Předmluva

Textem *Analýza metod pro maticové výpočty* s podtitulem *Základní metody* předkládáme laskavému čtenáři první výsledek naší snahy o vytvoření učebních textů pokrývajících oblast výpočtů s konečnými maticemi (či, jak bývá často ne zcela výstižně uváděno, oblast numerické lineární algebry). Chceme, aby obsah textu i způsob výkladu odpovídal vývoji oboru v posledních desetiletích. Ač je výběr materiálu a jeho zpracování vždy ovlivněn subjektivním pohledem, znalostmi (a neznalostmi) autorů, pokusili jsme se nevynechat žádnou ze součástí, která je obecně považována na úrovni učebního textu za podstatnou. Naším cílem přitom není popis jednotlivých algoritmů a metod. Především chceme ukázat jejich vnitřní souvislosti a vztahy s jinými oblastmi matematiky.

Maticové výpočty jsou součástí širšího oboru (či oborů) nazývaných numerická matematika, numerická analýza či výpočetní matematika, a spolu s nimi patří do matematiky aplikované. Je vhodné připomenout, že uvedené pojmy se mohou chápat v širším či užším významu. Nám je bližší jejich širší pojetí užívané řadou dnes již klasických autorů z poloviny 20. století. Vznik a rozšíření termínu numerická analýza je spojen s existencí Ústavu numerické analýzy při UCLA v letech 1947–1954, a jeho pojetí je věrně popsáno v historické monografii [62]. Aniž bychom se pokoušeli o důkladnější studii, přece jen považujeme za potřebné vymezit blíže místo maticových výpočtů či numerické lineární algebry v matematice, jak je chápáno v našich učebních textech. Použijeme přitom zejména esejů Baxtera a Iserlese [11] a Trefethena [130], [128], [51, sekce IV.21], které jsou rovněž veřejně přístupné prostřednictvím www stránek uvedených autorů.

Ve své slavné eseji *Obrana matematikova* [60] vyjadřuje jeden z nejslavnějších matematiků 20. století G. H. Hardy myšlenku, že podstatou a ospravedlněním matematiky je *krása a závažnost matematikem vytvořených myšlenkových vzorců*. Krásu přitom spojuje s nalézáním průzračné struktury a s její matematickou přesností. Aplikovanou matematiku či dokonce možnost jakékoliv užitečnosti „skutečné matematiky“ přitom zcela vylučuje (viz rovněž [129]). V polemice s ním Baxter a Iserles ukazují, že v numerické analýze se krása, průzračná struktura a matematická přesnost neztrácejí. V našich učebních textech se budeme laskavého čtenáře pokoušet přesvědčit, že se neztrácejí ani v maticových výpočtech a jejich analýze. Hardy by s velkou pravěpodobností byl překvapen, s jakou krásou a elegancí bylo například popsáno chování silně nelineárních metod Krylovových podporostorů v aritmetice s konečnou přesností jako *matematický problém aproximace speciální třídy Riemann-Stieltjesových distribučních funkcí*; viz [53] a přehledový článek [79]. Teorie krylovovských metod přitom velmi úzce souvisí s formulací a vyřešením problému momentů, se souvislostmi sahajícími hluboko do různých oblastí matematiky, viz [118, 140, 125].

S Hardyho pojetím souvisí i jeho teze o nutnosti popření pokory jakožto předpokladu dobré práce [60, str. 61–62] a nutnosti přehánění významu vlastního oboru. Domníváme se, že zde jde buď o nepochopení pojmu pokora či o omyl. Je-li matematikova práce spojena s uvědoměním si odpovědnosti, není bez pokory dost dobře možná, nehledě na pokoru vůči kráse objeveného.

Dle Baxtera a Iserlese je předmětem výpočetní matematiky vývoj a analýza metod a algoritmů pro vyřešení problémů prostřednictvím počítače tak, aby

> výpočet vedl spolehlivě, robustně a za přijatelnou cenu (v přijatelném čase) k přijatelné aproximaci řešení.

Pokusíme se vysvětlit jednotlivá klíčová slova.

- Požadavek *spolehlivosti* vyžaduje dostatečnou znalost metody a její algoritmické implementace, která je nutná k zaručení její správné funkce pro daná data.
- *Robustnost* znamená, že vlastnosti výpočtu by se měly zachovat při změnách dat uvnitř rozumného okolí nebo při omezení na data určená určitou třídou úloh.
- *Cena* omezuje použitelnost metod příliš drahých. Jde o problém matematický, nikoliv problém technický, který by byl řešitelný pouhým rozvojem výpočetní technologie. S výkonnějšími počítači budeme chtít řešit složitější úlohy, a problém omezení cenou či časem výpočtu zůstane.
- Konečně, nejde nám o dosažení přesného řešení, ale o jeho *přijatelnou aproximaci*.

V naprosté většině matematických úloh formulovaných na základě abstraktního popisu aplikačních problémů žádnou cestu k dosažení přesného řešení nemáme. Jak uvidíme v kapitole 2, z matematického hlediska neřešitelným problémem je například určení vlastních čísel matic (což odpovídá klasickému problému určení kořenů polynomu obecného stupně). Důležité však je, že i kdyby byla cesta k nalezení přesného řešení známa, bývá výhodné hledat pouze jeho dostatečně přesnou aproximaci. Vzhledem k tomu, že výpočet používá data, která jsou zpravidla zatížena chybami, nemá prostě smysl *řešit přesně nepřesnou úlohu*. Můžeme jít dokonce ještě dále. V inverzních úlohách vznikajících diskretizací tzv. ill-posed problémů existují často chyby v datech, které mohou být vzhledem k hodnotám dat nepatrné (mohou být například způsobeny šumy různého původu při provádění fyzikálních měření). Hledání přesných řešení zmíněných inverzních úloh nemá naprosto žádný smysl, neboť vzhledem k vlastnostem *problému* může dojít *nezávisle na použité metodě* (i za předpokladu přesného provádění všech výpočtů) k takovému zesílení náhodných chyb v *původních datech*, že přesné řešení o dané praktické úloze nic nevypovídá. Příkladem je rekonstrukce obrazové informace v počítačových tomografech či zpracování signálů v řadě inženýrských aplikací.

Zbývá nám všimnout si významu slov *maticové* či *lineární* v názvu oboru maticové výpočty či numerická lineární algebra. Jak uvidíme, lineární jsou prostředky, pomocí kterých jsou vlastní problémy formulovány, nikoliv metody pro jejich řešení. Úlohy lze s výhodou popsat prostřednictvím matic a metody výpočtu pomocí manipulací s maticemi. Metody pro řešení lineárních problémů však mohou být velmi silně nelineární. Metody Krylovových podprostorů jsou příkladem nelineárních metod a analýza jejich chování představuje velmi obtížný problém.

Učební texty týkající se numerické analýzy či numerické lineární algebry často začínají popisem prováděných numerických výpočtů na počítači, popisem aritmetiky s pohyblivou řádovou čárkou a diskusí numerické stability a šíření chyb v numerických výpočtech. Otázku chyb považujeme za naprosto zásadní, neboť

> uvádění výsledků numerického výpočtu bez současného uvádění odhadu jejich chyby (popřípadě bez uvádění relevantní informace odhad nahrazující) nemá smysl.

Přesto se v prvním dílu budeme snažit numerickou stabilitu a chyby výpočtu popisovat sice slovně přesně, ale bez podrobné matematické analýzy její kvantifikace. Ve shodě s [130] považujeme za důležitější při výkladu začít matematickým popisem metod a algoritmů, přičemž jejich numerické vlastnosti jsou vysvětlovány jen natolik, aby bylo možné porozumět zdrojům možných numerických nestabilit. Numerické stabilitě a teorii citlivosti bychom se rádi věnovali v zamýšleném druhém díle.

Základní otázkou, kterou by si měl každý autor položit předtím, než začne sdělovat své výdobytky okolí, je, *proč* se danou věcí vůbec zabývá. Každý text by měl mít svoji vnitřní logiku. Pokusíme se k dané věci v našem kontextu uvést několik poznámek.

Maticovými výpočty se zabýváme proto, protože nám připadají matematicky zajímavé. Navíc mohou být případné výsledky poměrně bezprostředně užitečné i mimo vlastní oblast matematiky, což působí také jako zdroj inspirace. Učební text vznikl z vnitřní potřeby napsat si pořádněji to, co chceme učit. Jeho vnitřní logika je poněkud jiná než u nám známých českých textů dotýkajících se matic, lineární algebry či numerické matematiky. Svojí filosofií je blízký důrazu na souvislosti, prezentovanému v pohledu na pěstování lineární algebry autorů Motla a Zahradníka [81]. Svým obsahem a zaměřením je však od [81] poněkud odlišný.

Necháme-li stranou shrnutí základních pojmů týkajících se konečných matic, je skutečným počátkem výkladu Schurova věta. Ilustruje základní matematický přístup – namísto snahy řešit problém hrubou silou je lépe nalézt takový popis problému, ze kterého je řešení zřejmé. Chceme-li určit vlastní čísla matice, musíme ji transformovat na tvar, ze kterého lze vlastní čísla přečíst. Hledaná transformace musí tedy převést obecnou matici nejlépe na matici horní (dolní) trojúhelníkovou, přičemž musí zachovávat vlastní čísla (musí jít o podobnostní transformaci). Ne každá podobnostní transformace je stejně vhodná. Aby nedocházelo ke zvětšování možných chyb v datech na úkor užitečné informace, musíme se omezit na unitární podobnostní transformace. Schurova věta ukazuje, že daný přístup je možný. Každou čtvercovou matici lze unitární podobnostní transformací převést na matici horní trojúhelníkovou s vlastními čísly na diagonále v libovolném předem zvoleném pořadí. Krása a elegance Schurovy věty je v tom, že silnější výsledek ve smyslu vynulování více prvků matice obecně dokázat nelze, a slabší výsledek, tj. opuštění požadavku unitární transformace, zaručující zachování normy chyby dat, nemá z důvodu možné numerické nestability smysl hledat.

Schurova věta je východiskem mnoha cest. Pro teorii matic podtrhuje význam normality, v popředí s existencí ortonormální báze celého prostoru složené z vlastních vektorů. Přes Schurovu větu můžeme vyjít k teorii citlivosti vlastních čísel matic, což máme v plánu v druhém díle. Filosoficky nám Schurova věta ukazuje základní obtíž, které se při maticových výpočtech často nemůžeme vyhnout – cílem výpočtu je aproximovat přesný výsledek, kterého se obecně žádným (konečným) výpočtem dosáhnout nedá. Na rozdíl od matematické abstrakce je každý skutečný výpočet konečný a nemožnost dosažení přesného výsledku je realitou, se kterou se musíme smířit a naučit se s ní pracovat. Příkladem je již dříve zmíněná nemožnost výpočtu vlastních čísel obecných matic netriviálního rozměru vyplývající z Abel-Galoisovy věty o kořenech polynomů. Výpočetně nám Schurova věta ukazuje zásadní důležitost ortogonality a ortogonálních transformací.

Přirozeně se tím dostaneme ke 3. kapitole, věnované Givensovým rotacím (které by se měly vlastně nazývat Jacobiho rotace, viz [125]), Householderovým reflexím, Gram-Schmidtovu ortogonalizačnímu procesu a použití toho všeho k výpočtu QR rozkladu matice. Uvidíme, že při výpočtu na počítači není ortogonalita jako ortogonalita. Zatímco Givensovy rotace a Householderovy reflexe ortogonalitu zachovávají na úrovni úměrné zaokrouhlovací jednotce počítače, Gram-Schmidtův ortogonalizační proces toho není (bez reortogonalizace) schopen. Není vše ztraceno, pokud jsme ovšem ochotni reortogonalizovat a tím ve smyslu ceny připlatit, někdy velmi podstatně. Nic nás ale nezachrání, jde-li o velmi rozsáhlý problém, který je většinou řídký. Matice může obsahovat velké množství nulových prvků jako důsledek aproximace lokality daného modelovaného jevu (např. fyzikálního původu) či jako důsledek lokality vzniklé při diskretizaci nekonečně rozměrného problému. Požadavek ortogonality spolehlivě zlikviduje (až na speciální případy) během několika kroků ortogonálního rozkladu řídkost faktorů rozkladu. Výpočet se beznadějně zpomalí a konečně zastaví na nedostatku času, paměti či obojího. V takovém případě nezbývá než od požadavku ortogonality ustoupit.

Namísto QR rozkladu můžeme použít LU rozklad (v obecném případě) či Choleského rozklad (v hermitovském pozitivně definitním případě). Jak ukazuje 4. kapitola, důsledky kompromisů ovšem někdy bývají devastující. Jak uvidíme, LU rozklad není obecně numericky stabilní. Kvalifikovaný numerický matematik naštěstí ví, jak poznat, že výpočet není numericky stabilní, a alespoň umí uživatele, vydaného jinak na milost a nemilost jeho problému a v dané souvislosti nebezpečným datům, varovat. Choleského rozklad je naopak bezpodmínečně (zpětně) numericky stabilní. Ovšem jen díky tomu, že se omezuje pouze na velmi speciální data. Pod dojmem spočtení příkladů s náhodně šťastným koncem jsou občas formulována tvrzení, že Choleského rozklad spolehlivě funguje i na obecné hermitovské matice, které nejsou pozitivně definitní. Snaha sestrojit podobné kvadratury kruhu bude i přes dávno existující velmi názorné protipříklady bohužel věčná.

Ani QR ani LU rozklad nám není schopen spolehlivě vyjevit, zda je matice (numericky) singulární, ani jak je blízko (a v jakém smyslu) k množině singulárních matic. Za podobným účelem musíme použít singulární rozklad (SVD) vyložený v 5. kapitole. Cena je ovšem úměrná síle použitého nástroje, což naznačíme na řadě teoretických i praktických aplikací.

Jednou z nich jsou úlohy nejmenších čtverců, které jsou obsahem 6. kapitoly. Je-li známa hodnost matice nebo je-li matice dostatečně vzdálena od matice s neúplnou hodností, lze standardní problém nejmenších čtverců (LS) většinou řešit pomocí mnohem lacinějších QR rozkladů. V případě úplného problému nejmenších čtverců (TLS) se použití singulárního rozkladu v teorii (a částečně i v praxi) nevyhneme. V případě diskrétních ill-posed problémů musíme navíc použít některou z regularizačních metod a vyhnout se zcela snaze co nejlépe aproximovat přesné řešení.

Schurova věta je zaměřena na celé spektrum matice. Ne vždy ale chceme určit všechna vlastní čísla matice. Navíc mnohdy nemáme matici vůbec k dispozici a provádění jejího rozkladu tedy nepřichází v úvahu. Namísto matice můžeme mít k dispozici pouze operaci matice × vektor (neboli působení operátoru na prvek vektorového prostoru). Dominantní část spektra matice (schválně neříkáme, která část to je – jde samo o sobě o velmi komplikovanou otázku) lze aproximovat s použitím Arnoldiho metody (v případě obecné matice) a Lanczosovy metody (v případě hermitovské matice), jak je vyloženo v 7. kapitole. Arnoldiho metoda je pozoruhodná tím, že je nesnadné určit, co jsme vlastně spočítali (nevěřte žádnému počítači, že jde jisto jistě o vlastní čísla či čísla zaručeně jim blízká). Lanczosova metoda používá k vynucení ortogonality generovaných vektorů tříčlennou rekurenci, jejíž původ sahá k mnoha vynikajícím matematikům mnoha minulých století (viz [125]). Pro nás je v prvním díle učebního textu důležitější, že tříčlenná rekurence vede v praxi často k velmi rychlé ztrátě nejen ortogonality, ale i (numerické) lineární nezávislosti generovaných vektorů. Vypočtené Jacobiho matice, s jejichž pomocí máme určit aproximace hledaných dominantních vlastních čísel původní matice, nemusí mít po pár krocích výpočtu (s použitím Lanczosova algoritmu) s odpovídajícími Jacobiho maticemi získanými hypotetickým přesným výpočtem zdánlivě téměř nic společného. Zdánlivě nemožné věci se však dějí. Pomocí velmi nepřesně vypočtených Jacobiho matic lze určit hledaná dominantní vlastní čísla s chybou úměrnou strojové přesnosti počítače, a přesnost výsledku lze zaručit z vypočtených hodnot! Neméně pozoruhodné je matematické pozadí Lanczosovy metody, související s řetězovými zlomky, Gauss-Christoffelovou kvadraturou, problémem momentů, ortonormálními polynomy, Padého aproximací atd., viz [61, 79].

Použijeme-li Lanczosův algoritmus na řešení soustavy lineárních algebraických rovnic s hermitovskou pozitivně definitní maticí, dostaneme metodu sdružených gradientů. Podobně jako u Lanczosovy metody pro výpočet vlastních čísel, ani teoretickou bohatost metody sdružených gradientů nelze v rozsahu našeho textu v nejmenším postihnout, viz např. [61, 79, 140] a podstatnou část knihy [76]. V 8. kapitole se omezíme jen na základní odvození metody sdružených gradientů vycházející z myšlenky minimalizace kvadratického funkcionálu. Přesto je naší ambicí ukázat její naprostou přirozenost vycházející ze základních principů. S metodou sdružených gradientů není spojena žádná agonizující bolest, jak se to některé veřejně šířené zdroje snaží namluvit (viz např. [111]) a jiné publikace papouškovat. Agonizující bolest vzniká zkreslenými způsoby jejího výkladu bez pochopení její podstaty a souvislostí. Aritmetika s konečnou přesností opět znamená ztrátu ortogonality a numerické lineární nezávislosti vypočtených směrových vektorů a reziduí. Podobně jako u Lanczosova algoritmu lze ztrátu ortogonality a její důsledky analyzovat, což například umožňuje vysvětlit spolehlivost případně nespolehlivost zastavovacích kritérií založených na odhadu energické normy chyby aproximace řešení diskretizované úlohy, která je rovna A-normě chyby aproximace řešení lineárního algebraického problému.

Není-li matice hermitovská pozitivně definitní, můžeme buď formulovat matematicky ekvivalentní úlohu s hermitovskou pozitivně definitní maticí a odvodit s použitím metody sdružených gradientů metody pro řešení původní úlohy, nebo můžeme zobecnit některé z principů, na kterých je metoda sdružených gradientů založena. Hledané aproximace řešení jsou prvky Krylovových podprostorů posunutých o počáteční aproximaci. Zachováme-li (myšleno ideálně, tj. v přesné aritmetice) ortogonalitu generujících vektorů, můžeme formulovat metody minimalizující na Krylovových podprostorech jistou normu chyby. Jak však bylo ukázáno ve slavné Faber-Manteuffelově větě, zachování ortogonality generujících vektorů (myšleno v přesné aritmetice) vyžaduje použití dlouhých rekurencí, což podstatně omezuje počet efektivně použitelných iterací při praktických výpočtech velmi rozsáhlých úloh. Volíme-li krátké rekurence například zobecněním Lanczosova algoritmu na nehermitovský případ a metody sdružených gradientů na metodu bikonjugovaných gradientů, nevyhnutelně obětujeme optimalitu. Navíc může dojít k selhání, neboť pro některé iterace nemusí hledaná aproximace řešení zvolené metody s krátkou rekurencí vůbec existovat. V 9. kapitole naznačíme uvedené otázky a popíšeme základní metody Krylovových podprostorů. Analýza principů krylovovských metod je předmětem monografie [76], která je uvažována rovněž jako učebnice pro pokročilé kurzy zaměřené na související oblast výpočetní matematiky.

Přes svoji stručnost a pouze úvodní charakter chce náš výklad v 9. kapitole vysvětlit zásadní rozdíl mezi klasickými iteračními metodami založenými na štěpení matice soustavy (Jacobi, Gauss-Seidel, SOR, atd.) případně na konvexní kombinaci jednotlivých aproximací (semiiterační metody, Čebyševova metoda) a metodami Krylovových podprostorů. Základem krylovovských metod je projekční proces na vnořených Krylovových podprostorech rostoucí dimenze, který lze zapsat jako (zobecněný) Vorobyevův problém momentů [140, 121, 126]. Na rozdíl od klasických iteračních metod jsou krylovovské metody silně nelineární, jejich konvergence má globální charakter a nemá smysl (s výjimkou velmi speciálních případů) popisovat ji asymptotickými nástroji. Zkoumání jejich konvergence nutně vede na silně nelineární problém na lineárních prostorech malé dimenze, který není možné popsat limitním přechodem.

V textu obecně uvažujeme (s několika výjimkami) vektorové prostory nad tělesem komplexních čísel, přičemž vysvětlujeme některé odlišnosti a obtíže, které mohou výpočty s komplexními maticemi přinést (například ve 3. kapitole věnované ortogonálním transformacím a QR rozkladům). Velká většina praktických úloh vede na reálné matice a reálné vektorové prostory. Při nahrazení hermitovské transpozice standardní transpozicí zůstanou všechny uvedené výsledky zachovány. V našem textu se soustředíme na matematický popis metod a případně na jejich základní algoritmické implementace a nezabýváme se více či méně technickými otázkami používání komplexní aritmetiky při provádění výpočtů na počítači.

Text předpokládá základní znalost matematické analýzy, lineární algebry a základů numerické matematiky. Výhodou, ne však nezbytností, je základní znalost funkcionální analýzy. Většina kapitol odděluje základní výklad, který svým obsahem zhruba odpovídá jednosemestrálnímu kurzu v rozsahu 2 hodiny přednášek a 2 hodiny cvičení týdně, od rozšířeného výkladu. Uvedená rozšíření mohou být využita různým způsobem při variabilním obsahu kurzů většího rozsahu. Logika textu předpokládá, že základní výklad bude veden v uvedeném řazení jednotlivých kapitol. Každá kapitola obsahuje cvičení teoretického i výpočetního charakteru (cvičení vyšší obtížnosti jsou označena *). Při výpočtech předpokládáme práci s Matlabem. Text je doplněn poměrně rozsáhlými odkazy na další doporučenou literaturu.

Předložený text vznikl na základě přednášek konaných v průběhu řady let na několika školách v České republice i v zahraničí. Je kolektivním dílem s výhodami i potížemi z toho plynoucími. Autoři děkují všem, kteří jim pomáhali výhody posílit a nedostatky odstraňovat. Jsme velmi zavázáni M. Rozložníkovi a M. Tůmovi za cenné rady a připomínky, a také všem studentům, především T. Gergelitsovi, J. Kohutkovi, M. Kubínové, J. Papežovi, J. Soukupovi a K. Tothové, kteří trpělivě text či jeho části pročítali. V neposlední řadě bychom rádi poděkovali H. Bílkové za pomoc při tvorbě obrázků a při úpravě textu. Budeme velmi vděčni všem případným učitelům, kteří by text či jeho části chtěli vyzkoušet při výuce, za jejich kritiku jak obsahu, tak formy. Laskavému čtenáři předem děkujeme za zpětnou vazbu, která by pomohla v nikdy nekončícím procesu možného zlepšování, oprav a doplnění pokračovat.

První vydání této učebnice vyšlo již v roce 2011. V tomto druhém přepracovaném vydání došlo kromě drobných úprav k rozšíření textu o mocninnou metodu a poznámku o QR algoritmu. Naši studenti T. Gergelits, M. Kubínová a J. Papež nám pomohli zjednodušit a zefektivnit řešení některých úloh cvičení, za což jim velmi děkujeme.

V Praze, 4. 10. 2022 Autoři