



Soustavy lineárních rovnic s hierarchickými maticemi řešené pomocí rozkladů

Diplomová práce

Studijní program: N1101 – Matematika
Studijní obory: 7504T089 – Učitelství matematiky pro střední školy
7503T009 – Učitelství anglického jazyka pro 2. stupeň základní školy
7504T – Profesní studium pro střední školy

Autor práce: **Barbora Košková**
Vedoucí práce: Martin Plešinger



Tento list nahrad'te
originálem zadání.

Prohlášení

Byla jsem seznámena s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé diplomové práce pro vnitřní potřebu TUL.

Užiji-li diplomovou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědoma povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Diplomovou práci jsem vypracovala samostatně s použitím uvedené literatury a na základě konzultací s vedoucím mé diplomové práce a konzultantem.

Současně čestně prohlašuji, že texty tištěné verze práce a elektronické verze práce vložené do IS STAG se shodují.

16. 4. 2020

Barbora Košková

Anotace

Při počítání praktických úloh např. z fyziky se často setkáváme s potřebou řešit soustavy lineárních rovnic s velkými řídkými maticemi. Ty se mohou řešit klasicky např. pomocí iteračních metod, kdy matice soustavy zůstává po celou dobu výpočtu nezměněná a využívá se pouze jejího součinu s vektorem. Druhou možností jsou metody založené na rozkladech matic, kde je ovšem potřeba pracovat chytře se zaplněním matice v průběhu úprav. Moderní přístup k těmto metodám mohou zprostředkovat tzv. hierarchické formáty uložení matice.

Tato diplomová práce si klade za cíl stručně zrekapitulovat základní manipulaci s hierarchickými maticemi a ukázat, jak lze řešit soustavu rovnic s takovou maticí pomocí rozkladových metod. Zaměříme se zejména na Gaußovu eliminaci, přesněji řečeno na její variantu pro symetrickou pozitivně definitní matici, tzv. Choleského rozklad. V závěru také ukážeme, jak pomocí hierarchického přístupu spočítat explicitně inverzi velké řídké matice.

Klíčová slova:

hierarchická matice; symetrická pozitivně definitní matice; LU rozklad; Gaußova eliminace; Choleského rozklad; maticové rovnice

Abstract

While dealing with practical real-world problems in, e.g., physics, it is often required to solve a large system of linear equations with a sparse matrix. Such linear problems can be solved in a classical way: By using iterative methods. In such methods, the matrix stays unchanged during the whole calculation, and it is employed repetitively only in evaluation matrix-vector products. The other way is represented by a bunch of methods based on matrix decompositions. During such decompositions, however, the matrix is changing and we have to use clever methods to avoid its fill-in with nonzeros. Hierarchical format for storing matrices is one of the modern approaches to do that.

The goals of this diploma thesis are to summarize basic arithmetics of hierarchical matrices, and to show how to solve a linear system with such matrix by using decompositions. We focus in particular on the Gaußian elimination method, more specifically, on its variant applicable to symmetric positive definite matrices, so-called Cholesky decomposition. Finally, we also show how to employ hierarchical approach to explicitly assemble an inverse to a large sparse matrix.

Key words:

hierarchical matrix; symmetric positive definite matrix; LU decomposition; Gaußian elimination; Cholesky decomposition; matrix equations

Poděkování

Ráda bych poděkovala, svému vedoucímu diplomové práce, panu Martinu Plešingerovi, za cenné rady, věcné připomínky a vstřícnost při konzultacích a vypracování diplomové práce.

Obsah

Anotace	4
Abstract	5
Seznam obrázků	9
Použité značení a zkratky	10
Úvod	12
1 Základní idea hierarchických matic	13
1.1 Hierarchické dělení	13
1.2 Příklad užití: Inverze	14
1.3 Náklady na uložení inverze	15
1.4 Obecná čtvercová matice	16
2 Aritmetika hierarchických matic	17
2.1 Sčítání a odečítání hierarchických matic	17
2.2 Násobení hierarchických matic	18
2.3 Inverze hierarchické matice	19
3 Choleského rozklad jako Gaußova eliminace symetrické pozitivně definitní matice	20
3.1 Základní operace Gaußovy eliminace	20
3.1.1 Násobení řádku nenulovým číslem	20
3.1.2 Prohození dvou řádků	21
3.1.3 Přičtení násobku řádku k jinému	21
3.2 Obecná struktura Gaußovy eliminace	22
3.2.1 Eliminace prvního sloupce	22
3.2.2 Eliminace ostatních sloupců a LU rozklad	23
3.3 Silně regulární matice	25
3.4 Symetrická pozitivně definitní matice	26
3.4.1 Vlastnosti SPD matic	27
3.4.2 LU rozklad SPD matice — první krok	28
3.4.3 LU rozklad SPD matice — druhý krok	29

4	Výpočet Choleského rozkladu a různé organizace výpočtu	31
4.1	Klasický výpočet Choleského rozkladu — výpočet po prvcích	31
4.1.1	Prvky prvního sloupce L	31
4.1.2	Prvky druhého sloupce L	32
4.2	Choleského rozklad — výpočet po sloupcích	33
4.3	Choleského rozklad — výpočet po blocích	35
5	Choleského rozklad matice v hierarchickém formátu	38
5.1	Formální hierarchická struktura rozkladu	38
5.2	Funkce $\text{Chol}(\cdot)$ v hierarchickém formátu	39
5.3	Mimodiagonální low-rank blok	40
5.4	Řešení soustavy s dolní trojúhelníkovou hierarchickou maticí	40
6	Řešení soustavy se symetrickou pozitivně definitní maticí v hierarchickém formátu	43
6.1	Převedení úlohy na řešení dvou trojúhelníkových soustav	43
6.2	Řešení soustavy s trojúhelníkovou hierarchickou maticí	44
7	Explicitní inverze řídké symetrické pozitivně definitní matice v hierarchickém formátu	45
7.1	Inverze rozložené matice	45
7.2	Inverze dolní trojúhelníkové hierarchické matice	45
7.3	Explicitní inverze celé matice	47
	Závěr	48
A	Apendix: Součiny a inverze trojúhelníkových matic	49
A.1	Součin dvou dolních trojúhelníkových matic je dolní trojúhelníkový .	49
A.2	Součin dvou horních trojúhelníkových matic je horní trojúhelníkový .	50
A.3	Inverze dolní trojúhelníkové matice je dolní trojúhelníková	50
A.4	Inverze horní trojúhelníkové matice je horní trojúhelníková	52
	Reference	53

Seznam obrázků

1.1	Hierarchické dělení třídiagonální matice	14
1.2	Matice v hierarchickém tvaru a její inverze	16
2.1	Schéma výpočtu inverze řídké matice v hierarchickém formátu	19
A.1	Součin dolních trojúhelníkových matic	50
A.2	Inverze dolní trojúhelníkové matice	51

Použité značení a zkratky

V textu značíme

vektory	pomocí malých písmen $u_1, u_2, u_r, v_1, v_2, v_r, x$, atd.,
matice a jejich bloky	pomocí velkých písmen (latinských i řeckých) A, B, C, D, E, F, U, V , atd.,

Pomocí malých písmen (latinských i řeckých) také značíme prvky matic a také skaláry. Speciální význam pak mají písmena i, j, ℓ , jimiž zpravidla indexujeme prvky matic, a k, m, n, r , která používáme k označení dimenze matice, resp. hodnosti (ranku) matice.

Matice a vektory

Značení	Význam
$A \in \mathbb{R}^{n \times m}$	reálná matice s rozměry n krát m , s prvky $a_{i,j}$
A^T	transpozice matice A
$\text{rank}(A)$	hodnost matice definovaná jako počet lineárně nezávislých řádků, resp. sloupců matice A
A^{-1}	inverze čtvercové regulární matice A
λ	vlastní číslo čtvercové matice A
$\det(A)$	determinant čtvercové matice A
I, I_n	jednotková matice, resp. jednotková matice řádu n
e_j	j -tý sloupec jednotkové matice I vhodného řádu
T_n	symetrická třídiagonální matice
$\alpha_1, \dots, \alpha_n$	prvky na diagonále T_n
β_2, \dots, β_n	prvky na první nad- a poddiagonále T_n
$\text{Chol}(\cdot)$	maticová funkce definovaná na SPD maticích, která vrací dolní trojúhelníkový Choleského faktor, $L = \text{Chol}(A)$, $A = LL^T$
$\text{Solve}(\cdot, \cdot)$	maticová funkce, která vrací řešení soustavy s dolní trojúhelníkovou maticí, $x = \text{Solve}(L, b)$, $Lx = b$

Použité zkratky a akronymy

Zkratka	Význam
SPD	symetrická pozitivně definitní matice
Ds	hustá (dense) matice
Sp	řídka (sparse) matice
LR	low-rank matice
Hi	hierarchická matice

Úvod

V dnešní době je nalezení řešení soustavy lineárních algebraických rovnic jedním z nezákladnějších úkolů v lineární algebře, a také problémem mnoha technických oborů. Toto téma se využívá nejen pro jeho aplikační část, ale také pro vysokou výpočetní složitost a širokou škálu metod a přístupů, které pomáhají řešit soustavy rychleji a přesněji, například v [12]. Na druhou stranu, o hierarchických maticích již z předchozích studií víme, že každá manipulace a použití je naopak náročné a zdoluhavé i pro moderní výpočetní techniky. Tato tvrzení jsou stěžejním bodem mnoha prací, například jsou podrobněji rozebrána v knize Stefana Pauliho [18]. Jeden konkrétní a přijatelný přístup pro řešení soustav s hierarchickou maticí $A \in \mathbb{R}^{n \times n}$, o které můžeme říci, že je symetricky pozitivně definitní, spočívá v jejím vhodném rozkladu. Konkrétně rozklad ve tvaru $A = LL^T$, kde L je dolní trojúhelníková matice, se nazývá Choleského rozklad. Při hledání takového rozkladu nesmí být opomenuto, že musí respektovat hierarchické dělení původní matice a musí s ní být kompatibilní. Z bakalářské práce [14] již víme, že bloky hierarchické matice jsou opět hierarchické, low-rank, nebo husté. Ke každému rozkladu odlišných typů bloků vede nejspíše rozdílná cesta. V této diplomové práci si představíme postup, jak k tomuto rozkladu dojít a dále ho využít.

V kapitole 1 představíme základní ideu hierarchické matice, konkrétně její dělení, inverzi a náklady na uložení. V kapitole 2 popíšeme základní aritmetiku hierarchických matic, kterou potřebujeme, abychom byli schopni s nimi dále manipulovat. Kapitola 3 se zabývá významem Choleského rozkladu z pozice Gaußovy eliminace symetrické pozitivně definitní matice. V této kapitole také ukážeme základní operace a vybrané vlastnosti zvolené matice. V následující kapitole 4 se budeme věnovat samotnému výpočtu Choleského rozkladu. Na tuto kapitolu přímo naváže kapitola 5, ve které bude hlavní myšlenkou získat Choleského rozklad v hierarchickém formátu. Získaný rozklad využijeme k řešení soustavy s hierarchickou maticí. V poslední kapitole 7 předvedeme, jak lze nalézt inverzi hierarchické matice právě pomocí jejího Choleského rozkladu.

1 Základní idea hierarchických matic

Uvedme, že jako „prototyp“ hierarchické matice budeme nyní uvažovat a již jsme dříve v bakalářské práci [14] uvažovali tzv. třídiagonální symetrickou matici respektive její inverzi, o které již víme, že ji lze chytře uložit pomocí bloků. Poznamenejme, že symetrická matice

$$T_n = \begin{bmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \beta_n & \\ & & & \beta_n & \alpha_n \end{bmatrix} \quad (1.1)$$

je určena jednoznačně právě $2n - 1$ čísly $\alpha_1, \dots, \alpha_n, \beta_2, \dots, \beta_n$. Při ukládání takovéto matice stačí uložit $(2n - 1) \sim n$ prvků namísto n^2 . Její inverze T_n^{-1} , jak lze snadno ověřit, je obecně hustá matice a na první pohled je potřeba uložit naopak právě n^2 (pomocí symetrie $\binom{n+1}{2} = \frac{1}{2}(n^2 + n) \sim \frac{1}{2}n^2$) prvků. Jak se ale ukazuje v [2], [7] a [14], paměť lze při vhodném způsobu ukládání výrazně ušetřit.

1.1 Hierarchické dělení

Předpokládejme symetrickou třídiagonální (a pro jednoduchost, vysvětlíme později, pozitivně definitní) matici o velikosti $n = 2^\ell$. V tento moment lze matici dělit na dvě poloviny následujícím způsobem

$$T_n = \begin{bmatrix} T_{n/2}^{(1)} & e_{n/2}\beta_{n/2+1}e_1^T \\ e_1\beta_{n/2+1}e_{n/2}^T & T_{n/2}^{(2)} \end{bmatrix} \in \mathbb{R}^{n \times n}. \quad (1.2)$$

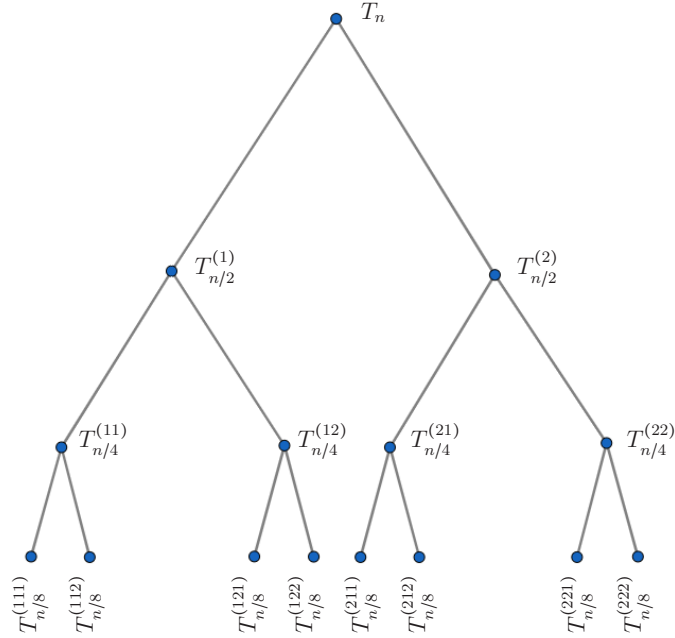
Každý diagonální blok $T_{n/2}^{(1)}$ a $T_{n/2}^{(2)}$ se dá rozdělit opět na půl, tedy

$$T_{n/2}^{(1)} = \begin{bmatrix} T_{n/4}^{(11)} & e_{n/4}\beta_{n/4+1}e_1^T \\ e_1\beta_{n/4+1}e_{n/4}^T & T_{n/4}^{(12)} \end{bmatrix} \in \mathbb{R}^{(n/2) \times (n/2)} \quad \text{a}$$

$$T_{n/2}^{(2)} = \begin{bmatrix} T_{n/4}^{(21)} & e_{n/4}\beta_{n/4+1}e_1^T \\ e_1\beta_{n/4+1}e_{n/4}^T & T_{n/4}^{(22)} \end{bmatrix} \in \mathbb{R}^{(n/2) \times (n/2)},$$

a tak dále. Obdobným způsobem budeme pokračovat v následujícím dělení, dokud to bude mít smysl. Nemá smysl dále dělit například matici 2×2 (matici 1×1 už dokonce dělit nelze). Matici tak rozdělíme na podmatice tvořící hierarchickou strukturu, kterou můžeme reprezentovat binárním stromem, viz obrázek 1.1. Více

podrobněji se tématice a pojmům dělení a tvorbě stromové struktury věnují české publikace [5], [6], [17], [19], [20], nebo obsáhlejší text [16]. Dále také můžeme doporučit [13]. Všechny větve stromu dále dělíme způsobem (pro dostatečně velké n ;



Obrázek 1.1: Hierarchické dělení symetrické třídiagonální matice (mimodiagonální bloky nejsou vyznačené). Obrázek převzat z [14].

zde konkrétně $n/256 \in \mathbb{N}$) například takto

$$T_{n/128}^{(1211212)} \longrightarrow \left(T_{n/256}^{(12112121)}, T_{n/256}^{(12112122)} \right).$$

1.2 Příklad užití: Inverze

Máme-li nyní symetrickou třídiagonální matici T_n , která je regulární (což je zde důsledkem pozitivní definitnosti), můžeme se ptát po její inverzi v hierarchickém tvaru

$$T_n^{-1} = \begin{bmatrix} E_{n/2}^{(1)} & F_{n/2} \\ F_{n/2}^T & E_{n/2}^{(2)} \end{bmatrix}. \quad (1.3)$$

Navíc hodnost mimodiagonálních bloků $F_{n/2}$ a $F_{n/2}^T$ je v tomto případě rovna jedné. Tyto vlastnosti můžeme na první pohled vidět, například napíšeme-li následující rovnost

$$\begin{bmatrix} I_{n/2} & 0 \\ 0 & I_{n/2} \end{bmatrix} = I_n = T_n T_n^{-1} = \begin{bmatrix} T_{n/2}^{(1)} & e_{n/2} \beta_{n/2+1} e_1^T \\ e_1 \beta_{n/2+1} e_{n/2}^T & T_{n/2}^{(2)} \end{bmatrix} \begin{bmatrix} E_{n/2}^{(1)} & F_{n/2} \\ F_{n/2}^T & E_{n/2}^{(2)} \end{bmatrix}.$$

Poté tedy platí, že

$$0 = (e_1 \beta_{n/2+1} e_{n/2}^T) E_{n/2}^{(1)} + T_{n/2}^{(2)} F_{n/2}^T,$$

a tedy po úpravě

$$\begin{aligned} F_{n/2}^T &= -(T_{n/2}^{(2)})^{-1} (e_1 \beta_{n/2+1} e_{n/2}^T) E_{n/2}^{(1)} \\ &= - \underbrace{[(T_{n/2}^{(2)})^{-1} e_1]}_u \beta_{n/2+1} \underbrace{[(E_{n/2}^{(1)})^T e_{n/2}]^T}_{v^T} \\ &= u \beta_{n/2+1} v^T. \end{aligned}$$

Tedy matice hodnosti jedna je napsána jako vnější součin dvou vektorů u a v . Samozřejmě toto můžeme provést za předpokladu regularity $T_{n/2}^{(2)}$ (regularita $T_{n/2}^{(1)}$ a $T_{n/2}^{(2)}$ zde opět plyne z pozitivní definitnosti). Poznamenejme, že detailněji viz [14]. Obdobným způsobem bychom se zaměřili na matice $E_{n/2}^{(1)}$ a $E_{n/2}^{(2)}$, které budou také opět hierarchické. V závěru zjistíme, že všechny mimodiagonální bloky matice mají hodnost jedna. V kapitole 2 se budeme podrobněji zabývat aritmetickými operacemi.

1.3 Náklady na uložení inverze

Co se týče paměťových nákladů na uložení inverze třídiagonální matice, standardně bychom potřebovali, jak už jsme zmínili, uložit n^2 prvků. Nyní si ukážeme, že náklady na uložení se dají podstatně ušetřit. Zaměříme se zejména na mimodiagonální bloky typu $F_{n/2}$. Za předchozích podmínek o velikosti jsou mimodiagonální bloky čtvercové a řádu $\frac{n}{2^j}$, $j \in \mathbb{N}$. První řádek největšího bloku má právě $\frac{n}{2}$ prvků, menší blok poté $\frac{n}{4}$ atd. To stejné lze tvrdit i o sloupcích. Uložení těchto bloků s hodnotou jedna obsahuje právě $n = \frac{n}{2} + \frac{n}{2}$, respektive $\frac{n}{2} = \frac{n}{4} + \frac{n}{4}$ a $\frac{n}{4} = \frac{n}{8} + \frac{n}{8}$ atd. čísel. Tím ale práce nekončí, z důvodu, že matic vzniká při dělení obecně více. Na první úrovni obdržíme pouze jednu mimodiagonální matici, protože druhou z důvodu symetrie zanedbáváme, na další úrovni již dvě, na třetí čtyři atd. Všechny mimodiagonální bloky inverze T_n^{-1} potřebují $1 \left(\frac{n}{2} + \frac{n}{2}\right) + 2 \left(\frac{n}{4} + \frac{n}{4}\right) + 4 \left(\frac{n}{8} + \frac{n}{8}\right) + \dots$ čísel. Jelikož původní matice T_n byla řádu $n = 2^\ell$ lze poslední sčítanec zapsat ve tvaru

$$2^{\ell-1} \left(\frac{n}{2^\ell} + \frac{n}{2^\ell} \right).$$

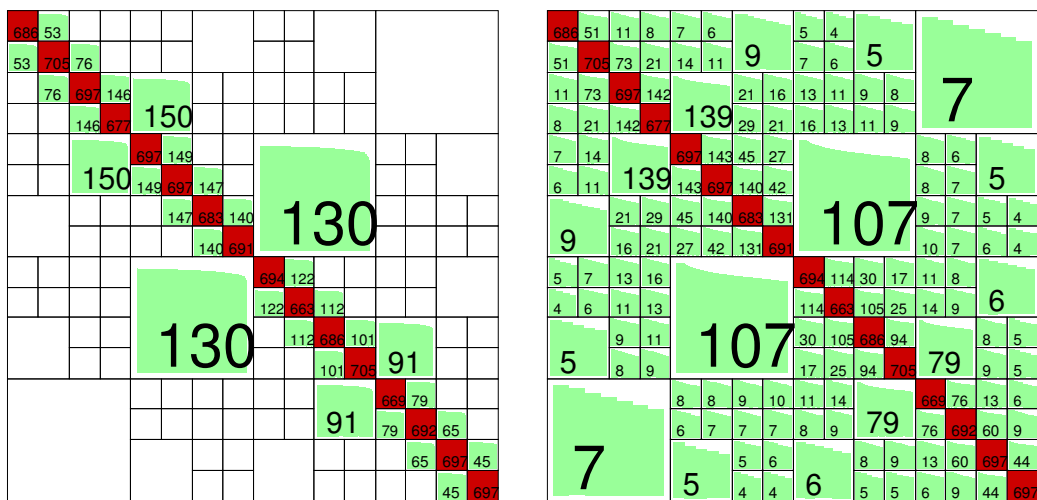
Počet sčítanců je roven ℓ a hodnota každého z nich je rovna n . Dohromady je potřeba

$$\underbrace{n + n + n + \dots + n + n}_\ell = n\ell + n = n(\log_2(n) + 1)$$

čísel. Součin $n\ell$ udává počet čísel v mimodiagonálních blocích, které mají hodnost rovnu jedné a zbývajících n čísel obsahuje prvky na diagonále.

1.4 Obecná čtvercová matice

Tento postup dělení nemůžeme ale vždy uplatňovat. Zejména na čtvercové matice, které mají obecné rozměry, viz [2], [3], [9] a [10]. V reálném případě dělíme množinu indexů na disjunktní podmnožiny, u kterých víme, že jejich sjednocením obdržíme počáteční množinu. Obdobným způsobem bychom postupovali u takto vytvořených podmnožin. Mimodiagonální bloky při takovémto dělení již nejsou čtvercové, princip ale zůstane zcela analogický. Poznamenejme, že ne každé dělení vede k ušetření paměti při ukládání. V principu lze vždy dělit, ale smysl to má jen když výrazně ušetříme místo (mimodiagonální bloky mohou být vyšší hodnoti viz obrázek 1.4). Právě z těchto důvodů je nejrozumnější rozdělit matici na půl, nebo při nejlepším zhruba na půl, neexistuje-li další východisko. Stromové dělení tedy nemusí být binární. Dodejme, že naše předešlá odvození byla pro symetrickou třídiagonální matici a její inverzi, ale ve skutečnosti jsou proveditelná pokaždé, když je možné hodnoti všech mimodiagonálních bloků určitým způsobem omezit.



Obrázek 1.2: Na obrázku vlevo můžeme vidět obecnou matici řádu $n = 11036$, vpravo pak její inverzi. Červené bloky na diagonále jsou uloženy přímo, nebo-li hustě. Číslo v každém červeném bloku označuje jeho řád a v zelenobílém bloku jeho hodnotu (značíme r). Tyto bloky jsou uloženy jako r lineárně nezávislých řádků a r lineárně nezávislých sloupců. Matice pocházejí z úlohy [15] a jsou vytvořené pomocí softwaru `hlib` [25] a Matlab.

2 Aritmetika hierarchických matic

Poznamenejme, že veškeré operace chceme provádět již přímo v hierarchickém tvaru. Z tohoto důvodu je vhodné připomenout, jak s těmito maticemi manipulovat. Dodejme, že hierarchická matice je matice, která je typicky rozdělena na 2×2 bloky. Důležité je kompatibilní dělení matice. Dodejme, že matici můžeme interpretovat i jako vektor s odpovídajícími vlastnostmi (např. musí být vhodně násobitelný).

Definice 1. *Matici nazveme hierarchickou, pokud je bloková a její bloky jsou buď:*

- *nízké hodnosti (značíme LR z anglického low-rank)*
- *hierarchické matice (značíme Hi)*
- *husté matice (značíme Ds z anglického dense)*

Poznámka 1. *Tato definice není zcela matematicky správná, protože k definování pojmu hierarchické matice používáme tentýž pojem. Tento fakt si uvědomujeme. Rigorózní definice by musela samotný pojem zavést hierarchicky. Taková definice by však byla nepřehledná. Proto dáváme přednost méně přesné a přehlednější, přesto však, jak doufáme dostatečně smysluplné definici.*

Poznámka 2. *Hustou maticí rozumíme matici, která je již na nejnižší úrovni a nepodléhá dalšímu dělení. Takovou matici tedy zpravidla ukládáme hustě, tzn. přesně „tak jak je“.*

2.1 Sčítání a odečítání hierarchických matic

Při sčítání (a odečítání; tj. přičítání opačného prvku) samozřejmě požadujeme stejné rozměry obou sčítaných matic. Ze stejného důvodu požadujeme i stejné rozměry jednotlivých bloků a do značné míry tedy i stejné stromové struktury. Pokud by stromové struktury byly zcela stejné (kompatibilní), neměla by nastat např. situace, kdy sčítáme $H_i + D_s$. Obecně však existuje šest možností, které při práci mohou nastat:

1. $H_i + H_i = H_i$
2. $H_i + LR = \dots$

3. $H_i + D_s = \dots$
4. $LR + LR = LR$
5. $LR + D_s = \dots$
6. $D_s + D_s = D_s$

U první možnosti předpokládáme, že matice mají kompatibilní stromovou strukturu dělení (situaci v opačném případě jsme popsali v [14]). Podrobněji se podíváme na čtvrtý případ, na součet typu $LR + LR$. Matice LR , která je nízké hodnoti, je zapsána ve tvaru UV^T . Navíc o jejich rozměrech víme, že matice $U \in \mathbb{R}^{n \times r}$ a matice $V \in \mathbb{R}^{m \times r}$, kde

$$r \ll \min\{n, m\}. \quad (2.1)$$

Součet dvou takových matic poté vypadá následujícím způsobem

$$(U_1V_1^T) + (U_2V_2^T) = [U_1U_2][V_1V_2]^T.$$

Příčemž blokové matice $[U_1U_2]$ a $[V_1V_2]^T$ můžeme tzv. komprimovat (např. pomocí ortogonalizace). Poslední případ, tj. součet dvou hustých matic je triviální.

Ve zbývajících situacích se jedná o smíšený součet. V případě takového součtu je potřeba si promyslet, jaký výsledek bychom rádi dostali. Například ve třetím případě je součet $H_i + D_s$, kde můžeme D_s převést na hierarchický tvar a sečíst jako $H_i + H_i$ nebo H_i můžeme převést na hustou matici a sečíst klasicky. Obdobně v případě součtu $LR + D_s$ můžeme na hustou matici nahlížet jako na matici nízké hodnoti, ale to je často nevýhodné. Pokud je jeden ze sčítanců hustý, nelze obecně nic usuzovat o jeho hodnoti, obecně tak může mít všechny řádky nebo sloupce lineárně nezávislé a jeho uložení v LR tvaru nic nešetříme. Typickým výsledkem tedy bude opět hustá matice. Analogicky by se vyřešila i druhá (vynechaná) situace.

2.2 Násobení hierarchických matic

U násobení, narozdíl od sčítání, mohou být velikosti matic odlišné, stačí pouze jeden stejný rozměr. Nemůžeme tedy zaručit ani stejné stromové struktury. Obecně existuje šest možností, které mohou nastat

1. $H_i \cdot H_i = H_i$
2. $H_i \cdot LR = LR$
3. $H_i \cdot D_s = D_s$
4. $LR \cdot LR = LR$
5. $LR \cdot D_s = LR$
6. $D_s \cdot D_s = D_s$

Detailněji si můžeme rozebrat součin $LR \cdot LR$, pro který platí, že jej lze rozepsat ve tvaru

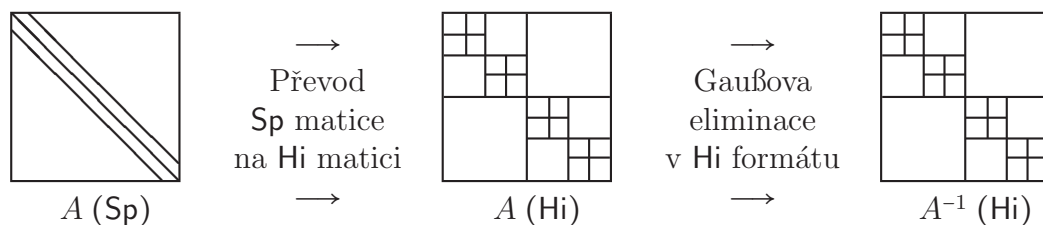
$$(U_1 V_1^T) \cdot (U_2 V_2^T) = (U_1 \underbrace{[V_1^T U_2]}_M) V_2^T,$$

kde matice M je typicky malých rozměrů, viz 2.1. Poté by v úpravě následoval krok komprese například podle [7] nebo [21, sekce 1.3.1].

Poznámka 3. *Uvědomme si, že speciálním případem je součin hierarchické matice a vektoru (matice o velikosti $n \times 1$), což je typický součin tvaru $Hi \cdot Ds$. Zpravidla nemá smysl vektor ukládat jinak, než jako hustou matici. Z tohoto důvodu je pak výsledkem takového součinu obecně hustá matice.*

2.3 Inverze hierarchické matice

K tomu, abychom uzavřeli kapitolu o aritmetice, čímž obvykle rozumíme soubor (respektive množinu čísel vybavenou souborem) operací *sčítání*, *odečítání*, *násobení* a *dělení*, bylo by vhodné umět spočítat inverzi v rámci hierarchického formátu. Ze základního kurzu lineární algebry víme, že inverzní matici můžeme nalézt např. pomocí Gaußovy eliminace. Takový postup v případě hierarchických matic je ilustrován na obrázku 2.1. O samotný převod původní matice do hierarchického tvaru se v tento moment nestaráme. Nás bude zajímat pouze získání inverze, která zde formálně nahrazuje dělení pomocí násobení zleva a zprava.



Obrázek 2.1: Řídkou (např. třídiagonální) matici $A (Sp)$ nejprve převedeme do hierarchického tvaru $A (Hi)$ a následně nalezneme její inverzi $A^{-1} (Hi)$, opět v hierarchickém tvaru. Obrázek převzat z [14].

Vidíme, že pro nalezení inverze bude v tomto případě klíčové zvládnout právě Gaußovu eliminaci v hierarchickém formátu.

3 Choleského rozklad jako Gaußova eliminace symetrické pozitivně definitní matice

Gaußova eliminace, také známá jako redukce řádků, je algoritmus v lineární algebře, který se hojně využívá pro řešení soustav lineárních rovnic. Obvykle se popisuje jako posloupnost kroků (operací) prováděných na odpovídající matici. Tuto metodu lze také využít k nalezení hodnoty matice, výpočtu determinantu i k nalezení inverzní čtvercové matice. Jak název navádí, metoda je pojmenována po matematikovi Johannu Carlu Friedrichovi Gaußovi, který jako první předložil potřebné důkazy její funkčnosti. V knize [4] je poznamenáno, že první zmínky o tomto algoritmu pochází z druhého století našeho letopočtu. Podrobněji se eliminaci věnují publikace například [7], [9] a [11].

Vycházejme z předpokladu, že daná matice A je regulární a označme její prvky obvyklým způsobem $a_{i,j}$

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}. \quad (3.1)$$

Na matici A budeme ilustrovat eliminaci. Poznamenejme, že Gaußova eliminace, která v nejběžnější podobě převádí matici A na horní trojúhelníkový tvar, úzce souvisí s tzv. LU rozkladem.

Nejprve se podíváme na obecnou strukturu Gaußovy eliminace, tedy pro obecnou čtvercovou regulární matici. Poté si objasníme význam silné regularity matice a její vliv na eliminaci. Nakonec se podíváme na (pro nás nejzajímavější) případ symetrické pozitivně definitní matice.

3.1 Základní operace Gaußovy eliminace

Obecně při eliminaci regulární matice provádíme tři typy kroků (operací), z nichž všechny lze vyjádřit pomocí násobení speciální maticí zleva.

3.1.1 Násobení řádku nenulovým číslem

První typickou operací je násobení j -tého řádku matice nenulovým číslem φ . Matice to provedeme tak, že naši matici A vynásobíme zleva skoro jednotkovou maticí,

kteřá má na příslušném řádku dané číslo φ . Schematicky

$$A \mapsto \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & \varphi & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{j,1} & \cdots & a_{j,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ \varphi a_{j,1} & \cdots & \varphi a_{j,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}.$$

Takovéto násobení se využívá zejména při ručních výpočtech, kdy je obtížné a časově náročné počítat například se zlomky a zápornými čísly. V našem případě není tento krok zásadní.

3.1.2 Prohození dvou řádků

Další typickou operací může být například prohození i -tého a j -tého řádku, $i < j$, které lze interpretovat pomocí maticového zápisu jako

$$A \mapsto \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & \cdots & 1 \\ & & \vdots & \ddots & \vdots \\ & & 1 & \cdots & 0 \\ & & & & \ddots \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & \ddots & \vdots \\ a_{j,1} & \cdots & a_{j,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{j,1} & \cdots & a_{j,n} \\ \vdots & \ddots & \vdots \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}.$$

Prohazování řádků se obecně nemůžeme vyhnout, stačí si uvědomit, jak by formálně probíhala eliminace (převedení na horní trojúhelníkový tvar) např. matice $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Nutnost prohazovat řádky úzce souvisí s tzv. silnou regularitou, jak uvidíme později, viz také [7, kapitola 4.2].

3.1.3 Přičtení násobku řádku k jinému

V poslední řadě nesmíme opomenout operaci přičtení α -násobku i -tého řádku k j -tému řádku, která je v eliminaci zpravidla nejdůležitější, protože nám umožní snadno vynulovat některý prvek v matici. Schematicky

$$A \mapsto \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \vdots & \ddots & \\ & & \alpha & \cdots & 1 \\ & & & & \ddots \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & \ddots & \vdots \\ a_{j,1} & \cdots & a_{j,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & \ddots & \vdots \\ \alpha a_{i,1} + a_{j,1} & \cdots & \alpha a_{i,n} + a_{j,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}.$$

Budeme-li chtít např. vynulovat první prvek v j -tém řádku, musí platit

$$\begin{aligned}\alpha a_{i,1} + a_{j,1} &= 0, \\ \alpha a_{i,1} &= -a_{j,1}, \\ \alpha &= -\frac{a_{j,1}}{a_{i,1}},\end{aligned}$$

samozřejmě za předpokladu $a_{i,1} \neq 0$. V případě, že daný předpoklad není splněn využijeme předešlý krok k nalezení vyhovujícího prvku.

Poznamenejme, že je možné uvažovat i další operace, my je ovšem v našem případě nebudeme potřebovat. Zejména z důvodu regularity dané matice, díky které např. nikdy nevznikne potřeba prohazovat sloupce. Shrneme-li předešlé kroky, jedná se vždy o manipulaci s maticí pomocí násobení různými maticemi. Násobení ovšem vždy aplikujeme zleva na původní matici.

3.2 Obecná struktura Gaußovy eliminace

Obecně chceme při eliminaci vynulovat všechny poddiagonální prvky. V prvním kroku budeme chtít vynulovat pouze ty, které jsou v prvním sloupci. Za předpokladu

$$a_{1,1} \neq 0$$

lze všechny tyto prvky vynulovat (eliminovat) jediným maticovým násobením, které bude sdružovat $(n - 1)$ výše zmíněných operací třetího typu.

Zjednodušeně lze o jednotlivém kroku eliminace říct, že matici A násobíme zleva příslušnou čtvercovou maticí, která je skoro jednotková, a navíc má ve vhodném sloupci pod hlavní diagonálou konkrétní prvky. Ty získáme jako podíl s opačným znaménkem prvků matice A v daném sloupci s prvkem na její hlavní diagonále, který náleží danému sloupci. Podstatou této práce není podrobné vysvětlení principu celé Gaußovy eliminace, uveďme proto pouze stručný náhled pro lepší přehlednost.

3.2.1 Eliminace prvního sloupce

Uvažujme matici G_1 , která je čtvercová a má rozměry jako matice A . Její strukturu lze zapsat v maticovém tvaru jako

$$G_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\frac{a_{2,1}}{a_{1,1}} & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \vdots \\ -\frac{a_{n,1}}{a_{1,1}} & 0 & 0 & 1 \end{bmatrix}, \quad \text{kde } a_{1,1} \neq 0. \quad (3.2)$$

Pokud vynásobíme matici A zleva touto maticí G , dojde k eliminaci prvků prvního sloupce pod hlavní diagonálou. Tedy:

- Prvek na hlavní diagonále zůstává nezměněn.
- Od druhého řádku odečítáme $(a_{2,1}/a_{1,1})$ násobek prvního řádku.

- Obecně od j -tého řádku, kde $j = 2, \dots, n$, odečítáme $(a_{j,1}/a_{1,1})$ násobek prvního řádku.

Maticově to lze zapsat ve tvaru

$$G_1 A = \begin{bmatrix} a_{1,1} & \cdots & \cdots & a_{1,n} \\ 0 & a_{2,2}^{(1)} & \cdots & a_{2,n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,2}^{(1)} & \cdots & a_{n,n}^{(1)} \end{bmatrix} \equiv A^{(1)}. \quad (3.3)$$

Výslednou matici jsme označili horním indexem (1) , čímž chceme vyjádřit, že se jedná o matici po prvním kroku eliminace (tj. po eliminaci prvního sloupce).

3.2.2 Eliminace ostatních sloupců a LU rozklad

Obdobným způsobem jako matici G_1 můžeme sestavit matice $G_2, G_3, G_4, \dots, G_{n-1}$. Například v druhém kroku eliminace odstraňujeme poddiagonální prvky druhého sloupce matice $A^{(1)}$ pomocí matice G_2 , která je v následujícím tvaru

$$G_2 = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & -\frac{a_{3,2}^{(1)}}{a_{2,2}^{(1)}} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & -\frac{a_{n,2}^{(1)}}{a_{2,2}^{(1)}} & 0 & \cdots & 1 \end{bmatrix}, \quad \text{kde } a_{2,2}^{(1)} \neq 0. \quad (3.4)$$

Dostali bychom matici $A^{(2)}$ a touto cestou bychom pokračovali až do $n - 1$ kroku. Po dokončení eliminace obdržíme horní trojúhelníkovou matici, kterou obecně značíme U ,

$$(G_{n-1} \cdots G_1) A = A^{(n-1)} \equiv U. \quad (3.5)$$

Všechny matice G_m , kde $m = 1, \dots, n - 1$, jsou dolní trojúhelníkové s jedničkami na diagonále, a tedy jsou regulární. Ke všem poté existuje inverze G_m^{-1} . Z rovnosti (3.5) můžeme vyjádřit matici A

$$\begin{aligned} G_{n-1}^{-1} \cdot \setminus \quad G_{n-1} \cdots G_1 A &= U, \\ G_{n-2}^{-1} \cdot \setminus \quad G_{n-2} \cdots G_1 A &= G_{n-1}^{-1} U, \\ &\vdots \\ G_1^{-1} \cdot \setminus \quad G_1 A &= G_2^{-1} \cdots G_{n-1}^{-1} U, \\ A &= G_1^{-1} \cdots G_{n-1}^{-1} U. \end{aligned}$$

Protože eliminační matice G_m je v dolním trojúhelníkovém tvaru, její inverze G_m^{-1} je také v dolním trojúhelníkovém tvaru, kde $m = 1, \dots, n - 1$. Důkaz tohoto tvrzení nalezneme např. v knize [1] a podrobněji je rozepsán v A.3. Konkrétně tedy platí

např.

$$G_1^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \frac{a_{2,1}}{a_{1,1}} & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \vdots \\ \frac{a_{n,1}}{a_{1,1}} & 0 & 0 & 1 \end{bmatrix}, \quad (3.6)$$

a obdobně u ostatních eliminačních matic.

Dále, součin dolních trojúhelníkových matic je opět dolní trojúhelníková matice viz appendix A.1, případně důkaz v [1]. Náš součin dolních trojúhelníkových matic budeme značit L (z anglického *lower triangular* podobně jako jsme dříve odvozenou horní trojúhelníkovou matici označili U , z anglického *upper triangular*). Platí tedy, že

$$L \equiv G_1^{-1} G_2^{-1} G_3^{-1} G_4^{-1} \cdots G_{n-1}^{-1}.$$

V důsledku pak lze matici A zapsat pomocí součinu dolní a horní trojúhelníkové matice jako

$$A = (G_1^{-1} \cdots G_{n-1}^{-1}) U = LU. \quad (3.7)$$

Podrobněji se Gaußově eliminaci a LU rozkladu věnuje [7].

V předešlém odstavci jsme vycházeli z předpokladu, že prvky matice A , kterými dělíme jsou nenulové. Obecně se ale mohou objevovat nuly, a tedy je potřeba aplikovat permutaci řádků. Existuje částečná a úplná permutace. První typ obsahuje záměnu řádků v matici $A^{(k-1)}$, kde $k = 1, \dots, n-1$. Druhý typ spočívá v přehození jak řádků, tak i sloupců matice $A^{(k-1)}$. Gaußovu eliminaci s částečnou permutací lze zapsat jako

$$PA = LU,$$

kde P je vhodná permutační matice. Pro regulární matici A lze Gaußovu eliminaci s částečnou permutací interpretovat jako Gaußovu eliminaci bez permutace aplikovanou na permutovanou matici PA , viz [7].

Poznámka 4. Vztahu $A = LU$ případně $PA = LU$ můžeme využít při řešení soustavy lineárních algebraických rovnic

$$\begin{aligned} Ax &= b, \\ P^T LUx &= b, \end{aligned}$$

kdy po případné permutaci

$$L \underbrace{Ux}_y = \underbrace{Pb}_c,$$

nejprve provedeme tzv. přímý chod Gaußovy eliminace, formálně vynásobíme rovnici zleva maticí L^{-1} , fakticky však řešíme soustavu s pravou stranou $c = Pb$, vektorem neznámých $y = Ux$ a dolní trojúhelníkovou maticí L . Což je opět jednoduché, jedná se pouze o dosazování. Získáme tedy rovnici ve tvaru

$$\begin{aligned} Ux &= y, \\ Ux &= L^{-1}c, \\ Ux &= L^{-1}Pb. \end{aligned}$$

Řešení této trojúhelníkové soustavy s maticí U se poté nazývá zpětný chod.

3.3 Silně regulární matice

V případě, že je matice tzv. silně regulární, v průběhu Gaußovy eliminace nenastávají situace, při kterých bychom museli provádět permutace řádků. To lze považovat za definici silné regularity. Formálně tedy lze takovou matici definovat tak, že prvky

$$a_{1,1} \equiv a_{1,1}^{(0)} \neq 0, \quad a_{2,2}^{(1)} \neq 0, \quad a_{3,3}^{(2)} \neq 0, \quad \dots \quad (3.8)$$

obecně

$$a_{j,j}^{(j-1)} \neq 0 \quad \text{pro} \quad j = 1, \dots, n, \quad (3.9)$$

vznikající v průběhu eliminace jsou nenulové; horní indexy zde opět značí jednotlivé kroky eliminace.

Je ale třeba vyjasnit, co silná regularita znamená. Předpokládejme tedy, že tato situace nastala, máme tedy LU rozklad matice A ve tvaru $A = LU$. Rozdělme si všechny tři matice A , L a U konzistentně na 2×2 bloky následujícím způsobem

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} = LU,$$

kde dále

$$A_{11}, U_{11}, L_{11} \in \mathbb{R}^{k \times k} \quad \text{pro nějaké } k = 1, 2, \dots, n-1.$$

Provedeme-li nyní blokové násobení obou matic, dostaneme

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11}U_{11} & L_{11}U_{12} \\ L_{21}U_{11} & L_{21}U_{12} + L_{22}U_{22} \end{bmatrix}.$$

Speciálně tedy platí rovnice

$$A_{11} = L_{11}U_{11} \in \mathbb{R}^{k \times k}$$

pro všechna $k = 1, \dots, n-1$. Protože matice L a U jsou dolní, respektive horní trojúhelníkové, pak také matice L_{11} a U_{11} jsou dolní, resp. horní trojúhelníkové. Tedy vidíme, že ze silné regularity matice A plyne, že všechny její tzv. hlavní rohové podmatice A_{11} jsou také (silně) regulární. Následně můžeme každou z nich napsat ve tvaru LU rozkladu. Tedy LU rozklad původní matice A je zároveň LU rozkladem všech jejich podmatic A_{11} .

Uvažujme nyní naopak regulární matici A , která má všechny podmatice $A_{11} \in \mathbb{R}^{k \times k}$ regulární. Z regularity A_{11} triviálně plyne nenulovost jejich determinantů. Pro všechny hlavní podmatice A_{11} naší matice A platí, že

$$D_k \equiv \det(A_{11}) = \det \left(\begin{bmatrix} a_{1,1} & \cdots & a_{1,k} \\ \vdots & \ddots & \vdots \\ a_{k,1} & \cdots & a_{k,k} \end{bmatrix} \right) \neq 0. \quad (3.10)$$

Podívejme se detailněji, co jsou čísla D_k zač:

- Pro $k = 1$ dostaneme triviálně

$$D_1 = a_{1,1} \neq 0.$$

Porovnáním s (3.8) vidíme, že $a_{1,1}^{(0)} \neq 0$.

- Pro $k = 2$ dostaneme

$$D_2 = a_{1,1} \cdot a_{2,2} - a_{2,1} \cdot a_{1,2} \neq 0.$$

S využitím nenulovosti $a_{1,1}$ můžeme nerovnost upravit následujícím způsobem

$$\frac{D_2}{D_1} = a_{2,2} - \frac{a_{2,1}}{a_{1,1}} \cdot a_{1,2} \neq 0, \quad \text{respektive}$$

$$\left[\begin{array}{cccccc} -\frac{a_{2,1}}{a_{1,1}} & 1 & 0 & \cdots & 0 & \\ & & & & & \end{array} \right] \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \\ \vdots \\ a_{n,2} \end{bmatrix} \neq 0.$$

Vidíme, že toto číslo můžeme napsat jako součin druhého řádku první eliminační matice G_1 (viz (3.2)) a druhého sloupce matice A , jedná se tedy právě o prvek v druhém řádku a druhém sloupci matice $A^{(1)}$, tedy o prvek $a_{2,2}^{(1)}$. Tedy $a_{2,2}^{(1)} \neq 0$.

- Obdobně lze analogicky ukázat, že

$$\frac{D_j}{D_{j-1}} = a_{j,j}^{(j-1)},$$

a tedy, že všechny prvky $a_{j,j}^{(j-1)} \neq 0$; viz [8].

3.4 Symetrická pozitivně definitní matice

Matici nazýváme symetrickou, když platí $A = A^T$. O matici dále říkáme, že je symetrická pozitivně definitní, když platí

$$A = A^T, \quad \text{a zároveň} \quad \forall x \neq 0, \quad x^T A x > 0.$$

Symetrická pozitivně definitní matice (dále SPD) má řadu zajímavých vlastností, například má kladná vlastní čísla nebo je silně regulární. Na tyto vlastnosti se blíže podíváme v následující podkapitole.

3.4.1 Vlastnosti SPD matic

Vybereme-li libovolné vlastní číslo λ a jemu odpovídající vlastní vektor v SPD matice A , platí vztah

$$Av = v\lambda, \quad v \neq 0.$$

Poté tuto rovnici vynásobíme transponovaným vlastním vektorem a pravou stranu vhodně uzavorkujeme. Získáme

$$v^T Av = v^T v\lambda = (v^T v)\lambda = \underbrace{\|v\|^2}_{> 0} \lambda > 0.$$

Kladnost normy plyne z nenulovosti vektoru a kladnost celého výrazu z pozitivní definitnosti matice. Z rovnice vyplývá vztah

$$\frac{v^T Av}{\|v\|^2} = \lambda > 0,$$

kde zlomek na levé straně je podíl dvou kladných čísel. Z tohoto vztahu tedy triviálně plyne, že vlastní čísla této matice A jsou kladná. Ve skutečnosti tato implikace platí i opačně. Přesněji řečeno symetrická matice s kladnými vlastními čísly je vždy pozitivně definitní, viz [7] a [8].

Víme, že determinant matice je součin vlastních čísel téže matice (viz opět [7] a [8]) v případě, že jsou všechna vlastní čísla kladná, pak platí

$$\det(A) = \prod_{j=1}^n \lambda_j \quad \implies \quad \det(A) > 0.$$

SPD matice má tedy i kladný determinant. Nyní si ukážeme, že totéž platí i pro její hlavní rohové poddeterminanty, tedy že je také silně regulární. Rozdělme si matici A , obdobně jako v silně regulárním případě, následujícím způsobem

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix} \quad \wedge \quad A_{11} = A_{11}^T, \quad A_{22} = A_{22}^T.$$

Z definice pozitivní definitnosti můžeme v kontextu tohoto dělení dostat například

$$x^T Ax = \underbrace{\begin{bmatrix} x_1^T & 0 \end{bmatrix}}_{x^T} \underbrace{\begin{bmatrix} A_{11} & A_{21}^T \\ A_{21} & A_{22} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ 0 \end{bmatrix}}_x = x_1^T A_{11} x_1 > 0. \quad (3.11)$$

Poznamenejme, že toto x je různé od nuly pokud platí

$$x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \neq 0 \quad \iff \quad x_1 \neq 0.$$

Z původního vztahu (3.11) následně vyplývá, že podmatice A_{11} (analogicky také A_{22}) je také symetrická pozitivně definitní. Následně i její determinant $\det(A_{11})$ je kladný.

Dostáváme tak velmi důležitou implikaci:

symetrická pozitivně definitní matice je silně regulární,

tedy všechny její hlavní rohové determinanty jsou nenulové, resp. dokonce *kladné*.

3.4.2 LU rozklad SPD matice — první krok

Gaußovu eliminaci lze tedy u symetrických pozitivně definitních matic provádět bez permutací řádků a sloupců. K takovéto matici A lze jednoznačně určit její LU rozklad tak, že

$$A = LU,$$

kde U je horní trojúhelníková regulární matice, z toho plyne, že prvky na diagonále $u_{1,1}, u_{2,2}, \dots, u_{n,n}$ jsou nenulové a L je dolní trojúhelníková s jednotkovou diagonálou. Tento rozklad lze přepsat do tvaru

$$A = L \underbrace{\text{diag}(u_{1,1}, u_{2,2}, \dots, u_{n,n})}_{D} \tilde{U}, \quad (3.12)$$

kde \tilde{U} je matice s jedničkami na hlavní diagonále, která vznikla vydělením j -tého řádku diagonálním prvkem $u_{j,j}$. Dostáváme tedy dvě rovnosti, které platí zároveň

$$A = LD\tilde{U} \quad \wedge \quad A = A^T,$$

musí pak také platit i

$$A^T = \tilde{U}^T D L^T.$$

Dosazením obou rozkladů do rovnice pro symetrii dostaneme

$$\begin{aligned} L^{-1} \cdot \setminus \quad LD\tilde{U} &= \tilde{U}^T D L^T, \\ D\tilde{U} &= L^{-1} \tilde{U}^T D L^T, \quad \setminus \cdot (L^T)^{-1} \\ D\tilde{U}(L^T)^{-1} &= L^{-1} \tilde{U}^T D, \quad \setminus \cdot D^{-1} \\ D\tilde{U}(L^T)^{-1} D^{-1} &= L^{-1} \tilde{U}^T. \end{aligned} \quad (3.13)$$

Podíváme-li se na rovnici po poslední úpravě, je zřejmé, že na levé straně rovnice jsou diagonální matice D a D^{-1} a horní trojúhelníkové matice \tilde{U} a $(L^T)^{-1}$. Součin na levé straně tedy dává horní trojúhelníkovou matici. Naopak, na pravé straně jsou dolní trojúhelníkové matice L^{-1} a \tilde{U}^T , jejichž součin tvoří opět dolní trojúhelníkovou matici. Rovnice bude platit tehdy a jen tehdy, pokud na levé i pravé straně bude právě diagonální matice.

Poznamenejme, že v úvaze výše jsme použili známá tvrzení, že

- (i) součin dvou horních (resp. dolních) trojúhelníkových matic je opět horní (resp. dolní) trojúhelníková matice (viz appendix A.2, resp. A.1) a
- (ii) inverze horní (resp. dolní) trojúhelníkové matice je opět horní (resp. dolní) trojúhelníková matice (viz appendix A.4, resp. A.3).

Dále je užitečné si všimnout, že

- (iii) součin dvou horních (resp. dolních) trojúhelníkových matic má na diagonále právě součiny odpovídajících diagonálních prvků jednotlivých součinitelů (viz

appendix A.1, respektive A.2). Tedy, že pro součin dolních (resp. horních) trojúhelníkových matic M a K

$$S = MK,$$

platí

$$s_{j,j} = m_{j,j}k_{j,j}, \quad \text{pro } j = 1, 2, \dots, n.$$

Připomeňme nyní, že matice \tilde{U} má na diagonále jedničky, které vznikly úpravou matice U . Matice L z LU rozkladu má také na diagonále jedničky. Z rovnice

$$LL^{-1} = I,$$

a výše uvedených vlastností (ii) a (iii) trojúhelníkových matic vyplývá, že matice L^{-1} má také na diagonále jedničky.

Podívejme se nyní podrobněji na pravou stranu posledního řádku rovnice (3.13). Součin matic L^{-1} a \tilde{U}^T musí být jednak diagonální maticí a za druhé bude mít jedničky na diagonále, tedy musí to být jednotková matice

$$L \cdot \begin{cases} L^{-1}\tilde{U}^T = I, \\ \tilde{U}^T = L. \end{cases}$$

Vidíme tedy, že SPD matici A lze vždy rozepsat ve tvaru

$$A = LDL^T,$$

kde L je dolní trojúhelníková matice s jednotkovou diagonálou.

3.4.3 LU rozklad SPD matice — druhý krok

V druhém kroku se v naší práci se symetrickou pozitivně definitní maticí zaměříme na diagonální matici $D = (u_{1,1}, u_{2,2}, \dots, u_{n,n})$. Využijeme vlastnosti

$$\forall x \neq 0, \quad x^T Ax > 0.$$

Pokud za matici A vhodně dosadíme, v našem případě součin LDL^T , dostaneme

$$x^T Ax = x^T (LDL^T)x = (x^T L)D(L^T x) = \underbrace{(L^T x)^T D (L^T x)}_y = y^T Dy > 0.$$

Protože L je dolní trojúhelníková s nenulovou (jednotkovou) diagonálou, je regulární. Tedy i L^T je regulární. Ze vztahu $y = L^T x$ a z regularity L^T triviálně plyne, že

$$x \neq 0 \iff y \neq 0.$$

Z toho vidíme, že matice A je pozitivně definitní, tehdy a jen tehdy, když je matice D pozitivně definitní. Protože D je diagonální, je také symetrická pozitivně definitní.

Navíc, tím že D je diagonální, má na diagonále přímo svá vlastní čísla. Z toho, že každá SPD matice má kladná vlastní čísla (jak jsme již ukázali), plyne, že matice D má kladná čísla na diagonále.

Konkrétně to můžeme vidět při speciální volbě vektoru x respektive y , kdy

$$x = (L^T)^{-1}y, \quad y = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = e_j.$$

Zřejmě tak dostáváme

$$x^T Ax = y^T Dy = e_j^T D e_j = u_{j,j} > 0,$$

viz (3.12). Znovu tedy vidíme, že diagonální matice D má na diagonále kladná čísla.

Z tohoto důvodu si matici D můžeme speciálně napsat jako

$$D = \sqrt{D}\sqrt{D},$$

kde matice \sqrt{D} je následujícího tvaru

$$\sqrt{D} = \begin{bmatrix} \sqrt{u_{1,1}} & & & \\ & \sqrt{u_{2,2}} & & \\ & & \ddots & \\ & & & \sqrt{u_{n,n}} \end{bmatrix}.$$

Navíc platí, že $(\sqrt{D})^T = \sqrt{D}$ je symetrická. Vrátime-li se zpět ke vztahu $A = LDL^T$, můžeme ho nyní upravit

$$A = LDL^T = L(\sqrt{D}\sqrt{D})L^T = (L\sqrt{D})(\sqrt{D}L^T) = \underbrace{(L\sqrt{D})(L\sqrt{D})^T}_{\tilde{L}} = \tilde{L}\tilde{L}^T.$$

Věta 1. *Nechť A je symetrická pozitivně definitní matice. Pak existuje dolní trojúhelníková matice \tilde{L} s kladnými čísly na diagonále taková, že*

$$A = \tilde{L}\tilde{L}^T.$$

Tento rozklad je jednoznačný a nazývá se Choleského rozklad.

Důkaz první části tvrzení plyne z předchozích úvah. Důkaz jednoznačnosti nalezneme například v [7].

4 Výpočet Choleského rozkladu a různé organizace výpočtu

V kapitole 3 jsme ukázali, že existuje Choleského rozklad symetrické pozitivně definitní matice A . Nyní se blíže podíváme na to, jak ho spočítat. Podíváme se, jak provést tzv. *klasický výpočet Choleského rozkladu*, tedy po jednotlivých prvcích matice. Podíváme se však i na další varianty, jak jej spočítat tzv. *po sloupcích*, a protože se chystáme na hierarchický formát matice, který je principiálně blokový, podíváme se také na výpočet tzv. *blokového Choleského rozkladu*, kdy máme matici A ve tvaru

$$A = \begin{bmatrix} A_{11} & A_{21}^T \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} L_{11}^T & L_{21}^T \\ 0 & L_{22}^T \end{bmatrix}.$$

Pro jednoduchost budeme v této a dalších kapitolách dolní trojúhelníkovou matici \tilde{L} značit již bez vlnky.

4.1 Klasický výpočet Choleského rozkladu — výpočet po prvcích

Nejprve se podíváme na klasický výpočet Choleského rozkladu matice. Matici A si rozepíšeme do tvaru

$$\begin{bmatrix} a_{1,1} & a_{2,1} & a_{3,1} & \dots & a_{n,1} \\ a_{2,1} & a_{2,2} & a_{3,2} & \ddots & \vdots \\ a_{3,1} & a_{3,2} & a_{3,3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n,1} & \dots & \dots & \dots & a_{n,n} \end{bmatrix} = \begin{bmatrix} l_{1,1} & 0 & \dots & \dots & 0 \\ l_{2,1} & l_{2,2} & \ddots & \ddots & \vdots \\ l_{3,1} & l_{3,2} & l_{3,3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ l_{n,1} & \dots & \dots & \dots & l_{n,n} \end{bmatrix} \begin{bmatrix} l_{1,1} & l_{2,1} & l_{3,1} & \dots & l_{n,1} \\ 0 & l_{2,2} & l_{3,2} & \ddots & \vdots \\ \vdots & \ddots & l_{3,3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & l_{n,n} \end{bmatrix}.$$

Maticová rovnost platí, pokud se prvky matice na levé straně rovnají odpovídajícím prvkům matice (součin matic) na pravé straně. Nyní se zaměříme na jednotlivé rovnosti a začneme v levém horním rohu.

4.1.1 Prvky prvního sloupce L

Z prostého vynásobení matic LL^T na pravé straně a porovnáním s maticí A ihned plyne, že pro první prvek v prvním sloupci platí

$$a_{1,1} = l_{1,1}l_{1,1},$$

tedy

$$l_{1,1} = \sqrt{a_{1,1}}.$$

Poznámka 5. Zároveň vidíme, že začínáme podnikat první kroky k důkazu jednoznačnosti Choleského rozkladu. Mohli bychom si také vybrat $-\sqrt{a_{1,1}}$, pak by ale matice L neměla kladné prvky na diagonále. Prvek $l_{1,1}$ je daný jednoznačně.

Dále budeme pokračovat v hledání prvků prvního sloupce L . Zde platí

$$a_{2,1} = l_{2,1}l_{1,1} \implies l_{2,1} = \frac{a_{2,1}}{l_{1,1}}.$$

Protože již známe $l_{1,1}$ můžeme vyjádřit prvek $l_{2,1}$. Stejným způsobem bychom pokračovali dál. V j -tém kroku dostáváme

$$a_{j,1} = l_{j,1}l_{1,1} \implies l_{j,1} = \frac{a_{j,1}}{l_{1,1}}, \quad j = 2, \dots, n.$$

První sloupec matice L máme celý vyjádřený. Postupně již vyjádřené prvky nám dále pomáhají.

4.1.2 Prvky druhého sloupce L

V druhém kroku, tedy v druhém sloupci, díky symetrii začínáme od druhého prvku, kde

$$a_{2,2} = l_{2,1}l_{2,1} + l_{2,2}l_{2,2} \implies l_{2,2} = \sqrt{a_{2,2} - l_{2,1}^2}.$$

Poznámka 6. Může zde zaznít otázka, zda číslo pod odmocninou je kladné (nezáporné, abychom ho mohli odmocnit, a navíc protože s odmocněným číslem budeme dále dělit, také nenulové). Předchozí prvky se ovšem počítaly jednoznačně a navíc jsme v předchozí kapitole ukázali, že rovnost existuje. Je možné také provést důkaz sporem: Nechtě

$$\begin{aligned} a_{2,2} - l_{2,1}^2 &\leq 0, \\ a_{2,2} &\leq l_{2,1}^2 = \left(\frac{a_{2,1}}{l_{1,1}}\right)^2 = \frac{a_{2,1}^2}{a_{1,1}}, & / \cdot a_{1,1} \\ a_{1,1}a_{2,2} &\leq a_{2,1}^2, & / - a_{2,1}^2 \\ \underbrace{a_{1,1}a_{2,2} - a_{2,1}^2}_{\det \begin{pmatrix} a_{1,1} & a_{2,1} \\ a_{2,1} & a_{2,2} \end{pmatrix}} &\leq 0, \end{aligned}$$

ovšem matice A je pozitivně definitní, a tak všechny její determinanty jsou kladné. \square

Dále bychom pokračovali zcela obdobným způsobem, a to

$$a_{3,2} = l_{3,1}l_{2,1} + l_{3,2}l_{2,2}.$$

Obecně v j -tém kroku

$$a_{j,2} = l_{j,1}l_{2,1} + l_{j,2}l_{2,2}, \quad j = 3, \dots, n.$$

Neznámý prvek $l_{j,2}$ bychom vyjádřili pomocí již známých prvků matice L , konkrétně pomocí $l_{j,1}$, $l_{2,1}$, $l_{2,2}$, kde

$$l_{j,2} = \frac{a_{j,2} - l_{j,1}l_{2,1}}{l_{2,2}}.$$

Analogicky bychom postupovali pro všechny poddiagonální prvky v druhém sloupci.

Tímto způsobem bychom mohli vyjádřit prvky postupně všech sloupců. Například u prvku $a_{3,3}$ máme již součet tří součinů

$$a_{3,3} = l_{3,1}l_{3,1} + l_{3,2}l_{3,2} + l_{3,3}l_{3,3} \implies l_{3,3} = \sqrt{a_{3,3} - l_{3,1}^2 - l_{3,2}^2},$$

analogicky bychom mohli diskutovat kladnost argumentu pod odmocninou, a následně výpočet poddiagonálního prvku. Všechny prvky matice L je tedy možné vyjádřit například pomocí algoritmu 1.

```

for  $i = 1, 2, \dots, n$  do
   $l_{i,i} = (a_{i,i} - \sum_{s=1}^{i-1} l_{i,s}^2)^{\frac{1}{2}}$ 
  for  $j = i + 1, \dots, n$  do
     $l_{j,i} = (a_{j,i} - \sum_{s=1}^{i-1} l_{j,s}l_{i,s})/l_{i,i}$ 
  end
end

```

Algoritmus 1: Choleského rozklad spočítaný klasicky, po prvcích.

4.2 Choleského rozklad — výpočet po sloupcích

Další možnost pro řešení je, pokud matice A a L rozdělíme na sloupce jako

$$L = [l_1, l_2, \dots, l_n],$$

$$A \equiv A^{(0)} = [a_1^{(0)}, a_2^{(0)}, a_3^{(0)}, \dots, a_n^{(0)}].$$

Následně si v prvním kroku vypočítáme $l_{1,1}$ jako $l_{1,1} = \sqrt{a_{1,1}^{(0)}}$. A poté celý sloupec l_1 získáme jako

$$l_1 = a_1^{(0)} \frac{1}{l_{1,1}} = a_1^{(0)} \frac{1}{\sqrt{a_{1,1}^{(0)}}}.$$

Můžeme tedy vytvořit matici $A^{(1)}$ pomocí již vypočítaného sloupce l_1 z prvního kroku, kde

$$A^{(1)} \equiv A - l_1 l_1^T. \quad (4.1)$$

Výsledná struktura $A^{(1)}$ bude

$$A^{(1)} = \begin{bmatrix} a_{1,1} - l_{1,1}l_{1,1} & a_{2,1} - l_{1,1}l_{2,1} & \cdots & a_{n,1} - l_{1,1}l_{n,1} \\ a_{2,1} - l_{2,1}l_{1,1} & a_{2,2} - l_{2,1}l_{2,1} & \cdots & a_{n,2} - l_{2,1}l_{n,1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} - l_{n,1}l_{1,1} & a_{n,2} - l_{n,1}l_{2,1} & \cdots & a_{n,n} - l_{n,1}l_{n,1} \end{bmatrix}.$$

Podíváme-li se nyní blíže na prvky v prvním sloupci (ze symetrie v prvním řádku), vidíme, že budou rovny nule, jelikož z 4.1.1 víme, že prvky

$$a_{n,1} = l_{k,1}l_{1,1}, \quad k = 1, \dots, n.$$

Výsledkem rozdílů v prvním sloupci a řádku, tedy budou nuly a matici $A^{(1)}$ lze po úpravě zapsat jako

$$A^{(1)} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & a_{2,2} - l_{2,1}l_{2,1} & \cdots & a_{n,2} - l_{2,1}l_{n,1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,2} - l_{n,1}l_{2,1} & \cdots & a_{n,n} - l_{n,1}l_{n,1} \end{bmatrix} = [0, a_2^{(1)}, a_3^{(1)}, \dots, a_n^{(1)}].$$

První sloupec bude nulový a zbývající sloupce jsou odlišné oproti původní matici A . Dále bychom pokračovali obdobným způsobem, tentokrát ale již s maticí $A^{(1)}$. Provedli bychom vyjádření druhého sloupce pomocí prvku $l_{2,2}$, který je roven $\sqrt{a_{2,2}^{(1)}}$. Tedy v druhém kroku vypočítáme

$$l_2 = a_2^{(1)} \frac{1}{\sqrt{a_{2,2}^{(1)}}},$$

a následně

$$A^{(2)} \equiv A^{(1)} - l_2 l_2^T = A - l_1 l_1^T - l_2 l_2^T.$$

Tímto způsobem bychom pokračovali dále a vytvořili $n - 2$ dalších kroků. Došli bychom až k matici

$$A^{(n-1)} \equiv A^{(n-2)} - l_{n-1} l_{n-1}^T,$$

kteřá je až na poslední diagonální prvek skoro nulovou maticí

$$A^{(n-1)} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n}^{(n-1)} \end{bmatrix}.$$

Platí tedy

$$A^{(n-1)} = l_n l_n^T,$$

kde poslední sloupec l_n matice L lze zapsat jako

$$l_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \sqrt{a_{n,n}^{(n-1)}} \end{bmatrix}.$$

Tím jsme vyjádřili všechny sloupce matice L .

Matici A lze nakonec zapsat jako součet vnějších součinů sloupců matice L , postupně

$$\begin{aligned} A^{(n-1)} &= l_n l_n^T, \\ A^{(n-2)} &= A^{(n-1)} + l_{n-1} l_{n-1}^T, \\ &\vdots \\ A^{(2)} &= A^{(3)} + l_3 l_3^T, \\ A^{(1)} &= A^{(2)} + l_2 l_2^T, \\ A^{(0)} &= A^{(1)} + l_1 l_1^T, \end{aligned}$$

tedy

$$A = A^{(0)} = l_1 l_1^T + l_2 l_2^T + l_3 l_3^T + \dots + l_n l_n^T = \sum_{i=1}^n l_i l_i^T.$$

Dostaneme sloupcový Choleského algoritmus, viz algoritmus 2.

```

for  $i = 1, 2, \dots, n$  do
   $l_i = a_i^{(i-1)} \cdot \left(1 / \sqrt{a_{i,i}^{(i-1)}}\right)$ 
   $A^{(i)} = A^{(i-1)} - l_i l_i^T$ 
end

```

Algoritmus 2: Choleského rozklad spočítaný po sloupcích.

4.3 Choleského rozklad — výpočet po blocích

Nyní se podíváme, jak bychom postupovali, kdybychom si matici A rozdělili již známým způsobem na bloky, konkrétně

$$A = \begin{bmatrix} A_{11} & A_{21}^T \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} L_{11}^T & L_{21}^T \\ 0 & L_{22}^T \end{bmatrix} = LL^T. \quad (4.2)$$

Pokud bychom naivně postupovali stejně, jako v klasickém Choleského rozkladu (tedy pokud bychom na všechny bloky na chvíli koukali jako na čísla) mělo by *formálně* platit jednak

$$„ L_{11} = \sqrt{A_{11}} „, \quad (4.3)$$

a pak také

$$L_{21} = \frac{A_{11}}{L_{11}} \quad (4.4)$$

Nastává ale otázka, jak *správně interpretovat* případně *definovat odmocninu* z matice v první rovnosti, případně *dělení dvou matic* v rovnosti druhé.

Pokud však rozklad (4.2) rozepíšeme pro jednotlivé bloky dostaneme rovnice

$$\begin{aligned} A_{11} &= L_{11}L_{11}^T, \\ A_{21} &= L_{21}L_{11}^T, \quad \text{a nakonec} \\ A_{22} &= L_{21}L_{21}^T + L_{22}L_{22}^T. \end{aligned}$$

Konkrétně v případě podmatice A_{11} víme, že je symetrická pozitivně definitní. Tedy odmocninu můžeme nahradit právě Choleského rozkladem této podmatice, ostatně

$$A_{11} = L_{11}L_{11}^T \sim L_{11}^2.$$

Odmocninu z SPD matice (4.3) tedy zadefinujeme pomocí maticové funkce

$$L_{11} = \text{Chol}(A_{11}),$$

kteřá vrátí Choleského rozklad (podmatice, resp. bloku), přesněji dolní trojúhelníkový faktor rozkladu; technicky ho může počítat právě klasickým algoritmem (po prvcích nebo po sloupcích). Následně si rozepíšeme vztah pro mimodiagonální blok, kde

$$\begin{aligned} A_{21} &= L_{21}L_{11}^T, \quad / \cdot (L_{11}^T)^{-1} \\ A_{21}(L_{11}^T)^{-1} &= L_{21}. \end{aligned}$$

Vidíme, že dělení ze vztahu (4.4) ve skutečnosti znamená násobení transpozice čitatele („dělece“) inverzí („dělitele“) zprava.

Poznámka 7. *Ve skutečnosti můžeme násobení inverzí nahradit řešením soustavy rovnic. Zřejmě když A_{21} bude low-rank. Pro přehlednost celou rovnost*

$$A_{21} = L_{21}L_{11}^T,$$

transponujeme (neboli vyjádříme druhý mimodiagonální blok)

$$L_{11}L_{21}^T = A_{21}^T = uv^T.$$

Pokud transponujeme součin, dostáváme součin transpozic v opačném pořadí. Rovnici můžeme vynásobit zleva maticí L_{11}^{-1} a dostáváme

$$\begin{aligned} L_{21}^T &= L_{11}^T A_{21}^T, \\ L_{21}^T &= L_{11}^{-1}(uv^T) = \underbrace{(L_{11}^{-1}u)}_x v^T. \end{aligned}$$

Díky asociativitě násobení matic jsme vhodně přezávorkovali součin a obsah v závorce označili jako vektor x . Platí tedy

$$\begin{aligned} L_{11} \cdot x &= L_{11}^{-1}(uv^T), \\ L_{11}x &= u. \end{aligned}$$

Poslední rovnice $L_{11}x = u$ je fakticky klasická soustava rovnic s vektorovou pravou stranou u , navíc s dolní trojúhelníkovou maticí, což je jednoduché řešit dosazovací metodou shora dolů. Z vypočteného x následně vidíme, že

$$L_{21}^T = xv^T.$$

V případě, že u a v nejsou vektory ale matice, tedy

$$A_{21}^T = UV^T, \quad \text{kde } U, V \in \mathbb{R}^{n \times r}$$

budou obsahovat obecně r ($r \geq 1$) sloupců, budeme řešit r soustav rovnic $L_{11}X = U$. Pokud $r \ll n$, tento postup se vyplatí více, než hledat celou inverzi L_{11}^{-1} .

V závěru práce s celými bloky bychom potřebovali vyjádřit L_{22} jako

$$L_{22} = \sqrt{A_{22} - L_{21}L_{21}^T},$$

tedy s využitím námi vytvořené funkce

$$L_{22} = \text{Chol}(A_{21} - L_{21}L_{21}^T).$$

Zcela analogicky bychom mohli pracovat s blokovou maticí s více než dvěma řádky a sloupci,

$$A = \begin{bmatrix} A_{11} & \cdots & A_{k1}^T \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{bmatrix} = A^T.$$

Dostali bychom blokovou variantu Choleského algoritmu, viz algoritmus 3.

```

for  $i = 1, 2, \dots, k$  do
   $L_{ii} = \text{Chol}(A_{ii} - \sum_{s=1}^{i-1} L_{is}L_{is}^T)$ 
  for  $j = i + 1, \dots, k$  do
     $L_{ji} = (A_{ji} - \sum_{s=1}^{i-1} L_{is}^T L_{js}) / L_{ii}^{-1}$ 
  end
end

```

Algoritmus 3: Choleského rozklad (velké matice) spočítaný po blocích. Poznamenejme, že funkce $\text{Chol}(\cdot)$ počítá Choleského rozklad (malých bloků) klasicky, po prvcích, tedy algoritmem 1.

5 Choleského rozklad matice v hierarchickém formátu

Nyní již budeme SPD matici A uvažovat v hierarchickém formátu. Proto budeme i matici L hledat v hierarchickém formátu tak, aby formáty matic A a L byly kompatibilní, tj. aby měly stejné dělení. Matice A i L jsou v první řadě blokové, přičemž hierarchické matice mohou být trojího typu, a to opět *hierarchické*, *low-rank*, a nebo to mohou být *husté* matice. Práci s hierarchickými maticemi jsme tak již svým způsobem nakousli v poslední sekci předchozí kapitoly, kde jsme hovořili o blokovém algoritmu.

5.1 Formální hierarchická struktura rozkladu

Začneme v podobném smyslu, jen si jednotlivé bloky přeznačíme (pro zpřehlednění hierarchické struktury) a sice následujícím způsobem,

$$A = A_0 = \begin{bmatrix} A_1 & B_0^T \\ B_0 & A_2 \end{bmatrix}, \quad L = L_0 = \begin{bmatrix} L_1 & 0 \\ M_0 & L_2 \end{bmatrix},$$

tedy

$$A = \begin{bmatrix} A_1 & B_0^T \\ B_0 & A_2 \end{bmatrix} = \begin{bmatrix} L_1 & 0 \\ M_0 & L_2 \end{bmatrix} \begin{bmatrix} L_1^T & M_0^T \\ 0 & L_2^T \end{bmatrix} = LL^T.$$

Bloky A_s , resp. L_s na diagonále jsou přitom *bud'* opět hierarchické matice *nebo* husté matice, a mimodiagonální bloky B_s , resp. M_s jsou matice nízké hodnoti (low-rank). Abychom objasnili značení, uděláme pár kroků v hierarchii.

Budou-li např. obě matice A_s , $s = 1, 2$, hierarchické, označíme jejich bloky takto

$$A_1 = \begin{bmatrix} A_{11} & B_1^T \\ B_1 & A_{12} \end{bmatrix}, \quad A_2 = \begin{bmatrix} A_{21} & B_2^T \\ B_2 & A_{22} \end{bmatrix}.$$

Budou-li dále všechny čtyři matice A_s , $s = 11, 12, 21, 22$, hierarchické, označíme jejich bloky takto

$$A_{11} = \begin{bmatrix} A_{111} & B_{11}^T \\ B_{11} & A_{112} \end{bmatrix}, \quad A_{12} = \begin{bmatrix} A_{121} & B_{12}^T \\ B_{12} & A_{122} \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} A_{211} & B_{21}^T \\ B_{21} & A_{212} \end{bmatrix}, \quad A_{22} = \begin{bmatrix} A_{221} & B_{22}^T \\ B_{22} & A_{222} \end{bmatrix},$$

atd. Obecně získáme

$$A_s = \begin{bmatrix} A_{s1} & B_s^T \\ B_s & A_{s2} \end{bmatrix}, \quad L_s = \begin{bmatrix} L_{s1} & 0 \\ M_s & L_{s2} \end{bmatrix}, \quad A_s = L_s L_s^T,$$

kde s je nějaká konečná posloupnost jedniček a dvojek dané délky, B_s je mimodiagonální low-rank blok matice z 2×2 blokového dělení A_s a $A_{s\ell}$, $\ell = 1, 2$, jsou bloky na diagonále tohoto dělení. Obdobně je tomu v případě matice L_s .

Z blokového algoritmu víme, že v principu budeme opět potřebovat operace odmocniny SPD matice (4.3) a operace podílu (4.4), definované v poslední sekci předchozí kapitoly pomocí funkce $\text{Chol}(\cdot)$, neboli pomocí násobení inverzí. V obou dvou případech je budeme muset být schopni provést s maticemi v příslušných formátech.

5.2 Funkce $\text{Chol}(\cdot)$ v hierarchickém formátu

Začneme rozkladem $A_s = L_s L_s^T$ pro nějaké s . Z blokového zápisu rozkladu opět plyne rovnost

$$A_{s1} = L_{s1} L_{s1}^T,$$

kde A_{s1} může být opět buď hierarchická nebo hustá matice. Záleží, jak hluboko jsme v hierarchii matic postoupili, resp. spíš, jak je blok malý (dostatečně malé matice, tj. řádu cca menšího než 8, se již nevyplatí ukládat hierarchicky; neušetří se žádná paměť; fakticky se to nevyplatí už u mnohem vyšších řádů z důvodu úspory práce, viz [14]). Funkci tedy definujeme opět takto

$$\text{Chol}(A_{s1}) = L_{s1},$$

příčemž:

1. V případě, že A_{s1} je hustá, je L_{s1} také hustá a je spočtená klasickým Choleského algoritmem (po prvcích či sloupcích), viz algoritmus 1 a 2.
2. V případě, že A_{s1} je hierarchická, budeme muset nejprve spočítat blokový Choleského rozklad A_{s1} pomocí algoritmu, který právě definujeme — postoupíme tedy v hierarchii o krok hlouběji.

Je dobré si uvědomit, že tento postup je zcela legitimní. Algoritmus se musí sám na sebe odvolávat, nebo sám do sebe vnořit tolikrát, schematicky Choleského rozklad $A = A_0$ volá (budeme značit \rightsquigarrow)

$$\text{Chol}(A_1) \rightsquigarrow \text{Chol}(A_{11}) \rightsquigarrow \dots \rightsquigarrow \text{Chol}(A_{11\dots 1}) \rightsquigarrow \text{Chol}(A_{11\dots 11}),$$

než dojdeme k nejmenší hlavní rohové podmatici $A_{11\dots 11}$, která je již uložena jako hustá. Její Choleského rozklad pak vypočítáme klasicky, jako jsme to provedli v kapitolách 4.1 a 4.2.

5.3 Mimodiagonální low-rank blok

Připomeňme, že pro mimodiagonální bloky platí

$$B_s = M_s L_{s1}^T$$

neboli

$$\begin{aligned} M_s L_{s1}^T &= B_s \\ L_{s1} M_s^T &= B_s^T. \end{aligned}$$

Nyní využijeme přesně té vlastnosti, kterou fakticky zmiňujeme již v poznámce 7. V našem případě pro matici B_s^T platí,

$$B_s^T = UV^T, \quad \text{kde} \quad U, V \in \mathbb{R}^{\star \times r},$$

kde \star je počet sloupců, resp řádků B_s a $r = \text{rank}(B_s)$ je typicky *výrazně menší* než \star . Označíme-li u_ℓ sloupce matice U ,

$$U = [u_1, \dots, u_r],$$

můžeme se pokusit najít řešení soustavy, resp. soustav

$$L_{s1} x_\ell = u_\ell, \quad \ell = 1, \dots, r.$$

Zřejmě pak

$$x_\ell = L_{s1}^{-1} u_\ell, \quad \text{neboli} \quad X = [x_1, \dots, x_r] = L_{s1}^{-1} U.$$

Celkově tedy dostáváme

$$M_s = B_s (L_{s1}^T)^{-1} = B_s (L_{s1}^{-1})^T = (L_{s1}^{-1} B_s^T)^T = (L_{s1}^{-1} UV^T)^T = (XV^T)^T = VX^T.$$

Protože B_s je ve tvaru součinu VU^T zadáno — mimodiagonální low-rank bloky u hierarchických matic máme uložené právě z důvodů úspory ve faktorizované podobě, jako uspořádanou dvojici (V, U) — to je ostatně klíčová myšlenka hierarchických matic, zbývá nám nalézt ono X . Jsme tedy prakticky hotovi, až na jednu maličkost.

5.4 Řešení soustavy s dolní trojúhelníkovou hierarchickou maticí

Naším úkolem je se naučit a ukázat jak řešit soustavu

$$L_s x_s = u_s,$$

kde L_s je dolní trojúhelníková hierarchická matice. Pomocí indexů u x a u si nyní budeme značit dělení těchto vektorů, které bude *konzistentní* s hierarchickým dělením matice

$$u_s = \begin{bmatrix} u_{s1} \\ u_{s2} \end{bmatrix}, \quad u_{s1} = \begin{bmatrix} u_{s11} \\ u_{s12} \end{bmatrix}, \quad u_{s2} = \begin{bmatrix} u_{s21} \\ u_{s22} \end{bmatrix},$$

atd., analogicky stejné dělení provedeme u x_s .

$$x_s = \begin{bmatrix} x_{s1} \\ x_{s2} \end{bmatrix}, \quad x_{s1} = \begin{bmatrix} x_{s11} \\ x_{s12} \end{bmatrix}, \quad x_{s2} = \begin{bmatrix} x_{s21} \\ x_{s22} \end{bmatrix}, \quad \dots$$

Nyní se podíváme, jak bude probíhat vlastní řešení v závislosti na struktuře L_s a vlastnostech jejich bloků. Nejprve si ale obecně, obdobným způsobem, jako jsme již vytvořili funkci $\text{Chol}(\cdot)$, stotožníme celý postup řešení úlohy $L_s x_s = u_s$ s funkcí $\text{Solve}(\cdot, \cdot)$,

$$\text{Solve}(L_s, u_s) = x_s = L_s^{-1} u_s.$$

Tato funkce pak provádí následující:

1. V případě, že L_s je hustá matice, snadno přijdeme na řešení soustavy rovnic pomocí tzv. *dopředného chodu* Gaußovy eliminace (L_s je dolní trojúhelníková). Tedy nejprve vyřešíme první rovnici o jedné neznámé a následně dosazujeme již známé kořeny shora dolů, až nalezneme všechny neznámé. Tedy obecně pro řešení rovnice $L_s x_s = u_s$ platí, že

$$\xi_1 = \frac{\mu_1}{l_{1,1}},$$

$$\xi_i = \frac{1}{l_{i,i}} \left(\mu_i - \sum_{j=1}^{i-1} l_{i,j} \xi_j \right), \quad i = 2, \dots, n_s,$$

kde ξ_i a μ_i jsou jednotlivé složky vektorů x_s resp. u_s , a kde n_s je rozměr husté matice L_s .

2. V případě, že L_s je hierarchická matice, soustava vypadá takto

$$L_s x_s = \begin{bmatrix} L_{s1} & 0 \\ M_s & L_{s2} \end{bmatrix} \begin{bmatrix} x_{s1} \\ x_{s2} \end{bmatrix} = \begin{bmatrix} u_{s1} \\ u_{s2} \end{bmatrix} = u_s.$$

Nejprve musíme řešit problém s L_{s1} , fakticky tedy hledání řešení této soustavy, tj. $\text{Solve}(L_s, u_s)$, vyžaduje nejprve nalezení řešení menší soustavy, schematicky:

$$\text{Solve}(L_s, u_s) \rightsquigarrow \text{Solve}(L_{s1}, u_{s1}) \rightsquigarrow \text{Solve}(L_{s11}, u_{s11}) \rightsquigarrow \dots$$

Tento proces vnořování opakujeme, dokud nedojdeme na úroveň, kde matice $L_{s11\dots1}$ bude již hustá. Soustava s touto maticí se pak vyřeší pomocí postupu v prvním bodě. Tím získáme $x_{s11\dots1}$.

- Abychom ale mohli „vycouvat“ ven těmito vnořenými funkcemi, po vyřešení té nejvnitřnější husté dolní trojúhelníkové soustavy, na kterou jsme aplikovali již zmíněný dopředný chod Gaußovy eliminace — bez újmy na obecnosti můžeme uvažovat, že tím nejvnitřnějším byl právě $L_{s1} x_{s1} = u_{s1}$, tedy x_{s1} již známe — musíme nalézt ještě druhou polovinu hledaného vektoru. Tedy musíme vyřešit rovnici

$$M_s x_{s1} + L_{s2} x_{s2} = u_{s2}.$$

Nicméně x_{s1} již známe a o M_s víme, že je low-rank, je tedy potřeba provést součin low-rank matice s vektorem $M_s x_{s1} = (UV^T)x_{s1} = U(V^T x_{s1})$, což lze provést snadno, jak je naznačeno uzávorkováním, a výsledek odečíst od obou stran. Dostaneme

$$L_{s2}x_{s2} = \underbrace{u_{s2} - M_s x_{s1}}_{\tilde{u}_{s2}}.$$

- Přesouváme se nyní na vyřešení soustavy s druhým blokem z diagonály matice L_s . Zde mohou nastat opět dvě možnosti. Buď L_{s2} je hustá (tedy opět řešíme přímo), nebo je v hierarchickém tvaru a musíme se ponořit do hierarchické struktury. Obecně řečeno zavoláme funkci $\text{Solve}(\cdot, \cdot)$, schematicky:

$$\text{Solve}(L_{s2}, \tilde{u}_{s2}) \rightsquigarrow \text{Solve}(L_{s21}, \tilde{u}_{s21}) \rightsquigarrow \text{Solve}(L_{s211}, \tilde{u}_{s211}) \rightsquigarrow \dots$$

Tímto způsobem nalezneme řešení celé soustavy $L_s x_s = u_s$ s dolní trojúhelníkovou hierarchickou maticí L_s .

Zároveň je ovšem také hotový celý Choleského rozklad

$$A = LL^T$$

symetrické pozitivně definitní matice A v hierarchickém formátu. Dolní trojúhelníkový faktor L jsme dostali opět v hierarchickém formátu.

6 Řešení soustavy se symetrickou pozitivně definitní maticí v hierarchickém formátu

V této kapitole se budeme věnovat hlavní otázce, která stojí v názvu této diplomové práce, a to jak řešit soustavy rovnic s hierarchickými maticemi pomocí rozkladů. Postupně připomeneme všechny předpoklady, které jsme již v předešlých kapitolách popisovali, a které jsou důležité pro odpověď na zmíněnou otázku.

6.1 Převedení úlohy na řešení dvou trojúhelníkových soustav

Prvním krokem je nalezení vhodného rozkladu. Protože pracujeme s SPD maticí A , bude vhodným rozkladem právě Choleského rozklad (viz sekce 3.4). Protože A bude navíc v hierarchickém formátu, budeme uvažovat Choleského rozklad v hierarchickém formátu (viz kapitola 5). Nejprve však budeme uvažovat obecný Choleského rozklad, pro který platí, že

$$A = LL^T.$$

Připomeňme, že SPD matice je vždy regulární. Chceme-li tedy řešit soustavu

$$Ax = b,$$

matici A nejprve nahradíme jejím rozkladem a dostaneme

$$LL^T x = b.$$

Tuto rovnici si můžeme vhodně uzávorkovat a poté úlohu rozdělit na dva samostatné kroky – dvě samostatné soustavy. Konkrétně

$$L \underbrace{(L^T x)}_y = b.$$

Označením uzávorkovaného součinu symbolem y (součin matice s vektorem je vektor), dostaneme *soustavu s dolní trojúhelníkovou maticí L* ,

$$Ly = b.$$

Její řešení je vektor y . Vrátime-li se k označení, které jsme provedli, tedy

$$L^T x = y,$$

vidíme, že zbývá vyřešit soustavu s horní trojúhelníkovou maticí L^T a pravou stranou y . Tím dostaneme vektor x , tedy řešení původní soustavy. Schematicky

$$Ax = b \quad \iff \quad Ly = b \quad \text{a zároveň} \quad L^T x = y.$$

Řešení jedné soustavy s relativně obecnou maticí je nahrazeno dvěma řešeními dvou soustav ovšem s trojúhelníkovými maticemi. Kdybychom pracovali s klasickými maticemi, tedy uloženými po prvcích, vyřešili bychom nejprve dolní trojúhelníkovou soustavu *dopředným chodem Gaußovy eliminace* – vyřešíme první rovnici o jedné neznámé, dosadíme do druhé, resp. všech ostatních rovnic, atd. – a pak horní trojúhelníkovou soustavu *zpětným chodem Gaußovy eliminace* – vyřešíme poslední rovnici o jedné neznámé, dosadíme do předposlední, resp. všech ostatních rovnic, atd.

6.2 Řešení soustavy s trojúhelníkovou hierarchickou maticí

Dalším krokem bude nejprve řešení soustavy $Ly = b$ s dolní trojúhelníkovou hierarchickou maticí. Tento krok jsme již vysvětlili v kapitole 5.4. Formálně použijeme zkonstruovanou funkci $\text{Solve}(\cdot, \cdot)$, s jejíž pomocí se dokážeme dostat hierarchií až na úroveň nejmenších, již hustých (dolních trojúhelníkových) podmatic, ty řešíme pomocí dopředného chodu Gaußovy eliminace a dopočítáváme jednotlivé části vektoru y .

Získáme tedy y a můžeme se posunout k řešení druhé soustavy $L^T x = y$. Řešení soustavy $L^T x = y$ s horní trojúhelníkovou hierarchickou maticí L^T je však zcela analogické řešení soustavy s dolní trojúhelníkovou maticí. Jen bychom adaptovali funkci $\text{Solve}(\cdot, \cdot)$ i na řešení horních trojúhelníkových soustav takto:

Místo u levého horního bloku L_{s1} matice L_s bychom začínali u pravého dolního bloku L_{s2}^T matice L_s^T

$$L_s = \begin{bmatrix} L_{s1} & 0 \\ M_s & L_{s2} \end{bmatrix}, \quad L_s^T = \begin{bmatrix} L_{s1}^T & M_s^T \\ 0 & L_{s2}^T \end{bmatrix}.$$

Podobně při řešení nejvnitřnějších hustých matic bychom postupovali zdola nahoru, neboli pomocí zpětného chodu Gaußovy eliminace.

Poznámka 8. *Stejně jako při dopředném chodu, i zde si uvedeme jednoduchý algoritmus zpětného chodu eliminace soustavy $L_{s2}^T x_s = y_s$,*

$$\xi_n = \frac{\nu_n}{l_{n,n}},$$

$$\xi_i = \frac{1}{l_{i,i}} \left(\nu_i - \sum_{j=i+1}^n l_{i,j} \xi_j \right), \quad i = n-1, \dots, 1$$

kde ξ_i a ν_i jsou jednotlivé složky vektorů x_s a y_s .

7 Explicitní inverze řídké symetrické pozitivně definitní matice v hierarchickém formátu

Jakmile jsme schopni vytvořit Choleského rozklad matice A , která je v hierarchickém tvaru (a tedy i umíme počítat soustavy rovnic s touto maticí a jejími faktory), můžeme tento rozklad využít například i pro výpočet *explicitní* inverze takovéto matice. Poznamenejme, že klasicky explicitní inverzi velké řídké matice řádu n spočítat nelze. Nelze to zpravidla z paměťových důvodů. Např. inverze třídiagonální matice, tedy matice na jejíž uložení potřebujeme uložit cca $3n$ čísel – vzorového příkladu řídké matice – je obecně hustá (jak jsme ukázali již v [14]) a na její uložení potřebujeme uložit n^2 . Stačí pak zvolit takové n , aby se $3n$ do paměti počítače ještě vešlo a n^2 již ne.

7.1 Inverze rozložené matice

Protože pracujeme se symetrickou pozitivně definitní maticí, uvažujme její Choleského rozklad

$$A = LL^T.$$

Pro inverzi matice (připomeňme, že SPD matice je vždy regulární) A^{-1} , použijeme-li známou poučku, že inverze součinu matice je součin inverzí v opačném pořadí, platí

$$A^{-1} = (LL^T)^{-1} = (L^T)^{-1}L^{-1}.$$

Potřebujeme tedy najít inverze jednotlivých součinitelů – faktorů. Respektive jen jednoho faktoru L , protože druhý je de-facto stejný, jen transponovaný.

7.2 Inverze dolní trojúhelníkové hierarchické matice

Označme si X hledanou inverzi, tedy $X = L^{-1}$. Uvažujme dělení matice L na 2×2 bloky a konzistentní dělení matice X . Pokud zřejmě platí

$$LL^{-1} = LX = \underbrace{\begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}}_{\text{Hi}} \underbrace{\begin{bmatrix} X_{11} & 0 \\ X_{21} & X_{22} \end{bmatrix}}_{\text{hledaná}} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = I.$$

Pak tedy, rozepsáním rovnosti po blocích, musí platit

$$\begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} X_{11} & 0 \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} L_{11}X_{11} & 0 \\ L_{21}X_{11} + L_{22}X_{21} & L_{22}X_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},$$

neboli

$$L_{11}X_{11} = I, \quad L_{22}X_{22} = I, \quad L_{21}X_{11} + L_{22}X_{21} = 0.$$

Matice L je přitom nejen bloková, ale i v hierarchickém formátu. Pro jednotlivé rovnosti dostáváme:

1. $L_{11}X_{11} = I$:

Obě matice L_{11} i X_{11} jsou v dolním trojúhelníkovém tvaru a jsou navzájem inverzní. Jestliže jsou tyto matice v hierarchickém formátu, musíme provést opět rekurentní krok – rozdělením na jednotlivé bloky –, dokud se nedostaneme na obecně husté matice.

Jestliže jsou tyto matice již husté, můžeme spočítat X_{11} např. pomocí Gaußovy eliminace. Tedy klasicky odeliminovat jednotlivé řádky

$$[L_{11}|I] \sim \dots \sim [I|L_{11}^{-1}] = [I|X_{11}].$$

2. $L_{22}X_{22} = I$:

Zde bude situace zcela analogická. Obě matice L_{22} i X_{22} jsou v dolním trojúhelníkovém tvaru a jsou navzájem inverzní. Buď jsou opět hierarchické, pak je rozdělíme na jednotlivé bloky, nebo jsou husté, pak např. pomocí Gaußovy eliminace dostaneme

$$[L_{22}|I] \sim \dots \sim [I|L_{22}^{-1}] = [I|X_{22}].$$

Výpočty bloků X_{11} a X_{22} lze zjevně provádět zcela nezávisle – paralelně.

3. $L_{21}X_{11} + L_{22}X_{21} = 0$:

Zde určitě známe matice L_{21} a L_{22} , pokud jsme již provedli první krok celý, známe i matici X_{11} . Hledáme tedy matici X_{21} . Rovnici si tak můžeme upravit do tvaru

$$L_{22}X_{21} = -L_{21}X_{11}.$$

Pravou stranu, kterou tvoří matice L_{21} , která má nízkou hodnot (low-rank, LR) a matice X_{11} , která je buď hustá (Ds) nebo v hierarchickém formátu (Hi). Připomeňme, že aritmetice, resp. násobení matic různých typů se stručně věnuje kapitola 2.2, detailněji viz bakalářská práce [14]. Každopádně součinem LR a Ds, resp. LR a Hi matic je vždy LR matice. Tedy matice $B = -L_{21}X_{11}$ je nízké hodnoti, můžeme ji tedy přepsat, jako obvykle, na součin $B = UV^T$, kde U a V mají $r = \text{rank}(B)$ sloupců.

- Zbývá tedy vyřešit rovnici

$$L_{22}X_{21} = B, \quad \text{resp.} \quad L_{22}X_{21} = UV^T.$$

Protože L_{22} je regulární, platí

$$X_{21} = L_{22}^{-1}B = L_{22}^{-1}(UV^T) = \underbrace{(L_{22}^{-1}U)}_{\tilde{U}} V^T.$$

Vidíme, jednak že X_{21} je také LR, neboť $X_{21} = \tilde{U}V^T$ a \tilde{U} a V mají r sloupců. Druhou věcí je, že umíme X_{21} nalézt. Vyjdeme z rovnosti

$$\begin{aligned} \tilde{U} &= L_{22}^{-1}U, \\ L_{22}\tilde{U} &= U. \end{aligned}$$

Když si označíme sloupce matic \tilde{U} a U takto

$$\tilde{U} = [\tilde{u}_1, \dots, \tilde{u}_r], \quad U = [u_1, \dots, u_r],$$

vidíme, že stačí vyřešit r soustav

$$L_{22}\tilde{u}_j = u_j, \quad j = 1, \dots, r$$

s dolní trojúhelníkovou maticí L_{22} , která je buď hustá, což umíme dopředným chodem Gaußovy eliminace, nebo je v hierarchickém formátu, což také umíme viz sekce 5.4.

7.3 Explicitní inverze celé matice

Explicitní inverzi velké řídké SPD matice A pak dostaneme již jednoduše

$$A^{-1} = (L^T)^{-1}L^{-1} = \begin{bmatrix} X_{11} & 0 \\ X_{21} & X_{22} \end{bmatrix}^T \begin{bmatrix} X_{11} & 0 \\ X_{21} & X_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} X_{11}^T & X_{21}^T \\ 0 & X_{22}^T \end{bmatrix}}_{\text{Hi}} \underbrace{\begin{bmatrix} X_{11} & 0 \\ X_{21} & X_{22} \end{bmatrix}}_{\text{Hi}},$$

což již umíme provést. Součinu dvou hierarchických matic (konzistentně dělených na bloky) se stručně věnuje kapitola 2.2, detailněji viz bakalářská práce [14]. Výsledkem součinu je opět matice v hierarchickém tvaru. Poznamenejme, že příklady takovýchto řídkých matic lze nalézt v Harwell–Boeing Collection (viz [22] nebo [23]).

Závěr

V této diplomové práci jsme nejdříve zopakovali případným čtenářům důležité informace o vlastnostech hierarchických matic. Popsali jsme základní aritmetické operace, které se při práci s těmito maticemi mohou objevit. Pojem hierarchická matice není natolik konkrétní, a tak jsme vhodně zvolili pro hlavní myšlenku této práce symetrickou pozitivně definitní matici. O SPD matici víme, že má řadu zajímavých vlastností, konkrétně v našem případě jsme využili zejména její silnou regularitu. Pomocí Gaußovy eliminace jsme našli LU rozklad SPD matice, který je možné zapsat jako $A = LL^T$, což je právě tzv. Choleského rozklad matice A . Hlavní část této práce byla zaměřená na výpočet Choleského rozkladu takovéto SPD matice. Poukázali jsme na několik možností, jak k tomuto rozkladu lze dojít, např. náročným výpočtem po prvcích. V tomto případě nám postupně vyjádřené prvky dále pomáhaly k získání prvků následujících. Druhý způsob, který jsme popsali je výpočet po sloupcích, kdy matici A lze nakonec zapsat jako součet vnějších součinů postupně nalezených sloupců L . Matici A jsme dále rozdělili na bloky a snažili se ukázat, zdali by bylo také možné tento rozklad získat. Jelikož jsme se setkali s několika nesrovnalostmi, například v podobě odmocniny z matice, kterou nelze interpretovat, museli jsme zadefinovat maticovou funkci $\text{Chol}(\cdot)$, jejímž úkolem bylo nalézt Choleského rozklad daného bloku.

Pokud je matice A v hierarchickém tvaru, poté musí být i její rozklad hierarchického formátu. V práci jsme ukázali, jak tohoto tvaru dosáhnout a následně využít k řešení soustav s hierarchickou maticí. Řešení takovéto soustavy $Ax = b$ jsme rozdělili na $Ly = b$, kde $L^T x = y$. Tímto způsobem lze využít Choleského rozklad pro výpočet soustav lineárních rovnic s hierarchickými maticemi.

Smysl využití Choleského rozkladu je tedy velmi zřejmý, jelikož vzhledem k tomu, že matice L a L^T jsou trojúhelníkového typu, je řešení jednotlivých rovnic jednoduché. Tato řešení nalezneme s pomocí dopředných a zpětných chodů Gaußovy eliminace. Choleského rozklad hierarchické matice A nám v neposlední řadě může pomoci k nalezení její explicitní inverze. To může být extrémně užitečné, zejména proto, že obecně, resp. klasicky pro velkou řídkou matici A její inverzi explicitně sestavit nelze, ne z matematických, ale z ryze praktických, fyzických důvodů.

A Apendix: Součiny a inverze trojúhelníkových matic

V této příloze jen vyslovíme a dokážeme několik užitečných pomocných vět o trojúhelníkových maticích.

A.1 Součin dvou dolních trojúhelníkových matic je dolní trojúhelníkový

Věta 2. *Nechť A a $B \in \mathbb{R}^{n \times n}$ jsou dolní trojúhelníkové matice stejného řádu. Pak jejich součin*

$$AB \in \mathbb{R}^{n \times n}$$

je opět dolní trojúhelníkový.

Důkaz. Označme si $a_{i,j}$ a $b_{i,j}$ prvky matice A , resp. B . Protože matice A a B jsou dolní trojúhelníkové, pro $i < j$ platí $a_{i,j} = b_{i,j} = 0$. Součin dvou matic, resp. (i, j) -tý prvek součinu je definovaný vztahem

$$(AB)_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}.$$

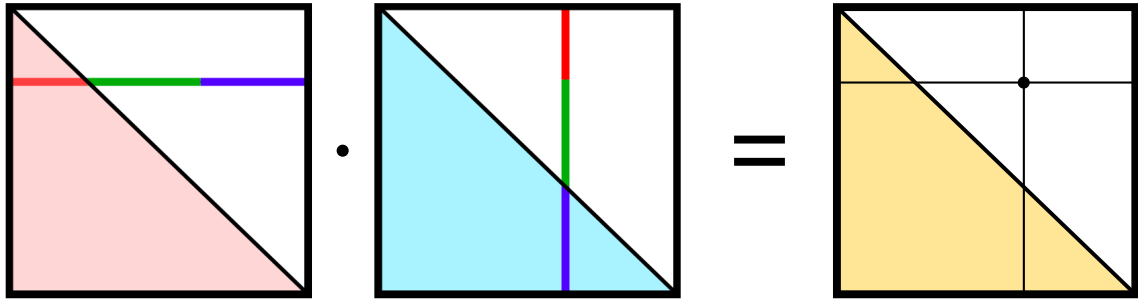
Pokud je $i < j$, můžeme sumu na pravé straně rozdělit podle indexů i a j obecně na tři části

$$\sum_{k=1}^n a_{i,k} b_{k,j} = \sum_{k=1}^i a_{i,k} b_{k,j} + \sum_{k=i+1}^{j-1} a_{i,k} b_{k,j} + \sum_{k=j}^n a_{i,k} b_{k,j},$$

přičemž v případě $i < i+1 = j$ prostřední suma neexistuje (obsahuje nula sčítanců) a pro $i < i+2 = j$, tedy $i+1 = j-1$ obsahuje jen jediný sčítanec. Viz také obrázek A.1.

Nyní se podívejme na hodnoty prvků v jednotlivých sumách. Víme, že když je řádkový index ostře menší než sloupcový, příslušný prvek je nula, tedy

$$\sum_{k=1}^i a_{i,k} \underbrace{b_{k,j}}_0 + \sum_{k=i+1}^{j-1} \underbrace{a_{i,k}}_0 \underbrace{b_{k,j}}_0 + \sum_{k=j}^n a_{i,k} \underbrace{b_{k,j}}_0.$$



Obrázek A.1: Výpočet (i, j) -tého prvku součinu dvou dolních trojúhelníkových matic, kde $i < j$; i -tý řádek první matice násobíme j -tým sloupcem druhé matice. Obecně se součin rozpadne na tři části naznačené barevně.

Tedy pro $i < j$ dostáváme

$$(AB)_{i,j} = \sum_{k=1}^i a_{i,k} \cdot 0 + \sum_{k=i+1}^{j-1} 0 \cdot 0 + \sum_{k=j}^n 0 \cdot b_{k,j} = 0.$$

Neboli součin AB je dolní trojúhelníková matice. □

A.2 Součin dvou horních trojúhelníkových matic je horní trojúhelníkový

Věta 3. *Nechť A a $B \in \mathbb{R}^{n \times n}$ jsou horní trojúhelníkové matice stejného řádu. Pak jejich součin*

$$AB \in \mathbb{R}^{n \times n}$$

je opět horní trojúhelníkový.

Důkaz. Protože A a B jsou horní trojúhelníkové matice, A^T a B^T jsou dolní trojúhelníkové matice. Jejich součin $B^T A^T$ je tedy podle věty 2 dolní trojúhelníková. Protože součin transpozic matic je transpozice součinu matic ale v opačném pořadí, dostáváme

$$B^T A^T = (AB)^T,$$

tedy $(AB)^T$ je dolní trojúhelníková matice a tedy AB horní trojúhelníková. □

A.3 Inverze dolní trojúhelníkové matice je dolní trojúhelníková

Věta 4. *Nechť $A \in \mathbb{R}^{n \times n}$ je dolní trojúhelníková regulární matice (tedy s nenulovými prvky na diagonále). Pak její inverze*

$$A^{-1} \in \mathbb{R}^{n \times n}$$

je opět dolní trojúhelníková matice.

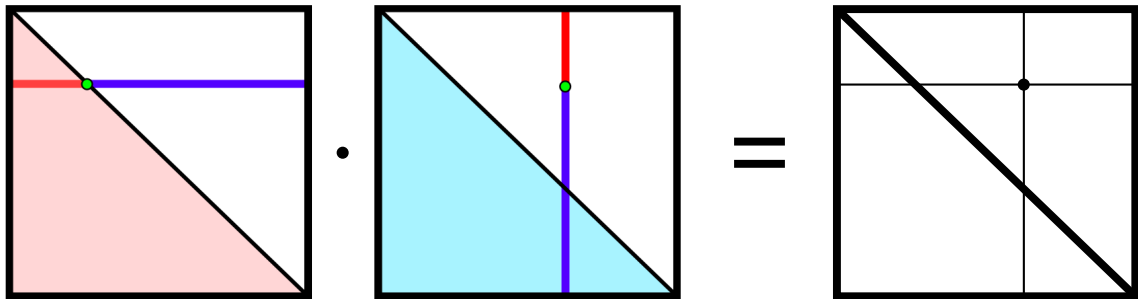
Důkaz. Označme si $a_{i,j}$ prvky matice A . Protože matice A je dolní trojúhelníková, pro $i < j$ platí $a_{i,j} = 0$. Protože matice A je regulární označme si dále $X = A^{-1}$ její inverzi a $x_{i,j}$ její prvky. Důkaz provedeme sporem.

Předpokládejme *naopak*, že X není v dolním trojúhelníkovém tvaru. Pak existují takové indexy i a j , $i < j$, že platí $x_{i,j} \neq 0$. Tedy matice X má nad diagonálou alespoň jeden nenulový prvek. Navíc je určitě možné zvolit indexy i a j takovým způsobem, že $x_{k,j} = 0$ pro všechny $k < i$. Jinými slovy, indexy volíme tak, že nenulový prvek $x_{i,j}$ je v j -tém sloupci prvním nenulovým prvkem, všude nad ním jsou již nuly.

Protože $AX = AA^{-1} = I$ je jednotková matice, platí $(AX)_{i,j} = 0$. Zároveň ale platí

$$(AX)_{i,j} = \sum_{k=1}^n a_{i,k}x_{k,j} = \sum_{k=1}^{i-1} a_{i,k}b_{k,j} + a_{i,i}b_{i,j} + \sum_{k=i+1}^n a_{i,k}b_{k,j},$$

kde jsme sumu napravo opět vhodným způsobem rozdělili. Pokud $i = 1$, pak první suma neexistuje. Viz také obrázek A.2.



Obrázek A.2: Součin dolní trojúhelníkové matice (vlevo) a její inverze (vpravo). V inverzi na pozici (i, j) , $i < j$, předpokládáme nenulový prvek (zelený bod). Následně spočítáme součin i -tého řádku matice s j -tým řádkem inverze. Obecně se součin rozpadne na tři části naznačené barevně. Nenulový prvek nad diagonálou inverze způsobí nenulový prvek nad diagonálou jednotkové matice.

Nyní zbývá rozmyslet, které z prvků jsou nulové z důvodů, že (i) A je dolní trojúhelníková resp. (ii) nad prvkem $x_{i,j}$ jsou již jen samé nuly. Zřejmě

$$\sum_{k=1}^{i-1} a_{i,k} \underbrace{x_{k,j}}_{0 \text{ (ii)}} + a_{i,i}x_{i,j} + \sum_{k=i+1}^n \underbrace{a_{i,k}}_{0 \text{ (i)}} x_{k,j}.$$

Dostáváme tak

$$0 = (AX)_{i,j} = \sum_{k=1}^{i-1} x_{i,k} \cdot 0 + a_{i,i}x_{i,j} + \sum_{k=i+1}^n 0 \cdot x_{k,j} = a_{i,i}x_{i,j}.$$

Nicméně o prvku $a_{i,i}$ víme, že je nenulový z regularity matice A , a o prvku $x_{i,j}$ víme, že je dle předpokladu nenulový (dle něj jsme volili indexy i a j), tedy také $a_{i,i}x_{i,j} \neq 0$. Dostáváme tak spor, tedy matice X musí být opět v dolním trojúhelníkovém tvaru. \square

A.4 Inverze horní trojúhelníkové matice je horní trojúhelníková

Věta 5. *Nechť $A \in \mathbb{R}^{n \times n}$ je dolní trojúhelníková regulární matice (tedy s nenulovými prvky na diagonále). Pak její inverze*

$$A^{-1} \in \mathbb{R}^{n \times n}$$

je opět dolní trojúhelníková matice.

Důkaz. Protože A je horní trojúhelníková matice, A^T je dolní trojúhelníková matice. Její inverze $(A^T)^{-1}$ je tedy podle věty 4 dolní trojúhelníková. Protože inverze a transpozice jsou navzájem zaměnitelné, dostáváme

$$(A^T)^{-1} = (A^{-1})^T,$$

tedy $(A^{-1})^T$ je dolní trojúhelníková matice, a tedy A^{-1} horní trojúhelníková. \square

Literatura

- [1] H. Anton, Ch. Rorres: *Elementary linear algebra: applications version* (Eleventh Edition), Hoboken, NJ, 2014.
- [2] M. Bebendorf: *Hierarchical Matrices*, Edice Lecture notes in computational science and engineering (LCNSE) 63, Springer-Verlag, Berlin, Heidelberg, 2008.
doi: 10.1007/978-3-540-77147-0
<http://www.springer.com/la/book/9783540771463>
- [3] S. Börm: *Efficient Numerical Methods for Non-local Operators: \mathcal{H}^2 -Matrix Compression, Algorithms and Analysis*, Edice EMS Tracts in Mathematics 14, European Mathematical Society, Zürich, 2010.
http://www.ems-ph.org/books/book.php?proj_nr=125
- [4] R. Calinger, J. E. Brown, T. R. West. *A contextual history of mathematics* NJ, Prentice Hall, 1999.
- [5] J. Demel: *Grafy a jejich aplikace*, Academia, Praha 2002.
- [6] J. Demel: *Grafy*, SNTL, Státní nakladatelství technické literatury (edice MVŠT, Matematika pro vysoké školy technické, sešit XXXIV), Praha, 1988, resp. 1989 (dotisk).
- [7] E. J. Duintjer Tebbens, I. Hnětynková, M. Plešinger, Z. Strakoš, P. Tichý: *Analýza metod pro maticové výpočty. Základní metody* Matfyzpress, Praha, 2012.
- [8] M. Fiedler: *Speciální matice a jejich použití v numerické matematice* Praha, Státní nakladatelství technické literatury, 1981.
- [9] G. H. Golub, C. F. Van Loan: *Matrix Computations* (Fourth Edition), Johns Hopkins University Press, Baltimore, MD, 2012.
<https://jhupbooks.press.jhu.edu/content/matrix-computations-0>
- [10] W. Hackbusch: *Hierarchische Matrizen: Algorithmen und Analysis*, Springer-Verlag, Berlin, Heidelberg, 2009.
doi: 10.1007/978-3-642-00222-9
<http://www.springer.com/la/book/9783642002212>
- [11] N. Higham: *Accuracy and stability of numerical algorithms* Society for Industrial and Applied Mathematics, Philadelphia, 2002.

- [12] S. Chandrasekaran, P. Dewilde, M. Gu, W. Lyons, T. Pals: *A fast solver for HSS representations via sparse matrices*, SIAM Journal on Matrix Analysis and Applications, Volume 29, Number 1 (2006), pp. 67-81. (15 pages)
doi: [10.1137/050639028](https://doi.org/10.1137/050639028)
<http://epubs.siam.org/doi/abs/10.1137/050639028>
- [13] G. Chartrand: *Introductory graph theory*, Dover Publications, New York, 1985.
- [14] B. Košková: *Hierarchické matice: Moderní přístup k práci s velkými hustými maticemi*, bakalářská práce TU v Liberci, 2018.
- [15] D. Kressner, M. Plešinger, C. Tobler: *A preconditioned low-rank CG method for parameter-dependent Lyapunov matrix equations*, Numerical Linear Algebra with Applications 21(5) (2014), str. 666–684.
- [16] J. Matoušek, J. Nešetřil: *Kapitoly z diskrétní matematiky*, Karolinum, Praha, 2009.
- [17] J. Nešetřil: *Teorie grafů*, SNTL, Státní nakladatelství technické literatury (edice MS, Matematický seminář, č. 13), Praha, 1979.
- [18] S. Pauli: *A numerical solver for Lyapunov equations based on the matrix sign function iteration in HSS arithmetic*, Semester Thesis, SAM, ETH Zürich, 2010.
- [19] J. Sedláček: *Kombinatorika v teorii a praxi. Úvod do teorie grafů*, Academia (edice CV, Cesta k vědění, č. 7), Praha, 1964.
- [20] J. Sedláček: *Úvod do teorie graů*, Academia (edice CV, Cesta k vědění, č. 25, resp. 29), Praha, 1977 (2. vydání), resp. 1981 (3. vydání).
- [21] J. Žáková: *Tenzory a kanonické tenzorové rozklady: Tuckerův rozklad*, bakalářská práce TU v Liberci, 2015.
- [22] HARWELL–BOEING SPARSE MATRIX COLLECTION,
ftp://ftp.numerical.rl.ac.uk/pub/harwell_boeing.
Sada standardních řídkých matic pro testování algoritmů numerické lineární algebry (původně sestavili Iain Duff, Roger Grimes, John Lewis). Tato sada matic je též k součásti [23] a [24].
- [23] MATRIX MARKET,
<http://math.nist.gov/MatrixMarket>.
Databáze obsahuje téměř 500 řídkých matic a mnoho nástrojů pro generování matic, jež pocházejí z různých reálných aplikací. Matrix Market je projekt Oddělení matematiky a informatiky (Mathematical and Computational Sciences Division) Laboratoří informatiky (Information Technology Laboratory, ITL) amerického Národního ústavu pro standardy a technologie (National Institute of Standards and Technology, NIST) při ministerstvu obchodu USA.

- [24] THE UNIVERSITY OF FLORIDA SPARSE MATRIX COLLECTION,
<https://sparse.tamu.edu>.
Sada testovacích matic (sestavili Tim Davis, Yifan Hu (AT&T Research)).
- [25] HIERARCHICAL MATRICES. HLIB PACKAGE,
<http://www.hlib.org>.
Software pro práci s hierarchickými maticemi.