Vectors

Tuesday, September 12, 2023 2:49 PM

I) Solve: $\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \\ 3 \end{bmatrix}$

II) If $u = \begin{bmatrix} 2 \\ -7 \\ 1 \end{bmatrix}$ $v = \begin{bmatrix} -3 \\ 0 \\ 4 \end{bmatrix}$ $w = \begin{bmatrix} 0 \\ 5 \\ -8 \end{bmatrix}$, Find

1) $3u - 4v = \begin{bmatrix} 18 \\ -21 \\ -13 \end{bmatrix}$

2) $2u + 3v - 5w = \begin{bmatrix} -5 \\ -39 \\ 54 \end{bmatrix}$

III) Find x & y if $\begin{bmatrix} x \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ x+y \end{bmatrix}$

$x = 2$; $3 = x + y \Rightarrow 3 = y + 2$

$y = 1$

1.8 Compute $u \cdot v$ for the following

(i) $u = \begin{bmatrix} 2 \\ -3 \\ 6 \end{bmatrix}$ $v = \begin{bmatrix} 8 \\ 2 \\ -3 \end{bmatrix}$ $u \cdot v = -8$

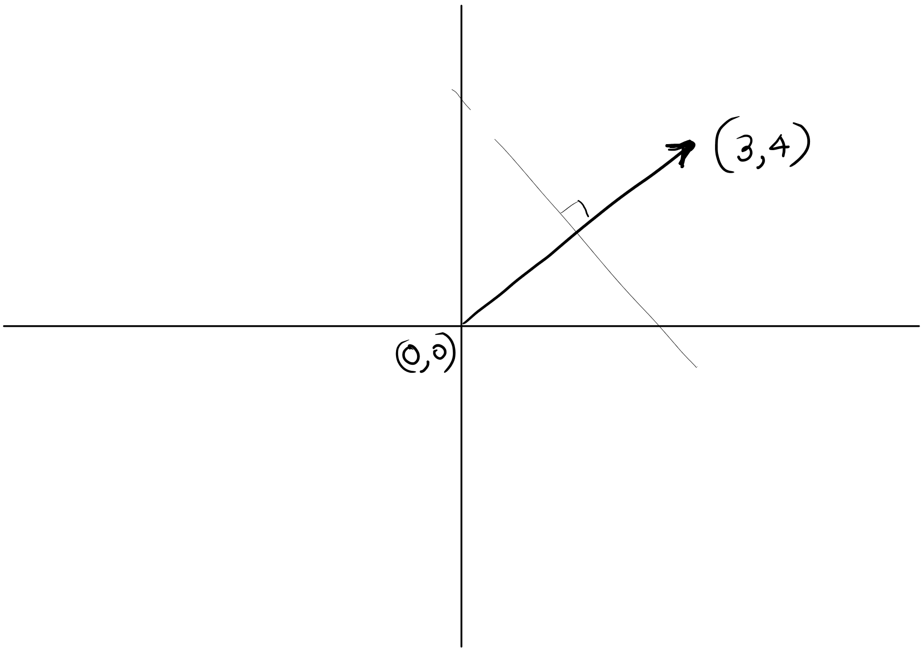
(ii) $u = \begin{bmatrix} 3 \\ -5 \\ 2 \\ 1 \end{bmatrix}$ $v = \begin{bmatrix} 4 \\ 1 \\ -2 \\ 5 \end{bmatrix}$ $u \cdot v = 8$

1.36 $u = \begin{bmatrix} 2 \\ 1 \\ -3 \\ 0 \\ 4 \end{bmatrix}$ $v = \begin{bmatrix} 5 \\ -3 \\ -1 \\ 2 \\ 7 \end{bmatrix}$

Find (i) $u \cdot v = 38$

(ii) $\|u\| = \sqrt{\sum_i u_i^2} = \sqrt{30}$

(iii) $\|v\| = \sqrt{\sum_i v_i^2} = 2\sqrt{22}$



1.37. Determine K so that the vectors u & v are orthogonal

(i) $u = \begin{bmatrix} 3 \\ k \\ -2 \end{bmatrix}$ $v = \begin{bmatrix} 6 \\ -4 \\ -3 \end{bmatrix}$

$u \cdot v = 0$

$3(6) + k(-4) + (-2)(-3) = 0$

$k = 6$

(ii) $u = \begin{bmatrix} 5 \\ k \\ -4 \\ 2 \end{bmatrix}$ $v = \begin{bmatrix} 1 \\ -3 \\ 2 \\ 2k \end{bmatrix}$ $k = 3$

(iii) $u = \begin{bmatrix} 1 \\ 7 \\ k+2 \\ -2 \end{bmatrix}$ $v = \begin{bmatrix} 3 \\ k \\ -3 \\ k \end{bmatrix}$ $k = 1.5$

l_2-norm Normalization of Vector:-

$$\|v\|_2 = \sqrt{6^2 + 4^2 + 3^2} = \sqrt{61}$$

$v = \begin{bmatrix} 6 \\ -4 \\ -3 \end{bmatrix}$

$$\left\| \frac{1}{\sqrt{61}} \begin{bmatrix} 6 \\ -4 \\ -3 \end{bmatrix} \right\| = 1$$

l_2-norm

$[v_1, v_2 \, . \, . \, . \, . \, v_k]$

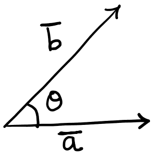
$$\|v\|_2 = \sqrt{\sum_{i=1}^{k} v_i^2}$$

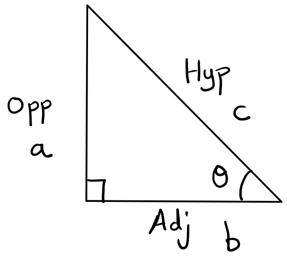
l_1-norm $\|v\|_1 = |6| + |-4| + |-3|$

$$\|v\|_1 = \sum^{k} |v_i|$$

l_1 - norm $\quad \|V\|_1 = |6| + |-4| + |-3| \qquad \|v\|_1 = \sum_{i=1}^{k} |v_i|$

Cosine Similarity





$\sin\theta = \frac{a}{c}$

$\cos\theta = \frac{b}{c}$

$\tan\theta = \frac{a}{b}$

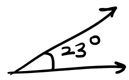
$\bar{a} \cdot \bar{b} = |a| \, |b| \cos\theta$

$\therefore \quad \cos\theta = \dfrac{\bar{a} \cdot \bar{b}}{|a| \, |b|}$

e.g. $\quad \bar{a} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \quad \bar{b} = \begin{bmatrix} 2 \\ 7 \\ 10 \end{bmatrix}$

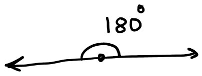
$$\cos\theta = \frac{\bar{a} \cdot \bar{b}}{|a| \, |b|} = \frac{1(2) + 3(7) + 2(10)}{\sqrt{1^2+3^2+2^2} \, \sqrt{2^2+7^2+10^2}} = 0.92$$

$$\theta = \cos^{-1}(0.92) = 23^\circ$$



$\bar{a} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \quad \bar{b} = \begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} \qquad \cos\theta = 1 \Rightarrow \theta = 0^\circ$

$\bar{a} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \quad \bar{b} = \begin{bmatrix} -1 \\ -3 \\ -2 \end{bmatrix} \qquad \cos\theta = -1 \Rightarrow \theta = 180^\circ$



User-Based Collaborative Filtering

	Item 1	Item 2	Item 3	Item 4	Item 5
User 6	1	4	10	2	NA

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	4	6	7	2	3
User 2	5	1	1	10	5
User 3	8	10	9	1	7
User 4	10	2	2	4	9
User 5	2	4	8	9	2

0.904
0.314
0.811
0.375
0.82

Similar
(1) Exp $\rightarrow 3$
(2) Exp $\rightarrow \frac{2+3}{2} = 2.5$
(3) Exp $\rightarrow \frac{2+3+7}{3} = 4$

Item-Based Collaborative Filtering

	Item 1	Item 2	Item 3	Item 4	Item 5
Item 1	1	0.37	-0.17	0.013	-0.57
Item 2	0.37	1	-0.22	0.908	0.443
Item 3	-0.17	-0.22	1	0.0543	-0.428
Item 4	0.013	0.908	0.0543	1	0.581
Item 5	-0.57	0.443	-0.428	0.581	1

Item 1	Item 2	Item 3	Item 4	Item 5
-0.89	0.2	0.3	0.4	0.34
0.2	0.4	-0.9	0.23	0.3
0.8	0.5	0.3	0.4	-0.1

thres = 0.1 rating > 0.1

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	4	6	?	?	3
User 2	5	?	1	?	5
User 3	8	10	?	1	7
User 4	?	2	2	4	?
User 5	2	?	8	9	2

	Item 1	Item 2	Item 3	Item 4	Item 5
User 6	1	4	?	2	?

(1) −0.17
(2) −0.22
(4) 0.0543

(1) −0.57
(2) 0.443 ✓
(4) 0.581 ✓

$$\text{Exp rating} = \frac{\sum r_i s_i}{\sum s_i} = \frac{4(0.443) + 2(0.581)}{0.443 + 0.581}$$

$$= 2.86 \approx 3$$

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	4	6	?	?	3

(1) −0.17 | 0.013 ✗
(2) −0.22 | 0.908 ✓
(5) −0.428 | 0.581 ✓

$$\text{Exp rating} = \frac{\sum r_i s_i}{\sum s_i} = \frac{6(0.908) + 3(0.581)}{(0.908 + 0.581)}$$

$$= 4.82 \approx 5$$

	Item 1	Item 2	Item 3	Item 4	Item 5
Item 1	1	0.37	-0.17	0.013	-0.57
Item 2	0.37	1	-0.22	0.908	0.443
Item 3	-0.17	-0.22	1	0.0543	-0.428
Item 4	0.013	0.908	0.0543	1	0.581
Item 5	-0.57	0.443	-0.428	0.581	1

	Item 1	Item 2	Item 3	Item 4	Item 5
User 2	5	?	1	?	5

Exp. rating = 5 ,

0.013 (1)
0.0543 (3)
✓ 0.581 (5)

$$\frac{0.581(5)}{0.581} = 5$$

	Item 1	Item 2	Item 3	Item 4	Item 5
User 5	2	?	8	9	2

Exp Rating = 5.7

1.1 Definition A *vector space* (over $\mathbb{R}$) consists of a set V along with two operations '+' and '·' subject to the conditions that for all vectors $\vec{v}, \vec{w}, \vec{u} \in V$ and all *scalars* $r, s \in \mathbb{R}$:

(1) the set V is closed under vector addition, that is, $\vec{v} + \vec{w} \in V$
(2) vector addition is commutative, $\vec{v} + \vec{w} = \vec{w} + \vec{v}$
(3) vector addition is associative, $(\vec{v} + \vec{w}) + \vec{u} = \vec{v} + (\vec{w} + \vec{u})$
(4) there is a *zero vector* $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$ for all $\vec{v} \in V$
(5) each $\vec{v} \in V$ has an *additive inverse* $\vec{w} \in V$ such that $\vec{w} + \vec{v} = \vec{0}$
(6) the set V is closed under scalar multiplication, that is, $r \cdot \vec{v} \in V$
(7) scalar multiplication distributes over scalar addition, $(r + s) \cdot \vec{v} = r \cdot \vec{v} + s \cdot \vec{v}$
(8) scalar multiplication distributes over vector addition, $r \cdot (\vec{v} + \vec{w}) = r \cdot \vec{v} + r \cdot \vec{w}$
(9) ordinary multiplication of scalars associates with scalar multiplication, $(rs) \cdot \vec{v} = r \cdot (s \cdot \vec{v})$
(10) multiplication by the scalar 1 is the identity operation, $1 \cdot \vec{v} = \vec{v}$.

[Linear combinations, span, and basis vectors | Chapter 2, Essence of linear algebra](#)

Matrices

Tuesday, September 12, 2023 3:42 PM

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

$ab_{21} = a_{21}b_{11} + a_{22}b_{21}$

3.6) $A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}_{2\times2}$, $B = \begin{bmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{bmatrix}_{2\times3}$ $AB = \begin{bmatrix} 11 & -6 & 14 \\ 1 & 2 & -14 \end{bmatrix}$

$ab_{ij} = \sum_k a_{ik} b_{kj}$

3.7) $A = \begin{bmatrix} 2 & 1 \end{bmatrix}_{1\times2}$ $B = \begin{bmatrix} 1 & -2 & 0 \\ 4 & 5 & -3 \end{bmatrix}_{2\times3}$ $AB = \begin{bmatrix} 6 & 1 & -3 \end{bmatrix}$

3.51) Let $B = \begin{bmatrix} 1 & 3 \\ 5 & 3 \end{bmatrix}$ $B * B = \begin{bmatrix} 1 & 3 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 5 & 3 \end{bmatrix}$

(i) If $f(x) = 2x^2 - 4x + 3$; find $f(B)$ $\begin{bmatrix} 31 & 12 \\ 20 & 39 \end{bmatrix}$

(ii) If $g(x) = x^2 - 4x - 12$; find $g(B)$ $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Determinant of Squared matrix :-

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ det = $ad - bc$ $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$

$\begin{bmatrix} a & b & c \\ e & f & g \\ h & i & j \end{bmatrix}$ $\text{del} = a\begin{vmatrix} f & g \\ i & j \end{vmatrix} - b\begin{vmatrix} e & g \\ h & j \end{vmatrix} + c\begin{vmatrix} e & f \\ h & i \end{vmatrix}$

$= a(fj - gi) - b(ej - gh) + c(ei - fh)$

Determinants :-

1) $\begin{bmatrix} 2 & 1 \\ 4 & 0 \end{bmatrix}$ 2) $\begin{bmatrix} 3 & 2 & 0 \\ 2 & 9 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ 3) $\begin{bmatrix} 2 & -7 & -2 \\ -3 & 4 & 3 \\ 14 & 2 & -2 \end{bmatrix}$

-4 45 -156

Determine value of k for which $\begin{vmatrix} k & k \\ 4 & 2k \end{vmatrix} = 0$

$2k^2 - 4k = 0 \Rightarrow k = 0$ or 2

Inverse of Squared Matrix :- Matrix $B_{n\times n}$ is inverse of $A_{n\times n}$

if $AB = BA = I_{n\times n}$ $I = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{n\times n}$

Inverse A is denoted by A^{-1}

$\begin{bmatrix} 3 & 5 \end{bmatrix}$

Inverse A is denoted by A^{-1}

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{n\times n}$$

1) $A = \begin{bmatrix} 3 & 5 \\ 2 & 3 \end{bmatrix}$;

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

$$\begin{bmatrix} 3 & 5 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3a+5c & 3b+5d \\ 2a+3c & 2b+3d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$3a+5c = 1$$
$$3b+5d = 0$$
$$2a+3c = 0$$
$$2b+3d = 1$$

$$A_{n\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$A = \begin{array}{c} \begin{matrix} a_{11} & a_{12} \end{matrix} \\ \begin{bmatrix} 3 & 5 \\ 2 & 3 \end{bmatrix} \\ \begin{matrix} a_{21} & a_{22} \end{matrix} \end{array}$$

$\text{minor}(a_{11}) = 3$ $\quad$ $\text{minor}(a_{12}) = 2$

$\text{minor}(a_{21}) = 5$ $\quad$ $\text{minor}(a_{22}) = 3$

$$\text{cofactor}(a_{ij}) = (-1)^{i+j} \text{minor}(a_{ij})$$

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

$\text{cof}(a_{11}) = 3$ $\quad$ $\text{cof}(a_{12}) = -2$

$\text{cof}(a_{21}) = (-1)^{2+1} 5 = -5$; $\text{cof}(a_{22}) = 3$

$$\text{Matrix of Cofactors} = \begin{bmatrix} 3 & -2 \\ -5 & 3 \end{bmatrix}$$

Adjoint matrix is transpose of matrix of cofactors.

$$\text{Adj}(A) = \begin{bmatrix} 3 & -5 \\ -2 & 3 \end{bmatrix}$$

$|A| = \det(A)$

Theorem :- For any Squared matrix $A_{n\times n}$;

$$A \cdot \text{Adjoint}(A) = \text{Adjoint}(A) \cdot A = |A| \, I_{n\times n}$$

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

$$A\,A^{-1} = I$$
$$\mathrm{Adj}(A)\,A\,A^{-1} = \mathrm{Adj}(A)$$
$$|A|\,I\,A^{-1} = \mathrm{Adj}(A)$$
$$A^{-1} = \frac{1}{|A|}\mathrm{Adj}(A)$$

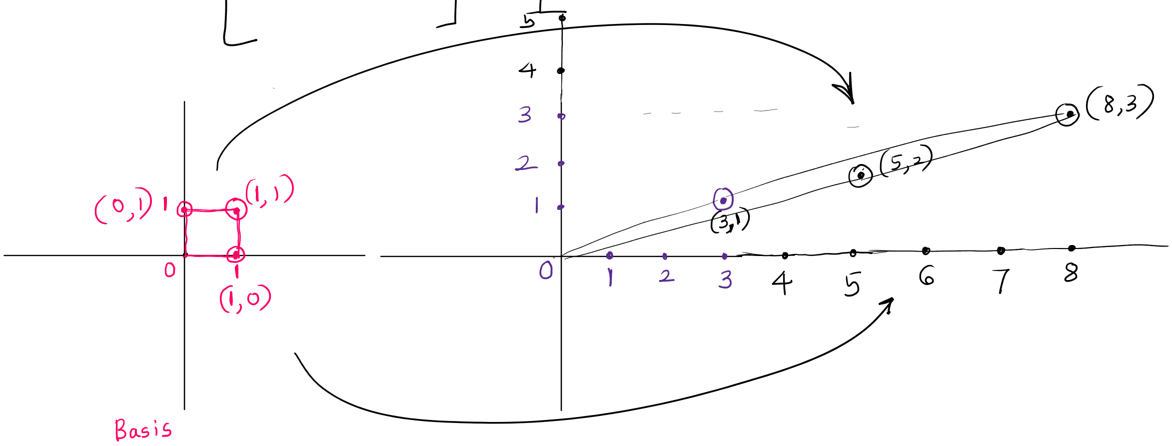
a) $\begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix}$

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

b) $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix}$

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

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



Basis

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

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

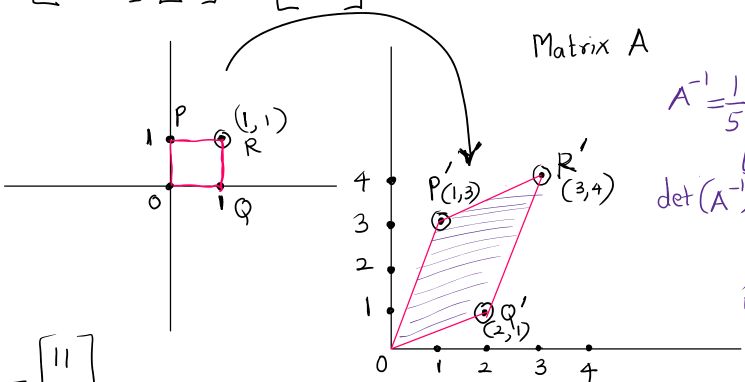
$$\begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$ $\det(A) = -5$

$(1,3)$ $(2,1)$

$(3,4)$

Matrix A



$A^{-1} = \frac{1}{5}\begin{bmatrix} -1 & 2 \\ 3 & -1 \end{bmatrix}$

$\det(A^{-1}) = -\frac{1}{5}$

BA

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

$\det(B) = -10$

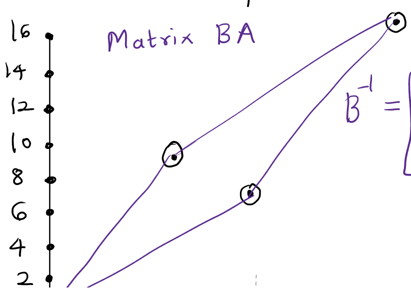
$$B\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 7 \end{bmatrix}$$

$$B\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$$

$$B\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 18 \\ 16 \end{bmatrix}$$

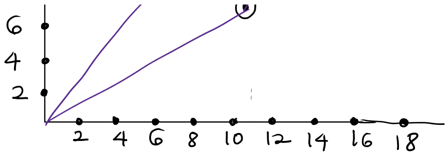
$$BA = \begin{bmatrix} 11 & 7 \\ 7 & 9 \end{bmatrix}$$

Matrix BA



$$B^{-1} = \left[\frac{1}{10}\begin{pmatrix} -1 & 3 \\ 4 & -2 \end{pmatrix}\right]$$

$$\begin{bmatrix} 11 & 7 \\ 7 & 9 \end{bmatrix} \qquad B\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 18 \\ 16 \end{bmatrix}$$



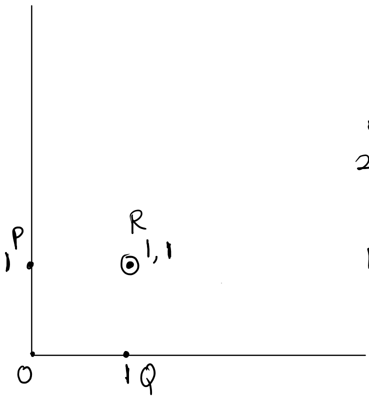
$$\det(A) = \frac{1}{\det(A^{-1})}$$

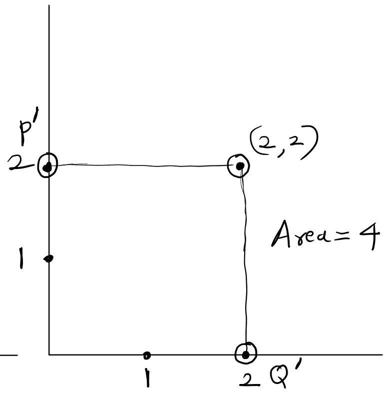
$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \det = 4$$

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

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

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





Hadamard Product

$$\underset{2\times3}{\begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & -1 \end{bmatrix}} \cdot \underset{2\times3}{\begin{bmatrix} 0 & 8 & 5 \\ 2 & -2 & 4 \end{bmatrix}} = \begin{bmatrix} 4(0) & 2(8) & 1(5) \\ 2(2) & 3(-2) & -1(4) \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 16 & 5 \\ 4 & -6 & -4 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 9 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 8 & 9 \\ 8 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 5 \\ 2 & 4 & -4 \end{bmatrix} \cdot \begin{bmatrix} 8 & 2 & 9 \\ 3 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 16 & 2 & 45 \\ 6 & -8 & -4 \end{bmatrix}$$

Reduced sum (M) = sum of all elements in M

Original Image

0	0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	0	0
0	0	0	0	0	0	1	0	0
0	0	0	0	0	1	0	0	0
0	0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0	0
0	0	1	1	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

Filter : 3 x 3

1	1	1
0	0	1
0	1	0

Convolution

Activation map

1	1	1	1	1	1	0
1	2	3	3	5	2	1
0	0	0	2	2	1	1
0	0	2	3	1	1	0
0	2	3	2	2	1	0
1	3	2	2	1	0	0
2	2	2	1	0	0	0

DEFINITION: The *rank* of a matrix A, written rank(A), is equal to the maximum number of linearly independent rows of A or, equivalently, the dimension of the row space of A.

no. of non-zero rows in a Echelon Form

DEFINITION: The *rank* of a matrix A, written rank(A), is equal to the maximum number of linearly independent rows of A or, equivalently, the dimension of the row space of A.

Rank = no. of non-zero rows in a Echelon Form

A	B	C
$4x + 2y = 15$ $2x + 3y = 18$	$2x + 2y = 13$ $4x + 4y = 26$	$2x + 3y = 13$ $2x + 3y = 10$
$\begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix}$	$\begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix}$
Rank = 2	Rank = 1	Rank = 1

System of Equations

Tuesday, September 12, 2023 5:53 PM

$$2x + y = 7 \qquad AX = B$$
$$3x - 5y = 4$$

$$A = \begin{bmatrix} 2 & 1 \\ 3 & -5 \end{bmatrix} \qquad B = \begin{bmatrix} 7 \\ 4 \end{bmatrix} \qquad X = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$AX = B$$
$$A^{-1}AX = A^{-1}B$$
$$I\,X = A^{-1}B$$
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}B = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Echelon Form

$$\begin{bmatrix} \bullet & \bullet \\ 0 & \bullet \end{bmatrix} \qquad \begin{bmatrix} \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet \\ 0 & 0 & \bullet \end{bmatrix}_{3\times 3}$$

$\bullet \Rightarrow$ Not all zeros

Reduced Echelon Form

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

$$4x + 2y = 15$$
$$2x + 3y = 18$$

$$\begin{array}{l} R_1 \rightarrow \\ R_2 \rightarrow \end{array}\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 15 \\ 18 \end{bmatrix}$$

$$R_2 \rightarrow R_2 - \frac{1}{2}R_1 \qquad \begin{bmatrix} 4 & 2 \\ 0 & 3-\frac{1}{2}(2) \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 15 \\ 18-\frac{1}{2}(15) \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 \end{bmatrix}\begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 15 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 15 \\ 10.5 \end{bmatrix}$$

$4x + 2y = 15$

$2y = 10.5 \Rightarrow y = 5.25$

$4x = 15 - 2(5.25) =$

2.16

(i) $2x + 3y = 1$
$5x + 7y = 3$

(ii) $2x + 4y = 10$
$3x + 6y = 15$

(iii) $4x - 2y = 5$
$-6x + 3y = 1$

Echelon

$$\begin{bmatrix} 4 & -2 \\ 0 & 0 \end{bmatrix}$$

(i) $\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$R_2 \rightarrow 2R_2 - 5R_1$ $\begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$2x + 3y = 1$

$-y = 1 \Rightarrow y = -1$

$2x = 1 - 3(-1)$

$2x = 4$

$x = 2$

$\det \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} = -1$

(ii) Infinite Sol^n.

$\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \end{bmatrix}$

$R_2 \rightarrow R_2 - \frac{3}{2} R_1$

$\begin{bmatrix} 2 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$

$2x + 4y = 10$

$\det \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix} = 0$

(iii)

$\det \begin{bmatrix} 4 & -2 \\ -6 & 3 \end{bmatrix} = 0$

Assignments: Consider the vectors {[3, 0, 4], [−1, 0, 7], [2, 9, 11]} .Check that the vectors are linearly independent or not ?

$$\det \begin{bmatrix} 3 & -1 & 2 \\ 0 & 0 & 9 \\ 4 & 7 & 11 \end{bmatrix} = 3 \begin{vmatrix} 0 & 9 \\ 7 & 11 \end{vmatrix} - (-1) \begin{vmatrix} 0 & 9 \\ 4 & 11 \end{vmatrix} + 2 \begin{vmatrix} 0 & 0 \\ 4 & 7 \end{vmatrix}$$

$= -225$

$\det \neq 0 \Rightarrow$ Linearly Independence

$\det = 0 \Rightarrow$ Dependence

$D = -2 \quad D_x = -4 \quad D_y = -2 \quad D_z = -1$

3.53. Solve

(a) $\begin{aligned} x + y + 2z &= 4 \\ 2x + 3y + 6z &= 10 \\ 3x + 6y + 10z &= 17 \end{aligned}$ (b) $\begin{aligned} x - 2y + 3z &= 2 \\ 2x - 3y + 8z &= 7 \\ 3x - 4y + 13z &= 8 \end{aligned}$ (c) $\begin{aligned} x + 2y + 3z &= 3 \\ 2x + 3y + 8z &= 4 \\ 5x + 8y + 19z &= 11 \end{aligned}$

3.53. (a) $(2, 1, \frac{1}{2})$, (b) no solution, (c) $u = (-7a - 1,\ 2a + 2,\ a)$.

Cramer's Rule :-

$$4x + 2y = 15$$
$$2x + 3y = 18$$

$$D = \begin{vmatrix} 4 & 2 \\ 2 & 3 \end{vmatrix} = 8$$

$$D_x = \begin{vmatrix} 15 & 2 \\ 18 & 3 \end{vmatrix} = 9$$

$$D_y = \begin{vmatrix} 4 & 15 \\ 2 & 18 \end{vmatrix} = 42$$

$$x = \frac{D_x}{D} = \frac{9}{8} \qquad y = \frac{D_y}{D} = \frac{42}{8}$$

3.57. Write v as a linear combination of u_1, u_2, u_3, where

(a) $v = (4, -9, 2)$, $u_1 = (1, 2, -1)$, $u_2 = (1, 4, 2)$, $u_3 = (1, -3, 2)$; $(2, -1, 3)$

(b) $v = (1, 3, 2)$, $u_1 = (1, 2, 1)$, $u_2 = (2, 6, 5)$, $u_3 = (1, 7, 8)$; $(4, -2, 1)$

(c) $v = (1, 4, 6)$, $u_1 = (1, 1, 2)$, $u_2 = (2, 3, 5)$, $u_3 = (3, 5, 8)$. $(\quad)$

Eigen values & vectors

Friday, September 15, 2023 3:42 PM

$A_{n\times n}$ exists a scaler λ & vector v

such that $Av = \lambda v$.

λ is called eigen value

v is called eigen vector

For $A_{n\times n}$, there can be n possible eigen values.

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} \qquad v, \lambda$$

s.t. $\underset{2\times 2}{A}\, v_{2\times 1} = \lambda_{1\times 1} v_{2\times 1}$

$$A\,v = \lambda\, I_{2\times 2}\, v_{2\times 1}$$

$$\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} v = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} v$$

$ax = 0$

$x \neq 0$

$a = 0$

$$\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} v - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} v = 0$$

$$\begin{bmatrix} 4-\lambda & 2 \\ 2 & 3-\lambda \end{bmatrix} v = 0$$

$$\begin{vmatrix} 4-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} = 0$$

$$(4-\lambda)(3-\lambda) - 4 = 0$$

$$\lambda^2 - 7\lambda + 12 - 4 = 0$$

$$\lambda^2 - 7\lambda + 8 = 0$$

$\lambda_1 = 5.5615528$

$\lambda_2 = 1.43844718$

$$\lambda^2 - 7\lambda + 8 = 0$$

$$ax^2 + bx + c = 0$$

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$\lambda_1 = 5.5615\ldots$

$\lambda_2 = 1.43844718$

$$\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1.438 \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\left.\begin{aligned} 4x + 2y &= 1.438x \\ 2x + 3y &= 1.438y \end{aligned}\right\} \quad \begin{aligned} 2.562x + 2y &= 0 \\ 2x + 1.562y &= 0 \end{aligned}$$

$$2x = -1.562y$$

$$x = -0.781y$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -0.781y \\ y \end{bmatrix}$$

Put $y = 1 \Rightarrow \begin{bmatrix} -0.781 \\ 1 \end{bmatrix}$

Put $y = 2 \Rightarrow \begin{bmatrix} -1.56 \\ 2 \end{bmatrix}$

9.9. Let $A = \begin{bmatrix} 3 & -4 \\ 2 & -6 \end{bmatrix}$.

(a) Find all eigenvalues and corresponding eigenvectors.

$$\begin{bmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{bmatrix}$$

first normalize $\vec{v_1} = [1, 0, 2, 1]$:

$$\vec{u_1} = [\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}].$$

Next, let

$$\vec{w_2} = \vec{v_2} - \vec{u_1} \cdot \vec{v_2} * \vec{u_1} = [2, 2, 3, 1] - [\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}] \cdot [2, 2, 3, 1] * [\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}]$$
$$= [2, 2, 3, 1] - (\frac{9}{\sqrt{6}}) * [\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}]$$
$$= [2, 2, 3, 1] - [\frac{3}{2}, 0, 3, \frac{3}{2}]$$
$$= [\frac{1}{2}, 2, 0, \frac{-1}{2}]$$

Normalize $\vec{w_2}$ to get

$$\vec{u_2} = [\frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}, 0, \frac{-\sqrt{2}}{6}]$$

Now compute $\vec{u_3}$ in terms of $\vec{u_1}$ and $\vec{u_2}$ as follows. Let

$$\vec{w_3} = \vec{v_3} - \vec{u_1} \cdot \vec{v_3} * \vec{u_1} - \vec{u_2} \cdot \vec{v_3} * \vec{u_2} = [\frac{4}{9}, \frac{-2}{9}, 0, \frac{-4}{9}]$$

and normalize $\vec{w_3}$ to get

$$\vec{u_3} = [\frac{2}{3}, \frac{-1}{3}, 0, \frac{-2}{3}]$$

More generally, if we have an orthonormal set of vectors $\vec{u_1}, .., \vec{u_{k-1}}$, then $\vec{w_k}$ is expressed as

$$\vec{w_k} = \vec{v_k} - \sum_{i=1}^{k-1} \vec{u_i} \cdot \vec{v_k} * \vec{u_i}$$

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

into orthonormal column vectors

$$A = \begin{bmatrix} \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{6} & \frac{2}{3} \\ 0 & \frac{2\sqrt{2}}{3} & \frac{-1}{3} \\ \frac{\sqrt{6}}{3} & 0 & 0 \\ \frac{\sqrt{6}}{6} & \frac{-\sqrt{2}}{6} & \frac{-2}{3} \end{bmatrix},$$

Singular Value Decomposition

SVD is based on a theorem from linear algebra which says that a rectangular matrix A can be broken down into the product of three matrices - an orthogonal matrix U, a diagonal matrix S, and the transpose of an orthogonal matrix V. The theorem is usually presented something like this:

$$A_{mn} = U_{mm} S_{mn} V_{nn}^T$$

where $U^T U = I, V^T V = I$; the columns of U are orthonormal eigenvectors of AA^T, the columns of V are orthonormal eigenvectors of $A^T A$, and S is a diagonal matrix containing the square roots of eigenvalues from U or V in descending order.

9.9. Let $A = \begin{bmatrix} 3 & -4 \\ 2 & -6 \end{bmatrix}$.

(a) Find all eigenvalues and corresponding eigenvectors.

$V_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ $V_2 = \begin{bmatrix} -4 \\ -6 \end{bmatrix}$

$$V_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \qquad V_2 = \begin{bmatrix} -4 \\ -6 \end{bmatrix}$$

$$\bar{u}_1 = \frac{1}{\sqrt{13}} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.832 \\ 0.554 \end{bmatrix}$$

$$w_2 = \bar{V}_2 - \bar{u}_1 \cdot \bar{V}_2 * \bar{u}_1$$

$$= \begin{bmatrix} -4 \\ -6 \end{bmatrix} - \frac{1}{\sqrt{13}}(-24)\frac{1}{\sqrt{13}} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 + \frac{24}{13}(3) \\ -6 + \frac{24}{13}(2) \end{bmatrix}$$

$$= \begin{bmatrix} 1.5384615385 \\ -2.3076923076 \end{bmatrix}$$

$$\bar{u}_2 = \begin{bmatrix} w_2 \end{bmatrix} \frac{1}{\sqrt{7.692307692}} = \begin{bmatrix} 0.5547 \\ -0.8323 \end{bmatrix}$$

$$\begin{array}{cc} u_1 & u_2 \\ \end{array}$$
$$\begin{bmatrix} 0.832 & 0.5547 \\ 0.554 & -0.8323 \end{bmatrix}$$

$-$ $*$

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	4	6	-	-	3
User 2	-	1	-	10	-
User 3	8	-	9	1	7
User 4	10	2	-	4	9
User 5	2	-	8	9	2

$m \times n$

R

r_{ij}

$$\underbrace{\begin{bmatrix} A \\ a_{ij} \end{bmatrix}}_{m \times k} \underbrace{\begin{bmatrix} B \\ b_{ij} \end{bmatrix}}_{k \times n}$$

$$r_{ij} = \sum_k a_{ik} b_{kj} \quad \text{ideally}$$

Error fn $\sum_{ij} \left(r_{ij} - \sum_k a_{ik} b_{kj} \right)^2$ is <u>minimized</u> to find

Error fn $\sum_{ij} \left(r_{ij} - \sum_{k} a_{ik} b_{kj} \right)$ is minimized

best (good) values a_{ij}, b_{ij} done by Gradient Descent method

Covariance-Variance Matrix

Saturday, September 23, 2023 11:30 AM

Deviation from mean

X_i	$X_i - 4.8$	$(X_i - 4.8)^2$
4	−0.8	0.64
9	4.2	17.64
2	−2.8	7.84
8	3.2	10.24
1	−3.8	14.44

Var = 10.16

Mean

$$\mu = \frac{\sum X_i}{N} = 4.8$$

Variance

$$\sigma^2 = \frac{\sum (X_i - \mu)^2}{N}$$

X_i
4
5
4
3

Var = 0.5

$$\text{Covariance}(X, Y) = \frac{\sum (X_i - \mu_X)(Y_i - \mu_Y)}{N}$$

X	y
2	9
5	7
7	5
8	2

μ_X : mean(X) μ_Y : mean(Y)

$$\underset{\text{mat}}{\text{Cov}}(X_1, X_2) = \Sigma = \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) \end{bmatrix}$$

X	y	X-mean	Y-mean	(X-mean)(Y-mean)
2	9	2-5.5=-3.5	9-5.75=3.25	-11.375
5	7	5.5.5=-0.5	7-5.75=1.25	-0.626
7	5	7-5.5=1.5	5-5.75=-0.75	-1.125
8	2	8-5.5=2.5	2-5.75=-3.75	-9.375

μ: 5.5 5.75 V_X V_Y Cov = −5.625

$$\Sigma = \begin{bmatrix} 5.25 , & -5.625 \\ -5.625 , & 6.6875 \end{bmatrix}$$

X	y	X-mean	Y-mean	(X-mean)(Y-mean)
2	3			
5	10			
7	18			
8	24			

Cov = 17.875

Var-Cov matrix gives us the relationships of all the variables

+ve cov ⇒ direct relation ↑ ↑

−ve cov ⇒ indirect relation ↑ ↓

Variables +ve cov $\Rightarrow$ direct relation $\uparrow$ $\uparrow$

-ve cov $\Rightarrow$ indirect relation $\uparrow$ $\downarrow$

& also the spread of data points

PCA

Saturday, September 23, 2023 11:30 AM

$$\text{Data} \Longrightarrow \underset{\text{Cov Matrix}}{\text{Variance}} \longrightarrow \text{Eigen values \& corresponding eigen vectors computed}$$

Eigen Val 1 val 2 val 3

Eigenvalues are descendingly arraged

Descending ordered eigenvalues

Val 2 > Val 3 > Val 1

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

Normalized Eigen vector matrix C

$\downarrow$

C * Data $\longrightarrow$ Principal Comp scores matrix

Variances of PCA Score Variables are nothing but Eigen values of Variance - Covariance matrix (descending order)

$X_1 \; X_2 \ldots X_m$

$$PC_1 = X_1 a_1 + X_2 a_2 \ldots + X_m a_m$$

Note: All the Eigenvectors (Components) are orthogonal to each other.

$$PC_2 = X_1 b_1 + X_2 b_2 + \ldots + X_m b_m$$

$\vdots$

$X_1 \; X_2 \ldots X_m$ $\quad [PC_1 \; PC_2 \ldots PC_m]$

$$PC_m = - \; - \; -$$

```
print(prcomp.explained_variance_)
```

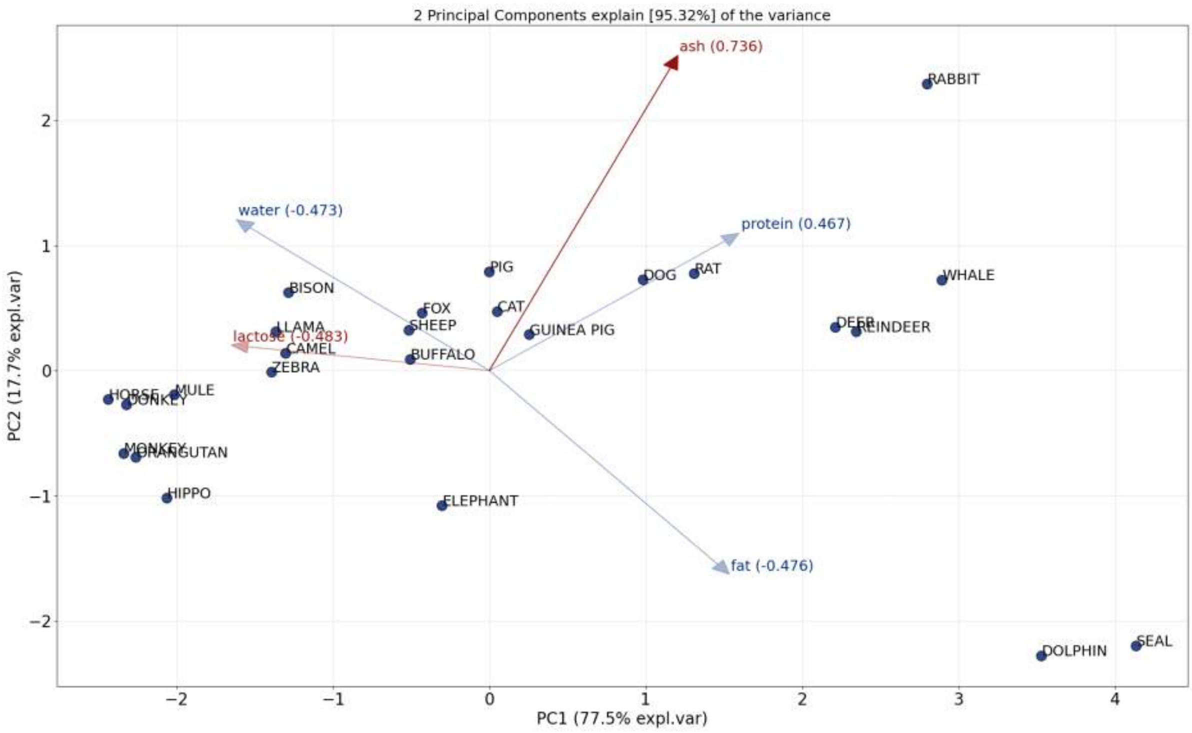
$$\underbrace{Var(PC_1) > Var(PC_2) > \cdots > Var(PC_m)}_{\text{Eigen values of Variance-Covariance matrix}}$$

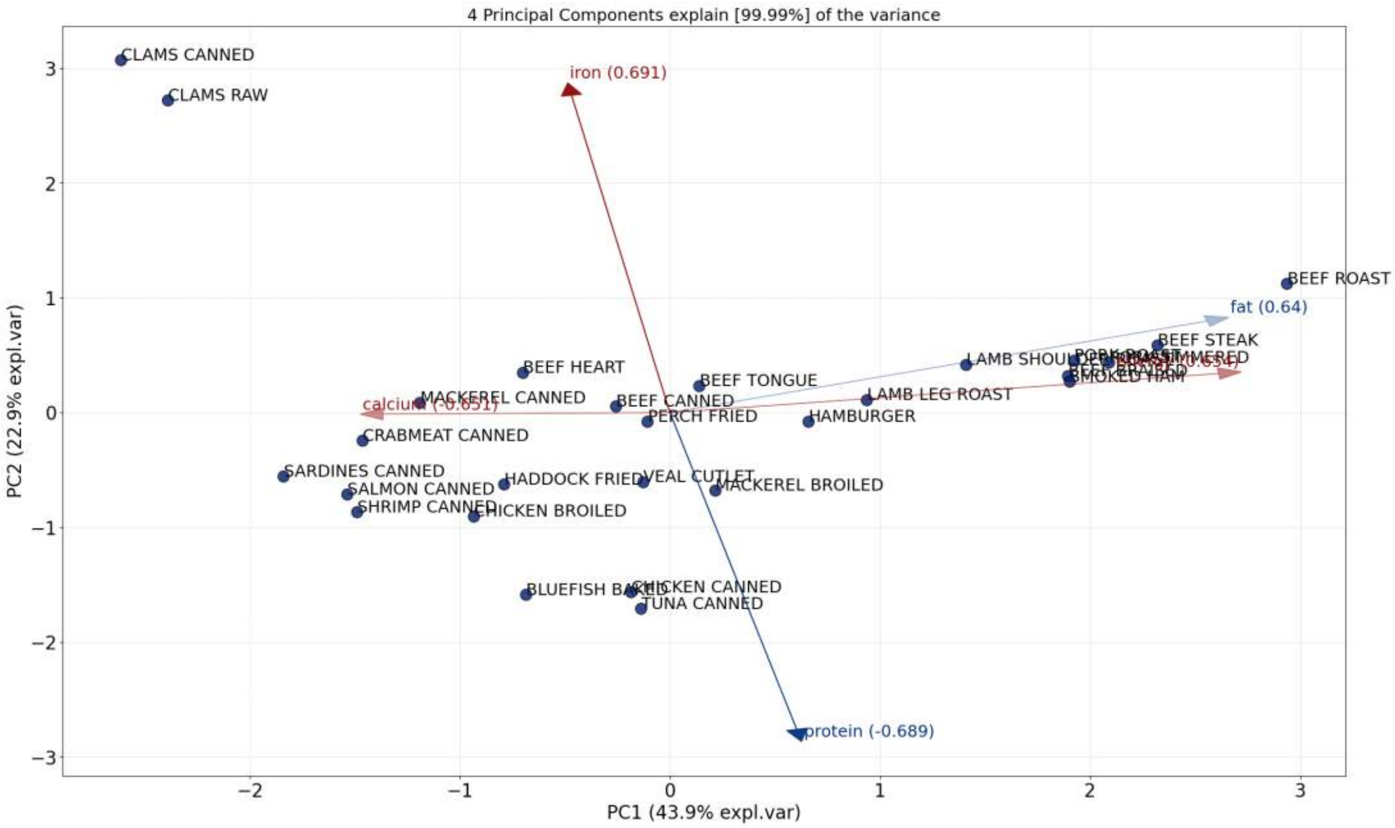
Good Practice: Scale the data, before you transform into PC scores.

$$\text{Total Variation} = Var(PC_1) + Var(PC_2) \ldots + Var(PC_m)$$

$$= \sum_{i=1}^{m} Var(PC_i)$$

$$\text{Percentage of Variation explained by } j^{th} \text{ PC} = \frac{Var(PC_j)}{\sum_{i=1}^{m} Var(PC_i)} \times 100$$


2 Principal Components explain [95.32%] of the variance
ash (0.736)
water (-0.473)
protein (0.467)
lactose (-0.483)
fat (-0.476)
RABBIT
WHALE
DEER
REINDEER
DOG
RAT
PIG
CAT
GUINEA PIG
BISON
FOX
SHEEP
LLAMA
CAMEL
BUFFALO
ZEBRA
HORSE
MULE
DONKEY
MONKEY
ORANGUTAN
HIPPO
ELEPHANT
DOLPHIN
SEAL
PC1 (77.5% expl.var)
PC2 (17.7% expl.var)


4 Principal Components explain [99.99%] of the variance
CLAMS CANNED
CLAMS RAW
iron (0.691)
BEEF ROAST
fat (0.64)
BEEF STEAK
LAMB SHOULDER
BEEF HEART
BEEF TONGUE
LAMB LEG ROAST
calcium (-0.651)
MACKEREL CANNED
BEEF CANNED
PERCH FRIED
HAMBURGER
CRABMEAT CANNED
SARDINES CANNED
SALMON CANNED
SHRIMP CANNED
HADDOCK FRIED
VEAL CUTLET
MACKEREL BROILED
CHICKEN BROILED
BLUEFISH BAKED
CHICKEN CANNED
TUNA CANNED
protein (-0.689)
PC1 (43.9% expl.var)
PC2 (22.9% expl.var)

Ex 2.3.5 $\lim\limits_{x\to 1} \dfrac{\sqrt{x+8}-3}{x-1}$ $= \dfrac{\sqrt{x+8}-3}{x-1} \times \dfrac{\sqrt{x+8}+3}{\sqrt{x+8}+3}$

$$= \frac{x+8-9}{(x-1)\left(\sqrt{x+8}+3\right)}$$

$$= \frac{\cancel{(x-1)}}{\cancel{(x-1)}\left(\sqrt{x+8}+3\right)}$$

$$= \frac{1}{6}$$

Ex 2.3.9 $\lim\limits_{x\to 0} \dfrac{4x-5x^2}{x-1}$

```
from sympy import limit, sqrt, pprint
from sympy.abc import x
```

```
   ...: expression = sqrt(x+1)
   ...: result = limit(expression, x, 0)
   ...: print("Limit of")
   ...: pprint(expression)
   ...: print("at x=0 is", result)
Limit of

  _______
╲╱ x + 1
at x=0 is 1
```

Ex 2.3.1 $\lim\limits_{x\to 3} \dfrac{x^2+x-12}{x-3}$

$f(x) = x^n \qquad \dfrac{df}{dx} = n\,x^{n-1}$

Ex 2.4.1 Find the derivative of $y = f(x) = \sqrt{169-x^2}$.

$$f(x) = \sqrt{169-x^2} = (169-x^2)^{1/2}$$

Put $u = 169 - x^2$; $\dfrac{du}{dx} = -2x$

$f(x) = u^{1/2}$; $\dfrac{df}{du} = \dfrac{1}{2}u^{-\frac{1}{2}} = \dfrac{1}{2\sqrt{u}}$

$$\frac{df}{dx} = \frac{df}{du}\,\frac{du}{dx} = \frac{-2x}{2\sqrt{169-x^2}} = \frac{-x}{\sqrt{169-x^2}}$$

$f(x) = e^x$ $\quad$ e : Euler's number $\quad$ $\frac{d e^x}{dx} = e^x$

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

n	1	2	3	4	5	100	100000
$\left(1+\frac{1}{n}\right)^n$	2	2.25	2.37037037	2.44140625	2.48832	2.704814	2.718268

$y = e^x \implies x = \log_e y$: Natural Log.

$y = \log x$; $\frac{dy}{dx} = \frac{1}{x}$

If for a point a $f'(a)=0$ then $f()$ is said to have extremum at point a

$f''(a) > 0 \Rightarrow f()$ has minimum at a
$f''(a) < 0 \Rightarrow f()$ has maximum at a

$$f_x(a,b)x + f_y(a,b)y - z = f_x(a,b)a + f_y(a,b)b - c$$
$$f_x(a,b)x + f_y(a,b)y - f_x(a,b)a - f_y(a,b)b + c = z$$
$$\star \quad f_x(a,b)(x-a) + f_y(a,b)(y-b) + c = z$$
$$f_x(a,b)(x-a) + f_y(a,b)(y-b) + f(a,b) = z$$

Ex 14.3.8 Find an equation for the plane tangent to $2x^2 + 3y^2 - z^2 = 4$ at $(1,1,-1)$. (answer)

$$\frac{d\sqrt{x}}{dx} = \frac{1}{2\sqrt{x}}$$

$$2x^2 + 3y^2 - z^2 = 4$$
$$z^2 = 2x^2 + 3y^2 - 4$$
$$f(x,y) = z = -\sqrt{2x^2+3y^2-4} \qquad f(1,1) = 1$$
$$f_x(x,y) = \frac{1}{\cancel{2}_1\sqrt{2x^2+3y^2-4}} \cancel{4}^2 x, \quad f_x(1,1) = -2$$
$$f_y(x,y) = \frac{1}{\cancel{2}_1\sqrt{2x^2+3y^2-4}} \cancel{6}^3 y, \quad f_y(1,1) = -3$$

$a=1, \ b=1, \ c=-1$

$$f_x(a,b)(x-a) + f_y(a,b)(y-b) + c = z$$
$$-2(x-1) + -3(y-1) - 1 = z$$

$(1,1,-1)$

$-2(1-1) + -3(1-1) - 1 = z \Rightarrow z = -1$

$$f_x(a,b)(x-a) + f_y(a,b)(y-b) + f(a,b) = z$$

Ex 14.3.10 Find an equation for the plane tangent to $f(x, y) = x^2 + y^3$ at $(3, 1, 10)$.

$f_x(3,1) = 6$, $f_y(3,1) = 3$ $c = 10$

$2x$ $3y^2$

$6(x-3) + 3(y-1) + 10 = z$

Directional Derivative
Tuesday, September 26, 2023 9:09 PM

Example 14.5.1 Find the slope of $z = x^2 + y^2$ at $(1,2)$ in the direction of the vector $\langle 3,4\rangle$.

We first compute the gradient at $(1,2)$: $\nabla f = \langle 2x, 2y\rangle$, which is $\langle 2,4\rangle$ at $(1,2)$. A unit vector in the desired direction is $\langle 3/5, 4/5\rangle$, and the desired slope is then
$\langle 2,4\rangle \cdot \langle 3/5, 4/5\rangle = 6/5 + 16/5 = 22/5$. □

Ex 14.5.1 Find $D_{\mathbf{u}} f$ for $f = x^2 + xy + y^2$ in the direction of $\mathbf{v} = \langle 2,1\rangle$ at the point $(1,1)$.

$f_x = 2x + y \qquad f_y = x + 2y$

$\nabla f_{(1,1)} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$; $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\quad \bar{n} = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$

l_2 norm $= \sqrt{5}$

$\nabla f_{(1,1)} \cdot \bar{n} = \dfrac{3 \times 2}{\sqrt{5}} + \dfrac{3}{\sqrt{5}} = \dfrac{9}{\sqrt{5}}$

Assignments: Let *f(x,y,z)=xyex2+z2−5f(x,y,z)=xyex2+z2−5*. Calculate the gradient of *ff* at the point (1,3,−2)(1,3,−2) and calculate the directional derivative *DufDuf* at the point (1,3,−2)(1,3,−2) in the direction of the vector v=(3,−1,4)v=(3,−1,4).

$f(x,y,z) = xye^{x^2} + z^2 - 5$, Calculate the gradient of f at point $(1,3,-2)$ & find directional derivative $D_u \cdot f$ / or $\nabla f \cdot u$ at point $(1,3,-2)$ in the direction of vector $v = \langle 3, -1, 4\rangle$

$f_x = y\left[x e^{x^2}(2x) + e^{x^2}(1)\right]$, $f_y = x e^{x^2}$, $f_z = 2z$

$\nabla f_{(1,3,-2)} = \begin{bmatrix} 9e \\ e \\ -4 \end{bmatrix}$ $\quad \bar{u} = \begin{bmatrix} 3/\sqrt{26} \\ -1/\sqrt{26} \\ 4/\sqrt{26} \end{bmatrix}$ $\quad \nabla f \cdot \bar{u} = \dfrac{26e - 16}{\sqrt{26}}$

https://mathinsight.org/directional_derivative_gradient_examples

Let $f(x,y,z) = xye^{x^2+z^2-5}$. Calculate the gradient of f at the point $(1,3,-2)$ and calculate the directional derivative $D_{\mathbf{u}} f$ at the point $(1,3,-2)$ in the direction of the vector $\mathbf{v} = (3,-1,4)$.

Solution: The gradient vector in three-dimensions is similar to the two-dimesional case. To calculate the gradient of f at the point $(1, 3, -2)$ we just need to calculate the three partial derivatives of f.

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = \left((y + 2x^2 y)e^{x^2+z^2-5}, xe^{x^2+z^2-5}, 2xyze^{x^2+z^2-5}\right)$$
$$\nabla f(1, 3, -2) = \left(3 + 2(1)^2 3e^0, 1e^0, 2(1)(3)(-2)e^0\right) = (9, 1, -12)$$

Just as for the above two-dimensional examples, the directional derivative is $D_{\mathbf{u}} f(x, y, z) = \nabla f(x, y, z) \cdot \mathbf{u}$ where $\mathbf{u}$ is a unit vector. To calculate $\mathbf{u}$ in the direction of $\mathbf{v}$, we just need to divide by its magnitude. Since $\|\mathbf{v}\| = \sqrt{3^2 + (-1)^2 + 4^2} = \sqrt{26}$,

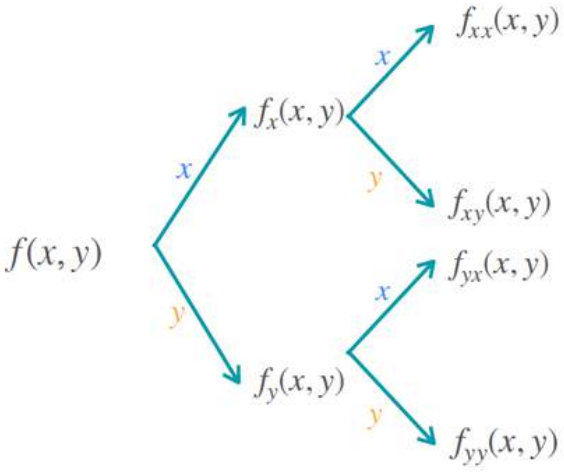
$$\mathbf{u} = \frac{\mathbf{v}}{\sqrt{26}} = \left(\frac{3}{\sqrt{26}}, \frac{-1}{\sqrt{26}}, \frac{4}{\sqrt{26}}\right)$$

and

$$\begin{aligned} D_{\mathbf{u}} f(1, 3, -2) &= \nabla f(1, 3, -2) \cdot \mathbf{u} \\ &= (9, 1, -12) \cdot \left(\frac{3}{\sqrt{26}}, \frac{-1}{\sqrt{26}}, \frac{4}{\sqrt{26}}\right) \\ &= \frac{9 \cdot 3 - 1 - 12 \cdot 4}{\sqrt{26}} = \frac{-22}{\sqrt{26}}. \end{aligned}$$

Hessian

Tuesday, September 26, 2023 9:12 PM



$$H = \begin{bmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{bmatrix}$$

Theorem 14.7.1 *Suppose that the second partial derivatives of $f(x, y)$ are continuous near (x_0, y_0), and $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$. We denote by D the* ***discriminant****:*

$$D(x_0, y_0) = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2.$$

If $D > 0$:

- if $f_{xx}(x_0, y_0) < 0$: there is a local maximum at (x_0, y_0);
- if $f_{xx}(x_0, y_0) > 0$: there is a local minimum at (x_0, y_0);

If $D < 0$: there is neither a maximum nor a minimum at (x_0, y_0);

If $D = 0$: the test fails.

Ex 14.7.1 Find all local maximum and minimum points of $f = x^2 + 4y^2 - 2x + 8y - 1$.

$f_x = 2x - 2$ $\qquad$ $f_y = 8y + 8$

$2x - 2 = 0 \Rightarrow x = 1$

$8y + 8 = 0 \Rightarrow y = -1$

$(x_0, y_0) = (1, -1)$

$D = (2)(8) - 0 = 16 > 0$

Has local min at $(1, -1)$

x_i	y_i
4	17
6	23
7	26
8	29
1	8
9	32

$$y = \overset{?}{b_0} + \overset{?}{b_1} x$$

$$\hat{y}_i = b_0 + b_1 x_i$$

$y_i - \hat{y}_i$: Residual error, we want as less as possible

$$Z = \sum_i (y_i - \hat{y}_i)^2 \text{ : error function}$$

To be minimized & b_0, b_1 to be found out

$$\boxed{\frac{d}{dx}(a - x^2) = -2x}$$

$$Z = \sum \left[y_i - (b_0 + b_1 x_i) \right]^2$$

$$\frac{\partial Z}{\partial b_0} = 0 \Rightarrow -2 \sum_{i=1}^{n} \left(y_i - (b_0 + b_1 x_i) \right) = 0$$

$$\therefore \sum_i y_i - \sum_i (b_0 + b_1 x_i) = 0$$

$$\bar{y} = \frac{\sum y_i}{n}$$

$$\therefore \sum y_i - n b_0 - b_1 \sum_i x_i = 0$$

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\therefore \frac{\sum y_i}{n} - b_0 - b_1 \frac{\sum x_i}{n} = 0$$

$$\therefore \bar{y} - b_0 - b_1 \bar{x} = 0$$

$$\therefore b_0 + b_1 \bar{x} = \bar{y} \quad \ldots\ldots (1)$$

$$\frac{\partial Z}{\partial b_1} = 0 \Rightarrow \frac{\partial}{\partial b_1} \sum_i \left[y_i - (b_0 + b_1 x_i) \right]^2 = 0$$

$$\boxed{\frac{d}{dx}(a - bx)^2 = -2b(a - bx)}$$

$$\therefore -2 \sum_i x_i (y_i - b_0 - b_1 x_i) = 0$$

$$\therefore \sum x_i y_i - b_0 \sum x_i - b_1 \sum x_i^2 = 0$$

$$b_0 \sum x_i + b_1 \sum x_i^2 = \sum x_i y_i \quad \ldots\ldots (2)$$

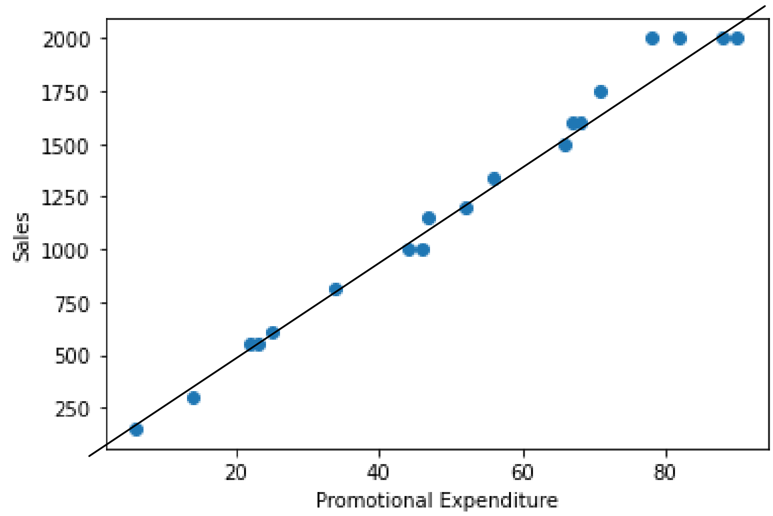
x_i	y_i	Xiyi	Xi^2
4	17	68	16
6	23	138	36
7	26	182	49
8	29	232	64
1	8	8	1
9	32	288	81
35	135	916	247

$$n b_0 + b_1 \sum x_i = \sum y_i \quad \ldots\ldots (1)$$

$$b_0 \sum x_i + b_1 \sum x_i^2 = \sum x_i y_i$$

$$\left. \begin{aligned} 6 b_0 + 35 b_1 &= 135 \\ 35 b_0 + 247 b_1 &= 916 \end{aligned} \right\} \Rightarrow \begin{aligned} b_0 &= 5 \\ b_1 &= 3 \end{aligned}$$

$y = b_0 + b_1 x$



X_1 X_2 y

Home	Automobile	Operating_Cost
400	1200	124000
350	360	71000
600	800	136000
800	1800	219000
900	1600	230000
200	1000	75000
120	900	56000
340	1100	110000
490	900	120000
700	800	144000

$$y = b_0 + b_1 x_1 + b_2 x_2$$

$$y^{(i)} = b_0 + b_1 x_1^{(i)} + b_2 x_2^{(i)}$$

$$Z = \sum_i \left(y^{(i)} - \hat{y}^{(i)}\right)^2$$

$$Z = \sum_i \left[y^{(i)} - \left(b_0 + b_1 x_1^{(i)} + b_2 x_2^{(i)}\right)\right]^2$$

$$\frac{\partial Z}{\partial b_0} = 0 \Rightarrow -2 \sum_i \left(y^{(i)} - b_0 - b_1 x_1^{(i)} - b_2 x_2^{(i)}\right) = 0$$

$$\sum y^{(i)} - n b_0 - b_1 \sum x_1^{(i)} - b_2 x_2^{(i)} = 0$$

$$n b_0 + \sum x_i^{(i)} b_1 + \sum x_2^{(i)} b_2 = \sum y^{(i)} \quad \text{---- } (1)$$

$$\frac{\partial z}{\partial b_1} = 0 \Rightarrow -2 \sum_i x_1^{(i)} \left[y^{(i)} - b_0 - b_1 x_1^{(i)} - b_2 x_2^{(i)}\right] = 0$$

$$\sum x_1^{(i)} y^{(i)} - b_0 \sum x_1^{(i)} - b_1 \sum x_1^{2(i)} - b_2 \sum x_1^{(i)} x_2^{(i)} = 0$$

$$\therefore \sum_i x_1^{(i)} b_0 + \sum_i x_1^{2(i)} b_1 + \sum_i x_1^{(i)} x_2^{(i)} b_2 = \sum x_1^{(i)} y^{(i)} \quad \text{---- } (2)$$

$$\frac{\partial z}{\partial b_2} = 0 \Rightarrow -2 \sum x_2^{(i)} \left[y^{(i)} - b_0 - b_1 x_1^{(i)} - b_2 x_2^{(i)}\right] = 0$$

$$\sum_i x_2^{(i)} y^{(i)} - b_0 \sum_i x_2^{(i)} - b_1 \sum_i x_1^{(i)} x_2^{(i)} - b_2 \sum_i x_2^{2(i)} = 0$$

$$\therefore \sum x_2^{(i)} b_0 + \sum_i x_1^{(i)} x_2^{(i)} b_1 + \sum_i x_2^{2(i)} b_2 = \sum_i x_2^{(i)} y^{(i)}$$

$$\therefore \quad \sum_i x_2^{(i)} b_0 + \sum_i x_1^{(i)} x_2^{(i)} b_1 + \sum_i x_2^{2^{(i)}} b_2 = \sum_i x_2^{(i)} y^{(i)} \quad \text{-----} (3)$$

$$n b_0 + \sum x_1^{(i)} b_1 + \sum x_2^{(i)} b_2 = \sum y^{(i)} \quad \text{----} (1)$$

$$\sum_i x_1^{(i)} b_0 + \sum_i x_1^{2^{(i)}} b_1 + \sum_i x_1^{(i)} x_2^{(i)} b_2 = \sum x_1^{(i)} y^{(i)} \quad \text{----} (2)$$

$$\sum_i x_2^{(i)} b_0 + \sum_i x_1^{(i)} x_2^{(i)} b_1 + \sum_i x_2^{2^{(i)}} b_2 = \sum_i x_2^{(i)} y^{(i)} \quad \text{-----} (3)$$

$\sum x_1^{(i)}$ | $\sum_i x_1^{2^{(i)}}$ | $\sum x_2^{(i)} y^{(i)}$

$\sum x_2^{(i)}$ | $\sum_i x_1^{(i)} x_2^{(i)}$

$\sum y^{(i)}$ | $\sum_i x_1^{(i)} y^{(i)}$ | $\sum_i x_2^{2^{(i)}}$

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$$

$$n b_0 + \sum x_1^{(i)} b_1 + \sum x_2^{(i)} b_2 + \sum x_3^{(i)} b_3 = \sum_i y^{(i)} \quad \text{-----} (1)$$

$$\sum_i x_1^{(i)} b_0 + \sum_i x_1^{2^{(i)}} b_1 + \sum_i x_1^{(i)} x_2^{(i)} b_2 + \sum x_1^{(i)} x_3^{(i)} b_3 = \sum x_1^{(i)} y^{(i)} \quad \text{----} (2)$$

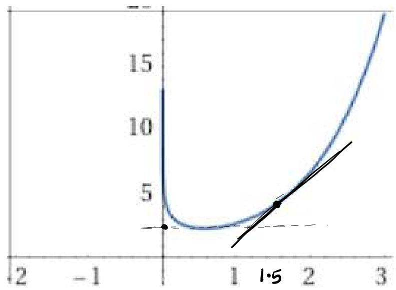
$$\sum_i x_2^{(i)} b_0 + \sum_i x_1^{(i)} x_2^{(i)} b_1 + \sum_i x_2^{2^{(i)}} b_2 + \sum x_2^{(i)} x_3^{(i)} b_3 = \sum_i x_2^{(i)} y^{(i)} \quad \text{----} (3)$$

$$\sum x_3^{(i)} b_0 + \sum_i x_1^{(i)} x_3^{(i)} b_1 + \sum x_2^{(i)} x_3^{(i)} b_2 + \sum x_3^{2^{(i)}} b_3 = \sum_i x_3^{(i)} y^{(i)} \quad \text{----} (4)$$

$$f(x) = e^{x} - \log x$$

$$f'(x) = e^{x} - \frac{1}{x}$$

$$f'(x) = 0 \Rightarrow e^{x} = \frac{1}{x}$$



$x_0 = 1.5$

$f'(1.5) = 3.815$, $f(1.5) = 4.076$

$x_1 = 1.5 - (0.01) f'(1.5) = 1.46$, $f(1.46) = 3.927$

$x_2 = 1.46 - 0.01 f'(1.46) = 1.423$

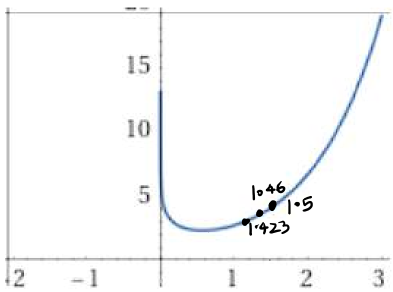
$$x_i = x_{i-1} - \eta f'(x_{i-1})$$

or

$$x_{i+1} = x_i - \eta f'(x_i)$$

$x_3 = 1.381$

$x_4 = 1.341$



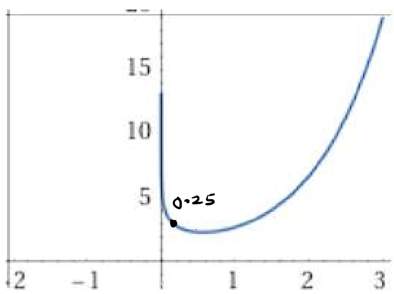
$\eta = 0.01$: learning rate

$x_0 = 0.25$

$$x_1 = x_0 - 0.01 f'(x_0)$$
$$= 0.277$$

$x_2 = 0.3$

$x_3 = 0.3198$



$$f(x) = x^3 - 6x^2 + 4x + 2$$

$y = b + w x$ $(-1, 1)$

? ?   say $b = 0.4$; $w = -0.6$

x	y
4	17
6	23
7	26
8	29
1	8
9	32

Steps:

1. Initialize values of weights (w) & bias (b) with some random values $(-1, 1)$
2. Calculate $\hat{y}$ as $\hat{y} = b + w x$ & also error as
$$L = \frac{1}{2n} \sum (y_i - \hat{y}_i)^2 = \frac{1}{2n} \sum \left[y_i - b - w x_i \right]^2$$
3. Update the values of b & w as :-
$$\text{new } b = \text{old } b - \eta \frac{\partial L}{\partial b}$$
$$\ldots = \text{old } w - \eta \frac{\partial L}{\ldots}$$

$$\text{new } b = \text{old } b - \eta \frac{\partial L}{\partial b}$$

$$\text{new } \omega = \text{old } \omega - \eta \frac{\partial L}{\partial \omega}$$

4. Calculate $\hat{y}$ & also L. Until $L < tol\ (0.0001)$, continue repeating steps from 2 to 4.

$$\frac{\partial L}{\partial b} = \frac{\partial}{\partial b} \frac{1}{2n} \sum_{i=1}^{n} (y_i - \hat{y})^2 \qquad \hat{y} = b + \omega x_i \qquad \frac{\partial \hat{y}}{\partial b} = 1$$

$$\frac{\partial \hat{y}}{\partial \omega} = x_i$$

$$L = \text{function}(\hat{y}) \quad ; \quad \hat{y} = \text{function}(b)$$

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} \quad ; \quad y = f(u), \quad u = g(x)$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial b} = \frac{1(-2)}{2n} \sum_{i=1}^{n} (y_i - \hat{y})(1) = -\frac{1}{n} \sum (y_i - \hat{y})$$

$$\frac{\partial L}{\partial \omega} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial \omega} = \frac{-2}{2n} \sum (y_i - \hat{y}) x_i = \frac{-1}{n} \sum (y_i - \hat{y}) x_i$$

$$\hat{y} = 0.4 + (-0.6) x_i$$

$$b_0 = 0.4 \quad , \quad \omega_0 = -0.6$$

$$b_1 = 0.4 - 0.01 \left[-\frac{1}{n} \sum (y_i - \hat{y}) \right] = 0.656$$

$$\omega_1 = -0.6 - 0.01 \left(-\frac{1}{n} \sum (y_i - \hat{y}) x_i \right) = 1.150$$

$$\hat{y} = 0.656 + 1.15 x$$

$$b_2 =$$

$$\omega_2 =$$

$$y = b + \omega_1 x_1 + \omega_2 x_2$$

$$y^{(i)} = b + \omega_1 x_1^{(i)} + \omega_2 x_2^{(i)}$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial b} = \frac{1(-2)}{2n} \sum_{i=1}^{n} (y^{(i)} - \hat{y})(1) = -\frac{1}{n} \sum (y^{(i)} - \hat{y})$$

$$\frac{\partial L}{\partial} = \frac{\partial L}{} \frac{\partial \hat{y}}{} = \frac{-2}{} \sum (y^{(i)} - \hat{y}) x = \frac{-1}{n} \sum (y^{(i)} - \hat{y}) x_1^{(i)}$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial w_1} = \frac{-2}{2n} \sum (y^{(i)} - \hat{y}) x = \frac{-1}{n} \sum (y^{(i)} - \hat{y}) x_1^{(i)}$$

$$\frac{\partial L}{\partial w_2} = \frac{-1}{n} \sum (y^{(i)} - \hat{y}) x_2^{(i)}$$

$$y = \frac{1}{1 + e^{-(b + wx)}} = \text{sigmoid}(x)$$

X	y
190	1
170	0
110	0
200	1
150	0
120	0
195	1

Steps:

1. Initialize values of w & b with some random values $(-1, 1)$
2. Calculate $\hat{y}$ as $\hat{y} = \text{sigmoid}(x)$ & also error as
   $L = \frac{-1}{n} \sum_i \left[y_i \ln \hat{y}_i + (1 - y_i) \ln (1 - \hat{y}_i) \right]$ (Log Loss f^n)
3. Update the values of b & w as:-
   new b = old b $- \eta \frac{\partial L}{\partial b}$
   new w = old w $- \eta \frac{\partial L}{\partial w}$
4. Calculate $\hat{y}$ & also L. Until $L <$ tol (0.0001), continue repeating steps from 2 to 4.

$$L = \text{function}(\hat{y}) \qquad \hat{y} = \text{sigmoid}(w, b) = \frac{1}{1 + e^{-(b + wx)}}$$

$$\frac{d}{dx} \ln x = \frac{1}{x}$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial b}$$

$$L = \frac{-1}{n} \sum_i \left[y_i \ln \hat{y} + (1 - y_i) \ln (1 - \hat{y}) \right]$$

$$\frac{\partial L}{\partial \hat{y}} = \frac{-1}{n} \sum \left[\left[\frac{\partial}{\partial \hat{y}} y_i \ln \hat{y} \right] + \left[\frac{\partial}{\partial \hat{y}} (1 - y_i) \ln (1 - \hat{y}) \right] \right]$$

$$= \frac{-1}{n} \sum \left[\frac{y_i}{\hat{y}} + \frac{(1 - y_i)}{(1 - \hat{y})} (-1) \right]$$

$$= \frac{-1}{n} \sum \left[\frac{y_i (1 - \hat{y}) - \hat{y} (1 - y_i)}{\hat{y} (1 - \hat{y})} \right]$$

$$= \frac{-1}{n} \sum \left(\frac{y_i - \hat{y}}{\hat{y} (1 - \hat{y})} \right) \quad \Big| \quad \sigma(z) = \frac{1}{1 + e^{-z}}$$

$$= -\frac{1}{n}\sum\left(\frac{y_i - \hat{y}}{\hat{y}(1-\hat{y})}\right)$$

$$\frac{\partial \hat{y}}{\partial b} = \frac{\partial}{\partial b}\,\sigma(wx_i + b)$$

$$= \sigma(wx_i+b)\left[1-\sigma(wx_i+b)\right](1)$$

$$\frac{\partial \hat{y}}{\partial b} = \hat{y}(1-\hat{y})$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial b}$$

$$= -\frac{1}{n}\sum\left(\frac{y_i-\hat{y}}{\cancel{\hat{y}(1-\hat{y})}}\right)\cancel{(\hat{y})(1-\hat{y})}$$

$$= -\frac{1}{n}\sum(y_i - \hat{y})$$

$$\frac{\partial L}{\partial w} = -\frac{1}{n}\sum\left[\frac{y_i-\hat{y}}{\hat{y}(1-\hat{y})}\right](\hat{y})(1-\hat{y})x_i$$

$$= -\frac{1}{n}\sum(y_i - \hat{y})x_i$$

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$$\frac{d}{dz}\sigma(z) = \frac{e^{-z}}{(1+e^{-z})^2}$$

$$= \frac{(1+e^{-z})-1}{(1+e^{-z})^2}$$

$$= \frac{1+e^{-z}}{(1+e^{-z})^2} - \frac{1}{(1+e^{-z})^2}$$

$$= \frac{1}{1+e^{-z}} - \frac{1}{(1+e^{-z})^2}$$

$$= \frac{1}{1+e^{-z}}\left[1 - \frac{1}{1+e^{-z}}\right]$$

$$= \sigma(z)\,(1-\sigma(z))$$

Riding mowers $\qquad x_1 = \text{income} \qquad x_2 = \text{Lot size}$

$$y = \frac{1}{1+e^{-(b_0 + w_1x_1 + w_2x_2)}}$$

$$\frac{\partial L}{\partial b} = -\frac{1}{n}\sum\left(y^{(i)} - \hat{y}^{(i)}\right) \qquad \frac{\partial L}{\partial w_1} = -\frac{1}{n}\sum\left(y^{(i)} - \hat{y}^{(i)}\right)x_1^{(i)}$$

$$\frac{\partial L}{\partial w_2} = -\frac{1}{n}\sum\left(y^{(i)} - \hat{y}^{(i)}\right)x_2^{(i)}$$

Stochastic & Mini-Batch

Wednesday, October 4, 2023 6:31 PM

Steps:

1. Initialize values of weights(w) & bias(b) with some random values $(-1, 1)$

2. Calculate $\hat{y}_i$ as $\hat{y}_i = b + wx$ & also error as
$L = \frac{1}{2n} (y_i - \hat{y}_i)^2 = \frac{1}{2n} \left[y_i - b - wx_i\right]^2$ for each obs

3. Update the values of b & w as:-

$$\left.\begin{aligned} \text{new } b &= \text{old } b - \eta \frac{\partial L}{\partial b} \\ \text{new } w &= \text{old } w - \eta \frac{\partial L}{\partial w} \end{aligned}\right\} \text{ for each obs}$$

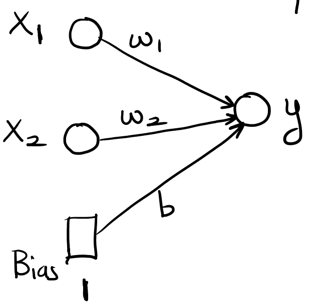
4. Calculate $\hat{y}$ & also L. Until $L < tol\ (0.0001)$, continue repeating steps from 2 to 4.

$y = b + w_1 x_1 + w_2 x_2$ y numerical

$$y = \frac{1}{1+e^{-(b+w_1x_1+w_2x_2)}}$$ categorical (0 or 1)

X_1 X_2 y

Perceptron



if y numerical,

$y = b + w_1x_1 + w_2x_2$

if y is binary,

$$y = \frac{1}{1+e^{-(b_0+w_1x_1+w_2x_2)}}$$

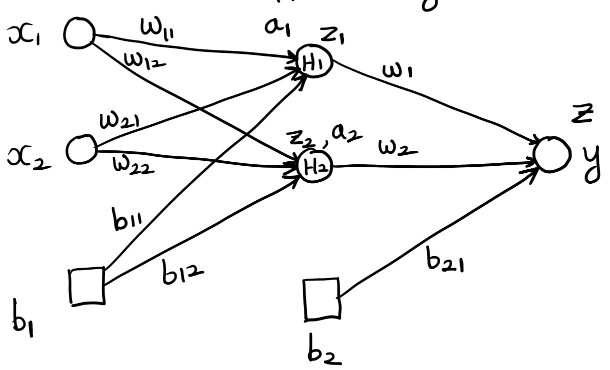
Activation Functions

y to be strictly non-negative, numerical

$y = \max(0, y)$

ReLU :- Rectified Linear Unit

Multi-Layer Perceptron



$z_1 = x_1 w_{11} + x_2 w_{21} + b_{11}$

$$a_1 = \frac{1}{1+e^{-z_1}} = \sigma(z_1)$$

$z_2 = x_1 w_{12} + x_2 w_{22} + b_{12}$

$a_2 = \sigma(z_2)$

$z = a_1 w_1 + a_2 w_2 + b_{21}$

$\hat{y} = \sigma(z)$

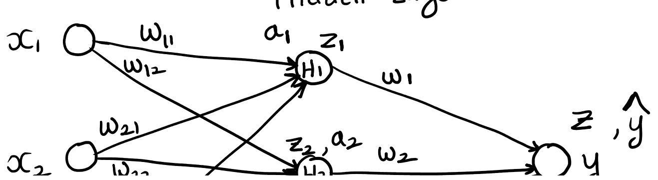
$L = \text{logloss}(y, \hat{y})$

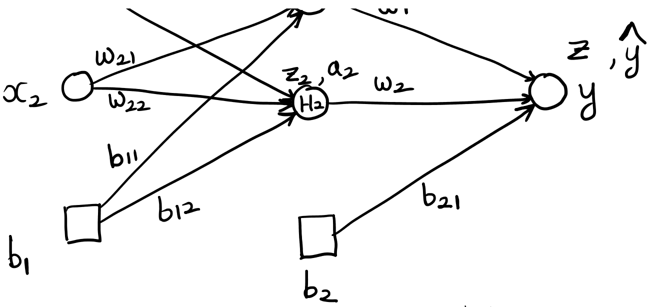
$$\frac{\partial L}{\partial w_{ij}} \qquad \frac{\partial L}{\partial w_i} \qquad \frac{\partial L}{\partial b_{ij}}$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial z} \frac{\partial z}{\partial w_1} = \frac{-(y_i - \hat{y}_i)}{\cancel{\hat{y}_i(1-\hat{y}_i)}} \cancel{\hat{y}_i(1-\hat{y}_i)}\, a_1 = -(y_i - \hat{y}_i) a_1$$

$$\frac{\partial L}{\partial w_2} = -(y_i - \hat{y}_i) a_2$$

$$\frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial z} \frac{\partial z}{\partial a_1} \frac{\partial a_1}{\partial z_1} \frac{\partial z_1}{\partial w_{11}} = -x_1 a_1 (1-a_1) w_1 (y_i - \hat{y}_i)$$





Forward Pass Calculations $\longrightarrow$

$a_1, a_2, \hat{y}$

Backpropagation (of errors) $\longleftarrow$

Updating w_{ij}, w_i, b_{ij}

- Find the closest point P (x, y, z) to the origin on the plane $2x + y - z = 5$

$$d(OP) = \sqrt{(x-0)^2 + (y-0)^2 + (z-0)^2}$$

$$f(x,y,z) = x^2 + y^2 + z^2 \quad \text{subject to} \quad 2x+y-z=5$$
$$\text{or} \quad 2x+y-z-5=0$$

$$L(x,y,z,\lambda) = x^2+y^2+z^2+\lambda(2x+y-z-5)$$

$$\frac{\partial L}{\partial x} = 2x + 2\lambda \qquad \frac{\partial L}{\partial y} = 2y + \lambda \qquad \frac{\partial L}{\partial z} = 2z - \lambda$$

$$x = -\lambda \qquad y = -\frac{\lambda}{2} \qquad z = \frac{\lambda}{2}$$

$$2x+y-z = 5 \Rightarrow 2(-\lambda) - \frac{\lambda}{2} - \frac{\lambda}{2} = 5 \Rightarrow \lambda = \frac{-5}{3}$$

$$x = \frac{5}{3} \qquad y = \frac{5}{6} \qquad z = \frac{-5}{6}$$

- Find the extreme values of the function $f(x, y) = 3x + 4y$ on the circle $x^2 + y^2 = 1$.

$$L(x,y,z,\lambda) = 3x+4y+\lambda(x^2+y^2-1)$$

$$\frac{\partial L}{\partial x} = 3 + \lambda(2x) \qquad \frac{\partial L}{\partial y} = 4 + \lambda(2y)$$

$$x = \frac{-3}{2\lambda} \qquad y = \frac{-4}{2\lambda}$$

$$\frac{9}{4\lambda^2} + \frac{16}{4\lambda^2} = 1 \Rightarrow 9 + 16 = 4\lambda^2 \Rightarrow \lambda^2 = \frac{25}{4} \Rightarrow \lambda = \pm\frac{5}{2}$$

1) $x = -\frac{3}{5}$, $y = -\frac{4}{5}$

Minima

2) $x = \frac{3}{5}$ $y = \frac{4}{5}$

Maxima

- Find the minimum value of $f(x,y) = x^2 + y^2$ subject to $x + 2y = 1$.

$$x = \frac{1}{5}, \quad y = \frac{2}{5}$$

- Find the point on the plane $x + 2y + 3z = 13$ closest to the point(1,1,1).

$$f(x,y,z) = (x-1)^2 + (y-1)^2 + (z-1)^2$$

$$L(x,y,z,\lambda) = (x-1)^2 + (y-1)^2 + (z-1)^2 + \lambda(x+2y+3z-13)$$

$$\frac{\partial L}{\partial x} = 2(x-1) + \lambda(1) \qquad \frac{\partial L}{\partial y} = 2(y-1) + \lambda(2) \qquad \frac{\partial L}{\partial z} = 2(z-1) + \lambda(3)$$

$$x = \frac{2-\lambda}{2} \;;\; y = 1-\lambda \;;\; z = \frac{2-3\lambda}{2} \Rightarrow \lambda = -1$$

$$x = \frac{3}{2}, \quad y = 2, \quad z = \frac{5}{2}$$

- Find the minimum value of $g(x,y) = x^2 + 4y^2$ subject to $x + y = 0$ and $x^2 + y^2 - 1 = 0$

$$L(x,y,\lambda_1,\lambda_2) = x^2 + 4y^2 + \lambda_1(x+y) + \lambda_2(x^2+y^2-1)$$

$$\frac{\partial L}{\partial x} = 0 \Rightarrow 2x + \lambda_1(1) + \lambda_2(2x) = 0 \qquad \frac{\partial L}{\partial y} = 8y + \lambda_1(1) + \lambda_2(2y) = 0$$

$$2x(1+\lambda_2) = -\lambda_1 \qquad y(8+2\lambda_2) = -\lambda_1$$

$$x = \frac{-\lambda_1}{2(1+\lambda_2)} \qquad y = \frac{-\lambda_1}{8+2\lambda_2}$$

1st constraint:- $x + y = 0$

$$\frac{-\lambda_1}{2(1+\lambda_2)} - \frac{\lambda_1}{8+2\lambda_2} = 0 \Rightarrow \lambda_1\left[\frac{1}{2(1+\lambda_2)} + \frac{1}{2(4+\lambda_2)}\right] = 0$$

As $\lambda_1 \neq 0$; $\frac{1}{1+\lambda_2} + \frac{1}{4+\lambda_2} = 0 \Rightarrow 4+\lambda_2 + 1+\lambda_2 = 0$

$$\therefore \lambda_2 = -\frac{5}{2}$$

2nd constraint:- $x^2 + y^2 = 1$

$$\frac{\lambda_1^2}{9} + \frac{\lambda_1^2}{9} = 1 \Rightarrow \lambda_1 = \pm\frac{3}{\sqrt{2}}$$

$$\left.\begin{aligned} &\lambda_1 = \frac{3}{\sqrt{2}}, \; \lambda_2 = -\frac{5}{2} \Rightarrow x = \frac{1}{\sqrt{2}}, \; y = \frac{-1}{\sqrt{2}} \\ &\lambda_1 = \frac{-3}{\sqrt{2}} ; \; \lambda_2 = -\frac{5}{2} \Rightarrow x = \frac{-1}{\sqrt{2}}, \; y = \frac{1}{\sqrt{2}} \end{aligned}\right\} \text{minimum.}$$

Taylor's Series

Friday, October 6, 2023 5:56 PM

https://math.libretexts.org/Bookshelves/Calculus/Supplemental_Modules_(Calculus)/Multivariable_Calculus/3%3A_Topics_in_Partial_Derivatives/Taylor__Polynomials_of_Functions_of_Two_Variables

For $f(x,y)$, whose first and second partial derivatives exist at point (a,b), the 2nd degree polynomial can be given as

$$f(x,y) \approx f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) + \frac{f_{xx}(a,b)}{2}(x-a)^2 + f_{xy}(a,b)(x-a)(y-b) + \frac{f_{yy}(a,b)}{2}(y-b)^2$$

$\frac{d}{dx}\sin x = \cos x$

$\frac{d}{dx}\cos x = -\sin x$

a. $f(x,y) = \sin 2x + \cos y$ for (x,y) near the point $(0,0)$

evaluate $x = 0.2$, $y = 0.1$

$f(0,0) = 1$, $f_x(x,y) = 2\cos 2x$ $\quad f_y(x,y) = -\sin y$

$f_{xx}(x,y) = -4\sin 2x$ $\quad f_{yy}(x,y) = -\cos y$

$f_{xy}(x,y) = 0$

$f_x(0,0) = 2$ $\quad f_y(0,0) = 0$, $f_x(0,0) = 0$, $f_{yy}(0,0) = -1$

Put $f(0.2, 0.1) = \sin(0.4) + \cos(0.1)$

$=$

$f(x,y) \approx 2x - \frac{1}{2}y^2 + 1$

b. $f(x,y) = xe^y + 1$ for (x,y) near the point $(1,0)$

$f_x(x,y) = e^y$ $\quad f_y(x,y) = xe^y$, $f_{xx}(x,y) = 0$, $f_{yy}(x,y) = xe^y$

$f_{xy}(x,y) = e^y$

$f(x,y) \approx 2 + (x-1) + y + \frac{y^2}{2} + xy - y = 1 + x + xy + \frac{y^2}{2}$

Test with $x = 0.92$, $y = 0.01$

$f(0.92, 0.01) = ?$ $\quad 1.929$

Approx $= 1 + 0.92 + 0.92(0.01) + \frac{0.01^2}{2} = 1.929$

Taylor series | Chapter 11, Essence of calculus

$$f(x,y) = \ln(x^2 + y^2 + 1), \quad P(0,0)$$

$f(x,y) \approx x^2 + y^2$

$0.1, \ 0.2$

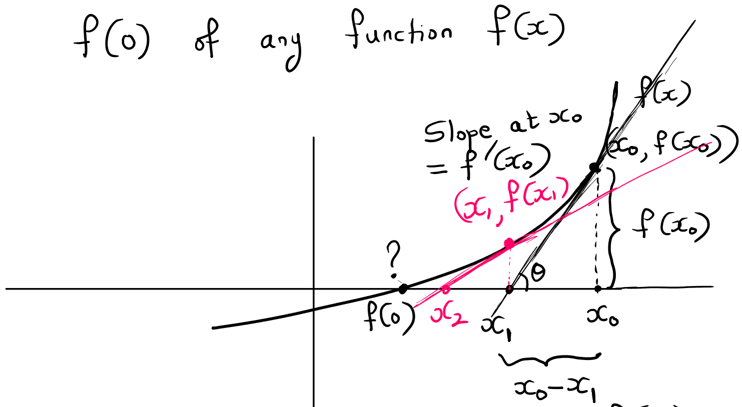
$\ln(0.1^2 + 0.2^2 + 1) = 0.048$

$0.1^2 + 0.2^2 = 0.05$

Newton's Method

Friday, October 6, 2023 6:37 PM

$f(0)$ of any function $f(x)$



$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

$$f'(x_0) = \frac{f(x_0)}{x_0 - x_1}$$

$$x_0 - x_1 = \frac{f(x_0)}{f'(x_0)}$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_{k+1})}$$

$$f(x) = x^3 - 6x^2 + 4x + 2$$

$$f'(x) = 3x^2 - 12x + 4$$

$$f''(x) = 6x - 12$$

$$x_0 = 0.2$$

$$x_1 = 0.2 - \frac{f'(x)}{f''(x)} = 0.359$$

$$x_2 = 0.366$$

$$f(x, y) = x^4 + 0.8y^4 + 4x^2 + 2y^2 - xy - 0.2x^2y$$

$$f_x = 4x^3 + 8x - y - 0.4xy \qquad f_y = 4(0.8)y^3 + 4y - x - 0.2x^2$$

$$f_{xx} = 12x^2 + 8 - 0.4y \qquad f_{yy} = 12(0.8)y^2 + 4$$

$$f_{xy} = -1 - 0.4x \qquad f_{yx} = -1 - 0.4x$$

$$x_0 = 1 \qquad y_0 = 1$$

$$H = \begin{bmatrix} 19.6 & -1.4 \end{bmatrix} \qquad \nabla f = \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 10.6 \\ 6 \end{bmatrix}$$

$$H = \begin{bmatrix} 19.6 & -1.4 \\ -1.4 & 13.6 \end{bmatrix} \qquad \nabla f = \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 10.6 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - H^{-1} \nabla f$$

$$f(x, y, z) = x^2 + y^2 + z^2 + 6xy + 2$$

Basics

Monday, October 30, 2023 2:10 PM

Classification Type	Possible Values
Binary	Y or N, 0 or 1 Either of them(only one)
Multi-Class	A or B or C or D, For k possible cases, A1,A2,.....Ak Either of them(only one)
Multi-Label	For k possible cases, A1,A2,.....Ak Some m (0<=m<=k) of them can occur

1. Human
2. Dog
3. Cat
4. Horse
5. Car
6. Scooter
7. Bike

Image	Hu	D	Cat	Ho	Car	S	B
Horse, Human	1	0	0	1	0	0	0
Car, Bike	0	0	0	0	1	0	1
D, Car	0	1	0	0	1	0	0

X, y

(1000, p) (1000, 1)

Dataset

1000 obs

$X_1 \; X_2 \ldots X_p \quad y \quad (1000, p+1)$

70% → Training 700 (700, p+1)

X_train (700, p) y_train (700, 1)

30% → Testing 300 (300, p+1)

X_test (300, p) y_test (300, 1) → y_test

model.fit(X_train, y_train) $\xrightarrow{\text{apply model}}$ X_test.predict( ) → Predictions (300, 1)

Compared

error

|(Existing − Predicted)|

$(\text{error})^2$

Existing Predicted

56 55.2

Existing	Predicted
56	55.2
70	78.9
43	45
58	50.2

mean = Mean Absolute error

mean = Mean Squared error

$$RMSE = \sqrt{MSE}$$

```
In [40]: train.index
Out[40]:
Index([174, 267, 351, 252, 473,  40,  18, 270, 304, 245,
       ...
       375, 114, 264, 128, 303, 134,  37, 249, 179, 453],
      dtype='int64', length=354)

In [41]: test.index
Out[41]:
Index([290, 222, 428, 189, 163, 482, 314, 208, 251, 471,
       ...
       479, 362,  87, 363, 393,  50, 169, 312, 456,  74],
      dtype='int64', length=152)
```

```
In [43]: train.index
Out[43]:
Index([486, 317, 434, 188,  70, 378,  94, 144, 396, 230,
       ...
       368,  71, 285, 313, 459,  92, 480, 403, 214, 318],
      dtype='int64', length=354)

In [44]: test.index
Out[44]:
Index([102, 322, 282, 342, 229, 191,  56, 258,  37, 320,
       ...
        81,  20, 494, 130, 295, 200, 166,  59, 197, 411],
      dtype='int64', length=152)
```

Parameter & Hyper-parameter
https://machinelearningmastery.com/difference-between-a-parameter-and-a-hyperparameter/

X, y

85 % Train

15 % Test

K-Fold CV

Train 1000 obs

A (200) B (200) C (200) D (200) E (200)

(800) Train .fit()	Validation .predict()	MSE
A B C D	E	m_1
B C D E	A	m_2
A C D E	B	m_3
A B D E	C	m_4
A B C E	D	m_5
		mean

Data 1000 obs

A C B D E

Train	Test	Error
A B C D	E	e_1
A B C E	D	e_2
A B D E	C	e_3
A C D E	B	e_4
B C D E	A	e_5

```
In [122]: print(y.value_counts(normalize=True)*100)
Class
Benign       65.522175
Malignant    34.477825
Name: proportion, dtype: float64
```

(458)
(241)

stratify = y

Train

y_train (490)

B 321

M 169

Test

y_test (209)

B 137

M 72

Actual \ Predicted	Benign	Malignant		Recall	Prec	F1
Benign	[[134	4]	138	$\frac{134}{138}$	$\frac{134}{140}$	
Malig	[6	66]]	72	$\frac{66}{72}$	$\frac{66}{70}$	
	140	70	210			

Regression case:

y_i y_test	y-pred $\hat{y}_i$
42	42.1
33	32.4
50	48.2

$$MSE = \frac{\sum_i (y_i - \hat{y}_i)^2}{n}$$

$$RMSE = \sqrt{MSE}$$

$$MAE = \frac{\sum_i |y_i - \hat{y}_i|}{n}$$

$$= \frac{|-0.1| + |0.6| + |1.8|}{3}$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\ldots}$$

$\bar{y}$: mean of y_test

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)}{\sum (y_i - \bar{y})^2}$$

R2 score

$\bar{y}$: mean of y_test

In case, .fit(X_train, y_train) ; y_pred = .predict(X_train)

$$0 \leq R^2(\text{y\_train}, \text{y\_pred}) \leq 1$$

In case, .fit(X_train, y_train) ; y_pred = .predict(X_test)

$$R^2 \leq 1$$

Scoring = in cross_val_score, GridSearchCV

1) For classification default is accuracy score
2) For Regression default is R^2 score

$$\frac{X - \min(X)}{\max(X) - \min(X)} \rightarrow (0, 1)$$

Min Max scaler

X_1 X_2

```
x_trn = np.array([[24000, 8],
                  [30000, 2],
                  [35000, 7],
                  [34500, 1]])
        min :      24000   1
        max :      35000   8
x_tst = np.array([[34000, 7],
                  [12000, 1]])
```

$$\frac{24000 - 24000}{35000 - 24000} = 0 \qquad \frac{8-1}{8-1} = 1$$

$$\frac{30000 - 24000}{35000 - 24000} = \frac{6}{11} \qquad \frac{2-1}{8-1} = \frac{1}{7}$$

$$\frac{35000 - 24000}{35000 - 24000} = 1 \qquad \frac{7-1}{8-1} = \frac{6}{7}$$

$$\frac{34500 - 24000}{35000 - 24000} = \frac{10.5}{11} \qquad \frac{1-1}{8-1} = 0$$

Polynomial

Monday, October 30, 2023 5:12 PM

X
promote

```
In [106]: poly = PolynomialFeatures(degree=2)

In [107]: X_poly = poly.fit_transform(X)

In [108]: X_poly
Out[108]:
array([[1.000e+00, 2.300e+01, 5.290e+02],
       [1.000e+00, 5.600e+01, 3.136e+03],
       [1.000e+00, 3.400e+01, 1.156e+03],
       [1.000e+00, 2.500e+01, 6.250e+02],
       [1.000e+00, 6.700e+01, 4.489e+03],
```

1 Promote Promote^2

b_0 b_1 b_2 b_3

```
-18.27491741665085 [ 0.00000000e+00  2.48006959e+01 -1.30947333e-02]
```

Inference with LR

Monday, October 30, 2023 6:33 PM

$H_0: \beta_0 = \beta_1 = 0$ (irrelevant)

H_1: Not all β_j zero (relevant)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  Sales   R-squared:                       0.989
Model:                            OLS   Adj. R-squared:                  0.988
Method:                 Least Squares   F-statistic:                     1503.
Date:                Mon, 30 Oct 2023   Prob (F-statistic):           4.97e-18
Time:                        18:32:20   Log-Likelihood:                -105.50
No. Observations:                  19   AIC:                             215.0
Df Residuals:                      17   BIC:                             216.9
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          5.4859     34.718      0.158      0.876     -67.763      78.734
Promote       23.5064      0.606     38.767      0.000      22.227      24.786
------------------------------------------------------------------------------
```

ANOVA for Regression $\rightarrow$ 4.97e-18 < 0.05

const: $b_0 = 5.4859$ — $H_0: \beta_0 = 0$, $H_1: \beta_0 \neq 0$

Promote: $b_1 = 23.5064$ — $H_0: \beta_1 = 0$, $H_1: \beta_1 \neq 0$

$$y = \beta_0 + \beta_1 x + \epsilon \leftarrow \text{random error}$$

$$y = b_0 + b_1 x$$

b_0 = estimate of β_0

b_1 = estimate of β_1

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   medv   R-squared:                       0.741
Model:                            OLS   Adj. R-squared:                  0.734
Method:                 Least Squares   F-statistic:                     108.1
Date:                Mon, 30 Oct 2023   Prob (F-statistic):          6.72e-135
Time:                        18:52:29   Log-Likelihood:                -1498.8
No. Observations:                 506   AIC:                             3026.
Df Residuals:                     492   BIC:                             3085.
Df Model:                          13
Covariance Type:            nonrobust
------------------------------------------------------------------------------
```

Prob (F-statistic): 6.72e-135 < 0.05 — Model is relevant

```
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         36.4595      5.103      7.144      0.000      26.432      46.487
crim          -0.1080      0.033     -3.287      0.001      -0.173      -0.043
zn             0.0464      0.014      3.382      0.001       0.019       0.073
indus          0.0206      0.061      0.334      0.738      -0.100       0.141
chas           2.6867      0.862      3.118      0.002       0.994       4.380
nox          -17.7666      3.820     -4.651      0.000     -25.272     -10.262
rm             3.8099      0.418      9.116      0.000       2.989       4.631
age            0.0007      0.013      0.052      0.958      -0.025       0.027
dis           -1.4756      0.199     -7.398      0.000      -1.867      -1.084
```

indus: 0.738 > 0.05 → Not Significant

age: 0.958 > 0.05 → Not Significant

nox	-17.7666	3.820	-4.651	0.000	-25.272	-10.262
rm	3.8099	0.418	9.116	0.000	2.989	4.631
age	0.0007	0.013	0.052	0.958 >0.05	-0.025	0.027
dis	-1.4756	0.199	-7.398	0.000	-1.867	-1.084
rad	0.3060	0.066	4.613	0.000	0.176	0.436
tax	-0.0123	0.004	-3.280	0.001	-0.020	-0.005
ptratio	-0.9527	0.131	-7.283	0.000	-1.210	-0.696
black	0.0093	0.003	3.467	0.001	0.004	0.015
lstat	-0.5248	0.051	-10.347	0.000	-0.624	-0.425

Significant

Step-wise Regression

$X_1 \; X_2 \ldots X_p \qquad y$

1) Forward Selection $(X_1), (X_1, X_2), (X_1, X_2, X_3) \ldots\ldots$

2) Backward Elimination $(X_1, X_2 \ldots X_p) \; (X_1 \, X_2 \ldots X_{p-1}) \ldots\ldots (X_1)$

3) Both (Step-wise selection)

Regularized (Shrinkage)

Tuesday, October 31, 2023 2:29 PM

$X_1 \quad X_2 \cdots\cdots X_K \qquad y$

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_K x_K$$

$X_1 \quad X_2 \cdots X_{15} \qquad y$

$\left.\begin{matrix} X_7 \\ X_9 \end{matrix}\right\}$ less signif.

have to reduce

$$y = b_0 + b_1 x_1 + b_2 x_2 \cdots + b_7 x_7 + \cdots + b_9 x_9 + b_{15} x_{15}$$

Ordinary Least Squares :-

$$Z = \sum_i \left[y^{(i)} - \left(b_0 + b_1 x_1^{(i)} + \cdots + b_K x_K^{(i)} \right) \right]^2$$

Ridge Regression :-

$$Z = \sum_i \left[y^{(i)} - \left(b_0 + b_1 x_1^{(i)} + \cdots + b_K x_K^{(i)} \right) \right]^2 + \alpha \left(b_1^2 + b_2^2 + \cdots + b_K^2 \right)$$

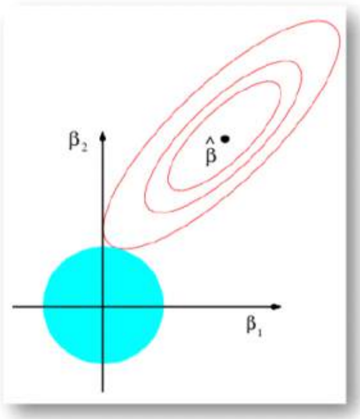
Regularization parameter ; $\alpha > 0$

Lasso Regression :-

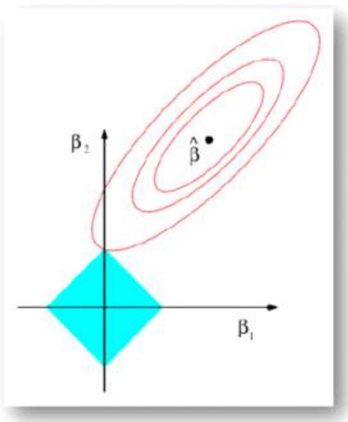
$$Z = \frac{1}{2n} \sum_i \left[y^{(i)} - \left(b_0 + b_1 x_1^{(i)} + \cdots + b_K x_K^{(i)} \right) \right]^2 + \alpha \left(|b_1| + |b_2| + \cdots + |b_K| \right)$$

Regularization parameter ; $\alpha > 0$

Ridge



Lasso



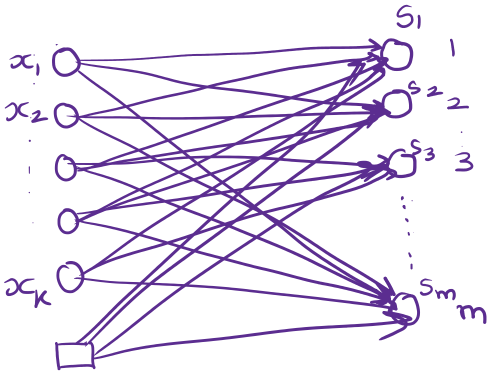
Logistic Regression

Thursday, November 2, 2023 8:15 AM

Type		Probability Function
Binary :	1 or 0	sigmoid $y = \dfrac{1}{1 + e^{-(b_0 + b_1 x_1 + \cdots + b_k x_k)}}$

Multiclass : 0 or 1 or 2 or ... m-1 (m categories)

(a) Multinomial : Softmax



Softmax Regression

$$P[y=0] = \frac{e^{s_1}}{e^{s_1} + e^{s_2} + \cdots + e^{s_m}}$$

$$P[y=1] = \frac{e^{s_2}}{e^{s_1} + e^{s_2} + \cdots + e^{s_m}} \cdots\cdots$$

$\vdots$

$$P[y=m-1] = \frac{e^{s_m}}{e^{s_1} + e^{s_2} + \cdots + e^{s_m}}$$

(b) OVR : One Vs Rest of All Categories

e.g. y = A or B or C

Train

$X_1 \cdots X_p$ y

A, A $\}$ 30

B, B $\}$ 40

C $\}$ 20

Trn A

$X_1 \cdots X_p$ y

A, A $\}$ 30

non-A $\}$ 60

Trn B

$X_1 \cdots X_p$ y

(40) B

non B (50)

Trn C

$X_1 \cdots X_p$ y

(20) C

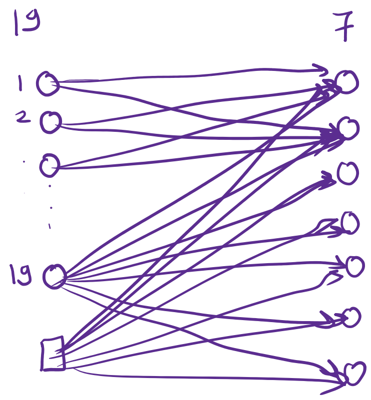
non - C (70)

Test

	$X_1\ X_2 \cdots X_p$	$P[y=A]$	$P[y=B]$	$P[y=C]$	y_pred
1.		0.8	0.6	0.4	A
		0.3	0.5	0.7	C

19   7   7x19 = 133

Img Seg.



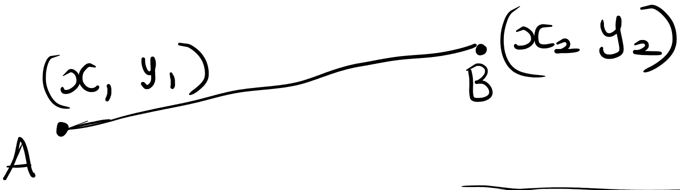
$7 \times 19 = 133$

$+ 7$

140

k-NN

Friday, November 3, 2023 11:50 AM



$$d(A,B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

$X_1 \quad X_2 \; \cdots \; X_p$

1) $(\quad\quad\quad\quad)$ •

2) $(\quad\quad\quad\quad)$ •

3) $(\quad\quad\quad\quad)$ •

Pipe

Friday, November 3, 2023 3:14 PM

X_train $\rightarrow$ scaling .fit_transform (Scaler) $\rightarrow$ Knn.fit() (Knn)

X_test $\rightarrow$ scaling .transform (scaler) $\rightarrow$ Knn.predict( ) (Knn)

Pipeline

Data $\rightarrow$ scaler $\rightarrow$ knn $\rightarrow$

trn

tst

model

prediction

X_train $\xrightarrow[\text{scaler .fit\_tran}]{}$ X_scl_trn $\rightarrow$ Knn

X_test $\xrightarrow[\text{scaler .transf}]{}$ X_scl_tst $\rightarrow$ Knn.predict ( )

Data 1000 obs

A C B D E

Train	Test	Error
A B C D	E	e_1
A B C E	D	e_2
A B D E	C	e_3
A C D E	B	e_4
B C D E	A	e_5

Scaler.fit_transf

Scaler

Scaler.fit_transf

↳ Knn.fit()

Scaler

.transf

↳

Knn.predict

params	plit0_test_scor	split1_test_score	plit2_test_scor	split3_test_score	split4_test_score	mean_test_score	std_test_score	ank_test_scor
{'KNN__n_neighbors': 1}	-4.29091	-4.29091	-5.14909	-7.72364	-1.75823	-4.64256	1.91397	10
{'KNN__n_neighbors': 2}	-4.34042	-3.53175	-4.38993	-4.37343	-0.963644	-3.51984	1.31858	9
{'KNN__n_neighbors': 3}	-1.8403	-3.54982	-3.57597	-2.76045	-0.999057	-2.54512	1.00055	8
{'KNN__n_neighbors': 4}	-1.86209	-3.57565	-3.566	-1.11383	-1.03254	-2.23002	1.13232	7
{'KNN__n_neighbors': 5}	-1.04585	-2.7987	-3.57773	-1.10108	-1.07235	-1.91914	1.06522	6
{'KNN__n_neighbors': 6}	-1.082	-2.01181	-2.78243	-1.12496	-1.08704	-1.61765	0.681656	5
{'KNN__n_neighbors': 7}	-1.12355	-1.21487	-2.79262	-1.15436	-1.09698	-1.47648	0.659242	4
{'KNN__n_neighbors': 8}	-1.13193	-1.25663	-2.81981	-0.365992	-1.13343	-1.34156	0.803921	3
{'KNN__n_neighbors': 9}	-1.1592	-0.456317	-2.01726	-0.335855	-1.15443	-1.02461	0.602938	2
{'KNN__n_neighbors': 10}	-0.34769	-0.481467	-2.03867	-0.348921	-1.18793	-0.880936	0.657505	1

Cases\Medical Cost Personal
Insurance.csv
Y = charges

Model	Best R2	Best Params
ElasticNet	0.7474	L1ratio=0.999, alpha=0.6324
knn	0.7929	n_neighbors=6, StandardScaler()

M Status

X_1

gender

X_2

X_3

Good 3

X_1

M

D

S

Sep

X_2

M

F

TG

Nominal Categorical

≈ Dummy / One Hot encoding

Good 3

Better 4

Best 5

Poor 2

V Poor 1

Ordinal Categorical

Vehicle Silhouette

y=class

Model	Best Params	Best Score
Logistic Regression	{'multi_class': 'multinomial', 'penalty': 'l2', 'solver': 'newton-cg'}	-0.4309818209374633
KNN	{'KNN__n_neighbors': 10, 'SCL': StandardScaler()}	-1.0711613659010997

Naïve Bayes

Saturday, November 4, 2023 8:44 AM

Talks for X_1 more than 100 min? (TT >= 100)	Gender X_2	Response
y	male	not bought
n	male	not bought
n	female	not bought
n	female	not bought
n	male	not bought
n	male	not bought
y	male	bought
y	female	bought
n	female	bought
y	female	bought

$$P(C_i|X_1,\ldots,X_p) = \frac{P(X_1,\ldots,X_p|C_i)P(C_i)}{P(X_1,\ldots,X_p|C_1)P(C_1)+\cdots+P(X_1,\ldots,X_p|C_m)P(C_m)}.$$

$C_1 = \text{Bought} \qquad C_2 = \text{not bought}$

$$P(\text{Response}=\text{bought} \,|\, [\underbrace{\text{Gdr}=\text{male}}_{X_2} \cap \underbrace{TT \geq 100 = y}_{X_1}])$$

$$= \frac{P(TT\geq 100 = y \cap \text{Gdr}=\text{male} \,|\, \text{Bought})\, P(\text{Bought})}{P(TT\geq 100 = y \cap \text{Gdr}=\text{male} \,|\, \text{Bought})\, P(\text{Bought}) + P(TT \geq 100 = y \cap \text{Gdr}=\text{male} \,|\, NB)\, P(NB)}$$

A & B indep.

$P(A \cap B) = P(A)P(B)$

Also,

$P(A \cap B | C) = P(A|C)P(B|C)$

Apriori Probabilities

$$= \frac{P(TT\geq 100 = y | B)\, P(\text{Gdr}=\text{male} | B)\, P(B)}{P(TT\geq 100 = y | B)\, P(\text{Gdr}=\text{male} | B)\, P(B) + P(TT\geq 100 = y | NB)\, P(\text{Gdr}=\text{male} | NB)\, P(NB)}$$

$$= \frac{(3/4)(1/4)(4/10)}{(3/4)(1/4)(4/10) + (1/6)(4/6)(6/10)} = 0.529$$

Posterior: getting calculated by Bayes' formula

Talks for more than 100 min? (TT >= 100)	Gender	Response
y	male	not bought
n	male	not bought
n	female	not bought
n	female	not bought
n	male	not bought
n	male	not bought
y	male	bought
y	female	bought
n	female	bought
y	female	bought

$$P(\text{Response}=\text{bought} \,|\, [\underbrace{\text{Gdr}=\text{female}}_{X_2} \cap \underbrace{TT \geq 100 = y}_{X_1}]) = 0.87$$

$P(\text{Gdr}=f | B) = 3/4 \qquad P(\text{Gdr}=f | NB) = 2/3$

Naive Bayes assumes independence betn the features

Gaussian NB assumes Normal Dist of every feature separately

Discriminant

Saturday, November 4, 2023 3:28 PM

$x_1 \quad x_2 \cdots \quad x_m$

Linear in X

$$\boxed{\delta_i(\bar{x}) = \underbrace{x^T}_{1\times m} \underbrace{\Sigma^{-1}}_{m\times m} \underbrace{\mu_i}_{m\times 1} - \frac{1}{2}\underbrace{\mu_i^T}_{1\times m} \underbrace{\Sigma^{-1}}_{m\times m} \underbrace{\mu_i}_{m\times 1} + \underbrace{\log(P(C_i))}_{1\times 1}}$$

Test $X_1 \ldots X_m$	$\delta_1(x)$	$\delta_2(x) \ldots$	$\delta_3(x)$	argmax (prediction)
1)	89	42	33	A
2)	109	202	32	B

max

SVM

Saturday, November 4, 2023 5:01 PM

X_1	X_2	y
		1
		0
		0
		0
		1
		0

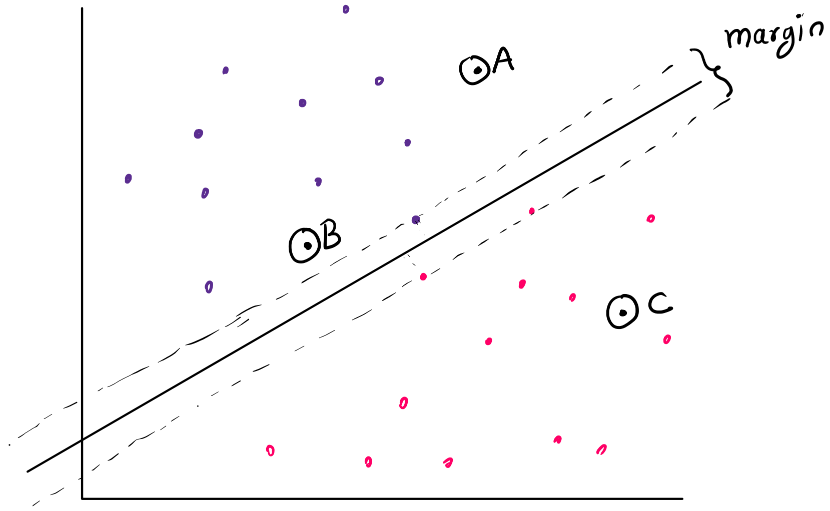
Test

A $\rightarrow$ 0

B $\rightarrow$ 0

C $\rightarrow$ 1

Maximum Margin Classifier



$C \propto \frac{1}{\text{margin}}$

Train

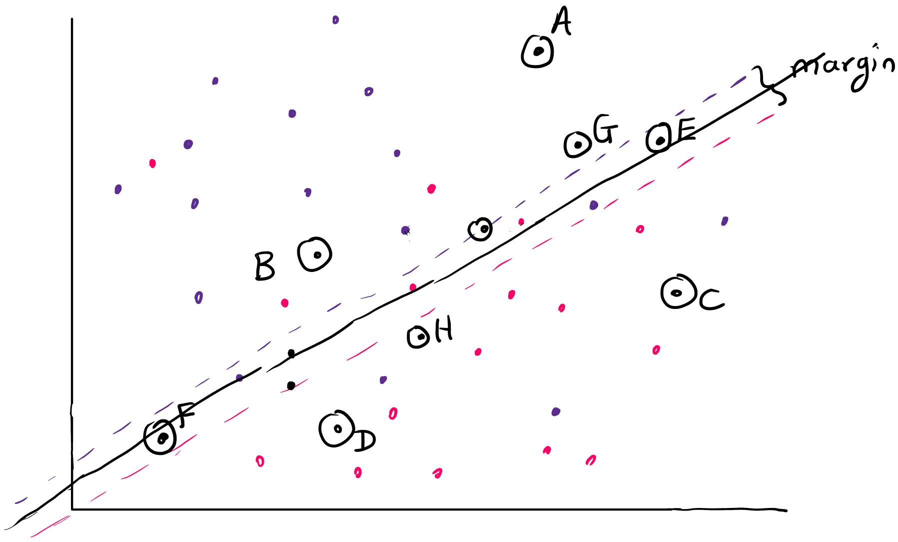
Purple 0

Pink 1

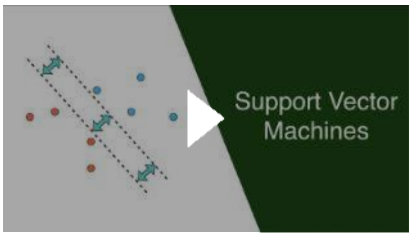
Test

F $\rightarrow$ 0

E $\rightarrow$ 1



[Support Vector Machines (SVMs): A friendly introduction]

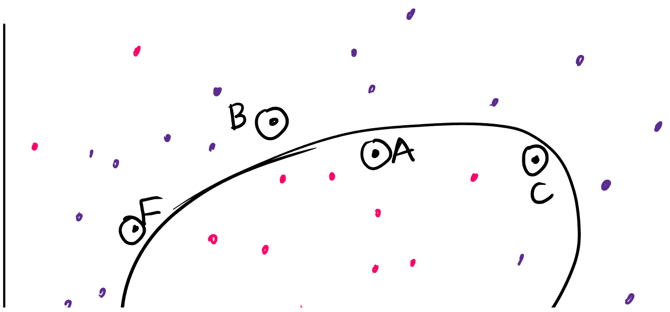


Radial (RBF)

$C \propto \frac{1}{\text{margin}}$

Test set

X_1	X	y
A		1
B		0

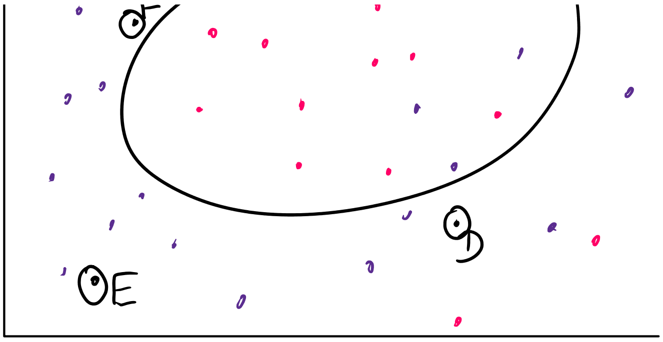


$\gamma \propto \frac{1}{\text{size}}$

(gamma)

$\frac{1}{\_\_} \approx \gamma$

B ⊙ 0
C ⊙ 0
D ⊙ 1
E ⊙ 0
F ⊙ 0



$$\frac{1}{2\sigma^2} \approx \gamma$$

Training

$X_1\ X_2 \dots X_p$	y
	$A \to 40$
	or
	$B \to 30$
	or
	$C \to 35$
	obs

OVO

A Vs B

$X_1 \dots X_p$	
(70)	$A \to 40$ $B \to 30$

SVM_{AB}

A Vs C

$X_1 \dots X_p$	
(75)	$A \to 40$ $B \to 35$

SVM_{AC}

B Vs C

$X_1 \dots X_p$	
(65)	$B \to 30$ $C \to 35$

SVM_{BC}

.predict

Test

$X_1 \dots X_p$	$\hat{y}_{AB}$	$\hat{y}_{AC}$	$\hat{y}_{BC}$	$\hat{y}_{final}$
	A	C	B	A
	B	A	B	B
	A	A	C	A

O.V R

Training

$X_1\ X_2 \dots X_p$	y
	$A \to 40$
	or
	$B \to 30$
	or
	$C \to 35$
	obs

A

$X_1 \dots X_p$	
	$A \to 40$ non-A $\to 65$

SVM_A

B

$X_1 \dots X_p$	
	B $\to 30$ non B $\to 75$

SVM_B

C

$X_1 \dots X_p$	
	C $\to 35$ non-C $\to 70$

SVM_C

.predict_proba

Test set

$X_1\ X_2 \dots X_p$	$P(y=A)$	$P(y=B)$	$P(y=C)$	$\hat{y}_{final}$
	0.87	0.7	0.3	A
	0.5	0.67	0.43	B

Satellite Imaging
Y = classes

Kernel	Params	Log Loss
Std scaler + Linear	{'SVM__decision_function_shape': 'ovo', 'SVM__C': 0.3544646464646464}	-0.33124457029711046
MM scaler + Linear	{'SVM__decision_function_shape': 'ovo', 'SVM__C': 4.8485151515151514}	-0.3311809989279939
Std scaler + RBF	{'SVM__gamma': 0.20297979797979795, 'SVM__decision_function_shape': 'ovr', 'SVM__C': 2.4247575757575754}	-0.2243476683647972
MM scaler + RBF	{'SVM__gamma': 4.545545454545454, 'SVM__decision_function_shape': 'ovr', 'SVM__C': 2.778222222222222}	-0.22215292831841907

Trees

Tuesday, November 7, 2023 2:13 PM

$X_1 \; X_2 \; y$

Test

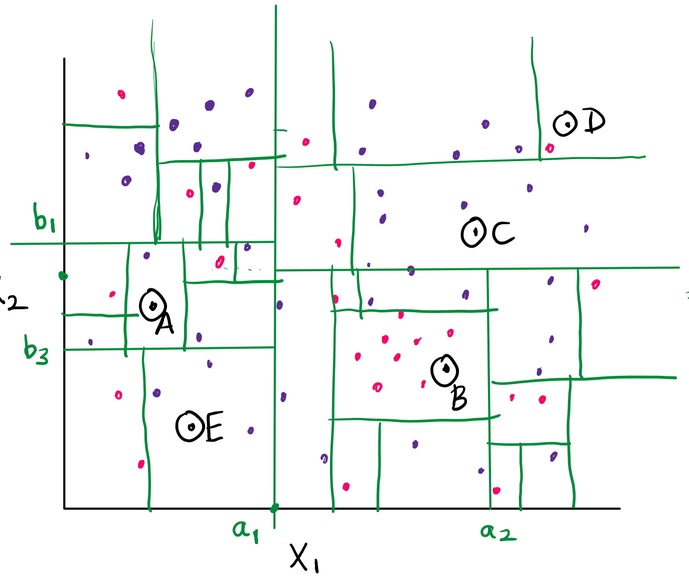
⊙A

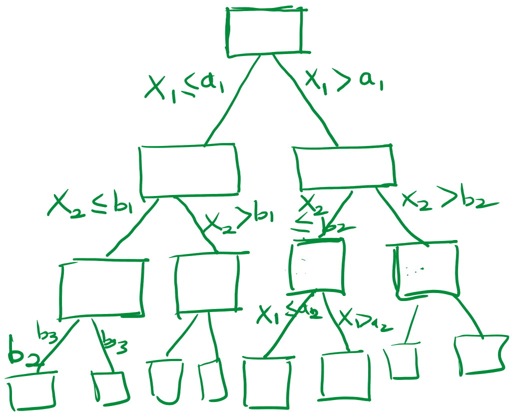
⊙B

⊙C

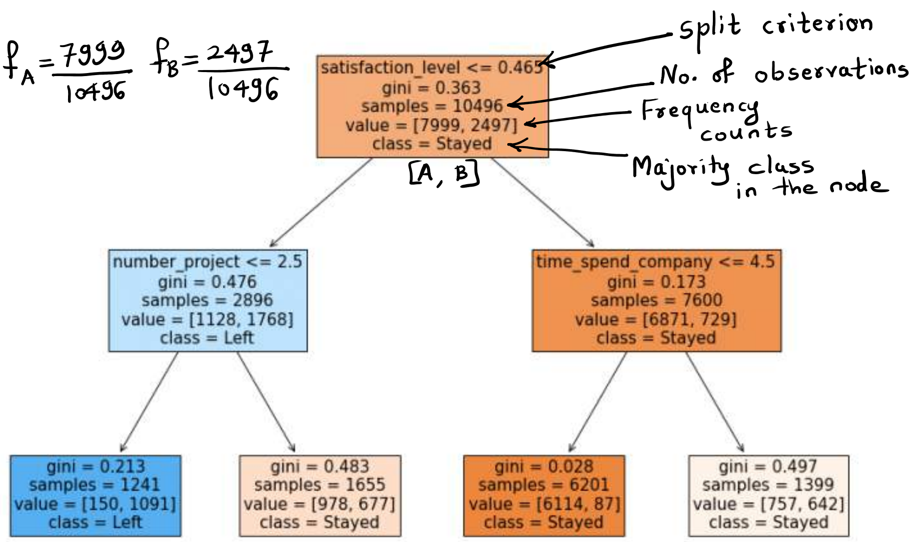
⊙D

⊙E





$f_A = \frac{7999}{10496}$ $\quad f_B = \frac{2497}{10496}$



A : 70%
B : 30%

Gini's Impurity Index

$f_A = \frac{70}{100} = 0.7$

$f_B = 0.3$

$$\begin{aligned}
\text{Gini} &= \sum f_i (1 - f_i) \\
&= f_A(1-f_A) + f_B(1-f_B) \\
&= 0.7(0.3) + 0.3(0.7) \\
&= 0.42
\end{aligned}$$

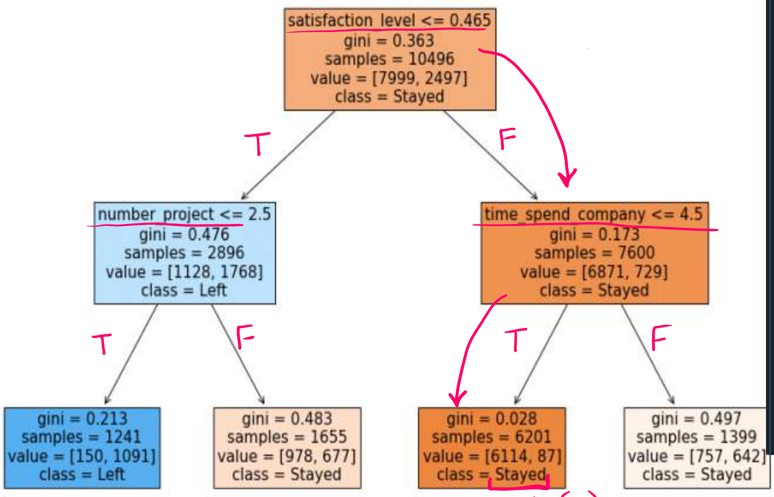
A : 90%
B : 10%

$f_A = 0.9$

$f_B = 0.1$

Gini = 0.18

$\left.\begin{array}{l} A: 100\% \\ B: 0\% \end{array}\right\}$ Gini = 0

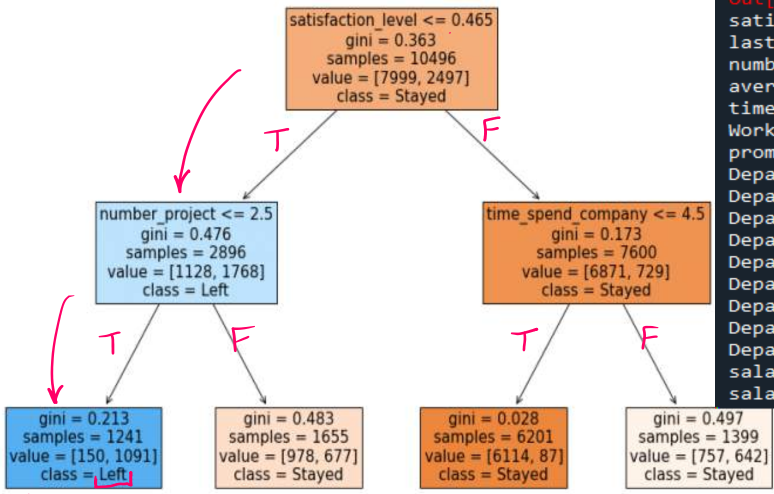



```
In [31]: X_test.loc[3049]
Out[31]:
satisfaction_level          0.54
last_evaluation             0.72
number_project                 3
average_montly_hours         172
time_spend_company             2
Work_accident                  0
promotion_last_5years          0
Department_RandD            True
Department_accounting      False
Department_hr              False
Department_management      False
Department_marketing       False
Department_product_mng     False
Department_sales           False
Department_support         False
Department_technical       False
salary_low                  True
salary_medium              False
```

left = ?

Predict_proba( )

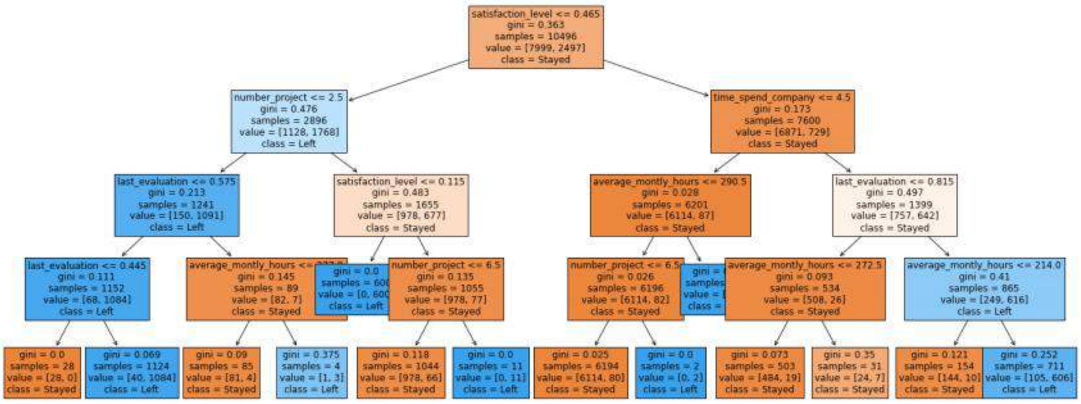
$P[0] = \frac{6114}{6201}$ $P[1] = \frac{87}{6201}$




```
In [33]: X_test.loc[256]
Out[33]:
satisfaction_level          0.46
last_evaluation             0.46
number_project                 2
average_montly_hours         154
time_spend_company             3
Work_accident                  0
promotion_last_5years          0
Department_RandD           False
Department_accounting       True
Department_hr              False
Department_management      False
Department_marketing       False
Department_product_mng     False
Department_sales           False
Department_support         False
Department_technical       False
salary_low                 False
salary_medium               True
```

$P[0] = \frac{150}{1241}$

$P[1] = \frac{1091}{1241}$



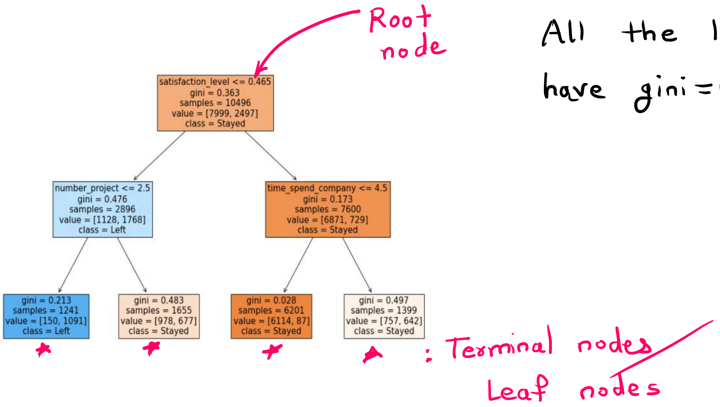
Fully Grown Tree :-

All the leaf nodes will

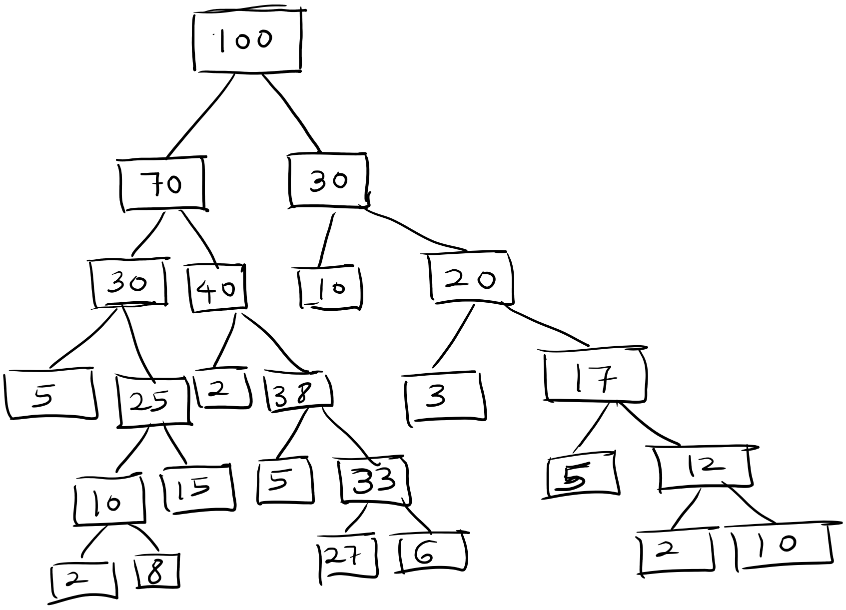
Root node

Fully Grown Tree :-

All the leaf nodes will have gini = 0



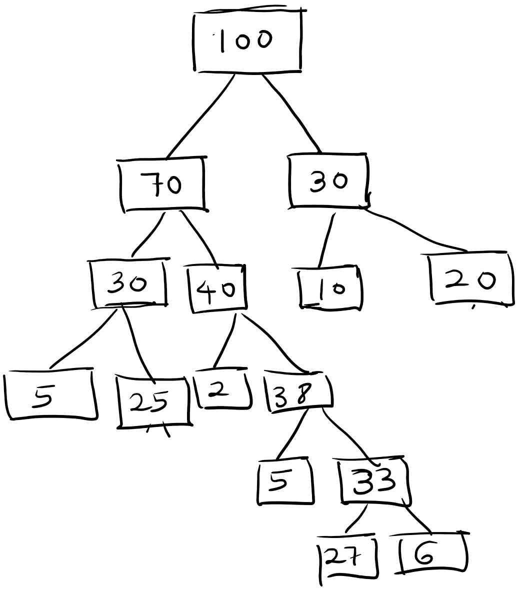
Tree Model



min_samples_split : *int or float, default=~~2~~ ~~15~~ ~~25~~ 30*

The minimum number of samples required to split an internal node

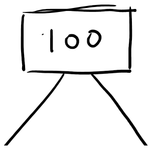
samples: obs/rows

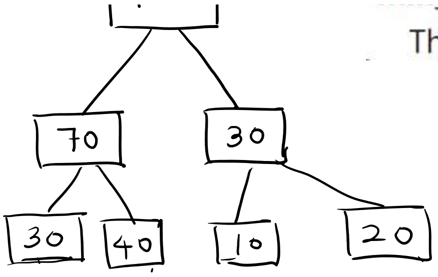


min_samples_leaf : *int or float, default=~~1~~ ~~5~~ 10*

The minimum number of samples required to be at a leaf node.

samples: rows/obs

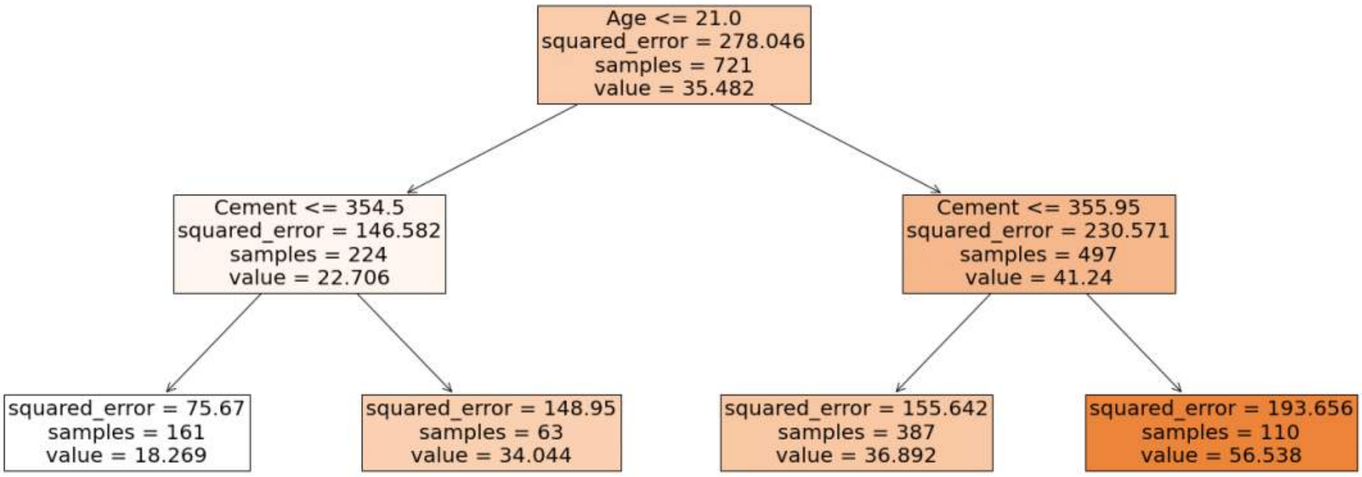




The minimum number of samples required to be at a leaf node.

rows/obs

Tree Regressor



Bagging & RF

Wednesday, November 8, 2023 6:28 PM

Simple Random Sampling
1) W/o Replacement
2) With Replacement → Bootstrap Sample

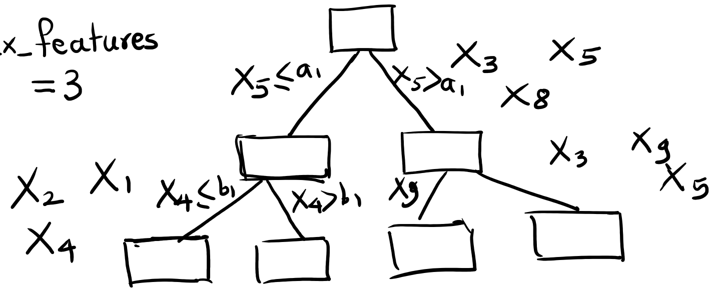
Data
- Train
  - Bootstrap Samples
- Test

$BS_1 \quad BS_2 \ldots \quad BS_{15}$

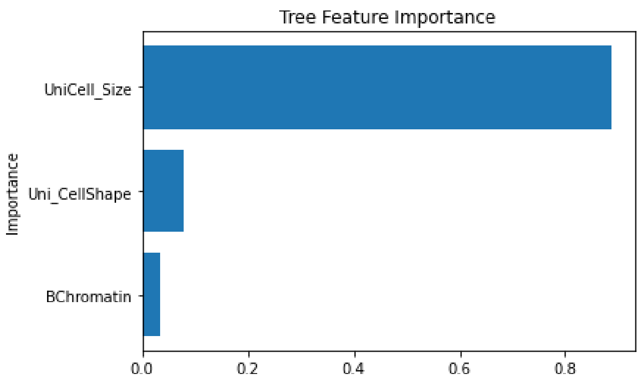
RF : estimator = decision tree

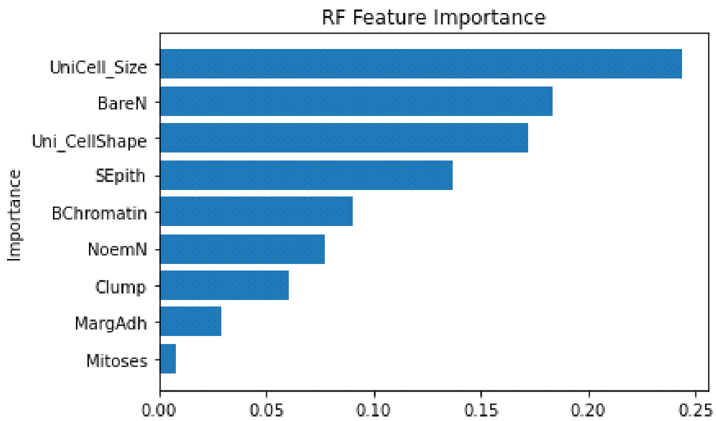
$X_1 \; X_2 \ldots X_{10}$: Bootstrap Sample

max_features = 3



For every split:-
1) At random, max_features no. of variables are chosen
2) The best split is chosen only among the chosen variables






```
In [17]: print(gcv_tree.best_score_)
-0.2456110095055712
```

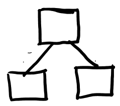
```
In [19]: print(gcv_rf.best_score_)
-0.1008663596688781
```

Boosting

Thursday, November 9, 2023 3:14 PM

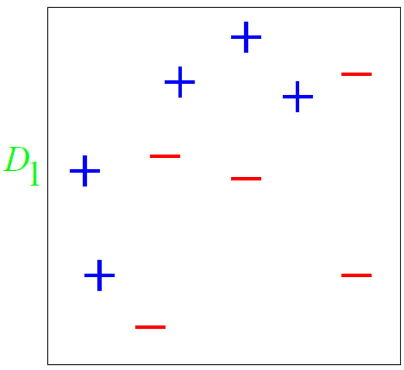
Training

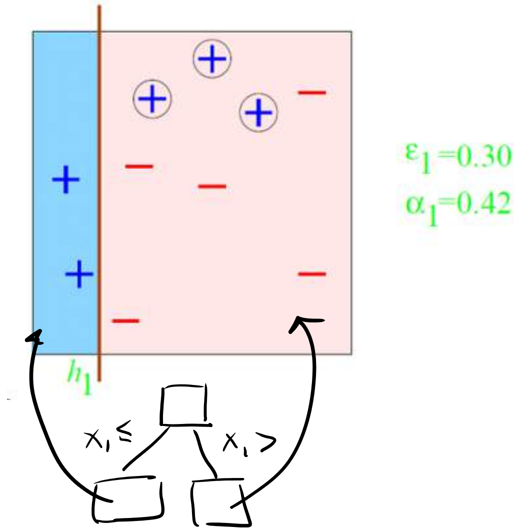
X_1	X_2	y	
		1	+
		0	-
		1	+
		0	-



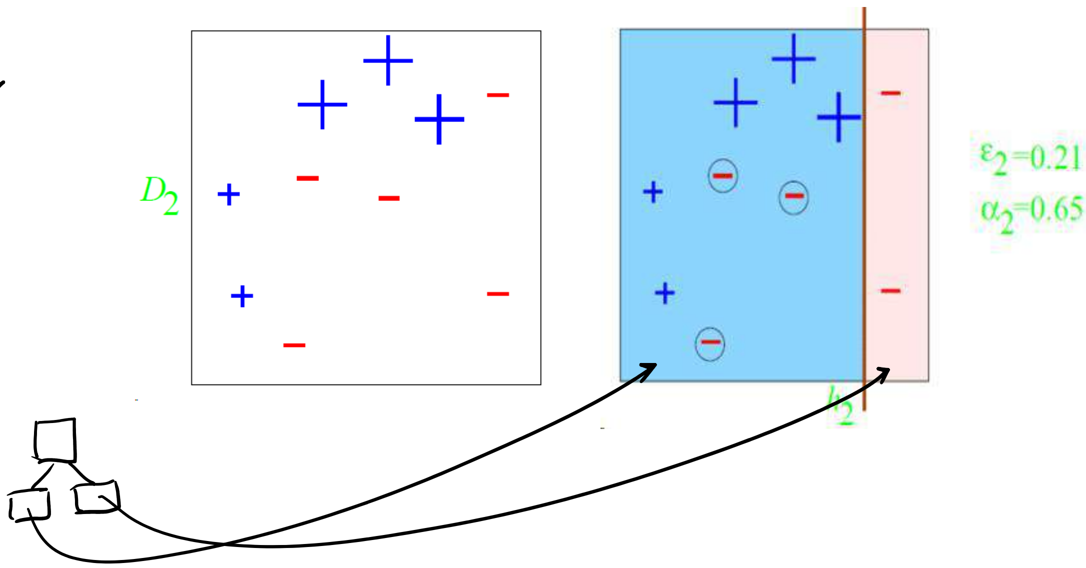
dtc.fit(x_train, y_train)

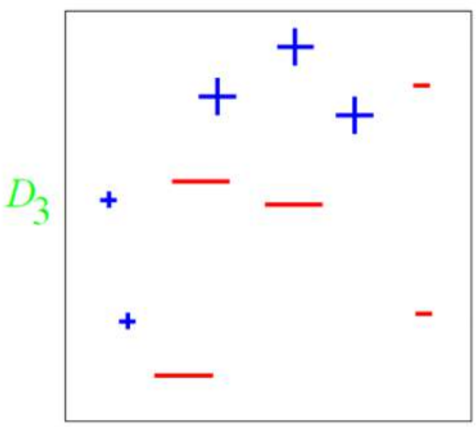
dtc.predict(x_train)

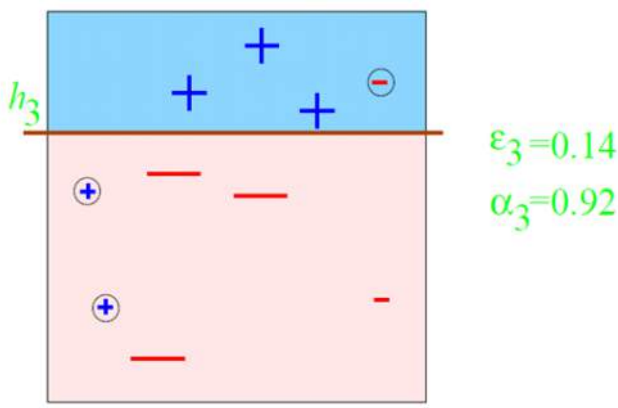


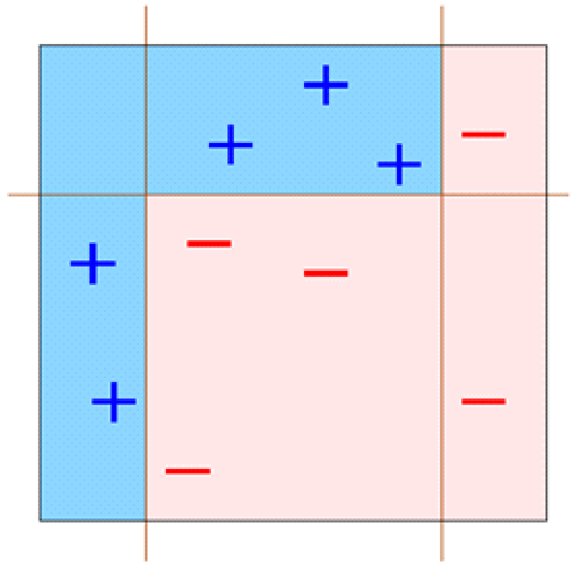


	X_1	X_2	y		
1)			1	+	*
2)			0	-	
3)			1	+	
4)			0	-	
5)			1	*	
6)			0		
7			0		
8			1	*	
9					
10			0		
1)			1		
5)			1		









Gradient Boosting

Train

X_1	X_2 ...	X_p	y_{train}	Pred	$error_1$	pred	err_2
			42	41.2	0.8	0.9	-0.1
			39	30	9	10	-1
			22	28	-6	-5	-1
			58	68	-10	-11	1
			49	41	8	4	4

dtr.fit(X_train, y_train)
dtr.predict(X_train)
m_1

dtr.fit(x_train, $error_1$)
m_2

dtr.fit(X_train, err_2)
m_3 ...... m_{10}

Test

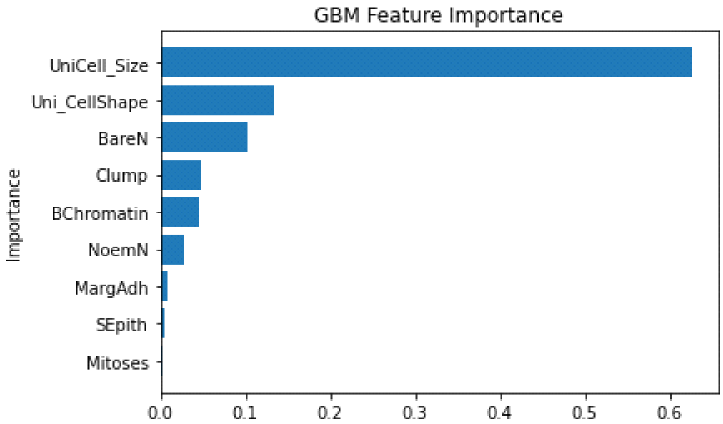
$0 < \eta \leq 1$

$X_1\ X_2 \ldots X_p$

prediction = m1.predict(X_test) + η m_2.predict(X_test) ....
+ η m_{10}.predict(X_test)

Params :-

1) Learning rate η learning_rate
2) Maximum depth max_depth
3) No. of trees n_estimators


GBM Feature Importance
Importance
UniCell_Size
Uni_CellShape
BareN
Clump
BChromatin
NoemN
MargAdh
SEpith
Mitoses
0.0
0.1
0.2
0.3
0.4
0.5
0.6

Stacking

Friday, November 10, 2023 10:03 AM

Image Segmentation

XG Boost

SVM-Linear LDA GaussianNB Logistic Reg

Randomized search

Concrete Strength

Final estimator : [xgb, rf, lgbm, catboost]

Elastic net [ ]

Pipe + K-NN [ ]

DT Regressor [ ]

t-SNE

Wednesday, November 15, 2023 12:01 PM

t-SNE [1] is a tool to visualize high-dimensional data. It converts similarities between data points to joint probabilities and tries to minimize the Kullback-Leibler divergence between the joint probabilities of the low-dimensional embedding and the high-dimensional data. t-SNE has a cost function that is not convex, i.e. with different initializations we can get different results.

It is highly recommended to use another dimensionality reduction method (e.g. PCA for dense data or TruncatedSVD for sparse data) to reduce the number of dimensions to a reasonable amount (e.g. 50) if the number of features is very high. This will suppress some noise and speed up the computation of pairwise distances between samples. For more tips see Laurens van der Maaten's FAQ [2].

High dimen

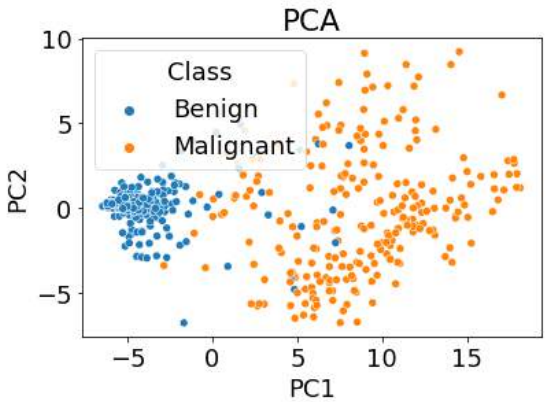
$\underbrace{X_1 \quad X_2 \ldots \quad X_{100}}$

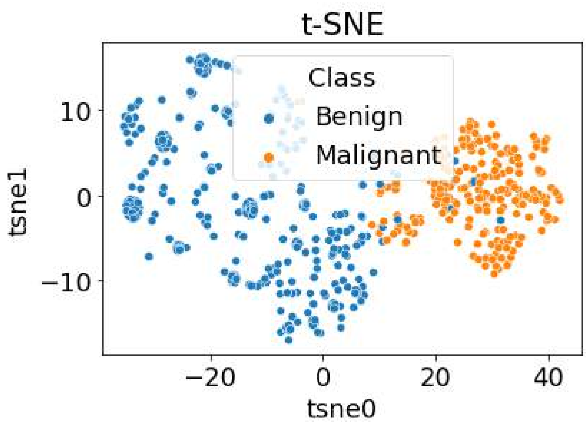
Lower dimensional

$\begin{bmatrix} C_1 & C_2 & C_3 \\ & & \end{bmatrix}$

Joint Prob Distribution

Joint Prob $\leftarrow$ KL Divergence $\rightarrow$ Comparison





https://www.enjoyalgorithms.com/blog/tsne-algorithm-in-ml

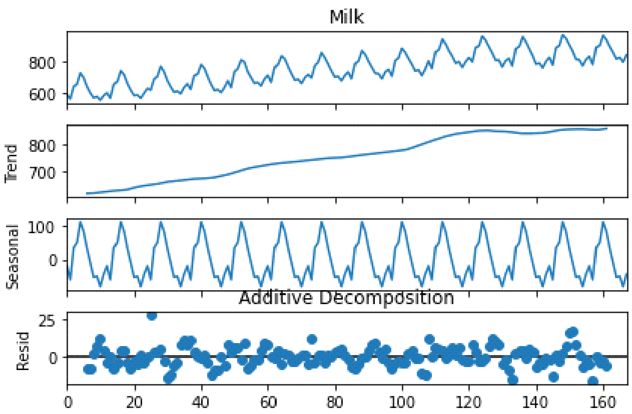
Imbalanced Classification

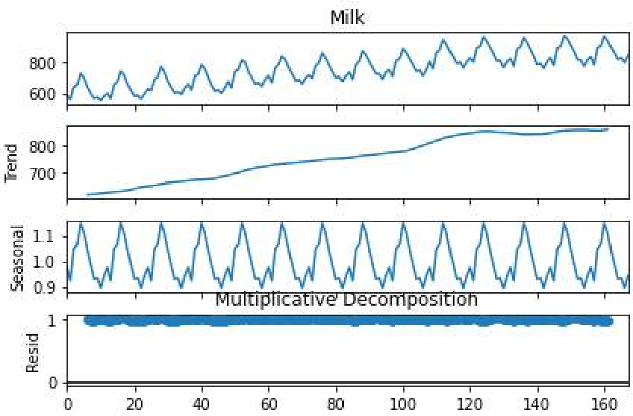
Wednesday, November 15, 2023 7:05 PM

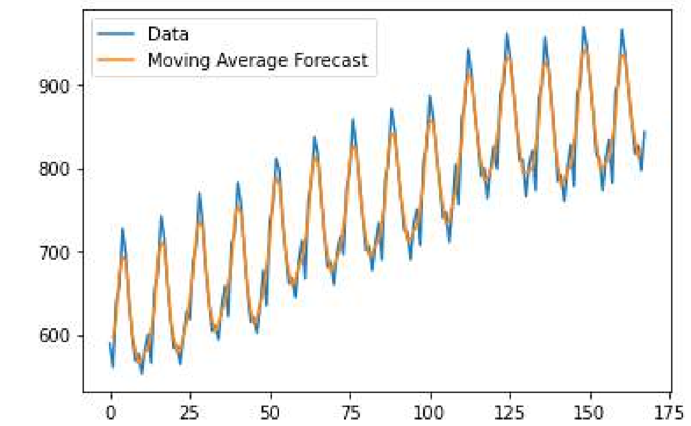
Technique		
Under Sampling	0 6599 1 220	Choose 220 at random from 6599 observations with '0' 0 - 220 1 - 220
Random Over Sampling	0 6599 1 220	Choose with replacement at random 6599 observations from 220 observations with '1' 0 - 6599 1 - 6599
SMOTE	0 6599 1 220	Synthetic observations are generated as linear combinations of existing '1' observations
ADASYN	0 6599 1 220	Synthetic observations with slight random deviations are generated as linear combinations of existing '1' observations

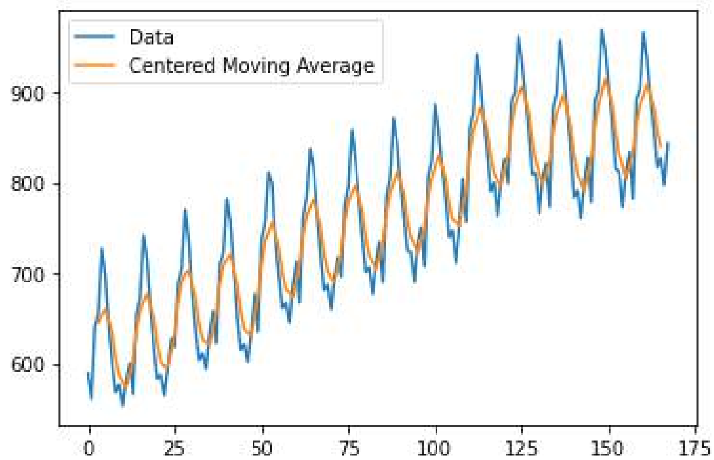
Smoothing

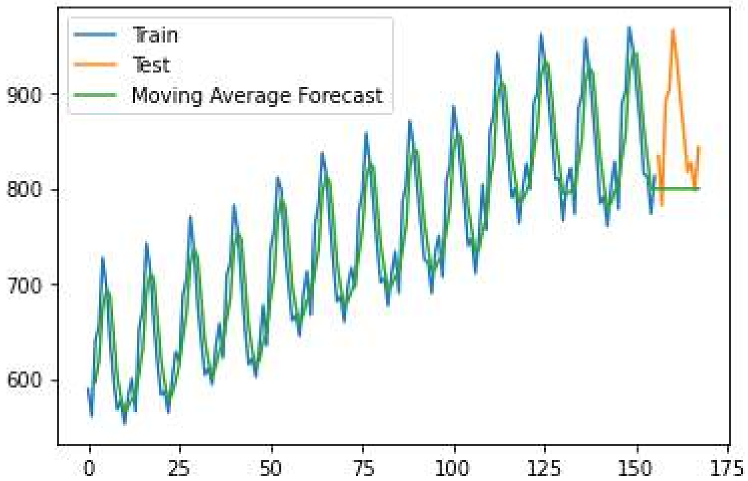
Thursday, November 16, 2023 12:33 PM

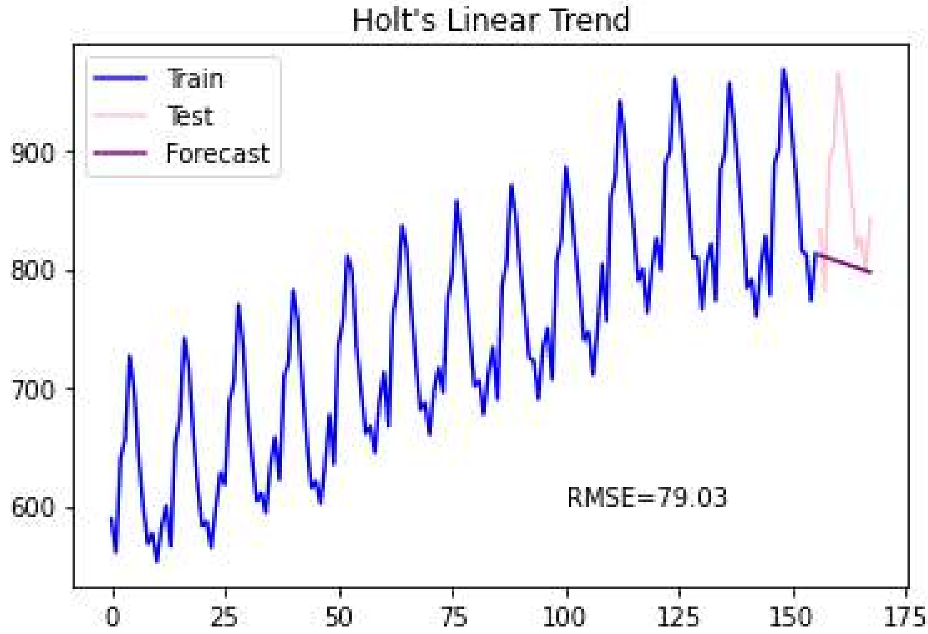











Holt's Linear Trend
Train
Test
Forecast
900
800
700
600
0
25
50
75
100
125
150
175
RMSE=79.03

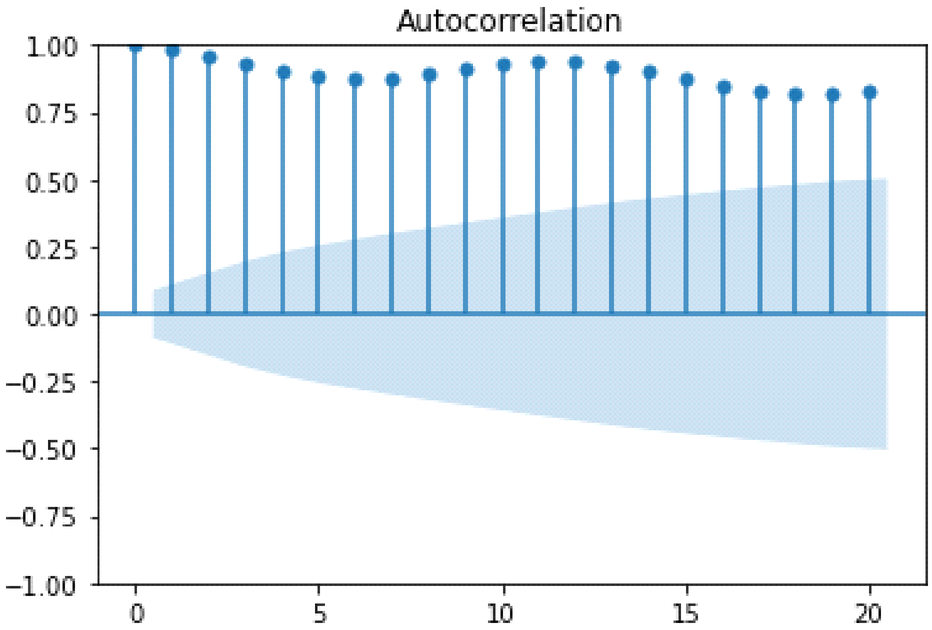
ARIMA

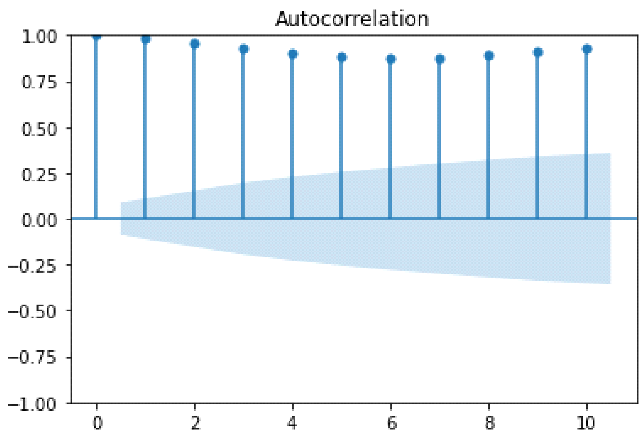
Thursday, November 16, 2023 7:48 PM

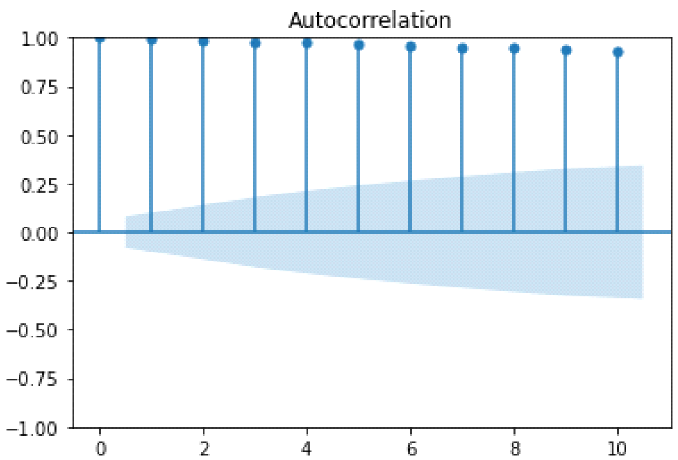
1st order

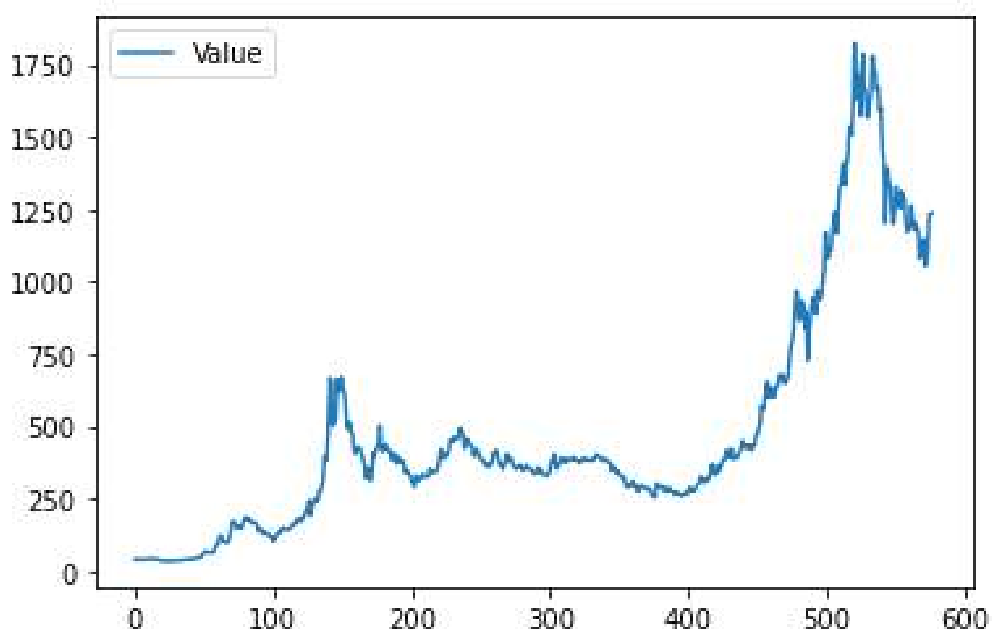
2nd order

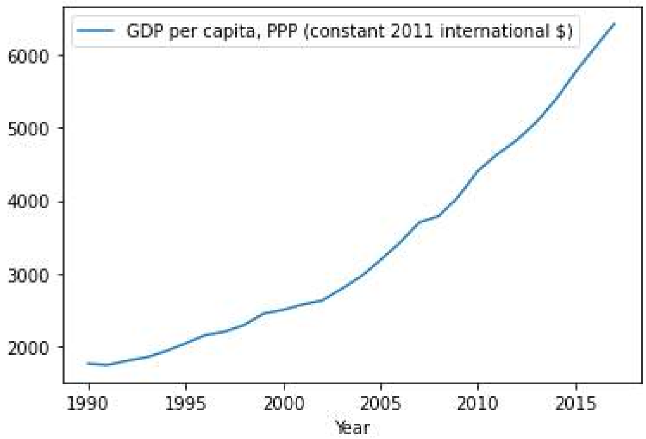
y_1	$[y_t - y_{t-1}]$	
y_2	$y_2 - y_1$	$(y_3 - y_2) - (y_2 - y_1)$
y_3	$y_3 - y_2$	$(y_4 - y_3) - (y_3 - y_2)$
y_4	$y_4 - y_3$	
y_5	$y_5 - y_4$	
y_6	$y_6 - y_5$	

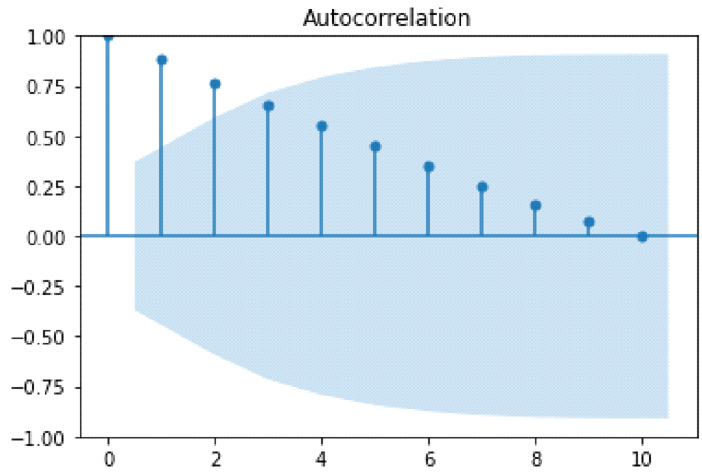












Auto-Regressive Models (AR)

1st ord

e.g. $y_t = b_0 + \phi_1 \, \text{Lag}\,1 + \epsilon_t$

Moving Average Models (MA)

1st ord

e.g $y_t = b_0 + \theta_1 \epsilon_{t-1} + \epsilon_t$

Auto-Reg + Moving Avg (ARMA)

Auto-Reg + Moving Avg (ARMA)

e.g. 1st ord 1st ord

$$y_t = b_0 + \phi_1 \, \text{Lag1} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

Differencing + ARMA $\longrightarrow$ ARIMA

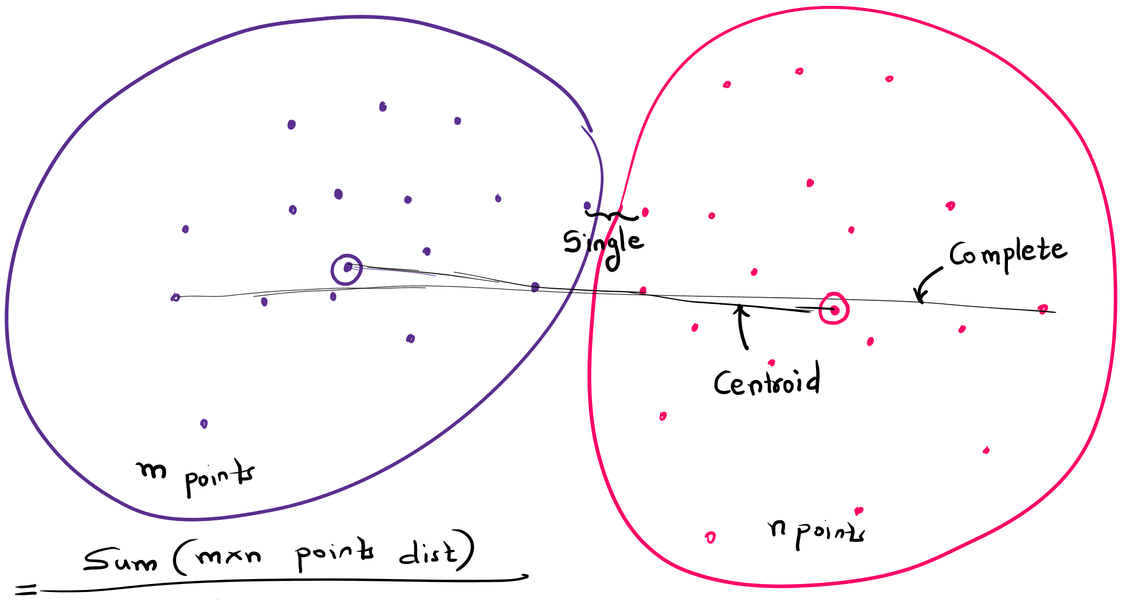
$$(\underbrace{p}_{\text{order of AR}}, \underbrace{d}_{\text{order of differencing}}, \underbrace{q}_{\text{order of MA}})$$

Seasonal ARIMA (SARIMA)

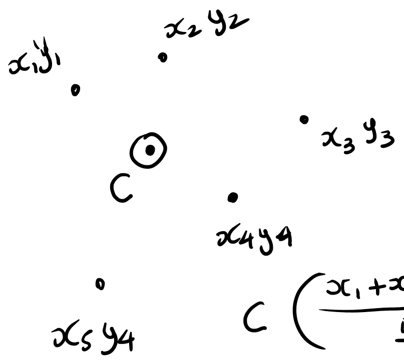
$$(p, d, q)(\underbrace{P}_{\text{order SAR}}, \underbrace{D}_{\text{order SD}}, \underbrace{Q}_{\text{order SMA}})[\underbrace{S}_{\text{Seasonal period}}]$$

Clustering

Friday, November 17, 2023 11:41 AM

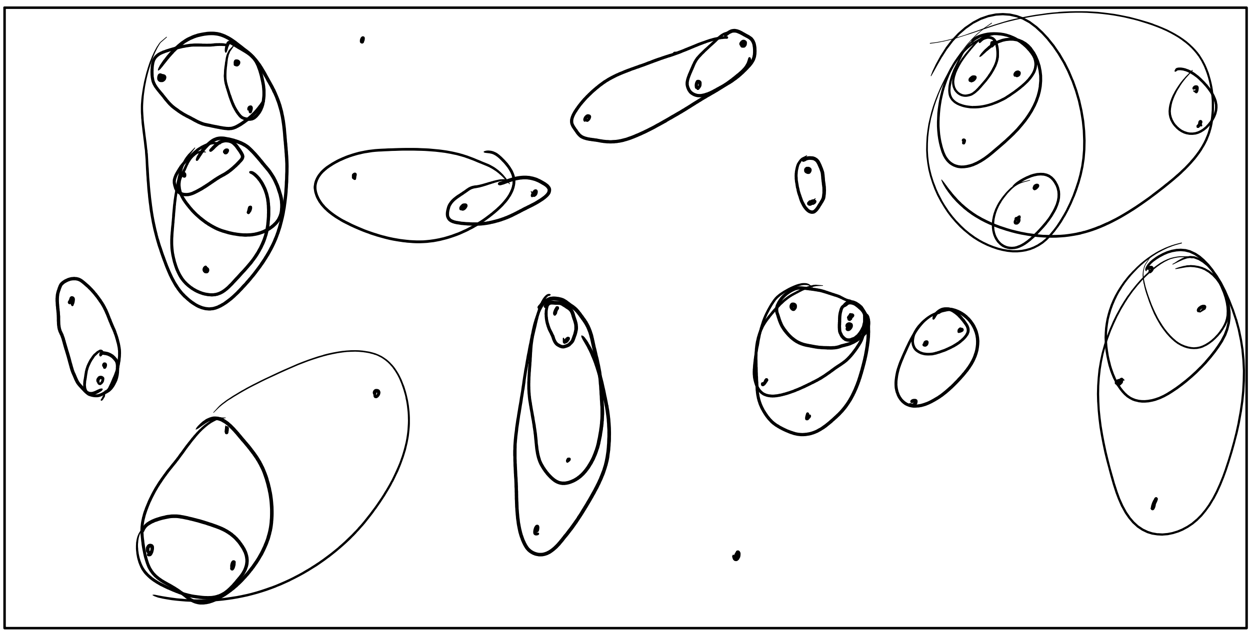


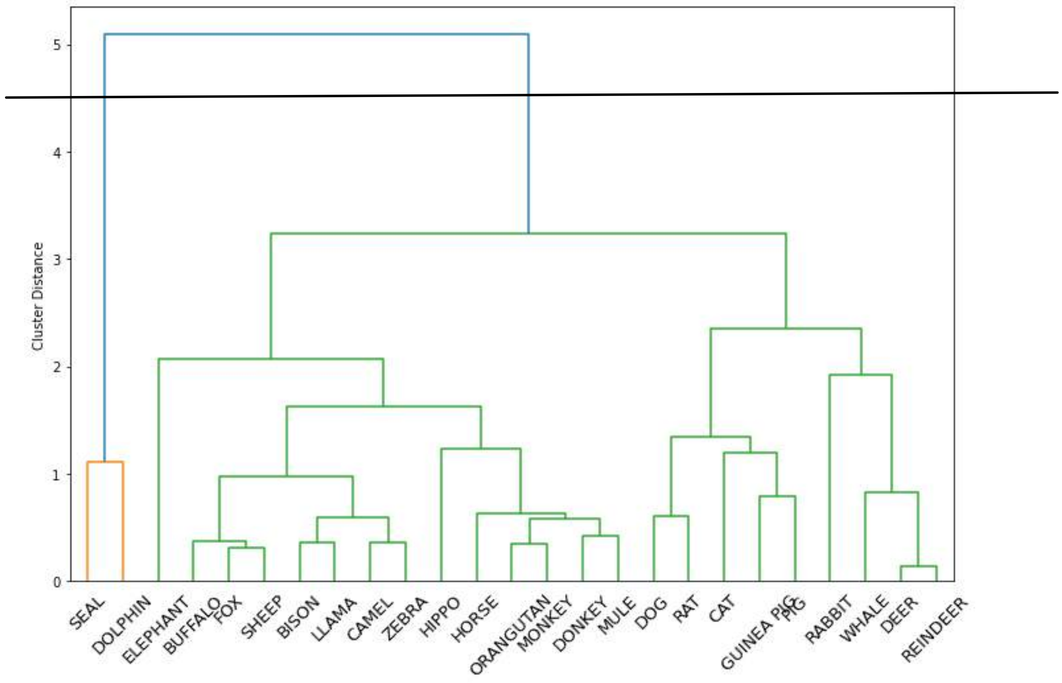
$$\text{Avg} = \frac{\text{Sum}\,(m \times n \text{ points dist})}{m \times n}$$



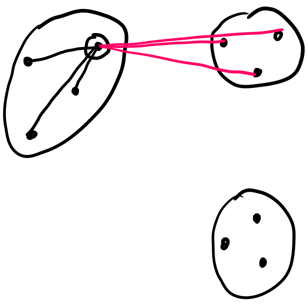
Centroid Formula

$$C\left(\frac{x_1+x_2+\cdots+x_5}{5}, \frac{y_1+y_2+\cdots y_5}{5}\right)$$





Silhouette Score

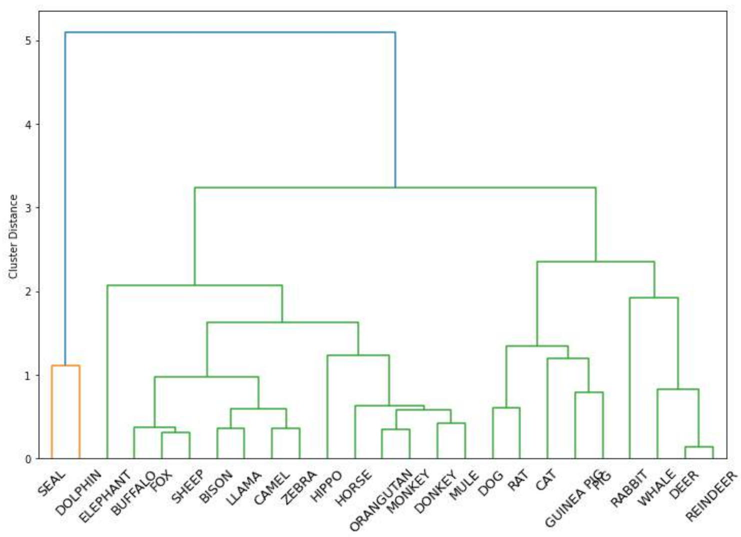


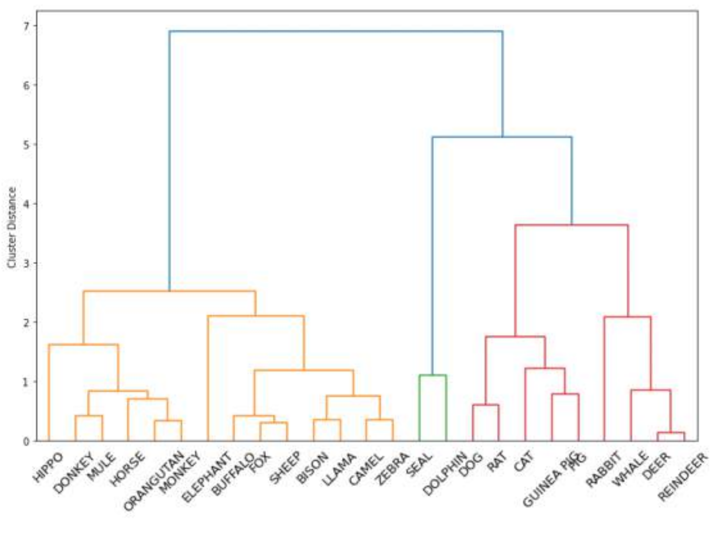
a: Average Intra-cluster distance

b: Average inter-nearest cluster distance

$$\text{sil coef} = \frac{b-a}{\max(a,b)}$$

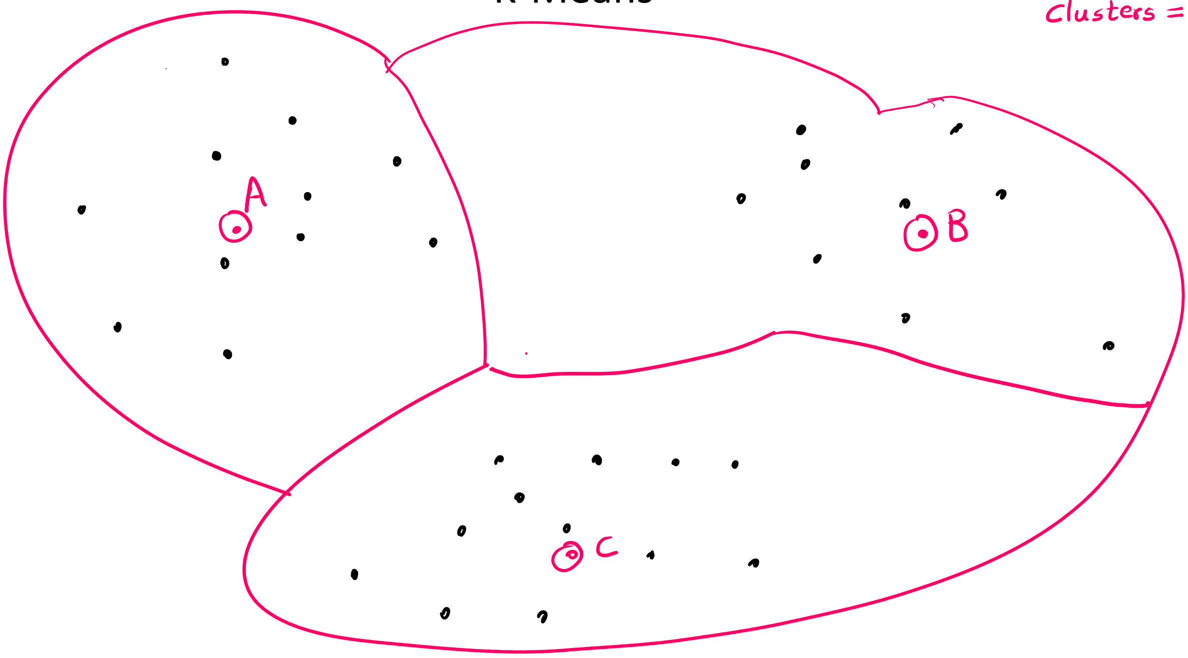
Silhouette Score : Arithmetic mean of sil. coefficients of all point

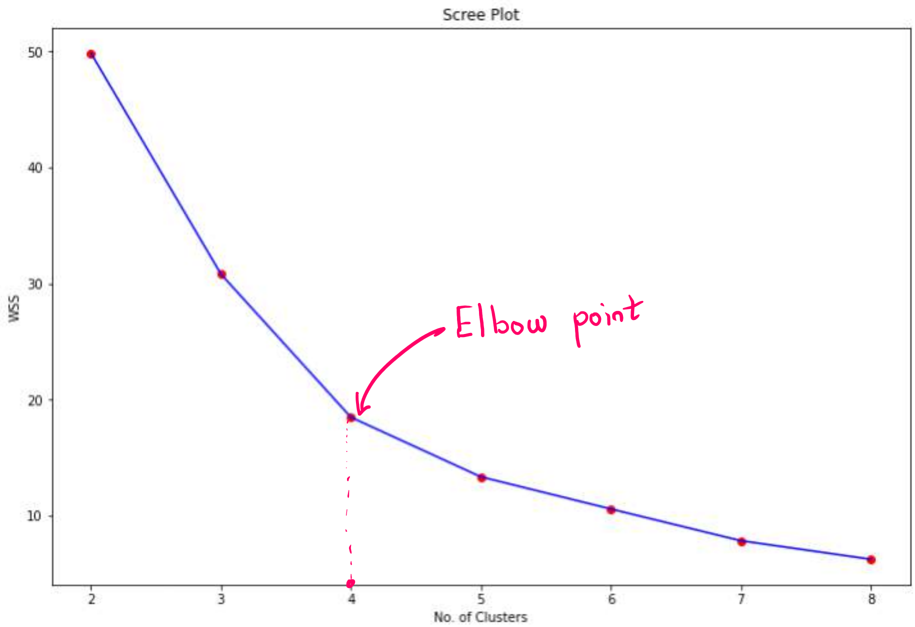




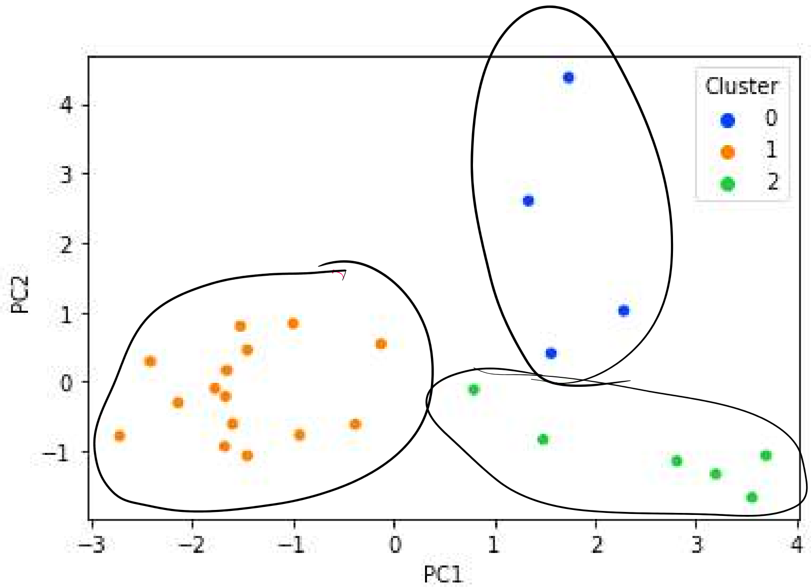
K-Means

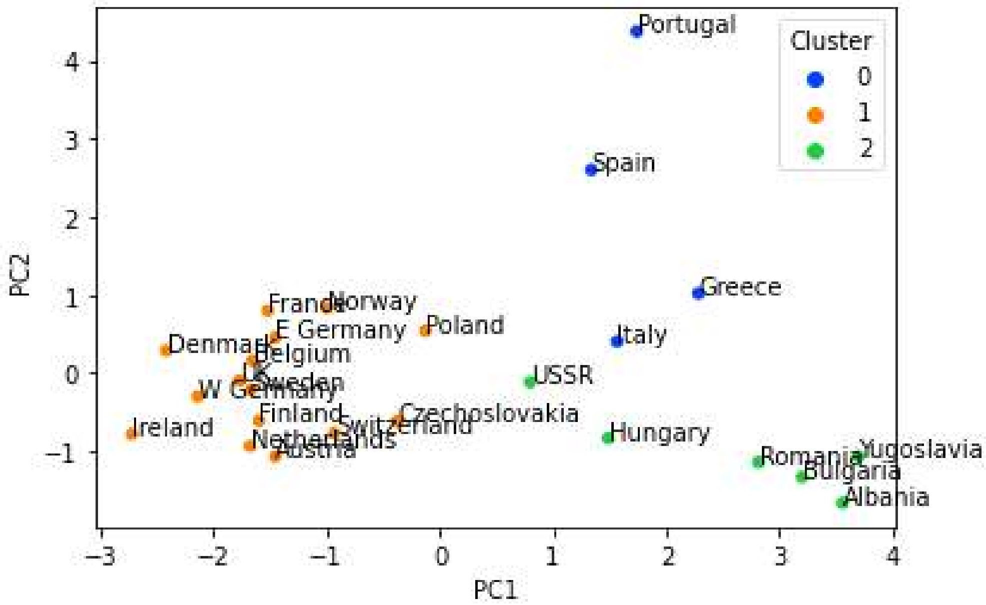
Clusters = 3

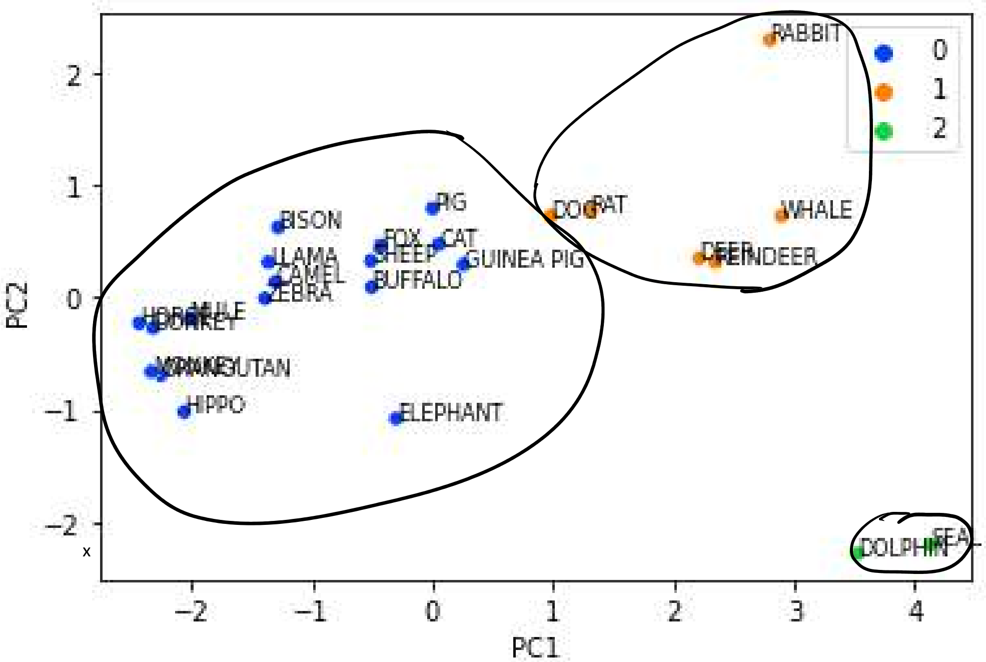


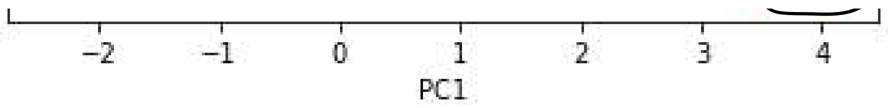


By elbow method with WSS Best = 4


Cluster
0
1
2
PC1
PC2


Cluster
0
1
2
Portugal
Spain
Greece
Italy
France
Norway
E Germany
Poland
Belgium
Sweden
W Germany
Finland
Switzerland
Czechoslovakia
Ireland
Netherlands
Austria
USSR
Hungary
Romania
Yugoslavia
Bulgaria
Albania
PC1
PC2


0
1
2
RABBIT
DOG
RAT
WHALE
DEER
REINDEER
PIG
BISON
FOX
CAT
SHEEP
LLAMA
GUINEA PIG
CAMEL
BUFFALO
ZEBRA
MULE
DONKEY
ORANGUTAN
HIPPO
ELEPHANT
DOLPHIN
PC1
PC2

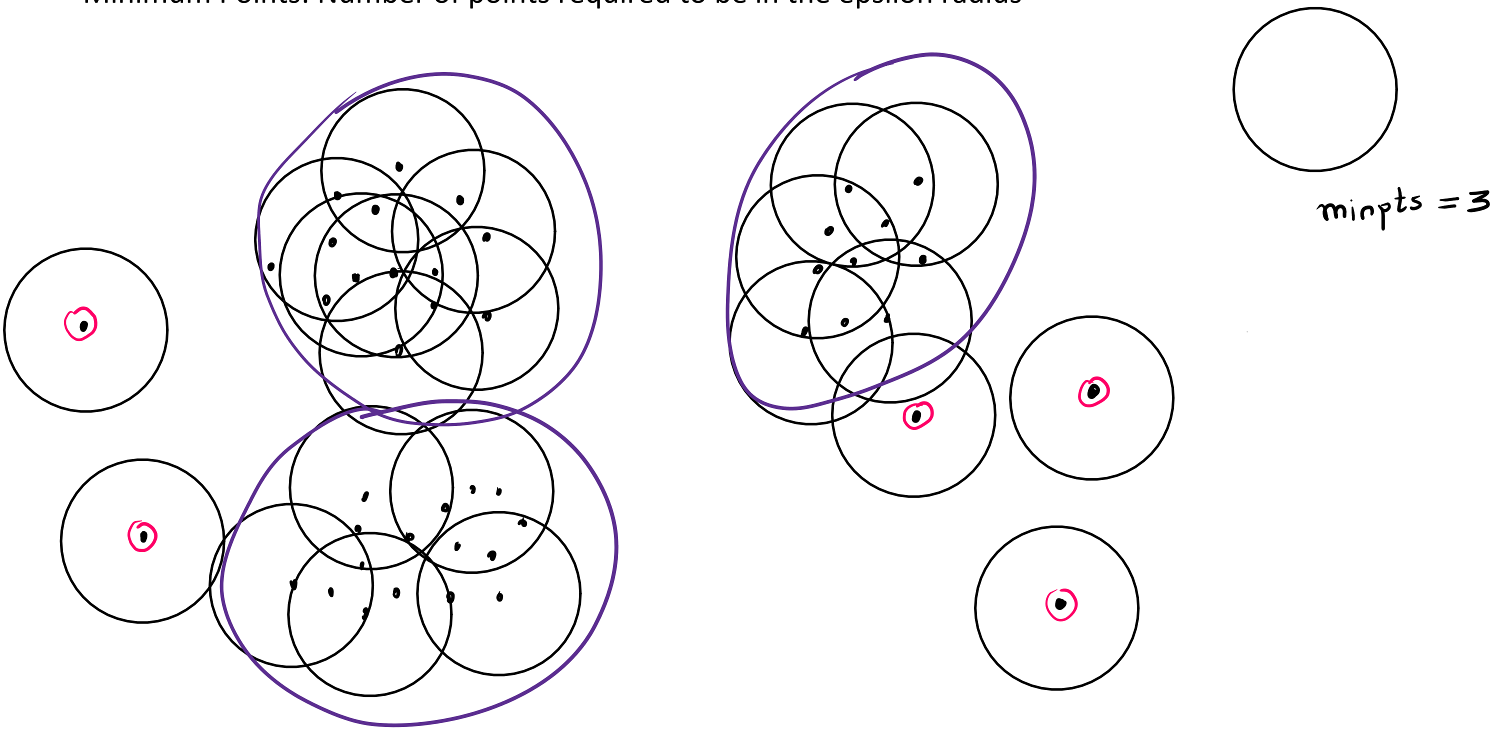

−2
−1
0
1
2
3
4
PC1

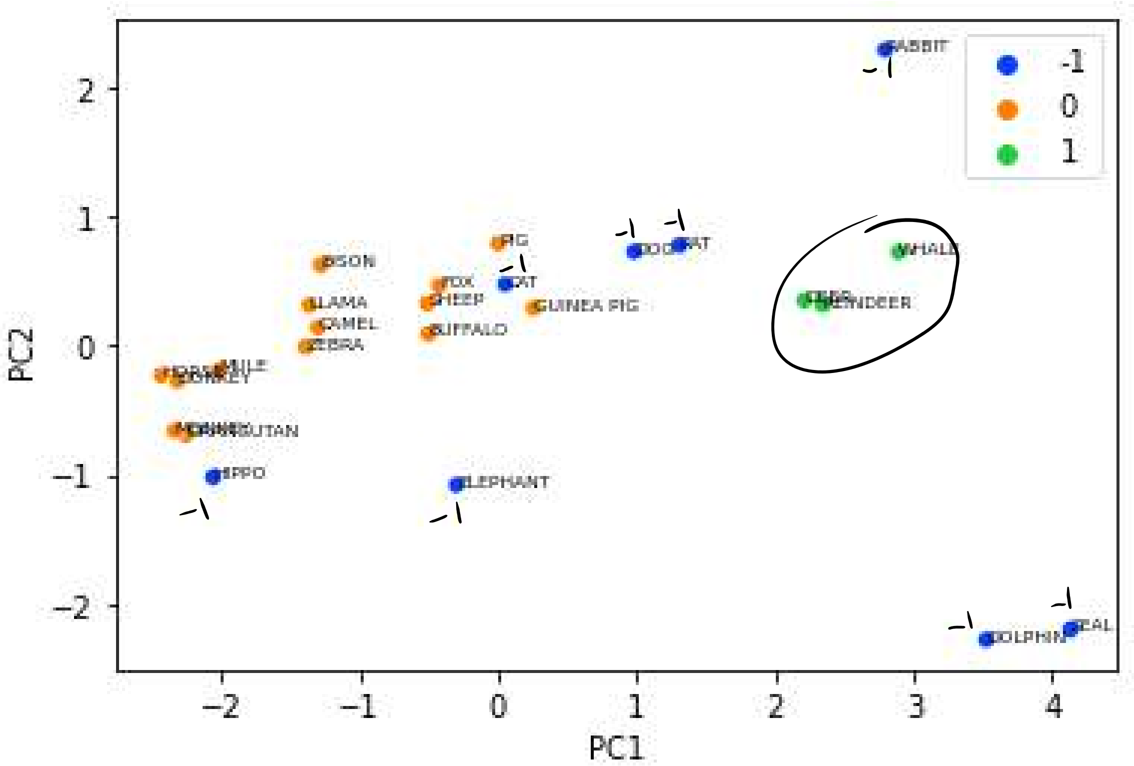
DBSCAN

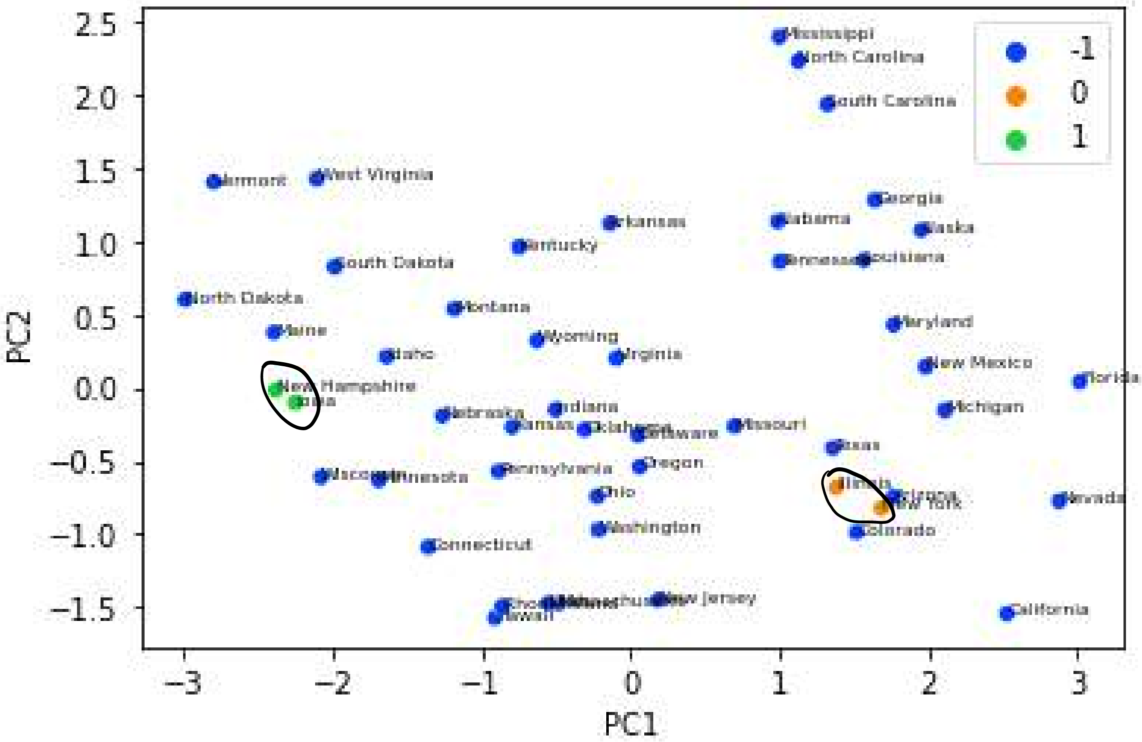
Saturday, November 18, 2023 10:45 AM

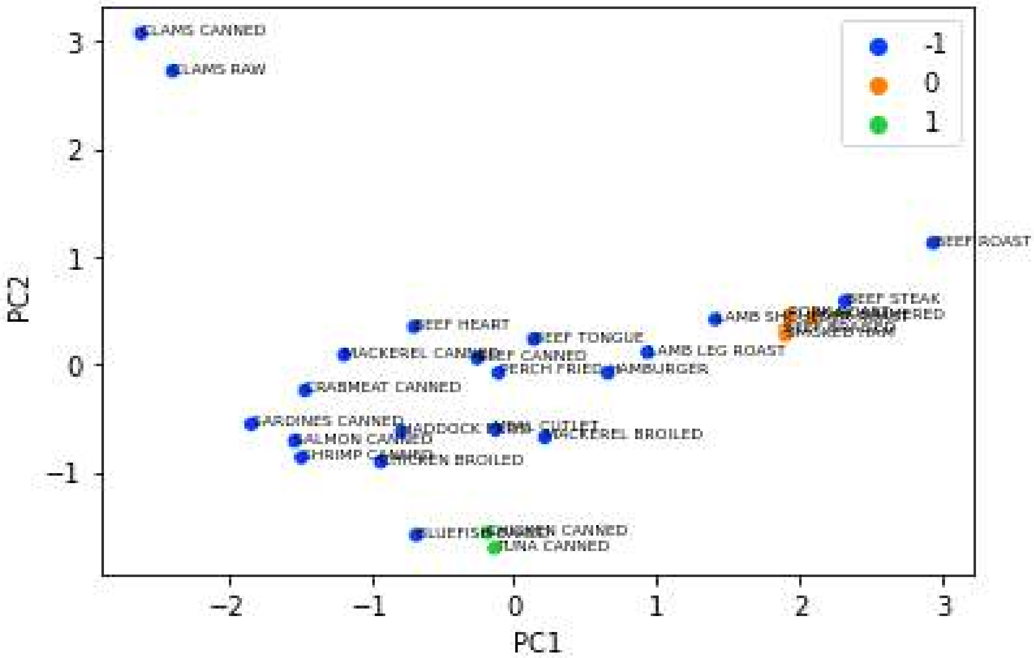
Density-based spatial clustering of applications with noise

- Epsilon: Radius for the neighbourhood of points
- Minimum Points: Number of points required to be in the epsilon radius






PC1
PC2
-1
0
1
North Carolina
South Carolina
West Virginia
South Dakota
North Dakota
New Hampshire
New Mexico
Pennsylvania
Washington
Connecticut
Minnesota
New Jersey
New York
California


PC1
PC2
-1
0
1
CLAMS CANNED
CLAMS RAW
BEEF ROAST
BEEF STEAK
BEEF HEART
BEEF TONGUE
LAMB LEG ROAST
CRABMEAT CANNED
SARDINES CANNED
SALMON CANNED
CHICKEN BROILED
MACKEREL BROILED
PERCH FRIED
HAMBURGER
BEEF CANNED
SMOKED HAM
CHICKEN CANNED
TUNA CANNED

Association Rules

Saturday, November 18, 2023 2:14 PM

Conf. > Benchmark Conf.

$$\text{Lift Ratio} = \frac{\text{Conf.}}{\text{B. Conf.}} > 1$$

For applying Association Rules on any data:

1. Make the data completely categorical. i.e. make all variables categorical. Those variables which are not categorical can be binned with functions like pandas.cut( )
2. Do the one hot encoding of the data
3. Apply the Apriori and Association Rules functions to see the associations within the data

Recommender Sys

Monday, November 20, 2023 4:03 PM

Hit Ratio :-

User A Already Bought :- $[I_8, I_9, I_{11}, I_{29}, I_{38}, I_{56}, I_{42}, I_3]$

Recommend Top 5 :- $[I_9, I_{10}, I_{50}, I_{56}, I_{100}]$

$$HR = \frac{2}{5}$$

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	4	6	?	?	3
User 2	5	?	1	?	5
User 3	8	10	?	1	7
User 4	?	2	2	4	?
User 5	2	?	8	9	2

Sparse Matrix Format

UID	IID	Rating
1	1	4
1	2	6
1	5	3
2	1	5
2	3	1
2	5	5
3	1	8
3	2	10
3	4	1
3	5	7
4	2	2
4	3	2
4	4	4
5	1	2
5	3	8
5	4	9
5	5	2

Matrix Factorization

Tuesday, November 21, 2023 2:46 PM

no. of
K : Latent factors

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	4	6	?	?	3
User 2	5	?	1	?	5
User 3	8	10	?	1	7
User 4	?	2	2	4	?
User 5	2	?	8	9	2

$$R_{m\times n} = \left[P_{m\times k}\right]\left[Q_{k\times n}\right]$$

$$r_{ij} = \sum_{l} p_{il}\, q_{lj}$$

$$\hat{r}_{12} = \sum_{l} \hat{p}_{1l}\, \hat{q}_{l2}$$

$R_{m\times n}$

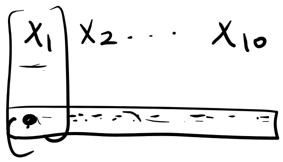
m users

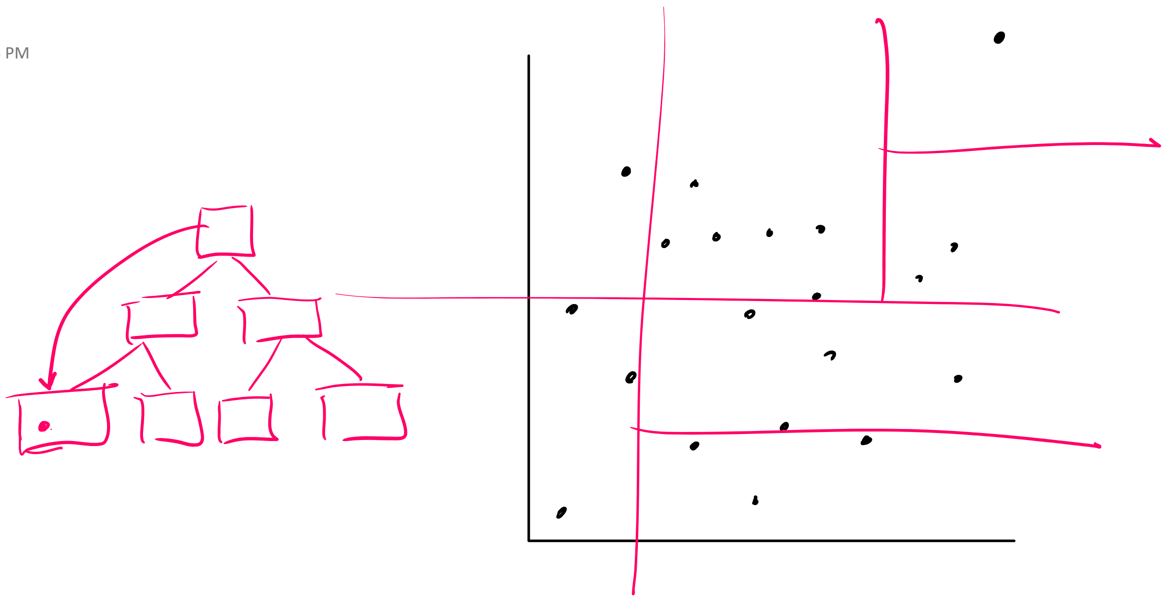
n items

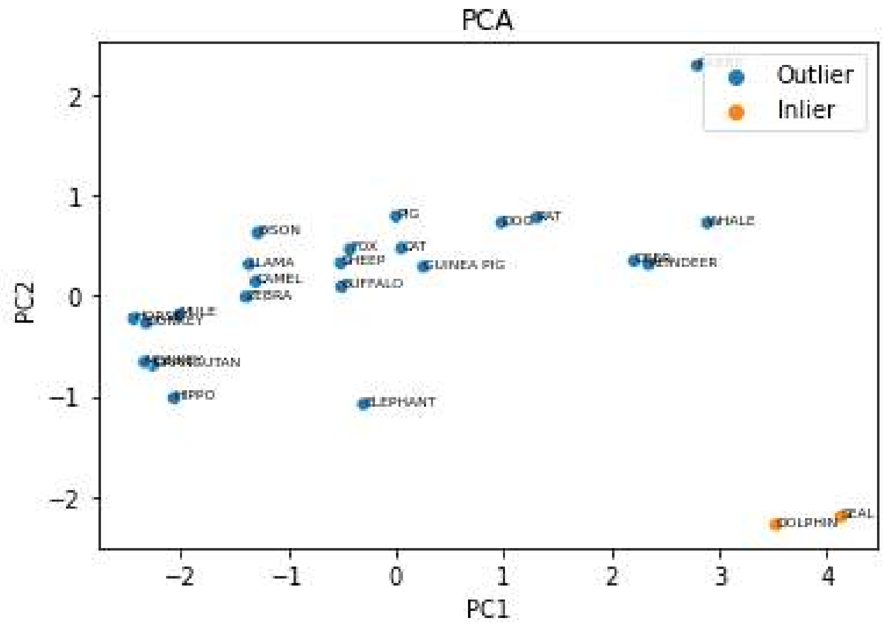
$$\sum_{i,j} (r_{ij} - \hat{r}_{ij})^2 \text{ : error } f^n$$

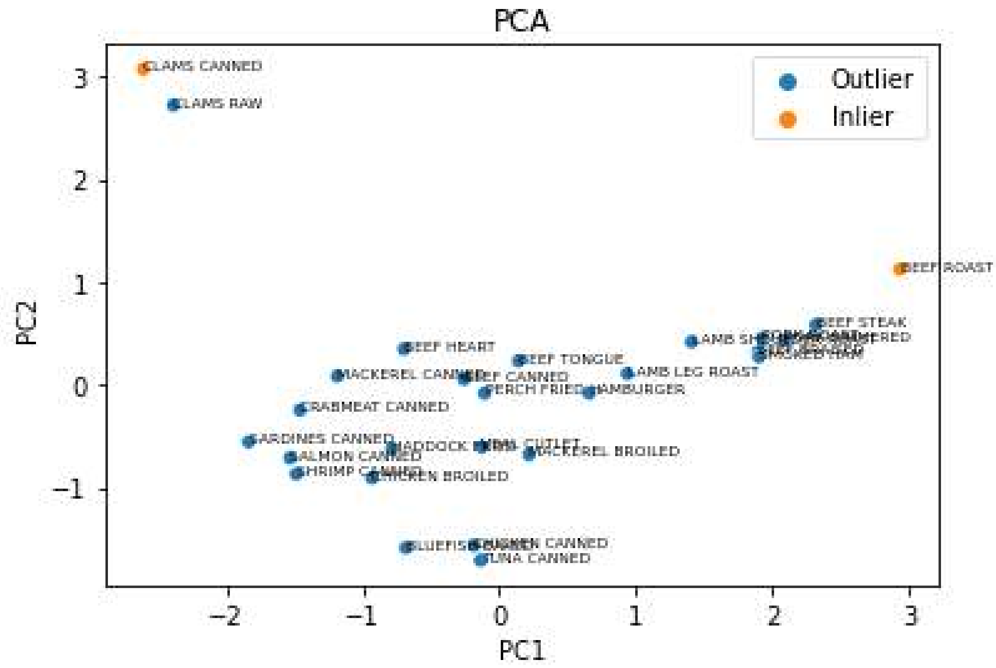
Isolation Forest

Tuesday, November 21, 2023 3:46 PM








PCA
Outlier
Inlier
PC1
PC2
CLAMS CANNED
CLAMS RAW
BEEF ROAST
BEEF STEAK
BEEF HEART
BEEF TONGUE
LAMB LEG ROAST
MACKEREL CANNED
PERCH FRIED
HAMBURGER
CRABMEAT CANNED
SARDINES CANNED
MACKEREL BROILED
CHICKEN BROILED
TUNA CANNED