

# Fundamentos de Aprendizaje Automático y Reconocimiento de Patrones

## Práctico 03

2024

# (Evaluación interactiva)

## Evaluación interactiva

A continuación se plantea realizar una evaluación interactiva del funcionamiento del clasificador de dígitos. Se reconocerán los dígitos dibujados sobre el *sketchpad* de la biblioteca *gradio*.

La función `preprocesar()` provista en la siguiente celda, realiza la conversión de las imágenes generadas con el *sketchpad* al formato de los datos de entrenamiento. A diferencia de los datos de entrenamiento, los dígitos en el *sketchpad* de *gradio* son negros sobre fondo blanco, tomando el blanco el valor 0 y el negro el 255.

```
1 def preprocesar(img):
2     # Se reescala la imagen al tamaño de los dígitos de entrenamiento
3     img = zoom(img, 16 / img.shape[1])
4     # Se normaliza de [0, 255] -> [0, 1]
5     # img = img / 255
6     img = img / np.max(img[:])
7     # Se invierten los valores (0,1) -> (1,0)
8     # img = 1 - img
9     img = img.reshape(1, 16 * 16).astype(float)
10    return img
```

# Tabla de contenido

Ejercicio 1: descenso por gradiente

Ejercicio 2: regresión lineal mediante mínimos cuadrados

Ejercicio 3: clasificación de flores de iris mediante regresión logística

Ejercicio 4: compromiso sesgo varianza

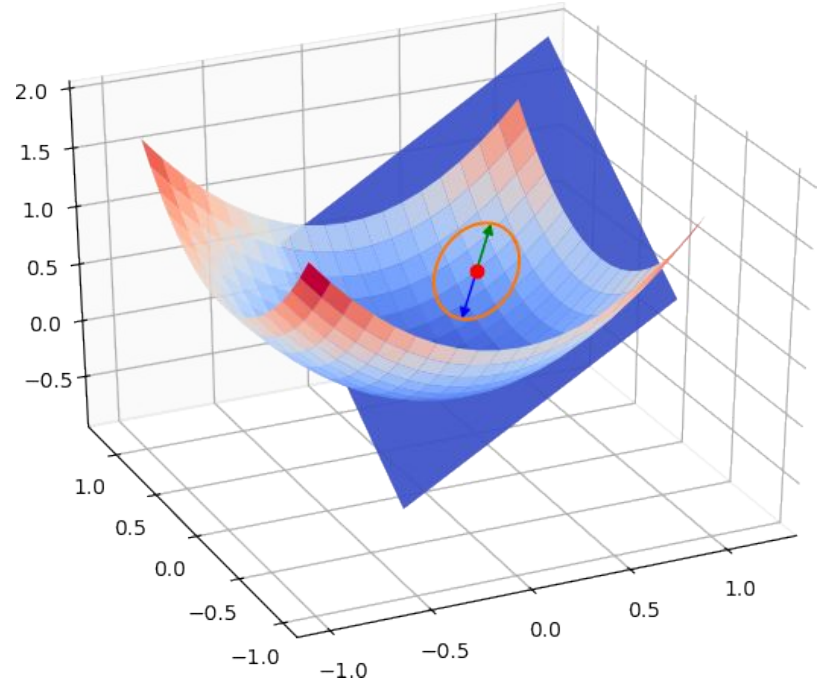
# Gradient descent (GD)

- Algoritmo (iterando):

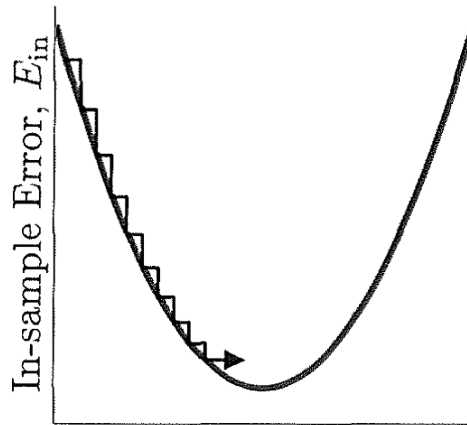
- Calcular el gradiente en el punto
- Dar «un paso» en la dirección opuesta al gradiente

$$x[k + 1] = x[k] - \eta \nabla f(x[k])$$

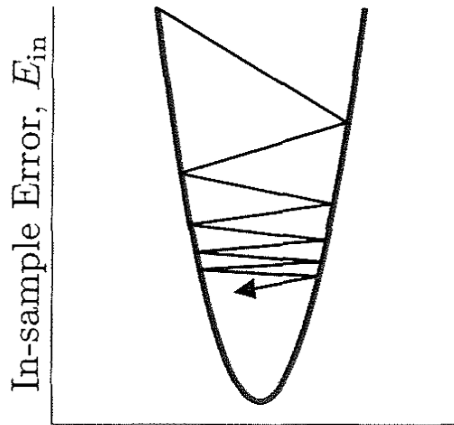
- Learning rate
- Criterio de parada
- Punto de partida
- Depending del tamaño de los datos:  
batch GD, stochastic GD, mini-batch  
GD.



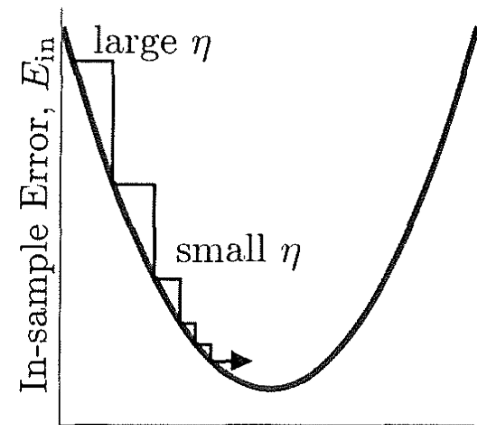
# Gradient descent: learning rate



Weights,  $w$   
 $\eta$  too small



Weights,  $w$   
 $\eta$  too large



Weights,  $w$   
variable  $\eta$  just right

# Gradiente

$$\begin{aligned} f &: \mathbb{R}^n \rightarrow \mathbb{R} \\ \nabla f &: \mathbb{R}^n \rightarrow \mathbb{R}^n \\ \mathbf{x} &\in \mathbb{R}^n \end{aligned} \quad \nabla f(\mathbf{x}) = \left( \frac{\partial f(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^\top$$

$$\frac{\partial(x^\top y)}{\partial x} = y$$

$$\frac{d(f(x, y))}{dx} = \frac{\partial(f(x, y))}{\partial x} + \frac{d(y^\top(x))}{dx} \frac{\partial(f(x, y))}{\partial y}$$

$$\nabla_{\mathbf{x}} (\mathbf{x}^\top \mathbf{b}) = \mathbf{b} \quad \nabla_{\mathbf{x}} (\mathbf{x}^\top \mathbf{x}) = 2\mathbf{x}$$

$$\nabla_{\mathbf{x}} (\mathbf{x}^\top \mathbf{A} \mathbf{x}) = (\mathbf{A} + \mathbf{A}^\top) \mathbf{x}$$