PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
INSTITUTO DE INGENIERÍA MATEMÁTICA Y COMPUTACIONAL
IMT2220-1 Cálculo para Ciencia de Datos, Semestre 2023-01
Profesor: EDUARDO CERPA
Ayudante: PABLO RADEMACHER *(pablo.rademacher@uc.cl)*

Tarea 1

Esta tarea consta de 2 problemas que deben resolver en grupos definidos por el profesor. Para completar cada problema el ayudante Pablo les hará entrega de archivos con datos. Ustedes, además de entregar esta tarea en formato PDF, deberán darle acceso a Pablo a los códigos utilizados en la resolución. Deben programar todas las funciones involucradas y no pueden usar comandos ya creados.

1. **Regresión Cuadrática.**
Este primer problema se trata de resolver un problema de regresión cuadrática como el de la Figura 1. En el archivo que se les hará llegar encontrarán una lista de datos de la forma $(x_i, y_i) \in \mathbb{R}^2$ con $i = 1, 2, \cdots, M$. Ustedes deben encontrar coeficientes $\beta_0, \beta_1, \beta_2 \in \mathbb{R}$ tales que el modelo

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

sea el de tipo cuadrático que mejor se ajuste a los datos entregados.

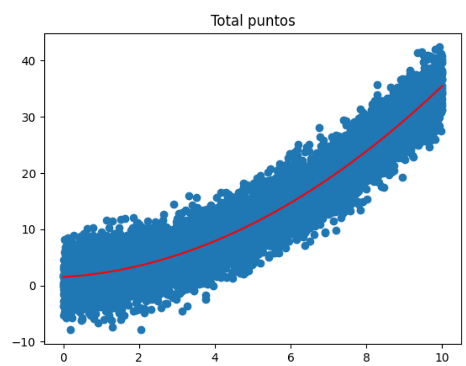

Figura 1: Ejemplo de Regresión Cuadrática.

   (a) Aplique el método de optimización visto en el curso con que resolvimos el problema de regresión lineal. Deduzca la fórmula para encontrar los coeficientes a partir de los datos.
   (b) Encuentre los valores de $\beta_0, \beta_1, \beta_2 \in \mathbb{R}$ buscados.
   (c) Grafique en conjunto los datos entregados y la regresión cuadrática encontrada.

2. **Red Neuronal.**
Este segundo problema se trata de generar una red neuronal que pueda aproximar una función sencilla. En el archivo que se les hará llegar encontrarán una lista de datos de entrenamiento de la forma $(u_i, v_i, w_i, z_i) \in \mathbb{R}^4$ con $i = 1, 2, \cdots, M$ en donde los tres primeros números corresponden a la entrada y el último a la salida. Usaremos la estructura de la Figura 2, en donde $(X_1, X_2, X_3) \in \mathbb{R}^3$ es la entrada, $Y \in \mathbb{R}$ es el valor en la capa intermedia, y $Z \in \mathbb{R}$ es la salida. Denotando σ una función de activación tendremos

$$Y = \sigma(aX_1 + bX_2 + cX_3 + e), \qquad Z = \sigma(dY + f),$$

en donde los seis parámetros de ajuste de la Red Neuronal son los pesos a, b, c, d y los sesgos e, f.

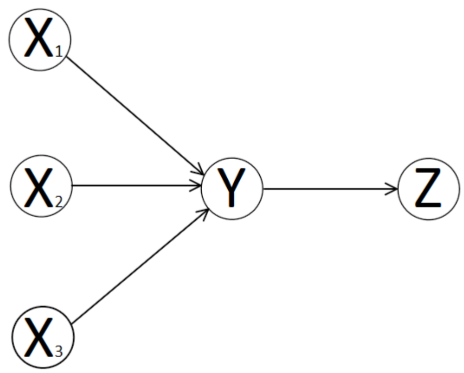

Figura 2: Estructura de la Red Neuronal.

(a) Considere la función de activación sigmoide

$$\sigma(t) = \frac{1}{1 + e^- t}$$

y encuentre una expresión para su derivada $\sigma'(t)$.

(b) Considerando los datos de entrenamiento $(u_i, v_i, w_i, z_i) \in \mathbb{R}^4$ construya $J = J(a, b, c, d, e, f)$ la función de pérdida del error cuadrático medio, como vimos en el ejemplo del curso, y calcule explícitamente su gradiente en un punto cualquiera $(a, b, c, d, e, f) \in \mathbb{R}^6$.

(c) Inicialice los pesos y sesgos de esta Red Neuronal asignándole valores en el intervalo $[-5, 5]$ a los parámetros $a_0, b_0, c_0, d_0, e_0, f_0$ y calcule $J(a_0, b_0, c_0, d_0, e_0, f_0)$.

(d) Para la búsqueda de un mínimo programe un algoritmo de descenso del gradiente

$$(a_{k+1}, b_{k+1}, c_{k+1}, d_{k+1}, e_{k+1}, f_{k+1})^T = (a_k, b_k, c_k, d_k, e_k, f_{k+1})^T - \varepsilon \nabla J(a_k, b_k, c_k, d_k, e_k, f_{k+1})$$

con un paso fijo $\varepsilon = 0,1$ y genere 2000 iteraciones. Muestre con un gráfico la evolución del error J usando los datos generados.

(e) Entregue los valores encontrados para $(a_{2000}, b_{2000}, c_{2000}, d_{2000}, e_{2000}, f_{2000}) \in \mathbb{R}^6$ y el valor

$$J(a_{2000}, b_{2000}, c_{2000}, d_{2000}, e_{2000}, f_{2000}).$$