

単語共起テンソルに内在する代数的不変構造 – 言語の対称性と PMI の数理構造の非自明な対応 –

前田晃弘^{1,3} 鳥居拓馬² 日高昇平¹

¹ 北陸先端科学技術大学院大学 ² 東京電機大学 ³ 日本学術振興会特別研究員
{akihiro.maeda,shhidaka}@jaist.ac.jp tak.torii@mail.dendai.ac.jp

概要

単語分散表現に線型的な性質が現れる機序を明らかにするため、本研究は、単語共起確率の相互情報量を成分とする PMI 行列の局所的なランク 1 構造に着目し、それが高階の確率テンソルに内在する代数的不変条件の射影であることを実証する。代数統計を用いた言語モデルを導入し、言語の組み合わせ構造に備わる対称性が、確率テンソル内に小行列式の形をした代数的制約をもたらすことを示す。PMI を単なる二語間の統計量ではなく高次の言語構造を反映する射影として捉えられる。

1 はじめに

1.1 線型構造仮説と単語共起分布の構造

単語分散表現に意味的關係が線型演算として現れる現象に対して、表現空間中の線型構造を仮定する一連の研究 (表現構造仮説 [1, 2, 3]) は、生成モデルの提案など、その数理的機序の解明を試みている。近年でも、大規模言語モデルの内部状態に線型的な性質が備わるとする線型表現仮説 (Linear Representation Hypothesis) が研究されている [4, 5]。

単語分散表現の解釈性研究の先駆けである [6] は、Word2Vec[7] が単語共起頻度の PMI (Pointwise Mutual Information) 行列の分解であることを示した。PMI は多様な意味タスクで高い有効性を示すことが知られており [8, 9]、分散表現の持つ線形性が、表現の学習モデルではなく、入力統計である単語共起分布の構造的性質に由来することを示唆する。

1.2 意味構造と数理構造の非自明な対応

著者らは、単語共起行列の数理構造の解明を試みた研究において [10]、PMI 行列より、近似的にランク 1 となる部分行列を抽出すると、意味的に解釈

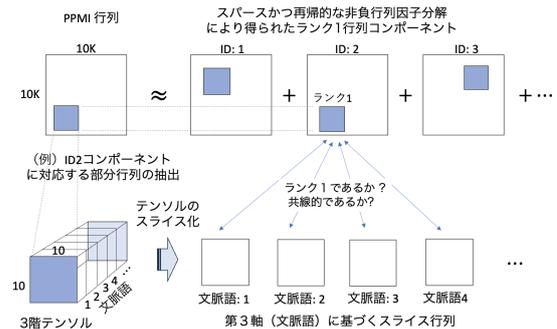


図 1 確率テンソルの局所的ランク 1 構造が、共起行列への射影において保存されていることを実証する。

可能な単語群 (概念) に対応することを明らかにした。すなわち、一つの単語の意味がその共起分布で表現される (分布仮説) だけではなく、単語と単語の意味的關係と、その部分行列の数理構造に対応が認められたのである。言語におけるこうした構造的対応は非自明であり、特に、ランク 1 となる部分行列の存在が、PMI 行列の分解として得られる単語分散表現に低ランク構造すなわち線型的な表現空間の構造をもたらしている可能性が考えられる。

1.3 本研究：不変制約の周辺化と PMI

本研究は、PMI 行列の局所的ランク 1 構造が文生成系における代数的制約の射影であることを実証する。先行研究 [11] での提案言語モデルに基づき、ランク 1 構造が、高階テンソル内の不変構造 (消失二項式) が周辺化を経て、保存されたものであるとの仮説を立て、3 階テンソルの分析によりこれを検証する。具体的には、共起行列から抽出したランク 1 構造をシードに 3 階部分テンソルを構築し、理論が予測する代数的制約が保持されているかを確認することで、その継承メカニズムを実証する (図 1)。本研究の核心は、PMI を単なる二語統計ではなく、高次の組合せ的制約が可視化された射影像と捉え直す点にある。この視点の転換により、PMI の局所構造

を、言語の組み合わせ構造を持つ対称性の代数的帰結として理解することが可能となる。

2 代数統計による言語モデル

2.1 文の確率テンソル

本研究では、言語の組み合わせ的な構造と単語の共起分布の対応に焦点を当てるため、文の確率を文を構成する単語の同時確率として定式化する。より正確には、文を構成する項構造が与える統語的な役割（例えば、依存文法における *subj*, *obj* など）を確率変数とみて、文中の特定の役割に当てられた単語をその実現値と考える。単純化のために、主語/述語/目的語の3語からなる文（例：*Ann likes fish*）の確率として式 (1) を考える

$$Pr(\text{sentence}) = Pr(S, V, O). \quad (1)$$

式中の S, V, O は確率変数で、それぞれの実現値の集合（状態空間） $\mathcal{S} = \{\text{ann}, \text{bob}\}$, $\mathcal{V} = \{\text{eats}, \text{likes}\}$, $\mathcal{O} = \{\text{fish}, \text{beef}\}$ とする。*Ann eats fish* という文の生起確率は、 $Pr(S = \text{ann}, V = \text{eats}, O = \text{fish})$ である。この確率分布は $2^3 = 8$ 状態（文）をもち、 $2 \times 2 \times 2$ の3階確率テンソルとして式 (2) で表される。

$$P = \begin{bmatrix} p_{000} & p_{001} \\ p_{010} & p_{011} \end{bmatrix} \begin{bmatrix} p_{100} & p_{101} \\ p_{110} & p_{111} \end{bmatrix} \in \mathbb{R}_{\geq 0}^{2 \times 2 \times 2} \quad (2)$$

なお、簡略化のために各変数の二状態を $0, 1$ とし、 $p_{ijk} = Pr(S = i, V = j, O = k)$, $i, j, k \in \{0, 1\}$ と表すと、 $\sum_{i,j,k} p_{ijk} = 1$ である。確率テンソル P をベクトル化した確率ベクトル $\mathbf{p} := \text{vec}(P) \in \mathbb{R}_{\geq 0}^8$ では、その成分は $S \succ V \succ O$ により辞書式順序で並ぶ。

2.2 確率テンソルの周辺化と PMI

次に確率テンソルと PMI 行列を関連づける。 n 階テンソル $P \in \mathbb{R}^{\prod_{k \in [n]} I_k}$ の n 個のモード（軸）に対して、その部分集合 $G \subseteq [n] := \{1, \dots, n\}$ による周辺化演算 $\mathcal{M} : \mathbb{R}^{\prod_{k \in [n]} I_k} \rightarrow \mathbb{R}^{\prod_{k \in G} I_k}$ を、補集合 $[n] \setminus G$ に属するモードの総和として式 (3) と定義する。

$$\mathcal{M}_G(P) := \sum_{i_k \in I_k} p_{i_1, \dots, i_k, \dots, i_n} \quad \text{for } k \in [n] \setminus G \quad (3)$$

2.1 節で与えた $2 \times 2 \times 2$ の確率テンソル P に対して、モードの集合を $\{S, V, O\}$ として、 $\mathcal{M}_G(P) = \mu_G$ と表記した上で、 $G = \{S, V\}, \{S\}$ の場合は式 (4) となる

$$\mu_{SV} := \sum_{k \in O} p_{ijk} \in \mathbb{R}^4, \quad \mu_S := \sum_{j \in V, k \in O} p_{ijk} \in \mathbb{R}^2. \quad (4)$$

μ_G は確率変数（の集合）に対する（一次）モーメントである。 μ_G は、 $|G| = 1$ の場合はユニグラム、 $|G| = 2$ の場合は文中の二単語の共起確率（文を超えないバイグラムとスキップグラム）を表している¹⁾。したがって、自然言語処理で用いられる単語共起確率は、確率テンソルを2つのモードずつ周辺化した上で、集約化したものと考えられる（図2）。

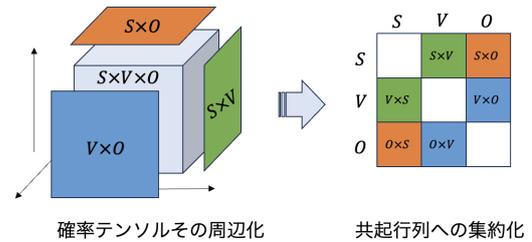


図2 共起行列は確率テンソルの2変数による周辺化

さらに式 (4) 中の二つの式は、追加的な周辺化を行う線型写像で対応づけられている（式 (5)）

$$\mu_S = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \mu_{SV}. \quad (5)$$

これは周辺化を行う写像で二つのモーメントが関係付けられており、これを順序関係とする束 (lattice) を構成することができる。この束上にモーメントに対するキュムラント [12] を式 (6) のように定義できる。

$$\kappa_{S,V} := \log \mu_{SV} - \log \mu_S - \log \mu_V \quad (6)$$

これは PMI の定義と同一であり、PMI が確率テンソルのモーメントがなす束上における2変数と1変数のモーメント間のキュムラントであり、二つの確率変数間の相互作用を評価していると解釈できる。

2.3 確率テンソル中のランク1部分構造

言語の組み合わせ構造を定式化する確率モデルとして、代数統計 [13] を用いた言語モデル [11] を考えると、パラメータに不変な構造として、確率テンソル中に次式 (7) の形をした代数的制約が現れる（導出は Appendix）。確率変数集合の分割 $A \sqcup B \sqcup C = [n] := \{1, \dots, n\}$ と状態 $i_1, i_2 \in \prod_{k \in A} I_k$, $j_1, j_2 \in \prod_{k \in B} I_k$, $k \in \prod_{k \in C} I_k$ に対し

$$p_{i_1, j_1, k} p_{i_2, j_2, k} - p_{i_1, j_2, k} p_{i_2, j_1, k} = 0 \quad (7)$$

が成り立つ。これは確率テンソル P 中の4つのセルの間に成り立つ代数的制約関係を意味する（代数幾

1) 異なる変数に現れる実現値を区別する。例えば S と O に現れる同一単語は別の実現値としてカウントする。

何的には、確率空間に Segre 多様体と呼ばれる低次元の幾何構造をなす). 式の形が 2×2 の小行列式の形 (2-minor), すなわち, $\det \begin{bmatrix} p_{i_1, j_1, k} & p_{i_1, j_2, k} \\ p_{i_2, j_1, k} & p_{i_2, j_2, k} \end{bmatrix} = 0$ と表されることに留意すれば, この代数的制約は確率テンソル中の局所的なランク 1 構造である.

3 数値分析

3.1 実験方法

目的 先行研究 [10] で抽出した単語共起行列における局所的なランク 1 構造が, 高次の局所的なランク 1 部分テンソルに由来するかを実データに基づき検証する. また, 抽出された構造が言語的に解釈可能な意味構造に対応するかを明らかにする.

手順 まず, 頻度上位 10,000 語を対象に, 窓幅 $w = 5$ で算出した PPMI 行列 $X \in \mathbb{R}_+^{n \times n}$ を作成する. 次に, 局所的なランク 1 構造を抽出するため, スパース制約を課した再帰的な非負行列因子分解を適用する. 残差行列 $X^{(m)}$ (ただし $X^{(1)} = X$) に対し, ランク $r = 100$ の分解 $X^{(m)} \approx WH^T$ を行い, 次回の入力として $X^{(m+1)} = X^{(m)} - WH^T$ を求める操作を 10 回繰り返す. L_1 ペナルティを 5×10^{-4} から 10^{-4} まで段階的に減衰させる. 分解された各基底ベクトルにおいて, 成分値の上位 10 位以内かつ累積寄与率 80% を満たす単語集合 (U_k, V_k) を特定し, 最大 10×10 のランク 1 部分行列を構成する (図 1 上段).

得られた 1,000 組の単語集合ペア (U_k, V_k) に対し, それらと同一文脈 ($w = 5$) で共起する第 3 の単語集合 W_k をカウントする. 具体的には, ペア $(u, v) \in U_k \times V_k$ と共起する単語のうち, 出現頻度上位 100 単語 (または累積頻度 80%) の集合を W_k とし, $10 \times 10 \times 100$ のテンソルを構築する (図 1 下段).

データおよび評価指標 コーパスには English Wikipedia dump (20171001) を使用した (約 25.8 億語²⁾). 評価として, 構成単語群の解釈可能性に関する定性分析に加え, 不変条件である 2-minor (式 (7)) の分布, およびスライス行列間での構成ベクトルの方向の一致度を用いた定量評価を行う.

3.2 定性分析結果

抽出された 3 階部分テンソルを構成する上位単語群の例を表 1 に示す. 単語群 (U_k, V_k) は PPMI 行列の局所的なランク 1 成分であり, 先行研究 [10] と

同様, 意味的に解釈可能である. また, 単語集合同士の関係は Syntagma (統括的關係) と Paradigms (連合的關係) の二種類に大別され³⁾, それぞれ三つ組 (U_k, V_k, W_k) が意味的・統語的な関係を示唆する. Syntagma の例として, 補文構造 (ID127) や格構造 (ID311) を反映した単語群が抽出されている. これは, 動詞の格関係や選択制限といった意味を反映した統語的制約が, 代数的制約の結果として局所的なランク 1 構造 (2-minor の消失) を形成していることを示唆する, 一方, 単語集合が重複する Paradigms の例 (ID004 や 502) では, 追加的な W_k の集合も多くの場合重複し, かつ同じ意味範疇に属している. これは, 意味的に置換可能な単語群が, 数理的な対称性を反映した代数的制約として捉えられていることを示す. 統語的な接続関係と意味的な類似性という異なる言語的性質のいずれもが共通の数値構造に帰着する点は興味深く, ランク 1 構造が確率テンソル中に広く存在する可能性を示す.

3.3 定量分析結果

代数的制約の充足度 図 3 は, 単語共起 $U_k \times V_k$ を含む 3 階テンソルを W_k の各単語でスライスした行列を分析している. 各行列のランク 1 純度を示す Sigma Ratio (第 1 特異値の全特異値の総和に対する比率) と, その全ての 2-minor (正規化した上で平均) の分布の関係を示す. 2-minor の分布は Sigma Ratio に対して逆 U 字型の傾向を示しており, Sigma Ratio が 1 に近い領域では式 (7) の消失が不変構造の存在を示す. ランク 1 の行列では全ての小行列式はゼロとなることは線形代数的には当然の帰結ではあるが, 実データにおいて特定の領域でこの制約が厳密に成立していることが確認された.

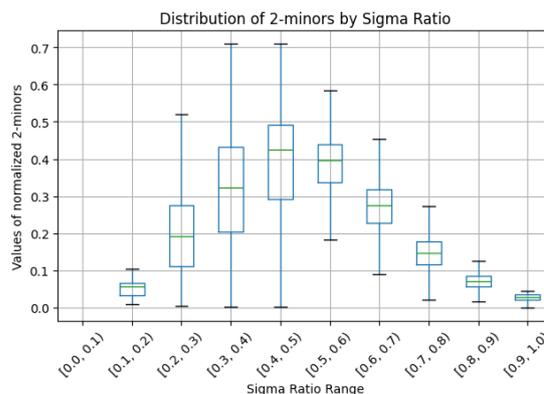


図 3 3 階テンソルにおける 2-minor の分布

2) <https://dumps.wikimedia.org/mowiki/20171001/>

3) 二つの集合間の Jaccard 係数が 0.5 以上か以下で分類.

表 1 抽出された 3 階部分テンソル (ランク 1) の構成単語例

| ID k | 関係性 | 集合 U_k の上位単語 | 集合 V_k の上位単語 | 集合 W_k の上位単語 |
|--------|----------|---|--|------------------------------------|
| 127 | Syntagma | <i>marry, sell, join, ...</i> | <i>decided, persuaded, tried, ...</i> | <i>to, him, but, ...</i> |
| 219 | Syntagma | <i>negotiations, feud, excavations, ...</i> | <i>subsequent, lasted, lengthy, ...</i> | <i>during, after, the, ...</i> |
| 311 | Syntagma | <i>criticized, impressed, warned, ...</i> | <i>strongly, readers, historians ...</i> | <i>that, some, many, ...</i> |
| 004 | Paradigm | <i>technique, methods, modeling, ...</i> | <i>technique, methods, modeling, ...</i> | <i>analysis, testing, methods,</i> |
| 502 | Paradigm | <i>contract, deal, sponsorship, ...</i> | <i>contract, deal, sponsorship, ...</i> | <i>contract, deal, signed, ...</i> |

周辺化における構造保存性 図 4 は, Sigma Ratio と, 周辺化された共起行列 (グローバル) の部分行列および 3 階テンソルの各スライス行列を構成するベクトルの類似度 (Cosine Similarity) の関係を表わす. 周辺化後の行列がランク 1 に近いほど, その 3 階テンソルのスライス行列も同様にランク 1 であり, かつその向きが一致していることがわかる. これは共起行列上に観測される局所的なランク 1 構造が, 周辺化の際の相殺ではなく, 文の生成構造における代数的関係を保存していることを示唆する. 周辺化後のランク 1 性と解釈可能性が対応しているという定性的観察と整合しており, PMI を高次構造の射影像として捉える本研究の視点の妥当性を支持するものである.

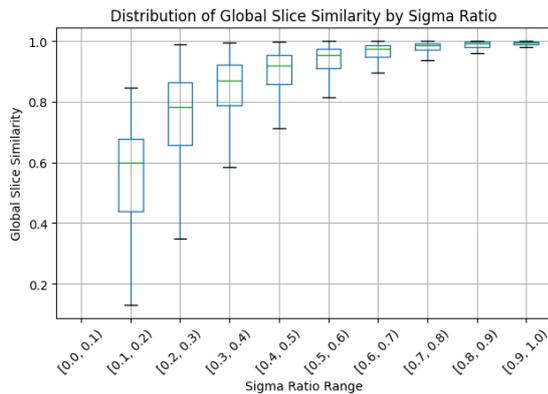


図 4 共起行列とテンソルのスライス行列の方向類似度

4 考察と結論

4.1 分析結果からの洞察

本研究の数値分析により, PMI 行列に観測される局所的なランク 1 構造は, 単なる統計上のアーティファクトではなく, 統語・意味的な対称性が高次元の確率空間において不変な構造を作り, それが周辺化を経て保存された代数的帰結である可能性が示された. 特に図 4 で示された周辺化後の行列とテンソルの各スライス行列間の方向の一致は, 高次元確率空間に内在する数理構造が射影後もその特性を維持していることを示唆する. PMI を単なる二語間の共

起統計としてではなく, 言語の組み合わせ構造の数理的性質を反映した射影像として捉え直す本研究の視点の妥当性を強く支持するものである.

4.2 先行研究

三単語以上の共起分布を分析した初期の先行研究として, [14] は, 構文解析済みコーパスから得られた (s, v, o) の三つ組共起確率 PMI を成分とするテンソルに対して, NMF のテンソル版である PARAFAC (Parallel Factor Analysis) を適用して, 本研究と同様解釈可能な三つ組を抽出している. [15] は複数のテンソル解析手法を比較して PARAFAC の優位性を確認している. [16] は, [1] の生成モデルをテンソルへ拡張しているほか, [17] は n 階テンソル言語モデルを再帰的に分解する手法を提案している.

テンソルを扱う先行研究の多くが用いている CP 分解や Tucker 分解 [18] は, 行列分解における SVD や NMF のように, テンソル全体を一体のモノリシックな構造として扱い一括して分解する. これに対して, 本研究の提案理論が示唆するのは, テンソル中に現れる統計的な規則性は, 局所的な組み合わせ構造を示すことであり, 大域的な分解手法では効果的に抽出することが困難である. 共起行列や PMI 行列を分解して得られる局所的なランク 1 成分をシードとして高次テンソルを探索的に再構成する本アプローチは, 大域的手法では見逃される代数的な不変構造を厳密に特定することを可能にする.

4.3 結び

本研究は, PMI 行列の局所ランク 1 構造を言語の構造に由来する対称性の代数的帰結と捉え, その理論的・実証的基盤を提供した. 3 階までの検証を拡張し, 高次の不変構造を検証することは今後の重要課題である. 本知見は, 分散表現の線形構造が言語データの代数的制約に起因することを示すとともに, 同定される不変構造を帰納バイアスとした学習手法を通じて, 大規模言語モデルの効率的訓練への道筋をも拓くものである.

謝辞

本研究は科研費基盤研究 B(一般) JP23H0369, JST さきがけ JPMJPR20C9, JST CREST JPMJCR23P4, JSPS KAKENHI 24KJ1202 の助成を受けて行われた。

参考文献

- [1] Sanjeev Arora, Yuanzhi Li, Yingyu Liang, Tengyu Ma, and Andrej Risteski. Linear algebraic structure of word senses, with applications to polysemy. **Transactions of the Association for Computational Linguistics**, Vol. 6, pp. 483–495, 2018.
- [2] Carl Allen and Timothy Hospedales. Analogies explained: Towards understanding word embeddings. In **International Conference on Machine Learning**, pp. 223–231. PMLR, 2019.
- [3] Daniel James Korchinski, Dhruva Karkada, Yasaman Bahri, and Matthieu Wyart. On the emergence of linear analogies in word embeddings. In **The Thirty-ninth Annual Conference on Neural Information Processing Systems**, 2025.
- [4] Kiho Park, Yo Joong Choe, and Victor Veitch. The linear representation hypothesis and the geometry of large language models. In **Causal Representation Learning Workshop at NeurIPS 2023**, 2023.
- [5] Yibo Jiang, Goutham Rajendran, Pradeep Kumar Ravikumar, Bryon Aragam, and Victor Veitch. On the origins of linear representations in large language models. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp, editors, **Proceedings of the 41st International Conference on Machine Learning**, Vol. 235 of **Proceedings of Machine Learning Research**, pp. 21879–21911. PMLR, 21–27 Jul 2024.
- [6] Omer Levy and Yoav Goldberg. Neural word embedding as implicit matrix factorization. **Advances in neural information processing systems**, Vol. 27, , 2014.
- [7] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. **Advances in neural information processing systems**, Vol. 26, , 2013.
- [8] Peter D Turney and Patrick Pantel. From frequency to meaning: Vector space models of semantics. **Journal of artificial intelligence research**, Vol. 37, pp. 141–188, 2010.
- [9] John A Bullinaria and Joseph P Levy. Extracting semantic representations from word co-occurrence statistics: A computational study. **Behavior research methods**, Vol. 39, No. 3, pp. 510–526, 2007.
- [10] Akihiro Maeda, Takuma Torii, and Shohei Hidaka. Decomposing co-occurrence matrices into interpretable components as formal concepts. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, **Findings of the Association for Computational Linguistics: ACL 2024**, pp. 4683–4700, Bangkok, Thailand, August 2024. Association for Computational Linguistics.
- [11] 前田晃弘, 鳥居拓馬, 日高昇平. 言語構造の数理分析のための代数統計的アプローチ試論. 言語処理学会第 31 回年次大会 発表論文集, 2025.
- [12] Martin J Wainwright, Michael I Jordan, et al. Graphical models, exponential families, and variational inference. **Foundations and Trends® in Machine Learning**, Vol. 1, No. 1–2, pp. 1–305, 2008.
- [13] Seth Sullivant. **Algebraic statistics**, Vol. 194. American Mathematical Society, 2023.
- [14] Tim Van de Cruys. A non-negative tensor factorization model for selectional preference induction. In Roberto Basili and Marco Pennacchiotti, editors, **Proceedings of the Workshop on Geometrical Models of Natural Language Semantics**, pp. 83–90, Athens, Greece, March 2009. Association for Computational Linguistics.
- [15] Márton Makrai. Three-order normalized pmi and other lessons in tensor analysis of verbal selectional preferences.
- [16] Abraham Frandsen and Rong Ge. Understanding composition of word embeddings via tensor decomposition. **arXiv preprint arXiv:1902.00613**, 2019.
- [17] Lipeng Zhang, Peng Zhang, Xindian Ma, Shuqin Gu, Zhan Su, and Dawei Song. A generalized language model in tensor space. In **Proceedings of the AAAI Conference on Artificial Intelligence**, Vol. 33, pp. 7450–7458, 2019.
- [18] Tamara G. Kolda and Brett W. Bader. Tensor decompositions and applications. **SIAM Review**, Vol. 51, No. 3, pp. 455–500, 2009.

A 代数統計的言語モデル

本研究は、著者らの先行研究 [11] で提案した代数統計的アプローチとしてトーリックモデルにより定式化された言語モデルに基づく。

A.1 トーリックモデル

トーリックモデルとは、指数型分布族のサブクラスであり、配置行列と呼ばれる二値行列 $A \in \{0, 1\}^{d \times N}$ と d 次元のパラメータベクトル $\theta \in \mathbb{R}^d$ に対して、その確率ベクトル $\mathbf{p} \in \mathbb{R}^N$ は式 (8) に従うとする

$$\mathbf{p}^\top = \frac{1}{Z} \exp(\theta^\top A); \quad Z = \sum_j \exp(\theta^\top A^j). \quad (8)$$

なお、 A^j は A の j 列ベクトルであり、 Z は確率の正規化項である。

配置行列の行はパラメータに、列は確率の状態に対応しており、配置行列の (i, j) 成分 A_{ij} が 1 であるとき、 i 番目のパラメータが j 番目の状態に対応する確率に関与していることを表す。言語モデルとしては、配置行列は単語や句あるいは隠れ状態としての文脈が文の生起確率に与える影響を記述していると考えることができる。

A.2 消失二項式の導出

確率ベクトルが式 (8) で与えられる時、その成分の間には配置行列の行空間の構造を反映した代数的制約が生じる。具体的には、行列 A の整数カーネル (核) を

$$\ker_{\mathbb{Z}} A := \{u \in \mathbb{Z}^N \mid Au = 0\} \quad (9)$$

により与える。この時、 $u - v \in \ker_{\mathbb{Z}} A$ となる $u, v \in \mathbb{Z}_{\geq 0}^N$ に対して、

$$\mathbf{p}^u - \mathbf{p}^v = 0 \quad (10)$$

が成り立つ。ただし、 $\mathbf{p}^u := p_1^{u_1} p_2^{u_2} \cdots p_N^{u_N}$ である (代数統計の基本定理 [13])。

式 (10) は、二項式の形をしておりゼロの値を取るため、代数統計において消失二項式 (vanishing binomial)、あるいはこれを生成式とするイデアルを消失イデアル (vanishing ideal) と呼ぶ。代数幾何的には、式 (10) を満たす確率ベクトル (p_1, p_2, \dots, p_N) の集合 (零点集合) は、 N 次元空間 (正確には N 次元シンプレックス) において、代数多様体と呼ばれる幾何構造をなす。特に、式 (7) に対応する代数多様体は、Segre 多様体と呼ばれる。

A.3 条件付き独立分布の例

2.1 節で例示したトイモデルとして、 S, V, O を三つの確率変数とする同時確率分布において、 SV, SO が交互作用を持つ確率モデルを考える。グラフィカルモデルを用いて、これは条件付き独立分布であり $V \perp\!\!\!\perp O \mid S$ である。この確率分布をトーリックモデルとして表現するための配置行列を示す。

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \quad (11)$$

式 (8) で与えられる確率ベクトルを $2 \times 2 \times 2$ のテンソルで表すことができるので、その座標を $\mathbf{p} = (p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111})$ とする。次の二つのベクトル

$$\mathbf{b}_1 := [1 \quad -1 \quad -1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0]^\top \quad (12)$$

$$\mathbf{b}_2 := [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad -1 \quad -1 \quad 1]^\top \quad (13)$$

は、 $A\mathbf{b}_1 = A\mathbf{b}_2 = 0$ より $\mathbf{b}_1, \mathbf{b}_2 \in \ker_{\mathbb{Z}} A$ である。行列 A のランクは 6、行空間は 8 次元であるので、二つの独立なベクトル $\{\mathbf{b}_1, \mathbf{b}_2\}$ は、 $\ker_{\mathbb{Z}} A$ の基底である。A.2 にて示した代数統計の基本定理により、 $\mathbf{u}_1 := [1, 0, 0, 1, 0, 0, 0, 0]^\top$, $\mathbf{v}_1 := [0, 1, 1, 0, 0, 0, 0, 0]^\top$ とすると、 $\mathbf{b}_1 = \mathbf{u}_1 - \mathbf{v}_1 \in \ker_{\mathbb{Z}} A$ より

$$\mathbf{p}^{\mathbf{u}_1} - \mathbf{p}^{\mathbf{v}_1} = p_{000}p_{011} - p_{001}p_{010} = 0 \quad (14)$$

が成り立つ。同様に、 $\mathbf{b}_2 = \mathbf{u}_2 - \mathbf{v}_2 \in \ker A$ となる $\mathbf{u}_2, \mathbf{v}_2$ を考えることができ、

$$\mathbf{p}^{\mathbf{u}_2} - \mathbf{p}^{\mathbf{v}_2} = p_{100}p_{111} - p_{101}p_{110} = 0 \quad (15)$$

が成り立つ。式 (14)(15) は確率テンソル中の小行列式をゼロとする (図 5)。

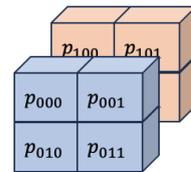


図 5 確率テンソルにおける局所的なランク 1 構造