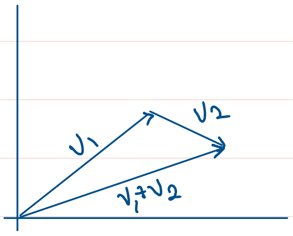
# Linear Algebra

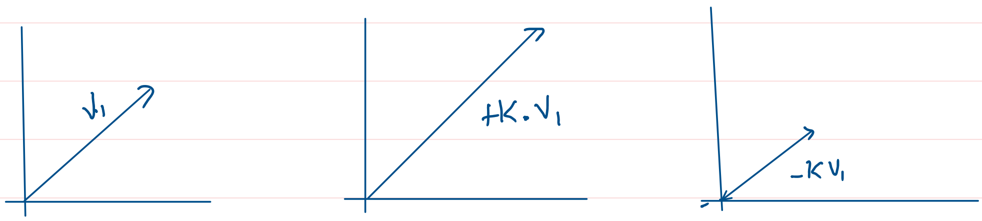
16 April 2024 22:00

- Vectors

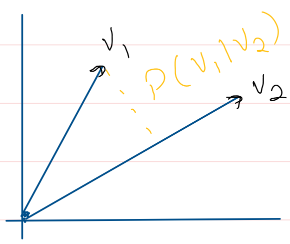
→ Vector Addition (Dot Product)




→ Scalar Multiplication

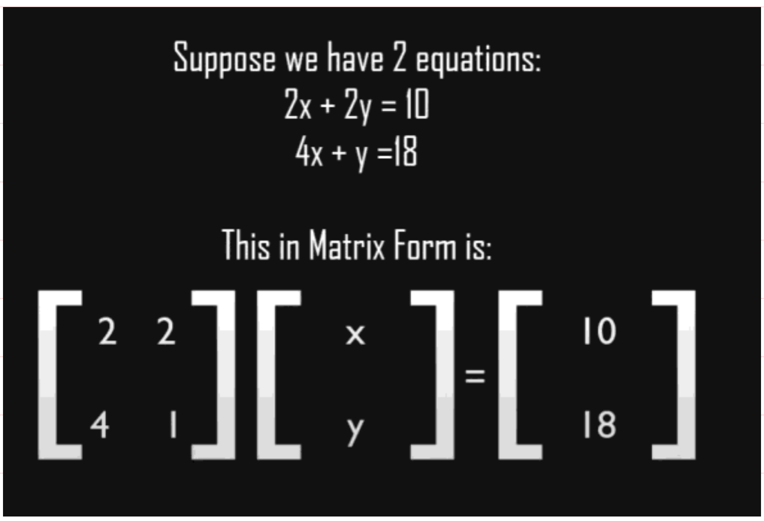



→ Projection




If we are able to project $v_1$ on $v_2$ then we are able to find unknown features of $v_1$.

- Matrices



→ Matrix Addition

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} a+A & b+B \\ c+C & d+D \end{bmatrix}$$

Same order $\Rightarrow$ must

→ Matrix Multiplication

$$\begin{bmatrix} 2 & 2 \\ 4 & 1 \end{bmatrix} \times \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 2*2+2*1 & 2*3+2*4 \\ 4*2+1*1 & 4*3+1*4 \end{bmatrix} \rightarrow \begin{bmatrix} 4+2 & 6+8 \\ 8+1 & 12+4 \end{bmatrix} \rightarrow \begin{bmatrix} 6 & 14 \\ 9 & 16 \end{bmatrix}$$

→ Transpose (Row $\rightleftharpoons$ Col)

$$x = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \qquad x^T = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$$

→ Determinant of Matrix

$$x = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$

it gives the scalar value of the Matrix

$$n = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad \text{Matrix}$$

$$Det(n) = a * \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b * \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

→ Inverse of Matrix

- matrix when multiplied with inverse gives identity Matrix.
- it helps to apply transformations easily.
- Some matrix do not have an inverse that mean they contain very noisy & un-reliable data.

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

$$Inv(A) = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

For order 3 and above

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$Inv(A) = \frac{1}{det(A)} \begin{bmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & \begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ \begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & \begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ \begin{vmatrix} b & c \\ e & f \end{vmatrix} & \begin{vmatrix} a & c \\ d & f \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix}$$

→ Scaling

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix}$$

→ Shearing

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ m & 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ m & 1 \end{bmatrix}$$

→ Rotation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

→ Row Echelon method

$$2x + y - z = 2$$
$$x + 3y + 2z = 1$$
$$x + y + z = 2$$

matrix

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

ans

$x = 2$

$y = -1$

$z = 1$

Step 1 $\quad R_1 = R_1/2$

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

Step 2 $\quad R_2 = R_2 - R_1$

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 1 \\ 1 & 3 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 5/2 & 5/2 & 0 \\ 1 & 1 & 1 & 2 \end{bmatrix}$$

Step 3 $\quad R_3 = R_3 - R_1$

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 5/2 & 5/2 & 0 \\ 1 & 1 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 5/2 & 5/2 & 0 \\ 0 & 1/2 & 3/2 & 1 \end{bmatrix}$$

Step 4 $R_2 = \frac{2}{5}(R_2)$

$$\begin{bmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 5/2 & 5/2 & 0 \\ 0 & 1/2 & 3/2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1/2 & 3/2 & 1 \end{bmatrix}$$

Step 5 $R_1 = R_1 - \frac{1}{2}(R_2)$

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

Step 6 $R_3 = R_3 - \frac{1}{2}(R_2)$

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

Step 7 $R_1 = R_1 + R_3$

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

Step 8 $R_2 = R_2 - R_3$

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

$x = 2$
$y = -1$
$z = 1$

only they should have values (1)
rest all should be zero.

∴ Transform matrix such that only diagonals have 1's rest all other should contain 0's.

→ Inverse

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

lets suppose $AB = C$

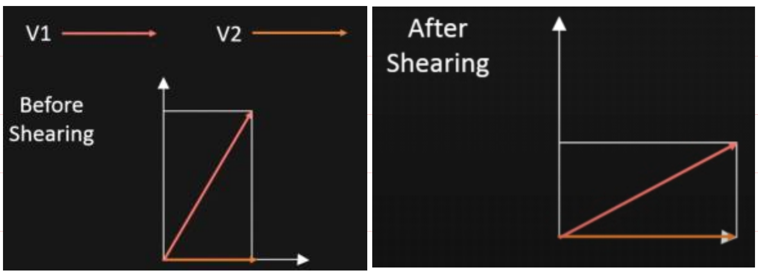
$\Rightarrow A A^{-1} B = C A^{-1}$

$\Rightarrow \underline{\underline{B = C A^{-1}}}$

$$A^{-1} = \frac{1}{ab - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

→ Eigen Vector

- doesnt change direction even after transformation is applied.




this is the reason even after applying shearing operation vectors remains the same

→ Uses of Linear algebra

a) PCA

a) PCA
b) Pictures
c) Dataset Encoding
d) Singular Value decomposing
e) latent Semantic Analysis for NLP

## → Differentiation :-

| | |
|---|---|
| Constant Rule | $\frac{d}{dx}[C] = 0$ |
| Power Rule | $\frac{d}{dx}x^n = nx^{n-1}$ |
| Product Rule | $\frac{d}{dx}[f(x)g(x)] = f'(x)g(x) + f(x)g'(x)$ |
| Quotient Rule | $\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}$ |
| Chain Rule | $\frac{d}{dx}[f(g(x))] = f'(g(x))\,g'(x)$ |

$x^n = n x^{n-1}$

$x^a + x^b = a x^{a-1} + b x^{b-1}$

$x^a \cdot x^b = a x^{a-1} x^b + b x^{b-1} x^a$

## → Partial Differentiation :-

*Partial Differentiations*

$f^1(x, y, z) =$

*With respect to x*

$(y, z).f^1(x)$

*With respect to y*

$(x, z).f^1(y)$

*With respect to z*

$(x, y).f^1(z)$

*Complete Differentiation*

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

Eg: $f = x^2 + 3y + 4z^2 x$

$x \to 2x + 4z^2$

$y \to 3$

$z \to 8zx$

$f(x,y,z) = 2x + 4z^2 + 3 + 8zx$

## → Applications -

a) Jacobian helps find global maxima of dataset.
Also helps linearizing a non-linear function to linear at a point.

b) Hessian helps minimize error.

c) Deep learning

d) Gradient Descent method for optimizing weights.

# Probability

17 April 2024 19:06

→ it is the measure of how likely an event will occur.

$P$ = desired / total

Events → Joint (Common Outcome)
→ Dis-Joint (un-Common OutCome)

Distribution → Probability Density function
⇒ it describes a Continious distribution
→ Normal Distribution
⇒ it depends on mean & Std deviation.

→ Central Limit Theorm
⇒ Sampling distribution of mean of any independent random variable will be normal, if Sample Size is big enough.

Probability → Marginal
⇒ probability of occurence of Single event.

→ Joint Probability
⇒ Prop. of two events happing at Same time

→ Conditional Probability

⇒ Prob. of event B is that the event will occur given that event A has alredy occured.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$P(A)$

→ Bayes Theorm

Shows relation b/w one Conditional probability & its inverse.

$$P(B|A) = \frac{P(A|B)\ P(B)}{P(B)}$$ → {Prior}

Prob of occurence of B given A
{Posterior}

→ Application

a) optimize Model
b) classification by algo requires Probability
c) Loss can be Calculated by probability

Statistics
18 April 2024 11:44

## Population

Collection or set of individual or object or events whose properties are to be analyzed.

## Sample

Subset of population

## Sampling

- Probability
  - Random
  - Systematic
  - Stratified
- Non-probability
  - Snowball
  - Quota
  - Judgement
  - Convenience

## Statistics

Descriptive

use to provide description of population

{ it focuses manily on the main charecterstics of data. To provide graphical Summary of data. }

Types :-

① Measures of Centre

mean, median & mode

② Measures of Spread

① Range ② Inter-Quartile Range ③ Variance

④ Standard Deviation

Inferential

makes iference and, predictions about a population based on Sample of data taken from population

{ generalize large data Set and apply propability to draw Conclusions }

# Statistics – Information Gain

Entropy

***Entropy measures the impurity or uncertainty present in the data.***

$$H(S) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

***where:***

- ***S – set of all instances in the dataset***
- ***N – number of distinct class values***
- ***pi – event probability***

Information Gain (IG)

***IG indicates how much "information" a particular feature/ variable gives us about the final outcome.***

$$Gain(A,S) = H(S) - \sum_{j=1}^{v} \frac{|S_j|}{|S|} \cdot H(S_j) = H(S) - H(A,S)$$

***where:***

***H(S) – entropy of the whole dataset S***

- ***|Sj| – number of instance with j value of an attribute A***
- ***|S| – total number of instances in dataset S***
- ***v – set of distinct values of an attribute A***
- ***H(Sj) – entropy of subset of instances for attribute A***
- ***H(A, S) – entropy of an attribute A***

# Statistics – Confusion Matrix

- There are two possible predicted classes: "yes" and "no"
- The classifier made a total of 165 predictions
- Out of those 165 cases, the classifier predicted "yes" 110 times, and "no" 55 times
- In reality, 105 patients in the sample have the disease, and 60 patients do not

| n=165 | Predicted: NO | Predicted: YES |
|---|---|---|
| **Actual: NO** | 50 | 10 |
| **Actual: YES** | 5 | 100 |

# Statistics – Confusion Matrix

Actual Values

| Predicted Values | Positive (1) | Negative (0) |
|---|---|---|
| Positive (1) | TP | FP |
| Negative (0) | FN | TN |

**true positives (TP):** These are cases in which we predicted yes (they have the disease), and they do have the disease

**true negatives (TN):** We predicted no, and they don't have the disease

**false positives (FP):** We predicted yes, but they don't actually have the disease. (Also known as a "Type I error.")

**false negatives (FN):** We predicted no, but they actually do have the disease. (Also known as a "Type II error.")