Section 5.3

Best Approximations

Objectives

- What is a Least Squares solution ?
- How to find the best approximate solution to inconsistent system?

Discussion 5.3.1

orthogonality $\xrightarrow{\text{applications}}$ study of approximations

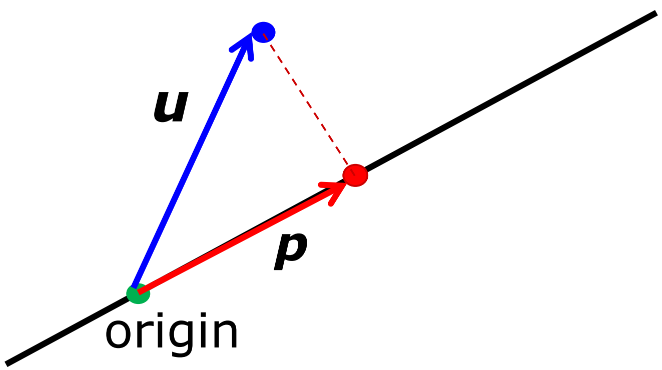
Linear system

consistent — get exact solutions

inconsistent — get approximate solutions

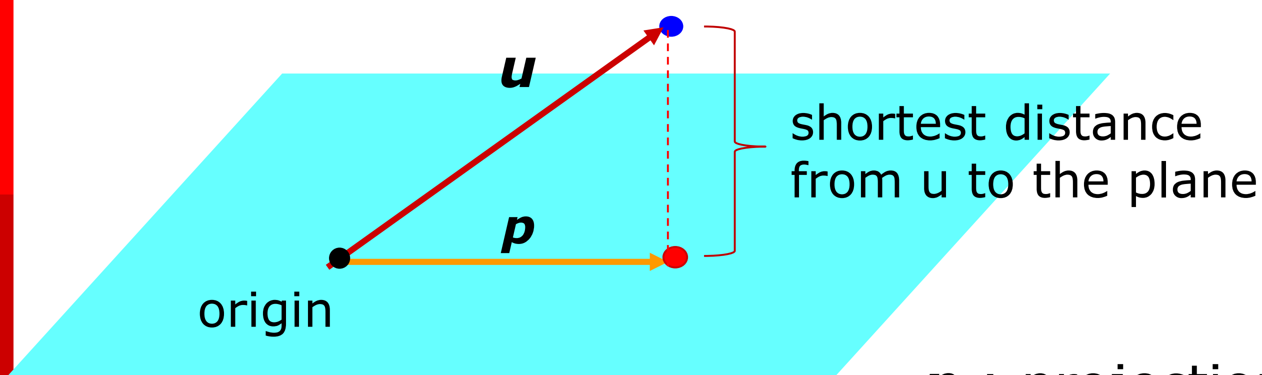
Finding the "best approximation" of a vector from a subspace

Nearest point



p : projection of ***u*** onto the line

We say: ***p*** is the best approximation of ***u*** in the line



Example 5.3.3

p : projection of ***u*** onto the plane

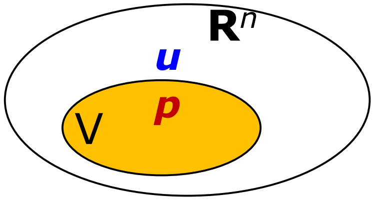
We say: ***p*** is the best approximation of ***u*** in the plane

Finding the "best approximation" of a vector from a subspace

Theorem 5.3.2

V : subspace in $\mathbf{R}^n$ and $\boldsymbol{u} \in \mathbf{R}^n$.

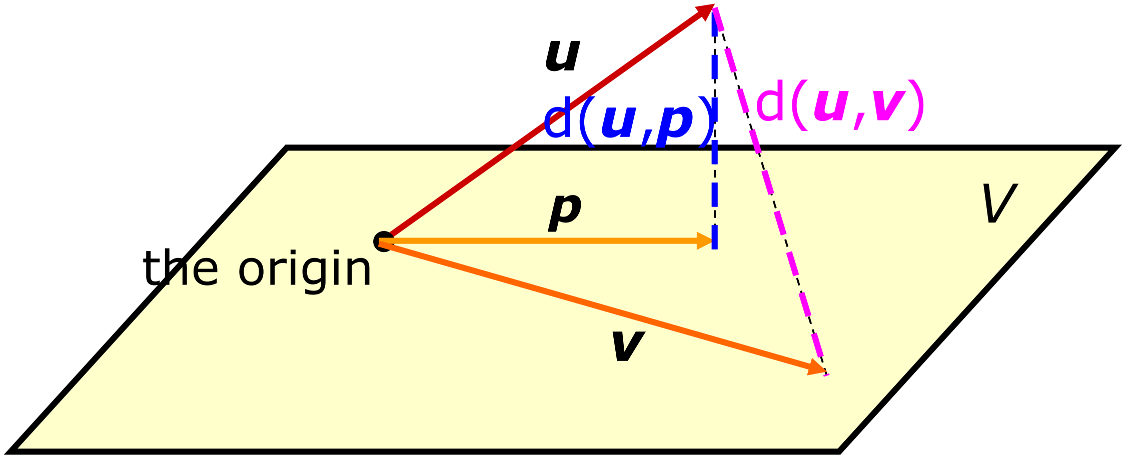
need not be a line or plane



$\boldsymbol{p}$: projection of $\boldsymbol{u}$ onto V

Then $d(\boldsymbol{u}, \boldsymbol{p}) \le d(\boldsymbol{u}, \boldsymbol{v})$ for any vector $\boldsymbol{v}$ in V

i.e. $\boldsymbol{p}$ is the best approximation of $\boldsymbol{u}$ in V.



Theorem 5.3.2

V : subspace in $\mathbf{R}^n$ and $\boldsymbol{u} \in \mathbf{R}^n$.

need not be a line or plane

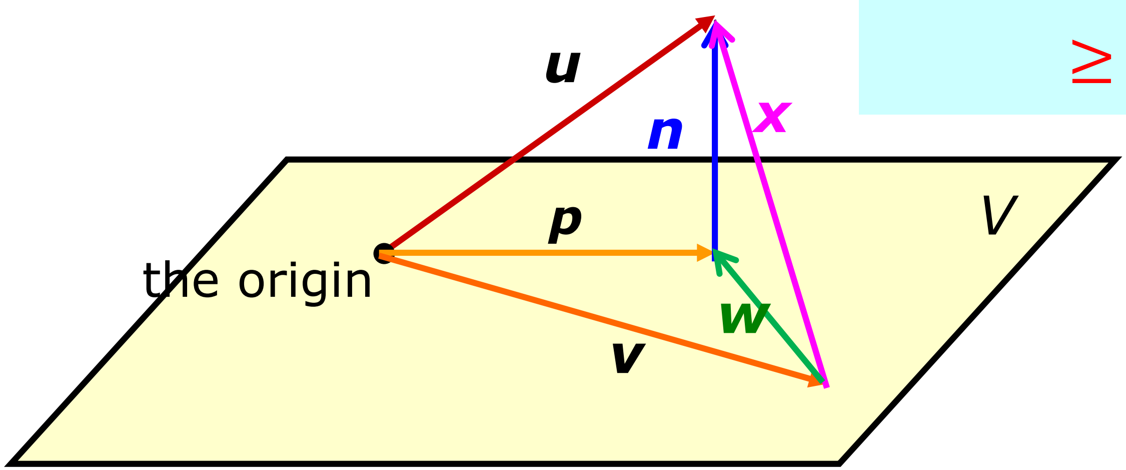
$\boldsymbol{p}$: projection of $\boldsymbol{u}$ onto V

Then $d(\boldsymbol{u}, \boldsymbol{p}) \leq d(\boldsymbol{u}, \boldsymbol{v})$ for any vector $\boldsymbol{v}$ in V

$$||\boldsymbol{n}|| \leq ||\boldsymbol{x}||$$

$$||\boldsymbol{x}||^2 = ||\boldsymbol{n} + \boldsymbol{w}||^2$$
$$= ||\boldsymbol{n}||^2 + ||\boldsymbol{w}||^2$$
$$\geq ||\boldsymbol{n}||^2$$

(see Ex 5 Q9)



Inconsistent system

$t = cr^2 + ds + e$

Example 5.3.5 experimental errors

6 equations
3 unknowns c, d, e

system $\boldsymbol{Ax} = \boldsymbol{b}$

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \\ 4 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} 0.5 \\ 1.6 \\ 2.8 \\ 0.8 \\ 5.1 \\ 5.9 \end{pmatrix}$$

This system is inconsistent $\boldsymbol{Ax} - \boldsymbol{b} \neq \boldsymbol{0}$

Find the best approximate solution see example 5.3.11.2

Find $\boldsymbol{x}_0$ such that $||\boldsymbol{Ax}_0 - \boldsymbol{b}||$ is the smallest

$\sqrt{\text{sum of squares}}$

Such an $\boldsymbol{x}_0$ is called
a least squares solution to the system $\boldsymbol{Ax} = \boldsymbol{b}$.

What is a least squares solution?

Definition 5.3.6

A least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$ ($\boldsymbol{A}$: m×n)

is a vector $\boldsymbol{u}$ in $\mathbf{R}^n$ that minimize $||\boldsymbol{b} - \boldsymbol{Ax}||$

i.e. $||\boldsymbol{b} - \boldsymbol{Au}|| \leq ||\boldsymbol{b} - \boldsymbol{Av}||$ for all $\boldsymbol{v}$ in $\mathbf{R}^n$

good for intuition, but not finding this approximation.

working definition

new linear system

is an actual solution of $\boxed{\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}}$

Theorem 5.3.10

Exercise 5 Q24

$$\begin{cases} x + y + z = 1 \\ y + z = 1 \\ x - y - z = 1 \\ z = 1 \end{cases}$$

$\boldsymbol{Ax} = \boldsymbol{b}$ inconsistent

does not satisfy

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{A}^T = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 1 \end{pmatrix}$$

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 3 \\ 0 & 3 & 4 \end{pmatrix} \quad \mathbf{A}^T\mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

$\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$ consistent

satisfies

$$\mathbf{u} = \begin{pmatrix} 1 \\ -\frac{2}{3} \\ 1 \end{pmatrix}$$

Theorem 5.3.10

$\mathbf{u}$ gives the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

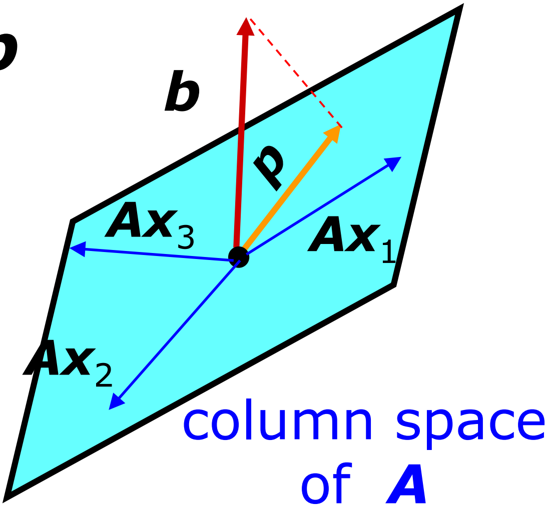
Projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$

Discussion 5.3.7

Find least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

Find $\boldsymbol{u}$ that minimize $||\boldsymbol{b} - \boldsymbol{Ax}||$

the projection $\boldsymbol{p}$ of $\boldsymbol{b}$
onto the column space of $\boldsymbol{A}$



Find $\boldsymbol{u}$ such that $\boldsymbol{Au} = \boldsymbol{p}$

This system is always consistent

$$\mathbf{A} = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3) \qquad \boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\text{: columns of } \boldsymbol{A} \qquad \mathbf{x} = \begin{pmatrix} c \\ d \\ e \end{pmatrix}$$

$\Rightarrow \boldsymbol{Ax} = c\boldsymbol{u_1} + d\boldsymbol{u_2} + e\boldsymbol{u_3}$ linear comb of columns of $\boldsymbol{A}$

All $\boldsymbol{Ax}$ belong to column space of $\boldsymbol{A}$ Discussion 4.1.16

Theorem 5.3.8

$\boldsymbol{u}$ is a least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

$\Leftrightarrow$ $\boldsymbol{u}$ is a solution of $\boldsymbol{Ax} = \boldsymbol{p}$

$\Leftrightarrow$ $\boldsymbol{Au} = \boldsymbol{p}$

$\boldsymbol{p}$: projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$

<u>Alternative way to find least squares solution:</u>
If we know
the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$,
then
we can find the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$.

Use projection to find least squares solution

Example 5.3.9

A(least squares solution) **=** projection

Find the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Solve $\boldsymbol{Ax} = \boldsymbol{p}$

$$\mathbf{x} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

column space of ***A***

$$V = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

projection of ***b*** onto the column space of ***A***

$$\mathbf{p} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

see example 5.3.3

Use least squares solution to find projection

Example 5.3.11. ***A***(least squares solution) **=** projection

Find the projection of (1,1,1,1) onto

V = span{(1,-1,1,-1), (1,2,0,1), (2,1,1,0)}

Form matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$ $\quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

First find the least squares solution of ***Ax* = *b***

Solve $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ Theorem 5.3.10

$$\mathbf{x} = \begin{pmatrix} -t + \frac{2}{5} \\ -t + \frac{4}{5} \\ t \end{pmatrix} \longrightarrow \text{Take } \mathbf{u} = \begin{pmatrix} \frac{2}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix} \longrightarrow \mathbf{Au} = \begin{pmatrix} \frac{6}{5} \\ \frac{6}{5} \\ \frac{2}{5} \\ \frac{2}{5} \end{pmatrix}$$

Solution of $A^TAx = A^Tb \Leftrightarrow$ least squares solution of $Ax = b$

V = column space of A

Theorem 5.3.10

$\mathbf{A} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3)$

u is the least squares solution to $Ax = b$

if and only if u is a solution of $A^TAx = A^Tb$

$\Leftrightarrow$ Au is the projection of b onto V

$\Leftrightarrow$ $b - Au$ is orthogonal to V

$\Leftrightarrow$ $b - Au$ is orthogonal to a_1, a_2, a_3

$$\Leftrightarrow \mathbf{A}^T(\mathbf{b} - \mathbf{A}\mathbf{u}) = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{pmatrix} (\mathbf{b} - \mathbf{A}\mathbf{u}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$\Leftrightarrow$ $A^Tb - A^TAu = 0$

$\Leftrightarrow$ $A^TAu = A^Tb$

Solution of $A^TAx = A^Tb \Leftrightarrow$ least squares solution of $Ax = b$

Theorem 5.3.10

always exists

u is the least squares solution to $Ax = b$

$\Leftrightarrow$ u is a solution of $A^TAx = A^Tb$ always consistent

$\Leftrightarrow$ u is a solution of $Ax = p$ always consistent

where p is the projection of b onto column space of A

Theorem 5.3.8

Section 5.4

Another usage of "orthogonal"

Orthogonal Matrices

Objective

- What is an orthogonal matrix?
- How is orthogonal matrix related to orthonormal basis?
- How is transition matrix related to orthogonal matrix?

What is an orthogonal matrix?

Definition 5.4.3 & Remark 5.4.4

A square matrix $\boldsymbol{A}$ is called an orthogonal matrix

if $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$

Equivalently (and more easily),

if $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}$ (or $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$).

See Ex 2.12

All orthogonal matrices are invertible.

What is an orthogonal matrix?

Example 5.4.5

These are orthogonal matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

rotation anticlockwise through angle θ

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

inverse of each other

(multiply them to check)

Their transposes:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

rotation clockwise through angle θ

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Their transposes are also orthogonal matrices

Orthogonal matrix vs orthonormal basis for $\mathbf{R}^n$

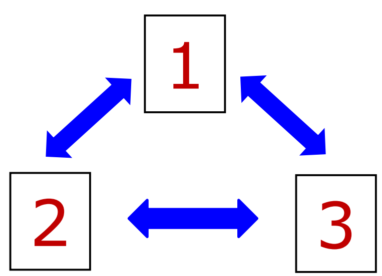
Theorem 5.4.6

Let $\boldsymbol{A}$ be a square matrix of order n.

The following statements are equivalent:

1. $\boldsymbol{A}$ is an orthogonal matrix.
2. The rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$.
3. The columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$.

Shall prove
(1) ⇔ (2)
and
(1) ⇔ (3)



The proof

1. $\boldsymbol{A}$ is orthogonal
2. The rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$

Theorem 5.4.6 (1 ⇔ 2)

For $i = 1, 2, \ldots, n$, let $\boldsymbol{a}_i$ be the ith row of $\boldsymbol{A}$.

$$\boldsymbol{A} = \begin{array}{cc} \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix} & \begin{array}{c} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \boldsymbol{a}_3 \end{array} \end{array} \qquad \boldsymbol{A}^T = \begin{array}{c} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\ \begin{array}{ccc} \boldsymbol{a}_1^T & \boldsymbol{a}_2^T & \boldsymbol{a}_3^T \end{array} \end{array}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{A}\boldsymbol{A}^T = \begin{pmatrix} \boldsymbol{a}_1\boldsymbol{a}_1^T & \boldsymbol{a}_1\boldsymbol{a}_2^T & \boldsymbol{a}_1\boldsymbol{a}_3^T \\ \boldsymbol{a}_2\boldsymbol{a}_1^T & \boldsymbol{a}_2\boldsymbol{a}_2^T & \boldsymbol{a}_2\boldsymbol{a}_3^T \\ \boldsymbol{a}_3\boldsymbol{a}_1^T & \boldsymbol{a}_3\boldsymbol{a}_2^T & \boldsymbol{a}_3\boldsymbol{a}_3^T \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{a}_1 & \boldsymbol{a}_1 \cdot \boldsymbol{a}_2 & \boldsymbol{a}_1 \cdot \boldsymbol{a}_3 \\ \boldsymbol{a}_2 \cdot \boldsymbol{a}_1 & \boldsymbol{a}_2 \cdot \boldsymbol{a}_2 & \boldsymbol{a}_2 \cdot \boldsymbol{a}_3 \\ \boldsymbol{a}_3 \cdot \boldsymbol{a}_1 & \boldsymbol{a}_3 \cdot \boldsymbol{a}_2 & \boldsymbol{a}_3 \cdot \boldsymbol{a}_3 \end{pmatrix}$$

orthogonal

$\boldsymbol{a}_i \cdot \boldsymbol{a}_i = 1$ for all i ⇔ $\|\boldsymbol{a}_i\| = 1$

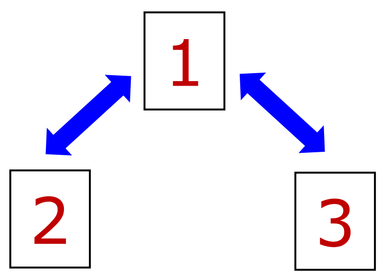
$\boldsymbol{a}_i \cdot \boldsymbol{a}_j = 0$ for $i \neq j$ ⇔ $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$ orthogonal

$\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is an orthonormal basis for $\mathbf{R}^n$

Theorem 5.4.6

1. $\boldsymbol{A}$ is an orthogonal matrix. $\Leftrightarrow \boldsymbol{A}^{\mathrm{T}}$ is orthogonal matrix
2. The rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$.
3. The columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbf{R}^n$.

We have proven
(1) $\Leftrightarrow$ (2)



Use $\boldsymbol{A}^{\mathrm{T}}$ to derive
(1) $\Leftrightarrow$ (3)

Discussion 5.4.1

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ be two bases for a vector space V.

Procedure to compute transition matrix $\boldsymbol{P}$ from S to T:

(i) write each $\boldsymbol{u}_i$ as a linear combination of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$.

(ii) use the coordinate vector $[\boldsymbol{u}_i]_T$ as the i th column $\boldsymbol{P}$.

$$\boldsymbol{P} = \left([\boldsymbol{u}_1]_T \ [\boldsymbol{u}_2]_T \cdots [\boldsymbol{u}_k]_T \right)$$

For any vector $\boldsymbol{w}$ in V, $[\boldsymbol{w}]_T = \boldsymbol{P}\,[\boldsymbol{w}]_S$.

Example 5.4.2

$S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$: the standard basis for $\mathbf{R}^3$

$$\mathbf{e}_1 = (1, 0, 0) \quad \mathbf{e}_2 = (0, 1, 0) \quad \mathbf{e}_3 = (0, 0, 1)$$

$T = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$: orthonormal basis for $\mathbf{R}^3$

$$\mathbf{u}_1 = \left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\right) \quad \mathbf{u}_2 = \left(\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}\right) \quad \mathbf{u}_3 = \left(\tfrac{1}{\sqrt{6}}, -\tfrac{2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}\right)$$

Transition matrix from T to S

$$\begin{cases} \mathbf{u}_1 = \frac{1}{\sqrt{3}}\mathbf{e}_1 + \frac{1}{\sqrt{3}}\mathbf{e}_2 + \frac{1}{\sqrt{3}}\mathbf{e}_3 \\ \mathbf{u}_2 = \frac{1}{\sqrt{2}}\mathbf{e}_1 - \frac{1}{\sqrt{2}}\mathbf{e}_3 \\ \mathbf{u}_3 = \frac{1}{\sqrt{6}}\mathbf{e}_1 - \frac{2}{\sqrt{6}}\mathbf{e}_2 + \frac{1}{\sqrt{6}}\mathbf{e}_3 \end{cases}$$

$$\boldsymbol{P} = \begin{array}{c} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \\ \begin{array}{ccc} \mathbf{u}_1^T & \mathbf{u}_2^T & \mathbf{u}_3^T \end{array} \end{array}$$

Transition matrix between orthonormal bases

Example 5.4.2

$S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$: the standard basis for $\mathbf{R}^3$

$$\mathbf{e}_1 = (1, 0, 0) \quad \mathbf{e}_2 = (0, 1, 0) \quad \mathbf{e}_3 = (0, 0, 1)$$

$T = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$: orthonormal basis for $\mathbf{R}^3$

$$\mathbf{u}_1 = \left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\right) \quad \mathbf{u}_2 = \left(\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}\right) \quad \mathbf{u}_3 = \left(\tfrac{1}{\sqrt{6}}, -\tfrac{2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}\right)$$

Transition matrix from S to T

$$\begin{array}{cccc} & \mathbf{e}_1\cdot\mathbf{u}_1 & \mathbf{e}_1\cdot\mathbf{u}_2 & \mathbf{e}_1\cdot\mathbf{u}_3 \end{array}$$

$$\begin{cases} \mathbf{e}_1 = \frac{1}{\sqrt{3}}\mathbf{u}_1 + \frac{1}{\sqrt{2}}\mathbf{u}_2 + \frac{1}{\sqrt{6}}\mathbf{u}_3 \\ \mathbf{e}_2 = \frac{1}{\sqrt{3}}\mathbf{u}_1 - \frac{2}{\sqrt{6}}\mathbf{u}_3 \\ \mathbf{e}_3 = \frac{1}{\sqrt{3}}\mathbf{u}_1 - \frac{1}{\sqrt{2}}\mathbf{u}_2 + \frac{1}{\sqrt{6}}\mathbf{u}_3 \end{cases}$$

Theorem 5.2.8.2

$$\boldsymbol{Q} = \begin{array}{c} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \end{array} \begin{array}{c} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{array}$$

Transition matrix between orthonormal bases

Example 5.4.2

Transition matrix from T to S

$$\begin{array}{c} \boldsymbol{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \\ \quad\quad \mathbf{u}_1^T \quad \mathbf{u}_2^T \quad \mathbf{u}_3^T \end{array}$$

Transition matrix from S to T

$$\boldsymbol{Q} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{array}{c} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{array}$$

By theorem 3.7.5

$$\left.\begin{array}{l} \boldsymbol{Q} = \boldsymbol{P}^{-1} \\ \boldsymbol{Q} = \boldsymbol{P}^{T} \end{array}\right\} \quad \boldsymbol{P}^{-1} = \boldsymbol{P}^{T}$$

S: orthonormal basis

T: orthonormal basis

So $\boldsymbol{P}$ is an orthogonal matrix

Transition matrix between orthonormal bases

Theorem 5.4.7

S and T: two orthonormal bases for a vector space.

The transition matrix $\boldsymbol{P}$ from S to T is orthogonal.

So $\boldsymbol{P}^T$ is the transition matrix from T to S.

Example 5.4.8.2

S : $\mathbf{u}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ $\mathbf{u}_2 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$ $\mathbf{u}_3 = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$

T : $\mathbf{v}_1 = (0, 0, 1)$ $\mathbf{v}_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)$ $\mathbf{v}_3 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$

transition matrix from S to T

$$\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{2} & \frac{3}{\sqrt{12}} \\ \frac{2}{\sqrt{6}} & \frac{1}{2} & -\frac{1}{\sqrt{12}} \end{pmatrix}$$

take transpose

transition matrix from T to S

$$\mathbf{P}^T = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{6}} & \frac{3}{\sqrt{12}} & -\frac{1}{\sqrt{12}} \end{pmatrix}$$

The proof

Theorem 5.4.7

S and T are orthonormal bases

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$.

By Theorem 5.2.8.2

$$\begin{cases} \mathbf{u}_1 = \boxed{(\mathbf{u}_1 \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}_1 \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u}_1 \cdot \mathbf{v}_k)\mathbf{v}_k} \\ \mathbf{u}_2 = \boxed{(\mathbf{u}_2 \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}_2 \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u}_2 \cdot \mathbf{v}_k)\mathbf{v}_k} \\ \quad \vdots \\ \mathbf{u}_k = \boxed{(\mathbf{u}_k \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}_k \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{u}_k \cdot \mathbf{v}_k)\mathbf{v}_k} \end{cases}$$

The transition matrix from S to T is

$$\mathbf{P} = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{v}_1 & \mathbf{u}_2 \cdot \mathbf{v}_1 & \cdots & \mathbf{u}_k \cdot \mathbf{v}_1 \\ \mathbf{u}_1 \cdot \mathbf{v}_2 & \mathbf{u}_2 \cdot \mathbf{v}_2 & \cdots & \mathbf{u}_k \cdot \mathbf{v}_2 \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_1 \cdot \mathbf{v}_k & \mathbf{u}_2 \cdot \mathbf{v}_k & \cdots & \mathbf{u}_k \cdot \mathbf{v}_k \end{pmatrix}$$

The proof

Theorem 5.4.7

S and *T* are orthonormal bases

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$.

By Theorem 5.2.8.2

$$\begin{cases} \mathbf{v}_1 = \boxed{(\mathbf{v}_1 \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_1 \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_1 \cdot \mathbf{u}_k)\mathbf{u}_k} \\ \mathbf{v}_2 = \boxed{(\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_2 \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_2 \cdot \mathbf{u}_k)\mathbf{u}_k} \\ \quad \vdots \\ \mathbf{v}_k = \boxed{(\mathbf{v}_k \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_k \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_k \cdot \mathbf{u}_k)\mathbf{u}_k} \end{cases}$$

The transition matrix from *T* to *S* is

$$\mathbf{Q} = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{u}_1 & \mathbf{v}_2 \cdot \mathbf{u}_1 & \cdots & \mathbf{v}_k \cdot \mathbf{u}_1 \\ \mathbf{v}_1 \cdot \mathbf{u}_2 & \mathbf{v}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{v}_k \cdot \mathbf{u}_2 \\ \vdots & \vdots & & \vdots \\ \mathbf{v}_1 \cdot \mathbf{u}_k & \mathbf{v}_2 \cdot \mathbf{u}_k & \cdots & \mathbf{v}_k \cdot \mathbf{u}_k \end{pmatrix}$$

The proof

Theorem 5.4.7

S and T are orthonormal bases

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$.

transition matrix from S to T — inverse of each other — transition matrix from T to S

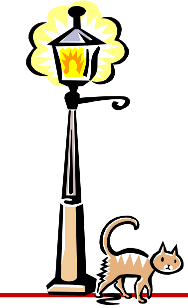
$$\boldsymbol{P} = \begin{pmatrix} \boxed{\begin{matrix} \boldsymbol{u}_1 \cdot \boldsymbol{v}_1 \\ \boldsymbol{u}_1 \cdot \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{u}_1 \cdot \boldsymbol{v}_k \end{matrix}} & \boxed{\begin{matrix} \boldsymbol{u}_2 \cdot \boldsymbol{v}_1 \\ \boldsymbol{u}_2 \cdot \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{u}_2 \cdot \boldsymbol{v}_k \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \\ \cdots \end{matrix} & \boxed{\begin{matrix} \boldsymbol{u}_k \cdot \boldsymbol{v}_1 \\ \boldsymbol{u}_k \cdot \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{u}_k \cdot \boldsymbol{v}_k \end{matrix}} \end{pmatrix} \qquad \mathbf{Q} = \begin{pmatrix} \boxed{\mathbf{v}_1 \cdot \mathbf{u}_1 \quad \mathbf{v}_2 \cdot \mathbf{u}_1 \quad \cdots \quad \mathbf{v}_k \cdot \mathbf{u}_1} \\ \boxed{\mathbf{v}_1 \cdot \mathbf{u}_2 \quad \mathbf{v}_2 \cdot \mathbf{u}_2 \quad \cdots \quad \mathbf{v}_k \cdot \mathbf{u}_2} \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \boxed{\mathbf{v}_1 \cdot \mathbf{u}_k \quad \mathbf{v}_2 \cdot \mathbf{u}_k \quad \cdots \quad \mathbf{v}_k \cdot \mathbf{u}_k} \end{pmatrix}$$

We have $\boldsymbol{Q} = \boldsymbol{P}^T$

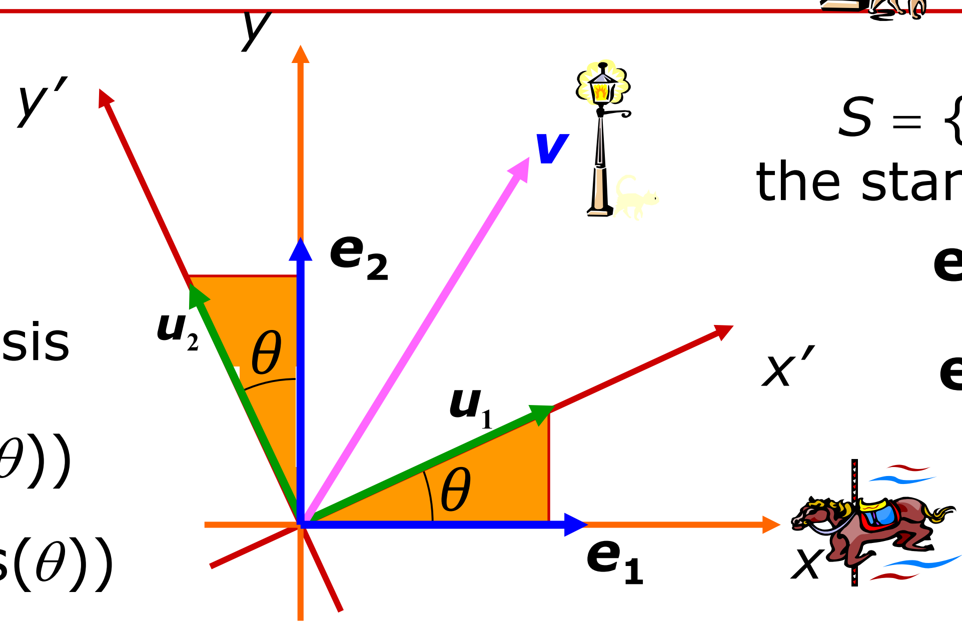
We also have $\boldsymbol{Q} = \boldsymbol{P}^{-1}$

So $\boldsymbol{P}^{-1} = \boldsymbol{P}^T$, i.e. $\boldsymbol{P}$ is orthogonal.

Rotation of *xy*-coordinates



Example 5.4.8.1



$S = \{\mathbf{e}_1, \mathbf{e}_2\}$
the standard basis

$\mathbf{e}_1 = (1, 0)$

$\mathbf{e}_2 = (0, 1)$

$T = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$
an orthonormal basis

$\boldsymbol{u}_1 = (\cos(\theta), \sin(\theta))$

$\boldsymbol{u}_2 = (-\sin(\theta), \cos(\theta))$

new *x′y′*-coordinate system

What is the coordinate of $\boldsymbol{v}$ w.r.t. the new coordinate system? Ans: $[\boldsymbol{v}]_T$

What is the transition matrix between *S* and *T* ?

Rotation of *xy*-coordinates

$$\mathbf{u}_1 = (\cos(\theta), \sin(\theta))$$
$$\mathbf{u}_2 = (-\sin(\theta), \cos(\theta))$$

Example 5.4.8.1

$S = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$
the standard basis

$T = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$
an orthonormal basis

transition matrix from T to S

$$\boldsymbol{P} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

transition matrix from S to T

$$\boldsymbol{P}^T = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

What is the coordinate of $\boldsymbol{v}$ w.r.t. the new coordinate system?

$$[\boldsymbol{v}]_T = \boldsymbol{P}^T[\boldsymbol{v}]_S$$

coordinates of $\boldsymbol{v}$ in the new x′y′-coordinate system

usual coordinates of $\boldsymbol{v}$

Rotation of *xy*-coordinates

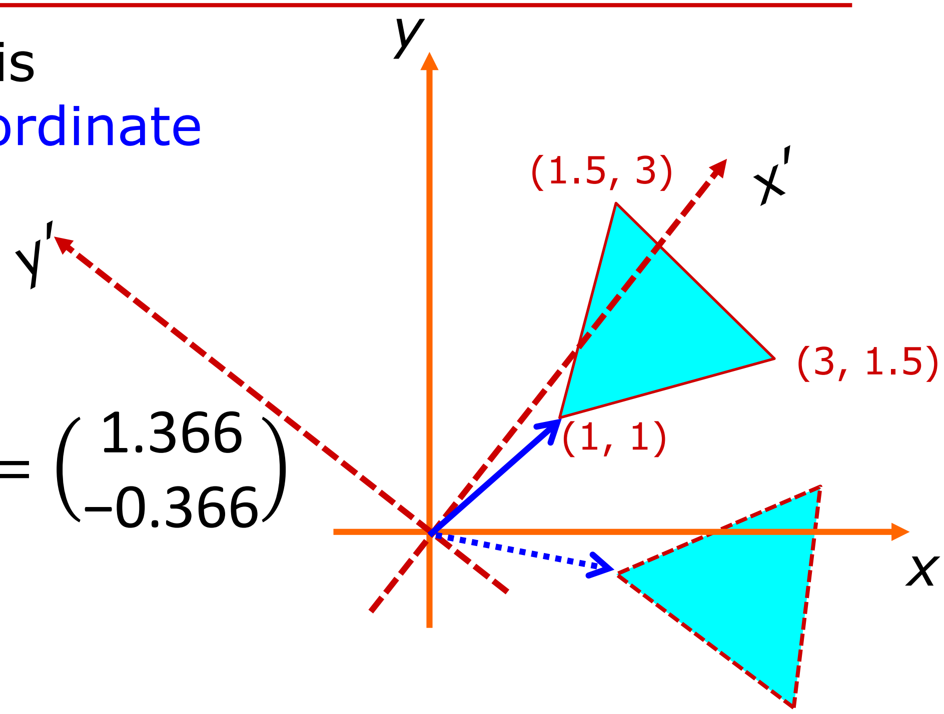
Quiz Time

A new $x'y'$-coordinate system is obtained by rotating the xy-coordinate anti-clockwise by 60°.

What is the $x'y'$-coordinates of vector (1,1)?

$$\boxed{\begin{pmatrix}\cos(\frac{\pi}{3}) & \sin(\frac{\pi}{3})\\ -\sin(\frac{\pi}{3}) & \cos(\frac{\pi}{3})\end{pmatrix}}\begin{pmatrix}1 & 3 & 1.5\\ 1 & 1.5 & 3\end{pmatrix} = \begin{pmatrix}\frac{1+\sqrt{3}}{2} & \frac{3\sqrt{3}+1.5}{2} & \frac{1.5+3\sqrt{3}}{2}\\ \frac{1-\sqrt{3}}{2} & \frac{1.5-3\sqrt{3}}{2} & \frac{3-1.5\sqrt{3}}{2}\end{pmatrix}$$

$$= \begin{pmatrix}1.366\\ -0.366\end{pmatrix}$$



Same effect as fixing the xy-coordinate and rotate the vector clockwise by 60°.

Section 6.1

Eigenvalues and Eigenvectors

Objectives

- What are Eigenvalues, Eigenvectors and Eigenspace?
- How to find eigenvalues and eigenvectors of a matrix?
- How is eigenvalue related to invertibility of matrix?

Google page rank

Google ranks webpages according to "hyperlinks"

e.g. we want to rank 4 webpages: A, B, C, D

Form a 4x4 matrix:

$$\begin{array}{c c} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 1 & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{1}{2} & 0 \end{pmatrix} \end{array} \xrightarrow{\text{eigenvector}} \begin{pmatrix} 0.446 \\ 0.223 \\ 0.743 \\ 0.446 \end{pmatrix} \quad \begin{array}{l} \text{page rank} \\ 2 \text{ (tie)} \\ 4 \\ 1 \\ 2 \text{ (tie)} \end{array}$$

A has a link to C, but not to B and D

B has a link to A, C, D

Power of matrices revisited

Example 6.1.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$$

$$\boxed{\boldsymbol{A}^n = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}^n}$$

"Factorize" $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}}_{\boldsymbol{P}} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}}_{\boldsymbol{D}} \underbrace{\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}}_{\boldsymbol{P}^{-1}}$$

$$\boldsymbol{A} = \boldsymbol{PDP}^{-1}$$

$$\boldsymbol{A}^n = (\boldsymbol{PDP}^{-1})^n \quad \neq \boldsymbol{P}^n\boldsymbol{D}^n\boldsymbol{P}^{-n}$$

$$= (\boldsymbol{PDP}^{-1})(\boldsymbol{PDP}^{-1})(\boldsymbol{PDP}^{-1}) \cdots (\boldsymbol{PDP}^{-1}) \quad (n \text{ times})$$

$$= \boldsymbol{PD\,D} \cdots \boldsymbol{DP}^{-1}$$

$$= \boldsymbol{PD}^n\boldsymbol{P}^{-1}$$

Power of matrices revisited

$$\boldsymbol{A}^n = \boldsymbol{P}\boldsymbol{D}^n\boldsymbol{P}^{-1}$$

Example 6.1.1

$$\boldsymbol{D}^n = \begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}^n = \begin{pmatrix} 1^n & 0 \\ 0 & 0.95^n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0.95^n \end{pmatrix}$$

$$\begin{aligned} \boldsymbol{A}^{100} &= \boldsymbol{P}\boldsymbol{D}^{100}\boldsymbol{P}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0.95^{100} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0.2047 & 0.1988 \\ 0.7953 & 0.8012 \end{pmatrix} \end{aligned}$$

Remark 6.1.2

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix} \boxed{\begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}} \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}$$

This is called "diagonalizing" a square matrix.

We need the concept of eigenvalues and eigenvectors.

What are eigenvalue and eigenvector?

Definition 6.1.3

$$\boldsymbol{A}\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}2\\4\\6\end{bmatrix}$$

relation?

Let $\boldsymbol{A}$ be a square matrix of order n.

Let $\boldsymbol{x}$ be a nonzero (column) vector in $\mathbf{R}^n$

If $\boldsymbol{Ax}$ = scalar multiple of $\boldsymbol{x}$ — $\boldsymbol{Ax}$ and $\boldsymbol{x}$ are parallel

$= \lambda\boldsymbol{x}$ for some scalar λ — lambda

then $\boldsymbol{x}$ is called an eigenvector of $\boldsymbol{A}$

The scalar λ is called an eigenvalue of $\boldsymbol{A}$ and $\boldsymbol{x}$ is said to be an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue λ.

Visual

What are eigenvalue and eigenvector?

Example 6.1.4.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} \qquad \boldsymbol{x} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \qquad \boldsymbol{y} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\boldsymbol{Ax} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}\begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \mathbf{x}$$

$\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ with the eigenvalue 1.

$$\boldsymbol{Ay} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0.95 \\ -0.95 \end{pmatrix} = 0.95\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0.95\boldsymbol{y}$$

$\boldsymbol{y}$ is an eigenvector of $\boldsymbol{A}$ with the eigenvalue 0.95.

$$\boxed{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1}}$$

Example 6.1.4.2

$$\mathbf{B}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\qquad \mathbf{x}=\begin{pmatrix}1\\1\\1\end{pmatrix}\qquad \mathbf{y}=\begin{pmatrix}1\\0\\-1\end{pmatrix}\qquad \mathbf{z}=\begin{pmatrix}1\\-2\\1\end{pmatrix}$$

$$\boxed{\mathbf{Bx}}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}3\\3\\3\end{pmatrix}=3\begin{pmatrix}1\\1\\1\end{pmatrix}=\boxed{3\mathbf{x}}$$

$\boldsymbol{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{B}(2\mathbf{x})=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\begin{pmatrix}2\\2\\2\end{pmatrix}=\begin{pmatrix}6\\6\\6\end{pmatrix}=3\begin{pmatrix}2\\2\\2\end{pmatrix}=3(2\mathbf{x})$$

$2\boldsymbol{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{B}(k\mathbf{x})=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\begin{pmatrix}k\\k\\k\end{pmatrix}=3\begin{pmatrix}k\\k\\k\end{pmatrix}=3(k\mathbf{x})$$

$k\boldsymbol{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}^{-1}$$

Example 6.1.4.2

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\mathbf{Bx} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 3\mathbf{x}$$

x is an eigenvector associated with eigenvalue 3

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

y is an eigenvector associated with eigenvalue 0

$$\mathbf{Bz} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0\mathbf{z}$$

z is an eigenvector associated with eigenvalue 0

Eigenvalues of triangular matrices

Theorem 6.1.9 & Example 6.1.10

If $\boldsymbol{A}$ is a triangular matrix, in particular, diagonal matrix

the eigenvalues of $\boldsymbol{A}$ are the diagonal entries of $\boldsymbol{A}$.

$$\begin{pmatrix} -1 & 3.5 & 14 \\ 0 & 5 & -26 \\ 0 & 0 & 2 \end{pmatrix}$$

The eigenvalues are -1, 5 and 2.

$$\begin{pmatrix} -2 & 0 & 0 \\ 99 & 0 & 0 \\ 10 & -4.5 & 10 \end{pmatrix}$$

The eigenvalues are -2, 0 and 10.

The proof

Theorem 6.1.9

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ 0 & & & a_{nn} \end{pmatrix} \quad \lambda\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ & \lambda - a_{22} & \cdots & -a_{2n} \\ & & \ddots & \vdots \\ 0 & & & \lambda - a_{nn} \end{pmatrix}$$

characteristic polynomial of $\boldsymbol{A}$

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

this polynomial is completely factorized

roots of the polynomial: $a_{11}, a_{22}, \ldots, a_{nn}$

eigenvalues of $\boldsymbol{A}$ $\Longleftrightarrow$ diagonal entries of $\boldsymbol{A}$

How to find eigenvalues?

$(\lambda - \boldsymbol{A})\,\boldsymbol{x} = \boldsymbol{0}$

Remark 6.1.5

Let $\boldsymbol{A}$ be a square matrix of order n.

λ is an eigenvalue of $\boldsymbol{A}$

$\Leftrightarrow$ $\boldsymbol{Ax} = \lambda\,\boldsymbol{x}$ for some nonzero column vector $\boldsymbol{x}$

$\Leftrightarrow$ $\lambda\,\boldsymbol{x} - \boldsymbol{Ax} = \boldsymbol{0}$ for some nonzero column vector $\boldsymbol{x}$

$\Leftrightarrow$ $\underbrace{(\lambda \boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{x} = \boldsymbol{0}}_{\text{homog. system}}$ for some nonzero column vector $\boldsymbol{x}$

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

has non-trivial solutions

$\Leftrightarrow$ $\underbrace{\det(\lambda\,\boldsymbol{I} - \boldsymbol{A})}_{\text{a polynomial in } \lambda} = 0$

Solve this equation to find the eigenvalues of $\boldsymbol{A}$

How to find eigenvalues?

Example 6.1.7.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$$

The eigenvalues of $\boldsymbol{A}$ are 1 and 0.95.

$$\begin{aligned}
\det(\lambda \boldsymbol{I} - \boldsymbol{A}) &= \det\left(\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}\right) \\
&= \begin{vmatrix} \lambda - 0.96 & -0.01 \\ -0.04 & \lambda - 0.99 \end{vmatrix} \\
&= (\lambda - 0.96)(\lambda - 0.99) - (-0.01)(-0.04) \\
&= \lambda^2 - 1.95\lambda + 0.95 \quad \text{polynomial of degree 2} \\
&= (\lambda - 1)(\lambda - 0.95) \quad \text{factorize the polynomial}
\end{aligned}$$

$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ if and only if $\lambda = 1$ or 0.95

What is characteristic polynomial?

Definition 6.1.6

Let $\boldsymbol{A}$ be a square matrix of order n.

The polynomial $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ degree n

is called the characteristic polynomial of $\boldsymbol{A}$.

λ is an eigenvalue of $\boldsymbol{A}$ $\Leftrightarrow$ $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$

$\Leftrightarrow$ λ is a root of the characteristic polynomial

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n1} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

Finding eigenvalues from characteristic polynomial

Example 6.1.7.3

$$\boldsymbol{C} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

The eigenvalues of $\boldsymbol{C}$ are 1, $\sqrt{2}$ and $-\sqrt{2}$

$$\det(\lambda \boldsymbol{I} - \boldsymbol{C}) = \begin{vmatrix} \lambda & 1 & 0 \\ 0 & \lambda & -2 \\ -1 & -1 & \lambda - 1 \end{vmatrix}$$

characteristic polynomial of $\boldsymbol{C}$

$$= \lambda^3 - \lambda^2 - 2\lambda + 2$$

one factor is (λ - 1)

$$= (\lambda - 1)(\lambda^2 - 2)$$

$$= (\lambda - 1)(\lambda - \sqrt{2})(\lambda + \sqrt{2})$$

guess one root

$\lambda = 1$

$\det(\lambda \boldsymbol{I} - \boldsymbol{C}) = 0$ if and only if $\lambda = 1$, $\sqrt{2}$ or $-\sqrt{2}$

A very^3 important theorem (revisited)

Theorem 6.1.8

1,2,3,4,5,6,7,8

9

A is an n x n matrix

The following statements are equivalent:

1. ***A*** is invertible.
2. The linear system $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.
3. The reduced row-echelon form of ***A*** is ***I***.
4. ***A*** can be expressed as a product of elementary matrices.
5. $\det(\boldsymbol{A}) \neq 0$.
6. The rows of ***A*** form a basis for $\mathbf{R}^n$.
7. The columns of ***A*** form a basis for $\mathbf{R}^n$.
8. $\mathrm{rank}(\boldsymbol{A}) = n$.

9. 0 is not an eigenvalue of ***A***.

5. $\det(\boldsymbol{A}) \neq 0$

9. 0 is not an eigenvalue of $\boldsymbol{A}$

The proof

Theorem 6.1.8

We are going to show "5 ⇔ 9".

Statement 9 0 is not an eigenvalue of $\boldsymbol{A}$

$\Leftrightarrow$ 0 is not a root of the char. poly. $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$

$\Leftrightarrow \quad \det(0\boldsymbol{I} - \boldsymbol{A}) \neq 0$

$\Leftrightarrow \quad \det(-\boldsymbol{A}) \neq 0$

$\Leftrightarrow \quad (-1)^n \det(\boldsymbol{A}) \neq 0$

$\Leftrightarrow \quad \det(\boldsymbol{A}) \neq 0$ Statement 5

How to find eigenvectors?

Remark 6.1.5

eigenvector

Let $\boldsymbol{A}$ be a square matrix of order n.

λ is an eigenvalue of $\boldsymbol{A}$

$\Leftrightarrow$ $\boldsymbol{Ax} = \lambda \boldsymbol{x}$ for some nonzero column vector $\boldsymbol{x}$

$\Leftrightarrow$ $\lambda \boldsymbol{x} - \boldsymbol{Ax} = \boldsymbol{0}$ for some nonzero column vector $\boldsymbol{x}$

$\Leftrightarrow$ $\underbrace{(\lambda \boldsymbol{I} - \boldsymbol{A}) \boldsymbol{x} = \boldsymbol{0}}_{\text{homog. system}}$ for some nonzero column vector $\boldsymbol{x}$

homog. system has non-trivial solutions

$\Leftrightarrow$ $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$

by solving this linear system

its solution space contains all the eigenvectors associated to λ

What is an eigenspace of a matrix?

Definition 6.1.11 (Eigenspace)

A : square matrix of order n

λ : an eigenvalue of ***A***

The solution space of the linear system

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

has nontrivial solutions

is called the eigenspace of ***A***

associated with the eigenvalue λ

denoted by E_λ

If ***u*** is a nonzero vector in E_λ,

then ***u*** is an eigenvector of ***A*** associated with the eigenvalue λ.

Eigenspace of a matrix

Example 6.1.12.1

$$\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$$

By Example 6.1.8.1,

the eigenvalues of $\boldsymbol{A}$ are 1 and 0.95.

$\boldsymbol{A}$ has two eigenspaces E_1 and $E_{0.95}$

How to find eigenspace?

Example 6.1.12.1 (Find E_1)

For $\lambda = 1$,

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 1-0.96 & -0.01 \\ -0.04 & 1-0.99 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

General solution

$$\begin{pmatrix} x \\ y \end{pmatrix} = t\begin{pmatrix} 0.25 \\ 1 \end{pmatrix}$$

t an arbitrary parameter

$$E_1 = \text{span}\boxed{\left\{\begin{pmatrix} 0.25 \\ 1 \end{pmatrix}\right\}}$$

any non-zero scalar multiple of $\begin{pmatrix} 0.25 \\ 1 \end{pmatrix}$ is an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue 1

Basis for the eigenspace E_1

How to find eigenspace?

Example 6.1.12.1 (Find $E_{0.95}$)

For $\lambda = 0.95$,

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 0.95-0.96 & -0.01 \\ -0.04 & 0.95-0.99 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

General solution

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

t an arbitrary parameter

$$E_{0.95} = \text{span}\boxed{\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}}$$

any non-zero scalar multiple of $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ is an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue 0.95

Basis for the eigenspace $E_{0.95}$

Example 6.1.12.2

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

By Example 6.1.8.2,
the eigenvalues of $\boldsymbol{B}$ are 3 and 0.

$\boldsymbol{B}$ has two eigenspaces E_3 and E_0

How to find eigenspace?

Example 6.1.12.2 (Find E_0)

For $\lambda = 0$,

$$(\lambda \boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 0-1 & -1 & -1 \\ -1 & 0-1 & -1 \\ -1 & -1 & 0-1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

General solution

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

s, t are arbitrary parameters

$$E_0 = \text{span}\boxed{\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}}$$

Basis for the eigenspace E_0

any non-zero linear combination of $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ is an eigenvector of $\boldsymbol{B}$ associated with the eigenvalue 0