MATH254: Linear Algebra
Lecture 1

Moira MacNeil

January 7, 2025

Today

1. Subspaces of $\mathbb{R}^n$
2. Spanning Sets

Vectors

- Let $\mathbb{R}$ denote the set of all real numbers

Vectors

$\mathbb{R}^2$ $\begin{bmatrix} a \\ b \end{bmatrix}$

- Let $\mathbb{R}$ denote the set of all real numbers

$\mathbb{R}^3$ $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$

- $\mathbb{R}^n$ denotes the set of all ordered n-tuples of real numbers.

$\mathbb{R}^n$ $\left.\begin{bmatrix} a \\ b \\ \vdots \\ \end{bmatrix}\right\}$ n entries

Vectors

- Let $\mathbb{R}$ denote the set of all real numbers
- $\mathbb{R}^n$ denotes the set of all ordered n-tuples of real numbers.
- **Vectors** (or n-vectors) are denoted as rows

$$(r_1, r_2, \ldots, r_n)$$

or columns

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$$

Subspace of $\mathbb{R}^n$

$\mathbf{0} = (0, 0, \cdots, 0)$

Subspace of $\mathbb{R}^n$

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

S1. The zero vector $\mathbf{0} \in U$.

S2. If $\mathbf{x} \in U$ and $\mathbf{y} \in U$, then $\mathbf{x} + \mathbf{y} \in U$.

S3. If $\mathbf{x} \in U$, then $a\mathbf{x} \in U$ for every real number a.

Subspace of $\mathbb{R}^n$

Subspace of $\mathbb{R}^n$

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

S1. The zero vector $\mathbf{0} \in U$.

S2. If $\mathbf{x} \in U$ and $\mathbf{y} \in U$, then $\mathbf{x} + \mathbf{y} \in U$.

S3. If $\mathbf{x} \in U$, then $a\mathbf{x} \in U$ for every real number a.

The subset U is

- **closed under addition** if S2 holds
- **closed under scalar multiplication** if S3 holds

Subspaces

- $\mathbb{R}^n$ is a subspace of itself

Subspaces

- $\mathbb{R}^n$ is a subspace of itself
- $U = \{\mathbf{0}\} \rightarrow$ also a subspace since S1 ✓
  $\underbrace{\mathbf{0} + \mathbf{0} = \mathbf{0}}_{S2}$ and $\underbrace{a\mathbf{0} = \mathbf{0}}_{S3}$ for each a in $\mathbb{R}$

Subspaces

- $\mathbb{R}^n$ is a subspace of itself
- $U = \{\mathbf{0}\} \rightarrow$ also a subspace since $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $a\mathbf{0} = \mathbf{0}$ for each a in $\mathbb{R}$
- $U = \{\mathbf{0}\} \rightarrow$ called the **zero subspace**

Subspaces

- $\mathbb{R}^n$ is a subspace of itself
- $U = \{\mathbf{0}\} \rightarrow$ also a subspace since
  $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $a\mathbf{0} = \mathbf{0}$ for each a in $\mathbb{R}$
- $U = \{\mathbf{0}\} \rightarrow$ called the **zero subspace**
- Any subspace of $\mathbb{R}^n$ other than $\{\mathbf{0}\}$ or $\mathbb{R}^n$ is called a **proper** subspace

Example 5.1.2

$L = \{t\mathbf{d} \mid t \in \mathbb{R}\}$

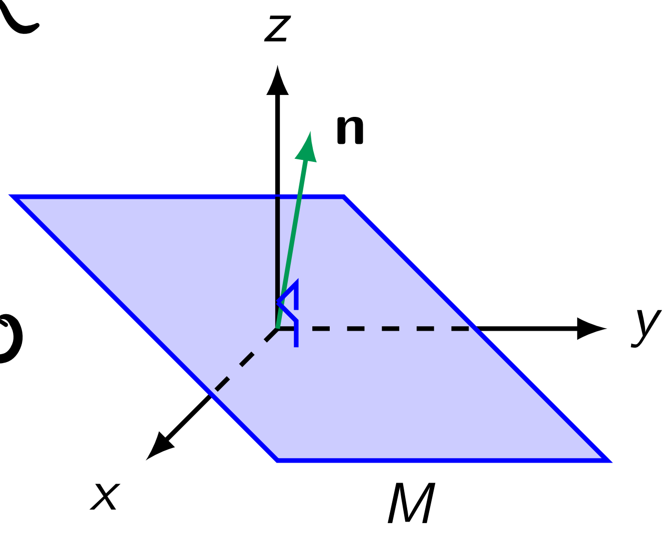
Planes and lines through the origin in $\mathbb{R}^3$ are all subspaces of $\mathbb{R}^3$.

M plane through origin

$ax + by + cz = 0$

a, b, c not all $= 0$

$\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is normal to M



$M = \{\mathbf{v} \in \mathbb{R}^3 \mid \mathbf{n} \cdot \mathbf{v} = 0\}$

$[\mathbf{n} \cdot \mathbf{v} = n_1 v_1 + n_2 v_2 + n_3 v_3]$

Show M satisfies S1, S2, S3

S1. $\mathbf{0} \in M$ $\quad \mathbf{n} \cdot \mathbf{0} = 0$

S2. If $v \in M$ $v_1 \in M$ then

$(n \cdot v = 0 \quad n \cdot v_1 = 0)$

$$n \cdot (v + v_1) = n \cdot v + n \cdot v_1 = 0$$

so $v + v_1 \in M$

S3. If $v \in M$, $a \in \mathbb{R}$

$$n \cdot (av) = a(n \cdot v) = a\,0 = 0$$

so $av \in M$

Subspaces of Matrices

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

Subspaces of Matrices

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

- null $A \rightarrow$ all solutions $\mathbf{x}$ in $\mathbb{R}^n$ of the homogeneous system $A\mathbf{x} = \mathbf{0}$

Subspaces of Matrices

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

- null $A \rightarrow$ all solutions $\mathbf{x}$ in $\mathbb{R}^n$ of the homogeneous system $A\mathbf{x} = \mathbf{0}$
- $\mathbf{x}$ is in null A if it satisfies the *condition* $A\mathbf{x} = \mathbf{0}$

Subspaces of Matrices

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

- null A $\rightarrow$ all solutions $\mathbf{x}$ in $\mathbb{R}^n$ of the homogeneous system $A\mathbf{x} = \mathbf{0}$
- $\mathbf{x}$ is in null A if it satisfies the *condition* $A\mathbf{x} = \mathbf{0}$
- **image space** of A

$$\text{im } A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$$

Subspaces of Matrices

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

- null $A \rightarrow$ all solutions $\mathbf{x}$ in $\mathbb{R}^n$ of the homogeneous system $A\mathbf{x} = \mathbf{0}$
- $\mathbf{x}$ is in null A if it satisfies the *condition* $A\mathbf{x} = \mathbf{0}$
- **image space** of A

$$\text{im } A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$$

- im $A \rightarrow$ set of all vectors $\mathbf{y}$ in $\mathbb{R}^m$ such that $A\mathbf{x} = \mathbf{y}$ *has* a solution $\mathbf{x}$

Subspaces of Matrices

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

- null $A \rightarrow$ all solutions $\mathbf{x}$ in $\mathbb{R}^n$ of the homogeneous system $A\mathbf{x} = \mathbf{0}$
- $\mathbf{x}$ is in null A if it satisfies the *condition* $A\mathbf{x} = \mathbf{0}$
- **image space** of A

$$\text{im } A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$$

- im $A \rightarrow$ set of all vectors $\mathbf{y}$ in $\mathbb{R}^m$ such that $A\mathbf{x} = \mathbf{y}$ *has* a solution $\mathbf{x}$
- im A consists of vectors of the *form* $A\mathbf{x}$ for some $\mathbf{x}$ in $\mathbb{R}^n$

Example 5.1.3

If A is an $m \times n$ matrix, then:

1. null A is a subspace of $\mathbb{R}^n$.
2. im A is a subspace of $\mathbb{R}^m$.

1. null $A = \{x \in \mathbb{R}^n \mid Ax = 0\}$

S1. $0 \in \mathbb{R}^n$ is in null A since $A0 = 0$

S2. x and x_1 in null A

$A(x + x_1) = Ax + Ax_1$
$= 0 + 0 = 0$

S3. $x \in$ null A $a \in \mathbb{R}$

$A(ax) = a(Ax)$
$= a0 = 0$

$\therefore$ null A is a subspace of $\mathbb{R}^n$

2. im $A = \{Ax \mid x \in \mathbb{R}^n\}$

S1. $0 \in \mathbb{R}^m$ $0 = A0$

S2. y $y_1 \in$ im A $y = Ax$ $y_1 = Ax_1$

for $x \in \mathbb{R}^n$ $x_1 \in \mathbb{R}^n$

$$y + y_1 = Ax + Ax_1 = A(\underbrace{\boxed{x + x_1}}_{\in \mathbb{R}^n})$$

S3. $y \in \operatorname{im} A \quad a \in \mathbb{R}$

$$ay = a(Ax) = A(\underbrace{\boxed{ax}}_{\in \mathbb{R}^n})$$

$\therefore$ im A is a subspace of $\mathbb{R}^m$

Example 5.1.4

$\lambda \to$ any number

Show $E_\lambda(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\} = \text{null}(\lambda I - A)$ is a subspace of $\mathbb{R}^n$ for each $n \times n$ matrix A and number λ.

$$x \in E_\lambda(A) \iff (\lambda I - A)x = 0$$

which is $\text{null}(\lambda I - A)$

$$Ax = \lambda x \iff (\lambda I - A)x = 0$$

Eigenspace of A

- $E_\lambda(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$ is called the **eigenspace** of A corresponding to λ

Eigenspace of A

- $E_\lambda(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$ is called the **eigenspace** of A corresponding to λ
- λ is an **eigenvalue** of A if $E_\lambda(A) \neq \{\mathbf{0}\}$

Eigenspace of A

- $E_\lambda(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$ is called the **eigenspace** of A corresponding to λ
- λ is an **eigenvalue** of A if $E_\lambda(A) \neq \{\mathbf{0}\}$
- The nonzero vectors in $E_\lambda(A)$ are the **eigenvectors** of A corresponding to λ

Caution: Not every subset of $\mathbb{R}^n$ is a subspace

$$U_1 = \left\{ \left[\begin{array}{c} x \\ y \end{array} \right] \middle| \, x \geq 0 \right\}$$ satisfies S1 ✓ and S2 ✓, but not S3;

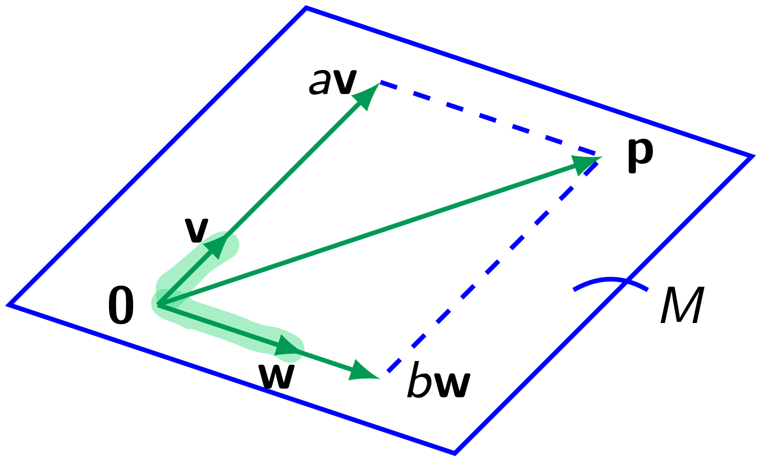
$a \begin{bmatrix} x \\ y \end{bmatrix}$ if $a < 0$ $\rightarrow$ $ax < 0$

$$U_2 = \left\{ \left[\begin{array}{c} x \\ y \end{array} \right] \middle| \, x^2 = y^2 \right\}$$ satisfies S1 and S3, but not S2;

Neither U_1 nor U_2 is a subspace of $\mathbb{R}^2$.

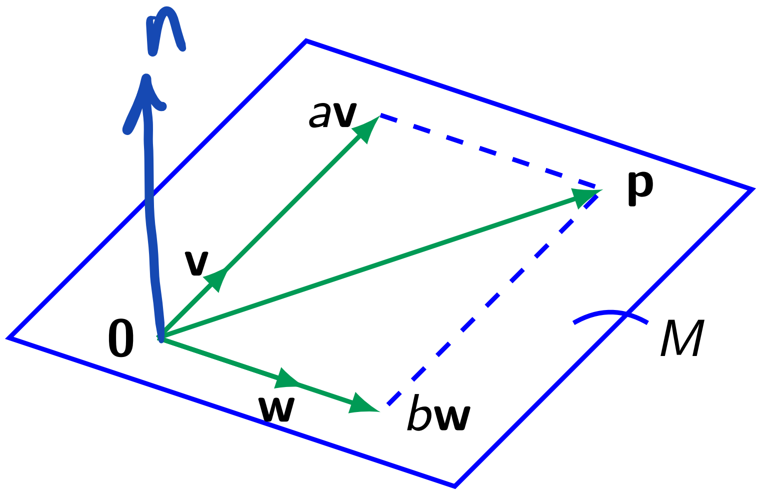
Spanning Sets

- $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3 \to$ nonzero, nonparallel, tails at the origin



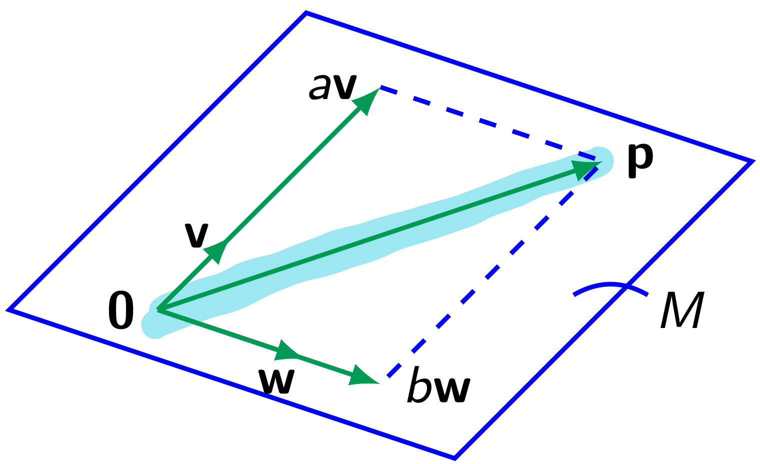
Spanning Sets

- **v** and **w** in $\mathbb{R}^3 \to$ nonzero, nonparallel, tails at the origin
- Plane M through the origin containing these vectors has normal $\mathbf{n} = \mathbf{v} \times \mathbf{w}$ and consists of all vectors **p** such that $\mathbf{n} \cdot \mathbf{p} = 0$



Spanning Sets

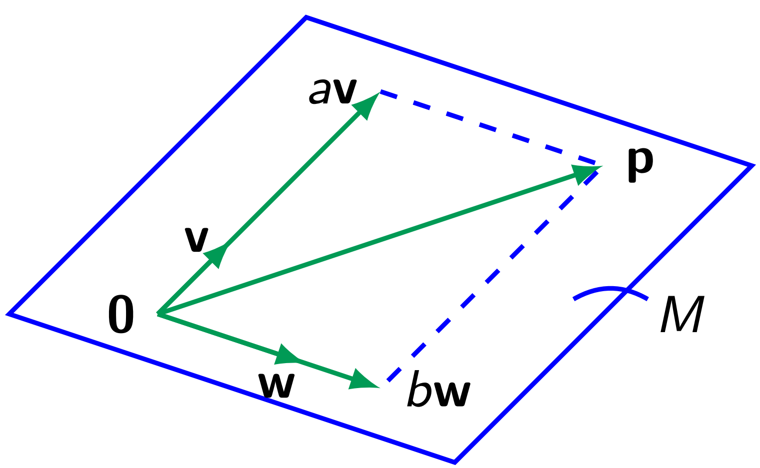
- $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$ $\to$ nonzero, nonparallel, tails at the origin
- Plane M through the origin containing these vectors has normal $\mathbf{n} = \mathbf{v} \times \mathbf{w}$ and consists of all vectors $\mathbf{p}$ such that $\mathbf{n} \cdot \mathbf{p} = 0$
- Vector $\mathbf{p}$ is in M if and only if it has the form $\mathbf{p} = a\mathbf{v} + b\mathbf{w}$ for certain real numbers a and b



Spanning Sets

p is linear comb of v and w

- $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3 \to$ nonzero, nonparallel, tails at the origin
- Plane M through the origin containing these vectors has normal $\mathbf{n} = \mathbf{v} \times \mathbf{w}$ and consists of all vectors $\mathbf{p}$ such that $\mathbf{n} \cdot \mathbf{p} = 0$
- Vector $\mathbf{p}$ is in M if and only if it has the form $\mathbf{p} = a\mathbf{v} + b\mathbf{w}$ for ~~certain~~ real numbers a and b
- Then $M = \{a\mathbf{v} + b\mathbf{w} \mid a, b \in \mathbb{R}\}$ and $\{\mathbf{v}, \mathbf{w}\}$ is a *spanning set* for M



Span in $\mathbb{R}^n$

Linear Combinations and Span in $\mathbb{R}^n$

The set of all such linear combinations is called the **span** of the $\mathbf{x}_i$ and is denoted

$$\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\} = \{t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k \mid t_i \text{ in } \mathbb{R}\}$$

If $V = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$, we say that V is **spanned** by the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$, and that the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ **span** the space V.

Span Examples

- $\text{span}\{\mathbf{x}\} = \{t\mathbf{x} \mid t \in \mathbb{R}\} = \mathbb{R}\mathbf{x}$

Span Examples

- $\text{span}\{\mathbf{x}\} = \{t\mathbf{x} \mid t \in \mathbb{R}\} = \mathbb{R}\mathbf{x}$
- $\text{span}\{\mathbf{x}, \mathbf{y}\} = \{r\mathbf{x} + s\mathbf{y} \mid r, s \in \mathbb{R}\}$

Span Examples

- $\text{span}\{\mathbf{x}\} = \{t\mathbf{x} \mid t \in \mathbb{R}\} = \mathbb{R}\mathbf{x}$
- $\text{span}\{\mathbf{x}, \mathbf{y}\} = \{r\mathbf{x} + s\mathbf{y} \mid r, s \in \mathbb{R}\}$
- $\mathbf{v}$ and $\mathbf{w}$ are two nonzero, nonparallel vectors in $\mathbb{R}^3$, $\implies$ $M = \text{span}\{\mathbf{v}, \mathbf{w}\}$ is the plane in $\mathbb{R}^3$ containing $\mathbf{v}$ and $\mathbf{w}$

Span Examples

- $\text{span}\{\mathbf{x}\} = \{t\mathbf{x} \mid t \in \mathbb{R}\} = \mathbb{R}\mathbf{x}$
- $\text{span}\{\mathbf{x}, \mathbf{y}\} = \{r\mathbf{x} + s\mathbf{y} \mid r, s \in \mathbb{R}\}$
- $\mathbf{v}$ and $\mathbf{w}$ are two nonzero, nonparallel vectors in $\mathbb{R}^3$, $\implies$ $M = \text{span}\{\mathbf{v}, \mathbf{w}\}$ is the plane in $\mathbb{R}^3$ containing $\mathbf{v}$ and $\mathbf{w}$
- $\mathbf{d}$ is any nonzero vector in $\mathbb{R}^3$ (or $\mathbb{R}^2$) $\implies$ $L = \text{span}\{\mathbf{v}\} = \{t\mathbf{d} \mid t \in \mathbb{R}\} = \mathbb{R}\mathbf{d}$ is the line with direction vector $\mathbf{d}$

Example 5.1.6

Let $\mathbf{x} = (2, -1, 2, 1)$ and $\mathbf{y} = (3, 4, -1, 1)$ in $\mathbb{R}^4$. Determine whether $\mathbf{p} = (0, -11, 8, 1)$ or $\mathbf{q} = (2, 3, 1, 2)$ are in $U = \text{span}\{\mathbf{x}, \mathbf{y}\}$.

Theorem 5.1.7

Span Theorem

Let $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ in $\mathbb{R}^n$. Then:

1. U is a subspace of $\mathbb{R}^n$ containing each $\mathbf{x}_i$.
2. If W is a subspace of $\mathbb{R}^n$ and each $\mathbf{x}_i \in W$, then $U \subseteq W$.

Example 5.1.8

If $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$, show that $\text{span}\{\mathbf{x}, \mathbf{y}\} = \text{span}\{\mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y}\}$.

Example 5.1.9

$\mathbb{R}^n = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ where $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the columns of I_n.

Note: Column j of the $n \times n$ identity matrix I_n is denoted $\mathbf{e}_j$ and called the jth **coordinate vector** in $\mathbb{R}^n$, and the set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is called the **standard basis** of $\mathbb{R}^n$.

Example 5.1.10

Given an $m \times n$ matrix A, let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ denote the basic solutions to the system $A\mathbf{x} = \mathbf{0}$ given by the gaussian algorithm. Then

$$\text{null } A = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$$

Example 5.1.11

Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ denote the columns of the $m \times n$ matrix A. Then

$$\text{im } A = \text{span}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$$

Recap

Today we saw:

- Subspaces of $\mathbb{R}^n$
- Spanning Sets

Next time: Independence and Dimension

MATH254: Linear Algebra
Lecture 2

Moira MacNeil

January 8, 2025

Last Time

Subspace of $\mathbb{R}^n$

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

S1. The zero vector $\mathbf{0} \in U$.

S2. If $\mathbf{x} \in U$ and $\mathbf{y} \in U$, then $\mathbf{x} + \mathbf{y} \in U$. closed under add

S3. If $\mathbf{x} \in U$, then $a\mathbf{x} \in U$ for every real number a.

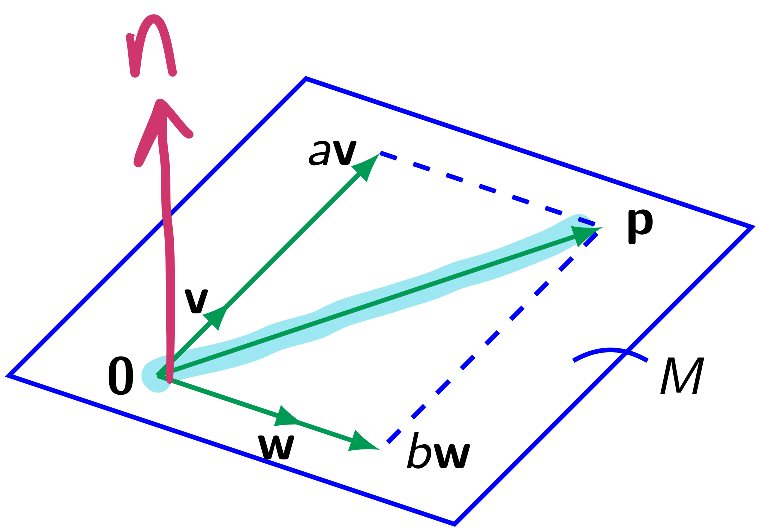
closed under scalar mult.

Today

1. Spanning Sets
2. Linear Independence

Spanning Sets

- $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3 \rightarrow$ nonzero, nonparallel, tails at the origin
- Plane M through the origin containing these vectors has normal $\mathbf{n} = \mathbf{v} \times \mathbf{w}$ and consists of all vectors $\mathbf{p}$ such that $\mathbf{n} \cdot \mathbf{p} = 0$
- Vector $\mathbf{p}$ is in M if and only if it has the form $\mathbf{p} = a\mathbf{v} + b\mathbf{w}$ for real numbers a and b
- Then $M = \{a\mathbf{v} + b\mathbf{w} \mid a, b \in \mathbb{R}\}$ and $\{\mathbf{v}, \mathbf{w}\}$ is a *spanning set* for M



Span in $\mathbb{R}^n$

Linear Combinations and Span in $\mathbb{R}^n$

The set of all such linear combinations is called the **span** of the $\mathbf{x}_i$ and is denoted

$$\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} = \{t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k \mid t_i \text{ in } \mathbb{R}\}$$

If $V = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$, we say that V is **spanned** by the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, and that the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ **span** the space V.

$V = \{w, v\}$

Span Examples

- $\text{span}\{\mathbf{x}\} = \{t\mathbf{x} \mid t \in \mathbb{R}\} \left[= \mathbb{R}\mathbf{x}\right]$

Span Examples

- $\text{span}\{\mathbf{x}\} = \{t\mathbf{x} \mid t \in \mathbb{R}\} = \mathbb{R}\mathbf{x}$
- $\text{span}\{\mathbf{x}, \mathbf{y}\} = \{r\mathbf{x} + s\mathbf{y} \mid r, s \in \mathbb{R}\}$

Span Examples

- $\text{span}\{\mathbf{x}\} = \{t\mathbf{x} \mid t \in \mathbb{R}\} = \mathbb{R}\mathbf{x}$
- $\text{span}\{\mathbf{x}, \mathbf{y}\} = \{r\mathbf{x} + s\mathbf{y} \mid r, s \in \mathbb{R}\}$
- $\mathbf{v}$ and $\mathbf{w}$ are two nonzero, nonparallel vectors in $\mathbb{R}^3$, $\implies$ $M = \text{span}\{\mathbf{v}, \mathbf{w}\}$ is the plane in $\mathbb{R}^3$ containing $\mathbf{v}$ and $\mathbf{w}$

Span Examples

- $\text{span}\{\mathbf{x}\} = \{t\mathbf{x} \mid t \in \mathbb{R}\} = \boxed{\mathbb{R}\mathbf{x}}$
- $\text{span}\{\mathbf{x}, \mathbf{y}\} = \{r\mathbf{x} + s\mathbf{y} \mid r, s \in \mathbb{R}\}$
- $\mathbf{v}$ and $\mathbf{w}$ are two nonzero, nonparallel vectors in $\mathbb{R}^3$, $\implies$ $M = \text{span}\{\mathbf{v}, \mathbf{w}\}$ is the plane in $\mathbb{R}^3$ containing $\mathbf{v}$ and $\mathbf{w}$ plane
- $\mathbf{d}$ is any nonzero vector in $\mathbb{R}^3$ (or $\mathbb{R}^2$) $\implies$ $L = \text{span}\{\mathbf{v}\} = \{t\mathbf{d} \mid t \in \mathbb{R}\} = \boxed{\mathbb{R}\mathbf{d}}$ is the line with direction vector $\mathbf{d}$

line

Example 5.1.6

Let $\mathbf{x} = (2, -1, 2, 1)$ and $\mathbf{y} = (3, 4, -1, 1)$ in $\mathbb{R}^4$. Determine whether $\mathbf{p} = (0, -11, 8, 1)$ or $\mathbf{q} = (2, 3, 1, 2)$ are in $U = \text{span}\{\mathbf{x}, \mathbf{y}\}$.

$p \in U \iff p = sx + ty$ for $s, t \in \mathbb{R}$

$$\left.\begin{aligned} 0 &= 2s + 3t \\ -11 &= -s + 4t \\ 8 &= 2s - t \\ 1 &= s + t \end{aligned}\right\} \quad \begin{aligned} s &= 3 \\ t &= -2 \end{aligned} \quad \Rightarrow p \in U$$

$q \in U \iff q = sx + ty$

$$\begin{aligned} 2 &= 2s + 3t \\ 3 &= -s + 4t \\ 1 &= 2s - t \end{aligned} \qquad 2 = s + t$$

$\rightarrow$ system has no sols.

$q \notin U$

Theorem 5.1.7

Span Theorem

Let $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ in $\mathbb{R}^n$. Then:

1. U is a subspace of $\mathbb{R}^n$ containing each $\mathbf{x}_i$.
2. If W is a subspace of $\mathbb{R}^n$ and each $\mathbf{x}_i \in W$, then $U \subseteq W$.

2. Span is the smallest subspace of $\mathbb{R}^n$ that contains each x_i

Example 5.1.8

If $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$, show that $\operatorname{span}\{\mathbf{x},\mathbf{y}\} = \operatorname{span}\{\mathbf{x}+\mathbf{y},\mathbf{x}-\mathbf{y}\}$.

$\underbrace{\operatorname{span}\{x,y\}}_{U}$, $\underbrace{\operatorname{span}\{x+y,x-y\}}_{W}$

$U \subseteq W$

$x = \frac{1}{2}(x+y) + \frac{1}{2}(x-y)$

$y = \frac{1}{2}(x+y) - \frac{1}{2}(x-y)$

are in $\operatorname{span}\{x+y, x-y\}$

$W \subseteq U$

since $x+y, x-y$ are in $\operatorname{span}\{x,y\}$

linear comb of x, y

$U \subseteq W$ and $W \subseteq U \Rightarrow U = W$

Example 5.1.9

$\mathbb{R}^n = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ where $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the columns of I_n.

Note: Column j of the $n \times n$ identity matrix I_n is denoted $\mathbf{e}_j$ and called the jth **coordinate vector** in $\mathbb{R}^n$, and the set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is called the **standard basis** of $\mathbb{R}^n$.

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \underbrace{e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\text{stand. basis } \mathbb{R}^3}$$

Any vector $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$

can be written as $x_1 e_1 + \cdots + x_n e_n = x$

Example 5.1.10

Given an $m \times n$ matrix A, let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ denote the basic solutions to $A\mathbf{x} = \mathbf{0}$ from the gaussian algorithm. Then null $A = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$.

null $\{x \mid Ax = 0\}$ if $x \in$ null A then $Ax = 0$, x is a linear comb. of basic sols.

null $A \subseteq$ span$\{x_1, \ldots, x_k\}$

$x \in$ span$\{x_1, \ldots, x_k\}$ then

$x = t_1 x_1 + \cdots + t_k x_k$

$Ax = t_1 A x_1 + \cdots + t_k A x_k =$

$= t_1 0 + \cdots + t_k 0 = 0$

$x \in$ null A

null $A \supseteq$ span$\{\ldots\}$

$\therefore$ equal

Example 5.1.11

Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ denote the columns of the $m \times n$ matrix A. Then

$$\text{im } A = \text{span}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$$

Some spanning sets are better than others

- $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a subspace of $\mathbb{R}^n \implies$ every vector in U can be written as a linear combination of the $\mathbf{x}_i$

Some spanning sets are better than others

- $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a subspace of $\mathbb{R}^n \implies$ every vector in U can be written as a linear combination of the $\mathbf{x}_i$
- Interested in spanning sets where each vector in U has *exactly one* representation as a linear combination

Some spanning sets are better than others

- $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a subspace of $\mathbb{R}^n \implies$ every vector in U can be written as a linear combination of the $\mathbf{x}_i$
- Interested in spanning sets where each vector in U has *exactly one* representation as a linear combination
- Given $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ in $\mathbb{R}^n$, suppose two linear combinations are equal:

$$r_1\mathbf{x}_1 + r_2\mathbf{x}_2 + \cdots + r_k\mathbf{x}_k = s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \cdots + s_k\mathbf{x}_k$$

Some spanning sets are better than others

- $U = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a subspace of $\mathbb{R}^n \implies$ every vector in U can be written as a linear combination of the $\mathbf{x}_i$
- Interested in spanning sets where each vector in U has *exactly one* representation as a linear combination
- Given $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ in $\mathbb{R}^n$, suppose two linear combinations are equal:

$$r_1\mathbf{x}_1 + r_2\mathbf{x}_2 + \cdots + r_k\mathbf{x}_k = s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \cdots + s_k\mathbf{x}_k$$

- Want a condition that guarantees the representation is *unique*; i.e., $r_i = s_i$ for each i

Some spanning sets are better than others

- $U = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a subspace of $\mathbb{R}^n \implies$ every vector in U can be written as a linear combination of the $\mathbf{x}_i$
- Interested in spanning sets where each vector in U has *exactly one* representation as a linear combination
- Given $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ in $\mathbb{R}^n$, suppose two linear combinations are equal:

$$r_1\mathbf{x}_1 + r_2\mathbf{x}_2 + \cdots + r_k\mathbf{x}_k = s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \cdots + s_k\mathbf{x}_k$$

- Want a condition that guarantees the representation is *unique*; i.e., $r_i = s_i$ for each i
- Rearrange to get: $(r_1 - s_1)\mathbf{x}_1 + (r_2 - s_2)\mathbf{x}_2 + \cdots + (r_k - s_k)\mathbf{x}_k = \mathbf{0}$

Some spanning sets are better than others

- $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a subspace of $\mathbb{R}^n \implies$ every vector in U can be written as a linear combination of the $\mathbf{x}_i$
- Interested in spanning sets where each vector in U has *exactly one* representation as a linear combination
- Given $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ in $\mathbb{R}^n$, suppose two linear combinations are equal:

$$r_1\mathbf{x}_1 + r_2\mathbf{x}_2 + \cdots + r_k\mathbf{x}_k = s_1\mathbf{x}_1 + s_2\mathbf{x}_2 + \cdots + s_k\mathbf{x}_k$$

- Want a condition that guarantees the representation is *unique*; i.e., $r_i = s_i$ for each i
- Rearrange to get: $(r_1 - s_1)\mathbf{x}_1 + (r_2 - s_2)\mathbf{x}_2 + \cdots + (r_k - s_k)\mathbf{x}_k = \mathbf{0}$
- Condition $\rightarrow$ all the coefficients $r_i - s_i$ are zero

Linear Independence

Linear Independence in $\mathbb{R}^n$

We call a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors **linearly independent** (or simply **independent**) if it satisfies the following condition:

$$\text{If } t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0} \text{ then } t_1 = t_2 = \cdots = t_k = 0$$

Linear Independence

Linear Independence in $\mathbb{R}^n$

We call a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors **linearly independent** (or simply **independent**) if it satisfies the following condition:

$$\text{If } t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0} \text{ then } t_1 = t_2 = \cdots = t_k = 0$$

A set of vectors in $\mathbb{R}^n$ is called **linearly dependent** (or **dependent**) if it is *not* linearly independent.

Theorem 5.2.2

Theorem

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an independent set of vectors in $\mathbb{R}^n$, then every vector in $\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ has a **unique** representation as a linear combination of the $\mathbf{x}_i$.

Independence Another Way

- A linear combination **vanishes** $\rightarrow$ equal to $\mathbf{0}$

Independence Another Way

- A linear combination **vanishes** $\rightarrow$ equal to $\mathbf{0}$
- A linear combination is **trivial** $\rightarrow$ if every coefficient is zero

Independence Another Way

- A linear combination **vanishes** $\rightarrow$ equal to $\mathbf{0}$
- A linear combination is **trivial** $\rightarrow$ if every coefficient is zero
- A set of vectors is independent if and only if the only linear combination that vanishes is the trivial one

$$-2\begin{bmatrix}1\\1\end{bmatrix}+1\begin{bmatrix}2\\2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

Independence Another Way

- A linear combination **vanishes** $\rightarrow$ equal to $\mathbf{0}$
- A linear combination is **trivial** $\rightarrow$ if every coefficient is zero
- A set of vectors is independent if and only if the only linear combination that vanishes is the trivial one
- A set of vectors is dependent if some nontrivial linear combination vanishes

$$-2\begin{bmatrix}1\\1\end{bmatrix}+1\begin{bmatrix}2\\2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

dependent

Theorem 5.2.3

Independence Test

To verify that a set $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ of vectors in $\mathbb{R}^n$ is independent, proceed as follows:

1. Set a linear combination equal to zero:
   $t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0}$.
2. Show that $t_i = 0$ for each i (that is, the linear combination is trivial).

If some nontrivial linear combination vanishes, the vectors are not independent.

Example 5.2.4

Determine whether $\{(1,0,-2,5),(2,1,0,-1),(1,1,2,1)\}$ is independent in $\mathbb{R}^4$.

Suppose

$$r(1,0,-2,5)+s(2,1,0,-1)+t(1,1,2,1)=0$$

$$\begin{aligned} r+2s+t&=0\\ s+t&=0\\ -2r+\quad+2t&=0\\ 5r-s+t&=0 \end{aligned}$$

only sol is trivial

$r=s=t=0$

$\Rightarrow$ vectors are indept.

Example 5.2.5

$e_j = \text{col}\ j\ I_n$

Show that the standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$ is independent.

$t_1 e_1 + t_2 e_2 + \cdots + t_n e_n = t$

components are $t_1, t_2, \cdots t_n$

linear comb vanishes $\Leftrightarrow$

each $t_i = 0$.

$$t = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$$

Example 5.2.6

If $\{\mathbf{x}, \mathbf{y}\}$ is independent, show that $\{2\mathbf{x} + 3\mathbf{y}, \mathbf{x} - 5\mathbf{y}\}$ is also independent.

$$s(2x+3y) + t(x-5y) = 0$$

$$(2s+t)x + (3s-5t)y = 0$$

$\{x, y\}$ indep. $\Rightarrow$ lin. comb. must be trivial

$2s+t=0 \quad 3s-5t=0 \rightarrow$ only trivial sol

$s=t=0$

Example 5.2.7

Show that the zero vector in $\mathbb{R}^n$ does not belong to any independent set.

$$\{0, x_1, \dots, x_k\}$$

$$a0 + 0x_1 + \cdots + x_k = 0$$

nontrivial

Example 5.2.8

Given $\mathbf{x}$ in $\mathbb{R}^n$, show that $\{\mathbf{x}\}$ is independent if and only if $\mathbf{x} \neq \mathbf{0}$.

$t x = 0 \quad t \in \mathbb{R}$

$t = 0$ since $x \neq 0$

Example 5.2.9

Show that the nonzero rows of a row-echelon matrix R are independent.

Recap

Today we saw:

- Spanning Sets
- Linear Independence

Next time: More Linear Independence, Dimension

MATH254: Linear Algebra
Lecture 3

Moira MacNeil

January 10, 2025

Last Time

Linear Combinations and Span in $\mathbb{R}^n$

The set of all such linear combinations is called the **span** of the $\mathbf{x}_i$ and is denoted

$$\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\} = \{t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k \mid t_i \text{ in } \mathbb{R}\}$$

If $V = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$, we say that V is **spanned** by the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$, and that the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ **span** the space V.

Last Time

Linear Independence in $\mathbb{R}^n$

We call a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors **linearly independent** (or simply **independent** if it satisfies the following condition:

$$\text{If } t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \underline{\mathbf{0}} \text{ then } \overbrace{t_1 = t_2 = \cdots = t_k = 0}^{\text{trivial}}$$

vanishes → 0

A set of vectors is independent if and only if the only linear combination that vanishes is the trivial one.

dependent → not indep.

Today

1. Linear Independence continued

2. Dimension

Note: Assignment 1 is posted. Due next Friday. (Only 3 short questions!)

Example 5.2.9

Show that the nonzero rows of a row-echelon matrix R are independent.

$$R = \begin{array}{c} \begin{bmatrix} 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \begin{array}{l} R_1 \\ R_2 \\ R_3 \\ \end{array}$$

$\to$ three leading 1s

$* =$ a nonspec. #

If $t_1R_1 + t_2R_2 + t_3R_3 = 0$ then we must have $t_1 = t_2 = t_3 = 0$

$(0, t_1, *, *, *, *) + (0, 0, 0, t_2, *, *) + (0, 0, 0, 0, t_3, *) = \underline{0}$

Equate entry 2 $\to$ $t_1 = 0$

$t_2R_2 + t_3R_3 = 0$ $\quad$ $t_2 = 0 \to t_3 = 0$

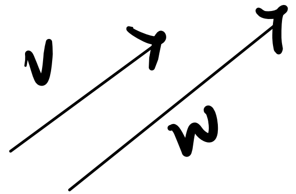
$t_3R_3 = 0$

nontriv. comb that vanishes

Example 5.2.10

If $\mathbf{v}$ and $\mathbf{w}$ are nonzero vectors in $\mathbb{R}^3$, show that $\{\mathbf{v}, \mathbf{w}\}$ is dependent if and only if $\mathbf{v}$ and $\mathbf{w}$ are parallel. $(\Leftarrow)$

v, w paral. $\to$ one is a scalar mult. of the other



$v = aw \quad a \in \mathbb{R}$

$v - \underbrace{aw}_{v} = 0$ nontriv. linear comb coeff $1, -a$ $\Rightarrow \{v, w\}$ dep.

$(\Rightarrow)$ $\{v, w\}$ dept., let $sv + tw = 0$ be nontriv.

$s \neq 0$ $\quad v = \dfrac{-t}{s} w \Rightarrow v$ scalar mult w

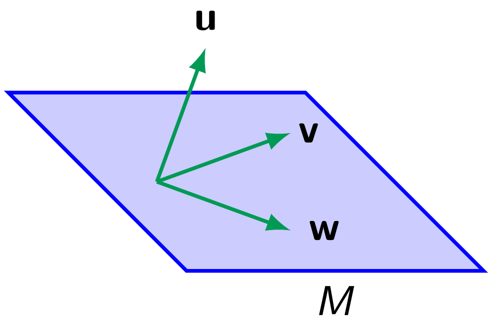
$\Rightarrow v, w$ parallel

Example 5.2.11

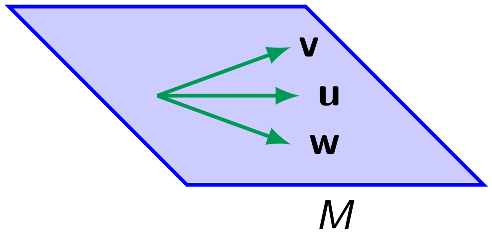
Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be nonzero vectors in $\mathbb{R}^3$ where $\{\mathbf{v}, \mathbf{w}\}$ independent. Show that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is independent if and only if $\mathbf{u}$ is not in the plane $M = \text{span}\{\mathbf{v}, \mathbf{w}\}$.

$\{u, v, w\}$ indep. and $u \in M$ $\quad u = av + bw$

$1u - av - bw = 0 \Rightarrow \Leftarrow \{u, v, w\}$ indep.



$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ independent



$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ not independent

let $u \notin M$ $\quad$ if $ru + sv + tw = 0$

$r = 0$ o/w $u = -\frac{s}{r} v + \frac{t}{r} w \rightarrow$ $sv + tw = 0$

$s = t = 0$ by assump.

assume $\{v, w\}$ independent

$\therefore \{u, v, w\}$ idep.

A $n \times n$

A invertible

$\Leftrightarrow$ $Ax = 0 \Rightarrow x = 0$

$\Leftrightarrow$ $Ax = b$ has a solution x for every $b \in \mathbb{R}^n$

$$A = \overset{\text{columns}}{[c_1, \dots, c_n]}$$

$$Ax = x_1 c_1 + x_2 c_2 + \dots + x_n c_n$$

Theorem 5.2.12

Theorem

If A is an $m \times n$ matrix, let $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ denote the columns of A.

1. $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ is independent in $\mathbb{R}^m$ if and only if $A\mathbf{x} = \mathbf{0}$, $\mathbf{x}$ in $\mathbb{R}^n$, implies $\mathbf{x} = \mathbf{0}$.
2. $\mathbb{R}^m = \text{span}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ if and only if $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x}$ for every vector $\mathbf{b}$ in $\mathbb{R}^m$.

Theorem 5.2.13

Theorem

The following are equivalent for an $n \times n$ matrix A:

1. A is invertible.
2. The columns of A are linearly independent.
3. The columns of A span $\mathbb{R}^n$.
4. The rows of A are linearly independent.
5. The rows of A span the set of all $1 \times n$ rows.

Example 5.2.14

WTS = want to show

Show that $S = \{(2,-2,5),(-3,1,1),(2,7,-4)\}$ is independent in $\mathbb{R}^3$.

$$A = \begin{bmatrix} 2 & -2 & 5 \\ -3 & 1 & 1 \\ 2 & 7 & -4 \end{bmatrix}$$

WTS rows of A indep.

$\Leftrightarrow$ A invertible

$\Leftrightarrow$ $\det(A) \neq 0$

$\det(A) = -117 \neq 0 \Rightarrow$ A invert.

$\Rightarrow$ S is indept.

also

$\mathbb{R}^3 = \operatorname{span} S$

Theorem 5.2.15

Fundamental Theorem

Let U be a subspace of $\mathbb{R}^n$. If U is spanned by m vectors, and if U contains k linearly independent vectors, then $k \leq m$.

Basis of $\mathbb{R}^n$

1 basis
multiple bases

Basis of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$, a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ of vectors in U is called a **basis** of U if it satisfies the following two conditions:

1. $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is linearly independent.
2. $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$.

Theorem 5.2.17

Invariance Theorem

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ and $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$ are bases of a subspace U of $\mathbb{R}^n$, then $m = k$.

By fund. thm we have $k \le m$
since $\{x_1, \ldots, x_m\}$ spans U and $\{y_1, \ldots, y_k\}$
indept.
Also have $m \le k$ (interchange x, y)
So $m = k$.

Dimension of a Subspace

Dimension of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$ and $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of U, the number, m, of vectors in the basis is called the **dimension** of U, denoted

$$\underline{\dim\ U} = m$$

Dimension of a Subspace

Dimension of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$ and $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of U, the number, m, of vectors in the basis is called the **dimension** of U, denoted

$$\dim\ U = m$$

- Define $\dim\{\mathbf{0}\} = 0$
- $\{\mathbf{0}\}$ has a basis containing *no* vector
- Recall: $\mathbf{0}$ cannot belong to *any* independent set.

Example 5.2.19

$\dim(\mathbb{R}^n) = n$ and $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis.

standard basis

↳ set of columns of Identity matrix

We have shown

$\mathbb{R}^n = \text{span}\{e_1, \ldots, e_n\}$

$\{e_1, \ldots, e_n\}$ is independent

$\Rightarrow$ basis of $\mathbb{R}^n$

$\dim \mathbb{R}^n = n$

Example 5.2.20

Let $U = \left\{ \begin{bmatrix} r \\ s \\ r \end{bmatrix} \middle| \, r, s \text{ in } \mathbb{R} \right\}$. Show that U is a subspace of $\mathbb{R}^3$, find a basis, and calculate $\dim U$.

$$\begin{bmatrix} r \\ s \\ r \end{bmatrix} = r u + s v \qquad u = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad v = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$\Rightarrow \quad U = \operatorname{span}\{u, v\} \Rightarrow U$ subspace $\mathbb{R}^3$

if $\quad ru + sv = 0 \Rightarrow \begin{bmatrix} r \\ s \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{matrix} r = 0 \\ s = 0 \end{matrix}$

so $\{u\ v\}$ indept and basis of $U \qquad \dim(U) = 2$

Example 5.2.21

Let $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis of $\mathbb{R}^n$. If A is an invertible $n \times n$ matrix, then $D = \{A\mathbf{x}_1, A\mathbf{x}_2, \ldots, A\mathbf{x}_n\}$ is also a basis of $\mathbb{R}^n$.

Theorem 5.2.22

Theorem

Let $U \neq \{\mathbf{0}\}$ be a subspace of $\mathbb{R}^n$. Then:

1. U has a basis and $\dim U \leq n$.
2. Any independent set in U can be enlarged (by adding vectors from any fixed basis of U) to a basis of U.
3. Any spanning set for U can be cut down (by deleting vectors) to a basis of U.

Example 5.2.23

Find a basis of $\mathbb{R}^4$ containing $S = \{\mathbf{u}, \mathbf{v}\}$ where $\mathbf{u} = (0, 1, 2, 3)$ and $\mathbf{v} = (2, -1, 0, 1)$.

Add vectors from stand. basis of $\mathbb{R}^4$ to S

try $e_1 \to \{e_1, u, v\}$ indep.

add $e_2 \to B = \{e_1, e_2, u, v\}$ indep.

B has $4 = \dim \mathbb{R}^4$ vectors

then B must span $\mathbb{R}^4$

B is basis of $\mathbb{R}^4$

Theorem 5.2.24

Theorem

Let U be a subspace of $\mathbb{R}^n$ where $\dim U = m$ and let $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ be a set of m vectors in U. Then B is independent if and only if B spans U.

Theorem 5.2.25

Theorem

Let $U \subseteq W$ be subspaces of $\mathbb{R}^n$. Then:

1. $\dim U \leq \dim W$.
2. If $\dim U = \dim W$, then $U = W$.

Proper Subspaces

- If U is a subspace of $\mathbb{R}^n$, then $\dim U$ is one of the integers $0, 1, 2, \ldots, n$

Proper Subspaces

- If U is a subspace of $\mathbb{R}^n$, then $\dim U$ is one of the integers $0, 1, 2, \dots, n$
- And:

$$\begin{array}{lll} \dim\ U = 0 & \text{if and only if} & U = \{\mathbf{0}\}, \\ \dim\ U = n & \text{if and only if} & U = \mathbb{R}^n \end{array}$$

Proper Subspaces

- If U is a subspace of $\mathbb{R}^n$, then $\dim U$ is one of the integers $0, 1, 2, \ldots, n$
- And:

$$\begin{array}{lll} \dim\ U = 0 & \text{if and only if} & U = \{\mathbf{0}\}, \\ \dim\ U = n & \text{if and only if} & U = \mathbb{R}^n \end{array}$$

- The other subspaces of $\mathbb{R}^n$ are called **proper**

Example 5.2.26

(a) If U is a subspace of $\mathbb{R}^2$ or $\mathbb{R}^3$, then $\dim U = 1$ if and only if U is a line through the origin

(b) If U is a subspace of $\mathbb{R}^3$, then $\dim U = 2$ if and only if U is a plane through the origin

Recap

Today we saw:

- Linear Independence
- Dimension

Next time: Orthogonality

MATH254: Linear Algebra
Lecture 4

Moira MacNeil

January 14, 2025

Last Time

Linear Independence in $\mathbb{R}^n$

We call a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors **linearly independent** (or simply **independent** if it satisfies the following condition:

$$\text{If } t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0} \text{ then } t_1 = t_2 = \cdots = t_k = 0$$

Last Time

Basis of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$, a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ of vectors in U is called a **basis** of U if it satisfies the following two conditions:

1. $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is linearly independent.
2. $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\} = \{t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k \mid t_i \text{ in } \mathbb{R}\}$

Dimension of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$ and $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of U, the number, m, of vectors in the basis is called the **dimension** of U, denoted

$$\dim\ U = m$$

Today

1. Extend Dot Product, Length, and Distance to $\mathbb{R}^n$

2. Orthogonal Sets and the Expansion Theorem

Reminder: Assignment 1 is due Friday

Office Hours on Thursday - cancelled

Dot Product

Dot Product

If $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ are two n-tuples in $\mathbb{R}^n$, recall that their **dot product** is the scalar:

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \qquad x \cdot y = x^T y$$

Vector Length

Length in $\mathbb{R}^n$

As in $\mathbb{R}^3$, the **length** $\|\mathbf{x}\|$ of the vector is defined by

$$\|\mathbf{x}\| = \sqrt{\underline{\mathbf{x} \cdot \mathbf{x}}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Where $\sqrt{(\quad)}$ indicates the positive square root.

Vector Length

Length in $\mathbb{R}^n$

As in $\mathbb{R}^3$, the **length** $\|\mathbf{x}\|$ of the vector is defined by

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Where $\sqrt{(\quad)}$ indicates the positive square root.

- Unit vector $\rightarrow$ vector $\mathbf{x}$ of length 1

Vector Length

Length in $\mathbb{R}^n$

As in $\mathbb{R}^3$, the **length** $\|\mathbf{x}\|$ of the vector is defined by

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}\cdot\mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Where $\sqrt{(\quad)}$ indicates the positive square root.

- Unit vector $\rightarrow$ vector $\mathbf{x}$ of length 1
- $\mathbf{x} \neq \mathbf{0} \implies \|\mathbf{x}\| \neq 0 \rightarrow \frac{1}{\|\mathbf{x}\|}\mathbf{x}$ is a unit vector

normalizing

Example 5.3.2

$\frac{1}{2}$ w

If $\mathbf{x} = (1, -1, -3, 1)$ and $\mathbf{y} = (2, 1, 1, 0)$ in $\mathbb{R}^4$, find unit vectors using $\mathbf{x}$ and $\mathbf{y}$.

$$\|x\| = \sqrt{x \cdot x} = \sqrt{1^2 + (-1)^2 + (-3)^2 + 1^2}$$
$$= \sqrt{12} = 2\sqrt{3}$$

$$\frac{1}{\|x\|} x = \frac{1}{2\sqrt{3}} x \rightarrow \text{unit vector}$$

$$\|y\| = \sqrt{y \cdot y} = \sqrt{2^2 + 1^2 + 1^2 + 0^2} = \sqrt{6}$$

$$\frac{1}{\|y\|} y = \frac{1}{\sqrt{6}} y \rightarrow \text{unit vector}$$

Properties of Dot Product and Length

Theorem 5.3.3

Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ denote vectors in $\mathbb{R}^n$. Then:

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.
2. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$.
3. $(a\mathbf{x}) \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x} \cdot (a\mathbf{y})$ for all scalars a.
4. $\|\mathbf{x}\|^2 = \mathbf{x} \cdot \mathbf{x}$.
5. $\|\mathbf{x}\| \geq 0$, and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
6. $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$ for all scalars a.

Example 5.3.4

Show that $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) + \|\mathbf{y}\|^2$ for any $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

$$\begin{aligned} \|x+y\|^2 &= (x+y)\cdot(x+y) \\ &= x\cdot x + 2(x\cdot y) + y\cdot y \\ &= \|x\|^2 + 2(x\cdot y) + \|y\|^2 \end{aligned}$$

$$\|x\|^2 = \sqrt{x\cdot x}^{\,2}$$

Example 5.3.5

Suppose that $\mathbb{R}^n = \text{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$ for some vectors $\mathbf{f}_i$. If $\mathbf{x} \cdot \mathbf{f}_i = 0$ for each i where $\mathbf{x}$ is in $\mathbb{R}^n$, show that $\mathbf{x} = \mathbf{0}$.

Show $x=0$ by showing $\|x\| = 0$
$(\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0)$

f_i span $\mathbb{R}^n \Rightarrow x = t_1 f_1 + t_2 f_2 + \cdots + t_k f_k$

$$\begin{aligned}
\|x\|^2 = x \cdot x &= x \cdot (t_1 f_1 + t_2 f_2 + \cdots + t_k f_k) \\
&= t_1 (x \cdot f_1) + t_2 (x \cdot f_2) + \cdots + t_k (x \cdot f_k) \\
&= t_1 (0) + t_2 (0) + \cdots + t_k (0) \\
&= 0
\end{aligned}$$

$\|x\|^2 = 0 \Rightarrow \|x\| = 0 \Rightarrow x = 0$

Theorem 5.3.6

Cauchy Inequality

If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, then

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$$

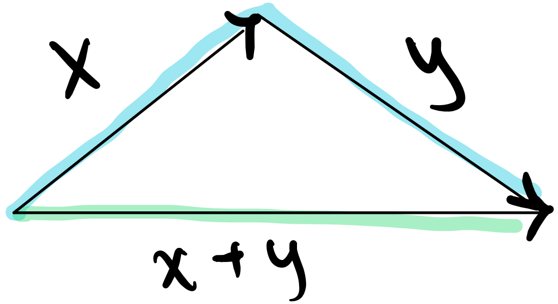
Moreover $|\mathbf{x} \cdot \mathbf{y}| = \|\mathbf{x}\|\|\mathbf{y}\|$ if and only if one of $\mathbf{x}$ and $\mathbf{y}$ is a multiple of the other.

Equiv. $(x \cdot y)^2 \leq \|x\|^2 \|y\|^2$

Corollary 5.3.7

Triangle Inequality

If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, then $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

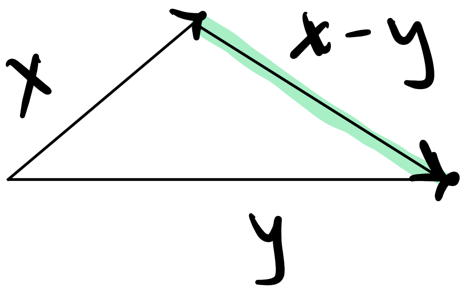


Distance

Distance in $\mathbb{R}^n$

If $\mathbf{x}$ and $\mathbf{y}$ are two vectors in $\mathbb{R}^n$, we define the **distance** $d(\mathbf{x}, \mathbf{y})$ between $\mathbf{x}$ and $\mathbf{y}$ by

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$



Properties of Distance

Theorem 5.3.9

If $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ are three vectors in $\mathbb{R}^n$ we have:

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}$ and $\mathbf{y}$.
2. $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$.
3. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}$ and $\mathbf{y}$.
4. $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$for all $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$. Triangle inequality.

$$\begin{aligned} d(x, z) = \|x - z\| &= \|(x-y)+(y-z)\| \\ &\leq \|x-y\| + \|y-z\| \\ &= d(x,y) + d(y,z) \end{aligned}$$

Orthogonal and Orthonormal Sets

Orthogonal and Orthonormal Sets

Two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ are **orthogonal** if $\mathbf{x} \cdot \mathbf{y} = 0$, extending the terminology in $\mathbb{R}^3$. More generally, a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors in $\mathbb{R}^n$ is called an **orthogonal set** if

$$\underbrace{\mathbf{x}_i \cdot \mathbf{x}_j = 0 \text{ for all } i \neq j}_{\text{pairwise}} \quad \text{and} \quad \mathbf{x}_i \neq \mathbf{0} \text{ for all } i$$

all vectors orthogonal

Note that $\{\mathbf{x}\}$ is an orthogonal set if $\mathbf{x} \neq \mathbf{0}$. A set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors in $\mathbb{R}^n$ is called **orthonormal** if it is orthogonal and, in addition, each $\mathbf{x}_i$ is a unit vector:

$$\|\mathbf{x}_i\| = 1 \text{ for each } i.$$

cols of I_n

The standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an orthonormal set in $\mathbb{R}^n$.

$$\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

Normalization

Normalizing an Orthogonal Set

Hence if $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an orthogonal set, then $\{\frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1, \frac{1}{\|\mathbf{x}_2\|}\mathbf{x}_2, \cdots, \frac{1}{\|\mathbf{x}_k\|}\mathbf{x}_k\}$ is an orthonormal set, and we say that it is the result of **normalizing** the orthogonal set $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}$.

Example 5.3.12

If $\mathbf{f}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{f}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{f}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{f}_4 = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 1 \end{bmatrix}$, verify $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ is an orthogonal set in $\mathbb{R}^4$.

Orthogonal? ✓

$f_1 \cdot f_2 = 1+1-2=0$

$f_1 \cdot f_3 = -1+1=0$

$f_1 \cdot f_4 = -1+3-1-1=0$

$f_2 \cdot f_3 = -1+1=0$

$f_2 \cdot f_4 = -1-1+2=0$

$f_3 \cdot f_4 = 1-1=0$

orthonormal? ✗

$\|f_1\| = \sqrt{4} = 2$

$\|f_2\| = \sqrt{6}$

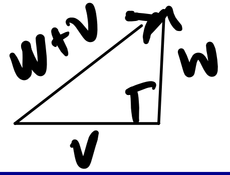
$\|f_3\| = \sqrt{2}$

$\|f_4\| = \sqrt{12}$

Normalize:

$\{ \frac{1}{2} f_1, \frac{1}{\sqrt{6}} f_2, \frac{1}{\sqrt{2}} f_3, \frac{1}{\sqrt{12}} f_4 \}$

Theorem 5.3.13



$$\|w+v\|^2 = \|v\|^2 + \|w\|^2$$

Pythagoras' Theorem

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an orthogonal set in $\mathbb{R}^n$, then

$$\|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k\|^2 = \|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 + \cdots + \|\mathbf{x}_k\|^2.$$

Proof: $\{x_1, \ldots, x_k\}$ orthogonal $\Rightarrow$ $x_i \cdot x_j = 0$ for $i \neq j$

$$\|x_1 + x_2 + \cdots + x_k\|^2 = (x_1 + \cdots + x_k) \cdot (x_1 + \cdots + x_k)$$

$$= (x_1 \cdot x_1 + x_2 \cdot x_2 + \cdots + x_k \cdot x_k) + \sum_{i \neq j} \cancel{x_i \cdot x_j}$$

$$= \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_k\|^2 \quad \square$$

Theorem

Theorem

Every orthogonal set in $\mathbb{R}^n$ is linearly independent.

Theorem 5.3.14

Expansion Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal basis of a subspace U of $\mathbb{R}^n$. If $\mathbf{x}$ is any vector in U, we have

$$\mathbf{x} = \left(\frac{\mathbf{x} \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\right)\mathbf{f}_1 + \left(\frac{\mathbf{x} \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\right)\mathbf{f}_2 + \cdots + \left(\frac{\mathbf{x} \cdot \mathbf{f}_m}{\|\mathbf{f}_m\|^2}\right)\mathbf{f}_m$$

Theorem 5.3.14

Expansion Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal basis of a subspace U of $\mathbb{R}^n$. If $\mathbf{x}$ is any vector in U, we have

$$\mathbf{x} = \boxed{\left(\frac{\mathbf{x}\cdot\mathbf{f}_1}{\|\mathbf{f}_1\|^2}\right)}\mathbf{f}_1 + \left(\frac{\mathbf{x}\cdot\mathbf{f}_2}{\|\mathbf{f}_2\|^2}\right)\mathbf{f}_2 + \cdots + \left(\frac{\mathbf{x}\cdot\mathbf{f}_m}{\|\mathbf{f}_m\|^2}\right)\mathbf{f}_m$$

- This expansion of $\mathbf{x}$ is called the **Fourier expansion** of $\mathbf{x}$,
- The coefficients are called the **Fourier coefficients**
- If $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is actually orthonormal, then $t_i = \mathbf{x}\cdot\mathbf{f}_i$ for each i

Example 5.3.15

Expand $\mathbf{x} = (a, b, c, d)$ as a linear combination of the orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ of $\mathbb{R}^4$ given in the previous example.

$f_1 = (1,1,1,-1)$ $f_2 = (1,0,1,2)$ $f_3 = (-1,0,1,0)$

$f_4 = (-1,3,-1,1)$

$\|f_1\| = \sqrt{4} = 2$

$\|f_2\| = \sqrt{6}$

$\|f_3\| = \sqrt{2}$

$\|f_4\| = \sqrt{12}$

Fourier Coeff:

$t_1 = 1/4(a+b+c-d)$

$t_2 = 1/6(a+c+2d)$

$t_3 = 1/2(-a+c)$

$t_4 = 1/12(-a+3b-c+d)$

$x = t_1 f_1 + t_2 f_2 + t_3 f_3 + t_4 f_4 = (a,b,c,d)$

Recap

Today we saw:

- Dot Product, Length, and Distance in $\mathbb{R}^n$
- Orthogonal Sets
- Expansion Theorem

Next time: Matrix Rank

MATH254: Linear Algebra
Lecture 5

Moira MacNeil

January 15, 2025

Last Time

1. Extend Dot Product, Length, and Distance to $\mathbb{R}^n$
2. Orthogonal Sets and the Expansion Theorem

Today

1. Rank of a Matrix

Reminders:

- Assignment 1 is due Friday
- Thursday office hours are cancelled

Column & Row Space

Column and Row Space of a Matrix

Let A be an $m \times n$ matrix.
The **column space**, col$A$, of A is the subspace of $\mathbb{R}^m$ spanned by the columns of A.
The **row space**, row$A$, of A is the subspace of $\mathbb{R}^n$ spanned by the rows of A.

Recall: Elementary Row Operations

Elementary Row Operations

i Swap two rows

ii Multiply one row by a nonzero number

iii Add a multiple of one row to a different row

Lemma 5.4.2

Lemma

Let A and B denote $m \times n$ matrices.

1. If $A \rightarrow B$ by elementary row operations, then row $A =$ row B.
2. If $A \rightarrow B$ by elementary column operations, then col $A =$ col B.

Lemma 5.4.3

Lemma

If R is a row-echelon matrix, then

1. The nonzero rows of R are a basis of row$(R.)$
2. The columns of R containing leading ones are a basis of col R.

PROOF (2) $c_{j_1}, \cdots, c_{j_r}$ denote the cols of R containing leading 1s. This set of cols is indept because all leading 1s are in different rows (zeros below + to left)

Let U be subspace of all cols of $\mathbb{R}^m$ in which last $m-r$ entries $= 0$

Then $\dim U = r$ (add extra 0s to $\mathbb{R}^r$)

Then $\{c_{j_1}, \dots, c_{j_r}\}$ is a basis of U (THM)

since each c_{j_i} is in col R then

col $R = U$ $\square$

Recall: Matrix Rank

Rank

The rank of matrix A is the number of leading 1s in any row-echelon matrix to which A can be carried by row operations.

NOT UNIQUE

Recall: Matrix Rank

Rank

The rank of matrix A is the number of leading 1s in any row-echelon matrix to which A can be carried by row operations.

- Carry matrix A to row-echelon matrix R by row operations $\rightarrow$ R is not unique

Recall: Matrix Rank

Rank

The rank of matrix A is the number of leading 1s in any row-echelon matrix to which A can be carried by row operations.

- Carry matrix A to row-echelon matrix R by row operations $\rightarrow$ R is not unique
- Rank does not depend on the choice of R

Recall: Matrix Rank

Rank

The rank of matrix A is the number of leading 1s in any row-echelon matrix to which A can be carried by row operations.

- Carry matrix A to row-echelon matrix R by row operations $\to$ R is not unique
- Rank does not depend on the choice of R
- Part 1. of the previous Lemma shows:

$A \to R \quad \text{row}A = \text{row}R$

$$\text{rank } A = \dim(\text{row } A) = \dim(\text{row} R)$$

$\implies$ rank$A$ independent of R

Example 5.4.4

Find a basis of $U = \text{span}\{(1,1,2,3),(2,4,1,0),(1,5,-4,-9)\}$.

U is row space of $A = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 2 & 4 & 1 & 0 \\ 1 & 5 & -4 & -9 \end{bmatrix}$

$A \to$ row echelon form
(non zero rows of RE are basis of row A)

$$\begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 2 & -3 & -6 \\ 0 & 4 & -6 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & -3/2 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so $\{(1,1,2,3), (0,1,-3/2,-3)\}$ is a basis of U

Theorem 5.4.5

Rank Theorem

Let A denote any $m \times n$ matrix of rank r. Then

$$\dim(\text{col } A) = \dim(\text{row } A) = r$$

Moreover, if A is carried to a row-echelon matrix R by row operations, then

1. The r nonzero rows of R are a basis of row$A$.
2. If the leading 1s lie in columns $j_1, j_2, \ldots, j_r$ of R, then columns $j_1, j_2, \ldots, j_r$ of A are a basis of col$A$.

Example 5.4.6

Compute the rank of $A = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 3 & 6 & 5 & 0 \\ 1 & 2 & 1 & 2 \end{bmatrix}$ and find bases for row$A$ and col$A$.

① $A \to RE$

$$\begin{bmatrix} 1 & 2 & 2 & -1 \\ 3 & 6 & 5 & 0 \\ 1 & 2 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 2 & -1 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & -1 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 2 & -1 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

rank $A = 2$ $\quad \{(1,2,2,-1),(0,0,1,-3)\}$ basis of row A

columns 1, 3 have leading 1's

$\Rightarrow$ cols 1, 3 of A are a basis of col A $\left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} \right\}$

Corollary 5.4.7

Corollary

If A is any matrix, then $\operatorname{rank} A = \operatorname{rank}(A^T)$.

rows of A indep. (span row A) $\Leftrightarrow$
transposes indep. (span col A)

Corollary 5.4.8

Corollary

If A is an $m \times n$ matrix, then rank $A \leq m$ and rank $A \leq n$.

$\text{row } A \subseteq \mathbb{R}^n \qquad \dim(\text{row } A) \leq \dim(\mathbb{R}^n) = n$

$\text{col } A \subseteq \mathbb{R}^m \qquad \dim(\text{col } A) \leq \dim(\mathbb{R}^m) = m$

Corollary 5.4.9

Corollary

$\text{rank} A = \text{rank}(UA) = \text{rank}(AV)$ whenever U and V are invertible.

PROOF $\text{rank}(A) = \text{rank}(UA)$ by Lem 5.2.2

$$\text{rank}(AV) = \text{rank}\,(AV)^T = \text{rank}\, V^T A^T$$
$$= \text{rank}\, A^T$$
$$= \text{rank}\, A$$

Lemma 5.4.10

Lemma

Let A, U, and V be matrices of sizes $m \times n$, $p \times m$, and $n \times q$ respectively.

1. $\text{col}(AV) \subseteq \text{col}A$, with equality if $VV' = I_n$ for some V'.
2. $\text{row}(UA) \subseteq \text{row}A$, with equality if $U'U = I_m$ for some U'.

Corollary

Corollary

If A is $m \times n$ and B is $n \times m$, then $\text{rank}AB \leq \text{rank}A$ and $\text{rank}AB \leq \text{rank}B$.

Prev. Lemma says $\text{col } AB \subseteq \text{col } A$

$\text{row } BA \subseteq \text{row } B$

Rank Theorem

Recall: Null Space and Image Space

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

Recall: Null Space and Image Space

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

- **image space** of A

$$\text{im } A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$$

Recall: Null Space and Image Space

- **null space** of an $m \times n$ matrix A

$$\text{null } A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

- **image space** of A

$$\text{im } A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$$

- If A has rank r, we have $\text{im }(A) = \text{col}(A)$, so $\dim[\text{im }(A)] = \dim[\text{col}(A)] = r$. Hence Rank Theorem provides a method of finding a basis of $\text{im }(A)$.

Theorem 5.4.12

Theorem

Let A denote an $m \times n$ matrix of rank r. Then

1. The $n - r$ basic solutions to the system $A\mathbf{x} = \mathbf{0}$ provided by the gaussian algorithm are a basis of null (A), so $\dim[\text{null }(A)] = n - r$.
2. Rank Theorem provides a basis of im $(A) = \text{col}(A)$, and $\dim[\text{im }(A)] = r$.

Example 5.4.13

If $A = \begin{bmatrix} 1 & -2 & 1 & 1 \\ -1 & 2 & 0 & 1 \\ 2 & -4 & 1 & 0 \end{bmatrix}$, find bases of null (A) and im (A), and so find their dimensions.

If $x \in$ null A then $Ax = 0$ find x by solving system $Ax = 0$

$$\overset{A}{\begin{bmatrix} 1 & -2 & 1 & 1 \\ -1 & 2 & 0 & 1 \\ 2 & -4 & 1 & 0 \end{bmatrix}} \to \overset{\text{RE}}{\begin{bmatrix} 1 & -2 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}} \to \overset{\text{RRE}}{\begin{bmatrix} 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}}$$

rank $(A) = 2$

im $(A) = $ col(A) basis $\to \left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$ (leading 1's in cols 1,3)

$\dim(\text{im}(A)) = 2 = r$

let $x_2 = s$ $\quad x_4 = t$

$x_3 = -2t$ $\quad x_1 = 2s + t$

General Solution:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2s+t \\ s \\ -2t \\ t \end{bmatrix} = s \overset{x_1}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}} + t \overset{x_2}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}}$$

x_1 x_2 are basic solutions so

$$\text{null} A = \text{span}\{x_1, x_2\}$$

$\{x_1, x_2\}$ are basis null A

$$\dim(\text{null} A) = 2$$
$$= n - r = 4 - 2 = 2$$

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

- Let A be an $m \times n$ matrix.
- We have $\text{rank}A \leq m$ and $\text{rank}A \leq n \rightarrow$ when do these hold at equality?

- Let A be an $m \times n$ matrix.
- We have $\text{rank}A \leq m$ and $\text{rank}A \leq n \rightarrow$ when do these hold at equality?
- If $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ are the columns of A

$$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\} \text{ spans } \mathbb{R}^m \iff A\mathbf{x} = \mathbf{b} \text{ is consistent for every } \mathbf{b} \in \mathbb{R}^m$$

- Let A be an $m \times n$ matrix.
- We have $\text{rank}A \leq m$ and $\text{rank}A \leq n \rightarrow$ when do these hold at equality?
- If $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ are the columns of A

$$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\} \text{ spans } \mathbb{R}^m \iff A\mathbf{x} = \mathbf{b} \text{ is consistent for every } \mathbf{b} \in \mathbb{R}^m$$

- $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ is independent $\iff A\mathbf{x} = \mathbf{0}$, $\mathbf{x}$ in $\mathbb{R}^n$, implies $\mathbf{x} = \mathbf{0}$.

Theorem 5.4.14

Theorem

The following are equivalent for an $m \times n$ matrix A:

1. $\text{rank}A = n$.
2. The rows of A span $\mathbb{R}^n$.
3. The columns of A are linearly independent in $\mathbb{R}^m$.
4. The $n \times n$ matrix $A^T A$ is invertible.
5. $CA = I_n$ for some $n \times m$ matrix C.
6. If $A\mathbf{x} = \mathbf{0}$, $\mathbf{x}$ in $\mathbb{R}^n$, then $\mathbf{x} = \mathbf{0}$.

Theorem 5.4.15

Theorem

The following are equivalent for an $m \times n$ matrix A:

1. $\text{rank} A = m$.
2. The columns of A span $\mathbb{R}^m$.
3. The rows of A are linearly independent in $\mathbb{R}^n$.
4. The $m \times m$ matrix AA^T is invertible.
5. $AC = I_m$ for some $n \times m$ matrix C.
6. The system $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^m$.

Example 5.4.16

Show that $\begin{bmatrix} 3 & x+y+z \\ x+y+z & x^2+y^2+z^2 \end{bmatrix}$ is invertible if x, y, and z are not all equal.

Recap

Today we saw:

- Row and Column Spaces
- Rank Theorem and its Consequences

Next time: Similar Matrices

MATH254: Linear Algebra
Lecture 6

Moira MacNeil

January 17, 2025

Last Time

1. Rank of Matrix
2. Column and Row Spaces
3. Rank Theorem

Today

1. Rank of a Matrix continued
2. Similar Matrices

Reminders:

- Assignment 2 will be released by Monday, due in two weeks

One last result on rank

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

One last result on rank

- Let A be an $m \times n$ matrix
- We have $\text{rank}A \leq m$ and $\text{rank}A \leq n \rightarrow$ when do these hold at equality?

One last result on rank

- Let A be an $m \times n$ matrix
- We have $\text{rank}A \leq m$ and $\text{rank}A \leq n \rightarrow$ when do these hold at equality?
- If $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ are the columns of A

$$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\} \text{ spans } \mathbb{R}^m \iff A\mathbf{x} = \mathbf{b} \text{ is consistent for every } \underline{\mathbf{b} \in \mathbb{R}^m}$$

One last result on rank

- Let A be an $m \times n$ matrix
- We have $\text{rank}A \leq m$ and $\text{rank}A \leq n \to$ when do these hold at equality?
- If $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ are the columns of A

$$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\} \text{ spans } \mathbb{R}^m \iff A\mathbf{x} = \mathbf{b} \text{ is consistent for every } \mathbf{b} \in \mathbb{R}^m$$

- $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ is independent $\iff A\mathbf{x} = \mathbf{0}$, $\mathbf{x}$ in $\mathbb{R}^n$, implies $\underline{\mathbf{x} = \mathbf{0}}$

Theorem 5.4.14

Theorem

The following are equivalent for an $m \times n$ matrix A:

1. $\text{rank} A = n$.
2. The rows of A span $\mathbb{R}^n$.
3. The columns of A are linearly independent in $\mathbb{R}^m$.
4. The $n \times n$ matrix $A^T A$ is invertible.
5. $CA = I_n$ for some $n \times m$ matrix C.
6. If $A\mathbf{x} = \mathbf{0}$, $\mathbf{x}$ in $\mathbb{R}^n$, then $\mathbf{x} = \mathbf{0}$.

Theorem 5.4.15

Theorem

The following are equivalent for an $m \times n$ matrix A:

1. $\text{rank} A = m$.
2. The columns of A span $\mathbb{R}^m$.
3. The rows of A are linearly independent in $\mathbb{R}^n$.
4. The $m \times m$ matrix AA^T is invertible.
5. $AC = I_m$ for some $n \times m$ matrix C.
6. The system $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^m$.

Example 5.4.16

Show that $\begin{bmatrix} 3 & x+y+z \\ x+y+z & x^2+y^2+z^2 \end{bmatrix}$ is invertible if x, y, and z are not all equal.

Observe that the given matrix has form A^TA where $A = \begin{bmatrix} 1 & x \\ 1 & y \\ 1 & z \end{bmatrix}$

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ x & y & z \end{bmatrix} \begin{bmatrix} 1 & x \\ 1 & y \\ 1 & z \end{bmatrix} = \begin{bmatrix} 3 & x+y+z \\ x+y+z & x^2+y^2+z^2 \end{bmatrix}$$

A has independent columns

because x, y, z are not all equal
(no nontrivial linear comb. vanishes)
$\Rightarrow$ columns of A are lin. indep. in
$\mathbb{R}^3$ $\Rightarrow$ A^TA is invertible

Similar Matrices

Similar Matrices

If A and B are $n \times n$ matrices, we say that A and B are **similar**, and write $A \sim B$, if $B = P^{-1}AP$ for some invertible matrix P.

Properties of Similarity

2. $B = P^{-1} A P$
$A = P B P^{-1}$ let $Q = P^{-1}$
$= Q^{-1} B Q$

- Similarity relation $\sim$ is an **equivalence relation** on the set of $n \times n$ matrices

1. *$A \sim A$ for all square matrices A.*
2. *If $A \sim B$, then $B \sim A$.*
3. *If $A \sim B$ and $B \sim C$, then $A \sim C$.*

- Compatible with inverses, transposes, and powers:

If $A \sim B$ then $\underline{A^{-1} \sim B^{-1}}$, $\underline{A^T \sim B^T}$, and $\underline{A^k \sim B^k}$ for all integers $k \geq 1$.

Example 5.5.2

If A is similar to B and either A or B is diagonalizable, show that the other is also diagonalizable.

$A \sim B$, assume A is diagonalizable,

say $A \sim D$ where D is diagonal

Since $B \sim A$ (by $\sim$ equiv #2) then $B \sim D$ (by #3)

$(B \sim A \quad A \sim D \rightarrow B \sim D)$

Therefore B is also diagonalizable.

Similar argument applies if we assume B is diag.

Trace

Trace of a Matrix

The **trace** tr$A$ of an $n \times n$ matrix A is defined to be the sum of the main diagonal elements of A.

$$\text{If } A = \left[a_{ij}\right], \text{ then tr } A = a_{11} + a_{22} + \cdots + a_{nn}.$$

Lemma 5.5.4

$\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$

$\mathrm{tr}(kA) = k\,\mathrm{tr}(A)$

Lemma

Let A and B be $n \times n$ matrices. Then $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Write $A = [a_{ij}]$ $B = [b_{ij}]$. For each i let

(i,i)-entry $= d_i$ $\to$ of matrix AB

$$d_i = a_{i1}b_{1i} + a_{i2}b_{2i} + \cdots + a_{in}b_{ni} = \sum_j a_{ij}b_{ji}$$

$$\mathrm{tr}(AB) = d_1 + d_2 + \cdots + d_n = \sum_i d_i = \sum_i \left(\sum_j a_{ij} b_{ji} \right)$$

$$\mathrm{tr}(BA) = \sum_i d_i = \sum_i \left(\sum_j b_{ij} a_{ji} \right) =$$

Theorem 5.5.5

Theorem

If A and B are similar $n \times n$ matrices, then A and B have the same determinant, rank, trace, characteristic polynomial, and eigenvalues.

PROOF Let $B = P^{-1}AP$ for some invert. P

Then $\det(B) = \det(P^{-1})\det(A)\det(P)$

$= \det(A)$

because $\det(P^{-1}) = \dfrac{1}{\det(P)}$

$\text{rank}(B) = \text{rank}(P^{-1}AP) = \text{rank}(A)$

$tr(B) =$

$tr(P^{-1}AP) = tr(P^{-1}[AP]) = tr(APP^{-1})$

$= tr(A)$

$c_B(x) = det(xI - B) = det[x(P^{-1}IP) - P^{-1}AP]$

$= det[P^{-1}(xI - A)P]$

$= det(xI - A)$

$= c_A(x)$

A and B have same eigenvalues

→ eigenvalues are roots of char. polynomial

same char. poly → same roots

□

Example 5.5.6

Are the matrices $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ similar?

$\operatorname{tr} A = \operatorname{tr} I = 2$

$\det A = \det I = 1$

$\operatorname{rank} A = \operatorname{rank} I = 2$

$$c_A(x) = \det(xI - A) = \begin{vmatrix} x-1 & 1 \\ 0 & x-1 \end{vmatrix} = (x-1)^2$$

Not similar

$P^{-1} I P = I \neq A$

for any invertible P

$c_I(x) = (x-1)^2$

eigenvalues $\lambda = 1$

Recall that a square matrix A is **diagonalizable** if there exists an invertible matrix P such that $P^{-1}AP = D$ is a diagonal matrix, that is if A is similar to a diagonal matrix D. Unfortunately, not all matrices are diagonalizable. Determining whether A is diagonalizable is closely related to the eigenvalues and eigenvectors of A. Recall that a number λ is called an **eigenvalue** of A if $A\mathbf{x} = \lambda\mathbf{x}$ for some nonzero column $\mathbf{x}$ in $\mathbb{R}^n$, and any such nonzero vector $\mathbf{x}$ is called an **eigenvector** of A corresponding to λ (or simply a λ-eigenvector of A). The eigenvalues and eigenvectors of A are closely related to the **characteristic polynomial** $c_A(x)$ of A, defined by

$$c_A(x) = \det(xI - A)$$

Theorem 5.5.7

Theorem

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

1. The eigenvalues λ of A are the roots of the characteristic polynomial $c_A(x)$ of A.
2. The λ-eigenvectors $\mathbf{x}$ are the nonzero solutions to the homogeneous system

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

of linear equations with $\lambda I - A$ as coefficient matrix.

Example 5.5.8

Show that the eigenvalues of a triangular matrix are the main diagonal entries.

Triangular matrix

Assume A is triangular $\begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix}$

$(xI - A)$ is also triangular and has diag.

$(x - a_{11}), (x - a_{22}), \dots, (x - a_{nn})$

det of a triangular matrix $\to$ product of diagonal entries.

$c_A(x) = (x - a_{11}) \cdots (x - a_{nn})$

result follows from eigenvalues roots of $c_A(x)$

Recap

Today we saw:

- Similar matrices
- Refresher on eigenvalues and eigenvectors

Next time: Diagonalization

MATH254: Linear Algebra
Lecture 7

Moira MacNeil

January 21, 2025

Last Time

$B \sim A$

$A = P^{-1} B P$

1. Similar matrices
2. Refresher on eigenvalues and eigenvectors

Today

1. Diagonalization
2. Complex Eigenvalues

Reminders:

- Assignment 2 is out, due January 31

Recall: Theorem 5.5.7

Theorem

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

1. The eigenvalues λ of A are the roots of the characteristic polynomial $c_A(x) = \det(Ix - A)$ of A.
2. The λ-eigenvectors $\mathbf{x}$ are the nonzero solutions to the homogeneous system

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

of linear equations with $\lambda I - A$ as coefficient matrix.

Theorem 5.5.9

$A = P^{-1} D P$ $\qquad$ $A^k = P^{-1} D^k P$

Theorem

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

1. A is diagonalizable if and only if $\mathbb{R}^n$ has a basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ consisting of eigenvectors of A.
2. When this is the case, the matrix $P = \left[\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{array}\right]$ is invertible and $P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where, for each i, λ_i is the eigenvalue of A corresponding to $\mathbf{x}_i$.

eigenvectors

Theorem 5.5.10

Theorem

Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ be eigenvectors corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ of an $n \times n$ matrix A. Then $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is a linearly independent set.

Eigenvectors corresponding to distinct eigenvalues are necessarily lin. indep.

Theorem 5.5.11

Theorem

If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable.

Choose an eigenvector for each of n eigenvalues then eigenvectors are lin. indep. and are a basis of $\mathbb{R}^n$ then A is diagonalizable.

Example 5.5.12

Show that $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 3 \\ -1 & 1 & 0 \end{bmatrix}$ is diagonalizable.

$$c_A(x) = \det(xI - A) = (x-1)\left[(x-2)(x) - 3\right]$$
$$= (x-1)(x-3)(x+1)$$

3 distinct eigenvalues $= 1, 3, -1$

A is diag.

Lemma 5.5.13

Lemma

Let $\{\overset{\text{eigenvect.}}{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k}\}$ be a linearly independent set of eigenvectors of an $n \times n$ matrix A, extend it to a basis $\{\underbrace{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k}_{\text{eig.}}, \ldots, \mathbf{x}_n\}$ of $\mathbb{R}^n$, and let

$$P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$$

be the (invertible) $n \times n$ matrix with the $\mathbf{x}_i$ as its columns. If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the (not necessarily distinct) eigenvalues of A corresponding to $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ respectively, then $P^{-1}AP$ has block form

$$P^{-1}AP = \begin{bmatrix} \boxed{\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)} & B \\ 0 & A_1 \end{bmatrix}$$

where B has size $k \times (n-k)$ and A_1 has size $(n-k) \times (n-k)$.

Eigenspace

Eigenspace of a Matrix

If λ is an eigenvalue of an $n \times n$ matrix A, define the **eigenspace** of A corresponding to λ by

$$E_\lambda(A) = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$$

Eigenspace

Eigenspace of a Matrix

If λ is an eigenvalue of an $n \times n$ matrix A, define the **eigenspace** of A corresponding to λ by

$$E_\lambda(A) = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$$

- Subspace of $\mathbb{R}^n$, eigenvectors corresponding to $\lambda \to$ nonzero vectors in $E_\lambda(A)$

Eigenspace

Eigenspace of a Matrix

If λ is an eigenvalue of an $n \times n$ matrix A, define the **eigenspace** of A corresponding to λ by

$$E_\lambda(A) = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$$

- Subspace of $\mathbb{R}^n$, eigenvectors corresponding to $\lambda \rightarrow$ nonzero vectors in $E_\lambda(A)$
- $E_\lambda(A)$ is the null space of the matrix $(\lambda I - A)$:

$$E_\lambda(A) = \{\mathbf{x} \mid (\lambda I - A)\mathbf{x} = \mathbf{0}\} = \text{null } (\lambda I - A)$$

Eigenspace

Eigenspace of a Matrix

If λ is an eigenvalue of an $n \times n$ matrix A, define the **eigenspace** of A corresponding to λ by

$$E_\lambda(A) = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$$

- Subspace of $\mathbb{R}^n$, eigenvectors corresponding to $\lambda \to$ nonzero vectors in $E_\lambda(A)$
- $E_\lambda(A)$ is the null space of the matrix $(\lambda I - A)$:

$$E_\lambda(A) = \{\mathbf{x} \mid (\lambda I - A)\mathbf{x} = \mathbf{0}\} = \text{null } (\lambda I - A)$$

- Basic solutions of $(\lambda I - A)\mathbf{x} = \mathbf{0}$ given by the gaussian algorithm form a basis for $E_\lambda(A)$

Eigenspace

Eigenspace of a Matrix

If λ is an eigenvalue of an $n \times n$ matrix A, define the **eigenspace** of A corresponding to λ by

$$E_\lambda(A) = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$$

- Subspace of $\mathbb{R}^n$, eigenvectors corresponding to $\lambda \rightarrow$ nonzero vectors in $E_\lambda(A)$
- $E_\lambda(A)$ is the null space of the matrix $(\lambda I - A)$:

$$E_\lambda(A) = \{\mathbf{x} \mid (\lambda I - A)\mathbf{x} = \mathbf{0}\} = \text{null } (\lambda I - A)$$

- Basic solutions of $(\lambda I - A)\mathbf{x} = \mathbf{0}$ given by the gaussian algorithm form a basis for $E_\lambda(A)$
- dim $E_\lambda(A)$ is the number of basic solutions $\mathbf{x}$ of $(\lambda I - A)\mathbf{x} = \mathbf{0}$

Multiplicity

- **Multiplicity** of an eigenvalue λ of $A \rightarrow$ number of times λ occurs as a root of the characteristic polynomial $c_A(x)$

Multiplicity

- **Multiplicity** of an eigenvalue λ of $A \to$ number of times λ occurs as a root of the characteristic polynomial $c_A(x)$

- Another way $\to$ multiplicity of λ is the largest integer $m \geq 1$ such that

$$c_A(x) = \underbrace{(x - \lambda)^m}g(x)$$

for some polynomial $g(x)$

Lemma 5.5.15

Lemma

Let λ be an eigenvalue of multiplicity m of a square matrix A. Then $\dim[E_\lambda(A)] \leq m$.

$\dim[E_\lambda(A)] = d$ WTS $c_A(x) = (x-\lambda)^d g(x)$ since m is highest power of $(x-\lambda)$ that divides $c_A(x)$.

Let $\{x_1, \dots, x_d\}$ be a basis of $E_\lambda(A)$, we know exists $n\times n$ P (invert.) s.t.

$$\underset{\hookrightarrow n\times n}{A'} = P^{-1}AP = \begin{bmatrix} \lambda I_d & B \\ 0 & A_1 \end{bmatrix} \begin{matrix} \} d \text{ rows} \\ \} n-d \text{ rows} \end{matrix}$$

$I_d \to n\times n$ identity

$$c_{A'}(x) = c_A(x) \qquad (A' \sim A)$$

$$c_A(x) = c_{A'}(x) = \det(xI_n - A') = \begin{vmatrix} (x-\lambda)I_d & -B \\ 0 & xI_{n-d} - A_1 \end{vmatrix}$$

$$= \det[(x-\lambda)I_d]\,\underbrace{\det[(xI_{n-d} - A_1)]}.$$

$$= (x-\lambda)^d \; g(x)$$

$$\square$$

When does $\dim[E_\lambda(A)] = m$ for each eigenvalue λ?

↳ multiplicity

- $\dim[E_\lambda(A)] = m \rightarrow$ characterizes the diagonalizable $n \times n$ matrices A for which $c_A(x)$ **factors completely** over $\mathbb{R}$
- That is $c_A(x) = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$, where the λ_i are *real* numbers (not necessarily distinct)
- In other words $\rightarrow$ every eigenvalue of A is real

Theorem 5.5.16

Theorem

The following are equivalent for a square matrix A for which $c_A(x)$ factors completely.

1. A is diagonalizable.
2. $\dim[E_\lambda(A)]$ equals the multiplicity of λ for every eigenvalue λ of the matrix A.

Example 5.5.17

If $A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & -2 \end{bmatrix}$ show that A is diagonalizable but B is not.

$c_A(x) = (x+3)^2(x-1) \rightarrow \lambda_1 = -3 \quad \lambda_2 = 1$

$\lambda_1 \rightarrow \quad x_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad x_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \qquad E_{\lambda_1}(A) = \operatorname{span}\{x_1, x_2\}$

$\lambda_2 \rightarrow \quad x_3 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \qquad E_{\lambda_2}(A) = \operatorname{span}\{x_3\}$

Clearly $\{x_1, x_2\}$ is independent

$\dim(E_{\lambda_1}(A)) = 2 \rightarrow$ multiplicity of λ_1

$\dim(E_{\lambda_2}(A)) = 1 \rightarrow$ mutl. of λ_2

$\Rightarrow A$ diag. with $P = [\, x_1 \quad x_2 \quad x_3 \,]$

$c_B(x) = (x+1)^2(x-3) \rightarrow \lambda_1 = -1 \quad \lambda_2 = 3$

$\lambda_1 \rightarrow y_1 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \qquad \lambda_2 \rightarrow y_2 = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix}$

$E_{\lambda_1}(B) = \text{span}\{y_1\} \qquad E_{\lambda_2}(B) = \text{span}\{y_2\}$

$\dim(E_{\lambda_1}(B)) = 1 <$ mult. $\Rightarrow$ B is <u>not</u> diag.

Complex Eigenvalues

- Eigenvalues need not be real numbers

Complex Eigenvalues

- Eigenvalues need not be real numbers
- $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has characteristic polynomial $c_A(x) = x^2 + 1$ which has no real roots

Complex Eigenvalues

- Eigenvalues need not be real numbers
- $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has characteristic polynomial $c_A(x) = x^2 + 1$ which has no real roots
- A is still diagonalizable $\rightarrow$ we need to use complex numbers $(a + bi)$ $a, b \in \mathbb{R}$

Complex Eigenvalues

- Eigenvalues need not be real numbers
- $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has characteristic polynomial $c_A(x) = x^2 + 1$ which has no real roots
- A is still diagonalizable $\rightarrow$ we need to use complex numbers $(a + bi)$
- Fundamental Theorem of Algebra:
  Every nonconstant polynomial with complex coefficients has a complex root, and hence factors completely as a product of linear factors

Example 5.5.18

Diagonalize the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

$$c_A(x) = \det(xI - A) = x^2 + 1$$
$$= (x - i)(x + i)$$
(where $i^2 = -1$)

$\lambda_1 = i \qquad \lambda_2 = -i$

$(iI - A)x = 0$

$$\begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} \Rightarrow x_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

A is diag. (2 dis. λ)

$$P = [x_1 \ \ x_2] = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}$$

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$
$$= \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

$$\lambda_2 \to \quad x_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

Symmetric Matrices

Recall a square matrix A is symmetric if $A = A^\top$

Theorem

Any eigenvalue of a symmetric real matrix is real.

Example

Verify that every 2×2 symmetric matrix has real eigenvalues.

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

Recap

Today we saw:

- Diagonalization
- Complex Eigenvalues

Next time: Vector Spaces

MATH254: Linear Algebra
Lecture 8

Moira MacNeil

January 22, 2025

Last Time

1. Diagonalization
2. Complex Eigenvalues

Today

1. Extending vector spaces beyond $\mathbb{R}^n$

Reminder:

- Assignment 2 is due January 31

Vector Spaces

Vector Spaces

A **vector space** consists of a nonempty set V of objects (called **vectors**) that can be added, that can be multiplied by a real number (called a **scalar** in this context), and for which certain axioms hold. If $\mathbf{v}$ and $\mathbf{w}$ are two vectors in V, their sum is expressed as $\mathbf{v} + \mathbf{w}$, and the scalar product of $\mathbf{v}$ by a real number a is denoted as $a\mathbf{v}$. These operations are called **vector addition** and **scalar multiplication**, respectively.

Axioms for vector addition

A1. If $\mathbf{u}$ and $\mathbf{v}$ are in V, then $\mathbf{u}+\mathbf{v}$ is in V. → closed under addition

A2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ for all $\mathbf{u}$ and $\mathbf{v}$ in V.

A3. $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$ for all $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V.

A4. An element $\mathbf{0}$ in V exists such that $\underline{\mathbf{v}+\mathbf{0}=\mathbf{v}=\mathbf{0}+\mathbf{v}}$ for every $\mathbf{v}$ in V.

A5. For each $\mathbf{v}$ in V, an element $-\mathbf{v}$ in V exists such that $\underline{-\mathbf{v}+\mathbf{v}=\mathbf{0} \text{ and } \mathbf{v}+(-\mathbf{v})=\mathbf{0}}$.

Axioms for scalar multiplication

S1. If $\mathbf{v}$ is in V, then $a\mathbf{v}$ is in V for all a in $\mathbb{R}$.

S2. $a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w}$ for all $\mathbf{v}$ and $\mathbf{w}$ in V and all a in $\mathbb{R}$.

S3. $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$ for all $\mathbf{v}$ in V and all a and b in $\mathbb{R}$.

S4. $a(b\mathbf{v}) = (ab)\mathbf{v}$ for all $\mathbf{v}$ in V and all a and b in $\mathbb{R}$.

S5. $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v}$ in V.

Vector Spaces

- Axiom A1 + Axiom S1 $\rightarrow$ V is **closed** under vector addition and scalar multiplication

Vector Spaces

- Axiom A1 $+$ Axiom S1 $\rightarrow$ V is **closed** under vector addition and scalar multiplication
- $\mathbf{0}$ (in axiom A4) $\rightarrow$ **zero vector**

Vector Spaces

- Axiom A1 + Axiom S1 $\rightarrow$ V is **closed** under vector addition and scalar multiplication
- $\mathbf{0}$ (in axiom A4) $\rightarrow$ **zero vector**
- $-\mathbf{v}$ (in axiom A5) $\rightarrow$ **negative** of $\mathbf{v}$.

Vector Spaces

- Axiom A1 + Axiom S1 $\to$ V is **closed** under vector addition and scalar multiplication
- $\mathbf{0}$ (in axiom A4) $\to$ **zero vector**
- $-\mathbf{v}$ (in axiom A5) $\to$ **negative** of $\mathbf{v}$.
- **Example**:
  $\mathbb{R}^n$ is a vector space using matrix addition and scalar multiplication. → SATISFIES ALL AXIOMS A1-A5 S1-S5

Vector Spaces

- Axiom A1 $+$ Axiom S1 $\rightarrow$ V is **closed** under vector addition and scalar multiplication
- $\mathbf{0}$ (in axiom A4) $\rightarrow$ **zero vector**
- $-\mathbf{v}$ (in axiom A5) $\rightarrow$ **negative** of $\mathbf{v}$.
- **Example**:
  $\mathbb{R}^n$ is a vector space using matrix addition and scalar multiplication.
- In a general vector space, vectors can be **any** kind of objects $\rightarrow$ need to define addition and scalar multiplication and satisfy axioms

Example 6.1.3

Show the set $\mathbf{M}_{mn}$ of all $m \times n$ matrices is a vector space using matrix addition and scalar multiplication.

Zero element: matrix $(m \times n)$ of zeros

negative: usual matrix negative

M_{mn} is just $\mathbb{R}^{mn}$

Example 6.1.4

Show that every subspace of $\mathbb{R}^n$ is a vector space in its own right using the addition and scalar multiplication of $\mathbb{R}^n$.

Axioms S1, A1 (closed under add. + scalar mult) are defining for a set $U \subset \mathbb{R}^n$ to be a subspace. Remaining axioms inherited from $\mathbb{R}^n$.

Eg. $x, y \in U$ and $a \in \mathbb{R}$ then

$a(x+y) = ax + ay$ since $x, y \in \mathbb{R}^n$

so S2 holds

Example 6.1.5

$(a,b)+(x,y)$
$=(a+x,\ b+y)$

Let V denote the set of all ordered pairs (x, y) and define addition in V as in $\mathbb{R}^2$. However, define a new scalar multiplication in V by

$$a(x, y) = (\underline{ay}, \underline{ax})$$

Determine if V is a vector space with these operations.

Addition axioms A1 - A5 hold for V since they hold for matrices.

S1: $a(x,y) = (ay, ax) \in \mathbb{R}^2$ ✓

S2: $v = (x,y)$ $w = (x_1, y_1)$ elements of V

$a(v+w) = a(x+x_1, y+y_1) = (a(y+y_1), a(x+x_1))$ ✓

$av + aw = (ay, ax) + (ay_1, ax_1) = (a(y+y_1), a(x+x_1))$

S3: ✓ (exercise)

S4: $a, b \in \mathbb{R}$ $v = (x,y)$

$$a(bv) = a(b(x,y)) = a(by, bx)$$
$$= (abx, aby) \neq$$

$$(ab)v = (ab)(x,y) = (aby, abx)$$

So V is not a vector space

S5: ✗

A Brief Review of Polynomials

- A **polynomial** in an indeterminate x is an expression

$$p(x) = \underline{a_0} + \underline{a_1} x + \underline{a_2} x^2 + \cdots + \underline{a_n} x^n$$

where $a_0, a_1, a_2, \ldots, a_n$ are real numbers called the **coefficients** of the polynomial

A Brief Review of Polynomials

- A **polynomial** in an indeterminate x is an expression

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

  where $a_0, a_1, a_2, \ldots, a_n$ are real numbers called the **coefficients** of the polynomial

- **zero polynomial**, $0 \to$ all coefficients are zero

A Brief Review of Polynomials

- A **polynomial** in an indeterminate x is an expression

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

  where $a_0, a_1, a_2, \ldots, a_n$ are real numbers called the **coefficients** of the polynomial

- **zero polynomial**, $0 \rightarrow$ all coefficients are zero
- **degree** of $p(x)$, deg $p(x) \rightarrow$ highest power of x with a nonzero coefficient (if $p(x) \neq 0$)

A Brief Review of Polynomials

- A **polynomial** in an indeterminate x is an expression

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

  where $a_0, a_1, a_2, \ldots, a_n$ are real numbers called the **coefficients** of the polynomial

- **zero polynomial**, $0 \rightarrow$ all coefficients are zero
- **degree** of $p(x)$, deg $p(x) \rightarrow$ highest power of x with a nonzero coefficient (if $p(x) \neq 0$)
- **leading coefficient** of $p(x) \rightarrow$ coefficient of highest power of x

A Brief Review of Polynomials

- A **polynomial** in an indeterminate x is an expression

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

  where $a_0, a_1, a_2, \ldots, a_n$ are real numbers called the **coefficients** of the polynomial
- **zero polynomial**, $0 \rightarrow$ all coefficients are zero
- **degree** of $p(x)$, deg $p(x) \rightarrow$ highest power of x with a nonzero coefficient (if $p(x) \neq 0$)
- **leading coefficient** of $p(x) \rightarrow$ coefficient of highest power of x
- Note: degree of the zero polynomial $\rightarrow$ not defined

Addition and Scalar Multiplication in Polynomials

- $\mathbf{P} \rightarrow$ set of all polynomials

Addition and Scalar Multiplication in Polynomials

- $\mathbf{P} \rightarrow$ set of all polynomials

- Take two polynomials in $\mathbf{P}$ (possibly of different degrees):

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots \quad a_n x^n$$
$$q(x) = b_0 + b_1 x + b_2 x^2 + \cdots \quad a_m x^m$$

$n < m$

Addition and Scalar Multiplication in Polynomials

- $\mathbf{P} \rightarrow$ set of all polynomials
- Take two polynomials in $\mathbf{P}$ (possibly of different degrees):

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots$$
$$q(x) = b_0 + b_1x + b_2x^2 + \cdots$$

- $p(x) = q(x) \iff$ coefficients are equal $a_0 = b_0,\ a_1 = b_1,\ a_2 = b_2,\ \cdots$

Addition and Scalar Multiplication in Polynomials

- $\mathbf{P} \to$ set of all polynomials
- Take two polynomials in $\mathbf{P}$ (possibly of different degrees):

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots$$
$$q(x) = b_0 + b_1x + b_2x^2 + \cdots$$

- $p(x) = q(x) \iff$ coefficients are equal $a_0 = b_0,\ a_1 = b_1,\ a_2 = b_2, \cdots$
- $a_0 + a_1x + a_2x^2 + \cdots = 0$ means that $a_0 = 0,\ a_1 = 0,\ a_2 = 0, \ldots$

Addition and Scalar Multiplication in Polynomials

- $\mathbf{P} \to$ set of all polynomials
- Take two polynomials in $\mathbf{P}$ (possibly of different degrees):

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots$$
$$q(x) = b_0 + b_1x + b_2x^2 + \cdots$$

- $p(x) = q(x) \iff$ coefficients are equal $a_0 = b_0,\ a_1 = b_1,\ a_2 = b_2, \cdots$
- $a_0 + a_1x + a_2x^2 + \cdots = 0$ means that $a_0 = 0,\ a_1 = 0,\ a_2 = 0, \ldots$
- Addition on $\mathbf{P}$: $p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots$

Addition and Scalar Multiplication in Polynomials

- $\mathbf{P} \to$ set of all polynomials
- Take two polynomials in $\mathbf{P}$ (possibly of different degrees):

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots$$
$$q(x) = b_0 + b_1x + b_2x^2 + \cdots$$

- $p(x) = q(x) \iff$ coefficients are equal $a_0 = b_0,\ a_1 = b_1,\ a_2 = b_2, \cdots$
- $a_0 + a_1x + a_2x^2 + \cdots = 0$ means that $a_0 = 0,\ a_1 = 0,\ a_2 = 0, \ldots$
- Addition on $\mathbf{P}$: $p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots$
- Scalar multiplication on $\mathbf{P}$: $ap(x) = aa_0 + (aa_1)x + (aa_2)x^2 + \cdots$

Addition and Scalar Multiplication in Polynomials

- $\mathbf{P} \rightarrow$ set of all polynomials
- Take two polynomials in $\mathbf{P}$ (possibly of different degrees):

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots$$
$$q(x) = b_0 + b_1x + b_2x^2 + \cdots$$

- $p(x) = q(x) \iff$ coefficients are equal $a_0 = b_0,\ a_1 = b_1,\ a_2 = b_2, \cdots$
- $a_0 + a_1x + a_2x^2 + \cdots = 0$ means that $a_0 = 0,\ a_1 = 0,\ a_2 = 0, \ldots$
- Addition on $\mathbf{P}$: $p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots$
- Scalar multiplication on $\mathbf{P}$: $ap(x) = aa_0 + (aa_1)x + (aa_2)x^2 + \cdots$
- These operations result in polynomials $\rightarrow$ $\mathbf{P}$ is closed under addition and scalar multiplication, called **pointwise** addition and scalar multiplication

Example 6.1.6

The set **P** of all polynomials is a vector space with addition and scalar multiplication as described previously. What is the zero vector and the negative vector in this vector space?

zero vector $\rightarrow$ zero polynomial

negative of $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots$

$-p(x) = -a_0 - a_1 x_1 - a_2 x^2 - \cdots$

(negate each coeff.)

Example 6.1.7

Given $n \geq 1$, let $\mathbf{P}_n$ denote the set of all polynomials of degree at most n, together with the zero polynomial. That is

$$\mathbf{P}_n = \{\underline{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n} \mid a_0, a_1, a_2, \ldots, a_n \text{ in } \mathbb{R}\}.$$

Then $\mathbf{P}_n$ is a vector space.

Sums and scalar multiples of $p(x)$ in P_n
are also in P_n → closed
Inherit other vector axioms from P
↳ zero vector
↳ negative
} same as in P

A Brief Review of Functions

- **interval**, $[a, b]$ $(a, b \in \mathbb{R},\ a < b) \rightarrow$ set of all real numbers x such that $a \leq x \leq b$

A Brief Review of Functions

- **interval**, $[a, b]$ $(a, b \in \mathbb{R},\ a < b) \to$ set of all real numbers x such that $a \leq x \leq b$

$x \to f(x)$

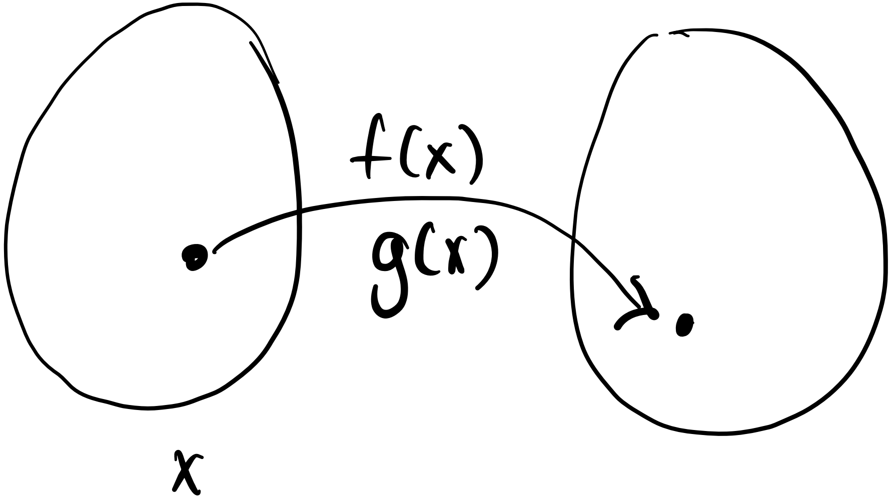
- (real-valued) **function** f on $[a, b] \to$ associates $f(x) \in \mathbb{R}$ to to every number x in $[a, b]$

A Brief Review of Functions

- **interval**, $[a, b]$ ($a, b \in \mathbb{R}$, $a < b$) $\rightarrow$ set of all real numbers x such that $a \leq x \leq b$
- (real-valued) **function** f on $[a, b]$ $\rightarrow$ associates $f(x) \in \mathbb{R}$ to to every number x in $[a, b]$
- $\mathbf{F}[a, b]$ $\rightarrow$ set of all functions on $[a, b]$

A Brief Review of Functions

- **interval**, $[a, b]$ $(a, b \in \mathbb{R},\ a < b) \rightarrow$ set of all real numbers x such that $a \leq x \leq b$
- (real-valued) **function** f on $[a, b] \rightarrow$ associates $f(x) \in \mathbb{R}$ to to every number x in $[a, b]$
- $\mathbf{F}[a, b] \rightarrow$ set of all functions on $[a, b]$
- f equals $g \iff f(x) = g(x)$ for every x in $[a, b]$ (we say f and g have the **same action**)



A Brief Review of Functions

- **interval**, $[a, b]$ ($a, b \in \mathbb{R}$, $a < b$) $\rightarrow$ set of all real numbers x such that $a \le x \le b$
- (real-valued) **function** f on $[a, b]$ $\rightarrow$ associates $f(x) \in \mathbb{R}$ to to every number x in $[a, b]$
- $\mathbf{F}[a, b]$ $\rightarrow$ set of all functions on $[a, b]$
- f equals $g \iff f(x) = g(x)$ for every x in $[a, b]$ (we say f and g have the **same action**)
- Addition and scalar multiplication in $\mathbf{F}[a, b]$:

$$\begin{aligned} (f+g)(x) &= \underline{f(x) + g(x)} && \text{for each } x \text{ in } [a, b] \\ (rf)(x) &= rf(x) && \text{for each } x \text{ in } [a, b] \end{aligned}$$

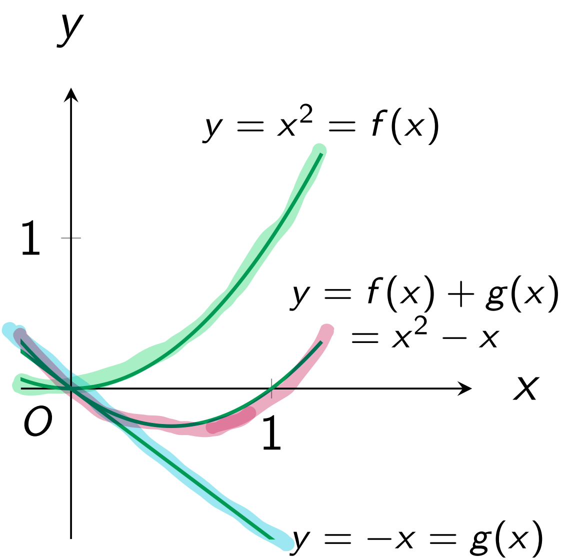
A Brief Review of Functions

- **interval**, $[a, b]$ $(a, b \in \mathbb{R},\ a < b) \to$ set of all real numbers x such that $a \leq x \leq b$
- (real-valued) **function** f on $[a, b] \to$ associates $f(x) \in \mathbb{R}$ to to every number x in $[a, b]$
- $\mathbf{F}[a, b] \to$ set of all functions on $[a, b]$
- f equals $g \iff f(x) = g(x)$ for every x in $[a, b]$ (we say f and g have the **same action**)
- Addition and scalar multiplication in $\mathbf{F}[a, b]$:

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \quad && \text{for each } x \text{ in } [a, b] \\ (rf)(x) &= rf(x) && \text{for each } x \text{ in } [a, b] \end{aligned}$$

- These operations on $\mathbf{F}[a, b] \to$ **pointwise addition and scalar multiplication** of functions

Example: Sum of $f(x) = x^2$ and $g(x) = -x$



$$(f+g)(x) = f(x)+g(x) = x^2 - x$$

Example 6.1.8

The set $\mathbf{F}[a, b]$ of all functions on the interval $[a, b]$ is a vector space using pointwise addition and scalar multiplication. What is the zero function and the negative function in this vector space?

zero function $\rightarrow$ 0
constant function

$0(x) = 0$ for each x in $[a,b]$

negative of f $-f$

$(-f)(x) = -f(x)$ for each x in $[a,b]$

A1, S1 satisfied, others are too (exercise)

Theorem 6.1.9

Cancellation

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in a vector space V. If $\mathbf{v}+\mathbf{u}=\mathbf{v}+\mathbf{w}$, then $\mathbf{u}=\mathbf{w}$.

PROOF

Given $\quad v+u = v+w$

$$-v+(v+u) = -v+(v+w) \qquad (A5)$$

$$(-v+v)+u = (-v+v)+w \qquad (A3)$$

$$0+u = 0+w \qquad (A5)$$

$$u = w \qquad (A4)$$

$\square$

Theorem 6.1.10

Theorem

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a vector space V, the equation $\mathbf{x} + \mathbf{v} = \mathbf{u}$ has one and only one solution $\mathbf{x}$ in V given by $\mathbf{x} = \mathbf{u} - \mathbf{v}$.

let $x = u - v$

$$
\begin{aligned}
x + v = (u - v) + v &= [u + (-v)] + v \\
&= u + ((-v) + v) \\
&= u + 0 \\
&= u
\end{aligned}
$$

$\therefore$ $x = u - v$ is a solution to the equation suppose x_1 is another sol to eq

$x_1 + v = u$

then $x + v = x_1 + v$

by cancellation $x = x_1$

$\therefore$ $x = u - v$ is the only sol. $\square$

Recap

Today we saw:

- Axioms of vector spaces
- Examples of vector spaces

Next time: Extend subspaces and spanning sets

MATH254: Linear Algebra
Lecture 9

Moira MacNeil

January 24, 2025

Last Time

1. Axioms of vector spaces
2. Examples of vector spaces

Today

1. A few more properties of vector spaces
2. Extending subspaces and spanning sets beyond $\mathbb{R}^n$

Reminder:

- Assignment 2 is due January 31

Theorem 6.1.11

A3: $u+(w+v)=(u+w)+v$

$u+w+v$

$v_1+v_2+\cdots+v_n$

Theorem

Let $\mathbf{v}$ denote a vector in a vector space V and let a denote a real number.

1. $0\mathbf{v}=\mathbf{0}$.
2. $a\mathbf{0}=\mathbf{0}$.
3. If $a\mathbf{v}=\mathbf{0}$, then either $a=0$ or $\mathbf{v}=\mathbf{0}$.
4. $(-1)\mathbf{v}=-\mathbf{v}$.
5. $(-a)\mathbf{v}=-(a\mathbf{v})=a(-\mathbf{v})$.

These properties hold in every vector space!

PROOFS USE AXIOMS, CANCEL. + SUBTRACTION

Example 6.1.12

If **u**, **v**, and **w** are vectors in a vector space V, simplify the expression

$$2(\mathbf{u}+3\mathbf{w})-3(2\mathbf{w}-\mathbf{v})-3[2(2\mathbf{u}+\mathbf{v}-4\mathbf{w})-4(\mathbf{u}-2\mathbf{w})]$$

$$= 2u + \cancel{6w} - \cancel{6w} + 3v - 3[\cancel{4u} + 2v - \cancel{8w} - \cancel{4u} + \cancel{8w}]$$

$$= 2u + 3v - 6v$$

$$= 2u - 3v$$

* treat arb. vector space like one we know

proceed as if vectors or variables

Zero Vector Space

A set $\{\mathbf{0}\}$ with one element becomes a vector space if we define

closed under addition

$$\mathbf{0} + \mathbf{0} = \mathbf{0} \quad \text{and} \quad a\mathbf{0} = \mathbf{0} \text{ for all scalars } a.$$

closed under scalar mult.

The resulting space is called the **zero vector space** and is denoted $\{\mathbf{0}\}$.

Subspaces

Subspaces of a Vector Space

If V is a vector space, a nonempty subset $U \subseteq V$ is called a **subspace** of V if U is itself a vector space using the addition and scalar multiplication of V

Theorem 6.2.2

$U \subseteq V$

Subspace Test

A subset U of a vector space is a subspace of V if and only if it satisfies the following three conditions:

1. $\mathbf{0}$ lies in U where $\mathbf{0}$ is the zero vector of V.
2. If $\mathbf{u}_1$ and $\mathbf{u}_2$ are in U, then $\mathbf{u}_1 + \mathbf{u}_2$ is also in U.
3. If $\mathbf{u}$ is in U, then $a\mathbf{u}$ is also in U for each scalar a.

closed under add & scalar mult.

Example 6.2.3

If V is any vector space, show that $\{\mathbf{0}\}$ and V are subspaces of V.

$U = \{0\}$

subspace Test

1 ✓

2 $0 + 0 = 0$ ✓

3 $a0 = 0$ ✓

$U = V$

1 $0 \in V$ o/w not vector space

2 $u + v \in V$

3 $av \in V$

Example 6.2.4

Let $\mathbf{v}$ be a vector in a vector space V. Show that the set

$$\mathbb{R}\mathbf{v} = \{a\mathbf{v} \mid a \text{ in } \mathbb{R}\}$$

of all scalar multiples of $\mathbf{v}$ is a subspace of V.

$d = 0$ in $\mathbb{R}^3$

$\mathbb{R}d$ is line through origin w/ direction vector d

Subspace Test

1. $0 = 0v$ $\quad 0 \in \mathbb{R}v$
2. Given av $\quad a_1 v$ in $\mathbb{R}v$

   $(av + a_1 v) = (a + a_1)v$ also in $\mathbb{R}v$
3. Given av, $\quad r(av) = (ra)v$ in $\mathbb{R}v$

Example 6.2.5

all nxn matrices

Let A be a fixed matrix in $\mathbf{M}_{nn}$. Show that $U = \{X \text{ in } \mathbf{M}_{nn} \mid AX = XA\}$ is a subspace of $\mathbf{M}_{nn}$.

Subspace Test

① $0 \in U$? let 0 be nxn zero matrix then

$0A = A0 = 0$ so $0 \in U$

② let $X, X_1 \in U$ $(AX = XA,\ AX_1 = X_1A)$

$A(X + X_1) = AX + AX_1 = XA + X_1A$

$= (X + X_1)A$

③ $A(aX) = a(AX) = a(XA) = (aX)A$

$X + X_1 \in U$ and $aX \in U$

A bit more polynomial notation

- Let $p(x)$ be a polynomial and a be a number
- **Evaluation** of $p(x)$ at $a \rightarrow$ the number $p(a)$
- **Root** of $p(x) \rightarrow$ the number a that gives $p(a) = 0$

Example 6.2.6

all polynomials w/ real coeff

Consider the set U of all polynomials in **P** that have 3 as a root:

$$U = \{p(x) \in \mathbf{P} \mid \underline{p(3) = 0}\}$$

Show that U is a subspace of **P**.

① 0 polynomial is in U

$0 + 0(3) + 0(3)^2 + \cdots + 0(3)^n = 0$

② let $p(x), q(x) \in U$ $\quad p(3) = 0 = q(3)$

$(p+q)(x) = p(x) + q(x) \quad \forall x$

$(p+q)(3) = p(3) + q(3) = 0 + 0 = 0$

③ $a\,p(3) = aa_0 + aa_1(3) + \cdots = a(\underbrace{a_0 + a_1(3) + \cdots}_{=0}) = a0 = 0$

Example 6.2.7

- Recall: $\mathbf{P}_n$ consists of all polynomials of the form

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

where $a_0, a_1, a_2, \ldots, a_n$ are real numbers

Example 6.2.7

- Recall: $\mathbf{P}_n$ consists of all polynomials of the form

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

  where $a_0, a_1, a_2, \ldots, a_n$ are real numbers

- $\mathbf{P}_n \rightarrow$ closed under the addition and scalar multiplication in $\mathbf{P}$

Example 6.2.7

- Recall: $\mathbf{P}_n$ consists of all polynomials of the form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

  where $a_0, a_1, a_2, \ldots, a_n$ are real numbers

- $\mathbf{P}_n \rightarrow$ closed under the addition and scalar multiplication in $\mathbf{P}$
- Zero polynomial is in $\mathbf{P}_n$

Example 6.2.7

- Recall: $\mathbf{P}_n$ consists of all polynomials of the form

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

  where $a_0, a_1, a_2, \ldots, a_n$ are real numbers
- $\mathbf{P}_n \rightarrow$ closed under the addition and scalar multiplication in $\mathbf{P}$
- Zero polynomial is in $\mathbf{P}_n$ $\quad a_0 = a_1 = \cdots = a_n = 0$
- Thus by the subspace test:

$$\mathbf{P}_n \text{ is a subspace of } \mathbf{P} \text{ for each } n \geq 0$$

Example 6.2.8

Show that the subset $\mathbf{D}[a, b]$ of all **differentiable functions** on $[a, b]$ is a subspace of the vector space $\mathbf{F}[a, b]$ of all functions on $[a, b]$.
(A function f defined on the interval $[a, b]$ is called differentiable if the derivative $f'(r)$ exists at every r in $[a, b]$.)

① constant 0 is diff. $\Rightarrow$ in $D[a,b]$
derivative of any constant func is 0

② f $g \in D[a,b]$ (f', g' exist)
$f+g$ is diff for any $r \in \mathbb{R}$
$(f+g)' = f' + g' \to \in D[a,b]$

③ similar for $(rf)' = rf' \in D[a,b]$

Linear Combinations and Spanning

Linear Combinations and Spanning

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a set of vectors in a vector space V. As in $\mathbb{R}^n$, a vector $\mathbf{v}$ is called a **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ if it can be expressed in the form

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$$

where $a_1, a_2, \ldots, a_n$ are scalars, called the **coefficients** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. The set of all linear combinations of these vectors is called their **span**, and is denoted by

$$V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \mid a_i \text{ in } \mathbb{R}\}$$

If $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, these vectors are called a **spanning set** for V.

Example 6.2.10

all poly with deg 2

Consider the vectors $p_1 = 1 + x + 4x^2$ and $p_2 = 1 + 5x + x^2$ in $\mathbf{P}_2$. Determine whether p_1 and p_2 lie in $\text{span}\{1 + 2x - x^2, 3 + 5x + 2x^2\}$.

a vector is in span if s, t exist s.t.

$p_1 = s(1+2x-x^2) + t(3+5x+2x^2)$?

(Equate powers of x)

$1 = s + 3t \quad 1 = 2s + 5t \quad 4 = -s + 2t$

$s = -2 \quad t = 1 \Rightarrow p_1$ is in span$\{\ldots\}$

Repeat for p_2:

$1 = s + 3t \quad 5 = 2s + 5t \quad 1 = -s + 2t$

$\rightarrow$ no solution $\quad p_2$ not in span$\{\ldots\}$

Example 6.2.11

$\mathbf{M}_{mn}$ is the span of the set of all $m \times n$ matrices with exactly one entry equal to 1, and all other entries zero.

cols of I_m

$$\mathbb{R}^m = \text{span}\{e_1, \cdots, e_m\} \qquad \mathbb{R}^m = M_{m1}$$

$m \times 1$ matrices

M_{mn} is analogous

E.g. 2×2 matrices

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{A} + b\underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{B} + c\underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{C} + d\underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}}_{D}$$

$$M_{22} = \text{span}\{A, B, C, D\}$$

Example 6.2.12

$\mathbf{P}_n = \text{span}\{1, x, x^2, \ldots, x^n\}$

Theorem 6.2.13

Theorem

Let $U = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V. Then:

1. U is a subspace of V containing each of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.
2. U is the "smallest" subspace containing these vectors in the sense that any subspace that contains each of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ must contain U.

Example 6.2.14

Show that $\mathbf{P}_3 = \text{span}\{x^2 + x^3, x, 2x^2 + 1, 3\}$.

Example 6.2.15

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in a vector space V. Show that

$$\mathrm{span}\{\mathbf{u}, \mathbf{v}\} = \mathrm{span}\{\mathbf{u} + 2\mathbf{v}, \mathbf{u} - \mathbf{v}\}$$

Recap

Today we saw:

- Subspaces and spanning sets of vector spaces

Next time: Linear Independence and Dimension

MATH254: Linear Algebra
Lecture 10

Moira MacNeil

January 28, 2025

Last Time

1. Subspaces and spanning sets of vector spaces

Today

1. One more theorem about spanning sets
2. Linear independence in vector spaces

Reminders:

- Assignment 2 is due Friday, January 31 → Typo in Q5
- Midterm 1 is Friday, February 7 → INFO on Moodle

Recall: Subspaces

Subspaces of a Vector Space

If V is a vector space, a nonempty subset $U \subseteq V$ is called a **subspace** of V if U is itself a vector space using the addition and scalar multiplication of V

Recall: Linear Combinations and Spanning

Linear Combinations and Spanning

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a set of vectors in a vector space V. As in $\mathbb{R}^n$, a vector $\mathbf{v}$ is called a **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ if it can be expressed in the form

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$$

where $a_1, a_2, \ldots, a_n$ are scalars, called the **coefficients** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. The set of all linear combinations of these vectors is called their **span**, and is denoted by

$$\underline{\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}} = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \mid a_i \text{ in } \mathbb{R}\}$$

If $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, these vectors are called a **spanning set** for V.

Theorem 6.2.13

Theorem

Let $U = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V. Then:

1. U is a subspace of V containing each of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.
2. U is the "smallest" subspace containing these vectors in the sense that any subspace that contains each of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ must contain U.

Example 6.2.14

P_n set of n degree polynomials $n=3$

Show that $\mathbf{P}_3 = \text{span}\{x^2 + x^3, x, 2x^2 + 1, 3\}$.

$$U = \text{span}\{\underbrace{x^2+x^3,\ x,\ 2x^2+1,}_{\text{all in } P_3}\ 3\} \Rightarrow U \subseteq P_3$$

Want to show $P_3 \subseteq U$, we know

$$P_3 = \text{span}\{1, x, x^2, x^3\} = t_1 1 + t_2 x + t_3 x^2 + t_4 x^3$$

x and $1 = \left[\tfrac{1}{3}\right]3$ are in U

$x^2 = \tfrac{1}{2}\left[(2x^2+1) - 1\right]$ is in U

$x^3 = (x^2 + x^3) - x^2$ is in U

$\Rightarrow P_3 \subseteq U \quad \therefore P_3 = U$ condition 2 of thm

Example 6.2.15

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in a vector space V. Show that $\text{span}\{\mathbf{u}, \mathbf{v}\} = \text{span}\{\mathbf{u} + 2\mathbf{v}, \mathbf{u} - \mathbf{v}\}$.

Show $\text{span}\{u+2v, u-v\} \subseteq \text{span}\{u,v\} = au+bv$

clearly $u+2v$ and $u-v$ in $\text{span}\{u,v\}$ ✓

Show $\text{span}\{u,v\} \subseteq \text{span}\{u+2v, u-v\}$ ✓

$$= a(u+2v) + b(u-v)$$

$$= (a+b)u + (2a-b)v$$

$u \rightarrow$ $a+b=1$, $2a-b=0$, $a = 1/3$ $b = 2/3$

$v \rightarrow$ $a+b=0$, $2a-b=1$, $a = 1/3$ $b = -1/3$

Linear Independence

Linear Independence and Dependence

As in $\mathbb{R}^n$, a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V is called **linearly independent** (or simply **independent**) if it satisfies the following condition:

$$\text{If} \quad \underbrace{s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_n\mathbf{v}_n = \mathbf{0}}_{\text{vanishes}}, \quad \text{then} \quad \underbrace{s_1 = s_2 = \cdots = s_n = 0}_{\text{trivial}}.$$

A set of vectors that is not linearly independent is said to be **linearly dependent** (or simply **dependent**).

Recall: The **trivial linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is the one with every coefficient zero: $0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n$

These vectors are linearly independent when this is the only way to express $\mathbf{0}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$

Example 6.3.2

Show that $\{1+x, 3x+x^2, 2+x-x^2\}$ is independent in $\mathbf{P}_2$.

Suppose

$$s_1(1+x) + s_2(3x+x^2) + s_3(2+x-x^2) = 0$$

$$s_1 + s_1 x + 3s_2 x + s_2 x^2 + 2s_3 + s_3 x - s_3 x^2 = 0$$

$$(*)\ (s_1 + 2s_3) + (s_1 + 3s_2 + s_3)x + (s_2 - s_3)x^2 = 0$$

If $s_1 = s_2 = s_3 = 0$ is only sol to $(*)$ $\rightarrow$ INDEP.

$$s_1 + 2s_3 = 0$$
$$s_1 + 3s_2 + s_3 = 0$$
$$s_2 - s_3 = 0$$

$$s_1 + 2s_2 = 0$$
$$s_1 + 4s_2 = 0$$

$s_2 = 0 \rightarrow s_3 = 0 \rightarrow s_1 = 0$

Example 6.3.3

Show that $\{\sin x, \cos x\}$ is independent in the vector space $\mathbf{F}[0, 2\pi]$ of functions defined on the interval $[0, 2\pi]$.

Suppose $s_1(\sin x) + s_2(\cos x) = 0$

This must hold for all $x \in [0, 2\pi]$

Let $x = 0 \rightarrow \sin(0) = 0 \quad \cos(0) = 1$

Then we must have $s_2 = 0$ (vanishes for all $x \in [0, 2\pi]$) Leaving $s_1 \sin(x) = 0$

Let $x = \pi/2 \quad \sin(\pi/2) = 1$ so $s_1 = 0$

$\rightarrow s_1 = s_2 = 0$ and indept. in $F[0, 2\pi]$

Example 6.3.4

Suppose that $\{\mathbf{u}, \mathbf{v}\}$ is an independent set in a vector space V. Show that $\{\mathbf{u} + 2\mathbf{v}, \mathbf{u} - 3\mathbf{v}\}$ is also independent.

Suppose $s(u+2v) + t(u-3v) = 0$

Collect terms $(s+t)u + (2s-3t)v = 0$

Since $\{u, v\}$ indept.

$s+t=0$

$2s-3t=0$

$s=t=0 \Rightarrow \{u+2v, u-3v\}$ indept.

Example 6.3.5

Eg $P_3 = \text{span}\underbrace{\{1, x, x^2, x^3\}}_{\text{indep}}$

Show that any set of polynomials of distinct degrees is independent.

Let $p_1, \dots, p_m$ be polynomials where $\deg(p_i) = d_i$, assume (by relabelling) that $d_1 > d_2 > \dots > d_m$

Suppose

$t_1p_1 + t_2p_2 + \dots + t_mp_m = 0 \quad (t_i \in \mathbb{R})$

let ax^{d_1} be highest deg term in p_1 $(a \neq 0)$

since $d_1 > d_2 > \dots > d_m$ it must be the only term of degree d_1 in the linear

combination. $\rightarrow$ we must have $t_1 = 0$
$(t_1 a x^{d_1} = 0, a \neq 0)$
same argument for $t_2 = 0$ and so on, conclude that all $t_i = 0$.

Example 6.3.6

$M_{nm} \to$ $\underline{n \times m}$ matrices

Suppose that A is an $n \times n$ matrix such that $\underline{A^k = 0}$ but $A^{k-1} \neq 0$. Show that $B = \{I, A, A^2, \ldots, A^{k-1}\}$ is independent in $\mathbf{M}_{nn}$.

Suppose $t_0 I + t_1 A + \cdots + t_{k-1} A^{k-1} = 0$

Multiply by A^{k-1}

$$t_0 A^{k-1} + t_1 \underbrace{A^k}_{=0} + \cdots + t_{k-1} A^{2k-2} = 0$$

if $A^k = 0$ any higher power is also 0

$$t_0 \underbrace{A^{k-1}}_{\neq 0} = 0 \Rightarrow r_0 = 0 \text{ and we have}$$

$$t_1 A + \cdots + t_{k-1} A^{k-1} = 0$$

mult. by A^{k-2} $\rightarrow$ $t_1 = 0$
repeat for each term $t_i = 0$
and B is indep. in M_{nm}

Example 6.3.7

Let V denote a vector space.

1. If $\mathbf{v} \neq \mathbf{0}$ in V, then $\{\mathbf{v}\}$ is an independent set.
2. No independent set of vectors in V can contain the zero vector.

Theorem 6.3.8

Theorem

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a linearly independent set of vectors in a vector space V. If a vector $\mathbf{v}$ has two (ostensibly different) representations

$$\mathbf{v} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_n\mathbf{v}_n$$
$$\mathbf{v} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_n\mathbf{v}_n$$

as linear combinations of these vectors, then $s_1 = t_1, s_2 = t_2, \ldots, s_n = t_n$. In other words, every vector in V can be written in a unique way as a linear combination of the $\mathbf{v}_i$.

Theorem 6.3.9

Fundamental Theorem

Suppose a vector space V can be spanned by n vectors. If any set of m vectors in V is linearly independent, then $m \leq n$.

Steinitz Exchange Lemma

- If $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, and if $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ is an independent set in V
- Then the Fundamental Theorem shows $m \leq n$
- We also have that m of the (spanning) vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ can be replaced by the (independent) vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ and the resulting set will still span V
- In this form the result is called the **Steinitz Exchange Lemma**

Basis

Basis of a Vector Space

As in $\mathbb{R}^n$, a set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of vectors in a vector space V is called a **basis** of V if it satisfies the following two conditions:

1. $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is linearly independent
2. $V = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$

Theorem 6.3.11

Invariance Theorem

Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be two bases of a vector space V. Then $n = m$.

Dimension

Dimension of a Vector Space

If $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is a basis of the nonzero vector space V, the number n of vectors in the basis is called the **dimension** of V, and we write

$$\dim V = n$$

The zero vector space $\{\mathbf{0}\}$ is defined to have dimension 0:

$$\dim \{\mathbf{0}\} = 0$$

Recap

Today we saw:

- Linear independence and dimension in vector spaces

Next time: Examples of Basis and Dimension

MATH254: Linear Algebra
Lecture 11

Moira MacNeil

January 29, 2025

Last Time

1. Linear independence and dimension in vector spaces

Today

1. More linear independence in vector spaces

2. Examples of basis and dimension

Reminders:

- Assignment 2 is due Friday, January 31
- Midterm 1 is Friday, February 7

Recall: Linear Independence

Linear Independence and Dependence

As in $\mathbb{R}^n$, a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V is called **linearly independent** (or simply **independent**) if it satisfies the following condition:

$$\text{If} \quad \underbrace{s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_n\mathbf{v}_n = \mathbf{0}}_{\text{vanishes}}, \quad \text{then} \quad \overbrace{s_1 = s_2 = \cdots = s_n = 0}^{\text{trivial}}.$$

A set of vectors that is not linearly independent is said to be **linearly dependent** (or simply **dependent**).

Example 6.3.7

Let V denote a vector space.

1. If $\mathbf{v} \neq \mathbf{0}$ in V, then $\{\mathbf{v}\}$ is an independent set.
2. No independent set of vectors in V can contain the zero vector.

1. Assume $tv = 0$ $t \in \mathbb{R}$.
   Assume $t \neq 0$ $v = \frac{1}{t}\underbrace{(tv)}_{=0} = \frac{1}{t}(0) = 0$
   contradicts $v \neq 0$ so $t = 0$ and $\{v\}$ indep.
2. If $\{v_1, v_2, \dots, v_k\}$ is indep. let $v_2 = 0$
   $0v_1 + 1v_2 + \cdots + 0v_k = 0$ is a nontrivial comb. that vanishes $\Rightarrow\Leftarrow$ $\{v_1, \dots, v_k\}$ indept.

Theorem 6.3.8

Theorem

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a linearly independent set of vectors in a vector space V. If a vector $\mathbf{v}$ has two (ostensibly different) representations

$$\mathbf{v} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_n\mathbf{v}_n$$
$$\mathbf{v} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_n\mathbf{v}_n$$

as linear combinations of these vectors, then $s_1 = t_1, s_2 = t_2, \ldots, s_n = t_n$. In other words, every vector in V can be written in a unique way as a linear combination of the $\mathbf{v}_i$.

PROOF Subtract two equations

$(s_1 - t_1)v_1 + \cdots + (s_n - t_n)v_n = 0$

$\{v_1, v_2, \ldots, v_n\}$ indep. $\Rightarrow (s_i - t_i) = 0 \Rightarrow s_i = t_i$
for each i $\square$

Theorem 6.3.9

Fundamental Theorem

Suppose a vector space V can be spanned by n vectors. If any set of m vectors in V is linearly independent, then $m \leq n$.

$V = \text{span}\{v_1, v_2, \cdots, v_n\}$

$\{u_1, u_2, \cdots, u_m\}$ indep. in V then $m \leq n$

Steinitz Exchange Lemma

- If $V = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, and if $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ is an independent set in V
- Then the Fundamental Theorem shows $m \leq n$
- We also have that m of the (spanning) vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ can be replaced by the (independent) vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ and the resulting set will still span V
- In this form the result is called the **Steinitz Exchange Lemma**

Basis

Basis of a Vector Space

As in $\mathbb{R}^n$, a set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of vectors in a vector space V is called a **basis** of V if it satisfies the following two conditions:

1. $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is linearly independent
2. $V = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$

Any vector v in V can be written as a unique linear comb. of vectors in V

Theorem 6.3.11

Invariance Theorem

Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be two bases of a vector space V. Then $n = m$.

PROOF $V = \text{Span}\{e_1, \cdots, e_n\}$ and $\{f_1, \cdots, f_m\}$ indep (by def of basis), Fund Thm $m \leq n$

Similarly $V = \text{span}\{f_1, \cdots, f_m\}$ and $\{e_1, \cdots e_n\}$ indept. Fund Thm $n \leq m \Rightarrow n = m$ $\square$

Dimension

Dimension of a Vector Space

If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis of the nonzero vector space V, the number n of vectors in the basis is called the **dimension** of V, and we write

$$\boxed{\dim\ V = n}$$

vectors in basis

The zero vector space $\{\mathbf{0}\}$ is defined to have dimension 0:

$$\dim\ \{\mathbf{0}\} = 0$$

$\{0\}$ has no basis $\quad \dim\{0\} = 0$

empty set is basis for $\{0\}$

Example 6.3.13

m×n matrices

The space $\mathbf{M}_{mn}$ has dimension mn, and one basis consists of all $m \times n$ matrices with exactly one entry equal to 1 and all other entries equal to 0. We call this the **standard basis** of $\mathbf{M}_{mn}$.

Eg. in M_{22} dim $= 2(2) = 4$

standard basis:

$$a\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}+b\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}+c\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}+d\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}a & b\\c & d\end{bmatrix}$$

Example 6.3.14

P_n = poly of deg at most n

Show that $\dim \mathbf{P}_n = n+1$ and that $\{1, x, x^2, \dots, x^n\}$ is a basis, called the **standard basis** of $\mathbf{P}_n$.

② $p(x) = \underbrace{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n}_{\text{linear comb. of } 1, x, x^2 \dots}$ in P_n

So $P_n = \operatorname{span}\{1, x, \dots, x^n\}$

① If linear comb vanishes

$a_0 + a_1 x + \cdots + a_n x^n = 0$ then $a_0 = a_1 = \cdots = a_n = 0$

since x is indeterminante

$\therefore \{1, x, \dots, x^n\}$ is a basis with $n+1$ vectors $\Rightarrow \dim P_n = n+1$

Example 6.3.15

If $\mathbf{v} \neq \mathbf{0}$ is any nonzero vector in a vector space V, show that $\text{span}\{\mathbf{v}\} = \mathbb{R}\mathbf{v}$ has dimension 1.

$\{v\}$ spans $\mathbb{R}v$ (can write any $u \in \mathbb{R}v$ as tv)

$\{v\}$ is lin indept (previous eg.)

$\therefore$ $\{v\}$ is a basis for $\mathbb{R}v$

and $\dim(\mathbb{R}v) = 1$

Example 6.3.16

Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and consider the subspace

$$U = \{X \text{ in } \mathbf{M}_{22} \mid AX = XA\}$$

of $\mathbf{M}_{22}$. Show that $\dim U = 2$ and find a basis of U.

We have shown U is a subspace for any A

if $X = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ is in U $AX = XA$

$$AX = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} x+z & y+w \\ 0 & 0 \end{bmatrix}$$

$$XA = \begin{bmatrix} x & y \\ z & w \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} x & x \\ z & z \end{bmatrix}$$

Equate $z = 0$ $x = y + w$

We can write

$$X = \begin{bmatrix} y+w & y \\ 0 & w \end{bmatrix} = y\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + w\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$B = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ then $U = \text{span } B$

B indept. $\therefore$ B basis of U

(show) $\dim U = 2$

Example 6.3.17

Show that the set V of all symmetric 2×2 matrices is a vector space, and find the dimension of V.

A symm. if $A = A^T$

If $A, B \in V$ using prop of symm. matrices transposes +

$(A+B)^T = A^T + B^T = A+B$

$(kA)^T = kA^T = kA$

V is a vector space

inherit from M_{22}

A is symmetric → entries directly across main diag are equal

2x2 symm matrices

$$\begin{bmatrix} a & c \\ c & b \end{bmatrix} = a\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{u_1} + b\underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}}_{u_2} + c\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{u_3}$$

$B = \{u_1, u_2, u_3\}$ $\quad V = \text{span } B$

(show) B is indep.

$\therefore$ B basis of V $\quad \dim V = 3$

Recap

Today we saw:

- Linear independence, basis and dimension in vector spaces

Next time: Finite dimensional spaces

MATH254: Linear Algebra
Lecture 12

Moira MacNeil

January 31, 2025

Last Time

1. Linear independence, basis and dimension in vector spaces

Today

1. Can we guarantee that an arbitrary vector space has a basis?
2. How can we find that basis?

Reminders:

- Midterm 1 is Friday, February 7
- Assignment 3 is now posted, due Friday, February 14

Lemma 6.4.1

Independent Lemma

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be an independent set of vectors in a vector space V. If $\mathbf{u} \in V$ but $\mathbf{u} \notin \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, then $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is also independent.

Converse is also true:
if $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is independent, then $\mathbf{u}$ is not in $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$.

PROOF

Let $tu + t_1 v_1 + \cdots + t_k v_k = 0$

WTS all coeff are zero

We have $t=0$, otherwise

$u = -\frac{t_1}{t} v_1 - \frac{t_2}{t} v_2 - \cdots - \frac{t_k}{t} v_k$ is in span $\{v_1, \cdots, v_k\}$ (contradiction)

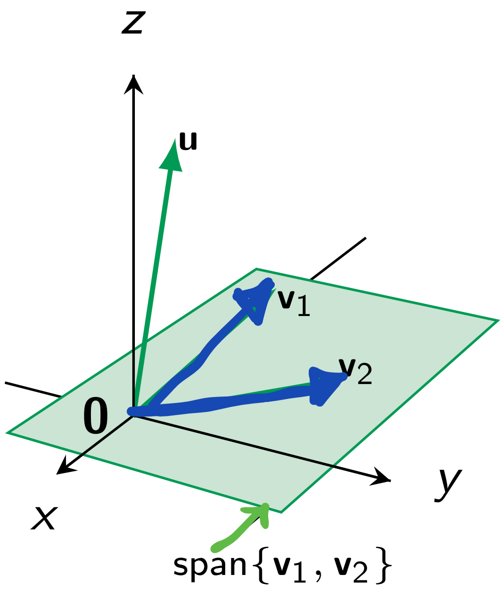
Then $t_1 v_1 + \cdots + t_k v_k = 0$

rest of t_i are zero since $\{v_1, \cdots, v_k\}$ independent.

$\square$

Independent Lemma Graphically

- Suppose $\{\mathbf{v}_1, \mathbf{v}_2\}$ independent in $\mathbb{R}^3$
- Then $\mathbf{v}_1$ and $\mathbf{v}_2$ are not parallel
- $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ is a plane through the origin
- By the Independent Lemma, $\mathbf{u}$ is not in this plane $\iff$ $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2\}$ is independent



Finite and Infinite Dimensional Spaces

Finite Dimensional and Infinite Dimensional Vector Spaces

A vector space V is called **finite dimensional** if it is spanned by a finite set of vectors. Otherwise, V is called **infinite dimensional**.

Zero vector space $\{\mathbf{0}\}$ is finite dimensional because $\{\mathbf{0}\}$ is a spanning set.

Lemma 6.4.3

Lemma

Let V be a finite dimensional vector space. If U is any subspace of V, then any independent subset of U can be enlarged to a finite basis of U.

Theorem 6.4.4

Theorem

Let V be a finite dimensional vector space spanned by m vectors.

1. V has a finite basis, and $\dim V \leq m$.
2. Every independent set of vectors in V can be enlarged to a basis of V by adding vectors from any fixed basis of V.
3. If U is a subspace of V, then
   (a) U is finite dimensional and $\dim U \leq \dim V$.
   (b) If $\dim U = \dim V$ then $U = V$.

V is finite dimensional $\iff$ V has finite basis (possibly empty), and that every subspace of a finite dimensional space is itself finite dimensional

Example 6.4.6 polynomials of deg at most 3

Find a basis of $\mathbf{P}_3$ containing the independent set $\{1+x, 1+x^2\}$.

Standard basis of P_3: $\{1, x, x^2, x^3\}$

dim $P_3 = 4$ so add two vectors from standard basis.

$\{1, 1+x, 1+x^2, x^3\}$ → show independent

$\underbrace{1, 1+x, 1+x^2, x^3}_{\text{distinct degrees}}$ distinct degrees ∴ indep. (see prev. ex)

This is a basis.

Note: adding $\{1, x\}$ or $\{1, x^2\}$ wouldn't work.

Example 6.4.7

Show that the space **P** of all polynomials is infinite dimensional.

For each $n \geq 1$, P has a subspace P_n with dim $n+1$.

Suppose P is finite dimensional, say dim $P = m$. Then $\dim P_n \leq \dim P$

$\Rightarrow n+1 \leq m$.

Contradiction since n is arbitrary (no single upper bound on $n+1$)

$\therefore$ P is infinite dimensional.

Theorem 6.4.9

Theorem

Let U and W be subspaces of the finite dimensional space V.

1. If $U \subseteq W$, then $\dim U \leq \dim W$.
2. If $U \subseteq W$ and $\dim U = \dim W$, then $U = W$.

Example 6.4.10

If a is a number, let W denote the subspace of all polynomials in $\mathbf{P}_n$ that have a as a root:

B=

$$W = \{p(x) \mid p(x) \in \mathbf{P}_n \text{ and } p(a) = 0\}$$

Show that $\{(x-a), (x-a)^2, \ldots, (x-a)^n\}$ is a basis of W.

First observe that each element of B is in W (each is a poly. with a as a root).
We also have that they are indep. since the degrees are distinct.

$U = \operatorname{span} B$ so we have $U \subseteq W \subsetneq P_n$

$\dim U = n$ (B has n elem indep.) $\dim P_n = n+1$

$\dim W$ is an integer and

$\dim U = n \leq \dim W \leq \dim P_n = n+1$

so $\dim W$ is either n or $n+1$

But then $W = U$ or $W = P_n$

Because $W \neq P_n$ then $W = U$ and

B is a basis W as required.

Lemma 6.4.11

Recall: A set of vectors is called **dependent** if it is *not* independent, that is if some nontrivial linear combination vanishes.

Dependent Lemma

A set $D = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in a vector space V is dependent if and only if some vector in D is a linear combination of the others.

Theorem 6.4.12

Theorem

Let V be a finite dimensional vector space. Any spanning set for V can be cut down (by deleting vectors) to a basis of V.

Example 6.4.13

Find a basis of $\mathbf{P}_3$ in the spanning set $S = \{1, x + x^2, 2x - 3x^2, \underline{1 + 3x - 2x^2}, x^3\}$.

dim $P_3 = 4 \to$ eliminate one poly. from S.

Can't remove x^3 since rest of S is in span P_2. Eliminating $1+3x-2x^2$ leaves a basis. Note that

$$1+3x-2x^2 = 1+(x+x^2)+(2x-3x^2)$$

* Show remaining vectors are indep. and Span P_3

Theorem 6.4.14

Theorem

Let V be a vector space with $\dim V = n$, and suppose S is a set of exactly n vectors in V. Then S is independent if and only if S spans V.

Example 6.4.15

Consider the set $S = \{p_0(x), p_1(x), \dots, p_n(x)\}$ of polynomials in $\mathbf{P}_n$. If $\deg p_k(x) = k$ for each k, show that S is a basis of $\mathbf{P}_n$.

S is indep. since all degrees are distinct.

S has $n+1 = \dim P_n$

So by Thm S is a basis P_n

Example 6.4.5

2x2 matrices

Enlarge the independent set $D = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \right\}$ to a basis of $\mathbf{M}_{22}$.

Standard basis M_{22} is: ($\dim M_{22} = 4$)

$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ so add one of these to D

Adding any standard basis vector gives indep. set:

$$a \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} + c \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Equate entries

$a + c = 0 \qquad a = -c$

$a + b = 0$

$a + b + c + d = 0$

$b + c = 0$

$$\begin{array}{r} b - c = 0 \\ b + c = 0 \\ \hline b = 0 \\ a = 0 \\ c = 0 \\ d = 0 \end{array}$$

$\therefore$ basis

Theorem 6.4.18

If U and W are subspaces of a vector space V, their **sum** $U+W$ and their **intersection** $U \cap W$ are

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U \text{ and } \mathbf{w} \in W\}$$
$$U \cap W = \{\mathbf{v} \in V \mid \mathbf{v} \in U \text{ and } \mathbf{v} \in W\}$$

Theorem

Suppose that U and W are finite dimensional subspaces of a vector space V. Then $U+W$ is finite dimensional and

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W).$$

Addition rule prob.

Recap

Today we saw:

- Finite dimensional spaces
- Relationship between span and linear independence

Next time: A new chapter: Linear Transformations

MATH254: Linear Algebra
Lecture 13

Moira MacNeil

February 4, 2025

Last Time

1. Finite dimensional spaces
2. Relationship between span and linear independence

Today

1. Linear transformations: properties and examples

Reminders:

- Midterm 1 is this Friday, February 7
- Assignment 3 is due Friday, February 14

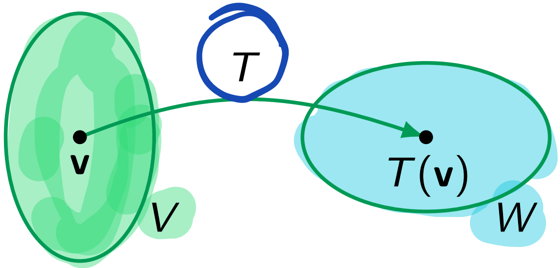
Linear Transformations

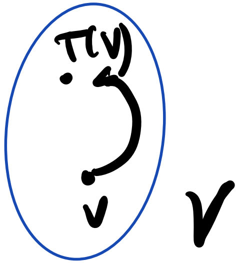
Linear Transformations of Vector Spaces

If V and W are two vector spaces, a function $T : V \to W$ is called a **linear transformation** if it satisfies the following axioms.

T1. $T(\mathbf{v} + \mathbf{v}_1) = T(\mathbf{v}) + T(\mathbf{v}_1)$ for all $\mathbf{v}$ and $\mathbf{v}_1$ in V.
T2. $T(r\mathbf{v}) = rT(\mathbf{v})$ for all $\mathbf{v}$ in V and r in $\mathbb{R}$.

A linear transformation $T : V \to V$ is called a **linear operator** on V.





Axioms of Linear Transformations

$$T1: \quad \underbrace{T(v+v_1)}_{V} = \underbrace{T(v)+T(v_1)}_{W}$$

- Axiom T1 $\rightarrow$ T **preserves** vector addition. $T: V \to W$
- That is the result $T(\mathbf{v}+\mathbf{v}_1)$ of adding $\mathbf{v}$ and $\mathbf{v}_1$ then applying T is the same as applying T to get $T(\mathbf{v})$ and $T(\mathbf{v}_1)$ and then adding
- Axiom T2 means that T **preserves** scalar multiplication.
- Note: the additions in axiom T1 are both denoted by $+$, but the addition on the left forming $\mathbf{v}+\mathbf{v}_1$ is carried out in V, whereas the addition $T(\mathbf{v})+T(\mathbf{v}_1)$ is done in W
- The scalar multiplications $r\mathbf{v}$ and $rT(\mathbf{v})$ in axiom T2 refer to the spaces V and W, respectively.

Example 7.1.2

If V and W are vector spaces, the following are linear transformations:

Identity operator $V \to V$	$1_V : V \to V$	where $1_V(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v}$ in V
Zero transformation $V \to W$	$0 : V \to W$	where $0(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$ in V
Scalar operator $V \to V$	$a : V \to V$	where $a(\mathbf{v}) = a\mathbf{v}$ for all $\mathbf{v}$ in V
		(Here a is any real number.)

The symbol 0 will be used to denote the zero transformation from V to W for *any* spaces V and W.

1_v satisfies axioms

let $v \in V$ and $u \in V$ and $r \in \mathbb{R}$

T1: $1_v(v+u) = v+u = 1_v(v) + 1_v(u)$

T2: $1_v(rv) = rv = r\,1_v(v)$

Example 7.1.3

sum elem on main diag.

Show that the transposition and trace are linear transformations. More precisely,

$$R : \mathbf{M}_{mn} \to \mathbf{M}_{nm} \qquad \text{where } R(A) = A^T \text{ for all } A \text{ in } \mathbf{M}_{mn}$$
$$S : \mathbf{M}_{mn} \to \mathbb{R} \qquad \text{where } S(A) = \text{tr } A \text{ for all } A \text{ in } \mathbf{M}_{nn}$$

are both linear transformations.

Transpose let $A, B \in M_{mn}$, $r \in \mathbb{R}$

T1: $R(A+B) = (A+B)^T = A^T + B^T = R(A) + R(B)$ ✓

T2: $R(rA) = (rA)^T = r(A^T) = r\,R(A)$ ✓

From properties of transpose

Trace

T1: $S(A+B) = tr(A+B) = tr(A) + tr(B) = S(A) + S(B)$ ✓

T2: $S(rA) = tr(rA) = r\,tr(A) = r\,S(A)$ ✓

Example 7.1.4

$\swarrow$ polynomials of deg at most n

If a is a scalar, define $E_a : \mathbf{P}_n \to \mathbb{R}$ by $E_a(p) = p(a)$ for each polynomial p in $\mathbf{P}_n$. Show that E_a is a linear transformation (called **evaluation** at a).

Let p, q be polynomials, $r \in \mathbb{R}$. Then

sum of $p+q$: $(p+q)(x) = p(x) + q(x)$

scalar mult.: $(rp)(x) = r\,p(x)$ for all x.

For all $p, q \in P_n$ and all $r \in \mathbb{R}$

T1: $E_a(p+q) = (p+q)(a) = p(a) + q(a) = E_a(p) + E_a(q)$

T2: $E_a(rp) = (rp)(a) = r\,p(a) = r\,E_a(p)$

$\therefore$ E_a is a linear transformation.

Theorem 7.1.5

Theorem

Let $T : V \to W$ be a linear transformation.

1. $T(\mathbf{0}) = \mathbf{0}$.
2. $T(-\mathbf{v}) = -T(\mathbf{v})$ for all $\mathbf{v}$ in V.
3. $T(r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \cdots + r_k\mathbf{v}_k) = r_1T(\mathbf{v}_1) + r_2T(\mathbf{v}_2) + \cdots + r_kT(\mathbf{v}_k)$ for all $\mathbf{v}_i$ in V and all r_i in $\mathbb{R}$.

Proof.

① $T(0) = T(0v) = 0\,T(v) = 0$ for all $v \in V$

② $T(-v) = T([-1]v) = (-1)T(v) = -T(v)$ for $v \in V$

③ induction on k

Property (3)

- Property (3): If $T : V \to W$ is a linear transformation and $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ are known, then $T(\mathbf{v})$ can be computed for *every* vector $\mathbf{v}$ in $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$
- If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ spans V, then $T(\mathbf{v})$ is determined for all $\mathbf{v}$ in V by the choice of $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$
- Two linear transformations $T : V \to W$ and $S : V \to W$ are called **equal** $(T = S)$ if they have the same **action** $\to$ if $T(\mathbf{v}) = S(\mathbf{v})$ for all $\mathbf{v}$ in V

Example 7.1.6

Let $T : V \to W$ be a linear transformation. If $T(\mathbf{v} - 3\mathbf{v}_1) = \mathbf{w}$ and $T(2\mathbf{v} - \mathbf{v}_1) = \mathbf{w}_1$, find $T(\mathbf{v})$ and $T(\mathbf{v}_1)$ in terms of $\mathbf{w}$ and $\mathbf{w}_1$.

$$T(v-3v_1) \overset{3}{=} T(v) - 3T(v_1) = w$$

$$T(2v-v_1) \overset{3}{=} 2T(v) - T(v_1) = w_1$$

Solve for $T(v)$ and $T(v_1)$!

$$T(v) - 6T(v) = w - 3w_1$$

$$T(v) = -\tfrac{1}{5}(w - 3w_1)$$

$$-T(v_1) + 6T(v_1) = w_1 - 2w$$

$$T(v_1) = \tfrac{1}{5}(w_1 - 2w)$$

Theorem 7.1.7

Theorem

Let $T : V \to W$ and $S : V \to W$ be two linear transformations. Suppose that $V = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. If $\mathrm{T}(\mathbf{v}_i) = S(\mathbf{v}_i)$ for each i, then $T = S$.

If we know what a linear transformation $T : V \to W$ does to each vector in a spanning set for V, then we know what T does to *every* vector in V.

Proof.

If $v \in V = \text{span}\{v_1, \cdots, v_n\}$, then let $v = a_1 v_1 + \cdots + a_n v_n$ where $a_i \in \mathbb{R}$. Since $T(v_i) = S(v_i)\ \forall i$ then

$$T(v) = T(a_1 v_1 + \cdots + a_n v_n)$$
$$= a_1 T(v_1) + \cdots + a_n T(v_n)$$

$$= a_1 S(v_1) + \cdots + a_n S(v_n)$$
$$= S(a_1 v_1 + \cdots + a_n v_n)$$
$$= S(v)$$

Since v arbitrary, $T = S$ $\square$

Example 7.1.8

Let $V = \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Let $T : V \to W$ be a linear transformation. If $T(\mathbf{v}_1) = \cdots = T(\mathbf{v}_n) = \mathbf{0}$, show that $T = 0$, the zero transformation from V to W.

$0 : V \to W$

$0(v) = 0 \quad \forall\, v \in V$

so $T(v_i) = 0 = 0(v_i)$ for all i

$\therefore \quad T = 0$

Theorem 7.1.9

Theorem

Let V and W be vector spaces and let $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ be a basis of V. Given any vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ in W (they need not be distinct), there exists a unique linear transformation $T : V \to W$ satisfying $T(\mathbf{b}_i) = \mathbf{w}_i$ for each $i = 1, 2, \ldots, n$. In fact, the action of T is as follows:
Given $\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n$ in V, v_i in $\mathbb{R}$, then

$$T(\mathbf{v}) = T(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = v_1\mathbf{w}_1 + v_2\mathbf{w}_2 + \cdots + v_n\mathbf{w}_n.$$

Linear Transformations and Basis

- This theorem shows that to define a linear transformation we can simply specify where the basis vectors go, and the rest of the action is dictated by the linearity
- Deciding whether two linear transformations are equal $\rightarrow$ determine whether they have the same effect on the basis vectors
- Given a basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of a vector space V, there is a different linear transformation $V \rightarrow W$ for every ordered selection $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ of vectors in W (not necessarily distinct)

Example 7.1.10

Find a linear transformation $T : \mathbf{P}_2 \to \mathbf{M}_{22}$ such that

$$T(1+x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad T(x+x^2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad T(1+x^2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

$\{1+x,\ x+x^2,\ 1+x^2\}$ → basis of P_2 (check)

Every $p \in P_2$ can be written as lin. comb:

$p = a + bx + cx^2$

$$p(x) = \tfrac{1}{2}(a+b-c)(1+x) + \tfrac{1}{2}(-a+b+c)(x+x^2) + \tfrac{1}{2}(a-b+c)(1+x^2)$$

To be cont.

Recap

Today we saw:

- Properties of linear transformations

Next time: Two important subspaces associated with linear transformations

MATH254: Linear Algebra
Lecture 14

Moira MacNeil

February 5, 2025

Last Time

1. Properties of linear transformations

Today

1. Finish the example from yesterday

2. Two important subspaces associated with linear transformations

Reminders:

- Midterm 1 is this Friday, February 7
- Assignment 3 is due Friday, February 14

Recall: Theorem 7.1.9

Theorem

Let V and W be vector spaces and let $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ be a basis of V. Given any vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ in W (they need not be distinct), there exists a unique linear transformation $T : V \to W$ satisfying $T(\mathbf{b}_i) = \mathbf{w}_i$ for each $i = 1, 2, \ldots, n$. In fact, the action of T is as follows:
Given $\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n$ in V, v_i in $\mathbb{R}$, then

$$T(\mathbf{v}) = T(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = v_1\mathbf{w}_1 + v_2\mathbf{w}_2 + \cdots + v_n\mathbf{w}_n.$$

Example 7.1.10 (again)

polynomials of deg. ≤ 2 $\longrightarrow$ 2×2 matrices

Find a linear transformation $T : \mathbf{P}_2 \to \mathbf{M}_{22}$ such that

$$T(1+x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad T(x+x^2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad T(1+x^2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Set $B = \{1+x, x+x^2, 1+x^2\}$ is a basis of P_2

Suppose there are scalars a, b, c such that

$(*)$ $a(1+x) + b(x+x^2) + c(1+x^2) = 0$

Collect our exponents of x to get system:

$a + c = 0$ Solve system:

$a + b = 0$ $a = -c$ $-2c = 0$

$b + c = 0$ $b = -c$ $c = 0$

$a = b = 0$

$\Rightarrow$ B is indept

dim $P_2 = 3$ and B has 3 vect.

$\Rightarrow$ B spans P_2

$\therefore$ B is a basis for P_2

Then every vector $p = a_0 + a_1 x + a_2 x^2$ in P_2 can be written as linear comb. of vectors in B.

System ⊛ but RHS coeff of p (not 0 vect).

$$p(x) = \tfrac{1}{2}(a+b-c)(1+x) + \tfrac{1}{2}(-a+b+c)(x+x^2) + \tfrac{1}{2}(a-b+c)(1+x^2)$$

Applying Thm

$$T(p(x)) = \tfrac{1}{2}(a+b-c)\,T(1+x) + \tfrac{1}{2}(-a+b+c)\,T(x+x^2) + \tfrac{1}{2}(a-b+c)\,T(1+x^2)$$

$$= \frac{1}{2}\left[(a+b-c)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (-a+b+c)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + (a-b+c)\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right]$$

$$= \frac{1}{2}\begin{bmatrix} a+b-c & -a+b+c \\ -a+b+c & a-b+c \end{bmatrix}$$

Given $p(x) = a + bx + cx^2$

$T(p(x)) =$

Kernel and Image

Kernel and Image of a Linear Transformation

The **kernel** of T (denoted $\ker T$) and the **image** of T (denoted $\operatorname{im} T$ or $T(V)$) are defined by

$$\ker T = \{\mathbf{v} \text{ in } V \mid T(\mathbf{v}) = \mathbf{0}\}$$
$$\operatorname{im} T = \{T(\mathbf{v}) \mid \mathbf{v} \text{ in } V\} = T(V)$$

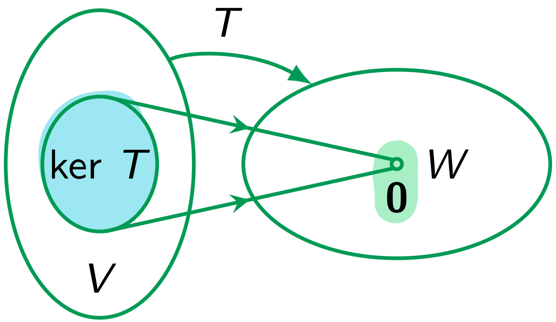
Nullspace and Range

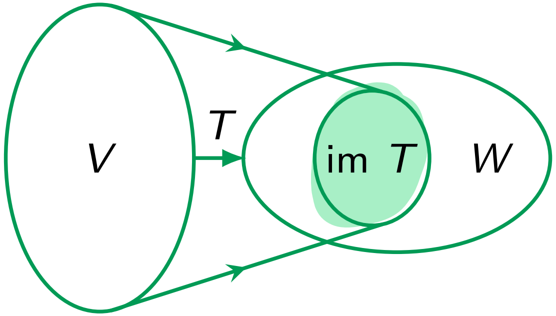
$T: V \to W$

$\ker T = \{ v \text{ in } V \mid T(v) = 0 \}$ in V

$\text{im } T = \{ T(v) \mid v \text{ in } V \}$ in W

- Kernel of T $\to$ called the **nullspace** of T because it consists of all vectors $\mathbf{v}$ in V satisfying the *condition* that $T(\mathbf{v}) = \mathbf{0}$
- Image of T $\to$ called the **range** of T and consists of all vectors $\mathbf{w}$ in W of the *form* $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V





Example 7.2.2

Let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation induced by the $m \times n$ matrix A, that is $T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$. Then

$$\ker\, T_A = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}\} = \text{null } A \quad \text{and}$$
$$\text{im } T_A = \{A\mathbf{x} \mid \mathbf{x} \text{ in } \mathbb{R}^n\} = \text{im } A$$

Theorem 7.2.3

Theorem

Let $T : V \to W$ be a linear transformation.

1. ker T is a subspace of V.
2. im T is a subspace of W.

Proof. Subspace Test

① $T(0) = 0$ $\therefore$ both ker T and im T contain vector of V and W respectively.

ker T let v $v_1 \in$ ker T, then $T(v) = 0 = T(v_1)$

② $T(v + v_1) = T(v) + T(v_1) = 0 + 0 = 0$

$v+v_1 \in \ker T$

3 let $r \in \mathbb{R}$

$T(rv) = rT(v) = r\,0 = 0$ true for any $r \in \mathbb{R}$

$rv \in \ker T$

$\therefore$ $\ker T$ is subspace of V

$\operatorname{im} T$ let $w, w_1 \in \operatorname{im} T$ so

$w = T(v)$ $w_1 = T(v_1)$ $v, v_1 \in V$

2 $w + w_1 = T(v) + T(v_1) = T(v+v_1)$

$w + w_1 \in \operatorname{im} T$

3 $rw = rT(v) = T(rv)$ $\forall r \in \mathbb{R}$

$rw \in \operatorname{im} T$

$\therefore$ $\operatorname{im} T$ is subspace of W

Nullity and Rank

Given a linear transformation $T : V \to W$.

- $\dim(\ker\ T)$ is called the **nullity** of T and denoted as $\text{nullity}(T)$
- $\dim(\text{im}\ T)$ is called the **rank** of T and denoted as $\text{rank}(T)$
- Recall: we defined the rank of a matrix A as dimension of col A, the column space of A

Example 7.2.4 $T_A : \mathbb{R}^n \to \mathbb{R}^m \quad T_A(x) = Ax$

Given an $m \times n$ matrix A, show that im $T_A =$ col A, so rank $T_A =$ rank A.

let $A = [c_1 \cdots c_n]$ (columns)

then im $T_A = \{Ax \mid x \in \mathbb{R}^n\}$

by def of matrix vect. multiplication $= \{x_1c_1 + x_2c_2 + \cdots + x_nc_n \mid x_i \in \mathbb{R}\}$

This is col space of A!

Example 7.2.5

Define a transformation $P : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ by $P(A) = A - A^T$ for all A in $\mathbf{M}_{nn}$. Show that P is linear and that:

(a) ker P consists of all symmetric matrices.

(b) im P consists of all skew-symmetric matrices.

P linear show satisfies axioms T1 and T2 (exercise)

(a) a matrix A is in ker P when

$0 = P(A) = A - A^T$ this holds when

$A = A^T \to$ i.e., A is symmetric.

(b) im P is all matrices $P(A)$ $A \in M_{nn}$

(skew sym $-A = A^T$)

$$P(A)^T = (A - A^T)^T = A^T - (A^T)^T$$
$$= A^T - A$$
$$= -(A - A^T)$$

Need to show all skew sym matrices are in im P

if S is skew sym then S lies in im P

$$P(\tfrac{1}{2}S) = (\tfrac{1}{2}S - \tfrac{1}{2}S^T) = \tfrac{1}{2}(S - S^T)$$
$$= \tfrac{1}{2}(S + S)$$
$$= S$$

Recap

Today we saw:

- Kernel and image
- One-to-one and onto transformations

Next time: Midterm 1

MATH254: Linear Algebra
Lecture 15

Moira MacNeil

February 11, 2025

Last Time

1. Midterm 1
2. Before that: kernel and image of transformations

Today

1. Review kernel and image
2. One-to-one and onto transformations
3. Review solutions for Midterm 1

Reminders:

- Assignment 3 is due this ~~Friday, February 14~~

Tuesday after break
Feb 25

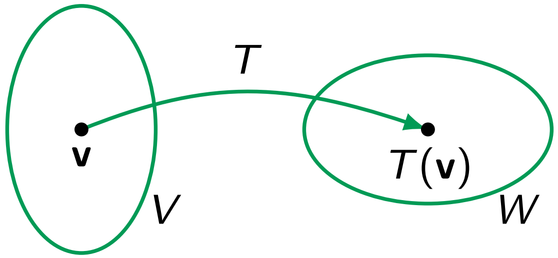
Recall: Linear Transformation

Linear Transformations of Vector Spaces

If V and W are two vector spaces, a function $T : V \to W$ is called a **linear transformation** if it satisfies the following axioms.

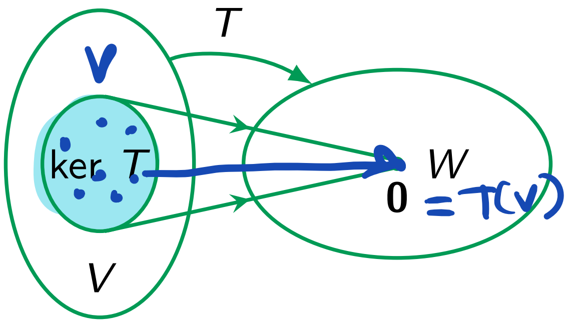
T1.	$T(\mathbf{v}+\mathbf{v}_1) = T(\mathbf{v}) + T(\mathbf{v}_1)$	for all $\mathbf{v}$ and $\mathbf{v}_1$ in V.
T2.	$T(r\mathbf{v}) = rT(\mathbf{v})$	for all $\mathbf{v}$ in V and r in $\mathbb{R}$.

A linear transformation $T : V \to V$ is called a **linear operator** on V.

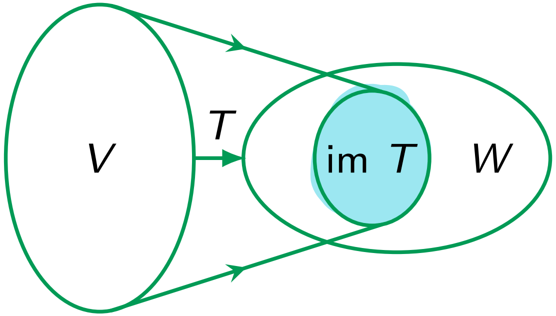


Recall: Kernel and Image of a Linear Transformation

- ker $T = \{\mathbf{v} \text{ in } V \mid T(\mathbf{v}) = \mathbf{0}\}$
- Kernel of T $\rightarrow$ called the **nullspace** of T
- Consists of all vectors $\mathbf{v}$ in V satisfying the *condition* that $T(\mathbf{v}) = \mathbf{0}$
- ker T is a subspace of V



- im $T = \{T(\mathbf{v}) \mid \mathbf{v} \text{ in } V\} = T(V)$
- Image of T $\rightarrow$ called the **range** of T
- Consists of all vectors $\mathbf{w}$ in W of the *form* $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V
- im T is a subspace of W



Recall: Nullity and Rank

Given a linear transformation $T : V \to W$.

- $\dim(\ker T)$ is called the **nullity** of T and denoted as $\text{nullity}(T)$
- $\dim(\text{im } T)$ is called the **rank** of T and denoted as $\text{rank}(T)$

Num vectors that are transformed to 0

size of image space

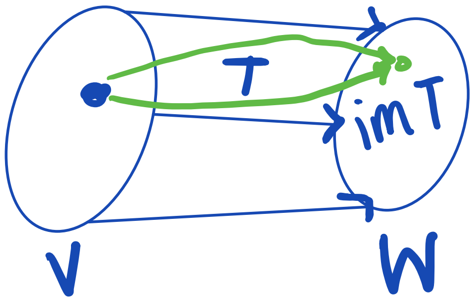
Num of vectors in W that are reached by T(v)

One-to-one and Onto Transformations

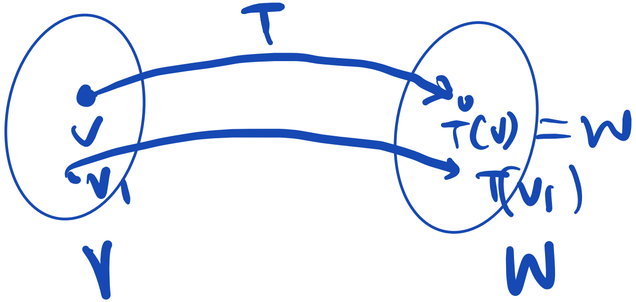
One-to-one and Onto Linear Transformations

Let $T : V \to W$ be a linear transformation.

1. T is said to be **onto** if $\text{im } T = W$.
2. T is said to be **one-to-one** if $T(\mathbf{v}) = T(\mathbf{v}_1)$ implies $\mathbf{v} = \mathbf{v}_1$.



ONTO



One-to-one and Onto Linear Transformations

- A vector $\mathbf{w}$ in W is **hit** by T if $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V
- T is onto if every vector in W is hit at least once
- T is one-to-one if no element of W gets hit twice
- Onto transformations T are those for which im $T = W$ is as large a subspace of W as possible → entire vector space
- One-to-one transformations T are the ones with ker T as small a subspace of V as possible

Theorem 7.2.7

Theorem

If $T : V \to W$ is a linear transformation, then T is one-to-one if and only if $\ker\ T = \{\mathbf{0}\}$.

Proof. ($\Rightarrow$) T is 1-1 let $v \in \ker(T)$ then $T(v) = 0$, so $T(v) = T(0)$ since T is 1-1. $\therefore$ $\ker T = \{0\}$

($\Leftarrow$) Assume $\ker T = \{0\}$ let $T(v) = T(v_1)$ where $v, v_1 \in V$. $T(v - v_1) = T(v) - T(v_1) = 0$ so $v - v_1 \in \ker T$. Then $v - v_1 = 0 \Rightarrow v = v_1$ $\therefore$ T is one-to-one. $\square$

Example 7.2.8

The identity transformation $1_V : V \to V$ is both one-to-one and onto for any vector space V.

$$v \to 1_V(v) = v$$

Example 7.2.9

→ Satisfy axioms T1, T2

Consider the linear transformations

$$S : \mathbb{R}^3 \to \mathbb{R}^2 \quad \text{given by } S(x, y, z) = (x + y, x - y)$$
$$T : \mathbb{R}^2 \to \mathbb{R}^3 \quad \text{given by } T(x, y) = (x + y, x - y, x)$$

Show that T is one-to-one but not onto, whereas S is onto but not one-to-one.

T 1-1 means ker T = {0} zero vector in $\mathbb{R}^2$ is (0,0)

Ker T = {(x,y) | x+y = x-y = x = 0} = {(0,0)}

Not onto (counter example)

(0,0,1) is not in im T because if

(0,0,1) = (x+y, x-y, x) →

$x+y = x-y = 0$ and $x=1$ $\rightarrow$ impossible.

S is not 1-1 since $(0,0,1)$ is in $\ker S$

Every element (s,t) in $\mathbb{R}^2$ is in $\operatorname{im} S$

$(s,t) = (x+y, x-y) = S(x,y,z)$ for some x,y,z

$\rightarrow$ $x = \frac{1}{2}(s+t)$ $y = \frac{1}{2}(s-t)$ $z=0$

$\therefore$ S is onto.

Theorem 7.2.11

Theorem

Let A be an $m \times n$ matrix, and let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation induced by A, that is $T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$.

1. T_A is onto if and only if rank $A = m$.
2. T_A is one-to-one if and only if rank $A = n$.

Recap

Today we saw:

- Kernel and image
- One-to-one and onto transformations

Next time: Dimension Theorem

MATH254: Linear Algebra
Lecture 16

Moira MacNeil

February 12, 2025

Last Time

1. One-to-one and onto transformations

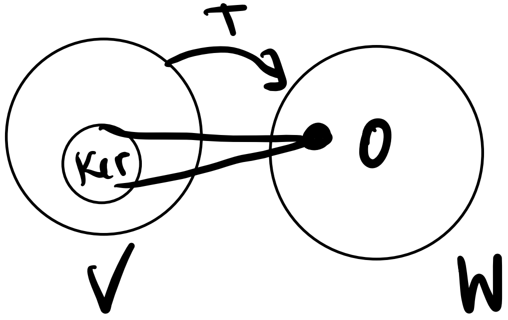
Today

1. Dimension Theorem
2. Isomorphisms

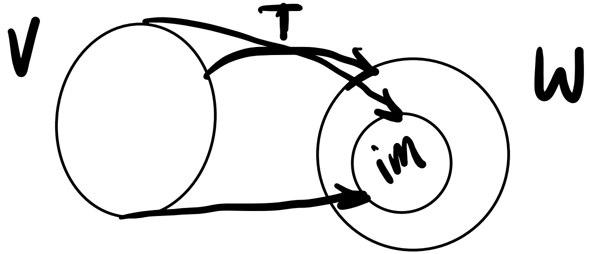
Reminders:

- Assignment 3 is now due Tuesday, February 25

Recall: Nullity and Rank



Given a linear transformation $T : V \to W$



- $\ker\ T = \{\mathbf{v} \text{ in } V \mid T(\mathbf{v}) = \mathbf{0}\}$
- $\text{im}\ T = \{T(\mathbf{v}) \mid \mathbf{v} \text{ in } V\} = T(V)$
- $\dim(\ker\ T)$ is called the **nullity** of T and denoted as nullity(T)
- $\dim(\text{im}\ T)$ is called the **rank** of T and denoted as rank(T)

One-to-one and Onto Transformations

One-to-one and Onto Linear Transformations

Let $T : V \to W$ be a linear transformation.

1. T is said to be **onto** if $\operatorname{im} T = W$.
2. T is said to be **one-to-one** if $T(\mathbf{v}) = T(\mathbf{v}_1)$ implies $\mathbf{v} = \mathbf{v}_1$.

- T is onto if every vector in W is hit at least once
- T is one-to-one if no element of W gets hit twice

Recall: Theorem 7.2.7

Theorem

If $T : V \to W$ is a linear transformation, then T is one-to-one if and only if $\ker T = \{\mathbf{0}\}$.

Theorem 7.2.12

Dimension Theorem

Let $T : V \to W$ be any linear transformation and assume that ker T and im T are both finite dimensional. Then V is also finite dimensional and

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T)$$

In other words, $\dim V = \text{nullity}(T) + \text{rank}(T)$.

Theorem 7.2.13

Theorem

Let $T : V \to W$ be a linear transformation, and let $\{\mathbf{e}_1, \ldots, \mathbf{e}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ be a basis of V such that $\{\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ is a basis of $\ker T$. Then $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_r)\}$ is a basis of $\operatorname{im} T$, and hence $r = \operatorname{rank} T$.

If either $\dim(\ker T)$ or $\dim(\operatorname{im} T)$ can be found, then the other is automatically known!

Example 7.2.15

If $T : V \to W$ is a linear transformation where V is finite dimensional, then

$$\dim(\ker T) \leq \dim V \quad \text{and} \quad \dim(\operatorname{im} T) \leq \dim V$$

Indeed, $\dim V = \dim(\ker T) + \dim(\operatorname{im} T)$ by Dimension Theorem. Of course, the first inequality also follows because $\ker T$ is a subspace of V.

Example 7.2.16

$P_n \rightarrow$ poly. of degree n

Let $D : \mathbf{P}_n \to \mathbf{P}_{n-1}$ be the differentiation map defined by $D[p(x)] = p'(x)$. Compute ker D and hence conclude that D is onto.

$p'(x) = 0$ then $p(x) =$ constant so

we have $\dim(\ker D) = 1$

Since $\dim(P_n) = n+1$

By dimension thm

$$\dim(\operatorname{im} D) = \dim(P_n) - \dim(\ker D)$$

$$n+1 \quad - \quad 1$$

$$= n = \dim(P_{n-1})$$

$\therefore \operatorname{im} D = P_{n-1}$ so D is onto

Example 7.2.17

$p(x) = a_0 + a_1 x + \cdots + a_n x^n$
$q(x) = b_0 + b_1 x + \cdots + b_n x^n$

Given a in $\mathbb{R}$, the evaluation map $E_a : \mathbf{P}_n \to \mathbb{R}$ is given by $E_a[p(x)] = p(a)$. Show that E_a is linear and onto, and hence conclude that $\{(x-a), (x-a)^2, \ldots, (x-a)^n\}$ is a basis of ker E_a, the subspace of all polynomials $p(x)$ for which $p(a) = 0$.

$\underline{E_a}$ Linear

T1: $T(v + v_1) = T(v) + T(v_1)$

$E_a(p(x) + q(x)) = E_a[(a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n]$
$= (a_0 + b_0) + (a_1 + b_1)a + \cdots + (a_n + b_n)a^n$
$= (a_0 + a_1 a + \cdots + a_n a^n) + (b_0 + b_1 a + \cdots + b_n a^n)$
$= E_a(p(x)) + E_a(q(x))$

T2: $T(rv) = rT(v)$ let $r \in \mathbb{R}$

$E_a(r\,p(x)) = E_a(ra_0 + ra_1x + \cdots + ra_nx^n)$

$= ra_0 + ra_1a + \cdots + ra_na^n$

$= r(a_0 + a_1a + \cdots + a_na^n)$

$= r\,E_a(p(x))$

<u>onto</u> $\dim(\operatorname{im} E_a) = \dim(\mathbb{R}) = 1$

so $\dim(\ker E_a) = \underset{P_n}{n+1} - \underset{\mathbb{R}}{1} = n$ by dim thm

Each of the n polynomials

$(x-a), (x-a)^2, \ldots, (x-a)^n$ is in $\ker E_a$

since $p(a) = 0$ for any of them

Linearly independent since they have distinct degrees.

n vectors, $\dim(\ker(E_a)) = n$

$\therefore$ these polynomials are a basis for the $\ker(E_a)$.

Example 7.2.18

If A is any $m \times n$ matrix, show that rank $A =$ rank $A^TA =$ rank AA^T.

rank A = rank $A^T A$ let $B = A^T A$

$T_A : \mathbb{R}^n \to \mathbb{R}^m$
$x \mapsto Ax$

$T_B : \mathbb{R}^n \to \mathbb{R}^n$
$x \mapsto Bx$

im T_A = col A
rank T_A = rank A

im T_B = col B
rank T_B = rank B

rank A = rank T_A = dim(im T_A)
= n − dim(ker T_A)

$\operatorname{rank} B = \operatorname{rank} T_B = \dim(\operatorname{im} T_B)$
$= n - \dim(\operatorname{Ker} T_B)$

it suffices to show $\ker(T_A) = \ker(T_B)$

$Ax = 0 \Rightarrow Bx = A^T A x = A^T(0) = 0$

$\operatorname{Ker} T_A \subseteq \ker T_B$

$Bx = 0 \Rightarrow A^T A x = 0$

$$\|Ax\|^2 = (Ax)^T (Ax) = x^T \overbrace{A^T A}^{B} x$$
$$= x^T(0) = 0$$

$\Rightarrow Ax = 0$

$\operatorname{Ker} T_B \subseteq \ker T_A$

$\therefore \ \ker T_A = \operatorname{Ker} T_B \Rightarrow \operatorname{rank} A = \operatorname{rank} B$

When are two vectors spaces the "same"?

- Consider the spaces

$$\mathbb{R}^2 = \{(a, b) \mid a, b \in \mathbb{R}\} \quad \text{and} \quad \mathbf{P}_1 = \{a + bx \mid a, b \in \mathbb{R}\}$$

- Compare the addition and scalar multiplication in these spaces:

$$\begin{aligned} (a, b) + (a_1, b_1) &= (a + a_1, b + b_1) & \quad (a + bx) + (a_1 + b_1 x) &= (a + a_1) + (b + b_1)x \\ r(a, b) &= (ra, rb) & r(a + bx) &= (ra) + (rb)x \end{aligned}$$

- These are the same vector space expressed in different notation: Change each (a, b) in $\mathbb{R}^2$ to $a + bx$, then $\mathbb{R}^2$ *becomes* $\mathbf{P}_1$

- The map $(a, b) \mapsto a + bx$ is a linear transformation $\mathbb{R}^2 \to \mathbf{P}_1$ that is both one-to-one and onto

Isomorphisms

Isomorphic Vector Spaces

A linear transformation $T : V \to W$ is called an **isomorphism** if it is both onto and one-to-one. The vector spaces V and W are said to be **isomorphic** if there exists an isomorphism $T : V \to W$, and we write $V \cong W$ when this is the case.

Example 7.3.2

The identity transformation $1_V : V \to V$ is an isomorphism for any vector space V.

$$v \mapsto v$$

Example 7.3.3

If $T : \mathbf{M}_{mn} \to \mathbf{M}_{nm}$ is defined by $T(A) = A^T$ for all A in $\mathbf{M}_{mn}$, then T is an isomorphism. Hence $\mathbf{M}_{mn} \cong \mathbf{M}_{nm}$.

m×n zero matrix

Verify: 1-1 $\ker(T) = \{0\}$

$A^T = 0 \to A = 0$

Verify: onto dimension thm

notice that $\dim(M_{mn}) = mn$

$= \dim(M_{nm})$

Example 7.3.4

Isomorphic spaces can “look” quite different. For example, $\mathbf{M}_{22} \cong \mathbf{P}_3$ because the map $T : \mathbf{M}_{22} \to \mathbf{P}_3$ given by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a + bx + cx^2 + dx^3$ is an isomorphism.

Verify linear transformation
verify 1-1 and onto

Theorem 7.3.5

Theorem

If V and W are finite dimensional spaces, the following conditions are equivalent for a linear transformation $T : V \to W$.

1. T is an isomorphism.
2. If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is any basis of V, then $\{T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)\}$ is a basis of W.
3. There exists a basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of V such that $\{T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)\}$ is a basis of W.

Isomorphisms are the linear transformations that preserve bases.

Recap

Today we saw:

- Dimension Theorem
- Isomorphisms

Next time: More on isomorphisms, composition

MATH254: Linear Algebra
Lecture 17

Moira MacNeil

February 14, 2025

Last Time

1. Dimension Theorem

2. Isomorphisms

Today

1. Isomorphisms
2. Composition of linear transformations

Reminders:

- Assignment 3 is now due Tuesday, February 25

Recall: Isomorphisms

Isomorphic Vector Spaces

A linear transformation $T : V \to W$ is called an **isomorphism** if it is both onto and one-to-one. The vector spaces V and W are said to be **isomorphic** if there exists an isomorphism $T : V \to W$, and we write $V \cong W$ when this is the case.

- An isomorphism $T : V \to W$ induces a pairing $\mathbf{v} \leftrightarrow T(\mathbf{v})$ between vectors $\mathbf{v}$ in V and vectors $T(\mathbf{v})$ in W that preserves vector addition and scalar multiplication
- The spaces V and W are identical except for notation, that is all vector space properties of either space are completely determined by those of the other

Theorem 7.3.5

Theorem

If V and W are finite dimensional spaces, the following conditions are equivalent for a linear transformation $T : V \to W$.

1. T is an isomorphism.
2. If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is any basis of V, then $\{T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)\}$ is a basis of W.
3. There exists a basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of V such that $\{T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)\}$ is a basis of W.

Isomorphisms are the linear transformations that preserve bases!

Theorem 7.3.6

Theorem

If V and W are finite dimensional vector spaces, then $V \cong W$ if and only if $\dim V = \dim W$.

Corollary 7.3.7

Corollary

Let U, V, and W denote vector spaces. Then:

1. $V \cong V$ for every vector space V.
2. If $V \cong W$ then $W \cong V$.
3. If $U \cong V$ and $V \cong W$, then $U \cong W$.

The relation $\cong$ is called an equivalence relation on the class of finite dimensional vector spaces.

Corollary 7.3.8

Corollary

If V is a vector space and $\dim V = n$, then V is isomorphic to $\mathbb{R}^n$.

Coordinate Isomorphism

- If V is a vector space of dimension n, there are important explicit isomorphisms $V \to \mathbb{R}^n$
- Fix a basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ of V and write the standard basis of $\mathbb{R}^n$ $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$
- There is a unique linear transformation $C_B : V \to \mathbb{R}^n$ given by

$$C_B(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

  where each v_i is in $\mathbb{R}$
- Moreover, $C_B(\mathbf{b}_i) = \mathbf{e}_i$ for each i so C_B is an isomorphism, called the **coordinate isomorphism** corresponding to the basis B

Example 7.3.9

Let V denote the space of all 2×2 symmetric matrices. Find an isomorphism $T : \mathbf{P}_2 \to V$ such that $T(1) = I$, where I is the 2×2 identity matrix.

$\{1, x, x^2\}$ is a basis of P_2

Want a basis of V containing I

The set $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is indept. in V (check)

it has dim 3 $\therefore$ a basis since $\dim V = 3$ (previous Eg)

Define Transformation $T: P_2 \to V$

$T(1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $T(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $T(x^2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

$$p(x) = a + bx + cx^2$$

$$M = a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{aligned} T(a+bx+cx^2) &= T(a) + T(bx) + T(cx^2) \\ &= a\,T(1) + b\,T(x) + c\,T(x^2) \\ &= \begin{bmatrix} a & b \\ b & a+c \end{bmatrix} \end{aligned}$$

Theorem 7.3.10

Theorem

If V and W have the same dimension n, a linear transformation $T : V \to W$ is an isomorphism if it is either one-to-one or onto.

Proof. Dimension Thm

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T) = n$$

so $\dim(\ker T) = 0 \iff \dim(\operatorname{im} T) = n$

thus T is one-to-one <u>iff</u> T is onto.

Composition

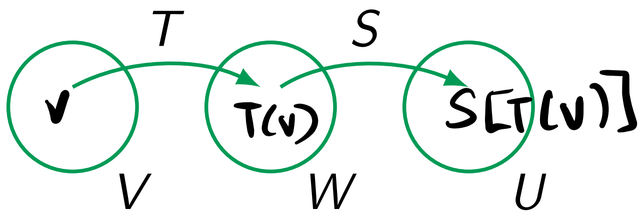
$T : V \to W \qquad S : W \to U$

Composition of Linear Transformations

Given linear transformations $V \xrightarrow{T} W \xrightarrow{S} U$, the **composite** $ST : V \to U$ of T and S is defined by

$$ST(\mathbf{v}) = S\,[T(\mathbf{v})] \quad \text{for all } \mathbf{v} \text{ in } V$$

The operation of forming the new function ST is called **composition** (sometimes denoted $S \circ T$).



Not all linear transformations can be composed

- If $T : V \to W$ and $S : W \to U$ are linear transformations then $ST : V \to U$ is defined, but TS *cannot* be formed unless $U = V$
- Even if ST and TS can both be formed, they may not be equal
- For example, if $S : \underbrace{\mathbb{R}^m}_{V} \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \underbrace{\mathbb{R}^m}_{U}$ are induced by matrices A and B respectively, then ST and TS can both be formed (they are induced by AB and BA respectively), but the matrix products AB and BA may not be equal (they may not even be the same *size*)

$T: \mathbb{R}^n \to \mathbb{R}^m \qquad S: \mathbb{R}^m \to \mathbb{R}^n$

$ST: x \mapsto ABx \qquad TS: x \mapsto BAx$

Example 7.3.12

Define: $S : \mathbf{M}_{22} \to \mathbf{M}_{22}$ and $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ by $S\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$ and $T(A) = A^T$ for $A \in \mathbf{M}_{22}$. Describe the action of ST and TS, and show that $ST \neq TS$.

$$S\left[T\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right] = S\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} b & d \\ a & c \end{bmatrix}$$

$$TS\begin{bmatrix} a & b \\ c & d \end{bmatrix} = T\begin{bmatrix} c & d \\ a & b \end{bmatrix} = \begin{bmatrix} c & a \\ d & b \end{bmatrix}$$

in general $ST \neq TS$

Theorem 7.3.13

$$1_V : \begin{array}{c} V \to V \\ v \mapsto v \end{array}$$

Theorem

Let $V \xrightarrow{T} W \xrightarrow{S} U \xrightarrow{R} Z$ be linear transformations.

1. The composite ST is again a linear transformation.
2. $T1_V = T$ and $1_W T = T$.
3. $(RS)T = R(ST)$.

Theorem 7.3.14

Theorem

Let V and W be finite dimensional vector spaces. The following conditions are equivalent for a linear transformation $T : V \to W$.

(a) T is an isomorphism.

(b) There exists a linear transformation $S : W \to V$ such that $ST = 1_V$ and $TS = 1_W$.

Moreover, in this case S is also an isomorphism and is uniquely determined by T:

$$\text{If } \mathbf{w} \text{ in } W \text{ is written as } \mathbf{w} = T(\mathbf{v}), \text{ then } S(\mathbf{w}) = \mathbf{v}.$$

Inverse of a Linear Transformation

- Given an isomorphism $T : V \to W$, the unique isomorphism $S : W \to V$ satisfying condition (b) of the previous Theorem is called the **inverse of T** and is denoted by T^{-1}.
- Hence $T : V \to W$ and $T^{-1} : W \to V$ are related by the **fundamental identities**:

$$T^{-1}\left[T(\mathbf{v})\right] = \mathbf{v} \text{ for all } \mathbf{v} \text{ in } V \quad \text{and} \quad T\left[T^{-1}(\mathbf{w})\right] = \mathbf{w} \text{ for all } \mathbf{w} \text{ in } W$$

- Each of T and T^{-1} reverses the action of the other.

Example 7.3.15

Define $T : \mathbf{P}_1 \to \mathbf{P}_1$ by $T(a + bx) = (a - b) + ax$. Show that T has an inverse, and find the action of T^{-1}.

First step is to check T is linear. Preserve addition and scalar multiplication.

$T(1+0x) = 1+x$ $\quad$ $T(0+1x) = -1$

$B = \{1, x\}$ is a basis of P_1 so T takes B to $D = \{1+x, -1\}$ T is isomorphism.

T^{-1} takes D back to B

$T^{-1}(1+x) = 1$ $\quad$ $T^{-1}(-1) = x$

$a + bx = b(1+x) + (b-a)(-1)$ so

$T^{-1}(a+bx) = b\,T^{-1}(1+x) + (b-a)\,T^{-1}(-1)$
$= b + (b-a)\,x$

Example 7.3.16

If $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is a basis of a vector space V, the coordinate transformation $C_B : V \to \mathbb{R}^n$ is an isomorphism defined by

$$C_B(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = (v_1, v_2, \ldots, v_n)^T$$

The way to reverse the action of C_B is clear: $C_B^{-1} : \mathbb{R}^n \to V$ is given by

$$C_B^{-1}(v_1, v_2, \ldots, v_n) = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n \quad \text{for all } v_i \text{ in } V$$

Example 7.3.17

Define $T : \mathbb{R}^3 \to \mathbb{R}^3$ by $T(x, y, z) = (z, x, y)$. Show that $T^3 = 1_{\mathbb{R}^3}$, and hence find T^{-1}.

$$T^2(x,y,z) = T[T(x,y,z)] = T(z,x,y)$$
$$= (y,z,x)$$

$$T^3(x,y,z) = T(y,z,x) = (x,y,z)$$

$$\therefore T^3 = 1_{\mathbb{R}^3} = T(T^2)$$

Thus $T^{-1} = T^2$

Example 7.3.18

Define $T : \mathbf{P}_n \to \mathbb{R}^{n+1}$ by $T(p) = (p(0), p(1), \ldots, p(n))$ for all p in $\mathbf{P}_n$. Show that T^{-1} exists.

$T(p) = 0$ $\quad p(k) = 0$ for $k = 0, \ldots, n$

so p has $n+1$ distinct roots

because p is deg at most n, $n+1$ distinct roots $\Rightarrow$ $p = 0$ $\therefore$ T is 1-1

$\dim P_n = n+1 = \dim \mathbb{R}^{n+1}$ so T is also onto

$\therefore$ T is an iso. and T^{-1} exists.

Recap

Today we saw:

- Isomorphisms
- Composition

Next time: A new chapter - Orthogonality

MATH254: Linear Algebra
Lecture 18

Moira MacNeil

February 25, 2025

Today

1. Orthogonal Complements and Projections

Reminders:

- Assignment 4 is posted and due Friday March 7

Lemma 8.1.1

Orthogonal Lemma

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal set in $\mathbb{R}^n$. Given $\mathbf{x}$ in $\mathbb{R}^n$, write

$$\mathbf{f}_{m+1} = \mathbf{x} - \frac{\mathbf{x}\cdot\mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}\cdot\mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\mathbf{x}\cdot\mathbf{f}_m}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Then:

1. $\mathbf{f}_{m+1}\cdot\mathbf{f}_k = 0$ for $k = 1, 2, \ldots, m$. orthog. to all vect in orthog. set
2. If $\mathbf{x}$ is not in span$\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$, then $\mathbf{f}_{m+1} \neq \mathbf{0}$ and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is an orthogonal set.

Theorem 8.1.2

Theorem

Let U be a subspace of $\mathbb{R}^n$.

1. Every orthogonal subset $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ in U is a subset of an orthogonal basis of U.
2. U has an orthogonal basis.

Gram-Schmidt Orthogonalization Algorithm

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of a subspace U of $\mathbb{R}^n$, construct $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m$ in U successively as follows:

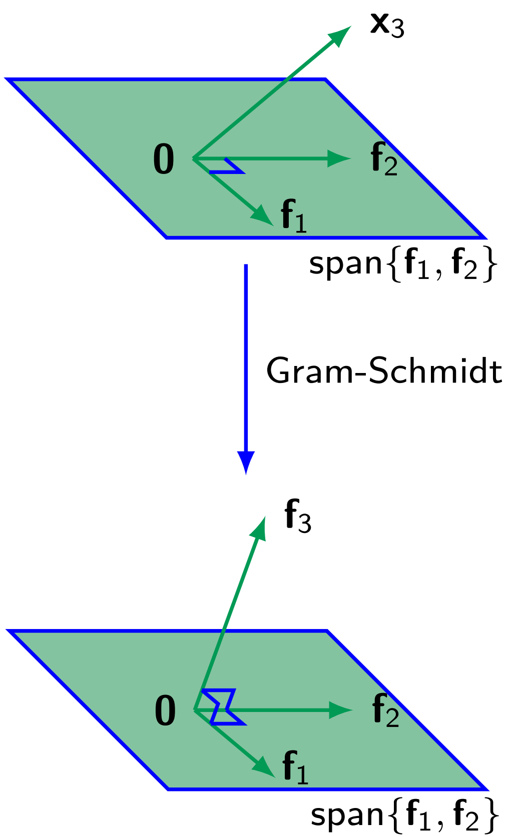
converts any basis to an orthog. basis

$$
\begin{aligned}
\mathbf{f}_1 &= \mathbf{x}_1 \\
\mathbf{f}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
&\vdots \\
\mathbf{f}_k &= \mathbf{x}_k - \frac{\mathbf{x}_k \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_k \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\mathbf{x}_k \cdot \mathbf{f}_{k-1}}{\|\mathbf{f}_{k-1}\|^2}\mathbf{f}_{k-1}
\end{aligned}
$$

for each $k = 2, 3, \ldots, m$. Then

1. $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is an orthogonal basis of U.
2. $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ for each $k = 1, 2, \ldots, m$.

A Remark on the Gram-Schmidt Orthogonalization Algorithm



- The vector $\frac{\mathbf{x}\cdot\mathbf{f}_i}{||\mathbf{f}_i||^2}\mathbf{f}_i$ is unchanged if a nonzero scalar multiple of $\mathbf{f}_i$ is used in place of $\mathbf{f}_i$
- If a newly constructed $\mathbf{f}_i$ is multiplied by a nonzero scalar at some stage of the Gram-Schmidt algorithm, the subsequent $\mathbf{f}$s will be unchanged
- Very useful in actual calculations

Example 8.1.4

Find an orthogonal basis of the row space of $A = \begin{array}{c c} \left[\begin{array}{rrrr} 1 & 1 & -1 & -1 \\ 3 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right] & \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \end{array}$

Observe $\{x_1, x_2, x_3\}$ is indep. (verify).

Use G.S.

$f_1 = x_1$

$f_2 = x_2 - \frac{x_2 \cdot f_1}{\|f_1\|^2} f_1 = (3,2,0,1) - \frac{4}{4}(1,1,-1,-1)$

$= (2, 2, -1, 2)$

$f_3 = x_3 - \frac{x_3 \cdot f_1}{\|f_1\|^2} f_1 - \frac{x_3 \cdot f_2}{\|f_2\|^2} f_2$

$$= (1,0,1,0) - \frac{0}{4} f_1 - \frac{1}{13}(2,2,-1,2)$$

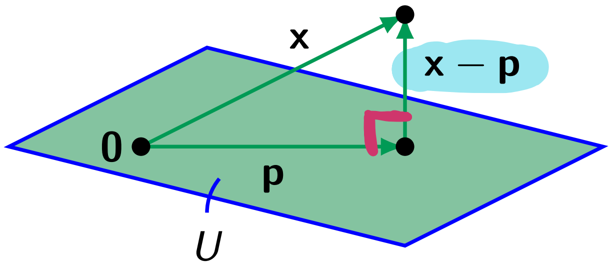
$$= (1,0,1,0) - \left(\frac{2}{13}, \frac{2}{13}, \frac{-1}{13}, \frac{2}{13}\right)$$

$$= \left(\frac{11}{13}, \frac{-2}{13}, \frac{14}{13}, \frac{11}{13}\right)$$

$$\Rightarrow (11, -2, 14, 11)$$

Point Closest to a Plane

- Given a point $\mathbf{x}$ and a plane U through the origin in $\mathbb{R}^3$, we want to find the point $\mathbf{p}$ in the plane that is closest to $\mathbf{x}$
- Geometrically, $\mathbf{p}$ must be chosen such that $\mathbf{x} - \mathbf{p}$ is perpendicular to the plane
- U is a subspace of $\mathbb{R}^3$ (it contains the origin)
- $\mathbf{x} - \mathbf{p}$ is perpendicular to U means that $\mathbf{x} - \mathbf{p}$ is orthogonal to every vector in U



Orthogonal Complement

Orthogonal Complement of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$, define the **orthogonal complement** $U^\perp$ of U (pronounced "U-perp") by

$$U^\perp = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{y} = 0 \text{ for all } \mathbf{y} \text{ in } U\}$$

The set of vectors that are orthogonal to every vector in U.

Lemma 8.1.6

Lemma

Let U be a subspace of $\mathbb{R}^n$.

1. $U^\perp$ is a subspace of $\mathbb{R}^n$.
2. $\{\mathbf{0}\}^\perp = \mathbb{R}^n$ and $(\mathbb{R}^n)^\perp = \{\mathbf{0}\}$.
3. If $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$, then $U^\perp = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{x}_i = 0 \text{ for } i = 1, 2, \ldots, k\}$.

Example 8.1.7

Find $U^\perp$ if $U = \text{span}\{(1,-1,2,0),(1,0,-2,3)\}$ in $\mathbb{R}^4$.

$x_1 = (x, y, z, w) \in U^\perp$ iff it is orthog. to both $x_2 = (1,-1,2,0)$ and $x_3 = (1,0,-2,3)$

$$x_1 \cdot x_2 = x - y + 2z = 0$$
$$x_1 \cdot x_3 = x - 2z + 3w = 0$$

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

let $z = s$
$w = t$

$y = 4s - 3t$
$x = -2s + 4s - 3t$

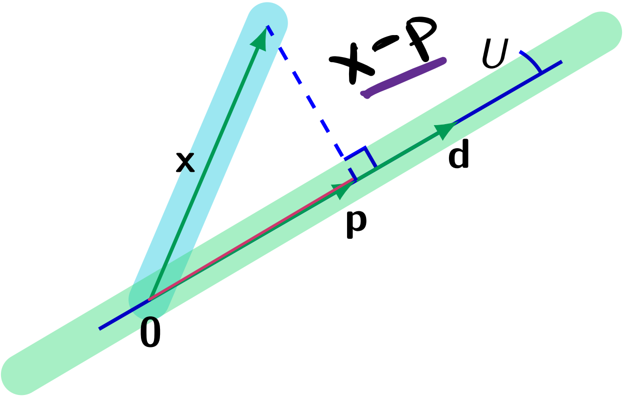
General Sol

$$s \begin{bmatrix} 2 \\ 4 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

$$U^{\perp} = \operatorname{span}\{(2,4,1,0), (-3,-3,0,1)\}$$

Recall: Projection in $\mathbb{R}^3$

- Consider $\mathbf{x}$ and $\mathbf{d} \neq \mathbf{0}$ in $\mathbb{R}^3$
- The projection $\mathbf{p} = \text{proj}_{\mathbf{d}}\mathbf{x}$ of $\mathbf{x}$ on $\mathbf{d}$ is $\text{proj}_{\mathbf{d}}\mathbf{x} = \left(\frac{\mathbf{x}\cdot\mathbf{d}}{\|\mathbf{d}\|^2}\right)\mathbf{d}$ where $\mathbf{x} - \mathbf{p}$ is orthogonal to $\mathbf{d}$
- The line $U = \mathbb{R}\mathbf{d} = \{t\mathbf{d} \mid t \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$, $\{\mathbf{d}\}$ is an orthogonal basis of U, and $\mathbf{p} \in U$ and $\mathbf{x} - \mathbf{p} \in U^{\perp}$
- We can generalize this for any vector $\mathbf{x}$ in $\mathbb{R}^n$ and any subspace U of $\mathbb{R}^n$



Projection onto a Subspace

Projection onto a Subspace of $\mathbb{R}^n$

Let U be a subspace of $\mathbb{R}^n$ with orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$. If $\mathbf{x}$ is in $\mathbb{R}^n$, the vector

$$\text{proj}_U \mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\mathbf{x} \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\mathbf{x} \cdot \mathbf{f}_m}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

is called the **orthogonal projection** of $\mathbf{x}$ on U. For the zero subspace $U = \{\mathbf{0}\}$, we define

$$\text{proj}_{\{\mathbf{0}\}} \mathbf{x} = \mathbf{0}$$

Theorem 8.1.9

Projection Theorem

If U is a subspace of $\mathbb{R}^n$ and $\mathbf{x}$ is in $\mathbb{R}^n$, write $\mathbf{p} = \text{proj}_U \mathbf{x}$. Then:

1. $\mathbf{p}$ is in U and $\mathbf{x} - \mathbf{p}$ is in $U^\perp$.
2. $\mathbf{p}$ is the vector in U closest to $\mathbf{x}$ in the sense that

$$\|\mathbf{x} - \mathbf{p}\| < \|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{y} \in U, \mathbf{y} \neq \mathbf{p}$$

Example 8.1.10

Let $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$ in $\mathbb{R}^4$ where $\mathbf{x}_1 = (1,1,0,1)$ and $\mathbf{x}_2 = (0,1,1,2)$. If $\mathbf{x} = (3,-1,0,2)$, find the vector in U closest to $\mathbf{x}$ and express $\mathbf{x}$ as the sum of a vector in U and a vector orthogonal to U.

$\{x_1, x_2\}$ indep. (verify), not orthog.

Use GS to find orthog. basis is $\{f_1, f_2\}$

$f_1 = x_1 = (1,1,0,1)$

$$f_2 = x_2 - \frac{x_2 \cdot f_1}{\|f_1\|^2} f_1 = (0,1,1,2) - \frac{3}{3}(1,1,0,1)$$

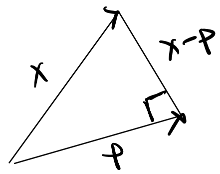
$$= (-1, 0, 1, 1)$$

Compute $\text{proj}_U x = p = \dfrac{x \cdot f_1}{\|f_1\|^2} f_1 + \dfrac{x \cdot f_2}{\|f_2\|^2} f_2$

p is pt in U
closest to x

$$= \frac{4}{3} f_1 + \frac{-1}{3} f_2$$

$$= \tfrac{1}{3}(5, 4, -1, 3)$$



$$x - p = \tfrac{1}{3}(4, -7, 1, 3)$$

$$x = p + (x-p) = \tfrac{1}{3}(5, 4, -1, 3) + \tfrac{1}{3}(4, -7, 1, 3)$$

Example 8.1.11

Find the point in the plane with equation $2x + y - z = 0$ that is closest to the point $(2, -1, -3)$. $= x$

Plane is subspace U of $\mathbb{R}^3$ whose pts satisfy $z = 2x + y$

$$U = \{(s, t, 2s+t) \mid s, t \in \mathbb{R}\}$$

$$= \text{span}\{(0,1,1), (1,0,2)\}$$

Use GS to get orthog. basis

$$f_1 = (0,1,1)$$

$$f_2 = (1,0,2) - \frac{2}{2}(0,1,1) = (1,-1,1)$$

$$\operatorname{proj}_U x = \frac{x \cdot f_1}{\|f_1\|^2} f_1 + \frac{x \cdot f_2}{\|f_2\|^2} f_2$$

$$= \frac{-4}{2} f_1 + 0 f_2$$

$$= (0, -2, -2) = \text{pt in } U \text{ closest to } x$$

Theorem 8.1.12

Theorem

Let U be a fixed subspace of $\mathbb{R}^n$. If we define $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(\mathbf{x}) = \text{proj}_U \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

1. T is a linear operator.
2. $\operatorname{im} T = U$ and $\ker T = U^\perp$.
3. $\dim U + \dim U^\perp = n$.

Recap

Today we saw:

- Gram-Schmidt Orthogonalization Algorithm
- Orthogonal Complements
- Projections

Next time: Orthogonal Diagonalization

MATH254: Linear Algebra
Lecture 19

Moira MacNeil

February 26, 2025

Last Time

1. Gram-Schmidt Orthogonalization Algorithm
2. Orthogonal Complements
3. Projections

Today

1. Finish Orthogonal Complements and Projections

2. Orthogonal Diagonalization

Reminders:

- Assignment 4 is due Friday March 7

Recall: Orthogonal Complement

Orthogonal Complement of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$, define the **orthogonal complement** $U^\perp$ of U (pronounced "U-perp") by

$$U^\perp = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{y} = 0 \text{ for all } \mathbf{y} \text{ in } U\}$$

The set of vectors that are orthogonal to every vector in U.

Theorem 8.1.12

Theorem

Let U be a fixed subspace of $\mathbb{R}^n$. If we define $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(\mathbf{x}) = \text{proj}_U \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

1. T is a linear operator.
2. $\text{im}\, T = U$ and $\ker T = U^\perp$.
3. $\dim U + \dim U^\perp = n$.

Proof of 1.

If $U = \{0\}$ then $U^\perp = \mathbb{R}^n$

$T(x) = \text{proj}_{\{0\}} x = 0$ for all x

$\therefore$ $T = 0$ so 1, 2, 3 hold

Assume $U \neq \{0\}$

1. If $\{f_1, f_2, \ldots, f_m\}$ is an orthonormal basis of U, $\|f_i\|^2 = 1$

$$T(x) = \frac{(x \cdot f_1)}{\cancel{\|f_1\|^2}} f_1 + (x \cdot f_2) f_2 + \cdots + (x \cdot f_m) f_m$$

for all $x \in \mathbb{R}^n$

by definition of projection

let $x, y \in \mathbb{R}^n$ $r \in \mathbb{R}$

$(x+y) \cdot f_i = x \cdot f_i + y \cdot f_i$ for all $i = 1, \ldots, m$

so addition preserved

$(rx) \cdot f_i = r(x \cdot f_i)$ for all $i = 1, \ldots, m$

so scalar multiplication preserved.

Recall: Diagonalization

- An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors
- The matrix P with these eigenvectors as columns is a diagonalizing matrix for A, that is

$$D = P^{-1}AP \text{ is diagonal.}$$

- The nice bases of $\mathbb{R}^n$ are the orthogonal ones
- A natural question: which $n \times n$ matrices have an orthogonal basis of eigenvectors?
- The symmetric matrices $\rightarrow$ main result today

Recall: Orthonormal Vectors

- An orthogonal set of vectors is orthonormal if $\|\mathbf{v}\| = 1$ for each vector $\mathbf{v}$ in the set,
- An orthogonal set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ can be "normalized" to an orthonormal set $\{\frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1, \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2, \ldots, \frac{1}{\|\mathbf{v}_k\|}\mathbf{v}_k\}$
- If matrix A has n orthogonal eigenvectors, they can (by normalizing) be taken to be orthonormal
- The diagonalizing matrix P has orthonormal columns $\rightarrow$ these matrices are very easy to invert!

Theorem 8.2.1

Theorem

The following conditions are equivalent for an $n \times n$ matrix P.

1. P is invertible and $P^{-1} = P^T$.
2. The rows of P are orthonormal.
3. The columns of P are orthonormal.

Orthogonal Matrices

Orthogonal Matrices

An $n \times n$ matrix P is called an **orthogonal matrix** if it satisfies one (and hence all) of the conditions:

1. P is invertible and $P^{-1} = P^T$.
2. The rows of P are orthonormal.
3. The columns of P are orthonormal.

Given (2) and (3) orthonormal matrix might be a better name. But orthogonal matrix is standard.

Example 8.2.3

The rotation matrix $\begin{array}{cc} x_1 & x_2 \\ \end{array}$ $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal for any angle θ.

3. $x_1 \cdot x_2 = -\cos\theta\sin\theta + \cos\theta\sin\theta = 0$

$\therefore$ orthogonal

$\|x_1\| = \sqrt{\cos^2\theta + \sin^2\theta} = \|x_2\|$

$= \sqrt{1} = 1$

Orthogonal matrices have many nice properties

- They are easy to invert $\rightarrow$ simply take the transpose
- If $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a linear operator, T is distance preserving if and only if its matrix is orthogonal
- The matrices of rotations and reflections about the origin in $\mathbb{R}^2$ and $\mathbb{R}^3$ are all orthogonal
- However, it is not enough that the rows of a matrix A are merely orthogonal for A to be an orthogonal matrix

orthonormal rows

Example 8.2.4

$x_2 \cdot x_3 = 1$

The matrix $\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \left[\begin{array}{rrr} 2 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{array}\right] \end{array}$ has orthogonal rows but the columns are not orthogonal.

However, if the rows are normalized, the resulting matrix $\left[\begin{array}{ccc} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right]$ is orthogonal (so the columns are now orthonormal).

Example 8.2.5

If P and Q are orthogonal matrices, then PQ is also orthogonal, as is $P^{-1} = P^T$.

P and Q are invertible $(P^{-1} = P^T,\ Q^{-1} = Q^T)$
PQ is also invertible

$$(PQ)^{-1} = Q^{-1}P^{-1} = Q^T P^T = (PQ)^T$$

Therefore PQ is orthog.
(Pt 1. of Thm)

Orthogonally Diagonalizable Matrices

Orthogonally Diagonalizable Matrices

An $n \times n$ matrix A is said to be **orthogonally diagonalizable** when an orthogonal matrix P can be found such that $P^{-1}AP = P^{T}AP$ is diagonal.

Theorem 8.2.7

Principal Axes Theorem

The following conditions are equivalent for an $n \times n$ matrix A.

1. A has an orthonormal set of n eigenvectors.
2. A is orthogonally diagonalizable.
3. A is symmetric.

Proof $(2) \Rightarrow (3)$

Assume A ortnog. diag. then

$D = P^T A P$ is diag for some

P where $P^{-1} = P^T$

Then $A = PDP^T$

But $D = D^T$

So $A^T = P^{TT} D^T P^T = PDP^T = A$

$\therefore$ A is symmetric

Principal Axes for a Matrix

- A set of orthonormal eigenvectors of a symmetric matrix A is called a set of **principal axes** for A
- Principal Axes Theorem is also called the **Real Spectral Theorem** because the eigenvalues of a (real) symmetric matrix are real,
- The set of distinct eigenvalues is called the **spectrum** of the matrix
- The spectral theorem is a more general result for matrices with complex entries

Example 8.2.8

Find an orthogonal matrix P such that $P^{-1}AP$ is diagonal, where $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}$.

$$c_A(x) = \det(xI - A)$$

$$= \begin{vmatrix} x-1 & 0 & 1 \\ 0 & x-1 & -2 \\ 1 & -2 & x-5 \end{vmatrix} = \begin{vmatrix} x-1 & 0 & 1 \\ 2(x-1) & x-1 & 0 \\ 1 & -2 & x-5 \end{vmatrix}$$

$$= (x-1)[(x-1)(x-5)] + -4(x-1) - (x-1)$$

$$= (x-1)[(x-1)(x-5) - 5]$$

$$= (x-1)(x^2 - 6x) = x(x-1)(x-6)$$

Eigenvalues $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 6$

$\lambda_1 = 0$ sols $(\lambda_1 I - A)x = 0$

$$\begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & -2 \\ 1 & -2 & -5 \end{bmatrix} \to \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & -2 \\ 0 & -2 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$x_3 = t$ $\quad x_2 = -2t$ $\quad x_1 = t$

Eigenvector $\quad x_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

$\lambda_2 = 1$

$x_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

$\lambda_3 = 6$

$x_3 = \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$

Orthogonal eigen vectors! (verify)

Normalize

$\|x_1\| = \sqrt{1^2 + (-2)^2 + 1^2} = \sqrt{6}$

$\|x_2\| = \sqrt{5}$

$\|x_3\| = \sqrt{30}$

Then $P = \left[\frac{1}{\sqrt{6}} x_1 \quad \frac{1}{\sqrt{5}} x_2 \quad \frac{1}{\sqrt{30}} x_3 \right]$

$$= \frac{1}{\sqrt{30}} \begin{bmatrix} \sqrt{5} & 2\sqrt{6} & -1 \\ -2\sqrt{5} & \sqrt{6} & 2 \\ \sqrt{5} & 0 & 5 \end{bmatrix}$$

is orthogonal

So $P^{-1} = P^T$

$$D = P^T A P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

Recap

Today we saw:

- Orthogonally Diagonalizable Matrices
- Principal Axes Theorem

Next time: More on orthogonal diagonalization

MATH254: Linear Algebra
Lecture 20

Moira MacNeil

February 28, 2025

Last Time

1. Orthogonal Diagonalization

Today

1. Orthogonal Diagonalization continued

Reminders:

- Assignment 4 is due next Friday March 7

Recall: Gram-Schmidt Orthogonalization Algorithm

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of a subspace U of $\mathbb{R}^n$, construct $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m$ in U successively as follows:

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{x}_1 \\
\mathbf{f}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{f}_1}{||\mathbf{f}_1||^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_1}{||\mathbf{f}_1||^2}\mathbf{f}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_2}{||\mathbf{f}_2||^2}\mathbf{f}_2 \\
&\vdots \\
\mathbf{f}_k &= \mathbf{x}_k - \frac{\mathbf{x}_k \cdot \mathbf{f}_1}{||\mathbf{f}_1||^2}\mathbf{f}_1 - \frac{\mathbf{x}_k \cdot \mathbf{f}_2}{||\mathbf{f}_2||^2}\mathbf{f}_2 - \cdots - \frac{\mathbf{x}_k \cdot \mathbf{f}_{k-1}}{||\mathbf{f}_{k-1}||^2}\mathbf{f}_{k-1}
\end{aligned}$$

for each $k = 2, 3, \ldots, m$. Then

1. $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is an orthogonal basis of U.
2. $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ for each $k = 1, 2, \ldots, m$.

Recall: Orthogonal Matrices

Orthogonal Matrices

An $n \times n$ matrix P is called an **orthogonal matrix** if it satisfies one (and hence all) of the conditions:

1. P is invertible and $P^{-1} = P^T$.
2. The rows of P are orthonormal.
3. The columns of P are orthonormal.

Recall: Theorem 8.2.7

Principal Axes Theorem

The following conditions are equivalent for an $n \times n$ matrix A.

1. A has an orthonormal set of n eigenvectors.
2. A is orthogonally diagonalizable.
3. A is symmetric.

Theorem 8.2.9

Theorem

If A is an $n \times n$ symmetric matrix, then

$$(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$$

for all columns $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. converse is also true :)

Proof. Recall $x \cdot y = x^T y$ for all columns x, y

$A^T = A$

$$(Ax) \cdot y = (Ax)^T y = x^T A^T y = x^T A y = x \cdot (Ay)$$

Theorem 8.2.10

Theorem

If A is a symmetric matrix, then eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof. Let $Ax = \lambda x$ $\quad Ay = \mu y$ where $\mu \neq \lambda$

$$\lambda(x\cdot y) = (\lambda x)\cdot y = (Ax)\cdot y = x\cdot(Ay) = x\cdot(\mu y)$$
$$= \mu(x\cdot y)$$

$$\lambda(x\cdot y) = \mu(x\cdot y) \rightarrow \underbrace{(\lambda-\mu)}_{\neq 0}(x\cdot y) = 0$$
$$\Rightarrow x\cdot y = 0$$

Diagonalizing a Symmetric Matrix

Procedure to diagonalize a symmetric $n \times n$ matrix:

1. Find the distinct eigenvalues
2. Find orthonormal bases for each eigenspace (using Gram-Schmidt algorithm if necessary)

Then the set of all these basis vectors is orthonormal and contains n vectors.

Example 8.2.11

Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$.

$$c_A(x) = \det(xI - A) = \begin{vmatrix} x-8 & 2 & -2 \\ 2 & x-5 & -4 \\ -2 & -4 & x-5 \end{vmatrix}$$

$$= \begin{vmatrix} x-8 & 2 & -2 \\ 0 & x-9 & x-9 \\ -2 & -4 & x-5 \end{vmatrix} = (x-8)\left[(x-9)(x-5)+4(x-9)\right] - 2\left[2(x-9)+2(x-9)\right]$$

$$= (x-9)\left[(x-8)(x-5)-4\right] = (x-9)\left[x^2-9x\right]$$
$$= x(x-9)^2$$

$\lambda_1 = 0$ $\quad\quad$ $\lambda_2 = 9$ (mult. = 2)

solve $(\lambda_1 I - A)\,x_1 = 0$

$$\begin{bmatrix} -8 & 2 & -2 \\ 0 & -9 & -9 \\ -2 & -4 & -5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1/4 & 1/4 \\ 0 & -9 & -9 \\ 0 & -4.5 & -4.5 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & -1/4 & 1/4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_3 &= t \\ x_2 &= -t \\ x_1 &= -1/2\,t \end{aligned}$$

$$x_1 = \begin{bmatrix} -1/2 \\ -1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \qquad E_0(A) = \text{span}\{x_1\}$$

solve $(\lambda_2 I - A)\,x = 0$

$$\begin{bmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ -2 & -4 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_2 &= s \\ x_3 &= t \\ x_1 &= -2s+2t \end{aligned}$$

$x_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ $\quad x_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ $\quad x_2 \cdot x_3 = -4$

$\to$ NOT ORTHOG!

$E_9(A) = \text{span}\{x_2, x_3\}$

Use GS to find orthog. basis

$$f_1 = x_2$$
$$f_2 = x_3 - \frac{x_3 \cdot f_1}{\|f_1\|^2} f_1 = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

Normalize: $\{ \frac{1}{3} x_1, \frac{1}{\sqrt{5}} f_1, \frac{1}{3\sqrt{5}} f_2 \}$
(orthonormal eigenvectors)

$$P = \begin{bmatrix} \frac{1}{3} x_1 & \frac{1}{\sqrt{5}} f_1 & \frac{1}{3\sqrt{5}} f_2 \end{bmatrix}$$

$$= \frac{1}{3\sqrt{5}} \begin{bmatrix} \sqrt{5} & -6 & 2 \\ 2\sqrt{5} & 3 & 4 \\ -2\sqrt{5} & 0 & 5 \end{bmatrix}$$ is orthog. st $P^{-1}AP$ is diag.

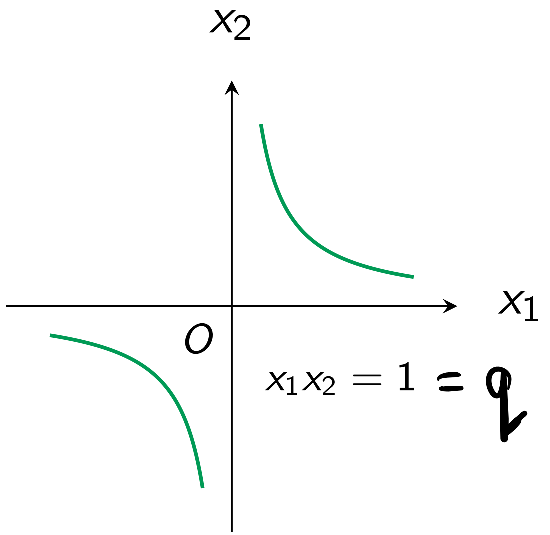
A nicer diag. matrix P

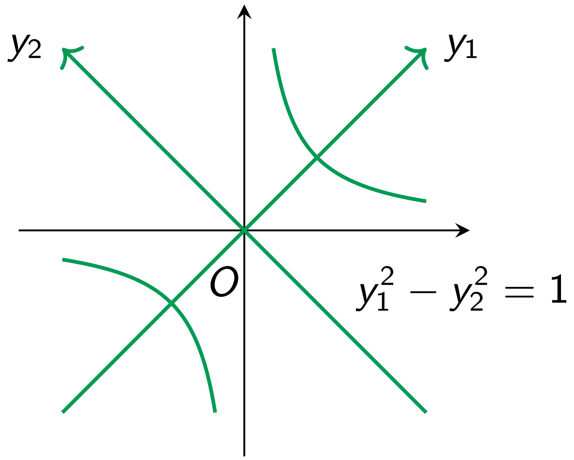
$y_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ $y_3 \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ are in $E_q(A)$ and orthog.

$Q = [\frac{1}{3} x_1 \quad \frac{1}{3} y_2 \quad \frac{1}{3} y_3]$ is also orthog. and $Q^{-1} A Q$ is diag.

Principle Axes and Quadratic Forms

- If A is symmetric and a set of orthogonal eigenvectors of A is given, the eigenvectors are called principal axes of A → from geometry
- An expression $q = ax_1^2 + bx_1x_2 + cx_2^2$ is called a **quadratic form** in the variables x_1 and x_2
- The graph of the equation $q = 1$ is called a **conic** in x_1 and x_2, e.g., $q = x_1x_2 = 1$
- Set $x_1 = y_1 + y_2$ and $x_2 = y_1 - y_2 \to q = y_1^2 - y_2^2$ (no cross term!)
- The y_1 and y_2 axes are the principal axes for the conic





Example 8.2.12

Find principal axes for the quadratic form $q = x_1^2 - 4x_1x_2 + x_2^2$.

First express q as a matrix

$$q = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

NOT SYMMETRIC

$$q = x_1^2 - 2x_1x_2 - 2x_2x_1 + x_2^2$$

$$q = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x^T A x$$

SYMMETRIC → A

Eigenvalues of A are $\lambda_1 = 3 \quad \lambda_2 = -1$

Eigenvectors (orthogonal) $x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\|x_1\| = \|x_2\| = \sqrt{2}$ so

$P = 1/\sqrt{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} = P^T A P$

$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ where $y = P^T x$

$\Leftrightarrow \quad x = Py$ (since $P^{-1} = P^T$)

solve for ys

Thus $y_1 = \frac{1}{\sqrt{2}}(x_1 - x_2)$

$y_2 = \frac{1}{\sqrt{2}}(x_1 + x_2)$

Rewrite q in terms of y

$$\begin{aligned} q = x^T A x &= (Py)^T A (Py) \\ &= y^T P^T A P y \\ &= y^T D y \\ &= 3y_1^2 - y_2^2 \end{aligned}$$

Theorem 8.2.13

Triangulation Theorem

If A is an $n \times n$ matrix with n real eigenvalues, an orthogonal matrix P exists such that $P^T AP$ is upper triangular.

Corollary 8.2.14

Corollary

If A is an $n \times n$ matrix with real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (possibly not all distinct), then $\det A = \lambda_1 \lambda_2 \ldots \lambda_n$ and $\mathrm{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Recap

Today we saw:

- Orthogonal diagonalization
- Principal axes and quadratic forms

Next time: Positive Definite Matrices

MATH254: Linear Algebra
Lecture 21

Moira MacNeil

March 4, 2025

Last Time

1. Orthogonal Diagonalization

Today

1. Positive Definite Matrices
2. Introduction to Complex Matrices

Reminders:

- Assignment 4 is due on Friday (March 7)
- Midterm 2 is next Friday, March 14 → MIDTERM INFO POSTED

→ PRACTICE QUESTIONS COMING SOON
↳ FOR SEC 8.3 and 8.7

Positive Definite Matrices

Positive Definite Matrices

A square matrix is called **positive definite** if it is symmetric and all its eigenvalues λ are positive, that is $\lambda > 0$.

Theorem 8.3.2

Theorem

If A is positive definite, then it is invertible and $\det A > 0$.

Proof.

If A is $n \times n$ and has eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ then $\det A = \lambda_1 \lambda_2 \cdots \lambda_n > 0$ by the principal axes thm and since A is PD.

Theorem 8.3.3

Theorem

A symmetric matrix A is positive definite if and only if $\underbrace{\mathbf{x}^T A\mathbf{x}}_{\text{quadratic forms}} > 0$ for every column $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$.

The positive definite matrices are the symmetric matrices A for which the quadratic form $q = \mathbf{x}^T A\mathbf{x}$ takes only positive values.

Example 8.3.4

(*) Every PD matrix can be factored as U^TU where U is upper Δ with pos on main diag.

If U is any invertible $n \times n$ matrix, show that $A = U^T U$ is positive definite. (The converse is also true!)

If $x \in \mathbb{R}^n$ where $x \neq 0$

$$x^TAx = x^T(U^TU)x = (x^TU^T)(Ux)$$
$$= (Ux)^TUx$$
$$= \|Ux\|^2 > 0$$

because $Ux \neq 0$ (U is invertible)

Principal Submatrices

- If A is any $n \times n$ matrix, let ${}^{(r)}A$ denote the $r \times r$ submatrix in the upper left corner of A
- ${}^{(r)}A$ is the matrix obtained from A by deleting the last $n - r$ rows and columns
- The matrices ${}^{(1)}A, {}^{(2)}A, {}^{(3)}A, \ldots, {}^{(n)}A = A$ are called the **principal submatrices** of A

Example 8.3.5

If $A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$ then:

$^{(1)}A = [10]$ delete 3-1 row/cols

$^{(2)}A = \begin{bmatrix} 10 & 5 \\ 5 & 3 \end{bmatrix}$ delete 3-2 rows/cols

$^{(3)}A = A$ delete 0 rows cols

Lemma 8.3.6

Lemma

If A is positive definite, so is each principal submatrix ${}^{(r)}A$ for $r = 1, 2, \ldots, n$.

Proof. $A = \begin{bmatrix} {}^{(r)}A & P \\ Q & R \end{bmatrix}$ If $y \neq 0 \in \mathbb{R}^r$ let $x = \begin{bmatrix} y \\ 0 \end{bmatrix} \in \mathbb{R}^n$ $(x \neq 0)$

A is PD $\therefore$

$$\underline{0 <} x^T A x = \begin{bmatrix} y^T & 0 \end{bmatrix} \begin{bmatrix} {}^{(r)}A & P \\ Q & R \end{bmatrix} \begin{bmatrix} y \\ 0 \end{bmatrix} = \underline{y^T \, {}^{(r)}A \, y}$$

$\therefore$ ${}^{(r)}A$ is PD for any $r = 1, \ldots, n$ $\square$

Theorem 8.3.7

Theorem

The following conditions are equivalent for a symmetric $n \times n$ matrix A:

1. A is positive definite.
2. $\det({}^{(r)}A) > 0$ for each $r = 1, 2, \ldots, n$.
3. $A = U^T U$ where U is an upper triangular matrix with positive entries on the main diagonal.

The factorization in (3) is unique (called the **Cholesky factorization** of A).

Cholesky Factorization Algorithm

Algorithm for the Cholesky Factorization

If A is a positive definite matrix, the Cholesky factorization $A = U^T U$ can be obtained as follows:

Step 1. Carry A to an upper triangular matrix U_1 with positive diagonal entries using row operations each of which adds a multiple of a row to a lower row.

Step 2. Obtain U from U_1 by dividing each row of U_1 by the square root of the diagonal entry in that row.

Example

Find the Cholesky factorization of $A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$.

Step 0: is A PD?

$\det({}^{(1)}A) = 10 > 0$

$\det({}^{(2)}A) = 5 > 0$

$\det({}^{(3)}A) = \det(A) = 3 > 0$

Step 1:

$$A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 10 & 5 & 2 \\ 0 & 1/2 & 1 \\ 0 & 1 & 13/5 \end{bmatrix} \to \begin{bmatrix} 10 & 5 & 2 \\ 0 & 1/2 & 1 \\ 0 & 0 & 3/5 \end{bmatrix} = U_1$$

Step 2:

$$U = \begin{bmatrix} \sqrt{10} & 5/\sqrt{10} & 2/\sqrt{10} \\ 0 & 1/\sqrt{2} & \sqrt{2} \\ 0 & 0 & \sqrt{3}/\sqrt{5} \end{bmatrix}$$

$U^T U = A$

Complex Eigenvalues

- Until now, examples have been chosen so that the roots of characteristic polynomials are real
- This need not be true, even when the characteristic polynomial has real coefficients
- For example, if $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ then $c_A(x) = x^2 + 1$ has roots i and $-i$

$i^2 = -1$

Complex Vector Space

- $\mathbb{C}^n \to$ the set of all n-tuples of complex numbers

- We define vector operations on $\mathbb{C}^n$ as follows:

$$(v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n) = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$$
$$u(v_1, v_2, \ldots, v_n) = (uv_1, uv_2, \ldots, uv_n) \quad \text{for } u \text{ in } \mathbb{C}$$

- $\mathbb{C}^n$ satisfies the axioms for a vector space (with complex scalars)

- Thus we can have spanning sets, linearly independent subsets and bases for $\mathbb{C}^n$

- Definitions are identical to the real case, except that the scalars are allowed to be complex numbers.

- The standard basis of $\mathbb{R}^n$ remains a basis of $\mathbb{C}^n$, called the **standard basis** of $\mathbb{C}^n$

e_i

Complex Matrices

- $A = [a_{ij}]$ is called a **complex matrix** if every entry a_{ij} is a complex number
- If $z = a + bi$ is a complex number, the **conjugate** of z is the complex number:

$$\overline{z} = a - bi \quad \text{where } z = a + bi$$

- Define the **conjugate** of $A = [a_{ij}]$ to be the matrix

$$\overline{A} = \left[\; \overline{a}_{ij} \; \right]$$

obtained from A by conjugating every entry

- For all (complex) matrices of appropriate size:

$$\overline{A + B} = \overline{A} + \overline{B} \quad \text{and} \quad \overline{AB} = \overline{A}\,\overline{B}$$

Standard Inner Product

Standard Inner Product in ~~$\mathbb{R}^n$~~ $\mathbb{C}^n$

Given $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ in $\mathbb{C}^n$, define their **standard inner product** $\langle \mathbf{z}, \mathbf{w} \rangle$ by

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n = \mathbf{z} \cdot \overline{\mathbf{w}}$$

where $\overline{w}$ is the conjugate of the complex number w.

Gives scalar in $\mathbb{C}^n$ and if $z, w \in \mathbb{R}^n$

then $\langle z, w \rangle = z \cdot w$

Example 8.7.2

$\overline{z} = (2, 1+i, -2i, 3+i)$ $\overline{w} = (1+i, -1, i, 3-2i)$

If $\mathbf{z} = (2, 1-i, 2i, 3-i)$ and $\mathbf{w} = (1-i, -1, -i, 3+2i)$, then

$$\langle \mathbf{z}, \mathbf{w} \rangle = 2(1+i) + (1-i)(-1) + (2i)(i) + (3-i)(3-2i) = 6-6i$$

$$\langle \mathbf{z}, \mathbf{z} \rangle = 2 \cdot 2 + (1-i)(1+i) + (2i)(-2i) + (3-i)(3+i) = 20$$

Theorem 8.7.3

Theorem

Let $\mathbf{z}$, $\mathbf{z}_1$, $\mathbf{w}$, and $\mathbf{w}_1$ denote vectors in $\mathbb{C}^n$, and let λ denote a complex number.

1. $\langle \mathbf{z}+\mathbf{z}_1, \mathbf{w}\rangle = \langle \mathbf{z}, \mathbf{w}\rangle + \langle \mathbf{z}_1, \mathbf{w}\rangle$ and $\langle \mathbf{z}, \mathbf{w}+\mathbf{w}_1\rangle = \langle \mathbf{z}, \mathbf{w}\rangle + \langle \mathbf{z}, \mathbf{w}_1\rangle$.
2. $\langle \lambda\mathbf{z}, \mathbf{w}\rangle = \lambda\langle \mathbf{z}, \mathbf{w}\rangle$ and $\langle \mathbf{z}, \lambda\mathbf{w}\rangle = \underline{\overline{\lambda}\langle \mathbf{z}, \mathbf{w}\rangle}$.
3. $\langle \mathbf{z}, \mathbf{w}\rangle = \overline{\langle \mathbf{w}, \mathbf{z}\rangle}$.
4. $\langle \mathbf{z}, \mathbf{z}\rangle \geq 0$, and $\langle \mathbf{z}, \mathbf{z}\rangle = 0$ if and only if $\mathbf{z} = \mathbf{0}$.

PROOF.

(3) $z = (z_1, \cdots, z_n)$ $w = (w_1, \cdots, w_n) \in \mathbb{C}^n$

$$\overline{\langle w, z\rangle} = \overline{(w_1\bar{z}_1 + \dots + w_n\bar{z}_n)} = \overline{w_1\bar{z}_1} + \dots + \overline{w_n\bar{z}_n}$$
$$= z_1\overline{w_1} + \dots + z_n\overline{w_n}$$
$$= \langle z, w\rangle$$

(4) let $w = z = (w_1, \dots, w_n)$

$$\langle w, w\rangle = w_1\overline{w_1} + \dots + w_n\overline{w_n}$$

$$\left[\begin{array}{l} |w_j| = \sqrt{a_j^2 + b_j^2} \geq 0 \text{ where } \quad w_j = a_j + b_j i \\ |w_j|^2 = w_j\overline{w_j} \geq 0 \end{array}\right]$$

$$\langle w, w\rangle = |w_1|^2 + \dots + |w_n|^2$$

nonnegative real

equal to zero iff $w = 0$

This will give us vector norm in $\mathbb{C}^n$ $\square$

Recap

Today we saw:

- Positive Definite Matrices
- Cholesky Factorization
- Standard Inner Product

Next time: Complex Matrices

MATH254: Linear Algebra
Lecture 22

Moira MacNeil

March 5, 2025

Last Time

1. Positive Definite Matrices
2. Cholesky Factorization
3. Standard Inner Product

Today

1. More on Complex Matrices

Reminders:

- Assignment 4 is due on Friday (March 7)
- Midterm 2 is next Friday, March 14

Recall: Standard Inner Product

$z = a + bi \in \mathbb{C}$

$\overline{z} = a - bi$

Standard Inner Product in $\mathbb{C}^n$

Given $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ in $\mathbb{C}^n$, define their **standard inner product** $\langle \mathbf{z}, \mathbf{w} \rangle$ by

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n = \underline{\mathbf{z}} \cdot \underline{\overline{\mathbf{w}}}$$

where $\overline{w}$ is the conjugate of the complex number w.

Norm and Length of Complex Vectors

Norm and Length in $\mathbb{C}^n$

As for the dot product on $\mathbb{R}^n$, property (4) enables us to define the **norm** or **length** $\|\mathbf{z}\|$ of a vector $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ in $\mathbb{C}^n$:

$$\|\mathbf{z}\| = \sqrt{\langle \mathbf{z}, \mathbf{z} \rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

For a complex number $z = a + bi$, we define $|z| = \sqrt{a^2 + b^2}$

Theorem 8.7.5

Theorem

If $\mathbf{z}$ is any vector in $\mathbb{C}^n$, then

1. $\|\mathbf{z}\| \geq 0$ and $\|\mathbf{z}\| = 0$ if and only if $\mathbf{z} = \mathbf{0}$.
2. $\|\lambda\mathbf{z}\| = |\lambda|\|\mathbf{z}\|$ for all complex numbers λ.

A vector $\mathbf{u}$ in $\mathbb{C}^n$ is called a **unit vector** if $\|\mathbf{u}\| = 1$.

Normalize $\frac{1}{\|v\|} v$

Example 8.7.6

$|z| = \sqrt{a^2 + b^2}$ $\quad z = a + bi$

In $\mathbb{C}^4$, find a unit vector **u** that is a positive real multiple of $\mathbf{z} = (1 - i, i, 2, 3 + 4i)$.

$$\|z\| = \sqrt{\langle z, z\rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

$$= \sqrt{(1^2 + 1^2) + (1^2) + 2^2 + (3^2 + 4^2)}$$

$$= \sqrt{32} = 4\sqrt{2}$$

$$u = \frac{1}{4\sqrt{2}} z$$

Conjugate Transpose

Conjugate Transpose in $\mathbb{C}^n$

The **conjugate transpose** A^H of a complex matrix A is defined by

$$A^H = (\overline{A})^T = \overline{(A^T)}$$

We have $A^H = A^T$ when A is real. (Other notations for A^H are A^* and $A^\dagger$.)

Example 8.7.8

$$\begin{bmatrix} 3 & 1-i & 2+i \\ 2i & 5+2i & -i \end{bmatrix}^H = \begin{bmatrix} 3 & -2i \\ 1+i & 5-2i \\ 2-i & i \end{bmatrix}$$

Theorem 8.7.9

Theorem

Let A and B denote complex matrices, and let λ be a complex number.

1. $(A^H)^H = A$.
2. $(A + B)^H = A^H + B^H$.
3. $(\lambda A)^H = \overline{\lambda} A^H$.
4. $(AB)^H = B^H A^H$.

Hermitian Matrices

Hermitian Matrices

A square complex matrix A is called **hermitian** if $A^H = A$, equivalently if $\overline{A} = A^T$.

Hermitian matrices have real entries on the main diagonal and the "reflection" of each nondiagonal entry must be the conjugate of that entry.

Example 8.7.11

Show $\begin{bmatrix} 3 & i & 2+i \\ -i & -2 & -7 \\ 2-i & -7 & 1 \end{bmatrix}$ is hermitian, whereas $\begin{bmatrix} 1 & i \\ i & -2 \end{bmatrix}$ and $\begin{bmatrix} 1 & i \\ -i & i \end{bmatrix}$ are not.

$$= \begin{bmatrix} 3 & i & 2+i \\ -i & -2 & -7 \\ 2-i & -7 & 1 \end{bmatrix}$$

$$\neq \begin{bmatrix} 1 & -i \\ -i & -2 \end{bmatrix}$$

$$\neq \begin{bmatrix} 1 & i \\ -i & -i \end{bmatrix}$$

Theorem 8.7.12

A sym. $\in \mathbb{R}$ $(Ax)\cdot y = x\cdot(Ay)$ for any $x, y \in \mathbb{R}^n$

Theorem

An $n \times n$ complex matrix A is hermitian if and only if

$$\langle A\mathbf{z}, \mathbf{w}\rangle = \langle \mathbf{z}, A\mathbf{w}\rangle$$

for all n-tuples $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^n$.

Proof. $(\Rightarrow)$ A is hermitian $A^T = \overline{A}$, if $z, w \in \mathbb{C}^n$

then $\langle z, w\rangle = z^T \overline{w}$

$\langle Az, w\rangle = (Az)^T \overline{w} = z^T A^T \overline{w} = z^T \overline{A} \overline{w}$

$= z^{\dagger}\overline{(Aw)} = \langle z, Aw\rangle$

$(\Leftarrow)$ e_j is col j of identity matrix

$A = [a_{ij}]$ then condition gives

$\bar{a}_{ij} = \langle e_i, Ae_j\rangle \boxed{=} \langle Ae_i, e_j\rangle = a_{ij}$

Thus $\bar{A} = A^T$ $\therefore$ A hermitian.

Theorem 8.7.13

Theorem

Let A denote a hermitian matrix.

1. The eigenvalues of A are real.
2. Eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof. let λ, μ be eigenvalues of A with eigenvectors z, w (non zero) then $Az = \lambda z$ and $Aw = \mu w$

$$\lambda \langle z, w \rangle = \langle \lambda z, w \rangle = \langle Az, w \rangle = \langle z, Aw \rangle$$

$$= \langle z, \mu w \rangle = \bar{\mu} \langle z, w \rangle \quad (*)$$

(1) If $\lambda = \mu$ and $z = w$ $(*)$ becomes

$$\lambda \langle z, z \rangle = \bar{\lambda} \langle z, z \rangle$$

$$\langle z, z \rangle = \|z\|^2 \neq 0 \Rightarrow \underbrace{\lambda = \bar{\lambda}}_{\text{only true if } \lambda \text{ is real}}$$

Similarly μ is real

□

(2) Since μ, λ are real $(*)$ becomes

$$\lambda \langle z, w \rangle = \mu \langle z, w \rangle$$

if $\lambda \neq \mu \Rightarrow \langle z, w \rangle = 0$

$\therefore$ z, w are orthogonal. □

Orthogonality in $\mathbb{C}^n$

Orthogonal and Orthonormal Vectors in $\mathbb{C}^n$

As in the real case, a set of nonzero vectors $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_m\}$ in $\mathbb{C}^n$ is called **orthogonal** if $\langle \mathbf{z}_i, \mathbf{z}_j \rangle = 0$ whenever $i \neq j$, and it is **orthonormal** if, in addition, $\|\mathbf{z}_i\| = 1$ for each i.

Theorem 8.7.15

Theorem

The following are equivalent for an $n \times n$ complex matrix A.

1. A is invertible and $A^{-1} = A^H$.
2. The rows of A are an orthonormal set in $\mathbb{C}^n$.
3. The columns of A are an orthonormal set in $\mathbb{C}^n$.

Unitary Matrices

Unitary Matrices

A square complex matrix U is called **unitary** if $U^{-1} = U^H$.

A real matrix is unitary if and only if it is orthogonal.

Example 8.7.17

The matrix $A = \begin{bmatrix} 1+i & 1 \\ 1-i & i \end{bmatrix}$ has orthogonal columns, but the rows are not orthogonal.

Normalizing the columns gives the unitary matrix $\frac{1}{2}\begin{bmatrix} 1+i & \sqrt{2} \\ 1-i & \sqrt{2}i \end{bmatrix}$.

Example 8.7.18

Consider the hermitian matrix $A = \begin{bmatrix} 3 & 2+i \\ 2-i & 7 \end{bmatrix}$. Find the eigenvalues of A, find two orthonormal eigenvectors, and so find a unitary matrix U such that $\underline{U^H AU}$ is diagonal.

$$C_A(x) = \det(xI - A) = \begin{vmatrix} x-3 & -2-i \\ -2+i & x-7 \end{vmatrix}$$

$$= (x-3)(x-7) - (-2+i)(-2-i)$$

$$= x^2 - 10x + 21 - [4 - i^2]$$

$$= x^2 - 10x + 16 = (x-2)(x-8)$$

So $\lambda_1 = 2 \qquad \lambda_2 = 8$

Eigenvectors: solve $(\lambda_i I - A)x = 0$

$$\begin{bmatrix} -1 & -2-i \\ -2+i & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 2+i \\ 0 & 0 \end{bmatrix} \qquad x_1 = \begin{bmatrix} 2+i \\ -1 \end{bmatrix}$$

Add $-(-2+i)$ Row 1 to Row 2

$$-5-(-2-i)(-2+i) = -5-(4-i^2)$$
$$= -5+5 = 0$$

$$\begin{bmatrix} 5 & -2-i \\ -2+i & 1 \end{bmatrix} \to \begin{bmatrix} 1 & \frac{-2-i}{5} \\ 0 & 0 \end{bmatrix} \qquad \begin{matrix} x_2 = t \\ x_1 = -\left(\frac{-2-i}{5}\right) \end{matrix}$$

$$x_2 = \begin{bmatrix} \frac{-2-i}{5} \\ 1 \end{bmatrix} = \begin{bmatrix} 2+i \\ 5 \end{bmatrix}$$

x_1, x_2 orthog. $\to$ normalize

$$\|x_1\| = \sqrt{2^2+1^2+1^2} = \sqrt{6}$$

$$\|x_2\| = \sqrt{(-2)^2+(-1)^2+5^2} = \sqrt{30}$$

$$U = \begin{bmatrix} \frac{1}{\sqrt{6}}x_1 & \frac{1}{\sqrt{30}}x_2 \end{bmatrix}$$

$$U^H A U = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}$$

Recap

Today we saw:

- Norm and Length in $\mathbb{C}^n$
- Conjugate Transpose
- Hermitian and Unitary Matrices

Next time: Unitary Diagonalization, Quadratic Forms

MATH254: Linear Algebra
Lecture 23

Moira MacNeil

March 7, 2025

Last Time

1. Norm and Length in $\mathbb{C}^n$
2. Conjugate Transpose A^H
3. Hermitian and Unitary Matrices

Today

1. Unitary Diagonalization
2. An Application to Quadratic Forms → NOT ON MIDTERM

Reminders:

- Midterm 2 is next Friday, March 14

Recall: Theorem 8.7.15

$$A^H = (\overline{A})^T$$

Theorem

The following are equivalent for an $n \times n$ complex matrix A.

1. A is invertible and $A^{-1} = A^H$.
2. The rows of A are an orthonormal set in $\mathbb{C}^n$.
3. The columns of A are an orthonormal set in $\mathbb{C}^n$.

Recall: Unitary Matrices

Unitary Matrices

A square complex matrix U is called **unitary** if $U^{-1} = U^H$.

A real matrix is unitary if and only if it is orthogonal.

Theorem 8.7.19

Schur's Theorem

If A is any $n \times n$ complex matrix, there exists a unitary matrix U such that

$$U^H A U = T$$

is upper triangular. Moreover, the entries on the main diagonal of T are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A (including multiplicities).

Corollary 8.7.20

Corollary

Let A be an $n \times n$ complex matrix, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues of A, including multiplicities. Then

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n \quad \text{and} \quad \operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

Schur's theorem

- Schur's theorem asserts that every complex matrix can be "unitarily triangularized"
- This does not mean that that every complex matrix can be "unitarily diagonalized"
- For example, if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, there is no invertible complex matrix U at all such that $U^{-1}AU$ is diagonal.

Theorem 8.7.21

Hermitian: $A^H = A$

Spectral Theorem

If A is hermitian, there is a unitary matrix U such that $U^H AU$ is diagonal.

A is called **unitarily diagonalizable** in this case.

Proof. By Schur's theorem, let $U^H A U = T$ where T is upper triangular and U is unitary.

$$\underset{LT}{T^H} = (U^H A U)^H = U^H A^H U^{HH} = U^H A U = \underset{UT}{T}$$

T is both upper and lower tri $\rightarrow$ diag.!

Example 8.7.22

Show that the non-hermitian matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is unitarily diagonalizable.

A is non-herm. $A^H = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

$c_A(x) = \begin{vmatrix} x & -1 \\ 1 & x \end{vmatrix} = x^2+1 = (x-i)(x+i)$

$\lambda_1 = i \quad \lambda_2 = -i$

Eigenvectors

$\lambda_1 = i, \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \to \begin{bmatrix} i & -1 \\ 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$ $x_2 = t$, $x_1 = -it$

$$x_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

$$\lambda_2 = -i, \quad \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \to \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_2 &= t \\ x_1 &= it \end{aligned}$$

$$x_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$$\begin{aligned} \langle x_1, x_2 \rangle &= x_1 \cdot \overline{x_2} \\ &= (-i)(-i) + 1(1) \\ &= i^2 + 1 = 0 \end{aligned}$$

this was the issue! we need to use inner prod since these vectors are in $\mathbb{C}^n$

A good reminder for all of us (especially me) :)

$$\|x_1\| = \|x_2\| = \sqrt{2}$$

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \text{ is unitary}$$

$$U^H A U = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \text{ is diagonal}$$

Normal Matrix

Normal Matrix

An $n \times n$ complex matrix N is called **normal** if $NN^H = N^H N$.

Every hermitian or unitary matrix is normal, as is, for example, the matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ from the previous example.

Theorem 8.7.23

$A^H A = A A^H$

Theorem

An $n \times n$ complex matrix A is unitarily diagonalizable if and only if A is normal.

Proof. ($\Longrightarrow$) Assume $U^H A U = D$ where U unitary D diagonal. D is normal $DD^H = D^H D$.

$$DD^H = (U^H A U)(U^H A U)^H = U^H A (U U^H) A^H U$$
$$= U^H (A A^H) U$$

$\Updownarrow =$

$$D^H D = U^H (A^H A) U$$

$$U^H (A A^H) U = U^H (A^H A) U$$

$$A A^H = A^H A \Rightarrow A \text{ is normal.}$$

$\square$

Theorem 8.7.24

Cayley-Hamilton Theorem

If A is an $n \times n$ complex matrix, then $c_A(A) = 0$; that is, A is a root of its characteristic polynomial.

Example: Cayley-Hamilton Theorem

Let $A = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$, find the characteristic polynomial $c_A(x)$, and compute $c_A(A)$.

$$c_A(x) = \det\begin{bmatrix} x-1 & -3 \\ 1 & x-2 \end{bmatrix} = (x-1)(x-2)+3$$
$$= x^2 - 3x + 5$$

$$c_A(A) = A^2 - 3A + 5I_2$$
$$= \begin{bmatrix} -2 & 9 \\ -3 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 9 \\ -3 & 6 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

MIDTERM TOPICS STOP HERE

Quadratic Forms

Quadratic Form

A **quadratic form** q in the n variables $x_1, x_2, \ldots, x_n$ is a linear combination of terms $x_1^2, x_2^2, \ldots, x_n^2$, and cross terms $x_1x_2, x_1x_3, x_2x_3, \ldots$.

$x_1 x_2$

$x_2 x_1 \quad \ldots$

Quadratic Forms as Matrix Products

- If $n = 3$, q has the form x_1, x_2, x_3

$$q = \underbrace{a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2} + \underbrace{a_{12}x_1x_2 + a_{21}x_2x_1 + a_{13}x_1x_3 + a_{31}x_3x_1 + a_{23}x_2x_3 + a_{32}x_3x_2}_{\text{cross terms}}$$

- In general, this can be written compactly as

$$q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} \qquad \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- If $i \neq j$, the separate terms $a_{ij}x_ix_j$ and $a_{ji}x_jx_i$ can be replaced by

$$\frac{1}{2}(a_{ij} + a_{ji})x_ix_j \quad \text{and} \quad \frac{1}{2}(a_{ij} + a_{ji})x_jx_i$$

without altering the quadratic form.

- May assume that x_ix_j and x_jx_i have the same coefficient in the sum for q $\rightarrow$ **assume that A is symmetric**.

Example 8.9.2

Write $q = x_1^2 + 3x_3^2 + 2x_1x_2 - x_1x_3$ in the form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, where A is a symmetric 3×3 matrix.

$a_{12} = 2 \quad a_{21} = 0$

CROSS TERMS $\quad 2x_1x_2 = x_1x_2 + x_2x_1$

$a_{13} = -1 \quad a_{31} = 0$

$$-x_1x_3 + 0x_3x_1 = -\tfrac{1}{2}x_1x_3 - \tfrac{1}{2}x_3x_1$$

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix} \rightarrow A = \begin{bmatrix} 1 & 1 & -1/2 \\ 1 & 0 & 0 \\ -1/2 & 0 & 3 \end{bmatrix}$$

$$q(x) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1/2 \\ 1 & 0 & 0 \\ -1/2 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x^T A x$$

Change of Variables

Can always assume

- Given a quadratic form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, where A is symmetric, the problem is to find new variables $y_1, y_2, \ldots, y_n$, related to $x_1, x_2, \ldots, x_n$, such that q has no cross terms $x_i x_j$
- We want to find $q = \mathbf{y}^T D\mathbf{y}$ where D is diagonal
- D is the matrix obtained when the symmetric matrix A is orthogonally diagonalized!
- By Principal Axes Theorem, we can find an orthogonal matrix P where $P^T AP = D$
- Define variables $\mathbf{y}$ by $\mathbf{x} = P\mathbf{y} \iff \mathbf{y} = P^T\mathbf{x}$ (these are the principal axes)
- Substituting into $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ gives

$$q = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T(P^T AP)\mathbf{y} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

which has no cross terms

Theorem 8.9.3

Diagonalization Theorem

Let $q = \mathbf{x}^T A\mathbf{x}$ be a quadratic form in the variables $x_1, x_2, \ldots, x_n$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and A is a symmetric $n \times n$ matrix. Let P be an orthogonal matrix such that $P^T AP$ is diagonal, and define new variables $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ by

$$\mathbf{x} = P\mathbf{y} \quad \text{equivalently} \quad \mathbf{y} = P^T\mathbf{x}$$

If q is expressed in terms of these new variables $y_1, y_2, \ldots, y_n$, the result is

$$q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A repeated according to their multiplicities.

Recap

Today we saw:

- Unitary Diagonalization
- An Introduction to Quadratic Forms

Next time: More on Quadratic Forms

MATH254: Linear Algebra
Lecture 24

Moira MacNeil

March 11, 2025

Last Time

1. Unitary Diagonalization
2. Introduction to Quadratic Forms

Today

1. Continue with Application to Quadratic Forms

Reminders:

- Midterm 2 is on Friday, March 14

FORMULA SHEET – DIY ☺

8.5 x 11 inches 1 sided

Recall: Quadratic Forms

Quadratic Form

A **quadratic form** q in the n variables $x_1, x_2, \ldots, x_n$ is a linear combination of terms $x_1^2, x_2^2, \ldots, x_n^2$, and cross terms $x_1x_2, x_1x_3, x_2x_3, \ldots$.

- Can be written compactly as
$$q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$
- If $i \neq j$, the separate terms $a_{ij}x_ix_j$ and $a_{ji}x_jx_i$ can be replaced by
$$\frac{1}{2}(a_{ij} + a_{ji})x_ix_j \quad \text{and} \quad \frac{1}{2}(a_{ij} + a_{ji})x_jx_i$$
without altering the quadratic form $\implies$ A is symmetric

Theorem 8.9.3

Diagonalization Theorem

Let $q = \mathbf{x}^T A\mathbf{x}$ be a quadratic form in the variables $x_1, x_2, \ldots, x_n$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and A is a symmetric $n \times n$ matrix. Let P be an orthogonal matrix such that $P^T AP$ is diagonal, and define new variables $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ by

$$\mathbf{x} = P\mathbf{y} \quad \text{equivalently} \quad \mathbf{y} = P^T\mathbf{x}$$

If q is expressed in terms of these new variables $y_1, y_2, \ldots, y_n$, the result is

$$q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A repeated according to their multiplicities.

Principal Axes

- Let $q = \mathbf{x}^T A\mathbf{x}$ be a quadratic form where A is a symmetric matrix and let $\lambda_1, \ldots, \lambda_n$ be the (real) eigenvalues of A repeated according to their multiplicities

- A corresponding set $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ of orthonormal eigenvectors for A is called a set of **principal axes** for the quadratic form q

- The orthogonal matrix P in Diagonalization Theorem is $P = \begin{bmatrix} \mathbf{f}_1 & \cdots & \mathbf{f}_n \end{bmatrix}$, so the variables X and Y are related by

$$\mathbf{x} = P\mathbf{y} = y_1\mathbf{f}_1 + y_2\mathbf{f}_2 + \cdots + y_n\mathbf{f}_n$$

Ch 5

- The coefficients are $y_i = \mathbf{x} \cdot \mathbf{f}_i$ by the expansion theorem. Hence q itself is easily computed:

$$q = q(\mathbf{x}) = \lambda_1(\mathbf{x} \cdot \mathbf{f}_1)^2 + \cdots + \lambda_n(\mathbf{x} \cdot \mathbf{f}_n)^2$$

Example 8.9.4

Find new variables y_1, y_2, y_3, and y_4 such that

$$q = 3(x_1^2 + x_2^2 + x_3^2 + x_4^2) + 2x_1x_2 - 10x_1x_3 + 10x_1x_4 + 10x_2x_3 - 10x_2x_4 + 2x_3x_4$$

has diagonal form, and find the corresponding principal axes.

$x = [x_1 \ x_2 \ x_3 \ x_4]^T$

$$A = \begin{bmatrix} 3 & 1 & -5 & 5 \\ 1 & 3 & 5 & -5 \\ -5 & 5 & 3 & 1 \\ 5 & -5 & 1 & 3 \end{bmatrix}$$

$c_A(x) = \det(xI - A)$

$= (x-12)(x+8)(x-4)^2$

Eigenvalues

$\lambda_1 = 12$, $\lambda_2 = -8$

$\lambda_3 = \lambda_4 = 4$

Orthonormal eigenvectors
(might involve using GS algo to orthog. x_3, x_4)

$$f_1 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \quad f_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \quad f_3 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$f_4 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \quad P = \begin{bmatrix} f_1 & f_2 & f_3 & f_4 \end{bmatrix}$$

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

P is orthogonal.
$P^{-1}AP = P^TAP$ is diagonal

New vars y are related to x vars

$$x = Py$$
$$y = P^Tx$$

$x \cdot f_1$

$$y_1 = \frac{1}{2}(x_1 - x_2 - x_3 + x_4)$$
$$y_2 = \frac{1}{2}(x_1 - x_2 + x_3 - x_4)$$
$$y_3 = \frac{1}{2}(x_1 + x_2 + x_3 + x_4)$$
$$y_4 = \frac{1}{2}(x_1 + x_2 - x_3 - x_4)$$

Substitute y's back into the $q(x)$

$$q = 12y_1^2 - 8y_2^2 + 4y_3^2 + 4y_4^2$$

$$= \lambda_1(x \cdot f_1)^2 + \lambda_2(x \cdot f_2)^2 + \lambda_3(x \cdot f_3)^2 + \lambda_4(x \cdot f_4)^2$$

Theorem 8.9.5

Theorem

If $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is a quadratic form given by a symmetric matrix A, then A is uniquely determined by q.

Proof. Let $q(x) = x^T B x$ and $q(x) = x^T A x$ for all x where $B^T = B$. If $C = A - B$, then $C^T = C$ $\quad x^T C x = 0$ for all x,

since $x^T(A-B)x = (x^T A - x^T B)x = x^T A x - x^T B x$

$= q - q = 0$

Want to show $A = B \Leftrightarrow C = 0$

Given $y \in \mathbb{R}^n$

$$0 = (x+y)^T C (x+y) = \underbrace{x^T C x}_{=0} + x^T C y + y^T C x + \underbrace{y^T C y}_{=0}$$

$$= x^T C y + y^T C x$$

$$y^T C x = (x^T C y)^T = x^T C y \quad (1 \times 1)$$

$$0 = 2(x^T C y) \rightarrow x^T C y = 0$$

for all $x, y \in \mathbb{R}^n$

If e_j is col j of I_n

then $e_i^T C e_j = 0$

this is the (i,j)th entry of C

this holds for any (i,j) pair

so all of $C = 0$.

$\square$

Different Ways to Express the Same Quadratic Form

- A quadratic form q in variables x_i can be written in several ways . For example, if $q = 2x_1^2 - 4x_1x_2 + x_2^2$ then

$$q = 2(x_1 - x_2)^2 - x_2^2 \quad \text{and} \quad q = -2x_1^2 + (2x_1 - x_2)^2$$

- Let $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, $\rightarrow$ new variables $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ are linear combinations of x_i, then $\mathbf{y} = A\mathbf{x}$ for $n \times n$ matrix A
- Since we want to solve for the x_i in terms of y_i $\rightarrow$ matrix A invertible
- Hence suppose U is an invertible matrix and that the new variables $\mathbf{y}$ are given by

$$\mathbf{y} = U^{-1}\mathbf{x}, \quad \text{equivalently } \mathbf{x} = U\mathbf{y}$$

- Then $q = q(\mathbf{x}) = (U\mathbf{y})^T A(U\mathbf{y}) = \mathbf{y}^T(U^T AU)\mathbf{y} \rightarrow q$ has matrix $U^T AU$ with respect to the new variables $\mathbf{y}$

Congruence

Congruent Matrices

Two $n \times n$ matrices A and B are called **congruent**, written $A \overset{c}{\sim} B$, if $B = U^T AU$ for some invertible matrix U.

1. $A \overset{c}{\sim} A$ *for all* A.
2. *If* $A \overset{c}{\sim} B$, *then* $B \overset{c}{\sim} A$.
3. *If* $A \overset{c}{\sim} B$ *and* $B \overset{c}{\sim} C$, *then* $A \overset{c}{\sim} C$.
4. *If* $A \overset{c}{\sim} B$, *then* A *is symmetric if and only if* B *is symmetric.*
5. *If* $A \overset{c}{\sim} B$, *then* $\text{rank} A = \text{rank} B$.

↳ converse is not always true

Example 8.9.6

Find the rank of the symmetric matrices $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Are they congruent?

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

A and B are not congruent

Assume $A \overset{c}{\sim} B$ then there is an invert.

U such that $B = U^T A U = U^T I U = U^T U$

$\det B = -1 = (\det U)^2$

contradiction since $(\det U)^2 \geq 0$

Theorem 8.9.7

Sylvester's Law of Inertia

If $A \overset{c}{\sim} B$, then A and B have the same number of positive eigenvalues, counting multiplicities.

Complete Diagonalization

- The **index** of a symmetric matrix A is the number of positive eigenvalues of A
- If $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is a quadratic form, the **index** and **rank** of q are defined to be, respectively, the index and rank of the matrix A. Hence the index and rank depend only on q and not on the way it is expressed.
- Now let $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be any quadratic form in n variables, of index k and rank r, where A is symmetric. We claim that new variables $\mathbf{z}$ can be found so that q is **completely diagonalized**—that is,

$$q(\mathbf{z}) = z_1^2 + \cdots + z_k^2 - z_{k+1}^2 - \cdots - z_r^2$$

$$D_3(1,1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- If $k \le r \le n$, let $D_n(k, r)$ denote the $n \times n$ diagonal matrix whose main diagonal consists of k ones, followed by $r - k$ minus ones, followed by $n - r$ zeros.

Complete Diagonalization continued

- We seek new variables $\mathbf{z}$ such that $q(\mathbf{z}) = \mathbf{z}^T D_n(k,r)\mathbf{z}$
- To determine $\mathbf{z}$: diagonalize A with an orthogonal matrix P_0 such that

$$P_0^T A P_0 = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0)$$

- By reordering the columns of P_0, assume that $\lambda_1, \ldots, \lambda_k$ are positive and $\lambda_{k+1}, \ldots, \lambda_r$ are negative.
- Let $D_0 = \mathrm{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_k}}, \frac{1}{\sqrt{-\lambda_{k+1}}}, \ldots, \frac{1}{\sqrt{-\lambda_r}}, 1, \ldots, 1\right)$

orthogonal columns

- Then $D_0^T D D_0 = D_n(k,r)$, so if $\mathbf{x} = (P_0 D_0)\mathbf{z}$, we obtain

$$q(\mathbf{z}) = \mathbf{z}^T D_n(k,r)\mathbf{z} = z_1^2 + \cdots + z_k^2 - z_{k+1}^2 - \cdots - z_r^2$$

Example 8.9.8

Completely diagonalize the quadratic form q in Example 8.9.4 and find the index and rank.

Eigenvals $12, -8, 4, 4$

index $A = 3$ (pos eigenvals) rank $A = 4$

Orthog. eigenvectors f_1, f_2, f_3, f_4 (see prev Eg)

$$P_0 = \begin{array}{cccc} 12 & 4 & 4 & -8 \\ \end{array}$$

$$P_0 = [f_1 \; f_3 \; f_4 \; f_2] \text{ orthogonal}$$

$$P_0^T A P_0 = \operatorname{diag}(12, 4, 4, -8)$$

$$D_0 = \operatorname{diag}\left(\frac{1}{\sqrt{12}}, 1/2, 1/2, -\frac{1}{\sqrt{8}}\right)$$

$$x = (D_0 P_0) z \quad \rightarrow \quad z = D_0^{-1} P_0^T x$$

$$z_1 = \sqrt{3}(x_1 - x_2 - x_3 + x_4)$$
$$z_2 = x_1 + x_2 + x_3\ x_4$$
$$z_3 = x_1 + x_2 - x_3 - x_4$$
$$z_4 = \sqrt{2}(x_1 - x_2 + x_3 - x_4)$$

$$q = z_1^2 + z_2^2 + z_3^2 - z_4^2$$

Theorem 8.9.9

Theorem

Let A and B be symmetric $n \times n$ matrices, and let $0 \leq k \leq r \leq n$.

1. A has index k and rank r if and only if $A \overset{c}{\sim} D_n(k, r)$.
2. $A \overset{c}{\sim} B$ if and only if they have the same rank and index.

Proof.

Recap

Today we saw:

- Diagonalization to find change of variables for quadratic forms
- Congruence
- Complete diagonalization

Next time: Chapter 9 - Change of Basis

MATH254: Linear Algebra
Lecture 25

Moira MacNeil

March 12, 2025

Last Time

1. Diagonalization to find change of variables for quadratic forms
2. Congruence
3. Complete diagonalization

Today

$T_A : \mathbb{R}^n \to \mathbb{R}^m$

$T_A(x) = Ax \qquad x \in \mathbb{R}^n$

1. A new chapter: Change of Basis
   Starting with the matrix of a linear transformation

Reminders:

- Midterm 2 is on Friday, March 14

The Matrix of a Linear Transformation

- Let $T : V \to W$ be a linear transformation where $\dim V = n$ and $\dim W = m$
- We would liketo describe the action of T as multiplication by an $m \times n$ matrix A
- Idea: convert $\mathbf{v}$ in V into a column in $\mathbb{R}^n$, multiply that column by A to get a column in $\mathbb{R}^m$, and convert this column back to get $T(\mathbf{v})$ in W
- Up to now the *order* of the vectors in a basis han't mattered
- Now $\to$ an **ordered basis** $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$, i.e., a basis where the order of the vectors is taken into account
- E.g., $\{\mathbf{b}_2, \mathbf{b}_1, \mathbf{b}_3\}$ is a different *ordered* basis from $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$

Coordinate Vectors

Coordinate Vector $C_B(\mathbf{v})$ of $\mathbf{v}$ for a basis B

The **coordinate vector** of $\mathbf{v}$ with respect to B is defined to be

$$C_B(\mathbf{v}) = (v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Note that $C_B(\mathbf{b}_i) = \mathbf{e}_i$ is column i of I_n.

Example 9.1.2

The coordinate vector for $\mathbf{v} = (2,1,3)$ with respect to the ordered basis $B = \{(1,1,0),(1,0,1),(0,1,1)\}$ of $\mathbb{R}^3$ is $C_B(\mathbf{v}) = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ because

$$\mathbf{v} = (2,1,3) = 0(1,1,0) + 2(1,0,1) + 1(0,1,1)$$

Theorem 9.1.3

Theorem

If V has dimension n and $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is any ordered basis of V, the coordinate transformation $C_B : V \to \mathbb{R}^n$ is an isomorphism. In fact, $C_B^{-1} : \mathbb{R}^n \to V$ is given by

$$C_B^{-1} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n \quad \text{for all} \quad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \text{ in } \mathbb{R}^n.$$

- $T : V \to W$ linear transformation where $\dim V = n$ and $\dim W = m$, $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ and D be ordered bases of V and W, respectively

- $C_B : V \to \mathbb{R}^n$ and $C_D : W \to \mathbb{R}^m$ are isomorphisms

- The composite $C_D T C_B^{-1} : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation $\implies$ a unique $m \times n$ matrix A exists such that $C_D T C_B^{-1} = T_A \iff C_D T = T_A C_B$

- The latter condition is $C_D[T(\mathbf{v})] = AC_B(\mathbf{v})$ for all $\mathbf{v}$ in V

$$\text{column } j \text{ of } A = AC_B(\mathbf{b}_j) = C_D[T(\mathbf{b}_j)]$$

$$A = \begin{bmatrix} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{bmatrix}$$

basis B

basis D

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "C_B"] & W \arrow[d, "C_D"] \\ \mathbb{R}^n \arrow[r, "T_A"] & \mathbb{R}^m \end{tikzcd}$$

Matrix of T corresponding to ordered bases

Matrix $M_{DB}(T)$ of $T : V \to W$ for bases D and B

This is called the **matrix of T corresponding to the ordered bases B and D**, and we use the following notation:

$$M_{DB}(T) = \left[\begin{array}{cccc} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{array} \right]$$

Theorem 9.1.5

Let $T : V \to W$ be a linear transformation where dim $V = n$ and dim $W = m$, and let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and D be ordered bases of V and W, respectively. Then the matrix $M_{DB}(T)$ just given is the unique $m \times n$ matrix A that satisfies

$$C_D T = T_A C_B$$

$T = C_D^{-1} T_A C_B$

Hence the defining property of $M_{DB}(T)$ is

$$\boxed{C_D[T(\mathbf{v})] = M_{DB}(T) C_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V}$$

The matrix $M_{DB}(T)$ is given in terms of its columns by

$$M_{DB}(T) = \begin{bmatrix} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{bmatrix}$$

Example 9.1.6

Define $T : \mathbf{P}_2 \to \mathbb{R}^2$ by $T(a + bx + cx^2) = (a + c, b - a - c)$ for all polynomials $a + bx + cx^2$. If $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ and $D = \{\mathbf{d}_1, \mathbf{d}_2\}$ where

$$\mathbf{b}_1 = 1, \mathbf{b}_2 = x, \mathbf{b}_3 = x^2 \quad \text{and} \quad \mathbf{d}_1 = (1,0), \mathbf{d}_2 = (0,1)$$

compute $M_{DB}(T)$ and verify Theorem 9.1.5.

$M_{DB}(T) = [C_D[T(b_i)] \cdots]$

$T(b_1) = d_1 - d_2$

$T(b_2) = d_2$

$T(b_3) = d_1 - d_2$

$C_D[T(b_1)] = [1 \quad -1]^T$

$C_D[T(b_2)] = [0 \quad 1]^T$

$C_D[T(b_3)] = [1 \quad -1]^T$

$$M_{DB}(T) = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \end{bmatrix}$$

$$v = a + bx + cx^2 = ab_1 + bb_2 + cb_3$$

$$T(v) = (a+c,\ b-a-c) = (a+c)d_1 + (b-a-c)d_2$$

$$C_D[T(v)] = \begin{bmatrix} a+c \\ b-a-c \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = M_{DB}(T)\,C_B(v)$$

Example 9.1.7

2x2 matrices with real entries

Suppose $T : \mathbf{M}_{22}(\mathbb{R}) \to \mathbb{R}^3$ is linear with matrix $M_{DB}(T) = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ where

$$B = \left\{ \overset{b_1}{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}, \overset{b_2}{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}, \overset{b_3}{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}, \overset{b_4}{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}} \right\} \text{ and } D = \{\overset{d_1}{(1,0,0)}, \overset{d_2}{(0,1,0)}, \overset{d_3}{(0,0,1)}\}$$

Compute $T(\mathbf{v})$ where $\mathbf{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. $= a b_1 + b b_2 + c b_3 + d b_4$

$$C_D[T(v)] = M_{DB}(T)\, C_B(v)$$

$$= \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a-b \\ b-c \\ c-d \end{bmatrix}$$

$$T(v) = (a-b)\, d_1 + (b-c)\, d_2 + (c-d)\, d_3$$

Example 9.1.8

Let A be an $m \times n$ matrix, and let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the matrix transformation induced by A: $T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$. If B and D are the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively (ordered as usual), then

$$M_{DB}(T_A) = A$$

In other words, the matrix of T_A corresponding to the standard bases is A itself.

$B = \{e_1, \dots, e_n\}$ D is standard basis of $\mathbb{R}^m$

$C_D(y) = y$ for all $y \in \mathbb{R}^m$

$$M_{DB}(T_A) = \begin{bmatrix} T_A(e_1) & T_A(e_2) & \cdots \end{bmatrix} = \begin{bmatrix} Ae_1 & Ae_2 & \cdots \end{bmatrix} = A$$

Example 9.1.9

Let V and W have ordered bases B and D, respectively. Let $\dim V = n$.

1. The identity transformation $1_V : V \to V$ has matrix $M_{BB}(1_V) = I_n$.
2. The zero transformation $0 : V \to W$ has matrix $M_{DB}(0) = 0$.

Theorem 9.1.10

Theorem

Let $V \xrightarrow{T} W \xrightarrow{S} U$ be linear transformations and let B, D, and E be finite ordered bases of V, W, and U, respectively. Then

$$M_{EB}(ST) = M_{ED}(S) \cdot M_{DB}(T)$$

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[rr, bend right, "ST"'] & W \arrow[r, "S"] & U \end{tikzcd}$$

Theorem 9.1.11

Theorem

Let $T : V \to W$ be a linear transformation, where $\dim V = \dim W = n$. The following are equivalent.

1. T is an isomorphism.
2. $M_{DB}(T)$ is invertible for all ordered bases B and D of V and W.
3. $M_{DB}(T)$ is invertible for some pair of ordered bases B and D of V and W.

When this is the case, $[M_{DB}(T)]^{-1} = M_{BD}(T^{-1})$.

Theorem 9.1.12

Theorem

Let $T : V \to W$ be a linear transformation where $\dim V = n$ and $\dim W = m$. If B and D are any ordered bases of V and W, then $\operatorname{rank} T = \operatorname{rank}[M_{DB}(T)]$.

The rank of a linear transformation $T : V \to W$ is $\operatorname{rank} T = \dim(\operatorname{im} T)$. Moreover, if A is any $m \times n$ matrix and $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is the matrix transformation, we showed that $\operatorname{rank}(T_A) = \operatorname{rank} A$. So it may not be surprising that $\operatorname{rank} T$ equals the rank of any matrix of T.

Example 9.1.13

$v = a b_1 + b b_2 + c b_3$

Define $T : \mathbf{P}_2 \to \mathbb{R}^3$ by $T(a + bx + cx^2) = (a - 2b, 3c - 2a, 3c - 4b)$ for a, b, $c \in \mathbb{R}$. Compute rank T.

rank T = rank $[M_{DB}(T)]$ for any bases $B \subseteq P_2$, $D \subseteq \mathbb{R}^3$ so choose the nicest ones

$$B = \{\underset{b_1}{1}, \underset{b_2}{x}, \underset{b_3}{x^2}\} \qquad D = \{\underset{d_1}{(1,0,0)}, \underset{d_2}{(0,1,0)}, \underset{d_3}{(0,0,1)}\}$$

$$M_{DB}(T) = \left[C_D[T(1)] \;\; C_D[T(x)] \;\; C_D[T(x^2)] \right]$$

$$T(1) = d_1 - 2d_2 \qquad T(x^2) = 3d_2 + 3d_3$$

$$T(x) = -2d_1 - 4d_3 \qquad C_D(T(v)) = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

$$M_{DB}(T) = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 0 & 3 \\ 0 & -4 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 0 \\ 0 & -4 & 3 \\ 0 & -4 & 3 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -3/4 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{rank} = 2$$

so $\operatorname{rank} T = 2$

Example 9.1.14

$\operatorname{im} T \subseteq W$

Let $T : V \to W$ be a linear transformation where dim $V = n$ and dim $W = m$. Choose an ordered basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_r, \mathbf{b}_{r+1}, \dots, \mathbf{b}_n\}$ of V in which $\{\mathbf{b}_{r+1}, \dots, \mathbf{b}_n\}$ is a basis of ker T, possibly empty. Then $\{T(\mathbf{b}_1), \dots, T(\mathbf{b}_r)\}$ is a basis of im T, so extend it to an ordered basis $D = \{T(\mathbf{b}_1), \dots, T(\mathbf{b}_r), \mathbf{f}_{r+1}, \dots, \mathbf{f}_m\}$ of W. Because $T(\mathbf{b}_{r+1}) = \cdots = T(\mathbf{b}_n) = \mathbf{0}$, we have

$$M_{DB}(T) = \begin{bmatrix} C_D[T(\mathbf{b}_1)] & \cdots & C_D[T(\mathbf{b}_r)] & C_D[T(\mathbf{b}_{r+1})] & \cdots & C_D[T(\mathbf{b}_n)] \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Incidentally, this shows that rank $T = r$ by the previous Theorem.

Recap

Today we saw:

- Coordinate vectors
- Matrices corresponding to ordered bases

Next time: Midterm II

MATH254: Linear Algebra
Lecture 26

Moira MacNeil

March 18, 2025

Last Time

1. Coordinate vectors
2. Matrix of a linear transformation

Today

1. Linear operators and similarity

Reminders:

- Assignment 5 (the last one!) is due on Friday March 28

Operators and Similarity

- Linear operator: a linear transformation $T : V \to V$ from a space V to itself
- If $T : V \to V$ is a linear operator where $\dim(V) = n$, it is possible to choose bases B and D of V such that the matrix $M_{DB}(T)$ has a very simple form: $M_{DB}(T) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ where $r = \text{rank } T$
- Finding bases B such that $M_{BB}(T)$ is as simple as possible tells us a lot about the operator T

The B-matrix of an Operator

Matrix $M_{BB}(T)$ of $T : V \to W$ for basis B

If $T : V \to V$ is an operator on a vector space V, and if B is an ordered basis of V, define $M_B(T) = M_{BB}(T)$ and call this the ***B*-matrix** of T.

Recall: If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator and $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$, then $C_E(\mathbf{x}) = \mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$

$\to M_E(T) = [T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)]$ is the **standard matrix** of the operator T

Theorem 9.2.2

Theorem

Let $T : V \to V$ be an operator where $\dim V = n$, and let B be an ordered basis of V.

1. $C_B(T(\mathbf{v})) = M_B(T)C_B(\mathbf{v})$ for all $\mathbf{v}$ in V.
2. If $S : V \to V$ is another operator on V, then $M_B(ST) = M_B(S)M_B(T)$.
3. T is an isomorphism if and only if $M_B(T)$ is invertible. In this case $M_D(T)$ is invertible for every ordered basis D of V.
4. If T is an isomorphism, then $M_B(T^{-1}) = [M_B(T)]^{-1}$.
5. If $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$, then
$M_B(T) = \left[\begin{array}{cccc} C_B[T(\mathbf{b}_1)] & C_B[T(\mathbf{b}_2)] & \cdots & C_B[T(\mathbf{b}_n)] \end{array} \right]$.

Change Matrix

Change Matrix $P_{D \leftarrow B}$ for bases B and D

With this in mind, define the **change matrix** $P_{D \leftarrow B}$ by

$$P_{D \leftarrow B} = M_{DB}(1_V) \quad \text{for any ordered bases } B \text{ and } D \text{ of } V$$

$T = 1_V$

$T(v) = v$

Theorem 9.2.4

Let $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ and D denote ordered bases of a vector space V. Then the change matrix $\underline{P_{D\leftarrow B}}$ is given in terms of its columns by

$$P_{D\leftarrow B} = \begin{bmatrix} C_D(\mathbf{b}_1) & C_D(\mathbf{b}_2) & \cdots & C_D(\mathbf{b}_n) \end{bmatrix} = M_{DB}(1_V) \tag{1}$$

and has the property that

$$\boxed{C_D(\mathbf{v}) = P_{D\leftarrow B}C_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V} \tag{2}$$

Moreover, if E is another ordered basis of V, we have

1. $P_{B\leftarrow B} = I_n$
2. $P_{D\leftarrow B}$ is invertible and $(P_{D\leftarrow B})^{-1} = P_{B\leftarrow D}$
3. $P_{E\leftarrow D}P_{D\leftarrow B} = P_{E\leftarrow B}$

Example 9.2.5 change matrix

In $\mathbf{P}_2$ find $P_{D \leftarrow B}$ if $B = \{\overset{b_1}{1}, \overset{b_2}{x}, \overset{b_3}{x^2}\}$ and $D = \{\overset{d_1}{1}, \overset{d_2}{(1-x)}, \overset{d_3}{(1-x)^2}\}$. Then use this to express $p = p(x) = a + bx + cx^2$ as a polynomial in powers of $(1-x)$.

To find $P_{D \leftarrow B}$ express B in terms of D

$b_1 = 1 = 1(1) + 0(1-x) + 0(1-x)^2 = 1d_1 + 0d_2 + 0d_3$

$b_2 = x = (1)1 + (-1)(1-x) + 0(1-x)^2 = 1d_1 - 1d_2 + 0d_3$

$b_3 = x^2 = (1)1 + (-2)(1-x) + 1(1-x)^2 = 1d_1 - 2d_2 + 1d_3$

$(1-x)^2 = x^2 - 2x + 1$

$$C_D(1) = [1\ 0\ 0]^T \qquad C_D(x^2) = [1\ -2\ 1]^T$$

$$C_D(x) = [1\ -1\ 0]^T$$

$$P_{D\leftarrow B} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$p = a + bx + cx^2 \qquad C_B(p) = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$C_D(p) = P_{D\leftarrow B}\, C_B(p)$$

$$= \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a+b+c \\ -b-2c \\ c \end{bmatrix}$$

$p(x)$ as polynomial of powers of $(1-x)$:

$$p(x) = (a+b+c) - (b+2c)(1-x) + c(1-x)^2$$

Relationship between matrices of an operator

- Let $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ and B_0 be two ordered bases of a vector space V
- Operator $T : V \to V$ has different matrices $M_B[T]$ and $M_{B_0}[T]$
- We have

$$C_{B_0}(\mathbf{v}) = P_{B_0 \leftarrow B} C_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V$$

and

$$C_B[T(\mathbf{v})] = M_B(T) C_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V$$

- Combining these (and writing $P = P_{B_0 \leftarrow B}$ for convenience) gives (for all $\mathbf{v}$ in V)

$$\begin{aligned} P M_B(T) C_B(\mathbf{v}) &= P C_B[T(\mathbf{v})] \\ &= C_{B_0}[T(\mathbf{v})] \\ &= M_{B_0}(T) C_{B_0}(\mathbf{v}) \\ &= \underline{M_{B_0}(T)} P C_B(\mathbf{v}) \end{aligned}$$

Relationship between matrices of an operator

$$P = P_{B_0 \leftarrow B}$$

- Because $C_B(\mathbf{b}_j)$ is the jth column of the identity matrix, it follows that

$$PM_B(T) = M_{B_0}(T)P$$

- Moreover P is invertible (in fact, $P^{-1} = P_{B \leftarrow B_0}$), so this gives

$$M_B(T) = P^{-1}M_{B_0}(T)P$$

- This asserts that $M_{B_0}(T)$ and $M_B(T)$ are similar matrices!

Theorem 9.2.6

Similarity Theorem

Let B_0 and B be two ordered bases of a finite dimensional vector space V. If $T : V \to V$ is any linear operator, the matrices $M_B(T)$ and $M_{B_0}(T)$ of T with respect to these bases are similar. More precisely,

$$M_B(T) = P^{-1} M_{B_0}(T) P$$

where $P = P_{B_0 \leftarrow B}$ is the change matrix from B to B_0.

$$P^{-1} = P_{B \leftarrow B_0}$$

Example 9.2.7

$B_0 = \left\{ b_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, b_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T(a, b, c) = (2a - b, b + c, c - 3a)$. If B_0 denotes the standard basis of $\mathbb{R}^3$ and $B = \{\underset{d_1}{(1, 1, 0)}, \underset{d_2}{(1, 0, 1)}, \underset{d_3}{(0, 1, 0)}\}$, find an invertible matrix P such that $P^{-1} M_{B_0}(T) P = M_B(T)$.

$$M_{B_0}(T) = \begin{bmatrix} C_{B_0}(2,0,-3) & C_{B_0}(-1,1,0) & C_{B_0}(0,1,1) \end{bmatrix}$$

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

$$M_B(T) = \begin{bmatrix} C_B(T(d_1)) & C_B(T(d_2)) & C_B(T(d_3)) \end{bmatrix}$$

$T(d_1) = (1, 1, -3)$

$T(d_2) = (2, 1, -2) \rightarrow C_B(T(d_2)) = [4 \ -2 \ -3]$

$T(d_3) = (-1, 1, 0) \rightarrow C_B(T(d_3)) = [-1 \ 0 \ 2]$

$$\begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (-3)\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + (-3)\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$C_B(T(d_1)) = [4 \ -3 \ -3]^T$

$$M_B(T) = \begin{bmatrix} 4 & 4 & 1 \\ -3 & -2 & 0 \\ -3 & -3 & 2 \end{bmatrix}$$

$$P = P_{B_0 \leftarrow B} = [C_{B_0}(d_1) \ \ C_{B_0}(d_2) \ \ C_{B_0}(d_3)]$$

$$= [C_{B_0}(1,1,0) \ \ C_{B_0}(1,0,1) \ \ C_{B_0}(0,1,0)]$$

$$= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

We can now verify $P^{-1}M_{B_0}(T)P = M_B(T)$

$\Rightarrow$ Verify that $PM_B(T) = M_{B_0}(T)P$:

$$PM_B(T) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 4 & -1 \\ -3 & -2 & 0 \\ -3 & -3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ -3 & -2 & 0 \end{bmatrix}$$

$$M_{B_0}(T)P = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 1 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ -3 & -2 & 0 \end{bmatrix}$$

MATH254: Linear Algebra
Lecture 27

Moira MacNeil

March 19, 2025

Last Time

1. Operators
2. Change matrix

Today

1. Linear operators and similarity

Reminders:

- Assignment 5 (the last one!) is due on Friday March 28

Recall: Coordinate Vectors

Coordinate Vector $C_B(\mathbf{v})$ of $\mathbf{v}$ for a basis B

The **coordinate vector** of $\mathbf{v}$ with respect to B is defined to be

$$C_B(\mathbf{v}) = (v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Note that $C_B(\mathbf{b}_i) = \mathbf{e}_i$ is column i of I_n.

Recall: Matrix of T corresponding to ordered bases

Matrix $M_{DB}(T)$ of $T : V \to W$ for bases D and B

This is called the **matrix of T corresponding to the ordered bases B and D**, and we use the following notation:

$$M_{DB}(T) = \left[\begin{array}{cccc} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{array} \right]$$

Recall: B-matrix of an Operator and Change Matrix

Matrix $M_{BB}(T)$ of $T : V \to V$ for basis B

If $T : V \to V$ is an operator on a vector space V, and if B is an ordered basis of V, define $M_B(T) = M_{BB}(T)$ and call this the ***B*-matrix** of T.

$$C_D(v) = P_{D\leftarrow B}\, C_B(v) \quad \forall v \in V$$

Change Matrix $P_{D\leftarrow B}$ for bases B and D

The **change matrix** $P_{D\leftarrow B}$ is

$$P_{D\leftarrow B} = M_{DB}(1_V) \quad \text{for any ordered bases } B \text{ and } D \text{ of } V$$

$$P_{D\leftarrow B} = [\,C_D(b_1) \quad C_D(b_2) \quad \cdots\,]$$

Recall: Theorem 9.2.6

Similarity Theorem

Let B_0 and B be two ordered bases of a finite dimensional vector space V. If $T : V \to V$ is any linear operator, the matrices $M_B(T)$ and $M_{B_0}(T)$ of T with respect to these bases are similar. More precisely,

$$M_B(T) = P^{-1} M_{B_0}(T) P$$

where $P = P_{B_0 \leftarrow B}$ is the change matrix from B to B_0.

$$P^{-1} = P_{B \leftarrow B_0}$$

Relationship to Diagonalization

- Recall: A square matrix is diagonalizable if and only if it is similar to a diagonal matrix
- Suppose an $n \times n$ matrix $A = M_{B_0}(T)$ is the matrix of some operator $T : V \to V$ with respect to an ordered basis B_0
- If another ordered basis B of V can be found such that $M_B(T) = D$ is diagonal, then Similarity Theorem shows how to find an invertible P such that $P^{-1}AP = D$

  $M_B(T) = P^{-1} M_{B_0}(T) P$
- The "algebraic" problem of finding P such that $P^{-1}AP$ is diagonal $\iff$ the "geometric" problem of finding a basis B such that $M_B(T)$ is diagonal
- Each $n \times n$ matrix A can be easily realized as the matrix of an operator

$$M_E(T_A) = A$$

where $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is given by $T_A(\mathbf{x}) = A\mathbf{x}$, and E is the standard basis of $\mathbb{R}^n$

Theorem 9.2.8

Theorem

Let A be an $n \times n$ matrix and let E be the standard basis of $\mathbb{R}^n$.

1. Let A' be similar to A, say $A' = P^{-1}AP$, and let B be the ordered basis of $\mathbb{R}^n$ consisting of the columns of P in order. Then $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is linear and
$$M_E(T_A) = A \quad \text{and} \quad M_B(T_A) = A'$$
2. If B is any ordered basis of $\mathbb{R}^n$, let P be the (invertible) matrix whose columns are the vectors in B in order. Then
$$\boxed{M_B(T_A) = P^{-1}AP}$$

Example 9.2.9

$T_A(x) = Ax$

Given $A = \begin{bmatrix} 10 & 6 \\ -18 & -11 \end{bmatrix}$, $P = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$, and $D = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$, verify that $P^{-1}AP = D$ and use this fact to find a basis B of $\mathbb{R}^2$ such that $M_B(T_A) = D$.

$P^{-1}AP = D \Leftrightarrow AP = PD$ (verify)

Let $B = \left\{ \overset{b_1}{\begin{bmatrix} 2 \\ -3 \end{bmatrix}}, \overset{b_2}{\begin{bmatrix} -1 \\ 2 \end{bmatrix}} \right\}$ (columns of P in order)

$M_B(T_A) = P^{-1}AP = D$

$M_B(T_A) = \left[C_B\left(T_A \begin{bmatrix} 2 \\ -3 \end{bmatrix}\right) \; C_B\left(T_A \begin{bmatrix} -1 \\ 2 \end{bmatrix}\right) \right]$

$$T_A(b_1) = Ab_1 = \begin{bmatrix} 10 & 6 \\ -18 & -11 \end{bmatrix}\begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

$$T_A(b_2) = Ab_2 = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ -3 \end{bmatrix} = 1\begin{bmatrix} 2 \\ -3 \end{bmatrix} + 0\begin{bmatrix} -1 \\ 2 \end{bmatrix} \qquad C_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ -4 \end{bmatrix} = 0\begin{bmatrix} 2 \\ -3 \end{bmatrix} + (-2)\begin{bmatrix} -1 \\ 2 \end{bmatrix} \qquad C_B = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

$$M_B(T_A) = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} = D$$

Another way to find P

- Let A be an $n \times n$ matrix
- We have a new way to find an invertible matrix P such that $P^{-1}AP$ is diagonal
- Idea: find a basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ of $\mathbb{R}^n$ such that $M_B(T_A) = D$ is diagonal and take $P = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$ to be the matrix with the $\mathbf{b}_j$ as columns
- Then, by Theorem 9.2.4
$$P^{-1}AP = M_B(T_A) = D$$

Similarity Invariants

- **Similarity invariant** $\to$ property of $n \times n$ A if when A has the property, every matrix similar to A also has the property.
- E.g., rank: If $T : V \to V$ is a linear operator, the matrices of T wrt various bases of V all have the same rank (they're similar)
- Hence the rank of T could be *defined* to be the rank of A, where A is *any* matrix of T
- If $T : V \to V$ is a linear operator on a finite dimensional space V, define the **determinant** of T (denoted $\det T$) by

  $$\det T = \det M_B(T), \quad B \text{ any basis of } V$$

  Independent of the choice of basis because $M_B(T)$ and $M_D(T)$ are similar
- The **trace** of T ($\operatorname{tr} T$) can be defined by

  $$\operatorname{tr} T = \operatorname{tr} M_B(T), \quad B \text{ any basis of } V$$

Example 9.2.11

$$M_B(ST) = M_B(S)M_B(T)$$

Let S and T denote linear operators on the finite dimensional space V. Show that

$$\det (ST) = \det S \det T$$

Let B be a basis of V

$$\begin{aligned}\det(ST) &= \det M_B(ST)\\ &= \det\left[M_B(S)M_B(T)\right]\\ &= \det M_B(S) \det M_B(T)\\ &= \det S \det T\end{aligned}$$

Characteristic polynomial is another similarity invariant

- If A and A' are similar matrices, then $c_A(x) = c_{A'}(x)$
- If $T : V \to V$ is a linear operator on the finite dimensional space V, define the **characteristic polynomial** of T by

$$c_T(x) = c_A(x) \text{ where } A = M_B(T),\ B \text{ any basis of } V$$

- I.e., the characteristic polynomial of an operator T is the characteristic polynomial of *any* matrix representing T.

Example 9.2.12

Compute the characteristic polynomial $c_T(x)$ of the operator $T : \mathbf{P}_2 \to \mathbf{P}_2$ given by $T(a + bx + cx^2) = (b + c) + (a + c)x + (a + b)x^2$.

Want to find char. poly of $M_B(T)$

Let $B = \{1, x, x^2\}$ $\quad M_B(T) = [C_B(T(b_1)) \dots]$

$T(1) = x + x^2 \rightarrow C_B = [0 \ 1 \ 1]^T$

$T(x) = 1 + x^2 \rightarrow C_B = [1 \ 0 \ 1]^T$

$T(x^2) = 1 + x \rightarrow C_B = [1 \ 1 \ 0]^T$

$$M_B(T) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$c_T(x) = \det(xI - M_B(T))$$

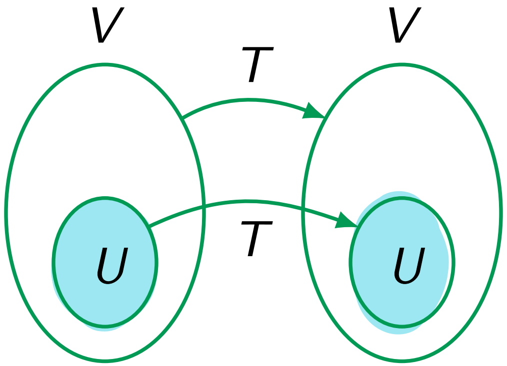
$$= \begin{vmatrix} x & -1 & -1 \\ -1 & x & -1 \\ -1 & -1 & x \end{vmatrix} = x^3 - 3x - 2$$

$$= (x+1)^2(x-2)$$

T-invariant Subspace

T-invariant Subspace

Let $T : V \to V$ be an operator. A subspace $U \subseteq V$ is called ***T*-invariant** if $T(U) = \{T(\mathbf{u}) \mid \mathbf{u} \text{ in } U\} \subseteq U$, that is, $T(\mathbf{u}) \in U$ for every vector $\mathbf{u} \in U$. Hence T is a linear operator on the vector space U.



Example 9.3.2

Let $T : V \to V$ be any linear operator. Then:

1. $\{\mathbf{0}\}$ and V are T-invariant subspaces.
2. Both $\ker T$ and $\operatorname{im} T = T(V)$ are T-invariant subspaces.
3. If U and W are T-invariant subspaces, so are $T(U)$, $U \cap W$, and $U + W$.

Example 9.3.3

Define $T : \mathbb{R}^3 \to \mathbb{R}^3$ by $T(a, b, c) = (3a + 2b, b - c, 4a + 2b - c)$. Then $U = \{(a, b, a) \mid a, b \text{ in } \mathbb{R}\}$ is T-invariant because

$$T(a, b, a) = (3a + 2b, b - a, 3a + 2b)$$

is in U for all a and b (the first and last entries are equal).

Example 9.3.4

Let $T : V \to V$ be a linear operator, and suppose that $U = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ is a subspace of V. Show that U is T-invariant if and only if $T(\mathbf{u}_i)$ lies in U for each $i = 1, 2, \ldots, k$.

($\Rightarrow$) Let $u \in U$ then we have

$$u = r_1 u_1 + r_2 u_2 + \cdots + r_k u_k$$ then

$$T(u) = T(r_1 u_1) + \cdots + T(r_k u_k)$$
$$= r_1 T(u_1) + \cdots + r_k T(u_k)$$

$T(u)$ is in U if each $T(u_i)$ is in U

U is T-invar if $T(u_i) \in U$ $\forall$ i

($\Leftarrow$) let $T(u_i) \in U$ then $T(u) \in U \Rightarrow U$ T-invar

Example 9.3.5

Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(a,b) = (b,-a)$. Show that $\mathbb{R}^2$ contains no T-invariant subspace except 0 and $\mathbb{R}^2$.

Suppose toward contradiction that $U \subset \mathbb{R}^2$ s.t. $U \neq \{0\}$, $U \neq \mathbb{R}^2$. Then U has dimension 1, so $U = \mathbb{R}x$ for $x \neq 0$. $T(x) \in U$ let $T(x) = rx$, $r \in \mathbb{R}$, $x = (a,b)$ this is $r(a,b) = (b,-a)$

$ra = b$ $\qquad r(ra) = -a$

$rb = -a$ $\qquad r^2 a = -a \to (r^2+1)a = 0$

$r^2 = -1$ $(\Rightarrow\Leftarrow)$ $\qquad a=0, b=0$ $(\Rightarrow\Leftarrow)$

Recap

Today we saw:

- Similar matrices of linear operators
- Similarity invariants
- Invariant subspaces

Next time: Restrictions, direct sums

MATH254: Linear Algebra
Lecture 28

Moira MacNeil

March 21, 2025

Last Time

1. Linear operators and similarity
2. Invariant subspaces

Today

1. Restrictions of linear operators
2. Direct sums
3. Reducible linear operators

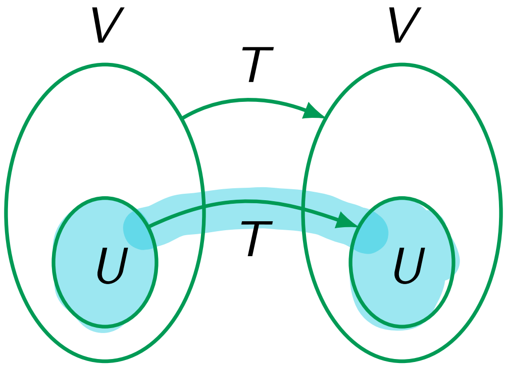
Reminders:

- Assignment 5 is due on Friday March 28

T-invariant Subspace

T-invariant Subspace

Let $T : V \to V$ be an operator. A subspace $U \subseteq V$ is called ***T*-invariant** if $T(U) = \{T(\mathbf{u}) \mid \mathbf{u} \text{ in } U\} \subseteq U$, that is, $T(\mathbf{u}) \in U$ for every vector $\mathbf{u} \in U$. Hence T is a linear operator on the vector space U.



Restriction

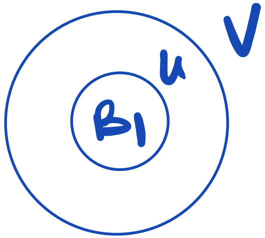
Restriction of an Operator

Let $T : V \to V$ be a linear operator. If U is any T-invariant subspace of V, then

$$T : U \to U$$

is a linear operator on the subspace U, called the **restriction** of T to U.

Theorem 9.3.7



Theorem

Let $T : V \to V$ be a linear operator where V has dimension n and suppose that U is any T-invariant subspace of V. Let $B_1 = \{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ be any basis of U and extend it to a basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ of V in any way. Then $M_B(T)$ has the block triangular form

$$M_B(T) = \begin{bmatrix} M_{B_1}(T) & Y \\ 0 & Z \end{bmatrix}$$

where Z is $(n-k) \times (n-k)$ and $M_{B_1}(T)$ is the matrix of the restriction of T to U.

Theorem 9.3.8

Theorem

Let A be a block upper triangular matrix, say

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & A_{22} & A_{23} & \cdots & A_{2n} \\ 0 & 0 & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & A_{nn} \end{bmatrix}$$

$A_{ii} \to$ square block

$A_{ij} \to$ matrix (block)

where the diagonal blocks are square. Then:

1. $\det A = (\det A_{11})(\det A_{22})(\det A_{33}) \cdots (\det A_{nn})$.
2. $c_A(x) = c_{A_{11}}(x) c_{A_{22}}(x) c_{A_{33}}(x) \cdots c_{A_{nn}}(x)$.

Example 9.3.9

$c_T(x) \Rightarrow$ char poly of $T \Rightarrow$ char poly of $M_B(T)$

Consider the linear operator $T : \mathbf{P}_2 \to \mathbf{P}_2$ given by

$$T(a + bx + cx^2) = (-2a - b + 2c) + (a + b)x + (-6a - 2b + 5c)x^2$$

Show that $U = \text{span}\{x, 1 + 2x^2\}$ is T-invariant, use it to find a block upper triangular matrix for T, and use that to compute $c_T(x)$.

$U = \text{span}\{u_1, \cdots, u_n\}$ is T invar $\Leftrightarrow$ $T(u_i) \in U$

$T(x) = -1 + x - 2x^2 = x - (1+2x^2) + 0x^2 \in U$

$T(1+2x^2) = 2 + x + 4x^2 = x + 2(1+2x^2) \in U$

$\therefore$ U is T-invarient

Note $\{x, 1+2x^2\}$ is independent, so let

$B_1 = \{x, 1+2x^2\}$ is a basis for U

Extend to a basis of P_2 in any way.

Let $B = \{x, 1+2x^2, x^2\}$

$$T(x^2) = 2 + 5x^2 = 2(1+2x^2) + x^2$$

$$M_B(T) = \left[C_B(T(x)) \;\; C_B(T(1+2x^2)) \;\; C_B(T(x^2)) \right]$$

$$= \left[C_B(-1+x-2x^2) \;\; C_B(2+x+4x^2) \;\; C_B(2+5x^2) \right]$$

$$= \begin{array}{cl} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ -1 & 2 & 2 \\ \hline 0 & 0 & 1 \end{array}\right] & \begin{array}{l} \leftarrow y \\ \\ \leftarrow z \end{array} \end{array}$$

$$c_T(x) = \det \left[\begin{array}{cc|c} x-1 & -1 & 0 \\ 1 & x-2 & -2 \\ \hline 0 & 0 & x-1 \end{array}\right]$$

$$= \begin{vmatrix} x-1 & -1 \\ 1 & x-2 \end{vmatrix} \, |x-1| = [(x-1)(x-2)+1](x-1)$$

$$= (x^2 - 3x + 3)(x-1)$$

Eigenvalues

$A v = \lambda v$

- A real number λ is called an **eigenvalue** of an operator $T : V \to V$ if

$$T(\mathbf{v}) = \lambda \mathbf{v}$$

holds for some $\mathbf{v} \neq 0$ in V $\to$ $\mathbf{v}$ is an **eigenvector** of T corresponding to λ

- The subspace

$$E_\lambda(T) = \{\mathbf{v} \text{ in } V \mid T(\mathbf{v}) = \lambda \mathbf{v}\}$$

is called the **eigenspace** of T corresponding to λ

- If A is an $n \times n$ matrix, $\lambda \in \mathbb{R}$ is an eigenvalue of $T_A : \mathbb{R}^n \to \mathbb{R}^n \iff \lambda$ is an eigenvalue of A and

$$E_\lambda(T_A) = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \lambda \mathbf{x}\} = E_\lambda(A)$$

$T_A(x) = Ax$

Theorem 9.3.10

Theorem

Let $T : V \to V$ be a linear operator where $\dim V = n$, let B denote any ordered basis of V, and let $C_B : V \to \mathbb{R}^n$ denote the coordinate isomorphism. Then:

1. The eigenvalues λ of T are precisely the eigenvalues of the matrix $M_B(T)$ and thus are the roots of the characteristic polynomial $c_T(x)$.
2. In this case the eigenspaces $E_\lambda(T)$ and $E_\lambda[M_B(T)]$ are isomorphic via the restriction $C_B : E_\lambda(T) \to E_\lambda[M_B(T)]$.

Example 9.3.11

Find the eigenvalues and eigenspaces for $T : \mathbf{P}_2 \to \mathbf{P}_2$ given by

$$T(a + bx + cx^2) = (2a + b + c) + (2a + b - 2c)x - (a + 2c)x^2$$

$B = \{1, x, x^2\}$

$M_B(T) = \begin{bmatrix} C_B(T(1)) & C_B(T(x)) & C_B(T(x^2)) \end{bmatrix}$

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

$c_T(x) =$
$\det(xI - M_B(T)) =$
$(x+1)^2(x-3)$

$\lambda_1 = -1$

$$\begin{bmatrix} -3 & -1 & -1 \\ -2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} -3 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$x_3 = t \qquad x_1 = t \qquad x_2 = -4t$

$$\underline{x} = \begin{bmatrix} 1 \\ -4 \\ 1 \end{bmatrix} \qquad E_{-1}[M_B(T)] = \mathbb{R}\begin{bmatrix} 1 \\ -4 \\ 1 \end{bmatrix}$$

$$= \operatorname{span}\{\underline{x}\}$$

$$\Rightarrow E_{-1}(T) = \operatorname{span}\{1 - 4x + x^2\}$$

$\lambda_2 = 3$ $\qquad E_3[M_B(T)] = \operatorname{span}\left\{\begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix}\right\}$

$$E_3(T) = \operatorname{span}\{5 + 6x - x^2\}$$

Theorem 9.3.12

Theorem

Each eigenspace of a linear operator $T : V \to V$ is a T-invariant subspace of V.

Vector Spaces as Sums of Subspaces

- Sometimes we can write vectors of V as a sum of vectors in two subspaces

- E.g., $\mathbf{M}_{nn}$ has subspaces

$$U = \{P \text{ in } \mathbf{M}_{nn} \mid P \text{ is symmetric } ((P^T = P))\}$$

$$W = \{Q \text{ in } \mathbf{M}_{nn} \mid Q \text{ is skew symmetric } (Q^T = -Q)\}$$

- Every matrix A in $\mathbf{M}_{nn}$ can be written as the sum of a matrix in U and a matrix in W:

$$A = \underbrace{\frac{1}{2}(A + A^T)}_{u \in U} + \underbrace{\frac{1}{2}(A - A^T)}_{w \in W}$$

where $\frac{1}{2}(A + A^T)$ is symmetric and $\frac{1}{2}(A - A^T)$ is skew symmetric

- We can generalize this!

Direct Sum of Subspaces

Direct Sum of Subspaces

A vector space V is said to be the **direct sum** of subspaces U and W if

$$U \cap W = \{\mathbf{0}\} \quad \text{and} \quad U + W = V$$

In this case we write $V = U \oplus W$. Given a subspace U, any subspace W such that $V = U \oplus W$ is called a **complement** of U in V.

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \text{ in } U \text{ and } \mathbf{w} \text{ in } W\}$$
$$U \cap W = \{\mathbf{v} \mid \mathbf{v} \text{ lies in both } U \text{ and } W\}$$

Example 9.3.14

In the space $\mathbb{R}^5$, consider the subspaces $U = \{(a, b, c, 0, 0) \mid a, b, \text{ and } c \text{ in } \mathbb{R}\}$ and $W = \{(0, 0, 0, d, e) \mid d \text{ and } e \text{ in } \mathbb{R}\}$. Show that $\mathbb{R}^5 = U \oplus W$.

If $x = (a, b, c, d, e) \in \mathbb{R}^5$ then

$x = (a, b, c, 0, 0) + (0, 0, 0, d, e)$

$\Rightarrow x \in U + W$

Let $x = (a, b, c, d, e) \in U \cap W$ then

$d = e = 0$ b/c $x \in U$ and

$a = b = c = 0$ b/c $\in W$

$\Rightarrow x = 0$ and this is the only vect. in $U \cap W$.

Example 9.3.16

Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be a basis of a vector space V, and partition it into two parts: $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ and $\{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$. If $U = \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ and $W = \text{span}\{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$, show that $V = U \oplus W$.

If $v \in U \cap W$ then $v = a_1 e_1 + \cdots + a_k e_k$ and
$v = b_{k+1} e_{k+1} + \cdots + b_n e_n$
for some $a_i, b_i \in \mathbb{R}$. e_i are independent (basis) $\Rightarrow a_i = b_i = 0 \Rightarrow v = 0$ $U \cap W = \{0\}$

$v \in V$ $\quad v = \underbrace{v_1 e_1 + \cdots + v_k e_k}_{u \in U} + \underbrace{v_{k+1} e_{k+1} + \cdots + v_n e_n}_{w \in W} = u + w \in U + W$

Theorem 9.3.17

$P_2 \to \mathbb{R}$ $a+b+c+d=0$

Theorem

Let U and W be subspaces of a finite dimensional vector space V. The following three conditions are equivalent:

1. $V = U \oplus W$.
2. Each vector $\mathbf{v}$ in V can be written uniquely in the form

$$\mathbf{v} = \mathbf{u} + \mathbf{w} \qquad \mathbf{u} \text{ in } U, \mathbf{w} \text{ in } W$$

3. If $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ are bases of U and W, respectively, then $B = \{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_1, \dots, \mathbf{w}_m\}$ is a basis of V.

(The uniqueness in (2) means that if $\mathbf{v} = \mathbf{u}_1 + \mathbf{w}_1$ is another such representation, then $\mathbf{u}_1 = \mathbf{u}$ and $\mathbf{w}_1 = \mathbf{w}$.)

Theorem 9.3.18

Theorem

If a finite dimensional vector space V is the direct sum $V = U \oplus W$ of subspaces U and W, then

$$\dim\ V = \dim\ U + \dim\ W$$

Theorem 9.3.19

Let $T : V \to V$ be a linear operator where V has dimension n. Suppose $V = U_1 \oplus U_2$ where both U_1 and U_2 are T-invariant. If $B_1 = \{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ and $B_2 = \{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ are bases of U_1 and U_2 respectively, then

$$B = \{\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$$

is a basis of V, and $M_B(T)$ has the block diagonal form

$$M_B(T) = \begin{bmatrix} M_{B_1}(T) & 0 \\ 0 & M_{B_2}(T) \end{bmatrix}$$

where $M_{B_1}(T)$ and $M_{B_2}(T)$ are the matrices of the restrictions of T to U_1 and to U_2 respectively.

Reducible Operator

Reducible Linear Operator

The linear operator $T : V \to V$ is said to be **reducible** if nonzero T-invariant subspaces U_1 and U_2 can be found such that $V = U_1 \oplus U_2$.

Example 9.3.21

Let $T : V \to V$ be a linear operator satisfying $T^2 = 1_V$ (such operators are called **involutions**). Define

$$U_1 = \{\mathbf{v} \mid T(\mathbf{v}) = \mathbf{v}\} \quad \text{and} \quad U_2 = \{\mathbf{v} \mid T(\mathbf{v}) = -\mathbf{v}\}$$

1. Show that $V = U_1 \oplus U_2$.
2. If dim $V = n$, find a basis B of V such that $M_B(T) = \begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix}$ for some k.
3. Conclude that, if A is an $n \times n$ matrix such that $A^2 = I$, then A is similar to $\begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix}$ for some k.

Example 9.3.22

Consider the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b \\ b \end{bmatrix}$. Show that $U_1 = \mathbb{R}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is T-invariant but that U_1 has no T-invariant complement in $\mathbb{R}^2$.

Recap

Today we saw:

- Restrictions of linear operators
- Direct sums
- Reducible linear operators

Next time: Operators and characteristic polynomials, invariant subspaces

MATH254: Linear Algebra
Lecture 29

Moira MacNeil

March 25, 2025

Last Time

1. Restrictions of linear operators

2. Direct sums

Today

1. Direct sums
2. Reducible linear operators
3. Chapter 10: Inner product spaces

Reminders:

- Assignment 5 is due on Friday March 28

T-invariant Subspace and Restrictions

T-invariant Subspace

Let $T : V \to V$ be an operator. A subspace $U \subseteq V$ is called **T-invariant** if $T(U) = \{T(\mathbf{u}) \mid \mathbf{u} \text{ in } U\} \subseteq U$, that is, $T(\mathbf{u}) \in U$ for every vector $\mathbf{u} \in U$. Hence T is a linear operator on the vector space U.

Restriction of an Operator

Let $T : V \to V$ be a linear operator. If U is any T-invariant subspace of V, then

$$T : U \to U$$

is a linear operator on the subspace U, called the **restriction** of T to U.

Recall: Direct Sum of Subspaces

Direct Sum of Subspaces

A vector space V is said to be the **direct sum** of subspaces U and W if

$$U \cap W = \{\mathbf{0}\} \quad \text{and} \quad U + W = V$$

In this case we write $V = U \oplus W$. Given a subspace U, any subspace W such that $V = U \oplus W$ is called a **complement** of U in V.

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \text{ in } U \text{ and } \mathbf{w} \text{ in } W\}$$
$$U \cap W = \{\mathbf{v} \mid \mathbf{v} \text{ lies in both } U \text{ and } W\}$$

Recall: Theorem 9.3.17

Theorem

Let U and W be subspaces of a finite dimensional vector space V. The following three conditions are equivalent:

1. $V = U \oplus W$.
2. Each vector $\mathbf{v}$ in V can be written uniquely in the form
$$\mathbf{v} = \mathbf{u} + \mathbf{w} \qquad \mathbf{u} \text{ in } U, \mathbf{w} \text{ in } W$$
3. If $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ and $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ are bases of U and W, respectively, then $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is a basis of V.

(The uniqueness in (2) means that if $\mathbf{v} = \mathbf{u}_1 + \mathbf{w}_1$ is another such representation, then $\mathbf{u}_1 = \mathbf{u}$ and $\mathbf{w}_1 = \mathbf{w}$.)

Theorem 9.3.18

Theorem

If a finite dimensional vector space V is the direct sum $V = U \oplus W$ of subspaces U and W, then

$$\dim\ V = \dim\ U + \dim\ W$$

$$M_B(T) = \begin{bmatrix} M_{B_1}(T) & Y \\ 0 & Z \end{bmatrix}$$

B_1 basis of U_1

U_1 is T invar.

Theorem 9.3.19

Let $T : V \to V$ be a linear operator where V has dimension n. Suppose $V = U_1 \oplus U_2$ where both U_1 and U_2 are T-invariant. If $B_1 = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ and $B_2 = \{\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$ are bases of U_1 and U_2 respectively, then

$$B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$$

is a basis of V, and $M_B(T)$ has the block diagonal form

$$M_B(T) = \begin{bmatrix} M_{B_1}(T) & 0 \\ 0 & M_{B_2}(T) \end{bmatrix}$$

where $M_{B_1}(T)$ and $M_{B_2}(T)$ are the matrices of the restrictions of T to U_1 and to U_2 respectively.

Reducible Operator

Reducible Linear Operator

The linear operator $T : V \to V$ is said to be **reducible** if nonzero T-invariant subspaces U_1 and U_2 can be found such that $V = U_1 \oplus U_2$.

Example 9.3.21

$1_V : V \to V \qquad 1_V(v) = v$

Let $T : V \to V$ be a linear operator satisfying $T^2 = 1_V$ (such operators are called **involutions**). Define

$$U_1 = \{\mathbf{v} \mid T(\mathbf{v}) = \mathbf{v}\} \quad \text{and} \quad U_2 = \{\mathbf{v} \mid T(\mathbf{v}) = -\mathbf{v}\}$$

$M_{B_i}(T)$

1. Show that $V = U_1 \oplus U_2$. ✓
2. If dim $V = n$, find a basis B of V such that $M_B(T) = \begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix}$ for some k.
3. Conclude that, if A is an $n \times n$ matrix such that $A^2 = I$, then A is similar to $\begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix}$ for some k.

1. $U \cap W = \{0\}$ $\quad U + W = V$

Assume $v \in U_1 \cap U_2$ then
$v = T(v) = -v$ so $v = 0 \Rightarrow$
$U_1 \cap U_2 = \{0\}$

Given $v \in V$

$$v = \frac{1}{2}\Big[\underbrace{[v + T(v)]}_{\in U_1} + \overbrace{[v - T(v)]}^{\in U_2}\Big]$$

$= 1v$

$$T[v + T(v)] = T(v) + T^2(v) = \underset{\in U_1}{T(v) + v}$$

$$T[v - T(v)] = T(v) - T^2(v) = \underset{\in U_2}{v - T(v)}$$

2. U_1 and U_2 are T-invarient
(check as exercise)

From prev. theorem if bases
$B_1 = \{b_1, \ldots, b_k\}$ and $B_2 = \{b_{k+1}, \ldots, b_n\}$
of U_1 U_2 resp. can be found

$$M_{B_1}(T) = I_k \qquad M_{B_2}(T) = -I_{n-k}$$

$$M_{B_1}(T) = \big[C_{B_1}[T(b_1)] \;\cdots\; C_{B_1}[T(b_k)]\big]$$
$$= \big[C_{B_1}(b_1) \cdots C_{B_1}(b_k)\big] = I_k$$

similar for $M_{B_2}(T) = -I_{n-k}$

Thus $B = \{b_1, \dots, b_k, b_{k+1}, \dots b_n\}$

3. Given A st $A^2 = I$ (A is $n \times n$)

$T_A : \mathbb{R}^n \to \mathbb{R}^n \qquad T_A(x) = Ax$

$$(T_A)^2(x) = A^2 x = Ix = x \quad \forall x \in \mathbb{R}^n$$

$\Rightarrow (T_A)^2 = 1_V$. By (2) there exists a basis B of $\mathbb{R}^n$ such that

$$M_B(T_A) = \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix}$$

$M_B(T_A) = P^{-1}AP$ for some inv. P $\therefore$ A is similar to $M_B(T_A)$

Example 9.3.22

$\circledast$ $U = \operatorname{span}\{u_i\}$, U T invar $\Leftrightarrow$ $T(u_i) \in U$

Consider the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b \\ b \end{bmatrix}$. Show that $U_1 = \mathbb{R} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is T-invariant but that U_1 has no T-invariant complement in $\mathbb{R}^2$. $\circledast$

$U_1 = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$ $\quad T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in U_1$ $\Rightarrow$ U_1 is T-invar.

Assume U_2 is T-invar. complement to U_1

Then $U_1 \oplus U_2 = \mathbb{R}^2$

$\dim \mathbb{R}^2 = \dim U_1 + \dim U_2 \Rightarrow \dim U_2 = 2-1=1$

Let $U_2 = \operatorname{span}\left\{ u_2 = \begin{bmatrix} p \\ q \end{bmatrix} \right\}$ Claim $u_2 \notin U_1$

since if $u_2 \in U_1 \Rightarrow$ $u_2 \in U_1 \cap U_2 = \{0\}$
$u_2 = 0$

then $U_2 = \{0\} \Rightarrow\Leftarrow$ $(\dim U_2 = 1)$

$u_2 \notin U_1 \Rightarrow q \neq 0$

$T(u_2) \in U_2 \rightarrow T(u_2) = \lambda u_2 = \lambda \begin{bmatrix} p \\ q \end{bmatrix}$

$$\begin{bmatrix} p+q \\ q \end{bmatrix} = T\begin{bmatrix} p \\ q \end{bmatrix} = \lambda \begin{bmatrix} p \\ q \end{bmatrix} \qquad \lambda \in \mathbb{R}$$

so $p + q = \lambda p$ $\qquad q = \lambda q$ $(q \neq 0)$

$\Rightarrow \lambda = 1$

$p + q = p$

$q = 0 \Rightarrow\Leftarrow$

$\therefore$ no T invar compl of U_1

Inner Product Spaces

Inner Product Spaces

An **inner product** on a real vector space V is a function that assigns a real number $\langle \mathbf{v}, \mathbf{w} \rangle$ to every pair $\mathbf{v}$, $\mathbf{w}$ of vectors in V in such a way that the following axioms are satisfied.

P1. $\langle \mathbf{v}, \mathbf{w} \rangle$ *is a real number for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V.

P2. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V.

P3. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$ *for all* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *in* V.

P4. $\langle r\mathbf{v}, \mathbf{w} \rangle = r\langle \mathbf{v}, \mathbf{w} \rangle$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V *and all* r *in* $\mathbb{R}$.

P5. $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ *for all* $\mathbf{v} \neq \mathbf{0}$ *in* V.

A real vector space V with an inner product $\langle\ ,\ \rangle$ will be called an **inner product space**.

Example 10.1.2

$\mathbb{R}^n$ is an inner product space with the dot product as inner product:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v} \cdot \mathbf{w} \quad \text{for all } \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$$

This is also called the **euclidean** inner product, and $\mathbb{R}^n$, equipped with the dot product, is called **euclidean** *n*-**space**.

Example 10.1.3

If A and B are $m \times n$ matrices, define $\langle A, B\rangle = \operatorname{tr}(AB^T)$ where $\operatorname{tr}(X)$ is the trace of the square matrix X. Show that $\langle\ ,\rangle$ is an inner product in $\mathbf{M}_{mn}$. Check P1 – P5

P1 clear tr(X) X real is also real.

P2 $\langle A, B\rangle = \operatorname{tr}(AB^T) = \operatorname{tr}[(AB^T)^T] = \operatorname{tr}(BA^T)$
$= \langle B, A\rangle$

P3, P4 follow since trace is lin. transform $M_{mn} \to \mathbb{R}$

P5 let r_i denote row i of A
then entry (i,j) of $AA^T = r_i \cdot r_j$

$\langle A, A \rangle = tr(AA^T) = r_1 \cdot r_1 + \cdots + r_m \cdot r_m$

$r_j \cdot r_j$ = sum of squares of entries of r_j

so $\langle A, A \rangle > 0$ sum of sum of squares

Theorem 10.1.4

Theorem

Let $\langle\ ,\rangle$ be an inner product on a space V; let $\mathbf{v}$, $\mathbf{u}$, and $\mathbf{w}$ denote vectors in V; and let r denote a real number.

1. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
2. $\langle \mathbf{v}, r\mathbf{w} \rangle = r\langle \mathbf{v}, \mathbf{w} \rangle = \langle r\mathbf{v}, \mathbf{w} \rangle$
3. $\langle \mathbf{v}, \mathbf{0} \rangle = 0 = \langle \mathbf{0}, \mathbf{v} \rangle$
4. $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$

Theorem 10.1.5

Theorem

If A is any $n \times n$ positive definite matrix, then

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A\mathbf{y} \text{ for all columns } \mathbf{x}, \mathbf{y} \text{ in } \mathbb{R}^n$$

defines an inner product on $\mathbb{R}^n$, and every inner product on $\mathbb{R}^n$ arises in this way.

Every inner product on $\mathbb{R}^n$ corresponds to a positive definite $n \times n$ matrix. In particular, the dot product corresponds to the identity matrix I_n.

Remark

If we refer to the inner product space $\mathbb{R}^n$ without specifying the inner product, we mean that the dot product is to be used.

Example 10.1.6

Let the inner product $\langle\ ,\rangle$ be defined on $\mathbb{R}^2$ by

$$\left\langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\rangle = 2v_1w_1 - v_1w_2 - v_2w_1 + v_2w_2$$

Find a symmetric 2×2 matrix A such that $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A\mathbf{y}$ for all $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^2$.

Recap

Today we saw:

- Direct sums
- Reducible linear operators
- Inner product spaces

Next time: Norms and distance, orthogonal sets of vectors (again!)

MATH254: Linear Algebra
Lecture 30

Moira MacNeil

March 26, 2025

Last Time

1. Direct sums
2. Reducible linear operators
3. Chapter 10: Inner product spaces

Today

1. More inner product spaces
2. Norms and distance
3. Orthogonal sets

Reminders:

- Assignment 5 is due on Friday March 28

Recall: Inner Product Spaces

Inner Product Spaces

An **inner product** on a real vector space V is a function that assigns a real number $\langle \mathbf{v}, \mathbf{w} \rangle$ to every pair $\mathbf{v}$, $\mathbf{w}$ of vectors in V in such a way that the following axioms are satisfied.

P1. $\langle \mathbf{v}, \mathbf{w} \rangle$ *is a real number for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V.

P2. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V.

P3. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$ *for all* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *in* V.

P4. $\langle r\mathbf{v}, \mathbf{w} \rangle = r\langle \mathbf{v}, \mathbf{w} \rangle$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V *and all* r *in* $\mathbb{R}$.

P5. $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ *for all* $\mathbf{v} \neq \mathbf{0}$ *in* V.

A real vector space V with an inner product $\langle \ , \rangle$ will be called an **inner product space**.

Theorem 10.1.4

Theorem

Let $\langle\ ,\ \rangle$ be an inner product on a space V; let $\mathbf{v}$, $\mathbf{u}$, and $\mathbf{w}$ denote vectors in V; and let r denote a real number.

1. $\langle \mathbf{u}, \mathbf{v}+\mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
2. $\langle \mathbf{v}, r\mathbf{w} \rangle = r\langle \mathbf{v}, \mathbf{w} \rangle = \langle r\mathbf{v}, \mathbf{w} \rangle$
3. $\langle \mathbf{v}, \mathbf{0} \rangle = 0 = \langle \mathbf{0}, \mathbf{v} \rangle$
4. $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$

$\langle 0, v \rangle = \langle 0+0, v \rangle = \langle 0, v \rangle + \langle 0, v \rangle$

$\Rightarrow \langle 0, v \rangle = 0$

Theorem 10.1.5

Theorem

If A is any $n \times n$ positive definite matrix, then

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A\mathbf{y} \text{ for all columns } \mathbf{x}, \mathbf{y} \text{ in } \mathbb{R}^n$$

defines an inner product on $\mathbb{R}^n$, and every inner product on $\mathbb{R}^n$ arises in this way.

Every inner product on $\mathbb{R}^n$ corresponds to a positive definite $n \times n$ matrix. In particular, the dot product corresponds to the identity matrix I_n.

Remark

If we refer to the inner product space $\mathbb{R}^n$ without specifying the inner product, we mean that the dot product is to be used.

Example 10.1.6

Let the inner product $\langle \ , \rangle$ be defined on $\mathbb{R}^2$ by

$$\left\langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\rangle = 2v_1w_1 - v_1w_2 - v_2w_1 + v_2w_2 \in \mathbb{R}$$

Find a symmetric 2×2 matrix A such that $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$ for all $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^2$.

entry (i,j) corresponds to coeff of $v_i w_j$

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

Take $x = \begin{bmatrix} x \\ y \end{bmatrix}$

$$\langle x, x \rangle = 2x^2 - xy - yx + y^2$$
$$= x^2 + (x-y)^2 \geq 0$$

$\forall x \in \mathbb{R}^2$ $\langle x, x \rangle = 0 \Rightarrow x = 0$

so $\langle , \rangle$ is an inner product so
A is positive definite.

A PD $\Rightarrow x^T A x > 0 \quad \forall \; x \neq 0 \in \mathbb{R}^n$

$\langle x, x \rangle = x^T A x \rightarrow$ this is a quadratic form!

Norm and Distance

Norm and Distance

As in $\mathbb{R}^n$, if $\langle\ ,\rangle$ is an inner product on a space V, the **norm** $\|\mathbf{v}\|$ of a vector $\mathbf{v}$ in V is defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

We define the **distance** between vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space V to be

$$\mathrm{d}(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|$$

Example 10.1.8

Show that $\langle \mathbf{u}+\mathbf{v}, \mathbf{u}-\mathbf{v}\rangle = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2$ in any inner product space.

$$\langle u+v, u-v\rangle = \langle u,u\rangle - \langle u,v\rangle + \langle v,u\rangle - \langle v,v\rangle$$

$$= \|u\|^2 - \langle u,v\rangle + \langle v,u\rangle - \|v\|^2$$

$$= \|u\|^2 - \cancel{\langle u,v\rangle} + \cancel{\langle u,v\rangle} - \|v\|^2$$

By P2, any inner product has $\langle v,w\rangle = \langle w,v\rangle$

Unit Vectors

- A vector $\mathbf{v}$ in an inner product space V is called a **unit vector** if $\|\mathbf{v}\| = 1$
- The set of all unit vectors in V is called the **unit ball** in V
- For example, if $V = \mathbb{R}^2$ (with the dot product) and $\mathbf{v} = (x, y)$, then

$$\|\mathbf{v}\|^2 = 1 \quad \text{if and only if} \quad x^2 + y^2 = 1$$

- Hence the unit ball in $\mathbb{R}^2$ is the **unit circle** $x^2 + y^2 = 1$ with centre at the origin and radius 1
- The shape of the unit ball varies with the choice of inner product

Example

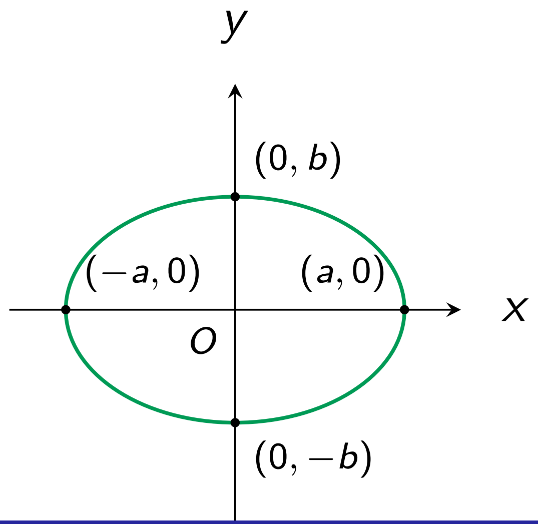
Let $a > 0$ and $b > 0$. If $\mathbf{v} = (x, y)$ and $\mathbf{w} = (x_1, y_1)$, define an inner product on $\mathbb{R}^2$ by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \frac{xx_1}{a^2} + \frac{yy_1}{b^2}$$

In this case

$$\|\mathbf{v}\|^2 = 1 \quad \text{if and only if} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

so the unit ball is the ellipse:



Theorem 10.1.10

Theorem

If $\mathbf{v} \neq \mathbf{0}$ is any vector in an inner product space V, then $\frac{1}{\|\mathbf{v}\|}\mathbf{v}$ is the unique unit vector that is a positive multiple of $\mathbf{v}$.

Theorem 10.1.11

Cauchy-Schwarz Inequality

If $\mathbf{v}$ and $\mathbf{w}$ are two vectors in an inner product space V, then

$$\langle \mathbf{v}, \mathbf{w} \rangle^2 \leq \|\mathbf{v}\|^2 \|\mathbf{w}\|^2$$

Moreover, equality occurs if and only if one of $\mathbf{v}$ and $\mathbf{w}$ is a scalar multiple of the other.

Theorem 10.1.12

Theorem

If V is an inner product space, the norm $\|\cdot\|$ has the following properties.

1. $\|\mathbf{v}\| \geq 0$ for every vector $\mathbf{v}$ in V.
2. $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.
3. $\|r\mathbf{v}\| = |r|\|\mathbf{v}\|$ for every $\mathbf{v}$ in V and every r in $\mathbb{R}$.
4. $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ for all $\mathbf{v}$ and $\mathbf{w}$ in V (**triangle inequality**).

PROOF OF (4)

By Cauchy-Schwarz $\langle v, w \rangle \leq \|v\|^2 \|w\|^2$

$$
\begin{aligned}
\|v+w\|^2 &= \langle v+w, v+w \rangle \\
&= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \\
&\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\
&= (\|v\| + \|w\|)^2
\end{aligned}
$$

take positive square root!

Example 10.1.13

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a spanning set for an inner product space V. If $\mathbf{v}$ in V satisfies $\langle \mathbf{v}, \mathbf{v}_i \rangle = 0$ for each $i = 1, 2, \ldots, n$, show that $\mathbf{v} = \mathbf{0}$.

$v = r_1 v_1 + r_2 v_2 + \cdots + r_n v_n$ for $r_i \in \mathbb{R}$

To show $v = 0$ use $\langle v, v \rangle = 0 \Leftrightarrow v = 0$

$$\langle v, v \rangle = \langle v, r_1 v_1 + \cdots + r_n v_n \rangle = \langle v, r_1 v_1 \rangle + \cdots + \langle v, r_n v_n \rangle$$
$$= r_1 \langle v, v_1 \rangle + \cdots + r_n \langle v, v_n \rangle$$
$$= r_1 0 + \cdots + r_n 0 = 0 \Rightarrow v = 0$$

Theorem 10.1.14

$d(v,w) = \|v-w\|$

Theorem

Let V be an inner product space.

1. $\mathrm{d}(\mathbf{v},\mathbf{w}) \geq 0$ for all $\mathbf{v}$, $\mathbf{w}$ in V.
2. $\mathrm{d}(\mathbf{v},\mathbf{w}) = 0$ if and only if $\mathbf{v} = \mathbf{w}$.
3. $\mathrm{d}(\mathbf{v},\mathbf{w}) = \mathrm{d}(\mathbf{w},\mathbf{v})$ for all $\mathbf{v}$ and $\mathbf{w}$ in V.
4. $\mathrm{d}(\mathbf{v},\mathbf{w}) \leq \mathrm{d}(\mathbf{v},\mathbf{u}) + \mathrm{d}(\mathbf{u},\mathbf{w})$ for all $\mathbf{v}$, $\mathbf{u}$, and $\mathbf{w}$ in V.

Orthogonal Sets of Vectors

- Two nonzero geometric vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ are orthogonal if and only if $\mathbf{x} \cdot \mathbf{y} = 0$
- In general, two vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space V are said to be **orthogonal** if

$$\boxed{\langle \mathbf{v}, \mathbf{w} \rangle = 0}$$

- A set $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ of vectors is called an **orthogonal set of vectors** if
  1. *Each* $\mathbf{f}_i \neq \mathbf{0}$
  2. $\langle \mathbf{f}_i, \mathbf{f}_j \rangle = 0$ *for all* $i \neq j$
- If we also have $\|\mathbf{f}_i\| = 1$ for each i, the set $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ is called an **orthonormal** set

Theorem 10.2.1

$f_i \in V$ $\quad V$ is inner product space

Pythagoras' Theorem

If $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ is an orthogonal set of vectors, then

$$\|\mathbf{f}_1 + \mathbf{f}_2 + \cdots + \mathbf{f}_n\|^2 = \|\mathbf{f}_1\|^2 + \|\mathbf{f}_2\|^2 + \cdots + \|\mathbf{f}_n\|^2$$

Theorem 10.2.2

Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ be an orthogonal set of vectors.

1. $\{r_1\mathbf{f}_1, r_2\mathbf{f}_2, \ldots, r_n\mathbf{f}_n\}$ is also orthogonal for any $r_i \neq 0$ in $\mathbb{R}$.
2. $\left\{\frac{1}{\|\mathbf{f}_1\|}\mathbf{f}_1, \frac{1}{\|\mathbf{f}_2\|}\mathbf{f}_2, \ldots, \frac{1}{\|\mathbf{f}_n\|}\mathbf{f}_n\right\}$ is an orthonormal set.

The process of passing from an orthogonal set to an orthonormal one is called **normalizing** the orthogonal set.

Theorem 10.2.3

Theorem

Every orthogonal set of vectors is linearly independent.

Example 10.2.4

Show that $\left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \right\}$ is an orthogonal basis of $\mathbb{R}^3$ with inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T A \mathbf{w}, \text{ where } A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\left\langle \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\rangle = \begin{bmatrix} 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 0$$

Verify that the other pairs also give $\langle , \rangle = 0$

THM: EXPANSION THEOREM

Let $\{f_1, \ldots, f_n\}$ be an orthogonal basis of an inner product space V.

If v is any vector in V, then

$$v = \underbrace{\frac{\langle v, f_1 \rangle}{\|f_1\|^2}} f_1 + \cdots + \underbrace{\frac{\langle v, f_n \rangle}{\|f_n\|^2}} f_n$$

is the expansion of v as a linear combination of the basis vectors.

The coefficients in expansion thm are called Fourier coefficients of v with respect to orthogonal basis $\{f_1, \ldots, f_n\}$.

Example:

If $a_0, \cdots, a_n$ are distinct numbers and $p(x), q(x) \in P_n$ define

$$\langle p(x), q(x)\rangle = p(a_0)q(a_0) + \cdots + p(a_n)q(a_n)$$

This is an inner product on P_n:

P1: $\langle p(x), q(x)\rangle \in \mathbb{R}$ for any $p, q \in P_n$

P2: $\langle p(x), q(x)\rangle =$

$$\begin{aligned} &p(a_0)q(a_0) + \cdots + p(a_n)q(a_n) \\ &= q(a_0)p(a_0) + \cdots + q(a_n)p(a_n) \\ &= \langle q(x), p(x)\rangle \end{aligned}$$

P3: $r(x) \in P_n$

$$\begin{aligned} &\langle p(x) + r(x), q(x)\rangle \\ &= (p(a_0) + r(a_0))q(a_0) + \cdots \\ &= p(a_0)q(a_0) + r(a_0)q(a_0) + \cdots \\ &= \langle p(x), q(x)\rangle + \langle r(x), q(x)\rangle \end{aligned}$$

P4: $\langle k\,p(x), q(x)\rangle = k\langle p(x), q(x)\rangle$

similar

Recap

Today we saw:

- Inner product spaces
- Norms and distance
- Orthogonal sets of vectors

Next time: More on orthogonal sets of vectors

MATH254: Linear Algebra
Lecture 31

Moira MacNeil

March 28, 2025

Last Time

1. Norms and distance
2. Orthogonal sets

Today

1. Orthogonal sets
2. Orthogonal projections
3. Orthogonal diagonalization

Posted Ch 10 practice Q
Exam review coming today!

Reminders:

- Final Exam is Monday April 7
- Course evaluations are available

Theorem 10.2.5

Expansion Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ be an orthogonal basis of an inner product space V. If $\mathbf{v}$ is any vector in V, then

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_n \rangle}{\|\mathbf{f}_n\|^2}\mathbf{f}_n$$

is the expansion of $\mathbf{v}$ as a linear combination of the basis vectors.

The coefficients $\frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}, \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}, \ldots, \frac{\langle \mathbf{v}, \mathbf{f}_n \rangle}{\|\mathbf{f}_n\|^2}$ in the expansion theorem are sometimes called the **Fourier coefficients** of $\mathbf{v}$ with respect to the orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$.

Example 10.2.6

If $a_0, a_1, \ldots, a_n$ are distinct numbers and $p(x)$ and $q(x)$ are in $\mathbf{P}_n$, define

$$\langle p(x), q(x)\rangle = p(a_0)q(a_0) + p(a_1)q(a_1) + \cdots + p(a_n)q(a_n)$$

This is an inner product on $\mathbf{P}_n$.
Last time we showed axioms P1-P4 hold. We show *P5.* $\langle \mathbf{v}, \mathbf{v}\rangle > 0$ *for all* $\mathbf{v} \neq \mathbf{0}$ *in* V.

0 is the only polynomial of deg. n
with n+1 distinct roots

Example 10.2.6

The **Lagrange polynomials** $\delta_0(x), \delta_1(x), \ldots, \delta_n(x)$ relative to the numbers $a_0, a_1, \ldots, a_n$ are defined as follows:

$$\delta_k(x) = \frac{\prod_{i\neq k}(x - a_i)}{\prod_{i\neq k}(a_k - a_i)} \quad k = 0, 1, 2, \ldots, n$$

where $\prod_{i\neq k}(x - a_i)$ means the product of all the terms

$$(x - a_0), (x - a_1), (x - a_2), \ldots, (x - a_n) \quad \text{except that the } k\text{th term is omitted.}$$

Then $\{\delta_0(x), \delta_1(x), \ldots, \delta_n(x)\}$ is orthonormal with respect to $\langle\ ,\rangle$ because $\delta_k(a_i) = 0$ if $i \neq k$ and $\delta_k(a_k) = 1$. These facts also show that $\langle p(x), \delta_k(x)\rangle = p(a_k)$ so the expansion theorem gives

$$p(x) = p(a_0)\delta_0(x) + p(a_1)\delta_1(x) + \cdots + p(a_n)\delta_n(x)$$

for each $p(x)$ in $\mathbf{P}_n$. This is the **Lagrange interpolation expansion** of $p(x)$, which is important in numerical integration.

Lemma 10.2.7

Orthogonal Lemma

Let $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ be an orthogonal set of vectors in an inner product space V, and let $\mathbf{v}$ be any vector not in $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$. Define

$$\mathbf{f}_{m+1} = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\langle \mathbf{v}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Then $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is an orthogonal set of vectors.

Gram-Schmidt Orthogonalization Algorithm

Let V be an inner product space and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any basis of V. Define vectors $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ in V successively as follows:

$$
\begin{aligned}
\mathbf{f}_1 &= \mathbf{v}_1 \\
\mathbf{f}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}_3, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
\vdots \quad & \quad \vdots \\
\mathbf{f}_k &= \mathbf{v}_k - \frac{\langle \mathbf{v}_k, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}_k, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\langle \mathbf{v}_k, \mathbf{f}_{k-1} \rangle}{\|\mathbf{f}_{k-1}\|^2}\mathbf{f}_{k-1}
\end{aligned}
$$

for each $k = 2, 3, \ldots, n$. Then

1. $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ is an orthogonal basis of V.
2. $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\} = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ holds for each $k = 1, 2, \ldots, n$.

Example 10.2.9

Consider $V = \mathbf{P}_3$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)dx$. If the Gram-Schmidt algorithm is applied to the basis $\{1, x, x^2, x^3\}$, show that the result is the orthogonal basis

$$\{1, x, \frac{1}{3}(3x^2 - 1), \frac{1}{5}(5x^3 - 3x)\}$$

$$f_1 = 1$$

$$f_2 = x - \frac{\langle x, f_1 \rangle}{\|f_1\|^2} f_1 = x - \frac{0}{2}(1) = x$$

$$f_3 = x^2 - \frac{\langle x^2, f_1 \rangle}{\|f_1\|^2} f_1 - \frac{\langle x^2, f_2 \rangle}{\|f_2\|^2} f_2 = x^2 - \frac{2/3}{2}(1) - \frac{0}{2/3} x$$

$= x^2 - 1/3$

f_4 is similar

Corollary 10.2.10

Corollary

If V is any n-dimensional inner product space, then V is isomorphic to $\mathbb{R}^n$ as inner product spaces. More precisely, if E is any orthonormal basis of V, the coordinate isomorphism

$$C_E : V \to \mathbb{R}^n \text{ satisfies } \langle \mathbf{v}, \mathbf{w} \rangle = C_E(\mathbf{v}) \cdot C_E(\mathbf{w})$$

for all $\mathbf{v}$ and $\mathbf{w}$ in V.

Theorem 10.2.11

Theorem

Let U be a finite dimensional subspace of an inner product space V.

1. $U^\perp$ is a subspace of V and $V = U \oplus U^\perp$.
2. If $\dim V = n$, then $\dim U + \dim U^\perp = n$.
3. If $\dim V = n$, then $U^{\perp\perp} = U$.

Let U be a subspace of an inner product space V. As in $\mathbb{R}^n$, the **orthogonal complement** $U^\perp$ of U in V is defined by

$$U^\perp = \{\mathbf{v} \mid \mathbf{v} \in V, \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \in U\}$$

Orthogonal Projection

Orthogonal Projection on a Subspace

The projection on U with kernel $U^{\perp}$ is called the **orthogonal projection** on U (or simply the **projection** on U) and is denoted $\underline{\mathrm{proj}_U} : V \to V$.

Theorem 10.2.13

Projection Theorem

Let U be a finite dimensional subspace of an inner product space V and let $\mathbf{v}$ be a vector in V.

1. $\text{proj}_U : V \to V$ is a linear operator with image U and kernel $U^\perp$.
2. $\text{proj}_U\mathbf{v}$ is in U and $\mathbf{v} - \text{proj}_U\mathbf{v}$ is in $U^\perp$.
3. If $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is any orthogonal basis of U, then

$$\text{proj}_U\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Note that there is no requirement that V is finite dimensional.

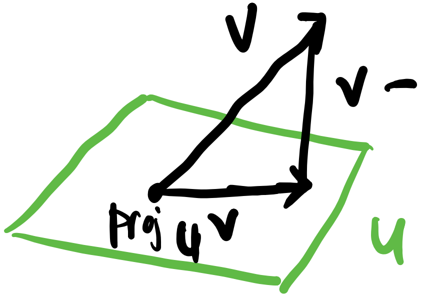
Example 10.2.14

Let U be a subspace of the finite dimensional inner product space V. Show that $\mathrm{proj}_{U^\perp}\mathbf{v} = \mathbf{v} - \mathrm{proj}_U\mathbf{v}$ for all $\mathbf{v} \in V$.

We have $V = U^\perp \oplus U^{\perp\perp}$

let $p = \mathrm{proj}_U v$ then $v = \underbrace{(v-p)}_{\in U^\perp} + \underbrace{p}_{\in U^{\perp\perp}}$

$\Rightarrow \mathrm{proj}_{U^\perp} v = v - p$



$v - \mathrm{proj}_U v \in U^\perp$

Theorem 10.2.15

Approximation Theorem

Let U be a finite dimensional subspace of an inner product space V. If $\mathbf{v}$ is any vector in V, then $\text{proj}_U\mathbf{v}$ is the vector in U that is closest to $\mathbf{v}$. Here **closest** means that

$$\|\mathbf{v} - \text{proj}_U\mathbf{v}\| < \|\mathbf{v} - \mathbf{u}\|$$

for all $\mathbf{u}$ in U, $\mathbf{u} \neq \text{proj}_U\mathbf{v}$.

Theorem 10.3.1

Theorem

Let $T : V \to V$ be a linear operator on a finite dimensional space V. Then the following conditions are equivalent.

1. V has a basis consisting of eigenvectors of T.
2. There exists a basis B of V such that $M_B(T)$ is diagonal.

Diagonalizable Linear Operators

Diagonalizable Linear Operators

A linear operator T on a finite dimensional space V is called **diagonalizable** if V has a basis consisting of eigenvectors of T.

Example 10.3.3

Let $T : \mathbf{P}_2 \to \mathbf{P}_2$ be given by

$$T(a + bx + cx^2) = (a + 4c) - 2bx + (3a + 2c)x^2$$

Find the eigenspaces of T and hence find a basis of eigenvectors.

$B_0 = \{1, x, x^2\}$, then find

$$M_{B_0}(T) = \left[C_{B_0}(T(1)) \quad C_{B_0}(T(x)) \quad C_{B_0}(T(x^2)) \right]$$

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

Find $c_T(x) = (x+2)^2(x-5)$

$\lambda_1 = -2 \quad \lambda_2 = 5$

Eigenvals of T

Eigenvectors for $\lambda_1 = -2$

$(-2I - A)x = 0$

$$\begin{bmatrix} -3 & 0 & -4 \\ 0 & 0 & 0 \\ -3 & 0 & -4 \end{bmatrix} \to \begin{bmatrix} -3 & 0 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$x_2 = s$

$x_3 = t$

$x_1 = -\frac{4}{3}t$

$$E_{-2}(x) = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 3 \end{bmatrix} \right\}$$

$$E_5(x) = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of eigenvect. of $M_{B_0}(T)$

$B = \{x, 3x^2 - 4, x^2 + 1\}$ is a basis of P_2 consisting of eigenvectors of T

Theorem 10.3.4

Theorem

Let $T : V \to V$ be a linear operator on an inner product space V. If $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is an orthogonal basis of V, then

$$M_B(T) = \left[\frac{\langle \mathbf{b}_i, T(\mathbf{b}_j) \rangle}{\|\mathbf{b}_i\|^2} \right]$$

Example 10.3.5

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$T(a, b, c) = (a + 2b - c, 2a + 3c, -a + 3b + 2c)$$

If the dot product in $\mathbb{R}^3$ is used, find the matrix of T with respect to the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ where $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, $\mathbf{e}_3 = (0, 0, 1)$.

B is orthonormal $\|b_i\|^2 = 1$

$$M_B(T) = \left[\frac{\langle b_i, T(b_j) \rangle}{\|b_i\|^2} \right] = \begin{bmatrix} e_1 \cdot T(e_1) & e_1 \cdot T(e_2) & e_1 \cdot T(e_3) \\ e_2 \cdot T(e_1) & e_2 \cdot T(e_2) & e_2 \cdot T(e_3) \\ e_3 \cdot T(e_1) & e_3 \cdot T(e_2) & e_3 \cdot T(e_3) \end{bmatrix}$$

$T(e_1) = (1, 2, -1)$
$T(e_2) = (2, 0, 3)$
$T(e_3) = (-1, 3, 2)$

$$M_B(T) = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 3 \\ -1 & 3 & 2 \end{bmatrix}$$

$n \times n$ matrix A symmetric

$\iff \quad x \cdot (Ay) = (Ax) \cdot y \quad \forall\, x, y \in \mathbb{R}^n$

Theorem 10.3.6

Theorem

Let V be a finite dimensional inner product space. The following conditions are equivalent for a linear operator $T : V \to V$.

1. $\langle \mathbf{v}, T(\mathbf{w}) \rangle = \langle T(\mathbf{v}), \mathbf{w} \rangle$ for all $\mathbf{v}$ and $\mathbf{w}$ in V.
2. The matrix of T is symmetric with respect to every orthonormal basis of V.
3. The matrix of T is symmetric with respect to some orthonormal basis of V.
4. There is an orthonormal basis $B = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ of V such that $\langle \mathbf{f}_i, T(\mathbf{f}_j) \rangle = \langle T(\mathbf{f}_i), \mathbf{f}_j \rangle$ holds for all i and j.

A linear operator T on an inner product space V is called **symmetric** if $\langle \mathbf{v}, T(\mathbf{w}) \rangle = \langle T(\mathbf{v}), \mathbf{w} \rangle$ holds for all $\mathbf{v}$ and $\mathbf{w}$ in V.

Example 10.3.7

If A is an $n \times n$ matrix, let $T_A : \mathbb{R}^n \to \mathbb{R}^n$ be the matrix operator given by $T_A(\mathbf{v}) = A\mathbf{v}$ for all columns $\mathbf{v}$. If the dot product is used in $\mathbb{R}^n$, then T_A is a symmetric operator if and only if A is a symmetric matrix.

If E is standard basis of $\mathbb{R}^n$, then E is orthonormal with dot product, and $M_E(T) = A$

Apply (3) of previous thm (A is symmetric)

Theorem 10.3.8

Theorem

A symmetric linear operator on a finite dimensional inner product space has real eigenvalues.

Theorem 10.3.9

Theorem

Let $T : V \to V$ be a symmetric linear operator on an inner product space V, and let U be a T-invariant subspace of V. Then:

1. The restriction of T to U is a symmetric linear operator on U.
2. $U^\perp = \{\mathbf{v} \text{ in } V \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \text{ in } U\}$ is also T-invariant.

Theorem 10.3.10

Principal Axes Theorem

The following conditions are equivalent for a linear operator T on a finite dimensional inner product space V.

1. T is symmetric.
2. V has an orthogonal basis consisting of eigenvectors of T.

Example 10.3.11

Let $T : \mathbf{P}_2 \to \mathbf{P}_2$ be given by

$$T(a + bx + cx^2) = (8a - 2b + 2c) + (-2a + 5b + 4c)x + (2a + 4b + 5c)x^2$$

Using the inner product $\langle a + bx + cx^2, a' + b'x + c'x^2 \rangle = aa' + bb' + cc'$, show that T is symmetric and find an orthonormal basis of $\mathbf{P}_2$ consisting of eigenvectors.

Recap

Today we saw:

- Orthogonal sets of vectors
- Diagonalizable linear operators

Next time: Isometries

MATH254: Linear Algebra
Lecture 32

Moira MacNeil

April 1, 2025

Last Time

1. Orthogonal sets of vectors

2. Orthogonal projections

3. Diagonalizable linear operators

Today

1. More on diagonalizable linear operators END OF MATERIAL FOR EXAM
2. Isometries

Reminders:

- Final Exam is next Monday April 7
- Exam info and practice questions for Chapter 10 are posted
- Course evaluations are available

Recall: Diagonalizable Linear Operators

Diagonalizable Linear Operators

A linear operator T on a finite dimensional space V is called **diagonalizable** if V has a basis consisting of eigenvectors of T.

Recall: Theorem 10.3.6

Theorem

Let V be a finite dimensional inner product space. The following conditions are equivalent for a linear operator $T : V \to V$.

1. $\langle \mathbf{v}, T(\mathbf{w}) \rangle = \langle T(\mathbf{v}), \mathbf{w} \rangle$ for all $\mathbf{v}$ and $\mathbf{w}$ in V.
2. The matrix of T is symmetric with respect to every orthonormal basis of V.
3. The matrix of T is symmetric with respect to some orthonormal basis of V.
4. There is an orthonormal basis $B = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ of V such that $\langle \mathbf{f}_i, T(\mathbf{f}_j) \rangle = \langle T(\mathbf{f}_i), \mathbf{f}_j \rangle$ holds for all i and j.

A linear operator T on an inner product space V is called **symmetric** if $\langle \mathbf{v}, T(\mathbf{w}) \rangle = \langle T(\mathbf{v}), \mathbf{w} \rangle$ holds for all $\mathbf{v}$ and $\mathbf{w}$ in V.

Theorem 10.3.8

Theorem

A symmetric linear operator on a finite dimensional inner product space has real eigenvalues.

Theorem 10.3.9

Theorem

Let $T : V \to V$ be a symmetric linear operator on an inner product space V, and let U be a T-invariant subspace of V. Then:

1. The restriction of T to U is a symmetric linear operator on U.
2. $U^{\perp} = \{\mathbf{v} \text{ in } V \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \text{ in } U\}$ is also T-invariant.

Theorem 10.3.10

Principal Axes Theorem

The following conditions are equivalent for a linear operator T on a finite dimensional inner product space V.

1. T is symmetric.
2. V has an orthogonal basis consisting of eigenvectors of T.

Example 10.3.11

Let $T : \mathbf{P}_2 \to \mathbf{P}_2$ be given by

$$T(a+bx+cx^2) = (8a-2b+2c) + (-2a+5b+4c)x + (2a+4b+5c)x^2$$

Using the inner product $\langle a+bx+cx^2, a'+b'x+c'x^2 \rangle = aa'+bb'+cc'$, show that T is symmetric and find an orthonormal basis of $\mathbf{P}_2$ consisting of eigenvectors.

$B_0 = \{1, x, x^2\}$

$T(1) = 8 - 2x + 2x^2$

$T(x) = -2 + 5x + 4x^2$

$T(x^2) = 2 + 4x + 5x^2$

$$M_{B_0}(T) = \begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$$

$\hookrightarrow$ symmetric

$\Rightarrow T$ is symmetric

Find orthogonal eigenvectors of $M_{B_0}(T)$

$$C_T(x) = \det(xI - M_{B_0}(T))$$

$$= \begin{vmatrix} x-8 & 2 & -2 \\ 2 & x-5 & -4 \\ -2 & -4 & x-5 \end{vmatrix} = x(x-9)^2$$

$\lambda_1 = 0 \qquad \lambda_2^2 = 9$

$E_0(T) = \text{span}\{x_1\} \qquad x_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$

$\lambda_2 = 9$

$(9I - M_{B_0}(T))x = 0$

$$\begin{bmatrix} 1 & 2 & -2 \\ 2 & 4 & -4 \\ -2 & -4 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$x_3 = t$
$x_2 = s$
$x_1 = -2s + 2t$

$$E_9(T) = \text{span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

both orthog. to x_1

Use G.S. to orthogonalize E_q

$$f_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

$$f_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \frac{(-4)}{5} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -8/5 \\ 4/5 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 2/5 \\ 4/5 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

Normalize to get orthonormal basis:
(of eigenvect)

$$\left\{ 1/3 \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{45}} \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} \right\}$$

Then the corresponding basis of P_2 (note B_0 is orthonormal)

$$B = \left\{ 1/3\,(1 + 2x - 2x^2),\ 1/\sqrt{5}\,(-2 + x),\ 1/\sqrt{45}\,(2 + 4x + 5x^2) \right\}$$

EXAM MATERIAL ENDS

Distance Preserving Transformations

- If V is an inner product space, a transformation $S : V \to V$ (not necessarily linear) is said to be **distance preserving** if the distance between $S(\mathbf{v})$ and $S(\mathbf{w})$ is the same as the distance between $\mathbf{v}$ and $\mathbf{w}$ for all vectors $\mathbf{v}$ and $\mathbf{w}$
- Formally, if
$$\|S(\mathbf{v}) - S(\mathbf{w})\| = \|\mathbf{v} - \mathbf{w}\| \quad \text{for all } \mathbf{v} \text{ and } \mathbf{w} \text{ in } V$$
- Distance-preserving maps need not be linear!
- For example, if $\mathbf{u}$ is any vector in V, the transformation $S_{\mathbf{u}} : V \to V$ defined by $S_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} + \mathbf{u}$ for all $\mathbf{v}$ in V is called **translation** by $\mathbf{u}$, and it is routine to verify that $S_{\mathbf{u}}$ is distance preserving for any $\mathbf{u}$. However, $S_{\mathbf{u}}$ is linear only if $\mathbf{u} = \mathbf{0}$ (since then $S_{\mathbf{u}}(\mathbf{0}) = \mathbf{0}$).

Lemma 10.4.1

Lemma

Let V be an inner product space of dimension n, and consider a distance-preserving transformation $S : V \to V$. If $S(\mathbf{0}) = \mathbf{0}$, then S is linear.

Isometries

Isometries

Distance-preserving linear operators are called **isometries**.

Theorem 10.4.3

Theorem

If V is a finite dimensional inner product space, then every distance-preserving transformation $S : V \to V$ is the composite of a translation and an isometry.

PROOF if $S: V \to V$ is dist preserving let
$S(0) = u$ and let $T: V \to V$ $T(v) = S(v) - u$
for all $v \in V$

$$\|T(v) - T(w)\| = \|S(v) - u - (S(w) - u)\|$$
$$= \|S(v) - S(w)\|$$
S is dist pres.
$$= \|v - w\|$$
$\Rightarrow$ T is also dist. preserv.

$$T(0) = S(0) - u = u - u = 0$$
$\Rightarrow$ T linear

$\therefore$ T is an isometry.

$$S(v) = u + T(v) = (S_u \circ T)(v)$$
ie., $S = S_u \circ T$

$\square$

Theorem 10.4.4

Theorem

Let $T : V \to V$ be a linear operator on a finite dimensional inner product space V. The following conditions are equivalent:

1. T is an isometry. (T preserves distance)
2. $\|T(\mathbf{v})\| = \|\mathbf{v}\|$ for all $\mathbf{v}$ in V. (T preserves norms)
3. $\langle T(\mathbf{v}), T(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$ for all $\mathbf{v}$ and $\mathbf{w}$ in V. (T preserves inner products)
4. If $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ is an orthonormal basis of V, then $\{T(\mathbf{f}_1), T(\mathbf{f}_2), \dots, T(\mathbf{f}_n)\}$ is also an orthonormal basis. (T preserves orthonormal bases)
5. T carries some orthonormal basis to an orthonormal basis.

Corollary 10.4.5

Corollary

Let V be a finite dimensional inner product space.

1. Every isometry of a finite dimensional V is an isomorphism.
2. a. $1_V : V \to V$ is an isometry.
   b. The composite of two isometries of V is an isometry.
   c. The inverse of an isometry of V is an isometry.

Part (2) asserts that the set of isometries of a finite dimensional inner product space forms an algebraic system called a **group**. More on this in MATH 354: Modern Algebra.

Example 10.4.6

Rotations of $\mathbb{R}^2$ about the origin are isometries, as are reflections in lines through the origin: They clearly preserve distance and so are linear . Similarly, rotations about lines through the origin and reflections in planes through the origin are isometries of $\mathbb{R}^3$.

Example 10.4.7

Let $T : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ be the transposition operator: $T(A) = A^T$. Then T is an isometry if the inner product is $\langle A, B \rangle = \underline{\operatorname{tr}(AB^T)} = \sum_{i,j} a_{ij} b_{ij}$. In fact, T permutes the basis consisting of all matrices with one entry 1 and the other entries 0.

Theorem 10.4.8

Theorem

Let $T : V \to V$ be an operator where V is a finite dimensional inner product space. The following conditions are equivalent.

1. T is an isometry.
2. $M_B(T)$ is an orthogonal matrix for every orthonormal basis B.
3. $M_B(T)$ is an orthogonal matrix for some orthonormal basis B.

It is important that B is *orthonormal*.

For example, $T : V \to V$ given by $T(\mathbf{v}) = 2\mathbf{v}$ preserves *orthogonal* sets but is not an isometry.

Corollary 10.4.9

Corollary

If $T : V \to V$ is an isometry where V is a finite dimensional inner product space, then $\det T = \pm 1$.

If P is an orthogonal square matrix, then $P^{-1} = P^T$. Taking determinants yields $(\det P)^2 = 1$, so $\det P = \pm 1$.

Example 10.4.10

$T_A(x) = Ax$

If A is any $n \times n$ matrix, the matrix operator $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry if and only if A is orthogonal using the dot product in $\mathbb{R}^n$. Indeed, if E is the standard basis of $\mathbb{R}^n$, then $M_E(T_A) = A$.

Theorem 10.4.11

Let $T : V \to V$ be an isometry on the two-dimensional inner product space V. Then there are two possibilities.

Either (1) There is an orthonormal basis B of V such that

$$M_B(T) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad 0 \leq \theta < 2\pi$$

→ rotations

or (2) There is an orthonormal basis B of V such that

$$M_B(T) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

→ reflections

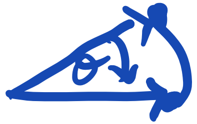
Furthermore, type (1) occurs if and only if $\det T = 1$, and type (2) occurs if and only if $\det T = -1$.

Corollary 10.4.12

Corollary

An operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ is an isometry if and only if T is a rotation or a reflection.

Isometries in $\mathbb{R}^2$



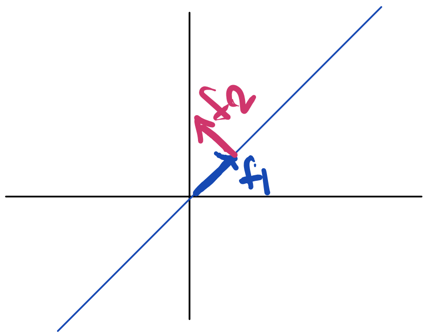
- If E is the standard basis of $\mathbb{R}^2$, then the clockwise rotation R_θ about the origin through an angle θ has matrix

$$M_E(R_\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

- If $S : \mathbb{R}^2 \to \mathbb{R}^2$ is the reflection in a line through the origin (called the **fixed line** of the reflection), let $\mathbf{f}_1$ be a unit vector pointing along the fixed line and let $\mathbf{f}_2$ be a unit vector perpendicular to the fixed line. Then $B = \{\mathbf{f}_1, \mathbf{f}_2\}$ is an orthonormal basis, $S(\mathbf{f}_1) = \mathbf{f}_1$ and $S(\mathbf{f}_2) = -\mathbf{f}_2$, so

$$M_B(S) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

- In this case, 1 is an eigenvalue of S, and any eigenvector corresponding to 1 is a direction vector for the fixed line


f2
f1

Example 10.4.13

In each case, determine whether $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is a rotation or a reflection, and then find the angle or fixed line:

$$\text{(a) } A = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{bmatrix} \qquad \text{(b) } A = \frac{1}{5}\begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}$$

Recap

Today we saw:

- Diagonalizable linear operators
- Isometries

Next time: More isometries