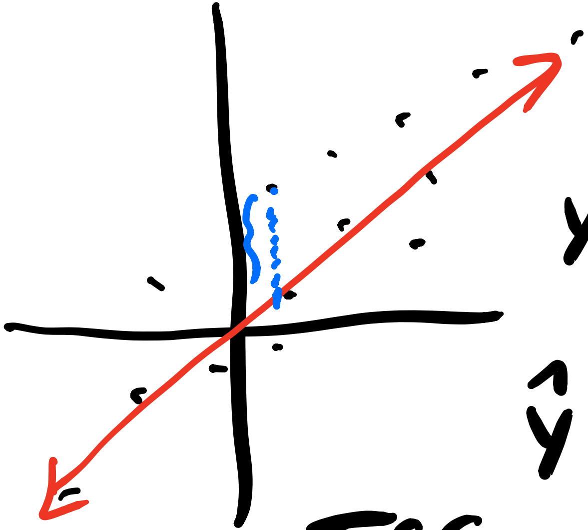
Linear Algebra

Linear Regression

Michael Ruddy

Line of Best Fit

Data: $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})$



$$\beta_0 + \beta_1 x = y$$

$y^{(i)}$ – truth

$$\boxed{y^{(i)} - (\beta_0 + \beta_1 x^{(i)})}$$

$$\hat{y}^{(i)}_{m} = \beta_0 + \beta_1 x^{(i)}$$

$$TSS = \sum_{i=1} \left(y^{(i)} - \hat{y}^{(i)}\right)^2$$

Line of Best Fit

$$X = \begin{pmatrix} 1 & x^{(1)} \\ 1 & x^{(2)} \\ \vdots & \vdots \\ 1 & x^{(m)} \end{pmatrix}$$

$$\hat{y} = X \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} \beta_0 + \beta_1 x^{(1)} \\ \beta_0 + \beta_1 x^{(2)} \\ \vdots \\ \beta_0 + \beta_1 x^{(m)} \end{pmatrix}$$

$$\vec{y} = \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{pmatrix}$$

$$\vec{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

Line of Best Fit

$$\vec{y} = X\vec{B}$$

$$TSS = (\vec{y} - X\vec{B}) \cdot (\vec{y} - X\vec{B})$$

$$\vec{y} - X\vec{B} = \begin{pmatrix} y^{(1)} - (\beta_0 + \beta_1 x^{(1)}) \\ y^{(2)} - (\beta_0 + \beta_1 x^{(2)}) \\ \vdots \end{pmatrix}$$

- $\vec{\beta}$ is unknown
- X $\vec{y}$ are our data

Line of Best Fit

$$TSS = \sum_{i=1}^{m} \left(y^{(i)} - (\beta_0 + \beta_1 x^{(i)}) \right)^2$$

$$\frac{\partial TSS}{\partial \beta_0} = \sum_{i=1}^{m} 2 \left(y^{(i)} - (\beta_0 + \beta_1 x^{(i)}) \right) = 0$$

$$\frac{\partial TSS}{\partial \beta_1} = \sum_{i=1}^{m} x^{(i)} \cdot 2 \left(y^{(i)} - (\beta_0 + \beta_1 x^{(i)}) \right) = 0$$

$$\nabla TSS = \begin{bmatrix} \partial TSS / \partial \beta_0 \\ \partial TSS / \partial \beta_1 \end{bmatrix} = 2X^T(\vec{y} - X\vec{\beta}) = 0$$

Linear Regression

$$\left(x_1^{(1)}, x_2^{(1)}, \ldots, x_N^{(1)}, y^{(1)}\right) \quad y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 +$$

$$\left(x_1^{(2)}, \ldots \ldots x_N^{(2)}, y^{(2)}\right) \qquad \ldots + \beta_N x_N$$

$$\vdots$$

$$X = \begin{pmatrix} 1 & x_1^{(1)} & \cdots & x_N^{(1)} \\ 1 & x_1^{(2)} & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_1^{(m)} & & x_N^{(m)} \end{pmatrix}$$

$$TSS = (y - X\beta)(y - X\beta)$$

$$\vec{y} = \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{pmatrix} \qquad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_N \end{pmatrix}$$

$$\nabla TSS = 2X^T(y - X\beta)$$

$$= 0$$

Linear Regression Gradient Descent!

(2, 3)
(-1, -7)
(5, 7)
(-3, -2)

$\beta_0 + \beta_1 x = y$

$\vec{\beta}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$X = \begin{pmatrix} 1 & 2 \\ 1 & -7 \\ 1 & 5 \\ 1 & -3 \end{pmatrix} \quad y = \begin{pmatrix} 3 \\ -1 \\ 7 \\ -2 \end{pmatrix}$$

$2X^T(y - X\beta)$

$\begin{bmatrix} ? \\ ? \end{bmatrix}$

$\vec{\beta}_2 = \vec{\beta}_1 - \lambda$