UNIVERSIDAD DE CONCEPCION
FACULTAD DE CIENCIAS
FISICAS Y MATEMATICAS
DEPARTAMENTO DE INGENIERIA MATEMATICA

Cálculo Numérico (521230) - Laboratorio 4

MINIMOS CUADRADOS

Ajuste Polinomial

Considere el problema de trazar la gráfica de un polinomio

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1} \in \mathcal{P}_{n-1}(\mathbb{R})$$

pasando por los puntos

$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_m, y_m)$$

con $n < m$. Es natural, entonces, exigir que

$$p(x_i) = y_i, \qquad \forall\, i \in \{1, 2, \ldots, m\},$$

o bien, evaluando en el polinomio, tenemos el siguiente sistema de ecuaciones:

$$\left.\begin{array}{c} c_0 + c_1 x_1 + c_2 x_1^2 + \cdots + c_{n-1} x_1^{n-1} = y_1 \\ c_0 + c_1 x_2 + c_2 x_2^2 + \cdots + c_{n-1} x_2^{n-1} = y_2 \\ \vdots \\ c_0 + c_1 x_m + c_2 x_m^2 + \cdots + c_{n-1} x_m^{n-1} = y_m \end{array}\right| \iff \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Dado que $n < m$, generalmente el sistema rectangular $\boldsymbol{Ac} = \boldsymbol{y}$ no tiene solución, lo que es equivalente a decir que no existe ningún polinomio $p \in \mathcal{P}_{n-1}(\mathbb{R})$ tal que $p(x_i) = y_i$.

Sin embargo, podemos encontrar un polinomio $p \in \mathcal{P}_{n-1}(\mathbb{R})$ que pase **lo más cerca posible** de los puntos (x_i, y_i), es decir, encontrar $p(x)$ tal que:

$$\sum_{i=1}^{m} |y_i - p(x_i)|^2 \qquad \text{sea mínima,}$$

(la suma de los cuadrados de los errores sea mínima) lo cual es equivalente a hallar $\boldsymbol{c} \in \mathbb{R}^m$ tal que

$$||\boldsymbol{y} - \boldsymbol{Ac}||_2^2 \qquad \text{sea mínima.}$$

En este caso decimos que el sistema $\boldsymbol{Ac} = \boldsymbol{y}$ se resuelve **en el sentido de los mínimos cuadrados**, y decimos que el polinomio $p(x)$ se ajusta a los puntos (x_i, y_i) en el sentido de los mínimos cuadrados.

En MATLAB el comando `c=A\y` que usted utiliza para resolver un sistema cuadrado también sirve para resolver un sistema rectangular en el sentido de los mínimos cuadrados.

Ejercicio 1. La tabla 1 relaciona la cantidad de cierto aditivo a un barniz con el tiempo de secado del mismo.

1.1 Escriba el sistema de ecuaciones lineales asociado al problema de encontrar el polinomio de grado menor o igual que 2 que mejor ajusta por cuadrados mínimos los datos en la tabla.

1.2 Escriba un rutero MATLAB en el que haga lo siguiente:

Tabla 1: Aditivo y tiempo de secado

Aditivo (en gramos)	**Tiempo de secado (en horas)**
0.0	12.0
1.0	10.5
2.0	10.0
3.0	8.0
4.0	7.0
5.0	8.0
6.0	7.5
7.0	8.5
8.0	9.0

- Construya la matriz $\boldsymbol{A}$ y parte derecha $\boldsymbol{y}$ del sistema escrito por usted en **1.1**.
- Resuelva el sistema $\boldsymbol{Ac} = \boldsymbol{y}$ en el sentido de los mínimos cuadrados.
- Grafique en un mismo gráfico los pares en la tabla y el polinomio obtenido (evaluado en 100 puntos entre 0 y 8 con ayuda de `polyval`).
- Basados en el polinomio resultante, ¿qué cantidad de aditivo resulta en tiempo mínimo de secado? ¿Cuál es el tiempo mínimo de secado?

Observación: Si `x` denota al vector que contiene la cantidad de aditivo en el barniz y `y` contiene los tiempos de secado asociados, según aparecen en la tabla anterior, el llamado `p = polyfit(x,y,2)` retorna en `p` los coeficientes del polinomio de grado menor o igual que 2 que mejor ajusta, en el sentido de los mínimos cuadrados, los pares de datos en `x` y `y`. Este polinomio puede evaluarse usando `polyval`.

Ajuste para otro tipo de funciones

La misma idea del ajuste polinomial se puede aplicar para el ajuste de otro tipo de funciones f. Estas también satisfacen que:

$$\sum_{i=1}^{m} |y_i - f(x_i)|^2 \qquad \text{sea mínima.}$$

Ejercicio 2. En las aguas de un lago hay tres clases de microorganismos provocadores de enfermedades. Se sabe que, en respuesta a un tratamiento aplicado a las aguas, los microorganismos están disminuyendo en forma exponencial de acuerdo al modelo:

$$p(t) = c_1 e^{-1.5t} + c_2 e^{-0.3t} + c_3 e^{-0.05t}, \qquad t \geq 0,$$

donde $p(t)$ da el número (en miles) de microorganismos. De una muestra de las aguas, en un laboratorio se obtuvieron los datos que se muestran en la tabla 2:

Tabla 2: Número de microorganismos de la muestra (en miles)

t (horas)	0.5	1	2	3	4
$p(t)$ (miles)	7	5.2	3.8	3.2	2.5

2.1 Escriba el sistema de ecuaciones lineales asociado al problema de encontrar la función exponencial $p(t)$ que mejor ajusta por cuadrados mínimos los datos en la tabla.

2.2 Escriba un rutero en MATLAB que haga lo siguiente:

- Construya la matriz $\boldsymbol{A}$ y parte derecha $\boldsymbol{y}$ del sistema escrito por usted en **2.1**.
- Resuelva el sistema $\boldsymbol{Ac} = \boldsymbol{y}$ en el sentido de los mínimos cuadrados.
- Grafique en un mismo gráfico los pares en la tabla y la función $p(t)$ obtenida.
- En base a la función obtenida, ¿cuál es el número de microorganismos que había en la muestra inicialmente? ¿y después de una hora y media? ¿y después de 5 horas y media?

Ajuste no lineal

En algunos casos se desean ajustar funciones, las cuales es imposible escribirlas como uns sistema lineal de ecuaciones (tal como lo hemos hecho hasta ahora). Sin embargo, es posible modificar estas funciones para poder escribirlas de forma lineal.

Suponga que se desea ajustar a los datos

$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_m, y_m)$$

la función del ejercicio 3

$$y_a(x) = \alpha\beta^x.$$

Note que si evaluamos esta función en los puntos, el sistema

$$\left.\begin{array}{r} \alpha\beta^{x_1} = y_1 \\ \alpha\beta^{x_2} = y_2 \\ \vdots \\ \alpha\beta^{x_m} = y_m \end{array}\right|$$

no puede ser escrito en forma matricial (ya que no es un sistema lineal). Si aplicamos la función $\ln(\cdot)$ a ambos lados de la igualdad $y_a = \alpha\beta^x$ tenemos:

$$\ln(\alpha\beta^x) = \ln(y_a) \qquad \Longrightarrow \qquad \ln(\alpha) + \ln(\beta)\cdot x = \ln(y_a).$$

Evaluando esta última igualdad en los datos (x_i, y_i) tenemos el sistema

$$\left.\begin{array}{r} \ln(\alpha) + \ln(\beta)\cdot x_1 = \ln(y_1) \\ \ln(\alpha) + \ln(\beta)\cdot x_2 = \ln(y_2) \\ \vdots \\ \ln(\alpha) + \ln(\beta)\cdot x_m = \ln(y_m) \end{array}\right| \quad \Longleftrightarrow \quad \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \begin{pmatrix} \ln(\alpha) \\ \ln(\beta) \end{pmatrix} = \begin{pmatrix} \ln(y_1) \\ \ln(y_2) \\ \vdots \\ \ln(y_m) \end{pmatrix},$$

el cuál ahora puede ser resuelto en el sentido de los mínimos cuadrados. Si el vector $(c_1, c_2)^t$ es solución en el sentido de los mínimos cuadrados del sistema anterior, entonces tenemos que:

$$\left\{\begin{array}{l} c_1 = \ln(\alpha) \\ c_2 = \ln(\beta) \end{array}\right. \qquad \Longrightarrow \qquad \left\{\begin{array}{l} \alpha = e^{c_1} \\ \beta = e^{c_2} \end{array}\right.$$

A pesar de haber modificado la función (este proceso se llama linealización), los coeficientes α y β obtenidos en este proceso hacen de que:

$$\sum_{i=1}^{m} |y_i - y_a(x_i)|^2 \qquad \text{sea mínima.}$$

La forma de linealizar una función dependerá de las funciones que estén involucradas en esta (como por ejemplo en el ejercicio 5).

Tabla 3: Porcentaje de llantas útiles de acuerdo a las millas recorridas.

Miles de Millas recorridas (x)	1	2	5	15	25	30	35	40
Porcentaje útil (y)	99	95	85	55	30	24	20	15

Ejercicio 3. Las cifras de la tabla 3 son datos sobre el porcentaje de llantas radiales producidas por cierto fabricante que aún pueden usarse después de recorrer cierto número de millas.

Se desea ajustar los datos de dicha tabla a los siguientes modelos en el sentido de los mínimos cuadrados:

$$y_a(x) = \alpha\beta^x \qquad \text{e} \qquad y_b(x) = \alpha(100 - x)10^{\beta x}$$

Escriba un rutero en Matlab que ejecute las siguientes tareas:

3.1 Determine los parámetros α y β que ajustan ambos modelos a los datos de la tabla en el sentido de los mínimos cuadrados. Su programa debe mostrar estos parámetros.

3.2 Para ambos modelos, muestre $||\boldsymbol{b} - \boldsymbol{Ax}||_2$ del sistema $\boldsymbol{Ax} = \boldsymbol{b}$ que su programa resuelve.

3.3 Dibuje en un mismo gráfico los datos de la tabla y ambos modelos ajustados.

3.4 Con el mejor modelo, estime qué porcentaje de las llantas radiales del fabricante durarán 45000 millas y 50000 millas. Su programa debe mostrar estas estimaciones.

Ejercicio 4. Para estimar la cantidad de vitamina A requerida para mantener el peso se dió a ratas de laboratorio una dieta básica exenta de vitamina A y se les administró raciones controladas de vitamina A en forma de tabletas. La tabla 4 muestra la relación entre la cantidad de vitamina A administrada y el aumento de peso de las ratas.

Tabla 4: Aumento de peso de las ratas al administrar vitamina A

Dosis de vitamina A (mg)	**Aumento de peso (g)**
0.25	-10.8
1.0	13.5
1.5	16.4
2.5	28.7
7.5	51.3

Escriba un rutero Matlab que,

- Encuentre la función de la forma

$$\textit{Aumento de peso} = a + b\log_{10}(\textit{dosis de vitamina A}),\ a, b \in \mathbb{R}$$

  que mejor ajusta por cuadrados mínimos los datos dados.
- Grafique en un mismo gráfico los pares en la tabla y la función obtenida (evaluada en 100 puntos entre 0.25 y 7.5).
- Basado en la función obtenida, ¿qué cantidad de vitamina A es requerida para no aumentar de peso?

Tabla 5: Concentración de iones a través del tiempo

Tiempo (seg)	$\boldsymbol{n(\times 10^{-4})}$
0	5.03
1	4.71
2	4.40
3	3.97
4	3.88
5	3.62
6	3.30
7	3.15
8	3.08
9	2.92
10	2.70

Ejercicio 5. La tabla 5 muestra la concentración de iones n como una función del tiempo transcurrido después de haber apagado a un agente de ionización.

Se sabe que se cumple la siguiente relación entre la concentración de iones y el tiempo

$$n = \frac{n_0}{1 + n_0\, \alpha\, t}, \tag{1}$$

donde n_0 es la concentración inicial de iones y α, el coeficiente de recombinación.

5.1 Muestre que existe una relación lineal entre n^{-1} y t.

5.2 Encuentre la función (1) que mejor ajusta por cuadrados mínimos a los datos en la tabla. Escriba las aproximaciones a n_0 y α obtenidas.

5.3 Grafique los pares ordenados en la tabla y la función n obtenida (evaluada en 110 puntos entre 0 y 10).