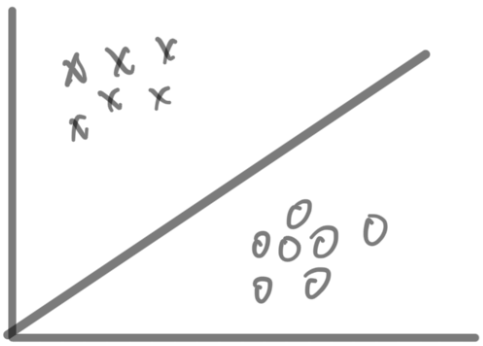
Stanford CS229 Lecture - 4

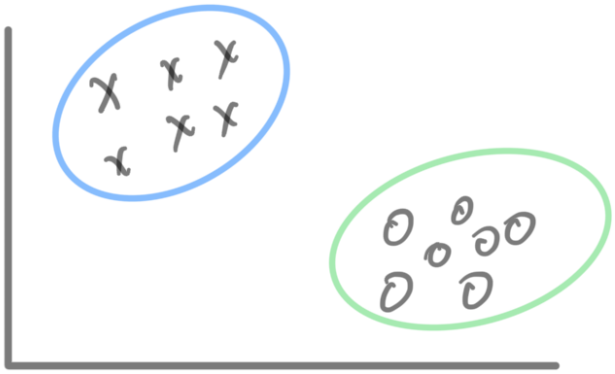
Generative Learning Algorithm (G.L.A) :-

- Discriminant Learning is all about finding the right decision boundary



↑

(Discriminant Learning algorithm)



G.L.A

↳ G.L.A isolates each class & studies it, when a new data point comes it adds the data point based on the comparision.

Formula :-

Learn $P(y|x)$

or

Learn $h_\theta(x) = \begin{cases} 0 \\ 1 \end{cases}$ directly

Formula :-

Learn $P(x|y)$

features, class

→ What are the features like, given its class

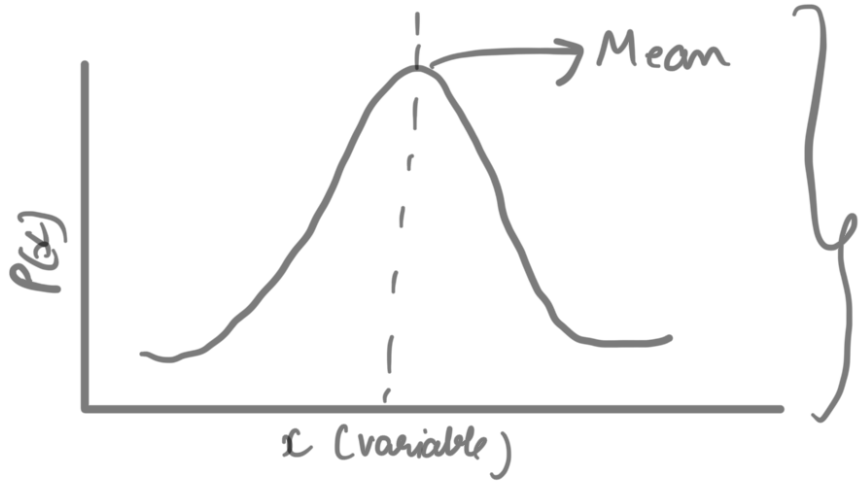
Baye's Rule

$$P(y=1|x) = \frac{P(x|y=1)\,P(y=1)}{P(x)}$$

} Framework for G.L.A

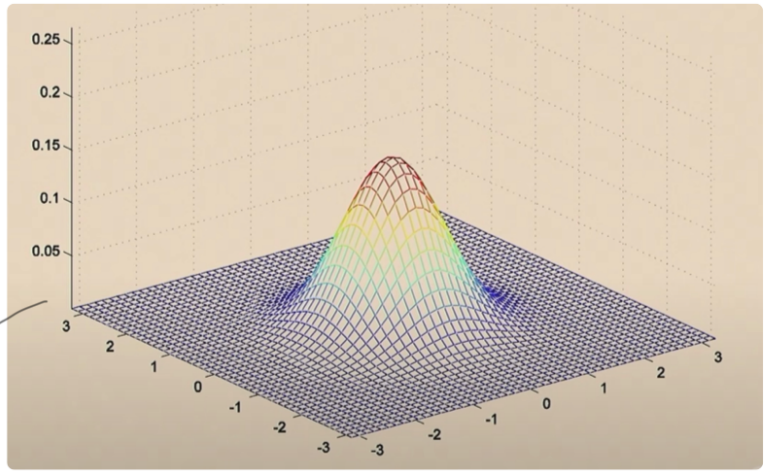
Gaussian Discriminant Analysis :-

Univariate gaussian distribution :-



→ Probability of a value near the mean is higher.

Multivariate Gaussian Distribution :-



Presence of multiple variables

↳ Dependent on μ, Σ

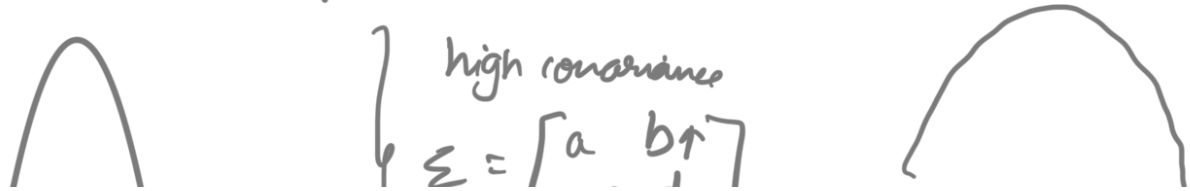
mean vector
(vector of mean of all variables)

↳ co-variance matrix.
(matrix of how variables are related to each other)

→ Probability Density function :- ($f(x)$)

$$f(x) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

→ to change bumps



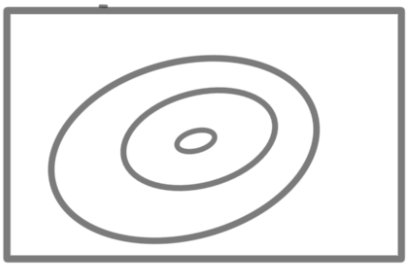
$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$

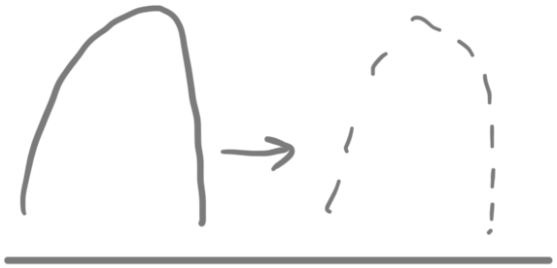
(a, d) = variance

(b, c) = co-variance

→ to move the bump

low co-variance





Change the mean

$\mu = \begin{bmatrix} x \\ y \end{bmatrix}$ x & y are mean of variables

G.D.A > Logistic Regression

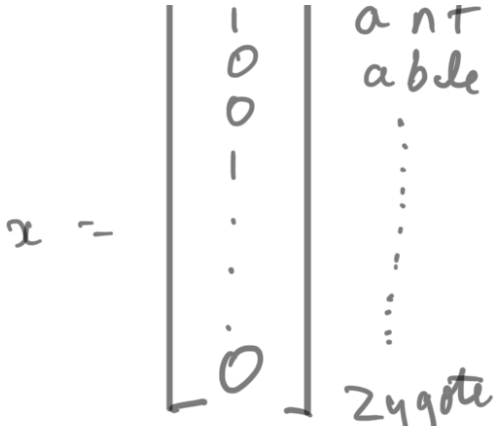
for datasets with normal distribution.

If you do not know much about the features, go for Logistic Regression.

Naive Bayes (G.L.A)

- Naive coz it assumes all features are independent.

Ex:-



↳ dictionary of words

↳ feature vector that shows 1 → spam
0 → not spam

Formula :-

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

→ Prior probability

X → features
C → class

↓ Posterior probability

↳ Marginal Likelihood (Probability of feature X across all possible classes)

$$P(X|C) = \prod_{i=1}^{n} P(x_i|C)$$