

CAPÍTULO 3

Matemática Numérica, 2da Edición

Manuel Álvarez, Alfredo Guerra, Rogelio Lau

SISTEMAS DE ECUACIONES LINEALES Y MATRICES

Objetivos

Al finalizar el estudio y la ejercitación de este capítulo, el lector debe ser capaz de:

- Describir las características de los métodos directos y los métodos iterativos para la resolución de sistemas de ecuaciones lineales.
- Describir los conceptos de norma de un vector y de norma de una matriz y explicar las propiedades (axiomas) de las mismas.
- Describir el algoritmo de Gauss para resolver sistemas de ecuaciones lineales.
- Describir las diferentes estrategias de pivote que se utilizan en el método de Gauss y compararlas entre sí.
- Utilizar el método de Gauss manualmente o mediante un programa personal en algún lenguaje computacional que conozca.
- Analizar el volumen de operaciones que requerirá resolver un sistema de ecuaciones por el método de Gauss.
- Describir el algoritmo de Gauss especializado en sistemas tridiagonales y computar el orden de la cantidad de operaciones que se requerirá para su aplicación.
- Utilizar el método de Gauss especializado en sistemas tridiagonales, ya sea manualmente o mediante un programa computacional en algún lenguaje que domine.
- Describir el algoritmo para calcular determinantes basado en el proceso directo del método de Gauss y utilizarlo para calcular determinantes, ya sea manualmente o mediante un programa que confeccione.
- Describir el algoritmo para invertir matrices basado en el método de Gauss y utilizarlo, ya sea manualmente o mediante un programa que confeccione.
- Describir el concepto de sistema lineal mal condicionado.
- Establecer el concepto de número de condición de una matriz y explicar por qué éste mide el mal condicionamiento de una matriz.
- Calcular el número de condición de la matriz de un sistema lineal y decidir si se trata o no de un sistema mal condicionado.
- Describir los métodos de Jacobi y de Seidel para resolver sistemas lineales en forma iterativa y las condiciones bajo las cuales los mismos convergen.
- Resolver sistemas de ecuaciones lineales mediante los algoritmos de Jacobi y de Seidel manualmente o mediante programas computacionales escritos al efecto.
- Realizar un análisis comparativo entre los algoritmos iterativos de Jacobi y de Seidel.
- Decidir, basándose en la cantidad de operaciones necesarias, cuál de los algoritmos estudiados es preferible para resolver un sistema de ecuaciones lineales.
- Describir los conceptos de valores y vectores propios de una matriz.
- Describir el concepto de polinomio característico y las dificultades de su obtención en problemas reales.
- Graficar grosso modo el polinomio característico de una matriz trazando una cantidad grande de puntos de la gráfica mediante un programa capaz de calcular determinantes eficientemente.
- Elaborar programas computacionales en algún lenguaje que usted domine, que permitan hallar valores propios resolviendo en forma numérica la ecuación característica de la matriz.

- Describir el método de la potencia para hallar el valor propio de mayor valor absoluto de una matriz y utilizarlo manualmente o mediante un programa computacional confeccionado al efecto.

3.1 Introducción

En la modelación de problemas reales surgen con frecuencia sistemas de ecuaciones lineales de orden apreciable. Los dos ejemplos que siguen muestran solamente dos situaciones concretas y pueden dar una idea de cuan grande puede ser el tamaño de estos sistemas.

Ejemplo 1

En la resolución de circuitos eléctricos, uno de los métodos empleados consiste en definir las llamadas corrientes de malla y plantear la ley de Kirchhoff para voltajes en las diferentes mallas del circuito. En la figura 1 se muestra un circuito con cuatro mallas donde se han definido las corrientes i_1 , i_2 , i_3 e i_4 . Al plantear la ley de Kirchhoff para cada malla se obtienen las ecuaciones:

$$\begin{aligned} R_1 i_1 + R_5 (i_1 - i_4) &= E_1 - E_2 \\ R_2 i_2 + R_3 (i_2 - i_3) &= E_2 \\ R_3 (i_3 - i_2) + R_6 i_3 + R_4 (i_3 - i_4) &= E_3 \\ R_5 (i_4 - i_1) + R_4 (i_4 - i_3) &= 0 \end{aligned}$$

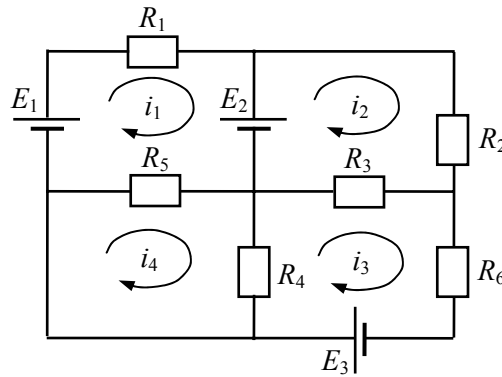


Figura 1

Si se conocen los valores de las resistencia y de las fuentes de voltaje, los valores de intensidad de corriente pueden ser obtenidos resolviendo el sistema anterior, que es un sistema de cuatro ecuaciones lineales con cuatro incógnitas i_1 , i_2 , i_3 e i_4 . Después de ordenado, el sistema toma la forma:

$$\begin{aligned} (R_1 + R_5)i_1 & & & -R_5i_4 &= E_1 - E_2 \\ & (R_2 + R_3)i_2 & -R_3i_3 & &= E_2 \\ & -R_3i_2 + (R_3 + R_4 + R_6)i_3 & & -R_4i_4 &= E_3 \\ -R_5i_1 & & -R_4i_3 + (R_4 + R_5)i_4 & &= 0 \end{aligned}$$

Es fácil imaginar que en un circuito real pueden aparecer decenas de mallas que dan lugar a sistemas con esa misma cantidad de ecuaciones y de incógnitas, los cuales no pueden ser resueltos por vía manual. ■

La importancia de los sistemas lineales no viene dada solamente porque ellos aparezcan directamente en la modelación matemática de infinitud de problemas del mundo real; además de esto, la solución de muchos problemas matemáticos conduce a resolver sistemas lineales; así sucede con el ajuste de curvas, en varias técnicas de interpolación, en la solución de ecuaciones diferenciales parciales y en muchos otros campos. El ejemplo que sigue muestra, en forma elemental, los sistemas lineales que surgen cuando se aplica la técnica de diferencias finitas en la solución de una ecuación diferencial parcial de tipo elíptico.

Ejemplo 2

Una lámina cuadrada de metal de un metro de lado es sometida en sus bordes a temperaturas fijas (ver figura 2) hasta que la temperatura alcanza valores estables en cada punto de la lámina. Se demuestra que si $u = u(x, y)$ es la temperatura en el punto (x, y) entonces, para los puntos de la lámina rige la ecuación diferencial:

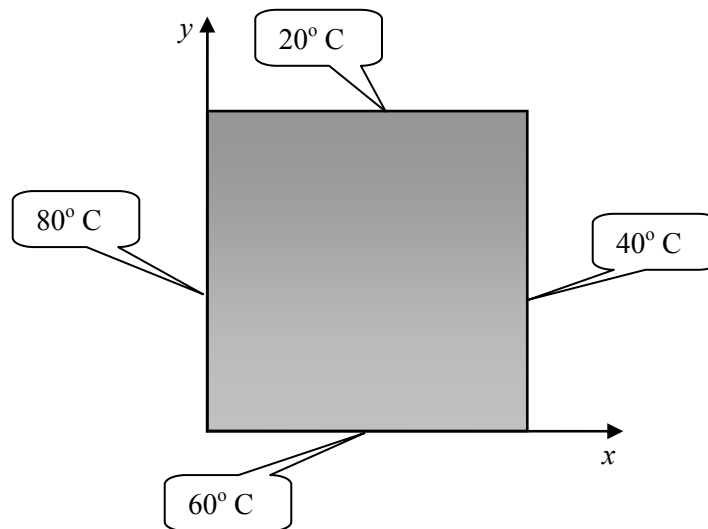


Figura 2

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

con las condiciones de frontera:

$$\begin{aligned} u(0, y) &= 80 \quad \text{para } 0 < y < 1 \\ u(1, y) &= 40 \quad \text{para } 0 < y < 1 \\ u(x, 0) &= 60 \quad \text{para } 0 < x < 1 \\ u(x, 1) &= 20 \quad \text{para } 0 < x < 1 \end{aligned}$$

Si se define sobre la lámina una red de puntos como la que muestra la figura 3 y se llama u_{ij} ($i = 0, 1, 2, 3, 4, 5$; $j = 0, 1, 2, 3, 4, 5$) a la temperatura en los puntos de la red, entonces la temperatura en cada uno de los puntos que no están en la frontera está relacionada aproximadamente con la temperatura de sus puntos vecinos mediante la ecuación lineal:

$$4u_{ij} - u_{i-1,j} - u_{i+1,j} - u_{i,j-1} - u_{i,j+1} = 0 \quad (1)$$

donde: $i = 1, 2, 3, 4$; $j = 1, 2, 3, 4$

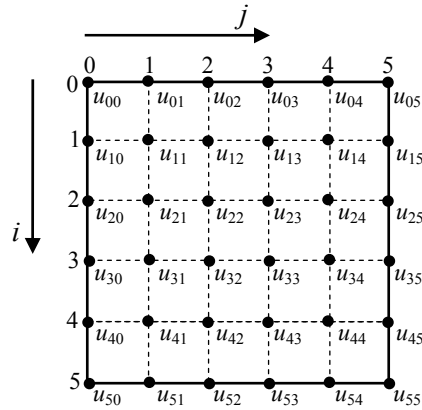


Figura 3

Nótese que la expresión (1) constituye realmente un sistema de 16 ecuaciones lineales con 16 incógnitas que corresponden a las temperaturas en los 16 puntos interiores, ya que las temperaturas de los puntos que están en el borde son conocidas por las condiciones de frontera. Resolviendo este sistema de 16 ecuaciones se obtienen las temperaturas deseadas.

La aproximación se hace mejor mientras más tupida sea la red de puntos, sin embargo, la cantidad de ecuaciones crece considerablemente. Por ejemplo, si en lugar de formar una red con distancias entre puntos de $h = 0,2$, se hubiera tomado $h = 0,01$, la exactitud mejoraría y la temperatura sería conocida casi en cualquier punto de la lámina; sin embargo, el sistema lineal que habría que resolver en ese caso sería de $(99)(99) = 9801$ ecuaciones con 9801 incógnitas.

Problemas a resolver

Relacionados con los sistemas lineales existen varios problemas numéricos importantes. El primero es la solución de grandes sistemas, formados por cientos y hasta decenas de miles de ecuaciones. Cuando el tamaño del problema crece en ese grado es fundamental analizar la eficiencia del método empleado, de otro modo la cantidad de operaciones aritméticas puede crecer a tal punto que ni aun en la máquina más eficiente sea factible su solución en un tiempo razonable. Por ejemplo, si se decide resolver un sistema lineal de 30 ecuaciones con 30 incógnitas por el método de Cramer, será necesario calcular 31 determinantes de orden 30; si para calcular cada uno de estos determinantes se utiliza el algoritmo recursivo conocido como desarrollo por menores, el cual transforma el determinante de orden 30 en una combinación lineal de 30 determinantes de orden 29, cada determinante de orden 29 en una combinación lineal de 29 determinantes de orden 28, etcétera, es obvio que cada determinante de orden 30 requerirá 30! operaciones de multiplicación. Como son 31 determinantes, se necesitarán $31 \cdot 30! = 31!$

multiplicaciones (sin contar las sumas). ¿Puede alguna máquina efectuar un algoritmo que requiera $31!$ operaciones? Note el lector que $31!$ es aproximadamente $8,22 \cdot 10^{33}$. Una máquina que efectúe 10^9 operaciones por segundo y que trabaje los 31 536 000 segundos que tiene un año será capaz de realizar unas $3,15 \cdot 10^{16}$ operaciones en el año. Si se contara con 10 000 millones de computadoras de este tipo (más de una computadora por cada habitante del planeta) y se pudiera tenerlas a todas trabajando en el mismo problema simultáneamente, en este descomunal sistema paralelo se podrían realizar $3,15 \cdot 10^{26}$ operaciones en un año. Se necesitaría entonces de

$$\frac{8,22 \cdot 10^{33}}{3,15 \cdot 10^{26}} = 2,61 \cdot 10^7 \text{ años}$$

o sea, unos 26 millones de años; claro, es de suponer que mucho antes, el incauto analista numérico de esta historia se habría percatado de que, con un método más razonable una computadora personal requiere menos de un segundo para resolver este problema.

Además de la solución de grandes sistemas lineales con algoritmos eficientes, en este capítulo se abordará el problema del cálculo de determinantes numéricos, operación que, aunque no es aconsejable como forma de resolver sistemas lineales, en ocasiones se requiere realizar con otros fines.

Otro problema menos frecuente pero que suele presentarse es la inversión de matrices. Aunque en los cursos de Álgebra Lineal se calculan las matrices inversas mediante el algoritmo usual de dividir la traspuesta de la matriz de cofactores entre el determinante de la matriz, este procedimiento, salvo para matrices de orden 3 ó 4, es extraordinariamente ineficiente. La inversión de matrices numéricas será abordada también en este capítulo.

El problema de hallar los valores y vectores propios de una matriz es uno de los más complejos de la matemática numérica y en ocasiones se necesita resolver, al menos parcialmente. Aquí solo se hará una introducción elemental a este tema y en la bibliografía recomendada se encuentran algunas referencias útiles.

Relacionados con los sistemas lineales se encuentran frecuentemente problemas inestables (que en este ámbito se suelen llamar, *mal condicionados*). Aunque el carácter elemental de este texto no permite un tratamiento profundo de este asunto, en la sección 3.3 se abordará esta cuestión.

Métodos directos y métodos iterativos

La mayor parte de los métodos utilizados para la solución de sistemas lineales están concentrados en dos grandes tipos: métodos directos o “exactos” y métodos iterativos o de aproximaciones sucesivas. La primera clase incluye a aquellos métodos en que, si se pudieran evitar todos los errores por redondeo, se obtendría la respuesta exacta en un número finito de pasos; muchos de los métodos elementales para resolver sistemas (Cramer, sustitución, suma y resta) y para calcular determinantes e invertir matrices pertenecen a esta categoría; aquí se estudiará el método de Gauss, que es también un método directo aplicable a varios de los problemas que se analizarán. Por otra parte, los métodos iterativos son todos aquellos que producen una sucesión infinita de respuestas aproximadas, sucesión que, bajo ciertas condiciones, converge hacia la solución exacta del problema; este tipo de método estará aquí representado por los métodos de Jacobi y de Seidel para resolver ciertos sistemas lineales y el método de la potencia para el cálculo de valores y

vectores propios. Más adelante, cuando se haya apreciado las características de cada uno de los métodos estudiados, se podrá valorar en que situaciones es cada uno más conveniente.

Además de los métodos directos y los iterativos, se pueden mencionar los llamados métodos de Montecarlo, en los que, con la utilización de procesos aleatorios, se puede llegar a una solución aproximada, en general con muy poca exactitud, de sistemas lineales. La utilización de la teoría de probabilidades que ello requiere impide su tratamiento en un libro como este.

Normas matriciales y vectoriales

Se supone que el lector tiene conocimientos acerca de las operaciones con matrices y vectores al nivel que se estudia en los cursos de pregrado de Álgebra Lineal. No obstante, como en muchos de estos cursos se obvia el tema de las normas matriciales, aquí se han incluido los elementos que más adelante se requieren.

En general, si se tiene un espacio vectorial V , una norma en V es cualquier función $\| \cdot \|$ que a cada vector v de V le haga corresponder un número real denotado como $\|v\|$, y que posea las siguientes propiedades, que se llaman axiomas de norma:

1. Para todo v de V , $\|v\| \geq 0$
2. Para todo v de V y todo k real, $\|kv\| = |k| \cdot \|v\|$
3. Para todos u y v de V , $\|u + v\| \leq \|u\| + \|v\|$
4. Si $\|v\| = 0$ entonces v es el vector nulo de V

En los espacios de tipo R^n se definen normas de diversas formas. Probablemente el lector esté familiarizado con la norma cuadrática que se define como la raíz cuadrada de la suma de los cuadrados de las componentes del vector, que en algunos temas es muy importante pero que aquí no resulta la más conveniente. En todo este capítulo se utilizará para los vectores de R^n la siguiente norma.

Definición 1

Si $\mathbf{x}$ es un vector de R^n : $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, se define la norma de $\mathbf{x}$ como: $\|\mathbf{x}\| = \max_i |x_i|$.

■

Se puede probar sin dificultad que esta definición satisface los cuatro axiomas de norma.

En el caso de las matrices cuadradas, ellas, como se sabe, constituyen un espacio vectorial y por tanto, su norma debe satisfacer los cuatro axiomas generales. No obstante, como en el espacio de las matrices cuadradas están definidas también las operaciones de producto entre matrices y producto de una matriz por un vector, se requiere de axiomas especiales que establezcan las propiedades que debe poseer la norma matricial respecto a estas operaciones. Así, para el espacio vectorial M_n de las matrices cuadradas de orden n , cualquier norma que se utilice deberá cumplir las siguientes 6 propiedades (axiomas de norma matricial):

1. Para toda matriz $\mathbf{A}$ de M_n $\|\mathbf{A}\| \geq 0$
2. Para todo $\mathbf{A}$ de M_n y todo k real, $\|k\mathbf{A}\| = |k| \cdot \|\mathbf{A}\|$
3. Para todos $\mathbf{A}$ y $\mathbf{B}$ de M_n , $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$
4. Si $\|\mathbf{A}\| = 0$ entonces $\mathbf{A}$ es la matriz nula de M_n
5. Para todos $\mathbf{A}$ y $\mathbf{B}$ de M_n , $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$
6. Si $\|\cdot\|_v$ es la norma de un vector de R^n , entonces, para toda $\mathbf{A}$ de M_n y todo $\mathbf{x}$ de R^n ,

$$\|\mathbf{Ax}\|_v \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\|_v$$

Nótese que el axioma 6 significa que el tipo de norma utilizado para las matrices debe ser compatible con la norma utilizada en R^n . La siguiente definición de norma matricial satisface estos 6 axiomas si se toma la norma en R^n de acuerdo con la definición 1.

Definición 2

Si $\mathbf{A}$ es una matriz cuadrada de orden n , $\mathbf{A} = [a_{ij}]$, se define la norma de $\mathbf{A}$ como:

$$\|\mathbf{A}\| = \max_i \sum_{j=1}^n |a_{ij}|$$

Aunque se pueden utilizar otras combinaciones de normas matriciales y vectoriales, la que aquí se ha seleccionado es la que más se emplea en el tema que se tratará, debido a la simplicidad de los cálculos que se requieren.

Ejemplo 3

Dadas las matrices $\mathbf{A} = \begin{bmatrix} 3 & -1 & 2 \\ 0 & 5 & 1 \\ 8 & -3 & 2 \end{bmatrix}$ y $\mathbf{B} = \begin{bmatrix} 0 & 3 & 7 \\ 1 & -1 & 3 \\ 2 & 0 & -2 \end{bmatrix}$ y los vectores $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$ y $\mathbf{y} = \begin{bmatrix} 0 \\ 2 \\ -5 \end{bmatrix}$,

calcule: a) $\|\mathbf{A}\|$, b) $\|\mathbf{B}\|$, c) $\|\mathbf{A} + \mathbf{B}\|$, d) $\|\mathbf{AB}\|$, e) $\|\mathbf{x}\|$, f) $\|\mathbf{y}\|$, g) $\|\mathbf{x} + \mathbf{y}\|$, h) $\|\mathbf{Ax}\|$ y verifique que se satisfacen los axiomas correspondientes.

Solución:

a) $\|\mathbf{A}\| = \max\{6, 6, 13\} = 13$

b) $\|\mathbf{B}\| = \max\{10, 5, 4\} = 10$

c) $\mathbf{A} + \mathbf{B} = \begin{bmatrix} 3 & 2 & 9 \\ 1 & 4 & 4 \\ 10 & -3 & 0 \end{bmatrix}$, así que $\|\mathbf{A} + \mathbf{B}\| = \max\{14, 9, 13\} = 14$. Se cumple el axioma 3 de la

norma matricial, es decir: $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\| = 23$

d) $\mathbf{AB} = \begin{bmatrix} 3 & 10 & 14 \\ 7 & -5 & 13 \\ 1 & 27 & 43 \end{bmatrix}$, por tanto, $\|\mathbf{AB}\| = \max\{27, 25, 71\} = 71$. Se satisface el axioma 5 de la

norma matricial, esto es, $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\| = 130$.

e) $\|\mathbf{x}\| = 4$

f) $\|\mathbf{y}\| = 5$

g) $\mathbf{x} + \mathbf{y} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$, luego, $\|\mathbf{x} + \mathbf{y}\| = 2$. Se cumple que $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| = 9$, que es el axioma 3 de

los vectores.

h) $\mathbf{Ax} = \begin{bmatrix} 15 \\ -1 \\ 27 \end{bmatrix}$, por lo tanto, $\|\mathbf{Ax}\| = 27$. Se satisface el axioma 6 de la norma matricial, esto es:

$$\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\| = (13)(4) = 52.$$

3.2 El método de Gauss

Aunque el método de Gauss puede ser aplicado a sistemas lineales de cualquier orden, aquí solo interesan los sistemas cuadrados, es decir, con igual número de ecuaciones e incógnitas. Se supone además que se trata de sistemas determinados, es decir, con solución única. Sea entonces el sistema a resolver:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \quad (1)$$

donde todos los coeficientes a_{ij} son números reales conocidos, al igual que los términos independientes. Este sistema, como se sabe, se puede representar en forma matricial de la forma:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (2)$$

Tomando el convenio de representar la matriz de los coeficientes por $\mathbf{A}$, el vector de incógnitas por $\mathbf{x}$ y el vector de términos independientes como $\mathbf{b}$, la ecuación matricial (2) se reduce a:

$$\mathbf{Ax} = \mathbf{b} \quad (3)$$

$$\text{donde } \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Como se supone que el sistema es determinado, eso significa que $\mathbf{A}$ es no singular, es decir, que su determinante no es cero.

El algoritmo de Gauss consta de dos etapas bien diferenciadas que se llaman *proceso directo* y *proceso inverso*.

Proceso directo

El proceso directo consiste en realizar sobre el sistema de ecuaciones transformaciones elementales de dos tipos:

- I. Intercambiar dos ecuaciones del sistema.
- II. Sumar a una ecuación, miembro a miembro, otra ecuación (pivote) del sistema multiplicada por un número real cualquiera.

Es claro que la transformación elemental tipo I no cambia la solución del sistema, aunque si afecta el determinante de $\mathbf{A}$, ya que equivale a la permutación de dos de sus filas, lo cual cambia el signo del determinante. La transformación elemental de tipo II no altera ni la solución del sistema ni el valor del determinante. Estas transformaciones elementales se realizan con el objetivo de transformar el sistema a la forma triangular:

$$\begin{aligned} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n &= d_1 \\ c_{22}x_2 + \cdots + c_{2n}x_n &= d_2 \\ &\vdots \\ c_{nn}x_n &= d_n \end{aligned} \tag{4}$$

Cuando se trabaja en forma manual (y con más razón cuando se realiza un algoritmo computacional) estas transformaciones elementales se efectúan sobre los coeficientes de las ecuaciones, los cuales se escriben en forma de una matriz de orden n por $n+1$ que agrupa los elementos de $\mathbf{A}$ y los de $\mathbf{b}$ y que suele llamarse *matriz ampliada* del sistema:

$$\mathbf{A}|\mathbf{b} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix} \tag{5}$$

la cual, efectuado el proceso directo, toma la forma escalonada:

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} & d_1 \\ 0 & c_{22} & \cdots & c_{2n} & d_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & c_{nn} & d_n \end{bmatrix} \quad (6)$$

Para pasar de la matriz inicial a su forma escalonada, primero se efectúan $n - 1$ transformaciones elementales de tipo II, utilizando la fila 1 como pivote y afectando a las filas debajo de ella, de modo que se anulen todos los elementos de la primera columna que están debajo de la diagonal, después $n - 2$ transformaciones tipo II con la fila 2 como pivote, para modificar a las filas debajo de ella de manera que se anulen todos los elementos de la segunda columna debajo de la diagonal; el proceso continua hasta llegar a la fila $n - 1$ como pivote para eliminar al coeficiente de la columna $n - 1$ debajo de la diagonal.

En el paso número i del proceso directo, en el cual la fila i -sima actúa como pivote, la fila k ($k > i$) se cambia de acuerdo con la fórmula:

$$\text{fila } k := \text{fila } k - \frac{a'_{ki}}{a'_{ii}} (\text{fila } i) \quad k = i + 1, i + 2, \dots, n \quad (7)$$

donde los coeficientes se han afectado con un apóstrofe para indicar que no son necesariamente los coeficientes originales. Nótese que puede suceder que el coeficiente a'_{ii} de la fila pivote tome valor cero o muy pequeño. En ese caso se procede a permutar la fila i por alguna de las inferiores, de acuerdo con algunas estrategias disponibles y que se analizarán más adelante.

Proceso inverso

Una vez obtenido el sistema triangular (4) o su forma compacta (6) se puede calcular x_n de la última ecuación, después x_{n-1} a partir de la penúltima ecuación y así sucesivamente hasta encontrar x_1 de la primera ecuación. La forma de obtener la solución, de abajo hacia arriba, le da nombre a esta parte del algoritmo. Resulta:

$$\begin{aligned} x_n &= \frac{1}{c_{nn}} d_n \\ x_{n-1} &= \frac{1}{c_{n-1,n-1}} (d_{n-1} - c_{n-1,n} x_n) \\ &\vdots \\ x_1 &= \frac{1}{c_{11}} (d_1 - c_{12} x_2 - c_{13} x_3 - \cdots - c_{1n} x_n) \end{aligned}$$

Estrategias de pivote

La estrategia de pivote es el criterio que se utiliza para seleccionar la fila que se ha de utilizar como pivote en cada paso del proceso directo. Se utilizan en la práctica tres estrategias llamadas *elemental*, *parcial* y *total*. En la estrategia *elemental*, cuando se va a realizar el paso número i del proceso directo, se utiliza la fila i -sima como pivote a menos que el elemento de la diagonal a'_{ii} sea cero o menor que un número ε muy pequeño; en ese caso, se procede a analizar las filas por debajo de la i -sima, se escoge la primera que contenga el elemento de la i -sima posición distinto de cero (o modularmente mayor que ε) y esta fila se intercambia con la fila i ; si todas las filas tuvieran nulo su elemento i , entonces la matriz original sería singular, contrario a lo que se

ha supuesto. En la estrategia parcial de pivote, al realizar el paso i del proceso directo, se analizan todas las filas desde la i -sima hasta la última para seleccionar aquella que posea el elemento en posición i con el mayor valor absoluto (no todos pueden ser cero) y esta fila se intercambia con la i -sima fila. En la estrategia total, en el paso número i del proceso directo se busca el elemento de mayor valor absoluto de la submatriz que aun no se encuentra escalonada y se intercambian dos filas y dos columnas de manera que ese elemento pase a ocupar la posición de a'_{ii} .

La estrategia elemental es muy fácil de implementar pero se corre el riesgo de que el elemento a'_{ii} que forma el denominador de la expresión (7) sea muy pequeño. En ese caso la fila i -sima puede ser multiplicada por un número muy grande y esto produce una ampliación de los errores absolutos que contiene esta fila debido a los redondeos y su propagación hasta ese momento. En sistemas un poco grandes esta acumulación y amplificación de errores puede llevar a resultados desastrosos.

La estrategia parcial es también fácil de implementar y garantiza que el factor que se utiliza para multiplicar a la fila pivote sea siempre menor que 1, con lo cual los errores acumulados en la fila i -sima, lejos de ampliarse, se disminuyen.

La estrategia total da aun mejores resultados en cuanto a la propagación de los errores de redondeo pero es mucho más difícil de implementar debido al intercambio de columnas; nótese que intercambiar dos filas no trae afectaciones a la solución del sistema, pero cuando se intercambian dos columnas esto significa que las variables correspondientes han cambiado su posición y se necesita mantener un registro de todos los intercambios de este tipo que se hayan efectuado a la hora de realizar el proceso inverso. Por esta razón, la estrategia total no se incluirá en el algoritmo en pseudo código para el método de Gauss.

Ejemplo 1

Resuelva el siguiente sistema lineal utilizando a) estrategia elemental de pivote y b) estrategia parcial de pivote. En ambos casos conserve solamente tres cifras exactas en los resultados.

$$\begin{aligned} 3x_1 + 3x_2 - x_3 - x_4 &= 4 \\ x_1 - x_2 + 3x_3 + x_4 &= 2 \\ -2x_1 - x_2 + x_3 + 3x_4 &= 4 \\ 2x_1 - 2x_2 - x_3 - 2x_4 &= -1 \end{aligned}$$

Solución:

Formando la matriz ampliada del sistema:

$$\left[\begin{array}{cccc|c} 3 & 3 & -1 & -1 & 4 \\ 1 & -1 & 3 & 1 & 2 \\ -2 & -1 & 1 & 3 & 4 \\ 2 & -2 & -1 & -2 & -1 \end{array} \right]$$

a) Utilizando estrategia elemental de pivote

Primer paso del proceso directo:

Como el elemento 1 de la fila 1 es $3.00 \neq 0$, se conserva esta fila como pivote.

fila 2 := fila 2 - (0,333) fila 1

fila 3 := fila 3 + (0,667) fila 1

fila 4 := fila 4 - (0,667) fila 1:

$$\begin{bmatrix} 3.00 & 3.00 & -1.00 & -1.00 & 4.00 \\ & -2.00 & 3.33 & 1.33 & 0.668 \\ & 1.00 & 0.333 & 2.33 & 6.67 \\ & -4.00 & -0.333 & -1.33 & -3.67 \end{bmatrix}$$

Segundo paso del proceso directo:

Como el segundo elemento de la fila 2 es $-2.00 \neq 0$, se conserva esta fila pivote.

fila 3 := fila 3 + (0,500) fila 2

fila 4 := fila 4 - (2,00) fila 2:

$$\begin{bmatrix} 3.00 & 3.00 & -1.00 & -1.00 & 4.00 \\ & -2.00 & 3.33 & 1.33 & 0.668 \\ & & 2.00 & 3.00 & 7.00 \\ & & -6.99 & -3.99 & -5.01 \end{bmatrix}$$

Tercer paso del proceso directo:

Como el tercer elemento de la fila 3 es $2.00 \neq 0$, se conserva esta fila pivote.

fila 4 := fila 4 + (3,50) fila 3:

$$\begin{bmatrix} 3.00 & 3.00 & -1.00 & -1.00 & 4.00 \\ & -2.00 & 3.33 & 1.33 & 0.668 \\ & & 2.00 & 3.00 & 7.00 \\ & & & 6.51 & 19.5 \end{bmatrix}$$

Resulta entonces el sistema escalonado:

$$\begin{aligned} 3.00x_1 + 3.00x_2 - 1.00x_3 - 1.00x_4 &= 4.00 \\ -2.00x_2 + 3.33x_3 + 1.33x_4 &= 0.668 \\ 2.00x_3 + 3.00x_4 &= 7.00 \\ 6.51x_4 &= 19.5 \end{aligned}$$

Proceso inverso: $x_4 = \frac{19.5}{6.51} = 3.00$

$$x_3 = \frac{7.00 - (3.00)(3.00)}{2.00} = -1.00$$

$$x_2 = \frac{0.668 - (3.33)(-1.00) - (1.33)(3.00)}{-2.00} = -0.004$$

$$x_1 = \frac{4.00 - (3.00)(-0.004) + (1.00)(-1.00) + (1.00)(3.00)}{3.00} = 2.00$$

b) Utilizando estrategia parcial de pivote

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

Primer paso del proceso directo:

Como el elemento de mayor valor absoluto esta en la fila 1, se conserva esta fila pivote.

fila 2 := fila 2 - (0,333) fila 1

fila 3 := fila 3 + (0,667) fila 1

fila 4 := fila 4 - (0,667) fila 1:

$$\begin{bmatrix} 3.00 & 3.00 & -1.00 & -1.00 & 4.00 \\ & -2.00 & 3.33 & 1.33 & 0.668 \\ & 1.00 & 0.333 & 2.33 & 6.67 \\ & -4.00 & -0.333 & -1.33 & -3.67 \end{bmatrix}$$

Segundo paso del proceso directo:

Como el elemento de mayor valor absoluto esta en la fila 4, se intercambian las filas 2 y 4:

$$\begin{bmatrix} 3.00 & 3.00 & -1.00 & -1.00 & 4.00 \\ & -4.00 & -0.333 & -1.33 & -3.67 \\ & 1.00 & 0.333 & 2.33 & 6.67 \\ & -2.00 & 3.33 & 1.33 & 0.668 \end{bmatrix}$$

fila 3 := fila 3 + (0,250) fila 2:

fila 4 := fila 4 - (0,500) fila 2:

$$\begin{bmatrix} 3.00 & 3.00 & -1.00 & -1.00 & 4.00 \\ & -4.00 & -0.333 & -1.33 & -3.67 \\ & & 0.250 & 2.00 & 5.75 \\ & & 3.50 & 2.00 & 2.50 \end{bmatrix}$$

Tercer paso del proceso directo:

Como el elemento de mayor valor absoluto esta en la fila 4, se intercambian las filas 3 y 4:

$$\begin{bmatrix} 3.00 & 3.00 & -1.00 & -1.00 & 4.00 \\ & -4.00 & -0.333 & -1.33 & -3.67 \\ & & 3.50 & 2.00 & 2.50 \\ & & 0.250 & 2.00 & 5.75 \end{bmatrix}$$

fila 4 := fila 4 – (0,0714) fila 3:

$$\begin{bmatrix} 3.00 & 3.00 & -1.00 & -1.00 & 4.00 \\ & -4.00 & -0.333 & -1.33 & -3.67 \\ & & 3.50 & 2.00 & 2.50 \\ & & & 1.86 & 5.57 \end{bmatrix}$$

Resulta entonces el sistema escalonado:

$$\begin{aligned} 3.00x_1 + 3.00x_2 - 1.00x_3 - 1.00x_4 &= 4.00 \\ -4.00x_2 - 0.333x_3 - 1.33x_4 &= -3.67 \\ 3.50x_3 + 2.00x_4 &= 2.50 \\ 1.86x_4 &= 5.57 \end{aligned}$$

Proceso inverso:

$$x_4 = \frac{5.57}{1.86} = 2.99$$

$$x_3 = \frac{2.50 - (2.00)(2.99)}{3.50} = -0.994$$

$$x_2 = \frac{-3.67 + (0.333)(-0.994) + (1.33)(2.99)}{-4.00} = 0.00608$$

$$x_1 = \frac{4.00 - (3.00)(0.00608) + (1.00)(-0.994) + (1.00)(2.99)}{3.00} = 1.99$$

Como en este ejemplo se trata de un sistema pequeño (cuarto orden) no hay una diferencia apreciable entre ambas estrategias de pivote y en ambos casos se obtienen pequeños errores debido a que solo se ha trabajado con tres cifras exactas.

Algoritmo en pseudo código

Se trata de resolver el sistema de n ecuaciones lineales con n incógnitas $\mathbf{Ax} = \mathbf{b}$ donde se supone que $\mathbf{A}$ es no singular. Se consideran datos del problema el número n y los coeficiente a_{ij} ($i = 1, 2, \dots, n; j = 1, 2, \dots, n$) de $\mathbf{A}$ y b_i ($i = 1, 2, \dots, n$) de $\mathbf{b}$. El algoritmo se ha acompañado de algunos comentarios entre llaves, para darle más claridad. La orden “Seleccionar la fila pivote i – sima” se especificará después.

```

C := [A|b] {Se forma la matriz ampliada C, de  $n$  filas y  $n + 1$  columnas}
{Aquí comienza el proceso directo}
for  $i = 1$  to  $n - 1$ 
    Seleccionar la fila pivote  $i$  – sima {Depende de la estrategia de pivote usada}
    for  $k = i + 1$  to  $n$ 
         $m := \frac{c_{ki}}{c_{ii}}$ 
         $fila(k) := fila(k) - m\,fila(i)$ 
    end
end
{Aquí comienza el proceso inverso}
 $i := n$ 
repeat
     $x_i := c_{i,n+1}$ 
    for  $j = i + 1$  to  $n$  {Cuando  $i = n$  este lazo no se ejecuta}
         $x_i := x_i - c_{ij}x_j$ 
    end
     $x_i := \frac{x_i}{c_{ii}}$ 
     $i := i - 1$ 
until  $i = 0$ 
La solución es  $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$  {El supraíndice  $T$  indica traspuesta}
Terminar

```

En el caso de usar la estrategia elemental de pivote, el algoritmo “Seleccionar la fila pivote i -sima” sería:

```

if  $c_{ii} = 0$  then
     $k := i$ 
    repeat
         $k := k + 1$ 
    until  $c_{ki} \neq 0$  or  $k > n$  {Si se llega a  $k > n$  y todos los  $c_{ki}$  han sido cero es
                                porque la matriz del sistema es singular}
    if  $k > n$  then
        El sistema no tiene solución única pues la matriz es singular
        Terminar
    else
        Intercambiar  $fila(k)$  con  $fila(i)$ 
    end
end

```

En el caso de usar la estrategia parcial de pivote, el algoritmo “Seleccionar la fila pivote i -sima” sería:

```

 $max := |c_{ii}|$  {La variable  $max$  guarda el mayor coeficiente obtenido hasta el
                momento}
 $Fila\ de\ max := i$  {La variable  $Fila\ de\ max$  guarda el número de la fila donde
                    ocurrió el máximo}

```

```

for  $k = i + 1$  to  $n$ 
    if  $|c_{ki}| > max$  then
         $max := |c_{ki}|$ 
         $Fila\ de\ max := k$ 
    end
end
if  $k > i$  then    {Si  $k = i$  el intercambio no tiene sentido}
    Intercambiar  $fila(k)$  con  $fila(i)$ 
end

```

Cantidad de operaciones del método de Gauss

Es muy importante conocer de que orden es el número de operaciones aritméticas que se necesita para ejecutar el algoritmo de Gauss para un sistema de orden n , ya que el interés principal ahora es obtener métodos eficientes para resolver sistemas con un número alto de ecuaciones. Primero se analizará el proceso directo y después el inverso.

Proceso directo:

Con vistas a simplificar y como el intercambio de filas no siempre se produce y solo sería necesario a lo más en $n - 1$ ocasiones, no se tomará en cuenta esta operación. Con el mismo objetivo, solo se contarán las multiplicaciones y divisiones, no las sumas, las cuales consumen mucho menos tiempo en la máquina.

En el paso 1 se realizan $n - 1$ transformaciones elementales de tipo II que requieren cada una un cociente (para hallar m) y n productos (recuérdese que las filas tienen $n + 1$ elementos). Un total de $(n - 1)(n + 1)$ multiplicaciones y divisiones.

En el paso 2 se realizan $n - 2$ transformaciones elementales de tipo II que requieren cada una un cociente y $n - 1$ productos. Un total de $(n - 2)(n)$ multiplicaciones y divisiones.

En general, en el paso i ($i = 1, 2, \dots, n - 1$) del proceso directo se necesitan $(n - i)(n - i + 2)$ multiplicaciones y divisiones.

Entonces, en todo el proceso directo la cantidad de operaciones de multiplicar y dividir será:

$$\sum_{i=1}^{n-1} (n-i)(n-i+2)$$

En los libros de Álgebra Superior se estudian formas de calcular este tipo de sumas. Aquí no se entrará en los detalles. El resultado de la suma es:

$$\frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n \quad (8)$$

Proceso inverso:

Para hallar la variable x_n se requiere una división y ninguna multiplicación.

Para hallar el valor de x_{n-1} hacen falta una multiplicación y una división.

Para hallar el valor de x_{n-2} hacen falta dos productos y una división.

En general, cada variable requiere de una división y la cantidad de productos aumenta desde 0 hasta $n - 1$ que se necesitan para calcular x_1 . El número total de multiplicaciones y divisiones del proceso inverso será de:

$$\sum_{i=1}^n i = 1 + 2 + 3 + \dots + n = \frac{n(n+1)}{2}$$

Es decir:
$$\frac{1}{2}n^2 + \frac{1}{2}n \quad (9)$$

Si se comparan las expresiones (8) y (9) se comprenderá que el proceso inverso requiere muy pocas operaciones en comparación con el proceso directo. Por ejemplo, un sistema de 30 ecuaciones requerirá, para el proceso directo, de:

$$\frac{1}{3}(30)^3 + \frac{1}{2}(30)^2 - \frac{5}{6}(30) = 9000 + 450 - 25 = 9425 \text{ productos y cocientes}$$

Mientras el proceso inverso necesitará:

$$\frac{1}{2}(30)^2 + \frac{1}{2}(30) = 450 + 15 = 465 \text{ productos y cocientes}$$

El total de operaciones es de unas 10000 y se realiza en una computadora personal en fracciones de segundo. Compárese con la astronómica cifra de $8,22 \cdot 10^{33}$ obtenida al principio de este capítulo mediante un procedimiento ineficiente.

La expresión para el número total de multiplicaciones y divisiones será la suma de (8) y (9):

$$\frac{1}{3}n^3 + n^2 - \frac{1}{3}n$$

Como usualmente no se necesita sino una idea del orden de este número y no su valor exacto, se suele utilizar solamente su sumando más significativo, que es el de exponente 3. Se concluye que:

La cantidad de multiplicaciones y divisiones que requiere el método de Gauss para resolver un sistema de n ecuaciones con n incógnitas es del orden de

$$\frac{1}{3}n^3$$

Aunque aquí no se hayan tenido en cuenta las sumas, su número es prácticamente el mismo que el de multiplicaciones, así que no aumenta significativamente cualquier estimado de tiempo que se haga con la fórmula obtenida.

Ejemplo 2

Haga un estimado del tiempo que requerirá una computadora capaz de realizar 50 000 multiplicaciones o divisiones por segundo para resolver un sistema lineal de 1000 ecuaciones con 1000 incógnitas utilizando el método de Gauss.

Solución:

La cantidad de productos y cocientes es del orden de $\frac{1}{3}n^3 = \frac{1}{3}(1000)^3 = 333\,000\,000$

El tiempo necesario será aproximadamente: $\frac{333000000}{50000} = 6600$ segundos, es decir, aproximadamente 2 horas.

Ejercicios

En los siguientes ejercicios, utilice un programa computacional del método de Gauss, preferiblemente confeccionado por usted. Si no cuenta con un programa adecuado, realice los cálculos mediante una calculadora trabajando con 5 cifras exactas.

1. Resuelva los siguientes sistemas lineales utilizando el método de Gauss con estrategia parcial de pivote.

$$\begin{aligned} &2,6x - 2,7y + 5,2z = 7,25 \\ \text{a)} \quad &3,1x + 0,5y - z = 4,33 \\ &2,9x - 5,2y - 4,3z = -5,6 \end{aligned}$$

$$\begin{aligned} &\frac{2}{5}x - \frac{1}{3}y + \frac{2}{7}z = \frac{1}{9} \\ \text{c)} \quad &\frac{1}{2}x + \frac{4}{5}y = \frac{3}{11}z \\ &\frac{2}{3}x - \frac{3}{4}y + \frac{4}{5}z = \frac{5}{6} \end{aligned}$$

$$\begin{aligned} &3x - 45 = u - v + z \\ &z = x + u \\ \text{b)} \quad &4u + v = 18 - z \\ &5 = x + z + u + v \end{aligned}$$

$$\begin{aligned} &e^x + 3y - \tan z = 22 \\ \text{d)} \quad &-2y = 2e^x + 3 \tan z \\ &\frac{1}{3}e^x = 5 - y \end{aligned}$$

$$\begin{aligned} &e^{2x}e^{3y}e^{-4z} - 3e^xe^{2y}e^z + 4e^{4x}e^ye^{-z} = -2,5 \\ \text{e)} \quad &2e^xe^{2y}e^z - e^{2x}e^{3y}e^{-4z} - 5e^{4x}e^ye^{-z} = -1,2 \\ &e^{4x}e^ye^{-z} + 2e^{2x}e^{3y}e^{-4z} + 3e^xe^ze^{2y} = 13,8 \end{aligned}$$

2. Se sabe que el polinomio $p(x) = ax^3 + bx^2 + cx + d$ satisface las condiciones:

$$\begin{aligned} p(1,3) &= 2,8 \\ p(2,1) &= 3,2 \\ p(2,7) &= 6 \\ p(3,1) &= 0 \end{aligned}$$

Halle a , b , c y d .

3. Halle el punto de intersección de los planos: $2x + y - z = 7$; $2x + y + z = 8$; $x - 3y + z = -2$.

4. Halle el punto del espacio donde la recta:

$$\frac{x-2}{3} = \frac{y-1}{4} = z+2$$

se encuentra con el plano cuyos interceptos con los ejes coordenados son $x = 1$, $y = 2$, $z = 3$.

5. La función $f(x, y, z) = kx^A y^B z^C$ satisface las condiciones:

x	y	z	$f(x, y, z)$
2	3	1	1,38
1	2	2	1,25
3	1	1	1,05
1	4	3	4,3

Determine los valores de los parámetros k , A , B y C .

6. Halle la ecuación del plano vertical que pasa por el origen de coordenadas y por el punto donde se intersecan las rectas

$$\frac{x-2}{2} = \frac{y+1}{-3} = \frac{z-2}{4} \quad \text{y} \quad \frac{x-3}{1} = \frac{y+2}{-2} = \frac{z-1}{5}$$

7. Halle una curva paramétrica del tipo: $x = a_1 t^2 + b_1 t + c_1$; $y = a_2 t^2 + b_2 t + c_2$ que pase por los tres puntos $P_1(3, 4)$, $P_2(1, 2)$, $P_3(1, 3)$ en ese mismo orden.
8. Elabore un algoritmo en pseudo código que permita hallar y dibujar curvas paramétricas del tipo anterior para tres puntos ordenados cualesquiera del plano.
9. Elabore un algoritmo en pseudo código que permita hallar la ecuación de la circunferencia que determinen tres puntos no colineales cualesquiera del plano.
10. El esquema de la figura 1 representa un sistema de cañerías de dos entradas con gastos de $Q_1 = 200$ litros por minuto y $Q_2 = 300$ litros por minuto y una salida con gasto Q_3 litros por minuto. Determine el gasto que circula por cada tramo de tubería si se sabe que, debido al diámetro y longitud de los tubos, x_2 es tres veces menor que x_1 y x_3 es dos veces mayor que x_4 .

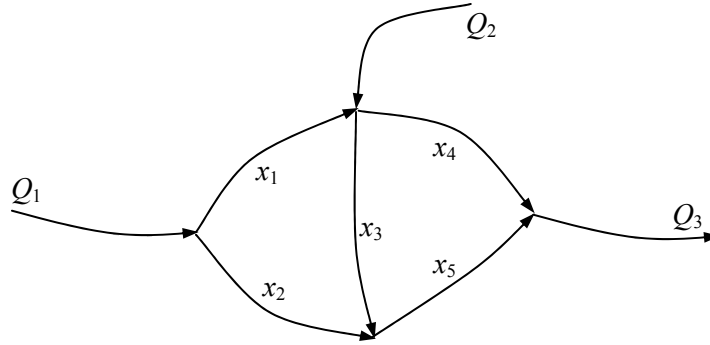


Figura 1

3.3 Consecuencias del método de Gauss

El método de Gauss puede ser adaptado para resolver otros problemas especiales. Aquí se analizarán tres de ellos: la solución de sistemas tridiagonales, el cálculo de determinantes y la inversión de matrices.

El método de Gauss para sistemas tridiagonales

En muchos problemas importantes aparecen sistemas de ecuaciones muy grandes con la característica especial de que la matriz del sistema es tridiagonal, esto es, solo la diagonal principal y las dos adyacentes, no son nulas. Para sistemas de este tipo se puede obtener un algoritmo de Gauss especializado sumamente eficiente. Ante todo, es obvio que en lugar de guardar todos los elementos de la matriz del sistema (la mayoría de los cuales son ceros) es preferible guardar solamente las tres diagonales no nulas y el vector de términos independientes. El sistema, por tanto se expresa matricialmente como:

$$\begin{bmatrix} a_1 & c_1 & & & \\ b_2 & a_2 & c_2 & & \\ & b_3 & a_3 & c_3 & \\ & & b_4 & \ddots & \ddots \\ & & & \ddots & a_{n-1} & c_{n-1} \\ & & & & b_n & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}$$

donde se ha tomado el convenio de no escribir los términos (todos ellos ceros) fuera de las tres diagonales. Para aplicar el método de Gauss al sistema, primero debe construirse la matriz ampliada:

$$\begin{bmatrix} a_1 & c_1 & & & & d_1 \\ b_2 & a_2 & c_2 & & & d_2 \\ & b_3 & a_3 & c_3 & & d_3 \\ & & b_4 & \ddots & \ddots & \vdots \\ & & & \ddots & a_{n-1} & c_{n-1} & d_{n-1} \\ & & & & b_n & a_n & d_n \end{bmatrix} \quad (1)$$

En el desarrollo que sigue se supondrá que los términos que se utilizan como denominadores no son ceros. Posteriormente se darán condiciones que garantizan que ello se cumpla. Dividiendo por a_1 toda la primera fila queda:

$$\begin{bmatrix} 1 & p_1 & & & & q_1 \\ b_2 & a_2 & c_2 & & & d_2 \\ & b_3 & a_3 & c_3 & & d_3 \\ & & b_4 & \ddots & \ddots & \vdots \\ & & & \ddots & a_{n-1} & c_{n-1} d_{n-1} \\ & & & & b_n & a_n d_n \end{bmatrix}$$

donde $p_1 = \frac{c_1}{a_1}$ y $q_1 = \frac{d_1}{a_1}$

Ahora se multiplicará la primera fila por $-b_2$ para sumársela a la segunda fila, de modo que se anule el segundo elemento de la columna 1 (que es único de esa columna por debajo de la diagonal que no era ya cero). La matriz quedará entonces:

$$\begin{bmatrix} 1 & p_1 & & & & q_1 \\ 0 & r_2 & c_2 & & & d_2 - b_2 q_1 \\ & b_3 & a_3 & c_3 & & d_3 \\ & & b_4 & \ddots & \ddots & \vdots \\ & & & \ddots & a_{n-1} & c_{n-1} d_{n-1} \\ & & & & b_n & a_n d_n \end{bmatrix}$$

donde: $r_2 = a_2 - b_2 p_1$

Ahora se divide toda la segunda fila por r_2 con el objetivo de hacer 1 el elemento de la diagonal principal. La matriz ampliada queda en la forma:

$$\begin{bmatrix} 1 & p_1 & & & & q_1 \\ 0 & 1 & p_2 & & & q_2 \\ & b_3 & a_3 & c_3 & & d_3 \\ & & b_4 & \ddots & \ddots & \vdots \\ & & & \ddots & a_{n-1} & c_{n-1} d_{n-1} \\ & & & & b_n & a_n d_n \end{bmatrix}$$

donde: $p_2 = \frac{c_2}{r_2}$ y $q_2 = \frac{d_2 - b_2 q_1}{r_2}$

Si ahora se usa la segunda fila como pivote para colocar un cero en lugar de b_3 y se definen r_3, p_3 y q_3 de manera análoga al paso anterior, la matriz se transforma en:

$$\begin{bmatrix} 1 & p_1 & & & q_1 \\ 0 & 1 & p_2 & & q_2 \\ & 0 & 1 & p_3 & q_3 \\ & & b_4 & \ddots & \vdots \\ & & & \ddots & a_{n-1} & c_{n-1} & d_{n-1} \\ & & & & b_n & a_n & d_n \end{bmatrix}$$

donde: $r_3 = a_3 - b_3 p_2$ $p_3 = \frac{c_3}{r_3}$ y $q_3 = \frac{d_3 - b_3 q_2}{r_3}$

Procediendo de esta manera, la matriz ampliada se transforma en una matriz equivalente con todos los elementos nulos por debajo de la diagonal:

$$\begin{bmatrix} 1 & p_1 & & & q_1 \\ & 1 & p_2 & & q_2 \\ & & 1 & p_3 & q_3 \\ & & & \ddots & \vdots \\ & & & & 1 & p_{n-1} & q_{n-1} \\ & & & & & 1 & q_n \end{bmatrix}$$

donde $p_1 = \frac{c_1}{a_1}$; $q_1 = \frac{d_1}{a_1}$ (2)

y $r_i = a_i - b_i p_{i-1}$ $p_i = \frac{c_i}{r_i}$ y $q_i = \frac{d_i - b_i q_{i-1}}{r_i}$ para $i = 2, 3, \dots, n$ (3)

El sistema de ecuaciones representado por esta matriz tiene la forma:

$$\begin{bmatrix} 1 & p_1 & & & \\ & 1 & p_2 & & \\ & & 1 & p_3 & \\ & & & \ddots & \ddots \\ & & & & 1 & p_{n-1} \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_{n-1} \\ q_n \end{bmatrix}$$

$$\text{Esto es: } \left\{ \begin{array}{l} x_1 + p_1 x_2 = q_1 \\ x_2 + p_2 x_3 = q_2 \\ x_3 + p_3 x_4 = q_3 \\ \vdots \\ x_{n-1} + p_{n-1} x_n = q_{n-1} \\ x_n = q_n \end{array} \right.$$

El proceso inverso se simplifica notablemente por el hecho de que en cada ecuación solamente aparecen dos incógnitas, salvo en la última que permite encontrar x_n . Despejando sucesivamente las incógnitas de abajo hacia arriba, se obtiene:

$$\begin{cases} x_n = q_n \\ x_{n-1} = q_{n-1} - p_{n-1}x_n \\ x_{n-2} = q_{n-2} - p_{n-2}x_{n-1} \\ \vdots \\ x_2 = q_2 - p_2x_3 \\ x_1 = q_1 - p_1x_2 \end{cases} \quad (4)$$

Obsérvese que para resolver un sistema tridiagonal se requiere de muy pocas operaciones aritméticas:

Ecuaciones (2) y (3) 6 operaciones n veces = $6n$ operaciones

Ecuaciones (4) 2 operaciones n veces $= 2n$ operaciones

Total: $8n$ operaciones

La cantidad de operaciones que requiere el método de Gauss especializado en sistemas tridiagonales para resolver un sistema de n ecuaciones con n incógnitas es del orden de

$8n$

Algoritmo en pseudo código para sistemas tridiagonales

Se trata de resolver un sistema tridiagonal para el cual se conocen las diagonales:

 $\{a_1, a_2, \dots, a_n\}$ (diagonal principal) $\{b_2, b_3, \dots, b_n\}$ (diagonal debajo de la principal) $\{c_1, c_2, \dots, c_{n-1}\}$ (diagonal encima de la principal)

y los términos independientes $\{d_1, d_2, \dots, d_n\}$.

Puede demostrarse que si la diagonal principal es predominante, es decir si en cada fila i el valor absoluto de a_i es mayor que la suma de los valores absolutos de b_i y c_i , entonces, ninguno de los denominadores que aparecen en el algoritmo que sigue se hacen cero.

```

 $p_1 := \frac{c_1}{a_1}$ 
 $q_1 := \frac{d_1}{a_1}$ 
for  $i = 2$  to  $n$ 
     $r_i = a_i - b_i p_{i-1}$ 
     $p_i = \frac{c_i}{r_i}$ 
     $q_i = \frac{d_i - b_i q_{i-1}}{r_i}$ 
end
 $x_n := q_n$ 
 $i := n - 1$ 
repeat
     $x_i := q_i - p_i x_{i+1}$ 
     $i := i - 1$ 
until  $i = 0$ 
La solución es  $x_1, x_2, \dots, x_n$ 
Terminar

```

Ejemplo 1

Haga un estimado del tiempo que requerirá una computadora capaz de realizar 50 000 multiplicaciones o divisiones por segundo para resolver un sistema tridiagonal de 1000 ecuaciones con 1000 incógnitas utilizando el método de Gauss especializado para sistemas tridiagonales.

Solución:

El total de operaciones es del orden de $8n$, es decir, 8000. El tiempo requerido será de:

$$\frac{8000}{50000} = 0,16 \text{ segundos}$$

Compárese con el tiempo de unas dos horas que se hubiera requerido (ejemplo 2 de la sección anterior) para resolver un sistema de este tamaño mediante el algoritmo general de Gauss en esta misma máquina.

Cálculo de determinantes

Como se vio en la sección anterior, en el proceso directo de Gauss para llevar la matriz del sistema a una forma escalonada, se utilizan transformaciones elementales de tipo I y de tipo II. La

transformación de tipo I (intercambiar dos filas) produce el efecto colateral de cambiar el signo del determinante de la matriz A del sistema; la transformación de tipo II no afecta el valor del determinante.

Por otra parte, es muy simple demostrar que el determinante de una matriz triangular es el producto de los elementos en la diagonal de la matriz. En efecto, utilizando el desarrollo por menores de un determinante de orden n respecto a su primera columna, que tiene un solo elemento no nulo, se tiene que:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}$$

Desarrollando el determinante de orden $n - 1$ respecto a su primera columna, que solo posee un elemento no nulo, se tiene que:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22} \begin{vmatrix} a_{33} & a_{34} & \cdots & a_{3n} \\ 0 & a_{44} & \cdots & a_{4n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}$$

Repitiendo la operación las veces necesarias se llega a:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}a_{33} \cdots a_{nn}$$

A partir de lo anterior, una forma eficiente de calcular el determinante de una matriz es realizar sobre ella transformaciones elementales de tipo I y II que la reduzcan a una matriz triangular (tal como hace el proceso directo del método de Gauss) tomando en cuenta que cada aplicación de una transformación de tipo I produce un cambio en el signo del determinante. Una vez que la matriz tiene forma triangular, el determinante se obtiene como el producto de los elementos de la diagonal principal.

El cálculo de un determinante de orden n por esta vía requiere una cantidad de operaciones que está determinada por el proceso directo de Gauss, que es mayor consumidor de tiempo de

cómputo. Por tanto en número de operaciones es del orden de $\frac{1}{3}n^3$.

En el algoritmo que sigue, se utiliza una variable llamada *Signo* a la que inicialmente se le da valor 1, y que cambia de signo cada vez que se realiza un intercambio de filas. Al final el producto de los elementos de la diagonal es afectada por la variable *Signo* para obtener el valor del determinante. En el proceso directo de Gauss se utiliza la estrategia de pivote parcial.

Algoritmo para calcular determinantes

El algoritmo que sigue utiliza como datos el entero $n > 1$ y los elementos de la matriz cuadrada **A** de orden n .

```
Signo := 1
for i = 1 to n - 1
    {En este sector se halla la fila pivote mediante la estrategia parcial}
    max := |aii|
    Fila de max := i
    for k = i + 1 to n
        if |aki| > max then
            max := |aki|
            Fila de max := k
        end
    end
    if max = 0 then
        El determinante es cero
        Terminar
    end
    if k > i then
        Intercambiar fila(k) con fila(i)
        Signo := - Signo
    end

    {En el sector que sigue se anulan los elementos de la columna i que se hallan
    debajo de la diagonal}
    for k = i + 1 to n
        m :=  $\frac{a_{ki}}{a_{ii}}$ 
        fila(k) := fila(k) - m fila(i)
    end
end
Determinante := Signo
for i = 1 to n
    Determinante := (aii)(Determinante)
end
El valor del determinante es Determinante
Terminar
```

Inversión de matrices mediante el método de Gauss

El problema de hallar la inversa de la matriz cuadrada **A** de orden n equivale a encontrar otra matriz **X** del mismo orden tal que:

$$\mathbf{AX} = \mathbf{I} \quad (5)$$

donde **I** es la matriz identidad de orden n . En forma desarrollada, la ecuación (5) es:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (6)$$

Si las matrices $\mathbf{X}$ e $\mathbf{I}$ se dividen en bloques por columnas, la ecuación matricial (6) toma la forma:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \cdot \left(\begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix} \cdots \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix} \right) = \left(\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \cdots \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right)$$

Esta ecuación puede ser separada en n ecuaciones más simples:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Cada una de estas n ecuaciones matriciales no es otra cosa que un sistema lineal de ecuaciones. Resolviendo cada uno de los sistemas mediante el método de Gauss, se tendrían las n columnas de la matriz inversa. Como los n sistemas poseen la misma matriz de coeficientes (lo que cambia es el vector de términos independientes) se puede hacer una adaptación del método de Gauss en la cual la matriz ampliada se forma añadiendo a la matriz $\mathbf{A}$ las n columnas de términos independientes y entonces el proceso directo (que es el que más operaciones requiere) se realiza con esta matriz ampliada. Por supuesto, el proceso inverso del método de Gauss habrá que realizarlo n veces, una con cada una de las columnas con que se amplió la matriz $\mathbf{A}$. Puede

probarse que la cantidad de operaciones que requiere todo el algoritmo es del orden de $\frac{4}{3}n^3$.

Algoritmo para invertir una matriz mediante el método de Gauss

El algoritmo en pseudo código que sigue utiliza la idea anterior para hallar la inversa de la matriz cuadrada **A** de orden n . En el proceso directo de Gauss se utiliza la estrategia parcial de pivote. Se supone que **A** es no singular. El algoritmo utiliza como datos el entero $n > 1$ y los elementos de la matriz **A**.

```
C := [A|I] {La matriz C se obtiene ampliando A con los elementos de la idéntica}
for  $i = 1$  to  $n - 1$ 
    {En este sector se halla la fila pivote mediante la estrategia parcial}
     $max := |c_{ii}|$ 
    Fila de  $max := i$ 
    for  $k = i + 1$  to  $n$ 
        if  $|c_{ki}| > max$  then
             $max := |c_{ki}|$ 
            Fila de  $max := k$ 
        end
    end
    if  $k > i$  then
        Intercambiar  $fila(k)$  con  $fila(i)$ 
    end

    {En el sector que sigue se anulan los elementos de la columna  $i$  que se hallan
    debajo de la diagonal}
    for  $k = i + 1$  to  $n$ 
         $m := \frac{c_{ki}}{c_{ii}}$ 
         $fila(k) := fila(k) - m fila(i)$ 
    end
end
{A partir de aquí se realiza el proceso inverso}

for  $j = 1$  to  $n$  {El proceso inverso se repetirá  $n$  veces}
     $i := n$ 
    repeat {En este ciclo se halla el elemento  $i$  de la columna  $j$  de la matriz inversa}
         $x_{ij} := c_{i,n+j}$  {El término independiente está en la columna  $n + j$ }
        for  $k = i + 1$  to  $n$  {Cuando  $i = n$  este lazo no se ejecuta}
             $x_{ij} := x_{ij} - c_{ik}x_{kj}$ 
        end
         $x_{ij} := \frac{x_{ij}}{c_{ii}}$ 
         $i := i - 1$ 
    until  $i = 0$ 
end
La matriz inversa es X =  $[x_{ij}]$ 
Terminar
```

Ejercicios

En los siguientes ejercicios, utilice programas computacionales, preferiblemente confeccionados por usted. De no contar con un programa adecuado, use una calculadora y trabaje con cinco cifras significativas en los cálculos.

1. Resuelva los siguientes sistemas de ecuaciones mediante el algoritmo de Gauss especializado en sistemas tridiagonales. Verifique previamente que en cada caso la diagonal es predominante.

$$\begin{array}{lll}
 7x_1 + 3x_2 = 5 & 5x_1 + 2x_2 = 6 & -3x_1 + x_2 = 6 \\
 2x_1 + 6x_2 - x_3 = 6 & -x_1 + 7x_2 - 3x_3 = 4 & 3x_1 + 7x_2 + 3x_3 = 2 \\
 \text{a) } x_2 + 8x_3 + 3x_4 = 2 & \text{b) } 2x_2 + 9x_3 + 6x_4 = 10 & \text{c) } x_2 + 6x_3 + 2x_4 = 3 \\
 3x_3 + 8x_4 - x_5 = 2 & x_3 + 6x_4 - 2x_5 = 1 & 3x_3 + 7x_4 + 2x_5 = 5 \\
 x_4 + 3x_5 = 4 & 3x_4 + 5x_5 = 6 & 4x_4 + 6x_5 = 2
 \end{array}$$

2. Como se verá en el próximo capítulo, para hacer pasar un tipo de curva llamado spline cúbico natural por n puntos del plano, se requiere resolver un sistema de ecuaciones de $n-2$ ecuaciones lineales, el cual es tridiagonal y con la diagonal predominante. Suponga una computadora que es capaz de realizar 50000 multiplicaciones por segundo y estime qué tiempo requiere esa máquina para calcular un spline que pasa por 100 puntos de la pantalla. Considere dos posibilidades: a) Utilizar el método general de Gauss, b) Utilizar el método de Gauss especializado en sistemas tridiagonales.
3. Para hallar un spline cúbico natural que pase por los $n+1$ puntos del plano: $P_0(x_0, y_0)$, $P_1(x_1, y_1)$, ..., $P_n(x_n, y_n)$ con $x_0 < x_1 < \dots < x_n$ se necesita resolver el sistema tridiagonal

$$\frac{h_{i-1}}{6} M_{i-1} + \frac{h_{i-1} + h_i}{3} M_i + \frac{h_i}{6} M_{i+1} = \frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}} \quad i = 1, 2, \dots, n-1$$

donde $h_i = x_{i+1} - x_i \quad i = 0, 1, \dots, n-1$
 $M_0 = M_n = 0$

Elabore un algoritmo en pseudo código que utilice el método de Gauss especializado en sistemas tridiagonales, para calcular los coeficientes M_i ($i = 1, 2, \dots, n-1$). Verifique previamente si este sistema tiene la diagonal predominante.

4. Calcule los siguientes determinantes mediante el algoritmo basado en el proceso directo de Gauss.

$$\begin{array}{lll}
 \text{a) } \begin{vmatrix} 3 & -4 & 3 & 2 \\ 5 & 7 & 1 & -4 \\ 3 & 2 & 2 & 1 \\ 0 & 2 & 7 & 1 \end{vmatrix} & \text{b) } \begin{vmatrix} 2 & 3 & 1 & 1 \\ 0 & 4 & 2 & 6 \\ 1 & 4 & -2 & 4 \\ 2 & 1 & 3 & 5 \end{vmatrix} & \text{c) } \begin{vmatrix} 3 & -2 & 0 & 2 \\ 1 & 2 & 3 & 4 \\ 3 & 2 & 3 & 7 \\ 3 & 5 & 6 & 2 \end{vmatrix}
 \end{array}$$

5. El siguiente algoritmo para calcular determinantes es una modificación del que se mostró en las páginas anteriores. Analice cómo funciona el mismo y detecte y rectifique un error que contiene.

```

Cambios := 0
for i = 1 to n - 1
    if  $a_{ii} = 0$  then
         $k := i$ 
        repeat
             $k := k + 1$ 
        until  $a_{ki} \neq 0$  or  $k > n$ 
        if  $k > n$  then
            El determinante es cero
            Terminar
        end
        if  $k < i$  then
            Intercambiar  $\text{fila}(k)$  con  $\text{fila}(i)$ 
             $\text{Cambios} := \text{Cambios} + 1$ 
        end
    end
    for  $k = i + 1$  to  $n$ 
         $m := \frac{a_{ki}}{a_{ii}}$ 
         $\text{fila}(k) := \text{fila}(k) - m \text{fila}(i)$ 
    end
end
Determinante := Cambios mod 2
for i = 1 to n
    Determinante :=  $(a_{ii})(\text{Determinante})$ 
end
El valor del determinante es Determinante
Terminar

```

6. Suponga que utilizando una calculadora de mano usted es capaz de multiplicar o sumar dos números cualesquiera en 5 segundos. Calcule que tiempo tardaría en calcular manualmente determinantes de orden 4, 6, 8 y 10 en dos variantes: a) Utilizando el método basado en el algoritmo de Gauss. b) Desarrollando el determinante por menores. En ambos casos suponga que los algoritmos requieren tantas sumas como multiplicaciones.
7. Halle la matriz inversa de cada una de las siguientes matrices mediante el método basado en el algoritmo de Gauss.

a)
$$\begin{bmatrix} 3 & 5 & 2 \\ 3 & -1 & 3 \\ 4 & 2 & 4 \end{bmatrix}$$

b)
$$\begin{bmatrix} 1 & 2 & -1 & 3 \\ 3 & 0 & 1 & 2 \\ 2 & 3 & 2 & -2 \\ 1 & 1 & 3 & 4 \end{bmatrix}$$

c)
$$\begin{bmatrix} 1 & -1 & 0 & 2 \\ 2 & 1 & 3 & 1 \\ 1 & 1 & 2 & 2 \\ 3 & 0 & 1 & 1 \end{bmatrix}$$

8. Suponga que usted necesita resolver un número grande p de sistemas lineales, todos con la misma matriz de coeficientes $\mathbf{A}$ pero con diferentes vectores de términos independientes. Elabore un algoritmo basado en el método de Gauss que utilice una idea similar a la que se

utilizó para hallar la inversa de una matriz: ampliar la matriz de los coeficientes con todos los vectores independientes, realizar una vez el proceso inverso y p veces el proceso inverso.

9. El esquema de la figura 2 representa un sistema real con cuatro causas x_1, x_2, x_3, x_4 y cuatro efectos y_1, y_2, y_3, y_4 . Como el sistema es lineal, las relaciones entre causas y efectos se pueden representar con una matriz $\mathbf{A}$, de modo que $\mathbf{y} = \mathbf{Ax}$ donde $\mathbf{y}$ es el vector de efectos y $\mathbf{x}$ el vector de causas. Determine la relación entre efectos y causas, de modo que a partir de un vector de efectos se pueda hallar el vector de causas que lo originó. Se sabe que:

$$y_1 = 3x_1 - 2x_2 + x_3 - 5x_4$$

$$y_2 = 2x_1 + 3x_2 - x_3 + 2x_4$$

$$y_3 = -x_1 + 2x_2 + 3x_3 + x_4$$

$$y_4 = x_1 - 3x_2 + 2x_3 - 4x_4$$

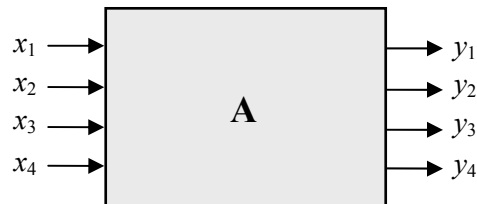


Figura 2

10. En el ejemplo 1 de la sección 3.1 se analizó un circuito de cuatro mallas por el método de la corrientes de malla. El sistema de ecuaciones obtenido se puede generalizar del siguiente modo:

$$\begin{aligned} R_{11}i_1 - R_{12}i_2 - \cdots - R_{1n}i_n &= V_1 \\ -R_{21}i_1 + R_{22}i_2 - \cdots - R_{2n}i_n &= V_2 \\ &\vdots \\ -R_{n1}i_1 - R_{n2}i_2 - \cdots + R_{nn}i_n &= V_n \end{aligned}$$

- Determine el sentido físico de los diferentes parámetros que aparecen en el sistema
- Expresa el sistema matricialmente como $\mathbf{V} = \mathbf{RI}$
- ¿Qué características tiene la matriz $\mathbf{R}$?
- ¿Cómo se puede expresar el vector $\mathbf{I}$ de corrientes de malla en función del vector $\mathbf{V}$ de voltajes de malla?

3.4 Sistemas mal condicionados

En la sección 1.7 fue tratado en general el problema de la estabilidad. Allí se dijo que un problema inestable es aquel en el cual pequeños cambios en los datos pueden producir grandes cambios en los resultados. Entre los sistemas lineales de ecuaciones se presentan a veces problemas inestables, que reciben más frecuentemente el nombre de mal condicionados (del inglés *ill-conditioned*). En el ejemplo 2 de aquella sección se mostró un pequeño sistema de dos

ecuaciones lineales mal condicionado, en el cual un pequeño cambio de 1% en uno de los datos causó un cambio del orden de 100% en la solución. En esta sección se verá un poco más profundamente el tema de los sistemas lineales mal condicionados, se estudiará la manera de medir el mal condicionamiento y se analizará la forma en que debe procederse en estos casos.

Ejemplo 1

Analice el mal condicionamiento de los sistemas lineales:

$$\begin{array}{ll} \text{a)} \quad \begin{cases} 4x_1 + 3x_2 + 8x_3 - 7x_4 = 8 \\ 5x_1 + 6x_2 - 3x_3 + 5x_4 = 13 \\ 2x_1 + 7x_2 + 4x_3 - 4x_4 = 9 \\ 3x_1 - 2x_2 + 5x_3 + 8x_4 = 14 \end{cases} & \text{b)} \quad \begin{cases} 5x_1 + 7x_2 + 6x_3 + 5x_4 = 23 \\ 7x_1 + 10x_2 + 8x_3 + 7x_4 = 32 \\ 6x_1 + 8x_2 + 10x_3 + 9x_4 = 33 \\ 5x_1 + 7x_2 + 9x_3 + 10x_4 = 31 \end{cases} \end{array}$$

Solución:

Ambos sistemas tienen como solución: $x_1 = 1$; $x_2 = 1$; $x_3 = 1$; $x_4 = 1$.

a) Al cambiar ligeramente los términos independientes la solución cambia de la siguiente forma:

$$\begin{array}{ll} \text{Para } \mathbf{b} = \begin{bmatrix} 8,1 \\ 13 \\ 9 \\ 14 \end{bmatrix} \text{ se obtiene } \mathbf{x} = \begin{bmatrix} 1,0199 \\ 0,9901 \\ 0,9984 \\ 0,9910 \end{bmatrix} & \text{Para } \mathbf{b} = \begin{bmatrix} 8 \\ 13,1 \\ 9 \\ 14 \end{bmatrix} \text{ se obtiene } \mathbf{x} = \begin{bmatrix} 1,0157 \\ 0,9999 \\ 0,9916 \\ 0,9993 \end{bmatrix} \\ \\ \text{Para } \mathbf{b} = \begin{bmatrix} 8 \\ 13 \\ 9,1 \\ 14 \end{bmatrix} \text{ se obtiene } \mathbf{x} = \begin{bmatrix} 0,9769 \\ 1,0188 \\ 1,0105 \\ 1,9968 \end{bmatrix} & \text{Para } \mathbf{b} = \begin{bmatrix} 8 \\ 13 \\ 9 \\ 14,1 \end{bmatrix} \text{ se obtiene } \mathbf{x} = \begin{bmatrix} 0,9961 \\ 1,0008 \\ 1,0091 \\ 1,0085 \end{bmatrix} \end{array}$$

En todos los casos que se analizaron, al realizar cambios del orden de 1% en el vector $\mathbf{b}$ se obtuvo una solución con cambios del orden de 1% del valor original.

$$\text{b) Para } \mathbf{b} = \begin{bmatrix} 23,1 \\ 32 \\ 33 \\ 31 \end{bmatrix} \text{ se obtiene } \mathbf{x} = \begin{bmatrix} 7,8000 \\ -3,1000 \\ -0,7000 \\ 2,0000 \end{bmatrix}$$

Con un cambio en $\mathbf{b}$ del orden de 0,5% se han producido en la solución cambios del orden de 700%.

Conclusión: Todo parece indicar que el sistema a) es bien condicionado. El sistema b) presenta mal condicionamiento.

■

El ejemplo 1 ilustra la forma más efectiva de probar que un sistema es mal condicionado: verificar que pequeños cambios en los datos provocan cambios mucho mayores en la solución. Esta prueba, en cambio no asegura el buen condicionamiento pues no es posible realizar toda la variedad de pequeñas alteraciones en los datos del problema; por eso en el caso a) la respuesta no es concluyente.

Una medida del mal condicionamiento

En lo que sigue se tratará de hallar una relación entre el cambio relativo del vector **b** y el cambio relativo del vector **x** en un sistema general

$$\mathbf{Ax} = \mathbf{b} \quad (1)$$

Como la norma de un vector mide el tamaño del vector, es natural definir el cambio sufrido por el vector **b** al pasar a **b₁** como la norma del vector diferencia **b – b₁**. El cambio relativo se definirá entonces dividiendo el resultado anterior por la norma de **b**. Esto es:

$$\text{Cambio relativo al cambiar } \mathbf{b} \text{ por } \mathbf{b}_1: \frac{\|\mathbf{b} - \mathbf{b}_1\|}{\|\mathbf{b}\|}$$

Sea ahora **x₁** el vector solución del sistema lineal obtenido al sustituir **b** por **b₁**, es decir:

$$\mathbf{Ax}_1 = \mathbf{b}_1 \quad (2)$$

$$\text{El cambio relativo sufrido por la solución será: } \frac{\|\mathbf{x} - \mathbf{x}_1\|}{\|\mathbf{x}\|}$$

Para establecer una relación entre estos cambios relativos, se restan miembro a miembro las ecuaciones (1) y (2) y se extrae **A** factor común:

$$\mathbf{A}(\mathbf{x} - \mathbf{x}_1) = \mathbf{b} - \mathbf{b}_1$$

Premultiplicando por **A⁻¹** ambos miembros de esta ecuación:

$$\mathbf{x} - \mathbf{x}_1 = \mathbf{A}^{-1}(\mathbf{b} - \mathbf{b}_1)$$

Ahora bien, según el axioma 6 de las normas matriciales, la norma del producto de una matriz por un vector es menor o igual que el producto de sus respectivas normas, así que:

$$\|\mathbf{x} - \mathbf{x}_1\| \leq \|\mathbf{A}^{-1}\| \cdot \|\mathbf{b} - \mathbf{b}_1\|$$

Dividiendo en ambos miembros por la norma de **x** se obtiene:

$$\frac{\|\mathbf{x} - \mathbf{x}_1\|}{\|\mathbf{x}\|} \leq \frac{\|\mathbf{A}^{-1}\| \cdot \|\mathbf{b} - \mathbf{b}_1\|}{\|\mathbf{x}\|} = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \cdot \frac{\|\mathbf{b} - \mathbf{b}_1\|}{\|\mathbf{A}\| \cdot \|\mathbf{x}\|} \quad (3)$$

Pero como $\mathbf{Ax} = \mathbf{b}$, aplicando de nuevo el axioma 6, se tiene que $\|\mathbf{A}\| \cdot \|\mathbf{x}\| \geq \|\mathbf{b}\|$, así que cambiando el producto $\|\mathbf{A}\| \cdot \|\mathbf{x}\|$ en el denominador de (3) por la norma de $\mathbf{b}$ la desigualdad se refuerza y se obtiene:

$$\frac{\|\mathbf{x} - \mathbf{x}_1\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \cdot \frac{\|\mathbf{b} - \mathbf{b}_1\|}{\|\mathbf{b}\|}$$

Como se ve, el término $\|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \cdot \frac{\|\mathbf{b} - \mathbf{b}_1\|}{\|\mathbf{b}\|}$ constituye una cota para el cambio relativo que sufrirá la solución del sistema. Esta cota se obtiene multiplicando el cambio relativo que haya sufrido $\mathbf{b}$ por el número

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

Valores pequeños de este número garantizan que el cambio relativo de $\mathbf{x}$ no será grande respecto al cambio relativo de $\mathbf{b}$.

Definición 1

El número de condición del sistema lineal $\mathbf{Ax} = \mathbf{b}$ se denota $\text{cond}(\mathbf{A})$ y se define como

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \quad \blacksquare$$

El número de condición de la matriz $\mathbf{A}$ permite medir el mal condicionamiento del sistema. Es fácil ver que:

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \geq \|\mathbf{AA}^{-1}\| = \|\mathbf{I}\| = 1$$

esto es, el número de condición siempre es mayor o igual que 1. Valores de $\text{cond}(\mathbf{A})$ próximos a 1 aseguran un buen condicionamiento del sistema lineal mientras valores mucho mayores que 1 indican un mal condicionamiento. Puede probarse que siempre se puede seleccionar un vector $\mathbf{b}$ y uno $\mathbf{b}_1$ de modo que

$$\frac{\|\mathbf{x} - \mathbf{x}_1\|}{\|\mathbf{x}\|} = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \cdot \frac{\|\mathbf{b} - \mathbf{b}_1\|}{\|\mathbf{b}\|}$$

así que un valor grande de $\text{cond}(\mathbf{A})$ significa que para determinados vectores $\mathbf{b}_1$ pudiera haber grandes cambios relativos en $\mathbf{x}$. En la práctica, valores de $\text{cond}(\mathbf{A})$ menores que 10 se toman como problemas bien condicionados mientras que valores mayores que 10 ya comienzan a ser considerados como mal condicionados.

Ejemplo 2

Halle el número de condición para cada uno de los sistemas lineales del ejemplo 1.

Solución:

a) En este caso el sistema es

$$\begin{cases} 4x_1 + 3x_2 + 8x_3 - 7x_4 = 8 \\ 5x_1 + 6x_2 - 3x_3 + 5x_4 = 13 \\ 2x_1 + 7x_2 + 4x_3 - 4x_4 = 9 \\ 3x_1 - 2x_2 + 5x_3 + 8x_4 = 14 \end{cases}$$

La matriz $\mathbf{A}$ del sistema se obtiene directamente. Su inversa se obtuvo mediante un programa confeccionado a partir del algoritmo mostrado en la sección anterior:

$$\mathbf{A} = \begin{bmatrix} 4 & 3 & 8 & -7 \\ 5 & 6 & -3 & 5 \\ 2 & 7 & 4 & -4 \\ 3 & -2 & 5 & 8 \end{bmatrix} \quad \mathbf{A}^{-1} = \begin{bmatrix} 0,1990 & 0,1569 & -0,2310 & -0,0394 \\ -0,0991 & -0,0006 & 0,1880 & 0,0077 \\ -0,0157 & -0,0839 & 0,1047 & 0,0910 \\ -0,0896 & -0,0065 & 0,0682 & 0,0848 \end{bmatrix}$$

$$\|\mathbf{A}\| = \max\{22, 19, 17, 18\} = 22 \quad \|\mathbf{A}^{-1}\| = \max\{0,6263, 0,2954, 0,2953, 0,2491\} = 0,6263$$

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| = (22)(0,6263) = 13,8$$

El sistema presenta solamente un ligero mal condicionamiento.

b) En este caso el sistema es

$$\begin{cases} 5x_1 + 7x_2 + 6x_3 + 5x_4 = 23 \\ 7x_1 + 10x_2 + 8x_3 + 7x_4 = 32 \\ 6x_1 + 8x_2 + 10x_3 + 9x_4 = 33 \\ 5x_1 + 7x_2 + 9x_3 + 10x_4 = 31 \end{cases}$$

La matriz $\mathbf{A}$ del sistema y su inversa son:

$$\mathbf{A} = \begin{bmatrix} 5 & 7 & 6 & 5 \\ 7 & 10 & 8 & 7 \\ 6 & 8 & 10 & 9 \\ 5 & 7 & 9 & 10 \end{bmatrix} \quad \mathbf{A}^{-1} = \begin{bmatrix} 68 & -41 & -17 & 10 \\ -41 & 25 & 10 & -6 \\ -17 & 10 & 5 & -3 \\ 10 & -6 & -3 & 2 \end{bmatrix}$$

$$\|\mathbf{A}\| = \max\{23, 32, 33, 31\} = 33 \quad \|\mathbf{A}^{-1}\| = \max\{136, 82, 35, 21\} = 136$$

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| = (33)(136) = 4488$$

El sistema es muy mal condicionado. ■

¿Qué hacer?

Una vez que se tiene la certeza o la sospecha de que un sistema puede presentar mal condicionamiento, es importante tomar medidas para disminuir los errores, ya que ellos representan pequeños cambios en los datos del problema. Para ello hay dos cosas que deben hacerse. Primero utilizar una estrategia de pivote parcial o total (si está disponible), de modo que no se produzca una ampliación de los errores de redondeo. En segundo lugar, disminuir los errores de redondeo trabajando con un número mayor de cifras exactas; la mayoría de los lenguajes de programación brindan la posibilidad de trabajar en precisión doble o extendida, lo cual hace más lentos los programas pero utilizan entre 12 y 18 cifras exactas, lo cual es suficiente para que la respuesta obtenida sea correcta.

Existen además algunas técnicas iterativas para mejorar la solución, una de las cuales es la que sigue:

Sea el sistema a resolver

$$\mathbf{Ax} = \mathbf{b} \quad (4)$$

Sea $\mathbf{x}_1$ una solución obtenida por el método de Gauss (o por otro) la cual se sospecha que posee errores debido al mal condicionamiento del sistema. Si $\mathbf{x}_1$ fuera exacta, la diferencia vectorial

$$\mathbf{r}_1 = \mathbf{b} - \mathbf{Ax}_1 \quad (5)$$

sería el vector nulo, de modo que el vector $\mathbf{r}_1$, llamado *residuo* es una manera de medir la discrepancia de una solución aproximada $\mathbf{x}_1$ con la verdadera solución $\mathbf{x}$. El residuo de cualquier solución aproximada se obtiene sin mucho esfuerzo computacional. El proceso que sigue tiene por objetivo disminuir el residuo hasta valores aceptables.

Sea
$$\text{error}(\mathbf{x}_1) = \mathbf{x} - \mathbf{x}_1$$

Sustituyendo en la ecuación (5) $\mathbf{b}$ por $\mathbf{Ax}$ se obtiene:

$$\mathbf{r}_1 = \mathbf{Ax} - \mathbf{Ax}_1 = \mathbf{A}(\mathbf{x} - \mathbf{x}_1)$$

Pero esto significa que
$$\mathbf{A} \text{ error}(\mathbf{x}_1) = \mathbf{r}_1 \quad (6)$$

Resolviendo la ecuación (6), que tiene la misma matriz de coeficientes que el sistema original, se obtiene $\text{error}(\mathbf{x}_1)$ en forma aproximada (debido al mal condicionamiento de $\mathbf{A}$), sin embargo, la solución aproximada:

$$\mathbf{x}_2 = \mathbf{x}_1 + \text{error}(\mathbf{x}_1)$$

está mucho más próxima a la solución exacta. Ahora puede hallarse el residuo de $\mathbf{x}_2$:

$$\mathbf{r}_2 = \mathbf{b} - \mathbf{Ax}_2$$

y el error de $\mathbf{x}_2$ resolviendo el sistema: $\mathbf{A} \text{ error}(\mathbf{x}_2) = \mathbf{r}_2$

y el proceso se continúa hasta que la diferencia entre dos aproximaciones sucesivas $\mathbf{x}_n$ y $\mathbf{x}_{n-1}$ sea suficientemente pequeña.

Ejercicios

1. Para cada uno de los sistemas lineales que siguen, halle el número de condición y analice su mal condicionamiento.

$$\begin{array}{ll}
 28x_1 + 17x_2 + 35x_3 + 29x_4 = 48 & 26x_1 + 15x_2 + 18x_3 + 28x_4 = 37 \\
 \text{a) } 23x_1 + 28x_2 + 32x_3 + 15x_4 = 31 & \text{b) } 45x_1 + 25x_2 + 34x_3 + 53x_4 = 62 \\
 43x_1 + 20x_2 + 29x_3 + 25x_4 = 36 & 42x_1 + 23x_2 + 27x_3 + 45x_4 = 19 \\
 31x_1 + 44x_2 + 18x_3 + 27x_4 = 55 & 28x_1 + 16x_2 + 19x_3 + 30x_4 = 44
 \end{array}$$

2. En los dos sistemas del ejercicio 1, realice pequeños cambios en el vector de términos independientes y analice los cambios que sufre la respuesta. Verifique las conclusiones a las que usted llegó en el ejercicio 1 acerca del mal condicionamiento de cada uno de ellos.
3. Suponga que usted ya tiene un algoritmo que, a partir de una matriz $\mathbf{A}$ cuadrada de orden n y no singular y un vector $\mathbf{b}$ le da como salida la solución del sistema $\mathbf{Ax} = \mathbf{b}$. Utilice este procedimiento para elaborar un algoritmo que calcule el vector de residuos y mejore iterativamente la respuesta hasta que la norma de la diferencia entre dos aproximaciones sucesivas no exceda de una cierta tolerancia ε .
4. La matriz $\mathbf{H}_4$ que se muestra a continuación se llama matriz de Hilbert de orden 4. Este tipo de matriz aparece en muchas cuestiones prácticas. La matriz $\hat{\mathbf{H}}_4$ es una aproximación de $\mathbf{H}_4$ obtenida redondeando sus elementos a tres cifras decimales exactas.

$$\mathbf{H}_4 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix} \quad \hat{\mathbf{H}}_4 = \begin{bmatrix} 1 & 0,5 & 0,333 & 0,25 \\ 0,25 & 0,333 & 0,25 & 0,2 \\ 0,333 & 0,25 & 0,2 & 0,167 \\ 0,25 & 0,2 & 0,167 & 0,143 \end{bmatrix}$$

Al calcular las inversas de estas matrices se obtienen resultados muy diferentes:

$$(\mathbf{H}_4)^{-1} = \begin{bmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & 1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & 1680 & -4200 & 2800 \end{bmatrix} \quad (\hat{\mathbf{H}}_4)^{-1} = \begin{bmatrix} 2,3 & 32,8 & -127,4 & 98,9 \\ 32,8 & -499,9 & 1381,8 & -971,8 \\ -127,4 & 1381,8 & -3314,1 & 2160,5 \\ 98,9 & -971,8 & 2160,5 & -1329,9 \end{bmatrix}$$

Explique a qué se debe este comportamiento.

3.5 Métodos iterativos para sistemas lineales

Del mismo modo que la solución de ecuaciones escalares puede obtenerse por métodos iterativos (bisección, Regula Falsi, Newton – Raphson, Secantes) los sistemas de ecuaciones lineales pueden ser resueltos también por procedimientos de este estilo, lo cual tiene a veces sus ventajas. Una de estas ventajas es la sencillez de estos algoritmos, lo cual significa programas computacionales más simples y seguros. Otra de sus ventajas es la no propagación de errores. En

efecto, como en los métodos iterativos se obtiene una sucesión de soluciones aproximadas que converge hacia la solución exacta, la aproximación número $n - 1$ solo sirve para encontrar una mejor solución en la aproximación número n , por tanto, los errores de la aproximación $n - 1$ no se propagan a la aproximación n .

A pesar de estos dos atractivos, la cuestión más importante al decidirse por uno u otro algoritmo es la eficiencia computacional, medida en tiempo de máquina, y, para sistemas muy voluminosos, la cantidad de memoria necesaria. En este sentido, los métodos iterativos no siempre compiten con el método de Gauss. Sobre este aspecto se volverá más adelante.

La diferencia entre los distintos métodos iterativos existentes está en la forma en que se genera la sucesión de soluciones aproximadas. De ellos se verá el método de Jacobi y el método de Seidel que son dos de los más empleados actualmente.

El método de Jacobi

Sea $\mathbf{Ax} = \mathbf{b}$ un sistema de n ecuaciones lineales. Se supone, como hasta ahora, que este sistema es cuadrado y de solución única. Escrito en forma desarrollada, el sistema es:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \quad (1)$$

Si todos los elementos a_{ii} ($i = 1, 2, \dots, n$) de la diagonal son no nulos, el sistema se puede escribir así:

$$\begin{aligned} x_1 &= \frac{b_1}{a_{11}} - 0x_1 - \frac{a_{12}}{a_{11}}x_2 - \cdots - \frac{a_{1n}}{a_{11}}x_n \\ x_2 &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}}x_1 - 0x_2 - \cdots - \frac{a_{2n}}{a_{22}}x_n \\ &\vdots \\ x_n &= \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}}x_1 - \frac{a_{n2}}{a_{nn}}x_2 - \cdots - 0x_n \end{aligned} \quad (2)$$

donde, como se aprecia, en la i -sima ecuación ($i = 1, 2, \dots, n$) se ha despejado x_i . Utilizando la forma matricial, el sistema puede escribirse:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \cdots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$

Llamando $\mathbf{M}$ y $\mathbf{c}$ a las matrices:

$$\mathbf{M} = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$

el sistema toma la forma

$$\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{c} \quad (3)$$

A partir de la ecuación (3) se puede generar un proceso iterativo matricial, partiendo de un vector inicial $\mathbf{x}^{(0)}$ y generando una sucesión de vectores $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \dots$ mediante la ecuación recursiva:

$$\mathbf{x}^{(k)} = \mathbf{M}\mathbf{x}^{(k-1)} + \mathbf{c} \quad k = 1, 2, 3, \dots \quad (4)$$

La sucesión generada puede o no converger pero, cuando lo hace, su límite es la solución del sistema (3) o (2) o (1) que son todos equivalentes. Este algoritmo recibe el nombre de método de Jacobi.

Ejemplo 1

Escriba los siguientes sistemas lineales en la forma $\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{c}$ y aplíqueles el algoritmo de Jacobi.

$$\begin{array}{lll} 10x_1 - x_2 + 2x_3 = 6 & 10x_1 + 8x_2 - 7x_3 = -3 & 4x_1 - 3x_2 + 5x_3 = 25 \\ \text{a) } x_1 - 5x_2 + x_3 = -10 & \text{b) } -3x_1 + 10x_2 + x_3 = -25 & \text{c) } 3x_1 + 2x_2 + 2x_3 = 11 \\ 2x_1 - x_2 + 8x_3 = -8 & 5x_1 - x_2 + 10x_3 = 22 & 4x_1 - 2x_2 + 3x_3 = -4 \end{array}$$

Solución:

a) Despejando x_1 en la primera ecuación, x_2 en la segunda y x_3 en la tercera, se tiene:

$$\begin{aligned} x_1 &= 0x_1 + 0,1x_2 - 0,2x_3 + 0,6 \\ x_2 &= 0,2x_1 + 0x_2 + 0,2x_3 + 2 \\ x_3 &= -0,25x_1 + 0,125x_2 + 0x_3 - 1 \end{aligned}$$

que en forma matricial resulta:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 0,1 & -0,2 \\ 0,2 & 0 & 0,2 \\ -0,25 & 0,125 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0,6 \\ 2 \\ -1 \end{bmatrix}$$

Tomando $\mathbf{x}^{(0)}$ como el vector nulo y generando la sucesión de vectores $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \dots$ mediante la ecuación recursiva:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^{(k)} = \begin{bmatrix} 0 & 0,1 & -0,2 \\ 0,2 & 0 & 0,2 \\ -0,25 & 0,125 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^{(k-1)} + \begin{bmatrix} 0,6 \\ 2 \\ -1 \end{bmatrix} \quad k = 1, 2, 3, \dots$$

se obtienen los resultados que se muestran en la tabla 1. Resulta evidente que en este caso el método de Jacobi converge hacia la solución del sistema de ecuaciones, que es $x_1 = 1$, $x_2 = 2$ y $x_3 = -1$.

Iteración	x_1	x_2	x_3
0	0,0	0,0	0,0
1	0,6	2,0	-1,0
2	1,0	1,92	-0,9
3	0,972	2,02	-1,01
4	1,004	1,9924	-0,9905
5	0,99734	2,0027	-1,00195
6	1,00066	1,999078	-0,998997
7	0,999707	2,000333	-1,00028
8	1,000089	1,999885	-0,999885
9	0,999966	2,000041	-1,000037

Tabla 1

- b) De manera similar al anterior, este sistema queda expresado en la forma $\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{c}$ como sigue:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -0,8 & 0,7 \\ 0,3 & 0 & -0,1 \\ -0,5 & 0,1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} -0,3 \\ -2,5 \\ 2,2 \end{bmatrix}$$

Al ser aplicado el algoritmo de Jacobi se obtienen los resultados que muestra la tabla 2. Por razones de espacio no se muestran todas las filas de la tabla. Como se observará, el método en este caso también converge hacia la solución del sistema, que es $x_1 = 2$, $x_2 = -2$ y $x_3 = 1$, pero ahora la convergencia se produce más lentamente: en el sistema a) se había logrado cuatro cifras decimales exactas en la novena iteración, sin embargo, en este con 40 iteraciones aun no se ha logrado esa exactitud. Más adelante se comprenderá la causa de este comportamiento.

Iteración	x_1	x_2	x_3
0	0,0	0,0	0,0
1	-0,3	-2,5	2,2
2	3,24	-2,81	2,1
3	3,418	-1,738	0,299
4	1,2997	-1,5045	0,3172
5	1,12564	-2,14181	1,3997
$\vdots$	$\vdots$	$\vdots$	$\vdots$
10	2,121377	-2,111874	1,161146
$\vdots$	$\vdots$	$\vdots$	$\vdots$
20	1,994785	-1,995258	0,986102
$\vdots$	$\vdots$	$\vdots$	$\vdots$
30	2,000055	-2,000701	1,001144
$\vdots$	$\vdots$	$\vdots$	$\vdots$
40	2,000024	-1,999948	0,99991

Tabla 2

c) El sistema queda expresado en la forma $\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{c}$ tal como se hizo en los casos anteriores:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 0,75 & -1,25 \\ -1,5 & 0 & -1 \\ -1,3333 & 0,6667 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 6,25 \\ 5,5 \\ -1,3333 \end{bmatrix}$$

Los resultados obtenidos al aplicar el algoritmo de Jacobi se muestran en la tabla 3. Obsérvese como aquí el proceso iterativo no converge hacia la solución del sistema, que es $x_1 = 3$, $x_2 = -1$ y $x_3 = 2$. En las páginas que sigue se hará claro el motivo de este comportamiento.

Iteración	x_1	x_2	x_3
0	0,0	0,0	0,0
1	6,25	5,5	-1,3333
2	12,0417	-2,5417	10,6667
3	-8,9896	-23,2292	13,0298
4	-27,4566	5,9566	-28,8056
5	46,7244	75,4905	-33,9711
6	105,3317	-30,6155	111,2928
7	-155,8277	-263,7907	118,6985

Tabla 3

Convergencia del método de Jacobi

Como se ha visto, no siempre el método de Jacobi da lugar a una sucesión convergente hacia la solución del sistema. En lo que sigue se establecerán algunas condiciones de uso práctico que garanticen la obtención de la solución.

Para estudiar la convergencia del método de Jacobi (y, más adelante, el de Seidel) es necesario definir la forma en que se medirá el error en la k -sima aproximación $\mathbf{x}^{(k)}$, respecto al vector solución $\mathbf{x}$.

Definición 1

Si $\mathbf{x}^{(k)}$ denota la k -sima aproximación de un proceso iterativo a la solución $\mathbf{x}$ de un sistema lineal, entonces el error absoluto de $\mathbf{x}^{(k)}$ se denota por $E(\mathbf{x}^{(k)})$ y se define como:

$$E(\mathbf{x}^{(k)}) = \|\mathbf{x} - \mathbf{x}^{(k)}\| \quad \blacksquare$$

Nótese que, según los axiomas de las normas matriciales, el error absoluto es cero cuanto los vectores $\mathbf{x}$ y $\mathbf{x}^{(k)}$ coinciden en todas sus componentes, es decir, cuando su diferencia es el vector nulo.

Ejemplo 2

En el sistema lineal del ejemplo 1 a) muestre $\mathbf{x}$, $\mathbf{x}^{(4)}$ y $E(\mathbf{x}^{(4)})$

Solución:

La solución $\mathbf{x}$ del sistema es $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$

La aproximación $\mathbf{x}^{(4)}$ obtenida en la cuarta iteración es $\mathbf{x}^{(4)} = \begin{bmatrix} 1.004 \\ 1.9924 \\ -0.9905 \end{bmatrix}$

El error absoluto de $\mathbf{x}^{(4)}$, de acuerdo con la definición que se acaba de dar es:

$$E(\mathbf{x}^{(4)}) = \|\mathbf{x} - \mathbf{x}^{(4)}\| = \left\| \begin{bmatrix} 1 - 1.004 \\ 2 - 1.9924 \\ -1 - (-0.9905) \end{bmatrix} \right\| = \left\| \begin{bmatrix} -0.004 \\ 0.0076 \\ -0.0095 \end{bmatrix} \right\| = \max\{0.004; 0.0076; 0.0095\}$$

$$E(\mathbf{x}^{(4)}) = 0.0095 \quad \blacksquare$$

Considérese ahora el sistema de ecuaciones cuya solución se desea hallar, escrito como en la ecuación (3):

$$\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{c}$$

Si de ella se resta miembro a miembro la ecuación recursiva del método de Jacobi (4)

$$\mathbf{x}^{(k)} = \mathbf{M}\mathbf{x}^{(k-1)} + \mathbf{c} \quad k = 1, 2, 3, \dots$$

Se obtiene:

$$\mathbf{x} - \mathbf{x}^{(k)} = \mathbf{M}(\mathbf{x} - \mathbf{x}^{(k-1)}) \quad k = 1, 2, 3, \dots$$

Tomando norma en cada miembro y aplicando el axioma 6 de las normas matriciales, se obtiene:

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \|\mathbf{M}\| \cdot \|\mathbf{x} - \mathbf{x}^{(k-1)}\|$$

Si se tiene en cuenta la definición de error absoluto:

$$E(\mathbf{x}^{(k)}) \leq \|\mathbf{M}\| \cdot E(\mathbf{x}^{(k-1)}) \quad k = 1, 2, 3, \dots \quad (5)$$

La norma de $\mathbf{M}$ juega un papel muy importante en la convergencia del algoritmo de Jacobi, por lo cual se le dará un nombre y una notación especiales.

Definición 2

Sea el sistema lineal $\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{c}$. Se llama *factor de convergencia* del método de Jacobi para este sistema a la norma de $\mathbf{M}$ y se denotará como α , es decir:

$$\alpha = \|\mathbf{M}\| \quad \blacksquare$$

De acuerdo con esta definición, la ecuación (5) toma la forma más simple:

$$E(\mathbf{x}^{(k)}) \leq \alpha E(\mathbf{x}^{(k-1)}) \quad k = 1, 2, 3, \dots \quad (6)$$

De la ecuación (6) se obtiene una consecuencia muy importante relacionada con la convergencia del algoritmo de Jacobi.

Teorema 1

Una condición suficiente para que el método de Jacobi converja hacia la solución del sistema $\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{c}$ independientemente de la aproximación inicial $\mathbf{x}^{(0)}$ es que el factor de convergencia α , sea menor que 1.

Demostración:

Aplicando k veces la desigualdad (6):

$$E(\mathbf{x}^{(k)}) \leq \alpha E(\mathbf{x}^{(k-1)}) \leq \alpha^2 E(\mathbf{x}^{(k-2)}) \leq \dots \leq \alpha^k E(\mathbf{x}^{(0)})$$

Como $E(\mathbf{x}^{(0)})$ no depende de k , tomando límites para k tendiendo hacia infinito, resulta:

$$\lim_{k \rightarrow \infty} E(\mathbf{x}^{(k)}) \leq E(\mathbf{x}^{(0)}) \lim_{k \rightarrow \infty} \alpha^k = 0$$

por ser $\alpha < 1$. Esto significa que $\lim_{k \rightarrow \infty} E(\mathbf{x}^{(k)}) = 0$ de donde resulta que

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$$

como se quería demostrar. ■

De la ecuación (6) también se puede ver con claridad que la rapidez de convergencia del método de Jacobi depende del valor del parámetro α de convergencia. En efecto, la ecuación (6) se puede escribir en términos de error absoluto máximo como:

$$E_m(\mathbf{x}^{(k)}) = \alpha E_m(\mathbf{x}^{(k-1)})$$

de manera que se trata de un método de convergencia lineal. Valores de α próximos a cero garantizan una alta rapidez en la convergencia, valores menores que 1 pero próximos a él, producen una convergencia lenta. Nótese que el teorema 1 establece una condición suficiente para la convergencia; esta condición no es necesaria y, de hecho, se pueden hallar sistemas lineales con $\alpha > 1$ para los cuales el método de Jacobi converge. No es difícil demostrar que una condición necesaria y suficiente para la convergencia es que el mayor valor propio de la matriz $\mathbf{M}$ sea menor que 1, sin embargo esta condición no posee gran valor práctico dado que hallar el mayor valor propio de $\mathbf{M}$ es un problema, en general, más complicado que resolver el sistema lineal.

Ejemplo 3

Halle el factor de convergencia α de cada uno de los sistemas del ejemplo 1 y explique la convergencia del método de Jacobi en cada uno de ellos.

Solución:

a) La matriz $\mathbf{M}$ del sistema es

$$\mathbf{M} = \begin{bmatrix} 0 & 0,1 & -0,2 \\ 0,2 & 0 & 0,2 \\ -0,25 & 0,125 & 0 \end{bmatrix}$$

su factor de convergencia se obtiene como:

$$\alpha = \|\mathbf{M}\| = \max\{0,3; 0,4; 0,375\} = 0,4$$

Como se ve, es próximo a cero. En este caso, el error absoluto máximo en cada iteración será menos de la mitad del de la iteración anterior y se obtendrá una alta velocidad de convergencia.

b) En este caso

$$\mathbf{M} = \begin{bmatrix} 0 & -0,8 & 0,7 \\ 0,3 & 0 & -0,1 \\ -0,5 & 0,1 & 0 \end{bmatrix}$$

y el parámetro de convergencia: $\alpha = \|\mathbf{M}\| = \max\{1,5; 0,4; 0,6\} = 1,5$

Como $\alpha > 1$, la convergencia no se puede garantizar por el teorema 1. En este caso el método de Jacobi converge, aunque la convergencia, como ya se había visto, es bastante lenta.

c) Para este sistema:

$$\mathbf{M} = \begin{bmatrix} 0 & 0,75 & -1,25 \\ -1,5 & 0 & -1 \\ -1,3333 & 0,6667 & 0 \end{bmatrix}$$

El parámetro de convergencia es: $\alpha = \|\mathbf{M}\| = \max\{2; 2,5; 2\} = 2,5$

Como $\alpha > 1$, el teorema 1 no garantiza que exista convergencia. Como se recordará del ejemplo 1, en este caso el método de Jacobi diverge. ■

El teorema que sigue es una consecuencia inmediata del anterior. Su importancia radica en que permite analizar la convergencia del método de Jacobi para un sistema lineal cuando este aun se encuentra en la forma $\mathbf{Ax} = \mathbf{b}$.

Teorema 2

Sea el sistema lineal de n ecuaciones con n incógnitas $\mathbf{Ax} = \mathbf{b}$. Una condición suficiente para que el método de Jacobi aplicado a dicho sistema, sea convergente, es que $\mathbf{A}$ tenga la diagonal predominante, esto es, que para cada fila $i = 1, 2, 3, \dots, n$, el elemento de la diagonal sea, en valor absoluto, mayor que la suma de los valores absolutos de los otros elementos. Es decir que:

$$|a_{ii}| > |a_{i1}| + |a_{i2}| + \dots + |a_{i,i-1}| + |a_{i,i+1}| + \dots + |a_{in}| \quad \text{para } i = 1, 2, 3, \dots, n \quad (7)$$

Demostración:

Como

$$\mathbf{M} = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{bmatrix}$$

entonces, si i es una fila cualquiera de $\mathbf{A}$ la desigualdad (7) implica, dividiendo en cada uno de sus términos por $|a_{ii}|$, que:

$$\frac{|a_{i1}|}{|a_{ii}|} + \frac{|a_{i2}|}{|a_{ii}|} + \dots + \frac{|a_{i,i-1}|}{|a_{ii}|} + \frac{|a_{i,i+1}|}{|a_{ii}|} + \dots + \frac{|a_{in}|}{|a_{ii}|} < 1$$

lo cual significa que la suma de los valores absolutos de los elementos de la i -sima fila de $\mathbf{M}$ es menor que 1. Como esto sucede para todas las filas de $\mathbf{M}$ entonces la norma de $\mathbf{M}$ será necesariamente menor que 1 y el algoritmo de Jacobi converge, según el teorema 1. ■

El error en el método de Jacobi

Al igual que en todos los demás métodos iterativos, se necesita una forma de acotar el error en cada iteración del método de Jacobi para poder detener el proceso iterativo cuando el error sea suficientemente pequeño. Para ello recuérdese que se definió:

$$E(\mathbf{x}^{(k)}) = \|\mathbf{x} - \mathbf{x}^{(k)}\|$$

por tanto,

$$E(\mathbf{x}^{(k-1)}) = \|\mathbf{x} - \mathbf{x}^{(k-1)}\|$$

Sumando y restando $\mathbf{x}^{(k)}$ a la diferencia del segundo miembro, se obtiene:

$$E(\mathbf{x}^{(k-1)}) = \|\mathbf{x} - \mathbf{x}^{(k)} + \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

Pero, según el axioma 3 de las normas, la norma de la suma de dos vectores es menor o igual a la suma de las normas de esos vectores, así que:

$$E(\mathbf{x}^{(k-1)}) \leq \|\mathbf{x} - \mathbf{x}^{(k)}\| + \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

El primer sumando del segundo miembro no es más que $E(\mathbf{x}^{(k)})$. Esto conduce a:

$$E(\mathbf{x}^{(k-1)}) \leq E(\mathbf{x}^{(k)}) + \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \quad (8)$$

El primer miembro de esta inecuación se puede modificar teniendo en cuenta la desigualdad (6),

$$E(\mathbf{x}^{(k)}) \leq \alpha E(\mathbf{x}^{(k-1)}) \text{ que implica } \frac{E(\mathbf{x}^{(k)})}{\alpha} \leq E(\mathbf{x}^{(k-1)})$$

Entonces la expresión (8) queda:

$$\frac{E(\mathbf{x}^{(k)})}{\alpha} \leq E(\mathbf{x}^{(k)}) + \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

Despejando de esta inecuación el error absoluto de $\mathbf{x}^{(k)}$, para lo cual se supondrá que $\alpha < 1$:

$$E(\mathbf{x}^{(k)}) \left(\frac{1}{\alpha} - 1 \right) \leq \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

$$E(\mathbf{x}^{(k)}) \leq \left(\frac{\alpha}{1 - \alpha} \right) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

Esto significa que el error absoluto máximo puede tomarse como el miembro de la derecha de esta desigualdad, es decir, que:

$$E_m(\mathbf{x}^{(k)}) = \left(\frac{\alpha}{1-\alpha} \right) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \quad (9)$$

La fórmula (9) permite establecer una condición práctica para la terminación del proceso iterativo de Jacobi.

Condición de terminación:

Si se desea obtener la solución de un sistema lineal con un error absoluto menor que ε , y el factor de convergencia del método de Jacobi es $\alpha < 1$, entonces el proceso iterativo de Jacobi se llevará a cabo hasta la aproximación $\mathbf{x}^{(k)}$ para la cual:

$$E_m(\mathbf{x}^{(k)}) = \left(\frac{\alpha}{1-\alpha} \right) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq \varepsilon$$

En los casos en que α es menor que 0,5, la condición de terminación se puede simplificar un poco. En efecto, si $\alpha < 0,5$:

$$\frac{\alpha}{1-\alpha} < \frac{0,5}{1-\alpha} < \frac{0,5}{1-0,5} = 1$$

y entonces:

$$\left(\frac{\alpha}{1-\alpha} \right) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

y puede tomarse como error absoluto máximo:

$$E_m(\mathbf{x}^{(k)}) = \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

De aquí resulta:

Condición de terminación:

Si se desea obtener la solución de un sistema lineal con un error absoluto menor que ε , y el factor de convergencia del método de Jacobi es $\alpha < 0,5$, entonces el proceso iterativo de Jacobi se llevará a cabo hasta la aproximación $\mathbf{x}^{(k)}$ para la cual:

$$E_m(\mathbf{x}^{(k)}) = \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq \varepsilon$$

Ejemplo 4

Complete la tabla 1 del ejemplo 1 a) con una columna que muestre el error absoluto máximo para cada iteración. Indique en que iteración debe detenerse el proceso si se desea obtener la solución con 3 cifras decimales exactas.

Solución:

En el ejemplo 3 se determinó que el factor de convergencia para este sistema es $\alpha = 0,4$. Por tanto, para cada iteración, el error absoluto máximo se puede tomar como:

$$E_m(\mathbf{x}^{(k)}) = \left(\frac{\alpha}{1-\alpha} \right) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| = \frac{0,4}{1-0,4} \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| = (0,67) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

En cada iteración, $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$ se obtiene como el mayor valor absoluto de las componentes del vector diferencia entre la iteración actual y la anterior. En la tabla 4 se muestra el error absoluto máximo para cada iteración. En cada iteración se ha marcado con * la componente donde ocurre la máxima diferencia con la iteración anterior.

Iteración	x_1	x_2	x_3	$E_m(\mathbf{x}^{(k)})$
0	0,0	0,0	0,0	-----
1	0,6	2,0*	- 1,0	1,34
2	1,0*	1,92	- 0,9	0,27
3	0,972	2,02	- 1,01*	0,074
4	1,004*	1,9924	- 0,9905	0,021
5	0,99734	2,0027	- 1,00195*	0,0077
6	1,00066	1,999078*	- 0,998997	0,0024
7	0,999707	2,000333	- 1,00028*	0,00086
8	1,000089	1,999885*	- 0,999885	0,00030
9	0,999966	2,000041*	- 1,000037	0,00010

Tabla 4

Para obtener 3 cifras decimales exactas el error absoluto debe ser menor que 0,0005, por tanto, habría que tomar como solución la iteración 8.

Algoritmo del método de Jacobi

Se desea hallar una solución del sistema $\mathbf{Ax} = \mathbf{b}$ con error absoluto menor que ε . Se supone que $\mathbf{A}$ tiene diagonal predominante y se conoce el valor del factor de convergencia α . El algoritmo utiliza como datos las matrices $\mathbf{A}$ y $\mathbf{b}$, el factor de convergencia α , la tolerancia ε y la aproximación inicial $\mathbf{x}^{(0)}$. De no conocerse $\mathbf{x}^{(0)}$ se puede tomar el vector nulo $\mathbf{0}$.

xv := **x**⁽⁰⁾ {El vector **xv** representa la solución vieja, es decir, la aproximación obtenida en la iteración anterior, el vector **xa** representa la solución nueva, es decir, la aproximación obtenida en la iteración actual}

repeat

Error := 0

for *i* = 1 **to** *n*

$xa_i := b_i$ {En esta sección se calcula la *i*-sima componente de **xa**}

for *j* := 1 **to** *n*

if *j* ≠ *i* **then**

$xa_i := xa_i - a_{ij}xv_j$

end

$xa_i := \frac{xa_i}{a_{ii}}$

end

if $|xa_i - xv_i| > Error$ **then** {En este lazo se va guardando la máxima diferencia entre las dos ultimas iteraciones}

$Error := |xa_i - xv_i|$

end

end

xv := **xa** {La solución obtenida se transfiere a la vieja}

$Error := Error \cdot \left(\frac{\alpha}{1 - \alpha} \right)$

until *Error* < ε

La solución del sistema es **xa** con error absoluto menor que *Error*

Terminar

Ejemplo 5

Resuelva con cuatro cifras decimales exactas el siguiente sistema de ecuaciones mediante el método de Jacobi.

$$9x_1 - x_2 + 2x_3 = 9$$

$$x_1 + 8x_2 + 2x_3 = 19$$

$$x_1 - x_2 + 11x_3 = 10$$

Solución:

Como la diagonal es predominante, se puede asegurar la convergencia del método de Jacobi. Para obtener el factor de convergencia se requiere hallar la matriz **M**:

$$\mathbf{M} = \begin{bmatrix} 0 & \frac{1}{9} & -\frac{2}{9} \\ -\frac{1}{8} & 0 & -\frac{1}{4} \\ -\frac{1}{11} & \frac{1}{11} & 0 \end{bmatrix}$$

de donde:

$$\alpha = \max \left\{ \frac{1}{3}, \frac{3}{8}, \frac{2}{11} \right\} = \frac{3}{8} = 0,375$$

Por ser $\alpha < 0,5$, se puede tomar el error absoluto máximo como la norma de $\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}$. Esto es, el proceso iterativo será detenido cuando

$$\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq 0,00005$$

La tabla 5 muestra las aproximaciones obtenidas y el error absoluto máximo de cada una. Aparecen marcadas con * las componentes de la solución donde ocurre la máxima diferencia con la iteración anterior.

Iteración	x_1	x_2	x_3	$E_m(\mathbf{x}^{(k)})$
0	0,0	0,0	0,0	-----
1	1,0	2,375*	0,909091	2,375
2	1,061869	2,022727*	1,034091	0,352273
3	0,994949*	1,983744	0,996442	0,066919
4	0,998984	2,001521*	0,998981	0,017777
5	1,000395*	2,000382	1,000231	0,001411
6	0,999991	1,999893*	0,999999	0,000489
7	0,999988	2,000001*	0,999991	0,000108
8	1.000002*	2.000004	1.000001	0.000014

Tabla 5

La solución aproximada, con cuatro cifras decimales exactas, es:

$$\begin{aligned}x_1 &= 1,000002 \\x_2 &= 2,000004 \\x_3 &= 1,000001\end{aligned}$$

Ejemplo 6

Se desea obtener la solución del siguiente sistema tridiagonal con 100 ecuaciones y 100 incógnitas:

$$\begin{aligned}8x_1 - x_2 &= 10 \\x_1 + 8x_2 - x_3 &= 11 \\x_2 + 8x_3 - x_4 &= 12 \\x_3 + 8x_4 - x_5 &= 13 \\&\vdots \\x_{98} + 8x_{99} - x_{100} &= 108 \\x_{99} + 8x_{100} &= 109\end{aligned}$$

Realice un algoritmo en pseudo código para hallar la solución mediante el método de Jacobi con 5 cifras decimales exactas. Este ejemplo ilustra una de las ventajas más importantes de los métodos iterativos: grandes sistemas de ecuaciones donde todas las ecuaciones poseen una misma estructura general.

Solución:

Se trata de un sistema con diagonal predominante. Es más, el factor de convergencia se puede calcular rápidamente como $\alpha = 0,25$. Salvo la primera y la última ecuación, todas presentan una estructura similar:

$$x_{i-1} + 8x_i - x_{i+1} = i + 9 \quad i = 2, 3, 4, \dots, 99$$

Despejando x_i en cada una de las ecuaciones, el sistema sería:

$$\begin{aligned} x_1 &= \frac{10 + x_2}{8} \\ x_i &= \frac{i + 9 - x_{i-1} + x_{i+1}}{8} \quad \text{para } i = 2, 3, 4, \dots, 99 \\ x_{100} &= \frac{109 - x_{99}}{8} \end{aligned}$$

Por ser $\alpha = 0,25$, se tomará el error absoluto máximo de cada aproximación como:

$$E_m(\mathbf{x}^{(k)}) = \left(\frac{0,25}{1 - 0,25} \right) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| = (0,34) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

Como la estructura del sistema de ecuaciones es especial, es ventajoso utilizar este hecho para no tener que guardar en memoria las matrices **A** y **b**. Por esta razón el algoritmo que se usará no es el mismo que el algoritmo general elaborado antes. El algoritmo en pseudo código es:

```

xv := x(0)
repeat
   $xa_1 = \frac{10 + xv_2}{8}$ 
   $Error := |xa_1 - xv_1|$ 
  for  $i = 2$  to 99
     $xa_i = \frac{i + 9 - xv_{i-1} + xv_{i+1}}{8}$ 
    if  $|xa_i - xv_i| > Error$  then
       $Error := |xa_i - xv_i|$ 
    end
  end
   $xa_{100} = \frac{109 - xv_{99}}{8}$ 
  if  $|xa_{100} - xv_{100}| > Error$  then
     $Error := |xa_{100} - xv_{100}|$ 
  end
  xv := xa {La solución obtenida se transfiere a la vieja}
   $Error := (0,34)Error$ 
until  $Error < 0,000005$ 
La solución del sistema es xa con error absoluto menor que  $Error$ 
Terminar

```

El método de Seidel

El método de Seidel, también llamado de Gauss – Seidel, es una variación del método de Jacobi que logra simplificar dicho algoritmo y mejorar la rapidez de la convergencia en la mayoría de los casos.

En el método de Jacobi, al calcular la variable x_i se utilizan los valores de las demás variables que se obtuvieron en la iteración anterior, sin embargo, en ese momento ya se han calculado los nuevos valores de $x_1, x_2, \dots, x_{i-1}$ que son, por lo general, mejores aproximaciones que los obtenidos en la iteración anterior. En el método de Seidel, una vez que se obtiene el nuevo valor de una variable, éste se utiliza para actualizar los valores de las variables que siguen; de esta forma, no se necesita guardar los valores de la iteración anterior, lo cual simplifica el algoritmo y ahorra memoria y, por otra parte, la velocidad de la convergencia mejora sustancialmente.

En términos más precisos, sea $\mathbf{Ax} = \mathbf{b}$ un sistema de n ecuaciones lineales. Se supone, como hasta ahora, que este sistema es cuadrado y que posee diagonal predominante. Escrito en forma desarrollada, el sistema es:

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\&\vdots \\a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n\end{aligned}$$

Como los elementos a_{ii} ($i = 1, 2, \dots, n$) de la diagonal son no nulos, el sistema, igual que antes, se puede escribir así:

$$\begin{aligned}x_1 &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}}x_2 - \dots - \frac{a_{1n}}{a_{11}}x_n \\x_2 &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}}x_1 - \dots - \frac{a_{2n}}{a_{22}}x_n \\&\vdots \\x_n &= \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}}x_1 - \frac{a_{n2}}{a_{nn}}x_2 - \dots - \frac{a_{nn}}{a_{nn}}x_n\end{aligned}$$

El proceso iterativo de Seidel queda entonces definido de la siguiente forma:

$$\begin{aligned}x_1^{(k)} &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}}x_2^{(k-1)} - \frac{a_{13}}{a_{11}}x_3^{(k-1)} - \dots - \frac{a_{1n}}{a_{11}}x_n^{(k-1)} \\x_2^{(k)} &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}}x_1^{(k)} - \frac{a_{23}}{a_{22}}x_3^{(k-1)} - \dots - \frac{a_{2n}}{a_{22}}x_n^{(k-1)} \\x_3^{(k)} &= \frac{b_3}{a_{33}} - \frac{a_{31}}{a_{33}}x_1^{(k)} - \frac{a_{32}}{a_{33}}x_2^{(k)} - \dots - \frac{a_{3n}}{a_{33}}x_n^{(k-1)} \\&\vdots \\x_n^{(k)} &= \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}}x_1^{(k)} - \frac{a_{n2}}{a_{nn}}x_2^{(k)} - \dots - \frac{a_{n,n-1}}{a_{nn}}x_{n-1}^{(k)}\end{aligned} \tag{10}$$

Aunque el método de Seidel se puede expresar en forma matricial, no se gana mucho en este caso, pues las matrices que aparecen ya no son tan simples.

Ejemplo 7

Los siguientes sistemas de ecuaciones fueron resueltos por el método de Jacobi en el ejemplo 1. Revuélvalos por el método de Seidel.

$$\begin{array}{lll} 10x_1 - x_2 + 2x_3 = 6 & 10x_1 + 8x_2 - 7x_3 = -3 & 4x_1 - 3x_2 + 5x_3 = 25 \\ \text{a) } x_1 - 5x_2 + x_3 = -10 & \text{b) } -3x_1 + 10x_2 + x_3 = -25 & \text{c) } 3x_1 + 2x_2 + 2x_3 = 11 \\ 2x_1 - x_2 + 8x_3 = -8 & 5x_1 - x_2 + 10x_3 = 22 & 4x_1 - 2x_2 + 3x_3 = -4 \end{array}$$

Solución:

Aplicando el algoritmo de Seidel se obtiene los resultados que muestran respectivamente las tablas 6, 7 y 8 para los sistemas a), b) y c). Al igual que sucedió en el ejemplo 1 para el algoritmo de Jacobi, el proceso iterativo converge en los sistemas a) y b) y diverge en el c).

Iteración	x_1	x_2	x_3
0	0,0	0,0	0,0
1	0,6	2,12	-0,885
2	0,989	2,0208	-0,99465
3	1,00101	2,001272	-1,000094
4	1,000146	2,000010	-1,000035
5	1.000008	1.999995	-1.000003

Tabla 6

Iteración	x_1	x_2	x_3
0	0,0	0,0	0,0
1	-0,3	-2,59	2,091
2	3,2357	-1,73839	0,408311
3	1,37653	-2,127872	1,298948
4	2,311561	-1,936426	0,850577
5	1,844545	-2,031694	1,074558
⋮	⋮	⋮	⋮
10	2,004801	-1,999021	0,997698
⋮	⋮	⋮	⋮
15	1,999852	-2,000074	1,000071
16	2,000074	-1,999985	0,999965
17	1,999963	-2,000008	1,000018

Tabla 7

Iteración	x_1	x_2	x_3
0	0,0	0,0	0,0
1	6,25	-3,875	4,4167
2	-2,1771	4,3490	-1,3368
3	11,1827	-9,9373	6,9521
4	-9,8931	13,3875	-5,5991
5	23,2895	-23,8352	13,8293
6	-28,9130	35,0402	-16,5238
7	53,1849	-57,7536	31,0775
8	-75,9121	88,2906	-43,6890

Tabla 8

Nótese, sin embargo, que el método de Seidel ha requerido menos iteraciones para llegar a la solución de a) y b) con una exactitud similar.

Convergencia del método de Seidel

Considérese la i -sima ecuación del algoritmo de Seidel:

$$x_i^{(k)} = \frac{b_i}{a_{ii}} - \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(k)} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j^{(k-1)} \quad i = 1, 2, \dots, n \quad (11)$$

La i -sima ecuación del sistema se puede escribir, agrupando convenientemente, como:

$$x_i = \frac{b_i}{a_{ii}} - \sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j \quad i = 1, 2, \dots, n \quad (12)$$

Restando miembro a miembro las ecuaciones (12) menos (11), se obtiene:

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

Tomando módulos en ambos miembros de esta igualdad y recordando que el módulo de una suma es menor o igual que la suma de los módulos, resulta:

$$\left| x_i - x_i^{(k)} \right| \leq \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \cdot \left| x_j - x_j^{(k)} \right| + \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \cdot \left| x_j - x_j^{(k-1)} \right| \quad (13)$$

Si en la ecuación (13) los términos $\left| x_j - x_j^{(k)} \right|$ y $\left| x_j - x_j^{(k-1)} \right|$ se sustituyen respectivamente por $\left\| \mathbf{x} - \mathbf{x}^{(k)} \right\| = E(\mathbf{x}^{(k)})$ y $\left\| \mathbf{x} - \mathbf{x}^{(k-1)} \right\| = E(\mathbf{x}^{(k-1)})$ la desigualdad se cumple con más razón pues las normas son los máximos valores que toman $\left| x_j - x_j^{(k)} \right|$ y $\left| x_j - x_j^{(k-1)} \right|$ para las diferentes j . Se obtiene:

$$\left| x_i - x_i^{(k)} \right| \leq \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \cdot E(\mathbf{x}^{(k)}) + \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \cdot E(\mathbf{x}^{(k-1)}) \quad (14)$$

Los errores absolutos, como son independientes de j se pueden extraer como factores comunes en las sumas respectivas, lo cual conduce a:

$$\left| x_i - x_i^{(k)} \right| \leq E(\mathbf{x}^{(k)}) \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| + E(\mathbf{x}^{(k-1)}) \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \quad (15)$$

Conviene ahora definir, $p_i = \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right|$ y $q_i = \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right|$ $i = 1, 2, \dots, n$ (16)

Obsérvese que p_i es la suma de los módulos de los elementos de la fila i de la matriz $\mathbf{M}$, sumando desde el primer elemento hasta el que antecede la diagonal, mientras que q_i es la suma desde el elemento después de la diagonal hasta el final de la fila. La desigualdad (15) queda ahora:

$$\left| x_i - x_i^{(k)} \right| \leq E(\mathbf{x}^{(k)}) p_i + E(\mathbf{x}^{(k-1)}) q_i \quad i = 1, 2, \dots, n \quad (17)$$

Como la desigualdad (17) es cierta para todas las ecuaciones $i = 1, 2, \dots, n$ del sistema, será también cierta, en particular, para aquella, llamémosla m , en la cual es máximo $\left| x_i^{(k)} - x_i \right|$.

Entonces para $i = m$ el primer miembro de la desigualdad es $E(\mathbf{x}^{(k)})$ y se tiene:

$$E(\mathbf{x}^{(k)}) \leq E(\mathbf{x}^{(k)}) p_m + E(\mathbf{x}^{(k-1)}) q_m$$

Despejando $E(\mathbf{x}^{(k)})$: $E(\mathbf{x}^{(k)}) \leq \left(\frac{q_m}{1 - p_m} \right) E(\mathbf{x}^{(k-1)})$ (18)

supuesto que $p_m < 1$, lo cual está garantizado, por ejemplo, si el sistema tiene diagonal predominante.

El valor de m puede variar de una iteración a otra. Sin embargo, si se define:

$$\beta = \max_i \frac{q_i}{1 - p_i}$$

la desigualdad (18) conduce a:

$$E(\mathbf{x}^{(k)}) \leq \beta \cdot E(\mathbf{x}^{(k-1)}) \quad (19)$$

Si se supone que la diagonal de $\mathbf{A}$ es predominante, entonces

$$p_i + q_i < 1 \quad i = 1, 2, 3, \dots, n$$

esto es:

$$q_i < 1 - p_i$$

$$y \quad \frac{q_i}{1-p_i} < 1 \quad i = 1, 2, 3, \dots, n$$

en ese caso el parámetro $\beta = \max_i \frac{q_i}{1-p_i}$ será menor que 1, lo cual prueba que el proceso iterativo converge hacia la solución del sistema.

El parámetro β juega en el método de Seidel un papel análogo al α del método de Jacobi. Por su importancia, la siguiente definición formaliza este concepto.

Definición 3

Sea el sistema lineal $\mathbf{Ax} = \mathbf{b}$ con diagonal predominante. Se llama *factor de convergencia* del método de Seidel para este sistema al número β definido como:

$$\beta = \max_i \frac{q_i}{1-p_i}$$

$$\text{donde } p_i = \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \quad \text{y} \quad q_i = \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \quad i = 1, 2, \dots, n \quad \blacksquare$$

Como las fórmulas (19) para el método de Seidel y (6) para el método de Jacobi tienen idéntica estructura con el único cambio de α por β , no es necesario repetir para el método de Seidel todas las deducciones realizadas en el método de Jacobi. A continuación se resumen los resultados fundamentales:

Teorema 3

Una condición suficiente para que el método de Seidel converja hacia la solución del sistema $\mathbf{Ax} = \mathbf{b}$ independientemente de la aproximación inicial $\mathbf{x}^{(0)}$ es que el sistema tenga diagonal predominante.

El error absoluto máximo en la iteración k -sima del algoritmo de Seidel puede tomarse como:

$$E_m(\mathbf{x}^{(k)}) = \left(\frac{\beta}{1-\beta} \right) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \quad (20)$$

y de aquí resulta la condición de terminación:

Condición de terminación:

Si se desea obtener la solución de un sistema lineal con un error absoluto menor que ε , y el factor de convergencia del método de Seidel es $\beta < 1$, entonces el proceso iterativo de Seidel se llevará a cabo hasta la aproximación $\mathbf{x}^{(k)}$ para la cual:

$$E_m(\mathbf{x}^{(k)}) = \left(\frac{\beta}{1-\beta} \right) \cdot \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq \varepsilon$$

En los casos en que β es menor que 0,5, se tiene:

$$\frac{\beta}{1-\beta} < 1$$

y la fórmula (20) implica:

$$E_m(\mathbf{x}^{(k)}) = \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$$

De aquí resulta:

Condición de terminación:

Si se desea obtener la solución de un sistema lineal con un error absoluto menor que ε , y el factor de convergencia del método de Seidel es $\beta < 0,5$, entonces el proceso iterativo de Seidel se llevará a cabo hasta la aproximación $\mathbf{x}^{(k)}$ para la cual:

$$E_m(\mathbf{x}^{(k)}) = \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq \varepsilon$$

La fórmula (19) puede ser escrita en términos del error absoluto máximo:

$$E_m(\mathbf{x}^{(k)}) = \beta \cdot E_m(\mathbf{x}^{(k-1)}) \quad (21)$$

La cual deja ver que el método de Seidel posee convergencia lineal y que valores pequeños de β (próximos a cero) aseguran una convergencia rápida hacia la solución, mientras que valores próximos y menores que 1, dan lugar a una convergencia más lenta. Cuando el factor de convergencia es mayor que 1, no se asegura la convergencia, aunque puede suceder.

Comparación entre la convergencia de los métodos de Jacobi y de Seidel

Se pueden encontrar sistemas en los cuales el método de Seidel converge y el de Jacobi no y también a la inversa, sistemas en los que el método de Seidel diverge y converge el de Jacobi. Sin

embargo, cuando la matriz **A** tiene diagonal predominante, de modo que $\alpha < 1$, la convergencia del método de Jacobi nunca es más rápida que la de Seidel. En efecto, nótese que:

$$p_i + q_i = \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| + \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| = \sum_{j=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \quad \text{para} \quad i = 1, 2, \dots, n$$

y como α es el máximo valor que alcanza la última suma, se tiene que:

$$p_i + q_i \leq \alpha \quad \text{para} \quad i = 1, 2, \dots, n$$

es decir;

$$q_i \leq \alpha - p_i$$

Entonces,

$$\frac{q_i}{1 - p_i} \leq \frac{\alpha - p_i}{1 - p_i} \leq \frac{\alpha - \alpha p_i}{1 - p_i} = \frac{\alpha(1 - p_i)}{1 - p_i} = \alpha$$

y como β es el máximo valor del primer miembro, se tiene que:

$$\text{Si } \alpha < 1 \text{ entonces } \beta \leq \alpha \quad (22)$$

Lo cual prueba que en los casos de diagonal predominante, el método de Seidel tiene una convergencia más rápida o igual que el de Jacobi. En la práctica la diferencia es bastante apreciable y, en casi todos los casos de diagonal predominante, el método de Seidel necesita aproximadamente la mitad de las iteraciones que requiere el método de Jacobi.

Como el parámetro β puede ser más complicado de calcular que α , en ocasiones, cuando la diagonal es predominante, se suele usar α en lugar de β al trabajar con el método de Seidel y esto no conduce a errores precisamente porque en estos casos $\beta \leq \alpha$.

Ejemplo 8

Determine el parámetro β para el sistema lineal a) del ejemplo 7. Compare con el valor de α obtenido.

$$\begin{aligned} 10x_1 - x_2 + 2x_3 &= 6 \\ x_1 - 5x_2 + x_3 &= -10 \\ 2x_1 - x_2 + 8x_3 &= -8 \end{aligned}$$

Solución:

La matriz **M** del sistema es

$$\mathbf{M} = \begin{bmatrix} 0 & 0,1 & -0,2 \\ 0,2 & 0 & 0,2 \\ -0,25 & 0,125 & 0 \end{bmatrix}$$

$$p_1 = 0$$

$$q_1 = 0,1 + 0,2 = 0,3$$

$$\frac{q_1}{1 - p_1} = \frac{0,3}{1 - 0} = 0,3$$

$$p_2 = 0,2$$

$$q_2 = 0,2$$

$$\frac{q_2}{1 - p_2} = \frac{0,2}{1 - 0,2} = 0,25$$

$$p_3 = 0,25 + 0,125 = 0,375 \quad q_3 = 0 \quad \frac{q_3}{1-p_3} = \frac{0}{1-0,375} = 0$$

y como $\beta = \max_i \frac{q_i}{1-p_i}$
se tiene $\beta = \max\{0,3; 0,25; 0\} = 0,3$

En este mismo sistema, el parámetro de convergencia de Jacobi es $\alpha = 0,4$. ■

Algoritmo del método de Seidel

Se desea hallar una solución del sistema $\mathbf{Ax} = \mathbf{b}$ con error absoluto menor que ε . Se supone que $\mathbf{A}$ tiene diagonal predominante y se conoce el valor del factor de convergencia β . El algoritmo utiliza como datos las matrices $\mathbf{A}$ y $\mathbf{b}$, el factor de convergencia β , la tolerancia ε y la aproximación inicial $\mathbf{x}^{(0)}$. De no conocerse $\mathbf{x}^{(0)}$ se puede tomar el vector nulo $\mathbf{0}$.

```

x := x(0)
repeat
    Error := 0
    for i = 1 to n
        prov := bi      {En esta sección se calcula la i-sima componente de x,
                           que se almacena provisionalmente en prov, de modo que
                           no se afecte xi para poder determinar la diferencia entre
                           la iteración actual y la anterior}
        for j = 1 to n
            if j ≠ i then
                prov := prov - aijxj
            end
            prov :=  $\frac{prov}{a_{ii}}$ 
        end
        if  $|prov - x_i| > Error$  then {En este lazo se va guardando la máxima
                                     diferencia entre las dos ultimas iteraciones}
            Error :=  $|prov - x_i|$ 
        end
        xi := prov {El antiguo valor de xi se cambia por el nuevo}
    end

    Error := Error ·  $\left(\frac{\beta}{1-\beta}\right)$ 

until Error <  $\varepsilon$ 
La solución del sistema es x con error absoluto menor que Error
Terminar

```

Ejemplo 9

Resuelva con cuatro cifras decimales exactas el siguiente sistema de ecuaciones mediante el método de Seidel. El sistema ya fue resuelto por el método de Jacobi en el ejemplo 5.

$$9x_1 - x_2 + 2x_3 = 9$$

$$x_1 + 8x_2 + 2x_3 = 19$$

$$x_1 - x_2 + 11x_3 = 10$$

Solución:

Como la diagonal es predominante, se puede asegurar la convergencia del método de Seidel. Para obtener el factor de convergencia se requiere hallar la matriz **M**:

$$\mathbf{M} = \begin{bmatrix} 0 & \frac{1}{9} & -\frac{2}{9} \\ -\frac{1}{8} & 0 & -\frac{1}{4} \\ -\frac{1}{11} & \frac{1}{11} & 0 \end{bmatrix}$$

$p_1 = 0$	$q_1 = 0,33$	$\frac{q_1}{1-p_1} = \frac{0,33}{1-0} = 0,33$
$p_2 = 0,125$	$q_2 = 0,25$	$\frac{q_2}{1-p_2} = \frac{0,25}{1-0,125} = 0,29$
$p_3 = 0,182$	$q_3 = 0$	$\frac{q_3}{1-p_3} = \frac{0}{1-0,182} = 0$

de donde: $\beta = \max\{0,33; 0,29; 0\} = 0,33$

Por ser $\beta < 0,5$, se puede tomar el error absoluto máximo como la norma de $\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}$. Esto es, el proceso iterativo será detenido cuando

$$\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq 0,00005$$

La tabla 9 muestra las aproximaciones obtenidas y el error absoluto máximo de cada una. Aparecen marcadas con * las componentes de la solución donde ocurre la máxima diferencia con la iteración anterior.

Iteración	x_1	x_2	x_3	$E_m(\mathbf{x}^{(k)})$
0	0,0	0,0	0,0	-----
1	1,0	2,25*	1,022727	2,25
2	1,022727	1,991477*	0,997159	0,258523
3	0,999684*	2,000750	1,000097	0,023043
4	1,000062	1,999968*	0,999991	0,000782
5	0,999998*	2,000002	1,000000	0,000063
6	1,000000	2,000000*	1,000000	0,000002

Tabla 9

La solución aproximada, con cinco cifras decimales exactas, es:

$$\begin{aligned}x_1 &= 1,000000 \\x_2 &= 2,000000 \\x_3 &= 1,000000\end{aligned}$$

Aunque la solución se pidió con cuatro cifras decimales exactas, fueron obtenidas cinco, ya que en la quinta iteración todavía el error era mayor que 0,00005 pero en la sexta iteración está por debajo de 0,000005. Como en este problema se sabe que la solución exacta es $x_1 = 1$; $x_2 = 2$; $x_3 = 1$ puede apreciarse que ya desde la quinta iteración se tenían 5 cifras decimales exactas, pero la cota del error, que es lo que permite detener el proceso, siempre es mayor que el verdadero error absoluto.

Ejemplo 10

En el ejemplo 2 del comienzo del capítulo, se llegó al sistema de 16 ecuaciones con 16 incógnitas:

$$4u_{ij} - u_{i-1,j} - u_{i+1,j} - u_{i,j-1} - u_{i,j+1} = 0$$

para $i = 1, 2, 3, 4$; $j = 1, 2, 3, 4$

donde u_{ij} representa la temperatura en el punto (i, j) de una red definida sobre una lamina metálica cuadrada sometida en sus bordes a temperaturas fijas. Como las temperaturas en el contorno son conocidas, se sabe que:

$$\begin{aligned}u_{01} &= u_{02} = u_{03} = u_{04} = 20 \\u_{51} &= u_{52} = u_{53} = u_{54} = 60 \\u_{10} &= u_{20} = u_{30} = u_{40} = 80 \\u_{15} &= u_{25} = u_{35} = u_{45} = 40\end{aligned}$$

Elabore un algoritmo en pseudo código basado en el método de Seidel, que permita obtener la solución del sistema con tres cifras decimales exactas.

Solución:

Si en cada ecuación del sistema se coloca en la diagonal la variable con coeficiente 4, es claro que la matriz del sistema (de 16 por 16) tiene la diagonal casi predominante, pues la suma de los elementos (en valor absoluto) que no están en la diagonal es exactamente 4. Aunque en este caso, la convergencia no se puede garantizar, es muy probable que el algoritmo converja ya que el parámetro α es exactamente 1 en lugar de la condición suficiente que es $\alpha < 1$. En estos casos, el algoritmo se utiliza a riesgo y se toma como condición de terminación $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq \varepsilon$ con una tolerancia mucho menor que la que se desea. Como en el problema se piden tres cifras decimales exactas, se tomará una tolerancia $\varepsilon = 0,00005$, diez veces menor que la necesaria.

En el algoritmo que sigue, se emplea la variable provisional *prov* para guardar el nuevo valor de cada incógnita sin borrar el de la iteración anterior, de manera que pueda ser hallada la diferencia entre ambas. La variable *Error* va actualizándose con la máxima diferencia entre las dos iteraciones hasta el momento en que se calcula la nueva u_{ij} .

```

for  $k = 1$  to 4           {Aquí se fijan las condiciones de frontera}
     $u_{0k} := 20$ 
     $u_{5k} := 60$ 
     $u_{k0} := 80$ 
     $u_{k5} := 60$ 
end
for  $i = 1$  to 4           {Todas las incógnitas se inicializan en cero}
    for  $j = 1$  to 4
         $u_{ij} := 0$ 
    end
end
repeat
     $Error := 0$ 
    for  $i = 1$  to 4
        for  $j = 1$  to 4
             $prov := (u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}) / 4$ 
            if  $|prov - u_{ij}| > Error$  then
                 $Error := |prov - u_{ij}|$ 
            end
             $u_{ij} := prov$ 
        end
    end
until  $Error < 0,00005$ 
La solución es  $u_{ij}$  ( $i = 1, 2, 3, 4$ ;  $j = 1, 2, 3, 4$ )
Terminar

```

El algoritmo anterior converge en unas 30 iteraciones con una baja rapidez debido a que la diagonal no es predominante.

Comentarios finales sobre los métodos iterativos

A diferencia de los métodos directos, en que la cantidad de operaciones aritméticas requeridas está determinada únicamente por el orden del sistema, en los métodos iterativos hay que tener en cuenta la cantidad de iteraciones que se necesitará; esta depende de la exactitud deseada y de la

rapidez de convergencia del método que se emplee. En la demostración del teorema 1 sobre la convergencia del método de Jacobi se obtuvo una fórmula que puede utilizarse para decidir:

$$E(\mathbf{x}^{(k)}) \leq \alpha^k E(\mathbf{x}^{(0)})$$

que se puede escribir en términos de errores absolutos máximos como:

$$E_m(\mathbf{x}^{(k)}) = \alpha^k E_m(\mathbf{x}^{(0)}) \quad (23)$$

Para el método de Seidel, análogamente, se tiene:

$$E_m(\mathbf{x}^{(k)}) = \beta^k E_m(\mathbf{x}^{(0)}) \quad (24)$$

Estas fórmulas permiten predecir cuantas iteraciones se necesitará para obtener una exactitud deseada si se puede hacer una estimación del error absoluto máximo en la iteración inicial, lo cual requiere tener una idea aproximada de los valores que tomará la solución. Por ejemplo, si se desea la solución de un sistema con tolerancia 0,00005 (cuatro cifras decimales exactas), se toma la aproximación inicial como **0** y se sabe que la solución debe tener todas sus componentes por debajo de 5, entonces $E_m(\mathbf{x}^{(0)}) = 5$ y $E_m(\mathbf{x}^{(k)}) = 0,00005$. Si se selecciona el método de Seidel, entonces la cantidad de iteraciones que se requerirá se puede despejar la fórmula (24):

$$k = \frac{\ln \left[\frac{E_m(\mathbf{x}^{(k)})}{E_m(\mathbf{x}^{(0)})} \right]}{\ln \beta}$$

Si, en este mismo ejemplo $\beta = 0,3$ entonces harán falta:

$$k = \frac{\ln \left[\frac{0,00005}{5} \right]}{\ln 0,3} = 9,56$$

es decir, unas 10 iteraciones. Sin embargo, si $\beta = 0,9$ entonces el número de iteraciones se incrementa a:

$$k = \frac{\ln \left[\frac{0,00005}{5} \right]}{\ln 0,9} = 109,27$$

o sea 110 iteraciones. Tanto en el método de Jacobi como en el de Seidel, se requieren n^2 operaciones de multiplicar y dividir en cada iteración (n operaciones para cada ecuación y son n ecuaciones), así que multiplicando se obtiene la cantidad de iteraciones necesarias.

Sin embargo, el único factor para decidir si se aplica un método directo o uno iterativo no es la cantidad de operaciones. Si el sistema tiene muchas ecuaciones, todas con la misma estructura, como en los ejemplos 6 y 10, seguramente es preferible escribir un pequeño programa que implemente el método de Seidel o el Jacobi, que tener que escribir todos los coeficientes de la

matriz A (que pueden ser decenas de miles) para utilizar un programa profesional de alguna biblioteca numérica con un método directo.

Como los métodos iterativos solo se utilizan para sistemas con diagonal predominante, muchas veces no hay nada que decidir, porque aunque teóricamente siempre es posible hacer transformaciones elementales en el sistema que hagan predominante la diagonal, esto puede resultar prácticamente imposible en sistemas de más de tres o cuatro ecuaciones.

En cuanto a la selección entre Seidel y Jacobi, el método de Seidel converge mucho más rápido, es más fácil de programar y requiere la mitad de la memoria. Solo en una cosa lleva ventaja el método de Jacobi: si se va a implementar el algoritmo en una máquina de cálculo paralelo, el algoritmo de Jacobi permite calcular en cada iteración todas las incógnitas (si la máquina posee esta posibilidad) simultáneamente, ya que cada una requiere los mismos datos y es independiente de la otra; el método de Seidel es, por naturaleza, un método secuencial, ya que para calcular una incógnita en una iteración hay que haber calculado todas las que le preceden.

Ejemplo 11

En un sistema lineal con diagonal predominante, se puede hallar una estimación inicial de la solución de tal modo que $E_m(\mathbf{x}^{(0)}) = 1$. Se necesita la solución con cuatro cifras decimales exactas y los parámetros de convergencia son $\alpha = 0,6$ y $\beta = 0,45$. Si el sistema tiene 70 ecuaciones con 70 incógnitas, calcule cuantas operaciones se necesitarán utilizando el método de Gauss, el de Jacobi y el de Seidel.

Solución:

El método de Gauss requeriría: $\frac{1}{3}n^3 = \frac{1}{3}70^3 = 114334$ operaciones

El método de Jacobi, necesita realizar

$$k = \frac{\ln \left[\frac{E_m(\mathbf{x}^{(k)})}{E_m(\mathbf{x}^{(0)})} \right]}{\ln \alpha} = \frac{\ln \left[\frac{0,00005}{1} \right]}{\ln 0,6} = 19,39 \text{ iteraciones}$$

es decir, 20. En cada iteración realizará $70^2 = 4900$ operaciones. Por tanto necesitará

$$20 \cdot 4900 = 98000 \text{ operaciones}$$

El método de Seidel, requiere:

$$k = \frac{\ln \left[\frac{E_m(\mathbf{x}^{(k)})}{E_m(\mathbf{x}^{(0)})} \right]}{\ln \beta} = \frac{\ln \left[\frac{0,00005}{1} \right]}{\ln 0,45} = 12,40$$

o sea, 13 iteraciones. A razón de 4900 operaciones por cada iteración, un total de

$$13 \cdot 4900 = 63700 \text{ operaciones}$$

En este ejemplo, el método de Seidel lleva la ventaja, aunque la diferencia no es realmente significativa.

Ejercicios

En los siguientes ejercicios, utilice programas computacionales, preferiblemente confeccionados por usted. De no contar con un programa adecuado, use una calculadora y trabaje con cinco cifras significativas en los cálculos.

- Para los siguientes sistemas de ecuaciones, determine si la diagonal es predominante, calcule el valor del factor α de convergencia del método de Jacobi y, si $\alpha < 1$, determine el valor del factor de convergencia β del método de Seidel. De cumplirse las condiciones suficientes de convergencia, establezca el criterio de terminación para obtener la solución con cuatro cifras decimales exactas si se utilizara cada uno de estos métodos.

$\begin{aligned} 10x_1 - 2x_2 + 3x_3 + 2x_4 &= 15 \\ 3x_1 - 10x_2 - 4x_3 + 2x_4 &= 4 \\ 5x_1 + 3x_2 + 10x_3 - x_4 &= 13 \\ 4x_1 - 3x_2 - 4x_3 + 10x_4 &= -2 \end{aligned}$	$\begin{aligned} 10x_1 + x_2 + 2x_3 - x_4 &= 3 \\ -2x_1 + 10x_2 + x_3 - x_4 &= 4 \\ 2x_1 + 3x_2 - 10x_3 - 2x_4 &= -7 \\ 3x_1 - 2x_2 + 4x_3 + 10x_4 &= 4 \end{aligned}$
$\begin{aligned} 10x_1 + 2x_2 - 3x_3 + 4x_4 &= 1 \\ -x_1 + 10x_2 + 3x_3 + 2x_4 &= -3 \\ 2x_1 - 3x_2 - 10x_3 + 2x_4 &= 2 \\ x_1 - 2x_2 - 4x_3 + 10x_4 &= -1 \end{aligned}$	

- Escriba los siguientes sistemas lineales de manera que se pueda garantizar la convergencia de los métodos de Jacobi y de Seidel y calcule los valores de los parámetros α y β .

$\begin{aligned} x + 3y - 9z &= 3 \\ 5x - y + 2z &= 25 \\ 4x - 15y + z &= -7 \end{aligned}$	$\begin{aligned} 9z + 3y - x &= 4 \\ 8x - y + 2z &= 14 \\ 3x - 7y - z &= 5 \end{aligned}$	$\begin{aligned} -x + 4y - 2z &= 5 \\ 3x + 2y + 9z &= 25 \\ 5x + 2y + z &= 4 \end{aligned}$
---	---	---

- Intente resolver los tres sistemas del ejercicio 1 (a pesar de que el primero no tiene la diagonal predominante) por los métodos de Jacobi y de Seidel. Compare la rapidez con que se obtiene la solución con 5 cifras decimales exactas para cada sistema y para cada método y analice su relación con los respectivos factores de convergencia.
- Elabore un algoritmo en pseudo código para resolver el siguiente sistema lineal mediante el método de Seidel con cuatro cifras decimales exactas.

$$8u_{ij} - u_{i-1,j} - 2u_{i+1,j} - u_{i,j-1} - u_{i,j+1} = 0 \quad i = 1, 2, 3; \quad j = 1, 2, 3$$

Se sabe que:

$$\begin{aligned} u_{01} &= u_{02} = u_{03} = 25 \\ u_{41} &= u_{42} = u_{43} = -3 \\ u_{10} &= u_{20} = u_{30} = 18 \\ u_{14} &= u_{24} = u_{34} = 12 \end{aligned}$$

- Las x_{ij} de la región de la figura 3 satisfacen la ecuación

$$4u_{ij} = u_{i-1,j} + u_{i+1,j} + \frac{u_{i,j-1} + u_{i,j+1}}{2}$$

siempre que (i, j) corresponda a un punto interior (no en la frontera) de la región triangular. Los valores de u_{ij} para puntos (i, j) de la frontera se muestran en la propia figura. Elabore un algoritmo en pseudo código basado en el método de Seidel para obtener la solución con cuatro cifras decimales exactas.

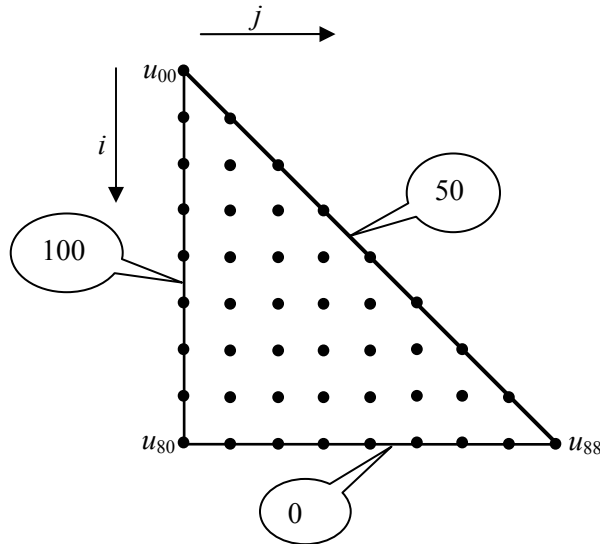


Figura 3

6. Se necesita resolver un sistema con diagonal predominante de n ecuaciones y n incógnitas con 4 cifras decimales exactas. Se puede hallar una aproximación inicial con error absoluto menor que 0,5. Determine la cantidad de operaciones que se necesitaría para hallar la solución mediante el método de Gauss y mediante el método iterativo de Seidel si el valor de n fuera 10, 100 ó 1000 y si el factor de convergencia fuera $\beta = 0,2$; $\beta = 0,5$ ó $\beta = 0,8$. Analice qué método sería más rápido en cada una de las nueve combinaciones de n y β .

3.6 Cálculo de valores y vectores propios

Definición 1

Si $\mathbf{A}$ es una matriz cuadrada de orden n se dice que el número λ es un valor propio de $\mathbf{A}$ si existe algún vector $\mathbf{x}$ no nulo de R^n tal que:

$$\mathbf{Ax} = \lambda \mathbf{x} \quad (1)$$

Si λ es un valor propio de $\mathbf{A}$, a todos los vectores $\mathbf{x}$ que satisfacen (1), incluido el vector nulo, se les llama vectores propios de $\mathbf{A}$ asociados al valor propio λ . ■

En la definición anterior y en todo lo que sigue, se supone que los elementos de la matriz $\mathbf{A}$ son números reales. No obstante, aun en ese caso, la ecuación (1) puede ser satisfecha para valores imaginarios de λ y de $\mathbf{x}$. Aunque el interés fundamental en este libro se centrará en el caso de valores y vectores propios reales, algunos resultados requieren que se consideren también los valores imaginarios; en esos casos se hará la advertencia explícitamente, de lo contrario, puede suponer que λ es un número real y que los elementos de $\mathbf{x}$ son números reales.

En algunas cuestiones teóricas y prácticas aparece la necesidad de calcular los valores y vectores propios de una matriz. Por ejemplo, la estabilidad de algunos procesos numéricos iterativos depende del valor propio de mayor valor absoluto de alguna matriz, la frecuencia con que puede oscilar un edificio alto y el modo en que puede producirse la oscilación depende de los valores y vectores propios de la matriz de rigidez de la estructura. Incluso en algunos procesos naturales, como el que se muestra en el siguiente ejemplo, aparecen valores y vectores propios.

Ejemplo 1

Para estudiar cómo cambia la estructura de una población en cuanto a las edades, suponga que esta se ha dividido en 5 grupos de edad del siguiente modo:

grupo 1:	de 0 a 9 años
grupo 2:	de 10 a 14 años
grupo 3:	de 15 a 20 años
grupo 4:	de 21 a 45 años
grupo 5:	más de 45 años

Se supone, además, que en todos los grupos de edades la cantidad de individuos de cada sexo es idéntico, lo cual es muy próximo a la realidad. Se define un periodo de tiempo, digamos de un año, para cuantificar la cantidad de individuos en cada grupo de edad. Para el año n , la población en cada grupo de edad se define como $p_{1n}, p_{2n}, p_{3n}, p_{4n}$ y p_{5n} respectivamente. Supóngase que la cantidad de individuos en cada grupo de edad en el año siguiente ($n + 1$) depende de las cifras del año actual mediante las siguientes relaciones:

Los individuos del grupo 1 en el año $n + 1$ proceden o de ese grupo o nacieron en el transcurso del año, a partir de individuos de otros grupos de edades. Se supone que cada grupo de edad tiene su propia tasa de procreación t_i ($i = 1, 2, 3, 4, 5$) que es 0 en el primer grupo, muy baja en los grupos 2 y 5 y mayor en los grupos 3 y 4. Así:

$$p_{1,n+1} = \beta_1 p_{1,n} + t_2 p_{2,n} + t_3 p_{3,n} + t_4 p_{4,n} + t_5 p_{5,n}$$

Para los demás grupos de edad, la cantidad de individuos en el año $n + 1$ viene dada por una parte (β) de los que estaban en ese grupo de edad y que todavía sigue estando, más una parte (α) de los del grupo más joven que arribaron durante el año al grupo mayor. Esto es:

$$\begin{aligned} p_{2,n+1} &= \alpha_2 p_{1,n} + \beta_2 p_{2,n} \\ p_{3,n+1} &= \alpha_3 p_{2,n} + \beta_3 p_{3,n} \\ p_{4,n+1} &= \alpha_4 p_{3,n} + \beta_4 p_{4,n} \\ p_{5,n+1} &= \alpha_5 p_{4,n} + \beta_5 p_{5,n} \end{aligned}$$

Si se define el vector población $\mathbf{p}_n$ como un vector cuyas cinco componentes representan la población en año n de cada grupo poblacional, entonces las relaciones anteriores se pueden representar en forma matricial como:

$$\mathbf{p}_{n+1} = \mathbf{A} \mathbf{p}_n \quad (2)$$

$$\text{donde: } \mathbf{p}_{n+1} = \begin{bmatrix} p_{1,n+1} \\ p_{2,n+1} \\ p_{3,n+1} \\ p_{4,n+1} \\ p_{5,n+1} \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} \beta_1 & t_2 & t_3 & t_4 & t_5 \\ \alpha_2 & \beta_2 & 0 & 0 & 0 \\ 0 & \alpha_3 & \beta_3 & 0 & 0 \\ 0 & 0 & \alpha_4 & \beta_4 & 0 \\ 0 & 0 & 0 & \alpha_5 & \beta_5 \end{bmatrix} \quad \mathbf{p}_n = \begin{bmatrix} p_{1,n} \\ p_{2,n} \\ p_{3,n} \\ p_{4,n} \\ p_{5,n} \end{bmatrix}$$

Surge entonces el siguiente problema: ¿existirá alguna estructura de edades estable? es decir, tal que para algún número positivo λ se satisfaga:

$$\mathbf{p}_{n+1} = \lambda \mathbf{p}_n$$

Si esta estructura poblacional existe, teniendo en cuenta la ecuación (2), se deberá cumplir que:

$$\mathbf{A} \mathbf{p}_n = \lambda \mathbf{p}_n$$

Comparando esta ecuación con (1), se ve que el problema equivale a buscar los valores y vectores propios de la matriz $\mathbf{A}$.

Ejemplo 2

Compruebe que $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ es un vector propio de la matriz $\mathbf{A} = \begin{bmatrix} 5 & 3 & -1 \\ 1 & -2 & 2 \\ 4 & 3 & 3 \end{bmatrix}$ y determine a qué valor propio corresponde.

Solución:

En efecto, el producto $\mathbf{A}\mathbf{x}$ da como resultado:

$$\begin{bmatrix} 5 & 3 & -1 \\ 1 & -2 & 2 \\ 4 & 3 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 12 \end{bmatrix} = 4 \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

Es decir, $\mathbf{A}\mathbf{x} = 4\mathbf{x}$, lo cual demuestra que $\mathbf{x}$ es un vector propio de $\mathbf{A}$ que corresponde al valor propio $\lambda = 4$. ■

Aunque se supone que el lector tiene los conocimientos básicos de un curso elemental de Álgebra Lineal, a continuación se tratan algunos de estos conceptos a modo de repaso.

Sub espacios propios

Es sencillo demostrar que todos los vectores propios de una matriz $\mathbf{A}$ de orden n , que corresponden a un mismo valor propio λ , forman un sub espacio vectorial $E(\lambda)$ de R^n con dimensión 1 o mayor. Esto significa que cuando se conoce un vector propio de una matriz, entonces todos los múltiplos de ese vector son también vectores propios de la matriz.

Si se conoce λ , el sub espacio propio $E(\lambda)$ se puede determinar sin dificultad resolviendo el sistema lineal homogéneo:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

el cual siempre es indeterminado, pues de lo contrario su única solución sería el vector nulo.

Polinomio característico

Teóricamente, los valores propios se hallan de la siguiente forma. Si la ecuación (1) se escribe:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{I}\mathbf{x}$$

donde $\mathbf{I}$ es la matriz idéntica de orden n , y se traspone el segundo miembro al primero:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \quad (3)$$

Para que el sistema homogéneo (3) pueda tener alguna solución no trivial es necesario y suficiente que el determinante de la matriz $(\mathbf{A} - \lambda\mathbf{I})$ sea nulo. Se demuestra que el determinante:

$$\det(\mathbf{A} - \lambda\mathbf{I})$$

es un polinomio de grado n en λ y se le llama *polinomio característico* de la matriz $\mathbf{A}$. Así pues, los valores propios de $\mathbf{A}$ son las n raíces de la ecuación algebraica de grado n :

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

que recibe el nombre de *ecuación característica*.

Para matrices de orden pequeño, no es difícil hallar el polinomio característico, pero si n excede de 4 o 5 es bastante complicado, por cuanto no se trata de un problema aritmético sino que se requiere calcular un determinante con términos literales, lo cual requiere de operaciones algebraicas simbólicas. Ningún método numérico con valor práctico supone que se halle en algún momento el polinomio característico de la matriz.

Ejemplo 3

Halle el polinomio característico de la matriz $\mathbf{A}$ del ejemplo anterior y calcule sus raíces para hallar los valores propios de $\mathbf{A}$.

Solución:

Como la matriz es

$$\mathbf{A} = \begin{bmatrix} 5 & 3 & -1 \\ 1 & -2 & 2 \\ 4 & 3 & 3 \end{bmatrix}$$

el polinomio característico es:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 5-\lambda & 3 & -1 \\ 1 & -2-\lambda & 2 \\ 4 & 3 & 3-\lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 + 6\lambda - 56$$

La gráfica del polinomio característico se muestra en la figura 1. Es evidente la presencia de tres raíces reales en los intervalos $[-3,5; -2,5]$, $[3,5; 4,5]$ y $[4,5; 5,5]$. En el segundo de estos intervalos se encuentra el valor propio $\lambda = 4$ analizado en el ejemplo 2.

$$-\lambda^3 + 6\lambda^2 + 6\lambda - 56$$

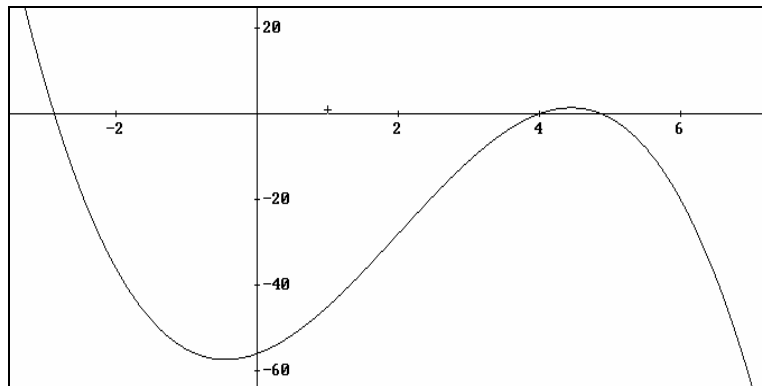


Figura 1

■

El hecho de que los valores propios son raíces de una ecuación algebraica de grado n explica por qué pueden aparecer valores propios imaginarios, aun cuando los elementos de la matriz $\mathbf{A}$ sean números reales. Se demuestra que si el valor propio λ es una raíz de multiplicidad k del polinomio característico, entonces la dimensión del sub espacio propio $E(\lambda)$ no puede ser mayor que k .

El caso especial de las matrices simétricas

La matriz $\mathbf{A}$ se llama simétrica si cada elemento a_{ij} coincide con su elemento simétrico a_{ji} . En muchas aplicaciones importantes aparecen matrices simétricas, para las cuales el problema de los valores y vectores propios se simplifica notablemente. En efecto, si $\mathbf{A}$ es simétrica entonces todos los valores propios son reales (aunque no necesariamente diferentes) y los vectores propios son ortonormales (ortogonales y con norma 1). Como se verá posteriormente, estas características permiten hallar los vectores y valores propios de manera más eficiente y segura.

Localización de los valores propios

Aunque los valores propios son las raíces de una ecuación polinomial, en la práctica no se conocen los coeficientes del polinomio (salvo algunos de ellos). Por esta causa, las reglas de Descartes y de Lagrange no serán aquí de utilidad. En su lugar, existen algunas propiedades interesantes que permiten localizar a groso modo los valores propios de la matriz.

Todos los valores propios λ (reales e imaginarios) de la matriz $\mathbf{A}$ satisfacen la desigualdad:

$$|\lambda| \leq \|\mathbf{A}\|$$

Esto significa que si se traza un círculo con centro en el origen del plano coordenado y radio $\|\mathbf{A}\|$, entonces todos los valores propios de la matriz $\mathbf{A}$ están dentro de este círculo.

Otros resultados importantes que permiten en algunos casos encontrar algunos de los valores propios de la matriz y que por lo menos, se pueden utilizar para comprobar los cálculos realizados, son los siguientes:

Si $\lambda_1, \lambda_2, \dots, \lambda_n$ son todos los valores propios (reales e imaginarios) de la matriz $\mathbf{A}$ de orden n , se cumple que:

$$\lambda_1 + \lambda_2 + \dots + \lambda_n = \text{Traza}(\mathbf{A}) = a_{11} + a_{22} + \dots + a_{nn}$$

$$\lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n = \det(\mathbf{A})$$

Ejemplo 4

Para la matriz $\mathbf{A}$ de los ejemplos 2 y 3, ya se sabe que uno de sus valores propios es 4. Halle los otros dos a partir de la propiedad anterior.

Solución:

Como $\mathbf{A} = \begin{bmatrix} 5 & 3 & -1 \\ 1 & -2 & 2 \\ 4 & 3 & 3 \end{bmatrix}$ se tiene que $\text{Traza}(\mathbf{A}) = 5 - 2 + 3 = 6$ y $\text{Det}(\mathbf{A}) = -56$, de aquí se tiene

que: $\lambda_1 + \lambda_2 + \lambda_3 = 6$

y $\lambda_1 \lambda_2 \lambda_3 = -56$

Como uno de los valores propios (digamos λ_3) tiene valor 4, los otros dos cumplirán que:

$$\lambda_1 + \lambda_2 = 2 \quad (4)$$

y $\lambda_1 \lambda_2 = -14$

Despejando λ_2 en la primera ecuación y sustituyendo en la segunda:

$$\lambda_1(2 - \lambda_1) = -14$$

esto es: $\lambda_1^2 - 2\lambda_1 - 14 = 0$

de donde: $\lambda_1 = 1 \pm \sqrt{15}$

Tomando para λ_1 el valor negativo, se tiene: $\lambda_1 = 1 - \sqrt{15}$ y, utilizando la ecuación (4), $\lambda_2 = 1 + \sqrt{15}$. Es obvio que al tomar λ_1 como el valor positivo se obtendría para λ_2 el valor negativo. En resumen, los tres valores propios son en este caso:

$$\lambda_1 = 1 - \sqrt{15} = -2,872983$$

$$\lambda_2 = 1 + \sqrt{15} = 4,872983$$

$$\lambda_3 = 4$$

■

Gráfica del polinomio característico

Aunque hallar el polinomio característico es un problema simbólico engorroso cuando el orden de la matriz es mayor que 4, es relativamente simple, con un pequeño programa computacional, evaluar el polinomio característico para valores numéricos específicos de la variable λ , lo cual solamente implica calcular determinantes *numéricos* de orden n .

Ejemplo 5

Para la matriz $\mathbf{M}$ de orden 5, determine una zona donde estén todos sus valores propios reales y haga la gráfica del polinomio característico evaluándolo para algunos valores de la variable λ .

$$\mathbf{M} = \begin{bmatrix} 2 & 1 & 0 & -1 & 3 \\ 1 & 0 & 3 & -2 & 1 \\ 3 & 1 & -1 & 2 & 1 \\ 2 & 3 & 0 & -1 & 2 \\ 2 & -1 & 1 & 3 & 0 \end{bmatrix}$$

Solución:

La norma de la matriz es $\|\mathbf{M}\| = \max\{7, 7, 8, 8, 7\} = 8$, así que todos los valores propios reales se encuentran en el intervalo $[-8, 8]$. Evaluando el determinante:

$$\det(\mathbf{M} - \lambda \mathbf{I})$$

para valores de λ desde -8 hasta 8 con paso $0,2$, se obtienen 81 puntos que aparecen en la gráfica de la figura 2.

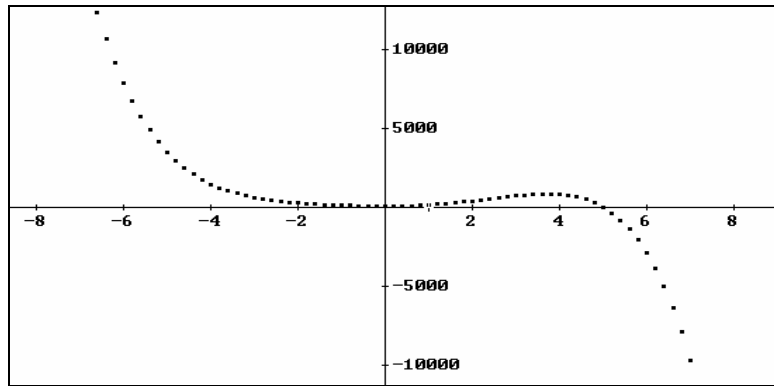


Figura 2

De la gráfica se aprecia que existe un valor propio cerca de 5. Para poder apreciar mejor qué sucede en el intervalo $[-2, 2]$, la figura 3 muestra la gráfica en ese intervalo a una escala mayor:

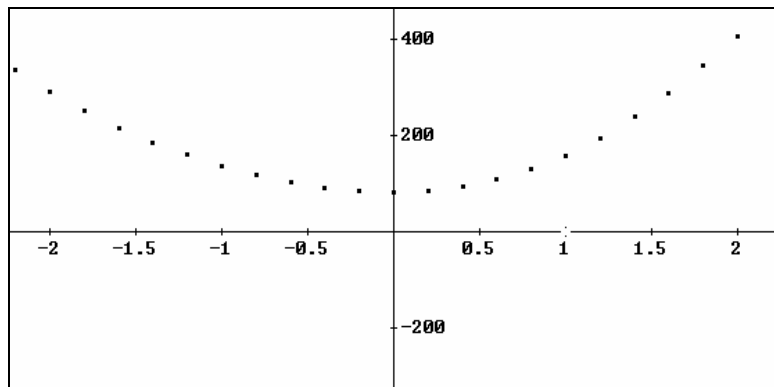


Figura 3

y ya de aquí resulta obvio que en ese intervalo no hay valores propios. En conclusión, la matriz $\mathbf{M}$ solo posee un valor propio real y se halla próximo a 5. Obsérvese que todo este trabajo gráfico se ha realizado sin utilizar el polinomio característico.

■

Para hallar uno o varios valores propios de una matriz existe una gran cantidad de técnicas, muchas de las cuales están más allá del carácter elemental de este texto. En la bibliografía recomendada al final de este capítulo pueden verse algunas referencias importantes. En lo que sigue se hace una descripción general de las estrategias de solución más empleadas.

Solución numérica de la ecuación característica

Como los valores propios de una matriz son las raíces de la ecuación característica, estos pueden hallarse por varios de los métodos estudiados en el capítulo 2. Si lo que se desea son los valores propios reales, los métodos de bisección, Regula Falsi y el método de las secantes pueden ser empleados. Como cada vez que se evalúa el polinomio característico, cuya expresión analítica se supone desconocida, se requiere calcular un determinante de orden n , aquí es importante utilizar un método eficiente, de manera que haya que evaluarlo pocas veces; por esta razón el método preferido en este caso es el de las secantes. En cuanto al método de Newton – Raphson, la necesidad de hallar la derivada del polinomio característico, lo hace inadecuado en este caso.

Para hallar los valores propios imaginarios no se aconseja utilizar el método de Newton – Bairstow, el cual requiere conocer los coeficientes del polinomio característico. Existe un método que no ha sido tratado en este texto, llamado método de Muller, que generaliza el método de las secantes. En el método de Muller en lugar de hallar la recta que pasa por dos puntos de la gráfica de la función, determinados por las dos aproximaciones anteriores, se hace pasar una parábola por tres puntos de la gráfica que corresponden a las tres últimas aproximaciones. Este procedimiento es sumamente eficiente, aunque el algoritmo es algo complicado, y permite hallar no solo los valores propios reales sino también los imaginarios. En la bibliografía recomendada del capítulo se dan referencias acerca del método.

Transformación de la matriz en una similar

Una gran cantidad de algoritmos están basados en la idea de realizar sobre la matriz **A**, cuyos valores propios se desea hallar, transformaciones que simplifiquen su estructura sin afectar los valores propios. Una vez simplificada la matriz, se hallan los valores propios de una manera más fácil. Las transformaciones que se realizan se llaman de similitud. El término matrices similares se refiere a matrices **A** y **B** relacionadas entre si por la ecuación:

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

donde **P** se llama matriz de transformación. Se puede demostrar sin mucha dificultad que dos matrices similares tienen los mismos valores propios.

Por ejemplo, para matrices tridiagonales y simétricas (ambas cosas a la vez) la evaluación del polinomio característico se realiza de modo muy eficiente y la ecuación característica puede ser resuelta numéricamente con muy pocas operaciones. Entonces, para hallar los valores propios de una matriz simétrica, se realizan sobre ella transformaciones de similitud que la transformen en una tridiagonal simétrica. Para ello se han usado diferentes matrices de transformación. Una de los algoritmos de este tipo más empleado es el que utiliza las matrices de transformación de Householder acerca del cual puede encontrar referencias al final del capítulo.

El método de la potencia

Este es un método muy simple basado en una ingeniosa idea, que permite hallar el valor propio (real) de mayor valor absoluto. Para comprender como el método funciona, suponga una matriz **A** de orden 2, que posee dos valores propios reales λ_1 y λ_2 tales que λ_1 es mayor que λ_2 en valor absoluto. Para concretar, se supondrá que ambos son positivos y que $\lambda_1 = k\lambda_2$ con $k > 1$. Si se

llama $\mathbf{x}_1$ y $\mathbf{x}_2$ a dos vectores propios unitarios (con norma 1) de la matriz $\mathbf{A}$ correspondientes a los valores propios λ_1 y λ_2 , se tiene que:

$$\begin{aligned} \mathbf{A}\mathbf{x}_1 &= \lambda_1\mathbf{x}_1 \\ \mathbf{A}\mathbf{x}_2 &= \lambda_2\mathbf{x}_2 \end{aligned}$$

y

Sea ahora un vector $\mathbf{z}$ no nulo cualquiera de R^2 . Como $\mathbf{x}_1$ y $\mathbf{x}_2$ forman una base de R^2 , el vector $\mathbf{z}$ se puede expresar como combinación lineal de ellos, es decir:

$$\mathbf{z} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 \quad (5)$$

Si se multiplica la matriz $\mathbf{A}$ por el vector $\mathbf{z}$, la imagen $\mathbf{Az}$ obtenida será un vector de R^2 dado por:

$$\begin{aligned} \mathbf{Az} &= \mathbf{A}(\alpha_1\mathbf{x}_1) + \mathbf{A}(\alpha_2\mathbf{x}_2) = \alpha_1\mathbf{Ax}_1 + \alpha_2\mathbf{Ax}_2 = \alpha_1\lambda_1\mathbf{x}_1 + \alpha_2\lambda_2\mathbf{x}_2 \\ \mathbf{Az} &= \lambda_1(\alpha_1\mathbf{x}_1) + \lambda_2(\alpha_2\mathbf{x}_2) \end{aligned} \quad (6)$$

Si se compara la expansión del vector $\mathbf{z}$ (5) con la de su imagen (6) se observa que las componentes sobre $\mathbf{x}_1$ y sobre $\mathbf{x}_2$ no se transformaron de igual modo, mientras la primera quedó multiplicada por λ_1 la segunda lo hizo por λ_2 . Bajo la suposición de que $\lambda_1 > \lambda_2$ la componente a lo largo de $\mathbf{x}_1$ ha adquirido una mayor preponderancia que la componente sobre $\mathbf{x}_2$. En la figura 4 se ilustra geoméricamente esta situación, en ella se ha tomado $\lambda_1 = 3$ y $\lambda_2 = 0,6$ es decir, $k = 5$.

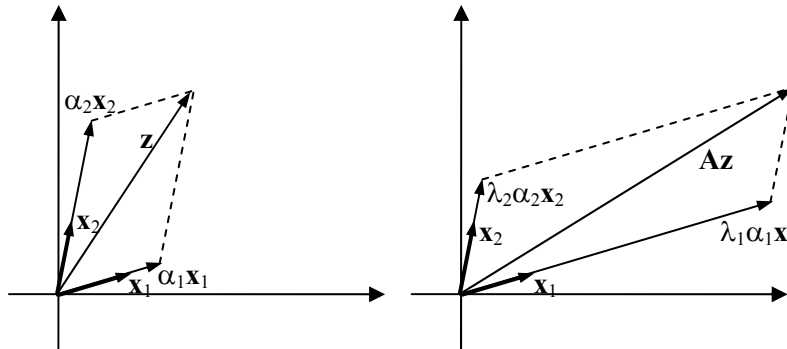


Figura 4

Como se ve, el vector $\mathbf{Az}$ está más cerca del vector propio $\mathbf{x}_1$ de lo que lo estaba $\mathbf{z}$. Si este proceso se repite, tomando ahora como vector de partida $\mathbf{Az}$ se obtiene un nuevo vector $\mathbf{A}^2\mathbf{z}$ que estará aun más próximo a la dirección de $\mathbf{x}_1$. Es fácil comprender que el vector $\mathbf{A}^n\mathbf{z}$ converge hacia la dirección de $\mathbf{x}_1$ cuando n tiende hacia infinito con una rapidez que depende del valor de k . Para evitar que en este proceso el vector imagen crezca desmesuradamente (si $\lambda_1 > 1$) o decrezca mucho (si $\lambda_1 < 1$), en cada paso del proceso el vector se normaliza (se divide por su norma, para hacerlo unitario) antes de hallarle su imagen. Con esta idea, el algoritmo de la potencia, consta esencialmente de los siguientes pasos:

1. Tomar un vector $\mathbf{z}$ cualquiera
2. Normalizarlo, para obtener $\mathbf{z}_0$
3. Hallar el vector imagen $\mathbf{w} = \mathbf{Az}_0$
4. Regresar a 2.

Para terminar el proceso iterativo, observe que, en la medida en que el vector $\mathbf{z}_0$ se va aproximando al vector propio $\mathbf{x}_1$ el vector $\mathbf{w}$ se va aproximando a $\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1$, así que, en cada iteración, se puede calcular una aproximación a λ_1 hallando cuánto cambian las componentes de $\mathbf{z}_0$ al aplicarle la matriz $\mathbf{A}$ para obtener $\mathbf{w}$. Si se llama λ_{aprox} a esta aproximación, el proceso se puede detener cuando su valor se va estabilizando, es decir, cuando el cambio sufrido entre dos iteraciones llegue a ser suficientemente pequeño.

Las ideas anteriores se generalizan y formalizan en el siguiente algoritmo en pseudo código, en el cual se supone que la matriz $\mathbf{A}$ posee un valor propio λ que es mayor en valor absoluto que todos los demás (nótese que esto descarta la posibilidad de que λ sea imaginario, pues en ese caso su conjugado sería también un valor propio y tendría el mismo valor absoluto). Se suponen conocidos la matriz cuadrada $\mathbf{A}$ de orden n , una aproximación inicial $\mathbf{x}_0$ al vector propio que corresponde a λ y la tolerancia ε con que se desea hallar el valor propio λ . En caso de no conocerse una aproximación $\mathbf{x}_0$ adecuada, se puede tomar un vector con todas sus componentes iguales a 1.

```

 $\mathbf{z} := \mathbf{x}_0$ 
 $\lambda_{anterior} := 10^8$  {Se toma un valor arbitrariamente grande para que se pueda evaluar
                        Error en la primera iteración del algoritmo}
repeat
     $\mathbf{w} := \mathbf{A}\mathbf{z}$ 
     $NormaW := 0$ 
     $imax := 0$ 
    for  $i = 1$  to  $n$  {En este lazo se hallan la norma de  $\mathbf{W}$  así como el número  $imax$ 
                    de la componente máxima de  $\mathbf{W}$ }
        if  $|w_i| > NormaW$  then
             $NormaW := |w_i|$ 
             $imax := i$ 
        end
    end
     $\lambda_{aprox} := \frac{w_{imax}}{z_{imax}}$  {Se aproxima el valor propio por el cambio que ha tenido una de
                                las componentes del vector  $\mathbf{z}$  al ser multiplicada por  $\mathbf{A}$ }
     $Error := |\lambda_{aprox} - \lambda_{anterior}|$ 
     $\lambda_{anterior} := \lambda_{aprox}$ 
     $\mathbf{z} := \frac{\mathbf{w}}{NormaW}$ 
until  $Error < \varepsilon$ 
El valor propio de mayor valor absoluto de  $\mathbf{A}$  es  $\lambda_{aprox}$  y  $\mathbf{z}$  es un vector propio asociado.
Terminar

```

Ejemplo 6

Halle con error menor que 0,001 el valor propio de mayor valor absoluto de la matriz $\mathbf{A}$ y un vector propio correspondiente, mediante el método de la potencia.

$$\mathbf{A} = \begin{bmatrix} 6 & 3 & -2 \\ 5 & 1 & 0 \\ 2 & 3 & -4 \end{bmatrix}$$

Solución:

Aplicando a la matriz un programa basado en el algoritmo anterior, se obtienen los resultados que muestra la tabla 1.

Iter	x_1	x_2	x_3	λ	$E_m(\lambda)$
0	1,000000	1,000000	1,000000	-----	-----
1	1,000000	0,857143	0,142857	7,000000	-----
2	1,000000	0,706897	0,482759	8,285714	1,2857
3	1,000000	0,797590	0,306024	7,155172	1,1305
4	1,000000	0,745122	0,407247	7,780723	0,6256
5	1,000000	0,774184	0,351223	7,420873	0,3598
6	1,000000	0,757756	0,383890	7,620107	0,1992
7	1,000000	0,766935	0,365197	7,507489	0,1126
8	1,000000	0,761773	0,375147	7,570412	0,0629
9	1,000000	0,764665	0,369571	7,535025	0,0354
10	1,000000	0,763041	0,372702	7,554853	0,0198
11	1,000000	0,763952	0,370946	7,543720	0,0111
12	1,000000	0,763441	0,371932	7,549964	0,0062
13	1,000000	0,763728	0,371379	7,546460	0,0035
14	1,000000	0,763567	0,371689	7,548426	0,0020
15	1,000000	0,763657	0,371515	7,547322	0,0011
16	1,000000	0,763607	0,371613	7,547941	0,0006

Tabla 1

Entonces, con error menor que 0,001 el mayor valor propio de la matriz $\mathbf{A}$ es $\lambda = 7,5479$ y un vector propio asociado a este valor propio es:

$$\mathbf{x} = \begin{bmatrix} 1,0000 \\ 0,7636 \\ 0,3716 \end{bmatrix}$$

■

Ejercicios

1. A continuación se muestran tres matrices $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$ y tres vectores $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$. Determine de qué matriz o matrices es vector propio cada uno de los vectores. Determine en cada caso a qué valor propio corresponde.

$$\mathbf{A} = \begin{bmatrix} 129 & -26 & 92 \\ 93 & -20 & 66 \\ 150 & -30 & -107 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 182 & -28 & -136 \\ 78 & -8 & 60 \\ 228 & -36 & -170 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} -8 & -4 & 28 \\ -22 & 22 & 22 \\ -40 & 2 & 63 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 5 \\ 3 \\ 6 \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}$$

2. En el ejercicio 1 de la sección anterior apareció el sistema

$$10x_1 - 2x_2 + 3x_3 + 2x_4 = 15$$

$$3x_1 - 10x_2 - 4x_3 + 2x_4 = 4$$

$$5x_1 + 3x_2 + 10x_3 - x_4 = 13$$

$$4x_1 - 3x_2 - 4x_3 + 10x_4 = -2$$

para el cual el método de Jacobi converge, a pesar de que su diagonal no es predominante. Esto no es contradictorio porque la condición de diagonal predominante es suficiente pero no necesaria para la convergencia. Una condición necesaria y suficiente es que el valor propio de mayor valor absoluto de la matriz $\mathbf{M}$ tenga módulo menor que 1. Compruebe que en este sistema esta condición se satisface.

3. Utilice un algoritmo para calcular determinantes basado en el método de Gauss, para trazar un número suficiente de puntos de la gráfica del polinomio característico de la matriz $\mathbf{M}$ correspondiente al sistema del ejercicio anterior. Compruebe que esta matriz posee sus cuatro valores propios reales, dos positivos y dos negativos todos en el intervalo $[-1, 1]$.
4. Cada una de las cuatro matrices que siguen, tienen un valor propio (real) que es el de mayor valor absoluto. Utilice el método de la potencia para hallar este valor propio de máximo valor absoluto y determine en cada caso los otros dos a partir de la propiedad de que la suma de todos los valores propios es igual a la traza de la matriz y su producto igual al determinante de ella.

$$\text{a) } \begin{bmatrix} 3 & 4 & -3 \\ 3 & 5 & 3 \\ 4 & 2 & 8 \end{bmatrix} \quad \text{b) } \begin{bmatrix} -1 & 3 & 2 \\ 1 & -2 & 2 \\ -1 & 3 & 5 \end{bmatrix} \quad \text{c) } \begin{bmatrix} 3 & 1 & -2 \\ 1 & 2 & 3 \\ -2 & 3 & -3 \end{bmatrix} \quad \text{d) } \begin{bmatrix} 3 & 1 & -2 \\ 1 & 2 & -3 \\ 2 & 3 & -3 \end{bmatrix}$$

5. Elabore un algoritmo basado en el método de bisección que permita hallar con una exactitud dada, un valor propio real de una matriz dada que se encuentre comprendido en un intervalo dado. Establezca qué condiciones deben cumplirse para que se garantice el resultado del algoritmo.
6. Elabore un algoritmo basado en el método de las secantes que permita hallar con una exactitud dada, un valor propio real de una matriz dada conociendo dos aproximaciones iniciales del valor propio.

Otras lecturas recomendadas

Para cualquiera de los temas vinculados con los métodos numéricos de los sistemas lineales, es obligada la referencia la obra clásica “Computational Methods of Linear Álgebra” de D. K. Faddeev y V. N. Fadееva, que ha sido impreso en Cuba desde la década del 70.

Para un estudio más pormenorizado sobre normas matriciales es recomendable consultar el libro “Introduction to Numerical Analysis” de K. E. Atkinson, editorial John Wiley and Sons, 1989 y también la obra clásica de E. Isaacson y H. B. Keller, “Analysis of Numerical Methods”, reproducida en Cuba por Ediciones R.

Acerca de los temas relacionados con valores y vectores propios, la parte conceptual está muy bien expuesta en el libro “Multivariable Calculus, Linear Álgebra and Differential Ecuations” de S. I. Grossman. Este mismo autor posee otras obras sobre Álgebra Lineal de gran valor didáctico y conceptual. Sobre el acotamiento de los valores propios en el plano complejo existen resultados muy interesantes, como el teorema de Gerschgorin, que no han sido tratados en el texto y que se encuentran explicados en el texto de Atkinson, antes citado, donde también aparece con detalle los métodos para calcular valores propios basados en las transformaciones con las matrices ortogonales de Householder. El método de Muller para hallar los valores propios reales e imaginarios de matrices reales está desarrollado, al menos en parte, en el libro “Elementary Numerical Analysis” de S. D. Conte, reproducido en Cuba a partir de una edición de McGraw-Hill.

Principales ideas del capítulo

- En la práctica suelen aparecer sistemas lineales de gran tamaño (cientos y miles de ecuaciones) para los cuales se necesita contar con algoritmos eficientes, pues de lo contrario, aun en una rápida computadora, el tiempo de solución podría ser muy largo.
- Los métodos para resolver sistemas de ecuaciones lineales son, en su casi totalidad, de dos tipos: métodos directos y métodos iterativos.
- El método de Gauss es un eficiente método directo en el cual la cantidad de operaciones para resolver un sistema lineal de n ecuaciones y n incógnitas es del orden de $\frac{1}{3}n^3$.
- En el método de Gauss existen tres estrategias de pivote: elemental, parcial y total, de las cuales la más empleada es la parcial porque es fácil de implementar y evita la propagación de los errores de redondeo.
- Para el caso especial de sistemas lineales tridiagonales con la diagonal predominante, el método de Gauss da lugar a un algoritmo muy compacto y eficiente. Este algoritmo solo requiere guardar en memoria los coeficientes no nulos y la cantidad de operaciones que realiza es del orden de $8n$.
- A partir del proceso directo del método de Gauss se elabora un algoritmo para calcular determinantes que resulta muy eficiente pues requiere una cantidad de operaciones de orden $\frac{1}{3}n^3$ para hallar un determinante de orden n .
- A partir de una modificación del método de Gauss se obtiene un algoritmo muy comodo y eficiente para hallar la inversa de una matriz, que consiste, en esencia en resolver simultáneamente n sistemas lineales de n ecuaciones y n incógnitas. Este algoritmo requiere una cantidad de operaciones del orden de $\frac{4}{3}n^3$ para invertir una matriz cuadrada de orden n .

- En el mismo tiempo que se necesita para hallar la inversa de la matriz de un sistema, se pueden resolver cuatro sistemas de ecuaciones de ese mismo tamaño utilizando el método de Gauss.
- Sistemas lineales mal condicionados son aquellos para los cuales pequeños cambios en los coeficientes producen grandes cambios en la solución del sistema.
- La cuantía del mal condicionamiento se mide mediante el número de condición de la matriz del $\mathbf{A}$ sistema que se define como $\|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|$.
- El número de condición es siempre mayor o igual que 1. Valores por debajo de 10 se consideran sistemas bien condicionados. Por encima de 10 ya se empiezan a considerar con mal condicionamiento.
- Una solución $\mathbf{x}_1$ de mala calidad de un sistema lineal se puede mejorar iterativamente resolviendo el sistema: $\mathbf{A} \text{error}(\mathbf{x}_1) = \mathbf{r}_1$.
- Los métodos iterativos de Jacobi y de Seidel son una alternativa útil para resolver sistemas lineales con diagonal predominante.
- El factor de convergencia del método de Jacobi se llama α y el del método de Seidel es β . Ambos son menores que 1 cuando la diagonal del sistema es predominante y en ese caso siempre es $\beta \leq \alpha$, por lo cual la convergencia del método de Seidel suele ser más rápida que la de Jacobi.
- Además de una convergencia más rápida el método de Seidel posee un algoritmo más simple y requiere menos memoria. El método de Jacobi, sin embargo, se adapta muy fácilmente al trabajo con cálculo paralelo.
- Los métodos iterativos son muy apropiados para grandes sistemas de ecuaciones con muchos coeficientes nulos en los cuales todas las ecuaciones obedecen a una misma estructura general.
- Calcular todos los valores propios de una matriz cualquiera es un problema muy complejo. Cuando la matriz es simétrica o cuando solo se desea hallar el valor propio de mayor valor absoluto, existen métodos no muy complicados y eficientes para resolver el problema.
- El método de la potencia es un algoritmo iterativo muy fácil de programar capaz de hallar el valor propio de mayor valor absoluto de una matriz y un vector propio asociado a este valor propio.
- Si se tiene un programa eficiente para calcular el determinante de una matriz numérica, el polinomio característico se puede graficar trazando un suficiente número de puntos aislados de la gráfica y la ecuación característica se puede resolver numéricamente por métodos como bisección y secantes.

Auto examen

1. En la solución de sistemas de ecuaciones lineales, ¿qué significado poseen los términos “métodos directos” y “métodos iterativos”?
2. Dado el siguiente sistema de ecuaciones:

$$\begin{aligned}x_2 + 2x_3 &= -3 \\3x_1 - 2x_2 + x_3 &= 6 \\2x_1 + 3x_2 - 2x_3 &= 5\end{aligned}$$

- a) Expréselo en la forma matricial $\mathbf{Ax} = \mathbf{b}$.
- b) Halle la norma de $\mathbf{A}$ y la norma de $\mathbf{b}$.

- c) Resuelva el sistema mediante el método de Gauss utilizando la estrategia parcial de pivote.
- d) Compruebe si $\mathbf{Ax} = \mathbf{b}$.
- e) Halle la norma de $\mathbf{x}$, analice si $\|\mathbf{A}\| \cdot \|\mathbf{x}\| \geq \|\mathbf{b}\|$ y justifique su respuesta.

3. Dado el sistema de 80 ecuaciones:

$$2x_{i-1} + 8x_i - x_{i+1} = b_i \quad i = 1, 2, \dots, 80$$

donde $x_0 = 6$, $x_{81} = 10$ y los coeficientes b_i ($i = 1, 2, \dots, 80$) son datos,

- a) Determine qué cantidad de operaciones se requiere para resolver el sistema mediante el método de Gauss especializado en sistemas tridiagonales.
 - b) Analice si es posible aplicar a este sistema los métodos iterativos de Jacobi y de Seidel y, de ser así, halle que cantidad de operaciones se necesitará en cada uno para lograr 4 cifras decimales exactas si el error inicial es del orden de 1,5.
4. Halle la inversa de la matriz $\mathbf{A}$ del sistema del ejercicio 2 mediante el método de Gauss y diga el valor del determinante de $\mathbf{A}$.
5. ¿Qué se entiende por un sistema mal condicionado y cómo se mide el mal condicionamiento?
6. Elabore un algoritmo en pseudo código para obtener la solución del sistema del ejercicio 3 con cuatro cifras decimales exactas mediante el método de Seidel.
7. ¿A qué se llama valores propios y vectores propios de una matriz?
8. Sin hallar el polinomio característico de la matriz $\mathbf{C}$ halle varios valores del mismo que le permitan graficarlo en el intervalo en el que están acotados todos los valores propios.

$$\mathbf{C} = \begin{bmatrix} 2 & 3 & -1 \\ 3 & 1 & 1 \\ -1 & 1 & 3 \end{bmatrix}$$

9. Halle el valor propio de mayor valor absoluto de la matriz $\mathbf{C}$.