2) ISLR 6.6.3

a) Training RSS

iv) Steadily decreases

Because as $s$ increases it minimizes the RSS. When $s$ becomes sufficiently large RSS will satisfy the constraint. $\beta$ also minimizes the RSS in this situation.

b) Test RSS

ii) Decrease initially and then eventually start increasing in a U shape

$\beta_p$ are all set to 0, when $s=0$. As $s$ starts increasing it starts to fit the model well on test data. so Test RSS decreases when it further increases, it will try to overfit the data and test RSS will start to increase.

c) Variance

iii) Steadily increase

When '$s$' is '0' the model will predict only the constant and variance will be less. When '$s$' starts to increase, it starts to include more $\beta_p$ and it will increase the variance.

d) Bias

iv) steadily decreases

same as point c When model becomes more flexible by increasing $s$ from zero it will include more $\beta_p$ so bias decreases

c) Irreducible error.

i) Remains constant

As the irreducible error is independant of the model, it will remain constant.

3) SLR 6.6.5

a) We need to minimize

$$\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2 + \lambda \sum_{j=1}^{p} R_j^2$$

We have,

$\hat{\beta}_0 = 0$ & $n = p = 2$.

minimize

$$\sum_{i=1}^{2} (y_i - \sum_{j=1}^{2} \beta_j x_j)^2 + \lambda \sum_{j=1}^{2} \beta_j^2$$

$$(y_1 - \beta_1 x_{11} - \beta_2 x_{12})^2 + (y_2 - \beta_1 x_{21} - \beta_2 x_{22})^2 + \lambda(\beta_1^2 + \beta_2^2)$$

b) From previous question we have

$f(\hat{\beta}_1, \hat{\beta}_2)$

$(y_1 - \hat{\beta}_1 x_{11} - \hat{\beta}_2 x_{12})^2 + (y_2 - \hat{\beta}_1 x_{21} - \hat{\beta}_2 x_{22})^2 + \lambda(\beta_1^2 + \beta_2^2)$

$= 2(y_1 - \hat{\beta}_1 x_{11} - \hat{\beta}_2 x_{11})^2 + \lambda(\beta_1^2 + \beta_2^2)$

$= 2(y_1^2 + \hat{\beta}_1^2 x_{11}^2 + \hat{\beta}_2^2 x_{11}^2 - 2y_1\hat{\beta}_1 x_{11} - 2y_1\hat{\beta}_2 x_{11} - 2\hat{\beta}_1 x_{11}^2 \hat{\beta}_2)$

$= 2y_1^2 + 2\hat{\beta}_1^2 x_{11}^2 + 2\hat{\beta}_2^2 x_{11}^2 - 4y_1\hat{\beta}_1 x_{11} - 4y_1\hat{\beta}_2 x_{11} - 4\hat{\beta}_1 x_{11}^2 \hat{\beta}_2$

Let us differentiate w.r.t $\hat{\beta}_1$ & $\hat{\beta}_2$ to minimize

$$\frac{\partial f(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_1} = -4y_1 x_{11} + 4x_{11}^2 \hat{\beta}_2 + 4x_{11}^2 \beta_1 + 2\lambda\hat{\beta}_1 = 0$$

$$\hat{\beta}_1(\lambda + 2x_{11}^2) = 2y_1 x_{11} - 2x_{11}^2 \hat{\beta}_2 = 0$$

$$\Rightarrow \hat{\beta}_1 = \frac{2y_1 x_{11} - 2x_{11}^2 \hat{\beta}_2}{\lambda + 2x_{11}^2}$$

$$\frac{\partial f(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_2} = -4y_1 x_{11} + 4x_{11}^2 \hat{\beta}_1 + 4x_{11}^2 \hat{\beta}_2 + 2\lambda\hat{\beta}_2 = 0$$

$$\hat{\beta}_2(\lambda + 2x_{11}^2) = 2y_1 x_{11} - 2x_{21}^2 \hat{\beta}_1$$

$$\Rightarrow \hat{\beta}_2 = \frac{2y_1 x_{11} - 2x_{11}^2 \hat{\beta}_1}{\lambda + 2x_{11}^2}$$

from above equations $\hat{\beta}_1 = \hat{\beta}_2$

c) Similar to Ridge

minimize $\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda \sum_{j=1}^{p}|\beta_j|$

$\Rightarrow \sum_{i=1}^{2}(y_i - \sum_{j=1}^{2}\beta_j x_{ij})^2 + \lambda \sum_{j=1}^{2}|\beta_j|$

$= (y_1 - \beta_1 x_{11} - \beta_2 x_{12})^2 + (y_2 - \beta_1 x_{21} - \beta_2 x_{22})^2 + \lambda(|\beta_1| + |\beta_2|)$

d) We have lasso eqtn

$\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2$ subject to $\sum_{j=1}^{p}|\beta_j| \le s$

so must minimize lasso coefficients

$2(y_1 - (\hat{\beta_1} + \hat{\beta_2})x_{11})^2 \ge 0$

subject to $|\hat{\beta_1}| + |\hat{\beta_2}| \le s$

The contours will touch the lasso diamond at many points & hence has many solutions

4) ISLR 8.4.5.

i) Based on majority vote

Red = 6, green = 4.

so class predicted is red

ii) Based on average probability

Average is 0.45.

so predicted class is green.

5) ISLR 9.7.3.

a) Refer Jupyter Notebook

b) Equation: $-0.5 + x_1 = x_2$.

$-0.5 + x_1 - x_2 > 0$.

c) classification rules

$-\frac{1}{2} + x_1 - x_2 > 0 \rightarrow$ Blue

$-\frac{1}{2} x_1 - x_2 \leq 0 \rightarrow$ Red

d) Refer Jupyter Notebook

e) Refer Jupyter Notebook

f) Refer Jupyter Notebook

g) Equation can be $-0.25 + x_1 - x_2 > 0$

h) Refer Jupyter Notebook