

CALCULUL VALORILOR PROPRII ȘI VECTORILOR PROPRII PENTRU MATRICE NESIMETRICE

Sistemele dinamice liniare formează o clasă importantă, cu aplicații largi în domenii variate ale științei și tehnicii. Acestea sunt reprezentate numeric de un set de matrice, iar proprietățile lor sunt exprimate, din punct de vedere cantitativ, în termenii elementelor acestor matrice. Astfel stabilitatea sistemelor liniare este condiționată exclusiv de plasarea valorilor proprii ale matricei de stare a sistemului.

În acest capitol sunt prezentate, în principal, cele mai performante metode de calcul ale valorilor proprii ale matricelor reale dense nesimetrice care se întâlnesc cel mai des în aplicații. Din motive de eficiență aceste metode fac apel, practic exclusiv, la o aritmetică reală deși valorile proprii ale unei matrice reale pot fi și complexe. Din motive de condiționare numerică calculul vectorilor proprii se face cu mai puțină acuratețe și, din acest motiv, practica numerică recomandă utilizarea în aplicații în locul lor a vectorilor Schur, definiți în acest capitol. Cazul matricelor complexe poate fi redus la cazul matricelor reale de ordin dublu.

Valori și vectori proprii

Prezentăm aici conceptele de valoare proprie și vector propriu ale unei matrice și principalele lor proprietăți. De asemenea, având în vedere faptul că valorile proprii, ca și elementele vectorilor proprii, ale unei matrice reale pot fi numere complexe, reamintim câteva noțiuni referitoare la spațiul vectorial C^n .

În acest capitol, prin vector complex n -dimensional x vom înțelege un vector coloană cu n elemente complexe

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad x_i \in C, \quad i = 1 : n.$$

Mulțimea tuturor vectorilor complecși n -dimensionali va fi notată C^n . Evident, orice vector $x \in C^n$ poate fi scris, în mod unic, sub forma

$$x = u + iv, \quad u, v \in R^n$$

unde i este unitatea imaginară. Mulțimea tuturor matricelor cu elemente complexe având m linii și n coloane va fi notată $C^{m \times n}$. Deși spațiile liniare R^n și C^n au foarte multe proprietăți comune, induse de axiomele comune ale corpurilor de numere complexe și reale, există și diferențe cum sunt cele datorate definirii diferite a produsului scalar standard a doi vectori și a normei vectoriale induse de acesta. Pentru evidențierea unora dintre aceste diferențe vom nota conjugatul transpusului (i.e. conjugatul hermitic al) unui vector $x \in C^n$ prin

$$x^H = \bar{x}^T,$$

unde bara superioară semnifică operația de conjugare complexă iar indicele T transpunerea.

Dacă $x, y \in C^n$, definim produsul scalar al celor doi vectori prin numărul complex

$$\langle x, y \rangle = y^H x.$$

Evident, avem

$$\langle y, x \rangle = \overline{\langle x, y \rangle}.$$

Cu ajutorul produsului scalar putem defini conceptul de ortogonalitate în C^n . Vom spune că doi vectori sunt ortogonali dacă produsul lor scalar este nul, i.e.

$$x^H y = y^H x = 0.$$

De asemenea, plecând de la observația că scalarul $x^H x$ este un număr real pozitiv, oricare ar fi vectorul nenul $x \in C^n$, se poate defini următoarea normă pe C^n

$$\|\cdot\|_2: C^n \rightarrow R_+, \quad \|x\|_2 = \sqrt{x^H x},$$

numită normă euclidiană.

Mai general, foarte multe rezultate referitoare la norme pe R^n pot fi extinse la norme corespunzătoare pe C^n .

Rolul matricelor reale simetrice este jucat în $C^{n \times n}$ de matricile *hermitice*. O matrice $A \in C^{n \times n}$ se numește *hermitică* dacă

$$A^H = A.$$

Dacă $A \in C^{n \times n}$ este hermitică, atunci scalarul $\alpha = x^H A x$ este real, oricare ar fi $x \in C^n$, fapt care ne permite definirea matricelor pozitiv-definite în $C^{n \times n}$. O matrice hermitică $A \in C^{n \times n}$ este *pozitiv-definită* dacă

$$x^H A x > 0 \quad \forall x \in C^n, \quad x \neq 0.$$

În sfârșit, în analogie cu matricile reale ortogonale, vom spune despre o matrice $Q \in C^{n \times n}$ că este *unitară* dacă are coloanele ortogonale și de normă euclidiană (7) unitară, respectiv dacă

$$Q^H Q = I_n.$$

Vom introduce acum conceptele de *valoare proprie* și *vector propriu asociat*. Deși vom considera mai ales cazul matricelor reale, aceste noțiuni pot fi definite în cadrul mai general al matricelor complexe.

Definiția 9.1 Fie $A \in C^{n \times n}$. Un număr $\lambda \in C$ se numește *valoare proprie* a matricei A dacă există un vector nenul $x \in C^n$, numit *vector propriu asociat* valorii proprii $\lambda \in C$, astfel încât

$$A x = \lambda x.$$

Sistemul liniar omogen (11) admite soluții nenule dacă și numai dacă

$$p(\lambda) = \det(\lambda I_n - A) = 0.$$

Polinomul monic $p(\lambda)$ de gradul n se numește *polinom caracteristic* al matricei $E(\lambda) \in C^n$, iar ecuația se numește *ecuație caracteristică*.

Prin urmare, valorile proprii ale unei matrice sunt zerourile polinomului caracteristic. Dacă luăm în considerare multiplicitățile, numărul valorilor proprii este egal cu ordinul matricei. Mulțimea

$$\lambda(A) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$$

a valorilor proprii ale matricei A poartă numele de *spectrul* (de valori proprii al) matricei A . Numărul real

$$\rho(A) = \max_{\lambda_i \in \lambda(A)} (|\lambda_i|)$$

se numește *raza spectrală* a matricei A .

Dacă A este o matrice reală atunci valorile proprii complexe apar în perechi complex conjugate.

Valorile proprii ale unei matrice satisfac relațiile

$$\sum_{i=1}^n \lambda_i = \sum_{i=1}^n A(i,i) = \text{tr}(A),$$

$$\prod_{i=1}^n \lambda_i = \det(A),$$

unde $\text{tr}(A)$ este, prin definiție, *urma* matricei A (vezi problema rezolvată R9.1).

Dacă $x \in C^n$ este un vector propriu asociat valorii proprii $\lambda \in \lambda(A)$ atunci oricare ar fi $0 \neq \alpha \in C$ vectorul $y = \alpha x$ este de asemenea un vector propriu al matricei A asociat aceleiași valori proprii. În consecință, vectorii proprii asociați unei valori proprii nu sunt unic determinați decât ca direcții. Mai mult, se arată că mulțimea vectorilor proprii asociați unei valori proprii $\lambda \in \lambda(A)$ generează un *subspațiu liniar* $E(\lambda) \subset C^n$ numit *subspațiul propriu* al valorii proprii λ . Subspațiile proprii sunt subspații A -invariante în sensul următoarei definiții.

Definiția 9.2 Fie o matrice $A \in C^{n \times n}$. Un subspațiu liniar $V \subset C^n$ se numește *subspațiu -invariant al matricei A sau, pe scurt, subspațiu A -invariant* dacă $AV \subset V$, i.e. $Ax \in V \quad \forall x \in V$.

O matrice $A \in C^{n \times n}$ care admite n vectori proprii liniar independenți se numește *simplă*. Se arată că o matrice care are cele n valori proprii distincte este simplă, dar reciproca nu este, în general, adevărată. Cei n vectori proprii liniar independenți ai unei matrice simple formează o bază a spațiului C^n , numită *baza proprie* asociată matricei A .

Ordinul de multiplicitate n_i al rădăcinii λ_i a polinomului caracteristic se numește *multiplicitate algebrică* a valorii proprii $\lambda_i \in \lambda(A)$. Dacă $n_i = 1$ valoarea proprie se numește *simplă*. Dimensiunea $\dim E(\lambda_i)$ a subspațiului propriu al valorii proprii $\lambda_i \in \lambda(A)$ se numește *multiplicitatea geometrică* a valorii proprii λ_i . În general, avem

$$\dim E(\lambda_i) \leq n_i.$$

Încheiem acest paragraf cu un set de proprietăți ale subspațiilor invariante și, în particular, ale subspațiilor proprii ale unei matrice.

Teorema 1 Fie matricea $A \in C^{n \times n}$.

a) Dacă x_1, x_2, \dots, x_p sunt vectori proprii ai matricei A și notăm $X = [x_1 \ x_2 \ \dots \ x_p] \in C^{n \times p}$, atunci subspațiul liniar

$$S = \text{Im} X = \left\{ y \in C^n \mid \exists z \in C^p \text{ astfel încât } y = Xz \right\}$$

este A -invariant, i.e. orice subspațiu generat de vectori proprii ai unei matrice este un subspațiu invariant al acelei matrice.

b) Dacă $S \subset C^n$ este un subspațiu A -invariant p -dimensional și vectorii x_1, x_2, \dots, x_p formează o bază a lui S , atunci există o matrice $B \in C^{p \times p}$ cu

$$\lambda(B) \subset \lambda(A)$$

astfel încât matricea $X = [x_1 \ x_2 \ \dots \ x_p] \in C^{n \times p}$ satisface condiția

$$AX = XB.$$

Matricea B se numește restricția matricei A la subspațiul A -invariant S și se notează $B = A|_S$.

c) Dacă are loc o relație de forma (20) atunci $S = \text{Im} X$ este un subspațiu A -invariant.

d) Complementul ortogonal $T = S^\perp \subset C^n$ al unui subspațiu A -invariant S este un subspațiu A^H -invariant.

Demonstrație. a) Pentru orice vector $y \in S$ există un vector $z \in C^p$ astfel încât $y = Xz$ și, deci,

$$Ay = AXz = [Ax_1 \ Ax_2 \ \dots \ Ax_p]z = [\lambda_1 x_1 \ \lambda_2 x_2 \ \dots \ \lambda_p x_p]z = [x_1 \ x_2 \ \dots \ x_p] \cdot \begin{bmatrix} \lambda_1 z_1 \\ \lambda_2 z_2 \\ \vdots \\ \lambda_p z_p \end{bmatrix} = Xw \in S, \text{ unde } \lambda_i, i = 1 : p \text{ sunt valorile}$$

proprie ale matricei A asociate vectorilor proprii $x_i, i = 1 : p$. Prin urmare, subspațiul S este A -invariant. b) Coloanele matricei X formând o bază a subspațiului S , orice vector $x \in S$ se scrie sub forma $x = Xy, y \in C^p$. Pe de altă parte S fiind un subspațiu A -invariant rezultă că $Ax \in S, \forall x \in S$, în particular pentru toți $j \in 1 : p$ avem $Ax_j \in S$, i.e. există vectorii $b_j \in C^p, j = 1 : p$ astfel încât $Ax_j = Xb_j$. Rezultă că matricea $B = [b_1 \ b_2 \ \dots \ b_p] \in C^{p \times p}$ satisface egalitatea (20). Mai mult, dacă $z \in C^p$ este un vector propriu al matricei B asociat valorii proprii $\mu \in \lambda(B)$ atunci din (20) avem $AXz = XBz = \mu Xz$. Cum $z \neq 0$ iar X este o matrice monică (i.e. cu coloanele liniar independente) rezultă că $y = Xz \neq 0$, adică y este vector propriu al matricei A (conținut în S) și, deci, $\mu \in \lambda(A)$. Rezultă $\lambda(B) \subset \lambda(A)$. c) Presupunem că are loc (20) pentru o matrice $X \in C^{n \times p}$ și o matrice $B \in C^{p \times p}$. Atunci avem $AXz = XBz, \forall z \in C^p$, i.e. $Ax = Xy \in \text{Im } X, \forall x = Xz \in \text{Im } X$, ceea ce înseamnă că $S = \text{Im } X$ este un subspațiu A -invariant. d) Fie $x \in S$ și $y \in T = S^\perp$ doi vectori arbitrari. Subspațiul S fiind A -invariant avem $Ax \in S$ și, prin urmare, $y^H Ax = (A^H y)^H x = 0$. Cum x este un vector arbitrar din S , rezultă $(A^H y) \perp S$, respectiv $A^H y \in T$ pentru toți $y \in T$, adică subspațiul T este A^H -invariant.

Transformări de asemănare. Forma Jordan

Vom fi interesați în transformări ale matricei A care conservă spectrul de valori proprii. Astfel de transformări sunt *transformările de asemănare*.

Vom spune că două matrice A și B sunt *asemenea* dacă există o matrice nesingulară $T \in C^{n \times n}$ astfel încât

$$B = T A T^{-1}.$$

Dacă matricea de transformare T este unitară (în cazul real ortogonală) atunci matricele A și B se numesc *unitar (ortogonal) asemenea*.

Dacă $A, B \in C^{n \times n}$ sunt asemenea atunci au același spectru de valori proprii

$$\lambda(A) = \lambda(B),$$

iar dacă $x \in C^n$ este un vector propriu al matricei A asociat valorii proprii λ , atunci vectorul $y \in C^n$, definit de

$$y = Tx,$$

este un vector propriu al matricei B din (21) asociat aceleiași valori proprii λ .

Evident, valorile proprii ale unei matrice triunghiulare (în particular diagonale) sunt date de elementele sale diagonale. De aceea suntem interesați de situațiile în care o matrice dată este asemenea cu o matrice având o astfel de structură. O matrice asemenea cu o matrice diagonală se numește *diagonalizabilă*.

Dacă pentru toate valorile proprii distincte ale unei matrice date $A \in C^{n \times n}$ avem

$$\dim E(\lambda_i) = n_i,$$

atunci matricea A este diagonalizabilă, respectiv există o matrice nesingulară $T = X^{-1} \in C^{n \times n}$ astfel încât

$$X^{-1}AX = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Într-un astfel de caz matricea A este simplă, coloanele matricei de transformare X formând un set de n vectori proprii liniar independenți ai matricei A .

În cazul general, cea mai simplă structură care poate fi obținută prin transformări de asemănare este așa numita *formă canonică Jordan* definită în următoarea teoremă.

Teorema 2 Oricare ar fi matricea $A \in C^{n \times n}$ există o matrice nesingulară $X \in C^{n \times n}$ astfel încât

$$X^{-1}AX = \text{diag}(J_1, J_2, \dots, J_q),$$

unde

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{bmatrix} \in C^{n_i \times n_i}$$

cu $\lambda_i \in \lambda(A)$ și $\sum_{i=1}^q n_i = n$. Matricele J_i se numesc *blocuri Jordan*.

Numărul și dimensiunile blocurilor Jordan asociate fiecărei valori proprii distincte din spectrul matricei A sunt unice, dar ordonarea blocurilor în (26) poate fi arbitrară.

Forma canonică Jordan conține maximul de informație structurală privitoare la o matrice dată. Totuși structura Jordan (respectiv numărul și dimensiunile blocurilor) este foarte sensibilă la perturbațiile numerice în elementele matricei și, din acest motiv, calculul numeric al formei canonice Jordan întâmpină dificultăți serioase (vezi [5]) și nu este recomandat pentru calculul valorilor proprii într-o aritmetică aproximativă.

9.1.3. Localizarea valorilor proprii. Cercurile lui Gershgorin

În unele aplicații o localizare a valorilor proprii într-un domeniu al planului complex este suficientă pentru a evidenția anumite proprietăți. Amintim, în acest context, problema aprecierii stabilității sistemelor dinamice liniare. De asemenea, o localizare a valorilor proprii permite stabilirea unor aproximații ale acestora utile pentru inițializările necesare în diverse metode iterative.

Există mai multe rezultate privitoare la localizarea valorilor proprii dintre care cel mai cunoscut este dat de *teorema cercurilor lui Gershgorin*.

Teorema 3 Spectrul de valori proprii al unei matrice $A \in C^{n \times n}$ satisface condiția

$$\lambda(A) \subset \bigcup_{i=1}^n D_i,$$

unde D_i sunt discurile (sau cercurile) lui Gershgorin definite de

$$D_i = \left\{ z \in C \mid |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right\}.$$

Demonstrație. Fie $\lambda \in \lambda(A)$ și x un vector propriu asociat lui λ . Fie x_j componenta de modul maxim a lui x , i.e. $\frac{|x_j|}{|x_i|} \leq 1$ pentru toți $j \in 1:n$. Atunci linia i a relației de definiție $Ax = \lambda x$ se poate scrie sub

$$\text{forma } (\lambda - a_{ii})x_i = \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j, \text{ de unde rezultă imediat inegalitățile } |\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \cdot \frac{|x_j|}{|x_i|} \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \text{ q.e.d.}$$

Se poate arăta că dacă unul din discurile lui Gershgorin este izolat de celelalte, atunci el conține o valoare proprie a matricei A și numai una, oferind un mijloc de separare a valorilor proprii.

Alte metode de localizare a valorilor proprii ale unei matrice, în general mai fine dar mult mai puțin utilizate în practica curentă, fac obiectul unor probleme rezolvate sau propuse (vezi secțiunile 9.8 și 9.9).

9.1.4. Condiționarea numerică a valorilor și vectorilor proprii

Așa cum s-a menționat în primul capitol al lucrării, acuratețea rezultatelor obținute prin rezolvarea unei probleme de calcul numeric într-un mediu de calcul aproximativ depinde esențial de doi factori. Pe de o parte precizia obținută depinde de “calitatea numerică a problemei”, i.e. de sensibilitatea rezultatelor la variațiile datelor problemei în ipoteza unor calcule exacte. Este vorba de așa numita *condiționare numerică a problemei*. Pe de altă parte calitatea rezultatelor depinde de calitatea algoritmului utilizat, exprimată prin așa numita *stabilitate numerică a algoritmului* respectiv. Întrucât condiționarea numerică a valorilor proprii este independentă de modul de calcul al acestora, vom prezenta aici rezultatele esențiale privitoare la aceasta, urmând ca discuția despre calitățile algoritmilor propuși să aibă loc în finalul capitolului.

Aprecierea condiționării numerice se face uzual prin stabilirea unor margini superioare pentru variațiile rezultatelor în raport cu variațiile datelor de intrare. Chiar dacă, de multe ori, aceste margini sunt supraevaluate ele oferă posibilitatea detectării situațiilor critice în care erorile au tendința de a ieși de sub control. Evaluarea sensibilității valorilor și vectorilor proprii se bazează pe proprietățile de continuitate ale acestora în raport cu elementele matricei. Precizăm de la început că ne vom ocupa în exclusivitate de cazul valorilor proprii simple. Valorile proprii multiple sunt, în general, mult mai sensibile decât cele simple, iar studiul analitic al sensibilității lor ridică dificultăți importante.

A. Condiționarea numerică a valorilor proprii

Fie o matrice $A \in C^{n \times n}$, o valoare proprie simplă $\lambda \in \lambda(A)$ și vectorii proprii de normă euclidiană unitară $x, y \in C^n$ la dreapta, respectiv la stânga, asociați valorii proprii λ . Prin urmare avem $Ax = \lambda x$ și, respectiv, $y^H A = \lambda y^H$. Considerăm acum o perturbație $E = \varepsilon G$ cu $\|G\| = 1$ a matricei A . Vom fi interesați de evaluarea valorii proprii $\lambda(\varepsilon)$ și a vectorului $x(\varepsilon)$ a matricei perturbate $\tilde{A} = A + E$. Conform unor rezultate clasice [10], valoarea proprie $\lambda(\varepsilon)$ și vectorul propriu $x(\varepsilon)$ admit dezvoltări în serie de puteri

$$\lambda(\varepsilon) = \lambda + \mu_1 \varepsilon + \mu_2 \varepsilon^2 + \dots$$

$$x(\varepsilon) = x + v_1 \varepsilon + v_2 \varepsilon^2 + \dots$$

convergente într-o vecinătate a punctului $\varepsilon = 0$. Evident, avem $\lambda(0) = \lambda$ și $x(0) = x$ iar funcțiile $\lambda(\varepsilon)$ și $x(\varepsilon)$ sunt continue și derivabile în domeniul de convergență. Vom considera în cele ce urmează că parametrul ε este suficient de mic pentru a putea neglija puterile superioare ale acestuia și a ne mărgini la aproximațiile liniare ale funcțiilor $\lambda(\varepsilon)$ și $x(\varepsilon)$. În aceste condiții putem aprecia că sensibilitatea valorii proprii $\lambda(\varepsilon)$ pe “direcția” definită de matricea G este dată de $|\mu_1|$ și că pentru toate direcțiile de perturbare sensibilitatea sa poate fi definită prin numărul de condiționare

$$\kappa_\lambda = \max_{\substack{G \in C^{n \times n} \\ \|G\|=1}} |\mu_1|.$$

Pentru a găsi o evaluare a acestui număr de condiționare considerăm relația $(A + \varepsilon G)x(\varepsilon) = \lambda(\varepsilon)x(\varepsilon)$ care derivată în raport cu ε ne conduce, pentru $\varepsilon = 0$, la

$$Gx + Av_1 = \mu_1 x + \lambda v_1.$$

Înmulțind această relație la stânga cu y^H rezultă $y^H Gx + y^H Av_1 = \mu_1 y^H x + \lambda y^H v_1$, i.e. $y^H Gx = \mu_1 y^H x$ întrucât $y^H A = \lambda y^H$. Ținând seama acum de faptul că pentru valorile proprii simple avem $y^H x \neq 0$ (vezi problema rezolvată R9.5) obținem

$$|\mu_1| = \frac{|y^H Gx|}{|y^H x|} \leq \frac{\|y\| \cdot \|G\| \cdot \|x\|}{|y^H x|} = \frac{1}{|y^H x|}.$$

Prin urmare, numărul de condiționare al unei valori proprii simple este

$$\kappa_\lambda = \max_{\substack{G \in C^{n \times n} \\ \|G\|=1}} |\mu_1| = \frac{1}{|y^H x|}.$$

Valoarea minimă a numărului de condiționare este 1 și se obține atunci când vectorii proprii la stânga și la dreapta, asociați valorii proprii de interes, sunt coliniari (aceasta se întâmplă, de exemplu, la matricele hermitice și, în particular, la matricele reale simetrice). Conform dezvoltărilor de mai sus, condiționarea κ_λ indică nivelul de amplificare al erorilor absolute din datele inițiale care se regăsesc în erorile absolute ale valorilor proprii. Prin urmare, pentru valori proprii cu același număr de condiționare erorile *relative* vor fi cu atât mai mari cu cât valorile proprii sunt de modul mai mic.

În aplicații prezintă interes de multe ori condiționarea unui grup de valori proprii sau a întregului spectru al unei matrice date. O cale de a defini condiționarea unui grup de valori proprii este de a considera o normă a vectorului condiționărilor numerice a valorilor proprii individuale care fac parte din grup. O altă cale face apel la conceptul de *proiector spectral* care exprimă condiționarea unui grup de valori proprii prin poziția relativă a subspațiilor A -invariante și A^H -invariante generate de vectorii proprii la dreapta și la stânga asociați valorilor proprii din grup. Pentru detalii recomandăm consultarea, de exemplu, a lucrărilor [13], [29], [30].

Încheiem considerațiile noastre privitoare la condiționarea numerică a valorilor proprii prin prezentarea teoremei Bauer-Fike care reprezintă un rezultat important ce permite exprimarea condiționării numerice a spectrului unei matrice simple într-un context general.

Teorema 4 Considerăm o matrice $A \in C^{n \times n}$ diagonalizabilă și fie $X \in C^{n \times n}$ o matrice nesingulară de vectori proprii ai matricei A . Dacă $E \in C^{n \times n}$ este o matrice de perturbație arbitrară și $\mu \in \lambda(A + E)$ este o valoare proprie a matricei perturbate $\tilde{A} = A + E$, atunci

$$e(\mu) = \min_{\lambda \in \lambda(A)} |\lambda - \mu| \leq \kappa(X) \cdot \|E\|$$

unde

$$\kappa(X) = \|X\| \cdot \|X^{-1}\|$$

este numărul de condiționare la inversare al matricei X a vectorilor proprii iar $\|\cdot\|$ este orice normă matriceală consistentă care satisface condiția

$$\|\text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)\| = \max_{i=1:n} (|\alpha_i|)$$

(în particular, $\|\cdot\|$ poate fi oricare din normele $\|\cdot\|_p$, $p = 1, 2, \infty$).

Demonstrația teoremei poate fi găsită în [13]. Aici ne vom mărgini la prezentarea unor succinte comentarii. În primul rând rezultatele specificate în teoremă sunt valabile nu numai pentru perturbații mici ci pentru orice nivel al perturbațiilor. În al doilea rând, trebuie evidențiat rolul jucat de numărul de condiționare la inversare al matricei vectorilor proprii în caracterizarea condiționării valorilor proprii.

Interpretând raportul $\frac{e(\mu)}{\|E\|}$ drept sensibilitate (condiționare) numerică a valorii proprii a matricei A

pentru care se atinge minimumul din (36) rezultă că numărul de condiționare la inversare al matricei vectorilor proprii este o margine superioară pentru numerele de condiționare ale tuturor valorilor proprii și, în consecință, poate fi folosit drept măsură pentru condiționarea întregului spectru. Se pot face și alte conexiuni între rezultatele din teorema Bauer-Fike și numerele de condiționare definite anterior [13]. Încheiem aceste scurte comentarii cu remarcarea faptului că teorema 9.4 confirmă observația că matricele hermitice (simetrice în cazul real) au un spectru perfect condiționat numeric întrucât, după cum vom vedea în capitolul 11, acestea au un set complet de vectori proprii ortogonali, i.e. matricea vectorilor proprii este unitară (ortogonală) și, deci, are în raport cu norma $\|\cdot\|_2$ numărul de condiționare la inversare egal cu 1 (cel mai mic posibil).

B. Condiționarea numerică a vectorilor proprii

Fie o matrice $A \in C^{n \times n}$ cu valorile proprii simple λ_k , $k = 1 : n$ și vectorii proprii asociați $x_k \in C^n$ de normă euclidiană unitară. Considerăm matricea perturbată $\tilde{A} = A + E$, unde $E = \varepsilon G$ cu $\|G\| = 1$ și fie $\lambda_k(\varepsilon)$, $x_k(\varepsilon)$, $k = 1 : n$ valorile proprii, respectiv vectorii proprii asociați ai matricei perturbate. La fel ca în cazul valorilor proprii, pe baza dezvoltării în serie de puteri (31) a funcției vectoriale $x(\varepsilon)$, în ipoteza unor variații ε mici, putem aprecia că sensibilitatea vectorului propriu $x(\varepsilon)$ pe “direcția” definită de matricea G este dată de $\|v_1\|$ și că pentru toate direcțiile de perturbare sensibilitatea sa poate fi definită prin numărul de condiționare

$$\kappa_x = \max_{\substack{G \in C^{n \times n} \\ \|G\|=1}} \|v_1\|.$$

Adaptând notațiile la noul context, relația (33) se poate scrie sub forma

$$Gx_k + Av_1^{(k)} = \mu_1^{(k)}x_k + \lambda_k v_1^{(k)}.$$

Deoarece, în ipotezele menționate, vectorii proprii x_k , $k = 1 : n$ formează o bază a spațiului C^n , rezultă că putem exprima vectorul $v_1^{(k)}$ în raport cu această bază prin relația $v_1^{(k)} = \sum_{i=1}^n \gamma_i^{(k)} x_i$ care, introdusă în (39), produce

$$\sum_{\substack{i=1 \\ i \neq k}}^n \gamma_i^{(k)} (\lambda_k - \lambda_i) x_i = (G - \mu_k I_n) x_k.$$

Înmulțind la stânga ultima relație cu y_i^H , unde y_i este vectorul propriu la stânga al matricei A asociat valorii proprii λ_i și, ținând seama (vezi problema rezolvată R9.5) de faptul că $y_j^H x_i = 0$, $\forall i \neq j$ și $y_i^H x_i \neq 0$, obținem

$$(42) \quad \gamma_i^{(k)} = \frac{y_i^H G x_k}{(\lambda_k - \lambda_i) y_i^H x_k}, \quad i = 1 : n, \quad i \neq k,$$

cu care numărul de condiționare (39) devine

$$(43) \quad \kappa_{x_k} = \max_{\substack{G \in C^{n \times n} \\ \|G\|=1}} \|v_1^{(k)}\| = \max_{\substack{G \in C^{n \times n} \\ \|G\|=1}} \left\| \sum_{\substack{i=1 \\ i \neq k}}^n \frac{y_i^H G x_k}{(\lambda_k - \lambda_i) y_i^H x_i} x_i \right\| \leq \sum_{\substack{i=1 \\ i \neq k}}^n \frac{1}{|\lambda_k - \lambda_i| |y_i^H x_i|}.$$

Se observă din evaluarea (43) că sensibilitatea unui vector propriu la variațiile elementelor matricei depinde atât de sensibilitățile tuturor celorlalte valori proprii cât și de depărtarea valorii proprii asociate de celelalte valori proprii. În consecință, mai ales în cazul existenței unor valori proprii apropiate, condiționarea numerică a vectorilor proprii este mult mai rea decât cea a valorilor proprii. De aceea în literatura de specialitate se recomandă evitarea calculului vectorilor proprii. Din fericire, în majoritatea aplicațiilor, aceștia pot fi înlocuiți cu succes de vectorii Schur (vezi secțiunea ce urmează).

Problema condiționării unor grupuri de vectori proprii se formulează mai general sub forma condiționării subspațiilor invariante. Condiționarea subspațiilor invariante este influențată decisiv de localizarea valorilor proprii asociate. Este însă posibil ca un subspațiu invariant generat de vectori proprii rău condiționați să aibă o condiționare foarte bună dacă grupul corespunzător de valori proprii este bine separat de restul valorilor proprii. Pentru detalii privitoare la cuantificarea condiționării subspațiilor invariante se poate consulta lucrarea [29].

9.2. Forma Schur reală. Tehnici de deflație

O modalitate de abordare a calculului valorilor proprii ce pare a fi naturală, respectiv determinarea polinomului caracteristic și rezolvarea ecuației caracteristice, este practic abandonată în calculul numeric actual pentru dimensiuni de matrice cât de cât semnificative, pe de o parte din cauza sensibilității ridicate a rădăcinilor în raport cu perturbațiile numerice în coeficienții ecuațiilor algebrice, iar pe de altă parte din cauza complexității relativ ridicate a algoritmilor disponibili pentru calculul coeficienților polinomului caracteristic. Din aceste motive, metodele numerice cele mai performante de calcul a valorilor proprii fac apel la structuri (cvasi)triunghiulare, cum este *forma Schur (reală)*.

Baza teoretică a calculului performant al valorilor proprii ale unei matrice este dată de următoarea teoremă.

Teorema 4. Pentru orice matrice $A \in C^{n \times n}$ (în particular $A \in R^{n \times n}$) există o matrice unitară $\tilde{Q} \in C^{n \times n}$ astfel încât matricea unitar asemenea cu A

$$(44) \quad \tilde{S} = \tilde{Q}^H A \tilde{Q} \in C^{n \times n}$$

este superior triunghiulară.

În cazul real, pentru orice matrice $A \in R^{n \times n}$ există o matrice ortogonală $Q \in R^{n \times n}$ astfel încât matricea ortogonal asemenea cu A

$$(45) \quad S = Q^T A Q = S \in R^{n \times n}$$

are o structură cvasisuperior triunghiulară

$$(46) \quad S = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1q} \\ 0 & S_{22} & \dots & S_{2q} \\ & & \dots & \\ 0 & 0 & \dots & S_{qq} \end{bmatrix}$$

unde blocurile diagonale $S_{ii}, i=1:q$ sunt matrice 1×1 sau 2×2 , cele de dimensiune 2×2 având valorile proprii complexe.

Matricea \tilde{S} din (27) poartă numele de *formă Schur (FS)* a matricei A dacă A este complexă și de *formă Schur complexă (FSC)* a lui A dacă A este reală. Matricea S din (45), (46) poartă numele de *formă Schur reală (FSR)* a matricei A . Coloanele $\tilde{q}_j = \tilde{Q} e_j, j=1:n$ ale matricei \tilde{Q} (respectiv coloanele $q_j = Q e_j, j=1:n$ ale matricei Q) se numesc *vectori Schur* ai lui A și formează o bază unitară a spațiului C^n (respectiv, o bază ortogonală a spațiului R^n) numită *baza Schur* asociată lui A . În multe aplicații

vectorii Schur pot înlocui cu succes vectorii proprii. Evident, elementele diagonale ale matricei \tilde{S} sunt valorile proprii ale matricei A , iar în cazul forme Schur reale avem

$$(47) \quad \lambda(A) = \lambda(S) = \bigcup_{i=1}^q \lambda(S_{ii})$$

și, în consecință, în acest caz calculul spectrului de valori proprii se reduce la rezolvarea unui set de ecuații algebrice de grad cel mult doi.

Demonstrația teoremei 9.4. Această demonstrație pune în evidență o tehnică procedurală numită *deflație*, care stă la baza algoritmului fundamental de reducere a unei matrice reale date la forma Schur reală printr-un șir de transformări ortogonale de asemănare.

a) Mai întâi arătăm cum se pune în evidență, printr-un pas de deflație, o valoare proprie reală. Dacă $\lambda \in \lambda(A)$ este reală, fie $x_1 \in R^n$ un vector propriu asociat. Fără restrângerea generalității putem presupune $\|x_1\|_2 = 1$. Considerăm o matrice ortogonală Q_1 a cărei primă coloană este x_1 , i.e.

$$(48) \quad Q_1 = [x_1 \ Y], \quad Y \in R^{n \times (n-1)}$$

(o astfel de matrice există întotdeauna: de exemplu Q_1 poate fi un reflector elementar care anulează elementele $2:n$ ale vectorului x_1). Evident, avem

$$(49) \quad x_1^T Y = 0$$

și, ținând seama de faptul că $Ax_1 = \lambda_1 x_1$, rezultă

$$(50) \quad x_1^T A x_1 = \lambda_1 x_1^T x_1 = \lambda_1, \quad Y^T A x_1 = \lambda_1 Y^T x_1 = 0,$$

de unde obținem matricea

$$(51) \quad A_1 = Q_1^T A Q_1 = \begin{bmatrix} \lambda_1 & x_1^T A Y \\ 0 & Y^T A Y \end{bmatrix} = \begin{bmatrix} \lambda_1 & g_1^T \\ 0 & B \end{bmatrix}$$

care este în formă Schur reală în prima coloană.

b) Un pas de deflație corespunzător unei perechi de valori proprii complex conjugate

$$(52) \quad \lambda_{1,2} = \alpha \pm i\beta, \quad \beta \neq 0$$

pune în evidență un bloc diagonal 2×2 din forma Schur reală. Într-adevăr nu este greu de arătat că vectorii proprii asociați valorilor proprii (52) pot fi aleși complex conjugăți

$$(53) \quad x_{1,2} = u \pm iv$$

cu $u, v \in R^n$ liniar independenți. Din relația de definiție

$$(54) \quad A(u \pm iv) = (\alpha \pm i\beta) \cdot (u \pm iv)$$

rezultă imediat

$$(55) \quad A \cdot [u \ v] = [u \ v] \cdot \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}.$$

Vectorii u și v fiind liniar independenți generează un subspațiu liniar $S = \text{Im}[u \ v]$ bidimensional. Fie $y_1, y_2 \in S$ o bază ortonormală a acestui subspațiu și $Z \in R^{n \times (n-2)}$ o completare până la o matrice ortogonală a matricei $Y = [y_1 \ y_2] \in R^{n \times 2}$ respectiv astfel încât matricea

$$(56) \quad Q_1 = [Y \ Z]$$

este ortogonală. În raport cu baza de mai sus vectorii $u, v \in S$ se pot scrie

$$(57) \quad u = t_{11}y_1 + t_{21}y_2, \quad v = t_{12}y_1 + t_{22}y_2,$$

respectiv

$$(58) \quad [u \ v] = YT$$

unde $T \in R^{2 \times 2}$ este nesusingulară, întrucât u și v sunt liniar independenți. În consecință, din (55) avem

$$(59) \quad AY = YT \cdot \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \cdot T^{-1} = YS_{11}$$

unde matricea

$$(60) \quad S_{11} = T \cdot \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \cdot T^{-1} \in R^{2 \times 2}$$

are valorile proprii complex conjugate $\lambda_{1,2}$. Prin urmare din (56), (59) avem

$$(61) \quad \begin{cases} Y^T AY = S_{11} \\ Z^T AY = 0 \end{cases},$$

de unde

$$(62) \quad A_1 = Q_1^T A Q_1 = \begin{bmatrix} S_{11} & Y^T AZ \\ 0 & Z^T AZ \end{bmatrix} = \begin{bmatrix} S_{11} & G \\ 0 & B \end{bmatrix}$$

inițiindu-se reducerea la forma Schur reală prin punerea în evidență a unui bloc diagonal 2×2 .

Procesul de deflație inițializat în (51) sau (62) poate fi continuat acționând ca mai sus asupra matricei B , ceea ce este echivalent cu acțiunea asupra matricei A_1 prin transformări ortogonale de asemănare definite

de matrice de forma $Q_2 = \text{diag}(I, \overline{Q}_2)$. În final, după q pași de deflație se obține forma Schur reală. Matricea cumulată a transformărilor ortogonale este dată de

$$(63) \quad Q = Q_1 Q_2 \dots Q_q.$$

Teorema 9.4 este demonstrată.

În continuare ne vom limita universul de investigație la clasa de matrice predominantă în practica numerică și anume cea a matricelor reale. Dacă în cazul complex apar numai deflații de tipul a), în cazul real avantajul principal constă în faptul că forma Schur reală a unei matrice reale poate fi calculată utilizând exclusiv o aritmetică reală. În acest fel, în procesul de calcul al valorilor proprii ale unei matrice reale numerele complexe apar numai în final, la calculul valorilor proprii ale blocurilor diagonale 2×2 .

Aplicarea procedurii de deflație presupune cunoașterea la fiecare pas al procedurii, a unui vector propriu de normă unitară (respectiv a unei perechi complex conjugate de vectori proprii) asociat valorii proprii corespunzătoare. Calculul unui vector propriu, în absența informației privitoare la valoarea proprie asociată, nu este posibil, pentru matrice de ordin superior lui patru, printr-o secvență finită de operații elementare, pentru că s-ar contrazice un rezultat fundamental din algebră conform căruia rezolvarea ecuațiilor de grad superior lui patru nu este posibilă printr-un număr finit de operații aritmetice elementare și extrageri de radical.

De aceea procedura de deflație trebuie să fie completată cu o procedură de determinare a unui vector propriu, fără cunoașterea valorii proprii asociate. O astfel de procedură va furniza, inerent, o aproximație mai mult sau mai puțin precisă a vectorului propriu, atât datorită erorilor de trunchiere, cât și erorilor de rotunjire. În acest context este important de știut care este cea mai bună aproximație a valorii proprii asociate unui vector propriu cunoscut aproximativ. Fie $x \in R^n$ un astfel de vector propriu aproximativ și

considerăm drept cea mai bună aproximație a valorii proprii asociate soluția în sensul celor mai mici pătrate a sistemului de n ecuații și necunoscuta scalară μ

$$(64) \quad x\mu = Ax,$$

i.e. acel μ care minimizează norma euclidiană a rezidului de ecuație $r(\mu) = Ax - \mu x$. Conform celor arătate în capitolul 3 avem

$$\mu = x^+ Ax = (x^T x)^{-1} x^T \cdot Ax.$$

Numărul

$$(65) \quad \mu = \frac{x^T Ax}{x^T x},$$

se numește *câțul Rayleigh* asociat perechii (A, x) și joacă un rol esențial în calculul (aproximativ) al valorilor proprii. Pentru început să observăm că, aplicarea procedurii de deflație, cu transformarea (48), în care x_1 este un vector propriu aproximativ, de normă unitară, conduce la

$$(66) \quad A_1 = Q_1^T A Q_1 = \begin{bmatrix} \mu & g_1^T \\ h & B \end{bmatrix}$$

având μ dat de (65) și vectorul $h = Y^T A x_1$ de normă euclidiană “mică”. Prin urmare, aplicarea transformării (48) reprezintă, în condițiile date, un pas de deflație optim întrucât μ este cea mai bună aproximație de care dispunem. O concluzie asemănătoare se poate stabili și în cazul unei perechi de valori proprii complex conjugate.

Observațiile de mai sus evidențiază necesitatea unei proceduri de calcul a unui vector propriu. Dintre metodele cele mai folosite în acest scop sunt tehnicile iterative cunoscute sub denumirile de *metoda puterii* și *metoda puterii inverse*.