Linear Regression

- output (y) is continuous
- Learned Model: linear function mapping input to output (a weight for each feature + bias)
- Goal: minimize the RSS or SSE (sum of squared errors)

Why are models useful?

- Visualizing data, drawing conclusions
- make predictions

Linear Models? (continuous, not discrete)

- which of the features (x) are associated with the response variable (y)?
- relationship b/w x & y?
- Given x can we predict y?
- is it enough?

~~Linear Regression: output is continuous~~

weights/params

Ex: $h_{\vec{w}} = w_0 + w_1 x = \hat{y}$; $\vec{w} = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$

minimize $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

actual predicted

Board:

cost function

$$J(w_0, w_1) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad \& \quad h_{\vec{w}} = w_0 + w_1 x = \hat{y}$$

$$= \sum_{i=1}^{n} (y_i - (w_0 + w_1 x))^2$$

Analytical solution: $\frac{\partial J}{\partial w_0} = 0$; $\frac{\partial J}{\partial w_1} = 0$

↓

J, with $\frac{1}{2} = J(w_0, w_1) = \frac{1}{2}\sum_{i=1}^{n}(y_i - (w_0 + w_1 x_i))^2$ 6

a) $\frac{\partial J}{\partial w_0} = 0$

$$2 \cdot \frac{1}{2}\sum_{i=1}^{n}(y_i - w_0 - w_1 x_i)(-1) = 0$$

$$= -\sum_{i=1}^{n} y_i + \underbrace{\sum_{i=1}^{n} w_0}_{n w_0} + \sum_{i=1}^{n} w_1 x_i = 0$$

$$n w_0 = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} w_1 x_i$$

$$w_0 = \frac{1}{n}\sum_{i=1}^{n} y_i - \frac{1}{n}\sum_{i=1}^{n} w_1 x_i$$

$$= \frac{1}{n}\sum_{i=1}^{n} y_i - w_1 \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$w_0 = \bar{y} - w_1 \bar{x}$$

b) $\frac{\partial J}{\partial w_1} = 0$

$$2 \cdot \frac{1}{2}\sum_{i=1}^{n}(y_i - w_0 - w_1 x_i)(-x_i) = 0$$

$$= \sum_{i=1}^{n}(y_i - \bar{y} + w_1 \bar{x} - w_1 x_i)(-x_i)$$

$$= -\sum_{i=1}^{n} y_i x_i + \sum_{i=1}^{n} \bar{y} x_i - \sum_{i=1}^{n} w_1 \bar{x}(x_i) + \sum_{i=1}^{n} w_1 x_i^2 = 0$$

$$= w_1\left(-\sum_{i=1}^{n} \bar{x} x_i + \sum_{i=1}^{n} x_i^2\right) - \sum_{i=1}^{n} y_i x_i + \sum_{i=1}^{n} \bar{y} x_i$$

$$w_1 = \frac{-\sum_{i=1}^{n}(y_i x_i) + \sum_{i=1}^{n}(\bar{y} x_i)}{\sum_{i=1}^{n} \bar{x} x_i - \sum_{i=1}^{n} x_i^2}$$

$$w_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})} = \frac{cov(x, y)}{var(x)}$$

for cov, we care about magnitude & sign, but only magnitude for var.