SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODJEL

Sanja Singer

Indefinitna QR dekompozicija i primjene

Disertacija

Zagreb, 1997.

Zahvaljujem se mojim suvoditeljima prof. dr. Krešimiru Veseliću i prof. dr. Vjeranu Hariju na konstruktivnoj pomoći pri izradi ove radnje.

Rad na ovoj temi započeo je za vrijeme mog boravka kod prof. Veselića u Hagenu. Moram istaknuti nepresušno vrelo ideja i novih rezultata, kojima me zasipao prof. Veselić. Prof. Hari je pokazivao nevjerojatnu strpljivost u poboljšavanju kvalitete izričaja radnje, posebno tehnički vrlo kompliciranih teorema.

Posebnu zahvalnost dugujem prof. dr. Beresfordu Parlettu, koji je na Internacionalnoj radionici u Splitu odvojio dio svog vremena i otkrio mi neobjavljene (uglavnom negativne) rezultate do kojih je došao zajedno sa svojim studentima na Sveučilištu Berkeley u Kaliforniji. Također, dao mi je smjernice za rješavanje nekih problema.

Zahvalnost dugujem kolegi i prijatelju dr. Ivanu Keglevicu, koji je nevjerojatnom brzinom pronalazio sve članke do kojih sama nisam uspjevala doći.

Specijalnu zahvalnost dugujem mojoj obitelji, suprugu Saši i sinu Deanu, koji su me poticali da ovu radnju što prije završim, strpljivošću su reagirali na moju nervozu i . . . tko zna što još, čega uopće nisam svjesna.

Sadržaj

1. Uvod

I don't mind lying, but I hate inaccuracy.
Samuel BUTLER (1835–1902)

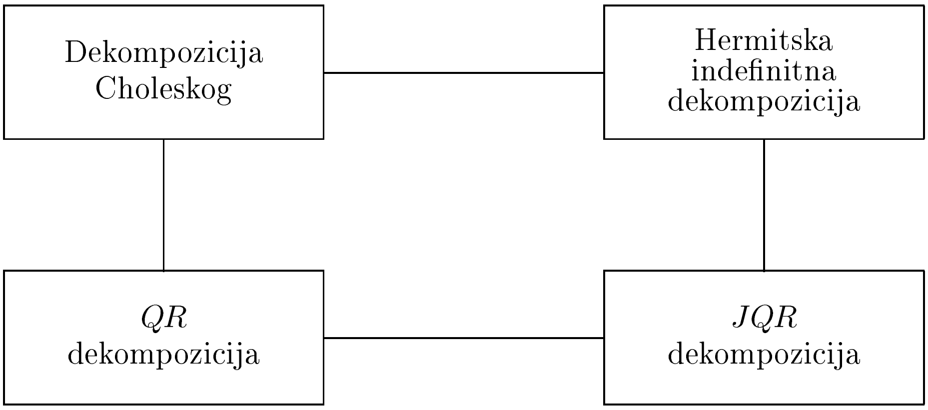
Razvojem računala razvile su se mnoge discipline numeričke matematike. Numerička linearna algebra doživjela je procvat, ali je donijela i svoje prokletstvo – točnost. Matematičari naviknuti na egzaktnost rezultata, jer, $1+\varepsilon$ je različito od 1, za svaki $\varepsilon \neq 0$, morali su, zbog konačnosti aritmetike računala, mijenjati pogled na matematički svijet. Različiti algoritmi koji u egzaktnoj aritmetici daju isti rezultat, ne moraju biti jednako točni u konačnoj aritmetici, pa se krajnji rezultati mogu (i to ne malo) razlikovati.

Danas, 35 i više godina nakon početaka numeričke linearne algebre, za većinu tradicionalnih algoritama provedena je obratna analiza grešaka zaokruživanja i dokazana točnost u apsolutnom smislu. Ipak, u nekim slučajevima, praksa pokazuje da možemo dobiti i točnije rezultate, no što to garantira ovakva analiza.

Za praksu je posebno važan problem računanja svojstvenih vrijednosti simetričnih matrica. Katkad su te matrice prirodno zadane u faktoriziranom obliku. Postavlja se pitanje, treba li pomnožiti faktore, pa zatim računati, ili ne. Već pozitivno definitni slučaj upućuje na pravo rješenje problema. Za pozitivno definitne matrice, svojstveni problem za zadanu matricu A, $A = G^*G$, ekvivalentan je dekompoziciji singularnih vrijednosti faktora G^*, odnosno G. Ako pomnožimo matrice G^*G, raspon elemenata u matrici A bit će znatno veći nego u G, što može dovesti do značajnijeg gubitka točnosti. Osim toga, ako je matrica G bila tipa (m, n), $m \gg n$, da bismo ubrzali računanje, prirodno je "skratiti" faktor G na kvadratnu matricu. Drmač je u [17] pokazao da postupak "skraćivanja" faktora optimalno pivotiranom QR dekompozicijom, ne samo što skraćuje računanje, nego i poboljšava točnost dobivenih singularnih vrijednosti.

U indefinitnom slučaju, situacija je nešto kompliciranija. Neka je zadana matrica A, $A = G^*JG$, gdje je J dijagonalna matrica s ± 1 na dijagonali. Slapničar je u [38] pokazao da je problem svojstvenih vrijednosti matrice A, ekvivalentan problemu svojstvenih vrijednosti para matrica (J, GG^*), odnosno hiperbolnoj dekompoziciji singularnih vrijednosti matrice G^*, obzirom na zadani J. Ako se za dijagonalizaciju para (J, GG^*) koristi Jacobijeva metoda, ona je provediva samo ako je matrica G^* punog stupčanog ranga, odnosno, matrica G punog ranga po recima. Polazeći od matrice A, takav faktor G možemo naći hermitskom indefinitnom dekompozicijom. S druge strane, ako je matrica A zadana faktorom G i matricom J, prirodno je takav faktor "skratiti" na kvadratni, "poštujući" zadani J za hiperbolnu dekompoziciju singularnih vrijednosti faktora, odnosno, skraćenog faktora. Intuitivno je jasno, da bi takva dekompozicija trebala biti indefinitni analogon QR dekompozicije.

U ovoj radnji precizno se definira indefinitna QR dekompozicija (ili, kraće, JQR dekompozicija) zadanog faktora G obzirom na zadani J i konstruktivno pokazuje njena egzistencija u slučaju da je matrica $A = G^*JG$ (Gramova matrica J–skalarnih produkata stupaca matrice G) regularna. Time se upotpunjuje "karika koja (djelomično) nedostaje" u slijedećem dijagramu.



Uobičajena QR dekompozicija ima vrlo široku primjenu kod, na primjer, rješavanja linearnih sistema, problema najmanjih kvadrata i dekompozicije singularnih vrijednosti. QR dekompozicija matrice $G \in \mathbf{C}^{m\times n}$, $m \geq n$, ima oblik

$$G = QR = [Q_1\ Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \quad ,$$

gdje je $Q \in \mathbf{C}^{m\times m}$ unitarna, a R_1 gornjetrokutasta matrica. Ako je matrica G punog ranga i ako su svi dijagonalni elementi matrice R_1 pozitivni, takva dekompozicija je jedinstvena. Jednako tako je poznato da svaka hermitska pozitivno definitna matrica A ima dekompoziciju Choleskog, tj. dekompoziciju oblika

$$A = R^*R \quad ,$$

gdje je R gornjetrokutasta matrica. Ako su dijagonalni elementi matrice R još i pozitivni, takva dekompozicija je jedinstvena. Veza između ove dvije dekompozicije je vrlo jednostavna. Neka je G bilo koja pravokutna matrica punog stupčanog ranga. Tada QR dekompozicija matrice G daje matricu R_1, jednaku matrici R iz dekompozicije Choleskog matrice $A = G^*G$.

Generalizacija dekompozicije Choleskog na hermitske indefinitne regularne matrice započela je sedamdesetih godina. Grupa autora: Bunch, Parlett i Kaufman ([7], [11], [8], [10] i [9]) napravila je algoritam i analizu grešaka za dekompoziciju hermitske indefinitne matrice u svrhu rješavanja hermitskih indefinitnih linearnih sistema. Pokazalo se da je hermitska indefinitna dekompozicija mnogo teži problem nego dekompozicija Choleskog, jer već vrlo jednostavan primjer 2×2 matrice

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

pokazuje da takva dekompozicija ne mora postojati, pa treba koristiti dijagonalne blokove dimenzije barem 2. Pokazuje se da je ta dimenzija blokova dovoljna, ako koristimo istovremene zamjene redaka i stupaca (pivotiranje). Svaka hermitska regularna matrica A ima dekompoziciju oblika

$$A = P^* R^* D R P \quad ,$$

gdje je R gornjetrokutasta matrica, D blok dijagonalna s 1×1 ili 2×2 dijagonalnim blokovima, a P je matrica permutacija. Različite pivotne strategije daju različite matrice R, D i P, pa ovakva dekompozicija nije jedinstvena. Izbor pivotiranja bitno utječe na stabilnost dekompozicije. Problem optimalne pivotne strategije nije zadovoljavajuće riješen. Ako koristimo, među poznatim strategijema najstabilnije, potpuno pivotiranje [11], ono ne mora biti dobar izbor za sve matrice. U Rutishauserovoj knjizi [35] (objavljenoj posthumno), kao primjer za problem pivotiranja u Gaussovim eliminacijama, navodi se matrica

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 10^{-10} & 0 \\ 1 & 0 & 10^{-10} \end{bmatrix} \quad .$$

Uočimo da je matrica A nesingularna. Isti primjer pokazuje probleme i u hermitskoj indefinitnoj dekompoziciji. Potpuno pivotiranje izabrat će 2 kao pivotni element. Nakon prvog koraka dekompozicije (u IEEE `single precision` aritmetici pomičnog zareza, s točnošću približno 10^{-7}), dobivena matrica bit će egzaktno singularna, jer će se elementima 10^{-10} pribrojiti $\sqrt{2}$, pa 10^{-10} ne utječe na konačni rezultat. Time smo uništili malu svojstvenu vrijednost matrice A. Ako, kao pivotni blok, izaberemo vodeću 2×2 podmatricu, dekompozicija će stabilno izračunati sve elemente u faktorima, a odgovarajući hiperbolni Jacobijev algoritam će relativno točno izračunati sve tri svojstvene vrijednosti matrice A.

U novije vrijeme, hermitska indefinitna dekompozicija postala je zanimljiva zbog točnog računanja svojstvenih vrijednosti. Veselić i Slapničar su u [38], [52] i

[40] pokazali da hermitska indefinitna dekompozicija, zajedno s hiperbolnim Jacobijevim algoritmom, uz određene pretpostavke, relativno točno računa svojstvene vrijednosti matrice. Slapničar je u [39] napravio komponentnu perturbacijsku analizu te dekompozicije. Razvoj paralelnih računala uvjetovao je potrebu za paralelnim algoritmima, u kojima je pivotiranje vrlo skupa operacija. Higham [24] daje analizu parcijalnog pivotiranja za hermitsku indefinitnu dekompoziciju.

Kao i u hiperbolnom Jacobijevom algoritmu, za realizaciju indefinitne QR dekompozicije, potrebne su unitarne matrice obzirom na indefinitni skalarni produkt zadan matricom J. Primjena takvih transformacija seže dvadesetak godina unatrag.

Veselić, kao jedan od pionira primjene neunitarnih transformacija na problem svojstvenih vrijednosti, već u svojim prvim radovima, na primjer u [48], spominje J–simetrične matrice i njihovu dijagonalizaciju, a njegov rad [49] o blok transformacijama, jedan je od temelja ove radnje.

Kako je primjena hiperbolnih matrica potekla iz više nezavisnih izvora, različitost je tolika da različiti autori za iste objekte koriste različita imena. Početkom osamdesetih godina, Bunse–Gerstner [12] koristi hiperbolne rotacije i hiperbolne Householderove matrice za konstrukciju indefinitne QR dekompozicije s gornjetrokutastom matricom R i pripadajućeg indefinitnog QR algoritma za računanje svojstvenih vrijednosti. Međutim, ako polazna matrica ima J–norme stupaca jednake 0, ovaj oblik dekompozicije ne postoji.

Približno pet godina kasnije, hiperbolne rotacije javljaju se u mnogim radovima, kao sredstvo za ažuriranje (downdating) dekompozicije Choleskog, ali i dalje ostaje isti problem sa stupcima J–norme jednake 0. Rader i Steinhardt [33] uvode J–Householderove matrice za problem najmanjih kvadrata.

S druge strane, krajem osamdesetih i početkom devedesetih godina, pronađena je primjena hiperbolnih matrica u obradi signala. Alexander, Pan i Plemmons [1] primjenjuju hiperbolne matrice za rješavanje problema brzih rekurzivnih filtera. Onn, Steinhardt i Bojanczyk [29] definiraju hiperbolnu dekompoziciju singularnih vrijednosti, a Bojanczyk i Steinhardt [4] uvode hiperbolne transformacije za izračunavanje dekompozicije Choleskog pozitivno definitne matrice X oblika

$$X = A^*A - B^*B \quad ,$$

zadane faktorima A i B.

Cybenko i Berry [15] koriste hiperbolne Householderove matrice za faktorizaciju Choleskog matrica s pomaknutom strukturom (tzv. displacement structure). Sličnom problematikom bave se Kailath i Sayed [28], te Chandrasekaran i Sayed [14].

Konačno, van der Veen [47] koristi hiperbolne rotacije kao sredstvo za izračunavanje aproksimacije smanjenog ranga $\widetilde{H}$ (u smislu tzv. numeričkog ranga) za zadanu matricu H. Ova konstrukcija, također, nije provediva, ako matrica H ima stupce J–norme nula.

Jednostavan je zaključak ovog povijesnog pregleda. Do sada je bio poznat samo nepotpuni indefinitni analogon QR dekompozicije, kad je matrica Q J–unitarna, a matrica R gornjetrokutasta. U slučaju da je matrica koju dekomponiramo takva da su joj (sve) J–norme stupaca jednake 0, dekompoziciju nije moguće napraviti samo trigonometrijskim i hiperbolnim rotacijama.

U ovoj radnji konstruira se potpuni analogon definitne QR dekompozicije, tj. dekompozicija oblika

$$G = P_1^* Q R P_2 = P_1^* Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} P_2 \quad , \quad Q^* J_1 Q = J_1 \quad , \quad J_1 = P_1^* J P_1,$$

gdje je R_1 blok gornjetrokutasta s dijagonalnim blokovima veličine 1×1 ili 2×2, J dijagonalna s ± 1 na dijagonali, Q J_1–unitarna, a P_1 i P_2 su matrice permutacija. Slično kao kod hermitske indefinitne dekompozicije, nužno je, u slučaju ove dekompozicije, koristiti obostrano pivotiranje — i po stupcima i po recima.

Radnja je podijeljena na sedam poglavlja.

Drugo poglavlje sadrži sažeti općeniti uvod o indefinitnim skalarnim produktima i, posebno, o onim svojstvima, poput degeneracije, koja ih bitno razlikuju od uobičajenog skalarnog produkta. Prema [22], uvode se pojmovi H–hermitske i H–unitarne matrice obzirom na skalarni produkt zadan matricom H.

Na početku trećeg poglavlja izložena su opća svojstva H–unitarnih matrica, a zatim se konstrukcija općih H–unitarnih matrica svodi na konstrukciju J–unitarnih matrica, gdje je J matrica inercije matrice H. Ostatak ovog poglavlja posvećen je efektivnoj konstrukciji tzv. elementarnih J–unitarnih matrica.

Poznato je da se QR dekompozicija može izvoditi Givensovim rotacijama ili Householderovim reflektorima. Slično vrijedi i za indefinitnu QR dekompoziciju. Iz [49] preuzeta je konstrukcija blok J–rotacija i njihova osnovna svojstva. Na kraju, dana je konstrukcija blok J–Householderovih matrica, koje su blok generalizacija do sada poznatih J–Householderovih matrica ([15], [12]) i blok Householderovih matrica ([5], [6], [16] i [37]).

Težište radnje je na četvrtom poglavlju, gdje je konstruktivno dokazan teorem egzistencije indefinitne QR dekompozicije u navedenom obliku, za proizvoljnu matricu G tipa (m, n), $m \geq n$, uz pretpostavku regularnosti odgovarajuće J–Gramove matrice $A = G^* J G$. Dokaz je algoritamski, primjenom transformacija Givensovog tipa, uz elementarno pivotiranje koje osigurava samo regularnost pivotnog bloka.

Osim toga, ako je matrica $A = G^* J G$ singularna, algoritam dekompozicije će tu singularnost, barem teoretski, detektirati. U tom slučaju, algoritam će u zadnjim stupcima radne matrice ostaviti (restrikcije) stupaca koji predstavljaju (općenito netrivijalnu) dekompoziciju 0–bloka. Naravno, dobra detekcija singularnosti bitno ovisi o izabranom pivotiranju. Sve razumne pivotne strategije za hermitsku indefinitnu dekompoziciju mogu se izravno koristiti i u algoritmu JQR dekompozicije (što odgovara analogiji između pivotiranja po stupcima u QR dekompoziciji i dijagonalnog pivotiranja u dekompoziciji Choleskog).

Nalaženje optimalne pivotne strategije za JQR dekompoziciju je, očito, barem jednako težak problem, kao i nalaženje optimalne pivotne strategije za hermitsku indefinitnu ili QR dekompoziciju. Za numeričko određivanje ranga zadane matrice, problem optimalne pivotne strategije nije zadovoljavajuće algoritamski riješen, ni za QR, niti za hermitsku indefinitnu dekompoziciju. Stoga je to otvoreni problem i za JQR dekompoziciju.

Indefinitnu QR dekompoziciju možemo, osim J–rotacijama, konstruirati i primjenom matrica Householderovog tipa, čemu je posvećen kraj četvrtog poglavlja. Jedini problem pri toj konstrukciji je moguća degeneracija blok J–Householderove matrice

$$H(W) = I - 2W(W^*JW)^+W^*J$$

tj. gubitak ranga u produktu W^*JW, obzirom na rang od W. U tom slučaju, nije moguće konstruirati blok J–Householderovu matricu za poništavanje bloka.

Direktna konstrukcija J–Householderove matrice koja bi poništavala dva puna stupca matrice G je vrlo komplicirana, jer dozvoljava previše mogućnosti za izgled bloka kojeg dobivamo nakon transformacije. Ako, kao predradnju, dva stupca matrice G dovedemo na isti skraćeni oblik kao kod Givensovih transformacija, a zatim iskoristimo rezultate o izgledu bloka nakon transformacije J–rotacijama, moguće je konstruirati odgovarajuće blok J–Householderove matrice, kod kojih, važno je naglasiti, ne dolazi do degeneracije.

U petom poglavlju analiziraju se greške zaokruživanja i dokazuju teoremi perturbacije za indefinitnu QR dekompoziciju. Analiza grešaka zaokruživanja sastoji se od nekoliko koraka. U prvom, analizira se pogreška nastala primjenom jedne izračunate trigonometrijske rotacije, što je klasični Wilkinsonov rezultat [53]. Slijedeći Gentlemanovu ideju [21], zatim se ocjenjuje pogreška za nizove rotacija, koje možemo nezavisno primjenjivati. Na kraju, ocjenjuje se pogreška jedne hiperbolne, odnosno, blok J–rotacije.

Pokazuje se da pogrešku u izračunatoj JQR dekompoziciji možemo interpretirati kao egzaktnu dekompoziciju malo perturbirane matrice. Ako je matrica takva da dekompoziciju provodimo samo rotacijama i hiperbolnim rotacijama, tj. ako su hiperbolni kutevi koje koristimo ograničeni odozgo, dokazuje se da je perturbacija u normama stupaca relativno mala obzirom na euklidske norme stupaca.

U protivnom, ako matrica G, ili neka podmatrica koja se javlja u realizaciji dekompozicije, ima neke J–norme stupaca jednake 0, (odnosno, ekvivalentno, ako matrica A, ili njene glavne podmatrice koje se dobivaju pri provođenju dekompozicije, imaju dijagonalne elemente jednake 0), relativna točnost izračunate J–norme stupaca nema smisla (relativna perturbacija nule je samo nula). Ovaj fenomen specifičan je za indefinitnu QR dekompoziciju i ne postoji analogon u pozitivno definitnom slučaju.

Na kraju, dokazujemo da se komponentni perturbacijski teoremi za QR dekompoziciju (Sun [44] i Zha [54]) mogu generalizirati i na indefinitnu QR dekompoziciju, uz dodatnu pretpostavku da je dekompoziciju moguće provesti samo trigonometrijskim i hiperbolnim rotacijama, tj. da je matrica R_1 trokutasta.

U šestom poglavlju dane su neke primjene indefinitne QR dekompozicije. Ovu dekompoziciju možemo koristiti za rješavanje linearnih sustava (analogno kao i QR dekompoziciju), jer se matrice Q i R_1 lako invertiraju, čak i kad R_1 ima 2×2 blokove na dijagonali. Međutim, zbog sporosti algoritma, ova primjena ima smisla samo za linearne sustave posebne strukture i nećemo je posebno razmatrati.

Pravu primjenu indefinitne QR dekompozicije nalazimo u problemima čija prirodna formulacija sadrži indefinitni skalarni produkt, a posebno kad je indefinitni skalarni produkt invarijanta problema. U terminu matrice J, to znači da su dozvoljene samo istovremene zamjene redaka i stupaca u J. Ova klasa problema obuhvaća računanje hiperbolne dekompozicije singularnih vrijednosti obzirom na zadani J. Numerički gledano, uz zahtjev relativne točnosti, problem svojstvenih vrijednosti simetrične matrice pripada istoj klasi, jer inercija polazne matrice mora (numerički) ostati nepromijenjena.

Na početku, izvodi se generalizacija implicitnog algoritma Choleskog (tzv. "flip–flap" algoritam, Fernando–Parlett [20]), za nalaženje hiperbolnih singularnih, odnosno, svojstvenih vrijednosti. Ovaj algoritam, uz neke uvjete, ima samo linearnu konvergenciju prema hiperbolnim singularnim vrijednostima. Kao otvoreni problem ostaje, kako treba dobro izabrati pomake (shiftove) u J–"flip–flap" algoritmu za ubrzavanje konvergencije. Za rješenje tog problema, potrebne su precizne ograde na veličinu hiperbolnih singularnih vrijednosti za specijalne tipove matrica.

U drugom odjeljku šestog poglavlja, izvodi se algoritam (indefinitne) bidijagonalizacije faktora G. Preciznije rečeno, za matricu G tipa (m, n), $m \geq n$, takvu da je matrica $A = G^* J G$ regularna, pokazuje se postojanje dekompozicije u obliku

$$G = P_1^* V B U^* P_2 = P_1^* V \begin{bmatrix} B_1 \\ 0 \end{bmatrix} U^* P_2 \quad ,$$

pri čemu je B_1 gornja bidijagonalna, U unitarna, za matricu V vrijedi

$$V^* J_1 V = J_1 \quad , \quad J_1 = P_1^* J P_1 \quad ,$$

a P_1 i P_2 su matrice permutacija. Takva dekompozicija uklapa se u današnje trendove brzog i točnog računanja (hiperbolnih) singularnih, odnosno svojstvenih vrijednosti (v. [32]).

U trećem odjeljku šestog poglavlja, razmatra se primjena indefinitne QR dekompozicije na iterativno profinjavanje svojstvenih vrijednosti. Poznata je činjenica da Jacobijev algoritam sporo, ali relativno točno računa svojstvene vrijednosti, barem za neke tipove matrica. Svojstvene vrijednosti možemo naći i nekim bržim, ali netočnijim algoritmom, a nakon toga vršiti iterativni popravak, za što možemo koristiti indefinitnu QR dekompoziciju, a nakon toga implicitni hiperbolni Jacobijev algoritam. Dokazuje se da indefinitna QR dekompozicija za η–skalirano dijagonalno dominantne matrice, uz dodatni uvjet na omjer euklidske i J–norme stupaca, daje "dobro" skalirani faktor za primjenu hiperbolne Jacobijeve metode.

Posljednje poglavlje u radnji posvećeno je numeričkom testiranju. U ovom poglavlju eksperimentalno se odgovara na pitanje, zbog čega je bolje koristiti indefinitnu QR dekompoziciju na faktoru, nego pomnožiti faktore, pa zatim primijeniti hermitsku indefinitnu dekompoziciju. Primjeri su vrlo pažljivo birani, tako da istaknu dobre, ali i negativne strane indefinitne QR dekompozicije. Kao uvodnu motivaciju, navedimo samo jedan od primjera, gdje se pokazuju značajne prednosti predložene dekompozicije.

Neka je zadan faktor

$$G = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 \cdot 10^{-11} & -1 \cdot 10^{-11} \\ 1 \cdot 10^{-11} & 1 \cdot 10^{-11} \end{bmatrix}$$

i

$$J = \operatorname{diag}(1, 1, -1, -1) \quad .$$

Matrica $A = G^* J G$ je regularna, ali ima jednu malu svojstvenu vrijednost. Ako koristimo `extended` točnost (64 binarne znamenke za mantisu i 15 za eksponent), tj. najviše što je arhitekturom procesora podržano na PC računalu, onda je izračunata matrica

$$f\ell(A) = f\ell(G^* J G) = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

egzaktno singularna, pa smo izgubili malu svojstvenu vrijednost.

2. Indefinitni skalarni produkt

Indefinitni skalarni produkt generalizacija je uobičajene definicije skalarnog produkta. Dobiva se ispuštanjem svojstva $(x, x) \geq 0$ u definiciji standardnog skalarnog produkta. U ovom poglavlju navodimo osnovna svojstva takvog skalarnog produkta koja su potrebna za konstrukciju pripadnih unitarnih operatora, odnosno matrica.

Na početku, pokazujemo da se svaki indefinitni skalarni produkt može reprezentirati nekom regularnom hermitskom matricom H i obratno. U tom smislu, standardni skalarni produkt odgovara jediničnoj matrici. Zatim, generaliziramo pojam hermitske i unitarne matrice na H–unitarne prostore i analiziramo odnos potprostora i njegovog ortogonalnog komplementa u takvom prostoru. Bitna razlika u odnosu na standardni skalarni produkt leži u činjenici, da nul–vektor ne mora biti jedini vektor okomit na samog sebe.

2.1. Osnovna svojstva

Prvo uvedimo standardne oznake. Malim slovima označavat ćemo vektore, a grčkim skalare. Matrice označavamo velikim slovima. Prema [22], uobičajeni skalarni produkt vektora označavat ćemo s $(\, , \,)$, a indefinitni skalarni produkt s $[\, , \,]$.

Uvedimo preciznu definiciju indefinitnog skalarnog produkta.

Definicija 2.1.1.

Funkcija $[\, , \,] : \mathbf{C}^m \times \mathbf{C}^m \to \mathbf{C}$ *je indefinitni skalarni produkt, ako vrijedi:*

(i) $[\alpha x_1 + \beta x_2, y] = \alpha [x_1, y] + \beta [x_2, y]$, $\forall \alpha, \beta \in \mathbf{C}$ *i* $\forall x_1, x_2, y \in \mathbf{C}^m$, *(linearnost u prvom argumentu),*

(ii) $[x, y] = \overline{[y, x]}$, $\forall x, y \in \mathbf{C}^m$, *(hermitičnost),*

(iii) $[x, y] = 0$, $\forall y \in \mathbf{C}^m \implies x = 0$, *(nedegeneriranost).*

■

Drugim riječima, indefinitni skalarni produkt je hermitski seskvilinearan funkcional koji zadovoljava još i dodatno svojstvo (iii).

Napomena 2.1.1.

Zamjenom $\mathbf{C}$ *s* $\mathbf{R}$ *u definiciji* 2.1.1.*, dobivamo definiciju indefinitnog skalarnog produkta na realnom prostoru* $\mathbf{R}^m$*, pri čemu se svojstvo* (ii) *pojednostavnjuje na*

(ii') $[x,y] = [y,x]$, $\forall x, y \in \mathbf{R}^m$, *(simetričnost).*

■

Postoji jednostavna veza indefinitnih skalarnih produkata i regularnih hermitskih matrica.

Teorem 2.1.1.

Svaka regularna hermitska matrica H *definira indefinitni skalarni produkt relacijom*

$$[x,y] = (Hx, y) \quad . \tag{2.1.1}$$

Dokaz:

Provjerimo sva svojstva iz definicije indefinitnog skalarnog produkta.

(i) Linearnost u prvom argumentu vrijedi zbog linearnosti množenja matrice vektorom i linearnosti u prvom argumentu običnog skalarnog produkta, tj.

$$\begin{aligned} [\alpha x_1 + \beta x_2, y] &= (H(\alpha x_1 + \beta x_2), y) \\ &= \alpha(Hx_1, y) + \beta(Hx_2, y) \\ &= \alpha[x_1, y] + \beta[x_2, y] \quad . \end{aligned}$$

(ii) Hermitičnost slijedi iz hermitičnosti običnog skalarnog produkta i hermitičnosti matrice H,

$$[x,y] = (Hx, y) = \overline{(y, Hx)} = \overline{(H^*y, x)} = \overline{(Hy, x)} = \overline{[y,x]} \quad .$$

(iii) Nedegeneriranost je posljedica regularnosti matrice H. Označimo $z = Hx$. Tvrdimo da $[x,y] = (z,y) = 0$, $\forall y \in \mathbf{C}^m$, povlači $x = 0$. Izaberimo, redom za vektor y, m vektora standardne baze prostora $\mathbf{C}^m$. Označimo te vektore s $e_1, \ldots, e_m$. Relacija $(z, e_i) = 0$ za svako i, povlači da su sve komponente vektora z jednake 0, dakle, da je z nul–vektor. Zaključak da je $x = 0$ slijedi iz činjenice da, ako je matrica homogenog linearnog sistema $Hx = 0$ regularna, onda on ima samo trivijalno rješenje $x = 0$.

■

Zanimljivo je da vrijedi i obrat prethodnog teorema.

Teorem 2.1.2.

Za svaki indefinitni skalarni produkt postoji regularna hermitska matrica za koju vrijedi (2.1.1).

Dokaz:

Izaberimo fiksni $y \in \mathbf{C}^m$. Tada je preslikavanje

$$x \mapsto [x, y]$$

linearni funkcional na $\mathbf{C}^m$. Prema Teoremu reprezentacije linearnih funkcionala, postoji jedan i samo jedan vektor $y_0 \in \mathbf{C}^m$, takav da vrijedi

$$[x, y] = (x, y_0) \quad ,$$

za sve $x \in \mathbf{C}^m$. Zbog toga je, definicijom $y_0 = Hy$, dobro definirano preslikavanje $H : \mathbf{C}^m \to \mathbf{C}^m$. To preslikavanje je linearno, zbog linearnosti indefinitnog skalarnog produkta. Hermitičnost i invertibilnost pripadne matrice H slijede izravno iz hermitičnosti i nedegeneriranosti skalarnog produkta. ■

Prethodna dva teorema opravdavaju slijedeću definiciju.

Definicija 2.1.2.

Neka je $[\ ,\ ] : \mathbf{C}^m \times \mathbf{C}^m \to \mathbf{C}$ *indefinitni skalarni produkt i neka je* H *regularna hermitska matrica za koju vrijedi* (2.1.1). *Takav skalarni produkt zvat ćemo* H*–skalarni produkt, a prostor* $\mathbf{C}^m$ *snabdjeven tim skalarnim produktom* H*–unitarni prostor.* ■

Uočimo da standardni pojam unitarnog prostora odgovara I–unitarnom prostoru, gdje je I jedinična matrica reda m.

Standardni skalarni produkt, zbog svojstva $(x, x) \geq 0$, generira normu vektora x relacijom

$$\|x\| = \sqrt{(x, x)} \quad , \quad \forall x \in \mathbf{C}^m \quad .$$

Ako je matrica skalarnog produkta H indefinitna, onda ona ima bar jednu negativnu svojstvenu vrijednost λ. Za pripadni svojstveni vektor x vrijedi

$$[x, x] = (Hx, x) = (\lambda x, x) = \lambda(x, x) < 0 \quad ,$$

što pokazuje da nije moguće definirati normu generiranu H–skalarnim produktom. Ipak, pokazat će se korisnim pojam "H–norme" definiran relacijom

$$\|x\|_H = \sqrt{|\,[x, x]\,|} \quad , \quad \forall x \in \mathbf{C}^m \quad . \tag{2.1.2}$$

Obzirom na H–skalarni produkt, možemo definirati adjungiranje matrica.

Definicija 2.1.3.

Matrica $A^{[*]}$ je H–adjungirana matrici A, ako vrijedi

$$[Ax, y] = [x, A^{[*]}y] \quad , \quad \forall x, y \in \mathbf{C}^m \quad . \tag{2.1.3}$$

■

U terminu uobičajenog skalarnog produkta, lijeva i desna strana relacije (2.1.3) glase

$$\begin{aligned} [Ax, y] &= (HAx, y) = (x, A^*Hy) \quad , \\ [x, A^{[*]}y] &= (Hx, A^{[*]}y) = (x, HA^{[*]}y) \quad . \end{aligned}$$

Uspoređivanjem desnih strana dobivamo matrično izraženu relaciju za H–adjungiranost

$$A^{[*]} = H^{-1}A^*H \quad . \tag{2.1.4}$$

Sada možemo definirati pojam H–hermitske i H–unitarne matrice.

Definicija 2.1.4.

(i) *Matrica A je H–hermitska, ako vrijedi $A^{[*]} = A$.*

(ii) *Matrica U je H–unitarna, ako vrijedi*

$$[Ux, Uy] = [x, y] \quad , \quad \forall x, y \in \mathbf{C}^m \quad ,$$

tj. ako U čuva H–skalarni produkt vektora. ■

U terminima matrice H, ova svojstva možemo izraziti na slijedeći način.

Propozicija 2.1.1.

U H–unitarnom prostoru $\mathbf{C}^m$ vrijedi

(i) *Matrica A je H–hermitska, ako i samo ako vrijedi*

$$A^* = HAH^{-1} \quad . \tag{2.1.5}$$

(ii) *Matrica U je H–unitarna, ako i samo ako vrijedi $U^{[*]}U = I$, ili, ekvivalentno,*

$$H = U^*HU \quad . \tag{2.1.6}$$

Dokaz:

(i) Relacija (2.1.5) izlazi izravno iz definicije i (2.1.4).

(ii) Prva tvrdnja posljedica je definicije 2.1.3., a (2.1.6) izlazi primjenom (2.1.4).

■

2.2. H–unitarnost

Prirodno je pokušati proširiti osnovna svojstva koja vrijede za unitarne prostore i na H–unitarne prostore.

U unitarnom prostoru $\mathbf{C}^m$, svaki potprostor $\mathcal{M}$ dekomponira prostor na direktnu sumu tog potprostora i njegovog ortogonalnog komplementa,

$$\mathcal{M} \oplus \mathcal{M}^{\perp} = \mathbf{C}^m \quad .$$

Posebno, vrijedi $\mathcal{M} \cap \mathcal{M}^{\perp} = \{0\}$ i

$$\dim \mathcal{M} + \dim \mathcal{M}^{\perp} = m \quad .$$

Pitanje je, koje od ovih tvrdnji vrijede i za H–unitarne prostore.

Neka je $\mathcal{S} \subseteq \mathbf{C}^m$ bilo koji neprazan skup vektora. Definirajmo skup vektora koji su H–ortogonalni na $\mathcal{S}$,

$$\mathcal{S}^{[\perp]} = \{x \in \mathbf{C}^m \mid [x, y] = 0 \ , \forall y \in \mathcal{S}\} \quad .$$

Očito je $\mathcal{S}^{[\perp]}$ potprostor u $\mathbf{C}^m$. Posebno interesantan je slučaj kad je i $\mathcal{S}$ potprostor. Tada, umjesto $\mathcal{S}$, koristimo oznaku $\mathcal{M}$.

Neka je $\mathcal{M}$ potprostor u H–unitarnom prostoru $\mathbf{C}^m$. Pitanje je da li $\mathcal{M}$ i $\mathcal{M}^{[\perp]}$ dekomponiraju prostor na direktnu sumu potprostora, tj. da li je jedan, direktni komplement drugog.

Slijedeći primjer pokazuje da to ne mora vrijediti.

Primjer 2.2.1.

Neka je $\{e_1, \ldots, e_m\}$ standardna baza u $\mathbf{C}^m$ i neka je matrica H skalarnog produkta zadana s

$$H = [e_m, \ldots, e_1] = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ & & \iddots & & \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \quad .$$

Nađimo skup svih vektora $x \in \mathbf{C}^m$, koji su H–ortogonalni na potprostor

$$\mathcal{M} = \operatorname{span}\{e_1\} \quad .$$

Iz uvjeta ortogonalnosti dobivamo

$$0 = [x, e_1] = (Hx, e_1) = (x, H^* e_1) = (x, He_1) = (x, e_m) \quad ,$$

odakle slijedi

$$\mathcal{M}^{[\perp]} = \text{span}\{e_1, e_2, \ldots, e_{m-1}\} \quad ,$$

a odavde je očito da je

$$\mathcal{M} \cap \mathcal{M}^{[\perp]} = \text{span}\{e_1\} = \mathcal{M} \quad .$$

Zaključujemo da je $\mathcal{M} \cap \mathcal{M}^{[\perp]} \neq \{0\}$, a to znači da se prostor ne može dekomponirati u direktnu sumu potprostora $\mathcal{M}$ i $\mathcal{M}^{[\perp]}$.

Osim toga, vektor e_m ne pripada ni potprostoru $\mathcal{M}$, niti potprostoru $\mathcal{M}^{[\perp]}$, što pokazuje da suma potprostora $\mathcal{M}$ i $\mathcal{M}^{[\perp]}$ nije čitav prostor.

Uočimo da za potprostor $\mathcal{M}$ vrijedi da je H–okomit na samog sebe. Kao što ćemo pokazati, upravo ta degeneriranost obzirom na H–skalarni produkt onemogućava rastav cijelog prostora. ■

U H–unitarnim prostorima uvijek vrijede slijedeće propozicije.

Propozicija 2.2.1.

Za svaki potprostor $\mathcal{M} \subseteq \mathbf{C}^m$ vrijedi

$$\dim \mathcal{M} + \dim \mathcal{M}^{[\perp]} = m \quad . \tag{2.2.1}$$

Dokaz:

Pokažimo da vrijedi

$$\mathcal{M}^{[\perp]} = H^{-1}(\mathcal{M}^{\perp}) \quad .$$

Neka je $x \in \mathcal{M}^{\perp}$ i $y \in \mathcal{M}$. Tada je

$$[H^{-1}x, y] = (HH^{-1}x, y) = (x, y) = 0 \quad ,$$

pa je $H^{-1}(\mathcal{M}^{\perp}) \subseteq \mathcal{M}^{[\perp]}$.

Obratno, ako je $x \in \mathcal{M}^{[\perp]}$ i $z = Hx$, tada za bilo koji $y \in \mathcal{M}$ vrijedi

$$0 = [x, y] = [H^{-1}z, y] = (z, y) \quad .$$

Odatle slijedi da je $z \in \mathcal{M}^{\perp}$. Iz $x = H^{-1}z$ dobivamo da je $x \in H^{-1}(\mathcal{M}^{\perp})$. ■

Propozicija 2.2.2.

Za svaki neprazan skup $\mathcal{S} \subseteq \mathbf{C}^m$ vrijedi

$$\mathcal{S} \subseteq (\mathcal{S}^{[\perp]})^{[\perp]} \quad . \tag{2.2.2}$$

Posebno, ako je $\mathcal{M}$ potprostor u $\mathbf{C}^m$, onda vrijedi

$$\mathcal{M} = (\mathcal{M}^{[\perp]})^{[\perp]} \quad . \tag{2.2.3}$$

Dokaz:

Neka je $\mathcal{S} \subseteq \mathbf{C}^m$ i $x \in \mathcal{S}$. Po definiciji skupa $\mathcal{S}^{[\perp]}$, za svaki $y \in \mathcal{S}^{[\perp]}$ vrijedi $[y,x] = 0$, odakle slijedi $x \in (\mathcal{S}^{[\perp]})^{[\perp]}$, što dokazuje (2.2.2).

Za potprostor $\mathcal{M}$, također, vrijedi $\mathcal{M} \subseteq (\mathcal{M}^{[\perp]})^{[\perp]}$. Primjenom propozicije 2.2.1. na $\mathcal{M}$ i $\mathcal{M}^{[\perp]}$, izlazi

$$\dim \mathcal{M} + \dim \mathcal{M}^{[\perp]} = \dim \mathcal{M}^{[\perp]} + \dim (\mathcal{M}^{[\perp]})^{[\perp]} = m \quad ,$$

odakle slijedi $\dim \mathcal{M} = \dim (\mathcal{M}^{[\perp]})^{[\perp]}$, pa vrijedi (2.2.3). ■

Preostale tvrdnje o vezi potprostora $\mathcal{M}$ i $\mathcal{M}^{[\perp]}$ vrijede samo uz dodatne pretpostavke na $\mathcal{M}$.

Definicija 2.2.1.

Potprostor $\mathcal{M}$ je nedegeneriran obzirom na indefinitni skalarni produkt, ako vrijedi

$$x \in \mathcal{M} \;\; i \;\; (\forall y \in \mathcal{M}) \; [x,y] = 0 \implies x = 0 \quad .$$

U ostalim slučajevima, potprostor $\mathcal{M}$ je degeneriran. ■

Drugim riječima, potprostor $\mathcal{M}$ je nedegeneriran, ako i samo ako je nul–vektor jedini vektor u $\mathcal{M}$ okomit na samog sebe.

U slučaju nedegeneriranih potprostora, vrijede sva lijepa svojstva na koja smo navikli u unitarnim prostorima. Slijedeća propozicija daje odgovarajuću karakterizaciju nedegeneriranih potprostora.

Propozicija 2.2.3.

Potprostor $\mathcal{M} \subseteq \mathbf{C}^m$ je nedegeneriran, ako i samo ako je $\mathcal{M}^{[\perp]}$ direktan komplement od $\mathcal{M}$ u $\mathbf{C}^m$.

Dokaz:

Prema definiciji, $\mathcal{M}$ je nedegeneriran, ako i samo ako je $\mathcal{M} \cap \mathcal{M}^{[\perp]} = \{0\}$, što, zajedno s propozicijom 2.2.1., daje traženi rezultat. ■

Ako je $\mathcal{M}$ nedegeneriran, možemo konstruirati i H–ortogonalnu bazu $\{x_i\}$ potprostora $\mathcal{M}$, tj. vrijedi

$$[x_i, x_j] = \begin{cases} \pm 1 & , \text{za } i = j \\ 0 & , \text{za } i \neq j \end{cases} \quad .$$

Konstrukcija ortogonalne baze je analogon Gram–Schmidtove ortogonalizacije vektora, a zasniva se na činjenici da je svaki potprostor nedegeneriranog potprostora $\mathcal{M}$, također, nedegeneriran. Zbog toga, u svakom koraku ortogonalizacije, možemo naći vektor x_i u "ostatku" potprostora, za koji je $[x_i, x_i] \neq 0$.

3. H–unitarne matrice

Unitarne matrice koriste se u različitim algoritmima za dovođenje zadane matrice u željeni oblik na numerički stabilan način. Najčešće se željeni oblik postiže poništavanjem elemenata matrice.

Po analogiji, prirodno je koristiti H–unitarne matrice za istu svrhu, u problemima koji su prirodno formulirani u nekom H–skalarnom produktu. Za tu svrhu, potrebne H–unitarne matrice treba brzo algoritamski izračunati.

Problem konstrukcije općenitih H–unitarnih matrica nije sasvim jednostavan, ako dozvolimo bilo kakav hermitski i regularan H. U općem slučaju, rješenje uključuje dijagonalizaciju matrice H, što nije algoritamski prihvatljivo, pa taj problem nećemo detaljno razmatrati.

Za praktičnu primjenu, interesantni su samo vrlo jednostavni oblici matrice H. Osim toga, dovoljno je naći tzv. elementarne H–unitarne matrice, koje mogu poslužiti za multiplikativnu reprezentaciju bilo koje H–unitarne matrice.

Najčešće korištene elementarne unitarne matrice su ravninske rotacije i Householderove matrice. U slijedećih nekoliko odjeljaka, definirat ćemo generalizacije tih elementarnih matrica u nekim, za praksu interesantnim, slučajevima matrice H.

3.1. Osnovna svojstva H–unitarnih matrica

Pokažimo neka svojstva H–unitarnih matrica.

Propozicija 3.1.1.

Ako je $U \in \mathbf{C}^{m\times m}$ H–unitarna matrica, onda je U regularna matrica i vrijedi $|\det(U)| = 1$.

Dokaz:

Prema propoziciji 2.1.1., za U vrijedi $U^{[*]}U = I$, pa je U regularna. Nadalje, iz (2.1.6) i Binet–Cauchyjevog teorema izlazi

$$\det(H) = \det(U^* H U) = \det(U^*) \det(H) \det(U) = |\det(U)|^2 \det(H) \quad .$$

Zbog regularnosti H, dijeljenjem prethodne relacije s $\det(H)$, dobivamo

$$|\det(U)|^2 = 1 \quad ,$$

odakle slijedi tvrdnja. ■

Propozicija 3.1.2.

Skup svih H–unitarnih matrica je (multiplikativna) podgrupa grupe regularnih matrica istog reda kao i H.

Dokaz:

Prema prethodnoj propoziciji, H–unitarne matrice su regularne, pa je dovoljno pokazati da za H–unitarne matrice U_1 i U_2 vrijedi, da je i matrica $U_1^{-1}U_2$ takva. Množenjem relacije

$$U_1^* H U_1 = H$$

slijeva s U_1^{-*} i zdesna s U_1^{-1} dobivamo

$$U_1^{-*} H U_1^{-1} = H \quad ,$$

što pokazuje da je i U_1^{-1} H–unitarna. Nadalje, množenjem prethodne relacije slijeva s U_2^* i zdesna s U_2, iz H–unitarnosti matrice U_2 izlazi

$$U_2^* U_1^{-*} H U_1^{-1} U_2 = U_2^* H U_2 = H \quad .$$

■

Slijedeća propozicija opisuje kako se ponaša grupa H–unitarnih matrica obzirom na unitarne sličnosti (i kongruencije) matrice H skalarnog produkta.

Propozicija 3.1.3.

Neka je H regularna hermitska matrica reda m i neka je V bilo koja unitarna matrica istog reda. Nadalje, neka je

$$\widetilde{H} = V^* H V \quad .$$

Matrica U je H–unitarna matrica, ako i samo ako je matrica

$$\widetilde{U} = V^* U V$$

$\widetilde{H}$–unitarna.

Dokaz:

Iz unitarnosti matrice V i $U^*HU = H$, dobivamo redom

$$\widetilde{U}^* \widetilde{H} \widetilde{U} = (V^* U^* V)(V^* H V)(V^* U V) = V^* U^* H U V = V^* H V = \widetilde{H} \quad ,$$

odakle slijedi da je $\widetilde{U}$ $\widetilde{H}$–unitarna. Obrat se dokazuje na isti način. ■

Slijedeća propozicija opisuje blok konstrukciju velikih unitarnih matrica koje nastaju spajanjem malih unitarnih matrica i pripadnih skalarnih produkata.

Propozicija 3.1.4.

Neka su H_1 i H_2 hermitske regularne matrice. Ako je matrica U_1 H_1–unitarna, a U_2 H_2–unitarna, onda je matrica

$$U = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}$$

unitarna obzirom na H–skalarni produkt definiran matricom

$$H = \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix} \quad .$$

Dokaz:

Matrica H je hermitska i regularna, što slijedi iz hermitičnosti i regularnosti matrica H_1 i H_2. Prema propoziciji 2.1.1., matrica U je H–unitarna, ako vrijedi $U^*HU = H$. Množenjem odgovarajućih blokova dobivamo

$$\begin{aligned} U^*HU &= \begin{bmatrix} U_1^* & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix} \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \\ &= \begin{bmatrix} U_1^*H_1U_1 & 0 \\ 0 & U_2^*H_2U_2 \end{bmatrix} = \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix} = H \quad . \end{aligned}$$

∎

Elementarne H–unitarne matrice, posebno H–rotacije, konstruiramo modifikacijom jedinične matrice na osnovu slijedeće propozicije.

Propozicija 3.1.5.

Neka je U H–unitarna matrica reda ℓ. Neka je $\widetilde{U}$ jednaka jediničnoj matrici, osim podmatrice reda ℓ na presjeku redaka i stupaca s indeksima $k_1, \ldots, k_\ell$, koja je jednaka U, tj. u Matlab *notaciji*

$$\widetilde{U}([k_1, \ldots, k_\ell], [k_1, \ldots, k_\ell]) = U \quad .$$

Neka je $\widetilde{H}$ bilo koja hermitska regularna matrica, čiji reci i stupci s indeksima $k_1, \ldots, k_\ell$ imaju nul–elemente, osim na presjeku tih redaka i stupaca, gdje vrijedi

$$\widetilde{H}([k_1, \ldots, k_\ell], [k_1, \ldots, k_\ell]) = H \quad .$$

Tada je $\widetilde{U}$ $\widetilde{H}$–unitarna matrica.

Dokaz:

Iz pretpostavki slijedi da postoji matrica permutacije $\widetilde{P}$ takva da je

$$\widehat{U} = \widetilde{P}^*\widetilde{U}\widetilde{P} = \begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} \quad , \quad \widehat{H} = \widetilde{P}^*\widetilde{H}\widetilde{P} = \begin{bmatrix} H & 0 \\ 0 & H_1 \end{bmatrix} \quad .$$

Iz regularnosti $\widetilde{H}$ slijedi da je i H_1 regularna hermitska matrica. Nadalje, jedinična matrica I je, očito, H_1–unitarna, pa, primjenom propozicije 3.1.4., zaključujemo da je $\widehat{U}$ $\widehat{H}$–unitarna matrica. Na kraju, matrica permutacije $\widetilde{P}$ je unitarna, pa je to i $\widetilde{P}^*$. Primjenom propozicije 3.1.3. na $\widehat{U}$ i $\widehat{H}$, uz $V = \widetilde{P}^*$, slijedi da je $\widetilde{U}$ $\widetilde{H}$–unitarna. ∎

Svaka hermitska matrica može se dijagonalizirati unitarnim sličnostima. Zbog propozicije 3.1.3., dovoljno je konstruirati H–unitarne matrice za dijagonalne H. Za takve matrice H vrijedi slijedeća propozicija.

Propozicija 3.1.6.

Neka je $\Psi \in \mathbf{C}^{m\times m}$ dijagonalna matrica oblika

$$\Psi = \mathrm{diag}(e^{i\psi_1}, \ldots, e^{i\psi_m}) \quad , \quad \psi_1, \ldots, \psi_m \in \mathbf{R} \quad . \tag{3.1.1}$$

i neka je H dijagonalna regularna matrica s realnim elementima. Matrica Ψ je H–unitarna. Ako je matrica U H–unitarna, onda su to i matrice

$$U_1 = \Psi U \quad i \quad U_2 = U\Psi \quad .$$

Dokaz:

Neka je H dijagonalna matrica, $H = \mathrm{diag}(h_{11}, \ldots, h_{mm})$. Onda je

$$\begin{aligned} \Psi^* H \Psi &= \mathrm{diag}(e^{-i\psi_1}, \ldots, e^{-i\psi_m})\ \mathrm{diag}(h_{11}, \ldots, h_{mm})\ \mathrm{diag}(e^{i\psi_1}, \ldots, e^{i\psi_m}) \\ &= \mathrm{diag}(h_{11}, \ldots, h_{mm}) = H \quad . \end{aligned}$$

Primjenom propozicije 2.1.1. slijedi da je Ψ H–unitarna matrica. Nadalje, prema propoziciji 3.1.2., H–unitarne matrice čine grupu, odakle slijedi druga tvrdnja. ∎

3.2. Elementarne matrice J–rotacija

Za praksu su najvažniji indefinitni skalarni produkti generirani matricama J oblika

$$J = \mathrm{diag}(j_{11}, \ldots, j_{mm}) \quad , \quad j_{ii} \in \{1, -1\} \quad . \tag{3.2.1}$$

Takve matrice su hermitske i regularne, a opisuju inerciju regularnih hermitskih matrica.

Za J–skalarni produkt, relacijom (2.1.2) definirana je "J–norma". U ovom jednostavnom slučaju, za svaki vektor $x \in \mathbf{C}^m$ je

$$\|x\|_J = \sqrt{|\,[x,x]\,|} = \sqrt{\left| \sum_{\substack{i=1 \\ j_{ii}=1}}^{m} |x_i|^2 - \sum_{\substack{i=1 \\ j_{ii}=-1}}^{m} |x_i|^2 \right|} \quad . \tag{3.2.2}$$

U nastavku ovog odjeljka konstruiramo matrice rotacija obzirom na indefinitni skalarni produkt zadan matricom (3.2.1). Takve matrice su najjednostavnije matrice koje čuvaju "J–norme" i J–skalarne produkte vektora.

Definicija i oznake za unitarne blok matrice i J–unitarne blok matrice preuzete su iz [49]. Zbog potpunosti, navodimo i osnovna svojstva dokazana, također, u [49].

3.2.1. Elementarne matrice trigonometrijskih i hiperbolnih rotacija

Najjednostavnije unitarne matrice su ravninske ili trigonometrijske rotacije. U realnom slučaju to su matrice oblika

$$U = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \quad . \tag{3.2.3}$$

Postoji nekoliko generalizacija ravninskih rotacija na kompleksan slučaj. Standardno se, u Jacobijevim metodama, koriste matrice oblika

$$U = \begin{bmatrix} \cos\varphi & e^{i\alpha}\sin\varphi \\ -e^{-i\alpha}\sin\varphi & \cos\varphi \end{bmatrix} \quad ,$$

ali je moguća i generalizacija kao u [30]

$$U = \begin{bmatrix} e^{i\alpha}\cos\varphi & e^{i\alpha}\sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \quad .$$

Za jednostrane transformacije, poput QR i JQR dekompozicije, jednostavnije je koristiti malo drugačiju generalizaciju ravninskih rotacija u obliku

$$U = \begin{bmatrix} e^{i\psi_1}\cos\varphi & e^{i\psi_1}\sin\varphi \\ -e^{i\psi_2}\sin\varphi & e^{i\psi_2}\cos\varphi \end{bmatrix} = \begin{bmatrix} e^{i\psi_1} & 0 \\ 0 & e^{i\psi_2} \end{bmatrix} \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \quad . \tag{3.2.4}$$

Korištenjem propozicije 3.1.6. izlazi da je ova matrica, također, unitarna.

Polazeći od elementarnih ravninskih rotacija, možemo konstruirati $m \times m$ matrice rotacija. Označimo s $U_G(k,\ell) \in \mathbf{C}^{m\times m}$ Givensovu ravninsku rotaciju, koja je jednaka jediničnoj matrici, osim u (k,ℓ) ravnini, gdje je

$$U_G(k,\ell)([k,\ell],[k,\ell]) = U \quad , \tag{3.2.5}$$

a U je definirana relacijom (3.2.4), odnosno (3.2.3), ako konstruiramo realne rotacije.

Također, označimo s

$$U_G = \prod_i U_G(k_i,\ell_i) \quad , \quad \{k_i,\ell_i\} \cap \{k_j,\ell_j\} = \emptyset \quad , \quad i \neq j \tag{3.2.6}$$

produkt ravninskih rotacija koje djeluju nezavisno.

Za matrice skalarnog produkta $J = \mathrm{diag}(1, -1)$ ili $J = \mathrm{diag}(-1, 1)$, analogon rotacija su takozvane hiperbolne rotacije. U realnom slučaju, one imaju oblik

$$U = \begin{bmatrix} \mathrm{ch}\, x & \mathrm{sh}\, x \\ \mathrm{sh}\, x & \mathrm{ch}\, x \end{bmatrix} \quad . \tag{3.2.7}$$

Kompleksna generalizacija takvih matrica slična je kompleksnoj generalizaciji ravninskih rotacija. U implicitnim Jacobijevim metodama koriste se matrice oblika

$$U = \begin{bmatrix} \mathrm{ch}\,\varphi & e^{i\alpha}\,\mathrm{sh}\,\varphi \\ e^{-i\alpha}\,\mathrm{sh}\,\varphi & \mathrm{ch}\,\varphi \end{bmatrix} \quad ,$$

ili

$$U = \begin{bmatrix} e^{i\alpha}\,\mathrm{ch}\,\varphi & e^{i\alpha}\,\mathrm{sh}\,\varphi \\ \mathrm{sh}\,\varphi & \mathrm{ch}\,\varphi \end{bmatrix} \quad .$$

Kao i kod kompleksnih rotacija, za potrebe jednostranih transformacija jednostavnije je koristiti matrice oblika

$$U = \begin{bmatrix} e^{i\psi_1}\,\mathrm{ch}\,\varphi & e^{i\psi_1}\,\mathrm{sh}\,\varphi \\ e^{i\psi_2}\,\mathrm{sh}\,\varphi & e^{i\psi_2}\,\mathrm{ch}\,\varphi \end{bmatrix} = \begin{bmatrix} e^{i\psi_1} & 0 \\ 0 & e^{i\psi_2} \end{bmatrix} \begin{bmatrix} \mathrm{ch}\,\varphi & \mathrm{sh}\,\varphi \\ \mathrm{sh}\,\varphi & \mathrm{ch}\,\varphi \end{bmatrix} \quad , \tag{3.2.8}$$

koje su, također, J–unitarne, prema propoziciji 3.1.6.

Hiperbolna rotacija $U_H(k, \ell) \in \mathbf{C}^{m \times m}$, u (k, ℓ) ravnini jednaka je jediničnoj matrici, osim na presjeku k–tog i ℓ–tog retka i stupca, gdje je

$$U_H(k, \ell)([k, \ell], [k, \ell]) = U \quad , \tag{3.2.9}$$

a matrica U je definirana relacijom (3.2.8) u kompleksnom, odnosno (3.2.7) u realnom slučaju.

3.2.2. Elementarne blok unitarne i blok J–unitarne matrice

Kao što ćemo poslije pokazati, da bismo mogli napraviti analogon QR dekompozicije, potrebne su nam blok J–unitarne matrice. Blok–generalizaciju unitarnih i J–unitarnih matrica dao je Veselić u [49]. U ovom paragrafu rezimirat ćemo neka njihova svojstva.

Za generalizaciju, potrebna je definicija pozitivnih potencija matrice.

Definicija 3.2.1.

Funkcija $A \mapsto A^\alpha$, $\alpha \in \mathbf{R}_0^+$, definirana je na skupu svih matrica $\mathcal{A}$ koje zadovoljavaju:

(i) *matrice iz $\mathcal{A}$ nemaju negativnih realnih svojstvenih vrijednosti;*

(ii) 0 *je u najgorem slučaju nedegenerirana svojstvena vrijednost, tj. pripadni elementarni divizori su linearni.*

Na podskupu $\mathcal{A}_0$ svih regularnih matrica iz $\mathcal{A}$, funkcija $A \mapsto A^\alpha$ definira se kao analitičko produljenje reda

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} (A-I)^k \quad ,$$

a zatim se proširuje na $\mathcal{A}$ po neprekidnosti. ■

Za dokaz unitarnosti i J–unitarnosti matrica koje ćemo definirati, potrebna nam je slijedeća lema.

Lema 3.2.1.

Neka su zadane matrice $X \in \mathbf{C}^{m\times n}$ i $Y \in \mathbf{C}^{n\times m}$. Za svaku analitičku funkciju f vrijedi

$$\begin{aligned} f(YX)Y &= Yf(XY) \quad , \\ f(XY)X &= Xf(YX) \quad , \end{aligned} \tag{3.2.10}$$

Dokaz:

Svaka analitička matrična funkcija definira se kao red

$$f(A) = \sum_{k=0}^{\infty} \alpha_k A^k \quad .$$

Ako funkciju $f(XY)$, odnosno $f(YX)$ razvijemo u red, nakon množenja toga reda slijeva s Y, odnosno X, slijedi tvrdnja. ■

Definirajmo elementarne blok unitarne i blok J–unitarne matrice.

Teorem 3.2.1. (Veselić)

Neka je matrica U oblika

$$U = \begin{bmatrix} (I-YX)^{1/2} & Y \\ -X & (I-XY)^{1/2} \end{bmatrix} \quad , \tag{3.2.11}$$

pri čemu su X i Y pravokutne matrice, takve da X i Y^* imaju istu dimenziju.

(i) *Matrica U je blok–unitarna matrica, ako je $\|X\|_2 \le 1$ i vrijedi*

$$Y = X^* \quad . \tag{3.2.12}$$

(*ii*) *Neka je J matrica indefinitnog skalarnog produkta unitarna, blok–dijagonalna i hermitska, oblika*

$$J = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \quad , \tag{3.2.13}$$

pri čemu dimenzije blokova J_1 i J_2 odgovaraju dimenzijama matrica YX i XY, respektivno. Ako odaberemo

$$Y = J_1 X^* J_2 \tag{3.2.14}$$

onda je matrica U definirana s (3.2.11) *J–unitarna.*

Dokaz:

U slučaju (i), treba pokazati da vrijedi $U^*U = I$, a u slučaju (ii), $U^*JU = J$. U oba slučaja dokaz se provodi množenjem matrica i korištenjem leme 3.2.1., pri čemu je funkcija $f(X) = \sqrt{X}$. ■

Elegancija formule (3.2.11) očituje se u inverzu sličnog oblika.

Teorem 3.2.2. (Veselić)

Inverz matrice definirane relacijom (3.2.11) *je*

$$U^{-1} = \begin{bmatrix} (I - YX)^{1/2} & -Y \\ X & (I - XY)^{1/2} \end{bmatrix} \quad . \tag{3.2.15}$$

Dokaz:

Treba pokazati da vrijedi $U^{-1}U = UU^{-1} = I$. Dokaz se, uz korištenje leme 3.2.1., provodi množenjem matrica. ■

Promatramo li blokove u matrici U, oni bi trebali biti blok generalizacija sinusa i kosinusa. Primijetimo da dijagonalni blokovi koji bi trebali odgovarati kosinusu nisu jednaki, a čak ne moraju biti niti jednake dimenzije. Označimo ta dva "sinusa" i dva "kosinusa" sa

$$U = \begin{bmatrix} (I - YX)^{1/2} & Y \\ -X & (I - XY)^{1/2} \end{bmatrix} = \begin{bmatrix} C_1 & S_2 \\ -S_1 & C_2 \end{bmatrix} \quad .$$

Ako dijeljenje interpretiramo kao množenje inverznom matricom i poštujemo dimenzije blokova, možemo na dva načina definirati matrične analogone tangensa

$$\begin{aligned} T_1 &= S_1 C_1^{-1} = X(I - YX)^{-1/2} = (I - XY)^{-1/2} X \quad , \\ T_2 &= S_2 C_2^{-1} = Y(I - XY)^{-1/2} = (I - YX)^{-1/2} Y \quad . \end{aligned} \tag{3.2.16}$$

Uz pomoć ta dva matrična tangensa, možemo izraziti dva matrična sinusa.

Teorem 3.2.3.

Uz oznake (3.2.16)*, podmatrice X i Y u relaciji* (3.2.11) *možemo izraziti kao*

$$\begin{aligned} X &= T_1(I+T_2T_1)^{-1/2} = (I+T_1T_2)^{-1/2}T_1 \quad , \\ Y &= T_2(I+T_1T_2)^{-1/2} = (I+T_2T_1)^{-1/2}T_2 \quad . \end{aligned} \tag{3.2.17}$$

Dokaz:

Izračunajmo $I + T_1T_2$ koristeći definicione relacije (3.2.16)

$$\begin{aligned} I + T_1T_2 &= I + (I-XY)^{-1/2}XY(I-XY)^{-1/2} \\ &= (I-XY)^{-1/2}((I-XY)+XY)(I-XY)^{-1/2} \\ &= (I-XY)^{-1} \quad . \end{aligned}$$

Odavde slijedi

$$(I+T_1T_2)^{-1/2} = (I-XY)^{1/2} \quad .$$

Množenjem zdesna s T_1 i slijeva s T_2 izlaze dvije tvrdnje. Korištenjem leme 3.2.1. dobivamo preostale dvije tvrdnje. ■

Ako je U blok unitarna, ili blok J–unitarna matrica, postoji veza između dva matrična tangensa.

Propozicija 3.2.1.

Neka je matrica U definirana relacijom (3.2.11)*, a matrični tangensi relacijama* (3.2.16)*.*

(i) *Ako je U unitarna matrica, onda je*

$$T_2 = T_1^* \quad . \tag{3.2.18}$$

(ii) *Ako je U J–unitarna matrica, obzirom na J definiran s* (3.2.13)*, onda je*

$$T_2 = J_1T_1^*J_2 \quad . \tag{3.2.19}$$

Dokaz:

(i) U slučaju unitarnosti matrice U, vrijedi

$$\begin{aligned} T_1 &= X(I-X^*X)^{-1/2} = (I-XX^*)^{-1/2}X \quad , \\ T_2 &= X^*(I-XX^*)^{-1/2} = (I-X^*X)^{-1/2}X^* \quad . \end{aligned}$$

Tvrdnja slijedi transponiranjem prve jednadžbe.

(ii) Ako je U J–unitarna matrica, obzirom na J zadan s (3.2.13), vrijedi

$$\begin{aligned} T_1 &= X(I - J_1 X^* J_2 X)^{-1/2} = (I - X J_1 X^* J_2)^{-1/2} X \quad , \\ T_2 &= J_1 X^* J_2 (I - X J_1 X^* J_2)^{-1/2} = (I - J_1 X^* J_2 X)^{-1/2} J_1 X^* J_2 \quad . \end{aligned}$$

Transponiranjem prve jednadžbe slijedi

$$T_1^* = (I - X^* J_2 X J_1)^{-1/2} X^* = X^* (I - J_2 X J_1 X^*)^{-1/2} \quad .$$

Množenjem prethodne jednakosti slijeva s J_1, i zdesna s J_2, dobivamo

$$J_1 T_1^* J_2 = J_1 (I - X^* J_2 X J_1)^{-1/2} X^* J_2 = J_1 X^* (I - J_2 X J_1 X^*)^{-1/2} J_2 \quad .$$

Promotrimo prvu jednakost. Koristeći propoziciju 3.1.1. tako da za jednu varijablu u spomenutoj propoziciji uzmemo $X^* J_2$, a za drugu $X J_1$, dobivamo

$$J_1 T_1^* J_2 = J_1 X^* J_2 (I - X J_1 X^* J_2)^{-1/2} = T_2 \quad .$$

■

3.3. Konstrukcija još nekih H–rotacija

U prošlom odjeljku konstruirali smo elementarne J–unitarne blok matrice, pri čemu je J unitaran, hermitski i blok–dijagonalan. U ovom odjeljku iskoristit ćemo J–unitarne matrice za konstrukciju nekih drugih H–unitarnih matrica.

U standardnom euklidskom prostoru matrica rotacija definira se kao unitarna matrica koja čuva orijentaciju prostora. Na sličan način možemo definirati rotaciju obzirom na bilo koji skalarni produkt.

Definicija 3.3.1.

H–unitarna matrica U je matrica H–rotacije, ako vrijedi

$$\det U = 1 \quad . \tag{3.3.1}$$

■

Svaka 2×2 regularna hermitska matrica može se dijagonalizirati jednim korakom Jacobijeve metode (2×2 rotacijom), a svojstvene vrijednosti tako dijagonalizirane matrice su dva realna broja d_{11} i d_{22}, različita od nule. Zbog toga je dovoljno promatrati realne unitarne matrice obzirom na dijagonalnu matricu

$$\Delta = \mathrm{diag}(d_{11}, d_{22}) \quad . \tag{3.3.2}$$

Pokažimo da su elementarne trigonometrijske i hiperbolne rotacije predstavnici jedina dva bitna tipa ravninskih rotacija.

Propozicija 3.3.1.

Neka je Δ dijagonalna matrica reda 2 definirana relacijom (3.3.2). Obzirom na Δ, postoje točno dvije bitno različite vrste Δ–rotacija, ovisno o tome da li su znakovi d_{11} i d_{22} jednaki ili različiti.

Dokaz:

Pretpostavimo da je U Δ–rotacija oblika

$$U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \quad .$$

Iz relacije (2.1.6) slijedi

$$U^{-1} = \Delta^{-1} U^* \Delta \quad ,$$

i vrijedi

$$\det U = u_{11}u_{22} - u_{12}u_{21} = 1 \quad ,$$

pa za U^{-1} dobivamo jednadžbu

$$\begin{bmatrix} u_{22} & -u_{12} \\ -u_{21} & u_{11} \end{bmatrix} = \begin{bmatrix} d_{11}^{-1} & 0 \\ 0 & d_{22}^{-1} \end{bmatrix} \begin{bmatrix} \bar{u}_{11} & \bar{u}_{21} \\ \bar{u}_{12} & \bar{u}_{22} \end{bmatrix} \begin{bmatrix} d_{11} & 0 \\ 0 & d_{22} \end{bmatrix} = \begin{bmatrix} \bar{u}_{11} & d_{22}d_{11}^{-1}\bar{u}_{21} \\ d_{11}d_{22}^{-1}\bar{u}_{12} & \bar{u}_{22} \end{bmatrix} \quad .$$

Odavde odmah slijede jednadžbe za elemente

$$\begin{aligned} u_{11} &= \bar{u}_{22} \\ u_{12} &= -\frac{d_{22}}{d_{11}}\bar{u}_{21} \end{aligned} \tag{3.3.3}$$

Korištenjem jednakosti (3.3.3), za determinantu rotacije izlazi

$$|u_{11}|^2 + \frac{d_{22}}{d_{11}}|u_{21}|^2 = 1 \quad ,$$

odnosno

$$|u_{21}|^2 = \frac{d_{11}}{d_{22}}(1 - |u_{11}|^2) \quad . \tag{3.3.4}$$

U ovoj relaciji imamo dvije bitno različite mogućnosti:

(a) d_{11} i d_{22} su istog znaka. U tom slučaju, zbog nenegativnosti lijeve strane u (3.3.4), nužno je $|u_{11}| \leq 1$. Zbog toga, $|u_{11}|$ možemo interpretirati kao kosinus nekog kuta, tj.

$$u_{11} = e^{i\alpha}\cos\varphi \quad .$$

Korištenjem (3.3.3) i (3.3.4), izlazi

$$|u_{21}|^2 = \frac{d_{11}}{d_{22}}\sin^2\varphi \quad ,$$

odnosno

$$u_{21} = \mp\sqrt{\frac{d_{11}}{d_{22}}}e^{-i\beta}\sin\varphi \quad .$$

U tom slučaju Δ–rotacija ima oblik

$$U = \begin{bmatrix} e^{i\alpha}\cos\varphi & \pm\sqrt{\frac{d_{11}}{d_{22}}}e^{i\beta}\sin\varphi \\ \mp\sqrt{\frac{d_{11}}{d_{22}}}e^{-i\beta}\sin\varphi & e^{-i\alpha}\cos\varphi \end{bmatrix} \quad ; \tag{3.3.5}$$

(b) d_{11} i d_{22} su različitog znaka. Ponovno, iz nenegativnosti lijeve strane u (3.3.4) slijedi da je $|u_{11}| \geq 1$. U tom slučaju $|u_{11}|$ možemo interpretirati kao hiperbolni kosinus nekog kuta, pa je

$$u_{11} = e^{i\alpha}\,\mathrm{ch}\,\varphi \quad .$$

Za u_{21} izlazi

$$u_{21} = \pm\sqrt{-\frac{d_{11}}{d_{22}}}e^{-i\beta}\,\mathrm{sh}\,\varphi \quad .$$

U tom slučaju Δ–rotacija ima oblik

$$U = \begin{bmatrix} e^{i\alpha}\,\mathrm{ch}\,\varphi & \pm\sqrt{-\frac{d_{11}}{d_{22}}}e^{i\beta}\,\mathrm{sh}\,\varphi \\ \pm\sqrt{-\frac{d_{22}}{d_{11}}}e^{-i\beta}\,\mathrm{sh}\,\varphi & e^{-i\alpha}\,\mathrm{ch}\,\varphi \end{bmatrix} \quad . \tag{3.3.6}$$

■

Prethodna propozicija pokazuje da se Δ–rotacije dobivaju na jednostavan način – skaliranjem odgovarajućih J–rotacija, ovisno o odnosu predznaka u Δ.

Prethodnu konstrukciju možemo generalizirati na proizvoljnu realnu regularnu dijagonalnu matricu Δ oblika

$$\Delta = \mathrm{diag}(d_{11}, \ldots, d_{mm}) \quad . \tag{3.3.7}$$

Propozicija 3.3.2.

Neka je Δ realna regularna dijagonalna matrica definirana relacijom (3.3.7) *i neka je matrica J definirana s*

$$J = \mathrm{diag}(j_{11}, \ldots, j_{mm}) \quad , \quad j_{ii} = \mathrm{sign}\, d_{ii} \quad .$$

Matrica U je Δ–unitarna matrica, ako i samo ako je matrica

$$\widetilde{U} = |\Delta|^{1/2}U|\Delta|^{-1/2}$$

J–unitarna.

Dokaz:

Uočimo da vrijedi

$$\Delta = |\Delta|^{1/2} J |\Delta|^{1/2} \quad . \tag{3.3.8}$$

Množenjem relacije $U^* \Delta U = \Delta$ slijeva i zdesna s $|\Delta|^{-1/2}$ i primjenom (3.3.8), dobivamo

$$|\Delta|^{-1/2} U^* \Delta U |\Delta|^{1/2} = (|\Delta|^{-1/2} U^* |\Delta|^{1/2}) J (|\Delta|^{1/2} U |\Delta|^{1/2}) = J \quad ,$$

pa je Δ–unitarnost matrice U ekvivalentna J–unitarnosti matrice $\widetilde{U}$. ■

Prethodna propozicija, zajedno s propozicijom 3.1.3., barem u principu, svodi konstrukciju H–unitarnih matrica za bilo kakav hermitski regularan H, na konstrukciju J–unitarnih matrica. Naime, svaki takav H može se dijagonalizirati unitarnom sličnošću, tj. postoje unitarna matrica V i dijagonalna matrica Δ, takve da je

$$V^* H V = \Delta \quad .$$

Propozicija 3.1.3. daje vezu između H–unitarnih i Δ–unitarnih matrica. Rastavom matrice Δ u obliku (3.3.8), konstrukcija Δ–unitarnih matrica, svodi se na konstrukciju J–unitarnih matrica, gdje je J inercija matrice H. To pokazuje da su J–unitarne matrice fundamentalne za reprezentaciju svih H–unitarnih matrica.

Veselićeva konstrukcija [49] J–unitarnih matrica obzirom na blok–dijagonalnu i unitarnu matricu J, može se generalizirati i na matrice oblika

$$\widehat{J} = \begin{bmatrix} 0 & \widehat{J}_1 \\ \widehat{J}_1^* & 0 \end{bmatrix} \quad , \tag{3.3.9}$$

pri čemu je $\widehat{J}_1$ kvadratna i unitarna matrica. Odmah je jasno da je uvjet kvadratičnosti matrice $\widehat{J}_1$ nužan, jer je u protivnom matrica $\widehat{J}$ singularna. Prema propoziciji 3.1.3., dovoljno je pronaći blok–unitarnu matricu koja blok–dijagonalizira matricu $\widehat{J}$ iz (3.3.9), a zatim primijeniti konstrukciju iz prethodnog odjeljka.

Propozicija 3.3.3.

Neka je matrica $\widehat{J}$ zadana relacijom (3.3.9). *Matricu $\widehat{J}$ možemo blok–dijagonalizirati unitarnom sličnošću iz teorema* 3.2.1.

Dokaz:

Uzmimo blok–unitarnu matricu Q oblika

$$Q = \begin{bmatrix} (I - X^* X)^{1/2} & X^* \\ -X & (I - X X^*)^{1/2} \end{bmatrix} \quad , \tag{3.3.10}$$

definiranu relacijama (3.2.11) i (3.2.12). Uz pomoć te matrice pokušajmo blok–dijagonalizirati zadani $\widehat{J}$. U tom slučaju matrica Q mora zadovoljavati

$$Q^{-1} \widehat{J} Q = J = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \quad . \tag{3.3.11}$$

Uočimo da, zbog unitarnosti matrice $\widehat{J}_1$, vrijedi

$$\widehat{J}_1^{-1} = \widehat{J}_1^* \quad . \tag{3.3.12}$$

Iz uvjeta za poništavanje bloka $(2,1)$ u relaciji (3.3.11) i relacije (3.3.12), dobivamo

$$(I - XX^*)^{1/2}\widehat{J}_1^{-1}(I - X^*X)^{1/2} = X\widehat{J}_1X \quad .$$

Množenjem te jednadžbe slijeva s $(I - XX^*)^{-1/2}$ i zdesna s $(I - X^*X)^{-1/2}$, te korištenjem leme 3.2.1., slijedi

$$\begin{aligned}\widehat{J}_1^{-1} &= (I - XX^*)^{-1/2}X\widehat{J}_1X(I - X^*X)^{-1/2} \\ &= (I - XX^*)^{-1/2}X\widehat{J}_1(I - XX^*)^{-1/2}X \quad .\end{aligned}$$

Primjenom matričnog tangensa T_1 iz (3.2.16)

$$T_1 = (I - XX^*)^{-1/2}X \quad ,$$

prethodnu relaciju možemo zapisati u obliku

$$\widehat{J}_1^{-1} = T_1\widehat{J}_1T_1 \quad ,$$

odnosno,

$$I = \widehat{J}_1T_1\widehat{J}_1T_1 \quad . \tag{3.3.13}$$

Jedno rješenje jednadžbe (3.3.13), očito je

$$T_1 = \widehat{J}_1^{-1} \quad .$$

Analogno, iz uvjeta za poništavanje bloka $(1,2)$ u relaciji (3.3.11) i relacije (3.3.12), dobivamo

$$(I - X^*X)^{1/2}\widehat{J}_1(I - XX^*)^{1/2} = X^*\widehat{J}_1^{-1}X^* \quad .$$

Množenjem te jednadžbe slijeva s $(I - X^*X)^{-1/2}$ i zdesna s $(I - XX^*)^{-1/2}$, te korištenjem leme 3.2.1., slijedi

$$\begin{aligned}\widehat{J}_1 &= (I - X^*X)^{-1/2}X^*\widehat{J}_1^{-1}X^*(I - XX^*)^{-1/2} \\ &= (I - X^*X)^{-1/2}X^*\widehat{J}_1^{-1}(I - X^*X)^{-1/2}X^* \quad .\end{aligned}$$

Primjenom matričnog tangensa T_2 iz (3.2.16)

$$T_2 = (I - X^*X)^{-1/2}X^* \quad ,$$

dobivamo

$$\widehat{J}_1 = T_2\widehat{J}_1^{-1}T_2 \quad ,$$

odnosno,

$$I = \widehat{J}_1^{-1} T_2 \widehat{J}_1^{-1} T_2 \quad .$$

Jedno rješenje ove jednadžbe je

$$T_2 = \widehat{J}_1 \quad .$$

Primjenom teorema 3.2.3., iz T_1 i T_2 možemo izračunati matricu X u (3.3.10)

$$X = (I + T_1 T_2)^{-1/2} T_1 = (2I)^{-1/2} \widehat{J}_1^{-1} = \frac{1}{\sqrt{2}} \widehat{J}_1^{-1} \quad ,$$

pa je

$$X^* = \frac{1}{\sqrt{2}} \widehat{J}_1 \quad .$$

Odatle slijedi da je

$$(I - X^* X)^{1/2} = (I - X X^*)^{1/2} = \frac{1}{\sqrt{2}} I \quad .$$

Dakle, unitarna matrica Q oblika (3.3.10), za koju vrijedi (3.3.11), jednaka je

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} I & \widehat{J}_1 \\ -\widehat{J}_1^* & I \end{bmatrix} \quad . \tag{3.3.14}$$

■

Primjenom propozicije 3.1.3., odmah dobivamo slijedeći rezultat.

Korolar 3.3.1.

Za matricu $\widehat{J}$ iz relacije (3.3.9), *$\widehat{J}$–unitarne matrice imaju oblik QUQ^{-1}, gdje je Q definirana relacijom* (3.3.14), *a U je J–unitarna, obzirom na J iz* (3.3.11). ■

Takve J–unitarne matrice U konstruirali smo u prethodnom odjeljku.

3.4. Indefinitne matrice Householderovog tipa

Većinu matričnih algoritama, koju možemo realizirati rotacijama, prirodno možemo realizirati i Householderovim reflektorima. Potpuna analogija vrijedi i za algoritme koji koriste J–rotacije.

U ovom odjeljku, navodimo do sada poznate rezultate o konstrukciji J–Householderovih i blok Householderovih matrica. Za realizaciju indefinitne QR dekompozicije, potrebne su još i blok J–Householderove matrice.

Blok J–Householderove matrice prirodna su, ali netrivijalna, generalizacija J–Householderovih i blok–Householderovih matrica. Pokazuje se da blok J–Householderove matrice imaju svojstva reflektora, ako i samo ako, ne dolazi do degeneracije, tj. smanjivanja ranga produkta matrica koje se koriste u konstrukciji blok J–reflektora. Taj specifičan problem nastaje zbog degeneracije indefinitnog skalarnog produkta.

3.4.1. J–Householderove matrice

Klasičnu definiciju Householderovih matrica

$$H = I - 2\frac{ww^*}{w^*w} \quad , \quad w \in \mathbf{C}^m \quad , \quad w \neq 0 \tag{3.4.1}$$

možemo generalizirati na J–Householderove matrice, tj. na Householderove matrice unitarne obzirom na J–skalarni produkt. Postoje dvije generalizacije tih matrica [12] i [15]. Između njih postoji jednostavna veza, ali usprkos jednostavnoj vezi, svojstva su im bitno različita.

Za korektnu definiciju J–Householderovih matrica, potrebna je definicija generaliziranog inverza zadane matrice. Generalizirani inverz je proširenje uobičajenog pojma inverza na pravokutne i singularne matrice. Za kvadratne, regularne matrice, generalizirani inverz jednak je običnom inverzu. Navodimo jednu od tri ekvivalentne definicije.

Definicija 3.4.1.

Generalizirani (Moore–Penroseov) inverz matrice $A \in \mathbf{C}^{m\times n}$ je jedinstvena matrica $A^+ \in \mathbf{C}^{n\times m}$ koja zadovoljava slijedeće jednakosti:

$$\begin{aligned} AA^+A &= A \quad , & (AA^+)^* &= AA^+ \quad , \\ A^+AA^+ &= A^+ \quad , & (A^+A)^* &= A^+A \quad . \end{aligned}$$

■

Dokaz egzistencije i jedinstvenosti matrice A^+ može se naći, na primjer, u [13].

Za matrice jednostavnog oblika, generalizirani inverzi su očiti. Ako je matrica reda 1, tj. matrica je broj a, njezin je generalizirani inverz

$$a^+ = \begin{cases} 1/a & \text{za } a \neq 0 \quad , \\ 0 & \text{za } a = 0 \quad . \end{cases}$$

Koristeći generalizirani inverz broja, za dijagonalnu matricu $A = \operatorname{diag}(a_{11}, \ldots, a_{mm})$ vrijedi

$$A^+ = \operatorname{diag}(a_{11}^+, \ldots, a_{mm}^+) \quad . \tag{3.4.2}$$

Pokažimo još jedno važno svojstvo generaliziranih inverza.

Lema 3.4.1.

Neka je $A \in \mathbf{C}^{m\times n}$ *i neka su* $U \in C^{m\times m}$ *i* $V^* \in C^{n\times n}$ *unitarne matrice. Tada vrijedi*

$$(UAV^*)^+ = VA^+U^* \quad .$$

Dokaz:

Pokažimo da matrica $X = VA^+U^*$ zadovoljava sve četiri relacije u definiciji generaliziranog inverza matrice UAV^*.

$$\begin{aligned}
UAV^*X^+UAV^* &= UAV^*VA^+U^*UAV^* = UAA^+AV^* \\
&= UAV^* \\
X^+UAV^*X^+ &= VA^+U^*UAV^*VA^+U^* = VA^+AA^+U^* \\
&= VA^+U^* = X \quad , \\
(UAV^*X^+)^* &= (UAV^*VA^+U^*)^* = (UAA^+U^*)^* \\
&= UAA^+U^* = UAV^*VA^+U^* = UAV^*X^+ \quad , \\
(X^+UAV^*)^* &= (VA^+U^*UAV^*)^* = (VA^+AV^*)^* \\
&= VA^+AV^* = VA^+U^*UAV^* = X^+UAV^* \quad .
\end{aligned}$$

Zaključak $X = (UAV^*)^+$ slijedi zbog jedinstvenosti generaliziranog inverza. ■

Slobodnija definicija Householderove matrice je da je to unitarna matrica, koja je modifikacija jedinične matrice matricom niskog ranga. Prenesemo li značenje na J–Householderove matrice, ta definicija se može interpretirati na jedan od slijedeća dva načina. U prvom slučaju, J–Householderova matrica je J–unitarna matrica koja je modifikacija jedinične matrice matricom niskog ranga (definicija Bunse–Gerstner), a u drugom (definicija Cybenko–Berry) je to J–unitarna matrica koja je modifikacija matrice J matricom niskog ranga.

Definicija 3.4.2. (Bunse–Gerstner)

Za zadani vektor $w \in \mathbf{C}^m$, $w \neq 0$ *i dijagonalnu matricu* $J = \mathrm{diag}(j_{11}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$, *J–Householderova matrica* $H \in \mathbf{C}^{m\times m}$ *je*

$$H = I - 2w(w^*Jw)^+w^*J \quad . \tag{3.4.3}$$

■

Definicija 3.4.3. (Cybenko, Berry)

Za zadani vektor $w \in \mathbf{C}^m$, $w \neq 0$ *i dijagonalnu matricu* $J = \mathrm{diag}(j_{11}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$, *J–Householderova matrica* $\widetilde{H} \in \mathbf{C}^{m\times m}$ *je*

$$\widetilde{H} = J - 2w(w^*Jw)^+w^* \quad . \tag{3.4.4}$$

■

Slijedeće dvije propozicije daju osnovna svojstva J–Householderovih matrica definiranih prethodnim dvjema definicijama.

Propozicija 3.4.1.

Matrica H definirana relacijom (3.4.3) *ima slijedeća svojstva:*

(i) *H je J–unitarna matrica,*

(ii) *H je J–hermitska matrica,*

(iii) $H^{-1} = H$.

Dokaz:

Pokažimo da za matricu H vrijedi relacija (2.1.6). Korištenjem

$$H^* = I - 2Jw(w^*Jw)^+w^* \quad ,$$

izlazi

$$\begin{aligned} H^*JH &= (I - 2Jw(w^*Jw)^+w^*)J(I - 2w(w^*Jw)^+w^*J) \\ &= J - 4Jw(w^*Jw)^+w^*J + 4Jw(w^*Jw)^+w^*Jw(w^*Jw)^+w^*J \\ &= J - 4Jw(w^*Jw)^+w^*J + 4Jw(w^*Jw)^+w^*J = J \quad , \end{aligned}$$

čime je dokazana J–unitarnost.

Prema propoziciji 2.1.1., matrica H je J–hermitska, ako vrijedi (2.1.5). Iz definicije matrice H izlazi

$$\begin{aligned} JHJ^{-1} &= JHJ = J(I - 2w(w^*Jw)^+w^*J)J \\ &= I - 2Jw(w^*Jw)^+w^* = H^* \quad , \end{aligned}$$

čime je dokazano (2.1.5).

Tvrdnja (iii), direktna je posljedica J–unitarnosti i J–hermitičnosti. Vrijedi

$$I = JJ = JH^*JH = H^2 \quad ,$$

pa odmah slijedi da je $H^{-1} = H$. ■

Propozicija 3.4.2.

Matrica $\widetilde{H}$ definirana relacijom (3.4.4) *ima slijedeća svojstva:*

(i) *$\widetilde{H}$ je J–unitarna matrica,*

(ii) *$\widetilde{H}$ je hermitska matrica,*

(iii) $\widetilde{H}^{-1} = J\widetilde{H}J$.

Dokaz:

Uočimo da je

$$\widetilde{H}^* = \widetilde{H} = J - 2w(w^*Jw)^+w^* \quad ,$$

što pokazuje hermitičnost zadane matrice.

Korištenjem svojstva hermitičnosti, pokažimo da vrijedi relacija (2.1.6), tj. da je matrica $\widetilde{H}$ J–unitarna. Vrijedi

$$\begin{aligned}
\widetilde{H}^*J\widetilde{H} &= \widetilde{H}J\widetilde{H} = (J - 2w(w^*Jw)^+w^*)J(J - 2w(w^*Jw)^+w^*) \\
&= J - 4w(w^*Jw)^+w^* + 4w(w^*Jw)^+w^*Jw(w^*Jw)^+w^* \\
&= J - 4w(w^*Jw)^+w^* + 4w(w^*Jw)^+w^* = J \quad .
\end{aligned}$$

Posljednja tvrdnja je posljedica J–unitarnosti i hermitičnosti matrice $\widetilde{H}$. Izlazi

$$I = JJ = J\widetilde{H}^*J\widetilde{H} = J\widetilde{H}J\widetilde{H} \quad ,$$

pa je $\widetilde{H}^{-1} = J\widetilde{H}J$. ∎

Uočimo da u definiciji J–Householderovih matrica postoji mogućnost degeneracije. Ako je $w^*Jw = 0$, onda je i $(w^*Jw)^+ = 0$, pa su Householderove matrice jednake $H = I$, odnosno $\widetilde{H} = J$. U tom slučaju, kao što ćemo pokazati, J–Householderove matrice nemaju svojstvo reflektora.

3.4.2. Blok Householderove matrice

Sredinom sedamdesetih godina su Brønlund i Lunde Johnsen [6], a zatim i Dietrich [16] promatrali blok reflektore u svrhu poništavanja blokova u matrici. Zaokruženu teoriju blok reflektora dali su Schreiber i Parlett u [37].

U ovom paragrafu, zbog potpunosti i generalizacije na indefinitni slučaj, ukratko navodimo definiciju i osnovna svojstva blok reflektora.

Definicija 3.4.4.

Za zadanu matricu $W \in \mathbf{C}^{m\times p}$, $m \geq p$, $W \neq 0$, reflektor koji zrcali $\mathcal{R}(W)$, sliku od W, obzirom na $\mathcal{R}(W)^\perp$, definiran je relacijom

$$H = H(W) = I - 2W(W^*W)^+W^* \quad . \tag{3.4.5}$$

∎

Pokažimo osnovna svojstva matrice H definirane relacijom (3.4.5).

Propozicija 3.4.3.

Matrica H definirana relacijom (3.4.5) *je unitarna i hermitska.*

Dokaz:

Posljednje svojstvo iz definicije generaliziranog inverza

$$((W^*W)^+)^* = (W^*W)^+$$

osigurava da je

$$H^* = I - 2W((W^*W)^+)^*W^* = I - 2W(W^*W)^+W^* = H$$

pa je H hermitska.

Pokažimo još i unitarnost matrice H. Vrijedi

$$\begin{aligned} HH^* &= (I - 2W(W^*W)^+W^*)(I - 2W(W^*W)^+W^*) \\ &= I - 4W(W^*W)^+W^* + 4W(W^*W)^+W^*W(W^*W)^+W^* \\ &= I - 4W(W^*W)^+W^* + 4W(W^*W)^+W^* = I \quad . \end{aligned}$$

■

Hermitičnost i unitarnost matrice H karakteriziraju reflektor, tj. za matricu H vrijedi

$$H^2 = I \quad .$$

Za dokaz osnovnog teorema o reflektiranju potprostora, potrebna je lema o "kraćenju" generaliziranog inverza produkta matrica.

Lema 3.4.2.

Za proizvoljnu matricu $W \in \mathbf{C}^{m\times p}$, vrijedi

$$W(W^*W)^+W^*W = W \quad . \tag{3.4.6}$$

Dokaz:

Dokažimo jednakost (3.4.6) primjenom leme 3.4.1. na matrice iz dekompozicije singularnih vrijednosti matrice W^*. Neka je

$$W^* = U\Sigma V^* \quad , \tag{3.4.7}$$

pri čemu su U i V unitarne matrice reda m, odnosno p, respektivno, a Σ je pravokutna matrica tipa (m, k), takva da je $\sigma_{ij} = 0$ za $i \neq j$. Označimo sa σ_i, $i = 1, \ldots \ell$, $\ell \leq \min\{m, p\}$, netrivijalne dijagonalne elemente matrice Σ (netrivijalne singularne vrijednosti matrice W^*).

Primjenom leme 3.4.1. dobivamo

$$\begin{aligned} W(W^*W)^+W^*W &= V\Sigma^*U^*(U\Sigma V^*V\Sigma^*U^*)^+U\Sigma V^*V\Sigma^*U^* \\ &= V\Sigma^*U^*(U\Sigma\Sigma^*U^*)^+U\Sigma\Sigma^*U^* \\ &= V\Sigma^*U^*U(\Sigma\Sigma^*)^+U^*U\Sigma\Sigma^*U^* \\ &= V\Sigma^*(\Sigma\Sigma^*)^+\Sigma\Sigma^*U^* \quad . \end{aligned}$$

Uočimo da vrijedi

$$\Sigma\Sigma^* = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_\ell^2, 0, \ldots, 0) \quad .$$

Korištenjem relacije (3.4.2) izlazi

$$(\Sigma\Sigma^*)^+ = \mathrm{diag}(1/\sigma_1^2, \ldots, 1/\sigma_\ell^2, 0, \ldots, 0) \quad .$$

Zbog toga je

$$(\Sigma\Sigma^*)^+\Sigma\Sigma^* = \begin{bmatrix} I_\ell & 0 \\ 0 & 0 \end{bmatrix} \in \mathbf{C}^{m\times m} \quad .$$

Odavde vrijedi zaključak

$$W(W^*W)^+W^*W = V\Sigma^* \begin{bmatrix} I_\ell & 0 \\ 0 & 0 \end{bmatrix} U^* = V\Sigma^*U^* = W \quad .$$

■

Dokažimo teorem o dekompoziciji prostora na dva potprostora: potprostora $\mathcal{R}(W)^\perp$ koji ostaje invarijantan na djelovanje reflektora i potprostora $\mathcal{R}(W)$, koji se zrcali djelovanjem reflektora H.

Teorem 3.4.1.

Neka je matrica H definirana relacijom (3.4.5). *Za sve vektore $w \in \mathcal{R}(W)$, vrijedi*

$$H(W)w = -w \quad ,$$

a za sve $y \in \mathcal{R}(W)^\perp$

$$H(W)y = y \quad .$$

Dokaz:

Neka je $w \in \mathcal{R}(W)$. Tada postoji x takav da je $w = Wx$. Iskoristimo definiciju matrice H i lemu 3.4.2. o "kraćenju" generaliziranog inverza. Dobivamo

$$\begin{aligned} Hw &= w - 2W(W^*W)^+W^*w = Wx - 2W(W^*W)^+W^*Wx \\ &= Wx - 2Wx = -w \quad . \end{aligned}$$

Drugi dio tvrdnje dobivamo primjenom činjenice da je $\mathcal{R}(W)^\perp = \mathcal{N}(W^*)$. Ako je $y \in \mathcal{R}(W)^\perp$, tada je $y \in \mathcal{N}(W^*)$, tj. $W^*y = 0$. Zbog toga je

$$Hy = y - 2W(W^*W)^+W^*y = y \quad .$$

■

3.4.3. Blok J–Householderove matrice

U ovom paragrafu uvodimo definiciju blok J–Householderovih matrica, koje su prirodna generalizacija J–Householderovih matrica. Standardni zadatak, za koji se takve matrice koriste, je poništavanje blokova neke matrice. U odjeljku o realizaciji indefinitne QR dekompozicije pokazat ćemo da je za poništavanje blokova prikladnija generalizacija definicije 3.4.2., tj. generalizacija definicije koju je dala Bunse–Gerstner u [12].

Definicija 3.4.5.

Za zadane matrice $W \in \mathbf{C}^{m \times p}$, $m \geq p$, $W \neq 0$ *i* $J = \mathrm{diag}(j_{11}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$, *blok J–Householderova matrica* $H \in \mathbf{C}^{m \times m}$ *je*

$$H = H(W) = I - 2W(W^*JW)^+W^*J \quad . \tag{3.4.8}$$

■

Pokažimo da vrijedi analogon propozicije 3.4.1.

Propozicija 3.4.4.

Matrica H definirana relacijom (3.4.8) *je J–unitarna, J–hermitska i vrijedi* $H^2 = I$.

Dokaz:

Korištenjem svojstva $(A^*)^+ = (A^+)^*$, (dokaz vidjeti, na primjer u [13]), izlazi da je

$$H^* = I - 2JW(W^*JW)^+W^* = JHJ \quad ,$$

pa je time, prema relaciji (2.1.5), dokazana J–hermitičnost matrice H.

Pokažimo J–unitarnost matrice H. Vrijedi

$$\begin{aligned} H^*JH &= (I - 2JW(W^*JW)^+W^*)J(I - 2W(W^*JW)^+W^*J) \\ &= J - 4JW(W^*JW)^+W^*J + 4JW(W^*JW)^+W^*JW(W^*JW)^+W^*J \\ &= J - 4JW(W^*JW)^+W^*J + 4JW(W^*JW)^+W^*J = J \quad , \end{aligned}$$

pa je, prema relaciji (2.1.6), matrica H J–unitarna.

Preostaje još pokazati da vrijedi $H^{-1} = H$. J–unitarnost i J–hermitičnost matrice H daje

$$H^{-1} = JH^*J = JJHJJ = H \quad ,$$

što je trebalo pokazati. ■

Za konstruktivni dokaz osnovnih svojstava J–unitarne Householderove matrice koristit ćemo teorem o hiperbolnoj dekompoziciji singularnih vrijednosti. Taj teorem se dobiva iz slijedećeg općeg teorema o dekompoziciji para matrica (A, B), dokazanog u [55].

Teorem 3.4.2. (Zha)

Neka su dane hermitska regularna matrica $A \in \mathbf{C}^{m\times m}$ i matrica $G \in \mathbf{C}^{m\times n}$, $m \geq n$. Tada postoje unitarna matrica U i nesingularna matrica Z, tako da vrijedi

$$Z^*AZ = \begin{bmatrix} \Delta_\ell & & \\ & \begin{bmatrix} 0 & I_j \\ I_j & 0 \end{bmatrix} & \\ & & \begin{bmatrix} I_s & 0 \\ 0 & -I_{k-s} \end{bmatrix} \end{bmatrix} \quad ,$$

$$Z^{-1}GU = \begin{bmatrix} I_\ell & & \\ & \begin{bmatrix} I_j \\ 0 \end{bmatrix} & \\ & & 0 \end{bmatrix} \quad ,$$

gdje je $\Delta_\ell = \mathrm{diag}(\lambda_1, \ldots, \lambda_\ell) \in \mathbf{R}^{\ell\times\ell}$ nesingularna i

$$\begin{aligned} j &= \mathrm{rang}(G) - \mathrm{rang}(G^*AG) \quad , \\ k &= \mathrm{rang}(A) + \mathrm{rang}(G^*AG) - 2\,\mathrm{rang}(G) \quad , \\ \ell &= \mathrm{rang}(G^*AG) \quad , \end{aligned}$$

a s je broj pozitivnih svojstvenih vrijednosti matrice A, koji odgovara nedegeneriranom dijelu defekta matrice G.

Dokaz:

Dokaz je konstruktivan i provodi se u nekoliko koraka. U svakom koraku, transformacija ima oblik

$$A^{(q+1)} = Z_{q+1}^* A^{(q)} Z_{q+1} \quad , \quad G^{(q+1)} = Z_{q+1}^{-1} G^{(q)} U_{q+1} \quad ,$$

gdje je Z_{q+1} regularna matrica reda m, a U_{q+1} unitarna reda n. Na početku, stavimo $A^{(0)} = A$, $G^{(0)} = G$.

Korak 1:

Neka je $r = \mathrm{rang}(G)$. Dekompozicija singularnih vrijednosti matrice G ima oblik

$$G = U_0 \Sigma_0 V_0^* = U_0 \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} V_0^* \quad ,$$

pri čemu, zbog $r = \mathrm{rang}(G)$, vrijedi

$$\Sigma_r = \mathrm{diag}(\sigma_1, \ldots, \sigma_r) \quad , \quad \sigma_1, \ldots, \sigma_r \neq 0 \quad ,$$

tj. $\Sigma_r \in \mathbf{C}^{r\times r}$ je regularna dijagonalna matrica. Definiramo

$$G^{(1)} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & I_{m-r} \end{bmatrix} U_0^* G V_0 = Z_1^{-1} G U_1 \quad ,$$

pri čemu je

$$Z_1^{-1} = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & I_{m-r} \end{bmatrix} U_0^* \quad , \quad U_1 = V_0 \quad .$$

Regularnost matrice Z_1 slijedi iz regularnosti obje matrice u produktu. Matrica U_1 je, očito, unitarna.

Pripadna transformacija matrice A ima oblik

$$A^{(1)} = Z_1^* A Z_1 = \begin{bmatrix} A_{11}^{(1)} & A_{12}^{(1)} \\ A_{21}^{(1)} & A_{22}^{(1)} \end{bmatrix} \quad ,$$

a particija matrice $A^{(1)}$ odgovara particiji matrice $G^{(1)}$, tj. matrica $A_{11}^{(1)}$ je reda r, a matrica $A_{22}^{(1)}$ je reda $m-r$. Uočimo da je matrica $A^{(1)}$ još i hermitska.

Korak 2:

Neka je svojstvena dekompozicija bloka $A_{11}^{(1)}$, tj. bloka koji odgovara regularnom dijelu matrice G, jednaka

$$\widetilde{U}_2^* A_{11}^{(1)} \widetilde{U}_2 = \begin{bmatrix} \Delta_\ell & 0 \\ 0 & 0 \end{bmatrix} \quad ,$$

gdje je $\Delta_\ell = \mathrm{diag}(\lambda_1, \ldots, \lambda_\ell)$, $\lambda_i \neq 0$ i

$$\mathrm{sign}(\lambda_i) = \begin{cases} -1 & \text{za } i = 1, \ldots, t, \\ 1 & \text{za } i = t+1, \ldots, \ell \end{cases} \quad . \tag{3.4.9}$$

Nadalje, množenjem matrica, dobivamo da je $\ell = \mathrm{rang}(G^* A G)$. Definiramo matrice $U_2 = \mathrm{diag}(\widetilde{U}_2, I_{n-r})$ i $Z_2 = \mathrm{diag}(\widetilde{U}_2, I_{m-r})$. Tada je

$$G^{(2)} = Z_2^{-1} G^{(1)} U_2 = \begin{bmatrix} \widetilde{U}_2^* & 0 \\ 0 & I_{m-r} \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{U}_2 & 0 \\ 0 & I_{n-r} \end{bmatrix} = G^{(1)} \quad .$$

Za matricu $A^{(2)}$, adekvatnom particijom, dobivamo

$$A^{(2)} = Z_2^* A^{(1)} Z_2 = \begin{bmatrix} \widetilde{U}_2^* & 0 \\ 0 & I_{m-r} \end{bmatrix} \begin{bmatrix} A_{11}^{(1)} & A_{12}^{(1)} \\ A_{21}^{(1)} & A_{22}^{(1)} \end{bmatrix} \begin{bmatrix} \widetilde{U}_2 & 0 \\ 0 & I_{m-r} \end{bmatrix} = \begin{bmatrix} \Delta_\ell & 0 & A_{13}^{(2)} \\ 0 & 0 & A_{23}^{(2)} \\ A_{31}^{(2)} & A_{32}^{(2)} & A_{33}^{(2)} \end{bmatrix} \quad .$$

Korak 3:

U matrici $A^{(2)}$, korištenjem bloka Δ_ℓ kao pivotnog bloka, poništavamo blok $A_{13}^{(2)}$. Zbog hermitičnosti matrice $A^{(2)}$, time smo poništili i blok $A_{31}^{(2)}$. Definiramo

$$Z_3 = \begin{bmatrix} I_\ell & 0 & Z_{13}^{(3)} \\ 0 & I_{r-\ell} & 0 \\ 0 & 0 & I_{m-r} \end{bmatrix} \quad ,$$

pri čemu je $Z_{13}^{(3)} = -\Delta_\ell^{-1} H_{13}^{(2)}$. Matrica Z_3 je gornjetrokutasta s jediničnom dijagonalom, pa je i regularna. Tada matrica $A^{(3)}$ ima oblik

$$A^{(3)} = Z_3^* A^{(2)} Z_3 = \begin{bmatrix} \Delta_\ell & 0 & 0 \\ 0 & 0 & A_{23}^{(3)} \\ 0 & A_{32}^{(3)} & A_{33}^{(3)} \end{bmatrix} \quad .$$

Invertiranjem matrice Z_3 dobivamo

$$Z_3^{-1} = \begin{bmatrix} I_\ell & 0 & -Z_{13}^{(3)} \\ 0 & I_{r-\ell} & 0 \\ 0 & 0 & I_{m-r} \end{bmatrix} \quad ,$$

pa, uz $U_3 = I_n$, vrijedi

$$G^{(3)} = Z_3^{-1} G^{(2)} U_3 \quad .$$

Množenjem odgovarajućih blokova u prethodnoj relaciji, izlazi $G^{(3)} = G^{(2)}$.

Primijetimo da je značenje broja $r - \ell = \mathrm{rang}(G) - \mathrm{rang}(G^* A G)$.

Korak 4:

Uočimo da singularnost matrice $A_{32}^{(3)}$ povlači singularnost matrice A, što je protivno pretpostavci teorema. Zbog toga, dekompozicija singularnih vrijednosti matrice $A_{32}^{(3)}$ ima oblik

$$A_{32}^{(3)} = \widehat{U}_0 \widehat{\Sigma}_0 \widehat{V}_0^* = \widehat{U}_0 \begin{bmatrix} \widehat{\Sigma}_{r-\ell} \\ 0 \end{bmatrix} \widehat{V}_0^* = \widetilde{Z}_4 \begin{bmatrix} I_{r-\ell} \\ 0 \end{bmatrix} \widetilde{U}_4^* \quad ,$$

i vrijedi

$$\widehat{\Sigma}_{r-\ell} = \mathrm{diag}(\hat{\sigma}_1, \ldots, \hat{\sigma}_{r-\ell}) \quad , \quad \hat{\sigma}_1, \ldots, \hat{\sigma}_{r-\ell} \neq 0 \quad .$$

Matrice Z_4 i U_4 dobivamo odgovarajućim sastavljanjem blokova,

$$Z_4 = \mathrm{diag}(I_\ell, \widetilde{U}_4, \widetilde{Z}_4) \quad , \quad U_4 = \mathrm{diag}(I_\ell, \widetilde{U}_4, I_{m-r}) \quad .$$

Poštujući dimenzije blokova matrice $A^{(4)}$, dobivamo

$$A^{(4)} = Z_4^* A^{(3)} Z_4 = \begin{bmatrix} \Delta_\ell & 0 & 0 & 0 \\ 0 & 0 & I_{r-\ell} & 0 \\ 0 & I_{r-\ell} & A_{33}^{(4)} & A_{34}^{(4)} \\ 0 & 0 & A_{43}^{(4)} & A_{44}^{(4)} \end{bmatrix} \quad .$$

Regularnost matrice $\widetilde{Z}_4$ povlači i regularnost matrice Z_4, a unitarnost matrice U_4 posljedica je unitarnosti matrice $\widetilde{U}_4$, pa definiramo

$$G^{(4)} = Z_4^{-1} A^{(3)} U_4 = G^{(3)} \quad .$$

Korak 5:

Iskoristimo blok $I_{r-\ell}$ za eliminaciju blokova $A_{33}^{(4)}$, $A_{34}^{(4)}$ i $A_{43}^{(4)}$. Množenjem odgovarajućih blokova, može se pokazati da za regularnu matricu

$$Z_5 = \begin{bmatrix} I_\ell & 0 & 0 & 0 \\ 0 & I_{r-\ell} & -A_{33}^{(4)}/2 & -A_{34}^{(4)} \\ 0 & 0 & I_{r-\ell} & 0 \\ 0 & 0 & 0 & I_{m-2r+\ell} \end{bmatrix}$$

vrijedi da je

$$A^{(5)} = Z_5^* A^{(4)} Z_5 = \begin{bmatrix} \Delta_\ell & 0 & 0 & 0 \\ 0 & 0 & I_{r-\ell} & 0 \\ 0 & I_{r-\ell} & 0 & 0 \\ 0 & 0 & 0 & A_{44}^{(5)} \end{bmatrix} .$$

Neka je $U_5 = I_n$. Onda je

$$G^{(5)} = Z_5^{-1} G^{(4)} U_5 \quad .$$

Množenjem odgovarajućih blokova u prethodnoj relaciji, možemo pokazati da je $G^{(5)} = G^{(4)}$.

Korak 6:

Uočimo da matrica $A_{44}^{(5)}$ mora biti regularna, jer je nastala kongruencijskim transformacijama regularne matrice. Svojstvena dekompozicija matrice $A_{44}^{(5)}$ ima oblik

$$\widetilde{U}_5^* A_{44}^{(5)} \widetilde{U}_5 = D_{m-2r+\ell} \quad ,$$

gdje je

$$D_{m-2r+\ell} = \operatorname{diag}(d_1, \ldots, d_{m-2r+\ell})$$

i vrijedi

$$\operatorname{sign}(d_i) = \begin{cases} 1 & \text{za } i = 1, \ldots, s, \\ -1 & \text{za } i = s+1, \ldots, m-2r+\ell \end{cases} ,$$

pa, očito, postoji regularna matrica $\widetilde{Z}_6$, takva da je

$$\widetilde{Z}_5^* A_{44}^{(5)} \widetilde{Z}_5 = \begin{bmatrix} I_s & 0 \\ 0 & -I_{m-2r+\ell-s} \end{bmatrix} .$$

Tada možemo definirati

$$Z_6 = \operatorname{diag}(I_\ell, I_{r-\ell}, I_{r-\ell}, \widetilde{Z}_6) \quad , \quad U_6 = I_n \quad .$$

Zbog regularnosti matrice $\widetilde{Z}_6$, možemo definirati

$$A^{(6)} = Z_6^* A^{(5)} Z_6 \quad , \quad G^{(6)} = Z_6^{-1} G^{(5)} U_6 \quad .$$

Množenjem matrica na desnoj strani relacije za $G^{(6)}$ izlazi $G^{(6)} = G^{(5)}$. ∎

Primjenom prethodnog teorema na matrice $A = J$ i $G = W$, možemo dokazati slijedeći teorem o hiperbolnoj dekompoziciji singularnih vrijednosti.

Teorem 3.4.3. (Zha)

Neka su dane matrica $J \in \mathbf{C}^{m\times m}$, $J = \mathrm{diag}(j_{11},\ldots,j_{mm})$, $j_{ii} \in \{1,-1\}$ i matrica $W \in \mathbf{C}^{m\times p}$, $m \geq p$. Tada postoje unitarna matrica $\widehat{U}$, J–unitarna matrica $\widehat{Z}$ i matrica permutacije P, takve da vrijedi

$$W^* = \widehat{U}\Sigma P\widehat{Z}^* = \widehat{U}\begin{bmatrix} \mathrm{diag}(|\lambda_1|^{1/2},\ldots,|\lambda_\ell|^{1/2}) & & \\ & \frac{1}{\sqrt{2}}\begin{bmatrix} I_j \ , \ -I_j \end{bmatrix} & \\ & & 0 \end{bmatrix} P\widehat{Z}^* \quad .$$

Dokaz:

Primijenimo teorem 3.4.2. redom na matrice J i W. Vrijedi

$$Z^*JZ = \begin{bmatrix} \Delta_\ell & & \\ & \begin{bmatrix} 0 & I_j \\ I_j & 0 \end{bmatrix} & \\ & & \begin{bmatrix} I_s & 0 \\ 0 & -I_{k-s} \end{bmatrix} \end{bmatrix} \quad , \qquad Z^{-1}WU = \begin{bmatrix} I_\ell & & \\ & \begin{bmatrix} I_j \\ 0 \end{bmatrix} & \\ & & 0 \end{bmatrix} \quad .$$

gdje je U unitarna, a Z nesingularna matrica.

Tvrdnju teorema dobivamo dijagonalizacijom blokova matrice Z^*JZ. Primjenom relacije (3.4.9), matricu Δ_ℓ možemo napisati u obliku

$$\begin{aligned} \Delta_\ell &= \mathrm{diag}(|\lambda_1|^{1/2},\ldots,|\lambda_\ell|^{1/2}) \begin{bmatrix} -I_t & 0 \\ 0 & I_{\ell-t} \end{bmatrix} \mathrm{diag}(|\lambda_1|^{1/2},\ldots,|\lambda_\ell|^{1/2}) \\ &= |\Delta_\ell|^{1/2} \begin{bmatrix} -I_t & 0 \\ 0 & I_{\ell-t} \end{bmatrix} |\Delta_\ell|^{1/2} \quad . \end{aligned}$$

Prema propoziciji 3.3.3., postoji unitarna matrica Q oblika (3.3.14), koja dijagonalizira blok

$$\begin{bmatrix} 0 & I_j \\ I_j & 0 \end{bmatrix} \quad .$$

Odmah slijedi

$$Q^* \begin{bmatrix} 0 & I_j \\ I_j & 0 \end{bmatrix} Q = \frac{1}{\sqrt{2}} \begin{bmatrix} I_j & -I_j \\ I_j & I_j \end{bmatrix} \begin{bmatrix} 0 & I_j \\ I_j & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} I_j & I_j \\ -I_j & I_j \end{bmatrix} = \begin{bmatrix} -I_j & 0 \\ 0 & I_j \end{bmatrix} \quad .$$

Treći blok u matrici Z^*JZ je već u dijagonalnoj formi. Definiramo matricu

$$T = \mathrm{diag}(|\Delta_\ell|^{1/2}, Q^*, I_{m-2r+\ell}) \quad .$$

Zbog regularnosti matrice T vrijedi

$$T^{-1}Z^*JZT^{-*} = \begin{bmatrix} \begin{bmatrix} -I_t & 0 \\ 0 & I_{\ell-t} \end{bmatrix} & & \\ & \begin{bmatrix} -I_j & 0 \\ 0 & I_j \end{bmatrix} & \\ & & \begin{bmatrix} I_s & 0 \\ 0 & -I_{k-s} \end{bmatrix} \end{bmatrix} , \tag{3.4.10}$$

pa matrica J i matrica s desne strane u (3.4.10) imaju istu inerciju. Drugim riječima, postoji matrica permutacije P, takva da je

$$T^{-1}Z^*JZT^{-*} = PJP^* \quad . \tag{3.4.11}$$

Definiramo

$$\widehat{Z} = ZT^{-*}P \quad , \tag{3.4.12}$$

što, zajedno s (3.4.11), pokazuje J–unitarnost matrice $\widehat{Z}$.

Promotrimo relaciju koja vrijedi za matricu W i iskoristimo definicione relacije za matrice Z i T.

$$W = Z \begin{bmatrix} I_\ell & & \\ & \begin{bmatrix} I_j \\ 0 \end{bmatrix} & \\ & & 0 \end{bmatrix} U^* = \widehat{Z}P^*T^* \begin{bmatrix} I_\ell & & \\ & \begin{bmatrix} I_j \\ 0 \end{bmatrix} & \\ & & 0 \end{bmatrix} U^* = \widehat{Z}P^* \begin{bmatrix} |\Delta_\ell| & & \\ & Q\begin{bmatrix} I_j \\ 0 \end{bmatrix} & \\ & & 0 \end{bmatrix} U^* .$$

Primijetimo da je

$$Q\begin{bmatrix} I_j \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} I_j & I_j \\ -I_j & I_j \end{bmatrix} \begin{bmatrix} I_j \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} I_j \\ -I_j \end{bmatrix} \quad ,$$

odakle, uvrštavanjem u relaciju za W i adjungiranjem, izlazi tvrdnja. ∎

Korolar 3.4.1.

Matrica $\Sigma \in \mathbf{C}^{m\times p}$ *definirana u teoremu* 3.4.3. *je dijagonalna, ako i samo ako iščezava blok*

$$\frac{1}{\sqrt{2}} \begin{bmatrix} I_j \ , \ -I_j \end{bmatrix} \quad ,$$

odnosno, ekvivalentno, ako i samo ako je

$$\operatorname{rang}(W^*JW) = \operatorname{rang}(W) \quad .$$

∎

Za matrice W i J iz definicije blok J–Householderovih matrica vrijedi slična relacija "kraćenja" kao i kod blok Householderovih matrica.

Teorem 3.4.4.

Jednakost

$$W(W^*JW)^+W^*JW = W \quad ,$$

vrijedi, ako i samo ako je

$$\mathrm{rang}(W^*JW) = \mathrm{rang}(W) \quad . \tag{3.4.13}$$

Dokaz:

Prema teoremu 3.4.3. matricu W^* možemo napisati u obliku $W^* = \widehat{U}\Sigma P\widehat{Z}^*$, gdje je $\widehat{U}$ unitarna, a $\widehat{Z}$ J–unitarna matrica. Množenjem izlazi

$$W^*JW = (\widehat{U}\Sigma P\widehat{Z}^*)J(\widehat{Z}P^*\Sigma^*\widehat{U}^*) = \widehat{U}\Sigma PJP^*\Sigma^*\widehat{U}^* = \widehat{U}\begin{bmatrix} 0 & & \\ & \Delta_\ell & \\ & & 0 \end{bmatrix}\widehat{U}^* \quad .$$

Korištenjem leme 3.4.1. i relacije (3.4.2) za generalizirani inverz dijagonalne matrice, dobivamo

$$(W^*JW)^+ = \widehat{U}\begin{bmatrix} 0 & & \\ & \Delta_\ell^{-1} & \\ & & 0 \end{bmatrix}\widehat{U}^* \quad ,$$

što daje

$$\begin{aligned} W(W^*JW)^+W^*JW &= \widehat{Z}P^*\Sigma^*\widehat{U}^*\widehat{U}\begin{bmatrix} 0 & & \\ & \Delta_\ell & \\ & & 0 \end{bmatrix}\widehat{U}^*\widehat{U}\begin{bmatrix} 0 & & \\ & \Delta_\ell^{-1} & \\ & & 0 \end{bmatrix}\widehat{U}^* \\ &= \widehat{Z}P^*\Sigma^*\begin{bmatrix} 0 & & \\ & I_\ell & \\ & & 0 \end{bmatrix}\widehat{U}^* = \widehat{Z}P^*\begin{bmatrix} 0 & & \\ & |\Delta_\ell|^{1/2} & \\ & & 0 \end{bmatrix}\widehat{U}^* \quad . \end{aligned}$$

Odavde odmah slijedi da je

$$W(W^*JW)^+W^*JW = W \quad ,$$

ako i samo ako ne dolazi do gubitka ranga matrice W^*JW, obzirom na rang W. ■

Teorem 3.4.1. o dekompoziciji prostora induciranoj blok reflektorom vrijedi i u indefinitnom slučaju, ali uz dodatni uvjet (3.4.13).

Teorem 3.4.5.

Neka je matrica H definirana relacijom (3.4.4). *Ako je*

$$\mathrm{rang}(W^*JW) = \mathrm{rang}(W) \quad ,$$

onda za sve $w \in \mathcal{R}(W)$ *vrijedi*

$$H(W)w = -w \quad ,$$

a za sve $y \in \mathcal{R}(W)^{[\perp]}$

$$H(W)y = y \quad .$$

Dokaz:

Dokažimo prvu tvrdnju. Neka je $w \in \mathcal{R}(W)$. Tada postoji x takav da vrijedi $w = Wx$. Korištenjem teorema 3.4.4. i definicije matrice H, imamo

$$\begin{aligned} Hw &= w - 2W(W^*JW)^+W^*Jw = Wx - 2W(W^*JW)^+W^*JWx \\ &= Wx - 2Wx = -w \quad . \end{aligned}$$

Da bismo dokazali drugi dio tvrdnje, potrebno je dokazati da vrijedi

$$\mathcal{R}(W)^{[\perp]} = \mathcal{N}(W^{[*]}) = \mathcal{N}(W^*J) \quad .$$

Dokažimo prvo da je $\mathcal{R}(W)^{[\perp]} \subseteq \mathcal{N}(W^{[*]})$. Neka je $y \in \mathcal{R}(W)^{[\perp]}$. Tada za svaki $w \in \mathcal{R}(W)$ vrijedi $[w, y] = 0$.

Po definiciji slike operatora, za svaki $w \in \mathcal{R}(W)$ postoji $x \in \mathbf{C}^p$, tako da je $w = Wx$. Također, za svaki $x \in \mathbf{C}^p$ je $Wx \in \mathcal{R}(W)$. Stoga za svako $x \in \mathbf{C}^p$ vrijedi

$$[Wx, y] = 0 \iff [x, W^{[*]}y] = 0 \quad .$$

Zbog svojstva nedegeneriranosti iz definicije indefinitnog skalarnog produkta, odmah slijedi $W^{[*]}y = 0$, tj. $y \in \mathcal{N}(W^{[*]})$.

Obratno, neka je $y \in \mathcal{N}(W^{[*]})$. Tada je $W^{[*]}y = 0$, pa je i

$$[x, W^{[*]}y] = 0 \iff [Wx, y] = 0 \quad ,$$

za svaki $x \in \mathbf{C}^p$. Dakle, y je okomit na $\mathcal{R}(W)$ obzirom na indefinitni skalarni produkt, pa je $y \in \mathcal{R}(W)^{[\perp]}$, čime je dokazano $\mathcal{N}(W^{[*]}) \subseteq \mathcal{R}(W)^{[\perp]}$.

Drugu jednakost dobivamo kao posljedicu relacije (2.1.4). Za $y \in \mathcal{N}(W^{[*]})$ vrijedi $W^{[*]}y = 0$, pa korištenjem (2.1.4) izlazi $JW^*Jy = 0$. Zbog regularnosti matrice J, prethodna jednakost ekvivalentna je s $W^*Jy = 0$. Odatle odmah zaključujemo da je $y \in \mathcal{N}(W^*J)$, ako i samo ako je $y \in \mathcal{N}(W^{[*]})$.

Zbog toga, za $y \in \mathcal{R}(W)^{[\perp]}$ vrijedi

$$Hy = y - 2W(W^*JW)^+W^*Jy = y \quad .$$

■

4. Indefinitna QR dekompozicija

Svaka pozitivno definitna hermitska matrica A ima dekompoziciju Choleskog

$$A = R^*R \quad ,$$

gdje je R gornjetrokutasta matrica. Ako se zahtijeva da R ima pozitivne dijagonalne elemente, takva je dekompozicija jedinstvena.

Dekompoziciju Choleskog matrice A možemo promatrati i na drugi način. Neka je G bilo koja pravokutna matrica za koju vrijedi

$$A = G^*G \quad .$$

Onda QR dekompozicija matrice G, $G = QR$, pri čemu je Q unitarna, daje faktor R iz dekompozicije Choleskog.

Ako je matrica A zadana faktorom G, očito, nije potrebno eksplicitno množiti faktore da bi se izračunala dekompozicija Choleskog, tj. matrica R.

4.1. Uvod

Za hermitsku indefinitnu matricu A moguće je napraviti analogon dekompozicije Choleskog — hermitsku indefinitnu dekompoziciju. Prvi ozbiljan problem je izračunavanje hermitske indefinitne dekompozicije, jer ona ne mora postojati. Na primjer, za matricu

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

ne postoji gornjetrokutasta matrica R i matrica $J = \text{diag}(\pm 1)$, tako da je

$$A = R^*JR \quad .$$

Početkom sedamdestih godina grupa autora [11], [7], [8], [10], [9], bavila se ovom dekompozicijom s ciljem rješavanja hermitskih indefinitnih linearnih sistema. Zbog sprečavanja pivotnog rasta, predložili su nekoliko varijanti potpunog i parcijalnog pivotiranja.

Pokazano je da se, uz razumnu pivotnu strategiju, može postići da matrica R bude blok–gornjetrokutasta s dijagonalnim blokovima dimenzija 1×1 ili 2×2. Naime, vrijedi slijedeća propozicija.

Propozicija 4.1.1.

Neka je A hermitska matrica. Ako su svi dijagonalni elementi matrice A jednaki 0 i sve 2×2 glavne minore jednake 0, tada je matrica A nul–matrica.

Dokaz:

Pretpostavimo da su svi dijagonalni elementi matrice A jednaki 0. U tom slučaju za sve glavne minore vrijedi

$$\det \begin{bmatrix} 0 & a_{ij} \\ \bar{a}_{ij} & 0 \end{bmatrix} = -|a_{ij}|^2 = 0 \quad ,$$

što povlači da su svi elementi $a_{ij} = 0$. ■

Dakle, dekompoziciju možemo započeti ako matrica A nije 0–matrica. Tada se sigurno može naći, ili barem jedan dijagonalni element koji je različit od 0, ili barem jedna 2×2 glavna podmatrica različita od 0. Istovremenim zamjenama redaka i stupaca, taj blok uvijek možemo dovesti u gornji lijevi kut matrice. Dakle, dovoljno je promatrati samo 1×1 ili 2×2 pivotne podmatrice. Ako je matrica A regularna, u svakom koraku dekompozicije možemo pronaći regularnu pivotnu podmatricu. Na taj način dobivamo hermitsku indefinitnu dekompoziciju matrice A oblika

$$A = P^* R^* J R P \quad ,$$

gdje je P matrica permutacija, R blok gornjetrokutasta matrica s dijagonalnim blokovima veličine 1×1 ili 2×2, a J dijagonalna matrica s ± 1 na dijagonali. Jasno je da taj oblik dekompozicije nije jedinstven. Osim toga, različite pivotne strategije mogu dovesti i do različitog broja 1×1 i 2×2 dijagonalnih blokova.

Ako je matrica A zadana faktorom G i matricom J, tako da je $A = G^* J G$, pitanje je, da li je moguće, bez eksplicitnog množenja matrica, izračunati faktor R u hermitskoj indefinitnoj dekompoziciji. Intuitivno je jasno, da bi na G trebalo slijeva djelovati J–unitarnom matricom i poništavati elemente u donjem trokutu matrice G.

Indefinitni analogon QR dekompozicije definiramo na slijedeći način.

Definicija 4.1.1.

Neka je $G \in \mathbf{C}^{m \times n}$, $m \geq n$ i neka je i $J \in \mathbf{C}^{m \times m}$, $J = \operatorname{diag}(j_{11}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$. Rastav matrice G oblika

$$G = P_1 Q \begin{bmatrix} R \\ 0 \end{bmatrix} P_2^* \quad , \quad Q^* J_1 Q = J_1 \quad , \quad J_1 = P_1^* J P_1 \quad , \qquad (4.1.1)$$

gdje su matrice P_1 i P_2 matrice permutacije, matrica Q je J_1–unitarna, a R je blok gornjetrokutasta s dijagonalnim blokovima dimenzija 1×1 i 2×2, zovemo indefinitna QR dekompozicija matrice G obzirom na zadani J. ■

Opravdanje ove definicije dano je u teoremu 4.2.1., gdje pokazujemo egzistenciju ove dekompozicije uz uvjet regularnosti matrice $A = G^*JG$. Napomenimo da u općem slučaju nije moguće postići da Q bude J–unitarna, obzirom na zadani J, već je nužno uvesti matricu permutacija P_1.

Slično kao i kod obične QR dekompozicije, indefinitnu QR dekompoziciju možemo realizirati uz pomoć (indefinitnih) rotacija, ili uz pomoć (indefinitnih) Householderovih transformacija. Obje konstrukcije razvijamo u slijedećim odjeljcima.

4.2. Izvod i realizacija dekompozicije rotacijama

U ovom odjeljku konstruktivno dokazujemo egzistenciju indefinitne QR dekompozicije, korištenjem rotacija, J–rotacija i elementarnih J–unitarnih blok matrica.

Za konstrukciju indefinitne QR dekompozicije u općem slučaju, prvo izvodimo specijalne slučajeve QR dekompozicije 4×2 i 3×2 matrica G, uz određene dodatne pretpostavke na G i J. Ti posebni slučajevi bit će elementarni koraci opće indefinitne QR dekompozicije.

Definicija 4.2.1.

Neka su $G_{42} \in \mathbf{C}^{4\times2}$ i $J_4 \in \mathbf{C}^{4\times4}$ matrice oblika

$$G_{42} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ 0 & g_{32} \\ 0 & g_{42} \end{bmatrix} \quad , \quad J_4 = \begin{bmatrix} j_{11} & & & \\ & -j_{11} & & \\ & & j_{33} & \\ & & & -j_{33} \end{bmatrix} \quad .$$

Kažemo da je G_{42}, J_4 elementarni par reda 4, ako vrijedi

$$g_{11}, g_{21}, g_{32}, g_{42} \neq 0 \quad , \quad j_{11}, j_{33} \in \{-1, 1\} \quad ,$$

barem jedan od elemenata g_{12}, g_{22} je različit od nule i matrica $A_2 = G_{42}^* J_4 G_{42}$ je indefinitna i regularna. ■

Za daljnju konstrukciju, korisno je promatrati G_{42} kao blok matricu oblika

$$G_{42} = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

pri čemu je

$$R_1 = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad , \quad R_2 = \begin{bmatrix} 0 & g_{32} \\ 0 & g_{42} \end{bmatrix} \quad .$$

Matrica A_2 iz definicije 4.2.1. ima oblik

$$\begin{aligned} A_2 &= G_{42}^* J_4 G_{42} \\ &= \begin{bmatrix} \bar{g}_{11} & \bar{g}_{21} & 0 & 0 \\ \bar{g}_{12} & \bar{g}_{22} & \bar{g}_{32} & \bar{g}_{42} \end{bmatrix} \begin{bmatrix} j_{11} & & & \\ & -j_{11} & & \\ & & j_{33} & \\ & & & -j_{33} \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ 0 & g_{32} \\ 0 & g_{42} \end{bmatrix} \\ &= \begin{bmatrix} j_{11}(|g_{11}|^2 - |g_{21}|^2) & j_{11}(\bar{g}_{11}g_{12} - \bar{g}_{21}g_{22}) \\ j_{11}(g_{11}\bar{g}_{12} - g_{21}\bar{g}_{22}) & j_{11}(|g_{12}|^2 - |g_{22}|^2) + j_{33}(|g_{32}|^2 - |g_{42}|^2) \end{bmatrix} \end{aligned} \tag{4.2.1}$$

Njena determinanta je

$$\begin{aligned} \det A_2 &= (|g_{11}|^2 - |g_{21}|^2)(|g_{12}|^2 - |g_{22}|^2) + j_{11}j_{33}(|g_{11}|^2 - |g_{21}|^2)(|g_{32}|^2 - |g_{42}|^2) \\ &\qquad - (\bar{g}_{11}g_{12} - \bar{g}_{21}g_{22})(g_{11}\bar{g}_{12} - g_{21}\bar{g}_{22}) \\ &= j_{11}j_{33}(|g_{11}|^2 - |g_{21}|^2)(|g_{32}|^2 - |g_{42}|^2) \\ &\qquad - |g_{12}|^2|g_{21}|^2 - |g_{11}|^2|g_{22}|^2 + g_{11}\bar{g}_{12}\bar{g}_{21}g_{22} + \bar{g}_{11}g_{12}g_{21}\bar{g}_{22} \\ &= -j_{11}j_{33}(|g_{21}|^2 - |g_{11}|^2)(|g_{32}|^2 - |g_{42}|^2) - |g_{11}g_{22} - g_{12}g_{21}|^2 \quad . \end{aligned}$$

Posljednju jednakost možemo napisati u obliku

$$\det A_2 = -j_{11}j_{33}(|g_{21}|^2 - |g_{11}|^2)(|g_{32}|^2 - |g_{42}|^2) - |\det R_1|^2 \quad . \tag{4.2.2}$$

Potpuno analogno definiramo elementarni par reda 3.

Definicija 4.2.2.

Neka su $G_{32} \in \mathbf{C}^{3\times 2}$ i $J_3 \in \mathbf{C}^{3\times 3}$ matrice oblika

$$G_{32} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ 0 & g_{32} \end{bmatrix} \quad , \quad J_3 = \begin{bmatrix} j_{11} & & \\ & -j_{11} & \\ & & j_{33} \end{bmatrix} \quad .$$

Kažemo da je G_{32}, J_3 elementarni par reda 3, ako je

$$g_{11}, g_{21}, g_{32} \neq 0 \quad , \quad j_{11}, j_{33} \in \{-1, 1\} \quad ,$$

barem jedan od elemenata g_{12}, g_{22} je različit od nule i matrica $A_2 = G_{32}^* J_3 G_{32}$ je indefinitna i regularna. ■

Uočimo da brisanjem posljednjeg retka u G_4, te posljednjeg retka i stupca u J_4, iz elementarnog para reda 4 dobivamo elementarni par reda 3. U tom smislu, G_{32} promatramo kao blok matricu oblika

$$G_{32} = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

pri čemu je

$$R_1 = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad , \quad R_2 = \begin{bmatrix} 0 & g_{32} \end{bmatrix} \quad .$$

Pripadna matrica A_2 iz definicije 4.2.2. ima isti oblik kao ona iz (4.2.1), ako uzmemo da je $g_{42} = 0$. Analogon relacije (4.2.2) je

$$\det A_2 = -j_{11}j_{33}(|g_{21}|^2 - |g_{11}|^2)|g_{32}|^2 - |\det R_1|^2 \quad . \tag{4.2.3}$$

Pretpostavku regularnosti i indefinitnosti matrica A_2 iz definicija 4.2.1. i 4.2.2. opravdava slijedeća lema.

Lema 4.2.1.

Neka su G_{42} i J_4 matrice oblika kao u definiciji 4.2.1. i neka je $\det R_1 \neq 0$. Definiramo

$$z = \frac{1}{\det R_1} = \frac{1}{g_{11}g_{22} - g_{12}g_{21}} \quad , \tag{4.2.4}$$

i

$$r = \frac{1}{\sqrt{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)(|g_{21}|^2 - |g_{11}|^2) + 1}} \quad . \tag{4.2.5}$$

Matrica $A_2 = G_{42}^* J_4 G_{42}$ je regularna i indefinitna, ako i samo ako je r realan broj.

Dokaz:

Matrica A_2 je regularna i indefinitna, ako i samo ako je $\det A_2 < 0$, ili, prema (4.2.2), ako i samo ako je

$$-j_{11}j_{33}(|g_{21}|^2 - |g_{11}|^2)(|g_{32}|^2 - |g_{42}|^2) < \frac{1}{|z|^2} \quad ,$$

odnosno

$$j_{11}j_{33}|z|^2(|g_{21}|^2 - |g_{11}|^2)(|g_{32}|^2 - |g_{42}|^2) > -1 \quad .$$

Dakle, izraz pod korijenom u r je veći od 0, ako i samo ako je A_2 regularna i indefinitna. ■

Napomena 4.2.1.

Analogna tvrdnja vrijedi i za matrice G_{32} i J_3 iz definicije 4.2.2., s tim da se r definira relacijom

$$r = \frac{1}{\sqrt{j_{11}j_{33}|z|^2|g_{32}|^2(|g_{21}|^2 - |g_{11}|^2) + 1}} \quad . \tag{4.2.6}$$

■

Realnost skalara r iz relacija (4.2.5), odnosno (4.2.6) ključna je za konstrukciju blok J–unitarnih transformacija. Prethodna lema pokazuje da za elementarne parove dobivamo realne r, ali uz dodatnu pretpostavku da je $\det R_1 \neq 0$.

Slijedeća lema pokazuje da se J–unitarnim transformacijama elementarnog para može osigurati $\det R_1 \neq 0$.

Lema 4.2.2.

Neka je G_{32}, J_3 elementarni par reda 3. Tada postoje matrice permutacije P_1 i P_2 i J_3'–unitarna matrica U, takve da je

$$G_{32}' = U^{-1} P_1^* G_{32} P_2 \quad , \quad U^* J_3' U = J_3' \quad , \quad J_3' = P_1^* J_3 P_1 \quad ,$$

i vrijedi

$$g_{11}' g_{22}' - g_{12}' g_{21}' \neq 0 \quad .$$

Drugim riječima, permutacijama stupaca matrice G_{32} i primjenom J_3'–unitarnih transformacija slijeva na G_{32} možemo postići da je gornji 2×2 blok u G_{32}' regularan.

Dokaz:

Ako je $\det R_1 \neq 0$, možemo uzeti da su U, P_1 i P_2 jedinične matrice, tj. $G_{32}' = G_{32}$ i tvrdnja trivijalno vrijedi.

Pretpostavimo stoga da je $\det R_1 = 0$. Relacija (4.2.2) i regularnost matrice A_2 povlače

$$\det A_2 = -j_{11} j_{33} (|g_{21}|^2 - |g_{11}|^2) |g_{32}|^2 \neq 0 \quad ,$$

tj. za elementarni par reda 3, ako je $\det R_1 = 0$, onda vrijedi $|g_{11}| \neq |g_{21}|$.

Zamjenom poretka stupaca dobivamo

$$\widetilde{G}_{32} = \begin{bmatrix} g_{12} & g_{11} \\ g_{22} & g_{21} \\ g_{32} & 0 \end{bmatrix} \quad .$$

Dovedimo ta dva stupca u standardnu formu, tako da je element u trećem retku prvog stupca jednak 0,

$$G_{32}' = \begin{bmatrix} g_{12}' & g_{11}' \\ g_{22}' & g_{21}' \\ 0 & g_{32}' \end{bmatrix} \quad .$$

Korištenjem jedne trigonometrijske rotacije, poništimo element g_{32}. Bez smanjenja općenitosti, možemo pretpostaviti da je $j_{11} = j_{33}$, a svi elementi u prvom stupcu realni (ako nisu, to uvijek možemo postići). Tada vrijedi

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} g_{12} \\ g_{32} \end{bmatrix} = \begin{bmatrix} g_{12}' \\ 0 \end{bmatrix} \quad .$$

Jednostavnim računom slijedi

$$s = \frac{-\operatorname{sign}(g_{12})g_{32}}{\sqrt{g_{12}^2+g_{32}^2}} \quad , \quad c = \frac{\operatorname{sign}(g_{12})g_{12}}{\sqrt{g_{12}^2+g_{32}^2}} \quad ,$$

pa odatle i

$$g'_{11} = \frac{\operatorname{sign}(g_{12})g_{11}g_{12}}{\sqrt{g_{12}^2+g_{32}^2}} \quad , \quad g'_{12} = \operatorname{sign}(g_{12})\sqrt{g_{12}^2+g_{32}^2} \quad .$$

Ponovno izračunajmo determinantu gornjeg 2×2 bloka.

$$\det R'_1 = g'_{12}g'_{21} - g'_{11}g'_{22} = \operatorname{sign}(g_{12})\left(g_{21}\sqrt{g_{12}^2+g_{32}^2} - g_{11}g_{12}g_{22}\frac{1}{\sqrt{g_{12}^2+g_{32}^2}}\right) \quad .$$

Ako je $\det R'_1 \neq 0$, završili smo postupak. U protivnom, ako je $\det R'_1 = 0$, onda vrijedi

$$g_{21}(g_{21}^2+g_{32}^2) = g_{11}g_{12}g_{22} \quad .$$

Grupiranjem pribrojnika dobivamo

$$0 = g_{12}(g_{12}g_{21} - g_{11}g_{22}) = -g_{21}g_{32}^2 \quad ,$$

što je moguće samo ako je $g_{21} = 0$, ili $g_{32} = 0$. Tvrdimo da u oba slučaja nismo imali elementarni par reda 3. Ako je $g_{21} = 0$, onda, zbog $\det R_1 = 0$ izlazi da je $g_{22} = 0$, pa nismo imali elementarni par reda 3. Ako je $g_{32} = 0$, po definiciji, nismo imali elementarni par reda 3. ■

Na sličan način dokazujemo lemu za elementarni par reda 4.

Lema 4.2.3.

Neka je G_{42}, J_4 elementarni par reda 4. Tada postoje matrice permutacije P_1 i P_2 i J'_4–unitarna matrica U, takve da je

$$G'_{42} = U^{-1}P_1^*G_{42}P_2 \quad , \quad U^*J'_4U = J'_4 \quad , \quad J'_4 = P_1^*J_4P_1 \quad ,$$

i vrijedi

$$g'_{11}g'_{22} - g'_{12}g'_{21} \neq 0 \quad .$$

Drugim riječima, permutacijama stupaca matrice G_{42} i primjenom J'_4–unitarnih transformacija slijeva na G_{42} možemo postići da je gornji 2×2 blok u G'_{42} regularan.

Dokaz:

Ako je $\det R_1 \neq 0$, možemo uzeti da su U, P_1 i P_2 jedinične matrice, tj. $G'_{42} = G_{42}$ i tvrdnja trivijalno vrijedi.

Pretpostavimo stoga da je $\det R_1 = 0$. Relacija (4.2.2) i regularnost matrice A_2 povlače

$$\det A_2 = -j_{11}j_{33}(|g_{21}|^2 - |g_{11}|^2)(|g_{32}|^2 - |g_{42}|^2) \neq 0 \quad ,$$

tj. tada vrijedi $|g_{11}| \neq |g_{21}|$ i $|g_{32}| \neq |g_{42}|$.

U slučaju $\det R_1 = 0$, zamijenimo poredak stupaca. Sada imamo

$$\widetilde{G}_{42} = \begin{bmatrix} g_{12} & g_{11} \\ g_{22} & g_{21} \\ g_{32} & 0 \\ g_{42} & 0 \end{bmatrix} \quad .$$

Dovedimo ta dva stupca u standardnu formu, tako da su elementi u trećem i četvrtom retku prvog stupca jednaki 0. Slično kao u dokazu leme 4.2.2., to uvijek možemo napraviti s dvije trigonometrijske rotacije

$$G'_{42} = \begin{bmatrix} g'_{12} & g'_{11} \\ g'_{22} & g'_{21} \\ 0 & g'_{32} \\ 0 & g'_{42} \end{bmatrix} \quad .$$

Bez smanjenja općenitosti, možemo pretpostaviti da je $j_{11} = j_{33}$, a svi elementi u prvom stupcu realni (ako nisu to uvijek možemo postići). Nakon prve rotacije dobivamo

$$g'_{11} = \frac{\operatorname{sign}(g_{12})g_{11}g_{12}}{\sqrt{g_{12}^2 + g_{32}^2}} \quad , \quad g'_{12} = \operatorname{sign}(g_{12})\sqrt{g_{12}^2 + g_{32}^2} \quad .$$

Na isti način dobivamo i relacije za preostale elemente:

$$g'_{21} = \frac{\operatorname{sign}(g_{22})g_{21}g_{22}}{\sqrt{g_{22}^2 + g_{42}^2}} \quad , \quad g'_{22} = \operatorname{sign}(g_{22})\sqrt{g_{22}^2 + g_{42}^2} \quad .$$

Determinanta bloka R'_1 u matrici G'_{42} je

$$\det R'_1 = g'_{12}g'_{21} - g'_{11}g'_{22} = \operatorname{sign}(g_{12}g_{22})\left(g_{21}g_{22}\frac{\sqrt{g_{12}^2 + g_{32}^2}}{\sqrt{g_{22}^2 + g_{42}^2}} - g_{11}g_{12}\frac{\sqrt{g_{22}^2 + g_{42}^2}}{\sqrt{g_{12}^2 + g_{32}^2}}\right) \quad .$$

Ako je $\det R'_1 \neq 0$, onda za matricu G'_{42} vrijedi tvrdnja, jer je dobivena iz G_{42} primjenom J–unitarnih transformacija i primjenom permutacija redaka i stupaca.

Pretpostavimo stoga da je i $\det R'_1 = 0$. Tada je

$$g_{21}g_{22}(g_{12}^2 + g_{32}^2) = g_{11}g_{12}(g_{22}^2 + g_{42}^2) \quad ,$$

odnosno, nakon grupiranja pribrojnika

$$g_{12}g_{22}(g_{12}g_{21} - g_{11}g_{22}) = g_{11}g_{12}g_{42}^2 - g_{21}g_{22}g_{32}^2 \quad ,$$

odakle, zbog $\det R_1 = 0$, slijedi

$$g_{11}g_{12}g_{42}^2 - g_{21}g_{22}g_{32}^2 = 0 \quad . \tag{4.2.7}$$

Uvažimo li odnose među g–ovima, tj. ako je $\det R_1 = 0$, onda je $g_{11} \neq g_{21}$ i vrijedi

$$R_1 = \begin{bmatrix} g_{11} & \gamma g_{11} \\ g_{21} & \gamma g_{21} \end{bmatrix}$$

za neki $\gamma \in \mathbf{C}$, tj. $g_{12} = \gamma g_{11}$ i $g_{22} = \gamma g_{21}$. Tada relacija (4.2.7) glasi

$$\gamma g_{11}^2 g_{42}^2 = \gamma g_{21}^2 g_{32}^2 \quad .$$

Zbog toga što G_{42} i J_4 čine elementarni par reda 4, γ ne može biti nula, jer bi istovremeno i g_{12} i g_{22} bili 0, pa zaključujemo da je $\det R_1' = 0$, ako je

$$\left|\frac{g_{11}}{g_{21}}\right| = \left|\frac{g_{32}}{g_{42}}\right| \quad .$$

U tom slučaju nismo postigli ništa samo zamjenom stupaca. Iskoristimo još činjenicu da je, zbog $\det R_1 = 0$, $|g_{32}| \neq |g_{42}|$. Upotrijebimo hiperbolnu transformaciju koja prebacuje J–normu na jedan od ta dva elementa (bez smanjenja općenitosti, možemo pretpostaviti da je to g_{32}). Zbog jednostavnosti pisanja, nazovimo taj element ponovno g_{32}.

Sada se problem svodi na elementarni par reda 3.

$$\widetilde{G}_{32} = \begin{bmatrix} g_{12} & g_{11} \\ g_{22} & g_{21} \\ g_{32} & 0 \end{bmatrix} \quad .$$

Korištenjem leme 4.2.2. izlazi tvrdnja. ■

U slijedeće četiri leme konstruiramo blok J–unitarne matrice koje poništavaju 4×2 i 3×2 blokove u matrici G.

Lema 4.2.4.

Neka je G_{42}, J_4 elementarni par reda 4 za koji vrijedi $g_{11}g_{22} - g_{12}g_{21} \neq 0$. Tada postoji blok–$J_4$ unitarna matrica U takva da je

$$U^{-1}G_{42} = G_{42}' = \begin{bmatrix} g_{11}' & g_{12}' \\ g_{21}' & g_{22}' \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad . \tag{4.2.8}$$

Matrica U ima oblik (3.2.11)

$$U = \begin{bmatrix} (I - YX)^{1/2} & Y \\ -X & (I - XY)^{1/2} \end{bmatrix} \quad ,$$

pri čemu je:

(i) *Ako je* $|g_{11}| \neq |g_{21}|$ *i* $|g_{32}| \neq |g_{42}|$

$$\begin{aligned} X &= rz \begin{bmatrix} g_{21}g_{32} & -g_{11}g_{32} \\ g_{21}g_{42} & -g_{11}g_{42} \end{bmatrix} \quad , \\ Y &= rj_{11}j_{33}\bar{z} \begin{bmatrix} \bar{g}_{21}\bar{g}_{32} & -\bar{g}_{21}\bar{g}_{42} \\ \bar{g}_{11}\bar{g}_{32} & -\bar{g}_{11}\bar{g}_{42} \end{bmatrix} \quad , \\ (I - XY)^{1/2} &= I - \frac{r-1}{|g_{32}|^2 - |g_{42}|^2} \begin{bmatrix} -|g_{32}|^2 & g_{32}\bar{g}_{42} \\ -\bar{g}_{32}g_{42} & |g_{42}|^2 \end{bmatrix} \quad , \\ (I - YX)^{1/2} &= I - \frac{r-1}{|g_{21}|^2 - |g_{11}|^2} \begin{bmatrix} -|g_{21}|^2 & g_{11}\bar{g}_{21} \\ -\bar{g}_{11}g_{21} & |g_{11}|^2 \end{bmatrix} \quad . \end{aligned}$$

(ii) *Ako je* $|g_{11}| = |g_{21}|$ *ili* $|g_{32}| = |g_{42}|$, *onda je*

$$\begin{aligned} X &= z \begin{bmatrix} g_{21}g_{32} & -g_{11}g_{32} \\ g_{21}g_{42} & -g_{11}g_{42} \end{bmatrix} \quad , \\ Y &= j_{11}j_{33}\bar{z} \begin{bmatrix} \bar{g}_{21}\bar{g}_{32} & -\bar{g}_{21}\bar{g}_{42} \\ \bar{g}_{11}\bar{g}_{32} & -\bar{g}_{11}\bar{g}_{42} \end{bmatrix} \quad , \\ (I - XY)^{1/2} &= I + \frac{j_{11}j_{33}|z|^2(|g_{21}|^2 - |g_{11}|^2)}{2} \begin{bmatrix} -|g_{32}|^2 & g_{32}\bar{g}_{42} \\ -\bar{g}_{32}g_{42} & |g_{42}|^2 \end{bmatrix} \quad , \\ (I - YX)^{1/2} &= I + \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)}{2} \begin{bmatrix} -|g_{21}|^2 & g_{11}\bar{g}_{21} \\ -\bar{g}_{11}g_{21} & |g_{11}|^2 \end{bmatrix} \quad , \end{aligned}$$

pri čemu je z definiran relacijom (4.2.4), *a r relacijom* (4.2.5).

Dokaz:

Iskoristimo blok J–unitarne matrice definirane teoremom 3.2.1. (ii), čiji je inverz U^{-1} definiran relacijom (3.2.15). Takve matrice su reda 4 s unutarnjim blokovima reda 2.

Uz pomoć tih matrica želimo poništiti elemente g_{32} i g_{42}. Matričnu relaciju za poništavanje blokova možemo zapisati kao

$$U^{-1} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} R_1' \\ 0 \end{bmatrix} \quad .$$

Ako iz te matrične jednadžbe ispišemo drugi blok–redak, korištenjem relacije (3.2.14), dobivamo

$$XR_1 + (I - XJ_1X^*J_2)^{1/2}R_2 = 0 \quad .$$

Množenjem slijeva s $(I - XJ_1X^*J_2)^{-1/2}$ i korištenjem definicije (3.2.16) za matrični tangens T_1, imamo

$$T_1 = (I - XJ_1X^*J_2)^{-1/2}X = -R_2R_1^{-1} \quad .$$

Zbog jednostavnosti pisanja, izostavimo indeks za T_1, tj. neka je $T = T_1$. Ako iskoristimo vezu (3.2.17) za matrice T i X, korištenjem propozicije 3.2.1. (ii), dobivamo

$$X = T(I + J_1T^*J_2T)^{-1/2} = (I + TJ_1T^*J_2)^{-1/2}T \quad , \tag{4.2.9}$$

uz pretpostavku da $(I + TJ_1T^*J_2)$ nema realnih svojstvenih vrijednosti manjih ili jednakih 0.

Raspišimo elemente matričnog tangensa.

$$T = -R_2R_1^{-1} = -\frac{1}{g_{11}g_{22} - g_{12}g_{21}} \begin{bmatrix} 0 & g_{32} \\ 0 & g_{42} \end{bmatrix} \begin{bmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{bmatrix} = z \begin{bmatrix} g_{21}g_{32} & -g_{11}g_{32} \\ g_{21}g_{42} & -g_{11}g_{42} \end{bmatrix} \quad ,$$

uz z definiran relacijom (4.2.4). Ovdje smo iskoristili pretpostavku o regularnosti gornjeg 2×2 bloka.

Za realizaciju algoritma potrebno je izračunati veličine X, Y, te $(I - XY)^{1/2}$ i $(I - YX)^{1/2}$. Blokovi J_1 i J_2 definirani relacijom (3.2.13) su

$$J_1 = \mathrm{diag}(j_{11}, -j_{11}) \quad , \quad J_2 = \mathrm{diag}(j_{33}, -j_{33}) \quad .$$

Da bismo pokazali egzistenciju algoritma, dovoljno je pokazati da su matrične potencije u relaciji (4.2.9) izračunljive. Uvrštavanjem dobivamo

$$\begin{aligned} J_1T^*J_2T &= \begin{bmatrix} j_{11} & \\ & -j_{11} \end{bmatrix} \bar{z} \begin{bmatrix} \bar{g}_{21}\bar{g}_{32} & \bar{g}_{21}\bar{g}_{42} \\ -\bar{g}_{11}\bar{g}_{32} & -\bar{g}_{11}\bar{g}_{42} \end{bmatrix} \begin{bmatrix} j_{33} & \\ & -j_{33} \end{bmatrix} z \begin{bmatrix} g_{21}g_{32} & -g_{11}g_{32} \\ g_{21}g_{42} & -g_{11}g_{42} \end{bmatrix} \\ &= |z|^2 j_{11}j_{33}(|g_{32}|^2 - |g_{42}|^2) \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11}g_{21} & -|g_{11}|^2 \end{bmatrix} \quad . \end{aligned} \tag{4.2.10}$$

U slučaju $|g_{32}| = |g_{42}|$, vrijedi

$$I + J_1T^*J_2T = I \quad ,$$

pa je

$$X = T \quad .$$

Ako je $|g_{32}| \neq |g_{42}|$, treba izračunati $(I + J_1T^*J_2T)^{-1/2}$. Uočimo da je matrica u (4.2.10) ranga 1, pa se može napisati kao skalarni produkt dva vektora w i u.

Zato možemo pisati

$$(I + J_1T^*J_2T)^{-1/2} = (I + wu^*)^{-1/2} = \sum_{k=0}^{\infty} \binom{-1/2}{k} (wu^*)^k \quad .$$

Primijenimo li trik

$$(wu^*)^k = w(u^*w)(u^*w)(u^*\cdots w)u^* = (u^*w)^{k-1}wu^* \quad ,$$

prethodnu relaciju možemo pisati kao

$$(I + wu^*)^{-1/2} = I + \sum_{k=1}^{\infty} \binom{-1/2}{k} (u^*w)^{k-1}wu^* \quad . \tag{4.2.11}$$

Treba razlikovati dva slučaja: $u^*w = 0$ i $u^*w \neq 0$.

U slučaju $u^*w = 0$, u (4.2.11) ostaje samo član uz $k = 1$, pa imamo:

$$(I + wu^*)^{-1/2} = I + \binom{-1/2}{1} wu^* = I - \frac{1}{2}wu^* \quad .$$

Ako je $u^*w \neq 0$, dobivamo sumu skalara, pa vrijedi

$$(I + wu^*)^{-1/2} = I + \frac{1}{u^*w}\left((u^*w + 1)^{-1/2} - 1\right) wu^* \quad .$$

Sada treba izračunati koji to u i w u produktu

$$wu^* = \begin{bmatrix} w_1\bar{u}_1 & w_1\bar{u}_2 \\ w_2\bar{u}_1 & w_2\bar{u}_2 \end{bmatrix} \quad ,$$

daju matricu iz relacije (4.2.10). Uočimo da je u^*w je trag matrice wu^*. Ako iskoristimo (4.2.10), imamo

$$u^*w = j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)(|g_{21}|^2 - |g_{11}|^2) \quad .$$

U slučaju da je $|g_{11}| = |g_{21}|$, očito je $u^*w = 0$, pa je

$$(I + J_1T^*J_2T)^{-1/2} = I - \frac{1}{2}j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2) \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11}g_{21} & -|g_{11}|^2 \end{bmatrix} \quad .$$

Tada je

$$\begin{aligned} X &= T(I + J_1T^*J_2T)^{-1/2} \\ &= T - \tfrac{1}{2}j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)z \begin{bmatrix} g_{21}g_{32} & -g_{11}g_{32} \\ g_{21}g_{42} & -g_{11}g_{42} \end{bmatrix} \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11}g_{21} & -|g_{11}|^2 \end{bmatrix} \\ &= T - \tfrac{1}{2}j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)z \\ &\qquad \begin{bmatrix} (|g_{21}|^2 - |g_{11}|^2)g_{21}g_{32} & -(|g_{21}|^2 - |g_{11}|^2)g_{11}g_{32} \\ (|g_{21}|^2 - |g_{11}|^2)g_{21}g_{42} & -(|g_{21}|^2 - |g_{11}|^2)g_{11}g_{42} \end{bmatrix} \quad . \end{aligned}$$

Kako je druga matrica u posljednjoj relaciji nul–matrica, slijedi da je, kao i u slučaju $|g_{32}| = |g_{42}|$, $X = T$.

Ako je $|g_{11}| \neq |g_{21}|$ i $|g_{32}| \neq |g_{42}|$, imamo

$$
\begin{aligned}
(I + J_1T^*J_2T)^{-1/2} &= I + \frac{1}{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)(|g_{21}|^2 - |g_{11}|^2)} \\
&\quad \cdot \left(\frac{1}{\sqrt{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)(|g_{21}|^2 - |g_{11}|^2) + 1}} - 1 \right) \\
&\quad \cdot j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2) \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11}g_{21} & -|g_{11}|^2 \end{bmatrix} \quad .
\end{aligned}
$$

Uz pokratu za r definiranu relacijom (4.2.5), dobivamo

$$
\begin{aligned}
X &= T(I + J_1T^*J_2T)^{-1/2} \\
&= T + (r-1)\frac{z}{|g_{21}|^2 - |g_{11}|^2} \begin{bmatrix} g_{21}g_{32} & -g_{11}g_{32} \\ g_{21}g_{42} & -g_{11}g_{42} \end{bmatrix} \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11}g_{21} & -|g_{11}|^2 \end{bmatrix} \quad . \\
&= T + (r-1)T = rT
\end{aligned}
$$

Prema lemi 4.2.1., veličina r je realan broj kad god je matrica A_2 indefinitna i regularna.

Dakle, X možemo izračunati iz formule

$$X = rT \quad , \tag{4.2.12}$$

uz $r = 1$, kad je $|g_{11}| = |g_{21}|$ ili $|g_{32}| = |g_{42}|$.

Za Y vrijedi

$$Y = J_1X^*J_2 = rJ_1T^*J_2 = rj_{11}j_{33} \begin{bmatrix} \bar{t}_{11} & -\bar{t}_{21} \\ -\bar{t}_{12} & \bar{t}_{22} \end{bmatrix} = rj_{11}j_{33}\bar{z} \begin{bmatrix} \bar{g}_{21}\bar{g}_{32} & -\bar{g}_{21}\bar{g}_{42} \\ \bar{g}_{11}\bar{g}_{32} & -\bar{g}_{11}\bar{g}_{42} \end{bmatrix} \quad ,$$

gdje su t_{ij} elementi matrice T. Odatle slijedi

$$
\begin{aligned}
XY &= r^2j_{11}j_{33} \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} \bar{t}_{11} & -\bar{t}_{21} \\ -\bar{t}_{12} & \bar{t}_{22} \end{bmatrix} \\
&= r^2j_{11}j_{33} \begin{bmatrix} |t_{11}|^2 - |t_{12}|^2 & -t_{11}\bar{t}_{21} + t_{12}\bar{t}_{22} \\ \bar{t}_{11}t_{21} - \bar{t}_{12}t_{22} & -|t_{21}|^2 + |t_{22}|^2 \end{bmatrix} \quad .
\end{aligned}
$$

Tvrdimo da je ova posljednja matrica ranga 1. Po definiciji je

$$
\begin{aligned}
t_{11} &= zg_{21}g_{32} & t_{12} &= -zg_{11}g_{32} \\
t_{21} &= zg_{21}g_{42} & t_{22} &= -zg_{11}g_{42} \quad ,
\end{aligned}
$$

pa je

$$
\begin{aligned}
XY &= r^2j_{11}j_{33}|z|^2(|g_{21}|^2 - |g_{11}|^2) \begin{bmatrix} |g_{32}|^2 & -g_{32}\bar{g}_{42} \\ \bar{g}_{32}g_{42} & -|g_{42}|^2 \end{bmatrix} \\
&= r^2j_{11}j_{33}|z|^2(|g_{21}|^2 - |g_{11}|^2) \begin{bmatrix} g_{32} & -g_{32} \\ g_{42} & -g_{42} \end{bmatrix} \operatorname{diag}(\bar{g}_{32}, \bar{g}_{42}) \quad .
\end{aligned}
$$

Ponovno možemo primijeniti trik s vektorima

$$\begin{aligned} I - XY &= I + wu^* = I + rTJ_1rT^*J_2 \\ &= I + r^2 j_{11} j_{33} \begin{bmatrix} -|t_{11}|^2 + |t_{12}|^2 & t_{11}\bar{t}_{21} - t_{12}\bar{t}_{22} \\ -\bar{t}_{11}t_{21} + \bar{t}_{12}t_{22} & |t_{21}|^2 - |t_{22}|^2 \end{bmatrix} \quad . \end{aligned}$$

Odatle slijedi

$$\begin{aligned} u^*w &= r^2 j_{11} j_{33} (-|t_{11}|^2 + |t_{12}|^2 + |t_{21}|^2 - |t_{22}|^2) \\ &= r^2 j_{11} j_{33} |z|^2 (-|g_{21}|^2 |g_{32}|^2 + |g_{11}|^2 |g_{32}|^2 + |g_{21}|^2 |g_{42}|^2 - |g_{11}|^2 |g_{42}|^2) \\ &= r^2 j_{11} j_{33} |z|^2 (|g_{21}|^2 - |g_{11}|^2)(|g_{42}|^2 - |g_{32}|^2) \quad . \end{aligned}$$

Ako je $|g_{11}| = |g_{21}|$ ili $|g_{32}| = |g_{42}|$, onda je $u^*w = 0$ i

$$\begin{aligned} (I - XY)^{1/2} &= (I + wu^*)^{1/2} = I + \binom{1/2}{1} wu^* \\ &= I + \frac{1}{2} wu^* \\ &= I + \frac{r^2 j_{11} j_{33}}{2} \begin{bmatrix} -|t_{11}|^2 + |t_{12}|^2 & t_{11}\bar{t}_{21} - t_{12}\bar{t}_{22} \\ -\bar{t}_{11}t_{21} + \bar{t}_{12}t_{22} & |t_{21}|^2 - |t_{22}|^2 \end{bmatrix} \\ &= I + \frac{j_{11} j_{33} |z|^2 (|g_{21}|^2 - |g_{11}|^2)}{2} \begin{bmatrix} -|g_{32}|^2 & g_{32}\bar{g}_{42} \\ -\bar{g}_{32} g_{42} & |g_{42}|^2 \end{bmatrix} \quad . \end{aligned}$$

U protivnom, imamo

$$\begin{aligned} (I - XY)^{1/2} &= I + \frac{1}{u^*w} \left((u^*w + 1)^{1/2} - 1 \right) wu^* \\ &= I - \frac{r - 1}{|g_{32}|^2 - |g_{42}|^2} \begin{bmatrix} -|g_{32}|^2 & g_{32}\bar{g}_{42} \\ -\bar{g}_{32} g_{42} & |g_{42}|^2 \end{bmatrix} \quad . \end{aligned}$$

Na kraju treba još izračunati $(I - YX)^{1/2}$:

$$\begin{aligned} YX &= r^2 j_{11} j_{33} \begin{bmatrix} \bar{t}_{11} & -\bar{t}_{21} \\ -\bar{t}_{12} & \bar{t}_{22} \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \\ &= r^2 j_{11} j_{33} \begin{bmatrix} |t_{11}|^2 - |t_{21}|^2 & \bar{t}_{11}t_{12} - \bar{t}_{21}t_{22} \\ -t_{11}\bar{t}_{12} + t_{21}\bar{t}_{22} & -|t_{12}|^2 + |t_{22}|^2 \end{bmatrix} \quad . \end{aligned}$$

Tvrdimo da je posljednja matrica ranga 1. Vrijedi

$$\begin{aligned} YX &= r^2 j_{11} j_{33} |z|^2 (|g_{32}|^2 - |g_{42}|^2) \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11} g_{21} & -|g_{11}|^2 \end{bmatrix} \\ &= r^2 j_{11} j_{33} |z|^2 (|g_{32}|^2 - |g_{42}|^2) \begin{bmatrix} \bar{g}_{21} & -\bar{g}_{21} \\ \bar{g}_{11} & -\bar{g}_{11} \end{bmatrix} \operatorname{diag}(g_{21}, g_{11}) \quad . \end{aligned}$$

Tada je

$$I - YX = I + wu^* = I + r^2 j_{11} j_{33} \begin{bmatrix} -|t_{11}|^2 + |t_{21}|^2 & -\bar{t}_{11} t_{12} + t_{21}\bar{t}_{22} \\ t_{11}\bar{t}_{12} - t_{21}\bar{t}_{22} & |t_{12}|^2 - |t_{22}|^2 \end{bmatrix} \quad .$$

Odatle slijedi

$$\begin{aligned} u^*w &= r^2 j_{11} j_{33} (-|t_{11}|^2 + |t_{21}|^2 + |t_{12}|^2 - |t_{22}|^2) \\ &= r^2 j_{11} j_{33} |z|^2 (-|g_{21}|^2 |g_{32}|^2 + |g_{11}|^2 |g_{32}|^2 + |g_{21}|^2 |g_{42}|^2 - |g_{11}|^2 |g_{42}|^2) \\ &= r^2 j_{11} j_{33} |z|^2 (|g_{21}|^2 - |g_{11}|^2)(|g_{42}|^2 - |g_{32}|^2) \quad . \end{aligned}$$

Ako je $|g_{11}| = |g_{21}|$ ili $|g_{32}| = |g_{42}|$, onda je $u^*w = 0$ i

$$\begin{aligned} (I - YX)^{1/2} &= (I + wu^*)^{1/2} = I + \binom{1/2}{1} wu^* \\ &= I + \frac{1}{2} wu^* \\ &= I + \frac{j_{11} j_{33}}{2} \begin{bmatrix} -|t_{11}|^2 + |t_{21}|^2 & -\bar{t}_{11} t_{12} - \bar{t}_{21} t_{22} \\ t_{11}\bar{t}_{12} - t_{21}\bar{t}_{22} & |t_{12}|^2 - |t_{22}|^2 \end{bmatrix} \\ &= I + \frac{j_{11} j_{33} |z|^2 (|g_{32}|^2 - |g_{42}|^2)}{2} \begin{bmatrix} -|g_{21}|^2 & g_{11}\bar{g}_{21} \\ -\bar{g}_{11} g_{21} & |g_{11}|^2 \end{bmatrix} \quad . \end{aligned}$$

U protivnom, imamo

$$\begin{aligned} (I - YX)^{1/2} &= I + \frac{1}{u^*w} \left((u^*w + 1)^{1/2} - 1 \right) wu^* \\ &= I - \frac{r - 1}{|g_{21}|^2 - |g_{11}|^2} \begin{bmatrix} -|g_{21}|^2 & g_{11}\bar{g}_{21} \\ -\bar{g}_{11} g_{21} & |g_{11}|^2 \end{bmatrix} \quad . \end{aligned}$$

■

Slijedeća lema daje formule za transformaciju elemenata matrice G_{42}.

Lema 4.2.5.

Neka su matrice G_{42}, J_4, U i G'_{42}, te skalari z i r, definirani kao u prethodnoj lemi. Tada je

(*i*) *Ako je $|g_{11}| \neq |g_{21}|$ i $|g_{32}| \neq |g_{42}|$, onda je*

$$\begin{aligned} g'_{11} &= g_{11} & g'_{12} &= g_{12} + \frac{\bar{g}_{21}(r - 1)}{(|g_{21}|^2 - |g_{11}|^2) r z} \\ g'_{21} &= g_{21} & g'_{22} &= g_{22} + \frac{\bar{g}_{11}(r - 1)}{(|g_{21}|^2 - |g_{11}|^2) r z} \quad . \end{aligned}$$

(ii) Ako je $|g_{11}| = |g_{21}|$ *ili* $|g_{32}| = |g_{42}|$, *onda je*

$$\begin{aligned} g'_{11} &= g_{11} & g'_{12} &= g_{12} - \frac{j_{11}j_{33}\bar{z}\bar{g}_{21}(|g_{32}|^2 - |g_{42}|^2)}{2} \\ g'_{21} &= g_{21} & g'_{22} &= g_{12} - \frac{j_{11}j_{33}\bar{z}\bar{g}_{11}(|g_{32}|^2 - |g_{42}|^2)}{2} \end{aligned} \quad .$$

Dokaz:

Elementi matrice U definirani su prethodnom lemom. Dokaz provodimo množenjem matrica $U^{-1}G_{42}$ iz relacije (4.2.8).

Prije dokaza, izvedimo pomoćnu relaciju. Neka je

$$a = j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)(|g_{21}|^2 - |g_{11}|^2) \quad . \tag{4.2.13}$$

Tada je

$$r = \frac{1}{\sqrt{a+1}} \quad ,$$

i vrijedi

$$r - 1 + ra = \frac{1 - \sqrt{a+1} + a}{\sqrt{a+1}} = \frac{1-r}{r} \quad . \tag{4.2.14}$$

Ako je $|g_{11}| \neq |g_{21}|$ i $|g_{32}| \neq |g_{42}|$, za elemente matrice G'_{42} dobivamo

$$\begin{aligned} g'_{11} &= g_{11} + \frac{(r-1)|g_{21}|^2 g_{11}}{|g_{21}|^2 - |g_{11}|^2} - \frac{(r-1)|g_{21}|^2 g_{11}}{|g_{21}|^2 - |g_{11}|^2} = g_{11} \\ g'_{21} &= g_{21} + \frac{(r-1)|g_{11}|^2 g_{21}}{|g_{21}|^2 - |g_{11}|^2} - \frac{(r-1)|g_{11}|^2 g_{21}}{|g_{21}|^2 - |g_{11}|^2} = g_{21} \\ g'_{12} &= g_{12} + \frac{(r-1)|g_{21}|^2 g_{12}}{|g_{21}|^2 - |g_{11}|^2} - \frac{(r-1)g_{11}\bar{g}_{21}g_{22}}{|g_{21}|^2 - |g_{11}|^2} - rj_{11}j_{33}\bar{z}\bar{g}_{21}(|g_{32}|^2 - |g_{42}|^2) \\ &= g_{12} - \frac{(r-1)\bar{g}_{21}}{(|g_{21}|^2 - |g_{11}|^2)z} - rj_{11}j_{33}\bar{z}\bar{g}_{21}(|g_{32}|^2 - |g_{42}|^2) \\ &= g_{12} - \frac{\bar{g}_{21}}{(|g_{21}|^2 - |g_{11}|^2)z}(r - 1 + ra) = g_{12} + \frac{\bar{g}_{21}(r-1)}{(|g_{21}|^2 - |g_{11}|^2)zr} \\ g'_{22} &= g_{22} + \frac{(r-1)\bar{g}_{11}g_{21}g_{12}}{|g_{21}|^2 - |g_{11}|^2} - \frac{(r-1)|g_{11}|^2 g_{22}}{|g_{21}|^2 - |g_{11}|^2} - rj_{11}j_{33}\bar{z}\bar{g}_{11}(|g_{32}|^2 - |g_{42}|^2) \\ &= g_{22} - \frac{(r-1)\bar{g}_{11}}{(|g_{21}|^2 - |g_{11}|^2)z} - rj_{11}j_{33}\bar{z}\bar{g}_{11}(|g_{32}|^2 - |g_{42}|^2) \\ &= g_{22} - \frac{\bar{g}_{11}}{(|g_{21}|^2 - |g_{11}|^2)z}(r - 1 + ra) = g_{22} + \frac{\bar{g}_{11}(r-1)}{(|g_{21}|^2 - |g_{11}|^2)zr} \quad . \end{aligned}$$

Ako je $|g_{11}| = |g_{21}|$ ili $|g_{32}| \neq |g_{42}|$, za elemente matrice G'_{42} dobivamo

$$\begin{aligned}
g'_{11} &= g_{11} - \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)|g_{21}|^2 g_{11}}{2} + \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)|g_{21}|^2 g_{11}}{2} \\
&= g_{11} \\
g'_{21} &= g_{21} - \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)|g_{11}|^2 g_{21}}{2} + \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)|g_{11}|^2 g_{21}}{2} \\
&= g_{21} \\
g'_{12} &= g_{12} - \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)|g_{21}|^2 g_{12}}{2} + \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)g_{11}\bar{g}_{21}g_{22}}{2} \\
&\quad - j_{11}j_{33}\bar{z}\bar{g}_{21}(|g_{32}|^2 - |g_{42}|^2) \\
&= g_{12} + \frac{j_{11}j_{33}\bar{z}\bar{g}_{21}(|g_{32}|^2 - |g_{42}|^2)}{2} - rj_{11}j_{33}\bar{z}\bar{g}_{21}(|g_{32}|^2 - |g_{42}|^2) \\
&= g_{12} - \frac{j_{11}j_{33}\bar{z}\bar{g}_{21}(|g_{32}|^2 - |g_{42}|^2)}{2} \\
g'_{22} &= g_{22} - \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)\bar{g}_{11}g_{21}g_{12}}{2} + \frac{j_{11}j_{33}|z|^2(|g_{32}|^2 - |g_{42}|^2)|g_{11}|^2 g_{22}}{2} \\
&\quad - j_{11}j_{33}\bar{z}\bar{g}_{11}(|g_{32}|^2 - |g_{42}|^2) \\
&= g_{22} + \frac{j_{11}j_{33}\bar{z}\bar{g}_{11}(|g_{32}|^2 - |g_{42}|^2)}{2} - rj_{11}j_{33}\bar{z}\bar{g}_{11}(|g_{32}|^2 - |g_{42}|^2) \\
&= g_{22} - \frac{j_{11}j_{33}\bar{z}\bar{g}_{11}(|g_{32}|^2 - |g_{42}|^2)}{2} \quad .
\end{aligned}$$

■

Slijedeće dvije leme su ekvivalentne lemama 4.2.4. i 4.2.5. za elementarni par reda 3.

Lema 4.2.6.

Neka je G_{32}, J_3 elementarni par reda 3, za koji vrijedi $g_{11}g_{22} - g_{12}g_{21} \neq 0$. Tada postoji blok–$J_3$ unitarna matrica U takva da je

$$U^{-1}G_{32} = G'_{32} = \begin{bmatrix} g'_{11} & g'_{12} \\ g'_{21} & g'_{22} \\ 0 & 0 \end{bmatrix} \quad . \tag{4.2.15}$$

Matrica U ima oblik (3.2.11)

$$U = \begin{bmatrix} (I - YX)^{1/2} & Y \\ -X & (I - XY)^{1/2} \end{bmatrix} \quad ,$$

pri čemu je:

(*i*) *Ako je* $|g_{11}| \neq |g_{21}|$, *onda je*

$$\begin{aligned} X &= rzg_{32}\begin{bmatrix} g_{21} & -g_{11} \end{bmatrix} \quad , \\ Y &= rj_{11}j_{33}\bar{z}\bar{g}_{32}\begin{bmatrix} \bar{g}_{21} \\ \bar{g}_{11} \end{bmatrix} \quad , \\ (I-XY)^{1/2} &= r \quad , \\ (I-YX)^{1/2} &= I - \frac{r-1}{|g_{21}|^2-|g_{11}|^2}\begin{bmatrix} -|g_{21}|^2 & g_{11}\bar{g}_{21} \\ -\bar{g}_{11}g_{21} & |g_{11}|^2 \end{bmatrix} \quad . \end{aligned}$$

(*ii*) *Ako je* $|g_{11}| = |g_{21}|$, *onda je*

$$\begin{aligned} X &= zg_{32}\begin{bmatrix} g_{21} & -g_{11} \end{bmatrix} \quad , \\ Y &= j_{11}j_{33}\bar{z}\bar{g}_{32}\begin{bmatrix} \bar{g}_{21} \\ \bar{g}_{11} \end{bmatrix} \quad , \\ (I-XY)^{1/2} &= 1 \quad , \\ (I-YX)^{1/2} &= I + \frac{j_{11}j_{33}|z|^2|g_{32}|^2}{2}\begin{bmatrix} -|g_{21}|^2 & g_{11}\bar{g}_{21} \\ -\bar{g}_{11}g_{21} & |g_{11}|^2 \end{bmatrix} \quad , \end{aligned}$$

pri čemu je z *definiran relacijom* (4.2.4), *a* r *relacijom* (4.2.6).

Dokaz:

Dokaz ove leme provodimo na isti način kao i kod leme 4.2.4., smatrajući da je $g_{42} = 0$.

$$G_{32} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ 0 & g_{32} \end{bmatrix} \quad , \tag{4.2.16}$$

uz

$$J_3 = \mathrm{diag}(j_{11}, -j_{11}, j_{33}) \quad . \tag{4.2.17}$$

Korištenjem 3×3 blok J–unitarnih matrica, koje redom imaju dijagonalne blokove dimenzija 2×2 i 1×1, dobivamo

$$T = -R_2R_1^{-1} = -z\begin{bmatrix} 0 & g_{32} \end{bmatrix}\begin{bmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{bmatrix} = z\begin{bmatrix} g_{21}g_{32} & -g_{11}g_{32} \end{bmatrix} \quad .$$

Odatle slijedi

$$\begin{aligned} J_1T^*J_2T &= \begin{bmatrix} j_{11} & \\ & -j_{11} \end{bmatrix}\bar{z}\begin{bmatrix} \bar{g}_{21}\bar{g}_{32} \\ -\bar{g}_{11}\bar{g}_{32} \end{bmatrix}j_{33}z\begin{bmatrix} g_{21}g_{32} & -g_{11}g_{32} \end{bmatrix} \\ &= |z|^2j_{11}j_{33}|g_{32}|^2\begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11}g_{21} & -|g_{11}|^2 \end{bmatrix} \quad . \end{aligned} \tag{4.2.18}$$

Ponovno, to je matrica ranga 1, pa vrijedi

$$(I + J_1 T^* J_2 T)^{-1/2} = I + wu^* \quad .$$

Korištenjem traga matrice u (4.2.18), lako je provjeriti da vrijedi

$$u^* w = j_{11} j_{33} |z|^2 |g_{32}|^2 (|g_{21}|^2 - |g_{11}|^2) \quad ,$$

pa opet imamo degeneraciju, ako je $|g_{11}| = |g_{21}|$. Dakle, u slučaju $|g_{11}| = |g_{21}|$, imamo

$$(I + J_1 T^* J_2 T)^{-1/2} = I - \frac{1}{2} wu^* = I - \frac{j_{11} j_{33} |z|^2 |g_{32}|^2}{2} \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11} g_{21} & -|g_{11}|^2 \end{bmatrix} \quad ,$$

pa je

$$\begin{aligned} X &= T(I + J_1 T^* J_2 T)^{-1/2} \\ &= T - \frac{j_{11} j_{33} z |z|^2 |g_{32}|^2}{2} \begin{bmatrix} g_{21} g_{32} & -g_{11} g_{32} \end{bmatrix} \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11} g_{21} & -|g_{11}|^2 \end{bmatrix} = T \quad . \end{aligned}$$

U slučaju nedegeneriranosti, imamo

$$(I + J_1 T^* J_2 T)^{-1/2} = I + \frac{r - 1}{|g_{21}|^2 - |g_{11}|^2} \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11} g_{21} & -|g_{11}|^2 \end{bmatrix} \quad ,$$

što za X daje

$$\begin{aligned} X &= T(I + J_1 T^* J_2 T)^{-1/2} \\ &= T + \frac{(r-1)z}{|g_{21}|^2 - |g_{11}|^2} \begin{bmatrix} g_{21} g_{32} & -g_{11} g_{32} \end{bmatrix} \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11} g_{21} & -|g_{11}|^2 \end{bmatrix} = rT \quad . \end{aligned}$$

Sad možemo izračunati Y.

$$Y = J_1 X^* J_2 = r j_{11} j_{33} \begin{bmatrix} \bar{t}_{11} \\ -\bar{t}_{12} \end{bmatrix} = r j_{11} j_{33} \bar{z} \bar{g}_{32} \begin{bmatrix} \bar{g}_{21} \\ \bar{g}_{11} \end{bmatrix} \quad .$$

Tada je produkt XY broj

$$XY = r^2 j_{11} j_{33} |z|^2 |g_{32}|^2 (|g_{21}|^2 - |g_{11}|^2) \quad .$$

Ako je $|g_{11}| = |g_{21}|$, jasno je da je $(I - XY)^{1/2} = 1$, a u netrivijalnom slučaju vrijedi $I - XY = r^2$. Objedinjavanjem te dvije formule dobivamo

$$(I - XY)^{1/2} = r \quad ,$$

bez obzira o kojem se slučaju radilo.

Obratni redoslijed množenja matrica X i Y kao rezultat daje 2×2 matricu ranga 1,

$$YX = r^2 j_{11} j_{33} |z|^2 |g_{32}|^2 \begin{bmatrix} |g_{21}|^2 & -g_{11}\bar{g}_{21} \\ \bar{g}_{11} g_{21} & -|g_{11}|^2 \end{bmatrix} \quad .$$

Ponovno, u slučaju $|g_{11}| = |g_{21}|$ vrijedi

$$(I - YX)^{1/2} = I + \frac{j_{11} j_{33} |z|^2 |g_{32}|^2}{2} \begin{bmatrix} -|g_{21}|^2 & g_{11}\bar{g}_{21} \\ -\bar{g}_{11} g_{21} & |g_{11}|^2 \end{bmatrix} \quad ,$$

dok je u ostalim slučajevima

$$(I - YX)^{1/2} = I - \frac{r-1}{|g_{21}|^2 - |g_{11}|^2} \begin{bmatrix} -|g_{21}|^2 & g_{11}\bar{g}_{21} \\ -\bar{g}_{11} g_{21} & |g_{11}|^2 \end{bmatrix} \quad .$$

■

Lema 4.2.7.

Neka su matrice G_{32}, J_3, U i G'_{32}, te skalari z i r, definirani kao u prethodnoj lemi. Tada za elemente matrice G'_{32} vrijedi

(i) *Ako je $|g_{11}| \neq |g_{21}|$*

$$\begin{aligned} g'_{11} &= g_{11} & g'_{12} &= g_{12} + \frac{\bar{g}_{21}(r-1)}{(|g_{21}|^2 - |g_{11}|^2) r z} \\ g'_{21} &= g_{21} & g'_{22} &= g_{22} + \frac{\bar{g}_{11}(r-1)}{(|g_{21}|^2 - |g_{11}|^2) r z} \end{aligned} \quad .$$

(ii) *Ako je $|g_{11}| = |g_{21}|$, onda je*

$$\begin{aligned} g'_{11} &= g_{11} & g'_{12} &= g_{12} - \frac{j_{11} j_{33} \bar{z} \bar{g}_{21} |g_{32}|^2}{2} \\ g'_{21} &= g_{21} & g'_{22} &= g_{12} - \frac{j_{11} j_{33} \bar{z} \bar{g}_{11} |g_{32}|^2}{2} \end{aligned} \quad .$$

Dokaz:

Elementi matrice U definirani su prethodnom lemom. Dokaz provodimo množenjem matrica $U^{-1} G_{32}$ iz relacije (4.2.15).

Slično, kao i prije, ali uz

$$a = j_{11} j_{33} |z|^2 |g_{32}|^2 (|g_{21}|^2 - |g_{11}|^2) \quad , \tag{4.2.19}$$

vrijedi relacija (4.2.14).

Ako je $|g_{11}| \neq |g_{21}|$, elementi matrice G'_{32} imaju oblik

$$\begin{aligned}
g'_{11} &= g_{11} + \frac{(r-1)|g_{21}|^2 g_{11}}{|g_{21}|^2 - |g_{11}|^2} - \frac{(r-1)|g_{21}|^2 g_{11}}{|g_{21}|^2 - |g_{11}|^2} = g_{11} \\
g'_{21} &= g_{21} + \frac{(r-1)|g_{11}|^2 g_{21}}{|g_{21}|^2 - |g_{11}|^2} - \frac{(r-1)|g_{11}|^2 g_{21}}{|g_{21}|^2 - |g_{11}|^2} = g_{21} \\
g'_{12} &= g_{12} + \frac{(r-1)|g_{21}|^2 g_{12}}{|g_{21}|^2 - |g_{11}|^2} - \frac{(r-1) g_{11}\bar{g}_{21} g_{22}}{|g_{21}|^2 - |g_{11}|^2} - r j_{11} j_{33} \bar{z} \bar{g}_{21} |g_{32}|^2 \\
&= g_{12} - \frac{(r-1)\bar{g}_{21}}{(|g_{21}|^2 - |g_{11}|^2) z} - r j_{11} j_{33} \bar{z} \bar{g}_{21} |g_{32}|^2 \\
&= g_{12} - \frac{\bar{g}_{21}}{(|g_{21}|^2 - |g_{11}|^2) z}(r - 1 + ra) = g_{12} + \frac{\bar{g}_{21}(r-1)}{(|g_{21}|^2 - |g_{11}|^2) z r} \\
g'_{22} &= g_{22} + \frac{(r-1)\bar{g}_{11} g_{21} g_{12}}{|g_{21}|^2 - |g_{11}|^2} - \frac{(r-1)|g_{11}|^2 g_{22}}{|g_{21}|^2 - |g_{11}|^2} - r j_{11} j_{33} \bar{z} \bar{g}_{11} |g_{32}|^2 \\
&= g_{22} - \frac{(r-1)\bar{g}_{11}}{(|g_{21}|^2 - |g_{11}|^2) z} - r j_{11} j_{33} \bar{z} \bar{g}_{11} |g_{32}|^2 \\
&= g_{22} - \frac{\bar{g}_{11}}{(|g_{21}|^2 - |g_{11}|^2) z}(r - 1 + ra) = g_{22} + \frac{\bar{g}_{11}(r-1)}{(|g_{21}|^2 - |g_{11}|^2) z r} \quad .
\end{aligned}$$

Ako je $|g_{11}| = |g_{21}|$, vrijedi

$$\begin{aligned}
g'_{11} &= g_{11} - \frac{j_{11} j_{33} |z|^2 |g_{32}|^2 |g_{21}|^2 g_{11}}{2} + \frac{j_{11} j_{33} |z|^2 |g_{32}|^2 |g_{21}|^2 g_{11}}{2} = g_{11} \\
g'_{21} &= g_{21} - \frac{j_{11} j_{33} |z|^2 |g_{32}|^2 |g_{11}|^2 g_{21}}{2} + \frac{j_{11} j_{33} |z|^2 |g_{32}|^2 |g_{11}|^2 g_{21}}{2} = g_{21} \\
g'_{12} &= g_{12} - \frac{j_{11} j_{33} |z|^2 |g_{32}|^2 |g_{21}|^2 g_{12}}{2} + \frac{j_{11} j_{33} |z|^2 |g_{32}|^2 g_{11} \bar{g}_{21} g_{22}}{2} - j_{11} j_{33} \bar{z} \bar{g}_{21} |g_{32}|^2 \\
&= g_{12} + \frac{j_{11} j_{33} \bar{z} \bar{g}_{21} |g_{32}|^2}{2} - r j_{11} j_{33} \bar{z} \bar{g}_{21} |g_{32}|^2 = g_{12} - \frac{j_{11} j_{33} \bar{z} \bar{g}_{21} |g_{32}|^2}{2} \\
g'_{22} &= g_{22} - \frac{j_{11} j_{33} |z|^2 |g_{32}|^2 \bar{g}_{11} g_{21} g_{12}}{2} + \frac{j_{11} j_{33} |z|^2 |g_{32}|^2 |g_{11}|^2 g_{22}}{2} - j_{11} j_{33} \bar{z} \bar{g}_{11} |g_{32}|^2 \\
&= g_{22} + \frac{j_{11} j_{33} \bar{z} \bar{g}_{11} |g_{32}|^2}{2} - r j_{11} j_{33} \bar{z} \bar{g}_{11} |g_{32}|^2 = g_{22} - \frac{j_{11} j_{33} \bar{z} \bar{g}_{11} |g_{32}|^2}{2} \quad .
\end{aligned}$$

■

Kao što je već spomenuto, indefinitna QR dekompozicija nije moguća bez permutacija redaka i stupaca matrice G. U slijedećoj propoziciji pokazuje se veza među matricama koje su unitarne obzirom na matricu J i matricama koje su unitarne obzirom na J', koji je dobiven permutacijama dijagonalnih elemenata matrice J.

Propozicija 4.2.1.

Neka su zadane matrice $G \in \mathbf{C}^{m\times n}$ *i* $J \in \mathbf{C}^{m\times m}$, $J = \mathrm{diag}(j_{11}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$ *i neka su* $G_0 = G, G_1, \ldots, G_k$ *i* $J_0 = J, J_1, \ldots, J_k$ *nizovi matrica dobiveni transformacijama oblika*

$$G_i = Q_i^{-1} P_{1,i}^* G_{i-1} P_{2,i} \quad , \quad J_i = P_{1,i}^* J_{i-1} P_{1,i} \quad , \quad Q_i^* J_i Q_i = J_i \quad , \tag{4.2.20}$$

pri čemu su $P_{1,i}$ *i* $P_{2,i}$ *matrice permutacija. Tada postoje matrice permutacija*

$$P_1 = P_{1,1} P_{1,2} \cdots P_{1,k} \quad , \quad P_2 = P_{2,1} P_{2,2} \cdots P_{2,k} \tag{4.2.21}$$

i J_k*–unitarna matrica*

$$Q = P_{1,k}^* \cdots P_{1,2}^* Q_1 P_{1,2} Q_2 P_{1,3} \cdots P_{1,k} Q_k \quad , \tag{4.2.22}$$

takve da vrijedi

$$Q^{-1} P_1^* G_0 P_2 = G_k \quad .$$

Dokaz:

Prema definiciji transformacija (4.2.20), za matricu G_k vrijedi

$$\begin{aligned} G_k &= Q_k^{-1} P_{1,k}^* G_{k-1} P_{2,k} \\ &= Q_k^{-1} P_{1,k}^* Q_{k-1}^{-1} P_{1,k-1}^* G_{k-2} P_{2,k-1} P_{2,k} \\ &= Q_k^{-1} P_{1,k}^* \cdots Q_1^{-1} P_{1,1}^* G_0 P_{2,1} \cdots P_{2,k} \\ &= Q_k^{-1} P_{1,k}^* \cdots Q_1^{-1} P_{1,2} \cdots P_{1,k} P_{1,k}^* \cdots P_{1,1}^* G_0 P_{2,1} \cdots P_{2,k} \quad , \end{aligned}$$

pa možemo definirati P_1 i P_2 relacijom (4.2.21), a Q relacijom (4.2.22). Preostaje još pokazati da je Q J_k–unitarna matrica. Vrijedi

$$\begin{aligned} Q^* J_k Q &= (Q_k^* P_{1,k}^* \cdots P_{1,2}^* Q_1^* P_{1,2} \cdots P_{1,k}) J_k (P_{1,k}^* \cdots P_{1,2}^* Q_1 P_{1,2} \cdots P_{1,k} Q_k) \\ &= (Q_k^* P_{1,k}^* \cdots P_{1,2}^* Q_1^* P_{1,2} \cdots P_{1,k-1}) J_{k-1} (P_{1,k-1}^* \cdots P_{1,2}^* Q_1 P_{1,2} \cdots P_{1,k} Q_k) \\ &= (Q_k^* P_{1,k}^* \cdots P_{1,2}^* Q_1^*) J_1 (Q_1 P_{1,2} \cdots P_{1,k} Q_k) \\ &= (Q_k^* P_{1,k}^* \cdots P_{1,2}^*) J_1 (P_{1,2} \cdots P_{1,k} Q_k) \\ &= (Q_k^* P_{1,k}^* \cdots Q_2^*) J_2 (Q_2 \cdots P_{1,k} Q_k) = J_k \quad . \end{aligned}$$

■

Dokaz teorema o indefinitnoj QR dekompoziciji matrice G je konstruktivan, i kao kod QR dekompozicije prebacuje se norma (ostatka) stupca na jedan element. U ovom slučaju, termin prebacivanje "J–norme", odgovarat će očuvanju apsolutne vrijednosti indefinitnog skalarnog produkta vektora sa samim sobom. Zbog kratkoće pisanja, a na neki način i tradicionalne paralele sa definitnim slučajem, zadržat ćemo ovu nešto slobodniju interpretaciju prebacivanja "norme".

Teorem 4.2.1.

Neka su dane matrice $G \in \mathbf{C}^{m\times n}$, $m \geq n$ i $J \in \mathbf{C}^{m\times m}$, $J = \mathrm{diag}(j_{11}, \dots j_{mm})$, $j_{ii} \in \{1, -1\}$. Ako je matrica $A = G^*JG$ regularna, postoji indefinitna QR dekompozicija matrice G, tj. rastav oblika

$$G = P_1 Q \begin{bmatrix} R \\ 0 \end{bmatrix} P_2^* \quad , \quad Q^* P_1^* J P_1 Q = P_1^* J P_1 \quad , \tag{4.2.23}$$

gdje su matrice P_1 i P_2 matrice permutacije, matrica R je blok gornjetrokutasta s dijagonalnim blokovima dimenzija 1×1 i 2×2, a Q je $P_1^*JP_1$–unitarna.

Dokaz:

Ako je $J = I$ ili $J = -I$, onda se dekompozicija svodi na standardnu QR dekompoziciju, matrica Q je unitarna, a P_1 i P_2 su jedinične matrice.

Ako propoziciju 4.1.1. interpretiramo u terminima matrice G, postoji, ili barem jedan stupac g_ℓ u matrici G čija je J–norma različita od nule, ili barem jedan par stupaca g_{ℓ_1}, g_{ℓ_2}, takvih da je matrica

$$A_2 = \begin{bmatrix} g_{\ell_1}^* \\ g_{\ell_2}^* \end{bmatrix} J \begin{bmatrix} g_{\ell_1} \, , \, g_{\ell_2} \end{bmatrix}$$

regularna. Primijetimo da pozitivna (ili negativna) definitnost matrice A_2 povlači da su J–norme oba stupca različite od nula, pa taj slučaj promatramo kao dva pojedinačna pivotna stupca. Drugim riječima, dva pivotna stupca imaju smisla, ako i samo ako je matrica A_2 indefinitna i regularna.

Osim toga, ako je J–skalarni produkt dva pivotna stupca jednak nuli, ta dva pivotna stupca smatramo pojedinačnim pivotnim stupcima. To uvijek možemo, jer su izvandijagonalni elementi matrice A_2 jednaki nuli, a zbog njene regularnosti slijedi da su dijagonalni elementi različiti od nule.

Nekom pivotnom strategijom izaberimo regularni pivotni stupac/stupce i primjenom permutacije $P_{2,1}$, dovedimo ih na prvo/prva dva mjesta u matrici G, tj. matricu G smo transformirali u $G_{0,1} = GP_{2,1}$.

Korak 1:

Označimo s g_1^+ elemente prvog stupca matrice $G_{0,1}$, koji odgovaraju $j_{ii} = 1$, a s g_1^- elemente koji odgovararaju $j_{ii} = -1$. Ako g_1^+ ima barem jedan netrivijalan element, permutacijama redaka dovedimo ga na prvo mjesto u vektoru g_1. Tada postoji niz trigonometrijskih rotacija $U_G^{-1}(i,j)$, takvih da poništavaju sve, osim jednog elementa (recimo onog u prvom retku) u g_1^+. Na sličan način postupimo i s g_1^-. Ako postoji netrivijalan element u g_1^-, dovedimo ga permutacijama redaka na drugo mjesto u prvom stupcu i nizom trigonometrijskih rotacija poništimo ostatak elemenata g_1^-.

Prema propoziciji 3.1.5., takve trigonometrijske rotacije su J_1–unitarne matrice, pri čemu je $J_1 = P_{1,1}^* J P_{1,1}$, a $P_{1,1}$ je matrica permutacija koja je na prvo mjesto u g_1 dovela neki element iz g_1^+, a na drugo mjesto neki element iz g_1^-. Prema propoziciji 3.1.2., J_1–unitarne matrice čine grupu, pa je njihov produkt J_1–unitarna matrica. Nakon ovog koraka transformacije, matrica G ima oblik

$$G_1 = Q_1^{-1} P_{1,1}^* G_{0,1} = Q_1^{-1} P_{1,1}^* G P_{2,1} = \begin{bmatrix} g_{11} & * & * & \cdots & * & * \\ g_{21} & * & * & \cdots & * & * \\ 0 & * & * & \cdots & * & * \\ \vdots & * & * & \cdots & * & * \\ 0 & * & * & \cdots & * & * \end{bmatrix} \quad ,$$

gdje $*$ predstavlja neke, općenito, netrivijalne elemente.

Korak 2 – jedan pivotni stupac:

Slučaj: jedan od elemenata g_{11}, g_{21} je različit od 0

Zbog regularnosti matrice G, g_{11} i g_{21} ne mogu istovremeno biti jednaki nula. Ako je $g_{11} = 0$, permutirajmo prvi i drugi redak matrice G_1,

$$G_2 = P_{1,2}^* G_1 \quad .$$

Ako je $g_{21} = 0$, onda je $G_2 = G_1$. Time je prvi stupac doveden na konačni oblik, pa nastavljamo raditi s matricom $G_2([2:m],[2:n])$.

Slučaj: $g_{11} \neq 0$ i $g_{21} \neq 0$

Bez smanjenja općenitosti, možemo smatrati da su g_{11} i g_{21} realni i veći od nule. Ako nisu, korištenjem propozicije 3.1.6., to možemo postići.

Ako je $g_{11} > g_{21}$, onda postoji hiperbolna rotacija $U_H^{-1}(1,2)$ takva da poništava element g_{21}, tj. za restrikciju $U_H^{-1}(1,2)([1,2],[1,2])$ vrijedi

$$\begin{bmatrix} \operatorname{ch}\varphi & -\operatorname{sh}\varphi \\ -\operatorname{sh}\varphi & \operatorname{ch}\varphi \end{bmatrix} \begin{bmatrix} g_{11} \\ g_{21} \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix} \quad ,$$

što vodi na jednadžbu

$$\operatorname{th}\varphi = \frac{g_{21}}{g_{11}} \quad .$$

Ako vrijedi obratno, $g_{11} < g_{21}$, onda zamjenom prva dva retka u matrici G_1 i zamjenom elemenata j_{11} i j_{22} u matrici J_1, možemo izračunati $\operatorname{th}\varphi$.

Nakon drugog koraka transformacija dobivamo matricu G_2 traženog oblika

$$G_2 = Q_2^{-1} P_{1,2}^* G_1 = \begin{bmatrix} g'_{11} & * & * & \cdots & * & * \\ 0 & * & * & \cdots & * & * \\ 0 & * & * & \cdots & * & * \\ \vdots & * & * & \cdots & * & * \\ 0 & * & * & \cdots & * & * \end{bmatrix} \quad .$$

Korak 2 – dva pivotna stupca:

Slučaj: $g_{11} \neq 0$ **i** $g_{21} \neq 0$

Ako su oba elementa g_{11} i g_{21} bila netrivijalna, prvi stupac restrikcije matrice G_1 tj. matrice $G_1([3:m],[2:n])$, poništimo kao u koraku 1. Time smo dobili matricu

$$G_3 = Q_3^{-1} P_{1,3}^* G_1 = \begin{bmatrix} g_{11} & g_{12} & * & \cdots & * & * \\ g_{21} & g_{22} & * & \cdots & * & * \\ 0 & g_{32} & * & \cdots & * & * \\ 0 & g_{42} & * & \cdots & * & * \\ \vdots & 0 & * & \cdots & * & * \\ 0 & \vdots & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \end{bmatrix}.$$

Postoje točno tri bitno različite mogućnosti za odnos elemenata g_{32} i g_{42}.

Ako su i $g_{32} = 0$ i $g_{42} = 0$, onda smo završili poništavanje prva dva stupca matrice G, pa definiramo $G_4 = G_3$.

U drugom slučaju, ako je točno jedan od elemenata g_{32}, g_{43} različit od nule, permutacijama redaka možemo ga dovesti na mjesto $(3,2)$ u matrici G_3. Tvrdimo da tada imamo elementarni par reda 3, jer g_{12} i g_{22} ne mogu istovremeno biti jednaki nula. Kad bi bili, onda bi J_1–skalarni produkt prva dva stupca bio jednak nuli. Prema lemi 4.2.2., postoji J_1'–unitarna matrica, $J_1' = P_{1,4}^* J_1 P_{1,4}$, koja množi matricu G_3 slijeva i osigurava da je $g_{11}g_{22} - g_{12}g_{21} \neq 0$. Korištenjem leme 4.2.6., tj. blok 3×2 transformacije, reduciramo matricu G_3 na tražni oblik.

U trećem slučaju, oba elementa g_{32}, g_{43} su različita od nule. Na isti način kao u prošlom slučaju, zaključujemo da g_{12} i g_{22} nisu istovremeno nula, pa imamo elementarni par reda 4. Prema lemi 4.2.3., postoji J_1'–unitarna matrica, $J_1' = P_{1,4}^* J_1 P_{1,4}$, koja slijeva množi matricu G_3 i osigurava da je $g_{11}g_{22} - g_{12}g_{21} \neq 0$. Matricu G_3 reduciramo primjenom blok 4×2 transformacije (lema 4.2.4.).

Bez obzira na odnos elemenata g_{32} i g_{42}, nakon transformacije dobivamo matricu oblika

$$G_4 = Q_4^{-1} P_{1,4}^* G_3 = \begin{bmatrix} g_{11}' & g_{12}' & * & \cdots & * & * \\ g_{21}' & g_{22}' & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \\ \vdots & \vdots & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \end{bmatrix}.$$

Algoritam nastavljamo s podmatricom $G_4([3:m],[3:n])$.

Slučaj: jedan od elemenata g_{11}, g_{21} je različit od 0

Ako je točno jedan od elemenata g_{11} i g_{21} različit od nule, poništavamo prvi stupac restrikcije matrice G_2 tj. matrice $G_2([2:m],[2:n])$. Time smo dobili matricu

$$G_3 = Q_3^{-1}P_{1,3}^* G_2 = \begin{bmatrix} g_{11} & g_{12} & * & \cdots & * & * \\ 0 & g_{22} & * & \cdots & * & * \\ 0 & g_{32} & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \\ \vdots & 0 & * & \cdots & * & * \\ 0 & \vdots & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \end{bmatrix} .$$

Uočimo da ne mogu istovremeno biti g_{22} i g_{32} jednaki nula, jer bi matrica G_3 imala proporcionalne stupce i bila singularna. Ako je točno jedan od ta dva elementa različit od nule, permutacijama redaka dovedimo ga na mjesto $(2,2)$ u matrici G_3, tj. $G_4 = P_{1,4}^* G_3$, čime smo završili transformaciju. Rad nastavljamo na podmatrici $G_4([3:m],[3:n])$.

Ako su oba elemanta g_{22} i g_{32} netrivijalna, onda je osiguran uvjet da je determinanta gornjeg 2×2 bloka netrivijalna i imamo elementarni par reda 3, koji transformiramo prema lemi 4.2.6., a zatim nastavljamo dekomponirati podmatricu $G_4([3:m],[3:n])$.

Primjenom koraka za dva pivotna stupca, reducirali smo problem indefinitne QR dekompozicije na matricu reda $(m-2) \times (n-2)$.

Tvrdnja teorema izlazi ponovljenim primjenama koraka 1, odnosno 2, na matricama dimenzija $(m-1) \times (n-1)$, odnosno $(m-2) \times (n-2)$. Prema propoziciji 4.2.1., nizovi matrica Q_i, $P_{1,i}$, $P_{2,i}$ i J_i definirani u algoritmu, definiraju matrice iz (4.2.23). ■

Napomena 4.2.2.

Ako je matrica G iz prethodnog teorema bila regularna, a matrica $A = G^* J G$ nije punog ranga, indefinitna QR dekompozicija matrice G, je rastav oblika

$$G = P_1 Q \begin{bmatrix} R & T_1 \\ 0 & T_2 \end{bmatrix} P_2^* \quad , \quad Q^* P_1^* J P_1 Q = P_1^* J P_1 = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \quad ,$$

gdje su matrice P_1 i P_2 matrice permutacije, matrica R je blok gornjetrokutasta s dijagonalnim blokovima dimenzija 1×1 i 2×2, a T_1 i T_2 su pune matrice. Dimenzije podmatrica J_1 i J_2, redom odgovaraju dimenzijama podmatrica T_1 i T_2. Osim toga, za matricu T_2 vrijedi $T_2^* J_2 T_2 = 0$, tj. to je općenito netrivijalna dekompozicija nul–matrice. Ovaj oblik dekompozicije je posljedica propozicije 4.1.1., jer nakon što smo izdvojili blok R, Gramova matrica J_2–skalarnih produkata "ostataka" stupaca $A_2 = T_2^* J_2 T_2$ je egzaktno singularna i potpuna redukcija se ne može provesti.

Stupce matrice T_2 ipak možemo "skratiti" permutacijama redaka i primjenama niza Givensovih transformacija na matrici T_2, te odgovarajućim permutacijama matrice J_2, tako da je

$$T_2' = P_1' U_{G_1} \cdots U_{G_p} T_2 = \begin{bmatrix} T_2'' \\ 0 \end{bmatrix} ,$$

gdje T_2'' oblika

$$T_2'' = \begin{bmatrix} * & * & \cdots & * & * \\ * & * & \cdots & * & * \\ & * & \cdots & * & * \\ & * & \cdots & * & * \\ & & & * & * \\ & & & * & * \\ & & & & * \\ & & & & * \end{bmatrix} .$$

∎

Napomena 4.2.3.

Primijetimo da u dokazu teorema 4.2.1., korak 2 za jedan stupac, možemo provoditi sve dok je J–norma odgovarajućeg stupca različita od nule. Numerički, ovo nije dobra odluka, jer može doći do kraćenja kod računanja hiperbolnog kuta. Zbog toga, pametnije je primijeniti neku pivotnu strategiju, koja će izabrati dva pivotna stupca i kad je J–norma stupaca mala prema njihovom skalarnom produktu. ∎

4.3. Realizacija dekompozicije Householderovim matricama

Kao što je poznato, QR dekompozicija realizirana ravninskim rotacijama sporija je od one realizirane Householderovim matricama, ali ima nešto manju pogrešku. Kod indefinitne QR dekompozicije, situacija će se još više promijeniti u korist realizacije rotacijama. Naime, javljaju se teškoće pri izračunavanju blok–Householderovih matrica, zbog nejedinstvenosti dijagonalnih 2×2 blokova matrice R.

Ipak, znajući kako izgledaju blokovi dobiveni rotacijama, možemo pretpostaviti isti takav izgled blokova. Treba još samo pokazati da u tom slučaju postoje odgovarajuće Householderove matrice.

U slučaju 1×1 dijagonalnih blokova, ovaj je algoritam potpuni analogon obične QR dekompozicije realizirane reflektorima. U slučaju 2×2 blokova, Householderovim i J–Householderovim matricama stupce "skratimo" na isti oblik kao kod dekompozicije realizirane rotacijama. Nakon toga, koristimo rezultat o obliku 2×2 blokova dobivenih iz dekompozicije rotacijama, te blok J–Householderove matrice.

Poznata je činjenica da Householderove matrice prevode vektor odgovarajuće norme u neki drugi vektor iste norme. Sličan rezulatat dokazat ćemo i za J–Householderove matrice.

Teorem 4.3.1.

Neka su $a,\ b \in \mathbf{C}^m$ takvi da vrijedi $a^*a = b^*b$. Tada postoji Householderova matrica

$$H = I - 2\frac{ww^*}{w^*w} \quad .$$

takva da vrijedi

$$Ha = -\sigma b \quad ,$$

pri čemu je

$$\sigma = \begin{cases} \dfrac{b^*a}{|a^*b|} & , \ a^*b \neq 0 \\ 1 & , \ a^*b = 0 \end{cases} \quad .$$

Dokaz:

Definirajmo vektor $w = a + \sigma b$. Izračunajmo prvo w^*w.

$$w^*w = (a^* + \bar{\sigma}b^*)(a + \sigma b) = a^*a + \sigma a^*b + \bar{\sigma}b^*a + |\sigma|^2 b^*b = 2a^*a + 2|a^*b| \quad .$$

Odmah slijedi

$$\begin{aligned} Ha &= a - 2\frac{ww^*a}{w^*w} = a - 2\frac{(a+\sigma b)(a^* + \bar{\sigma}b^*)a}{2(a^*a + |a^*b|)} \\ &= a - \frac{(a+\sigma b)(a^*a + \bar{\sigma}b^*a)}{a^*a + |a^*b|} \\ &= a - \frac{(a+\sigma b)(a^*a + |a^*b|)}{a^*a + |a^*b|} = -\sigma b \quad . \end{aligned}$$

■

Ako isključimo mogućnost degeneriranih vektora (onih koji su okomiti na same sebe), postoji potpuna analogija "prevođenja" jednog vektora u drugi iste J–norme. Ova tvrdnja vrijedi za Bunse–Gerstner oblik J–Householderove matrice.

Teorem 4.3.2.

Neka su $a,\ b \in \mathbf{C}^m$ takvi da vrijedi $a^*Ja = b^*Jb \neq 0$. Tada postoji J–Householderova matrica H takva da vrijedi

$$Ha = -\sigma Jb \quad ,$$

pri čemu je

$$\sigma = \begin{cases} \operatorname{sign}(a^*Ja)\dfrac{b^*a}{|a^*b|} & , \ a^*b \neq 0 \\ \operatorname{sign}(a^*Ja) & , \ a^*b = 0 \end{cases} \quad .$$

Dokaz:

Slično kao u prošlom teoremu definiramo vektor $w = a + \sigma Jb$. Izračunajmo prvo $w^* Jw$.

$$\begin{aligned} w^* Jw &= (a^* + \bar{\sigma} b^* J)J(a + \sigma Jb) = a^* Ja + \sigma a^* b + \bar{\sigma} b^* a + |\sigma|^2 b^* Jb \\ &= 2a^* Ja + 2\,\mathrm{sign}(a^* Ja)|a^* b| \quad . \end{aligned}$$

Odavde izlazi

$$\begin{aligned} Ha &= a - 2\frac{ww^* Ja}{w^* Jw} = a - 2\frac{(a + \sigma Jb)(a^* + \bar{\sigma} b^* J)Ja}{2(a^* Ja + \mathrm{sign}(a^* Ja)|a^* b|)} \\ &= a - \frac{(a + \sigma Jb)(a^* Ja + \bar{\sigma} b^* a)}{a^* Ja + \mathrm{sign}(a^* Ja)|a^* b|} \\ &= a - \frac{(a + \sigma Jb)(a^* Ja + \mathrm{sign}(a^* Ja)|a^* b|)}{a^* Ja + \mathrm{sign}(a^* Ja)|a^* b|} = -\sigma Jb \quad . \end{aligned}$$

■

Standardna zadaća blok Householderovih matrica je poništavanje blokova. Neka je

$$E = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} \in \mathbf{C}^{m \times p} \quad , \quad m > p \quad , \tag{4.3.1}$$

pri čemu je E_1 kvadratna matrica. Tražimo blok reflektor H tako da vrijedi

$$HE = F = \begin{bmatrix} F_1 \\ 0 \end{bmatrix} \quad , \tag{4.3.2}$$

pri čemu je F_1 kvadratna matrica. U teoremu 4.3.3., dokazanom u [37], dani su nužni i dovoljni uvjeti za egzistenciju takvog reflektora H. Za dokaz tog teorema potrebne su nam još neke definicije i jedna lema.

Definicija 4.3.1.

Neka su matrice E, F i F_1 definirane relacijama (4.3.1) *i* (4.3.2). *Reći ćemo da one zadovoljavaju svojstva*

(*i*) *izometrije, ako je*

$$F^* F = F_1^* F_1 = E^* H^* HE = E^* E \quad ,$$

(*ii*) *simetrije, ako je*

$$E_1^* F_1 = F_1^* E_1 = E^* HE = E^* H^* E \quad .$$

■

Za matricu H definiranu relacijom (3.4.3), vrijedi propozicija 3.4.3., pa E, F i F_1 ispunjavaju tražena svojstva, jer je $H = H^* = H^{-1}$.

Definicija 4.3.2.

Matrice S i D, pridruženene matricama E i F definirane su relacijama

$$S = F + E = \begin{bmatrix} F_1 + E_1 \\ E_2 \end{bmatrix} \quad , \quad D = F - E = \begin{bmatrix} F_1 - E_1 \\ -E_2 \end{bmatrix} \quad . \tag{4.3.3}$$

■

Lema 4.3.1.

Ako matrica F_1 zadovoljava svojstva izometrije i simetrije, vrijedi

$$D^* S = 0 \quad .$$

Dokaz:

Prema definiciji matrica S i D, te svojstvima izometrije i simetrije, imamo

$$D^* S = (F^* F - E^* E) + (F^* E - E^* F) = (F_1^* F_1 - E^* E) + (F_1^* E_1 - E_1^* F_1) = 0 \quad .$$

■

Izrecimo sad pripadni teorem.

Teorem 4.3.3. (Schreiber–Parlett)

Postoji blok reflektor H takav da je $HE = F$, ako i samo ako za pripadne matrice vrijede svojstva simetrije i izometrije.

Dokaz:

Izometrija je posljedica unitarnosti matrice H, a simetrija je posljedica hermitičnosti matrice H.

Da bismo dokazali dovoljnost, pokažimo da je

$$H(D)E = F \quad .$$

Iz definicija za S i D, izlazi da je

$$E = \frac{1}{2}(S - D) \quad , \quad F = \frac{1}{2}(S + D) \quad .$$

Korištenjem leme 4.3.1. i svojstva "kraćenja" generaliziranog inverza, slijedi

$$\begin{aligned} H(D)E &= \frac{1}{2}(H(D)S - H(D)D) \\ &= \frac{1}{2}[(S - 2D(D^*D)^+ D^* S) - (D - 2D(D^*D)^+ D^* D)] \\ &= \frac{1}{2}(S + D) = F \quad . \end{aligned}$$

Jednako tako, ako F_1 zadovoljava relacije simetrije i izometrije, zadovoljava ih i $-F_1$. Dapače, vrijedi

$$
\begin{aligned}
H(S)E &= \frac{1}{2}(H(S)S - H(S)D) \\
&= \frac{1}{2}[(S - 2S(S^*S)^+S^*S) - (D - 2S(S^*S)^+S^*D)] \\
&= \frac{1}{2}(-S - D) = -F \quad .
\end{aligned}
$$

■

Prethodni teorem daje nam motivaciju za generalizaciju na J–blok Householderove matrice. Neka je

$$
E = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} \in \mathbf{C}^{m\times p} \quad , \quad m > p \quad , \tag{4.3.4}
$$

pri čemu je E_1 kvadratna matrica. Tražimo blok J–Householderovu matricu H tako da vrijedi

$$
HE = F = \begin{bmatrix} F_1 \\ 0 \end{bmatrix} \quad , \tag{4.3.5}
$$

pri čemu je F_1 kvadratna matrica, a H (Bunse–Gerstner oblik) unitarna obzirom na indefinitni skalarni produkt definiran matricom

$$
J = \begin{bmatrix} J_1 & \\ & J_2 \end{bmatrix} \quad , \tag{4.3.6}
$$

koja zadovoljava $J = J^* = J^{-1}$. Svi teoremi koje ćemo izreći vrijedit će u ovom generalnijem slučaju, a primjenjivat ćemo ih na $J = \mathrm{diag}(j_{11}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$.

Analogon definicije 4.3.1. u indefinitnom slučaju je slijedeća definicija:

Definicija 4.3.3.

Neka su matrice E, F i F_1 definirane relacijama (4.3.4) *i* (4.3.5). *Reći ćemo da one zadovoljavaju svojstva*

(*i*) *J–izometrije, ako je*

$$
F_1^* J_1 F_1 = F^* J F = E^* H^* J H E = E^* J E \quad ,
$$

(*ii*) *simetrije obzirom na J, ako je*

$$
E_1^* J_1 F_1 = F_1^* J_1 E_1 = E^* J H E = E^* H^* J E \quad .
$$

■

Blok J–Householderova matrica H definirana relacijom (3.4.4) zadovoljava propoziciju 3.4.4., pa je za E, F i F_1, definicija trivijalno ispunjena.

Uočimo da je ovdje bitan izbor generalizacije, jer oblik Cybenko–Berry ne zadovoljava prethodnu definiciju.

Neka su pridružene matrice S i D definirane kao i prije, relacijom (4.3.3). Analogon leme 4.3.1. glasi

Lema 4.3.2.

Ako matrica F_1 zadovoljava svojstva J izometrije i simetrije obzirom na J, vrijedi

$$D^*JS = 0 \quad .$$

Dokaz:

Prema definiciji matrica S i D, te svojstvima J–izometrije i simetrije obzirom na J, imamo

$$\begin{aligned} D^*JS &= [F_1^* - E_1^*, -E_2^*] \begin{bmatrix} J_1 & \\ & J_2 \end{bmatrix} \begin{bmatrix} F_1 + E_1 \\ E_2 \end{bmatrix} \\ &= F_1^* J_1 F_1 - E_1^* J_1 F_1 + F_1^* J_1 E_1 - E_1^* J_1 E_1 - E_2^* J_2 E_2 \\ &= -E_1^* J_1 F_1 + F_1^* J_1 E_1 = 0 \quad . \end{aligned}$$

■

Generalizacija teorema 4.3.3. zakomplicirat će se mogućnošću gubitka ranga matrice W^*JW obzirom na rang od W (vidjeti teorem 3.4.4.).

Teorem 4.3.4.

Postoji blok reflektor H takav da je $HE = F$, ako za pripadne matrice vrijede svojstva simetrije obzirom na J i J–izometrije, a za D

$$\operatorname{rang}(D^*JD) = \operatorname{rang}(D) \quad .$$

Ekvivalentno, postoji blok reflektor takav da je $HE = -F$, ako za pripadne matrice vrijede svojstva simetrije obzirom na J i J–izometrije, a za S

$$\operatorname{rang}(S^*JS) = \operatorname{rang}(S) \quad .$$

Dokaz:

Dio iskaza o rangu u teoremu nužan je zbog "kraćenja" generaliziranih inverza. Neka D zadovoljava zahtjev o rangu. Da bismo dokazali dovoljnost, pokažimo da je

$$H(D)E = F \quad .$$

Iz definicija za S i D, izlazi da je

$$E = \frac{1}{2}(S - D) \quad , \quad F = \frac{1}{2}(S + D) \quad .$$

Korištenjem prethodne leme i svojstva "kraćenja" generaliziranog inverza, slijedi

$$\begin{aligned} H(D)E &= \frac{1}{2}(H(D)S - H(D)D) \\ &= \frac{1}{2}[(S - 2D(D^*JD)^+D^*JS) - (D - 2D(D^*JD)^+D^*JD)] \\ &= \frac{1}{2}(S + D) = F \quad . \end{aligned}$$

Jednako tako, ako F_1 zadovoljava relacije simetrije i izometrije, zadovoljava ih i $-F_1$. Na sličan način možemo dokazati da vrijedi

$$\begin{aligned} H(S)E &= \frac{1}{2}(H(S)S - H(S)D) \\ &= \frac{1}{2}[(S - 2S(S^*JS)^+S^*JS) - (D - 2S(S^*JS)^+S^*JD)] \\ &= \frac{1}{2}(-S - D) = -F \quad , \end{aligned}$$

ako S zadovoljava zahtjev o rangu. ■

Na ovom mjestu treba okarakterizirati formu matrice F_1. Naime, matrica F_1 ne mora biti jedinstvena, čak i kad je E punog stupčanog ranga, a matrica $J = I$. Na primjer, kad je $E_1 = 0$, svojstvo simetrije je suvišno, pa F_1 mora zadovoljavati samo svojstvo izometrije.

U drugoj krajnosti, ako je E_1 invertibilna, a $J = I$, F mora imati formu

$$F_1 = ME_1 \quad , \tag{4.3.7}$$

pri čemu je M hermitska matrica. Pokažimo da vrijedi relacija (4.3.7). Householderovu matricu H, napišimo u blok obliku, tako da blokovi odgovaraju dimenzijama blokova E_1 i E_2, tj. neka je

$$H = \begin{bmatrix} H_{11} & H_{12} \\ \bar{H}_{12} & H_{22} \end{bmatrix} \quad .$$

Tada vrijedi

$$HE = \begin{bmatrix} H_{11}E_1 + H_{12}E_2 \\ \bar{H}_{12}E_1 + H_{22}E_2 \end{bmatrix} = \begin{bmatrix} F_1 \\ 0 \end{bmatrix} \quad .$$

Zbog regularnosti matrice E_1 slijedi

$$F_1 = H_{11}E_1 + H_{12}E_2 = (H_{11} + H_{12}E_2E_1^{-1})E_1 = ME_1 \quad .$$

Preostaje još pokazati da je M hermitska matrica. Iz svojstva simetrije izlazi da je $E_1^* M E_1 = E_1^* M^* E_1$, pa množenjem slijeva s E_1^{-*} i zdesna s E_1^{-1} dobivamo traženo svojstvo. Na sličan način, ako je $J \neq I$, možemo pokazati da F_1 ima oblik (4.3.7) s tim da je M J–hermitska matrica.

Vratimo se zadatku poništavanja blokova, odnosno indefinitne QR dekompozicije realizirane Householderovim matricama. Ako odgovarajući stupac u matrici G ima J–normu različitu od 0, korištenjem J–Householderovih matrica kao posljedicu teorema 4.3.2., možemo tu normu "prebaciti" na samo jedan element. Potpuno je drugačija situacija, ako odgovarajuća pivotna strategija zahtjeva rad na dva stupca matrice G. Kao i kod realizacije rotacijama, takav slučaj možemo očekivati samo ako je odgovarajuća 2×2 podmatrica matrice $A = G^* J G$ indefinitna.

Pravi problem je, istovremeno poništavanje dva stupca. Zbog nejedinstvenosti J–Householderovih matrica i pitanja degeneriranosti, nije sigurno da uopće postoji blok J–Householderova matrica koja takve stupce svodi na 2×2 blok.

Zbog osiguranja egzistencije transformacije, moramo ta dva stupca dovesti u kondenziraniju formu. Uočimo da se Householderovim i J–Householderovim matricama takva dva stupca daju dovesti u isti oblik kao i kod realizacije dekompozicije rotacijama. Prvi se stupac može običnim reflektorima svesti na dva elementa. Ako je J–norma ostatka drugog stupca približno jednaka 0, onda i ostatak drugog stupca možemo Householderovim matricama svesti na samo dva elementa, tj. na elementarni par reda 4

$$E = G_{42} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ 0 & g_{32} \\ 0 & g_{42} \end{bmatrix} \quad , \tag{4.3.8}$$

uz

$$J_4 = \begin{bmatrix} j_{11} & & & \\ & -j_{11} & & \\ & & j_{33} & \\ & & & -j_{33} \end{bmatrix} \quad .$$

Ako je J–norma ostatka drugog stupca "dovoljno velika", ostatak drugog stupca možemo J–Householderovim matricama svesti na samo jedan element, tj. restrikcije odgovarajuće matrice E i pripadni J čine elementarni par reda 3

$$E = G_{32} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ 0 & g_{32} \end{bmatrix} \quad , \tag{4.3.9}$$

uz

$$J_3 = \begin{bmatrix} j_{11} & & \\ & -j_{11} & \\ & & j_{33} \end{bmatrix} \quad .$$

Očito smo odgovarajuća dva stupca "skratili" na najviše 6 elemenata. Tvrdimo da je u tom slučaju lako konstruirati blok J–Householderovu matricu koja poništava elemente g_{32} i g_{42}, odnosno, samo g_{32}.

Iskoristimo činjenicu da je iz realizacije dekompozicije rotacijama, poznata jedna od mogućih formi desne strane, tj. jedna od matrica F_1.

Korištenjem leme 4.2.5., za elementarni par reda 4 dobivamo

$$F_1 = \begin{bmatrix} g_{11} & g_{12} + \dfrac{\bar{g}_{21}(r-1)}{(|g_{21}|^2 - |g_{11}|^2) r z} \\ g_{21} & g_{22} + \dfrac{\bar{g}_{11}(r-1)}{(|g_{21}|^2 - |g_{11}|^2) r z} \end{bmatrix}, \tag{4.3.10}$$

ako je $|g_{11}| \neq |g_{21}|$ i $|g_{32}| \neq |g_{42}|$, odnosno

$$F_1 = \begin{bmatrix} g_{11} & g_{12} - \dfrac{j_{11} j_{33} \bar{z} \bar{g}_{21}(|g_{32}|^2 - |g_{42}|^2)}{2} \\ g_{21} & g_{12} - \dfrac{j_{11} j_{33} \bar{z} \bar{g}_{11}(|g_{32}|^2 - |g_{42}|^2)}{2} \end{bmatrix}, \tag{4.3.11}$$

ako je $|g_{11}| = |g_{21}|$ ili $|g_{32}| = |g_{42}|$.

Korištenjem leme 4.2.7., za elementarni par reda 3, za F_1 dobivamo

$$F_1 = \begin{bmatrix} g_{11} & g_{12} + \dfrac{\bar{g}_{21}(r-1)}{(|g_{21}|^2 - |g_{11}|^2) r z} \\ g_{21} & g_{22} + \dfrac{\bar{g}_{11}(r-1)}{(|g_{21}|^2 - |g_{11}|^2) r z} \end{bmatrix}, \tag{4.3.12}$$

ako je $|g_{11}| \neq |g_{21}|$, odnosno

$$F_1 = \begin{bmatrix} g_{11} & g_{12} - \dfrac{j_{11} j_{33} \bar{z} \bar{g}_{21} |g_{32}|^2}{2} \\ g_{21} & g_{12} - \dfrac{j_{11} j_{33} \bar{z} \bar{g}_{11} |g_{32}|^2}{2} \end{bmatrix}, \tag{4.3.13}$$

ako je $|g_{11}| = |g_{21}|$.

Preostaje nam samo još pokazati da je u tim slučajevima moguća konstrukcija J–blok Householderovih matrica. Uočimo da su ispunjeni svi uvjeti iz teorema 4.3.4., osim uvjeta o rangu matrica D, odnosno S.

Promotrimo prvo slučaj elementarnog para reda 4. Vrijedi slijedeća propozicija.

Propozicija 4.3.1.

Neka je matrica E oblika kao u relaciji (4.3.8).

(i) *Ako je* $|g_{11}| \neq |g_{21}|$ *i* $|g_{32}| \neq |g_{42}|$, *tada je*

$$D = \begin{bmatrix} 0 & \bar{g}_{21}p \\ 0 & \bar{g}_{11}p \\ 0 & -g_{32} \\ 0 & -g_{42} \end{bmatrix} \quad , \quad S = \begin{bmatrix} 2g_{11} & 2g_{12} + \bar{g}_{21}p \\ 2g_{21} & 2g_{22} + \bar{g}_{11}p \\ 0 & g_{32} \\ 0 & g_{42} \end{bmatrix} \quad , \tag{4.3.14}$$

gdje je

$$p = \frac{r-1}{(|g_{21}|^2 - |g_{11}|^2)zr} \quad .$$

(ii) *Ako je* $|g_{11}| = |g_{21}|$ *ili* $|g_{32}| = |g_{42}|$, *tada je*

$$D = \begin{bmatrix} 0 & -\bar{g}_{21}s \\ 0 & -\bar{g}_{11}s \\ 0 & -g_{32} \\ 0 & -g_{42} \end{bmatrix} \quad , \quad S = \begin{bmatrix} 2g_{11} & 2g_{12} - \bar{g}_{21}s \\ 2g_{21} & 2g_{22} - \bar{g}_{11}s \\ 0 & g_{32} \\ 0 & g_{42} \end{bmatrix} \quad . \tag{4.3.15}$$

gdje je

$$s = \frac{j_{11}j_{33}\bar{z}(|g_{32}|^2 - |g_{42}|^2)}{2} \quad .$$

Dokaz:

Tvrdnju dobivamo supstituiranjem matrica E i F iz relacija (4.3.8), odnosno (4.3.10) i (4.3.11), u definicionu relaciju (4.3.3) za D i S. ■

U slijedećoj propoziciji promotrimo odnose rang(D) i rang(D^*JD) i, ekvivalentno, za S.

Propozicija 4.3.2.

Za matrice D i S definirane prethodnom propozicijom vrijedi

(i) *Ako je* $|g_{32}| \neq |g_{42}|$, *tada je*

$$\text{rang}(D) = \text{rang}(D^*JD) = 1 \quad .$$

(ii) *Ako je* $|g_{32}| = |g_{42}|$, *tada je*

$$\text{rang}(D) = 1 \quad , \quad \text{rang}(D^*JD) = 0 \quad .$$

(iii) *Uvijek vrijedi*

$$\text{rang}(S) = \text{rang}(S^*JS) = 2 \quad .$$

Dokaz:

Dokažimo prvo relacije za rang(D). Zbog toga što E i J_4 čine elementarni par reda 4, elementi g_{32} i g_{42} ne mogu biti 0, pa je u oba slučaja matrica D ranga 1.

(i) Načinimo produkt D^*JD u slučaju kad je $|g_{11}| \neq |g_{21}|$ i $|g_{32}| \neq |g_{42}|$

$$D^*JD = \begin{bmatrix} 0 & 0 \\ 0 & j_{11}|p|^2(|g_{21}|^2 - |g_{11}|^2) + j_{33}(|g_{32}|^2 - |g_{42}|^2) \end{bmatrix} \quad .$$

Treba još pokazati da jedini netrivijalni element u matrici D^*JD, zaista nije 0. Za taj element, zovimo ga c, korištenjem relacije (4.2.13), izlazi

$$\begin{aligned} c &= j_{11}\frac{(r-1)^2}{(|g_{21}|^2 - |g_{11}|^2)|z|^2r^2} + j_{33}(|g_{32}|^2 - |g_{42}|^2) \\ &= \frac{j_{11}}{(|g_{21}|^2 - |g_{11}|^2)|z|^2}\left(\frac{(r-1)^2}{r^2} + j_{11}j_{33}|z|^2(|g_{21}|^2 - |g_{11}|^2)(|g_{32}|^2 - |g_{42}|^2)\right) \\ &= \frac{j_{11}}{(|g_{21}|^2 - |g_{11}|^2)|z|^2}\left((1-\sqrt{1+a})^2 + a\right) \\ &= \frac{j_{11}}{(|g_{21}|^2 - |g_{11}|^2)|z|^2}\left(2 - 2\sqrt{1+a} + 2a\right) \\ &= \frac{j_{11}}{(|g_{21}|^2 - |g_{11}|^2)|z|^2}2\left(\sqrt{1+a}(\sqrt{1+a} - 1)\right) \quad . \end{aligned}$$

Odatle je, zbog $a \neq 0$, i $a > -1$, $\mathrm{rang}(D^*JD) = 1$.

(ii) U slučaju kad je $|g_{11}| = |g_{21}|$ ili $|g_{32}| = |g_{42}|$, vrijedi

$$D^*JD = \begin{bmatrix} 0 & 0 \\ 0 & j_{11}|s|^2(|g_{21}|^2 - |g_{11}|^2) + j_{33}(|g_{32}|^2 - |g_{42}|^2) \end{bmatrix} \quad .$$

Uočimo da je

$$\begin{aligned} & j_{11}|s|^2(|g_{21}|^2 - |g_{11}|^2) + j_{33}(|g_{32}|^2 - |g_{42}|^2) \\ & \qquad = j_{11}(|g_{21}|^2 - |g_{11}|^2)\frac{|z|^2(|g_{32}|^2 - |g_{42}|^2)^2}{4} + j_{33}(|g_{32}|^2 - |g_{42}|^2) \\ & \qquad = j_{33}(|g_{32}|^2 - |g_{42}|^2) \quad . \end{aligned}$$

Zbog toga traženo svojstvo, očito vrijedi.

(iii) Označimo sa

$$S^*JS = \begin{bmatrix} s_{11} & s_{12} \\ \bar{s}_{12} & s_{22} \end{bmatrix} \quad .$$

U slučaju matrice S, kad je $|g_{11}| \neq |g_{21}|$ i $|g_{32}| \neq |g_{42}|$, množenjem i pojednostavljivanjem izraza dobivamo

$$\begin{aligned} s_{11} &= -4j_{11}(|g_{21}|^2 - |g_{11}|^2) \\ s_{12} &= 4j_{11}(\bar{g}_{11}g_{12} - \bar{g}_{21}g_{22}) \\ s_{22} &= 4j_{11}(|g_{12}|^2 - |g_{21}|^2) + j_{33}(|g_{32}|^2 - |g_{42}|^2) + \frac{j_{11}(r-1)(-3r-1)}{(|g_{21}|^2 - |g_{11}|^2)|z|^2r^2} \quad . \end{aligned}$$

Računanjem determinante, dobivamo

$$\det(S^*JS) = -\frac{8(r+1)}{|z|^2 r^2} < 0 \quad .$$

U slučaju $|g_{11}| = |g_{21}|$ ili $|g_{32}| = |g_{42}|$, množenjem i pojednostavljivanjem izraza dobivamo

$$\begin{aligned} s_{11} &= -4j_{11}(|g_{21}|^2 - |g_{11}|^2) \\ s_{12} &= 4j_{11}(\bar{g}_{11}g_{12} - \bar{g}_{21}g_{22}) \\ s_{22} &= 4j_{11}(|g_{12}|^2 - |g_{21}|^2) + 3j_{33}(|g_{32}|^2 - |g_{42}|^2) \quad . \end{aligned}$$

Računanjem determinante, dobivamo

$$\det(S^*JS) = -\frac{16}{|z|^2} < 0 \quad ,$$

što pokazuje regularnost (i indefinitnost).

■

Promotrimo sad slučaj elementarnog para reda 3. Vrijedi slijedeća propozicija:

Propozicija 4.3.3.

Neka je matrica E oblika kao u relaciji (4.3.9).

(*i*) *Ako je $|g_{11}| \neq |g_{21}|$, tada je*

$$D = \begin{bmatrix} 0 & \bar{g}_{21}p \\ 0 & \bar{g}_{11}p \\ 0 & -g_{32} \end{bmatrix} \quad , \quad S = \begin{bmatrix} 2g_{11} & 2g_{12} + \bar{g}_{21}p \\ 2g_{21} & 2g_{22} + \bar{g}_{11}p \\ 0 & g_{32} \end{bmatrix} \quad , \qquad (4.3.16)$$

gdje je

$$p = \frac{r-1}{(|g_{21}|^2 - |g_{11}|^2)zr} \quad ,$$

(*ii*) *Ako je $|g_{11}| = |g_{21}|$, tada je*

$$D = \begin{bmatrix} 0 & -\bar{g}_{21}s \\ 0 & -\bar{g}_{11}s \\ 0 & -g_{32} \end{bmatrix} \quad , \quad S = \begin{bmatrix} 2g_{11} & 2g_{12} - \bar{g}_{21}s \\ 2g_{21} & 2g_{22} - \bar{g}_{11}s \\ 0 & g_{32} \end{bmatrix} \quad . \qquad (4.3.17)$$

gdje je

$$s = \frac{j_{11}j_{33}\bar{z}|g_{32}|^2}{2} \quad .$$

Dokaz:

Propozicija je jednostavna posljedica definicija matrica D, S. Uvrštavanjem matrica F_1 iz (4.3.12) i (4.3.13) i E iz (4.3.8), u definicionu relaciju (4.3.3) za D i S, dobivamo traženu tvrdnju. ■

U slijedećoj propoziciji promatramo međusobne odnose brojeva $\mathrm{rang}(D)$ i $\mathrm{rang}(D^*JD)$, te $\mathrm{rang}(S)$ i $\mathrm{rang}(S^*JS)$.

Propozicija 4.3.4.

Za matrice D i S definirane relacijama (4.3.16) *i* (4.3.17) *vrijedi*

(i)

$$\mathrm{rang}(D) = \mathrm{rang}(D^*JD) = 1 \quad .$$

(ii) *Za matricu S vrijedi*

$$\mathrm{rang}(S) = \mathrm{rang}(S^*JS) = 2 \quad .$$

Dokaz:

Dokažimo prvo relacije za $\mathrm{rang}(D)$. Zbog toga što matrice E i J_3 čine elementarni par, g_{32} ne može biti 0, pa je matrica D ranga 1.

(i) U slučaju $|g_{11}| \neq |g_{21}|$ vrijedi

$$D^*JD = \begin{bmatrix} 0 & 0 \\ 0 & j_{11}|p|^2(|g_{21}|^2 - |g_{11}|^2) + j_{33}|g_{32}|^2 \end{bmatrix} \quad .$$

Treba još pokazati da jedini netrivijalni element u toj matrici, označimo ga s c, zaista nije 0. Vrijedi

$$\begin{aligned} c &= j_{11}\frac{(r-1)^2}{(|g_{21}|^2 - |g_{11}|^2)|z|^2 r^2)} + j_{33}|g_{32}|^2 \\ &= \frac{j_{11}}{(|g_{21}|^2 - |g_{11}|^2)|z|^2}\left(\frac{(r-1)^2}{r^2} + j_{11}j_{33}|z|^2(|g_{21}|^2 - |g_{11}|^2)|g_{32}|^2\right) \\ &= \frac{j_{11}}{(|g_{21}|^2 - |g_{11}|^2)|z|^2}\left((1-\sqrt{1+a})^2 + a\right) \\ &= \frac{j_{11}}{(|g_{21}|^2 - |g_{11}|^2)|z|^2}\left(2 - 2\sqrt{1+a} + 2a\right) \\ &= \frac{j_{11}}{(|g_{21}|^2 - |g_{11}|^2)|z|^2} 2\left(\sqrt{1+a}(\sqrt{1+a} - 1)\right) \quad . \end{aligned}$$

Odatle je, zbog $a \neq 0$ i $a > -1$, $\mathrm{rang}(D^*JD) = 1$.

U slučaju $|g_{11}| = |g_{21}|$, vrijedi

$$D^*JD = \begin{bmatrix} 0 & 0 \\ 0 & j_{11}|s|^2(|g_{21}|^2 - |g_{11}|^2) + j_{33}|g_{32}|^2 \end{bmatrix} \quad .$$

Uočimo da je

$$\begin{aligned} j_{11}|s|^2(|g_{21}|^2 - |g_{11}|^2) + j_{33}|g_{32}|^2 &= j_{11}(|g_{21}|^2 - |g_{11}|^2)\frac{|z|^2|g_{32}|^4}{4} + j_{33}|g_{32}|^2 \\ &= j_{33}|g_{32}|^2 \quad , \end{aligned}$$

zbog čega vrijedi traženo svojstvo.

(ii) Kao i prije, označimo

$$S^*JS = \begin{bmatrix} s_{11} & s_{12} \\ \bar{s}_{12} & s_{22} \end{bmatrix} \quad .$$

U slučaju matrice S, kad je $|g_{11}| \neq |g_{21}|$ množenjem i pojednostavljivanjem izraza dobivamo

$$\begin{aligned} s_{11} &= -4j_{11}(|g_{21}|^2 - |g_{11}|^2) \\ s_{12} &= 4j_{11}(\bar{g}_{11}g_{12} - \bar{g}_{21}g_{22}) \\ s_{22} &= 4j_{11}(|g_{12}|^2 - |g_{21}|^2) + j_{33}|g_{32}|^2 + \frac{j_{11}(r-1)(-3r-1)}{(|g_{21}|^2 - |g_{11}|^2)|z|^2r^2} \quad . \end{aligned}$$

Računanjem determinante, dobivamo

$$\det(S^*JS) = -\frac{8(r+1)}{|z|^2r^2} < 0 \quad .$$

U slučaju $|g_{11}| = |g_{21}|$, množenjem i pojednostavljivanjem izraza dobivamo

$$\begin{aligned} s_{11} &= -4j_{11}(|g_{21}|^2 - |g_{11}|^2) \\ s_{12} &= 4j_{11}(\bar{g}_{11}g_{12} - \bar{g}_{21}g_{22}) \\ s_{22} &= 4j_{11}(|g_{12}|^2 - |g_{21}|^2) + 3j_{33}|g_{32}|^2 \quad . \end{aligned}$$

Računanjem determinante, dobivamo

$$\det(S^*JS) = -\frac{16}{|z|^2} < 0 \quad ,$$

što pokazuje regularnost i indefinitnost matrice S^*JS.

■

Vidimo da je uvijek moguće konstruirati blok J–Householderovu matricu korištenjem pridružene matrice S. Rezultat poništavanja bloka bit će

$$H(S)E = -F \quad ,$$

tj. znakovi u F bit će suprotni onima iz realizacije dekompozicije rotacijama. Uočimo da je konstrukcija blok J–Householderovih matrica uz pomoć pridružene matrice D nemoguća u slučaju elementarnog para E, J_4 reda 4, ako vrijedi $|g_{32}| = |g_{42}|$ (propozicija 4.3.2., (ii)).

5. Analiza grešaka zaokruživanja i perturbacijska analiza

Analizu grešaka zaokruživanja u indefinitnoj QR dekompoziciji provodimo u tri koraka. Prvi korak je analiza grešaka zaokruživanja jedne realne trigonometrijske rotacije. Ovu analizu prezentirao je Wilkinson u [53], a pregled tih rezulata ovdje navodimo samo zbog potpunosti. Slijedeći korak je analiza grešaka zaokruživanja za jednu realnu hiperbolnu rotaciju, uz dodatni uvjet, kad ćemo umjesto hiperbolne rotacije koristiti blok J–rotaciju. Treći korak je analiza grešaka zaokruživanja u slučaju kad koristimo elementarnu blok J–rotaciju.

Posljednja dva koraka analize do sada nisu bila napravljena. Naime, analizu grešaka zaokruživanja hiperbolnih transformacija nemoguće je napraviti bez ograde na veličinu hiperbolnog tangensa, tj. na "kut" transformacije.

Na početku, ocijenimo greške zaokruživanja u slučaju realnih matrica G. Analiza grešaka zaokruživanja nastala indefinitnom QR dekompozicijom kompleksnih matrica vrlo je slična analizi za realne matrice. Dodatno, potrebno je analizirati greške nastale množenjem matrice dijagonalnom matricom Ψ definiranom u (3.1.1), tako da svi elementi u aktivnom dijelu odgovarajućeg stupca postanu realni, pozitivni brojevi. Nakon toga, primijenjene rotacije su realne, pa možemo jednostavnije ocijeniti pogrešku.

5.1. Greške zaokruživanja za realne trigonometrijske rotacije

Neka je $U_G(i,k)$ realna ravninska rotacija definirana relacijom (3.2.5). Označimo njenu 2×2 restrikciju u (i,k) ravnini s U. Zbog kratkoće pisanja, njene elemente označimo sa s i c, što odgovara, redom, sinusu i kosinusu kuta transformacije. Dakle, imamo

$$U^{-1} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \quad , \quad \begin{aligned} c &= \cos\varphi \\ s &= \sin\varphi \end{aligned} \quad . \tag{5.1.1}$$

Želimo li poništiti element g_{k1} u prvom stupcu matrice G, za poništavanje koristimo inverz trigonometrijske rotacije $U_G(i,k)$, tj. vrijedi

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} g_{i1} \\ g_{k1} \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix} \quad ,$$

što vodi na rješavanje jednadžbe

$$sg_{i1} + cg_{k1} = 0 \quad .$$

Odmah dobivamo

$$t = -\frac{g_{k1}}{g_{i1}} \quad , \tag{5.1.2}$$

gdje simbol t predstavlja tangens kuta. Poznata je relacija koja veže tangens i kosinus kuta:

$$c = \pm\frac{1}{\sqrt{1+t^2}} \quad .$$

Uvijek možemo izabrati da kosinus ima znak elementa g_{i1}. Uvrštavanjem t iz (5.1.2) u prethodnu relaciju dobivamo

$$c = \frac{g_{i1}}{\sqrt{g_{i1}^2 + g_{k1}^2}} \quad . \tag{5.1.3}$$

Za sinus vrijedi

$$s = ct = -\frac{g_{k1}}{\sqrt{g_{i1}^2 + g_{k1}^2}} \quad . \tag{5.1.4}$$

Za greške aritmetike pretpostavljamo IEEE standard, tj.

$$f\ell(a \circ b) = (a \circ b)(1 + \varepsilon_\circ) \quad , \quad |\varepsilon_\circ| \le \varepsilon \quad ,$$

gdje $\circ$ predstavlja bilo koju od četiri osnovne računske operacije, a relativna greška $\varepsilon_\circ$ ovisi o operandima a i b i operaciji $\circ$. Jednako tako, pretpostavimo da vrijedi

$$f\ell(\sqrt{a}) = \sqrt{a}(1 + \varepsilon_{\sqrt{a}}) \quad , \quad |\varepsilon_{\sqrt{a}}| \le \varepsilon \quad .$$

Formule (5.1.3) i (5.1.4) polazne su formule za analizu grešaka zaokruživanja. Uočimo da izračunavanje nazivnika tih relacija možemo obaviti na dva načina. Ako želimo kontrolirati preljev (overflow), nazivnik bismo mogli računati na slijedeći način:

$$\sqrt{x^2+y^2} = \begin{cases} |x|\sqrt{1+(y/x)^2} & \text{za } |x| \ge |y| \\ |y|\sqrt{1+(x/y)^2} & \text{za } |y| > |x| \end{cases} \quad .$$

Promatramo li relativnu grešku, ta će formula dati dvostruko veću relativnu grešku, nego da računamo direktno. Osim toga, ako su stupci matrice G dobro skalirani, onda je ta kontrola nepotrebna. Vrijedi

$$\begin{aligned} f\ell(g_{i1}^2 + g_{k1}^2) &= \Big((1+\varepsilon_1)g_{i1}^2 + (1+\varepsilon_2)g_{k1}^2\Big)(1+\varepsilon_3) \\ &= (g_{i1}^2 + g_{k1}^2)\left(1 + \frac{\varepsilon_1 g_{i1}^2 + \varepsilon_2 g_{k1}^2 + \varepsilon_3 g_{i1}^2 + \varepsilon_3 g_{k1}^2 + \varepsilon_1\varepsilon_3 g_{i1}^2 + \varepsilon_2\varepsilon_3 g_{k1}^2}{g_{i1}^2 + g_{k1}^2}\right) \\ &= (g_{i1}^2 + g_{k1}^2)(1+\varepsilon_4) \quad , \end{aligned} \tag{5.1.5}$$

pri čemu je

$$-2\varepsilon + \varepsilon^2 \leq \varepsilon_4 \leq 2\varepsilon + \varepsilon^2 \quad ,$$

odnosno

$$(1-\varepsilon)^2 \leq 1 + \varepsilon_4 \leq (1+\varepsilon)^2 \quad . \tag{5.1.6}$$

Izraženo jednostavnije, prethodna relacija glasi

$$|\varepsilon_4| \leq 2\varepsilon + \varepsilon^2 \quad .$$

Nazivnik u relacijama (5.1.3) i (5.1.4) je drugi korijen izraza iz (5.1.5)

$$\begin{aligned} f\ell(\sqrt{g_{i1}^2 + g_{k1}^2}) &= (1+\varepsilon_5)\sqrt{f\ell(g_{i1}^2 + g_{k1}^2)} \\ &= (1+\varepsilon_5)\sqrt{(1+\varepsilon_4)(g_{i1}^2 + g_{k1}^2)} \\ &= (1+\varepsilon_7)\sqrt{g_{i1}^2 + g_{k1}^2} \quad . \end{aligned}$$

Iz (5.1.6) slijedi

$$1 + \varepsilon_6 = \sqrt{1+\varepsilon_4} \quad , \quad |\varepsilon_6| \leq \varepsilon \quad ,$$

pa prema definiciji ε_7

$$1 + \varepsilon_7 = (1+\varepsilon_5)(1+\varepsilon_6) \quad ,$$

izlazi

$$-2\varepsilon + \varepsilon^2 \leq \varepsilon_7 \leq 2\varepsilon + \varepsilon^2 \quad ,$$

odnosno

$$(1-\varepsilon)^2 \leq 1 + \varepsilon_7 \leq (1+\varepsilon)^2 \quad . \tag{5.1.7}$$

Sada preostaje samo još analiza posljednjeg dijeljenja:

$$\begin{aligned} f\ell\left(\frac{g_{\ell 1}}{\sqrt{g_{i1}^2 + g_{k1}^2}}\right) &= (1+\varepsilon_8)\frac{g_{\ell 1}}{f\ell(\sqrt{g_{i1}^2 + g_{k1}^2})} \\ &= \frac{1+\varepsilon_8}{1+\varepsilon_7}\frac{g_{\ell 1}}{\sqrt{g_{i1}^2 + g_{k1}^2}} \\ &= (1+\varepsilon_9)\frac{g_{\ell 1}}{\sqrt{g_{i1}^2 + g_{k1}^2}} \quad , \end{aligned}$$

gdje $g_{\ell 1}$ može biti ili g_{i1} ili $-g_{k1}$. Korištenjem (5.1.7), izlazi

$$\frac{1-\varepsilon}{(1+\varepsilon)^2} \leq \frac{1+\varepsilon_8}{1+\varepsilon_7} = 1+\varepsilon_9 \leq \frac{1+\varepsilon}{(1-\varepsilon)^2} \quad . \tag{5.1.8}$$

Relaciju (5.1.8) je korisno linearizirati. Razvoj za $(1+x)^{-2}$ glasi

$$(1+x)^{-2} = 1 + \binom{-2}{1}x + \binom{-2}{2}x^2 + \cdots + \binom{-2}{n}x^n + \cdots \quad ,$$

gdje je

$$\binom{-2}{k} = \frac{(-2-k+1)(-1-k+2)\cdots(-2)}{k!} = (-1)^k(k+1) \quad .$$

Uvrštavanjem odgovarajućih koeficijenata izlazi

$$(1+x)^{-2} = 1 - 2x + 3x^2 - \cdots + (-1)^n(n+1)x^n + \cdots \quad . \tag{5.1.9}$$

Za $x = -\varepsilon$ u (5.1.9), dobivamo

$$(1-\varepsilon)^{-2} = 1 + 2\varepsilon + \mathcal{O}(\varepsilon^2) \quad ,$$

dok za $x = \varepsilon$ imamo

$$(1+\varepsilon)^{-2} = 1 - 2\varepsilon + \mathcal{O}(\varepsilon^2) \quad .$$

Kao posljedica prethodne dvije relacije, linearizirani oblik (5.1.8) je

$$1 - 3\varepsilon + \mathcal{O}(\varepsilon^2) \leq 1 + \varepsilon_9 \leq 1 + 3\varepsilon + \mathcal{O}(\varepsilon^2) \quad ,$$

odnosno

$$|\varepsilon_9| \leq 3\varepsilon + \mathcal{O}(\varepsilon^2) \quad . \tag{5.1.10}$$

Korištenjem (5.1.10) izlazi

$$fl(c) = (1+\varepsilon_c)\frac{g_{i1}}{\sqrt{g_{i1}^2 + g_{k1}^2}} = (1+\varepsilon_c)c = c + \delta c \quad , \quad |\varepsilon_c| \leq 3\varepsilon + \mathcal{O}(\varepsilon^2) \tag{5.1.11}$$

$$fl(s) = (1+\varepsilon_s)\frac{-g_{k1}}{\sqrt{g_{i1}^2 + g_{k1}^2}} = (1+\varepsilon_s)s = s + \delta s \quad , \quad |\varepsilon_s| \leq 3\varepsilon + \mathcal{O}(\varepsilon^2) \tag{5.1.12}$$

Promotrimo kako greške u izračunavanju elemenata unitarne matrice utječu na greške pri primjeni te transformacije na g_ℓ, ℓ–ti stupac matrice G. Neka je U^{-1} egzaktna unitarna matrica definirana relacijom (5.1.1) i neka je transformacija provedena egzaktno (bez grešaka zaokruživanja). Jedini elementi ℓ–tog stupca koji se mijenjaju su elementi u i–tom i k–tom retku. Tada vrijedi

$$\begin{bmatrix} g'_{i\ell} \\ g'_{k\ell} \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} g_{i\ell} \\ g_{k\ell} \end{bmatrix} \quad .$$

Ako u razmatranje uključimo greške zaokruživanja, umjesto elemenata matrice U^{-1} izračunali smo elemante matrice $\widehat{U}^{-1}$, pri čemu je

$$\widehat{U}^{-1} = \begin{bmatrix} c+\delta c & -s-\delta s \\ s+\delta s & c+\delta c \end{bmatrix} \quad .$$

Izračunate vrijednosti elemenata u stupcu g_ℓ označimo s $f\ell(\ )$, pri čemu smatramo da se u slijedećoj relaciji sve operacije izvode u aritmetici pomičnog zareza.

$$\begin{bmatrix} f\ell(g'_{i\ell}) \\ f\ell(g'_{k\ell}) \end{bmatrix} = \begin{bmatrix} c+\delta c & -s-\delta s \\ s+\delta s & c+\delta c \end{bmatrix} \begin{bmatrix} g_{i\ell} \\ g_{k\ell} \end{bmatrix} \quad . \tag{5.1.13}$$

Množenjem desne strane u prethodnoj relaciji dobivamo

$$\begin{bmatrix} f\ell(g'_{i\ell}) \\ f\ell(g'_{k\ell}) \end{bmatrix} = \begin{bmatrix} ((c+\delta c)(1+\varepsilon_1)g_{i\ell} - (s+\delta s)(1+\varepsilon_2)g_{k\ell})\,(1+\varepsilon_3) \\ ((s+\delta s)(1+\varepsilon_4)g_{i\ell} + (c+\delta c)(1+\varepsilon_5)g_{k\ell})\,(1+\varepsilon_6) \end{bmatrix} \quad .$$

Pri tome su greške u izračunatim elementima (do na $\mathcal{O}(\varepsilon^2)$) jednake

$$\begin{bmatrix} f\ell(g'_{i\ell}) - g'_{i\ell} \\ f\ell(g'_{k\ell}) - g'_{k\ell} \end{bmatrix} = \begin{bmatrix} \delta c & -\delta s \\ \delta s & \delta c \end{bmatrix} \begin{bmatrix} g_{i\ell} \\ g_{k\ell} \end{bmatrix} + \begin{bmatrix} h_{i\ell} \\ h_{k\ell} \end{bmatrix} \quad , \tag{5.1.14}$$

pri čemu je

$$\begin{bmatrix} h_{i\ell} \\ h_{k\ell} \end{bmatrix} = \begin{bmatrix} (\varepsilon_1+\varepsilon_3)cg_{i\ell} - (\varepsilon_2+\varepsilon_3)sg_{k\ell} + \mathcal{O}(\varepsilon^2) \\ (\varepsilon_4+\varepsilon_6)sg_{i\ell} + (\varepsilon_5+\varepsilon_6)cg_{k\ell} + \mathcal{O}(\varepsilon^2) \end{bmatrix} \quad .$$

Uočimo da je 2–norma prvog sumanda s desne strane u (5.1.14) manja ili jednaka 2–normi matrice pomnoženoj s 2–normom vektora. Označimo tu matricu s δU^{-1}. Vrijedi

$$\|\delta U^{-1}\|_2 = \sqrt{\lambda_{\max}((\delta U^{-1})^* \delta U^{-1})}$$

pri čemu je $\lambda_{\max}$ najveća svojstvena vrijednost matrice $(\delta U^{-1})^* \delta U^{-1}$. Jednostavnim računom dobivamo

$$(\delta U^{-1})^* \delta U^{-1} = \begin{bmatrix} (\delta c)^2 + (\delta s)^2 & 0 \\ 0 & (\delta c)^2 + (\delta s)^2 \end{bmatrix} \quad .$$

Odavde je očito da su obje svojstvene vrijednosti ove matrice jednake $(\delta c)^2 + (\delta s)^2$.

Za prvi član s desne strane, korištenjem (5.1.12), izlazi ocjena

$$\begin{aligned} \left\| \delta U^{-1} \begin{bmatrix} g_{i\ell} \\ g_{k\ell} \end{bmatrix} \right\|_2 &\leq \|\delta U^{-1}\|_2 \left\| \begin{bmatrix} g_{i\ell} \\ g_{k\ell} \end{bmatrix} \right\|_2 \\ &= \sqrt{(\delta c)^2 + (\delta s)^2}\,\sqrt{g_{i\ell}^2 + g_{k\ell}^2} \\ &\leq 3\varepsilon \sqrt{g_{i\ell}^2 + g_{k\ell}^2} \quad . \end{aligned} \tag{5.1.15}$$

Za drugi član s desne strane u (5.1.14) vrijedi

$$\begin{aligned}\left\| \begin{bmatrix} h_{i\ell} \\ h_{k\ell} \end{bmatrix} \right\|_2 &\leq \sqrt{((\varepsilon_1+\varepsilon_3)cg_{i\ell}-(\varepsilon_2+\varepsilon_3)sg_{k\ell})^2+((\varepsilon_4+\varepsilon_6)sg_{i\ell}+(\varepsilon_5+\varepsilon_6)cg_{k\ell})^2} \\ &\leq 2\varepsilon\sqrt{(|cg_{i\ell}|+|sg_{k\ell}|)^2+(|sg_{i\ell}|+|cg_{k\ell}|)^2} \\ &\leq 2\varepsilon\sqrt{g_{i\ell}^2+g_{k\ell}^2+4|sc|\,|g_{i\ell}|\,|g_{k\ell}|} \quad .\end{aligned}$$

Ako uočimo da je $2|sc| \leq 1$ i da vrijedi

$$2|g_{i\ell}|\,|g_{k\ell}| \leq g_{i\ell}^2+g_{k\ell}^2 \quad , \tag{5.1.16}$$

onda posljednju relaciju možemo ocijeniti s

$$\left\| \begin{bmatrix} h_{i\ell} \\ h_{k\ell} \end{bmatrix} \right\|_2 \leq 2\sqrt{2}\varepsilon\sqrt{g_{i\ell}^2+g_{k\ell}^2} \quad . \tag{5.1.17}$$

Korištenjem relacija (5.1.15) i (5.1.17), dobivamo

$$\begin{aligned}\sqrt{(f\ell(g'_{i\ell})-g'_{i\ell})^2+(f\ell(g'_{k\ell})-g'_{k\ell})^2} &\leq (3+2\sqrt{2})\varepsilon\sqrt{g_{i\ell}^2+g_{k\ell}^2} \\ &= e_G\sqrt{g_{i\ell}^2+g_{k\ell}^2} \quad ,\end{aligned} \tag{5.1.18}$$

gdje je $e_G = 5.83\varepsilon$.

Na kraju možemo, slično kao u [21], zaključiti da u prvom stupcu matrice G možemo nezavisno provoditi transformacije redaka $\{i,j\}$ i $\{k,\ell\}$, ako i samo ako je $\{i,j\} \cap \{k,\ell\} = \emptyset$, a odgovarajući elementi matrice J dozvoljavaju trigonometrijske transformacije. Drugim riječima, možemo primjenivati nizove nezavisnih rotacija U_G. Ocjena pogreške za maksimalan broj takvih nezavisnih transformacija zadržava oblik (5.1.18), tj.

$$\|f\ell(g'_\ell)-g'_\ell\|_2 \leq e_G\|g_\ell\|_2 \quad .$$

Jasno je da broj nezavisnih transformacija ovisi o broju elemenata u matrici J s pozitivnim i negativnim znakovima i broju nula u odgovarajućem stupcu. Neka je broj pozitivnih elemenata u matrici J koji odgovaraju ne-nula elementima u matrici G jednak m_1, a broj odgovarajućih negativnih m_2, pri čemu je $m \geq m_1+m_2$. Označimo sa $s = \max\{m_1, m_2\}$. U prvom koraku imamo $m_1/2+m_2/2$ nezavisnih rotacija, u drugom $m_1/4+m_2/4$, sve dok u posljednjem koraku nemamo samo jednu ili eventualno dvije nezavisne rotacije koje odgovaraju različitim znakovima u J. Pitanje je koliko nam treba koraka da prvi stupac matrice G svedemo na samo dva elementa (ili jedan, ako je J svagdje istog znaka).

Rotacije koje koristimo za poništavanja elemenata za različite znakove u J su međusobno nezavisne. Zbog toga je dovoljno promatrati koliko koraka treba za poništavanje $s-1$ elemenata, a sigurno je da broj potrebnih koraka raste s porastom s. Zbog toga promatramo broj koraka potrebnih za poništavanje $2^p - 1$, $p \in \mathbf{N}$ elemenata, gdje je

$$p = \lceil \lg s \rceil = \lceil \log_2 s \rceil \quad , \tag{5.1.19}$$

tj. 2^p je najmanja potencija broja 2 veća ili jednaka s. Nakon prvog koraka ostaje neponištenih 2^{p-1} elemenata, nakon drugog 2^{p-2} elemenata, sve dok u posljednjem koraku ne ostane samo jedan neponišteni element. Dakle, dubina binarnog stabla s 2^{p-1} listova jednaka je $p = \lceil \lg s \rceil$.

Promatrajmo greške zaokruživanja u ℓ–tom stupcu nakon primjene p nizova nezavisnih trigonometrijskih rotacija U_G. Skraćeno označimo te nizove rotacija indeksima koraka, tj. niz rotacija primijenjen u k–tom koraku označimo s U_k.

Neka je $f\ell(g_\ell')$ ℓ–ti stupac matrice G dobiven nakon primjene prvog niza nezavisnih rotacija, tj. nakon prvog koraka algoritma. Relacija (5.1.13) povlači

$$f\ell(g_\ell') = f\ell(\widehat{U}_1^{-1} g_\ell) \quad .$$

Desnu stranu prethodne relacije interpretirajmo kao malo pogrešnu unitarnu matricu $\widetilde{U}_1^{-1}$ egzaktno primijenjenu na stupac g_ℓ, tj. vrijedi

$$f\ell(\widehat{U}_1^{-1} g_\ell) = \widetilde{U}_1^{-1} g_\ell \quad . \tag{5.1.20}$$

Relacija (5.1.18) povlači

$$\begin{aligned}
\|\widetilde{U}_1^{-1} g_\ell - U_1^{-1} g_\ell\|_2 &\le e_G \|g_\ell\|_2 \\
&= e_G \|U_1^{-1} g_\ell\|_2 \\
\|\widetilde{U}_2^{-1}\widetilde{U}_1^{-1} g_\ell - U_2^{-1}\widetilde{U}_1^{-1} g_\ell\|_2 &\le e_G \|\widetilde{U}_1^{-1} g_\ell\|_2 \\
&= e_G \|U_2^{-1}\widetilde{U}_1^{-1} g_\ell\|_2 \\
\vdots &\le \vdots \\
\|\widetilde{U}_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1} g_\ell - U_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1} g_\ell\|_2 &\le e_G \|\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1} g_\ell\|_2 \\
&= e_G \|U_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1} g_\ell\|_2 \quad .
\end{aligned}$$

Izrazimo desne strane pomoću norme ℓ–tog stupca prije početka prve transformacije. Za desnu stranu druge nejednakosti vrijedi

$$\begin{aligned}
\|U_2^{-1}\widetilde{U}_1^{-1} g_\ell\|_2 &= \|\widetilde{U}_1^{-1} g_\ell\|_2 = \|\widetilde{U}_1^{-1} g_\ell - U_1^{-1} g_\ell + U_1^{-1} g_\ell\|_2 \\
&\le \|\widetilde{U}_1^{-1} g_\ell - U_1^{-1} g_\ell\|_2 + \|U_1^{-1} g_\ell\|_2 \le (1 + e_G)\|g_\ell\|_2 \quad .
\end{aligned}$$

Primjenom, redom, od treće do p–te nejednakosti, slijedi

$$\begin{aligned}
\|\widetilde{U}_1^{-1} g_\ell - U_1^{-1} g_\ell\|_2 &\le e_G \|g_\ell\|_2 \\
\|\widetilde{U}_2^{-1}\widetilde{U}_1^{-1} g_\ell - U_2^{-1}\widetilde{U}_1^{-1} g_\ell\|_2 &\le e_G (1 + e_G)\|g_\ell\|_2 \\
\vdots &\le \vdots \\
\|\widetilde{U}_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1} g_\ell - U_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1} g_\ell\|_2 &\le e_G (1 + e_G)^{p-1}\|g_\ell\|_2 \quad .
\end{aligned}$$

Primjenom nejednakosti trokuta dobivamo ukupnu pogrešku za ℓ–ti stupac nakon p koraka

$$\begin{aligned}
\|\widetilde{U}_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell \,-\, & U_p^{-1}U_{p-1}^{-1}\cdots U_1^{-1}g_\ell\|_2 \\
&= \|\widetilde{U}_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell - U_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell \\
&\quad + U_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell - U_p^{-1}U_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell \\
&\quad + \cdots + U_p^{-1}U_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell - U_p^{-1}U_{p-1}^{-1}\cdots U_1^{-1}g_\ell\|_2 \\
&\le \|\widetilde{U}_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell - U_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell\|_2 \\
&\quad + \|U_p^{-1}\widetilde{U}_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell - U_p^{-1}U_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell\|_2 \\
&\quad + \cdots + \|U_p^{-1}U_{p-1}^{-1}\cdots\widetilde{U}_1^{-1}g_\ell - U_p^{-1}U_{p-1}^{-1}\cdots U_1^{-1}g_\ell\|_2 \\
&\le e_G(1 + (1+e_G) + \cdots + (1+e_G)^{p-1})\|g_\ell\|_2 \\
&= ((1+e_G)^p - 1)\|g_\ell\|_2 \quad .
\end{aligned}$$

Time smo dokazali slijedeću lemu:

Lema 5.1.1.

Neka su zadane matrice $G \in \mathbf{R}^{m\times n}$ i $J = \operatorname{diag}(j_{11},\ldots,j_{mm})$, $j_{ii} \in \{1,-1\}$ i neka je p definiran relacijom (5.1.19). Nizom nezavisnih trigonometrijskih rotacija $U_1^{-1},\ldots,U_p^{-1}$, poništimo sve elemente prvog stupca matrice G, osim, najviše jednog koji odgovara pozitivnom i jednog koji odgovara negativnom znaku u matrici J. Ako su sve transformacije izračunate u aritmetici pomičnog zareza, za ℓ–ti stupac izračunate matrice G' vrijedi

$$\|fl(U_p^{-1}\cdots U_1^{-1}g_\ell) - U_p^{-1}\cdots U_1^{-1}g_\ell\|_2 \le ((1+e_G)^p - 1)\|g_\ell\|_2 \quad . \tag{5.1.21}$$

■

Zanimljivo je i pitanje, koliko se promijenila J–norma nekog stupca nakon niza ovakvih rotacija, tj. koliko se promijenio odgovarajući element u matrici G^*JG.

Promatrajmo odvojeno dio stupca g_ℓ^+, koji odgovara $j_{ii} = 1$, a posebno dio g_ℓ^- koji odgovara $j_{ii} = -1$. Označimo još i

$$p_1 = \lceil \lg m_1 \rceil \quad , \quad p_2 = \lceil \lg m_2 \rceil \quad ,$$

gdje su m_1 i m_2, redom, broj ne-nula elemenata u odgovarajućem stupcu koji odgovara plus (minus) znaku u matrici J.

Ako istovremeno primjenjujemo nizove trigonometrijskih rotacija V_i na g_1^+ i W_i na g_1^-, ocjene pogreške imat će isti oblik kao (5.1.21), tj. vrijedi

$$\begin{aligned}
\|\widetilde{V}_{p_1}^{-1}\widetilde{V}_{p_1-1}^{-1}\cdots\widetilde{V}_1^{-1}g_\ell^+ - V_{p_1}^{-1}V_{p_1-1}^{-1}\cdots V_1^{-1}g_\ell^+\|_2 &\le ((1+e_G)^{p_1} - 1)\|g_\ell^+\|_2 \quad , \\
\|\widetilde{W}_{p_2}^{-1}\widetilde{W}_{p_2-1}^{-1}\cdots\widetilde{W}_1^{-1}g_\ell^- - W_{p_2}^{-1}W_{p_2-1}^{-1}\cdots W_1^{-1}g_\ell^-\|_2 &\le ((1+e_G)^{p_2} - 1)\|g_\ell^-\|_2 \quad .
\end{aligned}$$

Iz prethodne relacije izvedimo ocjenu pogreške za perturbaciju J–norme. Vrijedi

$$\begin{aligned}
&\| \widetilde{V}_{p_1}^{-1} \cdots \widetilde{V}_1^{-1} g_\ell^+ - V_{p_1}^{-1} \cdots V_1^{-1} g_\ell^+ \|_2^2 - \| \widetilde{W}_{p_2}^{-1} \cdots \widetilde{W}_1^{-1} g_\ell^- - W_{p_2}^{-1} \cdots W_1^{-1} g_\ell^- \|_2^2 \,| \\
&\qquad \leq \max\{((1+e_G)^{p_1} - 1)^2 \|g_\ell^+\|_2^2, ((1+e_G)^{p_2} - 1)^2 \|g_\ell^-\|_2^2\} \\
&\qquad = \max\{\varepsilon_+^2 \|g_\ell^+\|_2^2, \varepsilon_-^2 \|g_\ell^-\|_2^2\} \quad . \qquad (5.1.22)
\end{aligned}$$

Nejednakost (5.1.22) pokazuje da relativna pogreška može biti velika, u slučaju da je J–norma blizu 0, ali je apsolutna pogreška mala, ako brojevi $\|g_\ell^+\|_2$ i $\|g_\ell^-\|_2$ nisu preveliki. Jednako tako je jasno da se ovaj rezultat ne može poboljšati, jer vektor J–norme 0 ne mora imati sve komponente jednake 0. Malom perturbacijom bilo koje od tih komponenti dolazi do perturbacije J–norme 0.

Mnogo je povoljnija situacija ako znamo da niti jedan stupac (odnosno dio stupca) matrice neće tokom transformacija poprimiti 0 J–normu. Takve su, na primjer, skalirano dijagonalno dominantne matrice.

Ako su norme $\|g_\ell^+\|_2$ i $\|g_\ell^-\|_2$ približno jednake, onda relacija (5.1.22) glasi

$$\| f\ell(g_\ell') - g_\ell' \|_J \leq \max\{\varepsilon_+, \varepsilon_-\} \|g_\ell^+\|_2 \quad .$$

Bez smanjenja općenitosti, možemo pretpostaviti da je matrica $G = \widetilde{G}D$ zdesna skalirana tako da je

$$\max\{\|g_\ell^+\|_2, \|g_\ell^-\|_2\} = 1 \quad . \qquad (5.1.23)$$

QR (ili indefinitna QR) dekompozicija matrice G jednaka je odgovarajućoj dekompoziciji matrice $\widetilde{G}$, pomnoženoj dijagonalnom matricom D.

Uz takvo skaliranje (zovimo ga standardno indefinitno skaliranje), vrijedi ocjena

$$\| f\ell(g_\ell') - g_\ell' \|_J \leq \max\{\varepsilon_+, \varepsilon_-\} \quad . \qquad (5.1.24)$$

Ako možemo odozdo ograditi J–normu stupca, može se pokazati da će i relativna pogreška u odgovarajućem stupcu biti mala. Neka vrijedi

$$\|g_\ell^+\|_2 \leq \alpha \|g_\ell^-\|_2 \quad , \quad \alpha < 1 \quad . \qquad (5.1.25)$$

Tada je

$$\begin{aligned}
\frac{\max\{\varepsilon_+^2 \|g_\ell^+\|_2^2, \varepsilon_-^2 \|g_\ell^-\|_2^2\}}{|\, \|g_\ell^+\|_2^2 - \|g_\ell^-\|_2^2 \,|} &\leq \frac{\max\{\varepsilon_+^2 \alpha^2, \varepsilon_-^2\} \|g_\ell^-\|_2^2\}}{(1-\alpha^2) \|g_\ell^-\|_2^2} \qquad (5.1.26) \\
&\leq \frac{\max\{\varepsilon_+^2 \alpha^2, \varepsilon_-^2\}\}}{1-\alpha^2} \quad .
\end{aligned}$$

Ako vrijedi suprotno, tj. ako je

$$\|g_\ell^-\|_2 \leq \alpha \|g_\ell^+\|_2 \quad , \quad \alpha < 1 \quad , \qquad (5.1.27)$$

tada je

$$\begin{aligned}\frac{\max\{\varepsilon_+^2\|g_\ell^+\|_2^2, \varepsilon_-^2\|g_\ell^-\|_2^2\}}{|\,\|g_\ell^+\|_2^2 - \|g_\ell^-\|_2^2\,|} &\leq \frac{\max\{\varepsilon_-^2\alpha^2, \varepsilon_+^2\}\|g_\ell^+\|_2^2}{(1-\alpha^2)\|g_\ell^+\|_2^2} \\ &\leq \frac{\max\{\varepsilon_-^2\alpha^2, \varepsilon_+^2\}\}}{1-\alpha^2} \quad .\end{aligned} \tag{5.1.28}$$

Uvrštavanjem relacija (5.1.26) i (5.1.28) u (5.1.22) dobivamo

$$\|f\ell(g_\ell') - g_\ell'\|_J \leq \frac{\max\{\varepsilon_+, \varepsilon_-\}}{1-\alpha^2}\|g_\ell\|_J \quad , \tag{5.1.29}$$

odnosno relativna pogreška ovisi o udaljenosti J–norme od 0.

5.2. Greške zaokruživanja za realne hiperbolne rotacije

U prošlom odjeljku smo analizirali greške zaokruživanja nastale "skraćivanjem" prvog stupca matrice G na najviše dva elementa koji odgovaraju različitim predznacima u matrici J. Ako su oba elementa netrivijalna, permutacijama redaka dovedimo ih na prva dva mjesta u prvom stupcu, a zatim poništimo primjenom jedne hiperbolne transforamcije.

Neka je $U_H(1,2)$ realna hiperbolna rotacija definirana relacijom (3.2.9) obzirom na J–skalarni produkt. Označimo s U njenu 2×2 restrikciju u $(1,2)$ ravnini, koja odgovara restrikciji $J_1 = \operatorname{diag}(j_{11}, -j_{11})$, $j_{11} \in \{1,-1\}$ matrice J. Zbog kratkoće pisanja, njene elemente označimo sa s i c, što odgovara, redom, sinusu i kosinusu hiperbolnom kuta transformacije. Dakle, imamo

$$U^{-1} = \begin{bmatrix} c & -s \\ -s & c \end{bmatrix} \quad , \quad \begin{aligned} c &= \operatorname{ch}\varphi \\ s &= \operatorname{sh}\varphi \end{aligned} \quad . \tag{5.2.1}$$

Poništavanje elementa g_{21} vodi na jednadžbu oblika

$$\begin{bmatrix} c & -s \\ -s & c \end{bmatrix}\begin{bmatrix} g_{11} \\ g_{21} \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix} \quad .$$

Iz drugog retka prethodne matrične jednadžbe izlazi

$$-sg_{11} + cg_{21} = 0 \quad ,$$

odnosno

$$t = \frac{g_{21}}{g_{11}} \quad , \tag{5.2.2}$$

gdje simbol t predstavlja hiperbolni tangens kuta, uz pretpostavku da je $|t| < 1$. U slučaju $|t| > 1$ zamijenit ćemo prvi i drugi redak u matrici G i dijagonalne elemente u J_1. Poznata je relacija koja veže hiperbolni tangens i kosinus:

$$c = \frac{1}{\sqrt{1-t^2}}$$

Prema propoziciji 3.1.6., uvijek možemo odgovarajuće retke u matrici G pomnožiti dijagonalnom J–unitarnom matricom, tako da elementi g_{11} i g_{21} budu veći ili jednaki 0. Uvrštavanjem (5.2.2) u prethodnu relaciju dobivamo

$$c = \frac{g_{11}}{\sqrt{g_{11}^2 - g_{21}^2}} \quad , \quad g_{11} > g_{21} \quad . \tag{5.2.3}$$

Za hiperbolni sinus vrijedi

$$s = ct = \frac{g_{21}}{\sqrt{g_{11}^2 - g_{21}^2}} \quad , \quad g_{11} > g_{21} \quad . \tag{5.2.4}$$

Formule (5.2.3) i (5.2.4) polazne su za analizu grešaka zaokruživanja. Prvo promotrimo grešku kod računanja nazivnika u tim relacijama.

$$\begin{aligned} f\ell(g_{11}^2 - g_{21}^2) &= \Big((1+\varepsilon_1)g_{11}^2 - (1+\varepsilon_2)g_{21}^2\Big)(1+\varepsilon_3) \\ &= (g_{11}^2 - g_{21}^2)\left(1 + \frac{\varepsilon_1 g_{11}^2 - \varepsilon_2 g_{21}^2 + \varepsilon_3 g_{11}^2 - \varepsilon_3 g_{21}^2 + \varepsilon_1\varepsilon_3 g_{11}^2 - \varepsilon_2\varepsilon_3 g_{21}^2}{g_{11}^2 - g_{21}^2}\right) \\ &= (g_{11}^2 - g_{21}^2)(1+\varepsilon_4) \quad . \end{aligned} \tag{5.2.5}$$

Veličinu ε_4 možemo ograditi samo uz pretpostavku da je

$$\frac{g_{21}}{g_{11}} \le \alpha < 1 \quad . \tag{5.2.6}$$

Tada za ε_4 vrijedi

$$|\varepsilon_4| \le \frac{(2\varepsilon+\varepsilon^2)(g_{11}^2 + g_{21}^2)}{g_{11}^2 - g_{21}^2} = \frac{(2\varepsilon+\varepsilon^2)(1 + \frac{g_{21}^2}{g_{11}^2})}{1 - \frac{g_{21}^2}{g_{11}^2}} \le \frac{(2\varepsilon+\varepsilon^2)(1+\alpha^2)}{1-\alpha^2} \quad . \tag{5.2.7}$$

Ako još dodatno označimo

$$\beta = \frac{1+\alpha^2}{1-\alpha^2} \quad , \tag{5.2.8}$$

onda (5.2.7) glasi

$$|\varepsilon_4| \le (2\varepsilon+\varepsilon^2)\beta \quad . \tag{5.2.9}$$

Dakle, možemo zaključiti da vrijedi

$$1-(2\varepsilon+\varepsilon^2)\beta \le 1+\varepsilon_4 \le 1+(2\varepsilon+\varepsilon^2)\beta \le (1+\beta\varepsilon)^2 \quad . \tag{5.2.10}$$

Nazivnik u relacijama (5.2.3) i (5.2.4) je drugi korijen izraza iz (5.2.5):

$$\begin{aligned} f\ell(\sqrt{g_{11}^2-g_{21}^2}) &= (1+\varepsilon_5)\sqrt{f\ell(g_{11}^2-g_{21}^2)} \\ &= (1+\varepsilon_5)\sqrt{(1+\varepsilon_4)(g_{11}^2-g_{21}^2)} \\ &= (1+\varepsilon_7)\sqrt{g_{11}^2-g_{21}^2} \quad . \end{aligned}$$

Ako označimo

$$1+\varepsilon_6 = \sqrt{1+\varepsilon_4} \quad ,$$

iz (5.2.10) slijedi

$$|\varepsilon_6| \le \beta\varepsilon + \mathcal{O}(\beta\varepsilon^2) \quad ,$$

pa uz oznaku

$$1+\varepsilon_7 = (1+\varepsilon_5)(1+\varepsilon_6) \quad ,$$

izlazi

$$1-(\beta+1)\varepsilon+\mathcal{O}(\beta\varepsilon^2) \le 1+\varepsilon_7 \le 1+(\beta+1)\varepsilon+\mathcal{O}(\beta\varepsilon^2) \quad . \tag{5.2.11}$$

Preostaje još samo analiza posljednjeg dijeljenja.

$$\begin{aligned} f\ell\left(\frac{g_{\ell 1}}{\sqrt{g_{11}^2-g_{21}^2}}\right) &= (1+\varepsilon_8)\frac{g_{\ell 1}}{f\ell(\sqrt{g_{11}^2-g_{21}^2})} \\ &= \frac{1+\varepsilon_8}{1+\varepsilon_7}\frac{g_{\ell 1}}{\sqrt{g_{11}^2-g_{21}^2}} \\ &= (1+\varepsilon_9)\frac{g_{\ell 1}}{\sqrt{g_{11}^2-g_{21}^2}} \quad , \end{aligned}$$

gdje $g_{\ell 1}$ može biti ili g_{11} ili g_{21}.

Korištenjem (5.2.11) i zanemarivanjem članova veličine $\mathcal{O}(\beta\varepsilon^2)$, izlazi ocjena

$$\frac{1-\varepsilon}{1+(\beta+1)\varepsilon} \le \frac{1+\varepsilon_8}{1+\varepsilon_7} = 1+\varepsilon_9 \le \frac{1+\varepsilon}{1-(\beta+1)\varepsilon} \quad . \tag{5.2.12}$$

Linearizacija posljednje relacije daje

$$1-(\beta+2)\varepsilon \le 1+\varepsilon_9 \le 1+(\beta+2)\varepsilon \quad ,$$

odnosno

$$|\varepsilon_9| \le (\beta+2)\varepsilon + \mathcal{O}(\beta\varepsilon^2) \quad . \tag{5.2.13}$$

Korištenjem (5.2.3) i (5.2.4), iz (5.2.13) dobivamo

$$f\ell(c) = (1+\varepsilon_c)\frac{g_{11}}{\sqrt{g_{11}^2 - g_{21}^2}} = (1+\varepsilon_c)c \quad , \quad |\varepsilon_c| \leq (\beta+2)\varepsilon \tag{5.2.14}$$

$$f\ell(s) = (1+\varepsilon_s)\frac{g_{21}}{\sqrt{g_{11}^2 - g_{21}^2}} = (1+\varepsilon_s)s \quad , \quad |\varepsilon_s| \leq (\beta+2)\varepsilon \quad , \tag{5.2.15}$$

gdje je ocjena za $|\varepsilon_c|$ i $|\varepsilon_s|$ linearizirana i vrijedi do na član reda veličine $\mathcal{O}(\beta\varepsilon^2)$.

Kao i u trigonometrijskom slučaju, označimo

$$\delta c = \varepsilon_c c \quad , \quad \delta s = \varepsilon_s s \quad .$$

Promotrimo kako greške u računanju elemenata J–unitarne matrice utječu na greške pri primjeni te transformacije na neki stupac matrice G. Neka je U^{-1} egzaktna J_1–unitarna matrica definirana relacijom (5.2.1) i neka je transformacija provedena egzaktno (bez grešaka zaokruživanja). U ℓ–tom stupcu matrice G mijenjaju se samo elementi u prva dva retka, tj. vrijedi

$$\begin{bmatrix} g'_{1\ell} \\ g'_{2\ell} \end{bmatrix} = \begin{bmatrix} c & -s \\ -s & c \end{bmatrix} \begin{bmatrix} g_{1\ell} \\ g_{2\ell} \end{bmatrix} \quad .$$

Ako u razmatranje uključimo greške zaokruživanja, umjesto elemenata matrice U^{-1} izračunali smo elemante matrice $\widehat{U}^{-1}$, pri čemu je

$$\widehat{U}^{-1} = \begin{bmatrix} c+\delta c & -s-\delta s \\ -s-\delta s & c+\delta c \end{bmatrix} \quad .$$

Izračunate vrijednosti stupca g_ℓ označimo s $f\ell(\ )$, pri čemu smatramo da se u slijedećoj relaciji sve operacije izvode u aritmetici pomičnog zareza.

$$\begin{bmatrix} f\ell(g'_{1\ell}) \\ f\ell(g'_{2\ell}) \end{bmatrix} = \begin{bmatrix} c+\delta c & -s-\delta s \\ -s-\delta s & c+\delta c \end{bmatrix} \begin{bmatrix} g_{1\ell} \\ g_{2\ell} \end{bmatrix} \quad . \tag{5.2.16}$$

Množenjem desne strane u prethodnoj relaciji dobivamo

$$\begin{bmatrix} f\ell(g'_{1\ell}) \\ f\ell(g'_{2\ell}) \end{bmatrix} = \begin{bmatrix} ((c+\delta c)(1+\varepsilon_1)g_{1\ell} - (s+\delta s)(1+\varepsilon_2)g_{2\ell})\,(1+\varepsilon_3) \\ (-(s+\delta s)(1+\varepsilon_4)g_{1\ell} + (c+\delta c)(1+\varepsilon_5)g_{2\ell})\,(1+\varepsilon_6) \end{bmatrix} \quad .$$

Pri tome su greške u izračunatim elementima (do na $\mathcal{O}(\varepsilon^2)$) jednake

$$\begin{bmatrix} f\ell(g'_{1\ell}) - g'_{1\ell} \\ f\ell(g'_{2\ell}) - g'_{2\ell} \end{bmatrix} = \begin{bmatrix} \delta c & -\delta s \\ -\delta s & \delta c \end{bmatrix} \begin{bmatrix} g_{1\ell} \\ g_{2\ell} \end{bmatrix} + \begin{bmatrix} h_{1\ell} \\ h_{2\ell} \end{bmatrix} \quad . \tag{5.2.17}$$

pri čemu je

$$\begin{bmatrix} h_{1\ell} \\ h_{2\ell} \end{bmatrix} = \begin{bmatrix} (\varepsilon_1+\varepsilon_3)cg_{1\ell} - (\varepsilon_2+\varepsilon_3)sg_{2\ell} + \mathcal{O}(\varepsilon^2) \\ (\varepsilon_4+\varepsilon_6)sg_{1\ell} + (\varepsilon_5+\varepsilon_6)cg_{2\ell} + \mathcal{O}(\varepsilon^2) \end{bmatrix} \quad .$$

Uočimo da je 2–norma prvog sumanda s desne strane u (5.2.17) manja ili jednaka 2–normi matrice pomnoženoj s 2–normom vektora. Označimo tu matricu s δU^{-1}. Matrica δU^{-1} je simetrična, pa vrijedi

$$\|\delta U^{-1}\|_2 = |\lambda_{\max}(\delta U^{-1})|$$

pri čemu je $\lambda_{\max}$ po apsolutnoj vrijednosti najveća svojstvena vrijednost matrice δU^{-1}. Primjenom jednog koraka Jacobijeve metode s kutem $\pi/4$, možemo dijagonalizirati matricu δU^{-1}, pa za njezine svojstvene vrijednosti dobivamo

$$\lambda_{1,2} = \delta c \pm \delta s \quad .$$

Korištenjem (5.2.14) i (5.2.15) izlazi

$$|\lambda_{\max}| \leq |\delta c| + |\delta s| \leq (\beta+2)\varepsilon(|c|+|s|) \quad . \tag{5.2.18}$$

Iz (5.2.2) i nejednakosti (5.2.6) slijedi da je

$$|c|+|s| \leq \frac{1+|t|}{\sqrt{1-t^2}} \leq \frac{1+\alpha}{\sqrt{1-\alpha^2}} = \sqrt{\frac{1+\alpha}{1-\alpha}} = \gamma \quad . \tag{5.2.19}$$

Uočimo da jednako tako vrijedi

$$s^2+c^2 = \frac{1+t^2}{1-t^2} \leq \frac{1+\alpha^2}{1-\alpha^2} = \beta \quad , \tag{5.2.20}$$

i

$$|sc| = \frac{|t|}{1-t^2} \leq \frac{\alpha}{1-\alpha^2} \quad . \tag{5.2.21}$$

Za prvi član s desne strane u (5.2.17), korištenjem (5.2.18) i (5.2.19) izlazi ocjena

$$\left\| \delta U^{-1} \begin{bmatrix} g_{1\ell} \\ g_{2\ell} \end{bmatrix} \right\|_2 \leq \|\delta U^{-1}\|_2 \left\| \begin{bmatrix} g_{1\ell} \\ g_{2\ell} \end{bmatrix} \right\|_2 \leq (\beta+2)\gamma\varepsilon\sqrt{g_{1\ell}^2+g_{2\ell}^2} \quad . \tag{5.2.22}$$

Za drugi član s desne strane u (5.2.17), korištenjem (5.1.16), (5.2.20) i (5.2.21) izlazi

$$\begin{aligned}
\left\| \begin{bmatrix} h_{1\ell} \\ h_{2\ell} \end{bmatrix} \right\|_2 &\leq \sqrt{((\varepsilon_1+\varepsilon_3)cg_{1\ell} - (\varepsilon_2+\varepsilon_3)sg_{2\ell})^2 + ((\varepsilon_4+\varepsilon_6)sg_{1\ell} + (\varepsilon_5+\varepsilon_6)cg_{2\ell})^2} \\
&\leq 2\varepsilon\sqrt{(|cg_{1\ell}|+|sg_{2\ell}|)^2 + (|sg_{1\ell}|+|cg_{2\ell}|)^2} \\
&\leq 2\varepsilon\sqrt{(s^2+c^2)(g_{1\ell}^2+g_{2\ell}^2) + 4|sc|\,|g_{1\ell}|\,|g_{2\ell}|} \\
&\leq 2\varepsilon\sqrt{\left(\beta+\frac{2\alpha}{1-\alpha^2}\right)(g_{1\ell}^2+g_{2\ell}^2)} \\
&\leq 2\gamma\varepsilon\sqrt{g_{1\ell}^2+g_{2\ell}^2} \quad ,
\end{aligned}$$

što zajedno s (5.2.22) daje

$$\| f\ell(g'_\ell) - g'_\ell\|_2 \leq (\beta+4)\gamma\varepsilon\sqrt{g_{1\ell}^2 + g_{2\ell}^2} = e_H\sqrt{g_{1\ell}^2 + g_{2\ell}^2} \leq e_H\|g_\ell\|_2 \quad .$$

Time je dokazana slijedeća lema:

Lema 5.2.1.

Neka je zadana matrica $G \in \mathbf{R}^{m\times n}$ takva da je $g_{11} > g_{21} > 0$ i to su jedini elementi u prvom stupcu različiti od 0. Neka je pripadna matrica indefinitnog skalarnog produkta zadana s $J = \operatorname{diag}(j_{11}, -j_{11}, j_{33}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$. Ako hiperbolnom rotacijom $U_H^{-1}(1,2)$ izračunatom u aritmetici pomičnog zareza, poništimo element g_{21}, za ℓ–ti stupac izračunate matrice G' vrijedi

$$\| f\ell(U_H^{-1}g_\ell) - U_H^{-1}g_\ell\|_2 \leq e_H\|g_\ell\|_2 \quad . \tag{5.2.23}$$

■

Analiza grešaka u izračunatom sinusu i kosinusu hiperbolnom dana formulama (5.2.14) i (5.2.15), može se koristiti i za ocjenu J–norme pogreške izračunatog ℓ–tog stupca. Ocijenimo prvo

$$|(f\ell(g'_{1\ell}) - g'_{1\ell})^2 - (f\ell(g'_{2\ell}) - g'_{2\ell})^2| \leq \max\{(f\ell(g'_{1\ell}) - g'_{1\ell})^2, (f\ell(g'_{2\ell}) - g'_{2\ell})^2\} \quad . \tag{5.2.24}$$

Prva jednakost u relaciji (5.2.17), uz korištenje $|c| \geq |s|$, povlači

$$\begin{aligned} | f\ell(g'_{1\ell}) - g'_{1\ell})| &\leq |\delta c|\,|g_{1\ell}| + |\delta s|\,|g_{2\ell}| + 2\varepsilon(|c|\,|g_{1\ell}| + |s|\,|g_{2\ell}|) \\ &\leq (\beta+4)\varepsilon(|c|\,|g_{1\ell}| + |s|\,|g_{2\ell}|) \\ &\leq (\beta+4)\varepsilon|c|(|g_{1\ell}| + |g_{2\ell}|) \quad . \end{aligned} \tag{5.2.25}$$

Na sličan način, druga jednakost u (5.2.17) daje

$$\begin{aligned} | f\ell(g'_{2\ell}) - g'_{2\ell})| &\leq |\delta s|\,|g_{1\ell}| + |\delta c|\,|g_{2\ell}| + 2\varepsilon(|s|\,|g_{1\ell}| + |c|\,|g_{2\ell}|) \\ &\leq (\beta+4)\varepsilon(|s|\,|g_{1\ell}| + |c|\,|g_{2\ell}|) \\ &\leq (\beta+4)\varepsilon|c|(|g_{1\ell}| + |g_{2\ell}|) \quad . \end{aligned} \tag{5.2.26}$$

Za ocjenu J–norme pogreške potrebna je još jedna elementarna nejednakost

$$|x| + |y| \leq \sqrt{2}\sqrt{|x|^2 + |y|^2} \quad .$$

Relacije (5.2.24), (5.2.25) i (5.2.26) zajedno, daju ocjenu

$$\begin{aligned} \sqrt{|(f\ell(g'_{1\ell}) - g'_{1\ell})^2 - (f\ell(g'_{2\ell}) - g'_{2\ell})^2|} &\leq \max\{|(f\ell(g'_{1\ell}) - g'_{1\ell})|, |(f\ell(g'_{1\ell}) - g'_{1\ell})|\} \\ &\leq (\beta+4)\varepsilon|c|(|g_{1\ell}| + |g_{2\ell}|) \\ &\leq \frac{\sqrt{2}(\beta+4)\varepsilon}{\sqrt{1-\alpha^2}}\sqrt{g_{1\ell}^2 + g_{2\ell}^2} \quad . \end{aligned} \tag{5.2.27}$$

Posljednja nejednakost je i ocjena pogreške (u apsolutnom smislu) za odgovarajući stupac. Naime, ako su stupci matrice G standardno indefinitno skalirani, onda je

$$\sqrt{g_{1\ell}^2 + g_{2\ell}^2} \leq \sqrt{2} \quad .$$

Relativni oblik pogreške možemo dobiti samo u slučaju da norma tog stupca nije 0, tj. norma ostatka aktivnog dijela stupca je dominirajući član u J–normi.

5.3. Greške zaokruživanja za realne blok J–rotacije

Temelj JQR dekompozicije je dekompozicija elementarnih parova reda 3, odnosno 4. Bez smanjenja općenitosti možemo pretpostaviti da je $g_{11} \geq g_{21} \geq 0$ i $g_{32} \geq g_{42} \geq 0$. To uvijek možemo postići korištenjem J–unitarnih matrica Ψ definiranih u (3.1.1) i permutacija redaka matrice G.

Analiza grešaka zaokruživanja za jednu blok J–unitarnu matricu bazira se na analizi greške zaokruživanja za izračunavanje veličine r definirane relacijama (4.2.5), odnosno (4.2.6). Ako pokažemo da je

$$a = j_{11} j_{33} |z|^2 (|g_{32}|^2 - |g_{42}|^2)(|g_{21}|^2 - |g_{11}|^2)$$

malen broj, i da se računa točno u smislu apsolutne vrijednosti, onda se relativno točno računa i r, koji je oblika $r = 1/\sqrt{a+1}$. Razlog leži u činjenici da su za broj koji je blizu 1 apsolutna i relativna točnost jedno te isto.

Ako ne vrijedi ograda (5.2.6) za hiperbolni tangens kuta, a u matrici G postoje dva stupaca čija Gramova matrica daje indefinitnu i regularnu 2×2 matricu, koristit ćemo blok transformacije. Prvo, ako je

$$\alpha < \frac{g_{21}}{g_{11}} \leq 1 \qquad \text{i} \qquad \frac{g_{42}}{g_{32}} \leq \alpha < 1 \quad , \tag{5.3.1}$$

primijenit ćemo jednu hiperbolnu rotaciju za poništavanje jednog od elemenata g_{42}, a zatim blok 3×3 transformaciju.

Ako je

$$\alpha < \frac{g_{21}}{g_{11}} \leq 1 \qquad \text{i} \qquad \alpha < \frac{g_{42}}{g_{32}} \leq 1 \quad , \tag{5.3.2}$$

onda koristimo blok 4×4 transformaciju. Tvrdimo da je u oba slučaja broj a po apsolutnoj vrijednosti malen, uz pretpostavku da je $|z|$ razumne veličine, što bi nam trebalo osigurati odgovarajuće pivotiranje.

Broj $r - 1$ javlja se samo u izrazima oblika $(r-1)/(|g_{32}|^2 - |g_{42}|^2)$, odnosno $(r-1)/(|g_{21}|^2 - |g_{11}|^2)$. Uočimo da vrijedi

$$r - 1 = \frac{1}{\sqrt{a+1}} - 1 = \frac{-a}{\sqrt{1+a}(1+\sqrt{1+a})} \quad ,$$

što možemo iskoristiti za računanje

$$\begin{aligned}
\frac{r-1}{|g_{21}|^2-|g_{11}|^2} &= \frac{-j_{11}j_{33}|z|^2(|g_{32}|^2-|g_{42}|^2)}{\sqrt{1+a}(1+\sqrt{1+a})} \quad , \\
\frac{r-1}{|g_{32}|^2-|g_{42}|^2} &= \frac{-j_{11}j_{33}|z|^2(|g_{21}|^2-|g_{11}|^2)}{\sqrt{1+a}(1+\sqrt{1+a})} \quad .
\end{aligned}$$

Nažalost, ove formule osigurat će nam apsolutnu točnost za prethodne izraze, ako je $|a|$ bio malen broj, bez obzira kako računali odgovarajuće razlike kvadrata.

Iz ove grube analize slijede relacije

$$\begin{aligned}
f\ell(r) &= (1+\varepsilon_r)r \quad , \\
f\ell\left(\frac{r-1}{|g_{21}|^2-|g_{11}|^2}\right) &= \frac{r-1}{|g_{21}|^2-|g_{11}|^2}+\varepsilon_a \quad , \\
f\ell\left(\frac{r-1}{|g_{32}|^2-|g_{42}|^2}\right) &= \frac{r-1}{|g_{32}|^2-|g_{42}|^2}+\varepsilon_b \quad ,
\end{aligned}$$

pri čemu su ε_r, ε_a i ε_b po apsolutnoj vrijednosti omeđeni s nekoliko ε.

Kod primjene blok transformacija, u svakom stupcu matrice G mijenjaju se samo četiri (tri) elementa. Norma pogreške u odgovarajućem stupcu ovisi ne samo o odnosu J–norme cijelog stupca i J–norme dijela tog stupca koji mijenjamo, nego i o veličini elemenata u odgovarajućem stupcu.

Ako je J–norma dijela stupca koji mijenjamo, reda veličine najvećeg elementa koji sudjeluje u transformaciji i ako je relativno mala prema čitavoj J–normi stupca, onda će transformacija napraviti relativno malu pogrešku obzirom na čitavu J–normu stupca.

U slučaju da su redovi veličina J–norme čitavog stupca, J–norme dijela stupca koji mijenjamo i najvećeg elementa koji sudjeluje u transformaciji približno jednaki, transformacija radi po apsolutnoj vrijednosti malu normu pogreške.

Postavlja se pitanje koje veličine mora biti α, tako da imamo optimalan rez za primjenu hiperbolnih, odnosno, blok transformacija. Zbog točnosti računanja, α ne bi trebao biti niti blizu nuli, niti blizu jedan. Nažalost, katkad je matrica G takva da ne postoje dva stupca čija Gramova matrica daje indefinitnu i regularnu 2×2 podmatricu, a α je ipak vrlo blizu jedan. Na primjer, neka je zadan par

$$G = \begin{bmatrix} b & c \\ b+\varepsilon_1 & bc/(b+\varepsilon_1) \\ b+\varepsilon_1 & 0 \\ b+\varepsilon_2 & 0 \end{bmatrix} \quad , \quad J=\mathrm{diag}(-1,1,-1,1) \quad ,$$

pri čemu su b i c veliki, a ε_1 i ε_2 mali brojevi. Matrica G^*JG je dijagonalna, pa moramo koristiti 1×1 transformacije. Zbog malih J–normi oba stupca, nastalih razlikama kvadrata velikih brojeva, α mora biti blizu jedan.

5.4. Greške zaokruživanja za množenje matricom $\Psi = \mathrm{diag}(e^{i\psi_1}, \ldots, e^{i\psi_n})$

U prethodnih nekoliko odjeljaka analizirali smo pogreške zaokruživanja za transformacije realnih matrica. Već smo pokazali da elemente prvog stupca možemo svesti na realne množenjem matricom $\Psi = \mathrm{diag}(e^{i\psi_1}, \ldots, e^{i\psi_n})$. Pitanje je kolika je tada pogreška koju smo napravili. To ćemo pokazati u ovom odjeljku.

Nakon primjene te transformacije, trigonometrijske i hiperbolne rotacije koje ćemo primjenivati bit će realne, ali će se primjenjivati na kompleksne brojeve u ostalim stupcima. Ova razlika povećat će, približno dvostruko, samo konstante u ocjenama, jer računalo izvodi kompleksne operacije kao realne na uređenim parovima brojeva. Slična situacija bit će i kod blok rotacija, samo što će u tom slučaju elementi g_{12} i g_{22} ostati kompleksni.

Promotrimo perturbaciju koja nastaje množenjem jednog elementa nekog stupca matrice G, ako se odgovarajući kompleksni element u prvom stupcu transformira u realan iste apsolutne vrijednosti. Vrijedi

$$g'_{k1} = e^{i\psi} g_{k1} \quad ,$$

pri čemu je g'_{k1} realan, pozitivan broj. Izjednačavenjem imaginarnih dijelova lijeve i desne strane ove jednakosti dobivamo

$$\sin\psi \,\mathrm{Re}(g_{k1}) + \cos\psi \,\mathrm{Im}(g_{k1}) = 0 \quad ,$$

odnosno

$$\mathrm{tg}\,\psi = -\frac{\mathrm{Im}(g_{k1})}{\mathrm{Re}(g_{k1})} \quad .$$

Iz prethodne relacije, možemo izračunati sinus i kosinus odgovarajućeg kuta. Znak sinusa izaberimo tako da je $g'_{k1} = |g_{k1}|$. Vrijedi

$$\cos\psi = \frac{\mathrm{Re}(g_{k1})}{|g_{k1}|} \tag{5.4.1}$$

$$\sin\psi = \frac{-\mathrm{Im}(g_{k1})}{|g_{k1}|} \quad . \tag{5.4.2}$$

Za početak, treba ocijeniti grešku nastalu računanjem apsolutne vrijednosti kompleksnog broja. Upotrijebimo stabilnu formulu za računanje apsolutne vrijednosti

$$\sqrt{x^2+y^2} = \begin{cases} |x|\sqrt{1+(y/x)^2} & \text{za } |x| \geq |y| \\ |y|\sqrt{1+(x/y)^2} & \text{za } |y| > |x| \end{cases} \quad .$$

Zbog simetrije formula, dovoljno je ocijeniti grešku za jednu od prethodnih formula.

$$\begin{aligned} f\ell(\sqrt{x^2+y^2}) &= (1+\varepsilon_1)(1+\varepsilon_2)|x|\sqrt{(1+\varepsilon_3)(1+(1+\varepsilon_4)(1+\varepsilon_5)(y/x)^2)} \\ &= (1+\varepsilon_1)(1+\varepsilon_2)|x|\sqrt{(1+\varepsilon_6)(1+(y/x)^2)} \\ &= (1+\varepsilon_7)|x|\sqrt{(1+(y/x)^2)} \quad , \end{aligned}$$

pri čemu je

$$\varepsilon_6 \leq 4\varepsilon + \mathcal{O}(\varepsilon^2) \quad ,$$

odnosno

$$\varepsilon_7 \leq 4\varepsilon + \mathcal{O}(\varepsilon^2) \quad .$$

Dakle, za pogreške u izračunatom kutu vrijedi

$$f\ell(\cos\psi) = (1+\varepsilon_c)\cos\psi \quad , \quad |\varepsilon_c| \leq 5\varepsilon + \mathcal{O}(\varepsilon^2) \tag{5.4.3}$$

$$f\ell(\sin\psi) = (1+\varepsilon_s)\sin\psi \quad , \quad |\varepsilon_s| \leq 5\varepsilon + \mathcal{O}(\varepsilon^2) \quad . \tag{5.4.4}$$

Ova transformacija promijenit će apsolutne vrijednosti odgovarajućih elemenata u recima matrice, što će uzrokovati promjenu u J–normi odgovarajućeg stupca.

Izračunati element u odgovarajućem retku je

$$\begin{aligned} f\ell(g'_{k\ell}) &= (1+\varepsilon_1)(\mathrm{Re}(g_{k\ell})\cos\psi(1+\varepsilon_c)(1+\varepsilon_2) - (\mathrm{Im}(g_{k\ell})\sin\psi(1+\varepsilon_s)(1+\varepsilon_3)) \\ &+ i(1+\varepsilon_4)(\mathrm{Im}(g_{k\ell})\cos\psi(1+\varepsilon_c)(1+\varepsilon_5) + (\mathrm{Re}(g_{k\ell})\sin\psi(1+\varepsilon_s)(1+\varepsilon_6)) \quad . \end{aligned}$$

Pogreška u izračunatom elementu, do na $\mathcal{O}(\varepsilon^2)$ je

$$\begin{aligned} f\ell(g'_{k\ell}) - g'_{k\ell} &= (\varepsilon_1+\varepsilon_2+\varepsilon_c)(\mathrm{Re}(g_{k\ell})\cos\psi) - (\varepsilon_1+\varepsilon_3+\varepsilon_s)(\mathrm{Im}(g_{k\ell})\sin\psi) \\ &+ i((\varepsilon_4+\varepsilon_5+\varepsilon_c)(\mathrm{Im}(g_{k\ell})\cos\psi) + (\varepsilon_4+\varepsilon_6+\varepsilon_s)(\mathrm{Re}(g_{k\ell})\sin\psi)) \quad . \end{aligned}$$

Koristeći činjenicu da je $|\sin\psi\cos\psi| \leq 1/2$, te da je $x^2+y^2 \geq 2|x||y|$ za x, $y \in \mathbf{R}$, izračunajmo kvadrat apsolutne vrijednosti pogreške za izračunati element

$$\begin{aligned} |f\ell(g'_{k\ell}) - g'_{k\ell}|^2 &\leq (7\varepsilon)^2 \Big((|\mathrm{Re}(g_{k\ell})\cos\psi| + |\mathrm{Im}(g_{k\ell})\sin\psi|)^2 \\ &+ (|\mathrm{Im}(g_{k\ell})\cos\psi| + |\mathrm{Re}(g_{k\ell})\sin\psi|)^2\Big) \leq 98\varepsilon^2|g_{k\ell}|^2 \quad . \end{aligned}$$

Ako sve elemente koji nisu realni, prevodimo u realne, pogreška je po normi manja ili jednaka

$$\|f\ell(g'_\ell) - g'_\ell\|_2 \leq 7\sqrt{2}\varepsilon\|g_\ell\| \quad ,$$

odnosno, za ℓ–ti stupac, došlo je do promjene J–norme za najviše

$$\|f\ell(g'_\ell) - g'_\ell\|_J \leq 7\sqrt{2}\varepsilon\max\{\|g_\ell^+\|, \|g_\ell^-\|\} \quad .$$

5.5. Teoremi o perturbaciji indefinitne QR dekompozicije

U ovom odjeljku navodimo neke rezultate o perturbaciji indefinitne QR dekompozicije.

Teorem 5.5.1.

Neka su dane matrice $G \in \mathbf{C}^{m\times n}$, $m \geq n$ i $J \in \mathbf{C}^{m\times m}$, $J = \mathrm{diag}(j_{11}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$. Neka je $G = P_1^* QRP_2$ egzaktna indefinitna QR dekompozicija matrice G obzirom na skalarni produkt određen matricom J, tj. vrijedi

$$\begin{bmatrix} R \\ 0 \end{bmatrix} = Q^{-1} P_1^* G P_2 \quad , \quad Q^* J_1 Q = J_1 \quad , \quad J_1 = P_1^* J P_1 \quad ,$$

pri čemu su P_1 i P_2 matrice permutacija. Neka je matrica $\widetilde{R}$ izračunata u aritmetici pomičnog zareza. Ako je $\widetilde{R}$ trokutasta, tj. ako se proces računanja može provesti samo korištenjem trigonometrijskih i hiperbolnih rotacija, onda izračunati $\widetilde{R}$ možemo interpretirati kao egzaktni faktor malo perturbirane matrice $\widetilde{G} = G + E$,

$$\begin{bmatrix} \widetilde{R} \\ 0 \end{bmatrix} = \widetilde{Q}^{-1} P_1^* (G + E) P_2 \quad , \quad \widetilde{Q}^* J_1 \widetilde{Q} = J_1 \quad , \quad J_1 = P_1^* J P_1$$

uz ocjenu

$$\|P_1^* E P_2 e_k\|_2 \leq \gamma^{k-1} \|P_1^* G P_2 e_k\|_2 \sum_{i=1}^{k} \mathrm{err}(p_i) \quad , \tag{5.5.1}$$

gdje je $\mathrm{err}(p_i)$ ocjena pogreške koja je nastala pri poništavanju elemenata jednog stupca matrice G,

$$\mathrm{err}(p_i) = e_H (1 + e_G)^{p_i} + \gamma \left[(1 + e_G)^{p_i} - 1 \right] \quad . \tag{5.5.2}$$

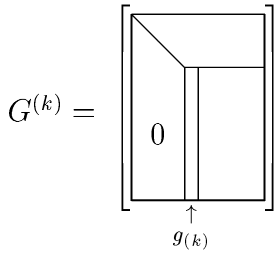
Pritom je p_i definiran relacijom (5.1.19) i označava broj koraka tj. nizova nezavisnih rotacija potrebnih za redukciju odgovarajućeg stupca na samo dva elementa. Konstante e_G, e_H, β i γ definirane su relacijama (5.1.18), (5.2.23), (5.2.8) i (5.2.19), respektivno.

Dokaz:

Permutacije stupaca matrice G ne utječu na analizu dekompozicije, ako u svakom koraku poništavanja za radni dio k–tog stupca vrijedi

$$\|g_{(k)}^{+}\| \leq \alpha \|g_{(k)}^{-}\| \quad \text{ili} \quad \|g_{(k)}^{-}\| \leq \alpha \|g_{(k)}^{+}\| \quad , \quad \alpha < 1 \quad , \quad k \geq 1 \quad ,$$

što znači da se dekompozicija provodi samo trigonometrijskim i hiperbolnim rotacijama. Zbog toga, bez smanjenja općenitosti, možemo pretpostaviti da smo prije



Slika 5.5.1. Oblik matrice G u k–tom koraku.

početka dekompozicije izvršili takvu permutaciju stupaca matrice G i odgovarajućih dijagonalnih elemenata u J. Oblik matrice G u k–tom koraku dan je slikom 5.5.1.

Uočimo da za poništavanje prvog stupca koristimo p_1 koraka, tj. nizove rotacija $\widetilde{U}_1^{-1}, \ldots, \widetilde{U}_{p_1}^{-1}$ izračuantih u aritmetici pomičnog zareza, a zatim najviše jednu hiperbolnu rotaciju $\widetilde{U}_H^{-1}(1,2)$. Greške zaokruživanja možemo interpretirati kao primjenu egzaktnih rotacija $U_{p_i}^{-1}$, odnosno, $U_H^{-1}(1,2)$ na malo perturbirane matrice $G_i^{(0)}$.

Uvedimo skraćenu oznaku za hiperbolnu rotaciju u k–tom koraku redukcije, tj. neka je $V_k = U_H(k,k+1)$ egzaktna, a $\widetilde{V}_k = \widetilde{U}_H(k,k+1)$ izračunata hiperbolna rotacija. Situacija nakon poništavanja prvog stupca, prikazana je dijagramom 5.5.2.

$$\begin{tikzcd}
G^{(0)} \arrow[r, "\widetilde{U}_1^{-1}"] \arrow[d] & G_1^{(0)} \arrow[r, "\widetilde{U}_2^{-1}"] \arrow[d] & G_2^{(0)} & \cdots & G_{p-1}^{(0)} \arrow[r, "\widetilde{U}_{p_1}^{-1}"] \arrow[d] & G_{p_1}^{(0)} \arrow[r, "\widetilde{V}_1^{-1}"] \arrow[d] & G^{(1)} \\
\widetilde{G}^{(0)} \arrow[ur, "U_1^{-1}"'] & \widetilde{G}_1^{(0)} \arrow[ur, "U_2^{-1}"'] & & & \widetilde{G}_{p_1-1}^{(0)} \arrow[ur, "U_{p_1}^{-1}"'] & \widetilde{G}_{p_1}^{(0)} \arrow[ur, "V_1^{-1}"'] &
\end{tikzcd}$$

Slika 5.5.2. Perturbacija matrice G u nakon poništavanja prvog stupca.

Primijetimo da za poništavanje elemenata radnog dijela stupca ne moramo koristiti hiperbolnu rotaciju, ako je $g_{(k)}^{+}$ ili $g_{(k)}^{-}$ trivijalan. U tom slučaju, u ocjeni pogreške neće biti člana koji proizlazi iz ocjene pogreške jedne hiperbolne rotacije.

Zbog jednostavnosti zapisa, uvedimo neke pokrate. Označimo nizove egzaktnih trigonometrijskih rotacija u k–tom koraku s $W_k = U_1 \cdots U_{p_k}$, a perturbirane rotacije s $\widetilde{W}_k = \widetilde{U}_1 \cdots \widetilde{U}_{p_k}$. Za perturbaciju stupaca matrice $G^{(1)}$ tada vrijedi

$$\begin{aligned}
\|\widetilde{V}_1^{-1}\widetilde{W}_1^{-1}g_\ell - V_1^{-1}W_1^{-1}g_\ell\|_2 &\leq \|\widetilde{V}_1^{-1}\widetilde{W}_1^{-1}g_\ell - V_1^{-1}\widetilde{W}_1^{-1}g_\ell\|_2 \\
&\quad + \|V_1^{-1}\widetilde{W}_1^{-1}g_\ell - V_1^{-1}W_1^{-1}g_\ell\|_2 \\
&\leq \|\widetilde{V}_1^{-1}\widetilde{g}_\ell - V_1^{-1}\widetilde{g}_\ell\|_2 \\
&\quad + \|V_1^{-1}\|_2\|\widetilde{g}_\ell - W_1^{-1}g_\ell\|_2 \quad , \qquad (5.5.3)
\end{aligned}$$

gdje je

$$\widetilde{g}_\ell = \widetilde{W}_1^{-1} g_\ell = \widetilde{U}_{p_1}^{-1} \cdots \widetilde{U}_1^{-1} g_\ell \quad .$$

Primijetimo da je relacija (5.5.3), zapravo dokazana u prethodnim odjeljcima.

Korištenjem nejednakosti trokuta i primjenom leme 5.1.1., zbog unitarnosti matrice W_1, izlazi

$$\|\widetilde{g}_\ell\|_2 = \|\widetilde{W}_1^{-1} g_\ell - W_1^{-1} g_\ell\|_2 + \|W_1^{-1} g_\ell\|_2 \leq (1+e_G)^p \|g_\ell\|_2 \quad . \tag{5.5.4}$$

Korištenjem relacija (5.2.18) i (5.2.19), za normu V_1^{-1} vrijedi ocjena $\|V_1^{-1}\|_2 \leq \gamma$.

Vratimo se ponovno na relaciju (5.5.3). Korištenjem relacija (5.5.4), (5.2.23) i (5.1.21) slijedi

$$\begin{aligned} \|\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell - V_1^{-1} W_1^{-1} g_\ell\|_2 &\leq e_H \|\widetilde{g}_\ell\|_2 + \gamma\left((1+e_G)^{p_1} - 1\right)\|g_\ell\|_2 \\ &\leq \left[e_H (1+e_G)^{p_1} + \gamma\left((1+e_G)^{p_1} - 1\right)\right]\|g_\ell\|_2 \\ &= \mathrm{err}(p_1)\|g_\ell\|_2 \quad . \end{aligned} \tag{5.5.5}$$

Prethodna relacija, zapravo daje perturbacijsku ocjenu za prvi stupac matrice G, jer se on nadalje više neće mijenjati. Za poništavanje drugog stupca koristimo matrice V_2^{-1} i W_2^{-1}. Korištenjem (5.5.5) izlazi

$$\begin{aligned} &\|\widetilde{V}_2^{-1}\widetilde{W}_2^{-1}\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell - V_2^{-1}W_2^{-1}V_1^{-1}W_1^{-1} g_\ell\|_2 \\ &\quad \leq \|\widetilde{V}_2^{-1}\widetilde{W}_2^{-1}\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell - V_2^{-1}W_2^{-1}\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell\|_2 \\ &\qquad + \|V_2^{-1}W_2^{-1}\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell - V_2^{-1}W_2^{-1}V_1^{-1}W_1^{-1} g_\ell\|_2 \\ &\quad \leq \|\widetilde{V}_2^{-1}\widetilde{W}_2^{-1}\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell - V_2^{-1}W_2^{-1}\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell\|_2 \\ &\qquad + \|V_2^{-1}W_2^{-1}\|_2 \, \|\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell - V_1^{-1}W_1^{-1} g_\ell\|_2 \\ &\quad \leq \mathrm{err}(p_2)\|\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell\|_2 + \gamma\,\mathrm{err}(p_1)\|g_\ell\|_2 \\ &\quad \leq \mathrm{err}(p_2)(\|\widetilde{V}_1^{-1}\widetilde{W}_1^{-1} g_\ell - V_1^{-1}W_1^{-1} g_\ell\|_2 \\ &\qquad + \|V_1^{-1}W_1^{-1} g_\ell\|_2) + \gamma\,\mathrm{err}(p_1)\|g_\ell\|_2 \\ &\quad \leq \mathrm{err}(p_2)(\mathrm{err}(p_1) + \gamma)\|g_\ell\|_2 + \gamma\,\mathrm{err}(p_1)\|g_\ell\|_2 \\ &\quad \leq \gamma(\mathrm{err}(p_1) + \mathrm{err}(p_2))\|g_\ell\|_2 + \mathcal{O}(\mathrm{err}(p)^2) \quad . \end{aligned}$$

Dokaz teorema dobivamo indukcijom po k. ■

Za teoreme perurbacije po komponentama, potrebna nam je definicija i neka svojstva matrice apsolutnih vrijednosti.

Definicija 5.5.1.

Neka je $A \in C^{m\times n}$, $A = [a_{ij}]$. Apsolutna vrijednost $|A|$ matrice A je

$$|A| = (|a_{ij}|) \quad .$$

Pisat ćemo $|A| \leq |B|$, ako vrijedi $|a_{ij}| \leq |b_{ij}|$, $i = 1, \ldots, m$, $j = 1, \ldots n$. ■

Propozicija 5.5.1.

Ako su A, B i C tri matrice takve da je $A = BC$, onda vrijedi $|A| \leq |B||C|$.

Dokaz:

Dokaz slijedi korištenjem nejednakosti trokuta za apsolutnu vrijednost elementa a_{ij}. Vrijedi

$$|a_{ij}| = \left|\sum_k b_{ik}c_{kj}\right| \leq \sum_k |b_{ik}||c_{kj}| \quad .$$

Posljednja suma je (i,j)–ti element produkta matrica $|B||C|$. ∎

Za indefinitnu QR dekompoziciju, u slučaju da je matrica R trokutasta, vrijede slične ocjene kao i za običnu QR dekompoziciju. Slijedeći korolar generalizacija je leme 18.8 iz [25].

Korolar 5.5.1.

Uz uvjete teorema 5.5.1., *i oznaku*

$$\gamma_c = \gamma^{n-1}\sum_{i=1}^{n} \mathrm{err}(p_n)$$

vrijede slijedeće ocjene

$$\begin{aligned} \|E\|_F &\leq \gamma_c\|G\|_F \\ |E| &\leq m\gamma_c K|G| \quad , \quad \|K\|_F = 1 \quad , \end{aligned}$$

gdje je K matrica definirana s

$$K = \frac{1}{m}ee^* \quad , \quad e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad .$$

Dokaz:

Prva tvrdnja izlazi kvadriranjem i zbrajanjem nejednakosti (5.5.1) iz teorema 5.5.1. Druga tvrdnja je posljedica činjenice da je apsolutna vrijednost pogreške jednog elementa u stupcu manja ili jednaka normi pogreške čitavog stupca i nejednakosti

$$\sqrt{\sum_{i=1}^{m} |g_{ik}|^2} \leq \sum_{i=1}^{m} |g_{ik}| \quad .$$

∎

Komponentnim ocjenama za perturbaciju QR dekompozicije bavili su se Stewart u [41], [42], Sun u [43], [44], [45], [46] i Zha u [54]. Vrlo slična perturbacijska analiza može se provesti i za indefinitnu QR dekompoziciju, ako znamo da se koriste samo obične i hiperbolne rotacije.

Prvo uvedimo neke oznake.

Definicija 5.5.2.

Neka je $A \in \mathbf{C}^{n\times n}$. Donjetrokutastu matricu $\mathrm{tril}(A)$ *i gornjetrokutastu matricu* $\mathrm{triu}(A)$ *definiramo s*

$$[\mathrm{tril}(A)]_{ij} = \begin{cases} a_{ij} & , \quad i \geq j \quad , \\ 0 & , \quad \textit{inače} \quad . \end{cases}$$

$$[\mathrm{triu}(A)]_{ij} = \begin{cases} a_{ij} & , \quad i \leq j \quad , \\ 0 & , \quad \textit{inače} \quad . \end{cases}$$

■

Za dokaz teorema o perturbaciji, potrebna nam je definicija nenegativnih matrica. Opširniji pregled svojstava takvih matrica, može se naći u [26], odakle je preuzeta slijedeća definicija.

Definicija 5.5.3.

Neka su $A, B \in \mathbf{C}^{m\times n}$. Matrična norma $\|\ \|$ je monotona, ako za svaki par matrica A, B, takav da je $|A| \leq |B|$, vrijedi $\|A\| \leq \|B\|$.

Matrična norma je apsolutna, ako vrijedi $\|A\| = \|\,|A|\,\|$. ■

Propozicija 5.5.2.

Euklidska norma matrice je monotona i apsolutna.

Dokaz:

Dokaz je trivijalna posljedica definicije euklidske (Frobenijusove) norme

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \quad .$$

■

Za dokaz teorema koji ćemo iskazati, potrebna je slijedeća definicija.

Definicija 5.5.4.

Matrica $A = [a_{ij}] \in \mathbf{R}^{n\times n}$ je M–matrica, ako je realni dio svih njezinih svojstvenih vrijednosti pozitivan i ako je

$$a_{ij} \leq 0 \quad , \quad 1 \leq i, j \leq n \quad , \quad i \neq j \quad .$$

■

Navest ćemo i nekoliko svojstava M–matrica. Dokazi se mogu pogledati u [27].

Propozicija 5.5.3. *Svaka M–matrica A se može zapisati u obliku $A = \alpha I - P$, pri čemu je $P \geq 0$, a za spektralni radijus matrice P vrijedi $\alpha \geq \rho(P)$. (U ovom slučaju $\geq$ znači da su svi elementi veći ili jednaki 0.)* ■

Propozicija 5.5.4. *Ako je A M–matrica oblika $A = I - E$ i spektralni radijus $\rho(E) < 1$, onda je A nesingularna i $A^{-1} \geq 0$.* ■

Za indefinitnu QR dekompoziciju vrijedi generalizacija teorema iz [44].

Teorem 5.5.2.

Neka su matrice $G, \widetilde{G} \in \mathbf{C}^{m\times n}$ takve da vrijedi $\mathrm{rang}(G) = \mathrm{rang}(\widetilde{G}) = n$ *i neka su*

$$G = P_1 Q R P_2^* = P_1 Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} P_2^* \quad , \quad Q^* J_1 Q = J_1 \quad , \quad J_1 = P_1^* J P_1$$

i

$$\widetilde{G} = P_1 \widetilde{Q} \widetilde{R} P_2^* = P_1 \widetilde{Q} \begin{bmatrix} \widetilde{R}_1 \\ 0 \end{bmatrix} P_2^* \quad , \quad \widetilde{Q}^* J_1 \widetilde{Q} = J_1 \quad , \quad J_1 = P_1^* J P_1$$

indefinitne QR dekompozicije matrica G i $\widetilde{G}$, obzirom na matricu J, gdje su P_1 i P_2 permutacije, a R_1 i $\widetilde{R}_1$ gornjetrokutaste matrice. Neka je $E = \widetilde{G} - G$, $F = \widetilde{R} - R$ i $W = \widetilde{Q} - Q$ i neka je

$$E_{ar} = |\widetilde{R}^{-*} P_2^* (E^* J \widetilde{G} + G^* J E) P_2 \widetilde{R}^{-1}| \quad .$$

Ako je spektralni radijus $\rho(E_{ar}) < 1$ i ako je $\rho(\mathrm{triu}(E_{ar}(I - E_{ar})^{-1})) < 1$, onda vrijedi

$$|F| \leq \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}) |\widetilde{R}| \quad ,$$

i

$$|W| \leq (P_1^* \, |E| \, P_2 - |\widetilde{Q}| \, \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}) |\widetilde{R}|) |\widetilde{R}^{-1}| (I - \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}))^{-1} \quad .$$

Dokaz:

Vrijedi jednakost

$$\begin{aligned} E^* J \widetilde{G} + G^* J E &= \widetilde{G}^* J \widetilde{G} - G^* J G \\ &= P_2 \widetilde{R}^* \widetilde{Q}^* P_1^* J P_1 \widetilde{Q} \widetilde{R} P_2^* - P_2 R^* Q^* P_1^* J P_1 Q R P_2^* \\ &= P_2 (\widetilde{R}^* J_1 \widetilde{R} - R^* J_1 R) P_2^* \\ &= P_2 (F^* J_1 \widetilde{R} + R^* J_1 F) P_2^* \quad . \end{aligned}$$

Množenjem ove jednakosti slijeva s $R^{-*}P_2^*$ i zdesna s $P_2\widetilde{R}^{-1}$, dobivamo

$$\begin{aligned} R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1} &= R^{-*}F^*J_1 + J_1F\widetilde{R}^{-1} \\ &= (J_1FR^{-1})^* + J_1F\widetilde{R}^{-1} \quad . \end{aligned} \tag{5.5.6}$$

Pretpostavimo li da R i $\widetilde{R}$ imaju iste znakove dijagonalnih elemenata, onda ih sigurno imaju i J_1FR^{-1} i $J_1F\widetilde{R}^{-1}$. Zbog toga što su i F, R^{-1} i $\widetilde{R}^{-1}$ gornjetrokutaste matrice iz relacije (5.5.6) slijedi

$$|F\widetilde{R}^{-1}| = |J_1F\widetilde{R}^{-1}| \leq \mathrm{triu}(|R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}|) \tag{5.5.7}$$

$$|(FR^{-1})^*| = |(J_1FR^{-1})^*| \leq \mathrm{tril}(|R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}|) \quad . \tag{5.5.8}$$

Uočimo da je

$$R^{-*}\widetilde{R}^* = I + R^{-*}(\widetilde{R}^* - R^*) = I + R^{-*}F^* \quad .$$

Množenjem prethodne relacije zdesna s $\widetilde{R}^{-*}$ dobivamo

$$R^{-*} = \widetilde{R}^{-*} + (FR^{-1})^*\widetilde{R}^{-*} \quad . \tag{5.5.9}$$

Raspišimo lijevu stranu relacije (5.5.6) korištenjem (5.5.9),

$$\begin{aligned} R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1} &= \widetilde{R}^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1} \\ &\quad + (FR^{-1})^*\widetilde{R}^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1} \quad . \end{aligned}$$

Korištenjem te relacije i relacije (5.5.6) dobivamo

$$\begin{aligned} |R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}| &\leq E_{ar} + |(FR^{-1})^*|\,|\widetilde{R}^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}| \\ &\leq E_{ar} + \mathrm{tril}(|R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}|)E_{ar} \\ &\leq E_{ar} + (|R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}|)E_{ar} \quad , \end{aligned}$$

odnosno,

$$|R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}|(I - E_{ar}) \leq E_{ar} \quad .$$

Prema pretpostavci je $\rho(E_{ar}) < 1$, pa je $I - E_{ar}$ regularna M–matrica i inverz joj je nenegativan. Odatle odmah slijedi

$$|R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}| \leq E_{ar}(1 - E_{ar})^{-1} \quad . \tag{5.5.10}$$

Primijetimo još da vrijedi $|F| \leq |F\widetilde{R}^{-1}|\,|\widetilde{R}|$. Ako to iskoristimo zajedno s (5.5.7) i (5.5.10), dobivamo prvu tvrdnju teorema

$$\begin{aligned} |F| &\leq |F\widetilde{R}^{-1}|\,|\widetilde{R}| \\ &\leq \mathrm{triu}(|R^{-*}P_2^*(E^*J\widetilde{G}+G^*JE)P_2\widetilde{R}^{-1}|)|\widetilde{R}| \\ &\leq \mathrm{triu}(E_{ar}(I - E_{ar})^{-1})|\widetilde{R}| \quad . \end{aligned} \tag{5.5.11}$$

Ocjenu za W dobivamo iz jednakosti

$$\begin{aligned} E &= \widetilde{G} - G = P_1(\widetilde{Q}\widetilde{R} - QR)P_2^* \\ &= P_1(\widetilde{Q}\widetilde{R} - \widetilde{Q}R + \widetilde{Q}R - QR)P_2^* = P_1(\widetilde{Q}F + WR)P_2^* \quad , \end{aligned}$$

tj. matrica W je jednaka

$$W = (P_1^* E P_2 - \widetilde{Q}F)R^{-1} \quad . \tag{5.5.12}$$

Dodatno, R možemo zapisati u obliku

$$R = (I - F\widetilde{R}^{-1})\widetilde{R} \quad .$$

Invertiranjem prethodne relacije

$$R^{-1} = \widetilde{R}^{-1}(I - F\widetilde{R}^{-1})^{-1} \quad , \tag{5.5.13}$$

a zatim uvrštavanjem u (5.5.12) dobivamo

$$W = (P_1^* E P_2 - \widetilde{Q}F)\widetilde{R}^{-1}(I - F\widetilde{R}^{-1})^{-1} \quad . \tag{5.5.14}$$

Posljednje dvije nejednakosti u relaciji (5.5.11) povlače

$$|F\widetilde{R}^{-1}| \leq \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}) \quad , \tag{5.5.15}$$

pa uvjet $\rho(\mathrm{triu}(E_{ar}(I - E_{ar})^{-1})) < 1$ znači da je matrica $I - |F\widetilde{R}^{-1}|$ nesingularna M–matrica, pa je $(I - |F\widetilde{R}^{-1}|)^{-1}$ matrica s nenegativnim elementima.

Da bismo dobili ocjenu za W, treba još primijetiti da je

$$\begin{aligned} |(I - F\widetilde{R}^{-1})^{-1}| &= |I + F\widetilde{R}^{-1} + \cdots + (F\widetilde{R}^{-1})^n + \cdots| \\ &\leq I + |F\widetilde{R}^{-1}| + \cdots + |F\widetilde{R}^{-1}|^n + \cdots \\ &\leq I + \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}) + \cdots + (\mathrm{triu}(E_{ar}(I - E_{ar})^{-1}))^n + \cdots \\ &= (I - \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}))^{-1} \quad . \end{aligned}$$

Zajedno s (5.5.13) i (5.5.15) izlazi

$$|R^{-1}| \leq |\widetilde{R}^{-1}|\,|(I - F\widetilde{R}^{-1})^{-1}| \leq |\widetilde{R}^{-1}|(I - \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}))^{-1} \quad , \tag{5.5.16}$$

i

$$\begin{aligned} |W| &\leq (P_1^*|E|P_2| - |\widetilde{Q}|\,|F|)|\widetilde{R}^{-1}|(I - \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}))^{-1} \\ &\leq (P_1^*|E|P_2 - |\widetilde{Q}|\,\mathrm{triu}(E_{ar}(I - E_{ar})^{-1})|\widetilde{R}|) \\ &\qquad \cdot |\widetilde{R}^{-1}|(I - \mathrm{triu}(E_{ar}(I - E_{ar})^{-1}))^{-1} \quad . \end{aligned}$$

■

Na sličan način možemo dobiti i relativnu perturbacijsku analizu indefinitne QR dekompozicije. Zha je u [54] primijetio da klasična perturbacijska analiza ne poštuje invarijantnost QR dekompozicije obzirom na skaliranje stupaca u matrici, tj. da za svaku regularnu dijagonalnu matricu D vrijedi

$$G = QR \implies GD = Q(RD) \quad .$$

Slična tvrdnja vrijedi i za indefinitnu QR dekompoziciju.

Definirajmo još Bauer–Skeelovu uvjetovanost matrice.

Definicija 5.5.5.

Za kvadratnu nesingularnu matricu S, Bauer–Skeelova uvjetovanost u odnosu na normu $\| \ \|$ je

$$\kappa_{BS}(S) = \| \, |S^{-1}| \, |S| \, \| \quad .$$

■

Teorem 5.5.3.

Neka su matrice $G, \widetilde{G} \in \mathbf{C}^{m\times n}$ takve da vrijedi $\mathrm{rang}(G) = \mathrm{rang}(\widetilde{G}) = n$ *i neka su*

$$G = P_1 Q R P_2^* = P_1 Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} P_2^* \quad , \quad Q^* J_1 Q = J_1 \quad , \quad J_1 = P_1^* J P_1$$

i

$$\widetilde{G} = P_1 \widetilde{Q} \widetilde{R} P_2^* = P_1 \widetilde{Q} \begin{bmatrix} \widetilde{R}_1 \\ 0 \end{bmatrix} P_2^* \quad , \quad \widetilde{Q}^* J_1 \widetilde{Q} = J_1 \quad , \quad J_1 = P_1^* J P_1$$

indefinitne QR dekompozicije matrica G i $\widetilde{G}$, obzirom na matricu J, takve da su matrice R_1 i $\widetilde{R}_1$ gornjetrokutaste, a matrice P_1 i P_2 su permutacije. Definiramo $E = \widetilde{G} - G$, $F = \widetilde{R} - R$ i $W = \widetilde{Q} - Q$. Pretpostavimo da vrijedi

$$|E| \le \varepsilon K |G| \quad ,$$

gdje je K neka matrica s nenegativnim elementima. Neka je $\| \ \|$ konzistentna i monotona norma,

$$\eta = \max\{ \| \, |Q^*| P_1^* K P_1 |Q| \, \|, \| \, |Q^*| P_1^* K^* P_1 |Q| \, \| \} \quad ,$$

i neka je

$$\varepsilon \eta (\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*)) < 1 \quad .$$

Tada vrijedi

$$\frac{\|F\|}{\|R\|} \le \varepsilon \eta (\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*)) + \mathcal{O}(\varepsilon^2) \quad ,$$

i

$$\frac{\|W\|}{\|Q\|} \le \frac{\varepsilon \left(\|P_1^* K P_1\| \kappa_{BS}(R^{-1}) + \eta (\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*)) \right)}{1 - \varepsilon \eta (\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*))} + \mathcal{O}(\varepsilon^2) \quad .$$

Dokaz:

Uočimo da je

$$\widetilde{G} = G + E = P_1 \widetilde{Q} \widetilde{R} P_2^* = P_1 (Q + W)(R + F) P_2^* \quad .$$

Iskoristimo još da je $G = P_1 Q R P_2^*$. Odatle odmah slijedi

$$P_1 Q R P_2^* + E = P_1 (QR + WR + QF + WF) P_2^* \quad ,$$

odnosno

$$P_1^* E P_2 = WR + QF + WF \quad .$$

Zbog regularnosti matrice R, imamo

$$W = P_1^* E P_2 R^{-1} - QFR^{-1} - WFR^{-1} \quad ,$$

pa je

$$|W| \leq |P_1^* E P_2|\, |R^{-1}| + |Q|\, |FR^{-1}| + |W|\, |FR^{-1}| \quad . \tag{5.5.17}$$

Promotrimo relaciju za $\widetilde{G}^* J \widetilde{G}$:

$$\widetilde{G}^* J \widetilde{G} = (G + E)^* J (G + E) = G^* J G + E^* J G + G^* J E + E^* J E \quad . \tag{5.5.18}$$

S druge strane, raspisivanjem dekompozicije matrice $\widetilde{G}$ izlazi

$$\begin{aligned} \widetilde{G}^* J \widetilde{G} &= P_2 \widetilde{R}^* \widetilde{Q}^* P_1^* J P_1 \widetilde{Q} \widetilde{R} P_2^* = P_2 \widetilde{R}^* J_1 \widetilde{R} P_2^* \\ &= P_2 (R + F)^* J_1 (R + F) P_2^* \\ &= P_2 (R^* J_1 R + R^* J_1 F + F^* J_1 R + F^* J_1 F) P_2^* \\ &= G^* J G + P_2 (R^* J_1 F + F^* J_1 R + F^* J_1 F) P_2^* \quad . \end{aligned} \tag{5.5.19}$$

Izjednačavanjem desnih strana u (5.5.18) i (5.5.19), izlazi

$$P_2 (R^* J_1 F + F^* J_1 R + F^* J_1 F) P_2^* = E^* J G + G^* J E + E^* J E \quad .$$

Množenjem prethodne jednakosti slijeva s $R^{-*} P_2^*$ i zdesna s $P_2 R^{-1}$, dobivamo

$$\begin{aligned} & J_1 F R^{-1} + R^{-*} F^* J_1 + R^{-*} F^* J_1 F R^{-1} \\ &\qquad = R^{-*} P_2^* (E^* J G + G^* J E + E^* J E) P_2 R^{-1} \quad . \end{aligned} \tag{5.5.20}$$

Prva dva člana s desne strane možemo napisati i u obliku

$$\begin{aligned} R^{-*} P_2^* E^* J G P_2 R^{-1} &= R^{-*} P_2^* E^* J P_1 Q \\ R^{-*} P_2^* G^* J E P_2 R^{-1} &= Q^* P_1^* J E P_2 R^{-1} \quad . \end{aligned} \tag{5.5.21}$$

Uvažimo li da su elementi $E^* J E$ i $F^* J_1 F$ reda veličine ε^2, tj. za red veličine manji od onih u E, odnosno F, dobivamo linearizirano ocjenu. Zbog toga što su obje matrice s lijeve strane u (5.5.20) trokutaste, a dijagonalni elementi su im realni i istog znaka, korištenjem relacija (5.5.20) i (5.5.21), te pretpostavki teorema, dobivamo

$$\begin{aligned}
|FR^{-1}| &= |J_1FR^{-1}| \leq |J_1FR^{-1} + (J_1FR^{-1})^*| \\
&\leq |R^{-*}P_2^*E^*JP_1Q| + |Q^*P_1^*JEP_2R^{-1}| \\
&\leq |R^{-*}||P_2^*||E^*||P_1|Q| + |Q^*||P_1^*||E|P_2||R^{-1}| \\
&\leq \varepsilon(|R^{-*}||P_2^*|G^*|K^*P_1|Q| + |Q^*|P_1^*K|G|P_2|R^{-1}|) \\
&\leq \varepsilon(|R^{-*}|\,|R^*|\,|Q^*|P_1^*K^*P_1|Q| + |Q^*|P_1^*KP_1|Q|\,|R|\,|R^{-1}|) \quad .
\end{aligned}$$

Ponovno, uočimo da je $|F| \leq |FR^{-1}|\,|R|$. Tada je zbog monotonosti i apsolutnosti norme

$$\begin{aligned}
\|F\| &\leq \|FR^{-1}\|\,\|R\| \\
&\leq \varepsilon(\|\,|R^{-*}|\,|R^*|\,\|\,\|\,|Q^*|P_1^*K^*P_1|Q|\,\| \\
&\qquad + \|\,|Q^*|P_1^*KP_1|Q|\,\|\,\|\,|R|\,|R^{-1}|\,\|)\|R\| \\
&= \varepsilon\eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*))\|R\| \quad .
\end{aligned}$$

Nakon dijeljenja s $\|R\|$ slijedi prva tvrdnja teorema. Primijetimo da smo ovdje dobili i ocjenu za $\|FR^{-1}\|$:

$$\|FR^{-1}\| \leq \varepsilon\eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*)) \quad .$$

Vratimo se ponovno na relaciju (5.5.17). Uzimanjem u obzir pretpostavke $|E| \leq \varepsilon K|G|$, izlazi

$$\begin{aligned}
|W| &\leq \varepsilon P_1^*K|G|P_2|R^{-1}| + |Q|\,|FR^{-1}| + |W|\,|FR^{-1}| \\
&\leq \varepsilon P_1^*KP_1|Q|\,|R|\,|R^{-1}| + |Q|\,|FR^{-1}| + |W|\,|FR^{-1}| \quad .
\end{aligned}$$

Uzimanjem norme i korištenjem ocjene za $\|FR^{-1}\|$, imamo

$$\begin{aligned}
\|W\| &\leq \varepsilon\|P_1^*KP_1\|\,\|Q\|\,\|\,|R|\,|R^{-1}|\,\| + \|Q\|\,\|FR^{-1}\| + \|W\|\,\|FR^{-1}\| \\
&\leq \varepsilon(\|P_1^*KP_1\|\,\|Q\|\,\|\,|R|\,|R^{-1}|\,\| \\
&\qquad + \eta\|Q\|(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*)) + \eta\|W\|(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*))) \quad .
\end{aligned}$$

Po pretpostavci teorema, vrijedi da je

$$\varepsilon\eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*)) < 1 \quad ,$$

pa smijemo posljednji član s desne strane prebaciti na lijevu, a zatim čitavu relaciju podijeliti s $1 - \varepsilon\eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*))$. Dobivamo

$$\|W\| \leq \frac{\varepsilon\|Q\|(\|P_1^*KP_1\|\kappa_{BS}(R^{-1}) + \eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*))}{1 - \varepsilon\eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*))} \quad ,$$

odakle dijeljenjem s $\|Q\|$ slijedi tvrdnja teorema. ■

Korolar 5.5.2.

Za indefinitnu QR dekompoziciju realiziranu samo trigonometrijskim i hiperbolnim rotacijama, uz uvjete teorema 5.5.3. *vrijedi*

$$\eta = \|Q^*\|_F \|Q\|_F \quad ,$$

$$\frac{\|F\|_F}{\|R\|_F} \leq m\gamma_c \eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*)) \quad ,$$

i

$$\frac{\|W\|_F}{\|Q\|_F} \leq \frac{m\gamma_c\left(\kappa_{BS}(R^{-1}) + \eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*))\right)}{1 - m\gamma_c\eta(\kappa_{BS}(R^{-1}) + \kappa_{BS}(R^*))} \quad .$$

Dokaz:

Iskoristimo li da je

$$\|P_1^* K P_1\|_F = \frac{1}{m}\|P_1^* e e^* P_1\|_F = 1 \quad ,$$

onda je dokaz direktna posljedica druge tvrdnje korolara 5.5.1. i tvrdnje prethodnog teorema. ∎

6. Primjene indefinitne QR dekompozicije

Obzirom da postoji mnogo primjena indefinitne QR dekompozicije, u ovom poglavlju navest ćemo samo neke. Prvi odjeljak bavi se generalizacijom algoritma iterirane dekompozicije Choleskog (flip–flap).

U drugom odjeljku pokazuje se da se korištenjem J–unitarnih i unitarnih transformacija može provesti bidijagonalizacija faktora G, obzirom na zadani J, bez eksplicitnog formiranja matrice G^*JG.

U trećem odjeljku pokazuje se primjena indefinitne QR dekompozicije na iterativno poboljšavanje netočno izračunatih svojstvenih vrijednosti. Posebno, za neke klase skalirano dijagonalno dominantnih matrica, pokazuje se da primjenom indefinitne QR dekompozicije radimo relativno malu perturbaciju u svojstvenim vrijednostima matrice.

6.1. Iterirana indefinitna QR dekompozicija

Rutishauser [34] je početkom šezdesetih godina opisao metodu za nalaženje svojstvenih vrijednosti simetričnih matrica. Algoritam je poznat kao LR algoritam i preteča je QR algoritma. Fernando i Parlett [20] su sredinom devedesetih godina napravili implicitnu varijantu ovog algoritma za pozitivno definitne matrice tj. algoritam za nalaženje singularnih vrijednosti neke matrice. Pokazali su ne samo linearnu konvergenciju algoritma, nego da se za ubrzanje konvergencije mogu koristiti pomaci (shiftovi), koji se jednostavno uklapaju u algoritam. Drmač [17] je pokazao da taj algoritam bez pomaka ima svojstvo da singularne vrijednosti računa na visoku relativnu točnost.

U ovom odjeljku taj se algoritam generalizira na algoritam za nalaženje svojstvenih vrijednosti hermitskih indefinitnih matrica i to korištenjem hermitske indefinitne dekompozicije matrice A ili JQR dekompozicije njenog faktora G, obzirom na J koji je inercija matrice A. Posebno je interesantna njegova implicitna varijanta, ali kao otvoreni problem ostaje procjena veličine pomaka za ubrzanje konvergencije.

6.1.1. Iterirana dekompozicija Choleskog

Na početku, promotrimo eksplicitni algoritam iterirane dekompozicije Choleskog.

Algoritam 6.1.1.

Slijedeći algoritam računa svojstvene vrijednosti hermitske pozitivno definitne matrice A, tako da je matrica G_k u k–tom koraku algoritma dobivena dekompozicijom Choleskog odgovarajuće matrice A_k.

start $A_1 = A$;

$$
\begin{array}{rcccc}
A_1 & = & G_1^* G_1 & & \\
A_2 & = & G_1 G_1^* & = & G_2^* G_2 \\
 & \vdots & & & \\
A_{2k} & = & G_{2k-1} G_{2k-1}^* & = & G_{2k}^* G_{2k} \\
A_{2k+1} & = & G_{2k} G_{2k}^* & = & G_{2k+1}^* G_{2k+1} \\
 & \vdots & & &
\end{array}
$$

■

Ako nam je umjesto matrice A zadan njezin faktor G, dobiven, na primjer, dekompozicijom Choleskog, onda dobivamo implicitni algoritam iterirane dekompozicije Choleskog.

Algoritam 6.1.2.

Slijedeći algoritam računa singularne vrijednosti matrice G, odnosno svojstvene vrijednosti matrice $A = G^* G$. U k–tom koraku algoritma, matrica G_k dobiva se QR dekompozicijom matrice G_{k-1}^*.

start: $G_1^* = G^*$;

$$
\begin{array}{rcccc}
G_1^* & = & Q_1 G_2 & , & Q_1^* Q_1 = I \\
G_2^* & = & Q_2 G_3 & , & Q_2^* Q_2 = I \\
 & \vdots & & & \\
G_{2k}^* & = & Q_{2k} G_{2k+1} & , & Q_{2k}^* Q_{2k} = I \\
G_{2k+1}^* & = & Q_{2k+1} G_{2k+2} & , & Q_{2k+1}^* Q_{2k+1} = I \\
 & \vdots & & &
\end{array}
$$

■

Fernando i Parlett [20] su pokazali da su dva koraka iterirane dekompozicije Choleskog ekvivalentna jednom koraku QR algoritma bez pomaka. Osim toga, ako se dekompozicija provodi rotacijama, postoje prirodne ocjene za najmanje singularne vrijednosti matrica G_k, koje se čitaju iz dijagonalnih elemenata. Tako se u dekompoziciju mogu na prirodan način uvesti pomaci.

6.1.2. Eksplicitni algoritam iterirane indefinitne dekompozicije

U ovom paragrafu konstruira se indefinitni analogon iterirane dekompozicije Choleskog. Prvo se konstruira eksplicitni oblik algoritma, koji je čisto teoretske prirode. Puna snaga algoritma iterirane indefinitne dekompozicije očituje se tek ako je zadan (po mogućnosti kvadratni) faktor odgovarajuće matrice.

Algoritam 6.1.3.

Neka je zadana hermitska indefinitna matrica A i neka je njena indefinitna dekompozicija $A = G^* JG$, pri čemu je G blok gornjetrokutasta matrica s dijagonalnim blokovima dimenzije 1×1 ili 2×2, a $J = \mathrm{diag}(\pm 1)$. U k–tom koraku algoritma matrice G_k i J_k dobivaju se hermitskom indefinitnom dekompozicijom matrice A_k. Slijedeći algoritam, u neparnim koracima, računa svojstvene vrijednosti matrice A.

start $A_1 = A$;

$$\begin{array}{rcccc}
A_1 & = & G_1^* J_1 G_1 & & \\
A_2 & = & G_1 G_1^* & = & G_2^* G_2 \\
 & \vdots & & & \\
A_{2k} & = & G_{2k-1} G_{2k-1}^* & = & G_{2k}^* G_{2k} \\
A_{2k+1} & = & G_{2k} J_{2k-1} G_{2k}^* & = & G_{2k+1}^* J_{2k+1} G_{2k+1} \\
 & \vdots & & &
\end{array}$$

Matrice J_k definirane ovim algoritmom su permutacijski slične matrici J. ∎

Da izbjegnemo eksplicitno množenje matrica $G^* JG$, koje može dovesti do gubitka točnosti, korisno je prethodni algoritam formulirati kao implicitni algoritam.

6.1.3. Implicitni algoritam iterirane indefinitne dekompozicije

Implicitna varijanta algoritma bazira se na indefinitnoj QR dekompoziciji. Nažalost, dobar izbor pomaka za ovaj algoritam temelji se na ocjenama za veličinu hiperbolnih singularnih vrijednosti specijalnih po dijelovima trokutastih matrica, koje se javljaju u algoritmu. Naravno, pomak se može birati i približno, pazeći da ne promijenimo matricu inercije J. Ipak, treba reći da se primjenom pomaka može izgubiti točnost, tako da u njihovom izboru treba biti vrlo oprezan.

Algoritam 6.1.4.

Neka su zadane matrica G i matrica $J = \mathrm{diag}(\pm 1)$. Slijedeći algoritam računa hiperbolne singularne vrijednosti matrice G, odnosno svojstvene vrijednosti matrice

$A = G^*JG$ naizmjeničnim primjenama QR i JQR dekompozicije. Hiperbolne singularne vrijednosti dobivaju se u neparnim koracima algoritma.

start $G_1^* = G^*$, $J_1 = J$;

$$\begin{array}{rcl c rcl}
G_1^* &=& Q_1 G_2 & , & Q_1^* Q_1 &=& I \\
G_2^* &=& Q_2 G_3 & , & Q_2^* J_1 Q_2 &=& J_3 \\
&\vdots& & & & & \\
G_{2k}^* &=& Q_{2k} G_{2k+1} & , & Q_{2k}^* J_{2k-1} Q_{2k} &=& J_{2k+1} \\
G_{2k+1}^* &=& Q_{2k+1} G_{2k+2} & , & Q_{2k+1}^* Q_{2k+1} &=& I \\
&\vdots& & & & &
\end{array}$$

U prethodnom algoritmu su sve matrice G_{2k+1} blok gornjetrokutaste s dijagonalnim blokovima veličine 1×1 ili 2×2, a matrice G_{2k} su gornjetrokutaste. Matrice J_k su permutacijski slične matrici J. ■

Pokažimo da transformacije u neparnim koracima prethodnog algoritam čuvaju svojstvene vrijednosti polazne matrice.

Teorem 6.1.1.

Matrice A_{2k+1} i A_1 generirane algoritmom 6.1.4. su unitarno slične.

Dokaz:

Dokaz teorema dobivamo uvrštavanjem jedne jednakosti u drugu. Vrijedi

$$\begin{aligned}
A_{2k+1} &= G_{2k+1}^* J_{2k+1} G_{2k+1} \\
&= G_{2k} Q_{2k}^{-*} J_{2k+1} Q_{2k}^{-1} G_{2k}^* \\
&= G_{2k} J_{2k-1} G_{2k}^* \\
&= Q_{2k-1}^{-1} G_{2k-1}^* J_{2k-1} G_{2k-1} Q_{2k-1}^{-*} \\
&= Q_{2k-1}^{-1} A_{2k-1} Q_{2k-1}^{-*} \\
&\ \vdots \\
&= Q_{2k-1}^{-1} \cdots Q_1^{-1} A_1 Q_1^{-*} \cdots Q_{2k-1}^{-*} \\
&= (Q_1 \cdots Q_{2k-1})^* A_1 (Q_1 \cdots Q_{2k-1}) \quad .
\end{aligned}$$

■

Za iteriranu dekompoziciju Choleskog, dokazano je da niz matrica A_k, uz neke uvjete, konvergira prema dijagonalnoj matrici. Slijedeći teorem dat će odgovor na pitanje da li to isto vrijedi i kod iterirane indefinitne dekompozicije, tj. da li vrijedi

$$\lim_{k\to\infty} A_{2k+1} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \quad ?$$

Teorem 6.1.2.

Ako je niz matrica G_{2k+1} generiran algoritmom 6.1.4. takav da za $k \geq k_0$, $k_0 \in \mathbf{N}$, ne dolazi do promjene strukture, tj. preklapanja 2×2 dijagonalnih blokova u matricama G_{2k+1}, onda niz matrica A_1, A_3, ..., A_{2k+1}, ..., konvergira kad konvergira i bazični QR algoritam i vrijedi

$$\lim_{k \to \infty} A_{2k+1} = D \quad ,$$

gdje je D blok–dijagonalna matrica s dijagonalnim blokovima veličine 1×1 ili 2×2.

Dokaz:

Dokaz se osniva na vezi iterirane indefinitne dekompozicije i bazičnog QR algoritma (bez pomaka).

Neka je $M_1 = A$. Bazični QR algoritam za matricu A generira niz matrica: hermitske M_k, unitarne U_k i gornjetrokutaste R_k, takve da vrijedi

$$\begin{aligned}
M_1 &= U_1 R_1 \\
M_2 &= R_1 U_1 = U_1^* M_1 U_1 = U_2 R_2 \\
M_3 &= R_2 U_2 = U_2^* M_2 U_2 = U_3 R_3 \\
&\vdots
\end{aligned} \tag{6.1.1}$$

Indefinitna dekompozicija počinje s $M_1 = G_1^* J_1 G_1$. Prvi korak algoritma 6.1.4. računa $G_1^* = Q_1 G_2$, gdje je Q_1 unitarna. Zbog toga što je matrica G_1 blok–gornjetrokutasta, a G_2 gornjetrokutasta i njihov produkt je blok–gornjetrokutast, pa za polaznu matricu M_1 vrijedi

$$M_1 = Q_1 G_2 J_1 G_1 = Q_1 \cdot \text{ blok gornjetrokutasta matrica } \quad .$$

Tvrdimo da postoji blok–dijagonalna unitarna matrica Q_1' takva da je

$$G_2 J_1 G_1 = Q_1' T_1 \quad ,$$

a matrica T_1 je gornjetrokutasta s pozitivnim dijagonalnim elementima, što slijedi zbog jedinstvenosti QR dekompozicije. Zbog toga za matrice R_1 i U_1 iz relacije (6.1.1) vrijedi

$$U_1 = Q_1 Q_1' \quad , \quad R_1 = T_1 \quad . \tag{6.1.2}$$

Treći korak algoritma, uz korištenje relacije (6.1.2) daje

$$\begin{aligned}
A_3 &= G_3^* J_3 G_3 = G_2 J_1 G_2^* \\
&= Q_1^* G_1^* J_1 G_1 Q_1 = Q_1^* M_1 Q_1 = Q_1' M_2 Q_1'^* \quad .
\end{aligned}$$

Ako krenemo od posljednje jednakosti u prethodnoj relaciji prema prvoj, a iz (6.1.1) uvažimo da je $G_3^* = Q_3 G_4$, dobivamo

$$M_2 = Q_1'^* A_3 Q_1' = Q_1'^* G_3^* J_3 G_3 Q_1' = Q_1'^* Q_3 G_4 J_3 G_3 Q_1' \quad .$$

Ako G_3 i Q_1' imaju nepresjecajuće 2×2 dijagonalne blokove, tvrdimo da postoji blok–dijagonalna unitarna matrica Q_3' takva da je

$$G_4 J_3 G_3 Q_1' = Q_3' T_2 \quad ,$$

pri čemu je T_2 gornjetrokutasta s pozitivnim dijagonalnim elementima. Ovdje smo iskoristili činjenicu da je red dijagonalnih blokova matrice Q_3' manji ili jednak 2. Dokaz slijedi iz jedinstvenost QR dekompozicije, pa za matrice U_2 i R_2 iz (6.1.1) dobivamo

$$U_2 = Q_1'^* Q_3 Q_3' \quad , \quad R_2 = T_2 \quad . \tag{6.1.3}$$

Pokažimo da postoji veza između matrice A_5 i matrice M_3. Korištenjem (6.1.3) izlazi

$$\begin{aligned} A_5 &= G_5^* J_5 G_5 = G_4 J_3 G_4^* \\ &= Q_3^* G_3^* J_3 G_3 Q_3 = Q_3^* Q_1' M_2 Q_1'^* Q_3 \\ &= Q_3' R_2 Q_1'^* Q_3 = Q_3' R_2 U_2 Q_3'^* = Q_3' M_3 Q_3'^* \quad . \end{aligned}$$

Uočimo da su matrice A_{2k-1} i M_k unitarno slične. Budući da su Q_k' blok–dijagonalne s dijagonalnim blokovim veličine 1×1, ili 2×2, ako matrice M_k konvergiraju prema dijagonalnoj, onda A_{2k-1} konvergiraju prema blok–dijagonalnoj s dimenzijama dijagonalnih blokova 1×1, ili 2×2. Prema tome, indukcijom po indeksu k pokazujemo da vrijedi zaključak o istovremenoj konvergenciji bazičnog QR algoritma i iterirane implicitne indefinitne dekompozicije. Uvjeti konvergencije bazičnog QR algoritma dani su u [31], poglavlje 8–6. ■

Promatramo li algoritam 6.1.4., uočavamo da su

(i) neparni koraci u algoritmu — QR dekompozicije

$$G_{2k+1}^* = Q_{2k+1} G_{2k+2} \quad , \quad Q_{2k+1}^* Q_{2k+1} = I \quad ;$$

(ii) parni koraci u algoritmu — JQR dekompozicije

$$G_{2k}^* = Q_{2k} G_{2k+1} \quad , \quad Q_{2k}^* J_{2k-1} Q_{2k} = J_{2k+1} \quad .$$

6.2. Bidijagonalizacija faktora G

Na konferenciji International Workshop on Accurate Eigensolving and Applications (Split, 1996.), Parlett i Dhillon izložili su svoje još nepublicirane rezultate o točnom rješavanju svojstvenog problema za tridijagonalne matrice.

Teorem 6.2.1. (Parlett)

Neka je T tridijagonalna matrica, $T = B^*JB$, takva da je

$$B = \begin{bmatrix} a_1 & b_1 & & & \\ & a_2 & b_2 & & \\ & & \ddots & \ddots & \\ & & & a_{n-1} & b_{n-1} \\ & & & & a_n \end{bmatrix} \quad , \quad a_i, b_i \neq 0 \quad ,$$

$J = \operatorname{diag}(j_{11}, \ldots, j_{nn})$. *Neka je*

$$\begin{aligned} T &= U\Lambda U^* \quad , \quad \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \quad , \quad U^*U = I \quad , \\ B^* &= U\Sigma V^* \quad , \quad \Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n) \quad , \quad V^*JV = J \quad , \end{aligned}$$

pri čemu su λ_i svojstvene vrijednosti matrice T, a $\sigma_i > 0$ hiperbolne singularne vrijednosti matrice B^*. Tada je

$$\lambda_i = j_{ii}\sigma_i^2 \quad ,$$

i mala relativna promjena u elementima a_i i b_i izaziva malu relativnu promjenu u hiperbolnim singularnim vrijednostima matrice B^*. ■

Za bidijagonalne matrice, Fernando i Parlett [18], [19] su pokazali da diferencijalni *qd* algoritam, relativno točno računa singularne vrijednosti matrice B^*. Diferencijalni *qd* algoritam, zapravo nije ništa drugo, nego pažljivo zapisani algoritam iterirane dekompozicije Choleskog.

Dhillon je u Splitu pokazao kako se taj algoritam modificira i u slučaju matrice B^*JB. Izabere se pomak tako da matrica $B^*JB + \tau I$ bude pozitivno ili negativno definitna. Zatim se primijeni diferencijalni *qd* algoritam. Za svaki skup svojstvenih vrijednosti koje imaju malu relativnu udaljenost, postupak se ponavlja, sve dok ne dobijemo velike relativne udaljenosti. Da bi se lako pronalazili odgovarajući pomaci, dobro je znati baš matricu B.

Neka je dana matrica G. G možemo bidijagonalizirati bez eksplicitnog formiranja matrice G^*JG slijedećim algoritmom.

Algoritam 6.2.1.

Neka su dane matrice $G \in \mathbf{C}^{m\times n}$ i $J = \operatorname{diag}(j_{11}, \ldots, j_{mm})$, $j_{ii} \in \{1, -1\}$, takve da je matrica G^*JG regularna. Slijedeći algoritam računa matrice

$$G = P_1^* VBU^*P_2 = P_1^*V \begin{bmatrix} B_1 \\ 0 \end{bmatrix} U^*P_2 \quad ,$$

pri čemu je B_1 gornja bidijagonalna, P_1 i P_2 su permutacije, U je unitarna, a za matricu V vrijedi

$$V^*J_1V = J_1 \quad , \quad J_1 = P_1^*JP_1 \quad .$$

start matrica G;

Ako je neka pivotna strategija odabrala 1×1 korak onda:

- *slijeva: reduciramo J–unitarnim transformacijama prvi stupac (nakon pivotiranja) matrice G;*
- *zdesna: reduciramo (unitarno) prvi redak (na dva elementa);*

Ako je neka pivotna strategija odabrala 2×2 korak onda:

- *slijeva: reduciramo J–unitarnim transformacijama prva dva stupca matrice G na 2×2 dijagonalni blok;*
- *zdesna: reduciramo (unitarno) prva dva retka na 2×2 dijagonalni blok;*
- *zdesna: unitarno reduciramo 2×2 dijagonalni blok na gornjetrokutastu matricu.* ■

6.3. Iterativno profinjavanje svojstvenih vrijednosti

Poznato je da je implicitna hiperbolna Jacobijeva metoda vrlo točna u slučajevima kad matrice imaju nisku skaliranu uvjetovanost. Međutim, metoda mnogo sporije konvergira, nego, na primjer, QR metoda. Veselić je u [50] pokazao da ne mora postojati simetrična permutacija redaka i stupaca hermitske matrice A, takva da QR metoda relativno točno računa svojstvene vrijednosti.

Problem je spajanje dobrih osobina te dvije metode: brzine QR metode i točnosti implicitne Jacobijeve metode. Neka je matrica U, matrica približnih svojstvenih vektora matrice A, izračunata, na primjer, QR metodom. Pretpostavimo da su svojstveni vektori ortonormalni do na točnost računala. Ako je skalirana uvjetovanost matrice mala, a uvjetovanost matrice velika, možemo očekivati da su svojstvene vrijednosti početne matrice loše izračunate.

Za definiciju problema, potrebna nam je definicija skalirano dijagonalno dominantne matrice, koju su uveli Barlow i Demmel u [2].

Definicija 6.3.1. *Hermitska matrica A oblika $A = D(J+N)D$ je η–skalirano dijagonalno dominantna, ako je D dijagonalna pozitivno definitna, $J = \operatorname{diag}(\pm 1)$, a N ima nule na dijagonali i vrijedi $\|N\|_2 \leq \eta < 1$.* ■

Za elemente η–skalirano dijagonalno dominantne matrice vrijedi slijedeća propozicija:

Propozicija 6.3.1.

Za η–skalirano dijagonalno dominantne matrice vrijedi

$$\max_{i \neq k} \frac{|a_{ik}|}{\sqrt{|a_{ii}|\,|a_{kk}|}} \leq \eta < 1 \quad .$$

Dokaz:

Za elemente matrice A vrijedi ocjena

$$\frac{|a_{ik}|}{\sqrt{|a_{ii}|\,|a_{kk}|}} = \frac{|d_{ii}d_{kk}n_{ik}|}{d_{ii}d_{kk}\sqrt{|j_{ii}|\,|j_{kk}|}} = \frac{|n_{ik}|}{\sqrt{|j_{ii}|\,|j_{kk}|}} = |n_{ik}| \quad . \tag{6.3.1}$$

Također, vrijedi

$$\max_{i=1,\ldots,n} |n_{ik}| \leq \sqrt{\sum_{i=1}^{n} |n_{ik}|^2} = \|Ne_k\|_2 \leq \sup_{\|x\|_2=1} \|Nx\|_2 = \|N\|_2 \quad ,$$

pri čemu je e_k, k–ti vektor standardne baze, što zajedno s (6.3.1) daje tvrdnju propozicije. ■

Propozicija 6.3.2.

Skalirano dijagonalno dominantna matrica je nesingularna.

Dokaz:

Ako pretpostavimo suprotno, tada postoji vektor x različit od nul–vektora, takav da vrijedi

$$(J + N)x = 0 \quad .$$

Odatle slijedi da je

$$\|x\|_2 > \|N\|_2\|x\|_2 \geq \|Nx\|_2 = \|Jx\|_2 = \|x\|_2 \quad ,$$

što je kontradikcija. ■

Slapničar je u [39] pokazao da za hermitske skalirane dijagonalno dominantne matrice vrijedi slijedeća propozicija:

Propozicija 6.3.3.

Ako je hermitska matrica A η–skalirano dijagonalno dominantna, hermitska indefinitna dekompozicija se može realizirati samo s 1×1 koracima. ■

Primijetimo da prethodna propozicija još uvijek ne kaže da će algoritam sigurno uzeti 1×1 pivot, već samo da on to može učiniti. Ipak, za neke pivotne strategije, ako je matrica A skalirano dijagonalno dominantna, hermitska indefinitna dekompozicija će sigurno uzeti 1×1 pivot. Jedna od takvih strategija je Bunch–Parlett starategija s potpunim pivotiranjem [11].

U Bunch–Parlett strategiji potpunog pivotiranja istaknute su tri veličine: konstanta $\alpha = (1 + \sqrt{17})/8$ i

$$\mu_0 = \max_{i \neq k} |a_{ik}| \quad , \quad \mu_1 = \max_{i} |a_{ii}| \quad .$$

Ta strategija izabrat će 1×1 pivot ako vrijedi

$$\mu_1 \geq \alpha \mu_0 \quad ,$$

a 2×2 pivot inače.

U slučaju skalirano dijagonalno dominantne matrice, neka su i, k indeksi gdje je postignuta maksimalna apsolutna vrijednost izvandijagonalnog elementa. Tada je

$$\alpha \mu_0 \leq \mu_0 = |a_{ik}| \leq \sqrt{|a_{ii}||a_{kk}|} \leq \max\{|a_{ii}|, |a_{kk}|\} \leq \mu_1 \quad .$$

Analogno, potpuno pivotiranje u indefinitnoj QR dekompoziciji faktora G skalirano dijagonalno dominantne matrice $A = G^* J G$, uvijek će birati 1×1 pivote. Problem s potpunim pivotiranjem na faktoru je samo broj aritmetičkih operacija. Ako je matrica $G \in \mathbf{C}^{m \times n}$, onda je za određivanje pivota potrebno $\mathcal{O}(m^2 n^2)$ aritmetičkih operacija (ako svaki puta računamo J skalarne produkte stupaca).

Jednom izračunate indefinitne skalarne produkte stupaca matrice G možemo pamtiti i ažurirati tijekom transformacija. Time se smanjuje broj aritmetičkih operacija za nalaženje pivota, ali postoji opasnost zbog kraćenja ažuriranih elemenata. Zbog toga treba povremeno obnavljati (ponovno računati) indefinitne skalarne produkte radnih dijelova stupaca matrice G.

Ipak, treba spomenuti da ovo nije elegantno rješenje problema. Otvoreni je problem pronalaženje pivotiranja koje bi korektno numerički detektiralo rang matrice (rank–revealing), koje zahtijeva reda veličine $\mathcal{O}(m^2 n)$ aritmetičkih operacija.

Vratimo se problemu iterativnog profinjavanja netočno izračunatih svojstvenih vrijednosti. Neka je zadana matrica A. Pretpostavimo da su nekom brzom, ali nedovoljno točnom metodom izračunati njeni (približni) svojstveni vektori. Pretpostavimo da je matrica izračunatih svojstvenih vektora U unitarna do na točnost računala.

Formirajmo matricu $A' = U^* A U$, za koju pretpostavljamo da je skalirano dijagonalno dominantna. Ako je poznata hermitska indefinitna dekompozicija matrice $A = G^* J G$, onda vrijedi

$$A' = U^* G^* J G U = G'^* J G' \quad .$$

Dakle, ne moramo eksplicitno formirati matricu A', nego odmah imamo njen faktor $G' = GU$.

Kao što smo već napomenuli, strategija potpunog pivotiranja birat će samo 1×1 pivote. Barlow i Demmel su u [2] pokazali da se malim perturabacijama η–skalirano dijagonalno dominantnih matrica malo mijenjaju i njihove svojstvene vrijednosti.

Teorem 6.3.1. (Barlow, Demmel)

Neka je $A = D(J+N)D \in \mathbf{C}^{n\times n}$ hermitska, η–skalirano dijagonalno dominantna matrica. Neka je δA njena hermitska perturbacija takva da vrijedi

$$\|D^{-1}\delta A D^{-1}\| = \nu \quad .$$

Pretpostavimo da je $A + \xi\delta A$ također η–skalirano dijagonalno dominantna matrica za sve $0 \le \xi \le 1$. Ako su $\lambda_1 \le \cdots \le \lambda_n$ svojstvene vrijednosti od A, a $\lambda_1' \le \cdots \le \lambda_n'$ svojstvene vrijednosti od $A + \delta A$, onda vrijedi

$$\exp\left(\frac{-\nu}{1-\eta}\right) \le \frac{\lambda_i'}{\lambda_i} \le \exp\left(\frac{\nu}{1-\eta}\right) \quad .$$

∎

Promotrimo pogreške u izračunatim elementima matrice A. Neka su g_k stupci u faktoru G matrice A, a $g_k' = g_k + \delta g_k$ stupci u faktoru $\widetilde{G}$ matrice $A+\delta A$. U teoremu 5.5.1. smo pokazali da, ako dekompoziciju možemo provesti samo trigonometrijskim i hiperbolnim rotacijama, onda je norma perturbacije svakog pojedinog stupca u matrici proporcionalna njegovoj euklidskoj normi

$$\|\delta g_k\| \le \varepsilon_k \|g_k\| \quad ,$$

gdje su ε_k definirani relacijom (5.5.1), tj. vrijedi

$$\varepsilon_k = \gamma^{k-1} \sum_{i=1}^{k} \mathrm{err}(p_i) \quad , \tag{6.3.2}$$

pri čemu je $\mathrm{err}(p_i)$ definiran relacijom (5.5.2). Za pogreške u indefinitnim skalarnim produktima stupaca dobivamo ocjenu

$$\begin{aligned} |[g_k', g_\ell'] - [g_k, g_\ell]| &= |[g_k, \delta g_\ell] + [\delta g_k, g_\ell] + [\delta g_k, \delta g_\ell]| \\ &\le |[g_k, \delta g_\ell]| + |[\delta g_k, g_\ell]| + |[\delta g_k, \delta g_\ell]| \\ &= |(J g_k, \delta g_\ell)| + |(J\delta g_k, g_\ell)| + |(J\delta g_k, \delta g_\ell)| \quad . \end{aligned}$$

Primjenom Cauchy–Schwarzove nejednakosti i zanemarivanjem člana $|(J\delta g_k, \delta g_\ell)|$ izlazi

$$\begin{aligned} |[g_k', g_\ell'] - [g_k, g_\ell]| &\le \|g_k\|\,\|\delta g_\ell\| + \|g_\ell\|\,\|\delta g_k\| \\ &\le \varepsilon_\ell \|g_k\|\,\|g_\ell\| + \varepsilon_k \|g_k\|\,\|g_\ell\| \\ &\le \varepsilon_{k\ell} \|g_k\|\,\|g_\ell\| \quad . \end{aligned}$$

Za skalirano dijagonalno dominantne matrice, uvijek možemo euklidsku normu napisati pomoću J–norme, jer J–norme stupaca nisu jednake 0. Dakle, postoji veličina $\omega_k > 1$, takva da je

$$\|g_k\| = \omega_k \sqrt{|\,\|g_k\|_J^2|} = \omega_k \|g_k\|_J \quad , \tag{6.3.3}$$

pa prethodna ocjena glasi

$$|[g_k', g_\ell'] - [g_k, g_\ell]| \leq \varepsilon_{k\ell}' \|g_k\|_J \, \|g_\ell\|_J = \varepsilon_{k\ell}' \sqrt{|a_{kk}| \, |a_{\ell\ell}|} \quad .$$

Da bismo dobili dobru ocjenu u prethodnoj relaciji, nužno je da je ω_k odozgo ograničen, jer ulazi u ocjenu (6.3.2), tj. vrijedi

$$\|\delta g_k\| \leq \varepsilon_k \|g_k\| = \varepsilon_k \omega_k \|g_k\|_J \quad . \tag{6.3.4}$$

Za matricu pogreške tada vrijedi

$$[|\delta A|]_{k\ell} = \varepsilon_{k\ell}' \sqrt{|a_{kk}| \, |a_{\ell\ell}|} \quad ,$$

pri čemu, ponovno, $\varepsilon_{k\ell}'$ ovisi o ω_k i ω_ℓ.

Skalirajmo matricu $|A|$, tako da joj dijagonalni elementi budu jednaki 1. Tada za skaliranu pogrešku, po elementima, vrijedi ocjena

$$[D^{-1} |\delta A| D^{-1}]_{k\ell} \leq \varepsilon_{k\ell}' \quad . \tag{6.3.5}$$

Uočimo da norma matrice definirane u relaciji (6.3.5) ovisi o brojevima ω_k.

Time smo dokazali korolar teorema 6.3.1.

Korolar 6.3.1.

Neka je $A = G^* J G$ skalirano dijagonalno dominantna matrica, a $G = P_1^* Q R P_2$ indefinitna QR dekompozicija matrice G. Neka za stupce faktora G vrijedi da su brojevi ω_k definirani relacijom (6.3.3) *odozgo omeđeni, tj. $\omega_k \leq \omega$ za $k = 1, \ldots, n$. Ako su matrice $\widetilde{R}$ i $\widetilde{Q}$ izračunate u aritmetici pomičnog zareza umjesto egzaktnih matrica R i Q, tj. ako je $\widetilde{G} = P_1^* \widetilde{Q} \widetilde{R} P_2$, onda je $\widetilde{A} = A + \delta A = \widetilde{G}^* J \widetilde{G}$. Ako su $\lambda_1 \leq \cdots \leq \lambda_n$ svojstvene vrijednosti od A, a $\lambda_1' \leq \cdots \leq \lambda_n'$ svojstvene vrijednosti od $\widetilde{A}$, onda vrijedi teorem* 6.3.1.*, pri čemu je ν norma matrice definirane relacijom* (6.3.5). ■

Dijagonalnim pivotiranjem, se u slučaju η–skalirano dijagonalno dominantnih matrica za $\eta \ll 1$, postiže i dobro skaliranje u recima matrice G. Pretpostavimo da smo nakon završetka dekompozicije izračunali faktor

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & r_{nn} \end{bmatrix} \quad .$$

Apsolutne vrijednosti J-skalarnih produkata prvog stupca sa svima ostalim jednake su $|r_{11}|\,|r_{1k}|$. Označimo s r_k, k–ti stupac u matrici R. Iz η–skalirane dijagonalne dominantnosti slijedi

$$\frac{|a_{1k}|}{\sqrt{|a_{11}||a_{kk}|}} = \frac{|r_{11}|\,|r_{1k}|}{|r_{11}|\,\|r_k\|_J} \leq \eta \quad ,$$

odnosno

$$|r_{1k}| \leq \eta \|r_k\|_J \leq \eta |r_{11}| \quad ,$$

pa prvi element u prvom retku dominira nad svim ostalima. Primjenjujući isto razmišljanje, na radne matrice manjih dimenzija koje se javljaju u algoritmu, slijedi da će dijagonalni elementi biti po apsolutnoj vrijednosti najveći elementi u svakom retku matrice R. Ovakav, skoro ortogonalan, faktor osigurat će brzinu i točnost implicitnog Jacobijevog algoritma.

7. Numeričko testiranje

U ovom poglavlju izlažu se rezultati numeričkog testiranja indefinitne QR dekompozicije. Svrha u koju se upotrebljava sama dekompozicija određuje i vrste testova, koji se provode na različitim skupinama matrica, a interesantne primjere izdvajamo u posebno poglavlje.

7.1. Testiranje dekompozicije

Test program pisan je u Turbo Pascalu 6.0 i testiran na osobnom računalu s Pentium procesorom na 120 MHz pod operacionim sistemom MS DOS 6.22. To je jedini široko rasprostranjen kompajler koji podržava IEEE `extended` aritmetiku.

IEEE `extended` tip sastoji se od 15–bitnog eksponenta i 64–bitne mantise, tj. preciznost računanja je $\varepsilon = 1.08420217248550443E{-}0019$. Taj tip podataka korišten je za kontrolu izračunatih vrijednosti. Za računanje, korišten je IEEE tip `single` (24–bitna mantisa, 8–bitni eksponent), tj. relativna preciznost računanja je $\varepsilon = 1.1920929E{-}07$.

Prva skupina testova osnovna je za samu dekompoziciju, Na primjer, testira se da li je finalna matrica Q J–unitarna, kontrolira se veličina

$$E_G = G - P_1\, f\ell(QR) P_2^* \quad ,$$

u relativnom i apsolutnom smislu i to po elementima i po pogreškama u J–normama stupaca. Slično, testira se i

$$E_A = A - P_2\, f\ell(R^* J_1 R) P_2^* \quad ,$$

u relativnom i apsolutnom smislu i to po elementima i po stupcima.

Uvjetovanost svake pojedine neunitarne transformacije mijenja uvjetovanost matrice Q. Kontroliraju se i ispisuju podaci o uvjetovanostima svake pojedinačne transformacije, ali i trenutno izračunate matrice Q. Testovi pokazuju da uvjetovanost matrice Q ne mora stalno rasti, već se može i blago smanjiti.

Za svaki primjer računa se i omjer normi matrica

$$\frac{\|G^* JG\|_E}{\|G^* G\|_E}\quad, \tag{7.1.1}$$

koji je uvijek manji od 1, što je pokazao Veselić u [51]. Što je on omjer bliži 1, to su vrijednosti J–normi stupaca bliže euklidskim, pa se može očekivati da su uvjetovanosti korištenih transformacija male. Osim toga, tijekom transformacija dolazi do promjene euklidske norme matrice $G^* G$, pa je uvijek naznačen takav omjer za polaznu matricu i za izračunati faktor.

Ako indefinitnu QR dekompoziciju koristimo kao korak u procesu nalaženja svojstvenih vrijednosti matrice $G^* JG$, onda je vrlo važno testirati koliko perturbacije u dekompoziciji mijenjaju svojstvene vrijednosti matrice.

Testovi svojstvenih vrijednosti provode se na slijedeći način. Zadajemo faktor G i matricu J. Smatramo da matrica $G^* JG$ nije egzaktno dostupna, nego je računamo iz G i J. U tom množenju može doći i do većeg gubitka točnosti, ako oduzimamo bliske velike brojeve, a konačan rezultat je mali broj. Nakon toga računamo i u `single` i u `extended` točnosti indefinitnu QR dekompoziciju i hermitsku indefinitnu dekompoziciju matrice $G^* JG$, te iz tako dobivenih faktora izračunamo svojstvene vrijednosti hiperbolnim Jacobijevim algoritmom, u sva četiri slučaja u `extended` točnosti. Odluka da se stalno koristi `extended` točnost motivirana je činjenicom da kontroliramo pogreške dekompozicije, pa tome ne treba još dodati i povećanu pogrešku dijagonalizacije.

7.2. Primjeri

U ovom odjeljku navodimo nekoliko interesantnih primjera. Ovakvi primjeri mnogo bolje pokazuju bît dekompozicije, nego, na primjer velika statistika sa slučajnim matricama. Naime, takva statistika, gotovo sigurno neće generirati primjere 4×4 transformacija.

Osim toga, ako je jedan primjer u takvoj statistici loš, na primjer daje veliku uvjetovanost matrice Q, on će podići prosjek za uvjetovanost matrice Q, čime neće dati realnu sliku o dekompoziciji, koja je dobra za mnogo drugih primjera. Edelman je pokazao da primjeri slučajnih matrica nisu dobri za testiranja. Mnogo je bolje za test primjere uzeti matrice koje će se takvom dekompozicijom obrađivati.

Primjer 7.2.1. Neka su zadane matrica

$$G = \begin{bmatrix} 2 & 4 & 1 & 2 \\ 1 & 3 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 5 & 1 & 1 \end{bmatrix}$$

i matrica skalarnog produkta

$$J = \mathrm{diag}(1, 1, -1, -1) \quad .$$

Uočimo da su sve J–norme stupaca jednake 0, pa će algoritam sigurno koristiti blok–transformacije.

Ako koristimo potpuno (Bunch-Parlett) pivotiranje, u prvom koraku izabiremo kao pivotne stupce, drugi i četvrti i dovodimo ih na prva dva mjesta. Nakon toga izabiru se 1×1 pivoti. Konačna matrica izračunata u `single` točnosti ima oblik

$$R_{\mathrm{jqr}} = \begin{bmatrix} 5.00000048E{+}00 & 6.00000322E{-}01 & 9.33333755E{-}01 & 2.00000510E{-}01 \\ 4.99999952E{+}00 & -6.00000024E{-}01 & 7.33333349E{-}01 & -1.99999824E{-}01 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 5.77350199E{-}01 & 5.77350140E{-}01 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 0.00000000E{+}00 & 5.77350259E{-}01 \end{bmatrix}$$

uz konačni

$$J_1 = \mathrm{diag}(1, -1, -1, 1) \quad .$$

Koristimo li množenje, pa hermitsku indefinitnu dekompoziciju, za R dobivamo

$$R_{\mathrm{sid}} = \begin{bmatrix} 1.73205090E{+}00 & 1.73205090E{+}00 & 5.77350259E{-}01 & 5.77350259E{-}01 \\ -1.73205090E{+}00 & 1.73205090E{+}00 & 0.00000000E{+}00 & 5.77350259E{-}01 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 5.77350259E{-}01 & 5.77350259E{-}01 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 0.00000000E{+}00 & 5.77350259E{-}01 \end{bmatrix}$$

uz isti J_1. Primijetimo razliku među dobivenim faktorima. Do takve razlike, koja ne mora biti posljedica grešaka ili kraćenja, može doći ako smo provodili blok–transformaciju. Naime, faktorizirana forma matrice čuva informaciju o pozitivnom i negativnom dijelu norme stupca, što u matrici $G^* J G$ nije vidljivo. Razlike proizlaze iz činjenice da svaki realan broj možemo na mnogo načina napisati kao razliku kvadrata dva realna broja.

U slijedećoj tablici imamo pregled uvjetovanosti korištenih hiperbolnih transformacija i blok transformacija. U prvom stupcu je red transformacije, u drugom uvjetovanost pojedinačne transformacije, u trećem uvjetovanost produkta svih dotad korištenih neunitarnih transformacija, a u četvrtom omjer uvjetovanosti matrica svih transformacija u dva susjedna koraka.

transf.	$\kappa(Q_k)$	$\kappa(\prod Q_k)$	omjer
2	$1.50000011920928955E{+}00$	$1.50000095367431641E{+}00$	$1.5000010E{+}00$
3	$1.97270812988281250E{+}01$	$2.02748908996582031E{+}01$	$1.3516585E{+}01$
2	$2.00000071525573730E{+}00$	$2.28166446685791016E{+}01$	$1.1253647E{+}00$

Označimo s nr početni omjer normi iz formule (7.1.1), a s nr_{sid} i nr_{jqr} završne omjere dobivene hermitskom indefinitnom dekompozicijom i indefinitnim QR algoritmom.

$$\begin{aligned} nr &= 1.33630633354187012E{-}01 \\ nr_{\mathrm{sid}} &= 9.81980443000793457E{-}01 \\ nr_{\mathrm{jqr}} &= 1.78261101245880127E{-}01 \quad . \end{aligned}$$

Kontroliramo još i po apsolutnoj vrijednosti najmanju ne-nula svojstvenu vrijednost. Algoritmi su dali slijedeće rezultate.

	λ_{min}	$\delta\lambda_{\mathrm{min}}/\lambda_{\mathrm{min}}$
Extended JQR	$1.97263881883089155E{-}01$	
Extended SID	$1.97263881883089155E{-}01$	$2.885506130778856E{-}18$
Single JQR	$1.97263894346219515E{-}01$	$6.317999139330042E{-}08$
Single SID	$1.97263876066934253E{-}01$	$2.948413489027556E{-}08$

Iz prethodne tablice vidljivo je da nema značajnije razlike u relativnoj točnosti izračunate najmanje svojstvene vrijednosti.

Primjer 7.2.2. Neka su zadane matrica

$$G = \begin{bmatrix} 2 & 1 & 4 & 1 \\ 1 & 1 & 3 & 2 \\ 1 & 1 & 0 & 2 \\ 2 & 1 & 5 & 1 \end{bmatrix}$$

i matrica skalarnog produkta

$$J = \mathrm{diag}(1, 1, -1, -1) \quad .$$

Zadana matrica G^*JG je ranga 2, što će otkriti oba algoritma. JQR kontrolira veličinu elemenata u matrici G i matrici G^*JG, tj. ako je red veličine elemenata u matrici G^*JG bitno manji nego u matrici G, onda se smatra da je došlo do degeneracije.

Faktor izračunat JQR dekompozicijom je

$$R_{\mathrm{jqr}} = \begin{bmatrix} 5.00000000E{+}00 & 5.00000000E{-}01 & 1.00000083E{-}01 & 2.00000226E{-}01 \\ 4.99999952E{+}00 & -5.00000060E{-}01 & -1.00000083E{-}01 & -1.99999928E{-}01 \end{bmatrix}$$

uz konačni

$$J_1 = \mathrm{diag}(1, -1) \quad .$$

Koristimo li množenje, pa hermitsku indefinitnu dekompoziciju, za R dobivamo

$$R_{\text{sid}} = \begin{bmatrix} 1.58113885E{+}00 & 1.58113885E{+}00 & 3.16227764E{-}01 & 6.32455528E{-}01 \\ -1.58113885E{+}00 & 1.58113885E{+}00 & 3.16227764E{-}01 & 6.32455528E{-}01 \end{bmatrix}$$

Uvjetovanost korištenih transformacija u JQR dekompoziciji je

transf.	$\kappa(Q_k)$	$\kappa(\prod Q_k)$	omjer
4	$2.85583648681640625E{+}01$	$2.85583896636962891E{+}01$	$2.8558390E{+}01$

Neka su oznake o omjerima normi iste kao u prethodnom primjeru:

$$\begin{aligned} nr &= 1.14108875393867493E{-}01 \\ nr_{\text{sid}} &= 2.93610095977783203E{-}01 \\ nr_{\text{jqr}} &= 1.54885903000831604E{-}01 \quad . \end{aligned}$$

Za najmanju (po apsolutnoj vrijednosti), ne-nula svojstvenu vrijednost algoritmi su dali slijedeće rezultate

	$\lambda_{\min}$	$\delta\lambda_{\min}/\lambda_{\min}$
Extended JQR	$-5.47722557505166114E{+}00$	
Extended SID	$-5.47722557505166116E{+}00$	$3.879767646873929E{-}18$
Single JQR	$-5.47722358910760084E{+}00$	$3.625821199225863E{-}07$
Single SID	$-5.47722569571193534E{+}00$	$2.202945132439253E{-}08$

što ponovno pokazuje da nije bilo značjnije razlike među algoritmima.

Primjer 7.2.3. Vratimo se matrici iz uvodnog primjera,

$$G = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1\cdot 10^{-11} & -1\cdot 10^{-11} \\ 1\cdot 10^{-11} & 1\cdot 10^{-11} \end{bmatrix}$$

i matrici skalarnog produkta

$$J = \operatorname{diag}(1, 1, -1, -1) \quad .$$

Zadanoj matrici, izračunavanjem G^*JG dolazi do gubitka ranga, pa je faktor iz JQR dekompozicije jednak

$$R_{\text{jqr}} = \begin{bmatrix} 1.41421354E{+}00 & 1.41421354E{+}00 \\ 0.00000000E{+}00 & 1.99999999E{-}11 \end{bmatrix}$$

uz konačni

$$J_1 = \mathrm{diag}(1, -1) \quad ,$$

a iz hermitske indefinitne dekompozicije

$$R_{\mathrm{sid}} = \begin{bmatrix} 1.41421354E{+}00 & 1.41421354E{+}00 \end{bmatrix}$$

uz $J = 1$.

Najmanju (po apsolutnoj vrijednosti) svojstvenu vrijednost, Jacobijev je algoritam pronašao na točnost računanja.

	$\lambda_{\min}$	$\delta\lambda_{\min}/\lambda_{\min}$
Extended JQR	$-2.00000000000000000E{-}22$	
Extended SID	$0.00000000000000000E{+}00$	$1.000000000000000E{+}00$
Single JQR	$-1.99999998000000005E{-}22$	$9.999999974772761E{-}09$
Single SID	$0.00000000000000000E{+}00$	$1.000000000000000E{+}00$

Primjer 7.2.4. U ovom primjeru pokazujemo da će, ako matrica GG^* skoro komutira s J, dekompozicija koristiti transformacije malih uvjetovanosti. Neka je

$$G = \begin{bmatrix} 1.0 & 1.0 \cdot 10^{-2} & 3.0 & 1.0 & 2.0 \cdot 10^{-3} \\ -1.0 \cdot 10^{-5} & 1.0 & 7.0 \cdot 10^{-5} & 1.0 \cdot 10^{-2} & -5.0 \\ -3.0 & 1.0 \cdot 10^{-4} & -1.0 & 1.0 & 4.0 \cdot 10^{-2} \\ 1.0 \cdot 10^{-7} & 1.0 & 8.0 \cdot 10^{-4} & -1.0 \cdot 10^{-4} & 4.0 \\ 1.0 & 2.0 \cdot 10^{-3} & 2.0 & 1.0 & 1.0 \cdot 10^{-7} \end{bmatrix}$$

i matrica skalarnog produkta

$$J = \mathrm{diag}(1, -1, 1, -1, 1) \quad .$$

transf.	$\kappa(Q_k)$	$\kappa(\prod Q_k)$	omjer
2	$1.00035738945007324E{+}00$	$1.01258850097656250E{+}00$	$1.0125885E{+}00$
2	$1.00030612945556641E{+}00$	$1.01258931488803711E{+}00$	$1.0000014E{+}00$
2	$1.00498723983764648E{+}00$	$1.01258969306945801E{+}00$	$9.9999976E{-}01$

Omjeri normi pokazuju vezu s uvjetovanostima matrice Q.

$$\begin{aligned} nr &= 9.99943017959594727E{-}01 \\ nr_{\mathrm{sid}} &= 9.99993562698364258E{-}01 \\ nr_{\mathrm{jqr}} &= 9.99993562698364258E{-}01 \quad . \end{aligned}$$

Kontroliramo još i najmanju (po apsolutnoj vrijednosti), ne-nula svojstvenu vrijednost. Algoritmi su dali slijedeće rezultate.

	$\lambda_{\min}$	$\delta\lambda_{\min}/\lambda_{\min}$
Extended JQR	$1.10413403397337956E{-}01$	
Extended SID	$1.10413403397337956E{-}01$	$0.000000000000000E{+}00$
Single JQR	$1.10413422558142460E{-}01$	$1.735369431031703E{-}07$
Single SID	$1.10413277840487874E{-}01$	$1.137152249805130E{-}06$

Primjer 7.2.5. U ovom primjeru zadana je matrica G kod koje dolazi do značajnog kraćenja pri računanju J–normi stupaca.

$$G = \begin{bmatrix} 10000.0 & 0.0 & 1.1 & 0.1 \\ 1.0 & -1.0 & 10000.0 & 1.0 \\ 10000.0 & 0.0 & 1.0 & 0.1 \\ 1.1 & -1.0 & 10000.0 & 1.0 \end{bmatrix}$$

i matrica skalarnog produkta

$$J = \mathrm{diag}(1, 1, -1, -1) \quad .$$

Najveća nevolja nastupa zbog J–norme prvog stupca koja je najveći element u matrici G^*JG, pa će taj stupac biti pivotni stupac. Zbog bliskih brojeva, hiperbolni tangens bit će vrlo blizu 1, a transformacije relativno loše uvjetovane. Ovo nam daje indikaciju da potpuno pivotiranje katkad nije idealno rješenje. Ali, jednako tako, taj prvi stupac u matrici uzrokuje kraćenje u računanju G^*JG, pa su tako dobiveni rezultati tek nešto bolji.

Faktor dobiven indefinitnom QR dekompozicijom je 0 matrica. Koristimo li množenje, pa hermitsku indefinitnu dekompoziciju, za R dobivamo

$$R_{\mathrm{sid}} = \begin{bmatrix} 1.10000002E{+}00 & 9.09226537E{-}02 & 2.16744156E{-}04 & -9.09091160E{-}02 \\ 0.00000000E{+}00 & 9.09226537E{-}02 & 1.18363164E{-}01 & -9.09091085E{-}02 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 1.18362971E{-}01 & -9.07427967E{-}02 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 0.00000000E{+}00 & 9.07428041E{-}02 \end{bmatrix}$$

uz konačni

$$J_1 = \mathrm{diag}(-1, 1, -1, 1) \quad .$$

Kontroliramo još i najmanju (po apsolutnoj vrijednosti), ne-nula svojstvenu vrijednost. Algoritmi su dali slijedeće rezultate.

	$\lambda_{\min}$	$\delta\lambda_{\min}/\lambda_{\min}$
Extended JQR	$-2.38418439797069585E{-}04$	
Extended SID	$-2.38418439820154388E{-}04$	$9.682473901171503E{-}11$
Single JQR	$0.00000000000000000E{+}00$	$1.000000000000000E{+}00$
Single SID	$-8.82152263818771864E{-}03$	$3.600016930609972E{+}01$

Primjer 7.2.6. U ovom primjeru pokazuje se da, iako reci u matrici G nisu dobro skalirani, indefinitna QR dekompozicija zajedno s Jacobijevim algoritmom relativno točno računa svojstvene vrijednosti. Neka je

$$G = \begin{bmatrix} 1.0 & 1.0 & 3.0 \\ -1.0\cdot 10^5 & 6.0\cdot 10^4 & 7.0\cdot 10^5 \\ -1.0\cdot 10^5 & 6.0\cdot 10^4 & -7.0\cdot 10^5 \\ 1.0 & 1.0 & 8.0 \\ 1.0 & 2.0 & 2.0 \end{bmatrix}$$

i matrica skalarnog produkta

$$J = \mathrm{diag}(1,-1,1,-1,1) \quad .$$

Razlog dobre dekompozicije leži u izboru transformacija niske uvjetovanosti

transf.	$\kappa(Q_k)$	$\kappa(\prod Q_k)$	omjer
4	$1.00002098083496094E{+}00$	$1.00002145767211914E{+}00$	$1.0000215E{+}00$
2	$3.20253968238830566E{+}00$	$3.20254111289978027E{+}00$	$3.2024724E{+}00$

Kontroliramo još i najmanju (po apsolutnoj vrijednosti), ne-nula svojstvenu vrijednost. Algoritmi su dali slijedeće rezultate.

	$\lambda_{\min}$	$\delta\lambda_{\min}/\lambda_{\min}$
Extended JQR	$4.97058823528287197E{+}00$	
Extended SID	$4.97058823528287198E{+}00$	$9.597434615633043E{-}19$
Single JQR	$4.97058785294118382E{+}00$	$7.692081300088443E{-}08$
Single SID	$3.08823504263356790E{+}00$	$3.786982754450955E{-}01$

Primjer 7.2.7. Svakako najzanimljiviji primjer je matrica sa skaliranim stupcima. Pokazat će se da algoritam množenja faktora ovdje uzrokuje potpuni raspad pri računanju svojstvenih vrijednosti. Nasuprot tome, korištenjem indefinitne QR dekompozicije, svojstvene vrijednosti izračunate su vrlo dobro. Neka je

$$G = \begin{bmatrix} 1.0 \cdot 10^{9} & 1.0 \cdot 10^{5} & 1.0 \cdot 10^{2} & 1.0 \cdot 10^{-1} \\ 1.0 \cdot 10^{5} & -1.0 \cdot 10^{4} & 1.0 \cdot 10^{-2} & 1.0 \cdot 10^{-3} \\ 1.0 \cdot 10^{3} & 1.0 \cdot 10^{2} & 1.0 \cdot 10^{-4} & 1.0 \cdot 10^{-5} \\ 1.0 \cdot 10^{-2} & -1.0 \cdot 10^{-1} & 0.0 & 1.0 \cdot 10^{-6} \end{bmatrix}$$

i matrica skalarnog produkta

$$J = \mathrm{diag}(1, 1, -1, -1) \quad .$$

Indefinitna QR dekompozicija ima oblik

$$R_{\mathrm{jqr}} = \begin{bmatrix} 1.00000000E{+}09 & 9.99990000E{+}04 & 1.00000098E{-}01 & 1.00000000E{+}02 \\ 0.00000000E{+}00 & -1.00095020E{+}04 & 9.90148168E{-}04 & 2.90926779E{-}11 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 1.98059697E{-}05 & -5.19612409E{-}11 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 0.00000000E{+}00 & -9.98651384E{-}10 \end{bmatrix}$$

uz konačni

$$J_1 = \mathrm{diag}(1, 1, -1, -1) \quad .$$

Koristimo li množenje, pa hermitsku indefinitnu dekompoziciju, za R dobivamo

$$R_{\mathrm{sid}} = \begin{bmatrix} 1.00000000E{+}09 & 9.99990000E{+}04 & 1.00000000E{+}02 & 1.00000106E{-}01 \\ 0.00000000E{+}00 & 1.00094570E{+}04 & 1.37908110E{-}05 & -9.90207889E{-}04 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 9.37757082E{-}03 & -1.03061029E{-}04 \\ 0.00000000E{+}00 & 0.00000000E{+}00 & 0.00000000E{+}00 & 1.05483712E{-}04 \end{bmatrix}$$

uz konačni

$$J_1 = \mathrm{diag}(1, 1, 1, -1) \quad ,$$

što znači da je izgubljena inercija matrice!

transf.	$\kappa(Q_k)$	$\kappa(\prod Q_k)$	omjer
2	$1.00000190734863281E{+}00$	$1.00000214576721191E{+}00$	$1.0000021E{+}00$

U ovom primjeru, razlike su velike čak i na velikim svojstvenim vrijednostima.

Zbog toga, evo potpune tablice svojstvenih vrijednosti.

	λ	$\delta\lambda/\lambda$
Extended JQR(1)	$1.00000001999880989E{+}18$	
Extended SID(1)	$1.00000001999880989E{+}18$	$0.000000000000000E{+}00$
Single JQR(1)	$1.00000000999981000E{+}18$	$9.998999685094406E{-}09$
Single SID(1)	$1.00000000999981000E{+}18$	$9.998999685094406E{-}09$
Extended JQR(2)	$1.00190117975918660E{+}08$	
Extended SID(2)	$1.00190117975918660E{+}08$	$0.000000000000000E{+}00$
Single JQR(2)	$1.00190129286123725E{+}08$	$1.128874313493379E{-}07$
Single SID(2)	$1.00189228432977734E{+}08$	$8.878549690302408E{-}06$
Extended JQR(3)	$-9.97500967359577651E{-}19$	
Extended SID(3)	$2.49498288390902419E{-}21$	$1.002501233548187E{+}00$
Single JQR(3)	$-9.97304586758241209E{-}19$	$1.968725923707817E{-}04$
Single SID(3)	$8.79494547162079642E{-}05$	$8.816979390908708E{+}13$
Extended JQR(4)	$-3.92276379471711299E{-}10$	
Extended SID(4)	$-3.92276379471328611E{-}10$	$9.755580773950742E{-}13$
Single JQR(4)	$-3.92276435760014222E{-}10$	$1.434914409035626E{-}07$
Single SID(4)	$-1.11254698921317278E{-}08$	$2.736130461669572E{+}01$

Primjer 7.2.8. Na kraju, specijalno se promatraju matrice koje nastaju kod rješavanja nekih realnih problema. Takve matrice su, na primjer matrice nastale primjenom konačnih elemenata za rješavanje problema rotirajućeg štapa. Za taj problem, poznat je eksplicitni faktor.

$$G = \begin{bmatrix} G' \\ D \end{bmatrix} \quad ,$$

gdje je $D = \eta I$, a

$$G' = (n+1)^2 \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix} \quad .$$

Pri tome je

$$J = \begin{bmatrix} & I \\ -I & \end{bmatrix} \quad .$$

Ako za η stavimo $\eta = \sqrt{100}$, što je otprilike najmanja svojstvene vrijednost, možemo pratiti porast uvjetovanosti matrica Q u dekompoziciji i točnost izračunavanja svojstvenih vrijednosti.

red mat.	$\kappa(\prod Q_k)$	JQR $\delta\lambda_{\min}/\lambda_{\min}$	SID $\delta\lambda_{\min}/\lambda_{\min}$
4×2	$1.979533195E{+}01$	$1.379134092563259E{-}07$	$6.627695397779842E{-}08$
6×3	$3.203159714E{+}01$	$1.158705606423956E{-}06$	$4.675001215865363E{-}07$
8×4	$4.360372162E{+}01$	$9.706378312613812E{-}07$	$8.326332931270635E{-}06$
10×5	$5.562741470E{+}01$	$2.379783349716876E{-}06$	$7.232162566855703E{-}06$
12×6	$6.695472717E{+}01$	$1.171652582334896E{-}05$	$4.833859271056691E{-}05$
14×7	$7.703354645E{+}01$	$8.655885862313956E{-}06$	$5.006407064179246E{-}06$
16×8	$8.599296570E{+}01$	$2.171218769530409E{-}05$	$8.170545673631054E{-}05$
18×9	$9.382825470E{+}01$	$7.531580501519689E{-}06$	$2.370512022900604E{-}04$
20×10	$1.006274109E{+}02$	$6.529128551636577E{-}06$	$1.219669071718377E{-}05$

Ako promatramo matrice 20×10 i mijenjamo η, onda vrijedi

η^2	$\kappa(\prod Q_k)$	JQR $\delta\lambda_{\min}/\lambda_{\min}$	SID $\delta\lambda_{\min}/\lambda_{\min}$
0	$1.0000000000E{+}00$	$3.174703636193478E{-}07$	$9.497003157842436E{-}06$
100	$1.0062741089E{+}02$	$6.529128551636577E{-}06$	$1.219669071718377E{-}05$
97	$4.2600036621E{+}02$	$1.062856433074080E{-}05$	$1.847551266282924E{-}03$
96.5	$9.4641204834E{+}02$	$8.620910995066906E{-}05$	$4.828088182873136E{-}03$

Literatura

[1] S. T. Alexander, C.-T. Pan, and R. J. Plemmons, *Analysis of a recursive least squares hyperbolic rotation algorithm for signal processing*, Linear Algebra Appl., 98 (1988), pp. 3–40.

[2] J. L. Barlow and J. W. Demmel, *Computing accurate eigensystems of scaled diagonally dominant matrices*, SIAM J. Numer. Anal., 27 (1990), pp. 762–791.

[3] A. W. Bojanczyk, R. Onn, and A. O. Steinhardt, *Existence of the hyperbolic singular value decomposition*, Linear Algebra Appl., 185 (1993), pp. 21–30.

[4] A. W. Bojanczyk and A. O. Steinhardt, *The hyperbolic transformations in signal processing and control*, ARO Report 92–1, U.S. Army Research Office, Transactions of the Ninth Army Conference on Applied Mathematics and Computing, University of Minnesota, Minneapolis, June 18–21, 1991., Research Triangle Park, Oct. 1991.

[5] K. A. Braun and T. Lunde Johnsen, *Hypermatrix generalization of the Jacobi–and Eberlein–method for computing eigenvalues and eigenvectors of Hermitian or non–Hermitian matrices*, Comput. Methods Appl. Mech. Engrg., 4 (1974), pp. 1–18.

[6] O. E. Brønlund and T. Lunde Johnsen, *QR factorization of partitioned matrices*, Comput. Methods Appl. Mech. Engrg., 3 (1974), pp. 153–172.

[7] J. R. Bunch, *Analysis of the diagonal pivoting method*, SIAM J. Numer. Anal., 8 (1971), pp. 656–680.

[8] ——, *Partial pivoting strategies for symmetric matrices*, SIAM J. Numer. Anal., 11 (1974), pp. 521–528.

[9] J. R. Bunch and L. C. Kaufman, *Some stable methods for calculating inertia and solving symmetric linear systems*, Math. Comp., 31 (1977), pp. 163–179.

[10] J. R. Bunch, L. C. Kaufman, and B. N. Parlett, *Decomposition of a symmetric matrix*, Numer. Math., 27 (1976), pp. 95–109.

[11] J. R. Bunch and B. N. Parlett, *Direct methods for solving symmetric indefinite systems of linear equations*, SIAM J. Numer. Anal., 8 (1971), pp. 639–655.

[12] A. Bunse-Gerstner, *An analysis of the HR algorithm for computing the eigenvalues of a matrix*, Linear Algebra Appl., 35 (1981), pp. 155–173.

[13] S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Surveys and Reference Works in Mathematics, Pitman, London, 1979.

[14] S. Chandrasekaran and A. H. Sayed, *Stabilizing the generalized Schur algorithm*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 950–983.

[15] G. Cybenko and M. W. Berry, *Hyperbolic Householder algorithms for factoring structured matrices*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 499–520.

[16] G. Dietrich, *A new formulation of the hypermatrix Householder–QR decomposition*, Comput. Methods Appl. Mech. Engrg., 9 (1976), pp. 273–280.

[17] Z. Drmač, *Computing the Singular and the Generalized Singular Values*, PhD thesis, FernUniversität–Gesamthochschule, Hagen, 1994.

[18] K. V. Fernando and B. N. Parlett, *Differential qd algorithms*, in Numerical Linear Algebra: Proceedings of the Conference on Numerical Linear Algebra and Scientific Computation, Kent (Ohio), March 13–14, 1992., L. Reichel, A. Ruttan, and R. S. Varga, eds., Berlin, New York, 1993, Walter de Gruyter, pp. 75–100.

[19] ——, *Accurate singular values and differential qd algorithms*, Numer. Math., 67 (1994), pp. 191–229.

[20] ——, *Implicit Cholesky algorithms for singular values and vectors of triangular matrices*, Numer. Linear Algebra Appl., 2 (1995), pp. 507–531.

[21] W. M. Gentleman, *Error analysis of QR decompositions by Givens transformations*, Linear Algebra Appl., 10 (1975), pp. 189–197.

[22] I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and Indefinite Scalar Products*, vol. 8 of Operator Theory: Advances and Applications, Birkhäuser, Basel, 1983.

[23] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 2nd ed., 1989.

[24] N. J. Higham, *Stability of the diagonal pivoting method with partial pivoting*, Numerical Analysis Report 265, University of Manchester, Department of Mathematics, July 1995.

[25] ——, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1996.

[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[27] ——, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[28] T. Kailath and A. H. Sayed, *Displacement structure: Theory and applications*, SIAM Rev., 37 (1995), pp. 297–386.

[29] R. Onn, A. O. Steinhardt, and A. W. Bojanczyk, *The hyperbolic singular value decomposition and applications*, IEEE Trans. Acoust. Speech Signal Process., ASSP–39 (1991), pp. 1575–1588.

[30] H. Park and V. Hari, *A real algorithm for the Hermitian eigenvalue decomposition*, BIT, 33 (1993), pp. 158–171.

[31] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice–Hall, Englewood Cliffs, N.J., 1980.

[32] B. N. Parlett and I. S. Dhillon, *Fernando's solution to Wilkinson's problem: an application of double dactorization.* Unpublished manuscript, University of California, Berkeley, Oct. 1996.

[33] C. M. Rader and A. O. Steinhardt, *Hyperbolic Householder transforms*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 269–290.

[34] H. Rutishauser, *Über eine kubisch konvergente Variante der LR–Transformation*, Z. Angew. Math. Mech., 40 (1960), pp. 49–54.

[35] ——, *Lectures on Numerical Mathematics*, Birkhäuser Verlag, Boston, 1990.

[36] R. S. Schreiber and B. N. Parlett, *Block reflectors: Computation and applications*, in Computing Methods in Applied Sciences and Engineering, VII, Versailles, 1985., R. Glowinski and J.-L. Lions, eds., North–Holland, New York, London, Sydney, 1986, pp. 71–79.

[37] ——, *Block reflectors: Theory and computation*, SIAM J. Numer. Anal., 25 (1988), pp. 189–205.

[38] I. Slapničar, *Accurate Symmetric Eigenreduction by a Jacobi Method*, PhD thesis, FernUniversität–Gesamthochschule, Hagen, 1992.

[39] ——, *Componentwise analysis of real symmetric and Hermitian decomposition.* Unpublished manuscript, University of Split, 1995.

[40] I. Slapničar and K. Veselić, *Perturbations of the eigenprojections of a factorised Hermitian matrix*, Linear Algebra Appl., 218 (1995), pp. 273–280.

[41] G. W. Stewart, *Perturbation bounds for the QR factorization of a matrix*, SIAM J. Numer. Anal., 14 (1977), pp. 509–518.

[42] ——, *On the perturbation of LU, Cholesky and QR factorizations*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1141–1145.

[43] J.-G. Sun, *Perturbation bounds for the Cholesky and QR factorizations*, BIT, 31 (1991), pp. 341–352.

[44] ——, *Componentwise perturbation bounds for some matrix decompositions*, BIT, 32 (1992), pp. 702–714.

[45] ——, *Rounding–error and perturbation bounds for the Cholesky and* LDL^{T} *factorizations*, Linear Algebra Appl., 173 (1992), pp. 77–97.

[46] ——, *On perturbation bounds for the QR factorization*, Linear Algebra Appl., 215 (1995), pp. 95–111.

[47] A.-J. van der Veen, *A Schur method for low–rank matrix approximation*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 139–160.

[48] K. Veselić, *Some convergent Jacobi–like procedures for diagonalising J–symmetric matrices*, Numer. Math., 27 (1976), pp. 67–75.

[49] ——, *On a new class of elementary matrices*, Numer. Math., 33 (1979), pp. 173–180.

[50] ——, *A note on the accuracy of symmetric eigenreduction*, Electr. Trans. on Numer. Anal., 4 (1996), pp. 37–45.

[51] ——, *Perturbation theory for the eigenvalues of indefinite symmetric matrices (draft).* Unpublished manuscript, FernUniversität Hagen, June 1996.

[52] K. Veselić and I. Slapničar, *Floating–point perturbations of Hermitian matrices*, Linear Algebra Appl., 195 (1993), pp. 81–116.

[53] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Monographs on Numerical Analysis, Oxford University Press, Oxford, 1965. (Reprinted with corrections, 1977.).

[54] H. Zha, *A componentwise perturbation analysis of the QR decomposition*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1124–1131.

[55] ——, *A note on the existence of the hyperbolic singular value decomposition*, Linear Algebra Appl., 240 (1996), pp. 199–205.

Sažetak

Osnovni cilj ovog rada je konstrukcija indefinitne QR dekompozicije zadane pravokutne matrice G obzirom na indefinitni skalarni produkt zadan matricom J oblika $J = \mathrm{diag}(\pm 1)$.

U uvodu daje se pregled do sada poznatih rezultata i pregled rezultata radnje. U drugom poglavlju definira se indefinitni skalarni produkt i ističu njegove razlike obzirom na standardni skalarni produkt. Treće poglavlje posvećeno je J–unitarnim matricama koje služe za dekompoziciju matrice G: i to generalizacijama rotacija i generalizacijama Householderovih matrica.

U četvrtom poglavlju konstruktivno se pokazuje egzistencija indefinitne QR dekompozicije. Peto poglavlje sadrži obratnu analiza grešaka zaokruživanja i neke perturbacione teoreme za dekompoziciju.

Šesto poglavlje posvećeno je primjenama: generalizaciji flip–flap algoritma za traženje svojstvenih vrijednosti hermitskih indefinitnih matrica, bidijagonalizaciju faktora i profinjavanju svojstvenih vrijednosti matrica.

U posljednjem poglavlju dani su rezultati testiranja za razne matrice koje se javljaju u primjenama.

Summary

The main purpose of this work is construction of the indefinite QR decomposition of a given matrix G according to an indefinite scalar product matrix J.

In foreword, summary of known results and results of this thesis are given. The second chapter contains a brief summary of known results about indefinite scalar products. In the third chapter construction of elementary J–unitary matrices – rotations and reflections is given.

In the fourth chapter we give constructive proofs of the existence of an indefinite QR decomposition, using both rotations and reflections. The next chapter contains backward rounding error analysis of this decomposition and componentwise perturbation estimates.

In the sixth chapter we consider some possible applications of the JQR decomposition: flip–flap algorithm for Hermitian indefinite matrices, J–bidiagonalization and iterative refinement of eigenvalues.

The final part contains a brief discussion of numerical results.

Biografija

Rođena sam u Zagrebu 27. 12. 1963., gdje sam završila osnovnu školu i Matematičko–informatički obrazovni centar. Godine 1981. upisala sam studij matematike na Prirodoslovno–matematičkom fakultetu Sveučilišta u Zagrebu. Diplomirala sam 1986. godine na smjeru inženjer matematike, struka matematička informatika i statistika. Slijedeće godine upisala sam Zajednički postdiplomski studij prirodnih znanosti iz područja matematike. 1993. obranila sam magistarski rad pod naslovom "Računanje spektralne dekompozicije pozitivno definitne matrice".

Od 20. 3. 1995. do 31. 12. 1995. bila sam na znanstvenom usavršavanju na Katedri za matematičku fiziku, FernUniversität-a u Hagenu, Njemačka, pod vodstvom prof. dr. Krešimira Veselića, gdje sam započela izradu doktorske disertacije.

Učestvovala sam na nekoliko međunarodnih znanstvenih skupova. Bila sam dugogodišnji član znanstvenog projekta "Numeričke metode linearne algebre" Ministarstva znanosti i tehnologije Republike Hrvatske, pod vodstvom glavnog istraživača prof. dr. sc. Vjerana Harija. Obzirom za interes, ne samo za numeričke metode linearne algebre, nego i za širu primjenu numeričkih algoritama, uključila sam se u znanstveni projekt "Konstrukcija i realizacija numeričkih algoritama tipa projekcije s primjenama", pod vodstvom doc. dr. Mladena Rogine.

Zaposlena sam kao znanstevna novakinja u suradničkom zvanju asistenta na Katedri za matematiku i nacrtnu geometriju Fakulteta strojarstva i brodogradnje Sveučilišta u Zagrebu.