CS 66: Machine Learning

Prof. Sara Mathieson

Spring 2019

SWARTHMORE

Lab 2: in-lab notes

Linear Regression so far

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

$\boxed{\text{want to minimize RSS}}$

$$RSS\Big|_{b_0} = 2 \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) = 0$$

$$\Rightarrow \quad n b_0 = \sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i$$

divide through by n

$$\Rightarrow \quad \boxed{\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}}$$

$$RSS\Big|_{b_1} \Rightarrow \cancel{2} \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)\, x_i = 0$$

$$\sum_{i=1}^{n} (y_i - \bar{y} + b_1 \bar{x} - b_1 x_i) x_i = 0$$

$$\sum_{i=1}^{n} (y_i x_i - \bar{y} x_i) - b_1 \sum_{i=1}^{n} (x_i^2 - \bar{x} x_i) = 0$$

$$b_1 = \frac{\sum_{i=1}^{n} (x_i y_i - x_i \bar{y}) + 0}{\sum_{i=1}^{n} (x_i^2 - \bar{x} x_i) + 0}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Note

$x_i) = 0$

$$\boxed{\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

<u>Note</u>: $\sum_{i=1}^{n}(\bar{x}^2 - x_i\bar{x})$

$$= \sum_{i=1}^{n}\bar{x}^2 - \bar{x}\sum_{i=1}^{n} x_i$$

$$= n\bar{x}^2 - \bar{x}(n\bar{x})$$

$$= 0$$

$$\sum_{i=1}^{n}(\bar{x}\bar{y} - y_i\bar{x}) = 0$$

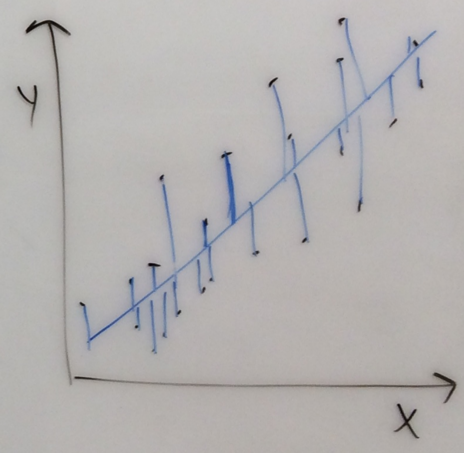
$$b_1 = \frac{\sum_{i=1}^{n}(x_i y_i - x_i\bar{y} + \bar{x}\bar{y} - y_i\bar{x})}{\sum_{i=1}^{n}(x_i^2 - \bar{x}x_i + x^2 - x_i\bar{x})}$$

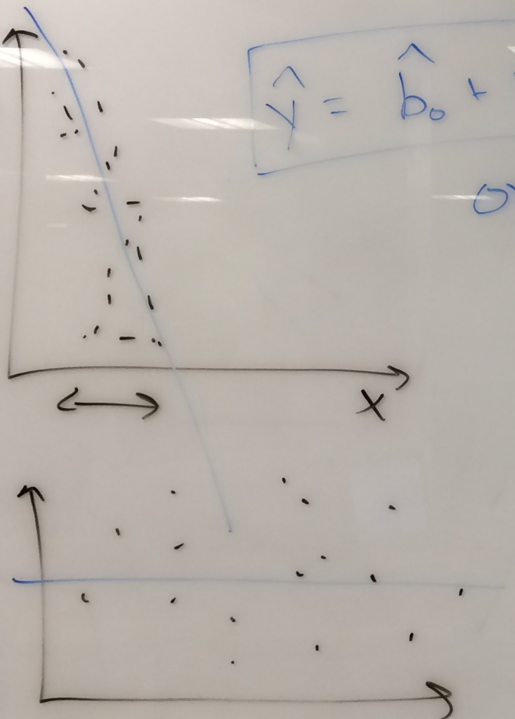
$$b_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\boxed{\hat{b}_1 = \frac{Cov(x,y)}{Var(x)}}$$

$$x_i(y_i - \bar{y}) - \bar{x}(y_i - \bar{y})$$

$$\boxed{x_i^2 - 2x_i\bar{x} + \bar{x}^2}$$





$$\boxed{\hat{y} = \hat{b}_0 + \hat{b}_1 x}$$

one feature model

D-Tree notes

1. multiple features have same entropy

2.