

Jan Brandts; Michal Křížek

Padesát let metody sdružených gradientů aneb Zvládnou počítače soustavy milionů rovnic o milionech neznámých?

*Pokroky matematiky, fyziky a astronomie*, Vol. 47 (2002), No. 2, 103--113

Persistent URL: <http://dml.cz/dmlcz/141120>

## Terms of use:

© Jednota českých matematiků a fyziků, 2002

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

# Padesát let metody sdružených gradientů

aneb

Zvládnou počítače soustavy milionů rovnic o milionech neznámých?

Jan Brandts, Amsterdam, a Michal Křížek, Praha

## 1. Úvod

V roce 1952 M. R. Hestenes a E. Stiefel publikovali fundamentální a zcela zásadní článek [12], v němž představili novou metodu pro řešení soustavy lineárních algebraických rovnic

$$(1) \quad Ax = b,$$

kde  $A$  je daná reálná symetrická a pozitivně definitní<sup>1)</sup> matice typu  $n \times n$ ,  $x \in \mathbb{R}^n$  je vektor neznámých a  $b \in \mathbb{R}^n$  je daná pravá strana. Metodu nazvali *the conjugate gradient method*, což se do češtiny překládá jako metoda konjugovaných nebo sdružených gradientů. Tato metoda dlouho zůstávala nepovšimnuta, protože pro „plné“ matice vyžaduje provedení  $2n^3 + \mathcal{O}(n^2)$  aritmetických operací (tj. sčítání a násobení), zatímco známá Gaussova eliminace pro symetrickou matici vyžaduje jen  $\frac{1}{3}n^3 + \mathcal{O}(n^2)$  operací<sup>2)</sup> pro  $n \rightarrow \infty$ . Metoda sdružených gradientů může být použita i na systémy rovnic, jejichž matice není pozitivně definitní, ale pak nemusí vždy dospět k řešení. Hestenes a Stiefel ji zprvu chápali jako přímou metodu<sup>3)</sup> pro řešení soustavy (1).

Přibližně o 15 let později se zjistilo, že metoda sdružených gradientů, kterou lze interpretovat též jako iterační metodu, může konvergovat velice rychle (viz [10], popř. [14]), pokud je matice  $A$  pozitivně definitní a má další příznivé vlastnosti. Takové

---

<sup>1)</sup> Matice  $A$  je pozitivně definitní, jestliže  $x^T Ax > 0$  pro každé  $x \neq 0$ .

<sup>2)</sup> Člen  $\mathcal{O}(n^2)$  zahrnuje ostatní sčítance, které odpovídají provedení řádově nejvýše  $n^2$  operací. Takový člen nižšího stupně dostaneme např. při zpětném chodu Gaussovy eliminace. Obecně řekneme, že  $f(n) = \mathcal{O}(n^\alpha)$  pro  $\alpha \in \mathbb{R}^1$  a  $n \rightarrow \infty$ , jestliže existují takové konstanty  $C$  a  $n_0$ , že pro všechna  $n > n_0$  platí  $|f(n)| \leq Cn^\alpha$ .

<sup>3)</sup> Přímé metody vedou k přesnému řešení soustavy (1) po konečném počtu kroků (za předpokladu, že se nedopouštíme žádných zaokrouhlovacích chyb). Naproti tomu u iteračních metod se konstruuje posloupnost  $\{x_k\}$ , která konverguje k řešení (1).

---

Dr. JAN BRANDTS (1968), Korteweg-de Vries Institute, Faculty of Science, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, Nizozemí, e-mail: brandts@math.cas.cz; doc. RNDr. MICHAL KRÍŽEK, DrSc. (1952), Matematický ústav Akademie věd ČR, Žitná 25, 115 67 Praha 1, e-mail: krizek@math.cas.cz

Předneseno na semináři Dějiny matematiky na Stavební fakultě ČVUT dne 26. února 2002.

matice vznikají při řešení řady technických úloh, např. při diskretizaci diferenciálních rovnic pomocí standardních numerických metod. To pak umožňuje na dnešních paralelních počítačích vyřešit soustavu (1) s dostatečnou přesností pro  $n \approx 1\,000\,000$  během několika minut. Pro srovnání uvedme, že pro inverzi matice soustavy (1) pro  $n = 10$  počítače v roce 1949 potřebovaly 8 hodin (viz [28, s. 462]).

V roce 1969 V. Strassen publikoval článek [25], v němž ukázal, že Gaussova eliminace není optimální přímou metodou pro řešení (1), jestliže ji posuzujeme podle počtu prováděných aritmetických operací. Navrhl přímou metodu vyžadující jen  $\mathcal{O}(n^{2,81})$  aritmetických operací. Později byl tento počet snížen dokonce na  $\mathcal{O}(n^{2,38})$  (viz např. [9], [26]). Poznamenejme však, že tyto metody se pro praktické počítání nepoužívají, protože jsou rekurzivní, složité a nestabilní. Dodnes ale není známo, jaký je minimální možný exponent u  $n$ . Zřejmě nemůže být menší než 2, protože plná obecně nesymetrická matice má  $n^2$  nenulových prvků a každý z nich se samozřejmě musí uplatnit při výpočtu řešení soustavy (1).

Je-li k dispozici dostatek času a paměti počítače, obvykle se doporučuje dát přímým metodám přednost před iteračními. Pak odpadnou problémy provázející iterační metody, např. potíže s volbou počátečního přiblížení, různých relaxačních parametrů, se zastavovacím kritériem apod. (viz [2]). Na dnešních nejvýkonnějších superpočítačích však lze přímými metodami vyřešit soustavy obsahující nejvýše sto tisíc rovnic (s „řádkou“ maticí), což je pro řešení skutečných technických trojrozměrných problémů nedostatečné. Řešit i „malou“ soustavu (1) bez znalosti teorie, např. Cramerovým pravidlem, může vést ke katastrofálním nárokům na výpočetní čas (srov. [17, s. 30]).

## 2. Hlavní myšlenka metody sdružených gradientů

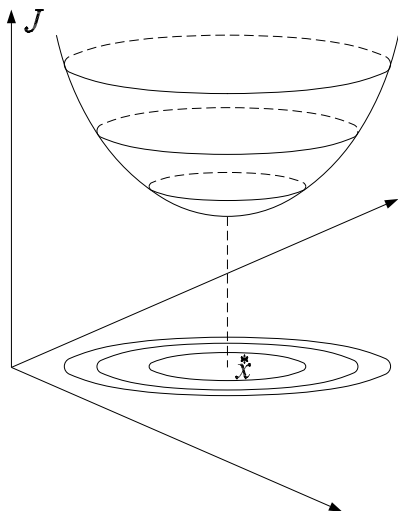
Snadno lze ukázat, že úloha (1) se symetrickou a pozitivně definitní maticí  $A$  je ekvivalentní úloze hledání minima kvadratického funkcionálu

$$(2) \quad J(x) = \frac{1}{2}x^T Ax - b^T x$$

v prostoru  $\mathbb{R}^n$ . Vrstevnice tohoto funkcionálu (srov. obr. 1) jsou hyperelipsoidy, jejichž společným středem je řešení soustavy (1) a jejichž osy jsou dány směry vlastních vektorů matice  $A$  (podrobnosti viz [15, s. 79]).

Metodu sdružených gradientů lze chápat jako iterační metodu, při níž se konstruuji tři posloupnosti vektorů — posloupnost iterací  $x_k$ , posloupnost reziduí  $r_k = b - Ax_k$  a posloupnost tzv. sdružených vektorů  $p_k$  pro  $k = 0, 1, \dots$ . Hestenes a Stiefel dokázali, že posloupnost  $x_k$  dokonverguje k přesnému řešení soustavy (1) po nejvýše  $n$  krocích algoritmu za předpokladu, že se nedopouštíme zaokrouhlovacích chyb.

Na obrázku 2 vidíme geometrickou interpretaci toho, jak jednotlivé iterace postupují pro  $n = 2$ , kdy vrstevnice  $J$  jsou elipsy. Nejprve zvolíme libovolnou počáteční aproximaci  $x_0 \in \mathbb{R}^n$  a funkcionál (2) minimalizujeme ve směru největšího spádu, tj. ve směru rezidua  $p_0 = r_0$ , které je kolmé na příslušnou vrstevnici. Tak získáme bod  $x_1$ . Dále minimalizujeme  $J$  ve směru *sdruženého vektoru*  $p_1$  (tj. směru, který spojuje bod



Obr. 1. Graf kvadratického funkcionálu  $J : \mathbb{R}^2 \rightarrow \mathbb{R}$ .

dotyku  $x_1$  a střed elipsy). Takto získáme řešení  $x_2$  po dvou krocích jednorozměrné minimalizace.

Pro obecné  $n$  je metoda sdružených gradientů pro soustavu (1) se symetrickou a pozitivně definitní maticí  $A$  definována takto:

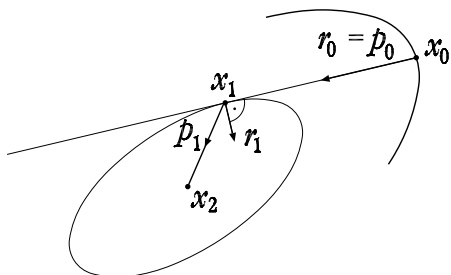
Nechť  $x_0$  je počáteční aproximace řešení soustavy (1) taková, že  $Ax_0 \neq b$ . Pak položíme

$$(3) \quad p_0 = r_0 = b - Ax_0$$

a definujeme

$$\left. \begin{aligned} (4) \quad \alpha_k &= \frac{r_k^T r_k}{p_k^T A p_k} \\ (5) \quad x_{k+1} &= x_k + \alpha_k p_k \\ (6) \quad r_{k+1} &= r_k - \alpha_k A p_k \\ (7) \quad \beta_k &= \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k} \\ (8) \quad p_{k+1} &= r_{k+1} + \beta_k p_k \end{aligned} \right\} k = 0, 1, \dots,$$

dokud  $r_k \neq 0$ .



Obr. 2. Geometrická interpretace metody sdružených gradientů.



Nespornou výhodou metody sdružených gradientů je také to, že se matice  $A$  po celý výpočet nemění, zatímco při eliminaci se matice soustavy mění a navíc se ztrácí její řídkost.

### 3. Rychlost konvergence metody sdružených gradientů

Zajímavým rysem metody sdružených gradientů je odhad její rychlosti konvergence (viz (10)). Jak známo, matice  $A$  má kladná vlastní čísla  $\lambda_n \geq \dots \geq \lambda_1$ , je-li symetrická a pozitivně definitní. Její číslo podmíněnosti se definuje vztahem

$$\kappa(A) = \max_j \lambda_j / \min_j \lambda_j = \lambda_n / \lambda_1.$$

Na základě této definice pro algoritmus metody sdružených gradientů (3)–(8) platí<sup>4)</sup>

$$(10) \quad \|x^* - x_k\|_A \leq 2 \left( \frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^k \|x^* - x_0\|_A, \quad k = 0, 1, \dots,$$

kde  $x_0$  je počáteční aproximace přesného řešení  $x^*$  soustavy (1) a

$$\|x\|_A = \sqrt{x^T A x}$$

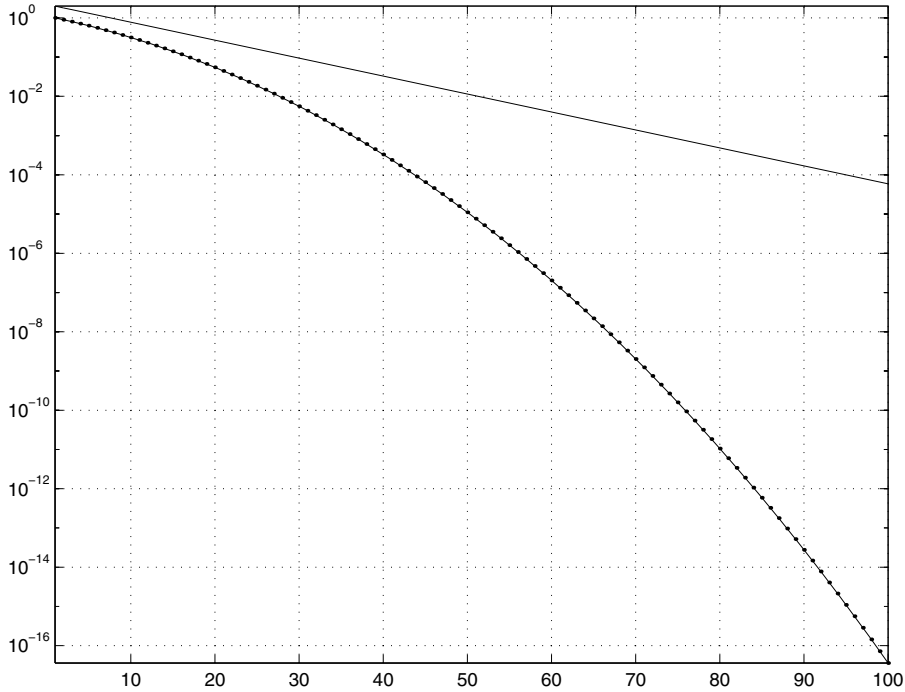
je tzv. *energetická norma*. Podrobný důkaz tohoto tvrzení je uveden např. v [16, kap. 5]. Zde je též ukázáno, že chyba monotónně klesá i ve standardní eukleidovské normě, což ovšem neplyne z (10), ale vyžaduje samostatnou analýzu.

Nerovnost (10) je ve většině případů (srov. [3, s. 15]) velmi pesimistická. Při praktických výpočtech pozorujeme, že konvergence není jen mnohem rychlejší, ale též jiného typu, než ukazuje vztah (10). Jestliže znázorníme závislost logaritmu pravé strany (10) na počtu iterací  $k$ , pak dostaneme přímku se zápornou směrnici. Na druhé straně pro logaritmus levé strany získáme křivku, jejíž směrnice s rostoucím  $k$  obecně monotónně klesá. Odtud je zřejmé, proč se takový typ konvergence nazývá *superlineární*, i když tento termín nebyl použit v článku [13] A. Jenningse, který zajímavé chování chyby matematicky analyzoval. Jak superlineární konvergence, tak i horní odhad z (10) jsou ilustrovány na obrázku 4 pro  $k = 1, \dots, 100$ .

Intuitivní výklad předchozího jevu je založen na ortogonalitě vlastních vektorů. Nechť  $v_1, \dots, v_n$  je ortonormální systém vlastních vektorů odpovídajících vlastním číslům  $\lambda_1, \dots, \lambda_n$  (tj.  $Av_i = \lambda_i v_i$ ). Uvažujme ortogonální doplněk  $V$  vektoru  $v_n$ , což je podprostor všech vektorů kolmých na  $v_n$ . Protože každé  $v \in V$  je lineární kombinací zbývajících vlastních vektorů, vidíme, že  $A$  lineárně zobrazuje vzájemně jednoznačně  $V$  na  $V$ . Odtud je patrné, že vlastní čísla každé matice  $A_r$ , která reprezentuje toto zúžené zobrazení, jsou  $\lambda_1, \dots, \lambda_{n-1}$ .

---

<sup>4)</sup> Tento odhad poprvé uvedený v [10] a [14] podstatným způsobem využívá vlastností Čebyševových polynomů. Základní myšlenka jeho důkazu je obsažena v článku MEINARDUS, G.: *Über eine Verallgemeinerung einer Ungleichung von L. V. Kantorovich*, Numer. Math. 5 (1963), 14–23.



Obr. 4. Lineární versus superlineární konvergence.

Předpokládejme na okamžik, že  $b \in V$ . Potom lze snadno ukázat, že  $x^*$  je také ve  $V$ . Použití metody sdružených gradientů s počáteční podmínkou  $x_0 = 0$  pak vede na metodu, jejíž iterace, rezidua i sdružené vektory patří do  $V$ , jak plyne z (3)–(8). Ve skutečnosti celá metoda může být interpretována tak, jako kdyby směr  $v_n$  neexistoval, přičemž číslo podmíněnosti  $\kappa(A_r) \leq \lambda_{n-1}/\lambda_1$  definuje konvergenci metody podobně jako v (10). Jestliže  $\lambda_{n-1} < \lambda_n$ , platí  $\kappa(A_r) < \kappa(A)$ , což vede ke zlepšení horního odhadu. Podobné úvahy platí, i pokud by  $b$  bylo v ortogonálním doplňku  $v_1$  místo  $v_n$ . V tomto případě „nové“ číslo podmíněnosti bude nejvýše  $\lambda_n/\lambda_2$ .

Lze vyslovit obecné pravidlo: pokud by  $b$  byl v ortogonálním doplňku vektorů  $v_1, \dots, v_q$  a  $v_p, \dots, v_n$  pro  $1 \leq q < p \leq n$ , pak horní odhad rychlosti konvergence metody sdružených gradientů bude dán číslem podmíněnosti  $\lambda_{p-1}/\lambda_{q+1}$  zúžené matice typu  $(p - q - 1) \times (p - q - 1)$ . Navíc lze ukázat, že metoda dá přesné řešení během nejvýše  $p - q$  iterací (pokud nedojde k zaokrouhlování).

Samozřejmě by se mohlo jen zcela výjimečně stát, že by taková čísla  $p$  a  $q$  existovala. Nicméně se dá ukázat (i když zde to dělat nebudeme), že iterační proces je možno interpretovat jako záměnu původní soustavy  $Ax = b$  posloupností jiných soustav lineárních algebraických rovnic, jejichž pravé strany se stávají čím dále tím více kolmé na vlastní vektory  $v_1, \dots, v_q$  a  $v_p, \dots, v_n$  a žádné jiné. Lze si to představit tak, že spektrum  $\lambda_1, \dots, \lambda_n$  je postupně „ujídáno“ z obou stran, což způsobuje neustálé zmenšování čísla podmíněnosti a odpovídající zrychlování konvergence. Více informací lze nalézt v [23] a v nedávno publikovaném článku [5].

#### 4. Metoda sdružených gradientů s předpodmíněním

Ze vztahu (10) je patrné, že rychlost konvergence metody sdružených gradientů bude velká, jestliže číslo podmíněnosti  $\kappa(A)$  bude blízké jedničce. Pak lze iterační proces zastavit mnohem dříve než po  $n$  krocích, což je velká přednost této metody. Číslo podmíněnosti lze zmenšit pomocí tzv. *předpodmiňování*, kdy místo soustavy (1) řešíme jinou soustavu

$$\tilde{A}\tilde{x} = \tilde{b},$$

kde  $\tilde{A} = P^{-1}A(P^{-1})^T$  je symetrická a pozitivně definitní matice,  $\tilde{x} = P^T x$ ,  $\tilde{b} = P^{-1}b$  a kde  $P$  je vhodně zvolená (viz např. [2], [4], [16]) regulární matice taková, že

$$\kappa(\tilde{A}) < \kappa(A).$$

V tomto případě hovoříme o *metodě sdružených gradientů s předpodmíněním*. Matice  $\tilde{A}$  se nazývá *předpodmíněná* a matice  $P$  *předpodmiňující matice*. Na obrázku 5 jsou znázorněna spektra matic  $A$  a  $\tilde{A}$  pro úlohu z [16, s. 110].

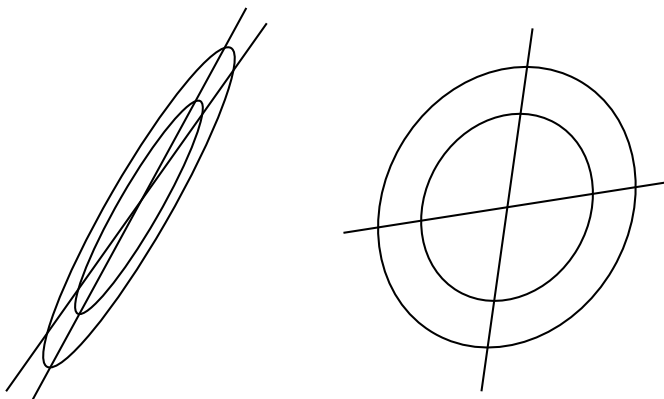


Obr. 5. Spektra matic  $A$  (vlevo) a  $\tilde{A}$  (vpravo).

Předpodmínění lze geometricky interpretovat takto: Místo funkcionálu (2) minimalizujeme na prostoru  $\mathbb{R}^n$  funkcionál

$$\tilde{J}(\tilde{x}) = \frac{1}{2}\tilde{x}^T \tilde{A}\tilde{x} - \tilde{b}^T \tilde{x},$$

jehož vrstevnice (viz obr. 6 vpravo) jsou mnohem méně protáhlé hyperelipsoidy než vrstevnice funkcionálu  $J$  (viz obr. 6 vlevo). Velikosti os hyperelipsoidů ve směru



Obr. 6. Geometrická interpretace předpodmínění.



vlastních vektorů  $v_i$  matice  $A$  jsou totiž přímo úměrné  $1/\sqrt{\lambda_i}$ ,  $i = 1, \dots, n$ . Přímkou na obrázku 6 odpovídají jednotlivým rovnicím soustav  $Ax = b$  a  $\tilde{A}\tilde{x} = \tilde{b}$  pro  $n = 2$ . Více podrobností o předpokládání lze nalézt např. v [4], [6].<sup>5)</sup>

Diskretizací jedno-, dvoj- a trojrozměrné eliptické okrajové úlohy (např. Poissonovy rovnice) metodou konečných prvků vzniká soustava (1) s pásovou maticí. Typická šíře pásu<sup>6)</sup> matice  $A$  pro  $d = 1, 2, 3$  je postupně  $\mathcal{O}(1)$ ,  $\mathcal{O}(n^{1/2})$ ,  $\mathcal{O}(n^{2/3})$  a při použití Gaussovy eliminace se tento pás zaplní obecně nenulovými prvky. Protože  $A$  má  $n$  řádků, je zapotřebí  $\mathcal{O}(n)$ ,  $\mathcal{O}(n^{3/2})$ ,  $\mathcal{O}(n^{5/3})$  paměťových buněk pro  $d = 1, 2, 3$ . Naproti tomu metoda sdružených gradientů nám umožňuje uchovávat v paměti počítače pouze nenulové prvky řídké matice  $A$  a informaci o jejich poloze v  $A$ . Nároky na paměť počítače pro Gaussovu eliminaci, metodu sdružených gradientů (CG) a její předpřipravenou verzi (PCG) jsou porovnány v tabulce 1.

Nároky na paměť počítače			
Metoda	$d = 1$	$d = 2$	$d = 3$
Gaussova	$\mathcal{O}(n)$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(n^{5/3})$
CG	$\mathcal{O}(n)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$
PCG	$\mathcal{O}(n)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$

TAB. 1.

Jak již bylo řečeno, při použití metody konečných prvků dostáváme řídkou pásovou maticí  $A$ , na jejíž sestavení je třeba jen  $\mathcal{O}(n)$ . Tabulka 2 (viz [4, s. 388]) uvádí srovnání nároků na počet operací pro výše diskutované metody řešení soustavy (1). Počet aritmetických operací pro Gaussovu eliminační metodu je úměrný čtverci šíře pásu násobeného počtem rovnic  $n$ , tj.  $\mathcal{O}(n)$ ,  $\mathcal{O}(n^2)$ ,  $\mathcal{O}(n^{7/3})$  pro  $d = 1, 2, 3$ . V důsledku velké rychlosti konvergence metody sdružených gradientů lze iterační proces ukončit mnohem dříve než po  $n$  krocích (když iterační chyba je zhruba rovna diskretizační chybě metody konečných prvků) a předpřipravená verze PCG konverguje ještě rychleji (podrobnosti viz [2]). Za dobré předpřipravení se považuje to, pro něž platí  $\kappa(\tilde{A}) \approx \sqrt{\kappa(A)}$ , což předpokládáme v tabulce 2. Za tohoto předpokladu jsou odhady uvedené v tabulkách 1 i 2 optimální v tom smyslu, že je obecně nelze zlepšit.

Počet aritmetických operací			
Metoda	$d = 1$	$d = 2$	$d = 3$
Gaussova	$\mathcal{O}(n)$	$\mathcal{O}(n^2)$	$\mathcal{O}(n^{7/3})$
CG	$\mathcal{O}(n^2)$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(n^{4/3})$
PCG	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(n^{5/4})$	$\mathcal{O}(n^{7/6})$

TAB. 2.

<sup>5)</sup> Za článek [6] získal v roce 2001 doc. ing. Miroslav Tůma, CSc., z Ústavu informatiky AV ČR cenu společnosti SIAM (Society for Industrial and Applied Mathematics).

<sup>6)</sup> Nechť  $A$  je nenulová matice typu  $n \times n$  a nechť  $s$  je nejmenší přirozené číslo takové, že  $a_{ij} = 0$  pro  $|i - j| \geq s$ ,  $i, j \in \{1, \dots, n\}$ . Pak číslo  $2s - 1$  nazýváme *šířkou pásu* matice  $A$ .

Opět vidíme, že použití Gaussovy eliminace není příliš efektivní, když  $d > 1$  a zejména když  $d = 3$ , což odpovídá velkým hodnotám  $n$ . Povšimněme si ještě, že s rostoucím  $d$  exponent u  $n$  pro Gaussovou eliminaci roste, zatímco pro metodu sdružených gradientů (i s předpokládáním) klesá! Exponent  $7/6$  v posledním řádku posledního sloupce tabulky 2 je hodně blízký jedničce. A právě v tom tkví obrovská přednost metody sdružených gradientů, která tak umožňuje rychle vyřešit soustavu (1) o milionech neznámých.

Jiná možnost, jak zlepšit konvergenci metody sdružených gradientů a příbuzných metod, je popsána např. v [7], [29].

## 5. Závěrečné poznámky

Po Gaussově eliminaci se pás řídké matice zaplní obecně nenulovými prvky. Při řešení skutečných technických problémů může být dimenze  $n$  tak velká, že nejsme schopni vyřešit soustavu (1) přímými metodami v důsledku omezené vnitřní paměti počítače (viz např. tabulku 1 pro  $d = 3$ ). Proto se pro řešení trojrozměrných problémů používají iterační metody, které umožňují uchovávat v paměti počítače pouze nenulové prvky matice  $A$  a informaci o jejich poloze v  $A$ .

Jestliže výchozí diferenciální rovnice má velké skoky v koeficientech (např. o několik řádů), pak je odpovídající číslo podmíněnosti  $\kappa(A)$  obrovské. Zvolíme-li pro řešení iterační metodu, pak je téměř nezbytné použít metodu sdružených gradientů (s předpokládáním) a její rozmanité modifikace (viz [2], [4]). Klasické iterační metody (Jacobiova metoda, Gaussova-Seidlova metoda, superrelaxační metoda, metoda největšího spádu, Kaczmarzova metoda ortogonálních projekcí apod.) totiž většinou globálně konvergují extrémně pomalu a navíc teoreticky potřebují obecně nekonečně mnoho iteračních kroků k dosažení přesného řešení. Lze je ale použít k dodatečnému „zhlazení numerického řešení“ získaného metodou sdružených gradientů, protože nejsou citlivé na zaokrouhlovací chyby a lokálně dobře konvergují (viz [16, s. 222]).

Metoda sdružených gradientů má další pozoruhodnou geometrickou interpretaci. Není těžké dokázat, že

$$\| \hat{x} - x_k \|_A = \min_{x \in x_0 + K_k} \| \hat{x} - x \|_A, \quad k = 1, 2, \dots,$$

kde  $K_k = \text{span}\{r_0, Ar_0, \dots, A^{k-1}r_0\}$  je tzv. *Krylovův prostor* a kde  $\text{span}$  označuje lineární obal. Lze si to představit tak, že se funkcionál  $J$  postupně minimalizuje na Krylovově prostoru  $K_k$ , jenž se s rostoucím  $k$  zvětšuje. Norma chyby  $\| \hat{x} - x_k \|_A$  tedy monotónně klesá, což neplyne z (10), neboť tento odhad jen udává horní odhad chyby.

Z definičních vztahů (3)–(8) lze snadno odvodit i jinou zajímavou vlastnost:

$$K_k = \text{span}\{p_0, p_1, \dots, p_{k-1}\} = \text{span}\{r_0, r_1, \dots, r_{k-1}\}.$$

Metoda sdružených gradientů se někdy používá i pro plné matice. Například matice vznikající z metody hraničních prvků jsou plné, ale obecně mají podstatně menší číslo podmíněnosti než matice vzniklé z metody konečných prvků.

Jistou nevýhodou metody sdružených gradientů je její citlivost na zaokrouhlovací chyby. V průběhu výpočtu na počítači totiž dochází ke ztrátě ortogonality reziduí (viz (9)) a iterační proces je občas zapotřebí „restartovat“, tj. začít s novou počáteční podmínkou. Přitom lze maximálně využít informaci získanou již v předchozím běhu (viz [22]). Zde je důležité si též uvědomit, že různé počítačové implementace stejného algoritmu mohou vést k poněkud odlišným numerickým výsledkům.

Metodu sdružených gradientů je možno zobecňovat mnoha směry. Například vztahy (3)–(8) lze uvažovat i pro hermitovské matice, pokud všechny transponované vektory zaměníme za hermitovsky sdružené. Pro obecnou komplexní matici  $A$  je v [16, s. 100] uvedena metoda bikonjugovaných gradientů (s předpokmáním), která za určitých předpokladů rovněž konverguje v konečném počtu kroků. Otevřeným problémem ale zůstává rychlost konvergence této metody. Rovněž se doporučuje používat blokovou verzi metody bikonjugovaných gradientů a dalších příbuzných metod, což často zlepšuje numerickou stabilitu (viz [8], [21]).

V [15, s. 211] uvádíme tvar zobecněné metody sdružených gradientů pro hledání minima hladkého ryze konvexního, koercivního a nekvadratického funkcionálu. Pokud by byl funkcionál nehladký třeba jen v jednom bodě, metoda sdružených gradientů vůbec nemusí konvergovat (srov. [30]).

Na závěr poznamenejme, že existují ještě další iterační metody používající Krylovovy prostory. Mezi nejdůležitější patří GMRES a BiCGStab pro nesymetrické soustavy a Lanczosova metoda pro indefinitní symetrické soustavy (viz [20], [27], [19]). Ve skutečnosti lze metodu sdružených gradientů chápat jako speciální případ posledně jmenované metody. Mírnou obměnou Lanczosova algoritmu [18] pro výpočet vlastních čísel symetrické matice dostaneme metodu sdružených gradientů (viz [11, s. 345]). Modifikace podobných algoritmů pro výpočet vlastních čísel vedla kdysi také k Arnoldiho metodě pro řešení (1) pro nesymetrické matice [1]. Efektivní restartovací strategie této metody se popisuje v [24].

**Poděkování.** Práce vznikla za podpory grantu A 1019201 GA AV ČR. Autoři děkují J. CHLEBOUNOVI a J. ZÍTKOVI za cenné připomínky.

#### L i t e r a t u r a

- [1] ARNOLDI, W.: *The principle of minimized iterations in the solution of the matrix eigenvalue problem*. Quart. Appl. Math. 9 (1951), 17–29.
- [2] AXELSSON, O.: *Iterative solution methods*. Cambridge Univ. Press, Cambridge 1994.
- [3] AXELSSON, O.: *Optimal preconditioners based on rate of convergence estimates for the conjugate gradient method*. LN of IMAMM 99 (eds. MÍKA, S., BRANDNER, M.), Univ. of West Bohemia, Pilsen 1999, 5–56.
- [4] AXELSSON, O., BARKER, V. A.: *Finite element solution of boundary value problems. Theory and computation*. Academic Press, New York 1984.
- [5] BECKERMANN, B., KUIJLAARS, A. B. J.: *Superlinear convergence of conjugate gradients*. SIAM J. Numer. Anal. 39 (2001), 300–329.
- [6] BENZI, M., TŮMA, M.: *A sparse approximate inverse preconditioner for nonsymmetric linear systems*. SIAM J. Sci. Comput. 19 (1998), 968–994.

- [7] BRANDTS, J.: *Explanation of a phenomenon witnessed in pre-processed GMRES*. ENUMATH 99 — Proc. of the 3rd European Conf. on Numer. Math. and Advanced Applications, Jyväskylä, Finland 1999, ed. by NEITTAANMÄKI, P. et al., World Scientific, Singapore 2000, 440–447.
- [8] BRANDTS, J.: *Riccati algorithms for eigenvalues and invariant subspaces of large and sparse matrices*. Accepted by Linear Algebra Appl. in 2001.
- [9] COPPERSMITH, D., WINOGRAD, S.: *Matrix multiplication via arithmetic progression*. J. Symbolic Comput. 9 (1990), 251–280.
- [10] DANIEL, J. W.: *The conjugate gradient method for linear and nonlinear operator equations*. SIAM J. Numer. Anal. 4 (1967), 10–26.
- [11] GOLUB, G. H., VAN LOAN, CH. F.: *Matrix computation*. The John Hopkins Univ. Press, Baltimore 1984.
- [12] HESTENES, M. R., STIEFEL, E.: *Methods of conjugate gradients for solving linear systems*. J. Res. Nat. Bur. Standards 49 (1952), 409–436.
- [13] JENNINGS, A.: *Influence of the eigenvalue spectrum on the convergence rate of the conjugate gradient method*. J. Inst. Maths Applics 20 (1977), 61–72.
- [14] KANIEL, S.: *Estimates for some computational techniques in linear algebra*. Math. Comp. 20 (1966), 369–378.
- [15] KRÍŽEK, M., NEITTAANMÄKI, P.: *Finite Element Approximation of Variational Problems and Applications*. Longman Scientific & Technical, Harlow; John Wiley & Sons, New York 1990.
- [16] KRÍŽEK, M., NEITTAANMÄKI, P.: *Mathematical and Numerical Modelling in Electrical Engineering: Theory and Applications*. Kluwer Academic Publishers, Dordrecht 1996.
- [17] KRÍŽEK, M., SEGETH, K.: *Numerické modelování problémů elektrotechniky*. Karolinum, Praha 2001.
- [18] LANCZOS, C.: *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*. J. Res. Nat. Bur. Standards 45 (1950), 255–282.
- [19] LANCZOS, C.: *Solution of systems of linear equations by minimized iterations*. J. Res. Nat. Bur. Standards 49 (1952), 33–53.
- [20] SAAD, Y., SCHULTZ, M. H.: *GMRES: a generalized minimum residual algorithm for solving nonsymmetric linear systems*. SIAM J. Sci. Stat. Comput. 7 (1986), 856–869.
- [21] SIMONCINI, V.: *A stabilized QMR version of block BiCG*. SIAM J. Matrix Anal. Appl. 18 (1997), 419–434.
- [22] SIMONCINI, V.: *On the convergence of restarted Krylov subspace methods*. SIAM J. Matrix Anal. Appl. 22 (2000), 430–452.
- [23] VAN DER SLUIS, A., VAN DER VORST, H.: *The rate of convergence of conjugate gradients*. Numer. Math. 48 (1986), 543–560.
- [24] SORENSEN, D. C.: *Implicit application of polynomial filters in a k-step Arnoldi method*. SIAM J. Matrix Anal. Appl. 13 (1992), 357–385.
- [25] STRASSEN, V.: *Gaussian elimination is not optimal*. Numer. Math. 13 (1969), 354–356.
- [26] STRASSEN, V.: *Algebraic complexity theory*. Handbook of Theoretical Computer Science, Vol. A (ed. VAN LEEUWEN, J.), Elsevier, Amsterdam 1990, 634–672.
- [27] VAN DER VORST, H.: *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*. SIAM J. Sci. Statist. Comput. 13 (1992), 631–644.
- [28] VERZUH, F. M.: *The solution of simultaneous linear equations with the aid of the 602 calculating punch*. Math. Comp. (Math. Tables and other Aids to Computation) 3 (1949), 453–462.
- [29] ZÍTKO, J.: *Combining the preconditioned conjugate gradient method and a matrix iterative method*. Appl. Math. 41 (1996), 19–39.
- [30] ZOWE, J.: *Nondifferentiable optimization — a motivation and short introduction in the subgradient and the bundle concept*. ASI Proc. Comp. Math. Progr. 1985, 323–356.