

MATEMATICAS VI (Métodos numéricos)





Objetivo General

EL ALUMNO EVALUARA LOS ALCANCES Y LIMITACIONES DE DIVERSOS ALGORITMOS EN LA RESOLUCION DE PROBLEMAS, HACIENDO ENFASIS EN SU IMPLANTACION A TRAVES DE SOFTWARE.



contenido

| | |
|---|-----------|
| 1. ARITMETICA DE PUNTO FLOTANTE | 5 |
| 1.1. OBJETIVO PARTICULAR | 5 |
| 1.2. SISTEMAS NUMERICOS DE PUNTO FLOTANTE | 5 |
| 1.3. SISTEMAS NUMERICOS DE MAQUINA EN PUNTO FLOTANTE | 5 |
| 1.4. PROPIEDADES CUALITATIVAS DEL SISTEMA DE NUMEROS DE MAQUINA EN PUNTO FLOTANTE | 6 |
| 1.4.1. MULTIPLICACIÓN | 7 |
| 1.4.2. DIVISIÓN | 8 |
| 1.4.3. SUMA Y RESTA | 9 |
| 1.5. ERRORES EN ARITMETICA DE PUNTO FLOTANTE | 11 |
| 2. SISTEMAS DE ECUACIONES ALGEBRAICAS LINEALES | 12 |
| 2.1. OBJETIVO PARTICULAR | 12 |
| 2.2. INTRODUCCION | 12 |
| 2.3. ELIMINACION DE GAUSS Y GAUSS-JORDAN PARA PROBLEMAS IDEALES SENCILLOS | 14 |
| 2.3.1. Método de eliminación de Gauss | 14 |
| 2.4. PIVOTEO Y ELIMINACION CONICA DE GAUSS | 19 |
| 2.5. PROBLEMAS SIN SOLUCION UNICA | 21 |
| 2.6. MATRICES Y VECTORES | 21 |
| CLASIFICACIÓN DE LOS MÉTODOS DIRECTOS | 24 |
| 2.7. INVERSION DE UNA MATRIZ | 25 |
| 2.8. DESCOMPOSICION LU | 27 |
| 2.9. DETERMINANTES | 28 |
| 2.10. PROBLEMAS MAL CONDICIONADOS | 33 |
| 2.11. SOLUCION DE N ECUACIONES CON M INCÓGNITAS | 34 |
| 3. INTERPOLACION | 36 |
| 3.1. OBJETIVO PARTICULAR | 36 |
| 3.2. INTRODUCCION | 36 |
| 3.3. INTERPOLACION LINEAL | 36 |
| 3.4. FORMULA DE INTERPOLACION DE LAGRANGE | 37 |
| 3.5. INTERPOLACIONES DE NEWTON HACIA DELANTE Y HACIA ATRÁS EN PUNTOS CON IGUAL SEPARACION | 40 |
| 3.5.1. INTERPOLACION DE NEWTON HACIA DELANTE | 40 |
| 3.5.2. INTERPOLACION DE NEUTON HACIA ATRÁS | 41 |
| 3.6. INTERPOLACION CON RAICES DE CHEBYSTEV | 42 |
| 3.7. POLINOMIOS DE INTERPOLACION DE HERMITE | 42 |
| 3.8. INTERPOLACION EN DOS DIMENCIONES | 42 |
| 3.9. EXTRAPOLACIONES | 44 |
| 4. ECUACIONES NO LINEALES | 45 |
| 4.1. OBJETIVO PARTICULAR | 45 |
| 4.2. INTRODUCCION | 45 |
| 4.3. METODO DE BISECCION | 45 |
| 4.4. METODO DE FALSA POSICION Y EL METODO DE LA FALSA POSICION MODIFICADO | 45 |
| 4.5. METODO DE NEWTON | 46 |
| 4.6. METODO DE LA SECANTE | 48 |
| 4.7. METODO DE SUCESSION SUCESSION | 48 |
| 4.8. METODO DE BAIRSTOW | 51 |
| 5. METODO DE MINIMOS CUADRADOS | 53 |
| 5.1. OBJETIVO PARTICULAR | 53 |
| 5.2. INTRODUCCIÓN | 53 |
| 5.3. REGRESION LINEAL | 53 |
| 5.4. AJUSTE DE CURVAS CON UN POLINOMIO DE ORDEN SUPERIOR | 60 |
| 5.5. AJUSTE DE CURVAS MEDIANTE UNA COMBINACION LINEAL DE FUNCIONES CONOCIDAS | 64 |
| 6. OPTIMIZACION DE FUNCIONES EN UNA DIMENSION | 66 |
| 6.1. OBJETIVO PARTICULAR | 66 |



| | |
|--|-----------|
| 6.2. INTRODUCCION | 66 |
| 6.3. MÉTODO SIMPLEX..... | 66 |
| 6.4. OPTIMIZACIÓN SIN RESTRICCIONES EN DIMENSIÓN..... | 68 |
| 6.5. MÉTODOS NUMÉRICOS PARA DIMENSIÓN..... | 69 |
| 6.5.1. <i>Método de búsqueda directa</i> | 69 |
| 6.5.2. <i>Búsqueda dicotómica (para máximos)</i> | 69 |
| 6.6. MÉTODO DE NEWTON | 69 |
| 6.7. MÉTODO DEL GRADIENTE..... | 71 |
| 6.8. MÉTODO DEL ASCENSO MÁS RÁPIDO | 72 |
| 7. BIBLIOGRAFÍA..... | 74 |



1. Aritmetica de punto flotante

1.1. Objetivo particular

El alumno identificara las dificultades al realizar diferentes operaciones con punto flotante.

1.2. Sistemas numericos de punto flotante¹

El sistema de números de punto flotante es la representación más común hoy en las computadoras.

Hay varias maneras de representar los números reales en computadoras. El sistema de coma fija, coloca una coma en algún lugar entre los dígitos, por convención y es equivalente a usar enteros que representan las porciones de alguna unidad. Por ejemplo, uno quizás represente un centésimo de una unidad; si usted tiene cuatro dígitos de decimal. Otro enfoque sería usar números racionales, y representar cada número como la razón de dos enteros.

La representación del punto flotante representa básicamente un número real en notación científica. La notación científica representa los números como un número base y un exponente.

1.3. Sistemas numericos de maquina en punto flotante

Estamos interesados en estimar el error en que se incurre al aproximar un número real positivo x mediante un número de máquina del MARC-32. Si representamos el número mediante:

$$x = (a_1 a_2 \cdots a_{24} a_{25} a_{26} \cdots)_2 \times 2^m$$

en donde cada a_i es 0 ó 1 y el bit principal es $a_1 = 1$. Un número de máquina se puede obtener de dos formas:

- **Truncamiento:** descartando todos los bits excedentes $a_{25} a_{26} \cdots$. El número resultante, x' es siempre menor que x (se encuentra a la izquierda de x en la recta real).
- **Redondeo por exceso:** Aumentamos en una unidad el último bit remanente a_{24} y después eliminamos el exceso de bits como en el caso anterior.

Todo lo anterior, aplicado al caso del MARC-32, se resume diciendo que si x es un número real distinto de 0 dentro del intervalo de la máquina, entonces el número de máquina x^* más cercano a x satisface la desigualdad:

$$\delta = \left| \frac{x - x^*}{x} \right| \leq 2^{-24} \quad (20)$$

que se puede escribir de la siguiente forma:

$$x^* = x(1 + \delta) \quad |\delta| \leq 2^{-24}$$

¹ Manual del PLC DL06, 1a. edición en español, 10/04 **J-5**



Ejemplo 6: ¿Cómo se expresa en binario el número $x = 2/3$? ¿Cuáles son los números de máquina x' y x'' próximos en el MARC-32?

El número $2/3$ en binario se expresa como:

$$\left(\frac{2}{3}\right)_{10} = (0.\overline{10})_2$$

Los dos números de máquina próximos, cada uno con 24 bits, son:

$$x' = (0.101010 \dots 1010)_2$$

$$x'' = (0.101010 \dots 1011)_2$$

en donde x' se ha obtenido por truncamiento y x'' mediante redondeo por exceso. Calculamos ahora las diferencias $x - x'$ y $x'' - x$ para estimar cual es el error cometido:

$$x - x' = \frac{2}{3} \times 10^{-24}$$

$$x'' - x = \frac{1}{3} \times 10^{-24}$$

Por tanto, el número más próximo es $fl(x) = x''$ y los errores de redondeo absoluto y relativo son:

$$|fl(x) - x| = \frac{1}{3} \times 10^{-24}$$

$$\left| \frac{fl(x) - x}{x} \right| = 2^{-25} < 2$$

1.4. Propiedades cualitativas del sistema de numeros de maquina en punto flotante

Vamos a analizar el resultado de operar sobre dos números en punto flotante normalizado de l -dígitos de longitud, x e y , que producen un resultado normalizado de l -dígitos. Expresaremos esta operación como:

$$fl(x \text{ op } y)$$



en donde op es $+$, $-$, \times ó \div . Supondremos que en cada caso la mantisa del resultado es primero normalizada y después redondeada (operación que puede dar lugar a un desbordamiento que requeriría renormalizar el número). El valor de la mantisa redondeada a p bits, q_r , se define como (de una forma más rigurosa que en el caso anterior):

$$q_r = \begin{cases} 2^{-p} \lfloor 2^p q + \frac{1}{2} \rfloor & q \geq 0 \\ 2^{-p} \lceil 2^p q - \frac{1}{2} \rceil & q < 0 \end{cases}$$

en donde la función *redondeo por defecto* $\lfloor x \rfloor$ es el mayor entero menor o igual a x y la función *redondeo por exceso* $\lceil x \rceil$ es el menor entero mayor o igual a x . Para números enteros, esta función se traduce en la bien conocida regla de sumar 1 en la posición $p + 1$. Teniendo en cuenta sólo la mantisa, redondear de este modo da lugar a un intervalo máximo del error de:

$$|\epsilon_a| \leq 2^{-p-1} \quad (21)$$

y un error relativo máximo en el intervalo:

$$|\epsilon_r| \leq \frac{2^{-p-1}}{1/2} = 2^{-p} \quad (22)$$

Analizaremos ahora el error generado por cada una de las operaciones básicas:

1.4.1. MULTIPLICACIÓN

La operación de multiplicar dos números expresados en punto flotante implica sumar los exponentes y multiplicar las mantisas. Si la mantisa resultante no está normalizada, se recurre a renormalizar el resultado ajustando adecuadamente el exponente. Después, es necesario redondear la mantisa a p bits. Para analizar el error de esta operación supongamos dos números:

$$x = q_x 2^{f_x}; \quad y = q_y 2^{f_y}$$

Tenemos entonces que el producto será:

$$xy = q_x q_y 2^{f_x + f_y}$$

en donde el valor de la mantisa se encontrará en el rango:

$$\frac{1}{4} \leq |q_x q_y| < 1$$



ya que tanto x como y satisfacen la ecuación (19). Por tanto, la normalización del producto $q_x q_y$ implica un desplazamiento a la derecha de, como máximo, una posición. La mantisa redondeada será entonces uno de estos dos posibles valores:

$$x = q_x q_y + \varepsilon \quad \text{ó} \quad x = 2q_x q_y + \varepsilon$$

en donde ε , que es el error de redondeo, cumple la ecuación (21). Tenemos entonces:

$$\begin{aligned} fl(x \times y) &= \begin{cases} (q_x q_y + \varepsilon) 2^{f_x + f_y} & |q_x q_y| \geq \frac{1}{2} \\ (2q_x q_y + \varepsilon) 2^{f_x + f_y} & \frac{1}{2} > |q_x q_y| \geq \frac{1}{4} \end{cases} \\ &= q_x q_y 2^{f_x + f_y} \begin{cases} 1 + \frac{\varepsilon}{q_x q_y} & |q_x q_y| \geq \frac{1}{2} \\ 1 + \frac{\varepsilon}{2q_x q_y} & \frac{1}{2} > |q_x q_y| \geq \frac{1}{4} \end{cases} \\ &= xy(1 + \varepsilon_q) \end{aligned}$$

en donde, de acuerdo con la ecuación (22), tenemos:

$$|\varepsilon_q| \leq 2|\varepsilon| \leq 2^{-p}$$

Por tanto, la cota del error relativo en la multiplicación es la misma que la que surge por redondear la mantisa.

1.4.2. **DIVISIÓN**

Para llevar a cabo la división en punto flotante, se divide la mitad de la mantisa del numerador por la mantisa del denominador (para evitar cocientes mayores de la unidad), mientras que los exponentes se restan. Esto es:

$$\frac{x}{y} = \frac{q_x / 2}{q_y} 2^{f_x - f_y + 1}$$

Puesto que ambas mantisas satisfacen la ecuación (18), el valor del cociente estará acotado entre los límites:

$$\frac{1}{4} \leq \left| \frac{q_x}{2q_y} \right| < 1$$

Aplicando un análisis similar al empleado en el caso de la multiplicación, obtenemos:

$$\begin{aligned} fl(x / y) &= \begin{cases} (q_x / 2q_y + \varepsilon) 2^{f_x - f_y + 1} & |q_x / 2q_y| \geq \frac{1}{2} \\ (q_x / q_y + \varepsilon) 2^{f_x - f_y} & \frac{1}{2} > |q_x / 2q_y| \geq \frac{1}{4} \end{cases} \\ &= q_x / 2q_y 2^{f_x - f_y + 1} \begin{cases} 1 + \frac{\varepsilon}{q_x / 2q_y} & |q_x / 2q_y| \geq \frac{1}{2} \\ 1 + \frac{\varepsilon}{q_x / q_y} & \frac{1}{2} > |q_x / 2q_y| \geq \frac{1}{4} \end{cases} \\ &= \frac{x}{y} (1 + \varepsilon_d) \end{aligned}$$



en donde, de acuerdo con la ecuación (22), tenemos:

$$|\epsilon_d| \leq 2|\epsilon| \leq 2^{-p}$$

Es decir, la cota máxima del error relativo en la división, como en el caso anterior, es la misma que la que surge por redondear la mantisa.

1.4.3. SUMA Y RESTA

La operación de suma o resta se realiza del siguiente modo: se toma la mantisa del operando de menor magnitud (supongamos que es y) y se desplaza $f_x - f_y$ posiciones a la derecha. La mantisa resultante es sumada (o restada) y el resultado se normaliza y después se redondea. Es decir:

$$x \pm y = (q_x \pm q_y 2^{f_y - f_x}) 2^{f_x}$$

El análisis del error cometido en esta operación es más complejo que los estudiados hasta ahora, por lo que no lo vamos a ver en detalle. Sin embargo, el resultado final indica que la cota máxima del error cometido en la adición y la sustracción viene dado por:

$$|\epsilon_a| \leq 2|\epsilon| \leq 2^{-p}$$

En conclusión, en todas las operaciones aritméticas elementales en punto flotante, el error absoluto del resultado es no mayor de 1 en el bit menos significativo de la mantisa.

Sin embargo, los errores de redondeo se acumulan a medida que aumenta el número de cálculos. Si en el proceso de calcular un valor se llevan a cabo N operaciones aritméticas es posible obtener, en el mejor de los casos, un error de

$$\sqrt{N}\epsilon_m$$

redondeo total del orden de ⁴ (que coincide con el caso en que los errores de redondeo están aleatoriamente distribuidos, por lo que se produce una cancelación parcial). Desafortunadamente, este error puede crecer muy rápidamente por dos motivos:

- Es muy frecuente que la regularidad del cálculo o las peculiaridades del computador den lugar a que el error se acumule preferentemente en una dirección; en cuyo caso el error de redondeo se puede aproximar a $N\epsilon_m$.
- En circunstancias especialmente desfavorables pueden existir operaciones que incremente espectacularmente el error de redondeo. Generalmente, este fenómeno se da cuando se calcula la diferencia entre dos números muy próximos, dando lugar a un resultado en el cual los únicos bits significativos que no se cancelan son los de menor orden (en los únicos en que difieren). Puede parecer que la probabilidad de que se de dicha situación es pequeña, sin embargo, algunas expresiones matemáticas fomentan este fenómeno.

Veamos con un ejemplo los problemas comentados anteriormente. Hay dos formas de calcular las soluciones de la familiar ecuación cuadrática:

$$ax^2 + bx + c = 0$$



que son:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (23)$$

$$x = \frac{2c}{-b \pm \sqrt{b^2 - 4ac}} \quad (24)$$

Cualquiera de estas dos expresiones da problemas cuando a , c o ambos, son pequeños. En estos casos, el valor del discriminante es muy próximo al valor de b :

$$\sqrt{b^2 - 4ac} \approx b$$

$$(|b| - \sqrt{b^2 - 4ac})$$

por lo que la diferencia viene afectada de un error de redondeo importante. En efecto, la ecuación (23) evalúa bien la raíz más grande en valor absoluto, pero da pésimos resultados al estimar la raíz menor en valor absoluto. Por otra parte, la ecuación (24) calcula bien la raíz menor (siempre en valor absoluto) pero no la raíz más grande.

La solución del problema pasa por emplear una expresión mejor condicionada. En este caso, es preferible calcular previamente:

$$q = -\frac{1}{2} \left[b + \text{sign}(b) \sqrt{b^2 - 4ac} \right] \quad (25)$$

y las dos raíces a partir de valor de q como:

$$x_1 = \frac{q}{a} \quad \text{y} \quad x_2 = \frac{c}{q} \quad (26)$$

Ejemplo: Calcular las raíces de la siguiente ecuación cuadrática:

$$ax^2 + bx + c = 0$$

siendo:

$$a = 1.0; \quad b = 1.343 \cdot 10^5; \quad 3.764 \cdot 10^{-6}$$



Solución: Empleando la ecuación (23), obtenemos:

$$\begin{aligned}x_1 &= \underline{-0.13430 \cdot 10^5} \\x_2 &= \underline{0.44676 \cdot 10^{-3}}\end{aligned}$$

Sin embargo, empleando la expresión (24):

$$\begin{aligned}x_1 &= \infty \text{ (overflow)} \\x_2 &= \underline{-0.28027 \cdot 10^{-10}}\end{aligned}$$

Por último, empleando las expresiones (25) y (26) se obtienen ambas soluciones correctas:

$$\begin{aligned}x_1 &= \underline{-0.13430 \cdot 10^5} \\x_2 &= \underline{-0.28027 \cdot 10^{-10}}\end{aligned}$$

1.5. Errores en aritmetica de punto flotante²

Al momento de aplicar las Matemáticas a situaciones del mundo real nos encontramos a menudo con problemas que no pueden ser resueltos analíticamente o de manera exacta y cuya solución debe ser abordada con ayuda de algún procedimiento numérico. A continuación consideramos algunos problemas típicos, ya formulados matemáticamente, para los cuales estudiaremos técnicas numéricas de solución.

Existen errores de redondeo, los cuales se originan debido al numero limitado de cifras que puede almacenar la computadora, estos se producen en cada operación aritmética, por esta misma situación se pueden ir creciendo.

Errores de truncamiento, este error se produce por la sustitución de un operación infinita, por una operación finita, o tambien por la sustitución de operaciones continuas por una discreta, estos aparecen en series de Taylor, integración numérica, y diferencias finitas.

² Manual del PLC DL06, 1a. edición en español, 10/04



2. Sistemas de ecuaciones algebraicas lineales

2.1. Objetivo particular

El alumno evaluara sistemas de ecuaciones lineales y realizara operaciones con matrices.

2.2. Introduccion

En este trabajo se omitirán los métodos para la resolución de sistemas con dos ecuaciones y dos incógnitas, se hablara de métodos para sistemas mas grandes.

Los sistemas de ecuaciones algebraicas lineales tienen un solo conjunto de soluciones, es decir, cada variable a resolver tiene un solo valor.

Los sistemas lineales pueden resolverse por métodos directos (método de Cramer o de determinantes, el método de eliminación de Gauss, Gauss-Jordan,) o por métodos iterativos (método de Gauss-Seidel).

El método directo es aquel que en un número finito de operaciones elementales (la adición, sustracción, multiplicación, división y la exponenciación) nos conduce a la solución exacta, salvo errores de redondeo de dicho sistema.

El método iterativo es menos susceptible a errores de redondeo y se utilizan para sistemas grandes de ecuaciones.

La solución de un sistema de ecuaciones consiste en encontrar los valores de las variables que satisfacen a todas las ecuaciones del sistema.

Algunos ejemplos de problemas que dan lugar a sistemas de ecuaciones lineales son: aplicaciones en termodinámica, en ciencias biológicas, redes eléctricas, estadística, ajuste polinomial de curvas y solución de ecuaciones diferenciales parciales.

Las características generales de los sistemas lineales son:

- Las ecuaciones deben ser linealmente independientes para que exista un conjunto único de soluciones, es decir, ninguna ecuación del sistema puede expresarse como una combinación lineal de dos o más ecuaciones del mismo sistema.
- La condición necesaria y suficiente para que un sistema de ecuaciones lineales sea linealmente independiente es que el determinante de la matriz de coeficientes $(a_{i,j})$ debe ser diferente de cero.

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} & a_{10} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & a_{29} & a_{20} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} & a_{38} & a_{39} & a_{30} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} & a_{49} & a_{40} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} & a_{58} & a_{59} & a_{50} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} & a_{68} & a_{69} & a_{60} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} & a_{78} & a_{79} & a_{70} \\ a_{81} & a_{82} & a_{83} & a_{84} & a_{85} & a_{86} & a_{87} & a_{88} & a_{89} & a_{80} \\ a_{91} & a_{92} & a_{93} & a_{94} & a_{95} & a_{96} & a_{97} & a_{98} & a_{99} & a_{90} \end{matrix}$$

■

ó en forma mas compacta

■

■

$$AX=B$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} & a_{10} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & a_{29} & a_{20} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} & a_{38} & a_{39} & a_{30} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} & a_{49} & a_{40} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} & a_{58} & a_{59} & a_{50} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} & a_{68} & a_{69} & a_{60} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} & a_{78} & a_{79} & a_{70} \\ a_{81} & a_{82} & a_{83} & a_{84} & a_{85} & a_{86} & a_{87} & a_{88} & a_{89} & a_{80} \\ a_{91} & a_{92} & a_{93} & a_{94} & a_{95} & a_{96} & a_{97} & a_{98} & a_{99} & a_{90} \end{matrix}$$



donde A : matriz de coeficientes. X : vector solución. B : vector de términos independientes.



2.3. Eliminación de gauss y gauss-Jordan para problemas ideales sencillos

2.3.1. Método de eliminación de Gauss

Se basa en la "triangulación" del sistema de ecuaciones mediante transformaciones elementales. Cuando se logra que el sistema se encuentre en forma triangular, una de las ecuaciones tiene sólo una variable o incógnita y cada una de las otras ecuaciones tiene una incógnita adicional. De esta manera, se pueden obtener los valores de las variables, uno a uno.

Las transformaciones elementales que pueden hacerse a un sistema de ecuaciones y que no alteran los valores de la solución del mismo son :

- a) Multiplicar una ecuación del sistema por una constante diferente de cero. (o dividir una ecuación del sistema entre una constante diferente de cero)
- b) Intercambiar la posición de dos ecuaciones del sistema.
- c) Sumar (o restar) a una ecuación del sistema un múltiplo de otra ecuación del sistema.

El método de eliminación de Gauss empieza "normalizando" la primera ecuación (el renglón 1 de la ecuación 3.1), dividiendo cada uno de sus coeficientes entre $a_{1,1}$. Luego, esta primera ecuación se multiplica por el coeficiente $a_{i,1}$ de cada una de las otras ecuaciones del sistema, y el producto se resta a cada ecuación correspondiente, en sucesión. El resultado será la eliminación de la primera variable de todas las ecuaciones, excepto la primera. Luego, usando las últimas $n-1$ ecuaciones y el mismo procedimiento, se elimina la segunda variable de las últimas $n-2$ ecuaciones. El proceso se repite, hasta que después de $n-1$ etapas, la forma triangular queda completa. Matemáticamente, el proceso puede describirse así: en el paso k del proceso de eliminación, los nuevos coeficientes normalizados de la ecuación k son:

y los nuevos coeficientes en las ecuaciones subsecuentes son:

$$c_{i,j} = a_{i,j} - a_{i,k}c_{k,j} \quad i > k \quad (3.3)$$

Al desarrollar este proceso, debe tenerse en cuenta que los coeficientes $a_{i,j}$ de las ecuaciones inferiores cambian durante cada etapa del proceso. Así, los coeficientes $c_{i,j}$ de una etapa se convierten en los coeficientes $a_{i,j}$ que van a usarse en la siguiente etapa.

Al final del proceso, la última ecuación contiene sólo la variable x_n , la cual puede ser re-suelta en dicha ecuación. Después, se sustituye el valor de x_n en la penúltima ecuación, que contiene sólo las variables x_n y x_{n-1} y se resuelve para x_{n-1} . Se continúa de la misma forma para encontrar los valores de todas las demás variables. A este procedimiento se le conoce como "sustitución inversa".

Ejemplo 3-1. Resuelve el siguiente sistema de ecuaciones por el método de eliminación de Gauss:

$$\begin{aligned} (1) \quad & 3x + 6y + 9z = 39 \\ (2) \quad & 2x + 5y - 2z = 3 \\ (3) \quad & x + 3y - z = 2 \end{aligned}$$

Normalizar la ecuación (1) dividiendo (1) entre 3:

$$\begin{aligned} (1') \quad & x + 2y + 3z = 13 \\ (2') \quad & 2x + 5y - 2z = 3 \\ (3') \quad & x + 3y - z = 2 \\ & \text{Restar dos veces (1') a (2') y (1') a (3')}: \\ (1'') \quad & x + 2y + 3z = 13 \end{aligned}$$



$$(2'') 0 + y - 8z = -23$$

$$(3') 0 + y - 4z = -11$$

Normalizar $(2'')$, lo cual no es necesario en este caso, y restar $(2'')$ a $(3'')$:

$$(1''') x + 2y + 3z = 13$$

$$(2''') 0 + y - 8z = -23$$

$$(3''') 0 + 0 + 4z = 12$$

Por sustitución inversa:

$$\text{de } (3'''), z = 12/4 = 3$$

$$\text{de } (2'''), y = -23 + 8z = -23 + 8(3) = 1$$

$$\text{de } (1'''), x = 13 - 2y - 3z = 13 - 2(1) - 3(3) = 2$$

Por lo tanto la solución del sistema de ecuaciones es:

$$x = 2$$

$$y = 1$$

$$z = 3$$



Se le puede agregar un refinamiento al método de eliminación de Gauss incorporando la operación de "pivoteo parcial", que consiste en seleccionar el coeficiente de mayor valor absoluto en la siguiente columna y usar la ecuación correspondiente como la base para el proceso de eliminación.

Ejemplo 3-2 Resuelve el siguiente sistema de ecuaciones por el método de eliminación de Gauss con pivoteo parcial:

$$(1) \quad x + y + z = 6$$

$$(2) \quad 2x - y + z = 3$$

$$(3) \quad 3x + 2y - z = 4$$

Intercambiar las posiciones de (1) y (3):

$$(1') \quad 3x + 2y - z = 4$$

$$(2') \quad 2x - y + z = 3$$

$$(3') \quad x + y + z = 6$$

Normalizar (1'), dividiendo (1') entre 3:

$$(1'') \quad x + \frac{2}{3}y - \frac{1}{3}z = \frac{4}{3}$$

$$(2'') \quad 2x - y + z = 3$$

$$(3'') \quad x + y + z = 6$$

Restar dos veces (1'') a (2'') y (1'') a (3''):

$$(1''') \quad x + \frac{2}{3}y - \frac{1}{3}z = \frac{4}{3}$$

$$(2''') \quad 0 - \frac{7}{3}y + \frac{5}{3}z = \frac{1}{3}$$

$$(3''') \quad 0 + \frac{1}{3}y + \frac{4}{3}z = \frac{14}{3}$$

En el siguiente paso no hay que hacer intercambio de posiciones, pues el elemento pivote (-7/3) es el de mayor valor absoluto en la columna 2 tomando en cuenta en la columna sólo de la posición del pivote hacia abajo.

Dividir (2''') entre -7/3:

$$(1^{IV}) \quad x + \frac{2}{3}y - \frac{1}{3}z = \frac{4}{3}$$

$$(2^{IV}) \quad 0 + y - \frac{5}{7}z = -\frac{1}{7}$$

$$(3^{IV}) \quad 0 + \frac{1}{3}y + \frac{4}{3}z = \frac{14}{3}$$

Restar 1/3 de (2^{IV}) a (3^{IV}):

$$(1^V) \quad x + \frac{2}{3}y - \frac{1}{3}z = \frac{4}{3}$$

$$(2^V) \quad 0 + y - \frac{5}{7}z = -\frac{1}{7}$$

$$(3^V) \quad 0 \quad 0 + \frac{33}{21}z = \frac{99}{21}$$

Por sustitución inversa:

$$z = \left(\frac{99}{21}\right) / \left(\frac{33}{21}\right) = 3$$

$$y = -\frac{1}{7} + \frac{5}{7}z = -\frac{1}{7} + \left(\frac{5}{7}\right)3 = 2$$

$$x = \frac{4}{3} - \frac{2}{3}y + \frac{1}{3}z = \frac{4}{3} - \left(\frac{2}{3}\right)2 + \left(\frac{1}{3}\right)3 = 1$$

La solución es:

$$x = 1$$

$$y = 2$$

$$z = 3$$



Hay dos ventajas al usar el pivoteo en el método de eliminación de Gauss: primero, si en una determinada etapa del proceso de eliminación el elemento pivote vale cero, se puede hacer un intercambio de posiciones en las ecuaciones, para que el elemento pivote sea diferente de cero y evitar así un error en la división (división entre cero) y, segundo, cuando los elementos pivotes son los coeficientes con mayor valor absoluto en su respectiva columna, el error por redondeo tiende a reducir



Eliminación de Gauss Jordán Este método, que constituye una variación del método de eliminación de Gauss, permite resolver hasta 15 o 20 ecuaciones simultáneas, con 8 o 10 dígitos significativos en las operaciones aritméticas de la computadora. Este procedimiento se distingue del método Gaussiano en que cuando se elimina una incógnita, se elimina de todas las ecuaciones restantes, es decir, las que preceden a la ecuación pivote así como de las que la siguen.

El método se ilustra mejor con un ejemplo. Resolvamos el siguiente conjunto de ecuaciones

$$3.0 X1 - 0.1 X2 - 0.2 X3 = 7.8500$$

$$0.1 X1 + 7.0 X2 - 0.3 X3 = - 19.3$$

$$0.3 X1 - 0.2 X2 + 10 X3 = 71.4000$$

Primero expresemos los coeficientes y el vector de términos independientes como una *matriz aumentada*.

$$\left[\begin{array}{ccc|c} 3.000 & -0.100 & -0.200 & 7.850 \\ 0.100 & 7.000 & -0.300 & -19.300 \\ 0.300 & -0.200 & 10.000 & 71.400 \end{array} \right]$$

Se normaliza el primer renglón dividiendo entre 3 para obtener:

$$\left[\begin{array}{ccc|c} 1.000 & -0.033 & -0.066 & 2.616 \\ 0.100 & 7.000 & -0.300 & -19.300 \\ 0.300 & -0.200 & 10.000 & 71.400 \end{array} \right]$$

El término $X1$ se puede eliminar del segundo renglón restando 0.1 veces el primero del segundo renglón. De una manera similar, restando 0.3 veces el primero del tercer renglón se elimina el término con $X1$ del tercer renglón.

$$\left[\begin{array}{ccc|c} 1 & -0.033 & -0.066 & 2.616 \\ 0 & 7.003 & -0.293 & -19.562 \\ 0 & -0.190 & 10.018 & 70.615 \end{array} \right]$$

En seguida, se normaliza el segundo renglón dividiendo entre 7.00333 :

$$\left[\begin{array}{ccc|c} 1 & -0.033 & -0.066 & 2.616 \\ 0 & 1.000 & -0.042 & -2.793 \\ 0 & -0.190 & 10.018 & 70.615 \end{array} \right]$$

Reduciendo los términos en $X2$ de la primera y la tercera ecuación se obtiene:



$$\left[\begin{array}{ccc|c} 1 & 0 & -0.068 & 2.524 \\ 0 & 1 & -0.042 & -2.793 \\ 0 & 0 & 10.010 & 70.084 \end{array} \right]$$

El tercer renglón se normaliza dividiéndolo entre 10.010:

$$\left[\begin{array}{ccc|c} 1 & 0 & -0.068 & 2.524 \\ 0 & 1 & -0.042 & -2.793 \\ 0 & 0 & 1.000 & 7.001 \end{array} \right]$$

Finalmente, los términos con X_3 se pueden reducir de la primera y segunda ecuación para obtener:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 3.000 \\ 0 & 1 & 0 & -2.499 \\ 0 & 0 & 1 & 7.001 \end{array} \right]$$

Nótese que no se necesita sustitución hacia atrás para obtener la solución.

Las ventajas y desventajas de la eliminación gaussiana se aplican también al método de Gauss-Jordan.

Aunque los métodos de *Gauss-Jordan* y de *eliminación de Gauss* pueden parecer casi idénticos, el primero requiere aproximadamente 50% menos operaciones. Por lo tanto, la eliminación gaussiana es el método simple por excelencia en la obtención de soluciones exactas a las ecuaciones lineales simultáneas. Una de las principales razones para incluir el método de Gauss-Jordan, es la de proporcionar un método directo para obtener la matriz inversa.

2.4. Pivoteo y eliminación conica de Gauss³

Esta operación se conoce con el nombre de **pivoteo parcial**, ya que se busca cambiar el elemento pivote. De hecho, se recomienda realizar el pivoteo parcial para reducir errores por redondeo. Los errores por redondeo se ven incrementados cuanto menor (en valor absoluto) sea el valor del elemento pivote. Por ello es recomendable, antes de comenzar el proceso de eliminación para cada columna, realizar el pivoteo parcial, llevando a la posición de la ecuación pivote aquella que posea el mayor coeficiente en valor absoluto, en esa columna. El intercambio se realiza entre la ecuación número i (cuyo elemento pivote es a_{ii}) y alguna de las $n-i$ ecuaciones que están por debajo de ella. Ilustramos esta técnica con un sistema genérico de 5×5 , en el que intercambiamos la ecuación 3 con la ecuación 5:

³ <http://bilbo.edu.uy/~cnyc/cn/sistecu/directos.html>



$$\left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} & b_2 \\ 0 & 0 & a_{33} & a_{34} & a_{35} & b_3 \\ 0 & 0 & a_{43} & a_{44} & a_{45} & b_4 \\ 0 & 0 & a_{53} & a_{54} & a_{55} & b_5 \end{array} \right)$$

Si efectuamos la operación $(E_3) \leftrightarrow (E_5)$ el sistema queda

$$\left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} & b_2 \\ 0 & 0 & a_{53} & a_{54} & a_{55} & b_5 \\ 0 & 0 & a_{43} & a_{44} & a_{45} & b_4 \\ 0 & 0 & a_{33} & a_{34} & a_{35} & b_3 \end{array} \right)$$

Los `` y los `` nos recuerdan que los coeficientes correspondientes son los obtenidos por el proceso de eliminación de Gauss de la primer y segunda columna. Vemos que luego del intercambio, el nuevo elemento a_{33} es el coeficiente que originalmente estaba en a_{53} , y viceversa.

Esta operación de pivoteo parcial, se puede combinar con otra: el intercambio de columnas de la matriz **A**. Esta última operación también está permitida, por propiedad conmutativa de la suma. Consiste, en forma similar a lo que hacemos en el pivoteo parcial, en intercambiar columnas de manera de llevar a la posición pivote a aquel elemento con mayor valor absoluto. La única precaución que debemos tomar es el recordar que al intercambiar columnas, estamos cambiando de lugar los elementos de la solución final **x**, y por lo tanto, al implementar este algoritmo en una computadora, conviene mantener un vector **p** en el que vayamos acumulando las permutaciones de columnas, de manera de poder reconstruir al final de proceso, el orden original de los elementos x_i de la solución. Esto es importante, ya que al resolver un problema real, las soluciones tienen un significado concreto y no debemos mezclar los resultados.

A la operación combinada de intercambio de filas y columnas se le conoce con el nombre de **pivoteo total**. Al igual que en el caso anterior, ilustraremos el procedimiento con una matriz generica de 5x5. Sea pues el sistema visto anteriormente, si efectuamos el intercambio de la columna 3 y la columna 4, tenemos

$$\left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{14} & a_{13} & a_{15} & b_1 \\ 0 & a_{22} & a_{24} & a_{23} & a_{25} & b_2 \\ 0 & 0 & a_{34} & a_{33} & a_{35} & b_3 \\ 0 & 0 & a_{44} & a_{43} & a_{45} & b_4 \\ 0 & 0 & a_{54} & a_{53} & a_{55} & b_5 \end{array} \right)$$



Observese que en el pivoteo parcial, se intercambian las filas de **A** y también los elementos de **b**, mientras que en el intercambio de columnas solo intervienen los elementos de **A**, nunca el vector **b**.

2.5. Problemas sin solución única

Podemos clasificar los sistemas atendiendo al número de sus soluciones:

Incompatible. No tiene solución.

Compatible. Tiene solución.

Compatible determinado. Única solución.

Compatible indeterminado. Infinitas soluciones.

$$\begin{cases} x + y = 3 \\ 2x + 2y = 6 \end{cases} \text{ incompatible. No tiene solución.}$$

$$\begin{cases} x + y = 3 \\ x - y = 1 \end{cases} \text{ compatible determinado. Única solución.}$$

$$\begin{cases} x + y = 3 \\ 2x + 2y = 6 \end{cases} \text{ compatible indeterminado. Infinitas soluciones.}$$

Para que un sistema no tenga solución única se trata de un sistema compatible indeterminado, el cual tiene infinitas soluciones, su característica es que se tratan las ecuaciones de rectas paralelas, planos, etc.

2.6. Matrices y vectores

Las matrices aparecen por primera vez hacia el año 1850, introducidas por [J.J. Sylvester](#). El desarrollo inicial de la teoría se debe al matemático [W.R. Hamilton](#) en 1853. En 1858, [A. Cayley](#) introduce la notación matricial como una forma abreviada de escribir un sistema de m ecuaciones lineales con n incógnitas.

Las matrices se utilizan en el cálculo numérico, en la resolución de sistemas de ecuaciones lineales, de las ecuaciones diferenciales y de las derivadas parciales. Además de su utilidad para el estudio de sistemas de ecuaciones lineales, las matrices aparecen de forma natural en geometría, estadística, economía, informática, física, etc.

La utilización de matrices (arrays) constituye actualmente una parte esencial en los lenguajes de programación, ya que la mayoría de los datos se introducen en los ordenadores como tablas organizadas en filas y columnas: hojas de cálculo, bases de datos,...

Definición de matriz

Se llama matriz de orden $m \times n$ a todo conjunto rectangular de elementos a_{ij} dispuestos en m líneas horizontales (filas) y n verticales (columnas) de la forma:



$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix}$$

Abreviadamente suele expresarse en la forma $A = (a_{ij})$, con $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$. Los subíndices indican la posición del elemento dentro de la matriz, el primero denota la fila (i) y el segundo la columna (j). Por ejemplo el elemento a_{25} será el elemento de la fila 2 y columna 5.

Dos matrices son iguales cuando tienen la misma dimensión y los elementos que ocupan el mismo lugar en ambas son iguales.

| Tipo de matriz | Definición | Ejemplo |
|----------------|--|---|
| FILA | Aquella matriz que tiene una sola fila, siendo su orden $1 \times n$ | $A_{1 \times 3} = (7 \quad 2 \quad -5)$ |
| COLUMNA | Aquella matriz que tiene una sola columna, siendo su orden $m \times 1$ | $A_{3 \times 1} = \begin{pmatrix} -7 \\ 1 \\ 6 \end{pmatrix}$ |
| RECTANGULAR | Aquella matriz que tiene distinto número de filas que de columnas, siendo su orden $m \times n$, $m \neq n$ | $A_{3 \times 4} = \begin{pmatrix} 1 & 3 & 2 & 9 \\ 5 & 7 & -1 & 8 \\ 0 & 3 & 5 & 1 \end{pmatrix}$ |
| TRASPUESTA | Dada una matriz A, se llama traspuesta de A a la matriz que se obtiene cambiando ordenadamente las filas por las columnas. Se representa por A^t ó A^T | Si es $A = (a_{ij})_{m \times n}$ su traspuesta es $A^t = (a_{ji})_{n \times m}$ $A = \begin{pmatrix} 1 & 2 & 5 \\ 3 & -4 & 7 \end{pmatrix}$; $A^t = \begin{pmatrix} 1 & 3 \\ 2 & -4 \\ 5 & 7 \end{pmatrix}$ |
| OPUESTA | La matriz opuesta de una dada es la que resulta de sustituir cada elemento por su opuesto. La opuesta de A es -A. | $A = \begin{pmatrix} 1 & 3 \\ 5 & -7 \\ -6 & 4 \end{pmatrix}$, $-A = \begin{pmatrix} -1 & -3 \\ -5 & 7 \\ 6 & -4 \end{pmatrix}$ |



| | | |
|-----------|---|--|
| NULA | Si todos sus elementos son cero. También se denomina matriz cero y se denota por $0_{m \times n}$ | $0_{3 \times 4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ |
| IDENTIDAD | Es una matriz cuadrada que tiene todos sus elementos nulos excepto los de la diagonal principal que son iguales a 1. También se denomina matriz unidad. | $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ |
| INVERSA | Decimos que una matriz cuadrada A tiene inversa, A^{-1} , si se verifica que : $A \cdot A^{-1} = A^{-1} \cdot A = I$ | $A = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} ; A^{-1} = \begin{pmatrix} -1 & 3 \\ 1 & -2 \end{pmatrix}$ |

Operaciones con Matrices

Entre las operaciones más sencillas podemos encontrar las siguientes:

Suma de Matrices

La suma de dos matrices $A = (a_{ij})_{m \times n}$ y $B = (b_{ij})_{p \times q}$ de la misma dimensión (equidimensionales) : $m = p$ y $n = q$ es otra matriz $C = A+B = (c_{ij})_{m \times n} = (a_{ij}+b_{ij})$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} ; B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$

$$A + B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{pmatrix}$$

p.ej.

$$A = \begin{pmatrix} 2 & -3 & 5 \\ 4 & 1 & -7 \end{pmatrix} ; B = \begin{pmatrix} 1 & 0 & 2 \\ -3 & 5 & 8 \end{pmatrix}$$

$$A + B = \begin{pmatrix} 3 & -3 & 7 \\ 1 & 6 & 1 \end{pmatrix}$$

Producto de dos matrices cualesquiera.

Definición: Sean A una matriz de orden $m \times n$, y B una matriz de orden $n \times p$, Entonces el producto de las matrices A y B es otra matriz C de orden $m \times p$ donde cada elemento c_{ij} es el producto del renglón i de A por la columna j de B, esto es:



$$c_{ij} = A_i B_j = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{pmatrix}$$

$$C = \begin{pmatrix} A_1 \cdot B_1 & A_1 \cdot B_2 & \dots & A_1 \cdot B_p \\ A_2 \cdot B_1 & A_2 \cdot B_2 & \dots & A_2 \cdot B_p \\ \dots & \dots & \dots & \dots \\ A_m \cdot B_1 & A_m \cdot B_2 & \dots & A_m \cdot B_p \end{pmatrix}$$

Obsérvese que para poder efectuar el producto AB es necesario que el número de columnas de la matriz A coincida con el número de filas de la matriz B. Esto implica que, en general, si existe el producto AB no necesariamente tiene por qué existir BA. Sin embargo, si las matrices son cuadradas y del mismo orden, siempre existen AB y BA.

Ejemplo:

$$\begin{pmatrix} 3 & 2 & 5 \\ 1 & 7 & 0 \end{pmatrix} \begin{pmatrix} 4 & 9 \\ 6 & 1 \\ 2 & 8 \end{pmatrix} = \begin{pmatrix} 12+12+10 & 27+2+40 \\ 4+42+0 & 9+7+0 \end{pmatrix} = \begin{pmatrix} 34 & 69 \\ 46 & 16 \end{pmatrix}$$

Métodos directos

Los métodos directos de resolución de sistemas lineales de ecuaciones son aquellos que permiten obtener la solución después de un número finito de operaciones aritméticas. Este número de operaciones es, obviamente, función del tamaño de la matriz.

En el cuadro siguiente se presenta un esquema con la clasificación de los procedimientos de cálculo más característicos.

Clasificación de los métodos directos

| | |
|-----------------------------------|---------------------------------------|
| | ➤ Matriz Diagonal, $A = D$ |
| • Sistemas con solución inmediata | ➤ Matriz Triangular Superior, $A = U$ |



| | |
|-----------------------------|--|
| | ➤ Matriz Triangular Inferior, $A = L$ |
| | |
| • Métodos de Eliminación | ➤ Método de Gauss |
| | ➤ Método de Gauss-Jordan |
| | |
| | ➤ Método de Doolittle, $A = LU$ |
| | ➤ Método de Crout, $A = LU$ |
| • Métodos de Descomposición | ➤ Método de Cholesky, $A = L.L^T$ |
| | ➤ Descomposición generalizada, $A=L.D.L^T$ |
| | ➤ Método de Thomas (A Tridiagonal) |
| | |

2.7. Inversion de una matriz⁴

Sea \mathbf{A} una matriz cuadrada *no singular*, es decir, que su determinante sea diferente de cero, $|\mathbf{A}| \neq 0$. Por definición de matriz inversa, se tiene que

$$\mathbf{A}^{-1}$$

es la inversa de \mathbf{A} si:

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I} \quad (13)$$

Haciendo $\mathbf{X} = \mathbf{A}^{-1}$ y sustituyendo en la ecuación anterior, se obtiene

$$\mathbf{A} \mathbf{X} = \mathbf{I} \quad (14)$$

Puede considerarse que esta ecuación matricial representa un sistema de ecuaciones simultáneas, en donde no hay un solo vector de términos independientes sino \mathbf{n} , los \mathbf{n} vectores básicos que forman la matriz unitaria \mathbf{I} . Además, no existe un solo vector de incógnitas, sino \mathbf{n} , los que corresponden a cada columna de la matriz unitaria.

Por lo anterior, es posible determinar la inversa de una matriz con el método de Gauss-Jordan de eliminación completa. Para lograrlo, bastará con aplicar las operaciones elementales sobre los

⁴ **Universidad Autónoma Metropolitana**

México

Departamento de Sistemas

Responsable del Proyecto

Lourdes Sánchez Guerrero lsg@luda.uam.mx



renglones de la matriz ampliada **(A, I)** de manera de transformar **A** en **I**. Cuando se haya hecho, se obtendrá la matriz ampliada **(I, A⁻¹)**, con lo que se tendrá la inversa buscada.

EJEMPLO

Invertir la matriz

$$A = \begin{bmatrix} 1 & -6 & 2 \\ 2 & -2 & -1 \\ 1 & -3 & -5 \end{bmatrix}$$

Auméntese la matriz de coeficientes con una matriz identidad

$$(A, I) = \left[\begin{array}{ccc|ccc} 1 & -6 & 2 & 1 & 0 & 0 \\ 2 & -2 & -1 & 0 & 1 & 0 \\ 1 & -3 & -5 & 0 & 0 & 1 \end{array} \right]$$

Usando *a11* como pivote, el renglón 1 se normaliza y se usa para eliminar a *X1* de los otros renglones.

$$\left[\begin{array}{ccc|ccc} 1 & -6 & 2 & 1 & 0 & 0 \\ 0 & 10 & -5 & -2 & 1 & 0 \\ 0 & 3 & -7 & -1 & 0 & 1 \end{array} \right]$$

En seguida, se usa *a22* como pivote y *X2* se elimina de los otros renglones.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & -\frac{1}{5} & \frac{3}{5} & 0 \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{5} & \frac{1}{10} & 0 \\ 0 & 0 & -\frac{11}{2} & -\frac{2}{5} & -\frac{3}{10} & 1 \end{array} \right]$$

Finalmente, se usa *a33* como pivote y *X3* se elimina de los renglones restantes:

$$(I, A^{-1}) = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{7}{55} & \frac{36}{55} & -\frac{2}{11} \\ 0 & 1 & 0 & -\frac{9}{55} & \frac{7}{55} & -\frac{1}{11} \\ 0 & 0 & 1 & \frac{4}{55} & \frac{3}{55} & -\frac{2}{11} \end{array} \right]$$

Por lo tanto, la inversa es:

$$A^{-1} = \begin{bmatrix} -\frac{7}{55} & \frac{36}{55} & -\frac{2}{11} \\ -\frac{9}{55} & \frac{7}{55} & -\frac{1}{11} \\ \frac{4}{55} & \frac{3}{55} & -\frac{2}{11} \end{bmatrix}$$

Se puede resolver un sistema de ecuaciones con la inversa de la matriz de coeficientes, de la siguiente manera:

$$X = A^{-1} C$$

donde **C** es el vector de términos independientes.



Comparando ambos métodos, es evidente que el método de inversión de matrices no es práctico para la solución de un sólo conjunto (o dos o tres conjuntos) de ecuaciones simultáneas, porque la cantidad de cálculos que intervienen para determinar la matriz inversa es muy grande. Sin embargo, si se desea resolver 20 conjuntos de 10 ecuaciones simultáneas que difieren únicamente en sus términos independientes, una matriz aumentada que contiene 20 columnas de constantes (que se utilizarían en el método de eliminación) sería difícil de reducir, y se podría usar con ventaja el método de inversión de matrices.

2.8. Descomposicion Lu

Factorización de matrices ($A = LU$)

Método de Doolittle

Programado en Java

El programa factoriza matrices de la forma:

$$[A] = [L][U]$$

para matrices cuadradas (N por N), utilizando para ello el siguiente algoritmo:

Inicia con el primer renglón de [U],

$$u_{1,i} = a_{1,i}, \text{ for } i = 1, 2, \dots, N;$$

Luego continua con la primera columna de [L],

$$l_{j,1} = a_{j,1} / u_{1,1}, \text{ for } j = 2, \dots, N;$$

Para el resto de los renglones y columnas:

for $i = 1, 2, \dots, N$:

Determinar el n^{th} renglón de [U],

$$u_{n,i} = a_{n,i} - \sum_{k=1}^{n-1} l_{n,k} u_{k,i} \text{ for } i = n, \dots, N;$$

Determinar la n^{th} columna de [L],



$$l_{j,n} = \left[a_{j,n} - \sum_{k=1}^{n-1} l_{j,k} u_{k,n} \right] / u_{n,n},$$

for $j = n+1, \dots, N$;until N^{th}

2.9. Determinantes⁵

Notación matemática formada por una tabla cuadrada de números, u otros elementos, entre dos líneas verticales; el valor de la expresión se calcula mediante su desarrollo siguiendo ciertas reglas. Los determinantes fueron originalmente investigados por el matemático japonés Seki Kowa alrededor de 1683 y, por separado, por el filósofo y matemático alemán Gottfried Wilhelm Leibniz alrededor de 1693. Esta notación se utiliza en casi todas las ramas de las matemáticas y en las ciencias naturales.

⁵ Nakamura, Shoichiro.

Métodos numericos aplicados con software.

México.

Prentice may, 1994.

<http://www.terra.es/personal2/jpb00000/tdeterminantes.htm>



El determinante es un número importante asociado con toda la matriz cuadrada. Un conjunto no homogéneo de ecuaciones lineales no tiene una solución única, a menos que el determinante de la matriz de coeficientes sea distinto de cero. Por otro lado un conjunto homogéneo de ecuaciones lineales tiene más de una solución solo cuando el determinante es igual a cero. Hay muchas ocasiones en las que es necesario evaluar el determinante de una matriz.

El determinante de una matriz A de orden N se denota por $\det(A)$ y se define como:

Sea $A = (a_{ij})$ una matriz cuadrada de orden n, se define el determinante de A como

$$\det A = \sum_{\sigma \in \text{Permutaciones de orden n}} \text{signo}(\sigma) \cdot a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}$$

Donde la suma se hace sobre todas las permutaciones de los primeros subíndices de a, y (+-) toma el signo más si la permutación es par y menos si la permutación es impar.

Para una matriz de 2×2 , el determinante de A se calcula como.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$$

En el caso de tres filas por tres columnas:

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$



Propiedades de los determinantes.

1º El determinante de una matriz es igual al determinante de su matriz transpuesta :

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{12} & a_{22} & a_{32} & a_{42} \\ a_{13} & a_{23} & a_{33} & a_{43} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{vmatrix}$$

2º Si los elementos de una fila (columna) de una matriz se multiplican por un número, el determinante de la matriz queda multiplicado por dicho número.

$$\begin{vmatrix} k \cdot a_{11} & k \cdot a_{12} & k \cdot a_{13} & k \cdot a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = k \cdot \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

3º si los elementos de una fila (columna) de una matriz se pueden descomponer en dos sumandos, su determinante es igual a la suma de dos determinantes que tienen iguales todas las filas (columnas) excepto dicha fila (columna) cuyos sumandos pasan, respectivamente, a cada uno de los determinantes:

$$\begin{vmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} & a_{14} + b_{14} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} & a_{24} + b_{24} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} & a_{34} + b_{34} \\ a_{41} + b_{41} & a_{42} + b_{42} & a_{43} + b_{43} & a_{44} + b_{44} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} & a_{24} + b_{24} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} & a_{34} + b_{34} \\ a_{41} + b_{41} & a_{42} + b_{42} & a_{43} + b_{43} & a_{44} + b_{44} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} & a_{24} + b_{24} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} & a_{34} + b_{34} \\ a_{41} + b_{41} & a_{42} + b_{42} & a_{43} + b_{43} & a_{44} + b_{44} \end{vmatrix}$$

4º El determinante de un producto de dos matrices cuadradas coincide con el producto de los determinantes de ambas matrices:

$$\det(A \cdot B) = \det(A) \cdot \det(B)$$

5º Si en una matriz cuadrada se permutan dos filas(columnas), su determinante cambia de signo.

6º Si los elementos de una fila (columna) de una matriz cuadrada son combinación lineal de las filas (columnas) restantes, es decir son el resultado de sumar los elementos de otras filas (columnas) multiplicadas por números reales, su determinante es cero.

7º Si a los elementos de una fila (columna) de una matriz cuadrada se le suma una combinación lineal de otras filas (columnas), su determinante no varía.

Menor complementario de un elemento

3

El menor complementario de un elemento de una matriz cuadrada es el determinante de la mat. que obtenemos al suprimir su fila y su columna. Lo representamos por M_{ij} .

Ejemplo: Hallar el menor complementario del elemento a_{23} en la matriz :

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix} \Rightarrow M_{23} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2$$

Adjunto de un elemento

Es el menor complementario con signo positivo o negativo según sea par o impar la suma de su número de fila y su número de columna. Lo representamos por A_{ij}



$$\begin{vmatrix}
 a_{11} & a_{12} & a_{13} & a_{14} \\
 a_{21} & a_{22} & a_{23} & a_{24} \\
 a_{31} & a_{32} & a_{33} & a_{34} \\
 a_{41} & a_{42} & a_{43} & a_{44}
 \end{vmatrix}
 =
 a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix}
 - a_{12} \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix}
 + a_{13} \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix}
 - a_{14} \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix}$$

Desarrollo de un determinante por los elementos de una línea.

El determinante de una matriz es igual a la suma de los productos de los elementos de una línea (fila o columna) por sus adjuntos.

Esta propiedad nos permite resolver determinantes de cualquier orden, puesto que al desarrollar por una línea, los determinantes que tenemos que calcular son de orden menor en una unidad y así, siempre podremos llegar a un determinante de orden 3, que ya sabemos calcular.

Para desarrollar por una línea es importante elegir la que más ceros tenga y utilizando la propiedad nº 7, hacer ceros todos los elementos de esa línea menos uno.



Ejemplo:

$$\begin{vmatrix} 2 & -1 & 3 & 1 \\ -1 & 3 & 1 & -2 \\ 1 & 2 & 0 & -2 \\ 1 & 2 & 2 & -1 \end{vmatrix}$$

lo más cómodo es desarrollar por la 3ª columna, haciendo en ella todos los elementos cero, menos el segundo. Es importante recordar aquí, que si multiplicamos la línea a la que sumamos la combinación lineal de las demás por un número real, el determinante queda multiplicado por ese número.

Para hacer ceros, procedemos así:

$$\begin{array}{l} f_1 \\ f_2 \\ f_3 \\ f_4 \end{array} \begin{vmatrix} 2 & -1 & 3 & 1 \\ -1 & 3 & 1 & -2 \\ 1 & 2 & 0 & -2 \\ 1 & 2 & 2 & -1 \end{vmatrix} = \begin{array}{l} f_1 - 3f_2 \\ f_2 \\ f_3 \\ f_4 - 2f_2 \end{array} \begin{vmatrix} 5 & -10 & 0 & 7 \\ -1 & 3 & 1 & -2 \\ 1 & 2 & 0 & -2 \\ 3 & -4 & 0 & 3 \end{vmatrix} = -1 \begin{vmatrix} 5 & -10 & 7 \\ 1 & 2 & -2 \\ 3 & -4 & 3 \end{vmatrix}$$

El determinante de tres por tres que queda, sabemos como desarrollarlo. Se puede comprobar como haciendo ceros los elementos de una línea, sólo tenemos que calcular un determinante de tres por tres, los demás desaparecen al estar multiplicados por cero.



2.10. Problemas mal condicionados

La obtención de la solución depende de la condición del sistema. En sentido matemático, los *sistemas bien condicionados* son aquellos en los que un cambio en uno o más coeficientes provoca un cambio similar en la solución. Los *sistemas mal condicionados* son aquellos en los que cambios pequeños en los coeficientes provocan cambios grandes en la solución.

Ejemplo: Resolver el siguiente sistema compatible determinado

$$\begin{cases} x + y + z = 11 \\ 2x - y + z = 5 \\ 3x + 2y + z = 24 \end{cases}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 11 \\ 2 & -1 & 1 & 5 \\ 3 & 2 & 1 & 24 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 11 \\ 0 & -3 & -1 & -17 \\ 0 & -1 & -2 & -9 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 11 \\ 0 & -1 & -2 & -9 \\ 0 & -3 & -1 & -17 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 11 \\ 0 & -1 & -2 & -9 \\ 0 & 0 & 5 & 10 \end{array} \right)$$

$$\begin{cases} x + y + z = 11 \\ -y - 2z = -9 \\ 5z = 10 \end{cases} \Rightarrow \begin{cases} 5z = 10 \\ -y - 4 = -9 \\ x + 5 + 2 = 11 \end{cases} \Rightarrow \begin{cases} z = 2 \\ y = 5 \\ x = 4 \end{cases}$$

Método de Cramer (por determinantes)

Es aplicable si el sistema tiene igual número de ecuaciones que de incógnitas $n=m$ y el determinante de la matriz de coeficientes es distinto de cero. Es decir, un sistema de Cramer es, por definición, compatible determinado y, por tanto, tiene siempre una solución única.

El valor de cada incógnita x_i se obtiene de un cociente cuyo denominador es el determinante de la matriz de coeficientes, y cuyo numerador es el determinante que se obtiene al cambiar la columna i del determinante anterior por la columna de los términos independientes.

Ejemplo:

Resolver el siguiente sistema compatible determinado

$$\begin{cases} x + y + z = 11 \\ 2x - y + z = 5 \\ 3x + 2y + z = 24 \end{cases} \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 11 \\ 2 & -1 & 1 & 5 \\ 3 & 2 & 1 & 24 \end{array} \right)$$

$$|A| = \begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 3 & 2 & 1 \end{vmatrix} = (-1 + 3 + 4) - (-3 + 2 + 2) = 5$$

$$|A_x| = \begin{vmatrix} 11 & 1 & 1 \\ 5 & -1 & 1 \\ 24 & 2 & 1 \end{vmatrix} = (-11 + 24 + 10) - (-24 + 22 + 5) = 20$$

$$|A_y| = \begin{vmatrix} 1 & 11 & 1 \\ 2 & 5 & 1 \\ 3 & 24 & 1 \end{vmatrix} = (5 + 33 + 48) - (15 + 24 + 22) = 25$$

$$|A_z| = \begin{vmatrix} 1 & 1 & 11 \\ 2 & -1 & 5 \\ 3 & 2 & 24 \end{vmatrix} = (-24 + 15 + 44) - (-33 + 10 + 48) = 10$$

$$x = \frac{|A_x|}{|A|} = \frac{20}{5} = 4; \quad y = \frac{|A_y|}{|A|} = \frac{25}{5} = 5; \quad z = \frac{|A_z|}{|A|} = \frac{10}{5} = 2$$

Por inversión de una Matriz

Es aplicable si el sistema tiene igual número de ecuaciones que de incógnitas $n=m$ y el determinante de la matriz de coeficientes es distinto de cero. Es decir, resuelve sistemas compatibles determinados (no-homogéneos).

Sean A y B dos matrices de $m \times n$ y se supone que $AB=BA=I$

Entonces B se llama la inversa de A y se denota por A^{-1} , entonces: $A \cdot A^{-1} = A^{-1} \cdot A = I$



Si una matriz es invertible entonces su inversa es única

Para encontrar la inversa de una matriz cuadrada se debe:

Se debe escribir la matriz aumentada $(A | I)$

Se utiliza la reducción por renglones para poner a la matriz A en su forma escalonada reducida por renglones

Se decide si A es invertible

Si la forma escalonada reducida por renglones de A es la matriz Identidad, entonces A^{-1} es la matriz que se tiene a la derecha

Si la reducción de A conduce a un renglón 0 a la izquierda de la barra vertical entonces A no es invertible

Es posible determinar la inversa de una matriz con el método de Gauss-Jordan de eliminación completa. Para lograrlo, bastará con aplicar las operaciones elementales sobre los renglones de la matriz ampliada (A, I) de manera de transformar A en I. Cuando se haya hecho, se obtendrá la

matriz ampliada (I, A^{-1}) , con lo que se tendrá la inversa buscada.

Propiedades de la matriz inversa.

1) La inversa de la matriz inversa vuelve a ser la matriz A, esto es:

$$(A^{-1})^{-1} = A$$

2) La transpuesta de la inversa es la inversa de la transpuesta:

$$(A^t)^{-1} = (A^{-1})^t$$

3) La inversa de un producto es el producto de las inversas cambiando el orden:

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$$

4) La inversa de la potencia k-ésima es la potencia k-ésima de la inversa:

$$(A^k)^{-1} = (A^{-1})^k$$

5) En una matriz A cuadrada de orden n se tiene que:

A es inversible si y solo si: $\det(A) \neq 0$ o lo que es lo mismo: $\text{rango}(A) = n$

2.11. Solucion de n ecuaciones con m incógnitas

Dado un sistema de "m" ecuaciones con "n" incógnitas se trata de obtener un sistema equivalente cuya 1ª ecuación tenga n incógnitas, la segunda n-1, la tercera n-2, y así sucesivamente hasta llegar a la última ecuación, que tendrá una sola incógnita. Hecho esto, resolvemos la última ecuación, a continuación la penúltima, y así hasta llegar a la primera. Es decir, el método de Gauss consiste en triangular la matriz de coeficientes. Para resolver un sistema de ecuaciones por este método y por otro se utiliza una matriz aumentada, pero ¿qué es una matriz aumentada? Una matriz aumentada es un arreglo que permite resolver un sistema de ecuaciones lineales. En lugar de escribir el sistema en cada paso de la eliminación gaussiana, solo se escribe el arreglo de números que muestran los coeficientes y los números independientes. La matriz aumentada de m ecuaciones y n variables es la siguiente:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & C_1 \\ a_{21} & a_{22} & \dots & a_{2n} & C_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & C_n \end{array} \right]$$

Para resolver un sistema de ecuaciones empleando la matriz aumentada se realizan las siguientes operaciones:

Intercambio de renglones

Multiplicar cualquier renglón de la matriz por un número diferente de 0

Sustituir cualquier renglón por el resultado de sumar el múltiplo de cualquier otro renglón.

A la acción de aplicar estos pasos a la matriz aumentada se le conoce como reducción de renglones y esto a su vez le permite obtener una matriz escalonada reducida.

DESVENTAJAS DEL MÉTODO DE ELIMINACIÓN



DIVISIÓN ENTRE CERO. Una de sus desventajas es que durante el proceso en las fases de eliminación y sustitución es posible que ocurra una división entre cero. Se ha desarrollado una estrategia del pivoteo para evitar parcialmente estos problemas. Ésta se deja como investigación al alumno.

ERRORES DE REDONDEO. La computadora maneja las fracciones en forma decimal con cierto número limitado de cifras decimales, y al manejar fracciones que se transforman a decimales que nunca terminan, se introduce un error en la solución de la computadora. Este se llama *error por redondeo*.

Cuando se va a resolver solamente un pequeño número de ecuaciones, el error por redondeo es pequeño y generalmente no se afecta sustancialmente la precisión de los resultados, pero si se van a resolver simultáneamente muchas ecuaciones, el efecto acumulativo del error por redondeo puede introducir errores relativamente grandes en la solución. Por esta razón el número de ecuaciones simultáneas que se puede resolver satisfactoriamente con el método de eliminación de Gauss, utilizando de 8 a 10 dígitos significativos en las operaciones aritméticas, se limita generalmente a 15 o 20.



3. Interpolacion

3.1. Objetivo particular

El alumno aprendera a evaluar con diferentes metodos y según sus necesidades.

3.2. Introduccion

Muchas veces, de una función sólo conocemos un conjunto de valores. Esto puede suceder, por ejemplo, porque son los resultados de un experimento gobernado por una ley que desconocemos. Si queremos calcular el valor de la función para una abscisa diferente de las conocidas, debemos utilizar otra función que la aproxime y, naturalmente, el valor que obtengamos será una aproximación del valor real. También puede suceder que sepamos la expresión analítica de la función, pero sea lo suficientemente complicada como para calcular aproximaciones a los valores de la función a partir de otros ya conocidos.

Existen varias formas de hacer esto, pero la más sencilla y una de las más utilizadas es la interpolación, que consiste en construir una función que pase por los valores conocidos (llamados polos) y utilizar ésta como aproximación de la función primitiva. Si se utilizan polinomios como funciones de aproximación, hablamos de interpolación polinómica.

Si la abscisa para la que queremos encontrar un valor aproximado de la función se encuentra fuera del mayor intervalo definido por las abscisas de los polos, se dice que estamos haciendo extrapolación.

Siempre que se utiliza un valor aproximado se está cometiendo un error. El estudio del error queda fuera de los límites del curso al que está dirigido esta unidad didáctica.

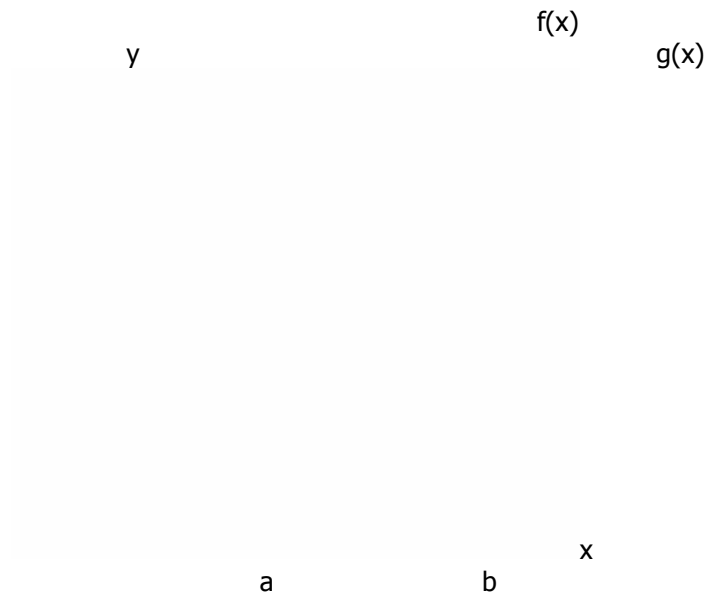
3.3. Interpolacion lineal

Esta interpolación es la base para varios modelos numéricos fundamentales. Al integrar la interpolación lineal, se define el modelo de integración llamado regla del trapecio. El gradiente de la interpolación lineal es una aproximación a la primera derivada de la función.

La interpolación lineal da como resultado una recta que se ajusta a dos puntos dados. La interpolación lineal esta dad por:

$$g(x) = (b-x/b-a)(f(a)) + (x-a/b-a)(f(b))$$

Donde $f(a)$ y $f(b)$ son valores conocidos de $f(x)$ en $x=a$ y $x=b$ respectivamente.



De este método podemos concluir:

- Por interpolación lineal se obtiene una recta que se ajusta a dos datos dados
- Si el signo de $f(x)$ no cambia en a menor o igual que x y x menor o igual que b , el error máximo de una interpolación lineal aparece aproximadamente en el punto medio y su magnitud es proporcional a la segunda derivada de la función aproximada.

3.4. Formula de interpolacion de lagrange

Nuevamente tenemos los datos :

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

El polinomio de interpolación de Lagrange se plantea como sigue:

$$P(x) = y_0 l_0(x) + y_1 l_1(x) + \dots + y_n l_n(x)$$



Donde los polinomios $l_i(\mathbf{x})$ se llaman los polinomios de Lagrange, correspondientes a la tabla de datos.

Como se debe satisfacer que $P(\mathbf{x}_0)=y_0$ esto se cumple si $l_0(\mathbf{x}_0)=1$ y $l_i(\mathbf{x}_0)=0$ para toda $i \neq 0$.
 Como se debe satisfacer que $P(\mathbf{x}_1)=y_1$ esto se cumple si $l_1(\mathbf{x}_1)=1$ y $l_i(\mathbf{x}_1)=0$ para toda $i \neq 1$.
 Y así sucesivamente, veremos finalmente que la condición $P(\mathbf{x}_n)=y_n$ esto se cumple si $l_n(\mathbf{x}_n)=1$ y $l_i(\mathbf{x}_n)=0$ para toda $i \neq n$.

Esto nos sugiere como plantear los polinomios de Lagrange. Para ser más claros, analicemos detenidamente el polinomio $l_0(\mathbf{x})$. De acuerdo al análisis anterior vemos que deben cumplirse las siguientes condiciones para $l_0(\mathbf{x})$:

$l_0(\mathbf{x}_0)=1$ y $l_0(\mathbf{x}_j)=0$, para toda $j \neq 0$.

Por lo tanto, planteamos $l_0(\mathbf{x})$ como sigue:

$$l_0(x) = c(x - x_1)(x - x_2) \cdots (x - x_n)$$

Con esto se cumple la segunda condición sobre $l_0(\mathbf{x})$. La constante c se determinará para hacer que se cumpla la primera condición:

$$l_0(x_0) = 1 \Rightarrow 1 = c(x_0 - x_1)(x_0 - x_2) \cdots (x_0 - x_n) \\ \Rightarrow c = \frac{1}{(x_0 - x_1)(x_0 - x_2) \cdots (x_0 - x_n)}$$

Por lo tanto el polinomio $l_0(x)$ queda definido como:

$$l_0(x) = \frac{(x - x_1)(x - x_2) \cdots (x - x_n)}{(x_0 - x_1)(x_0 - x_2) \cdots (x_0 - x_n)}$$

Análogamente se puede deducir que:

$$l_j(x) = \frac{\prod_{i \neq j} (x - x_i)}{\prod_{i \neq j} (x_j - x_i)} \quad \text{para } j=1, \dots, n$$



Ejemplo Calcular el polinomio de Lagrange usando los siguientes datos:

| | | | | |
|-----|----|---|---|----|
| x | 1 | 3 | 5 | 7 |
| y | -2 | 1 | 2 | -3 |

Tenemos que:

$$f(x) = y_0 l_0(x) + y_1 l_1(x) + y_2 l_2(x) + y_3 l_3(x)$$

$$f(x) = -2l_0(x) + l_1(x) + 2l_2(x) - 3l_3(x)$$

donde:

$$l_0(x) = \frac{(x-3)(x-5)(x-7)}{(-2)(-4)(-6)} = \frac{(x-3)(x-5)(x-7)}{-48}$$

$$l_1(x) = \frac{(x-1)(x-5)(x-7)}{(2)(-2)(-4)} = \frac{(x-1)(x-5)(x-7)}{16}$$

$$l_2(x) = \frac{(x-1)(x-3)(x-7)}{(4)(2)(-2)} = \frac{(x-1)(x-3)(x-7)}{-16}$$

$$l_3(x) = \frac{(x-1)(x-3)(x-5)}{(6)(4)(2)} = \frac{(x-1)(x-3)(x-5)}{48}$$

Sustituyendo arriba, el polinomio de Lagrange queda definido como sigue:

$$f(x) = \left[\frac{(x-3)(x-5)(x-7)}{24} \right] + \left[\frac{(x-1)(x-5)(x-7)}{16} \right] - \left[\frac{(x-1)(x-3)(x-7)}{8} \right] - \left[\frac{(x-1)(x-3)(x-5)}{16} \right]$$



3.5. Interpolaciones de newton hacia delante y hacia atrás en puntos con igual separación⁶

3.5.1. INTERPOLACION DE NEWTON HACIA DELANTE

La formula de interpolación de Newton hacia delante que pasa por $k+1$ puntos, $f_0, f_1, f_2, \dots, f_k$, se escribe como

$$g(x) = g(x_0 + sh) = \sum_{n=0}^k \left(\frac{s}{n} \right) \Delta^n f_0$$

Esta ecuación es un polinomio de orden k ya que $\left(\frac{s}{n} \right)$ es un polinomio de orden n y su máximo orden es k . los primeros $m+1$ términos de la ecuación forman un polinomio de interpolación de orden m ajustado a los $m+1$ puntos en $x_0, x_1, x_2, \dots, x_m$. De la misma forma, los primeros $m+2$ términos forman un polinomio de interpolación de orden $m+1$ ajustado a $m+2$ puntos. Así el orden de un polinomio de interpolación se puede cambiar fácilmente modificando el número de diferencias.

⁶ METODOS NUMERICOS APLICADOS

SHOICHIRO, NAKAMURA
ED. PRENTICE HALL



3.5.2. INTERPOLACION DE NEUTON HACIA ATRÁS

La interpolación de Newton hacia atrás es otra formula dentro del uso frecuente y se escribe en términos de las diferencias hacia atrás y los coeficientes binomiales. Consideremos puntos con igual separación

$x_0, x_{-1}, x_{-2}, \dots, x_{-k}$ con un espacio constante igual a $h = x_i - x_{i-1}$

Las diferencias hacia atrás se definen como

$$\nabla^0 f_i = f_i \quad (\text{Diferencia hacia atrás de orden cero})$$

$$\nabla f_i = f_i - f_{i-1} \quad (\text{Diferencia hacia atrás de orden uno})$$

$$\nabla^2 f_i = \nabla f_i - \nabla f_{i-1} \quad (\text{Diferencia hacia atrás de orden dos})$$

$$\nabla^3 f_i = \nabla^2 f_i - \nabla^2 f_{i-1} \quad (\text{Diferencia hacia atrás de orden tres})$$

...

$$\nabla^k f_i = \nabla^{k-1} f_i - \nabla^{k-1} f_{i-1} \quad (\text{Diferencia hacia atrás de orden k})$$

También se puede desarrollar una tabla de diferencias hacia atrás.

Los coeficientes binomiales que se utilizan en las interpolaciones son:

$$\binom{s+n-1}{n} = \frac{1}{n!} (s+n-1)(s+n-2)\dots(s+1)s$$

La interpolación de Newton hacia atrás ajustada a los puntos en $x = x_0, x = x_{-1}, x = x_{-k}$ se escribe como

$$g(x) = g(x_i + sh) = \sum_{n=0}^k \binom{s+n-1}{n} \nabla^n f_{i-n}, \quad -k \leq s \leq 0$$

Una relación de equivalencia entre la diferencia hacia delante y la diferencia hacia atrás está dada por

$$\nabla^n f_i = \Delta^n f_{i-n}$$

Por lo tanto, la ecuación se puede expresar en términos de las diferencias hacia delante como

$$g(x) = \sum_{n=0}^k \binom{s+n-1}{n} \Delta^n f_{i-n}, \quad -k \leq s \leq 0$$



3.6. Interpolacion con raices de chebystev

La interpolacion dcon raices de chebystev, es una tecnica que se deriva de la la tecnica de aproximacion de Padé, excepto que cada termino x^k en la aproximación de Padé se remplaza por el polinomio de Chebyhsey de k-ésimo grado.

3.7. Polinomios de interpolacion de Hermite

Los polinomios osculantes son una generalizacion de los polinomios de Taylor y de los polinomios de Lagrange. Estos polinomios osculantes son considerados como un caso especial a los cuales se les denomina polinomios de Hermite.

3.8. Interpolacion en dos dimensiones

Dados $n+1$ datos:

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

El polinomio de interpolación de Newton se define de la siguiente manera:

$$f(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1) + \dots + b_n(x - x_0)(x - x_1)\dots(x - x_{n-1})$$

donde :

$$b_0 = f(x_0)$$

$$b_1 = f[x_1, x_0]$$

$$b_2 = f[x_2, x_1, x_0]$$

\vdots

$$b_n = f[x_n, \dots, x_0]$$

Para calcular los coeficientes b_0, b_1, \dots, b_n , es conveniente construir una tabla de diferencias divididas como la siguiente:



| | | | | |
|--------------|------|------|------|-----|
| x | -2 | -1 | 2 | 3 |
| y | 4 | 6 | 9 | 3 |
| Δy | 2 | 3 | -6 | |
| $\Delta^2 y$ | 1 | -3 | | |
| $\Delta^3 y$ | -4 | | | |

Obsérvese que los coeficientes del polinomio de interpolación de Newton, se encuentran en la parte superior de la tabla de diferencias divididas.

Ejemplo Calcular la tabla de diferencias divididas finitas con los siguientes datos:

| | | | | |
|-----|------|------|-----|-----|
| x | -2 | -1 | 2 | 3 |
| y | 4 | 6 | 9 | 3 |

Y utilizar la información de dicha tabla, para construir el polinomio de interpolación de Newton.

Procedemos como sigue:

| | | | | |
|--------------|------|------|------|-----|
| x | -2 | -1 | 2 | 3 |
| y | 4 | 6 | 9 | 3 |
| Δy | 2 | 3 | -6 | |
| $\Delta^2 y$ | 1 | -3 | | |
| $\Delta^3 y$ | -4 | | | |

Por lo tanto el polinomio de interpolación de Newton es :

$$f(x) = 4 + 2(x+2) - 0.25(x+2)(x+1) - 0.3(x+2)(x+1)(x-2)$$



3.9. Extrapolaciones

La extrapolación es utilizada cuando se desean calcular valores, los cuales se encuentran fuera del intervalo del cual se tiene información, donde los valores a extrapolar pueden estar a la derecha o izquierda del intervalo.

| | | | | |
|---|---|----|----|----|
| x | 3 | 6 | 8 | 9 |
| y | 6 | 13 | 16 | 19 |

Supongase que se desea conocer el valor de $x=10$, como se sale del intervalo este valor vendría siendo una extrapolación.



4. Ecuaciones no lineales

4.1. Objetivo particular

El alumno podrá calcular raíces de diferentes funciones eligiendo la metodología que mas le convenga.

4.2. Introduccion⁷

Con frecuencia se tienen que estimar valores intermedios entre valores conocidos. El método mas común empleado para este propósito es la *interpolación polinomial*.

Recuerdese que la fórmula general de un polinomio de ***n-ésimo*** orden es:

$$f(X) = a_0 + a_1 X + a_2 X^2 + \dots + a_n X^n \quad (1)$$

Para ***n + 1*** puntos, existe uno y sólo un polinomio de ***n-ésimo*** orden o menor que pasa a través de todos los puntos. Por ejemplo, hay sólo una línea recta (es decir un polinomio de primer orden) que conecta dos puntos. El polinomio de interpolación consiste en determinar el único polinomio de ***n-ésimo*** orden que se ajusta a los ***n + 1*** puntos dados. Este polinomio proporciona una fórmula para calcular los valores intermedios.

Aunque existe uno y sólo un polinomio de ***n-ésimo*** orden que se ajusta a los ***n + 1*** puntos, existen una gran variedad de fórmulas matemáticas mediante las cuales se puede expresar este polinomio. En esta unidad se estudian dos técnicas alternativas que están bien condicionadas para implementarse en una computadora. Estos son los polinomios de *Newton* y de *Lagrange*.

4.3. Metodo de Biseccion

Para este método debemos considerar una función continua dentro de un intervalo $[a, b]$ tal que $f(a)$ tenga diferente signo $f(a)*f(b) < 0$.

El proceso de decisión para encontrar la raíz consiste en dividir el intervalo $[a, b]$ a la mitad $c = (a+b)/2$ y luego analizar las tres posibilidades que se pueden dar.

- 1.- Si $f(a)$ y $f(c)$ tienen signos opuestos, entonces hay un cero entre $[a, c]$.
- 2.- Si $f(c)$ y $f(b)$ tienen signos opuestos, entonces, hay un cero en $[a, b]$.
- 3.- Si $f(c)$ es igual a cero, entonces c es un cero.

4.4. Metodo de falsa posicion y el metodo de la falsa posicion modificado

Una de las razones de la introducción de este método es que la velocidad de convergencia del método de bisecciones es bastante baja. En el método de bisección se usa el punto medio del

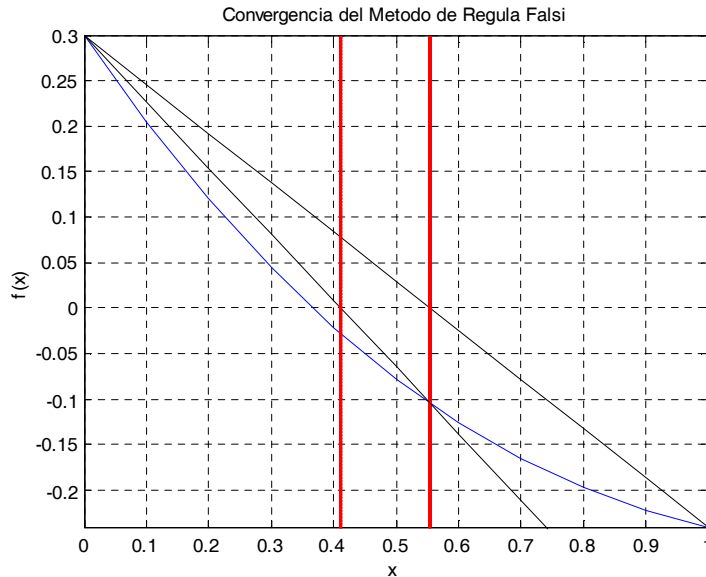
⁷ **Universidad Autónoma Metropolitana**
México
Departamento de Sistemas

Responsable del Proyecto
Lourdes Sánchez Guerrero lsq@luda.uam.mx



intervalo $[a,b]$ para llevar a cabo el siguiente paso. Suele conseguirse una mejor aproximación usando el punto $(c,0)$ en el que la recta secante L , que pasa por los puntos $[a, f(a)]$ y $[b, f(b)]$.

En la figura se puede ver como funciona el método. En esta figura en azul esta la función de la cual queremos calcular el cruce por cero y en negro dos líneas rectas que aproximan la solución.



Para calcula la ecuación de la línea secante hacemos

$$p_1 = [a, f(a)]$$

$$p_2 = [b, f(b)]$$

y sustituimos en la ecuación de la línea recta.

$$y - f(a) = (f(b) - f(a)) * (x - a) / (b - a)$$

El cruce por cero de esta ecuación esta en

$$c = a - f(a) * (b - a) / (f(b) - f(a))$$

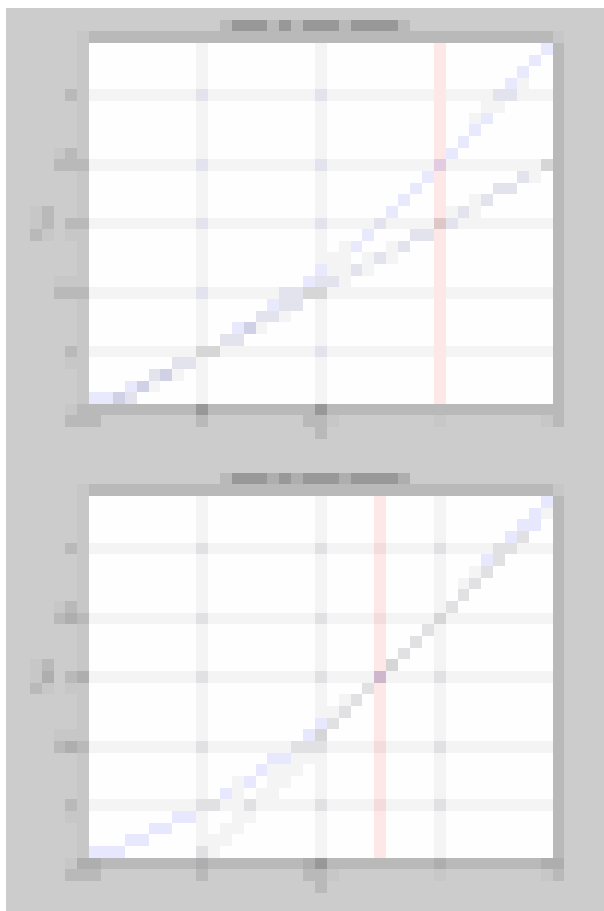
Entonces el método de bisecciones puede ser modificado, en lugar de calcular $c = (a+b)/2$ hacemos $c = a - f(a) * (b - a) / (f(b) - f(a))$ y aplicamos las mismas tres reglas de la bisección.

4.5. Metodo de newton⁸

Si tenemos una función $f(x)$ continua y cerca de una raíz p . Si la derivada $f'(x)$ existe, entonces puede utilizarse para desarrollar algoritmos que produzcan sucesiones $\{p_k\}$ que converjan a p más rápidamente que los algoritmos de bisección.

Consideremos el caso de una función como la que se muestra en la figura.

⁸ <http://scfie.fie.umich.mx/~calderon/Mnumericos/indice.html>



Dos iteraciones del método de Newton.

En estas figuras se muestra dos iteraciones del método. En la figura de la izquierda mostramos la línea recta que es tangente a la función $f(x)$ (en negro), note que la línea recta cruza el eje x en $x = 1$. A la derecha tenemos la segunda iteración tomando como valor inicial $x=1$. Note como poco a poco se acerca a la solución.

La sucesión que nos lleva a la solución esta dada por los puntos $\{p_0, p_1, p_2, \dots, p_k\}$. La pendiente de la línea recta es

$$m = (0 - f(p_0)) / (p_1 - p_0)$$

Por otro lado sabemos, del cálculo diferencial, que la pendiente de la línea tangente a una función es la primera derivada evaluada en ese punto. Así:

$$m = f'(p_0)$$

Uniendo las ecuaciones tenemos

$$f'(p_0) = (0 - f(p_0)) / (p_1 - p_0)$$



$$p_1 = p_0 - f(p_0)/f'(p_0)$$

De manera iterativa podemos hacer

$$p_2 = p_1 - f(p_1)/f'(p_1)$$

$$p_3 = p_2 - f(p_2)/f'(p_2)$$

$$p_{k+1} = p_k - f(p_k)/f'(p_k)$$

4.6. Metodo de la secante

Un problema en la implementación del método de Newton-Raphson es el de la evaluación de la derivada. Aunque esto no es un inconveniente para los polinomios y para muchas otras funciones, existen algunas funciones cuyas derivadas pueden ser en extremo difíciles de evaluar. En estos casos, la derivada se puede aproximar mediante una diferencia finita.

$$f'(x_i) \cong \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

Sustituyendo esta aproximación en la formula de Newton, tenemos la fórmula del método de la secante.

$$x_{i+1} = x_i - \frac{f(x_i)(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

4.7. Metodo de sucesion sucesiva⁹

El método de las aproximaciones sucesivas es uno de los procedimientos más importantes y más sencillos de codificar. Supongamos la ecuación

$$f(x) = 0$$

donde $f(x)$ es una función continua que se desea determinar sus raíces reales. Se sustituye $f(x)$ por la ecuación equivalente

$$x = \varphi(x)$$

Se estima el valor aproximado de la raíz x_0 , y se sustituye en el segundo miembro de la ecuación para obtener x_1 .

$$x_1 = \varphi(x_0)$$

⁹ Código fuente





Poniendo x_1 como argumento de $\phi(x)$, obtendremos un nuevo número x_2 y así sucesivamente. Este proceso se puede sintetizar en la fórmula.

$$x_n = \phi(x_{n-1}) \quad (n = 1, 2, \dots) \quad (1)$$

Si esta secuencia es convergente es decir, tiende hacia un límite, la solución α es

$$\alpha = \lim_{n \rightarrow \infty} x_n$$

■

■

■

El método de iteración se explica geométricamente mediante el gráfico de la figura. Se dibuja la curva $y=\phi(x)$, y la recta $y=x$, bisectriz del primer cuadrante. La abscisa α del punto de intersección es la raíz buscada.

Un ejemplo típico es la de encontrar la raíz de la ecuación

$$x = \cos(x)$$

Para encontrar la raíz, se comienza en el punto cualquiera de abscisa x_0 dentro del intervalo $(0, \pi/2)$, y se traza la línea vertical hasta que interseca la curva, luego, desde este punto, se traza una línea horizontal hasta que se alcanza la recta bisectriz, este punto tendrá por abscisa x_1 . Se traza de nuevo, una línea vertical hasta encontrar a la curva, y otra línea horizontal hasta encontrar la línea recta, el punto de intersección tiene de abscisa x_2 , y así sucesivamente. Como podemos apreciar en la figura, la sucesión x_1, x_2, x_3, \dots tiende hacia la raíz α de la ecuación buscada.

Tal como nos sugiere la representación gráfica de la función en la figura, la raíz buscada está en el intervalo 0 a $\pi/2$. Tomamos una aproximación inicial a la raíz x_0 en dicho intervalo y aplicamos la fórmula (1), su codificación no presenta grandes dificultades.

```
double x=0.5;
while(true){
    x=Math.cos(x);
}
```

La condición de finalización

Primero, introducimos el valor inicial x , la primera aproximación, calculamos el valor del coseno de x , el valor devuelto (segunda aproximación), lo guardamos de nuevo en la variable x , y repetimos el proceso indefinidamente. El código aunque correcto, necesita terminarse en algún momento, cumpliendo una determinada condición. Cuando el valor absoluto del cociente entre la diferencia de dos términos consecutivos de la sucesión y uno de los términos, sea menor que cierta cantidad ε

$$\left| \frac{x_{n+1} - x_n}{x_{n+1}} \right| < \varepsilon$$



Este criterio, no es completamente riguroso, pero es un buen punto de partida para el estudio de este método. Volvemos a escribir el código incluyendo la condición de terminación y dentro de una función denominada *raiz*, a la que se le pasa el valor inicial y que devuelve la raíz buscada final

En primer lugar, fijamos el valor de la constante ϵ o *ERROR*. Introducimos la primera aproximación a la raíz, y la guardamos en la variable x_0 , calculamos su coseno y obtenemos la segunda aproximación, la guardamos en la variable x_1 . Verificamos si se cumple la condición de terminación. En el caso de que no se cumpla, x_0 toma el valor de x_1 y se repite el proceso. En el momento en que se cumpla la condición de terminación, se sale del bucle y la función devuelve la raíz buscada. Como podemos observar las variables x_0 y x_1 guardan dos términos consecutivos de la sucesión que tiende hacia la raíz de la función.

La clase que describe el procedimiento

Queremos que el procedimiento numérico sea independiente de la función $f(x)$ cuya raíz deseamos averiguar. Para ello, creamos una [clase base abstracta](#) denominada *Ecuacion* con una función miembro *raiz* que defina el procedimiento numérico, y que deje sin definir la función $f(x)$, declarándola abstracta, dejando a las clases derivadas de *Ecuacion* la definición de la función $f(x)$ particular. Creamos objetos de las clases derivadas y llamamos desde ellos a la función *raiz* que describe el procedimiento numérico

En las dos primeras llamadas a la función *raiz* comprobamos que la solución buscada no depende del valor inicial de partida x_0 , que en el primer caso es 0.5 y en el segundo caso es 0.9.

El criterio de convergencia

No todas las ecuaciones pueden resolverse por este método, solamente si el valor absoluto de la derivada de la función $f'(x)$ en la vecindad de la raíz α es menor que la unidad (la pendiente de la recta bisectriz del primer cuadrante es uno). En la figura, podemos ver como es imposible encontrar la solución marcada por un puntito negro en la intersección entre la curva y la recta bisectriz del primer cuadrante, ya que la sucesión x_i diverge.



Por ejemplo, la ecuación

$$f(x) = x^3 - x - 1 = 0$$

tiene una raíz en el intervalo (1, 2) ya que $f(1) = -1 < 0$ y $f(2) = 5 > 0$. Esta ecuación puede escribirse de la forma

$$x = x^3 - 1$$



En este caso,

$$\phi(x) = x^3 - 1 \quad \text{y} \quad \phi'(x) = 3x^2$$

y por tanto,

$$\phi'(x) \geq 3 \quad \text{para} \quad 1 \leq x \leq 2$$

en consecuencia, no se cumplen las condiciones de convergencia del proceso de iteración. Si escribimos la ecuación en la forma

$$x = \sqrt[3]{x+1}$$

como podrá verificar fácilmente el lector, cumple las condiciones de convergencia, obteniéndose rápidamente un valor aproximado de la raíz buscada.

Ejemplos

Resolver por el método de aproximaciones sucesivas las siguientes ecuaciones. Primero hay que ponerlas de la forma $x=f(x)$.

$$\ln x = 4 - x^2$$

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

$$x^3 = \sin x$$

$$x = \sqrt[3]{x+2}$$

$$4 - x = \tan x$$

$$e^x + \frac{1}{3}x - 1 = 0$$

4.8. Metodo de Bairstow

El método de Bairstow es un proceso iterativo relacionado aproximadamente con los métodos de Müller y Newton–Raphson. Recuérdese la forma factorizada de un polinomio, ej:

Si se divide entre un factor que no es una raíz (por ejemplo, $x^2 + 1$), el cociente podría ser un polinomio de cuarto orden. Sin embargo, en este caso, podría haber residuo.

Con estas bases se puede elaborar un algoritmo para determinar la raíz de un polinomio: 1) suponiendo que el valor inicial de la raíz es a_0 , 2) al dividir el polinomio entre el factor $x^2 + 1$, y 3) determinando si existe un residuo. Si hay un residuo, el valor puede ajustarse en forma sistemática y el procedimiento repetirse hasta que el residuo desaparezca y la raíz sea localizada.

El método de Bairstow se basa por lo general en esta aproximación. El proceso matemático depende de dividir el polinomio entre un factor. Por ejemplo, el polinomio general

Puede dividirse entre el factor para producir un segundo polinomio que dé un orden más bajo, con un residuo, donde los coeficientes son calculados por la relación de recurrencia.

para $a = 0$

Obsérvese que si t fue una raíz del polinomio original, el residuo b_0 sería igual a cero.

Para permitir la evaluación de raíces complejas, el método de Bairstow divide el polinomio entre un factor cuadrático. El resultado es un nuevo polinomio con un residuo

Como con una división sintética normal, la simple relación de recurrencia puede usarse para realizar la división entre un factor cuadrático:

para $a = 0$

El factor cuadrático se introduce para permitir la determinación de las raíces complejas.

Los cambios, Δa y Δb , necesarios para mejorar nuestros valores iniciales se pueden estimar por medio de:



Bairstow muestra que las derivadas parciales pueden obtenerse por división sintética de las b en forma similar al camino en el cual las b en si mismas fueron derivadas:
para $a = 0$

Entonces, las derivadas parciales se obtienen por división sintética de las b . Así, las derivadas pueden sustituirse en las ecuaciones anteriores junto con las b para dar:

Estas ecuaciones pueden resolverse para y , las cuales pueden emplearse para mejorar los valores iniciales de r y s . en cada paso, el error aproximado en r y s puede ser estimado como en

Cuando ambos errores estimados fallan bajo un criterio especificado de paro, , los valores de las raíces pueden determinarse.



5. Metodo de minimos cuadrados

5.1. Objetivo particular

El alumno podra pronosticar valores con diferentes metodologias.

5.2. Introducción

El siguiente trabajo tiene por objeto exponer el Método de Mínimos Cuadrados, dentro del cual manejamos su concepto, características y un ejemplo de los mismos.

Para lograr este fin, se realizo la consulta de una bibliografía básica la cual permitió desarrollar los conceptos y ejemplos, como base para realizar una exposición adecuada del tema.

En este trabajo básicamente se habla de cómo desarrollar la aplicación de los métodos lineales y estimación por mínimos cuadrados.

El Método de Mínimos Cuadrados forma parte de los Modelos Lineales, de tal suerte que este es una de las técnicas estadísticas más utilizadas desde un punto de vista práctico. No obstante, el método de mínimos cuadrados, utilizado en la estimación de los parámetros mediante el modelo de regresión lineal es un método simple pero que proporciona unos estimadores muy sensibles a la posible presencia de datos extraños y a la no normalidad del error aleatorio.

5.3. Regresion lineal

El [procedimiento](#) mas [objetivo](#) para ajustar una recta a un conjunto de [datos](#) presentados en un [diagrama](#) de dispersión se conoce como "el [método](#) de los mínimos cuadrados". La recta resultante presenta dos [caracter](#)ísticas importantes:

1. Es nula la suma de las desviaciones verticales de los puntos a partir de la recta de ajuste

$$\sum (Y - \hat{Y}) = 0.$$

2. Es mínima la suma de los cuadrados de dichas desviaciones. Ninguna otra recta daría una suma menor de las desviaciones elevadas al cuadrado $\sum (Y - \hat{Y})^2 \rightarrow 0$ (mínima).

El [procedimiento](#) consiste entonces en minimizar los residuos al cuadrado C_i^2

$$\sum C_i^2 = \sum (Y^o - \hat{Y})^2$$

Re emplazando \hat{Y} nos queda

$$\sum C_i^2 = \sum [Y^o - (a + bx)]^2$$



La obtención de [los valores](#) de a y b que minimizan esta [función](#) es un problema que se puede resolver recurriendo a la derivación parcial de la [función](#) en términos de a y b: llamemos G a la función que se va a minimizar:

$$G = \sum (y - a - bx)^2$$

Tomemos las [derivadas](#) parciales de G respecto de a y b que son las incógnitas y las igualamos a cero; de esta forma se obtienen dos [ecuaciones](#) llamadas [ecuaciones](#) normales del [modelo](#) que pueden ser resueltas por cualquier [método](#) ya sea igualación o [matrices](#) para obtener los [valores](#) de a y b.

$$G = \sum (y - a - bx)^2$$

Derivamos parcialmente la ecuación respecto de a

$$\frac{dG}{da} = 2 \sum (y - a - bx) (-1) = 0$$

$$= -2 \sum (y - a - bx) = 0$$

$$= \sum (y - a - bx) = 0$$

$$= \sum y - na - b \sum x = 0$$

$$\sum y = na + b \sum x$$

Primera ecuación normal

Derivamos parcialmente la ecuación respecto de b

$$\frac{dG}{db} = 2 \sum (y - a - bx) (-x) = 0$$

$$= -2 \sum (y - a - bx) (x) = 0$$

$$= \sum (y - a - bx) (x) = 0$$

$$= \sum (xy - ax - bx^2) = 0$$

$$= \sum xy - a \sum x + b \sum x^2 = 0$$



$$\sum xy = a\sum x + b\sum x^2$$

Segunda ecuación normal

Los [valores](#) de a y b se obtienen resolviendo el [sistema](#) de ecuaciones resultante. Veamos el siguiente ejemplo:

En un estudio económico se desea saber la relación entre el nivel de instrucción de las personas y el ingreso.

Ejemplo:

Se toma una [muestra](#) aleatoria de 8 ciudades de una región geográfica de 13 departamentos y se determina por los [datos](#) del censo el porcentaje de graduados en [educación superior](#) y la mediana del ingreso de cada ciudad, los resultados son los siguientes:

Ciudad: 1 2 3 4 5 6 7 8 % de (X)

Graduados: 7.2 6.7 17.0 12.5 6.3 23.9 6.0 10.2

Ingreso (Y)

Mediana: 4.2 4.9 7.0 6.2 3.8 7.6 4.4 5.4 (0000)

Tenemos las ecuaciones normales

$$\sum y = na + b\sum x$$

$$\sum xy = a\sum x + b\sum x^2$$



Debemos encontrar los términos de las ecuaciones Σy , Σx , Σxy , Σx^2 Por tanto procedemos de la siguiente forma:

| Y | X | XY | X ² |
|------|------|--------|----------------|
| 4.2 | 7.2 | 30.24 | 51.84 |
| 4.9 | 6.7 | 32.83 | 44.89 |
| 7.0 | 17.0 | 119.00 | 289.00 |
| 6.2 | 12.5 | 77.50 | 156.25 |
| 3.8 | 6.3 | 23.94 | 39.69 |
| 7.6 | 23.9 | 181.64 | 571.21 |
| 4.4 | 6.0 | 26.40 | 36.00 |
| 5.4 | 10.2 | 55.08 | 104.04 |
| 43.5 | 89.8 | 546.63 | 1292.92 |

Sustituyendo en las ecuaciones los resultados obtenidos tenemos: $43.50 = 8a + 89.8b$

$546.63 = 89.8a + 1292.92b$ multiplicamos la primera ecuación por (-89.8) y la segunda por (8) así:

$$43.50 = 8a + 89.8b \quad (-89.8) \quad 546.63 = 89.8a + 1292.92b \quad (8)$$

$$-3906.30 = -718.4a - 8064.04b \quad 4373.04 = 718.4a + 10343.36b$$

$$466.74 = -0 - 2279.32b$$

$$b = \frac{466.74}{2279.32} = 0.20477$$

Este [valor](#) de b lo reemplazamos en cualquiera de las ecuaciones para obtener a así:

Reemplazando $b = 0.20477$ en la primera ecuación normal

$$43.5 = 8a + 89.8(0.20477) \quad 43.5 = 8a + 18.3880 \quad 43.5 - 18.3880 = 8a \quad 25.1120 = 8a$$



$$a = \frac{25.120}{8} = 3.139$$

Tenemos entonces que los coeficientes de regresión son: $a = 3.139$ y $b = 0.20477$.

Por tanto la ecuación de regresión nos queda:

$$\hat{Y} = 3.1390 + 0.20477(x)$$

Significa entonces que por cada incremento en una unidad en X el valor de \hat{Y} se aumenta en 0.20477

Esta ecuación permite estimar el valor de \hat{Y} para cualquier valor de X, por ejemplo: Una ciudad que tiene un porcentaje de graduados a nivel superior del 28% la mediana de ingreso para la ciudad será:

$$\hat{Y} = 3.1390 + 0.20477 (28)$$

$$\hat{Y} = 8.87 \text{ (decenas de miles de \$)}$$

Los valores a y b también se pueden obtener de la siguiente forma: partiendo de las ecuaciones normales tenemos:

$$\sum y = na + b\sum x$$

$$\sum xy = a\sum x + b\sum x^2$$

Si dividimos todos los términos de la ecuación (1) entre n nos queda:

$$\frac{\sum y}{n} = \frac{na}{n} + \frac{b\sum x}{n}$$

Tenemos entonces que el primer término es \bar{Y} el segundo término es la incógnita a y el tercer termino es la incógnita b multiplicada por \bar{X} por tanto nos queda:

$$\bar{Y} = a + b \bar{X} \text{ entonces.}$$



$$a = \bar{Y} - b \bar{X}$$

Reemplazando a en la ecuación (2) tenemos

$$\sum xy = (\bar{Y} - b \bar{X}) \sum x + b \sum x^2$$

$$b \sum x^2 = \sum xy - (\bar{Y} - b \bar{X}) \sum x$$

$$b \sum x^2 = \sum xy - \bar{Y} \sum x + b \bar{X} \sum x$$

$$b \sum x^2 = \sum xy - \frac{n \bar{Y} \sum x}{n} + \frac{nb \bar{X} \sum x}{n}$$

$$b \sum x^2 = \sum xy - n \bar{Y} \bar{X} + nb \bar{X}^2$$

$$b \sum x^2 - nb \bar{X}^2 = \sum xy - n \bar{Y} \bar{X}$$

$$b (\sum x^2 - n \bar{X}^2) = \sum xy - n \bar{Y} \bar{X}$$

$$b = \frac{\sum xy - n \bar{Y} \bar{X}}{\sum x^2 - n \bar{X}^2}$$

$$b = \frac{546.63 - 8 (5.4375) (11.2250)}{1292.92 - 8 (11.2250)^2} = \frac{58.3425}{284.9150} = 0.20477$$

$$a = 5.4375 - 0.20477 (11.2250) = 5.4375 - 2.2985 = 3.139$$

Se debe tener presente la diferencia entre el valor de \hat{Y} obtenido con la ecuación de regresión y el valor de Y observado. Mientras \hat{Y} es una estimación y su bondad en la estimación depende de lo estrecha que sea la relación entre las dos [variables](#) que se estudian; Y- es el valor efectivo,



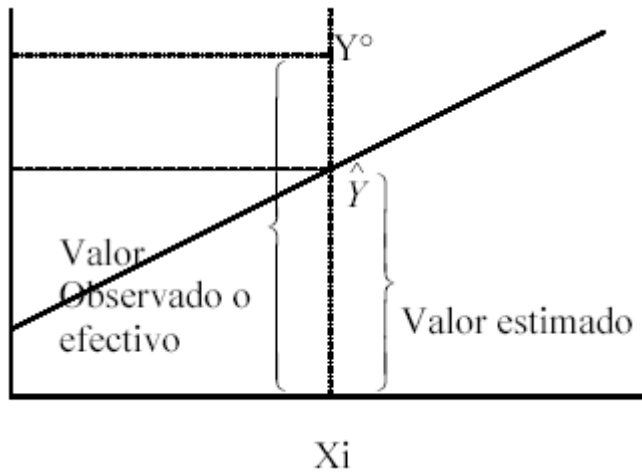
verdadero obtenido mediante la [observación](#) del investigador. En el ejemplo Y^- es el valor mediano del ingreso que obtuvo el investigador utilizando todos los [ingresos](#) observados en cada ciudad y \hat{Y} es el valor estimado con base en el [modelo](#) lineal utilizado para obtener la ecuación de regresión.

Los valores estimados y observados pueden no ser iguales, por ejemplo la primera ciudad tiene un ingreso mediano observado de $Y^- = 4.2$ al reemplazar en la ecuación el porcentaje de graduados obtenemos un \hat{Y} estimado de

$$\hat{Y} = 3.1390 + 0.20477 (1.2) = 4.61$$



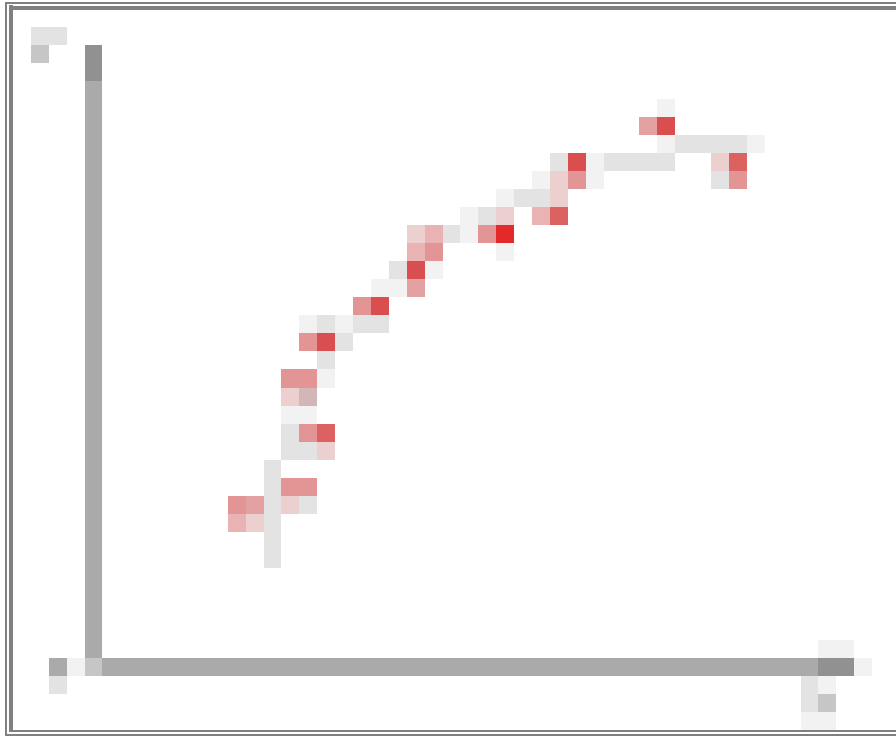
Gráficamente lo anterior se puede mostrar así:



Claramente se observa en la gráfica que hay una diferencia entre el valor efectivo de Y y el valor estimado; esta diferencia se conoce como error en la estimación, este error se puede medir.

5.4. Ajuste de curvas con un polinomio de orden superior

Supongamos que tenemos un conjunto de puntos que mostramos en la siguiente gráfica



De los puntos mostrados nos podemos dar cuenta que parece tener la forma de un polinomio de segundo grado de la forma:

$$Y = a_1 + a_2X + a_3X^2 \quad (1)$$

Esta ecuación (1) puede usarse para representar el conjunto de valores obtenidos experimentalmente para la cual debemos determinar los valores de a_1 , a_2 , a_3 , etc.

Para determinar estos valores utilizamos el siguiente procedimiento:

Establecer el criterio para determinar la ecuación que represente a los valores (obtenidos experimentalmente).

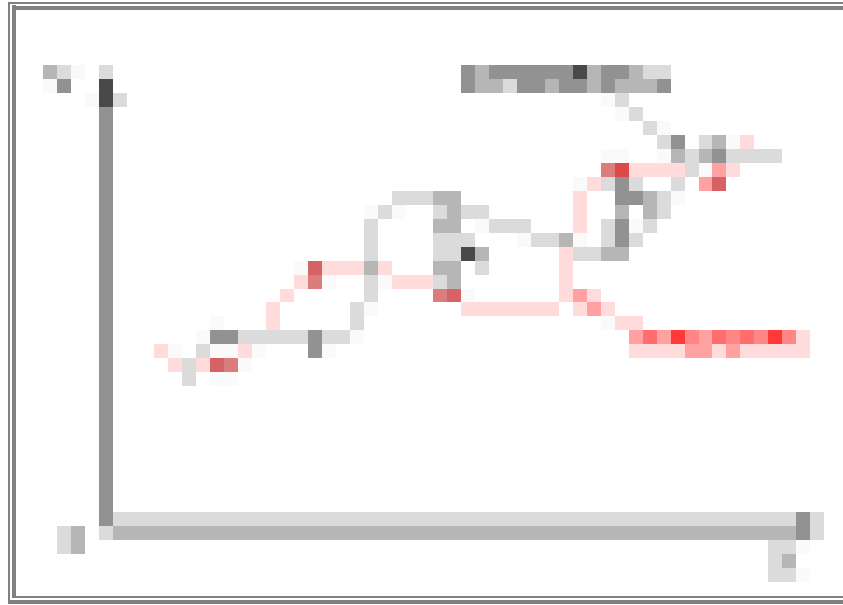
Escribir la ecuación que expresa el error o desviación entre el valor observado y los valores dados por la ecuación.

Habiendo obtenido la ecuación del error, minimizar dicho error.

EVALUACIÓN DEL ERROR

[Índice](#)

Si consideramos las parejas de datos, como se muestra en la gráfica



donde:

d = distancia = $Y_{\text{observada}} - Y_{\text{obtenida por la ecuación}}$

$Y_{\text{observada}}$ = Valor obtenido experimentalmente.

$Y_{\text{obtenida por la ecuación}}$ = valor de la función evaluada en cualquier valor X

Observando la gráfica, parece que esta distancia se puede usar para representar el error, pero habrá distancias positivas y negativas, (como se puede observar la distancia d_1 es positiva y la distancia d_2 es negativa) de modo que el error promedio para los puntos como los mostrados será pequeño aunque los errores individuales sean grandes.

Esta dificultad podría ser resuelta usando el valor absoluto de las distancias, sin embargo al derivar la función del valor absoluto se generan ciertos problemas.

La solución podría ser definir el error como el cuadrado de la distancia, esto elimina la dificultad del signo. Por esta razón el método se llama: *Método de Mínimos Cuadrados*.

$$S = d_1^2 + d_2^2 + d_3^2 + \dots + d_m^2 \quad (2)$$

en donde **S** es la suma de los cuadrados de las diferencias entre el valor calculado y el valor observado y por lo tanto es el valor que se debe minimizar

$$S = \sum_{i=1}^m (Y_{i\text{observada}} - Y_{i\text{calculada}})^2 \quad (3)$$

Siendo el caso de que la curva supuesta es una ecuación de segundo grado, se tiene la ecuación:

$$S = \sum_{i=1}^m (Y_i - a_1 - a_2 X_i - a_3 X_i^2)^2 \quad (4)$$

Para minimizar la función anterior, derivando parcialmente con respecto a a_1 , a_2 y a_3 e igualando a cero:

$$\frac{\partial S}{\partial a_1} = \frac{\partial S}{\partial a_2} = \frac{\partial S}{\partial a_3} = 0 \quad (5)$$



$$\frac{\partial S}{\partial \mathbf{a}_1} = -2 \sum_{i=1}^m (Y_i - \mathbf{a}_1 - \mathbf{a}_2 X_i - \mathbf{a}_3 X_i^2) = 0$$

$$\frac{\partial S}{\partial \mathbf{a}_2} = -2 \sum_{i=1}^m X_i (Y_i - \mathbf{a}_1 - \mathbf{a}_2 X_i - \mathbf{a}_3 X_i^2) = 0$$

$$\frac{\partial S}{\partial \mathbf{a}_3} = -2 \sum_{i=1}^m X_i^2 (Y_i - \mathbf{a}_1 - \mathbf{a}_2 X_i - \mathbf{a}_3 X_i^2) = 0$$

(Obsérvese que las variables son a_1 , a_2 y a_3 , mientras que Y_i , X_i son constantes)
Las ecuaciones se pueden expresar de acuerdo como sigue:

$$\begin{aligned} \sum_{i=1}^m \mathbf{a}_1 + \sum_{i=1}^m X_i \mathbf{a}_2 + \sum_{i=1}^m X_i^2 \mathbf{a}_3 &= \sum_{i=1}^m Y_i \\ \sum_{i=1}^m X_i \mathbf{a}_1 + \sum_{i=1}^m X_i^2 \mathbf{a}_2 + \sum_{i=1}^m X_i^3 \mathbf{a}_3 &= \sum_{i=1}^m X_i Y_i \\ \sum_{i=1}^m X_i^2 \mathbf{a}_1 + \sum_{i=1}^m X_i^3 \mathbf{a}_2 + \sum_{i=1}^m X_i^4 \mathbf{a}_3 &= \sum_{i=1}^m X_i^2 Y_i \end{aligned} \quad (6)$$

Lo anterior lo podemos expresar en forma matricial:

$$\begin{bmatrix} M & \sum_{i=1}^m X_i & \sum_{i=1}^m X_i^2 \\ \sum_{i=1}^m X_i & \sum_{i=1}^m X_i^2 & \sum_{i=1}^m X_i^3 \\ \sum_{i=1}^m X_i^2 & \sum_{i=1}^m X_i^3 & \sum_{i=1}^m X_i^4 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m Y_i \\ \sum_{i=1}^m X_i Y_i \\ \sum_{i=1}^m X_i^2 Y_i \end{bmatrix} \quad (7)$$

La fórmula general para un polinomio de grado n en donde hay m parejas de datos es:



$$\begin{bmatrix} M & \sum_{i=1}^m X_i & \sum_{i=1}^m X_i^2 & \dots & \sum_{i=1}^m X_i^n \\ \sum_{i=1}^m X_i & \sum_{i=1}^m X_i^2 & \sum_{i=1}^m X_i^3 & \dots & \sum_{i=1}^m X_i^{n+1} \\ \sum_{i=1}^m X_i^2 & \sum_{i=1}^m X_i^3 & \sum_{i=1}^m X_i^4 & \dots & \sum_{i=1}^m X_i^{n+2} \\ \dots & \dots & \dots & \dots & \dots \\ \sum_{i=1}^m X_i^n & \sum_{i=1}^m X_i^{n+1} & \sum_{i=1}^m X_i^{n+2} & \dots & \sum_{i=1}^m X_i^{2n} \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \dots \\ \mathbf{a}_{n+1} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m Y_i \\ \sum_{i=1}^m X_i Y_i \\ \sum_{i=1}^m X_i^2 Y_i \\ \dots \\ \sum_{i=1}^m X_i^n Y_i \end{bmatrix} \quad (8)$$

Como se puede observar el problema consiste en lo siguiente:

Obtener la matriz de coeficientes.

Resolver el sistema de ecuaciones resultantes.

Recordando que:

Si **n** es el grado del polinomio, hay **n+1** valores de la matriz de coeficientes y **n+1** ecuaciones.

El máximo exponente de *X* en los términos de la sumatoria de **2n** puede ser que los datos no representen un polinomio de 2o grado sino que representen uno de 3o y 4o grados.

El ajuste de curvas es un procedimiento de tanteo y error, si una curva no representa los datos, entonces se intenta con un polinomio de grado superior.

5.5. Ajuste de curvas mediante una combinacion lineal de funciones conocidas¹⁰

Al ajustar una función a puntos de datos, podemos utilizar una combinación de cualesquier funciones conocidas, incluidos polinomios.

$$g(x) = c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) + \dots + c_k f_k(x)_{4.1}$$

donde f_1, f_2, \dots son funciones preescritas, c_1, c_2, \dots son coeficientes no determinados y k es el número total de funciones preescritas. Si ajustamos la ecuación 4.1 a cada punto de datos, podremos escribir una función sobredeterminada así:

$$Ac = y_{4.2}$$

Con

¹⁰ <http://ciberia.ya.com/carloesime/2.htm>



$$\begin{bmatrix} x_1(1) & x_1(2) & \dots & x_1(L) \\ x_2(1) & x_2(2) & \dots & x_2(L) \\ \vdots & \vdots & \ddots & \vdots \\ x_k(1) & x_k(2) & \dots & x_k(L) \end{bmatrix} c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_l \end{bmatrix}; \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_l \end{bmatrix}$$

donde $L > k$. Los coeficientes están determinados por
 $c = A \backslash y$

Ejemplo 1.

Determine los coeficientes de la función

$$g(x) = c_1 + c_2 x + c_3 \text{sen}(x) + c_4 \exp(x)$$

ajustado a los datos de la siguiente tabla:

| x | y |
|-----|------|
| 0.1 | 0.61 |
| 0.4 | 0.92 |
| 0.5 | 0.99 |
| 0.7 | 1.52 |
| 0.7 | 1.47 |
| 0.9 | 2.03 |



6. Optimización de funciones en una dimensión

6.1. Objetivo particular

El alumno tendrá la capacidad de identificar la metodología que mejor le convenga para la optimización de modelos.

6.2. Introducción

Los primeros intentos de aplicar de manera formal la teoría de la evolución, a problemas prácticos de ingeniería, apareció en las áreas de control de procesos estadísticos, aprendizaje de máquina y optimización de funciones. Tal vez el primer intento serio de este tipo se dio en el trabajo que realizaron Box y sus colegas en 1957, en el desarrollo de una técnica que denominaron "operación evolutiva", la cual se aplicó a una planta de manufactura para manejarla, y que se implantó sobre la base de los votos de un comité de jefes técnicos. Bajo este esquema, la planta se veía como a una especie en evolución. La calidad del producto avanzaba a través de mutaciones aleatorias y la selección era determinada por el comité.

Aunque muchas de sus ideas se usan hoy en día, Bremermann cometió el error de tratar de optimizar funciones lineales y convexas, obteniendo resultados decepcionantes, pues sus algoritmos evolutivos tenían que ser complementados con otras heurísticas para converger en una solución. Hoy sabemos que los algoritmos evolutivos difícilmente pueden competir con las técnicas tradicionales de optimización en esos dominios.

Otra técnica evolutiva dirigida particularmente a la optimización de funciones continuas de alta complejidad se desarrolló en Alemania, en 1965, por Rechenberg y Schwefel. Esta técnica, llamada estrategia evolutiva, se utilizó inicialmente para resolver problemas ingenieriles que desafiaban a los métodos de optimización tradicionales, como el gradiente conjugado, y se basa en la modificación sistemática de un vector de números reales (representando las variables de decisión del problema) mediante operadores probabilísticos, usando ciertos criterios para decidir en qué dirección dirigir la búsqueda. La estrategia evolutiva utiliza como operador principal a la mutación, y en su versión más reciente usa la cruce como operador secundario.

6.3. Método Simplex

El procedimiento se debe a Nelder y Mead. Se trata de un algoritmo simple (de ahí su nombre) que sólo requiere la evaluación de la función y no de sus derivadas. Esto le confiere diversas características. En principio es un buen

método cuando la función a optimizar no es derivable (funciones de argumentos enteros, por ejemplo) o cuando la derivada no es calculable (o de

cálculo muy costoso). Para una función analítica el método no es de los más

eficientes pero permite obtener una respuesta al problema planteado. El

procedimiento también se utiliza en el diseño de experimentos: cuando las

variables son parámetros experimentales, se va evaluando la función objetivo a



medida que el procedimiento va requiriendo la evaluación de la función en diversos puntos.

Un simplex es una figura poliédrica de $n+1$ vértices inmersa en un espacio de dimensión n . Así, por ejemplo, en el plano un triángulo es un simplex y en el espacio tridimensional lo es un tetraedro. Estas figuras, en general, no tienen que ser regulares. Los simplex que se deben considerar son los que tienen área, volumen o hipervolumen no nulos.

Supondremos que queremos encontrar un máximo de la función. A continuación se va a exponer el algoritmo de forma sucinta y simple. Más información se puede encontrar en el libro de Press et al.

Algoritmo básico simplex para encontrar un punto óptimo (en este caso el máximo) de una función $f(x)$ de n variables:

1. Considerar $n+1$ puntos del espacio n dimensional. En todos ellos se evalúa la función f . Estos puntos deben generar un hipervolumen no nulo. Ello se

puede comprobar viendo como el determinante de la matriz formada por los $n-1$ vectores *posición de cada punto respecto a uno de ellos que se toma como origen no es nulo*.

2. Del conjunto de puntos, se considera el punto asociado al valor mínimo de f : Este peor punto se debe sustituir por otro nuevo. La forma habitual de

hacerlo es considerar el centro de masas de los restantes n puntos y hacer una reflexión del peor punto a través de ese centro de masas. En esta reflexión el punto ha "recorrido" una distancia $2d$ (siendo d la distancia inicial entre el punto y el centro de masas). Se evalúa la función en el nuevo punto obtenido.

2.1. Si el valor de f es mayor que el del mejor punto, se prueba de nuevo la reflexión pero avanzando un paso $3d$ (expansión). Se

elige la opción que nos de el mejor nuevo punto. Pasar a 3.

2.2. Si el nuevo punto tiene un valor peor (menor) que el del segundo peor punto, se hace una reflexión con un paso igual a $1.5d$ (contracción).

2.2.1. Si se continúa obteniendo un valor menor que el del segundo peor punto se hace una contracción del simplex (homotecia) con un factor de 0.5 manteniendo el mejor punto. Se evalúa la función en los nuevos puntos así obtenidos. Se pasa a 2.

2.2.2. Se pasa a 3.



2.3. Se acepta el nuevo punto calculado en 2.

3. Aplicar el criterio de convergencia: Si la distancia d recorrida es menor que un cierto valor predeterminado (figura simplex muy diminuta), se acaba el proceso. En caso contrario se pasa a 2.

6.4. Optimización sin restricciones en dimensión

Maximizar o minimizar $f(x)$ S.A. $x \in [a,b]$

Para encontrar la solución óptima de este problema buscamos primero todos los máximos (o mínimos) locales. La solución óptima será el máximo local (o mínimo) con el mayor (o menor) valor de $f(x)$.

Si $a=-8$ ó $b=8$ el problema puede no tener solución.

Hay tres tipos de puntos que pueden ser máximo o mínimos locales:

- Puntos estacionarios de f : $a < x < b$ y $f'(x)=0$
- Puntos donde no existe $f'(x)$
- Puntos extremos del intervalo $[a,b]$

Teorema

Si $f'(x_0)=0$ y $f''(x_0)<(>)0$ entonces x_0 es un máximo (mínimo) local.

Teorema

Si $f'(x_0)=0$ y

1. Si la primera derivada no nula en x_0 de f es de orden impar entonces x_0 no es un extremo local

2. Si la primera derivada no nula de f en x_0 es positiva y de orden par entonces x_0 es un mínimo local



3. Si la primera derivada no nula de f en x_0 es negativa y de orden par entonces x_0 es un máximo local.

6.5. Métodos numéricos para dimensión

6.5.1. Método de búsqueda directa

Identificar el intervalo de incertidumbre que incluye el óptimo a identificar.
Reducir el tamaño del intervalo de incertidumbre hasta encontrar el óptimo

Método de búsqueda de puntos críticos

Si la función objetivo es derivable, tratamos de localizar los puntos en los que se anula la derivada utilizando por ejemplo el método de Newton

6.5.2. Búsqueda dicotómica (para máximos)

Para funciones unimodales sobre un intervalo $[a,b]$ (funciones para las que existe un punto x^* en $[a,b]$ tal que f es creciente en $[a,x^*]$ y decreciente en $[x^*,b]$)

Definir dos puntos x_1, x_2 simétricamente con respecto a a y b de modo que los intervalos $[a,x_2]$ y $[x_1,b]$ se superpongan en un intervalo de longitud ϵ .
Evaluar $f(x_1)$ y $f(x_2)$

1. Si $f(x_1) > f(x_2)$, x^* debe estar entre a y x_2
2. Si $f(x_1) < f(x_2)$, x^* debe estar entre x_1 y b
3. Si $f(x_1) = f(x_2)$, x^* debe estar entre x_1 y x_2

Repetir el proceso en el intervalo en el que se encuentra x^* hasta que sea suficientemente pequeño.

6.6. Método de Newton



Para resolver la ecuación $g(x) = 0$, g derivable

Elegir un valor inicial x_0

Calcular para $k=0,1,\dots$

$$x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}$$

Terminar cuando $|x(k)-x(k-1)| < \epsilon$

Si g'' es continua el método converge cuadráticamente cuando x_0 está suficientemente próximo a una raíz simple de g .

Maximizar o minimizar $f(x_1, x_2, \dots, x_n)$ S. A. $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$

Suponemos que existen las derivadas parciales de primer y segundo orden de f y que son continuas.

Teorema

Si x^* es un extremo local del PNL entonces para $i=1, \dots, n$ se verifica

$$\frac{\partial f(x^*)}{\partial x_i} = 0$$

es decir, x^* es un punto estacionario.

Un punto estacionario que no es extremo local es un punto de silla.

Teorema

Si x^* es un punto estacionario de f y $H_k(x^*) > 0$, $k=1, 2, \dots, n$ (menor principal de orden k de la matriz hessiana de f) entonces x^* es un mínimo local para el PNL.

Teorema



Si x^* es un punto estacionario de f y $H_k(x^*)$ tiene el mismo signo que $(-1)^k$ $k=1,2,\dots,n$ entonces x^* es un máximo local para el PNL.

Teorema

Si x^* es un punto estacionario de f , $H_n(x^*) \neq 0$ y no se satisfacen las hipótesis de los teoremas anteriores, entonces x^* es un punto de silla para el PNL.

6.7. Método del gradiente

Válido para optimizar funciones que son diferenciables continuamente dos veces.

- Idea: generar puntos sucesivos en la dirección del gradiente de la función
- Métodos:
- Método de Newton
- Método del descenso más rápido

Resolver utilizando el método de Newton el sistema de ecuaciones

- Método de Newton para sistemas
- Para resolver la ecuación $F(X)=0$, F con derivadas parciales de primer orden
- Elegir un valor inicial X_0
- Calcular

$$X^{(k+1)} = X^{(k)} - JF^{-1}(X^{(k)})F(X^{(k)})$$

Repetir hasta que $\|X(k)-X(k-1)\| < \epsilon$



6.8. Método del ascenso más rápido

Para maximizar $f(X)$

- Elegir $X(0)$
- Calcular
- $X(k+1) = X(k) + t(k)\tilde{\nabla}f(X(k))$ siendo $t(k)$ la solución del problema
- maximizar $f(X(k) + t\tilde{\nabla}f(X(k)))$
- S.A. $t \geq 0$
- Repetir hasta que $\|X(k) - X(k-1)\| < \epsilon$.





7. BIBLIOGRAFÍA

- <http://stark.udg.com/~emili/docent/qtc/optimizacion.pdf>
- <http://eco-mat.ccees.org/Libro/FUNCIONES/Funciones3.htm>
- <http://html.rincondelvago.com/algoritmo-dual-simplex.html>
- <http://www.lania.mx/newsletters/1997-invierno/evolutiv.html>
- www1.us.es/pautadatos/publico/personal/pdi/1552/1696/Proyecto%20docente%202003-2004.doc
- www.itc.mx/reforma_curricular/Licenciatura%20en%20Informatica/Matematicas%20II_LI.pdf

<http://sai.uam.mx/apoyodidactico/c2/Unidad3/unidad3.html>

<http://luda.uam.mx/curso2/tema3/sistem02.html>

[Estadística para las ciencias sociales y el comportamiento](#). Haroldo Elorza.Edit.:Oxford
Pag.163,164

[Métodos y algoritmos básicos del álgebra numérica](#). Conde Lázaro Carlos. Edit. Reverte, S.A.

www.uv.es/~diaz/mn/node12.html

