Linear Algebra 1

Industrial AI Lab.

Changyun Choi / Iljeok Kim

POSTECH

Some materials from linear algebra review by Prof. Zico Kolter from CMU

Linear Equations

- Set of linear equations (two equations, two unknowns)

$$\begin{aligned} 4x_1 - 5x_2 &= -13 \\ -2x_1 + 3x_2 &= 9 \end{aligned}$$

Linear Equations

- Solving linear equations
  - Two linear equations

$$\begin{aligned} 4x_1 - 5x_2 &= -13 \\ -2x_1 + 3x_2 &= 9 \end{aligned}$$

  - In a vector form, $Ax = b$, with

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad b = \begin{bmatrix} -13 \\ 9 \end{bmatrix}$$

Linear Equations

- Solving linear equations
  - Two linear equations

$$
\begin{aligned}
4x_1 - 5x_2 &= -13 \\
-2x_1 + 3x_2 &= 9
\end{aligned}
$$

  - In a vector form, $Ax = b$, with

$$
A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad b = \begin{bmatrix} -13 \\ 9 \end{bmatrix}
$$

  - Solution using inverse

$$
\begin{aligned}
Ax &= b \\
A^{-1}Ax &= A^{-1}b \\
x &= A^{-1}b
\end{aligned}
$$

Linear Equations

- Solving linear equations
  - Two linear equations

$$\begin{aligned} 4x_1 - 5x_2 &= -13 \\ -2x_1 + 3x_2 &= 9 \end{aligned}$$

  - In a vector form, $Ax = b$, with

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad b = \begin{bmatrix} -13 \\ 9 \end{bmatrix}$$

  - Solution using inverse

$$\begin{aligned} Ax &= b \\ A^{-1}Ax &= A^{-1}b \\ x &= A^{-1}b \end{aligned}$$

  - Don’t worry here about how to compute matrix inverse
  - We will use a numpy to compute

Linear Equations in Python

$$\begin{aligned} 4x_1 - 5x_2 &= -13 \\ -2x_1 + 3x_2 &= 9 \end{aligned}$$

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad b = \begin{bmatrix} -13 \\ 9 \end{bmatrix}$$

```
import numpy as np
```

```
A = np.array([[4, -5],
              [-2, 3]])
b = np.array([[-13],
              [9]])

x = np.linalg.inv(A).dot(b)
print(x)
```

```
[[ 3.]
 [ 5.]]
```

```
A = np.asmatrix(A)
b = np.asmatrix(b)

x = A.I*b
print(x)
```

```
[[ 3.]
 [ 5.]]
```

System of Linear Equations

- Consider a system of linear equations

$$
\begin{aligned}
y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\vdots \\
y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\end{aligned}
$$

- Can be written in a matrix form as $y = Ax$, where

$$
y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
$$

Elements of a Matrix

- Can write a matrix in terms of its columns

$$A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}$$

- Careful, a_i here corresponds to an entire vector $a_i \in \mathbb{R}^m$
- Similarly, can write a matrix in terms of rows

$$A = \begin{bmatrix} - & b_1^T & - \\ - & b_2^T & - \\ & \vdots & \\ - & b_m^T & - \end{bmatrix}$$

- $b_i \in \mathbb{R}^n$

Vector-Vector Products

- Inner product: $x, y \in \mathbb{R}^n$

$$x^T y = \sum_{i=1}^{n} x_i \, y_i \qquad \in \mathbb{R}$$

```
x = np.array([[1],
              [1]])
y = np.array([[2],
              [3]])

print(x.T.dot(y))
```

```
[[5]]
```

```
x = np.asmatrix(x)
y = np.asmatrix(y)

print(x.T*y)
```

```
[[5]]
```

Matrix-Vector Products

- $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n \Leftrightarrow Ax \in \mathbb{R}^m$

- Writing A by rows, each entry of Ax is an *inner product* between x and a row of A

$$A = \begin{bmatrix} - & b_1^T & - \\ - & b_2^T & - \\ & \vdots & \\ - & b_m^T & - \end{bmatrix}, \qquad Ax \in \mathbb{R}^m = \begin{bmatrix} b_1^T x \\ b_2^T x \\ \vdots \\ b_m^T x \end{bmatrix}$$

Matrix-Vector Products

- $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n \Leftrightarrow Ax \in \mathbb{R}^m$

- Writing A by columns, Ax is a *linear combination* of the columns of A, with coefficients given by x

$$A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}, \qquad Ax \in \mathbb{R}^m = \sum_{i=1}^{n} a_i x_i$$

Norms (Strength or Distance in Linear Space)

- A vector norm is any function $f: \mathbb{R}^n \Longrightarrow \mathbb{R}$ with

  1. $f(x) \geq 0$ and $f(x) = 0 \iff x = 0$
  2. $f(ax) = |a| f(x)$ for $a \in \mathbb{R}$
  3. $f(x+y) \leq f(x) + f(y)$

$$\|x\|_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$$

- l_2 norm

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

- l_1 norm

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

- $\|x\|$ measures length of vector (from origin)

Norms in Python

```
x = np.array([[4],
              [3]])

np.linalg.norm(x, 2)
```

```
5.0
```

```
np.linalg.norm(x, 1)
```

```
7.0
```

Orthogonality

- Two vectors $x, y \in \mathbb{R}^n$ are *orthogonal* if

$$x^T y = 0$$

- They are *orthonormal* if

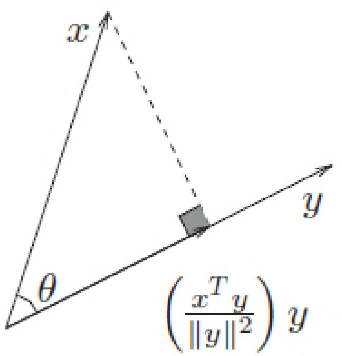
$$x^T y = 0 \quad \text{and} \quad \|x\|_2 = \|y\|_2 = 1$$

Angle between Vectors

- For any $x, y \in \mathbb{R}^n$,

$$|x^T y| \le \|x\| \, \|y\|$$

- (unsigned) angle between vectors in $\mathbb{R}^n$ defined as

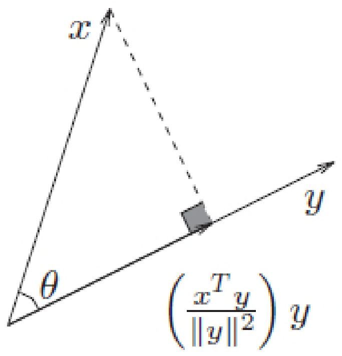


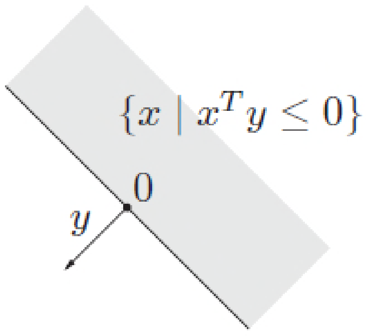
$$\theta = \angle(x, y) = \cos^{-1} \frac{x^T y}{\|x\| \|y\|}$$

$$\text{thus } x^T y = \|x\| \|y\| \cos \theta$$

Angle between Vectors

$$\theta = \angle(x, y) = \cos^{-1} \frac{x^T y}{\|x\| \|y\|}$$





- $\{x | x^T y \leq 0\}$ defines a half space with outward normal vector y, and boundary passing through 0

Linear Algebra 2

Industrial AI Lab.

Changyun Choi / Iljeok Kim

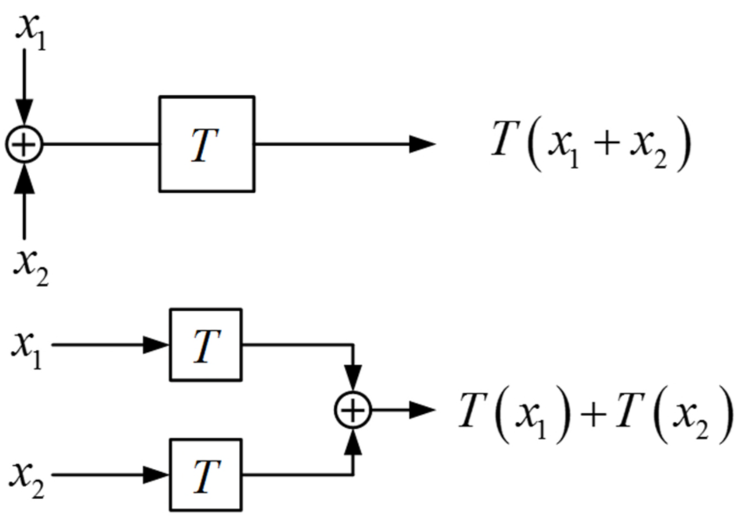
Linear Transformation

- See if the given transformation is linear
  - A linear system makes our life much easier

- Superposition
- Homogeneity

Linear Transformation

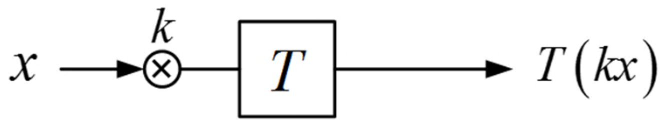
- Superposition

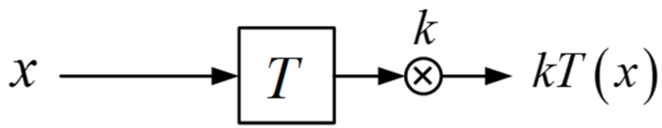


$$T(x_1 + x_2) = T(x_1) + T(x_2)$$

Linear Transformation

- Homogeneity





$$T(kx) = kT(x)$$

Matrix and (Linear) Transformation

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

$$\vec{y} = M\vec{x}$$

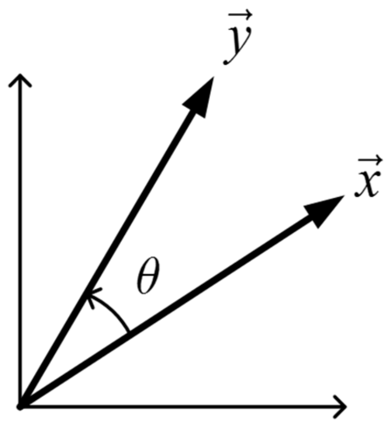
$$\begin{bmatrix} \\ \\ \\ \end{bmatrix} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}\begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

Given		Interpret
linear transformation	$\longrightarrow$	matrix
matrix	$\longrightarrow$	linear transformation

$\vec{x}$	linear transformation	$\vec{y}$
input	$\Longrightarrow$	output

Rotation

- Is a rotation operation linear?

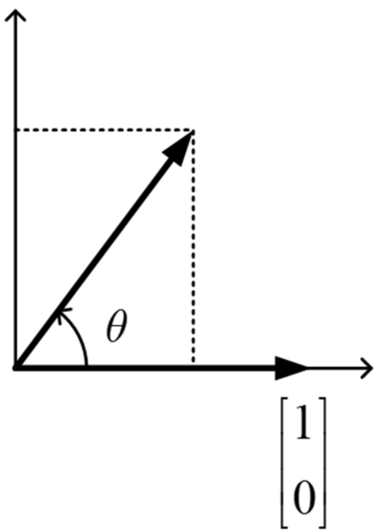


- Rotation matrix: $M = R(\theta)$
- Transformation: $\vec{y} = R(\theta)\vec{x}$

Rotation

- To find matrix $M = R(\theta)$

$$\vec{y} = R(\theta)\vec{x}$$

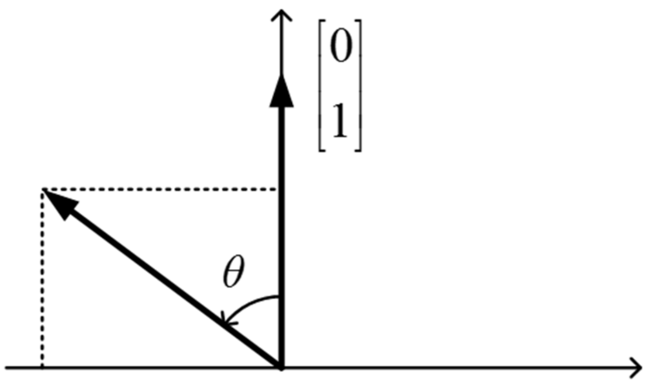


$$\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} = R(\theta) \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Rotation

- To find matrix $M = R(\theta)$

$$\vec{y} = R(\theta)\vec{x}$$



$$\begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix} = R(\theta) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Rotation

- To find matrix $M = R(\theta)$

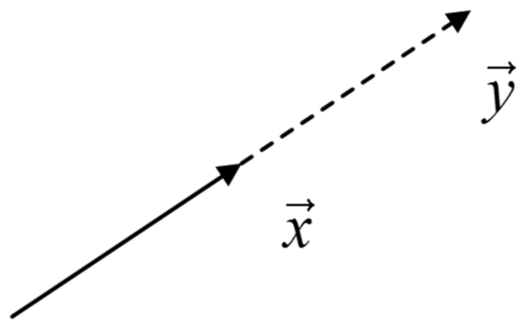
$$M\vec{x}_1 = \vec{y}_1 \qquad \qquad \Rightarrow \quad M\begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix} = \begin{bmatrix} \vec{y}_1 & \vec{y}_2 \end{bmatrix}$$
$$M\vec{x}_2 = \vec{y}_2$$

$$\Longrightarrow \quad \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = R(\theta)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = R(\theta)$$

- Note on how to find a matrix from two vectors and their linearly-transformed ones

Stretch/Compress

- Stretch/Compress
  - keep the direction



$$\vec{y} = \underset{\substack{\uparrow \\ \text{scalar (not matrix)}}}{k}\vec{x}$$

$$\vec{y} = kI\vec{x} \qquad \text{where } I = \text{ Identity martix}$$

$$\vec{y} = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \vec{x}$$

- Still represented by a matrix

Stretch/Compress: Example

- T: stretch by a along $\hat{x}$-direction & stretch by b along $\hat{y}$-direction
- Compute the corresponding matrix A

$$\begin{aligned} \begin{bmatrix} ax_1 \\ bx_2 \end{bmatrix} &= A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Longrightarrow A = ? \\ &= \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} A \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= \begin{bmatrix} a \\ 0 \end{bmatrix} \\ A \begin{bmatrix} 0 \\ 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ b \end{bmatrix} \\ A \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} &= A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \end{aligned}$$

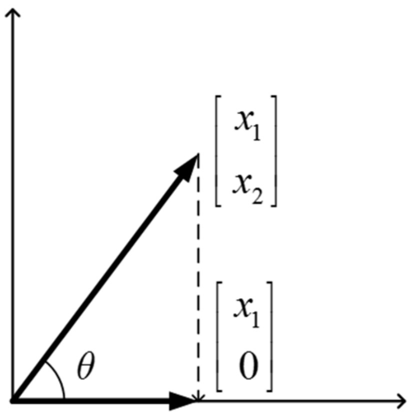
- More importantly, can you think of the corresponding transformation T by looking at $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$?

Projection

- Is a projection operation linear?

- Suppose P: Projection onto $\hat{x}$ - axis

$$\underset{\vec{x}}{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}} \overset{P}{\Longrightarrow} \underset{\vec{y}}{\begin{bmatrix} x_1 \\ 0 \end{bmatrix}}$$



$$\vec{y} = P\vec{x} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

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

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

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

Multiple Transformations

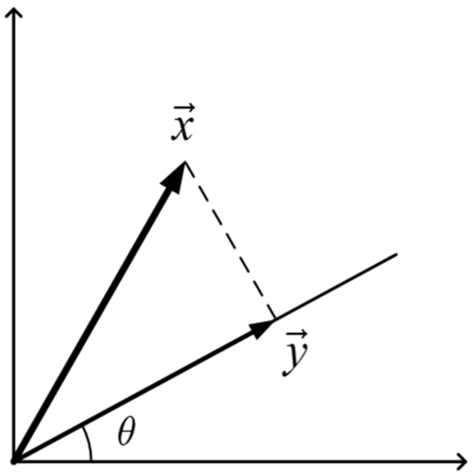
- T_1: transformation 1 of matrix M_1
- T_2: transformation 2 of matrix M_2
- T : Do transformation 1, followed by transformation 2

$$\vec{x} \xrightarrow{T_1} \vec{y} \xrightarrow{T_2} \vec{z}$$

$$\therefore\ M = M_2 M_1$$

$$\begin{aligned} \vec{y} &= M_1\vec{x} \\ \vec{z} &= M_2\vec{y} &&= M_2M_1\vec{x} \\ & &&= M\vec{x} \end{aligned}$$

Example: Projection onto Vector = $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$



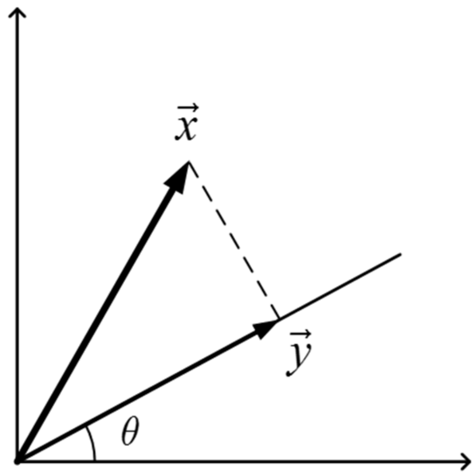
$$P\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos^2\theta \\ \cos\theta\sin\theta \end{bmatrix}$$

$$P\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\theta \\ \sin^2\theta \end{bmatrix}$$

$$P\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$$

Example: Projection onto Vector = $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$

- Another way to find this projection matrix



$$\vec{x} \xrightarrow{R(-\theta)} \vec{x}\,' \xrightarrow{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}} \vec{x}\,'' \xrightarrow{R(\theta)} \vec{y}$$

$$\begin{aligned}
\vec{y} &= R(\theta) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} R(-\theta)\vec{x} \\
&= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}
\end{aligned}$$

Eigenvalue and Eigenvector

$$A\vec{v} = \lambda\vec{v}$$

$A\vec{v}$ parallel to $\vec{v}$

$$\begin{aligned}
\lambda &= \begin{cases} \text{positive} \\ 0 \\ \text{negative} \end{cases} \\
\lambda\vec{v} &: \text{stretched vector} \\
&\quad (\text{same direction with } \vec{v}) \\
A\vec{v} &: \text{linearly transformed vector} \\
&\quad (\text{generally rotate} + \text{stretch})
\end{aligned}$$

How to Compute Eigenvalue and Eigenvector

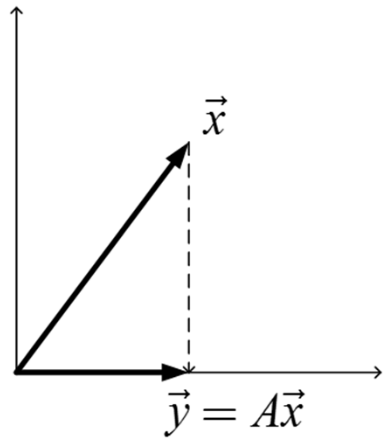
$$A\vec{v} = \lambda\vec{v} = \lambda I\vec{v}$$
$$A\vec{v} - \lambda I\vec{v} = (A - \lambda I)\vec{v} = 0$$

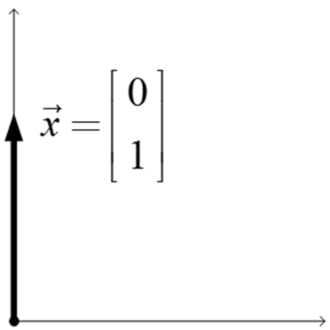
$$\Longrightarrow \quad \begin{array}{l} A - \lambda I = 0 \text{ or} \\ \vec{v} = 0 \text{ or} \\ (A - \lambda I)^{-1} \text{ does not exist} \end{array}$$

$$\Longrightarrow \quad \det(A - \lambda I) = 0$$

Example: Eigen Analysis of $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

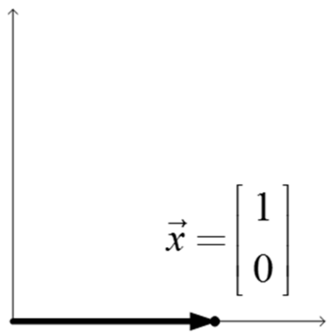
- $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$: projection onto $\hat{x}$ - axis
- Find eigenvalues and eigenvectors of A.





$$\vec{y} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = A\vec{x} = 0 \cdot \vec{x}$$

$$\lambda_1 = 0 \ \text{ and } \ \vec{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



$$\vec{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = A\vec{x} = 1 \cdot \vec{x}$$

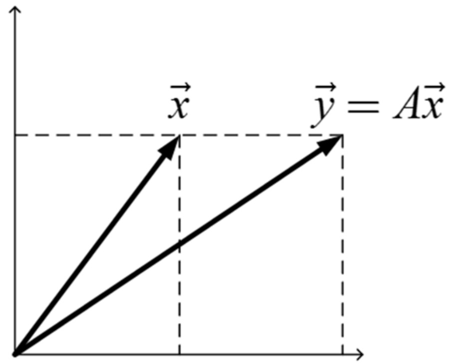
$$\lambda_2 = 1 \ \text{ and } \ \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

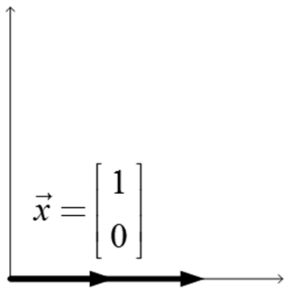
Example: Eigen Analysis of $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

- $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$: stretch by 2 along $\vec{x}$- axis

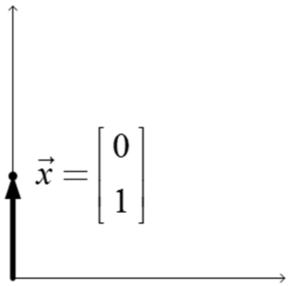
  stretch by 1 along $\vec{y}$- axis

- Find eigenvalues and eigenvectors.





$\lambda_1 = 2$ and $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$



$\lambda_2 = 1$ and $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

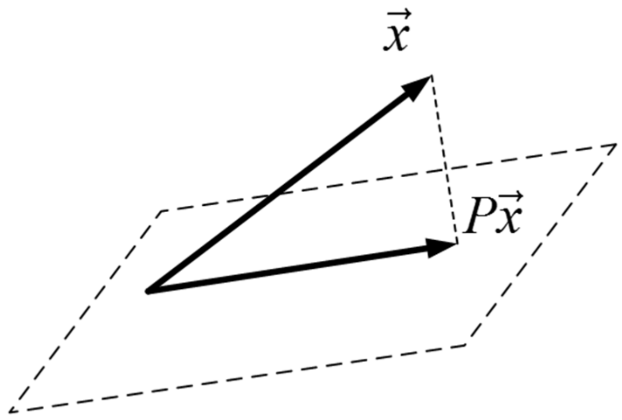
Eigen Analysis in Python

```
A = np.array([[2, 0],
              [0, 1]])
D, V = np.linalg.eig(A)

idx = np.argsort(-D)
D = D[idx]
V = V[idx]

print('D :', D)
print('V :', V)
```

```
D : [ 2.  1.]
V : [[ 1.  0.]
 [ 0.  1.]]
```

Example: Eigen Analysis of Projection

- Projection onto the plane
- Find eigenvalues and eigenvectors



- For any $\vec{x}$ in the plane, $P\vec{x} = \vec{x} \rightarrow \lambda = 1$
- For any $\vec{x}$ perpendicular to the plane, $P\vec{x} = \vec{0} \rightarrow \lambda = 0$

Linear Algebra 3

Industrial AI Lab.

Changyun Choi / Iljeok Kim

POSTECH

System of Linear Equations

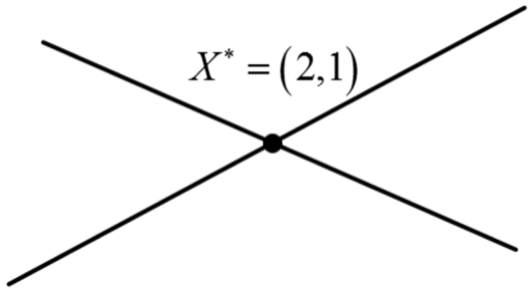
- Well-determined linear systems
- Under-determined linear systems
- Over-determined linear systems

Well-Determined Linear Systems

- System of linear equations

$$\begin{aligned} 2x_1 + 3x_2 &= 7 \\ x_1 + 4x_2 &= 6 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} x_1^* &= 2 \\ x_2^* &= 1 \end{aligned}$$

- Geometric point of view



Well-Determined Linear Systems

- System of linear equations

$$\begin{aligned} 2x_1 + 3x_2 &= 7 \\ x_1 + 4x_2 &= 6 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} x_1^* &= 2 \\ x_2^* &= 1 \end{aligned}$$

- Matrix form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \quad \overset{\text{Matrix form}}{\Longrightarrow} \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$AX = B$$

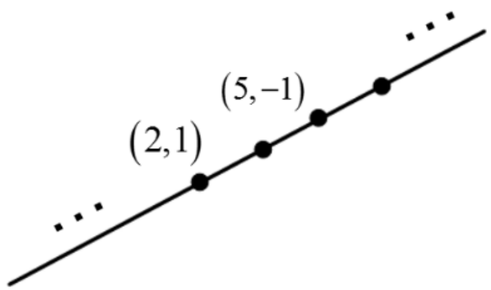
$$\therefore \quad X^* = A^{-1}B \qquad \text{if } A^{-1} \text{ exists}$$

Under-Determined Linear Systems

- System of linear equations

$$2x_1 + 3x_2 = 7 \quad \Longrightarrow \quad \text{Many solutions}$$

- Geometric point of view



Under-Determined Linear Systems

- System of linear equations

$$2x_1 + 3x_2 = 7 \quad \Longrightarrow \quad \text{Many solutions}$$

- Matrix form

$$a_{11}x_1 + a_{12}x_2 = b_1 \quad \xrightarrow{\text{Matrix form}} \quad \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = b_1$$

$$AX = B$$

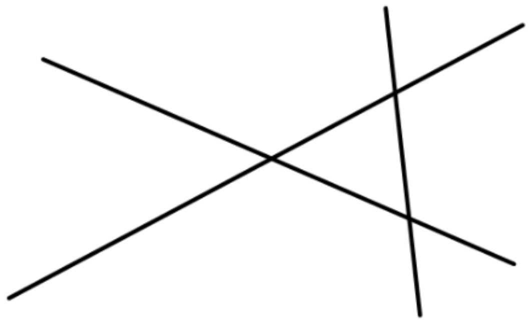
$$\therefore \text{ Many Solutions when } A \text{ is fat}$$

Over-Determined Linear Systems

- System of linear equations

$$\begin{aligned} 2x_1 + 3x_2 &= 7 \\ x_1 + 4x_2 &= 6 \\ x_1 + x_2 &= 4 \end{aligned} \implies \text{No solutions}$$

- Geometric point of view



Over-Determined Linear Systems

- System of linear equations

$$\begin{aligned} 2x_1 + 3x_2 &= 7 \\ x_1 + 4x_2 &= 6 \\ x_1 + x_2 &= 4 \end{aligned} \implies \text{No solutions}$$

- Matrix form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \\ a_{31}x_1 + a_{32}x_2 &= b_3 \end{aligned} \quad \overset{\text{Matrix form}}{\implies} \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$AX = B$$

$$\therefore \text{ No Solutions when } A \text{ is skinny}$$

Summary of Linear Systems

$$AX = B$$

- Square: Well-determined

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

- Fat: Under-determined

$$\begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = b_1$$

- Skinny: Over-determined

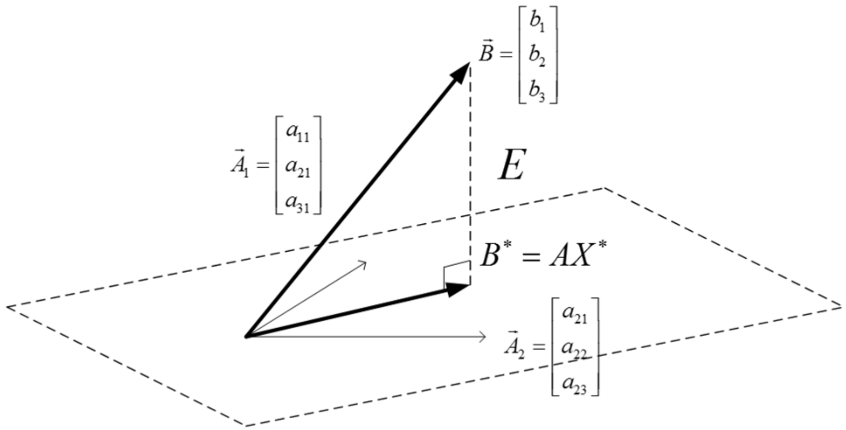
$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Least-Square Solution

- For over-determined linear system

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{or} \quad AX \neq B$$

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} \neq \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$



- Find X that minimizes $\|E\|$ or $\|E\|^2$ (error)

- *i.e.* optimization problem

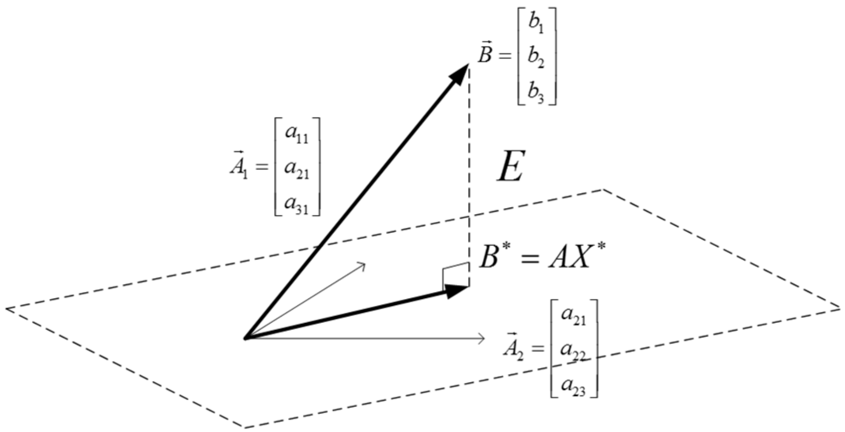
$$\min_X \|E\|^2 = \min_X \|AX - B\|^2$$

Least-Square Solution

- *i.e.* optimization problem

$$\min_X \|E\|^2 = \min_X \|AX - B\|^2$$
$$X^* = \left(A^T A\right)^{-1} A^T B$$
$$B^* = AX^* = A\left(A^T A\right)^{-1} A^T B$$

- Geometric interpretation



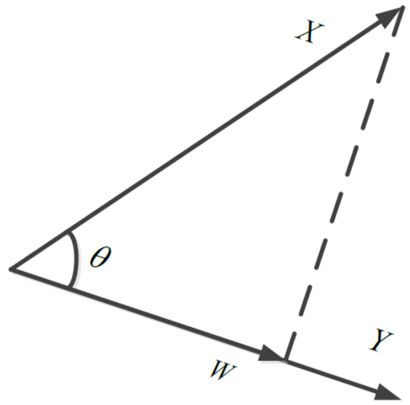
$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} \neq \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

- Often estimation problem

Vector Projection onto Y

- The vector projection of a vector X on (or onto) a nonzero vector Y is the orthogonal projection of X onto a straight line parallel to Y

$$
\begin{aligned}
& Y \perp (X - W) \\
\implies & Y^T (X - W) = Y^T \left(X - \omega \frac{Y}{\|Y\|} \right) = 0 \\
\implies & \omega = \frac{Y^T X}{Y^T Y} \|Y\| \\
& W = \omega \frac{Y}{\|Y\|} = \frac{Y^T X}{Y^T Y} Y = \frac{\langle X, Y \rangle}{\langle Y, Y \rangle} Y
\end{aligned}
$$

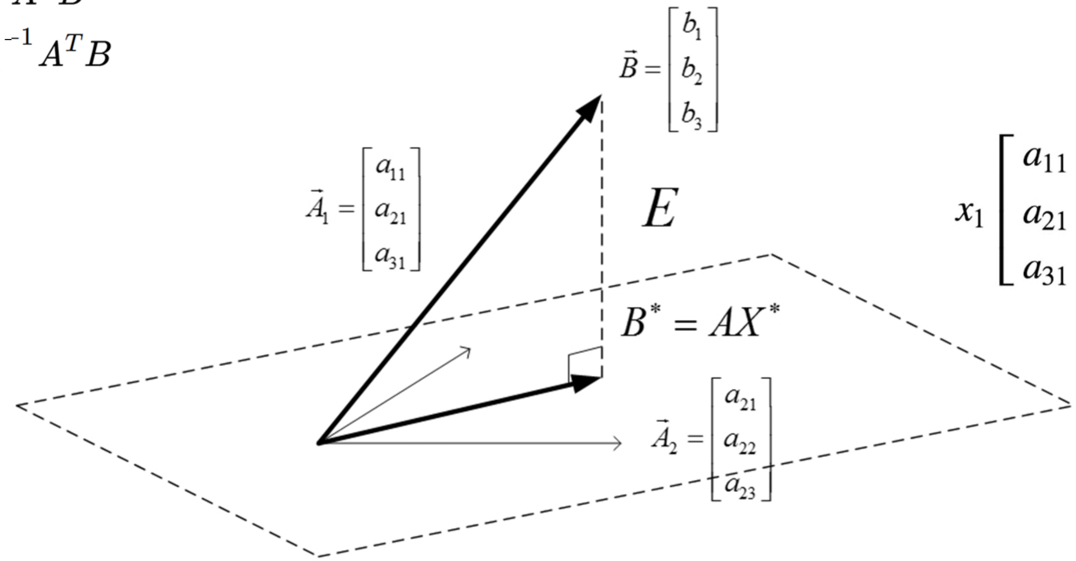


Orthogonal Projection onto a Subspace

- Projection of B onto a subspace U of span of A_1 and A_2
- Orthogonality

$$\min_X \|E\|^2 = \min_X \|AX - B\|^2$$
$$X^* = \left(A^T A\right)^{-1} A^T B$$
$$B^* = AX^* = A\left(A^T A\right)^{-1} A^T B$$

$$A \perp (AX^* - B)$$
$$A^T (AX^* - B) = 0$$
$$A^T AX^* = A^T B$$
$$X^* = \left(A^T A\right)^{-1} A^T B$$
$$B^* = AX^* = A\left(A^T A\right)^{-1} A^T B$$



$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} \neq \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$