Linear Regression

$x_3 = \ldots$

$x_2 = \text{humidity}$

$x_1 = \text{day}$

$y = \text{temperature}$ °F

$x^{(1)} = 1$

$y^{(1)} = 64$

May

	x_1(day)	y (temp)
1	5/1	64°
2	5/2	65°
3	5/3	58°
⋮	⋮	⋮
31	5/31	20°
32	6/1	?

31

"h": $x \rightarrow y$

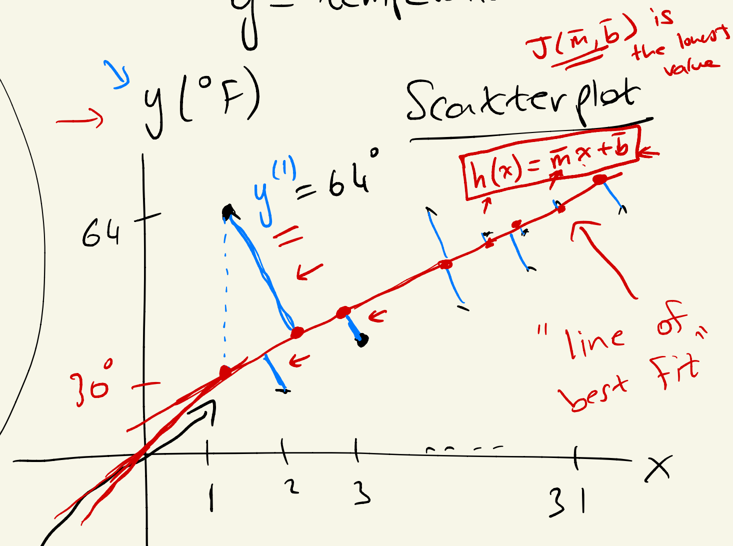
$\boxed{h(x) = mx + b}$

slope

y-int

$\boxed{y = mx + b}$

Scatterplot



y

$\hat{y}$ = predicted value

$(y - \hat{y})$ ← how different is the real value from predicted value

$$h(x) = \boxed{m}x + \boxed{b}$$

$x^{(i)}$

$y^{(i)}$

$(x^{(i)}, y^{(i)})$

$n = 31$ $\begin{cases} (x^{(1)}, y^{(1)}) \\ (x^{(2)}, y^{(2)}) \\ \vdots \\ (x^{(n)}, y^{(n)}) \end{cases}$

Training set — n different samples

learning algorithm

model

(input) $x \rightarrow \boxed{h} \rightarrow \hat{y}$ = prediction ← output

$28 \rightarrow 67°F$

Mean Squared Error

Mean Squared Error = Tells us the average difference between the real output and the predicted output.

$\frac{1}{n}\sum_i (\hat{y}^{(i)} - y^{(i)})^2$

error function "cost"

$$J(m, b) = \frac{1}{2n} \cdot \sum_{i=1}^{n} \left(\underbrace{h(x^{(i)})}_{\hat{y}^{(i)}} - y^{(i)}\right)^2$$

\# samples

sum

Squared sum of all differences for our predicted and real output

Cost function

$$\boxed{J(m,b) = \frac{1}{2n} \cdot \sum_{i=1}^{n} (mx^{(i)} + b - y^{(i)})^2}$$

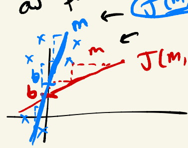
Goal of linear regression:

$$\boxed{\min_{m,b} J(m,b)}$$

We want this to be as small as possible

$(\bar{m}, \bar{b})$

$\boxed{h(x) = \bar{m}x + \bar{b}}$



$J(m,b) = 2$

$J(m,b) = 6$

How do we get the optimal slope (m) and y-intercept (b)?

Gradient Descent

$$\min_{m,b} J(m,b)$$

Initialize m and b randomly

steps

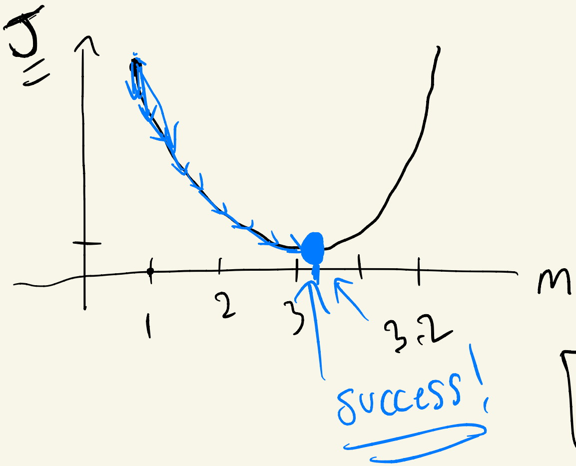
repeat:

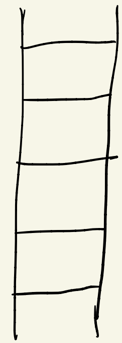
learning rate

$$m = m - \alpha \cdot \boxed{\frac{\partial J(m,b)}{\partial m}}$$

$$b = b - \alpha \cdot \frac{\partial J(m,b)}{\partial b}$$

$$J(m) = \uparrow\uparrow\uparrow$$





$$h(32) = (3.2)(32) = \boxed{102.4^\circ F}$$

$$\boxed{h(x) = \bar{m}x} = 3.2x$$

$$\boxed{\bar{m} = 3.2}$$

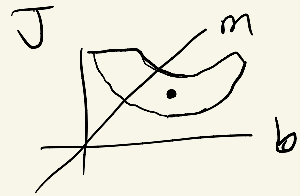
Multivariate Linear Regression

day x_1	humidity x_2	... x_3	y (temp)

$$\boxed{h(x_1, x_2, x_3) = m_1 x_1 + m_2 x_2 + m_3 x_3 + b}$$

"line of best fit"

Optimal parameter: $\boxed{\bar{m}_1, \bar{m}_2, \bar{m}_3, b}$



m, b $\quad$ $y = mx + b$

Linear Regression

1. Look at data/graphed
2. $h(x) = mx + b$ "line of best fit"
3. $J(m, b)$ defined
4. $\min\limits_{m,b} J(m, b)$
5. Gradient Descent to update m and b for a number of iterations to get our optimal $\bar{m}$ and $\bar{b}$.
6. $h(x) = \bar{m}x + \bar{b}$