I. Regresión Lineal o Polinomial de una variable.

A. Función de error (Mean Squared Error).

$$E(\theta|\mathcal{X}) = \frac{1}{N}\sum_{t=1}^{N}[r^t - g(x^t|\theta)]^2, \tag{1}$$

donde $x^t \in \mathcal{X}$ son las muestras y $r^t \in \mathcal{X}$, son las etiquetas correspondientes, $N = |\mathcal{X}|$, y la función $g(x^t|\theta) = w_0 + w_1x^t + w_2(x^t)^2 + ... + w_k(x^t)^k$ es nuestro estimador, que depende de los parámetros particulares del modelo $\theta : w_0, w_1, w_2, ..., w_k$.

B. Aprendizaje.

$$\mathbf{w} = (D^TD)^{-1}(D^T\mathbf{r}), \tag{2}$$

donde:
(a) $\mathbf{w} = [w_0\ w_1\ w_2\ ...\ w_k]^T$, válido si $(D^TD)^{-1}$ existe,
(b) $\mathbf{r} = [r^1\ r^2\ r^3 ...\ r^N]^T$,
(c) D es una matriz de tamaño $(N \times k+1)$ en que k representa el orden del polinomio para el modelo de regresión (i.e., $k = 1$ para el caso lineal, $k = 2$ para el caso cuadrático, y así sucesivamente):

$$D = \begin{bmatrix} 1 & x^1 & (x^1)^2 & \cdots & (x^1)^k \\ 1 & x^2 & (x^2)^2 & \cdots & (x^2)^k \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x^N & (x^N)^2 & \cdots & (x^N)^k \end{bmatrix}$$

C. Ejecución.

$$g(x^t|\theta) = \mathbf{w}^T[1\ x^t\ (x^t)^2\ ...\ (x^t)^k] \tag{3}$$

II. Regresión Lineal multi-variable.

A. *Función de error (Mean Squared Error).*

$$E(\theta|\mathcal{X}) = \frac{1}{N}\sum_{t=1}^{N}[r^t - g(\mathbf{x}^t|\theta)]^2, \tag{4}$$

donde los vectores $\mathbf{x}^t = [x_1^t \ x_2^t \ ... \ x_k^t \ ] \in \mathcal{X}$ son las muestras y los escalares $r^t \in \mathcal{X}$, son las etiquetas correspondientes, $N = |\mathcal{X}|$, y la función $g(\mathbf{x}^t|\theta) = w_0 + w_1x_1^t + w_2x_2^t + ... + w_kx_k^t$ es nuestro estimador, que depende de los parámetros particulares del modelo $\theta : w_0, w_1, w_2, ..., w_k$.

B. *Aprendizaje.*

$$\mathbf{w} = (D^TD)^{-1}(D^T\mathbf{r}), \tag{5}$$

donde:
(a) $\mathbf{w} = [w_0 \ w_1 \ w_2 \ ... \ w_k]^T$, válido si $(D^TD)^{-1}$ existe,
(b) $\mathbf{r} = [r^1 \ r^2 \ r^3 ... \ r^N]^T$,
(c) D es una matriz de tamaño $(N \times k + 1)$ en que k representa la cantidad de variables (también llamadas descriptores o *features*) para el modelo de regresión lineal:

$$D = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \cdots & x_k^1 \\ 1 & x_1^2 & x_2^2 & \cdots & x_k^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_1^N & x_2^N & \cdots & x_k^N \end{bmatrix}$$

C. *Ejecución.*

$$g(\mathbf{x}^t|\theta) = \mathbf{w}^T[1 \ x_1 \ x_2^t \ ... \ x_k^t] \tag{6}$$