

## CAPITOLO III

## PROBLEMI DI AUTOVALORI

Data una matrice  $A \in \mathcal{L}(C^n, C^n)$  la ricerca degli autovalori ed autovettori di  $A$  è, nel suo complesso, un problema non lineare per il quale si può sperare di avere, al più, delle soluzioni approssimate. Non ci si illuda di poter utilizzare le forme canoniche, dalle quali si potrebbe ottenere direttamente gli autovalori, inquanto l'esistenza di tali forme è dimostrata in maniera astratta oppure attraverso procedimenti iterativi.

Anche l'approccio diretto attraverso l'equazione caratteristica

$$\det(A - \lambda I) = 0$$

non è praticabile sul piano numerico in quanto, in generale, non si dispone dei coefficienti del polinomio caratteristico se non attraverso la manipolazione simbolica dei coefficienti che diventa molto complessa al crescere di  $n$ .

Il problema della ricerca degli autovalori potrà quindi essere risolto solo in maniera approssimata. Cominciamo con l'osservare che, se fosse nota una autosoluzione  $x$ , allora sarebbe molto facile trovare l'autovalore corrispondente  $\lambda$ , infatti:

$$\begin{aligned} Ax &= \lambda x \\ x^H Ax &= \lambda x^H x \\ \lambda &= \frac{x^H Ax}{x^H x} \end{aligned}$$

o equivalentemente, per  $y = \frac{x}{\|x\|_2}$ :

$$\lambda = y^H A y \quad (\text{dove } \|x\|_2 = 1).$$

Per un vettore arbitrario  $z$ , il numero  $\frac{z^H A z}{z^H z}$  è detto **quoziente di Rayleigh**. Il quoziente di Rayleigh di un autovettore è dunque l'autovalore associato. Se invece dispongo di una successione  $z_k$  di vettori convergente ad  $x$ , allora la successione di quozienti di Rayleigh

$$\lambda_k = \frac{z_k^H A z_k}{z_k^H z_k}$$

converge a  $\lambda$ . Questa asserzione è alla base del seguente metodo iterativo.

### Metodo delle potenze.

Supponiamo che la matrice  $A$  possieda un sistema completo di autovettori  $u_1, u_2, \dots, u_n$ , con i rispettivi autovalori  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Supponiamo inoltre che gli autovalori di modulo massimo, qualora siano più di uno, coincidano tra loro. Cioè:

$$|\lambda_1| = |\lambda_2| = \dots = |\lambda_r| > |\lambda_{r+1}| \geq \dots \geq |\lambda_n| \quad \text{con } \lambda_1 = \lambda_2 = \dots = \lambda_r$$

Preso ad arbitrio un vettore  $z_0$ , che si può sempre esprimere come

$$z_0 = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_r u_r + \alpha_{r+1} u_{r+1} + \dots + \alpha_n u_n,$$

supponiamo che il vettore  $v := \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_r u_r$ , cioè la sua componente nell'autospazio di  $\lambda_1$ , sia non nullo.

Applichiamo ora successivamente la matrice  $A$  in modo da ottenere la successione  $z_k = A^k z_0$   $k=1, 2, \dots$

Si avrà:

$$\begin{aligned} z_k &= A^k (\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_r u_r + \alpha_{r+1} u_{r+1} + \dots + \alpha_n u_n) = A^k (v + \alpha_{r+1} u_{r+1} + \dots + \alpha_n u_n) \\ &= A^k v + \alpha_{r+1} A^k u_{r+1} + \dots + \alpha_n A^k u_n = \lambda_1^k v + \alpha_{r+1} \lambda_{r+1}^k u_{r+1} + \dots + \alpha_n \lambda_n^k u_n \\ &= \lambda_1^k \left( v + \alpha_{r+1} \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^k u_{r+1} + \dots + \alpha_n \left( \frac{\lambda_n}{\lambda_1} \right)^k u_n \right) \end{aligned}$$

e moltiplicando per  $z_k^H$

$$z_k^H z_k = \lambda_1^k \left( z_k^H v + \alpha_{r+1} \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^k z_k^H u_{r+1} + \dots + \alpha_n \left( \frac{\lambda_n}{\lambda_1} \right)^k z_k^H u_n \right)$$

Analogamente,

$$\begin{aligned} A z_k &= z_{k+1} = \lambda_1^{k+1} v + \alpha_{r+1} \lambda_{r+1}^{k+1} u_{r+1} + \dots + \alpha_n \lambda_n^{k+1} u_n \\ &= \lambda_1^{k+1} \left( v + \alpha_{r+1} \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^{k+1} u_{r+1} + \dots + \alpha_n \left( \frac{\lambda_n}{\lambda_1} \right)^{k+1} u_n \right) \end{aligned}$$

e moltiplicando ancora per  $z_k^H$

$$z_k^H z_{k+1} = \lambda_1^k \left( z_k^H v + \alpha_{r+1} \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^{k+1} z_k^H u_{r+1} + \dots + \alpha_n \left( \frac{\lambda_n}{\lambda_1} \right)^{k+1} z_k^H u_n \right)$$

Si ottiene quindi il seguente quoziente di Rayleigh:

$$\frac{z_k^H A z_k}{z_k^H z_k} = \lambda_1 \frac{z_k^H v + \alpha_{r+1} \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^{k+1} z_k^H u_{r+1} + \dots + \alpha_n \left( \frac{\lambda_n}{\lambda_1} \right)^{k+1} z_k^H u_n}{z_k^H v + \alpha_{r+1} \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^k z_k^H u_{r+1} + \dots + \alpha_n \left( \frac{\lambda_n}{\lambda_1} \right)^k z_k^H u_n}$$

che converge a  $\lambda_1$  con la stessa rapidità con cui il termine  $\left( \frac{\lambda_{r+1}}{\lambda_1} \right)^k$  tende a zero.

L'algoritmo per l'esecuzione del metodo delle potenze è:

- 1 assegna  $z$
- 2 poni  $y = \frac{z}{\|z\|}$
- 3 poni  $z = Ay$
- 4 poni  $\lambda = y^H z$  e visualizza  $\lambda$
- 5 ritorna all'istruzione 2

Il programma visualizza una successione di valori convergenti a  $\lambda_1$

Nel caso che la matrice  $A$  sia hermitiana, non solo è garantita l'ipotesi sulla completezza del sistema di autovettori, ma tali autovettori risultano anche ortogonali. Ciò significa che, ottenuto:

$$z_k = \lambda_1^k \left( v + \alpha_{r+1} \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^k u_{r+1} + \dots + \alpha_n \left( \frac{\lambda_n}{\lambda_1} \right)^k u_n \right),$$

il prodotto scalare  $z_k^H z_k$  sarà dato da:

$$z_k^H z_k = \lambda_1^{2k} \left( v^H v + \alpha_{r+1}^2 \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^{2k} + \dots + \alpha_n^2 \left( \frac{\lambda_n}{\lambda_1} \right)^{2k} \right)$$

ed analogamente

$$z_k^H z_{k+1} = \lambda_1^{2k+1} \left( v^H v + \alpha_{r+1}^2 \left( \frac{\lambda_{r+1}}{\lambda_1} \right)^{2k+1} + \dots + \alpha_n^2 \left( \frac{\lambda_n}{\lambda_1} \right)^{2k+1} \right)$$

cosicché il quoziente di Rayleigh diventa in questo caso

$$\frac{z_k^H A z_k}{z_k^H z_k} = \lambda_1 \frac{v^H v + \alpha_{r+1}^2 \left(\frac{\lambda_{r+1}}{\lambda_1}\right)^{2k+1} + \dots + \alpha_n^2 \left(\frac{\lambda_n}{\lambda_1}\right)^{2k+1}}{v^H v + \alpha_{r+1}^2 \left(\frac{\lambda_{r+1}}{\lambda_1}\right)^{2k} + \dots + \alpha_n^2 \left(\frac{\lambda_n}{\lambda_1}\right)^{2k}}$$

la cui velocità di convergenza verso  $\lambda_1$  è doppia rispetto al caso generale.

Ritornando alle ipotesi fatte per la convergenza del metodo delle potenze, osserviamo che, quando sono verificate, esse garantiscono la convergenza della successione di quozienti di Rayleigh sempre all'autovalore di modulo massimo, qualunque sia il vettore iniziale  $z_0$  con la sola ipotesi che le sue componenti nelle direzioni di  $u_1, \dots, u_r$  non siano tutte nulle. D'altra parte è facile ricavare, ripercorrendo gli stessi passaggi, che nel caso si abbia

$$z_0 = \alpha_{r+1} u_{r+1} + \alpha_{r+2} u_{r+2} + \dots + \alpha_n u_n,$$

allora  $z_k = \alpha_{r+1} \lambda_{r+1}^k u_{r+1} + \dots + \alpha_n \lambda_n^k u_n$  ed il quoziente di Rayleigh assume la forma

$$\frac{z_k^H A z_k}{z_k^H z_k} = \lambda_{r+1} \frac{\alpha_{r+1} z_k^H u_{r+1} + \alpha_{r+2} \left(\frac{\lambda_{r+2}}{\lambda_{r+1}}\right)^{k+1} z_k^H u_{r+2} + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_{r+1}}\right)^{k+1} z_k^H u_n}{\alpha_{r+1} z_k^H u_{r+1} + \alpha_{r+2} \left(\frac{\lambda_{r+2}}{\lambda_{r+1}}\right)^k z_k^H u_{r+2} + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_{r+1}}\right)^k z_k^H u_n}$$

Se facciamo l'ipotesi  $|\lambda_{r+1}| > |\lambda_{r+2}| \geq \dots \geq |\lambda_n|$  (si osservi che non viene fatta alcuna ipotesi sui primi  $r$  autovalori) allora il metodo converge verso l'autovalore  $\lambda_{r+1}$ . Sfortunatamente, a parte le difficoltà di trovare un  $z_0$  che soddisfi la condizione richiesta, questo metodo risulta non praticabile a causa degli errori di troncamento dei calcoli. Infatti nel calcolare i vettori  $z_k$ , o meglio le sue componenti, si introdurranno inevitabilmente degli errori per i quali i vettori effettivamente calcolati avranno delle componenti non nulle nelle direzioni di  $u_1, \dots, u_r$ . Tali componenti, anche se inizialmente di piccolo modulo, riconducono il problema nei termini formulati all'inizio per cui la successione convergerà ancora all'autovalore di modulo massimo. In pratica si osserverà inizialmente una apparente convergenza verso l'autovalore  $\lambda_{r+1}$  e successivamente un allontanamento fino alla definitiva convergenza verso  $\lambda_1$ .

Poiché i reciproci degli autovalori di  $A$  sono autovalori di  $A^{-1}$ , nel caso di matrici non singolari il metodo delle potenze applicato alla matrice  $A^{-1}$  fornisce il reciproco

dell'autovalore minimo di A. Per avere la convergenza bisogna ovviamente che il modulo dell'autovalore minimo sia strettamente minore di tutti gli altri:

$$|\lambda_1| \geq \dots \geq |\lambda_{n-1}| > |\lambda_n|$$

Questa osservazione suggerisce il seguente metodo per il calcolo degli altri autovalori intermedi.

Metodo delle potenze inverse.

Poichè  $\lambda \in S(A) \Rightarrow \lambda - \alpha \in S(A - \alpha I)$ , allora per ogni autovalore  $\lambda_i \in S(A)$  si ottiene un autovalore di  $(A - \alpha I)^{-1}$  dato da

$$\sigma_i = \frac{1}{\lambda_i - \alpha} \in S(A - \alpha I)^{-1}.$$

Nota che sia un  $\sigma_i$ , si trova quindi un autovalore di A :

$$\lambda_i = \alpha + \frac{1}{\sigma_i}$$

In particolare per quello di modulo massimo:

$$\max_i |\sigma_i| = \max_i \frac{1}{|\lambda_i - \alpha|} = \frac{1}{\min_i |\lambda_i - \alpha|}$$

la formula precedente fornisce l'autovalore di A più vicino ad  $\alpha$ .

Applicando quindi il metodo delle potenze alla matrice  $(A - \alpha I)^{-1}$  per vari valori di  $\alpha$ , si possono calcolare tutti gli autovalori di A ammesso che siano verificate le ipotesi sulla completezza del sistema di autovettori per le matrici  $(A - \alpha I)^{-1}$ . Nel caso già considerato di matrici hermitiane, l'ipotesi è verificata per ogni  $\alpha$ .

Vale la pena di osservare che nel caso di una matrice A reale con una coppia di autovalori complessi, per ogni scelta di  $\alpha$  reale i due autovalori coniugati hanno la stessa distanza da  $\alpha$  e quindi non ci può essere un solo autovalore di modulo massimo per  $(A - \alpha I)^{-1}$ . La ricerca di tali autovalori va fatta con un  $\alpha$  complesso.

Il metodo delle potenze inverse risulta comunque macchinoso e richiede un uso interattivo del computer per cui non appare di grande utilità se non per casi molto speciali ed in particolare per il calcolo dell'autovalore minimo.

Metodo di deflazione.

Sappiamo che gli autovalori di una matrice di dimensione  $n$  sono le radici di un polinomio  $p(x)$  di grado  $n$ . Il teorema di Ruffini asserisce che, nota una radice  $\alpha$ , il polinomio  $p(x)$  è divisibile per  $(x-\alpha)$  e le rimanenti radici sono le radici del quoziente ottenuto dividendo  $p(x)$  per  $(x-\alpha)$  e così via riducendo successivamente il grado del polinomio. D'altra parte vi è uno stretto legame tra i polinomi di grado  $n$  e gli autovalori di una matrice  $n \times n$ . Infatti ogni matrice  $A$  dà luogo all'equazione caratteristica le cui radici sono gli autovalori di  $A$ , e viceversa ogni polinomio  $p_n$  è polinomio caratteristico di una matrice i cui autovalori sono quindi le radici di  $p_n$ . Tale matrice non è ovviamente unica perchè, come abbiamo già visto, ogni matrice ad essa simile ha lo stesso polinomio caratteristico. E' immediato verificare che il polinomio:

$$p(\lambda) = \lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_1\lambda - a_0$$

è il polinomio caratteristico associato della matrice:

$$B = (-1)^n \begin{vmatrix} a_{n-1} & a_{n-2} & \cdot & \cdot & \cdot & a_1 & a_0 \\ 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & 0 & & & & \cdot \\ \cdot & 0 & 1 & \cdot & & & \cdot \\ \cdot & & 0 & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 & 1 & 0 \end{vmatrix}$$

E' ragionevole quindi chiedersi se, noto un autovalore e la sua corrispondente autosoluzione, si riesce a ricavare, in analogia al teorema di Ruffini, una matrice di dimensione  $n-1$  il cui spettro sia dato dai rimanenti autovalori di  $A$ . In tal caso la ricerca di un secondo autovalore sarà meno costosa in quanto riferita ad una matrice di dimensione inferiore, e così via per gli autovalori successivi.

La risposta al quesito è positiva e la tecnica è nota come **deflazione** della matrice.

Siano dunque  $\lambda$  ed  $x$  tali che  $Ax = \lambda x$  e supponiamo di conoscere una matrice  $P$ , non singolare, tale che  $Px = ke_1$  con  $k \in \mathbb{C}$ . Allora:

$$\begin{aligned} Ax &= \lambda x \\ AP^{-1}Px &= \lambda x \\ PAP^{-1}Px &= \lambda Px \\ PAP^{-1}e_1 &= \lambda e_1 \end{aligned}$$

In virtù dell'ultima relazione, la matrice  $PAP^{-1}$ , simile ad  $A$ , è del tipo

$$PAP^{-1} = \begin{vmatrix} \lambda & b_2 & \dots & b_n \\ 0 & \boxed{B} & & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \end{vmatrix}$$

i cui autovalori sono, oltre  $\lambda$ , gli autovalori di  $B$ .

Il problema è ora quello di trovare una matrice  $P$  tale che  $Px = ke_1$ . A tale scopo introduciamo una classe di matrici che risulteranno molto utili anche in altre circostanze.

Dato un vettore  $w$  tale che  $w^H w = 1$  (cioè di norma 2 unitaria), definiamo la seguente matrice detta **matrice elementare di Householder**:

$$P := I - 2ww^H$$

e osserviamo che essa è hermitiana ed unitaria (e quindi involutoria):

$$P^H = I - 2(ww^H)^H = I - 2ww^H = P$$

$$P^H P = P^2 = (I - 2ww^H)(I - 2ww^H) = I - 4ww^H + 4ww^H ww^H =$$

$$= I - 4ww^H + 4ww^H = I$$

Osserviamo che se esiste un  $w$  per il quale  $Px = ke_1$ , allora necessariamente  $|k| = \|x\|_2$ .

Infatti  $(Px)^H Px = x^H x = \|x\|_2^2$  ed anche  $(Px)^H Px = \bar{k} e_1^T k e_1 = |k|^2$ .

Supponiamo ora che per un arbitrario vettore  $x = (x_1, x_2, \dots, x_n)^T \neq 0$  sia  $x_1 \neq 0$ , e definiamo il

vettore  $w = \frac{x - ke_1}{\|x - ke_1\|_2}$ , con  $k = \frac{x_1 \|x\|_2}{|x_1|}$ . Osserviamo che si ha:

$$\|x - ke_1\|_2^2 = (x - ke_1)^H (x - ke_1) = (x_1 - k, x_2, \dots, x_n)^H (x_1 - k, x_2, \dots, x_n)$$

$$= \bar{x}_1 x_1 - \bar{x}_1 k - \bar{k} x_1 + \bar{k} k + \bar{x}_2 x_2 + \dots + \bar{x}_n x_n$$

$$=|x_1|^2 + |x_2|^2 + \dots + |x_n|^2 - 2|x_1| \|x\|_2 + \|x\|_2^2$$

$$=2\|x\|_2^2 - 2|x_1| \|x\|_2 = 2\|x\|_2(\|x\|_2 - |x_1|)$$

Quindi:

$$P x = (I - 2 w w^H) x = x - 2 w w^H x$$

$$= x - 2 \frac{(x - k e_1)}{\|x - k e_1\|_2} \frac{(x - k e_1)^H}{\|x - k e_1\|_2} x = x - \frac{(x - k e_1)(x - k e_1)^H x}{\|x\|_2 (\|x\|_2 - |x_1|)}$$

$$= x - \frac{(x - k e_1)(x^H x - \bar{k} x_1)}{\|x\|_2 (\|x\|_2 - |x_1|)}$$

$$= x - \frac{(x - k e_1)\|x\|_2 (\|x\|_2 - |x_1|)}{\|x\|_2 (\|x\|_2 - |x_1|)} = k e_1$$

Nel caso che sia invece  $x_1=0$ , si pone  $k=\|x\|_2$ , e si ritrova ancora che  $Px=ke_1$ . Possiamo riassumere il risultato ottenuto nel seguente teorema:

**Teorema 3.1.** *Dato un vettore  $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^n$ ,  $x \neq 0$ , il vettore normalizzato  $w = \frac{x - k e_1}{\|x - k e_1\|_2}$ , con  $k = \frac{x_1 \|x\|_2}{|x_1|}$  se  $x_1 \neq 0$  e  $k = \|x\|_2$  se  $x_1 = 0$ , definisce una matrice  $P = I - 2 w w^H$  tale che  $Px = k e_1$ .*

Si osservi che se  $x$  è definito a meno di una costante moltiplicativa, come accade per un autovettore, allora si può sempre porre  $\|x\|_2 = 1$ . Si osservi infine che se  $x$  è già parallelo ad  $e_1$ , allora  $w=0$  e  $P=I$ .

Abbiamo quindi individuato una matrice  $P$ , che incidentalmente risulta unitaria, tale che da  $PAP^{-1}$  si può estrarre la matrice deflazionata  $B$ .

*Osservazione.* Il processo di deflazione consente di dare una dimostrazione diretta del teorema di Schur. Infatti detto  $\mu$  un altro autovalore di  $PAP^{-1}$ , e quindi di  $A$ , la cui autosoluzione sia  $y = (y_1, y_2, \dots, y_n)^T$ , si ha:

$$PAP^{-1}y = \mu y$$

$$By' = \mu y' \quad \text{dove } y' = (y_2, \dots, y_n)^T$$

Allora per la matrice deflazionata  $B$  ( $(n-1) \times (n-1)$ ) troviamo una trasformazione  $Q$  ( $(n-1) \times (n-1)$ ) di Householder tale che

$$QBQ^1 = \begin{vmatrix} \mu & c_2 & \dots & c_{n-1} \\ 0 & \boxed{D} & & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \end{vmatrix}$$

Allora si verifica facilmente che la matrice  $R$  ( $n \times n$ ):

$$R = \begin{vmatrix} 1 & & & \\ & \boxed{Q} & & \\ & & & \end{vmatrix}$$

che è ancora unitaria, realizza la seguente similitudine

$$R(PAP^{-1})R^{-1} = \begin{vmatrix} \lambda & \cdot & \dots & \dots & \cdot \\ 0 & \mu & \dots & \dots & \cdot \\ \cdot & 0 & \boxed{D} & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ 0 & 0 & & & \end{vmatrix}$$

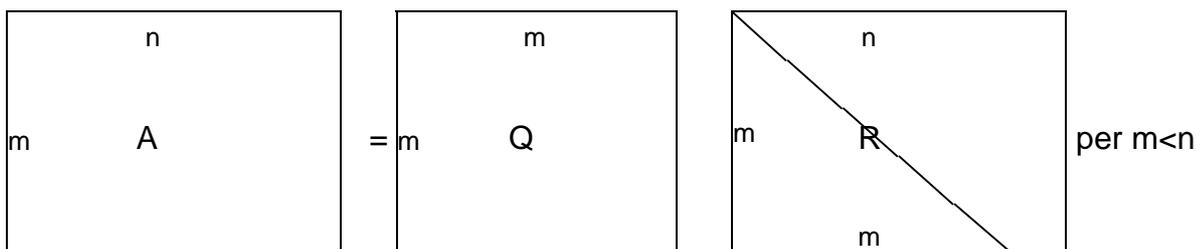
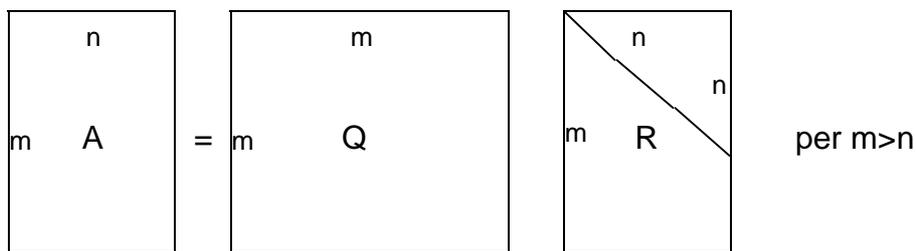
Continuando si perviene ad una matrice di forma triangolare superiore simile alla matrice A la cui matrice di trasformazione è data dal prodotto di matrici unitarie e quindi a sua volta unitaria.

La fattorizzazione QR e l'algoritmo QR di Francis-Kublanowskaya (1961).

La fattorizzazione che ora verrà descritta vale per una matrice di dimensione  $m \times n$  qualunque. Sia data dunque una matrice  $A \in \mathcal{L}(C^n, C^m)$  le cui colonne verranno indicate con  $a_1, a_2, \dots, a_n$ . Generalizziamo la definizione di matrice triangolare superiore al caso  $m \neq n$  chiamando ancora elementi diagonali di A gli elementi  $a_{11}, a_{22}, \dots, a_{rr}$ ,  $r = \min(m, n)$ , e dicendo che A è triangolare superiore se tutti gli elementi sotto la diagonale sono nulli. Esse avranno la seguente forma:

$$\left\| \begin{array}{cccc} a_{11} & \cdot & \cdot & \cdot & a_{1n} \\ & A_{22} & & & \cdot \\ & & \cdot & & \cdot \\ & & & \cdot & \cdot \\ & & & & a_{nn} \end{array} \right\| \text{ per } m > n, \quad \left\| \begin{array}{ccccccc} a_{11} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a_{1n} \\ & \cdot & & & & & & \cdot \\ & & \cdot & & & & & \cdot \\ & & & \cdot & & & & \cdot \\ & & & & \cdot & & & \cdot \\ & & & & & a_{mm} & \cdot & a_{mn} \end{array} \right\| \text{ per } m < n.$$

**Teorema 3.2.** Per ogni matrice  $A \in \mathcal{L}(C^n, C^m)$  esiste una matrice unitaria  $Q \in \mathcal{L}(C^m, C^m)$  ed una matrice triangolare superiore  $R \in \mathcal{L}(C^n, C^m)$  tali che  $A=QR$ .



Dim. Sia  $a_1$  la prima colonna di  $A$ . In base al teorema 3.1 la matrice elementare  $P_1 = I - 2ww^H \in \mathcal{L}(C^m, C^m)$  con

$$w = \frac{a_1 - ke_1}{\|a_1 - ke_1\|_2}, \quad \text{e } k = \frac{a_{11}\|a_1\|_2}{|a_{11}|}$$

è tale che  $P_1 a_1 = ke_1$ .

La matrice  $P_1 A$  è quindi del tipo:

$$P_1 A = \left\| \begin{array}{cccc} k_1 & \cdot & \cdot & \cdot & \cdot \\ 0 & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ 0 \end{array} & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \\ \cdot & & T_1 & & \\ \cdot & & & & \\ 0 & & & & \end{array} \right\|$$

Successivamente per la seconda colonna di  $P_1 A$ , si considera la matrice elementare di Householder  $P'_2$  relativa a  $T_1$ , entrambe in dimensione  $m-1$ , e con essa si costruisce la trasformazione  $P_2$

$$P_2 = \left\| \begin{array}{cccc} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ 0 \end{array} & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} & \\ \cdot & & P'_2 & & \\ \cdot & & & & \\ 0 & & & & \end{array} \right\|$$

che risulta ancora unitaria. Applicando quindi  $P_2$  a  $P_1 A$  si ottiene , ,

$$P_2 P_1 A = \left\| \begin{array}{cccc} k_1 & \cdot & \cdot & \cdot & \cdot \\ 0 & k_2 & \cdot & \cdot & \cdot \\ \cdot & 0 & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ 0 \end{array} & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} & \begin{array}{|c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \\ \cdot & \cdot & & T_2 & \\ 0 & 0 & & & \end{array} \right\|$$

e così via per tutte le  $n$  colonne. La matrice  $Q^{-1}=P_n P_{n-1} \dots P_1$  è ancora unitaria e tale che  $R:=Q^{-1}A$  è triangolare superiore. Si ha quindi  $A=QR$ .

Se la matrice  $T_1$ , o più in generale la sottomatrice da deflazionare, ha la prima colonna nulla, allora essa non ha bisogno di essere trasformata e si passa alla colonna successiva. In questo caso la matrice triangolare  $R$  avrà uno zero sulla diagonale.

Nel caso  $m=n$  la matrice  $R$  è singolare se e solo se tale è  $A$ . Infatti  $\det(A)=\det(Q)\det(R)$  dove  $\det(Q)=\pm 1$ . Quindi, per matrici quadrate, la presenza di zeri sulla diagonale di  $R$  coincide con la singolarità della matrice  $A$ .

L'algoritmo QR per il calcolo degli autovalori di  $A$  consiste nel calcolo della seguente successione di trasformazioni. Data  $A$ , che ora indicheremo con  $A_1$ , sia  $A_1=Q_1 R_1$  la sua fattorizzazione QR. Definiamo  $A_2:=R_1 Q_1$  e così via: , ,

$$A_1 := A \quad \text{e fattorizziamo } A_1 = Q_1 R_1$$

$$A_2 := R_1 Q_1 \quad \text{e fattorizziamo } A_2 = Q_2 R_2 \quad , ,$$

.....

$$A_{i+1} := R_i Q_i \quad \text{e fattorizziamo } A_{i+1} = Q_{i+1} R_{i+1} \quad , ,$$

....

Osserviamo che tutte le matrici  $A_i$  sono simili ad  $A$ . Ciò è vero per  $i=2$  :

$$Q_1 A_2 Q_1^H = Q_1 (R_1 Q_1) Q_1^H = Q_1 R_1 = A_1$$

ed inoltre ogni  $A_i$  è simile alla precedente  $A_{i-1}$  :

$$Q_{i-1} A_i Q_{i-1}^H = Q_{i-1} (R_{i-1} Q_{i-1}) Q_{i-1}^H = Q_{i-1} R_{i-1} = A_{i-1}$$

Per la successione  $A_i$  vale il seguente teorema:

**Teorema 3.3.** *Data una matrice  $A$ , la successione  $A_i$  definita attraverso l'algoritmo QR si comporta nel seguente modo a seconda che sia o non sia verificata la seguente ipotesi:*

a)  $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$  .

Se vale a), la successione converge ad una matrice triangolare superiore sulla cui diagonale si trovano gli autovalori di A. In altre parole la parte sottostante la diagonale tende a zero. Negli altri casi tende a zero la parte sottostante una diagonale a blocchi i cui blocchi sono di dimensione 1 e di dimensione 2. Gli autovalori dei blocchi diagonali superstiti convergono agli autovalori di A.

La forma limite della successione  $A_i$  è quindi, ad esempio, del tipo:

$$\begin{pmatrix} x & & & & & & & \\ & x & x & & & & & \\ & & x & x & & & & \\ & & & x & x & & & \\ & & & & x & x & & \\ & & & & & x & x & \\ & & & & & & x & x \\ & & & & & & & x & x \\ & & & & & & & & x & x \end{pmatrix}$$

Si capisce che il limite non può essere sempre una matrice triangolare superiore perchè, nel caso di matrici reali, le trasformazioni di Householder e quindi le fattorizzazioni QR nonchè i prodotti RQ sono ancora reali ad ogni passo della successione. Non si potrebbero quindi trovare gli eventuali autovalori complessi di A. In presenza di coppie di autovalori complessi coniugati non è infatti verificata l'ipotesi a). In tal caso compaiono dei blocchi diagonali a coefficienti reali di dimensione 2, i cui autovalori complessi coniugati approssimano quelli di A.

Il costo computazionale del metodo QR è notevolmente alto ( $\frac{2}{3}n^3$  per ogni fattorizzazione QR e  $\frac{n^3}{2}$  per ogni prodotto RQ, complessivamente  $\frac{7}{6}n^3$  per ogni iterazione) ed una tecnica standard per ridurlo consiste nel trasformare preliminarmente, per similitudine, la matrice A in una **matrice di Hessenberg**, cioè del tipo:

$$\begin{pmatrix} x & x & x & x & x & x & x \\ x & x & x & x & x & x & x \\ & x & x & x & x & x & x \\ & & x & x & x & x & x \\ & & & x & x & x & x \\ & & & & x & x & x \\ & & & & & x & x \\ & & & & & & x & x \end{pmatrix}$$

Ciò può essere fatto in n-2 passi utilizzando ancora le trasformazioni di Householder nel seguente modo. Individuata ad ogni passo la trasformazione  $P_i$ , come nella fattorizzazione

QR, con essa si trasforma per similitudine la matrice del passo precedente. Si vede che al passo  $i$ -esimo la colonna  $i$ -esima viene trasformata in modo da avere coefficienti nulli al disotto della sottodiagonale.

Se la matrice  $A$  è hermitiana, tale rimane la sua trasformata in forma di Hessenberg che risulta pertanto tridiagonale.

Dopo aver trasformato  $A$  in forma di Hessenberg o in forma tridiagonale, si applicano le trasformazioni dell'algorithmo QR che conservano, come è facile vedere, tale forma. In questo modo si risparmiano molti calcoli perchè molti dei coefficienti delle matrici coinvolte nel procedimento sono nulli e quindi non occorre calcolarli. La riduzione a forma di Hessenberg richiede  $\frac{5n^3}{3}$  ( $\frac{2n^3}{3}$  nel caso tridiagonale) operazioni. Tale costo iniziale viene ampiamente compensato dal fatto che, successivamente, l'applicazione del metodo QR richiede ad ogni iterazione un numero di operazioni che cresce quadraticamente con  $n$  (linearmente con  $n$  nel caso tridiagonale).

Condizionamento del problema di autovalori.

Concludiamo il capitolo dimostrando che, per matrici hermitiane, il problema degli autovalori è stabile. Per il corollario 1.5 sappiamo che la matrice  $A$  è diagonalizzabile ed ammette un sistema ortogonale di autosoluzioni. Detta  $X$  la matrice le cui colonne sono gli autovettori di  $A$ , e detta  $\Lambda$  la matrice diagonale degli autovalori, si ha:

$$AX = X\Lambda$$

con  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  e  $X^T X = I$ .

Siano ora  $\lambda_i(\epsilon)$  gli autovalori della matrice perturbata  $A(\epsilon) = A + \epsilon B$  con  $\|B\|_2 = 1$ , supponiamo cioè che la perturbazione sia di ampiezza  $\epsilon$ .

Si ha:

$$X^{-1}A(\epsilon)X = X^{-1}AX + \epsilon X^{-1}BX = \Lambda + \epsilon C$$

dove  $C = X^{-1}BX$ .

Per il teorema di Gerschgorin, ogni  $\lambda_i(\epsilon)$  sta nel disco di centro  $\lambda_i + \epsilon c_{ii}$  e raggio  $\epsilon$

$\sum_{j \neq i} |c_{ij}|$ , ossia:

$$|\lambda_i(\epsilon) - (\lambda_i + \epsilon c_{ii})| \leq \epsilon \sum_{j \neq i} |c_{ij}|$$

$$|\lambda_i(\epsilon) - \lambda_i| \leq \epsilon \sum_j |c_{ij}| \leq \epsilon \|C\|_\infty \quad \forall_i$$

Ora, per l'equivalenza delle norme,

$$\|C\|_{\infty} \leq M \|C\|_2$$

e, per l'ortogonalità della matrice X:

$$\|C\|_2 \leq \|X^{-1}BX\|_2 \leq \|X^{-1}\|_2 \|B\|_2 \|X\|_2 = \|B\|_2 = 1$$

In conclusione:

$$|\lambda_i(\varepsilon) - \lambda_i| \leq \varepsilon M$$

Poichè la costante M (che si dimostra essere uguale a  $\sqrt{n}$ ) non dipende dalla matrice A ma solo dalla norma nella quale ho misurato la perturbazione, abbiamo dimostrato che lo spostamento degli autovalori è proporzionale alla norma della perturbazione stessa. Ciò assicura che il problema è ben condizionato.