

非負値行列因子分解とその音響信号処理への応用

亀岡 弘和^{*†}

Non-negative Matrix Factorization and Its Variants with Applications to Audio Signal Processing

Hirokazu Kameoka^{*†}

実世界にはパワースペクトル, 画素値, 頻度など, 非負値で表されるデータが多い. このように非負値で表されたデータを有意な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) といい, 近年音響信号処理分野をはじめ様々な分野で注目を集めている. 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拡張モデルについて解説する.

In this paper, I will give a brief introduction to a data analysis technique called non-negative matrix factorization (NMF), which has attracted a lot of attention in the field of audio signal processing in recent years. I will mention some basic properties of NMF, effects induced by the non-negative constraints, how to derive an iterative algorithm for NMF, and some attempts that have been made to apply NMF to audio processing problems.

キーワード: 行列分解, 非負制約, スパース性, 補助関数法, 音響信号処理

1. はじめに

実世界には, パワースペクトル, 画素値, 頻度など, 非負値で表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. 例えば, 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「政治」

* 東京大学大学院情報理工学系研究科.

† 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.

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てることができる。同様に、購買ログから消費者の潜在的な購買欲求に該当する購買パターンを取り出すことができれば推薦システムに役立てることができる。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) (Lee and Seung (2000)) という。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、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。例えば、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す。NMF では、各観測ベクトルを M 個の基底ベクトルの重みつき和によって表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。パワースペクトルの非加法性については第 7 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)

$$\mathbf{y}_n \simeq \sum_{m=1}^M \mathbf{h}_m u_{m,n} \quad (n = 1, \dots, N) \quad (2.1)$$

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、式 (2.1) は $\mathbf{Y} \simeq \mathbf{H}\mathbf{U}$ を意味する。このように NMF は、観測ベクトルを並べた行列を二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

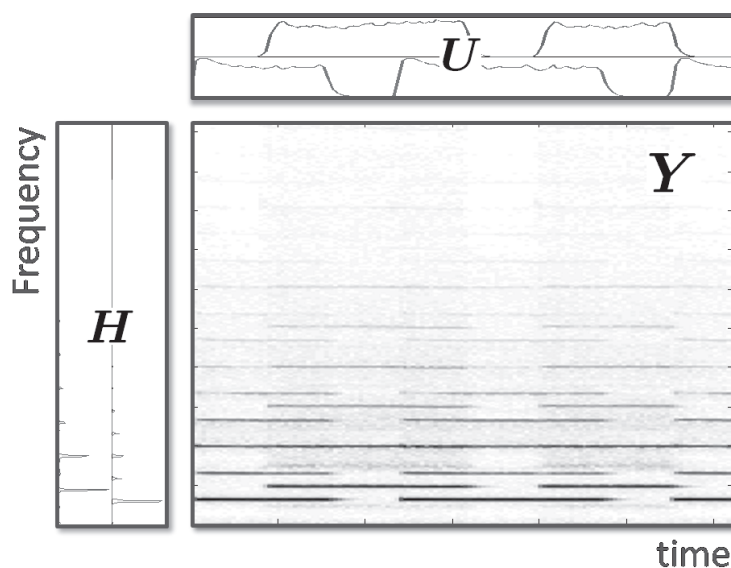


図1 NMFによるスペクトログラムの分解.

のイメージを掴んでもらうために、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。

3. NMFの基本性質

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。例えば $M = K$ の場合、 $\mathbf{H} = \mathbf{I}$ (\mathbf{I} は単位行列) であるような分解表現 $\mathbf{Y} = \mathbf{H}\mathbf{U}$ 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、 $\mathbf{U} = \mathbf{I}$ であるような分解表現 $\mathbf{Y} = \mathbf{H}\mathbf{U}$ 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 \mathbf{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、その場合に求まる基底行列と係数行列が重要な意味をもつ。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第8節で紹介する。

NMFでは非負制約の副次的効果により係数行列の要素がスパース(疎)になる傾向がある。このことは直感的には以下のように説明できる。まず、 $\mathcal{D}(\cdot, \cdot)$ をベクトル間の距離(な

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。

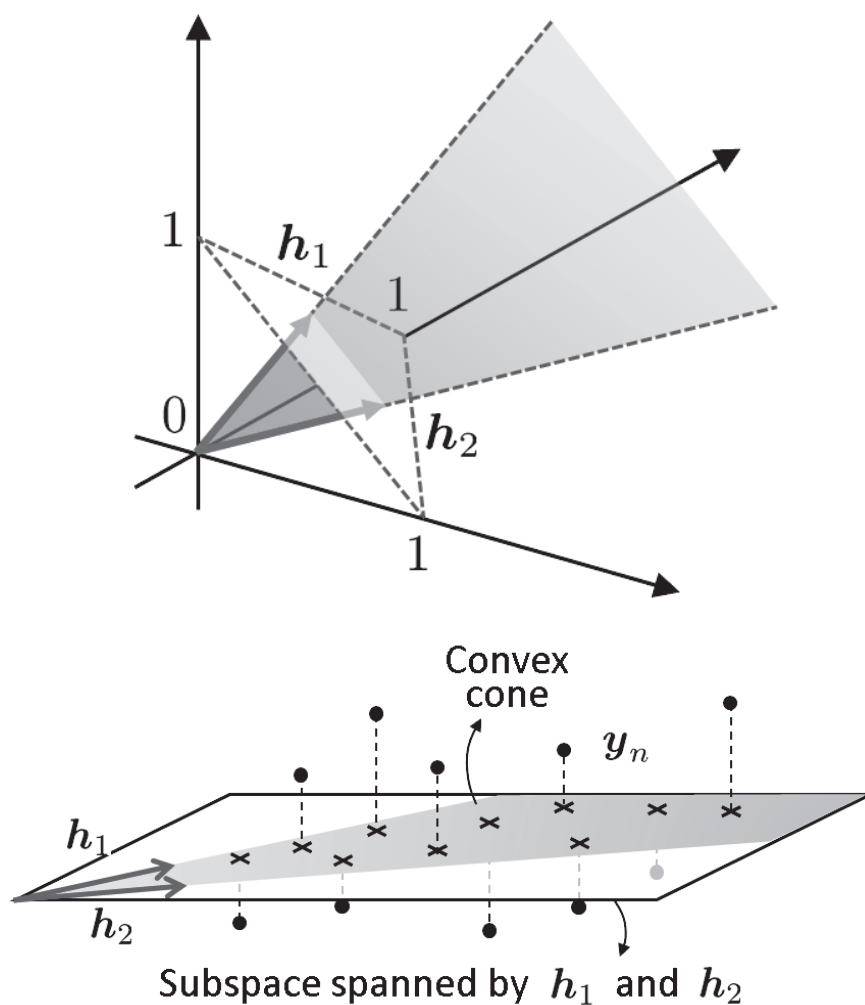


図2 NMFの幾何学的イメージ.

いし擬距離)として, 非負制約がない場合の $\hat{u} = \operatorname{argmin} \mathcal{D}(y|Hu)$ のような最適化問題を考えよう. このとき $H\hat{u}$ は, h_1, \dots, h_M が張る平面上で y から最も近い点となる. 例えば \mathcal{D} をベクトル間の差のノルムとすれば, この点は y から当該平面に下ろした垂線の足に対応する. ここで, 非負制約下でのこの最適化問題の解を \tilde{u} とすると, 無制約下での最適解 \hat{u} がたまたま非負制約を満たしている場合を除いて, $H\tilde{u}$ は図2で示した凸錐の中で \hat{u} に最も近い点, すなわち必ず凸錐の境界上の点になる. これは係数ベクトルの要素のうちいずれかが0になることを意味し, 非負制約下での最適解 \tilde{u} が無制約下での最適解 \hat{u} に比べてスパースになる (多くの要素が0になる) ことを示している. なお, スパース性は統計的独立性 (非ガウス性) と関連があり, 上述の傾向は係数行列の各行が似通いすぎないようにする効果を生んでいる. 以上のことは $Y \simeq HU$ の転置 $Y^T \simeq U^T H^T$ に対しても

まったく同様であり、 \mathbf{H} に対する非負制約により \mathbf{H} もスパースになる傾向がある。

NMF では、観測データの中で共起する成分をひとまとめにしたものが基底ベクトルの推定結果になる傾向がある。例えば、どのデータも成分 $\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c$ の重み付き和で表される時、 $\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c$ を基底にすれば全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 \mathbf{x}_a や成分 \mathbf{x}_b が生起するときいつも成分 \mathbf{x}_c も揃って生起するようなら、 $\mathbf{x}_a + \mathbf{x}_c$ および $\mathbf{x}_b + \mathbf{x}_c$ を基底としても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に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。先に述べたように、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。

4. NMF アルゴリズム

4.1 正値行列因子分解と NMF

観測行列を二つの非負値行列の積で表そうという NMF の基本概念自体は、Paatero and Tapper (1994) によって最初に提案されている。Paatero and Tapper (1994) は誤差行列 $\mathbf{Y} - \mathbf{H}\mathbf{U}$ の Frobenius ノルム (行列の要素の二乗和) で $\mathbf{H}\mathbf{U}$ の \mathbf{Y} からの乖離度 $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ 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、

$$B(\mathbf{H}, \mathbf{U}) = - \sum_{k,m} \log h_{k,m} - \sum_{m,n} \log u_{m,n} \quad (4.1)$$

のように各行列要素が 0 になる場合に無限大のペナルティを課す障壁関数を定義し、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 and Tapper (1994) 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 and Seung (2000) によって考案され、これをきっかけに NMF が広く知れ渡ることとなった。

4.2 行列の乖離度

NMF は、 $\mathbf{H}\mathbf{U}$ の \mathbf{Y} からの乖離度を表す規準の定義に応じて異なる最適化問題に帰着する。Lee and Seung (2000) は、前述の二乗誤差規準に加え、I ダイバージェンス規準 (Csiszár (1975)) の NMF アルゴリズムを導出している。当然ながら異なる乖離度規準を用いた場合で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。例えば、

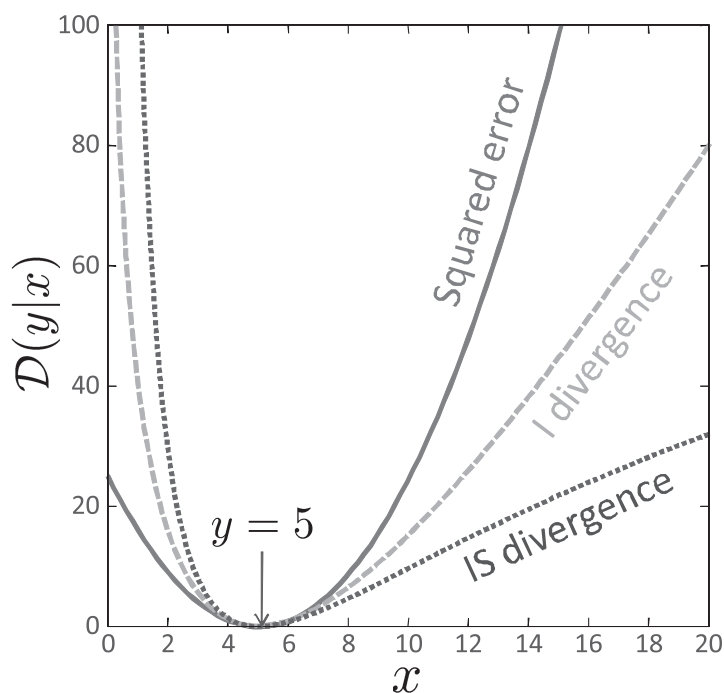


図3 $\mathcal{D}_{\text{EU/KL/IS}}(y|x)$ を x の関数と見たときのグラフ.

パワースペクトルを対象とした NMF では板倉齋藤擬距離 (板倉 (1972)) が乖離度規準として用いられることがある (Févotte *et al.* (2009)) が, これは 7.3 節で述べる信号波形の生成プロセスに関する仮定に基づいている.

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

$$\mathcal{D}_{\text{EU}}(y|x) = (y - x)^2 \quad (4.2)$$

$$\mathcal{D}_{\text{KL}}(y|x) = y \log \frac{y}{x} - y + x \quad (4.3)$$

$$\mathcal{D}_{\text{IS}}(y|x) = \frac{y}{x} - \log \frac{y}{x} - 1 \quad (4.4)$$

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 3 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齋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分かる. また, 板倉齋藤擬距離は y と x の比のみで表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. これらを用いれば, HU の Y からの乖離度を,

$$D(\mathbf{H}, \mathbf{U}) = \sum_{k,n} \mathcal{D} \left(y_{k,n} \left| \sum_m h_{k,m} u_{m,n} \right. \right)$$

で測ることができる (\cdot は EU/KL/IS のいずれかをさすものとする)。

4.3 補助関数法

NMF では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。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が、²⁾補助関数法と呼ぶ方法論 (Ortega and Rheinboldt (1970), Hunter and Lange (2000)) により局所解を得るための反復アルゴリズムを導くことができる。詳細は後述するが補助関数法とは、目的関数値がちょうど下界となっているような関数 (補助関数と呼ぶ) を設計し、目的関数の代わりにその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させていく方法である。なお、不完全データの下で確率モデルのパラメータを推定する方法として有名な Expectation-Maximization (EM) アルゴリズムはこの方法の一種である。

上述の最適化問題におけるポイントの一つは非負制約である。仮に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てくれれば、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大変都合であるが、残念ながらこれらの目的関数はそのようなになっていない。しかし、このような関数を補助関数の定義を満たすように設計することができれば、非負制約を満たしながら局所解を探索することが可能である。本章では、この方針に則って NMF のアルゴリズムを導く。

まず、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 4.1 $\theta = \{\theta_i\}_{1 \leq i \leq I}$ をパラメータとする目的関数 $D(\theta)$ に対し、 $D(\theta) = \min_{\alpha} G(\theta, \alpha)$ が成り立つとき、 $G(\theta, \alpha)$ を $D(\theta)$ の**補助関数**と定義する。また、 α を**補助変数**と呼ぶ。

定理 4.1 (補助関数法) 補助関数 $G(\theta, \alpha)$ を、 α 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

$$\alpha \leftarrow \underset{\alpha}{\operatorname{argmin}} G(\theta, \alpha) \quad (4.5)$$

$$\theta_i \leftarrow \underset{\theta_i}{\operatorname{argmin}} G(\theta, \alpha) \quad (i = 1, \dots, I) \quad (4.6)$$

を繰り返すと、目的関数 $D(\theta)$ の値は単調減少する。

証明 4.1 反復計算のステップ数を l とし、 $\theta = \theta^{(l)}, \alpha = \alpha^{(l)}$ から $\theta = \theta^{(l+1)}, \alpha = \alpha^{(l+1)}$ に更新されたときに、 $D(\theta)$ が増加しないことを示す。 $\alpha^{(l+1)} = \underset{\alpha}{\operatorname{argmin}} G(\theta^{(l)}, \alpha)$ なので、補助関数の定義より $D(\theta^{(l)}) = G(\theta^{(l)}, \alpha^{(l+1)})$ である。また、式 (4.6) より、明らかに $G(\theta^{(l)}, \alpha^{(l+1)}) \geq G(\theta^{(l+1)}, \alpha^{(l+1)})$ である。更に補助関数の定義より $G(\theta^{(l+1)}, \alpha^{(l+1)}) \geq D(\theta^{(l+1)})$ であるから、結局、 $D(\theta^{(l)}) \geq D(\theta^{(l+1)})$ である (図 4 参照)。

²⁾ Majorization-Minimization (または Minorization-Maximization) 法とも呼ばれている。

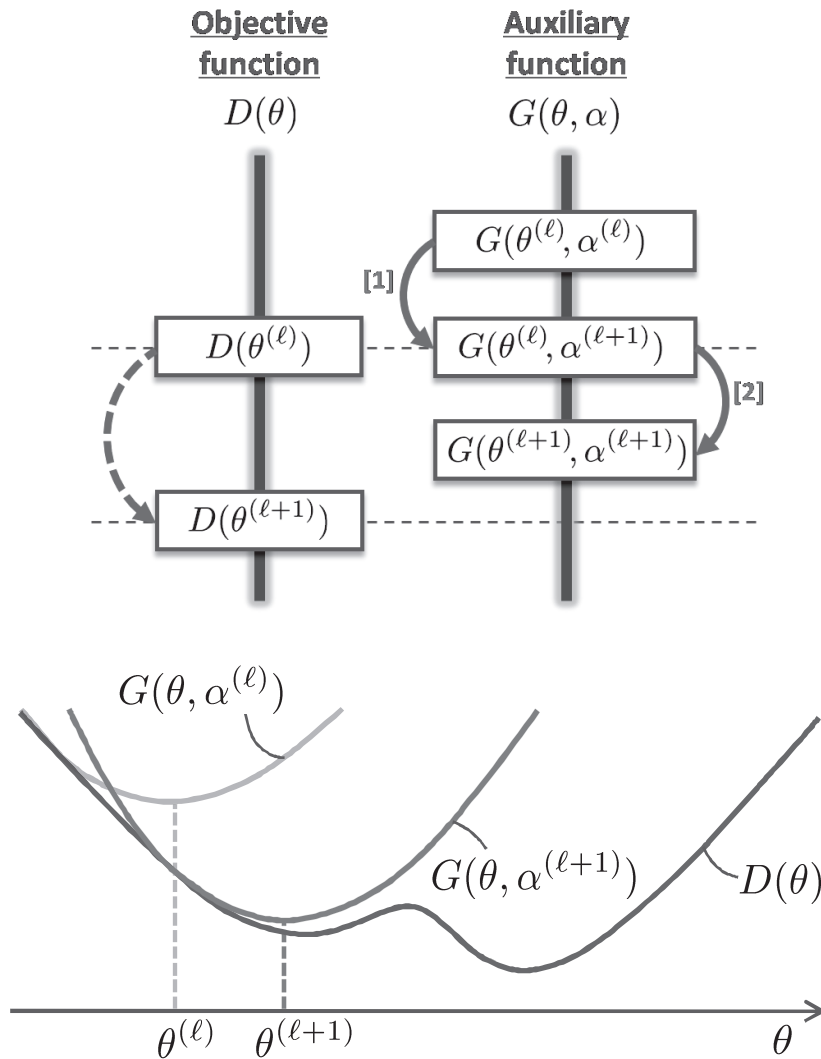


図4 補助関数法によるパラメータ更新のイメージ.

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 $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ を規準とした NMF アルゴリズムを導出する。以下の導出方法は Lee and Seung (2000) のものとは異なるがアルゴリズムは同形となる。

$D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ を展開し、 \mathbf{H} と \mathbf{U} に依らない項を省略すると、

$$D_{\text{EU}}(\mathbf{H}, \mathbf{U}) \stackrel{\mathbf{H}, \mathbf{U}}{=} \sum_{k,n} (-2y_{k,n}x_{k,n} + x_{k,n}^2) \quad (4.7)$$

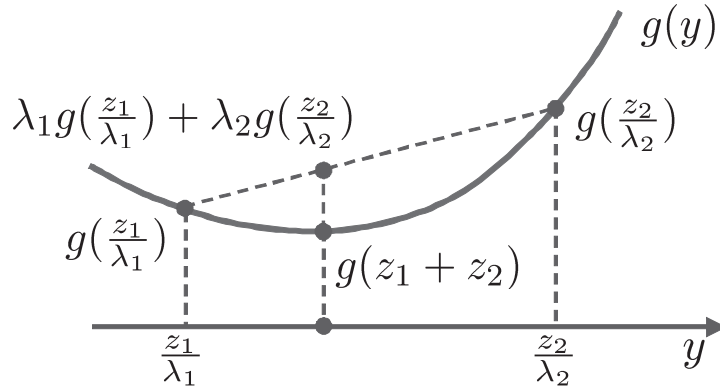


図5 Jensen の不等式 ($I = 2$ の場合).

と書ける。ただし、

$$x_{k,n} = \sum_m h_{k,m} u_{m,n} \tag{4.8}$$

である。また、 \hat{z} は z に関する項のみに関する等号を表す。式 (4.7) 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 4.2 (非負変数に対する Jensen の不等式 (図 5)) 任意の凸関数 g , I 個の非負の変数 z_1, \dots, z_I , $\sum_i \lambda_i = 1$ を満たす I 個の非負の重み係数 $\lambda_1, \dots, \lambda_I$ のもとで、

$$g\left(\sum_i z_i\right) \leq \sum_i \lambda_i g\left(\frac{z_i}{\lambda_i}\right) \tag{4.9}$$

が成り立ち、 $z_1/\lambda_1 = \dots = z_I/\lambda_I$, すなわち

$$\lambda_i = \frac{z_i}{\sum_j z_j} \tag{4.10}$$

のとき等号が成立する。

$h_{k,m} u_{m,n} \geq 0$ なのでこれを用いると、 $x_{k,n}^2$ に対し、

$$x_{k,n}^2 \leq \sum_m \lambda_{k,m,n} \left(\frac{h_{k,m} u_{m,n}}{\lambda_{k,m,n}}\right)^2 \tag{4.11}$$

のような不等式を立てることができる。ただし、 $\lambda_{k,m,n} \geq 0, \sum_m \lambda_{k,m,n} = 1$ である。さて、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

$$\lambda_{k,m,n} = \frac{h_{k,m}u_{m,n}}{x_{k,n}} \quad (4.12)$$

のときである。以上より、 $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ における $x_{k,n}^2$ の項を式 (4.11) の右辺に置き換えることにより得られる関数

$$G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}) = \sum_{k,n} \left(y_{k,n}^2 - 2y_{k,n} \sum_m h_{k,m}u_{m,n} + \sum_m \frac{h_{k,m}^2 u_{m,n}^2}{\lambda_{k,m,n}} \right) \quad (4.13)$$

は $D_{\text{EU}}(\mathbf{H}, \mathbf{U})$ 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 4.1 に従い、

$$\boldsymbol{\lambda} \leftarrow \underset{\boldsymbol{\lambda}}{\text{argmin}} G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}) \quad (4.14)$$

$$\mathbf{H} \leftarrow \underset{\mathbf{H}}{\text{argmin}} G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}) \quad (4.15)$$

$$\mathbf{U} \leftarrow \underset{\mathbf{U}}{\text{argmin}} G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}) \quad (4.16)$$

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。式 (4.14) については既に述べたとおり式 (4.12) で与えられる。式 (4.15) と式 (4.16) 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、先に述べたように $G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda})$ 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。例えば、 $G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda})$ 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 $h_{k,m}$ は

$$\hat{h}_{k,m} = \frac{\sum_n y_{k,n} u_{m,n}}{\sum_n u_{m,n}^2 / \lambda_{k,m,n}} \quad (4.17)$$

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

$$\hat{u}_{m,n} = \frac{\sum_k y_{k,n} h_{k,m}}{\sum_k h_{k,m}^2 / \lambda_{k,m,n}} \quad (4.18)$$

のとき $G_{\text{EU}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda})$ は最小となる。よって、式 (4.15) と式 (4.16) 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、式 (4.17), (4.18) 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、式 (4.14) を式 (4.15), (4.16) に代入すると、下記の 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 (Lee and Seung (2000))

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。

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

$D_{\text{KL}}(\mathbf{H}, \mathbf{U})$ を規準とした NMF アルゴリズムも以上と同様に導ける。 $D_{\text{KL}}(\mathbf{H}, \mathbf{U})$ は、

$$D_{\text{KL}}(\mathbf{H}, \mathbf{U}) \stackrel{H, U}{=} \sum_{k,n} (-y_{k,n} \log x_{k,n} + x_{k,n}) \quad (4.19)$$

と書けるが、この式の中で $-y_{k,n} \log x_{k,n}$ が $h_{k,1}, \dots, h_{k,M}, u_{1,n}, \dots, u_{M,n}$ を含む非線形項である。負の対数関数が凸関数であることに注意すると、この項に対し Jensen の不等式

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

が立てられる。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。この補助関数を用いて以下の NMF アルゴリズムを導ける。

I ダイバージェンス規準の NMF アルゴリズム (Lee and Seung (2000))

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。

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

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

次に、筆者らによる、板倉齋藤擬距離 $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ を規準とした NMF アルゴリズムの導出 (亀岡他 (2006)) 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

$$D_{\text{IS}}(\mathbf{H}, \mathbf{U}) \stackrel{H, U}{=} \sum_{k,n} \left(\frac{y_{k,n}}{x_{k,n}} + \log x_{k,n} \right) \quad (4.20)$$

と書ける. 先と同様の方針で補助関数法により当該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鍵となる. まず, 逆数関数は正領域において凸関数であるため, $1/x_{k,n}$ の項に関しては, Jensen の不等式

$$\frac{1}{x_{k,n}} \leq \sum_m \lambda_{k,m,n} \left(1 / \frac{h_{k,m} u_{m,n}}{\lambda_{k,m,n}} \right) \quad (4.21)$$

が成り立つ. ただし, $\lambda_{k,m,n}$ は先と同様, $\lambda_{k,m,n} > 0, \sum_m \lambda_{k,m,n} = 1$ を満たす. 次に, $\log x_{k,n}$ の項に関しては, 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. ここで, 任意の微分可能な凹関数 g に対し,

$$g(x) \leq g(\alpha) + (x - \alpha)g'(\alpha) \quad (4.22)$$

(等号成立は $x = \alpha$) が成り立つことを利用すると,

$$\log x_{k,n} \leq \log \alpha_{k,n} + \frac{1}{\alpha_{k,n}}(x_{k,n} - \alpha_{k,n}) \quad (4.23)$$

のような不等式が立てられ, 行列要素の 1 次式で表された上限関数を設計できる. 以上の二つの不等式の等号成立条件は $\lambda_{k,m,n}$ と $\alpha_{k,n}$ がそれぞれ

$$\lambda_{k,m,n} = \frac{h_{k,m} u_{m,n}}{x_{k,n}}, \quad \alpha_{k,n} = x_{k,n} \quad (4.24)$$

のときである. 以上より, $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ における $1/x_{k,n}$ の項を式 (4.21) の右辺に, $\log x_{k,n}$ の項を式 (4.23) の右辺に置き換えることにより得られる関数

$$G_{\text{IS}}(\mathbf{H}, \mathbf{U}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = \sum_{k,n} \left(\sum_m \frac{y_{k,n} \lambda_{k,m,n}^2}{h_{k,m} u_{m,n}} + \sum_m \frac{h_{k,m} u_{m,n}}{\alpha_{k,n}} - \log y_{k,n} + \log \alpha_{k,n} - 2 \right) \quad (4.25)$$

は $D_{\text{IS}}(\mathbf{H}, \mathbf{U})$ の補助関数の要件を満たす. あとは定理 4.1 に従って各更新式を求めれば板倉齋藤擬距離規準の NMF アルゴリズムが導かれる.

板倉齋藤距離規準の NMF アルゴリズム (亀岡他 (2006))

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.

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

4.6 $D_\beta(H, U)$ を規準とした NMF

式 (4.2)~(4.4) に 3 種類の規準を挙げたが、これらは β ダイバージェンスと呼ぶ規準 (Eguchi and Kano (2001))

$$D_\beta(y|x) = y \frac{y^{\beta-1} - x^{\beta-1}}{\beta-1} - \frac{y^\beta - x^\beta}{\beta} \quad (4.26)$$

で統一的に記述することができる。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、式 (4.26) は $\beta \rightarrow 0$ のとき板倉齋藤擬距離、 $\beta \rightarrow 1$ 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。筆者らはこれを規準とした一般化 NMF アルゴリズムを導いている (Nakano *et al.* (2010))。式 (4.26) 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。導出のポイントは前述の亀岡他 (2006) のアイデアと同様であり、凸関数の項に対しては式 (4.9) を、凹関数の項に対しては式 (4.22) の不等式を用いて上限関数を設計する、という方針で閉形式の更新式からなる反復アルゴリズムを補助関数法に基づき導くことができる。以上のアイデアで導出される具体的なアルゴリズムは以下のとおりである。

β ダイバージェンス規準の NMF アルゴリズム (Nakano *et al.* (2010))

1. H, U を非負値行列に初期設定する。また、 $\varphi(\beta)$ を以下のように設定する。

$$\varphi(\beta) = \begin{cases} 1/(2-\beta) & (\beta < 1) \\ 1 & (1 \leq \beta \leq 2) \\ 1/(\beta-1) & (\beta > 2) \end{cases} \quad (4.27)$$

2. 以下の更新を収束するまで繰り返す。

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

上記のアルゴリズムは、 $\beta = 0, 1, 2$ のとき 4.4 節と 4.5 節で示した板倉齋藤擬距離規準、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となる。

5. 生成モデルとしての解釈

5.1 二乗誤差, I ダイバージェンス, 板倉齋藤距離, β ダイバージェンス規準の NMF

ここでは, 以上のような規準で NMF を行うことが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場合に正当化されるかについて考える. 二乗距離, I ダイバージェンス, 板倉齋藤距離, β ダイバージェンスを規準とした NMF は, 観測行列の要素 $y_{k,n}$ が $x_{k,n}$ を平均とした正規分布, Poisson 分布, 指数分布, Tweedie 分布

$$y_{k,n} \sim \mathcal{N}(y_{k,n}; x_{k,n}, \sigma^2) \quad (5.1)$$

$$y_{k,n} \sim \text{Poisson}(y_{k,n}; x_{k,n}) \quad (5.2)$$

$$y_{k,n} \sim \text{Exponential}(y_{k,n}; x_{k,n}) \quad (5.3)$$

$$y_{k,n} \sim \text{Tweedie}(y_{k,n}; x_{k,n}, \phi) \quad (5.4)$$

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

$$\mathcal{N}(z; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-\mu)^2/2\sigma^2} \quad (5.5)$$

$$\text{Poisson}(z; \mu) = \mu^z e^{-\mu} / z! \quad (z \geq 0) \quad (5.6)$$

$$\text{Exponential}(z; \mu) = \frac{1}{\mu} e^{-z/\mu} \quad (z \geq 0) \quad (5.7)$$

$$\text{Tweedie}(z; \mu, \phi) = a(z, \phi) e^{\frac{1}{\phi}(z\rho(\mu) - \kappa(\mu))} \quad (5.8)$$

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

である. このことは, 以下により確かめられる. 式 (5.1)~(5.4) 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, $x_{k,n} = y_{k,n}$ のときにのみ 0 になる, $y_{k,n}$ と $x_{k,n}$ の近さを表す非負の尺度と見なせ, 式 (5.1)~(5.4) の場合の対数尤度差 $L(y_{k,n}) - L(x_{k,n})$ は, それぞれ式 (4.2)~(4.4) と式 (4.26) と等しい.

5.2 自然指数分布族と Bregman ダイバージェンス

以上のように, 観測値へのモデルフィッティング問題におけるフィッティング規準の仮定は観測値の確率分布の仮定に対応している. 上記では 4 つの例を示したが, ここでは自然指数分布族に含まれる確率分布のクラスと Bregman ダイバージェンス (Bregman (1967)) と呼ばれるフィッティング規準が対応していること, β ダイバージェンスが Bregman ダイバージェンスに含まれることを示す. 以降表記の簡単化のためインデックス k, n を省略し, 観測行列の要素 y が指数分布族の確率分布

$$y \sim \exp \{ \eta T(y) - \psi(\eta) + c(y) \} \quad (5.9)$$

に従うものとする。ただし、 ψ は無限微分可能な凸関数である。 η を自然パラメータといい、確率分布のパラメータの関数である。ここで、特に $T(y) = y$ というクラスを考える。このときの指数分布族を自然指数分布族という。

まず準備として、 ψ の Legendre 変換

$$\phi(z) = \max_v (vz - \psi(v)) \quad (5.10)$$

を導入する。 ψ は凸関数なので、Legendre 変換の性質より ϕ も凸関数である。 $vz - \psi(v)$ を最大にする v を v^* とすると、 v^* は $(vz - \psi(v))' = 0$ 、すなわち

$$\psi'(v^*) = z \quad (5.11)$$

を満たす。一方、 y のキュムラント母関数が $K(t) = \log \mathbb{E}[e^{yt}] = \psi(t + \eta) - \psi(\eta)$ となることから、 y の期待値を x とすると、

$$K'(0) = \psi'(\eta) = x \quad (5.12)$$

が言える。 ψ は凸関数より ψ' は一対一対応の関数であることが分かるので、 η と x は一対一対応する。式 (5.11), (5.12) より、 $\eta = \operatorname{argmax}_v (vx - \psi(v))$ が言えるので、 $\phi(x)$ は

$$\phi(x) = \eta(x)x - \psi(\eta(x)) \quad (5.13)$$

と書ける。ここで、式 (5.13) の両辺を微分すると、

$$\phi'(x) = \eta(x) + \eta'(x)x - \psi'(\eta(x))\eta'(x) \quad (5.14)$$

となるが、式 (5.12) より $\phi'(x) = \eta(x)$ となることが分かる。

以上で得た $\phi(x) = \eta x - \psi(\eta)$ および $\phi'(x) = \eta(x)$ という関係式を、自然指数分布族の確率密度関数 $\exp\{\eta y - \psi(\eta) + c(y)\} = \exp\{\eta x - \psi(\eta) + \eta(y - x) + c(y)\}$ に代入すると、

$$\exp\{\phi(x) + \phi'(x)(y - x) + c(y)\} \quad (5.15)$$

を得る。ここで、 x の対数尤度 $L(x) = \log p(y|x)$ は、 $(\phi(x) + \phi'(x)(y - x))' = \phi''(x)(y - x)$ より、 $x = y$ のときに最大となる。すなわち、 $L(y) \geq L(x)$ である。よって、前節と同様、対数尤度差 $L(y) - L(x)$

$$\mathcal{D}_\phi(y|x) = \phi(y) - \phi(x) - \phi'(x)(y - x) \quad (5.16)$$

は、 $x = y$ のときにのみ 0 になる、 y と x の近さを表す非負の尺度と見なせる。この尺度は Bregman ダイバージェンスと呼ばれる (Bregman (1967))。図 6 に示すとおり $\mathcal{D}_\phi(y|x)$

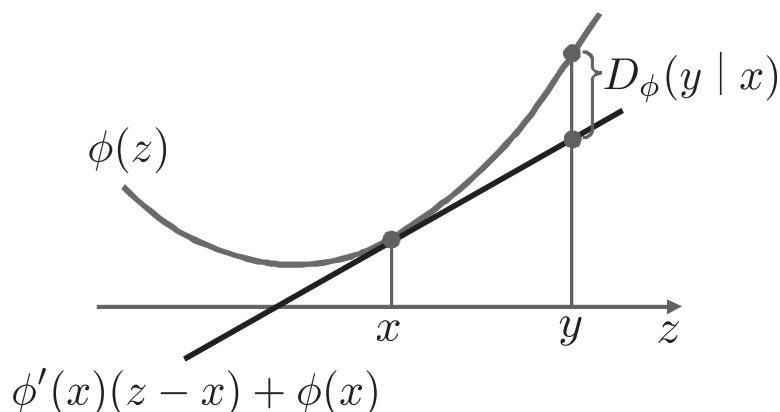


図6 Bregman ダイバージェンス.

は凸関数 ϕ と ϕ の接線の差に相当する. この図からも $D_\phi(y|x) \geq 0$, および x と y が等しいときにのみ $D_\phi(y|x) = 0$ となることが分かる.

先に登場した β ダイバージェンス (Eguchi and Kano (2001)) は, ϕ が

$$\phi(x) = \begin{cases} -\log x + x - 1 & (\beta = 0) \\ x \log x - x + 1 & (\beta = 1) \\ \frac{x^\beta}{\beta(\beta-1)} - \frac{x}{\beta-1} + \frac{1}{\beta} & (\text{otherwise}) \end{cases} \quad (5.17)$$

の場合の Bregman ダイバージェンスの特殊ケースに相当する (Hennequin *et al.* (2011)). 従って, β ダイバージェンスの特殊ケースに相当する二乗誤差, I ダイバージェンス, 板倉齋藤距離も当然 Bregman ダイバージェンスの特殊ケースに相当する.

なお, Dhillon and Sra (2006) は Bregman ダイバージェンスを規準とした NMF アルゴリズムを ϕ を特殊なクラスに限定して議論しているが, 任意の ϕ に対する補助関数法に基づくアルゴリズムはまだ見つかっていない.

6. 確率的潜在意味解析との関連

文書データのクラスタリングやインデキシングを目的とした確率的潜在意味解析 (probabilistic Latent Semantic Analysis; pLSA) (Hofmann (1999)) は, NMF と関連が深いので簡単に紹介する.

文書 n の中に含まれる単語 k の個数を $y_{k,n}$ とし, 全単語のヒストグラム $\mathbf{y}_n = (y_{1,n}, \dots, y_{K,n})^T$ を文書データと呼ぶ. 単語の出現回数は, 政治・経済・スポーツといったトピックに大きく依存するはずであるが, 文書データからトピックを推論することができれば文書検索のためのインデキシングなどに有用である. これを行うことが pLSA の目的である.

さて、トピックが m のとき単語 k が出現する確率を $p(k|m)$ 、文書 n がトピック m のものである確率を $p(m|n)$ とすると、文書 n に単語 k が出現する確率 $p(k|n)$ は

$$p(k|n) = \sum_m p(k|m)p(m|n) \quad (6.1)$$

となる。 $x_{k,n} = p(k|n)$, $h_{k,m} = p(k|m)$, $u_{m,n} = p(m|n)$ とし, $\mathbf{X} = (x_{k,n})_{K \times N}$, $\mathbf{H} = (h_{k,m})_{K \times M}$, $\mathbf{U} = (u_{m,n})_{M \times N}$ と置けば, 式 (6.1) は行列積 $\mathbf{X} = \mathbf{H}\mathbf{U}$ の形で書ける。文書データの各単語が上記の $p(k|n)$ に従って独立に生成されたものとする, 全文書データの全単語が生成される確率は $\prod_{k,n} p(k|n)^{y_{k,n}}$ となる。今, $H_{k,m} = p(k|m)$ と $U_{m,n} = p(m|n)$ が未知であるから, 文書データ $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_N)$ から \mathbf{H} と \mathbf{U} を最尤推定する問題は,

$$\log p(\mathbf{Y}|\mathbf{H}, \mathbf{U}) = \sum_{k,n} y_{k,n} \log x_{k,n} \quad (6.2)$$

を \mathbf{H} と \mathbf{U} に関して $h_{k,m} \geq 0$, $\sum_k h_{k,m} = 1$, $u_{m,n} \geq 0$, $\sum_m u_{m,n} = 1$ の条件の下で最大化する問題に帰着する。式 (4.19) (I ダイバージェンス) と見比べると, 上述の対数尤度は式 (4.19) の第一項の符号を逆転させたものと等しいことが分かる。また, 式 (4.19) の第二項が $h_{k,m}$ および $u_{m,n}$ に関する Lagrange 未定乗数項に対応すると考えれば, pLSA と I ダイバージェンス規準の NMF は同形の最適化問題となっていることが分かる。実際, 上述の最尤推定問題に対し, トピックインデックス m を潜在変数と見なすことで EM アルゴリズムを適用することができるが, このアルゴリズムは \mathbf{H} と \mathbf{U} を正規化するかどうかの違いを除けば 4.4 節で示した I ダイバージェンス規準の NMF アルゴリズムと等価である。

式 (6.2) の尤度関数は $h_{k,m}$ と $u_{m,n}$ を確率値と見なして立てられたものであり, 第 5 節での尤度関数とは立て方が異なっている点に注意が必要である。確率モデルのタイプとしては, pLSA は混合モデル (確率分布の和のモデル) に当てはまり, 第 5 節の解釈による NMF は因子モデル (確率変数の和の確率分布のモデル) に当てはまる。pLSA と I ダイバージェンス規準の NMF がアルゴリズム的に等価となるのは最尤推定の場合に限られるが, これは上記のように目的関数がたまたま同形になるからである。実際, 両者の³⁾Bayes 拡張は異なった姿になる。

7. 音響信号処理応用とその拡張

7.1 音源分離・自動採譜への応用

Smaragdīs and Brown (2003) は, 音楽音響信号の自動採譜を目的とし, 音楽音響信号の振幅 (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 を適用し, 楽音ごと

³⁾ pLSA の Bayes 拡張は Latent Dirichlet Allocation (LDA) (Blei *et al.* (2003)) と呼ばれる。因子モデルとしての I ダイバージェンス規準の NMF の Bayes 拡張は例えば Cemgil (2008) により議論されている。

のスペクトログラムに分解する方法を提案している。各時刻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の振幅比が時不変でパワー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（周波数成分の振幅比）と各時刻におけるパワーを推定する問題は、NMF と同形の問題と見なせることが本アプローチのポイントである。例えば、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、各音高の音の周波数成分の振幅比が時刻に依らずいつも同じであ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パワー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に即したものに改変した改良モデルを紹介する。

7.2 複素 NMF

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は実際には非加法的である。つまり、波形 A と波形 B の和の振幅スペクトルは波形 A の振幅スペクトルと波形 B の振幅スペクトルの和とは等しくなると限らない。従って、振幅スペクトログラムを加法的な成分に分解したところで信号を分解したことにはならないのである。

NMF を振幅スペクトログラムに適用するアプローチでは、周波数成分の振幅比さえ同じであれば波形（パワーや位相スペクトル）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が、著者らはこの考え方を基にした複素スペクトログラムのモデルを立て、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」と呼ぶ方法を提案している（亀岡他 (2008)）。以下にその概要を述べる。

音源 m の複素スペクトログラムを $a_{m,k,n} \in \mathbb{C}$ で表すと、 M 個の音源からなる多重音の複素スペクトログラム $f_{k,n} \in \mathbb{C}$ は

$$f_{k,n} = \sum_{m=1}^M a_{m,k,n} = \sum_m |a_{m,k,n}| e^{j\phi_{m,k,n}} \quad (7.1)$$

と表される。ここで、各音源の周波数成分の振幅比が時不変という拘束を設けると、 $|a_{m,k,n}| = h_{k,m} u_{m,n}$ と置けるため、

$$f_{k,n} = \sum_m h_{k,m} u_{m,n} e^{j\phi_{m,k,n}} \quad (7.2)$$

のような形の複素スペクトログラムモデルが立てられる。以上のモデルを観測複素スペクトログラム $y_{k,n} \in \mathbb{C}$ にフィットするように H, U, ϕ を求めることが目標である。また、NMF において非負制約によりもたらされる第3節で述べたような副次的効果と類似の効果を促すためにもできるだけ U がスパースになるような解を選びたい。筆者らは、

$$I(H, U, \phi) := \sum_{k,n} |y_{k,n} - f_{k,n}|^2 + 2\gamma \sum_{m,n} |u_{m,n}|^p \quad (7.3)$$

を H, U, ϕ に関して最小化する最適化問題を考え、補助関数法に基づく反復アルゴリズムを導いている。ただし、 $0 < p < 2$ と $\gamma \geq 0$ は定数である。この最適化問題では、目的関数が $\phi_{m,k,n}$ に関して非線形であること、 $u_{m,n}$ に関して微分不可能であることが最適解を解析的に得ることを困難にしている。式(7.3)の第一項については、4.4節と同様に \sum_m が $|\cdot|^2$ の外に出された形の関数を補助関数にすることができれば $\phi_{m,k,n}$ の更新則を閉形式で得ることができる。第4.4節と異なり、 $h_{k,m}u_{m,n}e^{j\phi_{m,k,n}}$ は複素数なので定理4.2は適用できないが、代わりに以下の不等式を適用することができる。

定理 7.1 (複素数変数に対する Jensen の不等式) 任意の凸関数 g , I 個の複素数変数 z_1, \dots, z_I , $\sum_i \alpha_i = 0$ を満たす I 個の複素数変数 $\alpha_1, \dots, \alpha_I$, $\sum_i \beta_i = 1$ を満たす I 個の任意の非負の重み係数 β_1, \dots, β_I のもとで、

$$g\left(\sum_i z_i\right) \leq \sum_i \beta_i g\left(\frac{z_i - \alpha_i}{\beta_i}\right) \quad (7.4)$$

が成り立ち、

$$\alpha_i = z_i - \beta_i \left(\sum_i z_i\right) \quad (7.5)$$

のとき等号が成立する。

また、式(7.3)の第二項は $u_{m,n}$ に関して微分不可能であるが、 $0 < p \leq 2$ のとき

$$|u_{m,n}|^p \leq \frac{p|v_{m,n}|^{p-2}}{2} u_{m,n}^2 + \frac{2-p}{2} |v_{m,n}|^p \quad (7.6)$$

が成り立つことを用いると、

$$I^+(H, U, \phi, \alpha, V) := \sum_{k,n,m} \frac{1}{\beta_{m,k,n}} \left| \alpha_{m,k,n} y_{k,n} - h_{k,m} u_{m,n} e^{j\phi_{m,k,n}} \right|^2 + \gamma \sum_{m,n} \left\{ p |v_{m,n}|^{p-2} u_{m,n}^2 + (2-p) |v_{m,n}|^p \right\} \quad (7.7)$$

のような補助関数を立てることができる。ただし、 $0 < \beta_{m,k,n} < 1$ は $\sum_m \beta_{m,k,n} = 1$ を満たす任意の定数、 $\alpha_{m,k,n}, v_{m,n}$ は補助変数である。これにより閉形式の更新則からなる反復アルゴリズムを導くことができる。

7.3 板倉齋藤擬距離規準の NMF

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、各構成音の信号を互いに統計的に独立な確率変数と見なしたとき、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。

多重音信号中の各構成音の信号が短時間区間毎に平均が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独立に平均が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 $s_{m,k,n}$ とすると、 $s_{m,k,n}$ は

$$s_{m,k,n} \sim \mathcal{N}_{\mathbb{C}}(s_{m,k,n}; 0, \nu_{m,k,n}) \quad (7.8)$$

に従う。ただし、 $\mathcal{N}_{\mathbb{C}}(z; \mu, \nu) = \frac{1}{\pi\nu} e^{-|z-\mu|^2/\nu}$ である。 $\nu_{m,k,n}$ は構成音 m のパワースペクトログラムの期待値 $\mathbb{E}[|s_{m,k,n}|^2]$ を表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 $y_{k,n} = \sum_m s_{m,k,n}$ 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、 $s_{m,k,n}$ と $s_{m',k,n}$ ($m \neq m'$) が互いに独立と仮定できるなら、 $y_{k,n}$ は

$$y_{k,n} \sim \mathcal{N}_{\mathbb{C}}(y_{k,n}; 0, \sum_m \nu_{m,k,n}) \quad (7.9)$$

に従う。よって、 $x_{k,n} = \sum_m \nu_{m,k,n}$ と置くと、 $y_{k,n}$ が与えられた下での $x_{k,n}$ の対数尤度は

$$L(x_{k,n}) = -\log \pi x_{k,n} - \frac{|y_{k,n}|^2}{x_{k,n}} \quad (7.10)$$

となる。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 $\mathcal{D}_{\text{IS}}(|y_{k,n}|^2 | x_{k,n})$ 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す $\nu_{m,k,n}$ に関し、周波数成分のパワー比が時不変であるような構造 $\nu_{m,k,n} = h_{k,m} u_{m,n}$ 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} に対し板倉齋藤擬距離規準の NMF を行うことと等価となる (Parry and Essa (2007), Févotte *et al.* (2009))。

7.4 可変基底 NMF

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の振幅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ことになる。この問題意識のもと、観測スペクトログラムを個々の持続音ごとに分解することを目的として、基底スペクトルが時間変化するよう改良された NMF の拡張モデルが提案されている (Smaragdis (2004), Ozerov *et al.* (2009), Nakano *et al.* (2011)).

7.5 その他の拡張モデル

NMF の楽音分離や自動採譜への応用においては、音楽スペクトログラムを各楽音のノートごとのスペクトログラムに分解することが目的となるが、その成功の鍵は、楽音がもつ構造、性質、傾向をいかに見つけ、NMF モデルにどう組み込めるか、にある。このような動機からさまざまな拡張モデルが提案されている。例えば、各基底スペクトルのパワーが時間方向に連続的となるよう制約が組み込まれたもの (Virtanen (2007)), 各基底スペクトルが調波構造をなすよう制約されたもの (Raczyński *et al.* (2007)), 各基底スペクトルに自己回帰系によるソースフィルタモデルの拘束が組み込まれたもの (Kameoka and Kashino (2009), Yoshii and Goto (2012b)), 基底スペクトルが音色特徴量空間においていくつかのクラスタを形成するよう基底のクラスタリングと NMF によるスペクトログラム分解が同時に行われるもの (Kameoka *et al.* (2012)), などが検討されている。

7.6 その他の音響信号処理応用

NMF は、音源分離や自動採譜の他、音声強調 (Smaragdis *et al.* (2007)), 帯域拡張 (Smaragdis and Raj (2007)), 音楽信号からのボーカル成分抽出 (Durrieu *et al.* (2010))・ドラム音抽出 (Helén and Virtanen (2005)), 音声認識のための音素特徴量抽出 (Hurmalainen *et al.* (2011)), フォルマントトラッキング (Durrieu and Thiran (2011)), エコーキャンセラ (戸上・川口 (2009)), 音声の時間分解 (Hiroya (2013)) など多岐にわたった音響信号処理問題に応用展開されており、筆者らも NMF のアルゴリズムをヒントにしたブラインド残響除去法を提案している (Kameoka *et al.* (2009))。また、NMF はモノラル信号の音源分離のための有効なアプローチとして認識されるようになって以来、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も積極的に進めている (Ozerov and Févotte (2010), Kitano *et al.* (2010), Sawada *et al.* (2011, 2012), Higuchi *et al.* (2014))。

8. ノンパラメトリック Bayes NMF

8.1 基底数の推定

NMF において基底数をいかに推定するかは重要課題の一つである。Cemgil (2008) や Schmidt *et al.* (2009) は、第 5 節で述べたような 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。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が、近年はノンパラトリック Bayes アプローチによりパラメータ (基底行列と係数行列) とともにモデルの複雑度 (基底数) を一挙に推論しようという方法も提案されている (Hoffman *et al.* (2010), Nakano *et al.* (2011), Yoshii and Goto (2012b)). 本章では、後者のアプローチの例を紹介する。

8.2 ノンパラメトリック Bayes 法

ノンパラメトリック Bayes における「ノンパラメトリック」はパラメータが無いということではなく、無数のパラメータを仮定することを意味する。ノンパラメトリック Bayes 法とは、無数のパラメータからなる確率モデルを仮定し、所与のデータに応じてモデルのパラメータとともに複雑度を推論するアプローチの総称である。

ノンパラメトリック Bayes 法では無数のパラメータをもつ生成モデル (無限モデルという) を確率過程を導入して記述することが多い。混合モデルや因子モデルなど、確率モデルのタイプに応じてさまざまな無限モデルがこれまで提案されている。例えば、混合正規分布、隠れ Markov モデル、確率文脈自由文法、先に登場した pLSA などの混合モデルに関しては、Dirichlet 過程と呼ばれる確率過程をベースとして潜在クラス数や状態数を無限化したモデルを構成することができる。Dirichlet 分布は総和が 1 となるような有限個の非負変数を生成する確率分布であるが、Dirichlet 過程はその変数を無限個にしたものと考えれば良い。従って、Dirichlet 過程は無限離散分布 ($\sum_{i=1}^{\infty} \pi_i = 1$ を満たす $\pi_1, \pi_2, \dots, \pi_{\infty} \in [0, 1]$) を生成する確率モデルであり、例えば混合モデルにおいて無限個の分布の混合重みに対する事前分布として使用できる。Dirichlet 過程が生成する無限離散分布はスパースになる傾向があるため、Dirichlet 過程を用いた無限混合モデルでは所与のデータを説明する上で必要な分布以外の混合重みは 0 になるよう推論結果が誘導される。混合モデルの複雑度 (混合数) がデータから自動的に得られるのはこの効果による。なお、第 6 節で NMF と pLSA が関連が深いと述べたが、pLSA の考え方をベースにしたスペクトログラムのモデル化も形式的には⁴⁾可能 (Smaragdis *et al.* (2006)) で、Dirichlet 過程を用いればその見方からの無限モデル化も可能である (Yoshii and Goto (2012a)). 一方、NMF や独立成分分析などの因子モデルに関しては、ベータ過程やガンマ過程を用いて因子数を無限化したモデルを構成することができる。端的に言えばベータ過程は無数の区間 $[0, 1]$ の変数 ($\pi_1, \pi_2, \dots, \pi_{\infty} \in [0, 1]$) の生成モデル、ガンマ過程は無数の非負変数 ($\theta_1, \theta_2, \dots, \theta_{\infty} \in [0, \infty)$) の生成モデルであり、NMF の基底数を無限化したモデルはこれらを用いて構成できる。ベータ過程を用いる場合は、各基底の寄

⁴⁾ 周波数を単語、時刻を文書、各時間周波数におけるパワーを単語の出現回数と見なせば、スペクトログラムの分解に pLSA を適用することができる。

与の有無を表したバイナリ変数 $z_{m,n} \in \{0, 1\}$ を導入して, $x_{k,n} = \sum_{m=1}^{\infty} z_{m,n} h_{k,m} u_{m,n}$ とし, $z_{m,n}$ が 1 となる確率値 $\pi_{m,n} = p(z_{m,n} = 1)$ をベータ過程から生成させるようにすれば良い (Knowles and Ghahramani (2007), Liang *et al.* (2013)). ベータ過程で生成される変数はスパース (ほとんどが 0 に近い値) になる傾向があるため, 所与のデータを説明する上で必要な基底以外に対応するバイナリ変数 z は 0 になるよう推論結果が誘導される. 一方, ガンマ過程を用いる場合は, 基底行列と係数行列に対しスケールに何らかの制約を置いた (例えば各行列要素の期待値が 1 になるように事前分布を設定するなどした) 上で, 各基底の寄与の度合いを表した非負の連続量の変数 $\theta_m \in \mathbb{R}^{\geq 0}$ を導入して, $x_{k,n} = \sum_{m=1}^{\infty} \theta_m h_{k,m} u_{m,n}$ とし, θ_m をガンマ過程から生成させるようにすれば良い (Hoffman *et al.* (2010), Nakano *et al.* (2011)). ガンマ過程で生成される変数もスパース (ほとんどが 0 に近い値) になる傾向があるため, 所与のデータを説明する上で必要な基底以外に対応する非負変数は 0 になるよう推論結果が誘導される.

9. まとめ

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. より理解を深めたい読者は例えば Cichocki *et al.* (2009), 亀岡 (2012), 澤田 (2012) の著書や解説記事も是非とも参照して頂きたい.

参考文献

- Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003). Latent Dirichlet allocation, *Journal of Machine Learning Research* 3, J. Lafferty ed., 993–1022.
- Bregman, L. M. (1967). The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming, *USSR Computational Mathematics and Mathematical Physics*, 7(3), 210–217.
- Cemgil, A. T. (2008). Bayesian inference for nonnegative matrix factorization models, *Tech. Rep. CUED/F-INFENG/TR.609*, University of Cambridge, 2008.
- Cichocki, A., Zdunek, R., Phan, A. H. and Amari, S. (2009). *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, John Wiley & Sons.
- Csiszár, I. (1975). I-divergence geometry of probability distributions and minimization problems, *The Annals of Probability*, 3(1), 146–158.
- Dhillon, I. S. and Sra, S. (2006). Generalized nonnegative matrix approximations with Bregman divergences, *Advances in Neural Information Processing Systems 19 (NIPS 2006)*, 283–290.
- Durrieu, J.-L. and Thiran, J.-P. (2011). Sparse non-negative decomposition of speech power spectra for formant tracking, *Proc. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2011)*, 5260–5263.
- Durrieu, J.-L., Richard, G., David, B. and Févotte, C. (2010). Source/filter model for unsupervised main melody extraction from polyphonic audio signals, *IEEE Transactions on Audio, Speech and Language Processing*, 18(3), 564–575.
- Eguchi, S. and Kano, Y. (2001). Robustifying maximum likelihood estimation, *Technical report, Institute of Statistical Mathematics*, Research Memo, 802.

- Févotte, C., Bertin, N. and Durrieu, J.-L. (2009). Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis, *Neural Computation*, **21**(3), 793–830.
- Helén, M. and Virtanen, T. (2005). Separation of drums from polyphonic music using non-negative matrix factorization and support vector machine, *Proc. The 13th European Signal Processing Conference (EUSIPCO 2005)*.
- Hennequin, R., David, B. and Badeau, R. (2011). Beta-divergence as a subclass of Bregman divergence, *IEEE Signal Processing Letters*, **18**(2), 83–86.
- Higuchi, T., Takeda, H., Nakamura, T. and Kameoka, H. (2014). A unified approach for underdetermined blind signal separation and source activity detection by multichannel factorial hidden Markov models, *Proc. The 15th Annual Conference of the International Speech Communication Association (Interspeech 2014)*, 850–854.
- Hiroya, S. (2013). Non-negative temporal decomposition of speech parameters by multiplicative update rules, *IEEE Transactions on Audio, Speech and Language Processing*, **21**(10), 2108–2117.
- Hoffman, M., Blei, D. and Cook, P. (2010). Bayesian nonparametric matrix factorization for recorded music, *Proc. The 27th International Conference on Machine Learning (ICML 2010)*, 439–446.
- Hofmann, T. (1999). Probabilistic latent semantic analysis, *Proc. The 15th Conference on Uncertainty in Artificial Intelligence (UAI 1999)*, 289–296.
- Hunter, D. R. and Lange, K. (2000). Quantile regression via an MM algorithm, *Journal of Computational and Graphical Statistics*, **9**, 60–77.
- Hurmala, A., Gemmeke, J. and Virtanen, T. (2011). Non-negative matrix deconvolution in noise robust speech recognition, *Proc. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2011)*, 4588–4591.
- 板倉文忠 (1972). 『統計的手法による音声分析合成系に関する研究』博士論文, 名古屋大学大学院工学研究科.
- Kameoka, H. (2007). *Statistical Approach to Multipitch Analysis*, Ph.D. Thesis, The University of Tokyo.
- 亀岡弘和 (2012). 「非負値行列因子分解」『計測と制御』**51**(9), 835–844.
- Kameoka, H. and Kashino, K. (2009). Composite autoregressive system for sparse source-filter representation of speech, *Proc. 2009 IEEE International Symposium on Circuits and Systems (ISCAS 2009)*, 2477–248.
- 亀岡弘和, 後藤真孝, 嵯峨山茂樹 (2006). 「スペクトル制御エンベロープによる混合音中の周期および非周期成分の選択的イコライザ」『情報処理学会研究報告』2006-MUS-66-13, 77–84.
- 亀岡弘和, 小野順貴, 柏野邦夫, 嵯峨山茂樹 (2008). 「複素 NMF: 新しいスパース信号分解表現と基底系学習アルゴリズム」『日本音響学会 2008 年秋季研究発表会』2-8-13, 657–660.
- Kameoka, H., Nakatani, T. and Yoshioka, T. (2009). Robust speech dereverberation based on non-negativity and sparse nature of speech spectrograms, *Proc. 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009)*, 45–48.
- Kameoka, H., Nakano, M., Ochiai, K., Imoto, Y., Kashino, K. and Sagayama, S. (2012). Constrained and regularized variants of non-negative matrix factorization incorporating music-specific constraints, *Proc. 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2012)*, 5365–5368.
- Kitano, Y., Kameoka, H., Izumi, Y., Ono, N. and Sagayama, S. (2010). A sparse component model of source signals and its application to blind source separation, *Proc. 2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010)*, 4122–4125.
- Knowles, D. and Ghahramani, Z. (2007). Infinite sparse factor analysis and infinite independent components analysis, *Proc. The 7th International Conference on Independent Component Analysis and Signal Separation (ICA 2007), Lecture Notes in Computer Science*, **4666**, 381–388.
- Lee, D. D. and Seung, H. S. (2000). Algorithms for nonnegative matrix factorization, *Advances in Neural Information Processing Systems 13 (NIPS 2000)*, 556–562.
- Liang, D., Hoffman, M. D. and Ellis, D. P. W. (2013). Beta process sparse nonnegative matrix factorization for music, *Proc. The 14th International Society for Music Information Retrieval Conference (ISMIR 2013)*, 375–380.
- Nakano, M., Kameoka, H., Le Roux, J., Kitano, Y., Ono, N. and Sagayama, S. (2010). Convergence-guaranteed multiplicative algorithms for non-negative matrix factorization with beta-divergence, *Proc. 2010 IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2010)*, 283–288.

- Nakano, M., Le Roux, J., Kameoka, H., Nakamura, T., Ono, N. and Sagayama, S. (2011). Bayesian nonparametric spectrogram modeling based on infinite factorial infinite hidden Markov model, *Proc. 2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA2011)*, 325–328.
- Ortega, J. M. and Rheinboldt, W. C. (1970). *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press, New York.
- Ozerov, A. and Févotte, C. (2010). Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation, *IEEE Transactions on Audio, Speech and Language Processing*, **18**(3), 550–563.
- Ozerov, A., Févotte, C. and Charbit, M. (2009). Factorial scaled hidden Markov model for polyphonic audio representation and source separation, *Proc. 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2009)*, 121–124.
- Paatero, P. and Tapper, U. (1994). Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics*, **5**, 111–126.
- Parry, R. M. and Essa, I. (2007). Phase-aware non-negative spectrogram factorization, *Proc. The 7th International Conference on Independent Component Analysis and Signal Separation (ICA)*, 536–543.
- Raczyński, S. A., Ono, N. and Sagayama, S. (2007). Multipitch analysis with harmonic nonnegative matrix approximation, *Proc. 8th International Conference on Music Information Retrieval (ISMIR 2007)*, 381–386.
- 澤田宏 (2012). 「非負値行列因子分解 NMF の基礎とデータ／信号解析への応用」『電子情報通信学会学会誌』**95**(9), 829–833.
- Sawada, H., Kameoka, H., Araki, S. and Ueda, N. (2011). New formulations and efficient algorithms for multichannel NMF, *Proc. 2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA2011)*, 153–156.
- Sawada, H., Kameoka, H., Araki, S. and Ueda, N. (2012). Efficient algorithms for multichannel extensions of Itakura-Saito nonnegative matrix factorization, *Proc. 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2012)*, 261–264.
- Schmidt, M. N., Winther, O. and Hansen, L. K. (2009). Bayesian non-negative matrix factorization, *Proc. The 10th International Conference on Independent Component Analysis and Blind Signal Separation (ICA 2009)*, 540–547.
- Smaragdis, P. (2004). Non-negative matrix factor deconvolution; extraction of multiple sound sources from monophonic inputs, *Proc. The 5th International Conference on Independent Component Analysis and Blind Signal Separation (ICA 2004)*, 494–499.
- Smaragdis, P. and Brown, J. C. (2003). Non-negative matrix factorization for music transcription, *Proc. 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2003)*, 177–180.
- Smaragdis, P. and Raj, B. (2007). Example-driven bandwidth expansion, *Proc. 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2007)*, 135–138.
- Smaragdis, P., Raj, B. and Shashanka, M. (2006). A probabilistic latent variable model for acoustic modeling, *Advances in Neural Information Processing Systems 19 (NIPS 2006)*.
- Smaragdis, P., Raj, B. and Shashanka, M. V. (2007). Supervised and semi-supervised separation of sounds from single-channel mixtures, *Proc. The 7th International Conference on Independent Component Analysis and Signal Separation (ICA 2007)*, 414–421.
- 戸上真人, 川口洋平 (2009). 「準ブラインド非負行列分解を用いたマルチチャンネル非線形エコーキャンセラ」『日本音響学会 2009 年秋季研究発表会音講論』3-Q-9, 757–758.
- Virtanen, T. (2007). Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria, *IEEE Transactions on Audio, Speech and Language Processing*, **15**(3), 1066–1074.
- Yoshii, K. and Goto, M. (2012a). A nonparametric Bayesian multipitch analyzer based on infinite latent harmonic allocation, *IEEE Transactions on Audio, Speech and Language Processing*, **20**(3), 717–730.
- Yoshii, K. and Goto, M. (2012b). Infinite composite autoregressive models for music signal analysis, *Proc. The 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, 79–84.