

Metodo di Jacobi-Davidson per il calcolo di autovalori

Hergert Gjoni

Calcolo Scientifico, 2019

Introduzione

In questo lavoro verrà presentato un algoritmo per il calcolo dei più piccoli autovalori di una data matrice $A \ n \times n$ dove con "più piccoli" si intende con segno (solo se gli autovalori sono reali) oppure in modulo.

L'algoritmo prende il nome di **Algoritmo di Jacobi-Davidson**

L'articolo a cui faremo riferimento sarà il testo di Arbenz-Kressner
Lecture Notes on Solving Large Scale Eigenvalue Problems
[Capitolo 11]

Spazio di ricerca

Siano $\mathbf{v}_1, \dots, \mathbf{v}_k$ vettori di \mathbb{C}^n ortonormali. Essi formeranno un sottospazio di \mathbb{C}^n che chiameremo *spazio di ricerca*.

Seguiremo l'approccio di Galerkin, ovvero cercheremo un vettore $\mathbf{s} \in \mathbb{C}^k$ tale che

$$AV_k \mathbf{s} - \vartheta \mathbf{s} \perp \mathbf{v}_1, \dots, \mathbf{v}_k \quad (1)$$

Dove $V_k = [\mathbf{v}_1, \dots, \mathbf{v}_k]$, e questo porta direttamente al seguente problema agli autovalori: (di piccola dimensione)

$$V_k^* AV_k \mathbf{s} = \vartheta V_k^* V_k \mathbf{s} = \vartheta \mathbf{s} \quad (2)$$

Le soluzioni le chiameremo $(\vartheta_j^{(k)}, \mathbf{s}_j^{(k)})$, con $j = 1, \dots, k$:

- ▶ $\vartheta_j^{(k)}$ si chiamerà *valore di Ritz*.
- ▶ $\mathbf{u}_j^{(k)} = V_k \mathbf{s}_j^{(k)}$ si chiamerà *vettore di Ritz*.

Per alleggerire la notazione scriveremo ϑ_j al posto di $\vartheta_j^{(k)}$, e faremo lo stesso usando le notazioni \mathbf{s}_j e \mathbf{u}_j .

Possiamo quindi definire il residuo

$$\mathbf{r}_j = A\mathbf{u}_j - \vartheta_j \mathbf{u}_j \quad (3)$$

Il nostro obiettivo è quello di ridurre il più possibile in norma il residuo \mathbf{r}_j all'aumentare di k (norma 2 che indicheremo con $\|\cdot\|$)
Se la norma di \mathbf{r}_j sarà sufficientemente piccola allora ϑ_j sarà una buona approssimazione ad un autovalore λ relativo ad A , e \mathbf{u}_j sarà una buona approssimazione per un autovettore relativo a λ .

Come migliorare \mathbf{r}_j ?

È naturale pensare di ampliare il nostro spazio di ricerca.

Un modo è il seguente:

Definendo D_A la parte diagonale di A Davidson propone il seguente modo:

- ▶ Si calcola un vettore \mathbf{t} che risolve il seguente sistema lineare

$$(D_A - \vartheta_j I)\mathbf{t} = \mathbf{r}_j \quad (4)$$

- ▶ Si ortogonalizza \mathbf{t} rispetto ai precedenti vettori $\mathbf{v}_1, \dots, \mathbf{v}_k$
- ▶ Si normalizza \mathbf{t}
- ▶ Si aggiunge alle colonne di V_k ottenendo V_{k+1}

Correction equation

Un altro modo proposto per ampliare lo spazio di ricerca fa uso di un metodo proposto da Jacobi. Tale metodo è il seguente:

- ▶ Sia λ un autovalore relativo ad A , sia \mathbf{x} un relativo autovettore.
- ▶ Sia \mathbf{u}_j un'approssimazione all'autovettore \mathbf{x} , e supponiamo che $\|\mathbf{u}_j\| = 1$
- ▶ Cerchiamo \mathbf{t} , con $\mathbf{t} \perp \mathbf{u}_j$ tale che

$$A(\mathbf{u}_j + \mathbf{t}) = \lambda(\mathbf{u}_j + \mathbf{t}) \quad (5)$$

Suddividiamo questa equazione nelle due seguenti equazioni:

$$\text{Parte parallela : } \mathbf{u}_j \mathbf{u}_j^* A (\mathbf{u}_j + \mathbf{t}) = \lambda \mathbf{u}_j \mathbf{u}_j^* (\mathbf{u}_j + \mathbf{t}) \quad (6)$$

$$\text{Parte ortogonale : } (I - \mathbf{u}_j \mathbf{u}_j^*) A (\mathbf{u}_j + \mathbf{t}) = \lambda (I - \mathbf{u}_j \mathbf{u}_j^*) (\mathbf{u}_j + \mathbf{t}) \quad (7)$$

- ▶ La parte parallela si trasformerà nella seguente equazione scalare:

$$\vartheta_j + \mathbf{u}_j^* A \mathbf{t} = \lambda \quad (8)$$

Dove $\vartheta_j = \mathbf{u}_j^* A \mathbf{u}_j$.

- ▶ La parte ortogonale si trasformerà nella seguente equazione:

$$(I - \mathbf{u}_j \mathbf{u}_j^*) (A - \lambda I) \mathbf{t} = -\mathbf{r}_j \quad (9)$$

Poiché deve valere che $\mathbf{t} \perp \mathbf{u}_j$ allora possiamo trasformare l'equazione (9) nell'equazione

$$(I - \mathbf{u}_j \mathbf{u}_j^*)(A - \lambda I)(I - \mathbf{u}_j \mathbf{u}_j^*)\mathbf{t} = -\mathbf{r}_j \quad (10)$$

Poiché λ non lo conosciamo allora lo sostituiamo con ϑ_j , e otteniamo la seguente equazione:

$$\boxed{(I - \mathbf{u}_j \mathbf{u}_j^*)(A - \vartheta_j I)(I - \mathbf{u}_j \mathbf{u}_j^*)\mathbf{t} = -\mathbf{r}_j} \quad (11)$$

L'equazione (11) si chiamerà **Jacobi-Davidson correction equation**.

- ▶ L'equazione (11) viene generalmente risolta con il MINRES oppure con il GMRES. Non è richiesta un'approssimazione accurata della soluzione.
- ▶ Una volta che \mathbf{t} è noto (approssimativamente) poniamo

$$\mathbf{u}_{j+1} = \mathbf{u}_j + \mathbf{t} \quad (12)$$

- ▶ Per via di (8), cioè $\vartheta_j + \mathbf{u}_j^* \mathbf{A} \mathbf{t} = \lambda$ possiamo porre

$$\vartheta_{j+1} = \vartheta_j + \mathbf{u}_j^* \mathbf{A} \mathbf{t} \quad (13)$$

Se A è simmetrica allora ϑ_{j+1} può essere posto come il quoziente di Rayleigh di \mathbf{u}_{j+1} (più consigliabile).

Che ruolo ha l'ortogonalità di \mathbf{t} ?

- ▶ L'articolo non scende molto nei dettagli a riguardo, ma possiamo osservare che se non richiedessimo l'ortogonalità, allora l'equazione che andremmo a risolvere è questa:

$$(I - \mathbf{u}_j \mathbf{u}_j^*)(A - \lambda I)\mathbf{t} = -\mathbf{r}_j \quad (14)$$

Per la quale $-\mathbf{u}_j$ è soluzione. Quindi non si farebbe alcun progresso nella ricerca dell'autovalore.

Cosa succede se \mathbf{t} fosse la soluzione esatta dell'equazione di correzione (11)?

Quello che succederà è che vale

$$(I - \mathbf{u}_j \mathbf{u}_j^*)(A - \vartheta_j I) \mathbf{t} = -\mathbf{r}_j \quad (15)$$

che porta immediatamente a

$$(A - \vartheta_j I) \mathbf{t} - \underbrace{\mathbf{u}_j \mathbf{u}_j^*(A - \vartheta_j I) \mathbf{t}}_{\alpha \in \mathbb{F}} = -\mathbf{r}_j \quad (16)$$

Se ϑ_j non è un autovalore di A possiamo scrivere allora

$$\mathbf{t} = \alpha(A - \vartheta_j I)^{-1} \mathbf{u}_j - (A - \vartheta_j I)^{-1} \mathbf{r}_j \quad (17)$$

Poichè $\mathbf{t} \perp \mathbf{u}_j$ vale allora

$$0 = \alpha \mathbf{u}_j^*(A - \vartheta_j I)^{-1} \mathbf{u}_j - \mathbf{u}_j^*(A - \vartheta_j I)^{-1} \mathbf{r}_j \quad (18)$$

Dalla quale

$$\alpha = \frac{\mathbf{u}_j^*(A - \vartheta_j I)^{-1} \mathbf{r}_j}{\mathbf{u}_j^*(A - \vartheta_j I)^{-1} \mathbf{u}_j} \quad (19)$$

Ponendo come abbiamo detto in precedenza $\mathbf{u}_{j+1} = \mathbf{u}_j + \mathbf{t}$ allora otteniamo

$$\mathbf{u}_{j+1} = \mathbf{u}_j + \mathbf{t} = \mathbf{u}_j + \alpha(A - \vartheta_j I)^{-1} \mathbf{u}_j - \underbrace{(A - \vartheta_j I)^{-1} \mathbf{r}_j}_{\mathbf{u}_j} = \alpha(A - \vartheta_j I)^{-1} \mathbf{u}_j$$

cioè

$$\mathbf{u}_{j+1} = \alpha(A - \vartheta_j I)^{-1} \mathbf{u}_j$$

che è un passo del quotient Rayleigh iteration. Quindi la convergenza è quadratica e, nel caso Hermitiano, è cubica.

Precondizionamento: Esistono anche tecniche di preconditionamento per l'equazione di correzione (11). Si usano preconditionatori della forma $\tilde{K} = (I - \mathbf{u}\mathbf{u}^*)K(I - \mathbf{u}\mathbf{u}^*)$, dove $K \approx A - \vartheta_j I$ ed è facilmente invertibile. I dettagli a riguardo sono scritti in [Appendice](#).

Algoritmo per l'autovalore più vicino a τ

Se vogliamo calcolare l'autovalore di A più vicino a un valore fissato τ si può procedere nel seguente modo:

- ▶ Fissiamo un vettore iniziale \mathbf{t} di norma 1 e $V := [\mathbf{t}]$
- ▶ Fissiamo un valore *tol* ed entriamo in un loop nel quale ad ogni iterazione:
 - ▶ Calcoliamo l'autovalore più vicino a τ di V^*AV , sia esso ϑ e sia \mathbf{s} il relativo autovettore con norma 1.
 - ▶ Calcoliamo i vettori $\mathbf{u} = V\mathbf{s}$, e il residuo $\mathbf{r} = A\mathbf{u} - \vartheta\mathbf{u}$
 - ▶ Risolviamo poi l'equazione di correzione (11)
 - ▶ Trovato \mathbf{t} soluzione (approssimata) dell'equazione lo ortogonalizziamo rispetto alle colonne di V , poi lo normalizziamo. (**Oss.** In questo modo \mathbf{t} sarà senza dubbio $\perp \mathbf{u}$.)
 - ▶ Assegnamo $V := [V, \mathbf{t}]$
- ▶ Usciamo dal loop quando $\|\mathbf{r}\| \leq \text{tol}$
- ▶ Fissiamo $\lambda = \vartheta$ e $\mathbf{x} = \mathbf{u}$

Algoritmo per la ricerca dei p autovalori più vicini a τ

Se vogliamo calcolare i primi p autovalori più vicini a τ , allora possiamo procedere così:

supponiamo di aver trovato i primi k autovalori e di avere effettuato una decomposizione di Schur parziale $AQ_k = Q_k T_k$, dove Q_k avrà le colonne ortonormali ($n \times k$) e T_k sarà triangolare superiore ($k \times k$).

Effettueremo lo stesso algoritmo precedente dove però la Correction equation sarà la seguente:

$$(I - \mathbf{u}^* \mathbf{u})(I - Q_k Q_k^*)(A - \vartheta I)(I - Q_k Q_k^*)(I - \mathbf{u}^* \mathbf{u})\mathbf{t} = -\mathbf{r} \quad (20)$$

Dove

- ▶ \mathbf{t} deve essere ortogonale a \mathbf{u} e a Q_k
- ▶ $\mathbf{r} = (I - Q_k Q_k^*)(A - \vartheta I)(I - Q_k Q_k^*)\mathbf{u}$

Deve anche valere che $\mathbf{u} \perp Q_k$, per questo motivo il vettore iniziale lo sceglieremo ortogonale a Q_k , così anche V lo sarà, e di conseguenza lo sarà \mathbf{u} .

Una volta trovata la coppia $(\lambda_{k+1}, \mathbf{x}_{k+1})$ Q_{k+1} sarà $[Q_k, \mathbf{x}_{k+1}]$,
mentre T_{k+1} sarà

$$\begin{bmatrix} T_k & Q_k^* A \mathbf{x}_{k+1} \\ & \lambda_{k+1} \end{bmatrix}$$

Precondizionatori: Esistono preconditionatori anche per l'equazione di correzione (20). Si useranno preconditionatori della forma

$$\tilde{K} = (I - \tilde{Q}_k \tilde{Q}_k^*) K (I - \tilde{Q}_k \tilde{Q}_k^*)$$

dove $\tilde{Q}_k = [Q_k, \mathbf{u}_j]$.

Per i dettagli guardare in [Appendice](#).

Spectral shifts Nell'equazione di correzione (11) compare ϑ_j (spectral shift). L'articolo propone, per via di risultati sperimentali, di sostituire ϑ_j nell'equazione (11) con un valore fissato che solitamente sarà τ , cioè il target. Quando \mathbf{r} sarà abbastanza piccolo si inserisce di nuovo ϑ_j . La motivazione di ciò sta nel fatto che inizialmente ϑ potrebbe essere molto lontano da τ nelle prime iterazioni, portando ad autovalori indesiderati.

Restarts Esistono tecniche di restarting per l'algoritmo proposto affinché la dimensione di V non diventi consistente. Per saperne di più guardare in [Appendice](#).

Problema agli autovalori generalizzati

Consideriamo il problema

$$A\mathbf{x} = \lambda M\mathbf{x} \quad (21)$$

Dove

- ▶ A è hermitiana
- ▶ M è hermitiana e definita positiva

Nel cercare l'autovalore più vicino a τ l'approccio che seguiremo sarà molto simile a quello esposto finora.

Il nostro spazio di ricerca V sarà fatto però di colonne

M -ortonormali, dunque i passaggi da effettuare saranno analoghi a quelli dell'algoritmo precedente, però con delle leggere modifiche.

Infatti:

- ▶ Come prima si risolve il problema agli autovalori standard $V_m^* A V_m \mathbf{s} = \vartheta \mathbf{s}$, $\|\mathbf{s}\| = 1$ (si cercherà l'autovalore più vicino a τ)
- ▶ Si costruiscono $\mathbf{u} = V_m \mathbf{s}$ ($\|\mathbf{u}\|_M = 1$) e il relativo residuo $\mathbf{r} = A\mathbf{u} - \vartheta M\mathbf{u}$
- ▶ Si risolve la Correction equation che in questo caso sarà

$$(I - M\mathbf{u}\mathbf{u}^*)(A - \vartheta M)(I - \mathbf{u}\mathbf{u}^*M)\mathbf{t} = -\mathbf{r}, \quad \mathbf{t} \perp_M \mathbf{u} \quad (22)$$

- ▶ Il \mathbf{t} trovato si M-ortogonalizza rispetto a V , si M-normalizza e poi si aggiunge a V

Precondizionatori: Si usano preconditionatori della forma

$$\tilde{K} = (I - M\mathbf{u}\mathbf{u}^*)K(I - M\mathbf{u}\mathbf{u}^*)$$

dove $K \approx A - \tau M$ e τ è il target value.

Algoritmo di J-D per autovalori interni

La procedura Rayleigh-Ritz può creare dei problemi se applicata alla ricerca degli autovalori interni.

Mostriamo il seguente esempio numerico:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad V = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{0.5} \\ 0 & \sqrt{0.5} \end{bmatrix}$$

V^*AV ci dà la matrice nulla.

Quindi ogni combinazione lineare delle colonne di V è un vettore di Ritz, e ovviamente 0 è un valore di Ritz.

Dunque anche

$$V \begin{bmatrix} \sqrt{0.5} \\ \sqrt{0.5} \end{bmatrix} = \begin{bmatrix} \sqrt{0.5} \\ 0.5 \\ 0.5 \end{bmatrix}$$

è un vettore di Ritz.

0 però è autovalore per A , e tale vettore è una pessima approssimazione.

Harmonic Ritz values and vectors

Un'alternativa per cercare gli autovalori più vicini a uno shift σ e i suoi relativi autovettori è la seguente:

- ▶ Sia $\text{span}(U)$ lo spazio di ricerca costruito relativo a $A - \sigma I$.
- ▶ Si effettua la ricerca su $(A - \sigma I)^{-1}$ degli autovalori di modulo più grande con i relativi autovettori. (Utilizzando uno spazio di ricerca $\text{span}(V)$)
- ▶ Si usa il solito approccio

$$V^*(A - \sigma I)^{-1}Vs = \mu V^*Vs \quad (23)$$

dove μ è autovalore generalizzato di modulo massimo.

- ▶ Come scegliere V ? Poniamo $V = (A - \sigma I)U$.
- ▶ Se poniamo $\tau = 1/\mu$ allora l'equazione diventa

$$V^*Vs = \tau V^*Us \quad (24)$$

- ▶ Si cercherà il τ più piccolo.
- ▶ Sia (τ, s) la coppia ottenuta: chiameremo allora **coppia Armonica di Ritz** la coppia $(\tau + \sigma, Us)$.

Nell'algoritmo \mathbf{u} sarà $U\mathbf{s}$ e ϑ sarà $\tau + \sigma$ (nel caso simmetrico si porrà $\vartheta = \rho(\mathbf{u})$).

Nel caso non simmetrico l'equazione diventa

$$\tilde{V}^* V\mathbf{s} = \tau \tilde{V}^* U\mathbf{s} \quad (25)$$

dove \tilde{V} è una base ortonormale di $\text{span}(V)$.

Refined Ritz vectors

Sia ϑ un valore di Ritz ristretto a un sottospazio V .
Cerchiamo di risolvere il seguente problema:

$$\min_{\mathbf{x} \in V, \|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x} - \vartheta\mathbf{x}\| \quad (26)$$

Una soluzione di questo problema è chiamata **Refined Ritz vector**.
Poichè \mathbf{x} può essere scritto come $\mathbf{x} = V\mathbf{z}$ allora il problema diventa

$$\min_{\|\mathbf{z}\|=1} \|(A - \vartheta I)V\mathbf{z}\| \quad (27)$$

Una soluzione viene data dal vettore singolare destro relativo al valore singolare più piccolo di $(A - \vartheta I)V$. quindi nell'algorithmo i passaggi che sostituiamo ,dopo aver trovato il valore di Ritz più piccolo relativo a V sono i seguenti:

- ▶ Si cerca il vettore singolare destro \mathbf{s} di $AV - \vartheta V$ relativo al valore singolare più piccolo.
- ▶ Si pone $\mathbf{u} = V\mathbf{s}$

Decomposizione di Schur generalizzata

Teorema (Decomposizione di Schur generalizzata) Se $A, B \in \mathbb{C}^{n \times n}$ allora esistono due matrici unitarie $Q, Z \in \mathbb{C}^{n \times n}$ tali che

$$Q^*AZ = T^A, \quad Q^*BZ = T^B \quad (28)$$

sono triangolari superiori. Se per qualche k $t_{kk}^A = t_{kk}^B = 0$ allora $\sigma(A, B) = \mathbb{C}$. Altrimenti

$$\sigma(A, B) = \{t_{ii}^A/t_{ii}^B \mid t_{ii}^B \neq 0\}$$

Ottenere una decomposizione QZ parziale con J-D

Ci chiediamo come usare l'algoritmo di Jacobi-Davidson per risolvere il problema

$$A\mathbf{x} = \lambda B\mathbf{x} \quad (29)$$

con A, B matrici arbitrarie. Sia $\alpha/\beta = \lambda$, allora il problema equivalente sarà

$$(\beta A - \alpha B)\mathbf{x} = \mathbf{0} \quad (30)$$

Una soluzione $\{(\alpha, \beta), \mathbf{x}\}$ la chiameremo *autocoppia* per $\{A, B\}$.

Definiamo **decomposizione di Schur parziale** per la coppia $\{A, B\}$ la seguente decomposizione:

$$AQ_k = Z_k R_k^A, \quad BQ_k = Z_k R_k^B \quad (31)$$

Dove Q_k è $n \times k$ unitaria e R_k^A, R_k^B triangolari superiori $k \times k$.

Valgono le seguenti relazioni:

- ▶ Se $((\alpha, \beta), \mathbf{y})$ sono un'autocoppia per $\{R_k^A, R_k^B\}$ allora $((\alpha, \beta), Q_k \mathbf{y})$ lo è per $\{A, B\}$
- ▶ Sia $\mathbf{q}_i = Q_k \mathbf{e}_i$, allora vale che $\beta_i A \mathbf{q}_i - \alpha_i B \mathbf{q}_i \perp \mathbf{z}_i$,
dove $(\alpha_i, \beta_i) = (R_k^A(i, i), R_k^B(i, i))$
 - ▶ Ciò ci ricorda la decomposizione di Schur, dove $A \mathbf{q}_i - \lambda_i \mathbf{q}_i \perp \mathbf{q}_i$

L'approccio che viene usato allora è il seguente (Petrov-Galerkin):

- ▶ Ad ogni passo:
 - ▶ L'approssimazione \mathbf{u} dell'autovettore apparterrà a $\text{span}(V_j)$ dove $\text{span}(V_j)$ sarà il nostro solito spazio di ricerca.
 - ▶ Cercheremo due valori η e ζ tali che $\eta \mathbf{A}\mathbf{u} - \zeta \mathbf{B}\mathbf{u} \perp \text{span}(W_j)$ per un ben selezionato W_j (in seguito vedremo come)
 - ▶ Giungiamo al seguente problema agli autovalori generalizzato di dimensione j :

$$\eta \underbrace{W_j^* A V_j}_{C_j^A} \mathbf{s} - \zeta \underbrace{W_j^* B V_j}_{C_j^B} \mathbf{s} = \mathbf{0} \quad (32)$$

Ci chiediamo come scegliere W_j e come trovare η, ζ

Sia

$$(S^L)^* \underbrace{(W_j^* A V_j)}_{C_j^A} S^R = T^A \quad (S^L)^* \underbrace{(W_j^* B V_j)}_{C_j^B} S^R = T^B \quad (33)$$

una decomposizione di Schur generalizzata.

Tale decomposizione può essere riordinata in modo tale che $S^R \mathbf{e}_1$ e $[T^A]_{1,1}$ e $[T^B]_{1,1}$ siano la soluzione cercata (Si chiamerà soluzione di Petrov, l'articolo non dà dettagli a riguardo al riordinamento).

Quindi:

- ▶ $\zeta := [T^A]_{1,1}$ e $\eta := [T^B]_{1,1}$
- ▶ Si pone $\mathbf{u} := V_j \mathbf{s} = V_j S^R \mathbf{e}_1$
- ▶ In modo analogo si pone $\mathbf{p} := W_j S^L \mathbf{e}_1$ (definito *vettore sinistro di Petrov*)

Come scegliere W_j ?

$V_j S^R$ approssima Q_j , $W_j S^L$ approssima Z_j .

Vale che $\text{span}(AQ_j) = \text{span}(BQ_j) = \text{span}(Z_j)$.

Quindi si procederà nel seguente modo:

- ▶ Si sceglie W_j in modo che $\text{span}(W_j) = \text{span}(\nu_0 AV_j + \mu_0 BV_j)$
dove $\nu_0 = 1/\sqrt{1 + |\tau|^2}$ e $\mu_0 = -\tau\nu_0$
- ▶ La Correction equation che andremo a risolvere è la seguente:

$$(I - \mathbf{p}\mathbf{p}^*)(\eta A - \zeta B)(I - \mathbf{u}\mathbf{u}^*)\mathbf{t} = -\mathbf{r}, \quad \mathbf{t} \perp \mathbf{u}, \quad \mathbf{r} := \eta A\mathbf{u} - \zeta B\mathbf{u}$$

- ▶ $\mathbf{v} = \mathbf{t}/\|\mathbf{t}\|$ sarà aggiunto a V_j , mentre $\mathbf{w} := \nu_0 A\mathbf{v} + \mu_0 B\mathbf{v}$ sarà aggiunto a W_j .
- ▶ Entrambi gli spazi saranno fatti di colonne ortonormali.

Deflazione

Supponiamo di avere una decomposizione di Schur parziale $\{(Q_{k-1}, Z_{k-1}), (T_{k-1}^A, T_{k-1}^B)\}$, e vogliamo espanderla ad una decomposizione $\{(Q_k, Z_k), (T_k^A, T_k^B)\}$ dove

$$Q_k = [Q_{k-1}, \mathbf{u}], \quad Z_k = [Z_{k-1}, \mathbf{p}], \quad T_k^A = \begin{bmatrix} T_{k-1}^A & \mathbf{a} \\ & \alpha \end{bmatrix}, \quad T_k^B = \begin{bmatrix} T_{k-1}^B & \mathbf{b} \\ & \beta \end{bmatrix}$$

È facile vedere che

$$Q_{k-1}^* \mathbf{u} = 0, \quad (\beta A - \alpha B) \mathbf{u} - Z_{k-1} (\beta \mathbf{a} - \alpha \mathbf{b}) = 0 \quad (34)$$

Questo si trasforma nell'equazione

$$Q_{k-1}^* \mathbf{u} = 0, \quad (I - Z_{k-1} Z_{k-1}^*) (\beta A - \alpha B) (I - Q_{k-1} Q_{k-1}^*) \mathbf{u} = 0 \quad (35)$$

Diventa quindi una ricerca di una coppia $((\alpha, \beta), \mathbf{u})$ per la coppia (M^A, M^B) dove

$$\blacktriangleright M^A = (I - Z_{k-1}Z_{k-1}^*)A(I - Q_{k-1}Q_{k-1}^*)$$

$$\blacktriangleright M^B = (I - Z_{k-1}Z_{k-1}^*)B(I - Q_{k-1}Q_{k-1}^*)$$

Questo problema può essere risolto con il processo QZ di Jacobi-Davidson sopra descritto. \mathbf{u} e \mathbf{p} saranno rispettivamente i vettori di Schur destro e sinistro. Una volta trovato \mathbf{u} possiamo trovarci anche \mathbf{a} e \mathbf{b} , ovvero

$$\mathbf{a} = Z_{k-1}^*A\mathbf{u}, \quad \mathbf{b} = Z_{k-1}^*B\mathbf{u} \quad (36)$$

Esperimenti numerici

Mostriamo adesso alcuni esperimenti numerici. Le matrici saranno reali di dimensione n , simmetriche e sparse, con densità

$1/n, 5/n, 10/n$.

$n = 10^3, 10^4, 10^5$

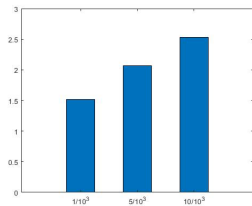
Abbiamo cercato i primi p autovalori più piccoli (secondo la relazione d'ordine standard su \mathbb{R}).

Le matrici le abbiamo generate con il comando

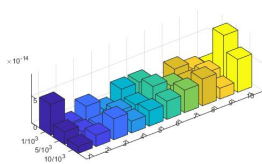
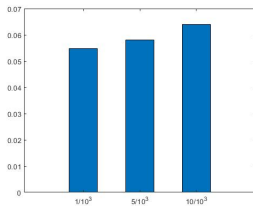
```
A=sprand(dim,0.5*density);A=A+A';
```

Mostriamo ora i grafici dei tempi di esecuzione e dei residui con Jacobi-Davidson (a sinistra) e con Matlab (a destra).

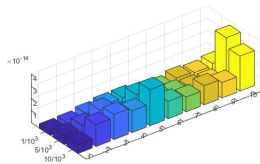
$$n = 10^3, p = 10$$



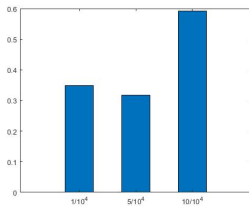
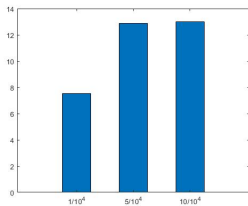
Tempi



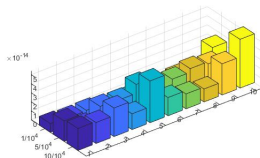
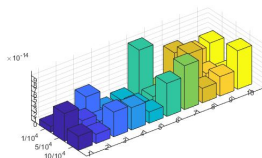
Residui



$$n = 10^4, p = 10$$

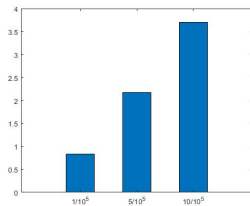
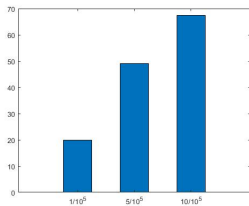


Tempi

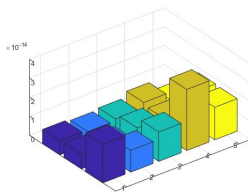
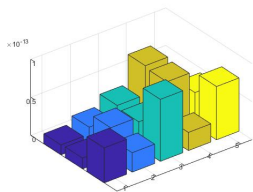


Residui

$$n = 10^5, p = 5$$



Tempi



Residui

Appendice

Per l'equazione (11) dei preconditionatori suggeriti sono della forma

$$\tilde{K} = (I - \mathbf{u}_j \mathbf{u}_j^*) K (I - \mathbf{u}_j \mathbf{u}_j^*), \quad K \approx A - \vartheta_j I \quad (37)$$

assumendo che K sia facilmente invertibile.

Supponiamo di avere un sistema di equazioni

$$\tilde{K} \mathbf{x} = \mathbf{b} \quad (38)$$

con $\mathbf{x} \perp \mathbf{u}_j$ e $\mathbf{b} \perp \mathbf{u}_j$. Chiameremo formalmente

$$\mathbf{x} = \tilde{K}^+ \mathbf{b} \quad (39)$$

la soluzione del sistema.

Invece di andare a risolvere l'equazione (11) andiamo a risolvere l'equazione

$$\tilde{K}^+ \tilde{A} \mathbf{t} = -\tilde{K}^+ \mathbf{r}_j \quad (40)$$

dove

$$\tilde{A} = (I - \mathbf{u}_j \mathbf{u}_j^*) (A - \vartheta_j I) (I - \mathbf{u}_j \mathbf{u}_j^*) \quad (41)$$

Possiamo risolvere il sistema (38) nel seguente modo:

Poiché deve valere che $\mathbf{x} \perp \mathbf{u}_j$ allora il sistema diventa

$$(I - \mathbf{u}_j \mathbf{u}_j^*) K \mathbf{x} = \mathbf{b} \quad (42)$$

Denotiamo

$$\alpha = -\mathbf{u}_j^* K \mathbf{x} \quad (43)$$

È facile vedere che la soluzione dell'equazione (42) è data da

$$\mathbf{x} = K^{-1} \mathbf{b} - \alpha K^{-1} \mathbf{u}_j \quad (44)$$

Con la condizione che $\mathbf{x} \perp \mathbf{u}_j$ allora

$$\alpha = \frac{\mathbf{u}_j K^{-1} \mathbf{y}}{\mathbf{u}_j K^{-1} \mathbf{u}_j} \quad (45)$$

Scegliendo come vettore iniziale $\mathbf{t}_0 = \mathbf{0}$ possiamo risolvere il sistema (40), infatti:

- ▶ Sappiamo calcolare $\tilde{K}^+ \mathbf{r}_j$ (che sarà ancora ortogonale a \mathbf{u}_j)
- ▶ Sappiamo calcolare i prodotti $\tilde{K}^+ A \mathbf{v}$ dove $\mathbf{v} \perp \mathbf{u}_j$

Spectral shifts

Come già detto nelle slides anche nel precodizzatore vanno posti gli stessi shift posti per A . Lo spectral shift di K è bene fissarlo affinché K^{-1} venga calcolato una volta sola.

Precondizionatore dopo aver calcolato k autovalori

Per il calcolo di più autovalori i preconditionatori scelti sono della forma

$$\tilde{K} = (I - \tilde{Q}_k \tilde{Q}_k^*) K (I - \tilde{Q}_k \tilde{Q}_k^*) \quad (46)$$

dove $\tilde{Q}_k = [Q_k, \mathbf{u}_j]$.

Si procede esattamente come in precedenza, considerando il seguente sistema:

$$\tilde{K} \mathbf{x} = \mathbf{b}, \quad \mathbf{x} \perp \tilde{Q}_k, \quad \mathbf{b} \perp \tilde{Q}_k \quad (47)$$

Come prima possiamo trasformarla nell'equazione

$$(I - \tilde{Q}_k^* \tilde{Q}_k) K \mathbf{x} = \mathbf{b} \quad (48)$$

Ponendo

$$\mathbf{a} = -\tilde{Q}_k^* K \mathbf{x} \quad (49)$$

allora

$$\mathbf{x} = K^{-1} \mathbf{b} - K^{-1} \tilde{Q}_k \mathbf{a} \quad (50)$$

è soluzione dell'equazione.

Sempre con la condizione che $\tilde{Q}_k \perp \mathbf{x}$ allora

$$\mathbf{a} = (\tilde{Q}_k^* K^{-1} \tilde{Q}_k)^{-1} \tilde{Q}_k^* K^{-1} \mathbf{b} \quad (51)$$

Il sistema che andremo a risolvere sarà

$$\tilde{K}^+ \tilde{A} \mathbf{t} = -\tilde{K}^+ \mathbf{r}_j \quad (52)$$

Dove

- ▶ $\tilde{A} = (I - \tilde{Q}_k \tilde{Q}_k^*)(A - \vartheta_j I)(I - \tilde{Q}_k \tilde{Q}_k^*)$
- ▶ $\mathbf{r}_j = (I - Q_k Q_k^*)(A\mathbf{u}_j - \vartheta_j \mathbf{u}_j)(I - Q_k Q_k^*)$

Tecniche di restarting

Affinché si voglia fare in modo che la dimensione dello spazio di ricerca non diventi eccessivo si possono effettuare delle tecniche di Restarting.

Si procede nel seguente modo:

- ▶ Si fissa un valore m_{max}
- ▶ Si fissa un valore q
- ▶ Quando $m = m_{max}$ effettuiamo il Restarting.
- ▶ $V_m = V_{m_{max}}$ viene rimpiazzato dai q vettori di Ritz corrispondenti ai valori di Ritz più vicini al target τ

Come si fa per rimpiazzare?

- ▶ Si può effettuare una decomposizione di Schur di $M = V_m^* A V_m$, sia essa $M = S T S^*$ con $|t_{11} - \tau| \leq |t_{22} - \tau| \leq \dots$
- ▶ Si pone $V_q = V_m S_{:,1:q}$
- ▶ $M := V_q^* A V_q = T_{1:q,1:q}$