Machine Learning and Computational Statistics

Theodoros Georgiopoulos

Homework 7

Exercises for Unit 7: Bayes Classifier

Solution for exercise 1

We consider a two-class 1-dim classification problem of two equiprobable classes ω_1 and ω_2 that are modeled by the normal distributions $\mathcal{N}(0,1)$ and $\mathcal{N}(0,5)$, respectively. So, $p(x|\omega_1) \sim \mathcal{N}(0,1)$ and $p(x|\omega_2) \sim \mathcal{N}(0,5)$. To determine the decision regions R_1 and R_2 we can use the Bayes rule.

$$\begin{aligned}
p(\omega_1|x) &> p(\omega_2|x) \iff \\
\frac{p(x|\omega_1)p(\omega_1)}{p(x)} &> \frac{p(x|\omega_2)p(\omega_2)}{p(x)} \iff \\
p(x|\omega_1) &> p(x|\omega_2) \iff \\
\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-0)^2}{2}} &> \frac{1}{\sqrt{2\pi 5}}e^{-\frac{(x-0)^2}{10}} \iff \\
-\frac{x^2}{2} + \frac{1}{2}\ln 5 &> -\frac{x^2}{10} \iff \\
-4x^2 &> -5\ln 5 \iff \\
x^2 &< \frac{5\ln 5}{4}
\end{aligned}$$

So, the decision regions are $R_1 \left\{-\sqrt{\frac{5\ln 5}{4}} < x < \sqrt{\frac{5\ln 5}{4}}\right\}$ and $R_2 \left\{x < -\sqrt{\frac{5\ln 5}{4}} \text{ and } x > \sqrt{\frac{5\ln 5}{4}}\right\}$.

Solution for exercise 2

We have $p(\boldsymbol{x}|\omega_1) \sim \mathcal{N}(\boldsymbol{\mu_1}, \boldsymbol{\Sigma})$ and $p(\boldsymbol{x}|\omega_2) \sim \mathcal{N}(\boldsymbol{\mu_2}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$ and the priors $p(\omega_1) = p(\omega_2)$.

(a) We can use the Bayes classifier as follows

$$\begin{aligned}
p(\omega_1|\boldsymbol{x}) &= p(\omega_2|\boldsymbol{x}) \iff \\
\frac{p(\boldsymbol{x}|\omega_1)p(\omega_1)}{p(\boldsymbol{x})} &= \frac{p(\boldsymbol{x}|\omega_2)p(\omega_2)}{p(\boldsymbol{x})} \iff \\
p(\boldsymbol{x}|\omega_1) &= p(\boldsymbol{x}|\omega_2) \iff \\
\frac{1}{\sqrt{(2\pi)^2|\boldsymbol{\Sigma}|}}e^{-\frac{(\boldsymbol{x}-\boldsymbol{\mu_1})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu_1})}{2}} &= \frac{1}{\sqrt{(2\pi)^2|\boldsymbol{\Sigma}|}}e^{-\frac{(\boldsymbol{x}-\boldsymbol{\mu_2})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu_2})}{2}} \iff \\
(\boldsymbol{x}-\boldsymbol{\mu_1})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu_1}) &= (\boldsymbol{x}-\boldsymbol{\mu_2})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu_2}) \iff \\
\frac{1}{2\sigma^2}||\boldsymbol{x}-\boldsymbol{\mu_1}||^2 &= \frac{1}{2\sigma^2}||\boldsymbol{x}-\boldsymbol{\mu_2}||^2 \iff \\
||\boldsymbol{x}-\boldsymbol{\mu_1}||^2 &= ||\boldsymbol{x}-\boldsymbol{\mu_2}||^2
\end{aligned}$$

(b) From (a) we have that

$$(\boldsymbol{x} - \boldsymbol{\mu_1})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x - \mu_1}) = (\boldsymbol{x} - \boldsymbol{\mu_2})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x - \mu_2}).$$

If $\Sigma \neq \sigma^2 I$,

$$\boldsymbol{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} - \boldsymbol{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu_1} - \boldsymbol{\mu_1}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} + \boldsymbol{\mu_1}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu_1} = \boldsymbol{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} - \boldsymbol{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu_2} - \boldsymbol{\mu_2}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} + \boldsymbol{\mu_2}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu_2}$$
$$2\boldsymbol{x}^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu_1} - \boldsymbol{\mu_2}) + \boldsymbol{\mu_1}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu_1} - \boldsymbol{\mu_2}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu_2} = 0$$

We can see that the decision boundary passes $\frac{1}{2}(\boldsymbol{\mu_1} - \boldsymbol{\mu_2})$.