補助関数法による識別的 NMF の基底学習アルゴリズム *

◎李 莉 [1], 亀岡 弘和 [2], 牧野 昭二 [1]
[1] 筑波大学 [2] 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所

1 はじめに

近年，非負値行列因子分解（Non-negative matrix factorization: NMF）はモノラル音響信号処理問題に対する有力な手法として注目されている [1]。各時刻で観測された振幅またはパワースペクトルを基底スペクトルの非負結合で近似することは，観測スペクトログラムを行列と見なし，二つの行列（基底行列とアクティベーション行列）の積で近似することに相当する。各々の行列の要素は非負値のため，非負制約のもと観測スペクトログラムに対し行列分解が行われることから NMF と呼ぶ。教師ありまたは半教師あり音源分離の問題設定においては，まず，各音源の学習サンプルのスペクトログラムに NMF を行い，基底行列を事前学習する。一方テスト時には，学習した基底行列を固定し，アクティベーション行列のみを推定する。このようにして求めた各音源のパワースペクトログラムを用い，Wiener フィルタにより混合信号から目的音源信号を得ることができる。

以上のアプローチ [1] では基底学習において学習サンプルのスペクトログラムと行列積との誤差が最適化規準として用いられるが，分離信号そのものが最適となるような規準とはなっていなかった。この点に着目し，Wiener フィルタの出力信号と目的音源の学習サンプルとの誤差を直接的に最適化規準として基底学習を行う，識別的 NMF（Discriminative non-negative matrix factorization: DNMF）[2] と呼ぶ枠組が提案されている。この方式では，学習時とテスト時に用いられる最適化規準が同一となるため，より高い分離能力をもった基底スペクトルが学習により得られるようになることが期待される。しかし，識別的 NMF の学習規準（後述）は従来の NMF の最適化規準に比べて解析的に複雑な形になる。このため [2] では乗法更新アルゴリズムと呼ぶ汎用的な手法を用いた最適化アルゴリズムが提案されているが，停留点への収束性が保証されておらず DNMF のポテンシャルを十分発揮できているとはいえなかった。そこで，本稿では，以上の問題を解決するため，補助関数法という呼ぶ原理に基づく，収束性が保証された識別的 NMF のアルゴリズムを提案する。

2 従来手法

2.1 教師あり NMF による音源分離

L 個の音源からなる混合信号のパワースペクトログラムを $\mathbf{Y} = (Y_{\omega,t})_{\Omega\times T} \in \mathbb{R}^{\geq 0,\Omega\times T}$ とする。ただし，ω と t は周波数および時刻のインデックスである。教師あり NMF では，事前学習した各音源の基底スペクトル $\mathbf{W}^l = (W^l_{\omega,k})_{\Omega\times K^l} \in \mathbb{R}^{\geq 0,\Omega\times K^l}$ を用いて，観測スペクトログラム $\mathbf{Y}$ を基底行列 $\mathbf{W} = [\mathbf{W}^1, \mathbf{W}^2, \ldots, \mathbf{W}^L]$ とアクティベーション行列 $\mathbf{H} = [\mathbf{H}^1; \mathbf{H}^2; \ldots; \mathbf{H}^L]$ の積で近似することで，Wiener フィルタにより各音源信号を混合信号から抽出するためのパワースペクトログラム推定値を得ることが目的である。[1] では $\mathbf{W}^l = (W^l_{\omega,k})_{\Omega\times K^l} \in \mathbb{R}^{\geq 0,\Omega\times K^l}$ の事前学習において，音源 l の学習サンプルのスペクトログラム $\mathbf{S}^l = (S^l_{\omega,t})_{\Omega\times T}$ との誤差

$$\{\hat{\mathbf{W}}^l, \tilde{\mathbf{H}}^l\} = \underset{\mathbf{W}^l,\mathbf{H}^l}{\mathrm{argmax}}\ \mathcal{D}(\mathbf{S}^l|\mathbf{W}^l\mathbf{H}^l) \tag{1}$$

を最適化規準として用いている。ただし，$\mathcal{D}$ は音源 l の学習サンプルのスペクトログラム $\mathbf{S}^l$ と行列積 $\mathbf{W}^l\mathbf{H}^l$ の誤差を測る関数である。テスト時においては，事前学習した基底行列 $\hat{\mathbf{W}} = [\hat{\mathbf{W}}^1, \hat{\mathbf{W}}^2, \ldots, \hat{\mathbf{W}}^L]$ を固定し，

$$\hat{\mathbf{H}} = \underset{\mathbf{H}}{\mathrm{argmax}}\ \mathcal{D}(\mathbf{Y}|\hat{\mathbf{W}}\mathbf{H}) \tag{2}$$

を最小にするアクティベーション行列 $\hat{\mathbf{H}}$ を推定することで，$\mathbf{Y}$ に含まれる各音源のパワースペクトログラムの成分 $\hat{\mathbf{W}}^l\hat{\mathbf{H}}^l$ を推定することができる。誤差関数 $\mathcal{D}$ として I ダイバージェンスを用いる場合，(2) は具体的に

$$\mathcal{D}_{\mathrm{KL}}(\mathbf{Y}|\mathbf{W}\mathbf{H}) = \sum_{\omega,t}\left(Y_{\omega,t}\log\frac{Y_{\omega,t}}{[\mathbf{W}\mathbf{H}]_{\omega,t}} - Y_{\omega,t} + [\mathbf{W}\mathbf{H}]_{\omega,t}\right) \tag{3}$$

となる。ただし，$[\cdot]_{i,j}$ は行列の $\{i,j\}$ 番目要素を表す。各音源のパワースペクトログラム $\hat{\mathbf{W}}^l\hat{\mathbf{H}}^l$ が求まれば，Wiener フィルタ

$$\hat{\mathbf{S}}^l = \frac{\hat{\mathbf{W}}^l\hat{\mathbf{H}}^l}{\hat{\mathbf{W}}\hat{\mathbf{H}}} \otimes \mathbf{Y} \tag{4}$$

*Auxiliary function approach to discriminative non-negative matrix factorization, Li Li (University of Tsukuba), Hirokazu Kameoka (NTT Communication Science Laboratories), Shoji Makino (University of Tsukuba)

により，足して矛盾なく $\mathbf{Y}$ になるよう保証された各音源信号のスペクトログラムを得ることができる。ただし，$\otimes$ と $\div$ は要素ごとの乗法と除法を表すものとする。しかし，上述のアプローチ [1] では，基底の学習規準において (1) が用いられていることから，(4) による分離信号が最適となるような規準になっていなかった。

2.2 識別的 NMF と乗法更新アルゴリズム

識別的 NMF は (1) の代わりに Wiener フィルタ出力と学習サンプルのスペクトログラムの誤差

$$\mathcal{J} = \sum_l \alpha_l \mathcal{D}_{\mathrm{KL}}(\mathbf{S}^l | \hat{\mathbf{S}}^l) \tag{5}$$

を規準として基底学習を行う教師あり NMF による音源分離の枠組である。ただし，$\alpha_l \geq 0$ は l 番目の分離信号の重要度を表すパラメータである。

以下では説明の簡略化のため，音声と雑音の二種類の音源 ($L = 2$) からなる音源分離問題を考える。音声強調が目的の場合は音声信号の分離精度がより重要となるので，重要度 α は，音声に対して 1，雑音に対して 0 とする。従って，クリーン音声の学習サンプルのスペクトログラムを $\mathbf{S}^{\mathrm{s}} = (S^{\mathrm{s}}_{\omega,t})_{\Omega\times T} \in \mathbb{R}^{\geq 0,\Omega\times T}$，雑音の学習サンプルのスペクトログラムを $\mathbf{S}^{\mathrm{n}} = (S^{\mathrm{n}}_{\omega,t})_{\Omega\times T} \in \mathbb{R}^{\geq 0,\Omega\times T}$ とし，その混合信号のスペクトログラムを $\mathbf{M} = (M_{\omega,t})_{\Omega\times T} \in \mathbb{R}^{\geq 0,\Omega\times T}$ とすると，識別的 NMF の基底学習問題は

$$\begin{aligned}&\text{minimize } f(\mathbf{W},\mathbf{H}) = \mathcal{D}_{\mathrm{KL}}\left(\mathbf{S}^{\mathrm{s}} \,\middle|\, \frac{\mathbf{W}^{\mathrm{s}}\mathbf{H}^{\mathrm{s}}}{\mathbf{W}\mathbf{H}} \otimes \mathbf{M}\right) \\ &\text{subject to } \ \forall k,\ \sum_\omega W_{\omega,k} = 1\end{aligned} \tag{6}$$

のような最適化問題として定式化される。ただし，基底行列 $\mathbf{W} = [\mathbf{W}^{\mathrm{s}}, \mathbf{W}^{\mathrm{n}}] \in \mathbb{R}^{\geq 0,\Omega\times K}$ は K^{s} 個の音声基底スペクトルと K^{n} 個の雑音基底スペクトルで構成される。

Weninger らは上述の最適化問題に対し乗法更新法を用いた最適化アルゴリズムを提案している [2]。Weninger らのアルゴリズムでは，まず通常の NMF（すなわち (2)）でアクティベーション行列 $\hat{\mathbf{H}}$ を求め，$\hat{\mathbf{H}}$ を固定した下で基底行列 $\mathbf{W}$ を

$$\mathbf{W}^{\mathrm{s}} \leftarrow \mathbf{W}^{\mathrm{s}} \otimes \frac{\frac{\mathbf{S}^{\mathrm{s}}\otimes\mathbf{W}^{\mathrm{s}}\mathbf{H}^{\mathrm{s}}}{\mathbf{W}\mathbf{H}\otimes\mathbf{W}^{\mathrm{s}}\mathbf{H}^{\mathrm{s}}}\mathbf{H}^{\mathrm{s}\top}}{\frac{\mathbf{M}\otimes\mathbf{W}^{\mathrm{n}}\mathbf{H}^{\mathrm{n}}}{(\mathbf{W}\mathbf{H})^2}\mathbf{H}^{\mathrm{s}\top}} \tag{7}$$

$$\mathbf{W}^{\mathrm{n}} \leftarrow \mathbf{W}^{\mathrm{n}} \otimes \frac{\frac{\mathbf{M}\otimes\mathbf{W}^{\mathrm{s}}\mathbf{H}^{\mathrm{s}}}{(\mathbf{W}\mathbf{H})^2}\mathbf{H}^{\mathrm{n}\top}}{\frac{\mathbf{S}^{\mathrm{s}}}{\mathbf{W}\mathbf{H}}\mathbf{H}^{\mathrm{n}\top}} \tag{8}$$

により更新する方法がとられている。上述の更新式は $f(\mathbf{W},\mathbf{H})$ の $\mathbf{W}$ に関する偏微分の負の項と正の項の商と $\mathbf{W}$ の要素ごとの積で与えられるが，各更新により目的関数が減少することが保証されない。このため，これらの更新式による反復アルゴリズムの収束性は保証されない。

3 補助関数法による基底学習アルゴリズム

前述の乗法更新アルゴリズムの問題を解決するため，本稿では，補助関数法の原理に基づいて導かれる，(6) の最適化問題の停留点への収束性が保証された最適化アルゴリズムを提案する。本節では，補助関数法とアルゴリズムの導出について述べる。

3.1 補助関数法

$F(\theta)$ を θ に関して最小化したい目的関数とすると，$F(\theta) = \min_\alpha F^+(\theta,\alpha)$ を満たす関数 $F^+(\theta,\alpha)$ を補助関数，α を補助変数と呼ぶ。このような補助関数を設計できれば，$\alpha \leftarrow \operatorname{argmin}_\alpha F^+(\theta,\alpha)$ と $\theta \leftarrow \operatorname{argmin}_\theta F^+(\theta,\alpha)$ を交互に繰り返すことで，目的関数 $F(\theta)$ の停留点を得ることができる。この最適化手法を補助関数法と呼ぶ。

3.2 補助関数の設計

以下で，目的関数 $f(\mathbf{W},\mathbf{H})$ の補助関数を設計する。まず，目的関数 $f(\mathbf{W},\mathbf{H})$ の中の

$$\frac{\sum_{k=1}^{K^{\mathrm{s}}} W^{\mathrm{s}}_{\omega,k} H^{\mathrm{s}}_{k,t}}{\sum_{k=1}^{K} W_{\omega,k} H_{k,t}} \tag{9}$$

の補助関数を次の不等式を用いて設計する。

補題 1. 任意の $a \in \mathbb{R}^{>0}$，$b \in \mathbb{R}^{>0}$ に対して，不等式

$$\frac{a}{b} \leq \frac{\lambda a^2}{2} + \frac{1}{2\lambda b^2} \tag{10}$$

が成り立ち，$\lambda = 1/(ab)$ のとき等号成立する。

証明: 任意の $a, b, \lambda \in \mathbb{R}^{>0}$ に対して，

$$\begin{aligned}&\lambda\left(a - \frac{1}{\lambda b}\right)^2 = \lambda\left(a^2 - 2\frac{a}{\lambda b} + \frac{1}{\lambda^2 b^2}\right) \geq 0 \\ &\Rightarrow \ \frac{a}{b} \leq \frac{\lambda a^2}{2} + \frac{1}{2\lambda b^2}\end{aligned}$$

□

$M_{\omega,t}$ は非負値のため，補題 1 より，

$$\begin{aligned}&\mathcal{D}_{\mathrm{KL}}\left(\mathbf{S}^{\mathrm{s}} \,\middle|\, \frac{\mathbf{W}^{\mathrm{s}}\mathbf{H}^{\mathrm{s}}}{\mathbf{W}\mathbf{H}} \otimes \mathbf{M}\right) \\ &\overset{c}{=} \sum_{\omega,t}\left(-S^{\mathrm{s}}_{\omega,t}\log G^{\mathrm{s}}_{\omega,t} + S^{\mathrm{s}}_{\omega,t}\log G_{\omega,t} + \frac{G^{\mathrm{s}}_{\omega,t}}{G_{\omega,t}} M_{\omega,t}\right) \\ &\leq \sum_{\omega,t}\Bigg(-S^{\mathrm{s}}_{\omega,t}\log G^{\mathrm{s}}_{\omega,t} + S^{\mathrm{s}}_{\omega,t}\log G_{\omega,t} \\ &\qquad + \frac{\lambda_{\omega,t} M_{\omega,t} {G^{\mathrm{s}}_{\omega,t}}^2}{2} + \frac{M_{\omega,t}}{2\lambda_{\omega,t} G^2_{\omega,t}}\Bigg)\end{aligned} \tag{11}$$

が成り立つ。ただし，$=^c$ はパラメータに依存する項のみに関する等号を表す。また，$G^{\mathrm{s}}_{\omega,t} = \sum_{k=1}^{K^{\mathrm{s}}} W^{\mathrm{s}}_{\omega,k} H^{\mathrm{s}}_{k,t}$

とし，$G_{\omega,t}=\sum_{k=1}^{K}W_{\omega,k}H_{k,t}$ とする。(11) の等号は

$$\lambda_{\omega,t}=\frac{1}{G^{\mathrm{s}}_{\omega,t}G_{\omega,t}} \tag{12}$$

のとき成立する。次に，(11) の各項の補助関数を設計する。$S^{\mathrm{s}}_{\omega,t}$ は正値であること，および負の対数関数は凸関数であることより，Jensen の不等式

$$-\log G^{\mathrm{s}}_{\omega,t}\le-\sum_{k=1}^{K^{\mathrm{s}}}\gamma_{k,\omega,t}\log\frac{W^{\mathrm{s}}_{\omega,k}H^{\mathrm{s}}_{k,t}}{\gamma_{k,\omega,t}} \tag{13}$$

が成り立つ。ただし，$\gamma_{k,\omega,t}$ は $\gamma_{k,\omega,t}>0$, $\sum_k\gamma_{k,\omega,t}=1$ を満たす変数であり，(13) の等号は

$$\gamma_{k,\omega,t}=\frac{W^{\mathrm{s}}_{\omega,k}H^{\mathrm{s}}_{k,t}}{\sum_{k'}W^{\mathrm{s}}_{\omega,k'}H^{\mathrm{s}}_{k',t}} \tag{14}$$

のとき成立する。$S^{\mathrm{s}}_{\omega,t}$ は正値のため，(6) の第二項の対数関数は凹関数である。凹関数は任意の点における接線により上から抑えることができるため，

$$\log G_{\omega,t}\le\sum_k\frac{W_{\omega,k}H_{k,t}}{\eta_{\omega,t}}+\log\eta_{\omega,t}-1 \tag{15}$$

が成り立つ。ここで，$\eta_{\omega,t}$ は正の変数であり，

$$\eta_{\omega,t}=G_{\omega,t} \tag{16}$$

のとき，(15) の等号は成立する。続いて，${G^{\mathrm{s}}_{\omega,t}}^2$ の補助関数を設計する。二次関数は凸関数なので，Jensen の不等式

$${G^{\mathrm{s}}_{\omega,t}}^2\le\sum_{k=1}^{K^{\mathrm{s}}}\frac{{W^{\mathrm{s}}_{\omega,k}}^2{H^{\mathrm{s}}_{k,t}}^2}{\beta_{k,\omega,t}} \tag{17}$$

が成り立つ。ただし，$\beta_{k,\omega,t}$ は $\sum_k\beta_{k,\omega,t}=1$ を満たす正数であり，(17) の等号は

$$\beta_{k,\omega,t}=\frac{W^{\mathrm{s}}_{\omega,k}H^{\mathrm{s}}_{k,t}}{\sum_{k'=1}^{K^{\mathrm{s}}}W^{\mathrm{s}}_{\omega,k'}H^{\mathrm{s}}_{k',t}} \tag{18}$$

のときに成立する。最後に，$1/G^2_{\omega,t}$ の補助関数を設計する。関数 $1/x^2$ は $x>0$ においては凸であるため，Jensen の不等式により

$$\frac{1}{G^2_{\omega,t}}=\frac{1}{(\sum_k W_{\omega,k}H_{k,t})^2}=\frac{1}{\left(\sum_k\theta_{k,\omega,t}\frac{W_{\omega,k}H_{k,t}}{\theta_{k,\omega,t}}\right)^2}$$
$$\le\sum_k\theta_{k,\omega,t}\frac{1}{\left(\frac{W_{\omega,k}H_{k,t}}{\theta_{k,\omega,t}}\right)^2}=\sum_k\frac{\theta^3_{k,\omega,t}}{W^2_{\omega,k}H^2_{k,t}} \tag{19}$$

が成り立つ。ただし，$\theta_{k,\omega,t}$ は $\theta_{k,\omega,t}>0$, $\sum_k\theta_{k,\omega,t}=1$ を満たす変数である。(19) の等号は

$$\theta_{k,\omega,t}=\frac{W_{\omega,k}H_{k,t}}{\sum_{k'}W_{\omega,k'}H_{k,t'}} \tag{20}$$

のとき成立する。

(11)，(13)，(15)，(17) と (19) により，目的関数 $f(\mathbf{W},\mathbf{H})$ の補助関数

$$f(\mathbf{W},\mathbf{H})\le f^+(\mathbf{W},\mathbf{H},\mathbf{\Gamma})$$
$$=-\sum_{\omega,t,k}S^{\mathrm{s}}_{\omega,t}\gamma_{k,\omega,t}\log\frac{W^{\mathrm{s}}_{\omega,k}H^{\mathrm{s}}_{k,t}}{\gamma_{k,\omega,t}}+\sum_{\omega,t,k}\frac{S^{\mathrm{s}}_{\omega,t}W_{\omega,k}H_{k,t}}{\eta_{\omega,t}}$$
$$+\sum_{\omega,t,k}\frac{\lambda_{\omega,t}M_{\omega,t}}{2\beta_{k,\omega,t}}{W^{\mathrm{s}}_{\omega,k}}^2{H^{\mathrm{s}}_{k,t}}^2+\sum_{\omega,t,k}\frac{M_{\omega,t}\theta^3_{k,\omega,t}}{2\lambda_{\omega,t}W^2_{\omega,k}H^2_{k,t}}+d$$

を得ることができる。ここで，$\mathbf{\Gamma}$ は補助変数 $\lambda_{\omega,t}$，$\gamma_{k,\omega,t}$，$\eta_{\omega,t}$，$\beta_{k,\omega,t}$ と $\theta_{k,\omega,t}$ の集合であり，d は定数項である。この補助関数を導いたことのポイントは，$\mathbf{W}$ と $\mathbf{H}$ に関する大域最適解は解析的に得ることができる点にある。

3.3 パラメータの更新式

以上の補助関数を最小にする補助変数の条件は各不等式の等号成立条件に他ならないので，(12), (14), (16), (18), (20) で与えられる。また，補助関数を最小にする $\mathbf{W}$, $\mathbf{H}$ は $\partial f^+/\partial\mathbf{W}=0$ と $\partial f^+/\partial\mathbf{H}=0$，すなわち，

$$\sum_t\frac{\lambda_{\omega,t}M_{\omega,t}}{2\beta_{k,\omega,t}}{H^{\mathrm{s}}_{k,t}}^2{W^{\mathrm{s}}_{\omega,k}}^4+\sum_t\frac{S^{\mathrm{s}}_{\omega,t}H^{\mathrm{s}}_{k,t}}{\eta_{\omega,t}}{W^{\mathrm{s}}_{\omega,k}}^3$$
$$-\sum_t S^{\mathrm{s}}_{\omega,t}\gamma_{k,\omega,t}{W^{\mathrm{s}}_{\omega,k}}^2-\sum_t\frac{M_{\omega,t}\theta^3_{k,\omega,t}}{2\lambda_{\omega,t}{H^{\mathrm{s}}_{k,t}}^2}=0$$
$$\sum_t\frac{S^{\mathrm{s}}_{\omega,t}H^{\mathrm{n}}_{k_n,t}}{\eta_{\omega,t}}{W^{\mathrm{n}}_{\omega,k}}^3-\sum_t\frac{M_{\omega,t}\theta^3_{k,\omega,t}}{2\lambda_{\omega,t}{H^{\mathrm{n}}_{k,t}}^2}=0$$
$$\sum_\omega\frac{\lambda_{\omega,t}M_{\omega,t}}{2\beta_{k,\omega,t}}{W^{\mathrm{s}}_{\omega,k}}^2{H^{\mathrm{s}}_{k,t}}^4+\sum_\omega\frac{S^{\mathrm{s}}_{\omega,t}W^{\mathrm{s}}_{\omega,k}}{\eta_{\omega,t}}{H^{\mathrm{s}}_{k,t}}^3$$
$$-\sum_\omega S^{\mathrm{s}}_{\omega,t}\gamma_{k,\omega,t}{H^{\mathrm{s}}_{k,t}}^2-\sum_\omega\frac{M_{\omega,t}\theta^3_{k,\omega,t}}{2\lambda_{\omega,t}{W^{\mathrm{s}}_{\omega,k}}^2}=0$$
$$\sum_\omega\frac{S^{\mathrm{s}}_{\omega,t}W^{\mathrm{n}}_{\omega,k}}{\eta_{\omega,t}}{H^{\mathrm{n}}_{k,t}}^3-\sum_\omega\frac{M_{\omega,t}\theta^3_{k,\omega,t}}{2\lambda_{\omega,t}{W^{\mathrm{n}}_{\omega,k}}^2}=0$$

のような四次方程式と三次方程式の正数解を解くことにより得られる。上記四次方程式の定数項と二次式の係数はいずれも負値であるため，必ず一つの正数解のみを持つことが示される。

4 評価実験

提案手法による音声強調効果を検証するため，ATR 音声データベース 503 文の音声データ [3] と ATR 環境音データベース（department noise, subway station noise の 2 種類）を用いて評価実験を行った。比較対象は従来の教師あり NMF 法（SNMF）と識別的 NMF の乗法更新アルゴリズム（DNMF_MU）とし，処理前と処理後の信号対歪み比 (SDR) および信号対干渉比 (SIR)[4] の改善値を評価した。

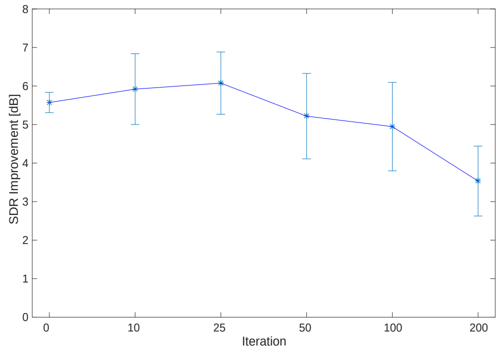

Fig. 1　提案アルゴリズムにおける反復回数が {0, 10, 25, 50, 100, 200} のときの SDR 改善量 [dB] の平均と分散

Table 1　各手法を 5 回試行して得られた SDR 改善量平均値 [dB]。上段は Department ノイズにおける音声強調結果であり，下段は Subway station ノイズにおける音声強調結果である。

Method	Input SNR				
	-6 dB	-3 dB	0 dB	3 dB	Avg
SNMF	5.58	5.53	5.18	4.64	5.23
DNMF_MU	5.88	5.68	5.11	**4.70**	5.34
DNMF_AU	**6.41**	**6.29**	**5.72**	**4.70**	**5.78**
SNMF	5.79	5.65	5.19	4.06	5.17
DNMF_MU	5.51	5.86	5.22	4.80	5.35
DNMF_AU	**6.82**	**7.20**	**6.50**	**4.89**	**6.35**

テストデータはクリーン音声に各雑音を-6, -3, 0, 3 dB の信号対雑音比 (SNR) で重畳させて作成した。実験に用いた音響信号はサンプリング周波数 16kHz のモノラル信号で，フレーム長 32ms，フレームシフト 16ms で短時間 Fourier 変換を行い，観測スペクトログラム $\mathbf{Y}$ を算出した。基底学習においては男性 2 名と女性 2 名の話者の計 200 文の音声を用いて音声基底の学習を行った。基底数は音声と雑音両方 40 とした。ランダムに決めた初期値で反復アルゴリズムを 5 回試行し，各試行における反復回数が 0, 10, 25, 50, 100, 200 の時の SDR 改善値の平均と分散をプロットしたものが図 1 である。図 1 の結果により，以下の実験では反復回数を 25 とした。テストデータセットは，ATR503 文データベースからランダムに選定した 40 文の音声データに雑音を重畳させて作成した。

以上の条件下で，提案法 (DNMF_AU) と従来法 (SNMF, DNMF_MU) を 5 回試行して得られた SDR および SIR の改善値の平均を表 1, 2 に示す。いずれの評価尺度においても全ての場合において提案手法の方が高い改善値を得られていることが確認できる。

Table 2　各手法を 5 回試行して得られた SIR 改善量平均値 [dB]。上段は Department ノイズにおける音声強調結果であり，下段は Subway station ノイズにおける音声強調結果である。

Method	Input SNR				
	-6 dB	-3 dB	0 dB	3 dB	Avg
SNMF	7.23	7.44	7.44	7.31	7.36
DNMF_MU	8.07	7.87	7.44	7.34	7.68
DNMF_MM	**9.76**	**9.66**	**10.16**	**9.74**	**9.83**
SNMF	7.78	8.04	8.10	8.16	8.02
DNMF_MU	8.04	8.67	7.95	8.29	8.24
DNMF_MM	**10.77**	**11.58**	**11.89**	**11.28**	**11.38**

5　おわりに

従来の教師あり NMF の基底学習においては、学習サンプルのスペクトルへのフィッティング規準を用いられることが多く，分離信号が最適となるような規準とはなっていなかった。そこで，より高い分離精度を得るには，分離信号そのものを最適化規準として基底を学習する識別的 NMF と呼ぶ枠組が提案されている。識別的 NMF の学習規準は解析的に複雑な形になるため，乗法更新アルゴリズムと呼ぶ汎用的な手法を用いた最適アルゴリズムが提案されている。しかし、乗法更新アルゴリズムは停留点への収束性が保証されていない。本稿では，補助関数法に基づく，収束性が保証された識別的 NMF の基底学習アルゴリズムを提案し、実験によりその効果を確認した。

謝辞　本研究は JSPS 科研費 26730100 と 26280060 の助成を受けて行われた。

参考文献

[1] P. Smaragdis, R. Bhiksha, and S. Madhusudana, “Supervised and semi-supervised separation of sounds from single-channel mixtures.”, In Proc. ICA, pp. 414–421, 2007.

[2] F. Weninger, J. L. Roux, J. R. Hershey, and S. Watanabe, “Discriminative NMF and its application to single-channel source separation.”, In Proc. INTERSPEECH, pp. 865–869, 2014.

[3] A. Kurematsu, K. Takeda, Y. Sagisaka, S. Katagiri, H. Kuwabara, and K. Shikano, “ATR Japanese speech database as a tool of speech recognition and synthesis,” Speech Communication, vol. 9, pp. 357–363, 1990.

[4] E. Vincent, R. Gribonval, and C. Fevotte, “Performance measurement in blind audio source separation.”, IEEE transactions on audio, speech, and language processing, vol. 14, no. 4, pp. 1462–1469, 2016.