Solving linear equations using the inverse matrix

a → apple b → banana c → carrot

Monday $1a + 1b + 1c = €15$

Tuesday $3a + 2b + 1c = €28$

Wednesday $2a + 1b + 1c = €23$

$$\begin{pmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 15 \\ 28 \\ 23 \end{bmatrix}$$

$\frac{7}{2} \cdot 3$

$\frac{7 \cdot 6}{2} = \frac{1}{2}$

$\frac{7}{4}$ $\frac{27}{4}$

$3/4$ $\frac{1}{4}$

2. Given another system $Br = t$

$$\begin{array}{c} ① \\ ② \\ ③ \end{array} \begin{pmatrix} 4 & 6 & 2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 2 \end{bmatrix}$$

$$①/4 \quad \begin{bmatrix} 1 & 3/2 & 1/2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 9/4 \\ 7 \\ 2 \end{bmatrix}$$

Next results in

$$\begin{bmatrix} 1 & 3/2 & 1/2 \\ 0 & 1 & 1 \\ 2 & 8 & 13 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 9/4 \\ -1/2 \\ 2 \end{bmatrix}$$

Find out what operation resulted in the new ②

0 1 -2 4

12 18 6 27

-9 -14 -5 -20

18 28 10 40

4 $-\frac{5}{4}$

$\frac{16-9}{4} = \frac{7}{4}$

-9 -14 -5 -20

$\frac{3}{4}$ 1 $\frac{1}{4}$ $\frac{7}{4}$

0 1 -2 4

8 12 4 18

-5 -8 -3 -11

① − 3 ⇒ 0 1 −2 4

② − 20 ⇒ −5 −8 −3 −11

② − 30 ⇒ −9 −14 −5 −20

② / 4 − ①

4 6 2 9

3 4 1 7

2 8 13 2

1 3/2 1/2 9/4

0 1 −2 4

2 8 13 2

4 6 2 9

−5 −8 −3 −11

8 12 4 18

3 4 1 7

12 18 6 27

3 4 1 7

−9 −14 −5 −20

4 6 2 9

$\frac{3}{4}$ 1 $\frac{1}{4}$ $\frac{7}{4}$

$\frac{7}{4} - 9$

$\frac{7 - 36}{4} =$

0 $\frac{1}{2}$ −1 2

$-\frac{1}{4}$ −5 $-\frac{3}{4}$ $-\frac{29}{4}$

$4 \cdot \frac{9}{4}$

$-\frac{7}{4}$

8 12 4 18

12 18 6 27

−9 −14 −5 −20

18 28 10 40

3/4 1 1/4 7/4

1/4 − 2

3/4 − 4

$-\frac{13}{4}$ −5 $-\frac{7}{4}$ $-\frac{29}{4}$

$\frac{7}{4} - 9$

$\frac{7 - 36}{4}$

$$\begin{array}{ccc} 1 & 3/2 & 1/2 \\ 3 & & \end{array}$$

$$\begin{pmatrix} 4 & 6 & 2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9 \\ 7 \\ 2 \end{pmatrix}$$

① ÷ 4

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \frac{9}{4} \\ 7 \\ 2 \end{pmatrix}$$

$$\begin{array}{cccc} 8 & 12 & 4 & 18 \\ -5 & -8 & -3 & -11 \end{array}$$

② '' = ② ' − 3① '

$$\begin{array}{cccc} 12 & 18 & 6 & 27 \\ -9 & -14 & -5 & -20 \end{array}$$

② '' = ② ' / 4 − ① ' $\frac{1}{4}$ − 2 =

$$\begin{array}{cccc} 3/4 & 1 & 1/4 & 7/4 \\ -1/4 & -5 & \frac{-7}{4} & \frac{-29}{4} \\ 0 & 1 & -2 & 4 \end{array}$$

② ' − 3① ' | × −2

$$\begin{array}{cccc} 12 & 18 & 6 & 27 \\ -9 & -14 & -5 & -20 \\ 18 & 28 & 10 & 40 \end{array}$$

② '' = ② ' − 2① '

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ -5 & -8 & -3 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \frac{9}{4} \\ -11 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ -1/4 & -5 & -7/4 \\ 2 & 8 & 13 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ 0 & 1 & -2 \\ 2 & 8 & 13 \end{pmatrix}$$

$$\begin{array}{cccc} 0 & 3 & 2 & 9 \end{array}$$

$$\begin{pmatrix} 1 & \frac{3}{2} & \frac{1}{2} \\ -5 & -8 & -3 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \frac{9}{4} \\ -11 \\ 2 \end{pmatrix}$$

$$-11 + \frac{9}{4}$$

$$\frac{-44+9}{4}$$

$$= \frac{-35}{4}$$

$$\begin{array}{ccc} 1 & 3/2 & 1/2 \\ -5 & -8 & -3 \\ 2 & 8 & 13 \end{array}$$

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ -3 & 0 & 10 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{bmatrix} 9/4 \\ -9 \\ 2 \end{bmatrix}$$

$$\begin{pmatrix} 4 & 6 & 2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9 \\ 7 \\ 2 \end{pmatrix}$$

$$\frac{3}{2} + \frac{4}{3} = \frac{9+8}{6} = \frac{15}{1}$$

$$\begin{pmatrix} 1 & \frac{3}{2} & 1/2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9/4 \\ 7 \\ 2 \end{pmatrix}$$

$3/2 \quad 5/3 \quad 9/6 \quad 8/6 \quad 1/6$

8 12 14 18

-5 -8 -3 -11

-7 -15 -16 -13

$2'' = 2' - 2\,1''$

$$\begin{pmatrix} 1 & \frac{3}{2} & \frac{1}{2} \\ -5 & -8 & -3 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9/4 \\ -11 \\ 2 \end{pmatrix}$$

$2''' = 2'' + 5\,3$

$$\begin{pmatrix} 1 & \frac{3}{2} & \frac{1}{2} \\ 0 & 5 & 15 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9/4 \\ -11 \\ 7 \end{pmatrix}$$

0 6 11

0 -1 -4 -16

$2''' = 2'' + 1''$

$$\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & 9/4 \\ -1 & -2 & -1 & -2 \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$-2 + \frac{3}{2}$

$-\frac{1}{2}$

$-2 + \frac{9}{4}$

$2 - 9/4$

$2'''' = 2''' +$

$$\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & 9/4 \\ 0 & -\frac{1}{2} & \frac{3}{2} & -1/4 \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & 9/4 \\ 1 & 2 & 1 & 2 \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & 9/4 \\ 0 & -\frac{1}{2} & \frac{3}{2} & -1/4 \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$1'' = 1' - 2$

$$\begin{pmatrix} 4 & 6 & 2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix}$$

$2'' = 1'' \times 3 - 2$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 3 & 4 & 1 & 7 \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & 2 & -1 \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & -1/2 \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & -1/2 \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$2'' = 2' - 2$

8 12 14 18

10 18 6 22

10 40 65 10

0 24 59 12

$2'' = 2'' - 2$

$$\begin{pmatrix} 4 & 6 & 2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9 \\ 7 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9/4 \\ 7 \\ 2 \end{pmatrix}$$

12 18 6 27

-9 -14 -5 -20

18 28 10 40

12 18 6 27

-9 -14 -5 -20

② - 3① × (-2)

-9

$$\cancel{\begin{pmatrix} 1 & 3/2 & 1/2 \\ -9 & -14 & -5 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9/4 \\ -20 \\ 2 \end{pmatrix}}$$

$$\begin{pmatrix} 4 & 6 & 2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix}$$

$$\left(\begin{array}{ccc|c} 1 & 3/2 & 1/2 & 9/4 \\ 18 & 28 & 10 & 40 \\ 2 & 8 & 13 & 2 \end{array}\right) \Rightarrow$$

②2 ①×3 - ②

$$\begin{pmatrix} 4 & 6 & 2 \\ 9 & 14 & 5 \\ 2 & 8 & 13 \end{pmatrix}$$

$\frac{1}{4} - 2$

$\frac{7}{4} - 9$

$\frac{9}{4} + 2$

$3/4 \quad 1 \quad 1/4 \quad 7/4$

$5/4 - 4$

$-13/4$

$1/4 - 2$

$\frac{7}{4} - 9$

- - - -

(2')/4 - ①

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ -13/4 & -5 & -7/4 \\ 2 & 8 & 13 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9/4 \\ -27/4 \\ 2 \end{pmatrix}$$

(2'')+1

$\frac{5}{4} + 2 \quad \frac{13}{4}$

$\frac{5}{4} + 8 \quad \frac{37}{4}$

$\frac{5}{4} + 13 \quad \frac{57}{4}$

$\frac{5}{4} + 2 \quad \frac{13}{4}$

$2 + 1/2$

$5/2$

$$\begin{pmatrix} 4 & 6 & 2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9 \\ 7 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3/2 & 1/2 \\ 3 & 4 & 1 \\ 2 & 8 & 13 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9 \\ 7 \\ 2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 3/2 & 1/2 \\ 0 & 1 & -2 \\ 2 & 8 & 13 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9/4 \\ -1/2 \\ 2 \end{pmatrix}$$

② - 3

$$\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & \frac{9}{4} \\ 0 & 1 & 1 & -\frac{1}{2} \\ 2 & 8 & 13 & 2 \end{array}\right)$$

$$\cancel{\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & \frac{9}{4} \\ 0 & 1 & 1 & -\frac{1}{2} \\ -1 & 4 & 12 & -5 \end{array}\right)}$$

$$\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & \frac{9}{4} \\ 0 & 1 & 1 & -\frac{1}{2} \\ 0 & 5 & 12 & -\frac{5}{2} \end{array}\right)$$

③ = ③ − 2 × ①

③ = ③ − 2 × ②

$2 - \frac{9}{2}$

$-\frac{5}{2}$

$$\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & \frac{9}{4} \\ 0 & 1 & 1 & -\frac{1}{2} \\ 0 & 0 & 7 & 0 \end{array}\right)$$

$c = 0$

$b + c = -\frac{1}{2} \quad b = -\frac{1}{2}$

$a + \frac{3}{2}b + \frac{1}{2}c = \frac{9}{4}$

$a + \frac{3}{2}\left(-\frac{1}{2}\right) = \frac{9}{4}$

$a + -\frac{3}{4} = \frac{9}{4} \Rightarrow a = \frac{9}{4} + \frac{3}{4}$

$= \frac{12}{4} = 3$

$$\left(\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{1}{2} & \frac{9}{4} \\ 0 & 1 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 3 & 2 & 1 & 28 \\ 2 & 1 & 1 & 23 \end{array}\right) \quad ② = ① \times 3 - ② \quad \left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 1 & 2 & 17 \\ 2 & 1 & 1 & 23 \end{array}\right)$$

$$\cancel{② = ② - ①} \quad \cancel{\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ -1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 23 \end{array}\right)}$$

$$③ = ① \times 2 - ③ \quad \left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 1 & 2 & 17 \\ 0 & 1 & 1 & 7 \end{array}\right) \quad ③ = ② - ③$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 1 & 2 & 17 \\ 0 & 0 & 1 & 10 \end{array}\right) \quad ② = ② - ③ \times 2 \quad \cancel{\begin{array}{cccc} 1 & 1 & 1 & 15 \\ 0 & 1 & 2 & 17 \end{array}}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 10 \end{array}\right) \quad ① = ① - ②$$

$$\left(\begin{array}{cccc} 1 & 0 & 1 & 18 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 10 \end{array}\right)$$

$$① = ① - ③$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 10 \end{array}\right)$$

$a + b + c = 15$

$b = -3$

$c = 10$

$a + 7 = 15$

$a = 8$

$Bx = t$

So every week you buy the same amount of apples, bananas and oranges on Monday, Tuesday and Wednesday and every week you get a new list of totals then you should solve the system in general

Same ~~of~~ 3 goods but prices vary

$$\begin{bmatrix} & & \\ & \text{Goods} & \\ & & \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{pmatrix} \\ \\ \end{pmatrix}$$

$$G p = V$$

$$p = G^{-1} V$$

Find G^{-1}

$$G G^{-1} = 1$$

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

② = ① × 3 - 2

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

③ = ① × 2 - ③

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

~~① = ① - ②~~

~~② = ② − ③ × 2~~

③ = ② − ③

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

2 -2 2

② = ② − ③ × 2

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

① = ① − ②

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

① = ① − ③

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} G^{-1} = \begin{pmatrix} -1 & 4 & 1 \\ 1 & -3 & -2 \\ 1 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 15 \\ 18 \\ 25 \end{pmatrix}$$

$\frac{15}{25}$

$(2) = (1) \times 2 - 2$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 1 & 2 & 17 \\ 3 & 1 & 2 & 25 \end{array}\right)$$

$(3) = (1) \times 3 - 3$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 1 & 2 & 17 \\ 0 & 1 & 0 & 7 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 7 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 13 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 7 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 7 \end{array}\right)$$

$(2) \to (3)$

$(3) \to (2)$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 1 & 0 & 7 \\ 0 & 1 & 2 & 17 \end{array}\right)$$

$(3) = (3) - (1)$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

$(1) = (1) - (2)$

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 8 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

$(1) = (1) - (3)$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

$6 + 7 + 2 = 15$

$6 + 7 = 13$

$6 + 7 = 13$

$6 + 7$

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

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

③ = ① × 2 − ③

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

② → ③
③ → ②

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

③ = ③ − ①

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & -2 \\ 2 & -1 & 0 \end{pmatrix}$$

① = ① − ②

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} G^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 2 & 0 & -2 \\ 2 & -1 & 0 \end{pmatrix}$$

① = ① − ③

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

Convert to echelon form

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 3 & 2 & 1 & 28 \\ 2 & 1 & 2 & 23 \end{array}\right)$$

$② = ② - 3 \times ①$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & -1 & -2 & -17 \\ 2 & 1 & 2 & 23 \end{array}\right)$$

28

$\cancel{③ = ③ + ②}$

$$\cancel{\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & -1 & -2 & -17 \\ 2 & 0 & 0 & 6 \end{array}\right)}$$

$③ = ③ - 2 \times ①$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & -1 & -2 & -17 \\ 0 & -1 & 0 & -7 \end{array}\right)$$

$③ = ③ - ②$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & -1 & -2 & -17 \\ 0 & 0 & 2 & 10 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 15 \\ 0 & -1 & -2 & -17 \\ 0 & 0 & 1 & 5 \end{array}\right)$$

$c = 5$

$b + 2c = 17$

$b + 10 = 17$

$b = 7$

$a + 7 + 5 = 15$

$a = 3$

$$\begin{pmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$(2') = (2) - 3 \times (1)$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 2 & 1 & 2 \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$(3) = (3) - 2 \times (1)$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & -1 & 0 \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$$

$(3) = (3) - (2)$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 2 \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$$

$(3) = (3)/2$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$(2) = -(2)$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$(1) = (1) - (3)$

$(2) = (2) - (3) \times 2$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad G^{-1} = \begin{pmatrix} -2 & 1 & 0 \\ 4 & 0 & -1 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$(1) = (1) + 3$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad G^{-1} = \begin{pmatrix} -\frac{5}{2} & \frac{1}{2} & \frac{1}{2} \\ 4 & 0 & -1 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

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

$② = ② - ① \times 3$

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

$③ = ③ - ① \times 2$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & -1 & 0 \end{pmatrix} G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

$③ = ③ - ②$

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

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix} G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix} G^{-1} = \begin{pmatrix} -2 & 1 & 0 \\ -3 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$\cdot 2 + \frac{1}{2}$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix} G^{-1} = \begin{pmatrix} -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\ -3 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} G^{-1} = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\ -2 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Determinants and Inverses:

Matrix multiplication:

1) $\begin{bmatrix} 2 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix} = 2 + 12 + 10 + 6 = 30$

$1 \times 4 \quad 4 \times 1$

2) $\begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 5 & 6 \end{bmatrix} = \cancel{2 + 12 + 10 + 6 = 30}$

$4 \times 1 \quad 1 \times 4 \qquad 4 \times 4$

$$\begin{matrix} 2 & 4 & 5 & 6 \\ 6 & 12 & 15 & 18 \\ 4 & 8 & 10 & 12 \\ 2 & 4 & 5 & 6 \end{matrix}$$

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

$\begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$

3) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{matrix} 4 & 3 & 5 \\ 5 & 4 & 1 \end{matrix}$

$2 \times 3 \quad 3 \times 3 = 2 \times 3$

4) $\begin{bmatrix} 2 & -1 \\ 0 & 3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 4 & -1 \\ -2 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 8 & -4 \\ -6 & 0 & 0 & 6 \\ 0 & 1 & 4 & -1 \end{bmatrix}$

$3 \times 2 \quad 2 \times 4 = 3 \times 4$

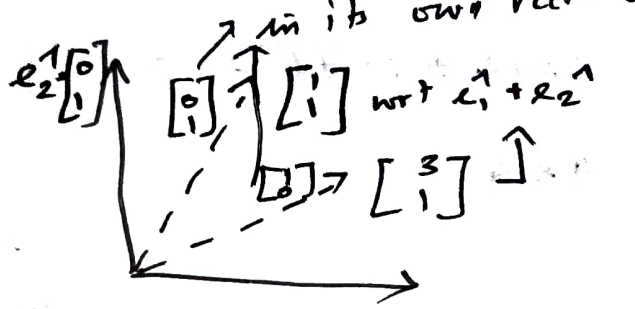
5) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{matrix}$

$2 \times 2 \quad 2 \times 3 = 2 \times 3$

Matrices changing basis:

Columns of a transforming matrix are the axes of the new basis vectors mapping in my coordinate system.

Diagram is not accurate



Bear is just an assumption. We assumed Panda's bear in its own vector space

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

Dotted vector is bear's vector

Bears basis vectors are $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in my space

I can write down bear's transformation matrix as

$\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \to$ columns of transformation matrix are the axes of new basis vector

So $\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$ acts as the basis vectors of the transformation matrix for bear vector

How to transform from our vector space to bear vector space?

$$\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$$

$\hookrightarrow$ Inverse

$$\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Using transformation matrix $\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$ when $\frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, a basis vector in basis space, when composed are get $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

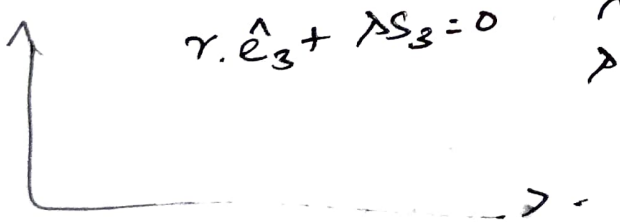
Doing transformation in a changed basis:

Shadows are an example of a transformation that reduces the number of dimensions. For example, 3D objects in the world cast shadows on surfaces that are 2D

Assignment Roughwork follows

Rough work

$r' = r + \lambda \hat{s}$ $\quad r'_3 = 0 \quad r' \cdot \hat{e}_3 = 0$



$r \cdot \hat{e}_3 + \lambda s_3 = 0$ $\quad \lambda s_3 = -r \cdot \hat{e}_3$ $\quad \lambda = -\frac{r \cdot \hat{e}_3}{s_3}$

$\lambda \hat{s} \cdot \hat{e}_3$ $\quad = -\frac{r \cdot \hat{e}_3}{\hat{s} \cdot \hat{e}_3}$

$\lambda \hat{s} = -r$ $\quad = \frac{-r}{\hat{s}}$

$r' = r + \frac{-r}{\hat{s}} \cdot \hat{s}$ $\quad \lambda s_3 = -r \cdot \hat{e}_3$

$\lambda = -\frac{r \cdot \hat{e}_3}{s_3}$

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

$r' = r + \left(\frac{-r \cdot \hat{e}_3}{s_3}\right) \hat{s}$

$r' = r - \hat{s}\left(\frac{r \cdot \hat{e}_3}{s_3}\right) = r\left(1 - \frac{\hat{s} \cdot \hat{e}_3}{s_3}\right)$

$$\begin{bmatrix} - & \sim & \sim \end{bmatrix}$$

$$r' = r - \hat{s}(r \cdot \hat{e_3})/s_3$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} -$$

$$r = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$r_1 = \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \quad r_2 \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} \quad r_3 = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$$

$$\hat{s} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$s_{ij}\, r_{ij} \cdot \hat{e_3}$$

$$s_{ij} \cdot [\hat{e_3}]$$

$$\hat{s_1} = \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \quad \hat{s_2} = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}$$

$$\hat{s_3} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

```
for int i=0 to n
    for int k=0 to n
        for int j=0 to n.
```

$$a_{11} \cdot b_{11} + a_{12} b_{21} + \cdots + a_{1n} b_{n1} \qquad a_{11} \cdot b_{12} + a_{12} b_{22} + \cdots + a_{1n} b_{n2}$$

$$\sum a_{ij} b_{jk} = a_{ij} b_{jk}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} -$$

$$\begin{array}{ccc} r_1 & r_2 & r_3 \end{array} \qquad \begin{array}{ccc} s_1 & s_2 & s_3 \end{array}$$

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} - \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix}$$

$$\begin{array}{cc} r & s \end{array}$$

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} - \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$$

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} - \begin{pmatrix} s_{11}l_{11} + s_{12}l_{21} + s_{13}l_{31} & s_{11}l_{12} + s_{12}l_{22} + s_{13}l_{32} & s_{11}l_{13} + s_{12}l_{23} + s_{13}l_{33} \end{pmatrix}$$

r_{11}

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \begin{pmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \cancel{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} s_{11} & \end{pmatrix}$$

Matrix A

$$\left[\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} - \hat{s}\cdot\hat{e_3} \Big/ s_3\right]$$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} - \underbrace{\begin{pmatrix}s_{11}&s_{12}&s_{13}\\s_{21}&s_{22}&s_{23}\\s_{31}&s_{32}&s_{33}\end{pmatrix}}_{3\times3}\cdot\underbrace{\begin{bmatrix}0\\0\\1\end{bmatrix}}_{3\times1} \Big/ s_3$$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} - \begin{pmatrix}s_{13}\\s_{23}\\s_{33}\end{pmatrix} \Big/ s_3$$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} - \begin{pmatrix}s_{13}/s_3\\s_{23}/s_3\\s_{33}/s_3\end{pmatrix} =$$

Rough Work

$r' = r + \lambda\hat{s}$

$r' = Ar$

↳ Transformation matrix

$r'\cdot\hat{e_3} = 0$

$\Rightarrow (r + \lambda\hat{s})\cdot\hat{e_3} = 0$

$\Rightarrow r\cdot\hat{e_3} + \lambda\,\hat{s}\cdot\hat{e_3}$

$\Rightarrow r\cdot\hat{e_3} + \lambda s_3$

$r\cdot\hat{e_3} + \lambda s_3 = 0$

$\lambda s_3 = -r\cdot\hat{e_3}$

$\lambda = -r\cdot\hat{e_3}/s_3$

$r + (-r\cdot\hat{e_3}/s_3)\hat{s}$

$r - r\cdot\hat{e_3}/s_3\,\hat{s}$

$r - \hat{s}\,r\cdot\hat{e_3}/s_3$

$r - \hat{s}\,(r\cdot\hat{e_3})/s_3$

$r' = \hat{s}\,(\hat{e_3}\cdot r)/s_3$

$(I - \hat{s}\,\hat{e_3}/s_3)\,r'$

A found

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} - \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \Big/ s_3$$

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ s_3 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \Big/ s_3$$

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ s_3 \end{pmatrix} \Big/ s_3 \right]$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{pmatrix} s_1/s_3 \\ s_2/s_3 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{bmatrix} s_1/s_3 & 0 & -\frac{s_1}{s_3} \\ 0 & s_2/s_3 & -s_2/s_3 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

$$s_1/s_3\, r_1 - \frac{s_1}{s_3} r_3$$

$$s/s_3\, r_2 - s_2/s_3\, r_3$$

Rough Work

$$\begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} \overset{\hat{e}_3}{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}$$

$$= \begin{bmatrix} s_{13} \\ s_{23} \\ s_{33} \end{bmatrix} = \underline{s_3}$$

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} r_{13} \\ r_{23} \\ r_{33} \end{bmatrix} =$$

$r' = r + \lambda \hat{s}$

$r' = Ar$

↳ Transformation matrix

$r' \cdot \hat{e}_3 = 0 \Rightarrow (r + \lambda \hat{s}) \cdot \hat{e}_3 = 0$

$r \cdot \hat{e}_3 + \lambda \hat{s} \hat{e}_3 = 0$

$\lambda \hat{s} \hat{e}_3 = -r \cdot \hat{e}_3$

$\lambda = \dfrac{-r \cdot \hat{e}_3}{\hat{s} \hat{e}_3}$

$\lambda = \dfrac{-r \cdot \hat{e}_3}{s_3}$

$$\Rightarrow r' = r - \frac{r \cdot \hat{e}_3 \hat{s}}{s_3}$$

$$= \frac{s_3 r - r \cdot \hat{e}_3 \hat{s}}{s_3}$$

$$= \frac{r(I_{ij} - \hat{s} \cdot \hat{e}_3)}{s_3}$$

$$= r(I_{ij} - \hat{s} \cdot \hat{e}_3) \frac{1}{s_3}$$

$$= r\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix}\right.$$

l_{3j}

$$[0\ 0\ 1] \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

$r_{31}\quad r_{23}\quad r_{33}$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \left(\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \cancel{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} [0\ 0\ 1]\right) \Big/ s_3$$

$3 \times 1 \qquad 1 \times 3$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & s_1 \\ 0 & 0 & s_2 \\ 0 & 0 & s_3 \end{bmatrix} \Big/ s_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & s_1/s_3 \\ 0 & 0 & s_2/s_3 \\ 0 & 0 & s_3/s_3 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & -s_1/s_3 \\ 0 & 1 & -s_2/s_3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\hat{s} = \begin{bmatrix} 4/13 \\ -3/13 \\ -12/13 \end{bmatrix} \begin{matrix} s_1 \\ s_2 \\ s_3 \end{matrix} \qquad A = \begin{bmatrix} 1 & 0 & -s_1/s_3 \\ 0 & 1 & -s_2/s_3 \end{bmatrix}$$

$$\Rightarrow A = \begin{bmatrix} 1 & 0 & 1/3 \\ 0 & 1 & -1/4 \end{bmatrix} \qquad x = \begin{bmatrix} 6 \\ 2 \\ 3 \end{bmatrix}$$

$$\underset{A_{ij}}{A_r} = 7 \qquad \underset{R_{jk}}{2\times 3} \qquad \underset{}{3\times 1}$$

$$2 - \frac{3}{4}$$

$$\frac{5}{4}$$

$$R'_{jk} = \begin{bmatrix} 1 & 0 & 1/3 \\ 0 & 1 & -1/4 \end{bmatrix} \begin{bmatrix} 5 & -1 & -3 & 7 \\ 4 & -4 & 1 & -2 \\ 9 & 3 & 0 & 12 \end{bmatrix}$$

$$4 - \frac{9}{4}$$

$$-4 - \frac{3}{4}$$

$$-19/4$$

$$= \begin{bmatrix} 8 & 0 & -3 & 11 \\ 7/4 & -19/4 & 1 & -5 \end{bmatrix}$$

Orthogonal matrices:

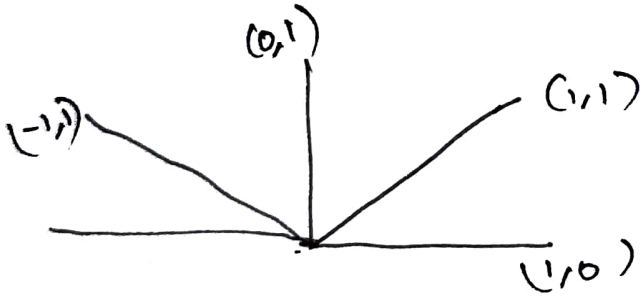
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 11 \\ 11 & 25 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 10 & 14 \\ 14 & 24 \end{bmatrix}$$

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

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



$$A \qquad \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

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

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

An orthogonal matrix is a square matrix whose columns and rows are orthogonal unit vectors.

i.e. $A^TA = AA^T = I$ where I is the identity matrix

$\Rightarrow$ A matrix A is orthogonal if its transpose is its inverse

In the case when A is composed of vectors normal to each other and have unit length then A^TA is the identity matrix. So effectively when the composing vectors are orthonormal and have unit length $A^TA = A^{-1}A$ where $A^T = A^{-1}$

In this case A is called an orthogonal matrix.

Orthogonal matrix when its composing vectors are of unit length then ~~the~~ the matrix must scale space by a factor of 1. So it doesn't make space bigger or smaller. So the determinant of an orthogonal matrix must be either ± 1.

$$|A| = \pm 1$$

In this case $A^TA = I$ $\rightarrow AA^T = I$. And that means by the logic that the rows of the orthogonal matrix are also orthonormal to each other.

$$A^TA = AA^T$$

$$\underbrace{\overset{A^T}{\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}} \overset{A}{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}}_{\text{Columns are orthonormal}} = \underbrace{\overset{A}{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}} \overset{A^T}{\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}}_{\text{Rows are orthonormal}} = 2\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

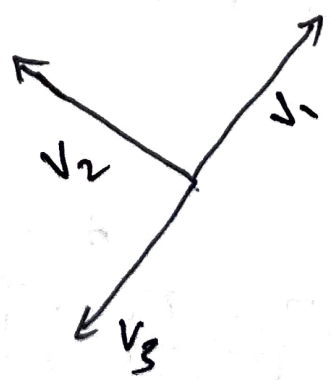
Gram-Schmidt process:

The Gram-Schmidt process is a method for constructing an orthonormal basis of a space that a set of given vectors span. It can also used to determine the dimension of that space, which may be different than the dimensions of the vectors themselves or the number of vectors provided to span the space.

Let $B = \{v_1, v_2 \dots v_n\}$

linearly independent vectors

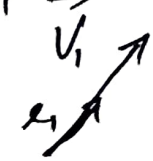
We need to create an orthonormal basis of a space with the given vectors. Assuming all the given vectors in the set B are linearly independent



Orthonormal basis of a span are a set of vectors that are unit vectors and are orthogonal to each other

To get the unit vector of V_1

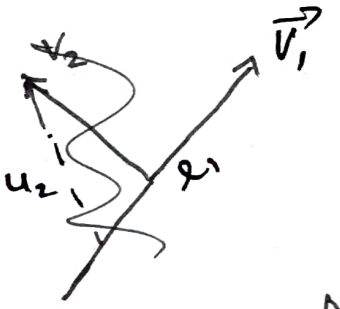
$$e_1 = \frac{V_1}{\|V_1\|}$$ e_1 is Normalised version of V_1 $\rightarrow$



To get the unit vector of V_2, we need to eliminate V_1's component

i.e

$$V_2 = (V_2 . e_1)\frac{e_1}{\underbrace{|e_1|}_{=1}} + u_2$$

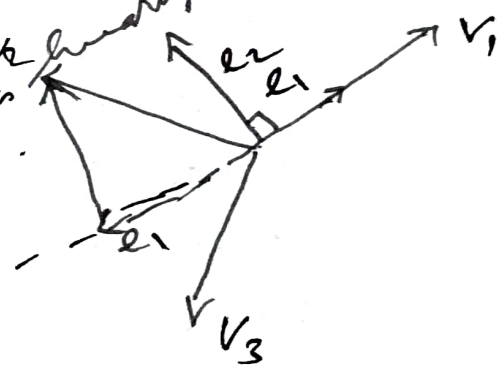


$\Rightarrow u_2 = V_2 - (V_2 . e_1) e_1$

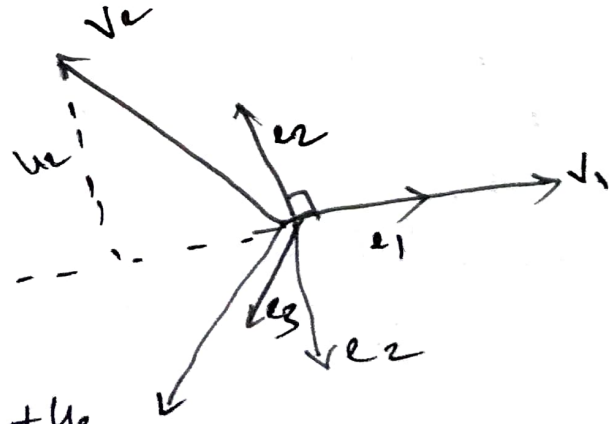
If there is anything remaining in V_2 linearly independent from e_1 then e_2 is

$e_2 = \frac{u_2}{\|u_2\|}$ Normalise u_2 to get e_2

$\hookrightarrow$ Normal to e_1



~~$V_3 = (V_3 . e_2) e_2 - V_2$~~



V_3 is not in the plane formed by V_1 + V_2

or e_1 + e_2

$$V_3 = (V_3 . e_1) e_1 + (V_3 . e_2) e_2 + u_3$$

$$u_3 = V_3 - (V_3 . e_1) e_1 + (V_3 . e_2) e_2$$

$$\& \quad e_3 = \frac{u_3}{\|u_3\|}$$

We can continue doing to get enough orthonormal span to other vectors

When we have orthonormal basis, a orthogonal matrix, then it becomes really simple

$$A^T A = A^{-1} A$$

Notes on numpy.linalg.norm

This is used to normalise a vector

see previous page $e_1 = \frac{V_1}{\|V_1\|}$

This can be written as numpy.linalg.norm(V_1)

numpy matrix

A[0,0] A[0,1] A[0,2] A[0,3]
A[1,0] A[1,1] A[1,2] A[1,3]
A[2,0] A[2,1] A[2,2] A[2,3]
A[3,0] A[3,1] A[3,2] A[3,3]

You can access a value in an element by

A[n, m]

you can also access a whole row at a time using

A[n]

You can select a whole column by

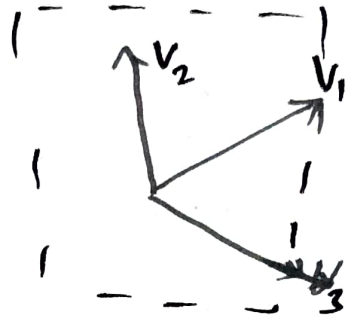
A[:, m] which will select the mth column

starting at 0.

Dot product between two vectors u, v can be obtained by u@v

Reflecting in a plane:

$$v_1 = \begin{pmatrix}1\\1\\1\end{pmatrix} \quad v_2 = \begin{pmatrix}2\\0\\1\end{pmatrix} \quad v_3 \begin{pmatrix}3\\1\\-1\end{pmatrix}$$



V_1 + V_2 lie in the same plane while V_3 is not

Use Gram-Schmidt process to find some orthonormal vectors describing the plane and its normal V_3

$$e_1 = \frac{V_1}{|V_1|} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}$$

dot product

$$u_2 = V_2 - (V_2 \cdot e_1) \cdot e_1 = \begin{bmatrix}2\\0\\1\end{bmatrix} - \left[\begin{pmatrix}2\\0\\1\end{pmatrix} \cdot \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}\right] \times \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$$

$$= \begin{pmatrix}2\\0\\1\end{pmatrix} - \left[\frac{1}{\cancel{\sqrt{3}}} \cdot \cancel{3}\right] \times \frac{1}{\cancel{\sqrt{3}}}\begin{pmatrix}1\\1\\1\end{pmatrix}$$

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

$$e_2 = \frac{u_2}{\|u_2\|} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}$$

$$u_3 = V_3 - (V_3 \cdot e_1)e_1 - (V_3 \cdot e_2)e_2$$

$$\begin{pmatrix}3\\1\\-1\end{pmatrix} - \left[\begin{pmatrix}3\\1\\-1\end{pmatrix} \cdot \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}\right] \times \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix} - \left[\begin{pmatrix}3\\1\\-1\end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}\right] \times \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}$$

$$\begin{pmatrix}3\\1\\-1\end{pmatrix} - [3] \times \frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} - [2] \times \frac{1}{2}\begin{pmatrix}1\\-1\\0\end{pmatrix}$$

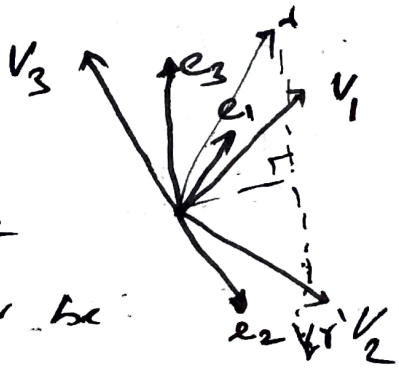
$$\begin{pmatrix}3\\1\\-1\end{pmatrix} - \begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}1\\-1\\0\end{pmatrix} = \begin{pmatrix}1\\1\\-2\end{pmatrix}$$

$$u_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

$$e_3 = \frac{u_3}{|u_3|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

$$e_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad e_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad e_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

These are orthonormal basis vectors of unit length



1) e_1 is the normalised version of V_1

2) e_2 is the perpendicular part of V_2 to e_1, normalised to be of unit length

3) e_3 is normal to the plane spanned by basis vectors e_1 + e_2. e_3 is the bit of V_3 that we can't make by projecting onto V_1 + V_2

$$E = \left[\begin{pmatrix} e_1 \end{pmatrix} \begin{pmatrix} e_2 \end{pmatrix} \begin{pmatrix} e_3 \end{pmatrix} \right] \Rightarrow \left[\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right]$$

Suppose we have a vector r and you want to reflect in the plane. How do we obtain it.

In the plane's basis we would represent the transformation $T_E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

But r is not in the plane's basis. r is in a basis of V_1, V_2, V_3 two ... i.e the basis on which V_1, V_2, V_3 were represented.

The orthonormal unit vectors derived out of Gram-Schmidt's process gives e_1, e_2, e_3 in terms of V_1, V_2, V_3.

If we inverse E we will be in the plane's basis. Remember V_1, V_2, V_3 aren't exactly in the plane's basis. i.e not expressed in the terms of the plane's basis

$r \cdot E^{-1} \rightarrow$ will give r in the plane's basis

Then $T_E \rightarrow$ will transform the vector

$E \rightarrow$ will transform r to the original basis

r'

$$\begin{tikzcd} \boxed{r} \arrow[d, "E^{-1}"'] & \boxed{r'} \\ \boxed{r_E} \arrow[r, "T_E"'] & \boxed{r'_E} \arrow[u, "E"'] \end{tikzcd}$$

$$r' = E T_E E^{-1} r$$

Because E is constructed orthonormal by using Gram-Schmidt's process

$$E^{-1} = E^T$$

Reflecting Bear:

Panda bear is confused. He is trying to work out how things should look when reflected in a mirror, but is getting the wrong results. As is the way with bears, his coordinate system is not orthonormal. So what he thinks is the direction perpendicular to the mirror isn't actually the direction the mirror reflects in.

Instruction:

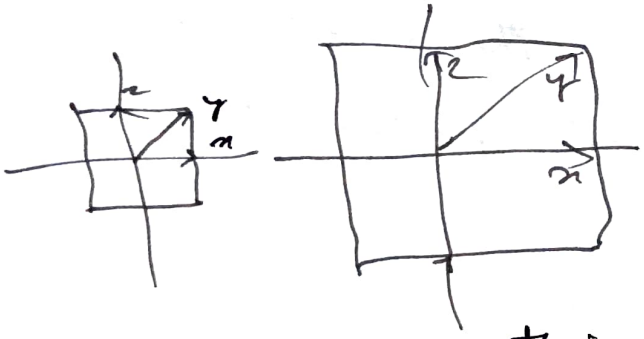
In this assignment you will write a python function that will produce a transformation matrix for reflecting vectors in an arbitrarily angled mirror

Completed

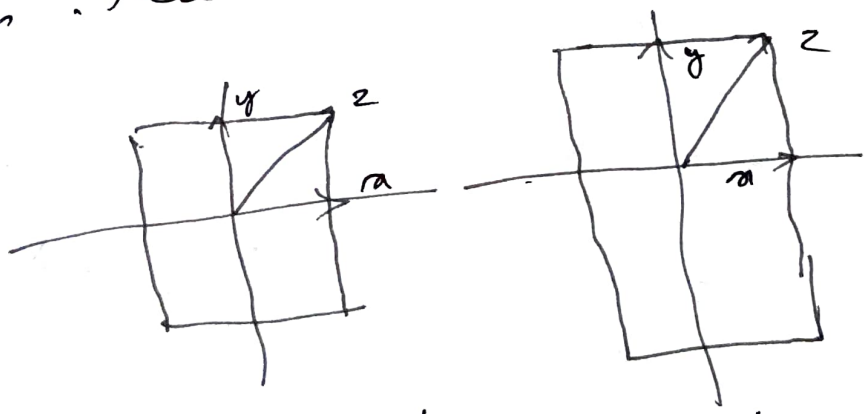
Eigen Vectors and Eigen Values:

Eigen vectors are the vectors that doesnt change their span or direction during a transformation. The factor by which the span of the eigen vector increased or decreased is called eigen value.

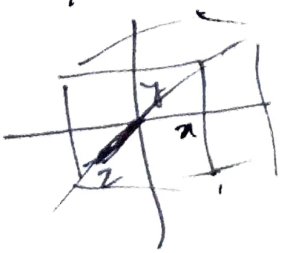
$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



Assume this is uniform transformation in all directions. Then all 3 vectors are eigen vectors.



Here x and y are the eigen vectors, the direction of $\vec{z}$ has changed. The eigen vector of a 3D volume is its axis



Calculating Eigen vectors and Eigen values algebraically

$$Ax = \lambda x$$

For ~~an~~ vector $\vec{x}$ to be eigen, both sides must be equal

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

This should evaluate to 0 for a non-trivial solution. For the matrix operation to be 0, the determinant of the resulting matrix must be 0.

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

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

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

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix}$$

$$(a-\lambda)(d-\lambda) - bc$$

$$ad - \lambda d - a\lambda + \lambda^2$$

$$\lambda^2 - (a+d)\lambda + ad - bc = 0$$

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

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

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

$$(1-\lambda)(1-\lambda) = 0$$

$$1 - \lambda - 1 + \lambda^2$$

Suppose lets scale vertically

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

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

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

$$(1-\lambda)(2-\lambda) = 0$$

$$\lambda = 1 \text{ or } \lambda = 2$$

When $\lambda = 1$ $\quad \begin{bmatrix} 1-1 & 0 \\ 0 & 2-1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$

$$= \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = 0$$

Here x_2 ~~of Comp~~ of $\vec{x}$ is 0

$\lambda = 2$ $\quad \begin{bmatrix} 1-2 & 0 \\ 0 & 2-2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 \\ 0 \end{pmatrix} = 0$

Here x_1 term must be 0

When $\lambda = 1$ $x = \begin{pmatrix} t \\ 0 \end{pmatrix}$

↳ can equal anything along the horizontal axis as long as

So $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ rotates 90°

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

$$\begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix}$$

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

$$\det\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix}\right) = 0 \Rightarrow \lambda^2 + 1 = 0$$

This has no real number solution.

So no eigen vectors in real number terms.

We can still calculate complex eigen vectors using imaginary numbers but not needed for this particular course.

Characteristic polynomials, eigenvalues and eigenvectors

For the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, what is the characteristic polynomial, and the solutions to the characteristic polynomial

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

$$Ax = \lambda x \Rightarrow \det(A - \lambda I)x = 0$$

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

$$\Rightarrow (1-\lambda)(2-\lambda) = 0 \Rightarrow \lambda = 1 \text{ or } \lambda = 2$$

$$2 - 2\lambda - \lambda + \lambda^2$$

$$\lambda^2 - 3\lambda + 2 = 0$$

characteristic polynomial

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

when $\lambda = 1$ $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} t \\ 0 \end{pmatrix}$

The x_2 component of $\vec{x}$ should be zero and x_1 can scale

@ $\lambda = 2$ $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = 0$

x_1 of $\vec{x}$ should be zero and x_2 can scale

$$\Rightarrow \begin{pmatrix} 0 \\ t \end{pmatrix}$$

$A = \begin{bmatrix} 3 & 4 \\ 0 & 5 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\begin{bmatrix} 3 & 4 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{matrix} 3 & 4 \\ 0 & 5 \end{matrix}$

$\hookrightarrow$ Ax should be equal to λx

$Ax = \lambda x \Rightarrow \underbrace{(A - \lambda I)}x = 0$

$\hookrightarrow$ Make this 0 to get a non trivial solution

$\hookrightarrow$ $\det(A - \lambda I) = 0$

$\det\left(\begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) \Rightarrow \det\begin{pmatrix} 3-\lambda & 4 \\ 0 & 5-\lambda \end{pmatrix} = 0$

So $(3-\lambda)(5-\lambda) = 0$ $\lambda = 3$ $\lambda = 5$

$15 - 3\lambda - 5\lambda + \lambda^2 = 0 \Rightarrow \lambda^2 - 8\lambda + 15 = 0$

@ $\lambda = 3$ $\begin{pmatrix} 0 & 4 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

$\div 2$ $\begin{pmatrix} 4x_2 \\ 2x_2 \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

~~$x_1 = 4x_2$~~

~~$x_1 = 4x_2$~~

$2x_2 = x_1$

$x_2 = \frac{1}{2}x_1$

@ $\lambda = 5$ $\begin{pmatrix} -2 & 4 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

$\begin{pmatrix} -2x_1 + 4x_2 \\ 0 \end{pmatrix} = 0$

$-2x_1 = -4x_2$

$x_1 = 2x_2$

$x_2 = \frac{1}{2}x_1$

5 :

$\det\left(\begin{bmatrix} 1 & 0 \\ -1 & 4 \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0 \Rightarrow \begin{bmatrix} 1 & 0 \\ -1 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$

$= \begin{bmatrix} 1-\lambda & 0 \\ -1 & 4-\lambda \end{bmatrix} \Rightarrow (1-\lambda)(4-\lambda) = 0 \Rightarrow 4 - 4\lambda - \lambda + \lambda^2 = 0$

$\lambda = 1$ $\lambda = 4$ $\lambda^2 - 5\lambda + 4 = 0$

@ $\lambda = 1$ $\begin{bmatrix} 0 & 0 \\ -1 & 3 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{pmatrix} 0 \\ -x_1 + 3x_2 \end{pmatrix} = 0$

$x_1 = 3x_2$ $x_2 = \frac{x_1}{3}$

@ $\lambda = 4$ $\begin{pmatrix} -3 & 0 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{pmatrix} -3x_1 \\ -x_1 \end{pmatrix} = 0$

$$\begin{bmatrix} 1 & 0 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

@ $\lambda\ 4\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$ @ $\lambda = 1$ $\quad 1\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

$$\cancel{\begin{bmatrix} 1 & 0 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix}} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -7 \end{bmatrix}$$

Matrix $A = \begin{bmatrix} -3 & 8 \\ 2 & 3 \end{bmatrix}$

$Ax = \lambda x$

$\underbrace{(A - \lambda I)}_{\text{make it zero}} x = 0$

$\det(A - \lambda I) = 0 \Rightarrow \begin{bmatrix} -3 & 8 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$

$= \cancel{\begin{pmatrix} -3-\lambda & 8\lambda \\ 2 & 3-\lambda \end{pmatrix}} \begin{pmatrix} -3-\lambda & 8 \\ 2 & 3-\lambda \end{pmatrix}$

$(-3-\lambda)(3-\lambda) - 16$

$\cancel{-9-3\lambda}\ -9 - 3\lambda + 3\lambda + \lambda^2 - 16$

$\cancel{\lambda^2 - 9 = 0 \quad \lambda = \pm\sqrt{9} = -3}$ $\quad \lambda^2 = 9$

$\lambda^2 - 25 = 0 \quad \lambda^2 = 25$

$\lambda = \pm\sqrt{25} \Rightarrow \pm 5$

~~@ $\lambda = +5$~~

$\cancel{\begin{bmatrix} 2 & 8 \\ 2 & -2 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{pmatrix} 2x_1 + 8x_2 \\ 2x_1 - 2x_2 \end{pmatrix} = 0}$

$\cancel{2x_1 = 2x_2}$

$\cancel{x_1 = x_2}$

$\cancel{2x_1 = -8x_2}$

$\cancel{x_1 = -4x_2}$

@ $\lambda = -5$

ⓐ $\lambda = +5$

$$\begin{bmatrix} -3-\lambda & 8 \\ 2 & 3-\lambda \end{bmatrix}$$

1 $$\begin{bmatrix} -8 & 8 \\ 2 & -2 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{pmatrix} -8x_1+8x_2 \\ 2x_1-2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$x_1 = x_2$

2 ⓐ $\lambda = -5$

$$\begin{bmatrix} 2 & 8 \\ 2 & 8 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{pmatrix} 2x_1+8x_2 \\ 2x_1+8x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$2x_1 = -8x_2$

$x_1 = -4x_2$

$x_2 = \frac{-1}{4}x_1$

Verification

$$\begin{bmatrix} -3 & 8 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -5 \\ -5 \end{bmatrix}$$

$\lambda = 5$ $\quad 5\begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{matrix} -5 \\ -5 \end{matrix}$

$$\begin{bmatrix} -3 & 8 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} -3 & 8 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

$$5\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

---

5 $A = \begin{bmatrix} 5 & 4 \\ -4 & -3 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$Ax = \lambda x$

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

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

$$\begin{bmatrix} 5 & 4 \\ -4 & -3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \Rightarrow \begin{bmatrix} 5-\lambda & 4 \\ -4 & -3-\lambda \end{bmatrix} = 0$$

$(5-\lambda)(-3-\lambda) + 16 = 0 \quad | \quad -15 + 3\lambda - 5\lambda + \lambda^2 + 16 = 0$

$(\lambda - 1)(\lambda - 1)$

$\lambda^2 - 2\lambda + 1 = 0$

$\lambda^2 - \lambda - \lambda + 1$

$(\lambda - 1)(\lambda - 1) = 0$

$\lambda = 1$

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

$$\begin{bmatrix} -2 & -3 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \Rightarrow \begin{bmatrix} -2-\lambda & -3 \\ 1 & 1-\lambda \end{bmatrix}$$

$$(-2-\lambda)(1-\lambda) + 3 = 0$$

$$-2 - \lambda + 2\lambda + \lambda^2 + 3 = 0 \Rightarrow \lambda^2 + \lambda + 1 = 0$$

$$\lambda^2 + \lambda = -1$$

$$\lambda(\lambda+1) = -1$$

Eigen Vectors - Changing basis

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}$$

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad A^3 = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}$$

When raising matrices to power diagonal matrices makes a lot easier.

Study undiagonalizable matric → Complex eigen vectors

Eigen basis conversion matrix:

What if matrix T is not a diagonal matrix.

To convert a non-diagonal matrix, we convert to a basis where the transformation T becomes diagonal.

This new basis is eigen basis.

Rough work:

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$

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

1) study undiagonalizable matrices
2) Complex eigen vectors

Diagonalisation and applications:

Given the matrix $T = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$ and change of basis matrix $C = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ (whose columns are eigen vectors of T), Calculate the diagonal matrix $D = C^{-1}TC$

$$C C^{-1} = I$$

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

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$$

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

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

$$\begin{bmatrix} 2 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 10 & 8 \\ -5 & -4 \end{bmatrix}$$

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

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

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

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\cancel{\begin{bmatrix} 1 & -1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 5 & 8 \end{bmatrix} \quad \begin{matrix} 0 & -4 \\ -15 & -20 \end{matrix}}$$

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

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

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \qquad T = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 5 & 8 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$$

Eigen Value λ_1 for $\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 5$

Eigen Value λ_2 for $\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 4$

Given the matrix $T = \begin{bmatrix} 2 & 7 \\ 0 & -1 \end{bmatrix}$ and change of basis matrix $C = \begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix}$ (whose columns are eigen vectors of T), calculate the diagonal matrix $D = C^{-1}TC$

$$C C^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ -3 & 0 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} -3 & 0 \\ 1 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 4 & 9 \\ 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 \\ 0 & -3 \end{bmatrix} C^{-1} = \begin{bmatrix} 4 & 9 \\ -3 & -7 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & -3 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & 1 \\ -3 & -7 \end{bmatrix}$$

$$C^{-1} = \begin{bmatrix} 1 & -0.3 \\ 1 & 2.3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -0.3 \\ 1 & 2.3 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -0.3 \\ 1 & 2.3 \end{bmatrix} \begin{bmatrix} -7 & 2 \\ 3 & 0 \end{bmatrix} = -8 \quad 2$$

$$\begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$R_1 \to R_2$
$R_2 \to R$

$$\begin{bmatrix} -3 & 0 \\ 7 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$R_1 \to \frac{R_1}{-3}$ $\quad \begin{bmatrix} 1 & 0 \\ 7 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & -\frac{1}{3} \\ 1 & 0 \end{bmatrix}$

$R_2 \to \frac{R_2}{7}$ $\quad \begin{bmatrix} 1 & 0 \\ 1 & 1/7 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & -\frac{1}{3} \\ \frac{1}{7} & 0 \end{bmatrix}$

$R_2 \to R_2 - R_1$ $\quad \begin{bmatrix} 1 & 0 \\ 0 & 1/7 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & -\frac{1}{3} \\ \frac{1}{7} & 1/3 \end{bmatrix}$

$R_2 \to R_2 \times 7$ $\quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & -\frac{1}{3} \\ 1 & 7/3 \end{bmatrix}$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & -0.33 \\ 1 & 2.33 \end{bmatrix}$$

$$\cancel{\begin{bmatrix} 1 & -0.33 \\ 1 & 2.33 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix}}$$

$$\cancel{= \begin{bmatrix} 1 & -0.33 \\ 1 & 2.33 \end{bmatrix} \begin{bmatrix} -7 & 2 \\ 3 & 0 \end{bmatrix} = \begin{matrix} -6 & 2 \\ -4 & 2 \end{matrix}}$$

$$\cancel{\begin{bmatrix} 1 & -0.33 \\ 1 & 2.33 \end{bmatrix} \begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ 0 & -1 \end{bmatrix}}$$

$$\cancel{= \begin{bmatrix} 1 & -0.33 \\ 1 & 2.33 \end{bmatrix} \begin{bmatrix} 14 & 48 \\ -6 & -21 \end{bmatrix} = [}$$

$$\begin{bmatrix} 0 & -0.33 \\ 1 & 2.33 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & -0.33 \\ 1 & 2.33 \end{bmatrix} \begin{bmatrix} -7 & 2 \\ 3 & 0 \end{bmatrix}$$

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

Given the matrix $T = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$ and change of basis matrix $C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ (whose columns are eigenvectors of T). Calculate the diagonal matrix $D = C^{-1}TC$

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$R_2 \to R_1$
$R_1 \to R_2$

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$R_1 \to R_1 - R_2$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} C^{-1} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

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

Given the diagonal matrix $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, and a change of basis matrix $C = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ with inverse $C^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$.
Calculate $T = CDC^{-1}$

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

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

125

Given that $T = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 125 & 0 \\ 0 & 64 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 250 & -125 \\ -64 & 64 \end{bmatrix}$$

$$= \begin{bmatrix} 186 & -61 \\ 186 & -186 \end{bmatrix}$$

$$= \begin{bmatrix} 186 & -61 \\ 122 & 3 \end{bmatrix}$$

250
64
186

125
64
61

250
128
122

250
64
186

-125
128
3

$T = \begin{bmatrix} 2 & 7 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & -1/3 \\ 1 & 7/3 \end{bmatrix}$ calculate T^3

$$\begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} 0 & -1/3 \\ 1 & 7/3 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & 1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1/3 \\ 8 & 56/3 \end{bmatrix} = \begin{bmatrix} 8 & 63/3 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 8 & 21 \\ 0 & -1 \end{bmatrix}$$

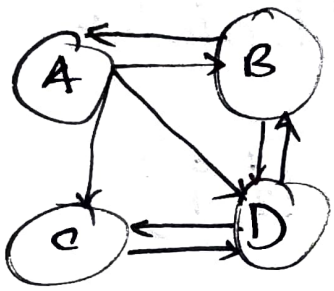
Given that $T = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$. calculate T^5

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

Page Rank algorithm

This algorithms decides which order pages are displayed.

The importance of a website was decided based on the links to and from other website



$$\begin{array}{lcccc} & A & B & C & D \\ L_A = & (0, & 1/3, & 1/3, & 1/3) \\ L_B = & (1/2, & 0, & 0, & 1/2) \\ L_C = & (0, & 0, & 0, & 1) \\ L_D = & (0, & 1/2, & 1/2, & 0) \end{array}$$

$$\text{Link matrix } L = \left[(L_A)(L_B)(L_C)(L_D)\right] = \begin{array}{c} \begin{array}{cccc} L_A & L_B & L_C & L_D \end{array} \\ \begin{bmatrix} 0 & 1/2 & 0 & 0 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 1/2 & 1 & 0 \end{bmatrix} \end{array}$$

The original page rank formula with summation:

$$PR(A) = (1-d) + d\left(\frac{PR(T_1)}{C(T_1)} + \frac{PR(T_2)}{C(T_2)} + \dots + \frac{PR(T_n)}{C(T_n)}\right)$$

$PR(A)$ page rank of page A ~ kind of recursive formula because it depends on other page's page rank

$PR(T_i)$ Page rank of pages T_i which link to page A

$C(T_i)$ number of outbounds links on a given T_i pg

d damping factor in the range of 0 and 1.

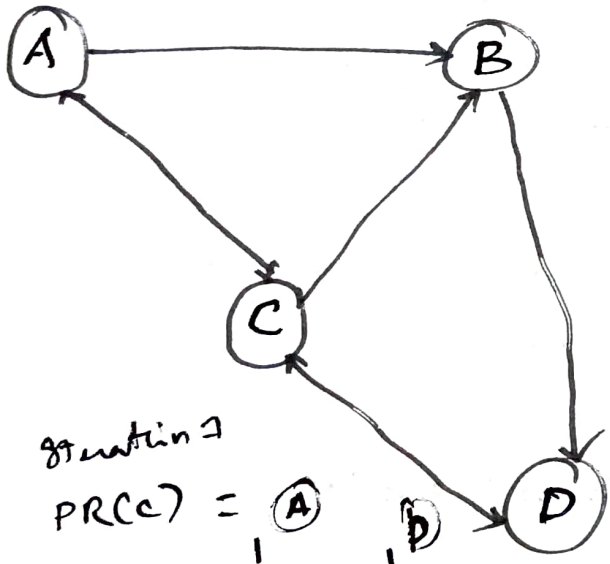
We have to initialize page ranks at the beginning.
All pages are given equal page rank $\frac{1}{n}$ where n is the number of pages.
Thats why we have to make several iterations until convergence.

The iterative formula:

$$PR_{t+1}(P_i) = \sum_{P_j} \frac{PR_t(P_j)}{C(P_j)}$$

→ Sum of the page ranks of pages pointing to the target page (i) in the previous iteration

÷ No of outgoing links in the page pointing to the target page in the previous iteration



iteration 1

$PR(A) = \frac{\frac{1}{4}}{3}$ → Page Rank of C in the previous iteration

C 3 outbound links

$= \frac{1}{12}$

~~$PR(B) = \frac{\frac{1}{4}}{3}$~~

$PR(B) = \frac{\frac{1}{4}}{2} + \frac{\frac{1}{4}}{3}$

$= \frac{1}{8} + \frac{1}{12} = \frac{20}{96} = \frac{2.5}{12}$

iteration 1

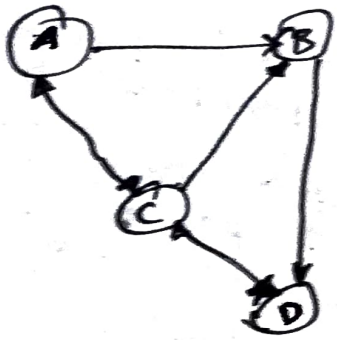
$PR(C) = \frac{\frac{1}{4}}{2} + \frac{\frac{1}{4}}{1} \Rightarrow \frac{1}{8} + \frac{1}{4}$

$= \frac{3}{8} = \frac{4.5}{12}$

	iteration 0	iteration 1	iteration 2	PageRank
A	1/4	1/12	4.5/12/3 ~~$\frac{1}{96}$~~ = 1.5/12	1
B	1/4	$\frac{2.5}{12}$	$\frac{\frac{1}{12}}{2} + \frac{\frac{4.5}{12}}{3} = \frac{2}{12}$	2
C	1/4	4.5/12	$\frac{\frac{1}{12}}{2} + \frac{\frac{4}{12}}{1}$	4
D	1/4	4/12	$\frac{4.5/12}{3} + \frac{2.5/12}{1}$	3
+	1	1	1	1

Matrix representation of Page Rank algorithm:

We can use matrix operations instead of the iterative approach. We updated values one by one: we can use matrix operations to do multiple calculations at the same time.



$$\begin{array}{c|cccc} & A & B & C & D \\ \hline A & 0 & 0 & \frac{1}{3} & 0 \\ B & \frac{1}{2} & 0 & \frac{1}{3} & 0 \\ C & \frac{1}{2} & 0 & 0 & 1 \\ D & 0 & 1 & \frac{1}{3} & 0 \end{array}$$

Columns are outgoing links
rows are incoming links

$$v = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

$$V_2 = Hv$$
$$V_3 = HV_2 = H(Hv) = H^2 v$$
$$V_n = H^n v$$

<u>Steady state approach</u>

We have the directed graph: we can construct a transition matrix ~ the steady state is where the eigenvalue is 1.

$$\begin{bmatrix} 0 & 0 & 1/3 & 0 \\ \frac{1}{2} & 0 & 1/3 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \\ 0 & 1 & 1/3 & 0 \end{bmatrix}$$

use damping factor in the random surfer model

d damping factor is in the range 0 and 1. Usually initialized to 0.15

$$M = (1-d)A + dB$$

M: this is the page rank matrix or google matrix

d: damping factor

A: transition matrix

B: identity matrix if items are all 1

$$B = \frac{1}{N}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

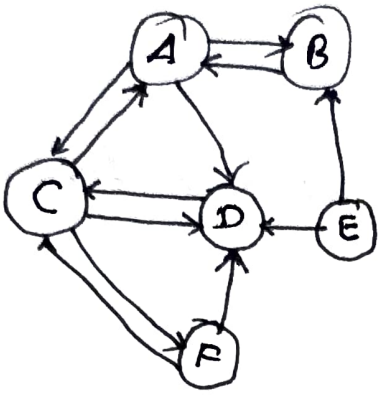
$N \rightarrow$ no of pages

Coursera

$$r_a = \sum_{j=1}^{n} L_{A,j}\, r_j$$

$$r^{i+1} = L r^{i}$$

$$r^{i+1} = d(L r^{i}) + \frac{1-d}{n}$$



A - Avocado
B - Bullseye
C - CatBabel
D - Dormeda
E - eTings
F - FaceSpace

The design principle of PageRank is that important websites will be linked to by important websites.

$$r = \begin{bmatrix} r_A \\ r_B \\ r_C \\ r_D \\ r_E \\ r_F \end{bmatrix}$$

We represent the number of Pats on each website with vector r

And say that number of Pats on each website in minute $i+1$ is related to those at minute i by the matrix transformation.

$r^{(i+1)} = L r^{i}$ with the matrix L taking the form

$$L = \begin{bmatrix} L_{A\to A} & L_{B\to A} & L_{C\to A} & L_{D\to A} & L_{F-A} \\ L_{A\to B} & L_{B-B} & L_{C\to B} & L_{D-B} & L_{F-B} \\ L_{A\to C} & L_{B\to C} & L_{C\to C} & L_{D-C} & L_{F-C} \\ L_{A\to D} & L_{B\to D} & L_{C\to D} & L_{D-D} & L_{F-D} \\ L_{A\to E} & L_{B\to E} & L_{C\to E} & L_{D-E} & L_{F-E} \\ L_{A\to F} & L_{B\to F} & L_{C\to F} & L_{D-F} & L_{F-F} \end{bmatrix}$$

While the columns represent the probability of leaving a website for any other website and sum to one. The rows determine how likely you are to enter a website going any other through these and not sum to one.

The long term behaviour of this system is when $r^{(k+1)} = r^{(k)}$, so we'll drop the superscript here and that allows us to write

$$Lr = r$$

which is an eigen value equation of the matrix L

~~A~~

Given the matrix $\begin{bmatrix} 3/2 & -1 \\ -1/2 & 1/2 \end{bmatrix}$ calculate its characteristic polynomial

~~$3/2 \quad -1$~~

$$Ax = \lambda x$$

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

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

$$\begin{bmatrix} 3/2 & -1 \\ -1/2 & 1/2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

$$\cancel{3/2\lambda - \lambda - 1/2 + 1/2\lambda}$$

$$\begin{bmatrix} 3/2 - \lambda & -1 \\ -1/2 & 1/2 - \lambda \end{bmatrix}$$

$$(3/2 - \lambda)(1/2 - \lambda) - 1/2 = 0$$

$$3/4 - \lambda/2 - 3\lambda/2 + \lambda^2 - 1/2 = 0$$

$$\lambda^2 - 2\lambda + \frac{1}{4} = 0$$

$$\cancel{(\lambda - \tfrac{1}{2})^2}$$

$$\lambda^2 - 2\lambda\tfrac{1}{2}$$

$$\times 4 \quad \cancel{4\lambda^2 - 8 + 1 = 0}$$

$$\cancel{(\lambda - \tfrac{1}{2})(\lambda - \tfrac{1}{2}) = 0}$$

$$(\lambda - \tfrac{1}{2})^2$$

$$\lambda^2 + \tfrac{1}{4}$$

$$\lambda + \tfrac{1}{2}$$

$$3/2 - \lambda = 1/2$$

$$-\lambda = 1/2 - 3/2$$

$$-\lambda = -1 \quad \lambda = 1$$

$$1/2 - \lambda = 1/2$$

$$3/4 - 1/2$$

$$\frac{6-4}{8} = \frac{2}{8}$$

$$\lambda^2 - 2\lambda + \frac{1}{4} = 0$$

$$\lambda^2 - 2\lambda = -\frac{1}{4}$$

$$\lambda(\lambda - 2) = -\frac{1}{4}$$

$$3/2 - \lambda$$

$$\lambda = \frac{1}{2}(\lambda - 2)$$

$$\lambda^2 - 2\lambda + \frac{1}{4} = 0$$

$$a \cdot \frac{1}{4} \cdot \lambda = 2\lambda$$

Rough Work

$$\begin{bmatrix} 3/2-\lambda & -1 \\ -1/2 & 1/2-\lambda \end{bmatrix}$$

When $\lambda = \frac{1}{2}$

$$\begin{bmatrix} 3/2-1/2 & -1 \\ -1/2 & 1/2-1/2 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

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

$$\begin{pmatrix} x_1 - x_2 \\ -1/2\, x_1 \end{pmatrix} = 0$$

$x_2 = -1/2\, x_1$

$x_1 + 1/2\, x_1 = \frac{3}{2} x_1$

$x_1 + \frac{1}{2} x_1 = 0$

$\frac{3}{2} x_1 = 0$

$(\lambda - \frac{1}{2})(\lambda - \frac{1}{2})$

$\lambda^2 - \frac{1}{2}\lambda - \frac{1}{2}\lambda + \frac{1}{4}$

$\lambda^2 - \lambda + \frac{1}{4}$

$\frac{3}{2} + \frac{1}{4}$

$\frac{14}{8} + \frac{1}{4}$

$\frac{14}{8}$

$\frac{5\lambda + 8}{32}$

$\frac{3}{2} \times \frac{1}{2} = \frac{4}{2}$

$\lambda^2 - 2\lambda + \frac{1}{4} = 0$

$(3/2 - \lambda)(1/2 - \lambda) - \frac{1}{2} = 0$

$3/2 - \lambda = 1/2$ $\quad$ $1/2 - \lambda = 1/2$

$-\lambda = 1/2 - 3/2$ $\quad$ $-\lambda = -1$

$-\lambda = -\frac{1}{2}$ $\quad$ $\lambda = \frac{1}{4}$

$\lambda = \frac{1}{2}$ $\quad$ $\lambda = \frac{1}{4}$

$\lambda^2 - 2\lambda + \frac{1}{4}$

$\lambda^2 - \frac{14}{8}\lambda - \frac{1\lambda}{4} + \frac{1}{4}$

$\lambda^2 - \frac{7}{4}\lambda - \frac{1}{4}\lambda + \frac{1}{4}$

$\frac{1}{4} \times \frac{8^2}{11}$

$\lambda^2 - \frac{1}{4}\lambda - \frac{14}{8}\lambda + \frac{1}{4}$

$\lambda(\lambda - \frac{1}{4}) - 1(\lambda + \frac{1}{4})$

$3/2 - 1/2$

$$\begin{bmatrix} 3/2-\lambda & -1 \\ -1/2 & 1/2-\lambda \end{bmatrix}$$

(a) $\lambda = 1/2$

$$\begin{bmatrix} 3/2-1/2 & -1 \\ -1/2 & 0 \end{bmatrix}$$

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

$\frac{3}{2}+\frac{1}{2}$ $\quad \frac{3}{2}+\frac{1}{2}$ $\quad \frac{3}{2}$

$\frac{2}{8}+1$ $\quad \frac{2}{8}+\frac{1}{2}$ $\quad 2\frac{18}{4}$

$\frac{10}{8}$ $\quad 2\frac{1}{8}$ $\quad \frac{12}{16}-\frac{1}{8}$

$\frac{1}{4}$ — $1\times\frac{1}{4}$

— $\frac{1}{2}\times\frac{1}{2}$

— $1\times\frac{2}{8}$

2

$\downarrow$

$\frac{3}{2}+\frac{1}{2}$

$\frac{16}{8} = 2$ $\quad \frac{2}{8}$ $\quad -2$

$-\frac{14}{8}$

(a) $\lambda =$

Rough Work

$= \frac{1}{4}$

$\frac{8}{4}$ $\quad \frac{1}{4}$

$Ax = \lambda x$

$\underbrace{(A-\lambda I)}_{\det = 0} x = 0$

$$\begin{bmatrix} 3/2+1+\sqrt{5}/2 & -1 \\ -1/2 & \frac{1}{2}+1-\frac{\sqrt{5}}{2} \end{bmatrix}$$

$$\begin{bmatrix} 5/2+\sqrt{5}/2 & -1 \\ -1/2 & \frac{3}{2}-\frac{\sqrt{5}}{2} \end{bmatrix}$$

$\left(5/2+\frac{\sqrt{5}}{2}\right)\left(\frac{3}{2}-\frac{\sqrt{5}}{2}\right)$

$-1/2 = 0$

$\cancel{\frac{15}{4}-\frac{5\sqrt{5}}{4}+\frac{3\sqrt{5}}{4}+\frac{15}{4}}$

$\frac{15}{4}-\frac{5\sqrt{5}}{4}+\frac{3\sqrt{5}}{4}-\frac{5}{4}$

$\frac{10}{4}-\frac{2\sqrt{5}}{4}$

$3/2 - \lambda \qquad -1$

$-1/2$

Rough Work

$$\begin{bmatrix} 3/2 - \lambda & -1 \\ -1/2 & 1/2 - \lambda \end{bmatrix}$$

$$\begin{bmatrix} 3/2 - 1 + \frac{\sqrt{5}}{2} & -1 \\ -1/2 & 1/2 - 1 - \frac{\sqrt{5}}{2} \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{2} + \frac{\sqrt{5}}{2} & -1 \\ -1/2 & -1/2 - \frac{\sqrt{5}}{2} \end{bmatrix}$$

$$\left(\frac{1}{2} + \frac{\sqrt{5}}{2}\right)\left(-1/2 - \frac{\sqrt{5}}{2}\right) - 1/2$$

$$-\frac{1}{4} - \frac{\sqrt{5}}{4} - \frac{\sqrt{5}}{4} - \frac{5}{4}$$

$$-\frac{6}{4} - \frac{2\sqrt{5}}{4}$$

$$\begin{bmatrix} \frac{3}{2} - 1 + \frac{\sqrt{3}}{2} & -1 \\ -\frac{1}{2} & 1/2 - 1 - \frac{\sqrt{3}}{2} \end{bmatrix}$$

$$\left(1/2 + \frac{\sqrt{3}}{2}\right)\left(1/2 - \frac{\sqrt{3}}{2}\right)$$

$$1/4 + \frac{\sqrt{3}}{4} - \frac{\sqrt{3}}{4} - \frac{3}{4}$$

$$-\frac{2}{4}$$

$$\begin{bmatrix} \frac{1}{2} + \frac{\sqrt{3}}{2} & -1 \\ -\frac{1}{2} & 1/2 - \frac{\sqrt{3}}{2} \end{bmatrix} =$$

$$\begin{bmatrix} \frac{3}{2}+1+\frac{\sqrt{3}}{2} & -1 \\ -1/2 & 1/2+1-\frac{\sqrt{3}}{2} \end{bmatrix}$$

Rough Work

$$\begin{bmatrix} \frac{5}{2}+\frac{\sqrt{3}}{2} & -1 \\ -\frac{1}{2} & \frac{3}{2}-\frac{\sqrt{3}}{2} \end{bmatrix} \rightarrow \left(\frac{5}{2}+\frac{\sqrt{3}}{2}\right)\left(\frac{3}{2}-\frac{\sqrt{3}}{2}\right)-\frac{1}{2}$$

$$= \frac{15}{4}-\frac{5\sqrt{3}}{4}+\frac{3\sqrt{3}}{4}-\frac{3}{4}$$

$$= \frac{12}{4}-\frac{2\sqrt{3}}{4}$$

$$\begin{bmatrix} \frac{3}{2}-\lambda & -1 \\ -1/2 & 1/2-\lambda \end{bmatrix} \rightarrow \begin{bmatrix} 3/2+1+\frac{\sqrt{5}}{2} & -1 \\ -1/2 & 1/2+1-\frac{\sqrt{5}}{2} \end{bmatrix} \rightarrow \begin{bmatrix} \frac{5}{2}+\frac{\sqrt{5}}{2} & -1 \\ -1/2 & \frac{3}{2}-\frac{\sqrt{5}}{2} \end{bmatrix}$$

$$\left(\frac{5}{2}+\frac{\sqrt{5}}{2}\right)\times\left(\frac{3}{2}-\frac{\sqrt{5}}{2}\right)-\frac{1}{2}=0$$

$$\frac{15}{4}-\frac{5\sqrt{5}}{4}+\frac{3\sqrt{5}}{4}-\frac{5}{4}-\frac{1}{2} = \frac{10}{4}-\frac{1}{2}-\frac{5\sqrt{5}}{4}+\frac{3\sqrt{5}}{4}$$

$$\cancel{\frac{8}{4}-\frac{2\sqrt{5}}{4}} \quad \frac{8}{4}-\frac{2\sqrt{5}}{4} = 2-1 = 1$$

$$\begin{bmatrix} \frac{3}{2}-1+\frac{\sqrt{3}}{2} & -1 \\ -\frac{1}{2} & 1/2-1-\frac{\sqrt{3}}{2} \end{bmatrix} \rightarrow \begin{bmatrix} \frac{1}{2}+\frac{\sqrt{3}}{2} & -1 \\ -\frac{1}{2} & 1/2-\frac{\sqrt{3}}{2} \end{bmatrix}$$

$$\left(\frac{1}{2}+\frac{\sqrt{3}}{2}\right)\left(\frac{1}{2}-\frac{\sqrt{3}}{2}\right)-\frac{1}{2}$$

$$\frac{1}{4}-\frac{3}{4}-\frac{1}{2}$$

$$-\frac{2}{4}-\frac{1}{2} = -1$$

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

$$\frac{15}{4}-\frac{5\sqrt{3}}{4}+\frac{3\sqrt{3}}{4}-\frac{3}{4}-1/2=0 \quad \Big| \quad \frac{12}{4}-1/2-\frac{2\sqrt{3}}{4}$$

$$\frac{10}{4}-\frac{2\sqrt{3}}{4}$$

$$\begin{bmatrix} \frac{3}{2} - \lambda & -1 \\ -1/2 & 1/2 - \lambda \end{bmatrix} = 0$$

$$\begin{bmatrix} \frac{3}{2} - 1 + \frac{\sqrt{5}}{2} & -1 \\ -\frac{1}{2} & 1/2 - 1 - \frac{\sqrt{5}}{2} \end{bmatrix} \rightarrow \left(\frac{3}{2} - 1 + \frac{\sqrt{5}}{2}\right)\left(\frac{1}{2} - 1 - \frac{\sqrt{5}}{2}\right) - \frac{1}{2}$$

$$\rightarrow \left(\frac{1}{2} + \frac{\sqrt{5}}{2}\right)\left(-\frac{1}{2} - \frac{\sqrt{5}}{2}\right) - 1/2$$

$$= -\frac{1}{4} - \frac{\sqrt{5}}{4} - \frac{\sqrt{5}}{4} - \frac{5}{4} - 1/2$$

$$-\frac{1}{4} - \frac{2\sqrt{5}}{4} - \frac{5}{4} - 1/2$$

$$\begin{bmatrix} \frac{3}{2} + 1 + \frac{\sqrt{5}}{2} & -1 \\ -1/2 & 1/2 + 1 - \frac{\sqrt{5}}{2} \end{bmatrix} \rightarrow \begin{bmatrix} \frac{5}{2} + \frac{\sqrt{5}}{2} & -1 \\ -1/2 & 3/2 - \frac{\sqrt{5}}{2} \end{bmatrix}$$

$$\left(\frac{5}{2} + \frac{\sqrt{5}}{2}\right)\left(3/2 - \sqrt{5}/2\right) - 1/2$$

$$\frac{15}{4} - \frac{5\sqrt{5}}{4} + \frac{3\sqrt{5}}{4} - \frac{5}{4} - 1/2$$

$$\frac{10}{4} - \frac{2\sqrt{5}}{4} - 1/2 = \frac{8}{4} - \frac{2\sqrt{5}}{4}$$

$$= 2 - 1$$

$$= 1$$

$\lambda^2 - 2\lambda + \frac{1}{4} = 0$

$$-\frac{1-\sqrt{3}}{2} + -\frac{1+\sqrt{3}}{2}$$

$$\left(\frac{1-\sqrt{3}}{2}\right)\left(\frac{1+\sqrt{3}}{2}\right)$$

$$1 - \frac{3}{4} = \frac{1}{4}$$

Select the two eigen vector of the matrix A. $\begin{bmatrix} 3/2 & -1 \\ -1/2 & 1/2 \end{bmatrix}$

$$D = C^{-1}AC = \overset{C}{\begin{bmatrix} 1-\sqrt{3} & 1+\sqrt{3} \\ 1 & 1 \end{bmatrix}}$$

$$CC^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1-\sqrt{3} & 1+\sqrt{3} \\ 1 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$R_1 \to \frac{R_1}{(1-\sqrt{3})} R_2$

$$\begin{bmatrix} 1 & \frac{1+\sqrt{3}}{1-\sqrt{3}} \\ & \end{bmatrix} C^{-1} =$$

$R_1 \to \frac{R_1}{1-\sqrt{3}}$

$$\begin{bmatrix} 1 & \frac{1+\sqrt{3}}{1-\sqrt{3}} \\ 1 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} \frac{1}{1-\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix}$$

$R_1 \to R_1 + \sqrt{3}$

$$\begin{bmatrix} 1 & 1+2\sqrt{3} \\ 1 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 1+\sqrt{3} & \sqrt{3} \\ 0 & 1 \end{bmatrix}$$

$\cancel{R_2 \to R}$

$R_1 \to R_1 + \sqrt{3}$

$$\begin{bmatrix} 1 & 1+2\sqrt{3} \\ 1 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 1+\sqrt{3} & \sqrt{3} \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 2\sqrt{3} \\ 1 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} \sqrt{3} & \sqrt{3}-1 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 2\sqrt{3} \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & 1 \\ \sqrt{3} & \sqrt{3}-1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} C^{-1} = \begin{bmatrix} 0 & 1 \\ \frac{\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \end{bmatrix}$$

$$1 - \frac{\sqrt{3}-1}{2\sqrt{3}}$$

$$\frac{2\sqrt{3}-\sqrt{3}-1}{2\sqrt{3}}$$

$$\frac{\sqrt{3}-1}{2\sqrt{3}}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} C^{-1}$$

$$= \begin{bmatrix} \frac{-\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \\ \frac{\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \end{bmatrix}$$

$$\cancel{\begin{bmatrix} 3/2 & -1 \\ -1/2 & 1/2 \end{bmatrix}}$$

$$C^{-1} = \begin{bmatrix} \frac{-\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \\ \frac{\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \end{bmatrix}$$

$$C = \begin{bmatrix} 1-\sqrt{3} & 1+\sqrt{3} \\ 1 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 3/2 & -1 \\ -1/2 & 1/2 \end{bmatrix}$$

$$\begin{bmatrix} \frac{-\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \\ \frac{\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \end{bmatrix} \begin{bmatrix} 3/2 & -1 \\ -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1-\sqrt{3} & 1+\sqrt{3} \\ 1 & 1 \end{bmatrix}$$

$$D = C^{-1} A C$$

$$\begin{bmatrix} 1+\frac{\sqrt{3}}{2} & 0 \\ 0 & 1-\frac{\sqrt{3}}{2} \end{bmatrix}$$

C^{-1}

$$\begin{bmatrix} -\frac{\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \\ \frac{\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \end{bmatrix} \begin{bmatrix} 1+\frac{\sqrt{3}}{2} & 0 \\ 0 & 1-\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1-\sqrt{3} & 1+\sqrt{3} \\ 1 & 1 \end{bmatrix}$$

$\cancel{\frac{7}{4}}$ 0

$\cancel{0}$

$1+\frac{3}{4}$

$$\begin{bmatrix} -\frac{\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \\ \frac{\sqrt{3}}{2\sqrt{3}} & \frac{\sqrt{3}-1}{2\sqrt{3}} \end{bmatrix} \begin{bmatrix} \frac{7}{4}+\sqrt{3} & 0 \\ 0 & \frac{7}{4}-\sqrt{3} \end{bmatrix} \begin{bmatrix} 1-\sqrt{3} & 1+\sqrt{3} \\ 1 & 1 \end{bmatrix} + \frac{\cancel{2}\sqrt{3}}{\cancel{2}}$$

$\frac{7}{4}+\sqrt{3}$

$\left(\frac{7}{4}+\sqrt{3}\right)(1-\sqrt{3})$

$\left(\frac{7}{4}+\sqrt{3}\right)(1+\sqrt{3})$

$\frac{7}{4}+\sqrt{3}+\frac{7}{4}\sqrt{3}+3$

$\frac{7}{4}-\frac{7}{4}\sqrt{3}+\sqrt{3}-3$

$\frac{7}{4}-\frac{3\sqrt{3}}{4}-3$

$\frac{7-3\sqrt{3}}{4}-3$

$\frac{7-3\sqrt{3}-12}{4}$

$\frac{-5-3\sqrt{3}}{4}$

2.732 1

1.366 -1

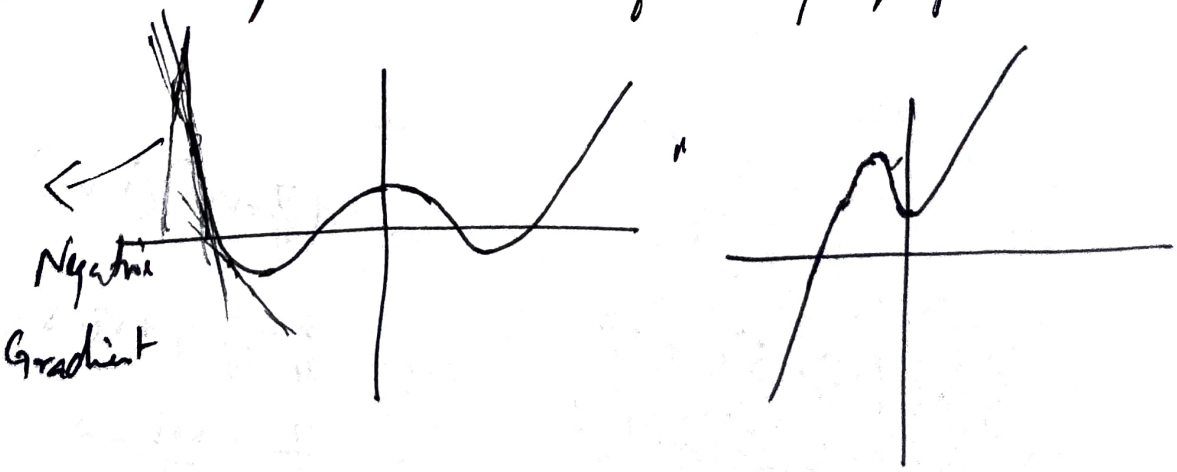
Multivariate calculus:

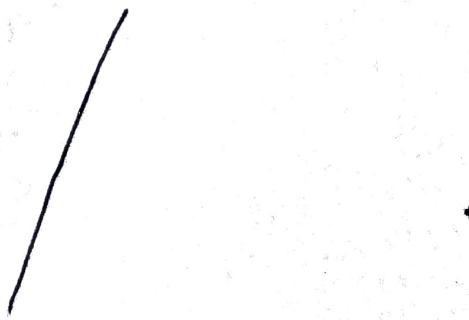
Leibniz Lagrange Newton Euler

Calculus is a set of tools that explains the relationship between a function and a change in its variables.

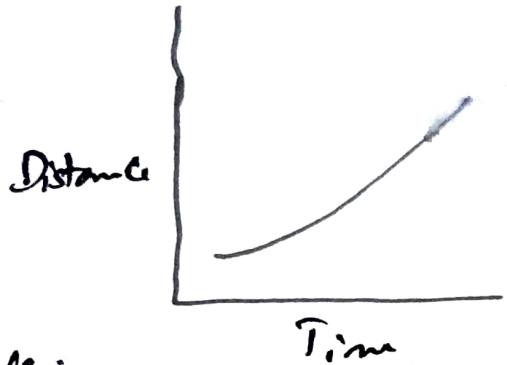
The local gradient of speed-time graph is acceleration

Which diagram best describes the differential of the function in the following graph





What is a derivative:



Assume this is a distance ~~x~~ time graph. Then the local gradient at every point gives the derivative speed.

Is there a function where it is equal to its own derivative?

i.e value of $f(x)$ is always equal to value of $f'(x)$

example when distance covered is equal to speed
when speed is equal to acceleration

$f(x) = f'(x)$

For such a function the value of the function and the value of its derivative should always be either positive or negative.

If the function + its derivative can take both positive and negative values then ~~they~~ the function would have to cross 0 where the derivative is undefined

e^x

$f(x) = e^x \quad f'(x) = e^x$

$e = 2.71828 \rightarrow$ Euler's number

Exercise

What function would differentiate to get

$$f'(x) = \frac{3\pi x^4}{4} + 11x^2 + \sqrt{2}$$

$$f(x) = \frac{3\pi x^5}{20} + \frac{11x^3}{3} + \sqrt{2}x + 2$$

$$= \frac{12\pi x^4}{20} + \frac{33x^2}{3} + \sqrt{2}$$

$$= \frac{3\pi x^4}{4} + 11x^2 + \sqrt{3}$$

When given distance as a function of time (i.e distance $= x$, $x = x(t)$)
one can calculate the rate of change of distance (that is, speed) as a function of time by differentiating $x(t)$ wrt t.

$$f(x) = x^3 + \frac{x^2}{3} + 2$$

$$3x^2 + \frac{2x}{3}$$

$x = 5$ $\quad 75 + \frac{50}{3}$ $\quad 75 + \frac{10}{3}$ $\quad$ 3) 225 (75
21
15

$$= \frac{235}{3}$$

$f(x) = x^3 + 27x^2 - 5x + 9$

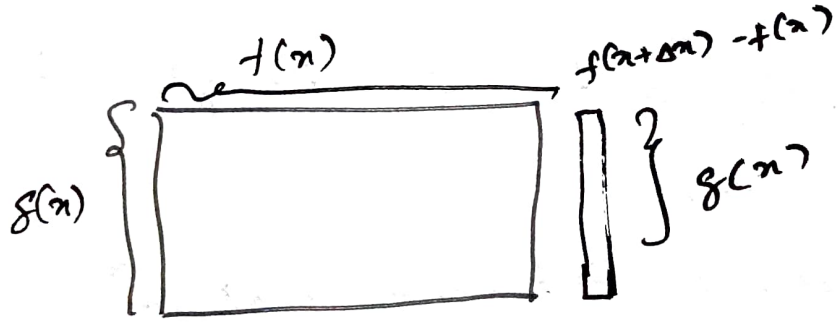
$f'(x) = 3x^2 + 54x - 5$

$x = -1$ $\quad 3 - 54 - 5$

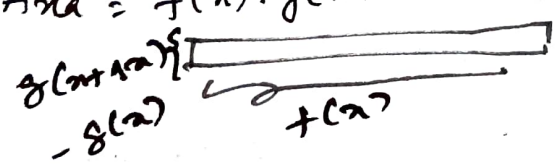
-56

Sum rule, Power rule, Product rule

Assume a rectangle



Area $= f(x) \cdot g(x)$



$\square \rightarrow$ Ignore this.

$A = f(x) \cdot g(x)$

Change in Area wrt x

$$\frac{\Delta A}{\Delta x} = \frac{f(x) \cdot (g(x+\Delta x) - g(x)) + g(x) \cdot (f(x+\Delta x) - f(x))}{\Delta x}$$

$$= f(x) \cdot g'(x) + g(x) \cdot f'(x)$$

$\Rightarrow$ Product rule

$A(x) = f(x)g(x)$

$A'(x) = f(x)g'(x) + g(x)f'(x)$

Using the product rule, differentiate wrt x the function

$f(x) = (x+2)(3x-3)$

$= (x+2)3 + (3x-3)$

$= 3x + 6 + 3x - 3$

$= 6x + 3$

Using the product rule, differentiate wrt x the function

$f(x) = e^x(x^2+7x-3-\sin(x))$

$= e^x(2x+7-\cos x) + (x^2+7x-3-\sin(x))e^x$

$= e^x(2x+7-\cos x + x^2+7x-3-\sin(x))$

$= e^x(9x+4+x^2-\cos(x)-\sin(x))$

$f(x) = \sin(x)$
$f'(x) = \cos(x)$
$f''(x) = -\sin(x)$
$f'''(x) = -\cos(x)$

Using the product rule, differentiate wrt x the function

$f(x) = (x^2-2x+5)(\sin(x)+\cos(x))$

$= (x^2-2x+5)(\cos(x)-\sin(x))$
$+ (\sin(x)+\cos(x))(2x-2)$

$= x^2\cos(x) - 2x\cos(x) + 5\cos(x) - x^2\sin(x) + 2x\sin(x)$
$-5\sin(x) + 2x\sin(x) - 2\sin(x) + 2x\cos(x) - 2\cos(x)$

$= \cancel{x^2\cos(x) + 3\cos(x) - x^2\sin(x) + 4x\sin}$

$\cos(x)(x^2-2x+5) - \sin(x)(x^2-2x+5)$

$+ \sin(x)(2x-2) + \cos(x)(2x-2)$

$x^2\cos x - 2x\cos x + 5\cos x + 2x\cos x - 2\cos x$

$\rightarrow x^2\cos x + 3\cos x$

$2x\sin x - 2\sin(x) - x^2\sin x + 2x\sin x - 5\sin(x)$

$4x\sin x - 7\sin(x) - x^2\sin x$

$x^2\cos x + 3\cos x + 4x\sin x - 7\sin(x) - x^2\sin(x)$

$x^2\cos x - x^2\sin x + 3\cos x + 4x\sin x - 7\sin x$

$f(x) = (x^2-2x+5)(\sin(x)+\cos(x))$

$f'(x) = (x^2-2x+5)(\cos(x)-\sin(x))$

$+ (2x-2)(\sin(x)+\cos(x))$

$x^2\cos x - 2x\cos(x) + 5\cos(x) + 2x\cos(x) - 2\cos(x)$

① $x^2\cos x + 3\cos(x)$

② $-x^2\sin(x) + 2x\sin(x) - 5\sin(x) + 2x\sin(x) - 2\sin(x)$

$-x^2\sin(x) + 4x\sin(x) - 7\sin(x)$

$(x^2+3)\cos(x) - (x^2-4x+7)\sin(x)$

Using the product rule, differentiate with respect to x the function

$$f(x) = e^x(3x-2)\sin(x)$$

Hint: Consider $f(x)g(x)h(x)$ as $[f(x)g(x)]h(x)$

$$f'(x) = \cancel{(e^x(3x-2)+\cos(x)}$$

$$e^x(3x-2)\cos x + \sin(x)\,(e^x 3 + (3x-2)e^x)$$

$$= e^x(3x-2)\cos x + \sin(x)e^x(3+3x-2)$$

$$= e^x(3x-2)\cos x + \sin(x)e^x(3x+1)$$

$$= e^x((3x-2)\cos x + (3x+1)\sin(x))$$

Chain Rule :-

$$h(P(m)) = h(m) = \frac{-1}{3}(e^m-1)^2 + (e^m-1) + \frac{1}{5}$$

calculate $\frac{dh}{dm}$: $\frac{-1}{3}(e^{2m}+1-2e^m) + e^m - 1 + \frac{1}{5}$

$$= \frac{-1}{3}e^{2m} - \frac{1}{3} + \frac{2}{3}e^m + e^m - 1 + \frac{1}{5}$$

$$= -\frac{2m}{3}e^{2m-1} + \frac{2m}{3}e^{m-1} + me^{m-1}$$

$$= -\frac{2m}{3}e^{2m-1} + \frac{5me^{m-1}}{3}$$

$$\frac{1}{3}(5me^{m-1} - 2me^{2m-1})$$

$$\frac{1}{3}e^m(5-2e^m)$$

$y^{-1/2}$ $\quad$ $y^{+1/2}$

e^{m-1} $\quad$ $e^{m+1/2}$

me^{m-1}

$4x^3 = 4$

$4e^{4x+1/2}$

Remember $\frac{d}{dm}e^a = e^a$

$$-\frac{1}{3}(e^m-1)^2+(e^m-1)+\frac{1}{5}$$

$$= -\frac{1}{3}(e^m-1)(e^m-1)+(e^m-1)+\frac{1}{5}$$

$$= \frac{-1}{3}\left[(e^m-1)(me^{m-1})+(e^m-1)(me^{m-1})\right]+me^{m-1}$$

Chain Rule :

$h(p(m))$

$$\frac{dh}{dp}\times\frac{dp}{dm}=\frac{dh}{dm}$$

$h(p) = -\frac{1}{3}p^2+p+\frac{1}{5}$ $\qquad$ $p(m) = e^m-1$

$\frac{dh}{dp} = 1-\frac{2}{3}p$ $\qquad$ $\frac{dp}{dm} = e^m$

$$\frac{dh}{dp}\times\frac{dp}{dm} = \left(1-\frac{2}{3}p\right)e^m$$

If $f(x) = g(h(x))$ what is the differential of f with respect to x in terms of g and h

$$\frac{dg}{dh}\times\frac{dh}{dx} \qquad g'(h'(x))$$

$f(x) = \sqrt{x^3-2x} = (x^3-2x)^{1/2}$

$\frac{1}{2}(3x^2-2)^{-\frac{1}{2}}$

$= \frac{1}{2\sqrt{3x^2-2}}$

$\cancel{\frac{1}{2\sqrt{x^3-2x}}}$

$\cancel{(x^3-2x)^2}$

$\cancel{2(3x^2-2)}$

$(x^3-2x)(3x^2-2)$

$+(x^3-2x)(3x^2-2)$

$(x^3-2x)(3x^2-2)$

2

Differentiate wrt x the function $f(x) = e^{x^3-3}$

$$e^{x^3-3} = \frac{e^{x^3}}{e^3}$$

$$\cancel{= 3}$$

$$e^{x^3-3}$$

$$\underline{(x^3-3)e^{x^3-3}} \rightarrow \frac{d}{dx}e^x = e^x$$

$= 64$

2^{2^3}

2^6

$2\times2\times2\times2\times2\times2$

$$f(x) = \sqrt{e^{x+2}}$$

$$(e^{x+2})^{1/2}$$

$$\cancel{\frac{1}{2\sqrt{e^{x+2}}}} \quad \frac{1}{2}\sqrt{e^{x+2}}$$

e^{3x}

$3xe^{3x-1}$ $\quad e^{-3}$

$3xe^{2x-4}$

$3xe^{3x-3}$

$f(t) = f(y(x(t)))$, what is the differential of f wrt t in terms of f, y, z and t?

$$\frac{df}{dy} \times \frac{dy}{dx} \times \frac{dx}{dt}$$

$\cancel{f'(y'(x'}$ $\quad f'(g$

$\cancel{\frac{3}{2}}$

$\frac{3}{2} - 1$

What is the derivative of the function

$$f(x) = x^{3/2} + \pi x^2 + \sqrt{7}$$

$$f'(x) = 3/2\, x^{1/2} + 2\pi x$$

$$x = 2 \Rightarrow \frac{3}{2}\sqrt{2} + 4\pi$$

$f(x) = x^3 \cos(x) e^x$

$\Rightarrow [x^3 \cos(x)] e^x$

$= x^3 \cos(x) e^x + e^x (3x^2 \cos(x) - x^3 \sin(x))$

$= x^3 \cos(x) e^x + 3x^2 e^x \cos(x) - x^3 \sin(x) e^x$

$= -x^3 \sin(x) e^x + x^3 \cos(x) e^x +$

$f(x) = e^{[(x+1)^2]}$

$\frac{de}{du} = e^u$

$u = (x+1)^2 = (x+1)(x+1)$

$\frac{du}{dv} = x+1 + x+1$

$= 2(x+1)$

$2(x+1) e^{[(x+1)^2]}$

$f(x) = x^2 \cos(x^3)$

$u = x^3$

$\frac{du}{dx} = 3x^2$

$t = \cos u$

$\frac{dt}{du} = -\sin(u)$

$-\sin x^3 . 3x^2$

$x^2 \times (-3x^2 \sin x^3) + \cos(x^3) 2x$

$f(x) = x^2 \cos(x^3)$

~~$u = \cos(x^3)$~~

~~$\frac{du}{dx} =$~~

$t = \cos(u)$

$\frac{dt}{du} = -\sin(u)$

$u = x^3$

$\frac{du}{dx} = 3x^2$

$-\sin(x^3) 3x^2$

$\frac{dt}{dx} \times \frac{du}{dx} = -\sin(u) . 3x^2$

$\frac{dt}{dx} = -\sin(x^3) . 3x^2$

$-3x^2 \sin(x^3)$

$u = x^2 \quad \frac{du}{dx} = 2x$

$x^2 x - 3x^2 \sin(x)^3 + \cos(x^3) 2x$

$= -3x^4 \sin(x^3) + 2x \cos(x^3)$

What is the derivative of the function

$f(x) = \sin(x)e^{\cos(x)}$ at the point $x = \pi$

$\sin(x)e^{\cos(x)}$

$\sin x - \sin x e^{\cos(x)}$

$+ e^{\cos(x)} \cos(x)$

$t = e^u$

$\frac{dt}{du} = e^u$

$u = \cos(x)$

$\frac{du}{dx} = -\sin(x)$

$\frac{dt}{du} \times \frac{du}{dx} = e^u \times -\sin(x)$

$= -e^{\cos(x)} \sin(x)$

$u = \sin(x)$

$\frac{du}{dx} = \cos(x)$

$\Rightarrow -\sin(x) e^{\cos(x)} \sin(x) + e \cos(x) \cdot \cos(x)$

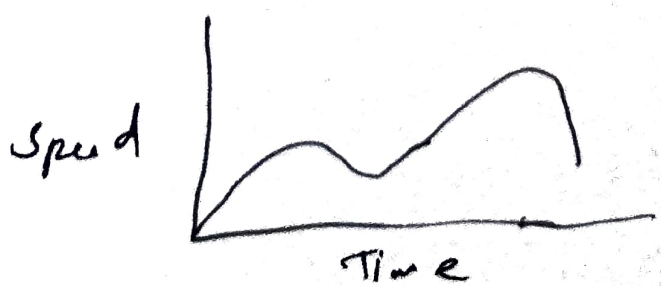
$= -e^{\cos(x)} \sin^2(x) + e^{\cos(x)} \cos(x)$

$= -e^{\cos(\pi)} \underbrace{\sin^2(\pi)}_{0} + e^{\cos(\pi)} \underbrace{\cos(\pi)}_{-1}$

What is a variable $y = f(x)$ $x = g(y)$

eg A vehicle speed is clearly a function of time.

At each time the vehicle can have one speed but at each speed there could be many time values. Hence speed is a dependent variable



~~Decision Trees~~

Machine learning classifier → study for sure

Typically we take functions with variables and constants, differentiate dependent variable wrt independent variables

Differentiate wrt anything

$f(x, y, z) = \sin(x) e^{yz^2}$

$\frac{\partial f}{\partial x} = \cos(x) e^{yz^2}$

$\frac{\partial f}{\partial y} = \sin x \, e^{yz^2} z^2$

$\frac{\partial f}{\partial z} = \sin x \, e^{yz^2} 2zy$

Chain rule

$h(x) = f[g(x)]$

$h'(x) = f'[g(x)] \cdot g'(x)$

Egs

$y = \sin(x^2)$

$y' = \cos(x^2) \cdot 2x$

$\frac{dz}{dt} = t^2 \quad z = e^{yz^2}$

$\frac{\partial z}{\partial y} = e^{yz^2} \cdot z^2$

Total derivative $dx/dt = 1$ $dy/dt = 2t$

Assume $x = t-1$ $y = t^2$ $z = \frac{1}{t}$

$df(x,y,z) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}$

$\Rightarrow \frac{df(x,y,z)}{dt} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial t} + \frac{\partial f}{\partial z} \cdot \frac{\partial z}{\partial t}$

$\Rightarrow \cos(x) e^{yz^2} \cdot 1 + \sin(x) e^{yz^2} z^2 + \sin(x) e^{yz^2} 2zy(-t^{-2})$

$\frac{df}{dt} = \frac{df}{dx} \cdot \frac{dx}{dt}$

f is a function of x. x is a function of t

$\frac{d}{dx} e^{f(x)} = e^{f(x)} \cdot f'(x)$

If e is raised to some complicated function then the derivative is the same

Practicing Partial differentiation

W

1) Given $f(x,y) = \pi x^3 + xy^2 + my^4$, with m a constant

What are $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$

$$\frac{\partial f}{\partial x} = 3x^2\pi + y^2$$

$$\frac{\partial f}{\partial y} = 2xy + 4y^3 m$$

2) $f(x,y,z) = x^2y + y^2z + z^2x$. What are $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

$$\frac{\partial f}{\partial x} = 2xy + z^2$$

$$\frac{\partial f}{\partial y} = x^2 + 2yz$$

$$\frac{\partial f}{\partial z} = y^2 + 2zx$$

5

3) Given $f(x,y,z) = e^{2x}\sin(y)z^2 + \cos(z)e^x e^y$ What are $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

$$\frac{\partial f}{\partial x} = 2e^{2x}\sin(y)z^2 + \cos(z)e^x e^y$$

$$\frac{\partial f}{\partial y} = e^{2x}\cos(y)z^2 + \cos(z)e^x e^y$$

$$\frac{\partial f}{\partial z} = e^{2x}\sin(y)2z - \sin(z)e^x e^y$$

Wh

Ap 4) Recall the formula for the total derivative that is for $f(x,y)$, $x = x(t)$ and $y = y(t)$. one can calculate

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

Va Given that $f(x,y) = \pi x^2 y$, $x(t) = t^2+1$ and $y(t) = t^2-1$ calculate the total derivative $\frac{df}{dt}$

$f(x,y) = \pi x^2 y$

$\frac{\partial z}{\partial x} = 2\pi x y$ $\qquad x(t) = t^2+1 \qquad y(t) = t^2-1$

$\frac{\partial z}{\partial y} = \pi x^2$ $\qquad \frac{dx}{dt} = 2t \qquad \frac{dy}{dt} = 2t$

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt}$$

$$\Rightarrow 2\pi x y \cdot 2t + \pi x^2 \cdot 2t$$

$$= 2\pi(t^2+1) \cdot 2t(t^2-1) + \pi(t^2+1)^2 \cdot 2t$$

$$= 4\pi t(t^2+1)(t^2-1) + 2\pi t(t^2+1)^2$$

5) Recall the formula for the total derivative, that is, for $f(x,y,z)$, $x = x(t)$, $y = y(t)$ and $z = z(t)$.

$$\frac{df}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} + \frac{\partial z}{\partial z}\frac{dz}{dt}$$

Given that $f(x,y,z) = \cos(x)\sin(y)e^{2z}$

$x(t) = t+1$, $y(t) = t-1$, $z(t) = t^2$. Calculate the total derivative df/dt

$f(x,y,z) = \cos(x)\sin(y)e^{2z}$ $\qquad \frac{dx}{dt} = 1$

$\frac{\partial z}{\partial x} = -\sin(x)\sin(y)e^{2z}$ $\qquad \frac{dy}{dt} = 1$

$\frac{\partial z}{\partial y} = \cos(x)\cos(y)e^{2z}$ $\qquad \frac{dz}{dt} = 2t$

$\frac{\partial z}{\partial z} = \cos(x)\sin(y)\,2e^{2z}$

$$-\sin(x)\sin(y)e^{2z} \times 1 + \cos(x)\cos(y)e^{2z} \times 1 + \cos(x)\sin(y)\,2e^{2z} \cdot 2t$$

$$-\sin(t+1)\sin(t-1)e^{2t^2} + \cos(t+1)\cos(t-1)e^{2t^2}$$

$$+ \cos(t+1)\sin(t-1)4e^{2t^2}t$$

The Jacobian :

Previously we saw that we can differentiate functions of multiple variables. The Jacobian brings in some of the ideas of linear algebra to partial differentiation and make this useful.

$$f(x, y, z) = x^2 y + 3z$$

$$\frac{\partial f}{\partial x} = 2xy \quad \frac{\partial f}{\partial y} = x^2 \quad \frac{\partial f}{\partial z} = 3$$

$$J = [2xy, x^2, 3]$$

Jacobian is a vector in an algebraic form that gives the maximum/steepest slope of the function for any given x, y, z.

Remember a derivative or differential value of a function is essentially the slope of the function at any given point.

Rephrasing :- $J = [2x, x^2, 3]$

i) Vector in an algebraic form/expression

ii) When given specific $[x, y, z]$ we get a vector pointing in the direction of steepest slope of the given function

The Jacobian is a vector that we can calculate for each location of a plot which points in the direction of the steepest uphill slope.

The steeper the slope the greater the magnitude of Jacobian at that point.

Calculating the Jacobian:

1) for $f(x,y) = x^2y + \frac{3}{4}xy + 10$, calculate the Jacobian row vector J.

$$\frac{\partial z}{\partial x} = 2xy + \frac{3}{4}y$$

$$\frac{\partial z}{\partial y} = x^2 + \frac{3}{4}x$$

$$J = \left[2xy + \frac{3}{4}y,\ x^2 + \frac{3}{4}x\right]$$

2) $f(x,y) = e^x\cos(y) + xe^{3y} - 2$. Calculate the Jacobian row vector J

$$\frac{\partial z}{\partial x} = e^x\cos(y) + e^{3y}$$

$$\frac{\partial z}{\partial y} = -e^x\sin(y) + 3xe^{3y}$$

3) $f(x,y,z) = e^x\cos(y) + x^2y^2z^2$, calculate the Jacobian row vector J

$$\frac{\partial z}{\partial x} = e^x\cos(y) + 2xy^2z^2$$

$$\frac{\partial z}{\partial y} = -e^x\sin(y) + 2x^2yz^2$$

$$\frac{\partial z}{\partial z} = 2x^2y^2z$$

4) $f(x,y,z) = x^2 + 3ye^z + \cos(x)\sin(z)$. Calculate the Jacobian row vector and evaluate at the point $(0,0,0)$

$$\frac{\partial f}{\partial x} = 2x - \sin(x)\sin(z)$$

$$J = [0, 3, 4]$$

$$\frac{\partial f}{\partial y} = 3e^y e^z$$

$$\frac{\partial f}{\partial z} = 3e^y e^z + \cos(x)\cos(z)$$

5) $f(x,y,z) = xe^y\cos(z) + 5x^2\sin(y)e^z$. Calculat the Jacobian row vector at the point $(0,0,0)$

$$\frac{\partial f}{\partial x} = e^y\cos(z) + 10x\sin(y)e^z$$

$$J = [1, 0, 0]$$

$$\frac{\partial f}{\partial y} = xe^y\cos(z) + 5x^2\cos(y)e^z$$

$$\frac{\partial f}{\partial z} = -xe^y\sin(z) + 5x^2\sin(y)e^z$$

Jacobian applied :

Vector Valued function :

A vector-valued function, also referred as a vector function, is a mathematical function of one or more variables whose range is a set of multi-dimensional vectors or infinite-dimensional vectors. The input of a vector valued function could be a scalar or vector (that is, the dimension of I could be 1 or greater than 1). The dimension of the domain is not defined by the dimension of the range.

X. By a theorem, if f is a differentiable function of x + y, then f has a directional derivative in the direction of any vector $u = \langle a, b \rangle$

$x = f(t)$ $y = g(t)$, $z = h(t)$

But vectors are defined by $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$

So we don't know x, y, z we can substitute the parametric functions $\vec{r}(t) = f(t)\hat{i} + g(t)\hat{j} + h(t)\hat{k}$

What variable the vector function is going to be based on? 't'.

$$\vec{r}(t) = \underset{x}{f(t)}\hat{i} + \underset{y}{g(t)}\hat{j} + \underset{z}{h(t)}\hat{k} \longrightarrow \langle f(t), g(t), h(t) \rangle$$

Every single value of t will give a vector

Parametric Equation:

$x = f(t)$ $y = g(t)$ 't' is the parameter for some interval 'I' on a common domain.

So $f(x,y) = e^{-(x^2+y^2)}$

$$\frac{\partial z}{\partial x} = e^{-(x^2+y^2)} \cdot -2x$$

$$\frac{\partial z}{\partial y} = e^{-(x^2+y^2)} \times -2y$$

$$J = \left[-2xe^{-(x^2+y^2)}, -2ye^{-(x^2+y^2)} \right]$$

We are going to look at the Jacobian and try to understand how the function looks.

~~By a thus~~

Directional Derivative:

Theorem: If f is a differentiable function of x and y, then f has a directional derivative in the direction of any unit vector $\bar{u} = \langle a, b \rangle$ and

$$D_u f(x,y) = f_x(x,y)a + f_y(x,y)b$$

Ex 1: Find the directional derivative of $f(x,y) = x^2y^3 - y^4$ at $(2,1)$ in the direction given by the angle $\theta = \pi/4$

1) Find a unit vector in the direction $\theta = \pi/4$

$$\Rightarrow \bar{u} = \langle \cos \pi/4, \sin \pi/4 \rangle$$
$$= \left\langle \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right\rangle$$

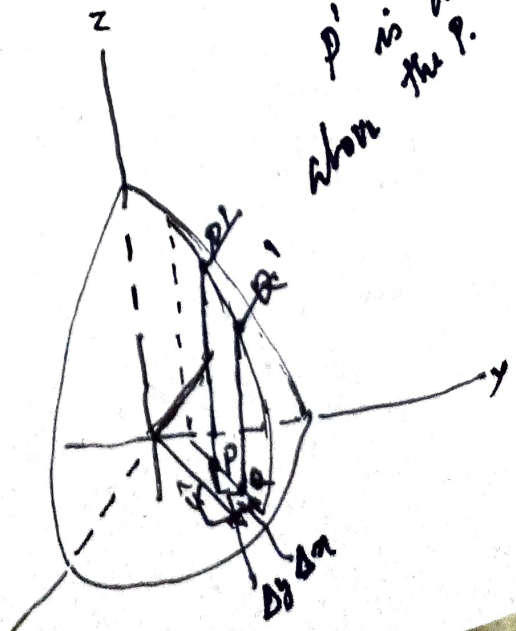
$f(x,y) = x^2y^3 - y^4$

$$\frac{\partial z}{\partial x} = 2xy^3 \qquad \frac{\partial z}{\partial y} = 3x^2y^2 - 4y^3$$
$$= -4y^3 + 3x^2y^2$$

$$D_u f(x,y) = (2xy^3)a + (3x^2y - 4y^3)b$$
$$D_u f(2,1) = (2 \cdot 2 \cdot 1^3)a + (3 \cdot 2^2 \cdot 1 - 4 \cdot 1^3)b$$
$$= 4a + 8b$$
$$= 4a\frac{\sqrt{2}}{2} + 8 \cdot \frac{\sqrt{2}}{2}$$
$$= 2\sqrt{2} + 4\sqrt{2} = 6\sqrt{2}$$ // is the value of the directional derivative

Prop Leonard: (direction derivative)

To give a direction we must have a vector. Lets call it a unit vector so that the slope isn't impacted.



P' is directly above the P.

Lets take a unit vector in the plane of the independent variables.

$\frac{\partial f}{\partial x}$ & $\frac{\partial z}{\partial y}$ are directional directions

Things we need:

We need a unit vector

We need a point for a specific rate of change

The point P is going to be in x, y plane

Make line through 'P' parallel to $\vec{u}$.

Create a plane parallel to the $\vec{u}$ & z plane that makes a curve in the surface in the plane.

P'Q' is a secant line in the plane. If Q is allowed to go to P, the secant approaches a tangent near the P'.

As $Q \to P$, the SECANT P'Q' approaches the tangent @ P'.

So both Δx and Δy must $\to 0$

Now $\vec{PQ} \parallel \hat{u}$ $\quad \vec{PQ} = \Delta x\hat{i} + \Delta y\hat{j}$

If this is true

$\vec{PQ} = h \cdot \hat{u} \Rightarrow h(u_1\hat{i} + u_2\hat{j}) = \Delta x\hat{i} + \Delta y\hat{j}$

$\Rightarrow \Delta x = hu_1 \quad \Delta y = hu_2$

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \cancel{f(x+\Delta x, y+\Delta y) - f(x,y)}$$

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{f(x+\Delta x, y+\Delta y) - f(x,y)}{\sqrt{(\Delta x^2 + \Delta y^2)}}$$

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{f(x+hu_1, y+hu_2) - f(x,y)}{h}$$

$\sqrt{h^2u_1^2 + h_2u_2^2}$

$\sqrt{h^2(u_1^2 + u_2^2)}$

$h\sqrt{u_1^2 + u_2^2}$

$h \cdot 1$

Jacobian applied:

We are going to extend the concept of Jacobian from vectors upto the matrices, which will allow you to describe the rates of change of vector valued functions.

$f(x,y) = e^{-(x^2+y^2)}$

$\frac{\partial z}{\partial x} = -2xe^{-(x^2+y^2)}$, $\frac{dz}{\partial y} = -2ye^{-(x^2+y^2)}$

$J = \left[-2xe^{-(x^2+y^2)}, 2ye^{-(x^2+y^2)}\right]$

$J(1,1) = [2e^{-2}, -2e^{-2}] = [-0.27, 0.27]$

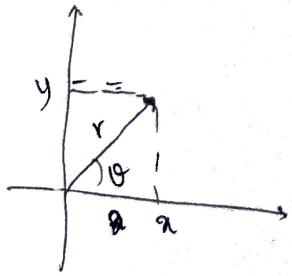
$J = \begin{pmatrix} \partial u/\partial x & \frac{\partial u}{\partial y} \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix}$ is the Jacobian Matrix

Jacobian matrix does ~~not~~ NOT contain any variables. This means that the functions u, v are linear functions $ax+y$ and the gradients are constant

$u(x,y) = x - 2y \quad v(x,y) = 3y - 2x$

$\Rightarrow \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}$

This matrix is the linear transformation from x,y space to u,v space



$x(r,\theta) = r\cos\theta$
$y(r,\theta) = r\sin\theta$

$J = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix}$

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

$|J| = r(\cos^2\theta + \sin^2\theta)$
$= r$

Bigger Jacobians

For the function $u(x,y) = x^2 - y^2$ and $v(x,y) = 2xy$ calculate the Jacobian matrix $J = \begin{bmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{bmatrix}$

$\frac{\partial u}{\partial x} = 2x \quad \frac{\partial u}{\partial y} = -2y$

$\frac{\partial v}{\partial x} = 2y \quad \frac{\partial v}{\partial y} = 2x$

$J = \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix}$

2) for the function $u(x,y,z) = 2x + 3y$, $v(x,y,z) = \cos(x)\sin(z)$ and $w(x,y,z) = e^x e^y e^z$

$\frac{\partial u}{\partial x} = 2 \quad \frac{\partial u}{\partial y} = 3 \quad \frac{\partial u}{\partial z} = 0$

$\frac{\partial v}{\partial x} = -\sin(x)\sin(z) \quad \frac{\partial v}{\partial y} = 0 \quad \frac{\partial v}{\partial z} = \cos(x)\cos(z)$

$\frac{\partial w}{\partial x} = e^x e^y e^z \quad \frac{\partial w}{\partial y} = e^x e^y e^z \quad \frac{\partial w}{\partial z}$

3) Consider the pair of linear equations $u(x,y) = ax+by$, $v(x,y) = cx+dy$ where a, b, c and d are all constants. Calculate the Jacobian and notice something kind of interesting

$$\frac{\partial u}{\partial x} = a \qquad \frac{\partial u}{\partial y} = b$$

$$\frac{\partial v}{\partial x} = c \qquad \frac{\partial v}{\partial y} = b$$

4) For the function $u(x,y,z) = 9x^2y^2 + 2e^x$, $v(x,y,z) = xy + x^2y^3 + 2z$ and $w(x,y,z) = \cos(x)\sin(z)e^y$. Calculate the Jacobian matrix and evaluate at the point $(0,0,0)$

$$\frac{\partial u}{\partial x} = 18xy^2 + 2e^x \qquad \frac{\partial u}{\partial y} = 18x^2y \qquad \frac{\partial u}{\partial z} = 0$$

$$\frac{\partial v}{\partial x} = y + 2xy^3 \qquad \frac{\partial v}{\partial y} = x + 3x^2y^2 \qquad \frac{\partial v}{\partial z} = 2$$

$$\frac{\partial w}{\partial x} = -\sin(x)\sin(z)e^y \qquad \frac{\partial w}{\partial y} = \cos(x)\sin(z)e^y \qquad \frac{\partial w}{\partial z} = \cos(x)\cos(z)e^y$$

$$J = \begin{bmatrix} 18xy^2+2e^x & 18x^2y & e^x \\ y+2xy^3 & x+3x^2y^2 & 2 \\ -\sin(x)\sin(z)e^y & \cos(x)\sin(z)e^y & \cos(x)\cos(z)e^y \end{bmatrix}$$

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

5) In the lecture we calculated the Jacobian of the transformation from Polar co-ordinates to Cartesian co-ordinates in 2D. In this question we will do the same but with spherical co-ordinates in 3D.

For the functions
$$x(r,\theta,\phi) = r\cos(\theta)\sin(\phi)$$
$$y(r,\theta,\phi) = r\sin(\theta)\sin(\phi)$$
$$z(r,\theta,\phi) = r\cos(\phi)$$

$$\frac{\partial x}{\partial r} = \cos(\theta)\sin(\phi) \qquad \frac{\partial x}{\partial \theta} = r(-\sin(\theta))\sin(\phi) \qquad \frac{\partial x}{\partial \phi} = r\cos(\theta)\cos(\phi)$$

$$\frac{\partial y}{\partial r} = \sin(\theta)\sin(\phi) \qquad \frac{\partial y}{\partial \theta} = r\cos(\theta)\sin(\phi) \qquad \frac{\partial y}{\partial \phi} = r\sin(\theta)\cos(\phi)$$

$$\frac{\partial z}{\partial r} = \cos(\phi) \qquad \frac{\partial z}{\partial \theta} = 0 \qquad \frac{\partial z}{\partial \phi} = r(-\sin(\phi))$$

Sandpit:

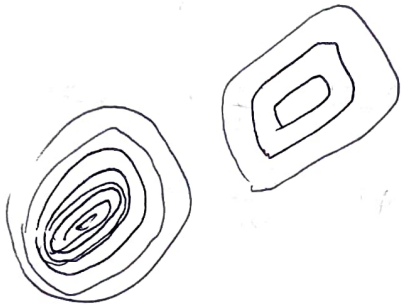
Jacobian:

The supervisor of a building site has dropped their mobile phone into a pit with an uneven floor, and it has rolled to the lowest point. To make matters worse the pit has been subsequently been filled with sand such that the phone is covered and cannot be seen. To find where in the pit their phone is, the supervisor has crafted a dipstick with a head designed to measure the slope of the floor if it is poked straight down through the sand.

By clicking on any point in the sandpit and thereby measuring the negative of the Jacobian $-J = -\nabla f(x)$ at that point, try and find the supervisor's phone. Try to do this with a few dips as possible

: Before trying the dip. [illegible]

Jacobian, the supervisor had constructed a dip-stick that can only measure the depth on each dip. Without information about the Jacobian try and find the supervisor's pole



The Hessian:

$f(x,y,z) = x^2yz$

$\frac{\partial f}{\partial x} = 2xyz \quad \frac{\partial f}{\partial y} = x^2z \quad \frac{\partial f}{\partial z} = x^2y$

$J = [2xyz, x^2z, x^2y]$

$$H = \begin{bmatrix} 2yz & 2xz & 2xy \\ 2xz & 0 & x^2 \\ 2xy & x^2 & 0 \end{bmatrix}$$

top right

bottom left

The Hessian matrix is symmetrical across the leading diagonal

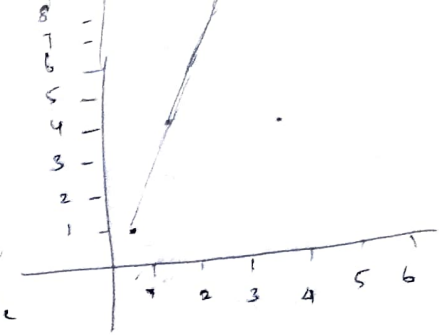
Once you get the top right, you get the bottom left. This will always be true if the function is continuous

$f(x,y) = x^2+y^2$

$J = [2x, 2y]$

$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

$Det(H) \rightarrow |H| = 4$



So the determinant is positive we know we are dealing with either a maximum or minimum

$J = 2x, 2y$

$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ So the first term is also positive its a minimum

So it is negative we have a maximum

$f(x,y) = x^2 - y^2$

$J = [2x, -2y]$

$H = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \Rightarrow |H| = -4$

So the Hessian is a negative we are not dealing with a minimum or maximum.

For a multivariable function the first order derivatives ~~of x~~ with a vector as Jacobian.

Second order derivatives with a matrix is Hessian.

A Hessian matrix will be a $N \times N$ square matrix, where N is also the number of the variables

Calculating Hessians:

In this quiz, you will calculate the Hessian for some functions of 2 variables and functions of 3 variables

1) For the function $f(x,y) = x^3y + x + 2y$. Calculate the Hessian matrix $H = \begin{bmatrix} \partial_{x,x} f & \partial_{x,y} f \\ \partial_{y,x} f & \partial_{y,y} f \end{bmatrix}$

$f(x,y) = x^3y + x + 2y$

$\frac{\partial f}{\partial x} = 3x^2y + 1 \qquad \frac{\partial f}{\partial y} = x^3 + 2$

$J = \begin{bmatrix} 3x^2y+1 & x^3+2 \end{bmatrix}$

$H = \begin{bmatrix} 6xy & 3x^2 \\ 3x^2 & 0 \end{bmatrix}$

2) for the function $f(x,y) = e^x \cos(y)$. Calculate the Hessian matrix

$\frac{\partial f}{\partial x} = e^x \cos(y) \qquad \frac{\partial f}{\partial y} = -e^x \sin(y)$

$J = \begin{bmatrix} e^x \cos(y) & -e^x \sin(y) \end{bmatrix}$

$H = \begin{bmatrix} e^x \cos(y) & -e^x \sin(y) \\ -e^x \sin(y) & -e^x \cos(y) \end{bmatrix}$

3) for the function $f(x,y) = \frac{x^2}{2} + xy + \frac{y^2}{2}$. Calculate the Hessian matrix. Notice something when you calculate $\frac{1}{2}[x,y] H \begin{bmatrix} x \\ y \end{bmatrix}$!

$\frac{\partial f}{\partial x} = x + y \qquad \frac{\partial f}{\partial y} = x + y$

$J = \begin{bmatrix} x+y & x+y \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

$\frac{1}{2}[x,y] = \frac{x^2}{4} + \frac{xy}{2} + \frac{y^2}{4} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

$\frac{x^2}{4} + \frac{xy}{4} + \frac{xy}{4} + \frac{y^2}{4}$

$\frac{x^2}{4} + \frac{xy}{2} + \frac{y^2}{4}$

$\frac{1}{2}[x,y] \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \frac{1}{2}[x+y] \begin{bmatrix} x+y \\ x+y \end{bmatrix}$

4) for the function $f(x,y,z) = x^2e^{-y}\cos(z)$. Calculate the Hessian matrix.

$H = \begin{bmatrix} \partial_{x,x} f & \partial_{x,y} f & \partial_{x,z} f \\ \partial_{y,x} f & \partial_{y,y} f & \partial_{y,z} f \\ \partial_{z,x} f & \partial_{z,y} f & \partial_{z,z} f \end{bmatrix}$

$f(x,y,z) = x^2e^{-y}\cos(z)$

$\frac{\partial f}{\partial x} = 2xe^{-y}\cos(z) \qquad \frac{\partial f}{\partial y} = -x^2e^{-y}\cos(z)$

$\frac{\partial f}{\partial z} = -x^2e^{-y}\sin(z)$

$J = \begin{bmatrix} 2xe^{-y}\cos(z) & -x^2e^{-y}\cos(z) & -x^2e^{-y}\sin(z) \end{bmatrix}$

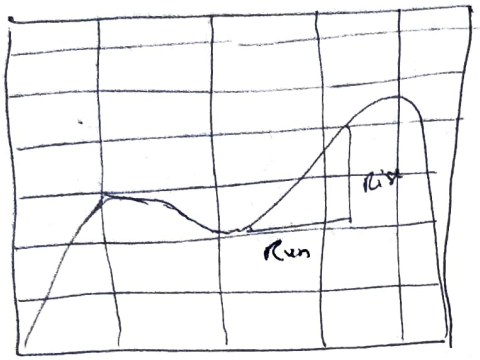
$H = \begin{bmatrix} 2e^{-y}\cos(z) & -2xe^{-y}\cos(z) & -2xe^{-y}\sin(z) \\ -2xe^{-y}\cos(z) & x^2e^{-y}\cos(z) & x^2e^{-y}\sin(z) \\ -2xe^{-y}\sin(z) & x^2e^{-y}\sin(z) & -x^2e^{-y}\cos(z) \end{bmatrix}$

5) for the function $f(x,y,z) = xe^y + y^2\cos(z)$. Calculate the Hessian matrix.

$\partial f/\partial x = e^y \qquad \frac{\partial f}{\partial y} = xe^y + 2y\cos(z) \qquad \partial f/\partial z = -y^2\sin(z)$

$J = \begin{bmatrix} e^y & xe^y + 2y\cos(z) & -y^2\sin(z) \end{bmatrix}$

$H = \begin{bmatrix} 0 & e^y & 0 \\ e^y & xe^y + 2\cos(z) & -2y\sin(z) \\ 0 & -2y\sin(z) & -y^2\cos(z) \end{bmatrix}$



$$\text{Gradient at } x \approx \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

Assessment: Jacobians and Hessians

1) Calculate the Jacobian of the function $f(x,y,z) = x^2\cos(y) + e^z\sin(y)$ and evaluate at the point $(x,y,z) = (\pi,\pi,1)$

$\frac{\partial f}{\partial x} = 2x\cos(y)$ $\quad \partial f/\partial y = -x^2\sin(y)$ $\quad \partial f/\partial z = e^z\cos(z)$

$J = [2x\cos(y) \quad -x^2\sin(y) \quad e^z\cos(z) + e^z\sin(z)]$

$= [2\pi\cos(\pi) \quad -\pi\sin(\pi) \quad e^{\pi}\cos(\pi) + e^{\pi}\sin(\pi)]$

$= [-2\pi \quad 0 \quad -e^{\pi}]$

$\cos(\pi) = -1$
$\sin(\pi) = 0$

$f(x,y,z) = x^2\cos(y) + e^z\sin(y)$

$\partial f/\partial x = 2x\cos(y)$ $\quad \partial f/\partial y = -x^2\sin(y) + e^z\cos(y)$

$\partial f/\partial z = e^z\sin(y)$

$J = [2x\cos(y) \quad -x^2\sin(y) + e^z\cos(y) \quad e^z\sin(y)]$

$= -2\pi \quad -e \quad 0$

2) Calculate the Jacobian of the vector value functions $u(x,y) = x^2y - \cos(x)\sin(y)$ $\quad v(x,y) = e^{x+y}$, evaluate at the point $(0,\pi)$

$\partial u/\partial x = 2xy + \sin(x)\sin(y)$ $\quad \partial u/\partial y = x^2 - \cos(x)\cos(y)$

$\partial v/\partial x = e^{x+y}$ $\quad \partial v/\partial y = e^{x+y}$

$0, \pi$ $\quad 0 \quad 1$

$e^{\pi} \quad e^{\pi}$

3) $f(x,y) = x^3\cos(y) - x\sin(y)$. Calculate the Hessian for the function

$J = \partial f/\partial x = 3x^2\cos(y) - \sin(y)$ $\quad \partial f/\partial y = -x^3\sin(y) - x\cos(y)$

$H = \begin{matrix} 6x\cos(y) & -3x^2\sin(y) - \cos(y) \\ -3x^2\sin(y) - \cos(y) & -x^3\cos(y) + x\sin(y) \end{matrix}$

4) $f(x,y,z) = xy + \sin(y)\sin(z) + z^3e^x$. Calculate the Hessian for the function

J $\quad \partial f/\partial x = y + z^3e^x$ $\quad \partial f/\partial y = x + \cos(y)\sin(z)$

$\partial f/\partial z = \sin(y)\cos(z) + 3z^2e^x$

H $\quad \begin{matrix} z^3e^x & 1 & 3z^2e^x \\ 1 & -\sin(y)\sin(z) & \cos(y)\cos(z) \\ 3z^2e^x & \cos(y)\cos(z) & -\cos(y)\sin(z) + 6ze^x \end{matrix}$

5) Calculate the Hessian for the function

$f(x,y,z) = xy\cos(z) - \sin(x)e^{y}z^3$ and evaluate at the point $(x,y,z) = (0,0,0)$

$\partial f/\partial x = y\cos(z) - \cos(x)e^{y}z^3$

$\partial f/\partial y = x\cos(z) - \sin(x)e^{y}z^3$

$\partial f/\partial z = -xy\sin(z) - 3\sin(x)e^{y}z^2$

$$\begin{matrix} \sin(x)e^{y}z^3 & \cos(z) - \cos(x)e^{y}z^3 & -y\sin(z) - \cos(x)e^{y}3z^2 \\ \cos(z) - \cos(x)e^{y}z^3 & -\sin(x)e^{y}z^3 & -x\sin(z) - \sin(x)e^{y}3z^2 \\ -y\sin(z) - 3\cos(x)e^{y}z^2 & -x\sin(z) - 3\sin(x)e^{y}z^2 & -xy\cos(z) - 3\sin(x)e^{y}2z \end{matrix}$$

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

Multivariate chain rule:

$f(x,y,z) = \sin(x)e^{yz^2}$

Total derivative:

When we have a multivariable function such as $f(x,y,z)$ and the variables x,y,z are themselves a function of another variable t

$x = t-1 \quad y = t^2 \quad z = \frac{1}{t}$

Then if we want to calculate the derivative of 'f' with respect to 't'

$$\frac{\partial f}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial f}{\partial y}\cdot\frac{dy}{dt} + \frac{\partial f}{\partial z}\cdot\frac{dz}{dt}$$

$\frac{df}{dt} = \cos(t-1)e$

$f(x_1, x_2, x_3, \dots, x_n) = f(x)$

$f(x) = f(x_1, x_2, x_3, \dots, x_n)$

$$\partial f/\partial x = \begin{bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \partial f/\partial x_3 \\ \vdots \\ \partial f/\partial x_n \end{bmatrix} \quad \frac{dx}{dt} = \begin{bmatrix} \frac{dx_1}{dt} \\ dx_2/dt \\ dx_3/dt \\ \vdots \\ dx_n/dt \end{bmatrix}$$

$$df/dt = \begin{bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \partial f/\partial x_3 \\ \vdots \\ \partial f/\partial x_n \end{bmatrix} \cdot \begin{bmatrix} dx_1/dt \\ dx_2/dt \\ dx_3/dt \\ \vdots \\ dx_n/dt \end{bmatrix}$$

$$df/dt = \frac{\partial f}{\partial x}\cdot\frac{dx}{dt}$$

Multivariate chain rule

5 $f(x(t))$

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt}$$

$f(x) = 5x$

$x(u) = 1-u$

$u(t) = t^2$

$$\frac{du}{dt} = 2t$$

$$\frac{dx}{du} = -1$$

$$\frac{df}{dx} = 5$$

$$\frac{df}{dx} \cdot \frac{dx}{du} \cdot \frac{du}{dt} = 5 \cdot -1 \cdot 2t$$

$$= -10t$$

What is $f(t)$

$x(t) = 1-t^2$

$f(x) = 5(1-t^2) = 5-5t^2$

$\frac{df}{dt} = -10t$

i) In this quiz, you will practice calculating the multivariate chain rule for various functions

For the following function, calculate the expression

$\frac{df}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt}$ in matrix form where

$x = (x_1, x_2)$. $f(x) = f(x_1, x_2) = x_1^2 x_2^2 + x_1 x_2$

$x_1(t) = 1-t^2$

$x_2(t) = 1+t^2$

$f(x) = f(x_1, x_2) = x_1^2 x_2^2 + x_1 x_2$

$\partial f/\partial x_1 = 2x_1 x_2^2 + x_2$ $\quad \partial f/\partial x_2 = 2x_1^2 x_2 + x_1$

$x_1(t) = 1-t^2$ $\quad x_2(t) = 1+t^2$

$\frac{dx_1}{dt} = -2t$ $\quad \frac{dx_2}{dt} = 2t$

$$\begin{bmatrix} 2x_1x_2^2 + x_2, & 2x_1^2x_2 + x_1 \end{bmatrix} \begin{bmatrix} -2t \\ 2t \end{bmatrix}$$

$\downarrow$

This is Jacobian

For the following function calculate the expression $df/dt = \frac{\partial f}{\partial x} \frac{dx}{dt}$ in matrix form, where $x = (x_1, x_2, x_3)$

$f(x) = f(x_1, x_2, x_3) = x_1^3 \cos(x_2) e^{x_3}$

$x_1(t) = 2t$

$x_2(t) = 1-t^2$

$x_3(t) = e^t$

$\partial f/\partial x_1 = 3x_1^2 \cos(x_2) e^{x_3}$ $\quad \partial f/\partial x_2 = x_1^3 - \sin(x_2) e^{x_3} = -x_1^3 \sin(x_2) e^{x_3}$

$\partial f/\partial x_3 = x_1^3 \cos(x_2) e^{x_3}$

$$J = \begin{bmatrix} 3x_1^2\cos(x_2)e^{x_3}, & -x_1^3\sin(x_2)e^{x_3}, & x_1^3\cos(x_2)e^{x_3} \end{bmatrix}$$

$dx_1/dt = 2$ $\quad dx_2/dt = -2t$ $\quad dx_3/dt = e^t$

$$\begin{bmatrix} 3x_1^2\cos(x_2)e^{x_3}, & -x_1^3\sin(x_2)e^{x_3}, & x_1^3\cos(x_2)e^{x_3} \end{bmatrix} \begin{bmatrix} 2 \\ -2t \\ e^t \end{bmatrix}$$

3) For the following functions, calculate the expression $df/dt = \partial f/\partial x \; \partial x/\partial u \; \partial u/\partial t$ in matrix form, where

$x = (x_1, x_2)$ and $u = (u_1, u_2)$

$f(x) = f(x_1, x_2) = x_1^2 - x_2^2$

$\partial f/\partial x_1 = 2x_1 \qquad \partial f/\partial x_2 = -2x_2$

$x_1(u_1, u_2) = 2u_1 + 3u_2$

$\partial x_1/\partial u_1 = 2 \qquad \partial x_1/\partial u_2 = 3$

$x_2(u_1, u_2) = 2u_1 - 3u_2$

$\partial x_2/\partial u_1 = 2 \qquad \partial x_2/\partial u_2 = -3$

$u_1(t) = \cos(t/2) \qquad u_2(t) = \sin(2t)$

$\frac{du_1}{dt} = \frac{-1}{2}\sin(t/2) \qquad \frac{du_2}{dt} = 2\cos(2t)$

$$\underset{\substack{\uparrow \\ J}}{[2x_1, -2x_2]} \begin{bmatrix} 2 & 3 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} -\frac{1}{2}\sin(t/2) \\ 2\cos(2t) \end{bmatrix}$$

$[\cos(x_1)\cos(x_2)e^{x_3}, -\sin(x_1)\sin(x_2)e^{x_3}, \sin(x_1)\cos(x_2)e^{x_3}]$

$$\begin{bmatrix} \cos(u_1) & -\sin(u_2) \\ -\sin(u_1) & -\cos(u_2) \\ e^{u_1} & e^{u_2} \end{bmatrix}$$

$$\begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix}$$

4) For the following function calculate the expression $\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u}\frac{du}{dt}$ where $x = (x_1, x_2)$ and $u = (u_1, u_2)$

$f(x) = f(x_1, x_2) = \cos(x_1)\sin(x_2)$

$\partial f/\partial x_1 = -\sin(x_1)\sin(x_2)$

$\partial f/\partial x_2 = \cos(x_1)\cos(x_2)$

$J = [-\sin(x_1)\sin(x_2), \cos(x_1)\cos(x_2)]$

$x_1(u_1, u_2) = 2u_1^2 + 3u_2^2 - u_2$

$\partial x_1/\partial u_1 = 4u_1 \qquad \partial x_1/\partial u_2 = 6u_2 - 1$

$x_2(u_1, u_2) = 2u_1 - 5u_2^3$

$\partial x_2/\partial u_1 = 2 \qquad \partial x_2/\partial u_2 = -15u_2^2$

$u_1(t) = e^{t/2} \qquad u_2(t) = e^{-2t}$

$du_1/dt = \frac{1}{2}e^{t/2} \qquad \frac{du_2}{dt} = e^{-2t} \cdot -2$

$$[-\sin(x_1)\sin(x_2), \cos(x_1)\cos(x_2)] \begin{bmatrix} 4u_1 & 6u_2 - 1 \\ 2 & -15u_2^2 \end{bmatrix} \begin{bmatrix} \frac{1}{2}e^{t/2} \\ -2e^{-2t} \end{bmatrix}$$

5) For the following function calculate the expression $df/dt = \partial f/\partial x \; \partial x/\partial u \; du/dt$ in matrix form where $x = (x_1, x_2)$ and $u = (u_1, u_2)$

$f(x) = f(x_1, x_2, x_3) = \sin(x_1)\cos(x_2)e^{x_3}$

$\partial f/\partial x_1 = \cos(x_1)\cos(x_2)e^{x_3} \qquad \partial f/\partial x_2 = \sin(x_1) - \sin(x_2)e^{x_3}$

$\partial f/\partial x_3 = \sin(x_1)\cos(x_2)e^{x_3}$

Soln

$J = [\cos(x_1)\cos(x_2)e^{x_3}, -\sin(x_1)\sin(x_2)e^{x_3}, \sin(x_1)\cos(x_2)e^{x_3}]$

$x_1(u_1, u_2) = \sin(u_1) + \cos(u_2) \qquad \partial x_1/\partial u_1 = \cos(u_1) \qquad \partial x_1/\partial u_2 = -\sin(u_2)$

$x_2(u_1, u_2) = \cos(u_1) - \sin(u_2) \qquad \partial x_2/\partial u_1 = -\sin(u_1) \qquad \partial x_2/\partial u_2 = -\cos(u_2)$

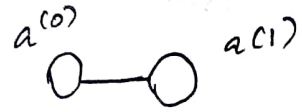
$x_3(u_1, u_2) = e^{u_1 + u_2} \qquad \partial x_3/\partial u_1 = e^{u_1+u_2} \qquad \partial x_3/\partial u_2 = e^{u_2}$

$u_1(t) = 1 + t/2 \qquad u_2(t) = 1 - t/2$

$du_1/dt = \frac{1}{2} \qquad du_2/dt = -\frac{1}{2}$

Simple Neural Networks :

Multivariate chain rule is very important for neural networks.



$$a^{(1)} = \sigma(Wa^{(0)} + b)$$

$$C = (a^{(1)} - y)^2 \rightarrow$$

$$\frac{\partial a^{(1)}}{\partial w}$$ Rate of change of activation with respect to the change of weights

Change of cost wrt to a is

$$\frac{\partial c}{\partial a^{(1)}}$$

So chang in weight w affects c

while $\frac{\partial c}{\partial a^{(1)}}$ affects c

$\frac{\partial a^{(1)}}{\partial w}$ affects $a^{(1)}$

$$\frac{\partial c}{\partial a^{(1)}} \frac{\partial a^{(1)}}{\partial w}$$

$$\frac{\partial c}{\partial w} = \frac{\partial c}{\partial a^{(1)}} \frac{\partial a^{(1)}}{\partial w}$$

$$\frac{\partial a^{(1)}}{\partial b}$$

$$\frac{\partial c}{\partial a^{(1)}} \frac{\partial a^{(1)}}{\partial b}$$

$$z^{(1)} = Wa^{(0)} + b$$

$$a^{(1)} = \sigma(z^{(1)})$$

$$\frac{\partial z^{(1)}}{\partial w}$$

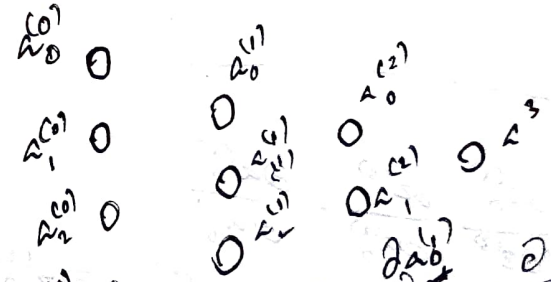
$$\frac{\partial z^{(1)}}{\partial b}$$

$$\frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial w}$$

$$\frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial b}$$

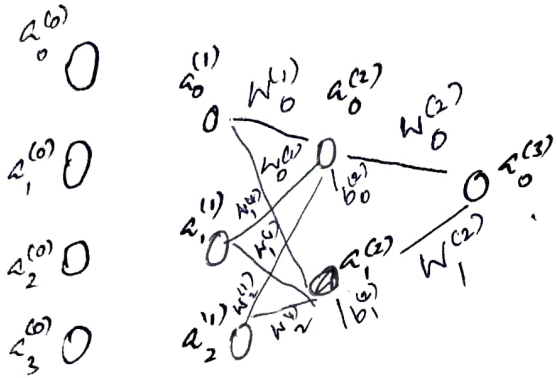
$$\frac{\partial c}{\partial a^{(1)}} \cdot \frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial w}$$

$$\frac{\partial c}{\partial a^{(1)}} \cdot \frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial b}$$



$$\frac{\partial c_0^{(1)}}{\partial a_0^{(1)}} \cdot \frac{\partial a_0^{(1)}}{\partial z_0^{(1)}} \cdot \frac{\partial z_0^{(1)}}{\partial w_0^{(0)}}$$

$$\frac{\partial c_0^{(1)}}{\partial a_0^{(1)}} \cdot \frac{\partial a_0^{(1)}}{\partial z_0^{(1)}} \cdot \frac{\partial z_0^{(1)}}{\partial b_0^{(0)}}$$



$$C = (a_0^{(3)} - y)^2$$

$$\frac{\partial C^{(3)}}{\partial a_0^{(3)}} \cdot \frac{\partial a_0^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial w_0^{(2)}}$$

$$\frac{\partial C^{(3)}}{\partial a_0^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial b_0^{(2)}}$$

$$\cancel{\frac{\partial C_0^{(2)}}{\partial a_0^{(2)}} \cdot \frac{\partial a_0^{(2)}}{\partial z_0^{(2)}} \cdot \partial z}$$

$$\frac{\partial C^{(3)}}{\partial a_0^{(3)}} \cdot \frac{\partial a_0^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a_0^{(2)}}$$

$$\frac{\partial C^{(3)}}{\partial a_0^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial}$$

Continued

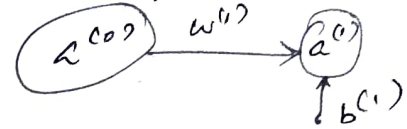
$$\frac{\partial C^{(3)}}{\partial a_0^{(3)}} \cdot \frac{\partial a_0^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a_0^{(2)}} \cdot \frac{\partial a_0^{(2)}}{\partial w_0^{(1)}}$$

$$\frac{\partial C^{(3)}}{\partial a_0^{(3)}} \cdot \frac{\partial a_0^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a_0^{(2)}} \cdot \frac{\partial a_0^{(2)}}{\partial z_0^{(2)}} \cdot \frac{\partial z_0^{(2)}}{\partial w_0^{(1)}}$$

$$\Rightarrow \frac{\partial C^{(3)}}{\partial w_0^{(1)}}$$

In this exercise we'll look in more detail about back-propagation, using the chain rule, in order to train our neural networks.

Lets look again at our two-node network



Recall the activation equations are

$$a^{(1)} = \sigma(z^{(1)})$$

$$z^{(1)} = w^{(1)}a^{(0)} + b^{(1)}$$

where we have introduce $z^{(1)}$ as the weighted sum of activation and bias.

We can formalise how good or bad our neural network is at getting the desired behaviour. For a particular input x and desired output y we can define the cost of that specific training example as the square of the difference between the network's output and the desired output, that is

$$C_k = (a^{(1)} - y)^2$$

where k labels the training example and $a^{(1)}$ is assumed to be the activation of output neuron when the input neuron $a^{(0)}$ is set to x.

We'll go into detail about how to apply this to an entire set of training data later on. But for now lets look at one by example

Recall our not genuine example from the previous quiz. For the input $x=1$ we would like that the network outputs $y=0$. For the starting weight and bias $w^{(1)} = 1.3$ and $b^{(1)} = -0.1$, the network actually outputs $a^{(1)} = 0.834$ so we work out the cost function for the example we get

$$C_1 = (0.834 - 0)^2 = 0.696$$

2) The cost function of a training set is the average of the individual cost functions of the data in the training set

$$C = \frac{1}{N}\sum_k C_k,$$

where N is the number of training sets in the example.

3) To improve the performance of the neural network on the training data, we can vary the weights and biases. We can calculate the derivative of the example cost with respect to these quantities using the chain rule

$$\frac{\partial C_k}{\partial w^{(1)}} = \frac{\partial C_k}{\partial a^{(1)}} \cdot \frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial w^{(1)}}$$

$$\frac{\partial C_k}{\partial b^{(1)}} = \frac{\partial C_k}{\partial a^{(1)}} \cdot \frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial b^{(1)}}$$

Individually, these derivatives take fairly simple form. Go ahead and calculate them. We'll repeat the defining equations for convenience.

$$a^{(1)} = \sigma(z^{(1)})$$

$$z^{(1)} = w^{(1)}a^{(0)} + b^{(1)}$$

$$C_k = (a^{(1)} - y)^2$$

$$\frac{\partial C_k}{\partial a}$$

$\sigma = \tanh$

$a = \sigma(z)$

$a = \tanh(z)$

$$\frac{\partial a}{\partial z} = \frac{1}{\cosh^2(z)}$$

Training this network is done by back-propagation because we start at the output layer and calculate derivatives backwards towards the input layer with the chain rule.

Let's see how this works

If we wanted to calculate the derivative of the cost with respect to the weight or bias of the final layer, then this is the same

$$\frac{\partial C_k}{\partial w^{(2)}} = \frac{\partial C_k}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial w^{(2)}}$$

$y = x^2$

$\frac{dy}{dx} = 2x$

derivative of y wrt x is $2x$.

With a similar form for the bias, if we want to calculate the derivative wrt to weights of the previous layers, we use the expression

$$\frac{\partial C_k}{\partial w^{(1)}} = \frac{\partial C_k}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial a^{(1)}} \cdot \frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial w^{(1)}}$$

where $\frac{\partial a^{(2)}}{\partial a^{(1)}}$ itself can be expanded to

$$\frac{\partial a^{(2)}}{\partial a^{(1)}} = \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial a^{(1)}}$$

This can be generalised to any layer

$$\frac{\partial C_k}{\partial w_i} = \frac{\partial C_k}{\partial a^{(N)}} \cdot \frac{\partial a^{(N)}}{\partial a^{(N-1)}} \cdot \frac{\partial a^{(N-1)}}{\partial a^{(N-2)}} \cdots \frac{\partial a^{(i+1)}}{\partial a^{(i)}} \frac{\partial a^{(i)}}{\partial z^{(i)}} \frac{\partial z^{(i)}}{\partial w^{(i)}}$$

By further application of the
choose the correct expression for the derivative

$$\frac{\partial a^{(j)}}{\partial a^{(j-1)}}$$

Remembering the activation equation

are

$$a^{(n)} = \sigma(z^{(n)})$$
$$z^{(n)} = w^{(n)} a^{(n-1)} + b^{(n)}$$

$$\frac{\partial a^{(1)}}{\partial z^{(1)}} \quad \frac{\partial z^{(1)}}{}$$

$$\frac{\partial a^{n}}{\partial a^{1}} = \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial a^{(1)}}$$

$$\frac{\partial a^{j}}{\partial a^{j-1}} ?$$

$$\frac{\partial a^{(n)}}{\partial z^{(n)}} = \sigma'(z^{(n)})$$
$$\frac{\partial z^{(n)}}{\partial a^{n-1}} = w^{(n)}$$

Back propagation:

In this assignment, you will train a neural network to draw a curve by implementing backpropagation by the chain rule to calculate the Jacobian of the cost function.

The neural network will be then ~~trained~~ trained using a Stochastic steepest descent method and will draw a series of curves to show the progress of training.

functions for the Jacobian of the cost function with respect to weights and biases. We will start with layer 3 and work backwards through the layers

$$J_{W^{(3)}} = \frac{\partial C}{\partial W^{(3)}}$$
$$J_{b^{(3)}} = \frac{\partial C}{\partial b^{(3)}}$$

Gives the maximum slope of the function at a given point

where C is the average cost function over the training set

$$C = \frac{1}{N} \sum_k C_k$$

$$\frac{\partial C}{\partial W^{(3)}} = \frac{\partial C}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial W^{(3)}}$$

$$\frac{\partial C}{\partial b^{(3)}} = \frac{\partial C}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial b^{(3)}}$$

NOTE:
Jacobian is a vector in an algebraic form that gives the maximum/steepest slope of the function for any x, y, z

With the partial derivatives taking the form

$$C = (a^{(3)} - y)^2$$
$$\frac{\partial C}{\partial a^{(3)}} = 2(a^{(3)} - y)$$
$$\frac{\partial a^{(3)}}{\partial z^{(3)}} = \sigma'(z^{(3)})$$
$$\frac{\partial z^{(3)}}{\partial W^{(3)}} = a^{(2)}$$
$$\frac{\partial z^{(3)}}{\partial b^{(3)}} = 1$$

$$z^{(3)} = W^{(3)} a^{(2)} + b^{(3)}$$

$$\frac{\partial c}{\partial w^{(3)}} = \frac{\partial c}{\partial a^{(3)}} \frac{\partial a^{(3)}}{\partial z^{(3)}} \frac{\partial z^{(3)}}{\partial w^{(3)}}$$

$$a^{(3)} = w^{(3)} a^{(2)} + b^{(3)}$$

$$\frac{\partial c}{\partial w^{(2)}} = \frac{\partial c}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial w^{(2)}}$$

$$a^{(2)} = \sigma(z^{(2)})$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = \sigma'(z^{(2)})$$

$$= \sigma'(w^{(2)} a^{(1)} + b^{(1)})$$

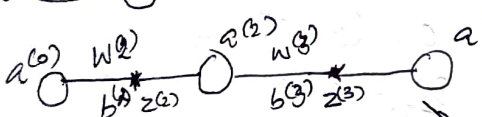
$$\cancel{\frac{\partial c}{\partial b^{(2)}} = \frac{\partial c}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial}}$$

?

$$\frac{\partial c}{\partial w^{(2)}} = \frac{\partial c}{\partial a^{(3)}} \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial w^{(2)}}$$

$$\frac{\partial c}{\partial b^{(2)}} = \frac{\partial c}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial b^{(2)}}$$

layer ②



y

$$c = (a^{(3)} - y)^2$$

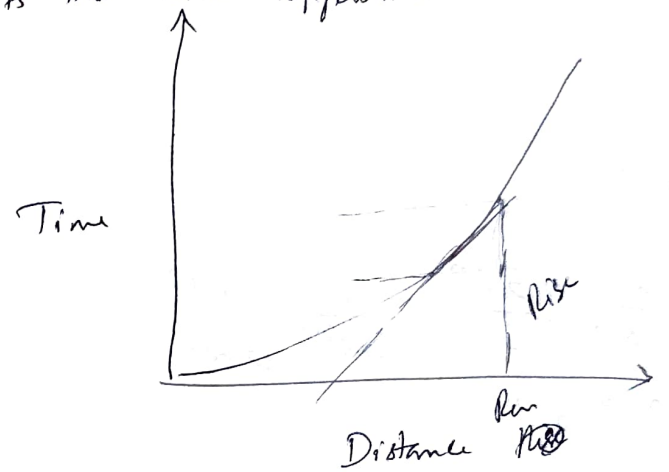
$$\frac{\partial c}{\partial a^{(3)}} = 2(a^{(3)} - y)$$

$$\frac{\partial a^{(3)}}{\partial z^{(3)}} = \sigma'(z^{(3)})$$

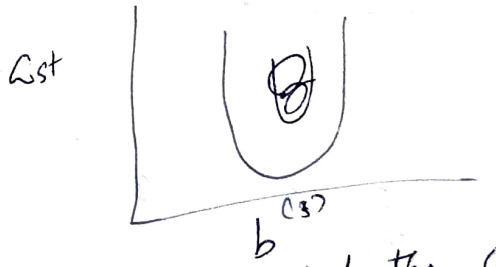
$$\frac{\partial z^{(3)}}{\partial w^{(3)}}$$

$$\frac{\partial z^{(3)}}{\partial b^{(3)}}$$

Lets think about differentiation



Similarly



Basically we try to find the change of cost wrt weight, bias, activation of the previous layer

The idea is to minimize the cost

For eg $\frac{\partial \text{cost}}{\partial w^{(3)}}$ should be minimum

$n_1 = 6 \quad n_2 = 7$

$W1 = 6$ rows 1 column $\quad W2 = 7$ rows 6 cl

Jacobian layer 2

$$\frac{\partial c}{\partial w^{(2)}} = \frac{\partial c}{\partial a^{(3)}}\left(\frac{\partial a^{(3)}}{\partial a^{(2)}}\right)\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial w^{(2)}}$$

$$\frac{\partial c}{\partial a^{(3)}} = 2(a^{(3)} - y)$$

$$\frac{\partial a^{(3)}}{\partial z^{(3)}} = \sigma'(z^{(3)})$$

$$z^{(3)} = W^{(3)}a^{(2)} + b^{(3)}$$

$$\frac{\partial z^{(3)}}{\partial a^{(2)}} = W^{(3)}$$

$$\frac{\partial z^{(3)}}{\partial a^{(2)}} = W^{(3)}$$

$$\frac{\partial a^{(3)}}{\partial a^{(2)}} = \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a^{(2)}} = \sigma'(z^{(3)}) \cdot W^{(3)}$$

$$a^{(2)} = \sigma(z^{(2)}) \qquad z^{(2)} = W^{(2)}a^{(1)} + b^{(2)}$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = \sigma'(z^{(2)})$$

$$\frac{\partial z^{(2)}}{\partial w^{(2)}} = a^{(1)} \qquad \frac{\partial z^{(2)}}{\partial b^{(2)}} = 1 \qquad 2(a^{(3)} - y) \cdot \sigma'(z^{(3)}) \cdot W^{(3)} \cdot \sigma'(z^{(2)}) \cdot a^{(1)}$$

W1

$$\sigma\left(\begin{bmatrix}1\\2\\3\\4\\5\\6\end{bmatrix}[a^0] + b\right) = \begin{bmatrix}a_0^{(1)}\\a_1^{(1)}\\a_2^{(1)}\\a_3^{(1)}\\a_4^{(1)}\\a_5^{(1)}\\a_6^{(1)}\end{bmatrix} \qquad \begin{bmatrix}1&2&3&4&5&6\\1&2&3&4&5&6\\1&2&3&4&5&6\\1&2&3&4&5&6\\1&2&3&4&5&6\\1&2&3&4&5&6\\1&2&3&4&5&6\end{bmatrix}$$

Layer 1 is similar to Layer 2 with an additional partial derivative

Layer 1 Jacobian

$$\frac{\partial c}{\partial w^{(1)}} = \frac{\partial c}{\partial a^{(3)}}\left(\frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial a^{(1)}}\right)\frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial w^{(1)}}$$

$$\frac{\partial c}{\partial b^{(1)}} = \frac{\partial c}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial a^{(1)}} \cdot \frac{\partial a^{(1)}}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial b^{(1)}}$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = \sigma'(z^{(2)}) \qquad z^{(2)} = W^{(2)}a^{(1)} + b^{2}$$

$$\frac{\partial z}{\partial z}$$

$$z^{(3)} = W^{(3)}a^{(2)} + b^{(3)}$$

$$\frac{\partial z^{(3)}}{\partial a^{(2)}} = W^{(3)}$$

$$\frac{\partial z^{(2)}}{\partial a^{(1)}} = W^{(2)} \qquad a_1 = \sigma(z^{(1)}) = \sigma(W^{(1)}a^{(0)} + b^{(1)})$$

$$\frac{\partial a^{(1)}}{\partial z^{(0)}} =$$

$$\frac{\partial z^{(1)}}{\partial w^{(1)}} = a^{(0)}$$

$$\frac{\partial a^{(2)}}{\partial a^{(1)}} = \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial a^{(1)}} = \sigma'(z^{(2)}) \cdot W^{(2)}$$

$$\frac{\partial a^{(1)}}{\partial z^{(1)}} = \sigma'(z^{(1)}) \qquad z^{1} = W^{1}a^{0} + b^{1}$$

$$\frac{\partial z^{1}}{\partial b^{1}} = 1$$

$$\frac{\partial a^{(3)}}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial a^{(1)}} = \sigma'(z^{(3)}) \cdot W^{(3)} \cdot \sigma'(z^{(2)}) \cdot W^{(2)}$$

$$\frac{\partial a^{(1)}}{\partial z^{(1)}} = \sigma'(z^{(1)})$$

Jacobian

Layer 3:

$$\frac{\partial c}{\partial a^{(3)}} = 2(a^{(3)} - y)$$

$$\frac{\partial a^{(3)}}{\partial z^{(3)}} = \sigma'(z^{(3)})$$

$$\frac{\partial z^{(3)}}{\partial w^{(3)}} = a^{(2)}$$

$$\frac{\partial z^{(3)}}{\partial b^{(3)}} = 1$$

Layer 2:

$$\frac{\partial c}{\partial a^{(3)}} = 2(a^{(3)} - y)$$

$$\frac{\partial a^{(3)}}{\partial a^{(2)}} = \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a^{(2)}}$$

$$= \sigma'(z^{(3)}) \cdot w^{(3)}$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = \sigma'(z^{(2)})$$

$$\frac{\partial z^{(2)}}{\partial w^{(2)}} = a^{(1)} \qquad \frac{\partial z^{(2)}}{\partial b^{(2)}} = 1$$

$2(a^{(3)} - y)$

$\sigma'(z^{(3)})$

$w^{(3)}$

$\sigma'(z^{(2)})$

$a^{(1)}$

Layer 1

$$\partial c / \partial a^{(3)} = 2(a^{(3)} - y)$$

$$\frac{\partial a^{(3)}}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial a^{(1)}} = \sigma'(z^{(3)}) \cdot w^{(3)}$$

$$= \sigma'(z^{(3)}) \cdot w^{(3)} \cdot \sigma'(z^{(2)}) \cdot w^{(2)}$$

$$\frac{\partial a^{(1)}}{\partial z^{(1)}} = \sigma'(z^{(1)}) \qquad \frac{\partial z^{(1)}}{\partial w^{(1)}} = a^{(0)}$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial a^{(1)}}$$

$\sigma'(z^{(2)}) \cdot w^{(2)}$

$z^{(1)} = w^{1} a^{(0)} + b^{1}$

$J = 2(a^{(3)} - y)$

$J = J * \text{d_sigma}(z^{(3)})$

$J = np.(J.T @ w3).T$

$J = J * \text{d_sig}(z^{(2)})$

$J = J * (J.T @ w2).T$

$J = J * \text{d_sigma}(z^{(1)})$

$J = J @ a^{(0)}.T$

Building approximate functions.

Taylor Series (Power Series):

$g(x) = a + bx + cx^2 + dx^3 + \dots$

Power Series Derivation.

Eg:

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \dots$$

Taylor series tells us that if we know everything about a function at one place, where "everything" is the value, first derivative, second derivative and so on we can use the information to reconstruct the function everywhere else.

This is true only for functions that are considered

- i) Well behaved
- ii) Continuous
- iii) You can differentiate as many times you want

There are many such function

$y = mx + c$

$\frac{dy}{dx} = m \qquad m = f'(P)$

$f'(P)x$

$y = mx + c = f(0)$

$\frac{dy}{dx} = m = f'(0)$

$y = mx + c = f(P)$

$\frac{dy}{dx} = m = f'(P)$

When $x = P$ $c =$

$mP + c = f(P)$

$c = f(P) - mP$

$= f(P) - f'(P)P$

$y = mx + c$

$= f'(P)x + f(P) - f'(P)P$

$f'(P)(x - P) + f(P)$

$g_0(x) = f(P)$

$g_1(x) = f'(P)(x-P) + f(P)$

$g_2(x) \rightarrow$ Surprising from Maclaurin Series

$f(P) + f'(P)(x-P) + \frac{1}{2}f''(P)(x-P)^2$

↳ This is Taylor Series

$y = ax^2 + bx + c$ $\quad c = f(0)$

$x=0$

$\frac{dy}{dx} = 2ax + b$ $\quad b = f'(0)$

$x=0$

$\frac{d^2y}{dx^2} = 2a = f''(0)$

$a = \frac{f''(0)}{2}$

$\frac{f''(0)}{2}x^2 + f'(0)x + f(0)$

↳ Maclaurine Series

Maclaurin Series

$$g(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

Taylor Series

$$g(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(p)}{n!} (x-p)^n$$

↓

One dimensional Taylor Series

1) For the function $f(x) = e^{x^2}$ about $x=0$, using the Maclaurin series formula, obtain an approximation upto the first three non zero terms.

Maclaurin formula:

$$\frac{f''(0)}{2} x^2 + f'(0)x + f(0)$$

$$\frac{d}{dx} e^u = e^u \cdot u'$$

$$f(x) = e^{x^2}$$

$$f'(x) = e^{x^2} \cdot 2x$$

$$f''(x) = e^{x^2} \cdot 2 + 2x \cdot 2x \cdot e^{x^2}$$

$$= 2e^{x^2} + 4x^2 e^{x^2}$$

$$f'''(x) = 2e^{x^2} \cdot 2x + (4x^2 \cdot e^{x^2} \cdot 2x + e^{x^2} 8x)$$

$$= 4xe^{x^2} + 8x^3 e^{x^2} + 8xe^{x^2}$$

$$= 12xe^{x^2} + 8x^3 e^{x^2}$$

$$f'''(0) = 0$$

$$f''(0) = 2$$

$$f'(0) = 0$$

$$f(0) = 1$$

$$f''''(x) = 12x \cdot 2x e^{x^2} + e^{x^2} \cdot 12$$

$$= 24x^2 e^{x^2} + 8x^3 2x e^{x^2} + e^{x^2} 24x^2 + 12e^{x^2} + 24x^2 e^{x^2} + 16x^3 e^{x^2}$$

$$f''''(0) = 0 + 12$$

2) Use the Taylor series formula to approximate the first three terms of the function $f(x) = 1/x$ expanded around the point $x = 4$

Taylor's Series

$$\frac{f''(p)}{2}(x-p)^2 + f'(p)(x-p) + f(p)$$

$$f(x) = 1/x \qquad x^{-1}$$

$$f'(x) = \frac{-1}{x^2} = -x^{-2}$$

$$f''(x) = \frac{2}{x^3} = 2x^{-3}$$

$$f'''(x) = \frac{-6}{x^4} = -6x^{-4}$$

$$f''''(x) = \frac{24}{x^5} = 24x^{-5}$$

$$\frac{24}{2 \cdot x^5}(x-4)^2 - \frac{1}{x^2}(x-4) + 1/x$$

$$\frac{(x-4)^2}{64} - \frac{1}{16}(x-4) + 1/4$$

$$\frac{1}{x^2}(x-4)$$

$$\frac{1}{x^2}(x-4)$$

$$\frac{1}{(x-4)}$$

$$\frac{1}{x-4} - \frac{1}{(x-4)^2}(x-4) + 2$$

$$(x-4)^2 - (x-4) + 2$$

$$(x-4)^3$$

$$x^2 + 16 - 8x - x + 4 + 2$$

$$\frac{1}{(x-4)^3}(x-4)^2$$

$$\frac{-1}{(x-4)^2} \cdot (x-4)$$

$$+ 1/x - 4$$

$$\frac{1}{x-4} - \frac{1}{x-4}$$

$$+ \frac{1}{x-4}$$

3) By finding the [illegible] shown above for the function $f(x) = \ln(x)$ about $x = 10$, determine the difference of using the second order taylor expansion against the first order Taylor expansion when approximating to find the value of $f(2)$.

Taylor series

$$\frac{f''(x)}{2}(x)^2 + f'(x) \cdot x + f(x)$$

$f(x) = \ln(x)$

$f'(x) = \frac{1}{x}$

$f''(x) = -\frac{1}{x^2}$

$$-\frac{1}{2x^2} \cdot x^2 + \frac{1}{x} \cdot x + x$$

Second order $-\frac{1}{2} + 1 + 2 = 2.5$

First order $\cancel{-\frac{1}{2} + 1 = 0.5}$ $1 + 2 = 3$

$\Delta = 2$

$f(x) = e^x$

$f'(x) = e^x$

$f''(x) =$

4) For some cases, a Taylor series can be expressed in a general equation that allows us to find a particular n^{th} term of our series. For example the function $f(x) = e^x$ has the general equation $f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Therefore if we want to find the 3rd term in our Taylor series substituting $n = 2$ in the general equation gives us the term $\frac{x^2}{2}$. We know the Taylor series of the function e^x is $f(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \dots$ Now let us try a further working example of using general equations with Taylor series.

By evaluating the function $f(x) = \frac{1}{(1-x)^2}$ about the origin $x = 0$, determine which general equation for the nth order correctly represents $f(x)$

$f(x) = \frac{1}{(1-x)^2}$

$f'(x) = \frac{1}{2x^2}$

$f''(x) = \frac{1}{2}$

$(1-x)^{-2}$

$(1-x)^{-3} \cdot (1-x)^{-1}$

$(1-x)^{-1} \cdot (-1)(-1)$

$\times (1-x)^{-1}(-1)(-1)$ or

$1 + x^2 - 2x$

$2x - 2$

$(1-x)(1-x)$

$(1-x)(-1) + (1-x)(-1)$

$-1 + x - 1 + x$

$2x - 2$

Wrong

Study the chain rule again

$\frac{f''(x)}{2} \cdot x^2 + f'(x) \cdot x + f(x)$

$1 \quad -\frac{1}{2} \times \frac{1}{2}$

$2f''(x) \cdot x^2 - 2f'(x) \cdot x$

$f'(x)(2x-2) = 1$

$f'(x)2x - f'(x)2 = 1$

$\frac{1}{2x-2}$

$\sum_{n=0}^{\infty} (1+n)(-x)^n$

$n=0$ $x=0$ 1

$n=1$ $x=0$ 2

1

$-\frac{1}{2}$

$\frac{1}{2}$

$\frac{1}{2}x^2 + \frac{1}{2x-2}x + \frac{1}{(1-x)^2}$

$f(x) = \frac{1}{(1-x)^2}$

$f(0) = 1$

$f(1) = 0$

$f(2) = 1$

$f(3) = \frac{1}{4}$

$f(4) = \frac{1}{9}$

$\frac{d}{dx}\left(\frac{C_1}{(1-x)^2}\right) = \frac{d}{dx}\left(\frac{1}{(1-x)(1-x)}\right)$

$= \frac{d}{dx}\left((1-x)^{-1} \cdot (1-x)^{-1}\right)$

$= (1-x^{-1}) - 1 \cdot$

$(1-x)^{-1} - 1 \cdot -x^{-2}$

$+ (1-x)^{-1} - 1 \cdot -x^{-2}$

$(1-x)^{-1} \cdot x^{-2}$

$(1-x)^{-1} \cdot (1-x)^{-1}$

$f'(x) \quad \frac{2}{(1-x)^3}$

$f''(x) = \frac{6}{(1-x)^4}$

$f'''(x) = \frac{-24}{(1-x)^5}$

Study the chain rule again

$\frac{1}{x^2-2x+1}$

$f'(x) \quad (1+2n)x^n$

$x=2$

$n=0 \quad 1$

$n=1 \quad 6$

$n=2 \quad 20$

$f(x) = \frac{1}{(1-x)^2}$

$f'(x) = \frac{2}{(1-x)^3}$

$f''(x) = \frac{6}{(1-x)^4}$

$f'''(x) = \frac{-24}{(x-1)^5}$

$f''''(x) = \frac{120}{(x-1)^6}$

$f(0) = 1$

$f'(0) = \frac{2}{1} = 2$

$f''(0) = 6$

$f'''(0) = 24$

$f''''(0) = 120$

$f(2) = 1 \quad n=0$

$f'(2) = -2 \quad 1$

$f''(2) = 6 \quad 2$

$f'''(2) = -24 \quad 3$

$\sum_{n=0}^{\infty} (1+n)(-x)^n$

$(1)(-2)^0 = 1$

$2 \cdot -2 = -4$

$f(x) = \sum_{n=0}^{\infty} (1+2n)(x)^n \qquad (2+n)x^n$

$f(2)$

$n=0 \quad 1$

$n=1 \quad 6$

$f(0)$

$n=0 \quad 1$

$n=1 \quad 0$

$n=2 \quad 0$

When $x=3$

$(1+n)(-x)^n$

$x=2$

$n=0 \quad 1(-x^0)$

1

$2 \cdot 2 = -4$

$f(x) = \sum_{n=0}^{\infty} (1+n)x^n$

$\sum_{n=0}^{\infty} (2+n)(x)^n$

$x=2$

$n=0 \quad 2$

$n=1 \quad 6$

$n=2 \quad 16$

$(1+n) - x^n$

$x=2$

$n=0 \quad 1$

$n=1 \quad -4$

$n=2 \quad -12$

1 2 6 24 120

$$1 + \frac{2x}{2!} + \frac{6x^2}{3!} + \frac{24x^3}{4!} + \frac{120x^4}{5!}$$

$x=2$

$1 + 2 + 4 + 8$

$\frac{24}{2}$ 3

$\frac{24}{192}$ 8

24) 192 (8

$x=1$

$1 + 1 + 24$

$$\sum_{n=0}^{\infty} (1+n)x^n$$

$x=2$

$(2+n)x^n$

$n=0 \quad x=2$

$n=0$ 1

$n=1$ 4

$n=2$ 27

1 2 6 24 120

1 2x

$$\frac{f''(0)}{2}x^2 + f'(0)x + f(0)$$

$6 + 2 + 1$

$$f(p) + f'(p)(x-p) + \frac{f''(p)}{2}(x-p)^2$$

24) 120 (5
120

2

32

1 2 6 24 120

$$1 + 2x + \frac{6x^2}{2!} + \frac{24x^3}{3!} + \frac{120x^4}{4!}$$

$1 + 2x + 3x^2 + 4x^3 + 5x^4$

0 1 2 3

By evaluating the function $f(x) = \sqrt{4-x}$ at $x=0$, find the quadratic equation that approximate the function when $x=0$

$f(x) = \sqrt{4-x}$ | 2

$f'(x) = -\dfrac{1}{2\sqrt{4-x}}$ | $-\dfrac{1}{4}$

$f''(x) = -\dfrac{1}{4(4-x)^{3/2}}$ | $-\dfrac{1}{4 \cdot 4^{3/2}} = -\dfrac{1}{32}$

$f'''(x) = -\dfrac{3}{8(4-x)^{5/2}}$ | $-\dfrac{3}{8 \cdot 4^{5/2}} = -\dfrac{3}{256}$

$n =$

$f(x) = 2 - x - \dfrac{x^3}{64}$

$f'(x) = -1 - \dfrac{3x^2}{64}$

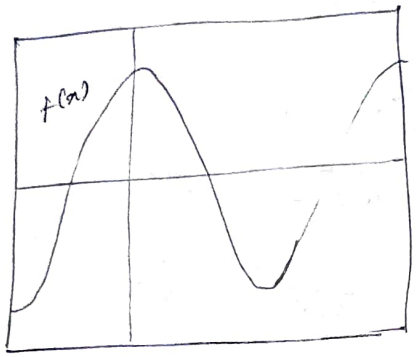
$f''(x) = \dfrac{6x}{64}$

$f'''(x) = \dfrac{6}{64}$

$\sqrt{2} - \dfrac{1}{2\sqrt{2}}x - \dfrac{1}{8 \cdot 4^{3/2}}x^2$

$2 - \dfrac{1}{4}x - \dfrac{1}{64}x^2$

McLaurin Series for Cosine function:



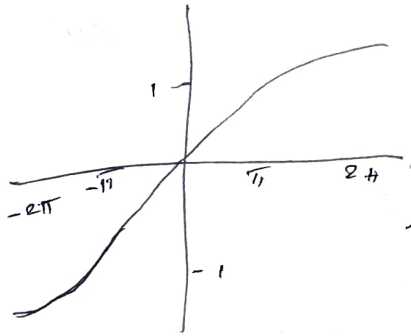
$f(x) = \cos(x)$

$$g(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

Cosine is an epitome of:
i) a well behaved function of Continuous
ii) infinitely differentiable

Linearisation:

The graph below shows the function $f(x) = \tan^{-1}(x)$



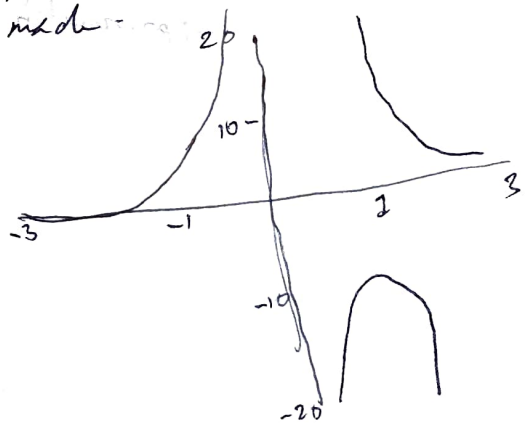
By using the Maclaurin series or otherwise, determine the function shown above is even, odd or neither

	At 0
$f(x) = \tan^{-1}(x)$	0
$f'(x) = \frac{1}{x^2+1}$	1
$f''(x) = \frac{2x}{(x^2+1)^2}$	0

$f(-x) = -f(x)$

odd

2) The graph below shows the discontinuous function $f(x) = \frac{2}{(x^2-x)}$. For this function, select the starting point that will allow a Taylor approximation to be made.



Multivariate Taylor

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(p)}{n!}(x-p)^n \qquad f(x+\Delta x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}\Delta x^n$$

Taylor's series can be expressed in terms of x & Δx

$\frac{dy}{dx}$ rate of change of y / rate of change of x

$\frac{dx}{dy}$ → rate of change of x / rate of change of y

$$f(x) = f(p) + f'(p)(x-p) + \frac{1}{2}f''(p)(x-p)^2 + \frac{1}{6}f^{(3)}(p)(x-p)^3 +$$

$$f(x+\Delta x) = f(x) + f'(x)\Delta x + \frac{1}{2}f''(x)\Delta x^2 + \frac{1}{6}f''(x)\Delta x^3 +$$

Two dimensional Taylor series

$$f(x+\Delta x, y+\Delta y) =$$

$f(x,y) \rightarrow 0^{th}$ order

$+ (\partial_x f(x,y)\Delta x + \partial_y f(x,y)\Delta y) \rightarrow$ first order $\xrightarrow{\text{Can be expressed as Jacobian}}$ $[\partial_x f(x,y), \partial_y f(x,y)]\begin{bmatrix}\Delta x \\ \Delta y\end{bmatrix} \Rightarrow J_f \Delta \mathbf{x}$

$+ \frac{1}{2}(\partial_{xx} f(x,y)\Delta x^2 + 2\partial_{xy} f(x,y)\Delta x \Delta y + \partial_{yy} f(x,y)\Delta y^2)$

↳ 2nd order approximation $\xrightarrow{\text{Can be expressed as Hessian matrix}}$

$$\frac{1}{2}[\Delta x, \Delta y]\begin{bmatrix}\partial_{xx} f(x,y) & \partial_{xy} f(x,y) \\ \partial_{xy} f(x,y) & \partial_{yy} f(x,y)\end{bmatrix}\begin{bmatrix}\Delta x \\ \Delta y\end{bmatrix}$$

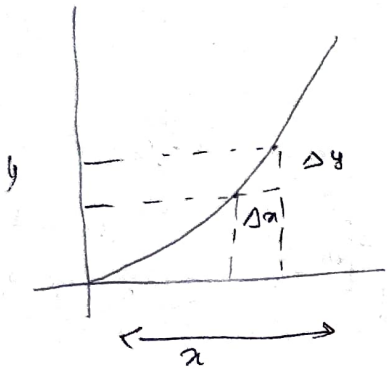
$$\Downarrow$$

$$\frac{1}{2}\Delta \mathbf{x}^t H_f \Delta \mathbf{x}$$

Finally

$f(x+\Delta x)$

Detour



2D Taylor series:

Now you have seen how to calculate the multivariate Taylor series, and what zeroth, first and second order approximations look like for a function of 2 variables. In this course we won't be considering anything higher than second order for functions of more than one variable.

In the following questions you will practise recognising these approximations by thinking about how they behave diagonal with x and y, then you will calculate the some terms in the multivariate Taylor series yourself.

The following plot features a surface and its Taylor approximation

Recall that upto second order the multivariate Taylor series is given by

$$f(x+\Delta x) = f(x) + J_f \Delta x + \frac{1}{2} \Delta x^T H_f \Delta x + .$$

Consider the function of 2 variables $f(x,y) = xy^2 e^{-x^4-y^2/2}$. which is the first order Taylor series expansion of f around the point $(-1,2)$?

$$f(-1,2) = -1 \cdot 2^2 e^{-1^4-4/2}$$
$$= -4e^{-3}$$

$$f'(x,y) = \left[\partial_x f(x,y), \partial_y f(x,y)\right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

$$= \cancel{y^2 \cdot 4x^3 e^{-x^4-y^2/2}}$$

$$= \left[y^2 \cdot -4x^3 - \frac{y^2}{2} e^{-x^4-y^2/2}\right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

$$x \cdot x^4 - y e^{-x^4-y^2/2}$$

$$f(-1,2) \quad 1 \cdot -4(-1) - \frac{4}{2} e^{-1}$$

$$-2e^{-1}$$

$$2e^{-1} \Delta x$$

Steps

$$= y^2 \frac{\partial}{\partial x} \left(x e^{-x^4-\frac{y^2}{2}}\right)$$

Apply product rule $(f \cdot g)' = f' \cdot g + f \cdot g'$

$$\partial/\partial x (x) = 1$$

$$\partial/\partial x \left(e^{-x^4-\frac{y^2}{2}}\right) = -4x^3 e^{-x^4-y^2/2}$$

$$= y^2 \left(1 \cdot e^{-x^4-y^2/2} + \left(-4e^{-x^4-y^2/2} x^3\right) x\right)$$

$$= y^2 \left(e^{-x^4-y^2/2} - 4e^{-x^4-\frac{y^2}{2}} x^4\right)$$

$$\frac{\partial}{\partial y}\left(xy^2 e^{-x^4-\frac{y^2}{2}}\right)$$

Apply the product rule

$$(f\cdot g)' = f'\cdot g + f\cdot g'$$

$$x\frac{\partial}{\partial y}\left(y^2 e^{-x^4-\frac{y^2}{2}}\right)$$

$$\frac{\partial}{\partial y}(y^2) = 2y$$

$$\frac{\partial}{\partial y}\left(e^{-x^4-\frac{y^2}{2}}\right) = e^{-x^4-\frac{y^2}{2}}\cdot y$$

$$2ye^{-x^4-\frac{y^2}{2}} + y^3 e^{-x^4-\frac{y^2}{2}}$$

$$\Rightarrow x\left(2ye^{-x^4-\frac{y^2}{2}} + y^3 e^{-x^4-\frac{y^2}{2}}\right)$$

$$\cancel{x e^{-x^4-\frac{y^2}{2}}(2y+y^3)}$$

$$x\left(2e^{-x^4-\frac{y^2}{2}}y - e^{-x^4-\frac{y^2}{2}}y^3\right)$$

Jacobian

$$y^2\left(e^{-x^4-y^2/2} - 4e^{-x^4-\frac{y^2}{2}}x^4\right) \quad x\left(2e^{-x^4-\frac{y^2}{2}} - e^{-x^4-\frac{y^2}{2}}y^3\right)$$

$f(-1,2)$

$$4\left(e^{-(-1)^4-2^2/2} - 4e^{-(-1)^4-\frac{2^2}{2}}(-1)^4\right) \quad -1\left(2e^{-(-1)^4-\frac{2^2}{2}} - e^{-(-1)^4-\frac{2^2}{2}}\cdot 8\right)$$

$$4\left(e^{-3} - 4e^{-3}\right) \quad -1\left(2e^{-3} - e^{-3}\cdot 8\right)$$

$$12e^{-3} \quad 8e^{-3} - 2e^{-3}$$

$$6e^{-3}$$

Now consider the function $f(x,y) = \sin(\pi x - x^2 y)$. What is the Hessian matrix H_f that is associated with the second order term in the Taylor expansion of f around $(1,\pi)$

$$f(x,y) = \sin(\pi x - x^2 y)$$

$$\frac{\partial}{\partial x}\left(\sin(\pi x - x^2 y)\right)$$

$$= (\pi - 2xy)\cos(\pi x - x^2 y)$$

$$\frac{\partial}{\partial y}\left(\sin(\pi x - x^2 y)\right)$$

$$= (\pi x - x^2)\cos(\pi x - x^2)$$

$$H = \begin{bmatrix} \partial_{xx}f & \partial_{xy}f \\ \partial_{yx}f & \partial_{yy}f \end{bmatrix}$$

$$\cancel{\partial_{xx}f = (\pi - 2xy)\sin(\pi x - 2xy)}$$

$$\partial_{xx}f = (\pi - 2xy)\cdot -\sin(\pi x - x^2 y)(\pi - 2xy) + (\pi - 2y)\cos(\pi x - x^2 y)$$

$$\cancel{\partial_{xx}f} = (1,\pi) \quad (1-2\pi) - \sin(\pi - \pi)(\pi - 2\pi) + (\pi - 2\pi)\cos(\pi - \pi)$$

$$(-\pi) + -\pi\cos(\pi - \pi)$$

$$-\pi \quad -\pi$$

$$\frac{\partial}{\partial y}\left((\pi x - x^2)\cos(\pi x - x^2)\right)$$

$$(\pi x - x^2)$$

Taylor Series assessment:

$f(x) = x\sin(x)$

$f'(x) = x\cos(x) + \sin(x)$

$f''(x) = -x\sin(x) + \cos(x) + \cos(x)$

$f'''(x) = -x\cos(x) - \sin(x) - \sin(x) - \sin(x)$

$f^{IV}(x) = -x\sin(x) - \cos(x) - \cos(x) - \cos(x) - \cos(x)$

$-x\sin(x) - 4\cos(x)$

$f^{5}(x) = -x\cos(x) - \cos(x) - 4\cos(x)$

$= -x\cos(x) - 5\cos(x)$

$f^{6}(x) = -x\sin(x) - 6\cos(x)$

$f(x) = x\sin(x)$ $\quad f'(x) = x\cos(x) + \sin(x)$

$f''(x) = -x\sin(x) + \cos(x) + \cos(x) = -x\sin(x) + 2\cos(x)$

$f'''(x) = -x\cos(x) - \sin(x) - 2\sin(x) = -x\cos(x) - 3\sin(x)$

$f^{4}(x) = x\sin(x) - \cos(x) - 3\sin(x) = x\sin(x) - 4\cos(x)$

$f^{5}(x) = x\cos(x) + \sin(x) + 4\sin(x) = x\cos(x) + 5\sin(x)$

$f^{6}(x) = -x\sin(x) + \cos(x) + 5\cos(x) = -x\sin(x) + 6\cos(x)$

$f(0) = 0$ $\quad f''(0) = 2$ $\quad f^{5}(0) = 0$

$f'(0) = 0$ $\quad f'''(0) = 0$ $\quad f^{6}(0) = 6$

$f^{4}(0) = -4$

2) Find the first four non-zero terms of the Taylor expansion for the function $f(x) = e^x + x + \sin(x)$ about $x=0$

$f(0) = e^0 + 0 + \sin(0) = 1$

$f'(x) = e^x + 1 + \cos(x)$

$\xrightarrow{x=0} 1 + 1 + 1 = 3$

$f^{2}(x) = e^x - \sin(x)$

$\xrightarrow{x=0} 1$

$f^{(3)}(x) = e^x - \cos(x)$

$x \to 0 \quad 1 - 1 = 0$

$f^{4}(x) = e^x + \sin(x)$

$x \to 0 \quad 1$

$$1 + 3x + \frac{x^2}{2} + \frac{x^4}{24}$$

3) The graph below shows the discontinuous function $f(x) = \frac{2}{(x^2 - x)}$. Approximate the section of this function that covers the domain $0 < x < 1$. Use the Taylor series formula and $x = 0.5$ as your starting point, find the first two non-zero terms

$f(x) = \frac{2}{x^2 - x} \Rightarrow -8$

$f'(x) = 2(x^2 - x)^{-1} = -2(2x-1)^{-2} = \frac{-2}{(2x-1)^2}$

$f^{2}(x) = \frac{-2}{(2x-1)^2} = -2(2x-1)^{-2} =$

$f'(x) = 2\frac{d}{dx}\left(\frac{1}{x^2 - x}\right)$

$= \frac{-2\cdot(2x-1)}{(x^2-x)^2}$

$\frac{-2(2\cdot .5 - 1)}{(.5^2 - .5)^2}$

$f^{2}(x) = -2\left[\frac{(2)\cdot -2(2x-1)}{(x^2-x)^3}\right] = \frac{-8(2x-1)}{(x^2-x)^3}$

$=$

$$f''(x) = \frac{d}{dx}\left[-\frac{2(2x-1)}{(x^2-x)^2}\right]$$

$$= -2\frac{d}{dx}\left[\frac{2x-1}{(x^2-x)^2}\right]$$

$$= \cancel{-2 \cdot \frac{d}{dx}(2x-1)\cdot(x^2-x)^2}$$

$$\frac{4(3x^2-3x+1)}{(x-1)^3x^3}$$

$$= \frac{4(.75-1.5+1)}{-(.125)(.125)}$$

$$= \frac{1}{.015625}$$

$$\frac{4(x-6+7)}{8}$$

$$\frac{2^2}{4}$$

$$4(3(x-0.5)^2 - 3(x-0.5)+1$$

$$12(x^2+0.25-x) - 3x+1.5)+1$$

5) $f(x) = e^{-2x}$ Linearise the function

$f'(x) = -2e^{-2x}$

$f''(x) = 4e^{-2x}$

$f'''(x) = -8e^{-2x}$

$f^{4}(x) = 16e^{-2x}$

$$-2e^{-2x} + 4e^{-2x}x - 8e^{-2x}x^2$$

$$\frac{1}{e^{2x}}(-2+4x-8x^2)$$

$x=2$ $\quad \frac{1}{e^4}(-2+8-32)$

$$= \frac{1}{e^4}(-26)$$

$$e^{-2x} - 2e^{-2x}x + 4e^{-2x}x^2 - 8e^{-2x}x^3 + 16e^{-2x}x^4)$$

$$\frac{1}{e^{2x}}(1-2x+4x^2-8x^3+16x^4)$$

$$\cancel{\frac{1}{e^{2x}}(1-}$$

$x=2$ $\quad \frac{1}{e^4}(1-4+16$

$$\frac{1}{e^4}[1-2(x-2)] + 4(x-2)^2 + O(\Delta x^3)$$

$$1-2x+4+4(x^2-4x+2)$$

$$1-2x+4+4x^2-16x+2)$$

$$1+3x+\frac{x^2}{2}+\frac{x^4}{24}$$

odd

$$\frac{1}{e^4}[1-2(x-2)] + O(\Delta x^2)$$

Module 5 :

Fitting as minimization problem :

Newton - Raphson method :

$y = x^3 - 2x + 2$

$\frac{dy}{dx} = 3x^2 - 2$

↳ This quadratic equation has two solutions so two turning points exist.

$$f(x_{i+1}) = f(x_i) - \frac{f(x_i)}{f'(x_i)}$$

i	x_i	$y(x_i)$	$\frac{\partial y(x_i)}{\partial x}$
0	-2	-2	10
1	-1.8	-0.23	7.7
2	-1.77	-0.005	7.4
3	-1.769	$2.3E^{-6}$	

$$f(x_1) = f(-2) - \frac{f(-2)}{f'(-2)}$$

$$= -2 - \frac{(-2)}{10}$$

$$= -2 + \frac{2}{10}$$

$$= -1.8$$

$$f(x_2) = f(-1.8) - \frac{f(-1.8)}{f'(-1.8)}$$

If we linearise about a particular point x_0, we can ask what the value of the function is at the point $x_0 + \delta x$ a short distance away

$$f(x_0 + \delta x) = f(x_0) + f'(x_0)\delta x$$

↳ $\|^{ly}$ $f(x_0) + f'(x_0)\Delta x$

Then if we assume that the function goes to zero somewhere nearby, we can rearrange to find how far away i.e. assume $f(x_0 + \delta x) = 0$ and solve for δx.

This became $\delta x = \frac{-f(x_0)}{f'(x_0)}$

Since the function $f(x)$ is not a line, this formula will try to get closer to the root, but won't exactly hit it. But this is ok because we can repeat the process from the new starting point to get even closer

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

This is the Newton-Raphson method used widely to find the roots of the function

$$f(x) = \frac{x^6}{6} - 3x^4 - \frac{2x^3}{3} + \frac{27x^2}{2} + 18x - 30$$

$$f'(x) = \frac{6x^5}{6} - 12x^3 - \frac{6x^2}{3} + \frac{54x}{2} + 18$$

$$= x^5 - 12x^3 - 2x^2 + 27x + 18$$

Location of the root near $x=1$

$$f(x_{i+1}) = f(x_i) - \frac{f(x_i)}{f'(x_i)}$$

$$f(x) = \frac{x^6}{6} - 3x^4 - \frac{2x^3}{3} + \frac{27x^2}{2} + 18x - 30$$

$$f'(x) = x^5 - 12x^3 - 2x^2 + 27x + 18$$

i	x	y	$f'(x)$
0	1	-2	42
	-1.95		

$f'(x)$

$1 - 12 - 2 + 27 + 18$

$= 1 - 14 + 27 + 18$

$= 1 + 4 + 27$

$= 32$

$f(x)$

$\frac{1}{6} - 3 - \frac{2}{3} + \frac{27}{2} + 18 - 30$

$\frac{1}{6} - \frac{2}{3} + \frac{27}{2} - 15$

$\frac{3-12}{18} + \frac{27}{2} - 15$

$\frac{-9}{18} + \frac{27}{2} - 15$

$-\frac{1}{2} + \frac{27}{2} - 15$

$\frac{26}{2} - 15$

-2

$f(1) = 1 - \frac{(-2)}{42}$

$= -2 + \frac{2}{42} = -2 + \frac{1}{21}$

$= \frac{-42+1}{21} = \frac{-41}{21}$

$= -1.95$

$f(1) = 1 - \frac{(-2)}{32}$

$= 1 + \frac{2}{32} = -2 + \frac{1}{16} = \frac{-31}{16}$

$= -1.9375$

$= 1 + \frac{1}{16} = \frac{17}{16}$

Gradient descent:

Jacobian is a vector in an algebraic form that gives the maximum/steepest slope of the function for any given x, y, z.

Eg: $f(x,y,z) = x^2y + 3z$

$$\frac{\partial z}{\partial x} = 2xy \qquad \frac{\partial z}{\partial y} = x^2 \qquad \frac{\partial z}{\partial z} = 3$$

$$J = [2xy, x^2, 3]$$

Constrained optimisation method of Lagrange Multipliers

Lagrange Multipliers:

$f(x,y) = xe^{-(x^2+y^2)}$

Let $f(x,y)$ be subject to the constraint $g(x,y) = 0$ and have a relative max/min at (a,b). Then there exists λ (constant) such that the partial derivatives of

$F(x,y,\lambda) = f(x,y) + \lambda g(x,y)$ all equals zero at some point (a,b,c)

$$\frac{\partial F}{\partial x} = 0 \quad \frac{\partial F}{\partial y} = 0 \quad \frac{\partial F}{\partial \lambda} = 0 \quad \text{at } (a,b,c)$$

Example: Maximise the function $f(x,y) = 2xy$ subject to the constraint $x+y=6$

$$F(x,y,\lambda) = f(x,y) + \lambda g(x,y)$$
$$= 2xy + \lambda(x+y-6)$$

$$\frac{\partial F}{\partial x} = 2y+\lambda = 0 \qquad \lambda = -2y$$
$$\frac{\partial F}{\partial y} = 2x+\lambda = 0 \qquad \lambda = -2x$$
$$\Big\} \; x = y$$
$$\frac{\partial F}{\partial \lambda} = x+y-6 = 0$$

$2x = 6$

$x = 3 \quad y = 3$

$f(3,3) = 2.3.3 = 18$

The maximum of $f(x,y) = 2xy$ subject to $x+y=6$ is 18 and occurs at $(3,3)$.

Consider the production function $f(x,y) = 60x^{3/4}y^{1/4}$. Suppose that labour costs 100\$ per unit and capital costs \$200 per unit. Assume that \$30000 is available to spend on costs. How many units of labour and capital will maximise production?

Constraint: $100x + 200y = 30000$

$$F(x,y,\lambda) = f(x,y) + \lambda g(x,y) = 60x^{3/4}y^{1/4} + \lambda(100x + 200y - 30000)$$

$$F_x = 45x^{-1/4}y^{1/4} + 100\lambda = 0 \qquad \lambda = \frac{9}{20}x^{-1/4}y^{1/4}$$
$$F_y = 15x^{3/4}y^{-3/4} + 200\lambda = 0 \qquad \lambda = \frac{3}{40}x^{3/4}y^{-3/4}$$
$$F_\lambda = 100x + 200y - 30000 = 0$$

Maximize $f(x,y) = x^2 y$

Constraint $g(x,y) = x^2 + y^2 = a^2$

$F(x,y,\lambda) = x^2 y + \lambda g(x,y)$

$= x^2 y + \lambda(x^2 + y^2 - a^2)$

$F_x : 2xy + 2\lambda x = 0$

$F_y = x^2 + 2\lambda y = 0$

$F_\lambda = x^2 + y^2 - a^2 = 0$

$F_x \quad 2xy + 2\lambda x = 0$

$2xy = -2\lambda x$

$y = -\lambda$

$F_y \quad x^2 + 2\lambda y = 0$

$x^2 - 2\lambda^2 = 0$

$x^2 = 2\lambda^2$

$x = \pm\lambda\sqrt{2}$

$F_y \quad 2\lambda^2 + \lambda^2 - a^2 = 0$

$a^2 = 3\lambda^2$

$a = \pm\sqrt{3}\lambda$

$\lambda = \pm\frac{a}{\sqrt{3}}$

Alternative way

$\nabla f = \lambda \nabla g$

$\nabla f = \begin{bmatrix} 2xy \\ x^2 \end{bmatrix}$

$\nabla g = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$

$\begin{bmatrix} 2xy \\ x^2 \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix}$

$2xy = \lambda 2x \Rightarrow y = \lambda$

$x^2 = \lambda 2y = 2y^2$

$x^2 + y^2 = a^2 \Rightarrow 2y^2 + y^2 = a^2$

$3y^2 = a^2$

$y^2 = \frac{a^2}{3}$

$y = \pm\frac{a}{\sqrt{3}}$

$x^2 = 2y^2 \Rightarrow x = \pm\sqrt{2}y$

$\Rightarrow x = \pm\sqrt{\frac{2a^2}{3}}$

$x = \pm\sqrt{2}\cdot\frac{a}{\sqrt{3}}$

Solution

$x^2 + y^2 - a^2 = 0$

$2y^2 + \sqrt{\frac{2a^2}{3}}$

$y = \pm\frac{a}{\sqrt{3}}$

$x = \pm\sqrt{2}\cdot\frac{a}{\sqrt{3}}$

$\frac{a}{\sqrt{3}}\begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix}$

$g(x) = x^2 + 3(y+1)^2 - 1 = 0$

$\frac{g(x)}{dx} = 2x$

$y^2 + 2y + 1$

$3y^2 + 6y + 3$

$6y + 6$

$g(x) = x^2 + 3(y+1)^2 - 1 = 0$

$(0.930942)^2 + 3(-1.21083+1)^2 - 1$

$0.8666 + 0.1333 - 1$

$f(x,y) = e^{-(x - y^2 + xy)}$

$\frac{df}{dx} = -1 - ye$

$\frac{dg}{dy} = (2y - x)$

Simple linear regression :

$y = y(x; a_i) = mx_i + c$

$\chi^2 = \sum_i r_i^2 = \sum_i (y_i - mx_i - c)^2$

$(y_i - mx_i - c)^2$

$(y_i - mx_i - c)(y_i - mx_i - c)$

$\frac{\partial \chi^2}{\partial m} = (y_i - mx_i - c)(-x_i) + (y_i - mx_i - c)(-x_i)$

$= -2x_i(y_i - mx_i - c)$

$\frac{\partial \chi^2}{\partial c} = (y_i - mx_i - c)(-1) + (-1)(y_i - mx_i - c)$

$= -2(y_i - mx_i - c)$

$\nabla \chi^2 = \begin{bmatrix} \partial \chi^2 / \partial m \\ \partial \chi^2 / \partial c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2\sum_i x_i (y_i - mx_i - c) \\ -2\sum_i (y_i - mx_i - c) \end{bmatrix}$

$c = \bar{y} - m\bar{x}$

Fitting by eye is

$(x, y) = \{(0.4, 0.1), (0.5, 0.25), (0.6, 0.55), (0.7, 0.75), (0.8, 0.85)\}$

$(0.1 - m \times 0.4 - c)^2$

$(0.25 - m \times 0.5 - c)^2$

$(0.55 - m \times 0.6 - c)^2$

$(0.75 - m \times 0.7 - c)^2$

$(0.85 - m \times 0.8 - c)^2$

$\bar{x} = \frac{3}{5} = .6$

$\bar{y} = \frac{2.5}{5} = .5$

$(0.1 - m \times 0.4 - c)^2$

$\frac{\partial \chi^2}{\partial m} = -2 \times 0.4 (0.1 - m \times 0.4 - c) = -0.8(0.1 - 0.4m - c)$

$\frac{\partial \chi^2}{\partial c} = -2(0.1 - m \times 0.4 - c)$

$(0.25 - 0.5m - c)^2$

$\frac{\partial \chi^2}{\partial m} = -2 \times 0.5 (0.25 - 0.5m - c)$

$= -1(0.25 - 0.5m - c)$

$\frac{\partial \chi^2}{\partial c} = = -2(0.25 - 0.5m - c)$

$m = \sum \frac{(x_i - \bar{x}) y_i}{(x_i - \bar{x})^2}$

$\frac{(0.4 - .6) \cdot 0.1}{(0.4 - .6)^2} = \frac{-.02}{.04} = -0.5$

-0.2×0.1

$\frac{(0.5 - 0.6) \times 0.1}{(0.5 - 0.6)^2} = \frac{-.01}{.01} = -.1$

$\frac{(0.6 - 0.6) \times 0.1}{(0.6 - 0.6)^2} = 0$

$\frac{(.7 - .6) \times 0.1}{(.7 - .6)^2} = \frac{0.01}{0.01} = 1$

$.8 - .6 \quad (.8 - .6)^2 \times$

$$m = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} \qquad c = \bar{y} - m\bar{x}$$

(0.4, 0.1) (0.5, 0.25) (0.6, 0.55) (0.7, 0.75) (0.8, 0.85)

$\bar{x} = 0.6 \qquad \bar{y} = 0.5$

$(0.4-0.6)\,0.1 = -.02 \qquad (0.4-0.6)^2 = .04$

$(0.5-0.6)\,0.25 = -.025 \qquad (0.5-0.6)^2 = .01$

$(0.6-0.6)\,0.55 = 0 \qquad (0.6-0.6)^2 = 0$

$(0.7-0.6)\,0.75 = .075 \qquad (0.7-0.6)^2 = .01$

$(0.8-0.6)\,0.85 = .17 \qquad (0.8-0.6)^2 = .04$

$0.2 \qquad 0.1$

$m = 2 \qquad c = 0.5 - 2 \cdot 0.6$

$= 0.5 - 1.2$

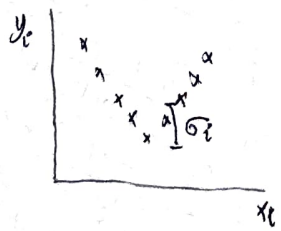
$= -0.7$

General non-linear least squares:

$y(x; a_k)$ where $k = 1 \ldots n$

$y(x; a_k) = (x - a_1)^2 + a^2 \Rightarrow$ we are going to do non linear least squares

(y_i, x_i, σ_i)



$$\chi^2 = \sum_{i=1}^{n} \frac{[y_i - y(x_i; a_k)]^2}{\sigma^2}$$

To minimize χ^2 we have to go in gradient of $\chi^2 = 0$

$$\nabla \chi^2 = 0$$

$$a_{next} = a_{cur} - cst\, \nabla \chi^2 = a_{cur} + \sum_{i=1}^{n} \frac{[y_i - y(x_i; a_k)]}{\sigma^2} \frac{dy}{da_k}$$

$$\frac{dy}{da_1} = -2(x - a_1) \qquad \frac{dy}{da_2} = 1$$

Fitting functions to data x_i $\Longrightarrow$ Learn the techniques

Simple linear regression: Understanding

Two main objectives

1) Establish if there is a relationship between two variables

One variable increases other decreases

Both increase or decrease

$$\chi^2 = \sum_i (y_i - m x_i - c)^2$$

$$\frac{\partial \chi^2}{\partial m} = (y_i - m x_i - c)(y_i - m x_i - c)$$

$$\Rightarrow 2\sum (y_i - m x_i - c) \cdot (-x_i)$$

$$\frac{\partial \chi_2^2}{\partial c} =$$

$$\chi^2 = |y - f(x;\mu,\sigma)|^2$$

$$\frac{\partial \chi^2}{\partial \mu} = (y - f(x;\mu,\sigma))(y - f(x;\mu,\sigma))$$

$$= (y - f(x;\mu,\sigma)) \cdot -\frac{\partial f}{\partial \mu}(x;\mu,\sigma) + -\frac{\partial f}{\partial \mu}(x;\mu,\sigma)(y - f(x;\mu,\sigma))$$

$$= -2(y - f(x;\mu,\sigma) \cdot \frac{\partial f}{\partial \mu}(x;\mu,\sigma)$$

$$\frac{\partial \chi^2}{\partial \sigma} = -2(y - f(x;\mu,\sigma) \cdot \frac{\partial f}{\partial \sigma}(x;\mu,\sigma)$$

$$f = \frac{e^{\left(\frac{-(x-\mu)^2}{2\cdot\sigma^2}\right)}}{\sigma\sqrt{2\pi}}$$

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$\frac{\partial f}{\partial \mu}(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \times$$

$$-2(x-\mu)(x-\mu)\cdot\sigma^{-2}$$

$$\frac{\partial f}{\partial \mu} \quad \cancel{-2\sigma^{-2}(x-\mu)}(x-\mu)$$

$$\sigma^{-2} - 2(x-\mu)(x)$$

$$h(g(x)) = \frac{\partial h}{\partial g} \cdot \frac{\partial g}{\partial x}$$

Use chain rule

$$f = e^v \qquad v = -\frac{(x-\mu)^2}{2\sigma^2}$$

$$\frac{\partial f}{\partial v} = e^v \qquad \frac{\partial v}{\partial \mu} = -\frac{(x-\mu)^2}{2\sigma^2}$$

$$= \left(\frac{x-\mu}{\sigma^2}\right)$$

$$= \frac{-1}{2\sigma^2}(x-\mu)^2$$

$$= \frac{-1}{2\sigma^2}(x-\mu)(x-\mu)$$

$$= \frac{-1}{2\sigma^2}\left((x-\mu)(-1) + (x-\mu)(-1)\right)$$

$$= \frac{-1}{2\sigma^2}(-2(x-\mu))$$

$$\frac{x-\mu}{\sigma^2}$$

$$\frac{1}{\sigma\sqrt{2\pi}}\left(\frac{x-\mu}{\sigma^2}\right) \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$= \frac{(x-\mu)}{\sigma\sqrt{2\pi}\cdot\sigma^2} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$= \frac{x-\mu}{\sigma^3\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)}{2\sigma^2}\right)$$

$$\frac{\partial z}{\partial \sigma} = \frac{-1}{\sigma^2\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) + \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$= \frac{-1}{\sigma^2\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) + \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{(x-\mu)^2}{\sigma^3}$$

$$\frac{\partial}{\partial \sigma} = \frac{-(x-\mu)^2}{2} \cdot \sigma^{-2} = -\frac{(x-\mu)^2}{2} \cdot -2\sigma^{-3} = \frac{2(x-\mu)^2}{2\sigma^3} = \frac{(x-\mu)^2}{\sigma^3}$$

$$= \frac{-1}{\sigma^2\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) + \frac{(x-\mu)^2}{\sigma^4\sqrt{2\pi}}$$

chain rule

$$f(x;\mu,\sigma) = \underbrace{\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}_{\text{product rule}}$$

$$\frac{\partial f}{\partial \sigma} \Rightarrow$$

$$\frac{\sigma^{-1}}{\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2} \cdot \sigma^{-2}\right) - \frac{\sigma^{-2}}{\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2} \cdot \sigma^{-2}\right) + \frac{1}{\sigma\sqrt{2\pi}} \frac{\partial}{\partial \sigma}\left(\exp\left(-\frac{(x-\mu)^2\sigma^{-2}}{2}\right)\right)$$

$$\frac{\partial}{\partial \sigma} \exp\left(-\frac{(x-\mu)^2 \sigma^{-2}}{2}\right)$$

$f(g(\sigma))$

$$\frac{\partial z}{\partial g} \cdot \frac{\partial g}{\partial \sigma}$$

$$\frac{\partial}{\partial \sigma} \exp^{v}$$

$$= \frac{\partial}{\partial v} \exp^{v} \frac{\partial v}{\partial \sigma} v$$

$$\frac{\partial v}{\partial \sigma}(v) = \frac{\partial}{\partial \sigma}\left(\frac{-(x-\mu)^2\sigma^{-2}}{2}\right)$$

$$= \frac{-1}{2}\frac{\partial}{\partial \sigma}\left((x-\mu)^2\sigma^{-2}\right)$$

$$= \frac{-1}{2} \times \frac{(x-\mu)^2 \cdot -2 \cdot \sigma^{-3}}{}$$

$$= \frac{(x-\mu)^2}{\sigma^3}$$

$$\Rightarrow \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \frac{(x-\mu)^2}{\sigma^3}$$

Finally

$$-\frac{1}{\sqrt{2\pi}\,\sigma^2} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) + \frac{1}{\sigma\sqrt{2\pi}} \times \frac{(x-\mu)^2}{\sigma^3} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$\Rightarrow \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\left(-\frac{1}{\sqrt{2\pi}\,\sigma^2} + \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{(x-\mu)^2}{\sigma^3}\right)$$

$$\Rightarrow \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi}\,\sigma}\left(\frac{-1}{\sigma} + \frac{(x-\mu)^2}{\sigma^3}\right)$$

$$\Rightarrow \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi}\,\sigma}\left(\frac{-\sigma^2 + (x-\mu)^2}{\sigma^3}\right)$$

$$\Rightarrow \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi}\,\sigma}\left(\frac{(x-\mu)^2-\sigma^2}{\sigma^3}\right)$$

$$\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \cdot \left(\frac{(x-\mu)^2-\sigma^2}{\sqrt{2\pi}\,\sigma^4}\right)$$

This can be written as

$$\underbrace{\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}_{} \cdot \left(\frac{(x-\mu)^2-\sigma^2}{\sigma^3}\right)$$

$$= f(x;\mu,\sigma)\left(\frac{(x-\mu)^2-\sigma^2}{\sigma^3}\right)$$

PCA – Machine learning:

1. What is the variance of the following data set

$D = \{1, 2, 3, 2\}$

Variance is Σ of square of the distance between the data points and the mean divided by the number of data points

$$\mu = \frac{1+2+3+2}{4} = \frac{8}{4} = 2$$

$$\sigma^2 = \frac{(1-2)^2 + (2-2)^2 + (3-2)^2 + (2-2)^2}{4}$$

$$= \frac{1+0+1+0}{4} = \frac{1}{2} = 0.5$$

$\sigma \Rightarrow$ standard deviation: $\sqrt{0.5}$

What would be the new variance if we added 1 to each element in the dataset $D = \{1, 2, 3, 2\}$ from question 1?

New $D = \{2, 3, 4, 3\}$

$$\mu = \frac{2+3+4+3}{4} = \frac{12}{4} = 3$$

$$\sigma^2 = \frac{(2-3)^2 + (3-3)^2 + (4-3)^2 + (3-3)^2}{3}$$

$$= \frac{1+0+1+0}{3} = \frac{2}{3}$$

SD =

$D = \{1, 2, 3, 2\}$ $\quad \mu \; \frac{1+2+3+2}{4} = 2$

$$\sigma^2 = \frac{(1-2)^2 + (2-2)^2 + (3-2)^2 + (2-2)^2}{4} = \frac{2}{4} = \frac{1}{2}$$

What would be the new variance if we multiplied each sample in a dataset by 2

$D = \{2, 4, 6, 4\}$

$$\mu = \frac{2+4+6+4}{4} = \frac{16}{4} = 4$$

$$\sigma^2 = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2 + (4-4)^2}{4}$$

$$= \frac{4+0+4+0}{4} = 2$$

$SD = \sqrt{2}$

Assume we have mean $\bar{x}_{n-1}$ and variance σ^2_{n-1} for some dataset D_{n-1} with $n-1$ samples. What would be the variance σ^2_n if we add a new elem x_n to the dataset

$D_{n-1} = \{2, 4, 6, 4\}$ $\quad n-1 = 4$

$\mu = 4 \quad \sigma^2 = 2$

$D_n = \{2, 4, 6, 4, 4\}$ $\quad n = 5$

$$\mu = \frac{2+4+6+4+4}{5} = \frac{20}{5} = 4$$

$$\sigma^2 = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2 + (4-4)^2 + (4-4)^2}{5}$$

$$= \frac{4+0+4+0+0}{5} = \frac{8}{5} = 1.6$$

$$= \frac{4}{5}\cdot 2 + \frac{1}{5}(4-4)(4-4)$$

$$= \frac{8}{5} + 0$$

$$\frac{4}{5}\cdot 2 + \frac{1}{4}(4-4$$

$D_{n-1} = \{2, 4, 6, 4\}$

$\mu_{n-1} = 4 \quad \sigma^2 = 2$

$D_n = \{2, 4, 6, 4, 5\}$

$$\mu = \frac{2+4+6+4-5}{5} = \frac{21}{5} = 4.2$$

$$\sigma^2 = \frac{(2-4.2)^2 + (4-4.2)^2 + (6-4.2)^2 + (5-4.2)^2 + (4-4.2)^2}{5}$$

$$= \cancel{(-1.8)^2 + \ldots (-2.2)^2 + (-0.2)^2 + (1.8)^2 + 0.8}$$

$$\frac{(2-4.2)^2 + (4-4.2)^2 + (6-4.2)^2 + (4-4.2)^2 + (5-4.2)^2}{5}$$

$$= \frac{(-2.2)^2 + (-0.2)^2 + (2.2)^2 + (.8)^2 + (-0.2)^2}{5}$$

$$(9.68 + 0.08 + 0.64)/5$$

$$= 2.08$$

$$\frac{4}{5}\cdot 2 + \frac{1}{5}(4.2-4)(4.2-5)$$

$$\frac{8}{5} + \frac{1}{5}(2.2)(-0.8)$$

$$1.6 - \frac{1}{5}(1.76)$$

$$1.6 - \frac{1.76}{5}$$

$$= 1.248$$

$$\frac{n-1}{n}\sigma^2_{n-1} + \frac{1}{n}(x_n - \bar{x}_{n-1})(x_n - \bar{x}_n)$$

$$\frac{4}{5}\cdot 2 + \frac{1}{5}(4.2-4)^2$$

$$\frac{8}{5} + \frac{4.84}{5}$$

$$= \frac{12.84}{5} = 2.568$$

$$\frac{n-1}{n}\sigma^2_{n-1} + \frac{1}{n}(x_n - \bar{x}_{n-1})^2$$

$$\frac{4}{5}\cdot 2 + \frac{1}{4}(4.2-4)(4.2-5)$$

$$\frac{8}{5} + \frac{.2 \cdot -(0.8)}{4}$$

$$= \frac{8}{5} - \frac{0.16}{4} = 1.56$$

$$\frac{n-1}{n}\sigma^2_{n-1} + \frac{1}{n-1}(x_n - \bar{x}_{n-1})(x_n - \bar{x}_n)$$

$$\frac{6}{4}\cdot 2 + \frac{1}{5}(4.2-4)(4.2-5)$$

$$3 - \frac{0.16}{4}$$

Definiteness of a matrix :

Not to be confused with positive matrix and totally positive matrix

In linear algebra, a symmetric $n \times n$ real matrix M is said to be positive definite if the scalar $z^T M z$ is strictly positive for every non-zero column vector z of n real numbers. Here z^T denotes the transpose of z

Let

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

Let $z = \begin{bmatrix} a & b & c \end{bmatrix}$

Then $z^T = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

$$\Rightarrow z^T M z \Rightarrow \begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} a & b & c \end{bmatrix}$$

More generally a complex $n \times n$ Hermitian matrix M is said to be positive definite if the scalar $z^* M z$ is automatically real since M is Hermitian.

Eigen Values and Eigen vectors

Let A be a $n \times n$ matrix. A scalar λ is called an eigenvalue of A if there is a non zero vector $\vec{x}$ such the $A\vec{x} = \lambda \vec{x}$. Such a vector $\vec{x}$ is called an eigen vector of A corresponding to λ.

$$\vec{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad A = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix} \qquad \lambda = 4$$

$$\begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

Then $\lambda = 4$ is an Eigen value

$\vec{x}$ is the eigen vector to matrix A corresponding to scalar value $\lambda = 4$

If λ is an eigen value of A and $\vec{x}$ is an eigen vector belonging to λ, Any non-zero multiple of $\vec{x}$ will be an eigen vector

Positive semi definite matrices are defined similarly except that the scalar $z^T M z$ or $z^* M z$ must be positive or zero (i.e non-negative).

Negative definite and negative semi-definite are defined analogously.

A matrix that is not positive semi-definite and not negative semi-definite is called indefinite

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ be an $n \times n$ matrix

and let $D_i = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{bmatrix}$ for $i = 1, 2 \ldots n$ Then

a) A is said to be positive definite if $D_i > 0$ for $i = 1, 2 \ldots n$

b) A is said to be negative definite if $D_i < 0$ for odd $i \in (1, 2 \ldots n)$ and $D_i > 0$ for even $i \in (1, 2 \ldots n)$

c) A is said to be indefinite if $\det(A) = D_n \neq 0$ and neither a) nor b) hold.

d) If $\det(A) = D_n = 0$, then A may be indefinite or what is known positive semidefinite or Negative semidefinite.

The values D_i for $i = 1, 2 \ldots n$ are the values of the determinants of the $i \times i$ top left submatrices of A

Note that $D_1 = a_{11}$, $D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ etc

Definition for real matrices

A $n \times n$ symmetric real matrix M is said to be positive definite if $x^T M x > 0$ for all non-zero x in R^n

M positive definite $\Leftrightarrow$ $x^T M x > 0$ for all $x \in R^n \setminus 0$

A $n \times n$ symmetric real matrix M is said to be positive semi-definite or non-negative definite if $x^T M x \geq 0$ for all non-zero x in R^n. Formally

M positive semi-definite $\Leftrightarrow$ $x^T M x \geq 0$ for all $x \in R^n \setminus 0$

A $n \times n$ symmetric real matrix M is said to be negative definite if $x^T M x < 0$ for all non-zero x in R^n. Formally

M negative definite $\Leftrightarrow$ $x^T M x < 0$ for all non-zero x in $R^n \setminus 0$

A $n \times n$ symmetric real matrix M is said to be negative semi-definite if $x^T M x \leq 0$ for all non-zero x in R^n. Formally

M negative semidefinite $\Leftrightarrow$ $x^T M x \leq 0$ for all $x \in R^n \setminus 0$

A $n \times n$ symmetric real matrix which is neither positive semi-definite nor negative semidefinite is called indefinite.

Definition for Complex matrices:

A $n \times n$ Hermitian Complex matrix M is positive definite if M positive definite $\iff x^* M x > 0$ for all $x \in \mathbb{C}^n \setminus 0$

A $n \times n$ Hermitian Complex matrix M is positive semi definite if M positive semi-definite $\iff x^* M x \geq 0$ for all $x \in \mathbb{C}^n \setminus 0$

A $n \times n$ Hermitian Complex matrix M is negative definite if M ~~positive~~ negative definite $\iff x^* M x < 0$ for all $x \in \mathbb{C}^n$

A $n \times n$ Hermitian Complex matrix M is negative semi definite if M negative semi definite $\iff x^* M x \leq 0$ for all $x \in \mathbb{C}^n \setminus 0$

A $n \times n$ Hermitian matrix which is neither positive semi definite nor negative semi definite is called indefinite

Consistency between Real and Complex definitions:

Since every real matrix is also a Complex matrix, the definition of definiteness for the two classes must agree. For Complex matrices, the most common definition says that M is positive definite if and only if $z^* M z$ is real and positive for all non-zero Complex column vectors z.

To see this consider the matrices $A = \frac{1}{2}(M + M^*)$ and $B = \frac{1}{2i}(M - M^*)$ so that $M = A + iB$ ~~and $z^* A z$~~ and $z^* M z = z^* A z + i z^* B z$. The matrices A and B are Hermitian. Therefore $z^* A z$ and $z^* B z$ are individually real. If $z^* A z$ is real then $z^* B z$ must be zero for all z. Then B is the zero matrix

Hermitian Matrix

A Hermitian matrix (or self-adjoint matrix) is a complex square matrix that is equal to its own conjugate transpose - that is, the element in the i-th row and j-th column is equal to the complex conjugate of the element

Covariance

$$\mathrm{Cov}[x, y] = E[(x - \mu_x)(y - \mu_y)]$$

$$\mu_x = E[x]$$

$$\mu_y = E[y]$$

Therefore for 2D data we get 4 quantities of interest

$\mathrm{Var}[x]$

$\mathrm{Var}[y]$

$\mathrm{Cov}[x, y]$

$\mathrm{Cov}[y, x]$

$$\begin{bmatrix} \mathrm{Var}[x] & \mathrm{Cov}[x, y] \\ \mathrm{Cov}[y, x] & \mathrm{Var}[y] \end{bmatrix}$$

Covariance Matrix: Summarizes the 4 quantities of interest of a 2D data set

If the Covariance of x & y is positive then an average y value increases if we increase x and the Covariance is negative when on an average y value decreases when x increases

If the Covariance between x & y is zero then x and y are uncorrelated

The covariance matrix is always symmetric positive definite matrix.

Eg

$$D = \{x_1, x_2 \dots x_n\}, \; x_i \in \mathbb{R}^D$$

$$Var[D] = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T$$

Exercise:

Compute the covariance matrix for the following dataset

$$D = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix} \right\}$$

Here every column vector represents a data point

$$\cancel{\mu_x = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \quad \mu_y = \quad \Rightarrow \mu_x = 3}$$

wrong

$\mu_y = 3$

Variance - Covariance Matrix:

Variance is a measure of the variability or spread in a set of data. Mathematically, it is the average squared deviation from the mean square. We use the following formula to compute variance.

$$Var(x) = \sum (x_i - \bar{x})^2 / N = \sum x_i^2 / N$$

where

N is the number of scores in a set of scores

$\bar{x}$ is the mean of N scores

X_i is the ith raw score in the set of scores

x_i is the ith deviation score in the set of scores

$Var(X)$ is the variance of all the scores in the set

Covariance:

Covariance is a measure of the extent to which corresponding elements from two sets of ordered data move in the same direction. We use the following formula to compute Covariance

$$Cov(x, y) = \sum (x_i - \bar{x})(y_i - \bar{y}) / N = \sum x_i y_i / N$$

Exercise:

Compute the covariance matrix for the following data set

$$D = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix} \right\}$$

Here every column vector represents a data point.

$$\mu_x = \frac{1+5}{2} = \frac{6}{2} = 3$$

$$\mu_y = \frac{2+4}{2} = \frac{6}{2} = 3$$

$$\sigma_x^2 \Rightarrow var[x] = \frac{(1-3)^2 + (5-3)^2}{2} = 4$$

$$var[y] = \frac{(2-3)^2 + (4-3)^2}{2} = \frac{2}{2} = 1$$

Covariance Matrix

$$\begin{bmatrix} var[x] & cov[x,y] \\ cov[y,x] & var[y] \end{bmatrix}$$

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

$$cov[x,y] = [(x - \mu_x)(y - \mu_y)] = (1-3)(2-3) = -2 \times -1 = +2$$

$$cov[y,x] = (5-3)(4-3) = 2 \times 1 = 2$$

2) Consider a data set D with covariance matrix $\begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}$. What is the covariance matrix if we multiply every vector in D by 2?

$$\begin{bmatrix} Var[x] & cov[x,y] \\ cov[y,x] & Var[y] \end{bmatrix}$$

$Var[x] = 3$, $cov[x,y] = 2$, $cov[y,x] = 2$, $Var[y] = 4$

$$\sigma_x^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

$$\sigma = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

Assume $x = 3$ and $\mu = 0$

$(3-0)^2 = 9$

$(6-0)^2 = 36$

if in 2 it becomes 4

$$cov[x,y] = E\left[(x - \mu_x)(y - \mu_y)\right]$$

if in 2 it becomes 4

3) Consider the data set $D = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ 4 \end{bmatrix} \right\}$ with covariance matrix $\begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}$. Compute the new covariance matrix when we add $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ to each element in D.

$$D = \left\{ \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 9 \\ 6 \end{bmatrix} \right\}$$

$\mu_x = 6$ $\mu_y = 5$

$$Var[x] = \frac{(3-6)^2 + (9-6)^2}{2} = 9$$

$$Var[y] = \frac{(4-5)^2 + (6-5)^2}{2} = \frac{2}{2} = 1$$

4. Provide a valid 2x2 covariance matrix by replacing the ? entries in the table below

5) We are looking at a data set D where every element in D consists of an x and y coordinate. The data covariance matrix is given by

$$\begin{bmatrix} 1 & 0.5 \\ 0.8 & 1 \end{bmatrix}$$

Which of the following statement is correct

Linear Transformation of Data sets.

Rough work

Effect on the mean.

$$E[D + a] = a + E[D]$$

If we shift/add the data set by a constant "a" then the mean also shifts by "a".

$x = \{1, 2, 3\}$

$\mu_x = \frac{6}{3} = 2$

$x' = \{3, 4, 5\}$

$\mu_{x'} = 4$

$x = \{1, 2, 3\}$

$\mu_x = 2$

$x' = \{2, 4, 6\}$

$= 4$

$$E[\alpha D] = \alpha \cdot E[D]$$

If we stretch/multiply the data set by a constant "α" then the mean also stretches by "α" times

$$E[\alpha D + a] = \alpha E[D] + a$$

Effect of linear transformation on variance:

Variance is the measure of spread of the data

$Var[D] = Var[D+a]$

When the mean is shifted there is no change in variance.

Since variance is the average squared distance of the data points from the mean. If we scale the data set by a factor 2, the distance of every data point to the mean is scaled by 2 and the squared distance is scaled by 4

$Var[\alpha . D] = \alpha^2 Var[D]$

For higher dimensional data-set like the one below

$D = \{x_1, \dots, x_N\} \quad x_i \in R^p$

The variance of the above data-set is given by a Covariance matrix.

If we perform a linear transformation of every data point

$Ax_i + b$ we get

$Var[AD+b] = A\, Var[D]\, A^T$

Shifting data has an effect only on mean and not the variance. Scaling/stretching/multiplying data has an effect on both mean + variance

Affine transformations:

When working with some datasets, before digging into further analysis, it is almost always useful to do a few things to understand your dataset

First of all, answer the following set of questions

1) What is the size of your data?
2) What is the dimensionality of your data?

The dataset we have are usually stored as 2D matrices, then it would be really important to know which dimension represents the dimension of the dataset, and which represents the datapoints in the dataset.

When you implement the functions for your assignment make sure you read the directory for what each dimension of your inputs represents the datapoints and which represents the dimensions of the dataset.

Find Covariance matrix

$X = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix}$

$a = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} \times \frac{1}{2}$ Should be $\frac{1}{3}$ & a should give the Covariance matrix

This is wrong

$= \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} - \begin{bmatrix} 2 & 5 & 7 \\ 3 & 5 & 7 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \end{bmatrix}$

$= \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} - \begin{bmatrix} 1.5 & 2.5 & 3.5 \\ 1.5 & 2.5 & 3.5 \end{bmatrix}$

$= \begin{bmatrix} -1.5 & -1.5 & -1.5 \\ 1.5 & 1.5 & 1.5 \end{bmatrix}$

$A \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix}$ mean $= \begin{bmatrix} 1 \\ 4 \end{bmatrix}$

$\mu_x = \frac{5}{2} = 1.5$ $\quad \mu_y = \frac{5}{2} = 2.5$ $\quad \mu_z = \frac{7}{2} = 3.5$

$Cov[x,y] = \sum (x-\bar{x})(y-\bar{y})/N$

Covariance Matrix: $\begin{bmatrix} Var[X] & Cov[x,y] \\ Cov[y,x] & Var[Y] \end{bmatrix}$

$A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix}$ Mean $= \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ Row wise

Mean $= [1.5 \quad 2.5 \quad 3.5]$ column wise

σ^2_x (or) $Var[X] = \frac{(0-1)^2 + (1-1)^2 + (2-1)^2}{3} = \cancel{\frac{1+0+4}{3} = \frac{5}{3}}$ $\quad \frac{2}{3} = 0.666666$

σ^2_y $Var[Y] = \frac{(3-4)^2 + (4-4)^2 + (5-4)^2}{3} = \frac{1+0+1}{3} = \frac{2}{3} = 0.666666$

$Var[A] = \begin{bmatrix} 0.6666 \\ 0.6666 \end{bmatrix}$

$Cov[xy] = \frac{(A[0][0]-M[0])(A[1][0]-M[1]) + (A[0][1]-M[0])(A[1][1]-M[1]) + (A[0][2]-M[0])(A[1][2]-M[1])}{3}$

$= \frac{(0-1)(3-4) + (1-1)(4-4) + (2-1)(5-4)}{3}$

$= \frac{1+0+1}{3} = 0.6666$

$Cov[yx] = \frac{(3-4)(0-1) + (4-4)(1-1) + (5-4)(2-1)}{3}$

$= 0.6666$

$\Longrightarrow \begin{bmatrix} 0.6666 & 0.6666 \\ 0.6666 & 0.6666 \end{bmatrix}$

Python code to calculate Covariance

$A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix}$ $M = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$

A.shape = (2, 3)

D, N = A.shape

for i in range(D):
  for j in range(D):
    ~~A[i][j]~~
    ~~for k in range(N)~~
    (A[i][j] - M[i])

for k in range(N):
  (A[i][k] - M[i])

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

$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix}$

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

$\Rightarrow \begin{matrix} -2 \\ 0 \\ 2 \end{matrix}$

A[0] = [0 1 2]

A[1] = [3 4 5]

for n in range(N):
  for i in range(D):
    for j in range(D)
      if (i==j): Cov[i][i] = Var[i]
      else:
        Cov[i][j] = (A[i][n] - M[i]) * [A[j][n] - M[j]]

Cov[i][j]

Cov[i][i]

for n in range(N):
  for i in range(D):
    for j in range(D):
      if i==j:
        Cov[i][j] = Var[i]
      else:
        Cov[i][j] = Cov[i][j] + (A[i][n] - M[i]) * [A[j][n] - M[j]]

$\begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix}$

Cov[0][0] (A[0][0] - M[0])(A[0][0] - M[0])

n = 0
i = 0 j = 0
i = 0 j = 1 skip
Cov[0][0] = A[0][0] A[0][0]

PTO → Not correct yet

This works ✓

```
for i in range(D):
  for j in range(D):
    for k in range(N):
      If j==i: cov[i][j] = Var[i]
      else:
        cov[i][j] = cov[i][j] + (A[i][k] - μ[i]) * (A[j][k] - μ[j]) * 1/N
```

Week 2:

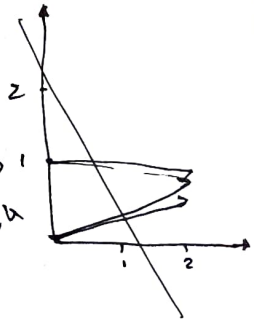
In order to measure lengths, angles and distance we need to equip vectors with inner products.

An eg of inner product is a dot product of two vectors.

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

The length of $\|x\| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{N} x_i^2}$

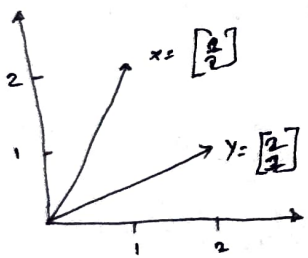
Eg



$x^T y$

$[1\ 2]\begin{bmatrix}2\\1\end{bmatrix}$

$2+2=4$



Length

$\|x\| = \sqrt{1^2+2^2} = \sqrt{5}$

$\|y\| = \sqrt{2^2+1^2} = \sqrt{5}$

Distance $d(x,y) = \|x-y\| = \sqrt{(x-y)^T(x-y)}$

$\cos\alpha = \dfrac{x^T y}{\|x\|\,\|y\|}$

1) Compute the angle $x = \begin{bmatrix}3\\4\end{bmatrix}$ $y = \begin{bmatrix}-1\\-1\end{bmatrix}$ ✗ wrong

Angle between x + y $\quad \cos\alpha = \dfrac{x^T y}{\|x\|\|y\|} = \dfrac{-7}{\sqrt{50}}$ ✗ wrong

2) Compute the distance $x = \begin{bmatrix}3\\4\end{bmatrix}$ $y = \begin{bmatrix}-1\\-1\end{bmatrix}$ $\quad \|x-y\| = \sqrt{16+25} = \sqrt{41}$ ✗ wrong

3) $x = \begin{bmatrix}1\\2\\3\end{bmatrix}$ $y = \begin{bmatrix}-1\\0\\8\end{bmatrix}$ Compute the angle (in rad) between x and $x-y$

$x-y = \begin{bmatrix}2\\2\\-5\end{bmatrix}$

$x^T(x-y) = [1\ 2\ 3]\begin{bmatrix}2\\2\\-5\end{bmatrix} = -11$

$\dfrac{-11}{\sqrt{14}\cdot\sqrt{33}} = \dfrac{-11}{21.4941}$ ✗ wrong

-0.5117

$x = \begin{bmatrix}3\\4\end{bmatrix}$ $y = \begin{bmatrix}1\\-1\end{bmatrix}$ $\quad \cos\alpha = \dfrac{x^T y}{\|x\|\|y\|} = \dfrac{-12}{\sqrt{50}} = -1.6970$

Distance

$\|x-y\| = \left\|\begin{bmatrix}2\\5\end{bmatrix}\right\| = \sqrt{2^2+5^2} = \sqrt{4+25} = \sqrt{29}$

$\begin{bmatrix}5\\4\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}$

$\begin{bmatrix}2\\3\end{bmatrix} = \sqrt{4+9} = \sqrt{13} = 3.6055$

$x = \begin{bmatrix}1\\2\\3\end{bmatrix}$ $y = \begin{bmatrix}-1\\0\\8\end{bmatrix}$ $\quad x-y = \begin{bmatrix}2\\2\\-5\end{bmatrix}$

$\begin{bmatrix}2\\5\end{bmatrix} = \sqrt{29}$

$\cos\alpha = \dfrac{x^T(x-y)}{\|x\|\,\|x-y\|} = \dfrac{-9}{\sqrt{14}\cdot\sqrt{33}} = \dfrac{-9}{21.4941} = -0.4189$

$\alpha = -\cos^{-1}(0.4189)$

$x = \begin{bmatrix}3\\4\end{bmatrix}$ $y = \begin{bmatrix}-1\\-1\end{bmatrix}$ $\quad \cos\alpha = \dfrac{x^T y}{\|x\|\|y\|} = \dfrac{-7}{\sqrt{50}} = -0.2899$

$x = \begin{bmatrix}3\\4\end{bmatrix}$ $y = \begin{bmatrix}1\\-1\end{bmatrix}$ $\Rightarrow x-y = \begin{bmatrix}2\\5\end{bmatrix} = \sqrt{2^2+5^2} = \sqrt{29} = 5.5851$

$x = \begin{bmatrix}3\\4\end{bmatrix}$ $y = \begin{bmatrix}1\\-1\end{bmatrix}$ $\quad \dfrac{-1}{\sqrt{25}\sqrt{2}} = \dfrac{-1}{\sqrt{50}}$

$\cos\alpha = \dfrac{-1}{\sqrt{50}}$ $\quad \alpha = \cos^{-1}\left(\dfrac{-1}{\sqrt{50}}\right) = -\cos^{-1}\left(\dfrac{1}{\sqrt{50}}\right) = -\cos^{-1}(0.1414)$

Inner product:

An inner product is to find symmetric positive definite bilinear mapping

The function $\beta(x,y) = x^T \begin{bmatrix}2 & -1\\-1 & 1\end{bmatrix} y$

Test for bilinear if $\langle x,y,z\rangle$ $\lambda \in \mathbb{R}$

$\langle \lambda x+y, z\rangle = \lambda\langle x,z\rangle + \langle y,z\rangle$

$\langle x, \lambda y+z\rangle = \lambda\langle x,y\rangle + \langle x,z\rangle$

Bilinear Transformation

$w = \frac{az+b}{cz+d}$, $\cancel{ad-bc \neq 0}$ is called bilinear transformation.

Useful

Definition : Bilinear form

Let X be a vector space over R. We call bilinear form on X a function $a : X \times X \to R$ such that for all u, v, w in X and λ, μ in R, we have

(i) $a(\lambda u + \mu v, w) = \lambda a(u,w) + \mu a(v,w)$

ii) $a(u, \lambda v + \lambda w) = \lambda a(u,v) + \mu a(u,w)$

It is symmetric if $a(u,v) = a(v,u)$ for all u,v in X

It is positive if $a(u,u) \geq 0$ for all u in X

It is definite positive if $a(u,u) > 0$ for all u in $X \setminus \{0\}$

Exercise

$X = R^2$ let $a([x_1, x_2], [y_1, y_2]) = 2x_1y_1 - x_2y_2$

Bilinear Maps and Scalar product : (Another explanation)

Def. Let $b : V \times V \to F$ where V is vector space over F

A bilinear map is linear in both entry.

1) $b(cx+y, z) = c\,b(x,z) + b(y,z)$

$b(cx+y, z) = b(cx, z) + b(y, z)$

2) $b(x, cy+z) = c\,b(x,y) + b(x,z)$

$b(x, cy+z) = b(x, cy) + b(x, z)$

for all $x, y, z \in V$ and $c \in F$

(a) If $b(x,y) = b(y,x)$ $\forall$ $x, y \in V$ then b is symmetric

(b) If $b(x,y) = -b(y,x)$ for all $\forall x, y \in V$ then b is ~~symmetric~~ anti symmetric

b is a bilinear form on V. b is an F value

$V \times V$ where V is R

$\text{function}(cx+y, z) = \cancel{\text{function}(cx+cy)}$

$\text{function}(cz x + z y)$

$= c\,\text{function}(xz) + \text{function}(zy)$

$\text{function}(x, cy+z) = c\,\text{function}(xy) + \text{fun}(xz)$

Let $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $y = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ $z = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ $c = 2$

$\text{fun}(cx+y, z) \Rightarrow (cx+y) \times z$

$= \left(2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix}\right) \times \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$= \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} \times \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$= \begin{bmatrix} 5 \\ 8 \end{bmatrix} \times \begin{bmatrix} 5 \\ 6 \end{bmatrix} = 25 + 48 = 73$

$c\,\text{function}(xz) + \text{fun}(yz)$

$2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$= 34 + 39 = 73$ Hence proved

$\text{fun}(x, cy+z) = 2 \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$= 22 + 17 = 39$

$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \times \left(2 \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 5 \\ 6 \end{bmatrix}\right)$

$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \times \begin{bmatrix} 6 \\ 8 \end{bmatrix} + \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \times \begin{bmatrix} 11 \\ 14 \end{bmatrix} = 39$

Hence proved

If $\gamma = \{v_1, v_2, \dots v_n\}$ is basis for V then $\exists M \in F^{n\times n}$

there exists

such that $b(x,y) = \underset{(1\times n)}{[x]_\gamma^T} \underset{(n\times n)}{M} \underset{(n\times 1)}{[y]_\gamma} : 1\times 1$

$\forall x, y \in V$. Moreover $M_{ij} = b(v_i, v_j) \ \forall i, j$

proof. $b(x,y) = b\left(\sum_{i=1}^{n} x^i v_i, \sum_{j=1}^{n} y^j v_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i y^j b(v_i, v_j)$

If $x = x^1 v_1 + \dots x^n v_n$ then $[x]_\gamma = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}$ for $x \in V$ with

basis $\gamma = \{v_1, v_2 \dots v_n\}$

$\sum_{i,j=1}^{n} x^i M_{ij} y^j$

Exercise:

Properties of inner products

$$\beta(x,y) = x^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} y$$

Bilinear Test is

$f(x+y, z) \to c f(x,z) + f(y,z)$

1st $f(x, c(y+z)) \to c f(y,x) + f(x,z)$

So this is bilinear

$\beta(x,y) \to$

$X = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ $Y = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

let $X = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$[3\ 3] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

$X = \overset{c=2}{\overset{a}{\begin{bmatrix} 2 \\ 4 \end{bmatrix}}} \overset{b}{\begin{bmatrix} 4 \\ 8 \end{bmatrix}}$

Assume $X = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \to \underset{x}{\begin{bmatrix} 2 \\ 2 \end{bmatrix}} + \underset{y}{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}$

let $X =$

let $Y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow X = 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}$

From the previous page:

Definition: Bilinear form

~~Ex 2~~

let $X = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $Y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$X^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} Y$

$[8\ 16] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\underset{1\times 2}{[0\ 8]} \underset{2\times 2}{\begin{bmatrix} 6 \\ 12 \end{bmatrix}} = 0\times 6 + 8\times 12 = 96$

X: let us express as $ca + b$

$\begin{bmatrix} 8 \\ 16 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}$ $Y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$f(x,y) = x^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} y$

$\left(2[2\ 4] + [4\ 8]\right) \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 96$

$\Rightarrow 2 f\left(\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \end{bmatrix}\right) + f\left(\begin{bmatrix} 4 \\ 8 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \end{bmatrix}\right)$

$\Rightarrow 2 [2\ 4] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [4\ 8] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow [4\ 8] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [0\ 4] \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow [0\ 4] \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow 48 + 48 = 96$

Previous page

$X = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $Y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$ let $Y = 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$\Rightarrow f(x, cy+z) = cf(y,x) + f(x,z)$

$f(x,y) = X^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} Y$

$= [8 \ 16] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \times \left(2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right)$

$= [8 \ 16] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \times \left(\begin{bmatrix} 4 \\ 8 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right)$

$= [0 \ 8] \times \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 0 + 96 = 96$

$2 f\left(\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 8 \\ 16 \end{bmatrix}\right) + f\left(\begin{bmatrix} 8 \\ 16 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right)$

$2 \cdot \begin{bmatrix} 2 \\ 4 \end{bmatrix}^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} + [8 \ 16] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$[4 \ 8] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} + [8 \ 16] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$[0 \ 4] \begin{bmatrix} 8 \\ 16 \end{bmatrix} + [0 \ 8] \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$64 + 32 = 96$

Hence bilinear

is $\beta(x,y) = X^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} Y$ is symmetric?

If $\beta(x,y) = \beta(y,x)$ then it is symmetric

let $X = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $Y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

So $\beta(x,y) = [8 \ 16] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= [0 \ 8] \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 96$

$\beta(y,x)$

$[6 \ 12] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$[0 \ 6] \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$= 96$

Hence symmetric

Is this an inner product:

If you ever want to show something is an inner product you need to show three things for all $f, g \in V$ and $\alpha \in \mathbb{R}$

i) Symmetry: $\langle f, g \rangle = \langle g, f \rangle$

ii) Linearity: $\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$

iii) Positive definite: $\langle f, f \rangle \geq 0$ with equality if and only if $f = 0$, the zero function

The first two properties follow directly from the definition of an integral. For the third property you have

$$\int_a^b f^2 + \int_a^b (f')^2 \geq 0$$

Now when is this equal to 0? Well, recall that if a continuous function is positive everywhere, then the integral is positive. So f^2 is continuous this mean $f = 0$ everywhere. Thus $f' = 0$ everywhere also so equality holds

Positive eq $a(u,u) \geq 0$

Positive definite if $a(u,u) > 0$

$\Rightarrow X = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $Y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 8 \\ 16 \end{bmatrix}^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} = [8 \ 16] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$= [0 \ 8] \begin{bmatrix} 8 \\ 16 \end{bmatrix} = 128$

So for this equation is

i) bilinear ii) Symmetric iii) Positive definite

Since all three holds good, it is an inner product

$x = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$ $\quad x^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} y$

let $x = 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}$

$\Rightarrow (2 \cdot [2\ 4] + [4\ 8]) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$2\times 2 \quad 2\times 1$

$\Rightarrow 2 \cdot [2\ 4] + [4\ 8]) \cancel{[-6\ 6]} \begin{bmatrix} -6 \\ 6 \end{bmatrix}$

$\Rightarrow [4\ 8] + [4\ 8] \times \begin{bmatrix} -6 \\ 6 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 8 \\ 16 \end{bmatrix} [8\ 16] \times \begin{bmatrix} -6 \\ 6 \end{bmatrix} = -48 + 96 = 48$

IIth

<u>let $y = 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}$</u>

$x^T \cdot \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} y$

$f(x, cy+z) = cf(x,y) + f(x,z)$

$\left[2 \cdot f\left(\begin{bmatrix} 8 \\ 16 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right) + f\left(\begin{bmatrix} 8 \\ 16 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right)\right.$

(1) $\Rightarrow 2 \times [8\ 16] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} + [8\ 16] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$= [16\ 32] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} + [-8\ 8] \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$= [16\ 16] \begin{bmatrix} 2 \\ 4 \end{bmatrix} + [-8\ 8] \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$= -32 + 64 - 16 + 32 = 48$

$[8\ 16] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \times \left(2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right)$

$[-8\ 8] \times 2\begin{bmatrix} 6 \\ 12 \end{bmatrix} = -48 + 96 = 48$

$\cancel{-48 + 96}$

$f(cx+y, z) = cf(x,z) + f(y,z)$

$\Rightarrow 2 \cdot \begin{bmatrix} 2 \\ 4 \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 12 \end{bmatrix} + \begin{bmatrix} 6 \\ 12 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

$2 \cdot [2\ 4] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + \begin{bmatrix} 6 \\ 12 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix}^T$

$= [4\ 8] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [4\ 8] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= [4\ 4] \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [-4\ 4] \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= -24 + 48 - 24 + 48$

$= 48$

The equation is bilinear

is it symmetric

$\beta(x,y) = \beta(y,x)$

$\beta(y,x) = [6\ 12] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} = [-6\ 6] \begin{bmatrix} 8 \\ 16 \end{bmatrix} = -48 + 96$
$= 48$

Hence symmetric

$A(x,x) = [8\ 16] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} = [-8\ 8] \begin{bmatrix} 8 \\ 16 \end{bmatrix} = -64 + 128$

Hence positive definite

It is inner product

Verifying an inner product:

Consider the function defined by $\langle p,q\rangle = 2a_1b_1 + a_2b_2$

where $p(x) = a_1 + a_2 x$

and $q(x) = b_1 + b_2 x$

Verify the $\langle p,q\rangle$ defines an inner product on R

i) $\langle p,q\rangle = \langle q,p\rangle$

$\langle p,q\rangle = 2a_1b_1 + a_2b_2$

$= a_2b_2 + 2a_1b_1 = \langle q,p\rangle$

(ii) $\langle p, q+z\rangle = \langle p,q\rangle + \langle p,z\rangle$

$z(x) = c_1 + c_2 x$

Answer: To verify something is an "inner product" you have to show that it satisfies the definition of "inner product". And that is 1) linear in the first variable

That is that $= a+b$

Definition of inner product:

Definition 1: Let V be a vector space over R. An inner product $(,)$ is a function $V \times V \to R$ with the following properties

for all 1) $\forall u \in V$, $(u,u) \geq 0$ and $(u,u) = 0 \Leftrightarrow u = 0$

2) $\forall u, v \in V$, holds $(u,v) = (v,u)$

3) $\forall u, v, w \in V$ and $\forall a, b \in R$ holds

$(au + bv, w) = a(u,w) + b(v,w)$

Notation V together with $(,)$ is called a inner product space.

3) The function $B(x,y) = x^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} y$

i) Test for bilinearity

ii) Test for symmetry

iii) Test for positive definite

iv) Test for inner product

i) Test for bilinearity:

$f(cx+y, z) = cf(x,z) + f(y,z)$

$f(x, cy+z) = cf(y,z) + f(x,z)$

Let solve for $f(cx+y, z) = cf(x,z) + f(y,z)$

Let $x = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $z = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

x can be written as $\overset{c}{2}\overset{x}{\begin{bmatrix} 2 \\ 4 \end{bmatrix}} + \overset{y}{\begin{bmatrix} 4 \\ 8 \end{bmatrix}}$

Hence $f(cx+y, z) =$

$$\left(2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}\right)^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$= \left(2\begin{bmatrix} 2 & 4 \end{bmatrix} + \begin{bmatrix} 4 & 8 \end{bmatrix}\right) \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$= 2\begin{bmatrix} 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + \begin{bmatrix} 4 & 8 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$= 2\begin{bmatrix} 0 & 6 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + \begin{bmatrix} 0 & 16 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$= 2 \times 72 + 192$$

$$= 144 + 192 = 336$$

$cf(x,z) \Rightarrow 2f\left(\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \end{bmatrix}\right) = 2 \times \begin{bmatrix} 2 \\ 4 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= 2 \times \begin{bmatrix} 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= 2 \times \begin{bmatrix} 0 & 6 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 144$

$f(y,z) \Rightarrow f\left(\begin{bmatrix} 4 \\ 8 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 8 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= \begin{bmatrix} 4 & 8 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= \begin{bmatrix} 0 & 12 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 144$

$= 288$

Hence

~~$f(cx+y$~~ $f(cx+y, z) \neq cf(x,z) + f(y,z)$ Not bilinear by the first condition

$f(x, cy+z) = cf(y,z) + f(x,z)$

~~$f(x, cy+z) =$~~

let $x = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$$y = 2\overset{cy}{\begin{bmatrix} 2 \\ 4 \end{bmatrix}} + \overset{z}{\begin{bmatrix} 2 \\ 4 \end{bmatrix}}$$

$$f(x, cy+z) = \begin{bmatrix} 8 \\ 16 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \left(2 \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right)$$

$$= [8 \ 16] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \left(2 \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right)$$

$$= 2 \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} \times [8 \ 16] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} + [8 \ 16] \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 8 \end{bmatrix} [0 \ 24] + [0 \ 24] \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 192 + 96$$

$$= 288 \ //$$

$$cf(y,z) = 2f\left(\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right) = 2 \times [2 \ 4] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$= 2 \times [0 \ 6] \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$= 48$$

$$f(x,z) = f\left(\begin{bmatrix} 8 \\ 16 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right) =$$

$$\begin{bmatrix} 8 \\ 16 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = [8 \ 16] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$= [0 \ 24] \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 96$$

$$288 \neq 48 + 96$$
$$\neq 144$$

Hence $f(x, cy+z) \neq cf(y,z) + f(x,z)$

Hence not bilinear

Symmetric

$\beta(x,y) = \beta(y,x)$

$$\beta(x,y) = \begin{bmatrix} 8 \\ 16 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = [8 \ 16] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$= [0 \ 24] \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 288$$

$$\beta(y,x) = \begin{bmatrix} 6 \\ 12 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} = [6 \ 12] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix}$$

$$= [0 \ 18] \begin{bmatrix} 8 \\ 16 \end{bmatrix} = 288$$

$$\begin{array}{r} 24 \\ 12 \\ \hline 48 \\ 24 \\ \hline 288 \end{array} \qquad \begin{array}{r} 18 \\ 16 \\ \hline 108 \\ 18 \\ \hline 288 \end{array} \quad 4$$

Test for positive definite:

positive if $a(u,u) \geq 0$

Definite positive if $a(u,u) > 0$

$$f\left(\begin{bmatrix} 8 \\ 16 \end{bmatrix}, \begin{bmatrix} 8 \\ 16 \end{bmatrix}\right) = [8 \ 16] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} = [0 \ 24] \begin{bmatrix} 8 \\ 16 \end{bmatrix} = 384$$

$$f\left(\begin{bmatrix} 6 \\ 12 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \end{bmatrix}\right) = [6 \ 12] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} [6 \ 12] = [0 \ 18] \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 216$$

$$\begin{array}{r} 24 \\ 16 \\ \hline 144 \\ 24 \\ \hline 384 \end{array} \quad 2 \qquad \begin{array}{r} 18 \\ 12 \\ \hline \end{array}$$

Hence definite positive

Hence not an inner product

$\beta(x,y) = x^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} y$ i) Test for bilinearity ii) Test for symmetric iii) Test for positive definite iv) Test for inner product

Bilinearity $x = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$f(cx+y, z)$

let $x = cx + y = 2 \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}$ $y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$$\rightarrow \left(2 \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}\right)^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 2 \times [2 \ 4] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [4 \ 8] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$= [4 \ 8] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [4 \ 8] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$= [4 \ 8] \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [4 \ 8] \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 120 + 120$$

$$= 240 \ //$$

$$cf(x,z) = 2f\left(\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \end{bmatrix}\right) = 2 \times [2 \ 4] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$= 2 \times [2 \ 4] \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 120$$

$$f(y,z) = f\left(\begin{bmatrix} 4 \\ 8 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \end{bmatrix}\right) = [4 \ 8] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = [4 \ 8] \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 24 + 96 = 120$$

$$f(cx+y, z) = cf(x,z) + f(y,z)$$

$$120 \ //$$

$f(x, cy+z)$

$x = \begin{bmatrix} 8 \\ 16 \end{bmatrix} \quad cy+z = 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}$

$\beta = x^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} y$

$f(x, cy+z) = \begin{bmatrix} 8 \\ 16 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \left(2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}\right)$

$= \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \left(2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}\right)$

$= \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix}$

$= 2\begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix}$

$= 160 + 160 = 320$

$$\begin{array}{r} 16 \\ 8 \\ \hline 128 \\ 32 \\ \hline 160 \end{array} \quad 4$$

$cf(y,z) = 2f\left(\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \end{bmatrix}\right)$

$= 2 \cdot \begin{bmatrix} 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix}$

$= 2 \cdot \begin{bmatrix} 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix} = 80$

$f(x, z) = f\left(\begin{bmatrix} 8 \\ 16 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \end{bmatrix}\right) = \begin{bmatrix} 8 & 16 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix}$

$= \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix} = 32 + 128 = 160$

Hence not bilinear

~~Symmetric~~

$f\left(\begin{bmatrix} 8 \\ 16 \end{bmatrix}, \begin{bmatrix} 8 \\ 16 \end{bmatrix}\right) = \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$= \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} = 64 + 256$

$= 320$

Symmetric

$f\left(\begin{bmatrix} 8 \\ 16 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \end{bmatrix}\right) = \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= \begin{bmatrix} 8 & 16 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 48 + 192$

$= 240$

$$\begin{array}{r} 12 \\ 12 \\ \hline 72 \\ 12 \\ \hline \end{array} \quad 1$$

$f\left(\begin{bmatrix} 6 \\ 12 \end{bmatrix}, \begin{bmatrix} 8 \\ 16 \end{bmatrix}\right) = \begin{bmatrix} 6 & 12 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$= \begin{bmatrix} 6 & 12 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} = 48 + 192 = 240$

// Hence symmetry

$\beta(x, y) = x^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} y$

[Previously calculated Wrong]

i) This is bilinear

ii) This is ~~not~~ symmetric

Let $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad y = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$

So $\beta(x, y) = \beta(y, x)$

$\begin{bmatrix} 1 \\ 2 \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} -2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$2 = 2$ ✓

iii) This is not positive definite

$\beta(x, x) > 0 \Rightarrow$ Positive definite

$\beta(x, x) \geq 0 \Rightarrow$ Positive

$\Rightarrow \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1$

So $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0$

So $x = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

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

It is not positive definite

Hence

It is not an inner product

In linear algebra a symmetric $n \times n$ real matrix M is said to be positive definite if the scalar $z^T M z$ is strictly positive for every non-zero column vector z of n real numbers.

The function

$\beta(x,y) = x^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} y$

Case 1: $f(cx+y, z) = cf(x,z) + f(y,z)$

Case 2: $f(x, cy+z) = cf(y,x) + f(x,z)$

Let $x = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

x can be written as

$x = 2 \cdot \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}$

$cx + y$

$f(cx+y, z) \Rightarrow \left(2 \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix}\right)^T \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow (2 \times [2\ 4] + [4\ 8]) \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow 2 \times [2\ 4] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [4\ 8] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow [4\ 8] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [4\ 8] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$\Rightarrow [0\ 12] \begin{bmatrix} 6 \\ 12 \end{bmatrix} + [0\ 12] \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= 144 + 144 = 288$

$\Rightarrow cf(x,z) + f(y,z)$

$c = 2$

Here $x = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $y = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$, $z = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$cf(x,z) = 2 \cdot [2\ 4] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$[4\ 8] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$= [0\ 12] \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 144$

$f(y,z) = [4\ 8] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = 144$

$\Rightarrow f(cx+y, z) = cf(x,z) + f(y,z)$

$288 = 144 + 144$

$f(x, cy+z)$ $x = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $cy+z = 2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \cdot \left(2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right)$

$= [0\ 24] \cdot \left(2\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}\right)$

$= [0\ 24] \begin{bmatrix} 4 \\ 8 \end{bmatrix} + [0\ 24] \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$= \cancel{192 + 192 = 384}$ $192 + 96 =$

$cf(y,x) = 2 \cdot [2\ 4] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$= [4\ 8] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix} = [0\ 12] \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$= 12 \times 16 = 192$

$f(x, z) = \cancel{[8\ 16] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}}\ [2\ 4] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$= [0\ 6] \begin{bmatrix} 8 \\ 16 \end{bmatrix} = 96$

$[8\ 16] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$[0\ 24] \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 96$

Hence $192 + 96 = 192 + 96$

Bilinear ✓

Symmetric

Let $x = [1\ 1]^T$ $y = [2\ -1]^T$

$\beta(x,y) = \beta(y,x)$

$[1\ 1] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = [2\ -1] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$[1\ 0] \begin{bmatrix} 2 \\ -1 \end{bmatrix} = [4\ 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$2 \neq 5$ Not symmetric

$x = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$ $y = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$

$[8\ 16] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = [6\ 12] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$[0\ 24] \begin{bmatrix} 6 \\ 12 \end{bmatrix} = [0\ 18] \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

$288 = 288$

Let $x = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ $y = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$[1\ 3] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = [2\ 4] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$[-1\ 4] \begin{bmatrix} 2 \\ 4 \end{bmatrix} = [0\ 6] \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$14 \neq 18$

Not symmetric

$\beta(x,x)$ $[1\ 3] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} =$

$= [-1\ -2] \begin{bmatrix} 1 \\ 3 \end{bmatrix} = -5$

$[2\ 4] \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$[0\ 3] \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 12$

Not an inner product

Not positive

Positive definit

~~$R(x,y) = x^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$~~

Inner Product Definition:

An inner product is a generalization of the dot product

Definition: ~~for~~ any $x, y \in V$

for any vector x & y in vector space V,

an inner product is defined as symmetric, positive definite bilinear mapping

$$\langle x, y \rangle : V \times V \longrightarrow R$$

- Symmetric
- positive definite
- bilinear

Bilinear: $x, y, z \in V$ $\lambda \in \mathbb{R}$

$\langle \lambda x + z, y \rangle = \lambda \langle x, y \rangle + \langle z, y \rangle$ This will prove linearity for the first argument of the function

Similarly linearity for the second argument of the function needs to be established for Bilinearity

$$\langle x, \lambda y + z \rangle = \lambda \langle x, y \rangle + \langle x, z \rangle$$

Positive definite means that an inner product of x with itself is greater than or equal to zero

$$\langle x, x \rangle \geq 0, \; 0 \text{ only when } x \text{ is a zero vector}$$

Symmetry $\langle x, y \rangle = \langle y, x \rangle$

Let

$\langle x, y \rangle = x^T I y$ where I is identity matrix. This is our regular dot product

$\langle x, y \rangle = x^T A y$ where $A \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$$x^T \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} y$$

$$\Rightarrow (2x_1 + x_2 \quad x_1 + 2x_2) y$$

$$\Rightarrow 2x_1y_1 + x_2y_2 + x_1y_2 + 2x_2y_2$$

And this inner product is different from dot product

Inner product: lengths of vectors & Distances between vector

$\|x\| = \sqrt{\langle x, x \rangle}$ The length of vector x is defined as the ~~the~~ square root of inner product with itself

Since inner product is positive definite

$\sqrt{\langle x, x \rangle} \geq 0$

length of the vector depends on the inner product and depending on the choice of the inner product the length of the vector can be quite different

Question:

What is the length of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ using the inner product defined as $\langle x, y \rangle = x^T \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix} y$

$$\Rightarrow [x_1 \; x_2] \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= [x_1 - 0.5x_2, \; -0.5x_1 + x_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= (x_1 - 0.5x_2)y_1 + (-0.5x_1 + x_2)y_2$$

$$= x_1y_1 - 0.5x_2y_1 - 0.5x_1y_2 + y_2x_2$$

$$= 1 - 0 - 0 + 0 = 1$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{matrix} x_1 \\ x_2 \end{matrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{matrix} y_1 \\ y_2 \end{matrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\|x\| = \sqrt{x_1^2 - \frac{1}{2}(x_1x_2 + x_2x_1) + x_2^2}$$

$$= \sqrt{x_1^2 - x_1x_2 + x_2^2}$$

$$\|x\|^2 = \langle x, x \rangle = 1 + 1 - 1 = 1$$

$$\|x\| = 1$$

$\|\lambda x\| = |\lambda| \|x\|$

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

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

Triangle inequality

$\|x+y\| \le \|x\| + \|y\|$

If we use dot product as the inner product

$\cancel{\begin{bmatrix} 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \end{bmatrix}}$

$\cancel{[1\ 0] \begin{bmatrix} 0 \\ 1 \end{bmatrix} =}$

$\|x\| = 1 = \|y\| \Rightarrow \sqrt{1^2+0} = \sqrt{0^2+1}$

$\|x+y\| = \sqrt{2}$

$\Rightarrow \|x+y\| \le \|x\| + \|y\|$

Cauchy-Schwartz inequality:

$|\langle x, y\rangle| \le \|x\| \|y\|$. The absolute value of inner product with $x+y$ is less than or equal to product of individual norms of these two vectors.

Inner product distances between vectors:

$x = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ $y = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ $d(x,y) = \|x-y\| = \sqrt{\langle x-y, x-y\rangle}$

$x - y = \begin{bmatrix} 2-4 \\ 3-1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

Dot product:

$\|x-y\| = \sqrt{4+4} = \sqrt{8}$ ✓

$\langle x, y\rangle = x^T \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} y = \cancel{[2\ 3] \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}}$

$\cancel{= [-0.5\ \ 2] \begin{bmatrix} 4 \\ 1 \end{bmatrix} = 0}$

$\langle x-y, x-y\rangle =$

$\cancel{[-2\ 0] \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = [-2\ \ 1] \begin{bmatrix} 0 \\ 2 \end{bmatrix} = 2}$

Wrong

Compute the length of

$x = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ using the inner product defined

$\langle a, b\rangle = a^T \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} b$

$= [a_1\ a_2\ a_3] \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

$= [2a_1 + a_2 \quad a_1 + 2a_2 - a_3 \quad -a_2 + 2a_3] \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

$= \cancel{2a_1b_1 + a_2b_1 + a_1b_1 + 2a_2b_2 - a_3b_3}$

$(2a_1 + a_2)b_1 + (a_1 + 2a_2 - a_3)b_2 - (a_2 + 2a_3)b_3$

$= 2a_1b_1 + a_2b_1 + a_1b_2 + 2a_2b_2 - a_3b_2 - a_2b_3 + 2a_3b_3$

$x - y = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

If $\langle x, y\rangle = x^T \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} y$

$\langle x-y, x-y\rangle = [-2\ 2] \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

$= [-2-1 \quad 1+2] \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

$= 6 + 6 = 12$

Compute the length of $x = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ using the inner product defined

$$\langle a, b \rangle = a^T \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} b \Rightarrow [1\ -1\ 3] \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$$

$$\Rightarrow [1\ -4\ 7] \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \Rightarrow 26 \Rightarrow \text{length} = \sqrt{26}$$

Compute the squared distance between $x = \begin{bmatrix} 1/2 \\ -1 \\ -1/2 \end{bmatrix}$ and $y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ using the inner product defined as

$$\langle a, b \rangle = a^T \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} b$$

$$x^2 = \begin{bmatrix} 1/4 \\ 1 \\ 1/4 \end{bmatrix} \quad y^2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad x^2 - y^2 = \begin{bmatrix} 1/4 \\ 0 \\ 1/4 \end{bmatrix}$$

$$[1/4\ 0\ 1/4] \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1/4 \\ 0 \\ 1/4 \end{bmatrix} \rightarrow [0.5\ 0\ 0.5] \begin{bmatrix} 1/4 \\ 0 \\ 1/4 \end{bmatrix}$$

$$\rightarrow 0.5 \times \frac{1}{4} + 0 + 0.5 \times \frac{1}{4}$$

$$\Rightarrow \frac{0.5}{4} + \frac{0.5}{4} \Rightarrow \frac{1}{4}$$

$$x - y = [1/2\ -1\ -1/2] \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1/2 \\ -1 \\ -1/2 \end{bmatrix}$$

$$= [0\ -1\ 0] \begin{bmatrix} 1/2 \\ -1 \\ -1/2 \end{bmatrix} = 1$$

$$\begin{bmatrix} \frac{1}{2} \\ -1 \\ -\frac{1}{2} \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -2 \\ -\frac{1}{2} \end{bmatrix} \Rightarrow [\tfrac{1}{2}\ -2\ -\tfrac{1}{2}] \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1/2 \\ -1 \\ -1/2 \end{bmatrix}$$

$$= [-1\ -3\ 1] \begin{bmatrix} 1/2 \\ -1 \\ -1/2 \end{bmatrix}$$

$$= -1/2 + 3 - 1/2 = -5$$

$$= [1/2\ -2\ -1/2] \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1/2 \\ -2 \\ -1/2 \end{bmatrix}$$

$$= [-1\ -3\ 1] \begin{bmatrix} 1/2 \\ -2 \\ -1/2 \end{bmatrix} = -1/2 + 6 - 1/2$$

$$= 5$$

Distance is $\sqrt{5}$

Squared distance between x & y is $(\sqrt{5})^2 = 5$

Compute the length of $x = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ using the inner product defined by $\langle a, b \rangle = a^T \frac{1}{2} \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix} b$

$$[-1\ 1] \frac{1}{2} \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \frac{1}{2} [-1\ 1] \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2} [-6\ 6] \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2} \times 12 = 6$$

Compute the distance between $x = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ and $y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ using the inner product defined as $\langle a, b \rangle = a^T \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} b$

$$x - y = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} \Rightarrow [1\ 1\ 4] \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$$

$$= [3\ -5\ 7] \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} = 3 - 5 + 28 = 26$$

$$\cancel{3 - 1 + 28 = 30}$$

Compute the length of $x = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$ using the inner product defined as $\langle a, b \rangle = a^T I b$ where I is the identity matrix

$$[-1\ -1\ -1] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

$$= [-1\ -1\ -1] \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} = 3$$

length $\|x\| = \sqrt{3}$

Generating sets and basis:

In a vector space V, we are particularly interested in sets of vectors A that possess the property that any vector $v \in V$ can be obtained by a linear combination of vectors in A.

Definition 1: Generating set / span. Consider a vector space V and set of vectors $A = \{x_1 \dots x_n\} \subseteq V$. If every vector $v \in V$ can be expressed as a linear combination of $x_1 \dots x_k$. A is called a generating set or span which spans the vector space V. In this case we write $V = [x_1, \dots, x_k]$

Generating sets are sets of vectors that span sub spaces i.e every vector can be represented as a linear combination of the vectors in the generating set. Now we will be more specific and characterise the smallest generating set that spans a vector (sub) space

Definition (Basis) Consider a real vector space V and $A \subseteq V$

1) A generating set A of V is called minimal if there exists no smaller set $A \subseteq A \subseteq V$ that spans V

2) Every linearly independent generating set of V is minimal and is called basis of V.

Let V be a real vector space and $B \subseteq V$, $B \neq 0$ Then the following statements are true

1) B is a basis of V
2) B is a minimum generating set
3) B is a maximal linearly independent subset of vectors in V
4) Every vector $x \in V$ is a linear combination of vectors from B and every linear combination is unique i.e with

$$x = \sum_{i=1}^{k} \lambda_i b_i = \sum_{i=1}^{k} \psi_i b_i$$

and $\lambda_i, \psi_i \in R$, $b_i \in B$ it follows that $\lambda_i = \psi_i$, $i = 1, \dots, k$

[NOTE 1 DEGREE = 0.0174533]

Exercise:

Angles between vectors using a non-standard inner product

1) Compute the angle between $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ using the inner product defined by $\langle x, y \rangle = x^T \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix} y$

$$\cos w = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

$$\langle x, y \rangle = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -1 + 3 = 2$$

$$\cos w = \frac{\cancel{2}}{\cancel{2}} = 1$$

$$\cancel{\|x\| = \sqrt{2} \quad \|y\| = \sqrt{2}}$$

$$w = \cos^{-1}(1) =$$

$$\cos w = \frac{\langle x, y \rangle}{\|x\| \|y\|} = \frac{\langle x, y \rangle}{\sqrt{\langle x, x \rangle \cdot \langle y, y \rangle}}$$

$$\langle x, x \rangle = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

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

$$\langle y, y \rangle = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 & 5 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 8$$

$$\cos w = \frac{2}{\sqrt{32}} = 2 \Rightarrow w = \cos^{-1}\left(\frac{2}{\sqrt{32}}\right) = 1.2 \text{ rad } (69^\circ)$$

Compute the angle between $x = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ and $y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ using the inner product defined by $\langle x, y \rangle = x^T \begin{bmatrix} 1 & -1/2 \\ -1/2 & 5 \end{bmatrix} y$

$$\cos w = \frac{\langle x, y \rangle}{\|x\| \|y\|} = \frac{\langle x, y \rangle}{\sqrt{\langle x, x \rangle \cdot \langle y, y \rangle}}$$

$$\begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1/2 \\ -1/2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1/2 - 5 = -4.5$$

$$\langle x, x \rangle = \begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1/2 \\ -1/2 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1/2 & -5 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 5$$

$$\langle y, y \rangle = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1/2 \\ -1/2 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.5 & 4.5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5$$

$$\cos w = \frac{-4.5}{\sqrt{5 \times 4.5}} \Rightarrow w = \cos^{-1}\left(\frac{-4.5}{\sqrt{5 \times 5}}\right) = \cancel{2.8984 \text{ rads}} \; 2.69056$$

1 Degree = 0.0174533 rad

180 degrees = π radians

1 degree = $\frac{\pi \text{ rad}}{180} \Rightarrow$ 1 radian = $\frac{180}{\pi}$

Compute the angle between $x = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $y = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$ using the inner product defined by $\langle x, y \rangle = x^T \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} y$

$$\cos w = \frac{\langle x, y \rangle}{\|x\| \|y\|} = \frac{\langle x, y \rangle}{\sqrt{\langle x, x \rangle \langle y, y \rangle}}$$

$$\langle x, y \rangle = \begin{bmatrix} 2 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ -2 \end{bmatrix} = -32$$

$$\langle x, x \rangle = \begin{bmatrix} 2 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 & 10 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 32$$

$$\langle y, y \rangle = \begin{bmatrix} -2 & -2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ -2 \end{bmatrix} = \begin{bmatrix} -6 & -10 \end{bmatrix} \begin{bmatrix} -2 \\ -2 \end{bmatrix} = 32$$

$$\cos^{-1}\left(\frac{-32}{\sqrt{32 \cdot 32}}\right) \Rightarrow 3.1415$$

Compute the angle between $x = \begin{bmatrix}1\\1\end{bmatrix}$ and $y = \begin{bmatrix}1\\-1\end{bmatrix}$ using the inner product defined by

$$\langle x,y\rangle = x^T \begin{bmatrix}1&0\\0&5\end{bmatrix} y \qquad \cos\theta = \frac{\langle x,y\rangle}{\|x\|\,\|y\|} = \frac{\langle x,y\rangle}{\sqrt{\langle x,x\rangle\langle y,y\rangle}}$$

$$\langle x,y\rangle = [1\ 1]\begin{bmatrix}1&0\\0&5\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = [1\ 5]\begin{bmatrix}1\\-1\end{bmatrix} = -4$$

$$\langle x,x\rangle = [1\ 1]\begin{bmatrix}1&0\\0&5\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = [1\ 5]\begin{bmatrix}1\\1\end{bmatrix} = 6$$

$$\langle y,y\rangle = [1\ -1]\begin{bmatrix}1&0\\0&5\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = [1\ -5]\begin{bmatrix}1\\-1\end{bmatrix} = 6$$

$$\cos\theta = \frac{-4}{6}$$

$$\theta = \cos^{-1}\left(-\frac{2}{3}\right) = 2.300$$

Compute the angle between $x = \begin{bmatrix}1\\1\\1\end{bmatrix}$ and $y = \begin{bmatrix}2\\-1\\0\end{bmatrix}$ using the inner product defined by

$$\langle x,y\rangle = x^T \begin{bmatrix}1&0&0\\0&2&-1\\0&-1&3\end{bmatrix} y$$

$$\cos\theta = \frac{\langle x,y\rangle}{\|x\|\,\|y\|} = \frac{\langle x,y\rangle}{\sqrt{\langle x,x\rangle\langle y,y\rangle}} = \cos^{-1}\left(\frac{\langle x,y\rangle}{\sqrt{\langle x,x\rangle\langle y,y\rangle}}\right)$$

$$\langle x,y\rangle = \begin{bmatrix}1\\1\\1\end{bmatrix}^T \begin{bmatrix}1&0&0\\0&2&-1\\0&-1&3\end{bmatrix}\begin{bmatrix}2\\-1\\0\end{bmatrix} = [1\ 1\ 2]\begin{bmatrix}2\\-1\\0\end{bmatrix} = 1$$

$$\langle x,x\rangle = [1\ 1\ 1]\begin{bmatrix}1&0&0\\0&2&-1\\0&-1&3\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = [1\ 1\ 2]\begin{bmatrix}1\\1\\1\end{bmatrix} = 4$$

$$\langle y,y\rangle = [2\ -1\ 0]\begin{bmatrix}1&0&0\\0&2&-1\\0&-1&3\end{bmatrix}\begin{bmatrix}2\\-1\\0\end{bmatrix} = [2\ -2\ 1]\begin{bmatrix}2\\-1\\0\end{bmatrix} = 6$$

$$\cancel{\cos\theta = \frac{1}{\sqrt{6\times 6}} = \frac{1}{6} \Rightarrow \cos^{-1}\left(\frac{1}{6}\right) = 1.4033}$$

$$\cos\theta = \frac{1}{\sqrt{6\times 4}} = \frac{1}{\sqrt{24}} \Rightarrow \theta = \cos^{-1}\left(\frac{1}{\sqrt{24}}\right) = 1.365227$$

MNIST Database:

Modified National Institute of Standards and Technology

The MNIST database is a large database of handwritten digits that is commonly used for training various image processing systems.

1) What does it mean for two digits in the MNIST dataset to be different by our distance function

2) Furthermore, how are different classes of digits different for MNIST digits?

The idea behind a kNN classifier is pretty simple. Given a training set $X \in \mathbb{R}^{N\times D}$ and $y \in \mathbb{Z}^N$, we predict the label of a new point $x \in \mathbb{R}^D$ as the label of the majority of its "k nearest neighbor" (hence the name kNN) by some distance measure (e.g. the euclidean distance). Here N is the number of data points in the dataset and D is the dimensionality of the data

$$X \in \mathbb{R}^{N\times D} \qquad y \in \mathbb{Z}^n$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \\ 1.0 & 0.1 & 43 & 2.3 \\ \vdots & & & \end{bmatrix} \qquad \begin{bmatrix}1\\2\\3\\\vdots\end{bmatrix}$$

2) How are different classes of digits different for MNIST digits?

We can compute a mean image for each class of image i.e we compute mean image for digits of 1, 2, 3 ... 9 then we compute pairwise distance between them

```
def pairwise_distance_matrix(x, y):
```

IRIS Data set: http://archive.ics.uci.edu/ml/datasets/iris

Sepal length	Sepal width	Petal length	Petal width	iris setosa
				iris versicolor
				iris virginica

$$\begin{bmatrix} \text{Sepal L} & \text{Sepal W} & \text{Petal L} & \text{Petal W} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \qquad \begin{bmatrix} \text{iris setosa} \\ \text{iris versicolour} \\ \text{iris virginica} \end{bmatrix}$$

The idea behind a KNN classifier is pretty simple. Given a training set $X \in \mathbb{R}^{N \times D}$ and $y \in \mathbb{Z}^N$, we predict the label of a new point $x \in \mathbb{R}^D$ as the label of the majority of its "k nearest neighbours" (hence the name KNN) by some distance measure.

Compute the pairwise distance between rows of x and rows of y.

```
def pairwise_distance_matrix:
    "Compute the pairwise distance between rows of x
    and rows of y"
    Arguments: x: ndarray of size (N, D)
               y: ndarray of size (M, D)

    Returns:
        distance matrix: matrix of shape (N, M)
        each entry distance_matrix[i, j] is the
        distance between ith row of x and the jth row of
        y. We use the dot product to compute the distance.
    N, D = x.shape
    M, _ = y.shape
```

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

$4 \times 4 \qquad 2 \times 4$

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$4 \times 4 - 1 \times 4$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{bmatrix}$$

Compute pairwise distance between rows of x and rows of y.

x: ndarray of size (N, D)

y: ndarray of size (M, D)

distance_matrix: Matrix of shape (N, M), each entry distance_matrix[i, j] is the distance between the ith row of x and j^{th} row of y.

Numpy argsort: numpy.argsort(a, axis=1, kind=None, order=None)

Returns the indices that would sort an array

Perform an indirect sort along the given axis using the algorithm specified by the kind keyword. It returns an array of indices of the same shape as "a" that index data along the given axis in sorted order.

Eg: x = np.array([3, 1, 2])

np.argsort(x) ⟹ array([1, 2, 0])

This means index 1 should be moved to index 0

index 2 should be moved to index 1

index 0 should be moved to index 2

x = np.array([[0, 3], [2, 2]])

Now if distance matrix is like this $\xrightarrow{\text{dist}}$ $\begin{bmatrix} 1 \\ \vdots \\ n \end{bmatrix}$ and labels $y = \begin{bmatrix} 5 \\ \vdots \\ n \end{bmatrix}$

np.argsort(dist) matrix would say which index should be at 0th index, 1st, 2nd and so on

dist $\begin{bmatrix} \vdots \end{bmatrix} = y \cdot \begin{bmatrix} \vdots \end{bmatrix}$

So when applied to y like this y[np.argsort(dist)][:k]

The corresponding index of dist at 0th position would be at y's 0th position

argsort array's

Remember $\begin{array}{c} y \\ \begin{bmatrix} 1 \\ a \\ b \\ c \\ d \end{bmatrix} \end{array} \begin{array}{c} x \\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \end{array}$ $\begin{array}{c} \text{Reg Distance from} \\ \begin{bmatrix} 4 & 8 \end{bmatrix} \end{array}$ gives the distance matrix dist

The distance matrix dist corresponds to y

$$\begin{array}{c} \text{dist} \\ \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix} \end{array} \begin{array}{c} y \\ \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix} \end{array}$$

& when y[np.argsort(dist)][:k] we get the k nearest neighbours.

Once y is populated, the majority in "k" ~~has~~ decides the label of the question.

```
Classes = y[np.argsort(dist)][:k]
ypred = np.zeros(num_classes) => num_classes = len(np.unique(y))
for c in np.unique(classes):
    ypred[c] =
```

Distance matrix

let $X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}$ $Y = \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \end{pmatrix}$

EUCLIDEAN DISTANCE $= \sqrt{(x-y)(x-y)}$

Difference matrix :

$$X_I = \begin{pmatrix} x_{11}-y_{11} & x_{12}-y_{12} & x_{13}-y_{13} \\ x_{21}-y_{11} & x_{22}-y_{12} & x_{23}-y_{13} \\ x_{31}-y_{11} & x_{32}-y_{12} & x_{33}-y_{13} \end{pmatrix}$$

$$X_{II} = \begin{pmatrix} x_{11}-y_{21} & x_{12}-y_{22} & x_{13}-y_{23} \\ x_{21}-y_{21} & x_{22}-y_{22} & x_{23}-y_{23} \\ x_{31}-y_{21} & x_{32}-y_{22} & x_{33}-y_{23} \end{pmatrix}$$

```
X.shape (3, 3)        Y.shape (2, 3)
m, n = X.shape
m1, n1 = Y.shape
X1 = X.reshape(1, m*n)
X2 = np.tile(X1, (m1, 1))
```

What to do with X:

1) flatten it
2) Tile it to match the number of rows in Y.

What to do with Y:

1) Tile the rows of Y to match the length of a row in X

```
m1, n1 = Y.shape
np.tile(Y, m)
```

~~$(x^2-y^2$ $(x-y)(x-y)$ x^2~~

Distance matrix continued:

$$N\left\{\overbrace{\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}}^{D}\right. \quad \left.\overbrace{\begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \\ y_{41} & y_{42} & y_{43} \end{pmatrix}}^{C}\right\} R$$

What do we do

x.reshape(1, N*D)

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{21} & x_{22} & x_{23} & x_{31} & x_{32} & x_{33} \end{pmatrix} \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ \cdots & \cdots & \cdots \end{pmatrix}$$

np.tile(x, R)

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{21} & x_{22} & x_{23} & x_{31} & x_{32} & x_{33} \\ x_{11} & x_{12} & x_{13} & x_{21} & x_{22} & x_{23} & x_{31} & x_{32} & x_{33} \\ x_{11} & x_{12} & x_{13} & x_{21} & x_{22} & x_{23} & x_{31} & x_{32} & x_{33} \\ x_{11} & x_{12} & x_{13} & x_{21} & x_{22} & x_{23} & x_{31} & x_{32} & x_{33} \end{pmatrix} \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \\ y_{41} & y_{42} & y_{43} \end{pmatrix}$$

np.tile(y, (1, N)) $\longrightarrow$

$$\begin{pmatrix} y_{11} & y_{12} & y_{13} & y_{11} & y_{12} & y_{13} & y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} & y_{21} & y_{22} & y_{23} & y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} & y_{31} & y_{32} & y_{33} & y_{31} & y_{32} & y_{33} \\ y_{41} & y_{42} & y_{43} & y_{41} & y_{42} & y_{43} & y_{41} & y_{42} & y_{43} \end{pmatrix}$$

X - Y

$$\begin{pmatrix} x_{11}-y_{11} & x_{12}-y_{12} & x_{13}-y_{13} & x_{21}-y_{11} & x_{22}-y_{12} & x_{23}-y_{13} & x_{31}-y_{11} & x_{32}-y_{12} & x_{33}-y_{13} \\ x_{11}-y_{21} & x_{11}-y_{22} & x_{11}-y_{23} & x_{21}-y_{21} & x_{22}-y_{22} & x_{23}-y_{23} & x_{31}-y_{21} & x_{32}-y_{22} & x_{33}-y_{23} \\ x_{11}-y_{31} & x_{11}-y_{32} & x_{11}-y_{33} & x_{21}-y_{31} & x_{22}-y_{32} & x_{23}-y_{33} & x_{31}-y_{31} & x_{32}-y_{32} & x_{33}-y_{33} \end{pmatrix}$$

Here # of columns is the # of rows in X
# of rows is the # of rows in Y

$$\begin{pmatrix} z_{1111} & z_{1212} & z_{1313} & z_{2111} & z_{2212} & z_{2313} & z_{3111} & z_{3212} & z_{3313} \\ z_{1121} & z_{1122} & z_{1123} & z_{2121} & z_{2222} & z_{2323} & z_{3121} & z_{3222} & z_{3323} \\ z_{1131} & z_{1132} & z_{1133} & z_{2131} & z_{2232} & z_{2333} & z_{3131} & z_{3232} & z_{3333} \end{pmatrix}$$

Inner products:

Inner product of functions and Inner product of random variables

Projection on to a 1-dimensional subspace:

1) Compute the projection matrix that allows us to project any vector $x \in \mathbb{R}^3$ on to the subspace spanned by the basis vector $b = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

Formula:

Projection of x on subspace U

$$\pi_U(x) = \underbrace{\frac{b b^t}{\|b\|^2}}_{\text{Projection matrix}} x$$

$$\Rightarrow \frac{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \cdot [1\ 2\ 2]}{9} x \Rightarrow \frac{9}{9} x$$

$3\times1 \quad 1\times3$

2) Given the projection matrix $\frac{1}{25}\begin{bmatrix} 9 & 0 & 12 \\ 0 & 0 & 0 \\ 12 & 0 & 16 \end{bmatrix}$ project $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ onto the corresponding subspace which is spanned by $b = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$

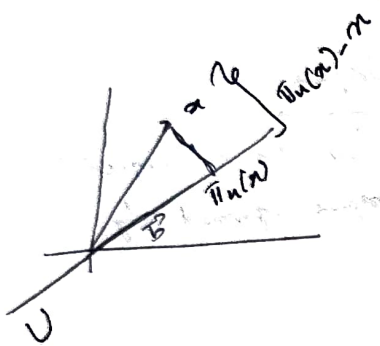
$$\Rightarrow \begin{bmatrix} 9 & 0 & 12 \\ 0 & 0 & 0 \\ 12 & 0 & 16 \end{bmatrix}$$

PTO

1) Compute the projection matrix that allows us to project any vector $x \in \mathbb{R}^3$ on to the subspace spanned by basis vector $b = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

$$\frac{x^T b}{\|b\|^2} b$$

$$\frac{5}{9}$$



i) $\Pi_u(x) = \lambda b$

ii) $\vec{b} \cdot (\Pi_u(x) - x) = 0$ They are orthogonal

$\rightarrow b\Pi_u(x) - bx = 0$

$b \cdot \lambda b - bx = 0$

$b \cdot \lambda b = bx$

$\lambda = \frac{bx}{\|b\|^2}$

$\Pi_u(x) = \frac{bx}{\|b\|^2} b$

$$\frac{bx}{(\sqrt{1^2+2^2+2^2})^2} b$$

$$\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}$$

$$\frac{1}{9} bx\, b$$

2) Given the projection matrix $\frac{1}{25}\begin{bmatrix} 9 & 0 & 12 \\ 0 & 0 & 0 \\ 12 & 0 & 16 \end{bmatrix}$ project $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ onto the corresponding subspace, which is spanned by $b = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$

$$\begin{bmatrix} 3 & 0 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 9 & 0 & 12 \\ 0 & 0 & 0 \\ 12 & 0 & 16 \end{bmatrix}$$

$$\underset{3\times 3}{\frac{1}{25}\begin{bmatrix} 9 & 0 & 12 \\ 0 & 0 & 0 \\ 12 & 0 & 16 \end{bmatrix}} \underset{1\times 3}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}$$

$$\frac{1}{25}\begin{bmatrix} 21 \\ 0 \\ 28 \end{bmatrix}$$

3) Now we compute the reconstruction error i.e the distance between the original point and its projection onto a lower-dimensional subspace.

Assume our original data point is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and its projection $\frac{1}{9}\begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix}$. What is the reconstruction error

$$\cancel{\frac{1}{9}\begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}} = \frac{1}{9}\Big[$$

$$\cancel{\Pi_u(x) = \frac{bx}{\|b\|^2}}$$

$$\Pi_u(x) = \frac{1}{9}\begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix}$$

$$x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\frac{1}{9}\begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \frac{5}{9} \\ \frac{10}{9} \\ 10/9 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{4}{9} \\ \frac{1}{9} \\ \frac{1}{9} \end{bmatrix}$$

$$\sqrt{\frac{16}{81} + \frac{1}{81} + \frac{1}{81}}$$

$$\sqrt{\frac{18}{81}} = \sqrt{\frac{2}{9}}$$

Orthogonal projection:

Projection onto - general subspaces:

General exx of orthogonal projection onto n dimensional sub spaces:

$\pi_u(x) = B\lambda$ $\quad x = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ $\quad b_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ $\quad b_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

$B = [b_1 | b_2]$ $\quad u = [b_1, b_2]$

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

$\lambda = (B^T B)^{-1} B^T x$

$B^T x = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$

$B^T B = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 0 & 0 \end{bmatrix}$

$\begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$

$B^T B \lambda = B^T x$

$\begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} \lambda = B^T x$

$\lambda = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$

$\lambda = \frac{b}{\|b\|^2} x$

$\Rightarrow \frac{b^T x}{b^T b}$

For a vector $x = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ and the subspace U spanned by the basis vectors $b_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$.

The projection matrix is symmetric -

$B = [b_1 | b_2] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$

$\lambda = (B^T B)^{-1} B^T x$

$B^T x = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$

$B^T B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$

$= \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$

$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \lambda = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$

Gaussian Elimination

$R1 \Rightarrow R2 + 1 - R1$

$\begin{bmatrix} 1 & 3 \\ 3 & 5 \end{bmatrix} \lambda = \begin{bmatrix} -5 \\ 0 \end{bmatrix}$

$R2 \Rightarrow R1 + 2 - R2$

$\begin{matrix} 1 & 3 \\ 0 & \end{matrix}$

The works

Attempt

$R_2 = R2 - R1$

$\begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix} \lambda = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$

$R_2 = R2 \div 2$

$\begin{bmatrix} 3 & 3 \\ 0 & 1 \end{bmatrix} \lambda = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$

$R_1 = R_1 \div 3$

$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \lambda = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

$R_1 = R_1 - R_2$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$

$\pi_u(x) = B\lambda$

$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \end{bmatrix}$

$3 \times 2 \quad 2 \times 1$

$= \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$

$b_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $b_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

Projection Matrix $(B^T B)^{-1} B^T$

$$B \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$$

Let $(B^T B)^{-1} = X$

$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$$

$R_2 = R_2 - R_1$

$$\begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix} x = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

$R2 = R2 \div 2$

$$\begin{bmatrix} 3 & 3 \\ 0 & 1 \end{bmatrix} x = \begin{bmatrix} 1 & 0 \\ -1/2 & 1/2 \end{bmatrix}$$

$R_1 = R_1 \div 3$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x = \begin{bmatrix} 1/3 & 0 \\ -1/2 & 1/2 \end{bmatrix}$$

$1/3 + 1/2$

$5/6$

$R_1 = R_1 - R_2$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x = \begin{bmatrix} 5/6 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 5/3 & -1 \\ -1 & 1 \end{bmatrix}$$

$$(B^T B)^{-1} = \frac{1}{2} \begin{bmatrix} 5/3 & -1 \\ -1 & 1 \end{bmatrix}$$

$$(B^T B)^{-1} B^T = \frac{1}{2} \begin{bmatrix} 5/3 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

$\frac{5}{3} - 1 = \frac{2}{3}$

$\frac{5}{3} - 2 = -\frac{1}{3}$

$$= \frac{1}{2} \begin{bmatrix} 5/3 & 5/3 - 1 & 5/3 - 2 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 5/3 & 2/3 & -1/3 \\ -1 & 0 & 1 \end{bmatrix} = \frac{3}{2} \begin{bmatrix} 5 & 2 & -1 \\ -3 & 0 & 3 \end{bmatrix}$$

$(B^T B)^{-1} (B^T B) = I$

$$\begin{bmatrix} 5/6 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} \frac{15}{6} - \frac{3}{2} & \frac{15}{6} - \frac{5}{2} \\ -\frac{3}{2} + \frac{3}{2} & -\frac{3}{2} + \frac{5}{2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{30-18}{12} & \frac{30-30}{2} \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 12/12 & 0 \\ 0 & 1 \end{bmatrix}$$

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

$$(B^T B)^{-1} B^T = \frac{1}{2} \begin{bmatrix} 5/3 & 2/3 & -1/3 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\lambda = (B^T B)^{-1} B^T x = \frac{1}{2} \begin{bmatrix} 5/3 & 2/3 & -1/3 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 10 \\ -6 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

Projection matrix is symmetric

$$\begin{bmatrix} 5/3 & 2/3 & -1/3 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5/3 & -1 \\ 2/3 & 0 \\ -1/3 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 5/3 & -1 \\ 2/3 & 0 \\ -1/3 & 1 \end{bmatrix} =$$

$2 \times 3 \quad 3 \times 2$

$3 \times 2 \quad 2 \times 3$

$$\begin{bmatrix} 5/3 & 2/3 & -1/3 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 25/9 + 4/9 + 1/9 & -5/3 + 0 + 1 \\ -5/3 + 0 - 1/3 & 2 \end{bmatrix} =$$

Project $\begin{bmatrix}3\\2\\2\end{bmatrix}$ on the subspace spanned by $\begin{bmatrix}1\\0\\0\end{bmatrix}$ $\begin{bmatrix}0\\1\\1\end{bmatrix}$

$b_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$ $b_2 = \begin{bmatrix}0\\1\\1\end{bmatrix}$

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

$\cancel{\pi_u(x)} = (B^TB)^{-1}B^Tx$ $\qquad \pi_u(x) = B\lambda$

$B^TB = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\\0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}$

Let $B^TB = x$

$x \cdot x^{-1} = I$

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

$R_2 \to R2/2$

$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} x^{-1} = \begin{bmatrix}1 & 0\\0 & 1/2\end{bmatrix}$

$(B^TB)^{-1} = \begin{bmatrix}1 & 0\\0 & 1/2\end{bmatrix} \Rightarrow (B^TB)^{-1}B^T \Rightarrow \begin{bmatrix}1 & 0\\0 & 1/2\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 1\end{bmatrix}$

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

$\lambda = \begin{bmatrix}1 & 0 & 0\\0 & 1/2 & 1/2\end{bmatrix}\begin{bmatrix}3\\2\\2\end{bmatrix} = \begin{bmatrix}3\\2\end{bmatrix}$

$2\times3 \quad 3\times1$

$\pi_u(x) = \cancel{\lambda B = \begin{bmatrix}3\\2\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\\0 & 1\end{bmatrix}} = B\lambda = \begin{bmatrix}1 & 0\\0 & 1\\0 & 1\end{bmatrix}\begin{bmatrix}3\\2\end{bmatrix}$

$2\times1 \quad 3\times2$

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

3) Project $\begin{bmatrix}12\\0\\0\end{bmatrix}$ onto the subspace U, spanned by $\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\2\end{bmatrix}$.

4) Project the result from i on to the subspace spanned by $\begin{bmatrix}-10\sqrt{6}\\-4\sqrt{6}\\2\sqrt{6}\end{bmatrix}$. What is the final projection

Hint: For step 2 you do not necessarily need to compute anything

$B = \begin{bmatrix}1 & 0\\1 & 1\\1 & 2\end{bmatrix}$ $\quad B^T = \begin{bmatrix}1 & 1 & 1\\0 & 1 & 2\end{bmatrix}$ $\quad$ Let $(B^TB)^{-1} = x$

$B^TB = \begin{bmatrix}1 & 1 & 1\\0 & 1 & 2\end{bmatrix}\begin{bmatrix}1 & 0\\1 & 1\\1 & 2\end{bmatrix} = \begin{bmatrix}3 & 3\\3 & 5\end{bmatrix}$

From the previous problem

$(B^TB)^{-1} = \frac{1}{2}\begin{bmatrix}5/3 & -1\\-1 & 1\end{bmatrix}$

$\lambda = \begin{bmatrix}5\\-3\end{bmatrix}$

$\pi_u(x) = B\lambda \Rightarrow \begin{bmatrix}1 & 0\\1 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}5\\-3\end{bmatrix} = \begin{bmatrix}5\\2\\-1\end{bmatrix}$

$\pi_u(x) = \frac{b\,x}{\|b\|^2}\,b$

$10^2 \cdot 6 + 4^2 \cdot 6 + 2^2 \cdot 6$

$600 + 96 + 24$

720

$\frac{\begin{bmatrix}-10\sqrt{6}\\-4\sqrt{6}\\2\sqrt{6}\end{bmatrix} \cdot \begin{bmatrix}5\\2\\-1\end{bmatrix}}{720}\begin{bmatrix}-10\sqrt{6}\\-4\sqrt{6}\\2\sqrt{6}\end{bmatrix}$

$\frac{\cancel{60}\sqrt{6}}{\cancel{720}\,12}\begin{bmatrix}-10\sqrt{6}\\-4\sqrt{6}\\2\sqrt{6}\end{bmatrix}$ $\quad \frac{1}{12}\begin{bmatrix}-60\\-24\\12\end{bmatrix} = \begin{bmatrix}-5\\-2\\1\end{bmatrix}$

8/December/2019 Groups and Vector spaces

In the following, we will have a closer look at vector spaces i.e. the space in which vectors live.

We can informally characterize vectors as objects that can be added together and multiplied by a scalar and they remain objects of the same type.

Group is a set of elements and an operation defined on these elements, that keeps some structure of the set intact.

Groups:

Groups play an important role in computer science. Besides providing a fundamental framework for operations on sets, they are heavily used in cryptography, coding theory and graphics.

Definition 1 (Group). Consider a set G and an operation $\otimes : G \to G$ defined on G.

Then $G := (G, \otimes)$ is called a group if the following hold:

1. Closure of G under $\otimes$: $\forall x, y \in G : x \otimes y \in G$
2. Associativity: $\forall x, y, z \in G : (x \otimes y) \otimes z = x \otimes (y \otimes z)$
3. Neutral element: $\exists e \in G \, \forall x \in G : x \otimes e = x$ and $e \otimes x = x$

   For some e is an element of G for all x is an element of G such that $x \otimes e = x$ and $e \otimes x = x$
4. Inverse element: $\forall x \in G \, \exists y \in G : x \otimes y = e$ and $y \otimes x = e$

   for all x in G some y in G such that $x \otimes y = e$ and $y \otimes x = e$. We often use x^{-1} to denote the inverse element of x.

If additionally $\forall x, y \in G : x \otimes y = y \otimes x$ then $G = (G, \otimes)$ is an Abelian group (Commutative)

For example, $\otimes$ could be $+, \cdot$ defined on $\mathbb{R}, \mathbb{N}, \mathbb{Z}$ or $\cup, \cap, \setminus$ defined on $P(B)$ or the power set of B

Remark: The inverse element is defined with respect to the operation $\otimes$ and does not necessarily mean $\frac{1}{x}$.

Example

$(\mathbb{Z}, +)$ is a group

Integers $\mathbb{Z} = \{\ldots -2, -1, 0, 1, 2, \ldots\}$

$(\mathbb{N}_0, +)$ is not a group. Although $(\mathbb{N}_0, +)$ possesses a neutral element (0), the inverse elements are missing

$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$

$(\mathbb{Z}, \cdot)$ is not a group. Although $(\mathbb{Z}, \cdot)$ contains a natural neutral element (1), the inverse elements for any $z \in \mathbb{Z}$, $z \neq \pm$ are missing

$(\mathbb{R}, \cdot)$ is not a group since 0 does not possess an inverse element

Vector spaces:

When we discussed groups, we looked at sets G and inner operations on G i.e. mappings $G \times G \to G$. In the following, we will consider sets that in addition to an inner operation $+$ also contain an outer operation $\cdot$, the multiplication by a scalar $\lambda \in \mathbb{R}$

Definition 2 (Vector space) A real-valued vector space (also called an $\mathbb{R}$-vector space) is a set V with two operations

$$+ : V \times V \to V$$
$$\cdot : \mathbb{R} \times V \to V$$

1. where $(V, +)$ is an Abelian group
2. Distributivity: a) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y \quad \forall \lambda \in \mathbb{R}, x, y \in V$

   b) $(\lambda + \psi) \cdot x = \lambda \cdot x + \psi \cdot x \quad \forall \lambda, \psi \in \mathbb{R}, x \in V$
3. Associativity (outer operation): $\lambda \cdot (\psi \cdot x) = (\lambda \psi) \cdot x \quad \forall \lambda, \psi \in \mathbb{R}, x \in V$
4. Neutral element with respect to the outer operation $1 \cdot x = x, \forall x \in V$

The elements $x \in V$ are called vectors. The neutral element of $(V, +)$ is the zero vector $0 = [0, \ldots, 0]^T$, and the inner product $+$ is called vector addition. The elements $\lambda \in \mathbb{R}$ are called

Remark: A vector multiplication defined
Theoretically, we could define an element-wise multiplication
such that $c = ab$ with $c_j = a_j b_j$. This array multiplication is
common to many programming languages

24/April/2020

Orthogonal complements / Alternate Coordinate System (bases)

V some subspace of $\mathbb{R}^n$

Orthogonal complement of V is $V^{\perp} = \{\vec{x} \in \mathbb{R}^n / \vec{x} \cdot \vec{v}$ for every $\vec{v} \in V\}$

is $V^{\perp}$ a subspace ($V^{\perp}$ is pronounced V-perp)

& $\vec{a}$ & $\vec{b}$ are members of $V^{\perp}$

$\Rightarrow$ 1) $\vec{a} + \vec{b} \in V^{\perp}$

2) $c\vec{a}, c\vec{b} \in V^{\perp}$

3) $c(\vec{a} + \vec{b}) \in V^{\perp}$

$\Rightarrow \vec{a} \cdot \vec{v} = 0 \quad \forall \vec{v} \in V$

$\vec{b} \cdot \vec{v} = 0 \quad \forall \vec{v} \in V$

scalars and the outer operation $\cdot$ is a multiplication by scalars.

June - 28 - 2020

Setting of PCA :

$$X = \{x_1 \dots x_N\}, \quad x_i \in \mathbb{R}^D$$

Our objective is to find lower dimensional representation of the dataset X, which is close to the dataset X in $\mathbb{R}^D$

i) Every vector in $\mathbb{R}^D$ can be represented as linear combination of basis vectors $x_n = \sum_{i=1}^{D} \beta_{in} b_i$

Consider a basis vector b_1 with $\|b_1\| = 1$
What is the orthogonal projection $\pi_U(x)$ of a vector x onto the subspace spanned by b_1 and what are the coordinates of that projected point with respect to b_1?
Hint: The projection formula for one dimensional subspace was $\pi_U(x) = \frac{bb^T}{\|b\|^2} x$

ii) $\beta_{in} = x_n^T b_i$

iii) $B = (b_1, \dots, b_M)$

$\tilde{x} = B B^T x$

Co-ordinates or Code

$$\tilde{x}_n = \underbrace{\sum_{i=1}^{M} \beta_{in} b_i}_{\text{Principal subspace}} + \sum_{i=M+1}^{D} \beta_{in} b_i \in \mathbb{R}^D$$

Average squared reconstruction

$$J = \frac{1}{N} \sum_{n=1}^{N} \|x_n - \tilde{x}_n\|^2$$

$$\frac{\partial J}{\partial \{\beta_{in}, b_i\}} = \frac{\partial J}{\partial \tilde{x}_n} \frac{\partial \tilde{x}_n}{\partial \{\beta_{in} b_i\}}$$

$$\frac{\partial J}{\partial \tilde{x}_n} = -\frac{2}{N} (x - \tilde{x}_n)^T$$

Multivariate chain rule :

In the multivariate case when $x \in \mathbb{R}^n$, the basic differentiation rules that we know from school (e.g. sum rule, product rule, chain rule) still apply. However we need to pay attention because now we have to deal with matrices where ~~multiplic~~ multiplication is no longer ~~commutat~~ commutative i.e., the order matters.

product rule : $\frac{\partial}{\partial x}(f(x) g(x)) = \frac{\partial f}{\partial x} f(x) + f(x) \frac{\partial g}{\partial x}$

equivalently $\Rightarrow A(x) = f(x) g(x) \quad A'(x) = f(x) g'(x) + g(x) f'(x)$

(or) $(fg)' = f'g + g'f$

Sum rule :- $\frac{\partial}{\partial x}(f(x) + g(x)) = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}$

or $\Rightarrow (f+g)' = f' + g'$

Chain rule : $\frac{\partial}{\partial x}(g \circ f)(x) = \frac{\partial}{\partial x}(g(f(x)))$

$\Rightarrow (g \circ f)' = g'(f) f'$

Ex $f(x) = g(h(x))$ What is the differential of f wrt x in terms of g and h

$$\frac{dg}{dh} \times \frac{dh}{dx}$$

Let us have a closer look at chain rule. The chain rule formula resembles to some degree the rules for matrix multiplication where ~~neighbouring~~ neighbouring dimensions have to match for matrix multiplication to be defined. So we go from left to right, the chain rule behaves similar properties. It shows up in denominator

of the first factor and in the numerator of the second factor. If we multiply the factors together multiplication is defined (the dimensions of ∂f match and ∂f cancels such that $\partial/\partial x$ remains

Consider a function $f: \mathbb{R}^2 \to \mathbb{R}$ of two variables x_1, x_2. Furthermore, $x_1(t)$ and $x_2(t)$ are themselves functions of t. To compute the gradient of f with respect to t we need to apply the chain rule(s) for multivariate functions as

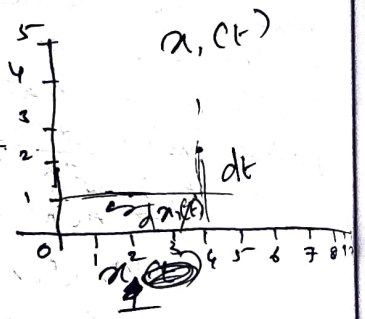
$$\frac{df}{dt} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1(t)}{\partial t} \\ \frac{\partial x_2(t)}{\partial t} \end{bmatrix} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t}$$

Where d denotes the gradient ∂ denotes the partial derivatives.

f is a function of x_1, x_2 are functions of t

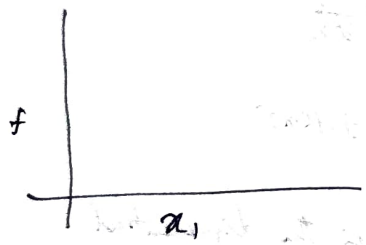
Assume $f = 2x_1$

$f = 2x_2$



$$\begin{cases} x_1 = 2t \\ x_2 = 3t \end{cases}$$ functions of t

Suppose $x_1(t) = 2t$, $x_2(t) = 3t$

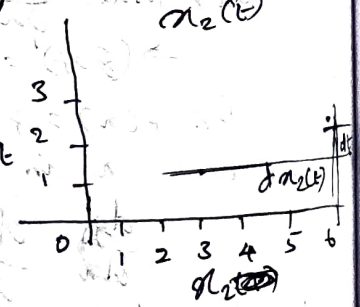


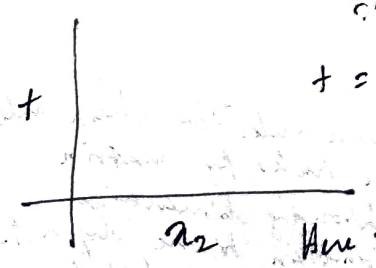
$f = 2x_1$

$x_1 = 2t$

$f = 2x_2$

$x_2 = 3t$





Here f is a function of

Consider $f = x_1^2 + 2x_2$ where $x_1 = \sin t$ and $x_2 = \cos t$ then

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t}$$

$$= 2x_1\frac{\partial x_1}{\partial t} + 2\frac{\partial x_2}{\partial t}$$

$$= 2\sin t\frac{\partial \sin t}{\partial t} + 2\frac{\partial \cos t}{\partial t}$$

$$= 2\sin t\cos t - 2\sin t = 2\sin t(\cos t - 1)$$

is the corresponding derivative of f with respect to t.

If $f(x_1, x_2)$ is a function of x_1 and x_2, when $x_1(s,t)$ and $x_2(s,t)$ are themselves functions of two variables s and t, the chain rule yields

$$\frac{df}{ds} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial s} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial s}$$

Which can be expressed as the matrix multiplication

$$\frac{df}{d(s,t)} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial(s,t)} = \underbrace{\begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}}_{\frac{\partial f}{\partial x}} \underbrace{\begin{bmatrix} \frac{\partial x_1}{\partial s} & \frac{\partial x_1}{\partial t} \\ \frac{\partial x_2}{\partial s} & \frac{\partial x_2}{\partial t} \end{bmatrix}}_{\frac{\partial x}{\partial(s,t)}}$$

$$\begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial s} & \frac{\partial x_1}{\partial t} \\ \frac{\partial x_2}{\partial s} & \frac{\partial x_2}{\partial t} \end{bmatrix}$$

Example: Gradient of linear model

Let us consider the linear model $y = \phi\theta$ where $\theta \in \mathbb{R}^D$ is a parameter vector, $\phi \in \mathbb{R}^{N \times D}$ are input features and $y \in \mathbb{R}^N$ are the corresponding observations.

$L(e) := \|e\|^2 \qquad e(\theta) := y - \phi\theta$

We seek $\frac{\partial L}{\partial \theta}$ and we will use the chain rule for this purpose.

Before we start any calculation, we determine the dimensionality of the gradient as $\frac{\partial L}{\partial \theta} \in \mathbb{R}^{1 \times D}$

The chain rule allows us to compute the gradient as

$$\frac{\partial L}{\partial \theta} = \frac{\partial L}{\partial e}\frac{\partial e}{\partial \theta}$$

We know that $\|e\|^2 = e^T e$ furthermore we obtain $\frac{\partial e}{\partial \theta} = -\phi \in \mathbb{R}^{N \times D}$

and determine $\frac{\partial L}{\partial e} = 2e^T \in \mathbb{R}^{1 \times N}$

$\frac{\partial L}{\partial \theta} = -2e^T\phi = 2(y^T - \theta^T\phi^T)\phi \in \mathbb{R}^{1 \times D}$

Consider the function $h: \mathbb{R} \to \mathbb{R}$ where $h(t) = (f \circ g)(t) = f(g(t))$

$g(t) = x = \begin{bmatrix} t\cos t \\ t\sin t \end{bmatrix}, t \in \mathbb{R}$

$f(x) = \exp(x_1 x_2^2), \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$

$g(t) = x = \begin{bmatrix} t\cos t \\ t\sin t \end{bmatrix}$

$\frac{dg}{dt} = dx = \begin{bmatrix} -t\sin t \\ t\cos t \end{bmatrix}$

$\cancel{\frac{\partial f(x)}{\partial(x_1 x_2)} \quad \frac{\partial f}{\partial g} \frac{\partial g}{\partial(x_1 x_2)}}$

$\frac{\partial f}{\partial g} \frac{\partial g}{\partial t}$

f is a function of g

g is a function of t

$\frac{df}{dt} = \frac{\partial f}{\partial g} \frac{\partial g}{\partial t}$

$\cancel{\frac{\partial g}{\partial t} = \begin{bmatrix} \cos t \\ t\cos t \end{bmatrix}}$

$\frac{\partial g}{\partial t} = \begin{bmatrix} \cos t - t\sin t \\ \sin t + t\cos t \end{bmatrix}$

$\frac{\partial f}{\partial g} = x_1 e(x_1 x_2^2) \quad 2x_2 e(x_1 x_2^2)$

$\frac{\partial f}{\partial t} = \begin{bmatrix} x_1 e(x_1 x_2^2) & 2x_2 e(x_1 x_2^2) \end{bmatrix} \begin{bmatrix} \cos t - t\sin t \\ \sin t + t\cos t \end{bmatrix}$

$\frac{\partial g}{\partial t} = \begin{bmatrix} \cos t - t\sin t \\ \sin t + t\cos t \end{bmatrix}$

$\frac{\partial f}{\partial g(x_1, x_2)} \Rightarrow \quad x_2^2 \exp(x_1 x_2^2) \quad 2x_1 x_2 \exp(x_1 x_2^2)$

$\frac{dh}{dt} = \begin{bmatrix} x_2^2 \exp(x_1 x_2^2) & 2x_1 x_2 \exp(x_1 x_2^2) \end{bmatrix} \begin{bmatrix} \cos t - t\sin t \\ \sin t + t\cos t \end{bmatrix}$

$x_2^2 \exp(x_1 x_2^2)(\cos t - t\sin t)$

$+ 2x_1 x_2 \exp(x_1 x_2^2)(\sin t + t\cos t)$

$\cos t\, x_2^2 \exp(x_1 x_2^2) - t\sin t\, x_2^2 \exp(x_1 x_2^2)$

$+ 2x_1 x_2 \exp(x_1 x_2^2) \sin t$

$+ 2x_1 x_2 \exp(x_1 x_2^2) \cos t$

2) Compute $\frac{df}{dx}$ of the following function using the chain rule

$a = x^2 \quad b = \exp(a) \quad c = a + b \quad d = \log(c) \quad e = \sin(c) \quad f = d + e$

$\cancel{\frac{db}{dx} = \frac{da}{dx} = 2x \quad \frac{db}{da} = \exp(a) \quad \frac{dc}{da} = b \quad \frac{dc}{db} = a \quad \frac{dd}{dc} = \frac{1}{c}\log(c)}$

$\cancel{\frac{de}{dc} = \cos(c)}$

$\cancel{\frac{df}{dd} = e \quad \frac{df}{de} = d}$

$a = x^2 \quad \frac{da}{dx} = 2x$

$\frac{db}{da} = \exp(a)$

$\frac{dc}{da} = 1 + b \quad \frac{dc}{db} = 1 + a$

$\frac{dd}{dc} = \frac{1}{c}\log(c)$

$\frac{de}{dc} = \cos(c)$

$\frac{df}{dd} = 1 + e$

$\frac{df}{de} = 1 + d$

compute $\frac{\partial x}{\partial x}$

$$\frac{\partial f}{\partial(s,t)} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial(s,t)} = \begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}\begin{bmatrix}\frac{\partial x_1}{\partial s} & \frac{\partial x_1}{\partial t}\\ \frac{\partial x_2}{\partial s} & \frac{\partial x_2}{\partial t}\end{bmatrix}$$

$$\frac{\partial f}{\partial d} = 1+e \qquad \frac{\partial f}{\partial e} = 1+d \qquad \frac{\partial e}{\partial c} = \cos(c)$$

$$\frac{\partial f}{\partial c} = \frac{\partial f}{\partial e}\cdot\frac{\partial e}{\partial c} = (1+d)\cdot\cos(c)$$

$$\frac{\partial f}{\partial c} = \frac{\partial f}{\partial d}\cdot\frac{\partial d}{\partial c} = (1+e)\cdot\frac{1}{c}\log(c)$$

$$\frac{\partial c}{\partial a} = 1+b \qquad \frac{\partial c}{\partial b} = 1+a$$

$$\frac{\partial b}{\partial x} = \exp(x) \qquad \frac{\partial a}{\partial x} = 2x$$

$$\cancel{\frac{\partial f}{\partial(d,e)} = \frac{\partial f}{\partial c}\frac{\partial c}{\partial(d,e)}} =$$

$$(1+d)\cdot\frac{1}{c}\log(c) = (1+e)\cdot\frac{1}{c}\log(c)$$

$$\frac{\partial c}{\partial x} = \frac{\partial c}{\partial a}\cdot\frac{\partial a}{\partial x} = (1+a)\,2x$$

$$\frac{\partial c}{\partial b}\cdot\frac{\partial b}{\partial a} = (1+a)\cdot\exp(x)$$

$$\cancel{\frac{df}{dt} = \begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}\begin{bmatrix}\frac{\partial x_1(t)}{\partial t}\\ \frac{\partial x_2(t)}{\partial t}\end{bmatrix}}$$

$$\cancel{\frac{\partial f}{\partial c} = \begin{bmatrix}\frac{\partial f}{\partial e} & \frac{\partial f}{\partial d}\end{bmatrix}\begin{bmatrix}\frac{\partial f}{\partial c}\\ \frac{\partial d}{\partial c}\end{bmatrix}}$$

$$\frac{\partial f}{\partial t} = \begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}\begin{bmatrix}\frac{\partial x_1(t)}{\partial t}\\ \frac{\partial x_2(t)}{\partial t}\end{bmatrix}$$

$$\frac{\partial f}{\partial x}\frac{\partial x}{\partial(s,t)} = \begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}\begin{bmatrix}\frac{\partial x_1}{\partial s} & \frac{\partial x_1}{\partial t}\\ \frac{\partial x_2}{\partial s} & \frac{\partial x_2}{\partial t}\end{bmatrix}$$

$$\frac{\partial f}{\partial c}\frac{\partial c}{\partial(d,e)} = \begin{bmatrix}\frac{\partial f}{\partial c} & \frac{\partial f}{\partial c}\end{bmatrix}\begin{bmatrix}\frac{\partial c}{\partial c}\\ \frac{\partial e}{\partial e}\end{bmatrix}$$

$$\cancel{\frac{\partial f}{\partial c}\frac{\partial d}{\partial c} + \frac{\partial f}{\partial c}\frac{\partial e}{\partial c}}$$

$$\frac{\partial f}{\partial c}\frac{\partial c}{\partial d} + \frac{\partial f}{\partial c}\frac{\partial c}{\partial e}$$

Start gradients next part

Compute the function $\frac{df}{dx}$ using the chain rule

$a = 2x^2$ $b = \exp(a)$ $c = a+b$ $d = \log(c)$ $e = \sin(c)$ $f = d+e$

$\frac{\partial a}{\partial x} = 2x$ $\frac{\partial b}{\partial a} = \exp(a)$ $\frac{\partial c}{\partial a} = 1+b$ $\frac{\partial c}{\partial b} = 1+a$ $\frac{\partial d}{\partial c} = \frac{1}{c}\log(c)$

$\frac{\partial e}{\partial c} = \cos(c)$

$$\frac{\partial f}{\partial t} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1(t)}{\partial t} \\ \frac{\partial x_2(t)}{\partial t} \end{bmatrix}$$

$$\frac{\partial f}{\partial x}\frac{\partial x}{\partial(s,t)} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial s} & \frac{\partial x_1}{\partial t} \\ \frac{\partial x_2}{\partial s} & \frac{\partial x_2}{\partial t} \end{bmatrix}$$

$\frac{\partial f}{\partial d} = 1+e$

$\frac{\partial f}{\partial e} = 1+d$

$$\cancel{\frac{\partial f}{\partial e} \cdot \frac{\partial e}{\partial c} \cdot \frac{\partial c}{\partial a} \cdot \frac{\partial a}{\partial x}}$$

$$\cancel{(1+d)\cos(c)\cdot(1+b)\cdot 2x}$$

$$\cancel{2x\cos(c)(1+d)(1+b)}$$

$\cancel{\frac{\partial e}{\partial t} \cdot \frac{\partial f}{\partial(d,e)}} =$

f is a function of $d+e$

e is a function of c

d is a function of c

$$\frac{df}{\partial(d,e)} = \frac{\partial f}{\partial d}\cdot\frac{\partial d}{\partial c} + \frac{\partial f}{\partial e}\cdot\frac{\partial e}{\partial c}$$

$$(1+e)\cdot\frac{1}{c}\cdot\log(c) + (1+d)\cos(c)$$

c is a function of a, b

a is a function of x

b is a function of a

$$\frac{\partial c}{\partial(a,b)} = \frac{\partial c}{\partial a}\cdot\frac{\partial a}{\partial x} + \frac{\partial c}{\partial b}\cdot\frac{\partial b}{\partial a}$$

$$(1+b)\cdot 2x + (1+a)\cdot\exp(a)$$

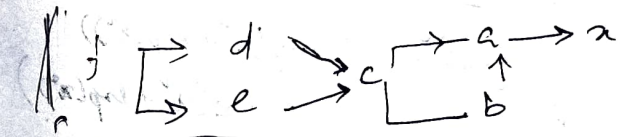
$$\begin{bmatrix} \frac{\partial f}{\partial d} & \frac{\partial f}{\partial e} \end{bmatrix} \begin{bmatrix} \frac{\partial d}{\partial c} \\ \frac{\partial e}{\partial c} \end{bmatrix} = \frac{\partial f}{\partial(d,e)}$$

$$\frac{\partial f}{\partial(d,e)} = \begin{bmatrix} \frac{\partial f}{\partial d} & \frac{\partial f}{\partial e} \end{bmatrix} \begin{bmatrix} \frac{\partial d}{\partial c} \\ \frac{\partial e}{\partial c} \end{bmatrix} \Rightarrow \frac{\partial f}{\partial d}\cdot\frac{\partial d}{\partial c} + \frac{\partial f}{\partial e}\cdot\frac{\partial e}{\partial c}$$

$$\frac{\partial c}{\partial(a,b)} = \begin{bmatrix} \frac{\partial c}{\partial a} & \frac{\partial c}{\partial b} \end{bmatrix} \begin{bmatrix} \partial a/\partial x \\ \partial b/\partial a \end{bmatrix} \Rightarrow \frac{\partial c}{\partial a}\cdot\frac{\partial a}{\partial x} + \frac{\partial c}{\partial b}\cdot\frac{\partial b}{\partial a}$$

$$\cancel{\frac{\partial f}{\partial(d,e)} = \begin{bmatrix} \frac{\partial f}{\partial a} & \frac{\partial c}{\partial b} \end{bmatrix} \begin{bmatrix} \partial a & \partial a \end{bmatrix}}$$

$\cancel{\frac{\partial c}{\partial(a,b)}}$



$\frac{\partial f}{\partial(d,e)}$

$\frac{\partial c}{\partial(a,b)}$

$\cancel{\frac{\partial c}{\partial a}\frac{\partial a}{\partial b}}$

$$\begin{bmatrix} \frac{\partial c}{\partial a} & \frac{\partial c}{\partial b} \end{bmatrix} \begin{bmatrix} \frac{\partial a}{\partial x} \\ \frac{\partial b}{\partial a} \end{bmatrix}$$

$$\frac{\partial f}{\partial(d,e)} = (1+e)\cdot\frac{1}{c}\cdot\log(c) + (1+d)\cos(c)$$

$$\frac{\partial f}{\partial d}\times\frac{\partial d}{\partial c} = \frac{\partial f}{\partial c}$$

$1+e$

$$\frac{\partial c}{\partial(a,b)}\cdot\frac{\partial f}{\partial(d,e)} =$$

$$\frac{\partial f}{\partial c} = (1+e)\cdot\frac{1}{c}\log(c)$$

$$\frac{\partial f}{\partial a} = \frac{\partial f}{\partial c}\cdot\frac{\partial c}{\partial a} = (1+e)\cdot\frac{1}{c}\log(c)(1+b)$$

$$\frac{\partial f}{\partial a}\cdot\frac{\partial a}{\partial x} = (1+e)\cdot\frac{1}{c}\log(c)(1+b)\cdot 2x$$

$$\frac{\partial f}{\partial x} = (1+e)\cdot\frac{1}{c}\log(c)\,(1+b)\cdot 2x$$

$$(1+\sin(c))\cdot\frac{1}{a+b}\log(a+b)\,(1+\exp(a))\cdot 2x$$

$$(1+\sin(c))\cdot\frac{1}{x^2+\exp(a)}\log(x^2+\exp(a))\,(1+\exp(a))\cdot 2x$$

Compute $\frac{df}{dx}$ of the following function using the chain rule

$a = x^2$

$b = \exp(a)$

$c = a+b$

$d = \log(c)$

$e = \sin(c)$

$f = d+e$

$f = \log(c) + \sin(c)$

$f = \log(a+b) + \sin(a+b)$

$f = \log(x^2+\exp(x^2)) + \sin(x^2+\exp(x^2))$

$\cancel{= \log(x^2) + \log(\exp(x^2))}$

$= \log(x^2+\exp(x^2)) + \sin(x^2+\exp(x^2))$

$f = \log(x^2) + \log(\exp(x^2)) + \sin(x^2) + \sin(\exp(x^2))$

$\cancel{\frac{\partial f}{\partial x} = \frac{1}{x^2}\log(x^2) + \frac{1}{\exp(x^2)}\log(\exp(x^2)) + \cos(x^2) + \cos(\exp(x^2))}$

$\cancel{\frac{\partial f}{\partial x} = \frac{1}{x^2}\log(x^2) + \exp(x^2)}$

$\frac{\partial f}{\partial x} = \cancel{\frac{2}{x} + 2x +}$

$\frac{2}{x} + 2x + 2x\cos(x^2) + 2xe^{x^2}\cos(e^{x^2})$

$f = \log(x^2+\exp(x^2)) + \sin(x^2+\exp(x^2))$

$= \log(x^2) + \log(\exp(x^2)) + \sin(x^2)\cdot 1\,\sin(\exp(x^2))$

$= 2\times\frac{1}{x} +$

$+ 2x\cos(x^2) +$

What is $\frac{df}{dx}$ where

$f = \sin(t^2)$ $\quad t = x^3$

$$\frac{\partial f}{\partial t} = -2t\sin(t^2)\frac{\partial t}{\partial x} = 3x^2$$

$$\frac{\partial f}{\partial t}\cdot\frac{\partial t}{\partial x} = -2t\sin(t^2)\cdot 3x^2$$

$$= -6x^5\sin(x^6)$$

(A) $\tilde{x}_n = \sum_{j=1}^{M} \beta_{jn} b_j$

(B) $J = \frac{1}{N}\sum_{n=1}^{N} \| x_n - \tilde{x}_n \|^2$

(C) $\frac{\partial J}{\partial \tilde{x}_n} = \frac{-2}{N}(x_n - \tilde{x}_n)^T$

(D) $\beta_{jn} = x_n^T b_j, \; j = 1, \dots, M$

Problem setting and PCA objective:

$X = \{x_1 \dots x_N\}$ $x_i \in \mathbb{R}^D$ our objective is to find the lower dimensional representation of data but as similar as x.

I: Every vector in $\mathbb{R}^D$ can be represented as linear combination of basis vectors

(1) $x_n = \sum_{i=1}^{D} \beta_{in} b_i$ where b_i will be orthonormal basis of $\mathbb{R}^D$

(2) $\beta_{in} = x_n^T b_i$

(3) $B = (b_1, \dots, b_M)$

$\tilde{x} = BB^T x$ $\quad$ $\tilde{x}$ is the orthogonal projection of x onto the subspace spanned by my basis vectors

---

$\|b_1\| = 1$

What is the orthogonal projection $\pi_u(x)$ of vector x onto the subspace U spanned by b_1 and what are the coordinates of that projected point with respect to b_1?

Hint: The projection formula for one dimensional subspaces is

$$\pi_u(x) = \frac{bb^T}{\|b\|^2}x$$

Orthogonal Decomposition Theorem

Definition: The orthogonal complement of a subspace W is

$$W^\perp = \{\vec{x} \mid \vec{x}\cdot\vec{y} = 0 \;\forall \vec{y} \in W\}$$

Suppose you have an orthogonal basis for your subspace W. Also suppose $\vec{x}\cdot\vec{y} = 0 \;\forall i$

$\vec{y} \in W$ $\quad \vec{y} = a_1\vec{u}_1 + \dots + a_n\vec{u}_n$

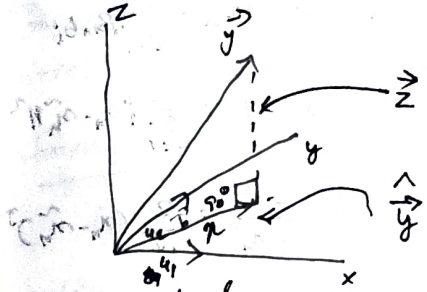
$\vec{x}\cdot\vec{y} = a_1\vec{x}\cdot\vec{u}_1 + \dots + a_n\vec{u}_n\vec{x}_n$

$= 0$

Let $u_1 \dots u_p$ be an orthogonal basis for $W \subset \mathbb{R}^n$

Let $\vec{y} \in \mathbb{R}^n$ $\quad \vec{y} = \hat{y} + \vec{z}$ $\quad \hat{y} \in W$ $\quad \vec{z} \in W^\perp$

Orthogonal decomposition states that a vector can be decomposed with something in W and something in $W^\perp$



Thing of $\vec{y}$ as something projected onto W

W = xy plane

$\vec{y} = \hat{y} + \vec{z}$

Orthogonal decomposition states that a vector $\vec{y}$ can be decomposed with vector in a plane W and $W^\perp \Rightarrow \hat{y} + \vec{z}$

let $\cancel{\hat{y} = \frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}u_1}$ $\in W$

$$\hat{y} = \frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \dots + \frac{\vec{y}\cdot\vec{u}_p}{\vec{u}_p\cdot\vec{u}_p}\vec{u}_p$$

Recall

Orthogonal projection of $\vec{y}$ on $\vec{u}_1$ basis vector

Orthogonal projection of $\vec{y}$ on $\vec{u}_p$ basis vector

$\hat{y}$ is W since $\vec{u}_1 \dots \vec{u}_p$ are the basis vectors in W

$\cancel{\vec{y} = \lambda x}$

$\vec{b}\cdot\vec{y}$

$$\pi_u(x) = \frac{\vec{b}x}{\vec{b}\vec{b}}\cdot\vec{b} \quad (\text{or}) \quad \frac{\vec{b}\cdot x}{\|b\|^2}\vec{b}$$

$\pi_u(x) = \hat{y}$

$\vec{z} = \vec{y} - \hat{y}$ $\quad$ is $\vec{z} \in W^\perp$

$\vec{z}\cdot\vec{u}_i = 0 \;\forall i$?

$$\vec{z}\cdot\vec{u}_i = \vec{y} - \hat{y}$$

$$= \vec{y}\cdot\vec{u}_i - \hat{y}\cdot\vec{u}_i$$

$$= \vec{y}\cdot\vec{u}_i - \left(\frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \dots + \frac{\vec{y}\cdot\vec{u}_p}{\vec{u}_p\cdot\vec{u}_p}\vec{u}_p\right)\cdot\vec{u}_i$$

$$= \vec{y}\cdot\vec{u}_i - \frac{\vec{y}\cdot\vec{u}_i}{\vec{u}_i\cdot\vec{u}_i}\cancel{\vec{u}_i\cdot\vec{u}_i}$$

$$= 0$$

Reason

$\frac{\vec{y}\cdot\vec{u}_p}{\vec{u}_p\cdot\vec{u}_p}\vec{u}_p\cdot\vec{u}_i = 0$ in the dot product will be all will be 0

Finding the co-ordinates of the projected ... $x_n = \sum_{i=1}^{N} \beta_{in} b_i$

$E(x) = 0$, $b_1 \dots b_N$ [basis vector from an orthonormal basis]

(B) $J = \sum_{i=1}^{N} \|x_n - \tilde{x}_n\|^2$

(C) $\frac{\partial J}{\partial \tilde{x}_n} = \frac{-2}{N}(x_n - \tilde{x}_n)$

$$\frac{\partial J}{\partial \beta_{in}} = \frac{\partial J}{\partial \tilde{x}_n} \cdot \frac{\partial \tilde{x}_n}{\partial \beta_{in}}$$

$$\frac{\partial \tilde{x}_n}{\partial \beta_{in}} = b_i, \quad \forall i = 1 \dots M$$

$$\frac{\partial J}{\partial \beta_{in}} = \frac{-2}{N}(x_n - \tilde{x}_n)^T b_i$$

$$= \frac{-2}{N}\left(x_n - \sum_{j=1}^{M} \beta_{jn} b_j\right)^T b_i = \frac{-2}{N}(x_n^T b_i - \beta_{in} b_i^T b_i)$$

$$= \frac{-2}{N}(x_n^T b_i - \beta_{in})$$

$\beta_{jn} b_j \cdot b_i = 0$ Only with i it will end up as 1

$= 0$ (D)

$\iff \beta_{in} = x_n^T b_i$ (The optimal co-ordinates of $\tilde{x}_n$ wrt basis b_i are the orthogonal projections of co-ordinates of our original basis vector that span the principal subspace

(A)

$$\tilde{x}_n = \sum_{j=1}^{N} \beta_{jn} b_j$$

$$= \sum_{j=1}^{M} x_n^T b_j = \sum_{j=1}^{M} b_j(b_j^T x_n) = \sum_{j=1}^{M} \underbrace{(b_j b_j^T)}_{\text{projection matrix}} x_n$$

$$x_n = \underbrace{\sum_{j=1}^{M} (b_j b_j^T) x_n}_{\text{projection on principal subspace}} + \underbrace{\sum_{j=M+1}^{M} (b_j b_j^T) x_n}_{\text{projection onto orthogonal complement}}$$

$$x_n - \tilde{x}_n = \sum_{j=M+1}^{D} (b_j b_j^T) x_n = \sum_{j=M+1}^{D} (b_j^T x_n) b_j \quad \text{(E)}$$

$$J = \frac{1}{N} \sum_{n=1}^{N} \|x_n - \tilde{x}_n\|^2 \quad \text{(Average square reconstruction error)}$$

(B) $$= \frac{1}{N} \sum_{n=1}^{N} \left\| \sum_{j=M+1}^{D} (b_j^T x_n) b_j \right\|^2 \overset{\text{orth}}{=} \frac{1}{N} \sum_{n=1}^{N} \sum_{j=M+1}^{D} (b_j^T x_n)^2 = \frac{1}{N} \sum_n \sum_j b_j^T x_n x_n^T b_j$$

$$= \sum_{j=M+1}^{D} b_j^T \underbrace{\frac{1}{N} \sum_{n=1}^{N} x_n x_n^T}_{\Rightarrow S \Rightarrow \text{Data covariance matrix}} b_j$$

$\sum_{j=M+1}^{D} b_j^T S b_j = \text{trace}\left(\sum_{j=M+1}^{D} b_j b_j^T S\right)$ projection matrix

projection matrix projects the covariance matrix on the orthogonal ~~subspace~~. Complement of the principal subspace.

Finding the basis vectors that span the principal subspace

Minimizing the average square reconstruction error is equivalent to minimizing the projection of the variance of the ~~variance of the~~ data by projecting onto the subspace that we ignore in PCA.

We can exploit this insight and determine an orthonormal basis of the M dimensional principal subspace.

Steps of PCA

1) Subtract the mean. That data is centered.
2) Divide by the standard deviation. Now the data is unit (cm or meter) free and has variance 1 along each axis. Now the correlations are still intact.
3) We compute the data covariance matrix, its eigen values and its corresponding eigen vectors. The eigen vectors are scaled by corresponding eigen values.
4) We can project any datapoint x^* onto principal subspace that is spanned by ~~largest~~ eigen vectors that belongs to the largest eigenvalues
5) We need to normalize x^* using the mean and standard deviation of the data set that is used to compute the data covariance matrix.

$$x^* = x_*$$

$$x_*^{(d)} \leftarrow \frac{x_*^{(d)} - \mu^{(d)}}{\sigma^{(d)}}$$

$$\tilde{x}_* = \pi_U(x_*) = B B_T x_*$$

B is the matrix that contains the eigen vectors with the largest eigen values as columns

$B_T x_*$ are the coordinates of the projections (wrt the basis) of principal subspace

PCA in high dimensions:

In order to do PCA we need to compute the data covariance matrix. In "d" dimensions the data covariance matrix is a $d \times d$ matrix. If 'd' is very high then computing eigen vectors and eigen values can become very expensive. It gets cubically in terms of rows and columns, in this case the number of dimensions

$x_1, \ldots, x_N \in \mathbb{R}^D$ we assume the data is centered i.e. $\mu = 0$

$$S = \frac{1}{N} X^T X$$

$$X = \begin{bmatrix} x_1^T \\ \vdots \\ x_N^T \end{bmatrix} \in \mathbb{R}^{N \times D} \Longrightarrow \text{Rank of } S = N$$

$\text{rank}(S) = N$

$\text{rank}(S) = D - N + 1$, if eigenvalues are

If N is significantly smaller than D, i.e. if the number of datapoints is significantly smaller than the dimensionality of the data.