Lecture 10: LDA

$\boxed{p=1}$

$$\delta_c(x) = \underbrace{\mathbb{P}(X=x \mid Y=c)}_{N(\mu_c,\sigma^2)}\underbrace{\mathbb{P}(Y=c)}_{\pi_c}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu_c)^2\right)\pi_c$$

only care up to increasing transf of δ_c

log

$$\equiv \underbrace{-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2)}_{\text{don't depend on } c} - \frac{1}{2\sigma^2}(x-\mu_c)^2 + \log(\pi_c)$$

$$\equiv -\frac{1}{2\sigma^2}(x^2 - 2\mu_c x + \mu_c^2) + \log(\pi_c)$$

$$\equiv \cancel{\frac{-x^2}{2\sigma^2}} + \frac{\mu_c x}{\sigma^2} - \frac{\mu_c^2}{2\sigma^2} + \log\pi_c$$

$$\delta_c(x) \equiv \underbrace{\frac{-\mu_c^2}{2\sigma^2} + \log\pi_c}_{\beta_{0c}} + \underbrace{\frac{\mu_c}{\sigma^2}}_{\beta_c} x$$

$$= \beta_{0c} + \beta_c x$$

LDA is a linear classifier

So to fit LDA, need to estimate

μ_c, σ^2, π_c

Way I do this!

MLEs

$\hat{\mu}_c$ = mean of training x_ns in class c

$\hat{\pi}_c$ = pct. of training in class c

$\hat{\sigma}^2$ = pooled variance of x_ns

pooled variance = calc. var in each class and take a weighted avg.

X_ns in class 1, Y_ns in class 2

$$\hat{\sigma}^2 = \frac{\sum_n (X_n - \bar{X})^2 + \sum_n (Y_n - \bar{Y})^2}{N_x + N_y - 2}$$

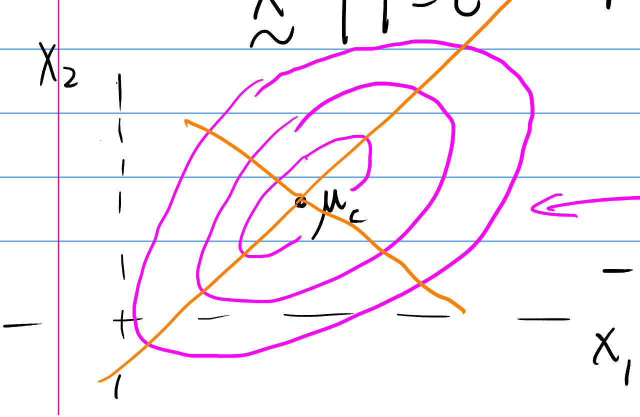
What about P > 1?

shared Cov; mtx p×p

$\Sigma_{ij} = \mathrm{Cov}(X_i, X_j)$

$$\underset{\sim}{X} \mid Y=c \sim N(\mu_c, \Sigma)$$

mean vector $\mu_c \in \mathbb{R}^p$



ellipses

- dir of axes are e-vects of Σ
- stretching of ellipses is det. by e-values

- Σ is symmetric : $\Sigma_{ij} = \Sigma_{ji}$
- Σ is positive-definite ($\sigma^2 > 0$)

→ for symmetric = e-vals > 0

→ for all mtx

$$x^T \Sigma x > 0 \quad \text{for } x \neq 0.$$

Density of MV Normal: mean μ and cov. Σ

$$f(\underset{\sim}{x}) = (2\pi)^{-p/2} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}(\underset{\sim}{x}-\mu)^T \Sigma^{-1} (\underset{\sim}{x}-\mu)\right)$$

Univariate:

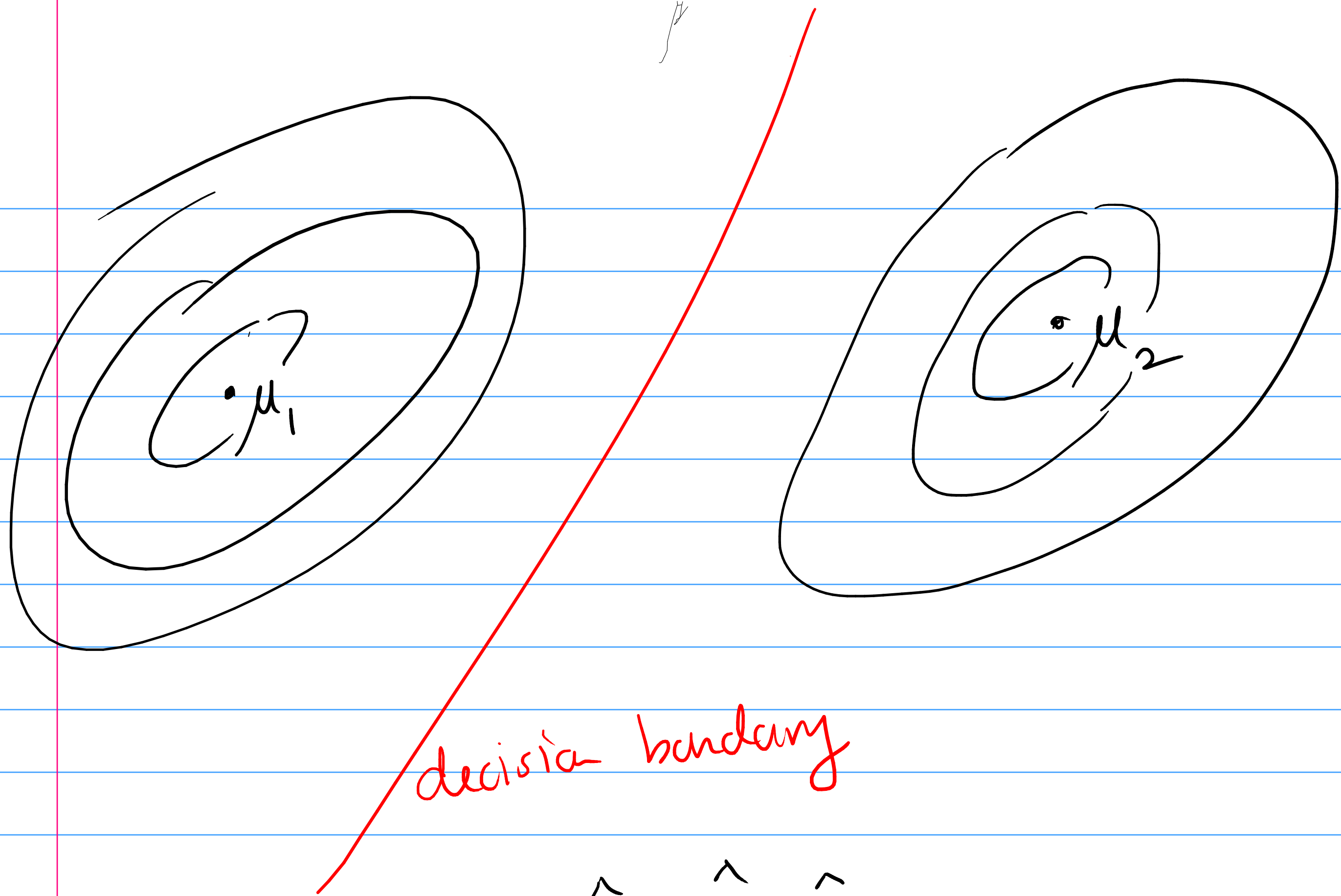
$$f(x) = (2\pi)^{-1/2} (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)\right)$$

$$\delta_c(\underset{\sim}{x}) = (2\pi)^{-p/2} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}(\underset{\sim}{x}-\mu_c)^T \Sigma^{-1} (\underset{\sim}{x}-\mu_c)\right) \pi_c$$

$$= \quad \vdots$$

$$\equiv \underbrace{\log \pi_c - \frac{1}{2}\mu_c^T \Sigma^{-1} \mu_c}_{\beta_{0c} \in \mathbb{R}} + \underbrace{\mu_c^T \Sigma^{-1}}_{\beta_c^T \in \mathbb{R}^p} \underset{\sim}{x}$$

$$\equiv \beta_{0c} + \beta_c^T \underset{\sim}{x}$$ ← also linear



Need to estimate $\hat{\mu}_c$, $\hat{\Sigma}$, $\hat{\pi}_c$

MLES
- $\hat{\mu}_c$ = mean vector of x_ns in class c
- $\hat{\pi}_c$ = pct. of training in class c
- $\hat{\Sigma}$ = pooled cov. mtx. (pool ests component-wise)

QDA: Quadratic Discrim. Analysis

LDA assumes that we have shared variance σ^2 or Σ

QDA relaxes, allows for diff. var. across classes σ_c^2 or Σ_c

Still assume

$$\delta_c(\underset{\sim}{x}) = \underbrace{\mathbb{P}(\underset{\sim}{X} = \underset{\sim}{x} \mid Y = c)}_{N(\mu_c, \Sigma_c)} \underbrace{\mathbb{P}(Y = c)}_{\pi_c}$$

Similar algebra:

$\boxed{P = 1}$ $$\delta_c(x) = -\log \sigma_c - \frac{\overbrace{(x - \mu_c)^2}^{\text{quadratic}}}{2\sigma_c^2} + \log \pi_c$$

$\boxed{P > 1}$ $$\delta_c(\underset{\sim}{x}) = -\log\det \Sigma_c - \frac{1}{2}\underbrace{(\underset{\sim}{x} - \mu_c)^T \Sigma_c^{-1} (\underset{\sim}{x} - \mu_c)}_{\text{quadratic}} + \log \pi_c$$

Count # of params: K classes, P vars,

LDA: $(K-1)(P+1) \approx KP$

QDA: $(K-1)\left(\frac{P(P+3)}{2} + 1\right) \approx KP^2$