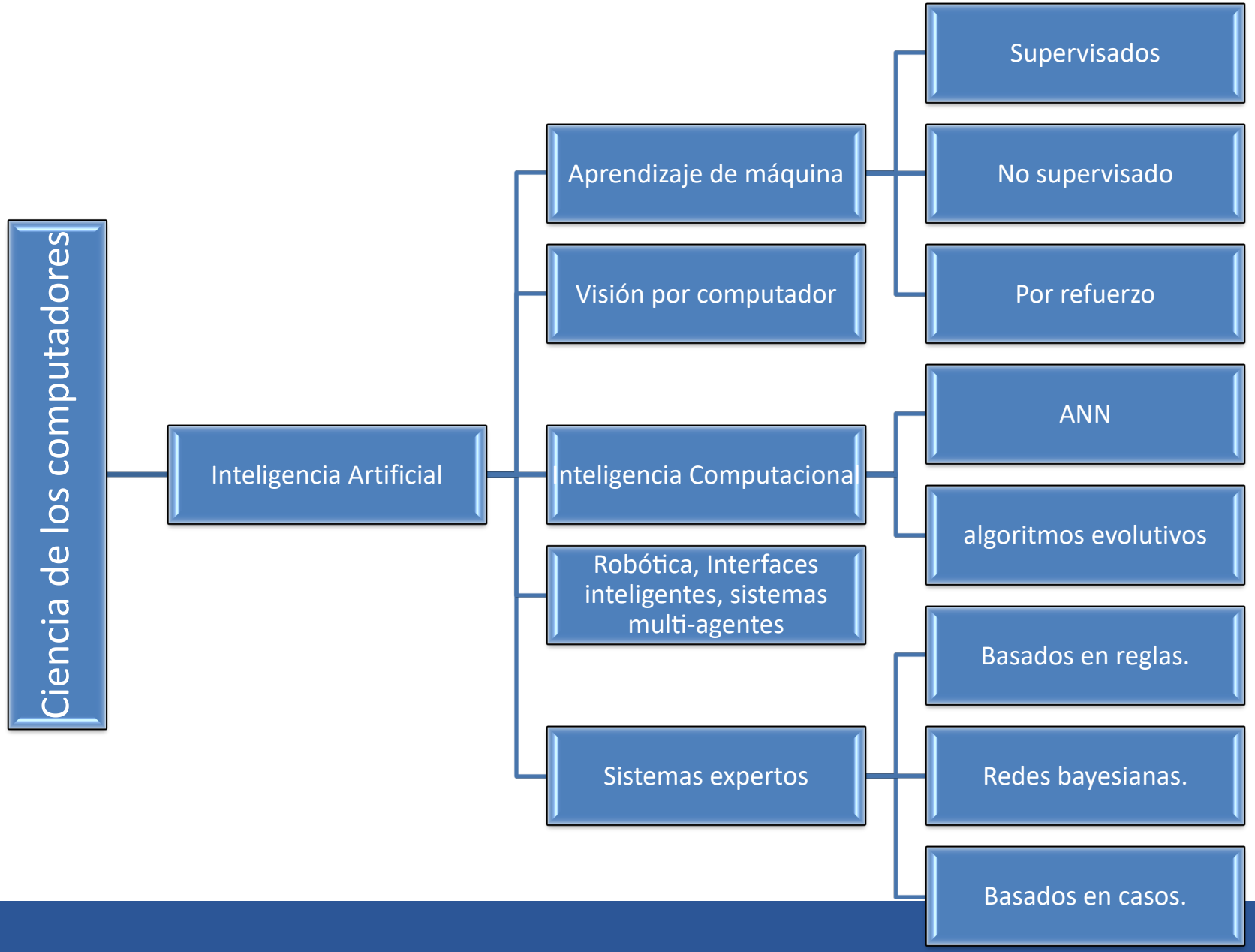
Regresión lineal univariada.

Pontificia Universidad Javeriana
Francisco Carlos Calderon Ph.D
2020


Ciencia de los computadores
Inteligencia Artificial
Aprendizaje de máquina
Visión por computador
Inteligencia Computacional
Robótica, Interfaces inteligentes, sistemas multi-agentes
Sistemas expertos
Supervisados
No supervisado
Por refuerzo
ANN
algoritmos evolutivos
Basados en reglas.
Redes bayesianas.
Basados en casos.

Aprendizaje de máquina, clasificación general

- Supervisados
  - Crean un modelo matemático que busca explicar unas "**etiquetas**" de entrada/salida a partir de un conjunto de "**características**" de entrada.
  - Se pueden dividir principalmente en:
    - Clasificación
    - Regresión
  - Existen otros sub-métodos como:
    - Aprendizaje activo.
    - "Similarity learning"
    - Recommender systems

No Supervisados

- Crean un modelo que busca explicar las **características** de entrada sin contar con etiquetas.
- Se pueden dividir en
  - Agrupamiento. "clustering"
  - Estimación de densidad (pdf).
  - Reducción dimensional

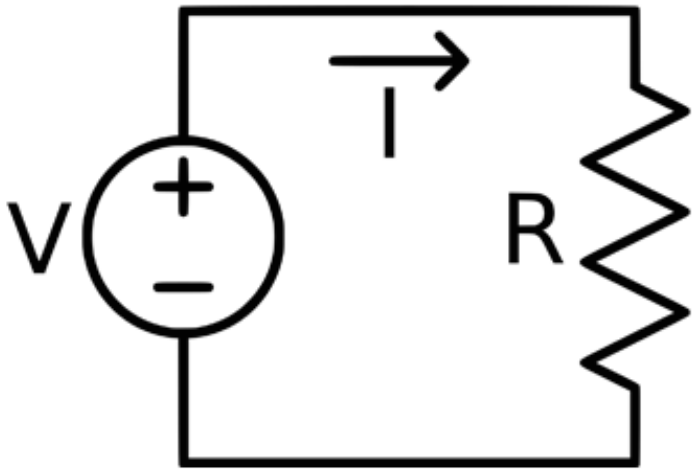
Regresión

Es un proceso estadístico para estimar las relaciones entre variables.

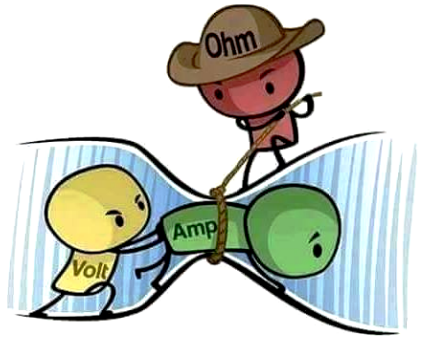
Nop, no es la película.

Caso de ejemplo

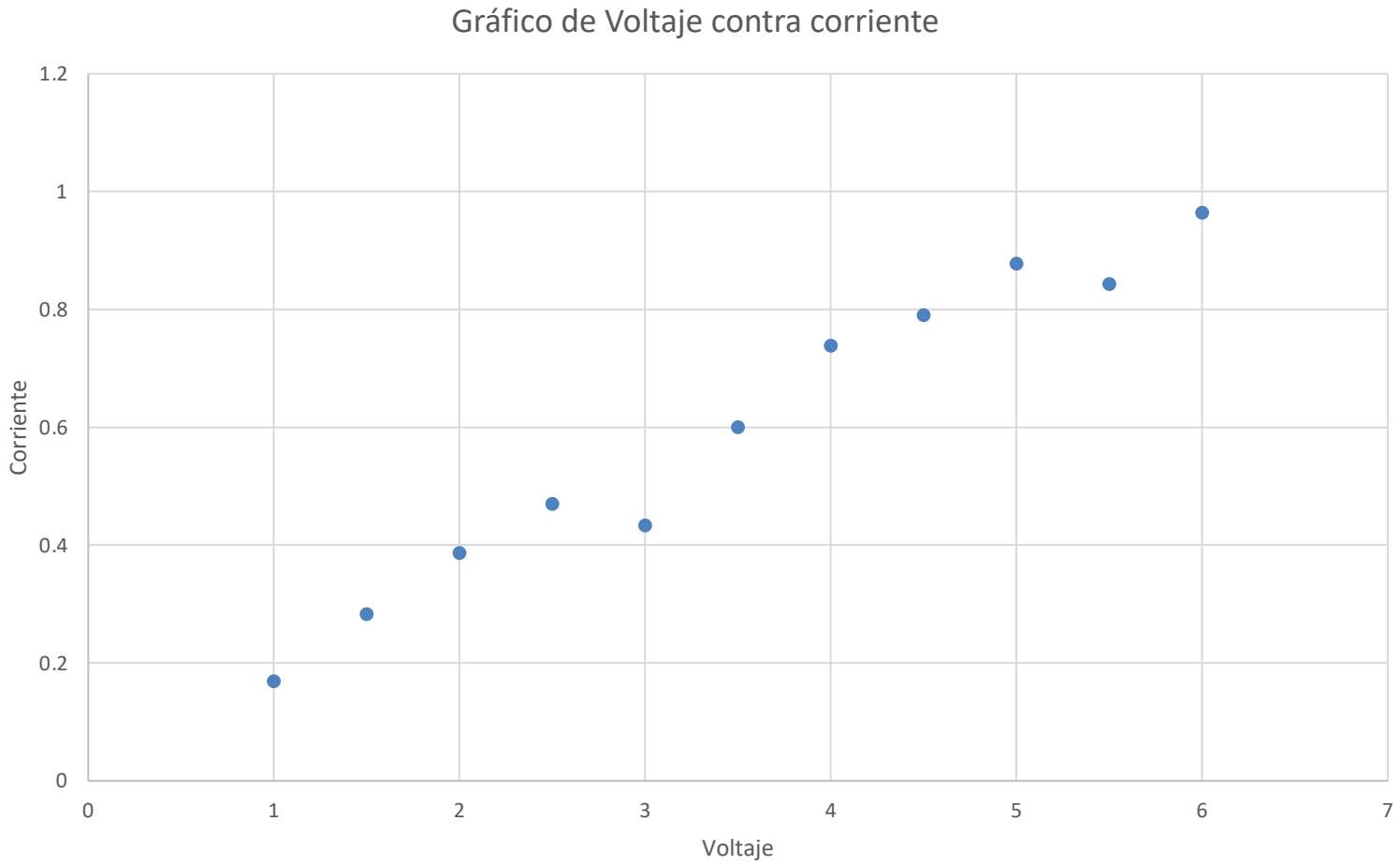
R=V/I



Voltaje (V)	Corriente (A)
1	0.16023453
1.5	0.27728321
2	0.36117187
2.5	0.48025391
3	0.44229119
3.5	0.59856803
4	0.79987169
4.5	0.82492895
5	0.79536605
5.5	0.87930235
6	0.90780985



Caso de ejemplo



Notación:

$$IR = V$$

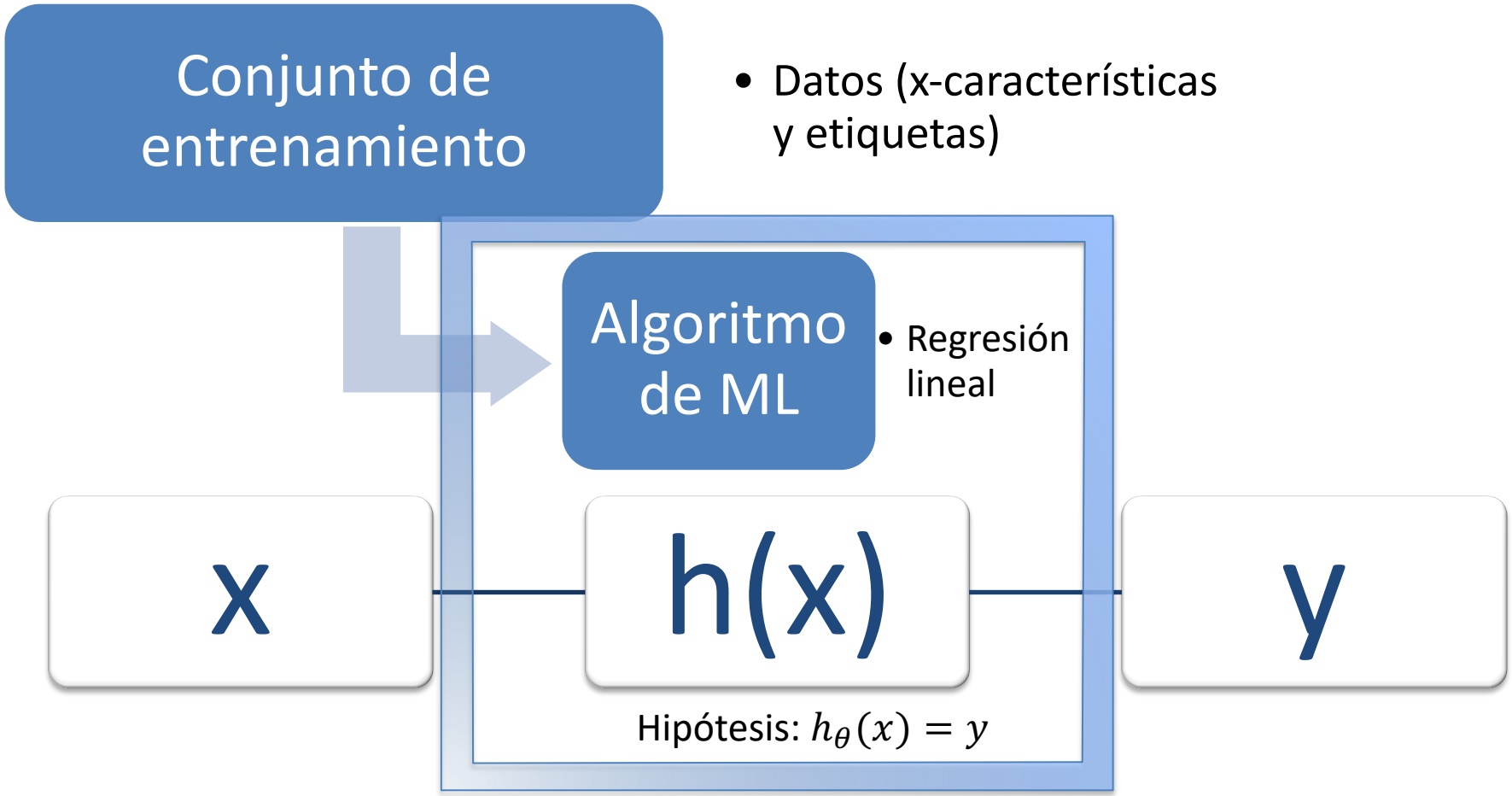
$I = \left(\frac{1}{R}\right)V + I_{ruido}$, donde I_{ruido} tiende a cero

$$y = \left(\frac{1}{R}\right)x + \theta$$

Si tenemos n mediciones

$(x^{(i)}, y^{(i)})$ corresponde a la i-ésima muestra

Idea de regresión



Para una regresión lineal

Definimos nuestra hipótesis como:

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

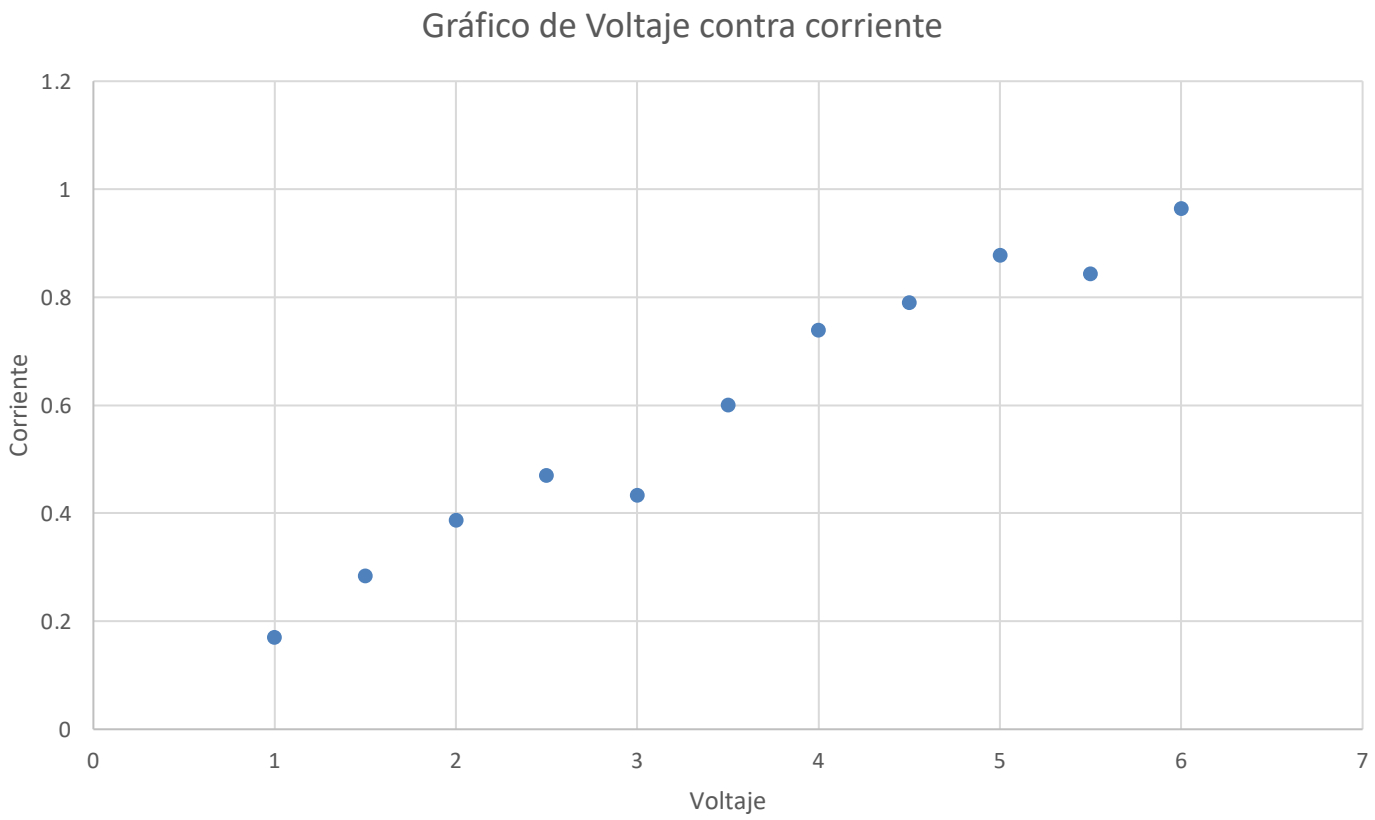
Donde θ_i se denominan los parámetros* del modelo.

* Si los mismos parámetros en los que usamos una estadística para estimarlos

¿Qué son los parámetros en nuestro caso?

Diferentes valores de parámetros representan en nuestra regresión lineal diferentes líneas.

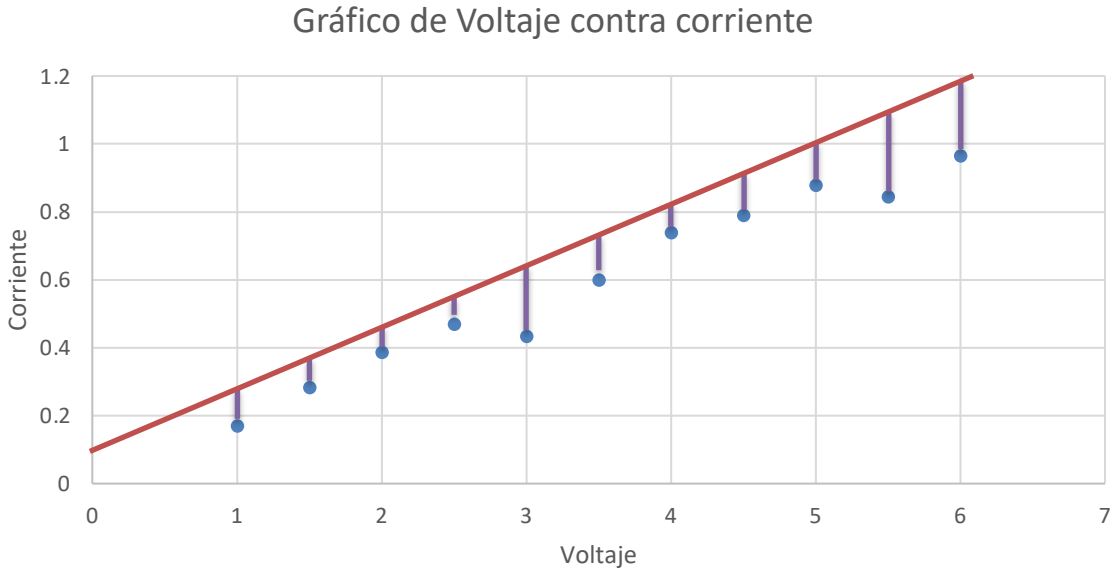
Debemos crear ahora una métrica para evaluar estos conjuntos de parámetros



Métrica de evaluación de parámetros

Definimos una función de costo $J(\theta_0, \theta_1)$.

IDEA: Definir y evaluar los (θ_0, θ_1) de tal manera que se minimize alguna métrica de error que ***tenga sentido en el problema***.



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

$$\min_{\theta_0,\theta_1} \sum_{i=1}^{n} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2$$

$$\min_{\theta_0,\theta_1} \frac{1}{n} \sum_{i=1}^{n} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2$$

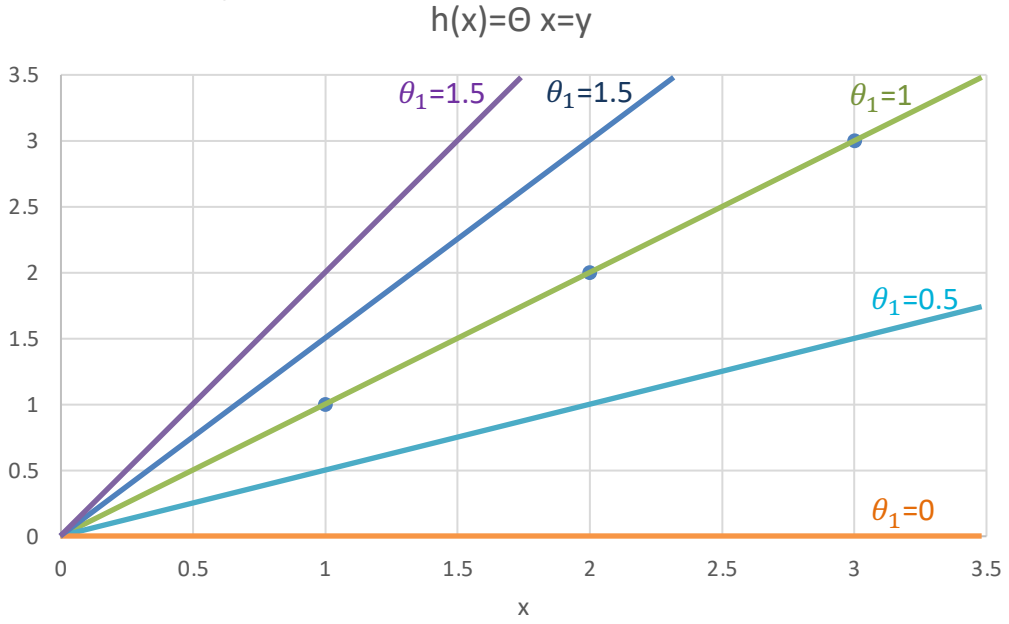
$$\min_{\theta_0,\theta_1} \frac{1}{2n} \sum_{i=1}^{n} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2 = \min_{\theta_0,\theta_1} J(\theta_0, \theta_1)$$

Caso de ejemplo suponiendo un solo parámetro

Si suponemos $h_\theta(x) = \theta_1 x = y$

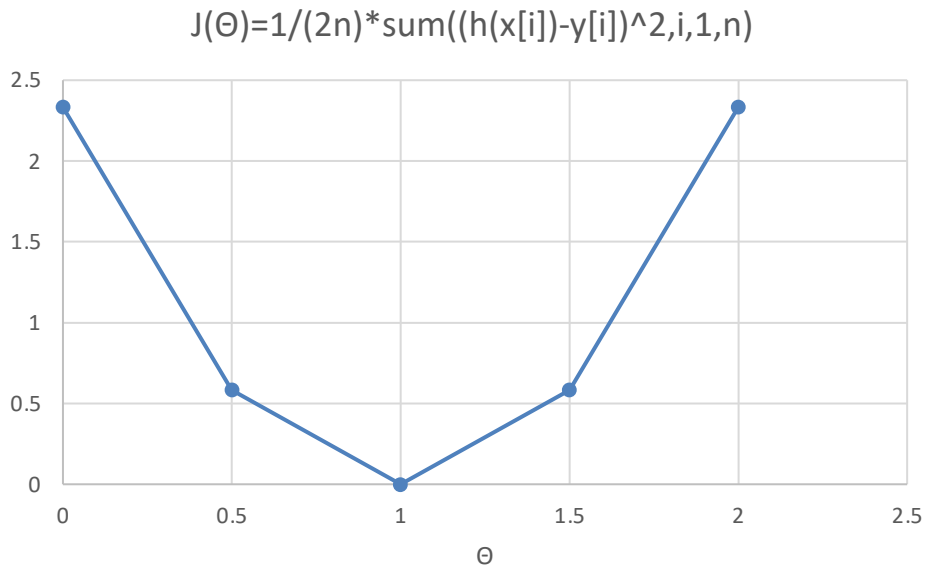
si n son 3 puntos:

Para un θ_1 fijo, es una función de x



Podemos evaluar cada parámetro θ_1 con $J(\theta_1)$

J es una función de los parámetros.



Ver el suplemento en excel

Caso general regresión lineal univariada

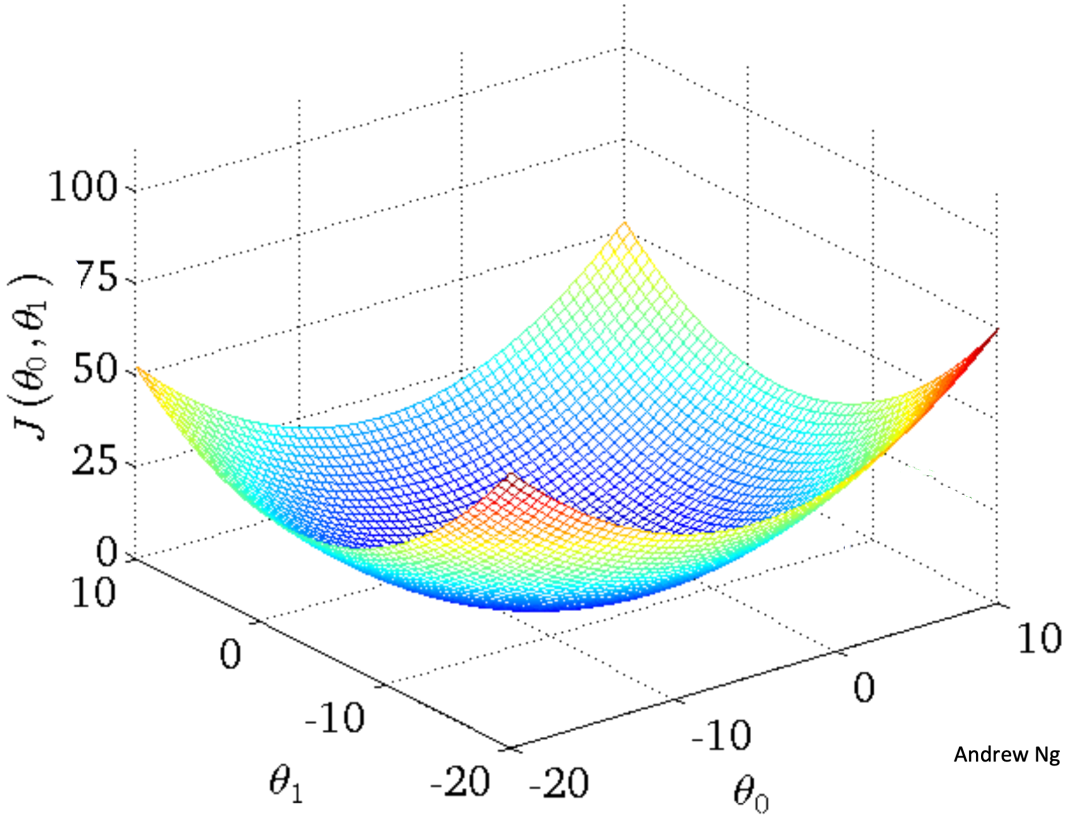


Imagen propiedad de Andrew Ng

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

$$\mathrm{J}(\theta_0,\theta_1) = \frac{1}{2n}\sum_{i=1}^{n}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2$$

$$\mathrm{J}(\theta_0,\theta_1) = \frac{1}{2n}\sum_{i=1}^{n}\left(\left(\theta_0 + \theta_1 x^{(i)}\right) - y^{(i)}\right)^2$$

Y como encuentro ese mínimo?

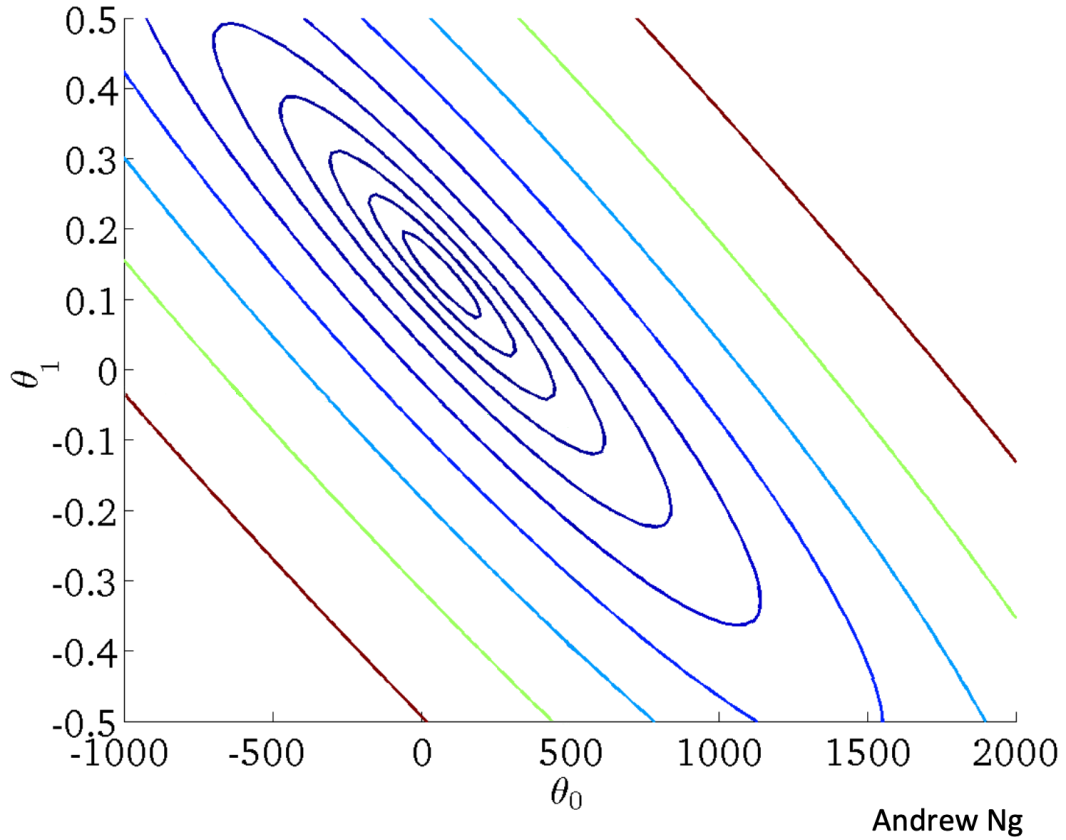


Imagen propiedad de Andrew Ng

Vamos a implementar en python una técnica llamada “grid search”.

También conocida como búsqueda a fuerza bruta.

https://en.wikipedia.org/wiki/Brute-force_search

Ejercicio para el resto de la clase

1. Implemente la función de costo anterior teniendo 3 entradas:
    1. Un vector “nparray” para los parámetros (θ_0, θ_1)
    2. Un vector “nparray” para las características “x”
    3. Un vector “nparray” para las etiquetas “y”

    La salida es un flotante.

https://numpy.org/doc/stable/reference/generated/numpy.array.html

Ejercicio para el resto de la clase

2. Implemente en dos bucles que busquen un mínimo en la función J, uno para cada parámetro.
3. Experimente con diferentes intervalos de búsqueda.
4. Cómo mejoraría este proceso?