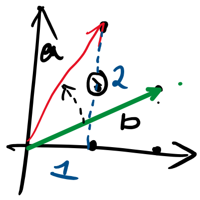
Dot product between two vectors

$$a = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \Rightarrow a \cdot b = 1*4 + 2*5 + 3*6$$
$$= 32$$

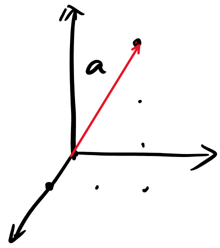
- What does the dot product tell us about the vectors?

- Geometric interpretation of a vector

$$a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$



$$a = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$



- Geometric length of a vector $\|a\|_2 = \sqrt{a_1^2 + a_2^2}$

(L2 norm of a) $= \sqrt{1^2 + 2^2}$

- Angle between vectors $\left\{ -1 \leq \dfrac{a \cdot b}{\|a\| \, \|b\|} \leq 1 \right.$ (Cauchy-Schwarz inequality)

$$\boxed{\cos(\angle a,b) = \frac{a \cdot b}{\|a\| \, \|b\|}}$$

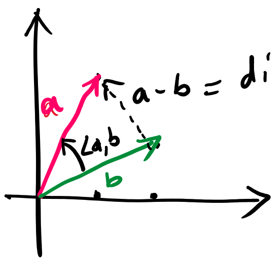
$\cos(0) = 1$
$\cos(\pi) = -1$
$\cos(\pi/2) = 0$

$a \cdot b = 0$
$\angle a,b = \pi/2$

$a \cdot b = -ve$
$\angle a,b > \pi/2$

$a \cdot b = +ve$
$\angle a,b < \pi/2$

- Distance between vectors



$a - b$ = difference vector $= \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

How big is the difference vector $= \|a - b\|_2$

$$= \left\| \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\| = \sqrt{(-1)^2 + 1^2}$$
$$= \sqrt{2}$$

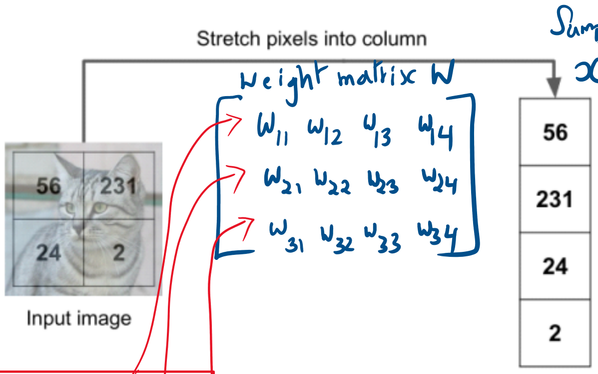
- An important relationship : $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$

$$\|a\|^2 = a \cdot a$$

$$\Downarrow \qquad \Downarrow$$

$$\left(\sqrt{a_1^2 + a_2^2}\right)^2 \Leftrightarrow a_1^2 + a_2^2$$

Example with an image with 4 pixels, and 3 classes (cat/dog/ship)



Sample vector x

Weight matrix W

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \end{bmatrix}$$

Weights for cats

$$w_1 = \begin{bmatrix} w_{11} \\ w_{12} \\ w_{13} \\ w_{14} \end{bmatrix}$$

weights for dogs

$$w_2 = \begin{bmatrix} w_{21} \\ w_{22} \\ w_{23} \\ w_{24} \end{bmatrix}$$

weights for ship

$$w_3 = \begin{bmatrix} w_{31} \\ w_{32} \\ w_{33} \\ w_{34} \end{bmatrix}$$

$$\boxed{\begin{aligned} &\text{Cat score} = w_1 \cdot x \\ &\text{Dog score} = w_2 \cdot x \\ &\text{ship score} = w_3 \cdot x \end{aligned}}$$

$\hookrightarrow$ matrix-vector product Wx

3x4-vector   4-vector

Why did we not do this?

$$\begin{matrix} w_1 \cdot x \\ w_2 \cdot x \\ w_3 \cdot x \end{matrix} \quad \overset{?}{=} \quad \underbrace{\begin{bmatrix} w_{11} & w_{21} & w_{31} \\ w_{12} & w_{22} & w_{32} \\ w_{13} & w_{23} & w_{33} \\ w_{14} & w_{24} & w_{34} \end{bmatrix}}_{W} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x} = \begin{bmatrix} w_1 \cdot x & w_2 \cdot x & w_3 \cdot x \end{bmatrix}$$

A matrix-vector product Wx = sequence of dot products

= dot products between the rows of matrix W (seen as vectors) and the vector x

Tensor-vector product

Matrix-vector product

$A \quad x$

Tensor-Vector product

E.g. $T \to (4, 3, 2)$ $\quad x \to (2,)$

Matrix ...

E.g. $A \to (3,4)$ $\quad x \to (4,)$

Linear combination of columns of a matrix A

1st row of A

2nd row of A

3rd row of A

$$\underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -2 & 3 \\ -1 & 1 & -2 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x} = \begin{cases} \begin{bmatrix} a^{(1)} \cdot x \\ a^{(2)} \cdot x \\ a^{(3)} \cdot x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ x_1 a_1 + x_2 a_2 + x_3 a_3 + x_4 a_4 = [\ ] \end{cases}$$

1st column of A

2nd column of A

3rd column of A

4th column of A

$x_1 a_1 + x_2 a_2 + x_3 a_3 + x_4 a_4$ = linear combination of columns of A using the elements of x as multipliers

Matrix-vector product

E.g. $A = \underbrace{\begin{bmatrix} \boxed{1 \quad 2 \quad 3} \\ \boxed{4 \quad 5 \quad 6} \end{bmatrix}}_{2\times 3}$, $x = \underbrace{\begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}}_{3}$ $= \underbrace{\begin{bmatrix} 1 \times 7 + 2 \times 8 + 3 \times 9 \\ 4 \times 7 + 5 \times 8 + 6 \times 9 \end{bmatrix}}_{2}$

$\parallel$

Ax is defined as $= \begin{bmatrix} (1\text{st row of } A) \cdot x \\ (2\text{nd row of } A) \cdot x \end{bmatrix} = \begin{bmatrix} a^{(1)} \cdot x \\ a^{(2)} \cdot x \end{bmatrix}$

Tensor-Vector product

E.g. $T = \begin{bmatrix} \boxed{\begin{bmatrix} 76 & 120 \\ 74 & 124 \\ 78 & 136 \end{bmatrix}} \\ \boxed{\begin{bmatrix} 78 & 132 \\ 80 & 124 \\ 90 & 120 \end{bmatrix}} \end{bmatrix}$ $\begin{matrix} \to \text{Timestamp-0} \\ \\ \to \text{Timestamp-1} \end{matrix}$, $x = \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{2}$

$2 \times 3 \times 2$ → features

Timestamps patients

$$\begin{bmatrix} 76 & 120 \\ 74 & 124 \\ 78 & 136 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 76 \\ 74 \\ 78 \end{bmatrix}$$

$$\begin{bmatrix} 78 & 136 \\ 80 & 124 \\ 70 & 120 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 78 \\ 80 \\ 70 \end{bmatrix}$$

$$\begin{bmatrix} 76 & 74 & 78 \\ 78 & 80 & 70 \end{bmatrix}$$

2×3

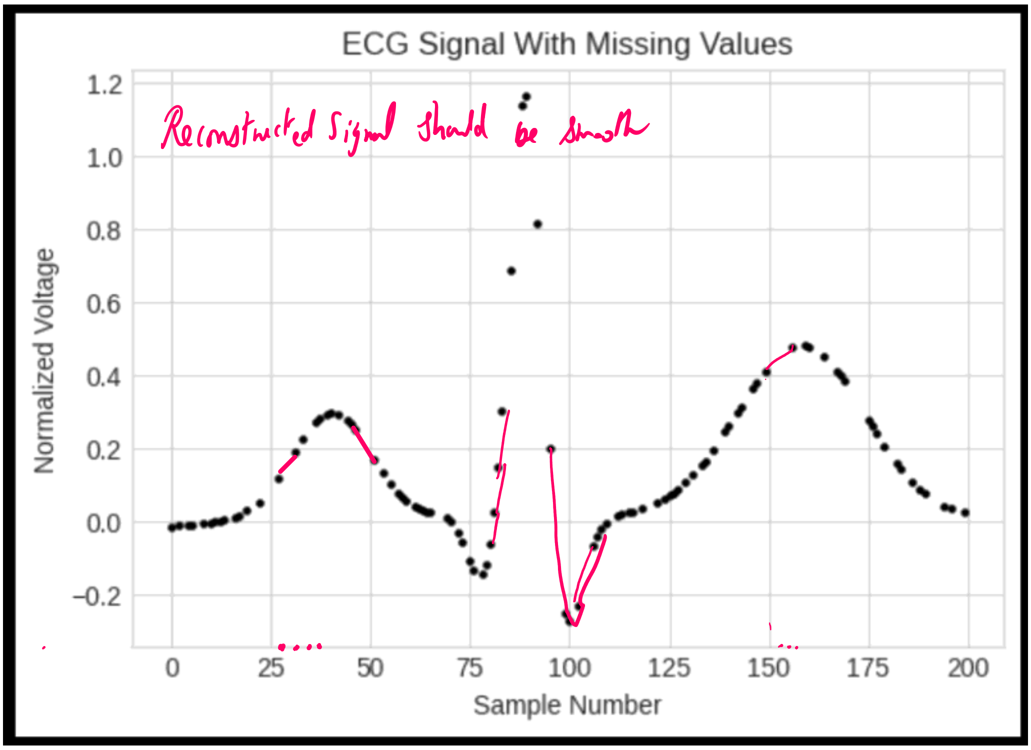
Matrix-Matrix Product

E.g. $A = \begin{bmatrix} \boxed{1 \quad 2 \quad 3} \\ \boxed{4 \quad 5 \quad 6} \end{bmatrix}$. $B = \begin{bmatrix} 7 & 10 \\ 8 & 11 \\ 9 & 12 \end{bmatrix}$ $= \begin{bmatrix} a^{(1)} \cdot b_1 & a^{(1)} \cdot b_2 \\ a^{(2)} \cdot b_1 & a^{(2)} \cdot b_2 \end{bmatrix}$

2×3 $\quad$ 3×2 $\quad$ 2×2

1st row of A $\quad$ 1st column of B

2nd row of A $\quad$ 1st column of B



$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} . \quad x = \begin{bmatrix} x_1 \\ x_2 \\ v_1 \\ x_4 \\ v_2 \\ v_3 \end{bmatrix}$$

missing ECG values

$$\begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ v_1 \\ x_4 \\ v_2 \\ v_3 \end{bmatrix}$$

Transposes

Filter matrix for known ECG values

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

ECG vector with unknown values turned to zeros

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \\ x_4 \\ 0 \\ 0 \end{bmatrix}$$

vector of known ECG values

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{S_1} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix}}_{x_{known}} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{S_2} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}}_{x_{unknown}} = \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ v_1 \\ x_4 \\ v_2 \\ v_3 \end{bmatrix}}_{x}$$

x Full ECG vector with missing values

$$\begin{bmatrix} x_1 \\ x_2 \\ 0 \\ x_4 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ v_1 \\ 0 \\ v_2 \\ v_3 \end{bmatrix}$$

$$\Rightarrow \boxed{\underbrace{x}_{\text{ECG vector}} = \underbrace{S_1 \, x_{known}}_{\text{known part of the ECG vector}} + \underbrace{S_2 \, x_{unknown}}_{\text{unknown part of the ECG vector}}}$$

$(y_1 - y_2) - (y_2 - y_3)$

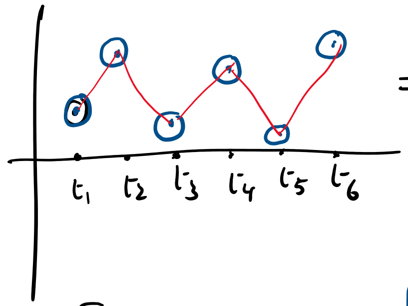
Aşile

vector vector $(y_1-y_2)-(y_2-y_3)$

Aside

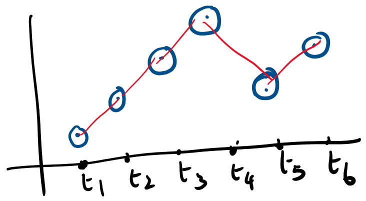
$$\underbrace{\begin{bmatrix} \boxed{1 & -2 & 1} & 0 & 0 & 0 \\ 0 & \boxed{1 & -2 & 1} & 0 & 0 \\ 0 & 0 & \boxed{1 & -2 & 1} & 0 \\ 0 & 0 & 0 & \boxed{1 & -2 & 1} \end{bmatrix}}_{D} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} \boxed{y_1 - 2y_2 + y_3} \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ y_4 - 2y_5 + y_6 \end{bmatrix}$$

$$y = \begin{bmatrix} 15 \\ 30 \\ 10 \\ 25 \\ 5 \\ 30 \end{bmatrix} \Rightarrow Dy = \begin{bmatrix} 15 - 60 + 10 \\ 30 - 20 + 25 \\ 10 - 50 + 5 \\ 25 - 10 + 30 \end{bmatrix} = \begin{bmatrix} -35 \\ 35 \\ -35 \\ 45 \end{bmatrix}$$

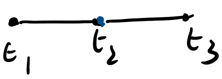


$\Rightarrow$ oscillatory signal

$$y = \begin{bmatrix} 15 \\ 20 \\ 25 \\ 30 \\ 20 \\ 25 \end{bmatrix}, \quad Dy = \begin{bmatrix} 15 - 40 + 25 \\ 20 - 50 + 30 \\ 25 - 60 + 20 \\ 30 - 40 + 25 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -15 \\ 15 \end{bmatrix}$$



Discrete version of second derivative



(Approx)

$$\frac{y_2 - y_1}{t_2 - t_1} = \frac{\text{change in temperature}}{\text{change in time}} = \frac{°c}{Hr}$$

Sensitivity of temperature w.r.t. time

$$y_3 - y_2 = \frac{°c}{}$$

$$\frac{y_3 - y_2}{t_3 - t_2} = \frac{\cdot c}{Hr}$$

Detour

$$x = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boxed{\frac{\|x\|}{\sqrt{n}}} \xrightarrow{} \mathrm{Rms}(x) \quad = \frac{\sqrt{1^2+1^2+1^2+1^2}}{\sqrt{4}} = \frac{\sqrt{4}}{\sqrt{4}} = 1$$

$$x = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad \frac{\|x\|}{\sqrt{n}} = 1$$

Reconstruction goal { the reconstructed ECG signal should be as smooth as possible

$$\mathrm{RMS}\left(D\left(\underbrace{S_1 x_{known} + S_2 x_{unknown}}_{\text{ECG vector}}\right)\right)$$ to be as

small as possible $\Rightarrow$ $$\frac{\|D(S_1 x_{known} + S_2 x_{unknown})\|_2}{\underbrace{\sqrt{n}}_{\text{constant}}}$$

Should be as small as possible

$\Rightarrow$ Find the vector $x_{unknown}$ s.t. $\boxed{\|D(S_1 x_{known} + S_2 x_{unknown})\|_2^2}$

is as small as possible.

$b = 4$-vector $\quad A = 4\times 3$-matrix $\quad$ unknown vector $\cancel{x} = 3$-vector

$$\Rightarrow \boxed{\underbrace{D S_1}_{4\times 3} x_{known}} + \boxed{D S_2}\,\boxed{x_{unknown}} \Rightarrow \|b + Ax\|_2^2$$

D: 4×6, S_1: 6×3, x_{known}: 3×1; D: 4×6, S_2: 6×3, $x_{unknown}$: 3×1

Not all linear systems of equations $Ax = b$ are

Solvable

$$\begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$
\begin{bmatrix}
1 & -2 & 1 & 0 & 0 & 0 \\
0 & 1 & -2 & 1 & 0 & 0 \\
0 & 0 & 1 & -2 & 1 & 0 \\
0 & 0 & 0 & 1 & -2 & 1 \\
0 & 0 & 0 & 0 & 1 & -2
\end{bmatrix}
$$

D matrix

Standard least squares problem

Find a vector x s.t. $\|Ax - b\|_2^2$ is minimized

A: Known matrix; b: Known vector

In Python, we say linalg.lstsq(A, b) to get the solution for the least squares problem.

For the ECG data, recall that we wanted to minimize

$$\| \underbrace{DS_1 x_{known}}_{-b} + \underbrace{DS_2}_{A} \underbrace{x_{unknown}}_{x} \|_2^2$$

Match A and b in the expression above with $\|Ax - b\|_2^2$

Solving systems of equations

Consider the following model for opinion formation among n individuals, each of whom interact with a certain number of individuals in the group. The numerical value of the ith person's opinion is denoted as x_i. The value of x_i is influenced by the following:

- The ith person's self opinion denoted as s_i
- The opinions of the remaining individuals x_j, where $j = 1, 2, \dots, n$ and $j \neq i$.

Assuming that the ith person gives a weightage w_{ij} to the jth person's opinion, we can compute x_i as follows:

$$x_i = \frac{s_i + \sum_{j \neq i} w_{ij} x_j}{1 + \sum_{j \neq i} w_{ij}}, \quad i = 1, \dots, n.$$

1. From the equation above, what do you see is the weightage that a person gives to their own opinion?
2. The equation above can be written as $(A + I)x = s$, where A is an $n \times n$-matrix and I represents the identity matrix. What are the elements of the matrix A, vectors x and s?

Scenario 4 people in a network

Self-opinion vector $= s = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix}$ Can be calculated based on individual history.

Opinion vector $= x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ that is built after interaction with others

1

weightage that the 1st individual gives to the 2nd individual's opinion

Weights matrix = W = $\begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \end{bmatrix}$

$\ldots \times 4]$ … weightage … gives to the 2nd individual's opinion

Weights matrix $= W = \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$

$$x_i = \frac{s_i + \sum_{j \neq i} w_{ij} x_j}{1 + \sum_{j \neq i} w_{ij}},$$

Suppose $i = 1$, the 1st individual

Opinion of 1st individual $x_1 = \dfrac{s_1 + w_{12}x_2 + w_{13}x_3 + w_{14}x_4}{\underbrace{w_{11}}_{=1} + w_{12} + w_{13} + w_{14}}$

$$\Rightarrow x_1 \underbrace{w_{11}}_{=1} + \boxed{x_1 w_{12}} + x_1 w_{13} + x_1 w_{14} = s_1 + w_{12}x_2 + w_{13}x_3 + w_{14}x_4$$

$$\Rightarrow \underbrace{x_1}_{\text{unknown}}(1 + \underbrace{w_{12} + w_{13} + w_{14}}_{\text{Known}}) + \underbrace{(-w_{12})}_{\text{Known}}\underbrace{x_2}_{\text{unknown}} + \underbrace{(-w_{13})}_{\text{Known}}\overset{\text{unknown}}{x_3} + \underbrace{(-w_{14})}_{\text{Known}}\overset{\text{unknown}}{x_4} = \underbrace{s_1}_{\text{Known}}$$

$$\Rightarrow \begin{cases} w_{12}(x_1 - x_2) + w_{13}(x_1 - x_3) + w_{14}(x_1 - x_4) + 1.x_1 = s_1 \\ w_{21}(x_2 - x_1) + 1.x_2 + w_{23}(x_2 - x_3) + w_{24}(x_2 - x_4) = s_2 \\ w_{31}(x_3 - x_1) + w_{32}(x_3 - x_2) + 1.x_3 + w_{34}(x_3 - x_4) = s_3 \\ w_{41}(x_4 - x_1) + w_{42}(x_4 - x_2) + w_{43}(x_4 - x_3) + 1.x_4 = s_4 \end{cases}$$

Given the weights matrix W and the self-opinion vector s, we get 4 equations in 4 unknowns

Recap of the matrix-vector product

A=
```
[[ 1  2 -1 -1]
 [ 2  4 -2  3]
 [-1  1 -2  4]]
```
, $x = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$

$$Ax = \begin{bmatrix} a^{(1)} \cdot x \\ a^{(2)} \cdot x \\ a^{(3)} \cdot x \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{s}$$

$a^{(1)} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \end{bmatrix}$

Linear combination of the columns of A using the elements of x as multipliers

$a_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, a_2 = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}, a_3 = \begin{bmatrix} -1 \\ -2 \\ -2 \end{bmatrix}, a_4 = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$

In general, if

In general, if

A: m×n-matrix

x: n-vector

Then

$$Ax = \begin{cases} \begin{bmatrix} a^{(1)}\cdot x \\ a^{(2)}\cdot x \\ \vdots \\ a^{(m)}\cdot x \end{bmatrix} & \text{m dot products} \\ x_1 a_1 + x_2 a_2 + \cdots + x_n a_n & \text{linear combination} \end{cases}$$

$a_1 = \begin{bmatrix} \vdots \\ 2 \\ -1 \end{bmatrix}, a_2 = \begin{bmatrix} \vdots \\ 4 \\ 1 \end{bmatrix}, a_3 = \begin{bmatrix} \vdots \\ -2 \\ -2 \end{bmatrix}, a_4 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$\Rightarrow x_1 a_1 + x_2 a_2 + x_3 a_3 + x_4 a_4$

$= -1\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + (1)\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} + (1)\begin{bmatrix} -1 \\ -2 \\ -2 \end{bmatrix} + (0)\begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$

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

Same answer

System of equations with

Solution exists means the System is consistent. otherwise the system is inconsistent

(1) no solution: $\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 5 \end{cases}$ no solution

(2) unique solution: $\begin{cases} x_1 + x_2 = 1 \\ x_1 + 2x_2 = 1 \end{cases} \Rightarrow \boxed{\begin{matrix} x_1 = 1 \\ x_2 = 0 \end{matrix}}$ unique solution

(3) Infinitely many Solutions: $\{ x_1 + x_2 = 1 \Rightarrow \boxed{\begin{matrix} x_1 = 1 \\ x_2 = 0 \end{matrix}}, \begin{bmatrix} x_1 = 0.5 \\ x_2 = 0.5 \end{bmatrix}, \begin{bmatrix} x_1 = 0.4 \\ x_2 = 0.6 \end{bmatrix}$ infinitely many Solutions

Elementary row operations

1. Divide/multiply a row by a non zero constant.
2. Subtract a scalar multiple of one row from another row.
3. Exchange two rows.

$\boxed{\begin{matrix} x_1 + x_2 = 1 \\ x_1 + 2x_2 = 1 \end{matrix}} \Longrightarrow \begin{matrix} 2(x_1 + x_2) = 2 \\ x_1 + 2x_2 = 1 \end{matrix} \Rightarrow \boxed{\begin{matrix} 2x_1 + 2x_2 = 2 \\ x_1 + 2x_2 = 1 \end{matrix}}$

$\boxed{\begin{matrix} x_1 + x_2 = 1 \\ x_1 + 2x_2 = 1 \end{matrix}} \Longrightarrow \boxed{\begin{matrix} x_1 + 2x_2 = 1 \\ x_1 + x_2 = 1 \end{matrix}}$

$\boxed{\begin{matrix} x_1 + x_2 = 1 \\ x_1 + 2x_2 = 1 \end{matrix}} \xrightarrow{Eq① = Eq① + 2 * Eq②} \boxed{\begin{matrix} 3x_1 + 5x_2 = 3 \\ x_1 + 2x_2 = 1 \end{matrix}}$

Reduced row echelon form (RREF)

1st column
2nd column
4th column

$\begin{bmatrix} 1 & 1 & 0 & 0 & | & 2 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & 1 & 0 & | & 1 \end{bmatrix}$

$$\underbrace{\left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 2 \\ 2 & 2 & 0 & 1 & 6 \\ 0 & 1 & -1 & 1 & 3 \\ -1 & -2 & 1 & 1 & -1 \end{array}\right]}_{\text{Original augmented matrix}} \longrightarrow \underbrace{\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]}_{\text{rref}}$$

$\Downarrow$

x_1, x_2, x_4 as the pivot variables and x_3 is the free variable

$$\Downarrow$$

$$\begin{cases} x_1 + x_2 = 2 \\ 2x_1 + 2x_2 + x_4 = 6 \\ x_2 - x_3 + x_4 = 3 \\ -x_1 - 2x_2 + x_3 + x_4 = -1 \end{cases} \Rightarrow \begin{cases} x_1 + x_3 = 1 \\ x_2 - x_3 = 1 \\ x_4 = 2 \\ \cancel{0 = 0} \end{cases}$$

anything about x_3 here? Nothing! x_3 can be anything

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 - x_3 \\ 1 + x_3 \\ x_3 \\ 2 \end{bmatrix} \Rightarrow \underset{\substack{\Downarrow \\ \text{unknown} \\ \text{vector}}}{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\Rightarrow \left\{ x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \;\middle|\; x_3 \in \mathbb{R} \right\}$$

such that

No solution case

$$x_1 + x_2 = 1$$
$$x_1 + x_2 = 5$$

$$\Rightarrow \text{Augmented matrix} = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 1 & 5 \end{array}\right]$$

$$\xrightarrow{Row2 = Row2 - Row1} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 4 \end{array}\right]$$

Infinitely many solutions case

$$x_1 + x_2 = 1$$
$$2x_1 + 2x_2 = 2$$

$$\Rightarrow \text{Augmented matrix} = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 2 & 2 \end{array}\right]$$

$\Rightarrow$ Augmented matrix $= \left[\begin{array}{cc|c} \cdot & \cdot & \cdot \\ 2 & 2 & 2 \end{array}\right]$

$RREF = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$

Example , Solving $Ax = b$

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 1 \\ 2 & 4 & -2 & 3 & 3 \\ -1 & 1 & -2 & 4 & 2 \end{array}\right]$$

A, b

Original augmented matrix

```
(Matrix([
[1, 0,  1, 0, -2/5],
[0, 1, -1, 0,  4/5],
[0, 0,  0, 1,  1/5]]), (0, 1, 3))
```

RREF using Sympy

Solution

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -x_3 - \frac{2}{5} \\ x_3 + \frac{4}{5} \\ x_3 \\ \frac{1}{5} \end{bmatrix} = \underbrace{x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}}_{\text{solution to } Ax=0} + \underbrace{\begin{bmatrix} -\frac{2}{5} \\ \frac{4}{5} \\ 0 \\ \frac{1}{5} \end{bmatrix}}_{\text{particular solution}}$$

Solution vector x

RHS vector b

$$A\left(x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -2/5 \\ 4/5 \\ 0 \\ 1/5 \end{bmatrix}\right) \stackrel{?}{=} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

$$A\left(x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}\right) + A \begin{bmatrix} -2/5 \\ 4/5 \\ 0 \\ 1/5 \end{bmatrix}$$

$$= x_3 * A \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + A \begin{bmatrix} -2/5 \\ 4/5 \\ 0 \\ 1/5 \end{bmatrix}$$

$$x_3 * \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

ALA Sessional-2 Review

(i) Matrices : matrix-vector and matrix-matrix products

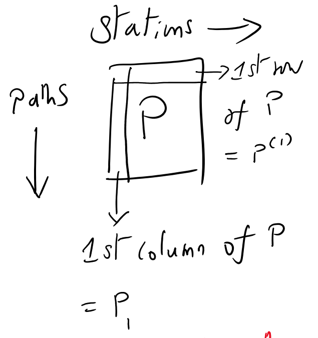
Questions will be scenario-based

Questions will be scenario-based

E.g. 5 stations and 8 paths

$S_1 \; S_2 \; S_3 \; S_4 \; S_5$

$\left\{ \begin{matrix} S_1 \; S_4 \; S_5 \\ S_2 \; S_3 \; S_4 \\ S_1 \; S_2 \; S_3 \; S_4 \end{matrix} \right\}$



P: 8x5-matrix

P_1: 8-vector, E.g $P_1 = \begin{array}{c} \text{station-1} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{array}$

station-1 is in the first path

station-2 is in the second path

station-1 is not in the third path

$P^{(1)}$: 5-vector, E.g $P^{(1)} = \begin{array}{c} \text{Path-1} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array}$

Path-1 contains station-1

Path-1 does not contain station-2

E.g.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \times 7 + 2 * 8 + 3 * 9 \\ 4 * 7 + 5 * 8 + 6 * 9 \end{bmatrix}$$

The 8x5-matrix P

$\mathbf{1}$-vector $= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $P\mathbf{1}$ is a special matrix-vector Product

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

P_1, $P^{(1)T}$

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

$$= \begin{bmatrix} 3 \\ 2 \\ 2 \\ 4 \\ 3 \\ 0 \\ 5 \\ 2 \end{bmatrix}$$

No. of stations in Path-1

No. of stations in Path-2

E.g. What about the matrix-vector Product $P^T \mathbf{1}$?

P^T: 5x8-matrix

$\mathbf{1}$: 8-vector

$$\underbrace{P^T}_{5\times 8\text{-matrix}} \; \underbrace{1}_{8\text{-vector}}$$

$(\underbrace{P^T 1}_{\text{vector}})_1$ = 1st component of the vector
= How many paths in which station-1 shows up

$(\underbrace{P 1}_{\text{vector}})_3$ = 3rd component of the vector
= How many stations are in the 3rd path

Recall the unit vectors: $e_1 = \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}$, $e_1 = \begin{bmatrix}1\\0\\0\\0\\0\\0\\0\\0\end{bmatrix}$

$$Pe_1 = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix}1 \checkmark\\0\\0\\0\\0\end{bmatrix} = 1 * \begin{bmatrix}1\\1\\0\\1\\0\\0\\1\\0\end{bmatrix} + 0 * \cdots$$

= Paths in which station-1 shows up $= P_1 =$ 1st column of P

Recall example

$$P = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad x = \begin{bmatrix}7\\8\\9\end{bmatrix}$$

$$Px = \begin{cases} \begin{bmatrix} \begin{bmatrix}1\\2\\3\end{bmatrix} \cdot \begin{bmatrix}7\\8\\9\end{bmatrix} \\ \begin{bmatrix}4\\5\\6\end{bmatrix} \cdot \begin{bmatrix}7\\8\\9\end{bmatrix} \end{bmatrix} & \text{Dot-product version} \\ 7\begin{bmatrix}1\\4\end{bmatrix} + 8\begin{bmatrix}2\\5\end{bmatrix} + 9\begin{bmatrix}3\\6\end{bmatrix} & \text{linear combination version} \end{cases}$$

$$e_1^T \underbrace{P e_3}_{} = \, ?$$

Pe_3: matrix-vector product $= P_3 =$ 3rd column of matrix P
= Paths in which station-3

matrix-vector product $= P_3 = $ 3rd column of matrix P

$=$ Paths in which station-3 shows up

$$\underbrace{e_1^T P_3}_{\text{matrix-vector product}} = P_3^T e_1 = \underbrace{P_3 \cdot e_1 = e_1 \cdot P_3}_{\text{Dot-product}} = \text{1st component of the vector } P_3$$

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

$$P_3 \cdot e_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0$$

In general
$$\begin{cases} Pe_i = P_i = i\text{th column of matrix } P \\ e_i^T P e_j = (i,j)\text{th element of matrix } P \\ e_i^T P = p^{(i)} = i\text{th row of matrix } P \end{cases}$$

Properties of matrix-vector product: $(AB)^T = B^T A^T$, $(A^T)^T = A$

$$e_i^T P = (\underset{A}{P^T} \underset{B}{e_i})^T = \underset{B^T}{e_i^T} \underset{A^T}{(P^T)^T} = e_i^T P$$

ith column of $P^T = i$th row of P

Build a matrix-vector product from a description

E.g. a signal sampled at 6 time stamps

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}.$$

Samples at only the even time stamps

$$= \underset{\text{Sampling matrix}}{S}\ \underset{\text{signal}}{x} = \begin{bmatrix} x_2 \\ x_4 \\ x_6 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_4 \\ x_6 \end{bmatrix}$$

What should be the matrix S such that we weight three successive timestamps equally to generate a new value.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} \rightarrow \begin{array}{l} \frac{1}{3} \cdot x_1 + \frac{1}{3} \cdot x_2 + \frac{1}{3} \cdot x_3 \\ \frac{1}{3} x_2 + \frac{1}{3} x_3 + \frac{1}{3} x_4 \\ ? \\ ? \end{array}$$

$$\underbrace{\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}}_{\Downarrow \; S} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}}_{\Downarrow \; x}$$

Matrix-Matrix Product

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & 10 \\ 8 & 11 \\ 9 & 12 \end{bmatrix}$$

2×3 $\qquad$ 3×2

matrix-vector product $\qquad$ matrix-vector product

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

2×2

$$P = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}, \quad P^T = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}$$

8 × 5 - matrix   5 × 8 - matrix

Which matrix- matrix product makes sense : P^TP ✓ or PP^T ✓ or P^2 ?

P^TP: 5×8   8×5 ⇓ 5×5 -matrix

PP^T: 8×5   5×8 ⇓ 8×8 -matrix

P^2: $P \cdot P$: 8×5   8×5 — inner dimensions do not match

$$[AB]_{i,j} = a^{(i)} \cdot b_j = a^{(i)T} b_j$$

$a^{(i)}$ ⇓ ith row of A b_j ⇓ jth column of B

$$[P^TP]_{ij} = (\text{ith row of } P^T) \cdot (\text{jth column of } P)$$

$$\Downarrow$$

$$(\text{ith column of } P) \cdot (\text{jth column of } P)$$

$$= P_i \cdot P_j = P_i^T P_j$$

for e.g $i=1, j=2 \Rightarrow [P^TP]_{1,2} = P_1 \cdot P_2$

$$\begin{array}{cc} P_1 & P_2 \\ \uparrow & \uparrow \end{array}$$

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

$P_1 \cdot P_2 = 1$ = no. of paths common to station-1 and station-2

How about $P_1 - P_2$? $P_1 - P_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

RREF

```
[1, 0,  1, 0,| -2/5],
[0, 1, -1, 0,|  4/5],
```

Always do this first

Is the system Consistent

or

$$\left[\begin{array}{cccc|c} 1, & 0, & 1, & 0, & -2/5 \\ 0, & 1, & -1, & 0, & 4/5 \\ 0, & 0, & 0, & 0, & 1/5 \end{array}\right]$$

Consistent or inconsistent?

$0 = 1/5$, not possible $\Rightarrow$ System is inconsistent

$$\begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & \\ \end{array}$$
$$\left[\begin{array}{cccc|c} 1, & 0, & 1, & 0, & -2/5 \\ 0, & 1, & -1, & 0, & 4/5 \\ 0, & 0, & 0, & 1, & 1/5 \end{array}\right]$$

x_1, x_2, x_4 = pivot variables

x_3 = free variable

$$\begin{aligned} x_1 + x_3 &= -2/5 \\ x_2 - x_3 &= 4/5 \\ x_4 &= 1/5 \end{aligned} \Rightarrow \begin{bmatrix} x_1 = \\ x_2 = \\ x_3 = \\ x_4 = \end{bmatrix} \begin{bmatrix} -2/5 - x_3 \\ 4/5 + x_3 \\ x_3 \\ 1/5 \end{bmatrix}$$

$$\Rightarrow x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 4/5 \\ 0 \\ 1/5 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \text{ where } x_3 \in \mathbb{R}$$

This system has infinitely many solutions, but there is an easy choice of x_3 which is equal to 0.

$$\begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -2 & 3 \\ -1 & 1 & -2 & 4 \end{bmatrix} \left(\begin{bmatrix} -2/5 \\ 4/5 \\ 0 \\ 1/5 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

Coefficient matrix A — Solution vector x — Right hand side vector b

$$\underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -2 & 3 \\ -1 & 1 & -2 & 4 \end{bmatrix} \begin{bmatrix} -2/5 \\ 4/5 \\ 0 \\ 1/5 \end{bmatrix}}_{\text{Matrix-vector product } = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}} + x_3 \underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -2 & 3 \\ -1 & 1 & -2 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}}_{\text{matrix-vector product } = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}$$

$$\underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -2 & 3 \\ -1 & 1 & -2 & 4 \end{bmatrix}}_{A} \left(x_3 \underbrace{\begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}}_{\text{solution to } Ax=0} + \underbrace{\begin{bmatrix} -\frac{2}{5} \\ \frac{4}{5} \\ 0 \\ \frac{1}{5} \end{bmatrix}}_{\text{particular solution}} \right)$$

Solution vector
and there are infinitely
many of them

$$= \underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -2 & 3 \\ -1 & 1 & -2 & 4 \end{bmatrix}}_{A} \left(x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right) + \underbrace{\underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -2 & 3 \\ -1 & 1 & -2 & 4 \end{bmatrix}}_{A} \begin{bmatrix} -2/5 \\ 4/5 \\ 0 \\ 1/5 \end{bmatrix}}$$

$$= x_3 \underbrace{\underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -2 & 3 \\ -1 & 1 & -2 & 4 \end{bmatrix}}_{A} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}} + \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

$$= x_3 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

Linear combination of vectors

$$a_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

One linear combination: $0 * \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0 * \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Multipliers

Another linear combination: $1 * \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0 * \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

One-more linear combination: $1*\begin{bmatrix}1\\2\end{bmatrix} + 1*\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}3\\3\end{bmatrix}$

One last linear combination: $\underbrace{?*\begin{bmatrix}1\\2\end{bmatrix} + ?*\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}}_{\text{Matrix-vector product}}$

$$\begin{bmatrix}1 & 2\\2 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

R.H.S. vector is special, it is the zero vector

Augmented matrix $= \left[\begin{array}{cc|c}1 & 2 & 0\\2 & 1 & 0\end{array}\right] \xrightarrow[\text{operations}]{\text{Elementary row}} \begin{array}{c} \begin{array}{cc} x_1 & x_2 \end{array} \\ \left[\begin{array}{cc|c}1 & 0 & 0\\0 & 1 & 0\end{array}\right] \end{array}$

RREF

$\Rightarrow \boxed{x_1 = 0 \quad \text{and} \quad x_2 = 0}$

The only multipliers that can result in a zero linear combination of the vectors $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}2\\1\end{bmatrix}$ are zeros. In other words,

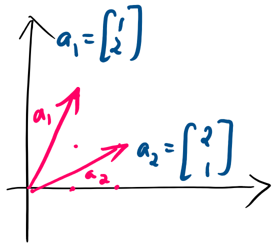
$A x = 0$ is possible only when $x = \begin{bmatrix}0\\0\end{bmatrix}$, the zero vector

$A \to \begin{bmatrix}1 & 2\\2 & 1\end{bmatrix}$, $0 \to \begin{bmatrix}0\\0\end{bmatrix}$

a_1 a_2

$\Rightarrow$ The vectors a_1 and a_2 (columns of matrix A) are said to be linearly independent.

Geometric perspective of linear independence



The two vectors are not aligned

E.g. $a_1 = \begin{bmatrix}72\\76\\\ldots\end{bmatrix}$ (HR), $a_2 = \begin{bmatrix}144\\152\\1\ldots\end{bmatrix}$ (BP)

If $a_1 = \begin{bmatrix} 72 \\ 76 \\ 80 \end{bmatrix}$, $a_2 = \begin{bmatrix} 144 \\ 152 \\ 160 \end{bmatrix}$

Are a_1 and a_2 linearly independent?

$$Ax = 0 \Rightarrow \begin{bmatrix} 72 & 144 \\ 76 & 152 \\ 80 & 160 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Augmented matrix

$$\left[\begin{array}{cc|c} 72 & 144 & 0 \\ 76 & 152 & 0 \\ 80 & 160 & 0 \end{array}\right] \xrightarrow[\text{row operations}]{\text{Elementary}} \begin{array}{c} \begin{array}{cc} x_1 & x_2 \end{array} \\ \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \end{array}$$

RREF

$$\begin{aligned} x_1 + 2x_2 = 0 &\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix} \\ x_2 = x_2 &\Rightarrow \end{aligned}$$

Choose the free variable $x_2 = 1 \Rightarrow x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$$\Rightarrow x_1 a_1 + x_2 a_2 = -2a_1 + 1 \cdot a_2 = 0$$

$$\Rightarrow \boxed{a_2 = 2a_1} \Rightarrow \boxed{\text{BP} = 2 * \text{HR is what we understand}}$$

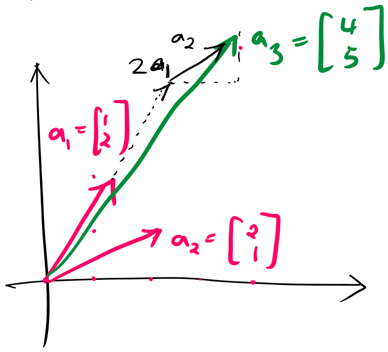
The vectors a_1 and a_2 are linearly dependent

Suppose we had HR (BPM), BP (mm Hg), and TEMP (°c). We observe that

$$\boxed{\text{BP} = 1.2 * \text{HR} + 1.05 * \text{TEMP}}$$

$$\boxed{\begin{array}{l} \text{Haemoglobin} = \beta_0 + \beta_1 * \text{HR} + \beta_2 * \text{BP} + \beta_3 * \text{TEMP} \\ \text{Linear model} \end{array}}$$

E.g. $\underbrace{\begin{bmatrix}1 & 2 & 4\\ 2 & 1 & 5\end{bmatrix}}_{A}$ with columns a_1, a_2, a_3; $\underbrace{\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}}_{x} = \underbrace{\begin{bmatrix}0\\0\end{bmatrix}}_{0 \text{ vector}} \Rightarrow x_1a_1 + x_2a_2 + x_3a_3 = 0$



$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 0\\ 0 & 1 & 1 & 0\end{array}\right] = \text{RREF of the augmented matrix}$$

x_1 = pivot, x_2 = pivot, x_3 = free

$$\Rightarrow \begin{aligned} x_1 + 2x_3 = 0 &\Rightarrow x_1 = -2x_3\\ x_2 + 1\cdot x_3 = 0 &\Rightarrow x_2 = -x_3\\ x_3 = x_3 &\Rightarrow x_3 = x_3 \end{aligned}$$

$$\Rightarrow \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}-2x_3\\-x_3\\x_3\end{bmatrix} = x_3\begin{bmatrix}-2\\-1\\1\end{bmatrix}$$

Easy choice for $x_3 = 1 \Rightarrow x = \begin{bmatrix}-2\\-1\\1\end{bmatrix} \Rightarrow -2a_1 + (-1)a_2 + a_3 = 0$

$\Rightarrow \boxed{a_3 = 2a_1 + a_2}$

E.g. $A = \begin{bmatrix}1 & 4 & 7 & 10\\ 2 & 5 & 8 & 11\\ 3 & 6 & 9 & 12\end{bmatrix}$, are the columns linearly dependent or independent?

Solving $Ax = 0$ (RHS vector is the zero vector)

$$\text{RREF} = \left[\begin{array}{cccc|c} 1 & 0 & -1 & -2 & 0\\ 0 & 1 & 2 & 3 & 0\\ 0 & 0 & 0 & 0 & 0\end{array}\right] \Rightarrow \begin{aligned} x_1 - x_3 - 2x_4 &= 0\\ x_2 + 2x_3 + 3x_4 &= 0\end{aligned}$$

$\underbrace{x_1 \quad x_2}_{\text{Pivots}} \quad \underbrace{x_3 \quad x_4}_{\text{free}}$

$$\Rightarrow \begin{aligned} x_1 &= x_3 + 2x_4\\ x_2 &= -2x_3 - 3x_4\\ x_3 &= x_3\\ x_4 &= x_4\end{aligned}$$

$$\Rightarrow x = \begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = x_3\begin{bmatrix}1\\-2\\1\\0\end{bmatrix} + x_4\begin{bmatrix}2\\-3\\0\\1\end{bmatrix}$$

Easy choices for x_3, x_4 are $x_3 = 1$ and $x_4 = 1$

$$\Rightarrow \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ -5 \\ 1 \\ 1 \end{bmatrix} \Rightarrow 3a_1 - 5a_2 + 1 \cdot a_3 + 1 \cdot a_4 = 0$$

$\Rightarrow$ Columns of A are linearly dependent

For an $m \times n$ - matrix

(i) $m > n$ (more rows than columns)

no. of pivots = at most n

E.g. $m = 5, n = 2$, RREF =
$$\begin{array}{cc|c} x_1 & x_2 & \\ \end{array}$$
$$\left[\begin{array}{cc|c} 1 & 0 & * \\ 0 & 1 & * \\ 0 & 0 & * \\ 0 & 0 & * \\ 0 & 0 & * \end{array}\right]$$

(2) $m < n$ (more columns than rows)

no. of pivots = at most m

E.g. $m = 2, n = 5$, RREF =
$$\begin{array}{cc} x_1 & x_2 \end{array}$$
$$\left[\begin{array}{ccccc|c} 1 & 0 & * & * & * & * \\ 0 & 1 & * & * & * & * \end{array}\right]$$

(3) $m = n$ (Same no. of rows and columns)

no. of pivots = at most m = at most n

E.g. $m = 5, n = 5$, RREF =
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \longrightarrow$$

```
[1, 0, -1, -2, 0]
[0, 1,  2,  3, 0]
[0, 0,  0,  0, 0]
```

Coefficient matrix

RREF

$$\begin{array}{l} x_1 - x_3 - 2x_4 = 0, \\ x_2 + 2x_3 + 3x_4 = 0, \end{array} \Rightarrow x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_3 + 2x_4 \\ -2x_3 - 3x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

$$a_1 - 2a_2 + a_3 = 0 \Rightarrow a_3 = -a_1 + 2a_2.$$

$$2a_1 - 3a_2 + a_4 = 0 \Rightarrow a_4 = -2a_1 + 3a_2.$$

Columns corresponding to the free variables can be written as a linear combination of columns corresponding to the pivot variables

$\Downarrow$

The only columns in A that matter are a_1 and a_2

Column space of matrix A denoted as C(A)

The set of all possible vectors that can be generated using a linear combination of the columns of A

$$A = \begin{array}{c} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \\ \begin{array}{cccc} \downarrow & \downarrow & \downarrow & \downarrow \\ a_1 & a_2 & a_3 & a_4 \end{array} \end{array}$$

$$C(A) = \{\alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3 + \alpha_4 a_4 \mid \alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}\}$$

$$= \{\alpha_1 a_1 + \alpha_2 a_2 \mid \alpha_1, \alpha_2 \in \mathbb{R}\}$$

Why do we need to define a 'column space' of a matrix?

$\boxed{\text{To solve } \underset{m\times n}{A}\, \underset{n}{x} = \underset{m}{b}}$ If this system is consistent, there is a set of values $x_1, x_2, \cdots, x_n$ s.t.

$$x_1 a_1 + x_2 a_2 + x_3 a_3 + \cdots + x_n a_n = b$$

E.g. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, solve $Ax = b$

$$\text{Augmented matrix} = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & \end{array} \\ \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right] \end{array} \Rightarrow \begin{array}{l} x_1 = 1 \\ x_2 = 2 \end{array}$$

Now we have $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, Solve $Ax = b$

Augmented matrix $= \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{array}\right]$ → Inconsistent

$\boxed{\text{Is } b \in C(A)?}$

$C(A) = \left\{ \alpha_1 a_1 + \alpha_2 a_2 \,\middle|\, \alpha_1, \alpha_2 \in \mathbb{R} \right\}$

$= \left\{ \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ 0 \end{bmatrix} \,\middle|\, \alpha_1, \alpha_2 \in \mathbb{R} \right\}$

<u>Can we find a 'compromise' solution</u>

- We want to solve $Ax = b$
- There is no s.t. $Ax = b$. Why? Because $b \notin C(A)$
- Let us find a compromise solution x s.t. $\|Ax - b\|^2$ is as small as possible

$$= \left\| \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\|^2$$

$$= \left\| \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \\ -3 \end{bmatrix} \right\|^2$$

$$\Rightarrow x_1 = 1, \; x_2 = 2$$

The least squares solution to $Ax = b$ is obtained by minimizing $\|Ax - b\|^2 \xrightarrow[\text{Calculation}]{\text{Gradient}} \boxed{x = A(A^TA)^{-1}A^Tb}$

E.g. $\|2x - 4\|^2$, $x = 2$

$$\left\| \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \end{bmatrix} \right\|^2, \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

- Projection of a vector onto the direction of another vector

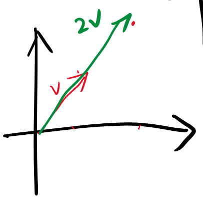
E.g. **Data matrix**

	HR	BP	Temp
Patient-1	76	126	38.0
Patient-2	74	120	38.0
Patient-3	72	118	37.5
Patient-4	78	136	37.0

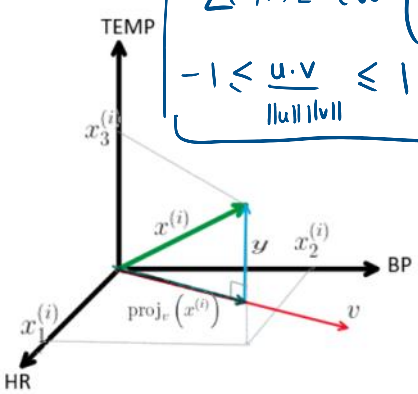
$$x^{(1)} = \begin{bmatrix} 76 \\ 126 \\ 38 \end{bmatrix}$$

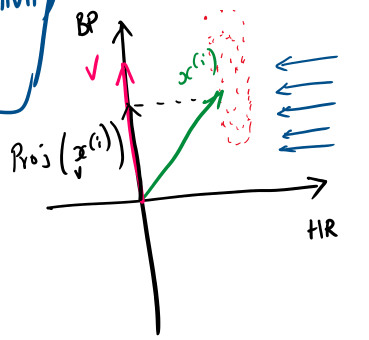
E.g.

$$v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad 2v = 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$



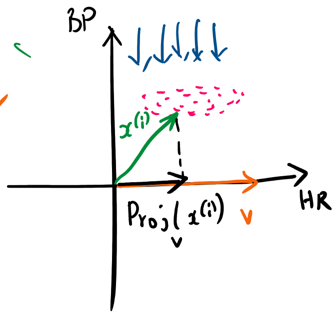
$$\boxed{\begin{aligned} &\angle(u,v) = \cos^{-1}\left(\frac{u \cdot v}{\|u\|\|v\|}\right) \\ &-1 \leq \frac{u \cdot v}{\|u\|\|v\|} \leq 1 \end{aligned}}$$

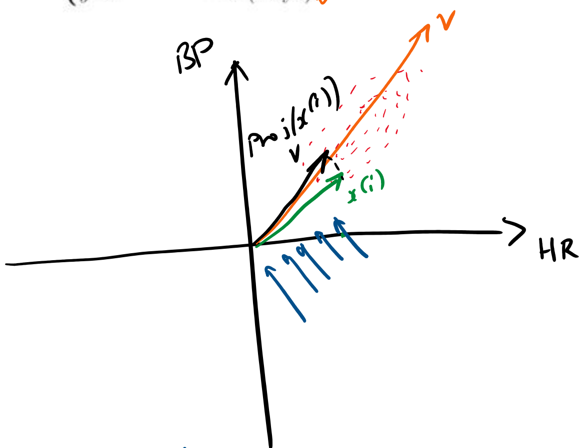




We can derive an expression for the projection as follows:

$$\begin{cases} \text{proj}_v\left(x^{(i)}\right) & = cv, \text{ for some unknown constant } c \text{ (why?)} \checkmark \\ y & = x^{(i)} - \text{proj}_v\left(x^{(i)}\right) \text{ (why?)} \checkmark \\ y \cdot v & = 0 \text{ (why?)} \checkmark \end{cases}$$





$y \cdot v = 0$ (y and v are orthogonal to each other)

$$\Rightarrow \left[x^{(i)} - \text{Proj}_v\left(x^{(i)}\right)\right] \cdot v = 0$$

$\rightarrow cv$

$$\Rightarrow x^{(i)} \cdot v - \overset{cv}{\mathrm{Proj}_v(x^{(i)})} \cdot v = 0$$

$$\Rightarrow x^{(i)} \cdot v - (cv) \cdot v = 0 \Rightarrow c(v \cdot v) = x^{(i)} \cdot v$$

$$\Rightarrow c = \frac{x^{(i)} \cdot v}{v \cdot v} \Rightarrow \boxed{\mathrm{Proj}_v\, x^{(i)} = cv = \left[\frac{x^{(i)} \cdot v}{v \cdot v}\right] v}$$

$$\boxed{v \cdot v = \|v\|^2}, \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$v \cdot v = v_1^2 + v_2^2$$

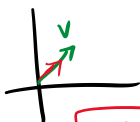
$$\|v\|^2 = \left(\sqrt{v_1^2 + v_2^2}\right)^2 = v_1^2 + v_2^2$$

$$\mathrm{Proj}_v(x^{(i)}) = \left(\frac{x^{(i)} \cdot v}{\|v\|}\right) \frac{v}{\|v\|}$$

$= \|v\|^2 = v \cdot v$

unit vector (direction)
length = 1

E.g. $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



$$\frac{v}{\|v\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad \left\|\frac{v}{\|v\|}\right\| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2} = 1$$

$$\mathrm{Proj}_v(x^{(i)}) = \boxed{\frac{x^{(i)} \cdot v}{\|v\|}}\ \boxed{\frac{v}{\|v\|}}$$

Magnitude of the projection (shadow length) — Direction of projection

Suppose the vector v is already a unit vector,

$$\boxed{\mathrm{Proj}_v(x^{(i)}) = (x^{(i)} \cdot v)\, v}$$

Block matrix-vector operations

$$\begin{bmatrix} 1 \cdot 5 \\ 2 \cdot 5 \\ 3 \cdot 5 \\ 4 \cdot 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} 5$$

$$\begin{bmatrix} \left(x^{(1)}\right)^{\mathrm{T}} v \\ \left(x^{(2)}\right)^{\mathrm{T}} v \\ \left(x^{(3)}\right)^{\mathrm{T}} v \\ \left(x^{(4)}\right)^{\mathrm{T}} v \end{bmatrix} = \begin{bmatrix} x^{(1)T} \\ x^{(2)T} \\ x^{(3)T} \\ x^{(4)T} \end{bmatrix} v = X v$$

$\Downarrow$ matrix

$\Downarrow$ vector

	HR	BP	Temp
Patient-1	76	126	38.0
Patient-2	74	120	38.0
Patient-3	72	118	37.5
Patient-4	78	136	37.0

Data matrix in terms of its rows

$$= X = \begin{bmatrix} x^{(1)T} \\ x^{(2)T} \\ x^{(3)T} \\ x^{(4)T} \end{bmatrix}$$

Data matrix in terms of its columns

$$= [x_1 \ x_2 \ x_3]$$

a = vector of heart rates

b = vector of BP

$\frac{a \cdot b}{\|a\|\|b\|}$ is guaranteed to be between -1 and 1

covariance between HR and BP

$$\boxed{\frac{a_{MC} \cdot b_{MC}}{\|a_{MC}\| \|b_{MC}\|}}$$

is also guaranteed to be between -1 and 1

Correlation coeff. between HR and BP

This tells us how HR and BP are correlated.

ALA Final Exam Review

(1) Given a vector, calculate the following for that vector: average, standard deviation, RMS.
(2) Given a pair of vectors, calculate the following between those vectors: angle, correlation, distance
(3) Given a data matrix, interpret dot-products, distances, and angles between vectors in terms of the data (like the MAHE registrar problem on the sessional).
(4) Interpret the dot product in terms of the data (like, what is the meaning of dot product with the ones-vector, with the unit vector etc.)
(5) Interpret linear combination of vectors intuitively in terms of the data (like, which feature is weighted more.)
(6) Understand scalar (shadow length) and vector projection, be able to calculate for simple vectors, and interpret the results (we had an assignment on this.)
(7) Calculate matrix-vector and matrix-matrix products and explain, in terms of the data, what the results mean. Also, explain matrix-vector products like Ae_1, A1 (like in the stations example on the sessional) in terms of the data.
(8) Build a matrix-vector product that describes the scenario. For example, calculating the raw scores for an image, the ECG problem that we had as an assignment, the material-movement example that I gave as a challenge problem, and the projection of samples which we had as an assignment.
(9) Intuitively explain what the rows and columns of a data matrix convey about the data.
(10) Understand the different ways a matrix-vector product can be computed, and the intuitive meaning behind those operations and the results.
(11) Understand how a matrix-matrix product is computed and interpret the results intuitively in terms of the data.
(12) Understand the RREF of an augmented matrix, check if a system of equations is consistent or not; if consistent, check if there are free variables or only pivot variables; write the solution as a set of vectors; write down the null space of a matrix.

E.g. $a = \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$

$$\text{avg}(a) = \frac{(-1) + 2 + (-3)}{3} = -\frac{2}{3}$$

$$= \frac{1}{3}(a \cdot \mathbb{1}) = \frac{1}{3} a^T \mathbb{1}$$

$$\text{std}(a) = \text{RMS}(a_M)$$

$$a_M = a - \underbrace{\left(\frac{1}{3} a^T \mathbb{1}\right)}_{\text{avg}(a)} \underbrace{\mathbb{1}}_{\text{ones vector}}$$

$$\text{RMS}(a) = \sqrt{\frac{1}{3}\left((-1)^2 + 2^2 + (-3)^2\right)}$$

Deviation vector

$\text{std}(a) = \text{RMS}(a_M)$

$$\Rightarrow a_M = \begin{bmatrix}-1\\2\\-3\end{bmatrix} - \underbrace{\left(\frac{-2}{3}\right)}_{\text{avg}(a)} \underbrace{\begin{bmatrix}1\\1\\1\end{bmatrix}}_{\text{ones vector}} = \overbrace{\begin{bmatrix}-1+2/3 \checkmark\\2+2/3 \checkmark\\-3+2/3 \checkmark\end{bmatrix}}^{\text{Deviation vector}} \Rightarrow \text{Std}(a) = \text{RMS}(a_M)$$

$$= \sqrt{\frac{1}{3}\left(\left(-\frac{1}{3}\right)^2 + \left(\frac{8}{3}\right)^2 + \left(\frac{-7}{3}\right)^2\right)}$$

Large standard deviation for a vector $\Rightarrow$ Components of the vector will be spread around the average value

$$\boxed{a = \begin{bmatrix}1.01\\0.99\\1.02\\0.98\end{bmatrix} \quad a_M \approx \begin{bmatrix}0\\0\\0\\0\end{bmatrix}}$$

E.g. $x^{(1)} = \begin{bmatrix}1.6\\4\end{bmatrix}$, $x^{(2)} = \begin{bmatrix}1.5\\4\end{bmatrix}$, $x^{(3)} = \begin{bmatrix}1.6\\2\end{bmatrix}$

area in units of sq. ft

no. of bedrooms

House-1 is more similar to house-2 or house-3?

Distance

Squared distances

$$\|x^{(1)} - x^{(3)}\|^2 = \left\|\begin{bmatrix}1.6\\4\end{bmatrix} - \begin{bmatrix}1.6\\2\end{bmatrix}\right\|^2 = \left\|\begin{bmatrix}0\\2\end{bmatrix}\right\|^2 = \sqrt{0^2+2^2} = 2$$

$$\|x^{(1)} - x^{(2)}\|^2 = \left\|\begin{bmatrix}1.6\\4\end{bmatrix} - \begin{bmatrix}1.5\\4\end{bmatrix}\right\|^2 = \left\|\begin{bmatrix}0.1\\0\end{bmatrix}\right\|^2 = \sqrt{(0.1)^2+0^2} = 0.1$$

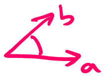
Standardize vectors before calculating distances.

Angle between vectors

$$\cos(\angle a, b) = \frac{a \cdot b}{\|a\| \, \|b\|}$$

$$a = \begin{bmatrix}-1\\2\\3\end{bmatrix}, \; b = \begin{bmatrix}1\\0\\2\end{bmatrix} \Rightarrow \cos(\angle a, b) = \frac{-1+0+6}{\sqrt{(-1)^2+2^2+3^2} \times \sqrt{1^2+0^2+2^2}}$$

$$a \cdot b = \begin{cases} 0, & \angle(a,b) = \pi/2 \\ +ve, & \angle(a,b) < \pi/2 \\ -ve, & \angle(a,b) > \pi/2 \end{cases}$$



$$0 \le \angle(a,b) \le \pi$$

$$0^\circ \le \angle(a,b) \le 180^\circ$$

between -1 and 1

Correlation between a and b $= \cos(\angle a_m, b_m) = \dfrac{a_m \cdot b_m}{\|a_m\| \, \|b_m\|}$ between -1 and 1

[10 points] [TLO 1.1, CO 1] The MAHE registrar has the complete list of courses taken by each graduating student in a program. This data is represented as a matrix X with m rows and n columns as follows:

$X =$

Student \ Course	1	2	⋯	n
1	1	1	⋯	0
2	0	1	⋯	0
⋮	⋮	⋮	⋮	⋮
m	1	0	⋯	1

$x^{(1)} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ $\quad x_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$

The entries of the data matrix are 1s and 0s representing whether a particular student has taken a particular course. For example, the red-highlighted entry 1 means that the 1st student has taken the 1st course and the blue-highlighted entry 0 means that the mth student has not taken the 2nd course. Recall that the ith student vector is represented as $x^{(i)}$ and the jth course vector is represented as x_j.

(a) The total number of students in the 2nd course is a dot product of two vectors. What are those vectors? $= x_2 \cdot \mathbb{1}$ = Sum of elements of vector x_2

$x^{(4)} \cdot \mathbb{1}$ = no. of courses taken by the 4th student

$x_k \cdot x_l$ = dot product of kth course vector and lth course vector
= no. of students who have taken both courses k and l

$x^{(k)} \cdot x^{(l)}$ = no. of courses taken in common by students k and l

$\|x^{(5)} - x^{(6)}\|$ = Difference between the 5th and the 6th student

[15 points] [TLO 1.2, CO 1] In a multiple channel marketing campaign, potential customers are divided into m market segments, which are groups of customers with similar demographics, e.g., male, female, married, college educated women aged 25–29 etc. A company markets its products by purchasing advertising in a set of n channels, i.e., specific TV or radio shows, magazines, web sites, blogs, direct mail, and so on. The ability of each channel to deliver impressions or views by potential customers is characterized by the data matrix X, which has m rows and n columns representing m customer segments and n advertising channels. (The entry in the ith row and jth column of the data matrix is the number of views of customers in segment i for each Rupee spent on channel j.) The n-vector c will denote the company's purchases of advertising, in Rupees, in the n channels. Finally, we introduce the m-vector a, where a_i gives the profit in Rupees per impression in market segment i.

$n = 4$

$X =$

	TV	Radio	Magazine	Internet
Male	0.1			
Female	0.3			
Married	1.5	0	0.7	5.
College	⋮			
25–29	⋮			

$m = 5$

$\to x_1$

0.1 = no. of male views per Rupee spent on TV advertisement

$c = \begin{bmatrix} \boxed{100000} \\ 50000 \\ 200000 \\ 500000 \end{bmatrix}$ → Total budget in Rupees for TV advertisement

(a) What do the values in the first column of the data matrix indicate? $x_1 = ?$

(b) What do the values in the third row of the data matrix indicate? $x^{(3)} = ?$

(c) Express the total amount of money the company spends on advertising using a dot product. $= \text{sum}(c) = c \cdot \mathbb{1}$, $\mathbb{1}$ = ones vector $= \begin{bmatrix} \vdots \end{bmatrix}$

(d) Consider the vector $c_1x_1 + c_2x_2 + \cdots + c_nx_n$. This is called a linear combination of the vectors $x_1, x_2, \ldots, x_n$. Interpret this vector in plain English; do this by starting with interpreting the under-highlighted terms below in plain English:

$$\underbrace{c_1}_{?} \times \underbrace{\begin{bmatrix} \underbrace{(x_1)_1}_{?} \\ \underbrace{(x_1)_2}_{?} \\ \vdots \\ \underbrace{(x_1)_m}_{?} \end{bmatrix}}_{?} + \cdots + \underbrace{c_n}_{?} \times \underbrace{\begin{bmatrix} \underbrace{(x_n)_1}_{?} \\ \underbrace{(x_n)_2}_{?} \\ \vdots \\ \underbrace{(x_n)_m}_{?} \end{bmatrix}}_{?} = \begin{bmatrix} 10000 \\ 30000 \\ 15000 \\ \vdots \end{bmatrix}$$

$c_1 x_1$: c_1 → Total TV advertisment budget (Rupees); x_1 → $\dfrac{\text{views on TV}}{\text{Rupee}}$

$c_1 x_1$ = Rupees × $\dfrac{\text{views}}{\cancel{\text{Rupee}}}$

→ = Total no. of views on TV across all market segments

(e) Consider the dot product $a \cdot (c_1x_1 + c_2x_2 + \cdots + c_nx_n)$. In two words, state what the resulting number indicates.

(e) Consider the dot product $a \cdot (c_1x_1 + c_2x_2 + \cdots + c_nx_n)$. In two words, state what the resulting number indicates.

→ = Total no. of views on TV across all market segments

$$\begin{bmatrix} \vdots \\ 200000 \\ 500000 \end{bmatrix} \quad \text{TV advertisement}$$

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = \text{Total no. of views across all market segments}$$

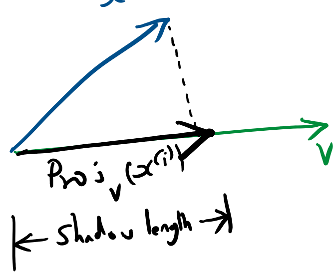
$$a = \begin{bmatrix} 0.1 \, \text{Rupees/view} \\ 0.2 \, R/V \\ 0.05 \, R/V \\ 0.2 \, R/V \end{bmatrix} = \text{earnings-per-view vector}$$

$$\underbrace{(c_1x_1 + c_2x_2 + \cdots + c_nx_n)}_{\text{Total views vector}} \cdot \underbrace{a}_{\text{Profit-per-view vector}} = \text{Total profit (or) earnings}$$

Vector projection

$$\text{proj}_v\left(x^{(i)}\right) = \left(\frac{x^{(i)} \cdot v}{v \cdot v}\right) v = \left(\frac{x^{(i)} \cdot v}{\|v\|^2}\right) v = \boxed{\underbrace{\left(\frac{x^{(i)} \cdot v}{\|v\|}\right)}_{\text{shadow length}} \underbrace{\frac{v}{\|v\|}}_{\text{direction}}}$$

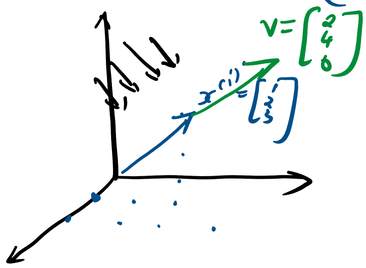
Scalar projection



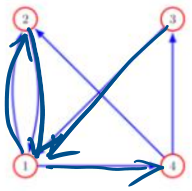
$$x^{(i)} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad v = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

$$\text{Scalar projection (shadow length)} = \frac{x^{(i)} \cdot v}{\|v\|} = \frac{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}}{\sqrt{2^2 + 4^2 + 6^2}} = \frac{2 + 8 + 18}{\sqrt{56}} = \frac{28}{\sqrt{56}}$$

$$\text{Vector projection} = \left(\frac{x^{(i)} \cdot v}{\|v\|}\right) \frac{v}{\|v\|} = \frac{28}{\sqrt{56}} * \frac{\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}}{\sqrt{56}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = x^{(i)}$$



[10 points] [CO 2, BT 3] Suppose we have four stations that are connected by train services as shown in the following graph:



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

The *adjacency matrix* associated with this graph has entries defined as

$$A_{ij} = \begin{cases} 1 & \text{if direct train service exists from station } j \text{ to station } i, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Write the adjacency matrix A associated with this graph clearly showing its entries.

$$Ae_1 = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$
= Stations which have a direct train from station-1

Sum of elements in each row $= A\mathbb{1} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 1 \end{bmatrix}$

Two stations from which we have a direct train to station-1

$(A\mathbb{1})_2$ = no. of stations from which we have a direct train to station-2

Sum of elements in each column $= A^T\mathbb{1} = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 2 \end{bmatrix}$

1 station for which we have a direct train from station-3

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

$$Ax = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} x_5 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$
= down-shifted version of the vector x

$$\begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ v_1 \\ x_4 \\ v_2 \\ v_3 \end{bmatrix}$$

$$\frac{1}{\|v\|} \begin{bmatrix} (x^{(1)})^T v \\ (x^{(2)})^T v \\ (x^{(3)})^T v \\ (x^{(4)})^T v \end{bmatrix}$$
= projection of all samples onto the direction of v

$= \cancel{X^T} v$ (or) Xv

$\|v\| = 1$ because v is a direction vector

$$= \begin{bmatrix} x^{(1)T} \\ x^{(2)T} \\ x^{(3)T} \\ x^{(4)T} \end{bmatrix} v = Xv$$

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

$$\underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}}_{A} * \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}}_{A} = \begin{bmatrix} 0 & 0 & \\ 0 & 0 & \\ 1 & 0 & \cdots \\ 0 & 1 & \\ 0 & 0 & \end{bmatrix}$$

$$\underset{m \times n}{A}\,\underset{n \times p}{B} = \begin{bmatrix} \underbrace{Ab_1}_{\text{matrix-vector product}} & Ab_2 & \cdots & Ab_p \end{bmatrix}$$