

## Část I

# Lineární algebra

## Organizační úvod

*Poznámka* (Organizační úvod)

Přednáška rozdělena na 2 části. Materiály na stránkách / studentském úložišti.

## Úvod

*Poznámka* (O čem vlastně Numerická matematika je)

Nejprve musíme problém diskretizovat (většinou bývá nekonečně dimenzionální/spojitý a my chceme konečnou dimenzi a diskrétní problém). Inverze k diskretizaci je relevance.

## 1 Chování algoritmů v počítačové aritmetice

### Definice 1.1 (Double precision)

Obsahuje 64 bitů – znaménko (1b) + exponent (11b) + mantisa (52b). Reprezentuje číslo v normalizovaném tvaru  $(\pm 1, \text{mantisa} \cdot 2^{e=\text{exponent}})$ , je-li  $-1022 = e_{\min} \leq e \leq e_{\max} = 1023$ .

Speciální hodnoty jsou  $e = e_{\min} - 1$ , kde pokud je mantisa 0, tak reprezentujeme 0, jinak máme  $\pm 0, \text{mantisa} \cdot 2^e$ . Pro  $e = e_{\max} + 1$  je při mantise nula hodnota  $\pm\infty$ , jinak tzv. NaN (not-a-number).

Toto „těleso“ označujeme  $\mathbb{F}$ .

┌

*Poznámka*

└

V intervalu mezi dvěma exponenty jsou čísla rozloženy rovnoměrně.

┌

└

*Poznámka*

Samozřejmě existují i jiné přesnosti, standardně binary32. binary128 je málo podporované, binary16 je spíše pro grafické karty.

### Definice 1.2 (Operace s $\mathbb{F}$ )

První, co potřebujeme je schopnost zaokrouhlit. V standardu je zaokrouhlení na nejbližší číslo (tj.  $\text{round}(x) = x(1 + \delta)$ , kde  $|\delta| \leq u$  je minimální a záleží na reprezentaci).

Standardní model aritmetiky s konečnou přesností předpokládá, že výsledek každé operace je rozmazán nějakým  $(1 + \delta)$ , kde  $|\delta| \leq u$ .

### **Definice 1.3** (Škálování, kancelace (vyrušení, ztráta informace))

Může se nám stát, že přičítáme tak malá čísla, že se výsledek neliší od původní hodnoty. Tomu se předchází škálováním dat.

Také se může stát, že odčítáme dvě téměř stejná čísla ( $|a - b| \ll |a| + |b|$ ), tak se odečtou úplně, protože drobná změna se ztratila v jejich reprezentaci. Takové výrazy je třeba vyjádřit jinak, aby se předešlo odčítání blízkých čísel.

## 1.1 Analýza chyb

### *Poznámka*

Na chyby při výpočtech se můžeme dívat 2 způsoby: jako na chybu výsledku, nebo jako na chybu zadání.

### **Lemma 1.1**

Nechť  $|\alpha_i| \leq u$ ,  $i \in [n]$  a nechť  $n \cdot u < 0.01$ . Potom platí

$$\prod_{i=1}^n (1 + \alpha_i) = 1 + \beta_n,$$

kde  $|\beta_n| \leq 1.01 \cdot n \cdot u$ .

┌

### *Důkaz*

Vyjádříme si  $\beta_n$ , omezíme  $|\beta_n| \leq (1 + u)^n - 1$ . Jelikož  $1 + x \leq e^x$  pro  $x \geq 0$ , pak tuto hodnotu dále omezíme  $e^{n \cdot u} - 1 < n \cdot u(1 + \frac{n \cdot u}{2} + (\frac{n \cdot u}{2})^2 + \dots) = \frac{n \cdot u}{1 - n \cdot u/2} < 1.01 \cdot n \cdot u$ .  $\square$

└

### **Definice 1.4** (Přímá analýza chyb)

Popis šíření zaokrouhlovacích chyb v algoritmu, odhad přímé chyby je

$$||a(x) - fl(a(x))||,$$

kde  $fl$  je strojové zaokrouhlování. Nezávislý (tzv. efektivní) odhad možný zřídka.

### **Definice 1.5** (Zpětná analýza chyb)

Snaha interpretovat zaokrouhlovací chyby pomocí změn vstupních dat, tj. odhad zpětné chyby, která je

$$||x - \hat{x}||,$$

kde  $\hat{x}$  je vstupní dato, se kterým by se dospělo stejného výsledku, ale přesnou cestou.

## 1.2 Stabilita algoritmu

### Definice 1.6 ( $O(u)$ )

$O(u)$  je neurčité číslo,

$$|O(u)| \leq K \cdot u,$$

kde  $K$  nezávisí na datech (jedině na dimenzi, na které může záviset libovolně).

### Definice 1.7 (Zpětně stabilní algoritmus)

Algoritmus  $a(x)$  je zpětně stabilní, pokud

$$\forall x \exists \hat{x} : fl(a(x)) = a(\hat{x}) \wedge \frac{\|x - \hat{x}\|}{\|x\|} = O(u).$$

(Tj. jestliže se chyby výpočtu způsobené zaokrouhlováním promítnou do dostatečně malých změn vstupních dat.)

### Definice 1.8 (Numerická stabilita algoritmu)

Algoritmus  $a(x)$  je numericky stabilní, pokud

$$\forall x \exists \hat{x} : \frac{\|x - \hat{x}\|}{\|x\|} = O(u) \wedge \frac{\|a(\hat{x} - fl(a(x)))\|}{a(\hat{x})} = O(u).$$

## 1.3 Podmíněnost problému

### Poznámka

Mějme matematický problém  $f : X \rightarrow Y$ , kde  $X, Y$  jsou normované lineární prostory,  $f$  je spojitě zobrazení, mající vlastnosti, že zkoumáme existenci, jednoznačnost řešení a že problém je citlivý na změny dat. Co to ale je?

### Definice 1.9 (Číslo podmíněnosti)

Označme  $\Delta x$  malou změnu dat (perturbace),  $\Delta f(x) = f(x + \Delta x) - f(x)$  změnu řešení.

Číslo podmíněnosti problému  $f$  v  $x$  ke

$$\kappa_f(x) \equiv \lim_{\delta \rightarrow 0} \sup_{\|\Delta x\| \leq \delta} \left( \frac{\|\Delta f(x)\|}{\|f(x)\|} / \frac{\|\Delta x\|}{\|x\|} \right).$$

Špatně či dobře podmíněný problém je, že malé změny  $x$  vedou na velké či malé změny

v  $f(x)$ .

### Definice 1.10 (Generovaná maticová norma)

Mějme vektorovou normu  $\|\cdot\|$  na  $\mathbb{C}^n$  a  $\mathbb{C}^m$ . Generovanou maticovou normou nazveme funkcionál  $g : \mathbb{C}^{n \times m} \rightarrow \mathbb{R}$ ,

$$g(A) \equiv \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|.$$

$g(A)$  je norma, značíme ji  $\|A\|$ . Je navíc „skoro multiplikativní“:  $\|AB\| \leq \|A\| \cdot \|B\|$ . A z definice  $\|Ax\| \leq \|A\| \cdot \|x\|$ . Je-li  $A$  čtvercová, pak  $\|I\| = 1$ .

### Definice 1.11 (Frobeniova norma)

$$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2}.$$

*Poznámka* (Podmíněnost  $Ax$ )

Pokud je problém  $f : x \mapsto Ax$ ,  $\kappa_f(x) = \frac{\|x\|}{\|Ax\|} \|A\|$ .

Pokud  $A$  je regulární, pak z  $\|x\| = \|A^{-1}Ax\| \leq \|A^{-1}\| \cdot \|Ax\|$  plyne  $\kappa_f(x) \leq \|A\| \cdot \|A^{-1}\|$ .

### Definice 1.12

Číslo z předchozí poznámky nazveme číslo podmíněnosti matice  $A$  a značíme ho  $\kappa(A) \equiv \|A\| \cdot \|A^{-1}\|$ .

*Poznámka*

Zřejmě  $1 = \|A^{-1}A\| \leq \kappa(A)$ .

### Definice 1.13 (Přesnost výpočtu)

Nechť  $a(x)$  je algoritmus. Pak (relativní) přesnost výpočtu je

$$\frac{\|a(x) - fl(a(x))\|}{\|a(x)\|}.$$

*Poznámka*

Je-li algoritmus  $a$  řešící problém  $f$  v  $\mathbb{F}$  zpětně stabilní, potom  $fl(a(x)) = a(\hat{x})$  pro nějaké  $\hat{x}$  splňující

$$\frac{\|x - \hat{x}\|}{\|x\|} = O(u) \equiv \delta.$$

Algoritmus řeší daný problém znamená  $f(x) = a(x)$ , proto

$$\frac{\|a(x) - a(\hat{x})\|}{\|a(x)\|} = \frac{\|f(x) - f(\hat{x})\|}{\|f(x)\|}.$$

Potom

$$\frac{\frac{\|f(x) - f(\hat{x})\|}{\|f(x)\|}}{\frac{\|x - \hat{x}\|}{\|x\|}} \leq \sup_{\frac{\|x - \tilde{x}\|}{\|x\|} \leq \delta} \frac{\frac{\|f(x) - f(\tilde{x})\|}{\|f(x)\|}}{\frac{\|x - \tilde{x}\|}{\|x\|}} \approx \kappa_f(x)$$

pro „rozumně“ spojitý problém  $f$ . Tj.

$$\frac{\|a(x) - fl(a(x))\|}{\|a(x)\|} \lesssim \kappa_f(x) O(u).$$

Tedy přesnost zpětně stabilního algoritmu je (může být) ovlivněna podmíněností problému a strojovou přesností.

## 2 Schurův rozklad

### Definice 2.1

V dalším bude  $\|x\|$  euklidovská norma, stejně tak  $\|A\|$  bude tzv. spektrální norma, která je odvozena právě od Eukleidovské.  $\kappa(S) = \|S\| \cdot \|S^{-1}\|$  je číslo podmíněnosti  $S$ .

#### Poznámka

Jestliže jsou vstupní data (matice  $A$ ) zatížena chybou  $E$ , tak transformace prováděné na  $A$  budeme provádět i na  $E$ .

#### Poznámka

Matice, které nám nezvětší chybu (tedy ty, co bychom chtěli používat) jsou unitární matice. Platí totiž

$$\|x\| = \sqrt{x^*x} = \sqrt{x^*U^*Ux} = \|Ux\|.$$

$$\kappa(U) = 1, \kappa(U^*EU) = \kappa(E).$$

Tomuto se říká unitární invariantnost spektrální normy. Unitární invariantnost Frobeniovy normy se dokáže také jednoduše  $\|A\|_F^2 = \text{trace}(A^*A) = \text{trace}(A^*U^*UA)$ .

Naopak platí i opačná implikace ( $\kappa = 1 \implies$  unitarita), ale tu dokážeme později.

### Věta 2.1 (Schurova)

Pro libovolnou čtvercovou matici  $A$  existuje unitární matice  $U$  tak, že

$$R = U^*AU,$$

kde  $R$  je horní trojúhelníková matice s vlastními čísly matice  $A$  na diagonále v libovolném předepsaném pořadí.

┌

*Důkaz*

Indukcí podle řádu matice  $A$ . Pro řád  $n = 1$  věta platí. Předpokládejme, že věta platí pro  $n$ : Nechť je dáno  $A \in \mathbb{C}^{(n+1) \times (n+1)}$  a uspořádání vlastních čísel. Označme  $\lambda$  první vlastní číslo v tomto uspořádání,  $x$  je příslušný vektor. Doplníme  $x$  na čtvercovou unitární  $H = (x|X)$ ,

$$H^*AH = \begin{pmatrix} x^*Ax & x^*AX \\ X^*Ax & X^*AX \end{pmatrix} = \begin{pmatrix} \lambda & b^* \\ 0 & C \end{pmatrix}.$$

Na matici  $C$  aplikujeme předpoklad indukčního kroku. Součinem  $H$  a získané matice (rozšířené o  $\text{id}$  na první souřadnici).

└

*Poznámka*

Hledání Schurova rozkladu lze převést na hledání kořenů polynomu (a zpět), tedy (z neřešitelnosti hledání kořenů polynomu řádu většího než 5) Schurův rozklad nelze spočítat obecně přesně v konečném množství kroků.

## Definice 2.2 (Kvazi-trojúhelníková matice)

Čtvercová matice  $T$  je horní kvazi-trojúhelníková, pokud je blokově horní trojúhelníková a na diagonále má bloky  $1 \times 1$  nebo  $2 \times 2$ .

## Věta 2.2

Nechť  $A$  je reálná čtvercová matice. Potom existují ortogonální  $U$  a reálná kvazi-trojúhelníková  $T$  takové, že

$$T = U^T A U,$$

Navíc vlastní čísla každého  $2 \times 2$  diagonálního bloku matice  $T$  tvoří komplexně sdružený pár.

┌

*Důkaz*

└ Bez důkazu

□

*Důsledek*

$A \in \mathbb{C}^{n \times n}$  je normální  $\Leftrightarrow \exists$  unitární  $U$  a diagonální  $D$  tak, že

$$U^* A U = D, \quad A = U D U^*.$$

*Důsledek (Důsledku)*

$$A u_j = \lambda_j u_j.$$

*Důsledek* (Dyadický rozvoj)

$$A = UDU^* = \sum_{i=1}^n \lambda_i u_i u_i^*.$$

*Poznámka*

$$\|\lambda_i u_i u_i^*\| = |\lambda_i|.$$

### Věta 2.3 (O hustotě diagonalizovatelných matic)

$\forall A \in \mathbb{C}^{n \times n}$  a pro libovolně malé  $\varepsilon > 0 \exists$  diagonalizovatelná  $A_\varepsilon \in \mathbb{C}^{n \times n}$  s navzájem různými vlastními čísly tak, že  $\|A - A_\varepsilon\| < \varepsilon$ .

*Důkaz*

Schurova věta nám dává rozklad  $A = URU^*$ . Není-li  $A$  diagonalizovatelná, pak musí mít alespoň jedno násobné vlastní číslo. Zvolme si

$$R_\varepsilon = R + D_\varepsilon,$$

aby na diagonále  $R_\varepsilon$  byly různé hodnoty a  $\|D_\varepsilon\| < \varepsilon$ . Uvažujme  $A_\varepsilon := UR_\varepsilon U^*$ .  $A_\varepsilon$  má různá vlastní čísla, tedy je diagonalizovatelná a

$$\|A - A_\varepsilon\| = \|URU^* - U(R + D_\varepsilon)U^*\| = \|UD_\varepsilon U^*\| = \|D_\varepsilon\| < \varepsilon.$$

□

## 3 Ortogonální transformace

### 3.1 Givensovy rotace

#### Definice 3.1 (Matice Givensovy rotace)

Matice

$$G(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

nazveme maticí Givensovy rotace.

Obdobně definujeme  $G_{ij}(\varphi)$  (matici elementární Givensovy rotace) pro více rozměrů, kde  $G_{ij}(\varphi)$  otáčí v rovině  $x_i x_j$ .

*Poznámka*

Vynásobením maticí  $G$  rotace lze např. vynulovat složku vektoru (resp. postupně všechny

až na 1).

$$\sin \varphi = -\frac{x_j}{\sqrt{x_i^2 + x_j^2}}, \cos \varphi = \frac{x_i}{\sqrt{x_i^2 + x_j^2}}.$$

## 3.2 Householderovy reflexe

### Definice 3.2

Nechť  $q \in \mathbb{R}^n$  a  $\|q\| = 1$ . Pak matici

$$H(q) = I - 2qq^T \in \mathbb{R}^{n \times n}$$

nazýváme maticí Householderovy reflexe vzhledem k nadrovině určené vektorem  $q$ .

*Poznámka*

Tato matice je zřejmě ortonormální symetrická a  $H^2(q) = I$ ,  $\det(H(q)) = -1$ .

### Věta 3.1 (O zrcadlení)

Nechť jsou dány dva různé vektory  $x \in \mathbb{R}^n$  a  $y \in \mathbb{R}^m$ ,  $\|x\| = \|y\|$ , a nechť

$$q_1 := \frac{x - y}{\|x - y\|}, \quad q_2 := \frac{x + y}{\|x + y\|}.$$

Potom

$$H(q_1)x = y, \quad H(q_2)x = -y.$$

*Důkaz*

Vektor  $x - y$  je kolmý k nadrovině zrcadlení vektoru  $x$  na vektor  $y$ . □

*Poznámka*

I pomocí Householderovy reflexe se dá nulovat, za  $y$  zvolíme  $\pm\|x\|e_1$  a najdeme  $H(q)$ . Můžeme zde narazit na zrušení (pokud je první složka  $x$  „nejvýraznější“), ale té se vyhneme tak, že správně zvolíme znaménko.

## 3.3 QR rozklad

### Definice 3.3 (QR rozklad, ekonomický tvar)

Buď  $A \in \mathbb{C}^{n \times m}$  obecně obdélníková. Rozklad

$$A = QR,$$

kde  $Q$  je matice s ortonormálními sloupci ( $Q^*Q = I$ ),  $R$  má všechny prvky pod hlavní diagonálou nulové.



Když odstraníme nulové řádky/sloupce, tak dostaneme tzv. ekonomický tvar QR rozkladu.

*Poznámka*

Lze ho dostat například Gram-Smidtovy ortogonalizace.

### Definice 3.4 (Hessenbergův tvar)

Matice v horním Hessenbergově tvaru je matice, která je „skoro horní trojúhelníková“ tak, že má ještě jednu diagonálu (pod hlavní) navíc potenciálně nenulovou.

*Poznámka* (Použití QR rozkladu)

Řešení soustav, hledání Schurova rozkladu pomocí QR algoritmu: Převědeme na Hessenbergův tvar a následně iteračně  $A_i = Q_i R_i$ ,  $A_{i+1} = R_i Q_i$ .

### Definice 3.5 (QR rozklad pomocí reflexí či rotací)

Vynulujeme poddiagonální prvky  $A$  tak, že nejprve vynulujeme téměř celý první sloupec, pak druhý (bez dvou prvků), pak třetí, ...

*Poznámka*

Existence plyne z konstrukce, jednoznačný obecně není (pouze pokud  $\text{diag}(R)$  jsou jen kladná čísla).

*Poznámka*

Dále se probírala projekce, viz LA. (Projektor = matice projekce, komplementární projektor =  $I -$  matice projekce). A Gram-Smidt + modifikovaná verze.

### Definice 3.6 (Iterační zpřesnění)

Postačuje iterovat dvakrát = projektujeme originální vektor a projektujeme výsledek (= rozdíl originálního vektoru a jeho projekce). (Zde klasický i modifikovaný GS dává podobné)

### Definice 3.7 (Cena výpočtu)

Cenu výpočtu měříme jako počet aritmetických operací (tzv. flops).

*Poznámka*

GS potřebuje  $O(n^3)$  operací.

### Definice 3.8 (Ztráta ortogonality)

Ztrátu ortogonality matic vypočtených v  $\mathbb{F}$  budeme měřit pomocí  $\|E_Q\|_2$ , kde  $E_Q := \hat{Q}^* \hat{Q} - I$ .

┌

*Například*

Householderův QR rozklad má ztrátu  $O(u)$  stejně jako Givensův a ICGS. CGS má  $\kappa^2(A)O(u)$  a MGS má  $\kappa(A)O(u)$ .

└

*Poznámka*

Householder, Given i MGS jsou zpětně stabilní.

### Definice 3.9 (Norma rezidua $A - \hat{Q}\hat{R}$ )

Matice  $\hat{Q}$  a  $\hat{R}$  vypočtené všemi uvažovanými variantami QR rozkladů splňují

$$\|A - \hat{Q}\hat{R}\| \approx u\|A\|.$$

## 4 LU rozklad

### Definice 4.1 (Eliminační matice)

Eliminační matice je  $n \times n$  matice  $M_k$ , která má na diagonále jedničky, v jednom sloupci má pod touto jedničkou nějaké prvky a jinde jsou nuly.

┌

*Poznámka*

Součin eliminačních matic je součet jich bez  $I$  sečtený s  $I$ .

└

### Definice 4.2 (LU rozklad)

LU rozklad je rozklad matice  $A \in \mathbb{C}^{n \times n}$  na součin  $L \cdot U$ , kde  $L$  je dolní trojúhelníková s jedničkami na diagonále a  $U$  je horní trojúhelníková.

*Poznámka*

S LU rozkladem lze počítat rovnice tak, že nejprve vyřešíme  $Ly = b$  a pak  $Ux = y$ .

### Věta 4.1 (O proveditelnosti GE)

Gausovu eliminaci (bez prohazování řádků) lze provést právě tehdy, je-li matice  $A$  silně regulární (tj. každá hlavní  $(11, 1122, 1133, 1144, \dots)$  podmatice  $k \times k$  je regulární).

┌  
Důkaz

Indukcí (podíváme se, co se děje s hlavními podmaticemi a zjistíme, že se upravují tak, že jsou regulární právě tehdy, když po úpravách prvek vpravo dole není nulový).  $\square$

## Lemma 4.2

*Je-li  $A \in \mathbb{C}^{n \times n}$  hermitovská pozitivně definitní (HPD), pak je silně regulární.*

## Definice 4.3 (Choleského rozklad)

Nechť  $A$  je (HPD). Pak lze provést GE a získat  $LU = A = A^* = U^*L^* \implies UL^{-1*}U^* = L^{-1}U^*$  a  $D := UL^{-1*}$  je diagonální. Tudíž

$$A = LDL^*.$$

Navíc  $\forall x \neq 0 : 0 < (L^*x)^*D(L^*x)$  neboli

$$A = LD^{\frac{1}{2}}D^{\frac{1}{2}}L^* = \tilde{L} \cdot \tilde{L}^*.$$

Tento jednoznačný rozklad nazveme Choleského.

*Poznámka*

Ch rozklad se používá např. pro určení pozitivní definitnosti matice.

## Definice 4.4 (GE s částečnou pivotací (GEPP))

Provádíme GE za pomoci eliminačních matic a zároveň prohazujeme řádky, aby aktuálně zpracovávaný pivot byl co největší (z ještě nepoužitých řádků). (Prohození řádků nám jen prohodí řádky v maticích, nemění to dosavadní hodnoty ani eliminovanou část.)

## Tvrzení 4.3 (Proveditelnost GEPP)

*GE s částečnou pivotací funguje pro všechny regulární matice.*

┌  
Důkaz

Zřejmý.  $\square$

## Tvrzení 4.4 (Zpětná stabilita GE)

$$A + E = \hat{L}\hat{U}$$

$$\|E\|_{\infty} \leq 2nu\|\hat{L}\|_{\infty}\|\hat{U}\|_{\infty} + O(u^2).$$

*GE jako taková stabilní není, ale s částečnou pivotací už je tzv. podmíněně stabilní,*

jelikož je to GE na  $P \cdot A$ , kde  $P$  je matice permutace a jelikož pro ni  $\hat{L} \leq 1$ , tedy

$$\frac{\|\Delta A\|_\infty}{\|A\|_\infty} \leq 6n^2 u \frac{\|\hat{U}\|_\infty}{\|A\|_\infty} + O(u^2),$$

tedy pokud  $\hat{U}$  příliš neroste (pozor, nesouvisí s podmíněností).

#### Definice 4.5 (Škálování)

Rovnici  $Ax = b$  lze přenásobit diagonální maticí, čímž lze zpřesnit algoritmus.

#### Tvrzení 4.5

Nechť  $\hat{L}$  je vypočtený faktor Choleského rozkladu HPD matice  $A \in \mathbb{C}^{n \times n}$  a nechť  $2n^{\frac{3}{2}}u < 1$ . Potom pro  $E$  takovou, že  $A + E = \hat{L}\hat{L}^*$ , platí

$$\|E\|_F \leq \left( \frac{2n^{\frac{3}{2}}}{1 - 2n^{\frac{3}{2}}u} \right) u \|A\|_F + O(u^2).$$

#### Definice 4.6

Iterační zpřesnění řešení  $Ax = b$  je to, že spočítáme  $\hat{x}$ , následně reziduum  $r = b - A\hat{x}$ , a s ním jako pravou stranou řešíme další soustavu a výsledek přičteme k  $\hat{x}$  a počítáme znovu reziduum...

Předpokládejme, že  $n\kappa(A) \ll 1$ . Navíc abychom dosáhli lepšího výsledku, je třeba reziduum spočítat ve vyšší přesnosti než ostatní operace.

#### Definice 4.7 (Řídká matice)

Matice, kterou je lepší uložit jako seznam nenulových prvků s jejich indexy.

#### Poznámka

Pro řídké matice je třeba matici nejprve správně zpermutovat než budeme provádět GE, aby nevznikalo moc nových prvků.

## 5 Singulární rozklad

#### Věta 5.1 (Vztahy mezi spektrálními rozklady $A^*A$ a $V^*$ )

Uvažujme spektrální rozklad  $A^*A$ . Potom jsou

$$u_j := \frac{Av_j}{\sqrt{\lambda_j}}, \quad j \in [r],$$

ortonormální vlastní vektory matice  $AA^*$  a platí

$$AA^*u_j = \lambda_j u_j, \|u_j\| = 1, \quad j \in [r].$$

*Důsledek*

Nenulová vlastní čísla  $A^*A$  a  $AA^*$  se rovnají.  $Av_j = u_j\sqrt{\lambda_j}$ .

$u_1, \dots, u_r$  je ortonormální báze  $R(AA^*) = R(A)$ . Doplníme-li  $u_1, \dots, u_r$  na ortonormální bázi prostoru  $\mathbb{C}^n$  pomocí  $u_{r+1}, \dots, u_n$ , dostaneme  $N(AA^*) = N(A^*)$ .

$U := [u_1, \dots, u_n]$  je unitární a  $AA^*U = U \operatorname{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$ .

### Definice 5.1 (Singulární čísla)

$Av_j = \sigma_j u_j$ ,  $\sigma_j := \sqrt{\lambda_j}$ ,  $j \in [r]$ .  $\sigma_j$  nazveme singulární čísla matice  $A$  a označíme

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0.$$

### Věta 5.2 (O singulárním rozkladu)

Matice  $A \in \mathbb{C}^{n \times m}$  hodnosti  $r$  lze rozložit na součin

$$A = U\Sigma V^*,$$

kde  $U \in \mathbb{C}^{n \times n}$  a  $V \in \mathbb{C}^{m \times m}$  jsou unitární,  $\Sigma \in \mathbb{R}^{n \times m}$ , kde  $\Sigma = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}$ ,  $\Sigma_r = \operatorname{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r \times r}$ .

### Tvrzení 5.3

Pro HPD matice je singulární rozklad totéž jako spektrální.

*Poznámka*

Singulární rozklad má také ekonomický tvar i dyadický rozvoj.

*Poznámka (Aplikace – normy)*

Spektrální a Frobeniovu normu díky tomu spočítáme jako

$$\|A\| = \|U\Sigma V^*\| = \|\Sigma\| = \sigma_1,$$

$$\|A\|_F = \|\Sigma\|_F = \sqrt{(\sigma_1^2 + \dots + \sigma_r^2)}.$$

Tedy  $\|A\| \leq \|A\|_F \leq \sqrt{r}\|A\|$  a  $\kappa(A) := \|A\| \cdot \|A^{-1}\| = \frac{\sigma_1}{\sigma_n}$ .  $\kappa(A) = 1 \Leftrightarrow \gamma U$ , kde  $\gamma \neq 0$  a  $U$  je unitární.

### Definice 5.2 (Zobecnění čísla podmíněnosti)

Zobecníme  $\kappa(A)$  pomocí veličin

$$\min \text{mag}(A) := \max_{\|x\|=1} \|Ax\|, \max \text{mag}(A) := \min_{\|x\|=1} \|Ax\|,$$

jako

$$\kappa(A) := \frac{\min \text{mag}(A)}{\max \text{mag}(A)}.$$

### Definice 5.3 (Pseudoinverze)

Pseudoinverzi (Moore-Penroseovu zobecněnou inverzi) matice  $A \in \mathbb{C}^{n \times m}$  definujeme jako  $A^\dagger := V\Sigma^\dagger U^*$ , kde

$$\Sigma^\dagger = \begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

### Tvrzení 5.4

*Pseudoinverze splňuje následující (a je tím jednoznačně určena):*

$$AA^\dagger A = A, \quad (AA^\dagger)^* = AA^\dagger, \quad A^\dagger AA^\dagger = A^\dagger, \quad (A^\dagger A)^* = A^\dagger A.$$

┌

*Důkaz*

Jednoznačnost se upočítá převodem jedné na druhou pomocí rovností z tvrzení. □

### Lemma 5.5 (O vyjádření pseudoinverze)

*Nechť  $A \in \mathbb{C}^{n \times m}$ . Je-li  $n \geq m$  a  $\text{rank}(A) = m$ , potom*

$$A^\dagger = (A^* A)^{-1} A^*.$$

*Je-li  $n < m$  a  $\text{rank}(A) = n$ , potom  $A^\dagger = A^* (A A^*)^{-1}$ .*

┌

*Důkaz*

Pro  $n \geq m = r$  je  $V_m$  unitární čtvercová,

$$A^\dagger = V_m \Sigma_m^{-1} U_m^* = (V_m \Sigma^{-2} V_m^*) (V_m \Sigma U_m^*) = (A^* A)^{-1} A^*.$$

└

Druhá analogicky. □

### Lemma 5.6 (O vzdálenosti nejlepší aproximace hodnoci $k$ )

Nechť  $A \in \mathbb{C}^{n \times m}$ . Potom  $\forall X \in \mathbb{C}^{n \times m}$  hodnosti  $k < r$  je

$$\|A - X\| \geq \sigma_{k+1}.$$

### Věta 5.7 (Eckart, Young, Mirsky)

Dána  $A \in \mathbb{C}^{n \times m}$  hodnosti  $r$

$$A = \sum_{j=1}^r \sigma_j u_j v_j^*.$$

Potom je pro  $k < r$  matice

$$A^{(k)} := \sum_{j=1}^k \sigma_j u_j v_j^*$$

nejlepší aproximace  $A$  hodnosti  $k$ ,  $\|A - A^{(k)}\| = \sigma_{k+1}$ .

Důsledek

Nejbližší singulární matice k  $A$  je  $\frac{\sigma_n}{\sigma_1} = \frac{1}{\kappa(A)}$ .

### Definice 5.4 (Numerická hodnota)

Numerická hodnota  $A$  je takové  $k$ , že

$$\sigma_1 \geq \dots \geq \sigma_k \gg \sigma_{k+1} \approx u.$$

Poznámka (Výpočet SVD)

Buď pomocí spektrálního rozkladu  $AA^*$  resp.  $A^*A$ . To má nevýhodu  $\kappa(A^*A) = \frac{\sigma_1^2}{\sigma_m^2} = \kappa(A)^2$ .

Nebo algoritmus Golub-Kahan-Reinsch – aplikace Householderových reflexí střídavě zprava a zleva, tím dostaneme tzv. bidiagonální tvar:

$$B = WAQ, W^*W = I, Q^*Q = I.$$

Následně iteračně pomocí varianty QR algoritmu spočítáme singulární rozklad  $B$ . Takto dostaneme přesnost až na úrovni  $u$ .

## 6 Problém nejmenších čtverců

### Definice 6.1 (Problém nejmenších čtverců (LS))

Nechť  $A \in \mathbb{C}^{n \times m}$ ,  $b \in \mathbb{C}^n$ . Problémem nejmenších čtverců (LS) budeme nazývat úlohu

určení  $x \in \mathbb{C}^m$ , které minimalizuje

$$\|b - Ax\|.$$

### **Definice 6.2** (Úplný problém nejmenších čtverců)

Stejně jako LS, jen budeme minimalizovat  $[\Delta b, \Delta A]_F$ , kde

$$(A + \Delta A)x = b + \Delta b.$$

### **Věta 6.1** (O řešení LS)

Nechť  $A \in \mathbb{C}^{n \times m}$ ,  $b \in \mathbb{C}^n$ .

$$x \text{ řeší LS} \Leftrightarrow Ax = b|_{R(A)}.$$

┌ Důkaz

└ V LA. □

### **Věta 6.2** (Řešení LS a soustava normálních rovnic)

Nechť  $A \in \mathbb{C}^{n \times m}$ ,  $b \in \mathbb{C}^n$ . Vektor  $x$  je řešením problému LS  $\Leftrightarrow$  je-li řešením soustavy normálních rovnic,

$$A^*Ax = A^*b.$$

┌ Důkaz

└ V LA. □

*Poznámka*

Má-li  $A$  plnou sloupcovou hodnotu, je  $A^*A$  regulární a soustava normálních rovnic má jednoznačné řešení,

$$x = A^\dagger b.$$

### **Věta 6.3** (Řešení LS minimální v normě)

Nechť  $A \in \mathbb{C}^{n \times m}$ ,  $b \in \mathbb{C}^n$ . Potom existuje právě jedno řešení  $x$  problému LS minimální v normě a je dáno vztahy

$$Ax = b|_{R(A)} \quad Ax \in R(A^*).$$

┌ Důkaz

└ Rozepíšeme z Pythagorovy věty a omezíme normu. □



*Poznámka (Řešení LS)*

Buď pomocí  $A^*Ax = A^*b$ , kde použijeme Choleského rozklad.

Nebo použijeme QR rozklad a řešíme  $\min_x \|b - QRx\| = \min_x \|Q^*b - Rx\|$  a

$$\begin{pmatrix} Q_m^* b \\ \tilde{Q}_m^* b \end{pmatrix} - \begin{pmatrix} R_m \\ 0 \end{pmatrix} x = \begin{pmatrix} Q_m^* b - R_m x \\ \tilde{Q}_m^* b \end{pmatrix}.$$

Rozšířenou soustavou rovnic (tzv. sedlobodová matice):

$$A^*Ax = A^*b \Leftrightarrow A^*(b - Ax) = 0,$$

$$\begin{pmatrix} I & A \\ A^* & 0 \end{pmatrix} \begin{pmatrix} b - Ax \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

### Lemma 6.4 (O sedlobodové matici)

*Uvažujme matici*

$$C = \begin{pmatrix} B & A \\ A^* & 0 \end{pmatrix},$$

kde  $B \in \mathbb{C}^{n \times n}$  je hermitovská pozitivně definitní a  $A \in \mathbb{C}^{n \times m}$ ,  $n \geq m$ , má plnou sloupcovou hodnot. Potom je  $C$  regulární, hermitovská a indefinitní.

┌

*Důkaz*

Hermitovskost zřejmá, regularita:

$$\begin{pmatrix} B & A \\ A^* & 0 \end{pmatrix} \begin{pmatrix} \delta \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \delta \\ x \end{pmatrix} = 0.$$

Protože pro první rovnici  $B\delta + Ax = 0$  přenásobíme  $A^*B^{-1}$  a s využitím  $A^*\delta = 0$  získáme  $A^*B^{-1}Ax = 0$ .  $A^*B^{-1}A$  je regulární, tj.  $x = 0$  a  $B\delta = 0 \implies \delta = 0$ .

Indefinitnost dostaneme ze spektrálního rozkladu  $C = U\Lambda U^*$ . Pro  $v \neq 0$ ,  $v = (0, \nu_{n+1}, \dots, \nu_{n+m})^T$  platí

$$v^*Cv = v^*U\Lambda U^*v = \sum_{j=1}^{m+n} \lambda_j |u_j^*v|^2,$$

└ tedy musí existovat kladná i záporná čísla. □

### Věta 6.5 (LS a rozšířená soustava rovnic)

$A \in \mathbb{C}^{n \times m}$ ,  $n \geq m$ ,  $b \in \mathbb{C}^n$ ,  $\text{rank}(A) = m$ . Řešení  $x$  problému LS lze nalézt řešením rozšířené soustavy

$$\begin{pmatrix} I & A \\ A^* & 0 \end{pmatrix} \begin{pmatrix} b - Ax \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

**Věta 6.6** (LS a singulární rozklad)

Nechť  $A \in \mathbb{C}^{n \times m}$ ,  $b \in \mathbb{C}^n$ ,  $A = U_r \Sigma_r V_r^*$  je ekonomický SVD, pak

$$x = A^\dagger b$$

je řešení LS minimální v normě.

┌

*Důkaz*

Řešení  $x$  minimální v normě je dáno jednoznačně

$$Ax = b|_{R(A)} \wedge x \in R(A^*).$$

Ze SVD známe ortogonální projektor na  $R(A)$ :

$$b|_{R(A)} = (U_r U_r^*)b.$$

Dosazením za  $A$  a  $b|_{R(A)}$  dostaneme

$$U_r \Sigma_r V_r^* x = U_r U_r^* b,$$

$$V_r^* x = \Sigma_r^{-1} U_r^* b.$$

Ve  $V_r$  je ortonormální báze  $R(A)$  a  $x \in R(A^*) \implies x = V_r y$ . Dosazením dostaneme  
 $y = V_r^* x = \Sigma_r^{-1} U_r^* b.$  □

└

## 7 Klasické iterační metody

**Definice 7.1** (Matrixfree computation)

Výpočet s maticí, kterou ani nemáme nikde uloženou (jen mohu např. dopočítávat její prvky).

*Poznámka*

Když provádíme matrixfree výpočty nebo výpočty s velmi řídkými metodami, tak rozklady nefungují a musíme použít iterační metody.

**Definice 7.2** (Částečný problém vlastních čísel)

Úkolem je nalézt část vlastních čísel. (Úplný problém vlastních čísel, tj. najít všechna čísla se musí řešit Schurovým rozkladem.)

### Definice 7.3 (Mocninná metoda)

Nechť  $A \in \mathbb{C}^{n \times n}$  je diagonalizovatelná,  $A = S\Lambda S^{-1}$ ,  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  vlastní čísla  $A$ ,  $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$ ,  $S = [s_1, \dots, s_n]$  normované vlastní vektory.

Pro jednoduchost nechť  $|\lambda_1| > |\lambda_2|$ . Nechť  $v$  je nenulový startovací vektor. Mocninná metoda je pak:  $v_0 = \frac{v}{\|v\|}$  následované opakováním  $w = Av_{k-1}$ ,  $v_k = \frac{w}{\|w\|}$ ,  $\mu_k = v_k^* Av_k$ .

*Poznámka*

$v_k = \frac{A^k v}{\|A^k v\|}$ , neboť  $v_k$  je  $A$   $k$ -krát aplikované na  $v_0$  dělené nějakými čísly, ale víme, že  $\|v_k\| = 1$ . Tedy (za předpokladu  $\zeta_1 \neq 0$ )

$$A^k v = \sum_{j=1}^n \zeta_j \lambda_j^k s_j, \quad v_k = \frac{A^k v}{\|A^k v\|} = \frac{\zeta_1 \lambda_1^k \cdot \sum_{j=2}^n \frac{\zeta_j}{\zeta_1} \left(\frac{\lambda_j}{\lambda_1}\right)^k s_j}{|\zeta_1 \lambda_1^k| \cdot \|\dots\|} \rightarrow e^{i\alpha_k} s_1,$$

tj.  $v_k \rightarrow e^{i\alpha_k} s_1$ ,  $\mu_k = v_k^* A v_k \rightarrow \lambda_1$ .

### Definice 7.4 (Modifikace mocninné metody)

$Ax = \lambda x \Leftrightarrow A^{-1}x = \frac{1}{\lambda}x$ . Tedy spočítáme LU rozklad a pomocí něho  $w = A^{-1}v_{k-1}$  jako řešení

$$Aw = v_{k-1}.$$

Navíc pokud známe aproximaci  $\mu$  vlastního čísla  $\lambda$ , pak můžeme pomocí

$$Ax = \lambda x \Leftrightarrow (A - \mu I)x = (\lambda - \mu)x \Leftrightarrow (A - \mu I)^{-1}x = \frac{1}{\lambda - \mu}x,$$

kde budu hledat  $\frac{1}{\lambda - \mu}$  jako největší vlastní číslo  $(A - \mu I)^{-1}$ .

### Definice 7.5 (Obecné iterační schéma)

Formálně rozštěpíme  $A$  na  $A = M - N$ , kde  $M$  je regulární a systém s  $M$  je lehce řešitelný. Potom

$$Ax = b \Leftrightarrow (M - N)x = b \Leftrightarrow Mx = Nx + b.$$

Iterační proces je pak, že  $x_0$  je dáno a  $Mx_k = Nx_{k-1} + b$ . Použijeme-li  $Mx = Nx + b$ , dostáváme

$$M(x - x_k) = N(x - x_{k-1}),$$

$$x - x_k = M^{-1}N(x - x_{k-1}) = \dots = (M^{-1}N)^k.$$

Tedy pokud  $(M^{-1}N)^k$  konverguje k 0, tak nám chyba konverguje k 0.

$M^{-1}N$  můžeme přepsat jako  $I - M^{-1}A$ . Ideálně tedy  $A \approx M$ . Stejně tak  $Mx_k = Nx_{k-1} + b$  můžeme přepsat na  $x_{k-1} + M^{-1}(b - Ax_{k-1})$ , kde  $r_{k-1} := b - Ax_{k-1}$  nazveme reziduum.

**Definice 7.6** (Spektrální poloměr)

Pro  $B \in \mathbb{C}^{n \times n}$  nazveme  $\varrho(B) := \max \{|\lambda| \mid \lambda \text{ je vlastní číslo } B\}$  spektrálním poloměrem matice  $B$ .

**Věta 7.1** (Oldenburgova)

Nechť  $B \in \mathbb{C}^{n \times n}$ . Platí

$$B^k \rightarrow 0 \Leftrightarrow \varrho(B) < 1.$$

┌

*Důkaz*

└ Jordanův rozklad. □

*Důsledek*

Posloupnost  $x_k$  generovaná

$$x - x_k = (M^{-1}N)^k(x - x_0)$$

konverguje k řešení  $\forall x_0$  právě tehdy, když

$$\varrho(M^{-1}N) < 1.$$

**Tvrzení 7.2** (Postačující podmínka konvergence iteračního schématu)

$\|M^{-1}N\| < 1 \implies \varrho(M^{-1}N) < 1$ , kde  $\|\cdot\|$  je generovaná nebo Frobeniova norma.

┌

*Důkaz*

$$|\lambda| = \|\lambda v\| = \|Bv\| \leq \|B\|.$$

└ □

┌

*Poznámka*

└ Platí pro každou multiplikativní normu.

**Definice 7.7** (Richardsonova metoda)

$M = \frac{1}{\omega}I$ ,  $N = \frac{1}{\omega}I - A$ , iterační matice je pak  $I - \omega A$  a iterační schéma

$$x_k = x_{k-1} + \omega(b - Ax_{k-1}).$$

Vhodná, pokud je  $A$  blízka násobku jednotkové matice.

**Definice 7.8** (Jacobiova iterační metoda)

Dále uvažujme štěpení ve tvaru  $A = D - L - U$ , kde  $D$  je diagonála,  $L$  je dolní trojúhelníková a  $U$  je horní. Pak (pokud jsme  $A$  zpermutovali tak, aby na diagonále nebyly 0)

$$M = D, N = L + U,$$

iterační matice je  $I - D^{-1}A$ , iterační schéma  $x_k = x_{k-1} + D^{-1}(b - Ax_{k-1})$ .

Vhodná, pokud je  $A$  blízká diagonální matici.

**Definice 7.9** (Gauss-Seidelova metoda)

Většinou Jacobiovu iterační metodu najdeme rozepsanou po složkách:

$$\zeta_{k,i} = \frac{1}{a_{i,i}} \left( - \sum_{j=1, j \neq i}^n A_{i,j} \zeta_{k-1,j} + b_i \right).$$

Tady můžeme dokonce využít i již spočtené složky ( $\zeta_{j,k}$  pro  $j < i$ ), tedy přesněji spočítané složky, pak dostaneme tzv Gauss-Seidelovu metodu. Navíc si stačí pamatovat jen 1 vektor.

**Definice 7.10** (Superrelaxační metoda)

$$A = \frac{1}{\omega} D - L - \left[ \left( \frac{1}{\omega} - 1 \right) D + U \right].$$

Tou se dále nebudeme zabývat.

**Definice 7.11** (Ostře diagonálně dominantní)

O matici  $A \in \mathbb{C}^{n \times n}$  budeme říkat, že je ostře diagonálně dominantní, pokud

$$\sum_{j=1, j \neq i}^n |a_{i,j}| < |a_{i,i}|, \forall i \in [n].$$

**Věta 7.3** (Ostře diagonálně dominantní matice)

*Je-li  $A \in \mathbb{C}^{n \times n}$  ostře diagonálně dominantní, pak aproximace  $x_k$  spočtené Jacobiho nebo Gauss-Seidelovou metodou konvergují k řešení  $Ax = b$  pro libovolné počáteční přiblížení  $x_0$ .*

┌ *Důkaz*

Stačí ukázat  $\|D^{-1}(L + U)\| < 1$  pro nějakou generovanou maticovou normu. Zvolme  $\|\cdot\|_\infty$ . Platí (jelikož  $A$  je ostře diagonálně dominantní).

$$\|D^{-1}(L + U)\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1, j \neq i}^n \frac{|a_{i,j}|}{|a_{i,i}|} < 1.$$

Ukažme  $\|(D - L)^{-1}U\|_\infty < 1$ : Generovaná norma  $\implies \exists v, \|v\|_\infty = 1$  :

$$\|(D - L)^{-1}U\|_\infty = \|(D - L)^{-1}Uv\|_\infty =: \|u\|_\infty$$

a platí

$$(D - L)u = Uv.$$

Nechť  $s$  je takové, že  $|u_s| = \|u\|_\infty$ . Pak

$$\sum_{i=1}^s a_{s,i}u_i = - \sum_{i=s+1}^n a_{s,i}v_i$$

a odtud

$$\|u\|_\infty = |u_s| \leq \|u\|_\infty \underbrace{\sum_{i=1}^{s-1} \left| \frac{a_{s,i}}{a_{s,s}} \right|}_{\alpha} + \underbrace{\sum_{i=s+1}^n \left| \frac{a_{s,i}}{a_{s,s}} \right|}_{\beta}.$$

Platí  $\alpha + \beta < 1$  (ostře diagonálně dominantní) a

$$\|u\|_\infty \leq \frac{\beta}{1 - \alpha} < 1.$$

└

□

### Věta 7.4 (Pozitivně definitní matice)

*Je-li  $A \in \mathbb{C}^{n \times n}$  hermitovská pozitivně definitní, pak  $x_k$  generované GS metodou konvergují k řešení  $Ax = b$  pro každé  $x_0$ .*

┌  
Důkaz

$$A = D - L - L^* = \sqrt{D}\sqrt{D} - L - L^*.$$

Chceme ukázat, že vlastní čísla iterační matice  $G := (D - L)^{-1}L^*$  leží uvnitř jednotkového kruhu.

$$G = \frac{1}{\sqrt{D}} \left( I - \underbrace{\frac{1}{\sqrt{D}} L \frac{1}{\sqrt{D}}}_Z \right) \underbrace{-1 \frac{1}{\sqrt{D}} L^* \frac{1}{\sqrt{D}}}_{Z^*} \sqrt{D},$$

tj.  $G$  je podobná  $(I - Z)^{-1}Z^*$ . Uvažujme vlastní pár  $(\lambda, v)$  této matice

$$(I - Z)^{-1}Z^*v = \lambda v, v^*v = 1.$$

Potom

$$v^*Z^*v = \lambda v^*(I - Z)v = \lambda(1 - v^*Zv),$$

$$\lambda = \frac{v^*Zv}{1 - v^*Zv},$$

Stačí ukázat (z rozepsání), že  $1 - 2\Re(v^*Zv) > 0$ . Matice

$$D^{-\frac{1}{2}}(D - L - L^*)D^{-\frac{1}{2}} = I - Z - Z^*$$

je pozitivně definitní a proto

$$0 < v^*(I - Z - Z^*)v = 1 - 2\Re(v^*Zv).$$

└

□

### Definice 7.12 (Přechodový jev)

Přechodový jev je jev, kdy se při iterační metodě (při  $\|I - M^{-1}A\| > 1$  a  $\varrho(I - M^{-1}A) < 1$ ) stane, že se na chvíli  $x_k$  začne zase vzdalovat od řešení.

## 8 Metody Krylovových podprostorů

### Definice 8.1 (Krylovův podprostor)

Prostor

$$\mathcal{K}_k(A, v) := \text{span} \{v, Av, A^2v, \dots, A^{k-1}v\},$$

se nazývá  $k$ -tý Krylovův podprostor generovaný  $A$  a  $v$ .

Platí  $w_{k-1} \in \mathcal{K}_k(A, w)$ ,  $x_k \in \mathcal{K}_k(A, b)$ .

**Definice 8.2** (Metoda sdružených gradientů)

$A \in \mathbb{R}^{n \times n}$  symetrická, pozitivně definitní, problém  $Ax = b$ . Označme  $x_*$  řešení.  $A, b$  definují kvadratický funkcionál  $\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\mathcal{F}(x) := \frac{1}{2}x^T Ax - x^T b,$$

jelikož je  $A$  pozitivně definitní a symetrická, tak definuje skalární součin a normu. Platí

$$\mathcal{F}(x) = \frac{1}{2}x(x_* - x)^T A(x_* - x) - \frac{1}{2}x_*^T Ax_* = \frac{1}{2}\|x_* - x\|_A^2 - \frac{1}{2}\|x_*\|_A^2 = \frac{1}{2}\|x_* - x\|_A^2 + \mathcal{F}(x_*).$$

Tudíž  $Ax_* = b \Leftrightarrow x_*$  minimalizuje  $\mathcal{F}$ .

Funkcionál budeme minimalizovat tak, že  $x_{k+1} = x_k + \gamma_k p_k$ , kde  $\gamma_k$  je správná velikost a  $p_k$  správný směr. Protože  $\mathcal{F}$  je kvadratický, můžeme volit  $\gamma_k = \frac{p_k^T r_k}{p_k^T A p_k}$ , kde  $r_k := b - Ax_k$  je reziduum.

O reziduu víme

$$r_{k+1} = r_k - \gamma_k A p_k,$$

tedy nemusíme pro jeho výpočet znovu násobit.

Volbou  $p_k = r_k$  dostaneme metodu největšího spádu. Tato metoda však může konvergovat velmi pomalu.

**Definice 8.3** (Volba směrového vektoru)

TODO.

**Věta 8.1** (O metodě sdružených směrů)

Bud'  $A \in \mathbb{R}^{n \times n}$  symetrická pozitivně definitní,  $b \in \mathbb{R}^n$ . Necht'  $\{p_i\}_{i=0}^{n-1}$  tvoří  $A$ -ortogonální bázi  $\mathbb{R}^n$  a necht'  $x_0 \in \mathbb{R}^n$  je libovolné počáteční přiblížení. Uvažujme posloupnost vektorů  $x_{k+1} = x_k + \gamma_k p_k$ , kde

$$\gamma_k = \frac{p_k^T r_k}{p_k^T A p_k}, \quad r_k = b - Ax_k.$$

Potom  $x_n = x_*$ , kde  $x_*$  je řešení systému  $Ax = b$ . Navíc platí, že  $x_k - x_0$  je  $A$ -ortogonální projekcí počáteční chyby  $x_* - x_0$  na prostor  $S_k := \text{span}\{p_0, \dots, p_{k-1}\}$ .



┌ *Důkaz*

Vyjádřeme  $x_* - x_0$  v bázi  $p_0, \dots, p_{n-1}$ ,

$$x_* - x_0 = \sum_{i=1}^{n-1} \alpha_i p_i, \quad \alpha_i = \frac{p_i^T A(x_* - x_0)}{p_i^T A p_i}.$$

Z definice metody plyne

$$x_k - x_0 = \gamma_0 p_0 + \dots + \gamma_{k-1} p_{k-1}$$

a pro další koeficient  $\gamma_k$  platí

$$\gamma_k = \dots = \alpha_k.$$

Koeficienty  $\gamma_i$  se rovnají koeficientům  $\alpha_i$  a  $x_n = x_0 + \sum_{i=0}^{n-1} \alpha_i p_i = x_*$ .

Navíc

$$x_* - x_0 = (x_k - x_0) + (x_* - x_k)$$

└ a  $x_* - x_0 \perp_A S_k$ . □

*Důsledek*

$$\mathcal{F}(x_k) = \min_{x \in x_0 + S_k} \mathcal{F}(x).$$

Navíc

$$x_* - x_k \perp_A S_k \Leftrightarrow r_k \in S_k.$$

*Poznámka* (Jak generovat  $A$ -ortogonální bázi, metoda sdružených gradientů)

Pokud  $r_{k+1}$  je nulové, tak  $r_{k+1}$  je řešení, jinak  $r_{k+1} \perp S_{k+1} = \text{span}\{p_0, \dots, p_k\}$ . K rozšíření ortogonální báze použijeme reziduum  $r_{k+1}$ :

$$p_{k+1} = r_{k+1} - \sum_{j=0}^k c_{k,j} p_j, \quad c_{k,j} = \frac{\langle r_{k+1}, p_j \rangle_A}{\langle p_j, p_j \rangle_A}.$$

Abychom si nemuseli pamatovat všechna  $p_j$ , tak si všimneme, že

$$c_{k,j} = \frac{r_{k+1}^T A p_j}{p_j^T A p_j} = 0, \quad j < k.$$

Tj.  $p_{k+1} = r_{k+1} - c_{k,k} p_k$ .

TODO kosmetické úpravy výpočtů.

**Lemma 8.2** (CG a Krylovovy podprostory)

Po  $k$  iteracích metody sdružených gradientů s  $r_j \neq 0$ ,  $j \in [k]$ , platí

$$S_{k+1} = \mathcal{K}_{k+1}(A, r_0).$$

┌ *Důkaz*

Indukcí: pro  $k = 0$  platí triviálně. Necht  $S_k = \mathcal{K}_k(A, r_0)$ . Pak z

$$r_k = r_{k-1} - \gamma_{k-1} A p_{k-1}, \quad p_k = r_k + \delta_k p_{k-1}$$

plyne  $p_k, r_k \in \mathcal{K}_{k+1}(A, r_0)$ , a proto

$$S_{k+1} \subseteq \mathcal{K}_{k+1}(A, r_0).$$

└ Z lineární nezávislosti  $r_0, \dots, r_k$  pak plyne rovnost. □

TODO.

### Definice 8.4 (Stupeň vektoru vzhledem k matici)

Maximální dimenzi, jež mohou Krylovy podprostory generované maticí  $A$  a vektorem  $v$  dosáhnout, budeme nazývat stupněm  $v$  vzhledem k  $A$  a značit  $d(A, v)$ .

### Definice 8.5 (Arnoldiho algoritmus)

Báze  $v, Av, \dots, A^{k-1}v$  Krylovových podprostorů není vhodná k výpočtům (je velmi podmíněná). Chtěli bychom ortogonální bázi. Necht  $k < d(A, v)$  a necht  $v_1, \dots, v_k$  splňují

$$v_{i_1} \in \mathcal{K}_{i+1}(A, v) \setminus \mathcal{K}_i(A, v), \text{span} \{v_{[i]}\} = \mathcal{K}_i(A, v), i \in [k-1].$$

Potom  $Av_k \in \mathcal{K}_{k+1}(A, v) \setminus \mathcal{K}_k(A, v)$ . Idea je projekce:

$$w = Av_k - \sum_{i=1}^k h_{i,k} v_i, \quad v_{k+1} = \frac{w}{\|w\|},$$

kde  $h_{i,k}$  jsou vhodně zvolené koeficienty, aby  $w \perp v_i$ .

Tomuto algoritmu se říká Arnoldiho algoritmus a najde ortogonální bázi celého invariantního prostoru  $\mathcal{K}_d(A, v)$ . Je to vlastně QR rozklad matice  $(v_1 | Av_1 | \dots | Av_k)$ .

### Definice 8.6 (Arnoldiho metoda)

TODO.

### Definice 8.7 (Lanczosův algoritmus)

Arnoldiho algoritmus na symetrickou  $A$  se nazývá Lanczosův algoritmus. Ortogonální bázi Krylovova prostoru lze pro symetrickou matici  $A$  počítat tříčlennou rekurencí.

## Část II

# Matematická analýza

## 9 Nelineární algebraické rovnice

### Definice 9.1 (Nelineární algebraické rovnice)

Základním problémem je, že máme *spojitou* funkci  $f : [a, b] \rightarrow \mathbb{R}$  a hledáme  $f(x_*) = 0$ .

### Definice 9.2 (Metoda bisekce (půlení intervalů))

Na začátku potřebujeme dva body, kde je funkce kladná a záporná. Dokud je interval větší než požadovaná přesnost, tak interval rozpůlíme a zahodíme jeden z původních bodů tak, aby funkce v zbylých bodech byla zase kladná a záporná.

Z Darbouxovy věty víme, že mezi dvěma body v libovolné iteraci algoritmu je  $x_*$ . Zároveň interval se zmenšuje vždy na polovinu (to je pomalé, ale jisté). Nedá se zobecnit na soustavy.

### Definice 9.3 (Regula falsi)

Jako bisekce, jen nepůlíme, ale dělíme interval v poměru funkčních hodnot (vezmeme kořen lineární funkce, která odpovídá funkční hodnotě v těchto 2 bodech). Obecně je rychlejší než bisekce, ale někdy je pomalejší. Stále vždy funguje.

### Definice 9.4 (Newtonova metoda (metoda tečen))

Pracujeme pouze s jedním bodem  $x_0$ . Vychází z  $0 = f(x_*) = f(x_0) + f'(x_0)(x_* - x_0) + R \implies x_* \approx x_0 - \frac{f(x_0)}{f'(x_0)}$ .

Problém je, že metoda konverguje jen lokálně, jinak se můžeme naopak vzdalovat nebo se zacyklit. Dokonce platí věta (Li-Yoshe), že pro polynom s alespoň 3 různými kořeny v  $\mathbb{R}$  a  $p \in \mathbb{N}$  existuje  $x_0$ , tak, že Newton má periodu  $p$ . Nebo může nastat úplný chaos.

### Definice 9.5 (Řád konvergence, rychlost konvergence)

Mějme posloupnost  $\{x_n\}_{n=0}^\infty$ ,  $x_n \rightarrow x_*$ ,  $e_n := x_n - x_*$  je chyba v  $n$ -té iteraci.

Nechť  $\exists p > 0$ ,  $c > 0$ :  $\lim_{n \rightarrow \infty} \frac{|e_{n+1}|^p}{|e_n|} = C$ , řekneme, že  $\{x_n\}_{n=0}^\infty$  má řád konvergence  $p$ . Metoda je řádu  $p$ , jestliže  $\forall$  posloupnost generovaná touto metodou má řád alespoň  $p$  a alespoň jedna posloupnost má řád přesně  $p$ .

### Věta 9.1 (Lokální kvadratická konvergence Newtona)

Předpokládejme  $f \in C^2([a, b])$ ,  $f'(x_*) \neq 0$ ,  $\exists x_* \in (a, b)$ ,  $x_0$  je dostatečně (dá se rigorózně napsat jak) blízko  $x_*$ . Pak  $\{x_n\}_{n=0}^\infty$  z Newtonovy metody konverguje kvadraticky k  $x_*$ .

┌

Důkaz

1. Nechť  $x_k \rightarrow x_*$ , dokážeme že konverguje kvadraticky:

$$0 = f(x_*) = f(x_k) + f'(x_k)(x_* - x_k) + \frac{1}{2}f''(\xi_k)(x_* - x_k)^2, \quad \xi_k \in [x_k, x_*].$$

$$-\frac{f(x_k)}{f'(x_k)} + x_k - x_* = \frac{1}{2} \frac{f''(\xi_k)}{f'(x_k)} (x_k - x_*)^2 \implies e_{k+1} = \frac{1}{2} \frac{f''(\xi_k)}{f'(x_k)} e_k^2 \implies$$

$$\lim_{k \rightarrow \infty} \frac{|e_{k+1}|}{|e_k|^2} = \lim_{k \rightarrow \infty} \frac{1}{2} \left| \frac{f''(\xi_k)}{f'(x_k)} \right| = \frac{1}{2} \frac{f''(x_*)}{f'(x_*)} = C > 0.$$

2.  $x_k \rightarrow x_*$ :  $e_{k+1} = \frac{1}{2} \frac{f''(\xi_k)}{f'(x_k)} e_k^2$ .  $|f''|$  omezená na  $[a, b]$ ,  $f'(x_*) \neq 0$ ,  $f'$  je spojitá  $\implies \exists U \in \mathcal{U}(x_*) : |f'| > 0$  na  $U$ . Tedy

$$\exists M < \infty : \left| \frac{f''(x)}{f'(y)} \right| \leq M \forall x, y \in U.$$

$$|e_{k+1}| \leq \frac{1}{2} M |e_k| \cdot |e_k| \leq \frac{1}{2} |e_k|$$

└ pro dostatečně malé  $e_k$ .

□

Poznámka (Cayley 1879)

Newton funguje i pro  $\mathbb{C}$ . Ale je nemožné uhádnout, ke kterému kořeni. Viz <https://demonstrations.wolfram.com/ComplexNewtonIterationForACubicPolynomial/>

### Definice 9.6 (Zastavovací kritéria)

Konvergence, splnění nerovnice,

Absolutní = nezávisí na hodnotě cíle, relativní = chybu měřím relativně vůči hodnotě cíle.

### Definice 9.7 (Modifikace Newtona)

Diference: místo derivace používáme diferenci.

Metoda sečen: použijeme  $x_{k+1} = x_k - \frac{f(x_k)}{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}}$ . (Řád =  $\frac{1+\sqrt{5}}{2}$ .)

## 9.1 Iterace funkcí

**Definice 9.8** (Iterace funkcí)

$f(x_*) = 0$ . Metoda  $x_{k+1} = g(x_k)$ . Nechť  $x_k \rightarrow x_*$ ,  $g$  spojitě, pak  $x_* = g(x_*)$ .

**Definice 9.9** (Kontrakce)

$g : [a, b] \rightarrow \mathbb{R}$  je kontrakce (kontrahující), pokud  $\exists L < 1$  :

$$\forall x, y \in [a, b] : |g(x) - g(y)| \leq L|x - y|.$$

Tj.  $g$  je  $L$ -lipschitzovská  $L < 1$ .

**Věta 9.2** (Banachova věta o kontrakci)

Nechť  $g : [a, b] \rightarrow [a, b]$  je kontrakce. Pak  $\exists!$  pevný bod  $x_* \in [a, b]$ . Dále  $\forall x_0 \in [a, b] : x_k \rightarrow x_*$ , kde  $x_{k+1} = g(x_k)$ . Platí odtud  $|x_k - x_*| \leq \frac{L^k}{1-L}|x_1 - x_0| \rightarrow 0$  pro  $k \rightarrow \infty$ .

┌  
Důkaz

$$|x_k - x_{k-1}| = |g(x_{k-1}) - g(x_{k-2})| \leq L|x_{k-1} - x_{k-2}| \leq \dots \leq L^{k-1}|x_1 - x_0|.$$

$\{x_k\}$  je cauchyovská:  $k > l : |x_k - x_l| \leq (L^{k-1} + \dots + L^l)|x_1 - x_0| \leq \frac{L^l}{1-L}|x_1 - x_0| \rightarrow 0$ .

Tj.  $\exists x_* : x_k \rightarrow x_*$ .  $g$  je kontrahující, tedy  $g$  je spojitá  $\implies g(x_*) = g(\lim_{n \rightarrow \infty} x_n) = \lim_{n \rightarrow \infty} g(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x_*$ . □

*Poznámka* (Soustavy nelineárních rovnic)

Jsou složitější, výpočetní náklady velmi rostou s  $N$ . Lze použít Newtonovu metodu ( $x_{k+1} = x_k - J_f(x_k)^{-1}F(x_k)$ ).

## 10 Optimalizace

**Definice 10.1** (Optimalizace, nepodmíněná a podmíněná optimalizace)

Máme  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  a chceme  $\bar{x} = \arg \min_{x \in M} f(x)$ , kde  $M \subseteq \mathbb{R}^n$ . Pokud  $M = \mathbb{R}^n$ , pak je to nepodmíněná, jinak podmíněná.

Minimum můžeme hledat i lokální (a většinou nic víc neumíme).

### 10.1 Minimalizace v $\mathbb{R}$

**Definice 10.2** (Metoda zlatého řezu)

Podobně jako metoda bisekce, jen se podíváme na funkční hodnoty ve dvou bodech a posuneme kraj intervalu, který je „u“ větší hodnoty do daného bodu.

Body budeme volit tak, abychom mohli „recyklovat“ spočtené funkční hodnoty. Tedy pokud jsou krajní body  $a, b$ , tak nové body zvolíme  $a + \varrho(b - a)$  a  $b - \varrho(b - a)$ , kde  $\varrho$  je 1-zlatý řez.

┌ *Poznámka*

└ Vždy funguje. (Ale hledá pouze lokální minimum.)

### Definice 10.3 (Newtonova metoda pro minimalizaci)

Jako Newtonova metoda. Aproximujeme  $f \approx p$ ,  $p$  kvadratický polynom,  $\bar{x} \approx \arg \min p$ . Pokud je  $f \in C^2(\mathbb{R})$ , zvolíme

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2.$$

Pak hledáme  $p'(x) = 0$ , tj.  $f'(x_0) + f''(x_0)(x - x_0) = 0$ . Tedy iterujeme  $x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$ . Je to vlastně Newtonova metoda pro hledání  $f'(x) = 0$ .

## 10.2 Optimalizace v $\mathbb{R}^n$

### Definice 10.4 (Spádový směr)

Dáno  $x \in \mathbb{R}^n$ , pak směr  $d \in \mathbb{R}^n$  je spádový, pokud  $\exists \bar{\alpha} > 0 : f(x + \alpha d) < f(x) \forall \alpha \in (0, \bar{\alpha})$ .

### Definice 10.5 (Metody spádových směrů)

Dáno  $x_0 \in \mathbb{R}^n$ , chceme  $\{x_k\}_{k \in \mathbb{N}}$ ,  $x_k \rightarrow \bar{x}$ . Hledáme tuto posloupnost tvaru  $x_{k+1} = x_k + \alpha_k d_k$ , kde  $d_k$  je směr a  $\alpha \in \mathbb{R}^+$  je koeficient tak, aby  $f(x_{k+1}) < f(x_k)$ .

### Lemma 10.1

Necht  $\exists \nabla f(x)$ , pak  $d$  je spádový směr v  $x \Leftrightarrow \nabla f(x) \cdot d < 0$ .

┌ *Důkaz*

└ Triviální. □

### Lemma 10.2 (Metoda největšího spádu)

Největší spád má  $f$  ve směru  $-\nabla f(x)$ .

┌  
Důkaz

BÚNO  $\|d\| = 1$ . Derivace  $f$  ve směru  $d$ , tj.  $\nabla f(x) \cdot d$ , chceme co nejzápornější. Z Cauchy-Swarze:

$$|\nabla f(x) \cdot d| \leq \|\nabla f(x)\| \cdot \|d\| \implies \nabla f(x) \cdot d \geq -\|\nabla f(x)\|.$$

Pro  $d = \frac{-\nabla f(x)}{\|\nabla f(x)\|}$  je  $\nabla f(x) \cdot d = -\|\nabla f(x)\|$  a větší spát z nerovnosti výše být nemůže.  $\square$   
└

*Poznámka* (Volba  $\alpha_k$ )

$\alpha_k = \alpha$  dostatečně malé.

$\alpha_k \rightarrow 0$ , ale  $\sum_k \alpha_k = \infty$ , abychom mohli kamkoliv dokonvergovat, tedy například  $\alpha_k = \frac{1}{k}$ .

Line search:  $g(\alpha) = f(x_k - \alpha d_k)$ , a minimalizujeme  $g(\alpha)$  – 1D problém a stačí přibližně.

Přesný line search  $\alpha_k = \arg \min_{\alpha > 0} g(\alpha)$ . (Špatná – zig-zag efekt, dá se dokázat, že  $d_k$  a  $d_{k+1}$  jsou na sebe kolmé, takže konverguje velmi pomalu).

### Definice 10.6 (Newtonova metoda)

Funguje úplně stejně i pro  $\mathbb{R}^n$ . Může se však stát, že nenajdeme (stejně jako v  $\mathbb{R}$ , ale tady je efekt markantnější, proto ho bereme zde) minimum, ale např. inflexní bod.

Netrpí zig-zag efektem.

### Lemma 10.3

Nechť  $\exists \nabla f(x) \neq 0$ . Nechť  $B$  je pozitivně definitní matice, pak  $-B\nabla f(x)$  je spádový směr v bodě  $x$ .

┌  
Důkaz

$\nabla f(x) \cdot (-B\nabla f(x)) = -\nabla f(x)^T B \nabla f(x) < 0$ , tj.  $-B\nabla f(x)$  je spádový směr.  $\square$   
└

*Důsledek*

Newton  $-(\nabla^2 f(x_k))^{-1} \nabla f(x_k)$  je spádový směr, pokud  $\nabla^2 f(x_k)$  je pozitivně definitní.

### Definice 10.7 (Obecný algoritmus)

$$x_{k+1} = x_k - \alpha_k M_k \nabla f(x_k),$$

kde  $M_k$  je pozitivně definitní a  $\alpha_k$  určíme pomocí line search.

┌ Poznámka

Volba  $M_k$ : id = metoda největšího spádu,  $(\nabla^2 f(x_k))^{-1}$  = Newton + line search,  $(\text{diag } \nabla^2 f(x_k))^{-1}$  = diagonální aproximace Newtona,  $(\beta_k \text{id} + \nabla^2 f(x_k))^{-1}$  = kompromis mezi největším spádem a Newtonem (užitečné, pokud  $\nabla^2 f(x_k)$  není pozitivně definitní).

## 11 Ortogonální polynomy

Viz MA3 nebo UFA. Konstrukce Gram-Schmidtem jako v UFA.

### Věta 11.1 (Tříčlenná rekurence)

$\exists A_n, B_n \in \mathbb{R} : \varphi_{n+1}(x) = (x + A_n)\varphi_n(x) + B_n\varphi_{n-1}(x)$ . ( $\varphi_n$  je polynomiální ortonormální báze prostoru polynomů.)

┌ Důkaz

$x\varphi_n(x) \in P_{n+1}$  (prostor polynomů stupně nejvýše  $n+1$ )  $\implies x\varphi_n(x) = \alpha_{n+1}\varphi_{n+1}(x) + \dots + \alpha_0\varphi_0$ . TODO  $\square$

### Definice 11.1 (Legenderovy polynomy)

$(a, b) = (-1, 1)$ ,  $w(x) = 1$ , tj.  $(f, g) = \int_{-1}^1 f g dx$ . Vyjde  $1, x, \frac{2}{3}3x^2 - 1, \dots$  Rekurence vyjde  $(n+1)\mathcal{L}_{n+1}(x) = (2n+1)x\mathcal{L}_n(x) - n\mathcal{L}_{n-1}(x)$ .

### Definice 11.2 (Čebyševovy polynomy)

Na  $(-1, 1)$ ,  $w(x) = \frac{1}{\sqrt{1-x^2}}$ .

$$T_n(x) = \cos(n \arccos x).$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

┌ Důkaz (Ortogonalita)

Zintegruje se a zasubstituuje.  $\square$

### Věta 11.2

Nechť  $p_n$  je monický polynom stupně  $n$ . Pak  $\max_{x \in [-1, 1]} |p_n(x)| \geq \max_{x \in [-1, 1]} |T_n(x)| = \frac{1}{2^{n-1}}$ .



┌  
Důkaz

1.  $|T_n|$  nabývá maxima v bodech  $x_k = \cos \frac{k\pi}{n}$ ,  $k = 0, \dots, n$ . Platí  $T_n(x_k) = \frac{(-1)^k}{2^{n-1}}$ .

2. Sporem: Necht  $\exists p_n : \max_{[-1,1]} |p_n| < \frac{1}{2^{n-1}}$ . Definujeme  $m = T_n - p_n$ . Pak  $m(x_k) = \frac{(-1)^k}{2^{n-1}} - p_n(x_k) \implies m(x_k) > 0$  pro  $k$  sudé a  $< 0$  pro  $k$  liché, tj.  $\forall k \in [n] \exists$  kořen polynomu  $m \in (x_k, x_{k+1})$ , ale  $m$  má stupeň nejvýše  $n-1$ , tedy  $m = 0$ , což je spor s  $p_n < T_n$ .  $\square$

└

## 12 Interpolace

### Definice 12.1 (Interpolace)

Máme zadanou funkční hodnoty v nějakých bodech a snažíme se najít funkci, která je tam má.

### Definice 12.2 (Lagrangeova interpolace)

Dáno  $f : [a, b] \rightarrow \mathbb{R}$  a  $a \leq x_0 < x_1 < \dots < x_n \leq b$ . Hledáme  $L_n \in P_n : L_n(x_i) = f(x_i)$ ,  $i \in [n]$ .

### Definice 12.3

Říkáme, že  $L_n$  interpoluje  $f$  v uzlech  $\{x_i\}_{i=0}^n$ .

### Definice 12.4 (Lagrangeovy polynomy)

$l_i(x_j) = \delta_{ij}$ , tj.  $l_i(x) = \prod_{k \in [n] \setminus \{i\}} \frac{(x - x_k)}{(x_i - x_k)}$ ,  $i \in [n]$ .

### Věta 12.1

Dány  $f$ ,  $\{x_i\}_{i=0}^n$  jako výše. Pak  $\exists! L_n \in P_n$  interpolující  $f$  v uzlech  $\{x_i\}$ , platí  $L_n(x) = \sum_{i=0}^n l_i(x) f(x_i)$ .

*Důkaz je triviální.*

$L_n$  se nazývá Lagrangeův interpolační polynom.

### Věta 12.2 (Odhad chyby L, Cauchy)

Necht  $f \in \mathbb{C}^{n+1}[a, b]$ ,  $a \leq x_0 < x_1 < \dots < x_n \leq b$ ,  $L_n =$  Lagrangeův interpolační polynom. Pak  $\forall x \in [a, b] \exists \xi_x \in [a, b] :$

$$f(x) - L_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^n (x - x_i).$$

┌  
Důkaz

Pokud  $x = x_i$ , pak  $f(x) - L_n(x) = 0 = \dots \cdot 0$ .

Pokud  $x \neq x_i \forall i$ , pak zafixujeme  $x$  a definujeme si pomocné funkce  $g(t) = f(t) - L_n(t) - (f(x) - L_n(x)) \prod_{i=0}^n \frac{t-x_i}{x-x_i}$ .  $g(x_i) = 0$  a  $g(x) = 0$ .  $g(t)$  má  $n+2$  různých kořenů  $x_0, \dots, x_n, x$ . Navíc  $g \in C^{n+1}[a, b]$ . Z Rolleovy věty má  $g'(t)$   $n+1$  různých kořenů (mezi každými  $x_i$  nebo  $x$ ).  $g''(t)$  má  $n$  různých kořenů...  $g^{(n+1)}(t)$  má 1 kořen v  $[a, b] \implies \exists \xi_x \in [a, b] : g^{(n+1)}(\xi_x) = 0$ .

└ Vyjádříme a dostaneme výraz ze znění. □

Poznámka (Konvergence)

$L_n \rightarrow f$  pro  $n \rightarrow \infty$ ? Ne. (Tzv Rungeho jev. Například  $f(x) = \frac{1}{1+25x^2}$ .)

### Definice 12.5 (Čebyševova interpolace)

Lepší volba  $x_i =$  uzly čebyševova polynomu, tedy  $\cos\left(\frac{2i-1}{2n}\pi\right)$  na  $[-1, 1]$ . Pak  $\prod_{i=0}^n (x - x_i)$  je nejmenší možný (co do maxima).

Zde  $\forall f \in C^1[a, b]$ , pak  $L_n \rightrightarrows f$ .

### Definice 12.6 (Extrapolace)

Interpolace na  $[a, b]$  a chceme odhad  $f(x)$  pro  $x \notin [a, b]$ .

Čebyševovy polynomy se na to absolutně nehodí (nejrychleji rostou mimo interval).

Hermiteova interpolace: Snažíme se najít polynom tak, aby měl stejné i derivace do  $j$ -tého stupně.

## 12.1 Spline funkce

### Definice 12.7 (Lineární interpolační spline)

Dáno  $f : [a, b] \rightarrow \mathbb{R}$  a  $a = x_0 < x_1 < \dots < x_n = b$ ,  $x_{i+1} - x_i = h$ . Hledáme funkci  $S_L$ :

$$S_L \in C[a, b], S_L|_{I_i} \text{ je lin. funkce, } S_L(x_i) = f(x_i).$$

Dostáváme  $S_L(x)|_{I_i} = \frac{x_{i+1}-x}{h}f(x_i) + \frac{x-x_i}{h}f(x_{i+1})$ .

Jelikož je to Lagrangeova interpolace (na jednotlivých intervalech), dostáváme  $\max_{x \in [a, b]} |f(x) - S_L(x)| \leq \frac{1}{8}h^2 \max_{x \in [a, b]} |f''(x)|$ .

**Definice 12.8** (Kubické interpolační spliny)

Hledáme  $S: S \in C^2[a, b]$ ,  $S|_{I_i}$  je kubická funkce,  $S(x_i) = f(x_i)$ . Doplněním 2 podmínek dostaneme jednoznačnost. Většinou se volí  $s'(a) = f'(a)$  a  $s'(b) = f'(b)$ . Nebo  $s''(a) = f''(a)$  a  $s''(b) = f''(b)$ . Nebo tzv. přirozený kubický spline:  $s''(a) = s''(b) = 0$ .

**Definice 12.9** (Momenty splinu)

$$M_i = S''(x_i).$$

**Lemma 12.3** (Soustava lineárních rovnic pro  $M_i$ )

Pro  $i \in [n-1]$  platí

$$M_{i-1} + 4M_i + M_{i+1} = \frac{6}{h^2}(f(x_{i+1}) - 2f(x_i) + f(x_{i-1})).$$

┌

*Důkaz*

Taylor pro  $S$ :  $S(x_{i+1}) = S(x_i) + S'(x_i)h + S''(x_i)\frac{h^2}{2} + S'''(x_i, +)\frac{h^3}{6}$ . Obdobně  $S(x_{i-1}) = S(x_i) - S'(x_i)h + S''(x_i)\frac{h^2}{2} - S'''(x_i, -)\frac{h^3}{6}$ . Součtem

$$f(x_{i+1}) + f(x_{i-1}) = 2f(x_i) + M_i h^2 + (S'''(x_i, +) - S'''(x_i, -))\frac{h^3}{6}.$$

Sečteme-li Taylory pro  $S''$  dostaneme (a to už dopočítáme)

$$f(x_{i+1}) + f(x_{i-1}) = 2f(x_i) + M_i h^2 + \left(\frac{M_{i+1} - 2M_i + M_{i-1}}{h}\right)\frac{h^3}{6}.$$

└

□

*Poznámka*

Je to tedy  $A \cdot \mathbf{M} = \frac{6}{h^2} \mathbf{g}$ , kde  $g_i = f(x_{i+1}) - 2f(x_i) + f(x_{i-1})$ , resp. od  $g_1$  a  $g_{n-1}$  jsou odečtené druhé derivace  $f$  v daných bodech. Matice  $A$  je ostře diagonálně dominantní, tedy regulární a  $\exists! S$ .

*Poznámka* (Konstrukce  $S$ )

$$S(x_i) = f(x_i), S(x_{i+1}) = f(x_{i+1}), S''(x_i) = M_i, S''(x_{i+1}) = M_{i+1}.$$

$$S(x) = \frac{(x-x_i)^3}{6h}M_{i+1} + \frac{(x_{i+1}-x)^3}{6h}M_i + \left(\frac{f(x_{i+1})-f(x_i)}{h} - \frac{h}{6}(M_{i+1}-M_i)\right)(x-x_i) + f(x_i) - \frac{h^2}{6}M_i.$$

**Věta 12.4**

Nechť  $f \in C^4[a, b]$ . Pak  $\exists C_0, C_1, C_2$  nezávislá na  $f$  a  $\{x_i\}_{i=0}^n$ :

$$\max_{x \in [a, b]} |f^{(j)}(x) - S^{(j)}(x)| \leq C_j \max_{x \in [a, b]} |f^{(4)}(x)| \cdot h^{4-j}.$$

## 13 Numerická kvadratura = integrace

### Definice 13.1 (Obdélníkové pravidlo)

Chceme integrál  $I(f)$  aproximovat  $Q(f) := \sum_{i=0}^u w_i f(x_i)$ . Například

$$I(f) \approx (b-a)f\left(\frac{a+b}{2}\right) =: Q_0(f),$$

což se nazývá obdélníkové pravidlo (midpoint rule).

### Definice 13.2 (Newton-Cotes, lichoběžníkové pravidlo, Simpsonovo pravidlo)

$f \approx L_n$ ,  $Q(f) := \int_a^b L_n(x) dx$ .

$$L_n(x) = \sum_{i=0}^n f(x_i) l_i(x), \quad l_i(x) = \prod_{k \neq i} \frac{x - x_k}{x_i - x_k},$$

$$\int_a^b f(x) dx \approx \int_a^b L_n(x) dx = \sum_i f(x_i) \int_a^b l_i(x) dx = \sum_{i=0}^n w_i f(x_i) =: Q(f).$$

*Poznámka*

Takto definované  $w$  nezávisí na  $f$ .

Metoda Newton-Cotes je tato metoda pro rovnoměrné rozložení uzlů. Pro  $n = 1$  je to tzv. Lichoběžníkové pravidlo. Pro  $n = 2$  je to Simpsonovo pravidlo (Simpsons' rules), což zintegruje správně i polynom třetího stupně (jelikož  $x^3$  je liché). Pro  $n = 3$  se nazývá Simpsonovo 3/8-pravidlo, které je přesné pro polynomy 3. stupně, ale pro 4. stupně ne (jelikož  $x^4$  je sudé).  $n = 4$  je Booleovo pravidlo, které je přesné pro polynomy 5. stupně.

### Definice 13.3

$Q(f)$  má (algebraický) řád  $N \in \mathbb{N}_0$ , pokud  $Q(p) = I(p) \forall p \in P_N$ . (Polynomy stupně nejvýše  $N$ ).

### Věta 13.1

Newton-Cotes s uzly rozmístěnými rovnoměrně má řád  $n$ , pokud  $n$  je liché, a  $n+1$ , pokud je sudé.

┌ Důkaz

$$\int_a^b \left(x - \frac{a+b}{2}\right)^{n+1} dx = 0 = Q(f),$$

jelikož aproximace je zřejmě symetrická podle osy  $x = \frac{a+b}{2}$ .

└ Ostatní polynomy jdou „šoupnout“ na tento. □

### Věta 13.2 (Odhad chyby)

$Q$  ... NC kvadratura s obecnými uzly  $x_0, \dots, x_n$ . Necht  $f \in C^{n+1}[a, b]$ . Pak  $|I(f) - Q(f)| \leq \frac{1}{(n+1)!} \max_{x \in [a, b]} |f^{(n+1)}(x)| \int_a^b |\prod_{i=0}^n (x - x_i)| dx$ .

┌ Důkaz

$$|I(f) - Q(f)| = |I(f) - I(L_n)| = \left| \int_a^b f(x) - L_n(x) dx \right| \leq \int_a^b \frac{1}{(n+1)!} |f^{(n+1)}(\xi_x)| \left| \prod_{i=0}^n (x - x_i) \right| dx \leq \dots$$

└ □

### Definice 13.4 (Gaussova kvadratura)

Chceme najít  $\{x_i\}$  a  $\{w_i\}$  tak, aby  $Q(t) = \sum_{i=0}^n w_i f(x_i)$  měla co nejvyšší řád (nejvýše  $2n + 1$ , jelikož máme  $2(n + 1)$  neznámých). Pro jednoduchost na  $[-1, 1]$ .

Pro  $n = 1$  vyjde (chceme, aby bylo přesné pro  $f = 1, x, x^2, x^3$ )  $Q(t) = f(-\frac{\sqrt{3}}{3}) + f(\frac{\sqrt{3}}{3})$ .

### Věta 13.3

Necht  $\{x_i\}_{i=0}^n$  jsou kořeny Legendrova polynomu  $\mathcal{L}_{n+1}$ . Pak příslušná  $Q$  je řádu  $2n + 1$ .

┌ Důkaz

Necht  $p \in P_{2n+1}$ , chceme  $Q(p) = I(p)$ . Vydělíme  $p/\mathcal{L}_{n+1}$  se zbytkem:  $p(x) = \mathcal{L}_{n+1}(x)p_n(x) + r_n(x)$ .

$$I(p) = \int_{-1}^1 p(x) dx = \int_{-1}^1 \mathcal{L}_{n+1} p_n dx + \int_{-1}^1 r_n dx = \int_{-1}^1 r_n dx = Q(r_n) = \sum_{i=0}^n w_i r_n(x_i) = \sum_{i=0}^n w_i p(x_i) = Q(p)$$

└ □

┌ Poznámka

Neexistuje způsob jak dostat přesně uzly G. kvadratury, existují jen aproximace.

### Věta 13.4 (Odhad chyby G. kvadratury na obecném intervalu)

$f \in C^{2n+2}[a, b]$ . Pak Gaussova kvadratura splňuje

$$|I(f) - Q(f)| \leq \frac{1}{(2n+2)!} (b-a)^{2n+3} \max_{x \in [a,b]} |f^{(2n+2)}(x)|.$$

┌  
Důkaz

K uzlům přidáme libovolné další  $x_{n+1}, \dots, x_{2n+1}$ .  $L_n$  Lagrangeova interpolace na  $x_0, \dots, x_n$  a  $L_{2n+1}$  Lagrangeova interpolace na  $x_0, \dots, x_{2n+1}$ . Pak  $Q(f) := \int_a^b L_n dx = \int_a^b L_{2n+1} dx$ , neboť:  $L_n(x_i) = L_{2n+1}(x_i) = f(x_i)$ . Tedy  $L_n - L_{2n+1}$  má kořeny  $x_0, \dots, x_n$ .

Tedy  $L_n(x) - L_{2n+1}(x) = (x - x_0) \dots (x - x_n) p_n(x) = \mathcal{L}_{n+1}(x) p_n(x)$ , tudíž  $\int_a^b L_n - L_{2n+1} dx = 0$ . □

*Poznámka* (Složené kvadratické vzorce)

Interpolace nemusí fungovat pro  $n \rightarrow \infty$ . Takže se používá např. složené lichoběžníkové pravidlo, kdy se l. pravidlo aplikuje na každý interval dělení. Stejně tak složené Simpsonovo pravidlo, kde například uděláme dělení s krokem  $2h$  a pak v těchto intervalech přidáme chtěný 3. bod.

### Věta 13.5 (Chyba složeného lichoběžníkového pravidla)

TODO

┌  
Důkaz

TODO (odhadne se normálním lichoběžníkovým a  $\sum_{i=1}^m h = a - b$ ). □

### Definice 13.5 (Metoda polovičního kroku)

Doteď jsme dělali apriori odhad chyby (před výpočtem). To má nevýhodu, že nemusíme znát konstanty a derivace v odhadech.

Nyní aposteriorní odhad chyby (počítán z vypočteného výsledku).  $Q_h$  složená kvadrurní formule s krokem  $h$ . Necht' platí  $I(f) - Q_k(f) \approx Ch^N$ . Spočítáme  $Q_{\frac{h}{2}} \approx C(\frac{h}{2})^N$ . Odečteme  $Q_{\frac{h}{2}}(f) - Q_h(f) \approx Ch^N(1 - 2^{-N}) = C(\frac{h}{2})^N(2^N - 1)$ . Tedy

$$I(f) - Q_{\frac{h}{2}}(f) \approx \frac{Q_{\frac{h}{2}}(f) - Q_h(f)}{2^N - 1}.$$

Není to odhad, ale spíš indikátor.

## 14 Numerické řešení ODR

*Poznámka* (Diskretizace)

Zavedeme si dělení  $a = x_0 < x_1 < \dots < x_N = b$ . Pro jednoduchost necht' je rovnoměrné, tedy  $x_n = a + nh$ ,  $h$  je krok metody.

Budeme aproximovat správně řešení  $y(x_n)$  hodnotou  $y_n \in \mathbb{R}$ .

### Definice 14.1 (Eulerova metoda)

$$y(x_{n+1}) = y(x_n + h) \approx y(x_n) + y'(x_n) \cdot h = y(x_n) + f(x_n, y(x_n)).$$

$y_0$  je počáteční podmínka, která je dána.

### Definice 14.2 (Explicitní metoda)

Explicitní metoda je metoda, kdy máme explicitní vzorec pro  $y_{n+1}$ . Implicitní Eulerova metoda by byla  $y_{n+1} = y_n + hf(x_{n+1}, y_{n+1})$ . Zde je třeba vyřešit nějaký nelineární problém.

Tento problém odpovídá  $\frac{y_{n+1} - y_n}{h} = f(x_n, y_n)$ , čemuž budeme říkat metoda 1. řádu (zanedbaný člen je  $O(h)$ ).

### Definice 14.3 (Heunova metoda (modifikovaná Eulerova metoda))

$y' = f(x, y)$  odpovídá  $y(x_{n+1} - y(x_n)) = \int_{x_n}^{x_{n+1}} f(x, y(x))dx$ , což je hrozná aproximace integrálu. Zkusme použít lichoběžníkové pravidlo:

$$y_{n+1} - y_n = \frac{h}{2}(f(x_n, y_n) + f(x_{n+1}, y_{n+1})) \approx \frac{h}{2}(f(x_n, y_n) + f(x_n, y_n + hf(x_{n+1}, y_n))).$$

┌ *Poznámka*

└ Heunova metoda je metoda 2. řádu.

### Definice 14.4 (Jednobodové metody (JKM))

Jednobodová metoda je metoda, kde počítáme  $y_{n+1}$  pouze z  $y_n$ . Tedy  $y_{n+1} = y_n + h\varphi(x_n, y_n, h)$ , kde  $\Phi$  je tzv. přírůstková funkce.

┌ *Například*

Euler:  $\Phi(x, y, h) = f(x, y)$ .

└ Heun:  $\Phi(x, y, h) = \frac{1}{2}[f(x, y) + f(x + h, y + h \cdot f(x, y))]$ .

O  $\Phi$  předpokládáme, (že souvisí s  $f$ ) že je spojitý, že je lipschitzovské vzhledem k  $y$  (jako  $f$ ).

### Definice 14.5

Metoda je konzistentní, pokud  $\Phi(x, y, 0) = f(x, y)$ .

### Definice 14.6 (Globální chyba metody)

$$e_n = y(x_n) - y_n.$$

### Definice 14.7 (Lokální diskretizační chyba)

$$\tau(x, h) = \frac{y(x+h) - y(x)}{h} - \Phi(x, y(x), h)$$

$$\left( \frac{y_{n+1} - y_n}{h} - \Phi(x_n, y_n, h) = 0 \right)$$

### Definice 14.8

JKM (jednokroková metoda) je řádu  $p$ , pokud  $\exists k, h_0 > 0 : |\tau(x, h)| \leq Kh^p, \forall x \in [a, b], \forall h \in (0, h_0)$ .

### Lemma 14.1

JKM je konzistentní  $\Leftrightarrow \lim_{h \rightarrow 0+} \tau(x, h) = 0, \forall x$ .

┌

Důkaz

„ $\Rightarrow$ “:  $\tau(x, h) = \frac{y(x+h) - y(x)}{h} - \Phi(x, y(x), h) \rightarrow y'(x) - f(x, y(x)) \rightarrow 0$ .

„ $\Leftarrow$ “:  $\tau(x, h) = \frac{y(x+h) - y(x)}{h} - \Phi(x, y(x), h) \rightarrow y'(x) - f(x, y(x)) \rightarrow 0 \Rightarrow f(x, y(x)) = \Phi(x, y(x), 0)$ . □

└

### Věta 14.2

Nechť  $\Phi$  je lipschitzovské vzhledem k  $y$ :  $|\Phi(x, y, h) - \Phi(x, \tilde{y}, h)| \leq L|y - \tilde{y}|, \forall x, y, \tilde{y}, h$ . Nechť  $|\tau(x, h)| \leq Kh^p$  (řád  $p$ ). Pak

$$\max_{a \leq x_n \leq b} |y(x_n) - y_n| \leq Kh^p \frac{e^{L(b-a)-1}}{L}.$$



┌  
Důkaz

$$\begin{aligned}
 y_{n+1} &= y_n + h\Phi(x_n, y_n, h), \\
 y(x_{n+1}) &= y(x_n) + h\Phi(x_n, y(x_n), h) + h\tau(x_n, h), \\
 e_{n+1} &= e_n + h(\Phi(x_n, y(x_n), h) - \Phi(x_n, y_n, h)) + h\tau(x_n, h) \implies |e_{n+1}| \leq (1 + hL)|e_n| + Kh^{p+1}, \\
 |e_{n+1}| &\leq \dots \leq (1 + hL)^{n+1}|e_0| + Kh^{p+1} \frac{(1 + hL)^n}{(1 + hL) - 1} \leq 0 + Kh^p \frac{e^{hL(n+1)} - 1}{L} \leq Kh^p \frac{e^{K(b-a)}}{L}.
 \end{aligned}$$

└

□

### Definice 14.9

JKM je konvergentní, pokud  $\max_{a \leq x_n \leq b} |y(x_n) - y_n| \rightarrow 0$  pro  $h \rightarrow 0$ .

### Věta 14.3

$\Phi$  lipschitzovská vzhledem k  $y$ . Pak JKM konzistentní  $\Leftrightarrow$  JKM konvergentní.

┌  
Důkaz

„ $\implies$ “ jako v minulé větě.

„ $\Leftarrow$ “ jinak bychom řešili jinou rovnici.

└

□

### Definice 14.10

JKM je  $A$ -stabilní, pokud  $|A| := |1 + h \frac{\partial f}{\partial y}| < 1$  pro problémy s  $\frac{\partial f}{\partial y} < 0$ .

Tzv. podmíněná stabilita.

┌  
Například

Implicitní Euler je nepodmíněně  $A$ -stabilní.

└

## 14.1 Konstrukce metod (JKM, vyšší řády)