



ENFIN761

# Business Intelligence para las Finanzas

## AYUDANTÍA 4

**Profesor:** *David Díaz S.*

**Ayudantes:** *Gabriel Cabrera G.*<sup>1</sup>

*03 septiembre 2019*

### Gradiente Descendiente

1. Genere una variable  $X$  e  $y$  que contengan 100 observaciones con las siguientes características:

- a.  $X = 3 * Aleatorio$
- b.  $y = 2 + 2 * X + Aleatorio$

El parámetro *Aleatorio* se refiere a la función `np.random()`. Utilice una semilla igual a 42.

2. Estime los parámetros  $\hat{\theta}$  utilizando la ecuación normal.

$$\hat{\theta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

3. Realice la predicción utilizando los parámetros  $\hat{\theta}$  encontrados en (2), luego gráfíquelos.
4. Compruebe que los parámetros  $\hat{\theta}$  obtenidos mediante la ecuación normal son equivalentes al utilizar la librería Scikit-Learn.
5. Programe los siguientes algoritmos:
  - a. Batch Gradient Descent
  - b. Stochastic Gradient Descent
  - c. Mini-batch gradient descent

Explique cada procedimiento y utilícelos para obtener los parámetros  $\hat{\theta}$ .

### Aplicación 1: Regresión lineal

El archivo `data_usd_clp.xlsx` contiene datos de divisas extranjeras, entre ellas, el dolar. Utilizando las variables de la base de datos, realice la predicción de la variación del precio del dolar utilizando una regresión lineal mediante la librería Scikit-Learn. Considere los siguientes pasos:

---

<sup>1</sup>✉:gcabrera@fen.uchile.cl

1. Importe la base de datos a su espacio de trabajo (*workspace*).
2. Construya hasta el tercer rezago (*lag*) de cada variable más la variación del precio del dolar. De ser necesario, programe una función que facilite la implementación de rezagos.
3. Utilice el año 2015-06-19 como corte para generar su muestra de *training* y de *testing*. Luego realice las transformaciones pertinentes.
4. Use las métricas MAE (*mean absolute error*) y RMSE (*root mean square error*) para medir el *accuracy* de su modelo.
5. Genere dos gráficos de linea, el primero que muestre la relación entre la muestra de *training* con su predicción y el segundola relación entre la muestra de *testing* con su predicción.