$X = [x_1, x_2, x_3 \ldots x_n]$ → Learning from Data → $\theta_0 = \bar{y} - \theta_1 \bar{x}$, $\theta_1 = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\overline{x^2} - \bar{x}}$

$y = \theta_0 + \theta_1 x$

Linear Regression

- Closed form solution $w = (X^T X)^{-1} X^T y$ → vector notation.

Learning from data

- Independent variable, Input, features
- dependent variable, Output, target (label)

Examples: stock price prediction, $[(x)] \rightarrow y$, $= \mathbb{R}$ → 235.35

- Quantitative variables, Qualitative variables ⇒ Categorical.

define the above terms with examples

Caesarian Delivery Predictor

Input: Number of delivery, 1, 2, 3 ⇒ $\mathbb{N}$

Patient Vitals = ?

Output: proceed with Caesarian delivery → True, False } = 1, = 0 (−1)

→ Cat / Dog } Qualitative Variable.

Convert to some meaningful numbers.

Dummy variable (one-hot encoding)

$x = \{$ cat, dog, zebra $\}$

⇓ one hot encoding $[x = 0, 1, 2]$

$$x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \text{cat} \\ \text{dog} \\ \text{zebra} \end{matrix}$$

Clas. $y =$ ⇓ [Gradient descent] → local min.

Computer Algorithm ← (Learn from data $D = \{(x_1, y_1) \cdots (x_n, y_n)\}$

$y = \theta_0 + \theta_1 x$ ← parameters

$y = mx + c$

Linear Regression: Worked out Example

$m = \frac{\text{rise}}{\text{run}}$

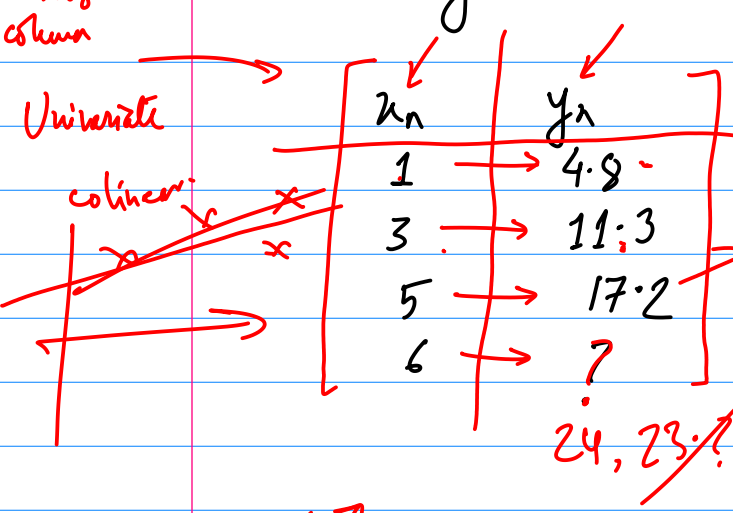
right column

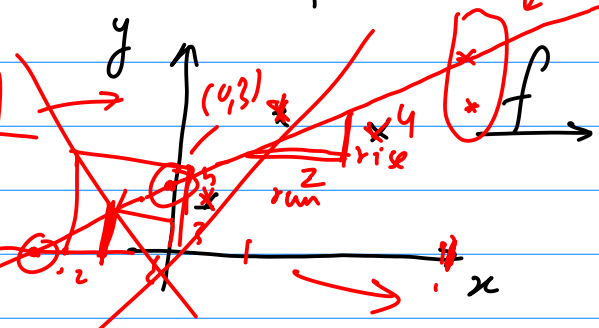
Univariate

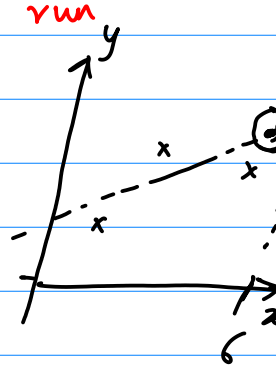
collinear

x_n	y_n
1	4.8
3	11.3
5	17.2
6	?

24, 23.?







We want to fit a line that closely passes through the given point.

$y = 2x + c$

$[y = 2x + 3]$

$y = mx - ?$

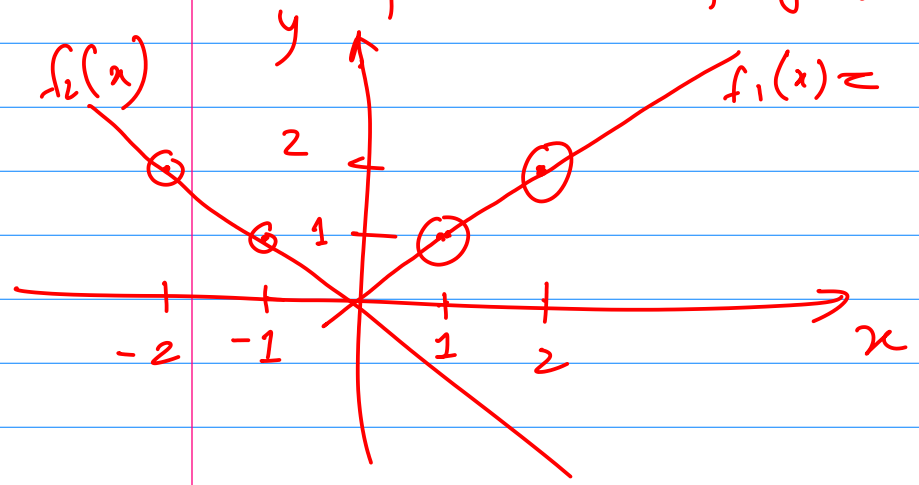
$y =$

The model has the form $y = f(x) = \theta_0 + \theta_1 x$

This is an example of Univariate linear regression.

Discuss possible solution for $y = f(x) = \theta_0 + \theta_1 x$ $[y = mx + c]$

$y = mx + c$



$f_1(x) = \boxed{1}\, x + \boxed{0}$

$f_2(x) = \boxed{-1}\, x + \boxed{0}$

m c

$f_1(x)$ $m = \frac{\text{rise}}{\text{run}} = \frac{1}{1} = 1$

$y = mx + c$

$\downarrow$

$y = \theta_1 x + \theta_0$

$f_2(x) = m$

$x_1, x_2, x_3, \cdots, x_n \longrightarrow y_n$

$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots \theta_n x_n$

multivariate.

Closed form solution for $y = f(x) = \theta_0 + \theta_1 x$ [Univariate Regression]

Under certain criteria, the best possible value of parameters in linear model.

Analytic solution (granular).

$$\left[\theta_1 = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\overline{x^2} - (\bar{x})^2}, \quad \theta_0 = \bar{y} - \theta_1 \cdot \bar{x}\right] \quad \text{---} \quad (1)$$

θ_1 = slope

$\theta \to ?$

$$\bar{x} = \frac{x_1 + x_2 + \cdots x_n}{n} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

x_n	y_n	$x_n \cdot y_n$	x^2
1	4.8	4.8	1
3	11.3	33.9	9
5	17.2	86	25
6	?		

$\theta_0 = f(x)_{x=0}$

$\bar{x} = 3 \quad \bar{y} = 11.1 \quad \overline{xy} = 41.57 \quad \overline{x^2} = 11.67$

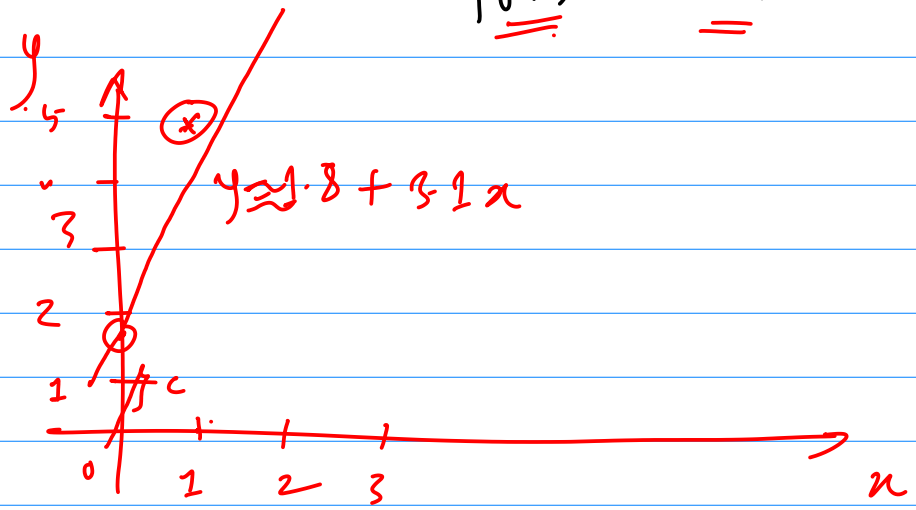
$y = mx + c = \theta_0 + \theta_1 x \to \theta$.

Putting values into eqn. (1), $\theta_1 = \frac{41.57 - 3 \times 11.1}{11.67 - 3 \times 3} = 3.1$, $\theta_0 = 11.1 - 3.1 \times 3 = 1.8$

(intercept, bias, DC value)

Hence, $\left[y = f(x;\theta) = f_\theta(x) = 1.8 + 3.1x\right]$

$f_\theta(1) = 20.4$



$y = f_\theta(x) = \theta_0 + \theta_1 x$

$\theta_1 = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\overline{x^2} - (\bar{x})^2}$ ✓

$\theta_0 = \bar{y} - \theta_1 \bar{x}$

[# Where did this magic formula come from?]

Given a dataset $D = \{(x_1, y_1), \dots (x_n, y_n)\}$, find a functional relationship using a linear model

$f(x;\theta) = f_\theta(x) = \theta_0 + \theta_1 x$ | $(\theta_0, \theta_1) \rightarrow$ find.

$(-1.7, 2) \leftrightarrow (2.1, 3)$

→ How do we quantify when the solution is good enough?

[Cost function:] → 0 means this fits the data very well

(loss function) @ Mean squared Error (MSE) or l2 norm

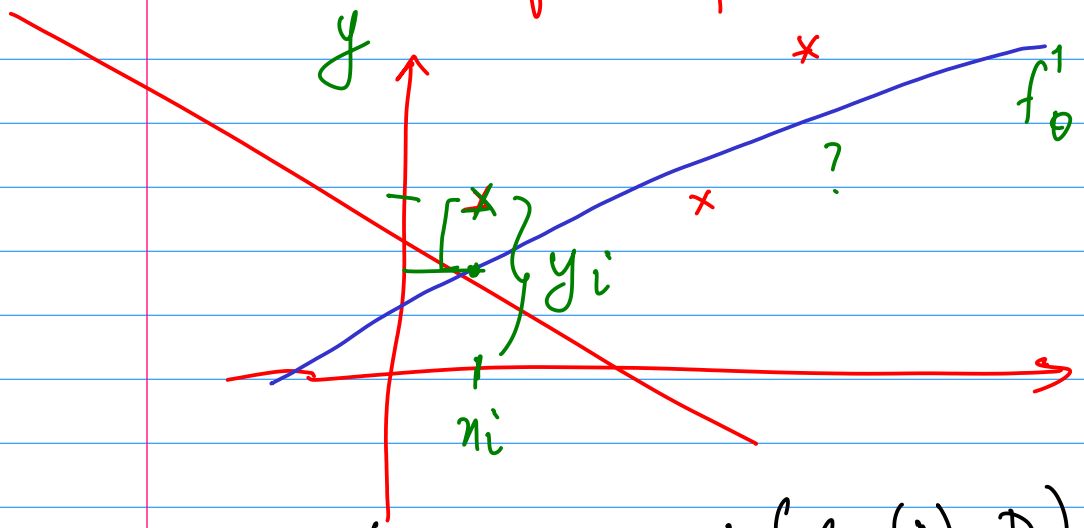
$= L(f_\theta(x), D) = \frac{1}{N}\sum_{i=1}^{N} (f_\theta(x_i) - y_i)^2 \rightarrow$ 1

predicted

ground truth

[Smooth, convex, differentiable]

Evaluate this cost function for some θ and some dataset D.

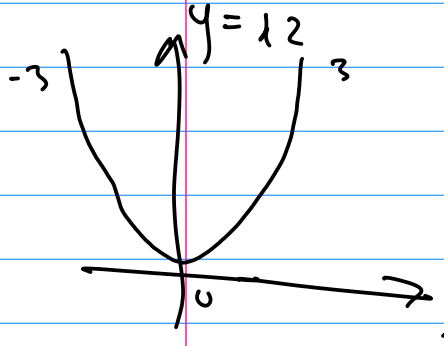


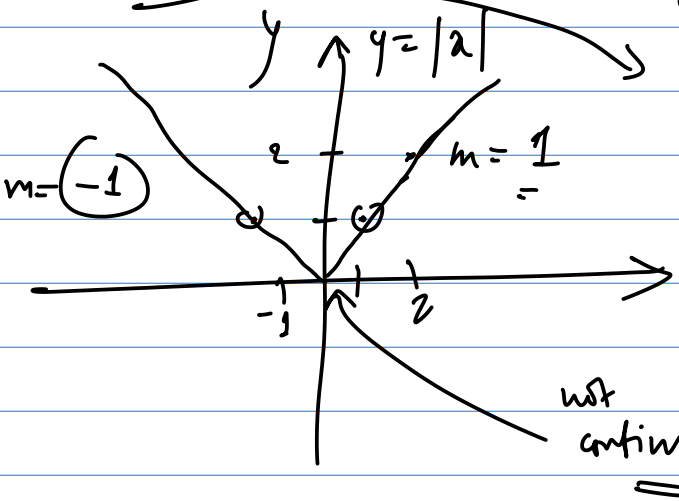
red line → 10.35

blue line → 35.6

cost function

cost function MAD $L(f_\theta(x), D) = \frac{1}{N}\sum_{i=1}^{N} |f_\theta(x_i) - y_i|$





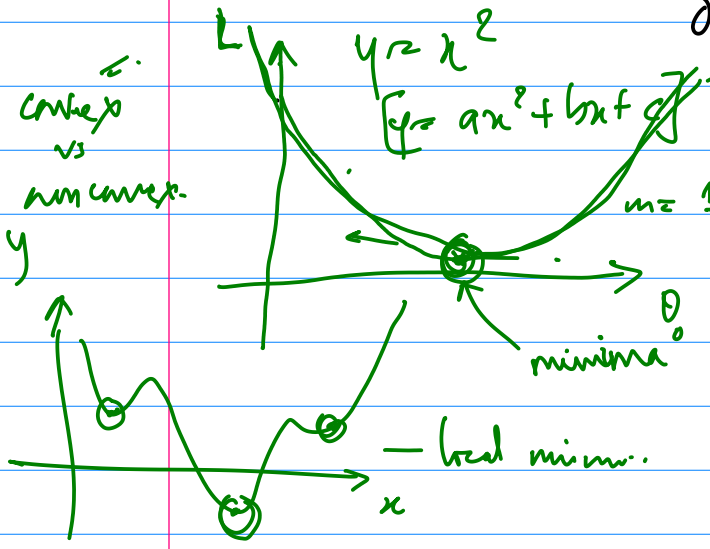
no analytic solution?

$\lim_{h\to 0^+} \frac{f(x+h) - f(x)}{h} \neq \lim_{h\to 0^-} \frac{f(x+h) - f(x)}{h}$

not continuous at this point

cost function (MSE) $\frac{dL}{c}$ $\frac{dy}{dx}$ $\frac{dL}{dslope}$ $\frac{dL}{dint}$ $\frac{\partial L}{\partial \theta_0}$ $\frac{\partial L}{\partial \theta_1}$

[Use the idea that $\frac{\partial L}{\partial \theta} \to 0$ should be 0 if θ is the best solution.



minima

$$L = \frac{1}{N}\sum_{i=1}^{N} L_n\left(f_\theta(x_i), y_i\right)$$

squared $(a-b)^2$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(y_i - (\theta_0 + \theta_1 x_i)\right)^2$$

mean error MSE.

$a^2 - 2ab + b^2$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(y_i^2 - 2(\theta_0 + \theta_1 x_i)y_i + (\theta_0 + \theta_1 x_i)^2\right)$$

$(a+b)^2$

$$L = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^2 - 2\theta_0 y_i - 2\theta_1 x_i y_i + \theta_0^2 + 2\theta_0\theta_1 x_i + \theta_1^2 x_i^2\right)$$

$a^2 + 2ab + b^2$

$\frac{dy}{dx} = 5x$

Differentiating the cost function,

$\frac{dL}{d\theta_0}$, $\frac{\partial L}{d\theta_1}$

$$\frac{dL}{d\theta_0} = \frac{1}{N}\sum_{i=1}^{N}\left(-2y_i + 2\theta_0 + 2\theta_1 x_i\right) = 0 \quad \text{(ii)}$$

$\frac{dL}{d\theta_0}$

$y = 3x + 12$

$\frac{dy}{dx} = 3$

$$\frac{dL}{d\theta_1} = \frac{1}{N}\sum_{i=1}^{N}\left(-2x_i y_i + 2\theta_0 x_i + 2\theta_1 x_i^2\right) = 0 \quad \text{(iii)}$$

Eqn (ii) $\frac{1}{N}\sum_{i=1}^{N}(-2y_i + 2\theta_0 + 2\theta_1 x_i) = 0$

$$\frac{1}{N}\sum_{i=1}^{N}(-y_i + \theta_0 + \theta_1 x_i) = 0$$

$$-\bar{y}_i + \theta_0 + \theta_1 \bar{x} = 0$$

$$\theta_0 = \bar{y}_i - \theta_1 \bar{x} \quad \text{(iv)}$$

Eqn (iii), $-\overline{x_i y_i} + \theta_0 \bar{x}_i + \theta_1 \overline{x_i^2} = 0$

$$\theta_0 \bar{x}_i = \overline{x_i y_i} - \theta_1 \overline{x_i^2}$$

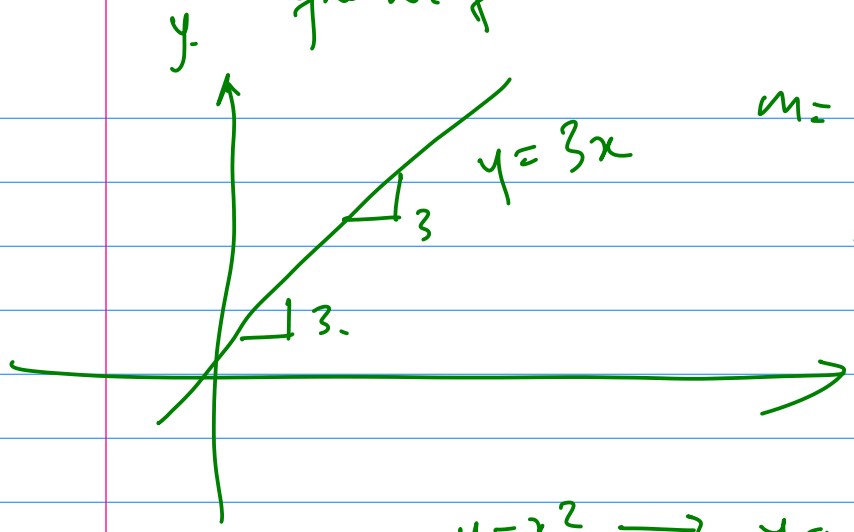
$$\theta_0 = \frac{\overline{x_i y_i} - \theta_1 \overline{x_i^2}}{\bar{x}_i}$$

θ_0, θ_1 2 unk., 2 eqn...

$\frac{1}{N}\sum_{i=1}^{N}\theta_1 x_i = \frac{\theta_1 x_1 + \theta_1 x_2 + \cdots \theta_1 x_N}{N} = \theta_1\left(\frac{x_1 + x_2 \cdots}{N}\right) = \theta_1 \bar{x}$

$\frac{1}{N}\sum_{i=1}^{N}\theta_0 = \frac{\theta_0 + \theta_0 + \theta_0 \cdots}{N} = \theta_0$

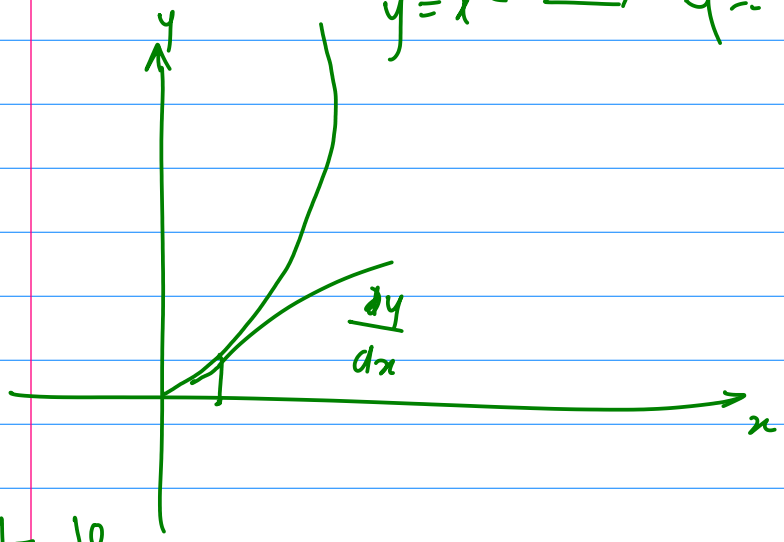
Refresher for Calculus 101 : derivative. (1st).



$$m = \frac{dy}{dx} = 3 \quad (\text{slope}) \; \theta_1 .$$

→ general form of slope

$$\frac{dy}{dx} = f'(x)$$



$y = x^2 \longrightarrow$ y - more fc. - $y = e^x$.

$$\frac{d \sin x}{dx} = \cos\theta$$

$$\frac{dx^2}{dx} = 2x$$

$$\frac{de^x}{dx} = e^x$$

$$\frac{d \log_{10} x}{dx} = ? \frac{1}{x}$$

$$\frac{d \ln x}{dx} = ? \frac{1}{x}$$

$$\frac{d \log_2 x}{dx} = ? \frac{1}{x}$$

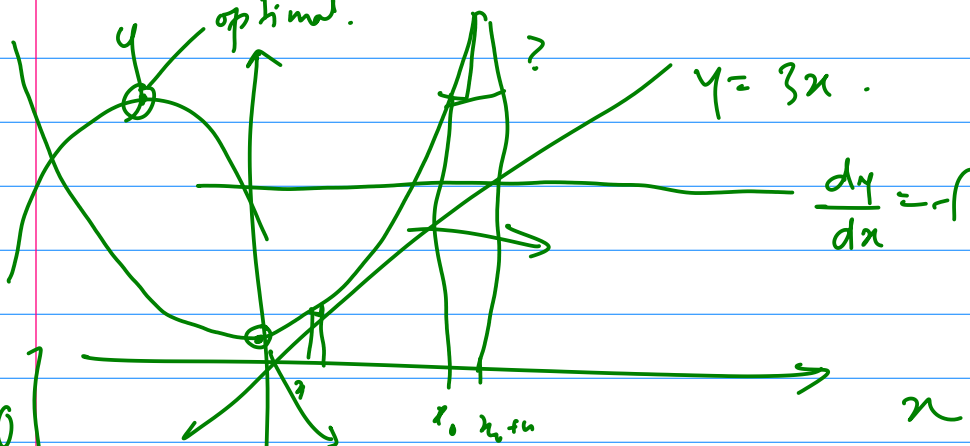
$\log_e 10$.

$\log_e 2$

$\log_{10} x = c \cdot \log_2 x$.

$= c' \cdot \log_e x$.

$$\frac{d\, c \cdot f(x)}{dx} = c \cdot \frac{d f(x)}{dx}$$



$$\frac{dy}{dx} = -f'(x)$$

$$\frac{d^2y}{dx^2}$$

$$\left[\frac{dy}{dx} = 0\right]$$ min

convex

$$\frac{dy}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

$$\frac{dy}{dx} = \frac{dx^2}{dx} = 2x$$

$$\left.\frac{dy}{dx}\right|_{x=0} = 2 \cdot 0 = 0$$