

独立性基準を用いた非負値行列因子分解の効果的な初期値決定法*

○北村大地 (総研大), 小野順貴 (NII/総研大)

1 はじめに

非負値行列因子分解 (nonnegative matrix factorization: NMF) [1] とは, 非負の入力データ行列を別の二つの非負行列の行列積に近似的に分解する手法であり, 全ての行列要素に非負の制約条件を与えた次元圧縮と捉えられる. NMF で分解される基底ベクトルは, データ行列の表現に必要な非負のパーツであり, それらの混ざり方は足し算のみ (非負の重み係数を伴う基底の線形結合) が許容される. そのため, NMF で得られる基底行列や係数行列は我々が直感的に解釈できる可能性が高い. 従って, データ行列中の重要な潜在パターンを得る教師無し学習法として広く活用されている. 音響信号に NMF を適用する場合, 短時間フーリエ変換 (short-time Fourier transform: STFT) を経て得られる振幅あるいはパワースペクトログラムをデータ行列と考えるのが一般的である.

NMF では, 観測データ行列と近似的に構成されるモデルの近さをコスト関数として定義し, 最適化問題を解くことで基底行列及び係数行列が推定される. 但し, NMF のコスト関数は通常, 推定すべきパラメータ (基底行列と係数行列の各要素) に対して凸関数ではなく, 局所的な解が無数に存在する. 基底行列と係数行列に初期値を与えた上で, 反復更新によって解を求める手法 [2, 3] が確立されているが, このような反復更新による推定結果も最適解である保証は無く, 初期値に依存して結果が左右されてしまう問題がある.

本稿では, NMF の初期値依存問題の解決を目的として, 独立成分分析 (independent component analysis: ICA) [4] を活用した効果的な初期値決定法を提案する. 提案法を初期値に用いた NMF は, 主成分分析 (principal component analysis: PCA) や特異値分解 (singular value decomposition: SVD) に基づく初期値よりも高速に良い解に収束することを実験的に示す.

2 NMF と従来の初期値決定法

2.1 定式化

非負の入力データ行列を $D \in \mathbb{R}_+^{M \times N}$, NMF で推定する基底行列を $F \in \mathbb{R}_+^{M \times K}$, 係数行列を $G \in \mathbb{R}_+^{K \times N}$ とする. M 及び N は D の行及び列の数, K は NMF で分解する際の基底数を示す. 通常 NMF は, 次元圧縮を経て限られた数 (K 個) の有意な特徴量を抽出することが目的であるため, 基底数 K は $\min(M, N)$ よりも十分小さな値に設定する. Figure 1 は, パワースペクトログラム D を $K=2$ で NMF によって分解した結果例を示している. 得られた基底行列 F は, スペクトログラム中に頻出するスペクトルパターンを示しており, 係数行列 G は, 各基底の時間的なアクティベーションを示している. このような性質から, NMF による分解は様々な用途に容易に活用できる利点がある. 特に音響信号に対しては, 音源分離における応用が盛んに研究されている [5]–[12].

NMF における最適化問題は次式で定式化される.

$$\arg \min_{F, G} \mathcal{D}(D \| FG) \quad \text{s.t. } f_{m,k}, g_{k,n} \geq 0 \quad \forall m, n, k \quad (1)$$

ここで, $\mathcal{D}(\cdot \| \cdot)$ はモデルと観測の近さを表すコスト関数, $f_{m,k} = [F]_{m,k}$, $g_{k,n} = [G]_{k,n}$, m, n , 及び k はそれぞれ

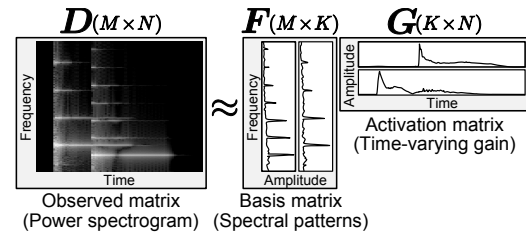


Fig. 1 Decomposition model of NMF ($K=2$).

れ行, 列, 及び基底のインデックスである. コスト関数には, 二乗ユークリッド距離 (EU), 一般化カルバック・ライブラダイバージェンス (KL), 板倉斎藤擬距離 (IS) 等がよく用いられるが, 変数 F 及び G の同時最適化を考えた場合はいずれのコスト関数も非凸であるため, F 及び G の初期値 ($F^{(0)}, G^{(0)}$) を何らかの手法で初期化して反復更新を行い, 局所解を得るのが一般的である. 尚, $F^{(i)}$ 及び $G^{(i)}$ は i 回更新後の基底及び係数行列を表す.

2.2 従来の初期値決定法

NMF における「良い初期値」とは, コスト関数の収束が速いことと, より小さい値まで目的関数が減少することの二点を満たすものである [13]. NMF 変数の初期化として最も単純な手法は, 擬似乱数を生成して ($F^{(0)}, G^{(0)}$) に与えることであるが, 試行のたびに異なる結果を導くため, 複数回試行した上で最良のものを選択する必要がある. 他には, 遺伝的アルゴリズムを用いてより効率的に良い初期値を探索する手法 [14, 15], D をクラスタリングし各クラスタのセントロイドを $F^{(0)}$ とする手法 [16]–[18] が提案されているが, いずれの手法も内部に乱数が必要なため, 根本的な解決にはなっていない. 乱数を全く用いないクラスタリングによる初期化法として, データ行列に対して減算クラスタリングを用いる手法 [19] も提案されているが, 良い初期値を決めるにはクラスタ間の許容距離等のハイパーパラメータの調整が必要となっている. その他, 乱数とハイパーパラメータのいずれも必要としない初期化法として, D に PCA を適用して得られる基底行列及び重み行列の絶対値を ($F^{(0)}, G^{(0)}$) とする手法 [20], D に非負二重 SVD (nonnegative double SVD: NNDSVD) を適用して得られる非負の左及び右特異ベクトルを ($F^{(0)}, G^{(0)}$) とする手法 [13] が提案されている. この二手法は X のみに依存する初期化法であるため, その後の応用タスクにおいても一意な結果を与えるものである.

3 独立性基準を用いた初期値決定法

3.1 提案法の動機と概要

PCA や SVD は基底の直交性に基づく次元圧縮であるため, 途中で非負化を施していたとしても, 基本的には基底間の直交性を最大化するような ($F^{(0)}, G^{(0)}$) を求めている. しかしながら, NMF で推定されるべき最良の基底は直交基底とは限らない. NMF は, 正の象限の凸多面錐 (convex polyhedral cone) 内に点在するデータを表現するため, その凸多面錐の端に沿った基底ベクトルが最も過不足のない理想的な解であ

* Statistical-independence-based effective initialization for nonnegative matrix factorization by Daichi Kitamura (SOK-ENDAI), Nobutaka Ono (NII/SOKENDAI)

る [21]. 従って, PCA や SVD のように直交性最大化基準で初期値を求める手法は, 最良な初期化とはいえない可能性がある.

本稿では, 乱数を用いない新しい初期化法として, ICA を用いた初期値決定法を提案する. ICA において独立成分に優ガウス分布を仮定した場合, 優ガウス分布とは原点 (O) 付近の鋭いピークと長い裾野をもつ分布であることから, 推定された独立成分は多数の O 近傍の値と少数の大きな値をもつ成分, すなわちスパースな成分が得られると期待される. NMF も非負性に起因するスパース分解と解釈できるため, ICA で推定される独立成分による初期化は良い分解結果をもたらす可能性が高い. 以上より, 直交性最大化基準ではなく, 独立性最大化基準で推定される ICA の基底及び独立成分を, それぞれ初期基底行列 $\mathbf{F}^{(0)}$ 及び初期係数行列 $\mathbf{G}^{(0)}$ とする. 但し, ICA 単体では次元を圧縮できないため, PCA と組み合わせることで, 限られた数 (K 個) の ICA 基底と独立成分を推定する¹. また, ICA は通常, 零平均の非ガウスな分布を独立成分の統計的な生成モデルとして仮定するため, 正負値を持つ信号を対象としている. NMF の初期値として活用するためには, 非負の独立成分を推定する必要があるため, 本稿では nonnegative ICA (NICA) [22] を適用する手法と, 非負独立成分の差分信号を通常の ICA で推定する手法の二つの初期化アルゴリズムを提案する.

3.2 PCA による次元圧縮と ICA の適用

任意の $M \times N$ の行列 \mathbf{X} を考える. 但し, \mathbf{X} の行は M 個の変数, 列は各変数に対応する N 個のデータサンプルである. PCA を適用して M 個の変数を K 個に次元圧縮する処理は, 次式のように表現できる.

$$\begin{cases} \mathbf{P}_1 \mathbf{X} = \mathbf{A} \mathbf{S} \\ \mathbf{P}_2 \mathbf{X} \approx \mathbf{0} \end{cases} \quad (2)$$

ここで, サイズ M の正方行列 $\mathbf{P} = (\mathbf{P}_1^T \ \mathbf{P}_2^T)^T$ を \mathbf{X} に対する PCA の変換行列とする. 但し, \mathbf{P} の行ベクトルは \mathbf{X} の分散共分散行列 $\mathbf{X} \mathbf{X}^T$ の固有ベクトルであり, 1 番目の行から固有値の高い順に並べている. 従って, \mathbf{P}_1 は最大固有値から K 番目の固有値までのそれぞれに対応する固有ベクトルを行に持つ $K \times M$ の行列であり, \mathbf{P}_2 は残りの固有ベクトルを行に持つ $(M-K) \times M$ の行列である. また, \mathbf{A} は $K \times K$ の混合行列であり, \mathbf{S} は K 個の独立成分 (ソース) を行に持つ $K \times N$ の行列である. さらに, $\mathbf{0}$ は $(M-K) \times N$ の零行列を示す. 従って, 独立成分 \mathbf{S} が混合行列 \mathbf{A} を経て $\mathbf{P}_1 \mathbf{X}$ として観測されていると仮定している. $\mathbf{P}_1 \mathbf{X}$ に ICA を適用することで分離行列 \mathbf{W} と分離信号が推定できる.

3.3 NICA を用いた NMF の初期値決定法

NICA [22] は, 非負の独立成分が混合された状態で観測される多次元信号から, 混合前の各成分を推定する手法である. その本質は, 中心化せずに白色化 (無相関化) したデータに対して, 全ての成分が非負となるような回転行列 \mathbf{W} を求めることである².

$$\mathbf{Z} = \mathbf{V} \mathbf{P}_1 \mathbf{X} = \mathbf{V} \mathbf{A} \mathbf{S} \quad (3)$$

$$\mathbf{Y} = \mathbf{W} \mathbf{Z} \quad (4)$$

¹ データ行列に対して PCA を適用した後で ICA を適用するため, 従来の直交性基準の基底の推定とは異なることに注意する.

² 回転変換を表す直交行列 \mathbf{W} の持つ NICA 基底は, 中心化せずに白色化されたデータに対する直交基底であるため, 白色化前の元データに対しては直交基底になっていないことに注意する.

ここで, \mathbf{V} は中心化せずに白色化を施す変換行列, \mathbf{Y} は分離信号を示し, $\mathbf{X} = \mathbf{D}$ とする. 従って, 実際に混合されている独立成分が全て非負である場合は, \mathbf{Y} のスケール及び順番の任意性はさておき, 最適解が原理的に推定可能である. しかしながら, 前節のような次元圧縮を施している場合は, いかなる回転行列を与えても全分離信号が非負となる解は基本的に存在しないため, 次式のように, \mathbf{Y} 中に残った負の成分のパワーが最小となるような分離行列 \mathbf{W} を推定することになる.

$$\arg \min_{\mathbf{W}} \sum_{k,n} \min(0, y_{k,n})^2 \quad (5)$$

ここで, $y_{k,n} = [\mathbf{Y}]_{k,n}$ である. この最適化問題に対して, 次式で表される最急降下法 [23] が提案されている.

$$\tilde{\mathbf{w}}_k = \mathbf{w}_k - 2\gamma \sum_n \min(0, y_{k,n}) z_{k,n} \quad (6)$$

$$\mathbf{W} = (\tilde{\mathbf{W}} \tilde{\mathbf{W}}^T)^{-1/2} \tilde{\mathbf{W}} \quad (7)$$

ここで, \mathbf{w}_k^T は \mathbf{W} の行ベクトル, γ はステップサイズパラメータ, $z_{k,n} = [\mathbf{Z}]_{k,n}$, $\tilde{\mathbf{W}}$ は $\tilde{\mathbf{w}}_k^T$ を行ベクトルに持つ行列である. その他, ステップサイズパラメータを持たない nonnegative FastICA [24] も提案されているが, 本稿では上記の最急降下法を用いる.

推定された分離信号 \mathbf{Y} は, 独立性を最大化した係数行列 $\mathbf{G} = \mathbf{Y}$ としてそのまま利用できる. 一方, 基底行列 \mathbf{F} は, ICA で推定される基底 \mathbf{A} から与えられる. 近似的に $\mathbf{X} = \mathbf{F} \mathbf{G}$, $\mathbf{S} = \mathbf{Y}$, 及び $\mathbf{A} = (\mathbf{W} \mathbf{V})^{-1}$ とおくと, 式 (2) より次式が成立する.

$$\mathbf{P} \mathbf{F} \mathbf{G} \approx \begin{bmatrix} (\mathbf{W} \mathbf{V})^{-1} \\ \mathbf{0} \end{bmatrix} \mathbf{G} \quad (8)$$

上式より, 独立性基準最大化によって推定される基底行列 \mathbf{F} は次のように得られる.

$$\mathbf{F} \approx \mathbf{P}^{-1} \begin{bmatrix} (\mathbf{W} \mathbf{V})^{-1} \\ \mathbf{0} \end{bmatrix} \quad (9)$$

しかしながら, 次元圧縮を施している関係上, \mathbf{G} は完全に非負行列にはならない. また, \mathbf{F} に至っては非負性が考慮されていない. 従って, 次の3通りの非負化法のいずれかを行う.

非負化処理 1: $\mathbf{F}^{(0)} = \text{abs}(\mathbf{F})$, $\mathbf{G}^{(0)} = \text{abs}(\mathbf{G})$

非負化処理 2: $\mathbf{F}^{(0)} = \text{abs}(\mathbf{F})$, $\mathbf{G}^{(0)} = \alpha_G^{(*)} \mathbf{F}^{(0)T} \mathbf{X}$

非負化処理 3: $\mathbf{G}^{(0)} = \text{abs}(\mathbf{G})$, $\mathbf{F}^{(0)} = \alpha_F^{(*)} \mathbf{X} \mathbf{G}^{(0)T}$

但し, $\text{abs}(\cdot)$ は要素毎の絶対値, $\alpha_F^{(*)}$ 及び $\alpha_G^{(*)}$ はスケールを \mathbf{X} に合わせるための係数であり, $*$ = {EU, KL, IS} である. 本初期化の後に行う NMF のコスト関数 \mathcal{D} を最小とするスケール $\alpha_F^{(*)}$ 及び $\alpha_G^{(*)}$ は次のように与えられる.

$$\alpha_F^{(\text{EU})} = \frac{\sum_{m,n} x_{m,n} b_{m,n}}{\sum_{m,n} b_{m,n}^2}, \quad \alpha_G^{(\text{EU})} = \frac{\sum_{m,n} x_{m,n} c_{m,n}}{\sum_{m,n} c_{m,n}^2} \quad (10)$$

$$\alpha_F^{(\text{KL})} = \frac{\sum_{m,n} x_{m,n}}{\sum_{m,n} b_{m,n}}, \quad \alpha_G^{(\text{KL})} = \frac{\sum_{m,n} x_{m,n}}{\sum_{m,n} c_{m,n}} \quad (11)$$

$$\alpha_F^{(\text{IS})} = \frac{1}{MN} \sum_{m,n} \frac{x_{m,n}}{b_{m,n}}, \quad \alpha_G^{(\text{IS})} = \frac{1}{MN} \sum_{m,n} \frac{x_{m,n}}{c_{m,n}} \quad (12)$$

$$b_{m,n} = [\mathbf{X} \mathbf{G}^{(0)T} \mathbf{G}^{(0)}]_{m,n}, \quad c_{m,n} = [\mathbf{F}^{(0)} \mathbf{F}^{(0)T} \mathbf{X}]_{m,n} \quad (13)$$

3.4 差分信号とICAを用いたNMFの初期値決定法

式(2)における独立成分 S と行列 X が正負の値を持つ場合は、通常のICAを用いて W を推定することができる。従って、本来は非負である独立成分の差分信号を S に与える手法を提案する。すなわち、前節の非負独立成分を推定するNICAとは異なり、 $S=G\Delta$ 及び $X=D\Delta$ として零平均の独立信号を通常のICAで推定する。ここで、 Δ は行方向の1階差分を取る $M \times (N-1)$ 行列である。

本稿では、独立成分の生成モデルとしてラプラス分布を仮定したICAを用いる。ラプラス分布を仮定することで、コントラスト関数が L_2 ノルムとなり、ICAのコスト関数全体が W に対して凸関数になるため一意な解が見つかる利点がある。さらに、補助関数法による高速な最適化手法[25]が提案されている。推定される分離信号 W を観測信号 X ではなく、差分をとる前の D に乗ずることで、係数行列を $G=WD$ として求められる。また、基底行列は式(9)の V を単位行列とした形で同様に求められる。本手法では、補助関数法による更新の1反復毎に W から式(9)を使って F を計算し、 $\text{abs}(F)$ に置き換えて W に戻す処理を加えることで、基底行列 F の非負性を保証する。

ICAを用いる本手法も、前節と同様に、PCAによる次元圧縮に起因して G が完全に非負にはならない。従って、前節と同様に非負化処理1から3までのいずれかを適用して最終的な初期値 $F^{(0)}$ 及び $G^{(0)}$ を得る。

4 評価実験

4.1 初期値としての性能評価

提案法の有用性を検証するため、MIDI音源のGarritan Personal Orchestra 4で生成したピアノ音(F3からE5の24音を重ならないように1音ずつ昇順で鳴らした信号)のパワースペクトログラムを入力データ行列 D とする。 D のサイズは $M=2049$, $N=1047$, 信号のサンプリング周波数は44.1 kHz, STFTの設定はハミング窓、窓長92 ms, 及びシフト長46 msである。従来法及び提案法で初期値を与えた後、乗法更新型のEU-NMF, KL-NMF, 及びIS-NMFをそれぞれ適用する。NICAを用いた提案法をNICA1, NICA2, NICA3と表記する。但し、後ろの数字は非負化手法の番号に対応している。同様に、差分信号とICAを用いた提案法をICA1からICA3と表記する。その他、PCAを用いた初期値[20](PCA)と、NNDSVDを用いた初期値[13](SVD)を比較する。さらに、 $[0, 1]$ の一様分布から異なるシードで生成した一様乱数10種(Rand1からRand10)と比較する。なお、NICAの更新回数は2000回、ICAの更新回数は30回、NMFの基底数は $K=30$ としている。

Figure 2は、提案法におけるコスト関数の収束の様子を示している。NICAは最急降下法であるため、比較的多い反復が要求されるが、ICAは補助関数法による最適化のため反復回数に対する収束が速い。ICAは更新の度に基底行列 F の絶対値処理を行っている為、単調減少性は失われているが、本来のコスト関数は W に対して凸関数であるため、一定の範囲内で収束していることが確認できる。

Figure 3は、NMFのコスト関数の収束の様子を初期値毎に示している。いずれのコスト関数においても、提案法は乱数初期値と比較して高速に降下し、より低い値で収束していることが確認できる。また、SVDよりもPCAの方が良質な初期値を与えており、PCAと提案法はEU-NMFで同程度、KL-NMFとIS-NMFで提案法が勝る結果となっている。NICAとICAの比較や非負化処理に関しては、本実験データではい

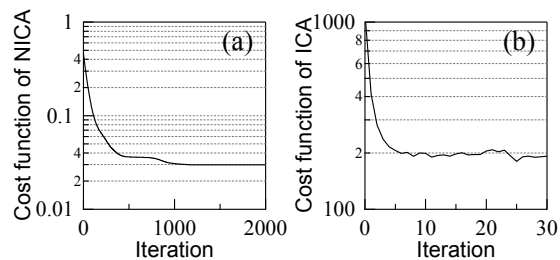


Fig. 2 Convergences of cost function in (a) NICA and (b) ICA.

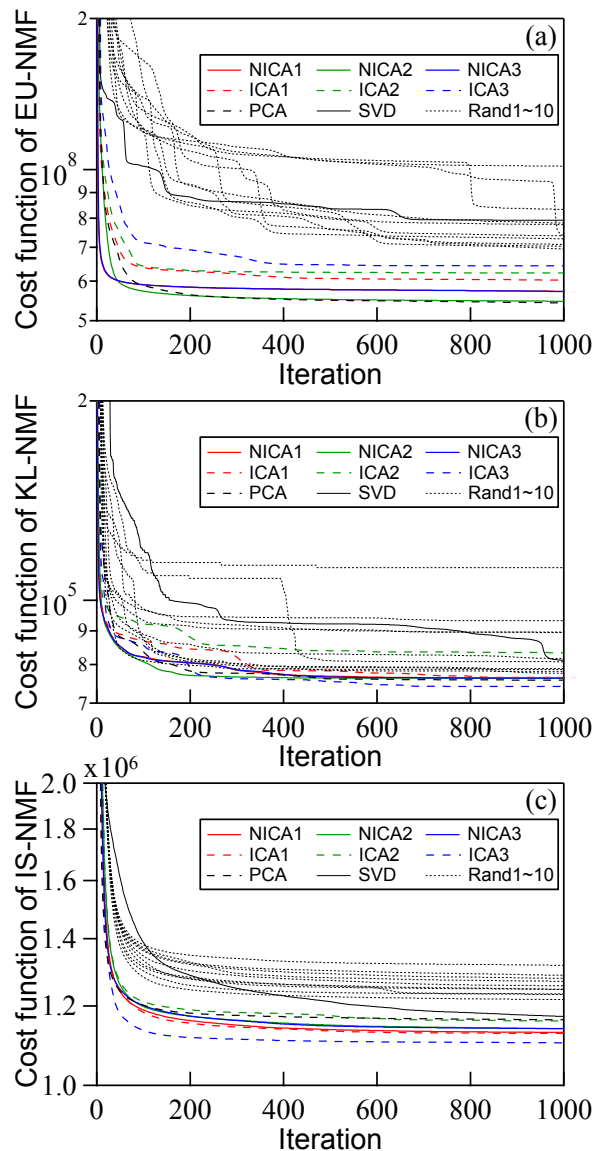


Fig. 3 Convergences of cost function in (a) EU-NMF, (b) KL-NMF, and (c) IS-NMF.

ずれかが一貫して良いという結果にはならなかった。その他、参考としてNICA, ICA, PCA, 及びSVDの初期値決定に要する時間はそれぞれ4.28 s, 10.77 s, 0.95 s, 及び1.31 sであった。NMFの1000回反復に要する時間はEU, KL, 及びISでそれぞれ18.39 s, 62.19 s, 202.96 sであったため、特にIS-NMFにおいては提案法を用いる利点が十分あると考えられる。

4.2 全教師有り音源分離タスクでの性能評価

全教師有り音源分離タスク[6]に対する各初期化手法の分離性能を比較する。評価に用いる信号はSiSEC2015[26]で公開されている音楽データセット

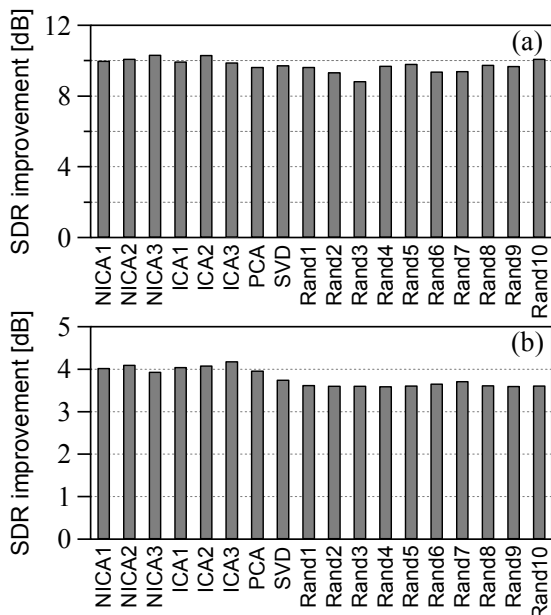


Fig. 4 SDR improvement of supervised NMF for (a) vocals and (b) other.

MSD100 の Test データをアルファベット順に並べた際の最初の 15 曲である。MSD100 は各楽曲に vocals, bass, drums, other の四音源が用意されているが、今回は vocals と other の二音源のソロ信号を各音源用の教師基底の学習に用い、同じ二音源の混合信号を分離テストに用いる。このとき、各楽曲の 40 秒から 1 分 40 秒の区間を 10 秒毎の 6 つの信号に切り分けて 6-fold cross-validation を行う。基底数は音源毎に 50 本とし、学習及び分離時の NMF はいずれも反復回数 200 回の IS-NMF とする。また、分離時の NMF の各音源の係数行列の初期値は、全ての手法において、音源毎に学習した教師基底 F_S を用いて $G^{(0)} = \alpha_G^{(IS)} F_S^T X$ としている。従って、教師基底を学習する際の初期値の違いのみから、その後の教師有り音源分離の精度がどのように変化するかを比較している。分離の評価尺度は、総合分離性能を示す signal-to-distortion ratio (SDR) [27] の改善量とする。その他の条件は前節と同様である。

Figure 4 は、15 曲の SDR 改善量の平均値を音源毎及び初期化手法毎に示している。結果を見ると、提案法が他の手法よりも高い分離精度を達成していることが確認できる。全教師有り NMF において高精度な分離を達成するためには、事前学習した教師基底 F_S が異なる音源の成分を表現しないように、過不足無く学習される必要がある。従って、直交性が最大化された表現力の高い基底よりも凸多面錘の端に沿う基底を得る必要がある。Figure 4 の結果は、提案法のようにより良い初期値を与えることで、そのような過不足の無い基底が得られることを示唆している。

5 まとめ

本稿では、NMF の初期値依存問題の解決を目的として、統計的独立性最大化に基づく初期値決定法を提案した。評価実験の結果、EU-NMF と IS-NMF において、乱数や従来の手法よりも高速に良い解に収束することを示した。また、全教師有り音源分離において、比較的良好な分離結果を与えることを確認した。

謝辞 本研究の一部は JSPS 特別研究員奨励費 26・10796 の助成を受けたものである。

References

- [1] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.
- [2] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Proc. Advances in Neural Information Processing Systems*, vol. 13, pp. 556–562, 2001.
- [3] A. Cichocki, R. Zdunek, A. H. Phan and S. Amari, "Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation," *John Wiley & Sons*, 2009.
- [4] P. Comon, "Independent component analysis, a new concept?," *Signal processing*, vol. 36, no. 3, pp. 287–314, 1994.
- [5] T. Virtanen, "Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria," *IEEE Trans. ASLP*, vol. 15, no. 3, pp. 1066–1074, 2007.
- [6] P. Smaragdakis, B. Raj and M. Shashanka, "Supervised and semi-supervised separation of sounds from single-channel mixtures," *Proc. International Conference on Independent Component Analysis and Signal Separation*, pp. 414–421, 2007.
- [7] H. Kameoka, M. Nakano, K. Ochiai, Y. Imoto, K. Kashino and S. Sagayama, "Constrained and regularized variants of non-negative matrix factorization incorporating music-specific constraints," *Proc. ICASSP*, pp. 5365–5368, 2012.
- [8] H. Sawada, H. Kameoka, S. Araki and N. Ueda, "Multichannel extensions of non-negative matrix factorization with complex-valued data," *IEEE Trans. ASLP*, vol. 21, no. 5, pp. 971–982, 2013.
- [9] D. Kitamura, H. Saruwatari, K. Yagi, K. Shikano, Y. Takahashi and K. Kondo, "Music signal separation based on supervised non-negative matrix factorization with orthogonality and maximum-divergence penalties," *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, vol. E97-A, no. 5, pp. 1113–1118, 2014.
- [10] D. Kitamura, H. Saruwatari, H. Kameoka, Y. Takahashi, K. Kondo and S. Nakamura, "Multichannel signal separation combining directional clustering and nonnegative matrix factorization with spectrogram restoration," *IEEE/ACM Trans. ASLP*, vol. 23, no. 4, pp. 654–669, 2015.
- [11] D. Kitamura, N. Ono, H. Sawada, H. Kameoka and H. Saruwatari, "Efficient multichannel nonnegative matrix factorization exploiting rank-1 spatial model," *Proc. ICASSP*, pp. 276–280, 2015.
- [12] D. Kitamura, N. Ono, H. Sawada, H. Kameoka and H. Saruwatari, "Relaxation of rank-1 spatial constraint in overdetermined blind source separation," *Proc. EUSIPCO*, pp. 1271–1275, 2015.
- [13] C. Boutsidis and E. Gallopoulos, "SVD based initialization: a head start for nonnegative matrix factorization," *Pattern Recognition*, vol. 41, pp. 1350–1362, 2008.
- [14] K. Stadlthanner, F. J. Theis, E. W. Lang, A. M. Tomé, C. G. Puntonet, P. G. Vilda, T. Langmann and G. Schmitz, "Sparse nonnegative matrix factorization applied to microarray data sets," *Proc. Independent Component Analysis and Blind Signal Separation*, pp. 254–261, 2006.
- [15] A. Janecek and Y. Tan, "Using population based algorithms for initializing nonnegative matrix factorization," *Proc. ICSI*, pp. 307–316, 2011.
- [16] Y. Xue, C. S. Tong, Y. Chen and W. S. Chen, "Clustering-based initialization for non-negative matrix factorization," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 525–536, 2008.
- [17] Z. Zheng, J. Yang and Y. Zhu, "Initialization enhancer for non-negative matrix factorization," *Engineering Applications of Artificial Intelligence*, vol. 20, no. 1, pp. 101–110, 2007.
- [18] M. Rezaei, R. Boostani and M. Rezaei, "An efficient initialization method for nonnegative matrix factorization," *Journal of Applied Sciences*, vol. 11, no. 2, pp. 354–359, 2011.
- [19] G. Casalino, N. D. Buono and C. Mencar, "Subtractive clustering for seeding non-negative matrix factorizations," *Information Sciences*, vol. 257, pp. 369–387, 2014.
- [20] L. Zhao, G. Zhuang and X. Xu, "Facial expression recognition based on PCA and NMF," *Proc. WCICA*, pp. 6826–6829, 2008.
- [21] D. Donoho and V. Stodden, "When does non-negative matrix factorization give a correct decomposition into parts?," *Advances in neural information processing systems*, 2003.
- [22] M. D. Plumbley, "Algorithms for nonnegative independent component analysis," *IEEE Trans. Neural Networks*, vol. 14, no. 3, pp. 534–543, 2003.
- [23] E. Oja and M. D. Plumbley "Blind separation of positive sources by globally convergent gradient search," *Neural Computation*, vol. 16, no. 9, pp. 1811–1825, 2004.
- [24] Z. Yuan and E. Oja "A FastICA algorithm for non-negative independent component analysis," *Proc. Independent Component Analysis and Blind Signal Separation*, pp. 1–8, 2004.
- [25] N. Ono and S. Miyabe, "Auxiliary-function-based independent component analysis for super-Gaussian sources," *Proc. LVA/ICA*, pp. 165–172, 2010.
- [26] N. Ono, Z. Rafii, D. Kitamura, N. Ito and A. Liutkus, "The 2015 signal separation evaluation campaign," *Proc. LVA/ICA*, 2015.
- [27] E. Vincent, R. Gribonval and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Trans. ASLP*, vol. 14, no. 4, pp. 1462–1469, 2006.