

AJUSTES DE CURVAS

Método de Regresión Potencial

Ing. Yamil Armando Cerquera¹
Esp Sistemas U. Nacional de Colombia
Facultad de Ingeniería
Universidad Surcolombiana

CONTENIDO

Preámbulo.....	2
Introducción	2
Objetivos.....	3
Desarrollo del modelo.....	3
Ejercicio	5
Cómo proponer la curva a ajustar?.....	7
Otra forma de ajuste similar a la potencial	7
COEFICIENTE DE DETERMINACIÓN (Γ^2):	9
COEFICIENTE DE CORRELACIÓN (Γ):.....	9
ERROR ESTANDAR DE LA ESTIMACIÓN (S_e):.....	9
Ejemplo	9
Ejemplo	12
Ejemplo aplicado a ingeniería.....	13
Cinética de una reacción	13
Aproximación de una función matemática complicada.....	18
Pronósticos del número de lectores de Supermán	20
Resumen.....	21
RECURSOS BIBLIOGRÁFICOS	23
Bibliografía Básica:.....	23
Bibliografía Complementaria:	23
Bibliografía OnLine:	24

¹ Docente de planta. Universidad Surcolombiana. Escalafón Asociado. Programa Ingeniería Electrónica

Preámbulo

A lo largo de la profesión de un ingeniero, un físico, un matemático, frecuentemente se presentan ocasiones en las que deben ajustar curvas a un conjunto de datos representados por puntos. Las técnicas desarrolladas para este fin pueden dividirse en dos categorías generales: interpolación y regresión. Se considerará aquí la primera de estas dos categorías. Más aún, como la teoría de aproximación potencial es una de las técnicas utilizadas, será la que se considere en este trabajo.

Cuando se asocia un error sustancial a los datos, la interpolación potencial es inapropiada y puede llevar a resultados no satisfactorios cuando se usa para predecir valores intermedios. Los datos experimentales a menudo son de ese tipo. Una estrategia más apropiada en estos casos es la de obtener una función aproximada que ajuste “adecuadamente” el comportamiento o la tendencia general de los datos, sin coincidir necesariamente con cada punto en particular.

Una curva potencial puede usarse en la caracterización de la tendencia de los datos sin pasar sobre ningún punto en particular. Una manera de determinar la curva, es inspeccionar de manera visual los datos graficados y luego trazar la “mejor” curva a través de los puntos. Aunque este enfoque recurre al sentido común y es válido para cálculos a “simple vista” es deficiente ya que es arbitrario. Es decir, cada analista trazará curvas diferentes.

La manera de quitar esta subjetividad es considerar un criterio que cuantifique la suficiencia del ajuste. Una forma de hacerlo es obtener una curva que minimice la diferencia entre los datos y la curva y el método para llevar a cabo este objetivo es al que se le llama *regresión potencial*.

Introducción

El presente trabajo forma parte de los objetivos y contenidos de aprendizaje de la cátedra MÉTODOS NUMÉRICOS, que pretende desarrollar las habilidades para la utilización de los métodos lineales y estimación de mínimos cuadrados.

En este trabajo básicamente se habla de cómo desarrollar la aplicación de los métodos lineales y estimación por mínimos cuadrados, además de inferencia, predicción y correlación para ajustar datos a una curva del tipo ax^b .

Se desarrollan una serie de ejemplos mediante los cuales se trata de presentar la manera más sencilla de usar estos métodos.

Si se sabe que existe una relación entre una variable denominada dependiente y otras denominadas independientes (como por ejemplo las existentes entre: la experiencia profesional de los trabajadores y sus respectivos sueldos, las estaturas y pesos de personas, la producción agraria y la cantidad de fertilizantes utilizados, etc.), puede

darse el problema de que la dependiente asuma múltiples valores para una combinación de valores de las independientes.

La dependencia a la que hace referencia es relacional matemática y no necesariamente de causalidad. Así, para un mismo número de unidades producidas, pueden existir niveles de costo, que varían empresa a empresa.

Si se da ese tipo de relaciones, se suele recurrir a los estudios de regresión en los cuales se obtiene una nueva relación pero de un tipo especial denominado función, en la cual la variable independiente se asocia con un indicador de tendencia central de la variable dependiente. Cabe recordar que en términos generales, una función es un tipo de relación en la cual para cada valor de la variable independiente le corresponde uno y sólo un valor de la variable dependiente.

Objetivos

Entre los objetivos propuestos en este apartado se puede citar los siguientes:

1. Que sea fácilmente comprensible para los alumnos con un conocimiento mínimo de matemáticas;
2. Capacitar a los alumnos para que practiquen los métodos numéricos en una computadora;
3. Elaborar programas simples que puedan usarse de manera sencilla en aplicaciones científicas;
4. Proporcionar software que resulte fácil de comprender.

La importancia de los métodos numéricos ha aumentado de forma drástica en la enseñanza de la ingeniería y la ciencia, lo cual refleja el uso actual y sin precedentes de las computadoras.

El desarrollo de un programa siempre es importante en el aprendizaje de métodos numéricos. La presentación de resultados calculados con gráficos utilizando algún software, por ejemplo MATLAB, motiva a los alumnos para aprender métodos matemáticos y numéricos que de otra forma podrían resultar tediosos.

Desarrollo del modelo

En el caso de que la esperanza propia de cada medición siga un modelo de la forma

$y = ax^b$, la expresión a minimizar se reduce a $S(A, B) = \sum_{i=1}^n \frac{(y_i - ax^{b_i})^2}{\sigma_{y_i}^2}$, La

minimización de $S(A, B)$ es un problema sin solución analítica, por esto mismo se debe realizar a través de métodos numéricos;

Desafortunadamente en la práctica no siempre los polinomios ajustan bien una serie de datos. Una curva que aparece frecuentemente en la práctica es la potencial, representada mediante la ecuación:

$$y = ax^b$$

Si se les aplica el método de mínimos cuadrados se obtiene la curva propuesta así:

$$y_i = ax_i^b$$

Una estrategia en el ajuste de una línea óptima es el criterio de *mínimas*. En este método, la línea se escoge de tal manera que minimice la distancia máxima a la que se encuentra un punto de la línea recta. Esta estrategia está mal condicionada para regresión ya que influye de manera indebida sobre un punto externo, aislado, cuyo error es muy grande. Se debe notar que el criterio de mínimas, algunas veces está bien condicionado para ajustar una función simple a una función complicada.

Una estrategia que ignora las restricciones anteriores es la de minimizar la suma de los cuadrados de los residuos, S_r , de la siguiente manera:

$S_r = \sum_{i=1}^n E_i^2 = \sum_{i=1}^n (y_i - ax_i^b)^2$	Ec 1
--	-------------

Si S^2 está dada por:

$$S^2 = \sum (y_i - Ax_i^B)^2$$

Se obtiene las dos ecuaciones normales, derivando la función respecto de A y B así:

$$\frac{\partial S^2}{\partial a} \equiv \frac{\partial}{\partial a} \sum (y_i - ax_i^b)^2 = 2 \sum (y_i - ax_i^b)(-x_i^b)$$

$$\frac{\partial S^2}{\partial b} \equiv \frac{\partial}{\partial b} \sum (y_i - Ax_i^b)^2 = 2 \sum (y_i - Ax_i^b)(-Ax_i^b \ln(x_i))$$

Igualando a 0 y simplificando

$$\frac{\partial S^2}{\partial a} = 2 \sum (y_i - ax_i^b)(-x_i^b) = 0 \Rightarrow \sum ax_i^{2b} = \sum y_i x_i^b$$

$$\frac{\partial S^2}{\partial b} = 2 \sum (y_i - ax_i^b)(-ax_i^b \ln(x_i)) = 0 \Rightarrow \sum ax_i^{2b} \ln(x_i) = \sum y_i x_i^b \ln(x_i)$$

Se puede observar que las ecuaciones normales no son lineales, lo cual hace difícil el resolverla. Por esta razón en la práctica se usa un cambio de variable antes de aplicar el método de mínimos cuadrados.

El error estándar está dado por:

$$\sigma_{xy} = \sqrt{\frac{S^2}{M-2}}$$

S^2 Se calcula con la expresión respectiva $S^2 = \sum (y_i - ax_i^b)^2$

Para la función potencial se tiene

$$y = ax^b$$

$$\ln(y) = \ln(a) + a \ln(x)$$

Si

$$z = \ln(y)$$

$$w = \ln(x)$$

$$a_0 = \ln(a)$$

$$a_1 = b$$

Se obtiene

$$z = a_0 + a_1 w$$

Es importante regresar a las variables originales, después de haber hecho la regresión, ya que nos interesa el ajuste de y en función de x , no de $\ln(y)$ en función de x ó de $\ln(y)$ en función de $\ln(x)$.

Las ecuaciones normales son

$$\begin{bmatrix} M & \sum w \\ \sum w & \sum w^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum z \\ \sum zw \end{bmatrix} \quad \text{Ec 2}$$

O si prefiere expresarla en los términos que son:

$$\begin{bmatrix} M & \sum \ln(x) \\ \sum \ln(x) & \sum (\ln(x))^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum \ln(y) \\ \sum \ln(x) \ln(y) \end{bmatrix} \quad \text{Ec 3}$$

Ejercicio

Los siguientes datos se determinaron de una curva a la que intencionalmente se le agregó error en los valores de y . Esto se hizo con el fin de emular lo que pasa en la vida real. En la práctica cuando se ajusta una curva a una tabla, no esta trabajando con los valores reales, sino tan solo con aproximaciones. Esto se debe al error inherente que pueden tener los datos. Se observará con este ejemplo si la curva obtenida se aproxima a la curva real. Los datos son:

Tabla 1

n	x	$\ln(x)$	$(\ln(x))^2$	y	$\ln(x)\ln(y)$
1	0.05	-2,99573227355399	8.97441185481296	0.956	0.13480006134362
2	0.11	-2,20727491318972	4.87206254239669	0.890	0.25722216916002
3	0.15	-1,89711998488588	3.59906423705341	0.832	0.34892369195203
4	0.31	-1,17118298150295	1.37166957616213	0.717	0.38962849652956
5	0.46	-0,77652878949900	0.60299696092078	0.571	0.43514038549013
6	0.52	-0,65392646740666	0.42761982477496	0.539	0.40415252301731
7	0.70	-0,35667494393873	0.12721701563370	0.378	0.34699517236851
8	0.74	-0,30110509278392	0.09066427690041	0.370	0.29937442301583
9	0.82	-0,19845093872384	0.03938277508037	0.306	0.23499968324033
10	0.98	-0,02020270731752	0.00040814938296	0.242	0.02866395575669
11	1.17	0,15700374880966	0.02465017714029	0.104	-0.35535669255726
$\sum 11$		-10.42119534399254	20.13014739025865		2.52454386931678

Los datos anteriores calculados mediante MatLab con el siguiente código

```
x=[0.05 0.11 0.15 0.31 0.46 0.52 0.7 0.74 0.82 0.98 1.17];
log(x);
y=[0.956 0.890 0.832 0.717 0.571 0.539 0.378 0.370 0.306 0.242 0.104];
log(x).*log(y)
sum(ans)
```

De la tabla anterior se tiene que:

$$\begin{aligned}\sum w &= \sum \ln(x) = -10.4211953439925 \\ \sum w^2 &= \sum (\ln(x))^2 = 20.13014739025865 \\ \sum wz &= \sum \ln(x)\ln(y) = 2.52454386931678\end{aligned}$$

Y reemplazando los valores anteriores en el sistema de ecuaciones normales se tiene:

$$\begin{bmatrix} 11 & -10.421195343992 \\ -10.421195343992 & 20.1301473902587 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} -8.69000470253179 \\ 2.52454386931678 \end{bmatrix}$$

$$\begin{aligned}11a_0 - 10.4211953439925a_1 &= -8.69000470253179 \\ -10.4211953439925a_0 + 20.1301473902587a_1 &= 2.52454386931678\end{aligned}$$

Resolviendo

$$a = \begin{bmatrix} -1.31721974461162 \\ -5.565016580885040e-01 \end{bmatrix} \equiv \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

Si se retoma la ecuación potencial original correspondiente a $y = ax^b$, y al haber simplificado como $a_0 = \ln(a)$ y $a_1 = b$, entonces se tendría:

$$a = e^{a_0} = e^{-1.31721974461162} = 2.678790397399640E - 01$$

$$b = -5.565016580885040E - 01$$

Por ende la ecuación potencial que ajusta los datos de la tabla 1 será:

$$y = 0.2678790397399640 x^{-0.5565016580885040}$$

El error estándar cuadrado es

$$\sigma_{xy} = \sqrt{\frac{s^2}{m-2}} = \sqrt{\frac{3.34985653953140E - 01}{11-2}} = 1.929264838957270E - 01$$

```
x=[0.05 0.11 0.15 0.31 0.46 0.52 0.7 0.74 0.82 0.98 1.17];
log(x);
y=[0.956 0.890 0.832 0.717 0.571 0.539 0.378 0.370 0.306 0.242 0.104];
log(x).*log(y)
sum(ans)
a=0.2678790397399640;
b=-0.5565016580885040;
Y=a*x.^b;
plot(x,y,'or',x,Y,'b')
grid on
```

Cómo proponer la curva a ajustar?

La curva propuesta puede determinarse de las siguientes maneras:

Considerando la teoría. A veces la naturaleza física de los datos nos dice o al menos propone la forma de la curva. Por ejemplo si ajustamos datos de voltaje contra corriente la ecuación apropiada es lineal por la ley del Ohm.

Graficando. Si no se tiene una teoría que indique la forma de la curva, viendo la gráfica se puede tener una idea.

Por tanteo. Probando diversas curvas, la que obtenga un σ_{xy} más pequeño será la apropiada.

Otra forma de ajuste similar a la potencial

Como en el caso de la curva exponencial y de la potencial en los que conviene cambiar las variables antes de aplicar el método de mínimos cuadrados, puede ocurrir que en otros casos se deba hacer también. Por ejemplo la siguiente ecuación se usa en óptica

$$y = \frac{x}{Ax + B}$$

Si se desea ajustar por el método de mínimos cuadrados el sistema de ecuaciones normales obtenido será no lineal. Esta curva puede expresarse de varias maneras. Por ejemplo si se considera los recíprocos

$$\frac{1}{y} = \frac{Ax + B}{x} \equiv A + \frac{B}{x}$$

Cambiando las variables

$$Radio1 = \frac{1}{y}$$

$$Sonidoz = \frac{1}{x}$$

$$A = a_0$$

$$B = a_1$$

Se obtiene una recta.

$$Radio1 = a_0 + a_1 Sonidoz$$

Por otro lado también se puede hacer

$$y(Ax + B) = x$$

$$Ax + B = \frac{x}{y}$$

Cambiando variables

$$Kebuena = \frac{x}{y}$$

$$B = a_0$$

$$A = a_1$$

Con lo que se obtiene otra recta.

$$Kebuena = a_0 + a_1 x$$

Es oportuno recordar que en cada caso, el error estándar cuadrado se calcula con las ecuaciones originales.

Existen casos donde es virtualmente imposible realizar un cambio de variable que simplifique el sistema. En estos casos no hay más remedio que resolver el sistema de ecuaciones normales no lineal que se obtiene. Este problema se denomina *regresión no lineal* y es difícil de resolver.

COEFICIENTE DE DETERMINACIÓN (Γ^2):

Este Coeficiente sirve para medir la relación entre las variables, medida de ajuste de modelo de Regresión y que corresponde al cuadrado del Coeficiente de correlación simple (Γ). El cálculo del Coeficiente de correlación puede efectuarse de manera directa, mediante la siguiente fórmula:

$$\Gamma^2 = \frac{[n\sum xy - \sum x \sum y]^2}{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}$$

COEFICIENTE DE CORRELACIÓN (Γ):

Se dice que existe correlación entre dos variables, cuando al variar una de ellas varía también la otra variable. Para que la Proyección sea más acertada es necesario que el número de observaciones (n) sea más amplio. A mayores años estudiados, tiene más relevancia estadística el valor de este Coeficiente.

$$\Gamma = \frac{[n\sum xy - \sum x \sum y]}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

El grado de aproximación entre variables es mayor cuando el Coeficiente de Correlación se acerca al valor máximo de 1. Entonces, en este caso se dice, existe una elevada Correlación entre X y Y.

ERROR ESTANDAR DE LA ESTIMACIÓN (S_e):

El error de la estimación es una medida que permite mostrar el nivel de confiabilidad que tiene la ecuación de predicción e indica hasta que punto los valores observados difieren de sus valores Históricos alrededor de la línea de Regresión. Cuando "Se" se aproxima a cero, entonces la ecuación de Regresión empleada será un estimador óptimo de la variable dependiente. Para el cálculo directo se puede utilizar la siguiente fórmula:

$$S_e = \sqrt{\frac{\sum y^2 - a\sum y - b\sum xy}{n-2}}$$

Ejemplo

Con los siguientes datos Históricos, Proyectar la Demanda mediante Regresión potencial

Año	x	y
1993	1	20000
1994	2	35000

1995	3	45000
1996	4	40000
1997	5	55000
1998	6	55000

Por el método no lineal de Regresión potencial se tiene las siguientes relaciones y se construye la siguiente tabla:

$$y = ax^b$$

$$\ln(y) = \ln(a) + b * \ln(x)$$

Ahora si se reemplaza,

$$Y = \ln(y)$$

$$A = \ln(a)$$

$$X = \ln(x)$$

Se tendría la siguiente ecuación

$$Y = A + bX$$

Año	x	y	X = ln(x)	Y = ln(y)	X ²	Y ²	X * Y
1993	1	20000	0.0000	9.9035	0.0000	98.0791	0.0000
1994	2	35000	0.6931	10.4631	0.4805	109.4765	7.2525
1995	3	45000	1.0986	10.7144	1.2069	114.7987	11.7710
1996	4	40000	1.3863	10.5966	1.9218	112.2887	14.6901
1997	5	55000	1.6094	10.9151	2.5903	119.1392	17.5672
1998	6	55000	1.7918	10.9151	3.2104	119.1392	19.5572
Σ			6.5793	63.5078	9.4099	672.9213	70.8379

Aplicando la fórmula de Regresión lineal (mínimos cuadrados):

$$A = \frac{\sum Y \sum X^2 - \sum X \sum XY}{n \sum X^2 - (\sum X)^2}; \quad b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2}$$

$$A = \frac{63.5078 * 9.4099 - 6.5793 * 70.8379}{6 * 9.4099 - 6.5793^2}; \quad b = \frac{6 * 70.8379 - 6.5793 * 63.5078}{6 * 9.4099 - 6.5793^2}$$

$$A = 9,9860415732118774639992096513185$$

$$b = 0,54588642571834924931295762346895$$

Reemplazando los valores de "A" y "b" en la ecuación general se tiene:

$$Y = 9.9860 + 0.54589 (\text{Log}(x))$$

Recuerde que los valores Históricos de "X", corresponden a los valores del logaritmo de

"x", por lo tanto para efectuar estimaciones, lo primero que se debe hacer es expresar en términos logarítmicos el valor del año de 1999 (cumpliendo así una exigencia de la ecuación lineal logarítmica), valor que corresponde al periodo 7. Entonces: $\text{Log } 7 = 0.8451$ y así se procede para proyectar los posteriores años.

$$Y = 4.3355 + 0.5485 (\text{Log } 7) = 4.7985$$

Por último, para cuantificar la Demanda del año 1999, se encuentra el antilogaritmo de 4,7985, cuyo resultado final es:

$$Y = \text{Anti log } (4.7985) = e^{4.7985}$$

$$Y = 62.878 \text{ unidades}$$

Comprobando el Coeficiente de determinación y el grado de Correlación entre las variables "X" y "Y" y calculando también el error estándar de estimación se tiene:

$$\Gamma^2 = \frac{[n \sum xy - \sum x \sum y]^2}{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]} = 0.924$$

$$\Gamma = \frac{[n \sum xy - \sum x \sum y]}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} = 0.96$$

$$S_e = \sqrt{\frac{\sum y^2 - a \sum y - b \sum xy}{n-2}} = 4.964$$

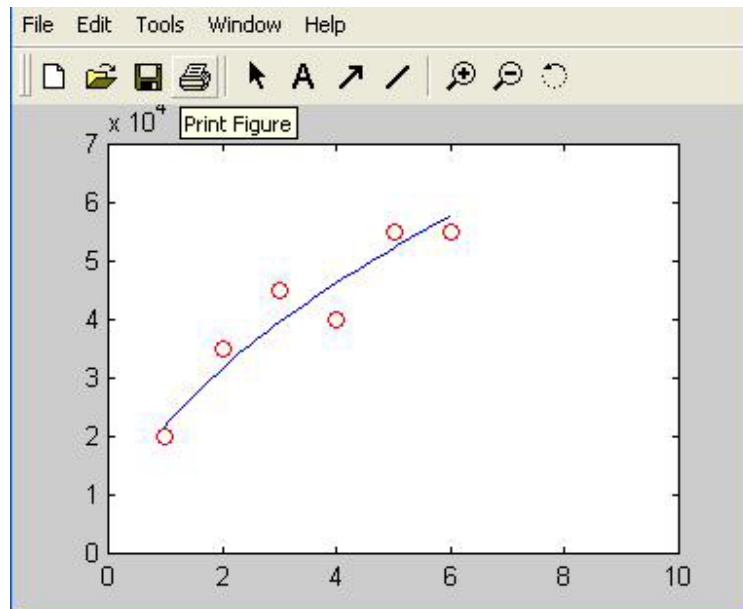
Como se puede apreciar ambos Coeficientes se aproximan a la unidad, lo que implica que la ecuación de Regresión potencial empleada es la que mejor se ajusta las variables.

En MatLab se tendría lo siguiente

```
x=1:6;
y=[20000 35000 45000 40000 55000 55000];
c=polyfit(log(x),log(y),1)
a=exp(c(2)); b=c(1);
xx=min(x):0.1:max(x);
yy=a*xx.^b;
plot(x,y,'or',xx,yy,'b')
axis([0 10 0 70000])
```

```
>a
a =
 2.1717e+004 igual a 21717
> b
b =
 0.5461
```

La ecuación potencial quedaría así: $21717x^{0.5461}$, que graficada junto con los datos discretos se vería de la siguiente forma:



Ejemplo

Con programación MatLab ajuste los siguientes datos para un ajuste de regresión potencial

En conjunto de datos está dado por:

```
x=[0.15 0.4 0.6 1.01 1.5 2.2 2.4 2.7 2.9 3.5 3.8 4.4 4.6 5.1 6.6 7.61];
y=[4.4964 5.1284 5.6931 6.2884 7.0989 7.5507 7.5106 8.0756 7.8708
  8.2403 8.5303 8.7394 8.9981 9.1450 9.5070 9.9115];
```

Ajustar los datos anteriores con la ecuación de potencia ax^b mediante el uso de MatLab.

```
>> x=[0.15 0.4 0.6 1.01 1.5 2.2 2.4 2.7 2.9 3.5 3.8 4.4 4.6 5.1 6.6 7.61];
>> y=[4.4964 5.1284 5.6931 6.2884 7.0989 7.5507 7.5106 8.0756 7.8708 8.2403 8.5303 8.7394 8.9981
9.1450 9.5070 9.9115];
>> c=polyfit(log(x),log(y),1)
c =
  0.2093  1.8588
```

Las constantes de la función de potencia serán:

$$a = \exp(c(2)) = e^{c(2)} \text{ y}$$

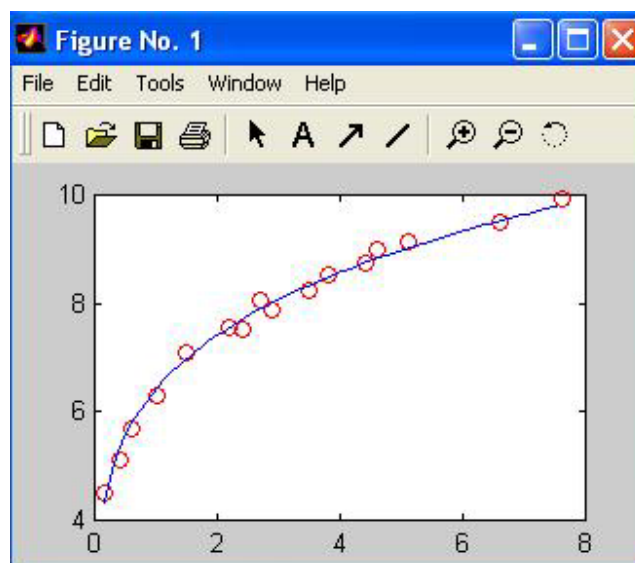
$$b = c(1)$$

Quedando de esta manera la ecuación de potencia que ajusta los datos anteriores así:

$$f(x) = ax^b = 6.4160x^{0.2093}$$

El conjunto de datos y la curva de potencia que ajusta los datos se grafican mediante el siguiente código en la figura

```
>> a=exp(c(2)); b=c(1);  
>> xx=min(x):0.1:max(x);  
>> yy=a*xx.^b;  
>> plot(x,y,'or',xx,yy,'b')
```



Los datos de la muestra se grafican de color rojo con círculos y los datos ajustados están representados mediante la línea de color azul.

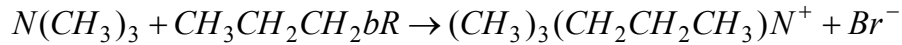
Ejemplo aplicado a ingeniería

A continuación se muestran algunas aplicaciones del ajuste de curvas.

Cinética de una reacción

En la industria química se elaboran productos que son de uso diario: lociones, jabones, perfumes, desodorantes, dulces, etc, etc, etc. Muchos de estos productos son sintéticos y se elaboran en equipos llamados reactores. Para diseñar y posteriormente construir un reactor se requiere información de como varía la concentración en función del tiempo de una reacción química determinada. Esta información es denominada cinética de la reacción. Para determinar dicha cinética se requiere medir en laboratorio datos de la concentración de algún reactivo para varios tiempos de reacción.

La trimetilamina y el bromuro de n-propilo es una reacción que se puede estudiar para diseñar reactores. La reacción es:



A continuación se muestra una tabla de la concentración de la trimetilamina en función del tiempo

Tabla 4 Concentración de trimetilamina en función del tiempo

t(min)	C(mol/l)
13	0.0888
34	0.0743
59	0.0633
120	0.0448

Para proponer el modelo apropiado se recurre a la teoría. La Fisicoquímica da la teoría necesaria. De acuerdo a la Fisicoquímica 2 modelos posibles en este caso son

$$C = Ae^{kt} \quad \text{Modelo 1}$$

$$C = \frac{1}{kt + A} \quad \text{Modelo 2}$$

Para el modelo 1

```
t=[13 34 59 120];
c=[0.0888 0.0743 0.0633 0.0448];
plot(t,c,'or')
axis([0 120 0 0.1])
A=0.09366307630583030;
k=-0.00625763329386314;
T=1:3:120;
C=A.*exp(k*T);
hold on
plot(T,C)
```

Se prueba ambos modelos para ver cual es más apropiado. Para el modelo exponencial $C = Ae^{kt}$ se hace un cambio de variable

$$C = Ae^{kt}$$

$$\ln(C) = \ln(A) + kt$$

Si se utiliza la siguiente nomenclatura

$$y = \ln(C)$$

$$a_0 = \ln(A)$$

$$a_1 = k$$

$$x = t$$

Se obtiene la ecuación

$$y = a_0 + a_1x$$

En estos términos las ecuaciones normales son

$$\begin{bmatrix} M & \sum x \\ \sum x & \sum x^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum yx \end{bmatrix}$$

O si lo prefiere

$$\begin{bmatrix} M & \sum t \\ \sum t & \sum t^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum \ln(C) \\ \sum t \ln(C) \end{bmatrix}$$

M	t(min)=x	$x^2 = t^2$	C(mol/l)	$\ln(C) = y$
1	13	169	0.0888	-2,421368628984012672056928546409
2	34	1156	0.0743	-2,5996443272584235712133469791029
3	59	3481	0.0633	-2,7598699498320065489347643163607
4	120	14400	0.0448	-3,1055471395611975744632203465452
$\sum 4$	226	19206		-10,886430045635640366668260188418

Calculando las sumatorias

$$\begin{aligned} M &= 4 \\ \sum x &= \sum t = 226 \\ \sum x^2 &= \sum t^2 = 19206 \\ \sum y &= \sum \ln(C) = -10.8864300456356 \\ \sum xy &= \sum t \ln(C) = -655.363683091011 \end{aligned}$$

Las ecuaciones son

$$\begin{bmatrix} 4 & 226 \\ 226 & 19206 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} -10.8864300456356 \\ -655.363683091011 \end{bmatrix}$$

$$4a_0 + 226a_1 = -10.8864300456356$$

$$226a_0 + 1920a_1 = -655.363683091011$$

Resolviendo

$$a = \begin{bmatrix} -2.36805123030564 \\ -6.257633293863140E-03 \end{bmatrix}$$

Regresando a la variable original

$$A = e^{a_0} = e^{-2.36805123030564} = 9.366307630583030E - 02$$

$$k = -6.257633293863140E - 03$$

La ecuación es

Si $C = Ae^{kt}$ entonces;

$$C = 9.366307630583030E - 02 e^{-6.257633293863140E - 03 t}$$

$$C = 0.09366307630583030 e^{-0.006257633293863140 t}$$

El error estándar cuadrado es

$$\sigma_{tC} = \sqrt{\frac{S^2}{M-2}} = \sqrt{\frac{1.047430671204260E - 05}{4-2}} = 2.288482762884900E - 03$$

Para el segundo modelo también se realiza cambios de variable

$$C = \frac{1}{kt + A}$$

$$\frac{1}{C} = kt + A$$

Si

$$y = \frac{1}{C}$$

$$k = a_1$$

$$A = a_0$$

$$t = x$$

Se obtiene

$$y = a_0 + a_1 x$$

Las ecuaciones normales son

$$\begin{bmatrix} M & \sum x \\ \sum x & \sum x^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum yx \end{bmatrix}$$

O si prefiere

$$\begin{bmatrix} M & \sum t \\ \sum t & \sum t^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum \frac{1}{C} \\ \sum \frac{t}{C} \end{bmatrix}$$

Calculando las sumatorias que faltan

$$\sum y = \sum \frac{1}{c} = 62.8394283442109$$

$$\sum xy = \sum \frac{1}{c} = 4214.64164210046$$

Las ecuaciones son

$$\begin{bmatrix} 4 & 226 \\ 226 & 19206 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 62.8394283442109 \\ 4214.64164210046 \end{bmatrix}$$

$$4a_0 + 226a_1 = 62.8394283442109$$

$$226a_0 + 19206a_1 = 4214.64164210046$$

Resolviendo

$$a = \begin{bmatrix} 9.879798841790472 \\ 1.031868790822660E - 01 \end{bmatrix}$$

Regresando a la variable original

$$A=9.87979841790472$$

$$k=1.031868790822660E-01$$

La ecuación es

$$C = \frac{1}{1.031868790822660E - 01 t + 9.87979841790472}$$

El error estándar cuadrado es

$$\sigma_{tC} = \sqrt{\frac{S^2}{M-2}} = \sqrt{\frac{0.0000002232941073824}{4-2}} = 6.013709784245E - 04$$

Se puede observar que el modelo 2 es mejor al modelo exponencial ya que su error estándar cuadrado es menor. Por esta razón se acepta que la curva ajustada es:

$$C = \frac{1}{1.031868790822660E - 01 t + 9.87979841790472}$$

Este modelo de acuerdo a su error estándar cuadrado da 3 decimales correctos o sea en este caso 2 cifras significativas. Con esto en mente se determina cual es la

concentración inicial al principio del experimento. Esto equivale a evaluar el modelo en $t=0$ min.

$$C(0 \text{ min}) = \frac{1}{9.87979841790472}$$

$$C(0 \text{ min}) = 0.1012166400266 \text{ mol/l}$$

$$C(0 \text{ min}) = 0.1 \text{ mol/l}$$

¿Cuánto reactivo queda a la media hora de iniciado el experimento?

$$C(30 \text{ min}) = \frac{1}{1.031868790822660E-01 * (30) + 9.87979841790472}$$

$$C(30 \text{ min}) = 0.07706888657084 \text{ mol/l}$$

$$C(30 \text{ min}) = 0.077 \text{ mol/l}$$

¿Cuánto reactivo queda a las 3 horas de iniciado el experimento?

$$C(180 \text{ min}) = \frac{1}{1.031868790822660E-01 * (180) + 9.87979841790472}$$

$$C(180 \text{ min}) = 0.03537163913639 \text{ mol/l}$$

$$C(180 \text{ min}) = 0.035 \text{ mol/l}$$

De los resultados anteriores en orden de confiabilidad tenemos: $C(30 \text{ min})$, $C(0 \text{ min})$, $C(180 \text{ min})$. Esto es porque en 30 minutos se esta interpolando. En los otros 2 valores se realizan extrapolaciones y como ya mencionamos anteriormente es mas seguro interpolar que extrapolar. La concentración en 0 min es mas confiable que la concentración en 180 min porque esta mas cerca del intervalo de tiempos que cubre la tabla.

Por último las unidades de las constantes del modelo son:

$$A [=] \text{mol/l}$$

$$k [=] \text{mol/l/min}$$

Esto es porque el modelo debe de ser dimensionalmente consistente.

Aproximación de una función matemática complicada

Existen funciones matemáticas que son difíciles de evaluar. De estas funciones muchas están tabuladas en manuales de matemáticas. Es común en aplicaciones, sobre todo en programas de computadora usar una ecuación que se aproxime al comportamiento de alguna de estas funciones. Una forma de obtener una función más simple de evaluar a partir de una función complicada, consiste en generar una tabla y posteriormente hallar una curva que se ajuste a la misma.

La función Gamma ó función factorial generalizada, se emplea en la solución de algunas ecuaciones diferenciales ó en el calculo de ciertas integrales. Esta función como su nombre lo indica generaliza el concepto de factorial. Esta definida por

$$\Gamma(x) = (x-1)! = \int_0^{\infty} t^{x-1} e^{-t} dt$$

Como puedes ver es difícil de evaluar.

Puede demostrarse que basta con tabular esta función en el intervalo [0,1] para determinar su valor en cualquier intervalo. La tabla siguiente es un extracto de una tabla de dicha función

Tabla 5 Función Factorial

X	Y
0.0000000000000000E+00	1.0000000000000000E+00
1.0000000000000000E-01	9.513507699000000E-01
2.0000000000000000E-01	9.181687424000000E-01
3.0000000000000000E-01	8.974706963000000E-01
4.0000000000000000E-01	8.872638175000000E-01
5.0000000000000000E-01	8.862269255000000E-01
6.0000000000000000E-01	8.935153493000000E-01
7.0000000000000000E-01	9.086387329000000E-01
8.0000000000000000E-01	9.313837710000000E-01
9.0000000000000000E-01	9.617658319000000E-01
1.0000000000000000E+00	1.0000000000000000E+00

La tabla original tiene 201 puntos. Con el fin de obtener una curva aproximada simple se ajustaron los datos a polinomios. La siguiente tabla muestra el error estándar cuadrado en función del grado

Tabla 6

n	σ_{xy}	nc
1	3.407921011880930E-02	1.251787073568060E-01
2	2.680405127657040E-03	1.351245089721490E-03
3	8.305981847918320E-04	1.115813123544800E-06
4	1.207116027524470E-04	6.672801758014870E-11
5	2.318866456123730E-05	2.801005330955970E-16
6	3.935621112303070E-06	8.087785268244310E-23
7	6.921022337989880E-07	1.583756736752350E-30
8	1.197875423941650E-07	2.081115955774190E-39

9	2.076761431477710E-08	1.819913769105280E-49
10	3.625215464623650E-09	1.049394671170170E-60
11	1.141214816768140E-09	3.621073490217890E-73
12	6.798503362032110E-09	8.429901698868790E-87

Se puede observar que el mejor polinomio es de grado 11, ya que tiene error estándar cuadrado mínimo. No conviene el polinomio de grado 12, ya que el error estándar cuadrado es mayor al de 11.

El polinomio de grado 11 obtenido es

$$\begin{aligned}
 y = & 9.99999992707950E - 01 - 5.77214491151253E - 01x \\
 & + 9.8013546886494E - 01x^2 - 9.068126293057010E - 01x^3 \\
 & + 9.75960153564640E - 01x^4 - 9.50895026807682E - 01x^5 \\
 & + 8.80664428706596E - 01x^6 - 7.08751938874818E - 01x^7 \\
 & + 4.57491864732919E - 01x^8 - 2.13432637183590E - 01x^9 \\
 & + 6.246418661118020E - 02x^{10} - 8.48745000252220E - 03x^{11}
 \end{aligned}$$

De acuerdo al error estándar cuadrado se tienen 8 cifras significativas. Este polinomio es más fácil de evaluar y es más recomendable que la función original, en un programa, ya que el tiempo de máquina para su evaluación es menor.

Pronósticos del número de lectores de Superman

Como se mencionó en la unidad anterior la extrapolación es menos confiable que la interpolación. Esto es principalmente por el fenómeno de oscilación. Este fenómeno es serio en los polinomios de colocación, pero no así en los polinomios de regresión. La diferencia entre ambos es que los de colocación pasan por 2 ó más puntos de la tabla, y los polinomios de regresión se *aproximan* a todos los puntos. Esto trae como consecuencia que su comportamiento sea más suave. Por esto se prefiere los polinomios de regresión para extrapolar.

En la unidad pasada en un problema se realizó la estimación de cuantos lectores potenciales tiene Superman, en base a datos del censo de USA. Los datos se dan nuevamente en la tabla 7.

Tabla 7. Censo de USA

Año	Población
1930	123'203.000
1940	131'669.000
1950	150'697.000
1960	179'323.000
1970	203'212.000

1980	226'505.000
------	-------------

Después de realizar diferentes ajustes el mejor polinomio es

$$y = 7.7677875657478090E + 11 + 1.971086459627230E + 09x \\ - 4.216781089597820E + 06x^2 + 2.355070630196930E + 03x^3 \\ - 4.176980570680360E - 01x^4$$

Con

$$\sigma_{xy} = 2.242942367069320E + 06$$

Por lo cual se tienen 3 cifras significativas.

Si se repite la interpolación y la extrapolación de ese ejemplo se tiene;

$$Y(1.9380000000000000E+03) = 1.288335516086430E+08$$

$$Y(1.9960000000000000E+03) = 2.366530072376710E+08$$

Redondeando a las cifras significativas que nos da el modelo tenemos

$$Y(1938) = 1.29E+08$$

$$Y(1996) = 2.37E+08$$

Los valores obtenidos con polinomios de colocación son respectivamente

$$Y(1938) = 1.291E+08$$

$$Y(1996) = 3E+08$$

Comparando los resultados obtuvimos mas cifras significativas con el polinomio de regresión que con el de colocación

Resumen

El ajuste de curvas ó regresión consiste en dada una tabla determinar una ecuación que se aproxime apropiadamente a los datos.

El método a utilizar se denomina mínimos cuadrados.

Consiste de los siguientes pasos:

Proponer una curva.

Formar la cantidad: $S^2 = \sum e_i^2$

Derivar parcialmente S^2 respecto de cada variable.

Igualar a 0.

Resolver las ecuaciones normales.

Calcular S^2 .

Calcular el error estándar cuadrado.

Las curvas mas usadas son los polinomios.

Para hallar el mejor polinomio se propone el grado y se va calculando el error estándar cuadrado hasta que sea menor o igual a una tolerancia, comience a subir de valor, ó se llegue al grado más alto posible.

Otras 2 curvas muy usadas son la exponencial y la potencial.

Para determinarlas se hace un cambio de variable con logaritmos y se aplica el método de mínimos cuadrados.

La curva propuesta puede determinarse considerando: Teoría, graficando ó por tanteo.

Si es necesario se debe de intentar de simplificar el modelo antes de aplicar el método de mínimos cuadrados. Si no es posible se aplica el método directamente.

El método de mínimos cuadrados NO implica necesariamente ajustar a polinomios.

RECURSOS BIBLIOGRÁFICOS

Bibliografía Básica:

- ✓ MATHEUS. John H. Fink Kurtis D. Métodos Numéricos con MATLAB. Editorial Prentice Hall

Bibliografía Complementaria:

- ✓ ALTZ, Franz L. Electronic. Digital. computers: Their use in science and Engineering. 1958 Academic Press inc. New York.
- ✓ BURDEN Richard L., J. Douglas Faires; Análisis numérico. tr. Efrén Alatorre Miguel; Revisión Técnica. Ildefonso. 1998 (Biblioteca USCO. Nro Topográfico: [515 / B949a.](#))
- ✓ CHAPRA Steven C., CANALE Raymond P, Numerical Methods for engineers. McGraw Hill, Inc. 1988. 839p. ISBN 0-07-909944-0.
- ✓ CHAPRA Steven C., CANALE Raymond P. Métodos numéricos para ingenieros: con aplicaciones en computadoras personales. 1988 (Biblioteca USCO Nro Topográfico: [519.5 / C467m](#))
- ✓ CONDE S. D, Carl de Boor. Análisis numérico elemental: Un enfoque algorítmico. Mc. Graw-Hill 1972, (Biblioteca USCO Nro Topográfico: [511.8 / C761](#) Biblioteca).
- ✓ CORMICK MC., John M. and SALVADOR M.C. Numerical Methods in FORTRAN. 1964. Prentice-Hall Inc Englewood Cliffs N:J.
- ✓ CURTIS, F. Gerald, WHEATLEY, O. Patrick. Análisis numérico con aplicaciones. Tr. Hugo Villagomez Vasquez. 6 Ed. Pearson Educación. 2000, 698p. ISBN 968-444-393-5
- ✓ FADDEEVA, V.N. Computational methods of linear algebra, Dover Publications. 1969, New York.
- ✓ GASTINEL Noël; Análisis numérico lineal. tr. Javier Ruiz Fernández de Pinedo. 1975. (Biblioteca USCO Nro Topográfico: [511.7 / G255](#)).
- ✓ GREENSPAN, D. Theory and solutions of Ordinary Differential Equations. 1960 The. Mc Millan Co. New York.
- ✓ KINCAID David y Ward Cheney; Análisis numérico: Las matemáticas del cálculo científico. tr. Rafael. 1994 (Biblioteca USCO Nro Topográfico: [515 / K51a](#)).
- ✓ LUTHE. Rodolfo, OLIVERA Antonio, SCHUTZ Fernando, Métodos numéricos. 1986 (Biblioteca USCO Nro Topográfico: [511.7 / L973m](#)).
- ✓ McCracken, Daniel D., Métodos numéricos y programación fortran: con aplicaciones en ingeniería y ciencias. 1986. Editorial Limusa. México. (Biblioteca USCO Nro. Topográfico: [001.6424 / M117](#)).
- ✓ NAKAMURA Shoichiro; Métodos numéricos aplicados con software. tr. Oscar Alfredo Palmas Velasco. Prentice Hall Hispanoamericana S.A. 1995. 570p. (Biblioteca USCO. Nro. Topográfico: [511.8 / N163m](#)) ISBN 968-880-263-8
- ✓ NAKAMURA Shoichiro; Análisis numérico y visualización gráfica con MatLab. tr. Roberto Escalona Garcia. Prentice Hall Hispanoamericana S.A. 1997. (Biblioteca USCO N ro Topográfico: [515.1 / N163a](#)). 465p. ISBN 968-880-980-1
- ✓ NIETO RAMIREZ José A., Métodos numéricos en computadoras digitales. Editorial Limusa 1980. (Biblioteca USCO Nro Topográfico: [001.64042 / N677](#)).

- ✓ RALSTON Anthony; Introducción al análisis numérico. tr. Carlos E. Cervantes de Gortari. Editorial Limusa. Mexico. 1978. 629p. (Biblioteca USCO Nro Topográfico: [511.7 / R164.](#))
- ✓ SCARBOROUGH, J.B Numerical mathematics analysis
- ✓ SIERRA ROMERO, Alberto. Manual de Métodos Numéricos. Universidad Tecnológica de Pereira.
- ✓ SMITH, W. Allen; Análisis numérico. tr. Francisco Javier Sánchez Bernabe; Rev. Téc. José Luis Turriza Pinto. Prentice Hall Hispanoamericana S.A. 1988. 608p. (Biblioteca USCO Nro Topográfico: [515 / S664a](#)) ISBN 968-880-119-4.
- ✓ STANTON, Ralph G. Numerical Methods for Science and Engineering. 1967. Prentice-Hall Inc. Englewood Cliffs N.J

Bibliografía OnLine:

http://anamat1.csi.ull.es/anamat_p/Titulaciones/matematicas.htm
<http://arxiv.org/>
<http://books.pdox.net/>
<http://luda.azc.uam.mx/curso2/cp2indic.html>
http://mailweb.udlap.mx/~ccastane/Analisis_Numerico_html/Lindley.html#RegresaGral1
<http://mathworld.wolfram.com/>
<http://omega.ilce.edu.mx:3000/sites/ciencia/html/fisica.htm>
<http://pessoal.sercomtel.com.br/matematica/resumos.htm>
<http://proton.ucting.udg.mx/posgrado/cursos/metodos/temario.html>
<http://sai.uam.mx/apoyodidactico/mn/>
<http://uprhmate01.upr.clu.edu/~pnm/notas4061/index.htm>
<http://webdiee.cem.itesm.mx/web/servicios/archivo/tutoriales/metodos/index.html>
<http://webdiee.cem.itesm.mx/web/servicios/archivo/tutoriales/metodos/algoritmos/index.html>
<http://www.ciencia-hoy.retina.ar/indice.htm>
<http://www.cnice.mecd.es/Descartes/>
<http://www.damp.cam.ac.uk/user/fdl/people/sd/lectures/nummeth98/contents.htm>
<http://www.elprisma.com/>
http://www.fortunecity.com/campus/earlham/850/metodos_numericos/indice.htm#
<http://www.geocities.com/SiliconValley/Pines/7894/metodos/>
http://www.iesrodeira.com/metodos_numericos/index-2.htm
<http://www.ii.uam.es/~pedro/ccii/teoria/>
<http://www.itlp.edu.mx/publica/tutors.htm>
<http://www.monografias.com/trabajos13/tumatlab/tumatlab.shtml>
<http://www.rinconmatematico.com/libros.htm>
<http://www.ucsc.cl/~kdt/numerico/index.htm>
<http://www.unalmed.edu.co/~ifasmar/libro.shtml>
<http://www.uv.es/~diaz/mn/fmn.html>
<http://www-gap.dcs.st-and.ac.uk/~history/index.html> (Biografías)