatorname{Im}(A)$, this means that $\vec{u} = A\vec{x} \in \mathbb{R}^m$ such that $x \in \mathbb{R}^n$ (note that $\vec{u} \in \mathbb{R}^m$ since A is $m \times n$ and $\vec{x}$ is $n \times 1$). In other words,

$$\begin{aligned}
\vec{u} &= A\vec{x} \\
&= \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{bmatrix} \\
&= \begin{bmatrix} a_{11}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \ldots + \begin{bmatrix} a_{1n}x_n \\ \vdots \\ a_{m1}x_n \end{bmatrix} \\
&= x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{m1} \\ \vdots \\ a_{mn} \end{bmatrix} \\
&= x_1\vec{c}_1 + \ldots + x_n\vec{c}_n \ ,
\end{aligned}$$

where $\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$ are the column vectors of the matrix A. Then this means that $\vec{u}$ is able to be expressed as a linear combination of the column vectors of the matrix A, and so $\vec{u} \in \text{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$. Now, since $\vec{u} \in \text{Im}(A)$ and $\vec{u} \in \text{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$, we have that $\text{Im}(A) \subseteq \text{span}\{\vec{c_1}, \ldots, \vec{c_n}\}$.

Now, we prove that $\text{span}\{\vec{c}_1, \ldots, \vec{c}_n\} \subseteq \text{Im}(A)$. Suppose $\vec{u} \in \text{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$. Then this means that $\vec{u}$ is a vector that can be expressed as a linear combination of all the vectors in the $\text{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$. That is,

$$\vec{u} = x_1\vec{c}_1 + \ldots + x_n\vec{c}_n \ ,$$

where $x_1, \ldots, x_n \in \mathbb{R}$. Since $\vec{c}_1, \ldots, \vec{c}_n$ are the columns of the $m \times n$ matrix A, this means that $\vec{u} = A\vec{x}$ where $\vec{x} \in \mathbb{R}^n$, and so $\vec{u} \in \text{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$. Now, since $\vec{u} \in \text{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$ and $\vec{u} \in \text{Im}(A)$, we have that $\text{span}\{\vec{c}_1, \ldots, \vec{c}_n\} \subseteq \text{Im}(A)$.

Thus, since $\text{Im}(A) \subseteq \text{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$ and $\text{span}\{\vec{c}_1, \ldots, \vec{c}_n\} \subseteq \text{Im}(A)$, we have shown that $\text{Im}(A) = \text{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$. $\square$

Corollary 3:

If A is an $m \times n$ matrix, then $\operatorname{Im}(A)$ is a subspace of $\mathbb{R}^m$.

Proof. Suppose A is an $m \times n$ matrix. We prove that $\operatorname{Im}(A)$ is a subspace of $\mathbb{R}^m$ by the Subspace Test.

By Theorem 2, $\operatorname{Im}(A) = \operatorname{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$, where $\vec{c}_1, \ldots, \vec{c}_n$ are the column vectors of the matrix A. Then by Theorem 1, since $\operatorname{Im}(A) = \operatorname{span}\{\vec{c}_1, \ldots, \vec{c}_n\}$ and $\vec{c}_1, \ldots, \vec{c}_n \in \mathbb{R}^m$, we get that $\operatorname{Im}(A)$ is a subspace of $\mathbb{R}^m$. $\square$

Null Spaces

Definition (Null Space)

The nullspace of an $m \times n$ matrix A is defined by

$$\text{null}(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\} .$$

Theorem 4

If $\{\vec{x}_1, \ldots, \vec{x}_n\}$ is the set of basic solutions to a homogeneous system $A\vec{x} = \vec{0}$, then $\text{null}(A) = \text{span}\{\vec{x}_1, \ldots, \vec{x}_n\}$.

Proof. Suppose $\{\vec{x}_1, \ldots, \vec{x}_n\}$ is the set of basic solutions to a homogeneous system $A\vec{x} = \vec{0}$. We show that $\text{null}(A) = \text{span}\{\vec{x}_1, \ldots, \vec{x}_n\}$.

First, we show that $\text{null}(A) \subseteq \text{span}\{\vec{x}_1, \ldots, \vec{x}_n\}$. Suppose $\vec{u} \in \text{null}(A)$. We want to prove that $\vec{u} \in \text{span}\{\vec{x}_1, \ldots, \vec{x}_n\}$; that is, $\vec{u} = c_1\vec{x}_1 + \ldots + c_n\vec{x}_n$ for $c_1, \ldots, c_n \in \mathbb{R}$. Since $\vec{u} \in \text{null}(A)$, this means that $\vec{u} \in \mathbb{R}^n$ such that $A\vec{u} = \vec{0}$. Then

$$\begin{aligned}
\vec{0} &= A\vec{u} \\
&= \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \\
&= \begin{bmatrix} a_{11}u_1 + \ldots + a_{1n}u_n \\ \vdots \\ a_{m1}u_1 + \ldots + a_{mn}u_n \end{bmatrix} \\
&= \begin{bmatrix} a_{11}u_1 \\ \vdots \\ a_{m1}u_1 \end{bmatrix} + \ldots + \begin{bmatrix} a_{1n}u_n \\ \vdots \\ a_{m1}u_n \end{bmatrix} \\
&= u_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \ldots + u_n \begin{bmatrix} a_{m1} \\ \vdots \\ a_{mn} \end{bmatrix}
\end{aligned}$$

Now, we show that $\text{span}\{\vec{x}_1, \ldots, \vec{x}_n\} \subseteq \text{null}(A)$. Suppose $\vec{u} \in \text{span}\{\vec{x}_1, \ldots, \vec{x}_n\}$. Then this means that $\vec{u} = c_1\vec{x}_1 + \ldots + c_n\vec{x}_n$ where $c_1, \ldots, c_n \in \mathbb{R}$. So, multiplying both sides by A, we get that

$$\begin{aligned} A\vec{u} &= A(c_1\vec{x}_1 + \ldots + c_n\vec{x}_n) \\ &= Ac_1\vec{x}_1 + \ldots + Ac_n\vec{x}_n \\ &= c_1(A\vec{x}_1) + \ldots + c_n(A\vec{x}_n) \\ &= c_1 \cdot 0 + \ldots + c_n \cdot 0 \\ &= 0 \ . \end{aligned}$$

□

Exercises

Let A be an $m \times n$ matrix.

1. Prove that if A is invertible, then $\text{null}(A) = \{\vec{0}\}$.

   Proof. Suppose A is invertible. Let $\vec{x} \in \text{null}(A)$. This means that $A\vec{x} = \vec{0}$. So it follows that (multiplying both sides of the equation by A^{-1})

$$\begin{aligned} A^{-1}A\vec{x} &= A^{-1}\vec{0} \\ I\vec{x} &= \vec{0} \\ \vec{x} &= \vec{0}\ . \end{aligned}$$

   Now, since $\vec{x} = \vec{0}$, this means that $\text{null}(A) \subseteq \{\vec{0}\}$. Indeed,

$$\text{null}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\} = \{\vec{0}\}$$

   So, we have that $\{\vec{0}\} \subseteq \text{null}(A) = \{\vec{0}\}$. Thus, since $\text{null}(A) \subseteq \{\vec{0}\}$ and $\{\vec{0}\} \subseteq \text{null}(A)$, $\text{null}(A) = \{\vec{0}\}$. □

2. Prove that if $\text{null}(A) \neq \{0\}$, then A is not invertible.

   Proof. This statement is the contrapositive of the statement in (1). The contrapositive is logically equivalent (has the same truth value) to the original statement. Hence, we are victorious. □

3. Is the converse of (1) true? Prove your answer.

The converse is: If $\text{null}(A) = \{\vec{0}\}$, then A is invertible.

Solution: This statement is false, so we prove the negation is true. The negation is: There exists a matrix A such that $\text{null}(A) = \{0\}$, but A is not invertible.

Proof. Choose $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Let $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Then $\text{null}(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$; that is, the null space of A is the set of solutions to the homogeneous system $A\vec{x} = \vec{0}$. Now,

$$\left(\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 1 & 0 \end{array}\right)$$

$$\xrightarrow{-1R1+R2} \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & -1 & 0 \end{array}\right)$$

$$\xrightarrow{-1R2} \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 1 & 0 \end{array}\right)$$

$$\xrightarrow{-2R2+R1} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right)$$

and so we get the trivial solution to this system of equations

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \vec{0}\ .$$

Then this means that $\text{null}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\} = \{0\}$. However,

$$\det(A) = (1 \cdot 2) - (2 \cdot 1) = 2 - 2 = 0\ .$$

Therefore, A is not invertible. $\square$

Linear Independence

Definition (Linear Independence)

A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_k\} \in \mathbb{R}^n$ is called linearly independent in $\mathbb{R}^n$ if the equation

$$c_1\vec{v}_1 + \ldots + c_k\vec{v}_k = \vec{0}$$

where $c_1, \ldots, c_k \in \mathbb{R}$, has only the **trivial solution**

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

Otherwise, the set is called linearly dependent.

Independence Test

To verify that a set $\{\vec{x_1}, \ldots, \vec{x_k}\}$ is linearly independent, proceed as follows:

1. Set a linear combination equal to zero: $t_1\vec{x_1} + \ldots + t_k\vec{v_k} = \vec{0}$.
2. Show that $t_i = 0$ for each i (that is, the linear combination is trivial).

Of course, if at least one of the $t_i \neq 0$, then the vectors are not linearly independent (i.e. they are **linearly dependent**).

1. $\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\2\end{pmatrix}\right\}$ is linearly dependent. This is because $c_1 = 2$ and $c_2 = -1$ is a non-trivial solution to $c_1\begin{pmatrix}1\\1\end{pmatrix} + c_2\begin{pmatrix}2\\2\end{pmatrix} = \vec{0}$.

2. $\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\2\end{pmatrix}\right\}$ is linearly independent. This is because $c_1 = c_2 = 0$ is the trivial solution to $c_1\begin{pmatrix}1\\1\\1\end{pmatrix} + c_2\begin{pmatrix}1\\1\\2\end{pmatrix} = \vec{0}$.

Example:

If a set of vectors S in $\mathbb{R}^n$ contains the zero vector, then S is linearly dependent.

Proof. Suppose $S = \{\vec{0}, \vec{v}_1, \ldots, \vec{v}_k\}$ is a set of vectors that contains the zero vector. We want to prove that S is linearly dependent. Then

$$(-69) \cdot \vec{0} + 0 \cdot \vec{v}_1 + \ldots 0 \cdot \vec{v}_k = \vec{0} \ .$$

So, $c_0 = -1$ and $c_1 = \ldots = c_k = 0$ is a non-trivial solution. So, we have shown that not all coefficients need to be zero for this linear combination to equal $\vec{0}$. Hence, S is linearly dependent. $\square$

Theorem 5

If $\{\vec{x}_1, \ldots, \vec{x}_k\}$ is a linearly independent subset of $\mathbb{R}^n$, then every vector $\vec{v} \in \text{span}\{\vec{x}_1, \ldots, \vec{x}_k\}$ can be written as a linear combination of $\vec{x}_1, \ldots, \vec{x}_k$ in exactly one way (i.e. the linear combination is unique).

Proof. Suppose $\{\vec{x}_1, \ldots, \vec{x}_k\}$ is a linearly independent subset of $\mathbb{R}^n$. We want to prove that any vector $\vec{v} \in \text{span}\{\vec{x}_1, \ldots, \vec{x}_k\}$ can be written as a unique linear combination of $\vec{x}_1, \ldots, \vec{x}_n$. Since $\{\vec{x}_1, \ldots, \vec{x}_n\}$ is linearly independent, this means that any vector $\vec{v} \in \text{span}\{\vec{x}_1, \ldots, \vec{x}_k\}$ can be written as

$$\vec{v} = c_1\vec{x}_1 + \ldots + c_k\vec{x}_k = 0 \cdot \vec{x}_1 + \ldots + 0 \cdot \vec{x}_k = \vec{0} \ ,$$

where $c_1 = \ldots = c_k = 0$ is the trivial solution. Now, let $\vec{u}, \vec{w} \in \text{span}\{\vec{x}_1, \ldots, \vec{x}_n\}$. Then it follows that

$$\vec{u} = a_1\vec{x}_1 + \ldots + a_k\vec{x}_k = \vec{0}$$

and

$$\vec{w} = b_1\vec{x}_1 + \ldots + b_k\vec{x}_k = \vec{0} \ .$$

So, we have that

$$\begin{aligned} \vec{u} - \vec{w} &= \vec{0} - \vec{0} \\ \vec{u} - \vec{w} &= \vec{0} \\ \vec{u} &= \vec{0} + \vec{w} \\ \vec{u} &= \vec{w} \\ (a_1\vec{x}_1 + \ldots + a_k\vec{x}_k) &= (b_1\vec{x}_1 + \ldots + b_k\vec{x}_k) \ , \end{aligned}$$

which means that $a_i = b_i$ for all $1 \leq i \leq k$. Thus, every vector $\vec{v} \in \text{span}\{\vec{x}_1, \ldots, \vec{x}_k\}$ can be written as a linear combination of $\vec{x}_1, \ldots, \vec{x}_k$ in exactly one way. $\square$

Example:

Let $\vec{v}, \vec{w} \in \mathbb{R}^n$ be non-zero vectors. Prove that $\{\vec{v}, \vec{w}\}$ is linearly dependent if and only if $\vec{v}$ and $\vec{w}$ are parallel (they are scalar multiples of each other).

Proof. ($\Longrightarrow$) Suppose $\{\vec{v}, \vec{w}\}$ are linearly dependent. Then $a\vec{v} + b\vec{w} = \vec{0}$, where $a, b \in \mathbb{R}$, and at least one of a and b is non-zero. Now, since $\vec{v}$ and $\vec{w}$ are both non-zero vectors, we conclude that both a and b are not zero. Then

$$\vec{v} = -\frac{b}{a}\vec{w}$$

which means that $\vec{v}$ is a scalar multiple of $\vec{w}$. Hence, $\vec{v}$ and $\vec{w}$ are parallel.

($\Longleftarrow$) Suppose $\vec{v}$ and $\vec{w}$ are parallel. This means that $\vec{v}$ and $\vec{w}$ are scalar multiples of each other. Then we get that

$$\vec{v} = k\vec{w}$$

where $k \in \mathbb{R}$. Then subtracting $k\vec{w}$ from both sides gives $\vec{v} - k\vec{w} = \vec{0}$, and so it follows that

$$\vec{0} = \vec{v} - k\vec{w} = 1\vec{v} + (-k)\vec{w}$$

and so we have a non-trivial solution to the homogeneous equation

$$x_1\vec{v} + x_2\vec{w} = \vec{0}\ .$$

That is, at least one of the coefficients of this linear combination is non-zero (in this case, both the coefficients 1 and $-k$ are non-zero). Hence, $\{\vec{v}, \vec{w}\}$ is linearly dependent. $\square$

Basis and Dimension

Definition (Basis)

Let U be a subspace of $\mathbb{R}^n$. A set of vectors $\{\vec{v}_1, \dots, \vec{v}_k\}$ is called a **basis** for U if

1. $\{\vec{v}_1, \dots, \vec{v}_k\}$ spans U.
2. $\{\vec{v}_1, \dots, \vec{v}_k\}$ is linearly independent in U.

1. $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is a basis for $\mathbb{R}^2$.
2. $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ is linearly dependent. For example, observe that the third vector is simply a linear combination of the first two vectors. However, it does indeed span $\mathbb{R}^2$. It's just that the third vector doesn't add anything to the span. Since this set of vectors fails property (2), it is therefore not a basis.
3. $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ doesn't span $\mathbb{R}^3$. This is because there exists some vector in $\mathbb{R}^3$ that cannot be written as a linear combination of the these two vectors. Since it doesn't span $\mathbb{R}^3$, it fails property (1), and is therefore not a basis.

Question:

Can there be more than one basis for a vector space?

Answer: Yes! For example, $\left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$ is also a basis for $\mathbb{R}^2$.

Theorem 6 (Uniqueness of Basis Representation)

If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is a basis for a subspace U, then every vector $\vec{w} \in U$ can be written in the form

$$\vec{w} = c_1\vec{v}_1 + \ldots + c_k\vec{v}_k,$$

where $c_1, \ldots, c_k \in \mathbb{R}$, in **exactly one** way

Proof. Let $\vec{w} \in U$. Since B is a basis for U, we have that $U = \text{span}(B) = \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$, meaning that every vector in the subspace U can be written as some linear combination of the basis vectors in B. Then since $w \in U = \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$, there exists $b_1, \ldots, b_k \in \mathbb{R}$ such that

$$\vec{w} = b_1\vec{v}_1 + \ldots + b_k\vec{v}_k\ .$$

Now suppose that there exists $d_1, \ldots, d_k \in \mathbb{R}$ such that

$$\vec{w} = d_1\vec{v}_1 + \ldots + d_k\vec{v}_k\ .$$

Then

$$\vec{w} = b_1\vec{v}_1 + \ldots + b_k\vec{v}_k = d_1\vec{v}_1 + \ldots + d_k\vec{v}_k\ .$$

From this we get that

$$\begin{aligned}
\vec{0} &= \vec{w} - \vec{w} \\
&= (b_1\vec{v}_1 + \ldots + b_k\vec{v}_k) - (d_1\vec{v}_1 + \ldots + d_k\vec{v}_k) \\
&= b_1\vec{v}_1 + \ldots + b_k\vec{v}_k - d_1\vec{v}_1 - \ldots - d_k\vec{v}_k \\
&= (b_1 - d_1)\vec{v}_1 + \ldots + (b_k - d_k)\vec{v}_k\ .
\end{aligned}$$

Now, since B is linearly independent (we know this because a basis is linearly independent from the definition of basis), we get that $(b_1 - d_1), \ldots, (b_k - d_k) = 0$, which means that $b_i = d_i$ for $i = 1, \ldots, k$. Therefore, the two linear combinations

$$\vec{w} = b_1\vec{v}_1 + \ldots + b_k\vec{v}_k$$

and

$$\vec{w} = d_1\vec{v}_1 + \ldots + d_k\vec{v}_k$$

are exactly the same. $\square$

Observation (Theorem 7. Proof Later)

1. A basis is a **minimal** spanning set.
2. A basis is a **maximal** linearly independent set.

Let $\{\vec{v}_1, \ldots, \vec{v}_k\}$ be a basis for U.

- Observation (1): If a set in U has less than k elements, then it cannot span U.
- If a set in U has more than k elements, then it is linearly dependent.

Corollary 8 (Invariance Theorem)

All bases of a subspace of $\mathbb{R}^n$ have **the same** number of elements.

1. $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is the standard basis for $\mathbb{R}^2$. This set has 2 elements.
2. $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ is also a basis. Notice how this set also has two elements.

Definition (Dimension)

The dimension of a subspace U of $\mathbb{R}^n$ is the number of elements of a basis for U. And we define $\dim(\{\vec{0}\}) = 0$.

Row Space and Column Space

Let A be an $m \times n$ matrix.

Definition (Row Space)

The row space of A, denoted $\mathrm{Row}(A)$, is the subspace of $\mathbb{R}^n$ spanned by the rows of A.

Definition (Column Space)

The column space of A, denoted $\mathrm{Col}(A)$, is the subspace of $\mathbb{R}^n$ spanned by the columns of A.

1. Let $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 5 & 2 \end{pmatrix}$. Then

$$\mathrm{Row}(A) = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 2 \end{pmatrix} \right\}$$

and

$$\mathrm{Col}(A) = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\} .$$

Notes:

- $\mathrm{Row}(A)$ is not a basis. Although it is linearly independent, it does not span.
- $\mathrm{Col}(A)$ is not a basis. Although it does span, it is linearly dependent.

2. Let $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$ and $R := \mathrm{RREF}(A) = \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$. Then

$$\begin{aligned}
\mathrm{Col}(A) &= \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \end{pmatrix} \right\} = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\} \\
\mathrm{Col}(R) &= \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix} \right\} = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \\
\mathrm{Row}(A) &= \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \end{pmatrix} \right\} = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\} \\
\mathrm{Row}(R) &= \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}
\end{aligned}$$

Row operations DO NOT affect the row space!

> **Theorem 9**
>
> Elementary row operations do not change the row space of a matrix.

1. Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 1 & 0 & 5 \end{pmatrix}$ and $R := \mathrm{RREF}(A) = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}$. Then a basis for $\mathrm{Row}(A)$ is

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix} \right\},$$

a basis for $\mathrm{Col}(A)$ is

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right\},$$

and a basis for $\mathrm{null}(A)$ is

$$\left\{ \begin{pmatrix} -5 \\ 3 \\ 1 \end{pmatrix} \right\}.$$

Theorem 10

Let B be a matrix obtained by performing elementary row operations on A. Then a given set of column vectors of A form a basis for $\mathrm{Col}(A)$ if and only if the corresponding column vectors of B form a basis for $\mathrm{Col}(B)$.

1. Let $A = \begin{pmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{pmatrix}$ and $R := \mathbf{REF}(A) = \begin{pmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

   We know that $A\vec{x} = \vec{0} \iff R\vec{X} = \vec{0}$. So, a basis for $\mathrm{Row}(A)$ is

$$\left\{ \begin{pmatrix} 1 \\ -3 \\ 4 \\ -2 \\ 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 3 \\ -2 \\ -6 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 5 \end{pmatrix} \right\},$$

   a basis for $\mathrm{Col}(A)$ (observe the positions of the pivot entries of R) is

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ 9 \\ 9 \\ -4 \end{pmatrix}, \begin{pmatrix} 5 \\ 8 \\ 9 \\ -5 \end{pmatrix} \right\},$$

   and a basis for $\mathrm{null}(A)$ is

$$\left\{ \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 14 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 37 \\ 0 \\ -4 \\ 0 \\ -5 \\ 1 \end{pmatrix} \right\}.$$

Rank

Definition (Rank)

The rank of a matrix A, denoted $\text{rank}(A)$, is the number of leading 1's in the RREF of A.

1. Let $A = \begin{pmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{pmatrix}$.

   Then we know from the previous example that the REF of A

$$R = \begin{pmatrix} \mathbf{1} & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & \mathbf{1} & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

   has 3 leading 1's. Hence, $\text{rank}(A) = 3$.

Theorem 11

If A is a matrix, then $\text{rank}(A) = \dim(\text{Row}(A)) = \dim(\text{Col}(A))$.

Proof. Let R be the RREF of A and suppose $\text{rank}(A) = r$. Then R has r nonzero rows and r leading 1's. By Theorem 9, $\text{rank}(A) = \dim(\text{Row}(A))$, and by Theorem 10, $\text{rank}(A) = \dim(\text{Col}(A))$. $\square$

Corollary 12

If A is an $m \times n$ matrix, then $\text{rank}(A) = \text{rank}(A^\top)$.

Proof. (Hint: $\text{Row}(A^\top) = \text{Col}(A)$). Suppose A is an $m \times n$ matrix. Then

$$\begin{aligned}\text{rank}(A) &= \dim(\text{Col}(A)) \\ &= \dim(\text{Row}(A^\mathsf{T})) \\ &= \dim(\text{Col}(A^\mathsf{T})) \\ &= \text{rank}(A^\mathsf{T})\end{aligned}$$

□

Definition (Nullity)

The nullity of a matrix A, denoted $\text{nullity}(A)$, is $\dim(\text{null}(A))$.

1. Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 1 & 0 & 5 \end{pmatrix}$. A basis for $\text{null}(A)$ is $\left\{ \begin{pmatrix} -5 \\ 3 \\ 1 \end{pmatrix} \right\}$.

   Hence, $\text{nullity}(A) = 1$ (the number of elements in the basis of $\text{null}(A)$).

Rank-Nullity Theorem

If a matrix A has n columns, then $\text{rank}(A) + \text{nullity}(A) = n$.

Some Useful Equivalent Statements

1. A is invertible.
2. $A\vec{x} = \vec{0}$.
3. The RREF of A is I_n ($n \times n$ identity matrix).
4. A is a product of elementary matrices.
5. $A\vec{x} = \vec{b}$ is consistent $\forall \vec{b} \in \mathbb{R}^n$.
6. $A\vec{x} = \vec{b}$ has one solution $\forall \vec{b} \in \mathbb{R}^n$.
7. $\det(A) \neq 0$.
8. $\text{rank}(A) = n$.
9. $\text{Im}(A) = \mathbb{R}^n$.
10. $\vec{x} \mapsto A\vec{x}$ is one-to-one (injective).
11. The columns of A are linearly independent.
12. The rows of A are linearly independent.
13. The columns of A span $\mathbb{R}^n$.
14. The rows of A span $\mathbb{R}^n$.
15. The columns of A form a basis for $\mathbb{R}^n$.
16. The rows of A form a basis for $\mathbb{R}^n$.

Dot Product

Definition (Norm)

The norm (or length) of a vector $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ is

$$||\vec{x}|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \ .$$

Fact: $||\vec{x}|| = 0$ if and only if $\vec{x} = \vec{0}$.

Definition (Unit Vector)

A vector with norm 1 is called a unit vector.

Remark: For any non-zero vector $\vec{v} \in \mathbb{R}^n$, $\frac{1}{||\vec{v}||}\vec{v}$ is a unit vector (parallel to $\vec{v}$).

Definition (Dot Product)

The dot product of two vectors $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ is

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n \ .$$

Properties of the Dot Product

For all $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ and for all $k \in \mathbb{R}$,

1. $\vec{x} \cdot \vec{x} = x_1^2 + x_2^2 + \ldots + x_n^2 = ||\vec{x}||^2$
2. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$ (Commutative)
3. $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ (Distributive)
4. $(k\vec{x}) \cdot \vec{y} = k(\vec{x} \cdot \vec{y}) = \vec{x} \cdot (k\vec{y})$
5. $\vec{0} \cdot \vec{x} = 0$

Definition (Distance)

The distance between two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ is $\mathrm{d}(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$.

Remark: If θ is the angle between two vectors $\vec{v}$ and $\vec{w}$, and $0 \leq \theta \leq \pi$, then

$$\cos\theta = \frac{\vec{v} \cdot \vec{w}}{||\vec{v}||\ ||\vec{w}||}$$

or

$$\vec{v} \cdot \vec{w} = ||\vec{v}||\ ||\vec{w}||\ \cos\theta\ .$$

Orthogonality

Definition (Orthogonal)

A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is called an orthogonal set if each $\vec{v}_i \neq \vec{0}$ and each pair of vectors are orthogonal (i.e. $\vec{v}_i \cdot \vec{v}_j = 0 \ \forall 1 \leq i \neq j \leq k$).

Theorem 13

If $S = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal set, then S is linearly independent.

Proof. Suppose S is orthogonal. Suppose $c_1\vec{v}_1 + \ldots + c_k\vec{v}_k = \vec{0}$ for some $c_i \in \mathbb{R}$ $(1 \leq i \leq k)$. We want to show that $c_1 = \ldots = c_k = 0$. Now, for each $i \in \{1, \ldots, k\}$ we have that

$$\begin{aligned}
0 &= \vec{0} \cdot \vec{v}_i \\
&= (c_1\vec{v}_1 + \ldots + c_k\vec{v}_k) \cdot \vec{v}_i \\
&= c_1\vec{v}_1 \cdot \vec{v}_i + \ldots + c_k\vec{v}_k \cdot \vec{v}_i \\
&= c_1(\vec{v}_1 \cdot \vec{v}_i) + \ldots + c_k(\vec{v}_k \cdot \vec{v}_i) \\
&= c_i(\vec{v}_i \cdot \vec{v}_i) \\
&= c_i \ ||\vec{v}_i||^2 \ .
\end{aligned}$$

Now, $||\vec{v}_i|| \neq 0$ because S is an orthogonal set that cannot contain the zero vector (this is from the definition of orthogonality), and $||\vec{v}_i|| = 0$ only when $\vec{v}_i = \vec{0}$. Then this implies that $c_i = 0$, since $c_i||\vec{v}_i||^2 = 0$. So, we have shown that each $c_i = 0$. Hence, S is linearly independent. □

Orthonormality

Definition (Orthonormal)

A set of vectors $S = \{\vec{v}_1, \ldots, \vec{v}_k\} \in \mathbb{R}^n$ is called orthonormal if it is an orthogonal set of **unit vectors**. That is, $\vec{v}_i \cdot \vec{v}_j = 0$ for all $i \neq j \in \{1, \ldots, k\}$ and $||\vec{v}_i|| = 1$ for all $i \in \{1, \ldots, k\}$.

Normalizing an Orthogonal Set

If $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is orthogonal, then

$$\left\{ \frac{1}{||\vec{v}_1||}\vec{v}_1,\ \frac{1}{||\vec{v}_2||}\vec{v}_2,\ \ldots,\ \frac{1}{||\vec{v}_k||}\vec{v}_k \right\}$$

is orthonormal.

Fourier Expansion

Question: Given an arbitrary basis B in $\mathbb{R}^n$, how can we write an arbitrary vector $\vec{v} \in \mathbb{R}^n$ as a linear combination of vectors in B?

Answer: When B is orthogonal, there is an easy way!

Fourier Expansion Theorem

Let $B = \{\vec{f}_1, \ldots, \vec{f}_k\}$ be an orthogonal basis for a subspace $U \subseteq \mathbb{R}^n$. Then for any $\vec{u} \in U$, we can write

$$\vec{u} = \left(\frac{\vec{u} \cdot \vec{f}_1}{||\vec{f}_1||^2}\right) \vec{f}_1 + \ldots + \left(\frac{\vec{u} \cdot \vec{f}_k}{||\vec{f}_k||^2}\right) \vec{f}_k \ .$$

Note: Each coefficient in front of each vector f_i is called a **Fourier Coefficient**.

Proof. Let $\vec{u} \in U$. We know that

$$\vec{u} = c_1 \vec{f}_1 + \ldots + c_k \vec{f}_k$$

for some $c_1, \ldots, c_k \in \mathbb{R}$, since $B = \{\vec{f}_1, \ldots, \vec{f}_k\}$ is a basis for U. Now, since B is orthogonal, we get that for each $1 \leq i \leq k$,

$$\begin{aligned} \vec{u} \cdot \vec{f}_i &= (c_1 \vec{f}_1 + \ldots + c_k \vec{f}_k) \cdot \vec{f}_i \\ &= c_i (\vec{f}_i \cdot \vec{f}_i) \\ &= c_i \ ||\vec{f}_i||^2 \ . \end{aligned}$$

Then, since we got that

$$\vec{u} \cdot \vec{f}_i = c_i \ ||\vec{f}_i||^2 \ ,$$

we can say that each Fourier coefficient c_i can be expressed as

$$c_i = \frac{\vec{u} \cdot \vec{f}_i}{||\vec{f}_i||^2} \ .$$

Hence,

$$\begin{aligned} \vec{u} &= c_1 \vec{f}_1 + \ldots + c_k \vec{f}_k \\ &= \left(\frac{\vec{u} \cdot \vec{f}_1}{||\vec{f}_1||^2}\right) \vec{f}_1 + \ldots + \left(\frac{\vec{u} \cdot \vec{f}_k}{||\vec{f}_k||^2}\right) \vec{f}_k \ . \end{aligned}$$

$\square$

Other Theorems

Cauchy-Schwarz Inequality

If $\vec{v}, \vec{w} \in \mathbb{R}^n$, then $|\vec{v} \cdot \vec{w}| \leq ||\vec{v}|| \; ||\vec{w}||$. (With equality when $\vec{v} = t\vec{w}$ for $t \in \mathbb{R}$.)

Triangle Inequality

If $\vec{v}, \vec{w} \in \mathbb{R}^n$, then $||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||$.

Pythagorean Theorem

If $\{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set, then

$$||\vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_k||^2 = ||\vec{x}_1||^2 + ||\vec{x}_2||^2 + \ldots + ||\vec{x}_k||^2$$

Similarity and Diagonalization

Eigenvalues and Eigenvectors

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

Definition (Eigenvector and Eigenvalue)

A non-zero vector $v \in \mathbb{R}^n$ is called an eigenvector of A if $A\vec{v} = \lambda\vec{v}$ for some scalar $\lambda \in \mathbb{R}$. The scalar λ is called an eigenvalue of A.

How to find eigenvalues and eigenvectors:

(1) Find the roots of the characteristic polynomial

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

The roots are the eigenvalues.

(2) Find solutions to the system $(A-\lambda I)\vec{v} = \vec{0}$ or $(\lambda I - A)\vec{v} = \vec{0}$. The solutions give the eigenvectors.

Note that

$$\begin{aligned} A\vec{v} = \lambda\vec{v} &\iff A\vec{v} - \lambda\vec{v} = \vec{0} \\ &\iff A\vec{v} - \lambda I\vec{v} = \vec{0} \\ &\iff (A - \lambda I\vec{v}) = \vec{0} \end{aligned}$$

and

$$\begin{aligned} A\vec{v} = \lambda\vec{v} &\iff \lambda\vec{v} - A\vec{v} = \vec{0} \\ &\iff \lambda I\vec{v} - A\vec{v} = \vec{0} \\ &\iff (\lambda I - A)\vec{v} = \vec{0}\ . \end{aligned}$$

Similarity

Definition (Similar)

Let A and B be $n \times n$ similar matrices. A and B are similar if there exists an invertible matrix $P \in M_{n\times n}(\mathbb{R})$ such that $B = P^{-1}AP$. In this case, $A \sim B$.

1. Let $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. We can find the eigenvalues of A via the characteristic polynomial. So,

$$\begin{aligned} C_A(\lambda) &= \det(\lambda I_2 - A) \\ &= \det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} \lambda - 1 & -1 \\ 2 & \lambda - 4 \end{bmatrix}\right) \\ &= (\lambda-1)(\lambda-4) - (-1)(2) \\ &= \lambda^2 - 4\lambda - \lambda + 4 + 2 \\ &= \lambda^2 - 5\lambda + 6 \\ &= (\lambda-3)(\lambda-2)\ . \end{aligned}$$

Here we get that $\lambda_1 = 3$ and $\lambda_2 = 2$ are eigenvalues of A. Now, we find the eigenvectors associated with these eigenvalues.

For $\lambda_1 = 3$, we solve the system $(3I - A)\vec{x} = \vec{0}$. So,

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

and

$$\left(\begin{array}{cc|c} 2 & -1 & 0 \\ 2 & -1 & 0 \end{array}\right) \to \ldots \to \left(\begin{array}{cc|c} 1 & -1/2 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

and so

$$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{2}t \\ t \end{bmatrix} = t\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

Then $\vec{x}_2 = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ is an eigenvector associated with the eigenvalue $\lambda_1 = 3$.

For $\lambda_2 = 2$, we solve the system $(2I - A)\vec{x} = \vec{0}$. So,

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

and

$$\left(\begin{array}{cc|c} 1 & -1 & 0 \\ 2 & -2 & 0 \end{array}\right) \xrightarrow{-2R1+R2} \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

and so

$$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \ .$$

Then $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a eigenvector associated with the eigenvalue $\lambda_1 = 2$.

Now,

$$P = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 1 & 1 \end{bmatrix}$$

and so

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

Indeed, we can check that $P^{-1}AP = B$ and $PBP^{-1} = A$:

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & -2 \\ -6 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \\ &= B \ . \end{aligned}$$

and

$$\begin{aligned} PBP^{-1} &= \begin{bmatrix} 1 & 1/2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 3/2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \end{aligned}$$

Hence, $A \sim B$.

2. Does $A \sim B$ imply $B \sim A$? In other words, is it true that if A is similar to B, then B is similar to A? Yes! We can even prove it.

   Proof. Suppose $A \sim B$. Then there exists an invertible matrix $P \in M_{n \times n}(\mathbb{R})$ such that $B = P^{-1}AP$. We want to show that $B \sim A$. So,

$$\begin{aligned} B &= P^{-1}AP \\ PB &= P(P^{-1}AP) \\ PB &= AP \\ PBP^{-1} &= (AP)P^{-1} \\ PBP^{-1} &= A \end{aligned}$$

   Hence, $A = PBP^{-1}$, and so $B \sim A$. $\square$

Trace

Definiton (Trace)

The trace of an $n \times n$ matrix A is the sum of the diagonal entries of A. That is,

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn} \ .$$

It is useful to refer to this $n \times n$ matrix A:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

Lemma (See Lemma 5.5.1 in Textbook)

For any $A, B \in M_{n\times n}(\mathbb{R})$, we have

$$\mathrm{tr}(AB) = \mathrm{tr}(BA) \ .$$

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

$$\begin{aligned} AB &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + \ldots + a_{1n}b_{n1} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \end{aligned}$$

$\square$

Theorem 14

If A and B are $n \times n$ matrices that are similar, then

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

(2) $\operatorname{rank}(A) = \operatorname{rank}(B)$

(3) $C_A(\lambda) = C_B(\lambda)$

(4) $\operatorname{tr}(A) = \operatorname{tr}(B)$

(5) A and B have the same eigenvalues

Let's prove these five properties.

Proof. Suppose $A, B \in M_{n\times n}(\mathbb{R})$ that are similar. Then there exists an invertible matrix $P \in M_{n\times n}(\mathbb{R})$ such that $B = P^{-1}AP$. Now,

(1) We show that $\det(B) = \det(A)$.

$$\begin{aligned}
\det(B) &= \det(P^{-1)AP} \\
&= \det(P^{-1})\det(AP) \\
&= \det(P^{-1})\det(A)\det(P) \\
&= \frac{1}{\det(P)}\det(A)\det(P) \\
&= \det(A)\ .
\end{aligned}$$

(2) Show that $\operatorname{rank}(A) = \operatorname{rank}(B)$.

(3) Show that $C_B(\lambda) = C_A(\lambda)$.

$$\begin{aligned}
C_B(\lambda) &= \det(\lambda I - B) \\
&= \det(\lambda I - P^{-1}AP) \\
&= \det(\lambda P^{-1}PI - P^{-1}AP) \\
&= \det(P^{-1}(\lambda PI - AP)) \\
&= \det(P^{-1}(\lambda I - A)P) \\
&= \det(P^{-1})\det(\lambda I - A)\det(P) \\
&= \frac{1}{\det(P)}\det(\lambda I - A)\det(P) \\
&= \det(\lambda I - A) \\
&= C_A(\lambda)\ .
\end{aligned}$$

(4) Show that $\operatorname{tr}(B) = \operatorname{tr}(A)$. Let

$$P^{-1}AP = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \dots & \vdots \\ c_{n1} & \dots & c_{nn} \end{bmatrix}\ .$$

Then we have that

$$\begin{aligned}
\operatorname{tr}(B) &= \operatorname{tr}(P^{-1}AP) \\
&= c_{11} + c_{22} + \ldots + c_{nn}
\end{aligned}$$

(5) Show that A and B have the same eigenvalues.

□

Diagonalization

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

Definition (Diagonalizable)

A is diagonalizable if A is similar to a diagonal matrix

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

Note from earlier the matrices $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$. Well, A is diagonalizable because $A \sim B$.

Recall: An $n \times n$ matrix A is diagonalizable if and only if A has n eigenvectors $\vec{x}_1, \dots, \vec{x}_n$ such that $P = \begin{bmatrix} \vec{x}_1 & \dots & \vec{x}_n \end{bmatrix}$ is invertible.

Question: How can we show that $P = \begin{bmatrix} \vec{x}_1 & \dots & \vec{x}_n \end{bmatrix}$ is invertible without computing P^{-1}?

Answer: Think about the rank! This is where Theorem 15 comes in.

Theorem 15

$A \in M_{n\times n}(\mathbb{R})$ is diagonalizable if and only if A has n linearly independent eigenvectors.

1. Is $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$ diagonalizable? Yes.

   We can find the eigenvalues of A via the characteristic polynomial $C_A(\lambda) =$ $\det(\lambda I_2 -$

2. Is $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ diagonalizable?

How to Diagonalize an $\mathbf{n \times n}$ matrix A

(1) Find n linearly independent eigenvectors, say $\vec{x}_1, \ldots, \vec{x}_n$.

(2) Let $P = \begin{bmatrix} \vec{x}_1 & \ldots & \vec{x}_n \end{bmatrix}$.

(3) Then $P^{-1}AP = D = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

> **Theorem 16** (See Theorem 5,5,4 and 5.5.5 in Textbook)
>
> If $A \in M_{n \times n}(\mathbb{R})$ has n distinct eigenvalues, then A is diagonalizable.

Note: Eigenvectors corresponding to distinct eigenvalues are linearly independent.

The converse is false.

Eigenspaces

Definition (Eigenspace)

The eigenspace of $A \in M_{n\times n}(\mathbb{R})$ corresponding to an eigenvalue λ is

$$\mathrm{E}_\lambda(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \lambda\vec{x}\} \ .$$

In other words, the eigenspace is the collection of eigenvectors associated with a particular eigenvalue.

Observation: Eigenspaces corresponding to eigenvalues of a matrix $A \in M_{n\times n}(\mathbb{R})$ are subspaces of $\mathbb{R}^n$.

Definition (Multiplicity)

The multiplicity of an eigenvalue λ_0 is the number of times it appears as a root of $C_A(\lambda)$.

Theorem 17

$A \in M_{n\times n}(\mathbb{R})$ is diagonalizable if and only if $\dim(\mathrm{E}_\lambda(A)) = \mathrm{mult}_A(\lambda)$ for each λ.