Linear Algebra Review

MLT

Karthik Thiagarajan

Vectors

Vectors

Features

- num of rooms
- size of house
- age of house
- parking facility

Data

ML: "Learning from data"

Vectors

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$\mathbf{x} \in \mathbb{R}^d$

feature-vector

Algebra

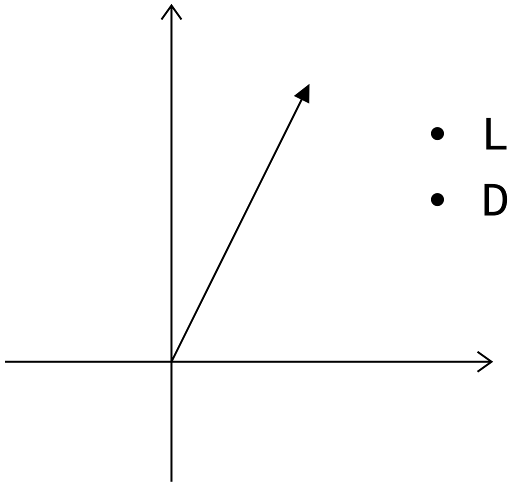
Features

- num of rooms
- size of house
- age of house
- parking facility

Data

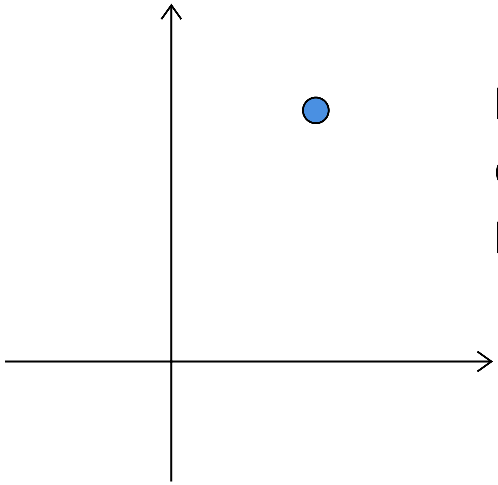
ML: "Learning from data"

Vectors



- Length (norm)
- Direction

Geometry



Data-point
or
Point

Features
- num of rooms
- size of house
- age of house
- parking facility

Data

ML: "Learning from data"

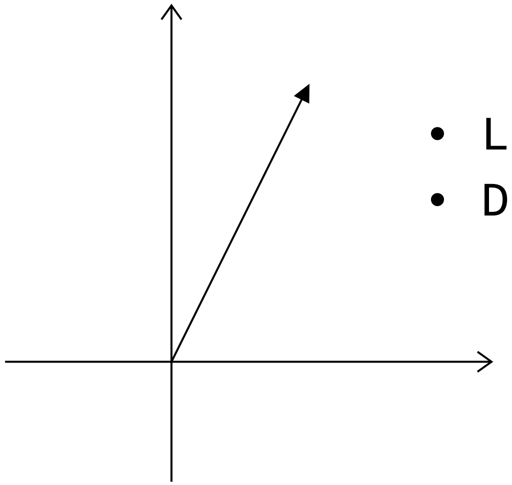
$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$\mathbf{x} \in \mathbb{R}^d$

feature-vector

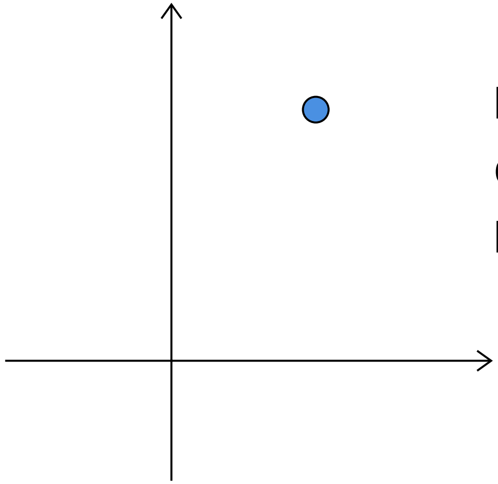
Algebra

Vectors



- Length (norm)
- Direction

Geometry



Data-point
or
Point

Features
- num of rooms
- size of house
- age of house
- parking facility

Data

ML: "Learning from data"

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$\mathbf{x} \in \mathbb{R}^d$

feature-vector

Algebra

[1, 2, 3, 4]

Array
or
List

Computation

$\mathbb{R}^d$: Conventions

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \qquad \mathbf{x}^T = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

- $\mathbf{x}$ is a column vector by default.
- Its shape is $(d,\ 1)$.
- $\mathbf{x}^T$ is a row-vector.
- This is a convention.

$\mathbb{R}^d$: Conventions

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}$$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \qquad \mathbf{x}^T = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

- $\mathbf{x}$ is a column vector by default.
- Its shape is $(d,\ 1)$.
- $\mathbf{x}^T$ is a row-vector.
- This is a convention.

- $\mathbf{X}$ is a data-matrix.
- Columns of $\mathbf{X}$ are feature vectors.
- Its shape is (d, n).
- This is a convention.

$\mathbb{R}^d$: Basis

$\mathbb{R}^d$: Basis

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$\mathbb{R}^d$: Basis

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- basis
  - not unique

$\mathbb{R}^d$: Basis

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- basis
  - not unique
  - $\{\mathbf{e_1}, \mathbf{e_2}\}$ is the standard basis

$\mathbb{R}^d$: Basis

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- basis
  - not unique
  - $\{\mathbf{e_1}, \mathbf{e_2}\}$ is the standard basis
- representation in a given basis

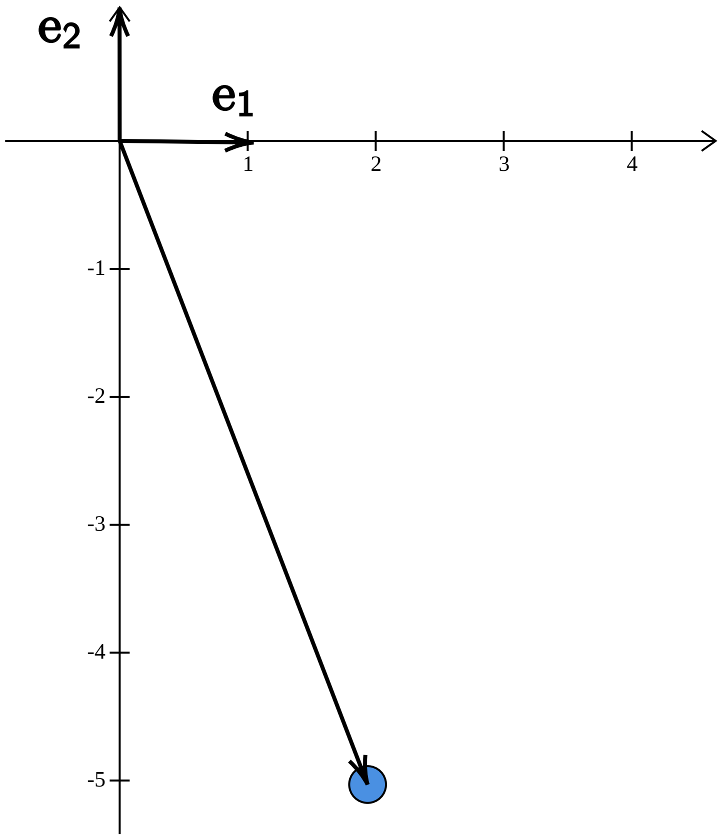
$\mathbb{R}^d$: Basis

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- basis
  - not unique
  - $\{\mathbf{e_1}, \mathbf{e_2}\}$ is the standard basis
- representation in a given basis
  - unique

$\mathbb{R}^d$: Basis

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

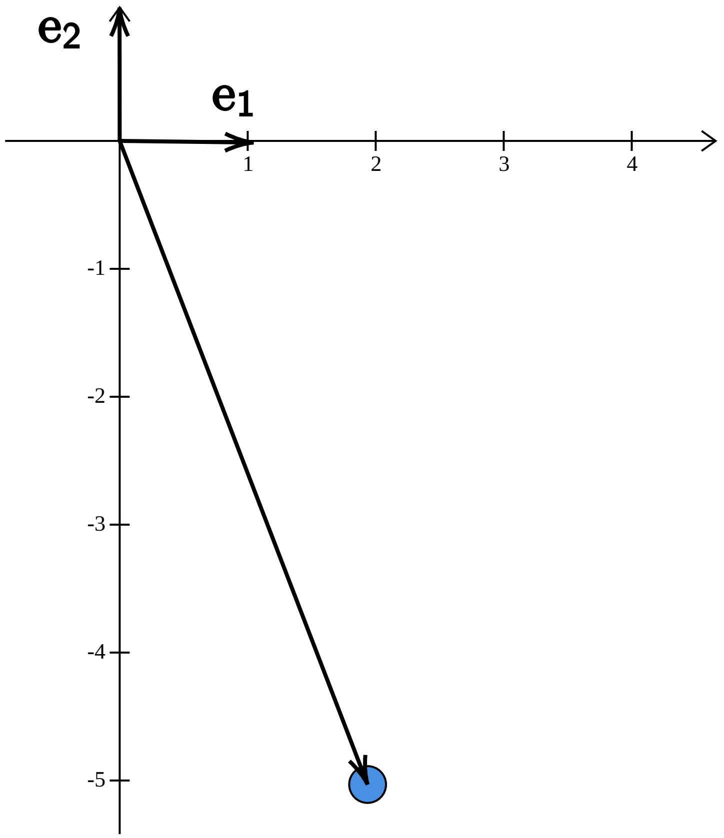


- basis
  - not unique
  - $\{\mathbf{e_1}, \mathbf{e_2}\}$ is the standard basis
- representation in a given basis
  - unique

$\mathbb{R}^d$: Basis

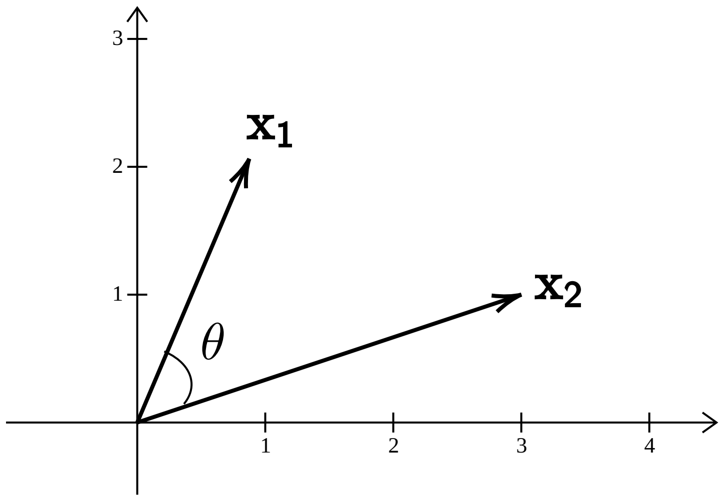
$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- basis
  - not unique
  - $\{\mathbf{e_1}, \mathbf{e_2}\}$ is the standard basis
- representation in a given basis
  - unique



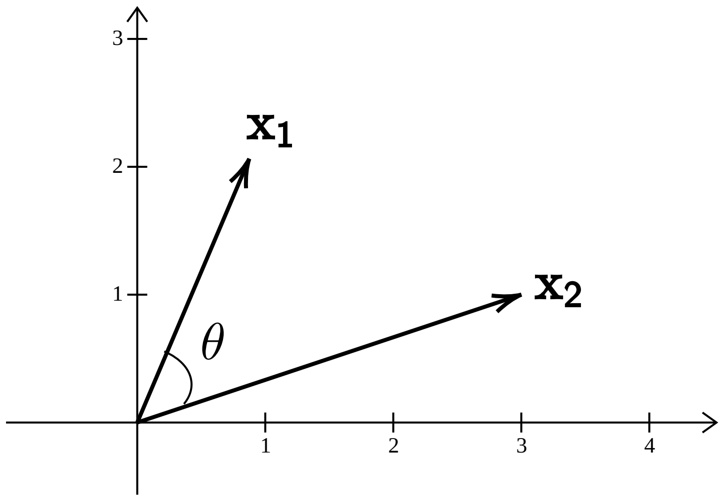
$$\begin{bmatrix} 2 \\ -5 \end{bmatrix} = 2 \cdot \mathbf{e_1} - 5 \cdot \mathbf{e_2}$$

$\mathbb{R}^d$: Dot Product



$\mathbb{R}^d$: Dot Product

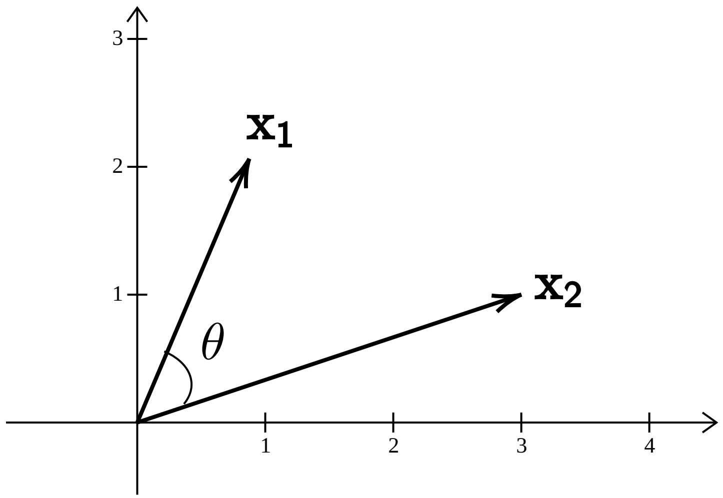
$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{x_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$



$\mathbb{R}^d$: Dot Product

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ \mathbf{x_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

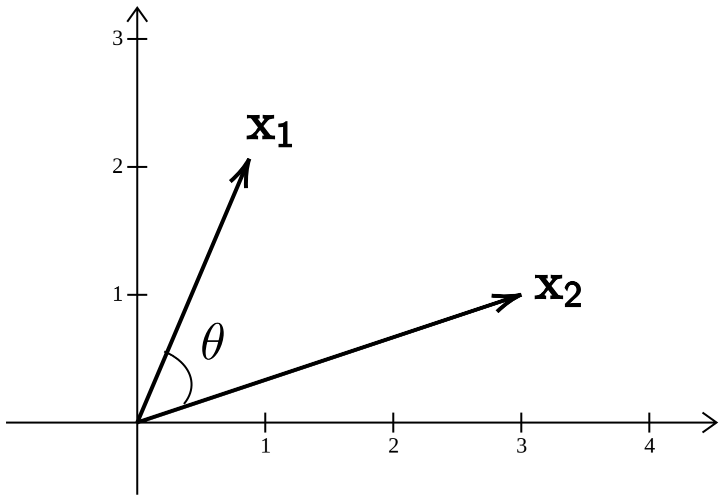
$$\mathbf{x_1} \cdot \mathbf{x_2}$$



$\mathbb{R}^d$: Dot Product

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ \mathbf{x_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

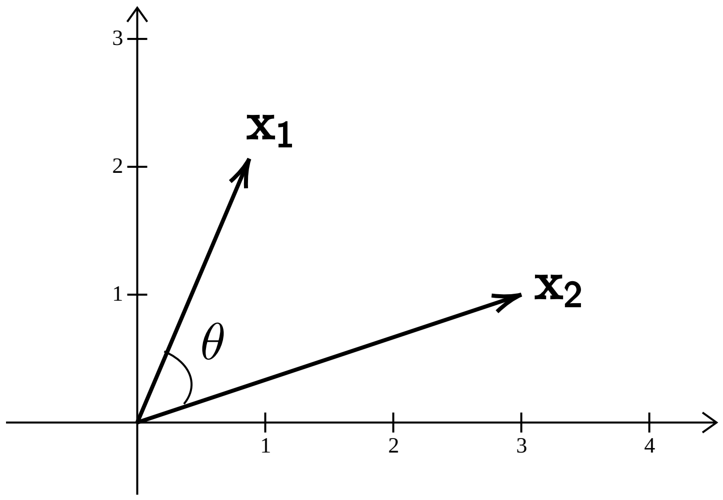
$$\mathbf{x_1} \cdot \mathbf{x_2} = \mathbf{x_1}^T \mathbf{x_2}$$



$\mathbb{R}^d$: Dot Product

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{x_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

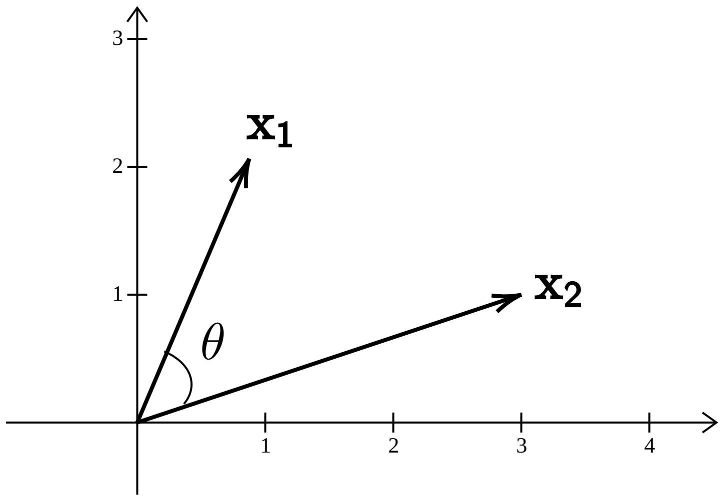
$$\mathbf{x_1} \cdot \mathbf{x_2} = \mathbf{x_1}^T \mathbf{x_2} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5$$



$\mathbb{R}^d$: Dot Product

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ \mathbf{x_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{x_1} \cdot \mathbf{x_2} = \mathbf{x_1}^T \mathbf{x_2} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5$$



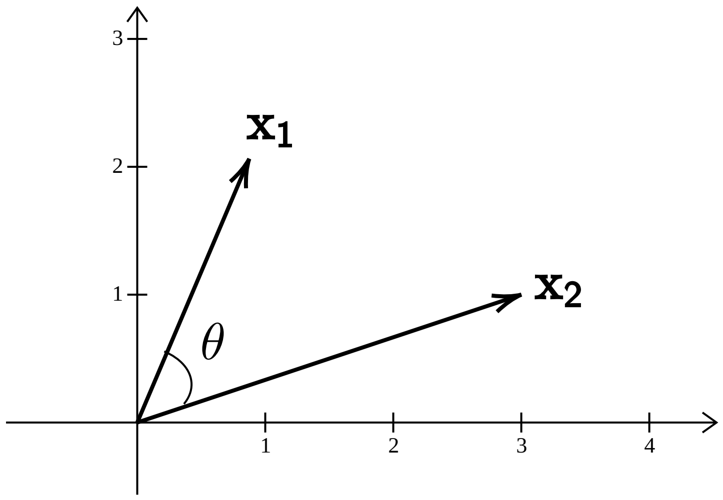
$$||\mathbf{x_1}||^2 =$$

$$||\mathbf{x_2}||^2 =$$

$\mathbb{R}^d$: Dot Product

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{x_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{x_1} \cdot \mathbf{x_2} = \mathbf{x_1}^T \mathbf{x_2} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5$$



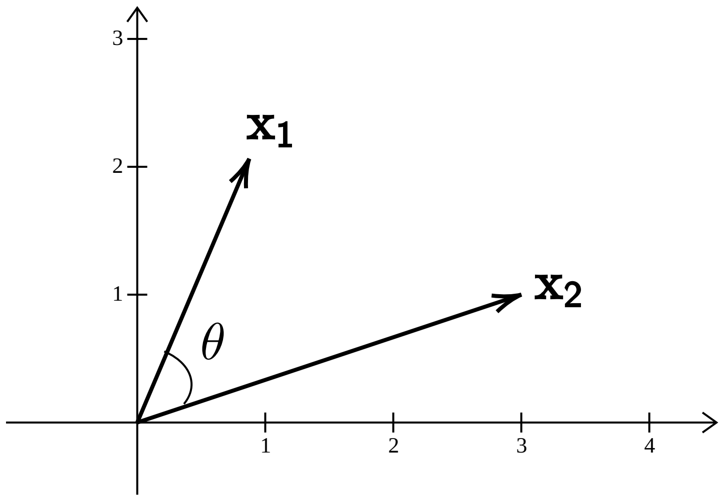
$$||\mathbf{x_1}||^2 = \mathbf{x_1}^T \mathbf{x_1} = 1^2 + 2^2 = 5$$

$$||\mathbf{x_2}||^2 = \mathbf{x_2}^T \mathbf{x_2} = 3^2 + 1^2 = 10$$

$\mathbb{R}^d$: Dot Product

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{x_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{x_1} \cdot \mathbf{x_2} = \mathbf{x_1}^T \mathbf{x_2} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5$$



$$||\mathbf{x_1}||^2 = \mathbf{x_1}^T \mathbf{x_1} = 1^2 + 2^2 = 5$$

$$||\mathbf{x_2}||^2 = \mathbf{x_2}^T \mathbf{x_2} = 3^2 + 1^2 = 10$$

$$\cos\theta = \frac{\mathbf{x_1}^T \mathbf{x_2}}{||\mathbf{x_1}|| \cdot ||\mathbf{x_2}||} = \frac{5}{\sqrt{5} \cdot \sqrt{10}} = \frac{1}{\sqrt{2}}$$

$$\implies \theta = 45^\circ$$

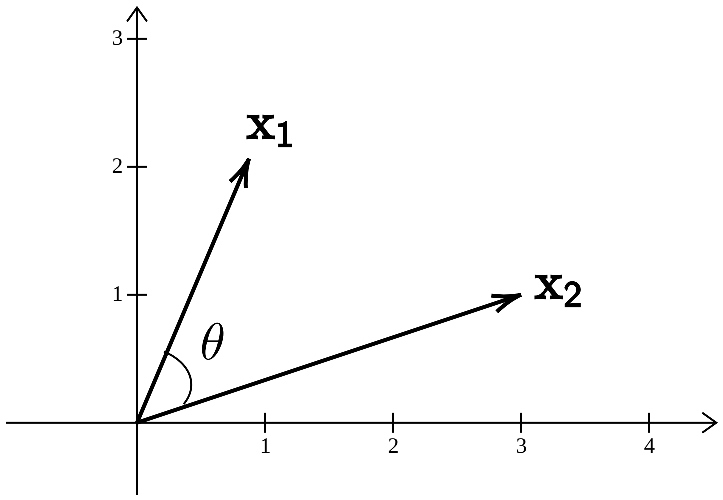
$\mathbb{R}^d$: Dot Product

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ \mathbf{x_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Geometry:

- Lengths
- Angles (directions)

$$\mathbf{x_1} \cdot \mathbf{x_2} = \mathbf{x_1}^T \mathbf{x_2} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5$$



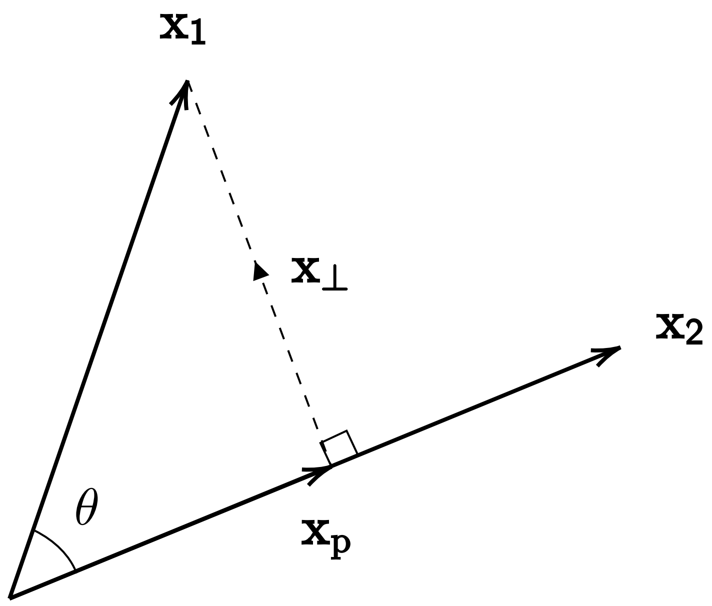
$$||\mathbf{x_1}||^2 = \mathbf{x_1}^T \mathbf{x_1} = 1^2 + 2^2 = 5$$

$$||\mathbf{x_2}||^2 = \mathbf{x_2}^T \mathbf{x_2} = 3^2 + 1^2 = 10$$

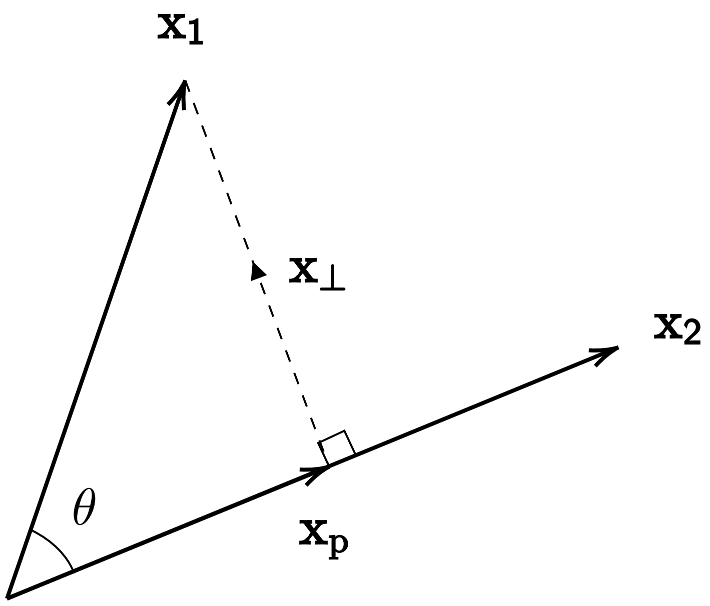
$$\cos\theta = \frac{\mathbf{x_1}^T \mathbf{x_2}}{||\mathbf{x_1}|| \cdot ||\mathbf{x_2}||} = \frac{5}{\sqrt{5} \cdot \sqrt{10}} = \frac{1}{\sqrt{2}}$$

$$\implies \theta = 45^\circ$$

$\mathbb{R}^d$: Projections



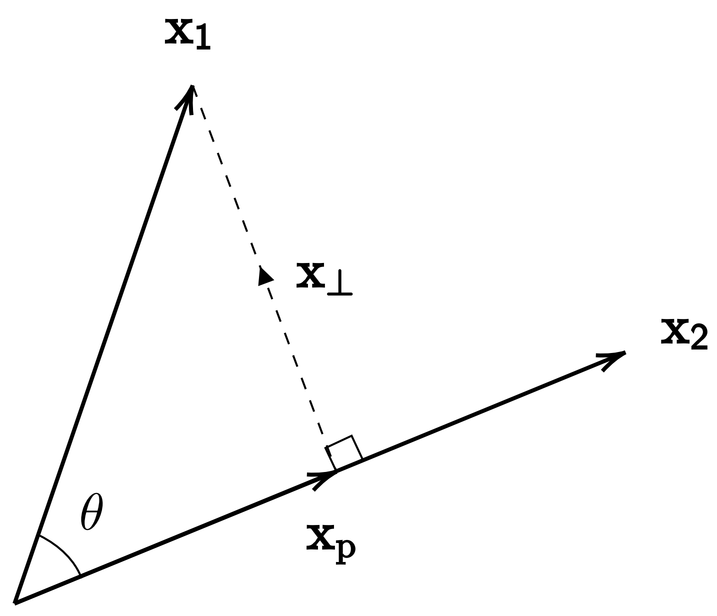
$\mathbb{R}^d$: Projections



$$\mathbf{x_p} = c \cdot \mathbf{x_2}$$

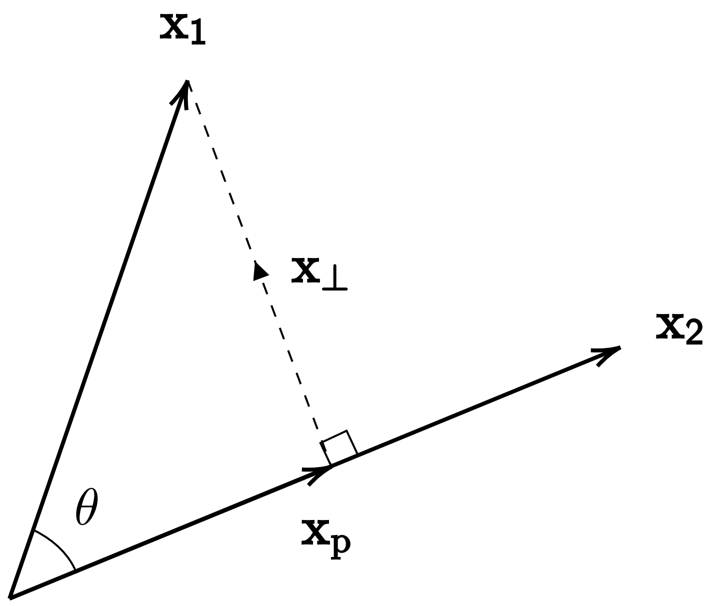
$\mathbb{R}^d$: Projections

$$\mathbf{x_1} = \mathbf{x_p} + \mathbf{x_\perp}$$



$$\mathbf{x_p} = c \cdot \mathbf{x_2}$$

$\mathbb{R}^d$: Projections

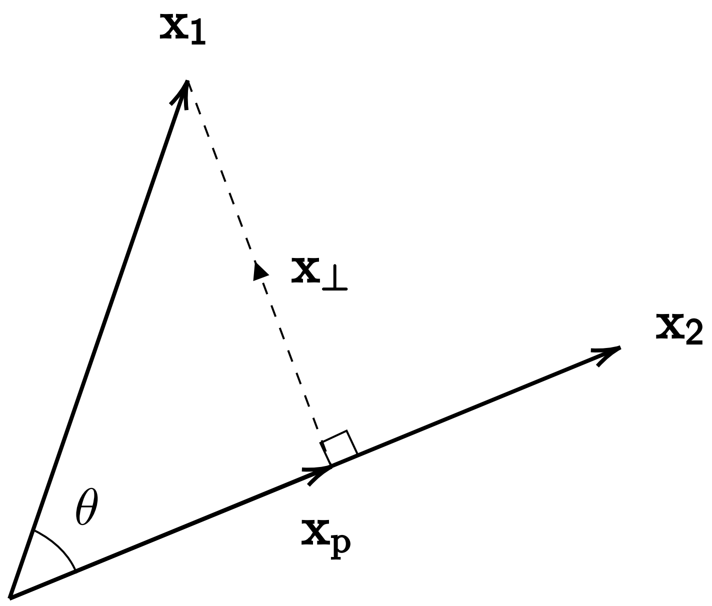


$$\mathbf{x_p} = c \cdot \mathbf{x_2}$$

$$\begin{aligned} \mathbf{x_1} &= \mathbf{x_p} + \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2} + \mathbf{x_\perp} \end{aligned}$$

$$\mathbf{x_2}^T \mathbf{x_1} = c \cdot \mathbf{x_2}^T \mathbf{x_2} + \mathbf{x_2}^T \mathbf{x_\perp}$$

$\mathbb{R}^d$: Projections

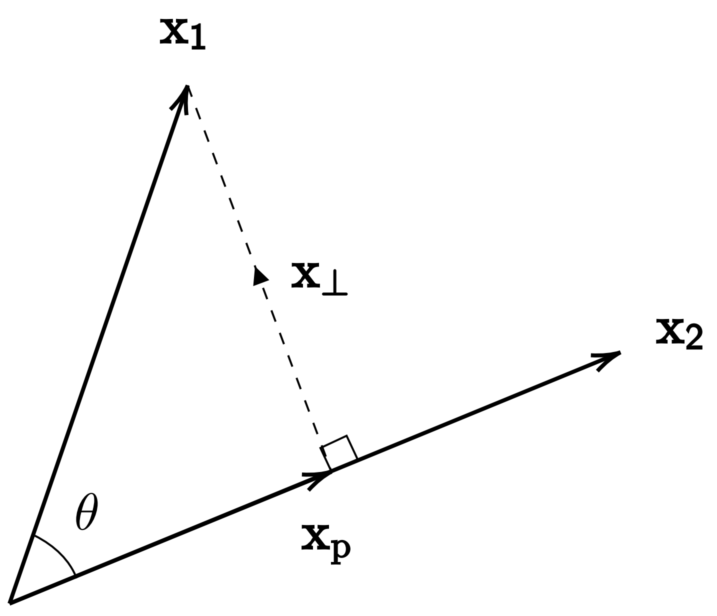


$$\mathbf{x_p} = c \cdot \mathbf{x_2}$$

$$\begin{aligned} \mathbf{x_1} &= \mathbf{x_p} + \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2} + \mathbf{x_\perp} \end{aligned}$$

$$\begin{aligned} \mathbf{x_2}^T \mathbf{x_1} &= c \cdot \mathbf{x_2}^T \mathbf{x_2} + \mathbf{x_2}^T \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2}^T \mathbf{x_2} \end{aligned}$$

$\mathbb{R}^d$: Projections



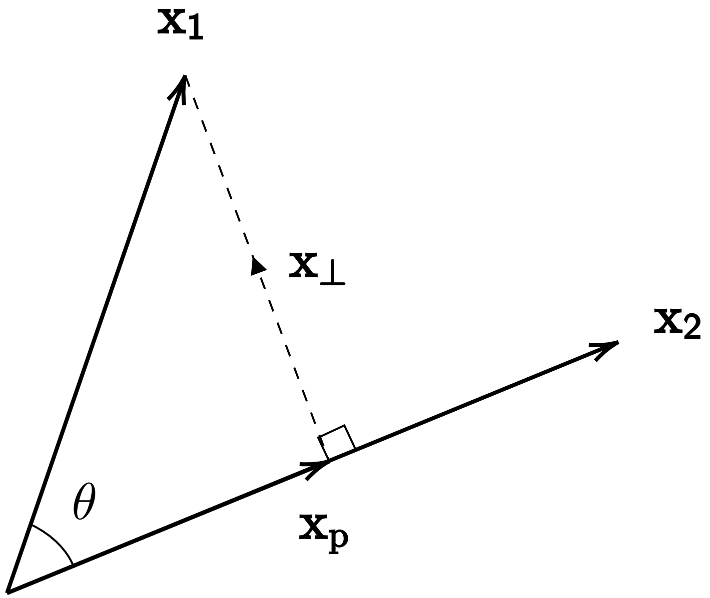
$$\mathbf{x_p} = c \cdot \mathbf{x_2}$$

$$\begin{aligned} \mathbf{x_1} &= \mathbf{x_p} + \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2} + \mathbf{x_\perp} \end{aligned}$$

$$\begin{aligned} \mathbf{x_2}^T \mathbf{x_1} &= c \cdot \mathbf{x_2}^T \mathbf{x_2} + \mathbf{x_2}^T \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2}^T \mathbf{x_2} \end{aligned}$$

$$c = \frac{\mathbf{x_2}^T \mathbf{x_1}}{\mathbf{x_2}^T \mathbf{x_2}}$$

$\mathbb{R}^d$: Projections



$$\mathbf{x_p} = c \cdot \mathbf{x_2}$$

$$\begin{aligned} \mathbf{x_1} &= \mathbf{x_p} + \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2} + \mathbf{x_\perp} \end{aligned}$$

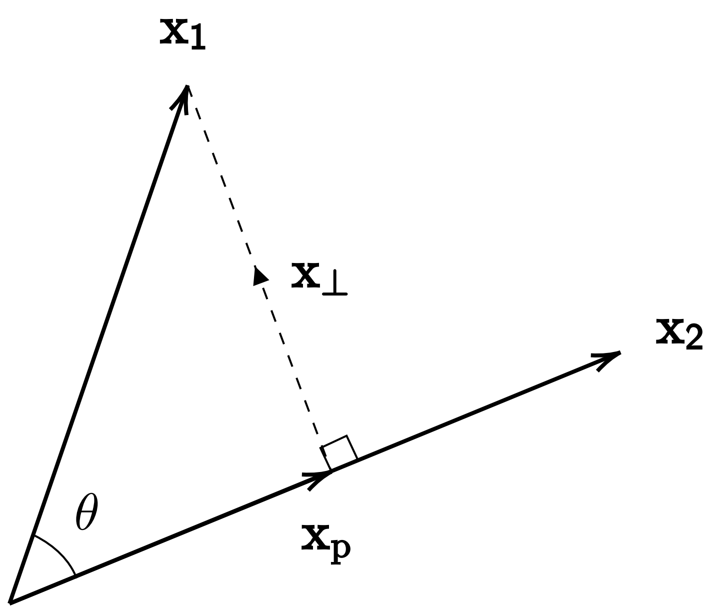
$$\begin{aligned} \mathbf{x_2}^T \mathbf{x_1} &= c \cdot \mathbf{x_2}^T \mathbf{x_2} + \mathbf{x_2}^T \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2}^T \mathbf{x_2} \end{aligned}$$

$$c = \frac{\mathbf{x_2}^T \mathbf{x_1}}{\mathbf{x_2}^T \mathbf{x_2}}$$

Vector Projection

$$\mathbf{x_p} = \frac{\mathbf{x_1}^T \mathbf{x_2}}{\mathbf{x_2}^T \mathbf{x_2}} \cdot \mathbf{x_2}$$

$\mathbb{R}^d$: Projections



$$\mathbf{x_p} = c \cdot \mathbf{x_2}$$

$$\begin{aligned} \mathbf{x_1} &= \mathbf{x_p} + \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2} + \mathbf{x_\perp} \end{aligned}$$

$$\begin{aligned} \mathbf{x_2}^T \mathbf{x_1} &= c \cdot \mathbf{x_2}^T \mathbf{x_2} + \mathbf{x_2}^T \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2}^T \mathbf{x_2} \end{aligned}$$

$$c = \frac{\mathbf{x_2}^T \mathbf{x_1}}{\mathbf{x_2}^T \mathbf{x_2}}$$

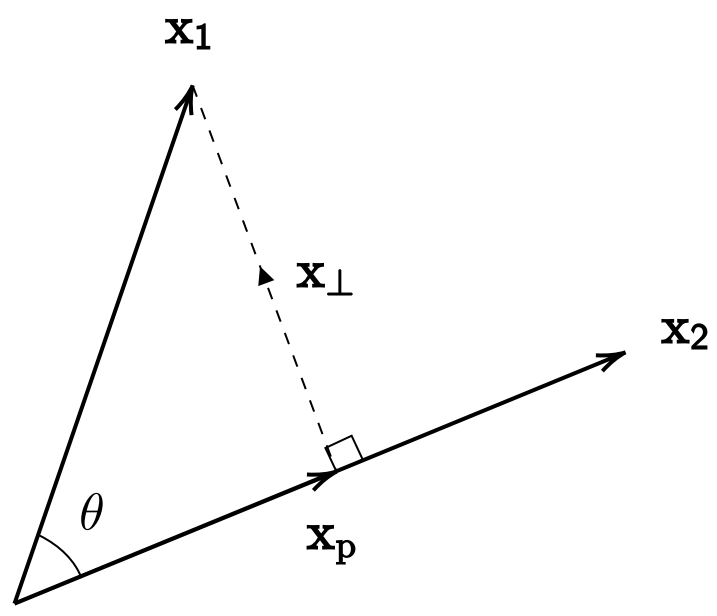
Vector Projection

$$\mathbf{x_p} = \frac{\mathbf{x_1}^T \mathbf{x_2}}{\mathbf{x_2}^T \mathbf{x_2}} \cdot \mathbf{x_2}$$

Scalar Projection

$$||\mathbf{x_p}|| = \frac{\mathbf{x_1}^T \mathbf{x_2}}{||\mathbf{x_2}||}$$

$\mathbb{R}^d$: Projections



$$\mathbf{x_p} = c \cdot \mathbf{x_2}$$

$$\begin{aligned} \mathbf{x_1} &= \mathbf{x_p} + \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2} + \mathbf{x_\perp} \end{aligned}$$

$$\begin{aligned} \mathbf{x_2}^T \mathbf{x_1} &= c \cdot \mathbf{x_2}^T \mathbf{x_2} + \mathbf{x_2}^T \mathbf{x_\perp} \\ &= c \cdot \mathbf{x_2}^T \mathbf{x_2} \end{aligned}$$

$$c = \frac{\mathbf{x_2}^T \mathbf{x_1}}{\mathbf{x_2}^T \mathbf{x_2}}$$

Vector Projection

$$\mathbf{x_p} = \frac{\mathbf{x_1}^T \mathbf{x_2}}{\mathbf{x_2}^T \mathbf{x_2}} \cdot \mathbf{x_2}$$

Scalar Projection

$$||\mathbf{x_p}|| = \frac{\mathbf{x_1}^T \mathbf{x_2}}{||\mathbf{x_2}||}$$

Projection on unit-norm vector

$$\left(\mathbf{x}^T \mathbf{w}\right) \mathbf{w}$$

$||\mathbf{w}|| = 1$

$\mathbb{R}^d$: Orthonormal Basis

$\mathbb{R}^d$: Orthonormal Basis

- Orthogonal
- Unit norm

$\mathbb{R}^d$: Orthonormal Basis

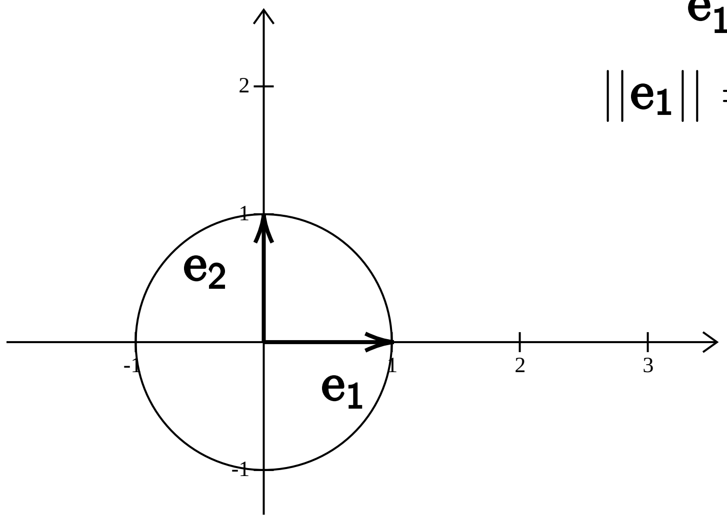
- Orthogonal
- Unit norm

Basis-1

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{e_1}^T \mathbf{e_2} = 0$$

$$||\mathbf{e_1}|| = ||\mathbf{e_2}|| = 1$$



$\mathbb{R}^d$: Orthonormal Basis

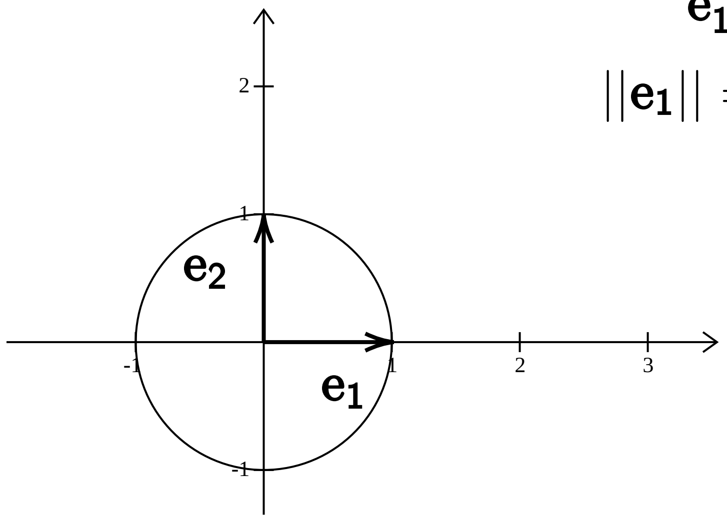
- Orthogonal
- Unit norm

Basis-1

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{e_1}^T \mathbf{e_2} = 0$$

$$||\mathbf{e_1}|| = ||\mathbf{e_2}|| = 1$$

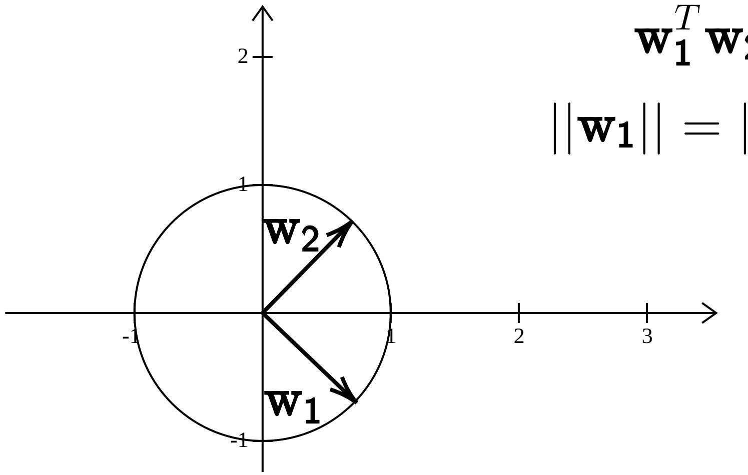


Basis-2

$$\mathbf{w_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix},\ \mathbf{w_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{w_1}^T \mathbf{w_2} = 0$$

$$||\mathbf{w_1}|| = ||\mathbf{w_2}|| = 1$$



$\mathbb{R}^d$: Orthonormal Basis

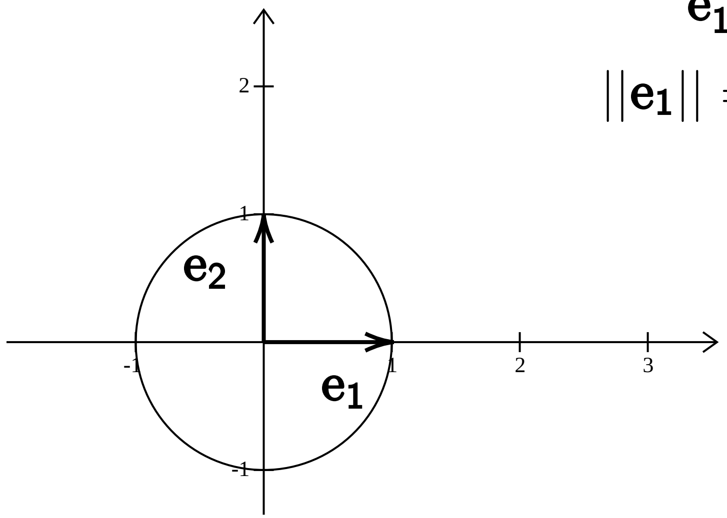
- Orthogonal
- Unit norm

Basis-1

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{e_1}^T \mathbf{e_2} = 0$$

$$||\mathbf{e_1}|| = ||\mathbf{e_2}|| = 1$$

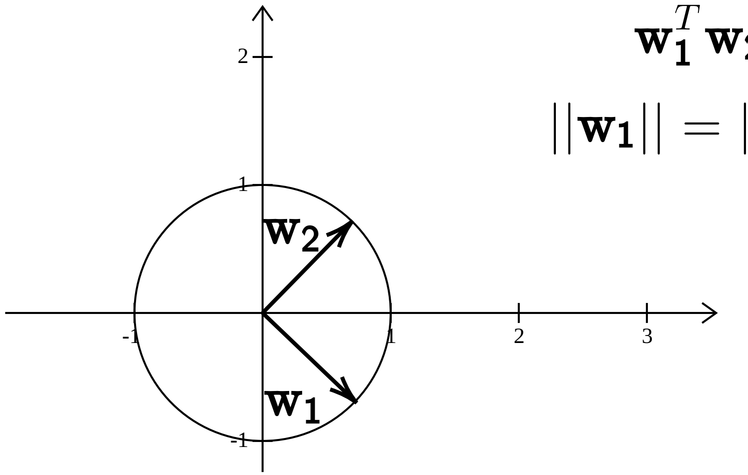


Basis-2

$$\mathbf{w_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix},\ \mathbf{w_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{w_1}^T \mathbf{w_2} = 0$$

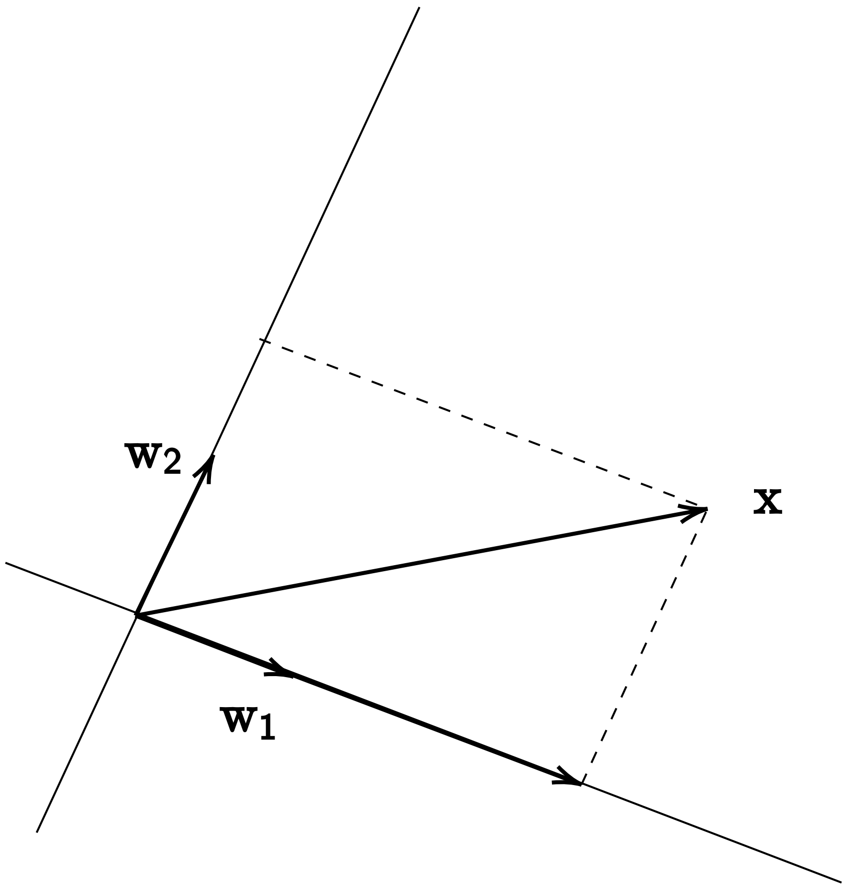
$$||\mathbf{w_1}|| = ||\mathbf{w_2}|| = 1$$



$$\text{Basis-1} \xleftrightarrow[\text{reflection}]{\text{rotation}} \text{Basis-2}$$

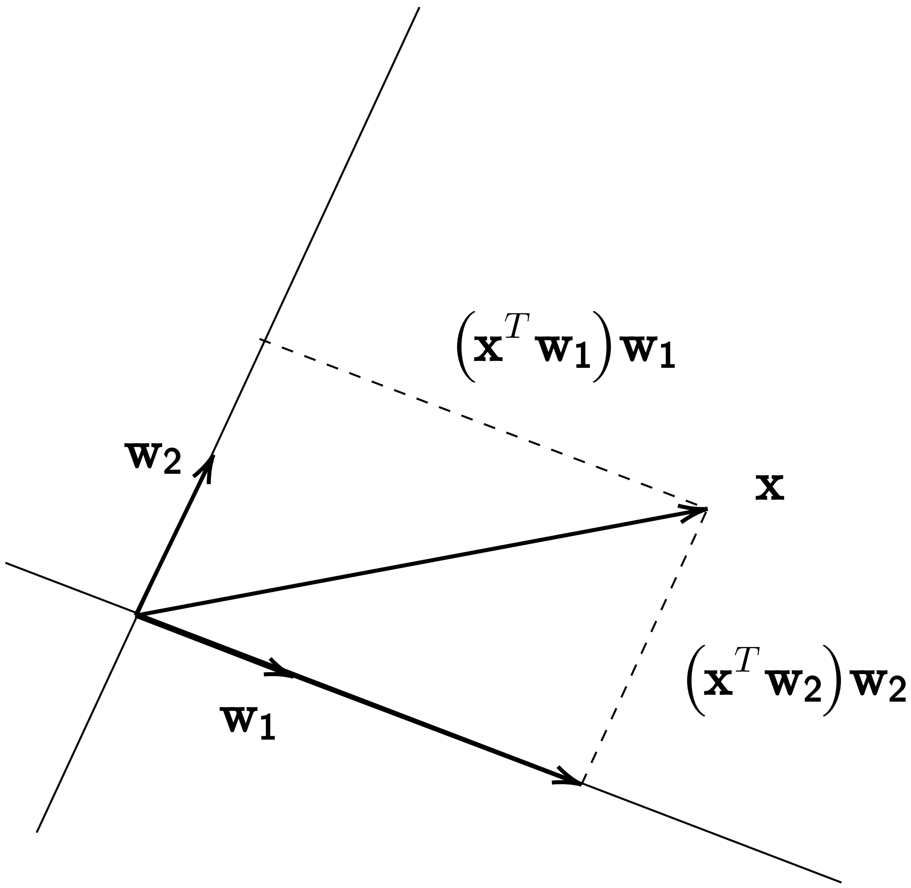
$\mathbb{R}^d$: Orthonormal Basis

- Orthogonal
- Unit norm



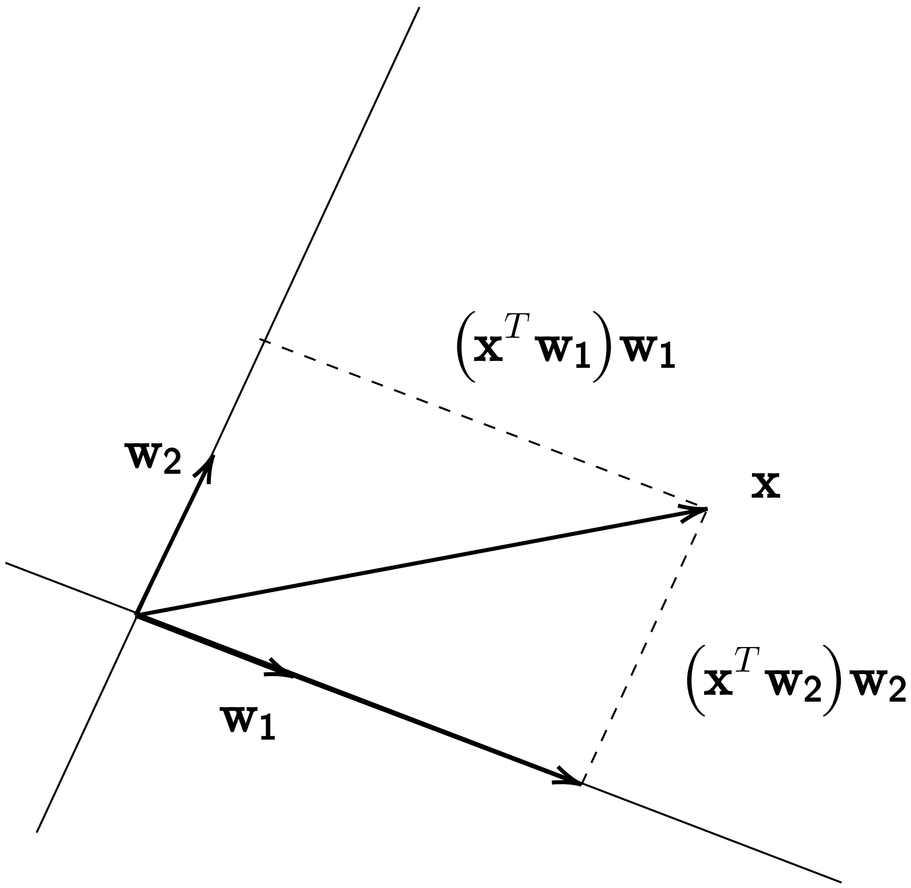
$\mathbb{R}^d$: Orthonormal Basis

- Orthogonal
- Unit norm



$\mathbb{R}^d$: Orthonormal Basis

- Orthogonal
- Unit norm



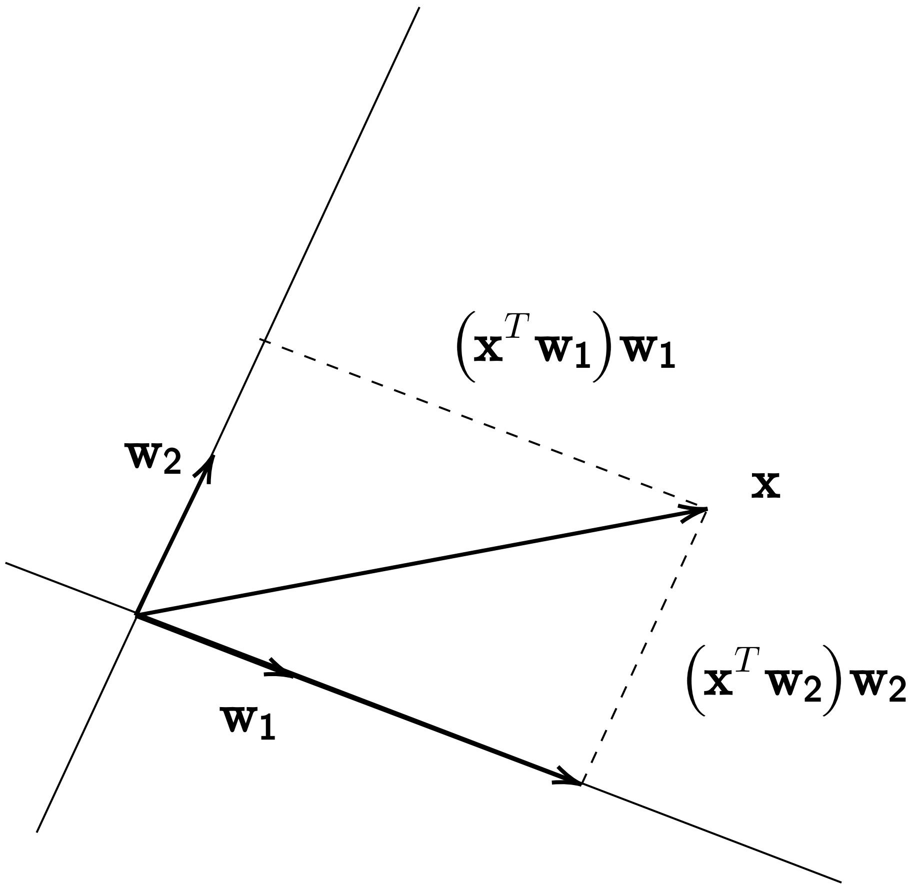
$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ 3\sqrt{2} \end{bmatrix}$$

$$\mathbf{w_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix},\ \mathbf{w_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$\mathbb{R}^d$: Orthonormal Basis

- Orthogonal
- Unit norm



$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

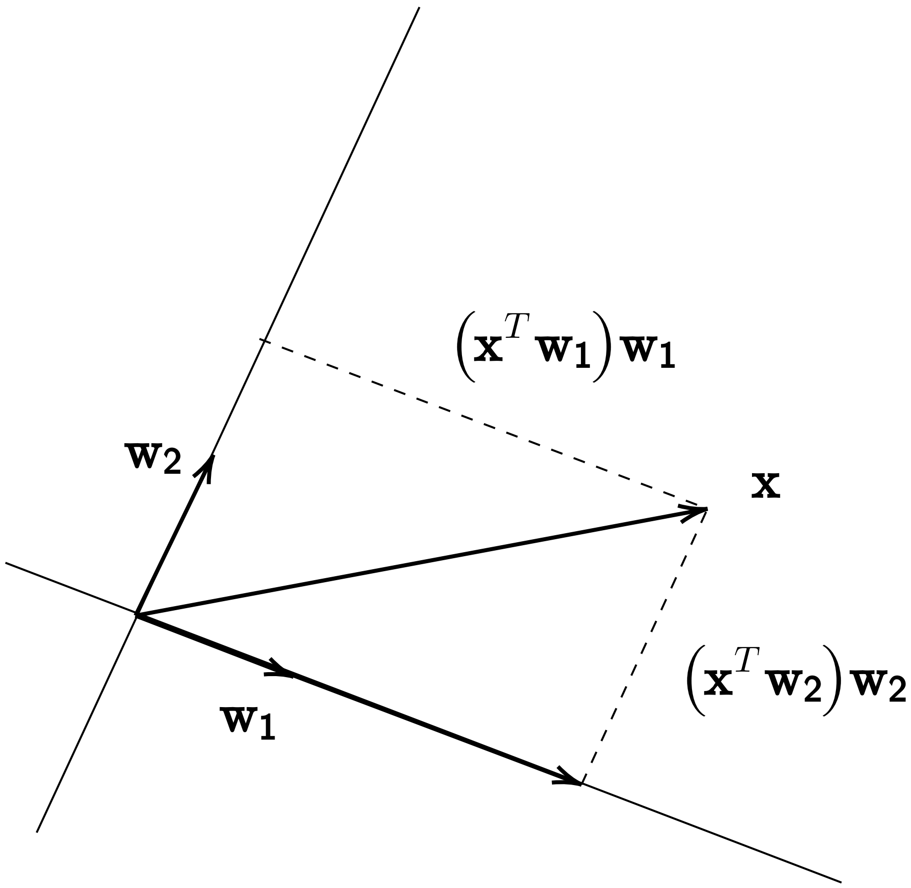
$$\mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ 3\sqrt{2} \end{bmatrix}$$

$$\mathbf{w_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix},\ \mathbf{w_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{x} =$$

$\mathbb{R}^d$: Orthonormal Basis

- Orthogonal
- Unit norm



$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

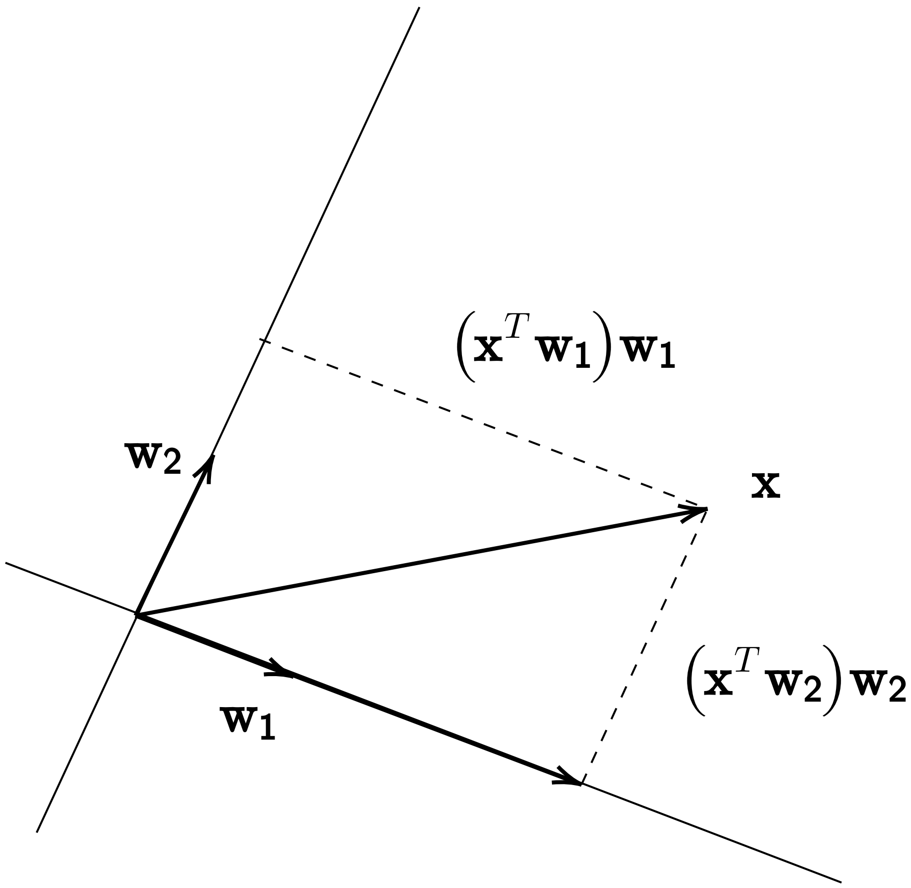
$$\mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ 3\sqrt{2} \end{bmatrix}$$

$$\mathbf{w_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix},\ \mathbf{w_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = 2\sqrt{2}\cdot\mathbf{e_1} + 3\sqrt{2}\cdot\mathbf{e_2}$$

$\mathbb{R}^d$: Orthonormal Basis

- Orthogonal
- Unit norm



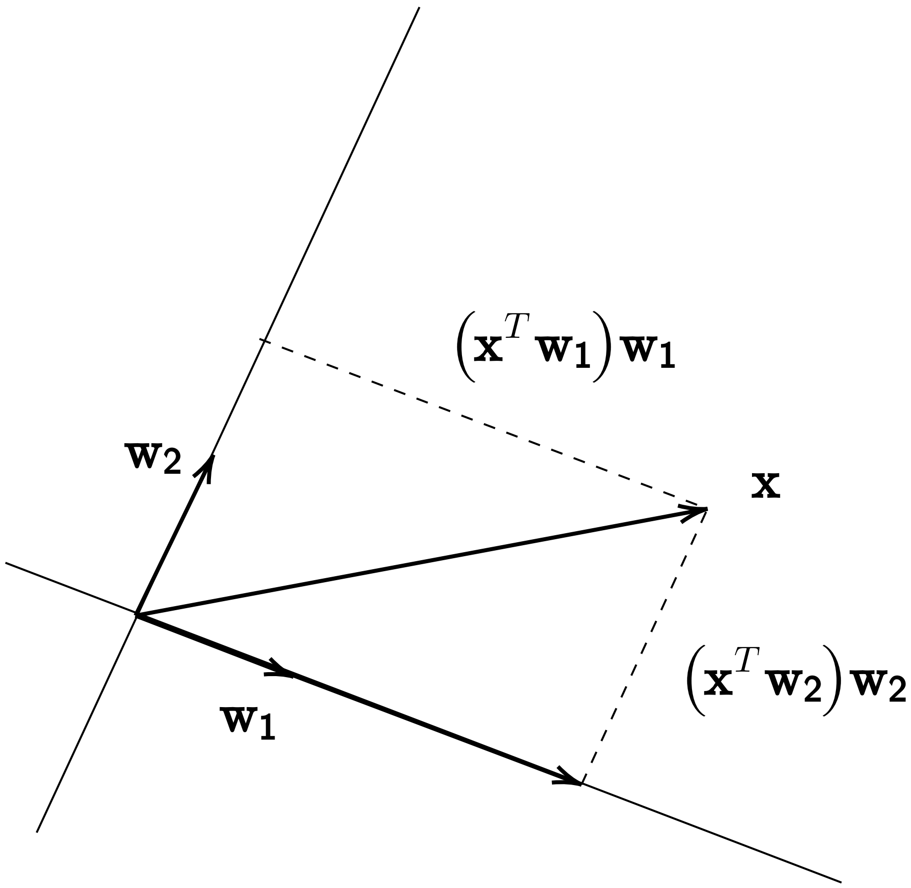
$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ 3\sqrt{2} \end{bmatrix}$$

$$\mathbf{w_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{w_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \mathbf{x} &= 2\sqrt{2} \cdot \mathbf{e_1} + 3\sqrt{2} \cdot \mathbf{e_2} \\ &= (\mathbf{x}^T\mathbf{w_1})\mathbf{w_1} + (\mathbf{x}^T\mathbf{w_2})\mathbf{w_2} \end{aligned}$$

$\mathbb{R}^d$: Orthonormal Basis

- Orthogonal
- Unit norm



$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ 3\sqrt{2} \end{bmatrix}$$

$$\mathbf{w_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{w_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \mathbf{x} &= 2\sqrt{2}\cdot\mathbf{e_1} + 3\sqrt{2}\cdot\mathbf{e_2} \\ &= \left(\mathbf{x}^T\mathbf{w_1}\right)\mathbf{w_1} + \left(\mathbf{x}^T\mathbf{w_2}\right)\mathbf{w_2} \\ &= -\mathbf{w_1} + 5\mathbf{w_2} \end{aligned}$$

Orthogonal Matrices

$$Q = \begin{bmatrix} | & | \\ \mathbf{w}_1 & \mathbf{w}_2 \\ | & | \end{bmatrix}$$

Orthogonal Matrices

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$

$$\mathbf{Q}^T\mathbf{Q} = \begin{bmatrix} - & \mathbf{w_1}^T & - \\ - & \mathbf{w_2}^T & - \end{bmatrix} \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$

Orthogonal Matrices

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$

$$\begin{aligned} \mathbf{Q}^T\mathbf{Q} &= \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix} \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{I} \end{aligned}$$

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

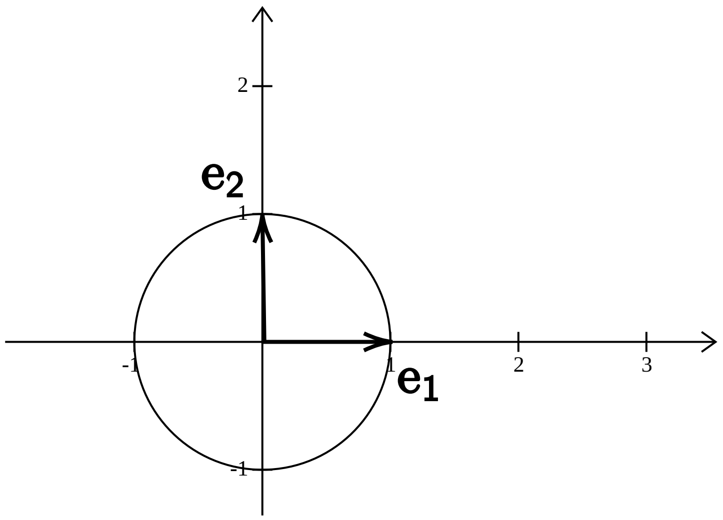
$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$

$$\begin{aligned} \mathbf{Q}^T\mathbf{Q} &= \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{I} \end{aligned}$$

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$



$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

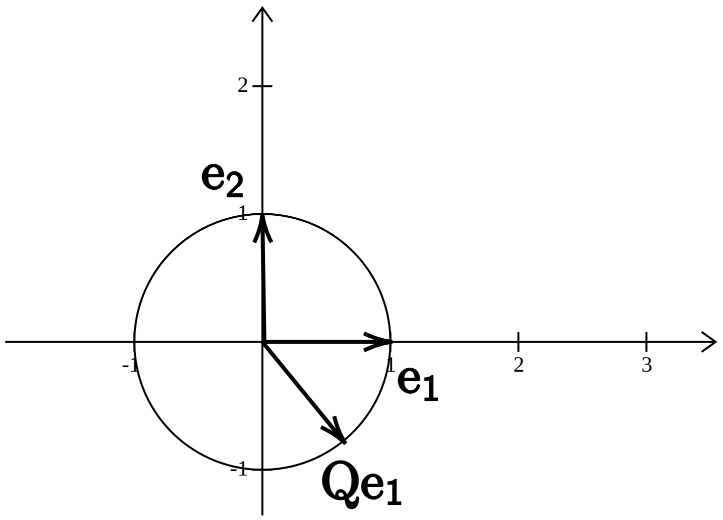
$$\begin{aligned} \mathbf{Q}^T\mathbf{Q} &= \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{I} \end{aligned}$$

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$

$$\mathbf{Q}^T\mathbf{Q} = \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \mathbf{I}$$



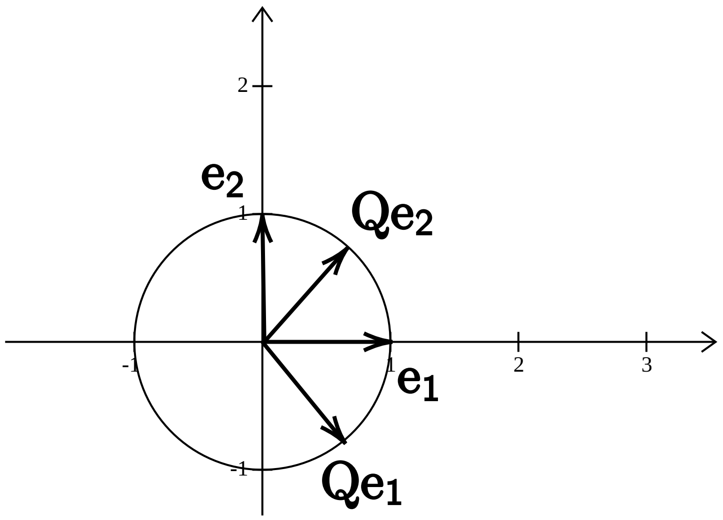
$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{Qe_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$



$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{Qe_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ \mathbf{Qe_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

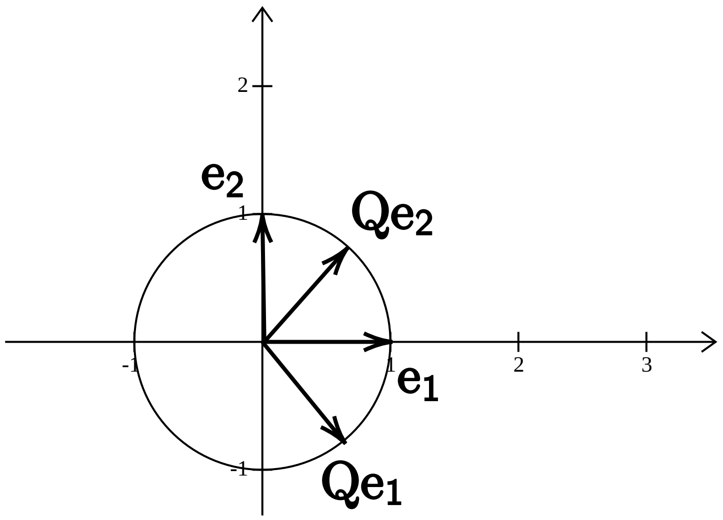
$$\begin{aligned}\mathbf{Q}^T\mathbf{Q} &= \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{I}\end{aligned}$$

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$

$$\mathbf{Q}^T\mathbf{Q} = \begin{bmatrix} - & \mathbf{w_1}^T & - \\ - & \mathbf{w_2}^T & - \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \mathbf{I}$$



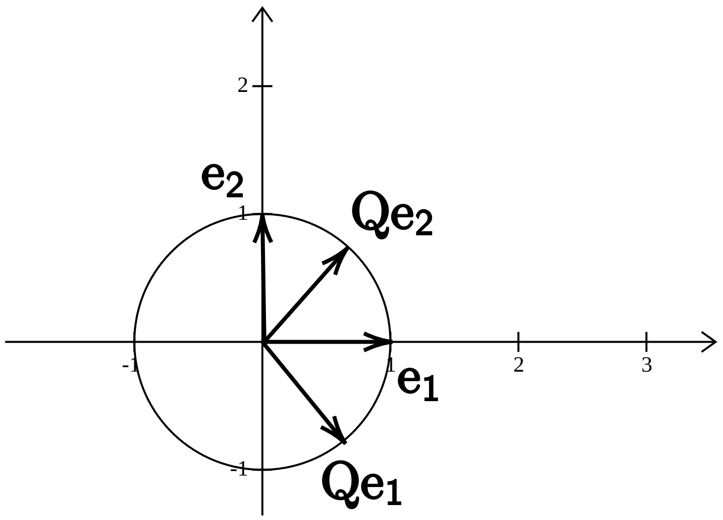
$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{Qe_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \mathbf{Qe_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$



$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{Qe_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ \mathbf{Qe_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \mathbf{Q}^T\mathbf{Q} &= \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{I} \end{aligned}$$

$$(\mathbf{Qx})^T(\mathbf{Qy})$$

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

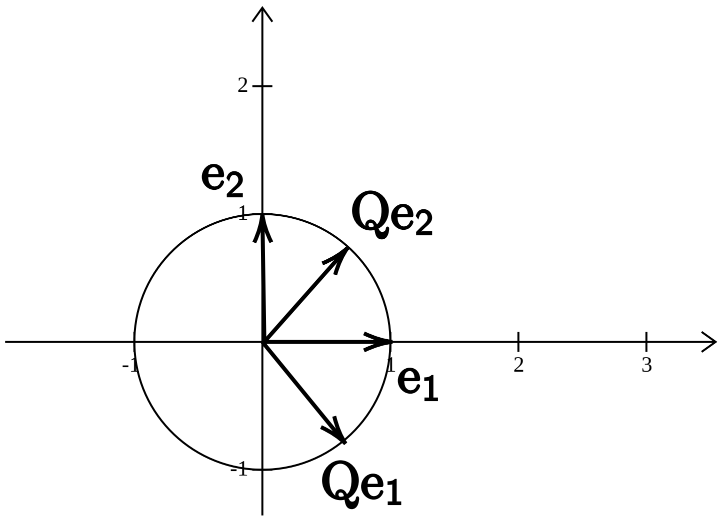
$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$

$$\mathbf{Q}^T\mathbf{Q} = \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \mathbf{I}$$

$$(\mathbf{Qx})^T(\mathbf{Qy}) = \mathbf{x}^T\mathbf{Q}^T\mathbf{Qy} = \mathbf{x}^T\mathbf{y}$$



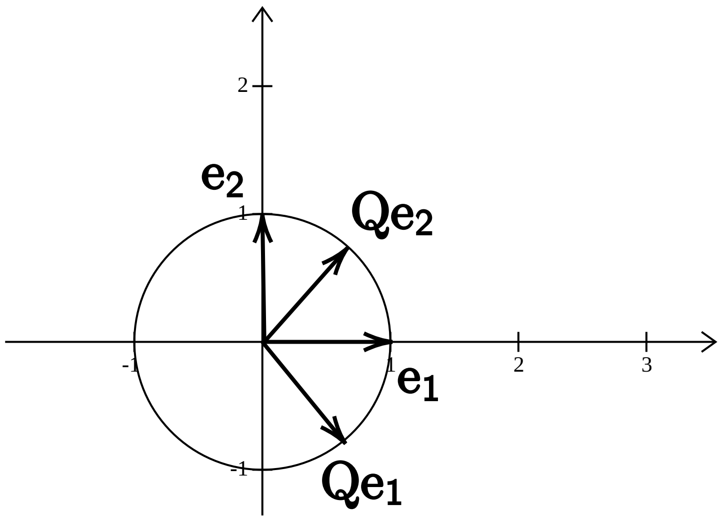
$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{Qe_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ \mathbf{Qe_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$



$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; \mathbf{Qe_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; \mathbf{Qe_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \mathbf{Q}^T\mathbf{Q} &= \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{I} \end{aligned}$$

$$(\mathbf{Qx})^T(\mathbf{Qy}) = \mathbf{x}^T\mathbf{Q}^T\mathbf{Qy} = \mathbf{x}^T\mathbf{y}$$

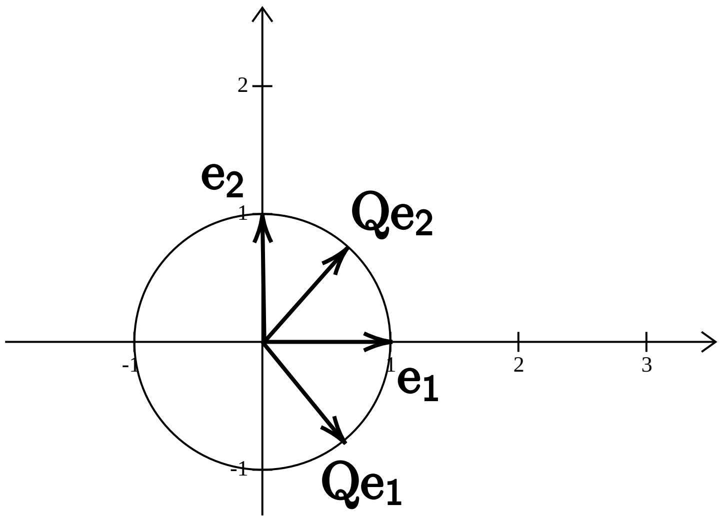
Orthogonal matrix

- Preserves inner products

Orthogonal Matrices

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix}$$



$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{Qe_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \mathbf{Qe_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \mathbf{Q}^T\mathbf{Q} &= \begin{bmatrix} — & \mathbf{w_1}^T & — \\ — & \mathbf{w_2}^T & — \end{bmatrix}\begin{bmatrix} | & | \\ \mathbf{w_1} & \mathbf{w_2} \\ | & | \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{I} \end{aligned}$$

$$(\mathbf{Qx})^T(\mathbf{Qy}) = \mathbf{x}^T\mathbf{Q}^T\mathbf{Qy} = \mathbf{x}^T\mathbf{y}$$

Orthogonal matrix

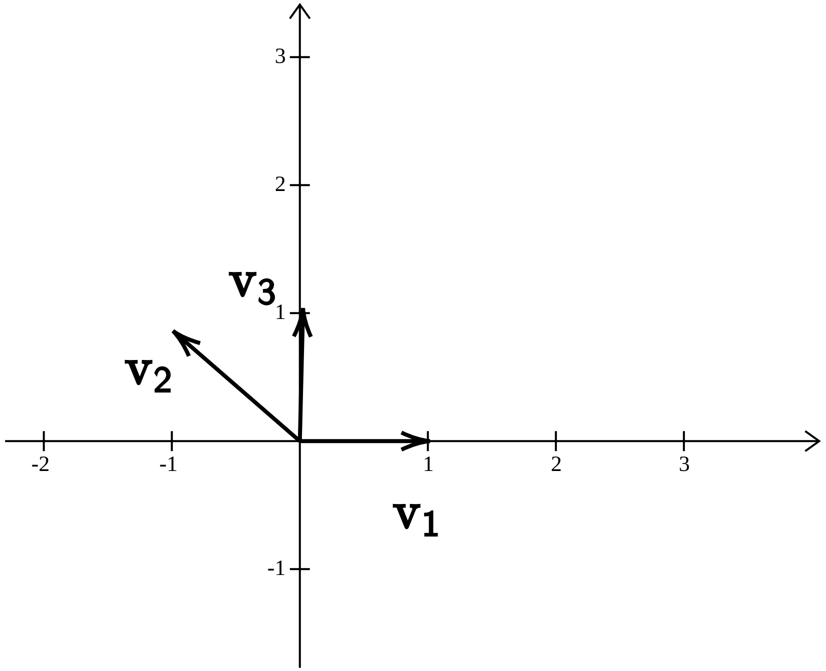
- Preserves inner products
  - Preserves lengths
  - Preserves angles

Eigenvalues and Eigenvectors

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

Eigenvalues and Eigenvectors

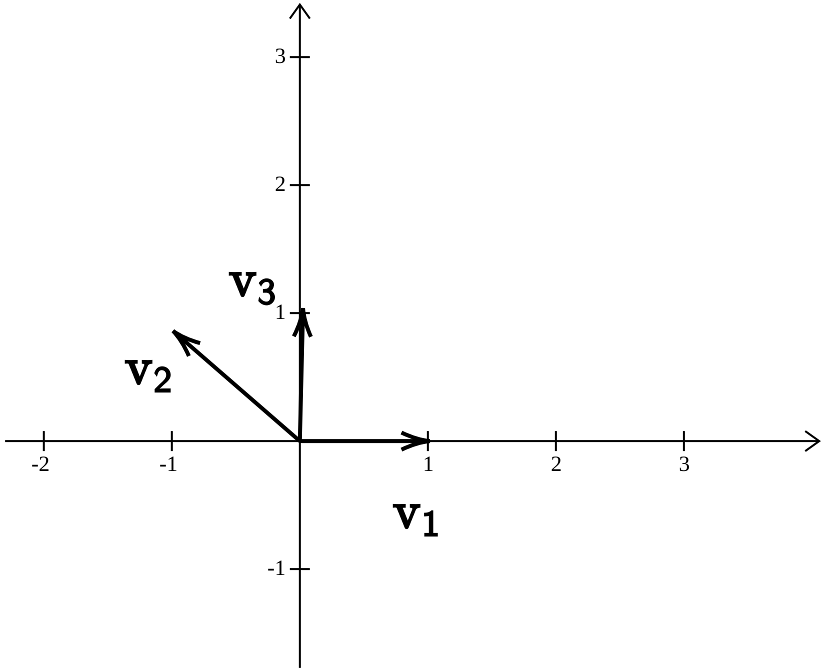
$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$



$$\mathbf{A}\mathbf{v_3} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Eigenvalues and Eigenvectors

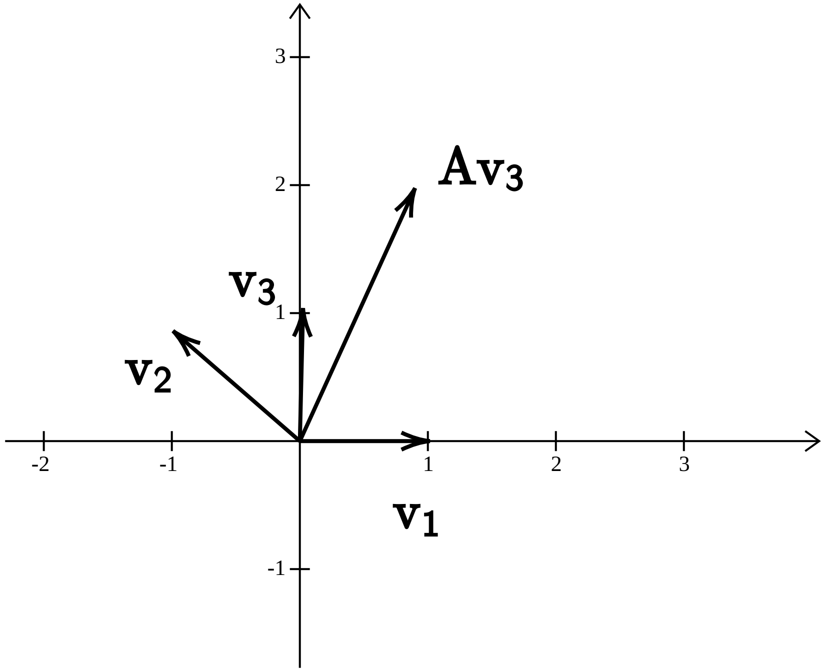
$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$



$$\begin{aligned} \mathbf{A}\mathbf{v}_3 &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$

Eigenvalues and Eigenvectors

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

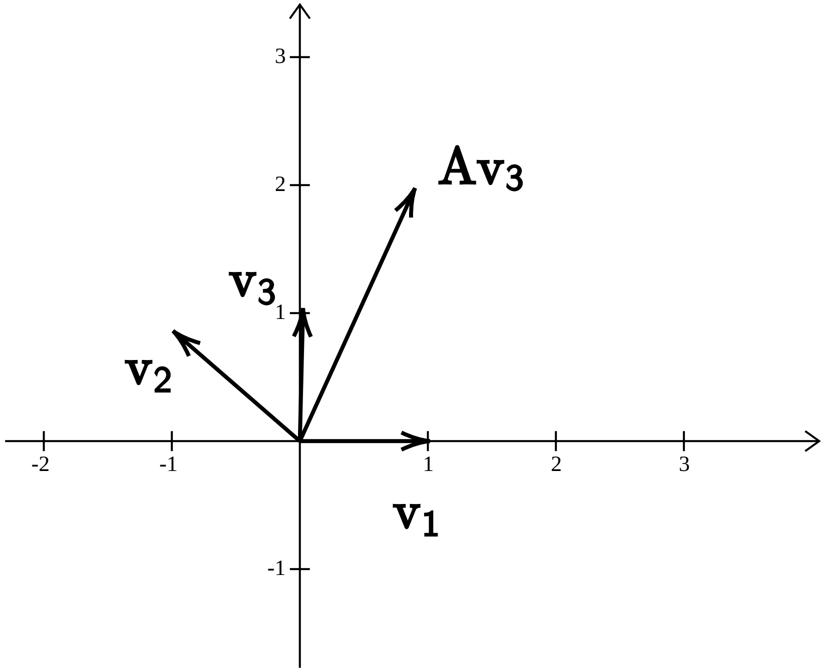


$$\mathbf{Av_3} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Eigenvalues and Eigenvectors

$$\mathbf{Av_1} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

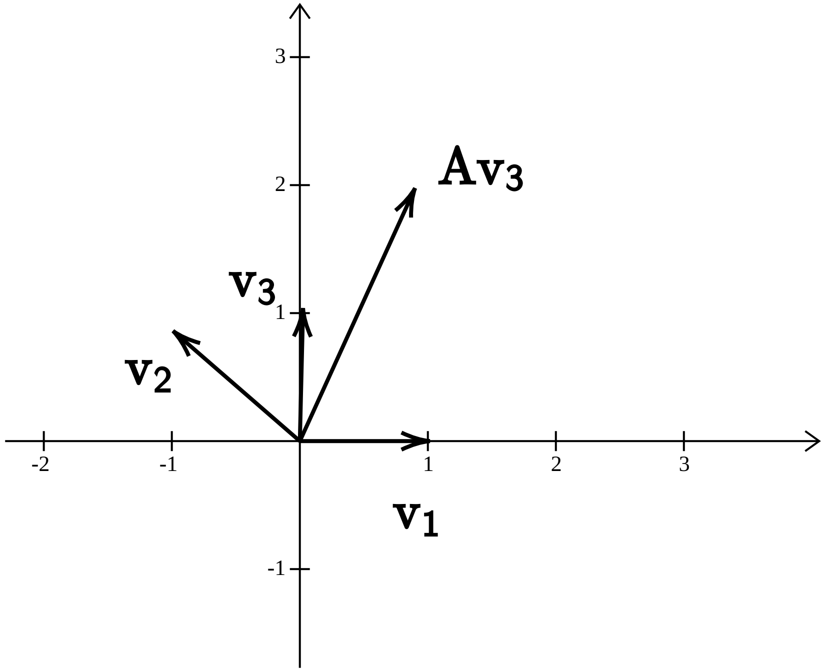
$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$



$$\mathbf{Av_3} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Eigenvalues and Eigenvectors

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

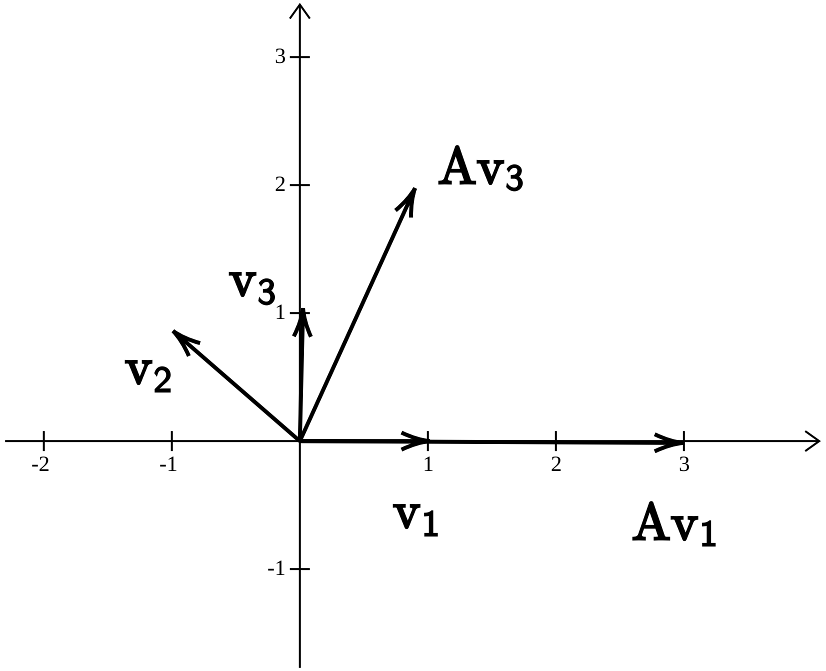


$$\mathbf{Av_1} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$$\mathbf{Av_3} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Eigenvalues and Eigenvectors

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

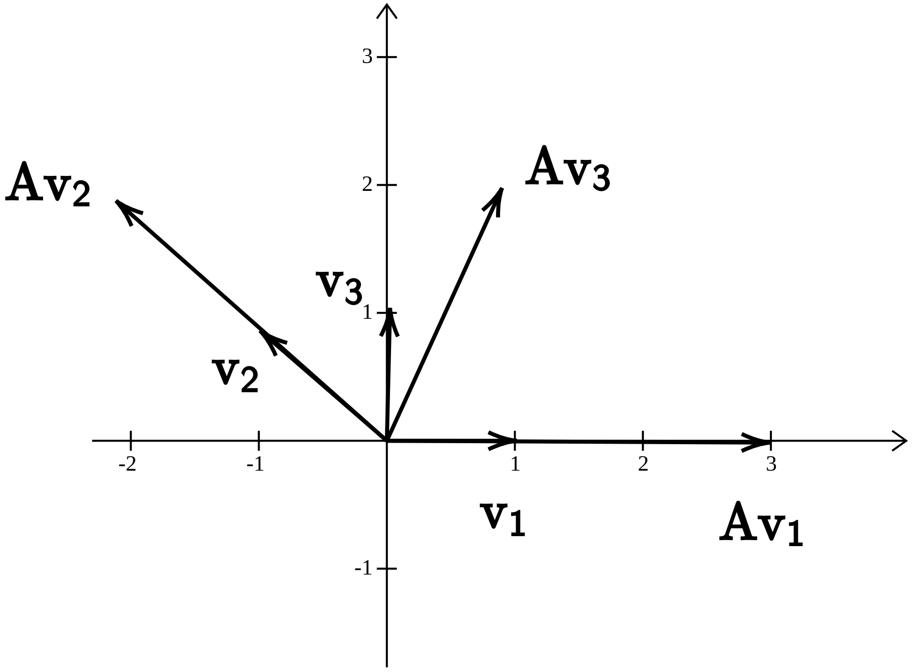


$$\begin{aligned} \mathbf{Av_1} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 0 \end{bmatrix} \\ &= 3 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= 3 \cdot \mathbf{v_1} \end{aligned}$$

$$\begin{aligned} \mathbf{Av_3} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$

Eigenvalues and Eigenvectors

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$



$$\begin{aligned} \mathbf{Av_1} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 0 \end{bmatrix} \\ &= 3 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= 3 \cdot \mathbf{v_1} \end{aligned}$$

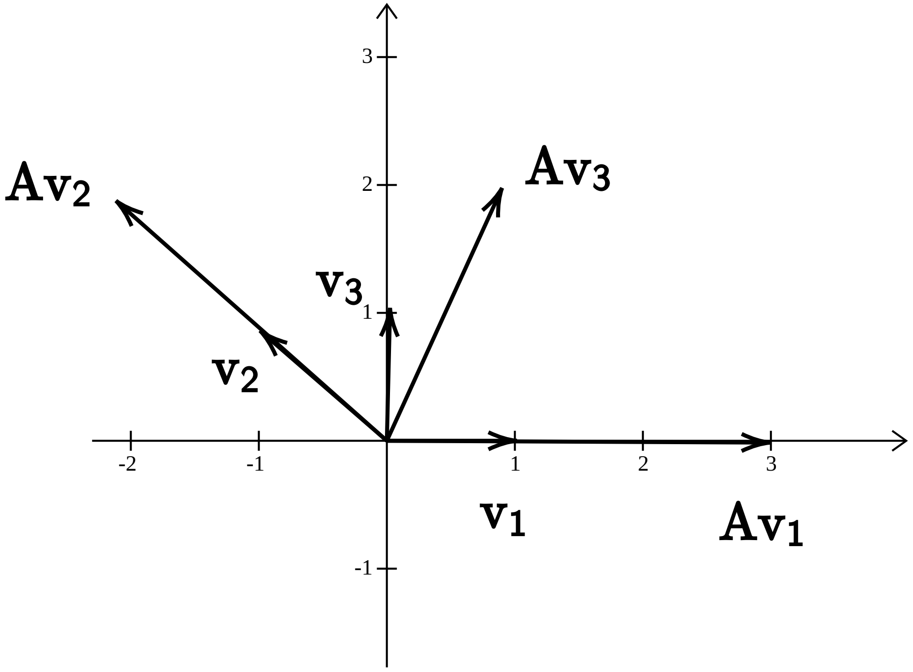
$$\begin{aligned} \mathbf{Av_2} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ 2 \end{bmatrix} \\ &= 2 \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= 2 \cdot \mathbf{v_2} \end{aligned}$$

$$\begin{aligned} \mathbf{Av_3} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$

Eigenvalues and Eigenvectors

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

$$\begin{aligned} \mathbf{A}\mathbf{v_2} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ 2 \end{bmatrix} \\ &= 2 \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= 2 \cdot \mathbf{v_2} \end{aligned}$$



$$\begin{aligned} \mathbf{A}\mathbf{v_1} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 0 \end{bmatrix} \\ &= 3 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= 3 \cdot \mathbf{v_1} \end{aligned}$$

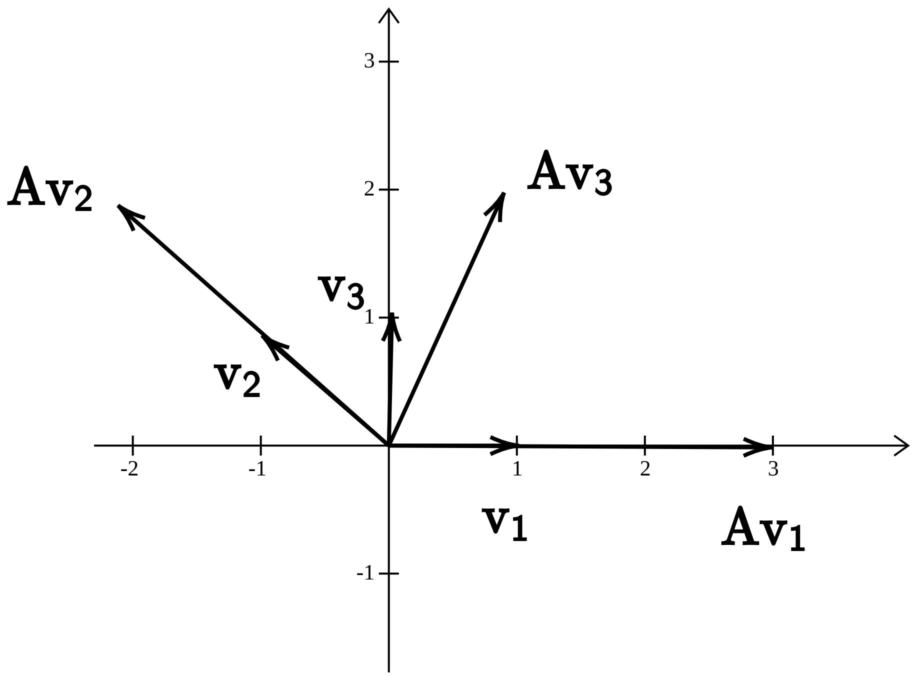
$$\begin{aligned} \mathbf{A}\mathbf{v_3} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &\neq \lambda \cdot \mathbf{v_3} \end{aligned}$$

for any $\lambda \in \mathbb{R}$

Eigenvalues and Eigenvectors

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

$$\begin{aligned}\mathbf{Av_1} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 0 \end{bmatrix} \\ &= 3 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= 3 \cdot \mathbf{v_1}\end{aligned}$$



$$\begin{aligned}\mathbf{Av_2} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ 2 \end{bmatrix} \\ &= 2 \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= 2 \cdot \mathbf{v_2}\end{aligned}$$

$$\begin{aligned}\mathbf{Av_3} &= \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &\neq \lambda \cdot \mathbf{v_3}\end{aligned}$$

for any $\lambda \in \mathbb{R}$

Definition: For a $d \times d$ square matrix $\mathbf{A}$, a non-zero vector $\mathbf{v} \in \mathbb{R}^d$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ if $\mathbf{Av} = \lambda\mathbf{v}$.

Matrix Multiplication Revisited

$$\mathbf{A} = \begin{bmatrix} | & & | \\ \mathbf{a_1} & \cdots & \mathbf{a_n} \\ | & & | \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} — & \mathbf{b_1}^T & — \\ & \vdots & \\ — & \mathbf{b_n}^T & — \end{bmatrix} \qquad \mathbf{AB} =$$

$$m \times n \qquad\qquad n \times p \qquad\qquad m \times p$$

Matrix Multiplication Revisited

$$\mathbf{A} = \begin{bmatrix} | & & | \\ \mathbf{a_1} & \cdots & \mathbf{a_n} \\ | & & | \end{bmatrix}$$

$$m \times n$$

$$\mathbf{B} = \begin{bmatrix} — & \mathbf{b_1}^T & — \\ & \vdots & \\ — & \mathbf{b_n}^T & — \end{bmatrix}$$

$$n \times p$$

$$\mathbf{AB} = \sum_{i=1}^{n} \mathbf{a_i} \mathbf{b_i}^T$$

$$m \times p$$

Matrix Multiplication Revisited

$$\mathbf{A} = \begin{bmatrix} | & & | \\ \mathbf{a_1} & \cdots & \mathbf{a_n} \\ | & & | \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} — & \mathbf{b_1}^T & — \\ & \vdots & \\ — & \mathbf{b_n}^T & — \end{bmatrix} \qquad \mathbf{AB} = \sum_{i=1}^{n} \mathbf{a_i} \mathbf{b_i}^T$$

$$m \times n \qquad n \times p \qquad m \times p$$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} d & e & f & g \end{bmatrix} =$$

Matrix Multiplication Revisited

$$\mathbf{A} = \begin{bmatrix} | & & | \\ \mathbf{a_1} & \cdots & \mathbf{a_n} \\ | & & | \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} — & \mathbf{b_1}^T & — \\ & \vdots & \\ — & \mathbf{b_n}^T & — \end{bmatrix} \qquad \mathbf{AB} = \sum_{i=1}^{n} \mathbf{a_i} \mathbf{b_i}^T$$

$$m \times n \qquad n \times p \qquad m \times p$$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} d & e & f & g \end{bmatrix} = \begin{bmatrix} d \cdot a & e \cdot a & f \cdot a & g \cdot a \\ d \cdot b & e \cdot b & f \cdot b & g \cdot b \\ d \cdot c & e \cdot c & f \cdot c & g \cdot c \end{bmatrix}$$

Spectral Theorem

Theorem: If $\mathbf{A}$ is a square *symmetric matrix* of shape $d \times d$, then:

Spectral Theorem

Theorem: If $\mathbf{A}$ is a square *symmetric matrix* of shape $d \times d$, then:

- All the eigenvalues of $\mathbf{A}$ are real

Spectral Theorem

Theorem: If $\mathbf{A}$ is a square *symmetric matrix* of shape $d \times d$, then:

- All the eigenvalues of $\mathbf{A}$ are real
- $\mathbb{R}^d$ has an orthonormal basis of eigenvectors of $\mathbf{A}$

Spectral Theorem

Theorem: If $\mathbf{A}$ is a square *symmetric matrix* of shape $d \times d$, then:

- All the eigenvalues of $\mathbf{A}$ are real
- $\mathbb{R}^d$ has an orthonormal basis of eigenvectors of $\mathbf{A}$
- $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$
  - the columns of $\mathbf{Q}$ are the eigenvectors of $\mathbf{A}$
  - $\mathbf{D}$ is a diagonal matrix of the corresponding eigenvalues.

Spectral Theorem

Theorem: If $\mathbf{A}$ is a square *symmetric matrix* of shape $d \times d$, then:

- All the eigenvalues of $\mathbf{A}$ are real
- $\mathbb{R}^d$ has an orthonormal basis of eigenvectors of $\mathbf{A}$
- $\mathbf{A} = \mathbf{QDQ}^T$
  - the columns of $\mathbf{Q}$ are the eigenvectors of $\mathbf{A}$
  - $\mathbf{D}$ is a diagonal matrix of the corresponding eigenvalues.

$$\mathbf{A} = \mathbf{QDQ}^T$$

$$= \begin{bmatrix} | & & | \\ \mathbf{w_1} & \cdots & \mathbf{w_d} \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_d \end{bmatrix} \begin{bmatrix} — & \mathbf{w_1}^T & — \\ & \vdots & \\ — & \mathbf{w_d}^T & — \end{bmatrix}$$

$$= \sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T$$

Positive Definiteness

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Positive Definiteness

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d}\sum_{j=1}^{d} A_{ij}x_ix_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Positive Definiteness

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d}\sum_{j=1}^{d} A_{ij}x_ix_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Test: A symmetric matrix is positive semi-definite if and only if all its eigenvalues are non-negative.

Positive Definiteness

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d}\sum_{j=1}^{d} A_{ij}x_i x_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Test: A symmetric matrix is positive semi-definite if and only if all its eigenvalues are non-negative.

$$\mathbf{v_i}^T\mathbf{A}\mathbf{v_i} =$$

Positive Definiteness

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{i=1}^{d} \sum_{j=1}^{d} A_{ij} x_i x_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T \mathbf{A} \mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Test: A symmetric matrix is positive semi-definite if and only if all its eigenvalues are non-negative.

$$\mathbf{v_i}^T \mathbf{A} \mathbf{v_i} = \mathbf{v_i}^T (\lambda_i \mathbf{v_i})$$

Positive Definiteness

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d}\sum_{j=1}^{d} A_{ij}x_ix_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Test: A symmetric matrix is positive semi-definite if and only if all its eigenvalues are non-negative.

$$\begin{aligned}\mathbf{v_i}^T\mathbf{A}\mathbf{v_i} &= \mathbf{v_i}^T(\lambda_i\mathbf{v_i}) \\ &= \lambda_i\left(\mathbf{v_i}^T\mathbf{v_i}\right) \\ &= \lambda_i \cdot ||\mathbf{v_i}||^2\end{aligned}$$

Positive Definiteness

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d}\sum_{j=1}^{d} A_{ij}x_ix_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Test: A symmetric matrix is positive semi-definite if and only if all its eigenvalues are non-negative.

$$\begin{aligned}\mathbf{v_i}^T\mathbf{A}\mathbf{v_i} &= \mathbf{v_i}^T(\lambda_i\mathbf{v_i}) \\ &= \lambda_i\left(\mathbf{v_i}^T\mathbf{v_i}\right) \\ &= \lambda_i \cdot ||\mathbf{v_i}||^2\end{aligned}$$

$$\mathbf{x} = \sum_{i=1}^{d} \alpha_i \mathbf{v_i}$$

Positive Definiteness

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d}\sum_{j=1}^{d} A_{ij}x_i x_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Test: A symmetric matrix is positive semi-definite if and only if all its eigenvalues are non-negative.

$$\begin{aligned}\mathbf{v_i}^T\mathbf{A}\mathbf{v_i} &= \mathbf{v_i}^T(\lambda_i\mathbf{v_i})\\ &= \lambda_i\left(\mathbf{v_i}^T\mathbf{v_i}\right)\\ &= \lambda_i \cdot ||\mathbf{v_i}||^2\end{aligned}$$

$$\mathbf{x} = \sum_{i=1}^{d}\alpha_i\mathbf{v_i}$$

$$\mathbf{A} = \sum_{i=1}^{d}\lambda_i\mathbf{v_i}\mathbf{v_i}^T$$

Positive Definiteness

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d}\sum_{j=1}^{d} A_{ij}x_ix_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Test: A symmetric matrix is positive semi-definite if and only if all its eigenvalues are non-negative.

$$\begin{aligned}\mathbf{v_i}^T\mathbf{A}\mathbf{v_i} &= \mathbf{v_i}^T(\lambda_i\mathbf{v_i}) \\ &= \lambda_i\left(\mathbf{v_i}^T\mathbf{v_i}\right) \\ &= \lambda_i \cdot ||\mathbf{v_i}||^2\end{aligned}$$

$$\mathbf{x} = \sum_{i=1}^{d}\alpha_i\mathbf{v_i}$$

$$\mathbf{x}^T\mathbf{A}\mathbf{x} =$$

$$\mathbf{A} = \sum_{i=1}^{d}\lambda_i\mathbf{v_i}\mathbf{v_i}^T$$

Positive Definiteness

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d}\sum_{j=1}^{d} A_{ij}x_ix_j$$

Definition: A symmetric matrix $\mathbf{A}$ of shape $d \times d$ is positive semi-definite if for every non-zero vector $\mathbf{x}$, $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$. If the inequality is strict, then the matrix is termed positive definite.

Test: A symmetric matrix is positive semi-definite if and only if all its eigenvalues are non-negative.

$$\begin{aligned}\mathbf{v_i}^T\mathbf{A}\mathbf{v_i} &= \mathbf{v_i}^T(\lambda_i\mathbf{v_i}) \\ &= \lambda_i\left(\mathbf{v_i}^T\mathbf{v_i}\right) \\ &= \lambda_i \cdot ||\mathbf{v_i}||^2\end{aligned}$$

$$\mathbf{x} = \sum_{i=1}^{d} \alpha_i\mathbf{v_i}$$

$$\mathbf{A} = \sum_{i=1}^{d} \lambda_i\mathbf{v_i}\mathbf{v_i}^T$$

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{d} \alpha_i^2\lambda_i$$

Covariance Matrix

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}, \ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

Covariance Matrix

$$D = \{\mathbf{x_1},\ \cdots, \mathbf{x_n}\},\ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

Covariance Matrix

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}, \ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathbf{x_i} - \boldsymbol{\mu})(\mathbf{x_i} - \boldsymbol{\mu})^T$$

Covariance Matrix

$$D = \{\mathbf{x_1},\ \cdots, \mathbf{x_n}\},\ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathbf{x_i} - \boldsymbol{\mu})(\mathbf{x_i} - \boldsymbol{\mu})^T$$

$$d \times d$$

Covariance Matrix

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}, \ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathbf{x_i} - \boldsymbol{\mu})(\mathbf{x_i} - \boldsymbol{\mu})^T$$

$$d \times d$$

$$C_{pq} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)(x_{qi} - \mu_q)$$

Covariance Matrix

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}, \; \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathbf{x_i} - \boldsymbol{\mu})(\mathbf{x_i} - \boldsymbol{\mu})^T$$

$$d \times d$$

$$C_{pq} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)(x_{qi} - \mu_q)$$

$$C_{pp} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)^2 = \sigma_p^2$$

Covariance Matrix

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}, \ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} \texttt{Var}(x_1) & \texttt{Cov}(x_1, x_2) & \texttt{Cov}(x_1, x_3) \\ \texttt{Cov}(x_2, x_1) & \texttt{Var}(x_2) & \texttt{Cov}(x_2, x_3) \\ \texttt{Cov}(x_3, x_1) & \texttt{Cov}(x_3, x_2) & \texttt{Var}(x_3) \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathbf{x_i} - \boldsymbol{\mu})(\mathbf{x_i} - \boldsymbol{\mu})^T$$

$$d \times d$$

$$C_{pq} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)(x_{qi} - \mu_q)$$

$$C_{pp} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)^2 = \sigma_p^2$$

Covariance Matrix

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}, \ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} \texttt{Var}(x_1) & \texttt{Cov}(x_1, x_2) & \texttt{Cov}(x_1, x_3) \\ \texttt{Cov}(x_2, x_1) & \texttt{Var}(x_2) & \texttt{Cov}(x_2, x_3) \\ \texttt{Cov}(x_3, x_1) & \texttt{Cov}(x_3, x_2) & \texttt{Var}(x_3) \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathbf{x_i} - \boldsymbol{\mu})(\mathbf{x_i} - \boldsymbol{\mu})^T$$

$$d \times d$$

- Symmetric
- Positive semi-definite

$$C_{pq} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)(x_{qi} - \mu_q)$$

$$C_{pp} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)^2 = \sigma_p^2$$

Covariance Matrix

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}, \ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} \mathtt{Var}(x_1) & \mathtt{Cov}(x_1, x_2) & \mathtt{Cov}(x_1, x_3) \\ \mathtt{Cov}(x_2, x_1) & \mathtt{Var}(x_2) & \mathtt{Cov}(x_2, x_3) \\ \mathtt{Cov}(x_3, x_1) & \mathtt{Cov}(x_3, x_2) & \mathtt{Var}(x_3) \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathbf{x_i} - \boldsymbol{\mu})(\mathbf{x_i} - \boldsymbol{\mu})^T$$

$$d \times d$$

- Symmetric $\mathbf{C}^T = \mathbf{C}$
- Positive semi-definite

$$C_{pq} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)(x_{qi} - \mu_q)$$

$$\mathbf{x}^T \mathbf{C} \mathbf{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} \left[(\mathbf{x_i} - \boldsymbol{\mu})^T \mathbf{x}\right]^T \left[(\mathbf{x_i} - \boldsymbol{\mu})^T \mathbf{x}\right]$$

$$\implies \mathbf{x}^T \mathbf{C} \mathbf{x} \geq 0, \ \mathbf{x} \neq \mathbf{0}$$

$$C_{pp} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)^2 = \sigma_p^2$$

Covariance Matrix

$$D = \{\mathbf{x_1}, \cdots, \mathbf{x_n}\}, \ \mathbf{x_i} \in \mathbb{R}^d$$

$$\mathbf{x_i} = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{di} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x_1} & \cdots & \mathbf{x_n} \\ | & & | \end{bmatrix}$$

$$d \times n$$

$$\boldsymbol{\mu} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} \texttt{Var}(x_1) & \texttt{Cov}(x_1, x_2) & \texttt{Cov}(x_1, x_3) \\ \texttt{Cov}(x_2, x_1) & \texttt{Var}(x_2) & \texttt{Cov}(x_2, x_3) \\ \texttt{Cov}(x_3, x_1) & \texttt{Cov}(x_3, x_2) & \texttt{Var}(x_3) \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathbf{x_i} - \boldsymbol{\mu})(\mathbf{x_i} - \boldsymbol{\mu})^T$$

$$d \times d$$

- Symmetric $\quad \mathbf{C}^T = \mathbf{C}$
- Positive semi-definite

Centered Dataset

$$\boxed{\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x_i} \mathbf{x_i}^T}$$

$$C_{pq} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)(x_{qi} - \mu_q)$$

$$\mathbf{x}^T \mathbf{C} \mathbf{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} \left[(\mathbf{x_i} - \boldsymbol{\mu})^T \mathbf{x} \right]^T \left[(\mathbf{x_i} - \boldsymbol{\mu})^T \mathbf{x} \right]$$

$$\implies \mathbf{x}^T \mathbf{C} \mathbf{x} \geq 0, \ \mathbf{x} \neq \mathbf{0}$$

$$C_{pp} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_{pi} - \mu_p)^2 = \sigma_p^2$$

Min-Max Theorem

For a symmetric matrix of size $d \times d$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$, the following inequality holds for all vectors $\mathbf{x} \in \mathbb{R}^d$ with $||\mathbf{x}|| = 1$:

$$\lambda_d \leq \mathbf{x}^T \mathbf{C} \mathbf{x} \leq \lambda_1$$

Min-Max Theorem

For a symmetric matrix of size $d \times d$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$, the following inequality holds for all vectors $\mathbf{x} \in \mathbb{R}^d$ with $||\mathbf{x}|| = 1$:

$$\mathbf{C} = \sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T$$

$$\lambda_d \leq \mathbf{x}^T \mathbf{C} \mathbf{x} \leq \lambda_1$$

$$\mathbf{x} = \sum_{i=1}^{d} \alpha_i \mathbf{w_i}$$

Min-Max Theorem

For a symmetric matrix of size $d \times d$ with eigenvalues $\lambda_1 \geq \quad \cdots \quad \geq \lambda_d$, the following inequality holds for all vectors $\mathbf{x} \in \mathbb{R}^d$ with $||\mathbf{x}|| = 1$:

$$\lambda_d \leq \mathbf{x}^T \mathbf{C} \mathbf{x} \leq \lambda_1$$

$$\mathbf{C} = \sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T$$

$$\mathbf{x} = \sum_{i=1}^{d} \alpha_i \mathbf{w_i}$$

$$\mathbf{x}^T \mathbf{x} = \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i} \right) = 1$$

Min-Max Theorem

For a symmetric matrix of size $d \times d$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$,
the following inequality holds for all vectors $\mathbf{x} \in \mathbb{R}^d$ with $||\mathbf{x}|| = 1$:

$$\lambda_d \leq \mathbf{x}^T \mathbf{C} \mathbf{x} \leq \lambda_1$$

$$\mathbf{C} = \sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T$$

$$\mathbf{x} = \sum_{i=1}^{d} \alpha_i \mathbf{w_i}$$

$$\mathbf{x}^T \mathbf{x} = \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i} \right) = 1$$

$$\sum_{i=1}^{d} \alpha_i^2 = 1$$

Min-Max Theorem

For a symmetric matrix of size $d \times d$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$,
the following inequality holds for all vectors $\mathbf{x} \in \mathbb{R}^d$ with $||\mathbf{x}|| = 1$:

$$\lambda_d \leq \mathbf{x}^T \mathbf{C} \mathbf{x} \leq \lambda_1$$

$$\mathbf{C} = \sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T$$

$$\mathbf{x} = \sum_{i=1}^{d} \alpha_i \mathbf{w_i}$$

$$\mathbf{x}^T \mathbf{C} \mathbf{x} = \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i} \right)$$

$$\mathbf{x}^T \mathbf{x} = \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i} \right) = 1$$

$$\sum_{i=1}^{d} \alpha_i^2 = 1$$

Min-Max Theorem

For a symmetric matrix of size $d \times d$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$,
the following inequality holds for all vectors $\mathbf{x} \in \mathbb{R}^d$ with $||\mathbf{x}|| = 1$:

$$\lambda_d \leq \mathbf{x}^T \mathbf{C} \mathbf{x} \leq \lambda_1$$

$$\mathbf{C} = \sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T$$

$$\mathbf{x} = \sum_{i=1}^{d} \alpha_i \mathbf{w_i}$$

$$\mathbf{x}^T \mathbf{x} = \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i} \right) = 1$$

$$\sum_{i=1}^{d} \alpha_i^2 = 1$$

$$\mathbf{x}^T \mathbf{C} \mathbf{x} = \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i} \right)$$

$$= \sum_{i=1}^{d} \alpha_i^2 \lambda_i$$

Min-Max Theorem

For a symmetric matrix of size $d \times d$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$,
the following inequality holds for all vectors $\mathbf{x} \in \mathbb{R}^d$ with $||\mathbf{x}|| = 1$:

$$\lambda_d \leq \mathbf{x}^T \mathbf{C} \mathbf{x} \leq \lambda_1$$

$$\mathbf{C} = \sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T$$

$$\mathbf{x} = \sum_{i=1}^{d} \alpha_i \mathbf{w_i}$$

$$\mathbf{x}^T \mathbf{x} = \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i} \right) = 1$$

$$\sum_{i=1}^{d} \alpha_i^2 = 1$$

$$\mathbf{x}^T \mathbf{C} \mathbf{x} = \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \lambda_i \mathbf{w_i} \mathbf{w_i}^T \right) \left(\sum_{i=1}^{d} \alpha_i \mathbf{w_i} \right)$$

$$= \sum_{i=1}^{d} \alpha_i^2 \lambda_i$$

$$\lambda_d \left(\sum \alpha_i^2 \right) \leq \sum_{i=1}^{d} \alpha_i^2 \lambda_i \leq \lambda_1 \left(\sum \alpha_i^2 \right)$$

END