FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
UNIVERZITY KOMENSKÉHO V BRATISLAVE

DIPLOMOVÁ PRÁCA

BRATISLAVA 2009 RADOVAN DOBÁK

Využitie blokových matíc pri riešení úloh lineárnej algebry

DIPLOMOVÁ PRÁCA

Radovan Dobák

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
KATEDRA APLIKOVANEJ MATEMATIKY A ŠTATISTIKY

9.1.9 Aplikovaná matematika
Ekonomická a finančná matematika

Vedúci diplomovej práce:
RNDr. Dušan Krajčovič, CSc.

BRATISLAVA 2009

Zadanie diplomovej práce

<u>Vedúci diplomovej práce:</u> RNDr. Dušan Krajčovič, CSc.

<u>Diplomant:</u> Radovan Dobák

<u>Názov diplomovej práce:</u> Využitie blokových matíc pri riešení úloh lineárnej algebry

<u>Popis zadania :</u>

V praxi sa väčšinou stretávame s maticami veľkých rozmerov, rádovo aj niekoľko miliónov, čo častokrát môže spôsobiť preplnenie pamäte počítača a pre úspešné riešenie je potrebné tok dát špeciálne organizovať. Jednou z takých možností organizácie je využitie blokových matíc.

Cieľom tejto diplomovej práce je popísať niektoré špeciálne metódy výpočtu vlastných čísel matice, ako aj riešenia systému lineárnych rovníc pomocou blokových operácií. Konkrétne ide o blokové schémy Householderovej metódy a tzv. blokovej Sylvestrovej rovnice, ktoré sa používajú na transformáciu matíc na blokovo diagonálny tvar.

Prehlasujem, že som túto diplomovú prácu vypracoval samostatne. Uvádzam všetky literárne pramene a publikácie, ktoré som použil.

Radovan Dobák

V Bratislave, 9.4.2009

Ďakujem vedúcemu mojej diplomovej práce,
RNDr. Dušanovi Krajčovičovi, CSc.
za odborné vedenie , pripomienky,
trpezlivosť a poskytnutú literatúru.

Abstrakt

Radovan Dobák; Využitie blokových matíc pri riešení úloh lineárnej algebry, Diplomová práca; Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; Vedúci diplomovej práce: RNDr. Dušan Krajčovič, CSc.; Bratislava, 2009; 61 strán

V práci sa zaoberáme využitím blokových matíc na zjednodušenie počítania s veľkými maticami. Na blokové zjednodušenie matice sme si zvolili dve metódy. Prvou metódou je bloková Householderova metóda pomocou ktorej vieme zjednodušiť štvorcovú symetrickú maticu na blokovo tridiagonálny tvar. Druhá metóda používa Sylvestrove rovnice na upravenie matice na hornú, alebo dolnú trojuholníkovú maticu. Najprv vynulujeme bloky štvorcovej matice pomocou blokovej Householderovej metódy a následne použijeme metódu so Sylvestrovými rovnicami. Celý postup nám zabezpečí blokovú diagonalizáciu matice. Využitie prvej metódy vidíme pri počítaní vlastných čísel a druhej pri zjednodušení počítania lineárnych systémov.

Kľúčové slová: Householderová metóda, Sylvestrove rovnice, blokové matice
vlastné čísla, lineárny systém

Obsah

Zoznam skratiek a symbolov

$\mathbb{C}$, $\mathbb{R}$, $\mathbb{N}$	množina komplexných, reálnych a prirodzených čísiel
u	vektor u
A	matica A
a_{ij}	ij-ty prvok matice A
A_{ij}	ij-ty blok matice A
A^T	transponovaná matica A
A^{-1}	inverzná matica k matici A
A^*	transponovaná a komplexne združená matica k matici A
$u^\perp$	kolmý vektor na vektor u
$\|u\|$	norma vektora u
λ	vlastné číslo matice
σ	singulárna hodnota matice
$\det(A)$	determinant matice A
$\mathrm{diag}(a_1 \dots a_n)$	hodnoty $a_1 \dots a_n$ na diagonále matice A
$\mathrm{hod}(A)$	hodnosť matice A
$\mathrm{cond}(A)$	číslo podmienenosti matice A

Zoznam obrázkov

Úvod

V lineárnej algebre a jej aplikáciách sa často stretávame s maticami. Tieto matice môžu byť menších, ale aj väčších rozmerov. Väčšie rozmery si môžeme predstaviť rádovo v tisíckach stĺpcov a riadkov. Pri počítaní takýchto matíc používame samozrejme počítače a príslušné programy, ale aj tak hlavne pri väčších maticiach sa nám môže stať, že sa preplní pamäť počítača. Vtedy nastáva problém ako organizovať tok údajov vstupujúcich do výpočtu. Jedným z riešení organizácie toku údajov je pomocou blokových schém. Táto práca je venovaná blokovým maticiam a algoritmom s blokovými schémami.

Práca je rozdelená na dve časti. V prvej časti chceme ukázať, ako pomocou blokovej Householderovej metódy vieme upraviť tvar veľkej štvorcovej symetrickej matice na blokový tridiagonálny tvar. Keďže táto metóda nemení vlastné čísla, tak jej využitím može byť zjednodušenie tvaru matice pred samotným výpočtom vlastných čísel. V druhej časti sme použili dve blokové metódy na zjednodušenie tvaru matice na blokový diagonálny tvar. Prvou metódou je už spomínaná bloková Householderová metóda, ktorá nám pretransformuje veľkú štvorcovú maticu na hornú blokovú trojuholníkovú maticu. Túto maticu následne pretransformujeme pomocou Sylvestrových rovníc na blokovú diagonálnu maticu. Výsledná bloková diagonálna matica má nezávislé bloky, takže využitie tohto postupu by sa dalo uplatniť pri počítaní lineárnych systémov. Obe metódy sme aj naprogramovali a overili ich funkčnosť na konkrétnych príkladoch.

Diplomová práca je rozdelená na teoretickú časť s tromi kapitolami a praktickú časť s dvomi kapitolami. V prvej kapitole uvádzame základné pojmy, ktoré využívame v celej práci. Taktiež tu uvádzame základné oboznámenie sa s vlastnými číslami a lineárnym systémom. V druhej kapitole popisujeme Sylvestrove rovnice, ich využitie a tiež metódu, ktorú použijeme na výpočet. Tretia kapitola je venovaná Householderovej metóde, jej využitiu a popísanie postupu blokovej Householderovej metódy. Štvrtá kapitola už patrí do praktickej časti a popisujeme v nej zjednodušenie tvarov matíc pomocou dvoch spomínaných metód. Popisujeme v nej aj programy, ktoré sme použili na demonštráciu funkčnosti našich postupov. V poslednej, piatej kapitole sa nachádzajú výpisy z programov, ktoré ukazujú transformáciu na blokovú tridiagonálnu a blokovú diagonálnu maticu.

Teoretická časť

1 Základné pojmy

V úvodnej kapitole sú uvedené rôzne základné pojmy z maticového počtu a lineárnej algebry, ktoré sú využité v nasledujúcich kapitolách.

1.1 Triedy matíc

Maticou A typu $m \times n$ ($m, n \in \mathbb{N}$) je vlastne tabuľka čísel

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \tag{1.1}$$

kde $a_{ij} \in \mathbb{R}$ ($i = 1, \dots, m$, $j = 1, \dots, n$) sú reálne prvky matice. V tejto práci sa zameriame práve na reálne matice a ich využitie. Komplexné matice nebudeme uvažovať.

Transpozícia matice je vymenenie prvkov matice. Meníme súradnicu riadku za súradnicu stĺpca a naopak

$$a_{ij}^T = a_{ji} \, .$$

Ak uvažujeme transpozíciu matice (1.1), potom

$$A^T = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}.$$

Identická matica I je štvorcová matica $n \times n$ ($n \in \mathbb{N}$) kde na diagonále sú jednotky a ostatné prvky sú nulové

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

Ďalej si popíšeme rôzne tvary matíc a operácie s nimi.

1.1.1 Blokové matice

Pod pojmom bloková matica si predstavujeme maticu A, napr. (1.1), ktorú rozdelíme na bloky A_{ij}

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

kde blok matice A_{ij} je tvorený prvkami matice (1.1) , čiže

$$A_{11} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,q} \\ \vdots & \ddots & \vdots \\ a_{p,1} & \cdots & a_{p,q} \end{bmatrix}, A_{12} = \begin{bmatrix} a_{1,q+1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{p,q+1} & \cdots & a_{p,m} \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} a_{p+1,1} & \cdots & a_{p+1,q} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,q} \end{bmatrix}, A_{22} = \begin{bmatrix} a_{p+1,q+1} & \cdots & a_{p+1,m} \\ \vdots & \ddots & \vdots \\ a_{n,q+1} & \cdots & a_{n,m} \end{bmatrix}.$$

Matica nemusí byť vždy rozdelená len symetricky. Bloky matice môžu byť aj nesymetrické, nesmú však obsahovať rovnaký prvok.

1.1.2 Symetrické matice

Symetrická matica A je štvorcová $n \times n$ ($n \in \mathbb{N}$) matica a má prvky $a_{ij} \in \mathbb{R}$ symetrické podľa diagonály. To znamená

$$a_{ij} = a_{ji}\,, \qquad\qquad (i \neq j\,, i,j \in \mathbb{N})$$

a na diagonále sa nachádzajú ľubovoľné prvky. Rovnako definujeme aj blokovo symetrickú maticu.

1.1.3 Diagonálne matice

Diagonálna matica A je štvorcová $n \times n$ ($n \in \mathbb{N}$) matica, ktorá má nenulové prvky a_{ii} ($a_{ii} \in \mathbb{R}$, $i \in \mathbb{N}$) na hlavnej diagonále

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}.$$

Menej sa vyskytujú aj obdĺžnikové diagonálne matice, ktoré majú tvar

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \end{bmatrix}, \text{ alebo } A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \\ 0 & 0 \end{bmatrix}, \quad (a_{ii} \in \mathbb{R}, i \in \mathbb{N}).$$

Analogicky vyzerajú blokové diagonálne matice. S obmenou, že namiesto prvkov na diagonále sú bloky na diagonále

$$A = \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{bmatrix}, \qquad \text{(prvky } A_{ii} \in \mathbb{R}, i \in \mathbb{N}).$$

Stretneme sa aj s bidiagonálnymi, alebo trojdiagonálnymi maticami. Bidiagonálna matica je diagonálna matica, ktorá má navyše ešte nenulové horné, alebo dolné susedné prvky (bloky) k diagonále

$$\begin{bmatrix} \times & \times & & \\ & \times & \times & \\ & & \times & \times \\ & & & \times \end{bmatrix}, \begin{bmatrix} \times & & & \\ \times & \times & & \\ & \times & \times & \\ & & \times & \times \end{bmatrix}.$$

Ak sú nenulové horné aj dolné susedné prvky (bloky) k diagonále, potom hovoríme o tridiagonálnom tvare matice

$$\begin{bmatrix} \times & \times & & \\ \times & \times & \times & \\ & \times & \times & \times \\ & & \times & \times \end{bmatrix}.$$

1.1.4 Trojuholníkové matice

Trojuholníkové matice sú také, že všetky prvky alebo bloky pod diagonálou sú nulové. Vtedy hovoríme, že je to horná trojuholníková matica

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{12} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix},$$

alebo naopak, všetky prvky alebo bloky nad diagonálou sú nulové. Takáto matica sa nazýva dolná trojuholníková matica

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

1.1.5 Inverzné matice

Inverzná matica A^{-1} k štvorcovej $n \times n$ ($n \in \mathbb{N}$) matici A je taká, ktorá spĺňa

$$A^{-1}A = AA^{-1} = I \,,$$

kde I je identická $n \times n$ matica. Aby k matici A existovala inverzná matica, musí byť determinant matice A nenulový.

Matica ku ktorej existuje inverzná matica sa nazýva regulárna matica.

1.1.6 Ortogonálne matice

Hovoríme, že $n \times n$ ($n \in \mathbb{N}$) matica Q je ortogonálna [1], ak $QQ^T = I$. Táto rovnica hovorí, že matica Q má inverznú maticu a platí $Q^T = Q^{-1}$. Z týchto dvoch vzťahov platí $Q^TQ = I$. Nasledujúce vzťahy sú podľa predchádzajúcich ekvivalentné a môžeme ich považovať ako definíciu ortogonálnej matice

$$QQ^T = I \qquad Q^TQ = I \qquad Q^T = Q^{-1} \,.$$

Determinant ortogonánej matice $\det(Q) = \pm 1$.

1.2 Systém lineárnych rovníc

Jednoducho popíšeme princíp počítania systému lineárnych rovníc [10]. Lineárna rovnica s n neznámymi má tvar

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \,,$$

kde a_i a b sú dané, x_i ($i = 1, \dots, n$) je neznáma, ktorej riešenie hľadáme. Zložením niekoľkých takýchto rovníc dostávame systém lineárnych rovníc.

Predpokladajme teda, že máme n takýchto rovníc

$$
\begin{gathered}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \;, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \;, \\
\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \;.
\end{gathered}
$$

Koeficienty a_{ij} a b_i $(i, j = 1, \dots, n)$ sú zadané väčšinou ako reálne čísla. Hľadáme riešenia $x_1, \dots, x_n$ ktoré spĺňajú tento systém. Pre zjednodušenie zápisu, môžeme systém prepísať ako

$$
Ax = b \;, \tag{1.2}
$$

kde

$$
A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \;,\; x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \;,\; b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} .
$$

Pomocou maticového počtu si vieme riešenie takejto rovnice zjednodušiť. Použijeme inverznú maticu k matici A a dostávame

$$
x = A^{-1}b \;.
$$

Veta 1.1: Nech matica A je štvorcová. Tento systém je riešiteľný, ak spĺňa tieto ekvivalentné podmienky:

(a) A^{-1} existuje.
(b) Neexistuje nenulové y také, že $Ay = 0$.
(c) Stĺpce matice A sú lineárne nezávislé.
(d) Riadky matice A sú lineárne nezávislé.
(e) $\det(A) \neq 0$.
(f) Pre daný vektor b, existuje práve taký vektor x, že $Ax = b$.

Ak sú splnené všetky tieto podmienky, potom matica A je regulárna, invertibilná a existuje práve jedno riešenie sústavy (1.2). Ak nie sú splnené tieto podmienky, tak matica A je singulárna a riešenie (1.2) neexistuje, alebo je riešení nekonečne veľa.

Využitie takéhoto systému lineárnych rovníc je široké v rôznych oblastiach a tak môžu vznikať veľmi veľké matice A s vysokou náročnosťou výpočtu. Preto sa v jednej z častí tejto práce zaoberáme zjednodušením matice A a to pomocou blokovej diagonalizácie matice A. Takéto zjednodušenie by malo následne uľahčiť výpočet rovnice typu (1.2).

1.3 Vlastné čísla

Vlastné čísla, alebo charakteristické čísla $n \times n$ ($n \in \mathbb{N}$) matice A označujeme λ [10]. Vypočítame ich pomocou rovnice

$$Av = \lambda v \, , \tag{1.3}$$

kde A je známa matica, skalár $\lambda \in \mathbb{R}$ je vlastné číslo matice A a vektor $v \in \mathbb{R}^n$ je vlastný vektor matice A. Dvojicu (λ, v) nazývame charakteristickou dvojicou matice A. Vlastný vektor v má tú vlastnosť, že vždy prislúcha práve k jednej vlastnej hodnote λ. Avšak naopak to neplatí, pretože jedno vlastné číslo môže prislúchať k viacerým vlastným vektorom. Našim cieľom je teraz vypočítať len vlastné čísla, bez použitia vlastných vektorov. Upravíme rovnicu (1.3) ako

$$(A - \lambda I)v = 0 \, , \tag{1.4}$$

kde I je identická matica, rovnako veľká ako matica A. Predpokladajme, že vlastný vektor v je nenulový, čo je splnené, pretože ak v je vlastný vektor s prislúchajúcou vlastnou hodnotou λ, potom v je nenulové riešenie rovnice (1.4). Z toho dôvodu podľa Vety 1.1 časti (b) je matica $(A - \lambda I)$ singulárna a zároveň podľa časti (e) tej istej Vety musí byť $\det(A - \lambda I) = 0$. Obrátením týchto argumentov môžeme vysloviť nasledujúcu Vetu.

<u>Veta 1.2:</u> λ je vlastné číslo matice A práve vtedy, keď

$$\det(A - \lambda I) = 0 \, . \tag{1.5}$$

Rovnica (1.5) sa tiež nazýva charakteristická rovnica matice A. Navyše polynóm neznámej λ rádu n, ktorý dostaneme z rovnice (1.5) sa nazýva charakteristický polynóm matice A.

Využitie vlastných čísel má zastúpenie v každej vedeckej oblasti. Napríklad v kvantovej mechanike, teórii vibrácií, oceánografii a podobne [1]. Široké využitie vlastných čísel je v samotnej matematike, pri lineárnej algebre, funkcionálnej analýze, aplikovanej matematike a veľa iných. V nasledujúcich častiach tejto práce sa budeme zaoberať možnosťou zjednodušenia veľkej matice na blokový diagonálny tvar, čo by malo uľahčiť následné počítanie vlastných hodnôt. Aby sa tieto vlastné hodnoty pod vplyvom výpočtu nemenili, ukážeme si, aká podmienka musí byť splnená pre ich zachovanie. Použijeme vzťah (1.3) ktorý prenásobíme ľubovoľnou regulárnou maticou P

$$PAv = \lambda Pv$$

a použijeme fakt, že $P^{-1}P = I$, čiže

$$\lambda Pv = PAv = PAP^{-1}Pv \; .$$

Zadefinovaním $Pv = w$ dostaneme

$$\lambda w = PAP^{-1}w \; .$$

Z tohto dôvodu λ je vlastné číslo a w je vlastný vektor matice PAP^{-1}. To znamená, že vlastné číslo matice A je zároveň vlastným číslom matice PAP^{-1}.

2 Sylvestrova rovnica a jej využitie

2.1 Sylvestrova rovnica

Pod pojmom Sylvestrova rovnica [2] budeme rozumieť rovnicu tvaru

$$AX - XB = C. \tag{2.1}$$

Premennou v tejto rovnici je obdĺžniková matica X, ktorá má M riadkov a N stĺpcov. Rovnaké rozmery má aj matica C na pravej strane rovnice. Ostatné dve matice sú štvorcové. Matica A má rozmer $M \times M$, a $N \times N$ sú rozmery matice B. Ak by mali matice iné rozmery, nebolo by možné vytvoriť súčin AX a XB, a teda aj vypočítať celú rovnicu.

Maticovú Sylvestrovu rovnicu môžeme považovať za systém lineárnych rovníc s $M \times N$ premennými x_{mn}, čo sú vlastne prvky matice X, zostavené z $M \times N$ rovníc

$$\sum_{k=1}^{M} a_{mk} x_{kn} - \sum_{l=1}^{N} x_{ml} b_{ln} = c_{mn} \ ,$$

$(1 \leq m \leq M, 1 \leq n \leq N)$.

<u>Lema 2.1:</u> Ak matice A a B majú jedno a to isté vlastné číslo λ, čiže $\det(A - \lambda I_M) = 0$ a $\det(B - \lambda I_N) = 0$, potom homogénna Sylvestrova rovnica

$$AX - XB = 0 \tag{2.2}$$

v ktorej jej pravej časti stojí $M \times N$ nulová matica, má rovnica netriviálne (rôzne od nuly) riešenie X.

<u>Dôkaz:</u> Dôkaz začneme tým, že si všimneme existenciu (M rozmerného) nenulového vektora y a (N rozmerného) vektora x, ktoré sú vlastnými vektormi matíc A a B, zodpovedajúce spoločnému vlastnému číslu λ t.j.

$$Ay = \lambda y \text{ a } B^T z = \lambda z \ .$$

Všimneme si, že z druhej rovnice $B^T z = \lambda z$ vyplýva rovnica $z^T B = \lambda z^T$, alebo v rozpísanom tvare

$$[z_1 \quad z_2 \quad \dots \quad z_n]\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1N} \\ b_{21} & b_{22} & \dots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N1} & \dots & b_{NN} \end{bmatrix} = [\lambda z_1 \quad \lambda z_2 \quad \dots \quad \lambda z_n] \, .$$

Ak položíme

$$X = yz^T = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}[z_1 \quad z_2 \quad \dots \quad z_n] = \begin{bmatrix} y_1 z_1 & y_1 z_2 & \dots & y_1 z_N \\ y_2 z_1 & y_2 z_2 & \dots & y_2 z_N \\ \vdots & \vdots & \ddots & \vdots \\ y_M z_1 & y_M z_2 & \dots & y_N z_N \end{bmatrix} .$$

Potom

$$AX = Ayz^T = \lambda yz^T = \lambda X \, , XB = yz^T B = y[\lambda z]^T = \lambda yz^T = \lambda X \, .$$

Teraz je zrejmé, že takýmto spôsobom zostrojená matica X je riešením homogénnej Sylvestrovej rovnice (2.2). Z teórie o systéme lineárnych rovníc vieme, že ak je počet rovníc rovný počtu premenných a homogénny systém je riešiteľný, potom nehomogénny systém môže byť neriešiteľný pre niektoré časti rovnice na pravej strane. Z tohto dôvodu ak existuje spoločné vlastné číslo matíc A a B, potom môžeme ukázať takú maticu C, že rovnica $AX - XB = C$ nemá riešenie.■

<u>Veta 2.2:</u> Ak matice A a B nemajú žiadne spoločné vlastné čísla, potom nehomogénna Sylvestrova rovnica (2.1) je riešiteľná pre ľubovoľnú pravú časť C. Zo všeobecnej teórie lineárnych rovníc vieme, že takáto riešiteľnosť nastáva pri jednoznačnosti riešenia.

<u>Dôkaz:</u> Pri dokazovaní tejto vety sa môžeme obmedziť na prípad trojuholníkových matíc A a B za predpokladu, že A je horná a B dolná trojuholníková matica. Na základe Schurovej vety vieme nájsť také unitárne matice P, Q $(P^*P = I_M, Q^*Q = I_N)$ s pomocou ktorých môžeme matice A a B napísať v nasledovnom tvare

$$A = PA_0P^* \ \ , B = QB_0Q^* \, ,$$

kde matice A_0 a B_0 sú trojuholníkové a majú na diagonálach rozmiestnené vlastné čísla pôvodných matíc A a B. Pripomenieme, že A_0 je horná trojuholníková a B_0 je dolná trojuholníková matica. Teraz môžeme prepísať Sylvestrovu rovnicu do tvaru

$$PA_0P^*X - XQB_0Q^* = C$$

a potom je možné túto rovnicu zameniť ekvivalentnou rovnicou, ktorú dostaneme násobením pôvodnej rovnice zľava a sprava regulárnymi maticami P^* a Q

$$A_0P^*XQ - P^*XQB_0 = P^*CQ.$$

Ak dokážeme nájsť riešenie Y rovnice

$$A_0Y - YB_0 = P^*CQ\ ,$$

potom riešenie pôvodnej rovnice môžeme vypočítať ako

$$X = PYQ^*.$$

Dôkaz riešiteľnosti Sylvestrovej rovnice (2.1) za predpokladu, že neexistujú spoločné vlastné čísla matíc A a B je vhodné vykonať pomocou indukcie.

Najskôr budeme riešiť rovnice rozmerov $M \leq 2$ a $N \leq 2$. Pri týchto obmedzeniach pre trojuholníkové matice A a B, Sylvestrova rovnica nadobúda jeden z nasledujúcich štyroch tvarov:

1. $M = 1, N = 1$:

$$a_{11}x_{11} - x_{11}b_{11} = c_{11}\ .$$

Podmienka neexistencie spoločných vlastných čísel matíc A a B nás privedie k nerovnosti $a_{11} \neq b_{11}$. Rovnica je riešiteľná a riešenie má tvar

$$x_{11} = \frac{c_{11}}{(a_{11} - b_{11})}.$$

2. $M = 1, N = 2$:

$$a_{11}\begin{bmatrix} x_{11} & x_{12} \end{bmatrix} - \begin{bmatrix} x_{11} & x_{12} \end{bmatrix}\begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \end{bmatrix},\ a_{11} \neq b_{11}\ ,\ a_{11} \neq b_{22}\ .$$

Riešenie existuje a má tvar

$$x_{12} = \frac{c_{12}}{(a_{11} - b_{22})}\ ,\ x_{11} = \frac{(c_{11} + x_{12}b_{21})}{(a_{11} - b_{11})}\ .$$

3. $M = 2, N = 1$:

$$\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} - \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} b_{11} = \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix},\ a_{11} \neq b_{11}\ ,\ a_{22} \neq b_{11}\ .$$

Riešenie existuje a má tvar

$$x_{21} = \frac{c_{21}}{(a_{22} - b_{11})}\ ,\ x_{11} = \frac{(c_{11} - a_{12}x_{21})}{(a_{11} - b_{11})}\ .$$

4. $M = 2, N = 2$:

$$\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} - \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix},$$

$$a_{22} \neq b_{22}\, ,\, a_{11} \neq b_{22}\, ,\, a_{22} \neq b_{11}\, ,\, a_{11} \neq b_{11}\, .$$

Riešenie existuje a vypočítame ho ako

$$x_{11} = \frac{(c_{11} + x_{12} b_{21} - a_{12} x_{21})}{(a_{11} - b_{11})}\, ,\, x_{22} = \frac{c_{22}}{(a_{22} - b_{22})}\, ,$$

$$x_{12} = \frac{(c_{12} - a_{12} x_{22})}{(a_{11} - b_{22})}\, ,\, x_{21} = \frac{(c_{21} + x_{22} b_{21})}{(a_{22} - b_{11})}\, .$$

Z tohto vyplýva, že pre $M \leq 2$ a $N \leq 2$ je tvrdenie správne. Toto tvrdenie rozšírime na M, N pomocou matematickej indukcie. Predpokladajme, že sme už dokázali tvrdenie Sylvestrovej rovnice s rozmermi M a N, kde $M \leq m_j$ a $N \leq m_j$ a s ľubovoľnými prípustnými maticami A a B, ktoré nemajú rovnaké vlastné čísla. Teraz ukážeme riešiteľnosť tohto typu rovníc $M \leq m_{j+1}$, $N \leq m_{j+1}$, $m_{j+1} \leq 2m_j$. Prípad $m_0 = 2$ máme už vyšetrený.

Ľubovoľná horná trojuholníková matica A s rozmermi $M \times M$ pre $M \leq 2m_j$ je buď $M \times M$ maticou A_{11} $(M \leq 2m_j)$, alebo ju môžeme rozdeliť na bloky

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

tak, že A_{11} je štvorcová matica rádu $M_1 \leq m_j$ a A_{22} je štvorcová matica rádu $M_2 \leq m_j$. Pritom množina vlastných čísel oboch matíc A_{11} a A_{22} je vlastne množinou všetkých vlastných čísel pôvodnej matice A.

Presne tak isto dolná trojuholníková matica B s rozmermi $N \times N$ je pre $N \leq 2m_j$ buď maticou B_{11} pre ktorú platí $N \leq m_j$, alebo ju môžeme rozdeliť na bloky

$$B = \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix},$$

kde diagonálne blokové matice B_{11} a B_{22}majú rozmery $N_1 \times N_1$ a $N_2 \times N_2$ $(N_1 \leq m_j$, $N_2 \leq m_j)$. Je zrejmé, že ak čo u len jednej z matíc A, B rádu M (alebo N) prevyšuje m_j , potom dostaneme jednu z nasledovných troch foriem Sylvestrovej rovnice :

(a) $\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} - \begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} B_{11} = \begin{bmatrix} C_{11} \\ C_{21} \end{bmatrix}.$

(b) $A_{11}[X_{11} \quad X_{12}] - [X_{11} \quad X_{12}]\begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix} = [C_{11} \quad C_{12}]$.

(c) $\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} - \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}\begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$.

V týchto prípadoch matica neznámych prvkov X a matica C na pravej strane rovnice sú rozdelené na bloky rovnakých rozmerov. Každú z týchto maticových rovníc môžeme transformovať na systém blokových rovníc, ktoré musia zodpovedať bloku X_{ij} :

(a) $A_{22}X_{21} - X_{21}B_{11} = C_{21}$,
$A_{11}X_{11} - X_{11}B_{11} = C_{11} - A_{12}X_{21}$.

(b) $A_{11}X_{12} - X_{12}B_{22} = C_{12}$,
$A_{11}X_{11} - X_{11}B_{11} = C_{11} - X_{12}B_{21}$.

(c) $A_{22}X_{22} - X_{22}B_{22} = C_{22}$,
$A_{22}X_{21} - X_{21}B_{11} = C_{21} + X_{22}B_{21}$,
$A_{11}X_{12} - X_{12}B_{22} = C_{12} - A_{12}X_{22}$,
$A_{11}X_{11} - X_{11}B_{11} = C_{11} - A_{12}X_{21} + X_{12}B_{21}$.

V každej variante (a), (b), (c) sú tieto rovnice postupne riešiteľné jedna po druhej. Každá z týchto rovníc má riešenie, pretože blokové matice A_{ii} , B_{jj} nemajú podľa predpokladu spoločné vlastné čísla a rozmery blokov nie sú väčšie ako $m_j \times m_j$. Týmto spôsobom môžeme pre všetky dané rovnice vypočítať im zodpovedné bloky X_{11} , X_{12} , X_{21} , X_{22} na základe indukčného predpokladu o riešiteľnosti Sylvestrových rovníc pre $M \leq m_j$, $N \leq m_j$. ■

2.2 Bloková diagonalizácia štvorcových matíc pomocou Sylvestrovej rovnice

V tejto časti budeme často používať Sylvestrove maticové rovnice pri riešení rôznych úloh. Príklad ktorý uvedieme, demonštruje úlohu riešenia Sylvestrových rovníc pri hľadaní nejakej podobnej transformácie, ktorá pretransformuje maticu na diagonálny, alebo blokovo-diagonálny tvar. Najskôr však vyslovíme vetu [10], potrebnú na ukázanie tejto transformácie.

Veta 2.3 (Schur): Nech $A \in \mathbb{C}^{n\times n}$. Potom existuje taká unitárna matica $U \in \mathbb{C}^{n\times n}$ a horná trojuhoníková matica $T \in \mathbb{C}^{n\times n}$, že $T = U^*AU$.

Ekvivalentne môžeme napísať $A = UTU^*$, čo nazývame Schurov rozklad matice A .

<u>Dôkaz:</u> Túto Vetu dokážeme indukciou. Ukážeme, že ak to platí pre $n = k$, tak potom to platí aj pre $n = k - 1$. Nech $A \in \mathbb{C}^{k \times k}$, nech λ je vlastné číslo matice A a v je prislúchajúci vlastný vektor zvolený tak, aby $\|v\|_2 = 1$. Zvoľme unitárnu maticu U_1 tak, že prvý stĺpec matice bude vlastný vektor v. To znamená, že si vyberieme ľubovoľnú ortonormálnu bázu $\mathbb{C}^k$, ktorej prvý stĺpec bude vektor v a ostatné vektory tvoriace maticu U_1 budú z tejto bázy. Maticu U_1 teraz rozdelíme na vektor v a blok W, čiže $U_1 = [v \quad W]$. Keďže stĺpce matice W sú ortogonálne k v, $W * v = 0$. Nech $A_1 = U_1^* A U_1$, potom

$$A_1 = \begin{bmatrix} v^* \\ W^* \end{bmatrix} A [v \quad W] = \begin{bmatrix} v^*Av & v^*AW \\ W^*Av & W^*AW \end{bmatrix}.$$

Ak $Av = \lambda v$, potom $v^*Av = \lambda$ a $W^*Av = \lambda W^*v = 0$. Nech $\tilde{A} = W^*AW$, maticu A_1 môžeme zapísať ako

$$A_1 = \begin{bmatrix} \lambda & | & \times & \cdots & \times \\ - & & - & - & - \\ 0 & | & & & \\ \vdots & | & & \tilde{A} & \\ 0 & | & & & \end{bmatrix}.$$

$\tilde{A} \in \mathbb{C}^{(k-1)\times(k-1)}$, takže indukčným predpokladom bude existencia unitárnej matice $\widetilde{U}_2$ a hornej trojuholníkovej matice $\tilde{T}$, takých, že $\tilde{T} = \widetilde{U}_2^* \tilde{A} \widetilde{U}_2$. Definujme $U_2 \in \mathbb{C}^{k \times k}$ ako

$$U_2 = \begin{bmatrix} 1 & | & 0 & \cdots & 0 \\ - & & - & - & - \\ 0 & | & & & \\ \vdots & | & & \widetilde{U}_2 & \\ 0 & | & & & \end{bmatrix},$$

potom U_2 je unitárna a

$$U_2^* A_1 U_2 = \begin{bmatrix} \lambda & | & \times & \cdots & \times \\ - & & - & - & - \\ 0 & | & & & \\ \vdots & | & & \widetilde{U}_2^* \tilde{A} \widetilde{U}_2 & \\ 0 & | & & & \end{bmatrix} = \begin{bmatrix} \lambda & | & \times & \cdots & \times \\ - & & - & - & - \\ 0 & | & & & \\ \vdots & | & & \tilde{T} & \\ 0 & | & & & \end{bmatrix}.$$

Posledná matica je horná trojuholníková a označíme ju T. Ďalej nech $U = U_1 U_2$, potom $T = U_2^* A_1 U_2 = U_2^* U_1^* A U_1 U_2 = U^* A U$. ■

Podľa Schurovej vety vieme, že existuje taká unitárna transformácia U kde $U^*U = I_N$, s pomocou ktorej môžeme štvorcovú ($N \times N$) maticu A zapísať ako

$$A = U \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ & \lambda_2 & & \vdots \\ & & \ddots & \times \\ 0 & & & \lambda_N \end{bmatrix} U^* .$$

V tomto rozklade je prostredná matica horná trojuholníková a na jej diagonále sú rozmiestnené vlastné čísla $\lambda_j = \lambda_j(A)$ matice A. Poradie v akom sú rozmiestnené vlastné čísla na diagonále je ľubovoľné, ale každému poradiu prislúcha špecificky zvolená transformácia U. Všetky vlastné čísla rozdelíme na skupiny :

$$\lambda_1, \lambda_2, \dots, \lambda_{k_1} ; \lambda_{k_1+1}, \lambda_{k_1+2}, \dots, \lambda_{k_1+k_2} ; \dots ;$$
$$\lambda_{k_1+k_2+\dots+k_{n-1}+1}, \lambda_{k_1+k_2+\dots+k_{n-1}+2}, \dots, \lambda_{k_1+k_2+\dots+k_n}$$
$$(k_1 + k_2 + \dots + k_n = N)$$

pričom sa zaujímame iba o prípad, ak má nejaké vlastné číslo násobnosť p, potom aby boli všetky tieto násobné vlastné čísla v jednej skupine. Inými slovami, v rôznych skupinách nesmú byť rovnaké vlastné čísla, čiže musí byť $\lambda_i \neq \lambda_j$ ak tieto čísla neležia v jednej skupine. Na základe Schurovej vety existuje taká matica U, pre ktorú má trojuholníková matica rozmiestnené vlastné čísla na hlavnej diagonále v nami zvolenom poradí. To znamená $\lambda_1, \lambda_2, \dots, \lambda_{k_1}$ z prvej skupiny, potom ďalšie $\lambda_{k_1+1}, \lambda_{k_1+2}, \dots, \lambda_{k_1+k_2}$ z druhej skupiny atď. Potom túto trojuholníkovú maticu rozdelíme na bloky

$$\begin{bmatrix} \lambda_1 & \times & \dots & \times \\ & \lambda_2 & & \vdots \\ & & \ddots & \times \\ 0 & & & \lambda_N \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ & A_{22} & & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix}$$

tak, aby diagonálne bloky A_{jj} boli štvorcové matice:

$$A_{11} = \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ & \lambda_2 & & \vdots \\ & & \ddots & \times \\ 0 & & & \lambda_N \end{bmatrix},$$

$$A_{jj} = \begin{bmatrix} \lambda_{k_1+k_2+\dots+k_{j-1}+1} & & \times \\ & \ddots & \\ 0 & & \lambda_{k_1+k_2+\dots+k_j} \end{bmatrix}, j \geq 2 .$$

Toto potvrdzuje aj Schurova veta umožňujúca existenciu blokovo-trojuholníkovej reprezentácie tvaru

$$A = U \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ & A_{22} & & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix} U^* ,$$

kde diagonálne štvorcové bloky A_{ii} a A_{jj} pre $i \neq j$ nemajú spoločné vlastné čísla. Vlastné čísla matice A sú rozmiestnené v týchto blokoch ľubovoľným, vopred zadaným spôsobom.

Ukazuje sa možnosť ďalšieho zjednodušenia tvaru blokovo-trojuholníkovej matice, ale za cenu, že vzájomne inverzné unitárne matice U ($U^* = U^{-1}$), je potrebné zameniť maticami T a T^{-1} ktoré vôbec nemusia byť unitárne. Podstatné je pritom, že matica T musí byť regulárna, aby bola zabezpečená existencia inverznej matice T^{-1}.

Najprv blokovo-trojuholníkovú maticu

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ & A_{22} & & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix},$$

v ktorej má blok A_{11} rozmer $k_1 \times k_1$, zapíšeme v tvare

$$\begin{bmatrix} A_{11} & C_1 \\ 0 & B_1 \end{bmatrix},$$

pričom matice B_1 a C_1 majú tvar

$$B_1 = \begin{bmatrix} A_{22} & A_{23} & \dots & A_{2n} \\ & A_{33} & & A_{3n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix}, C_1 = [A_{12} \quad A_{13} \quad \dots \quad A_{1n}] .$$

Je zrejmé, že A_{11} a B_1 nemajú spoločné vlastné čísla. Pokúsime sa nájsť takú $k_1 \times (N - k_1)$ maticu R_1 , aby platilo

$$\begin{bmatrix} I_{k_1} & -R_1 \\ 0 & I_{N-k_1} \end{bmatrix} \begin{bmatrix} A_{11} & C_1 \\ 0 & B_1 \end{bmatrix} \begin{bmatrix} I_{k_1} & R_1 \\ 0 & I_{N-k_1} \end{bmatrix} = \begin{bmatrix} A_{11} & C_1 + A_{11}R_1 - R_1B_1 \\ 0 & B_1 \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & B_1 \end{bmatrix}.$$

Vidíme, že matica R_1 musí byť riešením Sylvestrovej rovnice

$$A_{11}R_1 - R_1B_1 = -C_1 \,,$$

ktorá je jednoznačne riešiteľná, pretože A_{11} a B_1 nemajú spoločné vlastné čísla.

Ak použijeme rovnosť

$$\begin{bmatrix} I_{k_1} & -R_1 \\ 0 & I_{N-k_1} \end{bmatrix} \begin{bmatrix} I_{k_1} & R_1 \\ 0 & I_{N-k_1} \end{bmatrix} = I_N$$

a označíme

$$S_1 = \begin{bmatrix} I_{k_1} & R_1 \\ 0 & I_{N-k_1} \end{bmatrix} \ (\det\ S_1 = 1)\,,$$

dôjdeme k rovnici

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ & A_{22} & & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix} = S_1 \begin{bmatrix} A_{11} & 0 & \dots & 0 \\ 0 & A_{22} & \dots & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix} S_1^{-1}\,.$$

Túto novú blokovo-trojuholníkovú maticu

$$\begin{bmatrix} A_{11} & 0 & \dots & 0 \\ 0 & A_{22} & \dots & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix}$$

prepíšeme na tvar

$$\begin{bmatrix} A_2 & C_2 \\ 0 & B_2 \end{bmatrix}, \tag{2.3}$$

kde

$$A_2 = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}\,,\ B_2 = \begin{bmatrix} A_{33} & \dots & A_{3n} \\ & \ddots & \vdots \\ 0 & & A_{nn} \end{bmatrix}\,,\ C_2 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ A_{23} & A_{24} & \dots & A_{2n} \end{bmatrix}.$$

Je zrejmé, že bloky A_2 a B_2 znova nemajú spoločné vlastné čísla. Ak zopakujeme takmer rovnako všetky úvahy, ktoré sme použili pri konštrukcii matice S_1 , dostávame maticu S_2

$$S_2 = \begin{bmatrix} I_{k_2} & R_2 \\ 0 & I_{N-k_2} \end{bmatrix} \quad (\det\ S_2 = 1)\ ,$$

takú, kde R_2 je riešením rovnice

$$A_2R_2 - R_2B_2 = -C_2\ .$$

Maticu S_2 a jej inverznú maticu S_2^{-1} použijeme na vynulovanie bloku matice (2.3)

$$\begin{bmatrix} A_2 & C_2 \\ 0 & B_2 \end{bmatrix} = S_2 \begin{bmatrix} A_{11} & 0 & | & & & 0 \\ 0 & A_{22} & | & & & \\ - & - & & - & - & - \\ & & | & A_{33} & \dots & A_{3n} \\ & & | & & \ddots & \vdots \\ 0 & & | & 0 & & A_{nn} \end{bmatrix} S_2^{-1}$$

a teda

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ & A_{22} & & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix} = S_1S_2 \begin{bmatrix} A_{11} & 0 & | & & & 0 \\ 0 & A_{22} & | & & & \\ - & - & & - & - & - \\ & & | & A_{33} & \dots & A_{3n} \\ & & | & & \ddots & \vdots \\ 0 & & | & 0 & & A_{nn} \end{bmatrix} S_2^{-1}S_1^{-1}\ .$$

Ak použijeme tento postup, postupne vynulujeme bloky A_{ij} na pravej strane od hlavnej blokovej diagonály pomocou matíc S_3 , ... , S_{n-1} (det $S_j = 1$). Ako výsledok dostaneme rozklad

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ & A_{22} & & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix} = S_1S_2 \dots S_{n-1} \begin{bmatrix} A_{11} & & & 0 \\ & A_{22} & & \\ & & \ddots & \\ 0 & & & A_{nn} \end{bmatrix} S_{n-1}^{-1} \dots S_2^{-1}S_1^{-1}\ .$$

Označíme $T = US_1S_2 \dots S_{n-1}$ (je zrejmé, že $|\det\ T| = |\det\ U| = 1$, $\det\ T \neq 0$) čiže dostaneme konečný tvar

$$A = T \begin{bmatrix} A_{11} & & & 0 \\ & A_{22} & & \\ & & \ddots & \\ 0 & & & A_{nn} \end{bmatrix} T^{-1}$$

a z tohto zápisu vyslovíme nasledujúcu Vetu.

Veta 2.4: Pre ľubovoľnú maticu A existuje regulárna podobná transformácia, ktorá transformuje A do blokovo-diagonálneho tvaru a to tak, že vlastné čísla matice A sú rozmiestnené vopred zadaným spôsobom na diagonálnych blokoch A_{jj} .

2.3 Vynulovanie bloku matice pomocou špeciálne zvolenej matice

V predchádzajúcej časti sme použili Sylvestrovu rovnicu na blokovú diagonalizáciu matice A. Na potrebné vynulovanie ľavej dolnej časti matice A sme použili Schurov rozklad. Existujú však aj iné metódy, ktoré vynulujú ľavý dolný blok matice a zároveň zachovajú hodnoty vlastných čísel. Prvou takouto metódou je Householderova metóda, ktorú podrobne rozoberieme v nasledujúcej kapitole. Druhou metódou je počítanie s blokmi matice princípom, ktorý ukážeme v tejto časti.

Uvažujme maticu A, ktorej ľavý dolný blok budeme chcieť vynulovať pomocou nejakej transformácie

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = G \begin{bmatrix} \times & \times \\ 0 & \times \end{bmatrix} G^T \ ,$$

kde pre jednoduchosť výpočtu budeme považovať všetky bloky A_{ij} za štvorcové. Rozmery matíc A a G sú zhodné. Aby bola splnená podmienka nemennosti vlastných čísel pôvodnej matice, budeme požadovať, aby mali transformačné matice G a G^T vlastnosť $GG^T = I$ [11]. Z tohto dôvodu zvolíme bloky matice nasledovne

$$G = \begin{bmatrix} D & DE^T \\ -FE & F \end{bmatrix},$$

kde bloky D, E, F sú štvorcové, rovnaké ako bloky matice A.

Ukážeme za akej podmienky je takto zvolená matica G vyhovujúca podmienke $GG^T = I$, takže

$$GG^T = \begin{bmatrix} D & DE^T \\ -FE & F \end{bmatrix} \begin{bmatrix} D^T & -E^T F^T \\ ED^T & F^T \end{bmatrix} = \begin{bmatrix} DD^T + DE^T ED^T & -DE^T F^T + DE^T F^T \\ -FED^T + FED^T & FEE^T F^T + FF^T \end{bmatrix} =$$
$$= \begin{bmatrix} DD^T + DE^T ED^T & 0 \\ 0 & FEE^T F^T + FF^T \end{bmatrix}.$$

Vidíme, že prvky mimo diagonály sa nám vynulovali, takže ostáva už len ukázať podmienky pre diagonálne prvky. Začneme ľavým horným blokom

$$DD^T + DE^T ED^T = I$$
$$D(I + E^T E)D^T = I$$

$$I + E^T E = (D^T D)^{-1}$$
$$D^T D = (I + E^T E)^{-1}$$

aby táto rovnosť platila, matica D musí byť Choleského faktor z $(I + E^T E)^{-1}$. Analogický postup platí aj pre pravý dolný blok, len s malou obmenou, že matica F musí byť Choleského faktor z $(I + EE^T)^{-1}$.

Za platnosti týchto podmienok teda vieme, že matice G a G^T majú vlastnosť $GG^T = I$. To znamená, že pokojne môžeme pokračovať v postupe ako vynulovať prvok matice A, teda môžeme počítať

$$GAG^T = \begin{bmatrix} D & DE^T \\ -FE & F \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} D^T & -E^T F^T \\ ED^T & F^T \end{bmatrix} =$$

$$= \begin{bmatrix} DA_{11} + DE^T A_{21} & DA_{12} + DE^T A_{22} \\ -FEA_{11} + FA_{21} & -FEA_{12} + FA_{22} \end{bmatrix} \begin{bmatrix} D^T & -E^T F^T \\ ED^T & F^T \end{bmatrix} =$$

$$= \begin{bmatrix} (DA_{11} + DE^T A_{21})D^T + (DA_{12} + DE^T A_{22})ED^T & -(DA_{11} + DE^T A_{21})E^T F^T + (DA_{12} + DE^T A_{22})F^T \\ (-FEA_{11} + FA_{21})D^T + (-FEA_{12} + FA_{22})ED^T & -(-FEA_{11} + FA_{21})E^T F^T + (-FEA_{12} + FA_{22})F^T \end{bmatrix}.$$

Teraz sa budeme zaujímať len o ľavý dolný blok, pretože ten chceme vynulovať

$$(-FEA_{11} + FA_{21})D^T + (-FEA_{12} + FA_{22})ED^T = 0$$
$$-FEA_{11}D^T + FA_{21}D^T - FEA_{12}ED^T + FA_{22}ED^T = 0$$
$$F(-EA_{11} + A_{21} - EA_{12}E + A_{22}E)D^T = 0$$

matice F a D^T sú podľa predchádzajúcich výpočtov nenulové, môžeme ich vynechať a pokračujeme

$$-EA_{11} + A_{21} - EA_{12}E + A_{22}E = 0$$
$$A_{21} - EA_{11} + A_{22}E - EA_{12}E = 0.$$

Všimnime si, že v poslednej rovnici sa nachádzajú všetky bloky matice A a len jeden blok matice G. Inými slovami, v poslednej rovnici sa nachádza len jedna neznáma, čiže blok E z transformačnej matice G. Rovnicu takéhoto tvaru vieme riešiť, pretože je to tzv. Riccatiho algebraická rovnica.

Riccatiho algebraická rovnica sa však už nepočíta takýmto blokovo-maticovým spôsobom. Zvyčajne sa používajú numerické iteračné metódy (napr. Newtonova) [4].

Zhrnutím postupu teda dostávame, že v štvorcovej matici A sa dá vynulovať ľavý dolný blok pomocou Riccatiho algebraickej rovnice, odkiaľ dostaneme maticový blok E. Následným aplikovaním Choleského rozkladu dostaneme ostatné dva bloky D, F a vieme teda zostrojiť transformačnú maticu G.

3 Householderova metóda a jej využitie

Householderova metóda je ortogonálnou geometrickou transformáciou. Skôr ako však nahliadneme na samotnú Householderovu metódu, spomenieme podobnú ortogonálnu transformáciu, Givensove otočenie. Ako uvidíme, Householderova metóda vie vynulovať celý vektor okrem jedného prvku, alebo celý maticový blok. Givensove otočenie je jemnejšie, inými slovami, vie vynulovať jeden prvok matice. Nie vždy je totiž žiaduce transformovať iba veľké bloky.

3.1 Givensove otočenie

Otočenie je druh transformácie ktoré vidíme na Obrázku 1 a dá sa opísať ako zmena dvoch súradníc vektorov v rovine, pričom ostatné súradnice zanecháva nezmenené. Podstatu si vysvetlíme na jednoduchom dvojrozmernom príklade.

Uvažujme otočenie vektora $x = (x_1, x_2)$ o uhol θ na $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$. Dĺžka je zachovaná, takže $\|\tilde{x}\| = \|x\|$. Podľa Obrázku 1 môžeme napísať

$$\begin{aligned}\tilde{x}_1 &= \|x\| \cos(\Phi + \theta),\\ \tilde{x}_2 &= \|x\| \sin(\Phi + \theta).\end{aligned}$$

Z elementárnej trigonometrie využijeme

$$\begin{aligned}\cos(\Phi + \theta) &= \cos\Phi \cos\theta - \sin\Phi \sin\theta,\\ \sin(\Phi + \theta) &= \sin\Phi \cos\theta + \cos\Phi \sin\theta.\end{aligned}$$

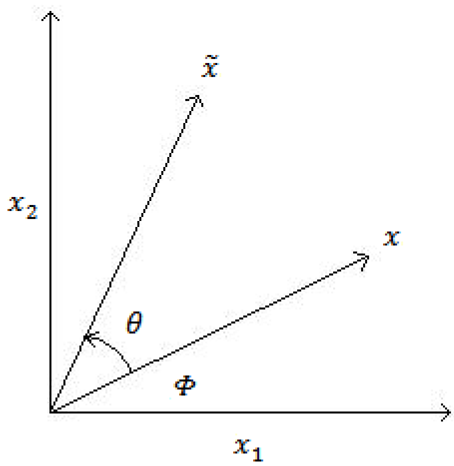
Pretože $\cos\Phi = x_1/\|x\|$ a $\sin\Phi = x_2/\|x\|$, ich skombinovaním dostaneme

$$\begin{aligned}\tilde{x}_1 &= x_1 \cos\theta - x_2 \sin\theta,\\ \tilde{x}_2 &= x_1 \sin\theta + x_2 \cos\theta.\end{aligned}$$

Z tohto postupu vidíme, že vynásobením x ortogonálnou maticou

$$\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

dostaneme otočenie vektora x o uhol θ.



Obrázok 1

V trojrozmernom priestore je otočenie podobné s dvojrozmerným priestorom, len nám pribudne jedna súradnica. Na ilustráciu uvedieme matice otočenia v trojrozmernom priestore podľa hlavných osí [7].

Otočenie okolo osi x

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

Otočenie okolo osi y

$$\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}.$$

Otočenie okolo osi z

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Podstatu Givensového otočenia [3] si tiež znázorníme v dvojrozmernom priestore. Máme ľubovoľný vektor $x = (x_1, x_2)$, ktorý chceme transformovať na vektor $\tilde{x} = (\tilde{x}_1, 0)$, ako vidíme na Obrázku 2. Transformáciu prevedieme za pomoci ortogonálnej matice otočenia tvaru

$$G = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},$$

kde môžeme funkcie sínus a kosínus vyjadriť pomocou zlomkov

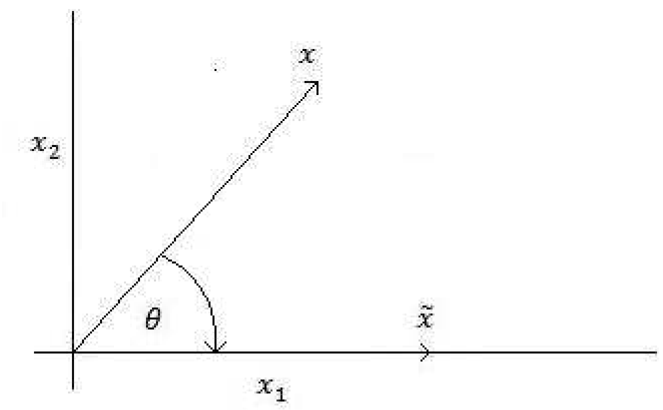
$$c = \cos\theta = \frac{x_1}{r},$$
$$s = \sin\theta = \frac{x_2}{r},$$

číslo r je norma vektora x

$$r = \|x\| = \sqrt{x_1^2 + x_2^2}\,.$$

Z tohto je potom ľahko vidieť, ako prebieha Givensove otočenie

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} cx_1 + sx_2 \\ -sx_1 + cx_2 \end{bmatrix} = \begin{bmatrix} \frac{x_1}{r}x_1 + \frac{x_2}{r}x_2 \\ -\frac{x_2}{r}x_1 + \frac{x_1}{r}x_2 \end{bmatrix} = \begin{bmatrix} \frac{x_1^2}{r} + \frac{x_2^2}{r} \\ \frac{x_1x_2}{r} - \frac{x_1x_2}{r} \end{bmatrix} = \begin{bmatrix} \frac{x_1^2+x_2^2}{\sqrt{x_1^2+x_2^2}} \\ 0 \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$



Obrázok 2

Vidíme, že pre vektor Givensové otočenie funguje a teraz si ukážeme aj aplikáciu na maticu [9]. Všeobecný tvar matice Givensového otočenia, ktorá vynuluje súradnicu $x_{j,i}$ v ľubovoľnej matici, je

$$G(i,j,\theta) = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & c_{i,i} & \cdots & s_{i,j} & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -s_{j,i} & \cdots & c_{j,j} & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}.$$

Pomocou tejto matice chceme vynulovať súradnicu a_{ij} v matici A. Givensova transformácia matice A má tvar

$$\tilde{A} = G^T A G \ , \tag{3.1}$$

kde

$$\tilde{A} = \begin{bmatrix} & \vdots & & \vdots & \\ \cdots & \tilde{a}_{ii} & \cdots & \tilde{a}_{ij} & \cdots \\ & \vdots & & \vdots & \\ \cdots & \tilde{a}_{ji} & \cdots & \tilde{a}_{jj} & \cdots \\ & \vdots & & \vdots & \end{bmatrix} .$$

Z rovnice (3.1) si vyjadríme niektoré prvky z matice $\tilde{A}$ kde ($k \neq i$, $k \neq j$)

$$\tilde{a}_{ii} = c^2 a_{ii} + s^2 a_{jj} - 2sca_{ij} \ , \tag{3.2}$$
$$\tilde{a}_{jj} = s^2 a_{ii} + c^2 a_{jj} + 2sca_{ij} \ , \tag{3.3}$$
$$\tilde{a}_{ij} = (c^2 - s^2) a_{ij} + sc\left(a_{ii} - a_{jj}\right) . \tag{3.4}$$

Podľa toho, že chceme $\tilde{a}_{ij} = 0$, rovnica (3.4) nám umožňuje vyjadriť uhol otočenia θ nasledovne

$$\omega = \operatorname{cotg} 2\theta = \frac{c^2 - s^2}{2sc} = \frac{a_{jj} - a_{ii}}{2a_{ij}} . \tag{3.5}$$

Nech $t = s/c$, a definíciu ω (3.5) môžeme napísať v tvare

$$t^2 + 2t\omega - 1 = 0 .$$

Najmenší koreň tejto rovnice zodpovedá uhlu otočenia menšiemu ako $\pi/4$. Štandardným zápisom koreňa kvadratickej rovnice máme

$$t = \frac{\operatorname{sgn}(\omega)}{|\omega| + \sqrt{\omega^2 + 1}} .$$

Zadefinovaním funkcií sínus a kosínus

$$c = \frac{1}{\sqrt{t^2 + 1}}$$

$$s = tc$$

a dosadením do rovníc (3.2) až (3.4) dostaneme konečné riešenie

$$\tilde{a}_{ii} = a_{ii} - ta_{ij} \ ,$$
$$\tilde{a}_{jj} = a_{jj} - ta_{ij} \ ,$$
$$\tilde{a}_{ij} = 0 \ .$$

Vidíme, že Givensove otočenie funguje aj na vynulovanie súradnice matice.

3.2 Householderova transformácia

Householderovu transformáciu si najskôr priblížime jednoduchým príkladom. Nech u a v sú ortonormálne vektory a nech vektor x leží medzi vektormi u a v, ktorý je daný ako kombinácia týchto susedných vektorov

$$x = c_1 u + c_2 v$$

kde c_1 a c_2 sú skaláry. Potom vektor

$$\tilde{x} = -c_1 u + c_2 v$$

je obrazom x, ak je definovaný vektorom v, alebo $u^{\perp}$.

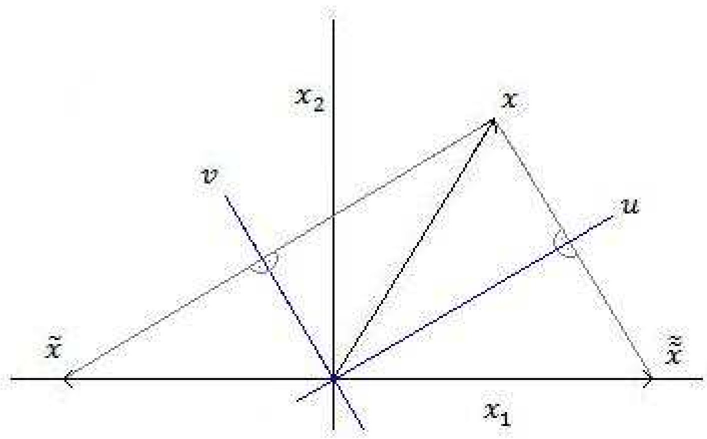
Uvažujeme obraz, ktorý transformuje vektor

$$x = (x_1, x_2, \dots, x_n)$$

na vektor kolineárny s jednotkovým vektorom,

$$\tilde{x} = (0, \dots, 0, \tilde{x}_i, 0, \dots, 0) = \pm \|x\|_2 e_i \ .$$

Ak sa obmedzíme len na dvojrozmerný priestor, môžeme toto znázorniť ako na Obrázku 3, kde $i = 1$. Vektor x je otočený podľa toho, aké je znamienko pri $\pm\|x\|_2$. Výber znamienok je na Obrázku 3 znázornený ako vektor $\tilde{x}$ a vektor $\tilde{\tilde{x}}$.



Obrázok 3

Teraz sa pozrieme na zobrazenie vektora x iným spôsobom. Uvažujme problém zobrazenia vektora x ($x \in \mathbb{R}^n$) cez vektor u ($u \in \mathbb{R}^n$). Použijeme Householderovu maticu transformácie H ($H \in \mathbb{R}^{n\times n}$) [3]

$$H = I - 2\frac{uu^T}{u^Tu},$$

ktorá transformuje vektor $x = (x_1, \dots, x_n)$ na jeho obraz $\tilde{x} = (\tilde{x}_1, 0, \dots, 0)$

$$Hx = \left(I - 2\frac{uu^T}{u^Tu}\right)x = x - 2\frac{uu^T}{u^Tu}x = x - 2\frac{u^Tx}{u^Tu}u = \tilde{x}\ .$$

Nastáva problém ako nájsť vektor u tak, aby spĺňal túto podmienku. Pomôžeme si Obrázkom 3. Vidíme, že vektor u sa dá vyjadriť za pomoci nášho známeho vektora x. Predpokladajme teda, že $u = x + \alpha e_1$ a počítajme výraz z matice H, t.j. začneme čitateľom

$$u^Tx = x^Tx + \alpha x_1\ ,$$

a menovateľ

$$u^Tu = x^Tx + 2\alpha x_1 + \alpha^2.$$

Ďalej počítajme

$$Hx = \left(1 - 2\frac{x^Tx+\alpha x_1}{x^Tx+2\alpha x_1+\alpha^2}\right)x - 2\alpha\frac{u^Tx}{u^Tu}e_1\ .$$

Aby bolo x vynulované, určíme koeficient α nasledovne $\alpha = \pm\|x\|_2$ a dostaneme

$$u = x \pm \|x\|_2 e_1 \;\; \rightarrow \;\; Hx = \left(I - 2\frac{uu^T}{u^T u}\right) x = \mp\|x\|_2 e_1 \,.$$

Takto určený vektor u sa nazýva Householderov vektor.

Takéto zobrazenie sa môže nazývať Householderove zobrazenie, alebo Householderova transformácia. Matica H má nasledujúce vlastnosti:

- $Hu = -u$.
- $Hv = v$ pre ľubovoľné v ortogonálne k u.
- $H = H^T$ (symetrickosť).
- $H^T = H^{-1}$ (ortogonálnosť).

Pretože H je ortogonálna, tak ak $Hx = \tilde{x}$, potom $\|x\|_2 = \|\tilde{x}\|_2$, čiže $\tilde{x}_1 = \pm\|x\|_2$. Navyše matica uu^T je symetrická, idempotentná a má hodnosť 1.

3.3 Vynulovanie prvkov vektora pomocou Householderovej metódy

Užitočnosť Householderovej transformácie spočíva v jednoduchosti skonštruovania zobrazenia, ktoré bude transformovať vektor x na taký vektor $\tilde{x}$, že bude mať na všetkých pozíciách nuly, okrem prvej. Uvedieme príklad, v ktorom ukážeme princíp Householderovej transformácie. Uvažujeme vektor

$$x = (3, 1, 2, 1, 1) \,,$$

ktorý by sme chceli transformovať na tvar

$$\tilde{x} = (\tilde{x}_1, 0, 0, 0, 0).$$

Dosadíme do vzorca

$$u = x \pm \|x\|_2 e_1$$

kde

$$\|x\|_2 = 4 \,,$$

čiže máme vektor

$$u = (7, 1, 2, 1, 1) \ .$$

Pomocou Householderového vektora u ďalej počítame maticu zobrazenia

$$H = I - 2\frac{uu^T}{u^Tu} \ .$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{28}\begin{bmatrix} 49 & 7 & 14 & 7 & 7 \\ 7 & 1 & 2 & 1 & 1 \\ 14 & 2 & 4 & 2 & 2 \\ 7 & 1 & 2 & 1 & 1 \\ 7 & 1 & 2 & 1 & 1 \end{bmatrix} =$$

$$= \frac{1}{28}\begin{bmatrix} -21 & -7 & -14 & -7 & -7 \\ -7 & 27 & -2 & -1 & -1 \\ -14 & -2 & 24 & -2 & -2 \\ -7 & -1 & -2 & 27 & -1 \\ -7 & -1 & -2 & -1 & 27 \end{bmatrix}$$

Vynásobením

$$Hx = \frac{1}{28}\begin{bmatrix} -21 & -7 & -14 & -7 & -7 \\ -7 & 27 & -2 & -1 & -1 \\ -14 & -2 & 24 & -2 & -2 \\ -7 & -1 & -2 & 27 & -1 \\ -7 & -1 & -2 & -1 & 27 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \\ 2 \\ 1 \\ 1 \end{bmatrix},$$

dostaneme $\tilde{x} = (-4, 0, 0, 0, 0)$, čo je v súlade s predpokladom všetkých nulových prvkov okrem prvého. Týmto sme ukázali funkčnosť Householderovej metódy pre vektory.

3.4 Bloková Householderová metóda

Bloková Householderová metóda [8] je prirodzeným rozšírením vektorovej Householderovej metódy so zmenou, kde sa vektory algoritmu nahradia blokmi matice. Matica transformácie ktorá je v tejto časti použitá, sa môže počítať pomocou polárneho rozkladu, alebo singulárneho rozkladu. Použijeme singulárny rozklad matice, pretože je viac zaužívaný a tiež sa mu venuje viac pozornosti. Pri počítaní blokovej Householderovej metódy sa používa ešte jeden rozklad. Nie je úplne špecifikovaný, len má byť ortogonálny. My sme si za tento ortogonálny typ zvolili QR rozklad matice. Na začiatok tejto časti uvedieme stručný náhľad na tieto dva spomínané rozklady a to QR rozklad a singulárny rozklad matice.

3.4.1 QR rozklad

QR rozklad [6] ľubovolnej $A \in \mathbb{R}^{n \times n}$ ($n \in \mathbb{N}$) matice má tvar

$$A = QR \,, \tag{3.6}$$

kde $Q \in \mathbb{R}^{n \times n}$ ($n \in \mathbb{N}$) je ortogonálna matica a matica $R \in \mathbb{R}^{n \times n}$ ($n \in \mathbb{N}$) je horná trojuholníková matica. Maticu Q získame napríklad pomocou Householderových vektorov, ktoré volíme tak, aby násobením matice A zľava vynulovali prvky pod hlavnou diagonálou

$$\tilde{Q}_n \tilde{Q}_{n-1} \dots \tilde{Q}_2 \tilde{Q}_1 A = R \,.$$

Maticu $\tilde{Q}$ vyjadríme ako $\tilde{Q} = \tilde{Q}_n \tilde{Q}_{n-1} \dots \tilde{Q}_2 \tilde{Q}_1$. Keďže jednotlivé matice $\tilde{Q}_i$ ($i = 1, \dots, n$) sú ortogonálne, potom aj matica $\tilde{Q}$ je ortogonálna. Využitím definície ortogonálnych matíc (pozri kapitolu 1.1.6) môžeme maticu Q označiť nasledovne $Q = \tilde{Q}^{-1} = \tilde{Q}^T$. Z toho vyplýva, že matica Q je tiež ortogonálna a potom QR rozklad (3.6) má tvar

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \cdots & q_{nn} \end{bmatrix} \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{bmatrix}.$$

V prípade obdĺžnikovej matice A rozmerov $m \times n$ ($m, n \in \mathbb{N}$) počítame rozklad analogicky ako pri štvorcovej matici s tým rozdielom, že počet jednotlivých matíc $\tilde{Q}_i$ nám určí menší rozmer matice. To znamená, že ak máme $m < n$ tak $i = 1, \dots, m$, alebo ak $m > n$ tak $i = 1, \dots, n$. Prípad $m > n$ potom vyzerá

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} q_{11} & \cdots & q_{1m} \\ \vdots & \ddots & \vdots \\ q_{m1} & \cdots & q_{mm} \end{bmatrix} \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \\ 0 & & \end{bmatrix}.$$

Nájdenie QR rozkladu pomocou Householderovej metódy je jedným z viacerých spôsobov. Túto metódu uvádzame logicky preto, že sme sa s ňou práve oboznámili a netreba ju opäť bližšie vysvetľovať. Príklady iných metód hľadania QR rozkladu sú uvedené v [6].

3.4.2 Singulárny rozklad matíc (SVD)

Princípom SVD [6] je rozklad reálnej, alebo komplexnej matice dvomi ortogonálnymi maticami. Uvažujme maticu A rozmerov $m \times n$ ($m, n \in \mathbb{N}$), potom jej singulárnym rozkladom bude

$$\underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mm} \end{bmatrix}}_{W} \underbrace{\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nn} \end{bmatrix}^T}_{V},$$

kde matica $D \in \mathbb{R}^{m\times n}$ je diagonálna matica a na jej diagonále sa nachádzajú singulárne hodnoty matice A usporiadané podľa veľkosti $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$ ($p = \min\{m, n\}$). Matica $V \in \mathbb{R}^{n\times n}$ je ortogonálna matica a $W \in \mathbb{R}^{m\times m}$ je tiež ortogonálna matica. Nasledujúce Vety bez dôkazov použijeme pri popísaní postupu výpočtu SVD.

<u>Veta 3.1:</u> Pre ľubovoľnú $A \in \mathbb{R}^{m\times n}$ ($m > n$, $m, n \in \mathbb{N}$) maticu existujú ortogonálne matice $W = (w_1, \ldots, w_m)$ a $V = (v_1, \ldots, v_n)$ také, že

$$W^T A V = D = diag(\sigma_1, \ldots, \sigma_n)$$

kde D je $m \times n$ matica a $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$.

<u>Veta 3.2:</u> Ak A je $m \times n$ matica a $A = WDV^T$ je singulárny rozklad matice A, potom stĺpcové vektory matice U sú normované vlastné vektory matice AA^T a stĺpcové vektory matice V sú normované vlastné vektory matice $A^T A$. Singulárne čísla matice A sa dajú nájsť ako druhé odmocniny vlastných čísel matice $A^T A$ alebo AA^T.

Využitím vlastností z týchto Viet uvedieme stručný postup výpočtu singulárneho rozkladu $m \times n$ ($m, n \in \mathbb{N}$) matice A .

Najprv nájdeme vlastné čísla λ_i matice $A^T A$ a zoradíme ich postupne od najväčšieho po najmenšie.

Ďalej nájdeme ortogonálne vlastné vektory matice $A^T A$, ktoré prislúchajú nájdeným vlastným číslam. Tieto vektory znormujeme a pomocou nich vytvoríme maticu V tak, že jednotlivé vektory budú stĺpcami matice.

Diagonálnu maticu D skonštruujeme tak, že na jej hlavnú diagonálu dáme $p = \min\{m, n\}$ hodnôt $\sigma_i = \sqrt{\lambda_i}$, usporiadaných od najväčšieho po najmenšie.

Zostáva nám už len skonštruovať maticu W. Prvé stĺpce matice W nájdeme podľa predpisu $w_i = \sigma_i^{-1} A v_i$, kde v_i sú stĺpce matice V a $i = 1, \ldots, p$. Ostatné stĺpce dostaneme použitím Gram-Schmidtovej ortogonalizačnej metódy a vektory znormujeme.

Tento postup je vcelku jednoduchý, ale jeho nedostatok je hlavne pri veľkých maticiach, pretože determinant veľkých matíc sa počíta ťažko. Iné metódy výpočtov nájdeme v [5] a [6].

3.4.3 Výpočet blokovej Householderovej metódy

Po oboznámení sa s potrebnými rozkladmi si ukážeme ako sa počíta bloková Householderova metóda [8]. To znamená, že nebudeme musieť nulovať stĺpce matice po jednom, ale bude to možné po blokoch, čo je samozrejme efektívnejšie pri veľkých maticiach. Zmení sa vlastne len metóda výpočtu vektora u, v tomto prípade teda matice U, ktorá sa používa na zostrojenie Householderovej matice $H = I - 2UU^T$.

Na začiatok si však ešte upravíme značenie matíc. Identická matica I bude mať označenie I_k ak sa bude jednať o štvorcovú maticu $k \times k$. Zovšeobecnená identická matica E_k bude označovať obdĺžnikovú maticu $m \times k$.

Uvažujme maticu $C \in \mathbb{R}^{m \times b}$ $(b \leq m)$ s nenulovou hodnosťou $\mathrm{hod}(C) = r$. Pre danú maticu existuje taká matica $U \in \mathbb{R}^{m \times r}$ $(U^T U = I_r)$, ktorá transformuje všetky prvky matice HC na nuly, okrem prvých r riadkov. Kde ako vieme, H je Householderova matica a matica U je tzv. bloková Householderova matica skonštruovaná z matice C. Zavedieme ešte označenie $\beta \in \mathbb{R}^{r \times b}$, je to matica označujúca prvých r nenulových riadkov v matici HC. Zapísané ako $HC = E_r \beta$.

Pokračovať budeme už samostatným postupom výpočtu blokovej Householderovej matice U.

<u>Krok 1:</u> Vyberieme si maticu C, ktorú chceme transformovať a použijeme ortogonálnu dekompozíciu, v našom prípade spomínaný QR rozklad

$$C = XZ \ ,$$

kde $X \in R^{m \times r}$ je ortogonálne a $Z \in R^{r \times b}$.

<u>Krok 2:</u> Označíme teraz $\hat{X} \in \mathbb{R}^{r \times r}$, čo vyjadruje časť matice X, konkrétne prvých r riadkov. Za ich pomoci vynulujeme ostatné prvky a oni samotné zostanú nenulové. Treba pripomenúť, že predpokladáme $\mathrm{hod}(C) = r$ nenulové. Ak by $r = 0$ potom výpočet končí v prvom kroku, keďže nemôžeme vybrať prvých r riadkov. Počítanie pokračuje SVD rozkladom matice $\hat{X}$

$$\hat{X} = WDV \ ,$$

kde $W, V \in \mathbb{R}^{r \times r}$ sú ortogonálne matice a $D \in \mathbb{R}^{r \times r}$ je matica so singulárnymi hodnotami na diagonále. Z tohto môžeme maticu β vyjadriť nasledovne

$$\beta = (-WV)Z \,.$$

Krok 3: Pokračujeme vypočítaním matice $Y \in \mathbb{R}^{m\times r}$

$$Y = X + E_r WV.$$

Krok 4: V poslednom kroku by sme mali hľadať nejaký ortogonálny rozklad matice Y, ktorý by nám dal už blokovú Householderovu maticu U takú, aby $U^T U = I_r$. Maticu U však vieme nájsť aj efektívnejšie, bez ortogonálneho rozkladu ako

$$U = Y\left[V^T\big(2(I_r + D)\big)^{-1/2}\right], \qquad (3.7)$$

kde U je špeciálne zvolená tak, aby $U^T U = I_r$ platilo. Zostáva dopočítať Householderovu maticu

$$H = I - 2UU^T \,,$$

ktorá nám vynuluje blok a platí

$$HC = E_r\beta \,.$$

Na zostavenie špeciálne zvoleného bloku U (3.7) si uvedieme vlastnosti niektorých matíc použitých v Krokoch 1 – 4. Predpokladáme, že $m, b, r \in \mathbb{N}$ a $0 \le r \le b \le m$.

Krok 1: Matica X pozostáva z ortonormálnych vektorov t.j. $X^T X = I_r$. Pre maticu Z platí $\mathrm{hod}(Z) = r$.

Krok 2: $I_r = X^T X = \hat{X}^T\hat{X} + \big(X - E_r\hat{X}\big)^T\big(X - E_r\hat{X}\big) = V^T D^2 V + \big(X - E_r\hat{X}\big)^T\big(X - E_r\hat{X}\big)$ a preto $I_r - D^2 = \big[\big(X - E_r\hat{X}\big)V^T\big]^T\big[\big(X - E_r\hat{X}\big)V^T\big]$. Keďže pravá časť tejto rovnice je kladne semidefinitná, tak existuje $I_r - D^2$. Z tohto zase vyplýva, že singulárne hodnoty nebudú väčšie ako 1, čo napomáha k stabilnejším ďalším krokom.

Krok 3: Vzťah $Y^T Y = V^T 2(I_r + D)V$ dostaneme z $X^T X = I_r$ a $X^T E_r = \hat{X}^T$. Y je riešiteľná bez komplikácií a $\mathrm{cond}(Y) \le \sqrt{2}$, pretože všetky prvky na diagonále $2(I_r + D)$ nadobúdajú hodnoty medzi 2 a 4. Fakt, že $\mathrm{cond}(Y) \le \sqrt{2}$ dostaneme z definície $\mathrm{cond}(Y) = \|Y\|\|Y^{-1}\|$ a z $Y = (I_r + D)V$.

Krok 4: Nech nejaký ortogonálny rozklad matice Y sa dá zapísať ako $Y = US$. Matica $S \in \mathbb{R}^{r \times r}$ je invertibilná, z čoho $\mathrm{cond}(S) \leq \sqrt{2}$. Matica U môže byť vynásobená ľubovoľnou ortogonálnou maticou veľkosti r sprava. Špeciálna voľba matice U (3.7), pochádza z Kroku 3, spĺňa túto podmienku a tiež je ortogonálna, $U^T U = I_r$.

Takýmto postupom zostrojená Householderova matica H spĺňa nasledujúcu Vetu.

Veta: Pre maticu $X \in \mathbb{R}^{m \times r}$ takú, že $X^T X = I_r$ existuje matica H skonštruovaná horeuvedeným spôsobom, ktorá transformuje maticu X na $HX = E_r(-WV)$. To znamená, že všetky riadky matice HX sú vynulované, až na prvých r riadkov. Matica prvých r riadkov HX má potom tvar $(-WV)$.

Dôkaz: Ak $Y = X + E_r WV$ a $\hat{X} = WDV$, potom $E_r^T Y = \hat{X} + WV = W(I_r + D)V$.

Zo vzťahu $(E_r WV)^T Y = V^T W^T (E_r^T Y) = V^T (I_r + D)V = \frac{1}{2} Y^T Y$ transponovaním dostaneme $Y^T E_r WV = \frac{1}{2} Y^T Y$ a podobne z $E_r WV = Y - X$ dostaneme $Y^T(Y - 2X) = 0$.

Pretože S nie je singulárne a $Y^T = S^T U^T$, $U^T(Y - 2X) = 0$; čiže $2U^T X = U^T Y$. Pretože $U^T Y = S$ a $US = Y$, $UU^T Y = Y$ a z toho $2UU^T X = U(2U^T X) = UU^T Y = Y$.

Z toho dôvodu $HX = X - 2UU^T X = X - Y = E_r(-WV)$, čo sme chceli dokázať. ■

Použitím tejto vety je matica H skonštruovaná z ortonormálnej matice X hodnosti r, preto $HX = E_r(-WV)$. Z toho dôvodu $HC = HXZ = E_r(-WV)Z = E_r\beta$, kde $\beta \in \mathbb{R}^{r \times b}$ je definované ako $\beta = (-WV)Z$, ktorá má plnú hodnosť t.j. $\mathrm{hod}(\beta) = r$. Vyplýva z toho, že z matice HC je matica v ktorej prvých r riadkov je matica β a ostatné riadky sú nulové.

3.5 Bloková tridiagonalizácia symetrickej matice

Aplikáciou blokovej Householderovej metódy z predchádzajúcej časti je napríklad bloková tridiagonalizácia symetrickej štvorcovej matice. Toto môže byť využité napríklad pri počítaní vlastných čísel vo veľkých maticiach, keďže Householderova metóda nemení vlastné čísla.

Uvažujme hustú symetrickú štvorcovú maticu A. Rozdelíme ju na menšie bloky rozmeru b. Označme ľavý horný blok s rozmermi $b \times b$ ako α. Potom blok ktorý sa nachádza pod maticou α je matica C ktorú budeme chcieť vynulovať

$$A = \begin{bmatrix} \alpha & | & & & \\ - & & - & - & - \\ & | & & & \\ C & | & & \tilde{A} & \\ & | & & & \end{bmatrix}.$$

Postup vynulovania matice C, čiže bloku matice A je analogický s postupom uvedeným v predchádzajúcej časti. Z matice C vynásobením príslušnou Householderovou maticou dostaneme maticu, v ktorej bude prvých r riadkov nenulových (matica β) a ostatné nulové

$$\begin{bmatrix} \\ C \\ \\ \end{bmatrix} \xrightarrow{\text{z } C \; na \; HC} \begin{bmatrix} \beta \\ - \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Vynásobením celej symetrickej matice A z oboch strán Householderovou maticou sa nám teda vynulujú bloky C smerom dole od bloku α a tiež smerom vpravo. Zároveň je týmto prenásobením zabezpečená nemennosť vlastných čísel (kapitola 1.3)

$$A \Rightarrow HAH = \begin{bmatrix} \alpha & | & \beta^T & | & 0 & \cdots & \\ - & & - & - & - & - & - \\ \beta & | & & & & & \\ - & | & & & & & \\ 0 & | & & & H\tilde{A}H & & \\ \vdots & | & & & & & \\ & | & & & & & \end{bmatrix}.$$

Opakovaním tohto postupu tridiagonalizujeme celú maticu A. Poznamenajme, že bloky nemusia mať rovnaký rozmer, t.j. každý diagonálny blok si môžeme rozmerovo prispôsobiť, či ľubovoľne zvoliť. Nakoniec dostaneme tridiagonalizovanú maticu tvaru

$$H^{(n-2)} \dots H^{(2)} H^{(1)} A H^{(1)} H^{(2)} \dots H^{(n-2)} =$$

$$
= \begin{bmatrix}
\alpha_1 & \beta_1^T & 0 & \cdots & & & & 0 \\
\beta_1 & \alpha_2 & \beta_2^T & 0 & \cdots & & & \\
0 & \beta_2 & \alpha_3 & \beta_3^T & & & & \\
\vdots & 0 & \beta_3 & \alpha_4 & \beta_4^T & & & \\
 & \vdots & & & \ddots & \ddots & & \\
 & & & & & \ddots & \ddots & \\
 & & & & & \beta_{n-2} & \alpha_{n-1} & \beta_{n-1}^T \\
0 & & & & & & \beta_{n-1} & \alpha_n
\end{bmatrix}.
$$

Využitie tejto metódy, aj ostatných predchádzajúcich si teraz ukážeme v nasledujúcej praktickej časti.

Praktická časť

4 Zjednodušenie tvaru matice

V praktickej časti chceme ukázať ako zjednodušiť vybrané matice. Vychádzame z toho, že v súčasnosti je široké využitie matíc v rôznych vedných disciplínach. Tieto matice bývajú väčšinou veľkých rozmerov a plné dát. Počítanie s takýmito maticami má svoju náročnosť. My chceme za pomoci blokov túto náročnosť zmenšiť pred samotným výpočtom. Uvažovať budeme však len štvorcové matice.

V prvom prípade chceme ukázať blokovú tridiagonalizáciu, ktorú spravíme za pomoci blokovej Householderovej metódy. Toto zjednodušenie sa dá využiť napríklad pri počítaní vlastných čísel, pretože úprava matice pomocou blokovej Householderovej metódy nemení hodnoty vlastných čísel. To znamená, že najskôr zjednodušíme tvar príslušnej matice a potom je oveľa jednoduchšie vypočítať vlastné čísla.

Druhé zjednodušenie bude využívať tiež blokovú Householderovu metódu a navyše Sylvestrove rovnice. Využitie tejto kombinácie by sa dalo uplatniť pri počítaní lineárnych systémov veľkých rozmerov. Tento postup zjednoduší maticu na blokový diagonálny tvar. Najskôr bloky ktoré sú pod hlavnou diagonálou vynulujeme pomocou blokovej Householderovej metódy a následne vynulujeme všetky bloky nad hlavnou diagonálou pomocou Sylvestrových rovníc. Je jednoduchšie potom počítať takýto lineárny systém napríklad paralelným výpočtom.

Pre oba prípady uvádzame aj krátke programy, ktoré ilustrujú zjednodušenie matíc.

4.1 Zjednodušenie tvaru veľkej matice pomocou blokovej Householderovej metódy pri počítaní vlastných čísel

V tejto časti ukážeme využitie blokovej Householderovej metódy. Predpokladajme, že máme ľubovoľnú štvorcovú symetrickú maticu $A \in \mathbb{R}^{n\times n}$ veľkých rozmerov a rozdelíme si ju na blokovú maticu

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{bmatrix}.$$

Pomocou blokovej Householderovej metódy by sme mali maticu A transformovať na tridiagonálny tvar

$$\begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{bmatrix} \Rightarrow \begin{bmatrix} \times & \times & & & 0 \\ \times & & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & & \times \\ 0 & & & \times & \times \end{bmatrix}.$$

Keďže máme štvorcovú a diagonálnu maticu, tak môžeme bez nejakých ďalších úprav použiť postup počítania ako v kapitole 3.4.3 na jednotlivé bloky, ktoré označíme C_i ($i \in \mathbb{N}$)

$$\left[\begin{array}{c|c|cccc} \times & & & C_1 & & \\ - & - & - & - & - & - \\ & \times & & & C_2 & \\ & - & & - & - & - \\ C_1 & & & & & \vdots \\ & C_2 & & & \ddots & \\ & & & \cdots & & \end{array}\right].$$

Budeme postupovať presne podľa postupu ako v kapitole 3.5, čiže budeme blokovo tridiagonalizovať maticu A pomocou vynulovaných blokov C_i . Schematicky naznačíme postup blokovej tridiagonalizácie

$$\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \end{bmatrix}}_{A} \Rightarrow \underbrace{\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{bmatrix}}_{H_1A} \Rightarrow \underbrace{\begin{bmatrix} \times & \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{bmatrix}}_{H_1AH_1} \Rightarrow$$

$$\Rightarrow \underbrace{\begin{bmatrix} \times & \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \end{bmatrix}}_{H_2H_1AH_1} \Rightarrow \underbrace{\begin{bmatrix} \times & \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \end{bmatrix}}_{H_2H_1AH_1H_2} \Rightarrow$$

$$\Rightarrow \ldots \Rightarrow \underbrace{\begin{bmatrix} \times & \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}}_{H_{n-2}\ldots H_2H_1AH_1H_2\ldots H_{n-2}}.$$

Kde matice H_i ($i \in \mathbb{N}$) sú Householderove blokové matice prislúchajúce k danému bloku C_i . Takže o funkčnosti metódy sme sa presvedčili a ukážeme ju aj na príklade. Konkrétne na matici rozmeru 16×16 , na ktorú sme použili program zostavený v programe MATLAB. Zdrojový kód je k nahliadnutiu v nasledujúcej časti a samotný príklad s maticou je v nasledujúcej kapitole 5.1 . Na základe poznatku, že vynásobenie ľubovoľnej matice ortogonálnou maticou zľava a inverznou orogonálnou sprava, nám zachováva vlastné čísla (kapitola 1.3), môžeme takúto zjednodušenú maticu blokového tridiagonálneho tvaru použiť na výpočet vlastných čísel.

4.1.1 Program a popis programu

```
function BlockHouseholder(A)

[m n]=size(A);
T=A;
for i=1:n/2-2
    C=T(2*i+1:n,2*i-1:2*i);
    [X,Z]=qr(C,0);
    Xhat=X(1:2,:);
    [W,D,V]=svd(Xhat);
    E=zeros(n-2*i,2);
    E(1,1)=1;
    E(2,2)=1;
    Y=X+E*W*V;
    U=Y*(V'*(2*(eye(2)+D))^(-0.5));
    H=eye(n-2*i)-2*U*U';
    Hpom=zeros(n);
    for j=1:2*i
        Hpom(j,j)=1;
    end
    Hpom(2*i+1:n,2*i+1:n)=H(:,:);
    Tnew=Hpom*T*Hpom;
    Tnew(abs(Tnew(:))<5e-12)=0;
    T=Tnew;
end
T
```

Uvedený program demonštruje postup opísaný v predchádzajúcej časti. Tento program počíta blokové Householderove matice podľa postupu v kapitole 3.4.3 a zároveň blokovo tridiagonalizuje pôvodnú maticu (kapitola 3.5). Vstupný údaj je len matica, ktorá musí byť štvorcová a symetrická. Výstupný údaj je blokovo tridiagonalizovaná matica. Program je skonštruovaný tak, že počíta po blokoch, kde šírka bloku je pevne daná hodnotou 2, čiže dva stĺpce. Z tohto vyplýva, že vstupná matica by mala mať párny počet stĺpcov. Použité sú predprogramované funkcie QR rozkladu a singulárneho rozkladu matice. Výstup tohto programu je uvedený v nasledujúcej kapitole 5.1 .

4.2 Zjednodušenie tvaru veľkej matice pomocou blokovej Householderovej metódy a Sylvestrových rovníc pri počítaní lineárneho systému

Zjednodušenie matice v tejto časti ukážeme pomocou dvoch metód. Na rozdiel od predchádzajúceho zjednodušenia, v tomto môžeme predpokladať ľubovoľnú štvorcovú maticu. Predpokladajme maticu $A \in \mathbb{R}^{n \times n}$ veľkých rozmerov, ktorú si rozdelíme na bloky

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{bmatrix}.$$

Najskôr použijeme blokovú Householderovu metódu, ktorá nám vynuluje všetky bloky pod diagonálou

$$\begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{bmatrix} \Rightarrow \begin{bmatrix} \times & \cdots & \times \\ & \ddots & \vdots \\ 0 & & \times \end{bmatrix}.$$

Keďže chceme vynulovať ľavú dolnú časť matice A, tak budeme vyberať bloky C_i ($i \in \mathbb{N}$) z ľavej dolnej časti matice vrátane diagonálnych blokov

$$\begin{bmatrix} & | & & & & \\ & | & - & & & \\ & | & & | & & \\ C_1 & | & & | & & \\ & | & C_2 & | & & \vdots \\ & | & & | & \ddots & \\ & | & & | & \cdots & \end{bmatrix}.$$

Postupom ako v kapitole 3.4.3 vypočítame k daným C_i príslušné Householderove matice H_i, ktoré zabezpečia vynulovanie pod hlavnou diagonálou

$$\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \end{bmatrix}}_{A} \Rightarrow \underbrace{\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{bmatrix}}_{H_1A} \Rightarrow \underbrace{\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \end{bmatrix}}_{H_2H_1A} \Rightarrow$$

$$\Rightarrow \ldots \Rightarrow \underbrace{\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix}}_{H_{n-1}\ldots H_2H_1A}.$$

Označme si maticu $F = H_{n-1} \ldots H_2 H_1 A$, čiže horná bloková trojuholníková matica. Našou nasledujúcou úlohou je ukázať, ako maticu T vynulujeme za pomoci Sylvestrových rovníc ako v kapitole 2.2 . Ukážeme si najskôr voľbu blokových matíc

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix} = \left[\begin{array}{c|c} A_{11} & C_1 \\ \hline 0 & B_1 \end{array}\right].$$

Kde A_{11} je štvorcová $k_1 \times k_1$ matica , B_1 je štvorcová $(n-k_1) \times (n-k_1)$ matica a C_1 je obdĺžniková $(n-k_1) \times k_1$ matica. Postupom ako je v kapitole 2.2 pomocou týchto matíc vynulujeme blok C_1. Zvolíme si nové bloky a pokračujeme až pokiaľ nedosiahneme blokový diagonálny tvar

$$\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix}}_{F} \Rightarrow \underbrace{\begin{bmatrix} \times & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix}}_{S_1FS_1} \Rightarrow \underbrace{\begin{bmatrix} \times & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix}}_{S_2S_1FS_1S_2} \Rightarrow$$

$$\Rightarrow \ldots \Rightarrow \underbrace{\begin{bmatrix} \times & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & 0 & 0 \\ 0 & 0 & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix}}_{S_{n-1} \ldots S_2 S_1 F S_1 S_2 \ldots S_{n-1}}$$

Matice S_i ($i \in \mathbb{N}$) sú matice transformácie, ktoré diagonalizujú maticu. Teoreticky by mal tento postup fungovať, prakticky sa presvedčíme v príklade so štvorcovou maticou 16×16 . Znovu sme použili program vytvorený v programe MATLAB, ktorého zdrojový kód uvádzame v ďalšej časti. Výsledok tohto programu, teda blokovo diagonalizovaná matica je k nahliadnutiu v kapitole 5.2 . Využitie tohto postupu by mohlo byť pri počítaní veľkých lineárnych systémoch. Zoberme si systém v tvare (1.2). Upravíme si najprv maticu A na blokový diagonálny tvar, nezabudneme príslušne upraviť aj pravú stranu rovnice b a môžeme počítať oveľa jednoduchší systém.

Takže namiesto systému

$$Ax = b\ ,$$

môžeme vďaka blokovej diagonalizácii použiť nezávislé bloky $\tilde{A}_{ii}$ transformovanej matice A na výpočet riešenia x pomocou paralelných výpočtov po blokoch

$$\tilde{A}_{11} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix}$$

$$\vdots$$

$$\tilde{A}_{88} \begin{bmatrix} x_{15} \\ x_{16} \end{bmatrix} = \begin{bmatrix} \tilde{b}_{15} \\ \tilde{b}_{16} \end{bmatrix}.$$

4.2.1 Program a popis programu

```
function BlockHouseholderAndSylvester(A)

[m n]=size(A);
T=A;
for i=1:n/2-1
    C=T(2*i-1:n,2*i-1:2*i);
    [X,Z]=qr(C,0);
    Xhat=X(1:2,:);
    [W,D,V]=svd(Xhat);
    E=zeros(n+2-2*i,2);
    E(1,1)=1;
    E(2,2)=1;
    Y=X+E*W*V;
    U=Y*(V'*(2*(eye(2)+D))^(-0.5));
    H=eye(n+2-2*i)-2*U*U';
    Hpom=zeros(n);
    for j=1:2*i
        Hpom(j,j)=1;
    end
    if i==1
        Tnew=H*T;
    else
        Hpom(2*i-1:n,2*i-1:n)=H(:,:);
        Tnew=Hpom*T;
    end
    Tnew(abs(Tnew(:))<5e-12)=0;
    T=Tnew;
end
T

ST=T;
for i=1:n/2-1
    SA=ST(1:2*i,1:2*i);
    SB=ST(2*i+1:n,2*i+1:n);
    SC=ST(1:2*i,2*i+1:n);
    SR=lyap(SA,-SB,SC);
    SIj=eye(n);
    SId=eye(n);
    SIj(1:2*i,2*i+1:n)=-SR(:,:);
    SId(1:2*i,2*i+1:n)=SR(:,:);
    STnew=SIj*ST*SId;
    STnew(abs(STnew(:))<5e-12)=0;
    ST=STnew;
end
ST
```

Tento program počíta blokovú diagonalizáciu opísanú v predchádzajúcej časti. Prvá časť je podobná s programom opísanom v kapitole 4.1.1 s tým rozdielom, že na konci programu násobí maticu len z ľavej strany a tým vynuluje všetky bloky pod diagonálou. V druhej časti program počíta diagonalizáciu pomocou Sylvestrových rovníc (kapitola 2.2). Vstupným údajom do programu je štvorcová matica a výstupom sú dve matice. Prvá je bloková horná trojuholníková a druhá je už bloková diagonálna. Aj tento program má pevne zadanú šírku blokov na 2, teda dva stĺpce. Znova teda predpokladáme , že vstupná matica bude mať párny počet stĺpcov. V programe sme okrem už spomínaných predprogramovaných funkcií QR rozkladu a singulárneho rozkladu použili funkciu Lyapunovej rovnice. Táto funkcia v programe MATLAB vie vypočítať práve Sylvestrove rovnice. Výstup tohto programu nájdeme v kapitole 5.2 .

5 Maticový experiment

V tejto kapitole uvádzame matice, na ktorých demonštrujeme blokovú tridiagonalizáciu a blokovú diagonalizáciu. Na začiatku je vždy uvedená matica, s ktorou sa počíta a následne sú uvedené výstupy programov už v upravených tvaroch.

5.1 Maticový experiment 1 (bloková tridiagonalizácia)

Matica, ktorá bude blokovo tridiagonalizovaná pomocou blokovej Householderovej metódy je štvorcová, rozmerov 16×16 a symetrická:

256	17	33	208	192	81	97	144	128	140	161	80	64	209	225	16
17	239	223	50	66	175	159	114	130	111	95	178	194	47	31	242
33	223	222	51	67	174	158	115	131	110	94	179	195	46	30	213
208	50	51	205	189	84	100	141	125	148	164	77	61	212	228	13
192	66	67	189	128	70	101	143	124	149	165	76	60	213	222	15
81	175	174	84	70	171	155	118	134	107	91	182	198	43	27	246
97	159	158	100	101	155	154	104	135	106	150	183	199	147	26	249
144	114	115	141	143	118	104	137	121	152	168	73	57	216	232	9
128	130	131	125	124	134	135	121	124	138	169	72	116	217	234	13
140	111	110	148	149	107	106	152	138	103	87	186	202	39	23	250
161	95	94	164	165	91	150	168	169	87	80	172	203	38	22	256
80	178	179	77	76	182	183	73	72	186	172	69	53	220	236	5
64	194	195	61	60	198	199	57	116	202	203	53	56	206	243	14
209	47	46	212	213	43	147	216	217	39	38	220	206	35	19	154
225	31	30	228	222	27	26	232	234	23	22	236	243	19	122	201
16	242	213	13	15	246	249	9	13	250	256	5	14	154	201	46

Matica, ktorá je blokovo tridiagonalizovaná programom popísaným v kapitole 4.1.1 :

$$
\begin{bmatrix}
256.0 & 17.0 & -140.6 & -514.9 & & & & & & & & & & & & \\
17.0 & 239.0 & -507.1 & -194.1 & & & & & & & & & & & & \\
-140.6 & -507.1 & 920.6 & 799.4 & -242.3 & 351.2 & & & & & & & & & & \\
-514.9 & -194.1 & 799.4 & 1095.1 & 126.3 & -316.2 & & & & & & & & & & \\
 & & -242.3 & 126.3 & -145.8 & 218.5 & -12.7 & 87.8 & & & & & & & & \\
 & & 351.2 & -316.2 & 218.5 & -371.9 & 131.3 & -13.5 & & & & & & & & \\
 & & & & -12.7 & 131.3 & 88.7 & -45.5 & -58.5 & 1.7 & & & & & & \\
 & & & & 87.8 & -13.5 & -45.5 & -3.5 & -12.9 & 50.7 & & & & & & \\
 & & & & & & -58.5 & -12.9 & -12.1 & 21.4 & -28.1 & 20.6 & & & & \\
 & & & & & & 1.7 & 50.7 & 21.4 & 31.3 & -29.6 & -46.8 & & & & \\
 & & & & & & & & -28.1 & -29.6 & 1.7 & 3.4 & -16.1 & 4.7 & & \\
 & & & & & & & & 20.6 & -46.8 & 3.4 & 32.7 & 3.5 & 16.1 & & \\
 & & & & & & & & & & -16.1 & 3.5 & 3.3 & -0.5 & -27.0 & 4.5 \\
 & & & & & & & & & & 4.7 & 16.1 & -0.5 & 14.2 & -7.6 & 8.0 \\
 & & & & & & & & & & & & -27.0 & -7.6 & -12.8 & 7.8 \\
 & & & & & & & & & & & & 4.5 & 8.0 & 7.8 & 10.5
\end{bmatrix}
$$

5.2 Maticový experiment 2 (bloková diagonalizácia)

Pôvodná matica, ktorá bude blokovo diagonalizovaná pomocou blokovej Householderovej metódy a Sylvestrových rovníc, je štvorcová s rozmermi 16×16 :

$$
\begin{bmatrix}
56 & 33 & 28 & 92 & 81 & 97 & 44 & 28 & 40 & 61 & 80 & 64 & 29 & 25 & 16 & 43 \\
17 & 39 & 50 & 66 & 75 & 59 & 14 & 30 & 11 & 95 & 78 & 94 & 47 & 31 & 42 & 46 \\
33 & 23 & 22 & 51 & 67 & 58 & 15 & 31 & 10 & 94 & 79 & 95 & 46 & 30 & 13 & 56 \\
38 & 50 & 51 & 25 & 19 & 84 & 10 & 41 & 48 & 64 & 77 & 61 & 12 & 28 & 13 & 72 \\
92 & 66 & 67 & 28 & 70 & 11 & 43 & 24 & 49 & 65 & 76 & 60 & 13 & 22 & 15 & 30 \\
81 & 74 & 84 & 70 & 71 & 55 & 18 & 34 & 17 & 91 & 82 & 98 & 43 & 27 & 46 & 21 \\
97 & 15 & 58 & 10 & 10 & 55 & 54 & 44 & 16 & 50 & 83 & 99 & 47 & 26 & 49 & 62 \\
44 & 14 & 15 & 43 & 18 & 14 & 37 & 21 & 52 & 62 & 73 & 57 & 26 & 32 & 69 & 88 \\
28 & 30 & 31 & 25 & 14 & 35 & 21 & 24 & 38 & 69 & 72 & 16 & 17 & 34 & 13 & 90 \\
40 & 11 & 10 & 48 & 19 & 57 & 16 & 52 & 38 & 13 & 86 & 22 & 39 & 23 & 50 & 17 \\
61 & 95 & 94 & 64 & 65 & 91 & 50 & 68 & 87 & 80 & 72 & 23 & 21 & 22 & 56 & 22 \\
80 & 18 & 79 & 77 & 76 & 82 & 83 & 73 & 72 & 86 & 72 & 69 & 20 & 36 & 85 & 46 \\
64 & 94 & 95 & 61 & 60 & 98 & 99 & 57 & 16 & 23 & 53 & 56 & 26 & 43 & 14 & 63 \\
49 & 47 & 46 & 12 & 21 & 43 & 47 & 16 & 17 & 39 & 38 & 20 & 26 & 35 & 54 & 43 \\
25 & 31 & 30 & 28 & 22 & 27 & 26 & 32 & 34 & 23 & 22 & 36 & 19 & 22 & 21 & 66 \\
55 & 41 & 10 & 23 & 12 & 74 & 56 & 52 & 64 & 73 & 92 & 62 & 29 & 26 & 51 & 26
\end{bmatrix}
$$

Prvá matica, ktorá je bloková horná trojuholníková, vypočítaná z predchádzajúcej matice programom popísaným v kapitole 4.2.1 :

-225.8	-131.3	-175.1	-149.1	-149.7	-193.9	-155.5	-139.4	-134.0	-209.3	-251.4	-217.7	-99.7	-95.6	-146.3	-156.4
-63.1	-150.6	-122.6	-86.8	-97.3	-121.3	-63.5	-65.6	-61.9	-101.5	-93.8	-65.8	-34.4	-48.4	-29.0	-67.1
		54.4	-39.6	-16.1	-31.0	5.0	-4.5	-29.3	30.9	-63.5	-23.0	-23.0	-10.0	-6.7	-14.5
		33.4	102.2	85.5	70.3	15.3	36.2	19.8	65.1	45.3	55.4	28.0	21.9	37.4	33.7
				-51.0	69.7	26.4	41.8	29.9	-15.1	27.8	-12.7	9.0	15.7	32.3	42.2
				23.1	60.8	11.1	14.9	-6.9	28.4	19.6	36.2	9.1	4.3	-22.5	-5.3
						32.4	-34.3	-40.9	-35.1	-39.6	-10.9	-10.6	3.6	-22.6	12.4
						-66.6	-25.4	-29.9	23.2	21.1	31.3	12.5	-8.3	-22.8	-5.6
								-64.1	-62.1	-21.9	11.9	18.8	-7.0	-27.0	-43.4
								-22.6	71.7	44.7	74.7	38.2	25.5	25.5	61.6
										31.3	36.9	7.1	12.4	-11.4	57.9
										27.2	-40.0	4.9	10.5	1.1	5.2
												21.1	-5.4	33.4	-18.0
												-13.4	-21.3	-15.3	-55.8
														-39.4	37.6
														18.0	-28.3

Druhá matica, teraz už blokovo diagonalizovaná, s nezávislými blokmi na diagonále:

$$
\begin{bmatrix}
-225.8 & -131.3 & & & & & & & & & & & & & & \\
-63.1 & -150.6 & & & & & & & & & & & & & & \\
 & & 54.4 & -39.6 & & & & & & & & & & & & \\
 & & 33.4 & 102.2 & & & & & & & & & & & & \\
 & & & & -51.0 & 69.7 & & & & & & & & & & \\
 & & & & 23.1 & 60.8 & & & & & & & & & & \\
 & & & & & & 32.4 & -34.3 & & & & & & & & \\
 & & & & & & -66.6 & -25.4 & & & & & & & & \\
 & & & & & & & & -64.1 & -62.1 & & & & & & \\
 & & & & & & & & -22.6 & 71.7 & & & & & & \\
 & & & & & & & & & & 31.3 & 36.9 & & & & \\
 & & & & & & & & & & 27.2 & -40.0 & & & & \\
 & & & & & & & & & & & & 21.1 & -5.4 & & \\
 & & & & & & & & & & & & -13.4 & -21.3 & & \\
 & & & & & & & & & & & & & & -39.4 & 37.6 \\
 & & & & & & & & & & & & & & 18.0 & -28.3
\end{bmatrix}
$$

Záver

Cieľom práce bolo popísať a ukázať dve metódy, ktoré transformujú matice na zjednodušené blokové matice. Ukázali sme princíp fungovania Householderovej metódy. Rozšírili sme ju na blokovú Householderovu metódu, ktorá vynuluje bloky matice. Ak máme ľubovoľnú štvorcovú maticu, pomocou tejto metódy vieme vynulovať bloky pod hlavnou diagonálou. Ak máme symetrickú štvorcovú maticu, tak vieme pomocou blokovej Householderovej metódy vynulovať bloky matice tak, že dostaneme blokový tridiagonálny tvar. Navyše v tomto prípade zostávajú zachované aj vlastné čísla, čo nám umožňuje využiť zjednodušený tvar matice na oveľa jednoduchší výpočet vlastných čísel. Fakt, že tieto dva postupy vynulovania blokov matice fungujú, sme sa presvedčili aj na konkrétnych príkladoch. Druhou metódou bola rovnako bloková metóda, ktorej výpočet bol na základe Sylvestrových rovníc. Oboznámili sme sa najprv so samotnými Sylvestrovými rovnicami. Rozšírená metóda blokovej Sylvestrovej rovnice nám umožňuje vynulovať bloky nad hlavnou diagonálou, avšak s tou podmienkou, že bloky pod diagonálou sú nulové. Túto podmienku vieme splniť spomínanou blokovou Householderovou metódou. Z toho nám vznikol postup, ktorým vieme maticu upraviť na blokový diagonálny tvar. Zo spomínaného vyplýva, že jednotlivé bloky budú nezávislé, čo môžeme využiť napríklad pri riešení lineárneho systému, ktorý riešime pomocou paralelných výpočtov. Blokovú diagonalizáciu matice sme tiež úspešne vyskúšali na konkrétnom príklade.

Použitá literatúra

[1] Bai Z. , Day D. , Demmel J. , Dongarra J. , *A test matrix collection for non-Hermitian eigenvalue problems* , release 1.0 , September 1996

[2] Godunov S.K. , *Lekcii po sovremennym aspektam linejnoj algebry* , Novosibirsk , Naučnaja Kniga , 2002

[3] Golub G.H. , Van Loan Ch.F. , *Matrix computations* , The Johns Hopkins University Press , 1996

[4] Guo C.-H. , Laub A.J. , *On a Newton-like method for solving algebraic Riccati equations* , SIAM J. Matrix Anal. Appl., 21 (2000), pp. 694-698

[5] Kotas P. , *Efektivní implementace singulárního rozkladu matice* , Ostrava , 2007

[6] Lacko V. , *Singulárny rozklad matíc a úskalia programovej realizácie Golubovho algoritmu na jeho určenie* , Bratislava , 2008

[7] Meyer C.D. , *Matrix Analysis and Applied Linear Algebra* , SIAM , 2001

[8] Murakami H. , *An Implementation of the Block Householder Method* , IPSJ Digital Courier Vol. 2 (2006) pp. 298-317

[9] Press W.H. , Teukolsky S.A. , Vetterling W.T. , Flannery B.P. , *Numerical recipes: The art of scientific computing* , Third edition , Cambridge University Press , 2007

[10] Watkins D.S. , *Fundamentals of Matrix Computations* , Second edition , John Wiley and Sons, 2002

[11] http://home.utah.edu/~nahaj/math/blockgivens.html