

# MÉTODOS NUMÉRICOS APLICADOS A LA INGENIERÍA

Responsable: Ing. Salomón Ortiz Quintanilla

Miembro: Ing. Jorge Segura Davila

Miembro: Ing. Romulo Cerdeña Velez

## RESUMEN

Es común que los datos se dan como valores discretos, y es posible que se requiera estimar el valor de un punto entre estos valores discretos, realizar predicciones o probar hipótesis. Además, se puede necesitar una versión simplificada de una función complicada. Las aplicaciones para esto se conocen como ajuste de curvas, porque en general no se conoce la curva exacta y tan solo se halla una curva aproximada. Existen dos métodos:

- Regresión por mínimos cuadrados.
- Interpolación.

## ABSTRAC

It is common that the data are given like discreet values, and it is possible that you require to be considered the value of a point among these discreet values, to carry out predictions or to prove hypothesis. One can also need a simplified version of a complicated function. The applications for this are known like adjustment of curved. This is because in general the exact and so alone curve is not known he/she is an approximate curve. Two methods exist:

- Regression for square minima.
- Interpolation.

## I. REGRESIÓN POR MÍNIMOS CUADRADOS

### Regresión Lineal

La regresión lineal es un técnica para determinar la mejor línea recta que pasa entre un conjunto de observaciones definidas por puntos  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ . La ecuación puede expresarse como:

$$y = a_0 + a_1x + e$$

Donde:

- Y** Es el valor verdadero
- $a_0$  y  $a_1$**  son la ordenada al origen y la Pendiente de la línea recta Respectivamente.
- e** es el error o diferencia entre el modelo y las observaciones, el cual se representa al reordenar la ecuación Como:  $e = y - a_0 - a_1x$
- $a_0 + a_1x$**  es el valor pronosticado de la variable Dependiente.

### Cuantificación del Error

La media aritmética ( $\bar{y}$ ) de una muestra se define como la suma de los datos individuales ( $y_i$ ) dividida entre el número de puntos ( $n$ ), o:

$$\bar{y} = \frac{\sum y_i}{n} \quad \text{de } i = 1 \dots n$$

La desviación estándar ( $S_y$ ) es la medida más común del espaciamiento de una muestra alrededor de la media: si las mediciones están muy espaciadas alrededor de la media, la desviación estándar será grande; si están agrupadas cerca de ella, será pequeña.

Donde  $S_t$  es la suma total de los cuadrados de los residuos entre los datos y una sola estimación de la medida de tendencia central (la media).

$$S_y = \sqrt{\frac{S_t}{n-1}} \quad \text{donde} \quad S_t = \sum (y_i - \bar{y})^2$$

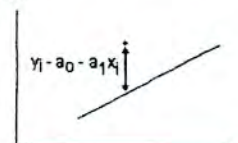
La varianza es el cuadrado de la desviación estándar.

$$S = \frac{S_t}{n-1}$$

### Cuantificación del Error de una Regresión Lineal

Se vio que la suma de los cuadrados de los residuos ( $S_r$ ) representa el cuadrado de la distancia vertical entre los datos y otra medida de tendencia central: la línea recta.

$$S_r = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - a_0 - a_1 X)^2$$



Una desviación estándar para la línea de regresión se puede determinar

$$S_{y/x} = \sqrt{\frac{S_r}{n-2}}$$

Donde  $S_{y/x}$  es llamado el error estándar del estimado, la notación del subíndice "y/x" designa que el error es para un valor predicho de  $y$  correspondiente a un valor particular de  $x$ ; asimismo, se divide entre  $n-2$  debido a los dos datos estimado ( $a_0$  y  $a_1$ ) que se usaron para calcular  $S_r$ ; el error estándar de la estimación cuantifica la dispersión de los datos. Sin embargo,  $S_{y/x}$  cuantifica la dispersión alrededor de la línea de regresión.

## Cuantificación de las Ventajas del Ajuste por Mínimos Cuadrados

El Coeficiente de correlación ( $r$ ) cuantifica la mejora o reducción del error originado por la representación de los datos por medio de una línea recta en vez de como un valor promedio. Antes de aplicar la regresión se calcula el  $S_t$  (suma total de los cuadrados alrededor de la media). Después de obtener la ecuación de la línea de regresión se calcula el  $S_r$  (suma de los cuadrados de los residuos alrededor de la recta de regresión), y que la magnitud de esta cantidad depende de la escala; la discrepancia se normaliza a  $S_t$  para obtener:

$$r = \sqrt{\frac{S_t - S_r}{S_t}}$$

$r^2$  = coeficiente de determinación. Para un ajuste perfecto  $S_r = 0$  y  $r^2 = 1$

Un coeficiente de correlación cercano a la unidad indica un buen ajuste, mientras que si es próximo a cero, el ajuste es pobre.

## Método de Mínimos Cuadrados para el Caso Polinomial

En la ingeniería, aunque algunos datos exhiben un patrón marcado, son pobremente representados por una línea recta, entonces una curva será la más adecuada para ajustarse a los datos; una alternativa es ajustar polinomios a los datos mediante regresión polinomial.

Como ya hemos mencionado anteriormente, los polinomios son muy usados en los cálculos numéricos, por sus propiedades. La ecuación de un polinomio de grado  $n$  es:

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n = \sum_{i=0}^n a_i x^i$$

Apliquémosle el método de mínimos cuadrados. La curva propuesta es:

$$y_p = a_0 + a_1x + a_2x^2 + \dots + a_nx^n + e$$

Donde  $a_i$  son coeficientes y  $e$  es el error. Una estrategia es minimizar la suma de los cuadrados de los residuos ( $S_r$ ), entre la  $y$  medida y la  $y$  calculada con el modelo lineal, está dada por:

$$S_r = \sum e_i^2 = \sum (y_{i, \text{medida}} - y_{i, \text{modelo}})^2 = \sum (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_nx_i^n)^2$$

Las derivadas parciales están dadas por:

Esto es:

$$\frac{\partial}{\partial a_0} S_r = \frac{\partial}{\partial a_0} \sum (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_nx_i^n)^2$$

$$\frac{\partial}{\partial a_0} S_r = -2 \sum (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_nx_i^n)$$

$$\frac{\partial}{\partial a_1} S_r = \frac{\partial}{\partial a_1} \sum (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_nx_i^n)^2$$

$$\frac{\partial}{\partial a_1} S_r = -2 \sum (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_nx_i^n) x_i$$

⋮

$$\frac{\partial}{\partial a_n} S_r = \frac{\partial}{\partial a_n} \sum (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_nx_i^n)^2$$

Y así sucesivamente hasta la  $n$  ecuación

Igualando a 0 las ecuaciones:

$$\frac{\partial}{\partial a_0} S_r = 0 \quad \frac{\partial}{\partial a_1} S_r = 0 \quad \dots \quad \frac{\partial}{\partial a_n} S_r = 0$$

Omitiendo los pasos siguientes, reordenando, para desarrollar el siguiente sistema de ecuaciones normales:

$$a_0 (m) + a_1 \sum x_i + a_2 \sum x_i^2 + \dots + a_n \sum x_i^n = \sum y_i$$

$$a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 + \dots + a_n \sum x_i^{n+1} = \sum y_i x_i$$

$$a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 + \dots + a_n \sum x_i^{n+2} = \sum y_i x_i^2$$

$$a_0 \sum x_i^n + a_1 \sum x_i^{n+1} + a_2 \sum x_i^{n+2} + \dots + a_n \sum x_i^{2n} = \sum y_i x_i^n$$

Todas las sumatorias son desde  $i = 1$  hasta  $m$  (donde  $m$  es el número de puntos). Los coeficientes de las incógnitas se pueden evaluar de manera directa a partir de los datos observados. El sistema es lineal y puede resolverse por los métodos conocidos. El error estándar del estimado se formula como:

$$S_{y/x} = \sqrt{\frac{S_r}{m - (n + 1)}}$$

Esta cantidad es dividida entre  $m - (n + 1)$ , ya que  $(n + 1)$  coeficientes obtenidos de los datos ( $a_0, a_1, \dots, a_m$ ) se usaron para calcular  $S_r$ ; así hemos perdido  $n + 1$  grados de libertad. Podemos escribir el sistema de ecuaciones normales obtenido en la forma:

$$S_x a = S_{xy}$$

donde:

$S_x$ : Matriz de sumatorias de potencias de  $x$ .

$$S_x = \begin{bmatrix} m & \sum x & \sum x^2 & \dots & \sum x^n \\ \sum x & \sum x^2 & \sum x^3 & \dots & \sum x^{n+1} \\ \sum x^2 & \sum x^3 & \sum x^4 & \dots & \sum x^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum x^n & \sum x^{n+1} & \sum x^{n+2} & \dots & \sum x^{2n} \end{bmatrix}$$

$$a = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad S_{xy} = \begin{bmatrix} \sum y \\ \sum yx \\ \sum yx^2 \\ \vdots \\ \sum yx^n \end{bmatrix}$$

$a$  : Vector de coeficientes. Las constantes del polinomio.

$S_{xy}$  : Vector de sumatorias de potencias de  $x$  con  $y$ 's.

No es necesario memorizar estas ecuaciones, pues son fáciles de construir para cada grado  $n$ .

Para construir el sistema para un grado, los pasos son:

- Se construye el primer renglón.
- Se construye la primera columna.
- Se llenan los renglones tomando en cuenta que cada uno tiene  $n + 1$  columnas.
- Se escribe el vector de términos independientes tomando en cuenta que la máxima potencia en  $x$  es  $n$ .

El ajustar un polinomio a una serie de datos se conoce

como regresión polinomial.

### Ejemplo de determinar el mejor polinomio

Un problema que se nos presenta es la determinación de los precios proyectados de los minerales. En este caso examinaremos los precios del cobre del mercado de Londres durante un lapso de 10 años y proyectaremos para dos años con la mejor curva de ajuste.

Tabla de Precios del Cobre en Ctps \$/libra:

| Año    | 1995    | 1996    | 1997    | 1998   | 1999  | 2000   | 2001   | 2002   | 2003   | 2004    |
|--------|---------|---------|---------|--------|-------|--------|--------|--------|--------|---------|
| precio | 133,198 | 103,894 | 103,224 | 74,974 | 71,38 | 82,294 | 71,566 | 70,647 | 80,734 | 130,106 |

**Solución: Para grado 2.** El sistema de ecuaciones normales es

$$\begin{bmatrix} m & \sum x & \sum x^2 \\ \sum x & \sum x^2 & \sum x^3 \\ \sum x^2 & \sum x^3 & \sum x^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum yx \\ \sum yx^2 \end{bmatrix}$$

Tabla: Cálculos para análisis

| i    | $X_i$   | $Fx$     | $x^2$    | $x^3$     | $x^4$      | $xy$      | $yx^2$     | $(y/\text{med})^2$ | $(S_i)^2$   |
|------|---------|----------|----------|-----------|------------|-----------|------------|--------------------|-------------|
| 1)   | 1.0000  | 133.1980 | 1.0000   | 1.0000    | 1.0000     | 133.1980  | 133.1980   | 1278.054810        | 13.495332   |
| 2)   | 2.0000  | 103.8940 | 4.0000   | 8.0000    | 16.0000    | 207.7880  | 415.5760   | 41.549526          | 186.782089  |
| 3)   | 3.0000  | 103.2240 | 9.0000   | 27.0000   | 81.0000    | 309.6720  | 929.0160   | 33.360954          | 1.035729    |
| 4)   | 4.0000  | 133.1980 | 16.0000  | 64.0000   | 256.0000   | 532.7920  | 2131.1680  | 1278.054810        | 1796.910767 |
| 5)   | 5.0000  | 74.9740  | 25.0000  | 125.0000  | 625.0000   | 374.8700  | 1874.3500  | 505.085419         | 70.425507   |
| 6)   | 6.0000  | 71.3800  | 36.0000  | 216.0000  | 1296.0000  | 428.2800  | 2569.6799  | 679.546204         | 72.253632   |
| 7)   | 7.0000  | 82.2940  | 49.0000  | 343.0000  | 2401.0000  | 576.0580  | 4032.4059  | 229.646927         | 3.776468    |
| 8)   | 8.0000  | 71.5660  | 64.0000  | 512.0000  | 4096.0000  | 572.5280  | 4580.2241  | 669.883240         | 174.541168  |
| 9)   | 9.0000  | 70.6470  | 81.0000  | 729.0000  | 6561.0000  | 635.8230  | 5722.4073  | 718.299072         | 506.852020  |
| 10)  | 10.0000 | 130.1060 | 100.0000 | 1000.0000 | 10000.0000 | 1301.0600 | 13010.6003 | 1066.538330        | 605.47528   |
| suma | 55.0000 | 974.4810 | 385.0000 | 3025.0000 | 25333.0000 | 5072.0690 | 35398.6254 | 6500.018555        | 3431.54808  |

Las sumatorias requeridas son:

$$m = 10 \quad \sum x_i = 55 \quad \sum x^2 = 385$$

$$\sum y = 974,481 \quad \sum yx = 5072,069$$

$$\sum x^3 = 3025 \quad \sum x^4 = 25\,333$$

$$\sum yx^2 = 35\,398,6254$$

El sistema de ecuaciones es:

$$\begin{bmatrix} 10 & 55 & 385 \\ 55 & 385 & 3\,025 \\ 385 & 3\,025 & 25\,333 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 974,481 \\ 5\,072,069 \\ 35\,398,625 \end{bmatrix}$$

La solución es:

$$a = \begin{bmatrix} 160,1386 \\ -25,2452 \\ 1,9781 \end{bmatrix}$$

La ecuación de la parábola es:  $y_p = 160,1386 - 25,2452x + 1,9781x^2$

El error estándar del estimado con base en la regresión de polinomios es:

$$s_{y/x} = \sqrt{\frac{S_r}{m - (n - 1)}}$$

$$s_{y/x} = \sqrt{\frac{3431,54809}{10 - (2 - 1)}} = 22,1409$$

Los demás valores estadísticos del modelo son:

- Sumatoria de los Errores al cuadrado ( $S_r$ )  $\Rightarrow$  3431.548096
- Sumatoria de los Errores a la media ( $S_t$ )  $\Rightarrow$  6500.018555
- Cálculo del Desviación estándar ( $S_y$ )  $\Rightarrow$  26.874231
- Cálculo del Error estándar ( $\hat{S}_{y/x}$ )  $\Rightarrow$  22.140938
- Coefficiente de determinación ( $r^2$ )  $\Rightarrow$  0.472071
- Coefficiente de correlación ( $r$ )  $\Rightarrow$  0.687074

Se ha efectuado una simulación para diferentes grados del polinomio, obteniendo los siguientes resultados:

Lineal  $y = 116,6199 - 3,4858x$

Grado 2:  $y = 160,1386 - 25,2452x + 1,9781x^2$

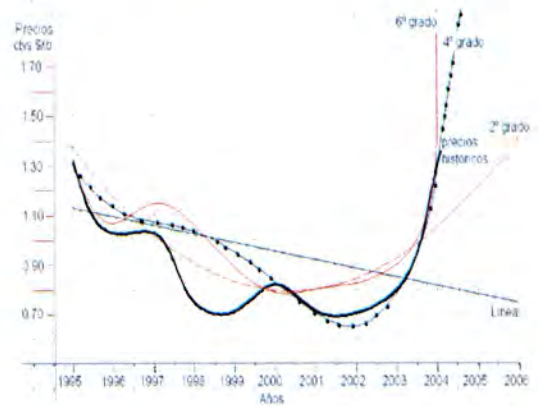
Grado 4  $y = 169,381 - 57,277x + 20,4536x^2 - 3,2716x^3 + 0,1758x^4$

Grado 6:  $y = 356,7043 - 434,68x + 283,5x^2 - 87,9229x^3 + 13,8692x^4 - 1,0806x^5 + 0,0331x^6$

Los factores de correlación y proyecciones para los dos años siguientes:

|                             | Lineal  | Grado 2  | Grado 4 | Grado 6 |
|-----------------------------|---------|----------|---------|---------|
| Año 2005                    | 78,2761 | 121,7915 | 233,6   | 519,25  |
| Año 2006                    | 74,79   | 142,0426 | 419,43  | 1573,71 |
| Coefficiente de correlación | 39,27 % | 68,7 %   | 84,8 %  | 90,75 % |

La gráfica correspondiente se aprecia



La curva que mejor está prediciendo los precios es la de cuarto grado, según se aprecia en especial en ajustar los 4 últimos años; sin embargo, a medida que aumenta el grado del polinomio el sistema de ecuaciones tiende a ser más inestable, y es más probable que más adelante existan oscilaciones.

Además, no es conveniente predecir para muchos años adelante, pues las respuestas serían demasiado erráticas.

## II. BIBLIOGRAFÍA

**Chapra, Steven y Canale, Raymond.** *Métodos numéricos para ingenieros.* McGraw-Hill Interamericana de México, S.A. De C.V. 4ta. edición 2003

**Gerald Curtis F.; Wheatley Patrick O.** *Applied numerical analysis.*, 4º edición, Addison Wesley Publishing Company. New York, 1984.

**Heath, Michael T.** *Scientific computing* McGraw-Hill, United States of America, 1997

**Hunt, B. R., Lipsman, R.L. and Rosenberg, J. M.** *A Guide to MATLAB. For Beginners and Experienced Users.* Cambridge University Press. 2000.