فرهنگ و اندیشهٔ ریاضی
سال ۴۰، شمارهٔ ۶۹ (پاییز و زمستان ۱۴۰۰) صص. ۶۱ تا ۷۸

رتبهٔ تانسور و مسئلهٔ بهترین تقریب رتبهٔ پایین

زهرا اردولالو، عفت گلپر رابوکی، و نظام‌الدین مهدوی امیری

چکیده

رتبه یکی از مشخصه‌های مهم هر ماتریس است. رتبهٔ ماتریس A عبارت است از کوچک‌ترین عدد صحیح r به‌طوری‌که A را بتوان با استفاده از مجموع r ماتریس رتبه‌یک نوشت. رتبهٔ ماتریس را می‌توان با استفاده از روش حذف گاوسی یا تجزیهٔ پلکانی به‌دست آورد. رتبهٔ تانسور A عبارت است از کوچک‌ترین عدد صحیح r به‌طوری‌که A به‌وسیلهٔ مجموع r تانسور رتبه‌یک ایجاد شود. برخلاف رتبهٔ ماتریس، رتبهٔ تانسور به‌راحتی قابل محاسبه نیست، چنان‌که به‌جز در مواردی خاص، مسئله‌ای NP−سخت است. تاکنون مطالعات گسترده‌ای در زمینهٔ محاسبه رتبهٔ تانسورهای $۲ \times ۲ \times ۲$ انجام گرفته است و چندین روش شامل محاسبهٔ ابردترمینان، بررسی ساختار درونی تانسور و نیز طبقه‌بندی تانسور به صورتهای کانونی ارائه شده‌اند. یک مسئلهٔ اساسی در کار با ماتریس‌ها و تانسورها، محاسبهٔ بهترین تقریب رتبهٔ پایین است. طبق قضیهٔ اِکارت−یانگ، بهترین تقریب رتبهٔ k در ماتریس‌ها با مجموع k جملهٔ اول از تجزیهٔ مقدار تکین قابل محاسبه است. به‌علاوه، برای یک ماتریس، محاسبهٔ بهترین تقریب رتبهٔ $k+۱$ با استفاده از بهترین تقریب رتبهٔ k، امکان‌پذیر است. اما برخلاف ماتریس‌ها، ممکن است بسیاری از تانسورها تقریب رتبهٔ پایین مشخص‌شده‌ای نداشته باشند. این مسئله‌ای تبهگن است و این تانسورها را می‌توان با دنباله‌ای از تانسورهای رتبهٔ پایین، به‌قدر کافی نزدیک، تقریب زد.

۱. مقدمه

در بسیاری از مسائل داده‌کاوی در علوم و مهندسی، داده‌هایی عددی مطرح هستند که معمولاً با ماتریس‌ها و تانسورها نمایش داده می‌شوند [۸]. تانسورها آرایه‌هایی چندبعدی هستند و با تعمیم بردارها

عبارات و کلمات کلیدی. جبر چندخطی عددی، تانسور، رتبهٔ تانسور، تقریب رتبهٔ پایین، تجزیهٔ تانسور.

و ماتریس‌ها به ابعاد بالاتر تعریف می‌شوند. تانسورها را اولین‌بار تولّیو لِوی[۱] و جرجیو ریتچی[۲] مطرح کردند. مفهوم رتبهٔ تانسور و تجزیه‌های تانسوری را هیچکاک[۳] در سال ۱۹۲۷ مطرح کرد [۳]. تجزیه‌های ماتریسی و تانسوری مهم‌ترین عملیاتی هستند که به محاسبهٔ رتبه می‌انجامد. تجزیهٔ CP[۴]، یک تانسور را به‌صورتِ مجموع تانسورهای رتبه‌یک بیان می‌کند و عملکردی شبیه به تجزیهٔ مقدار تکین (SVD) دارد. البته، محاسبه تجزیهٔ CP مستلزم روش‌های تکراری است که در نتیجهٔ آن، تقریبی از رتبهٔ تانسور به‌دست می‌آید. با توجه به پیچیدگی مسئلهٔ محاسبهٔ رتبهٔ تانسور، دسته‌های خاصی از آن‌ها، به‌ویژه، تانسورهای $۲ \times ۲ \times ۲$، بررسی شده‌اند. چند روش برای محاسبهٔ رتبهٔ تانسورهای $۲ \times ۲ \times ۲$ ارائه شده است. یک روش به بررسی ابردترمینان می‌پردازد که یک چندجمله‌ای مبتنی بر درایه‌های تانسور است. روش دیگر یک دسته‌بندی ارائه می‌کند به‌طوری‌که هر تانسور، با استفاده از تبدیل‌های ناتکین در یکی از این دسته‌ها قرار می‌گیرد. یک مسئلهٔ اساسی در کار با ماتریس و تانسور، محاسبهٔ بهترین تقریب رتبهٔ پایین است. برخلاف ماتریس، یک تانسور ممکن است تقریب رتبهٔ پایین با رتبهٔ مشخص‌شده‌ای نداشته باشد. حتی الگوریتمی که بهترین تقریب رتبه پایین را، در صورت وجود، برای تانسورها محاسبه کند، در دسترس نیست. ازاین‌رو، عموماً یک تقریب رتبهٔ پایین با دقت مناسب محاسبه می‌شود. روش‌های تکراری مانند تجزیهٔ CP برای محاسبهٔ یک تقریب رتبهٔ پایین با دقتی مناسب به‌کار می‌روند. در این مقاله، پس از بررسی رتبه در تانسورها، نخست عدم وجود بهترین تقریب رتبهٔ پایین را مطرح می‌کنیم و در ادامه راهکاری را برای محاسبهٔ یک تقریب رتبهٔ پایین مناسب مرور می‌کنیم.

۲. مفاهیم اولیه

فرض کنید $V_۱$، ...، V_n و U فضاهای برداری باشند. نگاشت φ از $V_۱ \times \cdots \times V_n$ به U یک نگاشت چندخطی (n–خطی) است اگر برای هر $v_i, v'_i \in V_i$ که در آن $i = ۱, \dots, n$ و برای هر $a \in \mathbb{R}$ داشته باشیم:

$$\varphi(v_۱, \dots, v_i + v'_i, \dots, v_n) = \varphi(v_۱, \dots, v_i, \dots, v_n) + \varphi(v_۱, \dots, v'_i, \dots, v_n),$$
$$\varphi(v_۱, \dots, a v_i, \dots, v_n) = a\varphi(v_۱, \dots, v_i, \dots, v_n).$$

تعریف ۱.۲. فرض کنید $V_۱$، ...، V_n و U فضاهای برداری و $\varphi : V_۱ \times \cdots \times V_n \longrightarrow U$ یک نگاشت چندخطی است. زوج (φ, U) را «حاصل‌ضرب تانسوری» $V_۱$، ...، V_n می‌نامیم. دراین‌صورت، U با $V_۱ \otimes \cdots \otimes V_n$ و $\varphi(v_۱, \dots, v_n)$ با $v_۱ \otimes \cdots \otimes v_n$ نشان داده می‌شوند [۷].

[۱]Tullio Levi [۲]Gergorio Ricci [۳]Hitchckock [۴]CANDECOMP/PARAFAC

عناصر $V_1 \otimes \cdots \otimes V_n$ تانسور نامیده می‌شوند. در این مقاله، $V_i = \mathbb{R}^{d_i}$ و ازاین‌رو، تانسورهای حقیقی فضای $\mathbb{R}^{d_n} \otimes \cdots \otimes \mathbb{R}^{d_1}$ بررسی می‌شوند. تانسور $\mathcal{A} \in \mathbb{R}^{d_1 \times \cdots \times d_n}$ از مرتبهٔ n یک آرایهٔ nگانه است. اسکالر، یک تانسور مرتبهٔ صفر، بردار، یک تانسور مرتبهٔ یک و ماتریس، یک تانسور مرتبهٔ دو است [۱]. در شکل ۱، نمایشی از اسکالر، بردار، ماتریس، و یک تانسور مرتبهٔ سه ارائه شده است.

شکل ۱. تانسورهایی از مرتبهٔ صفر تا سه

تانسورها از بردارها و ماتریس‌ها تشکیل می‌شوند. یک تار[۱] را می‌توان یک بردار یا قطعه‌ای تک‌بعدی از تانسور درنظر گرفت که با ثابت نگه‌داشتن همهٔ ابعاد به‌جز یکی از آن‌ها شکل می‌گیرد، در حالی‌که برش[۲]، یک ماتریس یا یک قطعه دو بعدی از تانسور است که با ثابت نگه‌داشتن همه ابعاد به‌جز دوتای آن‌ها به‌دست می‌آید. شکل ۲ تارها و برش‌های مختلف یک تانسور مرتبهٔ سه را نشان می‌دهد.

برای کار با تانسورها معمولاً تانسور را به یک صورتِ ماتریسی تبدیل می‌کنند.

تعریف ۲.۲ (صورت ماتریسی یک تانسور). صورتِ ماتریسی تانسور $\mathcal{A} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_n}$ نسبت به وجه iام که با $A_{(i)}$ نشان داده می‌شود، ماتریسی با اندازهٔ $d_i \times (d_1 \ldots d_{i-1} d_{i+1} \ldots d_n)$ است که از چینش مجدد تارهای یک تانسور ساخته می‌شود [۱].

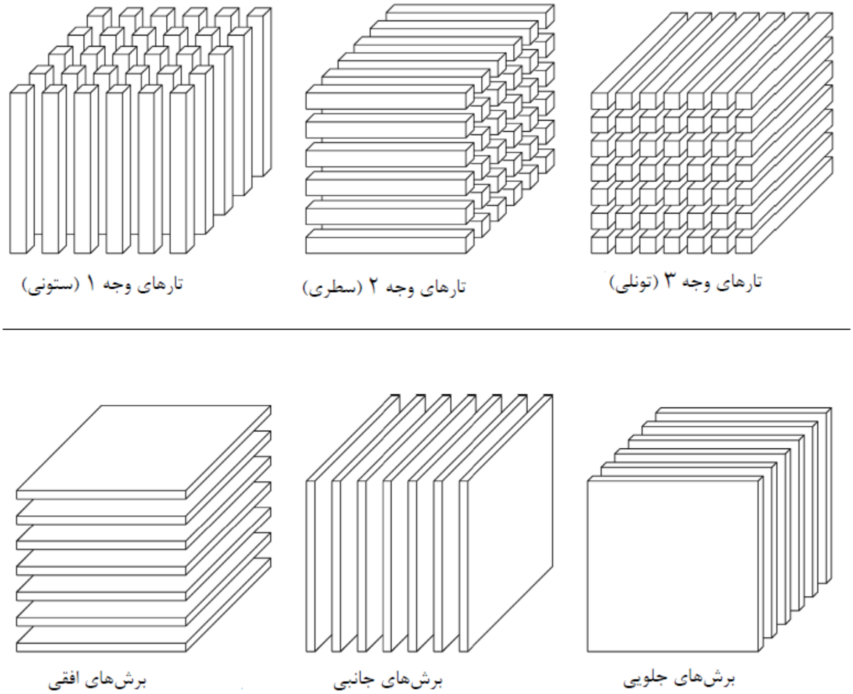
در شکل ۳ صورت ماتریسی یک تانسور $3 \times 4 \times 2$ نسبت به هر سه وجه آن نشان داده شده است.

۳. مفهوم رتبه

مفهوم رتبه در تانسورها را نخستین‌بار هیچکاک[۳] در سال ۱۹۲۷ مطرح کرد [۳]. پنجاه سال بعد، کراسکِل[۴] تعریف دیگری از تانسور را مستقلاً ارائه کرد [۵].

تعریف رتبهٔ تانسور شبیه به تعریف رتبهٔ ماتریس است، اما ویژگی‌های متفاوتی دارد. ابتدا به بیان مفهوم رتبهٔ تانسور می‌پردازیم و سپس برخی از این تفاوت‌ها را بیان می‌کنیم.

[۱]fiber [۲]slice [۳]Hitchcock [۴]Kruskal



شکل ۲. ردیف بالا: سه نوع تار از تانسور $\mathcal{A}$، ردیف پایین: سه نوع برش از تانسور $\mathcal{A}$ [۴]

۱.۳. رتبهٔ ماتریس. رتبهٔ ماتریس $A \in \mathbb{R}^{m\times n}$ عبارت است از تعداد سطرها یا ستون‌های مستقل خطی آن. یک تعریف دیگر برای رتبهٔ ماتریس به این صورت است: رتبهٔ ماتریس عبارت است از کمترین تعداد ماتریس‌های رتبه‌یک که مجموع آن‌ها ماتریس A را تولید کند، یعنی

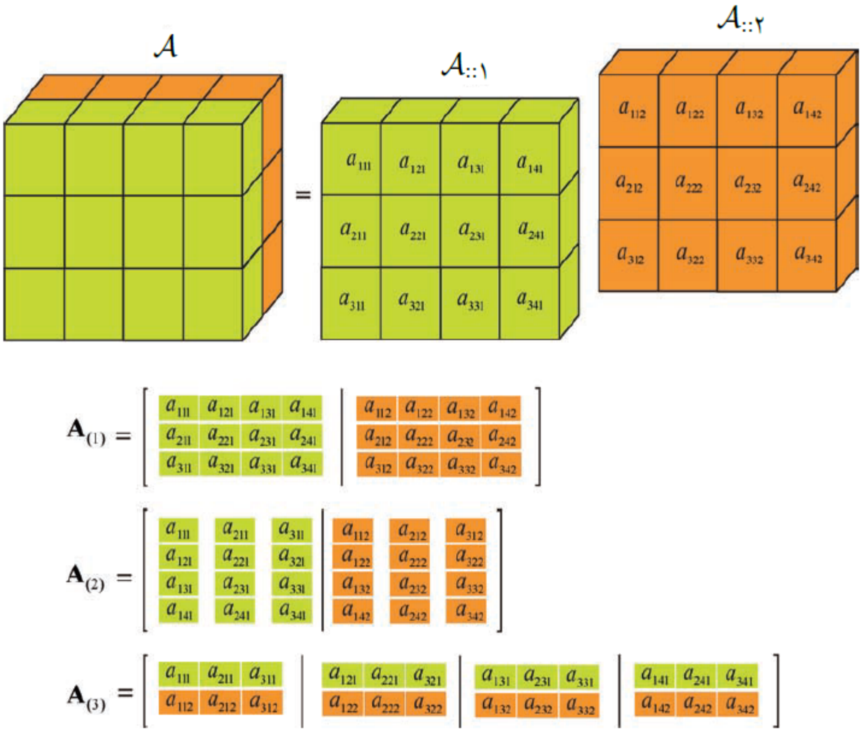
$$\operatorname{rank}(A) = \min\{r : A = \Sigma_{i=1}^{r} a_i \otimes b_i ،\text{به‌ازای برخی بردارهای } a_i \text{ و } b_i\},$$

که در آن $\otimes$ نماد حاصل‌ضرب خارجی است.

۲.۳. رتبهٔ تانسور. در این بخش رتبهٔ تانسورها و رتبهٔ چندخطی[۱] را تعریف می‌کنیم [۹].

تعریف ۱.۳. تانسور $\mathcal{A} \in \mathbb{R}^{d_1\times d_2\times\cdots\times d_n}$ تانسوری رتبه‌یک نامیده می‌شود اگر بتوان آن را به‌صورت حاصل‌ضرب خارجی n بردار نوشت، یعنی $\mathcal{A} = x_1 \otimes x_2 \otimes \cdots \otimes x_n$ که در آن $x_i \in \mathbb{R}^{d_i}$. تانسور $\mathcal{A}$ دارای رتبهٔ r است اگر بتوان آن‌را به‌صورت مجموع r تانسور رتبه‌یک نوشت و $\mathcal{A}$ را نتوان با کمتر از

[۱]multilinear rank



شکل ۳. صورت‌های ماتریسی یک تانسور [۱]

r تانسور رتبه‌یک تولید کرد. به‌بیان‌دیگر،

$$\operatorname{rank}(\mathcal{A}) = \min\left\{r : \mathcal{A} = \sum_{i=۱}^{r} x_۱^{(i)} \otimes \cdots \otimes x_n^{(i)}\right\}.$$

تعریف ۲.۳. رتبهٔ چندخطی تانسور $\mathcal{A} \in \mathbb{R}^{d_۱\times\cdots\times d_n}$ که با $\operatorname{mrank}(\mathcal{A})$ یا $\operatorname{rank}_{\boxplus}(\mathcal{A})$ نشان داده می‌شود، یک nتایی شامل رتبهٔ صورت ماتریسی تانسور نسبت به هر وجه آن است و به‌صورتِ $\operatorname{rank}_{\boxplus}(\mathcal{A}) = (r_۱(\mathcal{A}), \ldots, r_n(\mathcal{A}))$ نمایش داده می‌شود، که در آن $r_i(\mathcal{A})$ رتبهٔ $A_{(i)}$ است.

مثال ۳.۳. فرض کنید $\mathcal{A}_{::۱}$ و $\mathcal{A}_{::۲}$ به‌صورتِ زیر برش‌های جلویی تانسور $\mathcal{A}$ باشند

$$\mathcal{A}_{::۱} = \begin{pmatrix} ۰ & ۱ \\ ۰ & ۱ \end{pmatrix}, \quad \mathcal{A}_{::۲} = \begin{pmatrix} ۱ & ۰ \\ ۱ & ۰ \end{pmatrix}$$

در این‌صورت

$$\operatorname{rank}(\mathcal{A}) = ۲, \quad \operatorname{rank}_{\boxplus}(\mathcal{A}) = (۱, ۲, ۲).$$

قضیه ۴.۳ ([۹]). *فرض کنید* $\mathcal{A} \in \mathbb{R}^{d_1\times\cdots\times d_n}$*، در این صورت*

$$r_i(\mathcal{A}) \leq \min\{\operatorname{rank}(\mathcal{A}), d_i\}, \quad i = 1, \ldots, n.$$

تانسور $\mathcal{A} = (a_{i_1\cdots i_n}) \in \mathbb{R}^{d_1\times\cdots\times d_n}$ و ماتریس‌های $L_k = (\ell_{ij}^{(k)}) \in \mathbb{R}^{c_i\times d_i}$ را درنظر می‌گیریم. حاصل‌ضرب چندخطی $\mathcal{A}$ در L_1، ...، L_n که به‌صورتِ $(L_1, \ldots, L_n)\cdot\mathcal{A}$ نشان داده می‌شود، تانسوری مانند $\mathcal{A}' = (a'_{i_1\cdots i_n}) \in \mathbb{R}^{c_1\times\cdots\times c_n}$ است که درایه‌های آن به‌صورت

$$a'_{i_1\cdots i_n} = \sum_{j_1,\ldots,j_n=1}^{d_1,\ldots,d_n} \ell_{i_1j_1}^{(1)}\cdots\ell_{i_nj_n}^{(n)} a_{j_1\cdots j_n}$$

تعریف می‌شوند. نتایج زیر در مورد رتبهٔ تانسور و رتبهٔ چندخطی اهمیت بالایی دارند [۹].

قضیه ۵.۳. *فرض کنید* $\mathcal{A} \in \mathbb{R}^{d_1\times\cdots\times d_n}$ *و* $L_k \in \mathbb{R}^{c_i\times d_i}$*. دراین‌صورت*

$$\operatorname{rank}((L_1, \ldots, L_n)\cdot\mathcal{A}) \leq \operatorname{rank}(\mathcal{A}),$$
$$\operatorname{rank}_{\boxplus}((L_1, \ldots, L_n)\cdot\mathcal{A}) \leq \operatorname{rank}_{\boxplus}(\mathcal{A}).$$

و اگر L_i*ها ماتریس‌هایی ناتکین باشند، آنگاه*

$$\operatorname{rank}((L_1, \ldots, L_n)\cdot\mathcal{A}) = \operatorname{rank}(\mathcal{A}),$$
$$\operatorname{rank}_{\boxplus}((L_1, \ldots, L_n)\cdot\mathcal{A}) = \operatorname{rank}_{\boxplus}(\mathcal{A}).$$

با داشتن بردارهای مستقل خطی می‌توان تانسورهایی با رتبهٔ بالا به‌صورتِ زیر تعریف کرد.

لم ۶.۳. *فرض کنید برای* $1 \leq i \leq n$ *بردارهای* $x_1^{(i)}, \ldots, x_r^{(i)} \in \mathbb{R}^{d_i}$ *مستقل خطی باشند. دراین‌صورت، رتبهٔ تانسور* $\mathcal{A} = \sum_{j=1}^{r} x_j^{(1)}\otimes\cdots\otimes x_j^{(n)}$ *برابر با* r *است.*

رتبهٔ یک ماتریس و رتبهٔ یک تانسور ویژگی‌های متفاوتی دارند که به برخی از آن‌ها اشاره می‌کنیم.

- رتبهٔ یک تانسور با مقادیر حقیقی ممکن است روی میدان اعداد حقیقی و مختلط متفاوت باشد، درحالی‌که رتبهٔ ماتریس‌ها روی این میدان‌ها یکسان است. برای مثال، فرض کنید $x_1, y_1 \in \mathbb{R}^l$، $x_2, y_2 \in \mathbb{R}^m$ و $x_3, y_3 \in \mathbb{R}^p$ دوبه‌دو مستقل خطی باشند [۷]. برای $k = 1, 2, 3$ قرار دهید $z_k = x_k + iy_k$ و تانسور $\mathcal{A}$ را به‌صورتِ زیر تعریف کنید

$$\mathcal{A} = \frac{1}{2}((\overline{z}_1\otimes z_2\otimes\overline{z}_3) + (\overline{z}_1\otimes\overline{z}_2\otimes z_3))$$

دراین‌صورت، داریم

$$\operatorname{rank}_{\mathbb{C}}(\mathcal{A}) = ۲ < ۳ = \operatorname{rank}_{\mathbb{R}}(\mathcal{A}).$$

- الگوریتم مشخصی برای تعیین رتبهٔ تانسور وجود ندارد. درحقیقت، تعیین رتبهٔ تانسور مسئله‌ای NP–سخت است، درحالی‌که، رتبهٔ ماتریس با روش‌های گوناگون مانند حذف گاوسی و تجزیهٔ مقدار تکین در زمانیْ چندجمله‌ای نسبت به اندازهٔ ماتریس به‌دست می‌آید.
- ویژگی متفاوت دیگر تانسورها مربوط به ماکسیمم‌رتبه است. ماکسیمم‌رتبه، بزرگ‌ترین رتبه‌ای است که برای مجموعه‌ای از تانسورها در فضای $\mathbb{R}^{d_۱\times\cdots\times d_n}$ در دست است. برای ماتریس $A \in \mathbb{R}^{I\times J}$، این رتبه برابر است با $\min\{J, I\}$، درحالی‌که برای تانسور $\mathcal{A} \in \mathbb{R}^{d_۱\times\cdots\times d_n}$، ماکسیمم‌رتبه ممکن است بزرگ‌تر از $\min\{d_۱, \dots, d_n\}$ باشد.
- رتبهٔ قابل تعریف دیگری برای تانسورها رتبهٔ نوعی[۱] است. رتبهٔ نوعی هر رتبه‌ای است که احتمال وقوع آن روی یک مجموعه با اندازه مثبت، بزرگ‌تر از صفر باشد [۴]. برای مجموعه‌ای ازماتریس‌های $I \times J$ رتبهٔ نوعی و ماکسیمم‌رتبه برابر است با $\min(I, J)$، درحالی‌که در تانسورها معمولاً رتبهٔ نوعی و ماکسیمم‌رتبه با هم تفاوت دارند [۴].

۴. تعیین رتبه در تانسورهای ۲ × ۲ × ۲

هرچند برای محاسبهٔ دقیق رتبهٔ تانسورها روش خاصی وجود ندارد، نتایجی برای محاسبهٔ رتبهٔ دسته‌های مشخصی از تانسورها مانند تانسورهای ۲ × ۲ × ۲ ارائه شده است.

فرض کنید برش‌های جلویی تانسور $\mathcal{T}$ به‌صورت زیر در دست باشند

$$\mathcal{T}_{::۱} = \begin{bmatrix} t_{۱۱۱} & t_{۱۲۱} \\ t_{۲۱۱} & t_{۲۲۱} \end{bmatrix}, \qquad \mathcal{T}_{::۲} = \begin{bmatrix} t_{۱۱۲} & t_{۱۲۲} \\ t_{۲۱۲} & t_{۲۲۲} \end{bmatrix}$$

برای تانسور $\mathcal{T}$ ابردترمینان[۲] $\triangle(\mathcal{T})$ را که یک چندجمله‌ای به‌صورتِ زیر است، تعریف می‌کنیم

$$\begin{aligned}\triangle(\mathcal{T}) = {} & (t_{۱۱۱}^۲ t_{۲۲۲}^۲ + t_{۱۱۲}^۲ t_{۲۲۱}^۲ + t_{۱۲۱}^۲ t_{۲۱۲}^۲ + t_{۱۲۲}^۲ t_{۲۱۱}^۲) - ۲(t_{۱۱۱}t_{۱۱۲}t_{۲۲۱}t_{۲۲۲} \\ & + t_{۱۱۱}t_{۱۲۱}t_{۲۱۲}t_{۲۲۲} + t_{۱۱۱}t_{۱۲۲}t_{۲۱۱}t_{۲۲۲} + t_{۱۱۲}t_{۱۲۱}t_{۲۱۲}t_{۲۲۱} + t_{۱۱۲}t_{۱۲۲}t_{۲۲۱}t_{۲۱۱} \\ & + t_{۱۲۱}t_{۱۲۲}t_{۲۱۲}t_{۲۱۱}) + ۴(t_{۱۱۱}t_{۱۲۲}t_{۲۱۲}t_{۲۲۱} + t_{۱۱۲}t_{۱۲۱}t_{۲۱۱}t_{۲۲۲}).\end{aligned}$$

کراسکِل با بررسی این چندجمله‌ای نشان داد که رتبهٔ یک تانسور ۲ × ۲ × ۲ حداکثر ۳ است، و دسته‌بندی زیر را ارائه کرد [۸، ۹، ۱۰].

---

[۱] typical rank [۲] hyperdeterminant

قضیه ۱.۴. *فرض کنید $\triangle(\mathcal{T})$ ابردترمینان تانسور $\mathcal{T}$ از بعد ۲ × ۲ × ۲ باشد. دراین‌صورت*

(۱) *اگر $\triangle(\mathcal{T}) > 0$، آنگاه $\operatorname{rank}(\mathcal{T}) = 2$.*
(۲) *اگر $\triangle(\mathcal{T}) < 0$، آنگاه $\operatorname{rank}(\mathcal{T}) = 3$.*
(۳) *اگر $\triangle(\mathcal{T}) = 0$، آنگاه $0 \leq \operatorname{rank}(\mathcal{T}) \leq 3$.*

چون حالت خاص $\triangle(\mathcal{T}) = 0$ رتبهٔ تانسور را به‌صورتِ دقیق مشخص نمی‌کند، تن برگ[1] با بررسی برش‌های مختلف یک تانسور دسته‌بندی جدیدی برای تانسورهای ۲ × ۲ × ۲ به‌دست آورد. این دسته‌بندی در قضیهٔ زیر آمده است [۱۰].

قضیه ۲.۴. *فرض کنید $\mathcal{T}$ تانسوری ۲ × ۲ × ۲ باشد. دراین‌صورت*

(۱) *تانسور $\mathcal{T}$ رتبهٔ صفر دارد اگر و تنها اگر $\mathcal{T}$ تانسور صفر باشد.*
(۲) *تانسور $\mathcal{T}$ رتبهٔ یک دارد اگر و تنها اگر همهٔ ۶ برش ممکن $\mathcal{T}$ تکین باشند.*
(۳) *تانسور $\mathcal{T}$ رتبهٔ ۲ یا بیشتر دارد اگر و تنها اگر $\mathcal{T}$ دست‌کم یک برش ناتکین داشته باشد.*

در حالت سوم که تانسور $\mathcal{T}$ دست‌کم یک برش ناتکین دارد، بررسی ویژه‌مقدارهای یک ماتریس خاص نتیجه‌بخش است [۱۰].

قضیه ۳.۴. *فرض کنید $\mathcal{T}$ تانسوری ۲ × ۲ × ۲ و T_1 و T_2 برش‌های جلویی آن باشند به‌طوری‌که T_1 ناتکین است. دراین‌صورت*

(۱) *اگر $T_2T_1^{-1}$ دو ویژه‌مقدار حقیقی و متمایز داشته (و درنتیجه قطری‌شدنی) باشد، آنگاه رتبهٔ $\mathcal{T}$ برابر با ۲ است.*
(۲) *اگر $T_2T_1^{-1}$ دست‌کم دو ویژه‌مقدار مختلط داشته باشد، آنگاه رتبهٔ $\mathcal{T}$ برابر با ۳ است.*
(۳) *اگر $T_2T_1^{-1}$ دو ویژه‌مقدار حقیقی داشته (اما قطری‌ناشدنی) باشد، آنگاه رتبهٔ $\mathcal{T}$ برابر با ۳ است.*

۵. تعیین رتبه برای تانسورهای $L \times M \times N$

برای تعیین رتبهٔ تانسورهای $p \times p \times 2$ روشِ تن برگ، در قضیهٔ زیر، بسیار سودمند است [۱۰].

قضیه ۱.۵. *فرض کنید $\mathcal{T} \in \mathbb{R}^{p \times p \times 2}$ و T_1 و T_2 برش‌های جلویی از آن باشند به‌طوری‌که T_1 ناتکین است. دراین‌صورت*

[1] Ten Berge

(۱) *اگر ماتریس $T_2T_1^{-1}$ دارای p ویژه‌مقدار حقیقی و متمایز (و درنتیجه قطری‌شدنی) باشد، آنگاه* $\operatorname{rank}(\mathcal{T}) = p$.

(۲) *اگر $T_2T_1^{-1}$ دست‌کم دو ویژه‌مقدار مختلط داشته باشد، آنگاه* $\operatorname{rank}(\mathcal{T}) \geq p + 1$.

(۳) *اگر $T_2T_1^{-1}$ دارای p ویژه‌مقدار حقیقی (ولی قطری‌ناشدنی) باشد، آنگاه* $\operatorname{rank}(\mathcal{T}) \geq p + 1$.

(۴) *اگر $T_2T_1^{-1}$ دارای k جفت مزدوج ویژه‌مقدار مختلط باشد، آنگاه* $\operatorname{rank}(\mathcal{T}) \leq p + k$.

کراسکِل نشان داد که برای تانسورهای $p \times p \times 2$، ماکسیمم‌رتبهٔ ممکن برای pی زوج برابر با $3p/2$، و برای pی فرد برابر با $(3p-1)/2$ است [۵].

ملاحظه ۲.۵. رتبهٔ تانسور $p \times p \times 2$ با دست‌کم یک برش $p \times p$ ناتکین برابر p است. اگر همهٔ برش‌های ممکن $\mathcal{T}$ تکین باشند، آنگاه رتبهٔ $\mathcal{T}$ برابر با ۱ است.

محاسبهٔ رتبهٔ تانسورهای $L \times M \times N$، با شرط $N \neq 2$، و نیز تانسورهای از مرتبهٔ بالاتر به‌سادگی امکان‌پذیر نیست و تاکنون راه‌حل مناسبی برای آن ارائه نشده است. اگرچه هیچ الگوریتم ویژه‌ای برای محاسبهٔ رتبهٔ این نوع تانسورها وجود ندارد، می‌توان تقریبی از رتبه را با استفاده از روش تکراری تجزیهٔ CP محاسبه کرد. تجزیهٔ CP اغلب با استفاده از الگوریتم کمترین مربعات متناوب[1] محاسبه می‌شود که ماتریس‌های عامل را در هر مرحله محاسبه و بهنگام می‌کند. یک تجزیهٔ CP با m مولفه از تانسور $\mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$، با حل مسئلهٔ مینیمم‌سازی $\min_{\mathcal{G}} \|\mathcal{T} - \mathcal{G}\|$ که در آن

$$\mathcal{G} = \sum_{k=1}^{m} \lambda_k \, a_k \otimes b_k \otimes c_k$$

هم‌ارز است [۴].

۱.۵. محاسبه رتبه در متلب. در نرم‌افزار متلب، الگوریتم کمترین مربعات متناوب با استفاده از جعبهٔ ابزار تانسور[2] و با دستور `[P]=cp_als(T,n)` اجرا می‌شود. این دستور، تانسور $\mathcal{T}$ را به حاصل‌ضرب یک تانسورِ هسته و ماتریس‌های عامل تجزیه می‌کند و در سلول P قرار می‌دهد. دستور `full(P)` عوامل تجزیه را درهم ضرب می‌کند و تقریبی از تانسور $\mathcal{T}$ را بر اساس تجزیهٔ CP محاسبه می‌کند. برای سنجش نزدیکی تجزیه و تانسور $\mathcal{T}$ معیار زیر را به‌کار می‌گیرند

$$\mathtt{Final-fit} = 1 - \frac{\|\mathcal{T} - \mathtt{full(P)}\|_F}{\|\mathcal{T}\|_F}.$$

[1] alternating least squares (ALS) [2] http://www.sandia.gov/ ~tgkolda/TensorToolbox/index-2.5.html

هرقدر `Final-fit` به ۱ نزدیک‌تر باشد، تقریب بهتری در دست است. با اجرای الگوریتم برای nهای مختلف و با ملاحظهٔ `Final-fit` می‌توان تقریبی از رتبهٔ تانسور را به‌دست آورد. هرگاه به‌ازای nای مقدار `Final-fit` کاملاً به ۱ نزدیک باشد، مقدار آن n تقریباً برابر با رتبهٔ تانسور است [۸].

۶. مسئلهٔ بهترین تقریب رتبهٔ پایین

در مسئلهٔ بهترین تقریب رتبهٔ پایین برای $\mathcal{A} \in \mathbb{R}^{d_1 \times \cdots \times d_n}$، هدفْ محاسبهٔ یک تانسور رتبهٔ k با شرط $k < \operatorname{rank}(\mathcal{A})$ است که از میان همهٔ تانسورها با رتبهٔ نابیشتر از k کمترین فاصله را از $\mathcal{A}$ داشته باشد. به‌عبارت‌دیگر، هدف پیدا کردن تانسوری مانند $\mathcal{B}$ با رتبهٔ نابیشتر از k است به‌گونه‌ای که جواب مسئلهٔ مینیمم‌سازی زیر باشد

$$\min_{\operatorname{rank}(\mathcal{B}) \leq k} \|\mathcal{A} - \mathcal{B}\|.$$

مهم‌ترین نتایج موجود در این زمینه که در [۹] آمده‌اند عبارت است از

(۱) مسئلهٔ بهترین تقریب رتبه k برای بسیاری از kها بدحالت[۱] است؛
(۲) مسئلهٔ بهترین تقریب رتبه k برای بسیاری از تانسورها بدحالت است؛
(۳) جوابهای ضعیف[۲] راهکاری برای مقابله با بدحالتی مسئلهٔ بهترین تقریب رتبهٔ k است.

در ادامه، به بررسی این نتایج می‌پردازیم.

۱.۶. تقریب رتبهٔ پایین و فشرده‌سازی ماتریس.

در برخی ازکاربردهای عملی، اغلب با ماتریسی بسیار بزرگ سروکار داریم و می‌خواهیم آن‌را با یک ماتریس از رتبهٔ پایین‌تر، که فاصلهٔ کمی با آن دارد، تقریب بزنیم. برای این‌کار می‌توان تجزیهٔ مقدار تکین را به‌کار برد [۲].

قضیه ۱.۶ (تجزیهٔ مقدار تکین). *هر ماتریس A از مرتبهٔ $m \times n$ با شرط $m \geq n$ را می‌توان به‌صورت*

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T$$

تجزیه کرد که در آن $U \in \mathbb{R}^{m \times m}$ و $V \in \mathbb{R}^{n \times n}$ ماتریس‌هایی متعامد و $\Sigma \in \mathbb{R}^{n \times n}$ قطری است:

$$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n), \qquad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

بنابر تجزیهٔ مقدار تکین، ماتریس A را می‌توان به‌صورتِ مجموع حداکثر n ماتریس رتبه‌یک نوشت

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_n u_n v_n^T \tag{۱.۶}$$

[۱] ill-posed [۲] weak solutions

که در آن u_iها و v_iها به‌ترتیب ستون‌های U و V هستند. تجزیهٔ مقدار تکین، بهترین تقریب رتبهٔ k را برای ماتریس A به‌دست می‌دهد [۲].

قضیه ۲.۶ (اِکارت–یانگ). *گیریم A یک ماتریس $m \times n$ باشد و $k < \operatorname{rank}(A)$. دراین‌صورت*

$$\min_{\operatorname{rank}(B)=k} \|A-B\|_2 = \|A-A_k\|_2 = \sigma_{k+1}$$

که درآن A_k مجموع k جملهٔ اول در (۱.۶) است.

۲.۶. تقریب رتبهٔ پایین برای تانسور. همان‌گونه‌که بیان شد، برای ماتریسی با رتبهٔ r، بهترین تقریب رتبهٔ k با درنظر گرفتن k جملهٔ اول از مجموع (۱.۶) به‌دست می‌آید. به‌علاوه، بهترین تقریب رتبهٔ $k+1$ را می‌توان با استفاده از بهترین تقریب رتبهٔ k محاسبه کرد. تعمیم این مطلب به تانسورها به‌سادگی امکان‌پذیر نیست و در بسیاری از موارد تانسور مورد نظر بهترین تقریب رتبهٔ پایین با رتبهٔ مشخص شده ندارد و در صورت وجود چنین تقریبی، تقریب‌های رتبهٔ بالاتر را نمی‌توان با استفاده از تقریب‌های رتبهٔ پایین‌تر بهنگام کرد. نتیجهٔ زیر این مطلب را بیان می‌کند [۹].

قضیه ۳.۶. *فرض‌کنید $n \geq 3$ و به‌ازای هر i که $1 \leq i \leq n$، $d_i \geq 2$. در این صورت برای هر r که $2 \leq r \leq \min\{d_1, \ldots, d_n\}$، تانسور $\mathcal{A} \in \mathbb{R}^{d_1 \times \cdots \times d_n}$ با رتبهٔ r و ثابت $s < r$ وجود دارند به‌طوری‌که $\mathcal{A}$ بهترین تقریب رتبهٔ s ندارد.*

برای بررسی بیشتر این موضوع روی فضای $\mathbb{R}^{d_1 \times \cdots \times d_n}$، زیرفضای توپولوژیکی $\mathcal{S}_r$ به‌صورتِ

$$\mathcal{S}_r(d_1, \ldots, d_n) = \{\mathcal{A} \in \mathbb{R}^{d_1 \times \cdots \times d_n} : \operatorname{rank}(\mathcal{A}) \leq r\}$$

و بستار آن با نماد $\overline{\mathcal{S}}_r(d_1, \ldots, d_n)$ را درنظر بگیرید. این زیرفضا برای $n = 2$ بسته است، یعنی $\mathcal{S}_r = \overline{\mathcal{S}}_r$. اما برای $n \geq 3$ زیرفضای $\mathcal{S}_r$ لزوماً (یا حتی معمولاً) بسته نیست. علی‌رغم این، برای ماتریس‌ها، مجموعه $\{A \in \mathbb{R}^{m \times n} : \operatorname{rank}(A) \leq r\}$ بسته است. به‌بیان‌دیگر، اگر $\{A_i\}$ دنباله‌ای از ماتریس‌های $m \times n$ باشد که $\operatorname{rank}(A_i) \leq r$ و $A_i \to A$، آنگاه لزوماً $\operatorname{rank}(A) \leq r$. به‌این‌ترتیب، دنباله‌ای از ماتریس‌های رتبهٔ بالا به ماتریسی از همان رتبه یا رتبه‌ای پایین‌تر همگرا می‌شود. این گزاره برای تانسورها برقرار نیست و یک دنباله از تانسورها با رتبهٔ نابیشتر از r وجود دارد که می‌تواند به تانسوری با رتبهٔ بزرگ‌تر از r همگرا شود. مثال زیر را ببینید.

مثال ۴.۶. تانسور $\mathcal{A}_\varepsilon$ را به‌صورتِ زیر درنظر بگیرید

$$\mathcal{A}_\varepsilon = \left(\begin{array}{cc|cc} 0 & 1 & 1 & \frac{1}{\varepsilon} \\ 1 & \frac{1}{\varepsilon} & \frac{1}{\varepsilon} & \frac{1}{\varepsilon^2} \end{array}\right) = \varepsilon \begin{pmatrix} 1 \\ \frac{1}{\varepsilon} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \frac{1}{\varepsilon} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \frac{1}{\varepsilon} \end{pmatrix} - \varepsilon \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

به‌سادگی می‌توان نشان داد که $\text{rank}(\mathcal{A}_\varepsilon) = ۲$، درحالی‌که، دنبالهٔ تانسورهای $\{\mathcal{A}_\varepsilon\}$ به تانسور زیر با رتبهٔ ۳ همگراست

$$\mathcal{A} = \left(\begin{array}{cc|cc} ۰ & ۱ & ۱ & ۰ \\ ۱ & ۰ & ۰ & ۰ \end{array}\right).$$

برای تانسورها، وجود بهترین تقریب رتبهٔ پایین با رتبهٔ مشخص‌شده، تنها برای تانسورهای رتبه‌یک تضمین شده است و دنبالهٔ تانسورهای رتبه‌یک تنها به تانسوری با رتبه‌یک یا صفر همگرا می‌شود. به‌طور خلاصه، هر ماتریس بهترین تقریب رتبهٔ k دارد، و هر تانسور با رتبهٔ دلخواه بهترین تقریب رتبه‌یک دارد. برای تانسورهای با رتبهٔ بزرگ‌تر از یک، گزارهٔ زیر را داریم [۹].

گزاره ۵.۶. *فرض کنید $r \geq ۲$ و $n \geq ۳$. با توپولوژی نرمی مفروض روی $\mathbb{R}^{d_۱\times\cdots\times d_n}$، گزاره‌های زیر هم‌ارزند*

(الف) مجموعهٔ $\mathcal{S}_r(d_۱, \ldots, d_n)$ *بسته نیست.*

(ب) *دنبالهٔ $\{\mathcal{A}_i\}$ از تانسورهای $\mathcal{A}_i \in \mathbb{R}^{d_۱\times\cdots\times d_n}$ با شرط $\text{rank}(\mathcal{A}_i) \leq r$ وجود دارد که به تانسور $\mathcal{B} \in \mathbb{R}^{d_۱\times\cdots\times d_n}$ با شرط $\text{rank}(\mathcal{B}) > r$ همگراست.*

لم ۶.۶. *فرض کنید به‌ازای $i = ۱, ۲, ۳$، $x_i, y_i \in \mathbb{R}^{d_i}$ و تانسور $\mathcal{A}$ به‌صورتِ زیر تعریف شود*

$$\mathcal{A} = x_۱ \otimes x_۲ \otimes y_۳ + x_۱ \otimes y_۲ \otimes x_۳ + y_۱ \otimes x_۲ \otimes x_۳ \in \mathbb{R}^{d_۱\times d_۲\times d_۳}.$$

دراین‌صورت $\text{rank}(\mathcal{A}) = ۳$ اگر و تنها اگر به‌ازای $i = ۱, ۲, ۳$ مجموعهٔ $\{x_i, y_i\}$ مستقل خطی باشد.

در ادامه، رده‌ای از تانسورهای رتبهٔ ۳ را معرفی می‌کنیم که بهترین تقریب رتبهٔ ۲ ندارند [۹].

قضیه ۷.۶. *فرض کنید به‌ازای $i = ۱, ۲, ۳$، بردارهای $x_i, y_i \in \mathbb{R}^{d_i}$ مستقل خطی باشند. دراین‌صورت تانسور*

$$\mathcal{A} = x_۱ \otimes x_۲ \otimes y_۳ + x_۱ \otimes y_۲ \otimes x_۳ + y_۱ \otimes x_۲ \otimes x_۳$$

رتبهٔ ۳ دارد و آن‌را می‌توان به‌طور دلخواه با تانسورهای رتبهٔ ۲ تقریب زد. بنابراین، $\mathcal{A}$ بهترین تقریب رتبهٔ ۲ ندارد.

اثبات. طبق لم ۶.۶، تانسور $\mathcal{A}$ رتبهٔ ۳ دارد. برای $n \in \mathbb{N}$، دنبالهٔ $\mathcal{A}_n$ را به‌صورت

$$\mathcal{A}_n = n(x_۱ + \frac{۱}{n}y_۱) \otimes (x_۲ + \frac{۱}{n}y_۲) \otimes (x_۳ + \frac{۱}{n}y_۳) - nx_۱ \otimes x_۲ \otimes x_۳ \qquad (۲.۶)$$

تعریف کنید. روشن است که $\mathrm{rank}(\mathcal{A}_n) \leq ۲$. ازسوی‌دیگر،

$$\mathcal{A}_n - \mathcal{A} = \frac{۱}{n}(y_۱ \otimes y_۲ \otimes x_۳ + y_۱ \otimes x_۲ \otimes y_۳ + x_۱ \otimes y_۲ \otimes y_۳) + \frac{۱}{n^۲}(y_۱ \otimes y_۲ \otimes y_۳)$$

بنابراین، وقتی $n \to \infty$، داریم $\|\mathcal{A}_n - \mathcal{A}\|_F \to ۰$. ازاین‌رو، $\{\mathcal{A}_n\}$ به $\mathcal{A}$ همگراست و می‌توان $\mathcal{A}$ را به‌طور دلخواه نزدیک با $\mathcal{A}_n$ها تقریب زد. درنتیجه، $\mathcal{A}$ بهترین تقریب رتبهٔ ۲ ندارد. □

در قضیهٔ ۷.۶ دیدیم که یک دنباله از تانسورهای رتبهٔ ۲ می‌تواند به تانسوری از رتبهٔ ۳ همگرا شود. این مفهوم، پرش رتبه[۱] نام دارد و آن‌را می‌توان به تانسورهای با مرتبه‌های بالاتر تعمیم داد [۹].

قضیه ۸.۶. *فرض کنید $n \geq ۳$ و به‌ازای هر i که $۱ \leq i \leq n$، $d_i \geq ۲$. مسئلهٔ یافتن یک بهترین تقریب رتبهٔ k که $k = ۲, \ldots, \min\{d_۱, \ldots, d_n\}$ برای تانسوری در فضای $\mathbb{R}^{d_۱ \times \cdots \times d_n}$ عموماً جواب ندارد. همچنین، تانسور $\mathcal{A} \in \mathbb{R}^{d_۱ \times \cdots \times d_n}$ با $\mathrm{rank}(\mathcal{A}) = k + ۱$ وجود دارد به‌طوری‌که بهترین تقریب رتبهٔ k ندارد. این نتیجه به انتخاب نرم بستگی ندارد.*

۳.۶. واگرایی ضرایب. سوال مهمی که اینجا مطرح می‌شود این است که دلیل عدم موفقیت در رسیدن به یک تقریب رتبهٔ پایین مناسب چیست؟ یک پاسخ این است که هم‌زمان نمی‌توان همهٔ ضرایب یک دنباله از تانسورها را کراندار نگه داشت. برای روشن شدن این موضوع، به مثال زیر توجه می‌کنیم.

مثال ۹.۶. فرض کنید $\mathcal{A}_n$ به‌صورت (۲.۶) باشد. همان‌طورکه می‌بینیم، وقتی $n \to \infty$، جمله‌های برداری همگرا هستند، ولی ضرایب واگراست. در حقیقت، پرشِ رتبه زمانی رخ می‌دهد که برخی از ضرایب دنباله واگرا باشند.

نکتهٔ قابل توجه این است که برای تانسور $\mathcal{A}$ که بهترین تقریب رتبهٔ k ندارد، مینیمم‌سازی عبارت

$$\left\| \mathcal{A} - \sum_{i=۱}^{k} \lambda_۱ \, u_k \otimes v_k \otimes w_k \right\|$$

به واگرا شدن برخی از ضرایب می‌انجامد.

۷. جواب‌های ضعیف

اگر یک تانسور رتبهٔ r حد دنباله‌ای از تانسورهای رتبهٔ پایین‌تر باشد، بهترین تقریب رتبهٔ پایین برای این تانسور وجود ندارد. همان‌گونه‌که بررسی کردیم، دلیل اصلی بسته نبودن زیرفضای $S_r(d_۱, \ldots, d_n)$ است. برای محاسبهٔ یک جواب تقریبی می‌توان از عناصر زیرفضای $\overline{S}_r$ به جای S_r استفاده کرد. درواقع،

[۱] rank jumping

در پی راه‌حلی جایگزین برای محاسبه بهترین تقریب رتبهٔ پایین هستیم. برای تانسورهایی که بهترین تقریب رتبهٔ پایین ندارند، از یک جواب ضعیف استفاده می‌شود. بدین‌منظور، ابتدا مفهوم رتبهٔ مرزی[۱] را تعریف می‌کنیم [۶].

تعریف ۱.۷ (رتبهٔ مرزی). تانسور $\mathcal{A} \in \mathbb{R}^{d_1 \times \cdots \times d_n}$ دارای رتبهٔ مرزی r است اگر

$$\mathcal{A} \in \overline{\mathcal{S}}_r(d_1, \ldots, d_n) \quad \text{و} \quad \mathcal{A} \notin \overline{\mathcal{S}}_{r-1}(d_1, \ldots, d_n)$$

رتبهٔ مرزی تانسور $\mathcal{A}$ را با $\underline{\mathrm{rank}}(\mathcal{A})$ نشان می‌دهیم.

یک تعریف دیگر برای رتبهٔ مرزی به‌صورت زیر است [۷].

$$\underline{\mathrm{rank}}(\mathcal{A}) = \min\left\{r : \inf_{\mathrm{rank}(\mathcal{B}) \leq r} \|\mathcal{A} - \mathcal{B}\| = 0\right\}.$$

روشن است که برای ماتریس $A \in \mathbb{R}^{m \times n}$ داریم $\underline{\mathrm{rank}}(A) = \mathrm{rank}(A)$. بااین‌حال، برای تانسور $\mathcal{A} \in \mathbb{R}^{d_1 \times \cdots \times d_n}$ همواره $\underline{\mathrm{rank}}(\mathcal{A}) \leq \mathrm{rank}(\mathcal{A})$. چون $\overline{\mathcal{S}}_0 = \mathcal{S}_0$ و $\overline{\mathcal{S}}_1 = \mathcal{S}_1$، بنابراین وقتی $\mathrm{rank}(\mathcal{A}) \leq 2$، $\underline{\mathrm{rank}}(\mathcal{A}) = \mathrm{rank}(\mathcal{A})$.

مثال ۲.۷ ([۷]). فرض کنید $e_1 = (1, 0)^T$، $e_2 = (0, 1)^T$ و

$$\mathcal{A} = e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1 = \left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right).$$

دراین‌صورت $\mathrm{rank}(\mathcal{A}) = 3$. اکنون دنبالهٔ تانسوری $\{\mathcal{A}_\varepsilon\}$ را به‌صورت زیر درنظر بگیرید

$$\mathcal{A}_\varepsilon = \frac{1}{\varepsilon}\begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} - \frac{1}{\varepsilon} e_1 \otimes e_1 \otimes e_1 = \left(\begin{array}{cc|cc} 0 & 1 & 1 & \varepsilon \\ 1 & \varepsilon & \varepsilon & \varepsilon^2 \end{array}\right).$$

دراین‌صورت $\lim_{\varepsilon \to 0} \mathcal{A}_\varepsilon = \mathcal{A}$، و چون $\mathrm{rank}(\mathcal{A}_\varepsilon) \leq 2$، بنابراین $\underline{\mathrm{rank}}(\mathcal{A}) = 2$. به‌این‌ترتیب، دنباله‌ای از تانسورهای رتبهٔ ۲، به یک تانسور رتبهٔ ۳ همگرا شد و این نشان‌دهندهٔ شکافی بین رتبهٔ تانسور و رتبهٔ مرزی آن است.

دلیل اصلی بدحالت‌بودن مسئله بهترین تقریب رتبهٔ پایین، یکسان نبودن رتبه و رتبهٔ مرزی یک تانسور است [۷، ۹].

قضیه ۳.۷. *فرض کنید $\mathrm{rank}(\mathcal{A}) = r$ و $\underline{\mathrm{rank}}(\mathcal{A}) = s$. اگر $r < s$، آنگاه تانسور $\mathcal{A}$ بهترین تقریب رتبهٔ s ندارد.*

[۱] border rank

در روش جواب ضعیف برای محاسبهٔ تقریب رتبهٔ پایین از عناصر $\bar{S}_r$ استفاده می‌شود. فرض کنید $\mathcal{A} \in \mathbb{R}^{d_1 \times \cdots \times d_n}$ که در آن $d_i \geq 2$ و $n \geq 3$. دراین‌صورت راهی برای تضمین اینکه مسئلهٔ یافتن بهترین تقریب رتبهٔ k برای تانسور $\mathcal{A}$ همیشه جوابی سودمند دارد آن است که به جای حل مسئلهٔ تقریب رتبهٔ k به‌صورت

$$\min_{\mathrm{rank}(\mathcal{B}) \leq k} \|\mathcal{A} - \mathcal{B}\|$$

مسئلهٔ تقریب رتبهٔ مرزی k به‌صورتِ زیر را حل کرد

$$\min_{\underline{\mathrm{rank}}(\mathcal{B}) \leq k} \|\mathcal{A} - \mathcal{B}\|. \tag{۱.۷}$$

چون $\overline{S}_2(d_1, d_2, d_3) \subseteq \mathcal{S}_3(d_1, d_2, d_3)$، بنابراین تانسورهای با بعد ۳ و رتبهٔ ۲ به تانسوری با رتبهٔ حداکثر ۳ همگرا می‌شوند. هر تانسور در $\overline{S}_2(d_1, d_2, d_3)$ را می‌توان به یکی از دو صورت

$$y_1 \otimes x_2 \otimes x_3 + x_1 \otimes y_2 \otimes x_3 + x_1 \otimes x_2 \otimes y_3 \tag{۲.۷}$$

یا

$$x_1 \otimes x_2 \otimes x_3 + y_1 \otimes y_2 \otimes y_3 \tag{۳.۷}$$

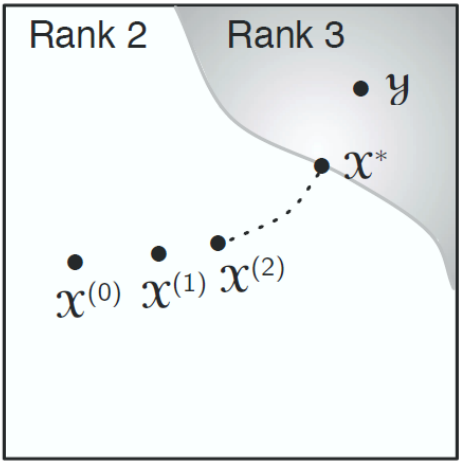
نوشت [۹].

فرض کنید $\mathcal{A} \in R^{d_1 \times d_2 \times d_3}$ به‌طوری‌که $d_i \geq 2$ و $\mathrm{rank}(\mathcal{A}) = 3$. یک جواب ضعیف برای مسئلهٔ بهترین تقریب رتبهٔ پایین از $\mathcal{A}$ را می‌توان با انتخاب $\mathcal{B}$ در مسئلهٔ (۱.۷) از عبارت (۲.۷) یا (۳.۷) به‌طوری‌که $\mathcal{B}$ کمترین فاصله را از $\mathcal{A}$ داشته باشد، به‌دست آورد. اگر $\mathcal{B}$ به‌صورتِ (۳.۷) انتخاب شود، یعنی $\mathrm{rank}(\mathcal{B}) = 2$، آنگاه $\mathcal{A}$ بهترین تقریب رتبهٔ ۲ دارد. اگر $\mathcal{B}$ به‌صورت (۲.۷) انتخاب شود، آنگاه طبق قضیهٔ ۷.۶، تانسور $\mathcal{B}$ حد دنباله‌ای از تانسورهای رتبهٔ ۲ است و بنابراین $\mathcal{A}$ بهترین تقریب رتبهٔ ۲ ندارد، درحالی‌که $\mathcal{B}$ بهترین تقریب رتبهٔ مرزی ۲ برای $\mathcal{A}$ است.

تانسورهایی که رتبهٔ مرزی آن‌ها کمتر از رتبهٔ خود تانسور است تبهگن[۱] نامیده می‌شوند. تانسورهای تبهگن بهترین تقریب رتبهٔ پایین ندارند [۸]. شکل ۴ تانسور $\mathcal{Y}$ را نشان می‌دهد که با دنباله‌ای از تانسورهای رتبهٔ ۲ تقریب زده شده است. دنبالهٔ $\{\mathcal{X}^{(n)}\}$ به‌دلخواه به $\mathcal{Y}$ نزدیک است و بهترین تقریب روی مرز بین تانسورهای رتبهٔ ۲ و رتبهٔ ۳ قرار دارد. به‌هرحال، چون فضای تانسورهای رتبه ۲ بسته نیست، این دنباله ممکن است به تانسور $\mathcal{X}^*$ با رتبهٔ ۳ همگرا شود [۴]. بهترین تقریب الزاماً روی مرز فضای تانسورهای رتبهٔ ۲ و رتبهٔ ۳ قرار دارد.

مجموعهٔ تانسورهایی که بهترین تقریب رتبهٔ k ندارند (دست‌کم برای برخی از kها) اندازهٔ لُبِگ مثبت دارد [۹]. به‌علاوه، تانسورهای $2 \times p \times p$ با رتبهٔ $p+1$ بهترین تقریب رتبهٔ p ندارند [۴].

[۱] degenerate



شکل ۴. تصویر دنباله‌ای از تانسورها که به تانسوری با رتبهٔ بالاتر همگرا می‌شود [۴].

تانسورهای $۲ \times ۲ \times ۲$ با رتبهٔ ۳ که ابردترمینان برابر با صفر دارند، در مرز بین مجموعهٔ تانسورهای رتبهٔ ۳ با $\triangle < ۰$ و تانسورهای رتبهٔ ۲ با $\triangle > ۰$ قرار دارند. بنابراین، این تانسورها می‌توانند به‌دلخواه به تانسورهای رتبهٔ ۲ نزدیک شوند و این بدین معنی است که تابع

$$f(A,B,C) = \left\|\mathcal{T} - \sum_{r=۱}^{۲} \lambda_r\, a_r \otimes b_r \otimes c_r\right\| \qquad (۴.۷)$$

مینیمم ندارد. برای یافتن یک تقریب رتبهٔ ۲ در این تانسورها، تجزیهٔ CP با استفاده از الگوریتم کمترین مربعات متناوب به شکست می‌انجامد.

۱.۷. دسته‌بندی تانسورهای $۲ \times ۲ \times ۲$. در این بخش، هم‌ارزی تانسورها در $\mathbb{R}^{۲\times۲\times۲}$ را تحت عمل ضربِ چندخطی ماتریس‌ها بیان می‌کنیم و یک دسته‌بندی از تانسورهای $۲\times۲\times۲$ ارائه می‌دهیم. این هم‌ارزی از طریق ضرب چندخطی ماتریس‌های ناتکین به‌دست می‌آید. چنین تانسورهایی را صورت‌های کانونی[۱] می‌نامند.

در این دسته‌بندی، رتبهٔ مرزی، رتبه، رتبهٔ چندخطی، و علامت ابردترمینان تانسورهای کانونی مشخص است. برای تانسورهای $۲ \times ۲ \times ۲$، هشت صورت کانونی مجزا وجود دارد و بقیهٔ تانسورهای متعلق به $\mathbb{R}^{۲\times۲\times۲}$ با یکی از این صورت‌های کانونی هم‌ارزند و مشخصات یکسانی دارند. قضیهٔ زیر به‌روشنی این مطلب را بیان می‌کند.

[۱] canonical forms

جدول ۱. صورت‌های کانونی تانسورهای $۲ \times ۲ \times ۲$

تانسور	علامت $\triangle$	$\text{rank}_{\boxplus}$	rank	$\overline{\text{rank}}$
$D_{۰} = \left(\begin{array}{cc\|cc} ۰ & ۰ & ۰ & ۰ \\ ۰ & ۰ & ۰ & ۰ \end{array}\right)$	۰	$(۰, ۰, ۰)$	۰	۰
$D_{۱} = \left(\begin{array}{cc\|cc} ۱ & ۰ & ۰ & ۰ \\ ۰ & ۰ & ۰ & ۰ \end{array}\right)$	۰	$(۱, ۱, ۱)$	۱	۱
$D_{۲} = \left(\begin{array}{cc\|cc} ۱ & ۰ & ۰ & ۰ \\ ۰ & ۱ & ۰ & ۰ \end{array}\right)$	۰	$(۱, ۲, ۲)$	۲	۲
$D'_{۲} = \left(\begin{array}{cc\|cc} ۱ & ۰ & ۰ & ۱ \\ ۰ & ۰ & ۰ & ۰ \end{array}\right)$	۰	$(۲, ۱, ۲)$	۲	۲
$D''_{۲} = \left(\begin{array}{cc\|cc} ۱ & ۰ & ۰ & ۰ \\ ۰ & ۰ & ۱ & ۰ \end{array}\right)$	۰	$(۲, ۲, ۱)$	۲	۲
$G_{۲} = \left(\begin{array}{cc\|cc} ۱ & ۰ & ۰ & ۰ \\ ۰ & ۰ & ۰ & ۱ \end{array}\right)$	+	$(۲, ۲, ۲)$	۲	۲
$D_{۳} = \left(\begin{array}{cc\|cc} ۱ & ۰ & ۰ & ۱ \\ ۰ & ۰ & ۱ & ۰ \end{array}\right)$	۰	$(۲, ۲, ۲)$	۳	۲
$G_{۳} = \left(\begin{array}{cc\|cc} ۱ & ۰ & ۰ & -۱ \\ ۰ & ۱ & ۱ & ۰ \end{array}\right)$	−	$(۲, ۲, ۲)$	۳	۳

قضیه ۴.۷ ([۹]). *هر تانسور در $\mathbb{R}^{۲\times۲\times۲}$ از طریق تبدیل‌های ناتکین دقیقاً با یک صورت کانونی در جدول ۱ هم‌ارز است.*

۸. جمع‌بندی

روش‌های چندخطی در فیزیک، روانشناسی، بینایی ماشین، پردازش صدا و تصویر، و بسیاری از زمینه‌های دیگر کاربرد دارند. با توجه به رشد داده‌ها، تشخیص روابط پنهان بین داده‌های ذخیره شده در تانسور و صرفه‌جویی در زمان پردازش و حافظهٔ موردنیاز از مهم‌ترین مسائل به‌شمار می‌روند. یکی از مهم‌ترین مشخصه‌های یک تانسور رتبه است که از آن در مسئلهٔ محاسبهٔ بهترین تقریب رتبهٔ پایین استفاده می‌شود. در این مقاله، ما به بررسی چگونگی محاسبهٔ رتبهٔ تانسورها پرداختیم. هرچند مسئلهٔ محاسبهٔ رتبهٔ تانسور مسئله‌ای NP–سخت است، روش‌های مناسبی برای محاسبهٔ رتبهٔ برخی از تانسورها وجود دارند. ارتباط بین درایه‌ها و تکینی یا ناتکینی هریک از برش‌ها نقش‌هایی اساسی در تعیین رتبهٔ یک تانسور دارند. محاسبهٔ تقریب رتبهٔ پایین برای ماتریس‌ها ازجملهٔ مسائل اساسی است که در بسیاری از زمینه‌ها ازجمله

پردازش صدا و تصویر و سیستم‌های توصیه‌گر[۱] کاربرد دارد. مسئلهٔ تقریب رتبهٔ پایین، به‌منظور صرفه‌جویی در زمان و حافظهٔ مورد نیاز و بررسی روابط پنهان داده‌ها، تقریبی مناسب از داده‌های حجیم ارائه می‌کند که موضوعی مورد توجه و البته چالش‌برانگیز است.

مراجع

[1] Cichocki, A., Zdunek, R. L. , Phan, A. H., Amari, S., *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, John Wiley and Sons, London, 2009.

[2] Golub, G. H., Van Loan, C. F., *Matrix Computations*, Johns Hopkins University Press, Baltimore, 2013.

[3] Hitchock, F. l., The expression of tensor or a polyadic as a sum of products, *J. Math. Phys.*, **6** (1927), 164-189.

[4] Kolda, T. G., Bader, B. W., Tensor decompositions and applications, *SIAM Rev.*, **51** (2009), 455-500.

[5] Kruskal, J. B., Rank, decomposition and uniqueness for 3-way and N-way arrays, in *Multiway Data Analysis*, North-Holland Publishing Co., Amsterdam, 1989, 7-18.

[6] Landsberg, J. M. , The border rank of the multiplication of 2×2 matrices is seven, *J. Amer. Math. Soc.*, **19** (2006), 447-459.

[7] Lim, L. H., Tensors and Hyprematrices, in *Handbook of Linear Algebra*, 2nd ed., L. Hogben, ed., CRC Press, Boca Raton, FL, 2013.

[8] Rovi, A., Analysis of $2 \times 2 \times 2$ Tensors, Master's thesis, Linköping University, 2010.

[9] Silva, V. D., Lim, L. H., Tensor rank and the ill-posedness of the best low-rank approximation problem, *SIAM J. Matrix Anal.*, **30** (2008), 1084-1127.

[10] Ten Berge, J. M. F., Kruskal's polynomial for $2\times2\times2$ arrays and a generalization to $2\times n\times n$ arrays, *Psychometrika*, **56** (1991), 631-636.

تاریخ ارسال: ۱۳۹۶/۹/۲۸؛ تاریخ بازنگری: ۱۳۹۷/۲/۱؛ تاریخ پذیرش: ۱۳۹۷/۲/۲

زهرا اردولالو: دانشگاه قم، دانشکدهٔ علوم پایه

رایانامه: ordolalo@chmail.ir

عفت گلپرابوکی: دانشگاه قم، دانشکدهٔ علوم پایه

رایانامه: g.raboky@qom.ac.ir

نظام‌الدین مهدوی امیری: دانشگاه صنعتی شریف، دانشکدهٔ علوم ریاضی

رایانامه: nezamm@sharif.edu

---

[۱] recommender system

Mathematical Culture and Thought
Volume 40, Number 69, Fall/Winter 2021, Pages 61–78
doi: 10.30504/mct.2022.323

Tensor Rank and the Best Low-rank Approximation Problem

Z. Ordolalo[1], E. Golpar Raboky[2]✉, N. Mahdavi-Amiri[3]

[1,2]Department of Mathematics, University of Qom, Iran
[3]Department of Mathematical Sciences, Sharif University of Technology, Iran

Abstract. The rank of a matrix is indeed an important characteristic, and it can be calculated using methods such as Gaussian elimination or stepwise decomposition. The rank of a tensor A, on the other hand, is more complex to calculate and is considered an NP-hard problem except in special cases. Various methods have been proposed for computing the rank of 2x2x2 tensors, including the calculation of hyperdeterminant, checking the internal structure of the tensor, and classifying the tensor into canonical forms. The Eckart-Young theorem provides a method for calculating the best low-rank approximation for matrices. However, many tensors may not have a specified low-rank approximation, which is a degenerate problem. These tensors can be approximated by a sequence of low-rank tensors.

Keywords: numerical multilinear algebra, tensor, tensor rank, tensor decomposition, low-rank approximation
Article history: Recieved 19 December 2017; Accepted 22 April 2018
Article type: review

References

[1] Cichocki, A., Zdunek, R. L. , Phan, A. H., Amari, S., *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, John Wiley and Sons, London, 2009.

[1]ordolalo@chmail.ir
[2]g.raboky@qom.ac.ir
[3]nezamm@sharif.edu

[2] Golub, G. H., Van Loan, C. F., *Matrix Computations*, Johns Hopkins University Press, Baltimore, 2013.

[3] Hitchock, F. l., The expression of tensor or a polyadic as a sum of products, *J. Math. Phys.*, **6** (1927), 164-189.

[4] Kolda, T. G., Bader, B. W., Tensor decompositions and applications, *SIAM Rev.*, 51 (2009), 455-500.

[5] Kruskal, J. B., Rank, decomposition and uniqueness for 3-way and N-way arrays, in *Multiway Data Analysis*, North-Holland Publishing Co., Amsterdam, 1989, 7-18.

[6] Landsberg, J. M. , The border rank of the multiplication of 2×2 matrices is seven, *J. Amer. Math. Soc.*, **19** (2006), 447-459.

[7] Lim, L. H., Tensors and Hyprematrices, in *Handbook of Linear Algebra*, 2nd ed., L. Hogben, ed., CRC Press, Boca Raton, FL, 2013.

[8] Rovi, A., Analysis of $2 \times 2 \times 2$ Tensors, Master's thesis, Linköping University, 2010.

[9] Silva, V. D., Lim, L. H., Tensor rank and the ill-posedness of the best low-rank approximation problem, *SIAM J. Matrix Anal.*, **30** (2008), 1084-1127.

[10] Ten Berge, J. M. F., Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays, *Psychometrika*, **56** (1991), 631-636.