Topic 3: Quadratic discriminant analysis

Case of different covariance matrices

Case of different covariance matrices ($\Sigma_1 \neq \Sigma_2$)

Theorem 5.2. *Now we assume that the two populations π_1 and π_2 are $N_p(\boldsymbol{\mu}_1, \Sigma_1)$ and $N_p(\boldsymbol{\mu}_2, \Sigma_2)$, respectively. Repeating the same steps as in Theorem 5.1 we get*

$$R_1 = \left\{ \boldsymbol{x} : -\frac{1}{2}\boldsymbol{x}^\top (\Sigma_1^{-1} - \Sigma_2^{-1})\boldsymbol{x} + (\boldsymbol{\mu}_1^\top \Sigma_1^{-1} - \boldsymbol{\mu}_2^\top \Sigma_2^{-1})\boldsymbol{x} - k \geq \log \left[\frac{c(1|2)}{c(2|1)} \times \frac{p_2}{p_1} \right] \right\}$$

$$R_2 = \left\{ \boldsymbol{x} : -\frac{1}{2}\boldsymbol{x}^\top (\Sigma_1^{-1} - \Sigma_2^{-1})\boldsymbol{x} + (\boldsymbol{\mu}_1^\top \Sigma_1^{-1} - \boldsymbol{\mu}_2^\top \Sigma_2^{-1})\boldsymbol{x} - k < \log \left[\frac{c(1|2)}{c(2|1)} \times \frac{p_2}{p_1} \right] \right\}$$

where $k = \frac{1}{2}\log(\frac{|\Sigma_1|}{|\Sigma_2|}) + \frac{1}{2}(\boldsymbol{\mu}_1^\top \Sigma_1^{-1} \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2^\top \Sigma_2^{-1} \boldsymbol{\mu}_2)$ and we see that the classification regions are ***quadratic*** *functions of the new observation in this case. One obtains the following rule:*

1. *Allocate $\boldsymbol{x}_0$ to π_1 if*
   $\frac{1}{2}\boldsymbol{x}_0^\top (\boldsymbol{S}_1^{-1} - \boldsymbol{S}_2^{-1})\boldsymbol{x}_0 + (\bar{\boldsymbol{x}}_1^\top \boldsymbol{S}_1^{-1} - \bar{\boldsymbol{x}}_2^\top \boldsymbol{S}_2^{-1})\boldsymbol{x}_0 - \hat{k} \geq \log \left[\frac{c(1|2)}{c(2|1)} \times \frac{p_2}{p_1} \right]$
   where $\hat{k}$ is the empirical analog of k.
2. *Allocate $\boldsymbol{x}_0$ to π_2 otherwise.*

When $\Sigma_1 = \Sigma_2$, the quadratic term disappears and we can easily see that the classification regions from Theorem 5.1 are obtained. Of course, the case considered in Theorem 5.2 is more general but we should be cautious when applying it in practice. It turns out that in more than two dimensions, classification rules based on quadratic functions do not always perform nicely and can lead to strange results. This is especially true when the data are not quite normal and when the differences in the covariance matrices are significant. The rule is very sensitive (non-robust) towards departures from normality. Therefore, it is advisable to try to first transform the data to more nearly normal by using some classical normality transformations. Also, tests discussed in Topic 1 of Week 4 can be used to check if equal variance assumption is valid.

Demonstration: Discriminant analysis

Please watch the following demonstration before proceeding to the task.

Transcript

This demonstration can be completed using the provided RStudio or your own RStudio.

To complete this task select the 'Discrim_Examples.demo.Rmd' in the 'Files' section of RStudio. Follow the demonstration contained within the RMD file.

If you choose to complete the example in your own RStudio, upload the following file:

Discrim_Examples.demo.Rmd

The output of the RMD file is also displayed below:

Challenge: Discriminant analysis

If you choose to complete this task in your own RStudio, upload the following file:

Discrim_Examples.challenge.Rmd

Click on the 'Discrim_Examples.challenge.Rmd' in the 'Files' section to begin. Enter your response to the tasks in the 'Enter your code here' section.

This activity and the solution will be discussed at the Collaborate session this week. In the meantime, share and discuss your results in the 'Tutorials' discussion forum.

The solution will also be available here on Friday of this week by clicking on the 'Solution' tab in the top right corner.

The output of the RMD file is also displayed below: