

TEMA 4º: RESOLUCION DE SISTEMAS DE ECUACIONES

El objetivo de este tema es resolver numéricamente sistemas de ecuaciones lineales que son complicados si los resolvemos por la regla de Cramer.

1. Consideraciones sobre matrices.

Denotaremos por \mathcal{A}_n el espacio vectorial de las matrices cuadradas de orden n sobre \mathbb{R} .

DEFINICION: Un escalar λ diremos que es un autovvalor de una matriz A de orden n si existe un vector $u \in \mathbb{R}^n$ tal que $A \cdot u = \lambda u$, o bien $(A - \lambda I) \cdot u = 0$.

A u se le llama autovector asociado al autovvalor λ .

Fácilmente se prueba que λ es autovvalor de A si, y solo si, $\det(A - \lambda I) = 0$.

Se puede hablar también de autovvalores complejos de A , en cuyo caso $u \in \mathbb{C}^n$. Si $\lambda \in \mathbb{C}$ es un autovvalor de A , su conjugado $\bar{\lambda}$ también es autovvalor de A .

$\det(A - \lambda I) = |A - \lambda I|$ es un polinomio de grado n que se llama polinomio característico asociado a la matriz A .

1.1. PROPOSICION: a) $\det(A - \lambda I) = (-1)^n \lambda^n + (-1)^{n-1} \text{Tr}(A) \lambda^{n-1} + \dots + \det A$
donde $\text{Tr}(A) = \sum_{i=1}^n a_{ii}$.

b) Si A es una matriz triangular entonces
 $\det(A - \lambda I) = (a_{11} - \lambda) \cdot (a_{22} - \lambda) \cdot \dots \cdot (a_{nn} - \lambda)$.

1.2. PROPOSICION: Sea $C \in \mathcal{A}_n$ una matriz no singular, es decir, $\det C \neq 0$.

Si $B = C^{-1}AC$, entonces A y B tienen los mismos autovvalores.

Demostr.: $|B - \lambda I| = 0 \Leftrightarrow |C^{-1}AC - \lambda I| = 0 \Leftrightarrow |C^{-1}AC - \lambda C^{-1}C| = 0 \Leftrightarrow$
 $\Leftrightarrow |C^{-1}(A - \lambda I)C| = 0 \Leftrightarrow |C^{-1}| \cdot |A - \lambda I| \cdot |C| = 0 \Leftrightarrow |A - \lambda I| = 0$. es γ d.

DEFINICIONES: a) Se llama radio espectral de A a
 $\rho(A) = \max_{0 \leq i \leq n} |\lambda_i|$

donde $\lambda_1, \dots, \lambda_n$ son los autovvalores de $A \in \mathcal{A}_n$ en \mathbb{C} .

b) A es simétrica si $A = A^t$.

c) A es definida positiva si $X^t A X > 0$, $\forall x \in \mathbb{R}^n - \{0\}$ donde X es la matriz unocolumna formada por las componentes de x .

d) A es semidefinida positiva si $X^t A X \geq 0, \forall x \in \mathbb{R}^n$.

e) A es ortogonal si $A^t = A^{-1}$.

• Consideremos el espacio vectorial \mathbb{K}^n , donde $\mathbb{K} = \mathbb{R}$ ó \mathbb{C} .

Vamos a definir en \mathbb{K}^n el producto interior siguiente:

$$\forall x, y \in \mathbb{K}^n, (x/y) = X^t \bar{Y},$$

$$\text{es decir, } (x/y) = (x_1, \dots, x_n) \begin{pmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_n \end{pmatrix}$$

DEFINICIÓN: Dada la matriz A con coeficientes en \mathbb{K} definimos la matriz adjunta A^* de A como aquella matriz que verifica

$$(Ax/y) = (x/A^*y), \forall x, y \in \mathbb{K}^n$$

1.3. PROPOSICIÓN: Dada una matriz A de orden n sobre \mathbb{K} , existe una única matriz adjunta de A y es $A^* = \bar{A}^t$

$$\text{Demostr.} - (x/\bar{A}^t y) = X^t (\bar{A}^t y) = X^t A^t \bar{Y} = (AX)^t \bar{Y} = (Ax/y),$$

- Supongamos que B es una matriz adjunta de A ,

$$\text{Entonces } (x/A^*y) = (x/By), \forall (x,y) \in \mathbb{R}^n$$

Luego $X^t \bar{A}^* \bar{Y} = X^t \bar{B} \bar{Y} \Rightarrow X^t (\bar{A}^* - \bar{B}) \bar{Y} = 0$. Siendo esto cierto para todo $x, y \in \mathbb{K}^n$ debe ser $\bar{A}^* - \bar{B} = 0$ y, portanto, $A^* = B$. ~~csqd.~~

pus si $\bar{A}^* - \bar{B} = (c_{ij})$, tomando para cada $i, j \in \{1, \dots, n\}$

$$X^t = (0, \dots, 0, 1, 0, \dots, 0) \text{ e } \bar{Y} = (0, \dots, 0, 1, 0, \dots, 0) \text{ tenemos}$$

$$0 = (0, \dots, 0, 1, 0, \dots, 0) \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = (c_{11} \dots c_{ij} \dots c_{in}) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = c_{ij} \cdot \text{csqd.}$$

DEFINICIÓN: Se dice que A es una matriz hermitiana si $A = A^*$.

1.4. TEOREMA: Si A es una matriz hermitiana, todos sus autovalores son reales.

Demostr.: Basta probar que si λ es autovalor de A , entonces $\bar{\lambda} = \lambda$.

Si λ es autovalor de A existe $u \in \mathbb{K}^n$ tal que $A \cdot u = \lambda u$.

$$\begin{aligned} \text{Entonces } (\lambda u/u) &= (Au/u) = (u/A^*u) = (u/Au) = u^t \bar{A} \cdot \bar{u} = \\ &= (u^t \bar{A} \bar{u})^t = \bar{u}^t \bar{A}^t u = (\bar{A} \bar{u})^t u = (\bar{A} \bar{u}/\bar{u}) = (\bar{\lambda} \bar{u}/\bar{u}) = \\ &= \bar{\lambda} (\bar{u}/\bar{u}) \end{aligned}$$

$$\text{Luego } \lambda (u/u) = \bar{\lambda} (\bar{u}/\bar{u}) \text{ y, siendo } (u/u) = (\bar{u}/\bar{u}), \text{ resulta}$$
$$\lambda = \bar{\lambda} \cdot \text{csqd.}$$

2. NORMAS MATRICIALES.

Sea \mathcal{A}_n el espacio vectorial de las matrices cuadradas de orden n sobre \mathbb{K} .

Una norma en \mathcal{A}_n es una aplicación

$$\|\cdot\| : \mathcal{A}_n \rightarrow \mathbb{R}$$

que verifica las siguientes propiedades

i) $\|A\| \geq 0, \forall A \in \mathcal{A}_n$.

ii) $\|A+B\| \leq \|A\| + \|B\|, \forall A, B \in \mathcal{A}_n$.

iii) $\|\lambda A\| = |\lambda| \cdot \|A\|, \forall \lambda \in \mathbb{K}, \forall A \in \mathcal{A}_n$.

iv) $\|A\| = 0 \Rightarrow A = 0$

Si además se verifica

v) $\|A \cdot B\| \leq \|A\| \cdot \|B\|, \forall A, B \in \mathcal{A}_n$.

se dice que es una norma matricial en \mathcal{A}_n .

DEFINICION: Se dice que una norma matricial $\|\cdot\|_M$ sobre \mathcal{A}_n y una norma vectorial $\|\cdot\|_V$ sobre \mathbb{R}^n son compatibles si se verifica que

$$\|A \cdot x\|_V \leq \|A\|_M \cdot \|x\|_V, \forall A \in \mathcal{A}_n, \forall x \in \mathbb{R}^n.$$

DEFINICION: Dada una norma vectorial $\|\cdot\|_V$ sobre \mathbb{R}^n se define la norma matricial asociada a $\|\cdot\|_V$ como

$$\|A\|_M = \sup_{\|x\|_V=1} \|A \cdot x\|_V = \sup_{x \neq 0} \frac{\|A \cdot x\|_V}{\|x\|_V}$$

2.1. PROPOSICION: Toda norma matricial $\|\cdot\|_M$ inducida por una norma vectorial $\|\cdot\|_V$ es compatible con ella.

Demostr.: $\forall A \in \mathcal{A}_n, \forall x \in \mathbb{R}^n - \{0\}, \frac{\|A \cdot x\|_V}{\|x\|_V} \leq \sup_{x \neq 0} \frac{\|A \cdot x\|_V}{\|x\|_V} = \|A\|_M$

Luego: $\|A \cdot x\|_V \leq \|A\|_M \cdot \|x\|_V, \forall A \in \mathcal{A}_n, \forall x \in \mathbb{R}^n - \{0\}$. c.s.q.d.

2.2. TEOREMA: Dada una norma matricial existen infinitas normas vectoriales compatibles con ella.

Demostr.: Dada una norma matricial $\|\cdot\|_M$ vamos a construir una norma vectorial compatible con ella.

Sea $u \in \mathbb{R}^n$ arbitrario, $u \neq 0$.

Dado $x \in \mathbb{R}^n$ tenemos $x u^t = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} (u_1 \dots u_n) = (x_i u_j)_{i,j=1, \dots, n}$

Definimos entonces $\|\cdot\|_V : x \in \mathbb{R}^n \mapsto \|x\|_V = \|x \cdot u^t\|_M \in \mathbb{R}$.

Veamos que $\|\cdot\|_V$ es una norma vectorial:

i) $\|x\|_V \geq 0, \forall x \in \mathbb{R}^n$

Si $\|x\|_V = 0$, entonces $\|xu^t\|_M = 0$ y por tanto $xu^t = 0$. Debe ser entonces $x = 0$, pues si $\exists i \in \{1, \dots, n\} / x_i \neq 0$, como $u \neq 0$, $\exists j \in \{1, \dots, n\} / u_j \neq 0$ y por tanto $x_i u_j \neq 0$, en contra de que $xu^t = 0$.

ii) $\|x+y\|_V = \|(x+y)u^t\|_M = \|xu^t + yu^t\|_M \leq \|xu^t\|_M + \|yu^t\|_M = \|x\|_V + \|y\|_V$

iii) $\|\lambda x\|_V = \|\lambda xu^t\|_M = |\lambda| \cdot \|xu^t\|_M = |\lambda| \cdot \|x\|_V$.

Problemas ahora que $\|\cdot\|_V$ es compatible con la norma matricial:

$$\forall A \in \mathcal{A}_n, \forall x \in \mathbb{R}^n, \|A \cdot x\|_V = \|A xu^t\|_M \leq \|A\|_{ij} \|xu^t\|_M = \|A\|_M \cdot \|x\|_V.$$

Luego para cada $u \in \mathbb{R}^n \setminus \{0\}$ podemos definir una norma vectorial compatible con la norma matricial. Que hay infinitas normas vectoriales compatibles con $\|\cdot\|_M$ es trivial, si tenemos en cuenta que dado $u \in \mathbb{R}^n \setminus \{0\}$ para cada $\lambda \in \mathbb{R}$ podemos construir por el procedimiento anterior una norma vectorial compatible con $\|\cdot\|_M$ y todas ellas distintas. csgd.

2.3. TEOREMA: Dada una matriz $A \in \mathcal{A}_n$ se tiene que $\rho(A) \leq \|A\|_M$ cualquier que sea la norma matricial $\|\cdot\|_M$.

Demostr.: Sea λ un autovvalor de A , x un autovector asociado a λ y $\|\cdot\|_M$ una norma matricial sobre \mathcal{A}_n .

Sea $\|\cdot\|_V$ una norma vectorial compatible con $\|\cdot\|_M$.

Se tiene que $Ax = \lambda x$ y, por tanto, $\|Ax\|_V = \|\lambda x\|_V = |\lambda| \cdot \|x\|_V$.

Por otra parte, $\|Ax\|_V \leq \|A\|_M \cdot \|x\|_V$.

Luego $|\lambda| \cdot \|x\|_V \leq \|A\|_M \cdot \|x\|_V$

y, por tanto, $|\lambda| \leq \|A\|_M$.

Por tanto $\rho(A) = \{\sup \{|\lambda| / \lambda \text{ autovvalor de } A\}\} \leq \|A\|_M$. csgd.

2.4. TEOREMA: Dada una matriz $A \in \mathcal{A}_n$ y $\varepsilon > 0$ existe una norma matricial $\|\cdot\|_\varepsilon$ tal que

$$\|A\|_\varepsilon \leq \rho(A) + \varepsilon$$

2.5. COROLARIO: Dada una matriz $A \in \mathcal{A}_n$, $\rho(A)$ es el inferior de $\|A\|_M$ cuando $\|\cdot\|_M$ recorre las normas matriciales sobre \mathcal{A}_n .

La demostración es consecuencia inmediata de los Teoremas 2.3 y 2.4.

Ejemplo: Una norma matricial sobre M_n esta definida por

$$\|(a_{ij})\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

Esta norma esta inducida por la norma vectorial $\|(x_1, \dots, x_n)\|_1 = \sum_{i=1}^n |x_i|$.

3. SUCESIONES MATRICIALES.

DEFINICION: Sea $\{A_r\}_{r \in \mathbb{N}}$ una sucesion de matrices cuadradas de orden n . Diremos que

$$\lim_{r \rightarrow \infty} A_r = A \quad \text{si y solo si} \quad \lim_{r \rightarrow \infty} \|A_r - A\| = 0$$

siendo $\|\cdot\|$ una norma matricial en M_n cualquiera.

Veamos que esta definicion es buena, es decir, que la convergencia de $\{A_r\}_r$ a A no depende de la norma matricial.

Toda matriz $A \in M_n$ tiene n^2 elementos. Por tanto podemos identificar M_n con $\mathbb{R}^{(n^2)}$. Entonces, toda norma matricial en M_n es una norma en \mathbb{R}^{n^2} . Siendo \mathbb{R}^{n^2} de dimension finita, todas las normas sobre \mathbb{R}^{n^2} son equivalentes. Se deduce de aqui que todas las normas matriciales sobre M_n son equivalentes. En consecuencia, si $\|\cdot\|$ y $\|\cdot\|_*$ son normas matriciales sobre M_n se tiene que $\lim_{r \rightarrow \infty} \|A_r - A\| = 0$ si y solo si $\lim_{r \rightarrow \infty} \|A_r - A\|_* = 0$, como fuieramos ver.

Podemos poner entonces que

$$\lim_{r \rightarrow \infty} A_r = A \quad \text{def} \quad \lim_{r \rightarrow \infty} \|A_r - A\| = 0$$

$$\text{Pero } \lim_{r \rightarrow \infty} \|A_r - A\| = 0 \Leftrightarrow \lim_{r \rightarrow \infty} \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}^r - a_{ij}| = 0 \Leftrightarrow$$

$$\Leftrightarrow \lim_{r \rightarrow \infty} |a_{ij}^r - a_{ij}| = 0, \forall i, j \in \{1, \dots, n\} \Leftrightarrow \lim_{r \rightarrow \infty} a_{ij}^r = a_{ij}, \forall i, j \in \{1, \dots, n\}.$$

3.1. TEOREMA: Sea $A \in M_n$. Es condicion necesaria y suficiente para que $\lim_{r \rightarrow \infty} A^r = 0$ que $\rho(A) < 1$.

Demostr.: \Rightarrow Supongamos que $\rho(A) \geq 1$. Entonces existe un autovalor λ de A tal que $|\lambda| \geq 1$.

Sea u un autovector de A asociado a λ .

Entonces $Au = \lambda u$ y, por tanto, $A^r u = A^{r-1} Au = \lambda A^{r-1} u = \dots = \lambda^r u$.

Sea $\|\cdot\|_M$ una norma matricial sobre M_n y $\|\cdot\|_V$ una norma vectorial.

A) SISTEMAS TRIANGULARES:

Son sistemas del tipo (I) de modo que $a_{ij} = 0$ si $j < i$. Se obtiene así un sistema triangular superior.

La matriz del sistema es

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & 0 & \dots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Si $a_{ij} = 0$, cuando $i < j$, se dice que el sistema es triangular inferior.

Para un sistema triangular se verifica que el determinante del sistema es el producto de los elementos de la diagonal, es decir

$$|A| = \prod_{i=1}^n a_{ii}$$

Para resolver un sistema triangular superior calculamos x_n en la n -ésima ecuación, $x_n = \frac{b_n}{a_{nn}}$

Conocido x_n , calculamos x_{n-1} en la $(n-1)$ -ésima ecuación y tenemos:

$$x_{n-1} = \frac{1}{a_{n-1,n-1}} (b_{n-1} - a_{n-1,n} x_n)$$

Así sucesivamente tendremos

$$x_k = \frac{1}{a_{kk}} \left(b_k - \sum_{r=k+1}^n a_{kr} x_r \right), \quad k = n-1, n-2, \dots, 2, 1.$$

B) METODO DE GAUSS:

Se trata de reducir un sistema de n ecuaciones lineales con n incógnitas a un sistema triangular superior.

Sea $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ la matriz del sistema, y suponemos $|A| \neq 0$.

Si fuese $a_{11} = 0$ cambiando las filas convenientemente obtendríamos una matriz $A^{(1)} = (a_{ij}^{(1)}) \in \mathbb{R}^{n \times n}$ tal que $a_{11}^{(1)} \neq 0$.

A $a_{11}^{(1)}$ le llamaremos pivote.

Definimos:

$$- u_{i1} = \frac{-a_{i1}^{(1)}}{a_{11}^{(1)}}, \quad i \geq 2.$$

$$- a_{ij}^{(2)} = a_{ij}^{(1)} + u_{i1} a_{1j}^{(1)}, \quad i \geq 2, j \leq n$$

$$- b_i^{(2)} = b_i^{(1)} + u_{i1} b_1^{(1)} \quad i \geq 2$$

De esta forma obtenemos a partir de $A^{(1)}$ una matriz $A^{(2)}$ de la forma

$$A^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2n}^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & a_{n2}^{(2)} & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix}$$

y la columna de los términos independientes es

$$B^{(2)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix}$$

Si definimos

$$M^{(1)} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ u_{21} & 1 & 0 & \dots & 0 \\ u_{31} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{n1} & 0 & 0 & \dots & 1 \end{pmatrix}$$

tenemos que $A^{(2)} = M^{(1)} A^{(1)}$, y $B^{(2)} = M^{(1)} B^{(1)}$

Además $|A^{(2)}| = |M^{(1)}| \cdot |A^{(1)}| = |A^{(1)}| = |A|$ y se verifica que

$$AX = B \Leftrightarrow A^{(1)}X = B^{(1)} \Leftrightarrow A^{(2)}X = B^{(2)}$$

$$\text{pues } A^{(2)}X = B^{(2)} \Leftrightarrow M^{(1)}A^{(1)}X = M^{(1)}B^{(1)} \Leftrightarrow A^{(1)}X = B^{(1)}$$

Supongamos que repitiendo este procedimiento hemos obtenido la matriz

$$A^{(r)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1r}^{(1)} & a_{1r+1}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2r}^{(2)} & a_{2r+1}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{rr}^{(r)} & a_{rr+1}^{(r)} & \dots & a_{rn}^{(r)} \\ 0 & 0 & \dots & a_{nr}^{(r)} & a_{nr+1}^{(r)} & \dots & a_{nn}^{(r)} \end{pmatrix}$$

$$\text{Entonces } |A^{(r)}| = a_{11}^{(1)} a_{22}^{(2)} \dots a_{(r-1)(r-1)}^{(r-1)} \cdot |A_*^{(r)}|$$

$$\text{donde } A_*^{(r)} = \begin{pmatrix} a_{rr}^{(r)} & a_{rr+1}^{(r)} & \dots & a_{rn}^{(r)} \\ a_{nr}^{(r)} & a_{nr+1}^{(r)} & \dots & a_{nn}^{(r)} \end{pmatrix}$$

Como $|A^{(r)}| \neq 0$, $|A_*^{(r)}| \neq 0$, y por tanto, existe $a_{ir}^{(r)} \neq 0$.

Para pasar a la matriz $A^{(r+1)}$ definiremos

$$- u_{ir} = \frac{-a_{ir}^{(r)}}{a_{rr}^{(r)}}, \quad i \geq r+1$$

$$- a_{ij}^{(r+1)} = a_{ij}^{(r)} + u_{ir} a_{rj}^{(r)}, \quad i \geq r+1, j \leq n.$$

$$- b_i^{(r+1)} = b_i^{(r)} + u_{ir} b_r^{(r)}, \quad i \geq r+1$$

Si definimos

$$M^{(r)} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & u_{r+1r} & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & u_{nr} & \dots & 1 \\ 0 & 0 & 0 & \dots & u_{nr} & \dots & 1 \end{pmatrix} \quad (1)$$

Se verifica que $A^{(r+1)} = M^{(r)} A^{(r)}$

Cuando se elige el pivote en la i -ésima columna se dice que la

elección es parcial; si se elige entre los elementos de toda la matriz se dice que la elección es total.

- Supongamos que podemos aplicar el método de Gauss sin permutar filas ni columnas. Entonces $A^{(1)} = A$ y $B^{(1)} = B$

Como $A^{(n)} = M^{(n-1)} M^{(n-2)} \dots M^{(1)} A^{(1)}$

tenemos que si $L = (M^{(n-1)} M^{(n-2)} \dots M^{(1)})^{-1}$ y $U = A^{(n)}$ entonces $A = L \cdot U$

U es una matriz triangular superior.

Siendo $L = [M^{(1)}]^{-1} \cdot [M^{(2)}]^{-1} \dots [M^{(n-1)}]^{-1}$ y $M^{(r)}$ de la forma (1)

se deduce que

$$[M^{(r)}]^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

y por tanto

$$L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -u_{21} & 1 & 0 & \dots & 0 & 0 \\ -u_{31} & -u_{32} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -u_{n1} & -u_{n2} & -u_{n3} & \dots & -u_{n(n-1)} & 1 \end{pmatrix}$$

Luego la matriz A ha quedado descompuesta en producto de dos matrices triangulares L y U, una triangular superior y otra triangular inferior donde los elementos de la diagonal son iguales a 1.

Damos a continuación dos teoremas sin demostración.

4.1. TEOREMA: Si una matriz A de orden n regular ($|A| \neq 0$) se puede descomponer como producto de dos matrices triangulares, una triangular inferior con diagonal unidad y otra triangular superior, entonces dicha descomposición es única.

DEFINICION: Dada una matriz $B \in \mathbb{R}^{n \times n}$ se llama menor fundamental de orden k a la matriz $B_k = (b_{ij})$, $i, j \in \{1, 2, \dots, k\}$.

4.2. TEOREMA: Es condición necesaria y suficiente para que la matriz $A \in \mathbb{R}^{n \times n}$ se pueda descomponer como producto de una matriz triangular superior y otra inferior con diagonal unidad y tales menores fundamentales tengan determinante distinto de cero.

DEFINICION Se llama matriz de permutación P_{lk} a la matriz que resulta de permutar la fila l por la fila k de la matriz unidad.

Ejemplo: $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. $P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

$$P_{23} A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Es decir, $P_{lk} A$ es una matriz que se obtiene a partir de A cambiando entre sí las filas l y k .

$$A \cdot P_{23} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31} & a_{33} & a_{32} \end{pmatrix}$$

Es decir, $A P_{lk}$ es una matriz que se obtiene a partir de A cambiando entre sí las columnas l y k .

Observese que $|P_{lk}| = -1$. Por tanto, $|A P_{lk}| = |P_{lk} A| = -|A|$

• Supongamos que estamos en la etapa k -ésima del método de Gauss y queremos pasar a la etapa $(k+1)$ -ésima.

Tenemos entonces que $A^{(k)} X = B^{(k)}$

Si para pasar a $A^{(k+1)}$ tenemos que permutar filas multiplicamos por la izquierda por la matriz de permutación: $P^{(k)} A^{(k)} X = P^{(k)} B^{(k)}$, y definimos

$$A^{(k+1)} = M^{(k)} P^{(k)} A^{(k)} \quad \text{y} \quad B^{(k+1)} = M^{(k)} P^{(k)} B^{(k)}$$

Tendríamos entonces que

$$A^{(n)} = M^{(n-1)} P^{(n-1)} M^{(n-2)} P^{(n-2)} \dots M^{(1)} P^{(1)} A^{(1)}$$

donde, eventualmente, alguna $P^{(r)}$ puede ser la matriz unidad, cuando en la etapa r no hay que permutar filas.

Se tiene entonces que

$$|A^{(n)}| = \left(\prod_{i=1}^{n-1} |M^{(i)}| \cdot |P^{(i)}| \right) \cdot |A^{(1)}|$$

Como $|M^{(i)}| = 1$, se tiene que $|A^{(n)}| = |A^{(1)}|$ ó $|A^{(n)}| = -|A^{(1)}|$ o de otra forma $|A^{(1)}| = (-1)^m |A^{(n)}|$ donde $m \leq n$ es el número de etapas en las que hemos de efectuar una permutación de filas.

Supongamos que para pasar de la etapa k a la etapa $k+1$ hay que permutar columnas

Si tenemos $A^{(k)} X = B^{(k)}$, como $P^{(k)} P^{(k)} = I$ se tiene $(A^{(k)} P^{(k)}) (P^{(k)} X) = B^{(k)}$.

Definimos entonces $A^{(k+1)} = M^{(k)} A^{(k)} P^{(k)}$,

Entonces $A^{(n)} = M^{(n-1)} M^{(n-2)} \dots M^{(4)} A^{(1)} P^{(1)} \dots P^{(n-2)} P^{(n-1)}$

Por tanto, $|A^{(1)}| = (-1)^{u'} |A^{(n)}|$ donde u' es el número de permutaciones de columnas efectuadas.

EJEMPLO: (Cálculo de la matriz inversa)

Sea $A \in \mathbb{A}^n$ una matriz de determinante no nulo. Se trata de hallar C tal que $AC = I$.

Sea $C = (C_1, \dots, C_n) \in \mathbb{A}^n$ donde $C_i = \begin{pmatrix} c_{1i} \\ \vdots \\ c_{ni} \end{pmatrix}$

Entonces $AC = I \Leftrightarrow A(C_1 \dots C_n) = (\delta_1 \dots \delta_n)$

donde $\delta_i = \begin{pmatrix} \delta_{1i} \\ \vdots \\ \delta_{ni} \end{pmatrix}$ siendo $\delta_{ij} = 1$ si $i=j$ y $\delta_{ij} = 0$ si $i \neq j$.

Luego encontrar C es equivalente a resolver el sistema

$$\begin{cases} AC_1 = \delta_1 \\ \vdots \\ AC_n = \delta_n \end{cases}$$

Hay que resolver n sistemas de n ecuaciones lineales con n incógnitas. Para n suficientemente grande es más sencillo calcular la matriz inversa resolviendo estos sistemas por el método expuesto que resolverla directamente.

5. Descomposición de una matriz cuadrada como producto de dos matrices triangulares una inferior y otra superior.

En el caso de que una matriz A se pueda descomponer como producto de dos matrices triangulares, una inferior L y otra superior U , $A = L \cdot U$, la resolución del sistema $AX = B$ se puede hacer resolviendo los sistemas triangulares $LY = B$ y $UX = Y$, pues

$$(UX = Y) \wedge (LY = B) \Rightarrow (LUX = B) \Rightarrow (AX = B)$$

Trataremos a continuación de obtener las matrices L y U .

A) Sean las matrices

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}, \quad L = \begin{pmatrix} l_{11} & & & & 0 \\ l_{21} & l_{22} & & & \\ \vdots & \vdots & \ddots & & \\ l_{n1} & l_{n2} & \dots & l_{nn} & \end{pmatrix} \quad \text{y} \quad U = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ & u_{22} & \dots & u_{2n} \\ & & \ddots & \\ 0 & & & u_{nn} \end{pmatrix}$$

Entonces

$$LU = \begin{pmatrix} l_{11}u_{11} & l_{11}u_{12} & l_{11}u_{13} & \dots & l_{11}u_{1n} \\ l_{21}u_{11} & \sum_{i=1}^2 l_{2i}u_{i2} & \sum_{i=1}^2 l_{2i}u_{i3} & \dots & \sum_{i=1}^2 l_{2i}u_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1}u_{11} & \sum_{i=1}^{n-1} l_{ni}u_{i2} & \sum_{i=1}^{n-1} l_{ni}u_{i3} & \dots & \sum_{i=1}^{n-1} l_{ni}u_{in} \end{pmatrix}$$

La matriz $\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{pmatrix}$ representa el número de sumandos que hay en

cada término de L.U. Si ha de ser $A = L \cdot U$ se tendrá que

$$(1) \quad a_{11} = l_{11} u_{11}$$

$$(2) \quad a_{1j} = l_{11} u_{1j}, \quad j \geq 2$$

$$(3) \quad a_{i1} = l_{i1} u_{11}, \quad i \geq 2$$

$$(4) \quad a_{kk} = \sum_{r=1}^k l_{kr} u_{rk} = l_{kk} u_{kk} + \sum_{r=1}^{k-1} l_{kr} u_{rk} \quad \text{y por tanto}$$

$$l_{kk} u_{kk} = a_{kk} - \sum_{r=1}^{k-1} l_{kr} u_{rk}, \quad k \geq 2.$$

$$(5) \quad \text{Si } i > k \geq 2, \quad a_{ik} = \sum_{r=1}^k l_{ir} u_{rk} = l_{ik} u_{kk} + \sum_{r=1}^{k-1} l_{ir} u_{rk} \quad \text{y por tanto}$$

$$l_{ik} = \frac{1}{u_{kk}} \left(a_{ik} - \sum_{r=1}^{k-1} l_{ir} u_{rk} \right)$$

$$(6) \quad \text{Si } j > k \geq 2, \quad a_{kj} = \sum_{r=1}^k l_{kr} u_{rj} = l_{kk} u_{kj} + \sum_{r=1}^{k-1} l_{kr} u_{rj} \quad \text{y por tanto}$$

$$u_{kj} = \frac{1}{l_{kk}} \left(a_{kj} - \sum_{r=1}^{k-1} l_{kr} u_{rj} \right).$$

Entonces, para calcular los coeficientes de las matrices L y U podemos seguir el proceso siguiente:

- Tomamos $l_{11} \neq 0$ arbitrario y hacemos $u_{11} = \frac{a_{11}}{l_{11}}$.

Entonces $u_{1j} = \frac{a_{1j}}{l_{11}} \quad j \geq 2$ y, $l_{i1} = \frac{a_{i1}}{u_{11}}, \quad i \geq 2$, supuesto $u_{11} \neq 0$.

- Tomamos $l_{22} \neq 0$ arbitrario y hacemos $u_{22} = \frac{1}{l_{22}} (a_{22} - l_{21} u_{12})$

Entonces $u_{2j} = \frac{1}{l_{22}} (a_{2j} - l_{21} u_{1j}) \quad j \geq 3$

y $l_{i2} = \frac{1}{u_{22}} (a_{i2} - l_{i1} u_{12}), \quad i \geq 3$, supuesto $u_{22} \neq 0$.

- Así sucesivamente tomamos $l_{kk} \neq 0$ arbitrario y hacemos $u_{kk} = \frac{1}{l_{kk}} \left(a_{kk} - \sum_{r=1}^{k-1} l_{kr} u_{rk} \right)$

Entonces $u_{kj} = \frac{1}{l_{kk}} \left(a_{kj} - \sum_{r=1}^{k-1} l_{kr} u_{rj} \right), \quad j \geq k+1$

y $l_{ik} = \frac{1}{u_{kk}} \left(a_{ik} - \sum_{r=1}^{k-1} l_{ir} u_{rk} \right), \quad i \geq k+1$, supuesto $u_{kk} \neq 0$.

OBSERVACIONES: 1) Si $a_{11} = 0$, entonces $u_{11} = 0$ y por tanto $|U| = 0$. Entonces sería $|A| = 0$, y el proceso anterior no "serviría".

2) Si para algún $k \geq 2$ fuese $a_{kk} - \sum_{r=1}^{k-1} l_{kr} u_{rk} = 0$ se tendría $u_{kk} = 0$, y el proceso anterior para calcular L y U tampoco serviría.

Para solucionar estos problemas veamos el siguiente teorema.

5.1. TEOREMA: Si $|A| \neq 0$ entonces existe una permutación de filas de la matriz A de forma que la matriz resultante se puede descomponer como producto de una matriz triangular inferior y una matriz triangular superior.

Demostri: Supongamos, en primer lugar, que podemos aplicar el método de Gauss a la matriz A sin permutar filas ni columnas.

Con las notaciones del apartado 4, si $L = [M^{(n-1)} \dots M^{(1)}]^{-1}$, entonces L es triangular inferior y $U = L^{-1}A$ es triangular superior, y se tiene que $A = LU$. En este caso L y U se pueden determinar por el proceso visto anteriormente.

Veamos ahora que las permutaciones que hacíamos en cada etapa del método de Gauss se pueden hacer al principio de una sola vez.

Sea A la matriz de orden n (a_{ij}^1) .

Sea $a_{i_1 1}^1$ el elemento de mayor módulo de la primera columna de la matriz A , que es distinto de cero, pues $|A| \neq 0$. Dejamos invariante la fila i_1 , y utilizando $a_{i_1 1}^1$ como pivote anulamos los restantes elementos de la primera columna. Obtenemos así una matriz

$$\begin{pmatrix} 0 & a_{12}^2 & \dots & a_{1n}^2 \\ \vdots & \vdots & \dots & \vdots \\ 0 & a_{i_1-1,2}^2 & \dots & a_{i_1-1,n}^2 \\ a_{i_1 1}^1 & a_{i_1 2}^1 & \dots & a_{i_1 n}^1 \\ \vdots & \vdots & \dots & \vdots \\ 0 & a_{n2}^2 & \dots & a_{nn}^2 \end{pmatrix}$$

Suprimimos en esta matriz la columna 1 y la fila i_1 . En la submatriz resultante, sea $a_{i_2 2}^2$ el elemento de mayor módulo de la primera columna. Dejando esta fila invariante y utilizando $a_{i_2 2}^2 \neq 0$ como pivote obtenemos la siguiente matriz

$$\begin{pmatrix} 0 & 0 & a_{13}^3 & \dots & a_{1n}^3 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & a_{i_2 2}^2 & a_{i_2 3}^2 & \dots & a_{i_2 n}^2 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & a_{i_1 2}^1 & a_{i_1 3}^1 & \dots & a_{i_1 n}^1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & a_{n3}^3 & \dots & a_{nn}^3 \end{pmatrix}$$

En la submatriz considerada, suprimiendo la primera columna y la fila i_2 , obtenemos otra submatriz a la cual aplicaremos de nuevo procedimiento.

Si repetimos repetido n veces este procedimiento, y la matriz resultante

Permutamos las filas de forma que i_1 quede en la primera fila, i_2 en la segunda, ..., i_n en la n -ésima, obtenemos una matriz de la forma

$$\begin{pmatrix} a_{i_1 i_1}^1 & a_{i_1 i_2}^1 & a_{i_1 i_3}^1 & \dots & a_{i_1 i_n}^1 \\ & a_{i_2 i_2}^2 & a_{i_2 i_3}^2 & \dots & a_{i_2 i_n}^2 \\ & & a_{i_3 i_3}^3 & \dots & a_{i_3 i_n}^3 \\ & & & \dots & \vdots \\ & & & & a_{i_n i_n}^n \end{pmatrix}$$

Por tanto, si en la matriz A permutamos todas las filas de forma que la i_1 quede en primer lugar, ..., la i_n quede en n -ésimo lugar, al aplicar el método de Gauss obtenemos la matriz triangular superior directamente, es qd.

B) METODO DE CHOLESKY PARA MATRICES SIMÉTRICAS DEFINIDAS POSITIVAS.

Se prueba que una matriz $A = (a_{ij})$ es definida positiva si y solo si los menores principales son no nulos, es decir si

$$\forall j \in \{1, \dots, n\}, \begin{vmatrix} a_{11} & \dots & a_{1j} \\ \vdots & \ddots & \vdots \\ a_{j1} & \dots & a_{jj} \end{vmatrix} \neq 0$$

Supongamos que $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ es una matriz simétrica definida positiva. Probaremos después que existe una matriz triangular inferior L con los elementos de la diagonal positivos tal que $A = LL^t$; aun más, probaremos que esta condición caracteriza a las matrices simétricas definidas positivas.

Sea entonces $A = LL^t$, es decir

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & \\ & & \ddots & & \\ l_{n1} & l_{n2} & \dots & & l_{nn} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & \dots & l_{n1} \\ & l_{22} & \dots & l_{n2} \\ & & \ddots & \vdots \\ 0 & & & l_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} l_{11}^2 & l_{11}l_{21} & \dots & l_{11}l_{n1} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 & \dots & l_{21}l_{n1} + l_{22}l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1}l_{11} & l_{n1}l_{21} + l_{n2}l_{22} & \dots & l_{n1}^2 + \dots + l_{nm}^2 \end{pmatrix}$$

La matriz $\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{pmatrix}$ representa en cada elemento el número de

de cada elemento de la matriz anterior.

Por un razonamiento análogo al que se hizo en el apartado A) tenemos

$$(I) \begin{cases} l_{11} = \sqrt{a_{11}} \\ l_{i1} = \frac{a_{i1}}{\sqrt{a_{11}}} \quad i \geq 2 \\ l_{kk} = \sqrt{a_{kk} - \sum_{r=1}^{k-1} l_{kr}^2} \quad k \geq 2 \\ l_{ik} = \frac{1}{l_{kk}} (a_{ik} - \sum_{r=1}^{k-1} l_{ir} l_{kr}) \quad k \geq 2 \end{cases}$$

Problemas ahora la siguiente

5.2. PROPOSICIÓN: Una matriz A de orden n es una matriz simétrica definida positiva si, y solo si, $A = L \cdot L^t$ donde L es una matriz triangular inferior con los elementos de la diagonal positivos.

Demostri: Las relaciones (I) tienen sentido cuando (a_{ij}) es una matriz simétrica definida positiva. Por tanto, la condición necesaria fue ya probada con lo dicho anteriormente.

Problemas ahora fue la condición también es suficiente.

Supongamos que $A = LL^t$ donde L es una matriz triangular inferior con los elementos de la diagonal positivos. Entonces

$$A^t = (LL^t)^t = (L^t)^t L^t = LL^t = A, \text{ que prueba que } A \text{ es simétrica.}$$

Sea $x \in \mathbb{R}^n$. Entonces

$$(Ax/x) = x^t A^t x = x^t A x = x^t L L^t x = (L^t x)^t L^t x = y^t y \text{ donde } y = L^t x.$$

Pero $y^t y \geq 0$, pues si $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $y^t y = \sum_{i=1}^n y_i^2$.

Luego $(Ax/x) \geq 0, \forall x \in \mathbb{R}^n$.

Además, $y^t y = 0 \Leftrightarrow L^t x = 0$.

Como L es triangular inferior con los elementos de la diagonal positivos se tiene que $|L| = \prod_{i=1}^n l_{ii} > 0$. Luego $|L^t| > 0$ y por tanto existe $(L^t)^{-1}$.

En consecuencia, $L^t x = 0 \Leftrightarrow x = 0$.

Luego, $(Ax/x) = 0 \Leftrightarrow x = 0$.

En definitiva, A es simétrica definida positiva. c.q.d.

Veamos algunas aplicaciones del Método de Cholesky para matrices simétricas definidas positivas:

① Sea $A \in \mathbb{R}^n$ una matriz simétrica. Para averiguar si es definida positiva se resuelve el sistema (I). Si se puede resolver sin que surja una división por cero, A será definida positiva y ya tendremos A como producto de una matriz triangular inferior y su transpuesta.

transpuesta.

② Si A es simétrica definida positiva, para resolver el sistema $Ax=B$ se resuelven los sistemas triangulares

$$\begin{cases} Ly = B \\ L^t x = y \end{cases}$$

Resolveremos primero $Ly=B$ y luego $L^t x=y$. Entonces

$$L^t x = y \Rightarrow \begin{cases} LL^t x = Ly \\ Ly = B \end{cases} \Rightarrow LL^t x = B \Rightarrow Ax = B.$$

③ Si $A=LL^t$, $A^{-1}=(LL^t)^{-1}=(L^t)^{-1}(L^{-1})=(L^{-1})^t(L^{-1})$. Por tanto, para calcular la matriz inversa de A basta calcular la matriz inversa de una matriz triangular inferior L.

④ Si $A=LL^t$, entonces $|A|=|LL^t|=|L| \cdot |L^t|=|L| \cdot |L| = \prod_{i=1}^n l_{ii}^2$.

⑤ Velocidad del método: Para hallar los l_{ik} hay que efectuar $\frac{2n^3+3n^2+n}{6}$ operaciones y para resolver $Ax=B$ por este método hay que efectuar $2n^2$ operaciones. Para n suficientemente grande el número de operaciones que hay que efectuar es $n^3/3$. Por el método de Gauss habría que hacer $\frac{4n^3+9n^2-7n}{6}$ operaciones, que para n grande es aproximadamente $2n^3/3$, que es el doble del anterior.

6. Error y residual de una solución aproximada.

Hasta ahora hemos estado buscando soluciones para la ecuación $Ax=B$, por medio de métodos finitos y casi siempre con un error debido a la imperfección de la máquina utilizada.

Si x es la solución del sistema y \hat{x} es la solución que nosotros obtenemos a $e = x - \hat{x}$ se le llama error de la solución aproximada y a $r = Ax - A\hat{x} = B - A\hat{x}$ se le llama residual de la solución aproximada.

Por ejemplo, consideremos el sistema

$$\begin{cases} 1.01x_1 + 0.99x_2 = 2 \\ 0.99x_1 + 1.01x_2 = 2 \end{cases}$$

La solución exacta es $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Supongamos que obtenemos la solución $\hat{x} = \begin{pmatrix} 1.01 \\ 1.01 \end{pmatrix}$. Entonces

$$e = \begin{pmatrix} -0.01 \\ -0.01 \end{pmatrix} \quad \text{y} \quad r = \begin{pmatrix} -0.02 \\ -0.02 \end{pmatrix}$$

Como vemos e y r son pequeños.

Si $\hat{x} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $e = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ y $r = \begin{pmatrix} -0.02 \\ 0.02 \end{pmatrix}$; el residual es pequeño y el error es grande.

Si tenemos el sistema

$$\begin{cases} 1.01x_1 + 0.99x_2 = 2 \\ 0.99x_1 + 1.01x_2 = -2 \end{cases}$$

la solución exacta es $x = \begin{pmatrix} 100 \\ -100 \end{pmatrix}$

Si $\hat{x} = \begin{pmatrix} 1.01 \\ -0.99 \end{pmatrix}$ se obtiene un error $e = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$, que es pequeño respecto de la solución, y un residual $r = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$, que es grande respecto de B. Veamos que esto depende de la matriz utilizada.

Tenemos $r = Ax - A\hat{x} = A(x - \hat{x}) = Ae$.

Luego $\|r\| = \|Ae\| \leq \|A\| \cdot \|e\|$, de donde $\|e\| \geq \frac{\|r\|}{\|A\|}$

Por otra parte, $e = A^{-1}r \Rightarrow \|e\| \leq \|A^{-1}\| \cdot \|r\|$

Obtenemos para e la acotación

$$\frac{\|r\|}{\|A\|} \leq \|e\| \leq \|A^{-1}\| \cdot \|r\| \quad (1)$$

Por ser $Ax = B$ se tiene que $\|B\| \leq \|A\| \cdot \|x\|$, y por tanto $\|x\| \geq \frac{\|B\|}{\|A\|}$.

Además $x = A^{-1}B \Rightarrow \|x\| \leq \|A^{-1}\| \cdot \|B\|$. Luego, obtenemos para x la acotación

$$\frac{\|B\|}{\|A\|} \leq \|x\| \leq \|A^{-1}\| \cdot \|B\| \quad (2)$$

Combinando (1) y (2) podemos acotar $\frac{\|e\|}{\|x\|}$:

$$\frac{1}{\|A\| \cdot \|A^{-1}\|} \cdot \frac{\|r\|}{\|B\|} \leq \frac{\|e\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|r\|}{\|B\|}$$

A $\frac{\|e\|}{\|x\|}$ se le llama error relativo y a $\frac{\|r\|}{\|B\|}$ residual relativo.

Sea $C(A) = \|A\| \cdot \|A^{-1}\|$. A $C(A)$ le llamaremos número de condición de A.

Entonces
$$\frac{1}{C(A)} \frac{\|r\|}{\|B\|} \leq \frac{\|e\|}{\|x\|} \leq C(A) \frac{\|r\|}{\|B\|}$$

Queremos ver cuando es pequeño el error si lo es el residual.

Si $C(A)$ está muy próximo a 1, $\frac{1}{C(A)}$ y $C(A)$ son "casi iguales" y por tanto $\frac{\|e\|}{\|x\|}$ varía en un intervalo muy pequeño. Si además $\frac{\|r\|}{\|B\|}$ es muy pequeño, $\frac{\|e\|}{\|x\|}$ está "muy próximo a cero".

Veamos que $C(A) \geq 1$ para cualquier matriz invertible:

$$C(A) = \|A\| \cdot \|A^{-1}\| \geq \|AA^{-1}\| = \|I\| \geq \rho(I) = 1.$$

Vamos a enunciar un teorema que nos da una acotación de $C(A)$, sin tener que calcular A^{-1} .

G.1. TEOREMA: Para cualquier matriz $A \in \mathbb{R}^n$ invertible y para cualquier norma matricial se tiene que $\frac{1}{C(A)} = \min \left\{ \frac{\|A - E\|}{\|A\|} \mid E \text{ es una matriz elemental} \right\}$.

Entonces si tenemos el sistema $Ax=B$ y C es una matriz no invertible se tiene que

$$\frac{1}{C(A)} \leq \frac{\|A-C\|}{\|A\|}$$

y por tanto, $C(A) \geq \frac{\|A\|}{\|A-C\|}$. Si $\frac{\|A\|}{\|A-C\|}$ es "grande", $C(A)$ no valdrá para calcular $\frac{\|e\|}{\|x\|}$.

6.7. COROLARIO: Si $A \in \mathbb{R}^{n \times n}$ es invertible y C es una matriz tal que $\|A-C\| < \frac{1}{\|A^{-1}\|}$ entonces C es invertible.

Demostr.: $\frac{1}{C(A)} = \frac{1}{\|A\| \cdot \|A^{-1}\|} \leq \frac{\|A-B\|}{\|A\|}$ si B es no invertible.

Luego para toda matriz B no invertible se cumple que $\frac{1}{\|A^{-1}\|} \leq \|A-B\|$.

Por tanto, C no puede ser no invertible, pues $\|A-C\| < \frac{1}{\|A^{-1}\|}$. c.q.d.

TEMA 5º: METODOS ITERATIVOS DE RESOLUCION DE SISTEMAS LINEALES

Trataremos en este tema de resolver el sistema $Ax=b$ donde $A \in \mathbb{A}^n$, $|A| \neq 0$ y $b \in \mathbb{R}^n$. Construiremos una función $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ y una sucesión $\{x^m\}_{m \in \mathbb{N}}$ donde x^0 es arbitrario y $x^m = F(x^{m-1})$ de forma que $\{x^m\}_m$ converja a la solución x del sistema.

Consideraremos funciones F de la forma $F(x) = Bx + C$ donde $B \in \mathbb{A}^n$ y $C \in \mathbb{R}^n$.

Por tanto tendremos $x^m = Bx^{m-1} + C$; diremos que el método así definido es convergente con respecto al problema de resolver el sistema $Ax=b$ si $\lim_{m \rightarrow \infty} x^m = x = A^{-1}b$ cualquiera que sea $x^0 \in \mathbb{R}^n$.

1. Condiciones de convergencia del método definido anteriormente.

Si la función F tomada es contractiva, es decir, si

$$\exists L \in]0, 1[\mid \|F(x') - F(x'')\| \leq L \|x' - x''\|, \forall x', x'' \in \mathbb{R}^n \quad (1)$$

podemos asegurar que la sucesión $\{x^m\}_m$ definida por $x^m = F(x^{m-1})$ converge a un punto fijo $x = F(x)$ cualquiera que sea x^0 .

Si tomamos $F(x) = Bx + C$, la condición (1), suficiente para que $\{x^m\}_m$ sea convergente a un punto fijo de F cualquiera que sea x^0 , queda así

$$\|Bx' - Bx''\| \leq L \|x' - x''\|, \forall x', x'' \in \mathbb{R}^n$$

$$\text{o bien } \|B(x' - x'')\| \leq L \|x' - x''\|, \forall x', x'' \in \mathbb{R}^n$$

$$\text{o lo que es equivalente } \|Bv\| \leq L \|v\|, \forall v \in \mathbb{R}^n.$$

Veamos entonces el siguiente

1.1. TEOREMA: El método definido anteriormente es convergente si, y solo si, se verifican las dos condiciones siguientes:

i) $C = (I - B)A^{-1}b$, donde I es la matriz unidad en \mathbb{A}^n .

ii) $\rho(B) < 1$

Demostr. \Rightarrow Si el método $x^m = Bx^{m-1} + C$ converge, se tiene que tomando x^0 arbitrario, la sucesión $\{x^m\}_m$ converge a $x = A^{-1}b$.

$$\text{Además } x^m = Bx^{m-1} + C, \forall m \in \mathbb{N} \Rightarrow x = Bx + C.$$

$$\text{Pero } x^m - x = B(x^{m-1} - x) = B \cdot B(x^{m-2} - x) = \dots = B^m(x^0 - x)$$

$$\text{Como } \{x^m - x\}_m \rightarrow 0, \text{ se tiene que } \lim_{m \rightarrow \infty} B^m(x^0 - x) = 0$$

Pero esto es cierto para cualquier $x^0 \in \mathbb{R}^n$. Luego se tiene que

$$\lim_{u \rightarrow \infty} B^u v = 0, \forall v \in \mathbb{R}^n.$$

Si tomamos v igual a $(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)$, sucesivamente se prueba que $\lim_{u \rightarrow \infty} b_{ij}^u = 0, i, j = 1, \dots, n$. Luego $\lim_{u \rightarrow \infty} B^u = 0$. Por tanto, según Teema 4, Teorema 3.1, se tiene que $\rho(B) < 1$.

$$\text{Además, } x = Bx + C \Rightarrow (I - B)x = C \Rightarrow (I - B)A^{-1}b = C$$

Luego se verifican i) e ii).

\Leftarrow Siendo $\rho(B) < 1$, existe una norma matricial $\|\cdot\|_M$ sobre A_n tal que $\|B\|_M < 1$. Sea $\|\cdot\|_V$ una norma vectorial compatible con $\|\cdot\|_M$. Entonces, $\forall v \in \mathbb{R}^n, \|Bv\|_V \leq \|B\|_M \|v\|_V$.

Luego tomando $L = \|B\|_M$ se verifica que $\|Bv\|_V \leq L \|v\|_V, \forall v \in \mathbb{R}^n$ y siendo $L < 1$, la sucesión $\{x^u\}_u$ definida por $x^u = Bx^{u-1} + C$, siendo x^0 arbitrario, converge a un punto x tal que $x = Bx + C$ y por tanto $(I - B)x = C$.

Pero por i), $(I - B)A^{-1}b = C$.

$$\text{Luego } (I - B)x = (I - B)A^{-1}b.$$

Como $|I - B| \neq 0$, pues siendo $\rho(B) < 1, 1$ no puede ser autovalor de B , se tiene que

$$x = A^{-1}b$$

que prueba que $\{x^u\}_u$ converge a la solución del sistema $Ax = b$. c.q.d.

OBSERVACION: Una condición equivalente a ii) es que exista una norma matricial $\|\cdot\|_M$ tal que $\|B\|_M < 1$.

2. CONSTRUCCION DE LA MATRIZ B Y EL VECTOR C.

A) Si encontramos dos matrices M y N tales que M es invertible y $A = M - N$, haciendo $B = M^{-1}N$ y $C = M^{-1}b$ se verifica la condición i) del apartado anterior, pues

$$I - B = I - M^{-1}N = M^{-1}M - M^{-1}N = M^{-1}(M - N) = M^{-1}A \quad \text{y por tanto}$$

$$(I - B)A^{-1}b = M^{-1}AA^{-1}b = M^{-1}b = C.$$

Si conseguimos además que $\rho(M^{-1}N) < 1$, tendríamos que el método definido anteriormente es convergente.

Una condición necesaria para la convergencia de este método la da el siguiente

2.1. TEOREMA: Si el método es convergente, existen matrices M y N tales que M es inversible, $A = M - N$, $B = M^{-1}N$ y $C = M^{-1}b$.

Demostr.: Siendo el método convergente, $\rho(B) < 1$ y por tanto 1 no es autovvalor de B. Luego $I-B$ es inversible. Entonces

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

Haciendo $M = A(I-B)^{-1}$ y $N = A(I-B)^{-1}B$ tenemos lo siguiente:

- M es inversible, pues A e $(I-B)^{-1}$ lo son.
- $A = M - N$ por construcción.
- $M^{-1}N = (I-B)A^{-1}A(I-B)^{-1}B = B$
- $M^{-1}b = (I-B)A^{-1}b = C$ por lo dicho anteriormente. c.s.q.d.

Entonces la sucesión $\{x^m\}$ la podríamos construir, partiendo de x^0 arbitrario, y tomando

$$x^m = M^{-1}N x^{m-1} + M^{-1}b$$

o bien $M x^m = N x^{m-1} + b$

Siendo M inversible podemos calcular x^m conocido x^{m-1} .

Vamos a estudiar tres métodos para tomar las matrices M y N.

B) METODO DE JACOBI: Partimos de una matriz $A = (a_{ij}) \in \mathbb{R}^n$, con $|A| \neq 0$ y tal que los elementos de la diagonal son distintos de cero.

Consideremos las matrices

$$D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ -a_{21} & 0 & \dots & 0 & 0 \\ -a_{31} & -a_{32} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & -a_{nn-1} & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & -a_{12} & \dots & -a_{1n} \\ 0 & 0 & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & -a_{n-1n} \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

Tenemos entonces $A = D - L - U$.

Sea $M = D$ y $N = L + U$. Entonces $A = M - N$ y M es inversible.

Haciendo $B = M^{-1}N = D^{-1}(L+U)$ tenemos

$$B = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \dots & -\frac{a_{2n}}{a_{22}} \\ -\frac{a_{31}}{a_{33}} & -\frac{a_{32}}{a_{33}} & 0 & \dots & -\frac{a_{3n}}{a_{33}} \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & -\frac{a_{n3}}{a_{nn}} & \dots & 0 \end{pmatrix}$$

Partiendo de x^0 arbitrario y tomando $M x^{m+1} = N x^m + b$ tenemos

$$\begin{pmatrix} a_{11} & & & & \\ & 0 & & & \\ & & \dots & & \\ & & & \dots & \\ 0 & & & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1^{m+1} \\ \vdots \\ x_n^{m+1} \end{pmatrix} = \begin{pmatrix} 0 & -a_{12} & \dots & -a_{1n} \\ -a_{21} & 0 & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & 0 \end{pmatrix} \begin{pmatrix} x_1^m \\ x_2^m \\ \vdots \\ x_n^m \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Por tanto

$$\begin{cases} a_{11}x_1^{u+1} = -a_{12}x_2^u - a_{13}x_3^u - \dots - a_{1n}x_n^u + b_1 \\ a_{21}x_1^{u+1} + a_{22}x_2^{u+1} = -a_{23}x_3^u - \dots - a_{2n}x_n^u + b_2 \\ a_{31}x_1^{u+1} + a_{32}x_2^{u+1} = -a_{33}x_3^u - \dots - a_{3n}x_n^u + b_3 \\ \dots \\ a_{nn}x_n^{u+1} = -a_{n1}x_1^u - a_{n2}x_2^u - \dots - a_{n,n-1}x_{n-1}^u + b_n \end{cases}$$

o bien

$$a_{ii}x_i^{u+1} = -\left(\sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j^u\right) + b_i, \quad i=1, \dots, n$$

C) METODO DE GAUSS-SEIDEL.

Si tomamos $M=D-L$ y $N=U$ se tiene que M es fácilmente invertible y $A=M-N$.

Haciendo $Mx^{u+1} = Nx^u + b$ tenemos

$$\begin{pmatrix} a_{11} & & & & \\ a_{21} & a_{22} & & & \\ \dots & \dots & \dots & \dots & \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1^{u+1} \\ x_2^{u+1} \\ \vdots \\ x_n^{u+1} \end{pmatrix} = \begin{pmatrix} 0 & -a_{12} & -a_{13} & \dots & -a_{1n} \\ & 0 & -a_{23} & \dots & -a_{2n} \\ & & & \dots & \\ 0 & & & & 0 \end{pmatrix} \begin{pmatrix} x_1^u \\ x_2^u \\ \vdots \\ x_n^u \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Por tanto:

$$\begin{cases} a_{11}x_1^{u+1} = -a_{12}x_2^u - a_{13}x_3^u - \dots - a_{1n}x_n^u + b_1 \\ a_{21}x_1^{u+1} + a_{22}x_2^{u+1} = -a_{23}x_3^u - \dots - a_{2n}x_n^u + b_2 \\ \dots \\ a_{n1}x_1^{u+1} + \dots + a_{nn}x_n^{u+1} = b_n \end{cases}$$

o bien

$$\sum_{j=1}^i a_{ij}x_j^{u+1} = -\left(\sum_{j=i+1}^n a_{ij}x_j^u\right) + b_i, \quad i=1, \dots, n$$

tenemos entonces

$$a_{ii}x_i^{u+1} = -\sum_{j=1}^{i-1} a_{ij}x_j^{u+1} - \sum_{j=i+1}^n a_{ij}x_j^u + b_i, \quad i=1, \dots, n$$

Como vemos, para calcular x_p^{u+1} solo utilizamos las $n-p$ últimas componentes de x^u y las p primeras las vamos calculando. El proceso que se sigue es el siguiente:

Sabemos que descomponiendo A como diferencia de una matriz invertible M y otra matriz N se verifica i). Falta ver que se cumple ii) para que el método sea convergente.

En el apartado anterior hemos puesto A en la forma $A=M-N$ en los métodos B), C) y D). Falta estudiar cuando es $\rho(B) < 1$ para saber cuando son convergentes dichos métodos.

B) Convergencia del método de Jacobi: Tenemos que

$$B = M^{-1}N = D^{-1}(L+U) = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \dots & -\frac{a_{2n}}{a_{22}} \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & -\frac{a_{n3}}{a_{nn}} & \dots & 0 \end{pmatrix}$$

Para ver que $\rho(B) < 1$, hemos de encontrar una norma matricial $\|\cdot\|_M$ tal que $\|B\|_M < 1$. Las normas más manejables son las normas $\|\cdot\|_\infty$ y $\|\cdot\|_1$ que están definidas como si fue

$$\|(a_{ij})\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|, \quad \|(a_{ij})\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

Por tanto

$$\|B\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \quad \text{y} \quad \|B\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n \left| \frac{a_{ij}}{a_{ii}} \right|$$

Si se verifica que $\|B\|_\infty < 1$ o $\|B\|_1 < 1$, el método será convergente. Si no se verifica esto hay que utilizar otro procedimiento para ver si $\rho(B) < 1$. Calculamos entonces los autovalores de la matriz B . Son trivialmente ciertas las equivalencias siguientes:
 $(\lambda \text{ es autovalor de } B) \Leftrightarrow |D^{-1}(L+U) - \lambda I| = 0 \Leftrightarrow |D| |D^{-1}(L+U) - \lambda I| = 0 \Leftrightarrow$
 $\Leftrightarrow |L+U - \lambda D| = 0 \Leftrightarrow |\lambda D - L - U| = 0 \Leftrightarrow$

$$\Leftrightarrow \begin{vmatrix} \lambda a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \lambda a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \lambda a_{nn} \end{vmatrix} = 0$$

Si $\lambda_1, \dots, \lambda_n$ son los autovalores de B , para que el método sea convergente debe ser $\rho(B) = \max_{1 \leq i \leq n} |\lambda_i| < 1$.

C) Convergencia del método de Gauss-Seidel: Calculamos los autovalores de la matriz B correspondiente a este método.

$$|B - \lambda I| = 0 \Leftrightarrow |(D-L)^{-1}U - \lambda I| = 0 \Leftrightarrow |D-L| |(D-L)^{-1}U - \lambda I| = 0 \Leftrightarrow$$

$$\Leftrightarrow |U - \lambda(D-L)| = 0 \Leftrightarrow |\lambda D - \lambda L - U| = 0 \Leftrightarrow$$

$$\Leftrightarrow \begin{vmatrix} \lambda a_{11} & a_{12} & \dots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \end{vmatrix} = 0$$

Habría que ver ahora si $\max_{1 \leq i \leq n} |\lambda_i| < 1$, donde $\lambda_i, i=1, \dots, n$, son las raíces de la ecuación anterior, es decir, los autovalores de B.

- Se prueban los siguientes resultados:
 - Si A es simétrica y definida positiva, entonces el método de Gauss-Seidel es convergente.
 - Si A es simétrica y la matriz

$$D+L+U = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$
 es definida positiva, entonces el método de Jacobi es convergente.
 - Si A es simétrica y definida positiva, el método de relajación es convergente si, y solo si, $0 < \omega < 2$.
 - Si A es simétrica, definida positiva y tridiagonal el ω para el que el método de relajación converge más rápidamente es $\omega = \frac{2}{1 + \sqrt{1 - \rho_J^2}}$ donde ρ_J es el radio espectral de la matriz B correspondiente al método de Jacobi.

• Supuesto que el método es convergente nos interesa saber cuando $x^{(m)}$ está suficientemente cerca de la solución x.

Como $x^m = Bx^{m-1} + C$, $\forall m$, tomando límites tenemos que $x = Bx + C$.

Luego $x^m - x = B(x^{m-1} - x) = \dots = B^m(x^0 - x)$.

Tomando una norma matricial y una norma vectorial compatible con ella tenemos que

$$\|x^m - x\| \leq \|B\| \cdot \|x^{m-1} - x\| \leq \dots \leq \|B\|^m \|x^0 - x\|$$

Por otra parte $\|x^{m-1} - x\| \leq \|x^{m-1} - x^m\| + \|x^m - x\| \leq \|x^{m-1} - x^m\| + \|B\| \cdot \|x^{m-1} - x\|$ de donde

$$\|x^{m-1} - x\| (1 - \|B\|) \leq \|x^{m-1} - x^m\| \text{ y por tanto } \|x^{m-1} - x\| \leq \frac{1}{1 - \|B\|} \|x^{m-1} - x^m\|$$

$$\text{Luego } \|x^m - x\| \leq \frac{\|B\|}{1 - \|B\|} \|x^{m-1} - x^m\|$$

Si conseguimos que el segundo término de la desigualdad sea menor que un cierto ϵ , sabemos que el error que estamos cometiendo es menor que ϵ , y si ϵ es suficientemente pequeño tomar x^m como solución aproximada.

TEMA 6°: CÁLCULO DE VALORES PROPIOS E INVERSIÓN DE MATRICES.

1. INTRODUCCIÓN

El conocimiento de los valores propios de una matriz A es uno de los puntos claves en muchos problemas matemáticos: acabamos de ver su relación con la convergencia de los métodos iterativos de resolución de sistemas lineales.

Sin embargo no es un problema sencillo, pues no es sencilla en general la determinación del polinomio característico ($\det(A - \lambda I)$) ni el cálculo de sus raíces.

No existen métodos directos para el cálculo de autavalores. Sin embargo, algunos autores llaman métodos directos a los que permiten calcular los coeficientes exactos del polinomio característico con un número finito de operaciones, y métodos iterativos al resto, a pesar de que en los métodos directos, una vez calculado el polinomio característico, habrá que acudir a algún método iterativo para calcular sus raíces.

Hay varios métodos para resolver este problema y la elección de uno u otro depende de lo que se intente calcular:

- se desean calcular todos los valores propios.
- se desean también los vectores propios.
- se desea solamente el valor propio dominante, es decir, el de mayor módulo.
- ¿tiene la matriz alguna forma especial que la haga más sencilla para el cálculo?

El polinomio característico $|A - \lambda I|$ puede ser escrito en la forma

$$P(\lambda) = (-1)^n \lambda^n + a_1 \lambda^{n-1} + \dots + a_n$$

Desarrollando se ve inmediatamente que

$$a_1 = (-1)^{n-1} \sum_{i=1}^n a_{ii}$$

y siendo $\sum_{i=1}^n \lambda_i = \frac{-a_1}{(-1)^{n-1}}$ se deduce que $\sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii} = \text{Tr}(A)$.

donde $\text{Tr}(A)$ es la traza de la matriz A y se define como $\sum_{i=1}^n a_{ii}$.

Se deduce también que

$$\frac{(-1)^n a_n}{(-1)^n} = \prod_{i=1}^n \lambda_i = P(0)$$

por lo que, por ser $P(0) = |A|$ resulta que

$$\prod_{i=1}^n \lambda_i = |A|.$$

Recordemos también que las matrices semejantes A y $B = C^{-1}AC$ tienen los mismos valores propios, y que si λ es autovector de A entonces $C\lambda$ es autovector de B .

autovector de A^n ya que si $Au = \lambda u$ entonces
 $A^n u = A^{(n-1)}(\lambda u) = \dots = \lambda^n u$.

2. METODOS DE DETERMINACION DEL POLINOMIO CARACTERISTICO

Vamos a recordar una serie de resultados que utilizaremos en lo siguiente:

2.1. TEOREMA: Se verifica en \mathbb{C} que el espectro (conjunto de autovalores) de la matriz A^n es $\sigma(A^n) = \{\lambda_1^n, \dots, \lambda_n^n\}$ siendo $\lambda_i, i=1, \dots, n$, los autovalores de A .

Una generalización de este teorema es el siguiente:

2.2. TEOREMA: Si $P(x)$ es un polinomio y $\lambda_1, \dots, \lambda_n$ son los autovalores de A , entonces los autovalores de $P(A)$ son $P(\lambda_1), \dots, P(\lambda_n)$.

2.3. TEOREMA: Si A es real y simétrica, todos los autovalores son reales. Además los autovectores correspondientes a autovalores distintos son ortogonales.

2.4. TEOREMA: Cualquier transformación del tipo $P^{-1}AP$ (semejanza) deja invariante los autovalores de A .

Probarémos el siguiente

2.5. TEOREMA: (de Cayley-Hamilton).

Toda matriz es raíz de su polinomio característico, es decir

$$P(A) = (-1)^n A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n I = 0.$$

Demostr.: Dada la matriz $A = (a_{ij})$ se define el adjunto A_{ik} del elemento a_{ik} como $A_{ik} = (-1)^{i+k} \det \alpha_{ik}$, donde α_{ik} es la matriz que se obtiene suprimiendo en la matriz A la fila i -ésima y la columna k -ésima. Denotamos por \tilde{A} la matriz adjunta (formada por los adjuntos) de A , es decir, $\tilde{A} = (A_{ik})$. Se verifica que $A\tilde{A} = \tilde{A}A = (\det A) \cdot I$.

Sea C la matriz adjunta de la matriz $A - \lambda I$. Los elementos de C son polinomios en λ de grado no mayor que $n-1$. Luego C puede expresarse como

$$C = C_1 \lambda^{n-1} + C_2 \lambda^{n-2} + \dots + C_{n-1} \lambda + C_n$$

donde C_1, \dots, C_n son matrices numéricas. Así

$$(A - \lambda I) \cdot (C_1 \lambda^{n-1} + C_2 \lambda^{n-2} + \dots + C_n) = |A - \lambda I| \cdot I = [(-1)^n \lambda^n + a_1 \lambda^{n-1} + \dots + a_n] \cdot I.$$

Identificando los coeficientes resulta

$$-C_1 = (-1)^n I$$

$$AC_1 - C_2 = a_1 I$$

$$AC_2 - C_3 = a_2 I$$

$$AC_{n-1} - C_n = a_{n-1} I$$

Multiplicando la primera ecuación por A^n , la segunda por A^{n-1} , y así sucesivamente obtenemos, sumando a continuación fue

$$(-1)^n A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n I = 0. \text{ c.s.q.d.}$$

Vamos a calcular los coeficientes del polinomio característico. El cálculo directo a partir de la definición del determinante solo es recomendado para matrices de orden muy bajo. Vamos a estudiar varios métodos para la determinación del polinomio característico.

A) METODO DE KRYLOV: Por el teorema de Cayley-Hamilton sabemos

$$\text{fue } (-1)^n A^n + \sum_{i=1}^n a_i A^{n-i} = 0$$

Entonces, para cualquier vector y se tiene fue

$$(-1)^n A^n y + \sum_{i=1}^n a_i A^{n-i} y = 0$$

que escrito en la forma

$$\sum_{i=1}^n a_i A^{n-i} y = -(-1)^n A^n y$$

representa un sistema de n ecuaciones lineales en las n incógnitas a_1, \dots, a_n , que puede ser resuelto por cualquiera de los métodos estudiados. Vamos a estudiar este sistema.

Sea $V_i = A^{n-i} y$. Entonces $\sum_{i=1}^n V_i a_i = -(-1)^n A^n y$

$$\text{Sea } V = (V_1, \dots, V_n) = \begin{pmatrix} v_{11} & \dots & v_{1n} \\ \dots & \dots & \dots \\ v_{n1} & & v_{nn} \end{pmatrix} \text{ y } a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

$$\text{Entonces } V \cdot a = -(-1)^n A^n y$$

Luego este método tiene la limitación de que ha de ser $\det V \neq 0$, para que tenga solución el anterior sistema, es decir, han de ser los V_i linealmente independientes.

B) METODO DE LEVERRIER: Veamos las relaciones de Newton entre las raíces de un polinomio $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ y los coeficientes. Sean x_1, \dots, x_n dichas raíces y sea $s_j = \sum_{i=1}^n x_i^j$, $j=1, \dots, n$.

Entonces se cumple que

$$a_n s_n + a_{n-1} s_{n-1} + \dots + a_1 s_1 + n a_0 = 0$$

$$a_n s_{n-1} + a_{n-1} s_{n-2} + \dots + a_2 s_1 + (n-1) a_1 = 0$$

$$a_n s_2 + a_{n-1} s_1 + 2 a_{n-2} = 0$$

$$a_n s_1 + a_{n-1} = 0$$

$$\text{Además } \sum_{i < j} x_i x_j = \frac{a_{n-2}}{a_n} \text{ y } \prod_{i=1}^n x_i = (-1)^n \frac{a_0}{a_n}.$$

Aplicando esto al polinomio característico $(-1)^n \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n$ se tiene que

$$(-1)^n s_n + a_1 s_{n-1} + \dots + a_{n-1} s_1 + n a_n = 0$$

$$(-1)^n s_2 + a_1 s_1 + 2 a_2 = 0$$

$$(-1)^n s_1 + a_1 = 0$$

de donde $a_1 = -(-1)^n s_1 = -(-1)^n \text{Tr}(A)$, pues $s_1 = \sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii} = \text{Tr}(A)$

En general, siendo

$$(-1)^n s_p + a_1 s_{p-1} + a_2 s_{p-2} + \dots + a_{p-1} s_1 + p a_p = 0$$

se deduce que

$$a_p = -\frac{1}{p} \left((-1)^n s_p + a_1 s_{p-1} + a_2 s_{p-2} + \dots + a_{p-1} s_1 \right)$$

y por tanto $a_p = -\frac{1}{p} \left((-1)^n \text{Tr}(A^p) + a_1 \text{Tr}(A^{p-1}) + \dots + a_{p-1} \text{Tr}(A) \right)$

pues $s_k = \sum_{i=1}^n \lambda_i^k = \text{Tr}(A^k)$, $k=1, \dots, n$.

C) METODO DE SOURIAU: Es una elegante modificación del método anterior. Se toma

$$A_1 = A$$

$$P_1 = \text{Tr}(A_1)$$

$$B_1 = A_1 - P_1 I$$

$$A_2 = B_1 A$$

$$P_2 = \frac{1}{2} \text{Tr}(A_2)$$

$$B_2 = A_2 - P_2 I$$

$$A_3 = B_2 A$$

$$P_3 = \frac{1}{3} \text{Tr}(A_3)$$

$$B_3 = A_3 - P_3 I$$

$$A_n = B_{n-1} A$$

$$P_n = \frac{1}{n} \text{Tr}(A_n)$$

$$B_n = A_n - P_n I$$

2.6. TEOREMA: Se verifica que $B_n = O$, (y por tanto $A_n = P_n I$ y $B_{n-1} A = P_n I$)

y que el polinomio característico de A es

$$(-1)^n \left[\lambda^n - P_1 \lambda^{n-1} - P_2 \lambda^{n-2} - \dots - P_n \right].$$

Demostr.:

$P(\lambda) = \det(A - \lambda I) = (-1)^n \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n = (-1)^n \left[\lambda^n + (-1)^n a_1 \lambda^{n-1} + \dots + (-1)^n a_{n-1} \lambda + (-1)^n a_n \right]$ es el polinomio característico de A .

Vamos a probar por inducción que $a_k = -(-1)^n P_k$.

Evidentemente $a_1 = -(-1)^n \text{Tr}(A) = -(-1)^n \text{Tr}(A_1) = -(-1)^n P_1$.

Supongamos que es cierto para cualquier $i \leq k-1$ y probémoslo para k .

$$A_k = B_{k-1} A = (A_{k-1} - P_{k-1} I) A = A_{k-1} A - P_{k-1} A.$$

Sustituyendo A_{k-1} por $A_{k-2} A - P_{k-2} A, \dots$ se obtiene

$$A_k = A^k - P_1 A^{k-1} - P_2 A^{k-2} - \dots - P_{k-1} A = A^k + (-1)^n a_1 A^{k-1} + (-1)^n a_2 A^{k-2} + \dots + (-1)^n a_{k-1} A.$$

Luego

$$\text{Tr}(A^k) = \text{Tr}(A^k) + (-1)^n a_1 \text{Tr}(A^{k-1}) + (-1)^n a_2 \text{Tr}(A^{k-2}) + \dots + (-1)^n a_{k-1} \text{Tr}(A) =$$

$$= s_k + (-1)^n a_1 s_{k-1} + (-1)^n a_2 s_{k-2} + \dots + (-1)^n a_{k-1} s_1 = -(-1)^n k a_k$$

y esta última igualdad es por las relaciones de Newton.

Luego

$$a_k = -(-1)^n \frac{\text{Tr}(A^k)}{k} = -(-1)^n p_k.$$

Se deduce entonces que

$$(-1)^n [\lambda^n - p_1 \lambda^{n-1} - \dots - p_n] = (-1)^n [\lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n], \text{ como fuéramos ver.}$$

Por otra parte

$$B_n = A^n - p_n I = B_{n-1} A - p_n I = (A_{n-1} - p_{n-1} I) A - p_n I =$$

$$= A^n - p_1 A^{n-1} - p_2 A^{n-2} - \dots - p_n I =$$

$$= A^n + (-1)^n a_1 A^{n-1} + (-1)^n a_2 A^{n-2} + \dots + (-1)^n a_n I.$$

Por el teorema de Cayley-Hamilton sabemos que

$$(-1)^n A^n + a_1 A^{n-1} + a_2 A^{n-2} + \dots + a_n I = 0$$

y por tanto

$$A^n + (-1)^n a_1 A^{n-1} + (-1)^n a_2 A^{n-2} + \dots + (-1)^n a_n I = 0$$

es decir $B_n = 0$.

Siendo $B_n = A^n - p_n I$ y $A^n = B_{n-1} A$ se deduce que $B_{n-1} A = p_n I$, esq.d.

OBSERVACION: Si A es no singular $a_n = \det A = -(-1)^n p_n \neq 0$. Luego $p_n \neq 0$.

$$\text{Entonces } A^{-1} = \frac{1}{p_n} B_{n-1}.$$

Luego el método de Sourian sirve para calcular la matriz inversa.

3. Cálculo del mayor autovalor en módulo: Método de la potencia iterada.

A) Vamos a recordar en primer lugar algunos resultados acerca de la forma canónica de Jordan. Sabemos que toda matriz A de orden n es semejante a una matriz diagonal por bloques de Jordan, es decir, existe una matriz no singular P , cuyos elementos pueden ser complejos, de forma que

$$PAP^{-1} = \begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ 0 & & J_s \end{pmatrix}$$

siendo cada J_k una matriz de la forma

$$J_k = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ & \lambda_i & 1 & \dots & 0 \\ & & & \ddots & 1 \\ 0 & & & & \lambda_i \end{pmatrix}$$

donde λ_i es un autovalor de A .

Un autovalor dado puede aparecer como el elemento diagonal de más de un J_k .

Si v_k es el orden de J_k se verifica que

$$\det(J_k - \lambda I) = (\lambda_i - \lambda)^{v_k}, \quad k=1, \dots, s.$$

A estos determinantes se les llama divisores elementales de A . Si $v_k=1$ decimos que el divisor elemental correspondiente es lineal.

Se deduce entonces que si todos los autovalores son distintos, todos los divisores elementales son lineales. Se puede demostrar que en una matriz simétrica los divisores elementales son lineales, sean o no distintos los autovalores.

Se prueba también que si los divisores elementales son lineales, entonces para cada autovalor de multiplicidad k existen k autovectores linealmente independientes.

3.1. TEOREMA: El conjunto de autovectores asociados a un autovalor tiene estructura de espacio vectorial de dimensión mayor o igual que 1 y menor o igual que la multiplicidad del autovalor.

Demostr.: El conjunto de autovectores asociados al autovalor λ es el núcleo del homomorfismo

$$x \in \mathbb{C}^n \mapsto (A - \lambda I)x \in \mathbb{C}^n.$$

cuya dimensión es igual a $n - \text{rang}(A - \lambda I)$.

Dado que toda matriz cuadrada es semejante a su forma canónica de Jordan J , cuyos elementos de la diagonal principal son los autovalores, repetidos tantas veces como indica su orden de multiplicidad, y que $\text{rang}(A - \lambda I) = \text{rang}(J - \lambda I)$ se sigue inmediatamente que la dimensión del núcleo está acotada en la forma indicada.

3.2. TEOREMA: Si λ_1 y λ_2 son autovalores distintos de A , cualesquiera que sean los autovectores x_1 y x_2 asociados respectivamente a λ_1 y λ_2 son linealmente independientes.

3.3. COROLARIO: Si $\lambda_1, \dots, \lambda_n$ son autovalores distintos de A , los autovectores asociados a ellos son linealmente independientes.

3.4. COROLARIO: Si todos los autovalores de una matriz son distintos el conjunto $\{x_1, \dots, x_n\}$ es una base del espacio vectorial \mathbb{C}^n sobre \mathbb{C} , donde cada x_i es autovector asociado a λ_i .

B) Vamos a estudiar ahora como calcular el mayor autovector en módulo de una matriz dada.

Supondremos que todos los divisores elementales de A son lineales. Sin embargo, el método que obtendremos será también aplicable aun cuando A tenga divisores elementales no lineales.

En estas hipótesis existe una base de \mathbb{C}^n formada por autovectores de A . Luego dado $v_0 \in \mathbb{C}^n$ existen $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ tales que

$$v_0 = \sum_{i=1}^n \alpha_i x_i$$

donde $\{x_i\}_{i=1}^n$ es una base de \mathbb{C}^n de autovectores de A .

Si λ_i es el autovector asociado a x_i , entonces

$$A v_0 = A \sum_{i=1}^n \alpha_i x_i = \sum_{i=1}^n \alpha_i \lambda_i x_i$$

Entonces si se considera el proceso iterativo $v_m = A v_{m-1}$, $m \in \mathbb{N}$, se tendrá $v_m = A v_{m-1} = \dots = A^m v_0 = \sum_{i=1}^n \alpha_i \lambda_i^m x_i$.

Ordeneemos los autovalores de forma que

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|.$$

Si λ_1 es real y dominante ($|\lambda_1| > |\lambda_2|$) tenemos el siguiente teorema que nos da el llamado método de la potencia iterada:

3.5. TEOREMA: (de Von Mises)

Si la matriz A tiene n autovectores linealmente independientes y si el autovector de mayor módulo λ_1 es dominante y v_0 tiene una componente en x_1 entonces

$$\lim_{m \rightarrow \infty} \frac{1}{\lambda_1^m} A^m v_0 = \alpha_1 x_1.$$

Demostr.: Si $v_0 = \sum_{i=1}^n \alpha_i x_i$, como $\alpha_1 \neq 0$, y dado que

$$A^m v_0 = \sum_{i=1}^n \alpha_i \lambda_i^m x_i$$

se tiene que $\frac{1}{\lambda_1^m} A^m v_0 = \sum_{i=1}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^m x_i$

que tiende a $\alpha_1 x_1$ cuando $m \rightarrow \infty$, por ser λ_1 dominante. esqcl.

Como consecuencia de este teorema observamos que si y es cualquier vector no ortogonal a x_1 entonces

$$\lim_{m \rightarrow \infty} \frac{y^t v_{m+1}}{y^t v_m} = \lambda_1,$$

pues $y^t v_{m+1} = y^t \sum_{i=1}^n \alpha_i \lambda_i^{m+1} x_i$ y $y^t v_m = y^t \sum_{i=1}^n \alpha_i \lambda_i^m x_i$ y por tanto

$$\lim_{m \rightarrow \infty} \frac{y^t v_{m+1}}{y^t v_m} = \lim_{m \rightarrow \infty} \frac{\lambda_1 \sum_{i=1}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{m+1} y^t x_i}{\sum_{i=1}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^m y^t x_i} = \lambda_1 \frac{y^t x_1}{y^t x_1} = \lambda_1$$

4. DEFLACIONES

Frecuentemente deseamos calcular, no justamente el autovvalor dominante, sino alguno, o quizas, todos los otros autovvalores. Veremos algunas técnicas para calcular uno o más de los autovvalores subdominantes después de que se ha usado el método de la potencia para calcular el autovvalor dominante.

El método de la potencia puede modificarse fácilmente para que puedan calcularse otros autovvalores. Si, por ejemplo, A tiene un valor propio λ , entonces $A - \lambda I$ tiene un valor propio 0 . De esta forma, si λ_1 es el autovvalor de mayor módulo de A , entonces

$$0, \lambda_2 - \lambda_1, \dots, \lambda_n - \lambda_1$$

son los autovvalores de $A - \lambda_1 I$ y mediante la potencia iterada podemos calcular el autovvalor $\lambda_i - \lambda_1$ de mayor módulo, y por tanto, el λ_i que más dista de λ_1 . De esta forma calcularemos los autovvalores extremos, lo cual ya nos da una cota para los demás autovvalores.

Más general resulta el método dado por el siguiente

4.1. TEOREMA: Sean λ_1 y x_1 un autovvalor de A y un autovector asociados. Sea v un vector tal que $v^t x_1 = 1$. Entonces la matriz

$$W_1 = A - \lambda_1 x_1 v^t$$

tiene los mismos autovvalores de A excepto λ_1 , que es sustituido por cero.

La demostración puede verse en Ralston (pg. 524)

A W_1 se le llame matriz defletada

La forma en que usaremos este teorema está clara; como el autovvalor dominante de W_1 es λ_2 , podemos aplicar el método de la potencia a W_1 para determinar λ_2 . Construyamos análogamente $W_2 = W_1 - \lambda_2 w_2 v_1^t$ donde $v_1^t w_2 = 1$. Este proceso de deflación puede repetirse para calcular λ_3 , y así sucesivamente.

Si λ_1 es un autovvalor de multiplicidad K , las primeras K aplicaciones de este proceso de deflación darán como resultado λ_1 .

Otro método para calcular los autovvalores subdominantes nos lo da el siguiente

4.2. TEOREMA: Si A tiene n autovvalores $\lambda_1, \dots, \lambda_n$ reales y distintos y con autovectores asociados u_1, \dots, u_n respectivamente, entonces la matriz

$$A_1 = A - \frac{\lambda_1}{v_1^t u_1} u_1 v_1^t$$

$0, \lambda_2, \dots, \lambda_n$ y los autovectores correspondientes son u_1, u_2, \dots, u_n .

Demostr.: Siendo v_1 autovector correspondiente a A^t y λ_1 , es ortogonal a cualquier u_i que sea autovector de A y λ_i ($\lambda_i \neq \lambda_1$)

Por consiguiente, como $\lambda_i \neq \lambda_j$, si $i \neq j$, se tendrá que $v_1 \perp u_i$, $i=2, \dots, n$, y por tanto v_1 no puede ser ortogonal a u_1 , es decir, $v_1^t u_1 \neq 0$.

Por ser $A_1 = A - \lambda_1 \frac{u_1 v_1^t}{v_1^t u_1}$ se tiene que

$$A_1 u_1 = A u_1 - \lambda_1 \frac{(u_1 v_1^t)}{v_1^t u_1} u_1 = A u_1 - \lambda_1 u_1 \frac{v_1^t u_1}{v_1^t u_1} = A u_1 - \lambda_1 u_1 = 0 \cdot u_1$$

Es decir, 0 es autovalor de A_1 y u_1 es autovector asociado a cero.

Por otra parte

$$A_1 u_i = A u_i - \lambda_1 \frac{u_1 v_1^t}{v_1^t u_1} u_i = A u_i - \lambda_1 u_1 \frac{v_1^t u_i}{v_1^t u_1}, \quad i=2, \dots, n.$$

Pero $v_1^t A = \lambda_1 v_1^t$, pues $A^t v_1 = \lambda_1 v_1$, por definición de v_1 .

Luego $v_1^t A u_i = \lambda_1 v_1^t u_i$, y de $A u_i = \lambda_i u_i$ se tiene que $v_1^t A u_i = \lambda_i v_1^t u_i$.

Luego $\lambda_i v_1^t u_i = \lambda_1 v_1^t u_i$, o lo que es lo mismo

$$(\lambda_i - \lambda_1) v_1^t u_i = 0$$

Como $\lambda_i \neq \lambda_1$, $i=2, \dots, n$, se deduce que $v_1^t u_i = 0$, $i=2, \dots, n$.

Por tanto $A_1 u_i = A u_i = \lambda_i u_i$. Es decir, λ_i es autovalor de A_1 y u_i es el autovector correspondiente, $\forall i \in \{2, \dots, n\}$. c.s.g.d.

Por consiguiente, si $\lambda_1 > \lambda_2 > \dots > \lambda_n > 0$, el método de la potencia iterada aplicado a la matriz A_1 nos da λ_2 , y a partir de él obtenemos v_2 .

Construyendo entonces la matriz

$$A_2 = A_1 - \lambda_2 \frac{u_2 v_2^t}{v_2^t u_2}$$

y repitiendo el proceso obtendremos λ_3 , y así sucesivamente todos los autovalores. Este método es bastante restrictivo y puede presentar problemas de estabilidad.

5. SITUACION Y COTAS DE LOS AUTOVALORES.

En esta sección consideraremos algunos de los teoremas importantes entre los muchos que se relacionan con la situación de los autovalores de una matriz. Estos teoremas pueden usarse, por ejemplo, para estimar la magnitud de los autovalores de mayor y menor magnitud y, de esta forma estimar la norma de la matriz, para saber si los métodos iterativos $x^m = B x^{m-1} + C$ definidos en el tema anterior son convergentes.

Tales estimaciones pueden utilizarse también para generar condiciones iniciales que serán usadas en métodos iterativos para la estimación

nación de autovalores.

5.1. TEOREMA: (de Gerschgorin)

Sea $A = (a_{ij})$ una matriz de orden n y $C_i, i=1, \dots, n$, los círculos de centro a_{ii} y radios

$$r_i = \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|$$

Considerese además que $D = \bigcup_{i=1}^n C_i$.

Entonces todos los autovalores de A están en D .

Demostr.: Es trivial que si x_0 es solución de $Bx=0$, donde $\det B=0$ y $x_0^{(i)}$ es la componente de x_0 de mayor módulo entonces

$$|b_{ii}| \leq \sum_{\substack{k=1 \\ k \neq i}}^n |b_{ik}|$$

Sea entonces λ autovalor de A . Llamemos $B = A - \lambda I$. Entonces

$$|\lambda - a_{ii}| \leq \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|$$

donde i es tal que si $x = (x_1, \dots, x_n)$ es el autovector asociado a λ , x_i es la componente de mayor módulo.

Ahora bien, como i puede variar para los distintos autovalores, se llega rápidamente a la tesis del teorema. csgd.

Como consecuencia del teorema anterior obtenemos los dos resultados siguientes.

5.2. COROLARIO: El radio espectral de A está acotado por

$$\max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}|$$

Este resultado ya lo conocíamos pues $\rho(A) \leq \|A\|_M$ para cualquier norma matricial y $\|A\|_\infty = \max_i \sum_{k=1}^n |a_{ik}|$ es una norma matricial.

5.3. COROLARIO: El radio espectral de A^{-1} es tal que

$$\frac{1}{\rho(A^{-1})} \geq \min_{1 \leq i \leq n} (|a_{ii}| - \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|)$$

Demostr.: $\frac{1}{\rho(A^{-1})} = \frac{1}{\max\{|\frac{1}{\lambda}| \mid \lambda \text{ autovalor de } A\}} = \min\{|\lambda| \mid \lambda \text{ autovalor de } A\} \geq$

$\geq \min_{1 \leq i \leq n} (|a_{ii}| - \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|)$, en virtud del teorema 5.1. csgd.

Estos dos corolarios nos dan respectivamente una cota superior e inferior del módulo de los autovalores.