ANÁLISIS NUMÉRICO

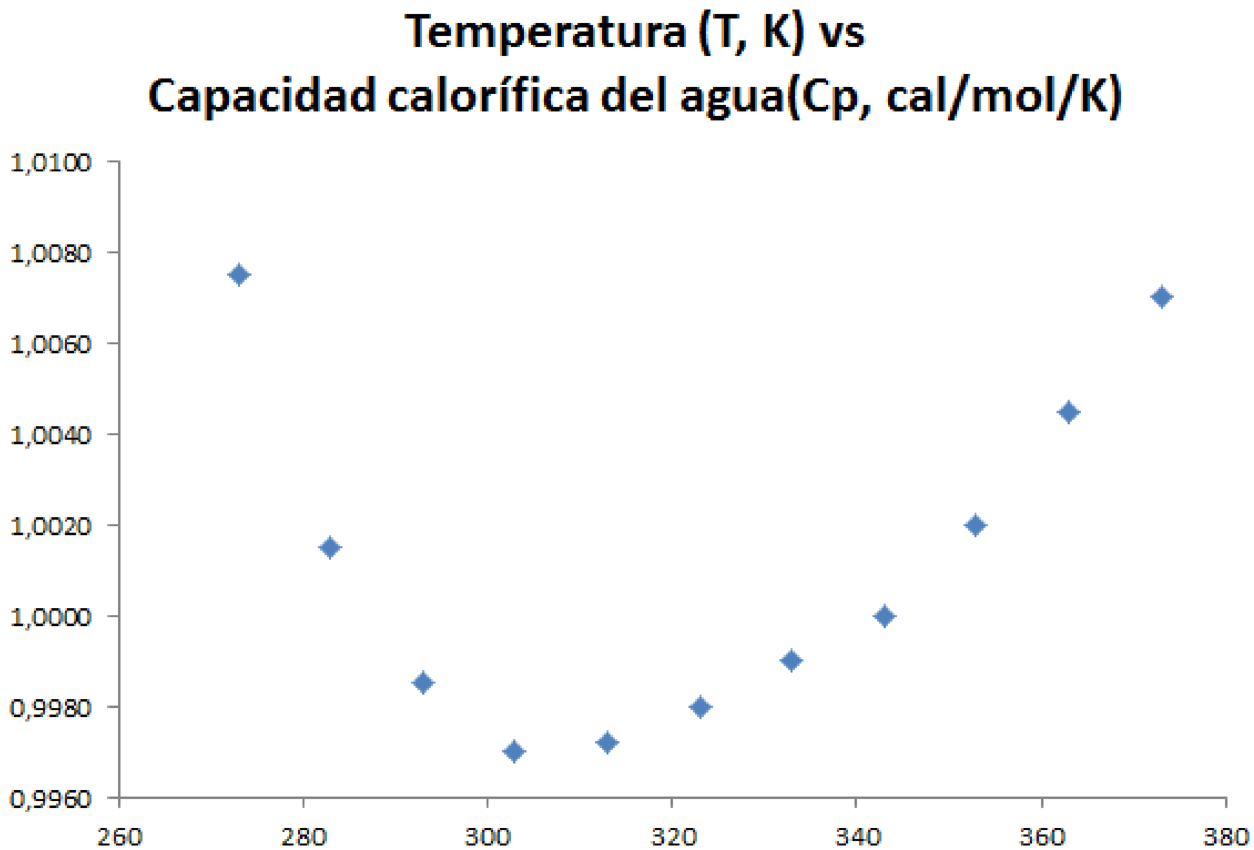
Mag. Carlos Alberto Ardila Albarracín

BLOQUE 2. AJUSTE DE CURVAS

2.3. REGRESIÓN CUADRÁTICA

REGRESIÓN CUADRÁTICA

Algunos datos científicos o de ingeniería, pueden presentar un patrón como este:



REGRESIÓN CUADRÁTICA

En estos casos, se ajusta mejor una curva a los datos y para ello se recomienda regresión polinomial

El procedimiento de mínimos cuadrados
se puede extender fácilmente y
ajustar datos a un polinomio de grado m.

$$Y = a_o + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_m x^m$$

REGRESIÓN CUADRÁTICA

En este caso, la suma de los cuadrados es:

$$Sr = \sum_{i=1}^{n} \left(Yi - a_o - a_1 x_i - a_2 x_i^2 - a_3 x_i^3 - \ldots - a_m x_i^m \right)^2$$

Que a la larga nos lleva al siguiente conjunto de ecuaciones :

$$\begin{array}{llllll}
a_o n & + \; a_1 \sum x_i & + \; a_2 \sum x_i^2 & + \; \ldots & + \; a_m \sum x_i^m & = \sum y_i \\
a_o \sum x_i & + \; a_1 \sum x_i^2 & + \; a_2 \sum x_i^3 & + \; \ldots & + \; a_m \sum x_i^{m+1} & = \sum x_i \, y_i \\
a_o \sum x_i^2 & + \; a_1 \sum x_i^3 & + \; a_2 \sum x_i^4 & + \; \ldots & + \; a_m \sum x_i^{m+2} & = \sum x_i^2 \, y_i \\
\;\;\vdots & \;\;\vdots & \;\;\vdots & & \;\;\vdots & \;\;\vdots \\
a_o \sum x_i^m & + \; a_1 \sum x_i^{m+1} & + \; a_2 \sum x_i^{m+2} & + \; \ldots & + \; a_m \sum x_i^{2m} & = \sum x_i^m \, y_i
\end{array}$$

REGRESIÓN CUADRÁTICA

Entonces, el problema de determinar polinomios de grado m con mínimos cuadrados es equivalente a resolver un sistema de m+1 ecuaciones lineales simultáneas

Así como en la regresión lineal, el error en la regresión polinomial se puede cuantificar mediante el error estándar de aproximación:

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

Donde m es el grado del polinomio que queremos ajustar

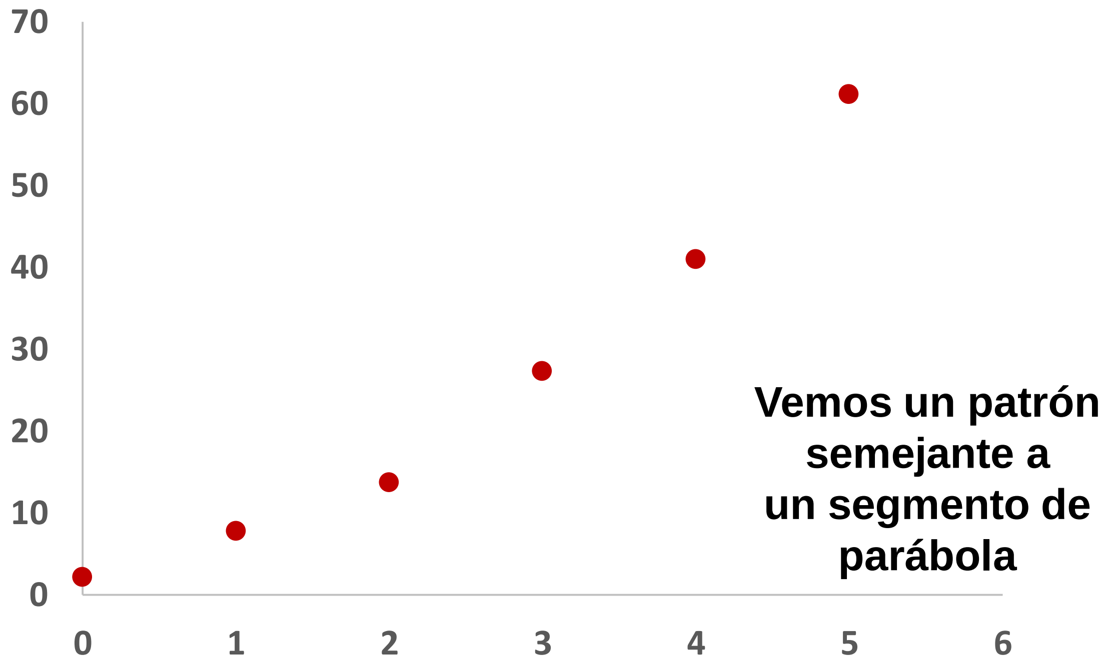
Además del error estándar, se puede calcular también el coeficiente de determinación, como en el caso lineal:

$$r^2 = \frac{S_t - S_r}{S_t}$$

REGRESIÓN CUADRÁTICA

Ejemplo: a partir de los datos de la tabla que se presenta a continuación, ajuste un polinomio de segundo grado, utilizando regresión polinomial.

Xi	Yi
0	2,1
1	7,7
2	13,6
3	27,2
4	40,9
5	61,1



Para el caso que nos ocupa,
m = 2 (el grado del polinomio que necesitamos)
n = 6 (la cantidad de datos)

REGRESIÓN CUADRÁTICA

Y el conjunto general de ecuaciones queda instanciado de la siguiente manera:

$$a_0 n + a_1 \sum x_i + a_2 \sum x_i^2 = \sum y_i$$

$$a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 = \sum x_i y_i$$

$$a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 = \sum x_i^2 y_i$$

REGRESIÓN CUADRÁTICA

Xi	**Yi**	**Xi^2**	**Xi^3**	**Xi^4**	**XiYi**	**Xi^2 Yi**
0	2,1	0	0	0	0	0
1	7,7	1	1	1	7,7	7,7
2	13,6	4	8	16	27,2	54,4
3	27,2	9	27	81	81,6	244,8
4	40,9	16	64	256	163,6	654,4
5	61,1	25	125	625	305,5	1527,5
$\sum Xi$	**$\sum Yi$**	**$\sum Xi^2$**	**$\sum Xi^3$**	**$\sum Xi^4$**	**$\sum XiYi$**	**$\sum Xi^2Yi$**
15	**152,6**	**55**	**225**	**979**	**585,6**	**2488,8**

REGRESIÓN CUADRÁTICA

Por lo tanto, las ecuaciones lineales simultáneas son:

$$a_0 \, 6 + a_1 \, 15 + a_2 \, 55 = 152{,}6$$

$$a_0 \, 15 + a_1 \, 55 + a_2 \, 225 = 585{,}6$$

$$a_0 \, 55 + a_1 \, 225 + a_2 \, 979 = 2488{,}8$$

REGRESIÓN CUADRÁTICA

O en un "formato" más familiar:

$$6\,a_0 + 15\,a_1 + 55\,a_2 = 152{,}6$$
$$15\,a_0 + 55\,a_1 + 225\,a_2 = 585{,}6$$
$$55\,a_0 + 225\,a_1 + 979\,a_2 = 2488{,}8$$

Resolviendo ese sistema con alguna técnica como la eliminación gaussiana, se obtiene:

$a_2 = 1.86071$ $a_1 = 2.35929$ $a_0 = 2.47857$

El polinomio es: $1.86071x^2 + 2.35929x + 2.47857$

REGRESIÓN CUADRÁTICA

Debemos calcular Sr y St

Sr → para calcular el error estándar de aproximación en la regresión polinomial

St → para calcular el coeficiente de determinación

Xtrazo	2,5000
Ytrazo	25,4333

$a_2 = 1.86071$ $a_1 = 2.35929$ $a_o = 2.47857$

Xi	Yi	Xi^2	Xi^3	Xi^4	XiYi	Xi^2 Yi	$(Yi - Ytrazo)^2$	$(Yi - ao - a1xi - a2xi^2)^2$
0	2,1	0	0	0	0	0	544,4444	0,14332
1	7,7	1	1	1	7,7	7,7	314,4711	1,00286
2	13,6	4	8	16	27,2	54,4	140,0278	1,08158
3	27,2	9	27	81	81,6	244,8	3,1211	0,80491
4	40,9	16	64	256	163,6	654,4	239,2178	0,61951
5	61,1	25	125	625	305,5	1527,5	1272,1111	0,09439
∑Xi	**∑Yi**	**$\sum Xi^2$**	**$\sum Xi^3$**	**$\sum Xi^4$**	**∑XiYi**	**$\sum Xi^2Yi$**	**St**	**Sr**
15	**152,6**	**55**	**225**	**979**	**585,6**	**2488,8**	**2513,3933**	**3,74657**

REGRESIÓN CUADRÁTICA

$$S_{y/x} = \sqrt{\frac{Sr}{n-(m+1)}} = \sqrt{\frac{3.74657}{6-3}} = 1.1175$$

$$S_y = \sqrt{\frac{St}{n-1}} = \sqrt{\frac{2513.3933}{5}} = 22.4205$$

Aquí aplica el mismo criterio que dice:

Si $(S_{y/x} < S_y)$ entonces la aproximación se considera aceptable

$$r^2 = \frac{St - Sr}{St} = \frac{2513.3933 - 3.74657}{2513.3933} = 0.99851$$

El resultado indica que el modelo cuadrático explica el 99.851% de la incertidumbre original

---------- *FIN DEL DOCUMENTO*