

Analiza performanței pachetului SLEPC în calculul valorilor și vectorilor proprii

Ionela RUSU, Ștefan-Gheorghe PENTIUC

Abstract—This paper presents an approach for performance analysis of SLEPC toolkit to solve the eigenvalue problem. SLEPC is a library for solving eigenvalue problems of large-scale low parallel computers. SLEPC is an extension of PETSc and can be used for standard or generalized eigenproblems, with real or complex arithmetic. In this work we present a parallel implementation of eigenvalue computation on IBM Cell B.E. processor.

Index Terms—eigenvalue, eigenvectors, parallel, SLEPC, PETSc, Cell/B.E.

I. INTRODUCERE

DETERMINAREA corectă a valorilor proprii (cunoscute și ca rădăcini ale polinomului caracteristic) ale unei matrici este extrem de importantă în domeniul fizicii și al ingineriei, unde este echivalentă cu diagonalizarea matricii, o operație comună în diverse aplicații precum rezolvarea ecuațiilor algebrice, teoria stabilității, analiza micilor oscilații într-un sistem care vibrează etc.

În această lucrare este prezentată o analiză a performanțelor colecției de biblioteci SLEPC, Scalable Library for Eigenvalue Problem Computation[1], în contextul problemelor de calcul al valorilor proprii a unei matrici de mari dimensiuni. Această bibliotecă oferă avantajul paralelizării aplicațiilor ce pot fi dezvoltate. Pentru analiza performanțelor s-a dezvoltat o aplicație pentru determinarea valorilor și vectorilor proprii pentru o matrice de dimensiuni mari, care va fi testată pe un cluster IBM miniRoadRunner dotat cu procesoare PowerXCell8i.

Lucrarea este organizată după cum urmează. În secțiunea II este prezentat contextul matematic în care este formulată problema ce se vrea a fi rezolvată. Secțiunea III prezintă o scurta descriere a colecției de biblioteci SLEPC. În secțiunea IV este prezentată aplicația care determină valorile și vectorii proprii pentru o matrice de dimensiuni mari. Rezultatele obținute în urma testelor efectuate sunt prezentate în secțiunea V. Lucrarea se încheie cu câteva concluzii referitoare la cele prezentate.

Ionela Rusu – Universitatea Ștefan cel Mare Suceava, str. Universității, nr.13, RO-720229 Suceava (e-mail: ionelar@eed.usv.ro)

Ștefan-Gheorghe Pentiu – Universitatea Ștefan cel Mare Suceava, str. Universității, nr.13, RO-720229 Suceava (e-mail: pentiu@eed.usv.ro)

II. FORMULAREA PROBLEMEI

A. Valori și vectori proprii

Fie A o matrice de dimensiune $n \times m$ cu coeficienții reali sau complecși. Scalarul λ se numește *valoare proprie* pentru matricea A dacă există vectorul x astfel încât[9]:

$$Ax = \lambda x \quad (1)$$

Vectorul x se numește *vector propriu* asociat valorii proprii λ .

Relația (1) $Ax = \lambda x$ poate fi scrisă echivalent:

$$(\lambda I_n - A)x = 0 \quad (2)$$

unde I_n este matricea unitatea de ordinul n .

Sistemul liniar omogen $(\lambda I_n - A)x = 0$ admite o soluție nenulă x dacă și numai dacă $\det(\lambda I_n - A) = 0$.

Polinomul $PA(\lambda) = \det(\lambda I_n - A)$ se numește *polinom caracteristic* al matricii A . Orice matrice cu coeficienți complecși are exact n valori proprii (reale sau complexe, nu neapărat distincte). Acestea reprezintă rădăcinile polinomului caracteristic. Pentru polinoame cu grad mai mare decât 4 nu pot fi întotdeauna calculate rădăcinile într-un număr finit de pași. Astfel calculul valorilor proprii se va face în mod iterativ. De obicei, calculul valorilor proprii exprimate ca și rădăcini ale polinomului caracteristic nu este recomandabil datorită:

- volumului de calcul necesar pentru găsirea coeficienților polinomului;
- volumului de calcul necesar găsirii rădăcinilor polinomului.

III. SLEPC

SLEPC – Scalable Library for Eigenvalue Problem Computation, este un pachet software ce oferă soluții pentru rezolvarea problemelor de calcul ale valorilor proprii pentru matrici rare de dimensiuni mari pe calculatoare paralele.

În afară de problema standard a valorilor proprii, această colecție de biblioteci este adresată și altor tipuri de probleme, precum problema generalizată a valorilor proprii sau descompunerea valorilor singulare. SLEPC permite calcule

aritmetice atât cu valori reale, cât și cu valori complexe, în simplă sau dublă precizie și nu este strict numai pentru probleme cu matrici simetrice (hermitiene). Ca și limbaje de programare suportă C,C++, FORTRAN și Python. Aplicabilitatea pachetului SLEPc include domenii precum ingineria nucleară, fizica plasmei, astrofizică etc.[2]

SLEPc furnizează o colecție de solveri pentru valori proprii, majoritatea dintre ei bazându-se pe paradigma proiecției subspațiului. În particular, pachetul include o implementare paralelă robustă și eficientă a metodei Lanczos [3], dar și o implementare a metodei Krylov-Schur[4] pentru variantele de probleme cu matrici nesimetrice. Solverul Lanczos include posibilitatea de a seta un parametru care să permită calcularea unui număr mare de perechi conjugate pentru valorile proprii. Această caracteristică lipsește în alte biblioteci precum ARPACK[5].

SLEPc este construit având la bază pachetul PETSc (Portable, Extensible Toolkit for Scientific Computation[6]), un framework pentru rezolvarea ecuațiilor diferențiale, a cărui abordare este de a încapsula algoritmi matematici

utilizând tehnici de programare orientate pe obiect în vederea gestionării în mod eficient codul pentru comunicația prin mesaje într-o execuție paralelă. Codul PETSc este disponibil gratuit și utilizat în foarte multe aplicații din diverse domenii. PETSc este orientat pe obiect în sensul că întreg codul este construit în jurul structurilor de date de tip obiect. Programatorul lucrează direct cu aceste obiecte, fără a se concentra asupra legăturii dintre aceste structuri de date. PETSc folosește trei tipuri de obiecte abstracte, și anume: seturile de indecși, vectori și matrici. Pe baza acestor tipuri de obiecte sunt construite diverse clase de obiecte solveri ce includ solveri liniari, neliniari etc. SLEPc moștenește toate aceste proprietăți ale pachetului PETSc, incluzând portabilitatea pentru o paletă largă de platforme paralele, scalabilitatea pentru un număr mare de procesoare și flexibilitatea în timpul rulării. În Fig. 1 este prezentat schematic funcționalitatea pachetului SLEPc în raport cu PETSc[6].

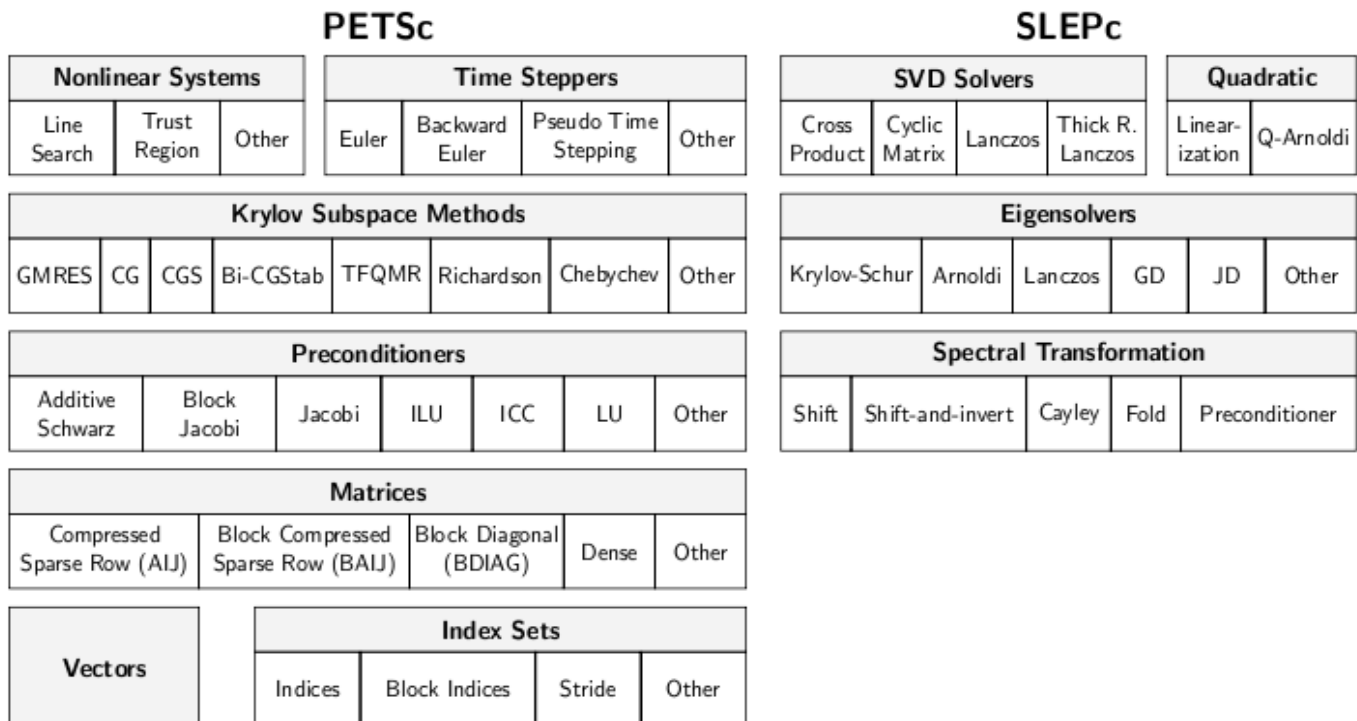


Fig. 1. Schema claselor în SLEPc în raport cu PETSc.

Solverii din PETSc și SLEPc au o implementare neutră în ceea ce privește structura datelor. Aceasta înseamnă că calculul poate fi efectuat cu diferite formate de stocare ale matricii, chiar și cu matrici care nu sunt stocate explicit. Implicit, o matrice în PETSc este memorată într-un format pentru matrici rare comprimat pe linii, paralel, unde fiecare procesor memorează un subset de linii. Alte formate includ varianta pentru matrici dense, cât și pentru matrici simetrice (sbaij), unde doar partea triunghiului superior este memorată.

IV. DETERMINAREA VALORILOR PROPRII UTILIZÂND SLEPc

Pentru calculul valorilor și vectorilor proprii a unei matrici a fost dezvoltată o aplicație constituită din trei module: un modul pentru pregătirea datelor asupra cărora se vor face prelucrările, modulul pentru calculul valorilor și vectorilor proprii, iar cel de-al treilea modul are drept scop verificarea datelor obținute. Primul modul constă în convertirea matricii pentru care se caută valorile proprii în formatul acceptat de către pachetul SLEPc, și anume formatul PETSc binar.

Formatul fișierelor de intrare suportat în acest moment este

formatul MatrixMarket [7]. Execuția acestui modul va genera un fișier binar în formatul PETSc în care se vor memora informații despre matricea asupra căreia se vor efectua calculele. Acest fișier va servi drept fișier de intrare pentru cel de al doilea modul. Constrângerile impuse de această aplicație se referă la caracteristicile matricei, și anume se impune ca matricea să fie matrice rară sau densă, cu valori reale.

Al doilea modul în cadrul acestei aplicații reprezintă modulul cel mai important. În această parte este construit obiectul ce va încapsula toate datele problemei ce trebuie rezolvată. Astfel, pe lângă matricea în cauză, se vor seta parametrii pentru algoritmul de determinare a valorilor proprii, și anume: metoda folosită în determinarea valorilor, numărul de valori proprii căutate, numărul de iterații ale algoritmului, condițiile de stop ale algoritmului etc. Modulul pentru calculul valorilor proprii este implementat astfel încât să suporte o execuție paralelă, pentru paralelizarea algoritmului folosindu-se standardul MPI – Message Passing Interface. Strategia de paralelizare adoptată aici, descompunerea domeniului, implică împărțirea matricei la toate cele N procese MPI și executarea algoritmului pe fiecare din cele N procese, dar la o dimensiune mai mică a problemei. În final, este utilizat un proces master care va colecta toate datele provenite de la celelalte $N-1$ procese MPI și va furniza soluția problemei, adică valorile proprii cu vectorii proprii asociați.

Cel de-al treilea modul constă în verificarea corectitudinii valorilor proprii și a vectorilor proprii determinați de cel de-al doilea modul.

V. REZULTATE EXPERIMENTALE

În această secțiune vor fi prezentate rezultatele experimentale obținute în calculul valorilor proprii pentru o matrice dată cu ajutorul colecției de biblioteci pentru rezolvarea problemelor cu valori proprii având la bază calcule științifice pe structuri paralele, SLEPc.

Pentru testarea performanței pachetului SLEPc în determinarea valorilor proprii s-au realizat o serie de teste pe clusterul IBM miniRoadRunner HPC dotat cu procesoare PowerXCell8i. Astfel, au fost utilizate matrici de diverse dimensiuni puse la dispoziție de Matrix Market Repository[8]. Din matricile puse la dispoziție s-a ales o matrice rară, simetrică, pozitiv semidefinită, de dimensiune 2003×2003 , având 30% elemente nenule. În Fig. 2 este prezentată diagrama structurii matricei.



Fig. 2. Digrama structurii pentru o matrice rară, simetrică, pozitiv semidefinită, de dimensiune 2003×2003 elemente.

Pentru înregistrarea timpilor de execuție s-a început cu o execuție secvențială a aplicației, prin utilizarea unui singur procesor PowerXCell, după care s-a mărit numărul de procesoare până la 60. În Fig. 3 este prezentat graficul timpilor de execuție obținuți.

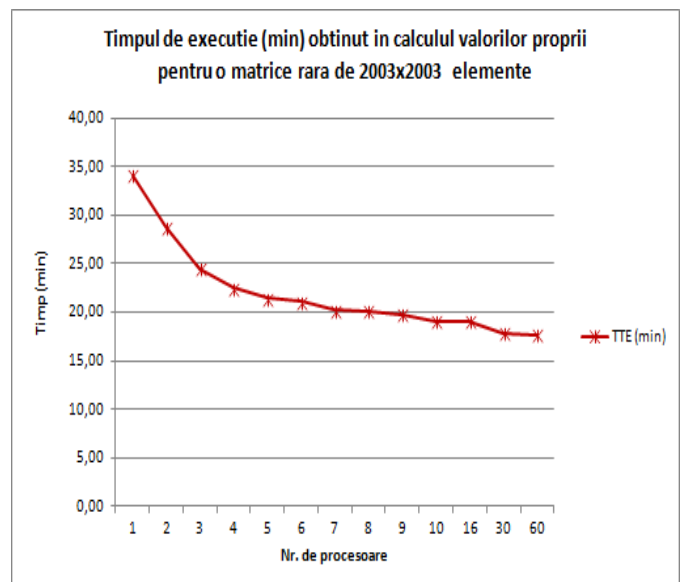


Fig. 3. Timpii de execuție obținuți în calculul valorilor proprii pentru o matrice rară de dimensiune 2003×2003 elemente.

Pe baza timpilor de execuție măsurați s-a calculat accelerarea cobținută prin folosirea celor $N=2,3,\dots, 60$ procesoare, față de execuția algoritmului pe un singur procesor. Rezultatele au fost exprimate sub forma unui grafic ce este prezentat în Fig. 4.

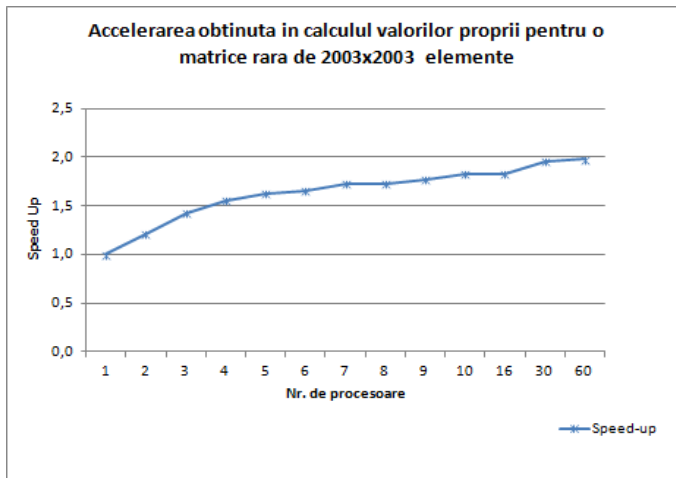


Fig. 4. Accelerarea obținută în urma execuției paralele a algoritmului de calcul al valorilor proprii pentru o matrice de 2003x2003 elemente.

CONCLUZII

În cadrul acestei lucrări a fost analizat pachetul SLEPc din punct de vedere al performanțelor în calculul valorilor proprii pentru matrici rare de dimensiuni mari. Din rezultatele obținute putem observa că pentru o matrice de dimensiune relativ mare de 2003x2003 elemente, s-a obținut o accelerare cu un factor de până la 2. Din graficul ce exprimă timpii de execuție (Fig.2) se observă că obținem o scădere a timpului de [10]

execuție la jumătate pentru un număr de procesoare ≥ 16 , față de timpul de execuție pe un singur procesor. Accelerarea obținută nu reprezintă un rezultat remarcabil, dar constituie punctul de plecare în optimizarea algoritmilor de determinare a valorilor proprii puși la dispoziție de pachetul SLEPc.

MULȚUMIRI

Această lucrare a beneficiat de suport financiar prin proiectul "Provocarile cunoașterii și dezvoltare prin cercetare doctorala PRO-DOCT Contract nr. POSDRU/88/1.5/S/52946", proiect cofinanțat din Fondul Social European prin Programul Operational Sectorial Dezvoltarea Resurselor Umane 2007-2013.

REFERINȚE BIBLIOGRAFICE

- [1] Pagina principală biblioteca SLEPc, <http://www.grycap.upv.es/slep/>
- [2] Aplicații SLEPc, <http://www.grycap.upv.es/slep/material/appli.htm>
- [3] Algoritmul Lanczos, http://en.wikipedia.org/wiki/Lanczos_algorithm
- [4] V. Hernandez, J. E. Roman, A. Tomas, V. Vidal, "Krylov-Schur Methods in SLEPc", SLEPc Technical Report STR-7, 2007
- [5] Rich Lehoucq, Kristi Maschhoff, Danny Sorensen, ARPACK Library, <http://www.caam.rice.edu/software/ARPACK/>
- [6] Pagina principală PETSc, <http://www.mcs.anl.gov/petsc/>
- [7] Formatul MatrixMarket, <http://math.nist.gov/MatrixMarket/formats.html>
- [8] Matrix Market Repository, <http://math.nist.gov/MatrixMarket/index.html>
- [9] Mădălina Roxana Buneci, Metode numerice, <http://www.utgjiu.ro/math/mbuneci/book/mn/fmnl13.pdf>