

マルチエージェントネットワークによる非負値行列因子分解の分散計算 Distributed Computation of Nonnegative Matrix Factorization by Multiagent Networks

道免 陽平[†] 右田 剛史[†] 高橋 規一[†]
Yohei Domen Tsuyoshi Migita Norikazu Takahashi

1 はじめに

非負値行列因子分解 (NMF: Nonnegative Matrix Factorization) とは、与えられた非負値行列 $\mathbf{X} \in \mathbb{R}^{M \times N}$ を二つの非負値行列 $\mathbf{W} \in \mathbb{R}^{M \times K}$, $\mathbf{H}^T \in \mathbb{R}^{K \times N}$ の積で近似する多変量解析の手法であり [1, 2], 信号処理, テキスト分類, ネットワーク分析, 推薦システムなどに広く利用されている. NMF は非負制約付き最適化問題として定式化され, その代表的解法として乗法型更新や階層的交互最小二乗 (HALS: Hierarchical Alternating Least Squares) 法 [3] などが知られている. しかしながら, \mathbf{X} が大規模行列の場合には \mathbf{X} 全体を記憶できる容量と高速な計算能力を有する高性能計算機が必要になる.

本稿では, 最も基本的なユークリッド距離に基づく NMF に着目し, その最適化問題をマルチエージェントネットワークによって分散的に解くアルゴリズムを提案する. 制御工学や信号処理の分野では, 近年, マルチエージェントネットワークによる分散計算が大きな注目を集めており, 様々な分散アルゴリズムが提案されている [4]. 本稿で提案するアルゴリズムは, HALS 法と合意アルゴリズム [5] を組み合わせたものであり, 単一の計算機で HALS 法を実行するのと同じ結果が得られる点に特徴がある.

2 非負値行列因子分解と HALS 法

ユークリッド距離に基づく NMF は次の最適化問題に定式化される.

$$\begin{aligned} & \text{minimize} && f(\mathbf{W}, \mathbf{H}) = \|\mathbf{X} - \mathbf{W}\mathbf{H}^T\|_F^2 \\ & \text{subject to} && \mathbf{W} \geq \mathbf{0}_{M \times K}, \mathbf{H} \geq \mathbf{0}_{N \times K} \end{aligned} \quad (1)$$

ただし, $\|\cdot\|_F$ はフロベニウスノルムを表す. すなわち

$$\|\mathbf{X} - \mathbf{W}\mathbf{H}^T\|_F^2 = \sum_{m=1}^M \sum_{n=1}^N (X_{mn} - (\mathbf{W}\mathbf{H}^T)_{mn})^2$$

である. また, $\mathbf{0}_{M \times K}$ ($\mathbf{0}_{N \times K}$) は $M \times K$ ($N \times K$) 零行列を表し, 制約条件の不等式は行列の要素ごとに適用されるものとする. すなわち, 制約条件は \mathbf{W} と \mathbf{H} のすべての要素が非負であることを意味する.

問題 (1) の高速解法の一つである HALS 法 [3] は, \mathbf{W} と \mathbf{H} をそれぞれ $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_K)$, $\mathbf{H} = (\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K)$ と K 個の列ベクトルに分け, $2K$ 個の列ベクトルの値を一つずつ, 目的関数の値が最も大きく減少するように更新するものである. しかし, HALS 法には変数の値によっては更新式が定義されないという問題がある. そこで変数の値がある正定数 ϵ より小さくならないように更新式を修正した修正 HALS 法が提案された [3]. その後, Kimura と Takahashi は, 修正 HALS 法が大域収束性を有する, すなわち, 修正 HALS 法によって生成される

アルゴリズム 1 修正 HALS 法

入力: $\mathbf{X} \geq \mathbf{0}_{M \times N}$, $\epsilon, \delta_1, \delta_2 > 0$, $\mathbf{W}_0 \geq \epsilon \mathbf{1}_{M \times K}$, $\mathbf{H}_0 \geq \epsilon \mathbf{1}_{N \times K}$
出力: $\mathbf{W} \geq \epsilon \mathbf{1}_{M \times K}$, $\mathbf{H} \geq \epsilon \mathbf{1}_{N \times K}$

- 1: $\mathbf{W} \leftarrow \mathbf{W}_0, \mathbf{H} \leftarrow \mathbf{H}_0$ とする.
- 2: 条件 (3)–(6) が満たされていれば \mathbf{W} と \mathbf{H} を出力して終了する.
- 3: $k = 1, 2, \dots, K$ の順に $\mathbf{R}_k \leftarrow \mathbf{X} - \sum_{\ell=1, \ell \neq k}^K \mathbf{w}_\ell \mathbf{h}_\ell^T$, $\mathbf{w}_k \leftarrow \left[\mathbf{R}_k \mathbf{h}_k / \|\mathbf{h}_k\|_2 \right]_{\epsilon+}$ とする.
- 4: $k = 1, 2, \dots, K$ の順に $\mathbf{R}_k \leftarrow \mathbf{X} - \sum_{\ell=1, \ell \neq k}^K \mathbf{w}_\ell \mathbf{h}_\ell^T$, $\mathbf{h}_k \leftarrow \left[\mathbf{R}_k^T \mathbf{w}_k / \|\mathbf{w}_k\|_2 \right]_{\epsilon+}$ とする.
- 5: ステップ 2 に戻る.

(\mathbf{W}, \mathbf{H}) の値の列は少なくとも一つの収束部分列をもち, すべての収束部分列の極限が最適化問題

$$\begin{aligned} & \text{minimize} && f(\mathbf{W}, \mathbf{H}) = \|\mathbf{X} - \mathbf{W}\mathbf{H}^T\|_F^2 \\ & \text{subject to} && \mathbf{W} \geq \epsilon \mathbf{1}_{M \times K}, \mathbf{H} \geq \epsilon \mathbf{1}_{N \times K} \end{aligned} \quad (2)$$

の停留点であることを証明した [6]. ここで, $\mathbf{1}_{M \times K}$ ($\mathbf{1}_{N \times K}$) はすべての要素が 1 の $M \times K$ ($N \times K$) 行列を表し, (\mathbf{W}, \mathbf{H}) が (2) の停留点であるとは, (2) に対する KKT (Karush-Kuhn-Tucker) 条件を満たすことを意味する. さらに彼らは, KKT 条件を正定数 δ_1, δ_2 を用いて緩めた

$$(\mathbf{W}\mathbf{H}^T - \mathbf{X})\mathbf{H} \geq -\delta_1 \mathbf{1}_{M \times K} \quad (3)$$

$$(\mathbf{H}\mathbf{W}^T - \mathbf{X}^T)\mathbf{W} \geq -\delta_1 \mathbf{1}_{N \times K} \quad (4)$$

$$W_{mk} - \epsilon \leq \delta_2 \text{ if } ((\mathbf{W}\mathbf{H}^T - \mathbf{X})\mathbf{H})_{mk} > \delta_1, \forall (m, k) \quad (5)$$

$$H_{nk} - \epsilon \leq \delta_2 \text{ if } ((\mathbf{H}\mathbf{W}^T - \mathbf{X}^T)\mathbf{W})_{nk} > \delta_1, \forall (n, k) \quad (6)$$

を終了条件に用いることにより, 修正 HALS 法が有限回の反復で終了することを示した [6].

終了条件付き修正 HALS 法をアルゴリズム 1 に示す.

3 非負値行列因子分解の分散計算法

単一エージェントによる HALS 法と同じ計算を複数のエージェントで分散的に実行する方法を提案する. 非負値行列 \mathbf{X} を $I \times J$ 個のブロックに分割し, それに応じて行列 \mathbf{W}, \mathbf{H}^T もそれぞれ I 個, J 個のブロックに分割する (図 1 参照). \mathbf{X} の上から i 番目, 左から j 番目のブロックを $\mathbf{X}^{(ij)} \in \mathbb{R}_+^{M_i \times N_j}$ で表し, \mathbf{W} の上から i 番目のブロックと \mathbf{H}^T の左から j 番目のブロックをそれぞれ $\mathbf{W}^{(i)} \in \mathbb{R}_+^{M_i \times K}$, $\mathbf{H}^{(j)T} \in \mathbb{R}_+^{K \times N_j}$ で表す. このとき, 問題 (2) は次のように書き換えられる.

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^I \sum_{j=1}^J \left\| \mathbf{X}^{(ij)} - \mathbf{W}^{(i)} \mathbf{H}^{(j)T} \right\|_F^2 \\ & \text{subject to} && \mathbf{W}^{(i)} \geq \epsilon \mathbf{1}_{M_i \times K}, \quad i = 1, 2, \dots, I \\ & && \mathbf{H}^{(j)} \geq \epsilon \mathbf{1}_{N_j \times K}, \quad j = 1, 2, \dots, J \end{aligned} \quad (7)$$

さらに変数 $\mathbf{W}^{(ij)} \in \mathbb{R}^{M_i \times K}$, $\mathbf{H}^{(ij)} \in \mathbb{R}^{N_j \times K}$ ($i = 1, 2, \dots, I$, $j = 1, 2, \dots, J$) を導入すると, 問題 (7) は次のように書き

[†] 岡山大学大学院自然科学研究科 Graduate School of Natural Science and Technology, Okayama University

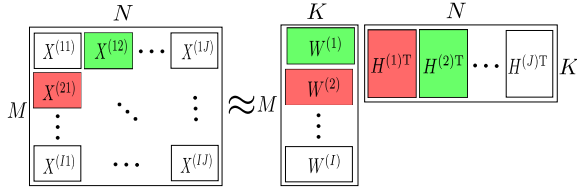


図1 非負値行列因子分解における行列の分割

アルゴリズム2 マルチエージェントネットワークによるNMFの分散計算

入力: $X \geq \mathbf{0}_{M \times N}$, $\epsilon, \delta_1, \delta_2 > 0$, $W_0^{(i)} \geq \epsilon \mathbf{1}_{M_i \times K}$ ($i = 1, 2, \dots, I$), $H_0^{(j)} \geq \epsilon \mathbf{1}_{N_j \times K}$ ($j = 1, 2, \dots, J$)

出力: $W \geq \epsilon \mathbf{1}_{M \times K}$, $H \geq \epsilon \mathbf{1}_{N \times K}$

- 各エージェントは自身の変数を $W^{(ij)} \leftarrow W_0^{(i)}$, $H^{(ij)} \leftarrow H_0^{(j)}$ と初期化する。
- 各エージェントは自身の変数が(3)–(6)の該当部分を満たしているかを分散的に判定し、満たしていれば $s^{(ij)} \leftarrow 1$ とし、そうでなければ $s^{(ij)} \leftarrow 0$ とする。
- 各エージェントは最小合意アルゴリズムを用いて $\{s^{(ij)}\}$ の最小値を求める。それが1ならば

$$W = \left(W^{(11)T}, W^{(21)T}, \dots, W^{(I1)T} \right)^T$$

$$H = \left(H^{(11)T}, H^{(12)T}, \dots, H^{(1J)T} \right)^T$$

を出力して終了する。

- 各エージェントはアルゴリズム3を用いて $W^{(ij)}$ の値を更新する。
- 各エージェントはアルゴリズム3と同様にして $H^{(ij)}$ の値を更新し、ステップ2に戻る。

換えられる。

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^I \sum_{j=1}^J \left\| X^{(ij)} - W^{(ij)} H^{(ij)T} \right\|_F^2 \\ & \text{subject to} && W^{(ij)} \geq \epsilon \mathbf{1}_{M_i \times K}, \quad i = 1, 2, \dots, I, j = 1, 2, \dots, J \\ & && H^{(ij)} \geq \epsilon \mathbf{1}_{N_j \times K}, \quad i = 1, 2, \dots, I, j = 1, 2, \dots, J \\ & && W^{(i1)} = W^{(i2)} = \dots = W^{(iJ)}, \quad i = 1, 2, \dots, I \\ & && H^{(1j)} = H^{(2j)} = \dots = H^{(Ij)}, \quad j = 1, 2, \dots, J \end{aligned} \quad (8)$$

また、条件(3)–(6)も $W^{(ij)}$, $H^{(ij)}$ ($i = 1, 2, \dots, I, j = 1, 2, \dots, J$) を用いて書き換えられる(紙面の都合上詳細は省略する)。

問題(8)の局所的最適解を $I \times J$ 個のエージェント $a^{(ij)}$ ($i = 1, 2, \dots, I, j = 1, 2, \dots, J$) に分散的に求めさせたい。エージェント $a^{(ij)}$ は変数 $W^{(ij)}$, $H^{(ij)}$ のみを扱い、エージェント $a^{(i-1,j)}$, $a^{(i+1,j)}$, $a^{(i,j-1)}$, $a^{(i,j+1)}$ と情報交換を行いながら自身の変数の値を更新するとする。この枠組みで $I \times J$ 個のエージェントが(8)の局所的最適解を求める方法として、アルゴリズム2を提案する。

アルゴリズム2が分散的であることは容易にわかる。それだけでなく、次の定理に示すように、アルゴリズム1と等価である(紙面の都合上証明は省略する)。

定理1 アルゴリズム1とアルゴリズム2の入力が $W_0 = \left(W_0^{(1)T}, W_0^{(2)T}, \dots, W_0^{(I)T} \right)^T$, $H_0 = \left(H_0^{(1)T}, H_0^{(2)T}, \dots, H_0^{(J)T} \right)^T$ を満たすならば、各反復にお

アルゴリズム3 エージェント $a^{(ij)}$ ($j = 1, 2, \dots, J$) による $W^{(ij)}$ ($j = 1, 2, \dots, J$) の更新

- $k \leftarrow 1$ とする。
- 各エージェントは次式により $p_k^{(ij)}$ を求める。

$$p_k^{(ij)} \leftarrow \left(X^{(ij)} - \sum_{\ell=1, \ell \neq k}^K w_\ell^{(ij)} h_\ell^{(ij)T} \right) h_k^{(ij)}$$

- 各エージェントは平均合意アルゴリズム[5]を用いて $p_k^{(i1)}, p_k^{(i2)}, \dots, p_k^{(iJ)}$ の平均値 $p_k^{(i)} = \sum_{j=1}^J p_k^{(ij)} / J$ を求める。
- 各エージェントは次式により $q_k^{(ij)}$ を求める。

$$q_k^{(ij)} \leftarrow \left\| h_k^{(ij)} \right\|_2^2$$

- 各エージェントは平均合意アルゴリズム[5]を用いて $q_k^{(i1)}, q_k^{(i2)}, \dots, q_k^{(iJ)}$ の平均値 $q_k^{(i)} = \sum_{j=1}^J q_k^{(ij)} / J$ を求める。
- 各エージェントは次式で $w_k^{(ij)}$ の値を更新する。

$$w_k^{(ij)} \leftarrow \left[p_k^{(i)} / q_k^{(i)} \right]_{\epsilon+}$$

- $k = K$ ならば終了する。そうでなければ $k \leftarrow k + 1$ としてステップ2に戻る。

いて $W = \left(W^{(1j)T}, W^{(2j)T}, \dots, W^{(Ij)T} \right)^T$ ($j = 1, 2, \dots, J$),

$H = \left(H^{(i1)T}, H^{(i2)T}, \dots, H^{(iJ)T} \right)^T$ ($i = 1, 2, \dots, I$) が成り立つ。また、任意の入力に対して、アルゴリズム2は有限回の反復で終了する。

4 おわりに

本稿では、ユークリッド距離に基づくNMFの最適解問題をマルチエージェントネットワークによって分散的に解く方法を提案し、それが修正HALS法と等価であることを示した。NMFにおける誤差評価には、ユークリッド距離だけでなく様々なダイバージェンスが用いられる。それらに対応可能な分散アルゴリズムを開発することは今後の課題である。

参考文献

- P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol.5, no.2, pp.111–126 (1994).
- D.D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol.401, no.6755, pp.788–791 (1999).
- A. Cichocki, R. Zdunek and S. Amari, "Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization," *Lecture Notes in Computer Science*, vol.4666, pp.169–176 (2007).
- N. Takahashi and K. Kawashima, "A simple sufficient condition for convergence of projected consensus algorithm," *IEEE Control Systems Letters*, vol.2, no.3, pp.537–542 (2018).
- S. Sundaram and C. N. Hadjicostis, "Finite-time distributed consensus in graphs with time-invariant topologies," *Proceedings of 2007 American Control Conference*, pp.711–716 (2007).
- T. Kimura and N. Takahashi, "Global convergence of a modified HALS algorithm for nonnegative matrix factorization," *Proceedings of 2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pp. 21–24 (2015).