
BOLLETTINO UNIONE MATEMATICA ITALIANA

Sezione A – La Matematica nella Società e nella Cultura

GIANNA M. DEL CORSO

Metodi probabilistici per il calcolo di autovalori ed autovettori

*Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. 1-A—La
Matematica nella Società e nella Cultura (1998), n.1S (Supplemento
Tesi di Dottorato), p. 185–188.*

Unione Matematica Italiana

http://www.bdim.eu/item?id=BUMI_1998_8_1A_1S_185_0

L'utilizzo e la stampa di questo documento digitale è consentito liberamente per motivi di ricerca e studio. Non è consentito l'utilizzo dello stesso per motivi commerciali. Tutte le copie di questo documento devono riportare questo avvertimento.

*Articolo digitalizzato nel quadro del programma
bdim (Biblioteca Digitale Italiana di Matematica)
SIMAI & UMI*

<http://www.bdim.eu/>

Metodi probabilistici per il calcolo di autovalori ed autovettori.

GIANNA M. DEL CORSO

1. - Introduzione e notazione.

Sia A una matrice $n \times n$ simmetrica e definita positiva con autovalori $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$. Sia inoltre r , $r < n$, la molteplicità dell'autovalore dominante λ_1 e denotiamo con Z_1 l'autospazio corrispondente, cioè, $Z_1 = \text{span} \{z: Az = \lambda_1 z\}$. Il problema affrontato è quello di stimare l'autovalore λ_1 e un autovettore di Z_1 . Per risolvere questo problema sono generalmente usati metodi iterativi basati sull'informazione di Krylov. Infatti, quando si vuole approssimare solo la coppia autovalore-autovettore dominante, i metodi iterativi sono da preferirsi a metodi basati sulla fattorizzazione della matrice in quanto risultano più convenienti dal punto di vista del costo computazionale. Tra i metodi che utilizzano l'informazione di Krylov quelli maggiormente utilizzati sono il metodo delle potenze ed il metodo di Lanczos. Per garantire la convergenza di questi metodi alle quantità desiderate, il vettore di partenza è generalmente scelto in modo casuale.

Per questi motivi è interessante studiare il comportamento di questi metodi quando il vettore di partenza \mathbf{b} è scelto con distribuzione uniforme sulla sfera unitaria di \mathbb{R}^n e effettuare l'analisi dell' »errore probabilistico«. Questo approccio è stato già seguito da Kuczyński e Woźniakowski [4] [5] per l'approssimazione all'autovalore dominante con il metodo delle potenze ed il metodo di Lanczos e ripreso successivamente da Leyk e Woźniakowski [6]. Nel seguito sarà usata la seguente notazione. Un *metodo polinomiale* è un metodo iterativo che restituisce al passo k -esimo un vettore $\mathbf{u}_k = \mathbf{u}(A, \mathbf{b}, k)$ tale che $\mathbf{u}_k = P(A)\mathbf{b}$ dove $P(\cdot)$ è un polinomio di grado al più $k - 1$ tale che $\|\mathbf{u}_k\| = 1$. Per ogni metodo polinomiale \mathcal{N} , $\xi_k^{\mathcal{N}}$ e $\mathbf{u}_k^{\mathcal{N}}$ denotano rispettivamente l'approssimazione di λ_1 e di un vettore $\mathbf{z} \in Z_1$ restituite da \mathcal{N} dopo k passi. Per un qualsiasi metodo polinomiale \mathcal{N} definiamo l'errore probabilistico nel senso di \mathcal{L}_p , $p \in [1, +\infty]$, per l'approssimazione di un autovettore come

$$(1) \quad e^{\text{prob}}(\mathbf{u}_k^{\mathcal{N}}, A, p) = \left(\int_{S_n} \inf_{\mathbf{v} \in Z_1} \|\mathbf{u}_k^{\mathcal{N}}(\mathbf{b}) - \mathbf{v}\| \mu(db) \right)^{1/p}.$$

Allo stesso modo, l'errore probabilistico nel senso di \mathcal{L}_p per l'approssimazione di λ_1 è

$$(2) \quad e^{\text{prob}}(\xi_k^{\mathcal{N}}, A, p) = \left(\int_{S_n} \left| \frac{\lambda_1 - \xi_k^{\mathcal{N}}(\mathbf{b})}{\lambda_1} \right| \mu(db) \right)^{1/p}.$$

Lo studio dell'errore probabilistico per ogni p enfatizza l'importanza della scelta della misura di errore per la convergenza del metodo. Infatti, il valore di p , ed in particolare la sua relazione con la molteplicità r di λ_1 , influenza la velocità di convergenza dei vari metodi.

2. - Alcuni risultati.

In questa nota possiamo illustrare solo alcuni risultati che saranno dati senza dimostrazione. I dettagli possono essere trovati in [1] [2] e i risultati dei test numerici in [3].

Analizzando il comportamento del metodo delle potenze e del metodo di Lanczos nel setting probabilistico Kuczyński e Woźniakowski [4] [5] hanno dimostrato che per l'approssimazione dell'autovalore massimo è possibile stimare l'errore in modo indipendente dagli autovalori della matrice. Infatti, nel caso di \mathcal{L}_1 è stato dimostrato che $e^{\text{prob}}(\xi_k^{\text{Pow}}, A, 1) \approx a(n)1/k$, dove la quantità $a(n)$ dipende solo dalla dimensione della matrice a cui è applicato il metodo delle potenze e non dai suoi autovalori. Una stima simile è stata trovata anche per il metodo di Lanczos per il quale abbiamo una velocità di convergenza proporzionale a $1/k^2$. In [2] abbiamo esteso queste stime indipendenti dagli autovalori di A al caso di \mathcal{L}_p , mostrando la dipendenza della velocità di convergenza dal parametro p . È importante sottolineare l'importanza di questo risultato che ci permette di affermare che nel setting probabilistico abbiamo sempre la convergenza di questi metodi allorché si consideri l'approssimazione di λ_1 mentre nel caso deterministico possiamo non avere convergenza. Un fatto molto interessante è che nel caso si consideri l'approssimazione di autovettori non è più possibile ottenere stime dell'errore probabilistico indipendenti dagli autovalori della matrice. In particolare, si è dimostrato [1] che per il metodo delle potenze esistono matrici «worst case» per le quali il metodo non risulta convergente nel setting probabilistico quando si considera l'approssimazione di autovettori. Questo risultato negativo è stato generalizzato in [2] dimostrando che per ogni metodo polinomiale basato sull'informazione di Krylov esistono matrici i cui autovalori sono così vicini tra loro che il metodo fallisce nell'approssimare un autovettore relativo a λ_1 a meno che non vengano effettuati n passi iterativi. Vale infatti il seguente

TEOREMA 1. - Sia \mathcal{C}_n la classe delle matrici $n \times n$ definite positive ed \mathcal{U}_k la classe dei metodi polinomiali che fanno k passi, allora, per $k < n$

$$(3) \quad \sup_{A \in \mathcal{C}_n} \inf_{\mathcal{M} \in \mathcal{U}_k} e^{\text{prob}}(\mathbf{u}_k^{\mathcal{M}}, A, p) = \left(\frac{\Gamma((n-k+1)/2) \Gamma((n+p-k)/2)}{\Gamma((n-k)/2) \Gamma((n+p-k+1)/2)} \right)^{1/p}.$$

Per $k \geq n$, $e^{\text{prob}}(\mathbf{u}_k^{\mathcal{M}}, A, p) = 0$.

Notiamo che la quantità (3) è decrescente con k , ma per $k = n - 1$ abbiamo una costante che non va a zero neppure per $n \rightarrow \infty$. Questo prova che nel setting probabilistico il problema della stima di un autovettore è intrinsecamente più difficile di quello dell'approssimazione di un autovalore.

La mancanza di stime indipendenti dalla distribuzione di autovalori ci ha fatto ricercare delle limitazioni dell'errore probabilistico che mostrassero la dipendenza dai λ_i . Per il metodo delle potenze abbiamo le seguenti limitazioni superiori ed inferiori.

TEOREMA 2. - Siano $x_i = \lambda_i/\lambda_1$, $i = 1, 2, \dots, n$. Per ogni $p \in [1, \infty)$, e per ogni

k abbiamo

$$e^{\text{prob}}(\mathbf{u}_k^{\text{Pow}}, A, p) \leq \begin{cases} \alpha(p, r, n-r) x_{r+1}^k & \text{per } p < r, \\ \beta(p, r, n-r) x_{r+1}^k k^{1/p} \ln(1/x_{r+1}) & \text{per } p = r, \\ \gamma(p, r, n-r) x_{r+1}^{kr/p} & \text{per } p > r, \end{cases}$$

e

$$e^{\text{prob}}(\mathbf{u}_k^{\text{Pow}}, A, p) \geq \begin{cases} \alpha(p, r, 1) x_{r+1}^k & \text{per } p < r, \\ \beta(p, r, 1) x_{r+1}^k k^{1/p} \ln(1/x_{r+1}) & \text{per } p = r, \\ \gamma(p, r, 1) x_{r+1}^{kr/p} & \text{per } p > r, \end{cases}$$

dove le quantità α, β, γ sono costanti dipendenti dalla dimensione della matrice e dai parametri p e r (si veda [1] per maggiori dettagli sul valore delle costanti).

In [1] abbiamo dimostrato che le limitazioni date nel Teorema 2 sono asintoticamente ottimali in quanto le costanti moltiplicative dipendono dalla molteplicità del secondo autovalore. Quando questa è $n - r$ troviamo le costanti date nell'upper bound; le costanti del limite inferiori sono ottenute quando la molteplicità del secondo autovalore è invece 1. Per quanto riguarda il metodo di Lanczos abbiamo dato solo limitazioni superiori [2] in quanto il problema è espresso come problema di massimo e risulta particolarmente difficile trovare delle stime inferiori dell'errore probabilistico (si noti tra l'altro che queste non esistono neppure nel caso deterministico).

TEOREMA 3. - Sia m il numero di autovalori distinti di A , per $k \geq m$, abbiamo $e^{\text{prob}}(\mathbf{u}_k^{\text{Lan}}, A, p) = 0$ e per $k < m$

$$e^{\text{prob}}(\mathbf{u}_k^{\text{Lan}}, A, p) \leq \begin{cases} \alpha'(p, r) f_k(\lambda_1, \lambda_{r+1}, \lambda_n) & \text{per } p < r, \\ \beta'(p, r) f_k(\lambda_1, \lambda_{r+1}, \lambda_n) \ln^{1/p}(1/f_k(\lambda_1, \lambda_{r+1}, \lambda_n)) & \text{per } p = r, \\ \gamma'(p, r) [f_k(\lambda_1, \lambda_{r+1}, \lambda_n)]^{r/p} & \text{per } p > r, \end{cases}$$

dove

$$f_k(\lambda_1, \lambda_{r+1}, \lambda_n) = \sqrt{\frac{\lambda_1 - \lambda_n}{\lambda_1 - \lambda_{r+1}}} \frac{1}{\left| T_k \left(1 + 2 \frac{\lambda_1 - \lambda_{r+1}}{\lambda_{r+1} - \lambda_n} \right) \right|},$$

T_k è il polinomio di Chebyshev di prima specie di grado $k - 1$, α', β' e γ' sono costanti indipendenti dagli autovalori di A (per il valore delle costanti si veda [2]).

Anche per l'approssimazione di autovalori con il metodo delle potenze ed il metodo di Lanczos abbiamo dato stime dell'errore simili a quelle date nei Teoremi 2 e 3. Si osservi che la velocità di convergenza di questi metodi dipende dagli autovalori dominanti nonché dai valori di p e r e che solo nel caso $p < r$ troviamo la stessa velocità di convergenza che avevamo nel caso deterministico allorché il metodo risultava convergente, ovvero il vettore di partenza \mathbf{b} non era ortogonale all'autospazio \mathbf{Z}_1 . Negli altri due casi, $p = r$ e $p > r$ abbiamo una convergenza più lenta. Questo è dovuto al fatto che nei Teoremi 2 e 3 è analizzato l'errore probabi-

listico e quindi occorre integrare su tutti i possibili vettori di partenza come sottolineato dalle equazioni (1) e (2), anche per i vettori per i quali i metodi sono non convergenti o hanno una velocità di convergenza molto bassa.

3. – Conclusioni.

In questa nota abbiamo presentato alcuni dei risultati ottenuti nella tesi di dottorato nella quale è stato analizzato il comportamento dell'errore probabilistico nel senso di \mathcal{L}_p di metodi polinomiali per il calcolo dell'autovalore dominante e di un suo autovettore. Un risultato generale è che l'errore probabilistico dipende dalla molteplicità r dell'autovalore dominante e dal parametro della norma p . Più precisamente, sembra che la relazione tra questi due parametri porti ad avere tre differenti velocità di convergenza. Per il calcolo di autovettori ed autovalori con il metodo delle potenze abbiamo limitazioni superiori ed inferiori strette che mostrano come l'esistenza di tre differenti velocità di convergenza sia intrinseca nel problema. Per il metodo di Lanczos abbiamo solo limitazioni superiori dell'errore probabilistico, quindi non possiamo affermare che anche per questo metodo i parametri p ed r sono così importanti, anche se i risultati sperimentali [3] mostrano un rallentamento della convergenza al crescere di p . I risultati ottenuti sono comunque molto incoraggianti e suggeriscono che sarebbe utile estendere tale analisi anche ad altri metodi, come al metodo di Arnoldi, per il calcolo di autovalori ed autovettori di matrici non simmetriche.

BIBLIOGRAFIA

- [1] DEL CORSO G.M., *Estimating an eigenvector by the power method with a random start*, SIAM J. Matrix Anal. Appl., 18(4) (1997), 913-937.
- [2] DEL CORSO G.M. e MANZINI G., *On the randomized error of polynomial methods for eigenvector and eigenvalue estimate*, Journal of Complexity, 13(4) (1997), 419-456.
- [3] DEL CORSO G.M. e MANZINI G., *On the randomized error of polynomial methods for eigenvector and eigenvalue estimate: Numerical tests*, Technical Report B4-97-01, Istituto di Matematica Computazionale, CNR, Pisa (1997).
- [4] KUCZYŃSKI J. e WOŹNIAKOWSKI H., *Estimating the largest eigenvalue by the power and Lanczos algorithms with a random start*, SIAM J. Matrix Anal. Appl., 13 (1992), 1094-1122.
- [5] J. KUCZYŃSKI e WOŹNIAKOWSKI H., *Probabilistic Bounds on the Extremal Eigenvalues and Condition Number by the Lanczos Algorithm*, SIAM J. Matrix Anal. Appl., 15 (1994), 672-691.
- [6] LEYK Z. e H. WOŹNIAKOWSKI, *Estimating a largest eigenvector by polynomial algorithms with a random start*, Technical report CUCS-026-96, Columbia University, New York, (1996).

IEI-CNR, Via Santa Maria 46 - 56126 Pisa

e-mail: delcorso@imc.pi.cnr.it

Dottorato in Matematica computazionale e ricerca operativa

(sede amministrativa: Milano) - Cielo IX

Direttori di ricerca: Bruno Codenotti e Henryk WOŹNIAKOWSKI