コサイン類似度罰則条件付き非負値行列因子分解に基づく音源分離の実験的評価 *

☆岩瀬佑太, 北村大地 (香川高専)

1 はじめに

音源分離とは，複数の音源が混合した状態で観測された信号から，混合前の個々の音源を分離・抽出する技術である．特に，非負値行列因子分解（nonnegative matrix factorization: NMF）[1, 2] を用いた音源分離手法は，音源分離の条件や用途等に応じてこれまで数多く提案されてきた [3]–[6]．NMF は Fig. 1 に示すように，観測された非負の時間周波数行列（振幅スペクトログラム等）を別の二つの非負行列（基底行列及びアクティベーション行列）の行列積で近似するアルゴリズムである．基底行列には，観測行列中の頻出スペクトルパターンが基底ベクトルとして含まれ，それらの時間的な強度変化がアクティベーション行列に含まれる．これらの特徴量を上手く活用することで，音源分離ができる．

近年では深層学習による音源分離手法が多くのタスクにおいて高い精度を達成している [7, 8]．しかし，分離対象となる音源（目的音源）の学習用データセットが極僅かしか用意できないような状況においては，省サンプルの学習データから特徴量を学習しそのまま活用する教師あり・半教師あり NMF [9] が依然として強力である．本稿では，半教師あり NMF（semi-supervised NMF: SNMF）に基づく音源分離のみを取り扱う．

Fig. 2 は SNMF における音源分離アルゴリズムを示しており，学習ステージと分離ステージから構成される．学習ステージでは，目的音源のサンプル信号のスペクトログラム $\boldsymbol{X}$ を NMF で分解し，頻出スペクトルパターンを含む教師基底行列 $\boldsymbol{F}$ を学習する．分離ステージでは，事前学習した教師基底行列を用いて混合信号のスペクトログラム $\boldsymbol{Y}$ を分解することで，目的音源成分 $\boldsymbol{FG}$ とその他の音源（非目的音源）成分 $\boldsymbol{HU}$ を推定する．

SNMF に基づく音源分離では，混合信号中の目的音源と非目的音源が類似したスペクトルを含む場合，分離精度が劣化する問題がある．これは，類似するスペクトルが教師基底行列 $\boldsymbol{F}$ 又は非目的音源の基底行列 $\boldsymbol{H}$ のいずれを用いても表現できることが原因である．この場合，目的音源の一部が $\boldsymbol{HU}$ に混入してしまうこととなり，音源分離精度が著しく劣化する．

この問題に対処するために，いくつかの罰則条件付き SNMF（penalized SNMF: PSNMF）が提案されてきた．その中の一つに，基底ベクトル間の内積最小化項を付与した PSNMF（内積型 PSNMF）がある [10]．この手法は，教師基底行列 $\boldsymbol{F}$ と非目的音源の基底行列 $\boldsymbol{H}$ ができるだけ直交化する罰則条件を付与したものであり，目的音源成分 $\boldsymbol{FG}$ の品質を大幅に向上させる効果がある．しかし，内積型 PSNMF は，NMF が持つ基底行列・アクティベーション行列間のスケール不定性により，罰則条件が正しく影響しない定式化となっている．同様の現象は，アクティベーション行列にスパース正則化を与える NMF でも確認されており，正しい制約条件を与えた最適化法が提案されている [11]．

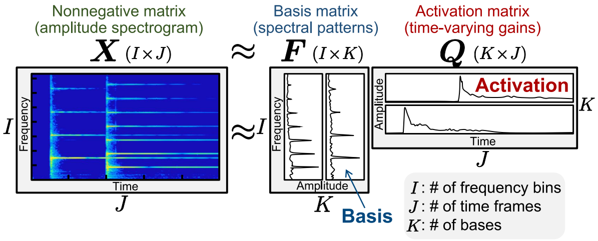

Fig. 1 Matrix decomposition by NMF.

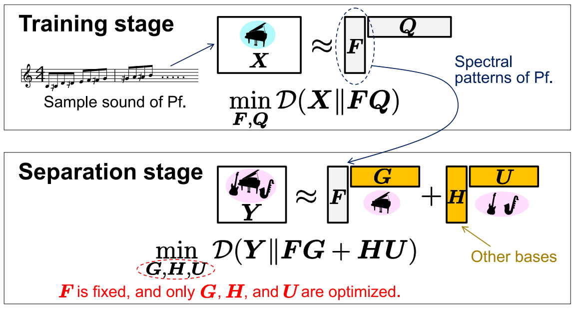

Fig. 2 Audio source separation based on SNMF.

内積型 PSNMF のスケール不定性の問題を解決するために，著者らは基底ベクトル間の対数コサイン類似度最小化項を付与した PSNMF（log-cos 型 PSNMF）を提案している [12]．Log-cos 型 PSNMF は基底ベクトル間の大きさに依存せず，ベクトル間の角度にのみ依存するコサイン類似度を罰則項に用いるため，内積型 PSNMF のスケール不定性の問題を解決できる．しかし，反復最適化更新式の導出を容易にするためにコサイン類似度の対数を罰則項として用いた結果，コスト関数や変数の値が発散する可能性が生じ，数値的安定性が損なわれる問題が生じた．文献 [12] では，数値的安定性を向上させるために，反復毎に $\boldsymbol{H}$ 中の各基底ベクトルを（コスト関数の値が変化しないように）正規化することで，場当たり的に対処していたが，最適化途中での変数値の発散を完全に防ぐことはできていない．

本稿では，スケール不定性及び数値的不安定性の両方の問題を同時に解決する手法として，対数を用いないシンプルなコサイン類似度を罰則条件として SNMF に付与する手法（cos 型 PSNMF）を新たに定式化し，最適化する手法を提案する．提案手法では，majorization-minimization（MM）アルゴリズム [13] による乗法型更新則を導出する．また，音楽信号を対象とした音源分離に適用し，従来手法よりも優れた音源分離が達成できることを実験的に示す．

* Experimental evaluation of source separation based on nonnegative matrix factorization with cosine similarity penalty. By Yuta Iwase and Daichi Kitamura (NIT Kagawa).

2　従来手法

2.1　NMF 及び SNMF

NMF は非負の観測行列 $\boldsymbol{X} \in \mathbb{R}_{\geq 0}^{I\times J'}$（本稿では音の振幅スペクトログラム）を別の 2 つの非負行列 $\boldsymbol{F} \in \mathbb{R}_{\geq 0}^{I\times K}$（基底行列）及び $\boldsymbol{Q} \in \mathbb{R}_{\geq 0}^{K\times J'}$（アクティベーション行列）の行列積で近似する低ランク分解である．$\boldsymbol{F}$ 及び $\boldsymbol{Q}$ は次式の最小化問題で推定される．

$$\min_{\boldsymbol{F},\boldsymbol{Q}} \mathcal{D}(\boldsymbol{X}\|\boldsymbol{F}\boldsymbol{Q}) \text{ s.t. } f_{ik}, q_{kj'} \geq 0\ \forall i,\ j',\ k \tag{1}$$

ここで f_{ik} 及び $q_{kj'}$ はそれぞれ $\boldsymbol{F}$ 及び $\boldsymbol{Q}$ の要素である．また，$i = 1, 2, \cdots, I$，$j' = 1, 2, \cdots, J'$，及び $k = 1, 2, \cdots, K$ はそれぞれ周波数，時間，及び基底ベクトルのインデクスを表す．$\mathcal{D}(\boldsymbol{X}\|\boldsymbol{F}\boldsymbol{Q})$ は，行列 $\boldsymbol{X}$ 及び $\boldsymbol{F}\boldsymbol{Q}$ 間の乖離度を表す関数である．

NMF に基づく音源分離手法として SNMF がある．SNMF では，Fig. 2 に示すように，学習ステージで目的音源のサンプル信号 $\boldsymbol{X}$ に NMF を適用し，教師基底行列 $\boldsymbol{F}$ を事前学習する．$\boldsymbol{F}$ は目的音源の頻出スペクトルを K 本の基底ベクトルとして含む．分離ステージでは，混合信号 $\boldsymbol{Y} \in \mathbb{R}_{\geq 0}^{I\times J}$ を $\boldsymbol{F}\boldsymbol{G} + \boldsymbol{H}\boldsymbol{U}$ に分解する．ここで，$\boldsymbol{G} \in \mathbb{R}_{\geq 0}^{K\times J}$ は $\boldsymbol{F}$ のアクティベーション行列，$\boldsymbol{H} \in \mathbb{R}_{\geq 0}^{I\times L}$ 及び $\boldsymbol{U} \in \mathbb{R}_{\geq 0}^{L\times J}$ はそれぞれ非目的音源の成分を表す基底行列及びアクティベーション行列である．理想的には，混合信号 $\boldsymbol{Y}$ 中の目的音源成分が $\boldsymbol{F}\boldsymbol{G}$，非目的音源成分が $\boldsymbol{H}\boldsymbol{U}$ と推定され，$\boldsymbol{F}\boldsymbol{G}$ のみを時間信号に戻すことで音源分離が達成される．しかし，目的音源と非目的音源が類似スペクトルを含む場合，目的音源の一部が $\boldsymbol{H}\boldsymbol{U}$ に混入し，分離精度が著しく劣化する．

2.2　内積型 PSNMF

内積型 PSNMF [10] は，教師基底行列 $\boldsymbol{F}$ と非目的音源の基底行列 $\boldsymbol{H}$ ができるだけ直交する罰則条件を付与した SNMF であり，目的音源成分（$\boldsymbol{F}\boldsymbol{G}$ で表現されるべき成分）が非目的音源成分 $\boldsymbol{H}\boldsymbol{U}$ に混入してしまう現象をある程度回避する効果がある．本手法では，分離ステージにおいて次の最小化問題を考える．

$$\min_{\boldsymbol{G},\boldsymbol{H},\boldsymbol{U}} \mathcal{D}(\boldsymbol{Y}\|\boldsymbol{F}\boldsymbol{G} + \boldsymbol{H}\boldsymbol{U}) + \mu\mathcal{P}_{\mathrm{orth}}(\boldsymbol{F},\boldsymbol{H})$$
$$\text{s.t. } g_{kj}, h_{il}, u_{lj} \geq\ 0\ \forall i,\ j,\ k,\ l \tag{2}$$

$$\mathcal{P}_{\mathrm{orth}}(\boldsymbol{F},\boldsymbol{H}) = \|\boldsymbol{F}^{\mathrm{T}}\boldsymbol{H}\|_{\mathrm{Fr}}^2 = \sum_{k,l}\left(\sum_i f_{ik}h_{il}\right)^2 \tag{3}$$

ここで，μ は罰則項の重み係数，g_{kj}，h_{il}，及び u_{lj} はそれぞれ $\boldsymbol{G}$，$\boldsymbol{H}$，及び $\boldsymbol{U}$ の要素，$j = 1, 2, \cdots, J$ 及び $l = 1, 2, \cdots, L$ はそれぞれ時間及び $\boldsymbol{H}$ 中の基底ベクトルのインデクス，$\|\cdot\|_{\mathrm{Fr}}$ はフロベニウスノルム，h_{il} 及び u_{lj} はそれぞれ $\boldsymbol{H}$ 及び $\boldsymbol{U}$ の要素である．式 (3) で表される罰則項 $\mathcal{P}_{\mathrm{orth}}(\boldsymbol{F},\boldsymbol{H})$ は，$\boldsymbol{F}$ の基底ベクトル（K 本）と $\boldsymbol{H}$ の基底ベクトル（L 本）の総組み合わせの内積の二乗和に対応する．従って，$\boldsymbol{H}$ を推定する際には，「$\boldsymbol{Y}$ と $\boldsymbol{F}\boldsymbol{G}+\boldsymbol{H}\boldsymbol{U}$ が近くなること」に加えて，「$\boldsymbol{H}$ の基底ベクトルが $\boldsymbol{F}$ の基底ベクトルとなるべく直交すること」の 2 点が考慮され，結果として単なる SNMF よりも音源分離が促進されることが実験的に確認されている．

2.3　Log-cos 型 PSNMF

内積型 PSNMF には罰則項の効果が直接反映されない本質的な問題がある．これは，NMF が本来持っている基底行列とアクティベーション行列間のスケール不定性が原因である．すなわち，内積型 PSNMF の罰則項（式 (3)）は，$\boldsymbol{H}$ に小さな値 c を乗じて $c\boldsymbol{H}$ とするだけで最小化できる．このとき，式 (2) の第一項は $\boldsymbol{U}$ を $(1/c)\boldsymbol{U}$ とすることで増加しない．従って，内積型 PSNMF は μ を大きくしても，理論的には $\boldsymbol{F}$ と $\boldsymbol{H}$ は直交化されない．そこで著者らは，$\boldsymbol{H}$ の大きさに非依存なコサイン類似度に基づく罰則項を用いて，log-cos 型 PSNMF を提案した [12]．本手法では，分離ステージにおいて次の最小化問題を考える．

$$\min_{\boldsymbol{G},\boldsymbol{H},\boldsymbol{U}} \mathcal{D}(\boldsymbol{Y}\|\boldsymbol{F}\boldsymbol{G} + \boldsymbol{H}\boldsymbol{U}) + \mu\mathcal{P}_{\mathrm{logcos}}(\boldsymbol{F},\boldsymbol{H})$$
$$\text{s.t. } g_{kj}, h_{il}, u_{lj}\ \geq\ 0\ \forall i,\ j,\ k,\ l \tag{4}$$

$$\mathcal{P}_{\mathrm{logcos}}(\boldsymbol{F},\boldsymbol{H}) = \sum_{k,l}\log\frac{\sum_i f_{ik}h_{il}}{\left(\sum_i f_{ik}^2\right)^{\frac{1}{2}}\left(\sum_i h_{il}^2\right)^{\frac{1}{2}}} \tag{5}$$

ここで，罰則項 $\mathcal{P}_{\mathrm{logcos}}(\boldsymbol{F},\boldsymbol{H})$ は，$\boldsymbol{F}$ の基底ベクトルと $\boldsymbol{H}$ の基底ベクトルの総組み合わせのコサイン類似度の対数和に対応する．コサイン類似度であるので，基底ベクトルの長さには依らず，ベクトル間の角度のみで直交性を測ることができ，NMF におけるスケールの不定性は最適化に影響しない．

3　提案手法

3.1　動機

Log-cos 型 PSNMF の最小化問題（式（4））では，最適化のための反復更新式の導出を簡単化するためにコサイン類似度の対数を用いている．しかし，教師基底行列 $\boldsymbol{F}$ と非目的音源の基底行列 $\boldsymbol{H}$ の基底ベクトルが 1 組でも完全に直交する場合，そのコサイン類似度は 0 になり得ることから，罰則項 $\mathcal{P}_{\mathrm{logcos}}(\boldsymbol{F},\boldsymbol{H})$ は $-\infty$ となる．従って，log-cos 型 PSNMF は新たに数値的不安定性の問題を孕んでいる．そこで本稿では，対数をとらない純粋なコサイン類似度を罰則項とした場合の反復更新式を導出し，これを新たに cos 型 PSNMF として提案することで，前述の問題の解決を図る．

3.2　Cos 型 PSNMF

一般に SNMF を用いた音源分離では，振幅スペクトルを一般化 Kullback-Leibler（KL）ダイバージェンスに基づく NMF で分解する手法が最も高性能であることが実験的に確認されている [10, 14]．従って本稿でも，乖離度 $\mathcal{D}(\cdot\|\cdot)$ に一般化 KL ダイバージェンス $\mathcal{D}_{\mathrm{KL}}(\cdot\|\cdot)$ を用いる場合のみを取り扱う．ここで，

$$\mathcal{D}_{\mathrm{KL}}(\boldsymbol{M}\|\boldsymbol{N}) = \sum_{i,j}\left(m_{ij}\log\frac{m_{ij}}{n_{ij}} - m_{ij} + n_{ij}\right) \tag{6}$$

であり，m_{ij} 及び n_{ij} はそれぞれ $\boldsymbol{M} \in \mathbb{R}_{\geq 0}^{I\times J}$ 及び $\boldsymbol{N} \in \mathbb{R}_{\geq 0}^{I\times J}$ の要素である．

Cos 型 PSNMF では，次の最小化問題を考える．

$$\min_{\boldsymbol{G},\boldsymbol{H},\boldsymbol{U}} \mathcal{D}(\boldsymbol{Y}\|\boldsymbol{FG}+\boldsymbol{HU})+\mu\mathcal{P}_{\cos}(\boldsymbol{F},\boldsymbol{H})$$
$$\text{s.t. } g_{kj}, h_{il}, u_{lj} \ \geq \ 0 \ \forall i,\ j,\ k,\ l \tag{7}$$

$$\mathcal{P}_{\cos}(\boldsymbol{F},\boldsymbol{H})=\sum_{k,l}\frac{\sum_i f_{ik}h_{il}}{\left(\sum_i f_{ik}^2\right)^{\frac{1}{2}}\left(\sum_i h_{il}^2\right)^{\frac{1}{2}}} \tag{8}$$

式 (8) は，$\boldsymbol{F}$ 中の基底ベクトルと $\boldsymbol{H}$ 中の基底ベクトルの総組み合わせのコサイン類似度の総和である．Log-cos 型 PSNMF の罰則項（式 (5)）と比較して，対数関数を省くことで，最適化アルゴリズムは複雑になるが，数値不安定性の問題を解消できる．

式 (7) の最適化問題のコスト関数は，MM アルゴリズム [13] を用いて最小化できる．まず，一般化 KL ダイバージェンス中の負対数項は凸なので，次式の Jensen の不等式を適用する．

$$\begin{aligned}&-\log\left(\sum_k f_{ik}g_{kj}+\sum_l h_{il}u_{lj}\right)\\&\quad\leq-\sum_k\alpha_{ijk}\log\frac{f_{ik}g_{kj}}{\alpha_{ijk}}-\sum_l\beta_{ijl}\log\frac{h_{il}u_{lj}}{\beta_{ijl}}\end{aligned} \tag{9}$$

ここで，$\alpha_{ijk}>0$ 及び $\beta_{ijl}>0$ はそれぞれ $\sum_k\alpha_{ijk}=1$ 及び $\sum_l\beta_{ijl}=1$ を満たす補助変数である．次に，罰則項 $\mathcal{P}_{\cos}(\boldsymbol{F},\boldsymbol{H})$ 中の $-1/2$ 次項は凸関数なので，次式の Jensen の不等式を適用する．

$$\begin{aligned}\left(\sum_i h_{il}^2\right)^{-\frac{1}{2}}&\leq\sum_i\gamma_{il}\left(\frac{h_{il}^2}{\gamma_{il}}\right)^{-\frac{1}{2}}\\&=\sum_i\gamma_{il}^{\frac{3}{2}}h_{il}^{-1}\end{aligned} \tag{10}$$

ここで，$\gamma_{il}>0$ は $\sum_l\gamma_{il}=1$ を満たす補助変数である．式 (9) 及び (10) の等号成立条件はそれぞれ次式である．

$$\alpha_{ijk}=\frac{f_{ik}g_{kj}}{\sum_{k'}f_{ik'}g_{k'j}+\sum_{l'}h_{il'}u_{l'j}} \tag{11}$$

$$\beta_{ijl}=\frac{h_{il}u_{lj}}{\sum_{k'}f_{ik'}g_{k'j}+\sum_{l'}h_{il'}u_{l'j}} \tag{12}$$

$$\gamma_{il}=\frac{h_{il}^2}{\sum_{i'}h_{i'l}^2} \tag{13}$$

式 (9) 及び (10) より，式 (7) の目的関数の上限関数 $\mathcal{J}^+$ は次式となる．

$$\begin{aligned}\mathcal{J}^+=&\sum_{i,j}\Bigg[y_{ij}\log y_{ij}-y_{ij}\sum_k\alpha_{ijk}\log\frac{f_{ik}g_{kj}}{\alpha_{ijk}}\\&-y_{ij}\sum_l\beta_{ijl}\log\frac{h_{il}u_{lj}}{\beta_{ijl}}-y_{ij}\\&+\sum_k f_{ik}g_{kj}+\sum_l h_{il}u_{lj}\Bigg]\\&+\mu\sum_{k,l}\left[\left(\sum_i f_{ik}^2\right)^{-\frac{1}{2}}\left(\sum_{i,i'}f_{ik}\gamma_{i'l}^{\frac{3}{2}}\frac{h_{il}}{h_{i'l}}\right)\right]\end{aligned} \tag{14}$$

ここで，y_{ij} は $\boldsymbol{Y}$ の要素である．式 (14) を各変数で偏微分し 0 とおくことで反復最適化更新式を導出できる．$\boldsymbol{G}$ 及び $\boldsymbol{U}$ の更新式は SNMF と同様であるため割愛する．$\partial\mathcal{J}^+/\partial h_{il}=0$ より

$$\begin{aligned}&\sum_j\left[\frac{-y_{ij}\beta_{ijl}}{h_{il}}+u_{lj}\right]+\mu\sum_k\Bigg[\left(\sum_i f_{ik}^2\right)^{-\frac{1}{2}}\\&\left(-\frac{1}{h_{il}^2}\sum_{i'\neq i}f_{i'k}\gamma_{il}^{\frac{3}{2}}h_{i'l}+\sum_{i'\neq i}f_{ik}\gamma_{i'l}^{\frac{3}{2}}\frac{1}{h_{i'l}}\right)\Bigg]=0\end{aligned}$$

を得る．上式は h_{il} に関して整理することで 2 次方程式となり，次式で解析的に解ける．

$$h_{il}=\frac{-b_{il}\pm\sqrt{b_{il}^2-4a_{il}c_{il}}}{2a_{il}} \tag{15}$$

ここで，a_{il}，b_{il}，c_{il} はそれぞれ以下の通りである．

$$\begin{aligned}a_{il}&=\sum_i u_{lj}+\mu\sum_k f_{ik}\left(\frac{\sum_{i'\neq i}\gamma_{i'l}^{\frac{3}{2}}\frac{1}{h_{i'l}}}{\sqrt{\sum_i f_{ik}^2}}\right)\\b_{il}&=-\sum_j y_{ij}\beta_{ijl}\\c_{il}&=-\mu\gamma_{il}^{\frac{3}{2}}\sum_k\left[\left(\sum_i f_{ik}^2\right)^{-\frac{1}{2}}\left(\sum_{i'\neq i}f_{i'k}h_{i'l}\right)\right]\end{aligned}$$

ここで，式 (15) の $\pm$ は $h_{il}\geq 0$ となる $+$ が正しい解である．式 (15) に，式 (11)–(13) の等号成立条件を代入したものが $\boldsymbol{H}$ の反復更新式となる．

3.3　変数更新毎の基底正規化

スケール不定性や数値的不安定性の問題を避けるために，文献 [10, 12] では，$\boldsymbol{H}$ を更新する度に全基底ベクトルを正規化する（正規化係数の逆数を $\boldsymbol{U}$ の行ベクトルに乗じて $\boldsymbol{HU}$ は不変とする）ように実装しているが，根本的な解決とはいえない．しかし，実験的には性能改善が見られたため本論文では前述の基底正規化（basis normalization: BN）をした場合としない場合の両方について評価を行う．$\boldsymbol{H}$ 及び $\boldsymbol{U}$ の基底ベクトル h_{il} 及び u_{lj} の BN の式を以下に示す．

$$\delta_l=\sum_i h_{il} \tag{16}$$

$$h_{il}\leftarrow\frac{h_{il}}{\delta_l}\ \forall i,\ l,\quad u_{lj}\leftarrow\delta_l u_{lj}\ \forall l,\ j \tag{17}$$

4　実験

4.1　実験条件

2 つの楽器音の音源分離性能を SNMF，内積型 PSNMF [10]，log-cos 型 PSNMF [12]，及び提案手法である cos 型 PSNMF の 4 手法で比較した．但し，各手法においてスケール不定性と数値的不安定性を解消するために，変数の更新毎に BN を行う場合と行わない場合の 2 通りを比較した．目的音源と非目的音源の基底ベクトル数はそれぞれ $K=27$ 及び $L=50$ とした．2 つの楽器音の混合信号は songKitamura [15] のデータセット用いて 90 曲作成し，45 曲ずつに分け

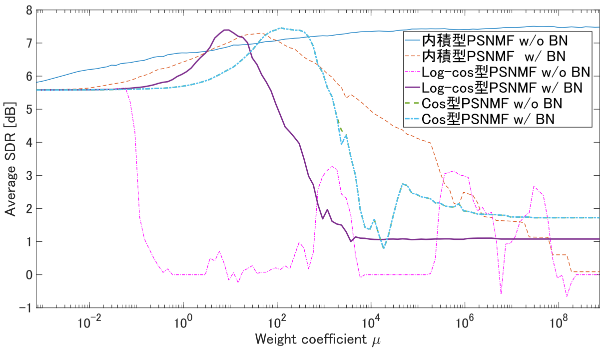

Fig. 3 Average SDR for searching optimal μ.

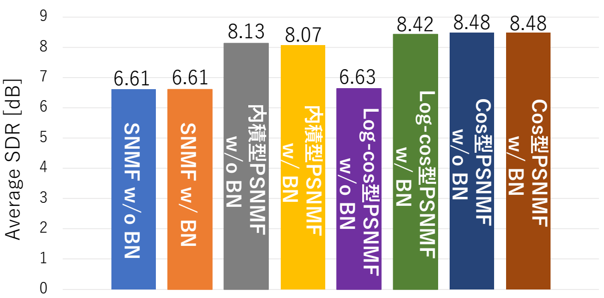

Fig. 4 Average SDR of test data for each method.

て開発データ及びテストデータとした．開発データで重み係数 μ の最適値を手法毎に決定し，最適な μ で固定してテストデータで性能を比較した．評価指標には，信号対歪み比（source-to-distortion ratio: SDR）[16] を用いた．なお，log-cos 型 PSNMF では，前述したとおり BN を行わなかった場合，数値的不安定性により変数値が発散することがある．この時，改善量をグラフに数値的にプロットするため，改善量がみられなかったことと同義である 0 dB とした．

4.2 実験結果

Fig. 3 に開発データにおける各手法の平均 SDR を示す．このグラフより，SDR が最大となる重み係数 μ を各手法で求めた．Cos 型 PSNMF では BN の有無が最適化に与える影響は（理論的には）無いため，Fig. 3 では 2 つのグラフがほぼ重なっている．手法毎の最適な μ でテストデータを音源分離した際の平均 SDR を Fig. 4 に示す．まず，内積型 PSNMF は，理論的には $\boldsymbol{F}$ と $\boldsymbol{H}$ の直交化がされない可能性があるが，それでも μ を大きくすことで罰則無しの SNMF よりも性能は向上している．また，BN 無しの log-cos 型 PSNMF は数値不安定性により一部の混合信号で変数値が発散する場合があり，平均性能は低くなった．BN 有りの log-cos 型 PSNMF と cos 型 PSNMF は他の手法よりも精度が高くなり，その有効性が示された．

5 まとめ

本稿では，内積型 PSNMF 及び log-cos 型 PSNMF が抱えるスケール不定性及び数値不安定性の 2 つの問題を避けるために cos 型 PSNMF を新たに提案し，反復更新式を導出した．提案手法では，前述の問題を解決できるとともに，高精度な音源分離が可能であることを評価実験で示した．

参考文献

[1] 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.

[2] D. D. Lee, and H. S. Seung, "Algorithms for non-negative matrix factorization," *Proc. Neural Information Processing Systems*, pp. 556–562, 2000.

[3] T. Virtanen, "Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 15, no. 3, pp. 1066–1074, 2007.

[4] A. Ozerov, C. Févotte, and M. Charbit, "Factorial scaled hidden Markov model for polyphonic audio representation and source separation," *Proc. Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 121–124, 2009.

[5] 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. International Conference on Acoustics, Speech and Signal Processing*, pp. 5365–5368. 2012.

[6] 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. Audio, Speech, and Language Processing*, vol. 23, no. 4, pp. 654–669, 2015.

[7] E. M. Grais, M. U. Sen, and H. Erdogan, "Deep neural networks for signal channel source separation," *Proc. International Conference on Independent Component Analysis and Signal Separation*, pp. 3734–3738. 2014.

[8] S. Uhilch, F. Giron, and Y. Mitsufuji, "Deep neural network based instrument extraction from music," *Proc. International Conference on Independent Component Analysis and Signal Separation*, pp. 2135–2139. 2015.

[9] P. Smaragdis, B. Raj, and M. Shashanka, "Supervised and semi-supervised separation of sounds from single-channel mixtures," *Proc. Independent Component Analysis and Signal Separation*, pp. 414–421, 2007.

[10] D. Kitamura, H. Saruwatari, K. Yagi, K. Shikano, Y. Takahashi, and K. Kondo, "Music signal separation based on supervised nonnegative 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.

[11] J. Le Roux, F. Weninger, and J. R. Hershey, "Sparse NMF - half-baked or well done?," *Mitsubishi Electric Research Lab. Technical Report*, TR2015-023, 2015.

[12] 岩瀬佑太, 北村大地, "コサイン類似度罰則条件付き半教師あり非負値行列因子分解," 日本音響学会 2020 年春季研究発表会講演論文集, 2-P-39, pp. 425–428, March 2020.

[13] D. R. Hunter and K. Lange, "Quantile regression via an MM algorithm," *Journal of computational and Graphical Statistics.*, vol. 9, no. 1, pp. 60–77, 2000.

[14] D. FitzGerald, M. Cranitch, and E. Coyle, "On the use of the beta divergence for musical source separation," *Proc. Irish Signals and Systems Conference*, 2009.

[15] D. Kitamura, "Open dataset: songKitamura" http://d-kitamura.net/dataset.html.

[16] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 14, no. 4, pp. 1462—1469, 2006.