

## Metodi basati su iterazioni in sottospazi di Krylov

L'idea di proiettare in un sottospazio di ordine inferiore un problema relativo ad una matrice  $A \in \mathbb{R}^{n \times n}$ , avente un numero di elementi non nulli dell'ordine di  $n$ , nasce dall'esigenza di ridurre la complessità computazionale degli algoritmi atti a risolvere onerosi problemi  $n$ -dimensionali quali la soluzione di sistemi lineari aventi  $A$  come matrice dei coefficienti, o l'approssimazione dello spettro di  $A$ .

Per quanto riguarda il primo dei due problemi computazionali, il *metodo GMRES* minimizza al  $k$ -esimo passo la norma euclidea del residuo in un sottospazio di Krylov di dimensione  $k$ . Sempre in questo ambito, è di interesse una rilettura del metodo del gradiente coniugato (CG) come metodo di Krylov. Come si è detto, il metodo CG si applica a matrici simmetriche definite positive e permette di arrivare alla soluzione del sistema al più in  $n$  passi. Un altro metodo di Krylov è una sua variante, il *metodo CGN*, che funziona anche nel caso che  $A$  non sia simmetrica definita positiva e che si applica alle equazioni normali  $A^T A \mathbf{x} = A^T \mathbf{b}$ .

Per il problema agli autovalori verranno trattati due metodi: il *metodo di Arnoldi*, basato - come il metodo GMRES - sull'iterazione di Arnoldi, e il *metodo di Lanczos*, specifico per matrici simmetriche sparse e di grandi dimensioni, basato sull'iterazione di Lanczos.

La matrice  $A$  - sparsa e di grandi dimensioni - potrebbe anche non essere esplicitamente a nostra disposizione. Ciò di cui dobbiamo disporre è un algoritmo, definito *black box*, che riceva in input il vettore  $\mathbf{x} \in \mathbb{R}^n$  e restituisca in output il vettore  $A\mathbf{x} \in \mathbb{R}^n$ , ovvero

$$\mathbf{x} \rightarrow \boxed{\text{black box}} \rightarrow A\mathbf{x}$$

Ipotizziamo ovviamente che questo prodotto matrice-vettore avvenga con un costo computazionale molto inferiore a  $2n^2$ .

### 4.1. Iterazione di Arnoldi

Si definisce *sottospazio di Krylov* di ordine  $k$ , con  $1 \leq k \leq n$ , relativo ad  $A$  ed al vettore  $\mathbf{v}$ , lo spazio definito come

$$K_k(A; \mathbf{v}) = \text{span}\{\mathbf{v}, A\mathbf{v}, \dots, A^{k-1}\mathbf{v}\}. \quad (4.1)$$

Ovviamente vale  $K_k(A; \mathbf{v}) \subseteq K_{k+1}(A; \mathbf{v})$ , e che  $\mathbf{z} \in K_k(A; \mathbf{v})$  è un vettore della forma  $\varphi_{k-1}(A)\mathbf{v}$  per qualche polinomio  $\varphi_{k-1}$  di grado al più  $k-1$ .

Se il problema computazionale assegnato è la soluzione del sistema lineare  $A\mathbf{x} = \mathbf{b}$ , a partire da un generico vettore iniziale  $\mathbf{x}^{(0)}$  e dal residuo corrispondente  $\mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$ , si pone

$$W_k = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} = \mathbf{x}^{(0)} + \mathbf{z}, \mathbf{z} \in K_k(A; \mathbf{r}^{(0)})\}. \quad (4.2)$$

Se il problema computazionale assegnato è la ricerca degli autovalori di  $A$ , a partire da un generico vettore iniziale  $\mathbf{v} \neq \mathbf{0}$ , si pone

$$W_k = K_k(A; \mathbf{v}).$$

Un *metodo di Krylov* per la soluzione per la soluzione del sistema lineare  $A\mathbf{x} = \mathbf{b}$  è un metodo che genera soluzioni approssimanti  $\mathbf{x}^{(k)}$  con la condizione che  $\mathbf{x}^{(k)}$  sia l'unico elemento in  $W_k$  a soddisfare un certo criterio di minima distanza dalla soluzione  $\mathbf{x}$  del sistema. Ogni metodo di Krylov ha un criterio associato. Il criterio prescelto è la caratteristica che distingue metodi di Krylov differenti. Per quanto riguarda la soluzione del sistema  $A\mathbf{x} = \mathbf{b}$ , il metodo GMRES (Generalized Minimal RESidual) calcola  $\mathbf{x}^{(k)}$  tale che  $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$  abbia minima norma euclidea in  $W_k$ . Il metodo MINRES (MINimal RESidual) adotta il medesimo criterio ma è relativo a matrici dei coefficienti simmetriche. Il metodo CG, come vedremo, di fatto risulta soddisfare il criterio di minimizzare globalmente la norma dell'energia dell'errore  $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$  in  $W_k$ .

Il primo obiettivo è generare una base ortonormale per lo spazio  $K_k(A; \mathbf{v})$ , con  $k \leq n$ . Per far ciò si sfrutta la seguente procedura alla Gram-Schmidt, che prende il nome di *iterazione di Arnoldi*:

```

 $\mathbf{v}_1 = \mathbf{v} / \|\mathbf{v}\|_2$ 
per  $k = 1, 2, \dots$ 
     $\mathbf{w} = A\mathbf{v}_k$ 
    per  $i = 1, \dots, k$ 
         $h_{ik} = \mathbf{v}_i^T \mathbf{w}$ 
    fine
     $\mathbf{w}_k = \mathbf{w} - \sum_{i=1}^k h_{ik} \mathbf{v}_i$ 
     $h_{k+1,k} = \|\mathbf{w}_k\|_2$ 
     $\mathbf{v}_{k+1} = \begin{cases} \mathbf{w}_k / h_{k+1,k} & \text{se } h_{k+1,k} \neq 0 \\ STOP & \text{se } h_{k+1,k} = 0 \end{cases}$ 
fine

```

Ad un certo passo  $m \leq n$ , si ha  $\mathbf{w}_m = \mathbf{0}$ , ovvero si incorre in un breakdown che forza il processo a terminare al passo  $m$ . Ad ogni passo  $k < m$ , i vettori  $\mathbf{v}_i$ , per  $i = 1, \dots, k+1$ , possono essere memorizzati come vettori colonna della matrice  $V_{k+1} \in \mathbb{R}^{n \times (k+1)}$ .

Si ha, per  $k < m$ ,

$$AV_k = V_{k+1} \hat{H}_k, \quad (4.3)$$

dove  $\hat{H}_k \in \mathbb{R}^{(k+1) \times k}$  è la matrice formata dagli elementi  $h_{i,j}$  restituiti dall'algoritmo (e zeri per  $i > j + 1$ ) e  $V_k$  è la matrice che ha per colonne i vettori della base ortonormale di  $K_k(A; \mathbf{v})$  che è stata costruita. Infatti, la  $j$ -esima colonna,  $j \leq k$ , della matrice che compare al primo membro della (4.3) soddisfa

$$A\mathbf{v}_j = h_{1,j}\mathbf{v}_1 + \cdots + h_{j,j}\mathbf{v}_j + h_{j+1,j}\mathbf{v}_{j+1},$$

ovvero è uguale alla  $j$ -esima colonna della matrice che compare al secondo membro della (4.3).

Vedremo che nel metodo GMRES si farà uso della formula (4.3).

Inoltre si ha, per  $k \leq m$ ,

$$V_k^T AV_k = H_k, \quad (4.4)$$

dove  $H_k$  è la matrice di Hessenberg superiore in  $\mathbb{R}^{k \times k}$  formata dagli  $h_{i,j}$  restituiti dall'algoritmo, ad esclusione dell'elemento  $h_{k+1,k}$ . Infatti, nel caso particolare in cui  $k = m$ , e quindi  $h_{m+1,m} = 0$ , si ha che  $H_m$  è tale che  $AV_m = V_m H_m$ , dunque  $V_m^T AV_m = H_m$ ; in generale, ovvero quando  $k$  è minore di  $m$ , dato che  $V_k^T V_{k+1} = I_{k,k+1}$ , dove con  $I_{k,k+1}$  denotiamo la matrice identità in  $\mathbb{R}^{k \times (k+1)}$  (formata dalla matrice identità  $I_k$  e a seguire una colonna di zeri), segue da (4.3)

$$V_k^T AV_k = I_{k,k+1} \hat{H}_k,$$

dove la matrice a secondo membro è chiaramente la restrizione di  $\hat{H}_k$  alle prime  $k$  righe e colonne, ovvero  $H_k$ .

Se  $k = m = n$ , la (4.4) è una trasformazione ortogonale di similitudine (in tal caso però - soprattutto se la matrice non è di dimensioni ingenti, può risultare più conveniente ricorrere al ben più stabile algoritmo di Householder, che discuteremo nel prossimo capitolo). Altrimenti, ovvero nel caso generale in cui  $k \leq m < n$ , la  $H_k$  è di fatto una proiezione di  $A$  in  $K_k(A; \mathbf{v})$  rappresentata nella base  $V_k$ .

La formula (4.4) sarà utilizzata nel metodo di Arnoldi nel contesto della ricerca di autovalori di  $A$ .

Ricordiamo che il ruolo del vettore  $\mathbf{v}$  che compare nell'iterazione di Arnoldi, nel caso dei sistemi lineari (metodo GMRES) è giocato da  $\mathbf{r}^{(0)}$  (o da  $\mathbf{b}$  se si assume  $\mathbf{x}^{(0)} = \mathbf{0}$ ), mentre nella ricerca degli autovalori di  $A$  (metodo di Arnoldi) viene assunto da un vettore arbitrario non nullo.

## 4.2. Sistemi lineari e metodo GMRES

Al passo  $k$ -esimo - dove  $k$  è assunto abbastanza piccolo da non essere incorsi in un breakdown nel generare la matrice  $V_{k+1}$  con colonne ortonormali associata al sottospazio di Krylov  $K_k(A; \mathbf{r}^{(0)})$  - il metodo GMRES (Saad, 1986)

definisce l'approssimazione  $\mathbf{x}^{(k)}$  della soluzione  $\mathbf{x}$  del sistema lineare assegnato, imponendo che la norma euclidea del residuo  $\mathbf{b} - A\mathbf{x}^{(k)}$  sia minima in  $W_k$ , ovvero

$$\mathbf{x}^{(k)} = \arg \min_{\mathbf{v} \in W_k} \|\mathbf{b} - A\mathbf{v}\|_2, \quad \|\mathbf{b} - A\mathbf{x}^{(k)}\|_2 = \min_{\mathbf{v} \in W_k} \|\mathbf{b} - A\mathbf{v}\|_2.$$

Dato che  $\mathbf{x}^{(k)} \in W_k$ , si ha  $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + \mathbf{z}$ , con  $\mathbf{z} \in K_k(A; \mathbf{r}^{(0)})$ . Possiamo rappresentare  $\mathbf{z}$  nella base ortonormale  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  fornitaci dall'iterazione di Arnoldi, assumendo che  $\mathbf{y} \in \mathbb{R}^k$  sia tale che  $\mathbf{z} = V_k \mathbf{y}$ . Quindi, per il residuo al passo  $k$  vale

$$\begin{aligned} \mathbf{r}^{(k)} &= \mathbf{b} - A\mathbf{x}^{(k)} = \mathbf{b} - A(\mathbf{x}^{(0)} + V_k \mathbf{y}) \\ &= \mathbf{r}^{(0)} - AV_k \mathbf{y} = \mathbf{r}^{(0)} - V_{k+1} \hat{H}_k \mathbf{y}, \end{aligned}$$

dove si è tenuto conto della (4.3). Inoltre, dato che

$$\mathbf{v}_1 = \frac{\mathbf{r}^{(0)}}{\|\mathbf{r}^{(0)}\|_2} = V_{k+1} \mathbf{e}_1,$$

dove  $\mathbf{e}_1$  è il primo vettore della base canonica di  $\mathbb{R}^{k+1}$ , si ha

$$\mathbf{r}^{(0)} = V_{k+1} \|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1.$$

Quindi, mettendo in evidenza  $V_{k+1}$ , per il residuo al passo  $k$  vale

$$\mathbf{r}^{(k)} = \mathbf{r}^{(0)} - V_{k+1} \hat{H}_k \mathbf{y} = V_{k+1} (\|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1 - \hat{H}_k \mathbf{y}).$$

Concludiamo osservando che il metodo GMRES, ovvero il minimizzare la norma del residuo  $k$ -esimo in  $W_k$ , si riduce al risolvere il problema ai minimi quadrati (di dimensione inferiore!)

$$\min_{\mathbf{y} \in \mathbb{R}^k} \|\|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1 - \hat{H}_k \mathbf{y}\|_2. \quad (4.5)$$

Infatti, ricordiamo che per  $\mathbf{w} \in \mathbb{R}^{k+1}$ , dato che  $V_{k+1}$  è una matrice in  $\mathbb{R}^{n \times (k+1)}$  con colonne ortonormali, vale  $\|V_{k+1} \mathbf{w}\|_2^2 = \mathbf{w}^T V_{k+1}^T V_{k+1} \mathbf{w} = \mathbf{w}^T I_{k+1} \mathbf{w} = \|\mathbf{w}\|_2^2$ .

**4.2.1. Algoritmo GMRES.** L'algoritmo risolve ad ogni passo un problema ai minimi quadrati:

$$\mathbf{v}_1 = \mathbf{r}^{(0)} / \|\mathbf{r}^{(0)}\|_2$$

**per**  $k = 1, 2, \dots$

si esegue il  $k$ -esimo passo dell'iterazione di Arnoldi

si cerca  $\mathbf{y}$  che minimizzi (4.5)

si aggiorna la soluzione  $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + V_k \mathbf{y}$

**fine**

Dato che  $\|\mathbf{r}^{(k)}\|_2$  è minima in  $W_k$ , allargando lo spazio da  $W_k$  a  $W_{k+1}$ , la norma del residuo della soluzione approssimata al passo  $(k+1)$ -esimo non può crescere, quindi la convergenza dell'algoritmo è monotona, ovvero

$$\|\mathbf{r}^{(k+1)}\|_2 \leq \|\mathbf{r}^{(k)}\|_2.$$

Inoltre, in aritmetica esatta al più in  $n$  passi (se l'iterazione di Arnoldi non ha subito un breakdown; si veda il Paragrafo 4.2.3) l'algoritmo termina restituendo la soluzione esatta: dato che  $W_n = \mathbb{R}^n$ , la norma dell' $n$ -esimo residuo deve essere nulla. In aritmetica finita, l'algoritmo GMRES viene fermato in genere per  $k \ll n$ , anche perché, oltre all'elevato costo computazionale (ricordiamo che al  $k$ -esimo passo il metodo deve risolvere un problema ai minimi quadrati di ordine  $k$ ), il metodo GMRES rivela instabilità.

Si può dimostrare che la velocità di convergenza dipende dalla localizzazione e dal condizionamento degli autovalori di  $A$ . Per garantire la convergenza in tempo breve si può, al solito, preconditionare il sistema con una matrice  $P$  la cui inversione sia bastevolmente economica e tale che  $P^{-1}A$  sia quasi-normale e con spettro clusterizzato lontano dall'origine. Notiamo esplicitamente che al sistema preconditionato  $P^{-1}Ax = P^{-1}\mathbf{b}$  è associato un differente sottospazio di Krylov,

$$K_k(P^{-1}A; \mathbf{z}^{(0)}) = \text{span}\{\mathbf{z}^{(0)}, P^{-1}A\mathbf{z}^{(0)}, \dots, (P^{-1}A)^{k-1}\mathbf{z}^{(0)}\},$$

avendo denotato con  $\mathbf{z}^{(0)}$  il residuo preconditionato  $P^{-1}\mathbf{r}^{(0)} = P^{-1}(\mathbf{b} - A\mathbf{x}^{(0)})$ .

**4.2.2. Strategia del restart: GMRES( $p$ ).** Il metodo GMRES necessita di un ingente spazio di memoria per mantenere tutti i vettori della base ortonormale del sottospazio di Krylov di ordine  $k$ . Un modo per evitare tale spreco di memoria è quello di definire a priori il numero massimo di vettori che si possono memorizzare, sia ad esempio  $p \ll n$  tale numero. La variante del metodo GMRES( $p$ ) consiste nel restart dopo  $p$  passi. In maggior dettaglio, dopo che sono state effettuate  $p$  iterazioni per ottenere  $\mathbf{x}^{(p)}$ , si cancellano tutti i vettori della base già memorizzati e si riparte con un nuovo sottospazio di Krylov avente come vettore generante  $\mathbf{r}^{(0)} := \mathbf{b} - A\mathbf{x}^{(p)}$  (analogamente al restart nel metodo CG; si veda Osservazione 3.6). Se da un lato si guadagna certamente in spazio memoria e in stabilità, dall'altro il restart può rallentare la convergenza alla soluzione.

**4.2.3. Soluzione esatta in  $m$  passi.** Se si è incorsi in un breakdown nell'iterazione di Arnoldi con  $m < n$ , si ha  $AV_m = V_m H_m$ , e vale

$$\mathbf{r}^{(m)} = V_m(\|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1 - H_m \mathbf{y}),$$

dove  $\mathbf{e}_1$  è il primo vettore della base canonica di  $\mathbb{R}^m$ . Quindi  $\mathbf{r}^{(m)}$  si annulla (ovvero  $\mathbf{x}^{(m)}$  coincide con la soluzione esatta del sistema!) se  $\mathbf{y}^{(m)}$  risolve il sistema lineare  $H_m \mathbf{y} = \|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1$ , avendosi  $\mathbf{x}^{(m)} = \mathbf{x}^{(0)} + V_m \mathbf{y}^{(m)} \equiv \mathbf{x} \in W_m$ .

### 4.3. Il metodo CG è un metodo di Krylov

Sia  $A \in \mathbb{R}^{n \times n}$  simmetrica definita positiva e sia  $K_k(A; \mathbf{r}^{(0)})$  lo spazio di Krylov di ordine  $k$  generato da  $A$  e  $\mathbf{r}^{(0)}$ . Vogliamo dimostrare che il metodo CG è un metodo di Krylov. Poniamo per comodità  $\mathbf{x}^{(0)} = \mathbf{0}$  (quindi  $\mathbf{r}^{(0)} = \mathbf{b}$ , ma per maggior chiarezza manteniamo la notazione  $\mathbf{r}^{(0)}$ ). Osserviamo quindi che, per  $k \geq 1$ ,

si ha  $\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + \alpha_{k-1} \mathbf{p}^{(k-1)} \in \text{span}\{\mathbf{p}^{(0)}, \dots, \mathbf{p}^{(k-1)}\}$ . Si possono dimostrare le seguenti identità che estendono (3.18),

$$\begin{aligned} K_k(A; \mathbf{r}^{(0)}) &= \text{span}\{\mathbf{r}^{(0)}, A\mathbf{r}^{(0)}, \dots, A^{k-1}\mathbf{r}^{(0)}\} = \\ \text{span}\{\mathbf{r}^{(0)}, \dots, \mathbf{r}^{(k-1)}\} &= \text{span}\{\mathbf{p}^{(0)}, \dots, \mathbf{p}^{(k-1)}\}. \end{aligned}$$

Perché il metodo CG sia un metodo di Krylov, in base alla definizione c'è bisogno che  $\mathbf{x}^{(k)}$  sia la migliore approssimazione di  $\mathbf{x}$ , secondo un qualche criterio, nel sottospazio di Krylov considerato. Il seguente teorema mostra che il metodo CG minimizza in  $K_k(A; \mathbf{r}^{(0)})$  la  $A$ -norma dell'errore.

**TEOREMA 4.1.** *Sia  $A \in \mathbb{R}^{n \times n}$  simmetrica definita positiva, applicando il metodo CG con  $\mathbf{x}^{(0)} = \mathbf{0}$ , la soluzione approssimata  $\mathbf{x}^{(k)}$  risulta l'unico vettore a rendere minima la norma dell'energia dell'errore, ovvero*

$$\|\mathbf{e}^{(k)}\|_A = \min_{\mathbf{v} \in K_k(A; \mathbf{r}^{(0)})} \|\mathbf{x} - \mathbf{v}\|_A.$$

Per  $k \geq 1$ , si ha  $\|\mathbf{e}^{(k)}\|_A \leq \|\mathbf{e}^{(k-1)}\|_A$ .

**DIMOSTRAZIONE.** Si abbia  $\hat{\mathbf{x}} = \mathbf{x}^{(k)} - \Delta\mathbf{x} \in K_k(A; \mathbf{r}^{(0)})$ . L'errore associato a  $\hat{\mathbf{x}}$  sarà  $\hat{\mathbf{e}} = \mathbf{x} - \hat{\mathbf{x}} = \mathbf{x} - \mathbf{x}^{(k)} + \Delta\mathbf{x}$ . Passando alla  $A$ -norma

$$\|\hat{\mathbf{e}}\|_A^2 = (\mathbf{e}^{(k)} + \Delta\mathbf{x})^T A (\mathbf{e}^{(k)} + \Delta\mathbf{x}) = \mathbf{e}^{(k)T} A \mathbf{e}^{(k)} + \Delta\mathbf{x}^T A \Delta\mathbf{x} + 2\mathbf{e}^{(k)T} A \Delta\mathbf{x}.$$

Dal Teorema 3.16,  $\mathbf{r}^{(k)}$  è ortogonale a  $\text{span}\{\mathbf{r}^{(0)}, \dots, \mathbf{r}^{(k-1)}\} \equiv K_k(A; \mathbf{r}^{(0)})$ , quindi si ha  $2\mathbf{e}^{(k)T} A \Delta\mathbf{x} = 2(\mathbf{Ae}^{(k)})^T \Delta\mathbf{x} = 2\mathbf{r}^{(k)T} \Delta\mathbf{x} = 0$ , e vale

$$\|\hat{\mathbf{e}}\|_A^2 = \mathbf{e}^{(k)T} A \mathbf{e}^{(k)} + \Delta\mathbf{x}^T A \Delta\mathbf{x},$$

con  $\Delta\mathbf{x}^T A \Delta\mathbf{x} \geq 0$ , essendo la matrice  $A$  definita positiva. Quindi  $\|\hat{\mathbf{e}}\|_A \geq \|\mathbf{e}^{(k)}\|_A$ , e l'uguaglianza si ha solo per  $\Delta\mathbf{x} = \mathbf{0}$ , ovvero per  $\hat{\mathbf{x}} = \mathbf{x}^{(k)}$ . Infine, la monotonia della convergenza del metodo CG è conseguenza del fatto che  $K_{k-1}(A; \mathbf{r}^{(0)}) \subseteq K_k(A; \mathbf{r}^{(0)})$ .  $\square$

**OSSERVAZIONE 4.1.** *A partire da  $\mathbf{x}^{(0)} = \mathbf{0}$ , il metodo CG genera quindi al passo  $k$ -esimo la soluzione approssimata  $\mathbf{x}^{(k)} \in K_k(A; \mathbf{r}^{(0)})$  che minimizza globalmente la  $A$ -norma dell'errore  $\mathbf{e}^{(k)}$  in  $K_k(A; \mathbf{r}^{(0)})$ ; in questo senso si può dire quindi che il metodo CG è un metodo di Krylov. Se si parte da un arbitrario  $\mathbf{x}^{(0)}$ , il ragionamento rimane inalterato, con  $W_k$  che prende il posto di  $K_k(A; \mathbf{r}^{(0)})$ . Al contrario, un generico metodo del gradiente (ad esempio il metodo steepest descent) genera al passo  $k$ -esimo la soluzione approssimata  $\mathbf{x}^{(k)}$  che risulta essere l'unica a minimizzare, solo lungo la direzione  $\mathbf{p}^{(k)}$ , la  $A$ -norma dell'errore  $\mathbf{e}^{(k)}$ ; si veda la Proposizione 3.11.*

#### 4.4. Metodo CG applicato alle equazioni normali (CGN)

Si consideri un sistema di equazioni lineari con matrice dei coefficienti  $A \in \mathbb{R}^{n \times n}$  invertibile. Oltre a metodo GMRES, riveste particolare importanza il metodo del gradiente coniugato applicato alle equazioni normali

$$A^T A \mathbf{x} = A^T \mathbf{b}, \quad (4.6)$$

dove  $A^T A$ , come è noto, è simmetrica definita positiva. Ad ogni iterazione il costo computazionale è ridotto, d'altro canto entra però in gioco anche la matrice  $A^T \neq A$ , che potrebbe non essere nota esplicitamente, ovvero potrebbe non essere disponibile un algoritmo che riceve in input il vettore  $\mathbf{x} \in \mathbb{R}^n$  e restituisce in output il vettore  $A^T \mathbf{x} \in \mathbb{R}^n$ .

Nel metodo CG applicato al sistema (4.6) (che viene in genere definito metodo CGN), a partire da  $\mathbf{x}^{(0)} = \mathbf{0}$ , le iterate  $\mathbf{x}^{(k)}$  appartengono al sottospazio di Krylov

$$K_k(A^T A; A^T \mathbf{b}) = \text{span}\{A^T \mathbf{b}, (A^T A)A^T \mathbf{b}, \dots, (A^T A)^{k-1} A^T \mathbf{b}\},$$

dato che

$$\mathbf{r}_{A^T A}^{(0)} = A^T (\mathbf{b} - A\mathbf{x}^{(0)}) = A^T \mathbf{b}.$$

Per il Teorema 4.1,  $\mathbf{x}^{(k)}$  è l'unico vettore in  $K_k(A^T A; A^T \mathbf{b})$  a minimizzare la norma  $A^T A$  dell'errore al passo  $k$ -esimo. Notiamo che vale

$$\|\mathbf{e}^{(k)}\|_{A^T A}^2 = \mathbf{e}^{(k)T} A^T A \mathbf{e}^{(k)} = \|A \mathbf{e}^{(k)}\|_2^2 = \|\mathbf{r}^{(k)}\|_2^2,$$

essendo  $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)} = A(\mathbf{x} - \mathbf{x}^{(k)})$ . Si deduce quindi che anche il metodo CGN - come il metodo GMRES - è un metodo di Krylov nel quale la norma euclidea del residuo al passo  $k$ -esimo viene minimizzata dall'iterata  $\mathbf{x}^{(k)}$ . D'altro canto, i due spazi di Krylov sono differenti quindi i due metodi non sono equivalenti.

Dato che il numero di condizionamento spettrale della matrice  $A^T A$  è uguale a  $\kappa_2(A)^2$ , si ha per il metodo CGN,

$$\|\mathbf{r}^{(k+1)}\|_2 \leq 2 \left( \frac{\kappa_2(A) - 1}{\kappa_2(A) + 1} \right)^{k+1} \|\mathbf{r}^{(0)}\|_2,$$

**OSSERVAZIONE 4.2.** *A differenza del metodo GMRES, nel caso del metodo CGN si dispone dunque di una maggiorazione esplicita per la norma euclidea del residuo al passo  $k$ -esimo. Poiché sistemi lineari malcondizionati - con  $\kappa_2(A)$  che cresce almeno come  $n^\alpha$ ,  $\alpha \geq 2$  - sono frequentissimi nei problemi differenziali, dalla maggiorazione precedente si deduce che il metodo CGN risulta sicuramente utile quando, oltre ad essere disponibile la routine  $\mathbf{x} \rightarrow \boxed{\text{black box}} \rightarrow A^T \mathbf{x}$ , è anche noto che la matrice  $A$  è sufficientemente ben condizionata. Altrimenti, pur con l'aggravio di costo computazionale e spazio memoria elevati, si preferisce il metodo GMRES. Aggiungiamo che se la matrice  $A$  è simmetrica, ma non definita positiva, si usa il metodo CGN se è sufficientemente ben condizionata, altrimenti viene preferito il metodo MINRES.*

#### 4.5. Problemi agli autovalori e metodo di Arnoldi

L'iterazione di Arnoldi oltre ad essere alla base di molti algoritmi di algebra lineare numerica, costituisce una tecnica per approssimare autovalori di matrici non simmetriche sparse e di grandi dimensioni. Si considera il problema agli autovalori

$$A\mathbf{x} = \lambda\mathbf{x}, \tag{4.7}$$

con  $A \in \mathbb{R}^{n \times n}$  e  $\mathbf{x} \neq \mathbf{0}$ . Si è detto che la matrice di Hessenberg superiore  $H_k$  che si ottiene dall'iterazione di Arnoldi a partire da un vettore  $\mathbf{v} \in \mathbb{R}^n$ ,  $\mathbf{v} \neq \mathbf{0}$  si veda (4.4), può essere interpretata come una proiezione di  $A$  in  $K_k(A; \mathbf{v})$  rappresentata nella base  $V_k$ . Siano  $\vartheta_i$ , per  $i = 1, \dots, k$ , gli autovalori di  $H_k$ , detti *valori di Ritz* di  $A$  (rispetto a  $K_k(A; \mathbf{v})$ ). Si può procedere al calcolo dei  $\vartheta_i$  usando il metodo *QR*, che discuteremo nel prossimo capitolo, con il vantaggio di avere una matrice già in formato Hessenberg superiore. I  $k$  valori di Ritz così ottenuti approssimano  $k$  autovalori di  $A$ . Andando avanti con le iterazioni, ovvero calcolando gli autovalori di matrici di Hessenberg di dimensioni sempre più grandi, un numero sempre maggiore di autovalori di  $A$  - euristicamente a partire dal bordo dello spettro verso l'interno - potrà essere approssimato, e la qualità dell'approssimazione degli autovalori (sempre a partire dal bordo dello spettro) migliorerà progressivamente.

Anche nel contesto dell'approssimazione dello spettro di  $A$ , essere incorsi in un breakdown nell'iterazione di Arnoldi con  $m < n$  si rivela una situazione vantaggiosa. Si dimostra che gli  $m$  valori di Ritz (calcolabili, come si è detto, con il metodo QR applicato a  $H_m$ ) sono essi stessi autovalori di  $A$ ! Per approssimare gli  $n - m$  rimanenti autovalori di  $A$ , si può procedere con l'iterazione di Arnoldi a partire da un nuovo vettore arbitrario che non appartenga a  $K_m(A; \mathbf{v})$ .

#### 4.6. Iterazione e metodo di Lanczos per matrici simmetriche

Nel caso particolare in cui  $A$  sia simmetrica, l'iterazione di Arnoldi si rivela computazionalmente inadatta a generare una base ortonormale per lo spazio  $K_k(A; \mathbf{v})$ , con  $k \leq n$ . Osserviamo infatti che, dato che  $V_k^T A V_k$  è simmetrica, la (4.4) implica che la matrice di Hessenberg superiore  $H_k$  è simmetrica, quindi tridiagonale. Il procedimento si può dunque semplificare, dato che

- $h_{i,k} = 0$  per  $i < k - 1$ , dunque il ciclo interno da 1 a  $k$  si riduce di fatto alle componenti  $k - 1$  e  $k$ ;
- $h_{k+1,k}$  coincide con  $h_{k,k+1}$ .

L'iterazione di Lanczos, dove la notazione viene alleggerita ponendo  $\alpha_k = h_{k,k}$  e  $\beta_k = h_{k+1,k} = h_{k,k+1}$ , consiste nella seguente procedura:

```

 $\beta_0 = 0$ 
 $\mathbf{q}_0 = \mathbf{0}$ 
 $\mathbf{q}_1 = \mathbf{v} / \|\mathbf{v}\|_2$ 
per  $k = 1, 2, \dots$ 
     $\mathbf{w} = A\mathbf{q}_k$ 
     $\alpha_k = \mathbf{q}_k^T \mathbf{w}$ 
     $\mathbf{w}_k = \mathbf{w} - \beta_{k-1}\mathbf{q}_{k-1} - \alpha_k\mathbf{q}_k$ 
     $\beta_k = \|\mathbf{w}_k\|_2$ 
     $\mathbf{q}_{k+1} = \begin{cases} \mathbf{w}_k / \beta_k & \text{se } \beta_k \neq 0, \\ STOP & \text{se } \beta_k = 0 \end{cases}$ 
fine

```

Ad ogni passo  $k$ , si ha la relazione a tre termini

$$A\mathbf{q}_k = \beta_{k-1}\mathbf{q}_{k-1} + \alpha_k\mathbf{q}_k + \beta_k\mathbf{q}_{k+1}.$$

Se i vettori  $\mathbf{q}_i$ ,  $i = 1, \dots, k$ , sono memorizzati come vettori colonna della matrice  $Q_k \in \mathbb{R}^{n \times k}$ , ragionando come per l'iterazione di Arnoldi, si ha

$$Q_k^T A Q_k = T_k, \quad (4.8)$$

dove

$$T_k = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & \dots & 0 \\ \beta_1 & \alpha_2 & \beta_2 & \ddots & \\ 0 & \beta_2 & \alpha_3 & \ddots & 0 \\ \vdots & & \ddots & \ddots & \beta_{k-1} \\ 0 & & & \beta_{k-1} & \alpha_k \end{pmatrix} \in \mathbb{R}^{k \times k} \quad (4.9)$$

è la matrice tridiagonale simmetrica restituita dall'algoritmo. Le operazioni da svolgere al  $k$ -esimo passo sono quindi notevolmente ridotte rispetto all'iterazione di Arnoldi, non dovendo calcolare  $k+1$  termini. Portando avanti fino all' $n$ -esimo passo l'iterazione di Lanczos, se non si è incorsi in un breakdown anticipato, da (4.8) si ha

$$A Q_n = Q_n T_n, \quad Q_n^T A Q_n = T_n,$$

ovvero la matrice tridiagonale  $T_n$  è ortogonalmente simile ad  $A$ . L'iterazione di Lanczos - oltre a essere alla base del metodo MINRES per la soluzione di sistemi lineari con matrice dei coefficienti  $A$  simmetrica, sul quale non ci soffermeremo - dà luogo al metodo di Lanczos (1950) per l'approssimazione dello spettro di  $A$ . In particolare, viene approssimata parte degli autovalori (che sappiamo essere reali e supponiamo denotati in ordine decrescente,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ ) calcolando con metodi standard al passo  $k$ -esimo gli autovalori di  $T_k$ ,  $\mu_1^{(k)} \geq \mu_2^{(k)} \geq \dots \geq \mu_k^{(k)}$ , ovvero i valori di Ritz di  $A$  (rispetto a  $K_k(A; \mathbf{v})$ ). Si può dimostrare che il metodo di Lanczos, se il vettore iniziale è opportunamente generico, approssima

piuttosto velocemente autovalori che sono ben separati dal resto dello spettro; in genere, si tratta degli autovalori estremali, dato che generalmente non ci sono clusters agli estremi dello spettro. La qualità dell'approssimazione migliora con le iterazioni. Per  $k \rightarrow n$  si ha

$$\{\mu_1^{(k)}\} \nearrow \lambda_1, \quad \{\mu_k^{(k)}\} \searrow \lambda_n,$$

ottenendo agevolmente, se  $\lambda_n > 0$ , una approssimazione per difetto del numero di condizionamento  $\kappa_2(A)$ .

Quasi come per l'algoritmo di Arnoldi, l'algoritmo di Lanczos, senza strategie di riortogonalizzazione completa o selettiva dei vettori della base, si rivela purtroppo instabile già per  $k$  moderatamente grande. La motivazione di tale perdita di ortogonalità dei vettori di Lanczos, nonostante i conti semplificati e la minore probabilità di avere cancellazione numerica nell'iterazione di Lanczos, è da imputarsi in questo caso alla costruzione della ricorrenza a tre termini facendo uso di una (implicita) identità matematica, e non - come nell'iterazione di Arnoldi - su una (esplicita) procedura alla Gram-Schmidt.