Linear Algebra Review

MLT

Karthik Thiagarajan



Vectors

Features

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

Data

ML: "Learning from data"

Vectors

 $egin{bmatrix} 1 \ 2 \ 3 \ 4 \end{bmatrix}$ $\mathbf{x} \in \mathbb{R}^d$

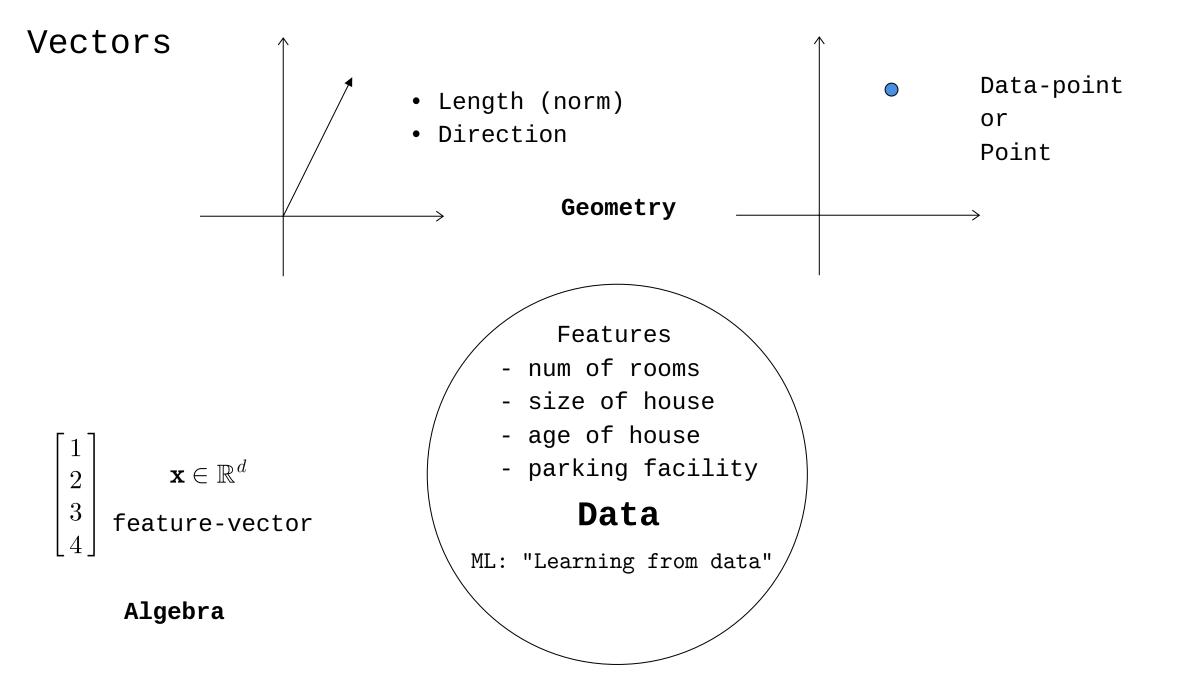
Algebra

Features

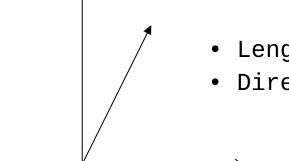
- num of rooms
- size of house
- age of house
- parking facility

Data

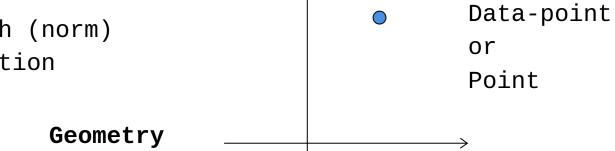
ML: "Learning from data"



Vectors



- Length (norm)
- Direction



- num of rooms
- size of house

Features

- age of house
- parking facility

Data

ML: "Learning from data"

[1, 2, 3, 4] Array or List

Computation

 $\mathbf{x} \in \mathbb{R}^d$ 2 3 feature-vector

Algebra

\mathbb{R}^d : Conventions

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

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

\mathbb{R}^d : Conventions

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

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

$$D = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$$

$$\mathbf{X} = \left[egin{array}{cccc} ert & \mathbf{x}_1 & \cdots & \mathbf{x}_n \ ert & & ert \end{array}
ight]$$

- X is a data-matrix.
- ullet Columns of X are data-points.
- Rows are the features.
- Its shape is (d, n).
- This is a convention.

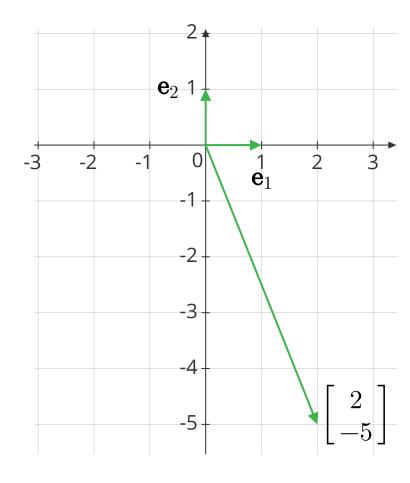
$$\mathbf{e}_1 = \left[egin{array}{c} 1 \ 0 \end{array}
ight], \mathbf{e}_2 = \left[egin{array}{c} 0 \ 1 \end{array}
ight]$$

$$\mathbf{e}_1 = egin{bmatrix} 1 \ 0 \end{bmatrix}, \mathbf{e}_2 = egin{bmatrix} 0 \ 1 \end{bmatrix}$$

- basis
 - not unique
 - $\{ {f e}_1, {f e}_2 \}$ is the standard basis
- representation in a given basis
 - unique

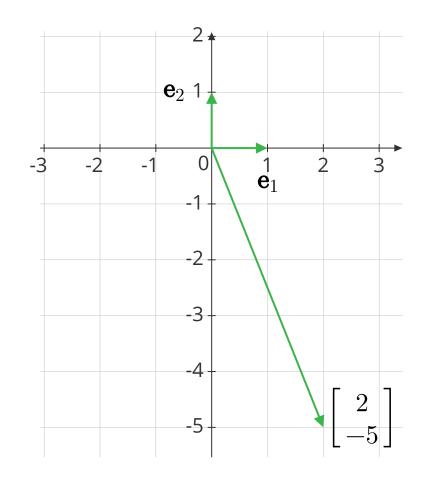
$$\mathbf{e}_1 = egin{bmatrix} 1 \ 0 \end{bmatrix}, \mathbf{e}_2 = egin{bmatrix} 0 \ 1 \end{bmatrix}$$

- basis
 - not unique
 - $\{ {f e}_1, {f e}_2 \}$ is the standard basis
- representation in a given basis
 - unique



$$\mathbf{e}_1 = egin{bmatrix} 1 \ 0 \end{bmatrix}, \mathbf{e}_2 = egin{bmatrix} 0 \ 1 \end{bmatrix}$$

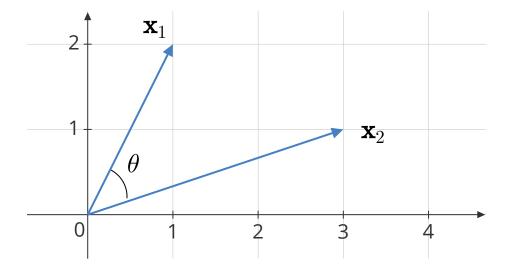
- basis
 - not unique
 - $\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$ is the standard basis
- representation in a given basis
 - unique



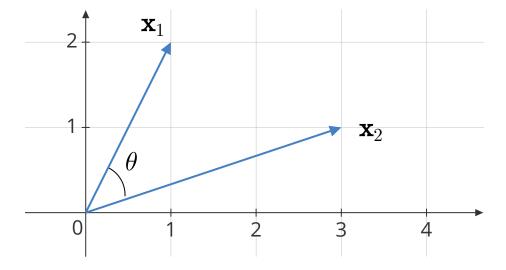
$$\begin{bmatrix} 2 \\ -5 \end{bmatrix} = 2 \cdot \mathbf{e_1} - 5 \cdot \mathbf{e_2}$$

- Lengths
- Angles (directions)

- Lengths
- Angles (directions)

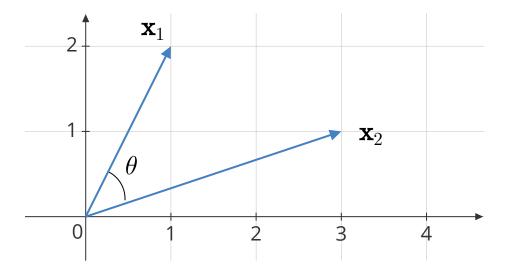


- Lengths
- Angles (directions)



$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

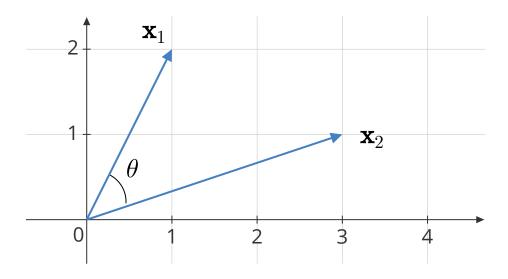
- Lengths
- Angles (directions)



$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = \mathbf{x}_1^T \mathbf{x}_2 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5$$

- Lengths
- Angles (directions)



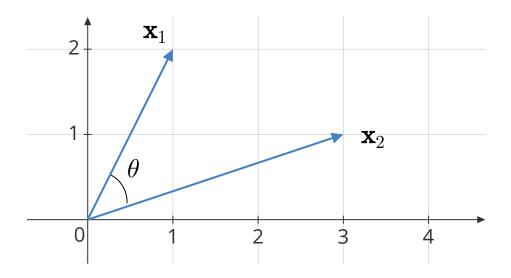
$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = \mathbf{x}_1^T \mathbf{x}_2 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5$$

$$||\mathbf{x}_1||^2 = \mathbf{x}_1^T \mathbf{x}_1 = 1^2 + 2^2 = 5$$

$$||\mathbf{x}_2||^2 = \mathbf{x}_2^T \mathbf{x}_2 = 3^2 + 1^2 = 10$$

- Lengths
- Angles (directions)



$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = \mathbf{x}_1^T \mathbf{x}_2 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5$$

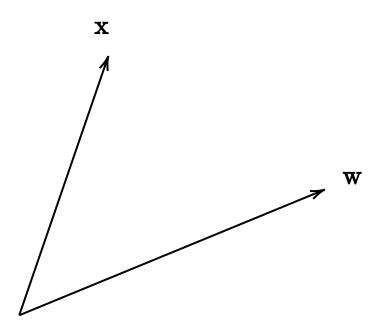
$$||\mathbf{x}_1||^2 = \mathbf{x}_1^T \mathbf{x}_1 = 1^2 + 2^2 = 5$$

$$||\mathbf{x}_2||^2 = \mathbf{x}_2^T \mathbf{x}_2 = 3^2 + 1^2 = 10$$

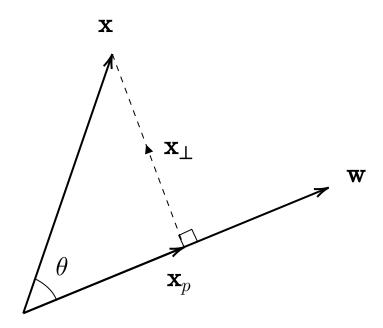
$$\cos \theta = \frac{\mathbf{x}_1^T \mathbf{x}_2}{||\mathbf{x}_1|| \cdot ||\mathbf{x}_2||} = \frac{5}{\sqrt{5} \cdot \sqrt{10}} = \frac{1}{\sqrt{2}}$$

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

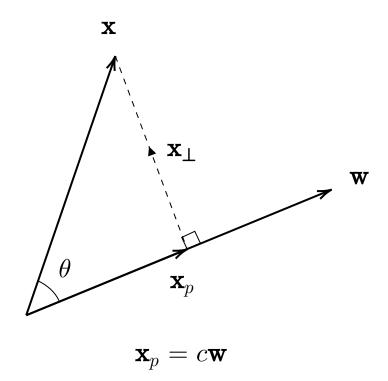
 \mathbb{R}^d : Projections



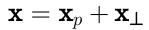
 \mathbb{R}^d : Projections

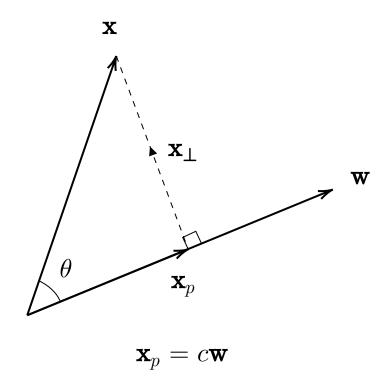


 \mathbb{R}^d : Projections

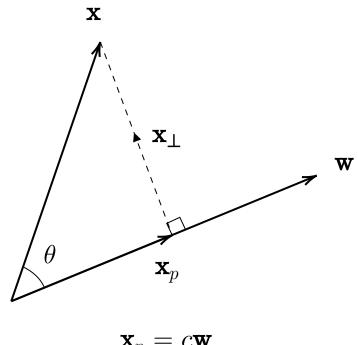


 \mathbb{R}^d : Projections

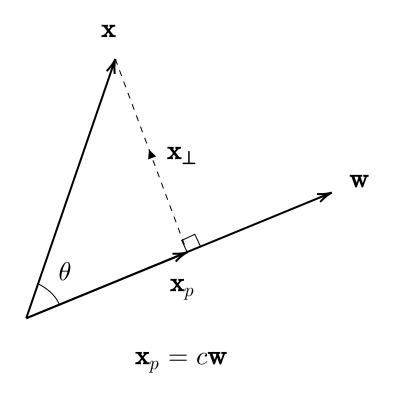




$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_{\perp}$$
$$= c\mathbf{w} + \mathbf{x}_{\perp}$$

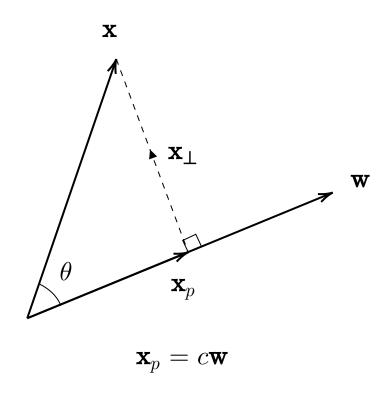


$$\mathbf{x}_p = c\mathbf{w}$$



$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_{\perp}$$
$$= c\mathbf{w} + \mathbf{x}_{\perp}$$

$$\mathbf{x}^T \mathbf{w} = \left(c \mathbf{w}^T + \mathbf{x}_{\perp}^T \right) \mathbf{w}$$

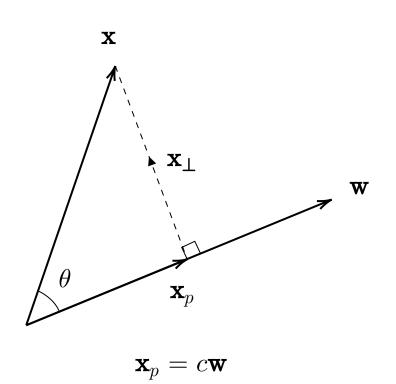


$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_{\perp}$$

$$= c\mathbf{w} + \mathbf{x}_{\perp}$$

$$\mathbf{x}^T \mathbf{w} = (c\mathbf{w}^T + \mathbf{x}_{\perp}^T)\mathbf{w}$$

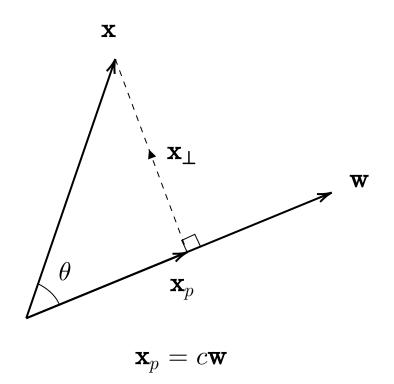
$$= c\mathbf{w}^T \mathbf{w}$$



$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_{\perp}$$
$$= c\mathbf{w} + \mathbf{x}_{\perp}$$

$$\mathbf{x}^T \mathbf{w} = \left(c \mathbf{w}^T + \mathbf{x}_{\perp}^T \right) \mathbf{w}$$
$$= c \mathbf{w}^T \mathbf{w}$$

$$\implies c = \frac{\mathbf{x}^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}}$$



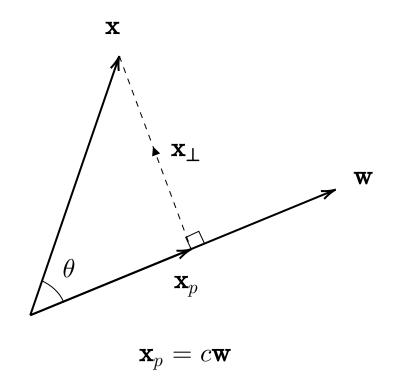
$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_{\perp}$$
$$= c\mathbf{w} + \mathbf{x}_{\perp}$$

$$\mathbf{x}^T \mathbf{w} = \left(c \mathbf{w}^T + \mathbf{x}_{\perp}^T \right) \mathbf{w}$$
$$= c \mathbf{w}^T \mathbf{w}$$

$$\implies c = \frac{\mathbf{x}^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}}$$

Vector Projection

$$\mathbf{x}_p = rac{\mathbf{x}^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \mathbf{w}$$



$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_{\perp}$$
$$= c\mathbf{w} + \mathbf{x}_{\perp}$$

$$\mathbf{x}^T \mathbf{w} = \left(c \mathbf{w}^T + \mathbf{x}_{\perp}^T \right) \mathbf{w}$$
$$= c \mathbf{w}^T \mathbf{w}$$

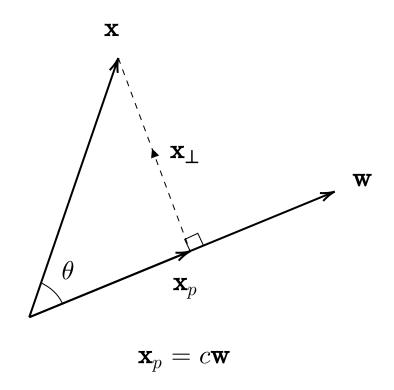
$$\Longrightarrow c = \frac{\mathbf{x}^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}}$$

Vector Projection

$$\mathbf{x}_p = rac{\mathbf{x}^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \mathbf{w}$$

Scalar Projection

$$||\mathbf{x}_p|| = rac{\mathbf{x}^T \mathbf{w}}{||\mathbf{w}||}$$



$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_{\perp}$$
$$= c\mathbf{w} + \mathbf{x}_{\perp}$$

$$\mathbf{x}^T \mathbf{w} = \left(c \mathbf{w}^T + \mathbf{x}_{\perp}^T \right) \mathbf{w}$$
$$= c \mathbf{w}^T \mathbf{w}$$

$$\implies c = \frac{\mathbf{x}^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}}$$

Vector Projection

$$\mathbf{x}_p = rac{\mathbf{x}^T \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \mathbf{w}$$

Scalar Projection

$$||\mathbf{x}_p|| = \frac{\mathbf{x}^T \mathbf{w}}{||\mathbf{w}||}$$

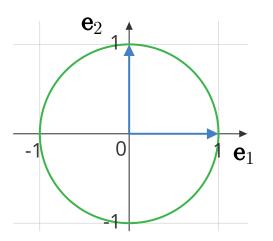
Projection on unit-norm vector

$$|\mathbf{x}^T \mathbf{w}| = 1$$

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

- Orthogonal
- Unit norm

Basis-1



$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

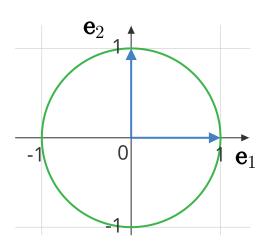
$$\mathbf{e}_1^T \mathbf{e}_2 = 0$$

$$||\mathbf{e}_1|| = ||\mathbf{e}_2|| = 1$$

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

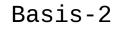
- Orthogonal
- Unit norm

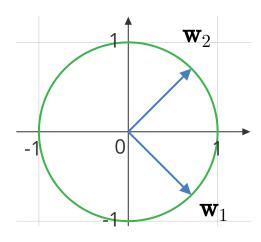
Basis-1



$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{e}_1^T \mathbf{e}_2 = 0$$
$$||\mathbf{e}_1|| = ||\mathbf{e}_2|| = 1$$





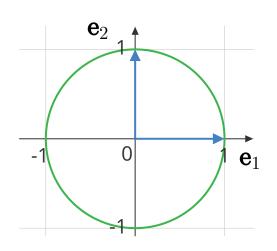
$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{w}_1^T \mathbf{w}_2 = 0$$
$$||\mathbf{w}_1|| = ||\mathbf{w}_2|| = 1$$

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

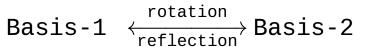
- Orthogonal
- Unit norm

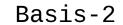
Basis-1

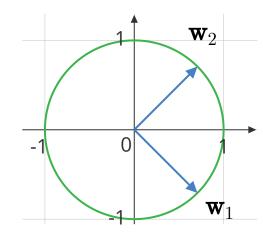


$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbf{e}_1^T \mathbf{e}_2 = 0$$
$$||\mathbf{e}_1|| = ||\mathbf{e}_2|| = 1$$



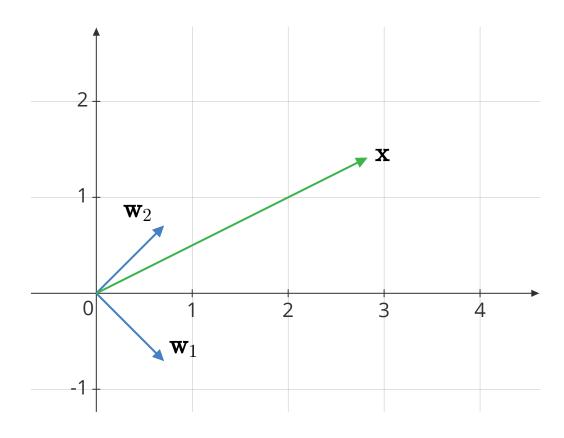




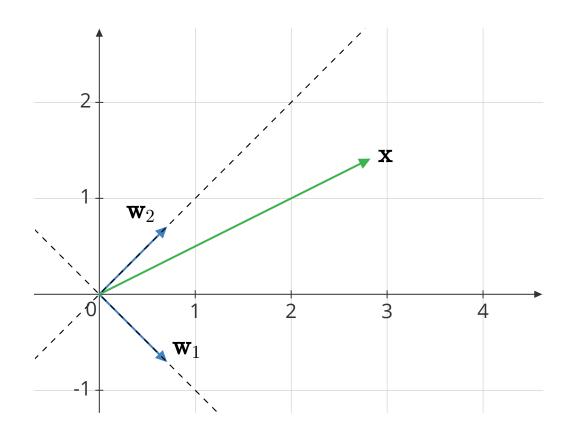
$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{w}_1^T \mathbf{w}_2 = 0$$
$$||\mathbf{w}_1|| = ||\mathbf{w}_2|| = 1$$

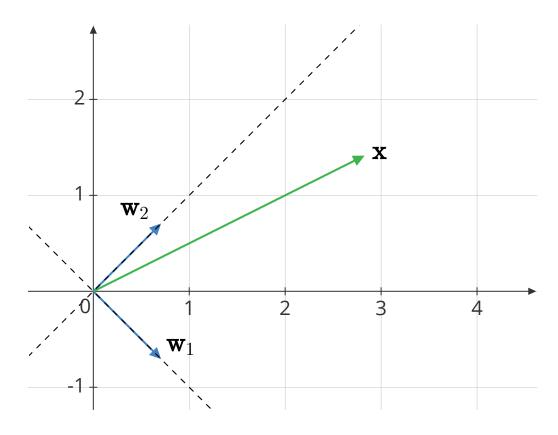
 \mathbb{R}^d : Orthonormal Basis



 \mathbb{R}^d : Orthonormal Basis



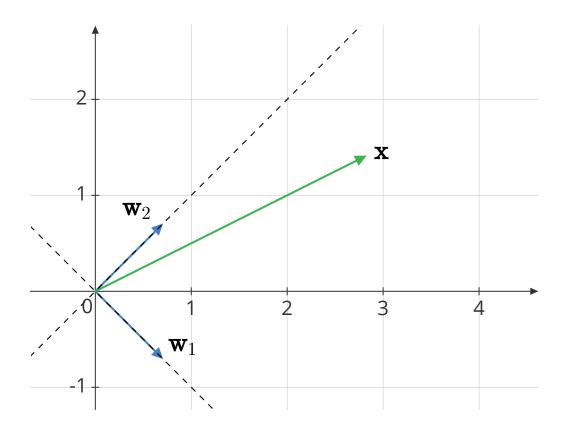
 \mathbb{R}^d : Orthonormal Basis



$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ \sqrt{2} \end{bmatrix}$$

$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

 \mathbb{R}^d : Orthonormal Basis

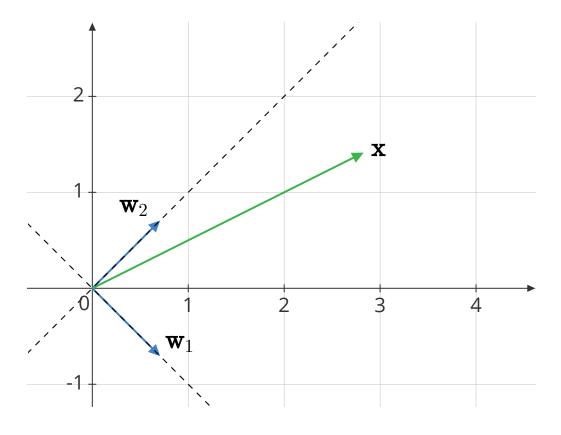


$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ \sqrt{2} \end{bmatrix}$$

$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = 2\mathbf{\sqrt{2}} \cdot \mathbf{e}_1 + \mathbf{\sqrt{2}} \cdot \mathbf{e}_2$$

 \mathbb{R}^d : Orthonormal Basis

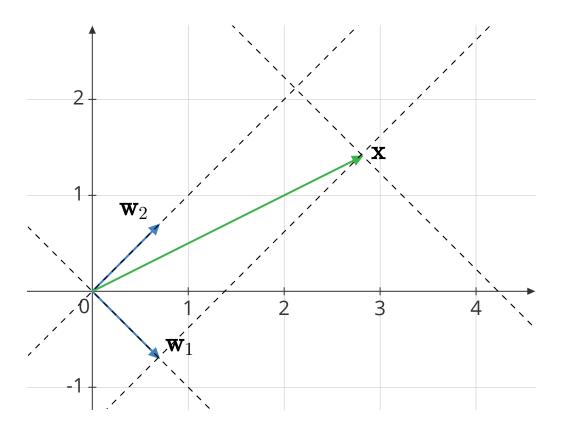


$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
 $\mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ \sqrt{2} \end{bmatrix}$

$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{x} = 2\mathbf{\sqrt{2}} \cdot \mathbf{e}_1 + \mathbf{\sqrt{2}} \cdot \mathbf{e}_2$$
$$= (\mathbf{x}^T \mathbf{w}_1) \mathbf{w}_1 + (\mathbf{x}^T \mathbf{w}_2) \mathbf{w}_2$$

 \mathbb{R}^d : Orthonormal Basis



$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
 $\mathbf{x} = \begin{bmatrix} 2\sqrt{2} \\ \sqrt{2} \end{bmatrix}$

$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

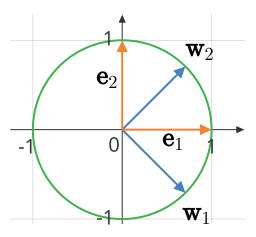
$$\mathbf{x} = 2\sqrt{2} \cdot \mathbf{e}_1 + \sqrt{2} \cdot \mathbf{e}_2$$

$$= (\mathbf{x}^T \mathbf{w}_1) \mathbf{w}_1 + (\mathbf{x}^T \mathbf{w}_2) \mathbf{w}_2$$

$$= 1\mathbf{w}_1 + 3\mathbf{w}_2$$

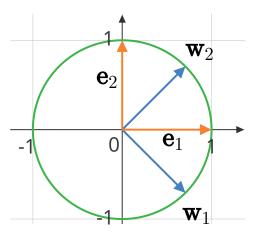
$$\mathbf{Q} = \left[egin{array}{ccc} ert & ert \ \mathbf{w}_1 & \mathbf{w}_2 \ ert & ert \end{array}
ight]$$

$$\mathbf{Q} = \left[egin{array}{ccc} ert & ert \ \mathbf{w}_1 & \mathbf{w}_2 \ ert & ert \end{array}
ight]$$



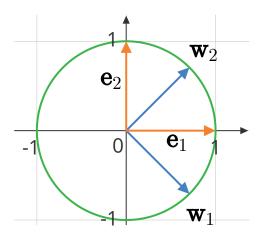
$$\mathbf{Q} = \left[egin{array}{ccc} ert & ert \ \mathbf{w}_1 & \mathbf{w}_2 \ ert & ert \end{array}
ight]$$

$$egin{aligned} \mathbf{Q}^T\mathbf{Q} &= egin{bmatrix} & - & \mathbf{w}_1^T & - \ - & \mathbf{w}_2^T & - \end{bmatrix} egin{bmatrix} & \mathbf{w}_1 & \mathbf{w}_2 \ & \mathbf{w}_1 & \mathbf{w}_2 \ & & \end{bmatrix} \ &= egin{bmatrix} 1 & 0 \ 0 & 1 \end{bmatrix} \ &= \mathbf{I} \end{aligned}$$



$$\mathbf{Q} = \left[egin{array}{ccc} ert & ert \ \mathbf{w}_1 & \mathbf{w}_2 \ ert & ert \end{array}
ight]$$

$$\mathbf{Q}^T \mathbf{Q} = egin{bmatrix} - & \mathbf{w}_1^T & - \ - & \mathbf{w}_2^T & - \end{bmatrix} egin{bmatrix} | & \mathbf{w}_1 & \mathbf{w}_2 \ | & | & | \end{bmatrix} \ = egin{bmatrix} 1 & 0 \ 0 & 1 \end{bmatrix} \ = \mathbf{I}$$

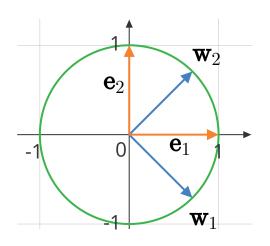


$$\mathbf{Q} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

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

$$\mathbf{Q} = \left[egin{array}{ccc} ert & ert \ \mathbf{w}_1 & \mathbf{w}_2 \ ert & ert \end{array}
ight]$$

$$\mathbf{Q}^T \mathbf{Q} = egin{bmatrix} - & \mathbf{w}_1^T & - \ - & \mathbf{w}_2^T & - \end{bmatrix} egin{bmatrix} | & \mathbf{w}_1 & \mathbf{w}_2 \ | & | & | \end{bmatrix} \ = egin{bmatrix} 1 & 0 \ 0 & 1 \end{bmatrix} \ = \mathbf{I}$$



$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{w}_1 = \mathbf{Q}\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; \mathbf{w}_2 = \mathbf{Q}\mathbf{e}_2 = rac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

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

$$\mathbf{Q} = \left[egin{array}{ccc} ert & ert \ \mathbf{w}_1 & \mathbf{w}_2 \ ert & ert \end{array}
ight]$$

$$\mathbf{Q}^T \mathbf{Q} = \begin{bmatrix} - & \mathbf{w}_1^T & - \\ - & \mathbf{w}_2^T & - \end{bmatrix} \begin{bmatrix} | & | \\ \mathbf{w}_1 & \mathbf{w}_2 \\ | & | \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \mathbf{I}$$

$$\mathbf{e}_2$$
 \mathbf{e}_2
 \mathbf{v}_2
 \mathbf{e}_1

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathbf{w}_1 = \mathbf{Q}\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \mathbf{w}_2 = \mathbf{Q}\mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Orthogonal matrix

- Preserves inner products
 - Preserves lengths
 - Preserves angles

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