Teoría de Ajuste por Mínimos Cuadrados

Introducción

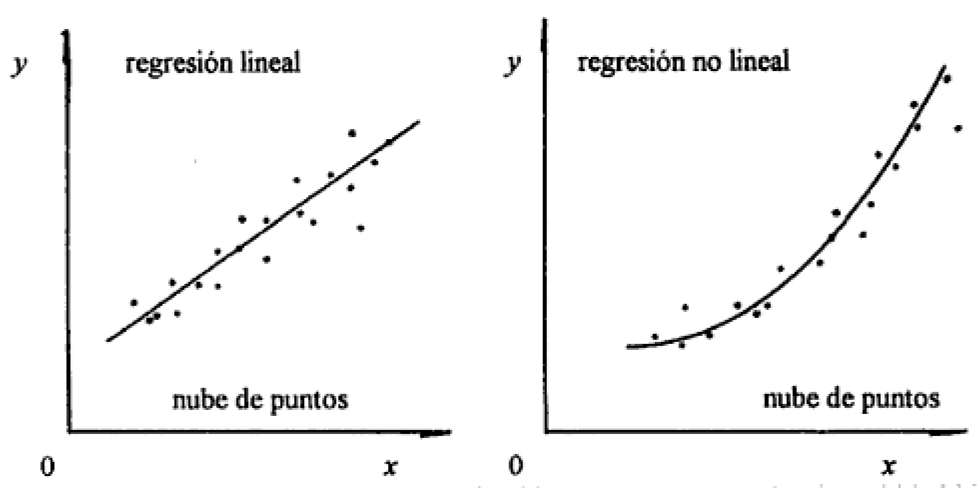
En muchas áreas de la ingeniería es necesario investigar si existe algún tipo de relación entre 2 o más variables. Esas relaciones se deben apoyar en fundamentos teóricos que establezcan una relación causa – efecto para el fenómeno en estudio. La búsqueda de la relación se realiza por medio de herramientas estadísticas como regresión y correlación.

Regresión: Se refiere a la obtención de la ecuación matemática que permite relacionar a la variable dependiente "y" con otra/s variable/s "x" (variable/s independiente/s) que son conocidas y por lo tanto permiten estimar valores de "y" a partir de los de "x".

Correlación: La correlación mide o cuantifica el grado de dependencia o asociación entre la variable dependiente "y" y las variables independientes "x". Se representa numéricamente por el llamado coeficiente de correlación.

Diagrama de Dispersión y modelos de regresión: Un diagrama de dispersión es una gráfica en el plano cartesiano "xy" en la que se muestran los valores experimentales del fenómeno en estudio. En base a ese diagrama, se puede detectar si los valores siguen un modelo de regresión de tipo Lineal o No Lineal. Un diagrama de dispersión permite identificar también aquellos valores más dispersos que por lo general se pueden deber a errores en la muestra.[1]

Diagramas de dispersión y modelos de regresión



[1]Introducción a los métodos numéricos: software en Basic y aplicaciones en Hidrología Superficial. Daniel Francisco Campos Aranda

Mínimos Cuadrados

El principio de Mínimos Cuadrados es hoy uno de los modelos matemáticos que se pueden usar para realizar Machine Learning (disciplina del campo de la Inteligencia Artificial que, a través de algoritmos, dota a las computadoras de la capacidad de identificar patrones en datos masivos y elaborar predicciones -análisis predictivo-) y establece que de todas las rectas o curvas que representan a una nube de puntos la que tiene la suma minima de los cuadrados de las distancias en cada punto a tal recta o curva es la de mejor ajuste posible. Se usan los cuadrados de las distancias porque de esa manera no importa si el punto está por arriba o por debajo de la curva propuesta.

Supongamos entonces tener un conjunto de datos experimentales

x	x_1	x_2	$\ldots$	x_{n-1}	x_n
y	$y(x_1)$	$y(x_2)$	$\ldots$	$y(x_{n-1})$	$y(x_n)$

1

y además suponemos conocida la *forma* de una función empírica:

$$y = f(x, a_1, a_2, ..., a_n) \quad (2)$$

Donde $a_1, a_2, ..., a_n$ son coeficientes que deseamos determinar.
Las desviaciones, o el error, entre la ordenada de la formula empírica **2** y el valor de la tabla **1** se indica como:

$$\varepsilon_i = f(x_i, a_1, a_2, ..., a_n) - y_i$$

Ahora, aplicando el método de los mínimos cuadrados, definimos los coeficientes $a_1, a_2, ..., a_n$ de manera tal que la suma del cuadrado de las desviaciones

$$\delta(a_1, a_2, \ldots, a_n) = \sum_{i=1}^{n} \left[f(x_i, a_1, a_2, \ldots, a_n) - y_i \right]^2 \quad (3)$$

Sea mínimo.

Aplicando las condiciones que se tiene que satisfacer para un valor extremo de la función de varias variables

$$\delta(a_1, a_2, ..., a_n)$$

en

$$\frac{\partial \delta}{\partial a_1} = 0 \ , \ \frac{\partial \delta}{\partial a_2} = 0 \ \ \ldots\ldots\ldots \ \frac{\partial \delta}{\partial a_n} = 0 \quad (4)$$

Si este sistema tiene solución y es única, esta es la buscada.

Si la función empírica es ***lineal*** respecto a los coeficientes $a_1, a_2, ..., a_n$ lo podemos expresar de la forma

$$f(x_1, a_1, a_2, ..., a_n) = a_1 \cdot \phi_1(x) + a_2 \cdot \phi_2(x) + ... + a_n \cdot \phi_n(x)$$

y la función $\delta(a_1, a_2, ..., a_n)$ toma la forma

$$\delta(a_1, a_2, \ldots, a_n) = \sum_{i=1}^{n} [\underbrace{a_1 \cdot \phi_1(x_i) + a_2 \cdot \phi_2(x_i) + \ldots. + a_n \cdot \phi_n(x_i)}_{f(x_i, a_1, a_2, ..., a_n)} - y_i]^2$$

y desarrollando el sistema **4** tenemos:

$$\begin{aligned}
\frac{\partial \delta}{\partial a_1} = 0 &\rightarrow 2 \cdot \sum_{i=1}^{n} [a_1 \cdot \phi_1(x_i) + a_2 \cdot \phi_2(x_i) + \ldots. + a_n \cdot \phi_n(x_i) - y_i] \cdot \phi_1(x_i) = 0 \\
\frac{\partial \delta}{\partial a_2} = 0 &\rightarrow 2 \cdot \sum_{i=1}^{n} [a_1 \cdot \phi_1(x_i) + a_2 \cdot \phi_2(x_i) + \ldots. + a_n \cdot \phi_n(x_i) - y_i] \cdot \phi_2(x_i) = 0 \qquad (5) \\
&\vdots \\
\frac{\partial \delta}{\partial a_n} = 0 &\rightarrow 2 \cdot \sum_{i=1}^{n} [a_1 \cdot \phi_1(x_i) + a_2 \cdot \phi_2(x_i) + \ldots. + a_n \cdot \phi_n(x_i) - y_i] \cdot \phi_n(x_i) = 0
\end{aligned}$$

Desarrollando, tenemos:

$$\begin{aligned}
\frac{\partial \delta}{\partial a_1} = 0 &\rightarrow a_1 \cdot \sum_{i=1}^{n} \phi_1(x_i) \cdot \phi_1(x_i) + \ldots.. + a_n \cdot \sum_{i=1}^{n} \phi_n(x_i) \cdot \phi_1(x_i) - \sum_{i=1}^{n} y_i \cdot \phi_1(x_i) = 0 \\
\frac{\partial \delta}{\partial a_2} = 0 &\rightarrow a_1 \cdot \sum_{i=1}^{n} \phi_1(x_i) \cdot \phi_2(x_i) + \ldots.. + a_n \cdot \sum_{i=1}^{n} \phi_n(x_i) \cdot \phi_2(x_i) - \sum_{i=1}^{n} y_i \cdot \phi_2(x_i) = 0 \qquad (6) \\
&\vdots \\
\frac{\partial \delta}{\partial a_n} = 0 &\rightarrow a_1 \cdot \sum_{i=1}^{n} \phi_1(x_i) \cdot \phi_n(x_i) + \ldots.. + a_n \cdot \sum_{i=1}^{n} \phi_n(x_i) \cdot \phi_n(x_i) - \sum_{i=1}^{n} y_i \cdot \phi_n(x_i) = 0
\end{aligned}$$

Análisis de los casos particulares

La función empírica es una recta

$$y = a_1 + a_2 x = y_{ajuste} = y_{pronostico}$$

entonces $\phi_1(x) = 1 \wedge \phi_2(x) = x$

En este caso, tenemos de **6**:

$$a_1 \cdot \sum_{i=1}^{n} 1 + a_2 \cdot \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$a_1 \cdot \sum_{i=1}^{n} x_i + a_2 \cdot \sum_{i=1}^{n} x_i^{\,2} = \sum_{i=1}^{n} y_i \cdot x_i$$

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^{\,2} \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i \cdot x_i \end{bmatrix}$$

Este sistema lo usaremos luego como referencia para todos los casos

donde $y = a_1 + a_2 \cdot x$ es la función de ajuste y de pronostico.

Ajuste polinómico de segundo orden

$$y = a_1 + a_2 x + a_3 x^2 = y_{ajuste} = y_{pronostico}$$

entonces $\phi_1(x) = 1$, $\phi_2(x) = x \;\; \wedge \; \phi_3(x) = x^2$

En este caso, de **6** tenemos:

$$a_1 \cdot \sum_{i=1}^{n} 1 + a_2 \cdot \sum_{i=1}^{n} x_i + a_3 \cdot \sum_{i=1}^{n} x_i^{\,2} = \sum_{i=1}^{n} y_i$$

$$a_1 \cdot \sum_{i=1}^{n} x_i + a_2 \cdot \sum_{i=1}^{n} x_i^{\,2} + a_3 \cdot \sum_{i=1}^{n} x_i^{\,3} = \sum_{i=1}^{n} y_i \cdot x_i$$

$$a_1 \cdot \sum_{i=1}^{n} x_i^{\,2} + a_2 \cdot \sum_{i=1}^{n} x_i^{\,3} + a_3 \cdot \sum_{i=1}^{n} x_i^{\,4} = \sum_{i=1}^{n} y_i \cdot x_i^{\,2}$$

$$\begin{bmatrix} n & \sum x_i & \sum (x_i)^2 \\ \sum x_i & \sum (x_i)^2 & \sum (x_i)^3 \\ \sum (x_i)^2 & \sum (x_i)^3 & \sum (x_i)^4 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i \cdot x_i \\ \sum y_i \cdot x_i^2 \end{bmatrix}$$

Ajuste polinómico de grado k

La función empírica es de la forma

$$y = a_1 + a_2 x + a_3 x^2 + ... + a_{k+1} x^k = y_{ajuste} = y_{pronostico}$$

entonces $\phi_1(x) = 1$, $\phi_2(x) = x$ .... $\phi_k(x) = x^{k-1} \wedge \ \phi_{k-1}(x) = x^k$

En este caso, tenemos de **6**:

$$a_1 \cdot \sum_{i=1}^{n} 1 + a_2 \cdot \sum_{i=1}^{n} x_i + ...+ a_k \cdot \sum_{i=1}^{n} x_i^{k-1} + a_{k+1} \cdot \sum_{i=1}^{n} x_i^k = \sum_{i=1}^{n} y_i$$

$$a_1 \cdot \sum_{i=1}^{n} x_i + a_2 \cdot \sum_{i=1}^{n} x_i^2 + ...+ a_k \cdot \sum_{i=1}^{n} x_i \cdot x_i^{k-1} + a_{k+1} \cdot \sum_{i=1}^{n} x_i \cdot x_i^k = \sum_{i=1}^{n} y_i \cdot x_i$$

$$\vdots$$

$$a_1 \cdot \sum_{i=1}^{n} x_i^k + a_2 \cdot \sum_{i=1}^{n} x_i \cdot x_i^k + ...+ a_k \cdot \sum_{i=1}^{n} x_i^k x_i^{k-1} + a_{k+1} \cdot \sum_{i=1}^{n} x_i^k \cdot x_i^k = \sum_{i=1}^{n} y_i \cdot x_i^k$$

$$\begin{bmatrix} n & \sum x_i & \cdots & \sum x_i^{k-1} & \sum x_i^k \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^k & \sum x_i^{k+1} \\ \vdots & \vdots & & \vdots & \vdots \\ \sum x_i^k & \sum x_i^{k+1} & \cdots & \sum x_i^{2k-1} & \sum x_i^{2k} \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i \cdot x_i \\ \vdots \\ \sum y_i \cdot x_i^k \end{bmatrix}$$

Ajuste exponencial

$$y = a \cdot e^{b \cdot x} = y_{pronostico}$$

En este caso la función empírica no es lineal, por lo que debemos linealizarla aplicando Ln a ambos lados de la ecuación

$$\ln y = \ln a + bx \cdot \ln e$$

$$\ln y = \ln a + bx = y_{ajuste} \quad (1)$$

Comparando la expresión **1** con la fórmula empírica de una recta:

$$y = a_1 + a_2 \cdot x \qquad \ln y = \ln a + bx$$

$$\begin{aligned} y &\longleftrightarrow \ln y \\ a_1 &\longleftrightarrow \ln a \\ a_2 &\longleftrightarrow b \\ x &\longleftrightarrow x \end{aligned}$$

Por lo que el sistema de ecuaciones queda:

$$\boxed{\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix} \cdot \begin{bmatrix} \ln a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \ln y_i \\ \sum_{i=1}^{n} x_i \cdot \ln y_i \end{bmatrix}}$$

Ajuste Potencial

$$y = a \cdot x^b = y_{pronostico}$$

En este caso la función empírica no es lineal como en el caso anterior, por lo que debemos linealizarla aplicando Ln a ambos lados de la ecuación

$$\ln y = \ln a + b \ln x = y_{ajuste}$$

Haciendo una analogía con la ecuación de una recta nos queda:

$$\boxed{\begin{bmatrix} n & \sum_{i=1}^{n} \ln x_i \\ \sum_{i=1}^{n} \ln x_i & \sum_{i=1}^{n} (\ln x_i)^2 \end{bmatrix} \cdot \begin{bmatrix} \ln a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \ln y_i \\ \sum_{i=1}^{n} \ln x_i \cdot \ln y_i \end{bmatrix}}$$

Ajuste del cociente

$$y = a \cdot \frac{x}{b+x} = y_{pron}$$

En este caso la función empírica no es lineal, por lo que debemos linealizarla en este caso invirtiendo las ecuaciones:

$$\frac{1}{y} = \frac{1}{a} \cdot \frac{b+x}{x} \longrightarrow \frac{1}{y} = \frac{b}{a} \cdot \frac{1}{x} + \frac{1}{a}$$

Ordenando

$$\frac{1}{y}=\frac{1}{a}+\frac{b}{a}\cdot\frac{1}{x}=y_{ajuste}$$

Comparando con la función lineal

$$\begin{bmatrix} n & \sum_{i=1}^{n}\left({}^{1}\!/_{x_i}\right) \\ \sum_{i=1}^{n}\left({}^{1}\!/_{x_i}\right) & \sum_{i=1}^{n}\left({}^{1}\!/_{x_i}\right)^2 \end{bmatrix}\cdot\begin{bmatrix} {}^{1}\!/_{a} \\ {}^{b}\!/_{a} \end{bmatrix}=\begin{bmatrix} \sum_{i=1}^{n}\left({}^{1}\!/_{y_i}\right) \\ \sum_{i=1}^{n}\left({}^{1}\!/_{x_i}\right)\cdot\left({}^{1}\!/_{y_i}\right) \end{bmatrix}$$

Bondad del Ajuste:

La bondad del ajuste es un parámetro que nos permite estimar si el ajuste realizado con respecto a los datos experimentales que teníamos ha sido efectivo o no. Se podría obtener arbitrariamente, por ejemplo, un ajuste lineal para una nube de puntos dada, y sin embargo podría suceder que los valores no tuvieran precisamente una distribución lineal, entonces con el cálculo de la bondad del ajuste se puede verificar si el ajuste realizado ha sido realmente efectivo.

Se calcula de la siguiente forma:

$$r^2=(ST-SR)/ST$$

Donde:

ST = Suma total alrededor de la media o del valor promedio
SR= Suma total de Regresión

La bondad del ajuste $r^2=(ST-SR)/ST$ va a ser un valor numérico siempre entre 0 y 1. Se considera que una bondad de ajuste mayor a 0.85 es un buen ajuste que representa a los datos analizados en el corto plazo.

De acuerdo al caso a resolver se plantean de diferentes maneras sus ecuaciones.

1) Caso Lineal y Polinómicos (todos los grados)

$$ST = \sum_{i=1}^{n} (\mathrm{y}_i - y_{media})^2$$

$$y_{media} = (\sum_{i=1}^{n} \mathrm{y}_i) / n$$

$$SR = \sum_{i=1}^{n} (\mathrm{y}_i - y_{Ajuste})^2$$

2) Casos Exponencial y Potencial

$$ST = \sum_{i=1}^{n} (\mathrm{Ln}(\mathrm{y}_i) - y_{media})^2$$

$$y_{media} = (\sum_{i=1}^{n} \mathrm{Ln}(\mathrm{y}_i)) / n$$

$$SR = \sum_{i=1}^{n} (\mathrm{Ln}(\mathrm{y}_i) - y_{Ajuste})^2$$

3) Caso del Cociente (o crecimiento saturado)

$$ST = \sum_{i=1}^{n} ((1/\mathrm{y}_i) - y_{media})^2$$

$$y_{media} = (\sum_{i=1}^{n} (1/\mathrm{y}_i)) / n$$

$$SR = \sum_{i=1}^{n} ((1/\mathrm{y}_i) - y_{Ajuste})^2$$