Note Title 19-11-2020

COL774

Minor Solutions

Semester I, 2020-21

Q-2. GDA with $2r$ classes.

Given:- $x^{(i)} \mid y^{(i)} = k \sim N(\mu_k, \Sigma_k)$

where $\Sigma_k = \Sigma_1$ if $k \leq r$

& $\Sigma_k = \Sigma_{r+1}$ if $k > r$

$y^{(i)} \sim \text{Multinoulli}(\Phi)$ $\Phi = (\phi_1 \dots \phi_{2r})$

$\sum_{k=1}^{2r} \phi_k = 1$

We would like to find the decision boundary for class 1.

Note that
$$p(y=k \mid x) = \frac{p(x \mid y=k)\, p(y=k)}{p(y)}.$$

For any class k, the decision boundary is boundary enclosing those points where $p(y=k \mid x) \geq p(y=k' \mid x)$ $\forall k' \in \{1 \ldots 2r\}$ [By definition]

[these are the points which are classified with label k]

(for class k)

Therefore, at the decision boundary it will be the case that

$$P(y=k \mid x) = P(y=k' \mid x) \quad [\text{for some } k \neq k']$$

① [& Also, $P(y=k \mid x) \geq P(y=k' \mid x) \; \forall k'$] will still hold

$\Rightarrow$ form of decision boundary will be :- Assume some k' satisfying ①

$$\frac{P(x \mid y=k)\, P(y=k)}{\cancel{P(x)}} = \frac{P(x \mid y=k')\, P(y=k')}{\cancel{P(x)}}$$

$$\Rightarrow \quad \underbrace{P(x \mid y=k)}_{\Downarrow \text{ Normal}(\mu_k, \Sigma_k)} \; \underbrace{P(y=k)}_{\Downarrow \; \phi_k} = \underbrace{P(x \mid y=k')}_{\Downarrow \text{ Normal}(\mu_{k'}, \Sigma_{k'})} \; \underbrace{P(y=k')}_{\Downarrow \; \phi_{k'}}$$

$$\Rightarrow \left[\frac{1}{(2\pi)^{n/2} |\Sigma_k|^{1/2}} \exp - \frac{(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)}{2} \right] \phi_k$$

$$= \left[\frac{1}{(2\pi)^{n/2} |\Sigma_{k'}|^{1/2}} \exp - \left[\frac{(x-\mu_{k'})^T \Sigma_{k'}^{-1} (x-\mu_{k'})}{2} \right] \right] \phi_{k'}$$

Taking log on both sides:-

$$\log \frac{1}{\cancel{(2\pi)^{n/2}}\,|\hat{\Sigma}_k|^{1/2}} - \frac{(x-\mu_k)^T \hat{\Sigma}_k^{-1} (x-\mu_k)}{2} + \log \phi_k$$

$$= \log \frac{1}{\cancel{(2\pi)^{n/2}}\,|\hat{\Sigma}_{k'}|^{1/2}} - (x-\mu_{k'})^T \hat{\Sigma}_{k'}^{-1} (x-\mu_{k'}) + \log \phi_{k'}$$

$$\Rightarrow 2\left[\log \frac{|\hat{\Sigma}_{k'}|^{1/2}}{|\hat{\Sigma}_k|^{1/2}} + \log\left[\frac{\phi_k}{\phi_{k'}}\right]\right] = \begin{array}{l}(x-\mu_k)^T \hat{\Sigma}_k^{-1} (x-\mu_k) \\ -(x-\mu_{k'})^T \hat{\Sigma}_{k'} (x-\mu_{k'})\end{array}$$

Now, the exact expression of this equation depends on two cases:-

(i) $K' \leq \gamma$ [$K = 1$ as given]

then $\hat{\Sigma}_k = \hat{\Sigma}_{k'} = \hat{\Sigma}_1$

$$\Rightarrow 2\left[\log \frac{\cancel{|\hat{\Sigma}_{k'}|^{1/2}}}{\cancel{|\hat{\Sigma}_k|^{1/2}}} + \log \frac{\phi_k}{\phi_{k'}}\right] = \cancel{x^T \hat{\Sigma}_k^{-1} x} - x^T \hat{\Sigma}_k^{-1} \mu_k - \mu_k^T \hat{\Sigma}_k^{-1} x + \mu_k^T \hat{\Sigma}_k^{-1} \mu_k - \left(\cancel{x^T \hat{\Sigma}_{k'}^{-1} x} - x^T \hat{\Sigma}_{k'}^{-1} \mu_{k'} - \mu_{k'}^T \hat{\Sigma}_{k'}^{-1} x - \mu_{k'}^T \hat{\Sigma}_{k'}^{-1} \mu_{k'}\right)$$

$\because \cancel{\hat{\Sigma}_k = \hat{\Sigma}_{k'}}$

— 1

Substituting $\hat{\Sigma}_k = \hat{\Sigma}_{k'} = \hat{\Sigma}_1 \Downarrow$

$$\Rightarrow 2\left(\log\left(\frac{\phi_k}{\phi_{k'}}\right)\right) = -x^T \hat{\Sigma}_1^{-1} \mu_k - \mu_k^T \hat{\Sigma}_1^{-1} x + \mu_k^T \hat{\Sigma}_1^{-1} \mu_k - \left[-x^T \hat{\Sigma}_1^{-1} \mu_{k'} - \mu_{k'}^T \hat{\Sigma}_1^{-1} x - \mu_{k'}^T \hat{\Sigma}_{k'}^{-1} \mu_{k'}\right]$$

Linear in x — (a)

On the other hand, if $k' > r$

then $\hat{\Sigma}_1 = \hat{\Sigma}_k \neq \hat{\Sigma}_{k'} = \hat{\Sigma}_{r+1}$

$\Rightarrow$ Quadratic terms in (1) don't cancel with each other $\Rightarrow$ we get a quadratic expression in x. — (b)

The overall boundary will be determined wrt value of k' ($k' \leq r$ or $k' > r$)

It will be a combination of piece-wise linear & quadratic. ~~so~~ The piece will change every time ~~the~~ value of 'k' changes.

Illustration:
(in 2-D)