Tema 2 - Operaciones matemáticas básicas

Técnicas de Interpolación

M. en C. Gustavo Contreras Mayén

10 de septiembre de 2012

Contenido

Contenido

Contenido

Contenido

Contenido

Una parte importante en el trabajo del físico es la interpretación de datos experimentales o cálculos teóricos.

Normalmente cuando hacemos mediciones, tenemos un conjunto discreto de puntos que representan nuestro experimento. Para facilitar el trabajo, suponemos que el experimento puede representarse por un par de valores: una *variable independiente* x la cual podemos controlar y una cantidad y, que se mide en el punto x.

Tomemos como ejemplo una fuente radioactiva y un detector, el cual contabiliza el número de decaimientos. Para determinar la vida media de la fuente, debemos de contar los decaimientos N_0, N_1, N_2, ..., N_k, en los tiempos t_0, t_1, t_2, ..., t_k

En este ejemplo, la variable independiente es t, siendo la forma apropiada para resolver el problema. Sin embargo, tenemos un conjunto discreto de pares de números (t_k, N_k) en el rango de (t_0, t_k)

Con la intención de obtener información del experimento, deberíamos de encontrar una función analítica que nos de el valor de N para cualquier punto arbitrario t.

Pero a veces, el tratar de encontrar una función analítica es imposible, o el pensar en utilizar una función conocida, nos podría llevar mucho tiempo para calcularla, más si nuestro interés se basa en una pequeña vecindad de la variable independiente.

Ejemplo

Supongamos que tenemos una fuente radioactiva de ^{241}Am, una fuente de rayos α. Su vida media es $\tau_{\frac{1}{2}} = 430$ años.

Obviamente no podríamos determinar su vida media midiéndola, ya que el decaimiento es lento y quizá lo que podríamos hacer es medir cada lunes durante algunos meses. Después de cinco meses (por ejemplo) podríamos detener las mediciones y revisar los datos.

Una pregunta que nos podemos plantear es: ¿cuál fue la actividad el miércoles de la tercera semana de mediciones? Ya que ese día está dentro del rango de mediciones (t, t)

Lo que podríamos hacer es usar técnicas de interpolación para determinar ese valor. Si lo que queremos, es el caso contrario, conocer la actividad luego de ocho meses posteriores a la última medición, lo que deberíamos de hacer es extrapolar a ese punto a partir de las medicones previas.

Objetivo de la interpolación.

La idea central de la interpolación es seleccionar una función $g(x)$ tal que $g(x_i) = f_i$ para cada dato i, es una buena aproximación para cualquier otro dato x entre el conjunto original de datos.

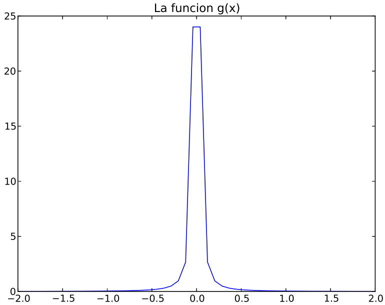
Pero ¿cómo podemos considerar una buena aproximación al conjunto de datos, si no tenemos la función original?

Dado que los puntos ser pueden interpolar por una familia infinita de funciones, para ello debemos de contar con algún criterio o guía para seleccionar una función razonable.

La regla para esos métodos se basa en la suavidad al ajuste de las funciones de interpolación.

Pero esto no podría funcionar para todo tipo de funciones, consideremos la función:

$$g(x) = \frac{1}{25x^2}$$

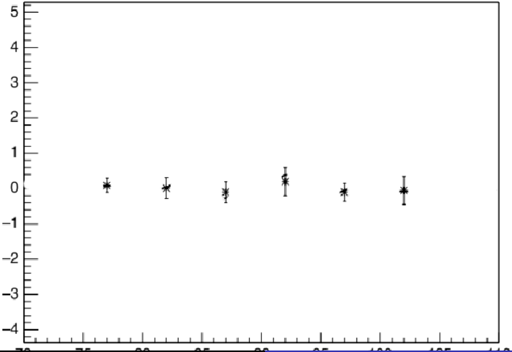


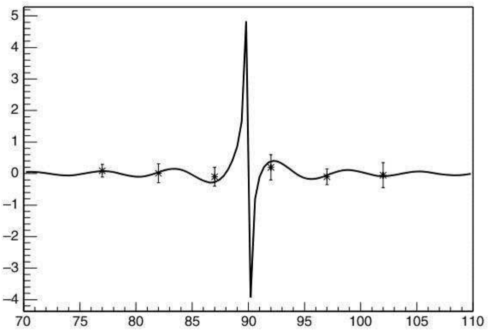
Consideración previa

Antes de entrar de lleno a la revisión de las técnicas de interpolación, es necesario mencionar lo siguiente: dado que contamos con un conjunto finito de puntos, debemos de tener cuidado en el espaciamiento de la variable independiente.

Si los puntos se alejan unos de otros, perderemos información para aquellos valores entre éstos puntos y la predicción de la interpolación ya no será la esperada.

Supongamos que tenemos seis mediciones como se indican en la figura, podemos ver claramente un comportamiento oscilatorio de la función, juzgando por los puntos y de acuerdo a las barras de error, una línea recta es la que probablemente nos ajustaría los puntos.




5
4
3
2
1
0
−1
−2
−3
−4
70
75
80
85
90
95
100
105
110

Interpolación de Lagrange

Este método se basa en el hecho de que en un intervalo finito $[a, b]$ una función $f(x)$ siempre puede representarse por un polinomio $P(x)$.

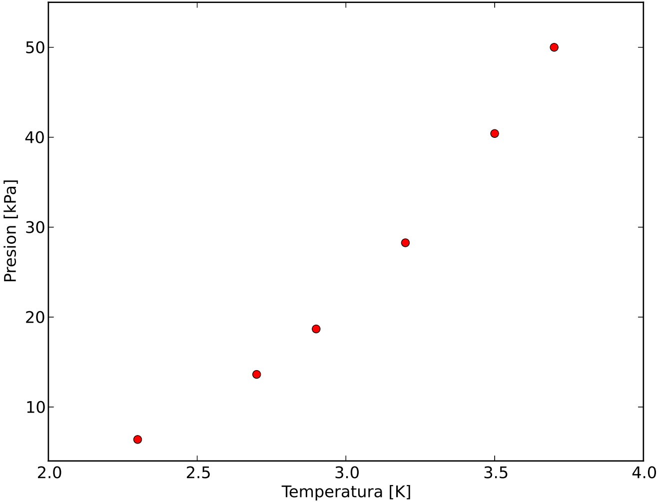
Nuestra tarea es encontrar ese polinomio $P(x)$ a partir del conjunto de datos $(x_i, f(x_i))$

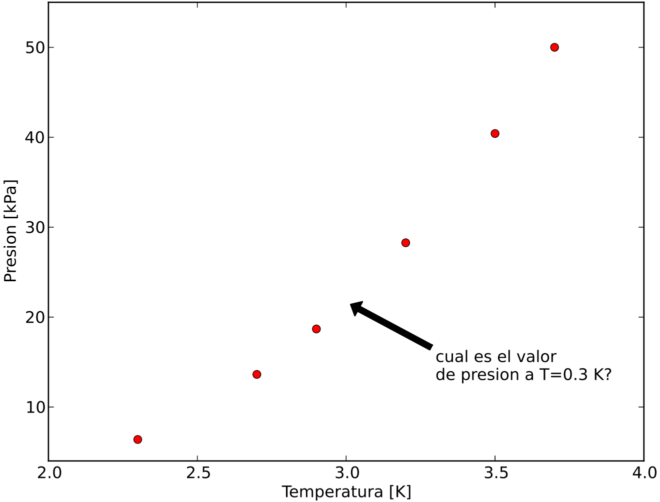
Si tenemos dos pares de datos, la interpolación se realiza hacia adelante, con los datos tabulados para determinar un punto entre los dos valores de la lista.

Ejemplo

Veamos el cambio de la presión del vapor de ^{4}He como función de la temperatura, de acuerdo a la literatura tenemos que:

Temperatura [K]	Presión de vapor [kPa]
2.3	6.38512
2.7	13.6218
2.9	18.6760
3.2	28.2599
3.5	40.4082
3.7	49.9945


Presion [kPa]
50
40
30
20
10
2.0
2.5
3.0
3.5
4.0
Temperatura [K]


Presion [kPa]
Temperatura [K]
cual es el valor
de presion a T=0.3 K?
50
40
30
20
10
2.0
2.5
3.0
3.5
4.0

Procedimiento

El procedimiento hacia adelante, es hacer una interpolación entre dos puntos, por lo que tendremos el conjunto de ecuaciones:

$$
\begin{aligned}
b &= \frac{y_2(x_2) - y_1(x_1)}{x_2 - x_1} \\
a &= y_1(x_1) - x_1 \frac{y_2(x_2) - y_1(x_1)}{x_2 - x_1}
\end{aligned}
$$

Para resolver la ecuación:

$$y(x) = a + bx$$

Combinando las dos primeras ecuaciones y ordenando términos, tenemos que:

$$y(x) = \frac{x - x_2}{x_1 - x_2} y_1(x_1) + \frac{x - x_1}{x_2 - x_1} y_2(x_2)$$

que es la ecuación de una recta que pasa por $(x_1, y_1(x_1))$ y $(x_2, y_2(x_2))$

Datos para el ejemplo

En nuestro ejemplo, tenemos que:

$$(x_1, y_1) = (2.9, 18.6760)$$
$$(x_2, y_2) = (3.2, 28.2599)$$

Datos para el ejemplo

En nuestro ejemplo, tenemos que:

$$
\begin{aligned}
(x_1, y_1) &= (2.9, 18.6760) \\
(x_2, y_2) &= (3.2, 28.2599)
\end{aligned}
$$

Con la interpolación, tenemos que para una temperatura de 3.0 K, la presión tiene un valor de 21.87 kPa.

Interpolación cuadrática

Con la intención de mejorar nuestro resultado, podemos usar un polinomio de segundo orden y utilizar una interpolación cuadrática. En este caso, nuestra función de interpolación, será:

$$
\begin{aligned}
y(x) = & \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} y_1(x_1) + \\
& + \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} y_2(x_2) + \\
& + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} y_3(x_3)
\end{aligned}
$$

Usando los valores de la tabla anterior:

$$\begin{aligned}(x_1, y_1) &= (2.7, 13.6218)\\(x_2, y_2) &= (2.9, 18.6760)\\(x_3, y_3) &= (3.2, 28.2599)\end{aligned}$$

El siguiente paso es usar cuatro puntos y construir el polinomio de orden 3.

Usando los valores de la tabla anterior:

$$
\begin{aligned}
(x_1, y_1) &= (2.7, 13.6218) \\
(x_2, y_2) &= (2.9, 18.6760) \\
(x_3, y_3) &= (3.2, 28.2599)
\end{aligned}
$$

A una temperatura de 3.0 K, la presión de vapor tiene un valor de 21.671 kPa.

El siguiente paso es usar cuatro puntos y construir el polinomio de orden 3.

Fórmula de interpolación de Lagrange

En general podemos escribir la interpolación en términos del polinomio:

$$P(x) = \sum_{k=1}^{N} \lambda_k(x) f(x_k)$$

donde

$$\lambda_k(x) = \frac{\prod_{l=1 \neq k}(x - x_l)}{\prod_{l=1 \neq k}(x_k - x_l)}$$

Esta es la fórmula de interpolación de Lagrange.

Cuando una función conocida $f(x)$ se aproxima mediante un polinomio de interpolación, lo que nos interesa es el error del polinomio.

El error se define como $e(x) = f(x) - g(x)$ donde $f(x)$ es la función de la cual se muestrean los datos $f_i = f(x_i)$

La distribución y magnitud de $e(x)$ se ven afectados por los siguientes parámetros:

- La distribución de las abscisas en los datos.
- El tamaño del dominio de interpolación.
- El orden del polinomio.

Haremos la suposición de que los puntos x_i, están distribuidos uniformemente, de tal forma que la magnitud de $e(x)$, es decir, $|e(x)|$ tiende a ser pequeña en los intervalos cercanos al centro del dominio y tiende a crecer rápidamente haca los extremos.

El tamaño del dominio de interpolación definido como

$$D = x_N - x_0$$

en general, el valor máximo de $|e(x)|$ tiende a cero cuando D disminuye, pero si D crece, el valor de $|e(x)|$ aumenta, inclusive tanto que puede dominar a $|g(x)|$ si N es muy grande.

El error de la fórmula de Lagrange es el siguiente:

$$e(x) = f(x) - g(x) = L(x) f^{N+1}(\xi), \qquad x_0 \leq \xi \leq x_N$$

donde $N+1$ es el número de datos, f^{N+1} es la $n+1$-ésima derivdaa de $f(x)$ y

$$L(x) = \frac{(x - x_0)(x - x_1) \ldots (x - x_N)}{(N+1)!}$$

A resolver el ejercicio

A partir de la siguiente tabla de datos,

x	P(x)
1	0.671
2	0.620
3	0.567
4	0.512

A resolver el ejercicio

A partir de la siguiente tabla de datos,

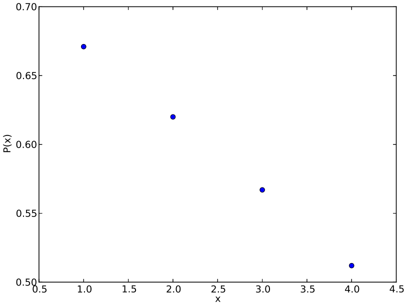
x	P(x)
1	0.671
2	0.620
3	0.567
4	0.512

Con el algoritmo de interpolación de Lagrange usando un polinomio de grado 3, estima el valor de P(x) para los siguientes puntos:
$x = 1.5, 2.5, 3.5,$

```python
import numpy as np
n = 3
x = np.array([1., 2., 3., 4.])
f = np.array([0.671, 0.620, 0.567, 0.512])
xa = eval(raw_input('Dame el valor de x: '))

yres = 0

for i in range(0,n+1):
    z = 1.0
    for j in range(0,n+1):
        if i != j:
            z = z * (xa-x[j])/(x[i]-x[j])
    yres = yres + z*f[i]

print 'El polinomio evaluado en P(',xa,') =', yres
```

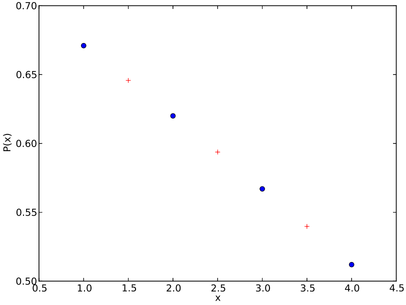
Solución

x	P(x)
1	0.671
1.5	0.64575
2	0.620
2.5	0.59375
3	0.567
3.5	0.53975
4	0.512



Solución

x	P(x)
1	0.671
1.5	0.64575
2	0.620
2.5	0.59375
3	0.567
3.5	0.53975
4	0.512



Ejercicios

1. Ajusta $x \sin(x)$ en $[0, \pi/2]$ con un polinomio de interpolación de Lagrange de orden 4, utilizando puntos con igual separación. Calcula el error de cada interpolación en cada incremento de $\pi/16$, muestra una gráfica.
2. Ajusta $\sin(x)$ en $[0, 2\pi]$ con el polinomio de interpolación de Lagrange de orden 4 y 8, utilizando puntos con igual separación (5 y 9 puntos respectivamente). Grafica los polinomios de interpolación junto con $\sin(x)$ y las distribuciones de sus errores.

Consideraciones importantes

La técnica de interpolación de Lagrange supone que el espaciamiento entre los puntos es la misma, por lo que cuando se presentan puntos que no cumplen ésta condición, la técnica ya no aplicaría.

A continuación se presentan algunas de las desventajas de la primera técnica de interpolación que hemos estudiado.

Desventajas

- La cantidad de cálculos necesarios para una interpolación es grande.
- La interpolación para otro valor de x necesita la misma cantidad de cálculos adicionales, ya que no se pueden utilizar partes de la aplicación previa.
- Cuando el número de datos tiene que incrementarse o decrementarse, no se pueden utilizar los resultados de los cálculos previos.
- La evaluación del error no es fácil.