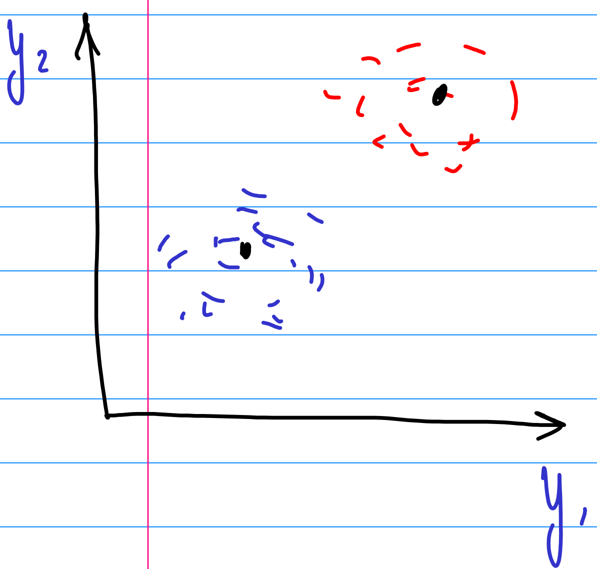
LDA

MANOVA $\longrightarrow$ LDA



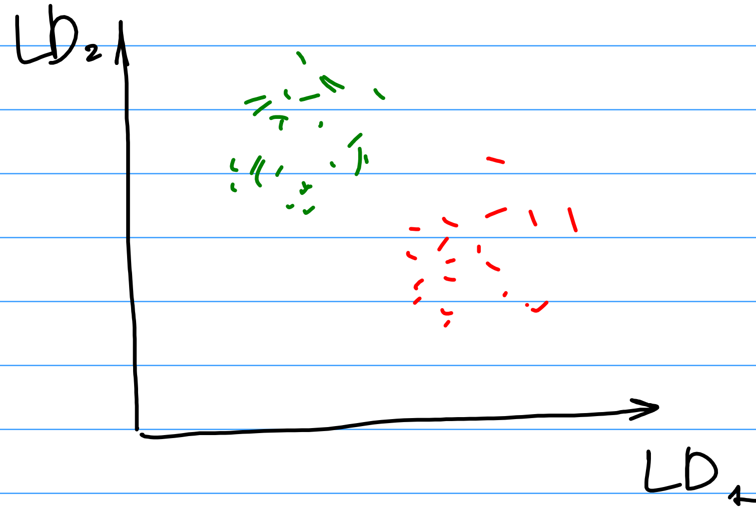
related to MANOVA by

$$\frac{SS\ between}{SS\ within}$$

PCA

$\hat{Var} \longrightarrow max$

LDA



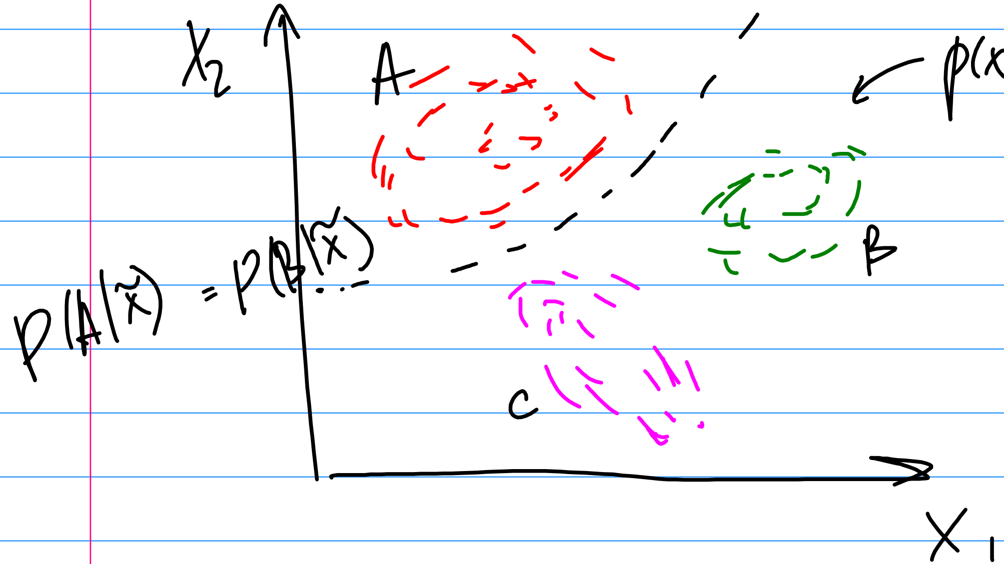
$$\frac{\hat{d}^2}{s_1^2+s_2^2}$$

Logit $\longleftrightarrow$ LDA

via bayesian

LDA

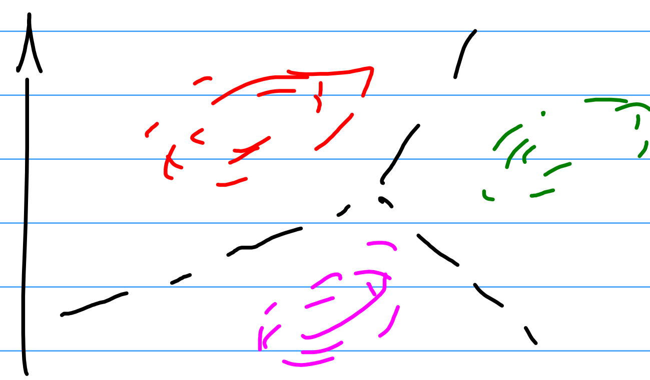
1) Multinomial Normal Distribution



$$p(x) = \frac{\exp\left(\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)}{\sqrt{(2\pi)^n \det \Sigma}}$$

If $\Sigma_A = \Sigma_B = \Sigma_C$

simplifies to linear case:



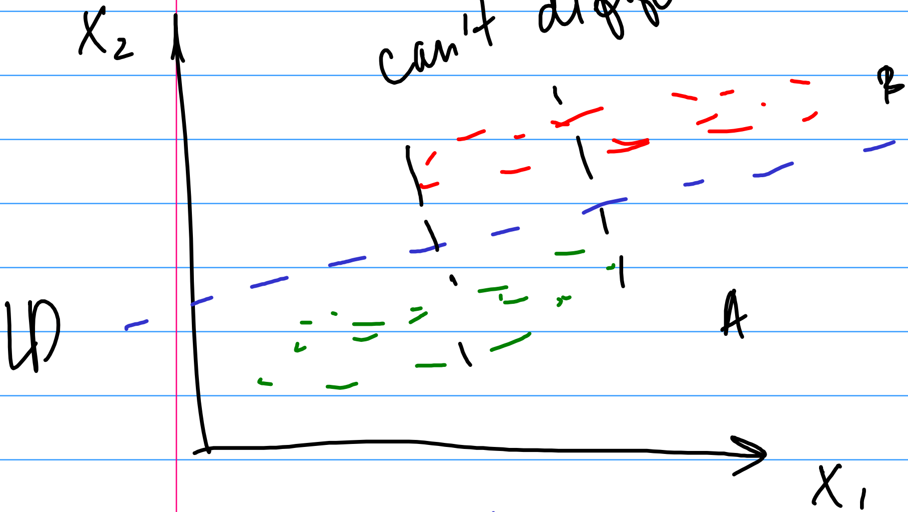
If assumption $\Sigma_A \neq \Sigma_B = \Sigma_B$

QDA (Quadratic DA)

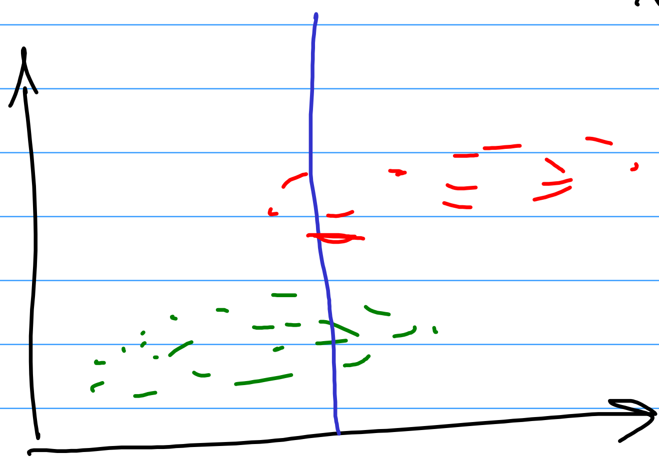
X_2
A
quadratic functions
B
C
X_1

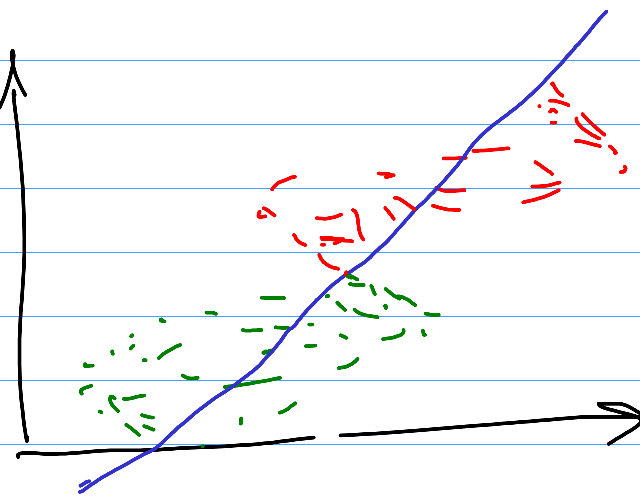
Example :

can't differentiate A from B



$\tilde{\mu}_1 - \tilde{\mu}_2$ not big enough

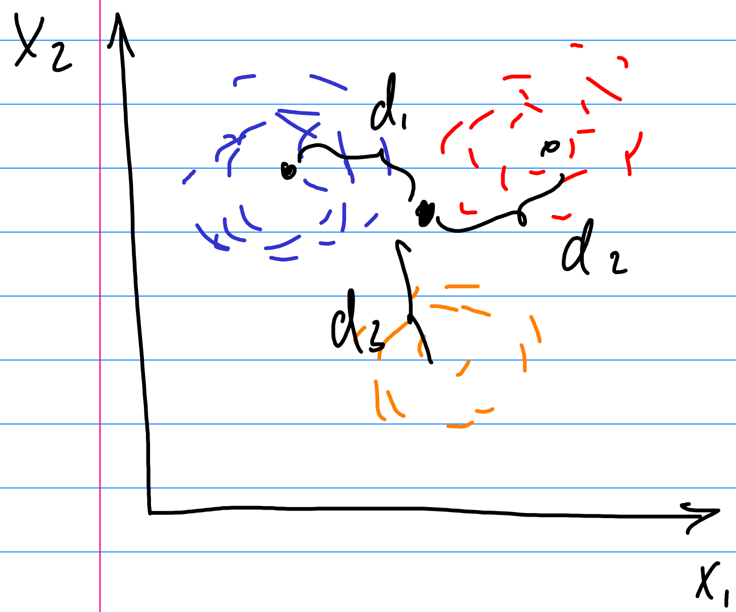




$$\frac{(\mu_1 - \mu_2)^2}{s_1^2 + s_2^2} \to \max$$

$s_1^2 + s_2^2$ is not small enough

Example: 3 categories



$$\frac{d_1^2 + d_2^2 + d_3^3}{s_1^2 + s_2^2 + s_3^3}$$

Assumptions:

1) Normality

2) (LDA) Homoscedasticity

(A) else => QDA

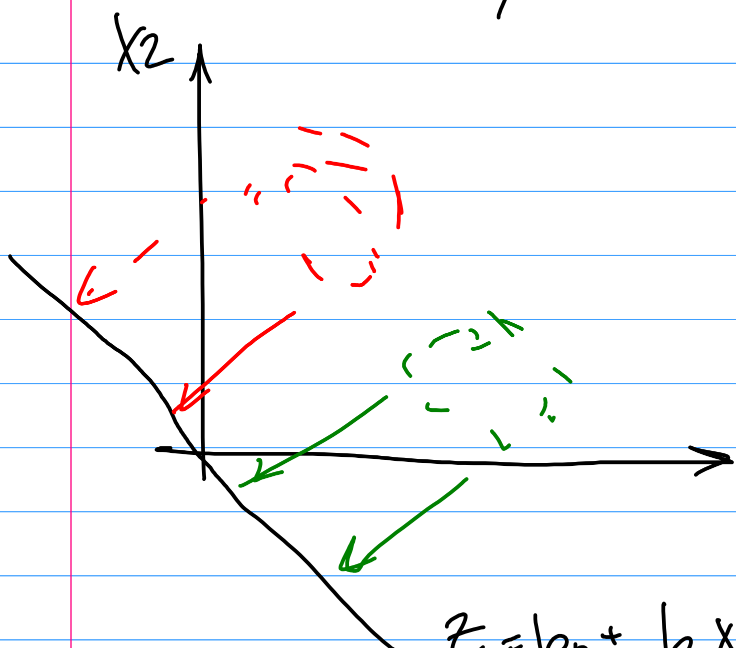
3) No multicollinearity

Modelling Aproach:

$$Z_{ik} = b_{0i} + b_{1i} X_{1k} + \dots + b_{ji} X_{jk}$$

↳ ith DF for object k
(kth object)

$i \in 1; G-1$



$z = b_0 + b_1 X_1 + b_2 X_2$

$$\max_b \gamma = \frac{SS_b}{SS_w} = \frac{\sum_{g=1}^{G} I_g (\bar{Z}_g - \bar{Z})^2}{\sum_{g=1}^{G} \sum_{i=1}^{I_g} (Z_{gi} - \bar{Z}_g)^2}$$

↑ eigenvalue (discriminant criteria

Measure of Importance for kth DF

$$EP = \frac{\gamma_k}{\gamma_1 + \ldots + \gamma_n}$$

Statistical significance of DF:

Wilk's Lambda:

$$\Lambda = \frac{1}{1+\gamma} = \frac{SS_W}{SS_b + SS_W} = \frac{UV}{TV}$$

Canonical correlation coefficient $(\sqrt{C})$

$$C = \frac{\gamma}{1+\gamma} = \frac{SS_b}{SS_b + SS_W} = \frac{EV}{TV}$$

$$\ln(\Lambda) \cdot \left(-\left(N - \frac{J+G}{2} - 1\right)\right) \sim \chi^2_{J \cdot (G-1)}$$