

スパース表現と凸最適化に基づく非負値行列因子分解と音楽信号処理への応用

Nonnegative Matrix Factorization Based on Sparse Representation and Convex Optimization: An Application to Music Signal Processing



湯川 正裕 (Masahiro YUKAWA, Dr. Eng.)

慶應義塾大学 理工学部 電子工学科 准教授

(Associate Professor, Department of Electronics and Electrical Engineering, Keio University)

電子情報通信学会 IEEE EURASIP

受賞：船井学術賞 (2015 年度) APSIPA Annual Summit and Conference 2015 Best Paper Award (2015 年度) KDDI 財団賞 (2014 年度) 科学技術分野の文部科学大臣表彰若手科学者賞 (2014 年度) 電気通信普及財団テレコムシステム技術賞 (2013 年度) 電子情報通信学会学術奨励賞 (2009 年度) エリクソン・ヤングサイエンティストアワード (2009 年度) 丹羽保次郎記念論文賞 (2006 年度) 電子情報通信学会論文賞 (2005 年度)
研究専門分野：信号処理 最適化 情報通信

あらまし 非負値行列因子分解 (NMF) は、音響信号や画像、文書データなどから頻出するパターンを抽出する技術として盛んに研究されてきた。しかし、NMF は一般に、(i) 分解の非一意性、(ii) 目的関数の非凸性、(iii) パターン数推定誤差への性能依存性の問題を抱えている。本研究では、これらの問題点を解決すべく、目的関数を損失関数 (Kullback-Leibler/Dual-Itakura-Saito ダイバージェンス)、非負値制約を満たすための指示関数、複数の正則化項の和として設計し、これを効率的に最適化する手法を与えた。正則化項には、全変動、ブロック L1 ノルム、(通常の) L1 ノルムが含まれる。ダイバージェンス関数の滑らかな近似として Moreau エンベロープを導入するとともに、「指示関数、ブロック L1 ノルム、L1 ノルムの和の近接写像が閉形式で与えられる」ことを示した。これにより、補助変数フリーな Proximal Forward-Backward Splitting (PFBS) 法を用いて、近似問題の解を求めることができる。計算機シミュレーションにより、提案法の有効性を例証した。

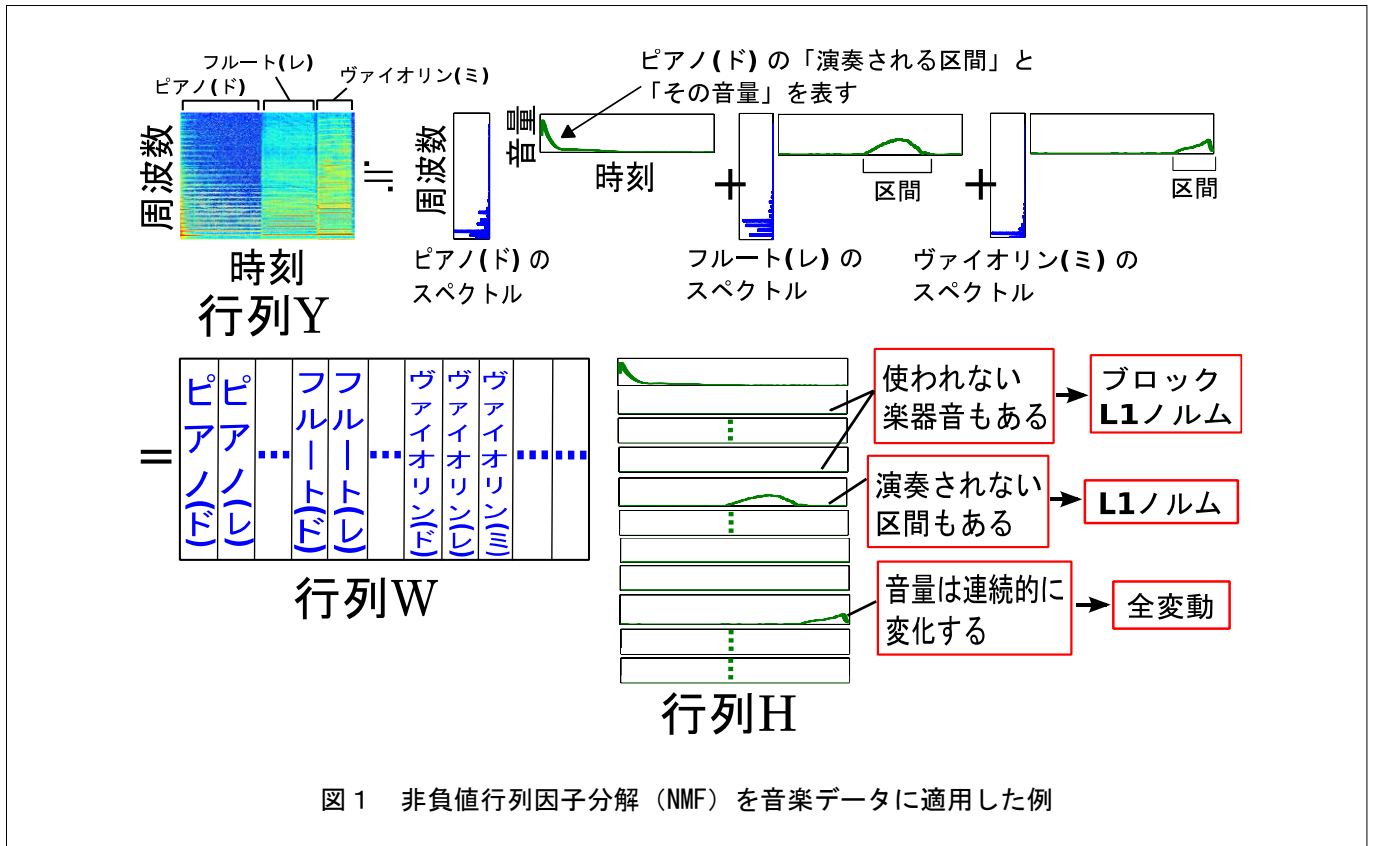
1. 研究背景と目的

非負値行列因子分解 (NMF) は、非負値行列 Y (各成分が 0 以上の実数値を取る行列) を 2 つの非負値行列 W と H の積に分解する数理問題である。この問題は、遅くとも 1970 年代には研究されていたが、20 世紀の終わりに学術誌 *Nature* に掲載された Lee, Seung の論文が火種となって広く研究されるようになった[1]。図 1 に、音楽データ解析への応用例を示す。行列 Y の列ベクトルは、楽曲を短い時間窓で短時間フーリエ変換した (局所) 周波数スペクトルの振幅 (または電力) を表している。簡単に言えば、楽曲に含まれる局所的な周波数成分の時間変化を図示している。これをスペクトログラムという。図の例では、3 つの楽器音、つまりピアノ (ド)、フルート (レ)、ヴァイオリン (ミ) が順に演奏される。右辺の第一項は、ピアノ (ド) に対応する音を「スペクトル (周波数情報)」と「音量変化」に分解したものである (楽器音にはハーモニクスと呼ばれる倍音 (高調波) 成分が含まれるため、複数のピークが立つようなスペクトルが見られる)。ピアノ (ド) は、前半部分のみ演奏されるため、後半部の音量はゼロとなっているのが分かる。この右辺のような分解を得ることが NMF の目的である。図の例では、分かりやすさのため単音の場合を描いているが、一般には、複数の楽器音が混ざった混合信号に対応するスペクトログラムが初めに得られる。しかし、複数のスペクトルが混ざった複雑なスペクトログラムを見ても、意味のある情報は得られにくい。右辺のように分解することで、自動採譜 (楽譜を自動生成する技術) や音源分離 (楽器音毎の音源に分離する技術) に役立てることができる。この右辺の分解を行列表現したものが、その下に描かれている。

各楽器音のスペクトル (列ベクトル) をまとめたものが行列 W 、音量変化 (行ベクトル) をまとめたものが行列 H である。本研究では冗長な行列 W を用いるため、実際には演奏されない楽器音に該当するスペクトルも W に含まれている (理由は後述する。) 同原理より、顔画像をパーツ (目、鼻、口、耳など) に分解できるという例も報告されている。他にも、文書・生体信号・遺伝子など、幅広いデータの解析に NMF を応用できるため、盛んに研究されてきた。

スパース表現と凸最適化に基づく非負値行列因子分解と音楽信号処理への応用

Nonnegative Matrix Factorization Based on Sparse Representation and Convex Optimization:
An Application to Music Signal Processing



NMF のアプローチは、教師なし学習と教師付き学習に大別される。教師なし学習では W と H が両方とも未知であると仮定するのに対して、教師付き学習では (予め、各楽器音単独のスペクトルの情報が入手可能である状況を想定して) W が既知であると仮定して H を推定する。教師なし学習は、後述のように非凸性に起因する問題や W の列数決定の正確性が求められるなど、様々な問題点を抱える。いずれにアプローチも、「適切な目的関数」と「最適化アルゴリズム」の設計が問題解決の鍵となる。

Y と WH の誤差を図る損失関数としては、ユークリッド距離を用いた 2 乗誤差関数の他、Kullback-Leibler (KL) ダイバージェンスや板倉斎藤 (IS) ダイバージェンスなどが用いられてきた。最適化アルゴリズムは、Lee, Seung らによって提案された Multiplicative updates (MU) [1]が広く用いられている。この手法は、非負値の実数を乗算することで係数を更新する (実数の減算によって係数を更新する最急降下法と対照的である)。初期推定行列を非負値に選んでおけば、非

負値性が自動的に担保される。このアプローチの基本原理は Majorization Minimization (MM) アルゴリズム (例えば、Hunter, Lange らの論文参照方[2]) であり、上界関数を求めるのに骨が折れる場合がある。他にも、Projected Gradient Method (PGM) や Proximal Forward-Backward Splitting (PFBS) Method, Alternating Direction Method of Multipliers (ADMM) などを用いた研究例も見られるが、NMF 研究の総数と比べて相当に限定的である。

NMF は幾つかの問題点を抱えている。以下、問題点を整理する。

- (1) 一般に解の一意性が保証されない (実際、強い条件の下、並び換えと定数倍の任意性を除いて、一意に分解できることが知られている [3]-[6])。
- (2) 非凸最適化問題として一般に定式化されるため、大域的な最適化が困難であり、初期値依存性や局所解 (または停留点) に収束してしまうなどの問題が生じる (同問題は NP 困難であることが知られている [7])。

スパース表現と凸最適化に基づく非負値行列因子分解と音楽信号処理への応用

Nonnegative Matrix Factorization Based on Sparse Representation and Convex Optimization: An Application to Music Signal Processing

(3) 行列 W の列数 (音楽データの場合、音源数に対応) の推定が重要であり、推定が不正確である場合、深刻な性能劣化を招く要因となる。

これらの問題点を解消するための新しい学習パラダイムの構築を目的として研究を行なった。本稿では、本研究で得られた主要成果の概要を述べる。基本的なアイデアは、「冗長な」行列 W (例えば、想定し得る全ての楽器音の周波数スペクトル) を用意し、NMF を「行列 H に関するスパース最適化問題」として定式化することである。ここでいう「想定し得る全ての楽器音」とは、必ずしも実際に楽曲で用いられているものと同じのものとは限らない (例えば、楽曲は市販の CD などのデータであり、 W の列ベクトルは別のデータベースの音源を用いる状況を想定している)。この点が、他の教師付き NMF とは異なる。

行列 H を高精度で推定するために、 WH と行列 Y の合致性を測る損失関数 (データ忠実項) と非負値制約に加えて、3 つの罰則項 (ブロック $L1$ ノルム、 $L1$ ノルム、全変動) を導入する (図 1 参照)。これにより、非負値行列因子分解は「微分可能な関数」と「複数の微分不可能な関数」の和の最小化問題として定式化される。損失関数は、2 乗誤差、KL ダイバージェンスに加えて、「双対板倉斎藤 (Dual-IS) ダイバージェンス」を検討した。ブロック $L1$ ノルムと $L1$ ノルムの利用はスパース最適化に関係しているので、興味のある読者は、拙稿 [8] とその参考文献を参照されたい。次節以降で、その詳細について述べ、自動採譜への応用で得られた結果を紹介する。

2. 複数の正則化項を用いた凸最適化アプローチ

幾つかの研究を経て、最終的に行きついた結果を学術論文 (Yukawa, Kagami, 投稿中 [14]) にまとめた。以下、その内容を説明する中で、途中経過で得られた成果についても触れる。

損失関数は、前節で述べたように、2 乗誤差、KL ダイバージェンス、Dual-IS ダイバージェンスの 3 種類を検討した。理由は、全て凸関数であるため、大域的最適化が可能であることと、最適化する際に利用する近接写像が計算できるからである。ここで、近接写

像とは劣微分のレゾルベントであり、滑らかでない (一般に不連続な) 凸関数の最適化における強力なツールとして現代の最適化数学で頻繁に利用されている (詳しくは、例えば、文献 [19])。前節で述べた問題点 (1) (分解の非一意性問題) に対処する手段として、先験情報の利用が有効であると考えられる。音楽データの場合、音量は連続的に変化するため、全変動 (時間的に隣接する係数同士の差分の 2 乗和) を正則化項として目的関数に加える。また、 W に含まれるスペクトルと Y に含まれる実際の楽器音のスペクトルの誤差や雑音の影響により、実際の楽器音のスペクトルと似た別のスペクトルが誤って得られてしまうことが実験で明らかになった。このため、正しいスペクトルを得るための手段として、行列 H の各行ベクトルのユークリッドノルム和 (ブロック $L1$ ノルムという) を第二の正則化項として加える。さらに、各楽器音は楽曲の中で部分的に演奏されるため、無音区間に誤差が生じないように通常の $L1$ ノルムも第三の正則化項として加える。指示関数 (非負値であれば 0、そうでなければ無限大を取る関数) を目的関数に加えることで、非負値性が保証される。

まとめると、NMF 問題は、以下の目的関数の最小化問題として定式化される。

損失関数 (2 乗誤差、KL ダイバージェンス、Dual-IS ダイバージェンスなど) + 全変動 + 指示関数 + ブロック $L1$ ノルム + $L1$ ノルム

KL ダイバージェンスと Dual-IS ダイバージェンスは、正の領域のみを定義域とすることが多いが、凸最適化の枠組みで扱いやすくするため、0 以下の領域で無限大を取るものとして、定義域を実数全体に拡張する。値域は実数全体の集合に無限大 ($+\infty$) を加えた集合である。このように定義すると、目的関数に含まれる全ての項は、(ユークリッド空間全体で定義された) 下半連続な真凸関数である。ここで、下半連続とは、関数のグラフを任意のレベルで横にスパッと切った切り口が閉集合になるという性質である。正確には、全ての実数 α に対して、 $f(x) \leq \alpha$ を満たす x 全体の集合 (レベル集合と呼ばれる) が閉集合になるとき、関数 $f(x)$ は下半連続であるという。レベル集合が常に凸

スパース表現と凸最適化に基づく非負値行列因子分解と音楽信号処理への応用

Nonnegative Matrix Factorization Based on Sparse Representation and Convex Optimization: An Application to Music Signal Processing

集合である（集合内の任意の2点を結ぶ線分がその集合に完全に含まれる）とき、関数 f は凸関数であるという。特に、関数 $f(x)$ が少なくとも一点で実数値を取るとき、 $f(x)$ は真凸関数であるという。

まず初めに、損失関数が2乗誤差の場合を検討した。損失関数 + 全変動は滑らかであり、指示関数、ブロック L1 ノルム、L1 ノルムのそれぞれの近接写像は容易に計算できる。実際、指示関数の近接写像は非負値領域への距離射影であり、行列の各成分に対して、それが負値であれば0にする、0以上の値であれば、そのままの値に保つという操作で計算できる。ブロック L1 ノルムの近接写像は、各行ベクトルの長さ(ノルム)を一定値 λ (>0) だけ縮め、元の長さが λ 以下であれば、0ベクトルとするという操作で計算できる。L1 ノルムの近接写像は、よく知られる shrinkage 作用素であり、各成分の大きさから一定値 γ (>0) 減じ、元の絶対値が γ 以下であれば0値とするという操作で計算できる。従って、この場合の最適化問題は、Generalized Forward-Backward Splitting (GFBS) Method を用いて解くことができる。GFBS は一般に、「滑らかな関数」と複数の「近接写像が計算可能な関数」の和の最小解に収束する点列を生成する逐次アルゴリズムである。ここで、近接写像が計算可能な関数は滑らかでなくてもよく、もっといって不連続であってもよい（指示関数は不連続関数である）。これが、Morikawa, Yukawa のアプローチ[9]である。さらに、音源が鳴っている区間を推定して、これに基づく局所的なブロック L1 ノルムに変更することで、1オクターブ上の音階に誤って採譜してしまうというエラーを抑止できることを示した。これが Morikawa, Yukawa, Kikuchi のアプローチ[10]である (APSIPA-ASC 2015 Best Paper Award を受賞)。

次に、KL ダイバージェンスと Dual-IS ダイバージェンスを目的関数として利用することを検討した。この場合、目的関数は滑らかでないため、GFBS をそのままの形で適用することはできない。目的関数は「近接写像が計算可能な関数」と「線形写像」の合成写像となっており、標準的な凸最適化アルゴリズム ADMM で解くことができる。これが、Kagami, Yukawa のアプローチ[11]である。

ADMM はポピュラーな手法であるが、補助変数を利用するため、メモリ効率が低い。補助変数を利用しない PFBS で解くことができれば、もっと効率良く問題を解くことができる。これを可能にする鍵は2つある。1つは、滑らかでない目的関数 (KL ダイバージェンスと Dual-IS ダイバージェンス) を Moreau エンペロープ[19]で近似することである。元の関数が下半連続な真凸関数であれば、Moreau エンペロープはいつでも滑らかな関数となり、勾配方向は近接写像を適用する前後の点の差分ベクトルで与えられる。KL ダイバージェンスと Dual-IS ダイバージェンスの近接写像は、閉形式で与えられるため、これらの Moreau エンペロープの勾配は簡単に計算できる。従って、2乗誤差のときと同様、損失関数 + 全変動は滑らかな関数となる。2つ目の鍵は、目的関数に含まれる残りの項（すなわち指示関数 + ブロック L1 ノルム + L1 ノルム）を一つの関数とみたときの近接写像が計算可能であることを示したことである。これらにより、「KL ダイバージェンス (または Dual-IS ダイバージェンス) の Moreau エンペロープ + 全変動」に対する最急降下シフト写像 (Forward ステップ) と、「指示関数 + ブロック L1 ノルム + L1 ノルム」の近接写像 (Backward ステップ) を交互に繰り返す PFBS によって、近似問題の解に収束する点列を生成することが可能となった。これが、学術論文 (Yukawa, Kagami 投稿中[14]) に記したアプローチである。

シミュレーション結果を図2~図4に示す。

図2は、合成データを用いた実験結果である。行列 W は一様分布からランダムに生成した。 H^* をスパース (全成分の50%が零値) かつ行スパース (行ベクトルのうち、20%が零ベクトル) な行列とし、ランダム雑音 N を加えて $Y = WH^* + N$ とした。 Y と W から H^* を推定し、推定誤差 $H - H^*$ の各成分の2乗和 (フロベニウスノルムの2乗) を評価した。ADMM は、提案法と同様の目的関数 (補助変数を利用して) 近似なしで最小化した場合の性能を示している[11]。MU と PGA は、損失関数を非負値制約の下で最小化した場合の性能を示している。KL ダイバージェンス規準、Dual-IS ダイバージェンス規準を用いた場合、提案法の優位性が確認できる。

スパース表現と凸最適化に基づく非負値行列因子分解と音楽信号処理への応用

Nonnegative Matrix Factorization Based on Sparse Representation and Convex Optimization: An Application to Music Signal Processing

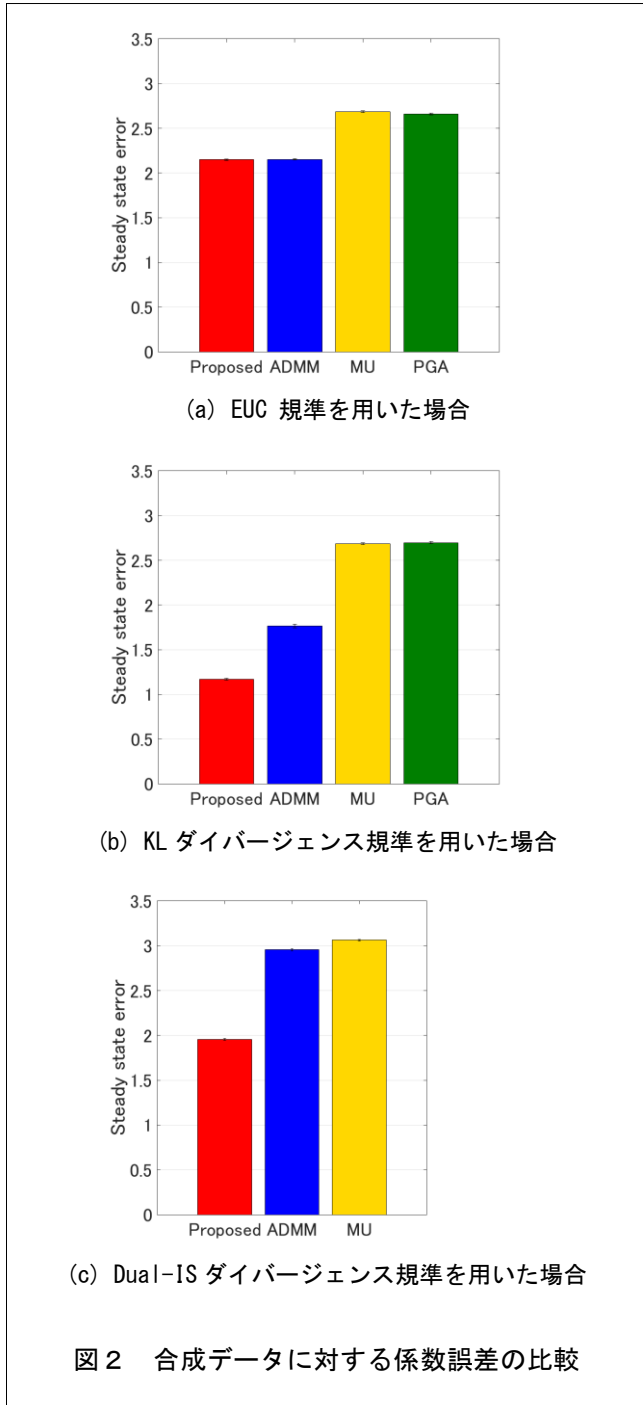
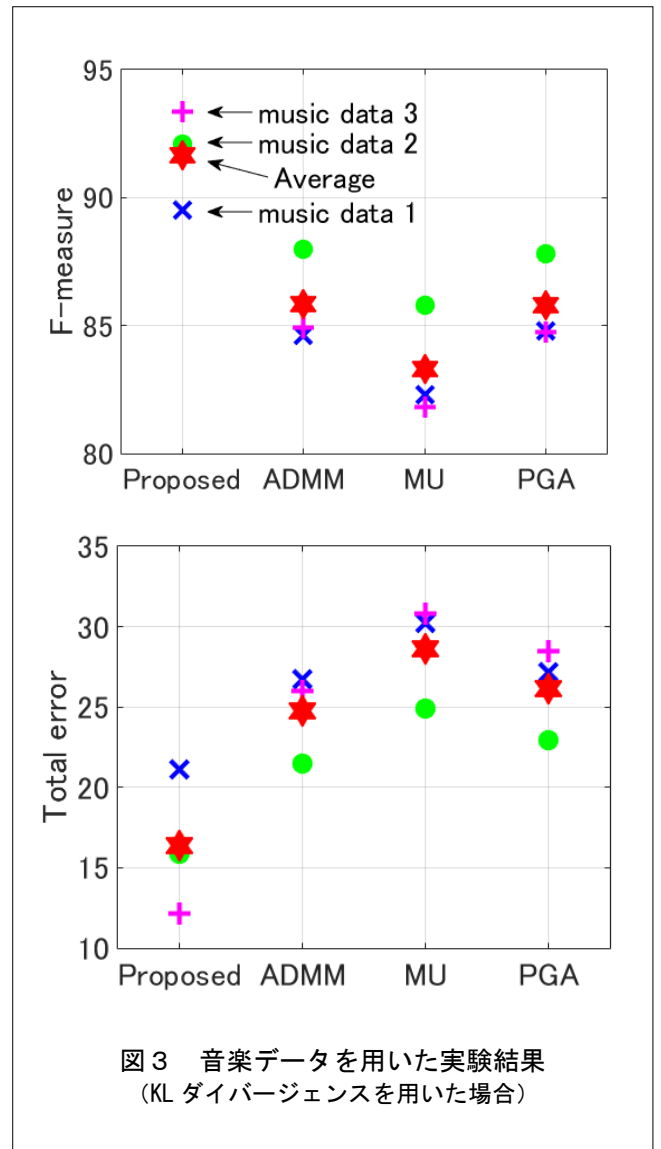


図3及び図4は、異なる3つの音楽データを用いて自動採譜の実験を行なった結果を図示している。F-measure が高いと高精度な自動採譜が達成できたことを表しており、Total error は低い方が良い。いずれのダイバージェンス規準を用いた場合も、提案法の優位性が確認できる。

Remark : ダイバージェンス関数を Moreau エンベロープで近似しているが、 W に含まれるスペクトルと実際の楽曲のスペクトルの間に誤差が存在すること、環境雑音が行列 Y に含まれることなどから、(近似なしの) ダイバージェンス関数を厳密に最小化することにどれだけの意義があるかという点に疑問が生じる。目的関数が滑らかな関数になったこと、補助変数を用いないことによって余計な誤差(双対変数の誤差が主変数に与える間接的誤差)を回避できることによる恩恵の方が近似誤差の影響よりも大きく、それがシミュレーション結果に示された高い性能に繋がったと考えられる。教師なし学習は、初期値への依存性が強く、音源数が既知である必要があるため、教師付き学習が有効なアプローチであるのは間違いない。NMFを種々の先験情報を取り入れた凸最適化問題として定式化し、最適解へ収束する点列を生成する数値解法を構築できたのは、大きな前進であるといえよう。



スパース表現と凸最適化に基づく非負値行列因子分解と音楽信号処理への応用

Nonnegative Matrix Factorization Based on Sparse Representation and Convex Optimization: An Application to Music Signal Processing

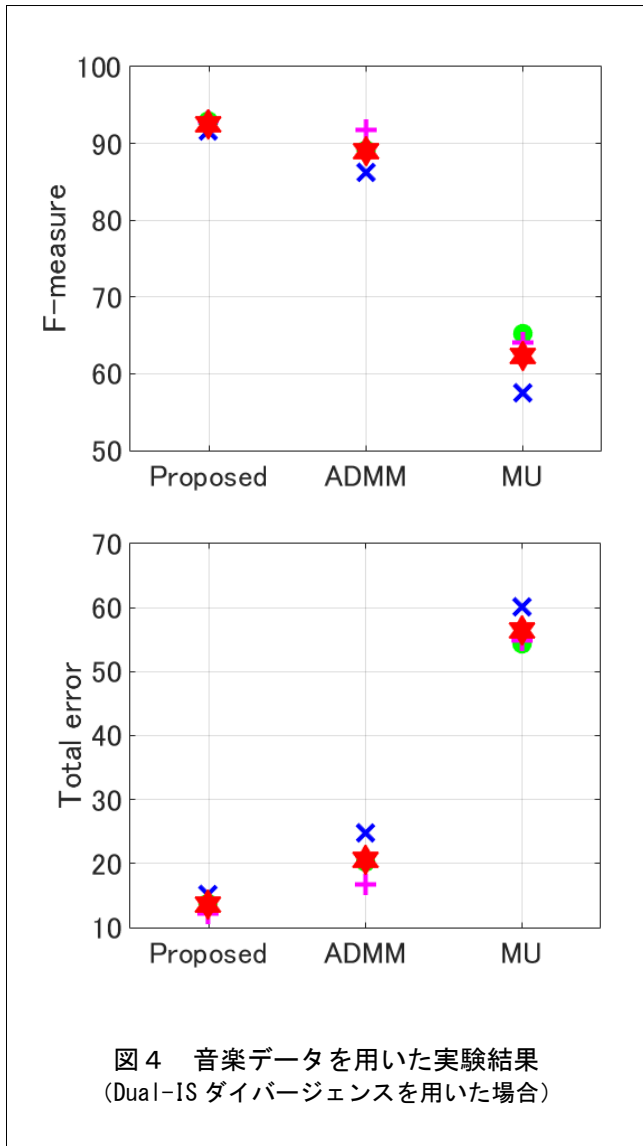


図4 音楽データを用いた実験結果
(Dual-IS ダイバージェンスを用いた場合)

3. 複素 NMF と時領域域アプローチ

前節で述べたアプローチとは異なる 2 つのアプローチについて述べる。

3.1. KL ダイバージェンス規準複素 NMF

第 1 節で紹介した音楽データに対する NMF では、スペクトルの加法性を仮定していた。しかし、振幅スペクトルやパワースペクトルは加法的ではない。実際に加法性が成り立つのは、時間領域もしくは複素スペクトル領域である。このため、NMF により観測スペクトログラムを加法的な成分に分解したとしても、必ずしも各音源に対応したスペクトログラムが得られる

保証はない。

この問題を解決するため、複素 NMF と呼ばれる枠組みが提案されている[15]。複素 NMF では、NMF における各スペクトルテンプレートに位相スペクトルを時変パラメータとして付加したモデルを用いている。これにより、NMF と類似した考え方により複素スペクトル領域での信号分離が可能になった。従来、複素 NMF におけるパラメータ推定問題 (NMF における行列分解に相当) は、観測複素スペクトログラムと複素スペクトログラムモデルとの 2 乗誤差 ($Y-WH$ の各成分の絶対値の 2 乗和) を規準とした最適化問題として定式化されており、効率的なパラメータ推定アルゴリズムが提案されている。2 乗誤差以外の規準でのパラメータ推定アルゴリズムは、まだ提案されていなかった。これは、複素 NMF で考えている問題が複素数値同士の乖離度を評価する必要があるためである。一方、NMF では、2 乗誤差、KL ダイバージェンス、板倉斎藤擬距離を規準とした場合のパラメータ推定アルゴリズムがそれぞれ提案されているが、経験的に KL ダイバージェンスを規準とした場合に高い音源分離性能が得られることが知られている[16]。このことから、複素 NMF においても、KL ダイバージェンスなどの 2 乗誤差以外の規準を用いた場合のパラメータ推定アルゴリズムを導出することができれば、より高い音源分離性能が得られる可能性がある。

そこで、まず 2 乗誤差規準の複素 NMF を解釈し直し、これまで提案されていた複素 NMF のパラメータ推定アルゴリズムが暗に最適化していた問題を双対形式と定義した。双対形式を用いても、既存の複素 NMF と同一のアルゴリズムが導かれる。ただし、複素数同士の距離が「線形制約下における実数同士の距離」に置き換わるため、2 乗誤差規準を KL ダイバージェンス規準で代用することが可能となる。これがポイントである。実験的に、2 乗誤差規準の複素 NMF の分離性能を上回ることを示した[12][13][17]。

3.2 時間領域低ランクスペクトログラム近似法

3.1 節で述べた通り、従来の NMF では、実際には成り立たない加法性を仮定している。複素 NMF では、加法性の問題を解決したものの、実はまだ問題が残っ

スパース表現と凸最適化に基づく非負値行列因子分解と音楽信号処理への応用

Nonnegative Matrix Factorization Based on Sparse Representation and Convex Optimization: An Application to Music Signal Processing

ている。複素 NMF では、位相スペクトログラムの各要素が独立なパラメータとして扱われているが、実際の時間周波数成分の位相は制約があり、互いに相関がある。なぜなら、時間周波数変換によって得られたスペクトログラムは、原信号の冗長表現になっているからである。例えば、STFT では信号の短時間フレームをオーバーラップさせてフーリエ変換し、それらを連結することでスペクトログラムを得ている。従って、得られたスペクトログラムの全ての要素は、オーバーラップした区間に含まれる波形の一貫性条件を満たす必要がある。複素 NMF は、このような冗長性を考慮していなかった。さらに、変換された信号の複素スペクトログラムは、元々の複素スペクトログラムと一致しないため、この冗長性を考慮しない限り、推定された複素スペクトログラムから生成された時間領域信号は一般に最適とは言えない。

スペクトログラムの冗長性をうまく扱うため、NMF に類する分離を時間領域で可能にする手法として、時間領域スペクトログラム分離法 (Time-domain Spectrogram Factorization : TSF) と呼ばれる手法が提案されている。TSF の最適化問題の定式化は、前節で述べた複素 NMF の「双対形式」に基づいており、「双対形式」における複素 NMF の推定パラメータを時間領域信号と時間周波数変換基底の積とすることで導かれる。このフレームワークでは、振幅スペクトログラムの基底を得るという特性を保持しつつ、構成音の時間領域信号を直接的に分離できる (既存の NMF、複素 NMF では、推定されたスペクトログラムから時間信号を取り出す後処理が必要である)。しかし、従来の TSF アルゴリズムでは、大規模行列の逆行列を求めるための計算時間が大きく、結果として、十分な性能を得ることができなかった。この課題を解決するため、大規模行列の逆行列演算を必要としない、補助関数法と射影勾配法ベースのアプローチを提案した。特に、時間周波数変換を STFT としたときのアルゴリズムを導き、分離性能が向上することを実証した [18]。

4. 将来の展望

本研究で開発した手法が、音楽データの自動採譜・音源分離のための基礎技術として広く利用されること

を期待する。画像や文書データなどに対しても、W を予め用意するための情報を取得できるので、同様のアプローチが適用できるであろう。今後もより精度を向上させ、応用範囲の裾野を広げていけるよう研究を続けたい。

参考文献

- [1] D. D. Lee, H. S. Seung, Learning the parts of objects by nonnegative matrix factorization, *Nature* 401 (1999) 788–791.
- [2] D. R. Hunter, K. Lange, A tutorial on MM algorithms, *The American Statistician* 58 (1) (2004) 30–37.
- [3] D. L. Donoho, V. C. Stodden, When does non-negative matrix factorization give a correct decomposition into parts?, in: *Advances in Neural Information Processing Systems (NIPS)*, Vol. 16, MIT Press, Cambridge, MA, USA, 2003, pp. 1141–1148.
- [4] S. Moussaoui, D. Brie, J. Idier, Non-negative source separation: Range of admissible solutions and conditions for the uniqueness of the solution, in: *Proc. ICASSP*, Vol. 5, 2005, pp. 289–292.
- [5] H. Laurberg, M. G. Christensen, M. D. Plumbley, L. K. Hansen, S. H. Jensen, Theorems on positive data: On the uniqueness of NMF, *Comput. Intell. Neurosci.* 2008.
- [6] K. Huang, N. D. Sidiropoulos, A. Swami, Non-negative matrix factorization revisited: uniqueness and algorithm for symmetric decomposition, *IEEE Trans. Signal Processing* 62 (1) (2014) 211–224.
- [7] S. A. Vavasis, On the complexity of nonnegative matrix factorization, *Siam J. Optim.* 20 (3) (2009) 1364–1377.
- [8] 湯川正裕、音響適応信号処理と最適化：スパース性の活用、*日本音響学会誌*, 71 巻 11 号, pp.615–623, 2015.

スパース表現と凸最適化に基づく非負値行列因子分解と音楽信号処理への応用

Nonnegative Matrix Factorization Based on Sparse Representation and Convex Optimization: An Application to Music Signal Processing

- [9] Yu Morikawa and Masahiro Yukawa, "A sparse optimization approach to supervised NMF based on convex analytic method," in Proceedings of 38th IEEE ICASSP, pp.6078--6082, Vancouver: Canada, May 2013.
- [10] Yu Morikawa, Masahiro Yukawa, and Hisakazu Kikuchi, "Supervised Nonnegative Matrix Factorization Using Active-Period-Aware Structured L1-Norm for Music Transcription," in Proceedings of APSIPA Annual Summit and Conference, pp.14--18, December 2015. [Award Winning]
- [11] Hideaki Kagami and Masahiro Yukawa, "Supervised Nonnegative Matrix Factorization with Dual-Itakura-Saito and Kullback-Leibler Divergences for Music Transcription," in Proceedings of EUSIPCO, Budapest: Hungary, pp. 1138--1142, August--September 2016.
- [12] Hideaki Kagami, Hirokazu Kameoka, and Masahiro Yukawa, "A majorization-minimization algorithm with projected gradient updates for time-domain spectrogram factorization," in Proceedings of 42nd IEEE ICASSP, pp.561--565, March 2017.
- [13] Hirokazu Kameoka, Hideaki Kagami, and Masahiro Yukawa, "Complex NMF with the generalized Kullback-Leibler divergence," in Proceedings of 42nd IEEE ICASSP, pp.56--60, March 2017.
- [14] Masahiro Yukawa and Hideaki Kagami, "Supervised Nonnegative Matrix Factorization via Minimization of Regularized Moreau-Envelope of Divergence Function with Application to Music Transcription", submitted for publication.
- [15] H. Kameoka, N. Ono, K. Kashino, and S. Sagayama, "Complex NMF: A new sparse representation for acoustic signals," in Proc. IEEE ICASSP, April 2009, pp. 3437--3440.
- [16] D. Fitzgerald, M. Cranitch, and E. Coyle, "On the use of the beta divergence for musical source separation," in Proc. Irish Signals Syst. Conf., 2009.
- [17] 鏡英章・亀岡弘和・湯川正裕, "I ダイバージェンス規準複素 NMF," 音講論 (秋), pp. 433--436, 2016.
- [18] Hirokazu Kameoka, Hideaki Kagami, and Masahiro Yukawa, "Complex NMF with the generalized Kullback-Leibler divergence," in Proceedings of 42nd IEEE ICASSP, pp.56--60, March 2017.
- [19] Isao Yamada, Masahiro Yukawa, and Masao Yamagishi, "Minimizing the Moreau envelope of nonsmooth convex functions over the fixed point set of certain quasi-nonexpansive mappings," in Fixed-Point Algorithms for Inverse Problems in Science and Engineering (H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, eds.), pp. 345--390, Springer, 2011.

この研究は、平成24年度SCAT研究助成の対象として採用され、平成25～27年度に実施されたものです。