



Tenzory a kanonické tenzorové rozklady: Tuckerův rozklad

Bakalářská práce

Studijní program: B1101 – Matematika
Studijní obory: 7504R015 – Matematika se zaměřením na vzdělávání
7507R036 – Anglický jazyk se zaměřením na vzdělávání

Autor práce: **Jana Žáková**
Vedoucí práce: Martin Plešinger





TECHNICAL UNIVERSITY OF LIBEREC
Faculty of Science, Humanities
and Education



Tensors and canonic tensor decompositions: Tucker decomposition

Bachelor thesis

Study programme: B1101 – Mathematics
Study branches: 7504R015 – Mathematics for Education
7507R036 – English for Education

Author: **Jana Žáková**
Supervisor: Martin Plešinger



ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Jana Žáková**
Osobní číslo: **P12000550**
Studijní program: **B1101 Matematika**
Studijní obory: **Matematika se zaměřením na vzdělávání**
Anglický jazyk se zaměřením na vzdělávání
Název tématu: **Tenzory a kanonické tenzorové rozklady: Tuckerův rozklad**
Zadávající katedra: **Katedra matematiky a didaktiky matematiky**

Z á s a d y p r o v y p r a c o v á n í :

Moderní numerická lineární algebra se stále více posouvá od maticových výpočtu k výpočtům tenzorovým. Tenzorové úlohy se přirozeně vyskytují např. ve výpočtech v kvantové chemii, ale i v jednodušších problémech, např. chceme-li řešit úlohu vedení tepla závislou na parametru tepelné vodivosti (nebo několika parametrech) pro široké spektrum hodnot parametru. Úlohy s tenzory se přirozeně vyskytují i v úlohách zpracování vícerozměrných dat, slouží např. k jejich kompresi (video), a/nebo k identifikování informace, která je v nějakém smyslu důležitá (data-mining).

Bakalářská práce čtenáře seznámí s popisem tenzoru tak, jak je v tenzorových výpočtech chápán, tj. jakožto vícerozměrného pole čísel. Zmíní a na jednoduchých příkladech ilustruje obtíže, které zde oproti práci s maticemi vyvstávají (zejm. problém zavedení hodnoty tenzoru). Klíčovým nástrojem maticové algebry je singulární rozklad (SVD), který lze použít právě v úlohách komprese matice nebo data-mining (analýza hlavních komponent). Práce objasní, jakým způsobem lze SVD použít pro zavedení tzv. Tuckerova rozkladu daného tenzoru.

Požadavky: Základní znalosti z lineární algebry, základní znalost anglického jazyka. Práce by měla být psána tak, aby mohla celá, nebo její části, sloužit jako materiál pro úvod studia dané problematiky. Práce by měla být psaná v LaTeXu, bude-li to v možnostech studenta.

Rozsah grafických prací:

Rozsah pracovní zprávy:

Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam odborné literatury:

T. G. Kolda, B. W. Bader: Tensor Decompositions and Applications. SIAM Review 51 (2009), issue 3, pp. 455–500.

L. Grasedyck: Hierarchical Singular Value Decomposition of Tensors. SIAM Journal on Matrix Analysis and Applications 31 (2010), issue 4, pp. 2029–2054.

W. Hackbusch, S. Kühn: A New Scheme for the Tensor Representation. Journal of Fourier Analysis and Applications 5 (2009), issue 5, pp. 706–722.

C. Tobler: Low-rank Tensor Methods for Linear Systems and Eigenvalue Problems. PhD Tesis, SAM, ETH Zurich, 2012.

Vedoucí bakalářské práce:

Ing. Martin Plešinger, Ph.D.

Katedra matematiky a didaktiky matematiky

Datum zadání bakalářské práce: **23. března 2014**

Termín odevzdání bakalářské práce: **24. dubna 2015**



doc. RNDr. Miroslav Brzezina, CSc.
děkan

L.S.



doc. RNDr. Jarošlav Mlýnek, CSc.
vedoucí katedry

dne

Prohlášení

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

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

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

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

Současně čestně prohlašuji, že tištěná verze práce se shoduje s elektronickou verzí, vloženou do IS STAG.

Datum:

Podpis:

Anotace

Soudobý vývoj v aplikované a numerické lineární algebře stále více směřuje, mimo jiné, od maticových výpočtů směrem k výpočtům tenzorovým. Tenzorové úlohy se přirozeně vyskytují např. ve výpočtech v kvantové chemii, ale i v jednodušších problémech, např. chceme-li řešit úlohu vedení tepla závislou na parametru tepelné vodivosti (nebo několika parametrech) pro široké spektrum hodnot parametru. Úlohy s tenzory se přirozeně vyskytují i v úlohách zpracování vícerozměrných dat. Používají se při jejich efektivním ukládání (např. při kompresi videa), nebo k identifikování informace, která je v nějakém smyslu důležitá (např. v úlohách data-mining), atd.

Úkolem této práce je seznámit čtenáře s popisem tenzoru tak, jak je v tenzorových výpočtech chápán, tj. jakožto vícerozměrného pole čísel (víceindexové matice, hypermatrix). Zavádí značení tenzorů a popisuje základní operace a nástroje pro práci s tenzory, tj. součiny tenzorů, rozvoj tenzoru v matici, atd. Poukazuje také na rozdíly a obtíže, které vyvstávají při tenzorových výpočtech, se kterými se v klasické lineární algebře a maticovém počtu nesetkáváme, zejména problém zavedení hodnoty tenzoru.

Klíčovým nástrojem maticové algebry je singulární rozklad (SVD), který lze použít právě v úlohách komprese dvourozměrných dat (matice), pod názvem analýza hlavních komponent (PCA) se s ním setkáváme ve statistice, třeba právě v úlohách data-mining. Důležité je s daty (maticí) takto komprimovanými pomocí singulárního rozkladu umět provádět alespoň některé základní maticové operace.

Jedním z možných zobecnění singulárního rozkladu pro tenzory vyššího řádu je tzv. Tuckerův rozklad (také nazývaný high-order SVD, HOSVD). Ten přirozeně zobecňuje některé z vlastností singulárního rozkladu, které se využívají ve výše uvedených aplikacích, pro vícerozměrná data. Zejména umožňuje daný tenzor dobře aproximovat tenzorem nižší hodnosti. Stejně jako v případě matic je i zde důležité umět provádět s tenzory ve tvaru Tuckerova rozkladu základní tenzorové operace.

Klíčová slova:

multilinerární algebra; tenzor; hodnota tenzoru; singulární rozklad (SVD); tenzorové rozklady; polyadický rozvoj (CP rozklad); Tuckerův rozklad; low-rank aritmetika; aproximace tenzoru tenzorem nižší hodnosti

Abstract

Contemporary applied and numerical linear algebra is expanding from matrix to tensor computations. Tensor problems naturally arise in, e.g., computational quantum chemistry, but also in more common areas. For example while solving the heat equation in a domain with parameter-dependent thermal conductivity for many particular values of the parameter (or several parameters) in a wide range. Tensor problems arise naturally also in multidimensional data processing. They are in the background of effective storing of big data (e.g., in video compression), they arise in problems of extracting important information hidden in data (e.g., in data-mining problems), etc.

The goal of this thesis is to introduce a tensor as a multidimensional array (hypermatrix) to the reader. In this way tensors are understood, studied, and analyzed in modern tensor computation. We summarize the standard notation, and introduce some important parts of tensor arithmetic, such as tensor products, matricization, etc. Differences between matrices and tensors are emphasized. In particular, we discuss difficulties linked with the concept of a tensor rank.

One of the most important tools in matrix algebra is the singular value decomposition (SVD). It can be used, e.g., for compression of two-dimensional data (a matrix). In statistics, the SVD is known as the principal component analysis (PCA), which can also be interpreted as a data-mining problem. In many applications, it is important to know how to do basic arithmetic with data compressed and stored by the SVD.

One of possible generalizations of the SVD to higher-order tensors, is the so-called Tucker decomposition (also known as high-order SVD, HOSVD). The Tucker decomposition naturally generalizes some of the properties of the SVD, which are used in applications mentioned above, to multidimensional data. In particular, it allows to approximate a given tensor by another one with lower rank. Similarly to the two-dimensional case, it is important to know how to do basic arithmetic with multidimensional data compressed and stored by the Tucker decomposition.

Key words:

multilinear algebra; tensor; tensor rank; singular value decomposition (SVD); tensor decompositions; CP decomposition; Tucker decomposition; low-rank arithmetic; low rank tensor approximation

Poděkování

Ráda bych poděkovala svému vedoucímu bakalářské práce, panu Martinu Plešingerovi, za odborné vedení, cenné rady a čas, který mi při tvorbě této práce věnoval.

Obsah

| | |
|---|-----------|
| Anotace | 5 |
| Abstract | 6 |
| Seznam obrázků | 10 |
| Seznam tabulek | 11 |
| Použité značení a zkratky | 12 |
| Úvod | 14 |
| 1 Singulární rozklad | 17 |
| 1.1 Základní tvar SVD | 17 |
| 1.2 Ekonomický tvar SVD | 18 |
| 1.3 Operace s maticí uloženou v ekonomickém tvaru SVD | 19 |
| 1.3.1 Sčítání dvou matic v ekonomickém tvaru | 19 |
| 1.3.2 Násobení matice v ekonomickém tvaru obecnou maticí | 21 |
| 1.4 Aproximace matice maticí nižší hodnoti | 22 |
| 2 Tenzory | 24 |
| 2.1 Značení a terminologie | 25 |
| 2.1.1 Vlákná tenzoru | 25 |
| 2.1.2 Řezy tenzoru | 26 |
| 2.1.3 Transpozice a symetrie tenzoru | 28 |
| 2.2 Základní operace s tenzory | 29 |
| 2.3 Rozvoj tenzoru v matici | 30 |
| 2.4 Hodnota tenzoru | 32 |
| 2.4.1 ℓ -rank a vektorový rank tenzoru | 32 |
| 2.4.2 Vnější součin tenzorů, polyadický rozvoj a polyadický rank tenzoru | 33 |
| 2.4.3 Norma tenzoru a aproximace tenzoru tenzorem nižšího polyadického ranku | 35 |
| 2.5 Zobecněný rozvoj tenzoru v matici, vektorizace a Kroneckerův součin matic | 36 |
| 2.6 Tenzorový součin a vztahy mezi jednotlivými součiny | 40 |

| | |
|---|-----------|
| 3 Tuckerův rozklad | 43 |
| 3.1 Odvození Tuckerova rozkladu | 43 |
| 3.2 Základní operace s tenzory v Tuckerově tvaru | 45 |
| 3.2.1 Sčítání tenzorů v Tuckerově tvaru | 46 |
| 3.2.2 Násobení tenzoru v Tuckerově tvaru maticí | 47 |
| 4 Aproximace tenzoru pomocí Tuckerova rozkladu | 50 |
| 4.1 Zobecněná Schmidtova–Eckartova–Youngova–Mirského věta | 50 |
| 4.2 Rozbor zobecněné věty na příkladech | 53 |
| 4.2.1 Nadhodnocení odhadu | 53 |
| 4.2.2 Rovnost odhadu a hodnost výsledného tenzoru | 55 |
| 4.2.3 Optimalita aproximace | 57 |
| Závěr | 59 |
| A Software | 60 |
| A.1 MATLAB | 60 |
| A.2 TENSOR TOOLBOX | 60 |
| A.3 HIERARCHICAL TUCKER TOOLBOX | 61 |
| A.4 TENSORLAB | 61 |
| A.5 N-WAY TOOLBOX & CUBATCH | 61 |
| A.6 PLS TOOLBOX | 61 |
| B Příklad tenzoru s malým Tuckerovým jádrem | 62 |
| C Aproximace tenzoru z přílohy B tenzorem nižšího ranku | 70 |
| Reference | 74 |

Seznam obrázků

| | | |
|-----|---|----|
| 2.1 | Vlákna tenzoru | 25 |
| 2.2 | Řezy tenzoru | 27 |
| 2.3 | Polyadický rozvoj tenzoru řádu 3 | 35 |
| 3.1 | Tuckerův rozklad tenzoru řádu 3 | 45 |
| B.1 | Příklad: Singulární čísla tenzoru \mathcal{A} | 63 |
| C.1 | Příklad: Aproximace tenzoru \mathcal{A} tenzorem hodnosti $(5, 6, 2)$ | 72 |
| C.2 | Příklad: Aproximace tenzoru \mathcal{A} tenzorem hodnosti $(5, 6, 1)$ | 72 |
| C.3 | Příklad: Aproximace tenzoru \mathcal{A} tenzorem hodnosti $(5, 2, 2)$ | 73 |

Seznam tabulek

| | | |
|-----|--|----|
| 1.1 | Počet prvků matice a jejího ekonomického SVD | 19 |
| 1.2 | Součet dvou matic v ekonomických SVD tvarech | 21 |
| 1.3 | Násobení matice v ekonomickém SVD tvaru obecnou maticí | 22 |
| 3.1 | Součet dvou tenzorů řádu 3 v Tuckerových tvarech | 48 |
| 3.2 | Násobení tenzoru v Tuckerově tvaru obecnou maticí | 49 |
| C.1 | Příklad: Singulární čísla tenzoru \mathcal{A} | 70 |
| C.2 | Příklad: Approximace tenzoru \mathcal{A} tenzorem nižší hodnosti | 71 |

Použité značení a zkratky

V textu značíme

| | |
|----------------------------|--|
| vektory (tenzory řádu 1) | pomocí malých písmen $u_1, u_2, u_r, v_1, v_2, v_r, x$, atd., |
| matice (tenzory řádu 2) | pomocí velkých písmen (latinských i řeckých) $A, B, C, D, E, F, U, V, \Sigma$, atd., |
| tenzory řádu $k, k \geq 3$ | pomocí velkých písmen psaných kaligraficky $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{T}, \mathcal{S}$, atd. (výjimečně Δ), |
| množiny | pomocí velkých písmen psaných Scriptem $\mathcal{D}, \mathcal{S}, \mathcal{T}, \mathcal{X}_j$, atd. |

Pomocí malých písmen (latinských i řeckých) také značíme prvky matic a tenzorů a také skaláry (tenzory řádu 0). Speciální význam pak mají písmena i, j, ℓ , jimiž zpravidla indexujeme prvky matic a tenzorů, a k, m, n, r , která používáme k označení řádu tenzoru, dimenze matice nebo tenzoru, resp. hodnoti (ranku) matice nebo tenzoru.

Matice a vektory

| Značení | Význam |
|---|---|
| $A \in \mathbb{R}^{n \times m}$ | reálná matice s rozměry n krát m , s prvky $a_{i,j}$ |
| $\text{vec}(A) \in \mathbb{R}^{nm}$ | vektORIZACE matice $A \in \mathbb{R}^{n \times m}$ |
| $A \otimes B$ | Kroneckerův součin dvou matic |
| A^T | transpozice matice A |
| $\text{rank}(A)$ | hodnota matice definovaná jako počet lineárně nezávislých řádků, resp. sloupců matice A |
| $\ x\ = (\sum_i x_i^2)^{1/2}$ | eukleidovská norma vektoru |
| $\ A\ = \max_{\ x\ =1} \ Ax\ $ | spektrální norma matice |
| $\ A\ _F = (\sum_i \sum_j a_{i,j}^2)^{1/2}$ | Frobeniova norma matice |

Tenzory

| Značení | Význam |
|---|--|
| $\mathcal{A} = (a_{i_1, i_2, i_3}) \dots \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ | tenzor třetího řádu o rozměrech n_1, n_2, n_3 |
| $\mathcal{A} = (a_{i_1, i_2, \dots, i_k}) \dots \in \mathbb{R}^{n_1 \times \dots \times n_k}$ | tenzor k -tého řádu o rozměrech n_1, \dots, n_k |
| $a_{:, i_2, i_3} \in \mathbb{R}^{n_1}$ | vláknو tenzoru třetího řádu v módu 1 |
| $A_{:, :, i_3} \in \mathbb{R}^{n_1 \times n_2}$ | řez tenzoru třetího řádu v módu (1, 2) |
| $\text{vec}(\mathcal{A}) \in \mathbb{R}^{n_1 n_2 \dots n_k}$ | vektorizace tenzoru $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}$ |
| $A^{[\ell]} \in \mathbb{R}^{n_\ell \times (n_1 \dots n_{\ell-1} n_{\ell+1} \dots n_k)}$ | rozvoj tenzoru do matice v ℓ -tém módu |
| $A^{[t_1, \dots, t_d]}$ | zobecněný rozvoj tenzoru do matice |
| $\mathcal{A} \times_\ell M$ | násobení tenzoru maticí v ℓ -tém módu; platí $(\mathcal{A} \times_\ell M)^{[\ell]} = M A^{[\ell]}$ |
| $\mathcal{A} \times_{\ell, s} \mathcal{B}$ | tenzorový součin v módech ℓ a s |
| $\mathcal{A} \times_{(\ell_1, \ell_2), (s_1, s_2)} \mathcal{B}$ | tenzorový součin ve dvojici módů |
| $\mathcal{A} \circ \mathcal{B}$ | vnější součin tenzorů |
| $\mathcal{A}^{\Pi_j}, j = 1, \dots, k!$ | (Π_j) -tá transpozice tenzoru řádu k |
| $\text{rank}_\ell(\mathcal{A})$ | ℓ -rank tenzoru; $\text{rank}_\ell(\mathcal{A}) = \text{rank}(A^{[\ell]})$ |
| $\text{rank}(\mathcal{A})$ | vektorový rank tenzoru řádu k , pro $k \geq 3$ platí $\text{rank}(\mathcal{A}) = (\text{rank}_1(\mathcal{A}), \dots, \text{rank}_k(\mathcal{A}))$ |
| $\ \mathcal{A}\ = (\sum_{i_1} \dots \sum_{i_k} a_{i_1, \dots, i_k}^2)^{1/2}$ | norma tenzoru |

Použité zkratky a akronymy

| Zkratka | Význam |
|-----------|---|
| QR | QR rozklad matice, $A = QR$ |
| SVD | singulární rozklad matice (singular value decomposition), $A = U \Sigma V^T$ |
| ek. SVD | ekonomická varianta SVD, $A = U_A \Sigma_A V_A^T$ |
| PCA | analýza hlavních komponent, varianta SVD (principal component analysis) |
| CanDecomp | polyadický rozvoj tenzoru (canonical decomposition) |
| ParaFac | ditto (parallel factors, ev. factorisation) |
| CP | ditto (CanDecomp-ParaFac) |
| HOSVD | Tuckerův rozklad tenzoru (high-order SVD) |
| MVp | součin matice s vektorem (matrix-vector product) |
| MMp | součin matice s maticí (matrix-matrix product) |
| TMp | součin tenzoru s maticí (tensor-matrix product) |

Úvod

Jedněmi ze základních a klasických objektů studia lineární algebry jsou vektory a matice, které mohou být chápány jako jednorozměrné, resp. dvourozměrné soubory dat (v tomto textu budeme pod pojmem data vždy rozumět výhradně reálná čísla, vše však lze bez problému zobecnit na čísla komplexní). Pokud jsou ale data uspořádána, nebo je potřebujeme uspořádat podle více než dvou parametrů, s maticemi už nevystačíme. Matematické objekty, pomocí kterých můžeme taková více-rozměrná data interpretovat, jsou *tenzory*. Pod pojmem tenzoru tedy budeme rozumět vícerozměrné pole čísel, počet těchto rozměrů nazýváme *řád tenzoru*, někdy se místo pojmu *řád tenzoru* používá pojem *stupeň* nebo *dimenze tenzoru*. My budeme pojmem *dimenze tenzoru* označovat dimenzi prostoru jehož je prvkem, tj. např. $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$ je tenzor řádu k , dimenze $n_1 n_2 \dots n_k$. Z tohoto úhlu pohledu klasická *lineární algebra* a *maticový počet* (*matrix calculus*), případně *maticové výpočty* (*matrix computations*) pracující se skaláry, vektory a maticemi formálně pracují s tenzory nultého, prvního, resp. druhého řádu. Rozšíříme-li lineární algebru i na práci s tenzory vyšších řádů, hovoříme také o *multilineární algebře*, *tenzorovém počtu* (*tensor calculus*) a *tenzorových výpočtech* (*tensor computations*).

Úkolem tohoto textu je mimo jiné seznámit čtenáře s terminologií a základními operacemi s tenzory vyšších řádů, přičemž vychází především z článku [28]. Jelikož počet prvků tenzoru roste exponenciálně s řádem tenzoru k (počet prvků tenzoru je zřejmě omezen čísly $\min_{\ell} n_{\ell}^k$ a $\max_{\ell} n_{\ell}^k$), nastává při práci s tenzory problém s jejich uložením a zpracováním vícerozměrných dat, které reprezentují. Obtíže vznikající při práci s takovými daty lapidárně shrnuje motto „*curse of dimensionality*“ často používané právě v kontextu práce s tenzory vysokých řádů (např. $k \sim 100$), ale i při jiných příležitostech. Tyto problémy vedly ke snaze nalézt způsob, jak data komprimovat, s cílem uložit tenzor ve tvaru, který není tak náročný na paměť počítače. Výsledkem tohoto snažení je hned několik způsobů tenzorových rozkladů, které podobně jako *singulární rozklad matic* (singular value decomposition, SVD) umožňují vyjádřit daný tenzor jako součin (případně součet součinů) několika jiných objektů, tenzorů nebo matic (výrazně) menších dimenzí.

Zajímavostí je, že potřeba nalézt takový aparát, aby se zmenšil objem dat tenzoru bez ztráty podstatných informací, se neobjevila jako první v matematice. Jak uvádí T. Kolda [28], tenzorové rozklady byly vynalezeny poprvé v psychologii (konkrétně v *psychometrice* zabývající se měřením různých psychických jevů, jako jsou např. osobnostní charakteristiky, postoje, schopnosti, atd.) a v chemii (přesněji v *chemometrii*, tj. zpracování a analýze chemických dat). První způsob rozkladu tenzoru popsal F. L. Hitchcock v pracích [22] a [23] už v roce 1927. Další různé způsoby

byly popsány ale až později v šedesátých a sedmdesátých letech dvacátého století v pracích L. R. Tuckera, viz [43], [44] a [45], J. D. Carrolla a J. J. Changa, viz [8], a R. A. Harshmana, viz [21], a to právě v literatuře zabývající psychometrikou. Následně tenzory a tenzorové rozklady začala využívat chemometrika a proto mnoho prací o tenzorech vznikalo právě v tomto odvětví. V posledních letech se tenzory a tenzorové výpočty používají i v mnoha dalších disciplínách, např. v analýze a zpracování signálů, v data-miningu, ve statistice, v neurovědách, atd., a přirozeně také v (aplikované a numerické) lineární algebře, která je svým způsobem přirozeným prostorem i pro základní výzkum v této oblasti.

Cílem tohoto textu je popsat zejména tzv. *Tuckerův rozklad tenzoru*, který je jedním z vícerozměrných zobecnění (ekonomické varianty) SVD. Proto se Tuckerův rozklad také někdy nazývá HOSVD (high-order SVD). Tomu je podřízena struktura textu a s ohledem na to je také veden výklad. Látka není vykládána encyklopedicky (např. různé typy součinů tenzorů nejsou popisovány na jednom místě), ale postupně, tehdy, kdy je daná látka potřeba. Drobnou výjimkou jsou podkapitoly 2.4.2 a 2.4.3, které rozvíjí diskuzi ohledně hodnoty tenzoru nad rámec toho, co je pro Tuckerův rozklad třeba. Stejně tak podkapitoly 2.5 a 2.6 se zabývají operacemi s tenzory, které nejsou pro výklad klasického Tuckerova rozkladu přímo potřeba. Jsou však v jistém smyslu přirozenou součástí světa tenzorů.

Text je strukturovaný následujícím způsobem. Po stručném úvodu je v kapitole 1 zaveden singulární rozklad, jeho ekonomický tvar a je ukázáno jeho využití při základních operacích s velkými maticemi malé hodnoty. Kapitola 2 je věnována tenzorům. Zavádí se zde značení a základní terminologie tenzorů a je také popsáno, jak provádět některé základní operace. V další části je definován *rozvoj tenzoru v matici*, který je užitečným nástrojem pro práci s tenzory. Velmi důležitým tématem je hodnota tenzoru, kde ukazujeme, že hodnota tak jak ji známe z maticové algebry nelze zobecnit pro tenzory. Ukazujeme dvě možnosti zavedení hodnoty tenzoru, tzv. *vektorový a polyadický rank*. Dále definujeme normu tenzoru a *vnější součin tenzorů*, abychom si následně mohli ukázat, jak je to s *aproximací tenzoru tenzorem nižší hodnoty ve smyslu polyadického ranku*. V této kapitole se věnujeme také *zobecněnému rozvoji tenzoru v matici*, čímž se dostaneme zpět do světa matic k tzv. *Kroneckerovu součinu matic*. Nakonec je definován *tenzorový součin* a jsou popsány vztahy mezi jednotlivými součiny. V kapitole 3 je odvozen *Tuckerův rozklad* jako zobecnění singulárního rozkladu matice a je ukázáno, jak je možné v tomto tvaru provádět základní operace s velkými tenzory malého (vektorového) ranku. V kapitole 4 je ukázáno využití Tuckerova rozkladu pro *aproximaci tenzoru tenzorem s nižším vektorovým rankem*.

Poznamenejme ještě, že klasické učebnice, ať už zabývající se přímo tenzory (viz např. [27], [24]), nebo jen využívající aparát tenzorového počtu zpravidla vykládají látku v nějakém fyzikálním kontextu (např. v rámci teorie pružnosti, viz [1], [7], kde má ostatně tenzorový počet i pojem *tenzor* původ¹; kvantové mechaniky, viz [24];

¹Anglické slovíčko *tension* (z lat. *tensio*) překládáme jako *napětí, pnutí*; tenzory napětí a deformace jsou ústřední stavební kameny teorie pružnosti.

teorie relativity, viz [48], [31]; atd.). To vede k nutnosti rozlišování tzv. *kontravariantních* a *kovariantních souřadnic* (složek, indexů) tenzoru (např. druhá kapitola knihy [27] věnovaná *Základům tenzorového počtu* začíná § 11 *Kontravariantní a kovariantní složky vektoru*, první kapitola je věnovaná *Základům vektorového počtu*; obdobně v knize [24] se hned v úvodu po zavedení základních vektorových operací dostáváme na str. 23 ke kapitole 10 *Das kovariante und kontravariante Grundsystem*). Rozlišování kontra- a kovariantních souřadnic je praktické, zejména pracujeme-li s tenzorem jako s multilineárním zobrazením (viz např. [6]), nebo když pracujeme v křivočarých souřadnicích, což je ve fyzice častý případ (např. v knize [7] je celá první kapitola *Kartézské tenzory* (...) věnovaná výkladu tenzorů obdobně, jako se mu budeme věnovat my; s kontra- a kovariantní souřadnicemi se zde setkáváme až v druhé kapitole *Tenzory v metrických prostorech*, v odstavci 2.1.* *Křivočaré souřadnice* (...); obdobný obrázek získáme např. porovnáním výkladů tenzorových partií speciální [48] a obecné [31] teorie relativity). To také často vede k potřebě rozlišovat např. mezi *tenzory druhého řádu* (stupně) a *maticemi* a pod., tj. zkoumat je jako různé matematické objekty; např. Babuška et al. v knize [1, str. 35] píší: „Z tenzorového počtu je známo, že uvažovaná matice určuje tenzor 2. stupně (...)“.

Předkládaný text převážně studuje vnitřní strukturu tenzorů (např. právě pojem hodnoti) a právě proto s tenzorem pracuje jakožto s pevně daným *vícerozměrným polem čísel* (v literatuře se v této souvislosti můžeme setkat s pojmem *víceindexová matice*, viz [38, kapitola II, § 4, str. 52], případně anglicky *hypermatrix*, viz [5, kapitola 2.5.2, str. 318]). To nám dovoluje nerozlišovat kontra- a kovariantní souřadnice a proto se ani tyto pojmy ani související rozlišující značení (zpravidla realizované pomocí horních, resp. dolních indexů) v dalším textu vůbec nevyskytují (souřadnice tenzorů jsou vždy značeny pouze pomocí dolních indexů). Přesněji řečeno, pro další výklad je zcela lhostejné, zda jsou některé souřadnice tenzoru kontra- či kovariantní. V případech, kdy se text věnuje interakci tenzoru s okolím, konkrétně jeho součinu (v klasické literatuře také nazývaném *úžení*, viz [31, str. 70]) s jiným tenzorem (nebo maticí), prosíme laskavého čtenáře aby indexy, přes které se sčítá, vždy vnímal jako kontra- či kovariantní u prvního resp. druhého tenzoru (matice), nebo naopak, vždy však tak, aby bylo možné součin v očích čtenáře provést (připomeňme první kapitolu knihy [7], kde se s tenzory a jejich součiny pracuje stejně). Rozumíme-li pod pojmem *tenzor* pouze vícerozměrné pole čísel, můžeme také snadno tenzor nultého, prvního a druhého řádu ztotožnit přímo se skalárem, vektorem, resp. maticí.

Tento (zjednodušený) pohled na tenzory, tj. práce s nimi jako by byly pouze vícerozměrná pole čísel, je v současnosti velmi běžný zejména v oblasti aplikované a numerické lineární algebry, jak lze snadno doložit velkou řadou referencí, viz např. přehledový článek [28] v prestižním časopise *SIAM Review* (který cituje 223 dalších článků povětšinou z posledních let), rešeršní článek [18] (který cituje dalších 272 obdobných článků), knihy [14], [5], a mnoho dalších. Proto se i v předkládaném textu držíme jak tohoto pohledu na tenzor, tak i pojmu *tenzor* samotného a záměrně se vyhýbáme např. termínu vícerozměrná matice, který se v komunitě zabývající se (aplikovanou a lineární) algebrou používá zřídka.

1 Singulární rozklad

V této kapitole se budeme věnovat jednomu z neužitečnějších rozkladů matice, *singulárnímu rozkladu matice*. Singulární rozklad (SVD, z anglického singular value decomposition) využíváme pro výpočet hodnoty, normy matice, báze oboru hodnot a nulového prostoru matice atd. Důležitou roli hraje SVD také v úlohách komprese matice nebo data-mining. Ve statistice se se singulárním rozkladem setkáváme jako s analýzou hlavních komponent (PCA, z anglického principal component analysis), viz např. [14, str. 66].

1.1 Základní tvar SVD

Jedním z nejdůležitějších výsledků lineární algebry je následující věta.

Věta 1 (Singulární rozklad). *Pro matici $A \in \mathbb{R}^{n \times m}$ hodnosti $r \equiv \text{rank}(A)$ existují ortogonální matice $U = [u_1, \dots, u_n] \in \mathbb{R}^{n \times n}$ a $V = [v_1, \dots, v_m] \in \mathbb{R}^{m \times m}$ a kladná čísla*

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

tak, že platí

$$A = U \Sigma V^T, \tag{1.1}$$

kde

$$\Sigma = \begin{bmatrix} \Sigma_A & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times m}, \quad \Sigma_A = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r \times r}.$$

Sloupce matice U , tedy vektory u_j , $j = 1, \dots, n$, se nazývají levé singulární vektory, sloupce matice V , tedy vektory v_j , $j = 1, \dots, m$, se nazývají pravé singulární vektory matice A . Čísla σ_j , $j = 1, \dots, r$, se nazývají singulární čísla matice A .

Důkaz. Důkaz neprovedeme úplně, ale pouze naznačíme. Podle věty o spektrálním rozkladu matice (viz např. [11, str. 196], [12, str. 40]) pro matici $B \in \mathbb{R}^{n \times n}$ hodnosti r , která je symetrická pozitivně semidefinitní existuje rozklad

$$B = G \Lambda G^T, \quad G^T G = 1, \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n), \tag{1.2}$$

přičemž

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0.$$

Rozklad (1.2) nazýváme spektrální rozklad matice. Věta o SVD ale pracuje s obecně obdélníkovou maticí $A \in \mathbb{R}^{n \times m}$, proto se k odvození SVD využívá matic AA^T a $A^T A$, které vlastnosti předpokládající spektrální rozklad splňují (jsou čtvercové symetrické pozitivně definitní). Je-li $A = U\Sigma V^T$, resp. $A^T = V\Sigma^T U^T$, potom platí

$$AA^T = U\Sigma\Sigma^T U^T, \quad A^T A = V\Sigma^T \Sigma V^T. \quad (1.3)$$

Srovnáním (1.2) a (1.3) vidíme, že v prvním případě $B = AA^T$, $G = U$ a $\Lambda = \Sigma\Sigma^T$. Uvědomme si, že $\sigma_j = \sqrt{\lambda_j}$, $j = 1, \dots, r$. Pro úplný důkaz odkazujeme na skripta [12, kap. 5], kde je proveden postupným odvozením. \square

Poznamenejme, že SVD má zajímavou geometrickou interpretaci, pokud se na matici díváme jako na lineární zobrazení $A : \mathbb{R}^m \mapsto \mathbb{R}^n$. Věta 1 říká, že pro každou matici $A \in \mathbb{R}^{n \times m}$ hodnoti r existuje ortonormální báze $\{v_1, \dots, v_n\}$, prostoru \mathbb{R}^m a ortonormální báze $\{u_1, \dots, u_n\}$ prostoru \mathbb{R}^n a kladná čísla $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ tak, že platí

$$Av_j = \begin{cases} \sigma_j u_j, & j = 1, \dots, r, \\ 0, & j = r + 1, \dots, m \end{cases},$$

$$A^T u_j = \begin{cases} \sigma_j v_j, & j = 1, \dots, r, \\ 0, & j = r + 1, \dots, n \end{cases}.$$

1.2 Ekonomický tvar SVD

Pro praktické použití je ale mnohem výhodnější tzv. *ekonomický tvar* singulárního rozkladu. Matice Σ obsahuje nulové bloky, které násobí příslušné sloupce, resp. řádky matice U , resp. V^T . Ekonomický tvar spočívá ve vynechání těchto částí matic, které vedou k násobení nulami. Získáme tedy matice, které budeme značit $U_A \equiv [u_1, \dots, u_r] \in \mathbb{R}^{n \times r}$ a $V_A \equiv [v_1, \dots, v_r] \in \mathbb{R}^{m \times r}$. Ekonomický tvar singulárního rozkladu se potom zapisuje

$$A = U_A \Sigma_A V_A^T. \quad (1.4)$$

Pro lepší představu můžeme tento tvar znázornit pomocí následujícího schématu (převzato z [12, str. 127]):

$$\begin{array}{c} A \\ \boxed{} \end{array} = \begin{array}{c} U \\ \boxed{} \end{array} \begin{array}{c} \Sigma \\ \begin{array}{|c|c|} \hline \Sigma_A & 0 \\ \hline 0 & 0 \\ \hline \end{array} \end{array} \begin{array}{c} V^T \\ \boxed{} \end{array} = \begin{array}{c} U_A \\ \boxed{} \end{array} \begin{array}{c} \Sigma_A \\ \boxed{} \end{array} \begin{array}{c} V_A^T \\ \boxed{} \end{array}.$$

Důsledek 1. Nesporným přínosem SVD v ekonomickém tvaru je možnost využití při kompresi dat. Matici A o rozměrech $n \times m$ jsme vztahem (1.4) vyjádřili pomocí tří matic $U_A \in \mathbb{R}^{n \times r}$, $\Sigma_A \in \mathbb{R}^{r \times r}$, $V_A \in \mathbb{R}^{m \times r}$. Rozměry těchto tří matic jsou $(m + n + r) \times r$. V případě, že matice A má malou hodnotu, $r \ll \min\{m, n\}$, je vyjádření v ekonomickém tvaru SVD z hlediska objemu dat úspornější, viz tabulka 1.1.

Tabulka 1.1: Počet prvků matice a jejího ekonomického SVD (vpravo s využitím faktu, že Σ_A je diagonální); $r = \text{rank}(A)$.

| Matice | Počet prvků |
|---------------------------------|---|
| $A \in \mathbb{R}^{n \times m}$ | $m \cdot n$ |
| U_A, Σ_A, V_A | $(m + n + r) \cdot r$ resp. $(m + n + 1) \cdot r$ |

Datová úspora může hrát roli např. když $m \cdot n$ je příliš velké v porovnání s pamětí počítače (cache paměť, operační paměť, ...). To se v případě matic nemusí zdát relevantní, nicméně později uvidíme, že u tenzorů to bude hrát roli. Ekonomický rozklad matice A můžeme vyjádřit také jako součet vnějších součinů,

$$A = \sum_{\ell=1}^r \sigma_{\ell} u_{\ell} v_{\ell}^T = \sum_{\ell=1}^r A_{\ell}, \quad A_{\ell} \equiv \sigma_{\ell} u_{\ell} v_{\ell}^T. \quad (1.5)$$

Tento tvar nazýváme *dyadický rozvoj* obecné matice A . Hodnota každé matice A_j je zřejmě jedna, pro které platí

$$\|A_j\| = \|A_j\|_F = \sigma_j. \quad (1.6)$$

Singulární čísla matice A byla uspořádána od největšího k nejmenšímu, proto jsou i normy matic A_j uspořádány $\|A_1\| \geq \|A_2\| \geq \dots \geq \|A_r\| > 0$.

1.3 Operace s maticí uloženou v ekonomickém tvaru SVD

V této podkapitole ukážeme, jak lze pracovat s maticemi uloženými v ekonomickém tvaru SVD, přičemž budeme chtít i výsledek získat ve stejném tvaru. Konkrétně se budeme zabývat operacemi sčítání a násobení. Poznamenejme, že provádění operací s maticemi tímto způsobem se nazývá *low-rank* aritmetika.

1.3.1 Sčítání dvou matic v ekonomickém tvaru

Nejprve ukážeme jak sčítat dvě matice vyjádřené v ekonomickém tvaru SVD. Mějme dvě matice $A, B \in \mathbb{R}^{n \times m}$, jejichž hodnosti jsou

$$r = \text{rank}(A), \quad p = \text{rank}(B), \quad r, p \ll \min\{m, n\}$$

a singulární rozklady jsou

$$\begin{aligned} A &= U_A \Sigma_A V_A^T, & U_A &\in \mathbb{R}^{n \times r}, \Sigma_A \in \mathbb{R}^{r \times r}, V_A \in \mathbb{R}^{m \times r}, \\ B &= U_B \Sigma_B V_B^T, & U_B &\in \mathbb{R}^{n \times p}, \Sigma_B \in \mathbb{R}^{p \times p}, V_B \in \mathbb{R}^{m \times p}. \end{aligned}$$

Matici $C = A + B$ v ekonomickém singulárním tvaru získáme ve třech krocích.

Krok 1 (Sestavení matic): Označme součet těchto matic $C \in \mathbb{R}^{n \times m}$, pak zřejmě

$$C = A + B = [U_A, U_B] \begin{bmatrix} \Sigma_A & 0 \\ 0 & \Sigma_B \end{bmatrix} [V_A, V_B]^T. \quad (1.7)$$

Pro hodnost matice C platí

$$q = \text{rank}(C), \quad \text{kde } 0 \leq q \leq p + r.$$

Poznamenejme, že součet matic ve tvaru (1.7) nás nestojí žádnou matematickou operaci, získáme jej pouze „poskládáním“ příslušných matic z rozkladů A a B matic vedle sebe. Matice $[U_A, U_B]$ a $[V_A, V_B]$ už ale nemají ortonormální sloupce. Tuto vlastnost můžeme někdy požadovat. Ukážeme, jakým způsobem získat matici C v ekonomickém tvaru SVD.

Krok 2 (Ortogonalizace matic): Nejprve musíme dosáhnout ortonormality. Prvním krokem je tedy QR rozklad matic $[U_A, U_B]$ a $[V_A, V_B]$; pro detailní popis QR rozkladu odkazujeme na [12, kapitola 3]. QR rozkladem získáme matici ve tvaru součinu matice s ortonormálními sloupci Q a matice horní trojúhelníkové R ,

$$\begin{aligned} [U_A, U_B] &= Q_U R_U, & Q_U &\in \mathbb{R}^{n \times (r+p)}, R_U \in \mathbb{R}^{(r+p) \times (r+p)}, \\ [V_A, V_B] &= Q_V R_V, & Q_V &\in \mathbb{R}^{m \times (r+p)}, R_V \in \mathbb{R}^{(r+p) \times (r+p)}. \end{aligned}$$

Matice C pak můžeme zapsat

$$C = Q_U \underbrace{\left(R_U \begin{bmatrix} \Sigma_A & 0 \\ 0 & \Sigma_B \end{bmatrix} R_V^T \right)}_W Q_V^T.$$

Dostáváme

$$C = Q_U W Q_V^T. \quad (1.8)$$

V této fázi máme rozklad matice C , kde Q_U a Q_V jsou matice s ortonormálními sloupci. Poznamenejme, že získání tohoto tvaru už je náročnější, stojí nás dva QR rozklady matic velikosti $n \times (p + r)$ a $m \times (p + r)$ a výpočet matice W (maticové násobení).

Krok 3 (SVD): Pokud chceme získat matici C ve tvaru ekonomického SVD, tj. rozklad na dvě matice s ortonormálními sloupci a diagonální matici se singulárními čísly, ještě nejsme hotovi. SVD rozklad matice $W \in \mathbb{R}^{(r+p) \times (r+p)}$ je

$$W = U_W \Sigma_W V_W^T, \quad \Sigma_W \in \mathbb{R}^{q \times q}, \quad \text{rank}(W) = q. \quad (1.9)$$

Vyjádríme-li matici C pomocí (1.8) a (1.9) a získáváme SVD rozklad

$$C = \underbrace{(Q_U U_W)}_{U_C} \Sigma_W \underbrace{(Q_V V_W^T)}_{V_C} = U_C \Sigma_W V_C^T. \quad (1.10)$$

K ceně výpočtu matice C ve tvaru (1.10) musíme k operacím, kterými jsme získali tvar (1.8), přidat ještě výpočet SVD matice W a další maticová násobení. Ceny jednotlivých kroků jsou uvedeny v tabulce 1.2.

Tabulka 1.2: Výpočetní cena při sčítání matic $A, B \in \mathbb{R}^{n \times m}$; ceny výpočtů QR rozkladu a SVD lze nalézt např. v [12, kapitoly 3.5.6 a 5.2.3]. Zkratka MMp značí součin dvou matic (matrix-matrix product).

| Operace | Výpočetní cena |
|-----------------|---|
| klasicky | $m \cdot n$ sčítání |
| ek. SVD, krok 1 | zdarma (sestavení matic) |
| ek. SVD, krok 2 | 1 QR rozklad typu $(n \times (r + p))$ + 1 QR rozklad typu $(m \times (r + p))$ + 2 MMp typu $((r + p) \times (r + p)) \cdot ((r + p) \times (r + p))$ |
| ek. SVD, krok 3 | 1 SVD typu $((r + p) \times (r + p))$ + 1 MMp typu $(n \times (r + p)) \cdot ((r + p) \times q)$, + 1 MMp typu $(m \times (r + p)) \cdot ((r + p) \times q)$ |

1.3.2 Násobení matice v ekonomickém tvaru obecnou maticí

Nyní popíšeme, jakým způsobem násobit dvě matice, je-li jedna z nich uložená ve tvaru SVD. Nechť $A = U_A \Sigma_A V_A^T \in \mathbb{R}^{n \times m}$ a matice $M \in \mathbb{R}^{m \times k}$ je taková, že buď $k \ll \min\{m, n\}$ nebo je M řídká. Podobné techniky se reálně používají v praktických výpočtech viz např. [29, str. 670], odkazy na další literaturu a aplikace je možné nalézt také v rešeršním článku [18]. Stejně jako jsme v předchozí části hledali tvar součtu matice v SVD, zde ukážeme jak takový tvar získat v případě násobení. Výsledek opět získáme ve třech krocích.

Krok 1: Součin matic A a M je matice $D \in \mathbb{R}^{n \times k}$ hodnosti $\text{rank}(D) = q \leq r$,

$$D = AM = U_A \Sigma_A \underbrace{(M^T V_A)}_K^T = U_A \Sigma_A K^T, \quad K \in \mathbb{R}^{k \times r}. \quad (1.11)$$

Cena výpočtu matice D ve tvaru (1.11) je dána pouze maticovým násobením. Matice K ale nemá ortogonální sloupce.

Krok 2 (ortogonalizace): Chceme-li dosáhnout ortonormality, musíme vypočítat QR rozklad matice K . Získáváme

$$K = Q_K R_K, \quad Q_K \in \mathbb{R}^{k \times r}, R_K \in \mathbb{R}^{r \times r}. \quad (1.12)$$

Dosadíme (1.12) do (1.11) a dostaneme

$$D = U_A \underbrace{(\Sigma_A R_K^T)}_W Q_K^T = U_A W Q_K^T, \quad W \in \mathbb{R}^{r \times r}. \quad (1.13)$$

Cenou za získání (1.13) oproti tvaru v (1.11) je QR rozklad matice o rozměrech $k \times r$ a maticové násobení.

Tabulka 1.3: Výpočetní cena při násobení matic $A \in \mathbb{R}^{n \times m}$ a $M \in \mathbb{R}^{m \times k}$; ceny výpočtů QR rozkladu a SVD lze nalézt např. v [12, kapitoly 3.5.6 a 5.2.3]. Zkratka MMp značí součin dvou matic (matrix-matrix product), MVp je součin matice a vektoru (matrix-vector product).

| Výpočet | Výpočetní cena |
|-----------------|---|
| klasicky | 1 MMp typu $(n \times m) \cdot (m \times k)$ = n MVp typu $(k \times m) \cdot m$ |
| ek. SVD, krok 1 | 1 MMp typu $(k \times m) \cdot (m \times r)$ |
| ek. SVD, krok 2 | 1 QR rozklad typu $(k \times r)$ + 1 MMp typu $(r \times r) \cdot (r \times r)$ |
| ek. SVD, krok 3 | 1 SVD typu $(r \times r)$ + 1 MMp typu $(n \times r) \cdot (r \times q)$ + 1 MMp typu $(m \times r) \cdot (r \times q)$ |

Krok 3 (SVD): Je zřejmé, že matice W není diagonální matice se singulárními čísly a tedy abychom získali matici D v ekonomickém SVD tvaru, budeme muset učinit ještě další kroky. Využijeme singulární rozklad matice W ,

$$W = U_W \Sigma_W V_W^T, \quad \text{rank}(W) = q. \quad (1.14)$$

Pokud potom dosadíme (1.14) do (1.13), získáme

$$D = \underbrace{(U_A U_W)}_{U_D} \Sigma_W \underbrace{(Q_K V_W^T)}_{V_D} = U_D \Sigma_W V_D^T. \quad (1.15)$$

Naším cílem byl právě tvar (1.15). Náročnost jeho získání stoupla oproti (1.13) o výpočet SVD matice velikosti $r \times r$ a další maticová násobení. Ceny jednotlivých kroků jsou přehledně uvedeny v tabulce 1.3.

1.4 Aproximace matice maticí nižší hodnosti

V této podkapitole ukážeme, jakou roli hraje singulární rozklad matice, pokud chceme aproximovat matici maticí nižší hodnosti. Uvažujme matici A hodnosti r , která má dyadický rozvoj získaný pomocí singulárního rozkladu

$$A = \sum_{\ell=1}^r \sigma_{\ell} u_{\ell} v_{\ell}^T,$$

viz (1.5). Pokud se pokusíme nalézt matici B takovou, že

$$\text{rank}(B) < \text{rank}(A) = r,$$

a zároveň $\|A - B\|_F$ nebo $\|A - B\|$ je minimální, pak

$$B = \sum_{\ell=1}^{r-1} \sigma_\ell u_\ell v_\ell^T, \quad \text{rank}(B) = r - 1$$

a

$$\|A - B\|_F = \|A - B\| = \sigma_r,$$

tj. minimum je dané nejmenším singulárním číslem σ_r matice A . Platí tedy následující věta, kterou uvedeme bez důkazu.

Věta 2 (Schmidt, Eckart, Young, Mirsky). *Nechť $A \in \mathbb{R}^{n \times m}$ je matice hodnosti r a (1.4)–(1.5) je její singulární rozklad. Nechť t je libovolné přirozené číslo, pro které platí $t < r$. Potom matice*

$$B = \sum_{\ell=1}^t \sigma_\ell u_\ell v_\ell^T = \sum_{\ell=1}^t A_\ell, \quad A_\ell \equiv \sigma_\ell u_\ell v_\ell^T,$$

je nejlepší aproximací matice A v následujícím smyslu:

$$\begin{aligned} \min_{X \in \mathbb{R}^{n \times m}, \text{rank}(X) \leq t} \|A - X\| &= \|A - B\| = \sigma_{t+1}, \\ \min_{X \in \mathbb{R}^{n \times m}, \text{rank}(X) \leq t} \|A - X\|_F &= \|A - B\|_F = \left(\sum_{j=t+1}^r \sigma_j^2 \right)^{1/2}. \end{aligned} \quad (1.16)$$

Pro podrobnější výklad a důkaz odkazujeme např. na knihu [40]. Původní důkaz lze nalézt v článku Schmidta [36]. Často je připisován Youngovi a Eckartovi [13], kteří výsledek dokázali pro Frobeniovu normu a Mirskému [33], který výsledek dokázal pro spektrální normu. Předchozí věta má následující zajímavý důsledek.

Důsledek 2. *V prostoru matic $\mathbb{R}^{n \times m}$ existuje kolem každé matice A hodnosti r otevřená koule o poloměru daném nejmenším singulárním číslem σ_r , uvnitř které se nacházejí pouze matice hodnosti r .*

Poznamenejme ještě, že v souvislosti s aproximací matice maticí nižší hodnosti se můžeme setkat s pojmem *truncated SVD* (TSVD), tedy „oříznutý“ singulární rozklad. Obvykle se však pod tímto pojmem skrývá metoda řešení soustav rovnic $Ax = b$, resp. aproximačních problémů $Ax \approx b$, kdy se matice A nahradí právě maticí B a výsledný problém se řeší metodou nejmenších čtverců, jedná se o tzv. regularizační metodu, viz [20].

2 Tenzory

V této kapitole se posuneme z lineární algebry do tzv. algebry multilineární. Zatímco lineární algebra pracuje s maximálně dvourozměrnými útvary (tj. s maticemi, vektory nebo čísly), my se nyní budeme zabývat vícerozměrnými objekty, tenzory. Zavedeme jejich značení a terminologii a popíšeme některé jejich vlastnosti a základní operace s tenzory. Na úvod poznamenejme, že za tenzor budeme považovat vícerozměrné pole čísel. Tenzor k -tého řádu je tedy uspořádaná množina čísel

$$\mathcal{A} = (a_{i_1, i_2, \dots, i_k})_{i_\ell=1, \dots, n_\ell, \ell=1, 2, \dots, k} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}.$$

Pokud bude z kontextu jasné, jaké jsou rozsahy indexů, budeme psát jen

$$\mathcal{A} = (a_{i_1, i_2, \dots, i_k})_{i_1, i_2, \dots, i_k}, \quad \text{případně jen } \mathcal{A} = (a_{i_1, i_2, \dots, i_k}) \dots$$

Speciálně tenzor nulého řádu je *skalár* (číslo), tenzor prvního řádu je *vektor*, tenzor druhého řádu je *matice*, tenzory řádu tři a výše nemají ustálené pojmenování.

Příklad 1. Pro lepší představu uvádíme jako příklad tenzor $\mathcal{A} \in \mathbb{R}^{4 \times 3 \times 2}$, který použijeme následně i při výkladu terminologie tenzorů:

$$\mathcal{A} = \begin{array}{|c|c|c|c|} \hline 6 & 4 & 1 & \\ \hline 6 & 3 & 2 & \\ \hline 7 & 1 & 0 & \\ \hline 7 & 7 & 0 & \\ \hline 3 & 0 & 7 & \\ \hline \end{array} \in \mathbb{R}^{4 \times 3 \times 2}. \quad (2.1)$$

□

Tenzor $\mathcal{A} = (a_{i_1, i_2, \dots, i_k}) \dots$ k -tého řádu si můžeme také představit jako funkci k proměnných

$$a(x_1, \dots, x_k) : \mathbb{R}^k \mapsto \mathbb{R}, \quad (2.2)$$

jejíž funkční hodnoty jsou však známy jen v několika málo diskrétních bodech. Konkrétně pro

$$x_j \in \{x_{j,1}, \dots, x_{j,n_j}\} \equiv \mathcal{X}_j, \quad j = 1, \dots, k,$$

tedy na „mřížce“, přesněji řečeno kartézském součinu $\mathcal{X}_1 \times \dots \times \mathcal{X}_k$. Tento přístup může v některých případech výrazně zjednodušit zápis při práci s tenzory, viz např. přednášku [47] nebo práci [15, kapitola 1.2]. Velmi zajímavý a rozsahem i obsahem neobvyklý úvod do práce s tenzory lze také nalézt v pozoruhodné knize R. Sikorskiho [38, kapitola II, §§ 4–10, str. 52–104], která je jinak věnována právě analýze funkcí více proměnných.

2.1 Značení a terminologie

Značení a terminologii tenzorů zavedeme v této podkapitole pro jednoduchost a názornost nejprve pro tenzor řádu 3, tj. $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Následně ji zobecníme i pro tenzory vyšších řádů. Vycházet budeme především z článku [28].

2.1.1 Vlákna tenzoru

Pokud zafixujeme jeden nebo více indexů, získáme určitou podmnožinu původního objektu (tenzoru). Např. v případě matic hovoříme o řádcích a sloupcích. Analogií řádků a sloupců jsou pro tenzory „vlákna“, v angličtině *fibres*. Obecně v tenzoru k -tého řádu vznikne vlákno zafixováním $(k - 1)$ indexů. Vlákna tenzoru třetího řádu znázorňuje obrázek 2.1. Následuje formální definice pojmu vlákno.

Definice 1 (Vlákno). *Nechť $\mathcal{A} = (a_{i_1, i_2, i_3})_{i_1, i_2, i_3} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ tenzor řádu 3. Vektory*

$$\begin{aligned} a_{:, i_2, i_3} &\equiv (a_{i_1, i_2, i_3})_{i_1} = [a_{1, i_2, i_3}, \dots, a_{n_1, i_2, i_3}]^T \in \mathbb{R}^{n_1}, \\ a_{i_1, :, i_3} &\equiv (a_{i_1, i_2, i_3})_{i_2} = [a_{i_1, 1, i_3}, \dots, a_{i_1, n_2, i_3}]^T \in \mathbb{R}^{n_2}, \\ a_{i_1, i_2, :} &\equiv (a_{i_1, i_2, i_3})_{i_3} = [a_{i_1, i_2, 1}, \dots, a_{i_1, i_2, n_3}]^T \in \mathbb{R}^{n_3} \end{aligned}$$

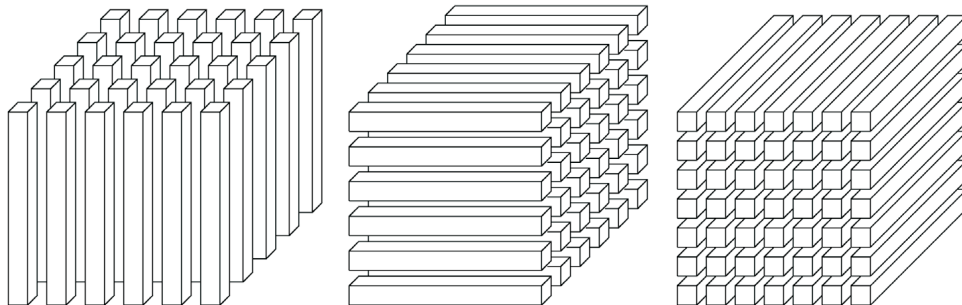
nazýváme vlákna v módu 1, 2, resp. 3.

Analogicky, pro tenzor k -tého řádu $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$ představují vektory

$$\begin{aligned} a_{:, i_2, \dots, i_k} &\equiv [a_{1, i_2, \dots, i_k}, \dots, a_{n_1, i_2, \dots, i_k}]^T \in \mathbb{R}^{n_1}, \\ a_{i_1, :, \dots, i_k} &\equiv [a_{i_1, 1, \dots, i_k}, \dots, a_{i_1, n_2, \dots, i_k}]^T \in \mathbb{R}^{n_2}, \\ &\vdots \\ a_{i_1, i_2, \dots, :} &\equiv [a_{i_1, i_2, \dots, 1}, \dots, a_{i_1, i_2, \dots, n_k}]^T \in \mathbb{R}^{n_k} \end{aligned}$$

vlákna v módu 1, 2, resp. k .

(Poznamenejme, že dvojtečka představuje všechny prvky daného směru a má tedy stejnou funkci jako v MATLABU.)



Obrázek 2.1: Vlákna postupně v módu 1, 2, resp. 3, převzato z článku [28].

Příklad 2. Význam pojmu vlákno si ukážeme na tenzoru z příkladu 1. Vlákna v módu 1 jsou prvky \mathbb{R}^4 :

$$a_{:,1,1} = a_{:,1,2} = a_{:,2,1} = a_{:,2,2} = a_{:,3,1} = a_{:,3,2} =$$

$$\begin{bmatrix} 6 \\ 7 \\ 7 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \\ 9 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 1 \\ 7 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \\ 7 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ 8 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 4 \\ 6 \end{bmatrix}.$$

Vlákna v módu 2 jsou prvky \mathbb{R}^3 :

$$a_{1,,:,1} = a_{1,,:,2} = a_{2,,:,1} = a_{2,,:,2} = a_{3,,:,1} = a_{3,,:,2} = a_{4,,:,1} = a_{4,,:,2} =$$

$$\begin{bmatrix} 6 \\ 6 \\ 2 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 7 \\ 7 \\ 0 \end{bmatrix}, \begin{bmatrix} 9 \\ 7 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 8 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \\ 6 \end{bmatrix}.$$

Vlákna v módu 3 jsou prvky \mathbb{R}^2 :

$$a_{1,1,:} = a_{1,2,:} = a_{1,3,:} = a_{2,1,:} = a_{2,2,:} = a_{2,3,:} =$$

$$\begin{bmatrix} 6 \\ 6 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix},$$

$$a_{3,1,:} = a_{3,2,:} = a_{3,3,:} = a_{4,1,:} = a_{4,2,:} = a_{4,3,:} =$$

$$\begin{bmatrix} 7 \\ 9 \end{bmatrix}, \begin{bmatrix} 7 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \end{bmatrix}, \begin{bmatrix} 8 \\ 6 \end{bmatrix}.$$

□

2.1.2 Řezy tenzoru

Zafixováním $(k - 2)$ indexů tenzoru k -tého řádu, získáme „řezy“ tenzoru, v anglické literatuře označované *slices*. Řezy tenzoru třetího řádu schematicky znázorňuje obrázek 2.2. Následuje formální definice pojmu řez.

Definice 2 (Řez). Necht' $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ je tenzor řádu 3. Matice

$$A_{i_1, :, :} \equiv (a_{i_1, i_2, i_3})_{i_2, i_3} \in \mathbb{R}^{n_2 \times n_3},$$

$$A_{:, i_2, :} \equiv (a_{i_1, i_2, i_3})_{i_1, i_3} \in \mathbb{R}^{n_1 \times n_3},$$

$$A_{:, :, i_3} \equiv (a_{i_1, i_2, i_3})_{i_1, i_2} \in \mathbb{R}^{n_1 \times n_2}$$

nazýváme řezy v módu $(2, 3)$, $(1, 3)$, resp. $(1, 2)$. Matice

$$A_{i_1, :, :}^T \in \mathbb{R}^{n_3 \times n_2}, \quad A_{:, i_2, :}^T \in \mathbb{R}^{n_3 \times n_1}, \quad A_{:, :, i_3}^T \in \mathbb{R}^{n_2 \times n_1}$$

jsou řezy v módu $(3, 2)$, $(3, 1)$, resp. $(2, 1)$.

Analogicky, pro tenzor k -tého řádu $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}$ je matice

$$A_{i_1, \dots, i_{\ell-1}, :, i_{\ell+1}, \dots, i_{s-1}, :, i_{s+1}, \dots, i_k} \equiv (a_{i_1, \dots, i_k})_{i_{\ell}, i_s} \in \mathbb{R}^{n_{\ell} \times n_s}, \quad \ell < s$$

řezem v módu \mathcal{A} (ℓ, s). Matice

$$A_{i_1, \dots, i_{\ell-1}, :, i_{\ell+1}, \dots, i_{s-1}, :, i_{s+1}, \dots, i_k}^T \in \mathbb{R}^{n_s \times n_\ell}$$

je řezem v módu (s, ℓ).

Příklad 3. Řezy tenzoru si opět ukážeme na tenzoru \mathcal{A} z příkladu 1. Řezy v módu (2, 3) jsou prvky $\mathbb{R}^{3 \times 2}$:

$$A_{1, :, :} = \begin{bmatrix} 6 & 6 \\ 6 & 4 \\ 2 & 1 \end{bmatrix}, A_{2, :, :} = \begin{bmatrix} 7 & 3 \\ 1 & 3 \\ 0 & 4 \end{bmatrix}, A_{3, :, :} = \begin{bmatrix} 7 & 9 \\ 7 & 7 \\ 0 & 4 \end{bmatrix}, A_{4, :, :} = \begin{bmatrix} 3 & 0 \\ 0 & 7 \\ 8 & 6 \end{bmatrix},$$

řezy v módu (1, 3) jsou prvky $\mathbb{R}^{4 \times 2}$:

$$A_{:, 1, :} = \begin{bmatrix} 6 & 6 \\ 7 & 3 \\ 7 & 9 \\ 3 & 0 \end{bmatrix}, A_{:, 2, :} = \begin{bmatrix} 6 & 4 \\ 1 & 3 \\ 7 & 7 \\ 0 & 7 \end{bmatrix}, A_{:, 3, :} = \begin{bmatrix} 2 & 1 \\ 0 & 4 \\ 0 & 4 \\ 8 & 6 \end{bmatrix},$$

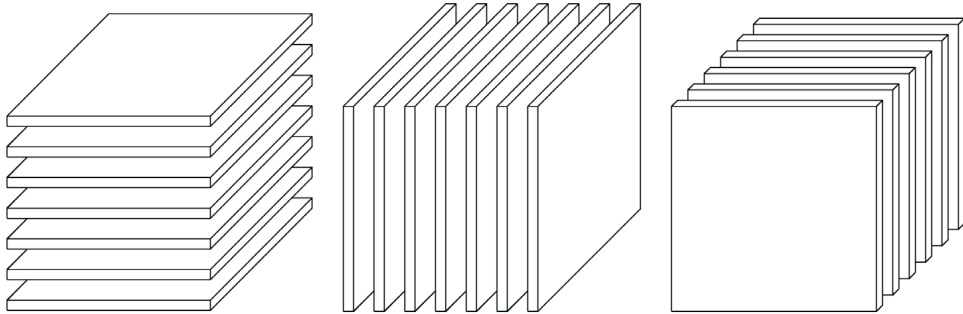
řezy v módu (1, 2) jsou prvky $\mathbb{R}^{4 \times 3}$:

$$A_{:, :, 1} = \begin{bmatrix} 6 & 6 & 2 \\ 7 & 1 & 0 \\ 7 & 7 & 0 \\ 3 & 0 & 8 \end{bmatrix}, A_{:, :, 2} = \begin{bmatrix} 6 & 4 & 1 \\ 3 & 3 & 4 \\ 9 & 7 & 4 \\ 0 & 7 & 6 \end{bmatrix}.$$

Transpozicemi těchto matic jsou řezy v módech (3, 2), (3, 1), resp. (2, 1). Řezy v módu (3, 2) jsou prvky $\mathbb{R}^{2 \times 3}$:

$$A_{1, :, :}^T = \begin{bmatrix} 6 & 6 & 2 \\ 6 & 4 & 1 \end{bmatrix}, A_{2, :, :}^T = \begin{bmatrix} 7 & 1 & 0 \\ 3 & 3 & 4 \end{bmatrix},$$

$$A_{3, :, :}^T = \begin{bmatrix} 7 & 7 & 0 \\ 9 & 7 & 4 \end{bmatrix}, A_{4, :, :}^T = \begin{bmatrix} 3 & 0 & 8 \\ 0 & 7 & 6 \end{bmatrix},$$



Obrázek 2.2: Řezy tenzoru řádu tři, převzato z článku [28]. Řezy tenzoru třetího řádu někdy nazýváme také horizontální, laterální, resp. frontální.

řezy v módu (3, 1) jsou prvky $\mathbb{R}^{2 \times 4}$:

$$A_{:,1,:}^T = \begin{bmatrix} 6 & 7 & 7 & 3 \\ 6 & 3 & 9 & 0 \end{bmatrix}, A_{:,2,:}^T = \begin{bmatrix} 6 & 1 & 7 & 0 \\ 4 & 3 & 7 & 7 \end{bmatrix}, A_{:,3,:}^T = \begin{bmatrix} 2 & 0 & 0 & 8 \\ 1 & 4 & 4 & 6 \end{bmatrix},$$

řezy v módu (2, 1) jsou prvky $\mathbb{R}^{3 \times 4}$:

$$A_{:,:,1}^T = \begin{bmatrix} 6 & 7 & 7 & 3 \\ 6 & 1 & 7 & 0 \\ 2 & 0 & 0 & 8 \end{bmatrix}, A_{:,:,2}^T = \begin{bmatrix} 6 & 3 & 9 & 0 \\ 4 & 3 & 7 & 7 \\ 1 & 4 & 4 & 6 \end{bmatrix}.$$

□

2.1.3 Transpozice a symetrie tenzoru

Dalším běžně používaným nástrojem při práci s maticemi je transpozice matice. Transpozicí matice (tj. tenzoru druhého řádu), která znamená záměnu řádků a sloupců, můžeme v kontextu tenzorů interpretovat jako záměnu vláken v módech 1 a 2. Analogicky potom můžeme za *transpozici tenzoru* považovat libovolnou záměnu (permutaci) jeho indexů. Uvažujme tenzor $\mathcal{A} = (a_{i_1, i_2, \dots, i_k}) \dots \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$ řádu k a permutaci

$$\Pi_j = \begin{pmatrix} 1 & 2 & \dots & k \\ \pi_1 & \pi_1 & \dots & \pi_k \end{pmatrix}, \quad j = 1, \dots, k!$$

Pak tenzor

$$\mathcal{A}^{\Pi_j} = (a_{i_{\pi_1}, i_{\pi_2}, \dots, i_{\pi_k}}) \dots \in \mathbb{R}^{n_{\pi_1} \times n_{\pi_2} \times \dots \times n_{\pi_k}}$$

nazveme (Π_j) -tou transpozicí tenzoru \mathcal{A} . Transpozice tenzoru je tedy de-facto *permutace indexů* tenzoru.

Stejně jako při práci s maticemi běžně používáme pojmy popisující specifické vlastnosti a tvary matice, např. matice transponovaná, symetrická, diagonální, čtvercová apod., existují takové pojmy i pro tenzory. Tenzor, který se nezmění při permutaci dvou nebo více indexů, se někdy nazývá symetrický tenzor, viz [28], někdy je též označován jako částečně symetrický, viz [9]. Pokud se tenzor nezmění při *libovolné* permutaci indexů, nazývá se supersymetrický, viz [28], někdy též symetrický, viz [9]. V tomto textu budeme při práci s tenzory používat výhradně pojmy *částečně symetrický*, resp. *supersymetrický*, bude-li to ovšem třeba; pojem *symetrický* bude rezervován pouze pro matice, aby nedošlo k nejasnostem.

Příklad 4. Tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_2}$ je symetrický v módech dva a tři, pokud pro všechny jeho horizontální řezy, tj. řezy v módu (2, 3) jsou symetrické matice

$$A_{i_1, :, :} = A_{i_1, :, :}^T, \quad i_1 = 1, \dots, n_1.$$

□

Příklad 5. Tenzor třetího řádu $\mathcal{A} \in \mathbb{R}^{n \times n \times n}$ je supersymetrický pokud platí

$$a_{i_1, i_2, i_3} = a_{i_1, i_3, i_2} = a_{i_2, i_1, i_3} = a_{i_2, i_3, i_1} = a_{i_3, i_1, i_2} = a_{i_3, i_2, i_1}$$

pro $i_1 = 1, \dots, n$, $i_2 = 1, \dots, n$, $i_3 = 1, \dots, n$.

□

Tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$, pro který platí $a_{i_1, i_2, \dots, i_n} \neq 0$ pouze pokud $i_1 = i_2 = \dots = i_n$, nazýváme *diagonální*. Tenzor $\mathcal{A} \in \mathbb{R}^{n \times n \times \dots \times n}$, který má všechny rozměry stejné, nazýváme *kubický tenzor*.

2.2 Základní operace s tenzory

Řada nástrojů běžně užívaných v maticovém počtu lze přímočaře zobecnit do tenzoru. Například můžeme tenzory stejného typu sčítat a násobit číslem; zřejmě tvoří *lineární vektorový prostor*. Nechť

$$\mathcal{A} = (a_{i_1, i_2, \dots, i_k})_{i_\ell=1, \dots, n_\ell, \ell=1, 2, \dots, k} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k},$$

$$\mathcal{B} = (b_{i_1, i_2, \dots, i_k})_{i_\ell=1, \dots, n_\ell, \ell=1, 2, \dots, k} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$$

a

$$\alpha, \beta \in \mathbb{R},$$

potom

$$\alpha \mathcal{A} + \beta \mathcal{B} = (\alpha a_{i_1, i_2, \dots, i_k} + \beta b_{i_1, i_2, \dots, i_k})_{\dots} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}.$$

Dále můžeme tenzory násobit maticí.

Definice 3 (Součin v ℓ -tém módu). *Uvažujme tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$ a matici $M \in \mathbb{R}^{m \times n_\ell}$. Pak tenzor*

$$\mathcal{D} \equiv \mathcal{A} \times_\ell M = \left(\sum_{i_\ell=1}^{n_\ell} a_{i_1, \dots, i_{\ell-1}, i_\ell, i_{\ell+1}, \dots, i_k} \cdot m_{j, i_\ell} \right)_{i_1, \dots, i_{\ell-1}, j, i_{\ell+1}, \dots, i_k}, \quad (2.3)$$

přičemž

$$\mathcal{D} \in \mathbb{R}^{n_1 \times \dots \times n_{\ell-1} \times m \times n_{\ell+1} \times \dots \times n_k},$$

se nazývá *součinem tenzoru \mathcal{A} a matice M v ℓ -tém módu (anglicky ℓ -mode product)*. Tedy násobíme-li tenzor maticí v ℓ -tém módu, násobíme maticí každé vlákno ℓ -tého módu.

Poznámka 1 (Einsteinova notace). *Často se násobení píše zjednodušeně pomocí tzv. Einsteinovy notace. Pokud*

$$\mathcal{D} = \mathcal{A} \times_\ell M,$$

pak jednotlivé prvky tenzoru \mathcal{D} zapisujeme jako

$$d_{i_1, \dots, i_{\ell-1}, j, i_{\ell+1}, \dots, i_k} = a_{i_1, \dots, i_{\ell-1}, i_\ell, i_{\ell+1}, \dots, i_k} \cdot m_{j, i_\ell},$$

tedy vynecháváme symbol sumy a v každém součinnu sčítáme přes opakující se indexy. V tomto textu nebudeme Einsteinovu notaci používat.

Uvažujme součin tenzoru řádu k se dvěma maticemi ve dvou *různých* módech

$$\left(\mathcal{A} \times_{\ell} M\right) \times_s P, \quad M \in \mathbb{R}^{m \times n_{\ell}}, \quad P \in \mathbb{R}^{p \times n_s}, \quad \ell \neq s. \quad (2.4)$$

Všimněme si, že v každém násobení stojí vždy na levé straně tenzor a na pravé matice tak, jak jsme zavedli v definici 3. Při zachování priority násobení zleva doprava můžeme zjednodušeně psát

$$\left(\mathcal{A} \times_{\ell} M\right) \times_s P = \mathcal{A} \times_{\ell} M \times_s P. \quad (2.5)$$

Poznamenejme, že výsledný tenzor je z prostoru

$$\mathbb{R}^{n_1 \times \dots \times n_{\ell-1} \times m \times n_{\ell+1} \times \dots \times n_{s-1} \times p \times n_{s+1} \times \dots \times n_k}, \quad \text{je-li } \ell < s,$$

případně

$$\mathbb{R}^{n_1 \times \dots \times n_{s-1} \times p \times n_{s+1} \times \dots \times n_{\ell-1} \times m \times n_{\ell+1} \times \dots \times n_k}, \quad \text{je-li } s < \ell.$$

Násobení tenzoru dvěma maticemi ve *shodném* módu se bude věnovat věta 4 v následující podkapitole.

2.3 Rozvoj tenzoru v matici

Nyní se dostáváme k popisu velmi důležitého nástroje pro práci s tenzory, jehož podstatou je převést tenzor do podoby matice. Tento proces je v anglické literatuře pojmenováván *matricization*, *unfolding*, nebo *flattening*.

Definice 4 (Rozvoj tenzoru v matici (matricization)). *Rozvoj tenzoru*

$$\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$$

podle módu ℓ je matice

$$A^{[\ell]} \in \mathbb{R}^{n_{\ell} \times (n_1 n_2 \dots n_{\ell-1} n_{\ell+1} \dots n_k)}, \quad (2.6)$$

jejíž sloupce jsou vlákna ℓ -tého módu tenzoru \mathcal{A} indexovaná uspořádanou $(k-1)$ -ticí indexů, tzv. multiindexem $(i_1, \dots, i_{\ell-1}, i_{\ell+1}, \dots, i_k)$, a seřazená v lexikografickém pořadí.

Je zjevné, že pro tenzor k -tého řádu existuje k způsobů rozvoje tenzoru, přičemž matice, které získáme, mohou mít různé rozměry. Ukážeme si to na příkladu 6.

Příklad 6. Uvažujme tenzor $\mathcal{A} \in \mathbb{R}^{4 \times 3 \times 2}$ z příkladu 1 (viz také příklad 2). Potom je možný rozvoj tohoto tenzoru třemi způsoby

$$A^{[1]} = [a_{\cdot,1,1}, a_{\cdot,1,2}, a_{\cdot,2,1}, a_{\cdot,2,2}, a_{\cdot,3,1}, a_{\cdot,3,2}] = \begin{bmatrix} 6 & 6 & 6 & 4 & 2 & 1 \\ 7 & 3 & 1 & 3 & 0 & 4 \\ 7 & 9 & 7 & 7 & 0 & 4 \\ 3 & 0 & 0 & 7 & 8 & 6 \end{bmatrix},$$

$$A^{[2]} = \begin{bmatrix} 6 & 6 & 7 & 3 & 7 & 9 & 3 & 0 \\ 6 & 4 & 1 & 3 & 7 & 7 & 0 & 7 \\ 2 & 1 & 0 & 4 & 0 & 4 & 8 & 6 \end{bmatrix},$$

$$A^{[3]} = \begin{bmatrix} 6 & 6 & 2 & 7 & 1 & 0 & 7 & 7 & 0 & 3 & 0 & 8 \\ 6 & 4 & 1 & 3 & 3 & 4 & 9 & 7 & 4 & 0 & 7 & 6 \end{bmatrix}.$$

□

Poznámka 2. Rozvoj tenzoru v módu ℓ , tj. matici $A^{[\ell]}$, resp. jeho permutaci $A^{[\ell]}\Pi$, kde Π je permutační matice, lze také získat zřetěžením řezů v módu (ℓ, j) pro libovolné pevně zvolené j . Zřejmě matici $A^{[1]}$ z příkladu 6 můžeme zapsat pomocí řezů $(1, 2)$ nebo $(1, 3)$ takto

$$A^{[1]} = [A_{:, :, 1}, A_{:, :, 2}] \Pi = [A_{:, 1, :}, A_{:, 2, :}, A_{:, 3, :}],$$

kde

$$\Pi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

viz příklad 3.

Následující věta ukazuje, jak lze interpretovat násobení tenzoru maticí pomocí rozvoje tenzoru.

Věta 3 (Součin v ℓ -tém módu (viz definici 3)). Uvažujme tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$ a matici $M \in \mathbb{R}^{m \times n_\ell}$. Součin tenzoru \mathcal{A} a matice M je

$$\mathcal{D} = \mathcal{A} \times_\ell M. \quad (2.7)$$

Uvažujeme-li rozvoj tenzorů \mathcal{A} i \mathcal{D} v matice, můžeme součin zapsat

$$D^{[\ell]} = MA^{[\ell]}. \quad (2.8)$$

Důkaz věty je elementární, stačí si uvědomit, že při násobení tenzoru \mathcal{A} maticí M v ℓ -tém módu se touto maticí násobí všechna vlákna ℓ -tého módu; právě tato vlákna však tvoří sloupce matice $A^{[\ell]}$. Nyní můžeme vyšetřit chování součinu tenzoru se dvěma maticemi ve shodném módu.

Věta 4. Uvažujme tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$, matici $M \in \mathbb{R}^{m \times n_\ell}$ a matici $P \in \mathbb{R}^{p \times m}$. Platí

$$(\mathcal{A} \times_\ell M) \times_\ell P = \mathcal{A} \times_\ell (PM), \quad (2.9)$$

kde $PM \in \mathbb{R}^{p \times n_\ell}$.

Důkaz. Označme jednotlivé tenzory

$$\mathcal{D} = \mathcal{A} \times_{\ell} M \quad \text{a} \quad \mathcal{E} = (\mathcal{A} \times_{\ell} M) \times_{\ell} P = \mathcal{D} \times_{\ell} P.$$

Rozvinutím tenzorů do matic podle módu ℓ dostaneme rovnosti

$$D^{[\ell]} = MA^{[\ell]} \quad \text{a} \quad E^{[\ell]} = KD^{[\ell]},$$

viz vztahy (2.7) a (2.8) z věty 3. Prostým dosazením získáme rovnost

$$E^{[\ell]} = PD^{[\ell]} = PMA^{[\ell]}. \quad (2.10)$$

Zřejmě tedy platí

$$\mathcal{E} = \mathcal{A} \times_{\ell} (PM),$$

což jsme chtěli dokázat. □

Věty 3 a 4 ukázaly, jak lze násobení tenzoru maticí (2.7), případně dvěma maticemi ve shodném módu (2.9), přeformulovat pomocí maticového násobení (s využitím rozvoje tenzoru v matici), viz (2.8) a (2.10). Násobení tenzoru dvěma maticemi v různých módech (2.4) lze také přeformulovat pomocí maticového násobení, jak ukážeme později (2.20) v podkapitole 2.5.

2.4 Hodnost tenzoru

Důležitou charakteristikou tenzoru je jeho hodnost. Hodnost (rank) matice je definovaná jako počet nezávislých řádků, resp. sloupců, tedy pro tenzor řádu dva (matici) platí, že má stejnou hodnost ve všech módech, tj. sloupcích a řádcích. Definice hodnosti tenzorů vyšších řádů je však mnohem komplikovanější, věnujeme jí následující podkapitolu.

2.4.1 ℓ -rank a vektorový rank tenzoru

Pro určení hodnosti tenzoru vycházíme z jeho rozvoje v matici, zjišťujeme vlastně hodnosti těchto rozvojevů.

Definice 5 (ℓ -rank). *Nechť je $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$ a $A^{[\ell]}$ je rozvoj tenzoru \mathcal{A} v ℓ -tém módu. Potom hodnost matice $A^{[\ell]}$ nazýváme ℓ -rank tenzoru \mathcal{A} a značíme $\text{rank}_{\ell}(\mathcal{A})$.*

Definice 6 (Vektorový rank). *Označme $r_{\ell} = \text{rank}_{\ell}(\mathcal{A})$ pro $\ell = 1, \dots, k$. Uspořádanou k -tici (r_1, r_2, \dots, r_k) nazýváme vektorovým rankem tenzoru \mathcal{A} . Píšeme*

$$\text{rank}(\mathcal{A}) = (r_1, r_2, \dots, r_k).$$

V literatuře se místo pojmu vektorový rank můžeme setkat také např. s pojmem *multilineární rank*, viz [29], nebo *multirank*, viz [5, str. 321]. Pojem vektorový rank ukážeme opět na příkladu 1.

Příklad 7. Uvažujme tenzor $\mathcal{A} \in \mathbb{R}^{4 \times 3 \times 2}$ z příkladu 1. Dále uvažujme jeho rozvoje z příkladu 6. Zřejmě

$$\text{rank}(A^{[1]}) = 4, \quad \text{rank}(A^{[2]}) = 3, \quad \text{rank}(A^{[3]}) = 2.$$

Tedy

$$\text{rank}(\mathcal{A}) = (4, 3, 2).$$

□

Ze základního kurzu lineární algebry víme, že matice (které lze považovat za tenzory druhého řádu) mají vždy stejný počet lineárně nezávislých řádků a sloupců. Vidíme, že tuto vlastnost nelze zobecnit pro tenzory vyšších řádů.

Počet lineárně nezávislých vláken tenzoru může být (narozdíl od matic) různý v různých módech. Proto zavádíme pojem *vektorový rank tenzoru*.

Tento jev lze ilustrovat i na jednodušších příkladech.

Příklad 8. Uvažujme tenzor

$$\mathcal{A} = \begin{array}{|c|c|c|c|} \hline & 1 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ \hline \end{array} \in \mathbb{R}^{2 \times 2 \times 2}. \quad (2.11)$$

Rozvoje tenzoru jsou matice $A^{[1]}, A^{[2]}, A^{[3]} \in \mathbb{R}^{2 \times 4}$ jsou

$$\begin{aligned} A^{[1]} &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, & \text{rank}(A^{[1]}) &= 2, \\ A^{[2]} &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, & \text{rank}(A^{[2]}) &= 2, \\ A^{[3]} &= \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, & \text{rank}(A^{[3]}) &= 1. \end{aligned}$$

Tenzor \mathcal{A} má tedy vektorový rank $(2, 2, 1)$. □

Hodnost tenzoru lze ale také zavést jiným způsobem, např. analogicky k (1.5) pomocí tzv. *polyadického rozvoje* tenzoru. V této práci budeme nicméně hodnotí tenzoru rozumět výhradně vektorový rank, pokud nebude explicitně uvedeno jinak.

2.4.2 Vnější součin tenzorů, polyadický rozvoj a polyadický rank tenzoru

Podobně jako dyadickým rozvojem vyjadřujeme matici jako *součet vnějších součinů vektorů*, můžeme to samé udělat i s tenzory. Nejprve definujeme tzv. *vnější součin tenzorů* a *rank-one tenzor*.

Definice 7 (Vnější součin tenzorů). *Nechť $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}$ s prvky a_{i_1, i_2, \dots, i_k} je tenzor řádu k a $\mathcal{B} \in \mathbb{R}^{m_1 \times \dots \times m_t}$ s prvky b_{i_1, i_2, \dots, i_t} je tenzor řádu s . Pak tenzor*

$$\mathcal{F} \in \mathbb{R}^{n_1 \times \dots \times n_k \times m_1 \times \dots \times m_t}$$

řádu $(k + s)$ s prvky

$$f_{i_1, \dots, i_k, j_1, \dots, j_t} = a_{i_1, \dots, i_k} \cdot b_{j_1, \dots, j_t} \quad (2.12)$$

nazýváme *vnější součin tenzorů \mathcal{A} a \mathcal{B}* , značíme $\mathcal{F} = \mathcal{A} \circ \mathcal{B}$.

Poznamenejme, že v některých textech se tento součin označuje jako tenzorový součin, viz např. [38, str. 56]; my budeme pojem tenzorového součinu používat v jiném kontextu (viz podkapitulu 2.6).

Vnější součin je tedy typem součinu „každý s každým“; každý prvek tenzoru \mathcal{A} je násoben každým prvkem tenzoru \mathcal{B} . Vnější součin je zřejmě asociativní $(\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C} = \mathcal{A} \circ (\mathcal{B} \circ \mathcal{C})$, není tedy třeba používat závorky. V dalším textu, budeme-li vnější součin potřebovat, budeme psát prostě $\mathcal{A} \circ \mathcal{B} \circ \mathcal{C}$. Vnější součin tenzorů však není komutativní. Tenzory $\mathcal{A} \circ \mathcal{B}$ a $\mathcal{B} \circ \mathcal{A}$ se liší permutací indexů. Zřejmě $\mathcal{A} \circ \mathcal{B} = (f_{i_1, \dots, i_k, j_1, \dots, j_t}) \dots$, ale $\mathcal{B} \circ \mathcal{A} = (f_{j_1, \dots, j_t, i_1, \dots, i_k}) \dots$.

Definice 8 (Rank-one tenzor). *Mějme k nenulových vektorů*

$$a_1 \in \mathbb{R}^{n_1}, \dots, a_k \in \mathbb{R}^{n_k}.$$

Tenzor k -tého řádu

$$\mathcal{A} = a_1 \circ a_2 \circ \dots \circ a_k \in \mathbb{R}^{n_1 \times \dots \times n_k} \quad (2.13)$$

nazýváme *rank-one tenzor*.

Vektorový rank rank-one tenzoru řádu k je zřejmě $(1, 1, \dots, 1)$.

Jakýkoli tenzor \mathcal{A} je možné vyjádřit ve tvaru *polyadického rozvoje*, tj. jako součet rank-one tenzorů. Například pro tenzor řádu tři $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ platí

$$\mathcal{A} = \sum_{\ell=1}^r a_\ell \circ b_\ell \circ c_\ell, \quad a_\ell \in \mathbb{R}^{n_1}, b_\ell \in \mathbb{R}^{n_2}, c_\ell \in \mathbb{R}^{n_3}, \quad (2.14)$$

pro nějaké $r \leq n_1 n_2 n_3$ (analogicky platí pro tenzor k -tého řádu). Pro tenzor řádu tři je polyadický rozvoj znázorněn na obrázku 2.3.

S myšlenkou vyjádření tenzoru ve formě polyadického rozvoje přišel jako první F. L. Hitchcock, viz [22], [23]. Později představili tuto formu i další vědci, např. J. D. Carrol and J. J. Chang zavedli ve článku [8] termín *CanDecomp* (*canonical decomposition*), R. A. Harshman v [21] používá *ParaFac* (*parallel factors*), H. A. L. Kiers v [26] zavádí *CP* jako *CanDecomp-ParaFac decomposition*. Pro přehled názvů polyadického rozvoje včetně jejich datace viz [28, str. 462–463]. Hodnost tenzoru potom lze definovat jako nejmenší možné číslo r ze vztahu (2.14), tedy nejmenší počet sčítanců polyadického rozvoje (rank-one tenzorů), pomocí kterých lze tenzor zapsat, viz [28, str. 464] nebo [41, str. 18]. Hodnost zavedenou takovýmto způsobem budeme nazývat *polyadický rank*.

2.4.3 Norma tenzoru a aproximace tenzoru tenzorem nižšího polyadického ranku

Polyadický rozvoj ale vede k tomu, že některé vlastnosti matic lze obtížně zobecnit na tenzory vyšších řádů. Jedním z takových problémů je aproximace matice maticí nižší hodnosti. Pro další výklad se nám bude hodit zavést pojem *norma tenzoru*. Nejjednodušším způsobem můžeme normu tenzoru zavést zobecněním eukleidovské normy vektoru, resp. Frobeniovy normy matice.

Definice 9 (Norma tenzoru). *Nechť $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}$, potom*

$$\|\mathcal{A}\| = \sqrt{\sum_{i_1=1}^{n_1} \dots \sum_{i_k=1}^{n_k} a_{i_1 \dots i_k}^2}, \quad (2.15)$$

tzn. norma tenzoru je druhá odmocnina součtu kvadrátů všech prvků.

U tenzorů lze najít příklady, kdy tenzor lze aproximovat tenzorem nižšího polyadického ranku s libovolně malou chybou. Ukážeme příklad uvedený v článku [28], původně z článku [34]. Nechť mají matice

$$[a_1, a_2] \in \mathbb{R}^{n_1 \times 2}, \quad [b_1, b_2] \in \mathbb{R}^{n_2 \times 2}, \quad [c_1, c_2] \in \mathbb{R}^{n_3 \times 2}$$

lineárně nezávislé sloupce, pak tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$

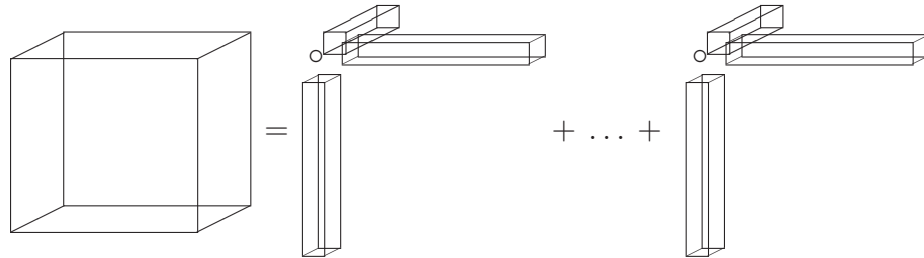
$$\mathcal{A} = a_1 \circ b_1 \circ c_2 + a_1 \circ b_2 \circ c_1 + a_2 \circ b_1 \circ c_1$$

má polyadický rank tři, viz [28]. Tento tenzor můžeme libovolně přesně aproximovat tenzorem

$$\mathcal{B}(\varepsilon) = \frac{1}{\varepsilon} (a_1 + \varepsilon a_2) \circ (b_1 + \varepsilon b_2) \circ (c_1 + \varepsilon c_2) - \frac{1}{\varepsilon} a_1 \circ b_1 \circ c_1$$

s polyadickým rankem dva, kde ε je libovolně malé kladné číslo. Potom

$$\|\mathcal{A} - \mathcal{B}(\varepsilon)\| = \varepsilon \|a_2 \circ b_2 \circ c_1 + a_2 \circ b_1 \circ c_2 + a_1 \circ b_2 \circ c_2 + \varepsilon a_2 \circ b_2 \circ c_2\|,$$



Obrázek 2.3: Polyadický rozvoj (ParaFac, CanDecomp, CP) tenzoru řádu 3. Převzato z [28].

a tedy v limitě platí

$$\lim_{\varepsilon \rightarrow 0} \|\mathcal{A} - \mathcal{B}(\varepsilon)\| = 0,$$

přičemž \mathcal{A} a $\mathcal{B}(\varepsilon)$ jsou tenzory s různým polyadickým rankem.

Vidíme, že tvrzení z důsledku 2 (viz str. 23), které platilo pro matice, nelze zobecnit do prostoru tenzorů řádu tři a více, pracujeme-li s polyadickým rankem.

2.5 Zobecněný rozvoj tenzoru v matici, vektorizace a Kroneckerův součin matic

Rozvoj tenzoru v matici podle módu ℓ , tak jak je zavedený podkapitole 2.3, lze zobecnit na rozvoj podle více než jednoho módu. Mějme tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}$. Množinu módů tohoto tenzoru $\{1, \dots, k\}$ rozdělíme do dvou disjunktních podmnožin

$$\{1, \dots, k\} = \mathcal{T} \cup \mathcal{S}, \quad \mathcal{T} \cap \mathcal{S} = \emptyset. \quad (2.16)$$

Označme jejich prvky

$$\mathcal{T} = \{t_1, \dots, t_d\} \quad \text{a} \quad \mathcal{S} = \{s_1, \dots, s_{k-d}\}. \quad (2.17)$$

U rozvoje tenzoru \mathcal{A} podle ℓ -tého módu byla množina indexů (multiindex) sloužící k indexování sloupců matice $A^{[\ell]}$ vždy uspořádaná, přičemž toto uspořádání bylo dáno uspořádáním indexů původního tenzoru, viz definici 4. Stejně tak i zde budeme striktně předpokládat

$$t_1 < t_2 < \dots < t_d \quad \text{a} \quad s_1 < s_2 < \dots < s_{k-d}. \quad (2.18)$$

Seřadíme-li prvky původního tenzoru \mathcal{A} do matice tak, že jako řádkové indexy budeme volit multiindexy $(i_{t_1}, \dots, i_{t_d})$ a jako sloupcové indexy budeme volit multiindexy $(i_{s_1}, \dots, i_{s_{k-d}})$, obojí seřazené v lexikografickém pořadí, získáme tzv. zobecněný rozvoj tenzoru v matici, viz následující definici.

Definice 10 (Zobecněný rozvoj tenzoru v matici). *Uvažujme obecný tenzor $\mathcal{A} = (a_{i_1, \dots, i_k}) \dots \in \mathbb{R}^{n_1 \times \dots \times n_k}$ a indexové množiny \mathcal{T} a \mathcal{S} splňující (2.16)–(2.18). Matici*

$$A^{[t_1, \dots, t_d]} \in \mathbb{R}^{(n_{t_1} n_{t_2} \dots n_{t_d}) \times (n_{s_1} n_{s_2} \dots n_{s_{k-d}})},$$

která obsahuje prvek a_{i_1, \dots, i_k} tenzoru \mathcal{A} na řádce s multiindexem $(i_{t_1}, \dots, i_{t_d})$ a ve sloupci multiindexem $(i_{s_1}, \dots, i_{s_{k-d}})$ nazveme zobecněným rozvojem tenzoru \mathcal{A} v matici (matricization) podle módů (t_1, \dots, t_d) .

Pro zobecněný rozvoj tenzoru v matici zřejmě platí

$$(A^{[t_1, \dots, t_d]})^T = A^{[s_1, \dots, s_{k-d}]} \quad (2.19)$$

Rozvoj tenzoru v matici módu ℓ z definice 4, viz (2.6), je zřejmě speciální případ, kdy množina řádkových indexů obsahuje pouze jeden prvek, tedy když $\mathcal{T} = \{\ell\}$, $\mathcal{S} = \{1, \dots, \ell - 1, \ell + 1, \dots, k\}$.

Další speciální případ nastane, když je jedna z množin prázdná, tedy když např. $\mathcal{T} = \{1, \dots, k\}$, $\mathcal{S} = \emptyset$. Pak zřejmě

$$A^{[1, \dots, k]} = (A^{[1]})^T \in \mathbb{R}^{(n_1 n_2 \dots n_k) \times 1}.$$

Výsledkem je tedy (sloupcový) vektor, který se zpravidla značí

$$\text{vec}(\mathcal{A})$$

a nazývá se *vektORIZACÍ* tenzoru, viz [28, str. 460] (případně vektorizací matice, jednalo-li se o tenzor druhého řádu, viz [29, str. 667]).

Na následujícím příkladu si ukážeme další využití zobecněného rozvoje.

Příklad 9. Uvažujme následující tenzor čtvrtého řádu, který je vnějším součinem dvou tenzorů druhého řádu (tedy matice),

$$\mathcal{T} = A \circ B,$$

kde

$$A \in \mathbb{R}^{n \times m}, \quad B \in \mathbb{R}^{u \times v}.$$

Označme a_{i_1, i_2} prvky matice A a b_{j_1, j_2} prvky matice B . Tedy $\mathcal{T} \in \mathbb{R}^{n \times m \times u \times v}$.

Ukážeme si, jak vypadá rozvoj tenzoru podle módů (1, 3), analogicky lze ukázat pro ostatní módy. Tímto rozvojem vznikne matice

$$T^{[1,3]} \in \mathbb{R}^{nu \times mv}.$$

Řádky této matice jsou indexovány multiindexem (i_1, j_1) a sloupce multiindexem (i_2, j_2) , tj.

$$T^{[1,3]} = \begin{bmatrix} a_{1,1}b_{1,1} & \dots & a_{1,1}b_{1,v} & & a_{1,m}b_{1,1} & \dots & a_{1,m}b_{1,v} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ a_{1,1}b_{u,1} & \dots & a_{1,1}b_{u,v} & & a_{1,m}b_{u,1} & \dots & a_{1,m}b_{u,v} \\ & & \vdots & \ddots & & & \vdots \\ a_{n,1}b_{1,1} & \dots & a_{n,1}b_{1,v} & & a_{n,m}b_{1,1} & \dots & a_{n,m}b_{1,v} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ a_{n,1}b_{u,1} & \dots & a_{n,1}b_{u,v} & & a_{n,m}b_{u,1} & \dots & a_{n,m}b_{u,v} \end{bmatrix}.$$

□

Jelikož tenzor \mathcal{T} z předchozího příkladu bylo možné vyjádřit jako vnější součin dvou tenzorů druhého řádu (matic), můžeme jeho rozvoj vyjádřit pomocí tzv. *Kroneckerova součinu* těchto matic, viz následující definici.

Definice 11 (Kroneckerův součin). *Mějme matice $A \in \mathbb{R}^{n \times m}$ a $B \in \mathbb{R}^{u \times v}$. Potom matici*

$$K = \begin{bmatrix} a_{1,1}B & \dots & a_{1,m}B \\ \vdots & & \vdots \\ a_{n,1}B & \dots & a_{n,m}B \end{bmatrix},$$

kteřou získáme tak, že (i_1, i_2) -tý prvek matice A nahradíme blokem $a_{i_1, i_2}B$, nazýváme *Kroneckerovým součinem* matic A a B , viz [46]. Tento součin značíme

$$K = A \otimes B$$

a jeho výsledkem je tedy matice o rozměrech $nu \times mv$.

Vidíme, že Kroneckerův součin matic je speciálním případem vnějšího součinu tenzorů. Platí pro něj tedy také asociativní zákon, tj. $(A \otimes B) \otimes C = A \otimes (B \otimes C) = A \otimes B \otimes C$. Kroneckerův součin ale není komutativní, výsledky součinů $A \otimes B$ a $B \otimes A$ se liší permutací řádků a sloupců. V tomto kontextu je vhodné poznamenat, že v některých textech se definice Kroneckerova součinu liší pořadím operandů. Příkladem může být dnes již klasická Fiedlerova učebnice [16, str. 122], kde je navíc Kroneckerův součin nazýván tenzorovým součinem matic; my budeme pojem tenzorového součinu používat v jiném kontextu (viz podkapitolu 2.6).

Příklad 10. *Ukážeme ještě příklad s konkrétními čísly. Mějme dva tenzory druhého řádu (matice) A a B , kde*

$$A = \begin{bmatrix} 2 & 3 & 5 \\ 7 & 11 & 13 \end{bmatrix} \in \mathbb{R}^{2 \times 3}, \quad B = \begin{bmatrix} 17 & 19 \\ 23 & 29 \\ 31 & 37 \\ 41 & 43 \end{bmatrix} \in \mathbb{R}^{4 \times 2}.$$

Vnějším součinem těchto tenzorů je tenzor čtvrtého řádu

$$\mathcal{T} = A \circ B \in \mathbb{R}^{2 \times 3 \times 4 \times 2}.$$

Uvažujme rozvoj tenzoru \mathcal{T} v módech $(1, 3)$.

$$T^{[1,3]} = \left[\begin{array}{cc|cc|cc} 34 & 38 & 51 & 57 & 85 & 95 \\ 46 & 58 & 69 & 87 & 115 & 145 \\ 62 & 74 & 93 & 111 & 155 & 185 \\ 82 & 86 & 123 & 129 & 205 & 215 \\ \hline 119 & 133 & 187 & 209 & 221 & 247 \\ 161 & 203 & 253 & 319 & 299 & 377 \\ 217 & 259 & 341 & 407 & 403 & 481 \\ 287 & 301 & 451 & 473 & 533 & 559 \end{array} \right] \in \mathbb{R}^{8 \times 6}.$$

Můžeme lehce ověřit, že platí rovnost

$$T^{[1,3]} = A \otimes B = \begin{bmatrix} 2B & 3B & 5B \\ 7B & 11B & 13B \end{bmatrix}.$$

□

V podkapitole 2.2, definici 3 jsme ukázali, jak lze násobit tenzor maticí v daném módu ℓ a následně v podkapitole 2.3, větě 3 jsme ukázali, jak je možné toto násobení zapsat pomocí rozvoje tenzoru v daném módu. Nyní si ukážeme, jak lze pomocí zobecněného rozvoje vyjádřit násobení tenzoru maticemi ve více různých módech.

Věta 5. *Mějme tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}$ a matice $M \in \mathbb{R}^{m \times n_\ell}$ a $P \in \mathbb{R}^{p \times n_s}$, přičemž pro čísla ℓ a s platí $1 \leq \ell < s \leq k$, a součin*

$$\mathcal{D} = \mathcal{A} \times_\ell M \times_s P \in \mathbb{R}^{n_1 \times \dots \times n_{\ell-1} \times m \times n_{\ell+1} \times \dots \times n_{s-1} \times p \times n_{s+1} \times \dots \times n_k}.$$

Pro zobecněné rozvoje tenzorů \mathcal{A} a \mathcal{D} v matice v módech (ℓ, s) platí

$$D^{[\ell,s]} = (M \otimes P) A^{[\ell,s]} \in \mathbb{R}^{(mp) \times (n_1 \dots n_{\ell-1} n_{\ell+1} \dots n_{s-1} n_{s+1} \dots n_k)}. \quad (2.20)$$

Důkaz věty je opět elementární. Stačí si uvědomit, jak jsou seřazeny prvky původních matic v jejich Kroneckerově součinu a jak jsou seřazeny prvky tenzoru v příslušných zobecněných rozvojiích.

Poznámka 3. *V lineární algebře se vyskytuje řada rovností zahrnujících Kroneckerův součin. Zde zmíníme jen několik nejdůležitějších resp. nejzajímavějších bez uvádění předpokladů, za kterých platí (a které jsou zřejmé). Platí např.*

$$A^T \otimes B^T = (A \otimes B)^T,$$

$$A^{-1} \otimes B^{-1} = (A \otimes B)^{-1}.$$

Některé maticové součiny můžeme přepsat s využitím Kroneckerových součinů a vektorizace matice. Toho lze využít např. pro řešení maticových rovnic. Platí např.

$$ABC = D \iff (C^T \otimes A)\text{vec}(B) = \text{vec}(D),$$

$$(A \otimes B)(C \otimes D) = (AC \otimes BD)$$

a mnoho dalších; viz např. [16], [46].

Poznámka 4. *Existují další součiny matic, které mají blízký vztah ke Kroneckerovu součinu a mohou být interpretovány jazykem tenzorů. Například tzv. Khatri–Rao product, viz [25], [32]; tzv. Tracy–Singh product, viz [42], [32]; nebo tzv. cross-product, viz [46].*

2.6 Tensorový součin a vztahy mezi jednotlivými součiny

V této podkapitole se budeme zabývat násobením tenzorů. Ukážeme, jak lze součin tenzoru a matice, který už známe z podkapitoly 2.2, zobecnit na součin dvou tenzorů. Někdy je takový součin nazýván také *úžení tenzorů*, viz [31, str. 70], *contracted product*, viz [2, str. 643] nebo také *tensor-tensor contraction*, viz [41, str. 31]. Násobení tenzorů můžeme provést, podobně jako v případě součinu tenzoru s maticí, v módech shodné délky (dimenze). V případě součinu tenzorů vyšších řádů můžeme násobení provádět ve více módech současně.

Definice 12 (Tensorový součin). *Mějme tenzory*

$$\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}, \quad \mathcal{B} \in \mathbb{R}^{m_1 \times \dots \times m_t}, \quad \text{přičemž } n_\ell = m_s \equiv \nu$$

pro nějaké $\ell, 1 \leq \ell \leq k$, a $s, 1 \leq s \leq t$. V tomto případě tedy můžeme tenzory \mathcal{A} a \mathcal{B} násobit v módech ℓ a s . Výsledkem bude tenzor řádu $k + t - 2$

$$\mathcal{F} = \mathcal{A} \times_{\ell,s} \mathcal{B} = \left(\sum_{\alpha=1}^{\nu} a_{i_1, \dots, i_{\ell-1}, \alpha, i_{\ell+1}, \dots, i_k} \cdot b_{j_1, \dots, j_{s-1}, \alpha, j_{s+1}, \dots, j_t} \right)_{i_1, \dots, i_{\ell-1}, i_{\ell+1}, \dots, i_k, j_1, \dots, j_{s-1}, j_{s+1}, \dots, j_t}, \quad (2.21)$$

kde

$$\mathcal{F} \in \mathbb{R}^{n_1 \times \dots \times n_{\ell-1} \times n_{\ell+1} \times \dots \times n_k \times m_1 \times \dots \times m_{s-1} \times m_{s+1} \times \dots \times m_t}.$$

Poznamenejme, že v jiných textech se pod pojmem tensorový součin může skrývat jiný význam, viz [38, str. 56], kde má význam vnějšího součinu tenzorů, nebo [16, str. 122], kde má význam Kroneckerova součinu matic; my používáme pojem tensorového součinu pro konzistentní rozšíření maticového součinu. Tensorový součin lze opět zapsat pomocí rozvoje tenzorů v matici; srovnej následující větu s větou 3.

Věta 6. *Mějme tenzory*

$$\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}, \quad \mathcal{B} \in \mathbb{R}^{m_1 \times \dots \times m_t}, \quad \text{přičemž } n_\ell = m_s$$

pro nějaké $\ell, 1 \leq \ell \leq k$, a $s, 1 \leq s \leq t$, a tenzor

$$\mathcal{F} = \mathcal{A} \times_{\ell,s} \mathcal{B} \in \mathbb{R}^{n_1 \times \dots \times n_{\ell-1} \times n_{\ell+1} \times \dots \times n_k \times m_1 \times \dots \times m_{s-1} \times m_{s+1} \times \dots \times m_t}.$$

Pak platí

$$F^{[1, \dots, \ell-1, \ell+1, \dots, k]} = (A^{[\ell]})^T B^{[s]}.$$

Důkaz. Důkaz věty je opět elementární, stačí si uvědomit, že při násobení tenzoru \mathcal{A} tenzorem \mathcal{B} se provádí skalární součiny vláken ℓ -tého módu tenzoru \mathcal{A} s vlákny s -tého módu tenzoru \mathcal{B} . Tato vlákna jsou sloupce matic $A^{[\ell]}$ a $B^{[s]}$ a jsou indexována multiindexy $(i_1, \dots, i_{\ell-1}, i_{\ell+1}, \dots, i_k)$, resp. $(j_1, \dots, j_{s-1}, j_{s+1}, \dots, j_t)$ a seřazena lexicograficky stejně jako řádky, resp. sloupce matice $F^{[1, \dots, \ell-1, \ell+1, \dots, k]}$; viz definice 4 a 10; viz také (2.19). \square

Nyní se ještě vrátíme k násobení tenzoru maticí v ℓ -tém módu. Matici můžeme chápat jako tenzor druhého řádu a tak zřejmě násobení, které jsme zavedli definicí 3 v podkapitole 2.2, se od obecného násobení tenzorů, které jsme popsali nyní, liší pouze permutací indexů výsledného tenzoru. Pro tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_\ell \times \dots \times n_k}$ a matici $M \in \mathbb{R}^{m \times n_\ell}$ tedy platí

$$(\mathcal{A} \times_\ell M)^\Pi = \mathcal{A} \times_{\ell,2} M,$$

kde

$$\Pi = \begin{pmatrix} 1 & \dots & k \\ 1, \dots, \ell-1, \ell+1, \dots, k, & \ell \end{pmatrix}.$$

V definici 7 podkapitoly 2.4.2 jsme zavedli vnější součin dvou tenzorů. Ten lze také zapsat pomocí klasického tenzorového součinu. Uvažujme tenzory

$$\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k} \quad \text{a} \quad \mathcal{B} \in \mathbb{R}^{m_1 \times \dots \times m_t}.$$

Podobně jako lze vektor interpretovat jako matici s jediným sloupcem, nebo např. matici jako tenzor třetího řádu s jediným frontálním řezem, lze i tenzory \mathcal{A} a \mathcal{B} řádů k a t interpretovat jako tenzory řádu $k+1$, resp. $t+1$,

$$\mathcal{A}' \in \mathbb{R}^{n_1 \times \dots \times n_k \times n_{k+1}}, \quad \mathcal{B}' \in \mathbb{R}^{m_1 \times \dots \times m_t \times m_{t+1}}, \quad \text{splňující} \quad n_{k+1} = m_{t+1} = 1.$$

Zřejmě platí

$$\mathcal{F} = \mathcal{A} \circ \mathcal{B} = \mathcal{A}' \times_{k+1, t+1} \mathcal{B}'. \quad (2.22)$$

Toto pozorování není překvapivé. Ze základního kurzu lineární algebry víme, že vnější součin dvou vektorů a a b (zde zobecněných na tenzory) lze zapsat pomocí klasického maticového násobení jako ab^T (zde zobecněného na tenzorový součin).

Na závěr ještě ukážeme, jak vypadá tenzorový součin, pokud násobíme ve více módech současně. Mějme tenzory

$$\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}, \quad \mathcal{B} \in \mathbb{R}^{m_1 \times \dots \times m_t}, \quad n_{\ell_1} = m_{s_1} \equiv \nu, \quad n_{\ell_2} = m_{s_2} \equiv \mu$$

kde např. $1 \leq \ell_1 < \ell_2 \leq k$ a $1 \leq s_1 < s_2 \leq t$. Součinem tenzorů \mathcal{A} a \mathcal{B} ve dvojici módů (ℓ_1, ℓ_2) a (s_1, s_2) je pak tenzor

$$\mathcal{F} = \mathcal{A} \times_{(\ell_1, \ell_2), (s_1, s_2)} \mathcal{B} = \left(\sum_{\alpha=1}^{\nu} \sum_{\beta=1}^{\mu} a_{i_1, \dots, i_{\ell_1-1}, \alpha, i_{\ell_1+1}, \dots, i_{\ell_2-1}, \beta, i_{\ell_2+1}, \dots, i_k} \cdot b_{j_1, \dots, j_{s_1-1}, \alpha, j_{s_1+1}, \dots, j_{s_2-1}, \beta, j_{s_2+1}, \dots, j_t} \right) \dots \quad (2.23)$$

řádu $k+t-4$. Pro zobecněné rozvoje tenzorů \mathcal{A} , \mathcal{B} a \mathcal{F} platí

$$F^{[1, \dots, \ell_1-1, \ell_1+1, \dots, \ell_2-1, \ell_2+1, \dots, k]} = (A^{[\ell_1, \ell_2]})^T (B^{[s_1, s_2]}).$$

Poznamenejme, že je-li naopak $\ell_1 < \ell_2$ a $s_2 < s_1$, jsou role indexů s_1 a s_2 přehozené; píšeme pak

$$\mathcal{F} = \mathcal{A} \times_{(\ell_1, \ell_2), (s_2, s_1)} \mathcal{B} = \left(\sum_{\alpha=1}^{\nu} \sum_{\beta=1}^{\mu} a_{i_1, \dots, i_{\ell_1-1}, \alpha, i_{\ell_1+1}, \dots, i_{\ell_2-1}, \beta, i_{\ell_2+1}, \dots, i_k} \cdot b_{j_1, \dots, j_{s_2-1}, \beta, j_{s_2+1}, \dots, j_{s_1-1}, \alpha, j_{s_1+1}, \dots, j_t} \right) \dots$$

a platí

$$F^{[1, \dots, \ell_1-1, \ell_1+1, \dots, \ell_2-1, \ell_2+1, \dots, k]} = (A^{[\ell_1, \ell_2]})^T (B^{[s_2, s_1]}).$$

Analogicky lze tenzory násobit v libovolném q -tici módů (ℓ_1, \dots, ℓ_q) a (s_1, \dots, s_q) pokud $n_{\ell_j} = m_{s_j}$, $j = 1, \dots, q$, kde $q \leq \min\{k, t\}$.

Poznámka 5 (Einsteinova notace). *Pro zápis tenzorového součinu, např. (2.21) nebo (2.23), lze opět použít Einsteinovy notace. Pokud $\mathcal{F} = \mathcal{A} \times_{\ell, s} \mathcal{B}$, pak jednotlivé prvky tenzoru \mathcal{F} zapisujeme jako*

$$f_{i_1, \dots, i_{\ell-1}, i_{\ell+1}, \dots, i_k, j_1, \dots, j_{s-1}, j_{s+1}, \dots, j_t} = a_{i_1, \dots, i_{\ell-1}, \alpha, i_{\ell+1}, \dots, i_k} \cdot b_{j_1, \dots, j_{s-1}, \alpha, j_{s+1}, \dots, j_t}.$$

Pokud $\mathcal{F} = \mathcal{A} \times_{(\ell_1, \ell_2), (s_1, s_2)} \mathcal{B}$, pak jednotlivé prvky tenzoru \mathcal{F} zapisujeme jako

$$f_{i_1, \dots, i_{\ell_1-1}, i_{\ell_1+1}, \dots, i_{\ell_2-1}, i_{\ell_2+1}, \dots, i_k, j_1, \dots, j_{s_1-1}, j_{s_1+1}, \dots, j_{s_2-1}, j_{s_2+1}, \dots, j_t} = \\ a_{i_1, \dots, i_{\ell_1-1}, \alpha, i_{\ell_1+1}, \dots, i_{\ell_2-1}, \beta, i_{\ell_2+1}, \dots, i_k} \cdot b_{j_1, \dots, j_{s_1-1}, \alpha, j_{s_1+1}, \dots, j_{s_2-1}, \beta, j_{s_2+1}, \dots, j_t}.$$

Opět tedy vynecháváme všechny sumy a v součinu sčítáme přes všechny opakující se indexy.

Všimněme si, že přepsáním vnějšího součinu $\mathcal{F} = \mathcal{A} \circ \mathcal{B}$ pomocí vztahu (2.22) získáme, s využitím Einsteinovy notace,

$$f_{i_1, \dots, i_k, j_1, \dots, j_t} = a_{i_1, \dots, i_k, \alpha} \cdot b_{j_1, \dots, j_t, \alpha},$$

přičemž sčítáme přes opakující se index α , ovšem $\alpha \in \{1\}$ (součet má jediný sčítanec). Zřejmě tedy platí

$$f_{i_1, \dots, i_k, j_1, \dots, j_t} = a_{i_1, \dots, i_k} \cdot b_{j_1, \dots, j_t},$$

čímž dostáváme přímo vztah (2.12) z definice vnějšího součinu. Vidíme tedy, že Einsteinova notace nám umožňuje snadno zapsat tenzorový i vnější součin tenzorů jedním způsobem (v předchozím vzorci se žádný index neopakuje, tedy neprovádíme žádné sčítání).

Poznámka 6. *Přístupujeme-li k tenzorům jako k funkcím více proměnných, můžeme tenzorový součin, např. (2.21), $\mathcal{F} = \mathcal{A} \times_{\ell, s} \mathcal{B}$, zapsat také s využitím funkcí. Odpovídá-li tenzor \mathcal{A} funkci $a(x_1, \dots, x_k)$, viz (2.2), a tenzor \mathcal{B} funkci $b(y_1, \dots, y_t)$, pak tenzor \mathcal{F} odpovídá funkci*

$$f(x_1, \dots, x_{\ell-1}, x_{\ell+1}, \dots, x_k, y_1, \dots, y_{s-1}, y_{s+1}, \dots, y_t) \equiv \\ \int a(x_1, \dots, x_{\ell-1}, \alpha, x_{\ell+1}, \dots, x_k) b(y_1, \dots, y_{s-1}, \alpha, y_{s+1}, \dots, y_t) d\alpha.$$

Využijeme-li navíc Einsteinovy notace, je dále možné vynechat symbol integrálu, viz např. [47].

3 Tuckerův rozklad

Z lineární algebry už známe různé maticové rozklady, které nám ukazují některé vlastnosti matic nebo nám umožňují získat matici ve tvaru vhodném pro určité výpočty. Podobně tomu je i u tenzorů. Je zřejmé, že početní nároky budou oproti maticím mnohem větší. Přesto ale uvidíme, že vyjádření tenzoru ve tvaru zobecněného ekonomického SVD (1.4) pro tenzor může být velmi výhodné.

3.1 Odvození Tuckerova rozkladu

V této podkapitole zobecníme SVD, který jsme popsali u matic, na tenzory. Takovému rozkladu říkáme *Tuckerův rozklad* (podle L. R. Tuckera, který jej poprvé popsal v [43]), tento termín je používán např. v článcích [28], [17]. Čím dál častěji je tento rozklad nazýván *singulárním rozkladem vyššího řádu*, anglicky *high-order SVD* (HOSVD). Tento pojem zavedli v článku [10] L. De Lathauwer, B. De Moor a J. Vandewalle. Existují i další názvy, viz např. [28, str. 474].

Při výkladu budeme vycházet z příkladu trojrozměrného tenzoru

$$\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

jehož rozvoje (matricization) označíme $A^{[\ell]}$. Singulární rozklady matic získaných pomocí matricization jsou

$$A^{[\ell]} = U^{(\ell)} \Sigma^{(\ell)} (V^{(\ell)})^T, \quad \text{rank}(A^{[\ell]}) = r_\ell, \quad (3.1)$$

kde matice

$$U^{(\ell)} = [U_A^{(\ell)}, U_\perp^{(\ell)}], \quad \Sigma^{(\ell)} = \begin{bmatrix} \Sigma_A^{(\ell)} & 0 \\ 0 & 0 \end{bmatrix}, \quad V^{(\ell)} = [V_A^{(\ell)}, V_\perp^{(\ell)}].$$

Kulaté závorky v horním indexu u matic $U^{(\ell)}$, $\Sigma^{(\ell)}$, $V^{(\ell)}$ používáme proto, aby bylo jasné, že se narázdíl od $A^{[\ell]}$ nejedná o rozvoj (matricization) nějakého tenzoru. Rozklad z (3.1) pak můžeme psát v ekonomickém tvaru SVD

$$A^{[\ell]} = U_A^{(\ell)} \Sigma_A^{(\ell)} (V_A^{(\ell)})^T,$$

viz (1.4). Dále označíme \mathcal{W} součin

$$\mathcal{W} = \mathcal{A} \times_\ell (U^{(\ell)})^T,$$

viz (2.3). Pokud tenzory na obou stranách rozvineme do matice podle módu ℓ , dostaneme

$$W^{[\ell]} = (U^{(\ell)})^T A^{[\ell]} \quad (3.2)$$

viz (2.8). Rozepsáním matice $A^{[\ell]}$ pomocí (3.1) dostaneme

$$W^{[\ell]} = (U^{(\ell)})^T U^{(\ell)} \Sigma^{(\ell)} (V^{(\ell)})^T.$$

Protože matice $U^{(\ell)}$ je ortogonální, můžeme psát

$$W^{[\ell]} = \Sigma^{(\ell)} (V^{(\ell)})^T = \begin{bmatrix} \Sigma_A^{(\ell)} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (V_A^{(\ell)})^T \\ (V_{\perp}^{(\ell)})^T \end{bmatrix} = \begin{bmatrix} \Sigma_A^{(\ell)} (V_A^{(\ell)})^T \\ 0 \end{bmatrix}.$$

Tato matice W má n_ℓ řádků, z nichž posledních $n_\ell - r_\ell$ je nulových. (Poznamenejme, že řádky mají nerostoucí normy dané singulárními čísly v matici $\Sigma_1^{(\ell)}$.) Tedy i vlákna tenzoru \mathcal{W} v módu ℓ mají posledních $n_\ell - r_\ell$ prvků nulových.

Vynásobíme-li tenzor \mathcal{A} podobně jako v (3.2) ve všech třech módech maticemi $U^{(\ell)}$, získáme tenzor

$$\mathcal{S} = \mathcal{A} \times_1 (U^{(1)})^T \times_2 (U^{(2)})^T \times_3 (U^{(3)})^T,$$

viz (2.5). Vlákna tenzoru \mathcal{S} mají stejně jako v případě (3.2) v každém módu posledních $n_\ell - r_\ell$ nulových, k dosažení nenulových prvků tohoto tenzoru tedy nepotřebujeme celé matice $U^{(\ell)}$, ale pouze matice $U_A^{(\ell)}$, tj.

$$\mathcal{S}_A = \mathcal{A} \times_1 (U_A^{(1)})^T \times_2 (U_A^{(2)})^T \times_3 (U_A^{(3)})^T, \quad \mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}, \quad \mathcal{S}_A \in \mathbb{R}^{r_1 \times r_2 \times r_3}.$$

Potom můžeme psát rozklad tenzoru \mathcal{A}

$$\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)},$$

resp. jeho redukovaný tvar, který je analogický k ekonomickému tvaru SVD matice,

$$\mathcal{A} = \mathcal{S}_A \times_1 U_A^{(1)} \times_2 U_A^{(2)} \times_3 U_A^{(3)}.$$

Definice 13 (Tuckerův rozklad (tvar) tenzoru). *Nechť $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$ je tenzor s vektorovým rankem (r_1, r_2, \dots, r_k) . Potom rozklad*

$$\mathcal{A} = \mathcal{S}_A \times_1 U_A^{(1)} \times_2 U_A^{(2)} \dots \times_k U_A^{(k)}, \quad U_A^{(\ell)} \in \mathbb{R}^{n_\ell \times r_\ell} \quad (3.3)$$

nazýváme Tuckerův rozklad tenzoru. Tenzor

$$\mathcal{S}_A \in \mathbb{R}^{r_1 \times r_2 \times \dots \times r_k}$$

nazýváme Tuckerovo jádro tenzoru \mathcal{A} .

Tuckerův rozklad znázorňuje obrázek 3.1. Všimněme si podobnosti mezi ekonomickým tvarem SVD matice (1.4), tj. $A = U_A \Sigma_A V_A^T$, a Tuckerovým tvarem tenzoru (3.3). Budeme-li matice A a Σ_A interpretovat jako tenzory druhého řádu, pak $A^{[1]} = A$, $A^{[2]} = A^T$ a platí $A = \Sigma_A \times_1 U_A \times_2 V_A$. Ekonomický tvar SVD je tedy speciálním případem Tuckerova rozkladu tenzoru pro tenzory řádu dva. Tuckerův rozklad se z tohoto důvodu také někdy nazývá *high-order SVD* (HOSVD).

Podstatným rozdílem je ovšem to, že Tuckerovo jádro, tj. tenzor \mathcal{S}_A , je pro tenzory řádu tři a více obecně husté, narozdíl od matic (tenzorů řádu dva), kde lze vždy volit Σ_A diagonální. To souvisí s tím, že pro matici (tenzor řádu dva) vždy platí $\text{rank}_1(A) = \text{rank}_2(A)$.

Příklad 11. Ukážeme si, jak vypadá jádro tenzoru z příkladu 1.

$$\mathcal{S}_A = \mathcal{A} \times_1 (U^{(1)})^T \times_2 (U^{(2)})^T \times_3 (U^{(3)})^T =$$

| | | | | | |
|------|-----|-----|------|------|------|
| 21.7 | 0.7 | 0.1 | 2.3 | 0.2 | 1.2 |
| 0.2 | 0.6 | 8.5 | 0.7 | -0.6 | -5.7 |
| -0.5 | 0.1 | 0.8 | 3.5 | 2.7 | 1.3 |
| 0.1 | 2.2 | 0.3 | -0.8 | 0.8 | 0.5 |

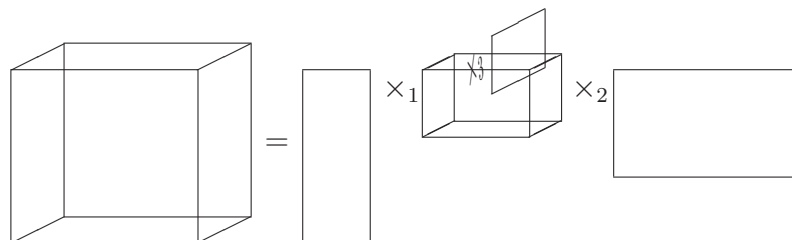
Výpočet jsme provedli v MATLABu. □

Vidíme, že v tomto příkladu nedojde k redukci rozměrů, neboť $r_\ell = n_\ell$ pro všechna ℓ . V praxi (u větších tenzorů) ale k redukci dochází, čehož lze využít, viz např. [18, obrázek 3.1]. Příklad většího tenzoru, kde jádro má menší rozměry je v příloze B.

Důsledek 3. Víme, že polyadický rank obecného tenzoru $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_k}$ lze omezit číslem $n_1 n_2 \dots n_k$, tj. dimenzí tenzoru. Pokud navíc víme, že má tenzor \mathcal{A} vektorový rank (r_1, r_2, \dots, r_k) , tj. že existuje jeho Tuckerův rozklad ve tvaru (3.3), pak lze jeho polyadický rank určitě omezit číslem $r_1 r_2 \dots r_k$, tj. dimenzí jeho Tuckerova jádra.

3.2 Základní operace s tenzory v Tuckerově tvaru

V této části se zaměříme na operace s tenzory v Tuckerově tvaru. Budeme chtít takové tenzory sčítat a násobit maticí tak jako v podkapitole 2.2. Výsledný tenzor



Obrázek 3.1: Tuckerův rozklad (HOSVD) tenzoru řádu 3. Převzato z [28].

což můžeme také zapsat

$$\mathcal{C} = \underbrace{(\mathcal{S}_{AB} \times_1 R_1 \times_2 R_2 \times_3 R_3)}_{\mathcal{W}} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3.$$

Protože matice Q_1 , Q_2 a Q_3 mají ortonormální sloupce, tenzory \mathcal{C} a \mathcal{W} mají stejnou hodnotu, tedy

$$\text{rank}(\mathcal{W}) = (q_1, q_2, q_3), \quad \text{kde } \mathcal{W} \in \mathbb{R}^{(r_1+p_1) \times (r_2+p_2) \times (r_3+p_3)}. \quad (3.6)$$

Krok 3 (Kompresce jádra): Zbývá nalézt Tuckerovo jádro tenzoru \mathcal{W} , které je zároveň jádrem tenzoru \mathcal{C} . Uvažujme Tuckerův rozklad

$$\mathcal{W} = \mathcal{S}_{\mathcal{W}} \times_1 U_W^{(1)} \times_2 U_W^{(2)} \times_3 U_W^{(3)}, \quad \mathcal{S}_{\mathcal{W}} \in \mathbb{R}^{q_1 \times q_2 \times q_3}. \quad (3.7)$$

Dosadíme do (3.6) za tenzor \mathcal{W} z (3.7) a získáme

$$\mathcal{C} = (\mathcal{S}_{\mathcal{W}} \times_1 U_W^{(1)} \times_2 U_W^{(2)} \times_3 U_W^{(3)}) \times_1 Q_1 \times_2 Q_2 \times_3 Q_3. \quad (3.8)$$

Označíme součiny matic, kterými násobíme tenzor $\mathcal{S}_{\mathcal{W}}$ v jednotlivých módech

$$U_C^{(1)} \equiv Q_1 U_W^{(1)}, \quad U_C^{(2)} \equiv Q_2 U_W^{(2)}, \quad U_C^{(3)} \equiv Q_3 U_W^{(3)},$$

připomeňme, že matice, kterými násobíme tenzor ve shodném módu se mezi sebou násobí v opačném pořadí, viz větu 4. Tyto matice mají ortonormální sloupce a po dosazení do (3.8) získáváme

$$\mathcal{C} = \mathcal{S}_{\mathcal{W}} \times_1 U_C^{(1)} \times_2 U_C^{(2)} \times_3 U_C^{(3)},$$

tj. Tuckerův rozklad tenzoru \mathcal{C} .

3.2.2 Násobení tenzoru v Tuckerově tvaru maticí

Další operací, kterou se budeme zabývat, je násobení tenzoru v Tuckerově tvaru maticí v ℓ -tém módu. Pro výklad použijeme jako příklad násobení tenzoru třetího řádu v prvním módu, lze ale zobecnit na tenzor řádu k a násobení v ℓ -tém módu, viz definici 3. Budeme chtít násobit tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_2}$,

$$\mathcal{A} = \mathcal{S}_{\mathcal{A}} \times_1 U_A^{(1)} \times_2 U_A^{(2)} \times_3 U_A^{(3)}, \quad \mathcal{S}_{\mathcal{A}} \in \mathbb{R}^{r_1 \times r_2 \times r_3}$$

a matici

$$M \in \mathbb{R}^{m \times n_1}.$$

Součin dostaneme opět ve třech krocích, analogicky jako v podkapitole 1.3.2.

Krok 1: Součin tohoto tenzoru a matice v prvním módu je

$$\mathcal{D} = \mathcal{A} \times_1 M = (\mathcal{S}_{\mathcal{A}} \times_1 U_A^{(1)} \times_2 U_A^{(2)} \times_3 U_A^{(3)}) \times_1 M, \quad (3.9)$$

Tabulka 3.1: Výpočetní cena při sčítání tenzorů řádu 3 v Tuckerově tvaru; cena výpočtu QR rozkladu lze nalézt např. v [12, kapitola 3.5.6]. Cena výpočtu HOSVD tenzoru řádu k je úměrná k -násobku ceny výpotu SVD; tu lze nalézt např. v [12, kapitoly 5.2.3]. Zkratka Tmp značí součin tenzoru a matice (tensor-matrix product), Mmp značí součin dvou matic (matrix-matrix product).

| Výpočet | Výpočetní cena |
|----------------|--|
| klasicky | $n_1 \cdot n_2 \cdot n_3$ sčítání |
| Tucker, krok 1 | zdarma (sestavení matic a tenzoru) |
| Tucker, krok 2 | 3 QR rozklady typu $(n_\ell \times (r_\ell + p_\ell))$, $\ell = 1, 2, 3$ + 3 Tmp typu $((r_1 + p_1) \times (r_2 + p_2) \times (r_3 + p_3)) \times_\ell \dots$ $\dots \times_\ell ((r_\ell + p_\ell) \times (r_\ell + p_\ell))$, $\ell = 1, 2, 3$ |
| Tucker, krok 3 | 1 HOSVD typu $((r_1 + p_1) \times (r_2 + p_2) \times (r_3 + p_3))$ + 3 Mmp $(n_\ell \times (r_\ell + p_\ell)) \cdot ((r_\ell + p_\ell) \times q_\ell)$, $\ell = 1, 2, 3$ |

hodnosti $\text{rank}(\mathcal{D}) = (q_1, q_2, q_3)$. Jinak můžeme psát

$$\mathcal{D} = \mathcal{A} \times_1 M = \mathcal{S}_A \times_1 \underbrace{(MU_A^{(1)})}_K \times_2 U_A^{(2)} \times_3 U_A^{(3)}, \quad (3.10)$$

tedy

$$\mathcal{D} = \mathcal{S}_A \times_1 K \times_2 U_A^{(2)} \times_3 U_A^{(3)}, \quad \mathcal{D} \in \mathbb{R}^{m \times n_2 \times n_3}. \quad (3.11)$$

Pro získání tenzoru \mathcal{D} ve tvaru (3.11) stačí pouze vynásobit matici, která násobí tenzor \mathcal{S}_A v příslušném (v našem případě prvním) módu maticí M .

Krok 2 (Ortogonalizace): Matice K ale zřejmě nemá ortonormální sloupce. Pokud chceme dosáhnout rozkladu tenzoru \mathcal{D} s ortogonálními maticemi, provedeme QR rozklad matice K a dosadíme jej do (3.11), dostaneme

$$K = Q_K R_K, \quad Q_K \in \mathbb{R}^{m \times r_1}, \quad R_K \in \mathbb{R}^{r_1 \times r_1},$$

a

$$\mathcal{D} = \mathcal{S}_A \times_1 (Q_K R_K) \times_2 U_A^{(2)} \times_3 U_A^{(3)}. \quad (3.12)$$

Pokud označíme součin

$$\mathcal{W} \equiv \mathcal{S}_A \times_1 R_K, \quad \mathcal{W} \in \mathbb{R}^{r_1 \times r_2 \times r_3},$$

můžeme psát rovnici (3.12) ve tvaru

$$\mathcal{D} = \mathcal{W} \times_1 Q_K \times_2 U_A^{(2)} \times_3 U_A^{(3)}. \quad (3.13)$$

Obecně ale tenzor \mathcal{W} nemusí Tuckerovým jádrem tenzoru \mathcal{D} a tedy ani rovnice (3.13) není Tuckerovým rozkladem tohoto tenzoru.

Tabulka 3.2: Výpočetní cena při násobení tenzoru maticí v prvním módu; cena výpočtu QR rozkladu lze nalézt např. v [12, kapitola 3.5.6]. Cena výpočtu HOSVD tenzoru řádu k je úměrná k -násobku ceny výpotu SVD; tu lze nalézt např. v [12, kapitoly 5.2.3]. Zkratka Tmp značí součin tenzoru a matice (tensor-matrix product), MMp značí součin dvou matic (matrix-matrix product), MVP značí součin matice a vektoru (matrix-vector product).

| Výpočet | Výpočetní cena |
|----------------|--|
| klasicky | 1 Tmp typu $(n_1 \times n_2 \times n_3) \times_1 (m \times n_1)$ = $(n_2 \cdot n_3)$ MVP typu $(m \times n_1) \cdot n_1$ |
| Tucker, krok 1 | 1 MMp typu $(m \times n_1) \cdot (n_1 \times r_1)$ |
| Tucker, krok 2 | 1 QR rozklad typu $(m \times r_1)$ + 1 Tmp typu $(r_1 \times r_2 \times r_3) \times_1 (r_1 \times r_1)$ |
| Tucker, krok 3 | 1 HOSVD typu $(r_1 \times r_2 \times r_3)$ + 1 MMp typu $(m \times r_1) \cdot (r_1 \times q_1)$ + 2 MMp typu $(n_\ell \times r_\ell) \cdot (r_\ell \times q_\ell)$, $\ell = 2, 3$ |

Krok 3 (Kompresce jádra): Abychom získali tento tvar, provedeme nejprve Tuckerův rozklad tenzoru \mathcal{W} , tj.

$$\mathcal{W} = \mathcal{S}_{\mathcal{W}} \times_1 U_W^{(1)} \times_2 U_W^{(2)} \times_3 U_W^{(3)}, \quad \text{rank}(\mathcal{W}) = (q_1, q_2, q_3),$$

který dosadíme do rovnice (3.13). Dostaneme

$$\mathcal{D} = (\mathcal{S}_{\mathcal{W}} \times_1 U_W^{(1)} \times_2 U_W^{(2)} \times_3 U_W^{(3)}) \times_1 Q_K \times_2 U_A^{(2)} \times_3 U_A^{(3)}. \quad (3.14)$$

Předchozí rovnici (3.14) můžeme ale také upravit tak, že předtím než provedeme násobení tenzoru maticemi, vynásobíme nejprve odpovídající matice, kterými bychom násobili jádro $\mathcal{S}_{\mathcal{W}}$ v příslušných módech, tj. získáme

$$U_D^{(1)} \equiv Q_K U_W^{(1)}, \quad U_D^{(2)} \equiv U_A^{(2)} U_W^{(2)}, \quad U_D^{(3)} \equiv U_A^{(3)} U_W^{(3)}.$$

Rovnice (3.14) má potom tvar

$$\mathcal{D} = \mathcal{S}_{\mathcal{W}} \times_1 U_D^{(1)} \times_2 U_D^{(2)} \times_3 U_D^{(3)},$$

tedy Tuckerův rozklad tenzoru \mathcal{D} .

4 Aproximace tenzoru pomocí Tuckerova rozkladu

V podkapitole 2.4 jsme se zabývali hodnotí tenzoru a mimo jiné jsme také došli k závěru, že poznatky o aproximaci matice maticí nižší hodnosti nelze zobecnit na tenzory, pokud pracujeme s polyadickým rankem. V této kapitole ukážeme, že pokud budeme jako hodnot tenzoru uvažovat vektorový rank, věta 2 o aproximaci matice maticí nižší hodnosti lze zobecnit na tenzory. Tohoto faktu se využívá např. v [37], [29], nebo [18].

4.1 Zobecněná Schmidtova–Eckartova–Youngova–Mirského věta

Nyní ukážeme, jak využít Tuckerův rozklad k aproximaci tenzoru tenzorem s nižším vektorovým rankem. Klíčovým nástrojem bude následující věta, viz také [41, str. 20] a [10, str. 1267].

Věta 7 (Aproximace tenzoru pomocí tenzoru s nižším vektorovým rankem). *Nechť tenzor $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_k}$ hodnosti (r_1, r_2, \dots, r_k) má Tuckerův rozklad*

$$\mathcal{A} = \mathcal{S}_A \times_1 U_A^{(1)} \cdots \times_k U_A^{(k)} \in \mathbb{R}^{n_1 \times \dots \times n_k},$$

kde tenzor

$$\mathcal{S}_A = (s_{i_1, \dots, i_k})_{i_\ell=1, \dots, r_\ell, \ell=1, 2, \dots, k} \in \mathbb{R}^{r_1 \times \dots \times r_k}$$

představuje Tuckerovo jádro a sloupce matice $U_A^{(\ell)} \in \mathbb{R}^{n_\ell \times r_\ell}$ jsou navzájem ortonormální levé singulární vektory tvořící bázi oboru hodnot matice $A^{[\ell]}$, $\ell = 1, \dots, k$. Označme dále $\sigma_1^{(\ell)} \geq \dots \geq \sigma_{r_\ell}^{(\ell)} > 0$ singulární čísla matice $A^{[\ell]}$. Nechť t_ℓ jsou čísla splňující

$$t_\ell \leq r_\ell, \quad \ell = 1, \dots, k.$$

Pak tenzor

$$\mathcal{B} = \mathcal{S}_B \times_1 U_B^{(1)} \cdots \times_k U_B^{(k)} \in \mathbb{R}^{n_1 \times \dots \times n_k}, \quad (4.1)$$

kde

$$\mathcal{S}_B \equiv (s_{i_1, \dots, i_k})_{i_\ell=1, \dots, t_\ell, \ell=1, 2, \dots, k} \in \mathbb{R}^{t_1 \times \dots \times t_k}$$

je podtenzor tenzoru \mathcal{S}_A a matice $U_B^{(\ell)} \in \mathbb{R}^{n_\ell \times t_\ell}$ je tvořena prvními t_ℓ sloupci matice $U_A^{(\ell)}$, má zřejmě hodnotu omezenou čísly t_ℓ , tj. $\text{rank}(\mathcal{B}) \leq (t_1, \dots, t_k)$ a aproximuje tenzor \mathcal{A} tak, že platí

$$\|\mathcal{A} - \mathcal{B}\| \leq \left(\sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2 \right)^{1/2}. \quad (4.2)$$

Důkaz. Protože rozklad (4.1) tenzoru \mathcal{B} má strukturu analogickou Tuckerovu rozkladu, tedy $\text{rank}(\mathcal{B}) \leq (t_1, \dots, t_k)$ platí triviálně. Abychom dokázali nerovnost (4.2), budeme uvažovat tenzory \mathcal{B}_ℓ , získané analogicky jako tenzor \mathcal{B} , takové, že

$$\mathcal{B}_\ell = \mathcal{S}_{\mathcal{B}_\ell} \times_1 U_A^{(1)} \cdots \times_{\ell-1} U_A^{(\ell-1)} \times_\ell U_B^{(\ell)} \times_{\ell+1} U_A^{(\ell+1)} \cdots \times_k U_A^{(k)},$$

kde tenzor

$$\mathcal{S}_{\mathcal{B}_\ell} \in \mathbb{R}^{r_1 \times \cdots \times r_{\ell-1} \times t_\ell \times r_{\ell+1} \times \cdots \times r_k}$$

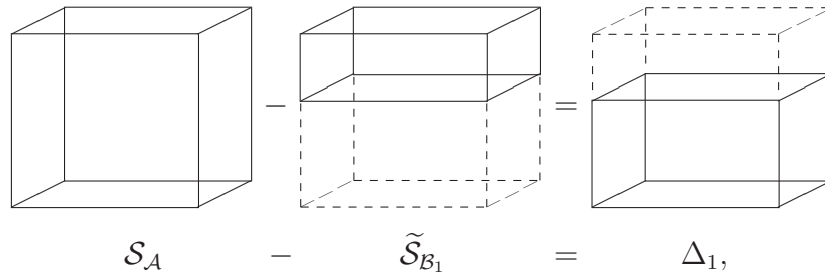
je podtenzor tenzoru \mathcal{S}_A a zároveň nadtenzor tenzoru \mathcal{S}_B . Dále budeme uvažovat tenzory

$$\begin{aligned} \tilde{\mathcal{S}}_B &\equiv \mathcal{B} \times_1 (U_A^{(1)})^T \cdots \times_k (U_A^{(k)})^T \in \mathbb{R}^{r_1 \times \cdots \times r_k}, \\ \tilde{\mathcal{S}}_{\mathcal{B}_\ell} &\equiv \mathcal{B}_\ell \times_1 (U_A^{(1)})^T \cdots \times_k (U_A^{(k)})^T \in \mathbb{R}^{r_1 \times \cdots \times r_k}, \quad \ell = 1, \dots, k. \end{aligned}$$

tj. tenzory \mathcal{S}_B a $\mathcal{S}_{\mathcal{B}_\ell}$ doplněné nulami. Protože tenzorová norma, kterou používáme, je ortogonálně invariantní (stejně jako eukleidovská norma vektoru a Frobeniova norma matice), pak zřejmě

$$\|\mathcal{A} - \mathcal{B}\| = \|\mathcal{S}_A - \tilde{\mathcal{S}}_B\| \quad \text{a} \quad \|\mathcal{A} - \mathcal{B}_\ell\| = \|\mathcal{S}_A - \tilde{\mathcal{S}}_{\mathcal{B}_\ell}\|. \quad (4.3)$$

Na levé straně nerovnosti (4.2) tak stačí místo tenzorů \mathcal{A} a \mathcal{B} uvažovat tenzory \mathcal{S}_A a $\tilde{\mathcal{S}}_B$. Nyní se podíváme jak vypadají rozdíly tenzorů $\Delta_\ell \equiv \mathcal{S}_A - \tilde{\mathcal{S}}_{\mathcal{B}_\ell}$, $\ell = 1, \dots, k$. Rozdíly můžeme schematicky znázornit následujícím způsobem (tato část důkazu je provedena kvůli názornosti pro tenzor třetího řádu, z konstrukce je však zřejmé, že lze použít na tenzor obecného řádu k ; čárkovaně jsou znázorněny bloky tenzorů, které obsahují pouze nuly, plnou čarou jsou pak znázorněny bloky s obecně nenulovými prvky):



$$\mathcal{S}_A - \tilde{\mathcal{S}}_{B_2} = \Delta_2,$$

$$\mathcal{S}_A - \tilde{\mathcal{S}}_{B_3} = \Delta_3.$$

Rozdíl tenzorů $\Delta \equiv \mathcal{S}_A - \tilde{\mathcal{S}}_B$ lze schematicky znázornit obdobně:

$$\mathcal{S}_A - \tilde{\mathcal{S}}_B = \Delta.$$

Označíme-li \mathcal{D} a \mathcal{D}_ℓ množiny *obecně nenulových* prvků (tj. prvků ležících v blocích vyznačených plnou čarou) tenzorů Δ a Δ_ℓ , pak, s využitím vztahu (4.3), dostaneme

$$\|\mathcal{A} - \mathcal{B}\|^2 = \sum_{\delta \in \mathcal{D}} \delta^2 \quad \text{a} \quad \|\mathcal{A} - \mathcal{B}_\ell\|^2 = \sum_{\delta \in \mathcal{D}_\ell} \delta^2,$$

pro $\ell = 1, \dots, k$. Protože

$$\mathcal{D} = \bigcup_{\ell=1}^k \mathcal{D}_\ell, \quad \text{pak} \quad \|\mathcal{A} - \mathcal{B}\|^2 \leq \sum_{\ell=1}^k \|\mathcal{A} - \mathcal{B}_\ell\|^2, \quad (4.4)$$

přičemž rovnost nastane tehdy a jen tehdy, když všechny průniky $\mathcal{D}_\ell \cap \mathcal{D}_s$, kde $\ell = 1, \dots, k$, $s = 1, \dots, k$, $\ell \neq s$, obsahují pouze nulové prvky.

Nakonec budeme uvažovat matice $A^{[\ell]}$ a $B_\ell^{[\ell]}$, tj. rozvoje tenzorů \mathcal{A} a \mathcal{B}_ℓ v matice v módu ℓ . Pak

$$A^{[\ell]} = \sum_{j=1}^{r_\ell} u_j^{(\ell)} \sigma_j^{(\ell)} (v_j^{(\ell)})^T, \quad B_\ell^{[\ell]} = \sum_{j=1}^{t_\ell} u_j^{(\ell)} \sigma_j^{(\ell)} (v_j^{(\ell)})^T,$$

představují singulární rozklady matic $A^{[\ell]}$ a $B_\ell^{[\ell]}$ ve tvaru dyadických rozvoje, viz (1.5). Věta 2 říká, že matice $B_\ell^{[\ell]}$ je nejlepší aproximací matice $A^{[\ell]}$ maticí hodnotnosti

t_ℓ , přičemž platí

$$\|A^{[\ell]} - B_\ell^{[\ell]}\|_F^2 = \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2. \quad (4.5)$$

S využitím faktu, že $\|\mathcal{T}\| = \|T^{[\ell]}\|_F$, tj. norma libovolného tenzoru je rovna Frobeniově normě rozvoje tohoto tenzoru v libovolném módu, a s využitím vztahů (4.4) a (4.5) ihned dostaneme

$$\|\mathcal{A} - \mathcal{B}\|^2 \leq \sum_{\ell=1}^k \|\mathcal{A} - \mathcal{B}_\ell\|^2 = \sum_{\ell=1}^k \|A^{[\ell]} - B_\ell^{[\ell]}\|_F^2 = \sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2.$$

Po odmocnění získáme nerovnost

$$\|\mathcal{A} - \mathcal{B}\| \leq \left(\sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2 \right)^{1/2},$$

kterou jsme chtěli dokázat. □

Poznamenejme, že aproximace tenzoru tenzorem nižší hodnosti se v praxi používá velmi často, viz např. [37], [29] nebo [18]. Kombinace výpočtu Tuckerova rozkladu daného tenzoru spolu s jeho aproximací tenzorem nižšího ranku se často nazývá *truncated Tucker decomposition* (příp. *truncated HOSVD*), tedy „oříznutý“ Tuckerův rozklad, viz např. [28, str. 477].

4.2 Rozbor zobecněné věty na příkladech

Nyní se podíváme na různé situace, které mohou v souvislosti s větou 7 nastat. Zejména nás bude zajímat, jak moc je odhad (4.2) nadhodnocený a kdy nastane rovnost.

4.2.1 Nadhodnocení odhadu

Nejprve se podíváme na následujícím příkladu, jak moc může být odhad nadhodnocený.

Příklad 12. *Uvažujme následující kubický, supersymetrický a navíc diagonální tenzor*

$$\mathcal{A} = \begin{array}{c} \begin{array}{|c|c|c|} \hline 1 & & \\ \hline & 1 & \\ \hline & & 1 \\ \hline \end{array} \end{array} \in \mathbb{R}^{6 \times 6 \times 6}. \quad (4.6)$$

Protože je supersymetrický, jsou všechny jeho rozvoje identické. Zřejmě

$$A^{[\ell]} = \left[\begin{array}{c|c|c|c|c|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \middle| \begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \middle| \begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \middle| \begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \middle| \begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right],$$

přičemž její singulární rozklad je $A^{[\ell]} = U_A^{(\ell)} \Sigma_A^{(\ell)} (V_A^{(\ell)})^T$, kde

$$U_A^{(\ell)} \equiv \Sigma_A^{(\ell)} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad V_A^{(\ell)} \equiv (A^{[\ell]})^T, \quad \ell = 1, 2, 3,$$

a tedy $\text{rank}(\mathcal{A}) = (6, 6, 6)$. Tensor \mathcal{A} je tedy sám svým Tuckerovým jádrem. Pokud se pokusíme tensor aproximovat tensorem \mathcal{B} hodnosti $(5, 5, 5)$ a tensor budeme konstruovat tak jak se postupuje v předchozím důkazu, pak dostaneme

$$\mathcal{B} = \begin{array}{c} \begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{array} \end{array} \cdot \quad (4.7)$$

Zřejmě $\|\mathcal{A} - \mathcal{B}\| = 1$, ale $(\sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2)^{1/2} = (\sum_{\ell=1}^3 (\sigma_6^{(\ell)})^2)^{1/2} = \sqrt{3}$. \square

Z předchozího příkladu je zřejmé, že obecně, pro analogický tensor řádu k , může dojít k nadhodnocení až (\sqrt{k}) -krát. Z konstrukce důkazu i předchozího příkladu se zdá, že k horšímu nadhodnocení nemůže dojít, tedy, že platí

$$\|\mathcal{A} - \mathcal{B}\| \leq \left(\sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2 \right)^{1/2} \leq \sqrt{k} \|\mathcal{A} - \mathcal{B}\|. \quad (4.8)$$

Druhá nerovnost ve vztahu (4.8) lze dokázat obdobně jako nerovnost z věty 7, viz také text psaný kurzívou pod rovnicí (4.4). Důkaz, s využitím značení z důkazu věty 7, jen naznačíme. Konkrétně budeme potřebovat množiny \mathcal{D} a \mathcal{D}_ℓ . Pokud množina $\mathcal{D} \setminus (\cap_{\ell=1}^k \mathcal{D}_\ell)$ obsahuje pouze nulové prvky, pak je každý nenulový prvek množiny \mathcal{D} obsažen ve všech množinách \mathcal{D}_ℓ , tj. Tuckerovo jádro $\mathcal{S}_\mathcal{A}$ je blokově diagonální tensor. Každý nenulový prvek „oříznutý“ z původního jádra je pak do odhadu na pravé straně nerovnosti (4.2), resp. jejího kvadrátu, započten právě k -krát, horší situace zřejmě nastat nemůže. Pak je ale odhad právě (\sqrt{k}) -krát nadhodnocen.

Libovolný tenzor druhého řádu, tedy matice A hodnosti r (formálně (r, r) ; řádková i sloupcová hodnota matice je vždy stejná), je takovým příkladem pro $k = 2$. Tuckerovo jádro matice A je *diagonální* matice $\Sigma_A \in \mathbb{R}^{r \times r}$ ze singulárního rozkladu (1.4). Platí $A^{[1]} = A$, $A^{[2]} = A^T$, $U_A^{(1)} = U_A$, $U_A^{(2)} = V_A$ a $\sigma_j^{(1)} = \sigma_j^{(2)}$. Budeme-li chtít hodnost matice A snížit z r na nějaké $t \leq r$ (formálně (t, t)) a aplikujeme-li odhad (4.2), pak bude započítáno každé singulární číslo matice dvakrát a celkový odhad bude právě $(\sqrt{2})$ -krát nadhodnocen.

4.2.2 Rovnost odhadu a hodnost výsledného tenzoru

Nyní za zaměříme na otázku, kdy ve vztahu (4.2) nastane rovnost. Ta je v podstatě již zodpovězena v důkazu v textu psaném kurzívou pod rovnicí (4.4). Taková situace zřejmě nastane např. tehdy, když ve vztazích $t_\ell \leq r_\ell$ nastane ostrá nerovnost jen pro jedno ℓ , např. $\ell = 1$, a ve všech ostatních případech nastane rovnost, tj. $t_1 < r_1$ a $t_\ell = r_\ell$, $\ell = 2, \dots, k$. Pak se odhad (4.2) zredukuje na rovnost (1.16) z věty 2 pro matice $A^{[1]}$ a $B^{[1]}$. To však s sebou může nést další zajímavé jevy, které budeme ilustrovat opět na tenzoru z předchozího příkladu.

Příklad 13. *Uvažujme tenzor \mathcal{A} (4.6) z přechozího příkladu. Pokusme se nyní tenzor aproximovat tenzorem \mathcal{B} např. hodnosti $(5, 6, 6)$. Nechť tedy nyní platí $t_1 = 5 < r_1 = 6$, $t_2 = r_2 = 6$, $t_3 = r_3 = 6$. Použijeme-li stejný postup jako dosud, dostaneme tenzor \mathcal{B} identický s tenzorem (4.7). Opět tedy platí $\|\mathcal{A} - \mathcal{B}\| = 1$, ale tentokrát $(\sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2)^{1/2} = \sigma_6^{(1)} = 1$. Odhad je tedy přesný. Tenzor \mathcal{B} , který jsme získali však není hodnosti $(5, 6, 6)$, jak jsme si přáli, ale hodnosti $(5, 5, 5)$. Získali jsem tedy tenzor hodnosti nižší, než bylo požadováno. \square*

Další příklady aproximace tenzoru tenzorem nižší hodnosti, kdy snižujeme hodnost jen v jediném módu nalezneme v příloze C, v tabulce C.2, řádky 1–4, 5–9, 10–11. Všimněme si, že, podobně jako v přechozím příkladu, i v tabulce C.2 dochází v řádcích 4, 9 a 11 k tomu, že je hodnost získaného tenzoru nižší než hodnost požadovaná. V těchto případech je tomu proto, že zde tenzor třetího řádu aproximujeme tenzorem, který má ℓ -rank v módu $\ell = 1, 2$, resp. 3 roven jedné. Tenzor třetího řádu tedy aproximujeme vnějším součinem matice a vektoru. Matice však má stejný počet lineárně nezávislých vláken v obou módech. Výsledný tenzor tak musí mít v obou módech různých od ℓ stejnou hodnost. Obecně tedy má nižší hodnost než bylo požadováno. Z podobného důvodu i při aproximaci *diagonálního* tenzoru tenzorem, u kterého předepíšeme různé ℓ -ranky v různých módech, dostaneme výsledný tenzor hodnosti nižší, než byla předepsaná, ne vyšší než $\min_\ell \text{rank}_\ell(\mathcal{A})$.

Na základě předchozích dvou příkladů by nás mohlo zajímat, zda je možné nalézt aproximaci tenzoru tenzorem nižší hodnosti s různými ℓ -ranky tak aby přitom byla zachována rovnost ve vztahu (4.2). V příloze C, v tabulce C.2 vidíme, že taková situace určitě nastala v řádcích 1–3, 5–8 a 10. V této příloze však pracujeme s obecně hustým tenzorem, který má různé rozměry a různé ranky v různých módech. Na základě těchto pozorování by se mohlo zdát, že hraje důležitou roli to, zda tenzor je kubický a supersymetrický, jako (4.6), nebo zda není, jako tomu je v příloze. Toto však nehraje roli, jak ukazuje následující příklad.

Příklad 14. Uvažujme následující kubický a supersymetrický tenzor

$$\mathcal{A} = \begin{array}{c} \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \\ \in \mathbb{R}^{6 \times 6 \times 6}. \end{array} \quad (4.9)$$

Protože je supersymetrický, jsou všechny jeho rozvoje identické. Zřejmě

$$A^{[\ell]} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & | & 1 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 1 & 0 & | & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 1 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 1 & 0 & | & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 1 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 1 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

přičemž její singulární rozklad je $A^{[\ell]} = U_A^{(\ell)} \Sigma_A^{(\ell)} (V_A^{(\ell)})^T$, kde

$$U_A^{(\ell)} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Sigma_A^{(\ell)} \equiv \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \ell = 1, 2, 3,$$

a tedy $\text{rank}(\mathcal{A}) = (6, 6, 6)$. Tenzor \mathcal{A} je tedy sám svým Tuckerovým jádrem. Budeme-li nyní aproximovat tento tenzor tenzorem hodnosti $(5, 6, 6)$, dostaneme tenzor

$$\mathcal{B} = \begin{array}{c} \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 0 \\ \hline \end{array} \\ \in \mathbb{R}^{6 \times 6 \times 6}. \end{array}$$

Snadno ověříme, že $\text{rank}(\mathcal{B}) = (5, 6, 6)$ a $\|\mathcal{A} - \mathcal{B}\| = 1$, což se přímo rovná odhadu $(\sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2)^{1/2} = \sigma_6^{(1)} = 1$. \square

Nyní se ještě podíváme, co se stane, pokud se tenzor z předchozího příkladu pokusíme aproximovat tenzorem nižší hodnosti ve více než jednom módu.

Příklad 15. Uvažujme tenzor \mathcal{A} (4.9) z příkladu 14. Pokusme se nyní tento tenzor aproximovat tenzorem \mathcal{B} hodnosti $(5, 5, 6)$. Dostaneme tak tenzor

$$\mathcal{B} = \begin{array}{c} \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 0 \\ \hline \end{array} \\ \in \mathbb{R}^{6 \times 6 \times 6}. \end{array} \quad (4.10)$$

Přestože jsme nyní modifikovali tenzor ve dvou módech, ve vztahu (4.2) i nadále platí rovnost, $\|\mathcal{A} - \mathcal{B}\| = \sqrt{2}$, $(\sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2)^{1/2} = ((\sigma_6^{(1)})^2 + (\sigma_6^{(2)})^2)^{1/2} = \sqrt{2}$ (viz text psaný kurzívou pod rovnicí (4.4)). Pokud se však podíváme na hodnotu výsledného tenzoru, zjistíme, že $\text{rank}(\mathcal{B}) = (5, 5, 5)$. Tím, že jsme z báze prostoru vláken prvního módu odebrali vektor $u_6^{(1)}$ (šestý sloupec matice $U_A^{(1)}$) a zároveň jsme z báze prostoru vláken druhého módu odebrali vektor $u_6^{(2)}$ (šestý sloupec matice $U_A^{(2)}$), zmizel tím automaticky i vektor $u_3^{(3)}$ z báze prostoru vláken třetího módu (třetí sloupec matice $U_A^{(3)}$). \square

Předchozí příklad ukazuje, že odebrání dvou nejmenších singulárních čísel ve dvou různých módech, může být ekvivalentní odebrání velkého singulárního čísla ve zbývajícím módu. To, že singulární čísla jednotlivých rozvoju nejsou nezávislá, ale provázaná také ilustruje příklad v příloze C v tabulce C.2, speciálně pak grafy na obrázcích C.1, C.2 a C.3, které ukazují pohyby singulárních čísel v jednotlivých módech při aproximaci tenzoru.

4.2.3 Optimalita aproximace

Na závěr se pokusíme tenzor z předchozích dvou příkladů aproximovat tenzorem s nižší hodnotou ve všech módech. Výsledek tohoto pokusu nám umožní diskutovat o optimalitě takové aproximace.

Příklad 16. Uvažujme opět tenzor \mathcal{A} (4.9) z příkladu 14. Pokusíme se tento tenzor aproximovat tenzorem \mathcal{B} hodnoty $(5, 5, 5)$. Dostaneme tak tenzor

$$\mathcal{B} = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 0 \\ \hline 1 & 1 & 0 \\ \hline \end{array} \in \mathbb{R}^{6 \times 6 \times 6}. \quad (4.11)$$

Tenzor jsme nyní modifikovali ve všech třech módech. Ve vztahu (4.2) i nadále platí rovnost, $\|\mathcal{A} - \mathcal{B}\| = \sqrt{3}$, $(\sum_{\ell=1}^k \sum_{j=t_\ell+1}^{r_\ell} (\sigma_j^{(\ell)})^2)^{1/2} = ((\sigma_6^{(1)})^2 + (\sigma_6^{(2)})^2 + (\sigma_6^{(3)})^2)^{1/2} = \sqrt{3}$. Podíváme-li se ale na hodnotu výsledného tenzoru, zjistíme, že $\text{rank}(\mathcal{B}) = (4, 4, 4)$. Tím, že jsme odebrali z báze prostoru vláken ℓ -tého módu vektor $u_6^{(\ell)}$ zároveň pro všechna ℓ , $\ell = 1, 2, 3$, z báze tohoto prostoru zmizel automaticky i vektor $u_3^{(\ell)}$. \square

Vidíme tedy, že tento postup nevede k aproximaci tenzoru \mathcal{A} (4.9) tenzorem hodnoty $(5, 5, 5)$, jak jsme požadovali a které jsme dosáhli v příkladu 14. Tenzor (4.10) z příkladu 14 navíc, narozdíl od tenzoru (4.11), aproximuje původní tenzor s menší chybou.

Tenzor \mathcal{B} (4.1) z věty 7 tedy obecně není optimální ve smyslu minimality normy, tak jak tomu bylo ve větě 2 (viz stranu 23).

Na druhou stranu aproximace původního tenzoru (4.9) tenzorem (4.11) může být někdy vhodnější, jelikož tenzor (4.11) je supersymetrický tak jako tenzor původní. Tenzor (4.10) je sice lepší aproximací hodnosti (5, 5, 5) ve smyslu normy, je ale jen částečně symetrický.

Závěr

Tato práce má čtenáři sloužit jako úvod do problematiky práce s tenzory. S tenzorem se zde pracuje výhradně jako s vícerozměrným polem čísel, díky čemuž je možné matici považovat za tenzor druhého řádu. V textu jsme se snažili poukazovat na analogie mezi maticemi a tenzory, ale také upozornit na takové vlastnosti matic, které naopak na tenzory zobecnit nelze, nebo to lze obtížně. Pozornost jsme věnovali zejména problému zavedení hodnoty tenzoru řádu k , $k \geq 3$, která je příkladem vlastnosti, která lze zobecnit jen s jistými obtížemi. Ukázali jsme dva možné přístupy pomocí tzv. polyadického a vektorového ranku tenzoru. První přístup s sebou přinesl nepříjemnosti topologického rázu (vždy existuje tenzor v jehož libovolně malém okolí je tenzor nižšího polyadického ranku), druhý pak nepříjemnosti se strukturou ranku jako takového (počet lineárně nezávislých vláken může být v různých módech různý, vektorový rank tenzoru tedy není číslo ale vektor).

Hlavním cílem práce bylo zavedení Tuckerova rozkladu tenzoru jako zobecnění singulárního rozkladu matic. Ukázali jsme, jak lze tento rozklad zkonstruovat a jak je možno s tenzory, máme-li je k dispozici pouze ve tvaru tohoto rozkladu, provádět některé základní operace. To je důležité např. tehdy, jsou-li tenzory, se kterými pracujeme, příliš velké na to, abychom s nimi mohli pracovat v počítači přímo, ale jsou malého (vektorového) ranku, je možné uchovávat jejich Tuckerovy rozklady a provádět výpočty v tzv. low-rank aritmetice.

Důležitým výstupem této práce je zformulování věty o aproximaci tenzoru tenzorem nižšího vektorového ranku, která je jistým zobecněním známé Schmidty–Eckartovy–Youngovy–Mirského věty o aproximaci matice maticí nižšího ranku, a její důkaz (který jsme provedli nezávisle na existující literatuře). Tato věta právě využívá struktury Tuckerova rozkladu tenzoru. Na řadě konkrétních příkladů jsme pak ilustrovali sílu i nedostatky modifikovaného tvrzení, zejména v porovnání s její maticovou předlohou.

A Software

V této příloze ještě zmíníme software, který je možné používat při práci a výpočtech s tenzory. Vycházet budeme mimo jiné z přehledu softwaru pro tenzory z článku [28]. První pokusy o programování základních operací v prostoru tenzorů se objevily už v sedmdesátých letech dvacátého století, viz článek [35], dnes už je ale dostupných mnoho mnoho programů a rozšiřujících *toolboxů*, které lze pro výpočty s tenzory používat.

A.1 MATLAB

Jedním z nejpoužívanějších programů v matematice je MATLAB[®] vyvíjený komerčně firmou *Math Works*[®]. Z hlediska tenzorů podporuje MATLAB tzv. vícerozměrná pole (multidimensional arrays, MDAs), v rámci nichž je možné provádět operace s prvky tenzorů, permutace indexů a vektorové operace (např. součet). Složitější operace (jako je např. součin tenzorů) samotný MATLAB nepodporuje, je ale možné v něm takové operace provádět přidáním externích *toolboxů*. Stručný výčet několika takových *toolboxů* následuje.

Podporu pro práci s tenzory mají kromě MATLABu i další matematické programy. Takovými programy jsou např. MATHEMATICA firmy *Wolfram Research, Inc.*, nebo MAPLE firmy *Waterloo Maple, Inc. (Maplesoft)*, kde lze s výhodou pracovat především s tenzory s řídkou strukturou.

A.2 TENSOR TOOLBOX

Dalším nástrojem pro MATLAB je TENSOR TOOLBOX, jehož autory jsou B. W. Bader and T. G. Kolda, viz [2], [3] a [4]. Tento software rozšiřuje základní funkce MATLABu tak, aby bylo možné provádět operace s tenzory jako je rozvoj tenzoru v matici nebo tenzorové násobení, obsahuje také programové třídy pro polyadický tvar tenzoru (*ktensor*) nebo pro Tuckerův tvar (*ttensor*). Tento nástroj také umožňuje efektivně ukládat tzv. řídké tenzory, stejně jako tenzory s nějakou speciální strukturou. Uživatel zde má možnost konstruovat také vlastní algoritmy.

Toolbox je volně dostupný, aktuálně (od 6. února 2015) ve verzi 2.6, na adrese <http://www.sandia.gov/~tgkolda/TensorToolbox>.

A.3 HIERARCHICAL TUCKER TOOLBOX

Další rozšiřující programovou třídou, inspirovanou právě třídami `ktensor` a `tten-
sor`, je třída `htucker`, kterou vyvinula C. Tobler pro práci s tenzory v tzv. *hierar-
chickém Tuckerově tvaru*, viz [19]. Tento nástroj ale umožňuje provádět operace i
s tenzory, které ve tvaru hierarchického Tuckerova rozkladu nejsou. Pro popis hierar-
chického Tuckerova rozkladu i nástroje `htucker` viz [41] a [30].

Toolbox je volně dostupná na adrese <http://anchp.epfl.ch/htucker>.

A.4 TENSORLAB

Pro práci s tenzorovými rozkladu v MATLABu je vhodný také nástroj TENSOR-
LAB autorů L. Sorbera, M. Van Barela a L. De Lathauwera, viz [39]. Tento pro-
gram je nástrojem pro výpočty tenzorových rozkladů, např. polyadický rozklad,
Tuckerův rozklad atd., a následnou práci s tenzory v těchto tvarech. TENSORLAB
nabízí algoritmy umožňující definovat tenzory a matice, jejich součiny a rozklady,
přičemž má uživatel možnost navrhnout vlastní struktury nebo si vybrat ze seznamu
předdefinovaných struktur, včetně takových, které splňují např. podmínku ortogo-
nality. Také zde mimo jiné najdeme algoritmus pro výpočet hodnoty tenzoru a s tím
související aproximaci pomocí nižšího vektorového ranku (low multilinear rank ap-
proximation). TENSORLAB umožňuje mnoho dalších výpočtů týkajících se problémů
optimalizace apod., kterými jsme se ale v této práci nezabývali.

Toolbox je volně dostupný na adrese <http://www.tensorlab.net>.

A.5 N-WAY TOOLBOX & CUBATCH

Nástroje N-WAY TOOLBOX a CUBATCH jsou volně dostupné. N-WAY TOOLBOX
obsahuje mimo jiné umožňuje výpočet tenzorových rozkladů, včetně polyadického
rozvoje a Tuckerova rozkladu, které jsme popsali v této práci. CUBATCH je grafické
uživatelské rozhraní, které v MATLABu umožňuje analyzovat data získaná v N-WAY
TOOLBOXu.

N-WAY TOOLBOX je volně dostupný na adrese

<http://www.models.life.ku.dk/nwaytoolbox>.

CUBATCH je volně dostupný na adrese

<http://www.models.life.ku.dk/cubatch>.

A.6 PLS TOOLBOX

PLS TOOLBOX je komerční software vyvíjený firmou *Eigenvector Research, Inc.*,
konstruovaný především pro využití v chemometrii.

Toolbox je k dispozici na adrese

http://www.eigenvector.com/software/pls_toolbox.htm.

B Příklad tenzoru s malým Tuckerovým jádrem

Výpočty v této a následující příloze C (stejně jako jádro v příkladu 11, viz str. 45) jsou provedeny v programu MATLAB[®] 8.1.0.604 (R2013a) bez použití výše toolboxů zmíněných v předchozí příloze, na počítači Dell Latitude 6430U s procesorem Intel[®] Core[™] i7-3687U, 2.10 GHz s 8.00 GB RAM a se 64 bitovým operačním systémem Windows[®] 7 Professional, SP1.

Tenzor $\mathcal{A} \in \mathbb{R}^{8 \times 10 \times 4}$ hodnosti (5, 6, 3)

```
A(:, :, 1) =
-0.7817 -0.6152 -0.5934 -0.1379 -0.5698 0.3926 -0.3953 -1.2317 0.5990 0.5995
-0.0559 -0.0433 -0.0408 -0.0100 -0.0406 0.0274 -0.0276 -0.0877 0.0400 0.0417
-0.5077 -0.3981 -0.3822 -0.0898 -0.3696 0.2537 -0.2553 -0.7992 0.3835 0.3869
0.6822 0.5392 0.5229 0.1200 0.4979 -0.3446 0.3472 1.0762 -0.5316 -0.5270
-0.6148 -0.4835 -0.4660 -0.1085 -0.4480 0.3084 -0.3106 -0.9685 0.4698 0.4709
-0.5373 -0.4235 -0.4093 -0.0947 -0.3919 0.2704 -0.2723 -0.8470 0.4143 0.4132
0.5367 0.4260 0.4153 0.0941 0.3922 -0.2726 0.2749 0.8476 -0.4251 -0.4175
-1.5256 -1.2040 -1.1654 -0.2685 -1.1130 0.7689 -0.7747 -2.4057 1.1819 1.1755

A(:, :, 2) =
-0.2729 -0.0763 -0.4582 -0.0616 -0.1092 0.2426 -0.1919 -0.6477 0.8182 -0.0296
-0.0170 -0.0037 -0.0306 -0.0059 -0.0079 0.0195 -0.0121 -0.0467 0.0571 -0.0070
-0.1724 -0.0461 -0.2934 -0.0428 -0.0710 0.1617 -0.1215 -0.4214 0.5285 -0.0286
0.2459 0.0721 0.4065 0.0493 0.0950 -0.2053 0.1724 0.5641 -0.7186 0.0110
-0.2135 -0.0592 -0.3594 -0.0492 -0.0859 0.1918 -0.1503 -0.5096 0.6429 -0.0255
-0.1898 -0.0540 -0.3169 -0.0411 -0.0750 0.1649 -0.1333 -0.4449 0.5637 -0.0161
0.1994 0.0610 0.3249 0.0353 0.0746 -0.1565 0.1395 0.4428 -0.5689 -0.0029
-0.5438 -0.1670 -0.9039 -0.1138 -0.2128 0.4642 -0.3818 -1.2624 1.6035 -0.0363

A(:, :, 3) =
-0.6973 0.2069 -1.1166 0.6581 -0.9868 0.1283 -0.0163 -0.2909 0.9948 0.6424
-0.0490 0.0171 -0.0794 0.0488 -0.0712 0.0085 0.0005 -0.0177 0.0697 0.0455
-0.4511 0.1388 -0.7241 0.4307 -0.6421 0.0822 -0.0074 -0.1831 0.6433 0.4163
0.6113 -0.1735 0.9762 -0.5691 0.8593 -0.1139 0.0193 0.2632 -0.8727 -0.5620
-0.5480 0.1638 -0.8780 0.5184 -0.7764 0.1006 -0.0120 -0.2274 0.7817 0.5049
-0.4801 0.1402 -0.7680 0.4509 -0.6778 0.0888 -0.0127 -0.2027 0.6851 0.4420
0.4831 -0.1311 0.7693 -0.4437 0.6745 -0.0911 0.0192 0.2143 -0.6901 -0.4432
-1.3649 0.3936 -2.1817 1.2768 -1.9232 0.2532 -0.0392 -0.5812 1.9481 1.2556

A(:, :, 4) =
-0.6657 -0.6279 -0.3473 -0.1407 -0.5149 0.2861 -0.3207 -0.9633 0.1643 0.6349
-0.0490 -0.0452 -0.0243 -0.0095 -0.0365 0.0186 -0.0231 -0.0684 0.0095 0.0468
-0.4350 -0.4083 -0.2244 -0.0903 -0.3338 0.1824 -0.2086 -0.6247 0.1024 0.4148
0.5767 0.5473 0.3048 0.1246 0.4503 -0.2552 0.2794 0.8423 -0.1502 -0.5502
-0.5242 -0.4940 -0.2729 -0.1103 -0.4048 0.2242 -0.2523 -0.7574 0.1282 0.5000
-0.4564 -0.4314 -0.2392 -0.0972 -0.3542 0.1983 -0.2203 -0.6627 0.1149 0.4353
0.4504 0.4299 0.2412 0.0994 0.3549 -0.2051 0.2195 0.6639 -0.1235 -0.4298
-1.2931 -1.2244 -0.6803 -0.2771 -1.0062 0.5664 -0.6253 -1.8824 0.3306 1.2335
```

Matice $U_A^{(1)} \in \mathbb{R}^{8 \times 5}$, $U_A^{(2)} \in \mathbb{R}^{10 \times 6}$, $U_A^{(3)} \in \mathbb{R}^{4 \times 3}$

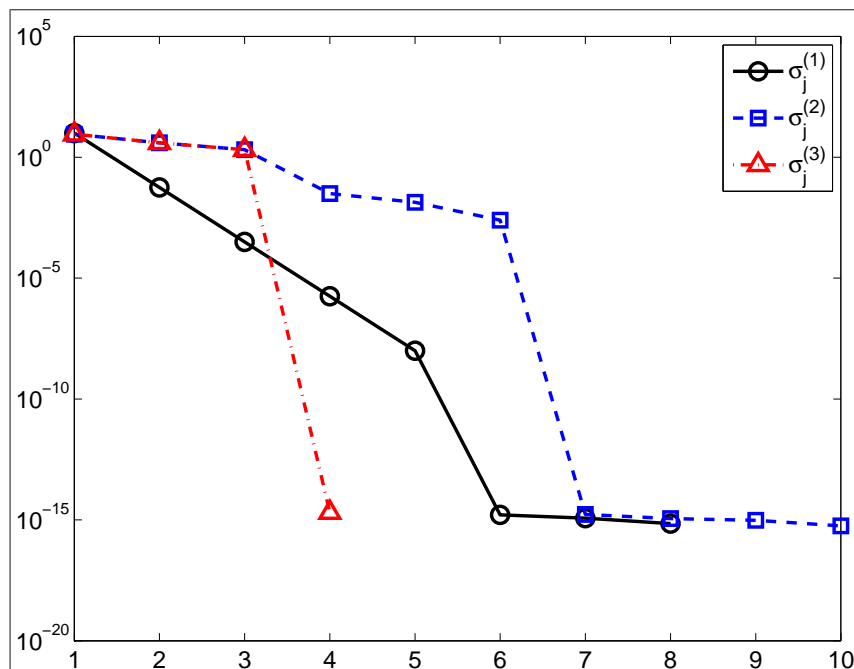
```

U1(:,1:r1) =
-0.3632 -0.3098 -0.2626 -0.3423 -0.0705
-0.0257 -0.1961 0.2902 -0.1392 0.5469
-0.2354 -0.5347 -0.4884 0.1221 -0.1051
0.3178 -0.2584 0.3343 -0.4789 0.1884
-0.2856 -0.3236 0.5416 0.6520 -0.0180
-0.2499 -0.0605 -0.2323 0.0643 0.7404
0.2506 -0.6152 0.1548 -0.1454 -0.2435
-0.7101 0.1615 0.3540 -0.4109 -0.2039

U2(:,1:r2) =
-0.3878 -0.0361 -0.1853 0.1003 -0.5892 0.0771
-0.1769 -0.4804 -0.1472 0.2405 0.3587 -0.2554
-0.3988 0.3480 0.0703 0.0853 0.0326 -0.4503
0.0616 -0.4454 0.1939 -0.0452 0.1571 -0.3152
-0.3605 0.2267 -0.4023 -0.5525 0.1586 -0.3348
0.1602 0.1267 -0.1139 0.4688 -0.4613 -0.3744
-0.1418 -0.1987 0.0771 0.2914 0.0329 -0.3613
-0.4812 -0.4762 0.2369 -0.2429 -0.3347 0.2043
0.3890 -0.3387 -0.6295 -0.2415 -0.2429 -0.1459
0.3112 0.0098 0.5180 -0.4423 -0.2973 -0.4253

U3(:,1:r3) =
-0.6087 0.3933 0.0531
-0.3169 -0.0085 0.8838
-0.5623 -0.7970 -0.2190
-0.4614 0.4583 -0.4101

```



Obrázek B.1: Singulární čísla $\sigma_j^{(\ell)}$, $j = 1, \dots, n_\ell$, jednotlivých rozvojevů $A^{[\ell]}$ (matricizations), $\ell = 1, 2, 3$, tenzoru $\mathcal{A} \in \mathbb{R}^{8 \times 10 \times 4}$ hodnosti $(5, 6, 3)$.

Tensor $\mathcal{A} \times_1 (U^{(1)})^T$

| | | | | | | | | | |
|----------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| AU1(:, :, 1) = | | | | | | | | | |
| 2.1493 | 1.6952 | 1.6397 | 0.3785 | 1.5678 | -1.0825 | 1.0904 | 3.3886 | -1.6613 | -1.6544 |
| 0.0032 | -0.0019 | -0.0071 | 0.0013 | 0.0010 | 0.0021 | -0.0026 | 0.0027 | 0.0143 | 0.0048 |
| -0.0000 | -0.0000 | -0.0001 | 0.0000 | -0.0000 | -0.0001 | -0.0000 | 0.0001 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 | -0.0000 |
| 0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AU1(:, :, 2) = | | | | | | | | | |
| 0.7628 | 0.2187 | 1.2705 | 0.1623 | 0.2998 | -0.6568 | 0.5357 | 1.7789 | -2.2570 | 0.0576 |
| -0.0134 | -0.0101 | -0.0106 | 0.0087 | 0.0009 | -0.0132 | -0.0085 | 0.0048 | 0.0053 | 0.0281 |
| -0.0000 | -0.0001 | -0.0001 | -0.0000 | -0.0000 | -0.0000 | -0.0001 | -0.0000 | -0.0000 | -0.0000 |
| 0.0000 | -0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 |
| -0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AU1(:, :, 3) = | | | | | | | | | |
| 1.9218 | -0.5576 | 3.0729 | -1.8010 | 2.7103 | -0.3559 | 0.0530 | 0.8149 | -2.7425 | -1.7684 |
| -0.0025 | -0.0141 | 0.0013 | -0.0125 | 0.0076 | 0.0031 | -0.0096 | -0.0164 | 0.0045 | 0.0001 |
| 0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0001 | -0.0000 | 0.0001 | -0.0000 | -0.0001 |
| 0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AU1(:, :, 4) = | | | | | | | | | |
| 1.8235 | 1.7253 | 0.9577 | 0.3896 | 1.4171 | -0.7956 | 0.8811 | 2.6513 | -0.4628 | -1.7395 |
| 0.0107 | 0.0040 | -0.0019 | -0.0028 | 0.0004 | 0.0093 | 0.0021 | 0.0009 | 0.0123 | -0.0100 |
| -0.0000 | 0.0000 | -0.0000 | 0.0001 | 0.0000 | -0.0001 | -0.0000 | 0.0001 | 0.0000 | 0.0001 |
| -0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 |
| 0.0000 | -0.0000 | 0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Tensor $\mathcal{A} \times_2 (U^{(2)})^T$

| | | | | | | | | | |
|----------------|---------|---------|---------|---------|---------|---|---|---|---|
| AU2(:, :, 1) = | | | | | | | | | |
| 1.9768 | 0.5673 | -0.0374 | 0.0023 | 0.0029 | 0.0001 | 0 | 0 | 0 | 0 |
| 0.1387 | 0.0414 | -0.0014 | 0.0014 | 0.0018 | 0.0001 | 0 | 0 | 0 | 0 |
| 1.2785 | 0.3701 | -0.0218 | 0.0040 | 0.0049 | 0.0001 | 0 | 0 | 0 | 0 |
| -1.7337 | -0.4925 | 0.0366 | 0.0018 | 0.0024 | 0.0001 | 0 | 0 | 0 | 0 |
| 1.5535 | 0.4465 | -0.0288 | 0.0023 | 0.0030 | 0.0001 | 0 | 0 | 0 | 0 |
| 1.3613 | 0.3892 | -0.0268 | 0.0005 | 0.0006 | -0.0000 | 0 | 0 | 0 | 0 |
| -1.3705 | -0.3854 | 0.0318 | 0.0045 | 0.0057 | 0.0002 | 0 | 0 | 0 | 0 |
| 3.8704 | 1.1033 | -0.0786 | -0.0012 | -0.0015 | -0.0000 | 0 | 0 | 0 | 0 |
| AU2(:, :, 2) = | | | | | | | | | |
| 1.0244 | -0.0103 | -0.6647 | 0.0089 | 0.0001 | 0.0001 | 0 | 0 | 0 | 0 |
| 0.0693 | 0.0003 | -0.0502 | 0.0057 | 0.0000 | 0.0001 | 0 | 0 | 0 | 0 |
| 0.6576 | -0.0048 | -0.4367 | 0.0155 | 0.0001 | 0.0002 | 0 | 0 | 0 | 0 |
| -0.9063 | 0.0121 | 0.5721 | 0.0076 | -0.0000 | 0.0002 | 0 | 0 | 0 | 0 |
| 0.8039 | -0.0076 | -0.5240 | 0.0093 | 0.0000 | 0.0002 | 0 | 0 | 0 | 0 |
| 0.7077 | -0.0080 | -0.4546 | 0.0017 | 0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| -0.7226 | 0.0120 | 0.4439 | 0.0179 | 0.0000 | 0.0003 | 0 | 0 | 0 | 0 |
| 2.0172 | -0.0247 | -1.2857 | -0.0049 | 0.0001 | -0.0000 | 0 | 0 | 0 | 0 |

AU2(:, :, 3) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---|---|---|---|
| 1.8251 | -1.1522 | 0.1666 | 0.0018 | 0.0005 | -0.0008 | 0 | 0 | 0 | 0 |
| 0.1274 | -0.0857 | 0.0137 | 0.0011 | 0.0003 | -0.0005 | 0 | 0 | 0 | 0 |
| 1.1793 | -0.7545 | 0.1115 | 0.0030 | 0.0009 | -0.0013 | 0 | 0 | 0 | 0 |
| -1.6025 | 0.9957 | -0.1401 | 0.0014 | 0.0005 | -0.0006 | 0 | 0 | 0 | 0 |
| 1.4340 | -0.9077 | 0.1318 | 0.0019 | 0.0006 | -0.0007 | 0 | 0 | 0 | 0 |
| 1.2574 | -0.7892 | 0.1130 | 0.0004 | 0.0001 | -0.0002 | 0 | 0 | 0 | 0 |
| -1.2683 | 0.7757 | -0.1061 | 0.0034 | 0.0010 | -0.0015 | 0 | 0 | 0 | 0 |
| 3.5760 | -2.2345 | 0.3175 | -0.0008 | -0.0003 | 0.0004 | 0 | 0 | 0 | 0 |

AU2(:, :, 4) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---|---|---|---|
| 1.5011 | 0.6601 | 0.3112 | -0.0024 | 0.0030 | 0.0000 | 0 | 0 | 0 | 0 |
| 0.1067 | 0.0477 | 0.0250 | -0.0016 | 0.0019 | 0.0000 | 0 | 0 | 0 | 0 |
| 0.9735 | 0.4297 | 0.2073 | -0.0042 | 0.0052 | 0.0001 | 0 | 0 | 0 | 0 |
| -1.3122 | -0.5745 | -0.2632 | -0.0021 | 0.0026 | 0.0000 | 0 | 0 | 0 | 0 |
| 1.1803 | 0.5193 | 0.2460 | -0.0026 | 0.0032 | 0.0001 | 0 | 0 | 0 | 0 |
| 1.0325 | 0.4533 | 0.2115 | -0.0004 | 0.0006 | 0.0000 | 0 | 0 | 0 | 0 |
| -1.0340 | -0.4508 | -0.2006 | -0.0050 | 0.0061 | 0.0001 | 0 | 0 | 0 | 0 |
| 2.9328 | 1.2859 | 0.5953 | 0.0014 | -0.0016 | -0.0000 | 0 | 0 | 0 | 0 |

Tensor $\mathcal{A} \times_3 (U^{(3)})^T$

AU3(:, :, 1) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1.2615 | 0.5720 | 1.2945 | -0.2017 | 1.1739 | -0.5200 | 0.4585 | 1.5631 | -1.2591 | -1.0097 |
| 0.0896 | 0.0387 | 0.0904 | -0.0151 | 0.0841 | -0.0362 | 0.0310 | 0.1097 | -0.0860 | -0.0703 |
| 0.8180 | 0.3673 | 0.8363 | -0.1323 | 0.7626 | -0.3360 | 0.2943 | 1.0112 | -0.8098 | -0.6520 |
| -1.1030 | -0.5060 | -1.1366 | 0.1739 | -1.0241 | 0.4566 | -0.4057 | -1.3704 | 1.1113 | 0.8871 |
| 0.9919 | 0.4489 | 1.0172 | -0.1590 | 0.9233 | -0.4085 | 0.3599 | 1.2283 | -0.9884 | -0.7932 |
| 0.8678 | 0.3952 | 0.8918 | -0.1380 | 0.8068 | -0.3583 | 0.3168 | 1.0763 | -0.8691 | -0.6958 |
| -0.8693 | -0.4033 | -0.8996 | 0.1352 | -0.8055 | 0.3614 | -0.3235 | -1.0831 | 0.8840 | 0.7026 |
| 2.4651 | 1.1262 | 2.5365 | -0.3906 | 2.2906 | -1.0188 | 0.9030 | 3.0597 | -2.4754 | -1.9792 |

AU3(:, :, 2) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| -0.0544 | -0.6940 | 0.5013 | -0.6427 | 0.3274 | 0.1812 | -0.2878 | -0.6885 | -0.4889 | 0.0150 |
| -0.0053 | -0.0513 | 0.0363 | -0.0471 | 0.0242 | 0.0124 | -0.0217 | -0.0513 | -0.0360 | 0.0016 |
| -0.0380 | -0.4539 | 0.3265 | -0.4196 | 0.2140 | 0.1165 | -0.1891 | -0.4511 | -0.3195 | 0.0107 |
| 0.0432 | 0.6005 | -0.4361 | 0.5574 | -0.2835 | -0.1599 | 0.2478 | 0.5947 | 0.4237 | -0.0115 |
| -0.0435 | -0.5466 | 0.3945 | -0.5060 | 0.2578 | 0.1422 | -0.2269 | -0.5425 | -0.3850 | 0.0121 |
| -0.0362 | -0.4756 | 0.3442 | -0.4408 | 0.2244 | 0.1251 | -0.1969 | -0.4715 | -0.3352 | 0.0099 |
| 0.0307 | 0.4685 | -0.3420 | 0.4358 | -0.2213 | -0.1273 | 0.1922 | 0.4630 | 0.3310 | -0.0079 |
| -0.1001 | -1.3471 | 0.9764 | -1.2493 | 0.6358 | 0.3563 | -0.5568 | -1.3348 | -0.9499 | 0.0272 |

AU3(:, :, 3) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 0.1430 | 0.1121 | -0.0495 | -0.1482 | 0.3005 | 0.0899 | -0.0555 | -0.1790 | 0.4697 | -0.3954 |
| 0.0128 | 0.0092 | -0.0019 | -0.0125 | 0.0215 | 0.0092 | -0.0028 | -0.0140 | 0.0334 | -0.0331 |
| 0.0978 | 0.0752 | -0.0290 | -0.0999 | 0.1951 | 0.0636 | -0.0338 | -0.1186 | 0.3046 | -0.2660 |
| -0.1169 | -0.0941 | 0.0482 | 0.1235 | -0.2624 | -0.0701 | 0.0520 | 0.1526 | -0.4106 | 0.3304 |
| 0.1136 | 0.0887 | -0.0382 | -0.1175 | 0.2363 | 0.0719 | -0.0432 | -0.1414 | 0.3694 | -0.3132 |
| 0.0960 | 0.0760 | -0.0355 | -0.1002 | 0.2066 | 0.0594 | -0.0392 | -0.1220 | 0.3231 | -0.2676 |
| -0.0858 | -0.0711 | 0.0419 | 0.0926 | -0.2065 | -0.0487 | 0.0436 | 0.1172 | -0.3236 | 0.2486 |
| 0.2676 | 0.2133 | -0.1040 | -0.2808 | 0.5866 | 0.1634 | -0.1135 | -0.3442 | 0.9177 | -0.7505 |

AU3(:, :, 4) =

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Tensor $\mathcal{A} \times_1 (U^{(1)})^T \times_2 (U^{(2)})^T$

| | | | | | | | | | |
|-----------------|---------|---------|---------|---------|---------|---|---|---|---|
| AU12(:, :, 1) = | | | | | | | | | |
| -5.4490 | -1.5555 | 0.1091 | -0.0000 | -0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| 0.0081 | -0.0073 | -0.0073 | -0.0073 | -0.0092 | -0.0003 | 0 | 0 | 0 | 0 |
| 0.0000 | -0.0001 | 0.0000 | -0.0001 | 0.0000 | 0.0001 | 0 | 0 | 0 | 0 |
| -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| 0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AU12(:, :, 2) = | | | | | | | | | |
| -2.8366 | 0.0334 | 1.8147 | 0.0002 | -0.0001 | 0.0000 | 0 | 0 | 0 | 0 |
| 0.0190 | -0.0058 | 0.0178 | -0.0291 | -0.0001 | -0.0004 | 0 | 0 | 0 | 0 |
| 0.0001 | 0.0000 | -0.0000 | -0.0001 | -0.0000 | 0.0001 | 0 | 0 | 0 | 0 |
| -0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| 0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AU12(:, :, 3) = | | | | | | | | | |
| -5.0339 | 3.1522 | -0.4495 | -0.0001 | 0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| 0.0109 | 0.0232 | -0.0106 | -0.0056 | -0.0017 | 0.0024 | 0 | 0 | 0 | 0 |
| -0.0001 | -0.0001 | -0.0000 | -0.0000 | 0.0001 | 0.0001 | 0 | 0 | 0 | 0 |
| 0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| -0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AU12(:, :, 4) = | | | | | | | | | |
| -4.1308 | -1.8123 | -0.8422 | -0.0001 | 0.0001 | 0.0000 | 0 | 0 | 0 | 0 |
| -0.0020 | -0.0056 | -0.0170 | 0.0080 | -0.0098 | -0.0002 | 0 | 0 | 0 | 0 |
| 0.0000 | -0.0001 | 0.0000 | -0.0001 | 0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| -0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Tensor $\mathcal{A} \times_1 (U^{(1)})^T \times_3 (U^{(3)})^T$

| | | | | | | | | | |
|-----------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| AU13(:, :, 1) = | | | | | | | | | |
| -3.4719 | -1.5837 | -3.5705 | 0.5511 | -3.2272 | 1.4342 | -1.2698 | -4.3078 | 3.4821 | 2.7858 |
| -0.0012 | 0.0104 | 0.0078 | 0.0048 | -0.0054 | -0.0032 | 0.0087 | 0.0056 | -0.0186 | -0.0072 |
| 0.0000 | -0.0000 | 0.0001 | -0.0000 | 0.0000 | 0.0002 | 0.0001 | -0.0001 | -0.0000 | 0.0000 |
| -0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | 0.0000 |
| -0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AU13(:, :, 2) = | | | | | | | | | |
| 0.1428 | 1.8999 | -1.3761 | 1.7615 | -0.8966 | -0.5011 | 0.7859 | 1.8832 | 1.3395 | -0.0389 |
| 0.0082 | 0.0124 | -0.0046 | 0.0091 | -0.0055 | 0.0027 | 0.0076 | 0.0145 | 0.0076 | -0.0030 |
| -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0001 | 0.0001 |
| -0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 |
| 0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

AU13(:, :, 3) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| -0.3803 | -0.3021 | 0.1443 | 0.3981 | -0.8264 | -0.2337 | 0.1584 | 0.4864 | -1.2926 | 1.0636 |
| -0.0155 | -0.0075 | -0.0093 | 0.0116 | -0.0010 | -0.0160 | -0.0064 | 0.0076 | -0.0006 | 0.0292 |
| 0.0000 | -0.0001 | -0.0001 | -0.0001 | -0.0000 | 0.0000 | -0.0001 | -0.0000 | 0.0000 | -0.0000 |
| 0.0000 | -0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 |
| -0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

AU13(:, :, 4) =

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Tensor $\mathcal{A} \times_2 (U^{(2)})^T \times_3 (U^{(3)})^T$

AU23(:, :, 1) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---|---|---|---|
| -3.2468 | 0.0013 | -0.0039 | -0.0041 | -0.0034 | 0.0003 | 0 | 0 | 0 | 0 |
| -0.2272 | 0.0009 | -0.0025 | -0.0025 | -0.0022 | 0.0002 | 0 | 0 | 0 | 0 |
| -2.0989 | 0.0022 | -0.0067 | -0.0071 | -0.0059 | 0.0006 | 0 | 0 | 0 | 0 |
| 2.8490 | 0.0011 | -0.0033 | -0.0033 | -0.0029 | 0.0002 | 0 | 0 | 0 | 0 |
| -2.5513 | 0.0014 | -0.0041 | -0.0042 | -0.0036 | 0.0002 | 0 | 0 | 0 | 0 |
| -2.2363 | 0.0002 | -0.0008 | -0.0008 | -0.0007 | 0.0001 | 0 | 0 | 0 | 0 |
| 2.2534 | 0.0026 | -0.0078 | -0.0081 | -0.0069 | 0.0006 | 0 | 0 | 0 | 0 |
| -6.3591 | -0.0006 | 0.0021 | 0.0021 | 0.0018 | -0.0002 | 0 | 0 | 0 | 0 |

AU23(:, :, 2) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---|---|---|---|
| 0.0020 | 1.4440 | 0.0008 | -0.0017 | 0.0021 | 0.0007 | 0 | 0 | 0 | 0 |
| 0.0013 | 0.1065 | 0.0005 | -0.0011 | 0.0013 | 0.0004 | 0 | 0 | 0 | 0 |
| 0.0035 | 0.9439 | 0.0013 | -0.0029 | 0.0036 | 0.0012 | 0 | 0 | 0 | 0 |
| 0.0017 | -1.2506 | 0.0006 | -0.0015 | 0.0018 | 0.0005 | 0 | 0 | 0 | 0 |
| 0.0022 | 1.1371 | 0.0008 | -0.0018 | 0.0022 | 0.0007 | 0 | 0 | 0 | 0 |
| 0.0004 | 0.9899 | 0.0002 | -0.0003 | 0.0004 | 0.0001 | 0 | 0 | 0 | 0 |
| 0.0041 | -0.9765 | 0.0014 | -0.0034 | 0.0042 | 0.0013 | 0 | 0 | 0 | 0 |
| -0.0010 | 2.8044 | -0.0003 | 0.0009 | -0.0011 | -0.0004 | 0 | 0 | 0 | 0 |

AU23(:, :, 3) =

| | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---|---|---|---|
| -0.0049 | 0.0026 | -0.7535 | 0.0086 | -0.0011 | 0.0002 | 0 | 0 | 0 | 0 |
| -0.0030 | 0.0017 | -0.0577 | 0.0055 | -0.0008 | 0.0002 | 0 | 0 | 0 | 0 |
| -0.0084 | 0.0044 | -0.4966 | 0.0150 | -0.0020 | 0.0004 | 0 | 0 | 0 | 0 |
| -0.0040 | 0.0021 | 0.6462 | 0.0073 | -0.0011 | 0.0003 | 0 | 0 | 0 | 0 |
| -0.0050 | 0.0027 | -0.5944 | 0.0090 | -0.0012 | 0.0003 | 0 | 0 | 0 | 0 |
| -0.0010 | 0.0005 | -0.5147 | 0.0016 | -0.0002 | 0.0000 | 0 | 0 | 0 | 0 |
| -0.0096 | 0.0051 | 0.4995 | 0.0174 | -0.0024 | 0.0006 | 0 | 0 | 0 | 0 |
| 0.0025 | -0.0013 | -1.4541 | -0.0047 | 0.0007 | -0.0001 | 0 | 0 | 0 | 0 |

AU23(:, :, 4) =

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$$\text{Tenzor } \mathcal{S} = \mathcal{A} \times_1 (U^{(1)})^T \times_2 (U^{(2)})^T \times_3 (U^{(3)})^T$$

| | | | | | | | | | |
|--------------|---------|---------|---------|---------|---------|---|---|---|---|
| S(:, :, 1) = | | | | | | | | | |
| 8.9522 | -0.0000 | -0.0000 | 0.0001 | 0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| -0.0162 | -0.0042 | 0.0126 | 0.0132 | 0.0111 | -0.0010 | 0 | 0 | 0 | 0 |
| 0.0000 | 0.0001 | -0.0000 | 0.0001 | -0.0001 | -0.0001 | 0 | 0 | 0 | 0 |
| 0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| S(:, :, 2) = | | | | | | | | | |
| -0.0000 | -3.9550 | -0.0001 | 0.0000 | 0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| -0.0066 | -0.0239 | -0.0024 | 0.0055 | -0.0068 | -0.0021 | 0 | 0 | 0 | 0 |
| 0.0001 | -0.0000 | 0.0000 | -0.0001 | -0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| -0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| 0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| S(:, :, 3) = | | | | | | | | | |
| 0.0000 | -0.0001 | 2.0534 | 0.0003 | -0.0001 | 0.0000 | 0 | 0 | 0 | 0 |
| 0.0157 | -0.0083 | 0.0246 | -0.0281 | 0.0039 | -0.0009 | 0 | 0 | 0 | 0 |
| 0.0001 | 0.0001 | -0.0000 | -0.0000 | -0.0000 | 0.0001 | 0 | 0 | 0 | 0 |
| -0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 | 0 | 0 | 0 | 0 |
| 0.0000 | 0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| S(:, :, 4) = | | | | | | | | | |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$$\text{Tuckerovo j\u00e1dro } \mathcal{S}_A \in \mathbb{R}^{5 \times 6 \times 3}$$

| | | | | | |
|---------------|---------|---------|---------|---------|---------|
| SA(:, :, 1) = | | | | | |
| 8.9522 | -0.0000 | -0.0000 | 0.0001 | 0.0000 | -0.0000 |
| -0.0162 | -0.0042 | 0.0126 | 0.0132 | 0.0111 | -0.0010 |
| 0.0000 | 0.0001 | -0.0000 | 0.0001 | -0.0001 | -0.0001 |
| 0.0000 | 0.0000 | 0.0000 | -0.0000 | 0.0000 | -0.0000 |
| 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | -0.0000 |
| SA(:, :, 2) = | | | | | |
| -0.0000 | -3.9550 | -0.0001 | 0.0000 | 0.0000 | 0.0000 |
| -0.0066 | -0.0239 | -0.0024 | 0.0055 | -0.0068 | -0.0021 |
| 0.0001 | -0.0000 | 0.0000 | -0.0001 | -0.0000 | -0.0000 |
| -0.0000 | 0.0000 | -0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | 0.0000 |
| SA(:, :, 3) = | | | | | |
| 0.0000 | -0.0001 | 2.0534 | 0.0003 | -0.0001 | 0.0000 |
| 0.0157 | -0.0083 | 0.0246 | -0.0281 | 0.0039 | -0.0009 |
| 0.0001 | 0.0001 | -0.0000 | -0.0000 | -0.0000 | 0.0001 |
| -0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | -0.0000 |
| 0.0000 | 0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 |

Ověření přesnosti $\|\mathcal{A} - \mathcal{S}_{\mathcal{A}} \times_1 U_A^{(1)} \times_2 U_A^{(2)} \times_3 U_A^{(3)}\|$

```
> norm(A(:)-kron(kron(U3(:,1:r3),U2(:,1:r2)),U1(:,1:r1))*SA(:))  
ans =  
9.7434e-15
```

C Aproximace tenzoru z přílohy B tenzorem nižšího ranku

Spočteme několik aproximací tenzoru \mathcal{A} tenzorem hodnosti (t_1, t_2, t_3) , spočteme normu chyby a porovnáme s odhadem (4.2) z věty 7. Výpočet aproximace a chyby lze v MATLABu provést následujícím způsobem:

```
> truncSA = SA(1:t1,1:t2,1:t3);
> approxA = reshape(kron(kron(U3(:,1:t3),U2(:,1:t2)),U1(:,1:t1))*truncSA(:),[n1,n2,n3]);
> norm(A(:)-approxA(:))
```

Tabulka C.1 obsahuje singulární čísla rozvoju $A^{[\ell]}$, $\ell = 1, 2, 3$, tenzoru \mathcal{A} , tabulka C.2 pak požadované hodnoty aproximací, skutečné hodnoty aproximací, odhady (4.2) a spočtené skutečné chyby aproximací. Z tabulky C.2 je vidět, že při postupu popsaném ve větě 7 může dojít ke snížení hodnoty více, než jsme zamýšleli. Dále vidíme, že (4.2) je skutečně pouze horním odhadem chyby aproximace. V některých případech však může nastat rovnost. Speciálně se tak stane tehdy, pokud chce hodnost redukovat jen v jednom módu. Pak se nerovnost (4.2) redukuje na rovnost a věta 7 na Schmidtovu–Eckartovu–Youngovu–Mirského větu (věta 2). Z toho, že se při aproximaci popsané větou 7 může snížit hodnost v některých módech více než jsme zamýšleli, lze tušit, že obecně mohou být aproximací zasažena různá singulární čísla v různých rozvojiích. Tento jev ilustrují obrázky C.1, C.2 a C.3.

Tabulka C.1: Singulární čísla $\sigma_j^{(\ell)}$, $j = 1, \dots, n_\ell$, jednotlivých rozvoju $A^{[\ell]}$, $\ell = 1, 2, 3$, tenzoru $\mathcal{A} \in \mathbb{R}^{8 \times 10 \times 4}$ hodnosti (5, 6, 3).

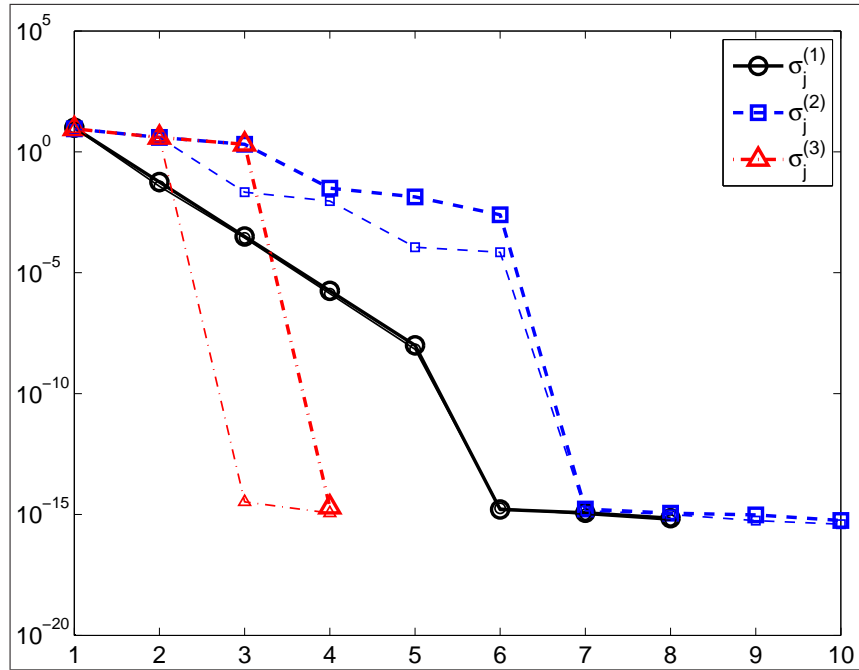
| ℓ | Singulární čísla $\sigma_j^{(\ell)}$ | | | | | |
|--------|--------------------------------------|--------|------------------------|------------------------|------------------------|--------|
| 1 | 10.0000 | 0.0562 | $3.1623 \cdot 10^{-4}$ | $1.7783 \cdot 10^{-6}$ | $1.0000 \cdot 10^{-8}$ | |
| 2 | 8.9522 | 3.9551 | 2.0536 | 0.0315 | 0.0136 | 0.0025 |
| 3 | 8.9522 | 3.9551 | 2.0538 | | | |

¹Přestože vektorový rank je v podstatě vektor celých čísel a zápis $v > w$ pro vektory v a $w \in \mathbb{Z}^k$

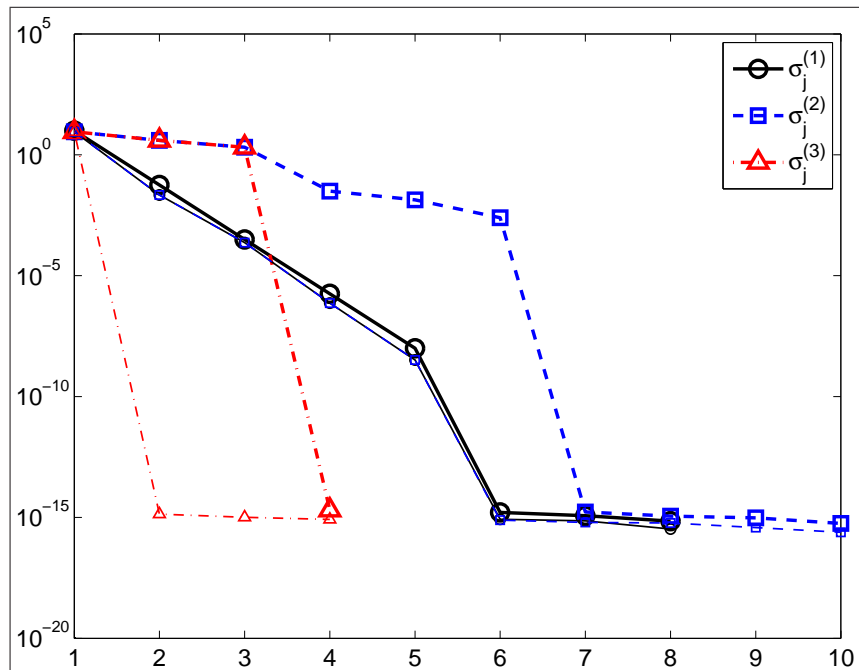
Tabulka C.2: Tenzor \mathcal{A} hodnosti $(5, 6, 3)$ aproximujeme tenzorem hodnosti (t_1, t_2, t_3) užitím věty 7, získaný aproximant \mathcal{B} má však obecně hodnost (ρ_1, ρ_2, ρ_3) . V případě, že chceme hodnost tenzoru snížit jen v jednom módu, odhad (4.2) se redukuje na Schmidtovu–Eckartovu–Youngovu–Mirského větu (věta 2) a je přesný, tj. identický (\equiv) s chybou. Viz také poznámku pod čarou na str. 71 a obrázky C.1, C.2 a C.3.

| $(t_1, t_2, t_3) \geq (\rho_1, \rho_2, \rho_3)$ | Odhad (4.2) \geq | Chyba $\ \mathcal{A} - \mathcal{B}\ $ |
|---|------------------------|---|
| $(4, 6, 3) = (4, 6, 3)$ | $1.0000 \cdot 10^{-8}$ | $\equiv 1.0000 \cdot 10^{-8} (\sigma_5^{(1)})$ |
| $(3, 6, 3) = (3, 6, 3)$ | $1.7783 \cdot 10^{-6}$ | $\equiv 1.7783 \cdot 10^{-6}$ |
| $(2, 6, 3) = (2, 6, 3)$ | $3.1623 \cdot 10^{-4}$ | $\equiv 3.1623 \cdot 10^{-4}$ |
| $(1, 6, 3) > (1, 3, 3)^1$ | 0.0562 | $\equiv 0.0562$ |
| $(5, 5, 3) = (5, 5, 3)$ | 0.0025 | $\equiv 0.0025 (\sigma_6^{(2)})$ |
| $(5, 4, 3) = (5, 4, 3)$ | 0.0138 | $\equiv 0.0138$ |
| $(5, 3, 3) = (5, 3, 3)$ | 0.0344 | $\equiv 0.0344$ |
| $(5, 2, 3) = (5, 2, 3)$ | 2.0539 | $\equiv 2.0539$ |
| $(5, 1, 3) > (3, 1, 3)$ | 4.4566 | $\equiv 4.4566$ |
| $(5, 6, 2) = (5, 6, 2)$ | 2.0538 | $\equiv 2.0538 (\sigma_3^{(3)}; \text{obr. C.1})$ |
| $(5, 6, 1) > (5, 5, 1)$ | 4.4565 | $\equiv 4.4565 (\text{obr. C.2})$ |
| $(3, 3, 3) = (3, 3, 3)$ | 0.0344 | $\gtrsim 0.0344$ |
| $(2, 2, 2) = (2, 2, 2)$ | 2.9046 | > 2.0540 |
| $(1, 1, 1) = (1, 1, 1)$ | 6.3028 | > 4.4566 |
| $(5, 2, 2) > (4, 2, 2)$ | 2.9046 | $> 2.0540 (\text{obr. C.3})$ |

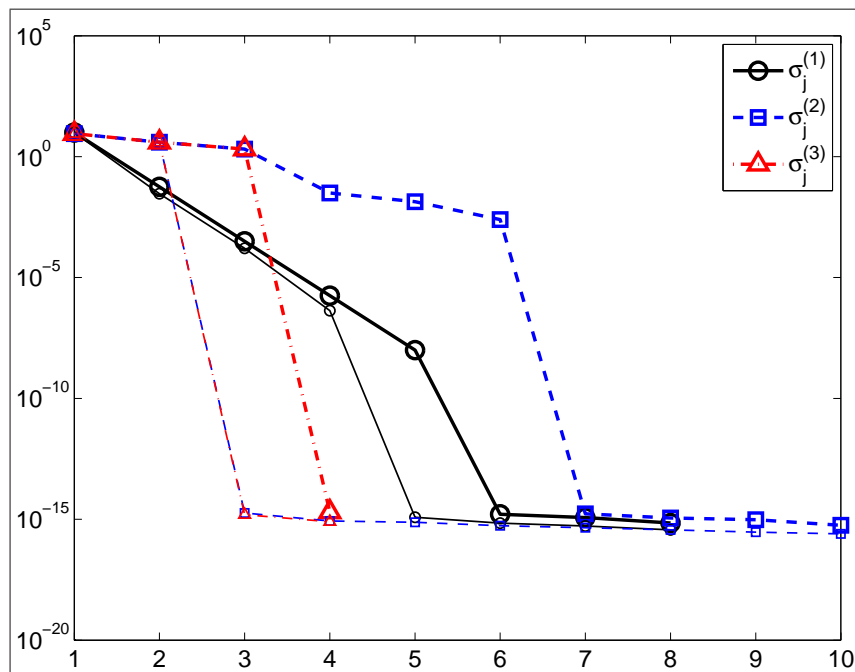
téměř výhradně znamená, že $v_j > w_j$, tj. každá složka vektoru v je ostře větší než odpovídající složka vektoru w , v tabulce C.2 používáme symbol „ $>$ “ v nepatrně jiném významu. Např. zápis $(1, 6, 3) > (1, 3, 3)$ používáme pro to, že se jedná o *hodnosti* tenzorů a abychom zdůraznili fakt, že hodnost tenzoru, který jsme získali, (vpravo) je ostře menší než hodnost požadovaná (vlevo).



Obrázek C.1: Singulární čísla jednotlivých rozvoju $A^{[\ell]}$, $\ell = 1, 2, 3$, tenzoru \mathcal{A} (tlusté čáry) a jeho aproximace tenzorem hodnosti (5, 6, 2) (tenké čáry).



Obrázek C.2: Singulární čísla jednotlivých rozvoju $A^{[\ell]}$, $\ell = 1, 2, 3$, tenzoru \mathcal{A} (tlusté čáry) a jeho aproximace tenzorem hodnosti (5, 6, 1) (tenké čáry).



Obrázek C.3: Singulární čísla jednotlivých rozvoju $A^{[\ell]}$, $\ell = 1, 2, 3$, tenzoru \mathcal{A} (tlusté čáry) a jeho aproximace tenzorem hodnosti $(5, 2, 2)$ (tenké čáry).

Literatura

- [1] I. Babuška, K. Rektorys, F. Vyčichlo: *Matematická teorie rovinné pružnosti*, Nakladatelství Československé akademie věd, Praha, 1955.
- [2] B. W. Bader, T. G. Kolda: *Algorithm 862: MATLAB tensor classes for fast algorithm prototyping*, ACM Transactions on Mathematical Software 32(4) (2006), str. 635–653.
- [3] B. W. Bader, T. G. Kolda: *Efficient MATLAB computations with sparse and factored tensors*, SIAM Journal on Scientific Computing 30(1) (2007), str. 205–231.
- [4] B. W. Bader, T. G. Kolda: *MATLAB Tensor Toolbox, version 2.6.*, 2015. Dostupné na: <http://www.sandia.gov/~tgkolda/TensorToolbox>.
- [5] Å. Björck: *Numerical methods in matrix computations*, Springer Verlag, Heidelberg, 2015.
- [6] L. Boček: *Tenzorový počet*, SNTL, Státní nakladatelství technické literatury (edice MS, Matematický seminář, č. 9), Praha, 1976.
- [7] M. Brdička, L. Samek, B. Sopko: *Mechanika kontinua*, Academia a Česká matice technická, Praha, 2000.
- [8] J. D. Carroll, J. J. Chang: *Analysis of individual differences in multidimensional scaling via an n -way generalization of “Eckart–Young” decomposition*, Psychometrika 35(3) (1970), str. 283–319.
- [9] P. Comon, G. Golub, L.-H. Lim, B. Mourrain: *Symmetric tensors and symmetric tensor rank*, SIAM Journal on Matrix Analysis and Applications 30(3) (2008), str. 1254–1279.
- [10] L. De Lathauwer, B. De Moor, J. Vandewalle: *A multilinear singular value decomposition*, SIAM Journal on Matrix Analysis and Applications 21(4) (2000), str. 1253–1278.
- [11] Z. Dostál, V. Vondrák: *Lineární algebra*, Vysoká škola báňská – Technická univerzita Ostrava, Fakulta elektrotechniky a informatiky, Ostrava, 2001.
- [12] J. Duintjer Tebbens, I. Hnětynková, M. Plešinger, Z. Strakoš, P. Tichý: *Analýza metod pro maticové výpočty: základní metody*, Matfyzpress, Praha, 2012.

- [13] C. Eckart, G. Young: *The approximation of the matrix by another of lower rank*, Psychometrika 1(3) (1936), str. 211–218.
- [14] L. Eldén: *Matrix methods in data mining and pattern recognition*, SIAM Publications, Philadelphia, 2007.
- [15] S. Etter: *Parallel tensor-formated numerics for the chemical master equation*, Master Thesis, ETH Zürich, Zürich, 2015. Dostupné na: <http://www.sam.math.ethz.ch/teaching/termproj/etterMSC.pdf>.
- [16] M. Fiedler: *Speciální matice a jejich použití v numerické matematice*, SNTL, Státní nakladatelství technické literatury (edice TKI, Teoretická knižnice inženýra), Praha, 1981.
- [17] L. Grasedyck: *Hierarchical singular value decomposition of tensors*, SIAM Journal on Matrix Analysis and Applications 31(4) (2010), str. 2029–2054.
- [18] L. Grasedyck, D. Kressner, C. Tobler: *A literature survey of low-rank tensor approximation techniques*, GAMM Mitteilungen 36(1) (2013), str. 53–78.
- [19] W. Hackbush, S. Kühn: *A new scheme for the tensor representation*, Journal of Fourier Analysis and Applications 5(5) (2009), str. 706–722.
- [20] P. C. Hansen: *The truncated SVD as a method for regularization*, BIT Numerical Mathematics 27(4) (1987), str. 534–553.
- [21] R. A. Harshman: *Foundations of the ParaFac procedure: Models and conditions for an “explanatory” multi-modal factor analysis*, UCLA Working Papers in Phonetics 16(1) (1970), str. 1–84.
- [22] F. L. Hitchcock: *The expression of a tensor or a polyadic as a sum of products*, Journal of Mathematics and Physics 6(1) (1927), str. 164–189.
- [23] F. L. Hitchcock: *Multilple invariants and generalized rank of a p -way matrix or tensor*, Journal of Mathematics and Physics 7(1) (1927), str. 39–79.
- [24] S. Kästner: *Vektoren, Tensoren, Spinoren*, Akademie Verlag, Berlin, 1960.
- [25] C. G. Khatri, C. R. Rao: *Solutions to some functional equations and their applications to characterization of probability distributions*, Sankhya: The Indian Journal of Statistics 30, Series A, Pt. 2 (1968), str. 167–180.
- [26] H. A. L. Kiers: *Towards a standardized notation and terminology in multiway analysis*, Journal of Chemometrics 14(3) (2000), str. 105–122.
- [27] N. A. Kilčevskij: *Základy tensorového počtu a jeho použití v mechanice*, SNTL, Státní nakladatelství technické literatury, Praha, 1956.
- [28] T. G. Kolda, B. W. Bader: *Tensor decompositions and applications*, SIAM Review 51(3) (2009), str. 455–500.

- [29] D. Kressner, M. Plešinger, C. Tobler: *A preconditioned low-rank CG method for parameter-dependent Lyapunov matrix equations*, Numerical Linear Algebra with Applications 21(5) (2014), str. 666–684.
- [30] D. Kressner, C. Tobler: *htucker — A MATLAB toolbox for tensors in hierarchical Tucker format*, Technical Report 2012-02, Seminar for Applied Mathematics, ETH Zurich, 2012. Dostupné na: <http://anchp.epfl.ch/htucker>.
- [31] K. Kuchař: *Základy obecné teorie relativity*, Academia, Praha, 1968.
- [32] S. Liu: *Matrix results on the Khatri–Rao and Tracy–Singh products*, Linear Algebra and its Applications 289(1–3) (1999), str. 267–277.
- [33] L. Mirsky: *Symmetric gauge functions and unitarily invariant norms*, The Quarterly Journal of Mathematics 11(1) (1960), str. 50–59.
- [34] P. Paatero: *Construction and analysis of degenerate ParaFac models*, Journal of Chemometrics 14(3) (2000), str. 285–299.
- [35] V. Pereyra, G. Scherer: *Efficient computer manipulation of tensor products with applications to multidimensional approximation*, Mathematics of Computation 27(123) (1973), str. 595–605.
- [36] E. Schmidt: *Zur Theorie der linearen un nichtlinearen Integralgleichungen. I. Tiel. Entwicklung willkürlichen Funkcionen nach System vorgechriebener*, Mathematische Annalen 63 (1907), str. 433–476.
- [37] M. Schuermans: *Weighted low rank approximation: Algorithms and applications*, PhD Thesis, KU Leuven, Leuven, 2006.
- [38] R. Sikorski: *Diferenciální a integrální počet. Funkce více proměnných*, Academia, Praha, 1973.
- [39] L. Sorber, M. Van Barel, L. De Lathauwer: *Tensorlab v2.0*, 2014. Dostupné na: <http://www.tensorlab.net>.
- [40] G. W. Stewart, J.-G. Sun: *Matrix perturbation theory*, Academic Press, Boston, 1990.
- [41] C. Tobler: *Low-rank tensor methods for linear systems and eigenvalue problems*, PhD Thesis, ETH Zürich, Zürich, 2012. Dostupné na: <http://sma.epfl.ch/~ctobler/diss.pdf>.
- [42] D. S. Tracy, R. P. Singh: *A new matrix product and its applications in matrix differentiation*, Statistica Neerlandica 26(4) (1972), str. 143–157.
- [43] L. R. Tucker: *Implications of factor analysis of three-way matrices for measurement of change*, v Problems in Measuring Change, C. W. Harris (editor), University of Wisconsin Press, 1963, str. 122–137.

- [44] L. R. Tucker: *The extension of factor analysis to three-dimensional matrices*, v Contributions to Mathematical Psychology, H. Gulliksen, N. Frederiksen (editoři), Holt, Rinehardt & Winston, New York, 1964, str. 110–127.
- [45] L. R. Tucker: *Some mathematical notes on three-mode factor analysis*, Psychometrika 31(3) (1966), str. 279–311.
- [46] D. A. Turkington: *Generalized vectorization, cross-products, and matrix calculus*, Cambridge University Press, Cambridge, 2013.
- [47] E. Tyrtysnikov: *Numerical methods with tensor representations of data*. Přednáška letní školy „Summer Supercomputing Academy“, 2012. Dostupné na: [http://academy2012.hpc-russia.ru/files/ \(...\)](http://academy2012.hpc-russia.ru/files/...)
(...) lectures/algebra/0703_1_tyrtysnikov.pdf.
- [48] V. Votruba: *Základy speciální teorie relativity*, Academia, Praha, 1969.