

Title

Liheng Cao

February 22, 2021

1

- $X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

- $y = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 7 \end{bmatrix}$

- The closed form solution is $(X^T X)^{-1} X^T y$

- $\begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = (X^T X)^{-1} X^T y = \left(\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 4 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 3/4 \\ 5/2 \\ 7/2 \end{bmatrix}$

- $\text{RSS} = \sum_{i=1}^4 \left(3/4 + 5/2 x_1^{(i)} + 7/2 x_2^{(i)} - y^{(i)} \right)^2$
 $= (3/4 + 5/2 \cdot 0 + 7/2 \cdot 0 - 1)^2 + \dots + (3/4 + 5/2 + 7/2 - 7)^2 = 0.25$

- $\text{TSS} = \sum_{i=1}^4 \left(y^{(i)} - \bar{y}^{(i)} \right)^2 =$
 $= (1 - 3.75)^2 + \dots + (7 - 3.75)^2 = 18.75$

- $R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} = 1 - \frac{0.25}{18.75} = 0.9867$

- This is the same thing as R^2 . 98.% of the variance in y is explained by x .

2

$$w = [8.29 \dots \quad -0.66 \dots \quad -0.90 \dots \quad 4.58 \dots \quad -0.24 \dots]^T, N = 506$$

- $\text{RSS} = \sum_{i=1}^N (w^T \cdot X^{(i)} - y^{(i)})^2 = 14517.66 \dots$
- $\text{TSS} = \sum_{i=1}^4 (y^{(i)} - \hat{y}^{(i)})^2 = 42716.30 \dots$
- $R^2 = 1 - \frac{14517.66 \dots}{42716.30 \dots} = 1 - \frac{0.25}{18.75} = 0.6601$
- 66.01% of the variance in y can be explained by x .

3

I would put the actual crop yields in a column vector y , and then I would put the amount of rainfall, the amount of fertilizer, the average temperature, and the number of sunny days into a matrix X . This matrix would have the first column be composed entirely of ones for the constant “trick”. If we use a linear model, then we can find the optimal solution using $w = (X^T X)^{-1} X^T y$.

4

$$w = (w_1, w_2)$$

- f_1

1. (7.0, 17.0)
2. (-85.0, -222.0)
3. (1202.0, 3106.0)
4. (-16725.0, -43265.0)

- f_1

1. (15.0, 20.0)
2. (-95.0, -460.0)
3. (2045.0, 10060.0)
4. (-44315.0, -219420.0)