```
  versicolor     0       30      1
  virginica      0        1     34
> (Training_error_rate <- mean(pred$class!=train$Species))
[1] 0.02
```

类似地，可在测试集中进行预测，并计算测试集的混淆矩阵与错误率：

```
> pred <- predict(fit,newdata=test)
> table(Predicted=pred$class,Actual=test$Species)
            Actual
Predicted    setosa versicolor virginica
  setosa        16          0         0
  versicolor     0         17         0
  virginica      0          2        15
> (Test_error_rate <- mean(pred$class!=test$Species))
[1] 0.04
```

结果显示，使用二次判别分析，其测试集错误率为4%，反而高于线性判别分析的测试集错误率2%。这表明，对于 iris 数据集，其决策边界更接近于线性，故使用线性判别分析更有效率。

附录

A7.1 总体中的多分类费雪判别分析

首先考虑在总体中进行多分类问题的费雪线性判别分析（假设所有总体参数为已知），然后再推广到样本数据中。记第 k 类数据的总体均值为 $\boldsymbol{\mu}_k=E(\boldsymbol{x}\mid y=k)$，其中 $k=1,\cdots,K$；而所有数据的总体均值为 $\boldsymbol{\mu}=\dfrac{1}{K}\sum_{k=1}^{K}\boldsymbol{\mu}_k$（简单起见，假设各类数据所占比重相同）。对于投影方向 $\boldsymbol{w}$，考虑线性组合 $z=\boldsymbol{w}'\boldsymbol{x}$。对于第 k 类数据，此线性组合的均值为

$$\mu_z^{(k)}\equiv E(z\mid y=k)=\boldsymbol{w}'E(\boldsymbol{x}\mid y=k)=\boldsymbol{w}'\boldsymbol{\mu}_k \tag{7.39}$$

对于所有数据，线性组合 z 的均值为

$$\bar{\mu}_z=\frac{1}{n}\sum_{k=1}^{K}\mu_z^{(k)}=\frac{1}{n}\sum_{k=1}^{K}\boldsymbol{w}'\boldsymbol{\mu}_k=\boldsymbol{w}'\left(\frac{1}{n}\sum_{k=1}^{K}\boldsymbol{\mu}_k\right)=\boldsymbol{w}'\boldsymbol{\mu} \tag{7.40}$$

假设各类数据的协方差矩阵均相等，即 $\boldsymbol{\Sigma}_1=\cdots=\boldsymbol{\Sigma}_K=\boldsymbol{\Sigma}$，则线性组合 $z=\boldsymbol{w}'\boldsymbol{x}$ 的协方差矩阵为（无论为哪类数据）

$$\mathrm{Var}(z)=\mathrm{Var}(\boldsymbol{w}'\boldsymbol{x})=\boldsymbol{w}'\mathrm{Var}(\boldsymbol{x})\boldsymbol{w}=\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} \tag{7.41}$$

作为对组间方差的度量，考虑投影后每类数据的中心($\mu_z^{(k)}$)与所有数据的中心($\bar{\mu}_z$)距离的平方和：