UNIVERZA V MARIBORU

FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO

Oddelek za matematiko in računalništvo

MAGISTRSKO DELO

Jure Karo

Maribor, 2020

UNIVERZA V MARIBORU

FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO

Oddelek za matematiko in računalništvo

Magistrsko delo

UNITARNE MATRIKE

na študijskem programu 2. stopnje Izobraževalna matematika

Mentorica:

doc. dr. Mateja Grašič

Kandidat:

Jure Karo

Maribor, 2020

ZAHVALA

Največja nagrada za človekov trud ni tisto,
kar bo zanj dobil,temveč tisto, kar bo postal.
John Ruskin (1819-1900)

Za pomoč, podporo in mnoge koristne nasvete ter usmeritve pri pisanju magistrskega dela se zahvaljujem mentorici dr. Mateji Grašič.

Do tega, da lahko pišem zahvalo v magistrskem delu enopredmetnega študijskega programa, je bilo potrebnih veliko administrativnih korakov. Za vodenje skozi zapleteno kolesje sistema in brezpogojno podporo se želim posebej zahvaliti gospe Moniki Šket in njenim sodelavkam v referatu za študentske zadeve.

Velika zahvala gre tudi vsem mojim, ki so me v času študija podpirali, mi pomagali in me prenašali tako v sončnih dneh kot tudi v dneh, ko je bliskalo in grmelo . . .

Vsem od srca hvala!

Program magistrskega dela za enopredmetni magistrski študijski program
Izobraževalna matematika

Unitarne matrike

JURE KARO

V magistrskem delu naj bodo obravnavane lastnosti unitarnih matrik. Kompleksna kvadratna matrika U je unitarna, če je $U^*U = I$. V primeru realne matrike U z zgornjo lastnostjo (torej $U^TU = I$) govorimo o realni ortogonalni matriki U.

Predstavljena naj bo unitarna podobnost matrik ter Schurov izrek o unitarni triangulaciji matrike, ki pravi, da za vsako kompleksno matriko $A \in M_n$ obstaja taka unitarna matrika $U \in M_n$, da je

$$U^*AU = T$$

zgoraj trikotna matrika, ki ima po diagonali lastne vrednosti matrike A. V posebnem primeru, ko je A realna matrika z realnimi lastnimi vrednostmi, obstaja realna ortogonalna matrika U, da je $U^TAU = T$ zgoraj trikotna matrika z zgoraj opisano lastnostjo diagonalnih elementov.

V zaključku dela naj bodo predstavljene še nekatere algebraične lastnosti matričnih grup U_n, SU_n, O_n in SO_n.

Osnovna literatura:

R. A. Horn, C. R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1990.

doc. dr. Mateja Grašič, 20. 4. 2020

KARO, J.: Unitarne matrike.
Magistrsko delo, Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Oddelek za matematiko in računalništvo, 2020.

IZVLEČEK

V magistrskem delu obravnavamo unitarne matrike in predstavljamo njihove osnovne lastnosti. Unitarne matrike služijo kot orodje za utemeljevanje splošnejših rezultatov s področja linearne algebre. Primer tega je Schurova triangulacija, ki zagotavlja, da lahko vsako matriko zapišemo kot zgornjetrikotno matriko s kompleksnimi elementi.

Magistrsko delo je razdeljeno v štiri večje sklope. V prvem podajamo osnovne pojme iz linearne algebre, ki so povezane z matrikami, vektorskimi prostori in z lastnimi vrednostmi ter lastnimi vektorji, ki jih povežemo s podobnostjo matrik. V drugem sklopu so predstavljene unitarne matrike, opisane in dokazane so nekatere osnovne lastnosti teh matrik. V nadaljevanju se posvetimo unitarni podobnosti matrik, ki je nadgradnja podobnosti matrik.

V tretjem delu dokažemo Schurov izrek in navedemo nekaj rezultatov linearne algebre, ki so lahko dokazani s pomočjo Schurovega izreka. Zadnji del magistrskega dela je namenjen kratki obravnavi grup, ki jih tvorijo unitarne (in ortogonalne) matrike skupaj z operacijo matričnega množenja. Pri tem so dokazane nekatere temeljne algebrske lastnosti teh grup.

Ključne besede: unitarne matrike, lastne vrednosti matrike, unitarna podobnost matrik, Schurova triangulacija matrike, ortogonalna grupa, unitarna grupa

Math. Subj. Class. (2020): 15A18 Lastne vrednosti in lastni vektorji,
15A21 Kanonične forme, redukcije, klasifikacije,
15A23 Faktorizacija matrik
15B10 Ortogonalne matrike,
20G20 Linearne grupe nad realnimi in kompleksnimi števili ter kvaternioni.

KARO, J.: Unitary matrices.
Master Thesis, University of Maribor, Faculty of Natural Sciences and Mathematics, Department of Mathematics and Computer Science, 2020.

ABSTRACT

This master's thesis discusses unitary matrices and presents their basic properties. Unitary matrices serve as a tool to justify more general results in the field of linear algebra. An example of this is the Schur triangulation, which ensures that each matrix can be written as an upper triangular matrix with complex elements.

The master's thesis is divided into four major sections. In the first, we present basic concepts from linear algebra that are related to matrices, vector spaces, and to eigenvalues and eigenvectors, which are linked with matrix similarity. In the second section, unitary matrices are presented; some basic properties of these matrices are described and proved. In the following section, we focus on the unitary similarity of matrices, which can be considered an upgrade of matrix similarity.

In the third part, we prove Schur's theorem and list some results of linear algebra that can be proved by Schur's theorem. The last part of the master's thesis is intended to briefly discuss the groups formed by unitary (and orthogonal) matrices together with the operation of matrix multiplication, whereby some fundamental algebraic properties of these groups are proved.

Keywords: unitary matrices, eigenvalues, eigenmatrices, unitary similarity of matrices, Schur triangulation, orthogonal group, unitary group

Math. Subj. Class. (2020): 15A18 Eigenvalues, singular values, and eigenvectors,
15A21 Canonical forms, reductions, classification,
15A23 Factorization of matrices
15B10 Orthogonal matrices,
20G20 Linear algebraic groups over the reals, the complexes, the quaternions.

Kazalo

Uvod

Tema magistrskega dela so unitarne matrike. Gre za matrike, katerih definicijo zasledimo v vsakem nekoliko obsežnejšem delu, ki obravnava linearno algebro. Tudi to priča o dejstvu, da je področje uporabe unitarnih in njihovih realnih ekvivalentov torej ortogonalnih matrik precej široko. V magistrskem delu bodo predstavljeni zgolj nekateri aspekti te uporabe.

Matematiki so uporabo unitarnih matrik odkrivali v prvih desetletjih 20. stoletja. Glavne prispevke k obravnavani tematiki so dodali v Rusiji rojeni matematik judovskega porekla Issai Schur, ki je v času do nacističnega preganjanja deloval na univerzi v Berlinu. Schurovo glavno delo se sicer osredotoča na teorijo reprezentacije grup, verjetno pa je danes najbolj znan ravno po dokazu po njem poimenovane Schurove triangulacije, ki jo bomo predstavili v petem poglavju. V delu bomo omenili še dva matematika, ki sta doktorirala pod mentorstvom Issaia Schura, in sicer Alfreda Brauerja, ki je bil prav tako judovskega porekla in Wilhelma Spechta, ki je dodal pomemben prispevek k unitarni podobnosti, ki jo predstavljamo v četrtem poglavju. [11]

Pri pisanju magistrskega dela je v prvih petih poglavjih temeljno oporo nudilo delo [6]. Za dopolnitve pa so služila dela [1], [2], [8] in [9]. V zadnjem poglavju pa so bila uporabljena dela [3], [7] in [10].

Delo je organizirano v šest poglavij. Prvo poglavje predstavlja osnovno teorijo matrik in vektorskih prostorov. V drugem poglavju je predstavljena osnovna teorija lastnih vrednosti in lastnih vektorjev matrik, v nadaljevanju pa se dotaknemo še podobnosti matrik in diagonalizabilnosti. To poglavje služi kot nabor splošnih pojmov, ki jih dopolnimo z ugotovitvami tretjega poglavja, kjer so predstavljene temeljne lastnosti unitarnih matrik in Householderjeve matrike, ki so uporabne za konstrukcijo unitarnih matrik pri dveh danih vektorjih. Ob koncu je predstavljena še QR-faktorizacija, ki je uporabljena pri dokazu realne Schurove forme.

Četrto poglavje nadgrajuje drugo, saj je unitarna podobnost matrik, ki jo obravnavamo v tem poglavju, zgolj nadgradnja podobnosti matrik. Predstavljena sta potrebni in zadostni pogoj, da sta matriki unitarno podobni. V petem poglavju je dokazan Schurov izrek o triangulaciji matrike. Ob koncu pa je podanih še nekaj primerov, kjer lahko pri dokazih izrekov linearne algebre uporabimo dejstvo, da za vsako matriko obstaja neka zgornjetrikotna matrika s kompleksnimi elementi, ki ima na diagonali lastne vrednosti začetne matrike. Ob koncu petega poglavja so predstavljene normalne matrike, ki so neke vrste posplošitev unitarnih matrik.

Glavni izrek, gre za spektralni izrek za normalne matrike, ki ga dokažemo pri podpoglavju o normalnih matrikah, je pomemben za dokaz zadnjega izreka v tem magistrskem delu, kjer obravnavamo center ortogonalne in unitarne grupe. To je tudi glavna tema zadnjega poglavja. V tem poglavju vpeljemo nekaj elementarnih trditev iz teorije grup in predstavimo unitarne ter ortogonalne matrike kot grupe. Pri tem se omejimo na osnovne algebrske lastnosti teh grup.

Poglavje 1

Osnovni pojmi

1.1 Matrike

Definicija 1.1 *Matrika je pravokotna tabela s števili. Števila v matriki imenujemo elementi matrike. Matrike običajno označujemo z velikimi tiskanimi črkami.*

V nadaljevanju besedila bomo pogosto srečevali oznaki $M_{n,m}(\mathbb{R})$ in $M_{n,m}(\mathbb{C})$. Prva oznaka predstavlja množico matrik, ki imajo n vrstic in m stolpcev (tem matrikam pravimo $n \times m$ matrike oziroma matrike dimenzije $n \times m$) elementi matrik pa so realna števila. Druga oznaka pa predstavlja množico matrik dimenzije $n \times m$ s to razliko, da so elementi teh matrik kompleksna števila. V nadaljevanju besedila je pogosto uporabljena oznaka $M_{n,m}$, ki pomeni množico matrik dimenzije $n \times m$ s kompleksnimi elementi.
Ob tem ločimo dva posebna primera matrik. Ko je $n = m$, govorimo o kvadratnih matrikah, ki jih bomo označevali z $M_n(\mathbb{R})$ in $M_n(\mathbb{C}) = M_n$. Če pa imamo matriko dimenzije $n \times 1$, jo imenujemo stolpčni vektor; če je matrika dimenzije $1 \times n$, pa jo imenujemo vrstični vektor. V nadaljevanju besedila, kjer ne bo določeno drugače, z besedo vektor ciljamo na stolpčni vektor.

Oznaka 1.2 *Naj bo $A \in M_{n,m}$ dana matrika. Z oznako a_{ij} ali $(A)_{ij}$ bomo označili element matrike, ki se nahaja v i-ti vrstici in j-tem stolpcu. Temu elementu pravimo tudi (i,j)-ti element matrike.*

Zgled. Matrika $A = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 2 \\ 1 & 3 & 1 \end{bmatrix}$ je kvadratna 3×3 matrika. $(A)_{12} = a_{12} = 2$. Naj bo dana

še matrika $B = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 2 \end{bmatrix}$. Ta matrika je dimenzije 2×3, saj ima dve vrstici in tri stolpce. Element b_{21} matrike B je 6.

Elementi $a_{11},\ a_{22},\ \ldots,\ a_{nn}$ matrike $A \in M_n$ predstavljajo glavno diagonalo kvadratne matrike. Pojem diagonale je mogoče še nekoliko posplošiti. Tako lahko o diagonali govorimo tudi v matrikah, ki niso kvadratne. V tem primeru velja, da so diagonalni elementi matrike $B \in M_{n,m}$ enaki $b_{11},\ b_{22},\ \ldots,\ b_{nn}$, če je $n < m$ in $b_{11},\ b_{22},\ \ldots,\ b_{mm}$, če je $m < n$. V zgornjem zgledu so diagonalni elementi matrike A števila 1, 5 in 1, v primeru matrike B pa elementa 1 in 5.

Omenimo še nekatere posebne primere kvadratnih matrik. Matriko, ki ima na glavni diagonali same enice, povsod drugod pa ničle imenujemo identična matrika ali identiteta in jo označimo z I oziroma z I_n, če želimo posebej poudariti njeno dimenzijo. Identična matrika je poseben primer diagonalne matrike, ki ima neničelne elemente samo na diagonali (pri tem ne zahtevamo, da so elementi na diagonali enaki 1). Matrika, ki ima neničelne elemente na diagonali in ima neničelne elemente tudi nad glavno diagonalo, pod glavno diagonalo pa same ničle, je zgornjetrikotna matrika. Matrika z neničelnimi elementi na glavni diagonali in pod njo ter z ničelnimi nad glavno diagonalo, se imenuje spodnjetrikotna matrika. Če ima matrika (ne nujno kvadratna) same ničelne elemente, potem jo imenujemo ničelna matrika in jo označimo z 0.

Oznaka 1.3 *Diagonalno matriko $D \in M_n$ označimo kot $D = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})$.*

Zgled. Spodaj so navedeni primeri kvadratnih 3×3 matrik, ki so bile opisane zgoraj.

$$0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} -8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3-2i \end{bmatrix} = \operatorname{diag}(-8, 5, 3-2i)$$

Ničelna matrika Identična matrika Diagonalna matrika

$$T_Z = \begin{bmatrix} 4 & 1-i & \pi \\ 0 & i & 5 \\ 0 & 0 & -7 \end{bmatrix} \qquad T_S = \begin{bmatrix} -1 & 0 & 0 \\ 6 & -3 & 0 \\ 9 & -2 & 7 \end{bmatrix}$$

Zgornjetrikotna matrika Spodnjetrikotna matrika

Definicija 1.4 *Matriki $A,\ B \in M_{n,m}$ sta enaki, ko imata isto dimenzijo in zanju velja, da je $a_{ij} = b_{ij}$ za vsak $i = 1,\ \ldots, n$ in $j = 1,\ \ldots, m$.*

Ob koncu uvodne vpeljave matrik, si oglejmo še primer bločne matrike:

$$A = \begin{bmatrix} 1 & 3 & 8 & 7 \\ 7 & 2 & 1 & 9 \\ 0 & 6 & 2 & 5 \\ 8 & 1 & 5 & 2 \end{bmatrix} \longrightarrow A' = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$
$$A_{12} = \begin{bmatrix} 1 & 3 \\ 7 & 2 \end{bmatrix},\ A_{22} = \begin{bmatrix} 8 & 7 \\ 1 & 9 \end{bmatrix},\ A_{21} = \begin{bmatrix} 0 & 6 \\ 8 & 1 \end{bmatrix} \text{ in } A_{22} = \begin{bmatrix} 2 & 5 \\ 5 & 2 \end{bmatrix}.$$

Dana je matrika A. Potem matrike A_{11}, A_{12}, A_{21} in A_{22} imenujemo podmatrike ali bloki matrike A. Matriko A' pa bločna matrika matrike A. Bločno matriko bomo označevali takole: $[A_{ij}]_{i,j}^{k} \in M_n$, kjer je $0 < i, j \leq k \leq n$.

Računske operacije z matrikami

Med matrikami iste dimenzije je definirano seštevanje, in sicer kot seštevanje istoležnih elementov matrik. Prav tako je definirano tudi množenje matrike s skalarjem, in sicer tako, da vsak element matrike pomnožimo z danim skalarjem. Če sta $A,\ B \in M_{n,m}$ in $\alpha \in \mathbb{C}$, potem je:

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij},\ \forall i \in \{1,\ \ldots,\ n\},\ \forall j \in \{1,\ \ldots,\ m\};$$
$$(\alpha A)_{ij} = \alpha\,(A)_{ij},\ \forall i \in \{1,\ \ldots,\ n\},\ \forall j \in \{1,\ \ldots,\ m\}.$$

Med matrikami je definirano tudi množenje. Naj bo matrika A dimenzije $n \times r$ in matrika B dimenzije $r \times m$. Potem je produkt $A \cdot B$ matrika dimenzije $n \times m$, kjer je (i, j)-ti element produkta vsota od prvega elementa i-te vrstice matrike A pomnoženega s prvim elementom j-tega stolpca matrike B do r-tega elementa i-te vrstice pomnoženega z r-tim elementom j-tega stolpca matrike B. S simboli:

$$(A \cdot B)_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj},\ \forall i \in \{1,\ \ldots,\ n\},\ \forall j \in \{1,\ \ldots,\ m\};$$

Zgled. Naj bosta dani matriki $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$ in $B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$. Izračunajmo $A \cdot B$

in $B \cdot A$.

$$A \cdot B = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

Produkta $B \cdot A$ ni mogoče izračunati, saj matriki A in B nista ustreznih dimenzij. Produkt matrik je mogoče izračunati le, če je število stolpcev prvega faktorja v produktu enako številu vrstic drugega faktorja. Na podlagi tega lahko vidimo, da množenje matrik ni komutativno.

Zgoraj navedene računske operacije z matrikami imajo nekaj lastnosti, ki bodo predstavljene v naslednjem izreku. Dokazana bo zgolj lastnost ascociativnosti množenja matrik, ostale lastnosti se dokažejo podobno. Dokaze pa lahko najdemo na primer v [1] ali [9].

Izrek 1.5 *Za računske operacije seštevanja in množenja matrik ter množenja matrik s skalarjem veljajo naslednje lastnosti:*

1. *seštevanje matrik je komutativno:* $\forall A, B \in M_{n,m} : A + B = B + A$,
2. *seštevanje matrik je asociativno:* $\forall A, B, C \in M_{n,m} : (A + B) + C = A + (B + C)$,
3. *nevtralni element za seštevanje je ničelna matrika:* $\exists 0_{m,n}, \forall A \in M_{n,m} : A + 0 = A$,
4. *vsaka matrika premore nasprotno matriko:* $\forall A \in M_{n,m}, \exists -A \in M_{n,m} : A + (-A) = 0_{n,m}$,
5. *produkt vsote skalarjev z matriko lahko porazdelimo:* $\forall A \in M_{n,m}, \forall \alpha, \beta \in \mathbb{C} : (\alpha + \beta) A = \alpha A + \beta A$,
6. *produkt skalarja z vsoto matrik lahko porazdelimo:* $\forall A, B \in M_{n,m}, \forall \alpha \in \mathbb{C} : \alpha (A + B) = \alpha A + \alpha B$,
7. *zaporedno množenje skalarjev z matriko lahko združimo:* $\forall A \in M_{n,m}, \forall \alpha, \beta \in \mathbb{C} : \alpha (\beta A) = (\alpha\beta) A$,
8. *množenje matrike s skalarjem* 1 *ne spremeni njene vrednosti:* $\forall A \in M_{n,m} : 1 \cdot A = A$,
9. *množenje matrik je asociativno:* $\forall A \in M_{m,n}, \forall B \in M_{n,p}, \forall C \in M_{p,q} : (AB) C = A (BC)$,
10. *velja desna distributivnost:* $\forall A \in M_{n,m}, \forall B, C \in M_{m,k} : A (B + C) = AB + AC$,
11. *velja leva distributivnost:* $\forall A, B \in M_{n,m}, \forall C \in M_{m,k} : (A + B) C = AC + BC$,
12. $\forall A \in M_{n,m}, \forall B \in M_{m,k}, \forall \alpha \in \mathbb{C} : \alpha (AB) = (\alpha A) B = A (\alpha B)$.

Dokaz. Dokazali bomo točko 9: Naj bodo A, B, C matrike kot je zapisano v izreku. Potem je:

$$((AB)\,C)_{ij} = \sum_{l=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kl}\right)c_{lj} = \sum_{l=1}^{p}\sum_{k=1}^{n} a_{ik}b_{kl}c_{lj} = \sum_{k=1}^{n}\sum_{l=1}^{p} a_{ik}b_{kl}c_{lj} =$$
$$= \sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{p} b_{kl}c_{lj}\right) = (A\,(BC))_{ij}\,,$$

kar smo želeli dokazati. □

O gornjih lastnostih lahko povemo še nekaj več. Prve štiri lastnosti, ki veljajo za seštevanje matrik: komutativnost, asociativnost, obstoj nevtralnega elementa za seštevanje in obstoj nasprotnih elementov, zagotavljajo, da je množica matrik dimenzije $m \times n$ skupaj z operacijo seštevanja Abelova grupa (več o tem v naslednjem podpoglavju in v zadnjem poglavju). Če obravnavamo kvadratne matrike, pa lahko zapišemo, da je množica kvadratnih matrik skupaj z notranjima binarnima operacijama seštevanja in množenja ter z zunanjo binarno operacijo množenja s skalarji, ki zadoščajo vsem lastnostim iz izreka 1.5, algebra nad poljem kompleksnih števil. Ker za operacijo matričnega množenja v množici kvadratnih matrik obstaja tudi enota za množenje (gre za identično matriko), kvadratne matrike z omenjenimi operacijami tvorijo algebro (nad poljem kompleksnih števili) z enoto. Hkrati pa vemo, da množenje matrik ni komutativna operacija, zato govorimo o nekomutativni algebri z enoto. O algebri kvadratnih matrik več v [2] na straneh 69–71, o algebrah kot algebrskih strukturah pa več v [3] začenši na straneh 22–23.

Oglejmo si še, kaj se dogaja pri množenju zgornjetrikotnih matrik.

Trditev 1.6 *Produkt dveh zgornjetrikotnih matrik je zgornjetrikotna matrika.*

Dokaz. Naj bosta $A, B \in M_n$ zgornjetrikotni matriki. Potem je produkt (i, j)-tega elementa te matrike enak:

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + \ldots + a_{ik-1}b_{k-1j} + a_{ik}b_{kj} + a_{ik+1}b_{k+1j} + \ldots + a_{in}b_{nj}.$$

Zanimajo nas zgolj elementi matrike, ko je $i > j$. V tem primeru morajo biti vsi elementi matrike AB enaki nič. Ločimo dva primera: (i) če je $i > k$, potem je $a_{ik} = 0$, saj je A zgornjetrikotna matrika in (ii) če je $j < k$, potem je $b_{kj} = 0$, saj je tudi B zgornjetrikotna matrika. Če obe ugotovitvi združimo dobimo:

$$(AB)_{ij} = 0 \cdot b_{1j} + \ldots + 0 \cdot b_{k-1j} + 0 \cdot 0 + a_{ik+1} \cdot 0 + \ldots + a_{in} \cdot 0 = 0, \text{ za vse } i > j.$$

S tem smo pokazali, da je produkt dveh zgornjetrikotnih matrik spet zgornjetrikotna matrika. □

Transponiranje in konjugiranje matrik

Operacija transponiranja pri matrikah zamenja vrstice in stolpce v matriki. To pomeni, da stolpci matrike postanejo vrstice transponirane matrike in obratno, pri tem pa se vrstni red stolpcev oziroma vrstic ne zamenja. Transponirano matriko matrike $A \in M_{n,m}$ označimo z A^T, njena dimenzija je $m \times n$. Za transponiranje velja: $\left(A^T\right)_{ij} = (A)_{ji}$.

Zgled. Poiščimo transponirano matriko matrike A:

$$A = \begin{bmatrix} 3 & -5 & 1 \\ -7 & 8 & -4 \end{bmatrix} \qquad A^T = \begin{bmatrix} 3 & -7 \\ -5 & 8 \\ 1 & -4 \end{bmatrix}$$

Izrek 1.7 *Za transponiranje veljajo naslednje lastnosti:*

1. $\forall A \in M_{n,m} : \left(A^T\right)^T = A$;
2. $\forall A \in M_{n,m}, \forall \alpha \in \mathbb{C} : (\alpha A)^T = \alpha A^T$;
3. $\forall A, B \in M_{n,m} : (A + B)^T = A^T + B^T$;
4. $\forall A \in M_{n,k}, \forall B \in M_{k,m} : (AB)^T = B^T A^T$.

Dokaz. Prve tri točke so očitne. Dokaz za točko 4: (j, i)-ti element produkta AB je enak $\sum_{k=1}^{n} a_{jk} b_{ki}$, kar je enako (i, j)-temu elementu matrike $(AB)^T$. Izračunajmo še (i, j)-ti element produkta $B^T A^T$:

$$\left(B^T A^T\right)_{ij} = \sum_{k=1}^{n} (B^T)_{ik} (A^T)_{kj} = \sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki},$$

saj je množenje kompleksnih števil komutativno. Če obe ugotovitvi združimo, dobimo: $(AB)^T = B^T A^T$. □

Definirajmo še operacijo konjugiranja matrik s kompleksnimi elementi. Konjugiranje na matriko s kompleksnimi elementi deluje tako, da konjugiramo vse elemente matrike.

Definicija 1.8 *Operacija $*$ predstavlja konjugiranje in transponiranje dane matrike. Matriko A^* bomo imenovali konjugirana-transponiranka matrike A.*

V primeru, da ima matrika A realne elemente, operacija $*$ nanjo deluje le kot transponiranje.

Zgled. Poiščimo konjugirano-transponiranko matrike A.

$$A = \begin{bmatrix} 5-2i & -3 \\ 9i & 1+2i \\ 1 & 6+7i \end{bmatrix} \qquad A^* = \begin{bmatrix} 5+2i & -9i & 1 \\ -3 & 1-2i & 6-7i \end{bmatrix}$$

Definicija 1.9 *Kvadratno matriko $A \in M_n$ imenujemo Hermitska, če velja $A^* = A$, poševno Hermitska, če velja $A^* = -A$ in esencialno Hermitska, če je matrika $e^{i\theta}A$ Hermitska za nek $\theta \in \mathbb{R}$.*

Zgled. Naj bo dana 2×2 kompleksna matrika. Ugotovimo katere lastnosti mora imeti ta matrika, da bo Hermitska.

Matrika A je oblike: $A = \begin{bmatrix} a_1 + a_2 i & b_1 + b_2 i \\ c_1 + c_2 i & d_1 + d_2 i \end{bmatrix}$. Ker želimo, da bo matrika A Hermitska mora veljati:

$$\begin{aligned} A - \overline{A^T} &= 0 \\ \begin{bmatrix} a_1 + a_2 i & b_1 + b_2 i \\ c_1 + c_2 i & d_1 + d_2 i \end{bmatrix} - \begin{bmatrix} a_1 - a_2 i & c_1 - c_2 i \\ b_1 - b_2 i & d_1 - d_2 i \end{bmatrix} &= 0 \\ \begin{bmatrix} 2a_2 i & b_1 + b_2 i - c_1 + c_2 i \\ c_1 + c_2 i - b_1 + b_2 i & 2d_2 i \end{bmatrix} &= 0 \end{aligned}$$

Veljati mora: $a_2 = d_2 = 0$ in $b_1 = c_1$ ter $b_2 = -c_2$, kjer so $a_1,\ b_1,\ b_2,\ d_1 \in \mathbb{R}$ poljubna števila. Ugotovili smo, da imajo Hermitske 2×2 matrike na glavni diagonali realne elemente, nediagonalna elementa pa sta konjugirani si kompleksni števili.

Opomba 1.10 *Prejšnji primer lahko posplošimo na $n \times n$ matrike. Ugotovimo, da je matrika $A \in M_n$ Hermitska, če za vsaka $i, j = 1,\ 2,\ \ldots,\ n$ velja: $a_{ij} = \overline{a_{ji}}$.*

Determinanta matrike

Definicija 1.11 *Naj bo $A \in M_n$ kvadratna matrika. Tedaj je determinanta matrike A enaka:*

$$\det A = \sum_{\pi \in S_n} \operatorname{sign}(\pi) a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}.$$

V naslednjem izreku bo navedenih nekaj osnovnih lastnosti determinante. Lastnosti ne bomo dokazovali, dokaze pa lahko najdemo v [1], [2], ali [9].

Izrek 1.12 *Naj bosta $A, B \in M_n$ in naj bo $\alpha \in \mathbb{C}$. Potem veljajo naslednje lastnosti:*

1. *determinanta produkta matrik je enaka produktu determinant:*
   $\det(AB) = \det A \det B$,
2. *determinanta produkta skalarja in matrike je enaka produktu n-te potence skalarja in determinante matrike:*
   $\det(\alpha A) = \alpha^n \det A$,
3. *determinanta transponirane matrike je enaka determinanti osnovne matrike:*
   $\det A^T = detA$,
4. *determinanta konjugirane matrike je enaka konjugirani determinanti matrike:*
   $\det \overline{A} = \overline{\det A}$.

Inverz matrike

Definicija 1.13 *Če za kvadratno matriko $A \in M_n$ obstaja kvadratna matrika $B \in M_n$, da velja: $AB = BA = I$, potem pravimo, da je matrika A obrnljiva oziroma nesingularna. Matriko B imenujemo inverz matrike A. Če taka matrika B ne obstaja, potem matriko A imenujemo singularna.*

Izrek 1.14 *Za vsako nesingularno matriko $A \in M_n$ obstaja natanko ena matrika $B \in M_n$, da velja: $AB = BA = I$.*

Dokaz. Recimo, da obstajata različni matriki B_1 in B_2, da velja: $AB_1 = B_1A = I$ in $AB_2 = B_2A = I$.
Potem je $AB_1 - AB_2 = 0$, kar je enako $A(B_1 - B_2) = 0$. Če celotno enakost z leve pomnožimo z B_1 dobimo: $B_1A(B_1 - B_2) = 0$. Ker je $B_1A = I$, velja: $B_1 - B_2 = 0$, kar pomeni, da je $B_1 = B_2$. To je protislovno temu, da sta matriki različni. □

Ugotovili smo, da za vsako obrnljivo matriko obstaja natanko en inverz, zato lahko vpeljemo novo oznako. Inverzno matriko matrike A bomo odslej označevali z A^{-1}. Velja: $AA^{-1} = A^{-1}A = I$.

Trditev 1.15 *Naj bosta dani matriki $A \in M_{m,n}$ in $B \in M_{n,r}$. Potem velja: $(AB)^T = B^TA^T$.*

Dokaz. Izračunajmo:
$(AB)^T_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} \left(B^T\right)_{ik} \left(A^T\right)_{kj} = \left(B^T A^T\right)_{ij}$.
Ker enakost velja za poljubni element matrike $(AB)^T$, je trditev dokazana. □

Trditev 1.16 *Naj bo $A \in M_n$ nesingularna matrika. Tedaj velja: $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$.*

Dokaz. Pokazati želimo, da je matrika $\left(A^{-1}\right)$ inverz matrike A^T. Zato mora po definiciji nesingularnosti in ob upoštevanju trditve 1.15 veljati: $A^T \left(A^{-1}\right)^T = \left(A^{-1}A\right)^T = I^T = I$.
Če to enakost pomnožimo z $\left(A^T\right)^{-1}$ z leve, dobimo: $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$. □

Naslednji izrek povezuje veliko osnovnih pojmov linearne algebre z nesingularnostjo matrik. Izrek bomo v nadaljevanju velikokrat potrebovali, kljub temu pa bomo njegov dokaz izpustili, saj bi pomenil prevelik odklon od teme naloge. Dokaz izreka lahko najdemo v [1] ali [9].

Izrek 1.17 *Naj bo $A \in M_n$, potem so naslednje trditve ekvivalentne:*

1. *A je nesingularna.*
2. *Sistem enačb $Ax = 0$ ima samo trivalno rešitev.*
3. $\det A \neq 0$.

Trditev 1.18 *Produkt dveh nesingularnih matrik je spet nesingularna matrika.*

Dokaz. Naj bosta $A, B \in M_n$ nesingularni matriki. Vemo, da je matrika nesingularna, če velja $\det A \neq 0$ in $\det B \neq 0$.
Potem velja: $\det(AB) = \det A \cdot \det B \neq 0$, saj je produkt dveh neničelnih števil vedno neničelno število. Velja, da je matrika AB nesingularna. □

Adjugirane matrike

Z $A_{i,j}$ bomo označevali matriko A, ki smo ji izbrisali i-to vrstico in j-ti stolpec. Ilustrirajmo zapisano s primerom.

Zgled. Naj bo dana matrika $A \in M_3(\mathbb{R})$ in poiščimo matriko $A_{2,3}$.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 2 \\ 1 & 3 & 1 \end{bmatrix} \qquad A_{2,3} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$$

Definicija 1.19 *Naj bo $A \in M_n$ kvadratna matrika in* $\det A$ *njena determinanta. Kofaktor matrike A je pravilno predznačena determinanta matrike $A_{i,j}$. Označimo ga z $C_{i,j}$.*

Zgornjo definicijo bi lahko zapisali tudi s simboli. Velja, da je $C_{i,j} = (-1)^{i+j} \det(A_{i,j})$, za $i, j = 1, 2, \ldots, n$.
V zgornjem zgledu bi bil kofaktor $C_{2,3}$ enak:

$$C_{2,3} = (-1)^{2+3} \det(A_{2,3}) = -\begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = -1.$$

Definicija 1.20 *Naj bo $A \in M_n$ kvadratna matrika. Potem matriko $\tilde{A}$ definiramo kot $(\tilde{A})_{ij} = C_{i,j}$, za vsak $i, j = 1,\ 2,\ \ldots,\ n$. Torej (i,j)-ti element matrike $\tilde{A}$ predstavlja (i,j)-ti kofaktor matrike A.*
Matriko $\mathrm{adj}(A) = \tilde{A}^T$ *imenujemo adjugirana matrika matrike A.*

Zgled. Vzemimo znova matriko A iz zgornjega zgleda in izračunajmo matriki $\tilde{A}$ in $\mathrm{adj}(A)$.

$$\tilde{A} = \begin{bmatrix} -1 & -4 & 13 \\ 7 & -2 & -1 \\ -11 & 16 & -7 \end{bmatrix} \qquad \mathrm{adj}(A) = \begin{bmatrix} -1 & 7 & -11 \\ -4 & -2 & 16 \\ 13 & -1 & -7 \end{bmatrix}$$

Izrek 1.21 *Naj bo $A \in M_n$ nesingularna matrika. Potem je njen inverz enak: $A^{-1} = \frac{1}{\det A} \mathrm{adj}(A)$.*

Dokaz. Pokazati moramo, da je $A \cdot \mathrm{adj}(A) = I \cdot \det A$:

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} C_{1,1} & \ldots & C_{n,1} \\ \vdots & \ddots & \vdots \\ C_{1,n} & \ldots & C_{n,n} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n a_{1k}C_{1,k} & \ldots & \sum_{k=1}^n a_{1k}C_{n,k} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^n a_{nk}C_{1,k} & \ldots & \sum_{k=1}^n a_{nk}C_{n,k} \end{bmatrix} =$$
$$= \begin{bmatrix} \det A & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \det A \end{bmatrix},$$

saj povsod dobimo vsote oblike $\sum_{k=1}^n a_{ik}C_{j,k}$. Če je $i = k$, potem je ta vsota enaka determinanti matrike A (velja po izreku o razvoju determinante, glej na primer [8] izrek 2.2), sicer pa je vsota enaka 0, saj vsota predstavlja determinanto, kjer sta i-ta in j-ta vrstica enaki. Za tako determinanto pa vemo, da je enaka 0 (za dokaz trditve glej na primer [2], posledica 5.9). □

Zgled. Izračunajmo inverzno matriko matrike A iz zgornjega zgleda.
Vemo, da je $\det A = 30$. Prav tako smo v zgornjem primeru izračunali adjugirano matriko matrike A. Če obe ugotovitvi povežemo z zgornjim izrekom dobimo:

$$A^{-1} = \frac{1}{\det A}\operatorname{adj}(A) = \frac{1}{30}\begin{bmatrix} -1 & 7 & -11 \\ -4 & -2 & 16 \\ 13 & -1 & -7 \end{bmatrix}$$

Trditev 1.22 *Naj bo $A \in M_n$ nesingularna matrika in $c \in \mathbb{C}$. Potem je* $\operatorname{adj}(cA) = c^{n-1}\operatorname{adj}(A)$.

Dokaz. Po izreku 1.21 velja $(cA)^{-1} = \frac{1}{\det(cA)} \cdot \operatorname{adj}(cA)$ in $A^{-1} = \frac{1}{\det A}\operatorname{adj}(A)$. Potem je: $\operatorname{adj}(cA) = (cA)^{-1}\det(cA) = c^{-1}A^{-1}c^n \det A = c^{n-1}\frac{1}{\det A}\det A \operatorname{adj}(A) = c^{n-1}\operatorname{adj}(A)$. □

Posledica 1.23 *Velja:* $\operatorname{adj}(cI_n) = c^{n-1}I_n$.

Dokaz. Posledica je očitna, saj smo namesto matrike A izbrali identično matriko, ki je sama sebi adjugirana. □

Sled matrike

Definicija 1.24 *Naj bo $A \in M_n$. Potem je sled matrike A vsota elementov na glavni diagonali matrike A:* $\operatorname{sled} A = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn}$.

Opomba 1.25 *Sled lahko definiramo tudi splošneje, če je $A \in M_{m,n}$, potem je sled matrike A:* $\operatorname{sled} A = \sum_{i=1}^{p} a_{ii}$, *kjer je* $p = \min\{m, n\}$.

Izrek 1.26 *Naj bosta $A, B \in M_n$ in $c \in \mathbb{C}$, potem za sled veljajo naslednje lastnosti:*

1. $\operatorname{sled}(A + B) = \operatorname{sled} A + \operatorname{sled} B$,
2. $\operatorname{sled}(cA) = c \cdot \operatorname{sled} A$,
3. $\operatorname{sled}(A^T) = \operatorname{sled} A$,
4. $\operatorname{sled}(AB) = \operatorname{sled}(BA)$.

Dokaz. Dokaz za točko 1:

$$\operatorname{sled}(A+B) = \sum_{i=1}^{n}(a_{ii}+b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \operatorname{sled} A + \operatorname{sled} B$$

Ob tem smo upoštevali komutativnost seštevanja kompleksnih števil. Podobno se dokaže tudi točka 2. Pri točki 3 vemo, da se diagonalni elementi matrike pri transponiranju ne spreminjajo, zato je lastnost očitna. Oglejmo si še točko 4:

$$\operatorname{sled}(AB) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{ji}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ji}a_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} = \operatorname{sled}(BA).$$

V tem primeru smo najprej upoštevali, da je množenje kompleksnih števil komutativno, kasneje pa, da je komutativno tudi seštevanje kompleksnih števil. □

Posledica 1.27 (Cikličnost sledi produkta matrik) *Naj bodo dane matrike A, B, C, D ustreznih dimenzij, da obstaja produkt $ABCD$. Potem je* $\operatorname{sled}(ABCD) = \operatorname{sled}(BCDA) = \operatorname{sled}(CDAB) = \operatorname{sled}(DABC)$.

Dokaz. Posledica sledi neposredno iz točke 4 prejšnjega izreka in iz dejstva, da je množenje matrik asociativno (glej izrek 1.5). □

Pregled ključnih matrik

Spodnja preglednica prikazuje različne kvadratne matrike, katerih imena se razlikujejo glede na to, ali imajo realne ali kompleksne elemente. Te matrike bomo pogosto srečevali v nadaljevanju.

Realni elementi	**Kompleksni elementi**
Simetrična matrika $A = A^T$ $a_{ij} = a_{ji}$	Hermitska matrika $A = A^*$ $a_{ij} = \overline{a_{ji}}$
Poševnosimetrična matrika $A = -A^T$ $a_{ij} = -a_{ji}$	Poševnohermitska matrika $A = -A^*$ $a_{ij} = -\overline{a_{ji}}$
Ortogonalna matrika $O^{-1} = O^T$	Unitarna matrika $U^{-1} = U^*$

1.2 Vektorski prostori

V poglavju bomo definirali nekaj osnovnih pojmov povezanih z vektorskimi prostori. Ob tem naj velja opomba, da povsod uporabljamo polje skalarjev $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Definicija 1.28 *Neprazno množico G skupaj z binarno operacijo $\circ : G \times G \to G$ imenujemo grupa, oznaka: $(G, \circ)$, če zadošča naslednjim trem pogojem:*

$(G1)$ *binarna operacija v grupi je asociativna: $\forall x, y, z \in G : (x \circ y) \circ z = x \circ (y \circ z)$;*

$(G2)$ *v grupi je vsebovan nevtralni element (enota) za operacijo: $\exists e \in G, \forall x \in G : e \circ x = x \circ e = x$;*

$(G3)$ *za vsak element grupe je v grupi vsebovan tudi njegov inverzni element: $\forall x \in G, \exists y \in G : x \circ y = y \circ x = e$.*

Ni težko pokazati, da ima vsak element v grupi natanko en inverzni element. Zato bomo inverz elementa x od sedaj naprej označevali z x^{-1}.

Opomba 1.29 *Grupo, ki poleg zgornjih treh pogojev zadošča še pogoju o komutativnosti svojih elementov, imenujemo komutativna ali Abelova grupa. Za Abelovo grupo torej še velja:*

$(G4)$ *poljubna elementa iz grupe sta komutativna: $\forall x, y \in G : x \circ y = y \circ x$.*

Definicija 1.30 *Množica V skupaj z operacijama seštevanja: $+ : V \times V \to V$ in množenja s skalarjem $\cdot : \mathbb{F} \times V \to V$ se imenuje vektorski prostor (nad $\mathbb{F}$), oznaka: $(V, +, \cdot_{\mathbb{F}})$, če zadošča naslednjim pogojem:*

$(V1)$ *množica V je skupaj z operacijo seštevanja Abelova grupa;*

$(V2)$ *produkt vsote skalarjev z vektorjem je mogoče porazdeliti: $\forall \alpha, \beta \in \mathbb{F}, \forall x \in V : (\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$;*

$(V3)$ *produkt skalarja z vsoto vektorjev je mogoče porazdeliti: $\forall \alpha \in \mathbb{F}, \forall x, y \in V : \alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$;*

$(V4)$ *zaporedno množenje vektorja z dvema skalarjema je mogoče združiti: $\forall \alpha, \beta \in \mathbb{F}, \forall x \in V : \alpha \cdot (\beta \cdot x) = (\alpha \cdot \beta) \cdot x$;*

($V5$) *množenje z enoto za množenje iz* $\mathbb{F}$ *ne spremeni vrednosti nobenega vektorja:* $\forall x \in V : 1 \cdot x = x$.

Ugotovimo lahko, da množica matrik $M_{n,m}$ skupaj z operacijo matričnega seštevanja in množenja s skalarjem ustreza pogojem vektorskega prostora, zato je množica matrik vektorski prostor nad obsegom kompleksnih števil (glej izrek 1.5). Ker so posebni primeri matrik tudi stolpčni vektorji, tudi ti tvorijo vektorske prostore, ki jih običajno označimo kot $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$, $\ldots$, $\mathbb{R}^n$ oziroma $\mathbb{C}^1$, $\mathbb{C}^2$, $\mathbb{C}^3$, $\ldots$, $\mathbb{C}^n$. Omenimo še, da bomo v nadaljevanju stolpčna vektorja $\mathbb{R}^1$ in $\mathbb{C}^1$ enačili kar z množico realnih oziroma kompleksnih števil.

Vpeljimo še pojem vektorskega podprostora, začnimo s formalno definicijo:

Definicija 1.31 *Neprazna podmnožica* U *vektorskega prostora* V *je vektorski podprostor vektorskega prostora* V*, če za poljubna* $x, y \in U$ *in* $\alpha, \beta \in \mathbb{F}$ *velja:* $\alpha x + \beta y \in U$.

Če nadaljujemo z zgornjim primerom, lahko zapišemo, da je vektorski podprostor prostora kvadratnih $n \times n$ matrik s kompleksnimi koeficienti množica diagonalnih $n \times n$ matrik ali množica zgornjetrikotnih $n \times n$ matrik.

Naj bodo dani vektorji $x_1,\ x_2,\ \ldots,\ x_n$ iz vektorskega prostora V. Potem vsak vektor oblike $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$ imenujemo linearna kombinacija danih vektorjev. Skalarji $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n \in \mathbb{C}$ se imenujejo koeficienti linearne kombinacije.

Definicija 1.32 *Vektorji* $x_1,\ x_2,\ \ldots,\ x_n \in V$ *so linearno odvisni, če obstajajo taki skalarji* $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n \in \mathbb{C}$*, ki niso vsi enaki* 0*, da velja:*

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0.$$

Če vektorji niso linearno odvisni, potem pravimo, da so linearno neodvisni. Množico, ki vsebuje linearno neodvisne vektorje, imenujemo linearno neodvisna množica.

Definicija 1.33 *Naj bo* M *podmnožica vektorskega prostora* V*. Potem množico vseh linearnih kombinacij vektorjev iz množice* M *imenujemo linearna lupina množice* M *in jo označimo* $\mathscr{L}(M)$.

Iz definicije ne sledi, da morajo biti vektorji v linearni lupini linearno neodvisni. To je pomembno poudariti zaradi definicije baze vektorskega prostora, ki sledi v nadaljevanju. Še preden definiramo bazo vektorskega prostora, namenimo nekaj besed o razsežnosti vektorskega prostora. Če v vektorskem prostoru V obstaja podmnožica M, ki ima končno število elementov in zanjo velja $\mathscr{L}(M) = V$, potem pravimo, da je vektorski prostor V končno razsežen. Če taka končno razsežna podmnožica ne obstaja, pa je vektorski prostor neskončno razsežen.

Definicija 1.34 *Podmnožica $B = \{b_1,\ b_2,\ \ldots,\ b_n\}$ končno razsežnega vektorskega prostora V se imenuje baza prostora V, če je linearno neodvisna in če velja $\mathscr{L}(B) = V$.*
Število elementov v množici B imenujemo dimenzija vektorskega prostora.

Primer baze prostora $\mathbb{R}^2$ sta vektorja $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, saj vidimo, da lahko poljuben vektor iz prostora $\mathbb{R}^2$ izrazimo kot linearno kombinacijo danih vektorjev. Množico, ki vsebuje vektorja e_1 in e_2, imenujemo standardna baza prostora $\mathbb{R}^2$. Podobno so definirane standardne baze vseh vektorskih prostorov $\mathbb{R}^n$ in $\mathbb{C}^n$. Pri tem oznaka e_i pomeni stolpčni vektor, ki ima na i-tem mestu enico, povsod drugod pa ničle.

Izrek 1.35 *Če je množica vektorjev $\{x_1,\ x_2,\ \ldots,\ x_k\}$ linearno neodvisna in se vektorja x_{k+1} ne da zapisati kot linearno kombinacijo vektorjev $x_1,\ x_2,\ \ldots, x_k$, potem je tudi množica vektorjev $\{x_1,\ x_2,\ \ldots, x_k,\ x_{k+1}\}$ linearno neodvisna.*

Dokaz. Vemo, da so vektorji $x_1,\ x_2,\ \ldots, x_k$ linearno neodvisni, zato zanje velja: $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0 \Leftrightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. Oglejmo si linearno kombinacijo vektorjev:

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k + \alpha_{k+1} x_{k+1} = 0. \tag{1.1}$$

Ločimo dve možnosti. (i) Če je $\alpha_{k+1} \neq 0$, potem lahko vektor x_{k+1} zapišemo kot: $x_{k+1} = \alpha_{k+1}^{-1} \sum_{i=1}^{k} \alpha_i x_i$. To pomeni, da smo vektor zapisali kot linearno kombinacijo preostalih vektorjev, kar je v nasprotju s predpostavko izreka. (ii) Če je $\alpha_{k+1} = 0$, potem so vsi koeficienti linearne kombinacije 1.1 enaki 0, zato so vektorji linearno neodvisni. Sledi, da je množica vektorjev $\{x_1,\ x_2,\ \ldots, x_k,\ x_{k+1}\}$ linearno neodvisna. □

Posledica 1.36 *Vsako linearno neodvisno podmnožico M končno razsežnega vektorskega prostora V je mogoče dopolniti do baze prostora V.*

Dokaz. Naj bo dana množica $M = \{x_1,\ x_2,\ \ldots, x_k\}$. Množica je linearno neodvisna, da pa bo množica M baza prostora V mora veljati, da je linearna lupina množice M enaka vektorskemu prostoru V. Ločimo dve možnosti: če je $\mathscr{L}(M) = V$, potem smo končali, sicer pa zgornji izrek zagotavlja, da lahko množico dopolnimo z vektorjem x_{k+1}. Ta vektor, ki je linearno neodvisen z vektorji iz množice M zagotovo obstaja, saj smo predpostavili, da je $\mathscr{L}(M) \neq V$. Tako dobimo množico $M' = \{x_1,\ x_2,\ \ldots, x_k,\ x_{k+1}\}$. Če je $\mathscr{L}(M') = V$ smo končali, če ne ponovimo zgornji postopek. Ker je vektorski prostor V končno razsežen, bo število korakov končno. □

Definicija 1.37 *Naj bo $(V, +, \cdot_{\mathbb{F}})$ vektorski prostor nad obsegom $\mathbb{F}$. Preslikavo $\langle \_, \_ \rangle : V \times V \to \mathbb{F}$ imenujemo skalarni ali notranji produkt, če zadošča naslednjim zahtevam:*

1. *aditivnost: $\forall x, y, z \in V : \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;*
2. *homogenost $\forall x, y \in V,\ \alpha \in \mathbb{F} :\ \langle \alpha \cdot x, y \rangle = \alpha \langle x, y \rangle$;*
3. *konjugirana simetričnost: $\forall x, y \in V :\ \langle x, y \rangle = \overline{\langle y, x \rangle}$*
4. *pozitivna definitnost: $\forall x \in V :\ \langle x, x \rangle \geq 0$ in $\langle x, x \rangle = 0$, natanko tedaj, ko je $x = 0$.*

Zgled. Naj bosta $x, y \in \mathbb{C}^n$. Pokažimo, da je s predpisom $\langle x, y \rangle = y^* x$ podan kompleksni ali **Hermitski skalarni produkt**.
Naj bodo $x, y, z \in \mathbb{C}^n$ poljubni vektorji, kjer je $x = [x_1 \ \ x_2 \ \ \ldots \ \ x_n]^T$ ter $\alpha \in \mathbb{C}$ poljuben skalar. Preveriti je potrebno vse štiri točke iz zgornje definicije:

1. $\langle x + y, z \rangle = z^* (x + y) = z^* x + z^* y = \langle x, z \rangle + \langle y, z \rangle$;
2. $\langle \alpha x, y \rangle = y^* \alpha x = \alpha y^* x = \alpha \langle x, y \rangle$;
3. $\overline{\langle y, x \rangle} = \overline{x^* y} = x^T \overline{y} = \left(\overline{y}^T x\right)^T = (y^* x)^T = \langle x, y \rangle$;
4. $\langle x, x \rangle = x^* x = \overline{x_1} x_1 + \overline{x_2} x_2 + \ldots + \overline{x_n} x_n = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 \geq 0$, in $\langle x, x \rangle = 0$, natanko tedaj, ko je $x_1 = x_2 = \ldots = x_n = 0$, oziroma ko je $x = 0$.

Iz zadnje točke zgornjega dokaza lahko razberemo še, da je Hermitski produkt vektorja s samim seboj enak realnemu številu, saj za kompleksno število $z = a + bi$, kjer sta $a, b \in \mathbb{R}$, velja: $z\overline{z} = (a + bi)(a - bi) = a^2 + b^2$. Torej $z\overline{z} \in \mathbb{R}$, za vsako kompleksno število z.

Definicija 1.38 *Za vektorja $x,\ y \in V$ pravimo, da sta pravokotna ali ortogonalna, če je $\langle x, y \rangle = 0$.*
Vektor $x \in V$ imenujemo enotski ali normiran vektor, če je $\langle x, x \rangle = 1$
Naj bo V vektorski prostor nad poljem $\mathbb{F}$. Baza $B = \{b_1,\ b_2,\ \ldots,\ b_n\}$ prostora V je ortonormirana, če za vsak $i, j = 1, \ldots, n;\ i \neq j$ velja: $\langle b_i, b_j \rangle = 0$ in ko je $i = j$ velja, da je $\langle b_i, b_i \rangle = 1$.

Opomba 1.39 *Ortonormirano bazo vektorjev lahko definiramo tudi s pomočjo Kroneckerjeve delte:*

$$\delta_{i,j} = \begin{cases} 1, & i = j; \\ 0 & i \neq j. \end{cases}$$

Velja torej: $\langle b_i, b_j \rangle = \delta_{ij}$.

Zgled. Naj bodo $y_1,\ y_2,\ \ldots,\ y_n \in \mathbb{C}^n$ ortogonalni neničelni vektorji. Pokažimo, da so vektorji $x_1,\ x_2,\ \ldots,\ x_n$ definirani kot $x_i = \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_i$, $i = 1,\ 2,\ \ldots,\ k$, ortonormirani. Preveriti je potrebno dvoje, in sicer, da je skalarni produkt poljubnih različnih vektorjev x_i in x_j, kjer $i \neq j$, enak 0 ter da je skalarni produkt poljubnega vektorja s samim seboj enak 1. Najprej preverimo, da so vektorji ortogonalni.

$$\begin{aligned}
&\langle x_i, x_j\rangle = \left(\langle y_j, y_j\rangle^{-\frac{1}{2}} \cdot y_j\right)^* \left(\langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_i\right) = \langle y_j, y_j\rangle^{-\frac{1}{2}} \cdot y_j^* \cdot \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_i = \\
&= \langle y_j, y_j\rangle^{-\frac{1}{2}} \cdot \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_j^* y_i = \langle y_j, y_j\rangle^{-\frac{1}{2}} \cdot \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot \langle y_i, y_j\rangle = \\
&\langle y_j, y_j\rangle^{-\frac{1}{2}} \cdot \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot 0 = 0,
\end{aligned}$$

saj so vektorji y_i, y_j, kjer $i \neq j$ in $i,\ j = 1,\ \ldots,\ k$, ortogonalni. Preverimo še drugo zahtevo za ortonormiranost vektorjev:

$$\begin{aligned}
&\langle x_i, x_i\rangle = \left(\langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_i\right)^* \left(\langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_i\right) = \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_i^* \cdot \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_i = \\
&= \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot \langle y_i, y_i\rangle^{-\frac{1}{2}} \cdot y_i^* y_i = \langle y_i, y_i\rangle^{-1} \cdot \langle y_i, y_i\rangle = 1.
\end{aligned}$$

Izrek 1.40 *Vsaka ortonormirana množica vektorjev v $\mathbb{C}^n$ je linearno neodvisna.*

Dokaz. Naj bo $\{x_1,\ x_2,\ \ldots,\ x_n\}$ ortonormirana množica vektorjev. Potem za vsak vektor x_i iz te množice velja $\langle x_i, x_i\rangle = 1$ in za poljubna različna vektorja x_i in x_j, da je $\langle x_i, x_j\rangle = 0$. Naj bo $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0$ linearna kombinacija danih vektorjev. Potem je:

$$\begin{aligned}
&0 = (\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n)^* (\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n) = \\
&= \sum_{i=1}^{n} \overline{\alpha}_i x_i^* \alpha x_i = \sum_{i=1}^{n} \overline{\alpha}_i \alpha_i x_i^* x_i = \sum_{i=1}^{n} \overline{\alpha}_i \alpha_i \langle x_i, x_i\rangle = \sum_{i=1}^{n} |\alpha_i|^2 .
\end{aligned}$$

Ker so seštevanci $|\alpha_i|^2 \geq 0$, za vsak i, bo vsota enaka nič samo tedaj, ko bodo vsi sumandi enaki 0. Torej, ko bo $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Ugotovili smo, da je samo trivialna linearna kombinacija danih vektorjev enaka 0, kar pomeni, da so vektorji linearno neodvisni. □

Definicija 1.41 *Preslikavo $\|\_\| : V \to \mathbb{R}$, imenujemo norma, če zadošča naslednjim zahtevam:*

1. *pozitivna definitnost: $\forall x \in V : \|x\| \geq 0$ in $\|x\| = 0 \Leftrightarrow x = 0$;*
2. *absolutna homogenost: $\forall x \in V, \forall \alpha \in \mathbb{F} : \|\alpha \cdot x\| = |\alpha|\,\|x\|$;*
3. *trikotniška neenakost: $\forall x, y \in V : \|x + y\| \leq \|x\| + \|y\|$.*

Definicija 1.42 *Evklidska norma vektorja $x \in V$ je definirana kot: $\|x\| = \sqrt{\langle x, x\rangle}$.*

Trditev 1.43 *Naj bosta $x,\ y \in V$, kjer je V vektorski prostor nad poljem $\mathbb{C}$. Potem velja:* $\|x+y\|^2 = \|x\|^2 + 2\,\mathrm{Re}\langle x,y\rangle + \|y\|^2$.

Dokaz. Izračunajmo: $\|x+y\|^2 = \langle x+y, x+y\rangle = (x+y)^*\,(x+y) = x^*x+x^*y+y^*x+y^*y = \langle x,x\rangle + \langle y,x\rangle + \langle x,y\rangle + \langle y,y\rangle = \|x\|^2 + \overline{\langle x,y\rangle} + \langle x,y\rangle + \|y\|^2 = \|x\|^2 + 2\,\mathrm{Re}\langle x,y\rangle + \|y\|^2$. □

Izrek 1.44 (Neenakost Cauchy-Schwarz-Bunjakovski) *Naj bo dan vektorski prostor $V_{\mathbb{C}}$, ki je opremljen s Hermitskim produktom. Potem za poljubna vektorja $x,\ y \in V$ velja:*

$$|\langle x,y\rangle| \le \|x\|\,\|y\|\,,$$

pri čemer je enakost dosežena, ko sta vektorja linearno odvisna.

Dokaz. Ločimo dva primera:
(i) Denimo, da je en izmed vektorjev enak 0. Brez škode za splošnost predpostavimo, da je $y = 0$, potem je $|\langle x,0\rangle| = 0^*x = 0$ in $\|x\|\,\|0\| = \|x\|\,0 = 0$. Izrek velja, če je eden izmed vektorjev enak 0.
(ii) Denimo, da je vektor y različen od 0. Naj bo $\alpha \in \mathbb{C}$ skalar za katerega velja: $\alpha = -\frac{\langle x,y\rangle}{\|y\|^2}$. Izračunajmo:

$$\begin{aligned}
&0 \le \|x+\alpha y\|^2 = \langle x+\alpha y, x+\alpha y\rangle = (x+\alpha y)^*\,(x+\alpha y) = \\
&= xx^* + x^*\alpha y + \overline{\alpha} y^* x + \overline{\alpha}\alpha y^* y = \langle x,x\rangle + \alpha\langle y,x\rangle + \overline{\alpha}\langle x,y\rangle + \overline{\alpha}\alpha\langle y,y\rangle = \\
&= \|x\|^2 + \alpha\overline{\langle x,y\rangle} + \overline{\alpha}\langle x,y\rangle + \overline{\alpha}\alpha\,\|y\|^2 = \\
&= \|x\|^2 - \frac{\langle x,y\rangle}{\|y\|^2}\cdot\overline{\langle x,y\rangle} - \frac{\overline{\langle x,y\rangle}}{\|y\|^2}\langle x,y\rangle + \frac{\overline{\langle x,y\rangle}\langle x,y\rangle}{\|y\|^2} = \|x\|^2 - \frac{|\langle x,y\rangle|^2}{\|y\|^2}
\end{aligned}$$

Od tod sledi: $0 \le \|x\|^2 - \frac{|\langle x,y\rangle|^2}{\|y\|^2}$ ali $|\langle x,y\rangle|^2 \le \|x\|^2\,\|y\|^2$. Če to neenakost korenimo, dobimo želeno enakost.
Če je $\|x+\alpha y\|^2 = 0$, potem mora biti $x+\alpha y = 0$ in vektorja x in y sta linearno odvisna. V tem primeru je $|\langle x,y\rangle| = \|x\|\,\|y\|$. □

Rang, jedro in ničelnost matrike

V zadnjem razdelku tega poglavja bomo natančneje predstavili še tri pojme, ki povezujejo prvo in drugo podpoglavje magistrskega dela, in sicer matrike ter vektorske prostore. Za začetek nekaj besed o delitvi matrike na vrstice/stolpce. Naj bo $A \in M_{n,m}$, potem zapis $A = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}$ imenujemo stolpčna sestava matrike. Zapis matrike $A = \begin{bmatrix} a^1 & a^2 & \dots & a^n \end{bmatrix}^T$ pa imenujemo vrstična sestava matrike.

Definicija 1.45 *Naj bo $A \in M_{n,m}$ matrika, ki ima vrstično sestavo, kot je predstavljeno zgoraj. Rang matrike A je enak razsežnosti vektorskega prostora $\mathscr{L}\{a^1,\ a^2,\ \ldots,\ a^n\}$. Oznaka:* $\operatorname{rang} A$.

Povedano nekoliko drugače, rang matrike A je število linearno neodvisnih vrstic matrike A. S to definicijo je bil definiran vrstični rang, stolpčni rang definiramo analogno. Izkaže se, da je vrstični rang enak stolpčnemu. Dokaz tega rezultata bomo izpustili, zapišemo pa lahko, da gre za enega izmed pomembnejših rezultatov linearne algebre, katerega dokaz je moč najti v [2] na straneh 96–97. Vpeljimo še ničelnost matrike, ki jo bomo vpeljali preko jedra matrike.

Definicija 1.46 *Jedro matrike $A \in M_{n,m}$ je množica vseh vektorjev $x \in \mathbb{C}^m$, za katere velja $Ax = 0$. Označimo:* $\operatorname{Ker} A = \{x \in \mathbb{C}^m : Ax = 0\}$.

Izrek 1.47 *Naj bo dana matrika $A \in M_{n,m}$. Jedro matrike A je podprostor vektorskega prostora $\mathbb{C}^m$.*

Dokaz. Naj bo matrika $A \in M_{n,m}$ poljubna. Da bo $\operatorname{Ker} A$ podprostor vektorskega prostora $\mathbb{C}^m$ je potrebno najprej preveriti, da množica $\operatorname{Ker} A$ ni prazna. To zagotovo velja, saj množica $\operatorname{Ker} A$ vsebuje ničelni vektor. Nadalje, naj bosta $x, y \in \operatorname{Ker} A$ in $\alpha, \beta \in \mathbb{C}$. Ker sta x, y iz jedra sledi: $Ax = 0$ in $Ay = 0$. Izračunajmo: $A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha \cdot 0 + \beta \cdot 0 = 0$. Pri tem smo pri prvem enačaju uporabili desno distributivnost matrik in povezavo med matričnim množenjem in množenjem matrike s skalarjem (glej izrek 1.5, točki 10 in 12), pri drugem enačaju pa predpostavko, ki je navedena zgoraj. Sledi, da je vektor $\alpha x + \beta y \in \operatorname{Ker} A$ in zato je $\operatorname{Ker} A$ vektorski podprostor prostora $\mathbb{C}^m$. □

Razsežnost jedra matrike označimo z $\mathrm{N}(A)$ in jo imenujemo ničelnost matrike. Za vsako matriko $A \in M_{n,m}$ velja naslednja enakost: $\operatorname{rang}(A) + \mathrm{N}(A) = m$. Dokaz tega izreka je mogoče najti v [1] na strani 250.

Poglavje 2

Lastni vektorji, lastne vrednosti in diagonalizabilnost matrik

2.1 Lastni vektorji in lastne vrednosti

Definicija 2.1 *Naj bo $A \in M_n$. Skalar $\lambda \in \mathbb{C}$ imenujemo lastna vrednost matrike A, če obstaja tak neničelni vektor $x \in \mathbb{C}^n$, da je $Ax = \lambda x$. Vektor x imenujemo lastni vektor pripadajoč lastni vrednosti λ.*

Lastna vrednost vedno nastopa skupaj z njej pripadajočim lastnim vektorjem. Pri tem je pomembno izpostaviti, da lastni vektor, ki pripada lastni vrednosti, ni enolično določen, pač pa lahko vsak lastni vektor normiramo in tako ustvarimo enotski vektor. Še več, če je x lastni vektor pripadajoč lastni vrednosti $\lambda \in \mathbb{C}$, je tudi vektor $e^{i\theta}\frac{x}{\|x\|}$, za vsak $\theta \in \mathbb{R}$, lastni vektor za lastno vrednost λ.

Trditev 2.2 *Skalar $\lambda \in \mathbb{C}$ je lastna vrednost matrike A natanko tedaj, ko ima homogen sistem enačb $(A - \lambda I)\, x = 0$ netrivialno rešitev.*

Dokaz. λ je lastna vrednost matrike $A \Leftrightarrow \exists x \in \mathbb{C}^n$, $x \neq 0$: $(A - \lambda I)\, x = 0 \Leftrightarrow$ homogen sistem enačb $(A - \lambda I)\, x = 0$ ima netrivialno rešitev. □

Povežimo zgornjo ugotovitev še z izrekom 1.17. Ugotovili smo, da ima homogen sistem enačb $(A - \lambda I)\, x = 0$ netrivialno rešitev. V prej navedenem izreku pa druga točka pravi, da je matrika nesingularna, če ima homogen sistem linearnih enačb samo trivialno rešitev. Od tod sledi sklep, da je matrika $(A - \lambda I)\, x = 0$ singularna.

Definicija 2.3 *Množico vseh lastnih vrednosti matrike $A \in M_n$ imenujemo spekter matrike A. Oznaka: $\sigma(A)$.*

Trditev 2.4 *Matrika $A \in M_n$ je singularna natanko tedaj, ko je $0 \in \sigma(A)$.*

Dokaz. Po izreku 1.17 je matrika nesingularna, če ima homogen sistem linearnih enačb $Ax = 0$ samo trivalno rešitev. Ker je $0 \in \sigma(A)$, je 0 lastna vrednost matrike A. Po definiciji lastne vrednosti obstaja tak vektor $x \neq 0$, da je $Ax = 0x = 0$. Torej ima homogen sistem linearnih enačb netrivialno rešitev, zato je matrika A singularna. □

Definicija 2.5 *Naj bo dan polinom p stopnje k s kompleksnimi koeficienti: $p(x) = a_k x^k + a_{k-1}x^{k-1} + \ldots + a_1 x + a_0$, kjer je $a_k \neq 0$ in naj bo $A \in M_n$. Potem polinom $p(A) = a_k A^k + a_{k-1}A^{k-1} + \ldots + a_1 A + a_0 I$ imenujemo matrični polinom. Če je $a_k = 1$, potem polinomu rečemo monični matrični polinom.*

Opomba 2.6 *Matrični polinom $p(A)$ lahko vedno zapišemo kot monični polinom, saj je $a_k \neq 0$ in zato je $a_k^{-1}p(A)$ monični polinom.*

Za dokazovanje naslednjih rezultatov bomo potrebovali osnovni izrek algebre, ki bo naveden brez dokaza.

Izrek 2.7 (Osnovni izrek algebre) *Polinom stopnje $n \geq 1$ s kompleksnimi koeficienti ima (šteto z večkratnostjo) natanko n kompleksnih ničel.*

Posledica 2.8 *Vsak monični polinom s kompleksnimi koeficienti stopnje $n \geq 1$ lahko zapišemo kot produkt linearnih faktorjev: $p(x) = (x - z_1)(x - z_2)\cdots(x - z_n)$, kjer so z_1, z_2, …, z_n kompleksna števila.*

Izrek 2.9 *Naj bo $p(s)$ polinom stopnje k. Če je λ lastna vrednost matrike $A \in M_n$ s pripadajočim lastnim vektorjem x, potem je x tudi lastni vektor pripadajoč lastni vrednosti $p(\lambda)$ matrike $p(A)$. Obratno: če je $k \geq 1$ in je μ lastna vrednost matrike $p(A)$, potem obstaja neka lastna vrednost λ matrike A, da je $\mu = p(\lambda)$.*

Dokaz. Naj bo dan matrični polinom stopnje k: $p(A) = a_k A^k + \ldots + a_1 A + a_0 I$, kjer je $a_k \neq 0$.
Potem je $p(A)x = a_k A^k x + \ldots + a_1 Ax + a_0 Ix = a_k A^{k-1}(Ax) + \ldots + a_1(Ax) + a_0 x$. Po definiciji

lastne vrednosti je $p(A)x = a_k A^{k-1}(\lambda x) + \ldots + a_1(\lambda x) + a_0 x = a_k\lambda^k x + \ldots + a_1\lambda x + a_0 x = \left(a_k\lambda^k + \ldots a_1\lambda + a_0\right)x = p(\lambda)x$.
Ker je $p(A)x = p(\lambda)x$, je $p(\lambda)$ lastna vrednost matrike $p(A)$ in x lastni vektor pripadajoč lastni vrednosti $p(\lambda)$.
Obratno: Če je μ lastna vrednost matrike $p(A)$, potem je matrika $p(A) - \mu I$ singularna. Ker je polinom $p(s)$ stopnje k, je tudi polinom $q(s) = p(s) - \mu$ stopnje k, kjer je $k \geq 1$. Po osnovnem izreku algebre velja: $q(s) = (s - \beta_1)\cdots(s - \beta_k)$, kjer so $\beta_1, \ldots, \beta_k$ kompleksna števila.
Ampak, ker je matrika $q(A)$ singularna, je za nek $i = 1, \ldots, k$ faktor $(A - \beta_i I)$ singularen. Na podlagi komentarja, ki sledi trditvi 2.2, je β_i lastna vrednost matrike A.
Sledi: $0 = q(\beta_i) = p(\beta_i) - \mu \Rightarrow p(\beta_i) = \mu$. □

Izrek 2.10 *Vsaka matrika $A \in M_n$ ima lastno vrednost. Še več, za vsak neničeln vektor $y \in \mathbb{C}^n$ obstaja polinom $q(s)$ stopnje največ $n-1$, da je $q(A)y$ lastni vektor matrike A.*

Dokaz. Definirajmo število m kot najmanjše število, da so vektorji $y,\ Ay,\ \ldots,\ A^m y$ linearno odvisni. Zagotovo velja, da je $m \geq 1$, saj, če bi bil $m = 0$, potem je $y = 0$, kar je v nasprotju s predpostavko izreka. Hkrati pa velja, da je $m \leq n$, saj je poljubnih $n+1$ vektorjev iz prostora $\mathbb{C}^n$ zagotovo linearno odvisnih (posledica tega, da je razsežnost prostora $\mathbb{C}^n$ (nad $\mathbb{C}$) enaka n).
Ker so vektorji linearno odvisni, obstajajo koeficienti $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_m$, ki niso vsi enaki 0, da zanje velja:

$$\alpha_1 I_n y + \alpha_2 Ay + \cdots + \alpha_m A^m y = 0. \tag{2.1}$$

Recimo, da je $\alpha_m = 0$. Potem so vektorji $y,\ Ay,\ \ldots,\ A^{m-1}y$ linearno odvisni, kar je v nasprotju z minimalnostjo m. Zato je $\alpha_m \neq 0$.
Definirajmo polinom: $p(s) := s^m + \frac{\alpha_{m-1}}{\alpha_m}s^{m-1} + \ldots + \frac{\alpha_1}{\alpha_m}s + \frac{\alpha_0}{\alpha_m}$. Potem je zaradi enakosti 2.1 $p(A)y = 0 = 0y$. Zato je y lastni vektor pripadajoč lastni vrednosti 0 za matriko $p(A)$. Prejšnji izrek zagotavlja, da je ena izmed m ničel polinoma $p(s)$ lastna vrednost matrike A. Predpostavimo, da je λ ničla polinoma $p(s)$ in lastna vrednost matrike A. Potem je $p(s) = (s - \lambda)\,q(s)$, kjer je $q(s)$ polinom stopnje $m-1$. Če bi veljalo: $q(A)y = 0$, bi bilo to v nasprotju s predpostavko o minimalnosti m.
Zato velja, da je $q(A)y \neq 0$ in sledi: $p(A)y = (A - \lambda I)q(A)y = 0 \Rightarrow Aq(A)y = \lambda q(A)y$. Torej je $q(A)y$ lastni vektor matrike A pripadajoč lastni vrednosti λ. □

Opomba 2.11 *Ker ima vsaka matrika $A \in M_n$ lastno vrednost, ni spekter matrike A nikoli prazna množica.*

Sedaj, ko smo ugotovili, da ima vsaka matrika vsaj kakšno lastno vrednost, si oglejmo še posebnost po kateri *slovijo* Hermitske matrike. Te smo predstavili že v uvodu, spomnimo, da gre za matrike s kompleksnimi elementi, za katere je $A^* = A$.

Trditev 2.12 *Če je $A \in M_n$ Hermitska matrika, potem so vse njene lastne vrednosti realna števila.*

Dokaz. Naj bo λ lastna vrednost matrike A in $x = \begin{bmatrix} a_1 + b_1 i & \dots & a_n + b_n i \end{bmatrix}^T \in \mathbb{C}^n$ pripadajoč ji lastni vektor. Potem velja: $Ax = \lambda x$. Če obe strani enakosti pomnožimo z x^*, dobimo:

$$x^* Ax = x^* \lambda x = \lambda x^* x = \lambda \left(a_1^2 + b_1^2 + \dots a_n^2 + b_n^2\right) = \lambda \sum_{i=1}^{n} \left(a_i^2 + b_i^2\right).$$

Po drugi strani pa velja, da je matrika x^*Ax Hermitska, saj je:

$$(x^*Ax)^* = x^* A^* \left(x^*\right)^* = x^* Ax.$$

To pomeni, da mora biti element 1×1 matrike x^*Ax realno število. Da bo veljala enakost matrik in ker je vsota $\sum_{i=1}^{n} \left(a_i^2 + b_i^2\right)$ realno število, mora biti tudi λ realno število. Ker je bila lastna vrednost matrike izbrana poljubno, velja, da so vse lastne vrednosti Hermitske matrike realna števila. □

Definicija 2.13 *Karakteristični polinom matrike $A \in M_n$ je definiran kot $p_A(\lambda) = \det(A - \lambda I)$. Enačbo $\det(A - \lambda I) = 0$ imenujemo karakteristična enačba matrike A.*

Izrek 2.14 *Skalar λ je lastna vrednost matrike $A \in M_n$ natanko tedaj, ko je λ ničla karakterističnega polinoma.*

Dokaz. λ je lastna vrednost matrike $A \Leftrightarrow \exists x \in \mathbb{C}^n, x \neq 0 : (A - \lambda I)\, x = 0 \Leftrightarrow$ homogen sistem linearnih enačb $(A - \lambda I)\, x = 0$ ima netrivialno rešitev $\Leftrightarrow$ matrika $A - \lambda I$ je singularna $\Leftrightarrow \det(A - \lambda I) = 0 \Leftrightarrow p_A(\lambda) = 0$. □

Definicija 2.15 *Večkratnost skalarja λ kot ničle karakterističnega polinoma $p(\lambda)$ imenujemo algebrajska večkratnost ali kar večkratnost lastne vrednosti λ.*

Zgled. Poiščimo lastne vrednosti simetrične matrike $A = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$ in določimo njihovo algebrajsko večkratnost.

$$\det(A-\lambda I) = \begin{vmatrix} 3-\lambda & -1 & -1 \\ -1 & 3-\lambda & -1 \\ -1 & -1 & 3-\lambda \end{vmatrix} = \begin{vmatrix} 4-\lambda & -1 & -1 \\ 0 & 3-\lambda & -1 \\ 0 & -2 & 2-\lambda \end{vmatrix} =$$
$$= (4-\lambda)\left(\lambda^2-5\lambda+4\right) = (4-\lambda)(\lambda-4)(\lambda-1) = 0$$

Lastne vrednosti matrike A so $\lambda_1 = 1$ in $\lambda_2 = 4$. Algebrajska večkratnost lastne vrednosti $\lambda_1 = 1$ je 1, lastne vrednosti $\lambda_2 = 4$ pa 2.

Naj bosta x in $y \in \mathbb{C}^n$ lastna vektorja pripadajoča lastni vrednosti λ matrike $A \in M_n$. Opazimo, da za poljubna skalarja α in $\beta \in \mathbb{C}$ velja: $A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha\lambda x + \beta\lambda y = \lambda(\alpha x + \beta y)$. Torej je tudi vektor $\alpha x + \beta y$ lastni vektor za lastno vrednost λ. S tem smo ugotovili, da vsi lastni vektorji lastne vrednosti λ skupaj z ničelnim vektorjem tvorijo vektorski podprostor prostora $\mathbb{C}^n$. Ugotovitev lahko strnemo v naslednjo definicijo.

Definicija 2.16 *Naj bo $A \in M_n$. Za vsako lastno vrednost $\lambda \in \sigma(A)$ je množica vseh vektorjev $x \in \mathbb{C}^n$, ki zadoščajo enačbi $Ax = \lambda x$ lastni podprostor matrike A za lastno vrednost λ. Oznaka: V_λ.*
Dimenzijo lastnega podprostora matrike A za lastno vrednost λ imenujemo geometrijska večkratnost lastne vrednosti.

Na lastni podprostor lahko gledamo tudi drugače. In sicer lahko lastni podprostor lastne vrednosti λ definiramo kot jedro matrike $(A-\lambda I) \in M_n$. Potem je geometrijska večkratnost lastne vrednosti enaka: $\mathrm{N}(A-\lambda I)$ oziroma $n - \operatorname{rang}(A-\lambda I)$.

Zgled. Vrnimo se na zgornji zgled in določimo geometrijsko večkratnost za lastni vrednosti $\lambda_1 = 1$ in $\lambda_2 = 4$.

$$A - I = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \qquad A - 4I = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Lastni podprostor matrike A za lastno vrednost λ_1 je enak: $V_1 = \{(x,x,x)\,; x \in \mathbb{R}\}$ in lastni podprostor matrike A za lastno vrednost λ_2 je enak: $V_4 = \{(x,y,-x-y)\,; x,y \in \mathbb{R}\}$. Dimenzija prostora V_1 je 1, dimenzija podprostora V_4 pa 2, to sta tudi geometrijski večkratnosti lastnih vrednosti matrike A.

2.2 Podobnost in diagonalizabilnost matrik

Definicija 2.17 *Naj bosta $A,\ B \in M_n$. Matrika B je podobna matriki A, če obstaja taka nesingularna matrika S, da velja $B = S^{-1}AS$.*

Trditev 2.18 *Podobnost matrik je ekvivalenčna relacija.*

Dokaz.

1. Refleksivnost: $\exists I \in M_n$, da je $I^{-1}AI = IAI = A$.
2. Simetričnost: Matrika B je podobna matriki A, zato je $B = S^{-1}AS$. Ker je S nesingularna, velja tudi $SBS^{-1} = A$. Če vpeljemo novo oznako $T = S^{-1}$ dobimo: $A = T^{-1}BT$.
3. Tranzitivnost: Matrika B je podobna matriki A, zato je $B = S^{-1}AS$ in matrika C je podobna matriki B, torej $C = T^{-1}BT$. Potem velja: $C = T^{-1}BT = T^{-1}S^{-1}AST = (ST)^{-1}\,A\,(ST)$. Matrika ST je tudi nesingularna, saj je dobljena kot produkt dveh nesingularnih matrik (glej trditev 1.18). Torej je tudi matrika C podobna matriki A.

□

Izrek 2.19 *Naj bosta $A,\ B \in M_n$. Če je matrika B podobna matriki A, potem imata matriki enak karakteristični polinom.*

Dokaz. Matriki A in B sta podobni, torej obstaja nesingularna matrika S, da je $B = S^{-1}AS$.

$$\begin{aligned} &p_B(\lambda) = \det(B - \lambda I) = \det\left(S^{-1}AS - \lambda S^{-1}S\right) = \det\left(S^{-1}(A - \lambda I)\,S\right) = \\ &= \det\left(S^{-1}\right)\det(A - \lambda I)\det S = \frac{1}{\det S}\det(A - \lambda I)\det S = \\ &= \det(A - \lambda I) = p_A(\lambda). \end{aligned}$$

□

Posledica 2.20 *Naj bosta $A,\ B \in M_n$ podobni matriki. Potem imata matriki iste lastne vrednosti.*

Dokaz. Naj bo λ neka poljubno izbrana lastna vrednost matrike B (njen obstoj jamči izrek 2.10). Po izreku 2.14 vemo, da je λ ničla karakterističnega polinoma $p_B(\lambda)$. Po prejšnjem izreku, sta karakteristična polinoma podobnih matrik enaka, torej je λ tudi ničla karakterističnega polinoma $p_A(\lambda)$ in zato je tudi (znova po izreku 2.14) lastna vrednost matrike A. □

Posledica 2.21 *Naj bosta A, $B \in M_n$ podobni matriki. Če je matrika B diagonalna, potem so elementi na njeni glavni diagonali lastne vrednosti matrike A.*

Dokaz. Zaradi prejšnje posledice zadošča dokazati, da so diagonalni elementi diagonalne matrike B lastne vrednosti te matrike.
Naj bodo $b_{11},\ b_{22},\ \ldots,\ b_{nn}$ diagonalni elementi matrike B. Potem je matrika $B - \lambda I$ tudi diagonalna, saj jo dobimo kot razliko diagonalnih matrik.
Potem je $\det(B - \lambda I) = (b_{11} - \lambda) \cdots (b_{nn} - \lambda)$, saj vemo, da je determinanta diagonalne matrike enaka produktu diagonalnih elementov. Ta determinanta pa je karakteristični polinom matrike B. Ničle tega polinoma so: $b_{11},\ \ldots,\ b_{nn}$ in zato so to tudi lastne vrednosti matrike B. □

Izrek 2.22 *Če sta A in $B \in M_n$ podobni matriki, potem imata enaki sledi.*

Dokaz. Ker sta matriki A in B podobni velja, da obstaja nesingularna matrika S, da je $B = S^{-1}AS$. Po posledici 1.27 vemo, da je $\text{sled}(ABC) = \text{sled}(BCA)$. Izračunajmo: $\text{sled}(B) = \text{sled}\left(S^{-1}AS\right) = \text{sled}\left(ASS^{-1}\right) = \text{sled}\left(AI\right) = \text{sled}(A)$. □

Kot smo opazili v posledici 2.21, so diagonalni elementi diagonalnih matrik kar lastne vrednosti te matrike. Zaradi te lepe lastnosti, so diagonalne matrike zelo priročne za izračun lastnih vrednosti. Zato se pojavi vprašanje, katere matrike so podobne diagonalnim matrikam. To vprašanje bo v ospredju v nadaljevanju poglavja.

Definicija 2.23 *Matriko $A \in M_n$, ki je podobna neki diagonalni matriki D, imenujemo diagonalizabilna matrika.*

Izrek 2.24 *Naj bo matrika $A \in M_n$. Potem velja: (i) Matrika A je podobna bločni matriki oblike $\begin{bmatrix} \Lambda & B \\ 0 & C \end{bmatrix}$, kjer je $\Lambda = \text{diag}\,(\lambda_1,\ \ldots,\ \lambda_k)$, $C \in M_{n-k}$ in $B \in M_{k,n-k}$, $1 \leq k < n$ natanko tedaj, ko obstaja k linearno neodvisnih vektorjev iz $\mathbb{C}^n$, ki so lastni vektorji pripadajoči lastnim vrednostim matrike A.*
(ii) Matrika A je diagonalizabilna natanko tedaj, ko obstaja n linearno neodvisnih vektorjev pripadajočih lastnim vrednostim matrike A. Če so $x_1,\ x_2,\ \ldots,\ x_n$ linearno neodvisni in je $S = [x_1\ x_2\ \ldots\ x_n]$, potem je $S^{-1}AS$ diagonalna matrika.
(iii) Če je A podobna matriki iz točke (i), potem so diagonalni elementi matrike Λ lastne vrednosti matrike A. Če je A podobna matriki $S^{-1}AS$ iz točke (ii), potem so diagonalni elementi te matrike vse lastne vrednosti matrike A.

Dokaz. Najprej si oglejmo točko (i): ($\Leftarrow$) Predpostavimo, da je $k < n$ in naj bodo $x_1,\ x_2,\ \ldots,\ x_k$ linearno neodvisni vektorji iz prostora $\mathbb{C}^n$, za katere je $Ax_i = \lambda_i x_i$, za vsak $i = 1,\ 2,\ \ldots,\ k$. Naj bo $\Lambda = \operatorname{diag}(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_k)$ in $S_1 = [x_1\ x_2\ \ldots\ x_k]$. Izberimo matriko $S_2 \in M_{n,n-k}$ tako, da bo matrika $[S_1\ S_2]$ nesingularna. To lahko naredimo, saj je matrika S_1 sestavljena iz k linearno neodvisnih stolpcev, teh je manj, kot je razsežnost vektorskega prostora. V bistvu so stolpci matrike S_2 vektorji s katerimi vektorje $x_1,\ x_2,\ \ldots,\ x_k$ dopolnimo do baze prostora $\mathbb{C}^n$.
Potem je:

$$\begin{aligned}
&S^{-1}AS = S^{-1}A\,[S_1\ S_2] = S^{-1}\,[AS_1\ AS_2] = S^{-1}\,[Ax_1\ Ax_2\ \ldots\ Ax_k\ \ AS_2] = \\
&= S^{-1}\,[\lambda_1 x_1\ \lambda_2 x_2\ \ldots\ \lambda_k x_k\ \ AS_2] = \left[\lambda_1 S^{-1}x_1\ \lambda_2 S^{-1}x_2\ \ldots\ \lambda_k S^{-1}x_k\ \ S^{-1}AS_2\right] = \\
&= \left[\lambda_1 e_1\ \lambda_2 e_2\ \ldots\ \lambda_k e_k\ \ S^{-1}AS_2\right] = \begin{bmatrix}\Lambda & B\\ 0 & C\end{bmatrix},
\end{aligned}$$

kjer je $\Lambda = \operatorname{diag}(\lambda_1,\ \ldots,\ \lambda_n)$ in $S^{-1}AS_2 = \begin{bmatrix}B\\ C\end{bmatrix} \in M_{n,n-k}$.

($\Rightarrow$) Naj bo S nesingularna matrika in $S^{-1}AS = \begin{bmatrix}\Lambda & B\\ 0 & C\end{bmatrix}$. Zapišimo S kot $S = [S_1\ \ S_2]$, kjer je $S_1 \in M_{n,k}$, ki ima k linearno neodvisnih stolpcev. Velja:

$$\begin{aligned}
S^{-1}AS &= \begin{bmatrix}\Lambda & B\\ 0 & C\end{bmatrix}\\
AS &= S\begin{bmatrix}\Lambda & B\\ 0 & C\end{bmatrix}\\
[AS_1\ \ AS_2] &= [S_1\ \ S_2]\begin{bmatrix}\Lambda & B\\ 0 & C\end{bmatrix}\\
[AS_1\ \ AS_2] &= [S_1\Lambda\ \ S_1B + S_2C]\,.
\end{aligned}$$

Da bo enakost matrik izpolnjena, mora biti $AS_1 = S_1\Lambda$, torej je vsak stolpec matrike S_1 lastni vektor matrike A.
(ii) ($\Leftarrow$) Naj bo $k = n$, potem množica vektorjev $\{x_1,\ x_2,\ \ldots,\ x_n\}$ predstavlja bazo prostora $\mathbb{C}^n$. Ker so $x_1,\ x_2,\ \ldots,\ x_n$ lastni vektorji velja: $Ax_i = \lambda_i x_i$, za vsak $i = 1,\ 2,\ \ldots,\ n$. Naj bo $\Lambda = \operatorname{diag}(\lambda_1,\ \ldots, \lambda_n)$ in $S = [x_1\ \ldots\ x_n]$. Matrika S je nesingularna, saj so njeni stolpci linearno neodvisni. Velja:

$$\begin{aligned}
&S^{-1}AS = S^{-1}\,[Ax_1\ Ax_2\ \ldots\ Ax_n] = S^{-1}\,[\lambda_1 x_1\ \lambda_2 x_2\ \ldots\ \lambda_n x_n] = \\
&= [\lambda_1 e_1\ \lambda_2 e_2\ \ldots\ \lambda_n e_n] = \operatorname{diag}(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n) = \Lambda.
\end{aligned}$$

($\Rightarrow$) Če je S nesingularna matrika, potem premore n linearno neodvisnih stolpcev. Zato velja: $S^{-1}AS = \Lambda$. Če enakost pomnožimo z matriko S z leve dobimo: $AS = S\Lambda$. Od tod, razmislek je podoben kot v točki (i), razberemo, da je vsak stolpec matrike S lastni vektor za pripadajočo lastno vrednost matrike A.
(iii) V izreku 2.14 smo pokazali, da so ničle karakterističnega polinoma lastne vrednosti matrike. Ločimo: (i) naj bo $k < n$, potem je $p_A(\lambda) = p_\Lambda(\lambda) \cdot p_C(\lambda)$. Torej so lastne vrednosti matrike Λ tudi lastne vrednosti matrike A. V točki (ii) je $k = n$ in velja $p_A(\lambda) = p_\Lambda(\lambda)$. Izrek 2.19 in posledica 2.20 jamčita, da so lastne vrednosti podobnih matrik A in Λ enake. □

Na dokaz zgornjega izreka lahko gledamo kot na algoritem za (delno) diagonalizacijo kvadratne matrike $A \in M_n$. Ta pravi, da je potrebno poiskati vse lastne vrednosti matrike A in njim pripadajoče lastne vektorje, ki morajo biti med seboj linearno neodvisni. S tem, da te vektorje vpišemo v stolpce matrike S, dobimo nesingularno matriko. Če je linearno neodvisnih vektorjev manj kot n, potem matriko dopolnimo s poljubnimi vektorji, ki so linearno neodvisni z že vpisanimi lastnimi vektorji. Naslednja lema olajša delo, saj podaja informacijo o linearni neodvisnosti lastnih vektorjev, ki pripadajo različnim lastnim vrednostim.

Lema 2.25 *Naj bodo* $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_k$ $(k \geq 2)$ *paroma različne lastne vrednosti matrike* $A \in M_n$ *in naj bo* x_i *lastni vektor pripadajoč lastni vrednosti* λ_i*, za vsak* $i = 1,\ 2,\ \ldots,\ k$*. Potem so vektorji* $x_1,\ x_2,\ \ldots,\ x_k$ *linearno neodvisni.*

Dokaz. Denimo, da so vektorji $x_1,\ x_2,\ \ldots,\ x_k$ linearno odvisni. Potem obstajajo taki koeficienti $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_k \in \mathbb{C}$, ki niso vsi enaki 0, da je:

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0. \tag{2.2}$$

Z B_1 označimo matriko: $(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_k I)$. Ker je x_i lastni vektor pripadajoč lastni vrednosti λ_i, velja: $B_1 x_i = (A - \lambda_2 I) \cdots (A - \lambda_i I) \cdots (A - \lambda_k I)\, x_i = (\lambda_i - \lambda_2) \cdots (\lambda_i - \lambda_i) \cdots (\lambda_i - \lambda_k)\, x_i = 0$, za $i = 2,\ \ldots, k$, in $B_1 x_1 = (\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_k)\, x_1 \neq 0$.
Če B_1 pomnožimo z enačbo 2.2 dobimo: $0 = B_1(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k) = \alpha_1 B_1 x_1 + \alpha_2 B_1 x_2 + \ldots + \alpha_k B_1 x_k = \alpha_1 (\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_k)\, x_1 + 0 + \ldots + 0 = \alpha_1 B_1 x_1$. Od tod sledi, da mora biti $\alpha_1 = 0$, saj vemo, da je produkt $B_1 x_1$ neničeln.
Naj bo $B_j = \prod_{i=1,\ i \neq j}^{n} (A - \lambda_i I)$, za $j = 2,\ \ldots, n$. Za vsak j ponovimo zgoraj opisani postopek, ki smo ga izvedli za $j = 1$. Tako dobimo, da je $\alpha_j = 0$, za vsak $j = 1,\ 2,\ \ldots,\ k$. Velja: $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, kar je v nasprotju s predpostavko, da so vektorji linearno odvisni. Vektorji $x_1,\ x_2,\ \ldots,\ x_k$ so torej linearno neodvisni. □

Izrek 2.26 *Če ima matrika $A \in M_n$ n različnih lastnih vrednosti, potem je A diagonalizabilna.*

Dokaz. Po lemi 2.25 velja, da so vektorji pripadajoči lastnim vrednostim linearno neodvisni, saj so vse lastne vrednosti matrike A med seboj različne. Potem po točki (ii) izreka 2.24 velja, da je matrika A podobna diagonalni matriki $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Torej je matrika A diagonalizabilna. □

Zgornji izrek ne da veliko informacij o diagonalizabilnih matrikah, saj implikacija velja samo v eno smer, torej, če ima matrika same različne lastne vrednosti, potem je diagonalizabilna. To pa še ne pomeni, da matrika, ki ima lastne vrednosti z algebrajsko večkratnostjo večjo od ena, ni diagonalizabilna. Zato bo naslednji izrek podal nekoliko drugačno karakterizacijo diagonalizabilnih matrik. Da ga bomo lahko dokazali, je potrebna naslednja trditev.

Trditev 2.27 *Naj bosta matriki $A \in M_n$ in $B \in M_n$ podobni. Tedaj je $\mathrm{N}(A) = \mathrm{N}(B)$.*

Dokaz. Naj bo $B = \{x_1, x_2, \ldots, x_k\}$ baza vektorskega prostora $\operatorname{Ker} B$, ki je dimenzije k. Matriki A in B sta podobni, zato obstaja neka nesingularna matrika S, da velja: $B = S^{-1}AS$ oziroma: $SB = AS$. Potem za poljuben vektor x_i iz baze jedra matrike B velja $ASx_i = SBx_i = S0 = 0$. Opazimo, da so vektorji Sx_i, za vsak $i \in \{1, \ldots, k\}$, v jedru matrike A.
S pomočjo protislovja pokažimo, da so vektorji $Sx_1, Sx_2, \ldots, Sx_k$ linearno neodvisni. Denimo, da so vektorji linearno odvisni. Potem obstajajo taki skalarji $\alpha_1, \alpha_2, \ldots, \alpha_k$, ki niso vsi enaki 0, da velja:

$$\begin{aligned} \alpha_1 Sx_1 + \alpha_2 Sx_2 + \ldots + \alpha_k Sx_k &= 0 \\ S(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k) &= 0 \\ \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k &= 0 \end{aligned}$$

Vemo, da vektorji $x_1, x_2, \ldots, x_k$ tvorijo bazo prostora $\operatorname{Ker} B$, zato so linearno neodvisni. Torej morajo biti vsi skalarji $\alpha_1, \ldots, \alpha_k$ enaki 0. To pa je v nasprotju s predpostavko o linearni odvisnosti vektorjev $Sx_1, \ldots, Sx_k$.
Sledi, da je množica $\{Sx_1, Sx_2, \ldots, Sx_k\}$ linearno neodvisna, hkrati pa velja, da vsak vektor iz te množice pripada jedru matrike A. Zato je množica $\{Sx_1, Sx_2, \ldots, Sx_k\}$ podmnožica baze jedra matrike A. Torej: $\mathrm{N}(B) \leq \mathrm{N}(A)$. Če obrnemo argumente, dobimo še obratno neenakost in zato velja: $\mathrm{N}(A) = \mathrm{N}(B)$. □

Izrek 2.28 *Matrika $A \in M_n$ je diagonalizabilna natanko tedaj, ko za vsako lastno vrednost λ matrike A velja, da je njena geometrijska večkratnost enaka algebrajski.*

Dokaz. ($\Rightarrow$) Denimo, da je matrika $A \in M_n$ diagonalizabilna. To pomeni, da obstaja nesingularna matrika S, da velja: $D = S^{-1}AS$, kjer je D diagonalna matrika. Po posledici 2.20 vemo, da imata matriki A in D enake lastne vrednosti. Naj bo λ_i neka lastna vrednost matrik A in D z algebrajsko večkratnostjo m_i. Preverimo, da sta matriki $A-\lambda_i I$ in $D-\lambda_i I$ podobni:

$$D - \lambda_i I = S^{-1}AS - \lambda_i S^{-1}S = S^{-1}\left(A - \lambda_i I\right)S.$$

Potem po prejšnji posledici velja, da je $\mathrm{N}\left(A - \lambda_i I\right) = \mathrm{N}\left(D - \lambda_i I\right)$. Ker so na diagonali matrike D njene lastne vrednosti, bo imela matrika $D - \lambda_i I$ $n - m_i$ neničelnih vrstic, torej bo njen rang enak $n - m_i$. Geometrijska večkratnost lastne vrednosti λ_i bo torej $n - \mathrm{rang}\left(A - \lambda_i I\right) = m_i$.

($\Leftarrow$) Naj bodo $\lambda_1, \ldots, \lambda_k$ paroma različne lastne vrednosti matrike $A \in M_n$. Naj bo λ_i neka lastna vrednost matrike A, za katero sta geometrijska in algebrajska večkratnost enaki m_i. Potem obstaja (po definiciji geometrijske večkratnosti) m_i linearno neodvisnih lastnih vektorjev, ki so pripadajoči tej lastni vrednosti. Označimo jih z $x_1^i,\ x_2^i,\ \ldots,\ x_{m_i}^i$. Po osnovnem izreku algebre velja, da je vsota vseh algebrajskih večkratnosti lastnih vrednosti matrike A enaka n, torej je taka tudi vsota geometrijskih večkratnosti lastnih vrednosti. Zato je v množici vseh lastnih vektorjev matrike A n lastnih vektorjev. Matrika A bo (po izreku 2.24) diagonalizabilna, če bo množica vektorjev $x_1^1,\ \ldots,\ x_{m_1}^1,\ \ldots x_1^k,\ \ldots,\ x_{m_k}^k$ linearno neodvisna.
Naj bodo $\alpha_1^1,\ \ldots,\ \alpha_{m_1}^1,\ \ldots,\ \alpha_1^k,\ \ldots,\ \alpha_{m_k}^k \in \mathbb{C}$ taki skalarji, da velja:

$$\alpha_1^1 x_1^1 + \ldots + \alpha_{m_1}^1 x_{m_1}^1 + \ldots + \alpha_1^k x_1^k + \ldots + \alpha_{m_k}^k x_{m_k}^k = 0$$
$$\sum_{j=1}^{m_1} \alpha_j^1 x_j^1 + \ldots + \sum_{j=1}^{m_k} \alpha_j^k x_j^k = 0$$

Linearna kombinacija lastnih vektorjev pripadajočih lastni vrednosti λ_i je bodisi lastni vektor za to lastno vrednost bodisi ničelni vektor. Ker pa po lemi 2.25 velja, da so lastni vektorji pripadajoči različnim lastnim vrednostim linearno neodvisni, mora veljati: $\sum_{j=1}^{m_i} \alpha_j^i x_j^i = 0$, za vsak $i = 1,\ \ldots,\ k$. Iz tega sledi, da so vsi skalarji $\alpha_1^1,\ \ldots,\ \alpha_{m_1}^1,\ \ldots,\ \alpha_1^k,\ \ldots,\ \alpha_{m_k}^k$ enaki 0.
Ugotovili smo, da so vektorji $x_1^1,\ \ldots\ x_{m_1}^1, \ldots x_1^k,\ \ldots\ x_{m_k}^k$ linearno neodvisni. Zato velja, da je matrika A diagonalizabilna. □

Posledica 2.29 *Matrika $A \in M_n$ je diagonalizabilna natanko tedaj, ko je vsota geometrijskih večkratnosti lastnih vrednosti matrike A enaka n.*

Dokaz. Posledica sledi neposredno iz zgornjega izreka in iz osnovnega izreka algebre. □

Zgled. Vrnimo se k simetrični matriki $A = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$, ki smo jo obravnavali v prejšnjem podpoglavju, in pokažimo, da je matrika diagonalizabilna.
Izrek 2.28 zagotavlja, da je matrika diagonalizabilna, če sovpadajo geometrijske in algebrajske večkratnosti vseh lastnih vrednosti matrike A. V prejšnjih zgledih smo preverili, da to drži, zato lahko zapišemo nesingularno matriko S, ki jo dobimo tako, da v stolpce matrike vpišemo lastne vektorje iz lastnih podprostorov. V lastnem podporstoru V_1 se nahaja vektor $(1, 1, 1)$, v lastnem podprostoru V_4 pa na primer vektorja $(1, 0, -1)$ in $(0, 1, -1)$.

$$S = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \qquad S^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

Potem je $S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

Diagonalizacija matrike je med drugim lahko uporabna za izračun potence dane matrike. S pomočjo zgornjih podatkov izračunajmo vrednost matrike A^{50}:

$$\begin{aligned} A^{50} &= \left(SDS^{-1}\right)^{50} \\ A^{50} &= SD^{50}S^{-1} \\ A^{50} &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4^{50} & 0 \\ 0 & 0 & 4^{50} \end{bmatrix} \cdot \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix} \\ A^{50} &= \frac{1}{3}\begin{bmatrix} 1+2^{101} & 1-2^{100} & 1-2^{100} \\ 1-2^{100} & 1+2^{101} & 1-2^{100} \\ 1-2^{100} & 1-2^{100} & 1+2^{101} \end{bmatrix} \end{aligned}$$

Poglavje 3

Unitarne matrike

3.1 Osnovne definicije in izreki

V tem podpoglavju bo predstavljena osnovna definicija unitarnih matrik in izrek, ki podaja ekvivalentno karakterizacijo unitarnih matrik. Za tem bodo predstavljeni nekateri temeljni rezultati povezani z unitarnimi matrikami.

Definicija 3.1 *Matrika $U \in M_n$ je unitarna, če velja $U^*U = I$. Matrika $O \in M_n(\mathbb{R})$ je ortogonalna, če je $O^TO = I$.*

Zgled. Naj bo dana matrika $A = \begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{bmatrix}$. Pokažimo, da je matrika unitarna.

$$A^* = \begin{bmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1-i}{2} \end{bmatrix} \longrightarrow A^*A = \begin{bmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1-i}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Kljub temu, da je definicija unitarne matrike precej elementarna, se izkaže, da imajo unitarne matrike veliko lepih lastnosti. Tako je na primer absolutna vrednost determinante unitarne matrike enaka 1. Zanimiva je tudi lastnost, da je direktna vsota dveh unitarnih matrik spet unitarna matrika. Direktna vsota matrik je definirana kot $A \oplus B = \operatorname{diag}(A, B)$, kjer sta A in B matriki poljubnih dimenzij. Zapis lahko razširimo tudi na več matrik, tedaj pišemo $\oplus_{k=1}^{n} A_k = \operatorname{diag}(A_1, \ \ldots, A_n)$. Ilustrirajmo zapisano z zgledom:

Zgled. Zapišimo direktno vsoto matrik A in B:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 7 & 8 \\ 9 & -1 \end{bmatrix} \qquad A \oplus B = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 & 0 \\ 0 & 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 9 & -1 \end{bmatrix}$$

Trditev 3.2 *Matriki $U \in M_n$ in $V \in M_m$ sta unitarni natanko tedaj, ko je matrika $U \oplus V \in M_{n+m}$ unitarna.*

Dokaz. ($\Rightarrow$) Vemo, da sta matriki U in V unitarni. Potem je:

$$(U \oplus V)^* (U \oplus V) = \begin{bmatrix} U^* & 0 \\ 0 & V^* \end{bmatrix} \cdot \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} = \begin{bmatrix} U^*U & 0 \\ 0 & V^*V \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_m \end{bmatrix} = I_{m+n}$$

($\Leftarrow$) Ker je matrika $U \oplus V$ unitarna, velja: $(U \oplus V)^* (U \oplus V) = I_{m+n}$. Da bo enakost izpolnjena mora veljati: $U^*U = I_n$ in $V^*V = I_m$. Torej sta matriki U in V unitarni. □

Trditev 3.3 *Če je matrika $U \in M_n$ unitarna, potem velja: $|\det U| = 1$.*

Dokaz. Ker je matrika U unitarna velja: $U^*U = I$. Od tod dobimo: $1 = \det I = \det(UU^*) = \det U \det U^* = \det U \det \overline{U^T} = \det U \overline{\det U} = |\det U|^2$. Sledi: $|\det U| = 1$. □

Izrek 3.4 *Če je $U \in M_n$, potem so naslednje trditve ekvivalentne:*

(a) U je unitarna.

(b) U je nesingularna in $U^* = U^{-1}$.

(c) $UU^* = I$.

(d) U^* je unitarna.

(e) Stolpci matrike U tvorijo ortonormirano bazo prostora $\mathbb{C}^n$.

(f) Vrstice matrike U tvorijo ortonormirano bazo prostora $\mathbb{C}^n$

(g) Za vse $x \in \mathbb{C}^n$ velja $\|x\| = \|Ux\|$, kar pomeni, da imata x in Ux isto Evklidsko normo.

Dokaz. $(a \Rightarrow b)$ Ker je matrika U unitarna, je $U^*U = I$, kar pomeni, da je matrika U^* inverz matrike U; velja: $U^* = U^{-1}$. Ker ima matrika U inverz, je nesingularna.

$(a \Leftarrow b)$ Vemo, da je U nesingularna in $U^* = U^{-1}$. Če to enakost pomnožimo z U iz desne dobimo: $U^*U = U^{-1}U = I$.

$(b \Leftrightarrow c)$ Sledi neposredno iz definicije inverzne matrike: $AB = BA = I$. Torej, če je U nesingularna, velja: $U^*U = I \Leftrightarrow UU^* = I$. Obrat te trditve je očiten.

$(c \Leftrightarrow d)$ Ker velja $(U^*)^* = U$, je matrika U^* po definiciji unitarnosti: $(U^*)^* U^* = UU^* = I$ unitarna. Obrat trditve je očiten.

$(a \Rightarrow e)$ Naj bodo $u_1,\ u_2,\ \ldots,\ u_n$ zaporedni stolpci matrike U. Ker je matrika $U = [u_1\ u_2\ \ldots\ u_n]$ unitarna velja:

$$U^*U = \begin{bmatrix} u_1^* \\ u_2^* \\ \ldots \\ u_n^* \end{bmatrix} \cdot \begin{bmatrix} u_1 & u_2 & \ldots & u_n \end{bmatrix} = \begin{bmatrix} u_1^*u_1 & u_1^*u_2 & \ldots & u_1^*u_n \\ u_2^*u_1 & u_2^*u_2 & \ldots & u_2^*u_n \\ \vdots & \vdots & \ddots & \vdots \\ u_n^*u_1 & u_n^*u_2 & \ldots & u_n^*u_n \end{bmatrix} = I$$

Opazimo, da velja $u_j^* u_i = \langle u_i, u_j \rangle = \begin{cases} 1, & i = j; \\ 0 & i \neq j, \end{cases}$ kar pomeni, da stolpci matrike U tvorijo ortonormirano bazo prostora $\mathbb{C}^n$.

$(a \Leftarrow e)$ Naj bodo $u_1,\ u_2,\ \ldots,\ u_n$ stolpci matrike U. Naj bo $A = U^*U$:

$$A = U^*U = \begin{bmatrix} u_1^*u_1 & u_1^*u_2 & \ldots & u_1^*u_n \\ u_2^*u_1 & u_2^*u_2 & \ldots & u_2^*u_n \\ \vdots & \vdots & \ddots & \vdots \\ u_n^*u_1 & u_n^*u_2 & \ldots & u_n^*u_n \end{bmatrix}.$$

Ker so stolpci matrike U ortonormirani vektorji, zanje po opombi 1.39 velja: $\langle u_i, u_j \rangle = u_j^* u_i = \begin{cases} 1, & i = j; \\ 0 & i \neq j, \end{cases}$ in zato je:

$$A = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}.$$

Produkt U^*U, kjer stolpci matrike U tvorijo ortonormirano bazo prostora $\mathbb{C}^n$, je enak identični matriki. Torej je matrika U unitarna.

($e \Leftrightarrow f$) Naj bo U matrika z ortonormiranimi vrsticami. Če vpeljemo novo matriko $V := U^T$, potem ima matrika V ortonormirane stolpce in zato zanjo po prejšnjem koraku dokaza velja, da je unitarna. Velja pa: $I = V^*V = \left(U^T\right)^* U^T = \left(UU^*\right)^T = I^T = I$. Vidimo, da je tudi matrika U z ortonormiranimi vrsticami unitarna.

($a \Rightarrow g$) Naj bo vektor $y \in \mathbb{C}^n$ definiran kot $y := Ux$, kjer je U unitarna matrika in $x \in \mathbb{C}^n$ vektor. Potem je: $\|y\|^2 = \langle y, y\rangle = \langle Ux, Ux\rangle = (Ux)^* Ux = x^*U^*Ux = x^*Ix = x^*x = \langle x, x\rangle = \|x\|^2$. Po lastnosti Evklidske norme je ta vedno večja ali enaka od 0, zato je: $\|Ux\| = \|x\|$.

($a \Leftarrow g$) Naj bo dana matrika $U \in M_n$, da je $U^*U = A$. Dokazati želimo, da je matrika A enaka identični matriki.
Naj bosta $z,\ w \in \mathbb{C}^n$ poljubna vektorja in definirajmo vektor $x \in \mathbb{C}^n$ kot vsoto vektorjev z in w. Podobno kot v prejšnjem koraku dokaza je definiran tudi $y := Ux$. Potem je:

$$\begin{aligned}
&\langle x, x\rangle = (z+w)^*(z+w) = z^*z + z^*w + w^*z + w^*w = \\
&= \langle z, z\rangle + \langle w, w\rangle + \langle z, w\rangle + \overline{\langle z, w\rangle} = \langle z, z\rangle + \langle w, w\rangle + 2\,\mathrm{Re}\left(\langle z, w\rangle\right) \text{ in} \\
&\langle y, y\rangle = (Ux)^* Ux = x^*U^*Ux = (z+w)^* U^*U(z+w) = \\
&= z^*Az + z^*U^*Uw + w^*U^*Uz + w^*Aw = z^*Az + z^*U^*Uw + \overline{z^*U^*Uw} + w^*Aw = \\
&= z^*Az + 2\,\mathrm{Re}\left(z^*Aw\right) + w^*Aw = \langle Az, z\rangle + \langle Aw, w\rangle + 2\,\mathrm{Re}\left(\langle Aw, z\rangle\right).
\end{aligned}$$

Po predpostavki izreka velja: $\|x\| = \|Ux\|$, za vsak $x \in \mathbb{C}^n$, torej mora veljati : $\langle x, x\rangle = \langle y, y\rangle$. Ker je $z \in \mathbb{C}^n$, po predpostavki velja: $\langle z, z\rangle = \langle Uz, Uz\rangle = z^*Az = \langle Az, z\rangle$. Podobno velja tudi za $w \in \mathbb{C}^n$. Od tod sledi, da mora biti

$$\mathrm{Re}\langle w, z\rangle = \mathrm{Re}\langle Aw, z\rangle. \tag{3.1}$$

Naj bosta $z := e_p$ in $w := ie_q$, kjer sta $p, q \in \{1,\ \ldots,\ n\}$ poljubna. Potem je:

$$\begin{aligned}
&\mathrm{Re}\langle w, z\rangle = \mathrm{Re}\left(e_p^T ie_q\right) = \mathrm{Re}\left(ie_p^T e_q\right) = 0, \\
&\mathrm{Re}\langle Aw, z\rangle = \mathrm{Re}\left(e_p^T Aie_q\right) = \mathrm{Re}\left(ia_{pq}\right) = -\,\mathrm{Im}\, a_{pq}.
\end{aligned}$$

Ker mora veljati enakost 3.1, so vsi elementi matrike A realni, saj smo ugotovili, da so imaginarne komponente vseh elementov v matriki enake nič. Predpostavimo še, da sta vektorja z in w realna. Naj bosta $z := e_p$ in $w := e_q$, kjer sta $p, q \in \{1,\ \ldots,\ n\}$. Potem je:

$$\begin{aligned}
&\mathrm{Re}\langle e_q, e_p\rangle = e_p^T e_q = \begin{cases} 1, & p = q; \\ 0 & p \neq q. \end{cases} \\
&\mathrm{Re}\langle Ae_q, e_p\rangle = \mathrm{Re}\left(e_p^T Ae_q\right) = a_{pq}.
\end{aligned}$$

Matrika A ima torej same realne elemente, ki so na glavni diagonali enaki 1, povsod drugod pa ima ničle. Matrika A je identična matrika. □

Posledica 3.5 *Naj bosta $U, V \in M_n$ unitarni matriki. Potem je tudi UV unitarna matrika.*

Dokaz. Ker sta matriki U in V unitarni velja po točki (b) izreka 3.4, da je $U^{-1} = U^*$ in $V^{-1} = V^*$. Potem je: $(UV)^* (UV) = V^*U^*UV = V^* \left(U^{-1}U\right) V = V^{-1}V = I$. □

Opomba 3.6 *Zgornjo trditev lahko tudi posplošimo: Naj bodo U_1, U_2, ... U_n unitarne matrike. Potem je tudi matrika $U_1U_2 \cdots U_n$ unitarna. Ideja dokaza te trditve je podobna zgornji, saj operacija $*$ obrne vrstni red faktorjev v produktu.*

Ugotovili smo, da je ena izmed lastnosti unitarnih matrik ta, da je njihov inverz enak konjugirani-transponiranki dane matrike, torej za unitarno matriko U je $U^{-1} = U^*$. V naslednjem izreku bomo na to zvezo pogledali nekoliko splošneje, saj bomo zahtevali, da je matrika U^* podobna matriki U^{-1}. Izkaže se, da je inverz matrike podoben konjugirani-transponiranki te matrike tedaj, ko obstaja neka nesingularna matrika, da lahko začetno matriko zapišemo kot produkt inverza in konjugirane-transponiranke te nesingularne matrike.

Izrek 3.7 *Naj bo $A \in M_n$ nesingularna matrika. Matrika A^{-1} je podobna matriki A^* natanko tedaj, ko obstaja taka nesingularna matrika $B \in M_n$, da velja $A = B^{-1}B^*$.*

Dokaz. ($\Leftarrow$) Denimo, da obstaja taka nesingularna matrika $B \in M_n$, da zanjo velja: $A = B^{-1}B^*$. Ker je matrika A nesingularna, velja: $A^{-1} = (B^*)^{-1} B$. Pokazali bomo, da je $A^{-1} = (B^*)^{-1} A^*B^*$. Izračunajmo:

$$B^*A^{-1} (B^*)^{-1} = B^* (B^*)^{-1} B (B^*)^{-1} = B (B^*)^{-1} = \left(B^{-1}B^*\right)^* = A^*.$$

Matrika A^{-1} je podobna matriki A^*.
($\Rightarrow$) Naj bo matrika A^{-1} podobna matriki A^*. Potem po definiciji podobnosti obstaja taka nesingularna matrika $S \in M_n$, da je $A^{-1} = S^{-1}A^*S$. Če enakost z leve pomnožimo s S in z desne z A, dobimo: $S = A^*SA$.
Naj bo matrika $S_\theta = e^{i\theta}S$, kjer je $\theta \in \mathbb{R}$, takšna, da je: $S_\theta = A^*S_\theta A$ in $S_\theta^* = (A^*S_\theta A)^* = A^*S_\theta^* A$. Če obe enakosti seštejemo, dobimo:

$$S_\theta + S_\theta^* = A^* (S_\theta + S_\theta^*) A. \tag{3.2}$$

Naj bo $H_\theta := S_\theta + S_\theta^*$. Potem je $H_\theta^* = (S_\theta + S_\theta^*)^* = S_\theta^* + S_\theta$, ker je seštevanje matrik komutativno, lahko pišemo $H_\theta^* = S_\theta + S_\theta^* = H_\theta$ in opazimo, da je matrika H_θ Hermitska. Denimo, da je matrika H_θ singularna. Po trditvi 2.4 bi to veljalo, če bi bila ena izmed lastnih vrednosti matrike H_θ enaka 0. Potem bi veljalo: $H_\theta x = 0$, za nek neničeln vektor x, od tod pa bi sledilo:

$$\begin{aligned} S_\theta x + S_\theta^* x &= 0 \\ -S_\theta x &= S_\theta^* x \\ -x = S_\theta^{-1} S_\theta^* x &= \left(e^{i\theta}\right)^{-1} S^{-1} \left(e^{i\theta}\right)^* S^* x = e^{-i\theta} S^{-1} e^{-i\theta} S^* x = e^{-2i\theta} S^{-1} S^* x \\ -e^{2i\theta} x &= S^{-1} S^* x. \end{aligned}$$

Če izberemo za θ neko vrednost $\theta_0 \in [0, 2\pi)$, da $-e^{2i\theta_0}$ ne bo lastna vrednost matrike $S^{-1}S^*$, potem Hermitska matrika $H := H_{\theta_0}$ ne bo singularna. Iz enačbe 3.2 sledi:

$$H = A^* H A. \tag{3.3}$$

Naj bo α poljubno kompleksno število, da zanj velja $|\alpha| = 1$ in da α ni lastna vrednost matrike A. Definirajmo matriko $B := \beta\left(\alpha I - A^*\right) H$, kjer je $\beta \in \mathbb{C}\backslash\{0\}$ poljuben. Matrika $(\alpha I - A^*)$ je nesingularna, ker α ni lastna vrednost matrike A. Ker tudi matrika H ni singularna, je matrika B, ki jo dobimo kot produkt dveh nesingularnih matrik (glej 1.18), nesingularna. Iščemo taka $\alpha, \beta \in \mathbb{C}$, $\beta \neq 0$, da bo izpolnjena enakost: $A = B^{-1}B^*$, ki jo lahko zapišemo tudi tako: $BA = B^*$. Izračunajmo:

$$B^* = \left(\beta\left(\alpha I - A^*\right) H\right)^* = \overline{\beta} H^* \left(\overline{\alpha} I - (A^*)^*\right) = H^* \left(\overline{\alpha}\overline{\beta} I - \overline{\beta} A\right) = H\left(\overline{\alpha}\overline{\beta} I - \overline{\beta} A\right),$$

saj je matrika H Hermitska. Izračunajmo še produkt BA:

$$BA = \left(\beta\left(\alpha I - A^*\right) H\right) A = \beta\alpha H A - \beta A^* H A = \beta\alpha H A - \beta H = H\left(\alpha\beta A - \beta I\right),$$

kjer smo uporabili enakost 3.3. Sledi:

$$\overline{\alpha}\overline{\beta} I - \overline{\beta} A = \alpha\beta A - \beta I \Leftrightarrow \left(\overline{\alpha}\overline{\beta} + \beta\right) I = \left(\alpha\beta + \overline{\beta}\right) A,$$

kar bo enako samo tedaj, ko bo $\alpha\beta = -\overline{\beta}$. Tej enakosti bo zadoščeno, če na primer izberemo za $\alpha = e^{i\gamma}$ in za $\beta = e^{i\frac{\pi-\gamma}{2}}$. Potem bo:

$$\alpha\beta = e^{i\left(\gamma + \frac{\pi-\gamma}{2}\right)} = e^{i\frac{\pi+\gamma}{2}} = e^{i\frac{2\pi-\pi+\gamma}{2}} = e^{i\pi + i\frac{-\pi+\gamma}{2}} = e^{i\pi} e^{i\frac{-\pi+\gamma}{2}} = -\overline{\beta}.$$

Našli smo taka $\alpha, \beta \in \mathbb{C}$, $\beta \neq 0$, da obstaja nesingularna matrika B, da velja $A = B^{-1}B^*$.

S tem je izrek dokazan. □

Ortogonalne matrike

Ob koncu razdelka se nekoliko ustavimo še pri ortogonalnih matrikah. Kot je bilo zapisano zgoraj vse dokazane lastnosti za unitarne matrike veljajo tudi za ortogonalne matrike. Ker bomo v zadnjem poglavju ortogonalne in unitarne matrike obravnavali na nivoju grup, pa vseeno vpeljimo še ortogonalne matrike in navedimo izrek, ki podaja ključne lastnosti ortogonalnih matrik. Pred tem pa še dve opombi. V tem podrazdelku bo skalarni produkt definiran nekoliko drugače, in sicer bo za poljubna vektorja $x, y \in \mathbb{R}^n$ (realni) skalarni produkt definiran kot: $\langle x, y\rangle = y^T x$. Druga opomba se nanaša na elemente v ortogonalni matriki, saj bomo vedno predpostavili, da so elementi ortogonalne matrike realna števila. Z ortogonalnimi matrikami s kompleksnimi elementi se ne bomo ukvarjali, vsekakor pa je potrebno poudariti, da ortogonalne matrike s kompleksnimi elementi niso enake unitarnim matrikam, kar je razvidno že iz začetne definicije ortogonalne matrike.

Zapisali smo že osnovno definicijo ortogonalne matrike, ki pravi, da je matrika $O \in M_n(\mathbb{R})$ ortogonalna, če je $O^T O = I$. Pogoj lahko ekvivalentno preoblikujemo v pogoj $O^T = O^{-1}$. Če matriko O razpišemo po stolpcih dobimo: $O = \begin{bmatrix} o_1 & o_2 & \dots & o_n \end{bmatrix}$. Od tod sledi:

$$O^T O = \begin{bmatrix} o_1^T \\ o_2^T \\ \vdots \\ o_n^T \end{bmatrix} \cdot \begin{bmatrix} o_1 & o_2 & \dots & o_n \end{bmatrix} = \begin{bmatrix} o_1^T o_1 & o_1^T o_2 & \dots & o_1^T o_n \\ o_2^T o_1 & o_2^T o_2 & \dots & o_2^T o_n \\ \vdots & \vdots & \ddots & \vdots \\ o_n^T o_1 & o_n^T o_2 & \dots & o_n^T o_n \end{bmatrix} = I.$$

Zato velja: $o_i^T o_j = \begin{cases} 0, & i \neq j; \\ 1, & i = j. \end{cases}$ Od tod, podobno kot pri unitarni matriki, ugotovimo, da je matrika ortogonalna, če stolpci te matrike tvorijo ortonormirano bazo vektorskega prostora $\mathbb{R}^n$.

Naslednji izrek pove, da je matrika ortogonalna natanko tedaj, ko ohranja Evklidsko normo in skalarni produkt. Nekoliko neformalno bi lahko rekli, da je matrika ortogonalna, ko se pri množenju vektorjev k matriki z desne ohranjajo dolžine in koti med vektorji. Pri dolžini ciljamo na normo, pri kotih med vektorji pa na znano zvezo, ki velja za vektorje v $\mathbb{R}^n$ in pravi: naj bosta $x, y \in \mathbb{R}^n$, potem je: $\langle x, y\rangle = \|x\| \, \|y\| \cos\phi$, kjer ϕ predstavlja kot med vektorjema x in y. Dokaz te trditve za geometrijske vektorje (torej za vektorje v $\mathbb{R}^3$) je mogoče najti v [2] na strani 27. Razširitev za $\mathbb{R}^n$ pa v delu [1], poglavje 3.2.

Izrek 3.8 *Naj bo $A \in M_n(\mathbb{R})$. Potem so naslednje trditve ekvivalentne:*

(a) A je ortogonalna matrika,

(b) $\|Ax\| = \|x\|$*, za vse* $x \in \mathbb{R}^n$*,*

(c) $\langle Ax, Ay\rangle = \langle x, y\rangle$*, za vse* $x, y \in \mathbb{R}^n$*.*

Dokaz tega izreka bomo izpustili, saj je podoben dokazu izreka 3.4.

3.2 Posebna primera realnih ortogonalnih in unitarnih matrik

V tem podpoglavju bosta predstavljena dva posebna tipa realnih ortogonalnih matrik oziroma unitarnih matrik, gre za rotacije ravnine in Householderjeve matrike. Slednje so uporabne za konstrukcijo unitarnih matrik pri danih vektorjih.

3.2.1 Rotacije ravnine

Definicija 3.9 *Naj bo* $1 \leq i < j \leq n$ *in* $\Theta \in [0, 2\pi)$*. Definirajmo kvadratno matriko* $O(\Theta; i, j)$*, ki jo dobimo iz identične matrike* $I_n, n \geq 2$ *tako, da elementa* $(I)_{ii}$ *in* $(I)_{jj}$ *zamenjamo s* $\cos\Theta$*, element* $(I)_{ij}$ *s* $-\sin\Theta$ *in element* $(I)_{ji}$ *s* $\sin\Theta$*.*

Če upoštevamo zgornjo definicijo, potem ima matrika $O(\Theta; i, j) \in M_n$ naslednjo obliko:

$$O(\Theta; i, j) = \begin{bmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \cos\Theta & \cdots & -\sin\Theta & & \\ & & & \ddots & & & \\ & & \vdots & 1 & \vdots & & \\ & & & & \ddots & & \\ & & \sin\Theta & \cdots & \cos\Theta & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix}$$

Takšno matriko imenujemo rotacija ravnine ali Givensova rotacija.

Trditev 3.10 *Matrika* $O(\Theta; i, j)$ *je realna ortogonalna matrika.*

Dokaz. Naj bosta i, j, takšna, da velja $1 \leq i < j \leq n$ in $\Theta \in [0, 2\pi)$ poljuben. Pokazati moramo, da je $O(\Theta; i, j)^T \cdot O(\Theta; i, j) = I$.
S k in l označimo vrstice in stolpce matrike $O(\Theta; i, j)$, ki so različni od i in j. Zanje velja: $\left(O(\Theta; i, j)^T \cdot O(\Theta; i, j)\right)_{kl} = \begin{cases} 1, & k = l; \\ 0, & k \neq l. \end{cases}$ Preostane še preveriti produkte i-te vrstice in i-tega stolpca. V tem primeru dobimo: $\cos\Theta\cos\Theta + \sin\Theta\sin\Theta = 1$. Podobno je v primeru j-te vrstice in j-tega stolpca.
V primeru i-te vrstice in j-tega stolpca dobimo: $-\cos\Theta\sin\Theta + \cos\Theta\sin\Theta = 0$. Podobno izračunamo tudi produkt j-te vrstice in i-tega stolpca, ki je prav tako enak 0.
Ugotovili smo, da so elementi produkta enaki 0, za vse elemente, ki niso na glavni diagonali in enaki 1, če so na glavni diagonali. Produkt je identična matrika in matrika rotacije ravnine je ortogonalna. □

Zgled. Naj bo dan vektor $x = [1 \;\; 2 \;\; 3]^T \in \mathbb{R}^3$ in naj bo $O\left(\frac{\pi}{2}; 1, 3\right)$ rotacijska matrika. Izračunajmo kam se z dano preslikavo preslika vektor x.

$$O\left(\tfrac{\pi}{2}; 1, 3\right) x = \begin{bmatrix} \cos\left(\frac{\pi}{2}\right) & 0 & -\sin\left(\frac{\pi}{2}\right) \\ 0 & 1 & 0 \\ \sin\left(\frac{\pi}{2}\right) & 0 & \cos\left(\frac{\pi}{2}\right) \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}.$$

Trditev 3.11 *Velja:* $O(\Theta; i, j)^{-1} = O(-\Theta; i, j)$.

Dokaz. V trditvi 3.10 smo pokazali, da so rotacijske matrike ortogonalne. Zato zanje velja: $O(\Theta; i, j)^{-1} = O(\Theta; i, j)^T$, kar pomeni, da je (i, j)-ti element matrike $O(\Theta; i, j)^T$ enak $\sin\Theta$, (j, i)-ti element pa je $-\sin\Theta$.
Funkcija sinus je liha, zato velja: $\sin(-\Theta) = -\sin\Theta$, funkcija kosinus pa soda, zato je $\cos(-\Theta) = \cos\Theta$. Potem sta (i, i)-ti in (j, j)-ti element matrike $O(-\Theta; i, j)$ enaka $\cos\Theta$, (i, j)-ti element je $\sin\Theta$, (j, i)-ti element matrike $O(-\Theta; i, j)$ pa je $-\sin\Theta$. Opazimo, da sta matriki $O(\Theta; i, j)^{-1}$ in $O(-\Theta; i, j)$ enaki, kar smo želeli pokazati. □

3.2.2 Householderjeve matrike

Definicija 3.12 *Naj bo $w \in \mathbb{C}^n$ neničelni vektor. Potem je Householderjeva matrika U_w definirana takole: $U_w = I - 2\langle w, w\rangle^{-1} ww^*$, kjer je $\langle w, w\rangle = w^*w$ Hermitski produkt.*

Posledica 3.13 *Če je w enotski vektor, potem velja $U_w = I - 2ww^*$.*

V nadaljevanju bomo dokazali nekaj trditev, ki veljajo za Householderjeve matrike in kažejo na njihovo splošno uporabnost.

Trditev 3.14 *Householderjeve matrike so unitarne in Hermitske.*

Dokaz. Najprej preverimo, da so Householderjeve matrike Hermitske. Ob tem bomo upoštevali, da je Hermitski produkt vektorja s samim seboj realno število.

$$U_w^* = \left(I - 2\left(w^*w\right)^{-1} ww^*\right)^* = I - 2\left(w^*w\right)^{-1}\left(w^*\right)^* w^* = I - 2\left(w^*w\right)^{-1} ww^* = U_w.$$

Preostane še, da pokažemo, da so Householderjeve matrike unitarne. Izračunajmo:

$$\begin{aligned}
&U_w^* U_w = U_w U_w = \left(I - 2\left(w^*w\right)^{-1} ww^*\right)\left(I - 2\left(w^*w\right)^{-1} ww^*\right) = \\
&= I - 2I\left(w^*w\right)^{-1} ww^* - 2\left(w^*w\right)^{-1} ww^* I + 2\left(w^*w\right)^{-1} ww^* 2\left(w^*w\right)^{-1} ww^* = \\
&= I - 4\left(w^*w\right)^{-1} ww^* + 4\left(w^*w\right)^{-2} w\left(w^*w\right) w^* = \\
&= I - 4\left(w^*w\right)^{-1} ww^* + 4\left(w^*w\right)^{-1} ww^* = I.
\end{aligned}$$

□

Definicija 3.15 *Če je S vektorski podprostor prostora $\mathbb{C}^n$, potem množico $S^\perp = \{x \in \mathbb{C}^n; \langle x, y\rangle = 0,$ za vse vektorje $y \in S\}$ imenujemo ortogonalni komplement podprostora S.*

Opazimo, da ima trivialni podprostor $\{0\}$ prostora S za ortogonalni komplement celoten prostor $\mathbb{C}^n$, po drugi strani pa je ortogonalni komplement prostora $\mathbb{C}^n$ podprostor $\{0\}$.

Trditev 3.16 *Naj bo $U_w \in M_n$ Householderjeva matrika in $S = \mathscr{L}(\{w\})$. Če vektorju iz vektorskega prostora $S^\perp$ na levi strani primnožimo matriko U_w, potem matrika deluje kot identična preslikava, na enodimenzionalnem prostoru porojenem z vektorjem w pa isto množenje deluje kot zrcaljenje. Torej: za vsak $x \in S^\perp : U_w x = x$ in za vsak $y \in \mathscr{L}(\{w\}) : U_w y = -y$.*

Dokaz. Naj bo $x \in S^\perp$ poljuben. Potem je: $U_w x = \left(I - 2\left(w^*w\right)^{-1} ww^*\right) x = x - 2\left(w^*w\right)^{-1} w\left(w^*x\right) = x - 2\left(w^*w\right)^{-1} w \cdot 0 = x$.
Ker je y iz enorazsežnega vektorskega prostora porojenega z vektorjem w lahko pišemo: $y = \alpha w$, kjer je $\alpha \in \mathbb{C}^n$. Potem velja: $U_w y = U_w(\alpha w) = \left(I - 2\left(w^*w\right)^{-1} ww^*\right)\alpha w = \alpha w - 2\alpha\left(w^*w\right)^{-1} w\left(w^*w\right) = \alpha w - 2\alpha w = -\alpha w = -y$. □

Izračunamo lahko tudi determinanto in lastne vrednosti za vse Householderjeve matrike. Izkaže se, da je determinanta Householderjeve matrike enaka -1, kar ustreza trditvi o unitarnih matrikah, ki smo jo izpeljali v prvem podpoglavju. Lastne vrednosti Householderjeve matrike pa so enake -1 in 1, slednja lastna vrednost ima algebrajsko večkratnost $n-1$. Da bomo lahko dokazali trditev o lastnih vrednostih in determinanti Householderjeve matrike, pa najprej potrebujemo pomožno naslednjo lemo. [4]

Lema 3.17 *Če je A nesingularna $n \times n$ matrika in sta $x,\ y \in \mathbb{R}^n$, potem je* $\det\left(A + xy^T\right) = \left(1 + y^T A^{-1} x\right) \det A$.

Dokaz. Trditev bomo dokazali, če jo bomo dokazali za $A = I$, saj, ker je A nesingularna, velja:

$$\left(A + xy^T\right) = A\left(I + A^{-1} xy^T\right), \tag{3.4}$$

oziroma: $\det\left(A + xy^T\right) = \det A \cdot \det\left(I + A^{-1} xy^T\right)$. V tem primeru lahko zapišemo diagonalno matriko: $\begin{bmatrix} I + xy^T & 0 \\ 0 & 1 \end{bmatrix}$, katere determinanta je enaka $\det(I + xy^T)$. S pomočjo zgornje- oziroma spodnjetrikotnih matrik z determinanto 1, bomo preoblikovali dano matriko v matriko, katere determinanta bo želen izraz. Izračunajmo:

$$\begin{bmatrix} I + xy^T & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ -y^T & 1 \end{bmatrix} = \begin{bmatrix} I + xy^T & 0 \\ -y^T & 1 \end{bmatrix}$$

$$\begin{bmatrix} I & x \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} I + xy^T & 0 \\ -y^T & 1 \end{bmatrix} = \begin{bmatrix} I & x \\ -y^T & 1 \end{bmatrix}$$

$$\begin{bmatrix} I & 0 \\ y^T & 1 \end{bmatrix} \cdot \begin{bmatrix} I & x \\ -y^T & 1 \end{bmatrix} = \begin{bmatrix} I & x \\ 0 & 1 + y^T x \end{bmatrix}$$

$$\begin{bmatrix} I & 0 \\ y^T & 1 \end{bmatrix} \cdot \begin{bmatrix} I & x \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} I + xy^T & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ -y^T & 1 \end{bmatrix} = \begin{bmatrix} I & x \\ 0 & 1 + y^T x \end{bmatrix}$$

Z upoštevanjem enačbe 3.4 dobimo: $\det\left(A + xy^T\right) = \det A \cdot \det\left(I + \left(A^{-1}x\right) y^T\right) =$
$= \left(1 + y^T A^{-1} x\right) \cdot \det A.$ □

Zgornjo lemo lahko z upoštevanjem izreka 1.21 nekoliko preoblikujemo:

$$\det\left(A + xy^T\right) = \left(1 + y^T A^{-1} x\right) \det A = \left(1 + y^T \operatorname{adj}(A) \frac{1}{\det A} x\right) \det A =$$
$$= \det A + y^T \operatorname{adj}(A) x.$$

Trditev 3.18 *Lastne vrednosti Hauseholderjeve matrike so vedno:* $-1,\ \overbrace{1,\ 1,\ldots,1}^{n-1}$, *determinanta Hauseholderjeve matrike pa je enaka* -1.

Dokaz. Na podlagi izreka 2.14 vemo, da je lastna vrednost matrike ničla karakterističnega polinoma. Oglejmo si karakteristični polinom Householderjeve matrike.

$$\begin{aligned}
&p_{U_w}(\lambda) = \det(U_w - \lambda I) = \det\left(I - 2(w^*w)^{-1} ww^* - \lambda I\right) = \\
&= \det\left((1-\lambda) I - 2(w^*w)^{-1} ww^*\right) = \\
&= \det((1-\lambda) I) - 2(w^*w)^{-1} w^* \operatorname{adj}((1-\lambda) I) w = \\
&= (1-\lambda)^n - 2(w^*w)^{-1} w^* (1-\lambda)^{n-1} Iw = \\
&= (1-\lambda)^n - 2(w^*w)^{-1} (w^*w)(1-\lambda)^{n-1} = \\
&= (1-\lambda)^{n-1}(1-\lambda-2) = (1-\lambda)^{n-1}(-1-\lambda) = 0
\end{aligned}$$

Kjer smo pri izračunu uporabili 3.17 in 1.23. Razberemo, da ima matrika U_w lastne vrednosti 1 in -1, pri čemer je algebrajska večkratnost prve enaka $n-1$, druge pa 1.
Dokazati moramo še, da je determinanta Householderjeve matrike enaka -1. Tudi v tem primeru si bomo pomagali z lemo 3.17.

$$\begin{aligned}
&\det U_w = \det\left(I - 2(w^*w)^{-1} ww^*\right) = \left(1 - 2(w^*w)^{-1} w^* I^{-1} w\right) \det I = \\
&= \left(1 - 2(w^*w)^{-1} (w^*w)\right) = 1 - 2 = -1.
\end{aligned}$$

□

Za tem, ko smo spoznali nekaj osnovnih lastnosti Householderjevih matrik, bomo naslednjim izrekom opredelili tudi njihovo uporabnost – Householderjeve matrike lahko uporabimo za konstrukcijo unitarnih matrik, če imamo podana dva neničelna vektorja v prostoru $\mathbb{C}^n$. Ob tem je potrebno upoštevati zgolj to, da imata enako dolžino, kar lahko dosežemo tako, da dana vektorja normiramo.

Izrek 3.19 *Naj bosta dana* $x,\ y \in \mathbb{C}^n$, *da zanju velja* $\|x\| = \|y\| > 0$. *Ločimo dva primera:*

(i) če je $y = e^{i\phi}x$, *za nek* $\phi \in \mathbb{R}$, *potem naj bo* $U(y,x) = e^{i\phi} I_n$,

(ii) naj bo $\theta \in [0, 2\pi)$ *takšen, da velja* $\langle y, x\rangle = e^{i\theta} |\langle y, x\rangle|$ *(v primeru da sta vektorja* x *in* y *ortogonalna (velja:* $x^*y = 0$*), izberemo* $\theta = 0$*) in definiramo* $w = e^{i\theta}x - y$ *ter* $U(y,x) = e^{i\theta} U_w$, *kjer je* U_w *Householderjeva matrika.*

Potem je matrika $U(y,x)$ *unitarna in esencialno Hermitska ter velja:* $U(y,x)\, x = y$ *in če je* $z \perp x$, *potem je* $U(y,x) z \perp y$.

Če sta $x,\ y \in \mathbb{R}^n$, tedaj je $U(y,x)$ realna ortogonalna matrika. V primeru $y = x$ je realna ortogonalna matrika oblike: $U(y,x) = I$, sicer pa oblike U_{x-y}.

Dokaz. V primeru (i), ko sta vektorja x in y linearno odvisna velja, da je matrika $e^{i\phi}I_n$ unitarna (velja: $\left(e^{i\phi}I_n\right)^* = e^{-i\phi}I_n$) in esencialno Hermitska, saj, če izberemo $\theta = -\phi$, dobimo identično matriko reda n, ki je Hermitska.
Izračunajmo: $U(y,x)\,x = e^{i\phi}I_n x = e^{i\phi}x = y$. Poleg tega velja: če je $z \perp x \Leftrightarrow \langle z,x\rangle = x^*z = 0$, potem je $\langle e^{i\phi}I_n z, y\rangle = y^*e^{i\phi}z = e^{i\phi}\left(e^{i\phi}x\right)^* z = e^{i\phi}e^{-i\phi}x^*z = x^*z = 0 \Leftrightarrow U(y,x)\,z \perp x$.

Oglejmo si še drugo točko izreka. Ker je matrika U_w po trditvi 3.14 unitarna in Hermitska in ker smo ugotovili, da je tudi matrika $e^{i\theta}I_n$ unitarna in Hermitska, velja, da je tudi matrika $U(y,x) = e^{i\theta}I_n U_w = e^{i\theta}I_n U_w$ unitarna, saj po posledici 3.5 velja, da je produkt unitarnih matrik unitarna matrika. Hkrati pa je matrika $U(y,x)$ tudi esencialno Hermitska, saj lahko izberemo tako vrednost $\Theta \in \mathbb{R}$, da bo matrika $e^{-i\Theta}U(y,x)$ Hermitska. Preverimo še ostale rezultate iz izreka.

Vemo, da sta x in y linearno neodvisna vektorja, zato po neenakosti Cauchy-Schwarz-Bunjakovski (glej 1.44) velja, da je $|\langle y,x\rangle| \neq \|x\| \cdot \|y\| = \|x\| \cdot \|x\| = \|x\|^2$. Torej: $|\langle y,x\rangle| \neq \|x\|^2$. Izračunajmo:

$$\begin{aligned}
&\langle w,w\rangle = \left(e^{i\theta}x - y\right)^* \left(e^{i\theta}x - y\right) = \left(e^{-i\theta}x^* - y^*\right)\left(e^{i\theta}x - y\right) = \\
&= x^*x - e^{-i\theta}x^*y - e^{i\theta}y^*x + y^*y = x^*x - e^{-i\theta}\langle y,x\rangle - \overline{e^{-i\theta}\langle y,x\rangle} + y^*y = \\
&= \|x\|^2 + \|y\|^2 - 2\,\mathrm{Re}\left(e^{-i\theta}\langle y,x\rangle\right) = 2\left(\|x\|^2 - \mathrm{Re}\,|\langle y,x\rangle|\right) = 2\left(\|x\|^2 - |\langle y,x\rangle|\right).
\end{aligned}$$

$$\begin{aligned}
&\langle x,w\rangle = \left(e^{i\theta}x - y\right)^* x = \left(e^{-i\theta}x^* - y^*\right)x = e^{-i\theta}\|x\|^2 - \langle x,y\rangle = \\
&= e^{-i\theta}\|x\|^2 - \overline{\langle y,x\rangle} = e^{-i\theta}\|x\|^2 - \overline{e^{i\theta}\,|\langle y,x\rangle|} = e^{-i\theta}\left(\|x\|^2 - |\langle y,x\rangle|\right).
\end{aligned}$$

Preverimo, da je $U(y,x)x = y$:

$$\begin{aligned}
&U(y,x)\,x = e^{i\theta}U_w x = e^{i\theta}\left(I - 2\left(\langle w,w\rangle\right)^{-1} ww^*\right)x = \\
&= e^{i\theta}x - 2e^{i\theta}\left(\langle w,w\rangle\right)^{-1} w\langle x,w\rangle = \\
&= e^{i\theta}x - e^{i\theta}\left(\|x\|^2 - |\langle y,x\rangle|\right)^{-1} we^{-i\theta}\left(\|x\|^2 - |\langle y,x\rangle|\right) = \\
&= e^{i\theta}x - w = e^{i\theta}x - \left(e^{i\theta}x - y\right) = y.
\end{aligned}$$

Če je $z \perp x$, potem je $\langle z,w\rangle = \left(e^{i\theta}x - y\right)^* z = e^{-i\theta}x^*z - y^*z = -\langle z,y\rangle$. Preverimo, da velja:

$U(y,x)\, z \perp y$:

$$\begin{aligned}
&\langle U(y,x)\, z, y\rangle = y^* e^{i\theta} U_w z = y^* e^{i\theta}\left(I - 2\langle w,w\rangle^{-1} ww^*\right) z = \\
&= e^{i\theta}\langle z,y\rangle - 2e^{i\theta} y^* \langle w,w\rangle^{-1} w\langle z,w\rangle = e^{i\theta}\langle z,y\rangle + 2e^{i\theta}\langle w,w\rangle^{-1} y^*\left(e^{i\theta}x - y\right)\langle z,y\rangle = \\
&= e^{i\theta}\langle z,y\rangle + 2e^{i\theta}\langle w,w\rangle^{-1}\left(e^{i\theta}y^*x - y^*y\right)\langle z,y\rangle = \\
&= e^{i\theta}\langle z,y\rangle + 2e^{i\theta}\langle w,w\rangle^{-1}\left(e^{i\theta}\langle x,y\rangle - \|y\|^2\right)\langle z,y\rangle = \\
&= e^{i\theta}\langle z,y\rangle + 2e^{i\theta}\langle w,w\rangle^{-1}\left(|\langle x,y\rangle| - \|y\|^2\right)\langle z,y\rangle = \\
&= e^{i\theta}\langle z,y\rangle - 2e^{i\theta}\langle w,w\rangle^{-1}\left(\|x\|^2 - |\langle y,x\rangle|\right)\langle z,y\rangle = \\
&= e^{i\theta}\langle z,y\rangle - e^{i\theta}\langle w,w\rangle^{-1}\langle w,w\rangle\langle z,y\rangle = e^{i\theta}\langle z,y\rangle - e^{i\theta}\langle z,y\rangle = 0.
\end{aligned}$$

S tem smo dokaz zaključili. □

Posledica 3.20 *Naj bo $y \in \mathbb{C}^n$ poljuben in naj bo vektor $e_1 \in \mathbb{C}^n$ enotski vektor. Prvi stolpec unitarne matrike $U(y,e_1)$ je enak vektorju y.*

Dokaz. Prejšnji izrek jamči obstoj unitarne matrike $U(y,e_1)$, da zanjo velja: $U(y,e_1)e_1 = y$. Torej je prvi stolpec matrike $U(y,e_1)$ res enak vektorju y. □

Oglejmo si še uporabo zgornjega izreka na naslednjem zgledu.

Zgled. Naj bosta dana vektorja $x = \left[\frac{5}{13},\ \frac{12}{13}\right]^T$ in $y = \left[\frac{3}{5},\ \frac{4}{5}\right]^T$. Konstruirajmo unitarno matriko, da bo $U(y,x)x = y$.
Opazimo, da velja $\|x\|^2 = \|y\|^2 = 1 > 0$. Ker vektorja x in y nista linearno odvisna, velja, da je matrika $U(y,x) = U_{x-y}$. Z nekaj računske spretnosti dobimo:

$$U_{x-y} = \begin{bmatrix} -\frac{33}{65} & \frac{56}{65} \\ \frac{56}{65} & \frac{33}{65} \end{bmatrix}.$$

Opazimo, da je dobljena matrika res realna ortogonalna matrika, še več je tudi simetrična, in zanjo velja, da je $U_{x-y}x = y$.

3.3 QR faktorizacija matrik

V zadnjem razdelku tega poglavja bomo predstavili QR faktorizacijo matrike. V dokazu si bomo pomagali z glavnim orodjem, ki smo ga razvili v prejšnjem podrazdelku, torej z izrekom, ki podaja konstrukcijo Householderjeve matrike.

Izrek 3.21 *Naj bo dana matrika $A \in M_{n,m}$. Potem velja:*

(a) Če je $n \geq m$, potem obstaja matrika $Q \in M_{n,m}$ z ortonormiranimi stolpci in zgornjetrikotna matrika $R \in M_m$ z nenegativnimi elementi na glavni diagonali, da velja: $A = QR$.

(b) Če je $\operatorname{rang} A = m$, potem sta matriki Q in R iz točke (a) enolično določeni in vsi elementi na glavni diagonali matrike R so neničelni.

(c) Če je $m = n$, potem je matrika Q iz točke (a) unitarna.

(d) Obstajata unitarna matrika $Q \in M_n$ in matrika $R \in M_{n,m}$, ki ima neničelno diagonalo in za katero je $(R)_{ij} = 0$, ko je $i > j$, da zanju velja $A = QR$.

(e) Če je A matrika z realnimi elementi, potem so vse matrike iz točk (a), (b), (c) in (d) realne.

Dokaz. (a) Naj bo $A \in M_{n,m}$ poljubna matrika, za katero je $m \leq n$, in naj bo $a_1 \in \mathbb{C}^n$ prvi stolpec te matrike. Definirajmo $r_1 := \|a_1\|$. Naj bo U_1 unitarna matrika, za katero velja $U_1 a_1 = r_1 e_1$, kjer e_1 predstavlja prvi stolpec identične matrike. Po izreku 3.19 taka unitarna matrika obstaja, gre za bodisi matriko oblike $e^{i\theta} I_n$ bodisi oblike $e^{i\theta} U_w$, kjer je U_w Householderjeva matrika.

Če matriko U_1 pomnožimo z matriko A dobimo: $U_1 A = \begin{bmatrix} r_1 & \diamondsuit \\ 0 & A_2 \end{bmatrix}$, kjer je $A_2 \in M_{n-1}$.

Postopek ponovimo: naj bo $a_2 \in \mathbb{C}^{n-1}$ prvi stolpec matrike A_2. Definiramo $r_2 := \|a_2\|$. Izrek 3.19 znova zagotavlja obstoj unitarne matrike, ki jo označimo z $V_2 \in M_{n-1}$, da zanjo velja $V_2 a_2 = r_2 e_1$, kjer je $e_1 \in \mathbb{C}^{n-1}$ enotski vektor. Definirajmo matriko $U_2 := I_1 \oplus V_2$. Potem je $U_2 U_1 A = \begin{bmatrix} r_1 & \diamondsuit & \diamondsuit \\ 0 & r_2 & \diamondsuit \\ 0 & 0 & A_3 \end{bmatrix}$, kjer je $A_3 \in M_{n-2}$. Opisan postopek ponavljamo in po m korakih dobimo naslednjo matriko:

$$U_m U_{m-1} \cdots U_1 A = \begin{bmatrix} r_1 & & & \diamondsuit \\ & r_2 & & \\ & & \ddots & \\ 0 & 0 & & r_m \\ & & \vdots & \\ & & 0 & \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix},$$

kjer je R zgornjetrikotna matrika. Za matriko R velja, da so njeni elementi na diagonali pozitivna števila oziroma število 0, če je bil prvi stolpec matrike A_i ničelni vektor. To dejstvo sledi iz definicije norme.

Naj bo $U := U_m U_{m-1} \cdots U_1$. Ker je matrika U produkt unitarnih matrik po posledici 3.5 velja, da je tudi matrika U unitarna. Zato velja, da je $UU^* = I$, matrika $U^* = U_1^* \cdots U_m^*$ je dimenzije $n \times n$ in ima (po izreku 3.4) ortonormirane stolpce. Definirajmo particijo matrike $U^* := [Q \;\; Q_2]$, kjer sta matriki $Q \in M_{n,m}$ in $Q_2 \in M_{n,n-m}$. Ker je imela matrika U^* ortonormirane stolpce, jih ima tudi matrika Q, saj ti stolpci predstavljajo prvih m stolpcev matrike U^*. Potem je:

$$\begin{bmatrix} Q & Q_2 \end{bmatrix} \cdot \begin{bmatrix} R \\ 0 \end{bmatrix} = \begin{bmatrix} QR + 0 \end{bmatrix} = A.$$

S tem je dokazana točka (a) izreka, saj ima matrika Q ortonormirane stolpce, matrika R pa je zgornjetrikotna matrika. Za matriki velja, da je $QR = A$.

(b) Predpostavimo, da je rang $A = m$. Potem velja, da ima matrika A m linearno neodvisnih vrstic. Iz točke (a) vemo, da lahko najdemo matriki Q in R, ki zadoščata predpisanim pogojem. Denimo, da lahko najdemo še neki dve matriki $\tilde{Q}$ in $\tilde{R}$, da zanju velja: $\tilde{Q}\tilde{R} = A$, in da je $\tilde{Q}$ matrika z ortonormiranimi stolpci, matrika $\tilde{R}$ pa zgornjetrikotna matrika.
Oglejmo si konjugirano-transponiranko matrike A: $A^* = (QR)^* = R^*Q^*$. Potem je $A^*A = R^* (Q^*Q) R = R^*R$, saj imata matriki Q in Q^* ortonormirane stolpce, torej sta matriki unitarni. Podobno dobimo tudi naslednjo enakost: $A^*A = \tilde{R}^*\tilde{Q}^*\tilde{Q}\tilde{R} = \tilde{R}^*\tilde{R}$. Iz teh enakosti sledi:

$$R^*R = \tilde{R}^*\tilde{R}. \tag{3.5}$$

Ker velja, da je rang $A = m$, matrika A ni imela ničelnega stolpca. Zato so, glede na konstrukcijo dokaza v točki (a), elementi na diagonalah matrik R in $\tilde{R} \in M_m$ neničelni. Ker sta obe matriki zgornjetrikotni, je njuna determinanta različna od 0, kar pomeni, da sta matriki nesingularni. Torej obstajata matriki R^{-1} in $\tilde{R}^{-1}$.
Iz enakosti 3.5 dobimo: $\left(\tilde{R}^*\right)^{-1} R^* = \tilde{R}R^{-1}$. Vidimo, da na levi strani enakosti dobimo spodnjetrikotno matriko, na desni pa zgornjetrikotno matriko. Matriki sta enaki zgolj takrat, ko sta obe diagonalni. Označimo z $D := \tilde{R}R^{-1}$ diagonalno matriko, ki ima vse svoje diagonalne elemente pozitivne.
Torej je: $\tilde{R} = DR$, hkrati pa velja:
$D = \tilde{R}R^{-1} = \left(\tilde{R}^*\right)^{-1} R^* = ((DR)^*)^{-1} R^* = (R^*D^*)^{-1} R^* = (D^*)^{-1} (R^*)^{-1} R^* = (D^*)^{-1}$.
Vidimo, da je $D = (D^*)^{-1} \Leftrightarrow D^*D = I \Leftrightarrow D = I$, in zato velja $\tilde{R} = R$ in posledično $\tilde{Q} = Q$.

(c) Če je $m = n$ je po konstrukciji v koraku (a) matrika $U^* = Q \in M_n$. Matrika U^* je unitarna, torej je unitarna tudi matrika Q, saj sta matriki enaki.

(d) Ločimo dve možnosti: če je $n \geq m$, potem smo vse potrebno že dokazali. Iz točke (a) vemo, da je matrika U^* unitarna. Naj bo iskana matrika $\tilde{Q} \in M_n$ enaka unitarni matriki

$U^* = \tilde{Q} = [Q \;\; Q_2]$ in naj bo matrika $\tilde{R} \in M_{n,m}$ definirana takole: $\tilde{R} = \begin{bmatrix} R \\ 0 \end{bmatrix}$. Potem je $A = \tilde{Q}\tilde{R} = QR$.
Preostane še, druga možnost, in sicer, če je $n < m$. V tem primeru lahko uberemo podobno strategijo, kot smo jo pri dokazu točke (a). Torej matriko A množimo z leve z zaporedjem unitarnih matrik $U_1,\ U_2,\ \ldots,\ U_n$, katerih obstoj nam jamči izrek 3.19. Postopek se bo ustavil po n korakih. Takrat bomo dobili naslednjo matriko: $U_n \cdots U_2 U_1 A = [R \;\; \diamondsuit]$, kjer $\diamondsuit$ predstavlja $m - n$ stolpcev produkta unitarne matrike $U := U_n \cdots U_1$ z matriko A.
Dobili smo torej unitarno matriko $Q := U^*$ dimenzije $n \times n$ in matriko R, za katero velja, da ima na diagonali nenegativne vrednosti, elementi $(R)_{ij}$, ki izpolnjujejo pogoj $i > j$, pa so enaki 0.

(e) Dokaz sledi neposredno iz izreka 3.19. □

Posledica 3.22 *Matriko $B \in M_n$, ki je oblike $B = A^*A$, $A \in M_n$, lahko zapišemo kot $B = LL^*$, kjer je $L \in M_n$ spodnjetrikotna matrika z nenegativnimi elementi na diagonali. Če je matrika A nesingularna, potem je tak zapis enoličen.*

Dokaz. Po izreku 3.21 velja, da obstajata taki matriki Q in $R \in M_n$, kjer je matrika R zgornjetrikotna matrika, matrika Q pa unitarna matrika, da zanju velja: $A = QR$. Potem je $B = A^*A = (QR)^* QR = R^*Q^*QR = R^*R$. Opazimo, da je matrika R^* spodnjetrikotna matrika. Naj bo $L := R^*$. Potem je $B = LL^*$, kar smo želeli pokazati.
Drugi del posledice sledi neposredno iz točke (b) izreka 3.21 in iz dejstva, da ima nesingularna matrika rang enak številu vrstic oziroma stolpcev. □

Poglavje 4

Unitarna podobnost matrik

Definicija 4.1 *Naj bosta $A,\ B \in M_n$ kvadratni matriki s kompleksnimi elementi. Pravimo, da je matrika A unitarno podobna matriki B, če obstaja taka unitarna matrika $U \in M_n$, da velja $A = UBU^*$. Če je U realna ortogonalna matrika, potem pravimo, da je A realno-ortogonalno podobna matriki B.*
Matrika A je unitarno diagonalizabilna, če je unitarno podobna diagonalni matriki, podobno je A realno-ortogonalno diagonalizabilna, če je realno-ortogonalno podobna diagonalni matriki.

Trditev 4.2 *Unitarna podobnost matrik je ekvivalenčna relacija.*

Dokaz. Naj bodo $A,\ B,\ C \in M_n$. Preveriti je potrebno refleksivnost, simetričnost in tranzitivnost:

(i) refleksivnost: obstaja matrika I, ki je unitarna in velja: $A = IAI^*$.

(ii) simetričnost: naj bo U unitarna matrika, da velja $B = UAU^*$. Potem je $U^*BU = A$. Naj bo $T := U^*$: $A = TBT^*$.

(iii) tranzitivnost: naj bosta U in T unitarni matriki, da zanju velja: $A = UBU^*$ in $B = TCT^*$. Potem je $A = UBU^* = UTCT^*U^* = UTC\,(UT)^*$. Vemo, da je produkt unitarnih matrik spet unitarna matrika (glej: 3.5), zato naj bo $V := UT$. Potem je $A = VCV^*$. □

Naslednji izrek bo podal potreben pogoj, da sta dve matriki unitarno podobni. Za dokaz izreka bomo potrebovali spodnjo trditev.

Trditev 4.3 *Za vsako matriko $A \in M_{n,m}$ velja:* $\operatorname{sled} AA^* = \operatorname{sled} A^*A = \sum_{i,j=1}^{n,m} |a_{ij}|^2$.

Dokaz. Naj bo $A \in M_{n,m}$. Potem je:

$$\operatorname{sled} AA^* = \sum_{i=1}^{n} \left(\sum_{k=1}^{m} a_{ik}\overline{a_{ik}} \right) = \sum_{k=1}^{m} \left(\sum_{i=1}^{n} \overline{a_{ik}} a_{ik} \right) = \operatorname{sled} A^*A.$$

Pri tem smo pri drugem enačaju uporabili komutativnost množenja kompleksnih števil. Velja še:

$$\operatorname{sled} AA^* = \sum_{i=1}^{n} \left(\sum_{k=1}^{m} a_{ik}\overline{a_{ik}} \right) = \sum_{i=1}^{n} \sum_{k=1}^{m} |a_{ik}|^2 .$$

□

Izrek 4.4 *Naj bosta dani matriki $A, B \in M_n$. Če je matrika A unitarno podobna matriki B, potem velja $\sum_{i,j=1}^{n} |b_{ij}|^2 = \sum_{i,j=1}^{n} |a_{ij}|^2$.*

Dokaz. Ker sta matriki A in B unitarno podobni obstaja taka unitarna matrika U, da je $B = UAU^*$. Da bo pogoj iz izreka izpolnjen, po prejšnji trditvi zadošča preveriti, da je $\operatorname{sled}(B^*B) = \operatorname{sled}(A^*A)$.

Izračunamo:

$$\operatorname{sled} A^*A = \operatorname{sled}((UBU^*)^* UBU^*) = \operatorname{sled}(UB^*U^*UBU^*) = \operatorname{sled}(U^*B^*BU) =$$
$$= \operatorname{sled}(B^*BUU^*) = \operatorname{sled} B^*B,$$

kjer smo upoštevali cikličnost sledi (glej posledico 1.27). □

Zgled. Pokažimo, da sta matriki: $A = \begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix}$ in $B = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ podobni, ampak nista unitarno podobni.
Da bosta matriki podobni je potrebno najti nesingularno matriko S, da zanjo velja: $B = S^{-1}AS$. Primer take matrike je matrika $S := \begin{bmatrix} 4 & -1 \\ -8 & 5 \end{bmatrix}$.
Matriki A in B pa nista unitarno podobni, saj je: $\sum_{i,j=1}^{2} |a_{ij}|^2 = 14$ in $\sum_{i,j=1}^{2} |b_{ij}|^2 = 6$. Matriki torej ne izpolnjujeta potrebnega pogoja za unitarno podobnost.

Opazimo, da je unitarna podobnost matrik močnejši pojem od podobnosti. Izkaže se, da izrek 4.4 ni zadosten pogoj za to, da bi bili dve matriki unitarno podobni. S tem problemom so se matematiki ukvarjali v zadnjem stoletju. Prvi rezultat, ki je vsaj teoretično zagotovil zadostni pogoj za unitarno podobnost matrik, je objavil nemški matematik Wilhelm Specht (1907–1985). Preden bomo navedli Spechtov izrek potrebujemo še nekaj predpriprave.

Definicija 4.5 *Naj bosta s, t nekomutativni spremenljivki. Poljuben končen formalen produkt potenc spremenljivk s, t z nenegativnimi potenčnimi eksponenti oblike:*

$$W(s,t) = s^{m_1}t^{n_1}s^{m_2}t^{n_2}\cdots s^{m_k}t^{n_k},\ \textit{kjer je}\ \ m_1, n_1, m_2, n_2, \ldots, m_k, n_k \geq 0$$

imenujemo beseda za spremenljivki s in t. Dolžina besede $W(s,t)$ je nenegativno celo število, ki ga dobimo kot vsoto potenčnih eksponentov spremenljivk s in t v besedi $W(s,t)$:

$$m_1 + n_1 + m_2 + n_2 + \ldots + m_k + n_k.$$

Podobno lahko definiramo besede za matrike. Naj bo $A \in M_n$, definirajmo besedo $W(A, A^*)$ kot: $W(A, A^*) = A^{m_1}(A^*)^{n_1}\cdots A^{m_k}(A^*)^{n_k}$. Ker potence matrik A in A^* nujno ne komutirajo, besede ni mogoče skrajšati z zamenjavo mest faktorjev v zgornjem produktu.

Trditev 4.6 *Naj bo $A \in M_n$ unitarno podobna matriki $B \in M_n$. Potem za vsako besedo $W(s,t)$ velja: $W(A, A^*) = UW(B, B^*)U^*$.*

Dokaz. Ker je matrika A unitarno podobna matriki B, obstaja taka unitarna matrika U, da velja: $A = UBU^*$. Naj bo $W(s,t)$ poljubna beseda. Potem je:

$$\begin{aligned} &W(A, A^*) = A^{m_1}(A^*)^{n_1}\cdots A^{m_k}(A^*)^{n_k} = \\ &= (UBU^*)^{m_1}(UB^*U^*)^{n_1}\cdots(UBU^*)^{m_k}(UB^*U^*)^{n_k} = \\ &= UB^{m_1}(B^*)^{n_1}\cdots B^{m_k}(B^*)^{n_k}U^* = UW(B, B^*)U^*. \end{aligned}$$

□

Ugotovili smo, da če sta matriki A in B unitarno podobni, potem sta matriki $W(A, A^*)$ in $W(B, B^*)$ prav tako unitarno podobni. Ker velja, da imajo podobne matrike enako sled, izrek 2.22, unitarna podobnost pa je zgolj bolj strog pojem podobnosti, za matriki velja: $\operatorname{sled} W(A, A^*) = \operatorname{sled} W(B, B^*)$. Če izberemo besedo $W(s,t) = ts$ dobimo ravno enakost, ki smo jo uporabili pri dokazu izreka 4.4.

Izkaže se, da je pogoj preverjanja enakosti sledi besed $W(A, A^*)$ in $W(B, B^*)$ zadosten pogoj za to, da sta matriki A in B unitarno podobni. To zagotavlja Spechtov izrek, ki ga na tem mestu zgolj navajamo, implikacija v desno tega izreka je bila dokazana v prejšnji trditvi, implikacija v levo pa presega okvir te naloge.

Izrek 4.7 (Specht, 1940) *Matriki A in $B \in M_n$ sta unitarno podobni natanko tedaj, ko je:*

$$\operatorname{sled} W(A, A^*) = \operatorname{sled} W(B, B^*), \tag{4.1}$$

za vsako besedo $W(s,t)$, kjer sta s in t nekomutativni spremenljivki.

Spechtov izrek je uporaben pri dokazovanju, da matriki nista unitarno podobni, pri dokazovanju, da sta matriki unitarno podobni pa se običajno zatakne, saj izrek ne podaja zgornje meje dolžine besed, ki jih moramo preveriti, da bosta dve matriki unitarno podobni. Matematiki so se s tem problemom aktivno ukvarjali [5], eno izmed prvih zgornjih mej je dokazal Percy:

Izrek 4.8 (Pearcy, 1962) *Da bosta $n \times n$ matriki s kompleksnimi elementi unitarno podobni, je potrebno preveriti pogoj 4.1 za besede $W(s,t)$ za spremenljivki s in t dolžine manjše ali enake $2n^2$.*

Za tem so zgornjo mejo še znižali. To nas pripelje do naslednjega izreka, ki ne podaja zgolj zgornje meje za število besed dveh nekomutativnih spremenljivk, ki jih moramo preveriti, da bi bili matriki unitarno podobni, temveč za majhne matrike podaja tudi sezname besed, ki jih moramo preveriti, da bosta matriki unitarno podobni.

Izrek 4.9 *Naj bosta $A,\ B \in M_n$.*

(a) Matriki A in B sta unitarno podobni natanko tedaj, ko zadoščata pogoju 4.1 za vsako besedo $W(s,t)$ za nekomutativni spremenljivki s in t dolžine največ

$$n \cdot \sqrt{\frac{2n^2}{n-1} + \frac{1}{4}} + \frac{n}{2} - 2.$$

(b) Če je $n = 2$, sta A in B unitarno podobni natanko tedaj, ko zadoščata pogoju 4.1 za naslednje tri besede: $W(s,t) = s,\ s^2$ in st.

(c) Če je $n = 3$, sta A in B unitarno podobni natanko tedaj, ko zadoščata pogoju 4.1 za naslednjih sedem besed: $W(s,t) = s,\ s^2,\ st,\ s^3,\ s^2t,\ s^2t^2 in\ s^2t^2st$.

(d) Če je $n = 4$, sta A in B unitarno podobni natanko tedaj, ko zadoščata pogoju 4.1 za naslednjih dvajset besed, ki so podane v spodnji tabeli (razvrstitev po dolžinah besed):

Dolžina besede	*Beseda*	*Dolžina besede*	*Beseda*
1	s	2	$s^2,\ st$
3	$s^3,\ s^2t$	4	$s^4,\ s^3t,\ s^2t^2,\ stst$
5	s^3t^2	6	$s^2ts^2t,\ s^2t^2st,\ t^2s^2ts$
7	s^3t^2st	8	$s^3t^2s^2t,\ s^3t^3st,\ t^3s^3ts$
9	$s^3ts^2tst,\ s^2t^2sts^2t$	10	$s^3t^3s^2t^2$

Izreka ne bomo dokazovali, saj bi to presegalo okvir te naloge. Vseeno pa si oglejmo njegovo uporabo. Obravnavali bomo unitarno podobnost matrik in njihovih transponiranih matrik. Najprej si oglejmo, kako je z unitarno podobnostjo matrik in njihovih transponirank dimenzije 2×2. Po Spechtovem izreku mora veljati: $\operatorname{sled} W\left(A^T, \left(A^T\right)^*\right) = \operatorname{sled} W\left(A^T, \overline{A}\right) = \operatorname{sled} W\left(A, A^*\right)$. Da bosta matriki A in A^T unitarno podobni moramo po točki (b) izreka 4.9 preveriti naslednje besede $W(s,t) = s$, s^2 in st:

$$\begin{aligned}
&\operatorname{sled} A^T = \operatorname{sled} A,\\
&\operatorname{sled}\left(A^T\right)^2 = \operatorname{sled}\left(A^T A^T\right) = \operatorname{sled}\left(AA\right)^T = \operatorname{sled}\left(A^2\right)^T = \operatorname{sled} A^2,\\
&\operatorname{sled}\left(A^T\overline{A}\right) = \operatorname{sled}\left(A^T\left(A^*\right)^T\right) = \operatorname{sled}\left(A^*A\right)^T = \operatorname{sled}\left(A^*A\right) = \operatorname{sled}\left(AA^*\right).
\end{aligned}$$

Preverili smo vse zahtevane pogoje, zato za poljubno 2×2 matriko velja, da je unitarno podobna svoji transponirani matriki.

Zgled. Poiščimo ortogonalno matriko O, da bosta matrika $A = \begin{bmatrix} 2 & -2 \\ -3 & 2 \end{bmatrix}$ in njena transponiranka ortogonalno podobni.

Preverimo, da je matrika $O = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ ustrezna:

$$OAO^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & -2 \\ -3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ -2 & 2 \end{bmatrix}.$$

Oglejmo si še, kaj se dogaja pri matrikah dimenzije 3×3. Točka (c) izreka 4.9 pove, da je potrebno preveriti sedem besed. Opazimo, da smo tri izmed teh preverili že zgoraj, saj pri preverjanju besed za matrike 2×2 nismo nikoli uporabili podatka o dimenziji matrik. Zato preostane da preverimo še ostale štiri besede, in sicer $W(s,t) = s^3$, s^2t, s^2t^2 in s^2t^2st:

$$\begin{aligned}
&\operatorname{sled}\left(A^T\right)^3 = \operatorname{sled}\left(A^3\right)^T = \operatorname{sled} A^3,\\
&\operatorname{sled}\left(\left(A^T\right)^2\overline{A}\right) = \operatorname{sled}\left(\left(A^2\right)^T\left(A^*\right)^T\right) = \operatorname{sled}\left(A^*A^2\right)^T = \operatorname{sled}\left(A^2A^*\right),\\
&\operatorname{sled}\left(\left(A^T\right)^2\left(\overline{A}\right)^2\right) = \operatorname{sled}\left(\left(A^2\right)^T\left(\left(A^*\right)^2\right)^T\right) = \operatorname{sled}\left(\left(A^*\right)^2A^2\right)^T =\\
&= \operatorname{sled}\left(\left(A^*\right)^2A^2\right) = \operatorname{sled}\left(A^2\left(A^*\right)^2\right).
\end{aligned}$$

Pri preverjanju zadnjega pogoja pa se nekoliko zatakne. Preveriti je potrebno, da je $\operatorname{sled} W\left(A^T, \overline{A}\right) = \operatorname{sled}\left(A, A^*\right)$, za besedi $\left(A^T\right)^2\overline{A}^2A^T\overline{A}$ oziroma $A^2\left(A^*\right)^2AA^*$. Najprej

si oglejmo levo stran enakosti:

$$\operatorname{sled}\left(\left(A^T\right)^2\overline{A}^2A^T\overline{A}\right)=\operatorname{sled}\left(\left(A^2\right)^T\left(\left(A^*\right)^2\right)^TA^T\left(A^*\right)^T\right)=$$
$$=\operatorname{sled}\left(A^*A\left(A^*\right)^2A^2\right)^T=\operatorname{sled}\left(A^*A\left(A^*\right)^2A^2\right).$$

Na desni strani dobimo: $\operatorname{sled}\left(A^2\left(A^*\right)^2AA^*\right)$. Dobljeni sledi morata biti enaki, torej:

$$\begin{aligned}\operatorname{sled}\left(A^2\left(A^*\right)^2AA^*\right)&=\operatorname{sled}\left(A^*A\left(A^*\right)^2A^2\right)\\ \operatorname{sled}\left(A^2\left(A^*\right)^2AA^*\right)-\operatorname{sled}\left(A^*A\left(A^*\right)^2A^2\right)&=0\\ \operatorname{sled}\left(A^2\left(A^*\right)^2AA^*-A^*A\left(A^*\right)^2A^2\right)&=0\\ \operatorname{sled}\left(A\left(A^*\right)^2AA^*A-AA^*A\left(A^*\right)^2A\right)&=0\end{aligned}$$

Od tod dobimo naslednjo zvezo:

$$\operatorname{sled}\left(AA^*\left(A^*A-AA^*\right)A^*A\right)=0. \tag{4.2}$$

Če sklenemo, smo ugotovili, da je vsaka matrika dimenzije 2×2 unitarno podobna svoji transponirani matriki. Po drugi strani pa za 3×3 matrike velja, da morajo izpolnjevati pogoj 4.2.

Zgled. Preverimo, da matrika $A=\begin{bmatrix}1&1&1\\-1&0&1\\-1&-1&-1\end{bmatrix}$ ni unitarno podobna svoji transponirani matriki.
Najprej izračunajmo produkta AA^T in A^TA:

$$AA^T=\begin{bmatrix}3&0&-3\\0&2&0\\-3&0&3\end{bmatrix}\qquad A^TA=\begin{bmatrix}3&2&1\\2&2&2\\1&2&3\end{bmatrix}$$

Izračunamo:

$$\operatorname{sled}\left(AA^T\left(A^TA-AA^T\right)A^TA\right)=\operatorname{sled}\begin{bmatrix}-24&0&24\\16&16&16\\24&0&-24\end{bmatrix}=-32\neq 0.$$

S tem smo pokazali, da matrika A ni unitarno podobna svoji transponirani matriki.

Poglavje 5

Schurov izrek in njegove posledice

V tem poglavju bomo najprej predstavili Schurov izrek, ki pravi, da za vsako kvadratno kompleksno matriko A obstaja taka unitarna matrika U, da je produkt matrik U^*AU zgornjetrikotna matrika. V nadaljevanju bo dokazana še realna Schurova forma. V drugem delu poglavja bo predstavljenih nekaj primerov uporabe Schurovega izreka.

5.1 Schurov izrek

Izrek 5.1 (Schurov izrek, Schurova triangulacija matrike) *Naj bodo $\lambda_1, \ldots, \lambda_n$ lastne vrednosti matrike $A \in M_n$, ki so podane v nekem predpisanem vrstnem redu in naj bo $x \in \mathbb{C}^n$ tak enotski vektor, da velja $Ax = \lambda_1 x$. Potem veljata naslednji trditvi:*

(a) Obstaja takšna unitarna matrika $U = [x\ u_2\ \ldots\ u_n]$, da je $U^*AU = T$, kjer je T zgornjetrikotna matrika z diagonalnimi elementi $t_{ii} = \lambda_i$ za vsak $i \in \{1, \ldots, n\}$.

(b) Če ima matrika $A \in M_n(\mathbb{R})$ samo realne lastne vrednosti, potem je vektor $x \in \mathbb{R}^n$ in obstaja taka realna ortogonalna matrika $Q = [x\ q_2\ \ldots\ q_n]$, da je matrika $Q^TAQ = T$ zgornjetrikotna matrika, ki ima na glavni diagonali lastne vrednosti matrike A, da velja $t_{ii} = \lambda_i$ za vsak $i = 1, \ldots, n$.

Dokaz. Naj bo $x_1 \in \mathbb{C}^n$ normiran lastni vektor pripadajoč lastni vrednosti λ_1, tako da velja: $\langle x_1, x_1 \rangle = 1$ in $Ax_1 = \lambda_1 x_1$. Naj bo $U_1 = [x_1\ u_2\ \ldots\ u_n] \in M_n$ unitarna matrika, katere prvi stolpec je vektor x_1. Obstoj take matrike jamči izrek 3.19. Z notacijo iz tega izreka je unitarna matrika U_1 oblike $U_1 := U(x_1, e_1)$, kjer e_1 predstavlja enotski vektor z 1

v prvi vrstici in ničlami povsod drugod. Potem velja:

$$U_1^* A U_1 = U_1^* A \begin{bmatrix} x_1 & u_2 & \dots & u_n \end{bmatrix} = U_1^* \begin{bmatrix} Ax & Au_2 & \dots & Au_n \end{bmatrix} =$$

$$= U_1^* \begin{bmatrix} \lambda_1 x_1 & Au_2 & \dots & Au_n \end{bmatrix} = \begin{bmatrix} x_1^* \\ u_2^* \\ \vdots \\ u_n^* \end{bmatrix} \begin{bmatrix} \lambda_1 x_1 & Au_2 & \dots & Au_n \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda_1 x_1^* x_1 & x_1^* A u_2 & \dots & x_1^* A u_n \\ \lambda_1 u_2^* x_1 & u_2^* A u_2 & \dots & u_2^* A u_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 u_n^* x_1 & u_n^* A u_2 & \dots & u_n^* A u_n \end{bmatrix}.$$

Ker je U_1 unitarna matrika, po izreku 3.4 velja, da stolpci te matrike predstavljajo ortonormirano bazo prostora $\mathbb{C}^n$. Ker je $x_1^* x_1 = 1$ in $u_i^* x_1 = 0$ za vsak $i = 2, \dots, n$, za dobljeno matriko velja:

$$\begin{bmatrix} \lambda_1 & \diamondsuit \\ 0 & A_1 \end{bmatrix}, \text{ kjer je } A_1 \in M_{n-1}.$$

Če je $n = 2$ smo končali, saj smo dobili zgornjetrikotno matriko. Sicer postopek nadaljujemo. Najprej pokažimo, da so $\lambda_2, \dots, \lambda_n$ lastne vrednosti matrike A_1. Ker je matrika U unitarna, velja: $\det U^* \cdot \det U = \det U^* U = \det I = 1$. Karakteristični polinom matrike A je:

$$\det(A - \lambda I_n) \cdot 1 = \det(A - \lambda I_n) \det U^* U = \det U^* \det(A - \lambda I_n) \det U =$$
$$= \det(U^* A U - U^* \lambda I_n U) = \det(U^* A U - \lambda I_n) = (\lambda_1 - \lambda) \det(A_1 - \lambda I_{n-1}).$$

Ker je λ_1 ničla karakterističnega polinoma matrike A, morajo tudi preostale ničle karakterističnega polinoma matrike A sovpadati z ničlami karakterističnega polinoma matrike A_1.

Za matriko A_1 izberimo lastno vrednost λ_2 in normirajmo lastni vrednosti pripadajoč lastni vektor. Tak normiran lastni vektor označimo z $x_2 \in \mathbb{C}^{n-1}$. Obstaja unitarna matrika U_2, saj nam njen obstoj jamči izrek 3.19 (na primer konstruiramo jo kot $U_2 := U(x_2, e_1)$, kjer je $e_1 \in \mathbb{C}^{n-1}$). Potem velja: $U_2^* A_1 U_2 = \begin{bmatrix} \lambda_2 & \diamondsuit \\ 0 & A_2 \end{bmatrix}$, kjer je $A_2 \in M_{n-2}$.
Naj bo $V_2 = [1] \oplus U_2$. Potem je matrika V_2 unitarna, saj je matrika $[1]$ unitarna, prav tako pa je unitarna matrika U_2. Po trditvi 3.2 sledi, da je direktna vsota dveh unitarnih matrik

unitarna matrika. Izračunajmo:

$$(U_1V_2)^*\, AU_1V_2 = V_2^*U_1^*AU_1V_2 = \begin{bmatrix} \lambda_1 & \diamondsuit & \diamondsuit \\ 0 & \lambda_2 & \diamondsuit \\ 0 & 0 & A_2 \end{bmatrix}$$

Če je $n = 3$ smo končali, sicer postopek induktivno ponavljamo. Izrek 3.19 jamči, da lahko na vsakem koraku konstruiramo unitarno matriko. Tako dobimo unitarne matrike $U_i \in M_{n-i+1}$, kjer je $i = 1, \ldots, n-1$ in unitarne matrike $V_j \in M_n$, kjer $j = 2, \ldots, n-1$. Matrika $U := U_1V_2V_3 \ldots V_{n-1}$ je unitarna, saj je dobljena kot produkt unitarnih matrik. Matrika U^*AU pa je zgornjetrikotna matrika.
Če so vse lastne vrednosti realne matrike A realne, potem lahko izberemo pripadajoče realne lastne vektorje in posledično namesto unitarnih dobimo realne ortogonalne matrike. □

Preden bomo ilustrirali uporabo gornjega izreka na zgledu, samo še opomba, da bomo v nadaljevanju včasih za gornji izrek uporabljali izraz Schurova forma. To je še tretje poimenovanje tega izreka.

Zgled. Naj bo dana matrika $A = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$. Izračunajmo A^{50}.
Najprej preverimo, ali je matrika A diagonalizabilna. Če je matrika diagonalizabilna, potem lahko uporabimo podobno idejo, kot smo jo v zgledu ob koncu drugega poglavja. Zato moramo najprej poiskati lastne vrednosti matrike A. Pri tem si bomo pomagali s karakterističnim polinomom:

$$p_A(\lambda) = \det(A - \lambda I) = \begin{bmatrix} 4-\lambda & 0 & -1 \\ 0 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{bmatrix} = (4-\lambda)(3-\lambda)(2-\lambda) + 3 - \lambda =$$
$$= -\lambda^3 + 9\lambda^2 - 27\lambda + 27 = 0$$

Izračunamo, da so ničle karakterističnega polinoma $\lambda_{1,2,3} = 3$. Torej je edina lastna vrednost te matrike enaka 3, njena algebrajska večkratnost je prav tako enaka 3. Poiščimo lastne vektorje, ki pripadajo tej lastni vrednosti:

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

Od tod razberemo, da je edini lastni vektor pripadajoč lastni vrednosti $\lambda = 3$ oblike $v_1 = [x,\ 0,\ x]^T$, kjer je $x \in \mathbb{R}$. Torej je geometrijska večkratnost lastne vrednosti enaka ena. Ker geometrijska in algebrajska večkratnost lastne vrednosti nista enaki, matrika ni diagonalizabilna (glej izrek 2.28). Kljub temu pa lahko s pomočjo Schurovega izreka poskusimo poiskati zgornjetrikotno matriko.
Izberimo $x = 1$. Potem je $v_1 = [1,\ 0,\ 1]^T$ lastni vektor za lastno vrednost $\lambda = 3$. Tedaj je $\|v_1\| = \sqrt{2}$ in vektor $u_1 = \frac{v_1}{\|v_1\|} = \left[\frac{\sqrt{2}}{2},\ 0,\ \frac{\sqrt{2}}{2}\right]^T$ je normiran vektor.
S tem smo našli normiran lastni vektor, pripadajoč lastni vrednosti matrike A. Sedaj lahko konstruiramo Householderjevo matriko za vektorja u_1 in e_1. Definirajmo vektor: $w = u_1 - e_1 = \left[\frac{\sqrt{2}-2}{2},\ 0,\ \frac{\sqrt{2}}{2}\right]^T$.

$$O = I - 2\langle w, w\rangle^{-1} w w^T = I - \left(2+\sqrt{2}\right)\begin{bmatrix} \frac{3-2\sqrt{2}}{2} & 0 & \frac{1-\sqrt{2}}{2} \\ 0 & 0 & 0 \\ \frac{1-\sqrt{2}}{2} & 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & -1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

Matrika O je ortogonalna, še več matrika O je tudi simetrična. Zato lahko izračunamo:

$$B = O^T A O = OAO = \begin{bmatrix} 3 & 0 & 2 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Opazimo, da za matriko $B_2 = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ velja, da je B_2^2 ničelna matrika. Zato lahko izračunamo:

$$B^{50} = \left(\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)^{50} =$$

$$= \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}^{50} + \binom{50}{1}\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}^{49}\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3^{50} & 0 & 100\cdot 3^{49} \\ 0 & 3^{50} & 0 \\ 0 & 0 & 3^{50} \end{bmatrix}.$$

Ker je matrika O ortogonalna je $A^{50} = (OBO)^{50} = OB^{50}O$. Do tega, da izračunamo A^{50} manjka le še zadnji korak. Matriko B je potrebno iz leve in desne množiti z matriko O:

$$A^{50} = 3^{49}\cdot O\begin{bmatrix} 3 & 0 & 100 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} O = 3^{49}\cdot \begin{bmatrix} 53 & 0 & -50 \\ 0 & 3 & 0 \\ 50 & 0 & -47 \end{bmatrix}.$$

Posledica 5.2 *Naj bo dana matrika $A \in M_n$, ki ima lastne vrednosti $\lambda_1, \ldots, \lambda_n$, da zanjo obstaja unitarna matrika U, da je matrika U^*A^TU zgornjetrikotna matrika. Naj bo $V = \overline{U}$, potem je V^*AV spodnjetrikotna matrika.*

Dokaz. Naj bo $B = A^T$. Potem izrek 5.1 jamči, da obstaja taka unitarna matrika U, da je U^*BU zgornjetrikotna matrika. Torej je zgornjetrikotna tudi matrika U^*A^TU. Potem je matrika $\left(U^*A^TU\right)^T = U^TA\overline{U} = \overline{U}^*A\overline{U} = V^*AV$ spodnjetrikotna matrika. □

Naslednji zgled bo ilustriral, zakaj je bilo v točki (b) izreka o Schurovi formi posebej poudarjeno, da morajo biti lastne vrednosti realne matrike realna števila in izpostavil težavo na katero lahko naletimo pri iskanju lastnih vrednosti realne matrike, to je, da se lahko kot lastne vrednosti pojavijo tudi kompleksna števila.

Zgled. Poiščimo lastne vrednosti matrike $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, kjer sta $a, b \in \mathbb{R}$.

Poiščimo karakteristični polinom dane matrike:

$$p(\lambda) = \begin{vmatrix} a-\lambda & b \\ -b & a-\lambda \end{vmatrix} = (a-\lambda)^2 + b^2 = \lambda^2 - 2a\lambda + a^2 + b^2 = 0.$$

Od tod izračunamo ničli karakterističnega polinoma, ki sta: $\lambda_1 = a + bi$ in $\lambda_2 = a - bi$.

Če ima realna matrika kompleksne lastne vrednosti, matrike ni mogoče reducirati na zgornjetrikotno matriko, ki bi imela samo realne elemente, saj vemo, da se na diagonali zgornjetrikotne matrike nahajajo lastne vrednosti prvotne matrike. Kljub temu lahko tako matriko reduciramo na bločno-zgornjetrikotno matriko. Pri tem so konjugirani pari kompleksnih lastnih vrednosti povezani z 2×2 matričnimi bloki. Še preden bomo predstavili realno Schurovo formo, si oglejmo nekaj rezultatov, ki so povezani z realnimi matrikami s kompleksnimi lastnimi vrednostmi in jih bomo potrebovali pri dokazu izreka o Schurovi realni formi.

Lema 5.3 *Naj bo $A \in M_n(\mathbb{R})$ matrika, ki ima kompleksno lastno vrednost $\lambda = a + bi$ in njej pripadajoč lastni vektor x. Potem veljata naslednji enakosti:*

$$A \cdot \mathrm{Re}(x) = a\,\mathrm{Re}(x) - b\,\mathrm{Im}(x)$$

$$A \cdot \mathrm{Im}(x) = b\,\mathrm{Re}(x) + a\,\mathrm{Im}(x)$$

Dokaz. Naj bo $x = \mathrm{Re}(x) + \mathrm{Im}(x)i$.
Potem je: $Ax = A \cdot \mathrm{Re}(x) + A \cdot \mathrm{Im}(x)i$. Ker je x lastni vektor pripadajoč lastni vrednosti λ matrike A, velja: $Ax = \lambda x$.
$A \cdot \mathrm{Re}(x) + A \cdot \mathrm{Im}(x)i = \lambda\,(\mathrm{Re}(x) + \mathrm{Im}(x)i) = a\,\mathrm{Re}(x) - b\,\mathrm{Im}(x) + i\,(a\,\mathrm{Im}(x) + b\,\mathrm{Re}(x))$.
Od tod dobimo želeni enakosti: $A\cdot\mathrm{Re}(x) = a\,\mathrm{Re}(x) - b\,\mathrm{Im}(x)$ in $A\cdot\mathrm{Im}(x) = b\,\mathrm{Re}(x) + a\,\mathrm{Im}(x)$.
□

Lema 5.4 *Naj bo $A \in M_n(\mathbb{R})$ matrika s kompleksno lastno vrednostjo $\lambda = a + bi$ in s pripadajočim lastnim vektorjem x. Potem sta vektorja* $\mathrm{Re}(x)$ *in* $\mathrm{Im}(x)$ *linearno neodvisna vektorja v $\mathbb{R}^n$.*

Dokaz. Vemo, da kompleksni lastni vrednosti realne matrike pripada kompleksni lastni vektor. Predpostavimo, da sta vektorja $\mathrm{Re}(x)$ in $\mathrm{Im}(x)$ linearno odvisna. Potem obstajata taka skalarja $\alpha_1, \alpha_2 \in \mathbb{R}$ ne oba enaka 0, da velja: $\alpha_1\,\mathrm{Re}(x) + \alpha_2\,\mathrm{Im}(x) = 0$. Brez škode za splošnost lahko predpostavimo, da je $\alpha_1 \neq 0$. Potem je $\mathrm{Re}(x) = -\frac{\alpha_2}{\alpha_1}\,\mathrm{Im}(x)$. Ob upoštevanju leme 5.3 izračunajmo:

$$
\begin{aligned}
&0 = A\,(\alpha_1\,\mathrm{Re}(x) + \alpha_2\,\mathrm{Im}(x)) = \alpha_1 A\,\mathrm{Re}(x) + \alpha_2 A\,\mathrm{Im}(x) = \\
&= \alpha_1\,(a\,\mathrm{Re}(x) - b\,\mathrm{Im}(x)) + \alpha_2\,(a\,\mathrm{Im}(x) + b\,\mathrm{Re}(x)) = \\
&= \alpha_1 a\,\mathrm{Re}(x) - \alpha_1 b\,\mathrm{Im}(x) + \alpha_2 a\,\mathrm{Im}(x) + \alpha_2 b\,\mathrm{Re}(x) = \\
&= -\alpha_1 a\frac{\alpha_2}{\alpha_1}\,\mathrm{Im}(x) - \alpha_1 b\,\mathrm{Im}(x) + \alpha_2 a\,\mathrm{Im}(x) - \alpha_2 b\frac{\alpha_2}{\alpha_1}\,\mathrm{Im}(x) = \left(-\alpha_1 b - \frac{\alpha_2^2}{\alpha_1} b\right)\mathrm{Im}(x)
\end{aligned}
$$

Ker je lastna vrednost kompleksno število velja, da $b \neq 0$. Zato mora veljati, da je: $\alpha_1 + \frac{\alpha_2^2}{\alpha_1} = 0$. Torej: $\alpha_1^2 = -\alpha_2^2$. Ker smo predpostavili, da je $\alpha_1 \neq 0$, velja, da je $\alpha_1^2 > 0$. Potem bi moralo biti $\alpha_2^2 < 0$, kar pa ni mogoče. Prišli smo do protislovja, torej sta vektorja $\mathrm{Re}(x)$ in $\mathrm{Im}(x)$ linearno neodvisna vektorja v $\mathbb{R}^n$. □

Naslednji izrek bo pokazal, da je realna matrika s kompleksno lastno vrednostjo podobna bločni-zgornjetrikotni matriki, katere 2×2 diagonalni blok vsebuje realno in imaginarno komponento lastne vrednosti.

Izrek 5.5 *Naj bo dana matrika $A \in M_n(\mathbb{R})$ in naj bo $\lambda = a + bi$ lastna vrednost te matrike ter x lastni vrednosti pripadajoči lastni vektor. Potem obstaja nesingularna matrika $S \in M_n(\mathbb{R})$, da zanjo velja: $S = [\mathrm{Re}(x)\ \mathrm{Im}(x)\ S_1]$ in $S^{-1}AS = \begin{bmatrix} B & D \\ 0 & A_1 \end{bmatrix}$, kjer je matrika $S_1 \in M_{n,n-2}(\mathbb{R})$, $B \in M_2(\mathbb{R})$ oblike: $B = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ in $A_1 \in M_{n-2}(\mathbb{R})$.*

Dokaz. Lema 5.4 zagotavlja, da sta vektorja $\mathrm{Re}(x)$ in $\mathrm{Im}(x)$ iz $\mathbb{R}^n$ linearno neodvisna. Vemo, da lahko izberemo linearno neodvisne vektorje $x_1,\ x_2,\ \ldots,\ x_{n-2} \in \mathbb{R}^n$, da bodo dopolnili vektorja $\mathrm{Re}(x)$, $\mathrm{Im}(x)$ do baze prostora $\mathbb{R}^n$. Dobljene vektorje vpišimo v stolpce matrike $S := \begin{bmatrix}\mathrm{Re}(x) & \mathrm{Im}(x) & x_1 & x_2 & \ldots & x_{n-2}\end{bmatrix}$. Ker so stolpci matrike linearno neodvisni, je matrika S nesingularna, zato obstaja njen inverz. Najprej zmnožimo inverz matrike S z matriko $[\mathrm{Re}(x)\ \mathrm{Im}(x)]$:

$$S^{-1}\begin{bmatrix}\mathrm{Re}(x) & \mathrm{Im}(x)\end{bmatrix} = \begin{bmatrix}S^{-1}\,\mathrm{Re}(x) & S^{-1}\,\mathrm{Im}(x)\end{bmatrix} = \begin{bmatrix}e_1 & e_2\end{bmatrix} = \begin{bmatrix}I_2\\ 0\end{bmatrix}.$$

Ob upoštevanju leme 5.3 izračunajmo:

$$\begin{aligned}
&S^{-1}AS = S^{-1}A\begin{bmatrix}\mathrm{Re}(x) & \mathrm{Im}(x) & S_1\end{bmatrix} = S^{-1}\begin{bmatrix}A\,\mathrm{Re}(x) & A\,\mathrm{Im}(x) & AS_1\end{bmatrix} = \\
&= S^{-1}\begin{bmatrix}\begin{bmatrix}\mathrm{Re}(x) & \mathrm{Im}(x)\end{bmatrix}\begin{bmatrix}a & b\\ -b & a\end{bmatrix} & AS_1\end{bmatrix} = \begin{bmatrix}S^{-1}\begin{bmatrix}\mathrm{Re}(x) & \mathrm{Im}(x)\end{bmatrix}B & S^{-1}AS_1\end{bmatrix} = \\
&= \begin{bmatrix}\begin{bmatrix}I_2\\ 0\end{bmatrix}B & S^{-1}AS_1\end{bmatrix} = \begin{bmatrix}B & D\\ 0 & A_1\end{bmatrix},
\end{aligned}$$

kjer je $B = \begin{bmatrix}a & b\\ -b & a\end{bmatrix}$ in $\begin{bmatrix}D\\ A_1\end{bmatrix} = S^{-1}AS_1 \in M_{n,n-2}(\mathbb{R})$. □

Izrek 5.6 (Realna Schurova forma) *Naj bo $A \in M_n(\mathbb{R})$. Potem veljata naslednji trditvi:*

(a) Obstaja realna nesingularna matrika $S \in M_n(\mathbb{R})$, da je matrika $S^{-1}AS$ realna bločno-zgornjetrikotna matrika oblike:

$$\begin{bmatrix}A_1 & & & \Diamond\\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_m\end{bmatrix}, \tag{5.1}$$

kjer za vsak A_i velja, da je bodisi 1×1 ali 2×2 matrični blok z naslednjimi lastnostmi:

(i) 1×1 bloki vsebujejo realne lastne vrednosti matrike A,

(ii) vsak izmed 2×2 diagonalnih blokov ima posebno obliko in pripada nerealnim lastnim vrednostim matrike A. Blok je oblike $\begin{bmatrix}a & b\\ -b & a\end{bmatrix}$, kjer sta $a, b \in \mathbb{R}$, $b > 0$. Pri tem je $a \pm bi$ lastna vrednost matrike A,

(iii) zaporedje diagonalnih blokov matrike oblike 5.1 je določeno z zaporedjem lastnih vrednosti matrike. Bloki se lahko pojavijo v kateremkoli želenem vrstnem redu.

(b) Obstaja realna ortogonalna matrika $Q \in M_n(\mathbb{R})$, da je matrika $Q^T AQ$ realna bločno-zgornjetrikotna matrika z naslednjimi lastnostmi:

(i) 1×1 diagonalni bloki pripadajo realnim lastnim vrednostim matrike A,

(ii) 2×2 diagonalni bloki pripadajo konjugiranim parom kompleksnih lastnih vrednosti, vendar ti bloki nimajo posebne oblike,

(iii) zaporedje diagonalnih blokov je določeno na naslednji način: če so realne lastne vrednosti in konjugirani pari lastnih vrednosti določeni v nekem vrstnem redu, potem se diagonalni bloki $A_1, \ldots, A_m$ matrike $Q^T AQ$, ki pripadajo realnim lastnim vrednostim in konjugiranim parom kompleksnih lastnih vrednosti, pojavijo v istem vrstnem redu.

Dokaz. Najprej bomo dokazali točko (a). Spomnimo se točke (b) Schurove forme, ki pravi, da, če ima matrika $A \in M_n(\mathbb{R})$ zgolj realne lastne vrednosti, potem obstaja takšna ortogonalna matrika $Q \in M_n(\mathbb{R})$, da je matrika $Q^T AQ$ zgornjetrikotna matrika. V dokazu Schurove forme smo za tem pokazali, da lahko matriko A z ustreznim množenjem z ortogonalno matriko Q_1 preoblikujemo v matriko $\begin{bmatrix} \lambda & \diamondsuit \\ 0 & A_1 \end{bmatrix}$, kjer je λ realna lastna vrednost matrike A, matrika A_1 pa je realna matrika dimenzije $(n-1) \times (n-1)$.
V izreku 5.5 je bilo pokazano, kako iz matrike A, ki ima kompleksno lastno vrednost dobimo matriko $\begin{bmatrix} B & \diamondsuit \\ 0 & A_1 \end{bmatrix}$, kjer je matrika B 2×2 matrika, ki ustreza obliki podani v drugi podtočki točke (a) zgornjega izreka.
Tako smo opisali oba postopka, kako postopamo v primeru, ko ima matrika A konjugiran par kompleksnih lastnih vrednosti oziroma ko ima matrika A realno lastno vrednost. Ker ima matrika dimenzije n, šteto s kratnostjo, natanko n lastnih vrednosti, bomo postopek konstrukcije matrike S končali po končnem številu korakov. Hkrati pa lahko lastne vrednosti izbiramo v poljubnem vrstnem redu, zato se lahko bloki pojavijo v kateremkoli želenem vrstnem redu.

(b) Predpostavimo, da je dano zaporedje lastnih vrednosti matrike A (realnih in konjugiranih kompleksnih) in naj bo dana taka nesingularna matrika S, da ima matrika $S^{-1}AS$ obliko 5.1. Ker je matrika S kvadratna, izrek 3.21 jamči, da obstaja realna ortogonalna matrika $Q \in M_n$ in zgornjetrikotna matrika $R \in M_n$, da je $S = QR$. Še več, ker je S nesingularna matrika velja, da je $\operatorname{rang} A = n$ in zato so elementi Q in R enolično določeni, velja pa tudi, da so vsi elementi na glavni diagonali matrike R pozitivni. Zato lahko pišemo:

$S^{-1}AS = (QR)^{-1}A(QR) = R^{-1}Q^{-1}AQR = R^{-1}Q^TAQR$. Če razbijemo matriko R na bloke, ki bodo ustrezali matriki 5.1, potem lahko pišemo:

$$Q^TAQ = R\begin{bmatrix} A_1 & & & \diamondsuit \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_m \end{bmatrix} R^{-1} =$$

$$= \begin{bmatrix} R_{11}A_1R_{11}^{-1} & & & \diamondsuit \\ & R_{22}A_2R_{22}^{-1} & & \\ & & \ddots & \\ 0 & & & R_{mm}A_mR_{mm}^{-1} \end{bmatrix}$$

Dobljena matrika je bločno-zgornjetrikotna. Poleg tega pa še velja, da so 1×1 bloki enaki blokom v matriki 5.1, saj je $\lambda_i = R_{ii}A_iR_{ii}^{-1} = R_{ii}\lambda_iR_{ii}^{-1} = \lambda_iR_{ii}R_{ii}^{-1} = \lambda_i$, za 2×2 bloke pa velja, da je: $B_{jj} = R_{jj}A_jR_{jj}^{-1}$, kar pomeni, da so dobljeni bloki podobni 2×2 blokom matrike 5.1. □

5.2 Posledice Schurovega izreka

Schurov izrek je bolj kot za iskanje zgornjetrikotnih matrik, ki so podobne danim matrikam, uporaben kot zagotovilo, da taka zgornjetrikotna matrika obstaja. V nadaljevanju bomo predstavili nekaj rezultatov, ki jih lahko dokažemo s pomočjo dejstva o obstoju (kompleksne) zgornjetrikotne matrike.

Sled in determinanta

V prvem poglavju smo dokazali nekaj osnovnih lastnosti sledi in determinante, sedaj pa bomo z uporabo Schurovega izreka povezali sled in determinanto matrike z njenimi lastnimi vrednostmi.

Trditev 5.7 *Naj bo $A \in M_n$ matrika z lastnimi vrednostmi $\lambda_1, \lambda_2, \ldots, \lambda_n$. Potem je* $\operatorname{sled} A = \sum_{i=1}^{n} \lambda_i$ *in* $\det A = \prod_{i=1}^{n} \lambda_i$.

Dokaz. Schurov izrek jamči, da lahko vsako matriko $A \in M_n$ zapišemo v obliki: $A = U^*TU$, kjer je U unitarna matrika, T pa zgornjetrikotna matrika, ki ima na diagonali lastne

vrednosti matrike A. Potem lahko z upoštevanjem lastnosti sledi in determinante matrike izračunamo naslednje:

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

$$\operatorname{sled} A = \operatorname{sled}(U^*TU) = \operatorname{sled}(TUU^*) = \operatorname{sled} T = \sum_{i=1}^{n} \lambda_i.$$

□

Rang matrike

Posredna aplikacija Schurovega izreka lahko posreduje nekaj informacij tudi o rangu matrike $A \in M_n$. Če ima matrika A natanko k, $k \geq 1$, neničelnih lastnih vrednosti, potem je rang matrike A vsaj k. Trditev velja, saj lahko po Schurovem izreku matriko A zapišemo v obliki $A = U^*TU$. To enakost lahko zapišemo tudi drugače, kot $UA = TU$. Zaradi enakosti morata imeti matriki UA in TU enaka ranga. Ker je matrika U nesingularna velja, da je njen rang enak n. Od tod sledi, da je $\operatorname{rang} A = \operatorname{rang} T$. Rang matrike T je zagotovo vsaj k, saj ima matrika T na glavni diagonali k neničelnih lastnih vrednosti matrike A. Vrstice, ki vsebujejo te neničelne vrednosti so linearno neodvisne, lahko pa se zgodi, da je linearno neodvisna še kakšna vrstica za lastno vrednost 0.

Trditev 5.8 *Naj bo $A \in M_n$, ki ima neničelne lastne vrednosti $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_k$. Potem je* $\operatorname{rang} A \geq \frac{|\operatorname{sled} A|^2}{\operatorname{sled}(A^*A)}$.

Dokaz. Znova uporabimo Schurov izrek in zapišemo matriko A v obliki produkta unitarne in zgornjetrikotne matrike: $A = U^*TU$. V prejšnji trditvi smo pokazali, da je sled matrike A enaka vsoti neničelnih lastnih vrednosti matrike A. Potem je:

$$|\operatorname{sled} A|^2 = \left|\sum_{i=1}^{n} \lambda_i\right|^2 = \left|\sum_{i=1}^{k} \lambda_i\right|^2 \leq k \cdot \sum_{i=1}^{k} |\lambda_i|^2 =$$

$$= k \cdot \sum_{i=1}^{k} |t_{ii}|^2 \leq k \cdot \sum_{i,j=1}^{n} |t_{ij}|^2 = k \cdot \sum_{i,j=1}^{n} |a_{ij}|^2 = k \cdot \operatorname{sled}(A^*A),$$

kjer smo uporabili trditev 4.3 in naslednje: $\operatorname{sled}(A^*A) = \operatorname{sled}((U^*TU)^* U^*TU) =$
$= \operatorname{sled}(U^*T^*UU^*TU) = \operatorname{sled}(T^*T)$. Iz zgornjega izračuna sledi:

$$|\operatorname{sled} A|^2 \leq k \cdot \operatorname{sled}(A^*A) \Leftrightarrow k \geq \frac{|\operatorname{sled} A|^2}{\operatorname{sled}(A^*A)} \Leftrightarrow \operatorname{rang}(A) \geq \frac{|\operatorname{sled} A|^2}{\operatorname{sled}(A^*A)},$$

saj je $\operatorname{rang} A \geq k$. □

Naslednji izrek bo pokazal, da lahko lastno vrednost matrike zamenjamo z drugo, ne da bi s tem vplivali na preostale lastne vrednosti matrike. V angleščini ta postopek imenujejo rank-one perturbation, saj gre za to, da matriki prištejemo matriko ranga 1.

Izrek 5.9 (Alfred Brauer) *Naj bodo* $\lambda,\ \lambda_2,\ \ldots,\ \lambda_n$ *lastne vrednosti matrike* $A \in M_n$ *in naj bo* x *lastni vektor pripadajoč lastni vrednosti* λ*. Potem so za vsak vektor* $v \in \mathbb{C}^n$ *lastne vrednosti matrike* $A + xv^*$ *enake* $\lambda + v^*x,\ \lambda_2,\ \ldots,\ \lambda_n$*.*

Dokaz. Naj bo $\xi = \frac{x}{\|x\|}$ normiran lastni vektor za lastno vrednost λ in naj bo $U = [\xi\ u_2\ \ldots\ u_n]$ unitarna matrika. V prvem koraku dokaza Schurovega izreka smo pokazali, da pri množenju matrike A s konjugirano-transponiranko matrike U z leve in matriko U z desne dobimo naslednjo matriko:

$$U^*AU = \begin{bmatrix} \lambda & \diamondsuit \\ 0 & A_1 \end{bmatrix},\ \text{kjer je } A_1 \in M_{n-1}.$$

Za matriko A_1 smo ugotovili, da so njene lastne vrednosti $\lambda_2,\ \ldots,\ \lambda_n$. Izračunajmo še naslednji produkt:

$$U^*xv^*U = \begin{bmatrix} \xi^* \\ u_2^* \\ \vdots \\ u_n^* \end{bmatrix} xv^* \begin{bmatrix} \xi & u_2 & \ldots & u_n \end{bmatrix} = \begin{bmatrix} \frac{x^*}{\|x\|}x \\ u_2^*x \\ \vdots \\ u_n^*x \end{bmatrix} \begin{bmatrix} v^*\xi & v^*u_2 & \ldots & v^*u_n \end{bmatrix} =$$

$$= \begin{bmatrix} \|x\| \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} v^*\frac{x}{\|x\|} & v^*u_2 & \ldots & v^*u_n \end{bmatrix} = \begin{bmatrix} v^*x & \|x\|\, v^*u_2 & \ldots & \|x\|\, v^*u_n \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix} = \begin{bmatrix} v^*x & \diamondsuit \\ 0 & 0 \end{bmatrix}$$

Če oba rezultata seštejemo dobimo:

$$U^*AU + U^*xv^*U = U^*\left(A + xv^*\right)U = U^*\left(\begin{bmatrix} \lambda & \diamondsuit \\ 0 & A_1 \end{bmatrix} + \begin{bmatrix} v^*x & \diamondsuit \\ 0 & 0 \end{bmatrix}\right)U =$$

$$= U^*\begin{bmatrix} \lambda + v^*u & \diamondsuit \\ 0 & A_1 \end{bmatrix}U$$

Matrika $A + xv^*$ ima lastne vrednosti: $\lambda + v^*x,\ \lambda_2,\ \ldots,\ \lambda_n$. □

Cayley-Hamiltonov izrek

Cayley-Hamiltonov izrek pravi, da je vsaka kvadratna matrika ničla svojega karakterističnega polinoma. Da pa bi lahko izrek dokazali, je potrebno pred tem dokazati še "tehnično"lemo.

Lema 5.10 *Naj bo $R \in M_n$ zgornjetrikotna matrika, za katero dodatno velja, $r_{ij} = 0$, za $1 \leq i, j \leq k < n$ in $T \in M_n$ zgornjetrikotna matrika, za katero velja še, da je $t_{k+1k+1} = 0$. Potem za matriko $S = RT$ velja: $s_{ij} = 0$, za $1 \leq i, j \leq k+1$. Poleg tega pa velja še, da je spodnja desna kvadratna podmatrika matrike S, ki je dimenzije $k-2$, zgornjetrikotna.*

Pred dokazom si oglejmo zgled, ki služi kot ilustracija zgoraj opisanih matrik:

Zgled. Naj bosta dani matriki:

$$R = \left[\begin{array}{cc:c} 0 & 0 & 1 \\ 0 & 0 & 2 \\ \hdashline 0 & 0 & 3 \end{array}\right] \text{ in } T = \left[\begin{array}{cc:c} 1 & 2 & 3 \\ 0 & 4 & 5 \\ \hdashline 0 & 0 & 0 \end{array}\right].$$

Matriki R in T ustrezata pogojem leme, saj ima za $k = 2$ matrika R prva dva stolpca ničelna, element t_{33} pa je enak 0. Če predpostavimo veljavnost leme, bi morala biti matrika $S = RT$ ničelna, kar, kot lahko preverimo s preprostim izračunom, tudi drži.

Dokaz. Naj bosta dani matriki R, $T \in M_n$, ki zadoščata pogojem iz leme. Matriko R razbijemo na podmatrike tako, da bo zgornja leva podmatrika dimenzije $k \times k$. Skladno z razbitjem matrike R razbijemo tudi matriko T. Potem dobimo:

$$R = \begin{bmatrix} 0 & R_{12} \\ 0 & R_{22} \end{bmatrix} \qquad T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix},$$

kjer so matrike $T_{11} \in M_k$, $T_{22} \in M_{n-k}$ in $R_{22} \in M_{n-k}$ zgornjetrikotne matrike. Po predpostavki leme velja, da je prvi stolpec matrike T_{22} ničeln. Zato lahko pišemo: $T_{22} = [0\ V]$, kjer je V zgornjetrikotna matrika dimenzije $n-k-1$. Izračunajmo produkt matrik R in T:

$$S = RT = \begin{bmatrix} 0 & R_{12} \\ 0 & R_{22} \end{bmatrix} \cdot \begin{bmatrix} T_{11} & T_{12} \\ 0 & \begin{bmatrix} 0 & V \end{bmatrix} \end{bmatrix} =$$

$$= \begin{bmatrix} 0 \cdot T_{11} + R_{12} \cdot 0 & 0 \cdot T_{12} + R_{12} \cdot \begin{bmatrix} 0 & V \end{bmatrix} \\ 0 \cdot T_{11} + R_{22} \cdot 0 & 0 \cdot T_{12} + R_{22} \cdot \begin{bmatrix} 0 & V \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & R_{12} \cdot \begin{bmatrix} 0 & V \end{bmatrix} \\ 0 & R_{22} \cdot \begin{bmatrix} 0 & V \end{bmatrix} \end{bmatrix}$$

Opazimo, da je prvi stolpec matrike $\begin{bmatrix} R_{12} \cdot \begin{bmatrix} 0 & V \end{bmatrix} \\ R_{22} \cdot \begin{bmatrix} 0 & V \end{bmatrix} \end{bmatrix} \in M_{n,n-k}$ ničeln, zato velja, da je prvih $k+1$ stolpcev matrike S ničelnih. Hkrati je tudi desna spodnja podmatrika matrike S dimenzije $n-k-1$ zgornjetrikotna matrika, saj je dobljena kot produkt zgornjetrikotnih matrik: matrike R_{22}, ki smo ji odvzeli prvo vrstico, in matrike V. □

Izrek 5.11 (Cayley-Hamilton) *Naj bo $p_A(t)$ karakteristični polinom matrike A. Potem velja $p_A(A) = 0$.*

Dokaz. Naj bodo $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n \in \mathbb{C}$ lastne vrednosti matrike A. Zapišimo karakteristični polinom: $p_A(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$. Schurov izrek jamči, da lahko vsako kvadratno matriko zapišemo kot $A = U^*TU$, kjer je U unitarna matrika, matrika T pa je zgornjetrikotna, njeni diagonalni elementi so lastne vrednosti matrike A. Potem je:

$$p_A(A) = p_A(U^*TU) = (U^*TU - \lambda_1 U^*U)(U^*TU - \lambda_2 U^*U)\cdots(U^*TU - \lambda_n U^*U) =$$
$$= U^*(T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_n I)U = U^* p_A(T) U.$$

Zadošča pokazati, da je polinom $p_A(T) = 0$. Najprej si oglejmo produkt prvih dveh matrik: $(T-\lambda_1 I)(T-\lambda_2 I)$. Obe matriki sta zgornjetrikotni, poleg tega pa velja še, da ima prva matrika ničeln prvi stolpec, druga matrika pa ima drugi element na glavni diagonali enak nič. Tako matriki ustrezata zgornji lemi, dobimo matriko, ki ima prva dva stolpca ničelna, spodnja desna podmatrika dimenzije $n-2$ pa je zgornjetrikotna. Postopek nadaljujemo. Dobljena matrika ustreza obliki matrike R v lemi za $k=2$, hkrati pa velja, da je $(T-\lambda_3 I)_{33} = 0$. Dobljeni matriki znova ustrezata lemi, zato dobimo matriko, ki bo imela prve tri stolpce ničelne, desna spodnja podmatrika dimenzije $n-3$ pa bo zgornjetrikotna. Postopek induktivno nadaljujemo. Na zadnjem koraku imamo matriko R, ki ima $n-1$ ničelnih stolpcev, ki jo dobimo kot produkt matrik $(T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_{n-1} I)$. Matrika $(T-\lambda_n I)$ pa je zgornjetrikotna in ustreza pogoju leme za matriko T. Z vnovično uporabo leme, dobimo (podobno kot v zgornjem zgledu) ničelno matriko. □

Ena izmed pomembnih aplikacij Cayley-Hamiltonovega izreka je, da lahko poljubno potenco matrike A: A^k, kjer je $k \geq n$, izrazimo kot linearno kombinacijo matrik $A^0 = I,\ A^1,\ \ldots,\ A^{n-1}$. Oglejmo si uporabo na naslednjem zgledu:

Zgled. Naj bo dana matrika $A = \begin{bmatrix} 2 & 1 \\ 7 & 4 \end{bmatrix}$. Potem je karakteristični polinom te matrike enak: $p_A(t) = t^2 - 6t + 1$. Po Cayley-Hamiltonovem izreku velja: $p_A(A) = A^2 - 6A + I = 0$.

Od tod lahko izrazimo $A^2 = 6A - I$.
Izračunajmo nekaj potenc matrike A:
$A^3 = A^2 \cdot A = (6A - I)\, A = 6A^2 - A = 6\,(6A - I) - A = 36A - 6I - A = 35A - 6I$
$A^4 = A^3 A = (35A - 6I)\, A = 35A^2 - 6A = 35\,(6A - I) - 6A = 204A - 35I$
$A^5 = A^4 A = (204A - 35I)\, A = 204A^2 - 35A = 204\,(6A - I) - 35A = 1189A - 204I$
Izračunamo lahko tudi negativne eksponente nesingularne matrike:

$$\begin{aligned} I &= 6A - A^2 \\ I &= A\,(6I - A) \\ A^{-1} &= 6I - A \end{aligned}$$

Ugotovimo, da je inverzna matrika matrike A enaka: $\begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix}$.

Izračunajmo še $A^{-2} = (6I - A)^2 = 36I - 12A + A^2 = 36I - 12A + 6A - I = 35I - 6A$.

Posledica 5.12 *Naj bo $A \in M_n$ nesingularna matrika in naj bo $p_A(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t^1 + a_0$ karakteristični polinom te matrike. Naj bo*

$$q(t) = -\frac{1}{a_0}\left(t^{n-1} + a_{n-1}t^{n-2} + \ldots + a_2 t + a_1\right).$$

Potem je $A^{-1} = q(A)$.

Dokaz. Po Cayley-Hamiltonovem izreku velja $p_A(A) = 0$. Ker je A nesingularna matrika lahko pišemo:

$$\begin{aligned} 0 &= A^n + a_{n-1}A^{n-1} + \ldots + a_2 A^2 + a_1 A + a_0 I \\ -a_0 I &= A\left(A^{n-1} + a_{n-1}A^{n-2} + \ldots + a_2 A + a_1 I\right) \\ A^{-1} &= -\frac{1}{a_0}\left(A^{n-1} + a_{n-1}A^{n-2} + \ldots + a_2 A + a_1 I\right) \\ A^{-1} &= q(A) \end{aligned}$$

□

Normalne matrike

Zadnji primer uporabe Schurovega izreka bo že povezan z naslednjim poglavjem, in sicer je cilj tega razdelka dokazati spektralni izrek za normalne matrike. Tega bomo dokazali v nekoliko širši obliki, ki vsebuje še nekatere druge lastnosti normalnih matrik. Še prej pa bomo definirali normale matrike, predstavili nekaj osnovnih lastnosti, ki bodo uporabljene pri dokazu izreka in trditve, ki bo svoje mesto dobila pred prej napovedanim izrekom.

Definicija 5.13 *Matrika $A \in M_n$ je normalna, če A komutira z matriko A^*, torej s svojo konjugirano-transponirano matriko. Zapisano s simboli: $AA^* = A^*A$.*

Oglejmo si nekaj lastnosti normalnih matrik:

(a) Če je matrika $A \in M_n$ normalna matrika in matrika $B \in M_n$ unitarno podobna matriki A, potem je matrika B normalna matrika.
Dokaz: Ker je B unitarno podobna A velja, da obstaja unitarna matrika U, da je $B = UAU^*$. Izračunajmo: $BB^* = UAU^*\,(UAU^*)^* = UAA^*U^* = UA^*AU^* = UA^*U^*UAU^* = (UAU^*)^*\,(UAU^*) = B^*B$. Matrika B je normalna.

(b) Matriki $A \in M_n$ in $B \in M_m$ sta normalni natanko tedaj, ko je normalna matrika $A \oplus B \in M_{n+m}$.
Dokaz: Ker sta matriki A in B normalni, zanju velja: $AA^* = A^*A$ in $BB^* = B^*B$. Izračunajmo:

$$(A \oplus B) \cdot (A \oplus B)^* = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \cdot \begin{bmatrix} A^* & 0 \\ 0 & B^* \end{bmatrix} = \begin{bmatrix} AA^* & 0 \\ 0 & BB^* \end{bmatrix} =$$
$$= \begin{bmatrix} A^*A & 0 \\ 0 & B^*B \end{bmatrix} = (A \oplus B)^* \cdot (A \oplus B)$$

Dokaz v drugo smer poteka podobno.

(c) Vsaka unitarna matrika je normalna.
Dokaz: Naj bo $U \in M_n$ unitarna matrika. Potem velja $U^*U = I$, hkrati pa velja $UU^* = I$, saj za unitarno matriko velja, da je $U^* = U^{-1}$ (glej izrek 3.4). Če obe zgornji enakosti združimo, dobimo: $UU^* = U^*U$. Ker je bila matrika U poljubna unitarna matrika, to pomeni, da je vsaka unitarna matrika normalna.

Preden bomo dokazali trditev, na katero se bomo oprli pri ključnem dokazu tega poglavja, je potrebno dokazati še naslednjo lemo, ki se nanaša na lastnost sledi v matriki.

Lema 5.14 *Naj bo $A \in M_n$. Če je* $\operatorname{sled}(AA^*) = 0$*, potem je $A = 0$.*

Dokaz. V dokazu trditve 4.3 smo pokazali, da je $\operatorname{sled}(AA^*) = \sum_{i=1}^{n}\sum_{k=1}^{n} |a_{ik}|^2$. Ker velja, da je $\operatorname{sled} AA^* = 0$, mora biti $\sum_{i=1}^{n}\sum_{k=1}^{n} |a_{ik}|^2 = 0$. Ker pa je absolutna vrednost kompleksnega števila nenegativno število, mora veljati, da je $a_{ik} = 0$ za vsaka $i, k = 1, \ldots, n$. To pa pomeni, da so vsi elementi matrike A ničelni in matrika A je ničelna matrika. □

Trditev 5.15 *Naj bosta $A, B \in M_n$ kvadratni matriki s kompleksnimi elementi.*
(a) Naj ima matrika A naslednjo obliko: $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$, kjer velja: $A_{11} \in M_k$, $A_{22} \in M_{n-k}$ in $A_{12} \in M_{n,n-k}$, za $1 \leq k < n$. Matrika A je normalna natanko tedaj, ko sta matriki A_{11} in A_{22} normalni in je matrika A_{12} ničelna.
(b) Naj bo B bločno-zgornjetrikotna matrika. Matrika B je normalna natanko tedaj, ko so njeni diagonalni matrični bloki normalne matrike, vsi preostali elementi pa so enaki 0.

Dokaz. Najprej si oglejmo točko (a). ($\Leftarrow$) V primeru, da sta matriki A_{11} in A_{22} normalni matriki ter da je matrika $A_{12} = 0$, velja: $A_{11} \oplus A_{22} = A$. Pri opisu lastnosti normalnih matrik smo v točki (b) dokazali, da je direktna vsota dveh normalnih matrik normalna matrika.
($\Rightarrow$) Vemo, da je matrika A normalna. Potem zanjo velja:

$$AA^* = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \cdot \begin{bmatrix} A_{11}^* & 0 \\ A_{12}^* & A_{22}^* \end{bmatrix} = \begin{bmatrix} A_{11}A_{11}^* + A_{12}A_{12}^* & A_{12}A_{22}^* \\ A_{22}A_{12}^* & A_{22}A_{22}^* \end{bmatrix},$$
$$A^*A = \begin{bmatrix} A_{11}^* & 0 \\ A_{12}^* & A_{22}^* \end{bmatrix} \cdot \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11}^*A_{11} & A_{11}^*A_{12} \\ A_{12}^*A_{11} & A_{12}^*A_{12} + A_{22}^*A_{22} \end{bmatrix}.$$

Da bo veljalo $AA^* = A^*A$ mora biti izpolnjena enakost $A_{11}A_{11}^* + A_{12}A_{12}^* = A_{11}^*A_{11}$. Vemo, da imajo enake matrike isto sled, zato:

$$\operatorname{sled}(A_{11}^*A_{11}) = \operatorname{sled}(A_{11}A_{11}^* + A_{12}A_{12}^*) = \operatorname{sled}(A_{11}^*A_{11}) + \operatorname{sled}(A_{12}A_{12}^*).$$

Pri tem smo pri drugem enačaju uporabili lastnosti sledi, da je sled vsote dveh matrik enaka vsoti sledi vsakega sumanda in cikličnost sledi (glej izrek 1.26). Od tod dobimo, da je $\operatorname{sled}(A_{12}A_{12}^*) = 0$. V prejšnji lemi smo dokazali, da je v tem primeru $A_{12} = 0$. Če upoštevamo to ugotovitev je matrika A oblike $A = A_{11} \oplus A_{22}$. Po predpostavki je ta matrika normalna, zato po točki (b) lastnosti normalnih matrik sledi, da sta tudi matriki A_{11} in A_{22} normalni matriki.
Preostane še dokaz točke (b), ki jo bomo dokazali s pomočjo matematične indukcije. Pri tem je baza indukcije točka (a). ($\Rightarrow$) Predpostavimo, da je matrika $B = [B_{ij}]_{i,j=1}^{k} \in M_n$, $k \leq n$, normalna, bločno-zgornjetrikotna matrika. To pomeni, da so podmatrike matrike B, ki ležijo pod glavno diagonalo ničelne, torej: $B_{ij} = 0$, če je $i > j$. Naj trditev velja za matriko, ki je razbita na $k-1$ diagonalnih blokov, dokazati želimo, da trditev velja tudi za matriko, ki je razbita na k diagonalnih blokov.
Oglejmo si naslednjo delitev matrike $B = \begin{bmatrix} B_{11} & E \\ 0 & \tilde{B} \end{bmatrix}$, kjer je $B_{11} \in M_m$, $1 \leq m < n$, in $\tilde{B} = [B_{ij}]_{i,j=1}^{k-1} \in M_{n-m,n-m}$ bločno-zgornjetrikotna matrika, matrika E pa ima obliko: $E = [B_{12}\ B_{13}\ \ldots\ B_{1k}] \in M_{m,n-m}$. Po točki (a) sledi, da sta matriki B_{11} in $\tilde{B}$ normalni,

matrika E pa je ničelna matrika. Ker je matrika $\tilde{B}$ normalna matrika, ki je razbita na $k-1$ diagonalnih blokov je trditev dokazana. Matrika B je torej sestavljena iz diagonalnih blokov, ki so normalne matrike, vse ostale nediagonalne podmatrike matrike B pa so ničelne.
($\Leftarrow$) Dokaz v drugo smer je očiten, saj smo pokazali, da je direktna vsota dveh normalnih matrik normalna matrika, da je direktna vsota k normalnih matrik normalna matrika, pa pokažemo s pomočjo matematične indukcije. □

Posledica 5.16 *Zgornjetrikotna matrika je normalna natanko tedaj, ko je diagonalna.*

Dokaz. Trditev sledi neposredno iz točke (b) zgornjega izreka, saj lahko zgornjetrikotno matriko razbijemo na podmatrike tako, da je vsak diagonalni element 1×1 diagonalni blok matrike. Potem po zgornjem izreku velja, da so vsi elementi matrike, ki niso vsebovani v diagonalnih blokih, ničelni. Od tod sledi, da je matrika diagonalna. Obrat trditve znova dokažemo z uporabo lastnosti normalnih matrik, da je direktna vsota normalnih matrik normalna matrika. Pri tem ni težko videti, da so 1×1 matrike normalne, saj je množenje teh matrik ekvivalentno množenju kompleksnih števil, ki je komutativno. □

S temi rezultati smo končno pripravili vse potrebno, da dokažemo izrek, ki smo ga napovedali v uvodu.

Izrek 5.17 *Naj bo $A \in M_n$ kvadratna matrika s kompleksnimi elementi in lastnimi vrednostmi $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$. Potem so naslednje trditve ekvivalentne.*

(a) A je normalna matrika.

(b) A je unitarno diagonalizabilna.

(c) A premore n ortonormiranih lastnih vektorjev.

Dokaz. Ker je matrika $A \in M_n$ kvadratna matrika s kompleksnimi elementi, lahko uporabimo Schurov izrek (izrek 5.1), ki pravi, da obstajata taka unitarna matrika U in zgornjetrikotna matrika T, da velja: $T = U^*AU$ oziroma $A = UTU^*$. Denimo, da je matrika A normalna matrika. Potem je normalna tudi matrika T, saj je matrika T unitarno podobna matriki A, po točki (a) lastnosti normalnih matrik pa vemo, da je matrika, ki je unitarno podobna normalni matriki, normalna. Ker je matrika T zgornjetrikotna in normalna matrika, po prejšnji posledici velja, da je matrika diagonalna. Od tod pa sledi, da je matrika A unitarno diagonalizabilna. S tem smo dokazali ($a \Rightarrow b$).
Pokažimo, da iz b sledi c. Vemo, da je matrika A unitarno diagonalizabilna, torej obstajata

diagonalna matrika $D = \text{diag}\,(\lambda_1,\ \ldots, \lambda_n) \in M_n$ in unitarna matrika $V = [v_1\ \ldots\ v_n] \in M_n$, da velja: $AV = VD$. Od tod sledi, da je $Av_i = d_{ii}v_i = \lambda_i v_i$ za vsak $i \in \{1,\ \ldots,\ n\}$. Ker je matrika V unitarna, zanjo velja, da njeni stolpci tvorijo ortonormirano bazo. Ker je vsak stolpec v_i matrike V pripadajoč neki lastni vrednosti λ_i velja, da lastni vektorji matrike A tvorijo ortonormirano bazo prostora $\mathbb{C}^n$.
Preostane še dokazati $(c \Rightarrow a)$. Ker ima matrika A n ortonormiranih lastnih vektorjev, so njeni lastni vektorji linearno neodvisni. Zato (po izreku 2.24, točka ii) obstaja taka nesingularna matrika S, da velja: $A = S^{-1}DS$, kjer je D diagonalna matrika. Ker pa so stolpci matrike S ortonormirani, je matrika S unitarna. To sledi po točki (e) izreka 3.4. Velja: $A = S^*DS$ oziroma $D = SAS^*$, kar pomeni, da je matrika A unitarno podobna diagonalni matriki D. Po prejšnji posledici je vsaka diagonalna matrika normalna. Vemo pa, da je matrika, ki je unitarno podobna normalni matriki normalna (točka (a) lastnosti normalnih matrik). Od tod sledi, da je matrika A normalna in izrek je dokazan. □

Trditev, da je matrika normalna natanko tedaj, ko je unitarno diagonalizabilna včasih imenujemo tudi spektralni izrek za normalne matrike.

Poglavje 6

Ortogonalna in unitarna grupa

V prejšnjih poglavjih, ko smo proučevali lastnosti unitarnih matrik, smo se pretežno ukvarjali s pojmi linearne algebre. Vse lastnosti, ki smo jih dokazali za unitarne matrike veljajo tudi za ortogonalne matrike, saj lahko na elemente ortogonalnih matrik gledamo kot na elemente unitarnih matrik, ki imajo ničeln imaginarni del. V tem poglavju bomo iz linearne algebre nekoliko posegli na polje abstraktne algebre, kamor uvrščamo grupe.
Ortogonalne in unitarne grupe skupaj s simplektično grupo, ki je ne bomo podrobneje predstavljali, imenujemo matrične grupe. Študij teh grup običajno pripelje do Liejeve teorije grup, ki združuje algebro, analizo in topologijo. Cilj tega poglavja ni študij Liejevih grup, temveč zgolj dokaz nekaterih preprostih algebrskih lastnosti ortogonalnih in unitarnih grup.

6.1 Osnovni pojmi in primeri

Grupo smo definirali že v prvem poglavju, in sicer kot množico G na kateri je definirana (notranja) binarna operacija. Par $(G, \circ)$ je grupa, če notranja binarna operacija $\circ$ zadošča trem lastnostim: izpolnjena mora biti asociativnost, v grupi mora biti vsebovan nevtralni element, poleg tega mora še veljati, da vsak element iz grupe premore inverzni element.

Navedimo nekaj primerov grup: Množica celih števil z operacijo seštevanja $(\mathbb{Z}, +)$ je grupa, podobno sta tudi množici realnih in kompleksnih števil z operacijo seštevanja grupi. Po drugi strani pa množica celih števil z operacijo odštevanja ni grupa, saj operacija ni asociativna. Prav tako je grupa tudi množica neničelnih realnih števil (označevali jo bomo: $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$, podobno bo veljalo tudi za kompleksna števila: $\mathbb{C}^* = \mathbb{C}\backslash\{0\}$) z operacijo množenja: $(\mathbb{R}^*, \cdot)$. Opazimo pa lahko, da množica celih števil brez števila 0 z operacijo množenja ni grupa, saj v množici celih števil, razen za elementov -1 in 1, ne moremo najti inverznih elementov za množenje.

Vse zgoraj omenjene grupe so neskončne, saj množice vsebujejo neskončno elementov. Če ima množica G končno število elementov, potem lahko definiramo red grupe. Gre za število elementov v množici G. Če je grupa končnega reda, jo lahko predstavimo s tako imenovano Cayleyevo tabelo. Množica G je z operacijo $\cdot$ grupa, če se vsak element iz množice G v vsaki vrstici in v vsakem stolpcu Cayleyeve tabele pojavi natanko enkrat.

Zgled. S pomočjo Cayleyjeve tabele preverimo, da je množica $G = \{1, -1, i, -i\} \subset \mathbb{C}$ z operacijo množenja kompleksnih števil grupa.

$\cdot$	1	-1	i	$-i$
1	1	-1	i	$-i$
-1	-1	1	$-i$	i
i	i	$-i$	-1	1
$-i$	$-i$	i	1	-1

Grupe tvorijo tudi matrike. Tako je na primer $(M_n(\mathbb{R}), +)$ grupa kvadratnih matrik dimenzije $n \times n$ z operacijo matričnega seštevanja. Na podobno težavo kot pri množici celih števil z operacijo množenja naletimo tudi pri matrikah, saj vemo, da vsaka kvadratna matrika ne premore inverza. Zato je za definiranje grupe kvadratnih $n \times n$ matrik z matričnim množenjem potrebno poiskati podmnožico kvadratnih matrik. To je na primer množica: $G = \{A \in M_n(\mathbb{R}) : \det A \neq 0\}$. Tako definirano množico skupaj z matričnim množenjem imenujemo splošna linearna grupa in jo označimo z $GL_n(\mathbb{R})$.
Spomnimo se še, da je grupa Abelova, če za operacijo v grupi velja komutativnost. Vidimo, da so vse navedene grupe, razen grupe $GL_n(\mathbb{R})$ (množenje matrik ni komutativna operacija), Abelove.
Glede splošne linearne grupe še opomba, da bomo običajno obravnavali omenjeno grupo z elementi iz polja realnih števil. Ker pa smo determinanto definirali kot preslikavo iz množice kvadratnih matrik s kompleksnimi elementi v množico kompleksnih števil, lahko definiramo tudi splošno linearno grupo, ki jo sestavljajo matrike s kompleksnimi elementi: $GL_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) : \det A \neq 0\}$.
Ob koncu opomba, da bomo takrat, ko bo jasno, katero operacijo v grupi imamo v mislih grupo označevali kar z G in ne z $(G, \circ)$, za operacijo v grupi pa bomo v teh primerih uporabljali standardno notacijo za množenje.

Podgrupa

Nekoliko podrobneje si oglejmo grupi $H = (\mathbb{Z}, +)$ in $G = (\mathbb{Q}, +)$. Opazimo, da je vsak element, ki je v grupi H, tudi v grupi G, saj je $\mathbb{Z} \subset \mathbb{Q}$. V obeh grupah pa je definirana ista operacija. To nas pripelje do naslednje definicije:

Definicija 6.1 *Podgrupa $(H, *)$ grupe $(G, *)$ je neprazna množica $H \subseteq G$, ki je grupa za operacijo $*$. Oznaka: $H \leq G$, če je $H \subseteq G$ in $H < G$, če je $H \subset G$.*

Z upoštevanjem zgornje definicije lahko zapišemo: $(\mathbb{Z}, +) < (\mathbb{Q}, +) < (\mathbb{R}, +) < (\mathbb{C}, +)$ in $(\mathbb{Q}^*, \cdot) < (\mathbb{R}^*, \cdot) < (\mathbb{C}^*, \cdot)$.

Trditev 6.2 *Naj bo H podmnožica v grupi G. Potem so naslednje trditve ekvivalente:*

(a) H je podgrupa grupe G.

(b) H zadošča naslednjim pogojem:

(i) identični element grupe G je vsebovan v H,

(ii) za poljubna $g, h \in H$ velja $gh \in H$,

(iii) za vsak $h \in H$ je $h^{-1} \in H$.

(c) H zadošča naslednjima pogojema:

(i') identični element grupe G je vsebovan v H,

(ii') za poljubna $g, h \in H$ velja $gh^{-1} \in H$.

Dokaz. (a) $\Rightarrow$ (b) Če je H podgrupa grupe G po definiciji velja, da je H grupa, torej je H zaprta za množenje (točka ii) in vsebuje inverze vseh svojih elementov (točka iii). Ker je H grupa premore enoto, ki jo označimo z $e_H \in H$, da za vsak $h \in H$ velja: $e_H h = h e_H = h$. Ker je $H \subseteq G$ je $h \in G$ in naj bo $e_G \in G$ enota grupe G. Potem je $e_G h = h e_G = h$. Če obe enakosti združimo dobimo: $e_G h = e_H h$. To enakost pomnožimo z desne s $h^{-1} \in H$. Dobimo $e_H = e_G$. S tem smo dokazali še točko i.
(b) $\Rightarrow$ (c) Pogoja i in i' sta enaka, zato ostane zgolj, da preverimo pogoj ii'. Ker za poljubna $g, h \in H : gh \in H$ in ker je $h^{-1} \in H$ (to zagotavljata predpostavki ii in iii), dobimo $gh^{-1} \in H$.
(c) $\Rightarrow$ (a) Vemo, da je H podmnožica v grupi G. Pokazati moramo, da je H grupa. Množica H ni prazna, saj H vsebuje vsaj enoto e (to zagotavlja točka i'). Preverimo še preostale pogoje, da bo H grupa: ker v grupi G velja asociativnost se le-ta prenese na vse podmnožice, torej tudi na H. Po točki i' je v H enota. Po točki ii' za poljubna $g, h \in H$ velja $gh^{-1} \in H$. Torej je H zaprta za množenje. Če izberemo $g = e$, potem je $eh^{-1} = h^{-1} \in H$. To pomeni, da imajo vsi elementi iz H v tej množici tudi svoje inverzne elemente. S tem so izpolnjeni vsi pogoji, da je H grupa. □

Ilustrirajmo zgoraj zapisano trditev na grupi $GL_n(\mathbb{R})$. Oglejmo si podmnožico $\mathcal{A} \subset GL_n(\mathbb{R})$, ki vsebuje vse matrike, ki imajo determinanto enako 1. S pomočjo točke (b) gornje trditve dokažimo, da je množica $\mathcal{A}$ podgrupa grupe $GL_n(\mathbb{R})$. Identična matrika I_n je vsebovana v množici $\mathcal{A}$, saj je $\det I_n = 1$. Naj bosta $A, B \in \mathcal{A}$ poljubni. Potem velja: $\det(AB) = \det A \cdot \det B = 1$ in zato $AB \in \mathcal{A}$. Preostane še pokazati, da je $A^{-1} \in \mathcal{A}$. Velja: $\det\left(A^{-1}\right) = \frac{1}{\det A} = 1$. S tem smo pokazali, da je množica $\mathcal{A}$ podgrupa v grupi $GL_n(\mathbb{R})$. To grupo imenujemo specialna linearna grupa in jo označimo s $SL_n(\mathbb{R})$.
Na podoben način pridemo tudi do specialne linearne grupe s kompleksnimi elementi; oznaka: $SL_n(\mathbb{C})$, ki je podgrupa grupe $GL_n(\mathbb{C})$. Poleg tega pa opazimo še, da so vse matrike iz $GL_n(\mathbb{R})$ vsebovane v $GL_n(\mathbb{C})$, zato je $GL_n(\mathbb{R}) < GL_n(\mathbb{C})$ in da velja: $SL_n(\mathbb{R}) < SL_n(\mathbb{C})$.

Ortogonalna in unitarna grupa

V tretjem poglavju smo dokazali veliko lastnosti ortogonalnih in unitarnih matrik, ki bodo pomagale pri dokazu, da sta množici ortogonalnih $n \times n$ matrik z realnimi elementi (oznaka O_n) oziroma unitarnih $n \times n$ matrik (oznaka U_n) skupaj z matričnim množenjem grupi. Omenjeni grupi imata še nekatere druge lastnosti, ki prav tako sledijo iz lastnosti dokazanih v tretjem poglavju, vendar o tem več v nadaljevanju. Najprej pokažimo, da sta omenjeni množici z matričnim množenjem res grupi.

Trditev 6.3 *Množici $O_n \subset M_n(\mathbb{R})$ in $U_n \subset M_n(\mathbb{C})$ tvorita skupaj z operacijo množenja matrik grupo.*

Dokaz. Dokaz za unitarno grupo: Najprej preverimo, da je operacija matričnega množenja notranja: torej $\forall U, V \in U_n : UV \in U_n$. To velja po posledici 3.5. Da bo množica U_n skupaj z operacijo matričnega množenja grupa, še mora zadoščati naslednjim trem pogojem:

G 1 Asociativnost: v množici vseh kvadratnih matrik velja asociativnost množenja. Potem asociativnost velja tudi v vsaki podmnožici množice kvadratnih matrik, torej tudi v U_n.

G 2 Enota za množenje je I. Matrika I je vsebovana v množici U_n, saj je: $I^*I = I \cdot I = I$.

G 3 Inverz za množenje: $\forall A \in U_n,\ \exists A^{-1} \in U_n : A^{-1}A = AA^{-1} = I$. Ker je A unitarna matrika velja po točki (b) izreka 3.4, da je $A^* = A^{-1}$ in zato $A^*A = AA^* = I$.

Za ortogonalno grupo dokaz poteka podobno. Vse omenjene lastnosti, ki veljajo za unitarne matrike veljajo tudi za ortogonalne matrike. □

Sedaj je že jasno, da množico unitarnih matrik z operacijo matričnega množenja imenujemo unitarna grupa in jo označimo z $U_n(\mathbb{C})$ oziroma U_n, množico realnih ortogonalnih matrik z operacijo matričnega množenja pa ortogonalna grupa; oznaka $O_n(\mathbb{R})$ oziroma O_n.

Zgled. Določimo vse elemente grupe O_2.
Naj bo matrika $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, kjer so $a, b, c, d \in \mathbb{R}$, element grupe O_2. Torej je matrika A 2×2 matrika, za katero velja $A^T A = I$. Izračunajmo:

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ba + dc & b^2 + d^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Ker velja komutativnost množenja realnih števil, dobimo naslednje enačbe: $a^2 + c^2 = 1$, $b^2 + d^2 = 1$ in $ab + cd = 0$. Poleg tega mora, zaradi ortogonalnosti matrike A, veljati še, da je $\det A = \pm 1$. Torej $ad - bc = \pm 1$. Od tod dobimo naslednjo množico matrik, ki rešijo dani sistem enačb:

$$O_2 = \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \begin{bmatrix} a & b \\ b & -a \end{bmatrix} : a^2 + b^2 = 1 \right\}.$$

Vemo, da velja $\sin^2 \phi + \cos^2 \phi = 1$, za vsak $\phi \in \mathbb{R}$, zato lahko pišemo:

$$O_2 = \left\{ \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{bmatrix} : \phi \in [0, 2\pi) \right\}.$$

V poglavju 3.2.1 smo zapisali, da so matrike zgornje oblike rotacije ravnine okoli koordinatnega izhodišča. Matrike iz grupe O_2 so torej rotacije ravnine $\mathbb{R}^2$ okoli točke $O(0, 0)$

Omenjena trditev je zanimiva, ob tem pa se takoj postavi vprašanje, ali velja tudi v prostoru $\mathbb{R}^n$. Kot običajno se izkaže, da kjer je dim, je tudi ogenj. Da bomo lahko trditev formalno utemeljili, potrebujemo dve definiciji. Prva je definicija linearne preslikave, ki je predmet linearne algebre in smo se ji skozi celotno dosedanje besedilo vztrajno izmikali.

Definicija 6.4 *Naj bosta U in V poljubna vektorska prostora nad poljem skalarjev $\mathbb{F}$. Preslikavi $\mathscr{A} : U \to V$ pravimo linearna preslikava iz vektorskega prostora U v vektorski prostor V, če zadošča naslednjima lastnostma:*

(a) $\mathscr{A}(x + y) = \mathscr{A}(x) + \mathscr{A}(y)$, *za poljubna* $x, y \in U$;

(b) $\mathscr{A}(\alpha x) = \alpha \mathscr{A}(x)$, *za poljubna* $\alpha \in \mathbb{F}$ *in* $x \in U$.

Še opomba glede zapisa: običajno v zapisu $\mathscr{A}(x)$ izpuščamo oklepaje in pišemo: $\mathscr{A}x$.

Zgled. Navedimo primer linearne preslikave. Gre za običajno množenje matrike z vektorjem iz desne. Naj bo dana matrika $A \in M_n$. Potem definiramo linearno preslikavo $\mathscr{A} : \mathbb{C}^n \to \mathbb{C}^n$ s predpisom $\mathscr{A}x = Ax$. Preverimo, da gre res za linearno preslikavo. Naj bodo $x, y \in \mathbb{C}^n$ ter $\alpha \in \mathbb{C}$ poljubni, potem je:

$$\mathscr{A}(x+y) = A \cdot (x+y) = Ax + Ay = \mathscr{A}x + \mathscr{A}y$$
$$\mathscr{A}\alpha x = A(\alpha x) = \alpha Ax = \alpha \mathscr{A}x$$

S tem smo dokazali, da gre za linearno preslikavo.

O gornjem zgledu lahko povemo še nekaj več. Pravimo, da je linearna preslikava obrnljiva natanko tedaj, ko je matrika A nesingularna. Dodatno v realnih vektorskih prostorih velja, da preslikava $\mathscr{B} : \mathbb{R}^n \to \mathbb{R}^n$, podana s predpisom $\mathscr{B}x = Bx$, za nek nesingularni $B \in M_n(\mathbb{R})$, ohranja orientacijo, če je determinanta matrike B pozitivna, oziroma spreminja orientacijo, če je determinanta matrike B negativna.

Definicija 6.5 *Linearna preslikava $\mathscr{A} : \mathbb{C}^n \to \mathbb{C}^n$, za katero velja $\|x\| = \|\mathscr{A}x\|$, za vse $x \in \mathbb{C}^n$, se imenuje Evklidska izometrija ali krajše izometrija.*
Če poleg zgornjega pogoja velja še, da linearna preslikava $\mathscr{B} : \mathbb{R}^n \to \mathbb{R}^n$ ohranja orientacijo, potem bomo takšno preslikavo imenovali rotacija v $\mathbb{R}^n$ okoli izhodišča O.

Premislimo še nekaj podrobnosti iz zgornje definicije. Rotacijo v $\mathbb{R}^n$ smo definirali kot zasuk okoli izhodišča. To pomeni, da mora preslikava $\mathscr{B}$ ohranjati izhodišče O. Enako velja tudi za zgoraj definirano Evklidsko izometrijo, saj je ena izmed lastnosti linearne preslikave, ki je izpeljana iz definicije, da ničelni vektor slika vase; $\mathscr{A}0 = 0$. To ni težko pokazati, saj $\mathscr{A}0 = \mathscr{A}(0+0) = \mathscr{A}0 + \mathscr{A}0$. Če pogledamo levo in desno stran enakosti in odštejemo $\mathscr{A}0$, dobimo: $0 = \mathscr{A}0$, kar pomeni, da linearna preslikava ničelni vektor vedno preslika vase.

Obe zgornji definiciji (6.4 in 6.5) nas pripeljeta do naslednjega izreka, ki smo ga v dobršni meri že dokazali v tretjem poglavju.

Izrek 6.6 *(a) Matrika $A \in M_n(\mathbb{R})$ predstavlja rotacijo v prostoru $\mathbb{R}^n$ natanko tedaj, ko velja $AA^T = I$ in $\det A = 1$.*
(b) Matrika $B \in M_n(\mathbb{C})$ ohranja Evklidsko normo v prostoru $\mathbb{C}^n$ natanko tedaj, ko je $BB^* = I$.

Dokaz. Matrika, ki predstavlja rotacijo v prostoru $\mathbb{R}^n$ po gornji definiciji ohranja Evklidsko normo. V izreku 3.8 smo pokazali, da velja: A ortogonalna matrika $\Leftrightarrow AA^T = I \Leftrightarrow$

množenje matrike A z vektorjem $x \in \mathbb{R}^n$ z desne ohranja normo. Hkrati vemo, da lahko ima ortogonalna matrika A determinanto enako 1 ali -1. Ker rotacija ohranja orientacijo, mora biti determinanta pozitivna, torej je edina možnost $\det A = 1$.

Točko (b) gornjega izreka smo že dokazali v izreku 3.4. Povzemimo: $B \in M_n(\mathbb{C})$ ohranja Evklidsko normo v $\mathbb{C}^n \Leftrightarrow$ matrika B je unitarna $\Leftrightarrow BB^* = I$ oziroma $B^*B = I$. □

Izrek je povedal, da matrike iz ortogonalne grupe, za katere velja, da je njihova determinanta enaka 1, predstavljajo rotacije v ravnini $\mathbb{R}^n$ okoli O. To nas privede do naslednjega vprašanja, ali matrike, ki predstavljajo rotacije v ravnini $\mathbb{R}^n$ okoli O tvorijo grupo. Odgovor se skriva v naslednji trditvi.

Trditev 6.7 *Z $SO_n \subset O_n$ označimo množico vseh ortogonalnih matrik, za katere velja, da je njihova determinanata enaka* 1*. Potem je SO_n podgrupa grupe O_n.*

Dokaz. Preveriti moramo, da veljajo vse zahteve iz točke (b) trditve 6.2. Naj bosta $A, B \in SO_n$ potem je: $\det(AB) = \det A \det B = 1$ in zato je množica SO_n zaprta za matrično množenje. Dokazati je potrebno še, da v SO_n leži identična matrika, kar je očitno. Za poljubno matriko $A \in SO_n$ je matrika $A^{-1} \in SO_n$, saj $\det A^{-1} = \frac{1}{\det A} = 1$. Zato je SO_n podgrupa grupe O_n. □

Grupo SO_n imenujemo specialna ortogonalna grupa. Preostale matrike v O_n, ki ne ležijo v SO_n ne tvorijo grupe, saj bi za dve matriki iz podmnožice matrik O_n z determinanto -1 veljalo, da je njuna determinanta enaka 1, kar pomeni, da matrično množenje ni notranja operacija te množice.
Če zgornji razmislek apliciramo na grupo O_2, jo lahko zapišemo kot disjunktno unijo dveh množic:

$$O_2 = \left\{ \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} : \phi \in [0, 2\pi) \right\} \cup \left\{ \begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{bmatrix} : \phi \in [0, 2\pi) \right\} =$$
$$= SO_2 \cup \left\{ \begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{bmatrix} : \phi \in [0, 2\pi) \right\}.$$

Na podoben način, kot smo vpeljali specialno ortogonalno grupo, lahko vpeljemo tudi specialno unitarno grupo. Pri tem se sklicujemo zgolj na determinanto, ki mora biti enaka 1 in ne na geometrijski pomen ohranjanja orientacije. Specialna unitarna grupa je definirana kot množica vseh unitarnih matrik z determinanto 1 z operacijo matričnega množenja. S simboli: $SU_n = \{A \in U_n : \det A = 1\}$.

Podgrupa edinka

Definicija 6.8 *Naj bo x element iz grupe G. Vsak element oblike $gxg^{-1} \in G$, kjer je $g \in G$, imenujemo konjugiran element elementa x.*

Trditev 6.9 *Naj bo H podgrupa grupe G. Za poljuben element $g \in G$ je gHg^{-1} podgrupa grupe G. Imenujemo jo podgrupa konjugirana grupi G.*

Dokaz. Preveriti moramo, da veljajo pogoji iz točke (b) trditve 6.2. Ker je H grupa vsebuje identični element, zato za poljuben $g \in G$ velja: $geg^{-1} = gg^{-1} = e$. Naj bosta gxg^{-1} in gyg^{-1} poljubna elementa iz gHg^{-1}. Potem je njun produkt enak: $gxg^{-1}gyg^{-1} = gxyg^{-1} \in gHg^{-1}$, saj je H grupa, zato vsebuje produkte vseh svojih elementov. Preverimo še inverze: naj bo $gxg^{-1} \in gHg^{-1}$. Potem je: $\left(gxg^{-1}\right)^{-1} = \left(g^{-1}\right)^{-1} x^{-1}g^{-1} = gx^{-1}g^{-1}$ in velja $gx^{-1}g^{-1} \in gHg^{-1}$, saj je $x^{-1} \in H$. □

Definicija 6.10 *Podgrupa N grupe G je edinka, če za vse $g \in G$ velja $gNg^{-1} = N$ (zapisano drugače: za vsak $g \in G$ in $n \in N$ velja: $gng^{-1} \in N$). Oznaka: $N \triangleleft G$.*

Opazimo, da sta podgrupi edinki v grupi G vedno grupa, ki vsebuje samo enoto, in celotna grupa G. Ti dve grupi imenujemo trivialni podgrupi edinki. Prav tako ni težko opaziti, da je vsaka podgrupa H Abelove grupe G edinka, saj za poljubna elementa $h \in H$ in $g \in G$ velja:$ghg^{-1} = gg^{-1}h = h \in H$. Ob koncu omenimo še definicijo enostavne grupe.

Definicija 6.11 *Netrivialno grupo G imenujemo enostavna, če ne vsebuje netrivialnih podgrup edink.*

V nadaljevanju se z enostavnimi grupami ne bomo poglobljeno ukvarjali. Definicija je namenjena zgolj temu, da bomo lahko ugotovili, ali so katere izmed obravnavanih matričnih grup enostavne.

6.2 Homomorfizem grup

Med grupami lahko definiramo tudi preslikave. Preslikavo med dvema grupama, ki ustreza pogoju da je slika produkta originalov enaka produktu slik, imenujemo homomorfizem grup. Podajmo definicijo še nekoliko bolj formalno.

Definicija 6.12 *Naj bosta $(G, \circ)$ in $(H, *)$ grupi. Homomorfizem iz grupe G v grupo H je preslikava $\phi : G \to H$, za katero za vsaka $x, y \in G$ velja: $\phi(x \circ y) = \phi(x) * \phi(y)$.*

Naslednjo trditev bomo uporabili za izpeljavo dveh pomembnih trditev o homomorfizmih, in sicer za to, da bomo pokazali, da homomorfizem enoto prve grupe preslika v enoto druge grupe in da se inverzni element iz prve grupe slika v inverzni element iz druge grupe. To trditev, dokažemo nekoliko bolj splošno.

Trditev 6.13 *Naj bosta $(G, \circ)$ in $(H, *)$ grupi in $\phi : G \to H$ homomorfizem. Potem za poljubna $x, y \in G$, veljata enakosti: $\phi\left(x \circ y^{-1}\right) = \phi(x) * \phi(y)^{-1}$ in $\phi\left(x^{-1} \circ y\right) = \phi(x)^{-1} * \phi(y)$.*

Dokaz. Dokazali bomo samo prvo enakost, druga se dokaže analogno. Izračunajmo:

$$\begin{aligned}
&\phi\left(x \circ y^{-1}\right) * \phi(y) = \phi\left(\left(x \circ y^{-1}\right) \circ y\right) = \phi\left(x \circ \left(y^{-1} \circ y\right)\right) = \phi(x) \qquad / * \phi(y)^{-1} \\
&\phi\left(x \circ y^{-1}\right) * \phi(y) * \phi(y)^{-1} = \phi(x) * \phi(y)^{-1} \\
&\phi\left(x \circ y^{-1}\right) = \phi(x) * \phi(y)^{-1}
\end{aligned}$$

Pri izračunu smo v prvi vrstici uporabili lastnost homomorfizma in asociativnost v grupi G. S tem je enakost dokazana. □

Z upoštevanjem prve zgornje enakosti dobimo rezultata, ki smo ju navedli pred trditvijo. Če je $x = y$, potem je $\phi\left(x \circ x^{-1}\right) = \phi(x) * \phi(x)^{-1} \Leftrightarrow \phi(e_G) = e_H$ in če je $x = e_G$ in $y = g$, potem je $\phi\left(g^{-1}\right) = \phi\left(e_G \circ g^{-1}\right) = \phi(e_G) * \phi(g)^{-1} = \phi(g)^{-1}$.

Oglejmo si, kako lahko na determinanto gledamo, kot na homomorfizem.

Zgled. Naj bosta dani grupi $GL_n(\mathbb{R})$ in $(\mathbb{R}^*, \cdot)$. Potem je z $\det : GL_n(\mathbb{R}) \to (\mathbb{R}^*, \cdot)$ podan homomorfizem.
Preveriti je potrebno, da velja pogoj iz definicije homomorfizma. Naj bosta $A, B \in GL_n(\mathbb{R})$. Potem po lastnosti determinante velja: $\det(AB) = \det A \cdot \det B$. S tem smo dokazali, da je preslikava det homomorfizem.

Homomorfizem $\phi : G \to H$, ki je surjektiven, to pomeni, da je vsak element iz grupe H slika nekega elementa iz grupe G, in injektiven, kar pomeni, da se različna elementa iz grupe G slikata v različna elementa iz grupe H, imenujemo bijektivni homomorfizem ali krajše izomorfizem. Če med dvema grupama obstaja izomorfizem, tedaj pravimo, da sta grupi izomorfni. To, da je grupa G izomorfna grupi H, označimo takole: $G \cong H$. V zvezi s homomorfizmi grup je potrebno izpostaviti še dva pojma, in sicer gre za jedro in sliko homomorfizma.

Definicija 6.14 *Naj bo $\phi : G \to H$ homomorfizem. Jedro homomorfizma ϕ je množica vseh elementov iz grupe G, ki jih homomorfizem ϕ slika v e_H. S simboli:* $\operatorname{Ker}\phi = \{g \in G : \phi(g) = e_H\}$. *Slika homomorfizma je množica vseh elementov iz grupe H, ki so slike nekih elementov iz grupe G:* $\operatorname{Im}\phi = \{h \in H : \ \phi(g) = h, \ \text{za } g \in G\}$.

Naslednji izrek združuje večino teoretičnih pojmov, ki so bili predstavljeni na predhodnih straneh.

Izrek 6.15 *Naj bo $\phi : G \to H$ homomorfizem. Potem veljata naslednji trditvi:*

(a) $\operatorname{Ker}\phi$ *je podgrupa edinka v G,*

(b) $\operatorname{Im}\phi$ *je podgrupa v grupi H.*

Dokaz. (a) Najprej preverimo, da je $\operatorname{Ker}\phi$ grupa. Po zgornjem razmisleku, ki je sledil trditvi 6.13 velja: $\phi(e_G) = e_H$ in zato je $e_H \in \operatorname{Ker}\phi$. Preveriti je potrebno še zaprtost za množenje in za inverze. Naj bosta $x, y \in \operatorname{Ker}\phi$. Potem je $\phi(xy) = \phi(x)\cdot\phi(y) = e_H \cdot e_H = e_H$, torej: $xy \in \operatorname{Ker}\phi$. Dalje velja: $\phi\left(x^{-1}\right) = \phi(x)^{-1} = e_H$ in zato: $x^{-1} \in \operatorname{Ker}\phi$. Pokažimo še, da je $\operatorname{Ker}\phi$ edinka. Naj bo $x \in \operatorname{Ker}\phi$. Potem za vsak $x \in G$ velja $\phi\left(gxg^{-1}\right) = \phi(g)\,\phi(x)\,\phi(g)^{-1} = e_H$, kar pomeni, da je grupa $\operatorname{Ker}\phi$ podgrupa edinka v grupi G.
(b) Vemo, da je $e_H \in \operatorname{Im}\phi$. Če sta $y_1, y_2 \in \operatorname{Im}\phi$, potem obstajata taka $x_1, x_2 \in G$, da velja $y_1 = \phi(x_1)$ in $y_2 = \phi(x_2)$. Izračunajmo: $y_1 y_2 = \phi(x_1)\,\phi(x_2) = \phi(x_1 x_2)$ in zato je $y_1 y_2 \in \operatorname{Im}\phi$. Ker je $y_1 = \phi(x_1)$, je tudi $y_1^{-1} = \phi(x_1)^{-1} = \phi\left(x_1^{-1}\right)$. Torej je $\operatorname{Im}\phi$ podgrupa grupe H. □

Uporabimo dokazan izrek. V prejšnjem zgledu smo preverili, da determinanta zadošča pogojem homomorfizma med grupama $GL_n(\mathbb{R})$ in $(\mathbb{R}^*, \cdot)$. Determinanta je prav tako homomorfizem med grupama O_n in $(\mathbb{R}^*, \cdot)$ ter med grupama U_n in $(\mathbb{C}^*, \cdot)$. Jedro prvega homomorfizma so vse matrike iz grupe O_n, ki imajo determinanto enako 1. Zapisali smo, da te matrike tvorijo grupo SO_n, zato je SO_n podgrupa edinka grupe O_n. Podoben razmislek nas pripelje tudi do tega, da je grupa SU_n podgrupa edinka grupe U_n. Oboje lahko povzamemo v naslednji trditvi.

Trditev 6.16 *Grupa SO_n je podgrupa edinka grupe O_n in grupa SU_n je podgrupa edinka grupe U_n.*

Če zapisano pogledamo v luči definicije enostavne grupe, lahko zaključimo, da ortogonalna in unitarna grupa nista enostavni.

6.3 Center grupe

Cilj zadnjega podpoglavja je poiskati največjo podmnožico grup O_n in U_n, ki bo vsebovala take matrike, ki komutirajo z vsemi preostalimi matrikami v grupi. Definirajmo to množico.

Definicija 6.17 *Množico elementov v grupi G, ki komutriajo z vsemi elementi iz grupe, imenujemo center grupe: $Z(G) = \{g \in G : gx = xg,\ \forall x \in G\}$.*

Center grupe ni nikoli prazna množica, saj vedno vsebuje enoto grupe G. To dejstvo izhaja iz definicije grupe, ki pravi, da za enoto e grupe G za vse $g \in G$ velja $eg = ge$. O centru pa lahko povemo še več:

Trditev 6.18 *Center grupe G je podgrupa grupe G, še več, je tudi podgrupa edinka grupe G.*

Dokaz. Najprej dokažimo, da je $Z(G) \leq G$. Da je center neprazna množica smo že dokazali. Naj bosta $x, y \in Z(G)$. Potem za vsak $g \in G$ velja: $xg = gx$ in $yg = gy$. Preverimo zaprtost za množenje: $gxy = xgy = xyg$, torej velja $xy \in Z(G)$. Kjer smo uporabili predpostavki. Zaprtost za inverze: če enakost $gx = xg$ pomnožimo z leve in desne z x^{-1}, ta element je gotovo v G, dobimo: $x^{-1}g = x^{-1}g$ in zato velja, da je $x^{-1} \in Z(G)$. Preostane še dokaz, da je $Z(G)$ podgrupa edinka. Dokazati moramo, da je $gxg^{-1} \in Z(G)$ za neka $g \in G$ in $x \in Z(G)$. Naj bosta $g \in G$ in $x \in Z(G)$ poljubna. Potem je: $gxg^{-1} = xgg^{-1} = xe = x \in Z(G)$. S tem smo dokazali, da je $Z(G)$ podgrupa edinka v grupi G. □

Preden se bomo bolj podrobno ukvarjali s centri matričnih grup, obravnavajmo še nekaj posebnih primerov matričnih grup. V grupi O_1 so matrike dimenzije 1×1. Te matrike identificiramo z realnimi števili. Za realna števila vemo, da je množenje komutativno, zato je grupa O_1 Abelova in velja, da je $Z(O_1) = O_1$. Podoben je razmislek pri grupi U_1. Elemente te grupe identificiramo s kompleksnimi števili, ker je množenje le-teh komutativno, je tudi center grupe U_1 enak tej grupi. Glede specialnih grup je zadeva podobna. Za grupi SO_1 in SU_1 pa velja: $SO_1 = SU_1 = \{1\}$. To je trivialna grupa, ki vsebuje zgolj enoto. Ta grupa je Abelova. Glede specialnih grup velja tudi naslednja trditev.

Trditev 6.19 *Grupa SO_2 je Abelova.*

Dokaz. V zgledu za trditvijo 6.7 smo navedli, da je:

$$SO_2 = \left\{ \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} : \phi \in [0, 2\pi) \right\}.$$

Izberimo $\phi,\ \theta \in [0, 2\pi)$. Izračunajmo:

$$\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(\phi+\theta) & -\sin(\phi+\theta) \\ \sin(\phi+\theta) & \cos(\phi+\theta) \end{bmatrix}$$
$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}$$

Pri tem smo pri izračunu uporabili adicijske izreke za funkciji kosinus in sinus. Ker je seštevanje realnih števil komutativno, sta dobljeni matriki enaki, zato velja, da je grupa SO_2 Abelova. □

Posledica gornje trditve je iz vidika centra grupe očitna, in sicer velja, da je $Z(SO_2) = SO_2$. Podatke o centrih grup višjega reda podaja naslednji izrek.

Izrek 6.20 *Naj bo $n \geq 2$. Potem za grupe O_n, SO_n, U_n in SU_n velja naslednje:*

(a) Center ortogonalne grupe je: $Z(O_n) = \{I_n, -I_n\}$.

(b) Center specialne ortogonalne grupe je: $$Z(SO_n) = \begin{cases} SO_2, & n = 2, \\ \{I_n, -I_n\}, & n \text{ sodo število}, \\ \{I_n\}, & n \text{ liho število}. \end{cases}$$

(c) Center unitarne grupe je: $Z(U_n) = \{e^{i\phi} I_n | \phi \in [0, 2\pi)\}$.

(d) Center specialne unitarne grupe je: $Z(SU_n) = \{e^{i\phi} I_n | \phi = \frac{2k\pi}{n},\ k = 0, 1, \ldots, n-1\}$.

Dokaz. (a) Naj matrika $A \in M_n(\mathbb{R})$ komutira z vsemi elementi iz grupe O_n. To pomeni, da matrika komutira tudi z elementarnimi matrikami $E_i(-1) \in M_n(\mathbb{R})$ za vsak $i = 1, \ldots, n$ in P_{ij} za vsaka $i, j = 1, \ldots, n$. Matrika $E_i(-1)$ predstavlja matriko, ki smo jo dobili iz identične matrike tako, da smo i-to vrstico identične matrike pomnožili z -1. Matriko $P_{ij} \in M_n(\mathbb{R})$ pa smo dobili iz identične matrike s tem, ko smo zamenjali i-to in j-to vrstico. Dokažimo, da se matrike $E_i(-1)$ in P_{ij} nahajajo v O_n. Determinanta matrike $E_i(-1)$ enaka produktu diagonalnih elementov matrike $E_i(-1)$. Diagonalni elementi te matrike so enice, teh je $n-1$, in -1. Produkt teh števil je enak -1. Pri matriki P_{ij} smo zamenjali i-to in j-to vrstico identične matrike, ki ima determinanto 1, zato je determinanta matrike P_{ij} enaka -1. Ker velja, da se v vsakem stolpcu in v vsaki vrstici matrike pojavi samo eno število 1 ali -1, stolpci/vrstice tvorijo ortonormirano bazo prostora $\mathbb{R}^n$, zato oba tipa matrik ustrezata pogoju, da sta ortogonalni. Ker sta ortogonalni, sta tudi obrnljivi. Inverza teh dveh matrik sta matriki sami.
Omenimo še to, da če matriko $E_i(-1)$ primnožimo k matriki A z leve to pomeni, da se

bo i-ta vrstica matrike A pomnožila z -1, če primnožimo matriko $E_i(-1)$ k matriki A z desne pa se bo z -1 pomnožil i-ti stolpec. Podobno velja tudi za matriko P_{ij}, in sicer, če jo k matriki A primnožimo z leve se bosta zamenjali i-ta in j-ta vrstica matrike A, če jo primnožimo z desne pa i-ti in j-ti stolpec. (Dokaze uporabljenih lastnosti elementarnih matrik je mogoče najti v [2] na straneh 89–91.)
Ker smo predpostavili, da matrika A komutira z vsemi matrikami iz grupe O_n velja:

$$AE_i(-1) = E_i(-1)A \Leftrightarrow A = E_i(-1)AE_i\,(-1)^{-1} \Leftrightarrow A \overset{(1)}{=} E_i(-1)AE_i(-1)$$
$$AP_{ij} = P_{ij}A \Leftrightarrow A = P_{ij}AP_{ij}^{-1} \Leftrightarrow A \overset{(2)}{=} P_{ij}AP_{ij}$$

Najprej razpišimo enakost (1):

$$\begin{bmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ii} & \dots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & -a_{1i} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ -a_{i1} & \dots & a_{ii} & \dots & -a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & -a_{ni} & \dots & a_{nn} \end{bmatrix}$$

Ker sta i-ta vrstica in i-ti stolpec matrike A pomnožena z -1, mora veljati $a_{ij} = 0$ za vsak $i \neq j$. Ker je bil i poljubno izbran, velja, da so vsi nediagonalni elementi matrike A enaki 0.
Iz enakosti (2) sledi, da morajo biti vsi elementi na glavni diagonali matrike enaki. Argumentacija je podobna kot zgoraj. Ker je matrika A ortogonalna pa mora veljati še to, da je $\|a_i\| = |a_{ii}| = 1$ za vsak $i = 1, \dots, n$. Zato so vsi elementi na diagonali matrike enaki bodisi 1 bodisi -1. Od tod sledi sklep: $Z(O_n) = \{I_n, -I_n\}$.

(b) Center grupe SO_n bomo določili podobno, kot smo določili center grupe O_2. Prej vpeljani matriki $E_i(-1) \in M_n(\mathbb{R})$ in $P_{ij} \in M_n(\mathbb{R})$ imata determinanto enako -1, zato matriki nista iz grupe SO_n. Oglejmo si, kaj se zgodi, če obe matriki pomnožimo. Potem dobimo matriko $B = P_{ij} \cdot E_i(-1)$. Ta matrika ima determinanto enako 1, hkrati pa velja, da ima matrika ortonormirane vrstice/stolpce, saj ima v vsaki vrstici natanko en element, ki je enak 1 ali -1, enako velja tudi za stolpce. Torej je matrika B iz grupe SO_n. Po pravilu za inverz produkta matrik vemo, da je $B^{-1} = (E_i(-1) \cdot P_{ij})^{-1} = E_i^{-1}(-1) \cdot P_{ij}^{-1} = E_i(-1) \cdot P_{ij}$. Naj bo matrika $A \in Z\,(SO_n)$. Potem je $AB = BA$. Če enakost pomnožimo z desne z inverzom matrike B, dobimo: $A = BAB^{-1} = E_i(-1)P_{ij}AE_i(-1)P_{ij}$. Razpišimo dano

enakost:

$$\begin{bmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \dots & \vdots \\ a_{i1} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \dots & \vdots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & -a_{1i} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \dots & \vdots \\ a_{j1} & \dots & a_{jj} & \dots & -a_{ji} & \dots & a_{jn} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \dots & \vdots \\ -a_{i1} & \dots & -a_{ij} & \dots & a_{ii} & \dots & -a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & -a_{ni} & \dots & a_{nn} \end{bmatrix}$$

Ker mora enakost veljati za vsaka $1 \leq i, j \leq n$ opazimo, da je $a_{ij} = 0$ za vsak $i \neq j$. Hkrati velja tudi, da je $a_{ii} = a_{jj}$ za vsak $i, j \in \{1, \dots, n\}$. To pomeni, da so edini neničelni elementi matrike A tisti, ki se nahajajo na glavni diagonali, prav tako pa so vsi elementi na glavni diagonali enaki.

Ker je matrika $A \in SO_n$, mora biti determinanta matrike A enaka 1. Označimo neničelni diagonalni element matrike A z a. Potem velja: $\det A = a^n = 1$. Ločimo dve možnosti: (i) če je $n = 2k$, $k \in \mathbb{N}\backslash\{1\}$ sodo število, potem je $\left(a^{2k} - 1\right) = \left(a^k - 1\right)\left(a^k + 1\right) = 0$. Od tod sledi, da je $a_1 = 1$ in $a_2 = -1$. Torej, če je n sodo število, različno od 2, center grupe SO_n vsebuje matriki I_n in $-I_n$. (ii) Če je $n = 2k + 1$, $n \in \mathbb{N}$ pa velja $\left(a^{2k+1} - 1\right) = (a - 1)\left(a^{k-1} + a^{k-2} + \ldots + a + 1\right) = 0$ Od tod sledi, da je edina realna rešitev gornje enačbe: $a = 1$ in zato je edina matrika v centru grupe SO_n, ko je n liho število, ki ni enako 1, matrika I_n.

Da je $Z(SO_2) = SO_2$, smo razmislili pred tem izrekom. Ugotovitve lahko združimo:

$$Z(SO_n) = \begin{cases} SO_2, & n = 2, \\ \{I_n, -I_n\}, & n \text{ sodo } \text{število}, \\ \{I_n\}, & n \text{ liho } \text{število}. \end{cases}$$

(c) Naj bo $B \in U_n(\mathbb{C})$ matrika, ki se nahaja v centru grupe U_n. Ker je matrika B unitarna, velja po lastnosti (c) normalnih matrik, da je tudi normalna. Dalje pa po spektralnem izreku za normalne matrike (izrek 5.17) velja, da obstajata taka unitarna matrika V in diagonalna matrika T, da je $B = VTV^*$. Če enakost z leve pomnožimo z V^* in z desne z V dobimo: $V^*BV = T$, kar pomeni, da je matrika V^*BV diagonalna.

Ker smo predpostavili, da je matrika B iz centra grupe U_n, za vsako matriko $U \in U_n$ velja $UB = BU$. Ker je matrika V unitarna, torej $V \in U_n$, prejšnja enakost velja tudi za matriko V, zato lahko pišemo: $VB = BV$, od koder dobimo: $B = V^*BV$. Če to ugotovitev združimo z ugotovitvijo iz konca prejšnjega odstavka, dobimo $B = T$. Torej je matrika B diagonalna.

Pokažimo, da mora biti matrika skalarna. Predpostavimo, da za matriko B iz centra grupe

U_n velja: $b_{ii} \neq b_{jj}$, za neka $i, j \in \{1, \ldots, n\}$. Naj bo matrika P_{ij} takšna, kot smo jo definirali v točki (a) dokaza tega izreka. Matrika P_{ij} je element grupe U_n, zato mora veljati: $P_{ij}B = BP_{ij}$. Matrika P_{ij} je obrnljiva in je sama sebi inverz. Veljati mora, da je matrika $C := P_{ij}BP_{ij}$ enaka matriki B. Zaradi predpostavke $b_{ii} \neq b_{jj}$, dobimo: $c_{ii} = b_{jj}$ in $c_{jj} = b_{ii}$ ter $b_{kk} = c_{kk}$, za $k \neq i, j$. Da bo matrika B iz centra grupe, mora biti $b_{ii} = b_{jj}$, kar je v nasprotju s predpostavko o različnosti elementov. Od tod sledi, da je matrika B skalarna. Naj bo $b_i = [0 \ \ldots \ b_{ii} \ \ldots \ 0\ ]^T$ i-ti stolpec matrike B. Ker je $B \in U_n$, stolpci matrike B tvorijo ortonormirano bazo. Velja: $\|b_i\| = |b_{ii}| = 1$. To pomeni, da so diagonalni elementi matrike B oblike: $e^{i\phi}$, za nek $\phi \in [0, 2\pi)$.
Dokazali smo, da je $Z(U_n) = \{e^{i\phi}I_n | \phi \in [0, 2\pi)\}$.

(d) Dejstvo, da je matrika $A \in SU_n$ iz centra grupe SU_n skalarna matrika, dokažemo analogno dokazu točke (b) tega izreka. Matrika $P_{ij} \cdot E_i(-1)$ je element grupe SU_n, saj so njene vrstice/stolpci ortogonalni vektorji, determinanta te matrike pa je enaka 1. Preostali argumenti, ki smo jih uporabili za dokaz, da je matrika A skalarna, so enaki kot zgoraj. Nadaljevanje razmisleka pa je prilagojeno dejstvu, da se ukvarjamo z matrikami s kompleksnimi elementi.
Denimo, da je matrika λI_n, kjer je $\lambda \in \mathbb{C}\backslash\{0\}$, element centra specialne unitarne grupe. Zato mora veljati, da je $\det(\lambda) = \lambda^n = 1$. Od tod dobimo enačbo: $\lambda^n = e^{2k\pi i}$, za $k \in \mathbb{Z}$. Rešitev te enačbe je $\lambda = e^{\frac{2k\pi}{n}i}$, pri čemer dobimo različne rešitve zgolj za $k \in \{0,\ 1,\ \ldots,\ n-1\}$.
Naj bo $\phi = \frac{2k\pi}{n}$, za $k \in \{0, 1, \ldots, n-1\}$, potem je center specialne unitarne grupe enak: $Z(SU_n) = \{e^{i\phi}I_n | \phi = \frac{2k\pi}{n}, k = 0, 1, \ldots, n-1\}$ □

Ob koncu se ozrimo le še na centra splošne linearne grupe in specialne linearne grupe z realnimi elementi. Izkaže se, da je potrebno točki (a) in (b) le malo spremeniti, da dobimo centra teh dveh grup. Center grupe $GL_n(\mathbb{R})$ predstavljajo skalarne matrike, se pravi matrike oblike λI_n, kjer je $\lambda \in \mathbb{R}\backslash\{0\}$. Dokaz tega dejstva poteka podobno kot dokaz točke (a). Matrike P_{ij} in $E_i(-1)$ so vsebovane tudi v grupi $GL_n(\mathbb{R})$, zato morajo elementi iz centra te grupe komutirati tudi s temi elementi. Od tod sledi, da je matrika, ki se nahaja v centru grupe $GL_n(\mathbb{R})$ zagotovo diagonalna matrika, ki ima na glavni diagonali same enake elemente. Ker je edini pogoj za determinanto matrike iz $GL_n(\mathbb{R})$ ta, da je determinanta različna od 0, lahko za diagonalne elemente matrike iz $GL_n(\mathbb{R})$ izberemo poljubno od 0 različno realno število. Torej so v centru grupe $GL_n(\mathbb{R})$ res skalarne matrike.
Pri specialni linearni grupi je začetni del razmisleka povsem enak, kot smo ga navedli v dokazu točke (b). To pomeni, da so edini kandidati za center grupe $SL_n(\mathbb{R})$ skalarne matrike, ki imajo determinanto enako 1. Zato podobno velja:

$$Z(SL_n(\mathbb{R})) = \begin{cases} \{I_n, -I_n\}, & n \text{ sodo število}, \\ \{I_n\}, & n \text{ liho število}. \end{cases}$$

Edina razlika glede centra med specialno linearno grupo $SL_n(\mathbb{R})$ in specialno ortogonalno grupo SO_n je v tem, da v primeru grupe $SL_2(\mathbb{R})$ ne moremo govoriti o komutativnosti, saj za poljubni 2×2 matriki z realnimi koeficienti in determinanto 1 ne velja, da sta komutativni. Na primer matriki:

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \quad \text{in} \quad B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

nista komutativni, saj je $AB = \begin{bmatrix} 4 & 3 \\ 5 & 4 \end{bmatrix}$, po drugi strani pa je $BA = \begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix}$.

Literatura

[1] H. Anton, C. Rorres, *Elementary Linear Algebra: applications version*, Wiley, Hoboken, 2014.

[2] D. Benkovič, *Vektorji in matrike*, Univerza v Mariboru: Fakulteta za naravoslovje in matematiko, Maribor, 2014.

[3] M. Brešar, *Uvod v algebro*, DMFA - založništvo, Ljubljana, 2018.

[4] J. Ding, A. Zhou, Eigenvalues of rank-one updated matrices with some applications, Applied Mathematics Letters 20 (2007) str. 1223–1226.

[5] D. Ž. Đoković, C. R. Johnson Unitarily achievable zero patterns and traces of words in A and A^*, Linear Algebra and its Applications (2007) str. 63–68.

[6] R. A. Horn, C. R. Johnson, *Matrix Analysis, Second Edition*, Cambridge University Press, Cambridge, 2013.

[7] J. F. Humphreys, *A Course in Group Theory*, Oxford University Press, Oxford, 2001.

[8] T. Košir, Linearna algebra za študente praktične matematike (online). (citirano 1. 12. 2019). Dostopno na naslovu: `https://www.fmf.uni-lj.si/ kosir/poucevanje/skripta/`.

[9] R. Larson, *Elementary Linear Algebra, Seventh Edition*, Brooks Cole Cengage Learning, Boston, 2013.

[10] J. Stillwell, *Naive Lee Theory*, Springer, New York, 2008.

[11] Mathematics Genealogy Project (online). (citirano 6. 1. 2020). Dostopno na naslovu: `https://www.genealogy.math.ndsu.nodak.edu/id.php?id=9179`.