

# TEMA 4:

## CALCULO NUMERICO

### DE AUTOVALORES

#### 1 INTRODUCCION

La determinación de autovalores y autovectores de una matriz cuadrada  $A$  de orden  $n$  es un problema que se presenta en numerosas ramas de la Matemática: teoría de E.D.P., determinación de ejes de cónicas y cuádricas, diseño de sistemas de información, estudio de las oscilaciones de ciertas estructuras, optimización lineal . . .

Dado que los autovalores son las raíces del polinomio característico

$$P(\lambda) = \det(A - \lambda I) = (-1)^n [\lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n],$$

es claro que los métodos de cálculo de autovalores no pueden ser más que iterativos (ya que según el Teorema de Abel es imposible resolver mediante un número finito de operaciones elementales los polinomios de grado mayor o igual que 5).

Una primera posibilidad para el cálculo de los autovalores es el cálculo (mediante los llamados *métodos clásicos*) de los coeficientes del polinomio característico. Una vez conocidos estos, el problema se reduce a la resolución de una ecuación algebraica por los métodos ya estudiados.

Sin embargo, estos métodos no se usan habitualmente debido a su gran inestabilidad numérica, derivada de la gran sensibilidad de los ceros de un polinomio a los cambios en sus coeficientes.

**Ejemplo 1** .- (*Wilkinson*)

*Sea  $A$  una matriz cuyo polinomio característico es*

$$P(\lambda) = (\lambda - 1)(\lambda - 2) \dots (\lambda - 20).$$

*Si en el cálculo de los coeficientes hubiésemos cometido un error en el coeficiente correspondiente a  $\lambda^{19}$  de orden  $10^{-7}$  de modo que tuviésemos*

$$Q(\lambda) = P(\lambda) - 2^{-23}\lambda^{19}$$

*los autovalores (que son  $1, 2, \dots, 20$ ) pasarían a ser complejos con parte imaginaria grande:  $\pm 2.81 i, \pm 2.51 i \dots$*

Es decir, una pequeña variación en los datos origina una gran variación en los autovalores. Aparece entonces la necesidad de medir el *condicionamiento* del problema de autovalores.

## 2 CONDICIONAMIENTO DEL PROBLEMA DE AUTOVALORES

**Ejemplo 2** .- (*Davis-Moller*)

$$A = \begin{pmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & -9 & -25 \end{pmatrix}$$
$$Sp(A) = \{1, 2, 3\}$$

$$A + \delta A = \begin{pmatrix} -149 & -50 & -154 \\ 537 & 180.01 & 546 \\ -27 & -9 & -25 \end{pmatrix}$$

$$Sp(A + \delta A) = \{0.2073, 2.3008, 3.5019\}$$

*Por tanto, A está mal condicionada.*

Para medir el condicionamiento se tiene el siguiente:

**Teorema 1** .- (*Bauer-Fike*)

*Sea A una matriz diagonalizable con autovalores  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Se considera P inversible tal que  $P^{-1}AP = D = \text{diag}(\lambda_i)$ . Sea  $\|\cdot\|$  una norma matricial tal que para toda matriz diagonal  $\text{diag}(\alpha_i)$  se verifique*

$$\|\text{diag}(\alpha_i)\| = \max_{i=1, \dots, n} |\alpha_i|.$$

*(Las normas usuales  $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$  verifican esta propiedad).*

Entonces, para toda matriz  $\delta A$  se tiene:

$$Sp(A + \delta A) \subset \cup_{i=1}^n D_i$$

donde:

$$D_i = \{z \in C : |z - \lambda_i| \leq cond(P) \cdot \|\delta A\|\}$$

$$con \quad cond(P) = \|P\| \cdot \|P^{-1}\|.$$

Como se puede observar, el condicionamiento del problema de autovalores no depende del número de condición de la matriz  $A$ , sino del de la matriz de paso  $P$  a una matriz diagonal.

**Corolario 1 .-**

$$Sp(A + \delta A) \subset \cup_{i=1}^n \{z \in C : |z - \lambda_i| \leq \Gamma(A) \cdot \|\delta A\|\}$$

donde:

$$\Gamma(A) = \inf\{cond(P) : P^{-1}AP = D\}.$$

$\Gamma(A)$  se denomina *número de condición de  $A$  para el cálculo de autovalores* y verifica que  $\Gamma(A) \geq 1$ .

Cuanto más pequeño sea el número de condición, mejor condicionado estará el problema.

### Observación 1 .-

1. Dado que toda matriz  $A$  normal es diagonalizable a través de una matriz unitaria (Teorema de Schur), si consideramos la norma matricial  $\|\cdot\|_2$  se tiene que  $\Gamma_2(A) = 1$ .

Por tanto, toda matriz normal es bien condicionada.

2. En el ejemplo de Davis-Moller, problema mal condicionado, se tiene  $\Gamma_2(A) = 1289$ .

## 3 METODOS CLASICOS DE CALCULO DEL POLINOMIO CARACTERISTICO

Presentaremos solamente los dos métodos más sencillos:

### 3.1 Método de Leverrier (1840).-

Sea  $A$  una matriz con polinomio característico

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

y con autovalores  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Sean:

$$s_k = \sum_{i=1}^n \lambda_i^k, \quad k = 1, \dots, n.$$

Por las fórmulas de Newton se tiene que:

$$s_r + a_1s_{r-1} + \dots + a_{r-1}s_1 + ra_r = 0, \quad r = 1, \dots, n.$$

Por otra parte,

$$\lambda_i \in Sp(A) \Rightarrow \lambda_i^k \in Sp(A^k) \Rightarrow s_k = tr(A^k).$$

Entonces, el método de Leverrier consiste en:

1. Calcular  $s_k = \text{tr}(A^k)$ ,  $k = 1, \dots, n$ .
2. Calcular los coeficientes  $a_k$  mediante:

$$a_1 = -s_1,$$

$$a_r = -\frac{s_r + a_1 s_{r-1} + \dots + a_{r-1} s_1}{r}, \quad r = 2, \dots, n.$$

### 3.2 Método de Krylov (1931).-

Dada una matriz  $A$ , por el Teorema de Cayley-Hamilton:

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

Sea  $u \neq 0$  un vector cualquiera. Entonces:

$$A^n u + a_1 A^{n-1} u + \dots + a_{n-1} A u + a_n u = 0.$$

Si denotamos  $v^i = A^{n-i} u$ ,  $i = 1, \dots, n$ , se tiene:

$$\sum_{i=1}^n a_i v^i = -A^n u.$$

Consideremos la matriz  $V$  que tiene como columnas  $i$ -ésimas los vectores  $v^i$  y el vector  $a$  que tiene como coordenadas los coeficientes  $a_i$ , entonces la igualdad anterior puede escribirse:

$$Va = -A^n u.$$

Si elegimos  $u$  de manera que el sistema  $\{v^1, \dots, v^n\}$  sea linealmente independiente y la matriz  $V$  fácilmente inversible, entonces los coeficientes del polinomio característico son solución del sistema lineal anterior.

Entonces, el método de Krylov consiste en:

1. Elegir  $u$  adecuado. (Por ejemplo,  $u = (1, 0, \dots, 0)$ .)

2. Calcular la matriz  $V = (v^1 | \dots | v^n)$  mediante:

$$v^n = u,$$

$$v^k = Av^{k+1}, \quad k = n-1, \dots, 1.$$

3. Calcular  $b = A^n u$  mediante:

$$b = Av^1.$$

4. Calcular los coeficientes  $(a_1, \dots, a_n)$  resolviendo el sistema de ecuaciones lineales  $Va = -b$ .

**Observación 2** .- *Los métodos que se utilizan en la práctica para el cálculo de autovalores son métodos que no precisan del polinomio característico (lo cual conduciría a inestabilidades numéricas).*

*De hecho, se utilizan los métodos de cálculo de autovalores para obtener las raíces de un polinomio cualquiera. Basta tener en cuenta que todo polinomio*

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

*es polinomio característico de su matriz de compañía*

$$A = \begin{pmatrix} -a_1 & -a_2 & \dots & -a_{n-1} & -a_n \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

## 4 EL METODO DE LA POTENCIA

Entre los métodos que no precisan del cálculo del polinomio característico destaca por su sencillez el *método de la potencia iterada* (Wielandt - 1944) y todas sus variantes.

### 4.1 El método de la potencia iterada.-

Se utiliza para calcular el autovalor de mayor módulo (*autovalor dominante*) de una matriz diagonalizable. Se considera  $A \in M_{n \times n}(R)$  una matriz diagonalizable con autovalores  $\lambda_1, \lambda_2, \dots, \lambda_n$ , que supondremos ordenados en la forma:

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

Esta hipótesis implica la existencia de una matriz inversible  $P = (p_1 | p_2 | \dots | p_n)$  tal que  $P^{-1}AP = \text{diag}(\lambda_i)$ . Por tanto, el sistema  $\{p_1, p_2, \dots, p_n\}$  es una base de  $C^n$  formada por autovectores:

$$Ap_i = \lambda_i p_i, \quad \forall i = 1, \dots, n,$$

donde  $p_i = (p_{i1}, p_{i2}, \dots, p_{in})$ .

El método de la potencia iterada permite calcular el autovalor dominante  $\lambda_1$  y se basa en la construcción de una sucesión  $\{u^k\}$ , con  $u^k = (u_1^k, u_2^k, \dots, u_n^k)$ , en la forma:

$$u^0 \in C^n \text{ arbitrario,}$$
$$u^{k+1} = Au^k, \quad k = 0, 1, \dots$$

Vamos a estudiar varios de los casos posibles:

**A)**  $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$ . (Por tanto,  $\lambda_1 \in R$ .)

Se tiene el siguiente resultado:

**Teorema 2** .- Si elegimos  $u^0$  adecuadamente (en concreto, si:

$$u^0 = \alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_n p_n$$

basta tomar  $\alpha_1 \neq 0$ ) entonces existe, al menos, un índice  $i \in \{1, \dots, n\}$  tal que:

$$\lim_{k \rightarrow \infty} \frac{u_i^{k+1}}{u_i^k} = \lambda_1.$$

**B)**  $|\lambda_1| = |\lambda_2| = \dots = |\lambda_r| > |\lambda_{r+1}| \geq \dots \geq |\lambda_n|$ ,  
 $\lambda_1 = \lambda_2 = \dots = \lambda_r \in R$ .

Se tiene el siguiente resultado:

**Teorema 3** .- Si elegimos  $u^0$  adecuadamente (en concreto, si:

$$u^0 = \alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_n p_n$$

basta tomar  $\alpha_1 p_1 + \dots + \alpha_r p_r \neq 0$ ) entonces existe, al menos, un índice  $i \in \{1, \dots, n\}$  tal que:

$$\lim_{k \rightarrow \infty} \frac{u_i^{k+1}}{u_i^k} = \lambda_1.$$

**C)**  $|\lambda_1| = |\lambda_2| > |\lambda_3| \geq \dots \geq |\lambda_n|$ ,  $\lambda_1 = -\lambda_2 \in \mathbb{R}$ .

Se tiene el siguiente resultado:

**Teorema 4** .- Si elegimos  $u^0$  adecuadamente (en concreto, si:

$$u^0 = \alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_n p_n$$

basta tomar  $\alpha_1 p_1 + \alpha_2 p_2 \neq 0$ ) entonces existe, al menos, un índice  $i \in \{1, \dots, n\}$  tal que:

$$\lim_{k \rightarrow \infty} \frac{u_i^{2k+2}}{u_i^{2k}} = \lambda_1^2 = \mu.$$

(Por tanto,  $\lambda_1 = \sqrt{\mu}$ ,  $\lambda_2 = -\sqrt{\mu}$ .)

**Observación 3** .- También, si se elige  $u^0$  tal que  $\alpha_1 p_1 - \alpha_2 p_2 \neq 0$ , entonces existe, al menos, un índice  $i \in \{1, \dots, n\}$  tal que:

$$\lim_{k \rightarrow \infty} \frac{u_i^{2k+3}}{u_i^{2k+1}} = \lambda_1^2.$$

**D)**  $|\lambda_1| = |\lambda_2| > |\lambda_3| \geq \dots \geq |\lambda_n|$ ,  $\lambda_1 = \bar{\lambda}_2 \in \mathbb{C}$ .

Se puede adaptar el método para calcular  $r$  y  $\theta$  tales que:

$$\lambda_1 = r e^{i\theta} \quad (\Rightarrow \lambda_2 = r e^{-i\theta})$$

**Observación 4 .-** (*Potencia normalizada*)

*Incluso en el caso más sencillo  $|\lambda_1| > |\lambda_j|$ ,  $j \neq 1$ , si  $\lambda_1$  es muy grande en módulo (o muy pequeño), al calcular  $u^k$  para  $k$  grande, este tiende en norma a  $\infty$  (ó a 0) con los consiguientes problemas numéricos.*

*Para evitar esta dificultad es conveniente trabajar con el vector normalizado  $y^k$ , esto es, para  $u^0 \in C^n$  arbitrario, se definen:*

$$y^0 = \frac{u^0}{\|u^0\|},$$

$$y^{k+1} = \frac{Ay^k}{\|Ay^k\|}, \quad k = 0, 1, \dots$$

*Se puede probar entonces que:*

$$y^k = \frac{u^k}{\|u^k\|}, \quad k = 0, 1, \dots$$

*(por tanto,  $\|y^k\| = 1$ ) y que:*

$$\lim_{k \rightarrow \infty} \|Ay^k\| = |\lambda_1|.$$

*Además:*

$$\lim_{k \rightarrow \infty} y^k = p, \quad (\text{para } \lambda_1 > 0),$$

$$\lim_{k \rightarrow \infty} (-1)^k y^k = p, \quad (\text{para } \lambda_1 < 0),$$

*donde  $p$  es un autovector unitario asociado a  $\lambda_1$ .*

*Por tanto, por el método de la potencia normalizada se calculan simultáneamente el autovalor dominante y un autovector asociado.*

## 4.2 Aceleración de la convergencia.-

La velocidad de convergencia del algoritmo depende, en el caso de un autovalor dominante, de la magnitud del cociente  $|\frac{\lambda_2}{\lambda_1}|$ , siendo mayor la velocidad cuanto menor sea el cociente.

Un método para acelerar la convergencia es el **método de la potencia trasladada** que consiste en aplicar el método de la potencia a la matriz  $(A - pI)$  en lugar de a la matriz  $A$ .

Dado que los autovalores de  $(A - pI)$  son  $\lambda_i - p$ ,  $i = 1, \dots, n$ , se elige  $p$  de forma que  $\lambda_1 - p$  sea el autovalor dominante y que el nuevo cociente  $|\frac{\lambda_2 - p}{\lambda_1 - p}|$  sea mucho menor que el antiguo. De esta forma, el método de la potencia aplicado a la matriz  $(A - pI)$  convergerá más rápidamente a  $\lambda_1 - p$ . Finalmente, a partir de  $\lambda_1 - p$  se determina  $\lambda_1$ .

### Ejemplo 3 .-

*Supongamos  $A \in M_{2 \times 2}(R)$  con autovalores*

$$\lambda_1 = 14, \lambda_2 = 12 \Rightarrow \left| \frac{\lambda_2}{\lambda_1} \right| = \frac{12}{14} = 0.857$$

*Si tomamos  $p = 11$  el autovalor dominante es  $\lambda_1 - p$ .*

*Entonces:*

$$\left| \frac{\lambda_2 - p}{\lambda_1 - p} \right| = \frac{12 - 11}{14 - 11} = \frac{1}{3} = 0.333 \ll 0.857$$

Por tanto, hay convergencia más rápida al autovalor dominante  $\lambda_1 - p = 3 \Rightarrow \lambda_1 = 3 + p = 14$ .

Pero si tomamos  $p = 15$  entonces el autovalor dominante pasa a ser  $\lambda_2 - p$ . En este caso:

$$\left| \frac{\lambda_1 - p}{\lambda_2 - p} \right| = \left| \frac{14 - 15}{12 - 15} \right| = \frac{1}{3} = 0.333 \ll 0.857$$

Por tanto, hay convergencia más rápida al autovalor dominante  $\lambda_2 - p = -3 \Rightarrow \lambda_2 = -3 + p = 12$ .

Así pues, mediante una elección adecuada de  $p$  se puede obtener por el método de la potencia trasladada también el autovalor de menor módulo.

Otra posibilidad para acelerar la convergencia del método de la potencia es utilizar el **método de Aitken** (ya explicado anteriormente) consistente en construir una nueva sucesión:

$$w_i^k = \frac{u_i^k u_i^{k+2} - (u_i^{k+1})^2}{u_i^k - 2u_i^{k+1} + u_i^{k+2}}, \quad k = 0, 1, \dots$$

Finalmente, si la matriz  $A$  es simétrica con autovalores  $\lambda_1 > \dots > \lambda_n$ , se puede utilizar el **método de Rayleigh**.

Para ello, dado un vector no nulo  $x \in R^n$ , se define el *cociente de Rayleigh* de  $x$  como el número:

$$\xi = \frac{x^t A x}{x^t x}.$$

Dado que  $\lambda_1 = \max_{x \neq 0} \frac{x^t A x}{x^t x}$ , se prueba que la sucesión  $\{\xi^k\}$  construida mediante:

$$\xi^k = \frac{(u^k)^t A u^k}{(u^k)^t u^k} = \frac{(u^k)^t u^{k+1}}{(u^k)^t u^k}, \quad k = 0, 1, \dots$$

converge a  $\lambda_1$  más rápidamente que la sucesión  $\left\{ \frac{u_i^{k+1}}{u_i^k} \right\}$ .

### 4.3 Cálculo de autovalores intermedios: deflación.-

Hemos visto que empleando el método de la potencia iterada podemos aproximar el autovalor dominante  $\lambda_1$ . Es natural preguntarse si se puede utilizar el conocimiento de  $\lambda_1$  y de un autovector asociado  $p_1$  para calcular el resto de los autovalores.

Una clase de métodos, que utilizan esta información para transformar  $A$  en otra matriz donde se ha eliminado el autovalor dominante  $\lambda_1$  sin variar el resto de los autovalores, son los llamados *métodos de deflación*, de los cuales veremos dos ejemplos:

#### Método de Hotelling.-

Si  $A$  es una matriz normal entonces el Teorema de Schur asegura que existe una base ortonormal de autovectores  $\{p_1, p_2, \dots, p_n\}$ , esto es, tal que:

$$p_i^* p_j = \delta_{ij}, \quad \forall i, j = 1, \dots, n.$$

Se construye entonces la matriz:

$$A_1 = A - \lambda_1 p_1 p_1^*$$

que verifica:

$$\begin{aligned} A_1 p_j &= 0, & \text{si } j = 1, \\ A_1 p_j &= \lambda_j p_j, & \text{si } j \neq 1. \end{aligned}$$

Por tanto,  $Sp(A_1) = \{0, \lambda_2, \dots, \lambda_n\}$ .

Aplicando el método de la potencia iterada a la matriz  $A_1$  se obtiene su autovalor dominante, que es el autovalor intermedio  $\lambda_2$ .

### **Método de transformación por semejanza.-**

Sea  $\lambda_1$  autovalor de  $A$  con autovector asociado  $p_1$ . Sea  $H$  la matriz de Householder tal que  $H p_1 = \|p_1\|_2 e_1$ . Entonces, teniendo en cuenta que  $H^t = H^{-1} = H$ , se verifica:

$$H A H^{-1} = \left( \begin{array}{c|c} \lambda_1 & c \\ \hline 0 & B \end{array} \right)$$

con  $B \in M_{(n-1) \times (n-1)}$  tal que  $Sp(B) = \{\lambda_2, \dots, \lambda_n\}$ .

Aplicando el método de la potencia iterada a la matriz  $B$  se obtiene su autovalor dominante, que es el autovalor intermedio  $\lambda_2$ .

### **4.4 El método de la potencia inversa.-**

Este método se utiliza para, dado un valor arbitrario  $q$ , calcular el autovalor de  $A$  más próximo a  $q$ .

Supongamos que dicho autovalor es  $\lambda_s$ , es decir:

$$|\lambda_s - q| < |\lambda_i - q|, \quad \forall i \neq s.$$

Entonces se tiene que:

$$\left| \frac{1}{\lambda_s - q} \right| > \left| \frac{1}{\lambda_i - q} \right|, \quad \forall i \neq s.$$

Dado que los autovalores de la matriz  $(A - qI)^{-1}$  son  $\frac{1}{\lambda_i - q}$ ,  $i = 1, \dots, n$ , se tiene que  $\frac{1}{\lambda_s - q}$  es el autovalor dominante de la matriz  $(A - qI)^{-1}$ .

Por tanto, para calcular  $\frac{1}{\lambda_s - q}$  basta aplicar el método de la potencia iterada a la matriz  $(A - qI)^{-1}$ . Entonces, eligiendo  $u^0 \in C^n$ , se contruye la sucesión:

$$u^{k+1} = (A - qI)^{-1}u^k, \quad k = 0, 1, \dots$$

Por lo visto anteriormente, si  $u^0$  está adecuadamente elegido, existe, al menos, un índice  $i \in \{1, \dots, n\}$  tal que:

$$\lim_{k \rightarrow \infty} \frac{u_i^{k+1}}{u_i^k} = \frac{1}{\lambda_s - q} = \eta.$$

Finalmente, a partir de  $\eta$  se obtiene  $\lambda_s$ :

$$\lambda_s = q + \frac{1}{\eta}.$$

Como es habitual, en cada iteración no se calcula la matriz inversa  $(A - qI)^{-1}$  sino que se resuelve el S.E.L.

$$(A - qI)u^{k+1} = u^k.$$

Como la matriz del sistema es la misma en todas las iteraciones, es conveniente utilizar un método de factorización (por ejemplo,  $LU$ ) para resolver el sistema, ya que dicha factorización se realiza una única vez.

Además, el método permite calcular un autovector asociado  $p$ , pues:

$$\lim_{k \rightarrow \infty} \frac{(\lambda_s - q)^k}{|\lambda_s - q|^k} \frac{u^k}{\|u^k\|} = p.$$

**Observación 5** .- *En el caso en que tanto  $\lambda_s$  como  $q$  son reales, esta expresión coincide con la obtenida por el método de la potencia normalizada:*

$$\lim_{k \rightarrow \infty} \frac{u^k}{\|u^k\|} = p, \quad (\text{si } \lambda_s > q),$$

$$\lim_{k \rightarrow \infty} (-1)^k \frac{u^k}{\|u^k\|} = p, \quad (\text{si } \lambda_s < q).$$

## 5 METODOS DE REDUCCION

La idea fundamental de estos métodos consiste en transformar la matriz  $A$  en otra semejante (por lo tanto, con los mismos autovalores) para la cual sea más sencillo calcular el polinomio característico.

Se estudiará un método de reducción a matrices más simples (Householder) y dos métodos para el cálculo del polinomio característico de las matrices simplificadas (Givens y Hyman).

### 5.1 El método de transformación de Householder.-

El método transforma una matriz  $A$  cualquiera en otra semejante *de tipo Hessenberg superior*, esto es, con ceros por debajo de la subdiagonal:

$$\begin{pmatrix} + & + & \dots & \dots & + \\ + & + & \dots & \dots & + \\ & + & \ddots & & + \\ & & \ddots & \ddots & \vdots \\ 0 & & & + & + \end{pmatrix}$$

Dado que son matrices semejantes, sus polinomios característicos son iguales, y por tanto también lo son sus autovalores.

Dada  $A \in M_{n \times n}(R)$ , se determinarán  $(n - 2)$  matrices de Householder  $H_1, \dots, H_{n-2}$  (ortogonales y simétricas) tales que, partiendo de  $A_1 = A$ , cada una de las matrices:

$$A_{k+1} = H_k^{-1} A_k H_k = H_k A_k H_k, \quad k = 1, \dots, n - 2,$$

tenga ceros en la columna  $k$ -ésima por debajo de la subdiagonal.

De esta forma, la matriz:

$$\begin{aligned} A_{n-1} &= H_{n-2}^{-1} \dots H_1^{-1} A H_1 \dots H_{n-2} \\ &= (H_1 \dots H_{n-2})^{-1} A (H_1 \dots H_{n-2}) \end{aligned}$$

es de tipo Hessenberg superior y semejante a  $A$ .

**Observación 6** .- Si la matriz  $A$  es simétrica, entonces también lo son todas las matrices  $A_k$  y, por tanto,  $A_{n-1}$  es una matriz tridiagonal simétrica, pues también tiene ceros por encima de la superdiagonal.

El paso de la matriz  $A_k$  a la matriz  $A_{k+1}$  se realiza de la siguiente manera:

Se toma el vector  $a_k \in R^{n-k}$  formado por los elementos de la columna  $k$ -ésima de  $A_k$  a partir del subdiagonal (inclusive). Se elige la matriz de Householder  $H(\tilde{v}_k) \in M_{(n-k) \times (n-k)}(R)$  tal que  $H(\tilde{v}_k)a_k = \|a_k\|_2 e_1$ . Se construye entonces la nueva matriz de Householder:

$$H_k = \left( \begin{array}{c|c} I_k & 0 \\ \hline 0 & H(\tilde{v}_k) \end{array} \right) = H(v_k), \quad \text{con } v_k = \begin{pmatrix} 0 \\ \tilde{v}_k \end{pmatrix} \in R^n.$$

Entonces la matriz  $A_{k+1} = H_k A_k H_k$  tiene ceros en la columna  $k$ -ésima por debajo de la subdiagonal.

**Observación 7** .- En la práctica no se calcula  $H_k$  sino directamente el producto  $H_k A_k H_k$ . Definiendo:

$$w_k = \frac{v_k}{\|v_k\|_2},$$

$$q_k = 2(I - w_k w_k^t) A_k w_k,$$

$$p_k = 2(I - w_k w_k^t) A_k^t w_k,$$

se tiene:

$$A_{k+1} = A_k - w_k p_k^t - q_k w_k^t.$$

Además, si  $A$  es simétrica, entonces  $p_k = q_k$ , y por tanto:

$$A_{k+1} = A_k - w_k p_k^t - (w_k p_k^t)^t.$$

Todo lo anterior puede resumirse en los siguientes resultados:

**Teorema 5** .- Dada una matriz  $A \in M_{n \times n}(R)$ , existe una matriz ortogonal  $H$ , producto de  $(n - 2)$  matrices de Householder, tal que  $H^t A H$  es de tipo Hessenberg superior.

**Corolario 2** .- Dada una matriz  $A \in M_{n \times n}(R)$  simétrica, existe una matriz ortogonal  $H$ , producto de  $(n - 2)$  matrices de Householder, tal que  $H^t A H$  es tridiagonal simétrica.

## 5.2 El método de Givens para matrices tridiagonales simétricas.-

Sea la matriz tridiagonal simétrica:

$$B = \begin{pmatrix} b_1 & c_1 & & & 0 \\ c_1 & b_2 & c_2 & & \\ & c_2 & \ddots & \ddots & \\ & & \ddots & b_{n-1} & c_{n-1} \\ 0 & & & c_{n-1} & b_n \end{pmatrix}$$

Supondremos que todos los elementos subdiagonales verifican:

$$c_i \neq 0, \quad \forall i = 1, \dots, n-1.$$

(Si algún  $c_i$  fuese nulo, se podría descomponer  $B$  en dos matrices en las condiciones anteriores).

Denotaremos por  $B_i$ ,  $i = 1, \dots, n$ , la submatriz principal de  $B$  formada por sus primeras  $i$  filas e  $i$  columnas. Sean  $p_i(\lambda)$ ,  $i = 0, \dots, n$ , los polinomios definidos por recurrencia:

$$p_0(\lambda) = 1,$$

$$p_1(\lambda) = b_1 - \lambda,$$

$$p_i(\lambda) = (b_i - \lambda)p_{i-1}(\lambda) - c_{i-1}^2 p_{i-2}(\lambda), \quad i = 2, \dots, n.$$

Se verifica entonces:

### **Teorema 6 .-**

$$p_i(\lambda) = \det(B_i - \lambda I), \quad \forall i = 1, \dots, n.$$

*Además, si  $q_i(\lambda) = (-1)^i p_i(\lambda)$ ,  $i = 0, 1, \dots, n$ , entonces la sucesión  $\{q_n, q_{n-1}, \dots, q_1, q_0\}$  es una sucesión de Sturm relativa al polinomio característico  $p_n(\lambda) = \det(B - \lambda I)$  y a cualquier intervalo.*

Por tanto, utilizando la sucesión de Sturm se pueden localizar y calcular las raíces de  $p_n(\lambda)$ , que son los autovalores de  $B$ , dentro del intervalo  $[-\|B\|, \|B\|]$ .

### 5.3 El método de Hyman para matrices Hessenberg superior.-

Sea  $B \in M_{n \times n}(R)$  una matriz Hessenberg superior tal que todos sus elementos subdiagonales verifiquen:

$$b_{i+1,i} \neq 0, \quad \forall i = 1, \dots, n-1.$$

Consideremos la matriz:

$$B - \lambda I = (p_1 | p_2 | \dots | p_n).$$

Si  $k(\lambda)$  es un polinomio cuyos coeficientes dependen de  $B$ , el sistema lineal:

$$(B - \lambda I) x = k(\lambda) e_1$$

se puede resolver de manera sencilla por sustitución retrógrada si elegimos arbitrariamente  $x_n$  (por ej:  $x_n = 1$ ), dado que  $B - \lambda I$  sigue siendo Hessenberg superior.

Si  $\lambda$  es una raíz de  $k(\lambda)$  entonces se tiene  $(B - \lambda I) x = 0$ , de modo que  $\lambda$  es un autovalor de  $B$  y  $x$  un autovector asociado. Por tanto  $k(\lambda)$  es, salvo un factor, el polinomio característico  $\det(B - \lambda I)$ .

Tomando  $x_n = 1$  es fácil calcular  $\det(B - \lambda I)$  en función de  $k(\lambda)$ :

$$\begin{aligned} (B - \lambda I) x &= k(\lambda) e_1 \\ \Leftrightarrow x_1 p_1 + x_2 p_2 + \dots + x_{n-1} p_{n-1} + p_n &= k(\lambda) e_1 \\ \Leftrightarrow p_n &= k(\lambda) e_1 - \sum_{i=1}^{n-1} x_i p_i \end{aligned}$$

Entonces:

$$\begin{aligned}
 \det(B - \lambda I) &= \det(p_1 | \dots | p_{n-1} | p_n) \\
 &= \det(p_1 | \dots | p_{n-1} | k(\lambda)e_1 - \sum_{i=1}^{n-1} x_i p_i) \\
 &= \det(p_1 | \dots | p_{n-1} | k(\lambda)e_1)
 \end{aligned}$$

$$\begin{aligned}
 &= \begin{vmatrix} b_{11} - \lambda & b_{12} & \dots & b_{1,n-1} & k(\lambda) \\ b_{21} & b_{22} - \lambda & \dots & b_{2,n-1} & 0 \\ & b_{32} & \ddots & b_{3,n-1} & 0 \\ & \ddots & \ddots & \vdots & \vdots \\ & & b_{n-1,n-2} & b_{n-1,n-1} - \lambda & 0 \\ 0 & & & b_{n,n-1} & 0 \end{vmatrix} \\
 &= (-1)^{n+1} k(\lambda) b_{21} b_{32} \dots b_{n,n-1}
 \end{aligned}$$

A fin de calcular  $k(\lambda)$  el método más sencillo es resolver por sustitución retrógrada el sistema lineal:

$$\left\{ \begin{array}{l} (b_{11} - \lambda)x_1 + b_{12}x_2 + \dots + b_{1n}x_n = k(\lambda) \\ b_{21}x_1 + (b_{22} - \lambda)x_2 + \dots + b_{2n}x_n = 0 \\ \quad b_{32}x_2 + \dots + b_{3n}x_n = 0 \\ \quad \quad \dots \quad \quad \dots \quad \quad \dots \\ \quad \quad \quad \quad \quad b_{n,n-1}x_{n-1} + (b_{nn} - \lambda)x_n = 0 \end{array} \right.$$

con  $x_n = 1$ .

Se tiene entonces el siguiente resultado:

**Teorema 7 .-** *Sea  $B \in M_{n \times n}(R)$  una matriz Hessenberg superior con todos sus elementos subdiagonales sean no nulos. Sea la sucesión de polinomios:*

$$x_n(\lambda) = 1,$$

$$x_i(\lambda) = -\frac{1}{b_{i+1,i}}[(b_{i+1,i+1} - \lambda)x_{i+1}(\lambda) + b_{i+1,i+2}x_{i+2}(\lambda) + \dots + b_{i+1,n}x_n(\lambda)], \quad i = n-1, n-2, \dots, 1.$$

*Entonces:*

$$\det(B - \lambda I) = (-1)^{n+1}k(\lambda)b_{21}b_{32} \dots b_{n,n-1},$$

*donde:*

$$k(\lambda) = (b_{11} - \lambda)x_1(\lambda) + b_{12}x_2(\lambda) + \dots + b_{1n}x_n(\lambda).$$

De esta forma se puede obtener  $k(\lambda)$  y aproximar sus raíces (los autovalores de  $B$ ), por los métodos ya vistos para polinomios.

## 6 METODOS DE FACTORIZACION

Son métodos basados en las factorizaciones ya estudiadas para la matriz  $A$ : la factorización  $LU$  y la  $QR$ , y que permiten calcular simultáneamente todos los autovalores de  $A$ . Estudiaremos los dos métodos mas simples:

### 6.1 El método $LR$ (Rutishauser - 1952) .-

Se basa en la factorización  $LU$  de una matriz. El algoritmo es como sigue:

Sea  $A$  una matriz que admita factorización  $LU$ .

Se toma  $A_1 = A$  y, a partir de la factorización  $LU$  de  $A_k = L_k R_k$ , se construye  $A_{k+1} = R_k L_k$ .

De esta la forma se tiene la sucesión de matrices  $\{A_k\}$  todas semejantes a  $A$ :

$$\begin{aligned} A_{k+1} &= R_k L_k = L_k^{-1} A_k L_k = \dots \\ &= L_k^{-1} \dots L_1^{-1} A_1 L_1 \dots L_k = (L_1 \dots L_k)^{-1} A (L_1 \dots L_k). \end{aligned}$$

**Teorema 8 .-** *Sea  $A \in M_{n \times n}(R)$  una matriz inversible con todos sus autovalores de módulo diferente:*

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0.$$

*Existe, al menos, una matriz inversible  $P$  tal que  $P^{-1}AP = \text{diag}(\lambda_i)$ .*

*Supongamos que las matrices  $P$  y  $P^{-1}$  admiten factorización  $LU$ . Entonces, la sucesión  $\{A_k\}$  generada por el método  $LR$  converge a una matriz triangular superior. En particular:*

$$\begin{aligned} \lim_{k \rightarrow \infty} (A_k)_{ii} &= \lambda_i, \quad 1 \leq i \leq n, \\ \lim_{k \rightarrow \infty} (A_k)_{ij} &= 0, \quad 1 \leq j < i \leq n. \end{aligned}$$

## 6.2 El método $QR$ (Francis, Kublanovskaya - 1960) .-

Se basa en la factorización  $QR$  de una matriz. El algoritmo es como sigue:

Sea  $A$  una matriz cualquiera.

Se toma  $A_1 = A$  y, a partir de la factorización  $QR$  de  $A_k = Q_k R_k$ , se construye  $A_{k+1} = R_k Q_k$ .

De esta la forma se tiene la sucesión de matrices  $\{A_k\}$  todas semejantes a  $A$ :

$$\begin{aligned} A_{k+1} &= R_k Q_k = Q_k^{-1} A_k Q_k = \dots \\ &= Q_k^{-1} \dots Q_1^{-1} A_1 Q_1 \dots Q_k = (Q_1 \dots Q_k)^{-1} A (Q_1 \dots Q_k). \end{aligned}$$

**Teorema 9 .-** *Sea  $A \in M_{n \times n}(R)$  una matriz inversible con todos sus autovalores de módulo diferente:*

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0.$$

*Existe, al menos, una matriz inversible  $P$  tal que  $P^{-1}AP = \text{diag}(\lambda_i)$ .*

*Supongamos que la matriz  $P^{-1}$  admite factorización  $LU$ . Entonces, la sucesión  $\{A_k\}$  generada por el método  $QR$  verifica:*

$$\begin{aligned} \lim_{k \rightarrow \infty} (A_k)_{ii} &= \lambda_i, \quad 1 \leq i \leq n, \\ \lim_{k \rightarrow \infty} (A_k)_{ij} &= 0, \quad 1 \leq j < i \leq n. \end{aligned}$$

**Observación 8** .- *No se puede hablar en el caso del algoritmo QR de la convergencia de la sucesión  $\{A_k\}$ , ya que no se puede asegurar la existencia de los límites:*

$$\lim_{k \rightarrow \infty} (A_k)_{ij}, \quad 1 \leq i < j \leq n.$$

*En cambio, si la matriz  $A$  es simétrica, también lo son todas las  $A_k$  y, por tanto:*

$$\exists \lim_{k \rightarrow \infty} (A_k)_{ij} = 0, \quad 1 \leq i < j \leq n,$$

*esto es, la sucesión  $\{A_k\}$  converge a la matriz diagonal de autovalores.*

*(Esto último no se verifica para el algoritmo LR.)*

