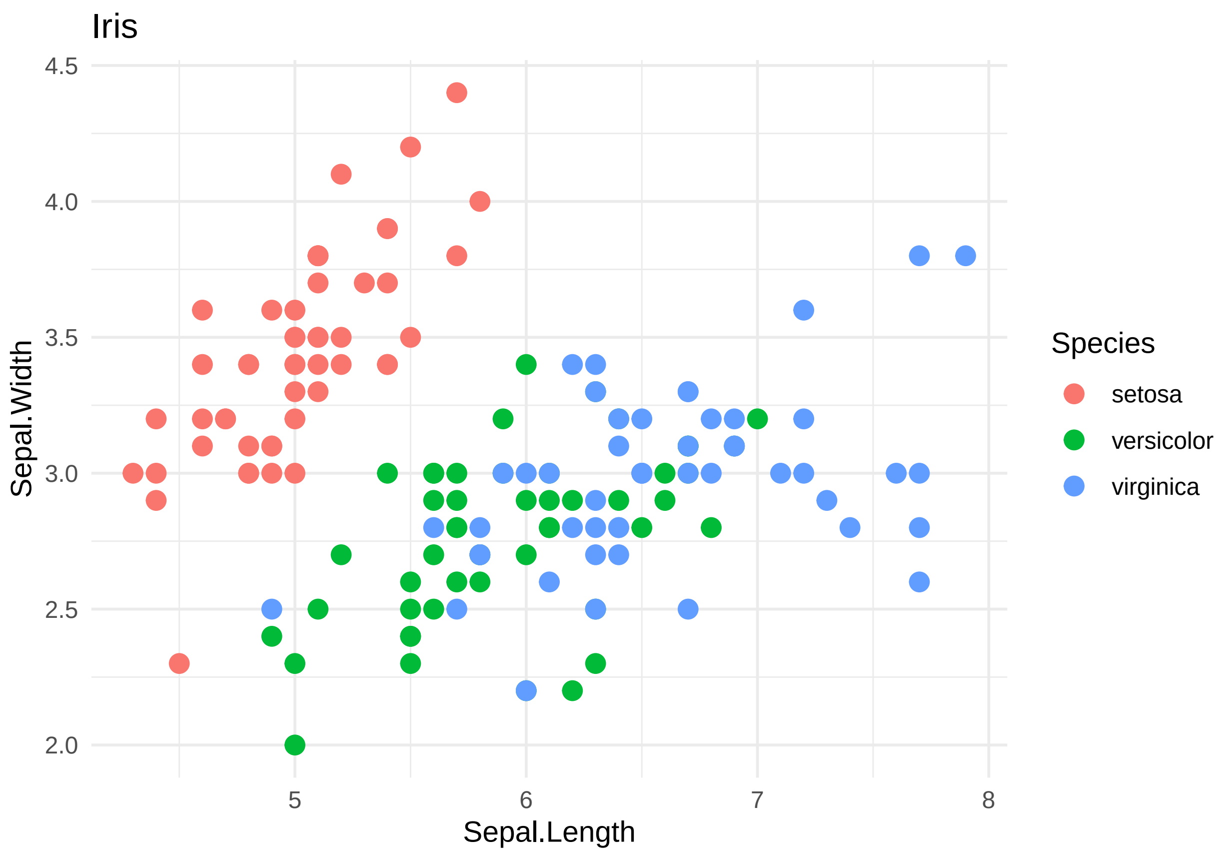
Assignment 1

Martynas Lukosevicius

19/11/2020

Assignment 1

1.



It is not easy to classify by LDA because as we can see from scatter plot versicolor overlay virginica. I expect that misclasification rate will be high because part of the virginica data will be predicted as versicolor and vice versa

2.

a)

Setosa: mean - $\begin{bmatrix} 5.006 \\ 3.428 \end{bmatrix}$, covariance - $\begin{bmatrix} 0.124 & 0.099 \\ 0.099 & 0.144 \end{bmatrix}$, $\pi_{setosa} - 0.3333333$

Virginica: mean - $\begin{bmatrix} 6.588 \\ 2.974 \end{bmatrix}$, covariance - $\begin{bmatrix} 0.404 & 0.094 \\ 0.094 & 0.104 \end{bmatrix}$, $\pi_{virginica} - 0.3333333$

Versicolor: mean - $\begin{bmatrix} 5.936 \\ 2.77 \end{bmatrix}$, covariance - $\begin{bmatrix} 0.266 & 0.085 \\ 0.085 & 0.098 \end{bmatrix}$, $\pi_{versicolor} - 0.3333333$

b)

Pooled covariance - $\begin{bmatrix} 0.219 & 0.09 \\ 0.09 & 0.114 \end{bmatrix}$

c)

Probabilistic model for LDA:

$$P(y = C_i|X, w) \propto P(X|Y = C_i, w)P(Y = C_i|w)$$
$$P(X|Y = C_i, w) \sim N(\mu_i, \Sigma)$$
$$P(Y = C_i|w) = \pi_i$$

$$P(y = C_i|X, w) \propto exp[(\Sigma^{-1}\mu_i)^T X - \frac{1}{2}\mu_i^T \Sigma^{-1}\mu_i + log(\pi_i)] = exp[w_i X + w_{0i}]$$

Where $w_i - (\Sigma^{-1}\mu_i)^T$ and $w_{0i} - -\frac{1}{2}\mu_i^T \Sigma^{-1}\mu_i + log(\pi_i)$

d)

discriminant function $\delta_k(x) = x^T \Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^T + log(\pi_k)$

```
discrim <- function(x,a){
  constant <- (-1/2) * t(a$mean) %*% solve(pcov) %*% a$mean + log(a$prior)
  nonconstant <- t(x) %*% solve(pcov) %*% a$mean
  return(nonconstant+constant)
}
```

e)

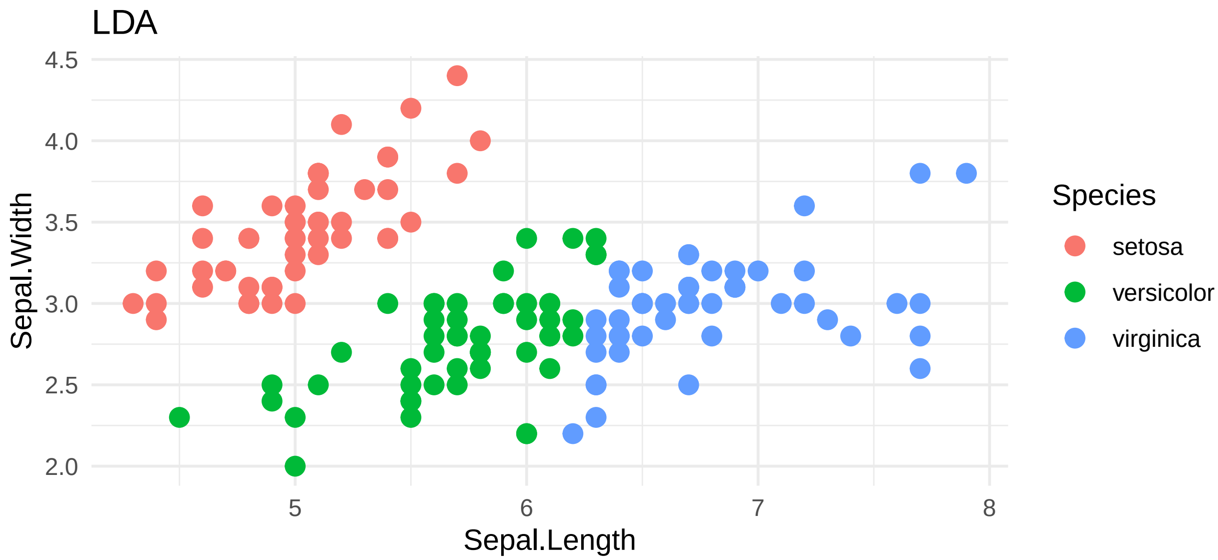
$$(w_i - w_k)x + (w_{0i} - w_{0k}) = 0$$

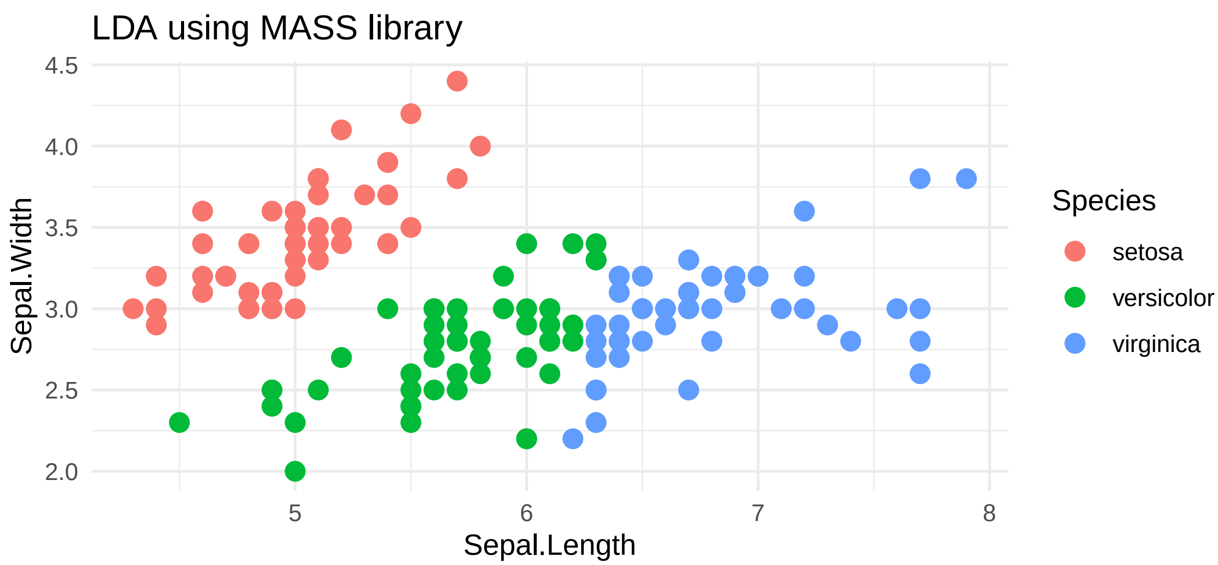
decision boundaries:

- Setosa - Versicolor: $\begin{pmatrix} -7.6573989 \\ 11.8556978 \end{pmatrix} x + (5.1528216) = 0$
- Virginica - Versicolor: $\begin{pmatrix} 2.5620509 \\ -0.2908121 \end{pmatrix} x + (-15.2083507) = 0$
- Setosa - Versicolor: $\begin{pmatrix} -10.2194498 \\ 12.1465099 \end{pmatrix} x + (20.3611723) = 0$

LDA assume that $\Sigma_i = \Sigma$, However it is not the case in this situation

3.



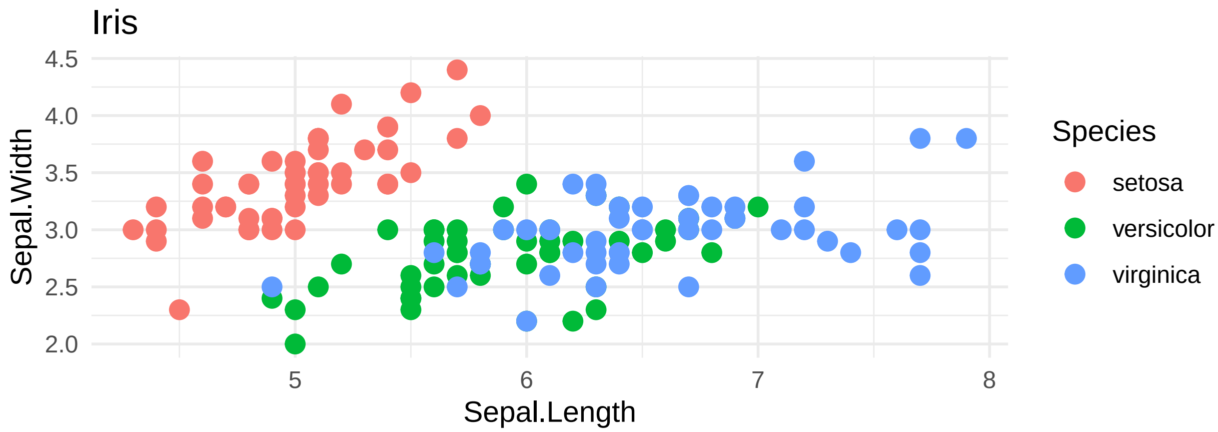


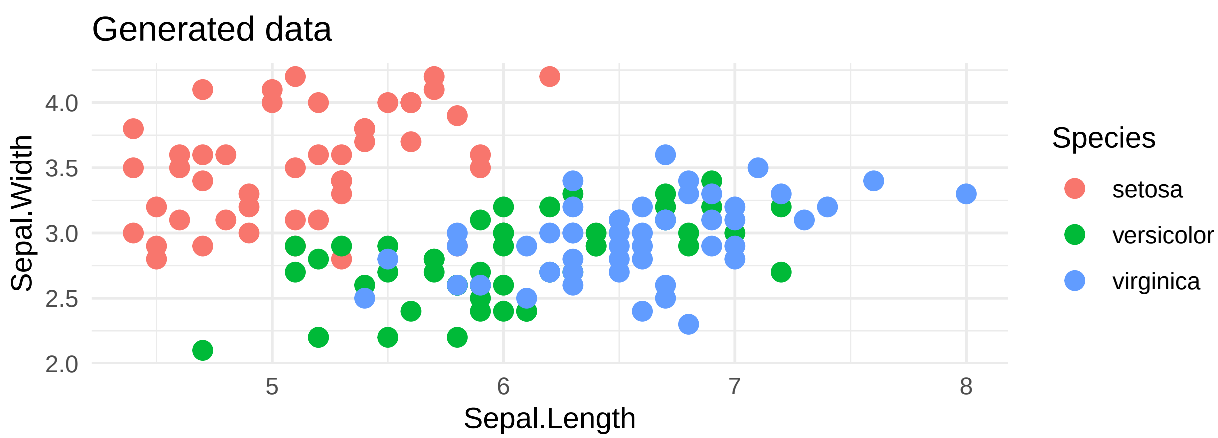
Missclasification rate of LDA: 0.2

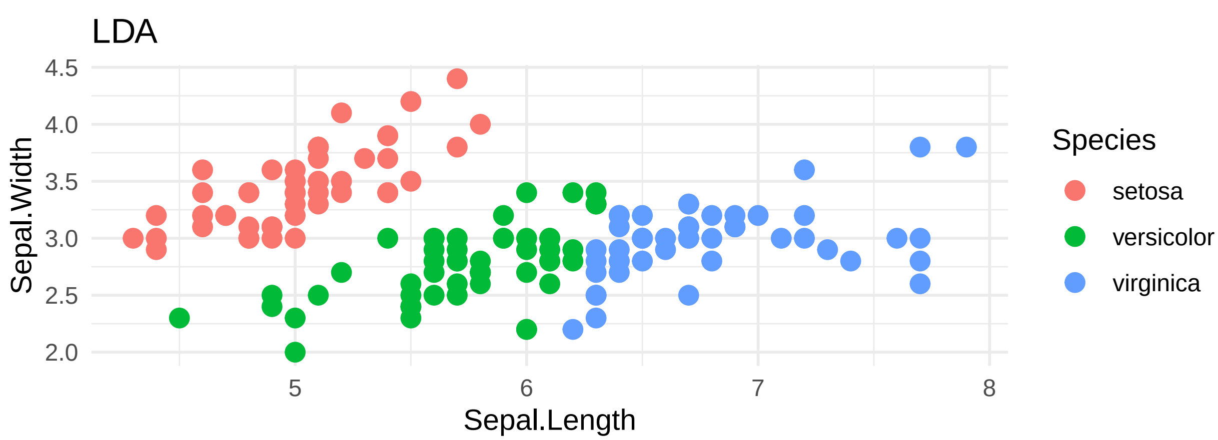
Missclasification rate of LDA using MASS library: 0.2

Test errors are the same, Classification methods are identical so and the results are identical

4.

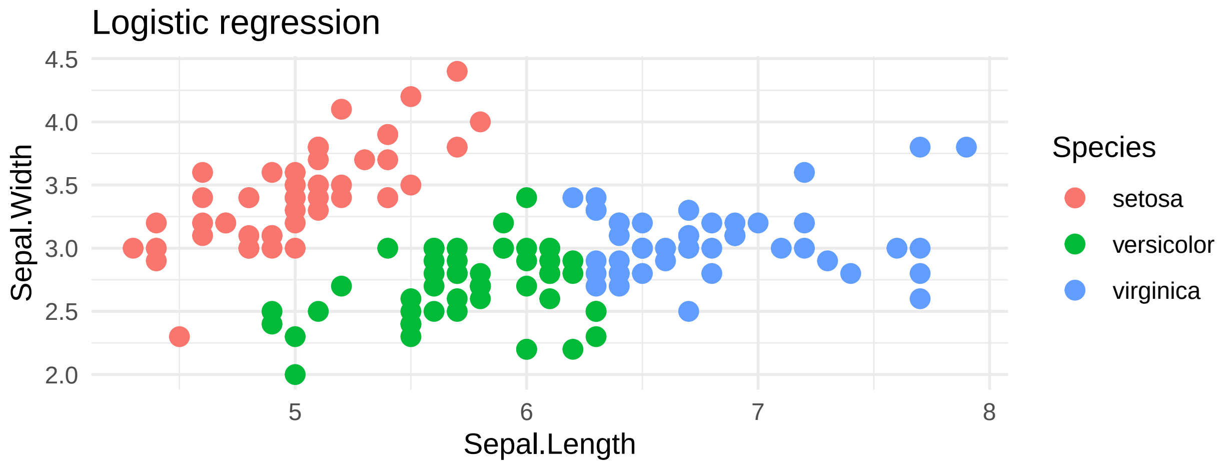


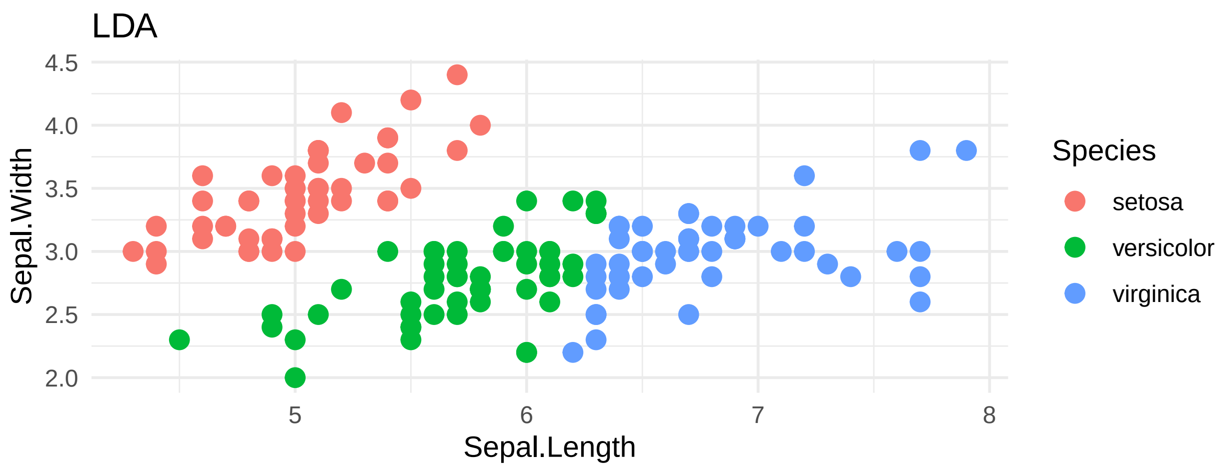




From Plots we can see generated data is spread equally, it is because of LDA assumption that covariances are equal. We can also notice that LDA can not distinguish classes when data overlay.

5.





	Misclasification rate
LDA	0.2000000
LR	0.1666667

From misclasification rate we can see that logistic regresion performed slightly better than LRA.