Presentación AA-1-2

LIA
IA PARA LA HUMANIDAD

Lección-1

Material preparado por: Daniel Boza

Descenso de gradiante para regresión mutiple

	Anotación Anterior	Notación de vector
parametros	$w_1, \cdots, w_n$ b	$\vec{w} = [w_1 \quad \cdots \quad w_n]$ b
modelo	$f_{\vec{w},b}(\vec{x}) = w_1 x_1 + \cdots + w_n x_n + b$	$f_{\vec{w},b}(\vec{x}) = \vec{w} \cdot \vec{x} + b$
función de costos	$J(w_1, \cdots, w_n, b)$	$J(\vec{w}\ b)$

Descenso de Gradiante

repetir {

$$w_j = w_j - \alpha \frac{\partial}{\partial w_j} J(w_1, \cdots, w_n, b)$$

$$b = b - \alpha \frac{\partial}{\partial b} J(w_1, \cdots, w_n, b)$$

}

repetir {

$$w_j = w_j - \alpha \frac{\partial}{\partial w_j} J(\vec{w}, b)$$

$$b = b - \alpha \frac{\partial}{\partial b} J(\vec{w}, b)$$

}

Gradiante de descenso

repetir { Una característica

$$\underbrace{w}_{} = w - \alpha \boxed{\frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}\left(x^{(i)}\right) - y^{(i)}\right)x^{(i)}}$$

$$\frac{\partial}{\partial w}J(w,b)$$

$$b = b - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}\left(x^{(i)}\right) - y^{(i)}\right)$$

Actualizar simultaneamente w, b

}

n característica $(n \geq 2)$

repetir {

$$w_1 = w_1 - \alpha \boxed{\frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{\mathrm{w}},b}\left(\vec{\mathrm{x}}^{(i)}\right) - y^{(i)}\right)x_1^{(i)}}$$

$$\vdots \qquad \frac{\partial}{\partial w_1}J(\vec{\mathrm{w}},b)$$

$$w_n = w_n - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{\mathrm{w}},b}\left(\vec{\mathrm{x}}^{(i)}\right) - y^{(i)}\right)x_n^{(i)}$$

$$b = b - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{\mathrm{w}},b}\left(\vec{\mathrm{x}}^{(i)}\right) - y^{(i)}\right)$$

Actualizar simultaneamente

w_j (for $j = 1, \cdots, n$) and b

}

Presentación AA-1-2

Lección-2

Escalamiento de caracteristicas

Material preparado por: Daniel Boza

Múltiples características (variables)

Tamaño y alimentación	Número de dormitorios	Número de pisos	Edad de la casa en años	Precio ($) en $1000
x1	**x2**	**x3**	**x4**	
2104	5	1	45	460
1416	3	2	40	232
1534	3	2	30	315
852	2	1	36	178
...	...	...	...	...

- $x_j = j^{th}$ característica j
- n = cantidad de características
- $\vec{x}^{(i)}$ = caracteristícas del ejemplo de entrenamiento i^{th}
- $x_j^{(i)}$ = valor de la característica j en el ejemplo de entrenamiento i^{th}

Modelo

Previamente: $f_{w,b}(x) = wx + b$

$$f_{w,b}(x) = w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 + b$$

$$f_{w,b}(x) = 0.1\, x_1 + 4 x_2 + 10\, x_3 + -2\, x_4 + 80$$

Tamaño # de cuartos pisos años precio base

$$f_{w,b}(x) = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b$$

$$f_{\vec{w},b}(\vec{x}) = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b$$

$$\vec{w} = [w_1 \;\; w_2 \;\; w_3 \; \ldots \; w_n]$$

b is a number

vector $\vec{x} = [x_1 \;\; x_2 \;\; x_3 \; \ldots \; x_n]$

$$f_{\vec{w},b}(\vec{x}) = \vec{w} \cdot \vec{x} + b = w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n + b$$

regresión lineal multiple

PARAMETROS Y CARACTERISTICAS

$\vec{w} = [w_1 \quad w_2 \quad w_3]$

b is a number

$\vec{x} = [x_1 \quad x_2 \quad x_3]$

```
w = np.array([1.0,2.5,-3.3])
b = 4
x = np.array([10,20,30])
```

SIN VECTORIZACIÓN

$$f_{\vec{w},b}(\vec{x}) = \left(\sum_{j=1}^{n} w_j x_j\right) + b$$

```
f = 0
for j in range(0,n):
    f = f + w[j] * x[j]
f = f + b
```

SIN VECTORIZACIÓN

$$f_{\vec{w},b}(\vec{x}) = w_1x_1 + w_2x_2 + w_3x_3 + b$$

```
f = w[0] * x[0] +
    w[1] * x[1] +
    w[2] * x[2] + b
```

VECTORIZACIÓN

$$f_{\vec{w},b}(\vec{x}) = \vec{w} \cdot \vec{x} + b$$

```
f = np.dot(w,x) + b
```

SIN VECTORIZACIÓN

```
for j in range(0,16):
    f = f + w[j] * x[j]
```

t_0

```
f + w[0] * x[0]
```

t_1

```
f + w[1] * x[1]
```

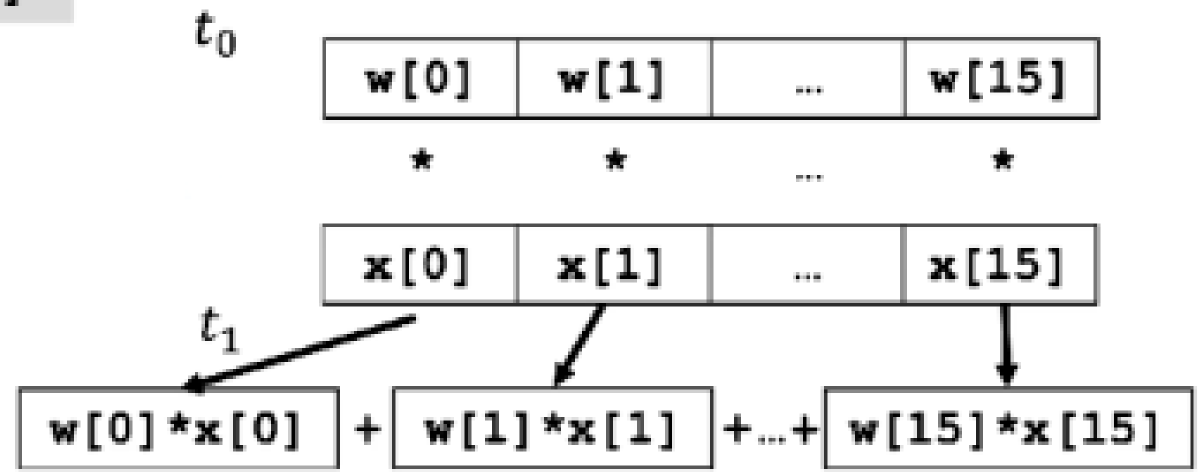
...

t_{15}

```
f + w[15] * x[15]
```

CON VECTORIZACIÓN

```
np.dot(w,x)
```



DESCENSO DE GRADIANTE

$$\vec{w} = (w_1 \quad w_2 \quad \dots \quad w_{16})$$

$$\vec{d} = (d_1 \quad d_2 \quad \dots \quad d_{16})$$

```
w = np.array([0.5, 1.3, … 3.4])
d = np.array([0.3, 0.2, … 0.4])
```

Calcular $w_j = w_j - 0.1 d_j \text{ for } j = 1 \dots 16$

SIN VECTORIZACIÓN

$$w_1 = w_1 - 0.1 d_1$$
$$w_2 = w_2 - 0.1 d_2$$
$$\vdots$$
$$w_{16} = w_{16} - 0.1 d_{16}$$

```
for j in range(0,16):
    w[j] = w[j] - 0.1 * d[j]
```

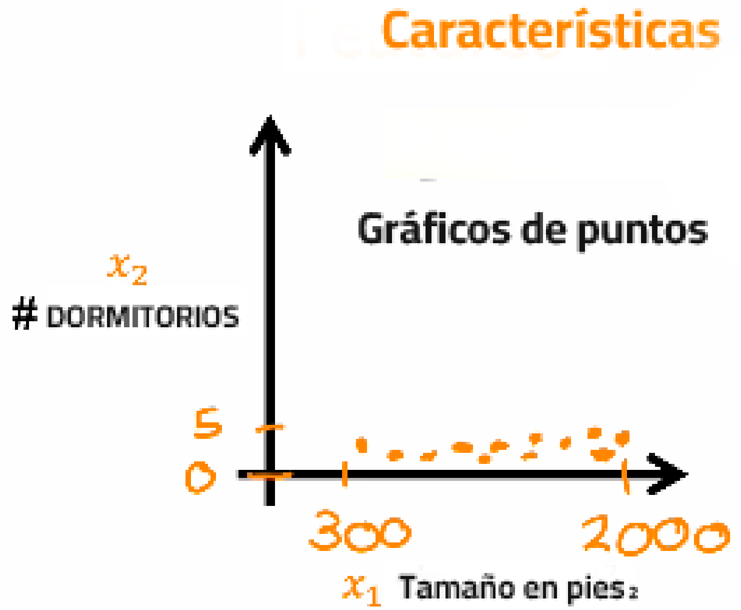
CON VECTORIZACIÓN

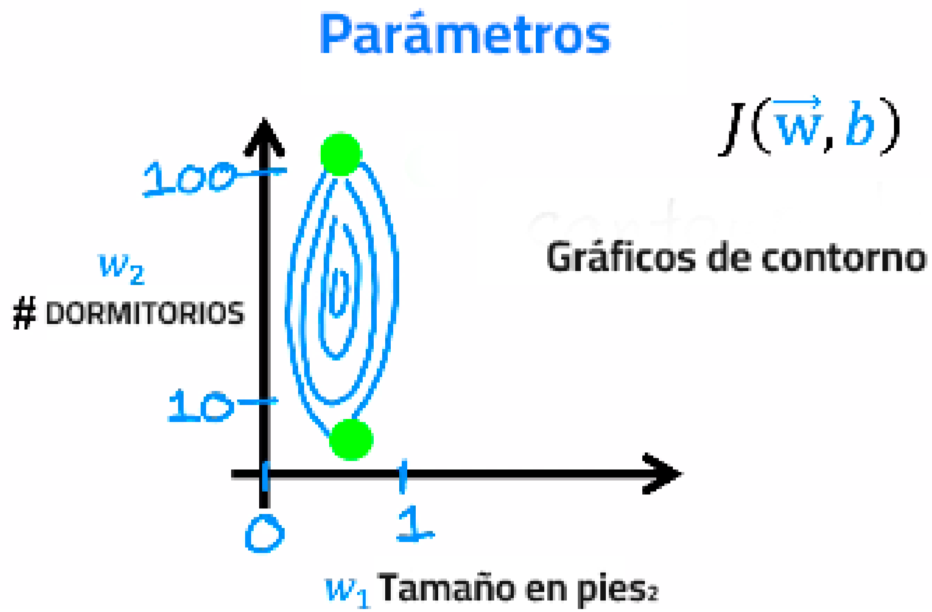
$$\vec{w} = \vec{w} - 0.1 \vec{d}$$

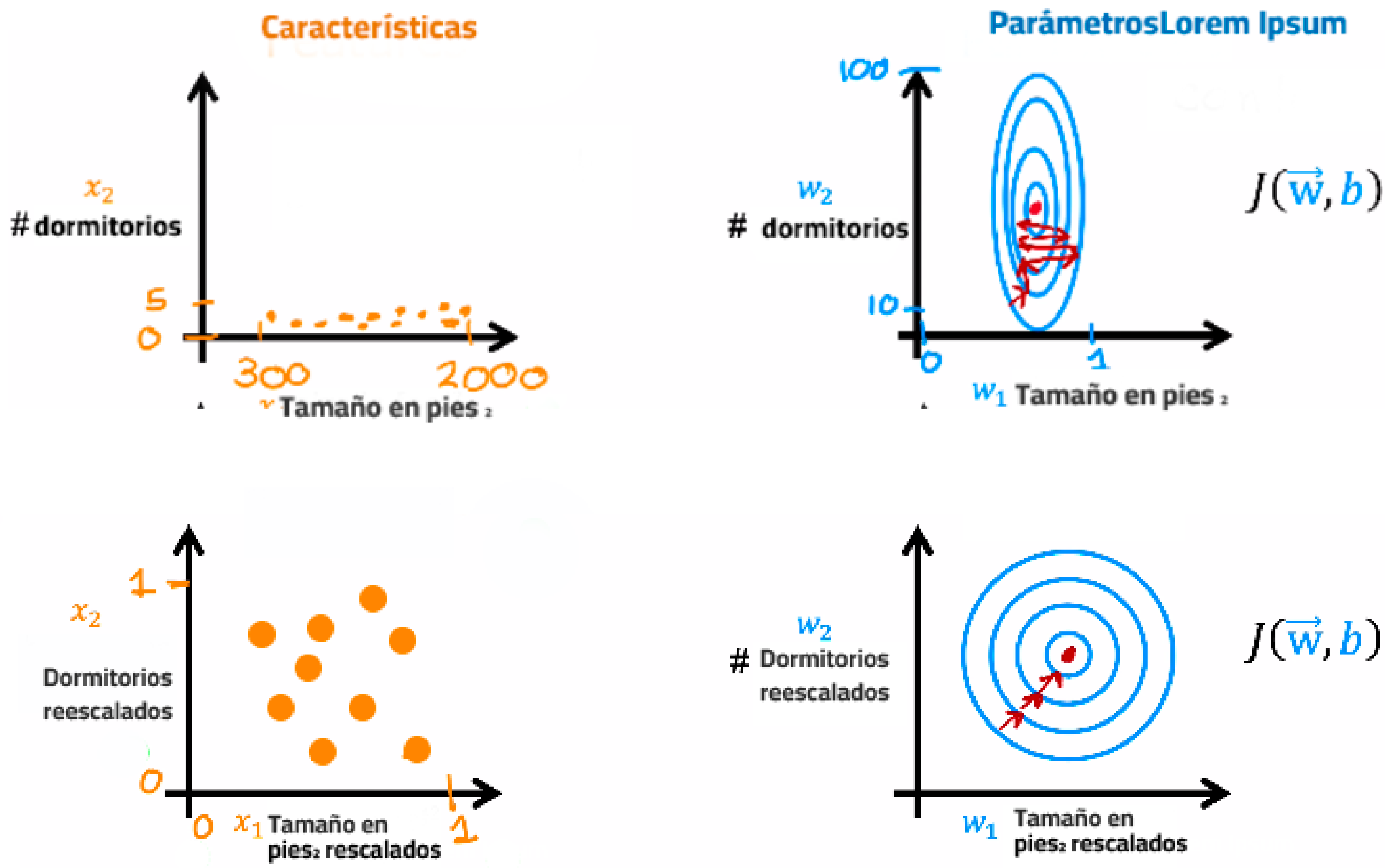
```
w = w - 0.1 * d
```

Tamaño de caracteristica y tamaño de parámeto

	tamaño de la característica x_j	tamaño de parámetro w_j
Tamaño en pies₂	⟷	⟷
# Dormitorios	⟷	⟷

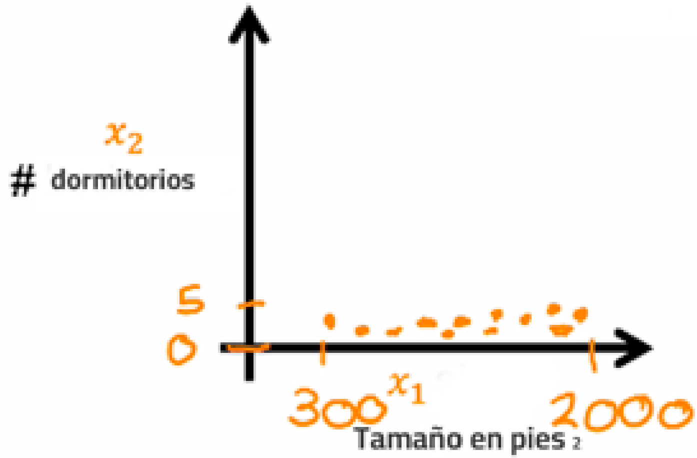


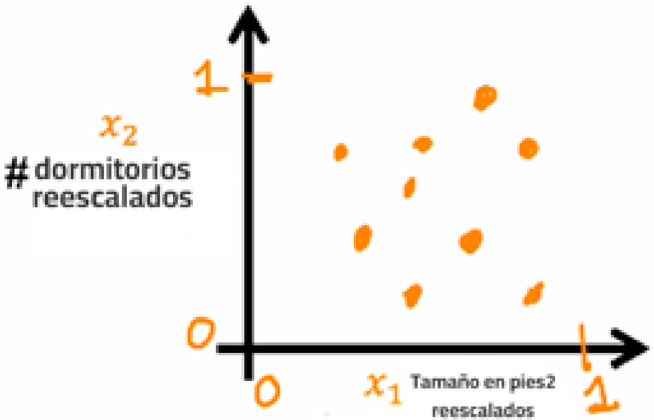




Caracteristicas
ParámetrosLorem Ipsum
x_2
dormitorios
5
0
300
2000
Tamaño en pies 2
100
10
0
1
w_2
dormitorios
w_1 Tamaño en pies 2
$J(\vec{w}, b)$
1
x_2
Dormitorios
reescalados
0
0
x_1 Tamaño en
pies2 rescalados
1
w_2
Dormitorios
reescalados
w_1 Tamaño en
pies2 rescalados
$J(\vec{w}, b)$

Escalado de características



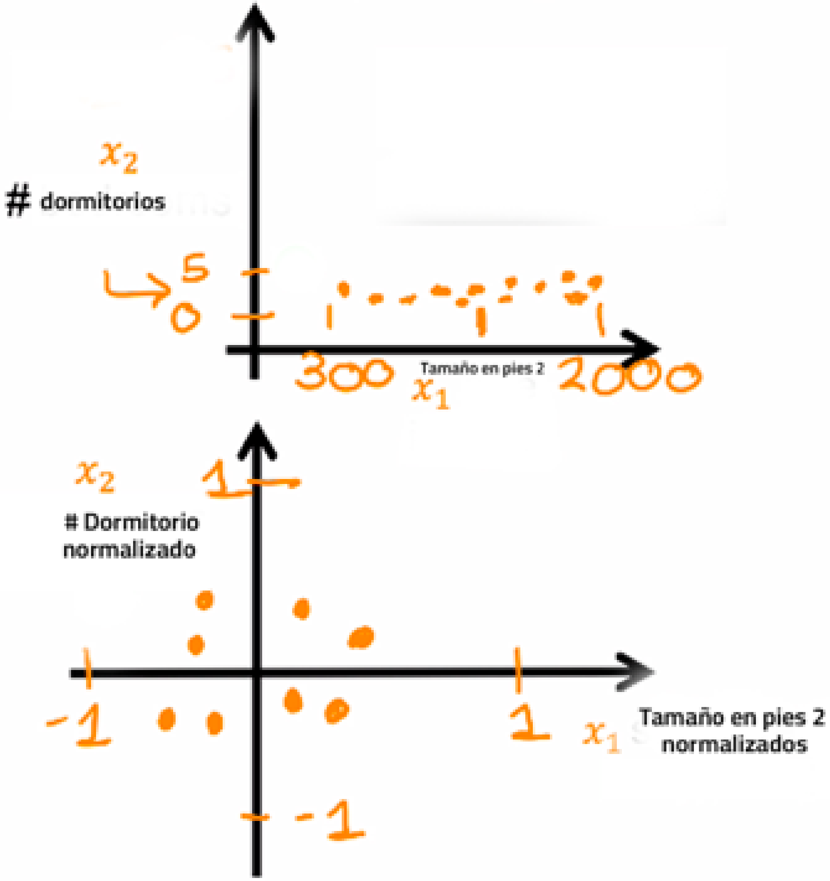


$$300 \leq x_1 \leq 2000 \qquad 0 \leq x_2 \leq 5$$

$$x_{1,scaled} = \frac{x_1}{2000} \qquad x_{2,scaled} = \frac{x_2}{5}$$

$$0.15 \leq x_{1,scaled} \leq 1 \qquad 0 \leq x_{2,scaled} \leq 1$$

Normalización de la media

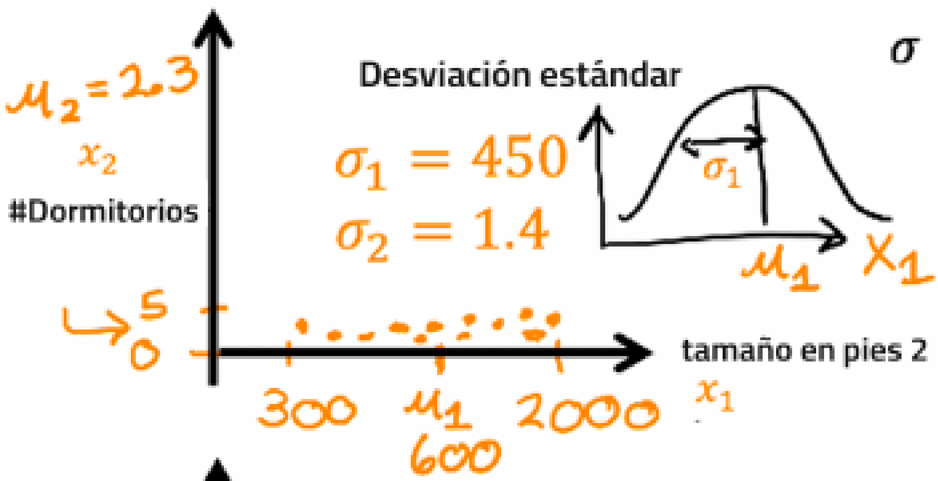


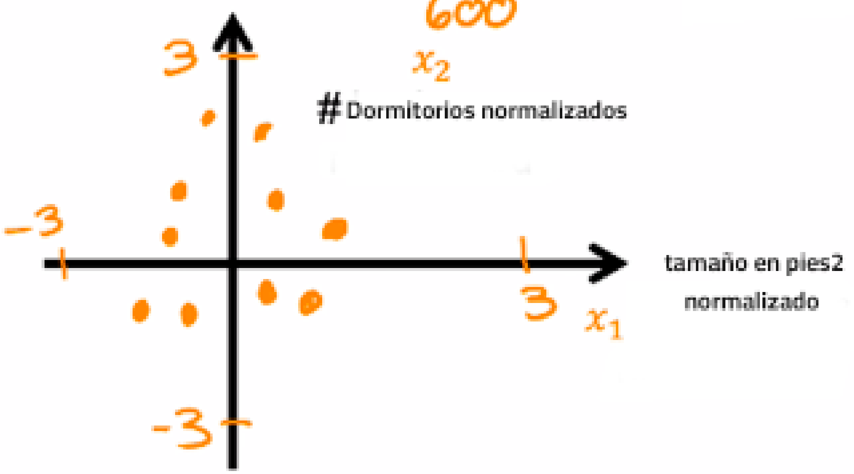
$$300 \le x_1 \le 2000 \qquad 0 \le x_2 \le 5$$

$$x_1 = \frac{x_1 - \mu_1}{2000-300} \qquad x_2 = \frac{x_2 - \mu_2}{5-0}$$

$$-0.18 \le x_1 \le 0.82 \qquad -0.46 \le x_2 \le 0.54$$

Normalización Z-score





$$300 \le x_1 \le 2000 \qquad 0 \le x_2 \le 5$$

$$x_1 = \frac{x_1 - \mu_1}{\sigma_1} \qquad x_2 = \frac{x_2 - \mu_2}{\sigma_2}$$

$$-0.67 \le x_1 \le 3.1 \quad -1.6 \le x_2 \le 1.9$$

Redimensionar características

Apunte aproximadamente $-1 \leq x_j \leq 1$ para cada característica x_j

$$-3 \leq x_j \leq 3$$

$$-0.3 \leq x_j \leq 0.3$$

rangos aceptables

$0 \leq x_1 \leq 3$ okey, no redimensionar

$-2 \leq x_2 \leq 0.5$ okey, no redimensionar

$-100 \leq x_3 \leq 100$ demasiado largo → redimensionar

$-0.001 \leq x_4 \leq 0.001$ demasiado pequeño → redimensionar

$98.6 \leq x_5 \leq 105$ demasiado largo → redimensionar

Presentación AA-1-2

Lección-3

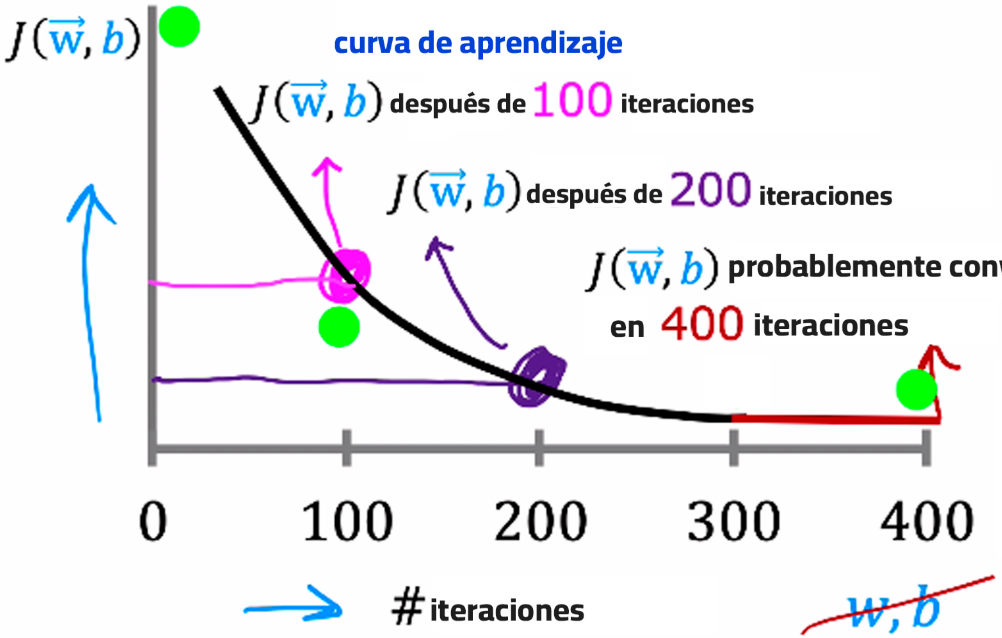
Tips prácticos para regresión lineal

Material preparado por: Daniel Boza

Asegúrese que el desenso de gradiante funcione correctamente

Objetivos : $\min_{\vec{w},b} J(\vec{w}, b)$

$J(\vec{w}, b)$ **debería disminuir después de cada iteración.**



~~w, b~~

 # Las iteraciones necesarias varian 30 1,000 100,000

Prueba de convergencia automática

Siendo ε "epsilon" 10^{-3}. 0.001

Sí $J(\vec{w}, b)$ **disminuye en** $\leq \varepsilon$ **una interación, declara convergencia**

Presentación AA-1-2

Lección-4

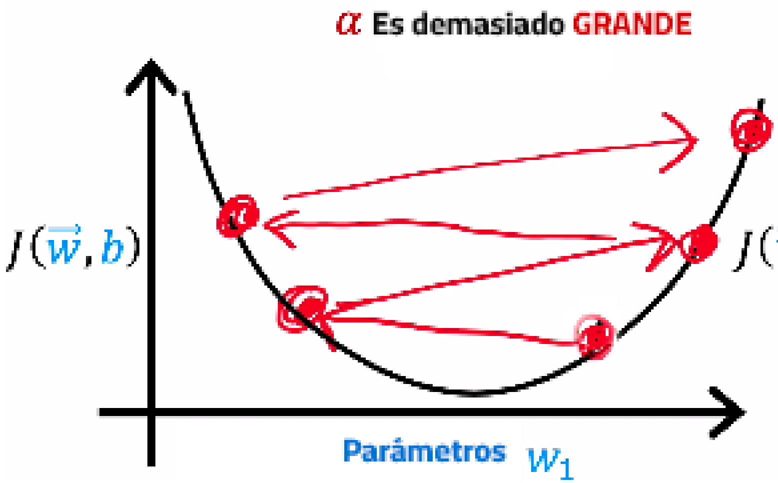
Escogiendo la tasa de aprendizaje

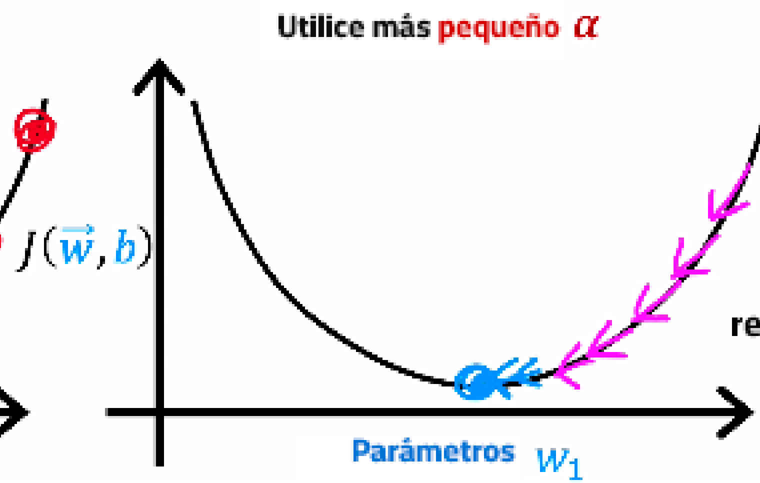
Material preparado por: Daniel Boza

Identificar el problema con el descenso de gradiante

La tasa de aprendizaje es demasiado grande

Ajustar tasa de aprendizaje





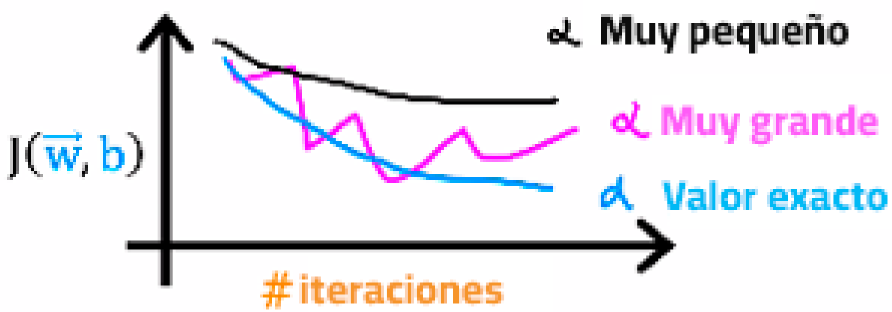
Si es suficientemente pequeño α, $J(\vec{w}, b)$ debería disminuir en cada iteración.

Si α es demasiado pequeño, el descenso del gradiante requiere muchas más itereaciones para converger.

Valores de α a intentar:

... 0.001 0.003 0.01 0.03 0.1 0.3 1 ...

3X ≈3X 3X ≈3X 3X ≈3X



Presentación AA-1-2

Lección-5

Ingeniería de características

Material preparado por: Daniel Boza

Ingeniería de las características

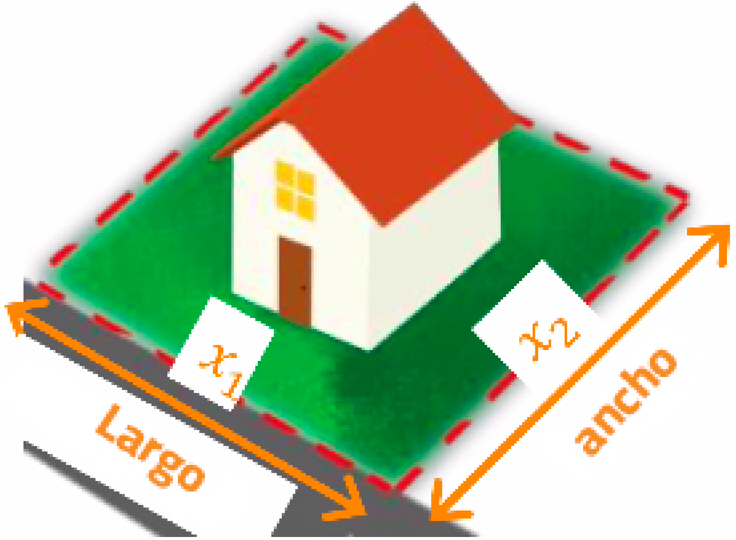
$$f_{\vec{w},b}(\vec{x}) = w_1 \underbrace{x_1}_{\text{Largo}} + w_2 \underbrace{x_2}_{\text{Ancho}} + b$$

$$area = \text{Largo} \times \text{Ancho}$$

$$x_3 = x_1 x_2$$

Nueva característica

$$f_{\vec{w},b}(\vec{x}) = \underbrace{w_1}x_1 + \underbrace{w_2}x_2 + \underbrace{w_3}x_3 + b$$



Ingenierías de la características

Utilizar la intuición para diseñar nuevos elementos transformando o combinando elementos originales.

Presentación AA-1-2

OLIA

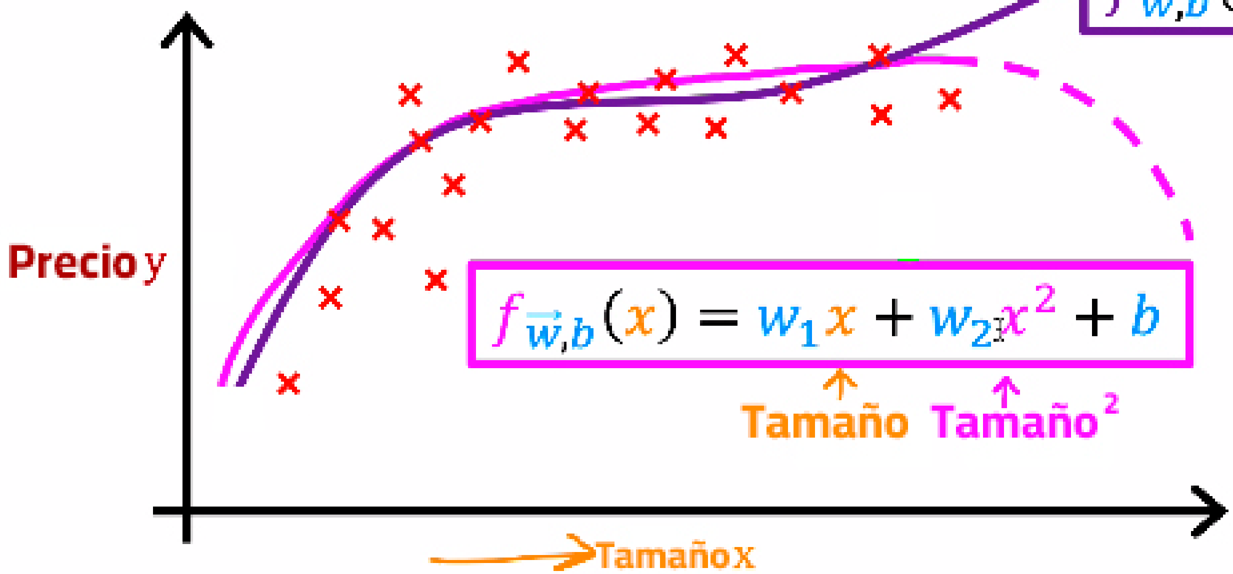
IA PARA LA HUMANIDAD

Lección-6

Regresión Polinomial

Material preparado por: Daniel Boza

Regresión Polinomial



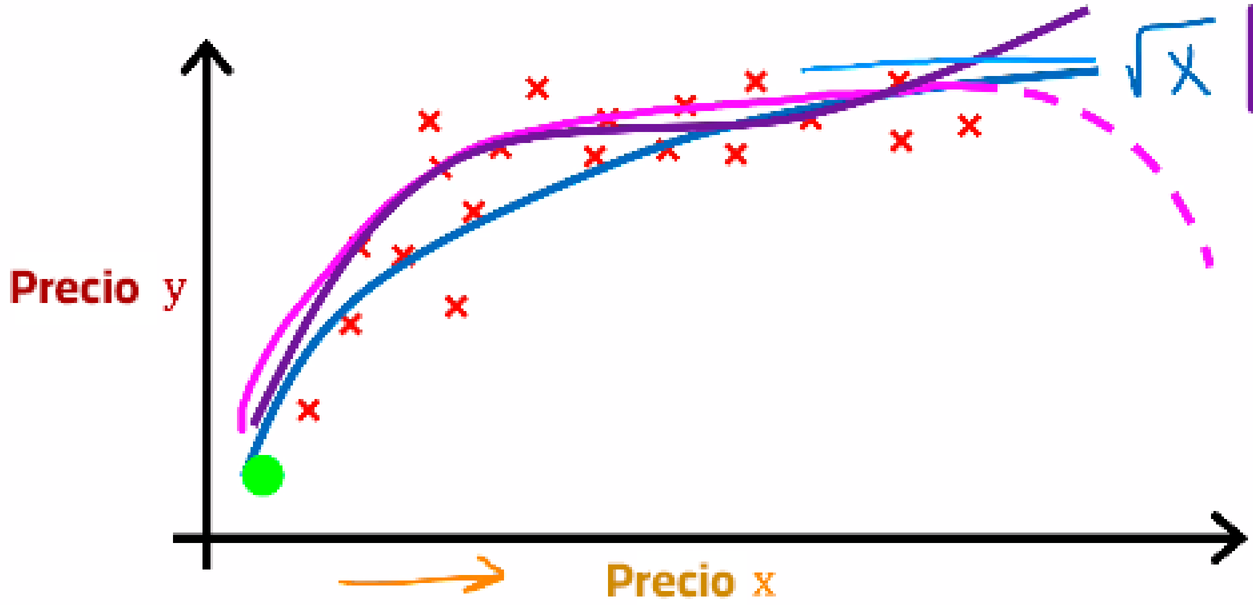
$$f_{\vec{w},b}(x) = w_1 x + w_2 x^2 + w_3 x^3 + b$$

Tamaño Tamaño^2 Tamaño^3

$1-10^3$ $1-10^6$ $1-10^9$

Escalado de características

Escoger características



$$\boxed{f_{\vec{w},b}(x) = w_1 x + w_2 \sqrt{x} + b}$$

tamaño

$\sqrt{\text{tamaño}}$

¿ Qué características usar?

¡Segundo Curso!