Vol.2017-MUS-116 No.5
2017/8/24

複素ラプラス分布に基づく非負値行列因子分解

丹治 寛樹[1,a)]　村上 隆啓[1]　鎌田 弘之[1]

概要：本稿では，信号分離への適用を目指し，実環境の音声信号のスペクトルが従う確率分布を考慮した非負値行列因子分解（nonnegative matrix factorization; NMF）の評価関数と最適化アルゴリズムを提案する．信号分離で対象となる音響信号は，調波構造を持つことが多いため，音響信号のスペクトルは複素領域で優ガウス性を持つ．そこで，本稿では，複素ラプラス分布と呼ばれる原点付近に鋭いピークを持つ分布に基づく NMF の評価関数を提案する．提案する評価関数は特殊関数を含むため，従来の NMF の評価関数と比較して最適化は非常に困難である．しかし，NMF の評価関数を最適化する majorization-minimization（MM）アルゴリズムの導出に従来から用いられている不等式に加え，確率分布に対する Jensen の不等式を利用することで，MM アルゴリズムを導出できることを示す．シミュレーションの結果，提案する NMF は従来の NMF と比較して遜色ない分離性能を持つことを確認した．

Nonnegative Matrix Factorization Based on Complex Laplace Distribution

Hiroki Tanji[1,a)]　Takahiro Murakami[1]　Hiroyuki Kamata[1]

1. はじめに

非負値行列因子分解（nonnegative matrix factorization; NMF）とは，観測された非負の行列を非負の基底と重みの積に分解する手法である [1,2]．NMF が初めて適用されたのは画像からの特徴抽出 [1] であるが，近年，モノラル音源の信号分離の手法として盛んに用いられるようになった [3–10]．例えば，文献 [3] では，楽音の振幅スペクトログラムに NMF を適用することで，楽音を構成する単一音のスペクトルとその混合比を推定できることが報告されている．文献 [3] のように，音響信号に NMF を適用する場合，観測行列には音響信号の振幅またはパワースペクトログラムを用いるのが一般的である．このとき，NMF は観測行列をランク 1 の非負行列の和で近似するため，観測信号の振幅またはパワースペクトルを音源の振幅またはパワースペクトルの和で近似できると仮定する必要がある．

NMF は，観測行列と基底および重みから再構成した行列との乖離度を評価関数とし，この評価関数を基底および重みについての非負制約の下で最小化する最適化問題に帰着する．古典的には，乖離度としてユークリッド距離の二乗 [1] や Kullback-Leibler（KL）divergence [1] がよく用いられている．NMF はパラメータに非負制約があるため，評価関数の最適化のために majorization-minimization（MM）アルゴリズム [11] あるいは補助関数法 [12] と呼ばれる手法を用いて，乗法形の更新式を導出するのが一般的である．MM アルゴリズムを導出するためには，評価関数に対して一定の条件を満たす上限を設定する必要がある．上限の設定のために，従来の NMF では凸関数に対する Jensen の不等式や凹関数の 1 次のテイラー展開が広く用いられている [4,5,12]．

近年，信号分離における NMF の発展を背景に，複素スペクトルの統計的な生成モデルに基づいた NMF が盛んに研究されている．文献 [6] では，Itakura-Saito（IS）divergence に基づいて観測信号のパワースペクトルを分解する IS-NMF が提案されている．IS-NMF では，観測信号の複素スペクトルが複素正規分布に従うと仮定し，複素正規分布の分散を最尤推定する問題として NMF を定式化する．このように定式化すると，複素正規分布の再生性から，音源の複素スペクトルの分散の和が観測信号の分散となるため，パワースペクトルの加法性を正当化できる．さらに，文献 [4]

1　明治大学理工学部 電気電生命学科
Meiji University, Tama-ku, Kawasaki-shi, Kanagawa 214–8571, Japan
a)　htanji@meiji.ac.jp

では，観測信号の複素スペクトルが複素コーシー分布に従うと仮定し，複素コーシー分布の再生性に基づいて，振幅スペクトルの加法性を正当化する Cauchy-NMF が提案されている．

一方，確率分布の再生性を利用して振幅またはパワースペクトルの加法性を正当化することは，必ずしも必要ではない．例えば，文献 [5] では，t-NMF と呼ばれる，複素 t 分布に基づく NMF が提案されている．t-NMF では，観測信号の複素スペクトルが複素 t 分布に従うと仮定する．複素 t 分布は，その自由度が 1 のとき複素コーシー分布に，∞ のとき複素正規分布と等しくなるため，t-NMF は IS-NMF と Cauchy-NMF の一般化となる．分布の加法性は自由度が 1 または ∞ のときにしか成立しないが，文献 [5] では，自由度が 2 のときに t-NMF が高い分離性能をもつことが報告されている．

さらに，音響信号の複素スペクトルにとって物理的な意味を持たないような乖離度でも，信号分離のタスクにおいては高い性能を持つことがある．例えば，KL divergence に基づく NMF（KL-NMF）は，観測信号の振幅スペクトルがポアソン分布に従うと仮定し，ポアソン分布の母数を最尤推定する問題と等価になる．ポアソン分布は離散的な確率変数上で定義される分布であるため，振幅スペクトルやパワースペクトルのような連続量のモデリングには適さないように思える．しかし，信号分離のタスクでは，観測行列を振幅スペクトログラムとした場合，KL-NMF がもっとも高い性能を持つことが報告されている [4,7]．

近年，NMF の評価関数には，以下の 3 点が重視されている．
- 尤度関数に複素分布を用いるか
- 確率分布の再生性が成り立つか
- 複数の評価関数の一般化になるか

本稿では，1 つ目の観点に加え，さらに異なる観点から NMF の評価関数を構成する．すなわち，実環境の音響信号がどのような分布に従うかを考慮し，より自然な観測信号の統計モデルを構築することで，NMF を定式化することを試みる．

信号分離の対象となる音響信号の複素スペクトルは調波構造を持つことが多い．このようなスペクトルは周波数領域上で局所的なピークを持つことから，優ガウス性を持つと推測できる．実際に，音響信号の複素スペクトルの実部および虚部は，正規分布よりもラプラス分布の方がよりよくフィットすることが実験的に示されている [13]．したがって，観測信号の複素スペクトルは複素数に拡張したラプラス分布に従うと仮定するのが自然である．

ラプラス分布は正規分布の分散が指数分布に従うと仮定し，分散を積分消去することで導出できる．文献 [14] では，このラプラス分布の特性に注目し，複素正規分布と指数分布を用いてラプラス分布を複素数に拡張している．本稿では，文献 [14] の複素ラプラス分布に着目し，観測信号の複素スペクトルが複素ラプラス分布に従うと仮定する Laplace-NMF を提案する．複素ラプラス分布は再生性が成り立たないため，振幅またはパワースペクトルの加法性を正当化できないが，観測信号のモデリングという視点では，提案法の仮定は極めて妥当である．

文献 [15] で提案されている球状ラプラス分布もラプラス分布の複素数への拡張として知られている．複素ラプラス分布が指数分布によって分散のスパース性を仮定するのに対し，球状ラプラス分布では，ガンマ分布を用いて分散に対し近似的なスパース性を仮定しているため，両者は異なる分布関数を与える．本稿では，分散のスパース性を厳密に仮定した方が観測信号を適切にモデリングできると考え，複素ラプラス分布を用いて NMF の評価関数を構築する．

複素ラプラス分布の分布関数は特殊関数を含むため，Laplace-NMF の評価関数の最適化は非常に困難である．なぜなら，従来 NMF において広く用いられている MM アルゴリズムのアプローチでは，アルゴリズムの導出が可能な上限を設定できないためである．本稿では，従来よく用いられている不等式に加えて，確率分布についての Jensen の不等式を利用することで評価関数の上限を設定し，収束の保証された最適化アルゴリズムを導出する．このアプローチは期待値の計算を必要とするが，Laplace-NMF においては計算機で評価可能な期待値を求めることができる．

2. 非負値行列因子分解

本節では，NMF の一般的な定式化について述べる．NMF では，全ての要素が非負の観測行列 $\boldsymbol{Y} = [y_{mn}] \in \mathbb{R}_+^{M\times N}$ を基底 $\boldsymbol{W} = [w_{mk}] \in \mathbb{R}_+^{M\times K}$ と重み $\boldsymbol{H} = [h_{kn}] \in \mathbb{R}_+^{K\times N}$ の積 $\hat{\boldsymbol{Y}} = \boldsymbol{W}\boldsymbol{H}$ で近似する．ここで，$\hat{\boldsymbol{Y}} = [\hat{y}_{mn}] \in \mathbb{R}_+^{M\times N}$ は観測行列の近似，$\mathbb{R}_+$ は半直線 $\mathbb{R}_+ = [0, \infty)$ である．

NMF は，一般的に，w_{mk} と h_{nk} の非負制約の下で，以下の評価関数を最小化することで，最適な $\boldsymbol{W}$ および $\boldsymbol{H}$ を推定する．

$$F(\boldsymbol{W}, \boldsymbol{H}) = \sum_{m=1}^{M}\sum_{n=1}^{N} f(y_{mn}; \hat{y}_{mn}) \tag{1}$$

$f(x; \hat{x})$ は，x と $\hat{x}$ の乖離度で，以下の条件を満たす．
(1) 全ての x，$\hat{x}$ について，$f(x; \hat{x}) \geq 0$
(2) $x = \hat{x}$ ならば，$f(x; \hat{x}) = 0$
1 つ目の条件は，パラメータと無関係な定数を足すことで容易に満たすことができるため，この定数は省略して書かれることがある [4,5]．

NMF を振幅スペクトルに適用する場合，$\hat{y}_{mn} = \sum_{k=1}^{K} w_{mk} h_{kn}$ であることから，以下の振幅スペクトルの加法性を仮定する必要がある．

$$|y_{mn}^{\mathbb{C}}| = \sum_{k=1}^{K} \varsigma_{mnk} \tag{2}$$

ここで，$\{y^{\mathbb{C}}_{mn}\}_m$ は時刻 n における観測信号の複素スペクトルの系列，ς_{mnk} は，振幅スペクトルの推定値で，$\varsigma_{mnk} = w_{mk}h_{kn}$ である．パワースペクトルに適用する場合は, パワースペクトルの加法性を仮定する．$\varsigma^2_{mnk} = w_{mk}h_{kn}$ をパワースペクトルの推定値とすれば，この仮定は次式で書かれる．

$$|y^{\mathbb{C}}_{mn}|^2 = \sum_{k=1}^{K} \varsigma^2_{mnk} \tag{3}$$

3. 複素ラプラス分布

本節では，文献 [14] で導出されている複素ラプラス分布について述べる．

定義 1 $\boldsymbol{\mu}, \boldsymbol{y} \in \mathbb{C}^d$，$z \in \mathbb{R}_+$ とし，$\boldsymbol{\Sigma}$ を正定値対称行列とする．$\boldsymbol{y}|z \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, z\boldsymbol{\Sigma})$，$z \sim \mathcal{E}(\lambda)$ であるとき，$\boldsymbol{y}$ の分布 $p(\boldsymbol{y})$ を多変量複素ラプラス分布と定義し，$\mathcal{L}_{\mathbb{C}}(\boldsymbol{y}; \boldsymbol{\mu}, \lambda, \boldsymbol{\Sigma})$ と書く．

$\mathcal{N}_{\mathbb{C}}$ は複素正規分布，$\mathcal{E}$ は指数分布である．本稿で用いる確率分布は付録 A.1 に記載する．1 変量の複素ラプラス分布の分布関数は，定義より次式で与えられる．

$$\mathcal{L}_{\mathbb{C}}(y; \mu, \lambda) = \frac{2}{\lambda\pi}\mathcal{K}_0\left(2\sqrt{\frac{|y-\mu|^2}{\lambda}}\right) \tag{4}$$

ここで，$\mathcal{K}_\nu(t), t \in \mathbb{R}_+$ は第二種変形ベッセル関数で，本稿ではその積分形を次式で定義する．

$$\mathcal{K}_\nu(t) = \frac{1}{2}\int_{\mathbb{R}_+} x^{-\nu-1}\exp\left(-\frac{t}{2}\left(x + \frac{1}{x}\right)\right)dx \tag{5}$$

$\mathcal{K}_\nu(t)$ は，一般の ν において解析的に解くことはできないが，計算機で数値解を評価することはできる．

4. 複素ラプラス分布に基づく NMF

本節では，観測信号の複素スペクトル $y^{\mathbb{C}}_{mn}$ が複素ラプラス分布に従うと仮定する Laplace-NMF を提案する．Laplace-NMF では，複素ラプラス分布の定義より，$y^{\mathbb{C}}_{mn}$ に対して以下を仮定することと等価になる．

(1) $y^{\mathbb{C}}_{mn}$ は平均 0，分散 z_{mn} の複素正規分布に従う
(2) z_{mn} は期待値 λ の指数分布に従う

二つ目の仮定では，分散 z_{mn} のスパース性を考慮している．Laplace-NMF では，さらに，あらゆる分散を考慮して，z_{mn} を積分消去する．

4.1 振幅スペクトルの分解問題としての定式化

本節では，観測信号の振幅スペクトルを分解する最尤推定問題として Laplace-NMF を定式化する．式 (4) を尤度関数とすることから，Laplace-NMF の評価関数 $F(\boldsymbol{W}, \boldsymbol{H})$ は次式で与えられる．

$$F(\boldsymbol{W}, \boldsymbol{H}) = \sum_{m,n}\left\{2\log\hat{y}_{mn} - \log\mathcal{K}_0\left(\frac{2y_{mn}}{c\hat{y}_{mn}}\right)\right\} \tag{6}$$

ここで，$y_{mn} = |y^{\mathbb{C}}_{mn}|$，$\hat{y}_{mn} = \sum_{k=1}^{K} w_{mk}h_{kn}$ である．c は評価関数の補正係数で，$y_{mn} = \hat{y}_{mn}, \forall m, n$ のときに評価関数が最小となるために必要である．c は以下の方程式の解で与えられる．

$$\left.\frac{\partial F(\boldsymbol{W}, \boldsymbol{H})}{\partial \hat{y}_{mn}}\right|_{\hat{y}_{mn}=y_{mn}} = 0 \tag{7}$$

最適な基底 $\boldsymbol{W}$ および重み $\boldsymbol{H}$ は，以下の非負制約つき最適化問題の解である．

$$\min_{\boldsymbol{W},\boldsymbol{H}} F(\boldsymbol{W}, \boldsymbol{H}), \text{subject to } w_{mk}, h_{kn} \geq 0, \forall m, n, k \tag{8}$$

4.2 最適化アルゴリズムの導出

式 (6) の最小化問題は，第二種変形ベッセル関数を含むため，式 (6) の上限を従来から NMF において広く用いられているアプローチと同様に設計しても，更新式を導出できない．そこで，本節では，確率分布に対する Jensen の不等式を利用することで評価関数の上限を求め，この上限に Jensen の不等式および凹関数の 1 次のテイラー展開を適用して上限を求めることで，式 (8) の最小化問題を解く．

確率分布に対する Jensen の不等式を利用することで，以下の不等式を得る．

$$\begin{aligned} F(\boldsymbol{W}, \boldsymbol{H}) &\overset{c}{=} -\sum_{m,n}\log p(y_{mn}; \hat{y}_{mn}) \\ &\leq -\sum_{m,n}\int_{\mathbb{R}_+} p(z_{mn}|y_{mn}; \tilde{\hat{y}}_{mn})\log\frac{p(y_{mn}, z_{mn}; \hat{y}_{mn})}{p(z_{mn}|y_{mn}; \tilde{\hat{y}}_{mn})}dz_{mn} \\ &\overset{c}{=} Q(\boldsymbol{W}, \boldsymbol{H}) \end{aligned} \tag{9}$$

ここで，$\tilde{\hat{y}}_{mn}$ は更新前の $\hat{y}_{mn}$，$z_{mn} \in \mathbb{R}_+$ は隠れ変数，$\overset{c}{=}$ は定数項を除いて等しいことを意味する．複素ラプラス分布の定義より，$p(y_{mn}, z_{mn}; \hat{y}_{mn})$ は以下のように与えられる．

$$p(y_{mn}, z_{mn}; \hat{y}_{mn}) = p(y_{mn}|z_{mn})p(z_{mn}; \hat{y}_{mn}) \tag{10}$$

$$p(y_{mn}|z_{mn}) = \frac{1}{z_{mn}\pi}\exp\left(-\frac{y^2_{mn}}{z_{mn}}\right) \tag{11}$$

$$p(z_{mn}; \hat{y}_{mn}) = \frac{1}{c^2\hat{y}^2_{mn}}\exp\left(-\frac{z_{mn}}{c^2\hat{y}^2_{mn}}\right) \tag{12}$$

式 (10)–(12) より，$Q(\boldsymbol{W}, \boldsymbol{H})$ は以下のようになる．

$$Q(\boldsymbol{W}, \boldsymbol{H}) = \sum_{m,n}\left(2\log\hat{y}_{mn} + \frac{\mathbb{E}_{p(z_{mn}|y_{mn}; \tilde{\hat{y}}_{mn})}[z_{mn}]}{c^2\hat{y}^2_{mn}}\right) \tag{13}$$

式 (13) を最小化することで，$F(\boldsymbol{W}, \boldsymbol{H})$ を最小化できる．式 (13) は $\hat{y}_{mn}$ についての初等関数のみから構成されているため，$F(\boldsymbol{W}, \boldsymbol{H})$ と比較して，最適化は容易になる．パラメータの更新の過程で z_{mn} の事後分布の期待値 $\mathbb{E}_{p(z_{mn}|y_{mn}; \hat{y}_{mn})}[z_{mn}] = \int_{\mathbb{R}_+} z_{mn}p(z_{mn}|y_{mn}; \hat{y}_{mn})dz_{mn}$ を計算する必要がある．この期待値は次式で与えられる．

$$\mathbb{E}_{p(z_{mn}|y_{mn}; \hat{y}_{mn})}[z_{mn}] = cy_{mn}\hat{y}_{mn}\frac{\mathcal{K}_1\left(\frac{2y_{mn}}{c\hat{y}_{mn}}\right)}{\mathcal{K}_0\left(\frac{2y_{mn}}{c\hat{y}_{mn}}\right)} \tag{14}$$

式 (14) の導出は付録 A.2 に載せる．

式 (13) は expectation-maximization（EM）アルゴリズムにおいて最適化される条件付き期待値と見なせる．文献 [16] で提案されている EM アルゴリズムを用いた NMF の評価関数の最適化手法では，隠れ変数を $w_{mk}h_{kn}$ としているのに対し，提案法では，観測信号のパワースペクトルの期待値を隠れ変数としていることから，文献 [16] と提案法のアプローチは異なる．

評価関数に対して確率分布についての Jensen の不等式を適用することにより，式 (8) の最小化問題はより容易な $Q(\boldsymbol{W},\boldsymbol{H})$ の最小化問題に置き換わった．$Q(\boldsymbol{W},\boldsymbol{H})$ は $\hat{y}_{mn}$ についての非線形関数を含むことから，以下の不等式を適用することで，さらに $Q(\boldsymbol{W},\boldsymbol{H})$ の上限を最小化する問題に置き換える．

$$\log \hat{y}_{mn} \leq \frac{1}{\varphi_{mn}}(\hat{y}_{mn}-\varphi_{mn})+\log\varphi_{mn} \tag{15}$$

$$\frac{1}{\hat{y}_{mn}^2} \leq \sum_{k=1}^{K}\frac{\rho_{mnk}^3}{(w_{mk}h_{kn})^2} \tag{16}$$

ここで，$\varphi_{mn},\rho_{mnk}\geq 0,\forall m,n,k$ で，ρ_{mnk} は $\sum_{k=1}^{K}\rho_{mnk}=1$ を満たす．式 (15) および式 (16) はそれぞれ 1 次のテイラー展開，および Jensen の不等式である．式 (15) および式 (16) はそれぞれ $\varphi_{mn}=\hat{y}_{mn}$，$\rho_{mnk}=\frac{w_{mk}h_{kn}}{\hat{y}_{mn}}$ のときに等号が成立する．式 (15) および式 (16) を式 (13) に代入することで，$Q(\boldsymbol{W},\boldsymbol{H})$ の上限 $Q^{+}(\boldsymbol{W},\boldsymbol{H},\boldsymbol{\varphi},\boldsymbol{\rho})$ を得る．

$$Q^{+}(\boldsymbol{W},\boldsymbol{H},\boldsymbol{\rho},\boldsymbol{\varphi})=\sum_{m,n}\left[2\left\{\frac{1}{\varphi_{mn}}(\hat{y}_{mn}-\varphi_{mn})+\log\varphi_{mn}\right\}\right.$$
$$\left.+\frac{1}{c^2}\mathbb{E}_{p(z_{mn}|y_{mn};\tilde{y}_{mn})}[z_{mn}]\sum_{k=1}^{K}\frac{\rho_{mnk}^3}{(w_{mk}h_{kn})^2}\right] \tag{17}$$

ここで，$\boldsymbol{\varphi}=\{\varphi_{mn}\}$，$\boldsymbol{\rho}=\{\rho_{mnk}\}$ である．$Q(\boldsymbol{W},\boldsymbol{H})$ の代わりに $Q^{+}(\boldsymbol{W},\boldsymbol{H},\boldsymbol{\varphi},\boldsymbol{\rho})$ を最小化することで，$Q(\boldsymbol{W},\boldsymbol{H})$ を最小化できる．したがって，$Q^{+}(\boldsymbol{W},\boldsymbol{H},\boldsymbol{\varphi},\boldsymbol{\rho})$ を最小化することで $F(\boldsymbol{W},\boldsymbol{H})$ を最小化できる．

w_{mk} および h_{kn} の更新式は，$Q^{+}(\boldsymbol{W},\boldsymbol{H},\boldsymbol{\varphi},\boldsymbol{\rho})$ を w_{mk} および h_{kn} についてそれぞれ偏微分して 0 とおいた方程式を解くことで求められる．t 回目の更新における w_{mk} および h_{kn} を $w_{mk}^{(t)}$，$h_{kn}^{(t)}$ とし，$\hat{y}_{mn}^{(t)}=\sum_{k=1}^{K}w_{mk}^{(t)}h_{kn}^{(t)}$ とすると，更新式は以下のように書ける．

$$w_{mk}^{(t+1)}=w_{mk}^{(t)}\left(\frac{\displaystyle\sum_{n=1}^{N}\frac{\mathbb{E}_{p(z_{mn}|y_{mn};\hat{y}_{mn}^{(t)})}[z_{mn}]}{c^2\hat{y}_{mn}^{(t)\,3}}h_{kn}^{(t)}}{\displaystyle\sum_{n=1}^{N}\frac{h_{kn}^{(t)}}{\hat{y}_{mn}^{(t)}}}\right)^{\frac{1}{3}} \tag{18}$$

$$h_{kn}^{(t+1)}=h_{kn}^{(t)}\left(\frac{\displaystyle\sum_{m=1}^{M}\frac{\mathbb{E}_{p(z_{mn}|y_{mn};\hat{y}_{mn}^{(t)})}[z_{mn}]}{c^2\hat{y}_{mn}^{(t)\,3}}w_{mk}^{(t)}}{\displaystyle\sum_{m=1}^{M}\frac{w_{mk}^{(t)}}{\hat{y}_{mn}^{(t)}}}\right)^{\frac{1}{3}} \tag{19}$$

4.3 パワースペクトルの分解問題としての定式化

Laplace-NMF は，パワースペクトルを分解する問題としても定式化することができる．y_{mn} を観測信号のパワースペクトルを $y_{mn}=|y_{mn}^{\mathbb{C}}|^2$ とすると，式 (4) より，評価関数 $F(\boldsymbol{W},\boldsymbol{H})$ は次式となる．

$$F(\boldsymbol{W},\boldsymbol{H})=\sum_{m,n}\left\{\log\hat{y}_{mn}-\log\mathcal{K}_0\left(\frac{2}{c}\sqrt{\frac{y_{mn}}{\hat{y}_{mn}}}\right)\right\} \tag{20}$$

w_{mk} および h_{kn} の更新式は 4.2 節と同様のアプローチで導出できる．

4.4 球状ラプラス分布との関係

本節では，文献 [15] で提案されている球状ラプラス分布について述べる．さらに，球状ラプラス分布に基づく NMF が観測行列を振幅スペクトログラムとした IS-NMF と等価になることを示す．

複素ベクトルに対する球状ラプラス分布は以下のように定義される．*1

定義 2 $\boldsymbol{\mu},\boldsymbol{y}\in\mathbb{C}^d$，$z\in\mathbb{R}_+$ とし，$\boldsymbol{\Sigma}$ を正定値対象行列とする．$\boldsymbol{y}|z\sim\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu},z\boldsymbol{\Sigma})$，$z\sim\mathcal{G}(\frac{2d+1}{2},\lambda^{-1})$ であるとき，$\boldsymbol{y}$ の分布 $p(\boldsymbol{y})$ を球状ラプラス分布と定義し，$\mathcal{SL}_{\mathbb{C}}(\boldsymbol{y};\boldsymbol{\mu},\lambda,\boldsymbol{\Sigma})$ と書く．

ここで，$\mathcal{G}$ はガンマ分布である．定義 2 のように，分散の分布を $\mathcal{G}(\frac{2d+1}{2},\lambda^{-1})$ とすると，λ が小さいとき，近似的に分散のスパース性を仮定することになる．しかし，$z=0$ 付近の密度は小さいため，一般の λ の場合には分散のスパース性の仮定にはならないことに注意が必要である．

球状ラプラス分布の分布関数も第二種変形ベッセル関数を含むが，この場合は式 (5) の積分を解析的に求めることができる．複素数 $y\in\mathbb{C}$ に対する球状ラプラス分布の分布関数は，次式となる．

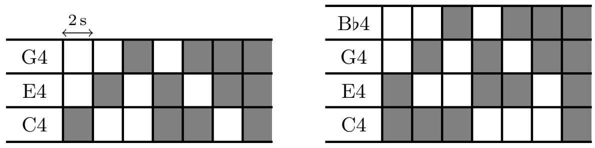
$$\mathcal{SL}_{\mathbb{C}}(y;\mu,\lambda)=\frac{2}{\lambda\pi}\exp\left(-2\sqrt{\frac{|y-\mu|^2}{\lambda}}\right) \tag{21}$$

観測信号の複素スペクトル $y_{mn}^{\mathbb{C}}$ が $\mathcal{SL}_{\mathbb{C}}(y_{mn}^{\mathbb{C}};0,\hat{y}_{mn}^2)$ に従うと仮定すると，観測信号の振幅スペクトル $|y_{mn}^{\mathbb{C}}|=y_{mn}$ を分解する問題として，NMF を定式化できる．評価関数 $F(\boldsymbol{W},\boldsymbol{H})$ は，式 (21) の対数を用いれば，次式のようになる．

$$F(\boldsymbol{W},\boldsymbol{H})=\sum_{m,n}\left(2\log\hat{y}_{mn}+\frac{2y_{mn}}{\hat{y}_{mn}}\right) \tag{22}$$

式 (22) は IS divergence の 2 倍から定数部を無視したものと等しい．したがって，球状ラプラス分布に基づく NMF は，振幅スペクトルを IS divergence に基づいて分解する IS-NMF と等価になる．

*1 文献 [15] で提案されている球状ラプラス分布は，厳密には実ベクトルに対する分布である．複素ベクトルに拡張する場合，複素ベクトルが 2 つの実ベクトルで構成できることから，次元 d を 2 倍する必要がある．



(a) $K = 3$　(b) $K = 4$

図 1　観測信号のピアノロール（網掛けの部分で音源がアクティブになる）
Fig. 1　Piano-roll of the observed signals. (Each source is active at the filled cells.)

5. シミュレーション

信号分離における Laplace-NMF と従来の NMF の性能をシミュレーションにより比較する．シミュレーションには，RWC 音楽データベース [17] に収録されたピアノ（011PFNOM）およびエレキギター（131EGLPM）を用いる．各音声ファイルから C4，E4，G4，B♭4 の音高に相当する 2 秒の区間を切り出し，図 1 のように重畳することで合計で 4 種類の観測信号を作成する．観測信号は 11025Hz でリサンプリングし，短時間フーリエ変換することでスペクトログラムを得る．このとき，窓関数をハミング窓，フレーム長を 512 点，フレーム周期を 128 点とし，1024 点で FFT を行う．NMF の観測行列は，得られたスペクトログラムから折り返し成分と直流成分およびナイキスト周波数成分を除外し，振幅またはパワーを算出することで作成する．

比較対象は，Eu-，KL-，IS-，Cauchy-，t-NMF とし，Eu-，KL-，IS-NMF，の観測行列には振幅およびパワースペクトログラムを用い，それぞれ 2 通りのシミュレーションを行う．t-NMF の自由度は $\nu = 2, 5$ とし，Laplace-NMF の補正係数 c はニュートン法により算出した小数点第 14 位までの値を用いる．基底の数 K は図 1(a) により作成した音源の場合は 3，図 1(b) の場合は 4 とする．基底および重みは 100 通りの乱数を用いて初期化し，3000 回更新した後の基底および重みから一般化ウィーナーフィルタ [18] を用いて分離信号を求める．Cauchy-NMF の最適化アルゴリズムには，naive multiplicative update アルゴリズム [4] を用い，他の従来の NMF は MM アルゴリズム [5, 12] により最適化する．分離結果は，source-to-distortion ratio（SDR）および source-to-interferences ratio（SIR）[19] により評価する．

図 2 にシミュレーション結果を示す．図 2 より，Laplace-NMF の分離性能は他の NMF と比較して遜色ないことがわかる．各 NMF の分離性能は少なからず音源に依存するため，全体的な傾向を把握するのは困難である．Laplace-NMF のみに着目すると，パワースペクトログラムよりも振幅スペクトログラムに適用した方が SDR および SIR の外れ値が少なくなる傾向にあることがわかる．これは，振幅スペクトログラムに適用する方が評価関数が最適解付近でより急峻になり，少ない更新回数で良好な分離結果が得られやすくなるためと思われる．Eu-NMF および KL-NMF が他の NMF と比較して際立って外れ値が少ないのもそのためである．

複素ラプラス分布は球状ラプラス分布と比較してより強いスパース性を仮定している．そのため，Laplace-NMF と振幅スペクトログラムを IS divergence に基づいて分離する NMF の間に分離性能の違いが見られることが期待されたが，このシミュレーションにおいては両者の分離性能に明確な差は見られなかった．

6. まとめ

本稿では，信号分離への適用を目指し，実環境の音響信号を適切に表現できる複素ラプラス分布を用いて評価関数を構築する Laplace-NMF を提案した．さらに，シミュレーションによって Laplace-NMF がモノラル音源を分離できることを確認した．今後は，Laplace-NMF の性能や初期値依存性についてさらに検証を進める．

付　　録

A.1 本稿で用いる確率分布

本稿で用いる確率分布を以下に列挙する．

複素正規分布　$\boldsymbol{y}, \boldsymbol{\mu} \in \mathbb{C}^d, \boldsymbol{\Sigma} > 0$

$$\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \pi^{-d}|\boldsymbol{\Sigma}|^{-1} \exp\left(-(\boldsymbol{y} - \boldsymbol{\mu})^\dagger \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})\right) \tag{A.1}$$

指数分布　$x, \lambda \in \mathbb{R}_+$

$$\mathcal{E}(\lambda) = \lambda^{-1} \exp(-\lambda^{-1}x) \tag{A.2}$$

ガンマ分布　$x, \beta \in \mathbb{R}_+, \eta > 1$

$$\mathcal{G}(\eta, \beta) = \frac{\beta^\eta}{\Gamma(\eta)} x^{\eta-1} \exp(-\beta x) \tag{A.3}$$

A.2 期待値の導出

式 (14) を導出する．複素ラプラス分布の定義より，$p(y_{mn}; \hat{y}_{mn})$ は以下のように書ける．

$$p(y_{mn}; \hat{y}_{mn}) = \int_{\mathbb{R}_+} p(y_{mn}|z_{mn}) p(z_{mn}; \hat{y}_{mn}) dz_{mn} \tag{A.4}$$

式 (A.4) の両辺を $(c\hat{y}_{mn})^{-2}$ で微分すると，右辺に y_{mn} と z_{mn} の同時分布の期待値 $\mathbb{E}_{p(y_{mn}, z_{mn}; \hat{y}_{mn})}[z_{mn}]$ が出現する．これを次式に代入することにより，式 (14) が導かれる．

$$\mathbb{E}_{p(z_{mn}|y_{mn}; \hat{y}_{mn})}[z_{mn}] = \frac{\mathbb{E}_{p(y_{mn}, z_{mn}; \hat{y}_{mn})}[z_{mn}]}{p(y_{mn}; \hat{y}_{mn})} \tag{A.5}$$

参考文献

[1] Lee, D. and Seung, H.: Learning the parts of objects with non-negative matrix factorization, *Nature*, Vol. 401, pp. 788–791, (1999).

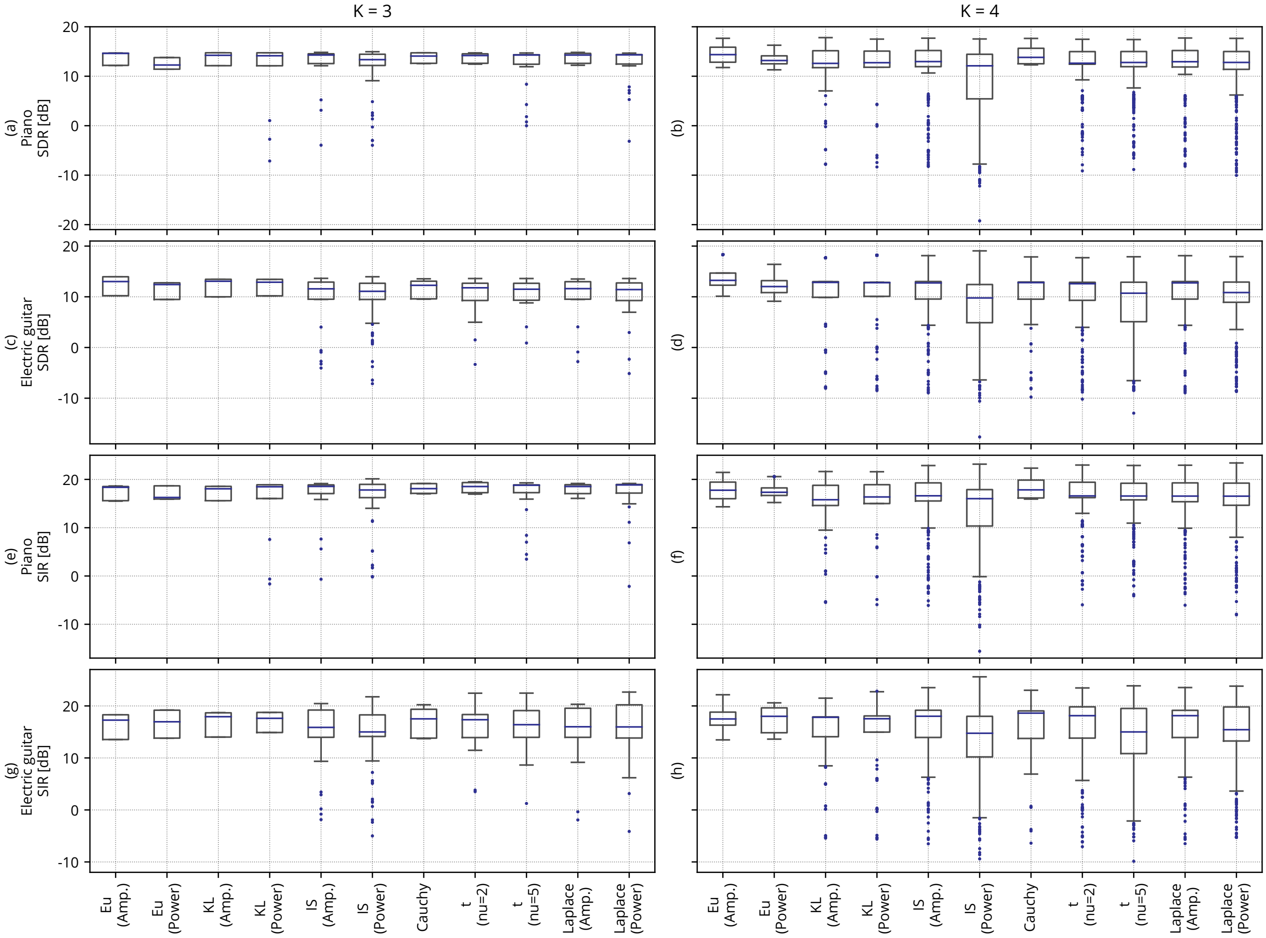


図 2　分離性能の評価結果

Fig. 2　Evaluation results of the separation performance.

[2] Lee, D. and Seung, H.: Algorithms for non-negative matrix factorization, *NIPS 2002*, pp. 556–562, (2002).
[3] Smaragdis, P. and Brown, J.: Non-negative matrix factorization for polyphonic music transcription, *WASPAA 2003*, pp. 177–180, (2003).
[4] Liutkus, A., Fitzgerald, D. and Badeau, R.: Cauchy nonnegative matrix factorization, *WASPAA 2015*, pp. 1–5, (2015).
[5] Yoshii, K., Itoyama, K. and Goto, M.: Student's T nonnegative matrix factorization and positive semidefinite tensor factorization for single-channel audio source separation, *ICASSP 2016*, pp. 51–55, (2016).
[6] Fevotte, C., Bertin, N. and Durrieu, J. L.: Nonnegative matrix factorization with the Itakura-Saito divergence: with application to music analysis, *Neural Computation*, Vol. 21, No. 3, pp. 793–830, (2008).
[7] FitzGerald, D., Cranitch, M. and Coyle, E.: On the use of the beta divergence for musical source separation, *ISSC 2009*, pp. 1–6, (2009).
[8] Smaragdis, P.: Convolutive speech bases and their application to supervised speech separation, *IEEE ASLP*, Vol. 15, No. 1, pp. 1–12, (2007).
[9] Tanji, H., Tanaka, R., Tabata, K., Iseki, Y., Murakami, T. and Ishida, Y.: Derivation of update rules for convolutive NMF based on squared Euclidean distance, KL divergence, and IS divergence, *IEICE Trans. Fundamentals*, Vol. E97-A, No. 11, pp. 2121-2129, (2014).
[10] Ozerov, A. and Fevotte, C.: Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation, *IEEE ASLP*, Vol. 18, No. 3, pp. 550–563, (2010).
[11] Hunter, D. and Lange, K.: A tutorial on MM algorithms, *The American Statistician*, Vol. 58, No. 1, pp. 30–37, (2004).
[12] Nakano, M., Kameoka, H., Le Roux, J., Kitano, Y., Ono, N. and Sagayama, S.: Convergence-guaranteed multiplicative algorithms for nonnegative matrix factorization with β-divergence, *MLSP 2010*, pp. 283–288, (2010).
[13] Martin, R.: Speech enhancement using MMSE short time spectral estimation with gamma distributed speech priors, *ICASSP 2002*, pp. 253–256, (2002).
[14] Lee, B., Kaler, T. and Schafer, R.: Maximum-likelihood sound source localization with a multivariate complex Laplacian distribution, *IWAENC 2008*, (2008).
[15] Kim, T., Attias, H., Lee, S. and Lee, T.: Blind source separation exploiting higher-order frequency dependencies, *IEEE ASLP*, Vol. 15, No. 1, pp. 70–79, (2007).
[16] Fevotte, C. and Cemgil, A.: Nonnegative matrix factorisations as probabilistic inference in composite models, *EUSIPCO 2009*, pp. 1913–1917, (2009).
[17] Goto, M., Hashiguchi, H., Nishimura, T. and Oka, R.: RWC music database: popular, classical, and jazz music databases, *ICMIR 2002*, pp. 287–288, (2002).
[18] Liutkus, A. and Badeau, R.: Generalized Wiener filtering with fractional power spectrograms, *ICASSP 2015*, pp. 266–270, (2015).
[19] Vincent, E., Gribonval, R. and Fevotte, C.: Performance measurement in blind audio source separation, *IEEE ASLP*, Vol. 14, No. 4, pp. 1462–1469, (2006).