

Autovalori

Sia ^(*) $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{C}$ è *autovalore* di A se $\exists \underline{x} \in \mathbb{R}^n$, $\underline{x} \neq 0$:

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

Il vettore \underline{x} si dice *autovettore* di A associato a λ .

Per ogni autovalore λ , l'insieme degli autovettori ad esso associati, compreso l'autovettore nullo, forma un sottospazio di \mathbb{R}^n , noto come l'*autospazio* associato a λ .

^(*) Siano V e W due spazi vettoriali sullo stesso campo K . Una funzione $f : V \rightarrow W$ è una **trasformazione lineare** se soddisfa le seguenti proprietà

- $f(x + y) = f(x) + f(y)$ (*additività*)
- $f(ax) = af(x)$ (*omogeneità*)

per ogni coppia di vettori x e y in V e per ogni scalare a in K .

Equivalentemente, f è lineare se preserva le combinazioni lineari, ovvero se

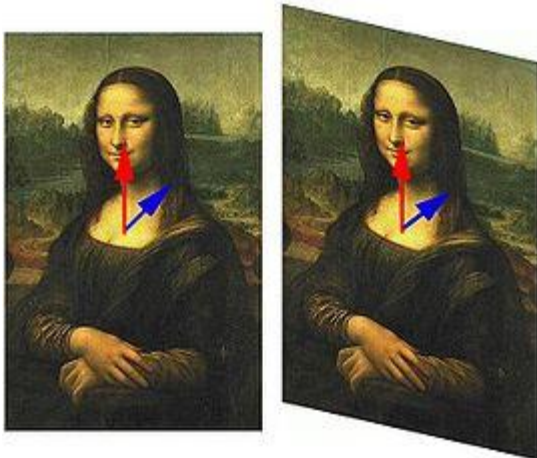
$$f(a_1x_1 + \dots + a_mx_m) = a_1f(x_1) + \dots + a_mf(x_m)$$

per ogni intero positivo m e ogni scelta dei vettori x_1, \dots, x_m e degli scalari a_1, \dots, a_m

Una matrice A di tipo $m \times n$ con valori reali definisce una trasformazione lineare

$$L_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$$
$$L_A(v) = Av$$

dove Av è il prodotto di A e v .



In questa trasformazione lineare della Gioconda, l'immagine è modificata ma l'asse centrale verticale rimane fisso. Il vettore blu ha cambiato lievemente direzione, mentre quello rosso no. Quindi il vettore rosso è un **autovettore** della trasformazione e quello blu no. Inoltre, poiché il vettore rosso non è stato né allungato, né compresso, né ribaltato, il suo **autovalore** è 1. Tutti i vettori sull'asse verticale sono multipli scalari del vettore rosso, e sono tutti autovettori: assieme all'origine formano l'**autospazio** relativo all'autovalore 1.

Un autovettore è sempre $\neq 0$. Un autovalore e' nullo sse A è singolare.

Spettro di A : $\sigma(A)$ insieme degli autovalori di A .

Raggio spettrale: $\rho = \max_{1 \leq j \leq n} |\lambda_j|$

Polinomio caratteristico: $\det(A - \lambda I)$ di grado n in λ .

Equazione caratteristica: $\det(A - \lambda I) = 0$.

Gli autovalori di A sono tutte e sole le radici dell'equazione caratteristica. Poiché il polinomio caratteristico è di grado n

in λ , A ha n autovalori non necessariamente distinti. A ha almeno una coppia autovalore-autovettore e poiché:

$Ax = \lambda x \Leftrightarrow A\alpha x = \lambda\alpha x$, A ha infiniti autovettori.

Il problema è quindi quello di determinare gli autovettori linearmente indipendenti.

Ricordiamo che i vettori v_1, \dots, v_n appartenenti ad uno spazio vettoriale si dicono **linearmente indipendenti** se nessuno di questi può essere espresso come una combinazione lineare degli altri.

Si dice *molteplicità algebrica* di un autovalore, la sua molteplicità come radice del polinomio caratteristico.

Si dice *molteplicità geometrica* di un autovalore, il numero di vettori linearmente indipendenti associati ad esso.

Se gli autovalori sono n , ci sono n autovettori linearmente indipendenti. Se la molteplicità algebrica di qualche autovalore è maggiore di uno, può essere che il numero di autovettori linearmente indipendenti sia minore di n .

Esempio.

Siano: $A = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}$, $B = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$. Per entrambe, $\lambda = 1$ con molteplicità 2

ma: $\begin{vmatrix} 1 \\ 0 \end{vmatrix}$ è autovettore di A e $\begin{vmatrix} 0 \\ 1 \end{vmatrix}$, $\begin{vmatrix} 1 \\ 0 \end{vmatrix}$ sono autovettori di B .

Ricordiamo alcune proprietà degli autovalori.

- A ed A^T hanno gli stessi autovalori.
- $\det(A) = 0 \Leftrightarrow \lambda = 0$
- Se $\det(A) \neq 0 \Rightarrow \exists A^{-1}$ e se μ è autovalore di $A \Rightarrow \mu^{-1}$ autovalore di A^{-1}
- Se μ è autovalore di A , μ^s è autovalore di $A^s \forall s \in \mathbb{Z}$

Siano A e $B \in \mathbb{R}^{n \times n}$, $\exists B^{-1}$ e sia $C = B^{-1}AB$. Allora A e C si dicono *simili*.

Teorema. Due matrici simili hanno lo stesso polinomio caratteristico e quindi gli stessi autovalori.

Teorema. Se A e C sono simili, allora A^s e C^s sono simili per $\forall s \in \mathbb{N}$.

Teorema. Se A e C sono simili e $\det(A) \neq 0$ allora anche $\det(C) \neq 0$ e inoltre A^{-1} , C^{-1} sono simili.

Teorema di Gerschgorin.

$$A \in \mathbb{R}^{n \times n}, \rho_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad \gamma_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq \rho_i\} \quad i = 1, \dots, n$$

$$\gamma = \bigcup_{i=1}^n \gamma_i. \quad \text{Allora, se } \lambda \in \sigma(A) \quad \Rightarrow \quad \lambda \in \gamma$$

Teorema di Hermite.

Se $A \in \mathbb{C}^{n \times n}$, $A = \overline{A}^T$, allora gli autovalori di A sono tutti reali.

Teorema. Se A è simmetrica definita positiva gli autovalori sono tutti reali positivi.

Una matrice si dice *diagonalizzabile* se e' simile ad una matrice diagonale.

Teorema. Una matrice $A \in \mathbb{C}^{n \times n}$ è diagonalizzabile se e solo se ha n autovettori linearmente indipendenti.

◦

Matrice *unitaria*: $\overline{U}^T U = U \overline{U}^T = I$

Matrice *ortogonale*: $U^T U = U U^T = I$

Teorema di Schur.

$A \in \mathbb{C}^{n \times n} \Rightarrow \exists U \text{ unitaria} : T = \overline{U}^T A U$, dove T è triangolare superiore.

NB: Se A è reale allora U è ortogonale.

Teorema. Se A è hermitiana essa è diagonalizzabile.

Teorema. Condizione necessaria e sufficiente perché $A \in \mathbb{C}^{n \times n}$ sia convergente è che $\rho(A) < 1$.

Teorema. Condizione necessaria e sufficiente perché A sia convergente è che sia infinitesima: $\|A^k\| \rightarrow 0$.

Teorema. $\rho(A) \leq \|A\|$

Condizionamento

Propagazione, sugli autovalori, delle perturbazioni degli elementi della matrice.

Per semplificare, sia A diagonalizzabile. Allora $\exists P, P^{-1}$:

$$P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = D.$$

Inoltre la norma matriciale sia tale che:

$\|D\| = \max |d_i|$ che è verificata per le norme matriciali indotte dalle p-norme: $\|x\|_p, 1 \leq p \leq \infty$.

Teorema di Bauer-Fike.

Sia A diagonalizzabile con norma soddisfacente la condizione enunciata. Sia E una matrice di perturbazione per A e siano λ un autovalore di $(A+E)$ e $\lambda_i \in \sigma(A)$.

Allora si ha:

$$\min_{\lambda_i \in \sigma(A)} |\lambda - \lambda_i| \leq \|P\| \|P^{-1}\| \|E\| \quad \square$$

Se A è una matrice normale, ovvero $A^T A = A A^T$, come ad es.: le matrici ortogonali, simmetriche, unitarie, allora la matrice P può essere scelta unitaria (ortogonale nel caso reale, v. teorema di Schur) per cui: $\|P\|_p = \|P^{-1}\|_p = 1$

e quindi:

$$\min_{\lambda_i \in \sigma(A)} |\lambda - \lambda_i| \leq \|E\|_p$$

e in questo caso il problema è ben condizionato.

Per le matrici non normali:

$$K(A) = \|P\| \|P^{-1}\|$$

è il numero di condizionamento del problema degli autovalori di A .

Esempio: matrice di Hilbert.

H_n è simmetrica e quindi normale. Mostriamo come variano l'errore assoluto e quello relativo calcolati con la norma 2.

n	E_A	E_r
1	10^{-3}	10^{-3}
2	$1.6 \cdot 10^{-4}$	$1.4 \cdot 10^{-3}$
4	$5 \cdot 10^{-7}$	$2.2 \cdot 10^{-3}$
8	$1.1 \cdot 10^{-12}$	$3.4 \cdot 10^{-3}$
10	$1.3 \cdot 10^{-15}$	$4 \cdot 10^{-3}$

Vediamo adesso il metodo delle potenze e delle potenze inverse che sono tecniche per ricavare, rispettivamente, l'autovalore di modulo massimo e quello di modulo minimo. Le tecniche per ricavare simultaneamente tutti gli autovalori sono invece basate sulle iterazioni QR, che a loro volta si basano sulle trasformazioni per similarità di Householder e di Givens.

Metodo delle potenze

Supponiamo che la matrice A abbia un autovalore dominante λ_1 tale che:

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

e che abbia n autovettori linearmente indipendenti $\{u_i\}_{i=1}^n$:

$$Au_i = \lambda_i u_i \quad i=1, \dots, n$$

Il metodo delle potenze fornisce un metodo iterativo per approssimare l'autovalore dominante λ_1 .

Ogni vettore $x \in \mathbb{R}^n$ può essere espresso come combinazione lineare degli u_i poiché essi sono linearmente indipendenti.

Sia $x_0 \in \mathbb{R}^n$ un vettore arbitrario che scriviamo come combinazione lineare degli u_i :

$$x_0 = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n \quad (2)$$

Poiché

$$Au_i = \lambda_i u_i$$

si ha:

$$A(Au_i) = \lambda_i Au_i$$

ovvero:

$$A^2 u_i = \lambda_i^2 u_i$$

$$\dots \quad \dots$$

$$A^k u_i = \lambda_i^k u_i \quad (3)$$

Poniamo allora:

$$x_1 = Ax_0$$

$$x_2 = Ax_1 \quad \rightarrow \quad x_2 = A^2 x_0$$

$$\dots \quad \dots$$

$$x_k = Ax_{k-1} \quad \rightarrow \quad x_k = A^k x_0$$

Dalla (2) $x_0 = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n$ e dalla (3) $A^k u_i = \lambda_i^k u_i$ si ha:

$$x_k = A^k x_0 = \alpha_1 A^k u_1 + \dots + \alpha_n A^k u_n = \alpha_1 \lambda_1^k u_1 + \dots + \alpha_n \lambda_n^k u_n$$

che riscriviamo:

$$x_k = \alpha_1 (\lambda_1)^k u_1 + \alpha_2 (\lambda_2)^k u_2 + \dots + \alpha_n (\lambda_n)^k u_n$$

e mettendo in evidenza $\alpha_1 (\lambda_1)^k$, si ha:

$$x_k = \alpha_1 (\lambda_1)^k \left[u_1 + \frac{\alpha_2}{\alpha_1} \left(\frac{\lambda_2}{\lambda_1} \right)^k u_2 + \dots + \frac{\alpha_n}{\alpha_1} \left(\frac{\lambda_n}{\lambda_1} \right)^k u_n \right]$$

E poiché $\lambda_i < \lambda_1$, per $k \rightarrow \infty$ la direzione di:

$$\tilde{x}_k = \frac{x_k}{\lambda_1^k \alpha_1} = u_1 + \sum_{i=2}^n \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1} \right)^k u_i \quad k=1,2,\dots$$

tende ad u_1 al rate con cui $\left(\frac{\lambda_2}{\lambda_1} \right)^k \rightarrow 0$.

Pertanto, poiché: $\lim_{k \rightarrow \infty} \tilde{x}_k = u_1$ e:

$$\tilde{x}_{k+1} = A \tilde{x}_k \rightarrow \lambda_1 u_1$$

il *quoziente di Rayleigh*, definito da:

$$\beta_k = \frac{\tilde{x}_k^T \tilde{x}_{k+1}}{\|\tilde{x}_k\|_2^2} \rightarrow \lambda_1 \frac{u_1 u_1}{u_1 u_1} = \lambda_1$$

tende a λ_1 .

Per evitare problemi di overflow o underflow si usa la versione normalizzata del metodo delle potenze, nella quale tutti i vettori sono divisi per la loro norma in modo che la norma dei vettori così ottenuti sia sempre unitaria.

Caso particolare: esistenza di due autovalori dominanti di modulo uguale (ovvero $|\lambda_1| = |\lambda_2|$).

Si hanno 3 casi:

1) $\lambda_2 = \lambda_1$, il metodo risulta ancora convergente per k sufficientemente grande.

2) $\lambda_2 = -\lambda_1$, il metodo viene applicato alla matrice A^2 poiché $\lambda_i(A^2) = [\lambda_i(A)]^2$ in modo che $\lambda_1^2 = \lambda_2^2$ e si ricade nel caso precedente.

3) $\lambda_2 = \overline{\lambda_1}$ il metodo in questo caso non converge.

Metodo delle potenze inverse

E' un metodo per approssimare autovalori diversi da quello di modulo massimo ed in particolare per approssimare l'autovalore di modulo minimo. Si tratta di applicare il metodo delle potenze alla matrice $(M_\alpha)^{-1} = (A - \alpha I)^{-1}$.

Essendo: $Au_i = \lambda_i u_i$

allora per α diverso da λ_i , autovalore di A , si ha:

$$(A - \alpha I)u_i = (\lambda_i - \alpha)u_i$$

inoltre: $(A - \alpha I)^{-1}u_i = (\lambda_i - \alpha)^{-1}u_i$

pertanto i $\xi_i = (\lambda_i - \alpha)^{-1}$ sono gli autovalori di $(M_\alpha)^{-1}$.

Supponiamo allora che esista un intero j tale che:

$$|\lambda_j - \alpha| < |\lambda_i - \alpha| \quad i=1,2,\dots,n, \quad i \neq j.$$

L'autovalore di modulo massimo di $(M_\alpha)^{-1}$ corrisponde quindi a ξ_j . In particolare, se $\alpha = 0$ questo metodo porta alla stima dell'autovalore di A di modulo minimo.

Trasformazioni per similarità.

Queste sono tecniche per determinare contemporaneamente tutto lo spettro di A.

Siano A e B simili $\Rightarrow \exists S, S^{-1}: A = SBS^{-1}$

pertanto A e B avranno gli stessi autovalori.

Trasformando A in B, perché il metodo sia efficiente è necessario che il calcolo degli autovalori di B sia più semplice, ovvero meno costoso, di quello di A. Ad esempio, se B è triangolare o diagonale i suoi autovalori sono gli elementi diagonali.

Dal teorema di Schur sappiamo che: $\exists U$:

$$A = UTU^T$$

dove T è triangolare superiore.

Il problema sarebbe risolto se fosse possibile determinare U con un numero finito di operazioni. Purtroppo, una conseguenza indiretta del teorema di Abel è che, per $n \geq 5$, la matrice U non può essere calcolata in modo elementare e si deve pertanto ricorrere a metodi iterativi. A tale scopo, l'algoritmo di riferimento è l'iterazione QR.

Metodo QR

Sia $A \in \mathbb{R}^{n \times n}$, data $Q_0 \in \mathbb{R}^{n \times n}$ ortogonale e posto:

$$T_0 = (Q_0)^T A Q_0,$$

per $k=1,2,\dots$ sino a convergenza, l'iterazione QR consiste nel determinare Q_k, R_k tali che:

$$Q_k R_k = T_{k-1} \quad (\text{fattorizzazione QR})$$

dove R_k è triangolare superiore, Q_k è ortogonale, e porre:

$$T_k = R_k Q_k.$$

Si osservi che:

$$T_k = R_k Q_k = (Q_k^T Q_k) R_k Q_k = Q_k^T T_{k-1} Q_k$$

Quindi T_k è ortogonalmente simile T_{k-1} e, ripetendo i passaggi, T_k è ortogonalmente simile a T_0 e quindi ad A .

$$T_k = R_k Q_k = (Q_k)^T (Q_k R_k) Q_k = (Q_k)^T T_{k-1} Q_k = (Q_0 Q_1 \dots Q_k)^T A (Q_0 Q_1 \dots Q_k)$$

Nella versione più elementare del metodo, si pone: $Q_0 = I_n$ in modo che: $T_0 = A$. Se la fattorizzazione QR è ottenuta tramite il metodo di ortogonalizzazione di Gram-Schmidt, il costo di ogni passo sarà dell'ordine di n^3 flops.

Nel caso in cui A abbia autovalori reali e distinti, il limite di T_k è una matrice triangolare superiore. Nel caso generico ciò non si avrà. Poiché l'implementazione del metodo QR presentato prima richiederebbe un costo computazionale di $\approx n^3$, illustriamo una variante dal costo ridotto, basata sulla riduzione della matrice A ad una forma quasi triangolare o di Hessenberg:

$$B = \begin{vmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \end{vmatrix}$$

In questo modo, per determinare ogni T_k nell'algoritmo QR, il costo sarà $\approx n^2$. Per fare questo, utilizziamo il metodo di Householder per ridurre A ad una forma quasi triangolare e poi usiamo il metodo di Givens per eseguire la fattorizzazione QR di T_k .

Metodo di Householder

La matrice di Householder è definita da:

$$H = I - 2uu^T$$

dove $u \in \mathbb{R}^n$ è un vettore colonna.

La matrice di Householder è simmetrica: $H = H^T$ e
ortogonale: $HH^T = I$. Vediamo come procedere per
determinare H.

Supponiamo di voler azzerare gli ultimi $q-1$ elementi di un
vettore colonna.

Sia: $v \in \mathbb{R}^q$ e definiamo: $\tilde{P} = I_q - 2 \frac{vv^T}{v^T v}$

dove: I_q è la matrice identità $q \times q$, $v^T v \in \mathbb{R}$, $vv^T \in \mathbb{R}^{q \times q}$,

$$\text{rank}(vv^T) = 1$$

\tilde{P} è una matrice di Householder e lo è pure P definita da:

$$P = \begin{vmatrix} I_{n-q} & 0 \\ 0 & \tilde{P} \end{vmatrix}$$

dove $P \in \mathbb{R}^{n \times n}$, $\tilde{P} \in \mathbb{R}^{q \times q}$.

Dato $x \in \mathbb{R}^q$ troviamo $v \in \mathbb{R}^q$ tale che:

$$\tilde{P}x \equiv Ke_1, \quad K \in \mathbb{R}, \quad e_1 = \begin{pmatrix} 1 \\ 0 \\ \dots \\ \dots \\ 0 \end{pmatrix} \in \mathbb{R}^q.$$

$$\tilde{P}x = (I_q - 2 \frac{vv^T}{v^T v})x = x - 2 \frac{v^T x}{v^T v} v \equiv Ke_1 \quad \Rightarrow \quad v = x + \alpha e_1$$

dove α è da determinare. Vediamo come determinarlo.

$$v^T x = (x + \alpha e_1)^T x = x^T x + \alpha e_1^T x = x^T x + \alpha x_1$$

$$v^T v = (x + \alpha e_1)^T (x + \alpha e_1) = x^T x + 2\alpha x_1 + \alpha^2$$

$$\tilde{P}x = x - 2 \frac{v^T x}{v^T v} (x + \alpha e_1) = \frac{v^T v - 2v^T x}{v^T v} x - 2\alpha \frac{v^T x}{v^T v} e_1 \equiv Ke_1$$

$$\Rightarrow v^T v - 2v^T x \equiv 0 \quad \text{ovvero:} \quad x^T x + 2\alpha x_1 + \alpha^2 - 2x^T x - 2\alpha x_1 = 0$$

$$\text{Da cui:} \quad \alpha^2 = x^T x \quad \Rightarrow \quad \alpha = \pm \|x\|_2$$

Sia ora $x \in \mathbb{R}^n$ e costruiamo $P_q \in \mathbb{R}^{q \times q}$ in modo tale da annullare gli ultimi $q-1$ elementi di x .

$$P x = \begin{bmatrix} I_{n-q} & 0 \\ 0 & \tilde{P} \end{bmatrix} \begin{pmatrix} x_1 \\ \dots \\ x_{n-q} \\ x_{n-q+1} \\ \dots \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \dots \\ x_{n-q} \\ -\alpha \\ 0 \\ \dots \\ 0 \end{pmatrix}$$

dove: $\alpha = \sqrt{x_{n-q+1}^2 + \dots + x_n^2}$

Data $A \in \Re^{n \times n}$ costruiamo $P_{(1)}, P_{(2)}, \dots, P_{(k)}$ tali che:

$$A_{(k)} = P_{(k)}^T A_{(k-1)} P_{(k)} = (P_{(k)} P_{(k-1)} \dots P_{(1)})^T A (P_{(1)} \dots P_{(k-1)} P_{(k)}) = Q_{(k)}^T A Q_{(k)}$$

dove il vettore x è sostituito dal k -esimo vettore colonna di $A_{(k-1)}$.

Per evitare problemi di cancellazione nel calcolo di

$$v = x \pm \|x\|_2 e_1 \text{ si pone } \alpha = \text{sign}(x_1) \|x\|_2.$$

Metodo di Givens

Analogo al metodo di Householder è il metodo di Givens, più adatto per le matrici sparse in quanto, anziché annullare gli elementi di una colonna, tale metodo permette di annullare un elemento alla volta.

Le matrici di Givens sono matrici di rotazioni ortogonali che hanno la proprietà di annullare selettivamente singoli elementi di una matrice o di un vettore. Fissati due indici i e k e un certo angolo ϑ ,

$$J(i, k, \vartheta) = \begin{pmatrix} 1 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & c & \dots & s & \dots & \dots \\ \dots & \dots & \dots & 1 & \dots & \dots & \dots \\ \dots & \dots & -s & \dots & c & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & 1 & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & 1 \end{pmatrix} \quad \begin{array}{l} \text{---} > i \\ \text{---} > k \end{array} \quad i < k$$

i
 k

dove: $c = \cos \vartheta$ $s = \sin \vartheta$ cosicchè: $J^T J = I$. Infatti:

$$J^T J = \begin{pmatrix} 1 & & & & & 0 \\ & \dots & & & & \\ & & c^2 + s^2 & & & \\ & & & \dots & & \\ & & & & c^2 + s^2 & \\ & & & & & \dots \\ 0 & & & & & & 1 \end{pmatrix} = I$$

Sia $x \in \mathbb{R}^n$, allora $y = J(i, k, \theta)x$ dà:

$$\begin{cases} y_i = cx_i + sx_k \\ y_k = -sx_i + cx_k \\ y_j = x_j \quad j \neq i, k \end{cases}$$

Imponendo che: $c = \frac{x_i}{\sqrt{x_i^2 + x_k^2}} \quad s = \frac{x_k}{\sqrt{x_i^2 + x_k^2}}$

si ha che l'elemento di posto k va a zero:

$$\begin{cases} -sx_i + cx_k = 0 \\ c^2 + s^2 = 1 \end{cases}$$

La fattorizzazione QR eseguita con matrici di Givens richiede un costo di $\approx n^2$.

Teorema (Cayley-Hamilton)

Sia A una matrice avente:

$$\lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n = 0$$

come equazione caratteristica. Allora la matrice definita da:

$$A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n I \text{ è la matrice zero.}$$

◦

Tale teorema può essere messo sotto la forma:
 “ogni matrice soddisfa la propria equazione caratteristica”.
 Su di esso si basa il seguente metodo.

Metodo di Krylov

E' un metodo elementare per trovare il polinomio caratteristico di una matrice in forma di Hessenberg. Supponiamo che tutti gli elementi subdiagonali della matrice H siano non zero e sia:

$$w_0 = e_1 = (1, 0, \dots, 0)^T$$

A partire da esso, definiamo $\{w_k\}_{k=1}^n$ con l'algoritmo:

$$w_{k+1} = Hw_k \quad k = 0, 1, \dots, n-1$$

Si può far vedere che la $(k+1)$ -esima componente di w_k è: $h_{21}h_{32}\dots h_{k+1,k}$ per $1 \leq k \leq n-1$ e zero per $r > k+1$.

Pertanto la matrice: $W = [w_0, w_1, \dots, w_{n-1}]$ sarà triangolare superiore con elementi diagonali non nulli. Il sistema:

$$Wx = -w_n$$

ha quindi un'unica soluzione la cui soluzione x dà i coefficienti del polinomio caratteristico. Infatti esso può essere scritto come:

$$a_n w_0 + a_{n-1} w_1 + \dots + a_1 w_{n-1} = -w_n$$

Inoltre:

$$w_{k+1} = Hw_k = H^2 w_{k-1} = \dots = H^{k+1} w_0 = H^{k+1} e_1$$

Calcolo del polinomio caratteristico per matrici non simmetriche

Nel caso di matrici non simmetriche le trasformazioni per similarità danno luogo a matrici di Hessemberg. In tal caso non si calcola il polinomio caratteristico ma una sua riformulazione che ha le stesse radici. Partiamo da una matrice B in forma di Hessemberg

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & \cdots & b_{1n} \\ a_1 & b_{22} & \ddots & \cdots & b_{2n} \\ 0 & a_2 & \ddots & \ddots & \cdots \\ 0 & 0 & \ddots & \ddots & \cdots \\ 0 & 0 & 0 & a_{n-1} & b_{nn} \end{pmatrix}$$

e siano gli $a_i \neq 0$ perché in tal caso ci ricondurremmo a due matrici di Hessemberg. Il problema sarebbe quindi suddiviso in due blocchi B_{11} , B_{22} e si avrebbe

$$\sigma(B) = \sigma(B_{11}) \cup \sigma(B_{22})$$

Calcoliamo quindi, per $a_i \neq 0$, il

$$\det(B - \lambda I) =$$

$$\det \begin{pmatrix} b_{11} - \lambda & b_{12} & \dots & \dots & b_{1n} \\ a_1 & b_{22} - \lambda & \ddots & \dots & b_{2n} \\ 0 & a_2 & \ddots & \ddots & \dots \\ 0 & 0 & \ddots & \ddots & \dots \\ 0 & 0 & 0 & a_{n-1} & b_{nn} - \lambda \end{pmatrix}$$

Consideriamo il sistema lineare

$$(B - \lambda I)\rho = K(\lambda)e_1$$

con $\rho \in \Re^n$, $K(\lambda) \in \mathbb{P}_n(\lambda)$ e con $\lambda \in \Re$ fissato ovvero:

$$\begin{pmatrix} b_{11} - \lambda & b_{12} & \dots & \dots & b_{1n} \\ a_1 & b_{22} - \lambda & \ddots & \dots & b_{2n} \\ 0 & a_2 & \ddots & \ddots & \dots \\ 0 & 0 & \ddots & \ddots & \dots \\ 0 & 0 & 0 & a_{n-1} & b_{nn} - \lambda \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \dots \\ \dots \\ \rho_n \end{pmatrix} = \begin{pmatrix} K(\lambda) \\ 0 \\ \dots \\ \dots \\ 0 \end{pmatrix}$$

dove, essendoci $n+1$ incognite $\rho_1, \rho_2, \dots, \rho_n, K(\lambda)$ ed n equazioni, possiamo porre $\rho_n = 1$. (i=n-1,2)

$$\left\{ \begin{array}{l} \rho_n = 1 \\ a_{n-1}\rho_{n-1} + (b_{nn} - \lambda)\rho_n = 0 \\ \vdots \\ a_{i-1}\rho_{i-1} + (b_{ii} - \lambda)\rho_i + \sum_{j=i+1}^n b_{ij} \rho_j = 0 \\ \vdots \\ (b_{11} - \lambda)\rho_1 + \sum_{j=2}^n b_{1j} \rho_j = K(\lambda) \end{array} \right.$$

$i=n-1,2$

da cui

$$\begin{array}{l} \rho_n = 1 \\ \vdots \\ \rho_{i-1} = -\frac{1}{a_{i-1}} \left[(b_{ii} - \lambda)\rho_i + \sum_{j=i+1}^n b_{ij} \rho_j \right] \\ \vdots \\ K(\lambda) = (b_{11} - \lambda)\rho_1 + \sum_{j=2}^n b_{1j} \rho_j \end{array}$$

$i=n,2$ dove la sommatoria con $j=n+1$ deve essere posta uguale a zero.

Quindi mostriamo che $K(\lambda)$ e $\det(B - \lambda I)$ hanno le stesse radici.

Sia c_i l' i -esima colonna di $(B - \lambda I)$. Allora:

$$\det(B - \lambda I) = \det[c_1, c_2, \dots, c_n]$$

Da

$$(B - \lambda I)\rho = K(\lambda)e_1$$

segue che

$$\rho_1 c_1 + \rho_2 c_2 + \dots + \rho_n c_n = K(\lambda)e_1$$

$$\rho_n c_n = K(\lambda)e_1 - \sum_{j=1}^{n-1} \rho_j c_j$$

$$\det[c_1, c_2, \dots, c_n] = \det \left[c_1, \dots, c_{n-1}, K(\lambda)e_1 - \sum_{j=1}^{n-1} \rho_j c_j \right] = \det[c_1, \dots, c_{n-1}, K(\lambda)e_1]$$

poiché $\sum_{j=1}^{n-1} \rho_j c_j$ è una combinazione lineare delle altre colonne.

$$\det \begin{pmatrix} (b_{11} - \lambda) & b_{12} & \dots & K(\lambda) \\ a_1 & (b_{22} - \lambda) & \ddots & 0 \\ 0 & a_2 & \ddots & \vdots \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & a_{n-1} & 0 \end{pmatrix} =$$

$$(-1)^{n+1} K(\lambda) \det \begin{pmatrix} a_1 & (b_{22} - \lambda) & \dots & b_{2,n-1} \\ 0 & a_2 & \ddots & \vdots \\ 0 & 0 & \dots & \vdots \\ \vdots & \vdots & \ddots & (b_{n-1,n-1} - \lambda) \\ 0 & 0 & \dots & a_{n-1} \end{pmatrix}$$

$$= (-1)^{n+1} K(\lambda) \prod_{i=1}^{n-1} a_i = 0$$

Per cui:

$$\det(B - \lambda I) \sim K(\lambda) \prod_{i=1}^{n-1} a_i \rightarrow K(\lambda) = 0.$$

Matrici simmetriche

Accenniamo infine al calcolo del polinomio caratteristico, ottenuto in maniera ricorsiva, nel caso in cui la matrice di partenza sia *simmetrica* e ad essa venga applicato il metodo di Householder. In tal caso, la trasformazione darà luogo ad una matrice tridiagonale. Indicata con B la matrice così trasformata, con $a_i, i=1, \dots, n$ i suoi elementi diagonali e con $c_i, i=1, \dots, n-1$ gli elementi delle due subdiagonali, si ha:

$$\det(\lambda I - B) = \det \begin{vmatrix} \lambda - a_1 & -c_1 & & & 0 \\ -c_1 & \lambda - a_2 & \dots & & \\ & & \dots & \dots & \\ & & & \dots & -c_{n-1} \\ 0 & & & -c_{n-1} & \lambda - a_n \end{vmatrix}$$

Sia $p_i(\lambda)$ il polinomio caratteristico del minore formato con le prime i righe ed i colonne.

$$p_0(\lambda) \equiv 1, \quad p_1(\lambda) = \lambda - a_1,$$

$$p_2(\lambda) = (\lambda - a_1)(\lambda - a_2) - c_1^2 = p_1(\lambda)(\lambda - a_2) - c_1^2 p_0(\lambda),$$

...

$$p_i(\lambda) = (\lambda - a_i)p_{i-1}(\lambda) - c_{i-1}^2 p_{i-2}(\lambda)$$

Nonostante sia abbastanza semplice trovare il polinomio caratteristico di una matrice con queste tecniche, in realtà la determinazione dei coefficienti di tale polinomio è sempre un'operazione molto delicata poiché anche un piccolo errore di arrotondamento può portare ad un errore molto grande.