Schurova dekompozicija matrice

Mrkojević, Ana

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku
Sveučilišni preddiplomski studij matematike

Ana Mrkojević

Schurova dekompozicija matrice

Završni rad

Osijek, 2019.

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku
Sveučilišni preddiplomski studij matematike

Ana Mrkojević

Schurova dekompozicija matrice

Završni rad

Mentor: doc. dr. sc. Ivana Kuzmanović Ivičić

Osijek, 2019.

Sažetak

Faktorizacije matrica od velikog su značaja u teoriji matrica i uopće u numeričkoj linearnoj algebri. Jedan od osnovnih alata u analiziranju i numeričkom rješavanju problema svojstvenih vrijednosti je Schurova dekompozicija. Upravo nam ova dekompozicija daje odgovor na pitanje koliko najviše možemo unitarnom transformacijom sličnosti pojednostaviti matricu $\mathbf{A} \in \mathbb{C}^{n \times n}$.
Svojstvene vrijednosti i svojstveni vektori matrice $\mathbf{A}$ jednostavno se prenose na njoj sličnu matricu $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, ako je $\mathbf{B}x = \lambda x$, onda je $\mathbf{A}\mathbf{S}x = \lambda \mathbf{S}x$. Dakle, za računanje svojstvenih vrijednosti i svojstvenih vektora matrice $\mathbf{A}$ praktično je tražiti transformacije sličnosti koje će dati matricu $\mathbf{B}$ koja je jednostavnije strukture.
Unitarna transformacija sličnosti matrice $\mathbf{A}$ gornjetrokutastoj matrici dana je Schurovim postupkom dekompozicije matrice. Unitarna transformacija sličnosti $\mathbf{S}$ ima prednost zbog nekih teoretskih i numeričkih svojstava, kao što je jednostavnije računanje inverza matrice transformacije, budući vrijedi $\mathbf{S}^{-1} = \mathbf{S}^*$ te ne povećava normu pertubacije, čuva neke klase matrice. Primjerice, ako je $\mathbf{A}$ normalna, hermitska, antihermitska, unitarna, onda je matrica $\mathbf{B}$ ($\mathbf{B} = \mathbf{S}^*\mathbf{A}\mathbf{S}$) normalna, hermitska, antihermitska, unitarna, respektivno.
U ovom radu dani su rezultati o Schurovoj dekompoziciji. Postupak Schurove dekompozicije (realne i kompleksne) zasnovan na upotrebi Gram-Schmidtovog postupka ortogonalizacije ilustriran je kroz primjere, a značaj Schurovog teorema dodatno je istaknut njegovim posljedicama predstavljenim u obliku teorema s kojima smo se već imali priliku susresti.

Ključne riječi: svojstvene vrijednosti, dekompozicija matrice, Schurova dekompozicija, unitarna sličnost, Gram-Schmidtov postupak ortogonalizacije

Abstract

Matrix factorization has a great significance in matrix theory, as well as in Numerical Linear Algebra. Schur decomposition is considered as one of the basic tools, when it comes to analyzes and finding a numerical solution to an eigenvalue problem. This decomposition in particular is proven to give an answer to the question, to which level a matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ can be simplified by the use of unitary transformations of similarity. Eigenvalues and eigenvectors of matrix $\mathbf{A}$ are transferred to the matrix $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, which is similar to the matrix $\mathbf{A}$. So, if matrices $\mathbf{A}$ and $\mathbf{B}$ are similar, following equation $\mathbf{B}x = \lambda x$ implies that $\mathbf{A}\mathbf{S}x = \lambda \mathbf{S}x$. This process implies that it can be practical to use a matrix $\mathbf{B}$, which has a simplified structure in order to calculate the eigenvalues and eigenvectors of matrix $\mathbf{A}$.
The process of unitary transformation of similarity which transforms a matrix $\mathbf{A}$ to an upper triangular matrix is given by the Schur decomposition. Unitary transformation of similarity $\mathbf{S}$ is proven to have some advantages when it comes to its theoretical and numerical properties. This particular case allows as to calculate the inverse of matrix of transformation more easily, as $\mathbf{S}^{-1} = \mathbf{S}^*$, and the norm of perturbations is not enlarged during the process, as a class of matrix is preserved. For instance, if matrix $\mathbf{A}$ is normal, hermitian, antihermitian or unitary, than matrix $\mathbf{B}$ ($\mathbf{B} = \mathbf{S}^*\mathbf{A}\mathbf{S}$) is also normal, hermitian, antihermitian or unitary, respectively.
In this paper the results of Schur decomposition are represented. The process of Schur decomposition (real and complex) is based on the use of the Gram-Schmidt orthogonalization process. Examples represented in this paper have covered both real and complex cases, so that the process can be fully illustrated. Significance of Schur theorem is stated in terms of well known theorems which are all results of Schur decomposition process.

Key words: eigenvalues, eigenvectors, unitary matrix, unitary similarity, Gram-Schmidt process of ortogonalization

Sadržaj

1 Uvod

Schurova dekompozicija jedan je od alata koji nam omogućuje da pri numeričkom računanju matricu prevedemo u jednostavniji, gornjetrokutasti oblik iz kojeg se onda lako iščitavaju svojstvene vrijednosti matrice. Stoga ćemo najprije navesti osnovne pojmove vezane za svojstveni problem. Nadalje, kako je Schurova dekompozicija bazirana na unitarnoj sličnosti koja se numerički računa pomoću Gram-Schmidtovog postupka ortogonalizacije, ukratko ćemo navesti rezultate vezane za navedene pojmove a koji će nam biti potrebni.

1.1 Svojstvene vrijednosti

Definicija 1. *Za kompleksan skalar λ, $\lambda \in \mathbb{C}$, kažemo da je svojstvena vrijednost kvadratne matrice $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, ako postoji vektor $\boldsymbol{u} \in \mathbb{C}^n$, $\boldsymbol{u} \neq 0$, takav da je*

$$\boldsymbol{A}\boldsymbol{u} = \lambda \boldsymbol{u}.$$

Vektor $\mathbf{u}$ nazivamo svojstveni vektor matrice $\mathbf{A}$ za svojstvenu vrijednost λ. Skup svih svojstvenih vrijednosti matrice $\boldsymbol{A}$ nazivamo spektar matrice $\boldsymbol{A}$ i označavamo s $\sigma(\boldsymbol{A})$.

Dodatno, ako je $\mathbf{u}$ svojstveni vektor za svojstvenu vrijednost λ, onda je i svaki vektor $\alpha\mathbf{u}$, $\alpha \neq 0$, također svojstveni vektor za λ.
Jednakost iz definicije $\mathbf{Au} = \lambda\mathbf{u}$ možemo zapisati u sljedećem obliku:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = 0, \quad \mathbf{u} \neq 0$$

što povlači da matrica $\mathbf{A}$ - λ $\mathbf{I}$ mora biti singularna, tj. da je $det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

Definicija 2. *Neka je $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Polinom*

$$k_A(\lambda) = det(\boldsymbol{A} - \lambda \boldsymbol{I})$$

naziva se svojstveni ili karakteristični polinom matrice $\boldsymbol{A}$.

Iz izraza $det(\mathbf{A} - \lambda\mathbf{I}) = 0$ slijedi da su nultočke karakterističnog polinoma svojstvene vrijednosti matrice $\mathbf{A}$. Stoga iz Osnovnog teorema algebre[1] slijedi sljedeći teorem:

Teorem 1. *Svaka matrica $\boldsymbol{A}\in \mathbb{C}^{n\times n}$ ima n, ne nužno različitih, svojstvenih vrijednosti.*

Primjer 1. *Odredimo svojstvene vrijednosti matrice $\boldsymbol{A} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$.*

Rješenje. *Kako bismo odredili svojstvene vrijednosti matrice $\boldsymbol{A}$ rješavamo jednadžbu*

$$det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0.$$

Najprije trebamo odrediti svojstveni vektor $\boldsymbol{u} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ koji zadovoljava sljedeću jednakost:

$$\boldsymbol{A}\boldsymbol{u} = \lambda \boldsymbol{u}.$$

[1]Osnovni teorem algebre tvrdi da svaki polinom $P(x)$ stupnja $n \geq 1$ ima točno n kompleksnih korijena $x_1, \ldots, x_n$ i može se na jedinstven način (do na poredak faktora) zapisati u obliku $P(x) = a_n(x - x_1)$ $(x - x_2) \cdots (x - x_n)$.

$$\begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \lambda \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$\begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{bmatrix}$$

Sada nam još ostaje izračunati determinantu matrice koju smo dobili prethodnim računom, tj. riješiti jednadžbu $det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$.

$$\begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} = 10 + 7\lambda + \lambda^2 - 4 = 0$$

$$\lambda^2 + 7\lambda + 6 = 0$$

$$\lambda_1 = -1 \quad \lambda_2 = -6$$

Spektar matrice $\boldsymbol{A}$ *je* $\{-1, -6\}$.
Svojstveni vektor koji pripada svojstvenoj vrijednosti $\lambda_1 = -1$ *dobiva se rješenjem sljedećeg sustava:*

$$\begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = (-1) \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

odakle dobivamo da je

$$a_2 = 2a_1$$

pa je svojstveni vektor za svojstvenu vrijednost $\lambda_1 = -1$ *:*

$$\boldsymbol{u}_1 = \begin{bmatrix} a_1 \\ 2a_1 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Svojstveni vektor za $\lambda_2 = -6$ *je:*

$$\boldsymbol{u}_2 = \begin{bmatrix} -2a_2 \\ a_2 \end{bmatrix} = a_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

1.2 Sličnost matrica

Vektorski prostor na kojem imamo definiran skalarni produkt nazivamo unitarni prostor.

Definicija 3. *Za matricu* $\boldsymbol{U} \in \mathcal{M}_n$ *kažemo da je unitarna ako vrijedi:*

$$\boldsymbol{U}\boldsymbol{U}^* = \boldsymbol{I}$$

Ako je $\boldsymbol{U}$ *realna matrica (u tom slučaju je* $\boldsymbol{U}^* = \boldsymbol{U}^T$*), tada* $\boldsymbol{U}$ *nazivamo ortogonalna matrica.*

Definicija 4. *Neka su* $\boldsymbol{A}$, $\boldsymbol{B} \in \mathcal{M}_n$. *Kažemo da je matrica* $\boldsymbol{B}$ *slična matrici* $\boldsymbol{A}$ *ako postoji regularna matrica* $\boldsymbol{S} \in \mathcal{M}_n$ *takva da je*

$$\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}.$$

Budući da je sličnost relacija ekvivalencije[2] na skupu $\mathcal{M}_n$ slične matrice iako nisu jednake imaju ista mnoga važna svojstva kao što su:

[2]**Sličnost je relacija ekvivalencije** Neka su **A**, **B**, **C** $\in \mathcal{M}_n$. Tada vrijede sljedeća svojstva:

1. **Refleksivnost:** Matrica **A** je slična matrici **A**.
2. **Simetričnost:** Ako je matrica **A** slična matrici **B**, onda je i matrica **B** slična matrici**A**.
3. **Tranzitivnost:** Ako je matrica **A** slična matrici **B** i matrica **B** slična matrici **C**, onda je matrica **A** slična matrici **C**.

- **Slične matrice imaju jednake tragove.**
  Neka je matrica $\mathbf{A}$ slična matrici $\mathbf{B}$. Tada vrijedi:

$$tr(\mathbf{A}) = tr(\mathbf{B}).$$

Dokaz. Kako su matrice $\mathbf{A}$ i $\mathbf{B}$ slične, $\mathbf{A}, \mathbf{B} \in \mathcal{M}_n$, tada postoji regularna matrica $\mathbf{S}$ $\in \mathcal{M}_n$ takva da je $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$.

$$\begin{aligned} tr(\mathbf{B}) &= tr(\mathbf{S}^{-1}\mathbf{AS}) = tr(\mathbf{S}^{-1}(\mathbf{AS})) = tr((\mathbf{S})(\mathbf{S}^{-1}\mathbf{A})) \\ &= tr((\mathbf{SS}^{-1})\mathbf{A}) = tr(\mathbf{IA}) = tr(\mathbf{A}) \end{aligned}$$

□

- **Slične matrice imaju jednake determinante.**
  Neka je matrica $\mathbf{A}$ slična matrici $\mathbf{B}$. Tada vrijedi:

$$det(\mathbf{A}) = det(\mathbf{B}).$$

Dokaz. Kako su matrice $\mathbf{A}$ i $\mathbf{B}$ slične, $\mathbf{A}, \mathbf{B} \in \mathcal{M}_n$, tada postoji regularna matrica $\mathbf{S}$ $\in \mathcal{M}_n$ takva da je $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$.

$$\begin{aligned} det(\mathbf{B}) &= det(\mathbf{S}^{-1}\mathbf{AS}) = det(\mathbf{S})^{-1}\, det(\mathbf{A})\, det(\mathbf{S}) \\ &= det(\mathbf{S})^{-1}\, det(\mathbf{S})\, det(\mathbf{A}) = det(\mathbf{A}) \end{aligned}$$

□

- **Slične matrice imaju jednake svojstvene vrijednosti.**
  Neka je matrica $\mathbf{A}$ slična matrici $\mathbf{B}$. Tada su svojstveni polinomi matrica $\mathbf{A}$ i $\mathbf{B}$ jednaki, tj. $k_A(x) = k_B(x)$.

Dokaz. Neka su matrice $\mathbf{A}, \mathbf{B} \in \mathcal{M}_n$. Budući da su matrice $\mathbf{A}$ i $\mathbf{B}$ slične postoji regularna matrica $\mathbf{S} \in \mathcal{M}_n$ takva da je $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$. Tada vrijedi:

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

□

Budući da smo pokazali da slične matrice imaju jednake karakteristične polinome, a svojstvene vrijednosti matrice su nultočke njenog karakterističnog polinoma, stoga je očito da su i svojstvene vrijednosti sličnih matrica jednake.

Definicija 5. *Neka su dane matrice $\boldsymbol{A}$, $\boldsymbol{B} \in \mathcal{M}_n$. Kažemo da je matrica $\boldsymbol{A}$ unitarno slična matrici $\boldsymbol{B}$, ako postoji unitarna matrica $\boldsymbol{U} \in \mathcal{M}_n$ takva da vrijedi:*

$$\boldsymbol{A} = \boldsymbol{UBU}^*.$$

Ako u obzir uzimamo matricu $\boldsymbol{U}$ koja je realna, stoga i ortogonalna, tada kažemo da je matrica $\boldsymbol{A}$ ortogonalno slična matrici $\boldsymbol{B}$.
Kažemo da je matrica $\boldsymbol{A}$ unitarno dijagonalizabilna, ako je unitarno slična dijagonalnoj matrici. Za matricu $\boldsymbol{A}$ kažemo da je ortogonalno dijagonalizabilna, ako je ortogonalno slična dijagonalnoj matrici.

1.3 Gram-Schmidtov postupak ortogonalizacije

Gram-Scmidtov postupak ortogonalizacije nam osigurava egzistenciju ortonormirane baze u svakom konačnodimenzionalnom prostoru.
Ideja Gram-Schmidtovog postupka ortogonalizacije se temelji na tome da od linearno nezavisnog skupa sljedećim algoritmom dobijemo ortonormiranu bazu. Prvi vektor linearno nezavisnog skupa zadržimo u početnom obliku, sada taj vektor određuje pravac na koji ortogonalno projiciramo drugi vektor skupa i zatim tu projekciju odbijemo od drugog vektora. Time smo dobili vektor koji je okomit na prvi vektor. Sada ta dva vektora razapinju ravninu. Treći vektor projiciramo ortogonalno na ravninu, i odbijemo projekciju od njega. Time smo dobili vektor okomit na tu ravninu, odnosno, vektor okomit na prva dva konstruirana vektora, itd. Normiranjem svakog od tih međusobno okomitih vektora, dobijemo ortonormiranu bazu.

Definicija 6. *Neka je V unitaran prostor. Kažemo da su vektori $\boldsymbol{x}$, $\boldsymbol{y}$ iz V međusobno okomiti ili ortogonalni (oznaka $\boldsymbol{x} \perp \boldsymbol{y}$) ako je:*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0.$$

Konačan skup vektora $\{e_1, \dots, e_k\}$ je ortogonalan ako je $e_i \perp e_j$, $\forall i \neq j$. Skup $\{e_1, \dots, e_k\}$ je ortonormiran ako je ortogonalan i ako je $\|e_i\| = 1$, $\forall i = 1, \dots, k$.

Teorem 2. (Gram-Schmidtov postupak ortogonalizacije)
Neka je dan linearno nezavisan skup $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\}, k \in \mathbb{N}$, u unitarnom prostoru V. Tada postoji ortonormiran skup $\{e_1, \dots, e_k\}$ u V takav da je

$$[\{e_1, \dots, e_k\}] = [\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\}].$$

Dokaz. Konstrukciju skupa $\{e_1, \dots, e_k\}$ provodimo induktivno.
Za bazu indukcije stavimo da je

$$e_1 = \frac{1}{\|x_1\|} x_1$$

što je dobro definirano jer je $x_1 \neq 0$. Očito su e_1 i x_1 kolinearni pa razapinju isti potprostor.
Pretpostavimo sada da je skup $\{e_1, \dots, e_j\}$ ortonormiran i da vrijedi:

$$[\{e_1, \dots, e_j\}] = [\{x_1, \dots, x_j\}]$$

i sada konstruiramo e_{j+1}. Uvest ćemo pomoćni vektor:

$$f_{j+1} = x_{j+1} - \sum_{i=1}^{j} \langle x_{j+1}, e_i \rangle e_i.$$

Iz definicije okomitosti vidi se da je $f_{j+1} \perp e_i$, $\forall i = 1, \dots, j$.
Nadalje treba pokazati da vrijedi $[\{e_1, \dots, e_j, f_{j+1}\}] = [\{x_1, \dots, x_j, x_{j+1}\}]$.
Kako bismo pokazali da ova tvrdnja vrijedi, dovoljno je pokazati da generatori s jedne strane jednakost pripadaju potprostoru s druge strane i obratno.
Prema pretpostavci indukcije lako je uočiti da je $e_1, \dots, e_j \in [\{x_1, \dots, x_j, x_{j+1}\}]$, a prema definiciji vektora f_{j+1} lako je učiti da je $f_{j+1} \in [\{x_1, \dots, x_j, x_{j+1}\}]$.
Sada je očit i obrat, odnosno da generatori $x_1, \dots, x_j$ pripadaju potprostoru $[\{e_1, \dots, e_j, f_{j+1}\}]$.
Sada uočimo da je

$$x_{j+1} = f_{j+1} + \sum_{i=1}^{j} \langle x_{j+1}, e_i \rangle e_i.$$

Očito je da je skup $\{e_1, \ldots, e_j, f_{j+1}\}$ ima gotovo sva tražena svojstva, još je potrebno odrediti kolika je norma vektora f_{j+1}. Zaključujemo da za svaki skalar $\lambda \neq 0$ i vektor λf_{j+1} može uzeti umjesto f_{j+1}. Naime,

$$\langle f_{j+1}, e_i \rangle = 0 \Longrightarrow \langle \lambda f_{j+1}, e_i \rangle = 0, \quad \forall i = 1, \ldots, j$$

iz jednakosti $[\{e_1, \ldots, e_j, f_{j+1}\}] = [\{x_1, \ldots, x_j, x_{j+1}\}]$
dobivamo $[\{e_1, \ldots, e_j, \lambda f_{j+1}\}] = [\{x_1, \ldots, x_j, x_{j+1}\}]$, za svaki skalar $\lambda \neq 0$.
Uzmimo specijalno $\lambda = \|f_{j+1}\|^{-1}$, tj. definiramo

$$e_{j+1} = \frac{1}{\|f_{j+1}\|} f_{j+1}.$$

Primjetimo da je $f_{j+1} \neq 0$, jer kada bi bilo $f_{j+1} = 0$, vrijedilo bi:

$$x_{j+1} = \sum_{i=1}^{j} \langle x_{j+1}, e_i \rangle e_i \in [\{e_1, \ldots, e_j\}] = [\{x_1, \ldots, x_j\}]$$

a to je u kontradikciji s nezavisnošću polaznog skupa $\{x_1, \ldots, x_k\}$. □

2 Schurova dekompozicija

U prethodnom odjeljku smo već predstavili značaj i korisna svojstva sličnih matrica. Sada možemo nešto više reći i o samoj Schurovoj dekompoziciji matrice. Riječ je o unitarnoj transformaciji sličnosti koja matrici $\mathbf{A}$ pridružuje njoj sličnu gornjetrokutastu matricu. Gornjetrokutasta matrica $\mathbf{T}$, koja je oblika $\mathbf{T} = \mathbf{U}^*\mathbf{A}\mathbf{U}$, dobije se djelovanjem unitarne matrice $\mathbf{U}$ na polaznu matricu $\mathbf{A}$.
Dokaz Schurove dekompozicije ćemo provesti metodom matematičke indukcije po dimenziji matrice $\mathbf{A}$, a sam postupak ćemo provesti koristeći Gram-Schmidtovog postupka ortogonalizacije, koji će nam omogućiti izgradnju unitarne matrice $\mathbf{U}$ čijim ćemo djelovanjem dobiti gornjetrokutastu matricu.

Teorem 3. (Schurova dekompozicije matrice)
Neka je dana matrica $\boldsymbol{A} \in \mathcal{M}_n$ čije su svojstvene vrijednosti $\lambda_1, \ldots, \lambda_n$ u proizvoljnom poretku i neka je $\boldsymbol{x} \in \mathbb{C}^n$ jedinični vektor, takav da vrijedi: $\boldsymbol{Ax} = \lambda_1 \boldsymbol{x}$.
Postoji unitarna matrica $\boldsymbol{U} = [\boldsymbol{x} \quad u_2 \ldots u_n] \in \mathcal{M}_n$ i gornje trokutasta matrica $\boldsymbol{T} = [t_{ij}]$ gdje je $t_{ii} = \lambda_i$, $i = 1, \ldots, n$ tako da je $\boldsymbol{U^*AU} = \boldsymbol{T} = [t_{ij}]$.
Ako matrica $\boldsymbol{A} \in \mathcal{M}_n(\mathbb{R})$ ima realne svojstvene vrijednosti, tada je $\boldsymbol{x}$ realan vektor i postoji realna ortogonalna matrica $\boldsymbol{Q} = [x \quad q_2 \ldots q_n] \in \mathcal{M}_n(\mathbb{R})$ takva da vrijedi: $\boldsymbol{Q}^T\boldsymbol{AQ} = \boldsymbol{T} = [t_{ij}]$ gdje je $\boldsymbol{T}$ gornje trokutasta matrica sa dijagonalnim vrijednostima $t_{ii} = \lambda_i, i = 1, \ldots, n$. Zapis $\boldsymbol{U^*AU} = \boldsymbol{T}$ zovemo Schurova dekompozicija od $\boldsymbol{A}$, a matrica $\boldsymbol{T}$ se naziva Schurova forma od A.

Dokaz. Dokaz provodimo indukcijom po dimneziji matrice.
Za $n = 1$, tvrdnja slijedi trivijalno.
Pretpostavimo da tvrdnja vrijedi za n-1 te promotrimo što se događa za matricu dimenzije n. Neka je $A = [a_{ij}] \in \mathcal{M}_n$ proizvoljna matrica sa svojstvenim vrijednostima $\lambda_1, \ldots, \lambda_n$.
Neka je $\mathbf{u}$ svojstveni vektor matrice $\mathbf{A}$ čija je odgovarajuća svojstvena vrijednost λ, $\mathbf{Au} = \lambda\mathbf{u}$. Bez smanjenja općenitosti možemo pretpostaviti da je $\mathbf{u}$ normiran vektor. Koristeći

Gram-Schmidtov postupak ortogonalizacije možemo doći do $n-1$ ortogonalnih vektora $u_2, \ldots, u_n$ koji su okomiti na vektor $\mathbf{u}$. Sada možemo konstruirati unitarnu matricu $U = [\mathbf{u} \quad u_2 \ldots u_n] = [\mathbf{u} \quad \mathbf{V}]$ čiji je prvi stupac svojstveni vektor $\mathbf{u}$ matrice $\mathbf{A}$, a matrica $\mathbf{V}$ je matrica dimenzije $n \times n-1$. Sada pogledajmo što se događa kada provedemo sljedeći račun:

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{U}^*\begin{bmatrix}\mathbf{A}\mathbf{x} & \mathbf{A}\mathbf{u}_2 & \ldots & \mathbf{A}\mathbf{u}_n\end{bmatrix} = \mathbf{U}^*\begin{bmatrix}\lambda\mathbf{x} & \mathbf{A}\mathbf{u}_2 & \ldots & \mathbf{A}\mathbf{u}_n\end{bmatrix}$$
$$= \begin{bmatrix}\mathbf{x}^* \\ \mathbf{u}_2{}^* \\ \vdots \\ \mathbf{u}_n{}^*\end{bmatrix}\begin{bmatrix}\lambda\mathbf{x} & \mathbf{A}\mathbf{u}_2 & \ldots & \mathbf{A}\mathbf{u}_n\end{bmatrix} = \begin{bmatrix}\lambda\mathbf{x}^*\mathbf{x} & \mathbf{x}^*\mathbf{A}\mathbf{u}_2 \ \ldots \ \mathbf{x}^*\mathbf{A}\mathbf{u}_n \\ \lambda\mathbf{u}_2{}^*\mathbf{x} & \\ \vdots & \mathbf{A}_1 \\ \lambda\mathbf{u}_n{}^*\mathbf{x} & \end{bmatrix}$$
$$= \begin{bmatrix}\lambda & * \\ 0 & \mathbf{A}_1\end{bmatrix}.$$

Budući da prema pretpostavci indukcije tvrdnja vrijedi za matrice dimenzije $n-1$, tada tvrdnja vrijedi i za podmatricu $\mathbf{A}_1 \in \mathcal{M}_{n-1}$. Definiramo unitarnu matricu $\mathbf{W} \in \mathcal{M}_{n-1}$ te dolazimo do sljedećeg izraza:

$$\mathbf{A}_1 = \mathbf{W}\begin{bmatrix}\lambda_2 & * & \ldots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n\end{bmatrix}\mathbf{W}^*$$

što je ekvivalentno sljedećem izrazu:

$$\mathbf{W}^*\mathbf{A}_1\mathbf{W} = \begin{bmatrix}\lambda_2 & * & \ldots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n\end{bmatrix}$$

Uočimo sljedeće:

$$\left[\begin{array}{c|c}1 & 0 \ \ldots \ 0 \\ \hline 0 & \\ \vdots & \mathbf{W} \\ 0 & \end{array}\right]^* \left[\begin{array}{c|c}\lambda & * \ \ldots \ * \\ \hline 0 & \\ \vdots & \mathbf{A}_1 \\ 0 & \end{array}\right] \left[\begin{array}{c|c}1 & 0 \ \ldots \ 0 \\ \hline 0 & \\ \vdots & \mathbf{W} \\ 0 & \end{array}\right]$$
$$= \left[\begin{array}{c|c}1 & 0 \ \ldots \ 0 \\ \hline 0 & \\ \vdots & \mathbf{W} \\ 0 & \end{array}\right] \left[\begin{array}{c|c}\lambda & * \ \ldots \ * \\ \hline 0 & \\ \vdots & \mathbf{A}_1\mathbf{W} \\ 0 & \end{array}\right] = \left[\begin{array}{c|c}\lambda & * \ \ldots \ * \\ \hline 0 & \\ \vdots & \mathbf{W}^*\mathbf{A}_1\mathbf{W} \\ 0 & \end{array}\right] = \left[\begin{array}{c|ccc}\lambda & * & \ldots & * \\ \hline 0 & \lambda_2 & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n\end{array}\right].$$

S obzirom da je $\mathbf{U}$ unitarna matrica, $\mathbf{U} := \left[\begin{array}{c|c}1 & 0 \\ \hline 0 & \mathbf{W}\end{array}\right]$, tada vrijedi:

$$\left[\begin{array}{c|c}\lambda & * \ \ldots \ * \\ \hline 0 & \\ \vdots & \mathbf{A}_1 \\ 0 & \end{array}\right] = \mathbf{W}\begin{bmatrix}\lambda & * & \ldots & * \\ 0 & \lambda_2 & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n\end{bmatrix}\mathbf{W}^*$$

te je tvrdnja Schurovog teorema dokazana. $\square$

Unitarna transformacija sličnosti svake kvadratne matrice **A** gornjetrokutastoj matrici dana je Schurovom dekompoziciji, ali ono što treba uočiti je da takva dekompozicija nije jedinstvena. Odnosno, tako dobivena matrica **T** nije jedinstvena.

2.1 Primjeri Schurove dekompozicije

Primjer 2. *Odredimo Schurovu dekompoziciju matrice s realnim svojstvenim vrijednostima. Neka je* $\boldsymbol{A} = \begin{bmatrix} 7 & -2 \\ 12 & -3 \end{bmatrix}$.

Rješenje. *Najprije odredimo svojstveni polinom matrice* $\boldsymbol{A}$*:*

$$det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \begin{vmatrix} 7-\lambda & -2 \\ 12 & -3-\lambda \end{vmatrix} = (7-\lambda)(-3-\lambda) + 24 = \lambda^2 - 4\lambda - 21 + 24$$
$$= \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3).$$

Svojstvene vrijednosti matrice $\boldsymbol{A}$ *su realne i iznose 1 i 3.*
Sada možemo izračunati svojstveni vektor za svojstvenu vrijednost $\lambda = 3$.

$$\boldsymbol{A} - \lambda \boldsymbol{I} = \boldsymbol{A} - 3\boldsymbol{I} = \begin{bmatrix} 4 & -2 \\ 12 & -6 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}$$

Tada dobijemo svojstveni vektor za svojstvenu vrijednost $\lambda = 3$ *dan s* $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$*, pripadnim računom:*

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{u}_1 = 0$$
$$\begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$2a_1 - a_2 = 0$$
$$2a_1 = a_2.$$

Sada vektor $\boldsymbol{u}$ *možemo normalizirati te dobijemo vektor* $\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ *Nadalje trebamo pronaći ortonormiranu bazu* $\{v_2\}$ *koja je okomita na* $[\{u_1\}]^{\perp} = \left[\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}\right]^{\perp} = \left[\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}\right]$.
Odaberimo $v_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} -2 \\ 1 \end{bmatrix}$
Na taj način smo dobili ortonormiranu matricu $\boldsymbol{U}$*:*

$$U = \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}.$$

Sada pogledamo zapis matrice $\boldsymbol{A}$ *u bazi* u_1, v_2*:*

$$U^{-1}\boldsymbol{A}U = U^T\boldsymbol{A}U = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 7 & -2 \\ 12 & -3 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -14 \\ 0 & 1 \end{bmatrix}.$$

Budući je matrica $\boldsymbol{A}$ *dimenzije* 2×2*, možemo uzeti da je* $u_2 = v_2$*, stoga je Schurova dekompozicija matrice* $\boldsymbol{A}$ *dana izrazom:* $\boldsymbol{A} = UTU^*$*, gdje je:*

$$T = \begin{bmatrix} 3 & -14 \\ 0 & 1 \end{bmatrix}$$

gornjetrokutasta matrica.

$$U = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$$

je unitarna matrica.

Primjer 3. *Odredimo Schurovu dekompoziciju matrice* $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix}$. *čije svojstvene vrijednosti su kompleksno konjugirane.*

Rješenje. *Prvo pronađimo svojstvene vrijednosti matrice* $\boldsymbol{A}$.

$$det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \begin{vmatrix} 1-\lambda & 1 \\ -2 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) + 2 = \lambda^2 - 4\lambda + 5$$

$$\lambda = \frac{4 \pm \sqrt{16 - 4(5)}}{2} = 2 \pm \sqrt{-1} = 2 \pm i$$

Odaberimo svojstvenu vrijednost $\lambda_1 = 2 + i$ *i pronađimo njoj pripadni svojstveni vektor.*

$$\boldsymbol{A} - \lambda_1 \boldsymbol{I} = \boldsymbol{A} - (2+i)\boldsymbol{I} = \begin{bmatrix} -1-i & 1 \\ -2 & 1-i \end{bmatrix} \sim \begin{bmatrix} 1+i & -1 \\ 0 & 0 \end{bmatrix}.$$

Tada dobijemo svojstveni vektor za svojstvenu vrijednost λ_1, $\boldsymbol{u} = \begin{bmatrix} 1 \\ 1+i \end{bmatrix}$, *pripadnim računom:*

$$(\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{u} = 0$$

$$\begin{bmatrix} 1+i & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(1+i)a_1 - a_2 = 0$$

$$a_2 = (1+i)a_1$$

Uočima da kod kompleksnih matrica nulprostor nije ortogonal na prostor redaka, već na konjugirani prostor redaka. Normalizacijom svojstvenog vektora $u = \begin{bmatrix} 1 \\ 1+i \end{bmatrix}$ *dobijemo vektor* $u_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1+i \end{bmatrix}$.

Sada trebamo pronaći ortonormiranu bazu $\{v_2\}$ *od prostora razapetog s vektorom u.*

$$\left[\left\{\begin{bmatrix} 1 \\ 1+i \end{bmatrix}\right\}\right]^{\perp} = Ker \begin{bmatrix} 1 \\ 1-i \end{bmatrix} = \left[\left\{\begin{bmatrix} 1-i \\ -1 \end{bmatrix}\right\}\right]$$

Sada uzmimo da je normalizirani vektor $v_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1-i \\ -1 \end{bmatrix}$. *Nadalje možemo izraziti matricu* $\boldsymbol{A}$ *u novoj ortonormiranoj bazi* u_1, v_2, *što možemo zapisati na sljedeći način:*

$$U = \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix}$$

a matrica $\boldsymbol{A}$ *u ortonormiranoj bazi ima sljedeći zapis:*

$$U^{-1}\boldsymbol{A}U = U^*\boldsymbol{A}U = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix} = \begin{bmatrix} 2+i & -1+2i \\ 0 & 2-i \end{bmatrix}$$

Zato što je matrica $\boldsymbol{A}$ dimenzije 2×2 uzimamo da je $u_2 = v_2$. Pronašli smo Schurovu dekompoziciju matrice $\boldsymbol{A} = UTU^*$ gdje je

$$T = \begin{bmatrix} 2+i & -1+2i \\ 0 & 2-i \end{bmatrix}$$

gornjetrokutasta matrica. Unitarna matrica je oblika

$$U = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix}.$$

3 Posljedice Schurove dekompozicije

Važnost Schurove dekompozicije upravo je vidljiva u posljedicama ovoga teorema, neke od značajnih navest ćemo u sljedećim retcima.

- **Trag i determinanta matrice**
  Neka je $\mathbf{A} \in \mathcal{M}_n$ matrica sa svojstvenim vrijednostima $\lambda_1, \ldots, \lambda_n$. Trag matrice $\mathbf{A}$ definiramo na sljedeći način: $tr\mathbf{A} = \sum_{i=1}^{n} \lambda_i$.
  Determinantu matrice $\mathbf{A}$ definiramo na sljedeći način: $det\,\mathbf{A} = \prod_{i=1}^{n} \lambda_i$.
  Kako slične matrice imaju isti trag i istu determinantu, umjesto traga i determinante matrice $\mathbf{A}$ možemo računati trag i determinantu matrice $\mathbf{T}$.
  Stoga je povoljan zapis gornjetrokutaste matrice sa svojstvenim vrijednostima na glavnoj dijagonali kakav daje upravo Schurova dekompozicija matrice. Neka je $\mathbf{T} = [t_{ij}]$ gornjetrokutasta matrica sa svojstvenim vrijednostima na glavnoj dijagonali (upravo takva da odgovara matrici iz Schurovog teorema). Tada je trag matrice $\mathbf{T}$ definiran izrazom $tr\mathbf{T} = \sum_{i=1}^{n} t_{ii}$, a determinanta matrice $\mathbf{T}$ je dana izrazom; $det\,\mathbf{T} = \prod_{i=1}^{n} t_{ii}$.

- **Karakteristični polinom matrice**
  Neka je dana matrica $\mathbf{A} \in \mathcal{M}_n$ sa svojstvenim vrijednostima $\lambda_1, \ldots, \lambda_n$ i neka je $k_A(\lambda)$ karakteristični polinom matrice $\mathbf{A}$.
  Neka je $\mathbf{A} = UTU^*$, gdje je U unitarna matrica, a T $= [t_{ij}]$ je gornjetrokutasta matrica sa vrijednostima na glavnoj dijagonali $t_{11} = \lambda_1, t_{22} = \lambda_2, \ldots, t_{nn} = \lambda_n$. Tada je $k_A(\lambda) = k_{UTU^*}(\lambda) = Uk_T(\lambda)U^*$. Vrijednosti na glavnoj dijagonali od $k_T(\lambda)$ su $k_T(\lambda_1), \ldots, k_T(\lambda_n)$, stoga su ove vrijednosti također svojstvene vrijednosti ne samo matrice $\mathbf{T}$, već i matrice $\mathbf{A}$.
  Posebno, za svaki $k = 1, 2, \ldots,$ svojstvene vrijednosti od $\mathbf{A}^k$ su $\lambda_1^k, \ldots, \lambda_n^k$ i vrijedi:

  $$tr\mathbf{A}^k = \lambda_1^k + \cdots + \lambda_n^k$$

  .

- **Hamilton-Cayleyev teorem**
  Neka je $\mathbf{A} \in \mathcal{M}_n$. Tada je

  $$k_A(\mathbf{A}) = 0.$$

  Činjenica da svaka kvadratna matrica poništava vlastiti karakteristični polinom, izravna je posljedica Schurove dekompozicije te je pokazana u sljedećem teoremu.

  Teorem 4. *(Hamilton-Cayleyev teorem)*
  Neka je $k_A(t)$ karakteristični polinom matrice $\boldsymbol{A} \in \mathcal{M}_n$. Tada je $k_A(\boldsymbol{A}) = 0$.

Dokaz. Prema teoremu o Schurovoj dekompoziciji matrice postoje unitarna matrica $\mathbf{U}$ i gornjetrokutasta matrica $\mathbf{T}$ takve da je $\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{T}$.
Neka su $\lambda_1, \lambda_2, \ldots, \lambda_k$ svojstvene vrijednosti matrice $\mathbf{A}$ sa pripadnim kratnostim m_i. Sada prema Schurovoj dekompozicji dolazimo do matrice $\mathbf{T}$ koja je sljedećeg oblika:

$$\mathbf{T} = \mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} T_1 & * & \ldots & * \\ 0 & T_2 & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & T_k \end{bmatrix}$$

gdje je matrica T_i matrica dimenzije $m_i \times m_i$ oblika:

$$T_i = \begin{bmatrix} \lambda_i & * & \ldots & * \\ 0 & \lambda_i & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_i \end{bmatrix}.$$

Karakteristični polinom matrice T_i je $k_i(t) = (t - \lambda_i)^{m_i}$.
Matrica $T_i - \lambda_i\mathbf{I}$ je nilpotentna matrica sa indeksom nilpotentnosti $\leq m_i$. Stoga je $k_i(T_i) = (T_i - \lambda_i\mathbf{I})^{m_i} = 0$.
Sada je karakteristični polinom matrice $\mathbf{T}$ oblika:

$$k_T(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k}$$

sada uvrštavanjem matrice $\mathbf{T}$ u njezinim karakteristični polinom elementi na dijagonali se ponište, pa je $k_T(T) = 0$.
Uočimo:

$$\mathbf{U} * k_T(A)\mathbf{U} = \mathbf{U}^*(\mathbf{A} - \lambda_1\mathbf{I})^{m_1} \cdots (\mathbf{A} - \lambda_k)^{m_k}\mathbf{U}$$

nadalje, ako između svakog od faktora umetnemo produkt $\mathbf{U}^*\mathbf{U}$ dolazimo do sljedećeg izraza:

$$\begin{aligned} &\mathbf{U}^*k_T(A)\mathbf{U} \\ &= \underbrace{\mathbf{U}^*(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{U}\mathbf{U}^* \cdots \mathbf{U}\mathbf{U}^*(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{U}}_{m_1 faktora \quad \mathbf{U}^*(\mathbf{A}-\lambda_i\mathbf{I})\mathbf{U}} \cdots \underbrace{\mathbf{U}^*(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{U}\mathbf{U}^* \cdots \mathbf{U}\mathbf{U}^*(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{U}}_{m_k faktora \quad \mathbf{U}^*(\mathbf{A}-\lambda_k\mathbf{I})\mathbf{U}} \\ &= (\mathbf{U}^*\mathbf{A}\mathbf{U} - \lambda_1\mathbf{U}^*\mathbf{U})^{m_1} \cdots (\mathbf{U}^*\mathbf{A}\mathbf{U} - \lambda_k\mathbf{U}^*\mathbf{U})^{m_k} \\ &= (T - \lambda_1\mathbf{I})^{m_1} \cdots (T - \lambda_k\mathbf{I})^{m_k} = k_T(T) = 0. \end{aligned}$$

Iz ovog izraza slijedi tvrdnja teorema. □

- **Sylvesterov teorem za linearne matrične jednadžbe**
  Jednadžbu oblika $\mathbf{A}X - X\mathbf{B} = \mathbf{C}$ nazivamo Sylvesterova jednadžba, gdje su zadane matrice $\mathbf{A} \in \mathcal{M}_m$, $\mathbf{B} \in \mathcal{M}_n$ i $\mathbf{C} \in \mathcal{M}_{m\times n}$, a matrica $\mathbf{X} \in \mathcal{M}_{m\times n}$ je nepoznata matrica. Sylvesterova jednadžba ima jedinstveno rješenje, ako se spektri matrica $\mathbf{A}$ i $\mathbf{B}$ ne podudaraju niti u jednom elementu.
  Prema tvrdnji Schurove dekompozicije znamo da postoji unitarna matrica $\mathbf{U}$ i gornjetrokutasta matrica $\mathbf{T}$ tako da je matrica $\mathbf{A}$ dana u obliku: $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*$.

S obzirom da je matrica $\mathbf{A}$ unitarno slična matrici $\mathbf{T}$, one imaju iste svojstvene vrijednosti. Prema Schurovo dekompoziciju za svojstvenu vrijednost $\lambda \in \sigma(\mathbf{A}) \cap \sigma(\mathbf{B})$ postoje unitarne matrice $\mathbf{U}_1$ i $\mathbf{U}_2$ takve da je:

$$\mathbf{U}_1^*\mathbf{A}\mathbf{U}_1 = \begin{bmatrix} \lambda\mathbf{I}_a & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix} : \mathbf{U}_\lambda \otimes \mathbf{U}_\lambda^\perp \to \mathbf{U}_\lambda \otimes \mathbf{U}_\lambda^\perp$$

$$\mathbf{U}_2^*\mathbf{B}\mathbf{U}_2 = \begin{bmatrix} \lambda\mathbf{I}_b & \mathbf{B}_{12} \\ 0 & \mathbf{B}_{22} \end{bmatrix} : \mathbf{V}_\lambda \otimes \mathbf{V}_\lambda^\perp \to \mathbf{V}_\lambda \otimes \mathbf{V}_\lambda^\perp$$

gdje su $\mathbf{U}_\lambda i \mathbf{V}_\lambda$ pripadni svojstveni potprostori.
Uočimo, budući da svojstvena vrijednost $\lambda \notin \sigma(\mathbf{A}_{22}) \cup \sigma(\mathbf{B}_{22})$, sada je Sylesterova jednadžba oblika:

$$\mathbf{U}_1 \begin{bmatrix} \lambda\mathbf{I}_a & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix} \mathbf{U}_1^*\mathbf{X} - \mathbf{X}\mathbf{U}_2 \begin{bmatrix} \lambda\mathbf{I}_b & \mathbf{B}_{12} \\ 0 & \mathbf{B}_{22} \end{bmatrix} \mathbf{U}_2^* = \mathbf{C}$$

$$\Leftrightarrow \begin{bmatrix} \lambda\mathbf{I}_a & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix} \mathbf{U}_1^*\mathbf{X}\mathbf{U}_2 - \mathbf{U}_1^*\mathbf{X}\mathbf{U}_2 \begin{bmatrix} \lambda\mathbf{I}_b & \mathbf{B}_{12} \\ 0 & \mathbf{B}_{22} \end{bmatrix} = \mathbf{U}_1^*\mathbf{C}\mathbf{U}_2$$

stoga ako uvedemo oznake: $\mathbf{U}_1^*\mathbf{X}\mathbf{U}_2 = \mathbf{Y}$ i $\mathbf{U}_1^*\mathbf{C}\mathbf{U}_2 = \mathbf{D}$ Sylvesterova jednadžba je oblika:

$$\begin{bmatrix} \lambda\mathbf{I}_a & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} - \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} \lambda\mathbf{I}_b & \mathbf{B}_{12} \\ 0 & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}.$$

Prethodna jednadžba je ekvivalentna sljedećem sustavu:

$$\mathbf{A}_{12}Y_{21} = D_{11} \tag{1}$$
$$Y_{12}(\lambda\mathbf{I} - \mathbf{B}_{22}) - Y_{11}\mathbf{B}_{12} = D_{12} - \mathbf{A}_{12}Y_{22} \tag{2}$$
$$(\mathbf{A}_{22} - \lambda\mathbf{I})Y_{21} = D_{21} \tag{3}$$
$$\mathbf{A}_{22}Y_{22} - Y_{22}\mathbf{B}_{22} = D_{22} + Y_{21}\mathbf{B}_{12} \tag{4}$$

Iz izraza (1) i (3) dolazimo do izraza:

$$\mathbf{A}_{12}(\mathbf{A}_{22} - \lambda\mathbf{I})^{-1}D_{21} = D_{11}$$

a iz izraza (3) slijedi

$$Y_{21} = (\mathbf{A}_{22} - \lambda\mathbf{I})^{-1}D_{21}.$$

Pogledajmo sada slučaj u kojemu je jedina zajednička svojstvena vrijednost matrica $\mathbf{A}$ i $\mathbf{B}$ λ, tj. $\sigma(\mathbf{A}) \cap \sigma(\mathbf{B}) = \{\lambda\}$.
Sada prema (4) imamo jedinstvenu matricu Y_{22}, pa sada prema (2) možemo Y_{12} izraziti preko Y_{11}:

$$Y_{12} = (D_{12} - \mathbf{A}_{12}Y_{22} + Y_{11}\mathbf{B}_{12})(\lambda\mathbf{I} - \mathbf{B}_{22})^{-1}.$$

Na ovaj način je dobijena familija rješenja.
U drugom slučaju, ako postoji neka druga zajednička svojstvena vrijednost marica $\mathbf{A}$ i $\mathbf{B}$, koju možemo označiti sa μ, tj. $\mu \in \sigma(\mathbf{A}_{22}) \cap \sigma(\mathbf{B}_{22})$. Primjenom već opisanog računa na jednadžbu (4) dobivamo:

$$\mathbf{A}_{22}Y_{22} - Y_{22}\mathbf{B}_{22} = D_{22} + (\mathbf{A}_{22} - \lambda\mathbf{I})^{-1}D_{21}\mathbf{B}_{12}.$$

Nakon što provedemo postupak supstitucije unazad, jednadžba je riješena u potpunost.
Budući da kvadratna matrica n-tog reda ima najviše n različitih svojstvenih vrijednosti, ovaj algoritam mora terminirati u najviše n koraka.

Literatura

[1] D. Bakić, Linearna algebra, Školska knjiga, Zagreb, 2008.

[2] R. A. Horn and C. R. Johnson, Matrix analysis (Second edition), Cambridge University Press, 1994.

[3] Y. Saad, Numerical methods for large eigenvalue problems, Manchester University Press Series in Algorithms and Architectures for Advanced Scientific Computing, 1991.

[4] M. Frankland, Abstract Linear ALgebra - Schur decomposition, 2011.
Dostupno na: http://www.home.uni-osnabrueck.de/mfrankland/Math416/Math416_SchurDecomposition.pdf

[5] Z. Drmač, Numerička Analiza 1, 2008.
Dostupno na://web.math.pmf.unizg.hr/ drmac/NA1-2.pdf

[6] Engineering Department - Harvey Mudd College
Dostupno na: http://fourier.eng.hmc.edu/e176/lectures/ch1/node5.html

[7] Lecture 1: Schur's Unitary Triangularization Theorem
Dostupno na: http://www.math.drexel.edu/ foucart/TeachingFiles/F12/M504Lect1.pdf