
ATTI ACCADEMIA NAZIONALE DEI LINCEI
CLASSE SCIENZE FISICHE MATEMATICHE NATURALI

RENDICONTI

PIERCARLO FANTUCCI, STEFANO POLEZZO

Su un programma di calcolo per la determinazione di alcuni autovalori di una matrice simmetrica secondo il metodo di Davidson

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, Vol. 63 (1977), n.5, p. 408–412.
Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1977_8_63_5_408_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/>

Chimica. — *Su un programma di calcolo per la determinazione di alcuni autovalori di una matrice simmetrica secondo il metodo di Davidson* (*). Nota di PIERCARLO FANTUCCI e STEFANO POLEZZO, presentata (**) dal Socio G. SEMERANO.

SUMMARY. — A computer implementation of E. R. Davidson's algorithm for iterative evaluation of the few lowest eigenvalues and eigenvectors of large (symmetric) matrices is presented. The performance of the method is tested on a 61×61 matrix with satisfactory results regarding both convergence rate and numerical precision. Comparisons are made with results obtained by Givens and Nesbet-deflation methods.

INTRODUZIONE

In meccanica quantistica molecolare e in particolare nei calcoli di interazione fra configurazioni (CI) si presenta il problema di calcolare alcuni degli autovalori più bassi, con i relativi autovettori associati, di matrici reali e simmetriche.

Poiché le dimensioni delle matrici CI sono molto grandi (oltre 1000×1000) non è conveniente usare metodi di diagonalizzazione che prevedano la trasformazione di similitudine della matrice di partenza (Jacobi, Householder e Givens, ecc.) dato l'elevatissimo numero di moltiplicazioni da eseguire. Si ravvisa pertanto la necessità di disporre di un procedimento che possieda le caratteristiche: *a*) di calcolare direttamente solo gli autovalori e gli autovettori desiderati, *b*) di eseguire solo prodotti del tipo

$$\mathbf{A}\mathbf{c} \quad \text{e} \quad \bar{\mathbf{c}}\mathbf{c}$$

dove \mathbf{A} è una matrice simmetrica e \mathbf{c} un vettore colonna, *c*) di non richiedere una conoscenza precisa degli autovettori e autovalori che precedono quello che si vuol calcolare.

I principali metodi attualmente in uso sono quello denominato MOR [1] che ricorre allo spostamento delle radici con ottimizzazione del rilassamento; quello di Nesbet [2] per la valutazione dell'autovalore più basso, accoppiato, per il caso di autovalori più alti, con una delle tecniche di rilassamento [6]; quello di Lanczos [3], di Feler [4], di Coope e Sabo [8] ed infine quello di Davidson [5].

(*) Lavoro eseguito con il contributo dell'Accademia Nazionale dei Lincei (Centro Linceo Interdisciplinare di Scienze Matematiche e loro applicazioni), nell'Istituto di Chimica Generale e nell'Istituto di Chimica Fisica dell'Università di Milano.

Gli Autori ringraziano l'Accademia Nazionale dei Lincei per il contributo finanziario.

(**) Nella seduta del 18 novembre 1977.

Il procedimento che in questa Nota verrà preso in considerazione è quello proposto da Davidson perché possiede tutte le caratteristiche prima auspiccate, e perché la convergenza è garantita in quanto, per raggiungere i valori estremali dei quozienti di Rayleigh, esso fa uso di un algoritmo a gradiente. Inoltre la semplicità del metodo rende facile la sua programmazione, anche se altri procedimenti (come ad esempio [8]) possono forse richiedere una minore occupazione di memoria.

Scopo di questa Nota è presentare brevemente sia il metodo di Davidson nella forma in cui è stato da noi programmato, sia i risultati delle prove numeriche eseguite per saggiare il metodo stesso.

DESCRIZIONE DEL METODO DI CALCOLO

L'esposizione dettagliata della teoria è contenuta nella nota originale di Davidson [5]: qui di seguito ci si limita ad un riassunto concettuale del procedimento.

Se \mathbf{A} è una matrice simmetrica $n \times n$ di cui si vogliono trovare i primi autovalori (ε_i e \mathbf{a}_i denoteranno rispettivamente gli autovalori e gli autovettori di \mathbf{A}), detta \mathbf{C} una matrice $n \times m$ formata da m vettori colonna \mathbf{c}_i ($i = 1, \dots, m$) ortonormali e arbitrari, si può formare la matrice simmetrica $m \times m$

$$\mathbf{B} = \tilde{\mathbf{C}}\mathbf{A}\mathbf{C} \quad (\text{con } \tilde{\mathbf{C}}\mathbf{C} = \mathbf{I}_m)$$

e quindi diagonalizzarla con una trasformazione ortogonale \mathbf{U} (λ_i e \mathbf{b}_i denoteranno rispettivamente gli autovalori e gli autovettori di \mathbf{B}). Allora, detta λ la matrice degli autovalori, si ha

$$\tilde{\mathbf{U}}\mathbf{B}\mathbf{U} = \lambda = \tilde{\mathbf{U}}\tilde{\mathbf{C}}\mathbf{A}\mathbf{C}\mathbf{U};$$

per cui, orlando successivamente la \mathbf{B} , i vettori \mathbf{c}_i possono essere calcolati in modo da approssimare sempre più l'autovalore (e l'autovettore associato) della \mathbf{A} che si vuol determinare (in generale, all'aumentare delle dimensioni di \mathbf{B} si avrà che $\lambda \rightarrow \varepsilon$ e $\mathbf{c}_i \rightarrow \mathbf{a}_i$).

È noto che il quoziente di Rayleigh

$$\frac{\tilde{\mathbf{a}}_i \mathbf{A} \mathbf{a}_i}{\tilde{\mathbf{a}}_i \mathbf{a}_i} = \varepsilon_i$$

gode di proprietà estremali in corrispondenza ad autovalori ed autovettori quando siano verificate le necessarie condizioni di ortonormalità [6]. Scelto allora un vettore $\tilde{\mathbf{a}}_k$ che approssimi il vettore \mathbf{a}_k associato all'autovalore ε_k , il valore di minimo viene raggiunto iterativamente e in modo selettivo usando il vettore gradiente ottenuto variando le componenti di $\tilde{\mathbf{a}}_k$ (nei casi che interessano si tratta di autovalori negativi).

Per quanto riguarda lo schema di calcolo esso può essere riassunto nei punti seguenti:

1. Volendo calcolare il k -esimo autovalore ε_k ed autovettore \mathbf{a}_k della matrice \mathbf{A} occorre costruire innanzitutto un sottospazio sotteso da l vettori ortonormali $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_l$ ($l \geq k$). Questi vengono ottenuti applicando ad \mathbf{A} la teoria delle perturbazioni al primo ordine e quindi ortonormalizzando i vettori ottenuti con il procedimento di Löwdin [7].

2. Si forma la matrice \mathbf{B} di dimensioni $l \times l$ avente elementi definiti da

$$B_{ij} = \bar{\mathbf{c}}_i \mathbf{A} \mathbf{c}_j \quad (i, j = 1, \dots, l).$$

La diagonalizzazione di \mathbf{B} fornisce gli l vettori $\mathbf{b}_i^{(1)}$ e gli autovalori $\lambda_i^{(1)}$. \mathbf{B} è la matrice che successivamente verrà orlata ed l è la sua dimensione iniziale.

3. Detta m ($m \geq l$) la dimensione corrente della matrice \mathbf{B} e scelto il vettore $\mathbf{b}_k^{(m)}$ e l'autovalore $\lambda_k^{(m)}$, si forma il vettore gradiente

$$\mathbf{q}_m = \sum_{i=1}^m b_{ik}^{(m)} (\mathbf{A} \mathbf{c}_i) - \sum_{i=1}^m b_{ik}^{(m)} \lambda_k^{(m)} \mathbf{c}_i$$

ove con $b_{ik}^{(m)}$ si è indicata l' i -esima componente del vettore $\mathbf{b}_k^{(m)}$. Il test di convergenza viene fatto sul valore del modulo $\|\mathbf{q}_m\|$ e sulla variazione dell'autovalore $|\lambda_k^{(m-1)} - \lambda_k^{(m)}|$. Come valori di soglia si sono scelti $\|\mathbf{q}_m\| \leq 2 \cdot 10^{-4}$ e $|\lambda_k^{(m-1)} - \lambda_k^{(m)}| \leq 1 \cdot 10^{-5}$.

4. Si forma il vettore \mathbf{x}_{m+1} di elementi

$$x_{s,m+1} = (\lambda_k^{(m)} - A_{ss})^{-1} q_{sm} \quad (s = 1, \dots, m).$$

5. Si forma il proiettore \mathbf{P} dato dalla matrice $n \times n$

$$\mathbf{P} = \mathbf{I}_n - \sum_{i=1}^m \mathbf{c}_i \bar{\mathbf{c}}_i$$

e si ortonormalizza il vettore \mathbf{x}_{m+1} rispetto ai vettori \mathbf{c}_i ($i = 1, \dots, m$)

$$\mathbf{c}_{m+1} = \mathbf{P} \mathbf{x}_{m+1} / \|\mathbf{P} \mathbf{x}_{m+1}\|.$$

6. Si forma il vettore $\mathbf{A} \mathbf{c}_{m+1}$ e si orla la matrice \mathbf{B} con i nuovi elementi

$$B_{i,m+1} = B_{m+1,i} = \bar{\mathbf{c}}_i \mathbf{A} \mathbf{c}_{m+1} \quad (i = 1, \dots, m+1)$$

passando così dalla dimensione m alla dimensione $m+1$.

7. Si diagonalizza la nuova matrice \mathbf{B} , se ne ordinano gli autovalori e si ritorna al punto 2 con l'autovalore $\lambda_k^{(m+1)}$ e l'autovettore $\mathbf{b}_k^{(m+1)}$.

8. Verificato il test di convergenza di cui al punto 3, il k -esimo autovalore di \mathbf{B} coincide con il k -esimo autovalore di \mathbf{A} e il corrispondente autovettore è dato dalla combinazione lineare

$$\mathbf{a}_k = \sum_{i=1}^m b_{ik}^{(m)} \mathbf{c}_i.$$

PROVE NUMERICHE

Le prove numeriche sono state eseguite su una matrice 61×61 (proveniente da un calcolo CI) che è stata diagonalizzata interamente con il metodo di Givens per ottenere gli autovalori e gli autovettori di confronto. Sono state eseguite due serie di prove: nella prima gli autovalori venivano determinati in modo indipendente uno dall'altro, nella seconda gli autovalori successivi venivano determinati partendo dai vettori (autovettori) relativi agli autovalori che precedevano quello in considerazione (data la forma del procedimento iterativo, nelle iterazioni successive questi autovettori rimangono inalterati). Mentre nella prima serie di prove si sono incontrate difficoltà di convergenza per la terza e quarta radice, nella seconda serie di prove ciò non si è verificato ma anzi ogni autovalore è stato raggiunto con un numero di cicli sensibilmente inferiore. Si è comunque notato che la velocità di convergenza dipende dal numero l di vettori di partenza (oltreché naturalmente dal tipo) e, contrariamente a quanto si può forse presumere, essa può non migliorare pur aumentando il numero l . Non sembra possibile ricavare un criterio che permetta di stabilire a priori il numero più conveniente di vettori con cui iniziare l'iterazione e ciò costituisce una limitazione, certo non determinante, del metodo.

TABELLA I

*Autovalori, velocità di convergenza e tempo di calcolo
per una matrice 61×61 .*

	Autovalori	N. di iterazioni ⁽¹⁾	Tempo di calcolo (s) ⁽²⁾
Givens	-27.542349		
	-26.502309		
	-26.500290		
	-26.336195		
Nesbet	-27.542331	300	1.26
	-26.501916	31200	120.10
Davidson	-27.542348	5 (1)	3.27
	-26.502293	24 (4)	19.45
	-26.500276	2 (6)	2.83
	-26.336137	17 (7)	20.54

(1) In parentesi è riportato il numero 1 dei vettori iniziali.

(2) Tempi CPU su elaboratore UNIVAC 1106, registrati in «time sharing».

In Tabella I sono messi a confronto gli autovalori ottenuti nella seconda serie di prove con quelli ottenuti secondo Givens. Va notato che il calcolo di confronto secondo Givens è stato condotto in doppia precisione, mentre gli altri calcoli sono stati eseguiti con aritmetica in doppia precisione ma con memorizzazione delle Tabelle in precisione semplice. Un confronto significativo per quanto riguarda la velocità di convergenza risulta dall'applicazione del metodo di Nesbet al calcolo del più basso autovalore della matrice \mathbf{A} e del più basso autovalore della matrice $\mathbf{A}' = \mathbf{A}(\mathbf{I} - \mathbf{a}_1 \bar{\mathbf{a}}_1)$, che, come è noto, coincide con il secondo autovalore di \mathbf{A} .

Dalla Tabella I si vede che per la determinazione del primo autovalore di \mathbf{A} il metodo di Nesbet è discretamente più veloce di quello di Davidson mentre per la determinazione del primo autovalore di \mathbf{A}' è estremamente più lento e lo sarebbe ancora di più nella determinazione degli autovalori superiori al secondo.

La lista FORTRAN del programma di calcolo utilizzabile sugli elaboratori UNIVAC della serie 1100 può essere richiesta agli Autori.

BIBLIOGRAFIA

- [1] I. SHAVITT, C. BENDER, A. PIPANO e R. P. HOSTENY (1973) - « J. Comput. Phys. », *11*, 90.
- [2] R. K. NESBET (1965) - « J. Chem. Phys. », *43*, 311.
- [3] C. LANCZOS (1950) - « J. Res. Nat. Bur. Stand. », *45*, 255.
- [4] M. G. FELER (1974) - « J. Comput. Phys. », *14*, 341.
- [5] E. R. DAVIDSON (1975) - « J. Comput. Phys. », *17*, 87.
- [6] E. BODEWIG (1956) - *Matrix Calculus*, North-Holland Publ. Comp., Amsterdam, p. 54 e p. 304.
- [7] P. O. LØWDIN (1950) - « J. Chem. Phys. », *18*, 365.
- [8] J. A. R. COOPE e D. W. SABO (1977) - « J. Comput. Phys. », *23*, 404.