

非負値行列因子分解

亀岡弘和^{*,**}

* 東京大学 大学院情報理工学系研究科 東京都文京区本郷 7-3-1
 ** 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 神奈川県厚木市森の里若宮 3-1
 * Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
 ** NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, 3-1 Morinosato Wakamiya, Atsugi, Kanagawa, Japan

* E-mail: kameoka@hil.t.u-tokyo.ac.jp, kameoka.hirokazu@lab.ntt.co.jp
 キーワード：多変量解析 (multivariate analysis), 次元圧縮 (dimensionality reduction), 信号分解 (signal decomposition), 補助関数 (auxiliary function).

JL 0009/12/5109-0835 © 2012 SICE

1. はじめに

実世界には、パワースペクトル、画素値、頻度、個数など、非負値で表わされるデータが多い。主成分分析や独立成分分析などの多変量解析では、所与のデータを複数の加法的な成分に分解することを目的とするが、これと同様に上述のような非負値のデータから構成成分を抽出することが役立つ場面が多い。たとえば、複数の音源の音響信号が混在する多重音のパワースペクトルから個々の音源のパワースペクトルをうまく取り出すことができれば、雑音除去や音源分離などに役立つことができるし、顔画像データを目や鼻などの顔のパーツに該当する画像データにうまく分解することができれば、顔認証や顔画像合成などに役立つことができる。また、文書データ (文書に出現する各単語の個数のデータ) から「時事」や「スポーツ」や「経済」といった潜在的なトピックに該当するような単語ヒストグラムの構成成分をうまく取り出すことができれば、各文書に対し自動インデキシングを行うことが可能となり、文書検索に大いに役立つことができる。以上のように、非負値のデータを加法的な構成成分に分解することを目的とした多変量解析手法を非負値行列因子分解 (Non-negative Matrix Factorization; NMF)¹⁾ といい、さまざまな分野で近年注目を集めている。本稿では、NMF の定式化、基本的な性質、アルゴリズムの導出方法、さまざまな視点からの解釈、おもに音響信号処理の問題に焦点を当てた改良・拡張のアイデアについて解説する。

2. NMF とは

以後、データをベクトルで表記することとする。たとえば、画像データであれば各ピクセルの画素値がベクトル要素となり、パワースペクトルであれば各周波数におけるパワー値がベクトル要素となる。今、 N 個の非負値データベクトル $\mathbf{y}_1, \dots, \mathbf{y}_N \in \mathbb{R}^{\geq 0, K}$ が与えられたとしよう。これらを観測ベクトルと呼ぶ。ただし、 $\mathbb{R}^{\geq 0, K}$ は K 次元の非負値ベクトル全体の集合を表わす。NMF では、これら観測ベクトルが、いずれも M 個の基底ベクトルの適当な重みつき和によって表わされたものと見なし、すべての観測ベクトルを最も良く説明するような M 個の基底ベクトルおよ

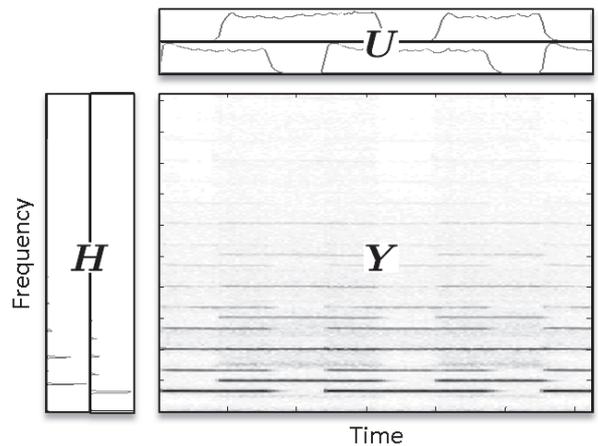


図1 スペクトログラムを行列と見なしてNMFを行った場合の例。図のスペクトログラム \mathbf{Y} を $\mathbf{H}\mathbf{U}$ で表わそうとすることは、各時刻の観測スペクトルを、2種類の基底スペクトルの適当な重みつき和で表わそうとしていることに相当する。

び重み係数を推定することが目的である。よって、NMF では加法性が成り立つ量のみが対象となる。画素値やパワースペクトルはいずれも厳密には加法性が成り立つ量ではないが、NMF を適用する場面では近似的に加法性が成り立つと仮定されているということに注意が必要である。パワースペクトルに対してNMFを適用する上でのパワースペクトルの非加法性についての議論は6.1節で詳しく述べる。以上の加法性に関する仮定の他に、基底ベクトルと重み係数がいずれも非負値であるという仮定が置かれる点もNMFにおける重要なポイントである。すなわち、NMFは、観測ベクトル \mathbf{y}_n を基底ベクトル $\mathbf{h}_1, \dots, \mathbf{h}_M \in \mathbb{R}^{\geq 0, K}$ の非負結合 (結合係数 $u_{1,n}, \dots, u_{M,n}$ が非負値の線形結合)

$$\mathbf{y}_n \simeq \sum_{m=1}^M \mathbf{h}_m u_{m,n} \quad (n = 1, \dots, N) \quad (1)$$

で近似する問題と見なせる。ここで、観測ベクトルを並べた行列を $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N] = (y_{k,n})_{K \times N}$, 基底ベクトルを並べた行列を $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_M] = (h_{k,m})_{K \times M}$, 結合係数 $u_{m,n}$ を m 行 n 列の要素とした行列を $\mathbf{U} = (u_{m,n})_{M \times N}$ とすると、(1) 式は

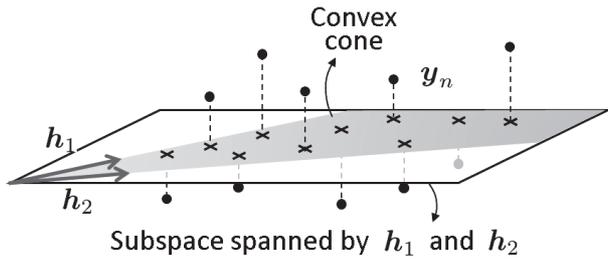


図2 非負値基底の非負結合

$$Y \simeq HU \quad (2)$$

と同じ意味である。このようにNMFは、観測ベクトルを並べた行列を2つの非負値行列の積に分解する問題と捉えることができる。これが、非負値行列因子分解と呼ばれる所以である。NMFによる分解表現のイメージを掴んでもらうために、^(注1)スペクトログラムを行列と見なしてNMFを適用した例を図1に示す。

基底ベクトルの非負性の仮定は、観測データを構成する成分も観測データと同じ属性の物理量であり、それゆえ同様に非負値であるべきという考えに基づいている。たとえば、パワースペクトルを構成する成分もパワースペクトルであるべきであり、負値をとるパワースペクトルは物理的に意味をなさない、という考え方である。

重み係数の非負性の仮定は、次章で詳しく述べるが、観測データ行列 Y から基底行列 H と係数行列 U を得る際に、行列 U の要素がスパース(疎)になるよう誘導する効果をもたらす。係数行列がスパースになるということは、それぞれの観測ベクトルを少数の基底ベクトルだけで表わそうとしていることを意味し、どのような基底ベクトルが得られるかということに大きく関係する。

3. NMFの基本性質

3.1 観測行列の低ランク近似

NMFでは通常、基底数 M は観測ベクトルの次元 K やデータの個数 N より小さく設定される。たとえば $M = K$ の場合、 $H = I$ (ただし、 I は単位行列) であるような分解表現 $Y = HU$ を得ることができるが、明らかにこの分解は意味をなさない。また、 $M = N$ の場合、 $U = I$ であるような分解表現 $Y = HU$ を得ることができるが、この分解からも意味を見出せない。 $M < \min(K, N)$ のとき、NMFは観測行列 Y を低いランクの行列で近似しようとしていることに相当し、その場合の基底行列と係数行列を求めることが重要である。幾何学的には、主成分分析(特異値分解)では観測データが属する部分空間を見つけようとしているのに対し、NMFでは観測データに良くフィットするような凸錐(図2参照)を見つけようとしていると解釈

^(注1)各時刻周辺における信号のスペクトルを時間順に並べて表示したものをスペクトログラムと呼ぶ。

できる。基底数 M は観測データが属する凸錐の次元に対応し、本来は未知であるため、観測データから適切な基底数を決定する問題はNMFにおいて重要課題の一つである。NMFにおける基底数の推定方法については6.4節で紹介する。

3.2 共起成分のグルーピング効果

NMFでは、観測データの中で共起する成分をひとまとめたものが基底ベクトルの推定結果になる傾向がある。たとえば、図3のようにデータが3種類の成分 a, b, c から成っているとき、成分 a, b, c を基底と置けば当然データを完全に表現することは可能である。ここでもし、成分 a や成分 b が生起するときいつも成分 c も揃って生起するようなら、成分 a と成分 c 、および、成分 b と成分 c をひとまとめたものを基底と置いても当該データを同様に表現可能である。この方が、少ない基底でデータを表現可能な分、より「節約的」である。このように、共起する成分をひとまとめてして基底と置いた方が、基底を節約でき、節約した分を他の成分にフィットするのに充てられるようになるわけである。3.1節で述べたように、NMFでは少ない基底で観測データをできるだけ良く表わす必要があるため、上記のような基底が自然に得られる傾向にある。

3.3 係数行列をスパースにする効果

NMFでは、係数行列の非負性の制約により、係数行列の要素がスパース(疎)になる傾向がある。これはベクトルの合成を考えるとイメージしやすい。観測ベクトルを近似するには、観測ベクトルと近い方向を向いた基底ベクトル以外の係数はできるだけ小さくした方が良く、ということである。スパース性は優ガウス性と近い概念であり、それゆえ統計的独立性とも概念的に関連が深い。3.2節で述べた、共起する成分をひとまとめたものを基底にする働きは、なるべく各基底の係数が互いにコヒーレントにならないように(相関をもたないように)基底を決定しようとする働きと見なすことができる。独立成分分析では、係数の互いの独立性を規準として基底を推定するのが目的であるが、NMFでは、非負制約による副次的な効果として係数が互いに独立になるように基底が求まる傾向にある。

3.4 分解の一意性について

一般に、 $X = HU$ となる非負値行列 H と U は一意に決まらない。 HU を、正則行列 D を用いて $H' = HD^{-1}$ と $U' = DU$ の積の形で表わすことができるが、 $H' \neq H$ 、 $U' \neq U$ 、かつ、 H' と U' がいずれも非負値となるような D が存在すれば、 $X = HU$ のような分解は一意に決まらない、ということになる。これを満たす D の中で自明なものは、置換行列、および、要素が正値の対角行列である。前者は、 H と U に対し適当な列置換と行置換を行ったものが H' と U' である場合に相当し、後者は、 H の各列に定数を乗じ、 U の各行にその逆数を乗じたものが H' と U' である場合に相当する。NMFにおいて、後者のようなスケールの任意性を除く目的で、 H の各列ベクトルのノルムが等しくなるような制約が置かれる場合が多い。

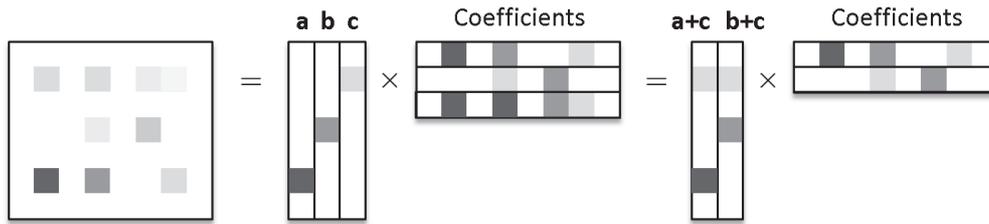


図3 NMFにおいて共起成分がグルーピングされる効果

4. NMF アルゴリズム

4.1 正値行列因子分解と NMF

観測行列を2つの非負値行列の積で表わそうという NMF の基本概念自体は, Paatero らによって最初に提案されている²⁾. Paatero らはこの中で, 誤差行列 $\mathbf{Y} - \mathbf{H}\mathbf{U}$ の^(注2) Frobenius ノルム (行列の要素の二乗和)

$$D_{\text{EU}}(\mathbf{H}, \mathbf{U}) = \|\mathbf{Y} - \mathbf{H}\mathbf{U}\|_F^2 = \sum_{k,n} \left| y_{k,n} - \sum_m h_{k,m} u_{m,n} \right|^2 \quad (3)$$

で $\mathbf{H}\mathbf{U}$ の \mathbf{Y} からの乖離度を定義し, さらに, \mathbf{H} と \mathbf{U} の各行列要素が非負であることを保証する目的で

$$B(\mathbf{H}, \mathbf{U}) = - \sum_{k,m} \log h_{k,m} - \sum_{m,n} \log u_{m,n} \quad (4)$$

なる対数障壁関数 (各行列要素が0になろうとすると無限大のペナルティを課す関数) を定義し, これに適当な係数を掛けて乖離度 $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ に足したものを目的関数とした最適化アルゴリズムを提案している. Paatero らの方法は, 最終的に得られる \mathbf{H} と \mathbf{U} が上記の障壁関数の性質より正値行列に限られることから, 正値行列因子分解 (Positive Matrix Factorization; PMF) と呼ばれる. 一方, 上述のような障壁関数を用いずとも行列要素の非負性を保証しながら $\mathbf{Y} \simeq \mathbf{H}\mathbf{U}$ となる \mathbf{H} と \mathbf{U} を効率的に得る反復アルゴリズムが Lee らによって考案され^{1), 3)}, これをきっかけに NMF が広く知れ渡るようになった. 本章では, Lee らの方法がベースにする補助関数法と呼ぶ最適化手法の原理をまず紹介し, これに基づいて NMF アルゴリズムがどのようにして導かれるかを解説する.

4.2 行列の乖離度

NMF は, $\mathbf{H}\mathbf{U}$ の \mathbf{Y} からの乖離度を表わす規準の定義に応じて異なる最適化問題に帰着する. Lee らは, 前述の二乗誤差規準に加え, ^(注3) I ダイバージェンス⁴⁾ 規準の NMF アルゴリズムを導出している³⁾. 当然ながら異なる乖離度規準を用いた場合には異なる解が最適解として得られる. このため, 最適解が所望の解であるためには, 背後

(注2) 正確には重みつき二乗誤差規準が採用されている²⁾.

(注3) 一般化 Kullback-Leibler ダイバージェンスとも呼ばれる.

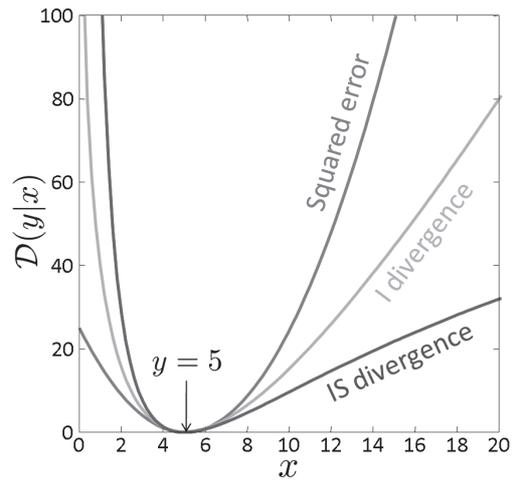


図4 二乗誤差 $D_{\text{EU}}(y|x)$, I ダイバージェンス $D_{\text{KL}}(y|x)$, 板倉斎藤擬距離 $D_{\text{IS}}(y|x)$ を x の関数と見たときのグラフ

にあるデータの生成プロセスに合った適切な規準を設定することが重要である. たとえば, パワースペクトルを対象とした NMF ではしばしば板倉斎藤擬距離⁵⁾ が乖離度規準として用いられる⁶⁾ が, これは信号波形の生成プロセスの仮定から導かれる. この理論的裏付けは 6.2 節 で詳しく述べる.

$y, x \in \mathbb{R}$ とすると, y の x からの二乗誤差, I ダイバージェンス, 板倉斎藤擬距離はそれぞれ

$$D_{\text{EU}}(y|x) = (y - x)^2 \quad (5)$$

$$D_{\text{KL}}(y|x) = y \log \frac{y}{x} - y + x \quad (6)$$

$$D_{\text{IS}}(y|x) = \frac{y}{x} - \log \frac{y}{x} - 1 \quad (7)$$

で与えられる. いずれも, $x = y$ のときにのみ 0 となり, x が y から離れるほど増大する関数である. 図 4 にそれぞれを x の関数と見たときのグラフを示す. 二乗誤差は y を中心に対称であるのに対し, I ダイバージェンスと板倉斎藤擬距離は非対称であり, x が y を下回る場合に, より過大なペナルティを課す関数であることがわかる. また, 板倉斎藤擬距離は y と x の比のみで表わされる関数であるため, y と x のスケールに非依存である (両者を α 倍しても乖離度は等しいと解釈される).

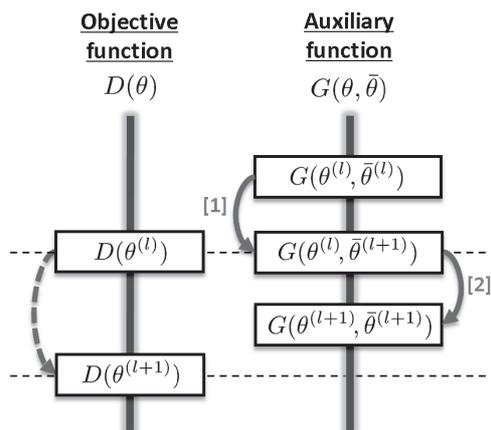


図5 補助関数法の原理

それぞれの規準を採用することがどのようなデータの生成プロセスを背後に仮定したことになるかは5章で詳しく述べることにして、ひとまずこれらを用いれば、 $\mathbf{H}\mathbf{U}$ の \mathbf{Y} からの乖離度を、

$$D(\mathbf{H}, \mathbf{U}) = \sum_{k,n} D(y_{k,n} \mid \sum_m h_{k,m} u_{m,n}) \quad (8)$$

で測ることができる。ただし、 \cdot はEU, KL, ISを表わす。

4.3 補助関数法

これらを最小化する非負値の \mathbf{H} と \mathbf{U} を求めるのが目的であるが、いずれも非負制約つき非線形最適化問題であり、解析的に解くことはできない。もし、目的関数が行列要素 $h_{k,m}, u_{m,n}$ ごとの関数の和に分離した形をしていれば、各行列要素ごとに非負制約の中で解を探索することができて好都合であるが、残念ながらこれらの目的関数ではそうはなっていない。Leeらは、目的関数の上限となる補助関数を反復的に降下させることで目的関数を間接的に降下していく方法をベースにし、その補助関数として行列要素 $h_{k,m}, u_{m,n}$ ごとの関数の和に分離した形をとるものをうまく設計することで、当該制約付き非線形最適化問題の解を見通し良く探索することができると考えた。補助関数の反復降下による目的関数の降下方法を補助関数法と呼び、音響信号処理分野で近年さまざまな最適化問題に適用されている⁷⁾⁻¹³⁾。補助関数の定義と補助関数法の原理は以下のとおりである。

定義 1 $\theta = \{\theta_i\}_{1 \leq i \leq I}$ を変量とする目的関数 $D(\theta)$ に対し、

$$D(\theta) = \min_{\bar{\theta}} G(\theta, \bar{\theta}) \quad (9)$$

が成り立つとき、 $G(\theta, \bar{\theta})$ を $D(\theta)$ の補助関数 (Auxiliary function), $\bar{\theta}$ を補助変数と定義する。

定理 1 (補助関数法) 補助関数 $G(\theta, \bar{\theta})$ を、 $\bar{\theta}$ に関して最小化するステップと、 $\theta_1, \dots, \theta_I$ に関して最小化するステップ

$$\bar{\theta} \leftarrow \operatorname{argmin}_{\bar{\theta}} G(\theta, \bar{\theta}) \quad (10)$$

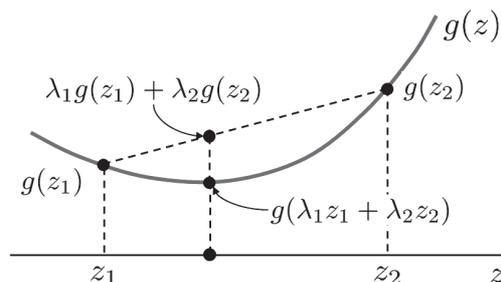


図6 Jensenの不等式

$$\theta_i \leftarrow \operatorname{argmin}_{\theta_i} G(\theta, \bar{\theta}) \quad (i = 1, \dots, I) \quad (11)$$

を繰り返すと、目的関数 $D(\theta)$ の値は単調減少する。

4.4 $D_{\text{EU/KL}}(\mathbf{H}, \mathbf{U})$ を規準とした NMF

上述の補助関数法の原理に基づき、まず Lee らによる $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ を規準とした NMF アルゴリズム³⁾ の導出を以下に示す。 $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ を展開し、 \mathbf{H} と \mathbf{U} に依らない項を省略すると、

$$D_{\text{EU}}(\mathbf{H}, \mathbf{U}) \stackrel{\mathbf{H}, \mathbf{U}}{=} \sum_{k,n} (-2y_{k,n}x_{k,n} + x_{k,n}^2) \quad (12)$$

と書ける。ただし、

$$x_{k,n} = \sum_m h_{k,m} u_{m,n} \quad (13)$$

である。また、 \hat{z} は z に関する項のみに関する等号を表わす。(12)式の中で、 $x_{k,n}^2$ が \mathbf{H} と \mathbf{U} の行列要素 $h_{k,1}, \dots, h_{k,M}, u_{1,n}, \dots, u_{M,n}$ を含んだ非線形関数項であることに気付く。この項に対し、行列要素 $h_{k,m}, u_{m,n}$ ごとの関数の和に分離した形をした上限関数を設けたい。2次関数は凸関数であるため、以下の Jensen の不等式を用いることで、上述のような上限関数が設計可能である。

定理 2 (Jensen の不等式 (図6)) 任意の凸関数 g , I 個の実数 x_1, \dots, x_I , $\sum_i \lambda_i = 1$ を満たす I 個の正値の重み係数 $\lambda_1, \dots, \lambda_I$ のもつて、

$$g\left(\sum_i \lambda_i z_i\right) \leq \sum_i \lambda_i g(z_i) \quad (14)$$

が成り立ち、 $z_1 = \dots = z_I$ のとき等号が成立する。

これを用いると、 $x_{k,n}^2$ に対し、

$$\begin{aligned} x_{k,n}^2 &= \left(\sum_m \lambda_{k,m,n} \frac{h_{k,m} u_{m,n}}{\lambda_{k,m,n}}\right)^2 \\ &\leq \sum_m \lambda_{k,m,n} \left(\frac{h_{k,m} u_{m,n}}{\lambda_{k,m,n}}\right)^2 \end{aligned} \quad (15)$$

のような不等式を立てることができる。ただし、 $\lambda_{k,m,n} > 0$, $\sum_m \lambda_{k,m,n} = 1$ である。ここで重要なのは、右辺が行列要

素 $h_{k,m}, u_{m,n}$ ごとの関数の和に分離した形をしているということである。また、もう一つ重要なポイントは、右辺を $\lambda_{k,m,n}$ に関して最小化したとき等号が成立するということであり、等号が成立するのは $\frac{h_{k,1}u_{1,n}}{\lambda_{k,1,n}} = \dots = \frac{h_{k,M}u_{M,n}}{\lambda_{k,M,n}}$, すなわち、 $\lambda_{k,m,n}$ が

$$\lambda_{k,m,n} = \frac{h_{k,m}u_{m,n}}{x_{k,n}} \quad (16)$$

のときである。

以上より、 $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ における $x_{k,n}^2$ の項を (15) 式の右辺に置き換えることにより得られる関数

$$G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}) = \sum_{k,n} \left(y_{k,n}^2 - 2y_{k,n} \sum_m h_{k,m} u_{m,n} + \sum_m \frac{h_{k,m}^2 u_{m,n}^2}{\lambda_{k,m,n}} \right) \quad (17)$$

はそれぞれ $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ の補助関数の要件を満たす。ただし、 $\boldsymbol{\lambda} = \{\lambda_{k,m,n}\}_{K \times M \times N}$ を表わすものとする。 $G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda})$ が得られれば、あとは定理 1 に従い、

$$\boldsymbol{\lambda} \leftarrow \underset{\boldsymbol{\lambda}}{\text{argmin}} G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}) \quad (18)$$

$$\mathbf{H} \leftarrow \underset{\mathbf{H}}{\text{argmin}} G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}) \quad (19)$$

$$\mathbf{U} \leftarrow \underset{\mathbf{U}}{\text{argmin}} G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}) \quad (20)$$

を反復的に行えば $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ を単調に降下させていくことができる。(18)式についてはすでに述べたとおり (16)式で与えられる。(19)式と(20)式は $h_{k,m} \geq 0, u_{m,n} \geq 0$ の制約のもとで解かれなければならない点に注意が必要だが、ありがたいことに $G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda})$ は行列要素ごとの関数の和に分かれた形をしているので、これらはいずれも行列要素ごとの一変数関数の非負制約つき最小化問題に帰着し、容易に解くことができる。たとえば、 $G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda})$ は $h_{k,m}$ ごとに見れば二次関数であり、これを最小化する $\hat{h}_{k,m}$ は

$$\hat{h}_{k,m} = \frac{\sum_n y_{k,n} u_{m,n}}{\sum_n u_{m,n}^2 / \lambda_{k,m,n}} \quad (21)$$

のように解析的に求まる。この値がもし負であれば $h_{k,m} \geq 0$ の制約下で $G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda})$ が最小となるのは明らかに $h_{k,m} = 0$ のときである。また、 $u_{m,n}$ については

$$\hat{u}_{m,n} = \frac{\sum_k y_{k,n} h_{k,m}}{\sum_k h_{k,m}^2 / \lambda_{k,m,n}} \quad (22)$$

のとき $G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda})$ は最小となる。よって、(19)式と(20)式は、それぞれ $h_{k,m} = \max\{\hat{h}_{k,m}, 0\}$, $u_{m,n} =$

$\max\{\hat{u}_{m,n}, 0\}$ で与えられる。ところで、(21), (22)式は、 $\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}$ の要素がすべて非負であればいずれも非負値となるため、 \mathbf{H} と \mathbf{U} の初期値が非負であることが前提となっている場合は $h_{k,m} = \hat{h}_{k,m}$, $u_{m,n} = \hat{u}_{m,n}$ である。

以上より、(18)式を(19), (20)式に代入すると、以下の NMF アルゴリズムが導かれる。

二乗誤差規準の NMF アルゴリズム

1. \mathbf{H}, \mathbf{U} を非負値行列に初期設定する。
2. 以下の更新を収束するまで繰り返す。

$$h_{k,m} \leftarrow h_{k,m} \frac{\sum_n y_{k,n} u_{m,n}}{\sum_n x_{k,n} u_{m,n}}$$

$$u_{m,n} \leftarrow u_{m,n} \frac{\sum_k y_{k,n} h_{k,m}}{\sum_k x_{k,n} h_{k,m}}$$

$D_{\text{KL}}(\mathbf{H}, \mathbf{U})$ を規準とした NMF アルゴリズムも以上と同様にして導かれる。 $D_{\text{KL}}(\mathbf{H}, \mathbf{U})$ は、

$$D_{\text{KL}}(\mathbf{H}, \mathbf{U}) \stackrel{\mathbf{H}, \mathbf{U}}{=} \sum_{k,n} (-y_{k,n} \log x_{k,n} + x_{k,n}) \quad (23)$$

と書けるが、この式の中で $-y_{k,n} \log x_{k,n}$ が $h_{k,1}, \dots, h_{k,M}, u_{1,n}, \dots, u_{M,n}$ を含んだ非線形関数項である。この項に関し、負の対数関数の凸性を利用して Jensen の不等式

$$-\log x_{k,n} \leq -\sum_m \lambda_{k,m,n} \log \left(\frac{h_{k,m} u_{m,n}}{\lambda_{k,m,n}} \right) \quad (24)$$

が立てられ、これにより行列要素 $h_{k,m}, u_{m,n}$ ごとの関数の和に分離した形の補助関数を設けることができる。これを用いて以下の NMF アルゴリズムを導くことができる。

I ダイバージェンス規準の NMF アルゴリズム

1. \mathbf{H}, \mathbf{U} を非負値行列に初期設定する。
2. 以下の更新を収束するまで繰り返す。

$$h_{k,m} \leftarrow h_{k,m} \frac{\sum_n y_{k,n} u_{m,n} / x_{k,n}}{\sum_n u_{m,n}}$$

$$u_{m,n} \leftarrow u_{m,n} \frac{\sum_k y_{k,n} h_{k,m} / x_{k,n}}{\sum_k h_{k,m}}$$

4.5 $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ を規準とした NMF

つぎに、著者らによる、板倉斎藤擬距離 $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ を

規準とした NMF アルゴリズムの導出⁷⁾を以下に示す。 \mathbf{H} , \mathbf{U} に依らない項を省略すると, $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ は

$$D_{\text{IS}}(\mathbf{H}, \mathbf{U}) \stackrel{H, U}{=} \sum_{k, n} \left(\frac{y_{k, n}}{x_{k, n}} + \log x_{k, n} \right) \quad (25)$$

と書ける。先と同様の方針で補助関数法により $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ を目的関数とした非負制約つき最適化問題を解決するためには, $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ に対し, 行列要素 $h_{k, m}, u_{m, n}$ ごとの関数の和に分離した形をした上限関数を設計できるかどうか⁸⁾がやはり鍵となる。まず, 逆数関数が正領域において凸関数であることに注意すると, 第一項の $1/x_{k, n}$ に関しては, Jensen の不等式

$$\begin{aligned} \frac{1}{x_{k, n}} &= 1 / \left(\sum_m \lambda_{k, m, n} \frac{h_{k, m} u_{m, n}}{\lambda_{k, m, n}} \right) \\ &\leq \sum_m \lambda_{k, m, n} \left(1 / \frac{h_{k, m} u_{m, n}}{\lambda_{k, m, n}} \right) \end{aligned} \quad (26)$$

が成り立つ。ただし, 先と同様, $\lambda_{k, m, n} > 0, \sum_m \lambda_{k, m, n} = 1$ である。つぎに, 第二項の $\log x_{k, n}$ の上限関数を設計することが目標であるが, 第一項と異なり, 正の対数関数は凹関数なので, Jensen の不等式では上限関数が作れない。しかしながら, 任意の微分可能な凹関数 g に対し,

$$g(x) \leq g(\alpha) + (x - \alpha)g'(\alpha) \quad (27)$$

(等号成立は $x = \alpha$) が成り立つことを利用すると, 第二項の $\log x_{k, n}$ に対し,

$$\log x_{k, n} \leq \log \alpha_{k, n} + (x_{k, n} - \alpha_{k, n}) / \alpha_{k, n} \quad (28)$$

なる不等式が立てられる。ここでやはり重要になるのは, いずれの不等式も右辺が行列要素 $h_{k, m}, u_{m, n}$ ごとの関数の和に分離しているということ, また, $\lambda_{k, m, n}$ と $\alpha_{k, n}$ が

$$\lambda_{k, m, n} = \frac{h_{k, m} u_{m, n}}{x_{k, n}} \quad (29)$$

$$\alpha_{k, n} = x_{k, n} \quad (30)$$

のとき等号成立するということである。

以上より, $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ における $1/x_{k, n}$ の項を (26) 式の右辺に, $\log x_{k, n}$ の項を (28) 式の右辺に置き換えることにより得られる関数

$$\begin{aligned} G_{\text{IS}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) &= \sum_{k, n} \left(\sum_m \frac{y_{k, n} \lambda_{k, m, n}^2}{h_{k, m} u_{m, n}} + \sum_m \frac{h_{k, m} u_{m, n}}{\alpha_{k, n}} \right. \\ &\quad \left. - \log y_{k, n} + \log \alpha_{k, n} - 2 \right) \end{aligned} \quad (31)$$

は $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ の補助関数の要件を満たす。補助関数がひとたび得られれば, あとは定理 1 に従い,

$$\boldsymbol{\lambda}, \boldsymbol{\alpha} \leftarrow \underset{\boldsymbol{\lambda}, \boldsymbol{\alpha}}{\operatorname{argmin}} G_{\text{IS}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) \quad (32)$$

$$\mathbf{H} \leftarrow \underset{\mathbf{H}}{\operatorname{argmin}} G_{\text{IS}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) \quad (33)$$

$$\mathbf{U} \leftarrow \underset{\mathbf{U}}{\operatorname{argmin}} G_{\text{IS}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) \quad (34)$$

を反復的に行えば $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ を単調に減少させることができる。(32) 式は (29), (30) 式のとおりであり, あとは先と同様に \mathbf{H} と \mathbf{U} の更新式を求めれば, 以下の板倉斎藤擬距離規準の NMF アルゴリズムが導かれる。

板倉斎藤擬距離規準の NMF アルゴリズム

1. \mathbf{H}, \mathbf{U} を非負値行列に初期設定する。
2. 以下の更新を収束するまで繰り返す。

$$\begin{aligned} h_{k, m} &\leftarrow h_{k, m} \left(\frac{\sum_n y_{k, n} u_{m, n} / x_{k, n}^2}{\sum_n u_{m, n} / x_{k, n}} \right)^{1/2} \\ u_{m, n} &\leftarrow u_{m, n} \left(\frac{\sum_k y_{k, n} h_{k, m} / x_{k, n}^2}{\sum_k h_{k, m} / x_{k, n}} \right)^{1/2} \end{aligned}$$

4.6 $D_{\beta}(\mathbf{H}, \mathbf{U})$ を規準とした NMF

(5)~(7) 式に 3 種類の乖離度規準を挙げたが, これらの乖離度規準をパラメータ β で統一的に記述した β ダイバージェンスと呼ぶ規準¹⁴⁾

$$\mathcal{D}_{\beta}(y|x) = y \frac{y^{\beta-1} - x^{\beta-1}}{\beta-1} - \frac{y^{\beta} - x^{\beta}}{\beta} \quad (35)$$

がある。ただし, β は $\beta \neq 0, \beta \neq 1$ の実数である。 $\lim_{\beta \rightarrow 0} (x^{\beta} - y^{\beta}) / \beta = \log(x/y)$ であることを利用すれば, (35) 式は $\beta \rightarrow 0$ のとき板倉斎藤擬距離, $\beta \rightarrow 1$ のとき I ダイバージェンス, $\beta = 2$ のとき二乗誤差となることが確認できる。これを (8) 式のように \mathbf{H}, \mathbf{U} の \mathbf{Y} からの乖離度の規準 $D_{\beta}(\mathbf{H}, \mathbf{U})$ とした NMF アルゴリズムが Kompass らや中野らによって提案されている^{15), 16)}。Kompass らは $1 \leq \beta \leq 2$ においてのみ, 中野らはすべての β において収束性が保証される NMF アルゴリズムを導出しており, 以下に中野らによる導出¹⁶⁾の概略を示す。

(35) 式の第一項の $(y^{\beta-1} - x^{\beta-1}) / (\beta-1)$ は, x の関数と見れば $\beta \leq 2$ のとき凸関数, $\beta \geq 2$ のとき凹関数となる。一方, 第二項の $-(y^{\beta} - x^{\beta}) / \beta$ は $\beta \leq 1$ のとき凹関数, $\beta \geq 1$ のとき凸関数となる。前述の著者らのアイデア⁷⁾と同様, β の条件に応じて凸関数の項に対しては (14) 式を, 凹関数の項に対しては (27) 式の不等式を適用することで $D_{\beta}(\mathbf{H}, \mathbf{U})$ の補助関数を設計することができる。詳細は省略するが, あとは前述の板倉斎藤擬距離規準の NMF アルゴリズムの導出に従えば, 以下の β ダイバージェンス規準の NMF アルゴリズムが導かれる。

β ダイバージェンス規準の NMF アルゴリズム

1. \mathbf{H} , \mathbf{U} を非負値行列に初期設定する.
2. 以下の更新を収束するまで繰り返す.

$$h_{k,m} \leftarrow h_{k,m} \left(\frac{\sum_n y_{k,n} x_{k,n}^{\beta-2} u_{m,n}}{\sum_n x_{k,n}^{\beta-1} u_{m,n}} \right)^{\varphi(\beta)}$$

$$u_{m,n} \leftarrow u_{m,n} \left(\frac{\sum_k y_{k,n} x_{k,n}^{\beta-2} h_{k,m}}{\sum_k x_{k,n}^{\beta-1} h_{k,m}} \right)^{\varphi(\beta)}$$

$$\varphi(\beta) = \begin{cases} 1/(2-\beta) & (\beta < 1) \\ 1 & (1 \leq \beta \leq 2) \\ 1/(\beta-1) & (\beta > 2) \end{cases}$$

以上のアルゴリズムは、 $\beta = 0, 1, 2$ のとき、前述の板倉斎藤距離規準、I ダイバージェンス規準、二乗誤差規準の NMF アルゴリズムと等価である。

5. 観測行列の生成プロセス

以上のような規準で NMF を行うことは、観測データの背後にどのような生成プロセスを仮定したことになるだろうか。実は、二乗誤差、I ダイバージェンス、板倉斎藤距離、 β ダイバージェンスを規準とした NMF は、観測行列 \mathbf{Y} の各行列要素 $y_{k,n}$ が、 $x_{k,n}$ を平均とした正規分布、Poisson 分布、指数分布、Tweedie 分布

$$y_{k,n} \sim \text{Normal}(y_{k,n}; x_{k,n}, \sigma^2) \quad (36)$$

$$y_{k,n} \sim \text{Poisson}(y_{k,n}; x_{k,n}) \quad (37)$$

$$y_{k,n} \sim \text{Exponential}(y_{k,n}; x_{k,n}) \quad (38)$$

$$y_{k,n} \sim \text{Tweedie}(y_{k,n}; x_{k,n}, \phi) \quad (39)$$

に従って独立に生成されたと仮定した場合の \mathbf{H} , \mathbf{U} の最尤推定問題とおのおの等価である。ただし、

$$\text{Normal}(z; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/2\sigma^2} \quad (40)$$

$$\text{Poisson}(z; \mu) = \frac{\mu^z e^{-\mu}}{z!} \quad (z \geq 0) \quad (41)$$

$$\text{Exponential}(z; \mu) = \frac{1}{\mu} e^{-z/\mu} \quad (z \geq 0) \quad (42)$$

$$\text{Tweedie}(z; \mu, \phi) = a(z, \phi) e^{\frac{1}{\phi}(z\rho(\mu) - \kappa(\mu))} \quad (43)$$

$$\rho(\mu) = \begin{cases} \frac{\mu^{\beta-1}-1}{\beta-1} & (\beta \neq 1) \\ \log \mu & (\beta = 1) \end{cases}, \quad \kappa(\mu) = \begin{cases} \frac{\mu^{\beta}-1}{\beta} & (\beta \neq 0) \\ \log \mu & (\beta = 0) \end{cases}$$

である。このことは、以下により確かめられる。(36)~(39) 式により立てられる $x_{k,n}$ の対数尤度 $L(x_{k,n}) =$

$\log p(y_{k,n}|x_{k,n})$ はいずれも $x_{k,n} = y_{k,n}$ のときに最大となる。すなわち、 $L(y_{k,n}) \geq L(x_{k,n})$ である。よって、対数尤度差 $L(y_{k,n}) - L(x_{k,n})$ は、 $x_{k,n} = y_{k,n}$ のときのみ 0 になる、 $y_{k,n}$ と $x_{k,n}$ の近さを表わす非負の尺度と見なすことができる。実は、(36)~(39) 式の場合の対数尤度差 $L(y_{k,n}) - L(x_{k,n})$ は、それぞれ (5)~(7) 式、(35) 式と等しい。

6. 音響信号処理への応用

(注4) モノラル音源分離のアプローチとして多重音信号のスペクトログラム (振幅スペクトログラムまたはパワースペクトログラム) を観測行列と見なして NMF を適用する方法^{17), 18)} が Smaragdis らと Virtanen によって提案され、それ以降 NMF は、自動採譜¹⁷⁾、音声強調・分離¹⁹⁾、帯域拡張²⁰⁾ などさまざまな問題への有効なアプローチとして近年応用展開されている。

瞬時瞬時の観測スペクトルが構成音のスペクトルの重みつき和によって表される (スペクトルは加法的である)、という仮定と、各構成音は周波数成分比が時不変でゲインのみが時間変化する、という仮定の下で、観測スペクトログラムから個々の構成音のスペクトル (周波数成分比) と各時刻におけるゲインを推定する問題は、NMF と同形の問題と見なすことができる。たとえば、音楽は平均律音階に該当する限られた種類の音高の音で構成されるため、各音高の音が時刻に依らずいつも同じ周波数成分比を有すると仮定できる場合、音楽音響信号のスペクトログラムに NMF を適用すると各基底ベクトルが一つの音階の音のスペクトルとなり、結合係数が各時刻におけるそれぞれの音階の音のゲインとなるように得られる。以上のアプローチの有効性は経験的に広く知られているが、上述の、核となる 2 つの仮定は実際には成り立たない。本章では、これらの仮定をより現実的に即したものに改変した改良モデルを紹介する。

6.1 複素 NMF²¹⁾

音の信号波形は加法的である。短時間 Fourier 変換やウェーブレット変換などの時間周波数分解は基本的には信号波形の線形変換であるため、時間周波数分解により得られる複素スペクトログラムも加法的である。しかし、複素スペクトログラムから振幅 (またはパワー) スペクトログラムへの変換は非線形であるため、振幅スペクトログラムもパワースペクトログラムも実際には非加法的である。端的に言えば、波形 A と波形 B の和のスペクトルは波形 A のスペクトルと波形 B のスペクトルの和とは等しくなると限らない。したがって、振幅 (ないしパワー) スペクトログラムを加法的な成分に分解したところで、それぞれの成分が物

(注4) 異なる位置に配置された複数のマイクロホンで取得した観測信号から個々の音源信号を分離することを目的とした多チャンネル音源分離では、音源の空間的な位置の違いを手がかりにすることができるが、モノラル信号からの音源分離は、これを手がかりにすることができないため、より難しい不良設定問題となる。

理的に何に対応するかは必ずしも定かではない。

NMF をスペクトログラムに適用するアプローチではスペクトログラムが行列の積でモデル化されるが、このモデルでは、各構成音の周波数成分比は時不変という仮定が置かれていることになる。つまり、周波数成分比さえ同じであれば波形(ゲインや位相スペクトル)が異なっていたとしても同一音と見なそう、という考え方がベースとなっている。著者らはこの考え方を基に複素スペクトログラムをモデル化し、観測複素スペクトログラムを加法的な成分に分解する「複素 NMF」と呼ぶ方法を考案した²¹⁾。以下、その原理について概説する。

複素スペクトログラムの加法的性より、 M 個の音から成る多重音の複素スペクトログラムは

$$x_{k,n} = \sum_{m=1}^M s_{k,n}^{(m)} \quad (44)$$

のように各構成音の複素スペクトログラム $s_{k,n}^{(m)}$ の和で表わすことができる。各構成音の周波数成分比が時不変と仮定することは、構成音の振幅スペクトログラム $|s_{k,n}^{(m)}|$ を

$$|s_{k,n}^{(m)}| = h_{k,m} u_{m,n} \quad (45)$$

と表わすことと等しく、これを(44)式に代入すると、多重音の複素スペクトログラムのモデル

$$x_{k,n} = \sum_m h_{k,m} u_{m,n} e^{j\phi_{k,n}^{(m)}} \quad (46)$$

が得られる。ただし、 j は虚数単位であり、 $\phi_{k,n}^{(m)} = \arg(s_{k,n}^{(m)})$ である。以上のモデルを観測複素スペクトログラム $y_{k,n}$ (本節における $y_{k,n}$ は、5章までの $y_{k,n}$ と異なり複素数であることを注意されたい。) にフィットするように未知パラメータ $h_{k,m}$, $u_{m,n}$, $\phi_{k,n}^{(m)}$ を推定することが目的であるが、NMF と同様に各基底スペクトル $h_{1,m}, \dots, h_{K,m}$ が独立に生起する個々の音のスペクトルに対応するようになるためには係数 $u_{m,n}$ がスパースになるかどうかは鍵となる。したがって、以上のモデルのもとで、できるだけ係数 $u_{m,n}$ をスパースに、かつ、 $x_{k,n}$ が $y_{k,n}$ にフィットするような $h_{k,m}$, $u_{m,n}$, $\phi_{k,n}^{(m)}$ を推定することができれば、NMF と同様の基底スペクトル獲得機能を実現できる可能性がある。ただし、3.3 節で述べたように NMF では非負制約の副次的効果により係数 $u_{m,n}$ が自然にスパース化される傾向があったのに対し、上記のモデルでは NMF とは違い実部と虚部における構成成分同士の減算を許すため、そのままでは係数 $u_{m,n}$ がスパース化される効果は望めないことに注意が必要である。著者らは、 $y_{k,n}$ と $x_{k,n}$ の二乗誤差に係数 $u_{m,n}$ に関するスパース正則化項を付加した

$$D(\mathbf{H}, \mathbf{U}, \phi) = \sum_{k,n} |y_{k,n} - x_{k,n}|^2 + 2\eta \sum_{k,t} |u_{m,n}|^\gamma \quad (47)$$

という目的関数を最小化する問題を考え、これを局所的に最小化する反復アルゴリズムを補助関数法に基づき導出した。ただし、 $0 < \gamma < 2$, $\sum_k h_{k,m} = 1$ とする。導出の詳細や実験結果についてはこちらの文献²¹⁾を参照されたい。

6.2 板倉齋藤擬距離規準の NMF⁶⁾

先述のとおりパワースペクトル自体は加法的ではないが、各構成音の信号が互いに統計的に独立のときそれぞれのパワースペクトルの期待値は加法的となる。これは、NMF において置かれるスペクトルの加法的性の仮定が、期待値の意味では正当化される場合があるということを示唆している。

平均が 0 の定常 Gauss 過程に従う信号は、Fourier 変換領域において各周波数成分が独立に 0 平均の複素正規分布に従う。したがって、多重音信号中の各構成音の信号が短時間区間ごとに平均が 0 の(巡回)定常 Gauss 過程に従って生成されたと仮定すると、各区間における離散 Fourier 変換の各周波数成分は平均が 0 の複素正規分布に従う。すなわち、区間 n における m 番目の構成音の周波数 k の成分(構成音 m の複素スペクトログラム)を $s_{k,n}^{(m)}$ とすると、 $s_{k,n}^{(m)}$ は

$$s_{k,n}^{(m)} \sim \text{Normal}_{\mathbb{C}}(s_{k,n}^{(m)}; 0, \nu_{k,n}^{(m)}) \quad (48)$$

に従う。ただし、 $\text{Normal}_{\mathbb{C}}(z; \mu, \nu) = (1/\pi\nu)e^{-|z-\mu|^2/\nu}$ である。 $\nu_{k,n}^{(m)}$ は構成音 m のパワースペクトログラムの期待値 $\mathbb{E}[|s_{k,n}^{(m)}|^2]$ を表わすパラメータである。ここで、観測信号の複素スペクトログラム $y_{k,n}$ が

$$y_{k,n} = \sum_m s_{k,n}^{(m)} \quad (49)$$

のように構成音の複素スペクトログラムの和で与えられ、 $s_{k,n}^{(m)}$ と $s_{k,n}^{(m')}$ ($m \neq m'$) が互いに独立と仮定できるなら、 $y_{k,n}$ は(48)、(49)式より

$$y_{k,n} \sim \text{Normal}_{\mathbb{C}}\left(y_{k,n}; 0, \sum_m \nu_{k,n}^{(m)}\right) \quad (50)$$

に従う。よって、 $x_{k,n} = \sum_m \nu_{k,n}^{(m)}$ と置くと、 $y_{k,n}$ が与えられたもとの $x_{k,n}$ の対数尤度は

$$L(x_{k,n}) = -\log \pi x_{k,n} - \frac{|y_{k,n}|^2}{x_{k,n}} \quad (51)$$

と書かれる。この対数尤度は $x_{k,n} = |y_{k,n}|^2$ のとき最大となるので $L(|y_{k,n}|^2) \geq L(x_{k,n})$ である。実は、対数尤度差 $L(|y_{k,n}|^2) - L(x_{k,n}) \geq 0$ は、 $|y_{k,n}|^2$ と $x_{k,n}$ の板倉齋藤擬距離 $D_{\text{IS}}(|y_{k,n}|^2 | x_{k,n})$ と等しい。従って、各構成音のパワースペクトログラムの期待値を表わす $\nu_{k,n}^{(m)}$ に関し、周波数成分比が時不変であるような構造

$$\nu_{k,n}^{(m)} = h_{k,m} u_{m,n} \quad (52)$$

を仮定すれば、 $\mathbf{H} = (h_{k,m})_{K \times M}$ と $\mathbf{U} = (u_{m,n})_{M \times N}$ の最尤推定問題は、観測パワースペクトログラム $|y_{k,n}|^2$ を

要素にもつ行列 \mathbf{Y} に対し板倉斎藤擬距離標準の NMF を行うことと等価となる⁶⁾ (4章における $y_{k,n}$ は本節における $|y_{k,n}|^2$ に対応していることに注意されたい). 以上のような生成モデルの見地に立てば, $s_{k,n}^{(m)}$ を潜在変数と見なして Expectation-Maximization (EM) アルゴリズムの考え方に従って \mathbf{H} と \mathbf{U} の最尤解を局所探索するアルゴリズムを導くことができる⁶⁾ が, 4章で示したアルゴリズムの方が収束が早いことが実験的に確認されている¹⁶⁾.

6.3 NMFD²²⁾・可変基底 NMF^{24)~27)}

NMF をスペクトログラムに適用するアプローチでは基本的に, 分解したい各構成音の周波数成分比は時不変であるという仮定が置かれるが, 音楽音響信号を楽音ごとに分解する問題を扱う上では必ずしもこの仮定は成り立たない. 歌声や音声はもちろん, ピアノやヴァイオリンの各音符に対応する音響信号のスペクトルは時々刻々と変化する. このように同一音源の音が持続していたとしてもスペクトルが時間変化する場合, 通常の NMF を適用しても, 時刻ごとのスペクトルが別の音としてばらばらに分解されてしまう. そこで, 観測スペクトログラムを各楽音の持続音ごとに分解することを目的として, 楽音スペクトルの時変構造を取り入れた NMF の拡張モデルが提案されている.

各楽音の振幅 (またはパワー) スペクトログラム $s_{k,n}^{(m)}$ を時不変なスペクトルに時変のゲインを乗じたものとしてモデル化する代わりに, ある時間幅をもったスペクトログラムに時変なゲインを畳み込んだもの

$$s_{k,n}^{(m)} = \sum_{n'=0}^{N'-1} h_{k,m,n-n'} u_{m,n'} \quad (53)$$

としてモデル化し, 観測振幅 (またはパワー) スペクトログラム $y_{k,n}$ が与えられたもとの $y_{k,n} \simeq \sum_m s_{k,n}^{(m)}$ となるように $\mathbf{H} = (h_{k,m,n'})_{K \times M \times N'}$ と $\mathbf{U} = (u_{m,n})_{M \times N}$ を推定する, Non-negative Matrix Factor Deconvolution (NMFD) と呼ぶ方法が提案されている²²⁾. また, 複素版も提案されており, 音声と背景音楽を分離する技術に応用されている²³⁾.

NMFD では, 各楽音スペクトルの時間変化の仕方ほどの発音時においても同じである, という仮定が暗に置かれるが, この仮定も必ずしも実際には成り立たない. たとえばビブラートのレート (揺れの周期) や深さは発音のたびに変わるかもしれない. そこで, 各楽音スペクトルの変化のタイミングはいつも同じとは限らないが, 変化の順序だけは保たれている, という仮定のもと, 基底スペクトルが内部状態の遷移に応じて時間変化するよう改良された NMF の拡張モデルが近年提案されている^{24)~27)}. このモデルでは, M 個の楽音スペクトログラムがそれぞれ隠れマルコフモデル (Hidden Markov Model; HMM) からの出力系列と見なされ, 観測スペクトログラムはそれらの重ね合わせと解釈される. このモデルの下で観測スペクトログラムを楽音ごとに分離する問題は, 観測スペクトログラムを最も良

く説明する, 各 HMM の状態ごとの出力スペクトル, 状態遷移系列, 時変ゲインを推定する問題として定式化される. なお, このモデルは, 生成モデルの観点で見れば Factorial HMM²⁸⁾ の一種と見なすことができる.

6.4 その他の拡張モデル・他手法との関連

NMF において基底数をいかに決定するかは応用と基礎の両面で重要課題の一つである. Schmidt らや Cemgil らは, 5章で述べたような NMF の生成モデルとしての解釈を通して, NMF の基底数決定問題を周辺尤度 (\mathbf{H} と \mathbf{U} に関して周辺化したデータ行列の生成確率 $p(\mathbf{Y})$) に基づくモデル選択の問題として定式化している^{29), 30)}. モデル選択により NMF の基底数を決定するためには, さまざまな基底数のもとで NMF を実行し, 情報量規準や周辺尤度を算出して比較する手続きが必要となる. これに対し, パラメータの推論と同時にモデルの複雑度の推論を行える枠組であるノンパラトリックベイズアプローチ³¹⁾により, NMF の基底数を推論する方法も提案されている^{27), 32)~34)}.

NMF をスペクトログラムに適用するアプローチは, モノラル信号の音源分離のための有効なアプローチであるため, そのモデル化の考え方は多チャンネル信号の音源分離にも効果的であろうという期待から, NMF の多チャンネル拡張に関する検討が近年進められている^{35)~39)}.

ところで, 文書データのクラスタリングを目的として開発された probabilistic Latent Semantic Analysis (pLSA)⁴⁰⁾ は, NMF と類似したアイディアに基づくデータ解析手法である. 両者の関連については紙面の都合上割愛せざるをえなかったが, 実は pLSA は I ダイバージェンス規準の NMF と \mathbf{H} と \mathbf{U} を正規化するかどうかの違いを除けばアルゴリズム的に等価である.

7. まとめ

NMF は, 音響信号処理やデータマイニングや画像処理や機械学習などの分野で市民権を得始めているデータ解析ツールであり, 本稿では, NMF の基本性質, アルゴリズムの導出方法, 生成モデルとしての解釈, 音響信号処理への応用とそのための拡張モデルについて解説した. NMF に関する解説記事は澤田氏のもの⁴¹⁾ もあるのでより理解を深めたい読者はぜひとも参照していただきたい.

(2012年7月27日受付)

参考文献

- 1) D.D. Lee and H.S. Seung: Learning the parts of objects with nonnegative matrix factorization, *Nature*, **401**, 788/791 (1999)
- 2) P. Paatero and U. Tapper: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics*, **5**, 111/126 (1994)
- 3) D.D. Lee and H.S. Seung: Algorithms for nonnegative matrix factorization, in *Adv. NIPS*, 556/562 (2000)
- 4) I. Csiszár: I-divergence geometry of probability distributions and minimization problems, *The Annals of Probability*

