

Meta 3.2 Caso de estudio: Modelo de regresión multivariada.

El modelo de regresión multivariada relaciona más de un predictor y más de una respuesta.

La regresión multivariada es un método utilizado para medir el grado en que más de una variable independiente (predictores) y más de una variable dependiente (respuestas) están relacionadas linealmente. El método se usa ampliamente para predecir el comportamiento de las variables de respuesta asociadas a los cambios en las variables predictoras, una vez que se ha establecido un grado de relación deseada.

Sean q observaciones $\{x_{p,i}; y_{p,j}\}_{p=1}^q$ con n predictores (\mathbf{X}) y m respuestas (\mathbf{Y}). Sea \mathbf{Y} la matriz de respuestas $q \times m$, \mathbf{A} es la matriz de diseño de tamaño $q \times \varrho$ tal que las n primeras columnas son los predictores (\mathbf{X}) y la última columna sean 1's, es decir $\mathbf{A} = [\mathbf{X} | \mathbf{1}_{q \times 1}]$. Sea $\boldsymbol{\Theta}$ una matriz $\varrho \times m$ de parámetros fijos, $\boldsymbol{\Xi}$ una matriz $q \times m$ tal que $\boldsymbol{\Xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ (multivariada normalmente distribuida con matriz de covarianza $\boldsymbol{\Sigma}$) y $\varrho = n + 1$ es el número de parámetros.

Dado el siguiente conjunto de datos de entrenamiento (challenge00_dataset22.txt) de dos atributos predictores ($\mathbf{X}=[x_{p,i}]$ para $i=1,2$ y $p=1\dots4$) y dos atributos de respuesta ($\mathbf{Y}=[y_{p,j}]$ para $j=1,2$ y $p=1\dots4$).

$$\mathbf{X} = \begin{bmatrix} 4.7 & 6.0 \\ 6.1 & 3.9 \\ 2.9 & 4.2 \\ 7.0 & 5.5 \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} 3.52 & 4.02 \\ 5.43 & 6.23 \\ 4.95 & 5.76 \\ 4.70 & 4.28 \end{bmatrix}$$

Realizar un análisis para demostrar lo siguiente:

1. Función de la hipótesis del modelo de regresión lineal múltiple con múltiples salidas.
2. Función de costo (Métricas de regresión SSE, MSE, RMSE)
3. Función gradiente, $\nabla E(\boldsymbol{\Theta}) \equiv \frac{\partial E}{\partial \boldsymbol{\Theta}}$
4. Función jacobiana, $\mathbf{J}_{\boldsymbol{\Xi}} \equiv \frac{\partial \boldsymbol{\Xi}}{\partial \boldsymbol{\Theta}}$
5. Encontrar los parámetros del modelo *lineal múltiple con múltiples salidas*.
6. Implementar las funciones de los puntos 1-4.

1. Función de la hipótesis del modelo de regresión multivariada.

La hipótesis de regresión para este caso de estudio se define como;

$$\hat{y}_{p,j} = \theta_{1,j}x_{p,1} + \theta_{2,j}x_{p,2} + \theta_{3,j} \text{ para } j=1,2 \text{ y } p=1\dots4$$

$$e_{p,j} = y_{p,j} - \hat{y}_{p,j}$$

Calcular para todas las $p=1,2,3,4$ y $j=1,2$

$$\hat{y}_{1,1} = \theta_{1,1}x_{1,1} + \theta_{2,1}x_{1,2} + \theta_{3,1}$$

$$\hat{y}_{2,1} = \theta_{1,1}x_{2,1} + \theta_{2,1}x_{2,2} + \theta_{3,1}$$

$$\hat{y}_{3,1} = \theta_{1,1}x_{3,1} + \theta_{2,1}x_{3,2} + \theta_{3,1}$$

$$\hat{y}_{4,1} = \theta_{1,1}x_{4,1} + \theta_{2,1}x_{4,2} + \theta_{3,1}$$

$$\hat{y}_{1,2} = \theta_{1,2}x_{1,1} + \theta_{2,2}x_{1,2} + \theta_{3,2}$$

$$\hat{y}_{2,2} = \theta_{1,2}x_{2,1} + \theta_{2,2}x_{2,2} + \theta_{3,2}$$

$$\hat{y}_{3,2} = \theta_{1,2}x_{3,1} + \theta_{2,2}x_{3,2} + \theta_{3,2}$$

$$\hat{y}_{4,2} = \theta_{1,2}x_{4,1} + \theta_{2,2}x_{4,2} + \theta_{3,2}$$

$$\begin{bmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} \\ \hat{y}_{2,1} & \hat{y}_{2,2} \\ \hat{y}_{3,1} & \hat{y}_{3,2} \\ \hat{y}_{4,1} & \hat{y}_{4,2} \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & 1 \\ x_{2,1} & x_{2,2} & 1 \\ x_{3,1} & x_{3,2} & 1 \\ x_{4,1} & x_{4,2} & 1 \end{bmatrix} \begin{bmatrix} \theta_{1,1} & \theta_{1,2} \\ \theta_{2,1} & \theta_{2,2} \\ \theta_{3,1} & \theta_{3,2} \end{bmatrix}$$

$$\hat{Y} = A\Theta$$

donde \hat{Y} es la hipótesis del modelo de regresión lineal múltiple con múltiples salidas, A es la matriz de diseño y Θ es la matriz de parámetros del modelo de regresión lineal. A continuación se muestran las variables matriciales.

$$\hat{Y} = \begin{bmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} \\ \hat{y}_{2,1} & \hat{y}_{2,2} \\ \hat{y}_{3,1} & \hat{y}_{3,2} \\ \hat{y}_{4,1} & \hat{y}_{4,2} \end{bmatrix}, A = \begin{bmatrix} x_{1,1} & x_{1,2} & 1 \\ x_{2,1} & x_{2,2} & 1 \\ x_{3,1} & x_{3,2} & 1 \\ x_{4,1} & x_{4,2} & 1 \end{bmatrix} = [X | 1_{4 \times 1}], \Theta = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} \\ \theta_{2,1} & \theta_{2,2} \\ \theta_{3,1} & \theta_{3,2} \end{bmatrix}$$

Las medidas de error se muestran a continuación:

$$e_{1,1} = y_{1,1} - \hat{y}_{1,1}$$

$$e_{2,1} = y_{2,1} - \hat{y}_{2,1}$$

$$e_{3,1} = y_{3,1} - \hat{y}_{3,1}$$

$$e_{4,1} = y_{4,1} - \hat{y}_{4,1}$$

$$e_{1,2} = y_{1,2} - \hat{y}_{1,2}$$

$$e_{2,2} = y_{2,2} - \hat{y}_{2,2}$$

$$e_{3,2} = y_{3,2} - \hat{y}_{3,2}$$

$$e_{4,2} = y_{4,2} - \hat{y}_{4,2}$$

La matriz de las medidas de error es:

$$\begin{bmatrix} e_{1,1} & e_{1,2} \\ e_{2,1} & e_{2,2} \\ e_{3,1} & e_{3,2} \\ e_{4,1} & e_{4,2} \end{bmatrix} = \begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \\ y_{3,1} & y_{3,2} \\ y_{4,1} & y_{4,2} \end{bmatrix} - \begin{bmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} \\ \hat{y}_{2,1} & \hat{y}_{2,2} \\ \hat{y}_{3,1} & \hat{y}_{3,2} \\ \hat{y}_{4,1} & \hat{y}_{4,2} \end{bmatrix}$$

donde

$$\mathbf{E} = \begin{bmatrix} e_{1,1} & e_{1,2} \\ e_{2,1} & e_{2,2} \\ e_{3,1} & e_{3,2} \\ e_{4,1} & e_{4,2} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \\ y_{3,1} & y_{3,2} \\ y_{4,1} & y_{4,2} \end{bmatrix}, \quad \hat{\mathbf{Y}} = \begin{bmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} \\ \hat{y}_{2,1} & \hat{y}_{2,2} \\ \hat{y}_{3,1} & \hat{y}_{3,2} \\ \hat{y}_{4,1} & \hat{y}_{4,2} \end{bmatrix}$$

$$\mathbf{E} = \mathbf{Y} - \hat{\mathbf{Y}}$$

2. Función de costo (Métricas de regresión SSE, MSE, RMSE)

La suma de los errores al cuadrado (SSE) se define como:

$$SSE = e_{1,1}^2 + e_{2,1}^2 + e_{3,1}^2 + e_{4,1}^2 + e_{1,2}^2 + e_{2,2}^2 + e_{3,2}^2 + e_{4,2}^2$$

$$SSE \equiv \sum_{p=1}^q \sum_{j=1}^m e_{p,j}^2$$

Entonces la media de la suma de los errores al cuadrado (MSE) es:

$$MSE \equiv SSE/(mq)$$

y la raíz de la media de la suma de los errores al cuadrado (RMSE) se evalúa como:

$$RMSE \equiv \sqrt{MSE}$$

Si las matrices \mathbf{E} y $\mathbf{\Theta}$ son vectorizadas en columna denotadas como:

$$\vec{\mathbf{E}} = \begin{bmatrix} e_{1,1} \\ e_{2,1} \\ e_{3,1} \\ e_{4,1} \\ e_{1,2} \\ e_{2,2} \\ e_{3,2} \\ e_{4,2} \end{bmatrix}, \quad \vec{\mathbf{\Theta}} = \begin{bmatrix} \theta_{1,1} \\ \theta_{2,1} \\ \theta_{3,1} \\ \theta_{1,2} \\ \theta_{2,2} \\ \theta_{3,2} \end{bmatrix}$$

entonces el SSE se evalúa como:

este es el que vamos a optimizar el SSE

$$SSE = [e_{1,1} \quad e_{2,1} \quad e_{3,1} \quad e_{4,1} \quad e_{1,2} \quad e_{2,2} \quad e_{3,2} \quad e_{4,2}] \begin{bmatrix} e_{1,1} \\ e_{2,1} \\ e_{3,1} \\ e_{4,1} \\ e_{1,2} \\ e_{2,2} \\ e_{3,2} \\ e_{4,2} \end{bmatrix}$$

$$SSE = \vec{e}'\vec{e}$$

teta - todas las variables

3. Función gradiente, $\nabla E(\Theta) \equiv \frac{\partial E}{\partial \theta_{i,j}}$

Asumiendo que $E = SSE$, se evalúan las derivadas $\frac{\partial E}{\partial \theta_{i,j}}$ para $i=1,\dots,n+1$ y $j=1,\dots,m$

Matriz gradiente: $\frac{\partial E}{\partial \theta}$

$$\frac{\partial E}{\partial \theta_{1,1}} = \frac{\partial E}{\partial e_{1,1}} \frac{\partial e_{1,1}}{\partial \hat{y}_{1,1}} \frac{\partial \hat{y}_{1,1}}{\partial \theta_{1,1}} + \frac{\partial E}{\partial e_{2,1}} \frac{\partial e_{2,1}}{\partial \hat{y}_{2,1}} \frac{\partial \hat{y}_{2,1}}{\partial \theta_{1,1}} + \frac{\partial E}{\partial e_{3,1}} \frac{\partial e_{3,1}}{\partial \hat{y}_{3,1}} \frac{\partial \hat{y}_{3,1}}{\partial \theta_{1,1}} + \frac{\partial E}{\partial e_{4,1}} \frac{\partial e_{4,1}}{\partial \hat{y}_{4,1}} \frac{\partial \hat{y}_{4,1}}{\partial \theta_{1,1}}$$

$$\frac{\partial E}{\partial \theta_{1,1}} = (2e_{1,1}) (-1) (x_{1,1}) + (2e_{2,1}) (-1) (x_{2,1}) + (2e_{3,1}) (-1) (x_{3,1}) + (2e_{4,1}) (-1) (x_{4,1})$$

$$\frac{\partial E}{\partial \theta_{1,1}} = -2e_{1,1}x_{1,1} - 2e_{2,1}x_{2,1} - 2e_{3,1}x_{3,1} - 2e_{4,1}x_{4,1}$$

$$\frac{\partial E}{\partial \theta_{1,1}} = -2 \sum_{p=1}^4 e_{p,1}x_{p,1}$$

$$\frac{\partial E}{\partial \theta_{2,1}} = \frac{\partial E}{\partial e_{1,1}} \frac{\partial e_{1,1}}{\partial \hat{y}_{1,1}} \frac{\partial \hat{y}_{1,1}}{\partial \theta_{2,1}} + \frac{\partial E}{\partial e_{2,1}} \frac{\partial e_{2,1}}{\partial \hat{y}_{2,1}} \frac{\partial \hat{y}_{2,1}}{\partial \theta_{2,1}} + \frac{\partial E}{\partial e_{3,1}} \frac{\partial e_{3,1}}{\partial \hat{y}_{3,1}} \frac{\partial \hat{y}_{3,1}}{\partial \theta_{2,1}} + \frac{\partial E}{\partial e_{4,1}} \frac{\partial e_{4,1}}{\partial \hat{y}_{4,1}} \frac{\partial \hat{y}_{4,1}}{\partial \theta_{2,1}}$$

$$\frac{\partial E}{\partial \theta_{2,1}} = (2e_{1,1}) (-1) (x_{1,2}) + (2e_{2,1}) (-1) (x_{2,2}) + (2e_{3,1}) (-1) (x_{3,2}) + (2e_{4,1}) (-1) (x_{4,2})$$

$$\frac{\partial E}{\partial \theta_{2,1}} = -2e_{1,1}x_{1,2} - 2e_{2,1}x_{2,2} - 2e_{3,1}x_{3,2} - 2e_{4,1}x_{4,2}$$

$$\frac{\partial E}{\partial \theta_{2,1}} = -2 \sum_{p=1}^4 e_{p,1} x_{p,2}$$

$$\frac{\partial E}{\partial \theta_{3,1}} = \frac{\partial E}{\partial e_{1,1}} \frac{\partial e_{1,1}}{\partial \hat{y}_{1,1}} \frac{\partial \hat{y}_{1,1}}{\partial \theta_{3,1}} + \frac{\partial E}{\partial e_{2,1}} \frac{\partial e_{2,1}}{\partial \hat{y}_{2,1}} \frac{\partial \hat{y}_{2,1}}{\partial \theta_{3,1}} + \frac{\partial E}{\partial e_{3,1}} \frac{\partial e_{3,1}}{\partial \hat{y}_{3,1}} \frac{\partial \hat{y}_{3,1}}{\partial \theta_{3,1}} + \frac{\partial E}{\partial e_{4,1}} \frac{\partial e_{4,1}}{\partial \hat{y}_{4,1}} \frac{\partial \hat{y}_{4,1}}{\partial \theta_{3,1}}$$

$$\frac{\partial E}{\partial \theta_{3,1}} = (2e_{1,1}) (-1) (1) + (2e_{2,1}) (-1) (1) + (2e_{3,1}) (-1) (1) + (2e_{4,1}) (-1) (1)$$

$$\frac{\partial E}{\partial \theta_{3,1}} = -2e_{1,1} - 2e_{2,1} - 2e_{3,1} - 2e_{4,1}$$

$$\frac{\partial E}{\partial \theta_{3,1}} = -2 \sum_{p=1}^4 e_{p,1}$$

$$\frac{\partial E}{\partial \theta_{1,2}} = \frac{\partial E}{\partial e_{1,2}} \frac{\partial e_{1,2}}{\partial \hat{y}_{1,2}} \frac{\partial \hat{y}_{1,2}}{\partial \theta_{1,2}} + \frac{\partial E}{\partial e_{2,2}} \frac{\partial e_{2,2}}{\partial \hat{y}_{2,2}} \frac{\partial \hat{y}_{2,2}}{\partial \theta_{1,2}} + \frac{\partial E}{\partial e_{3,2}} \frac{\partial e_{3,2}}{\partial \hat{y}_{3,2}} \frac{\partial \hat{y}_{3,2}}{\partial \theta_{1,2}} + \frac{\partial E}{\partial e_{4,2}} \frac{\partial e_{4,2}}{\partial \hat{y}_{4,2}} \frac{\partial \hat{y}_{4,2}}{\partial \theta_{1,2}}$$

$$\frac{\partial E}{\partial \theta_{1,2}} = (2e_{1,2}) (-1) (x_{1,1}) + (2e_{2,2}) (-1) (x_{2,1}) + (2e_{3,2}) (-1) (x_{3,1}) + (2e_{4,2}) (-1) (x_{4,1})$$

$$\frac{\partial E}{\partial \theta_{1,2}} = -2e_{1,2}x_{1,1} - 2e_{2,2}x_{2,1} - 2e_{3,2}x_{3,1} - 2e_{4,2}x_{4,1}$$

$$\frac{\partial E}{\partial \theta_{1,2}} = -2 \sum_{p=1}^4 e_{p,2} x_{p,1}$$

$$\frac{\partial E}{\partial \theta_{2,2}} = \frac{\partial E}{\partial e_{1,2}} \frac{\partial e_{1,2}}{\partial \hat{y}_{1,2}} \frac{\partial \hat{y}_{1,2}}{\partial \theta_{2,2}} + \frac{\partial E}{\partial e_{2,2}} \frac{\partial e_{2,2}}{\partial \hat{y}_{2,2}} \frac{\partial \hat{y}_{2,2}}{\partial \theta_{2,2}} + \frac{\partial E}{\partial e_{3,2}} \frac{\partial e_{3,2}}{\partial \hat{y}_{3,2}} \frac{\partial \hat{y}_{3,2}}{\partial \theta_{2,2}} + \frac{\partial E}{\partial e_{4,2}} \frac{\partial e_{4,2}}{\partial \hat{y}_{4,2}} \frac{\partial \hat{y}_{4,2}}{\partial \theta_{2,2}}$$

$$\frac{\partial E}{\partial \theta_{2,2}} = (2e_{1,2}) (-1) (x_{1,2}) + (2e_{2,2}) (-1) (x_{2,2}) + (2e_{3,2}) (-1) (x_{3,2}) + (2e_{4,2}) (-1) (x_{4,2})$$

$$\frac{\partial E}{\partial \theta_{2,2}} = -2e_{1,2}x_{1,2} - 2e_{2,2}x_{2,2} - 2e_{3,2}x_{3,2} - 2e_{4,2}x_{4,2}$$

$$\frac{\partial E}{\partial \theta_{2,2}} = -2 \sum_{p=1}^4 e_{p,2} x_{p,2}$$

$$\frac{\partial E}{\partial \theta_{3,2}} = \frac{\partial E}{\partial e_{1,2}} \frac{\partial e_{1,2}}{\partial \hat{y}_{1,2}} \frac{\partial \hat{y}_{1,2}}{\partial \theta_{3,2}} + \frac{\partial E}{\partial e_{2,2}} \frac{\partial e_{2,2}}{\partial \hat{y}_{2,2}} \frac{\partial \hat{y}_{2,2}}{\partial \theta_{3,2}} + \frac{\partial E}{\partial e_{3,2}} \frac{\partial e_{3,2}}{\partial \hat{y}_{3,2}} \frac{\partial \hat{y}_{3,2}}{\partial \theta_{3,2}} + \frac{\partial E}{\partial e_{4,2}} \frac{\partial e_{4,2}}{\partial \hat{y}_{4,2}} \frac{\partial \hat{y}_{4,2}}{\partial \theta_{3,2}}$$

$$\frac{\partial E}{\partial \theta_{3,2}} = (2e_{1,2}) (-1) (1) + (2e_{2,2}) (-1) (1) + (2e_{3,2}) (-1) (1) + (2e_{4,2}) (-1) (1)$$

$$\frac{\partial E}{\partial \theta_{3,2}} = -2e_{1,2} - 2e_{2,2} - 2e_{3,2} - 2e_{4,2}$$

$$\frac{\partial E}{\partial \theta_{3,2}} = -2 \sum_{p=1}^4 e_{p,2}$$

El gradiente matricial $\frac{\partial E}{\partial \Theta} = \left[\frac{\partial E}{\partial \theta_{i,j}} \right]$ para $i=1,...,3$ y $j=1,2$ del caso de estudio se puede simplificar con las siguientes sumatorias obtenidas anteriormente;

$$\frac{\partial E}{\partial \theta_{i,j}} = -2 \sum_{p=1}^4 e_{p,j} x_{p,i} \quad i = \text{parametro}, j = \text{modelo}$$

$$\frac{\partial E}{\partial \theta_{3,j}} = -2 \sum_{p=1}^4 e_{p,j}$$

O bien en forma matricial, sustituyendo los resultados anteriores, se expresa como:

$$\frac{\partial E}{\partial \Theta} = \begin{bmatrix} \frac{\partial E}{\partial \theta_{1,1}} & \frac{\partial E}{\partial \theta_{1,2}} \\ \frac{\partial E}{\partial \theta_{2,1}} & \frac{\partial E}{\partial \theta_{2,2}} \\ \frac{\partial E}{\partial \theta_{3,1}} & \frac{\partial E}{\partial \theta_{3,2}} \end{bmatrix} = \begin{bmatrix} -2 \sum_{p=1}^4 e_{p,1} x_{p,1} & -2 \sum_{p=1}^4 e_{p,2} x_{p,1} \\ -2 \sum_{p=1}^4 e_{p,1} x_{p,2} & -2 \sum_{p=1}^4 e_{p,2} x_{p,2} \\ -2 \sum_{p=1}^4 e_{p,1} & -2 \sum_{p=1}^4 e_{p,2} \end{bmatrix}$$

Expandiendo las sumatorias se obtiene

$$\frac{\partial E}{\partial \Theta} = \begin{bmatrix} -2e_{1,1}x_{1,1} - 2e_{2,1}x_{2,1} - 2e_{3,1}x_{3,1} - 2e_{4,1}x_{4,1} & -2e_{1,2}x_{1,1} - 2e_{2,2}x_{2,1} - 2e_{3,2}x_{3,1} - 2e_{4,2}x_{4,1} \\ -2e_{1,1}x_{1,2} - 2e_{2,1}x_{2,2} - 2e_{3,1}x_{3,2} - 2e_{4,1}x_{4,2} & -2e_{1,2}x_{1,2} - 2e_{2,2}x_{2,2} - 2e_{3,2}x_{3,2} - 2e_{4,2}x_{4,2} \\ -2e_{1,1} - 2e_{2,1} - 2e_{3,1} - 2e_{4,1} & -2e_{1,2} - 2e_{2,2} - 2e_{3,2} - 2e_{4,2} \end{bmatrix}$$

Factorizando la matriz anterior, obtenemos

$$\frac{\partial E}{\partial \vec{\Theta}} = 2 \begin{bmatrix} -x_{1,1} & -x_{2,1} & -x_{3,1} & -x_{4,1} \\ -x_{1,2} & -x_{2,2} & -x_{3,2} & -x_{4,2} \\ -1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} e_{1,1} & e_{1,2} \\ e_{2,1} & e_{2,2} \\ e_{3,1} & e_{3,2} \\ e_{4,1} & e_{4,2} \end{bmatrix}$$

Entonces $\frac{\partial E}{\partial \vec{\Theta}}$ es 2 veces la matriz de diseño transpuesta negativa ($-A'$) multiplicado por la matriz $\vec{\Xi}$, es decir;

$$\frac{\partial E}{\partial \vec{\Theta}} = 2(-A')\vec{\Xi}$$

4. Función jacobiana, $J_{\vec{\Xi}} \equiv \frac{\partial \vec{\Xi}}{\partial \vec{\Theta}}$

Desarrollando el vector gradiente $\frac{\partial E}{\partial \vec{\Theta}}$, obtenemos:

$$\begin{aligned} \frac{\partial E}{\partial \vec{\Theta}} = & 2 e_{1,1} \frac{\partial e_{1,1}}{\partial \vec{\Theta}} + 2 e_{2,1} \frac{\partial e_{2,1}}{\partial \vec{\Theta}} + 2 e_{3,1} \frac{\partial e_{3,1}}{\partial \vec{\Theta}} + 2 e_{4,1} \frac{\partial e_{4,1}}{\partial \vec{\Theta}} + \\ & 2 e_{1,2} \frac{\partial e_{1,2}}{\partial \vec{\Theta}} + 2 e_{2,2} \frac{\partial e_{2,2}}{\partial \vec{\Theta}} + 2 e_{3,2} \frac{\partial e_{3,2}}{\partial \vec{\Theta}} + 2 e_{4,2} \frac{\partial e_{4,2}}{\partial \vec{\Theta}} \\ \frac{\partial E}{\partial \vec{\Theta}} = & 2 \left[\frac{\partial e_{1,1}}{\partial \vec{\Theta}} \quad \frac{\partial e_{2,1}}{\partial \vec{\Theta}} \quad \frac{\partial e_{3,1}}{\partial \vec{\Theta}} \quad \frac{\partial e_{4,1}}{\partial \vec{\Theta}} \quad \frac{\partial e_{1,2}}{\partial \vec{\Theta}} \quad \frac{\partial e_{2,2}}{\partial \vec{\Theta}} \quad \frac{\partial e_{3,2}}{\partial \vec{\Theta}} \quad \frac{\partial e_{4,2}}{\partial \vec{\Theta}} \right] \begin{bmatrix} e_{1,1} \\ e_{2,1} \\ e_{3,1} \\ e_{4,1} \\ e_{1,2} \\ e_{2,2} \\ e_{3,2} \\ e_{4,2} \end{bmatrix} \end{aligned}$$

$$\frac{\partial E}{\partial \vec{\Theta}} = 2 J'_{\vec{\Xi}} \vec{\Xi}$$

Entonces el vector gradiente $\frac{\partial E}{\partial \vec{\Theta}}$ es 2 veces la jacobiana traspuesta de $\vec{\Xi}$ ($J'_{\vec{\Xi}}$) multiplicada por el vector $\vec{\Xi}$. La deriva parcial del vector $\vec{\Xi}$ con respecto al vector $\vec{\Theta}$ denotada como $\frac{\partial \vec{\Xi}}{\partial \vec{\Theta}}$, es la matriz jacobiana de $\vec{\Xi}$ definida como $J_{\vec{\Xi}} \equiv \frac{\partial \vec{\Xi}}{\partial \vec{\Theta}}$, es decir;

$$J_{\vec{\Xi}} \equiv \frac{\partial \vec{\Xi}}{\partial \vec{\Theta}} = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial \vec{\Theta}} \\ \frac{\partial e_{2,1}}{\partial \vec{\Theta}} \\ \frac{\partial e_{3,1}}{\partial \vec{\Theta}} \\ \frac{\partial e_{4,1}}{\partial \vec{\Theta}} \\ \frac{\partial e_{1,2}}{\partial \vec{\Theta}} \\ \frac{\partial e_{2,2}}{\partial \vec{\Theta}} \\ \frac{\partial e_{3,2}}{\partial \vec{\Theta}} \\ \frac{\partial e_{4,2}}{\partial \vec{\Theta}} \end{bmatrix} = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial \theta_{1,1}} & \frac{\partial e_{1,1}}{\partial \theta_{2,1}} & \frac{\partial e_{1,1}}{\partial \theta_{3,1}} & 0 & 0 & 0 \\ \frac{\partial e_{2,1}}{\partial \theta_{1,1}} & \frac{\partial e_{2,1}}{\partial \theta_{2,1}} & \frac{\partial e_{2,1}}{\partial \theta_{3,1}} & 0 & 0 & 0 \\ \frac{\partial e_{3,1}}{\partial \theta_{1,1}} & \frac{\partial e_{3,1}}{\partial \theta_{2,1}} & \frac{\partial e_{3,1}}{\partial \theta_{3,1}} & 0 & 0 & 0 \\ \frac{\partial e_{4,1}}{\partial \theta_{1,1}} & \frac{\partial e_{4,1}}{\partial \theta_{2,1}} & \frac{\partial e_{4,1}}{\partial \theta_{3,1}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial e_{1,2}}{\partial \theta_{1,2}} & \frac{\partial e_{1,2}}{\partial \theta_{2,2}} & \frac{\partial e_{1,2}}{\partial \theta_{3,2}} \\ 0 & 0 & 0 & \frac{\partial e_{2,2}}{\partial \theta_{1,2}} & \frac{\partial e_{2,2}}{\partial \theta_{2,2}} & \frac{\partial e_{2,2}}{\partial \theta_{3,2}} \\ 0 & 0 & 0 & \frac{\partial e_{3,2}}{\partial \theta_{1,2}} & \frac{\partial e_{3,2}}{\partial \theta_{2,2}} & \frac{\partial e_{3,2}}{\partial \theta_{3,2}} \\ 0 & 0 & 0 & \frac{\partial e_{4,2}}{\partial \theta_{1,2}} & \frac{\partial e_{4,2}}{\partial \theta_{2,2}} & \frac{\partial e_{4,2}}{\partial \theta_{3,2}} \end{bmatrix}$$

$$J_{\vec{\Xi}} \equiv \frac{\partial \vec{\Xi}}{\partial \vec{\Theta}} = \begin{bmatrix} -x_{1,1} & -x_{1,2} & -1 & 0 & 0 & 0 \\ -x_{2,1} & -x_{2,2} & -1 & 0 & 0 & 0 \\ -x_{3,1} & -x_{3,2} & -1 & 0 & 0 & 0 \\ -x_{4,1} & -x_{4,2} & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -x_{1,1} & -x_{1,2} & -1 \\ 0 & 0 & 0 & -x_{2,1} & -x_{2,2} & -1 \\ 0 & 0 & 0 & -x_{3,1} & -x_{3,2} & -1 \\ 0 & 0 & 0 & -x_{4,1} & -x_{4,2} & -1 \end{bmatrix} = - \begin{bmatrix} \mathbf{A}_{4 \times 3} & \mathbf{0}_{4 \times 3} \\ \mathbf{0}_{4 \times 3} & \mathbf{A}_{4 \times 3} \end{bmatrix}$$

matrz de disenio negativo

Entonces la jacobiana $J_{\vec{\Xi}}$ para m variables dependientes y q parámetros es:

$$J_{\vec{\Xi}} \equiv \frac{\partial \vec{\Xi}}{\partial \vec{\Theta}} = - \begin{bmatrix} \mathbf{A}_{q \times q} & \mathbf{0}_{q \times q} & \cdots & \mathbf{0}_{q \times q} \\ \mathbf{0}_{q \times q} & \mathbf{A}_{q \times q} & \cdots & \mathbf{0}_{q \times q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{q \times q} & \mathbf{0}_{q \times q} & \cdots & \mathbf{A}_{q \times q} \end{bmatrix}$$

Calculando el gradiente, usando la matriz jacobiana, obtenemos:

$$\frac{\partial E}{\partial \vec{\Theta}} = 2 \mathbf{J}'_{\Xi} \vec{\Xi} = 2 \begin{bmatrix} -x_{1,1} & -x_{2,1} & -x_{3,1} & -x_{4,1} & 0 & 0 & 0 & 0 \\ -x_{1,2} & -x_{2,2} & -x_{3,2} & -x_{4,2} & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -x_{1,1} & -x_{2,1} & -x_{3,1} & -x_{4,1} \\ 0 & 0 & 0 & 0 & -x_{1,2} & -x_{2,2} & -x_{3,2} & -x_{4,2} \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} e_{1,1} \\ e_{2,1} \\ e_{3,1} \\ e_{4,1} \\ e_{1,2} \\ e_{2,2} \\ e_{3,2} \\ e_{4,2} \end{bmatrix}$$

Multiplicando las matrices, obtenemos;

$$\frac{\partial E}{\partial \vec{\Theta}} = 2 \mathbf{J}'_{\Xi} \vec{\Xi} = \begin{bmatrix} -2e_{1,1}x_{1,1} - 2e_{2,1}x_{2,1} - 2e_{3,1}x_{3,1} - 2e_{4,1}x_{4,1} \\ -2e_{1,1}x_{1,2} - 2e_{2,1}x_{2,2} - 2e_{3,1}x_{3,2} - 2e_{4,1}x_{4,2} \\ -2e_{1,1} - 2e_{2,1} - 2e_{3,1} - 2e_{4,1} \\ -2e_{1,2}x_{1,1} - 2e_{2,2}x_{2,1} - 2e_{3,2}x_{3,1} - 2e_{4,2}x_{4,1} \\ -2e_{1,2}x_{1,2} - 2e_{2,2}x_{2,2} - 2e_{3,2}x_{3,2} - 2e_{4,2}x_{4,2} \\ -2e_{1,2} - 2e_{2,2} - 2e_{3,2} - 2e_{4,2} \end{bmatrix}$$

Si el vector gradiente $\frac{\partial E}{\partial \vec{\Theta}}$ se desvectoriza obtenemos la matriz gradiente $\frac{\partial E}{\partial \Theta}$, denotada como:

$$\frac{\partial E}{\partial \Theta} = \begin{bmatrix} -2e_{1,1}x_{1,1} - 2e_{2,1}x_{2,1} - 2e_{3,1}x_{3,1} - 2e_{4,1}x_{4,1} & -2e_{1,2}x_{1,1} - 2e_{2,2}x_{2,1} - 2e_{3,2}x_{3,1} - 2e_{4,2}x_{4,1} \\ -2e_{1,1}x_{1,2} - 2e_{2,1}x_{2,2} - 2e_{3,1}x_{3,2} - 2e_{4,1}x_{4,2} & -2e_{1,2}x_{1,2} - 2e_{2,2}x_{2,2} - 2e_{3,2}x_{3,2} - 2e_{4,2}x_{4,2} \\ -2e_{1,1} - 2e_{2,1} - 2e_{3,1} - 2e_{4,1} & -2e_{1,2} - 2e_{2,2} - 2e_{3,2} - 2e_{4,2} \end{bmatrix}$$

Factorizando la matriz anterior, es decir;

$$\frac{\partial E}{\partial \Theta} = 2 \begin{bmatrix} -x_{1,1} & -x_{2,1} & -x_{3,1} & -x_{4,1} \\ -x_{1,2} & -x_{2,2} & -x_{3,2} & -x_{4,2} \\ -1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} e_{1,1} & e_{1,2} \\ e_{2,1} & e_{2,2} \\ e_{3,1} & e_{3,2} \\ e_{4,1} & e_{4,2} \end{bmatrix}$$

Entonces $\frac{\partial E}{\partial \Theta}$ es 2 veces la matriz de diseño transpuesta negativa ($-\mathbf{A}'$) multiplicado por la matriz Ξ , es decir;

$$\frac{\partial E}{\partial \Theta} = 2(-\mathbf{A}')\Xi$$

Ambas demostraciones nos llevan al mismo resultado del gradiente en forma matricial.

5. Encontrar los parámetros del modelo de regresión multivariada.

Optimización de la función de costo, $E(\Theta)$:

$$\hat{\Theta} = \arg \min_{\Theta} E$$

esto solo funciona cuando el problema es lineal

$$\nabla E(\hat{\Theta}) = -2A'\Xi = 0$$

$$A'\Xi = 0$$

$$A'(Y - A\hat{\Theta}) = 0$$

$$A'Y - A'A\hat{\Theta} = 0$$

$$A'Y = A'A\hat{\Theta}$$

$$A'A\hat{\Theta} = A'Y$$

$$\hat{\Theta} = (A'A)^{-1}A'Y$$

$\hat{\Theta} = (A'A)^+ A'Y$, donde la pseudoinversa de Moore-Penrose, $(A'A)^+$ de una matriz $A'A$ es una generalización de la matriz inversa.

6. Implementar las funciones de los puntos 1-4.

Resumen

Hipótesis del modelo de regresión multivariada

$$\hat{y}_{p,j} = \sum_{i=1}^n \theta_{i,j} x_{p,i} + \theta_{n+1,j}$$

$$\hat{Y} = A\Theta$$

funcion para la hipotesis

donde la matriz de diseño para este caso es, $A = [X \mid \mathbf{1}]$

Término de error:

$$\mathbf{E} = Y - \hat{Y}$$

el error puede estar dentro o fuera en otra funcion

Funciones de costo para regresión multivariada

$$SSE = \mathbf{E}'\mathbf{E}$$

funcion separada la de costum

$$MSE = \frac{SSE}{qm}$$

$$RMSE = \sqrt{MSE}$$

Función gradiente para regresión multivariada

$\frac{\partial E}{\partial \theta_{i,j}} = -2 \sum_{p=1}^q e_{p,j} x_{p,i}$, para $i=1,\dots,n+1$; $j=1,\dots,m$. Asimismo $x_{p,n+1} = 1$

$$\frac{\partial E}{\partial \Theta} = 2(-A')\mathbf{E}$$

$$\frac{\partial E}{\partial \Theta} = 2 J'_{\mathbf{E}} \mathbf{E}$$

$$\text{donde } J_{\mathbf{E}} = - \begin{bmatrix} A_{q \times \varrho} & \mathbf{0}_{q \times \varrho} & \cdots & \mathbf{0}_{q \times \varrho} \\ \mathbf{0}_{q \times \varrho} & A_{q \times \varrho} & \cdots & \mathbf{0}_{q \times \varrho} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{q \times \varrho} & \mathbf{0}_{q \times \varrho} & \cdots & A_{q \times \varrho} \end{bmatrix}$$

este es para mi algoritmo, el calcular la matrzi jabociana

metricas que miden que la regesion se realizo correctamente

Métricas de regresión para evaluar la precisión

- a. Coeficiente de determinación (R^2)

$$R^2 = 1 - \frac{\sum_{p=1}^q (y_p - \hat{y}_p)^2}{\sum_{p=1}^q (y_p - \bar{y})^2}$$

- b. Coeficiente de determinación (R_{ajus}^2)

$$R_{ajus}^2 = 1 - \left(\frac{q - 1}{q - \varrho - 1} \right) (1 - R^2)$$

donde q es el número de datos de entrenamiento y ϱ es el número de parámetros del modelo.

funcion de hipotesi, funcion de costo, dependiendo del algoritmo utilizamos el gradiente o la jacobiana pero usamos la matriz