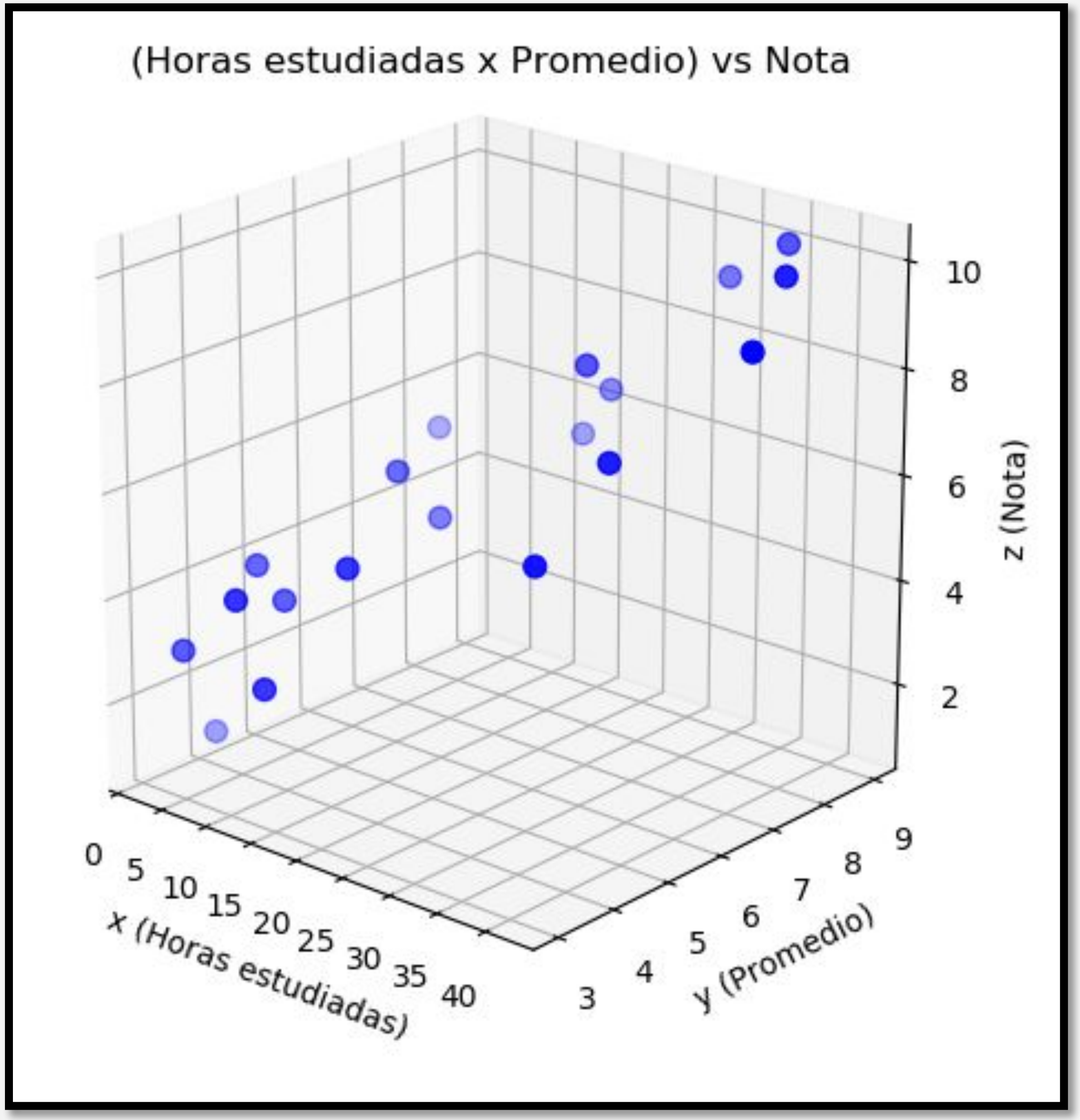
Aprendizaje Automático Profundo (Deep Learning)

Dr. Facundo Quiroga - Dr. Franco Ronchetti

Regresión Lineal con Múltiples Variables

Datos con múltiples variables

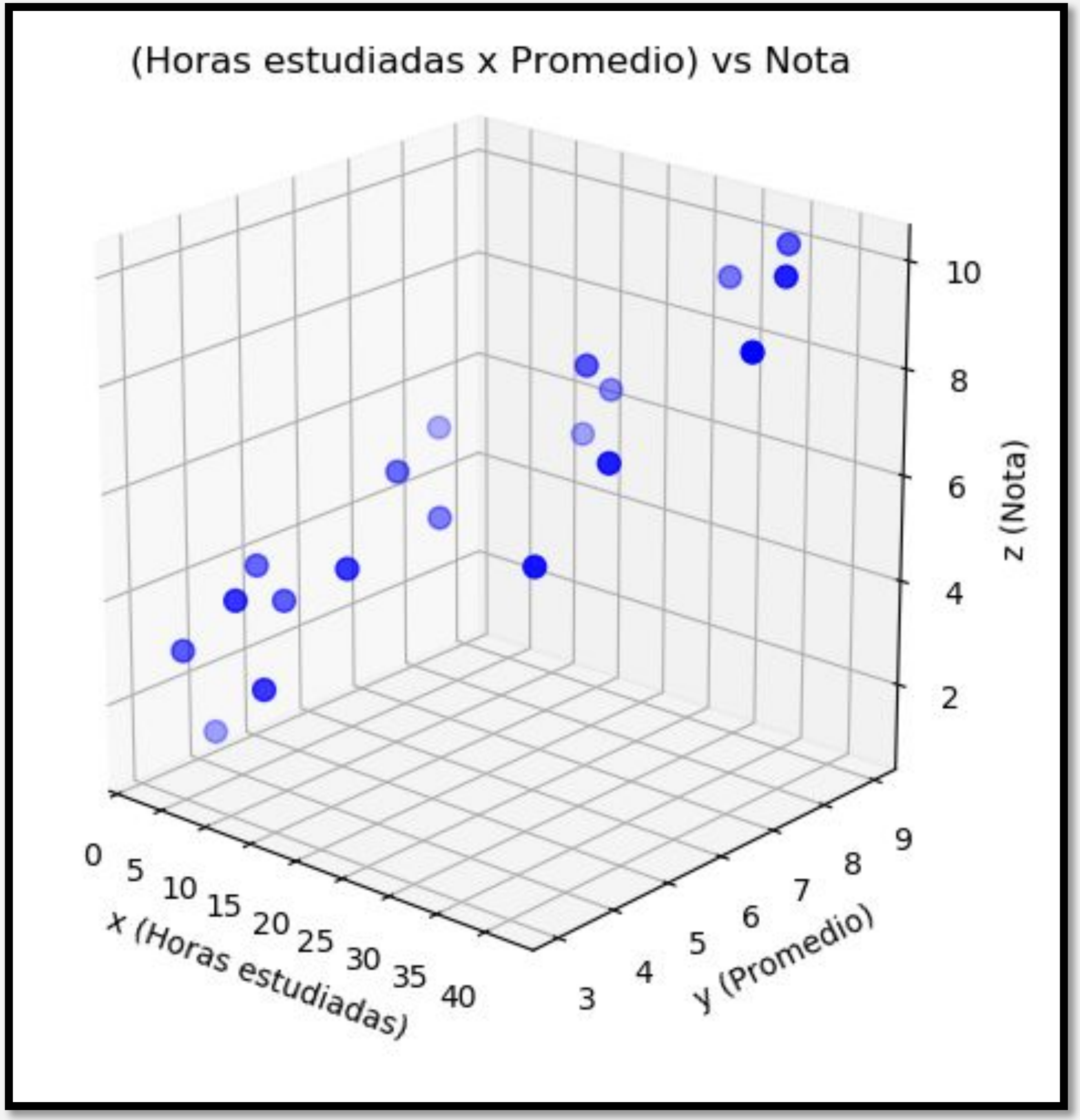
- Hasta ahora
  - Datos con una sola variable
    - 1-D
    - $x \in \mathbf{R}^1$
  - Ejemplo
    - Tiempo de estudio
- ¿Qué sucede si agregamos más información de cada ejemplo?
  - **m** variables
  - **m**-D
  - $\mathbf{x} \in \mathbf{R}^m$
  - Ejemplo
    - **Horas estudiadas** y **promedio**



Datos con 2 variables

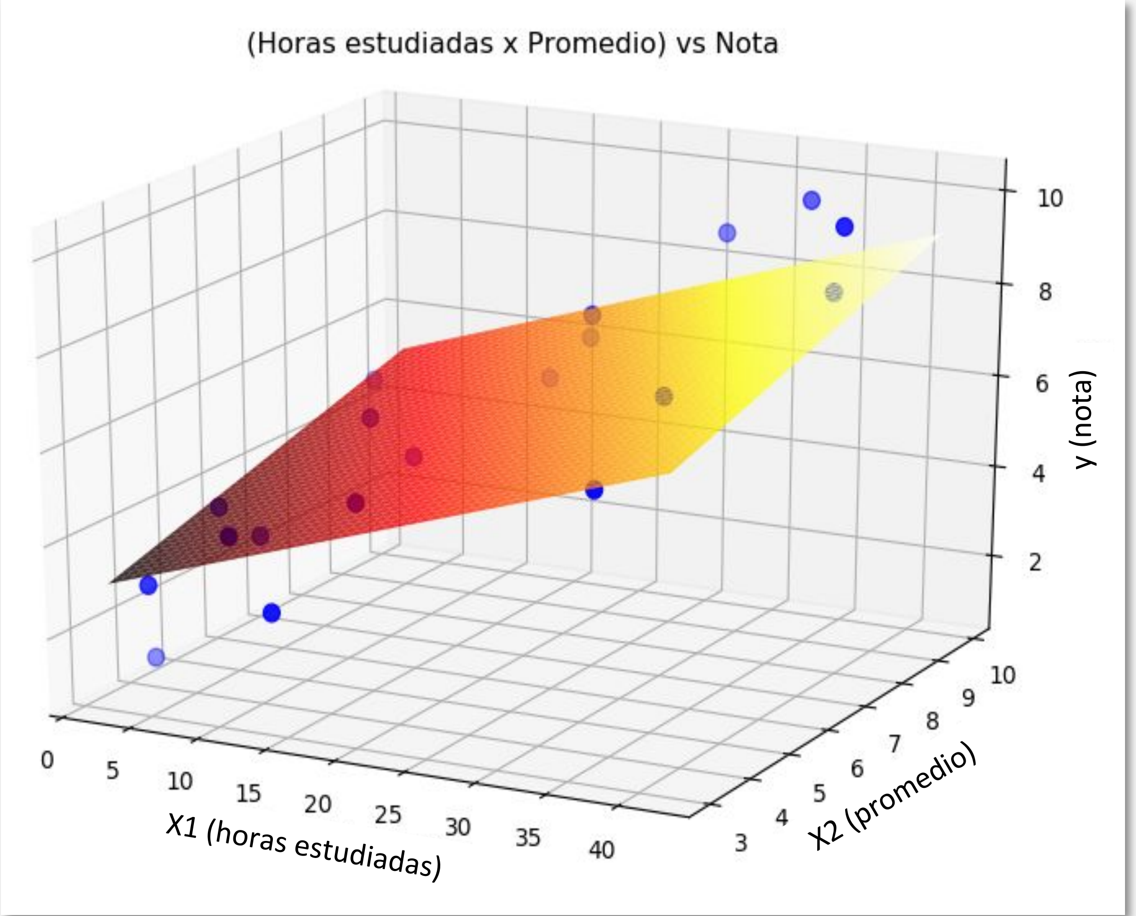
- x con subíndices
  - x_1 variable 1
  - x_2 variable 2

Estudio (x_1)	Promedio (x_2)	Nota (y)
2	4	1
5	3	3,2
7	4	4,5
9	7	6
10	4	4
11	3	4,5
13,4	5	5,5
14	3	3



Función del modelo con 2 variables

- Modelo: **plano 2D** en un **espacio 3D**
- Dos pendientes: w_1 y w_2
  - Coeficientes (o Pesos o **W**eights)



- Función del modelo
  - $f(x_1,x_2) = x_1\,w_1 + x_2\,w_2 + b$
- En forma vectorial
  - $\mathbf{x} = (x_1,x_2)$
  - $\mathbf{w} = (w_1,w_2)$
  - $f(\mathbf{x}) = \mathbf{x} \cdot \mathbf{w} + b$
    - $\cdot$ es el **producto escalar**
    - $\mathbf{x} \cdot \mathbf{w} = x_1w_1+x_2w_2$
      - $\mathbf{x} \cdot \mathbf{w} \in \mathbf{R}$

Función del modelo con 2 variables

- Ejemplo
  - $\mathbf{x} = (-2, 3)$
  - $\mathbf{w} = (4, 2)$
  - $b = 5$
  - $f(-2,3) = -2\, w_1 + 3\, w_2 + b$

    $= -2\ 4 + 3\ \ 2 + 5$

    $= -8+6+5$

    $= 3$
- Forma vectorial
  - $f(\,(-2, 3)\,) = (-2, 3) \cdot \mathbf{w} + b$

    $= (-2, 3) \cdot (4, 2) + 5$

    $= -2 + 5$

    $= 3$

- Los elementos de **w** son **coeficientes o pesos**
  - Indican el peso relativo de cada variable
  - Pesos negativos indican relación negativa
- b término constante
  - sigue siendo la ordenada al origen en 3D

Función de error para 2 variables

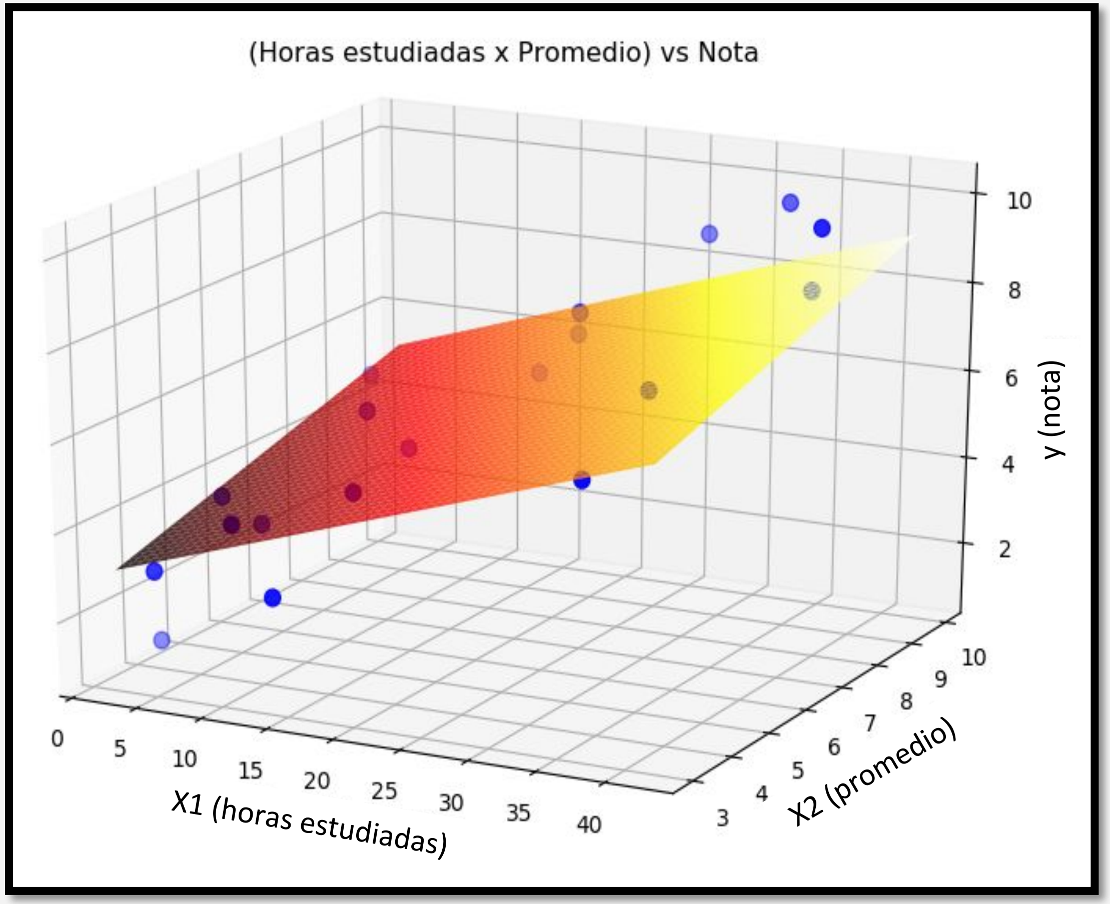
- La forma de **E** se mantiene
  - Depende de 3 parámetros
  - $E(w_1, w_2, b) = 1/n\ \Sigma_i^n\ E_i(w_1, w_2, b)$
- La de $\mathbf{E_i}$ también
  - Pero nueva notación
    - $x_{i,j}$ = variable **j** del ejemplo **i**
  - $E_i(w_1, w_2, b) = (y_i - f(x_1, x_2))^2$

    $= (y_i - (x_{i,1}\ w_1 + x_{i,2}\ w_2 + b))^2$
  - Viendo a $\mathbf{x}_i$ y **w** como vectores
    - $E_i(\mathbf{w}, b) = (y_i - f(\mathbf{x}_i))^2$

      $= (y_i - \mathbf{x}_i \cdot \mathbf{w}\ + b)^2$

Optimización con **1 vs 2** variables

- ¿Qué cambia?
  - Prácticamente nada
  - Ahora tenemos **3** parámetros a optimizar
    - **w_1, w_2, b**
      - También **3** Derivadas
  - Ya no es posible graficar el error
  - Muy difícil “adivinar” los mejores parámetros probando manualmente



Derivadas de **E** con 2 variables

- Derivada parcial respecto de $\mathbf{w_1}$ o $\mathbf{w_2}$ es similar a la de **m**

$$\frac{\partial E}{\partial w_1} = \frac{2}{n}\sum_{i}^{n}(y_i - f(x_i))x_{i,1}$$

$$\Delta E = (\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \frac{\partial E}{\partial b})$$

$$\frac{\partial E}{\partial w_2} = \frac{2}{n}\sum_{i}^{n}(y_i - f(x_i))x_{i,2}$$

$$\frac{\partial E}{\partial b} = \frac{2}{n}\sum_{i}^{n}y_i - f(x_i)$$

Función del modelo con **m** variables

- Modelo: **hiperplano** en $\mathbf{R}^{\mathbf{m}}$
- **m** coeficientes: $\mathbf{w_1}, \mathbf{w_2}, \ldots \mathbf{w_m}$
- Función del modelo
  - $f(x_1,x_2, \ldots, x_m) = x_1 w_1 + x_2 w_2 + \ldots + x_m w_m + b$
- En forma vectorial
  - $\mathbf{x} = (x_1,x_2, \ldots, x_m)$
  - $\mathbf{w} = (w_1,w_2, \ldots, w_m)$
  - $f(\mathbf{x}) = \mathbf{x} \cdot \mathbf{w} + b$
    - $\cdot$ es el **producto escalar**
    - $\mathbf{x} \cdot \mathbf{w} = x_1 w_1 + x_2 w_2 + \ldots + x_m w_m = \sum^{n}_{j=1} x_j w_j$
      - $\mathbf{x} \cdot \mathbf{w} \in \mathbf{R}$

Derivadas de **E** con **m** variables

- Derivada parcial respecto de $\mathbf{w_j}$ (igual para las m variables) y **b**

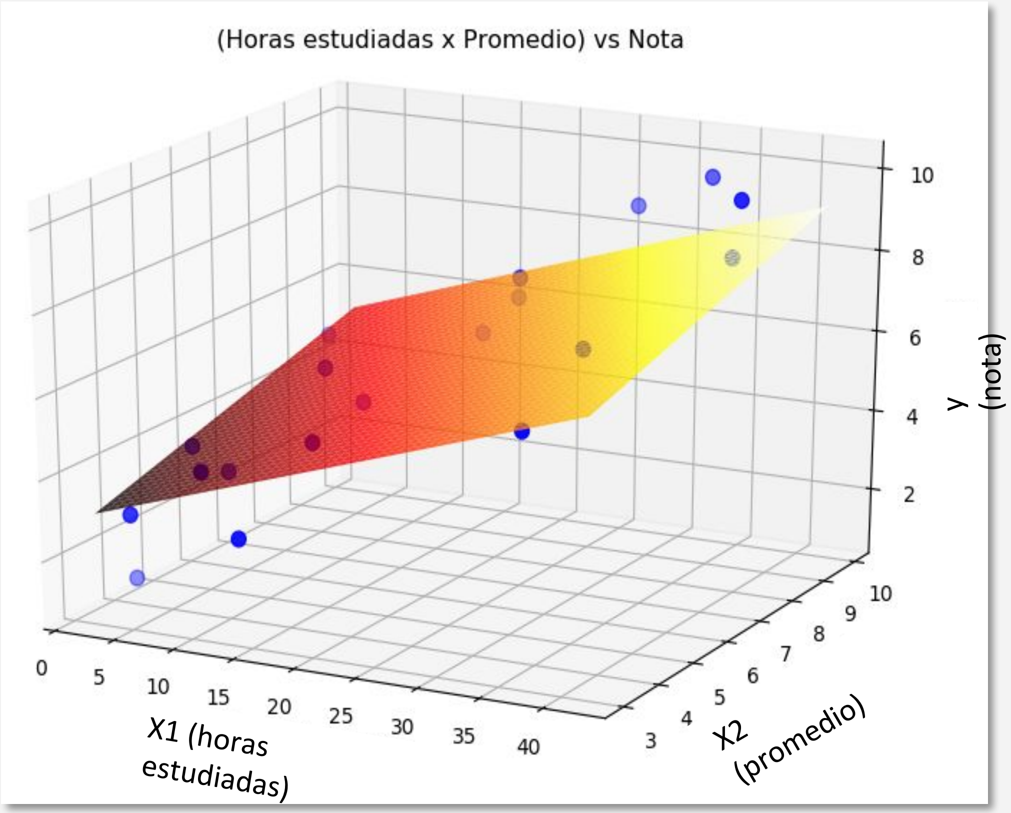
$$\Delta E = \left(\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \ldots, \frac{\partial E}{\partial w_m}, \frac{\partial E}{\partial b}\right)$$

$$\frac{\partial E}{\partial w_j} = \frac{2}{n} \sum_{i}^{n} (y_i - f(x_i)) x_{i,j}$$

$$\frac{\partial E}{\partial b} = \frac{2}{n} \sum_{i}^{n} y_i - f(x_i)$$

Resumen

- Regresioń lineal para **m** variables de entrada
  - Más información de los ejemplos
  - **m** coeficientes o pesos w_1 a w_m
  - No se puede visualizar la función de error
    - Tampoco **f** si m>2
  - Derivadas similares
    - Apto para descenso de gradiente
  - Notación vectorial
    - $f(\mathbf{x}) = \mathbf{x} \cdot \mathbf{w} + b$
    - Producto escalar •



Ejercicio: Archivo **Regresión Lineal Múltiples Variables - Modelo.ipynb**

1. Modificar los parámetros mx , my y b
   a. Observar como cambia el plano y el error

Ejercicio: Archivo **Regresión Lineal Múltiples Variables - Aprendizaje.ipynb**

1. Experimentar con varios valores iniciales de los parámetros y α.

Ejercicio: Archivo **Regresión Lineal Múltiples Variables - Aprendizaje - Boston Housing.ipynb**

1. Entrenar el modelo. Intentar bajar el error lo más posible cambiando la cantidad de iteraciones, el α y los valores iniciales.
2. Luego analizar los valores de los parámetros wi y b . ¿Qué significan?
3. Clasificar nuevos datos, cambiar los valores y observar cómo cambia el precio.
4. ¿Hay variables que no importan? ¿Cómo lo veo desde los valores del nuevo dato? ¿y desde el modelo? ¿cómo se refleja eso en las derivadas?
5. Repetir con Regresion Lineal Multiple -Aprendizaje - Vinos.ipynb.