2. SOLUCIÓN DE SISTEMAS DE ECUACIONES ALGEBRAICAS LINEALES Y NO LINEALES

2.1. Álgebra matricial y teoría de los sistemas lineales

Algunas Definiciones

Los sistemas de ecuaciones lineales tienen muchas aplicaciones en todos los campos y ciencias y ya desde a. C. se tenían métodos para resolver los sistemas. Estudiaremos sobre todo el método llamado de eliminación gaussiana o de Gauss, porque es la base de los procedimientos que se utilizan para resolver un sistema con el ordenador y asimismo para el estudio de los temas que siguen, Espacios Vectoriales y Aplicaciones Lineales.

2.1 INTRODUCCIÓN Y DEFINICIONES

El término lineal proviene de línea recta que es la expresión más simple de una ecuación y que puede escribirse de la forma

$$a_1 \cdot x + a_2 \cdot y = b$$

donde a_1, a_2 (***coeficientes***) y b (***término independiente***) son ctes. tal que a_1 y a_2 no son simultáneamente cero. Dicha ecuación se llama ***ecuación lineal de incógnitas x** e **y***.

En general una ecuación lineal es cualquiera de la forma

$$a_1 \cdot x_1 + a_2 \cdot x_2 + ... + a_n \cdot x_n = b$$

donde las variables x_1, x_2, ... x_n **(*incógnitas*)** aparecen elevadas a la primera potencia y no son funciones trascendentes (lnx, cosx, e^x etc.) ni existen productos, ni raíces de las variables.

A menudo tenemos necesidad de resolver varias ecuaciones lineales al mismo tiempo, una colección finita de **m** ecuaciones lineales con **n** incógnitas x_1, x_2, ... x_n se llama un ***sistema de ecuaciones lineales o sistema lineal de m ecuaciones con n incógnitas:***

$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n =$
b_1 $a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n$
$= b_2$ $a_{31}x_1 + a_{32}x_2 + ... +$
$a_{3n}x_n = b_3$
....
....
$a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m$

donde los coeficientes y términos independientes pertenecen en nuestro caso a

Ejemplo 2.1

$$3x_1 + 2x_2 - 5x_3 = -4$$
$$-x_1 - 2x_2 + 5x_3 = 2$$

Sistema lineal de dos ecuaciones con tres incógnitas, donde las variables vienen relacionadas con operaciones de suma y resta.

Ejemplo 2.2

$$3x - \ln y = 1 -$$
$$2x + 4\,e^{y} = -2$$

Sistema no lineal, ya que las ecuaciones tienen funciones trascendentes ($\ln y$, e^{y}), es un sistema de ecuaciones sin más.

Ejemplo 2.3

$$\begin{aligned} x - y + 2z &= 0 \\ 3x + y - 3z &= 0 \\ -2x + 3y - 4z &= 0 \end{aligned}$$

Sistema lineal de tres ecuaciones con tres incógnitas, términos independientes todos nulos (***sistema homogéneo***).

Ejemplo 2.4

$$\begin{aligned} 2x - y + 3z &= 2 \\ -3x + \frac{2}{y} - z &= -1 \\ y + 2z &= 4 \end{aligned}$$

No es un sistema lineal, pues aparece una incógnita dividiendo.

Sistema en forma matricial

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots\ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots\ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots\ldots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix} \Leftrightarrow \mathbf{A \cdot X = B}$$

La matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots\ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots\ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots\ldots & a_{mn} \end{pmatrix}$$

Matriz

Una matriz es un arreglo rectangular de números encerrados por un par de corchetes. Algunos ejemplos de matrices son:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 3 & 7 \\ 1 & -1 & 5 \end{bmatrix}$$

Se pueden representar en una matriz datos agrupados por columnas, los pagos de un conjunto de activos en cada estado de la naturaleza, los coe ficientes de un sistema de ecuaciones, etc.
Por convención las matrices se representan con letras en mayúsculas.

l Ejemplo 1.1.

La matriz B puede ser la matriz de coeficientes del siguiente sistema de ecuaciones.

$$2x + 3y + 7z = 0$$

$$x \quad y + 5z = 0$$

Formalmente, se dice que la matriz

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(1)$$

es de orden m x n ya que tiene m las y n columnas, donde cada elemento a_{ij} es un número o una función que pertenece a los números reales o complejos. Por ejemplo, la matriz A es de orden 2 x 2 y la matriz B es de 2 x 3 y sus elementos pertenecen a los números reales.

Matrices cuadradas.

Cuando m = n, La matriz () tiene la misma cantidad de las que de columnas y por lo tanto es cuadrada y la llamaremos matriz cuadrada de orden n. Si una matriz es cuadrada y es de orden 1 la llamamos escalar. Un ejemplo de una matriz cuadrada es la matriz A.

Igualdad de Matrices.

Las matrices A y B son iguales, si y solo si, tienen el mismo orden y cada elemento de A es igual al elemento de correspondiente B, formalmente:

$$A = B \iff a_{ij} = b_{ij} \qquad \forall \quad i, j$$

Vectores

Un vector es un conjunto ordenado de números dispuestos en una la o una columna, también pueden entenderse como una matriz de una la o de una columna. Denominamos vector columna a una matriz de orden m x 1 y vector la a una matriz de orden 1 x n.

Suma de Matrices y Multiplicación por un Escalar

Suma y Resta de Matrices

Sean A y B son dos matrices de orden m x n, su suma (o resta) es una matriz C de orden m x n donde cada elemento de C es la suma (o resta) de los elementos correspondientes en A y B; es decir:

$$A \pm B = C \iff c_{ij} = a_{ij} \pm b_{ij} \quad i = 1, \ldots, m \quad j = 1, \ldots, n$$

Ejemplo 1.2.

Si $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 3 & 0 \\ -1 & 2 & 5 \end{bmatrix}$ entonces:

$$A + B = \begin{bmatrix} 1+2 & 2+3 & 3+0 \\ 0+(-1) & 1+2 & 4+5 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 3 \\ -1 & 3 & 9 \end{bmatrix}$$

$$A - B = \begin{bmatrix} 1-2 & 2-3 & 3-0 \\ 0-(-1) & 1-2 & 4-5 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 3 \\ 1 & -1 & -1 \end{bmatrix}$$

Cuando dos matrices tienen el mismo orden se les llama conformables para la suma.

Multiplicación de una Matriz por un Escalar

Si se suma k veces una matriz se obtendrá una matriz en que cada uno de sus elementos será k veces el inicial. A esto se le llama multiplicación por un escalar, donde el escalar es k.

$$B = kA = Ak \iff b_{ij} = ka_{ij} = a_{ij}k \quad i = 1, \ldots, m \quad j = 1, \ldots, n$$

Ejemplo

Sea $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, entonces:

$$B = A + A + A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix} = 3A = A3$$

Teorema 1.1. *Sean A, B, C matrices de un mismo orden y k un escalar, entonces se cumplen las siguientes propiedades:*

$$\begin{aligned} A+B &= B+A \\ A+(B+C) &= (A+B)+C \\ k(A+B) &= kA+kB=(A+B)k \end{aligned}$$

Multiplicación de Vectores y Matrices

Multiplicación de Vectores

Sea $A = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}$ un vector de orden 1 x m y $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ un vector m x 1, entonces la multiplicación $C = AB$ (en ese orden) está definida como la suma de las multiplicaciones de los elementos i de la matriz A con los elementos i de la matriz B, esto es:

$$AB = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = [a_1 b_1 + a_2 b_2 + \dots + a_m b_m] = \sum_{i=1}^{m} a_i b_i$$

Ejemplo 1.4.

$$(a) \quad \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = [2(1)+3(-1)+4(2)] = [7]$$

$$(b) \quad \begin{bmatrix} 3 & -1 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 6 \\ 3 \end{bmatrix} = [-6-6+12] = [0]$$

Nótese que lo que se hizo fue, simplemente, multiplicar una la con una columna, esto hay que tenerlo muy presente en la sección que sigue.

Multiplicación de Matrices

Sea A una matriz de orden m x p y B una matriz de orden p x n, es decir, la cantidad de columnas de A es igual a la cantidad de las de B, entonces cada elemento c_{ij} de la matriz C = AB (en ese orden) se obtiene multiplicando la la i de la matriz A con la columna j de la matriz B:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pn} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{bmatrix}$$

donde

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \dots + a_{ip} b_{pj} = \sum_{k=1}^{p} a_{ik} b_{jk}$$

Ejemplo 1.5

(a) si $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$ y $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, entonces AB es igual a:$AB =$
$$\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} \end{bmatrix}$$

(b) si $C = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 0 & 2 \end{bmatrix}$ y $D = \begin{bmatrix} 3 & -4 \\ 1 & 5 \\ -2 & 2 \end{bmatrix}$, entonces CD es igual a:

$$CD = \begin{bmatrix} 1(3)+2(1)+1(-2) & 1(-4)+2(5)+1(2) \\ 4(3)+0(1)+2(-2) & 4(-4)+0(5)+2(2) \end{bmatrix} = \begin{bmatrix} 3 & 8 \\ 8 & -12 \end{bmatrix}$$

Teorema 1.2. *Sean las matrices A, B y C compatibles*[1] *para la multiplicación y k un escalar, entonces se cumple que:*

$$\begin{aligned} A(BC) &= (AB)C \\ A(B+C) &= AB+AC \\ (A+B)C &= AC+BC \\ k(AB) &= (kA)B = A(kB) = (AB)k \end{aligned}$$

Sin embargo, hay que tener en cuenta que:

(a) *generalmente* $AB \neq BA$.
(b) *si* $AB = 0$ *no necesariamente implica que* $A = 0$ o $B = 0$.
(c) *si* $AB = AC$ *no necesariamente implica que* $B = C$.

Dado que $AB \neq BA$, se pueden identificar dos tipos de multiplicaciones. Por ejemplo, si tenemos la multiplicación AB podemos decir que A está premultiplicando a B pero por otro lado, se podría pensar de otra forma y decir que B está postmultiplicando a A: Una vez que ya aprendimos a sumar, restar y multiplicar matrices podemos pasar a de nir matrices con características especiales.

[1] Cuando dos matrices son compatibles para la multiplicación también se les llama conformables para la multiplicación.

Más Definiciones

Matriz Transpuesta.

La transpuesta de una matriz A se obtiene intercambiando las las de A por sus columnas, y la denotamos como A^0. Si denotamos B = A^0 podemos de nir la transpuesta de A como:

$$B = A' \Longleftrightarrow b_{ij} = a_{ji} \qquad \forall \quad i, j$$

Ejemplo 1.6

Las traspuestas de las matrices A y B del ejemplo 1, es decir A' y B', son respectivamente:

$$A' = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \quad B' = \begin{bmatrix} 2 & 1 \\ 3 & -1 \\ 7 & 5 \end{bmatrix}$$

Teorema 1.3. *Si A' y B' son las traspuestas de A y B respectivamente, y si k es un escalar cualquiera, entonces:*

$$\begin{aligned} (A')' &= A \\ (A+B)' &= A' + B' \\ (kA)' &= kA' \\ (AB)' &= B'A' \end{aligned}$$

Matriz Identidad

La matriz identidad es una matriz diagonal de unos, es decir, una matriz cuadrada donde todos los elementos fuera de la diagonal principal son ceros y los elementos dentro de la diagonal son unos.

Una matriz identidad de orden 3: $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ y una de orden 2: $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Una de las propiedades de la matriz identidad es que $AI = IA = A$.

Matrices Idempotentes

Son aquellas matrices que multiplicadas por sí mismas son ellas mismas, es decir, A^n = A 8n 2 N: Por ejemplo, si A^2 = AA = A entonces A sería una matriz idempotente. . Además, si la matriz es idempotente y simétrica se cumple que A^0A = A: Por ejemplo, la matriz identidad es idempotente y simétrica.

I Ejemplo 1.8.

La siguiente matriz es idempotente $A = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$ ya que:

$$A^2 = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix} \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix} = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$$

La inversa de una matriz

Si A y B son dos matrices cuadradas tales que AB = BA = **I**, entonces se dice que B es la inversa de A y la denotamos $B = A^{-1}$: También se puede decir que A es la inversa de B.

Ejemplo 1.9.

Sea $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$, podemos decir que una es la inversa de la otra ya que:

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 4 \end{bmatrix}\begin{bmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Si una Matriz no tiene inversa se le llama matriz singular. Análogamente, si la matriz tiene inversa se le llama matriz no singular.

La Traza de una Matriz

La traza de una matriz (tr ( )) sólo está de nida en matrices cuadradas y es la suma de los elementos de la diagonal principal de una matriz. En el caso de nuestra matriz () es la suma de $a_{11} + a_{22} + \cdots + a_{nn} = \sum a_{ii}$:

Ejemplo 1.10.

Sean A y B las matrices del ejemplo anterior, entonces:

$$(a)\ tr(A) = 1 + 3 + 4 = 8$$

$$(b)\ tr(B) = 6 + 1 + 1 = 8$$

Teorema 1.4. *Sean A, B, C y D matrices cuadradas del mismo orden y k un escalar, entonces se cumple que:*

$$\begin{aligned} tr(kA) &= ktr(A) \\ tr(A) &= tr(A') \\ tr(A+B) &= tr(A) + tr(B) \\ tr(I_n) &= n \\ tr(ABCD) &= tr(BCDA) = tr(CDAB) = tr(DABC) \end{aligned}$$

Nota Esta última propiedad es muy importante y se le llama "propiedad circular de la traza."

Matrices Particionadas

Definición

Una matriz particionada es una matriz de matrices, ésta puede representar divisiones reales o imaginarias dentro de una matriz.

Ejemplo 1.11.

Podemos particionar la matriz $B = \begin{bmatrix} 2 & 3 & 7 & 0 \\ 1 & -1 & 5 & 0 \end{bmatrix}$ de la siguiente forma:

$$\begin{bmatrix} \begin{bmatrix} 2 & 3 & 7 \\ 1 & -1 & 5 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{bmatrix}$$

o de la siguiente

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

para ser una mejor representación del siguiente sistema de ecuaciones:

$$\begin{gathered} 2x + 3y + 7z = 0 \\ x - y + 5z = 0 \end{gathered}$$

Formalmente, si se toma la matriz ()

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

y se *particiona* tomando s grupos de filas $(m_1, m_2, ..., m_s)$ y t grupos de columnas $(n_1, n_2, ..., n_t)$, entonces se podrá escribir () de la siguiente forma:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{st} \end{bmatrix}$$

donde cada elemento A_{ij} de A es una submatriz de orden m_i x n_j.

Ejemplo 1.12.

Sea la matriz particionada $A = \left[\begin{array}{cc|c} 1 & 4 & 5 \\ 2 & 9 & 3 \\ \hline 8 & 9 & 6 \end{array}\right]$, entonces se puede representar de la siguiente forma $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ donde $A_{11} = \begin{bmatrix} 1 & 4 \\ 2 & 9 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$, $A_{21} = \begin{bmatrix} 8 & 9 \end{bmatrix}$ y $A_{22} = \begin{bmatrix} 6 \end{bmatrix}$.

Un caso especial es el de la matriz diagonal por bloques

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

esta matriz es de suma importancia en econometría y por eso la estudiaremos con espe-cial atención en éste y otros capítulos.

Suma de Matrices Particionadas

Sean A y B matrices de igual orden y particionadas de la misma forma, entonces la suma de estas matrices es de la forma:

$$C = A + B = \begin{bmatrix} A_{11} + B_{11} & \dots & A_{1t} + B_{1t} \\ \vdots & \ddots & \vdots \\ A_{s1} + B_{s1} & \dots & A_{st} + B_{st} \end{bmatrix}$$

y la suma de submatrices se realiza igual que la suma de matrices.

Ejemplo 1.13.

Sea $A = \left[\begin{array}{cc|c} 1 & 4 & 5 \\ 2 & 9 & 3 \\ \hline 8 & 9 & 6 \end{array}\right]$ y $B = \left[\begin{array}{cc|c} 4 & 3 & 3 \\ 2 & 0 & 2 \\ \hline 6 & 8 & 7 \end{array}\right]$, entonces $A + B$ es igual a:

$$\left[\begin{array}{cc|c} 1 & 4 & 5 \\ 2 & 9 & 3 \\ \hline 8 & 9 & 6 \end{array}\right] + \left[\begin{array}{cc|c} 4 & 3 & 3 \\ 2 & 0 & 2 \\ \hline 6 & 8 & 7 \end{array}\right] = \begin{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 9 \end{bmatrix} + \begin{bmatrix} 4 & 3 \\ 2 & 0 \end{bmatrix} & \begin{bmatrix} 5 \\ 3 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 8 & 9 \end{bmatrix} + \begin{bmatrix} 6 & 8 \end{bmatrix} & \begin{bmatrix} 6 \end{bmatrix} + \begin{bmatrix} 7 \end{bmatrix} \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 1+4 & 4+3 & 5+3 \\ 2+2 & 9+0 & 3+2 \\ \hline 8+6 & 9+8 & 6+7 \end{array}\right] = \left[\begin{array}{cc|c} 5 & 7 & 8 \\ 4 & 9 & 5 \\ \hline 14 & 17 & 13 \end{array}\right]$$

Multiplicación de Matrices Particionadas

El estudio de *matrices particionadas* nació a partir del estudio de la multiplicación de matrices, recordemos el ejemplo 1.5. Si $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$ y $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, entonces AB es igual a $AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} \end{bmatrix}$.

La pregunta que surge es ¿Qué pasa si cada uno de los elementos de las matrices A y B son a su vez una sola matriz? La respuesta es que abría que multiplicar cada uno de estos elementos nuevamente usando la multiplicación de matrices. Pero para que esta operación esté bien de nida cada una de las submatrices de la columna i de la matriz A tienen que ser conformables para la multiplicación con cada una de las submatrices de la la i de la matriz B.

Ejemplo 1.14.

(a) Sea $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|c} 3 & 1 & 0 \\ 3 & 2 & 0 \\ \hline 1 & 0 & 1 \end{array}\right]$ y $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 1 & 1 & 0 \\ \hline 3 & 2 & 1 & 2 \end{array}\right]$,

entonces AB es igual a:

$$
\begin{aligned}
AB &= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}\begin{bmatrix} 2 & 3 & 1 \end{bmatrix} & \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}\begin{bmatrix} 2 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 \end{bmatrix}\begin{bmatrix} 2 & 3 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \end{bmatrix}\begin{bmatrix} 2 \end{bmatrix} \end{bmatrix}
\end{aligned}
$$

$$
= \begin{bmatrix} \begin{bmatrix} 4 & 3 & 3 \\ 7 & 5 & 5 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} + \begin{bmatrix} 2 \end{bmatrix} \end{bmatrix} = \left[\begin{array}{ccc|c} 4 & 3 & 3 & 0 \\ 7 & 5 & 5 & 0 \\ \hline 3 & 4 & 2 & 2 \end{array}\right].
$$

(b) Sea $A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$, entonces $A'A$ es igual a:

$$
A'A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}' \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} = \begin{bmatrix} A'_{11}A_{11} & 0 \\ 0 & A'_{22}A_{22} \end{bmatrix}.
$$

2.2. Métodos de solución de sistemas de ecuaciones lineales sustitución de Gauss

2.2.1. Eliminación Gaussiana

Eliminación Gaussiana.

Considerar un sistema lineal.

1. Construya una matriz argumentada del sistema.
2. Use operaciones elementales en los renglones para transformar la matriz argumentada en una triangular
3. Escriba abajo el nuevo sistema lineal para cada matriz asociada al a matriz argumentada.
4. Resuelva el nuevo sistema. Quizás necesites asignar algunos valores para métricos a algunos desconocidos y entonces aplicar el método de sustitución para resolver el nuevo sistema.

Como se puede observar ambos métodos de gauss son en si una variación uno de otro.

Ejemplo

$$\begin{cases} x - y & = & 4 \\ 2x - 2y & = & -4 \end{cases}$$

Por lo tanto la matriz argumentada será

$$\left(\begin{array}{cc|c} 1 & -1 & 4 \\ 2 & -2 & -4 \end{array}\right).$$

Mantenemos el primer renglón y sustraemos el primer renglón multiplicando por 2 el segundo renglón y obtenemos.

$$\left(\begin{array}{cc|c} 1 & -1 & 4 \\ 0 & 0 & -12 \end{array}\right).$$

Esta es la matriz triangular. El sistema asociado será

$$\begin{cases} x & - & y & = & 4 \\ & & 0 & = & -12 \end{cases}$$

Y observamos que el sistema no tiene solución

Ejercicios

El alumno:

- Resolver el sistema por eliminación gaussiana

$$\begin{cases} 2x- & 3y - z+ & 2w & +3v & = & 4 \\ 4x- & 4y - z+ & 4w & +11v & = & 4 \\ 2x- & 5y - 2z+ & 2w & -v & = & 9 \\ & 2y + z & & +4v & = & -5 \end{cases}$$

- Resuelva el siguiente sistema de ecuaciones por medio de eliminación gauss jordan

$$
\begin{aligned}
2x_1 + 4x_2 + 6x_3 &= 18 \\
4x_1 + 5x_2 + 6x_3 &= 24 \\
2x_1 + 7x_2 + 12x_3 &= 30
\end{aligned}
$$

2.2.2. Matriz inversa

La inversa de una matriz $m \times n$. Suponga que

$AB = BA = I$

Entonces B se llama inversa de A y se denota A^{-1} Entonces se tiene

$AA^{-1} = A^{-1}A = I$ donde I es la matriz identidad

Sean A y B dos matrices de

Si existe la matriz inversa de **A**, se dice que la matriz **A** es ***inversible*** o r***egular***. En caso contrario, se dice que la matriz **A** es ***singular***.

Una matriz **A** de ***orden*** **n** (n filas y n columnas) tiene inversa cuando su **rango** es **n**, es decir, cuando el ***rango*** de dicha matriz coincide con su **orden**.

Básicamente hay tres procedimientos para calcular la inversa de una matriz. Son los siguientes:

- ✓ Por el ***método de Gauss***.
- ✓ Por ***determinantes y adjuntos*** (que describiremos en la unidad de determinantes).

2.2.3. Gauss Jordan

Para calcular la ***inversa*** de una matriz cuadrada **A**, aplicando el ***método de Gauss***, construimos, en primer lugar, la matriz **(A | I)**, siendo **I** la matriz identidad del mismo orden que **A**. Después de realizar diversas operaciones sobre las filas de ésta nueva matriz, tendremos que conseguir que se transforme en la siguiente **(I | B)**. La matriz **B** será la ***inversa*** de la matriz **A**, es decir: $\mathbf{B = A^{-1}}$.

Las operaciones que podemos realizar con las filas de la citada matriz son:

a) Multiplicar o dividir una fila por un número distinto de cero.

b) Sumarle a una fila otra fila multiplicada por un número distinto de cero.

2.2.4. Método iterativo de Gauss Seidel

Introducción.

Este método es iterativo o de aproximación y es similar a las técnicas que se usan en los métodos anteriores para obtener raíces. Aquellos métodos consisten en la determinación de un valor inicial a partir del cual, mediante una técnica sistemática se obtiene una mejor aproximación a la raíz. La razón por la cual los métodos iterativos son útiles en la disminución de los errores de redondeo en sistemas, se debe a que un método de aproximación se puede continuar hasta que converja dentro de alguna tolerancia de error previamente especificada.

Las técnicas iterativas se emplean rara vez para resolver problemas de dimensiones pequeñas ya que el tiempo requerido para lograr una precisión suficiente excede al de las técnicas directas. Sin embargo, para sistemas grandes con un gran porcentaje de ceros, ésta técnica es eficiente.

Los sistemas de este tipo surgen frecuentemente en la solución numérica de problemas de valores frontera y de ecuaciones diferenciales parciales.

Algoritmo

Se debe despejar da cada ecuación la variable sobre la diagonal principal.

Dar un valor inicial a las incógnitas (generalmente se establecen ceros).

Sustituir los valores iniciales en la primera ecuación para obtener un nuevo valor para la primera incógnita.

Ese nuevo valor es usado para obtener el valor de la siguiente incógnita. Este procedimiento se repite hasta obtener los nuevos valores de todas las incógnitas despejadas.

Se evalúa la aproximación relativa de todas las incógnitas hasta que la solución converja bastante cerca de la solución real, según la tolerancia establecida para el método.

La **iteración de Gauss-Seidel** se define al tomar Q como la parte triangular inferior de A incluyendo los elementos de la diagonal

$$Q = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

Si, como en el caso anterior, definimos la matriz $R=A-Q$

$$R = \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ 0 & 0 & 0 & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

y la ecuación se puede escribir en la forma:

$$Qx^{(k)} = -Rx^{(k-1)} + b$$

Un elemento cualquiera, i, del vector $Qx^{(k)}$ vendrá dado por la ecuación:

$$\sum_{j=1}^{n} a_{ij} x_j^{(k)} = -\sum_{j=1}^{n} a_{ij} x_j^{(k-1)} + b_i$$

Si tenemos en cuenta la peculiar forma de las matrices Q y R, resulta que todos los sumandos para los que $j > i$ en la parte izquierda son nulos, mientras que en la parte derecha son nulos todos los sumandos para los que $j \leq i$. Podemos escribir entonces:

$$\sum_{j=1}^{i} a_{ij} x_j^{(k)} = -\sum_{j=i+1}^{n} a_{ij} x_j^{(k-1)} + b_i$$

$$a_{ii} x_i^{(k)} + \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} = -\sum_{j=i+1}^{n} a_{ij} x_j^{(k-1)} + b_i$$

de donde despejando $x_i^{(k)}$, obtenemos:

$$x_i^{(k)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k-1)} \right) / a_{ii}$$

Obsérvese que en el método de Gauss-Seidel los valores actualizados de x_i sustituyen de inmediato a los valores anteriores, mientras que en el método de Jacobi todas las componentes nuevas del vector se calculan antes de llevar a cabo la sustitución. Por contra, en el método de Gauss-Seidel los cálculos deben llevarse a cabo por orden, ya que el nuevo valor x_i depende de los valores actualizados de x_1, x_2, ..., x_{i-1}.

2.2.5. Teoría de sistemas de ecuaciones no lineales

Introducción

Las bases de Gröbner son un conjunto finito de polinomios de múltiples variables. En el presente trabajo se hará uso de las bases de Gröbner para el caso particular de la resolución de las ecuaciones de consistencia de un método de Runge-Kutta explícito de orden 2 y 3.

Los sistemas de ecuaciones a resolver son los siguientes:

- Orden 2 con 2 etapas:

$$\begin{cases} b_1 + b_2 = 1 \\ a_{21} b_2 = 1/2 \end{cases}$$

Sistema con dos ecuaciones y tres incógnitas.

- Orden 3 con 3 etapas:

$$\begin{cases} b_1 + b_2 + b_3 = 1 \\ b_2 a_{21} + b_3 (a_{31} + a_{32}) = 1/2 \\ a_{32} b_3 a_{21} = 1/6 \\ b_2 a_{21}^2 + b_3 (a_{31} + a_{32})^2 = 1/3 \end{cases}$$

Sistema con cuatro ecuaciones y seis incógnitas.

Ambos sistemas de ecuaciones cuentan con más incógnitas que ecuaciones, por lo que en principio, tienen infinitas soluciones. La opción más común para la resolución de estos sistemas podría ser resolverlo por sustitución y dejar el resultado en función de un parámetro en el primer caso y de dos en el segundo.

El resultado obtenido haciendo uso de las bases de Gröbner tiene que ser equivalente a cualquier otro resultado obtenido mediante cualquier otra técnica.

En general, el procedimiento de las bases de Gröbner consiste en generar un ideal que es un conjunto de ecuaciones que tiene la misma solución que el sistema original pero de más fácil solución.

Matriz de Silvester

Sean los siguientes polinomios de una variable y coeficientes en un cuerpo ™,

$$f(x) = a_0 + a_1 \cdot x + \ldots + a_k \cdot x^k$$
$$g(x) = b_0 + b_1 \cdot x + \ldots + b_l \cdot x^l$$

con a_k y b_k no nulos, interesa conocer cuándo el sistema:

$$\begin{cases} f(x) = 0 \\ g(x) = 0 \end{cases}$$

tiene solución. Una respuesta inmediata nos dice que el sistema admite solución si y sólo si el grado del máximo divisor de f(x) y g(x) es mayor o igual a uno. En efecto, si la solución x^* existe, entonces x será primo común de f(x) y g(x). Además el polinomio $x-x^*$ dividirá a cada ecuación del sistema.

Se define la matriz de Silvester para los polinomios f(x) y g(x) de la siguiente manera :

$$Syl(f,g) = \begin{pmatrix} a_0 \cdots\cdots\cdots a_k & & \\ \ddots & \ddots & \\ & a_0 \cdots\cdots\cdots a_k & \\ b_0 \cdots\cdots\cdots b_k & & \\ \ddots & \ddots & \\ & b_0 \cdots\cdots\cdots b_k & \end{pmatrix}$$

Esta matriz es cuadrada y, por tanto, se puede calcular su determinante. De hecho, el determinante de esta matriz es conocido como el Resultante de los polinomios f(x) y g(x) y se expresa como Res(f,g).

Se puede demostrar que si Syl(f,g) es invertible, entonces f(x) y g(x) son primos entre sí. Por lo tanto no tienen raíces comunes. En este caso, el sistema no admite solución.

Si Syl(f,g) es singular, se pueden obtener dos polinomios p(x) y q(x) con grad(p)<grad(g) y grad(q)<grad(f) tales que p(x)f(x)+q(x)g(x) = 0. En este caso sí hay solución.

En general, el sistema $\{f_i(x)=0\}_{i=1}^{m}$ no tiene solución si y sólo si existen los polinomios $p_i(x)$ tales que $\sum_{i=1}^{m} f_i(x)p_i(x)=1$.

Esto se puede generalizar para un sistema de ecuaciones polinómicas de varias variables de la siguiente manera:

Dado un sistema de polinomios de la forma $\{f_i(x)=0\}_{i=1}^{m}$, es necesario y suficiente que existan unos polinomios $\{p_i(x)=0\}_{i=1}^{m}$, tales que:

$$\sum_{i=1}^{m} p_i(x_i,\dots,x_n)f_i(x_1,\dots,x_n)=1$$

para que el sistema no tenga solución.

Ordenación de monomios

El cálculo de las bases de Gröbner varía sustancialmente cuando se usan diferentes ordenaciones de los monomios que integran el sistema de ecuaciones original. Consideraremos polinomios en las variables $x_1,x_2,\dots,x_n$. Para ordenar dichas variables, asumiremos que:

$$x_1 > x_2 > x_3 > \dots > x_n$$

La ordenación lexicográfica (lex) se define como $\alpha >_{lex} \beta$ si y solo si el elemento más a la izquierda no nula en el vector $\alpha-\beta$ es positivo. Se tomarán α y β como los vectores que contienen las potencias de los monomios en las variables incluidas en el sistema.

Ejemplo:

El monomio $x^5y^2z^3$ tiene un vector $\alpha=(5,2,3)$, el monomio $x^3y^4z^3$ tiene un vector β=(3,4,3), por lo que $\alpha-\beta$ = (2,-2,0). Como la primera componente del vector resta distinto de cero es mayor que cero, el primer monomio es mayor en una ordenación lexicográfica.

Reducción polinómica

La reducción polinómica es la piedra angular del algoritmo de las bases de Gröbner y es, de hecho, la parte que mayor complejidad computacional acarrea.

Se dice que un polinomio g reduce a otro polinomio h módulo algún conjunto de polinomios F (expresado como $g \rightarrow_F h$) si y sólo si LT(g) puede eliminarse por la sustracción de un múltiplo apropiado de: un polinomio *f* en F, un monomio *u* donde u=LM(g)/LM(*f*) y un escalar b=LC(g)/LC(*f*) obteniéndose h, donde:

LM(g) es el monomio principal de g: LM(g)=$x^{multigrado(g)}$, siendo multigrado(g)= máx $(\alpha \in N^n : a_\alpha \neq 0)$.

LC(*f*) es el coeficiente principal de *f*, definido como LC(*f*)= $a_{multigrado(f)}$..

LT(g) es el término principal de g, definido como LM(g)·LC(g)

Un ejemplo de todos estas definiciones es:

Sea:

$$f(x,y,z) = 2x^2y^8 - 3x^5yz^4 + xyz^3 - xy^4$$

Si se ordena el polinomio lexicográficamente, se obtiene:

$$f(x,y,z) = -3x^5yz^4 + 2x^2y^8 - xy^4 + xyz^3$$
$$multigrado(f) = (5,1,4)$$
$$LM(f) = x^5yz^4$$
$$LC(f) = -3$$
$$LT(f) = -3x^5yz^4$$

Ocurre que $g \rightarrow_F h$ si y sólo si existe un $f \in$ F, b, y u tales que h=g-buf. En cualquier otro caso g es irreducible módulo F. Esto se puede expresar de la siguiente manera:

> g es irreducible módulo G si ningún monomio principal de un elemento de F divide al monomio principal de g.

Por otra parte, si g es irreducible módulo F, entonces se puede restar de él un múltiplo de un elemento de F para eliminar su monomio principal y para obtener un nuevo monomio principal menor que el monomio principal de g. Este nuevo polinomio es equivalente a g con respecto al ideal generado por F.

Un **ideal** es un subconjunto I del anillo de polinomios K[x_1, x_2,..., x_n] si satisface lo siguiente:

1.0 es un elemento de I
2.Si f y g son dos elementos cualesquiera de I, entonces f+g es también un elemento de I
3.Si f es un elemento de I, entonces para cualquier h en K[x_1, x_2,..., x_n] h·f es también un elemento de I.

Un ejemplo de esto es:

Sea : $F = (f_1, f_2)$

Con:

$$\begin{cases} f_1 = xy^2 - x \\ f_2 = x - y^3 \end{cases}$$

Además:

$$g(x, y) = x^7 y^2 + x^3 y^2 - y + 1$$

Estos polinomios están ya ordenados lexicográficamente con x>y. Si se elige:

$$\begin{cases} f = f_1 = xy^2 - x \\ u = \dfrac{LM(g)}{LM(f)} = \dfrac{x^7 y^2}{xy^2} = x^6 \\ b = \dfrac{LC(g)}{LC(f)} = 1/1 = 1 \end{cases}$$

Entonces se obtiene el polinomio h=g-buf=$x^7 + x^3 y^2 - y + 1$, con lo que $g \to_F h$.

La definición de reducción de polinomio implica sólo al termino principal de g. Sin embargo es posible eliminar otros monomios de g para hacer la combinación lineal más pequeña. Esto conduce a la siguiente definición:

Un polinomio g es completamente reducible con respecto a F si ningún término de g es divisible por ninguno de los LT(f_i) para todos los $f_i \in$ F. Con un ejemplo:

Sea F={f_1,f_2) con:

$$\begin{cases} f_1 = xy - 1 \\ f_2 = y^2 - 1 \end{cases}$$

Si g= $x^2 y + xy^2 + y^2$, estando los polinomios ordenados lexicográficamente (con x>y), se ve claramente que g es reducible módulo F. Si se continúa con el proceso de reducción, se puede llegar fácilmente a $h = x + y + 1$ que es irreducible módulo F. En este punto se dice que h es la *forma normal* de g.

Como se puede ver, una forma de reducir un conjunto de polinomios módulo polinómico consiste en generalizar el algoritmo de la división para todos los polinomios. Esto se puede expresar de la siguiente manera:

Si f,g $\in k[x_1]$ y g≠0, entonces existe unos únicos q y r $\in k[x_1]$ tales que f=qg+r y ó r=0 o bien grad(r)<grad(g).

Para generalizar el algoritmo de la división se necesita dividir un polinomio *f* perteneciente al conjunto inicial de polinomios por un conjunto de polinomios F=$f_1,...,f_s \subset k[x_1,...x_n]$. Formalmente:

Sea F=($f_1,...f_s$) una s-tupla ordenada de polinomios en $k[x_1,...,x_n]$. Entonces si *f* es un polinomio en $k[x_1,...,x_n]$, entonces existen $q_1,...,q_x$, r $\in k[x_1,...,x_n]$ tales que f=$q_1 f_1+...+q_s f_s$+r y, o r=0 o r es un polinomio completamente reducido.

Un término que cumple un rol esencial en la teoría de las bases de Gröbner es el concepto de los S-polinomios. Es sabido que el mínimo común múltiplo (m.c.m.) de dos monomios es el producto de todas las variables, cada una elevada a la potencia máxima que aparezca en ambos monomios. Entonces, dados dos polinomios, *f* y g, si J=m.c.m.{LM(*f*), LM(f)}, se definen los S-polinomios de *f* y g como:

$$\text{S-poli}(f,g) = \frac{J}{LT(f)} \cdot f - \frac{J}{LT(g)} \cdot g$$

Como J/LT(*f*) y J/LT(g) son monomios, entonces el S-polinomio(*f*,g) es una combinación lineal con los coeficientes de *f* y g y pertenece al mismo ideal generado por *f* y g. Los S-polinomios son productos en cruz de los términos principales y se construyen para cancelar dichos términos: los términos principales de los dos componentes de S-poli(*f*,g) son iguales y se cancelan. Ejemplo:

Sea:

$$F = \{f_1, f_2\}$$

Con:

$$\begin{cases} f_1 = xy^2z - xyz \\ f_2 = x^2yz - z^2 \end{cases}$$ Ordenados lexicográficamente con x>y>z

Entonces LM(f_1)=xy^2z y LM(f_2)=x^2yz, por lo que J=m.c.m.{ xy^2z, x^2yz}=x^2y^2z. Con esto:

$$S - poli(f_1, f_2) = \frac{J}{LT(f_1)} \cdot f_1 - \frac{J}{LT(f_2)} \cdot f_2 = \frac{x^2y^2z}{xy^2z} \cdot f_1 - \frac{x^2y^2z}{x^2yz} \cdot f_2 =$$
$$= xf_1 - yf_2 = -x^2yz + yz^2$$

2.2.5.1. Método iterativo simple de punto fijo

METODO DEL PUNTO FIJO

Un **punto fijo** de una función g, es un número p tal que $g(p) = p$. El problema de encontrar las soluciones de una ecuación $f(x) = 0$ y el de encontrar los puntos fijos de una función $h(x)$ son equivalentes en el siguiente sentido: dado el problema de encontar las soluciones de una ecuación $f(x) = 0$, podemos definir una función g con un punto fijo p de muchas formas; por ejemplo, $f(x) = x - g(x)$. En forma inversa, si la función g tiene un punto fijo en p, entonces la función definida por $f(x) = x - g(x)$ posee un cero en p.

El método de punto fijo inicia con una aproximación inicial x_0 y $x_{i+1} = g(x_i)$ genera una sucesión de aproximaciones la cual converge a la solución de la ecuación $f(x) = 0$. A la función g se le conoce como función **iteradora**. Se puede demostrar que dicha sucesión $\langle x_n \rangle$ converge siempre y cuando $|g'(x)| < 1$.

Ejemplo

Usando el método de punto fijo vamos a aproximar la solución de la ecuación $x^3 + 4x^2 - 10 = 0$ dentro del intervalo $[1, 2]$.

Lo primero es buscar una función $g(x)$ adecuada

$$\begin{aligned} x^3 + 4x^2 - 10 &= 0 \\ x^2(x+4) &= 10 \\ x &= \pm\sqrt{\frac{10}{x+4}} \end{aligned}$$

Y claramente elegimos como función iteradora a

$$g(x) = \sqrt{\frac{10}{x+4}}$$

además observe que

$$|g'(x)| = \frac{\sqrt{10}}{2(x+4)^{3/2}} \leq g(2) < 1$$

para toda $x \in [1, 2]$, lo cual garantiza que la sucesión que vamos a construir va a ser convergente.

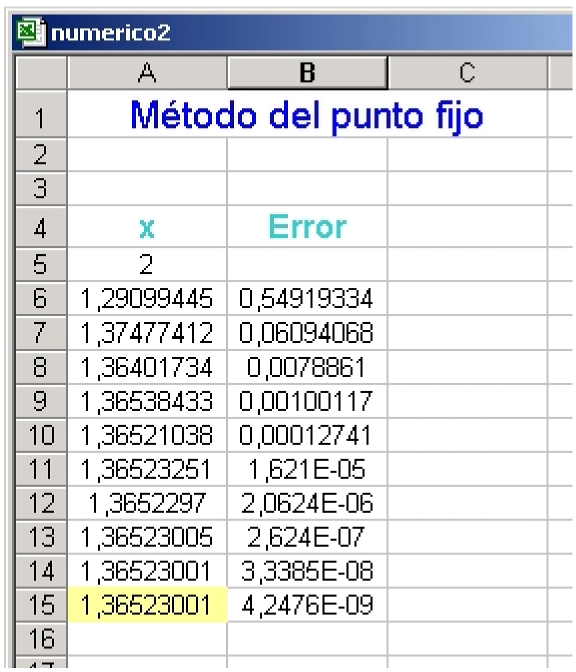
La implementación de este método en *Excel* es realmente simple, como veremos.

1. En la celda **A5** escribimos nuestra aproximación inicial, en este caso 2.
2. En la celda **A6** escribimos la fórmula que calculará las aproximaciones:

$$= \mathtt{raiz(10/(A5+4))}.$$

3. Por último arrastramos la celda **A6** para generar las restantes aproximaciones.

En la figura 10 se muestran los resultados generados por este método.



numerico2

	A	B	C
1	Método del punto fijo		
2			
3			
4	x	Error	
5	2		
6	1,29099445	0,54919334	
7	1,37477412	0,06094068	
8	1,36401734	0,0078861	
9	1,36538433	0,00100117	
10	1,36521038	0,00012741	
11	1,36523251	1,621E-05	
12	1,3652297	2,0624E-06	
13	1,36523005	2,624E-07	
14	1,36523001	3,3385E-08	
15	1,36523001	4,2476E-09	
16			

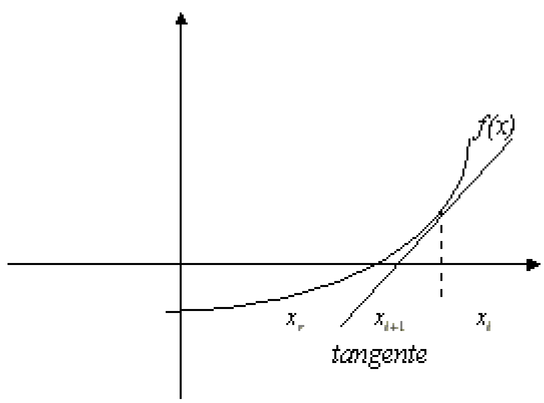
Figura 10: Iteración de punto fijo.

Una desventaja potencial del método de punto fijo es que la elección de la función iteradora $g(x)$ no siempre es fácil.

2.2.5.2. Método de Newton Raphson

Este método, el cual es un método iterativo, es uno de los más usados y efectivos. A diferencia de los métodos anteriores, el método de Newton-Raphson no trabaja sobre un intervalo sino que basa su fórmula en un proceso iterativo.

Supongamos que tenemos la aproximación x_i a la raíz x_r de $f(x)$,



Trazamos la recta tangente a la curva en el punto $(x_i, f(x_i))$; ésta cruza al eje x en un punto x_{i+1} que será nuestra siguiente aproximación a la raíz x_r.

Para calcular el punto x_{i+1}, calculamos primero la ecuación de la recta tangente. Sabemos que tiene pendiente

$$m = f'(x_i)$$

Y por lo tanto la ecuación de la recta tangente es:

$$y - f(x_i) = f'(x_i)(x - x_i)$$

Hacemos $y = 0$:

$$-f(x_i) = f'(x_i)(x - x_i)$$

Y despejamos x:

$$x = x_i - \frac{f(x_i)}{f'(x_i)}$$

Que es la fómula iterativa de Newton-Raphson para calcular la siguiente aproximación:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}, \quad \text{si} \quad f'(x_i) \neq 0$$

Note que el método de Newton-Raphson no trabaja con intervalos donde nos asegure que encontraremos la raíz, y de hecho no tenemos ninguna garantía de que nos aproximaremos a dicha raíz. Desde luego, existen ejemplos donde este método no converge a la raíz, en cuyo caso se dice que el método diverge. Sin embargo, en los casos donde si converge a la raíz lo hace con una rapidez impresionante, por lo cual es uno de los métodos preferidos por excelencia.

También observe que en el caso de que $f'(x_i) = 0$, el método no se puede aplicar. De hecho, vemos geométricamente que esto significa que la recta tangente es horizontal y por lo tanto no intersecta al eje x en ningún punto, a menos que coincida con éste, en cuyo caso x_i mismo es una raíz de $f(x)$!

Ejemplo 1

Usar el método de Newton-Raphson, para aproximar la raíz de $f(x) = e^{-x} - \ln x$, comenzando con $x_0 = 1$ y hasta que $|\in_a| < 1\%$.

Solución

En este caso, tenemos que

$$f'(x) = -e^{-x} - \frac{1}{x}$$

De aquí tenemos que:

$$x_{i+1} = x_i - \frac{e^{-x_i} - \ln(x_i)}{-e^{-x_i} - \frac{1}{x_i}} = x_i + \frac{e^{-x_i} - \ln(x_i)}{e^{-x_i} + \frac{1}{x_i}}$$

Comenzamos con $x_0 = 1$ y obtenemos:

$$x_1 = x_0 + \frac{e^{-x_0} - \ln(x_0)}{e^{-x_0} + \frac{1}{x_0}} = 1.268941421$$

En este caso, el error aproximado es,

$$|\epsilon_a| = \left|\frac{1.268941421 - 1}{1.268941421} \times 100\%\right| = 21.19\%$$

Continuamos el proceso hasta reducir el error aproximado hasta donde se pidió.

Resumimos los resultados en la siguiente tabla:

Aprox. a la raíz	Error aprox.
1	
1.268941421	21.19%
1.309108403	3.06%
1.309799389	0.052%

De lo cual concluímos que la aproximación obtenida es:

$$x_3 = 1.309799389$$

2.2.5.3. Método división sintética

Sea $f(x)$ cualquier polinomio y sea $g(x)$ un polinomio no nulo. Existen 2 únicos polinomios $q(x)$ y $r(x)$ que satisfacen las condiciones siguientes:

1.-) $f(x) = g(x)q(x) + r(x)$

2.-) grado de $r(x) <$ grado de $g(x)$

Teorema 1 (Teorema del residuo).

"Si el polinomio $f(x)$ se divide entre $x - r$ siendo r una constante independiente de *x*, el residuo es igual a $f(r)$ ".

Teorema 2 (Teorema del factor).

"r es una raíz de la ecuación entera $f(x)=0$ *,si y solo si* $x-r$ es un factor del polinomio $f(x)$ ".

El teorema del residuo nos permite obtener el valor de $f(x)$ para valores de *x* sin hacer la sustitución directa, pero esto requiere la división de un polinomio entre un binomio. El método para efectuar rápidamente esta división se conoce como **división sintética**. Así, al realizar la división de nuestro polinomio $f(x)$ por el binomio $x-r$, habremos encontrado una raíz de la ecuación si el residuo es cero.

Reglas para la división sintética.

Para dividir un polinomio $f(x)$ entre $x-r$, se procede como sigue:

1.- En la primera línea se escriben los coeficientes $a_0, a_1, \dots, a_n$ del dividendo $f(x)$ y el número *r* separado y a la derecha. Si alguna potencia de *x* no aparece en $f(x)$ su coeficiente se escribe como cero.

2.- Se escribe el coeficiente principal a_0 como primer término de la tercera linea y se multiplica por *r*, escribiendo a_0r en la segunda linea debajo de a_1 .

3.- Se suma a_1 con el producto a_0r y se escribe la suma a_0r+a_1 en la tercera linea.

4.- Se multiplica a_0r+a_1 por *r*, se escribe le producto en la segunda linea debajo de a_2 y se suma con a_2 escribiéndose la suma en la tercera linea.

5.- Se continua de esta manera hasta que se usa como sumando a_n escribiéndose la suma en la tercera linea. El último número de la tercera linea es el residuo; los números anteriores son los coeficientes del cociente correspondientes a potencias descendentes de *x*.

Bibliografía

1. L. V. Atkinson y P. J. Harley. *An Introduction to Numerical Methods with Pascal*. Adison-Wesley, 1983.
2. R. L. Burden y J. D. Faires. *Análisis Numérico*. Grupo Editorial Iberoamérica, 1985.
3. B. W. Char, K. O. Geddes, G. H. Gonnet, B. L. Leong, M. B. Monagan, y S. Watt. *Maple V Language Reference Manual*. Springer-Verlag, 1991.
4. B. W. Char, K. O. Geddes, G. H. Gonnet, B. L. Leong, M. B. Monagan, y S. Watt. *First Leaves: A Tutorial Introduction to Maple V*. Springer-Verlag, 1992.
5. Francis Sheid Rosa Elena Di Costanzo. *Métodos Numéricos*. McGraw Hill, 1989.
6. Stanley I. Grossman. *Algebra Lineal*. Grupo Editorial Iberoamericana, 1987.
7. Thomas Richard McCalla. *Introduction to Numerical Methods and FORTRAN Programming*. Wiley, 1967.
8. Antonio Nieves y Federico C. Domínguez. *Métodos Numéricos aplicados a la ingeniería*. CECSA, 1995.
9. Ben Noble y James W. Daniel. *Algebra Lineal*. Prentice Hall, tercera edition, 1989.
10. W. Allen Smith. *Análisis Numérico*. Prentice Hall, 1988.

Actividades Complementarias

1..- Un paracaidista, con una masa de 68.1 kgs salta de un globo aerostático fijo. Con la ayuda de la ecuación (9), calcule la velocidad antes de abrir el paracaídas, coeficiente de resistencia = 12 kg/seg.

Datos:

m = 68.1
c = 12.5
g = 9.8 m/s

$$v(t) = gm/c\ (\ 1 - e^{-(c/m)t}\ )$$

2.- Resolver el ejemplo anterior mediante una solución numérica para calcular la velocidad. Emplear un tamaño del paso de 2 segundos.

Datos:

m = 68.1 kg
c = 12.5 kg/s
g = 9.8 m/s