Numerička analiza
23. predavanje

Saša Singer

singer@math.hr

web.math.hr/~singer

PMF – Matematički odjel, Zagreb

Sadržaj predavanja

- Svojstveni problem – osnovna sredstva za rad:
  - Kanonske forme.
    - Gornjetrokutasta i blok-gornjetrokutaste forma.
    - Jordanova forma.
    - Schurova forma i realna Schurova forma.
- QR algoritam:
  - Svođenje na Hessenbergovu formu. Trodijagonalizacija.
  - Osnovni QR algoritam.
  - Nesimetrični QR algoritam.

Kanonske forme

Svojstvene vrijednosti

Prvo uvedimo malo netipičnu definiciju svojstvenih vrijednosti matrice, koja ne uključuje svojstvene vektore.

Definicija. Neka je A kvadratna matrica reda n. Polinom

$$p(\lambda) = \det(A - \lambda I)$$

zove se karakteristični polinom matrice A.

Nultočke (korijeni karakterističnog polinoma) su svojstvene vrijednosti matrice A.

Standardna definicija uključuje samo desne svojstvene vektore, dok ova definicija omogućava i definiciju lijevih svojstvenih vektora.

Svojstveni vektori

Sad možemo definirati i svojstvene vektore.

Definicija. Vektor $x \neq 0$ koji zadovoljava

$$Ax = \lambda x$$

je desni svojstveni vektor matrice A za svojstvenu vrijednost λ, a vektor $y \neq 0$ koji zadovoljava

$$y^* A = \lambda y^*$$

je lijevi svojstveni vektor matrice A za svojstvenu vrijednost λ.

Ove definicije nisu u suprotnosti sa standardnom definicijom.

Standardna definicija

Standardna definicija. Za par (λ, x), reći ćemo da je svojstveni par (tj. x je svojstveni vektor, a λ je svojstvena vrijednost) matrice A ako vrijedi

$$Ax = \lambda x, \quad x \neq 0.$$

Jednostavnim prebacivanjem desne strane na lijevu dobivamo

$$(A - \lambda I)x = 0,$$

pa će taj linearni sustav imati netrivijalno rješenje x, ako i samo ako je matrica sustava singularna, što možemo provjeriti računanjem determinante.

Još malo o svojstvenim vrijednostima

Uočmo neke činjenice vezane uz svojstvene vrijednosti:

- svaka matrica A reda n ima točno n svojstvenih vrijednosti (osnovni teorem algebre);
- realna matrica može imati kompleksne svojstvene vrijednosti i takve uvijek dolaze u konjugirano-kompleksnim parovima.

Primjer. Matrica

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

je realna, ali su njezine svojstvene vrijednosti $\pm i$.

Desni i lijevi svojstveni vektori mogu biti jednaki. Na primjer, jednaki su za hermitske matrice.

Kako računati svojstvene vrijednosti i vektore?

Iz uvodnih definicija nije jasno kako treba računati svojstvene vrijednosti, ni svojstvene vektore.

- Naizgled jednostavno rješenje je rješavanje nelinearne jednadžbe

  $$p(\lambda) = 0.$$

  Međutim to se nikad tako ne radi, jer umjesto realnih svojstvenih vrijednosti možemo dobiti kompleksne.

Primjer. Pokušajte numerički izračunati nultočke polinoma

$$p(x) = (x - 1) \cdot (x - 2) \cdots (x - 19) \cdot (x - 20),$$

napisanoga po potencijama od x.

“Dobre” forme

Postavlja se pitanje u koju formu treba dovesti matricu A da joj lako pročitamo svojstvene vrijednosti.

Najjednostavnija forma je dijagonalna forma, ali ne mogu se sve matrice dijagonalizirati. Na primjer, matrica

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

sigurno se ne može dijagonalizirati, jer bi njezina dijagonalna forma trebala biti

$$\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

“Dobre” forme (nastavak)

Ipak i matrica A je u “dobroj formi”!

- Svojstvene vrijednosti trokutastih matrica pišu na dijagonali!

Ideja. Treba pronaći transformacije, koje općenitu matricu

- dovode u trokutastu formu,
- ne mijenjaju joj svojstvene vrijednosti.

Čak ni to nije dovoljno za “sreću”:

- Ako matrica ima kompleksnih svojstvenih vrijednosti, one će morati pisati na njezinoj dijagonali.
- Moramo koristiti kompleksne transformacije za realne matrice!

Blok-gornjetrokutasta forma

Umjesto trokutaste forme, matrica se može dovesti i na realnu blok-gornjetrokutastu formu koja ima

- 1×1 blokove na dijagonali za realne svojstvene vrijednosti,
- 2×2 blokove na dijagonali za konjugirano-kompleksne parove svojstvenih vrijednosti.

Takva blok-gornjetrokutasta matrica izgleda ovako:

$$T = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1b} \\ & A_{22} & \cdots & A_{2b} \\ & & \ddots & \vdots \\ & & & A_{bb} \end{bmatrix}.$$

Sličnosti

Determinanta blok-gornjetrokutaste matrice T je

$$\det(T - \lambda I) = \prod_{k=1}^{b} \det(A_{kk} - \lambda I),$$

pa smo izbjegli korištenje kompleksne aritmetike.

Osnovne transformacije kojima se matrice dovode na takve kanonske forme su sličnosti.

Definicija: Neka je S proizvoljna nesingularna matrica, a A i B su takve da vrijedi

$$B = S^{-1}AS.$$

Matrice A i B zovemo sličnim matricama, a S sličnošću.

Svojstva sličnosti

Propozicija. Ako su matrice A i B slične, tj. ako je

$$B = S^{-1}AS,$$

matrice A i B imaju jednake svojstvene vrijednosti.

Vektor x je desni svojstveni vektor matrice A, ako i samo ako je $S^{-1}x$ desni svojstveni vektor matrice B.

Nadalje, y je lijevi svojstveni vektor matrice A, ako i samo ako je S^*y lijevi svojstveni vektor matrice B.

Svojstva sličnosti

Dokaz. Za kvadratne matrice vrijedi Binet–Cauchyjev teorem

$$\det(X \cdot Y) = \det X \cdot \det Y,$$

pa imamo

$$\begin{aligned}
\det(B - \lambda I) &= \det(S^{-1}AS - \lambda S^{-1}S) \\
&= \det(S^{-1}(A - \lambda I)S) \\
&= \det(S^{-1}) \det(A - \lambda I) \det S \\
&= \det(A - \lambda I).
\end{aligned}$$

Time smo pokazali da A i B imaju isti svojstveni polinom, pa im i svojstvene vrijednosti moraju biti jednake.

Svojstva sličnosti

Pokažimo još i vezu svojstvenih vekora. Ako je

$$Ax = \lambda x,$$

onda iz $B = S^{-1}AS$ izlazi da je $A = SBS^{-1}$, pa mora vrijediti

$$SBS^{-1}x = \lambda x.$$

Množenjem prethodne relacije slijeva sa S^{-1}, dobivamo

$$B(S^{-1}x) = \lambda(S^{-1}x).$$

Čitanjem “natraške” dobivamo i drugu implikaciju.

Za lijeve svojstvene vektore rezultat se dobiva na sličan način.

Svojstva sličnosti

Neka je

$$y^* A = \lambda y^*.$$

Tada je

$$y^* S B S^{-1} = \lambda y^*.$$

Množenjem zdesna sa S dobivamo

$$(y^* S) B = \lambda (y^* S),$$

pa zaključujemo da je $S^* y$ lijevi svojstveni vektor za B.

Obratnim čitanjem dobijemo drugi smjer dokaza. ■

Vrijedi li obrat?

Problem. Jesu li matrice koje imaju iste svojstvene vrijednosti slične, tj. vrijedi li na neki način obrat gornjeg teorema?

Obrat ne vrijedi! Na primjer, matrice

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{i} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

imaju dvostruku svojstvenu vrijednost 1.

- A ima samo jedan desni svojstveni vektor e_1 i samo jedan lijevi svojstveni vektor e_2.
- B ima dva desna (e_1 i e_2) i dva lijeva svojstvena vektora (e_1 i e_2).

Zaključak. Matrice A i B ne mogu biti slične.

Jordanova forma

Jedan od osnovnih teorema koji daje kanonsku strukturu matrice je Jordanov teorem, koji, gotovo u potpunosti,

- karakterizira formu matrice najbližu dijagonalnoj,

za matrice koje se ne mogu dijagonalizirati.

Teorem (Jordanova forma). Neka je A kvadratna matrica reda n. Tada postoji nesingularna matrica S, takva da je

$$S^{-1}AS = J,$$

pri čemu je J tzv. Jordanova forma, tj. J je blok-dijagonalna,

$$J = \operatorname{diag}(\, J_{n_1}(\lambda_1),\ J_{n_2}(\lambda_2),\ \ldots\ ,\ J_{n_k}(\lambda_k)\,),$$

gdje je $J_{n_i}(\lambda_i)$ matrica reda n_i, ...

Jordanova forma (nastavak)

… oblika

$$J_{n_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}.$$

Matrica J je jedinstvena, do na permutaciju dijagonalnih blokova.

Svaki J_{n_i} zove se Jordanov blok i njegova svojstvena vrijednost λ_i ima algebarsku kratnost n_i.

Zbroj svih redova n_i za iste vrijednosti λ_i zove se algebarska kratnost svojstvene vrijednosti $\lambda = \lambda_i$.

Dijagonalizibilnost, defektne matrice, ...

Navedimo još neke činjenice koje proizlaze iz Jordanove forme.

- Ako je $n_i = 1$ i λ_i se pojavljuje u samo jednom bloku, onda je λ_i jednostruka svojstvena vrijednost.
- Ako su sve svojstvene vrijednosti u matrici J jednostruke, onda je J dijagonalna, tj. A se može dijagonalizirati.
- Ako postoji barem jedan blok J_{n_i} reda barem 2, A je defektna matrica.
- Bitno svojstvo defektnih matrica je da nemaju punu bazu svojstvenih vektora.
- Broj svojstvenih vektora za svojstvenu vrijednost λ zove se geometrijska kratnost svojstvene vrijednosti.

Koje se matrice mogu dijagonalizirati?

Ipak, mnoge se matrice mogu dijagonalizirati.

Posljedica Jordanovog teorema je da se sigurno mogu dijagonalizirati

- matrice koje imaju samo različite svojstvene vrijednosti.

Nadalje, postoje i matrice, koje možda imaju iste svojstvene vrijednosti, a još uvijek se mogu dijagonalizirati.

- Najšira klasa takvih matrica su normalne matrice, tj. matrice za koje vrijedi

$$N^* N = N N^*.$$

Primjer

Za matricu

$$A = \begin{bmatrix} \frac{7}{4} & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 2 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{5}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{5}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{11}{4} & \frac{1}{4} \\ -\frac{3}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{7}{2} \end{bmatrix}$$

su ...

Primjer

... matrice J i S jednake

$$J = \left[\begin{array}{cc|ccc|c} 2 & 1 & & & & \\ & 2 & & & & \\ \hline & & 3 & 1 & & \\ & & & 3 & 1 & \\ & & & & 3 & \\ \hline & & & & & 2 \end{array}\right], \quad S = \begin{bmatrix} 1 & & & & & 1 \\ & 1 & & & & 1 \\ & & 1 & & & 1 \\ & & & 1 & & 1 \\ & & & & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Primijetite da 2 nije jednostruka, nego trostruka svojstvena vrijednost matrice A.

Puna baza svojstvenih vektora

Propozicija. Jordanov blok $J_{n_i}(\lambda_i)$ ima

- samo jedan desni svojstveni vektor, i to je e_1,
- i samo jedan lijevi svojstveni vektor, i to je e_{n_i}.

Dokaz se provodi računanjem x iz

$$(J_{n_i}(\lambda_i) - \lambda I)x = 0$$

za desni svojstveni vektor. Na sličan način se dobije i za lijevi svojstveni vektor. ■

Prema prethodnoj propoziciji odmah slijedi da se, čim postoji Jordanov blok veličine barem 2, matrica se ne može dijagonalizirati!

Jordanova forma — da ili ne?

Ipak, Jordanova forma je nekorisna u numeričkom računanju.

- Neka je $J_n(0)$ Jordanov blok reda n za svojstvenu vrijednost 0,

$$J_n(0) = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

Ako, za $i = 1, \ldots, n$, na i-tom mjestu na dijagonali dodamo $i \cdot \varepsilon$, pri čemu je $0 < \varepsilon \ll 1$, onda smo tom malom perturbacijom potpuno promijenili karakter matrice.

Jordanova forma — da ili ne?

- Umjesto $J_n(0)$, imamo

$$J_n'(0) = J_n(0) + \delta J_n(0) = \begin{bmatrix} \varepsilon & 1 & & & \\ & 2\varepsilon & 1 & & \\ & & \ddots & \ddots & \\ & & & (n-1)\varepsilon & 1 \\ & & & & n\varepsilon \end{bmatrix}.$$

Jordanova forma te malo perturbirane matrice je dijagonalna,

$$J_n'(0) = \operatorname{diag}(\,\varepsilon,\, 2\varepsilon,\, \ldots\,,\, (n-1)\varepsilon,\, n\varepsilon\,).$$

Jordanova forma — da ili ne?

Drugi razlog što se ne koristi Jordanova forma je:

- ona se ne može stabilno izračunati, tj. kad završimo s računanjem S i J, ne može se garantirati da vrijedi

$$S^{-1}(A + \delta A)S = J$$

za neki mali δA.

Umjesto Jordanove forme, u numeričkom se računanju koristi Schurova forma. Glavno svojstvo:

- Umjesto nesingularne matrice S, koja može biti proizvoljno loše uvjetovana,
- koristi se unitarna matrica Q, samo dobivena forma više nije kompaktna.

Schurova forma

Teorem (Schurova kanonska forma). Neka je dana matrica A reda n. Tada postoje unitarna matrica Q i gornjetrokutasta matrica T, takve da vrijedi

$$Q^* A Q = T.$$

Svojstvene vrijednosti matrice A su dijagonalni elementi matrice T.

Dokaz. Provodi se indukcijom po redu matrice A. Za bazu indukcije uzmemo $n = 1$. Tada, očito, možemo uzeti $Q = Q^* = 1$.

Schurova forma (nastavak)

Korak indukcije. Neka je λ proizvoljna svojstvena vrijednost, a u pripadni svojstveni vektor takav da je $\|u\|_2 = 1$. Svojstveni vektor u se može dopuniti matricom $\widetilde{U}$ do unitarne matrice U,

$$U = [u, \widetilde{U}].$$

Tada vrijedi

$$U^* A U = \begin{bmatrix} u^* \\ \widetilde{U}^* \end{bmatrix} A \, [u, \widetilde{U}] = \begin{bmatrix} u^* A u & u^* A \widetilde{U} \\ \widetilde{U}^* A u & \widetilde{U}^* A \widetilde{U} \end{bmatrix}.$$

Budući da je u jedinični svojstveni vektor za λ, imamo

$$u^* A u = u^* \lambda u = \lambda \|u\|_2^2 = \lambda.$$

Schurova forma (nastavak)

Nadalje, vrijedi

$$\widetilde{U}^* A u = \widetilde{U}^* \lambda u = \lambda \cdot 0 = 0,$$

jer su stupci matrice $\widetilde{U}$ ortogonalni na u. Prema tome, relaciju možemo pisati kao

$$U^* A U = \begin{bmatrix} \lambda & u^* A \widetilde{U} \\ 0 & \widetilde{U}^* A \widetilde{U} \end{bmatrix}.$$

Ako vrijedi pretpostavka indukcije, da svaku matricu reda $n-1$ možemo napisati u traženoj formi, onda to možemo napraviti i s matricom $\widetilde{U}^* A \widetilde{U}$.

Schurova forma (nastavak)

Dakle,

$$\widetilde{U}^* A \widetilde{U} = P^* \widetilde{T} P,$$

pri čemu je P unitarna matrica, a $\widetilde{T}$ gornjetrokutasta.
Uvrštavanjem u prethodnu jednadžbu, imamo

$$U^* A U = \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} \lambda & u^* A \widetilde{U} P \\ 0 & \widetilde{T} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & P^* \end{bmatrix},$$

čime je teorem dokazan. ■

Uočimo da Schurova forma nije jedinstvena, jer svojstvene vrijednosti se na dijagonali mogu pojaviti u bilo kojem poretku.

Realna Schurova forma

Nadalje, Schurova forma može voditi na kompleksnu gornjetrokutastu matricu T, čak i onda kad je A realna.

Zbog efikasnosti računanja uvodi se realna Schurova forma.

Teorem (Realna Schurova forma). Ako je A realna matrica reda n, postoji ortogonalna matrica V takva da je

$$V^T AV = T,$$

gdje je T realna blok-gornjetrokutasta s dijagonalnim blokovima 1×1 ili 2×2. Njezine svojstvene vrijednosti su svojstvene vrijednosti dijagonalnih blokova, s time da realne svojstvene vrijednosti odgovaraju 1×1 blokovima, a konjugirano-kompleksni parovi odgovaraju 2×2 blokovima.

Realna Schurova forma

Dokaz. Ide indukcijom, slično kao u prethodnom teoremu.

Ako je λ realna svojstvena vrijednost, korak u dokazu ekvivalentan je onome u prethodnom teoremu.

Prema tome, pretpostavimo da je λ kompleksna svojstvena vrijednost. Tada je njezin svojstveni vektor u nužno kompleksan. Neka je

$$Au = \lambda u.$$

Iz prethodne relacije, konjugiranjem dobivamo

$$\begin{aligned} \overline{Au} &= \overline{A}\bar{u} = (A \text{ je realna}) = A\bar{u} \\ &= \overline{\lambda u} = \bar{\lambda}\bar{u}. \end{aligned}$$

Iz desnih strana slijedi da je $(\bar{\lambda}, \bar{u})$, također, svojstveni par.

Realna Schurova forma

Neka je u_R realni, a u_I imaginarni dio vektora u:

$$u_R = \frac{1}{2}(u + \bar{u}), \quad u_I = \frac{1}{2i}(u - \bar{u}).$$

Očito, tada vrijedi da

- u i $\bar{u}$ razapinju isti potprostor kao i u_R i u_I.

To je dvodimenzionalni potprostor za koji tvrdimo da je invarijantni potprostor.

Za potprostor X reći ćemo da je invarijantan za A ako vrijedi

$$AX \subseteq X,$$

tj. ako za svaki $x \in X$ vrijedi da je $Ax \in X$.

Realna Schurova forma

Svaki vektor $x \in X$, $X = \operatorname{span}\{u, \bar{u}\}$, možemo napisati kao

$$x = \alpha u + \beta \bar{u},$$

pri čemu su α i β skalari. Tada vrijedi

$$Ax = A(\alpha u + \beta \bar{u}) = \alpha \lambda u + \beta \bar{\lambda} \bar{u},$$

pa je to opet linearna kombinacija vektora iz X.

Označimo realnu matricu

$$\widetilde{U} = [u_R, u_I],$$

i napravimo (skraćenu) QR faktorizaciju matrice $\widetilde{U}$ (onu koja ima samo dva stupca).

Realna Schurova forma

Tada vrijedi

$$\mathrm{span}\{Q\} = \mathrm{span}\{u_R, u_I\},$$

pa je i to invarijantan potprostor. Izaberimo $\widetilde{Q}$ tako da Q dopunimo do ortogonalne baze U, tj.

$$U = [Q, \widetilde{Q}].$$

Nadalje,

$$U^T A U = \begin{bmatrix} Q^T \\ \widetilde{Q}^T \end{bmatrix} A \, [Q, \widetilde{Q}] = \begin{bmatrix} Q^T A Q & Q^T A \widetilde{Q} \\ \widetilde{Q}^T A Q & \widetilde{Q}^T A \widetilde{Q} \end{bmatrix}.$$

Realna Schurova forma

Budući da Q razapinje invarijantni potprostor, postoji 2×2 matrica B takva da vrijedi

$$AQ = QB.$$

Onda blokove u relaciji za $U^T AU$ možemo napisati kao

$$Q^T AQ = Q^T QB = B, \qquad \widetilde{Q}^T AQ = \widetilde{Q}^T QB = 0.$$

Prema tome, izlazi

$$U^T AU = \begin{bmatrix} B & Q^T A\widetilde{Q} \\ 0 & \widetilde{Q}^T A\widetilde{Q} \end{bmatrix}.$$

Sada iskoristimo korak indukcije na matrici $\widetilde{Q}^T A\widetilde{Q}$. ■

Čitanje svojstvenih vektora

Ako smo matricu doveli u Schurovu formu, kako ćemo pronaći njezine svojstvene vektore?

Iz sličnosti znamo da su svojstveni vektori matrica T i A jednostavno vezani, tj. ako je

$$Tx = \lambda x,$$

onda je

$$AQx = QTx = \lambda Qx,$$

pa je Qx svojstveni vektor matrice A.

Ostaje nam samo pročitati svojstvene vektore matrice T iz njezine trokutaste forme.

Čitanje svojstvenih vektora

Jednostavnosti radi, pretpostavimo da tražimo svojstveni vektor za jednostruku svojstvenu vrijednost $\lambda = t_{ii}$.

Linearni sustav za svojstvene vektore glasi $(T - \lambda I)x = 0$, odnosno, u blok formi

$$0 = \begin{bmatrix} T_{11} - \lambda I & T_{12} & T_{13} \\ 0 & 0 & T_{23} \\ 0 & 0 & T_{33} - \lambda I \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

pri čemu je

- T_{11} matrica reda $i - 1$,
- $T_{22} = 0$ reda 1,
- T_{33} reda $(n - i)$.

Čitanje svojstvenih vektora

Blok vektori x_1, x_2 i x_3, redom, odgovaraju dimenzijama matrica T_{11}, T_{22} i T_{33}.

Budući da je λ jednostruka, onda su matrice $T_{33} - \lambda I$ i $T_{11} - \lambda I$ nesingularne, pa iz posljednje blok-jednadžbe slijedi da mora biti

$$x_3 = 0,$$

a iz prve blok jednadžbe slijedi da je

$$(T_{11} - \lambda I)x_1 = -T_{12}x_2.$$

Broj x_2 možemo proizvoljno odabrati, tako da vektor x nije nul-vektor.

Čitanje svojstvenih vektora

Recimo, uzmimo $x_2 = 1$. Tada izlazi da je

$$x_1 = -(T_{11} - \lambda I)^{-1} T_{12}.$$

Prema tome, svojstveni vektor za svojstvenu vrijednost λ je

$$x = \begin{bmatrix} -(T_{11} - \lambda I)^{-1} T_{12} \\ 1 \\ 0 \end{bmatrix}.$$

Drugim riječima, da bismo dobili svojstveni vektor x, moramo riješiti trokutasti linearni sustav za x_1.

Računanje svojstvenog vektora za kompleksnu svojstvenu vrijednost zahtijeva rješavanje blok-gornjetrokutastog sustava.

Svođenje na Hessenbergovu i trodijagonalnu formu

Podjela svojstvenog problema

Ima mnogo podjela algoritama prema vrsti svojstvenog problema. Navedimo neke od njih. Postoje algoritmi za:

- (puni) simetrični svojstveni problem,
- (puni) nesimetrični svojstveni problem,
- za matrične parove (engl. matrix pairs, pencils)

  $$Kx = \lambda Mx,$$

  (najčešće je jedna od matrica, M, pozitivno definitna),
- svojstveni problem za velike rijetke matrice (i simetrične i nesimetrične),
- matrice sa specijalnom strukturom (antisimetične, palindromne, hamiltonovske, ...).

Podjela svojstvenog problema

Nadalje, algoritmi se dijele i prema tome

- koliko svojstvenih vrijednosti želimo izračunati.

Postoje algoritmi koji računaju:

- sve svojstvene vrijednosti,
- najmanju/najveću svojstvenu vrijednost,
- ili nekoliko najmanjih/najvećih svojstvenih vrijednosti (po apsolutnoj vrijednosti),
- te svojstvene vrijednosti najbliže nekom zadanom broju.

Za hermitske matrice:

- sve svojstvene vrijednosti iz nekog intervala — tipična primjena je na matricama u tridijagonalnoj formi.

Svojstveni problem za pune matrice

Algoritmi i za (puni) simetrični i za nesimetrični svojstveni problem, najčešće, zbog efikasnosti

- rade na matricama u "kondenziranoj" formi.

Iznimka je

- Jacobijev algoritam u hermitskom slučaju, odnosno, norm-reducirajući algortimi tipa Eberlein u nesimetričnom slučaju.

Kondenzirana forma za nesimetrični problem je

- Hessenbergova forma,

koja se u simetričnom slučaju svodi na

- trodijagonalnu formu.

Hessenbergova forma

Definicija. Za matricu A ćemo reći da je u gornjoj Hessenbergovoj formi ako vrijedi

$$a_{ij} = 0, \quad \text{za} \quad i - j \geq 2,$$

tj. matrica A je gornjetrokutasta i ima još jednu sporednu dijagonalu ispod glavne.

Svođenje na Hessenbergovu formu provodi se

- korištenjem unitarnih sličnosti, tj. traži se unitarna matrica Q takva da je QAQ^* gornja Hessenbergova.
- Ideja redukcije vrlo je slična QR faktorizaciji, samo se norma stupca, s jedne strane, umjesto na dijagonalni element, "nabacuje" na element ispod glavne dijagonale.

Kako do Hessenbergove forme?

Primjer. Pokažimo to na primjeru realne 4×4 matrice A:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

Korištenjem Householderovog reflektora H_1

$$Q_1 = \begin{bmatrix} 1 & \\ & H_1 \end{bmatrix},$$

možemo poništiti elemente prvog stupca u matrici A, od mjesta 3 do n (ovdje je $n = 4$),

Kako do Hessenbergove forme?

pa imamo

$$Q_1 A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ 0 & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{bmatrix}.$$

Nakon toga primijenimo matricu Q_1^T zdesna, pa dobijemo

$$Q_1 A Q_1^T = \begin{bmatrix} a_{11} & a_{12}^{(2)} & a_{13}^{(2)} & a_{14}^{(2)} \\ a_{21}^{(1)} & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} & a_{34}^{(2)} \\ 0 & a_{42}^{(2)} & a_{43}^{(2)} & a_{44}^{(2)} \end{bmatrix}.$$

Kako do Hessenbergove forme?

Nakon toga biramo ortogonalnu matricu

$$Q_2 = \begin{bmatrix} I_2 & \\ & H_2 \end{bmatrix},$$

takvu da poništi elemente drugog stupca koji se nalaze od mjesta 4 do n (ovdje je $n = 4$),

$$Q_2 Q_1 A Q_1^T = \begin{bmatrix} a_{11} & a_{12}^{(2)} & a_{13}^{(2)} & a_{14}^{(2)} \\ a_{21}^{(1)} & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ 0 & a_{32}^{(3)} & a_{33}^{(3)} & a_{34}^{(3)} \\ 0 & 0 & a_{43}^{(3)} & a_{44}^{(3)} \end{bmatrix}.$$

Kako do Hessenbergove forme?

Transformaciju Q_2^T moramo primijeniti i zdesna

$$Q_2 Q_1 A Q_1^T Q_2^T = \begin{bmatrix} a_{11} & a_{12}^{(2)} & a_{13}^{(4)} & a_{14}^{(4)} \\ a_{21}^{(1)} & a_{22}^{(2)} & a_{23}^{(4)} & a_{24}^{(4)} \\ 0 & a_{32}^{(3)} & a_{33}^{(4)} & a_{34}^{(4)} \\ 0 & 0 & a_{43}^{(4)} & a_{44}^{(4)} \end{bmatrix}.$$

Time smo završili postupak.

Postupak smo mogli, umjesto Householderovim reflektorima, provesti i Givensovim rotacijama.

Simetrična Hessenbergova forma

Ako je matrica A bila simetrična, onda je simetrična i njezina Hessenbergova forma

$$Q_{n-2} \cdots Q_1 A Q_1^T \cdots Q_{n-2}^T,$$

pa je svođenje na Hessenbergovu formu, zapravo, trodijagonalizacija.

Nadalje, Householderovi reflektori su i sami simetrične matrice, pa je

$$Q_{n-2} \cdots Q_1 A Q_1^T \cdots Q_{n-2}^T = Q_{n-2} \cdots Q_1 A Q_1 \cdots Q_{n-2}.$$

Svojstvo Hessenbergove forme

Teorem. Ako je A gornja Hessenbergova, onda je Q u njezinoj QR faktorizaciji

$$A = QR,$$

također, gornja Hessenbergova matrica.

Dokaz. To se najlakše vidi ako QR faktorizaciju provodimo rotacijama, pa ona redom zahvaća

- retke 1 i 2 za poništavanje elementa u 1. stupcu,
- retke 2 i 3 za poništavanje elementa u 2. stupcu, ... ∎

Prehodni teorem garantira i da je matrica RQ gornja Hessenbergova, jer je to produkt gornjetrokutaste i gornje Hessenbergove matrice — bitno je za čuvanje strukture matrica pri korištenju QR algoritma.

Osnovni QR algoritam

Osnovni QR algoritam

QR algoritam je

- jedna je od najpoznatijih metoda za računanje svojstvenih vrijednosti matrica,
- otkrili su je nezavisno John Francis i Vera Kublanovskaja 1961–1962.

QR algoritam bez pomaka. Za danu nesimetričnu matricu A_0 iteriramo:

$i = 0$
ponavljaj
  faktoriziraj $A_i = Q_i R_i$ /* QR faktorizacija */
  $A_{i+1} = R_i Q_i$
  $i = i + 1$
do konvergencije

QR algoritam i sličnost

Propozicija. Sve matrice A_i dobivene prethodnim algoritmom su unitarno slične matrici A_0.

Dokaz. Budući da je

$$Q_i^* A_i = R_i,$$

uvrštavanjem u relaciju za A_{i+1} dobivamo

$$A_{i+1} = Q_i^* A_i Q_i. \qquad \blacksquare$$

QR algoritam je "najsretniji" kad mu damo singularnu matricu, jer je QR faktorizacija otkriva u jednom koraku. Nakon toga radi deflaciju.

Algoritmu dodati pomake (engl. shifts).

QR algoritam s pomacima

Ako su svojstvene vrijednosti matrice A jednake λ_i, onda su svojstvene vrijednosti matrice $A - \sigma I$ jednake $\lambda_i - \sigma$.

Ideja: pomak σ birati "blizu" svojstvenim vrijednostima.

Posljedica. Ubrzavanje konvergencije!

QR algoritam s pomacima. Za danu nesimetričnu matricu A_0 iteriramo:

$i = 0$
ponavljaj
  izaberi pomak σ_i /* blizu svojstvene vrijednosti */
  faktoriziraj $A_i - \sigma_i I = Q_i R_i$ /* QR faktorizacija */
  $A_{i+1} = R_i Q_i + \sigma_i I$
  $i = i + 1$
do konvergencije

QR algoritam s pomacima i sličnost

Propozicija. Matrice A_{i+1} i A_i dobivene QR iteracijama s pomakom su unitarno slične.

Dokaz. Iz

$$Q_i R_i = A_i - \sigma_i I$$

dobivamo da je

$$R_i = Q_i^* (A_i - \sigma_i I),$$

pa onda izlazi

$$A_{i+1} = R_i Q_i + \sigma_i I = (Q_i^* (A_i - \sigma_i I)) Q_i + \sigma_i I = Q_i^* A_i Q_i.$$

■

Problemi QR algoritma

Problemi koje treba uočiti:

- QR faktorizacija pune matrice traje $\mathcal{O}(n^3)$ operacija po faktorizaciji. Ako napravimo pomak jednak svojstvenoj vrijednosti, možemo ju otkriti u jednom koraku – ukupno trajanje QR iteracija bilo bi $\mathcal{O}(n^4)$.

  Rješenje: raditi na matrici u Hessenbergovoj formi.
- Kako treba izabrati pomak σ_i da bi se ubrzala konvergencija u slučaju kompleksne svojstvene vrijednosti, ako je matrica bila realna? Kompleksni pomak nije dobar izbor!
- Kako treba birati pomak u slučaju hermitskih matrica?
- Kako prepoznati konvergenciju prema realnoj Schurovoj formi?

Nesimetrični QR algoritam

Implicitni Q teorem

Za implementaciju pomaka u QR metodi i za simetrične i za nesimetrične matrice, potreban nam je sljedeći teorem.

Teorem (Implicitni Q teorem). Neka je $Q^T AQ = H$ nereducirana gornja Hessenbergova matrica. Stupci matrice Q, od drugog do n-tog, jedinstveno su određeni (do na znak) prvim stupcem matrice Q. ■

Nereducirana $=$ H nema nula na najdonjoj dijagonali.

Dokaz teorema, recimo, pogledati u knjizi: J. W. Demmel, "Applied Numerical Linear Algebra" (str. 168).

Prethodni teorem kaže kako se računa matrica Q_i u QR algoritmu.

Implicitni Q teorem — primjena

Iz A_i se izračuna sljedeća iteracija $A_{i+1} = Q_i^T A_i Q_i$ ovako:

- prvi stupac matrice Q_i je paralelan s prvim stupcem matrice $A_i - \sigma_i I$, jedino mu je norma jednaka 1, tj.
  - treba samo normirati prvi stupac od $A_i - \sigma_i I$;
- ostali stupci matrice Q_i računaju se tako da Q_i bude ortogonalna, a A_{i+1} je nereducirana Hessenbergova.

Prema implicitnom Q teoremu, tada znamo da smo A_{i+1} korektno izračunali, jer je Q_i jedinstvena do na predznak stupaca.

Jednostruki pomak

Ako su svojstvene vrijednosti jednostruke, onda bismo mogli raditi tzv. jednostruke pomake (engl. single shift).

Promotrimo algoritam na realnoj matrici reda 6. Neka je, zasad, Q_1 rotacija

$$Q_1 = \begin{bmatrix} c_1 & -s_1 & & & & \\ s_1 & c_1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix},$$

i neka je A gornja Hessenbergova.

Jednostruki pomak

Djelovanjem s Q_1^T slijeva i Q_1 zdesna, transformirana matrica A, u oznaci A_1, poprima sljedeći oblik

$$A_1 = Q_1^T A Q_1 = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ \star & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}.$$

Pritom, $\star$ označava element koji je prije transformacije bio 0, međutim, djelovanjem transformacije prestao je biti 0.

Jednostruki pomak

“Izbočinu” (engl. bulge) $\star$ treba “istjerivati” (engl. chasing the bulge) izvan matrice i to tako da slijeva uzmemo rotaciju Q_2^T u $(2,3)$ ravnini, koja će poništiti taj novostvoreni element.

Naravno, istom rotacijom (samo transponiranom) moramo djelovati i zdesna. Novi Q_2 izgleda ovako

$$Q_2 = \begin{bmatrix} 1 & & & & & \\ & c_2 & -s_2 & & & \\ & s_2 & c_2 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix}.$$

Jednostruki pomak

Nakon djelovanja s Q_2^T slijeva, dobivamo

$$A_2 = Q_2^T A_1 = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}.$$

Jednostruki pomak

Međutim, nakon primjene Q_2 zdesna, imamo

$$A_2 = Q_2^T A_1 Q_2 = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ & * & * & * & * & * \\ & \star & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix},$$

što znači da se "izbočina" pomaknula jedno mjesto dolje i desno.

Jednostruki pomak

U sljedećem koraku, poništimo tu izbočinu korištenjem Givensove rotacije u ravnini $(3, 4)$,

$$Q_3 = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & c_3 & -s_3 & & \\ & & s_3 & c_3 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix}.$$

Algoritam ponavljamo dok posljednje izbočenje ne izbacimo izvan matrice. To će se dogoditi djelovanjem rotacije Q_5 u $(5, 6)$ ravnini i matrica $A_5 = Q_5^T A_4 Q_5$ će ponovno biti u gornjoj Hessenbergovoj formi.

Jednostruki pomak

Primijetimo da je djelovanje na matricu A bilo ovim redom

$$Q^T A Q = (Q_1 Q_2 Q_3 Q_4 Q_5)^T A Q_1 Q_2 Q_3 Q_4 Q_5,$$

i da je dobivena matrica $Q^T A Q$ opet gornja Hessenbergova. Uočimo još i izgled konačne matrice Q,

$$Q = \begin{bmatrix} c_1 & * & * & * & * & * \\ s_1 & * & * & * & * & * \\ & s_2 & * & * & * & * \\ & & s_3 & * & * & * \\ & & & s_4 & * & * \\ & & & & s_5 & * \end{bmatrix}.$$

Jednostruki pomak

Prvi stupac od Q određuje, do na znak, sve ostale stupce u Q (implicitni Q teorem). Prvi stupac matrice Q izabrat ćemo tako da je

- proporcionalan s prvim stupcem $A - \sigma I$.

To znači da je naš Q jednak onome koji bi bio dobiven QR faktorizacijom od $A - \sigma I$.

Ipak, u praksi se obično ne koriste jednostruki pomaci, jer dvostruki pomaci imaju bolja svojstva (brža konvergencija).

Dvostruki pomak

Ako je matrica $A = A_0$ realna, a ima kompleksnih svojstvenih vrijednosti, treba napraviti pomak za σ i $\bar{\sigma}$

$$
\begin{aligned}
A_0 - \sigma I &= Q_1 R_1, \\
A_1 &= R_1 Q_1 + \sigma I, \\
A_1 - \bar{\sigma} I &= Q_2 R_2, \\
A_2 &= R_2 Q_2 + \bar{\sigma} I.
\end{aligned}
$$

Odatle odmah izlazi da je $A_2 = Q_2^T Q_1^T A_0 Q_1 Q_2$.

Lema. Matrice Q_1 i Q_2 možemo izabrati tako da vrijedi

- $Q_1 Q_2$ je realna, pa je i A_2 realna,
- prvi stupac matrice $Q_1 Q_2$ se lako računa.

Dvostruki pomak

Dokaz. Redom, iz prve dvije transformacije izlazi

$$\begin{aligned} Q_1Q_2R_2R_1 &= Q_1(A_1 - \bar{\sigma}I)R_1 = Q_1(R_1Q_1 + (\sigma - \bar{\sigma})I)R_1 \\ &= Q_1R_1Q_1R_1 + (\sigma - \bar{\sigma})Q_1R_1 \\ &= (A_0 - \sigma I)^2 + (\sigma - \bar{\sigma})(A_0 - \sigma I) \\ &= A_0^2 - (\sigma + \bar{\sigma})A_0 + |\sigma|^2 I =: M. \end{aligned}$$

Zbog $\sigma + \bar{\sigma} = 2\,\mathrm{Re}\,\sigma \in \mathbb{R}$, prethodna matrica M je realna.

- Onda je $Q_1Q_2R_2R_1$ QR faktorizacija realne matrice. Zato Q_1Q_2 i R_2R_1 možemo izabrati kao realne matrice.
- Prvi stupac matrice Q_1Q_2 je proporcionalan prvom stupcu matrice M, a ostali stupci se računaju korištenjem implicitnog Q teorema. ∎

Dvostruki pomak

I ovaj algoritam natjerava “izbočinu”, samo je ona, zbog dvostrukog pomaka, 2×2 izbočina.

Razlog: matrica $M = (A - \sigma I)^2 + (\sigma - \bar{\sigma})(A - \sigma I)$ ima jednu donju sporednu dijagonalu više od A, zbog kvadriranja.

Na primjer, za matricu A reda 6, to izgleda ovako:

$$A = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}, \quad M = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \end{bmatrix}.$$

Dvostruki pomak — “gonjenje izbočine”

Ovdje zahtijevamo da je prvi stupac od $Q_1 Q_2$ proporcionalan prvom stupcu matrice M (lako se računa iz A).

Neka je A ponovno matrica reda 6. Prvu ortogonalnu transformaciju $\widetilde{Q}_1^T$ biramo tako da je njezin prvi stupac proporcionalan prvom stupcu matrice M (isto kao i ranije). Dobivamo

$$\widetilde{Q}_1^T A = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ \star & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}$$

Dvostruki pomak — "gonjenje izbočine"

i

$$\widetilde{Q}_1^T A \widetilde{Q}_1 = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ \star & * & * & * & * & * \\ \star & \star & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}.$$

Sljedeća transformacija $\widetilde{Q}_2^T$ poništava elemente $(3, 1)$ i $(4, 1)$, "nabacujući" njihovu normu na element $(2, 1)$.

Ona mijenja samo retke 2, 3 i 4 u gornjoj matrici.

Dvostruki pomak — "gonjenje izbočine"

Transformaciju $\widetilde{Q}_2^T$ možemo realizirati kao

- par Givensovih rotacija, ili
- "kratki" Householderov reflektor (netrivijalni elementi u vektoru w su w_2, w_3 i w_4).

Dobivamo

$$\widetilde{Q}_2^T \widetilde{Q}_1^T A \widetilde{Q}_1 = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ & * & * & * & * & * \\ & \star & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}$$

Dvostruki pomak — "gonjenje izbočine"

Nakon primjene $\widetilde{Q}_2$ zdesna, izlazi

$$\widetilde{Q}_2^T \widetilde{Q}_1^T A \widetilde{Q}_1 \widetilde{Q}_2 = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ & * & * & * & * & * \\ & \star & * & * & * & * \\ & \star & \star & * & * & * \\ & & & & * & * \end{bmatrix}.$$

Dakle, trokutasta 2×2 "izbočina" se pomaknula jedno mjesto dolje i desno.

Sada, vrlo slično kao u jednostrukom pomaku, "izbočinu" pomičemo prema kraju matrice, sve dok ju ne izbacimo.

Izbor pomaka

Najčešće se pomak (engl. shift) bira kao svojstvene vrijednosti 1×1 ili 2×2 podmatrice u donjem desnom kutu matrice A:

- ako imamo jednostruki pomak, onda uzmemo $\sigma = a_{nn}$,
- ako imamo dvostruki pomak, onda uzmemo tzv. Francisov pomak, tj. σ i $\overline{\sigma}$ su svojstvene vrijednosti matrice
$$\begin{bmatrix} a_{n-1,n-1} & a_{n-1,n} \\ a_{n,n-1} & a_{n,n} \end{bmatrix}.$$

Izbor pomaka

Nažalost, ovakav dvostruki pomak, iako je izvrstan u praksi, može ne raditi. Na primjer, matricu

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

ostavlja na miru.

Izbor pomaka za simetrični trodijagonalni QR

Za simetrični trodijagonalni QR algoritam, pomak se bira kao

- a_{nn} (kubična konvergencija, ali ne konvergira uvijek),
- svojstvena vrijednost bloka

$$\begin{bmatrix} a_{n-1,n-1} & a_{n-1,n} \\ a_{n-1,n} & a_{n,n} \end{bmatrix}$$

koja je najbliža $a_{n,n}$. Taj izbor pomaka zove se Wilkinsonov pomak. Može se pokazati da su simetrične QR iteracije s tim pomakom kubično konvergentne za gotovo sve matrice A.