Logit vs LDA

$P(Y=1|X=x) = x^T\beta$ LPM doesn't work

$\Rightarrow$ We need a model that assumes

① $P(Y=1|X=x) \in [0;1]$

② Relation with $x^T\beta$

log-odds ratio

Assume:

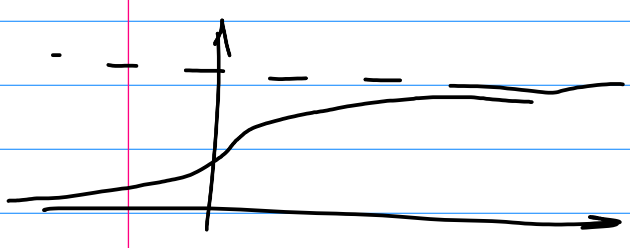
$$\text{logit}(P(Y=1|X=x)) = \log \frac{P(Y=1|X=x)}{P(Y=0|X=x)}$$

$$\boxed{= x^T\beta}$$

Equivalent to introduction of

$f(x) = \frac{e^x}{1+e^x}$ - logistic function

$$P(Y=1|X=x) = \frac{e^{x^T\beta}}{1-e^{x^T\beta}}$$



$P(Y=1|X) = \Lambda(x^T\beta)$

In case of Binary Classes:

$y|X \sim \text{Bernoulli}(p)$ ← Exponential family

$\text{logit}(p) = \text{logit}(E(y|X=x)) = \underbrace{X^T\beta}_{\eta \text{ linear predictor}}$

Special case of GLM

$g(E(y|X=x)) = \eta$ – link function

$P(y_i = 1) = p^y (1-p)^{1-y}, \quad y \in \{0;1\}$

$L(p) = \prod P(y_i = 1) = \prod p^{y_i} (1-p)^{1-y_i}$

$\ell(p) = \sum (y_i \cdot \log p + (1-y_i) \cdot \log(1-p)) =$

$= \sum y_i \cdot \log p(x_i^T\beta) + \sum (1-y_i) \cdot \log(1 - p(x_i^T\beta)) =$

$= \sum y_i (x_i^T\beta - \log(1 + x^T\beta)) +$

$+ \sum (1-y_i) \cdot \log(1 + e^{x^T\beta}) \longrightarrow \max_\beta$

In case of multiple classes k

1) One vs All : k models

2) One vs One : $\frac{k(k-1)}{2}$ pairs of model

+ majority of votes

In case of multinomial logit:

$Y|X \sim \text{Categorical}(p_1, \dots, p_k)$, $\sum p_i = 1$

$$p(Y = i \mid X = x) = \frac{e^{x^T \beta_i}}{\sum_{i=1}^{k} e^{x^T \beta_i}}$$

LDA

easy to model

Likelihood

Prior

$$\underbrace{P(Y = i \mid X = x)}_{\text{Posterior}} = \frac{P(X = x \mid Y = i)\, P(Y = i)}{\underbrace{P(X = x)}_{\text{Data}}}$$

hard to model

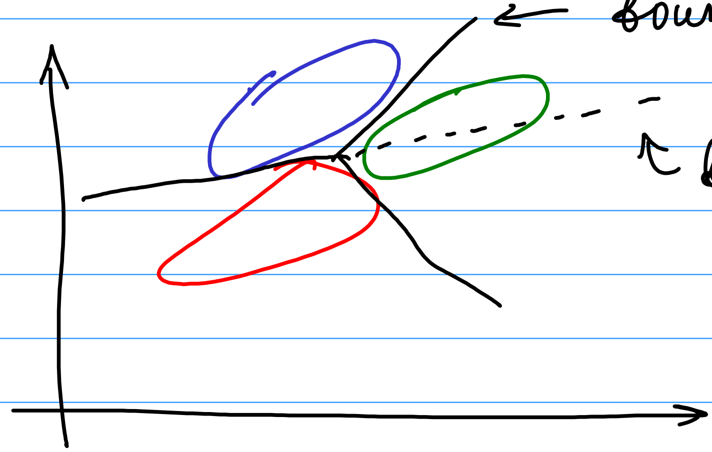
LDA: main difference X - normally distributed + $\Sigma_i = \Sigma$, $\forall\, i$, $\overline{i, k}$

$$P(y=i \mid X=x) = \frac{P(X=x \mid y=i)\, P(y=i)}{P(X=x)} =$$

$$= \frac{P(X=x \mid y=i)\, P(y=i)}{\sum_{j=1}^{k} P(X=x \mid y=j)\, P(y=j)} = \frac{f_i(x) \cdot \pi_i}{\sum f_j(x) \cdot \pi_j}$$

LDA: $\Sigma_j = \Sigma$

QDA: Σ_j can be distinct



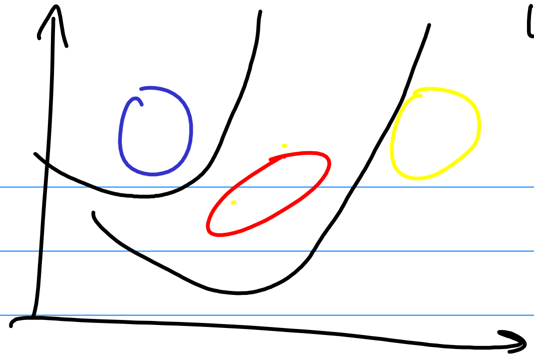
boundary for all 3 classes

boundary for each pair of classes

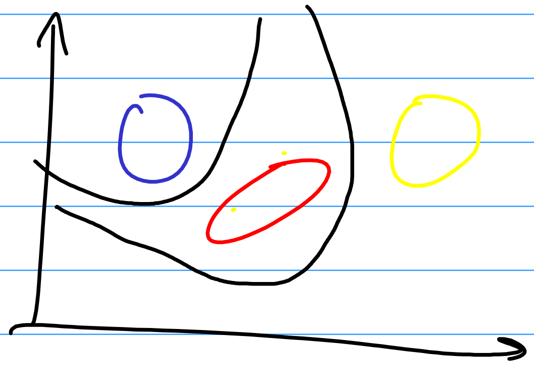
$$\log \frac{P(y=l \mid X=x)}{P(y=m \mid X=x)} = \log \frac{f_l(x)}{f_m(x)} + \log \frac{\pi_l}{\pi_m} =$$

$$\underbrace{\log \frac{\pi_l}{\pi_m} - \frac{1}{2}(\mu_l + \mu_m)^T \Sigma^{-1} (\mu_l - \mu_m)}_{\beta_0}$$

$$+\, x^T \underbrace{\Sigma^{-1} (\mu_l - \mu_m)}_{\beta} = \beta_0 + x^T \beta$$



LDA $x_1, x_2, x_1x_2, x_1^2, x_2^2$



QDA