

## Aprendizaje Estadístico.

Máster Interuniversitario en Tecnologías de Análisis de Datos Masivos: Big Data

Curso 2017/2018

Segunda actividad de evaluación continua

### Método de descenso de gradiente

Considera un modelo de regresión lineal simple  $Y = \beta_0 + \beta_1 X_1 + \epsilon$ . En primer lugar, simula una muestra con  $n = 100$  observaciones,  $(y_i, x_i)$ , del modelo de regresión lineal simple con parámetros  $\beta_0$  y  $\beta_1$  dados. Para ello, genera en primer lugar los valores  $x_i$  a partir de una distribución uniforme. A continuación genera los errores del modelo,  $\epsilon_i$ , a partir de una distribución normal de media 0 y varianza  $\sigma^2$ . Los valores  $y_i$  se calcularán entonces como

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i.$$

Aproxima los parámetros del modelo mediante el método de descenso de gradiente (también conocido como batch gradient descent). Para ello debes minimizar la función:

$$J(\beta_0, \beta_1) = \frac{1}{2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2.$$

El método de gradiente para minimizar la función  $J(\beta_0, \beta_1)$  se puede escribir:

---

**Algoritmo:** Método de descenso de gradiente para  $J(\beta_0, \beta_1)$

---

**Dado**  $(\hat{\beta}_0, \hat{\beta}_1)$  y  $t > 0$

**repite**

$$\hat{\beta}_0 = \hat{\beta}_0 + t \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)$$

$$\hat{\beta}_1 = \hat{\beta}_1 + t \sum_{i=1}^n x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)$$

**hasta** que se cumpla el criterio de parada

---

Programalo en R y haz una comparativa de los resultados obtenidos con los que devuelve la función `lm`.