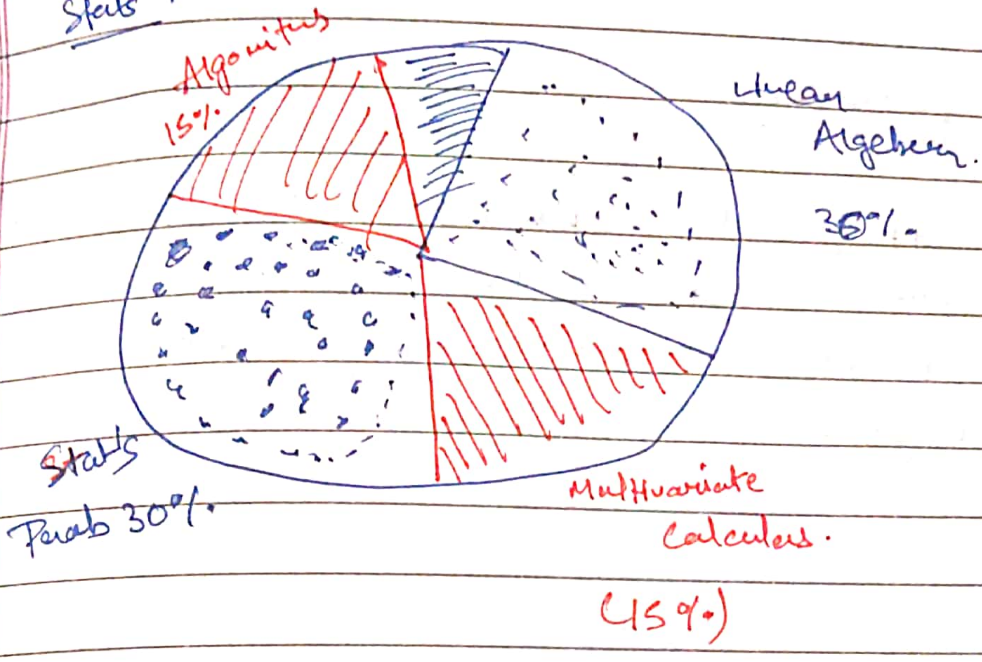
Mathematics for Machine Learning.

① Why Maths in Machine Learning.

→ It's not always about solving the problem mathematically rather understanding when to apply math to the given data or problem.

Stats ML overview.



Multivariate Calculus

- Helps to optimize the Machine Learning Models
- Real World problem uses Calculus for many operations.

Differentiation :-

- Breaking down of function with in-depth Analysis.
- Here It helps us to better find the information.

Derivation formula.

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

(1) Power Rule.

$$y'(x^n) = nx^{n-1}$$

$$\lim_{x \to 0} \frac{3(x+\Delta x)^2 - 3(x)^2}{\Delta x}$$

$$\lim_{x \to 0} \frac{3(x^2+\Delta x^2 + 2x\Delta x) - 3x^2}{\Delta x}.$$

$$\lim_{x\to 0} \frac{3\cancel{\Delta x}(\Delta x + 2x)}{\cancel{\Delta x}}$$

$$= \lim_{x\to 0} 3(\Delta x + 2x)$$

$$3(2x) = 6x$$

Now with Power Rule

$$3x^2 = 3\cdot(2x)^{3-1}$$
$$= 6x.$$

L.H.S = R.H.S.

b) Chain Rule :-

$$f'(g'(x)) = f'(g(x)).g'(x)$$

(c) Product Rule :-

$$f'(x_1 \cdot x_2) = x_1' \cdot x_2 + x_2' \cdot x_1$$

We will be more Concerned with Partial Derivatives.

$$f'(x,y,z) = (y,z)\cdot f'(x) + (x,z)\cdot f'(y) + (x,y)\cdot f'(z)$$

Scanned with CamScanner

AI/ML/DL Applications of Multivariate Calculus.

- Gradient descent method for optimizing weights.
- Hessian helps minimize the Error.
- Used in Deep Learning Models.
- Jacobian helps find the global maxima.

Probability in ML/AI.

With using Probability we make assumptions, hypotheses and more. So it plays a really important role when it comes to Mathematics for Machine Learning.

'Probability is a measure of how likely an event will occur'

$$\text{Probability} = \frac{\text{Event Desired}}{\text{Total Outcomes}}$$

Terminology in Probability :-

- Random Experiment An experiment or a process for ~~per~~ which the outcome cannot be predicted with certainty.
- Sample Space The entire possible set of outcomes.
- Event
  - ↳ Disjoint Event
  - ↳ Joint Event

Date
DELTA Pg No

② **Joint Probability :-** Measure of event happening at the same time.

③ **Conditional Probability :-**

It is the measure of an event that will occur if some other event has already occurred.

Baye's Theorem :-

It is used to calculate the conditional probability. It is the probability of an event occurring based on prior knowledge of conditions that might be related to the event.

formula :-

$$\underbrace{P(A|B)}_{\text{Probability of A when B has happened}} = \frac{\overbrace{P(B|A)}^{\text{Probability of B when A has happened}} \cdot \underbrace{P(A)}_{\text{Probability of A}}}{\underbrace{P(B)}_{\text{Probability of B.}}}$$

Applications of Probability for Machine Learning :-

- Probability helps us optimize our model.
- Classification by our algorithms require Probability
- Loss can also be calculated using Probability
- Models are built on Probability.

Distributions in Probability :-

for Machine Learning, we will only tackle only 3 distributions :-

(i) Probability Density function :-

It is concerned with relative likelihood for a continuous random variable to take on a given value.

(ii) Normal Distribution :-

It is a probability distribution that denotes the symmetric property of the mean.

(iii) Central Limit Theorem :-

It states that sampling distribution of the means of any independent random variable will be normal.

Types of Probability :-

① Marginal Probability :-

means that an event will occur without any intervention or dependency.

Scanned with CamScanner

Linear Algebra :-

→ Helps in Optimization.

→ Performs operations on the data.

① Scalars.

A Value which represents something.

All basic arithmetics funtions can be applied to two scalar values (+, -, /, *).. etc.

Eg. 10 + 13 = 23

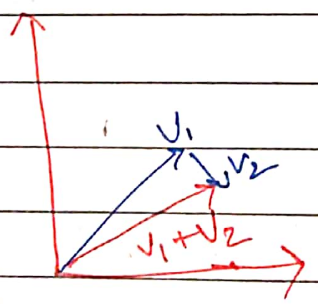
② Vectors :-

(i) Computer Science :- People interpret vectors as a list of numbers that represent something

Vector operation.

↳ To apply vector operation we first should know which type of ML dataset we are working with & then apply vector operation accordingly.

① Addition.

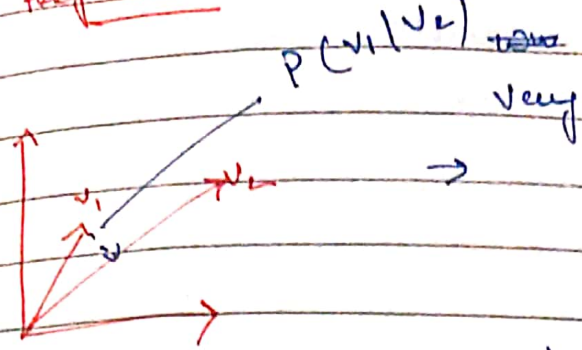


V_1+V_2 = Displacement.

② Scalar Multiplication.

Vector grows with +ve scalar value multiplied & for -ve vice versa.

③ Projection.



Very useful in Deep learning.

→ Suppose I know much about V_2 feature projecting V_1 on V_2 will help me know the unknown features of V_1.

③ Matrix Op

↳ • Composition of number, symbols or expressions in a rectangular array.

• We use Matrices to convert our expressions of the vector form to array.

• We generally do this to make our operations easier & helpful.

Some Basic operation of Matrices such as ~~Addition~~ & Subtraction & Multiplication we already did in +2.

Lets dive but what's imp. w.r.t. M.L.

Matrix Operations :-

① Transpose :- (i) Interchanging of rows & columns.

Date / DELTA Pg No.

- We generally use it to change the dimensionality of the given data / problem.

Eg. $A = \begin{bmatrix} 2 & 2 \\ 4 & 1 \end{bmatrix}$ $A^T \begin{bmatrix} 2 & 4 \\ 2 & 1 \end{bmatrix}$

(5) Determinant of Matrix.

- It is the scalar value of Matrix.
- It gives you the product of Eigen Values of the Matrix.

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

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

~~[illegible]~~

NOTE: Vector as a Matrix L.A.

- Vectors can be easily translate to Matrix.
- We already talked It is easy to apply operations on vector in Data

ML Pov :-

Known operations such as Scaling, Rotation & Shearing, The above operation can be ~~too~~ applied on our data to better extract information from it.

Eg.

Scaling:-

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

Shearing:-

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

Rotation:-

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

* Matrices can also help us solve the equations by

→ Inverse Method
→ Row Echelon Method.

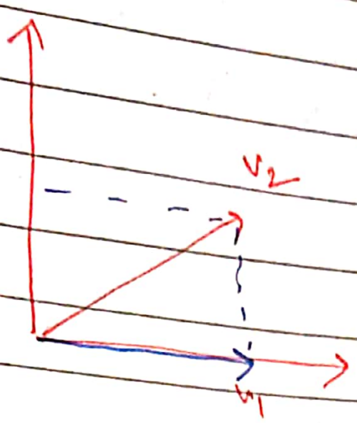
Eigen Vectors (ML POV)

- It doesn't change direction even if transformation is applied to it.
- These are generally used for analysis of ~~text~~ the data.

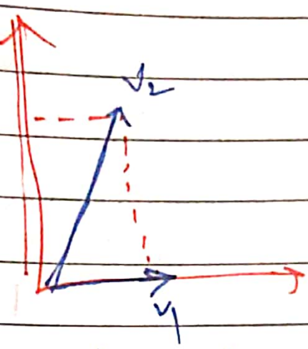
Eg.

$V_1 \rightarrow \quad V_2 \rightarrow$

Before



After Shearing.



My V_2 vector has completely change but my V_1 Vector is been applied with -ve Scalar.

After Shearing we are still able to extract the information that it is rectangle.

Applications of Linear Algebra in ~~ML/AI~~ AI/ML

- PCA for dimensionality reduction.
- Transformation of Pixel data to work comp with Images.
- Encoding of the Dataset.
- SVD
- Optimization for DL models.