

テンソル分解のマーケティングへの適用可能性

朝野熙彦

概要

Osgood ら(1957,1962) は「個人」「評価尺度」「概念」という3相構造のデータを分析する semantic differential scales、略称 SD 法を提唱した。これが社会科学におけるテンソルの最初の提唱だった。

今日では3相データはより一般的な N 階テンソルの枠組みで分析できるようになっている。 N 階テンソルは信号処理、医療画像解析など技術系の情報処理で活躍している。本ノートではテンソル分解がSD法から派生してきたという歴史をたどる。そのうえでマーケティング分野へのテンソルデータ解析の適用可能性を検討する。

キーワード:SD 法、3相因子分析、テンソル分解、HOSVD

1. テンソルの概念

具体例で説明した方が分かりやすいので、ここでは SD 法のデータ構造をテンソルとして記述しよう。

サンプル $i = 1, 2, \dots, I$ 、変数 $j = 1, 2, \dots, J$ 、概念 $k = 1, 2, \dots, K$

i, j, k の3つの添え字をインデックスにしたデータを X_{ijk} と書き、データを一括して書記体でテンソル $\mathcal{X}$ と書く。インデックスはデータの「所番地」と理解すればよい。データサイエンスおよび機械学習ではこのインデックスの数を階数と呼んでいる。ベクトルは1階のテンソル、行列は2階のテンソル、 N 個の添え字が必要なら N 階のテンソルと呼ぶ。

またランクは線形空間を張る一次独立なベクトルの最大数だが、テンソルでは次元を打ち切り数という意味で使うことがある。

X_{ijk} は3相から成るが、Tucker (1966) は各相をモード(mode)と呼んだ。3つのモードにそってテンソルを展開(unfolding)した行列を X_1, X_2, X_3 と書く。

X_1 : サンプル i を行にあてて列は変数と概念を組み合わせた $J \times K$ 列

X_2 : 変数 j を行にあてて列はサンプルと概念を組み合わせた $I \times K$ 列

X_3 : 概念 k を行にあてて列はサンプルと変数を組み合わせた $I \times J$ 列

それぞれモード 1, モード 2, モード 3 にそった展開と呼ぶ。図 1 に示すように展開行列はテンソル $\mathcal{X}$ の全要素を行列に並べなおしたものである。展開行列の狙いは行に選んだ相にフォーカスして残りの 2 相の情報を反復(repetition)として扱うことにある。

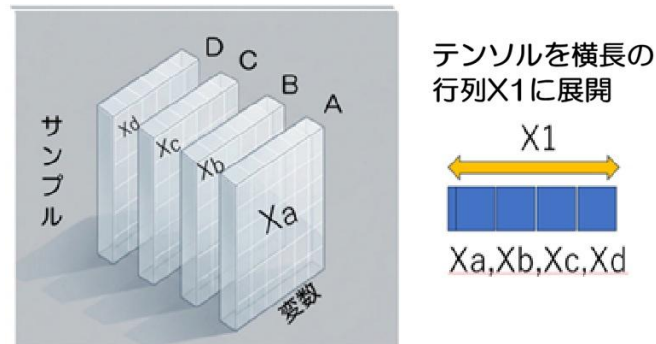


図 1 テンソルのモード 1 展開

展開行列は実行列なので①のように特異値分解(singular value decomposition)ができる。特異値分解を SVD と略する。 Λ は特異値を主対角要素とした対角行列であり、正の特異値が大きい順に並んでいるとする。 U と V は正規直交行列である。

$$X_1 = U_1 \Lambda_1 V_1^T, X_2 = U_2 \Lambda_2 V_2^T, X_3 = U_3 \Lambda_3 V_3^T \quad \dots\dots\dots ①$$

3つの展開行列の打ち切りランクを R_1, R_2, R_3 とする。モードごとにランクは変わりうるのでモードランクと呼ぶ。そして左特異行列 $U_n (n = 1, 2, 3)$ からモードランクだけ特異ベクトルを抽出し列に関して正規直交化したのが行列 A, B, C である¹。 A, B, C は原データの factorization (分解)という意味で因子行列(factor matrix)と呼ぶ。因子と言っても因子分析という因子とは異なる。

A: サンプルの因子行列 I 行 R_1 列 特異値は大きい順に R_1 個

B: 変数の因子行列 J 行 R_2 列 特異値は大きい順に R_2 個

C: 概念の因子行列 K 行 R_3 列 特異値は大きい順に R_3 個

このように各モードについて対等の処理を行い A, B, C を導く。なお実数の矩形行列が最小二乗の意味で低ランク近似できることは Eckart と Young (1936)が証明した。今から約 90 年前の彼らの定理が今日のテンソル分解の数学的基礎になっている。次に 3階テンソルを解析する3通りのアプローチを説明しよう。

2. HOSVD

まず Lathauwer ら(2000)の高階特異値分解 HOSD(higher order SVD)を紹介する。テンソル $\mathcal{X}$ を縮約²して得られるコアテンソルが②の $\mathcal{G}$ である。この $\mathcal{G}$ はテンソル $\mathcal{X}$ のミニチュア模型だと解釈すればよい。

$$\mathcal{G} = \mathcal{X} \times_1 A \times_2 B \times_3 C \quad \dots\dots\dots ②$$

②の意味を説明すると、右辺最初の $\mathcal{X} \times_1 A$ の部分がモード 1 積 $\times_1$ から得られるテ

ンソルである。さらに n モード積 $\times_n$ を繰り返すことでテンソルを縮約している。なおここでは第1、第2、第3の順にモード積をとったが、この順番を入れ換えても同じコアテンソルに到達する。

ここで②のアルゴリズムを整理しよう。モード n 積をとる関数を mnf として、その引数を(対象テンソル, 因子行列, モード)とすれば、次の再帰的な適用によってコアテンソル $\mathcal{G}$ が導かれる。

$$\begin{aligned} Y1 &= \text{mnf}(X, A, 1) \\ Y2 &= \text{mnf}(Y1, B, 2) \\ Y3 &= \text{mnf}(Y2, C, 3) \\ G &= Y3 \end{aligned}$$

テンソル $Y1, Y2, Y3$ の次数は順に $(R1, J, K) \Rightarrow (R1, R2, K) \Rightarrow (R1, R2, R3)$ へと縮約される。線形代数に慣れた人なら次の行列計算の方が理解しやすいだろう。

$$G_1 = A^T X_1 (C \otimes B), \quad G_2 = B^T X_2 (C \otimes A), \quad G_3 = C^T X_3 (B \otimes A) \quad \cdots \cdots ③$$

③の最初の式は、モード 1 にそったテンソルの展開行列 X_1 から出発して行列 G_1 を導くまでの計算過程を示している。記号 $\otimes$ はクロネッカー積である。図 2 にモード 1 にそった展開と畳み込みを図解した。このように G_1 を畳み込むだけでコアテンソルが導かれるので、③に書いた G_2 と G_3 の計算式は冗長である。

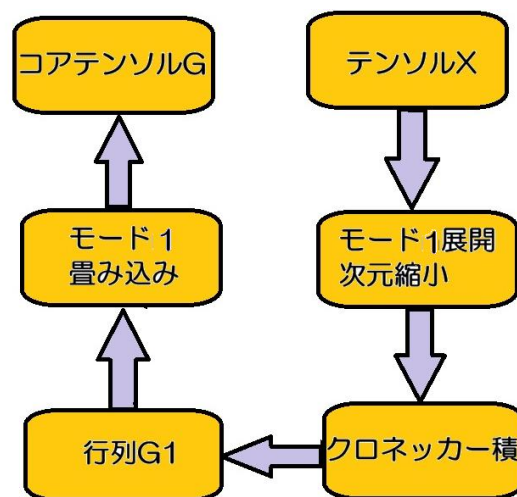


図 2 モード 1 にそった展開と畳み込み

最後に $\mathcal{X}$ と $\mathcal{G}$ の関係を述べよう。Tucker(1966,p293)は入力テンソルとコアテンソル

ルの要素の二乗和は等しいと述べている。④は Tucker の主張をフロベニウスノルム (Frobenius norm) の二乗で書き直したものである³。

$$\|X\|_F^2 = \|G\|_F^2 \dots\dots\dots \textcircled{4}$$

④が成り立つのは、入力テンソルが exact に分解できた場合に限る。実際にはテンソルを低ランクで近似するので④は近似的な関係 $\approx$ にとどまる。

■ HOSVD の特徴

- 1) HOSVD は 3 階から N 階に拡張できる
- 2) 因子行列は列に関して直交化されているので解釈しやすい
- 3) コアテンソルはモード間の交互作用を意味するが解釈しにくい
- 4) 因子行列とコアテンソルから元のテンソル $\mathcal{X}$ を近似的に復元できる
- 5) テンソル分解の結果はデータ生成器として利用できる

3) の短所と 4) の長所は二律背反といえる。コアテンソルは各モードの直交基底に射影したテンソルであり、モード間の交互作用を意味する。しかし数値はテンソルのスケールに依存するので大きさの評価が難しい。

4) は展開の逆操作のことである。テンソルではそれを **畳み込み (folding)** と呼んでいい。具体的な計算式は⑤式の通りである。

$$X_{ijk} = \sum_p \sum_q \sum_r a_{ip} b_{jq} c_{kr} g_{pqr} \dots\dots\dots \textcircled{5}$$

⑤の総和記号の添え字は順に $p=1,2,\dots,R_1$, $q=1,2,\dots,R_2$, $r=1,2,\dots,R_3$ と変化する。なぜ復元できるのかをここで説明しておこう。まず③の G_1 の右から $(C \otimes B)^T$ を掛けて⁴

$$\begin{aligned} G_1(C \otimes B)^T &= A^T X_1 (C \otimes B)(C \otimes B)^T \\ &= A^T X_1 (CC^T \otimes BB^T) = A^T X_1 (I_K \otimes I_J) = A^T X_1 I_{JK} = A^T X_1 \end{aligned}$$

次に左から A を掛けると A の張る空間への射影子 AA^T が得られて

$$\begin{aligned} AG_1(C \otimes B)^T &= AA^T X_1 = X_1 \\ X_1 &= AG_1(C \otimes B)^T \dots\dots\dots \textcircled{6} \end{aligned}$$

したがって⑥で展開行列 X_1 を求めて、それを畳み込めば元のテンソル $\mathcal{X}$ が復元できる。

テンソル全体を記録するよりもテンソル分解した結果を記録する方が情報圧縮になる。 $I=J=K=100$ で $R_1=R_2=R_3=10$ という小さなケースで単純に記録するデータ数を比べるとテンソル分解による圧縮比 (compression ratio) は $\frac{10 \times 10 \times 10 + 3(100 \times 10)}{100^3} = 0.004$ なので、テンソルの 0.4% で済む。大規模データほどストレージの効率が上がる。

⑤のテンソルの復元は簡単なので、シンセティックデータの生成器としてテンソル分

解が利用できる。その際は適当な乱数項を⑤に追加すればよい。

さてテンソルの復元精度の評価法であるが、 $\mathcal{X}$ が入力したテンソルで、 $\hat{\mathcal{X}}$ が復元されたテンソルだとすれば再構成の誤差は $\|\mathcal{X} - \hat{\mathcal{X}}\|_F$ であり、標準化した精度は⑦の通り。

$$Accuracy = 1 - \frac{\|\mathcal{X} - \hat{\mathcal{X}}\|_F}{\|\mathcal{X}\|_F} \dots\dots\dots ⑦$$

R1,R2,R3 を大きく設定すれば精度を高められる。しかし低ランクで近似することがテンソル分解の目的なので、100%の Accuracy を目標にするわけではない。

さて HOSVD で SD 法の問題はすべて解決できたと思うかもしれないが、そうではない。HOSVD によって $\|\mathcal{X} - \hat{\mathcal{X}}\|_F^2$ が最小化できるという保証がないからである。Eckart と Young が主張したのは行列の場合には誤差を最小化できるということであり、テンソルを対象にした主張ではない。HOSVD の解を改善する方法については後でふれる。

3. タッカーの3相因子分析

Tucker(1964,1966)は古典的な SD 法が個人差を扱ってこなかった欠点に着目して3相因子分析(Tucker3)を提唱した。彼の提案は今日のテンソル分解とほぼ同じだった。

ただし Tucker(1964)の three-dimensional matrices という用語は適切ではなく多重配列あるいはテンソルと呼ぶべきだった。行列なら2階だからである。また推定のためのランクの決定法が未解決だった。さらに Tucker は固有値分解を用いたが、固有値分解は SVD よりも迂遠な方法だった。その後 Kroonenberg ら(1980)が、因子行列 A,B,C のうち2つに初期値を与え、残る 1 つを推定し、順次推定パラメータを交代していくという ALS(交互最小二乗法)を用いた Tuckals3 を提案した。ただし Tuckals3 には初期値依存性があるので、その初期値をどう定めるかという問題が残った。このように多少の問題はあったにせよ、Tucker(1966)の提案はほぼ正しかった。

近年 Andrzej (2016)はタッカー分解を含めて複数のテンソル分解を組み合わせた柔軟なモデリングを提案している。横田 (2024)は HOOI (Higher-Order Orthogonal Iteration) といって ALS の初期値に HOSVD の結果を用いて解を改善する方法を示している。また Todorov ら(2026)は Tucker3 のパラメータを推定する rrcov3way という R のパッケージを公開している。このようにテンソルデータの解析は近年研究が進んでいる。

さてSD法のデータにHOSVDを適用するには何通りかの手順が考えられる。たとえば次のような割り当てが可能だろう。

モード1:ブランドI

モード2:イメージ変数J

モード3:条件の違いK

条件の違いとして、マーケティングでは広告キャンペーンの前か後かという時点の違いを使うことがある。

4. 因子分析

Osgoodら(1957)はSD法のデータを因子分析してバリマックス回転した。Greenら(1967)によれば、アメリカでは1964年からマーケティング導入が始まったという。また志津野(1984)は1960年代の日本はアメリカから新しい調査技法が次々と輸入された華やかな時代だったという。日本ではSD法をイメージ調査に借用したことを率直に述べている(同書47-48頁)。

この時代にOsgoodにならって両極尺度のデータを因子分析する一連の手順が慣用化された。Worcester(1972)のハンドブックその他を介してSD法が制度化された時代だった。SD法の運用にあたっては、すでに述べたモード2の展開を行い、次にイメージ変数の相関行列を求めることが多かった⁵。

行:J個のイメージ変数

列:サンプルIとブランドKの組み合わせ

さらに簡単に処理するために、あらかじめサンプルに関して平均をとってJ行K列の行列を作ってから分析する慣用もあった。

いずれにしても変数間の相関行列に情報を集約するのだから個人差の情報は行列分解に反映されない。そのため個人によるイメージ空間の違いを知ることがマーケティング課題の場合は、あらかじめ消費者をグループに分けてデータ行列を作り、グループごとに因子分析することもあった。このアプローチは因子空間の比較というさらなる問題を引き起こす。

■ テンソル分解と因子分析の相違

1) 因子分析は回転が任意なので、分析者は自分の仮説を支持する回転結果を選ぶことができる。概念の普遍性という仮説検証が目的だったにもかかわらず探索的な因子分析を用いたオズグッドの研究法は適切でなかった。それに対してHOSVDの解には回転の自由がないのでその点では任意性が低い⁶。

2) HOSVDは非対称テンソルの直交基底を求める方法なので、各モード間の次元は異なる。仮に次元数を同数に指定したとしてもA,B,Cの次元はそれぞれ別空間に属する。そのためHOSVDによるA,B,Cの推定値を同一空間にジョイント・プロットする表現には数理的な根拠がない。それに対して、因子分析から得られる因子得点と因子負荷量は同次元に属する。なぜなら因子分析は一つの実対称行列の固有値問題に帰着するので、固有ベクトルを一組抽出するだけだからである。従って解釈性という面では因子分析に有利さがある。

3) SD 法の3相データ構造はテンソルそのものだった。HOSVD であれば分析結果から元のテンソルを Accuracy の範囲で復元できる。それに対して 2 相データ用の因子分析から3相の原データを復元することはできない。データの復元性という意味では HOSVD が優れている。

5. マーケティングにおけるテンソル分解の応用

Zhang ら(2021)は POS データにテンソル分解を適用した dynamic tensor recommender system を提案している。Zhang らは IRI のマーケティングデータを CP 分解(canonical polyadic decomposition)で分析し売り上げ予測にテンソル回帰を使用した。なお CP 分解にはコアテンソルが無く各相のランクは等しい。したがって CP 分解は HOSVD の特殊なケースとみなせる。

さて消費行動はしばしば次のような多相的な構造を持っている。

誰が(ユーザー／セグメント)

何を(商品／コンテンツ)

どこで(チャネル／媒体)

いつ(時間帯／キャンペーン期間)

どういう状況で(文脈／オケーション)

変数が多いなら多変量解析を使えば済むと思うかもしれないがそうとも言いきれない。たとえば評価変数が多いので次元を縮約する事態を考えよう。購入時期や購入場所というモードの異なる変数を評価変数と一緒に次元縮約して何の意味があるだろうか。テンソル分解ならモードの違いを区別しつつデータ解析できる。

6. 討論

本稿では今日のテンソル分解の一つの源流がオズグッドの SD 法にあったことを指摘し、SD 法⇒3 相因子分析⇒HOSVD という研究の連鎖を述べた。もう一つの源流である多重配列データの特異値分解に関しては日本の吉澤(1976)が先駆的な研究を行っている。それにもかかわらず信号処理や機械学習のコミュニティで吉澤の研究が引用されることはほぼない。一部の学会に偏ることなくテンソルデータ解析を広く展望したレビューが望まれる。

最後にマーケティングでテンソルデータ解析を応用するために解決すべき課題と疑問点を指摘しよう。マーケティング固有の目的と制約の中で、テンソルをどう設計すればよいかは共通した課題である。

1) 3階のテンソルはごく単純化していえば記録単位 i について多変量データ j と何らかの状況情報 k が付随したデータである。その記録単位は個人に限らず都道府県や企業、自社の店舗、施設などを記録単位にとってよいはずである。そして記録の年度を k とすればテンソルデータが構成できる。

2) 本稿では基準変数を持たない HOSVD を紹介した。一方で、売上や購買行動を基準変数としてテンソルを説明変数に用いるテンソル回帰 (scalar on tensor regression) も重要な方法である。テンソル回帰は Zhou ら(2013) などいくつも提案がなされている。もともと医療画像解析から発展した手法だが、POS データや広告効果分析などマーケティング領域へも応用されはじめている。つまりテンソルデータ解析は次元縮約から予測と意思決定へと広がりつつある。

3) テンソル分解できる尺度水準は何か。購入金額のような連続量、そして購入の有無のような2値データはそのまま分析できる。しかし名義尺度と順序尺度のデータには数量としての意味がないので何らかの前処理が必要になる。順序埋め込み(ordinal embedding)による変換、そして名義尺度の one-hot ベクトル表現などが必要になる。

4) テンソルの設計にあたっては相内は同質で、相の間は異質な情報を配置するのが通例である。しかし相と相の因果関係はコアテンソルからは分からない。多相間の関係を表現する既存の方法として SEM がある。相を潜在変数とみて SEM で表現できる可能性がある。SEM とテンソル分解の統合ができれば解釈上も有用だろう。

5) 社会科学は実験科学と異なり観測数にアンバランスが生じる可能性がある。「ユーザー、ブランド、オケーション」の組み合わせに偏りが生じるのは当然である。それに加えてスパースなデータへの対応も必要になる。

6) テンソルのモードごとに特異値が求められるが、テンソルにおける各モードの重要度をどう評価すればよいか。要因の寄与率に相当する指標化ができるだろうか。

7) マーケティング分野には従来からID付POSデータやシングルソースデータが存在していた。購買や取引を単位にした多相的なデータが得られているのでテンソルとして設計するのは自然だろう。すでに各企業ではテンソルデータ解析を実験し実務に導入しているのだろうか。

(2026年 10 月 20 日)

引用文献

- Andrzej C. (2016) Tensor networks for dimensionality reduction and large-scale optimization, Part 1, Low-rank tensor decompositions, *Foundations and Trends in Machine Learning*, **9**, No.4-5, 249-429.
- Eckart, C. and Young, G. (1936) The approximation of one matrix by another of lower rank. *Psychometrika*, **1**, No.3, 211-218.
- Green, P.E. and Frank, R.E. (1967) “*A Manager’s Guide to Marketing Research: Survey of Recent Developments.*” John Wiley & Sons, inc.
- Kroonenberg, P.M. and de Leeuw, J. (1980) Principal component analysis of three mode data by means of alternating least squares algorithm. *Psychometrika*, **46**, No.1, 69-97.

- Lathauwer, L.De, Moor, B. De and Vandewalle, J. (2000) A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, **21**, No.4, 1253-1278.
- Osgood, C.E., Susi, G.J. and Tannenbaum, P.H. (1957) “*The Measurement of Meaning*”, Urbana: Univ. Illinois Press.
- Osgood, C.E. (1962) Studies on the generality of affective meaning systems, *Amer. Psychologist*, **17**, 10-28.
- 志津野知文 (1984) マーケティング・リサーチの概念. 金子泰雄 (編著) 「現代マーケティングリサーチ」創成社, 37-74.
- Todorov, V., Simonacci, V., Palma, A. and Gallo, M. (2026) Robust tools for three-way component analysis of compositional data. *Behaviormetrika*, **53**, 265-296.
- Tucker, L.R. (1964) The extension of factor analysis to three-dimensional matrices, In Gulliksen, H. and Frederiksen, N. (eds.) “*Contributions to Mathematical Psychology*”, New York: Holt, Rinehart & Winston, 110-127.
- Tucker, L.R. (1966) Some mathematical notes on three-mode factor analysis. *Psychometrika*, **31**, No.3, 279-311.
- Worcester, R.M, Eds. (1972) “*Consumer Market Research Handbook*”, McGraw-Hill (UK).
- 横田達也 (2024) 「テンソルデータ解析の基礎と応用」次世代信号処理シリーズ 7, コロナ社
- 吉澤正 (1976) 交互作用概念の一般化と多重配列の特異値分解. 行動計量学, **4**(1), 32-43.
- Zhang, Y., Xuan Bi, Tang, N. and Qu, A. (2021) Dynamic tensor recommender systems. *Journal of Machine Learning Research*, **22**, 1-35.
- Zhou, L., Li, Y., and Zhu, H. (2013) Tensor regression with applications in neuroimaging. *JASA*, 108(502), 540-552.

¹ 行列 A, B, C を因子負荷量と呼ぶのは因子分析と混同されるおそれがある。因子分析は共通性の推定と因子の回転を核心とするので、SVD とは別物である。同じく Worcester (1972, p92) は、PCA を因子分析と呼ぶ口語用法がマーケティング・リサーチ関係者にあったと指摘している。

² ここでの A'_1 は I 次元のベクトルを R1 次元のベクトルに写像する線形変換を意味する。R1 次元の空間への射影子は $A(A^T A)^{-1} A^T = AA^T$ である。

³ Frobenius norm はテンソルの各要素の二乗和の平方根で定義される。

⁴ ここでは $(A \otimes B)(C \otimes D) = AC \otimes BD$ 、 $(A \otimes B)^T = A^T \otimes B^T$ 、列直交行列に関しては $A^T A = I$ を用いて展開すればよい。

⁵ 因子分析は相関行列の主対角要素を共通性の推定値に置き換える。単なる相関行列の固有値分解ではない。

⁶ 同じ分析を繰り返すと次元ごとに正負が変わるので符号に関する不定性がある。