

PREDICCIÓN DE LA CONCENTRACIÓN EN UNA REACCIÓN QUÍMICA

Lázaro-Camasca¹, Ponce-Pinedo², Sarria-Palacios³, Loayza-Pizarro⁴

Facultad de Ciencias¹, Universidad Nacional de Ingeniería¹

Email: elazaroc@uni.pe¹, victor.ponce.p@uni.pe², esarriap@uni.pe³, fernando.loayza.p@uni.pe⁴

Resumen

En diversos campos de la ingeniería y ciencia se necesita obtener datos a partir de casos experimentales. Las reacciones químicas no son ajenas a esta, por ello conocer las concentraciones de estas ayuda a muchos profesionales. Los químicos y biólogos miden las cantidades de agentes contaminantes para determinar los niveles de **contaminación en el ambiente**. En la industria farmacéutica los laboraristas miden las cantidades de sustancias necesarias para preparar medicamentos; todas estas de concentración determinada y de cuya exacta preparación **depende de la vida y la pronta recuperación de cientos de miles de enfermos**. En las industrias alimentarias los ingenieros miden las cantidades de sustancias, con el propósito de **incrementar sus ingresos económicos**. Este documento proporciona una comparación entre los métodos de ajuste exacto (Interpolación) y mínimos cuadrados (Aproximación Funcional), además sugiere que método o la combinación de ambos predice con mayor exactitud la concentración del producto.

Palabras Clave:

Casos experimentales, Concentración, Contaminación Ambiental, Industrias, Interpolación, Aproximación Funcional.

Abstract

In various fields of engineering and science you need to obtain data from experimental cases. Chemical reactions are not unrelated to this, so knowing the concentrations of these helps many professionals. Chemists and biologists measure the amounts of contaminating agents to determine the levels of contamination in the environment. In the pharmaceutical industry, the laborists measure the quantities of substances necessary to prepare medicines; all of these of determined concentration and whose exact preparation depends on the life and the quick recovery of hundreds of thousands of patients. In the food industries, engineers measure the quantities of substances, with the purpose of increasing their economic income. This document provides a comparison between the methods of exact adjustment (Interpolation) and least squares (Functional Approximation), also suggests that the method or the combination of both predicts with greater accuracy the concentration of the product.

Keywords:

Experimental cases, Concentration, Environmental Pollution, Industries, Interpolation, Functional Approach.

1. INTRODUCCIÓN

En las reacciones químicas la velocidad de la reacción dependen de diversos parametros tales como la temperatura, la presión, la concentracion de los reactivos estos al ser variados aumentan o disminuyen la velocidad.

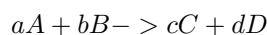
Las industrias se interesan grandemente en que las reacciones químicas se lleven a cabo rápidamente, para **ahorrar tiempo y dinero**. Es por esto que encontrar las concentraciones en un tiempo especifico es muy complicado, una posible solución seria contar con un dispositivo que registre en tiempo real las concentraciones, pero esto seria muy costoso. Una alternativa para esto es el uso de la matematica en concreto el uso de metodos numericos.

Por ello en presente trabajo se describe la obtención de una función que caracteriza la concentración de producto de una reacción química en función del tiempo. Esta función será calculada mediante metodos de interpolación polinomial y minimos cuadrados a partir puntos tabulados que han sido obtenidos de manera experimental.

2. CONCEPTOS PREVIOS

2.1. Reacciones Química

Es un proceso en el que dos o más sustancias (reactantes) se transforman en una o más sustancias (productos). Estas reacciones deben satisfacer la ley de la conservación de la materia.



2.1. Velocidad Instantánea de Reacción

Expresa el cambio de la concentración de un reactivo o de un producto por unidad de tiempo (concentración/tiempo).

$$v = \frac{-d[A]}{a * d[t]} = \frac{-d[B]}{b * d[t]} = \frac{-d[C]}{c * d[t]} = \frac{-d[D]}{d * d[t]}$$

2.2. Interpolación polinomial

Planteamos tres problemas relacionados con la aproximación de funciones

Primero, suponga que tenemos una tabla de valores numéricos de una función:

x	x_0	x_1	\dots	x_n
y	y_0	y_1	\dots	y_n

¿Es posible encontrar una fórmula simple y conveniente que reproduzca los puntos dados exactamente?

El segundo problema es similar, pero se supone que la tabla de valores numéricos dada está contaminada por errores, como puede ocurrir si los valores provienen de un experimento de física. Ahora nos preguntamos por una fórmula que represente los datos (aproximadamente) y, si es posible, filtre los errores.

Como un tercer problema, una función f está dada, quizás en la forma de un procedimiento computacional, pero es costoso evaluarla. En este caso, nos preguntamos por otra función g que sea simple de evaluar y produzca una aproximación razonable a f . A veces en este problema, queremos que g se aproxime a f con precisión total de máquina.

En todos estos problemas, se puede obtener una función simple p que represente o aproxime a la tabla o función f dadas. La representación p siempre se puede tomar como un polinomio, aunque también se pueden usar muchos otros tipos de funciones simples. Una vez que se ha obtenido una función simple p , se puede usar en lugar de f en muchas situaciones. Por ejemplo, la integral de f se podría estimar con la integral de p y, generalmente, esta última debe ser más fácil de evaluar

2.2.1 Polinomios de Lagrange

Suponga que queremos interpolar funciones arbitrarias en un conjunto de nodos fijos x_0, x_1, \dots, x_n . Primero definimos un sistema de $n+1$ polinomios especiales de grado n conocidos como **polinomios cardinales** en la teoría de interpolación. Éstos se denotan por L_0, L_1, \dots, L_n y tienen la propiedad:

$$L_i(x_j) = \gamma_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Una vez que están disponibles, podemos interpolar cualquier función f por la **forma de Lagrange de interpolación polinomial**:

$$p_n(x) = \sum_{i=0}^n L_i(x) f(x_i)$$

Esta función p_n , al ser una combinación lineal de los polinomios L_i , es en sí misma un polinomio de grado a lo más n . Además, cuando evaluamos p_n en x_i obtenemos $f(x_j)$:

$$p_n(x_j) = \sum_{i=0}^n L_i(x_j) f(x_i) = L_j(x_j) f(x_j) = f(x_j)$$

Por tanto, p_n es el polinomio de interpolación para la función f en los nodos x_0, x_1, \dots, x_n . Ahora sólo resta escribir la fórmula para el polinomio cardinal i , que es

$$L_i(x) = \prod_{j \neq i, j=0}^n \left(\frac{x - x_j}{x_i - x_j} \right)$$

Esta fórmula indica que $L_i(x)$ es el producto de n factores lineales:

$$L_i(x) = \left(\frac{x - x_0}{x_i - x_0} \right) \left(\frac{x - x_1}{x_i - x_1} \right) \dots \left(\frac{x - x_{i-1}}{x_i - x_{i-1}} \right) \left(\frac{x - x_{i+1}}{x_i - x_{i+1}} \right) \dots \left(\frac{x - x_n}{x_i - x_n} \right)$$

Los denominadores son sólo números; la variable x se presenta únicamente en los numeradores. Por tanto, i es un polinomio de grado n . Observe que cuando $L_i(x)$ se evalúa en $x = x_i$, cada factor en la ecuación anterior será 1. Por tanto, $L_i(x_i) = 1$. Pero cuando

se evalúa $L_i(x)$ en cualquier otro nodo, digamos, x_j , uno de los factores en la ecuación anterior será 0 y $L_i(x_j) = 0$ para $i \neq j$.

2.2.2 Método de Newton

Supóngase que se tiene una función dada en forma tabular como se presenta a continuación

Puntos	0	1	...	n
x	x_0	x_1	...	x_n
$f(x)$	$f[x_0]$	$f[x_1]$...	$f[x_n]$

y que se desea aproximarla preliminarmente con un polinomio de primer grado que pasa, por ejemplo, por los puntos (0) y (1). Sea además dicho polinomio de la forma:

$$p(x) = a_0 + a_1(x - x_0) \dots (*)$$

donde x_0 es la abscisa del punto (0) y a_0, a_1 son constantes por determinar. Para encontrar el valor de a_0 se hace $x = x_0$, de donde $a_0 = p(x_0) = f[x_0]$, y fin de encontrar el valor de a_1 se hace $x = x_1$, de donde $a_1 = (f[x_1] - f[x_0]) / (x_1 - x_0)$, o sea la primera diferencia dividida $f[x_1, x_0]$.

Al sustituir los valores de estas constantes en la ecuación (*) ésta queda

$$p(x) = f[x_0] + (x - x_0)f[x_0, x_1]$$

o sea un polinomio de primer grado en términos de diferencias divididas. Y si ahora se desea aproximar la función tabular con un polinomio de segundo grado que pase por los puntos (0), (1) y (2) y que tenga la forma:

$$p_2(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1)$$

generalizando se puede formar para p_n .

Una observación muy importante acerca de p_n es que los coeficientes a_0, a_1, \dots , no dependen de n . En otras palabras, p_n se obtiene de p_{n-1} sumando un término

más, sin alterar los coeficientes que ya están en p_{n-1} . Este es porque comenzamos esperando que pn pudiera expresarse en la forma:

$$p_n(x) = p_{n-1}(x) + a_n(x - x_0) \dots (x - x_{n-1})$$

y descubrimos que esto de hecho era posible.

Una forma de determinar sistemáticamente los coeficientes desconocidos a_0, a_1, \dots, a_n es hacer x igual a x_0, x_1, \dots, x_n en la forma de Newton y a continuación escribimos la forma compacta de las ecuación:

$$f(x_k) = \sum_{i=0}^k a_i \prod_{j=0}^{i-1} (x_k - x_j)$$

donde $(0 \leq k \leq n)$.

Las ecuaciones se pueden resolver para las a_i correspondientes, iniciando con a_0 . Entonces vemos que a_0 depende de $f(x_0)$, que a_1 depende de $f(x_0)$ y $f(x_1)$, y así sucesivamente. En general, a_k depende de $f(x_0), f(x_1), \dots, f(x_n)$. En otras palabras, a_k depende de los valores de f en los nodos x_0, x_1, \dots, x_k . La notación tradicional es:

$$a_k = f[x_0, x_1, \dots, x_k]$$

Esta ecuación define $f[x_0, x_1, \dots, x_k]$. La cantidad $f[x_0, x_1, \dots, x_k]$ se **llama diferencia dividida de orden k** para f .

2.3. Mínimos Cuadrados

Hasta ahora se ha enfocado en encontrar un polinomio de aproximación que pase **por los puntos** dados en forma tabular. Sin embargo, a veces la información (dada en la tabla) tiene errores significativos. En estas circunstancias no tiene sentido pasar un polinomio de aproximación por los puntos dados, sino sólo cerca de ellos.

No obstante, esto crea un problema, ya que se puede pasar un número infinito de curvas **entre los**

puntos. Para determinar **la mejor** curva se establece un criterio que la fije y una metodología que la determine. El criterio más común consiste en pedir la suma de distancia calculadas entre el valor de la función que aproxima $p(x_i)$ y el valor de la función $f(x_i)$ dada en la tabla, sea mínima, además para evitar problemas de derivabilidad más adelante, se acostumbra utilizar las distancias d_i elevadas al cuadrado; es decir, que

$$\sum_{i=1}^m [p(x_i) - f(x_i)]^2 = \sum_{i=1}^m [d_i]^2 = \text{mínimo}$$

si queremos aproximar una función en forma tabular con un polinomio de grado n , el procedimiento es minimizar la función:

$$\sum_{i=1}^m [a_0 + a_1 x_i + \dots + a_n x_i^n - f(x_i)]^2$$

el cual se obtiene derivándola parcialmente con respecto a cada coeficiente a_j , $(0 < j < n)$ e igualando a cero cada una de estas derivadas, con esto se llega la siguiente sistema lineal:

$$\begin{aligned} a_0 m + a_1 \sum x + \dots + a_n \sum x^n &= \sum y \\ a_0 \sum x + a_1 \sum x^2 + \dots + a_n \sum x^{n+1} &= \sum xy \\ a_0 \sum x^2 + a_1 \sum x^3 + \dots + a_n \sum x^{n+2} &= \sum x^2 y \\ &\vdots \\ a_0 \sum x^n + a_1 \sum x^{n+1} + \dots + a_n \sum x^{n+n} &= \sum x^n y \end{aligned}$$

3. ANÁLISIS

Para nuestro problema se tiene una tabla T que contiene la variación de la concentración del producto C_B con respecto al tiempo.

t	0.00	0.10	0.40	0.60	0.80	1.00
C_B	0.00	0.30	0.55	0.80	0.10	0.15

El objetivo es **calcule la concentración** C_B cuando $t = 0,82$ con los diferentes metodos mencionados y cual de ellos es **el mas preciso** para dicho problema.

Los metodos Lagrange, Newton y Mínimos Cuadrados fueron implemtados en el lenguaje pyhton, por ser este multiparadigma y robusto. Se utilizo en especial la libreria Numpy para el procesamiento matematico, aunque tambien existen otras como SysPy.

3.1. Análisis de Eficiencia

La eficiencia de un método se mide por medio del error, este error nos dira cuan confiable es el método en comparación de otro.

Para hallar el error se implemento un método creativo llamado **reemplazo de intervalo de tiempo mínimo** consiste en calcular la concentración C_B en un tiempo t_a usando un método X y una tabla de datos T , ahora reemplazamos t_a y su concentración C_B con un t_b tal que $|t_a - t_b| = \min$ donde \min es la minima diferencia entre t_a y todos los valores t de la tabla T .

Una vez reemplazado t_a en T se tiene la nueva tabla T' , ahora calculamos con el método X el valor de C_B en el tiempo t_b . Ahora se tiene dos concentraciones $C_{B.Exp}$ de la tabla T hallado experimentalmente y $C_{B.Aprox}$ de la tabla T' hallado aproximadamente con el método X .

El error se calculara de la siguiente manera:

$$error = \left| \frac{C_{B.Aprox} - C_{B.Exp}}{C_{B.Exp}} \right| 100\%$$

Ejemplo

A manera de ejemplo usemos el método de Newton y calculemos la C_B en el tiempo $t_a = 0,38$.

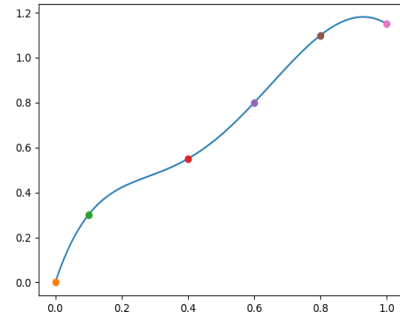


Fig. Polinomio de Newton a partir de la tabla T

Calculando se obtiene que en t_a la $C_B = 0,5338970444444444$. Ahora reemplazamos por el tiempo $t_b = 0,40$ tiempo minimo y $C_{B.Exp} = 0,55$.

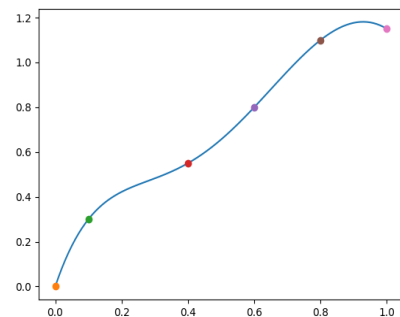


Fig. Polinomio de Newton a partir de la tabla T'

Las graficas de la tabla T y T' parecen ser las mismas sin embargo al calcular la concentracion encontramos que $C_{B.Aprox} = 0,5500000000000002$

Hallando el $error = 2,0185873175002846e - 14$, entonces nuestra interpolación para el tiempo $t_a = 0,38$ es aceptable.

¿Qué ocurre si no reemplazamos en un intervalo de tiempo minimo?

Remplazando $t_a = 0,38$ con $C_B = 0,5338970444444444$ en $t_b = 0,8$.

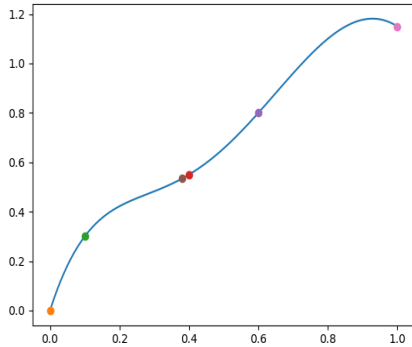


Fig. Polinomio de Newton a partir de la tabla T' sin tiempo minimo.

En la grafica se puede notar que hay dos puntos muy cercanos como si fueran uno, además que sigue describiendo una curva muy similar a las anteriores. Al calcular la concentracion encontramos que $C_{B.Aprox} = 1,099999999999998$ en comparacion del $C_{B.exp} = 1,1$.

Ahora vamos a resolver el problema calcular la concentraci3n en $t = 0,82$.

M3todo de Lagrange

Calculando se obtiene que en $t_a = 0,82$ la $C_B = 1,1217764000000001$. Remplazamos por $t_b = 0,8$ y $C_{B.Exp} = 1,10$ para obtener la tabla T' .

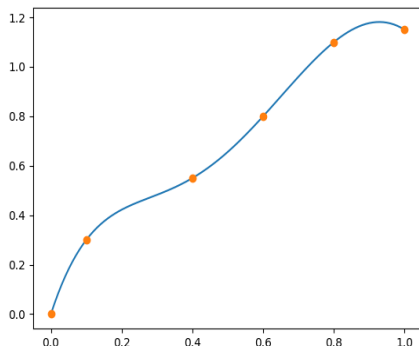


Fig. Polinomio de Lagrange a partir de la tabla T .

Calculando la concentraci3n con la tabla T' se obtiene $C_{B.Aprox} = 1,1$, entonces el $error = 0,0\%$.

M3todo de Newton

Calculando se obtiene que en $t_a = 0,82$ la $C_B = 1,1217764000000006$. Remplazamos por $t_b = 0,8$ y $C_{B.Exp} = 1,10$ para obtener la tabla T' .

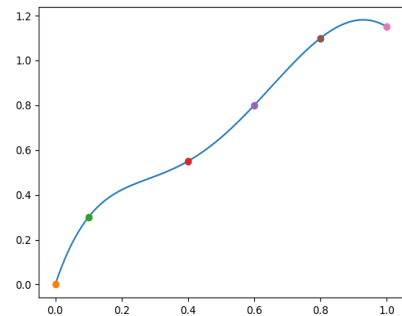


Fig. Polinomio de Newton a partir de la tabla T'

Calculando la concentraci3n con la tabla T' se obtiene $C_{B.Aprox} = 1,1000000000000001$, entonces el $error = 0,0\%$.

Aproximaci3n por M3nimos Cuadrados

En este caso se puede aproximar a polinomios de diferentes grados.

Polinomio de grado 2

Calculando se obtiene que en $t_a = 0,82$ la $C_B = 1,03986632$.

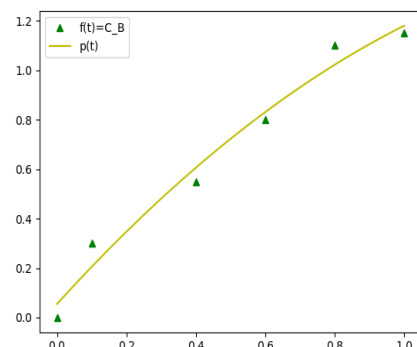


Fig. Polinomio minimos cuadrados de grado 2 a partir de la tabla T .

Remplazamos por $t_b = 0,8$ y $C_{B.Exp} = 1,10$ para obtener la tabla T' .

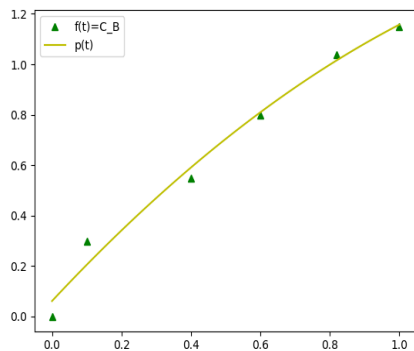


Fig. Polinomio minimo cuadrado de grado 2 a partir de la tabla T' .

Calculando la concentración con la tabla T' se obtiene $C_{B.Aprox} = 0,99858752$, entonces el $error = 9,2193163636375 \%$.

Polinomio de grado 5

Como en los anteriores metodos se genera un polinomio de grado 5, en este tambien tambien sera del mismo grado, para poder realizar la comparativa entre metodos con las mismas condiciones.

Calculando se obtiene que en $t_a = 0,82$ la $C_B = 1,12177642$.

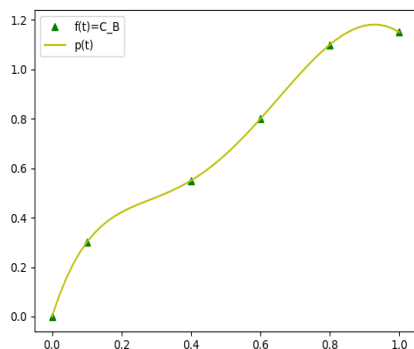


Fig. Polinomio minimo cuadrado de grado 5 a partir de la tabla T .

Remplazamos por $t_b = 0,8$ y $C_{B.Exp} = 1,10$ para obtener la tabla T' .

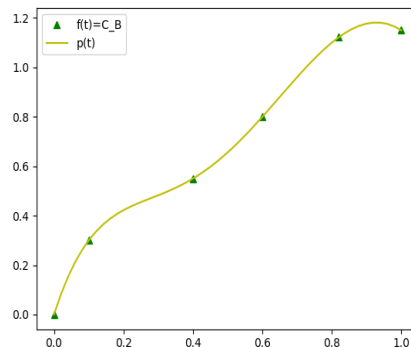


Fig. Polinomio minimos cuadrados de grado 5 a partir de la tabla T' .

Calculando la concentración con la tabla T' se obtiene $C_{B.Aprox} = 1,10000002$, entonces el $error = 1,8181818071318707e - 06 \%$.

Notar que en la aproximacion por minimos cuadrados, la grafica del polinomio no pasa por los puntos de la tabla esto se puede visualizar en la siguiente imagen:

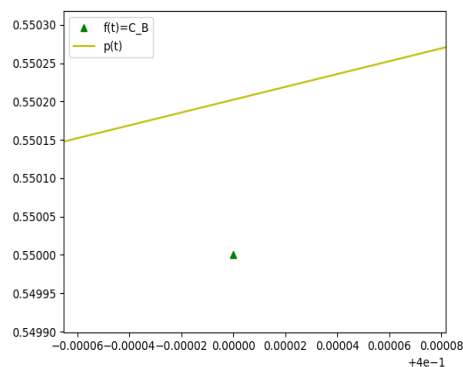


Fig. Polinomio minimo cuadratico de grado 5 no pasa por el punto $t_a = 4,0$.

En el método de Newton y Lagrange ocurre algo similar, pero en este caso el punto esta mucho mas cerca de la grafica del polinomio.

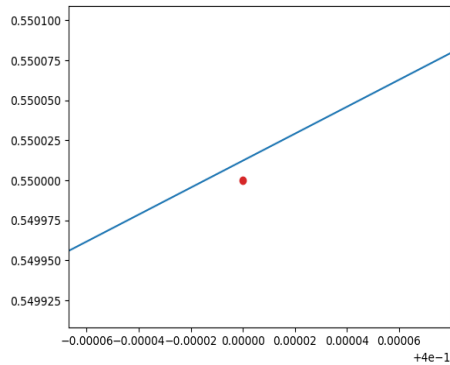


Fig. Polinomio Newton grado 5 no pasa por el punto $t_a = 4,0$ pero esta muy cerca.

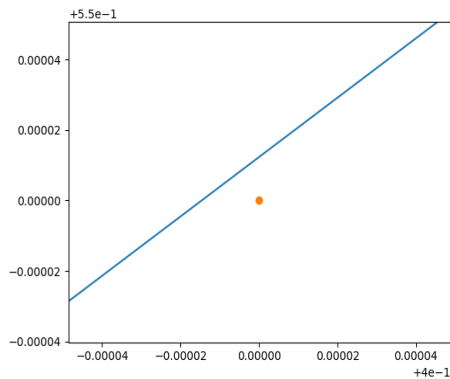


Fig. Polinomio Lagrange grado 5 no pasa por el punto $t_a = 4,0$ pero esta muy cerca.

En resumen se obtuvo los siguientes tabla, donde se muestra la concentracion en el tiempo $t = 0,82$

Método	C_B	Error
Lagrange	1.12177640+1e-16	0.0 %
Newton	1.12177640+6e-16	0.0 %
Minimos Cuadrados	1.12177642	1.82 e-06 %

4. OBSERVACIONES

Dentro del analisis se pudo observar fenomenos curiosos que pueden beneficiar o perjudicar significa-

tivamente a la predicción de la concentración, en lo siguiente se mencionan algunas.

Puntos Caracteristicos

Los puntos caracteristicos son aquellos puntos que al no ser considerados en el proceso de predicción generan un error considerable, hasta se puede perder valores en ciertos intervalos ejemplo si se no se considera los extremos en este caso $t = 0,00$ y $t = 1,00$ no se podra predecir en el intervalo $[t = 0,00, t = 0,10]$ y $[t = 0,80, t = 1,00]$.

Sin embargo no consideramos puntos caracteristicos aquellos puntos que al ser eliminados estos no perjudican a la aproximación, es decir la grafica que se describe sigue siendo casi la misma, esto se puede notar en la parte de Análisis en la sección del ejemplo, al eliminar el punto 0.8 la grafica y la aproximación no se altera considerablemente.

Los puntos caracteristicos se pueden obtener de replicar muchas veces el experimento, uno de los propositos es que al tener estos puntos puedes eliminar aquellos que no lo son y asi reducir la cantidad de data y acelerar el proceso de predicción con otros puntos fuera de la data.

Intervalos Proximos

Si se tiene 2 intervalos $A = [a', b']$, $B = [a, b]$ donde $|b' - a'| < |b - a|$ si se realiza la predicción de un $c' \in A$ y un $c \in B$, entonces la predicción de c' es mucho mejor que c .

Cantidad de puntos

Dentro del analisis la cantidad de data o puntos es un parametro muy relevante ya que cuando hay mayor cantidad de puntos la predicción sera mejor. Sin embargo cuando se tiene muchos puntos se realiza más operaciones y por se requiere de mayor capacidad de computo.

Comparación de metodos

Los resultados de los metodos son iguales hasta la

septima cifra decimal, pero en los metodos de interpolación se obtiene un adicional de $1e-16$ para Lagrange y $6e-16$ para Newton el cual es insignificante para el problema.

5. CONCLUSIONES

Cuando la información tabular es aproximada hasta cierto número de cifras significativa, por ejemplo las tablas de logaritmos, se recomienda el método de interpolación. En cambio cuando los datos son obtenidos experimentalmente, no tiene sentido encontrar un polinomio que pase por los puntos sino más bien que pase entre ellos; entonces, el método de mínimos cuadrados es el aplicable.

Mejor método

Se concluye que para el problema el mejor método es mínimos cuadrados porque los datos son hallados experimentalmente sin embargo si se cuenta con poca capacidad de computo y se realizan demasiados calculos se recomienda un método de interpolación.

Una solución para la no realizar demasiadas operación es utilizar la programación dinamica, este almacena las sumatorias en variables para su posterior uso, la programación dinamica aumenta la complejidad de almacenamiento y disminuye la complejidad

computacional.

Además la interpolacion y Minimos cuadrados son la base para muchos investigaciones en ingerieria y ciencias como la fisica, quimica, hasta en aplicaciones mas recientes en el Machine Learning, como las Redes neuronales donde se usa el método de mínimos cuadrados para encontrar el valor de la conexión entre neuronas al momento del entrenamiento.

Interfaz grafica

Para resolver el problema se implemento una **interfaz gráfica** donde el usuario puede elegir el método numérico para resolver el problema, además podrá hallar la C_B en el cualquier tiempo t .

Con esto es usuario podrá comparar la eficiencia de los diferentes metodos numéricos ver cual es el mejor y el peor, también podrá insertar sus datos para un mayor aprendizaje, ya que la **eficiencia de los metodos** depende de cada problema y por ende de los datos. Se puede descargar los script de los metodos y la interfaz grafica en [7].

Agradecimientos

Los autores agradecen a las autoridades de la Facultad de Ciencias de la Universidad Nacional de Ingeniería por su apoyo.

-
1. Jaan Kiusalaas, *Numerical Methods in Engineering with python*. Cambrige University Press **37** (2010).
 2. L.Héctor Juaréz V., *Análisis Numérico* Universidad Autónoma Metropolitana **2008** (2010).
 3. W. Kincaid, D. Cheney, *metodos Numéricos y Computación*, sexta edición. Cengage Learning, 300-305.
 4. Walter Mora F., *Introducción a los metodos Numéricos*, primera edición. Instituto Tecnológico de Costa Rica, 2013.
 5. Dr. Herón Morales Marchena, *Interpolación, Diferenciación e Integración Numérica*, <http://biblioteca.uns.edu.pe/saladocentes/archivoz/cuoroz/cuaderno.2.pdf>.
 6. Pontificia Universidad Católica del Perú, 2011 *Química General*, <http://corinto.pucp.edu.pe/quimicageneral/unidades-q2/unidad-2-cinetica-quimica.html>.
 7. Repositorio del proyecto en github <https://github.com/nicklc007/Proyecto-2-Numerico>