



卒業研究論文

論文題目

正則化非負値行列因子分解による
非負低ランク行列補完

| | |
|----------|---------------|
| 提出年月日 | 令和 6年 2月 22日 |
| 学 科 | 電 気 情 報 工 学 科 |
| 氏 名 | 和 気 佑 弥 印 |
| 指導教員(主査) | 北 村 大 地 講 師 印 |
| 副 査 | 柿 元 健 准 教 授 印 |
| 学 科 長 | 漆 原 史 朗 教 授 印 |

香川高等専門学校

Nonnegative low-rank matrix completion using regularized nonnegative matrix factorization

Yuya Wake

Department of Electrical and Computer Engineering
National Institute of Technology, Kagawa College

Abstract

Real-world data often include missing values, resulting from collection or subsequent processing, influenced by various factors such as human errors or technical limitations. As the dataset size increases, missing values become more prevalent. Directly handling missing values in data analysis is often impractical, leading to reduced usable data size. Imputation becomes crucial for effective utilization of the dataset. In particular, matrix completion methods are essential and have widely been investigated. Nonnegative matrix factorization (NMF) is a matrix decomposition algorithm that focuses on the negativity of real-world data. NMF is an unsupervised learning and shows high matrix completion accuracy resulting in a cost-reduced application compared to deep-learning-based or other supervised approaches. The matrix decomposition property in NMF can be controlled by introducing regularization. However, the suitable regularization type in matrix completion has not extensively been explored. In this thesis, I investigate the performance of NMF-based matrix completion using two regularization types; cosine similarity regularization and L_1 -norm regularization. The former emphasizes orthogonality between basis vectors, while the latter induces sparsity of the coefficient matrix in NMF. Experiments are conducted using artificial random matrices and audio spectrograms with missing values, and the performance of the matrix completion is compared. Results indicate that the proposed methods achieve a significant accuracy improvement for the matrix completion. Both regularizations provide stable estimation of missing values even when the number of basis vectors in NMF and the oracle rank of the input matrix are mismatched. In the audio spectrogram experiment, artificial distortions due to the missing values are mitigated, which is substantial enhancement compared to simple NMF without regularization.

Keywords: nonnegative matrix factorization, matrix completion, missing value, regularization

(和訳)

実世界で取り扱われるデータには、収集段階、またはその後のデータ処理により、データ内に欠損値が発生する。欠損値の発生要因は人為的なミスや技術的な限界など多様であり、データサイズが大きくなるほどその数は増す。多くの場合、データ解析において欠損値をそのまま扱うことはできず、結果として使用可能なデータサイズが減少してしまう。このため、情報資源を有効活用する上で欠損値に対し本来あるべき適切な値を推定して置き換える補完操作、特に対象データが 2 次元データであれば行列補完が重要であり、様々な手法が提案されている。行列補完手法の一つとして、現実において取り扱うべきデータがもつ非負性に着目した非負値行列因子分解 (NMF) と呼ばれる行列分解アルゴリズムがある。NMF は事前学習不要のアルゴリズムでありながら高精度な補完性能を持ち、コスト面で深層学習やその他教師あり学習を必要とする手法に勝る。NMF は正則化によって分解行列の特性が変化することが知られており、様々な正則化が提案されている。一方で、これらの正則化の行列補完への有効性はあまり検討されていない。本研究では、基底行列間の直交化を目的としたコサイン類似度正則化及び係数行列要素のスパース化を目的とした L_1 ノルム正則化を行列補完 NMF に適用することで、補完性能の変化を調査する。人工欠損乱数行列と音声スペクトログラムを用いた補完実験を行い、正則化を用いない NMF による補完手法と比較した。実験により、人工欠損乱数行列の補完において、正則化項を加えることで従来法よりコスト値が大きく改善することが示された。また、従来法では性能が安定しない基底数の設定が実際のデータと乖離した条件において、いずれの正則化も安定した結果をもたらすことを確認した。音声スペクトログラムによる実験では、僅かな歪みの改善が見られ、正則化を用いない NMF による補完手法に対し大きな改善を得た。

目次

| | | |
|--------------|---|-----------|
| 第 1 章 | 緒言 | 1 |
| 1.1 | 本論文の背景 | 1 |
| 1.2 | 本論文の目的 | 3 |
| 1.3 | 本論文の構成 | 4 |
| 第 2 章 | NMF を用いた行列補完 | 5 |
| 2.1 | はじめに | 5 |
| 2.2 | NMF | 5 |
| 2.2.1 | NMF の概要 | 5 |
| 2.2.2 | NMF の定式化 | 6 |
| 2.3 | NMF の行列補完への応用 | 10 |
| 2.3.1 | バイナリマスク行列を用いた NMF に基づく行列補完 | 10 |
| 2.3.2 | バイナリマスク行列を用いた NMF に基づく行列補完の課題 | 11 |
| 2.4 | 本章のまとめ | 12 |
| 第 3 章 | NMF の正則化項付与 | 13 |
| 3.1 | はじめに | 13 |
| 3.2 | CSRSNMF | 13 |
| 3.3 | L_1 -RNMF | 15 |
| 3.4 | 本章のまとめ | 16 |
| 第 4 章 | 正則化付き NMF に基づく行列補完 | 17 |
| 4.1 | はじめに | 17 |
| 4.2 | 動機 | 17 |
| 4.3 | CSRNMF に基づく行列補完 | 18 |
| 4.3.1 | コサイン類似度正則化の NMF への適用 | 18 |
| 4.3.2 | 補助関数法に基づく上限関数の導出 | 18 |
| 4.3.3 | CSRNMF の更新式の導出 | 19 |
| 4.3.4 | CSRNMF の行列補完への適用 | 21 |
| 4.4 | L_1 -RNMF に基づく行列補完 | 22 |
| 4.4.1 | L_1 -RNMF に基づく行列補完の目的関数 | 22 |

| | | |
|--------------|---------------------------------------|-----------|
| 4.4.2 | 補助関数法に基づく上限関数の導出 | 23 |
| 4.4.3 | L_1 -RNMF に基づく行列補完の更新式の導出 | 23 |
| 4.5 | 正則化項による NMF の挙動変化 | 25 |
| 4.6 | 本章のまとめ | 28 |
| 第 5 章 | 人工欠損行列の補完実験 | 29 |
| 5.1 | はじめに | 29 |
| 5.2 | 実験方法 | 29 |
| 5.3 | 実験結果 | 30 |
| 5.4 | 本章のまとめ | 36 |
| 第 6 章 | 欠損音声スペクトログラムの補完実験 | 38 |
| 6.1 | はじめに | 38 |
| 6.2 | 実験背景 | 38 |
| 6.2.1 | STFT | 38 |
| 6.2.2 | 時間周波数マスク | 40 |
| 6.3 | 実験方法 | 41 |
| 6.4 | 実験結果 | 46 |
| 6.5 | 本章のまとめ | 49 |
| 第 7 章 | 結言 | 50 |
| | 謝辞 | 51 |
| | 参考文献 | 51 |
| 付録 A | 補助関数法で利用される不等式 | 54 |
| A.1 | Jensen の不等式 | 54 |
| 付録 B | 人工欠損行列の補完実験結果 | 55 |
| 付録 C | 欠損音声スペクトログラムの補完実験結果 | 65 |

第1章

緒言

1.1 本論文の背景

今日のデータ解析では情報資源の肥大化によりビッグデータを扱う機会が増加している。取り扱われるデータは音声データ、画像データ、顧客の購買データなど、その種類は多様であり、それぞれの収集背景を持つ。このようなデータには往々にして欠損値が発生する。

欠損値とは、データの一部が破損して本来の値がわからない状態を言い、データの収集背景などに左右されて様々な要因により生じる。例として、Fig. 1.1 に音声信号、画像データ及び顧客の購買データにおける欠損の生成過程を示す。音声信号の場合、入力される音声は録音機器のダイナミックレンジを超えると、それらの音は最大値、または最小値で表されることとなるため、本来の大きさが表現できず欠損となる。画像データの場合、撮影段階において被写体が明るすぎる、または暗すぎると白飛びや黒潰れが発生し、これらが発生した部分の本来の画素値を取得できず欠損となる。顧客の購買データの場合、会員情報収集時において記入が不十分であると、一部の情報が不明になることで欠損が発生する。その他のデータにおいても、使用する機器の特性や不備、人為的なミスなど、あらゆる場面で欠損が発生することがある。これらの欠損は、その後続く処理の中で望ましくない影響を及ぼすことや、そもそも欠損を含むデータを活用することができなくなる等の問題を引き起こす。

以上のように、データにおける欠損値の発生は少なくなく、データサイズが大きくなるほどにその数は増加する。一方で、データ分析に用いる手法には欠損値を含むデータを扱えないものも多くあり、分析前に何らかの方法によって欠損値を削除する必要がある。欠損値の削除方法としては、欠損値を含む要素を切り捨てる方法と欠損部分に相当する値を予測し挿入する方法の2通りがあり、後者は欠損代入 (missing value imputation)、欠損補間 (missing value interpolation)、欠損補完 (missing value completion) 等と呼ばれる。欠損データの切り捨てるを行う場合、当然データ量が削減されるため、欠損の割合によっては解析結果やデータ自体が使い物にならない場合が生じる。このように、欠損値の対処において元データの性質を損なわないことは最も重要であり、補完が検討される場合も多い。

欠損データの補完を検討する場合、その手法には様々なものが提案されている。中でも、最も手軽な方法は単一代入法に分類される手法である。単一代入法は複数の欠損値部分に共通し

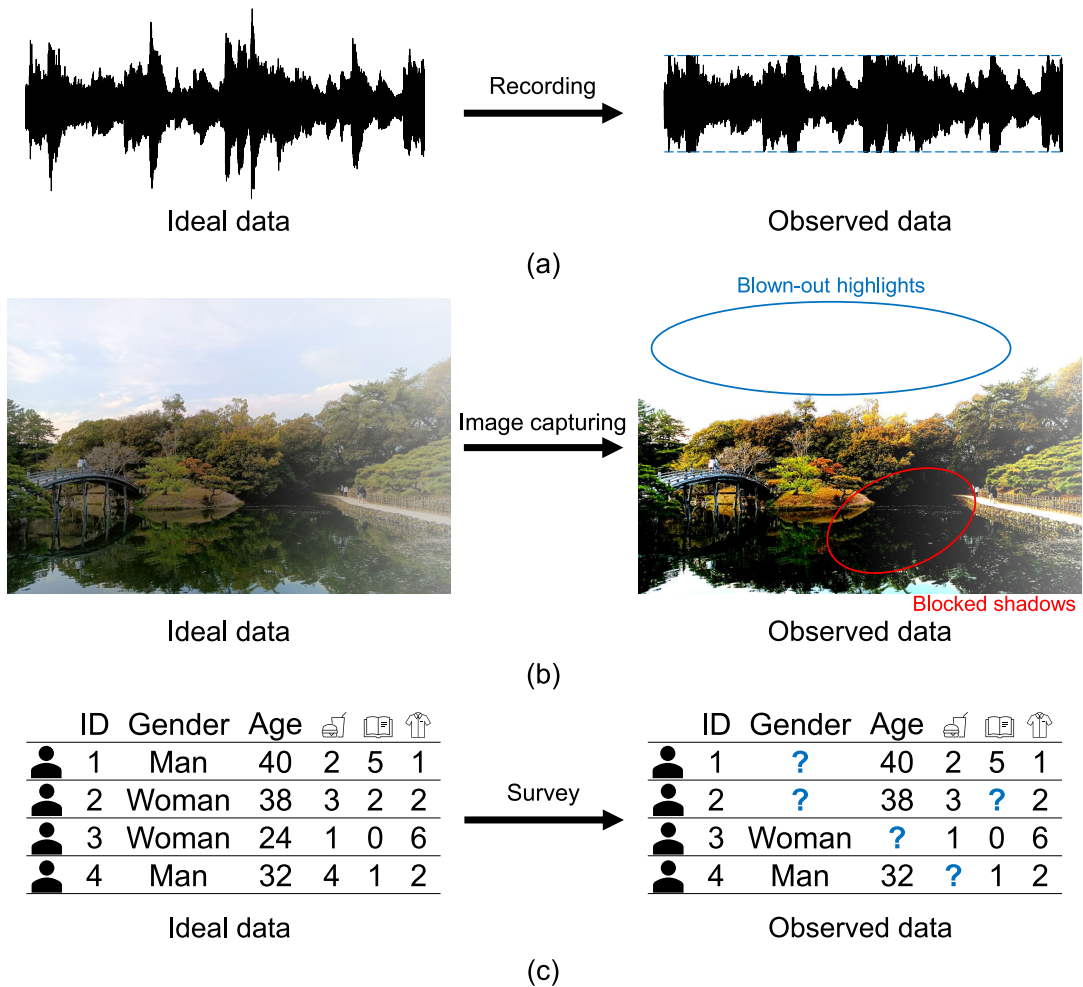


Fig. 1.1. Examples of missing values: (a) audio recording, (b) image capturing, and (c) surveying customer data.

て一つの値を代入する方法であり、データ全体の平均値を挿入する平均値補完や中央値を挿入する中央値補完などの種類がある。このような手法は手軽に適用できる一方、局所的な変化を持つ変数の一部が欠損した場合や特定の要因に従って欠損が発生する場合においては、一定値の挿入によってデータの特性を変質させる恐れがある。

そこで、より本来のデータに近い値で補完するために、データの本質的構造に基づく補完手法も多く提案されている [1]。近年、データの本質的構造に即した補完には、深層学習を代表とする学習データを用いた教師あり手法が多く提案されている [2]。これらの手法は、補完対象のデータを適切に学習することにより非常に高精度な補完を行うことができる。一方で、一定の精度を得るには莫大なコストがかかるため、小規模なデータ解析やサンプルデータを多く用意できない場合においては適用が困難となる。このため、コストを抑えつつデータの本質的構造を捉えた補完が可能な手法が求められている。

1.2 本論文の目的

本論文では、非負値のみを含む2次元データの欠損補完問題を取り扱う。この問題は欠損を含む非負行列の行列補完問題として解釈できる。非負値のみを取り扱うことは、本論文が適用できる対象を限定することになるが、この制約が実際のデータにおいて大きな問題となることは比較的少ない。その理由として、実際に観測される様々なデータは非負値であることが多い点が挙げられる。

顧客データを例にあげると、5人の顧客が3種類の商品の内、何をいくつ購入したかというデータは Fig. 1.2(a) のように 3×5 の非負値行列で表される。この時、 -1 回購入したという負のデータは通常あり得ない。また、8ビット白黒画像は Fig. 1.2(b) のように0から255の輝度を持つ非負値行列として表される。音響信号も例外ではなく、後述する短時間フーリエ変換 (short-time Fourier transform: STFT) を用いて1次元データの時間信号を2次元データの時間周波数信号に変換することで複素行列が得られ、その各要素の振幅値を取った振幅行列を定義すれば、Fig. 1.2(c) のように非負値行列が現れる。この行列は多くの音響信号処理で加工の対象となる特徴量である。このように、実世界において非負値行列で表されるものは多く、このような非負値行列の構造抽出において有効なアルゴリズムとして、非負値行列因子分解 (nonnegative matrix factorization: NMF) [3, 4] が知られている。

NMF は対象とするデータに非負制約を課し、低ランク性を仮定する行列分解アルゴリズムであり、補完手法にも応用される。NMF に基づく行列補完は、学習不要で高精度な補完手法であり、データの本質的構造を捉えながらコストを抑えることができる。上述の通り、NMF の非負制約は実世界において比較的問題になることはなく、また、NMF の仮定する低ランク性についても同様である。変数間に相関を持ったデータである場合、データサイズが巨大であればあるほど、相対的に少数の構造やパターンによって構成されることが多い。即ち、低ランク性を持つデータとなる。例えば、1万人の購買データがあるとき、その中から「食料品をよく買うグループ」や「娯楽品をよく買うグループ」といった特徴的なグループに分類することができる。このように、1万人の顧客データをカテゴリライズしていくことで、実際には100種類程度の傾向に分類でき、この傾向の組み合わせによって顧客のほとんどの特性を説明できる場合、行列においては低ランクであると言える。このようなパターンや傾向を見つけることができないデータとは、即ち、完全にランダムな要素を持つデータであり、実問題で扱うことは少ない。

以上より、実世界には非負かつ低ランクなデータは多く、NMF は幅広いデータに対して適用することができる。また、これらの制約により、NMF は学習データが不要でありながら高精度な補完を可能としている。一方で、高精度な補完を行うには、低ランク性が保証され、かつそのランクが判明しているという理想的な状態である必要がある。実際には、観測されるデータはすでに欠損していることや観測中に付加されるノイズの影響によって、このような条件を満たすことは困難である。そして、本質的に低ランクなデータがその低ランク性を損なう

4 第1章 緒言



Fig. 1.2. Nonnegative matrices appearing in real situations: (a) customer purchase data, (b) monochrome image data, (c) time-frequency amplitude of audio signal.

場合や元データのランクが不明な場合には補完精度が低下してしまう。そこで、本論文では NMF に基づく行列補完について改善を加えることで、より現実的な条件においても高精度な補完品質を安定させることを目的とする。

1.3 本論文の構成

2 章では、本論文で主に扱う NMF の定式化と行列補完への適用手法について説明し、従来手法の持つ問題点について述べる。3 章では、NMF に適用することで補完性能の向上が期待できる正則化項について説明する。ここでは、従来手法とともにコサイン類似度正則化項と $L1$ ノルム正則化項について説明し、その効果について述べる。4 章では、コサイン類似度正則化付き NMF 及び $L1$ ノルム正則化付き NMF の定式化を行い、その特性について確認する。5 章では、人工欠損行列の補完実験によって、コサイン類似度正則化付き NMF 及び $L1$ ノルム正則化付き NMF の補完精度を従来手法と比較し、その性能を評価する。6 章では、音響信号の補完実験を行い、より実用的な条件下における補完性能を評価する。7 章で本論文をまとめる。

第 2 章

NMF を用いた行列補完

2.1 はじめに

本章では、2.2 節で NMF の基礎理論について説明し、定式化を行う。2.3 節では NMF を用いた行列補完の手法について説明し、またその課題についても述べる。2.4 節で、本章についてまとめる。

2.2 NMF

2.2.1 NMF の概要

NMF とは非負の観測行列を 2 つの非負行列の行列積で近似的に分解する数値アルゴリズムである。このとき、2 つの非負行列の行列積による近似行列を低ランクに制約することで、観測行列の低ランク性を仮定した分解が可能になる。NMF は、この非負制約及び低ランク近似の両性質により、観測行列に頻出する少数の潜在的なパターンを抽出することが可能な教師なし学習である。実世界では、音の振幅時間周波数行列（振幅スペクトログラム）や画像等の非負行列で表されるデータが多く存在する。そのような非負値のデータを、非負値の構成要素に分解するために NMF を適用する。得られる非負行列は解釈が容易であるため、商品の購買データパターン抽出 [5] や音源分離 [6] 等のパターン認識を活用する分野で広く用いられている。

NMF による行列の近似分解は次の式で表される。

$$\mathbf{X} \approx \mathbf{W}\mathbf{H} \quad (2.1)$$

ここで、 $\mathbf{X} \in \mathbb{R}_{\geq 0}^{I \times J}$ は全要素が非負の観測行列であり、 $\mathbf{W} = [\mathbf{w}_1 \ \mathbf{w}_2 \ \cdots \ \mathbf{w}_K] \in \mathbb{R}_{\geq 0}^{I \times K}$ 及び $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \cdots \ \mathbf{h}_K]^T \in \mathbb{R}_{\geq 0}^{K \times J}$ は NMF で推定すべき非負変数行列で、 \mathbf{W} は基底行列、 \mathbf{H} は係数行列と呼ばれる。 I 及び J はそれぞれ観測行列の行数と列数である。また、 \mathbf{W} の列ベクトルである \mathbf{w}_k は基底ベクトルと呼ばれ、 K は基底ベクトルの個数を示す。 $k = 1, 2, \dots, K$ は基底ベクトルのインデックスを示す。NMF は基底ベクトルの数 K を $K \ll \min(I, J)$ と

$$\begin{array}{ccccccc}
 \text{Nonnegative} & & \text{Basis} & & \text{Coefficient} & & \text{Estimated} \\
 \text{observed matrix} & & \text{matrix} & & \text{matrix} & & \text{matrix} \\
 \mathbf{X}^{(I \times J)} & \approx & \mathbf{W}^{(I \times K)} & \mathbf{H}^{(K \times J)} & = & \widehat{\mathbf{X}}^{(I \times J)} \\
 \begin{array}{c} I \\ \left(\begin{array}{ccccc} 1 & 2 & 0 & 0 & 1 \\ 1 & 3 & 1 & 2 & 2 \\ 0 & 0 & 3 & 5 & 3 \end{array} \right) \\ J \end{array} & \approx & \begin{array}{c} \text{Basis} \\ I \\ \left(\begin{array}{cc} 1 & 0 \\ 1 & 1 \\ 0 & 3 \end{array} \right) \\ K \end{array} & K & \begin{array}{c} \left(\begin{array}{ccccc} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 \end{array} \right) \\ J \text{ Coefficient} \end{array} & = & \begin{array}{c} \left(\begin{array}{ccccc} 1 & 2 & 0 & 0 & 1 \\ 1 & 2 & 1 & 2 & 2 \\ 0 & 0 & 3 & 6 & 3 \end{array} \right) \\ \end{array}
 \end{array}$$

Fig. 2.1. Decompose observation matrix using NMF.

るように設定することで、観測行列 \mathbf{X} を式 (2.1) のように低ランク近似する。Fig. 2.1 は、NMF による行列の分解近似の例である。このように、観測行列を非負値の基底行列と係数行列で表現することで、ランクを基底ベクトルの数 K に抑えた近似行列 $\widehat{\mathbf{X}}$ を式 (2.1) のように得ることができる。

NMF を適用するにあたって注意すべき点として、NMF で得られる \mathbf{W} と \mathbf{H} には原理的に次の2つの任意性が存在する。

- 順序の任意性 (permutation ambiguity)

\mathbf{W} に含まれる K 本の基底ベクトル $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_K$ の順序を入れ替えても、 \mathbf{H} に含まれる係数ベクトル $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K$ を同様の順序に入れ替えれば、 \mathbf{WH} は一致する。

$$\text{例: } [\mathbf{w}_1 \quad \mathbf{w}_2] \begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{h}_2^T \end{bmatrix} = [\mathbf{w}_2 \quad \mathbf{w}_1] \begin{bmatrix} \mathbf{h}_2^T \\ \mathbf{h}_1^T \end{bmatrix}$$

- 大きさの任意性 (scale ambiguity)

\mathbf{W} に含まれる K 本の基底ベクトル $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_K$ のそれぞれに対して C_1, C_2, \dots, C_K というスカラーを乗じても、 \mathbf{H} に含まれる係数ベクトル $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K$ のそれぞれに対して $C_1^{-1}, C_2^{-1}, \dots, C_K^{-1}$ を乗じれば、 \mathbf{WH} は一致する。

$$\text{例: } [\mathbf{w}_1 \quad \mathbf{w}_2] \begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{h}_2^T \end{bmatrix} = [C_1 \mathbf{w}_1 \quad C_2 \mathbf{w}_2] \begin{bmatrix} C_1^{-1} \mathbf{h}_1^T \\ C_2^{-1} \mathbf{h}_2^T \end{bmatrix}$$

また、 \mathbf{W} 及び \mathbf{H} の推定問題は解析的に解けるわけではなく、次項で説明する最適化問題を解く必要がある。さらに、局所解の存在により常に同じ結果を得られるとは限らず。与える初期値や用いる最適化アルゴリズムによって異なる分解となる。

2.2.2 NMF の定式化

NMF では、次の最適化問題を解くことで変数行列 \mathbf{W} 及び \mathbf{H} を推定する。

$$\underset{\mathbf{W}, \mathbf{H}}{\text{Minimize}} \mathcal{D}(\mathbf{X} | \mathbf{WH}) \text{ s.t. } w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k \tag{2.2}$$

ここで, x_{ij} , w_{ik} , 及び h_{kj} はそれぞれ \mathbf{X} , \mathbf{W} , 及び \mathbf{H} の要素であり, $i = 1, 2, \dots, I$ 及び $j = 1, 2, \dots, J$ はそれぞれ観測行列の行インデックス及び列インデックスを示す. また, $\mathcal{D}(\cdot|\cdot)$ は2つの入力行列間の乖離度を測る関数である. 本論文では, 次式で表される一般化 Kullback–Leibler (KL) ダイバージェンス [7] $\mathcal{D}_{\text{KL}}(\cdot|\cdot)$ を用いる.

$$\mathcal{D}_{\text{KL}}(\mathbf{A}|\mathbf{B}) = \sum_i \sum_j \left[a_{ij} \log \frac{a_{ij}}{b_{ij}} - (a_{ij} - b_{ij}) \right] \quad (2.3)$$

ここで, a_{ij} 及び b_{ij} はそれぞれ $\mathbf{A} \in \mathbb{R}^{I \times J}$ 及び $\mathbf{B} \in \mathbb{R}^{I \times J}$ の (i, j) 要素を表す.

一般化 KL ダイバージェンスに基づく NMF (KLNMF) の反復最適化更新式を求める. また, 本論文では NMF の反復更新式の導出に補助関数法 (majorization-minimization algorithm) [8] と呼ばれる最適化アルゴリズムを用いる. まず, KLNMF の最適化問題は式 (2.3) より, 次式のようになる.

$$\underset{\mathbf{W}, \mathbf{H}}{\text{Minimize}} \mathcal{J}(\mathbf{W}, \mathbf{H}) \text{ s.t. } w_{ik}, h_{kj} \geq 0 \forall i, j, k \quad (2.4)$$

$$\mathcal{J}(\mathbf{W}, \mathbf{H}) = \sum_i \sum_j \left[x_{ij} \log \frac{x_{ij}}{\sum_k w_{ik} h_{kj}} - \left(x_{ij} - \sum_k w_{ik} h_{kj} \right) \right] \quad (2.5)$$

この目的関数を変形し次式を得る.

$$\mathcal{J}(\mathbf{W}, \mathbf{H}) = \sum_i \sum_j \left(x_{ij} \log x_{ij} - x_{ij} \log \sum_k w_{ik} h_{kj} - x_{ij} + \sum_k w_{ik} h_{kj} \right) \quad (2.6)$$

式 (2.6) の目的関数の w_{ik} 及び h_{kj} に関する偏微分を計算する場合, 第2項の k に関する総和を含む対数関数の偏微分が原因で \mathbf{W} や \mathbf{H} の停留点を求めることができない. そこで, 負対数関数は凸関数であることから, 付録 (A.1) に示す Jensen の不等式を適用することで, 式 (2.6) を上から抑える補助関数を設計する. 式 (2.6) の第2項に対して, 次の不等式を考える.

$$\begin{aligned} -\log \sum_k w_{ik} h_{kj} &= -\log \sum_k \delta_{ijk} \frac{w_{ik} h_{kj}}{\delta_{ijk}} \\ &\leq -\sum_k \delta_{ijk} \log \frac{w_{ik} h_{kj}}{\delta_{ijk}} \end{aligned} \quad (2.7)$$

ここで, $\delta_{ijk} \geq 0$ は補助変数であり $\sum_k \delta_{ijk} = 1$ を満たす. 式 (2.7) は, Jensen の不等式において, 凸関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(\cdot) = -\log(\cdot)$ として適用した不等式である. この不等式を式 (2.6) の第2項に適用することで, 補助関数を次式のように定義できる.

$$\begin{aligned} \mathcal{J}(\mathbf{W}, \mathbf{H}) &\leq \mathcal{J}^+(\mathbf{W}, \mathbf{H}, \mathbf{\Delta}) \\ &= \sum_i \sum_j \left(x_{ij} \log x_{ij} - x_{ij} \sum_k \delta_{ijk} \log \frac{w_{ik} h_{kj}}{\delta_{ijk}} - x_{ij} + \sum_k w_{ik} h_{kj} \right) \end{aligned} \quad (2.8)$$

ここで, $\mathbf{\Delta} \in \mathbb{R}_{\geq 0}^{I \times J \times K}$ は補助変数 δ_{ijk} で構成される3階のテンソルである.

8 第2章 NMF を用いた行列補完

次に、補助関数 $\mathcal{J}^+(\mathbf{W}, \mathbf{H}, \mathbf{\Delta})$ に対する補助変数 $\mathbf{\Delta}$ の最小化について考える。いま、補助変数には等式制約条件 $\sum_k \delta_{ijk} = 1$ が課せられているため、次の等式制約条件付き最適化問題を考える。

$$\underset{\mathbf{\Delta}}{\text{Minimize}} \mathcal{J}^+(\mathbf{W}, \mathbf{H}, \mathbf{\Delta}) \quad \text{s.t.} \quad \sum_k \delta_{ijk} = 1 \quad \forall i, j \quad (2.9)$$

最適化問題 (2.9) の Lagrange 関数は次式で与えられる。

$$\mathcal{L}(\mathbf{W}, \mathbf{H}, \mathbf{\Delta}, \lambda) = \mathcal{J}^+ - \lambda \left(\sum_k \delta_{ijk} - 1 \right) \quad (2.10)$$

従って、 δ_{ijk} 及び λ の偏微分をそれぞれ 0 とおくと次式が得られる。

$$\frac{\partial \mathcal{L}}{\partial \delta_{ijk}} = -x_{ij} \left(\log \frac{w_{ik} h_{kj}}{\delta_{ijk}} - 1 \right) - \lambda = 0 \quad (2.11)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = - \sum_k \delta_{ijk} + 1 = 0 \quad (2.12)$$

式 (2.12) は等式制約条件 $\sum_{k=1}^K \delta_{ijk} = 1$ そのものである。また、式 (2.11), $\delta_{ijk} > 0$, 及び $w_{ik}, h_{kj} \geq 0$ より次式となる。

$$\delta_{ijk} = w_{ik} h_{kj} \exp \left(\frac{\lambda}{x_{ij}} - 1 \right) \quad (2.13)$$

式 (2.13) の両辺に対して k について総和をとると次式を得る。

$$\sum_k \delta_{ijk} = \sum_k w_{ik} h_{kj} \exp \left(\frac{\lambda}{x_{ij}} - 1 \right) \quad (2.14)$$

よって、等式制約条件 $\sum_k \delta_{ijk} = 1$ より次式を得る、

$$\sum_k w_{ik} h_{kj} \exp \left(\frac{\lambda}{x_{ij}} - 1 \right) = 1 \quad (2.15)$$

未定乗数 λ は k に依らない定数であるため、結局、次式となる。

$$\exp \left(1 - \frac{\lambda}{x_{ij}} \right) = \sum_k w_{ik} h_{kj} \quad (2.16)$$

これを式 (2.13) に代入することで、補助変数の最小解が次式を得る。

$$\delta_{ijk} = \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}} \quad \forall i, j, k \quad (2.17)$$

従って、式 (2.17) が補助関数 (2.9) の補助変数に関する最小解を与える。今、変数の更新の反復回数を t とし、各変数の上付き文字として表示すると、次のようになる。

$$\delta_{ijk}^{(t+1)} = \frac{w_{ik}^{(t)} h_{kj}^{(t)}}{\sum_{k'} w_{ik'}^{(t)} h_{k'j}^{(t)}} \quad \forall i, j, k \quad (2.18)$$

続いて、設計した補助関数 $\mathcal{J}^+(\mathbf{W}, \mathbf{H}, \Delta)$ を本来の変数 \mathbf{W} 及び \mathbf{H} について最小化する。初めに、補助関数 (2.8) を w_{ik} で偏微分して 0 とおき、次式を得る。

$$\frac{\partial \mathcal{J}^+}{\partial w_{ik}} = \sum_j \left(-x_{ij} \delta_{ijk} \frac{\delta_{ijk}}{w_{ik} h_{kj}} \frac{h_{kj}}{\delta_{ijk}} + h_{kj} \right) = 0 \quad (2.19)$$

式 (2.19) を整理すると次式のようになる。

$$\begin{aligned} -\sum_j x_{ij} \frac{\delta_{ijk}}{w_{ik}} + \sum_j h_{kj} &= 0 \\ w_{ik} \sum_j h_{kj} &= \sum_j x_{ij} \delta_{ijk} \\ w_{ik} &= \frac{\sum_j x_{ij} \delta_{ijk}}{\sum_j h_{kj}} \end{aligned} \quad (2.20)$$

式 (2.20) が補助関数 (2.8) を最小化する w_{ik} となる。反復回数 t を記述し、次式を得る。

$$w_{ik}^{(t+1)} = \frac{\sum_j x_{ij} \delta_{ijk}^{(t+1)}}{\sum_j h_{kj}^{(t)}} \quad (2.21)$$

同様に、補助関数 (2.8) を h_{kj} で偏微分して 0 とおくと、 h_{kj} の最小解も次のように得られる。

$$\begin{aligned} \frac{\partial \mathcal{J}^+}{\partial h_{kj}} &= \sum_i \left(-x_{ij} \delta_{ijk} \frac{\delta_{ijk}}{w_{ik} h_{kj}} \frac{w_{ik}}{\delta_{ijk}} + w_{ik} \right) = 0 \\ -\sum_i x_{ij} \frac{\delta_{ijk}}{h_{kj}} + \sum_i w_{ik} &= 0 \\ h_{kj} \sum_i w_{ik} &= \sum_{i=1} x_{ij} \delta_{ijk} \\ h_{kj} &= \frac{\sum_i x_{ij} \delta_{ijk}}{\sum_i w_{ik}} \end{aligned} \quad (2.22)$$

反復回数 t を考慮して記述すると、次のようになる。

$$h_{kj}^{(t+1)} = \frac{\sum_i x_{ij} \delta_{ijk}^{(t+1)}}{\sum_i w_{ik}^{(t+1)}} \quad (2.23)$$

従って、式 (2.7) で補助変数を更新し、式 (2.21) 及び式 (2.23) で本来の変数を更新する。この 2 つの手順を繰り返すことで、目的関数 (2.6) の値を最小化する \mathbf{W} 及び \mathbf{H} を求めることができる。

なお、補助変数の反復更新式 (2.18) を式 (2.21) 及び式 (2.23) にそれぞれ代入することで、

次式のような統合された反復更新式に書き換えることができる。

$$\begin{aligned} w_{ik}^{(t+1)} &= \frac{\sum_j x_{ij} \frac{w_{ik}^{(t)} h_{kj}^{(t)}}{\sum_{k'} w_{ik'}^{(t)} h_{k'j}^{(t)}}}{\sum_j h_{kj}^{(t)}} \\ &= w_{ik}^{(t)} \frac{\sum_j \frac{x_{ij}}{\sum_{k'} w_{ik'}^{(t)} h_{k'j}^{(t)}} h_{kj}^{(t)}}{\sum_j h_{kj}^{(t)}} \end{aligned} \quad (2.24)$$

$$\begin{aligned} h_{kj}^{(t+1)} &= \frac{\sum_i x_{ij} \frac{w_{ik}^{(t+1)} h_{kj}^{(t)}}{\sum_{k'} w_{ik'}^{(t+1)} h_{k'j}^{(t)}}}{\sum_i w_{ik}^{(t+1)}} \\ &= h_{kj}^{(t)} \frac{\sum_i \frac{x_{ij}}{\sum_{k'} w_{ik'}^{(t+1)} h_{k'j}^{(t)}} w_{ik}^{(t+1)}}{\sum_i w_{ik}^{(t+1)}} \end{aligned} \quad (2.25)$$

式 (2.24) 及び式 (2.25) は、次のように行列形式で表現することができる。

$$\mathbf{W} \leftarrow \mathbf{W} \odot \frac{\mathbf{X} \mathbf{H}^T}{\mathbf{1}^{(I \times J)} \mathbf{H}^T} \quad (2.26)$$

$$\mathbf{H} \leftarrow \mathbf{H} \odot \frac{\mathbf{W}^T \mathbf{X}}{\mathbf{W}^T \mathbf{1}^{(I \times J)}} \quad (2.27)$$

ここで、見やすさのために反復回数 t の表記は省略し、変数更新を表す演算子 \leftarrow を用いている。また、 $\mathbf{1}^{(I \times J)}$ 全ての要素が 1 の $I \times J$ の行列であり、演算子 \odot は要素毎の積 (Hadamard product)、行列の分数は要素毎の商を表す。以上より、KLNMF の反復更新式が導出された。

2.3 NMF の行列補完への応用

2.3.1 バイナリマスク行列を用いた NMF に基づく行列補完

NMF に基づく行列補完には、ナイーブな手法として欠損要素のインデックスを行列とした、バイナリマスク行列を用いる方法がある。この手法の概要は Fig. 2.2 の通りである。NMF に基づく行列補完では、初めに、観測行列の欠損要素を 0、その他を 1 とするバイナリマスク行列 $\mathbf{M} \in \{0, 1\}^{I \times J}$ を生成する。生成されたバイナリマスク行列を観測行列及び近似行列のそれぞれに要素積することで、欠損要素を無視した行列補完が可能となる。即ち、観測可能な要素についてのみ、観測行列と近似行列の目的関数値を最小化することで、観測不能な要素についても行列全体の代数構造に基づいて何らかの非負値を補完する手法である。

NMF に基づく行列補完は、次式の最適化問題で表される。

$$\underset{\mathbf{W}, \mathbf{H}}{\text{Minimize}} \mathcal{J}(\mathbf{W}, \mathbf{H}) \text{ s.t. } w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k \quad (2.28)$$

$$\mathcal{J}(\mathbf{W}, \mathbf{H}) = \sum_i \sum_j m_{ij} \left[x_{ij} \log \frac{x_{ij}}{\sum_k w_{ik} h_{kj}} - \left(x_{ij} - \sum_k w_{ik} h_{kj} \right) \right] \quad (2.29)$$

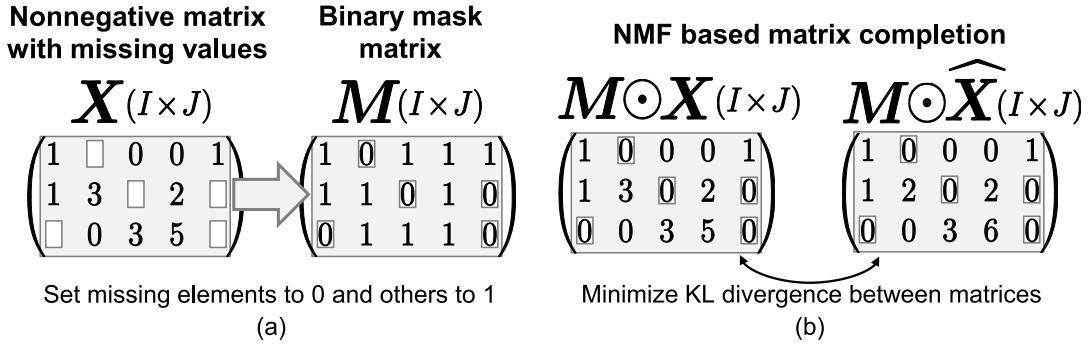


Fig. 2.2. NMF-based matrix completion using binary masks. Missing values are ignored in calculation of KL divergence during estimation of approximated matrix $\widehat{\mathbf{X}}$.

ここで、 m_{ij} はバイナリマスク行列 \mathbf{M} の要素である。バイナリマスク行列は定数行列であるため、2.2.2 項と同様の手順で次式の更新式を導出できる。

$$w_{ik}^{(t+1)} = w_{ik}^{(t)} \frac{\sum_j \frac{m_{ij} x_{ij}}{\sum_{k'} w_{ik'}^{(t)} h_{k'j}^{(t)}} h_{kj}^{(t)}}{\sum_j m_{ij} h_{kj}^{(t)}} \quad (2.30)$$

$$h_{kj}^{(t+1)} = h_{kj}^{(t)} \frac{\sum_i \frac{m_{ij} x_{ij}}{\sum_{k'} w_{ik'}^{(t+1)} h_{k'j}^{(t)}} w_{ik}^{(t+1)}}{\sum_i m_{ij} w_{ik}^{(t+1)}} \quad (2.31)$$

式 (2.30) 及び式 (2.31) は行列形式で表現すると次式となる。

$$\mathbf{W} \leftarrow \mathbf{W} \odot \frac{(\mathbf{M} \odot \mathbf{X}) \mathbf{H}^T}{\mathbf{W} \mathbf{H} \mathbf{H}^T} \quad (2.32)$$

$$\mathbf{H} \leftarrow \mathbf{H} \odot \frac{\mathbf{W}^T (\mathbf{M} \odot \mathbf{X})}{\mathbf{W}^T \mathbf{M}} \quad (2.33)$$

以上で、バイナリマスク行列を用いた NMF に基づく行列補完の更新式が導出された。

2.3.2 バイナリマスク行列を用いた NMF に基づく行列補完の課題

バイナリマスク行列を用いた NMF に基づく行列補完の適用例を Fig. 2.3 に示す。Fig. 2.3 は、ランク 5 の 25 行 25 列の正解行列の内、50% の要素を欠損させた観測行列に対し、バイナリマスク行列を用いた NMF に基づく行列補完を適用したときの近似行列である。このように、NMF に基づく行列補完は補完精度が高く、Fig. 2.3 においても、目視では正解行列と近似行列の差異を確認できない精度である。しかしながら、これは理想的な状態における欠損補完の例である。即ち、欠損前の観測行列のランクが 5 であることを既知とし、NMF の基底ベクトル数を $K = 5$ に設定した結果を示している。さらに、観測行列にノイズも付与されておらず、完全な最適化が達成された場合には式 (2.29) の目的関数の値を 0 にすることができるケースである。

NMF において、欠損前の観測行列（正解行列）のランクの推測精度（即ち K の設定値）が推定結果である $\widehat{\mathbf{X}}$ の近似精度に与える影響は大きい。NMF が低ランク近似するうえで、基

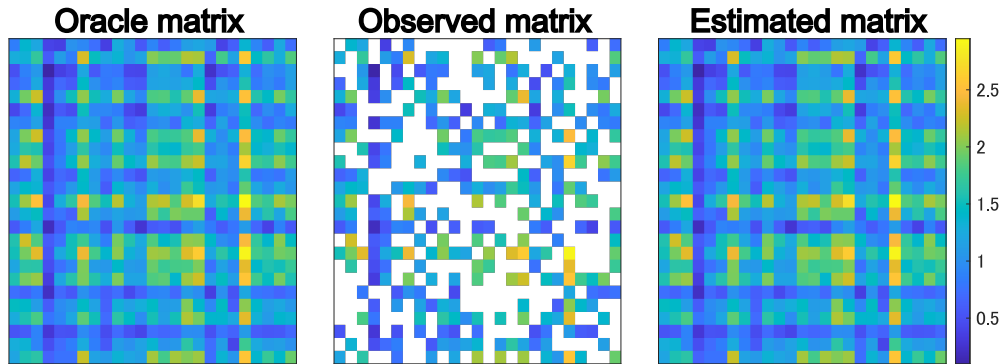


Fig. 2.3. Example of performing NMF-based matrix completion using binary mask. This method can estimate highly accurate completion results when number of bases is set to ideal value.

底行列の各基底ベクトルの役割は大きく、その本数となるランクは非常に重要な設定値となる。このため、NMF に用いるランク K が実際の正解行列のランク（欠損前の観測行列のランク）と 1 異なるだけでも、行列補完の精度が大きく低下し、Fig. 2.3 のような結果は得られなくなる。また、Fig 2.3 の例では、正解行列と観測行列の間に欠損は発生しているが、観測行列にノイズは与えられていない。実際にはいかなる観測系においてもノイズは発生してしまうため、多くの場合で観測行列は正解行列のランクによらずフルランクとなる。よって、NMF の仮定する低ランク性は成立せず、結果として NMF の近似精度に影響を与える。

バイナリマスク行列を用いた NMF に基づく行列補完は、低ランクかつ正解行列の基底数が既知の時、高精度な補完を可能とする。一方で、実際の観測行列は観測時のノイズによってフルランクになり、またノイズと欠損が生じる前の行列のランクの推定も困難であるため、その性能を十分に発揮することが難しい。そのため、バイナリマスク行列を用いた NMF に基づく行列補完は、観測行列に含まれるノイズやランクの推定誤差（NMF で用意する基底ベクトルの個数に関するモデル誤差）に対して頑健になるような工夫が求められる。

2.4 本章のまとめ

本章では、NMF の基礎理論とバイナリマスク行列を用いた NMF に基づく行列補完の手法について説明した。この手法は、本論文における従来手法であり、この手法で達成することのできる行列補完の性能がベースラインとなる。また、バイナリマスク行列を用いた NMF に基づく行列補完が抱える問題について述べた。次章では、観測行列に含まれるノイズや用意する基底ベクトルの個数への頑健性を持つことが期待できる正則化について説明する。

第 3 章

NMF の正則化項付与

3.1 はじめに

本章では、NMF に対してこれまで提案されている正則化項について説明する。3.2 節では、音源分離の分野で用いられる半教師あり NMF (semi-supervised NMF : SNMF) について示し、SNMF の改良型であるコサイン類似度正則化条件付き SNMF (cosine-similarity-regularized SNMF : CSRSNMF) に用いられているコサイン類似度正則化項について述べる。3.3 節では、 L_1 ノルム正則化付き NMF (L_1 -norm-regularized NMF : L_1 -RNMF) について説明し、 L_1 ノルム正則化項がもたらす効果について述べる。3.4 節では本章についてまとめる。

3.2 CSRSNMF

NMF には、音源分離の分野等で用いられる SNMF [9, 10] という手法がある。これは、混合信号の振幅スペクトログラムを、目的音源及び非目的音源の振幅スペクトログラムの和に近似的に分解する手法である。目的音源と非目的音源の振幅スペクトログラムはいずれも基底行列と係数行列の行列積で表される。より具体的に、SNMF では、混合信号の振幅スペクトログラム $\mathbf{Y} \in \mathbb{R}_{\geq 0}^{I \times J}$ を次の最適化により分離する。

$$\underset{\mathbf{G}, \mathbf{H}, \mathbf{U}}{\text{Minimize}} \mathcal{D}(\mathbf{Y} | \mathbf{F}\mathbf{G} + \mathbf{H}\mathbf{U}) \text{ s.t. } g_{kj}, h_{il}, u_{lj} \geq 0 \forall i, j, k, l \quad (3.1)$$

ここで、非目的音源の基底行列 \mathbf{H} のランクは L であり、 $l = 1, 2, \dots, L$ は非目的音源の基底行列 \mathbf{H} に含まれる基底ベクトルのインデックスを示す。 $\mathbf{F} \in \mathbb{R}_{\geq 0}^{I \times J}$ は事前に分離対象である目的音源の学習データから特徴量を学習した教師基底行列であり、 $\mathbf{G} \in \mathbb{R}_{\geq 0}^{K \times J}$ は目的音源の係数行列である。また、 $\mathbf{H} \in \mathbb{R}_{\geq 0}^{I \times L}$ 及び $\mathbf{U} \in \mathbb{R}_{\geq 0}^{L \times J}$ はそれぞれ非目的音源を表すための基底行列と係数行列であり、 g_{kl} , h_{il} 及び u_{lj} はそれぞれ \mathbf{G} , \mathbf{H} 及び \mathbf{U} の非負要素である。

SNMF は混合信号中に目的音源と非目的音源が類似する要素を含むときに分離精度が劣化する問題がある [10]。これは、類似する要素が教師基底行列 \mathbf{F} または非目的音源の基底行列 \mathbf{H} のいずれを用いても表現できてしまうためである。即ち、類似成分を \mathbf{F} と \mathbf{H} が取り合う

り、NMFにおけるコサイン類似度の最小値は0となるため、式(3.5)における正則化項は変数 \mathbf{H} の推定において、できるだけ \mathbf{F} と直交することを誘導する効果がある。

一般に、最適化問題における目的関数をデータ近似項 (data fidelity term) と呼ぶ。また、変数の推定において何らかの性質や構造を誘導するために追加する項を正則化項 (regularization term) と呼ぶ。式(3.5)は、第一項がデータ近似項であり、第二項が正則化項に該当する。正則化項の強さは重み係数 μ で制御され、この値が大きいほど最適化に正則化が強く作用することになる。

3.3 L_1 -RNMF

NMFの文脈では、係数行列の要素をスパース (sparse, 疎) に誘導することで、観測行列の特徴や潜在的な構造をより強く反映した基底行列を頑健に推定する手法が提案されている。この手法はスパース NMF と呼ばれ、独自のスパース性基準を定義して正則化に用いる手法 [12] や係数行列の L_1 ノルムを正則化に用いる手法 [13, 14] 等が提案されている。特に L_1 ノルムを用いた手法は、有名な least absolute shrinkage and selection operator (lasso) 回帰 [15] の正則化を NMF に導入したものであり、一般的な正則化といえる。

L_1 ノルムは $\|\cdot\|_1$ と表され、行列 \mathbf{X} において、次式のように定義される。

$$\|\mathbf{X}\|_1 = \sum_i \sum_j |x_{ij}| \quad (3.6)$$

これは即ち、行列の全要素の絶対値総和を表す。この値が小さくなることは、行列 \mathbf{X} 中の非零要素の個数が少ないことを意味する。但し、本来は L_0 擬ノルムが行列の非零要素の個数を表したものであり、 L_1 ノルムは L_0 擬ノルムの最良な凸緩和であるため、最適化の正則化として L_0 擬ノルムの代わりに用いられるものである。

NMF等のパラメータの更新によって目的関数を最小化するアルゴリズムにおいて、 L_1 ノルム正則化項はそのパラメータをスパースにする効果を持つことが知られている [16]。このような L_1 ノルム正則化項を NMF の係数行列に対して付与したものが L_1 -RNMF である。 L_1 -RNMF は次式によって定義される。

$$\underset{\mathbf{W}, \mathbf{H}}{\text{Minimize}} \mathcal{J}_{L_1\text{-RNMF}} \text{ s.t. } w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k \quad (3.7)$$

$$\mathcal{J}_{L_1\text{-RNMF}} = \mathcal{D}(\mathbf{X}|\mathbf{W}\mathbf{H}) + \nu\|\mathbf{H}\|_1 \quad (3.8)$$

ここで、 ν は L_1 ノルム正則化項の重み係数である。

L_1 -RNMF は NMF の持つ低ランク性に加え、係数行列にスパース性を持つことになる。この結果、観測行列を特徴的な値によって少ないパラメータで表そうとすることにより、より強力に特徴量の抽出が可能となる。

3.4 本章のまとめ

本章ではコサイン類似度正則化と $L1$ ノルム正則化の2つの正則化について説明した。いずれの正則化も NMF の特徴抽出をノイズの発生や基底ベクトルの個数のモデル誤差に対して頑健にする効果が期待できる手法である。次章では、本論文の提案手法として、この2つの正則化を付与した NMF に基づく行列補完の定式化を行う。

第4章

正則化付き NMF に基づく行列補完

4.1 はじめに

本章では、コサイン類似度正則化及び L_1 ノルム正則化の2つの正則化を、NMF に基づく行列補完へと適用する。4.2 節では、NMF に基づく行列補完に正則化を加える動機について述べる。4.3 節では、SNMF ではなく通常の NMF におけるコサイン類似度正則化付き NMF (cosine-similarity-regularized NMF: CSRNMF) の定式化と行列補完への適用について説明する。4.4 節では、 L_1 -RNMF に基づく行列補完の定式化について説明する。4.5 節では、正則化項が NMF にもたらす影響について確認し、4.6 節で本章についてまとめる。

4.2 動機

2.3 節に記載の通り、通常の NMF に基づく行列補完は、欠損やノイズが発生する前の観測行列である正解行列のランクが既知で、かつ低ランクである場合において高精度な補完を可能とする。一方、実問題においてこの条件を満たす場合は少ない。実際には正解行列が低ランクである場合においても、ノイズが発生することによりフルランク化することで低ランク性を失ってしまう。さらに、欠損が生じて観測行列となった後に、正解行列のランクを正確に推定することは難しい。そのため NMF の実用上は基底ベクトル数 K を適当な値や経験的な値で定めることが多く、当然正解行列のランクと適当に定めた K の間にはモデル誤差が生じてしまう。観測行列の低ランクな近似行列を得ようとする NMF において、基底行列の各基底ベクトルは非常に重要な役割を果たすにもかかわらず、正解行列のランクと基底ベクトル数 K の設定値が異なるならば、近似精度に多大な影響を与える問題がある。以上から、NMF に基づく行列補完の精度向上には、観測行列に含まれるノイズや正解行列のランクと基底ベクトル数のモデル誤差に対し、より頑健な性能を持つ必要があると考えられる。

本論文では、コサイン類似度及び L_1 ノルムの正則化を用いることで、観測行列に含まれるノイズの影響を受けにくくなるように改良された NMF に基づく行列補完を提案する。これらの正則化はいずれも、 K 個の基底ベクトルの個々の独立性を高める誘導を行うものであり、ノ

イズのような低ランク性に寄与しない要素を無視する効果が期待できる。そのため、観測信号中のノイズに対する頑健性が向上することが期待される。また、前述の効果は正解行列のランクと基底ベクトル数 K 間のモデル誤差に対する頑健性ももたらす可能性がある。

3.2 節において説明したコサイン類似度正則化は、CSRSNMF においては、目的音源と非目的音源の基底行列間を直交させるために活用される。本論文の提案手法である行列補完の用途に応用するためには、2つの基底行列間ではなく、1つの基底行列内に含まれる各基底ベクトルを相互に直交させることを考える。

4.3 CSRNMF に基づく行列補完

4.3.1 コサイン類似度正則化の NMF への適用

コサイン類似度による正則化項を KLNMF に適用する。目的関数は次式となる。

$$\text{Minimize } \mathcal{J}_{\text{CSRNMF}}(\mathbf{W}, \mathbf{H}) \text{ s.t. } w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k \quad (4.1)$$

$$\mathcal{J}_{\text{CSRNMF}}(\mathbf{W}, \mathbf{H}) = \mathcal{D}_{\text{KL}}(\mathbf{X} | \mathbf{W}\mathbf{H}) + \mu \sum_{k=1}^K \sum_{k'=k+1}^K \mathcal{R}_{\text{cos}}(\mathbf{w}_k, \mathbf{w}_{k'}) \quad (4.2)$$

$$\begin{aligned} &= \sum_i \sum_j \left(x_{ij} \log \frac{x_{ij}}{\sum_k w_{ik} h_{kj}} - x_{ij} + \sum_k w_{ik} h_{kj} \right) \\ &+ \mu \sum_{k=1}^K \sum_{k'=k+1}^K \frac{\sum_i w_{ik} w_{ik'}}{(\sum_i w_{ik}^2)^{\frac{1}{2}} (\sum_i w_{ik'}^2)^{\frac{1}{2}}} \end{aligned} \quad (4.3)$$

ここで、 $k' = k + 1, k + 2, \dots, K$ は基底インデクスであり、常に $k < k'$ の関係にある。

4.3.2 補助関数法に基づく上限関数の導出

前項で示した式 (4.3) から、 $\mathcal{J}_{\text{CSRNMF}}$ の上限関数を求める。煩雑になるのを避けるため、式 (4.2) の第一項と第二項を分けて考える。

式 (4.2) の第一項は、2.2.2 項と同様に、Jensen の不等式を利用して次のように変形できる。

$$\sum_i \sum_j \left(x_{ij} \log x_{ij} - x_{ij} \sum_k \delta_{ijk} \log \frac{w_{ik} h_{kj}}{\delta_{ijk}} - x_{ij} + \sum_k w_{ik} h_{kj} \right) \quad (4.4)$$

ここで、 $\delta_{ijk} > 0$ は補助変数であり、 $\sum_k \delta_{ijk} = 1$ を満たす。等式条件は次式となる。

$$\delta_{ijk} = \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}} \quad (4.5)$$

式 (4.2) の第二項は次のように変形できる.

$$\begin{aligned}
\mu \sum_{k=1}^K \sum_{k'=k+1}^K \mathcal{R}_{\cos}(\mathbf{w}_k, \mathbf{w}_{k'}) &= \mu \sum_{k=1}^K \sum_{k'=k+1}^K \frac{\sum_i w_{ik} w_{ik'}}{(\sum_i w_{ik}^2)^{\frac{1}{2}} (\sum_i w_{ik'}^2)^{\frac{1}{2}}} \\
&= \mu \sum_{k=1}^K \sum_{k'=k+1}^K \frac{\sum_i w_{ik} w_{ik'}}{\left(\sum_i \epsilon_{ik} \frac{w_{ik}^2}{\epsilon_{ik}}\right) (\sum_i w_{ik'}^2)^{\frac{1}{2}}} \\
&\leq \mu \sum_{k=1}^K \sum_{k'=k+1}^K \left[\left(\sum_{i'} w_{i'k'}^2\right)^{-\frac{1}{2}} \sum_{i'} \sum_{i''} w_{i'k'} \epsilon_{i''k}^{\frac{3}{2}} \frac{w_{i'k}}{w_{i''k}} \right] \quad (4.6)
\end{aligned}$$

式 (4.6) は Jensen の不等式を利用して変形している. ここで, $\epsilon_{ik} > 0$ は $\sum_i \epsilon_{ik} = 1$ を満たす補助変数であり, 式 (4.6) の等号成立条件は次式で与えられる.

$$\epsilon_{ik} = \frac{w_{ik}^2}{\sum_{i'} w_{i'k}^2} \quad (4.7)$$

以上より, 式 (4.2) の第一項及び第二項を式 (4.4) 及び式 (4.6) に置き換えた $\mathcal{J}_{\text{CSRNMF}}$ の上限関数 $\mathcal{J}_{\text{CSRNMF}}^+$ は次式となる.

$$\begin{aligned}
\mathcal{J}_{\text{CSRNMF}} \leq \mathcal{J}_{\text{CSRNMF}}^+ &= \sum_i \sum_j j \left(x_{ij} \log x_{ij} - x_{ij} \sum_k \delta_{ijk} \log \frac{w_{ik} h_{kj}}{\delta_{ijk}} - x_{ij} + \sum_k w_{ik} h_{kj} \right) \\
&\quad + \mu \sum_{k=1}^K \sum_{k'=k+1}^K \left[\left(\sum_{i'} w_{i'k'}^2\right)^{-\frac{1}{2}} \sum_{i'} \sum_{i''} w_{i'k'} \epsilon_{i''k}^{\frac{3}{2}} \frac{w_{i'k}}{w_{i''k}} \right] \quad (4.8)
\end{aligned}$$

これにより, 式 (4.8) を最小化することで最適化を行う. 次項では基底行列及び係数行列の更新式を導出する.

4.3.3 CSRNMF の更新式の導出

基底行列 \mathbf{W} の更新式は, 式 (4.8) を w_{ik} で偏微分し 0 と置いた次式から得られる.

$$\begin{aligned}
\frac{\partial \mathcal{J}_{\text{CSRNMF}}^+}{\partial w_{ik}} &= - \sum_j \frac{x_{ij} \delta_{ijk}}{w_{ik}} + h_{kj} \\
&\quad + \mu \sum_{k'=k+1}^K \left[\left(\sum_{i'} w_{i'k'}^2\right)^{-\frac{1}{2}} \left(-\frac{1}{w_{ik}^2} \epsilon_{ik}^{\frac{3}{2}} \sum_{i' \neq i} w_{i'k'} w_{i'k} + w_{ik'} \sum_{i' \neq i} \epsilon_{i'k}^{\frac{3}{2}} \frac{1}{w_{i'k}} \right) \right] \\
&= 0 \quad (4.9)
\end{aligned}$$

式 (4.9) の両辺に w_{ik}^2 を乗じると、次式となる。

$$\begin{aligned} -w_{ik} \sum_j x_{ij} \delta_{ijk} + w_{ik}^2 \sum_j h_{kj} - \mu \sum_{k'=k+1} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \epsilon_{ik}^{\frac{3}{2}} \sum_{i' \neq i} w_{i'k'} w_{i'k} \\ + \mu \sum_{k'=k+1} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} w_{ik}^2 w_{ik'} \sum_{i' \neq i} \epsilon_{i'k}^{\frac{3}{2}} \frac{1}{w_{i'k}} = 0 \end{aligned} \quad (4.10)$$

式 (4.10) を w_{ik} の次数ごとにまとめると、次式が得られる。

$$\begin{aligned} w_{ik}^2 \left[\sum_j h_{kj} + \mu \sum_{k'=k+1} w_{ik'} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \sum_{i' \neq i} \epsilon_{i'k}^{\frac{3}{2}} \frac{1}{w_{i'k}} \right] \\ + w_{ik} \left[-\sum_j x_{ij} \delta_{ijk} \right] + \left[-\mu \epsilon_{ik}^{\frac{3}{2}} \sum_{k'=k+1} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \left(\sum_{i' \neq i} w_{i'k'} w_{i'k} \right) \right] = 0 \end{aligned} \quad (4.11)$$

ここで、次式を用いて式 (4.11) の w_{ik}^2 , w_{ik} 及び第三項の係数をそれぞれ置き換える。

$$a_{ik} = \sum_j h_{kj} + \mu \sum_{k'=k+1} w_{ik'} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \left(\sum_{i' \neq i} \epsilon_{i'k}^{\frac{3}{2}} \frac{1}{w_{i'k}} \right) \quad (4.12)$$

$$b_{ik} = -\sum_j x_{ij} \delta_{ijk} \quad (4.13)$$

$$c_{ik} = -\mu \epsilon_{ik}^{\frac{3}{2}} \sum_{k'=k+1} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \left(\sum_{i' \neq i} w_{i'k'} w_{i'k} \right) \quad (4.14)$$

式 (4.12), 式 (4.13), 及び式 (4.14) を式 (4.11) に適用すると、次式が得られる。

$$a_{ik} w_{ik}^2 + b_{ik} w_{ik} + c_{ik} = 0 \quad (4.15)$$

式 (4.15) は w_{ik} の二次方程式となるため、上限関数 $\mathcal{J}_{\text{CSRNMF}}^+$ の w_{ik} に関する停留点は二次方程式の解の公式で次のように与えられる。

$$w_{ik} = \frac{-b_{ik} \pm \sqrt{b_{ik}^2 - 4a_{ik}c_{ik}}}{2a_{ik}} \quad (4.16)$$

ここで、式 (4.16) の \pm は w_{ik} が非負になるように決定する。 w_{ik} の更新則は等号成立条件の式 (4.5) 及び式 (4.7) を式 (4.16) に代入することで求めることができる。しかし、非常に複雑

になるため、代わりに a_{ik} , b_{ik} 及び c_{ik} の更新式をそれぞれ示す.

$$a_{ik} \leftarrow \sum_j h_{kj} + \mu \sum_{k'=k+1} w_{ik'} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \left(\sum_{i' \neq i} w_{i'k}^2 \left(\sum_{i''} w_{i''k}^2 \right)^{-\frac{3}{2}} \right) \quad (4.17)$$

$$b_{ik} \leftarrow - \sum_j x_{ij} \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}} \quad (4.18)$$

$$c_{ik} \leftarrow -\mu \left(\frac{w_{ik}^2}{\sum_{i'} w_{i'k}^2} \right)^{\frac{3}{2}} \sum_{k'=k+1} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \left(\sum_{i' \neq i} w_{i'k'} w_{i'k} \right) \quad (4.19)$$

以上より、更新後の w_{ik} は式 (4.17), 式 (4.18), 及び式 (4.19) によって更新した a_{ik} , b_{ik} , 及び c_{ik} を式 (4.16) の計算に用いることで得られる. なお, $\mu = 0$ のとき, 式 (4.16), 式 (4.17), 式 (4.18), 及び式 (4.19) は 2.2.2 項で示した KLNMF の基底行列 \mathbf{W} の更新式 (2.24) と一致する.

係数行列 \mathbf{H} の更新式は, 式 (4.8) を h_{kj} で偏微分し 0 と置いた次式から得られる.

$$\frac{\partial \mathcal{J}_{\text{CSRNMF}}^+}{\partial h_{kj}} = \sum_i -\frac{x_{ij} \delta_{ijk}}{h_{kj}} + w_{ik} = 0 \quad (4.20)$$

式 (4.20) を h_{kj} について解き, 等式条件の式 (4.5) を代入すると次式が得られる.

$$\begin{aligned} \frac{1}{h_{kj}} \sum_i -x_{ij} \delta_{ijk} + h_{kj} w_{ik} &= 0 \\ h_{kj} \sum_i w_{ik} &= \sum_i x_{ij} \delta_{ijk} \\ h_{kj} &= \frac{\sum_i x_{ij} \delta_{ijk}}{\sum_i w_{ik}} \\ &= \frac{\sum_i x_{ij} \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}}}{\sum_i w_{ik}} \\ &= h_{kj} \frac{\sum_i \frac{x_{ij}}{\sum_{k'} w_{ik'} h_{k'j}} w_{ik}}{\sum_i w_{ik}} \end{aligned} \quad (4.21)$$

これより, h_{kj} の更新式は次式となる.

$$h_{kj} \leftarrow h_{kj} \frac{\sum_i \frac{x_{ij}}{\sum_{k'} w_{ik'} h_{k'j}} w_{ik}}{\sum_i w_{ik}} \quad (4.22)$$

従って, CSRNMF の係数行列 \mathbf{H} は式 (4.22) によって更新される. これは 2.2.2 項で示した通常の KLNMF における係数行列 \mathbf{H} の更新式 (2.25) と一致する.

4.3.4 CSRNMF の行列補完への適用

CSRNMF に基づく行列補完は, 2.3 節で述べた手法と同様のアイデアによって実現する. 従って, 観測行列の欠損要素に対応する要素を 0, その他の要素を 1 とするバイナリマスク行

列 \mathbf{M} を定義することで, CSRNMF に基づく行列補完は次式で表される.

$$\text{Minimize}_{\mathbf{W}, \mathbf{H}} \mathcal{J}_{\text{CSRNMF}}(\mathbf{W}, \mathbf{H}) \text{ s.t. } w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k \quad (4.23)$$

$$\mathcal{J}_{\text{CSRNMF}}(\mathbf{W}, \mathbf{H}) = \sum_i \sum_j m_{ij} \mathcal{D}_{\text{KL}}(x_{ij} | \sum_k w_{ik} h_{kj}) + \mu \sum_{k=1} \sum_{k'=k+1} \mathcal{R}_{\text{cos}}(\mathbf{w}_k, \mathbf{w}_{k'}) \quad (4.24)$$

ここで, バイナリマスク行列は \mathbf{W} や \mathbf{H} に依らない定数行列であるため, 基底行列及び係数行列の更新式の導出は 4.3.3 項と同様に行うことができる. よって, 基底行列の更新式は式 (4.16) 同様に a_{ik} , b_{ik} 及び c_{ik} を用いることで次式となる.

$$w_{ik} = \frac{-b_{ik} \pm \sqrt{b_{ik}^2 - 4a_{ik}c_{ik}}}{2a_{ik}} \quad (4.25)$$

このとき, a_{ik} , b_{ik} 及び c_{ik} はそれぞれ次式となる.

$$a_{ik} \leftarrow \sum_j m_{ij} h_{kj} + \mu \sum_{k'=k+1} w_{ik'} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \left(\sum_{i' \neq i} w_{i'k}^2 \left(\sum_{i''} w_{i''k}^2 \right)^{-\frac{3}{2}} \right) \quad (4.26)$$

$$b_{ik} \leftarrow - \sum_j m_{ij} x_{ij} \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}} \quad (4.27)$$

$$c_{ik} \leftarrow -\mu \left(\frac{w_{ik}^2}{\sum_{i'} w_{i'k}^2} \right)^{\frac{3}{2}} \sum_{k'=k+1} \left(\sum_{i'} w_{i'k'}^2 \right)^{-\frac{1}{2}} \left(\sum_{i' \neq i} w_{i'k'} w_{i'k} \right) \quad (4.28)$$

また, 係数行列の更新式は次式となる.

$$h_{kj} \leftarrow h_{kj} \frac{\sum_i m_{ij} x_{ij} w_{ik}}{\sum_i m_{ij} w_{ik}} \quad (4.29)$$

以上で, 基底行列及び係数行列の更新式が示された.

4.4 L_1 -RNMF に基づく行列補完

4.4.1 L_1 -RNMF に基づく行列補完の目的関数

L_1 -RNMF に基づく行列補完は, 観測行列の欠損要素に対応する要素を 0, その他の要素を 1 とするバイナリマスク行列 \mathbf{M} を定義することで, 次式で表される.

$$\text{Minimize}_{\mathbf{W}, \mathbf{H}} \mathcal{J}_{L_1\text{-RNMF}}(\mathbf{W}, \mathbf{H}) \text{ s.t. } w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k \quad (4.30)$$

$$\mathcal{J}_{L_1\text{-RNMF}}(\mathbf{W}, \mathbf{H}) = \sum_i \sum_j m_{ij} \mathcal{D}_{\text{KL}}(x_{ij} | \sum_k w_{ik} h_{kj}) + \nu \|\mathbf{H}\|_1 \quad (4.31)$$

$$= \sum_i \sum_j m_{ij} \left(x_{ij} \log \frac{x_{ij}}{\sum_k w_{ik} h_{kj}} - x_{ij} + \sum_k w_{ik} h_{kj} \right) + \nu \sum_k \sum_j h_{kj} \quad (4.32)$$

4.4.2 補助関数法に基づく上限関数の導出

前項で示した式 (4.32) から, $\mathcal{J}_{L_1\text{-RNMF}}$ の上限関数を求める. 式 (4.31) の第一項と第二項を分けて考えると, 第二項については偏微分は容易であるため, 第一項のみについて導出すればよい.

式 (4.31) の第一項は, 2.2.2 項及び 4.3.2 項と同様に, Jensen の不等式を利用して次のように変形できる.

$$\sum_i \sum_j m_{ij} \left(x_{ij} \log x_{ij} - x_{ij} \sum_k \delta_{ijk} \log \frac{w_{ik} h_{kj}}{\delta_{ijk}} - x_{ij} + \sum_k w_{ik} h_{kj} \right) \quad (4.33)$$

ここで, δ_{ijk} は補助変数であり, $\sum_k \delta_{ijk} = 1$ を満たす. 等式条件は次式となる.

$$\delta_{ijk} = \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}} \quad (4.34)$$

よって, 式 (4.31) の第一項を式 (4.33) に置き換えた \mathcal{J}_2 の上限関数 \mathcal{J}_2^+ は次式となる.

$$\begin{aligned} \mathcal{J}_{L_1\text{-RNMF}} &\leq \mathcal{J}_{L_1\text{-RNMF}}^+ \\ &= \sum_i \sum_j m_{ij} \left(x_{ij} \log x_{ij} - x_{ij} \sum_k \delta_{ijk} \log \frac{w_{ik} h_{kj}}{\delta_{ijk}} - x_{ij} + \sum_k w_{ik} h_{kj} \right) \\ &\quad + \nu \sum_k \sum_j h_{kj} \end{aligned} \quad (4.35)$$

これにより, 最小化すべき上限関数が求まったため, 次項では基底行列及び係数行列の更新式を導出する.

4.4.3 L_1 -RNMF に基づく行列補完の更新式の導出

基底行列 \mathbf{W} の更新式は式 (4.35) を w_{ik} で偏微分し 0 と置いた次式から得られる.

$$\frac{\partial \mathcal{J}_{L_1\text{-RNMF}}^+}{\partial w_{ik}} = - \sum_j \left(\frac{m_{ij} x_{ij} \delta_{ijk}}{w_{ik}} + m_{ij} h_{kj} \right) = 0 \quad (4.36)$$

式 (4.36) を w_{ik} について解き, 等式条件の式 (4.34) を代入すると次式が得られる.

$$\begin{aligned}
 -\frac{1}{w_{ik}} \sum_j m_{ij} x_{ij} \delta_{ijk} + \sum_j m_{ij} h_{kj} &= 0 \\
 \frac{1}{w_{ik}} &= \frac{\sum_j m_{ij} h_{kj}}{\sum_j m_{ij} x_{ij} \delta_{ijk}} \\
 w_{ik} &= \frac{\sum_j m_{ij} x_{ij} \delta_{ijk}}{\sum_j m_{ij} h_{kj}} \\
 w_{ik} &= \frac{\sum_j m_{ij} x_{ij} \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}}}{\sum_j m_{ij} h_{kj}} \\
 w_{ik} &= w_{ik} \frac{\sum_j \frac{m_{ij} x_{ij}}{\sum_{k'} w_{ik'} h_{k'j}} h_{kj}}{\sum_j m_{ij} h_{kj}} \quad (4.37)
 \end{aligned}$$

これより, w_{ik} の更新式は次式となる.

$$w_{ik} \leftarrow w_{ik} \frac{\sum_j \frac{m_{ij} x_{ij}}{\sum_{k'} w_{ik'} h_{k'j}} h_{kj}}{\sum_j m_{ij} h_{kj}} \quad (4.38)$$

従って, L_1 -RNMF に基づく行列補完の基底行列 \mathbf{W} は式 (4.38) によって更新され, これは 2.3.1 項で示した通常の NMF に基づく行列補完における基底行列 \mathbf{W} の更新式 (2.32) と一致する.

係数行列 \mathbf{H} の更新式は式 (4.35) を h_{kj} で偏微分し 0 と置いた次式から得られる.

$$\frac{\partial \mathcal{J}_{L_1\text{-RNMF}}^+}{\partial h_{kj}} = \sum_i \left(-\frac{m_{ij} x_{ij} \delta_{ijk}}{h_{kj}} + m_{ij} w_{ik} \right) + \nu = 0 \quad (4.39)$$

式 (4.39) を h_{kj} について解くと, 次式を得る.

$$\begin{aligned}
 \frac{1}{h_{kj}} \sum_i -m_{ij} x_{ij} \delta_{ijk} + \sum_i m_{ij} w_{ik} + \nu &= 0 \\
 \frac{1}{h_{kj}} \sum_i m_{ij} x_{ij} \delta_{ijk} &= \sum_i m_{ij} w_{ik} + \nu \\
 h_{kj} &= \frac{\sum_i m_{ij} x_{ij} \delta_{ijk}}{\sum_i m_{ij} w_{ik} + \nu} \quad (4.40)
 \end{aligned}$$

補助変数 δ_{ijk} に等式条件の式 (4.34) を代入すると次式が得られる.

$$\begin{aligned}
 h_{kj} &= \frac{\sum_i m_{ij} x_{ij} \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}}}{\sum_i m_{ij} w_{ik} + \nu} \\
 &= h_{kj} \frac{\sum_i \frac{m_{ij} x_{ij} w_{ik}}{\sum_{k'} w_{ik'} h_{k'j}}}{\sum_i m_{ij} w_{ik} + \nu} \quad (4.41)
 \end{aligned}$$

これより, h_{kj} の更新式は次式となる.

$$h_{kj} = h_{kj} \frac{\sum_i m_{ij} x_{ij} w_{ik}}{\sum_i m_{ij} w_{ik} + \nu} \quad (4.42)$$

従って, L_1 -RNMF に基づく行列補完の基底行列 \mathbf{H} は式 (4.42) によって更新される. なお, $\nu = 0$ のとき, 式 (4.42) は 2.3.1 項で示した通常の NMF に基づく行列補完における係数行列 \mathbf{H} の更新式 (2.33) と一致する.

注意すべき点として L_1 -RNMF において, 2.1 節で述べたスケールの任意性は大きな問題となる. これは, 最小化問題の目的関数の式 (4.31) において, 正則化項 $\|\mathbf{H}\|_1$ は係数行列 \mathbf{H} のスケールに依存して小さくなるためである. 例えば, 限りなく 0 に近い値 $\varepsilon > 0$ を \mathbf{H} に乗じて $\varepsilon\mathbf{H}$ と更新だけで $\|\mathbf{H}\|_1$ を限りなく 0 に近づけることができる. このとき, 基底行列 \mathbf{W} は $\varepsilon^{-1}\mathbf{W}$ と更新することで \mathbf{WH} の値が不変となり, 式 (4.31) の第 1 項の値は変わらない. 従って, 最小化問題 (4.30) は厳密には最小解の無い問題であり解くことはできない [17]. この問題に対して, 基底ベクトルを正規化した最適化問題を考えて対処する手法 [14, 17] も提案されているが, 本論文では NMF の文脈でよく用いられてきたヒューリスティックな対処法である更新毎の \mathbf{W} 及び \mathbf{H} の正規化を用いた L_1 -RNMF を取り扱う. 具体的には, 式 (4.38) 及び式 (4.42) で \mathbf{W} 及び \mathbf{H} を更新した後に, 正規化係数 C_1, C_2, \dots, C_K を用いて次式の正規化処理を適用する.

$$C_k = \sum_i w_{ik} \quad (4.43)$$

$$w_{ik} = \frac{w_{ik}}{C_k} \quad \forall i, k \quad (4.44)$$

$$h_{kj} = C_k h_{kj} \quad \forall k, j \quad (4.45)$$

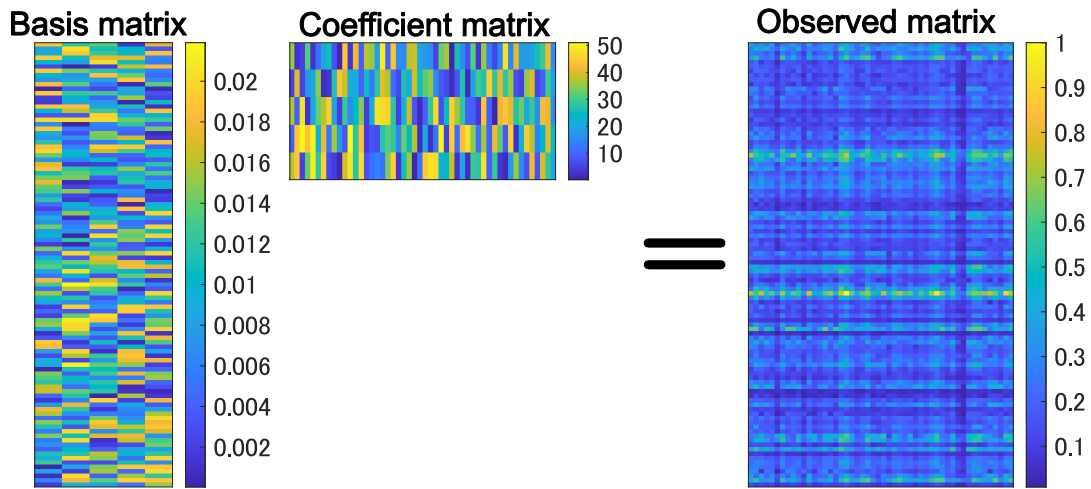
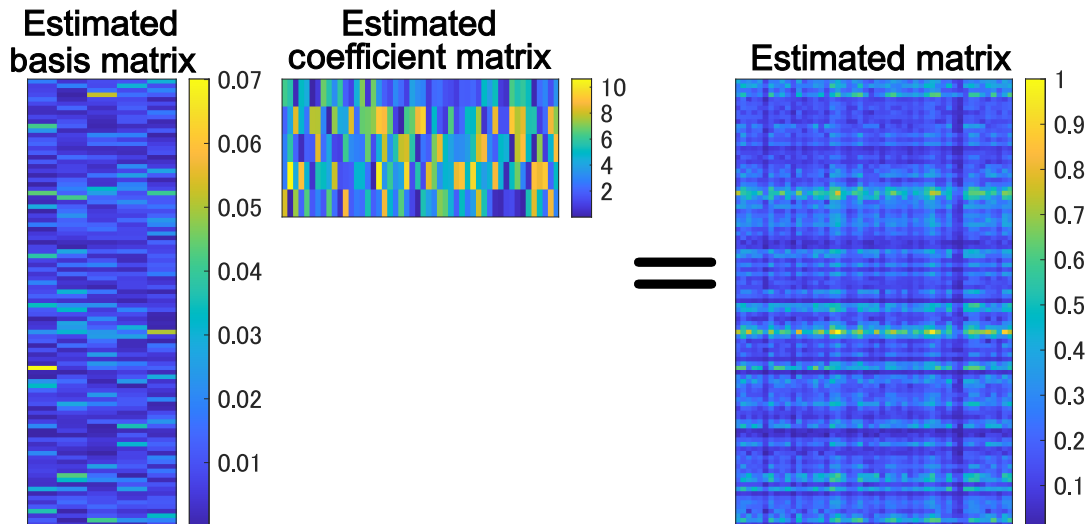
この正規化式の前後で目的関数の式 (4.31) の値は不変であるため, この処理により間接的に L_1 ノルムの正則化を変数の最適化に影響させ続けることができる.

4.5 正則化項による NMF の挙動変化

4.3 節及び 4.4 節によって, 正則化付き NMF の反復更新式を導出した. ここでは, コサイン類似度正則化項及び L_1 ノルム正則化項が NMF に与える影響について確認する.

Fig. 4.1 に区間 (0, 1) の一様分布乱数を用いて生成した観測行列を示す. また, この観測行列を KLNMF, CSRNMF, 及び L_1 -RNMF によって基底行列と係数行列に分解した結果を Figs. 4.2–4.4 にそれぞれ示す. Figs. 4.2–4.4 は, 基底ベクトル数 K を 5 としてそれぞれの NMF によって分解した結果である. ここで, CSRNMF の重み係数 μ は 100000, L_1 -RNMF の重み係数 ν は 100000 としている.

Figs. 4.2–4.4 より, それぞれの NMF における近似行列を比較すると, 通常の NMF が最も正確に観測行列の細かな値の変化まで近似できていることがわかる. 一方で, 正則化付き

Fig. 4.1. Observed matrix of size 100×50 and rank 5.Fig. 4.2. Example of matrix decomposition by KLNMF when size of observed matrix is 100×50 and its rank is 5.

NMF による近似行列は、いずれも観測行列の各行、各列の値を平滑化したような近似となっており、観測行列のパターン、特徴を強調するような結果となっていることがわかる。

Fig. 4.3 より、コサイン類似度正則項付き NMF の基底行列に注目すると、0 に近い値を持つ要素が増加し、各行において値を持つ要素がほとんど 1 つになっていることがわかる。これより、基底行列の列ベクトルを基底ベクトルとして、各基底ベクトル同士のコサイン類似度を考えると、内積が 0 になることから、それぞれの基底ベクトル同士において直交していると考えられる。よって、コサイン類似度正則項によって基底行列の基底ベクトル間を直交できることが確認でき、これにより補完後の推定行列はパターンが著しく強調された行列となっていることが見て取れる。実際には欠損していない要素は観測値をそのまま利用できる為、補完結果は Fig. 4.3 の推定行列ほど改変されたものにはならないが、それでもこの結果は CSRNMF

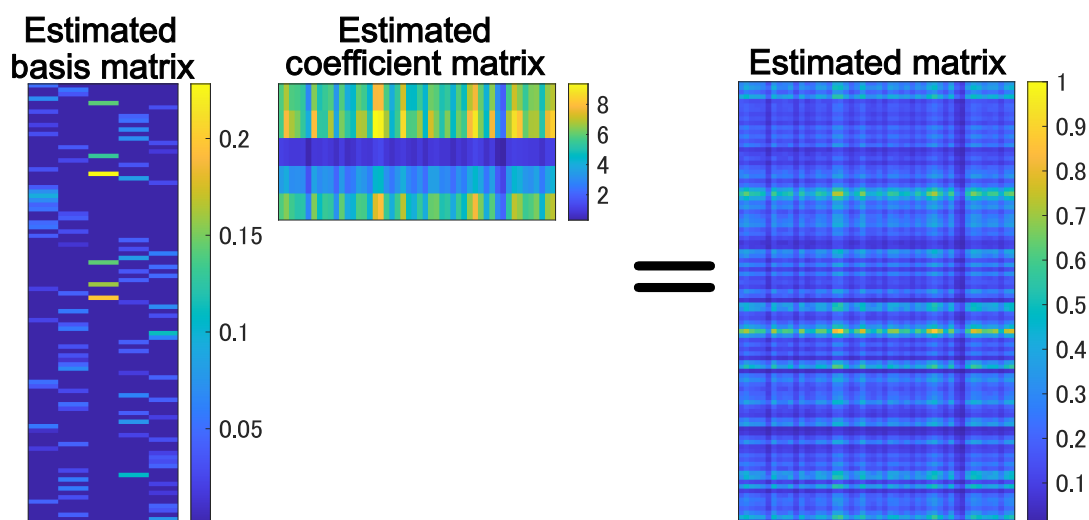


Fig. 4.3. Example of matrix decomposition by CSRNMF when size of observed matrix 100×50 , its rank is 5, and regularization coefficient is set to $\mu = 100000$.

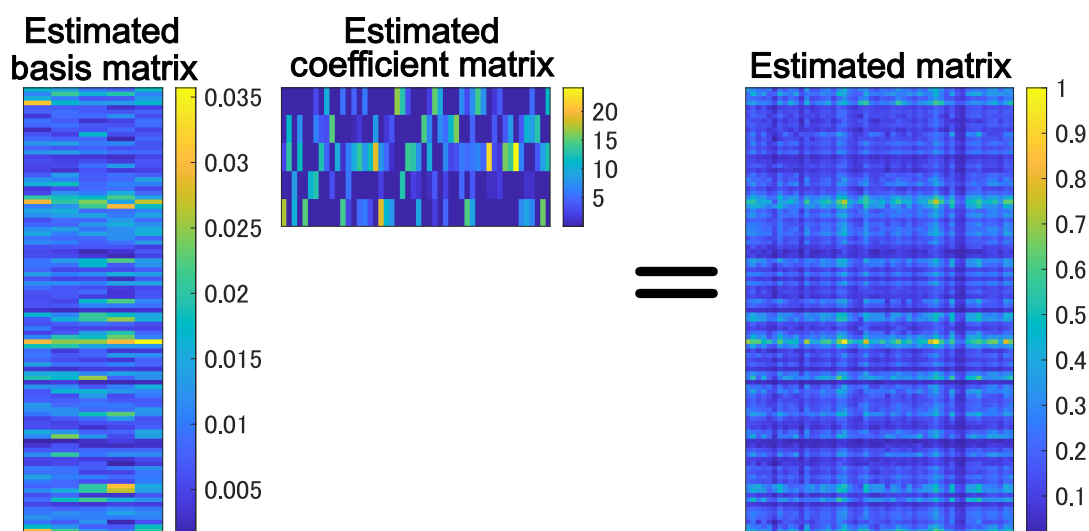


Fig. 4.4. Example of matrix decomposition by L_1 -RNMF when size of observed matrix 100×50 , its rank is 5, and regularization coefficient is set to $\nu = 10000$.

が強力な特徴抽出（基底ベクトルの推定）をもたらすことを示唆している。

Fig. 4.4 より、 L_1 -RNMF の基底行列に注目すると、係数行列の値を持つ要素が Fig. 4.2 に比べ少なくなっていることがわかる。コサイン類似度正則項付き NMF が基底行列要素が減少していたのに対し、 L_1 -RNMF においては係数行列の要素が減少しており、正解行列の各要素を少ない要素で表現できるように分解していることがわかる。これより、 L_1 ノルム正則化によって、値を持つ要素を抑制できていることが確認でき、観測行列の持つパターンを鮮明にした近似行列を求められていることがわかる。

以上の結果は、理想的な場合において KLNMF が最良であることも表している。正則化の

恩恵は、4.2 節で示したように、観測行列にノイズが含まれている場合や、正解行列のランクと基底ベクトル数 K 間で誤差が生じた場合にもたらされることを期待している。次章では、欠損補完というタスクにおいて、そのような条件で各正則化項がどの程度性能改善に寄与するかについて調査する。

4.6 本章のまとめ

本章では、CSRNMF 及び L_1 -RNMF に基づく行列補完の定式化をそれぞれ示した。4.5 節で述べた通り、それぞれの正則化項はおおむね目的の通りの挙動を示すことを確認できた。次章では、人工的に欠損行列を生成し、欠損行列に対してそれぞれの NMF に基づく行列補完を実験することで、性能を評価する。

第 5 章

人工欠損行列の補完実験

5.1 はじめに

本章では、4章で示した行列補完正則化付き NMF の性能について、人工的に作成した欠損乱数データを用いた実験により検証する。5.2 節では、補完実験の方法について説明する。5.3 節では、補完実験の結果について解説し、各正則化付き NMF の性能を比較する。5.4 節で本章についてまとめる。

5.2 実験方法

正則化項の付与による行列補完の性能変化を確認するために、通常の KLNMF, CSRNMF, 及び L_1 -RNMF の 3 手法を対象とした人工欠損行列の補完実験を行う。本実験では、乱数によって生成した正解行列 \mathbf{T} に対しノイズを加算した後に欠損を加えた観測行列 \mathbf{X} を用意し、調査対象とする NMF によって行列補完を行い、その補完性能を調査する。

観測行列の生成過程を、Fig. 5.1 に示す。初めに、ランク K_{true} の正解行列 $\mathbf{T} \in \mathbb{R}^{I \times J}$ をサイズが $I \times K_{\text{true}}$ 及び $K_{\text{true}} \times J$ の乱数行列の行列積により生成した。乱数には区間 $(0, 1)$ の一様分布乱数を用いた。次に、正解行列 \mathbf{T} に上乘せするノイズ行列 $\mathbf{N} \in \mathbb{R}_{\geq 0}^{I \times J}$ を作成した。ノイズ行列の生成には区間 $(0, 1)$ の一様分布及び母数 $\sigma = 1.0$ のレイリー分布に従う乱数の 2 種類を用いた。レイリー分布とは、時間領域において発生した正規分布に従うノイズ（白色雑音）が、STFT と絶対値処理により振幅スペクトログラムに変換されたときに従う確率分布であり、次章で述べる音響信号の振幅スペクトログラムにおける欠損補完を応用先に考慮した実験条件である。次に、正解行列 \mathbf{T} とノイズ行列 \mathbf{N} の大きさが所望の信号雑音比 (signal-to-noise ratio: SNR) となるような係数 d を次式で求めた。

$$d = 10^{\frac{\text{SNR} - r}{20}} \quad (5.1)$$

$$\text{SNR} = 10 \log_{10} \left(\frac{\|\mathbf{T}\|_2}{\|\mathbf{N}\|_2} \right) \quad (5.2)$$

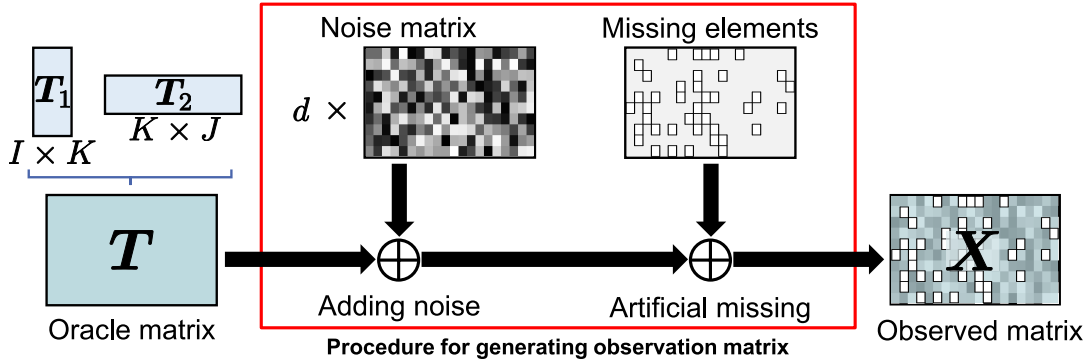


Fig. 5.1. Process flow of producing observed matrix.

ここで、 r は所望の SNR dB, $\|\cdot\|_2$ は次式で定義される L_2 ノルムである.

$$\|\mathbf{X}\|_2 = \sqrt{\sum_i \sum_j x_{ij}^2} \quad (5.3)$$

従って、所望の SNR r を持つ行列は $\mathbf{T} + d\mathbf{N}$ となる. このようにノイズを正解行列 \mathbf{T} に付加した後、全ての要素を確率 $\alpha \in [0, 1]$ で欠損させた.

以上の手順により、NMF の補完対象となる観測行列 \mathbf{X} を作成した. 観測行列 \mathbf{X} をもとに、どれだけ正解行列 \mathbf{T} に近い補完ができるかを評価することで、各手法の性能を比較する. 正解行列 \mathbf{T} にどれだけ近づけることができたかという評価基準には欠損を与えた全要素ごとの一般化 KL ダイバージェンスの総和を用いた. 即ち、値が小さいほど、より補完性能が高いと評価できる.

実験手順と各処理で使用するパラメータを Fig. 5.2 に示す. Fig. 5.2 に示す通り、行列の補完処理においては NMF に用いるランクの推定値と正則化項の重み係数を調整する. 本実験では、正解行列のサイズを $I = 100$ かつ $J = 200$, ランク $K_{\text{true}} = 10$ に固定し、Fig. 5.2 の各パラメータについて網羅的に条件を変更して実験を行った. 詳細な実験条件は Table 5.1 に示す.

5.3 実験結果

Figs. 5.3–5.8 は、ノイズにレイリー分布を用いたときの、SNR r を 0.88 dB, 5.99 dB, 及び 21.54 dB と設定した場合の各正則化付き NMF による行列補完の実験結果をそれぞれ示したヒートマップ図である. また、その他の条件における結果を付録 B に示す. ヒートマップの縦軸は欠損率 α , 横軸は NMF の基底ベクトル数 K (即ち、正解行列 \mathbf{T} のランクの予測値であり $K = 10$ で誤差の無い理想的な条件) を示している. 各セルは対応する α 及び K の条件下において、正解行列及び補完行列間の誤差に関する評価値を、赤色から青色のグラデーションで示している. ヒートマップ上の各色が示す具体的な値は、ヒートマップ右側のカラーバーによって確認することができる. このカラーバーの値は、正則化項の無い NMF に基づく

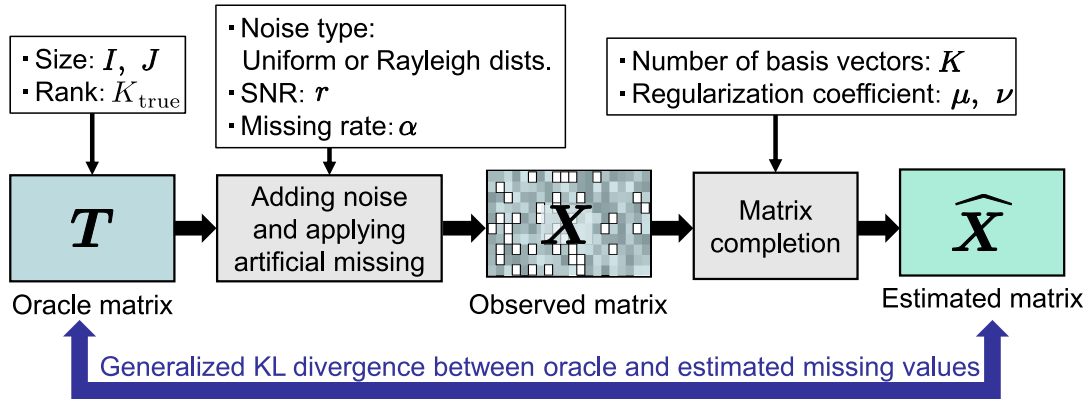


Fig. 5.2. Experimental conditions changed in evaluation and performance criterion of matrix completion.

Table 5.1. Experimental conditions

| | |
|---|--|
| Noise type | Uniform dist./Rayleigh dist. |
| Number of basis vectors K | 6, 7, \dots , 20 |
| SNR r [dB] | 10^{-2} to 10^3 divided by 40 on a logarithmic scale |
| Missing rate α | 0 to 0.8 divided by 20 on a linear scale |
| Regularization coefficient μ and ν | 10^{-3} to 10^5 divided by 40 on a logarithmic scale |
| Initialization of \mathbf{W} and \mathbf{H} | Uniformly distributed random values in the range (0,1) |
| Number of iterations for parameter updates | 500 |
| Number of trials | 10 with different random seeds |

行列補完（従来手法）の結果の値との差分を表している。従って、0の色は提案手法と従来手法が全く同じ性能であることを表し、負値になるほど提案手法の方が高精度な補完を達成できていることを表している。即ち、ヒートマップにおいてはより赤いほど性能が向上している。

Fig. 5.3 はノイズにレイリー分布を用いた時の、SNR $r = 0.88$ dB の条件下の各欠損率 α 及び基底ベクトル数 K における CSRNMF の補完実験結果である。この時、重み係数 $\mu = 61584.82$ である。Fig. 5.3 の結果より、欠損率 α がおよそ 50% より大きい範囲について、CSRNMF の補完精度が KLNMF を上回っている。一方で、欠損率 α が 50% を下回る範囲のほとんどにおいては改善が見られなかった。

Fig. 5.4 はノイズにレイリー分布を用いた時の、SNR $r = 5.99$ dB の条件下の結果である。

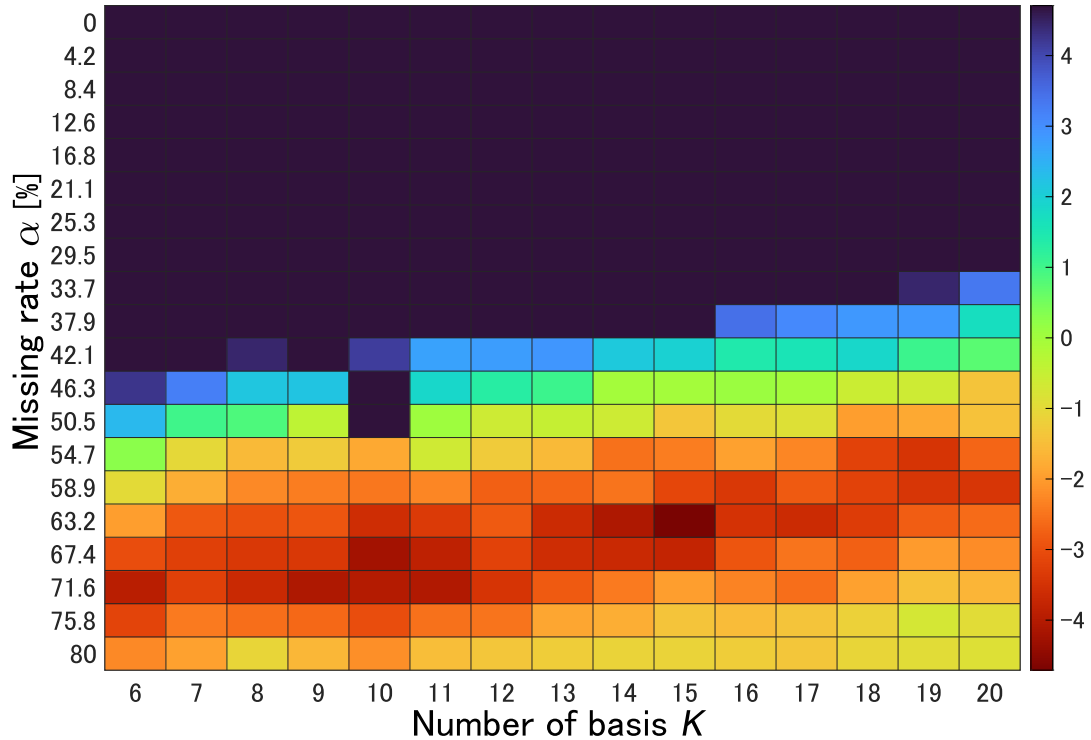


Fig. 5.3. Improvement value obtained by CSRNMf: noise type is Rayleigh dist., $r = 0.88$ dB, and $\mu = 61584.82$.

この時、重み係数 $\mu = 0.55$ である。Fig. 5.4 より、基底ベクトル数 K が正解行列のランクである 10 に対して小さく、かつ欠損率 α が小さい範囲で改善を示している。補完精度が改善した範囲について、SNR $r = 0.88$ dB の条件下における改善量と比べると小さくなっている。一方で、その他の広い範囲については目立つ改善は得られなかった。

Fig. 5.5 はノイズにレイリー分布を用いた時の、SNR = 21.54 dB の条件下の各欠損率 α 及び基底ベクトル数 K における CSRNMf の補完実験結果である。この時、重み係数 $\mu = 0.55$ である。Fig. 5.5 より、基底ベクトル数 K が正解行列のランクである 10 に対して小さく、かつ欠損率 α が小さい範囲で改善を示していることがわかる。一方で、この他の範囲については目立つ改善は得られなかった。

以上より、CSRNMf はノイズが大きい時、欠損率 α が大きい範囲において基底ベクトル数 K によらず同程度の改善を示し、ノイズが小さい時、欠損率 α 及び基底ベクトル数 K が正解行列のランクより小さい時において改善を示す。ノイズ及び欠損率 α が大きい条件とは、正解行列の特徴抽出が最も困難な状態であると言える。また、ノイズが小さく欠損率 α 及び基底ベクトル数 K が正解行列のランクより小さい条件というのは、正解行列に近い値を取る要素の多くが既知の中で、正解行列を構成する特徴の中でより重要な特徴を抽出する必要がある条件であると言える。このような条件において改善を示したことより、コサイン類似度正則化項が観測行列から正解行列の構造を抽出する効果を NMF に与えていると考えられる。つまり、コサイン類似度正則化項により、KLNMF よりも強力な特徴抽出が可能であると言える。

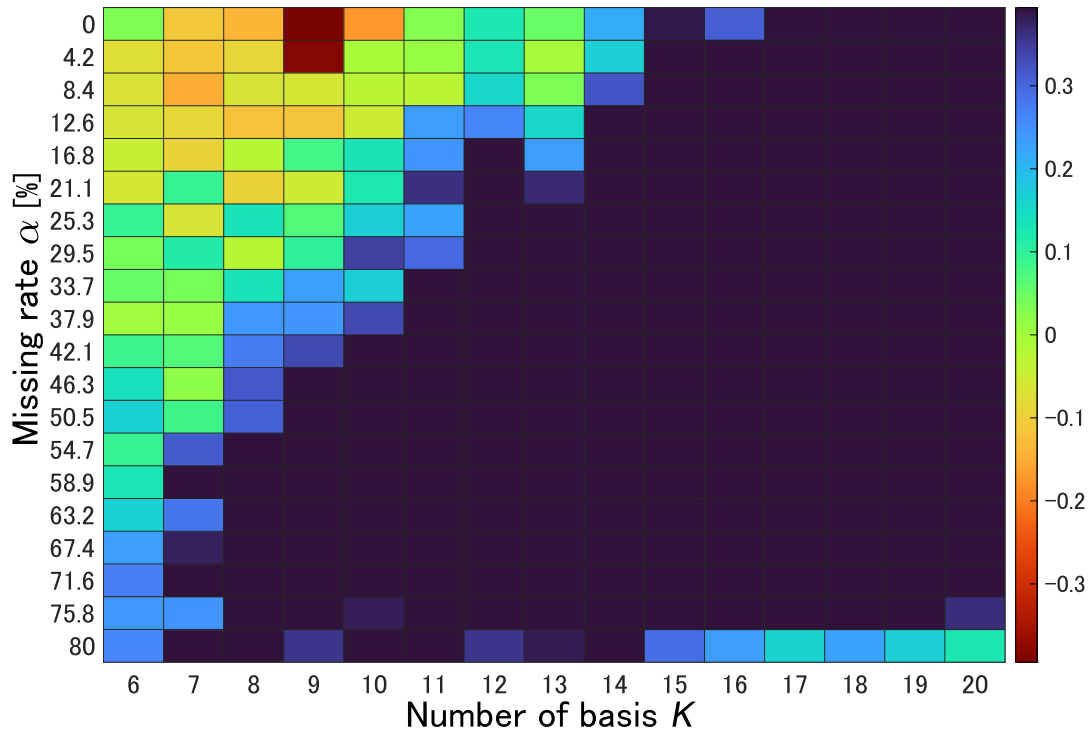


Fig. 5.4. Improvement value obtained by CSRNMf: noise type is Rayleigh dist., $r = 5.99$ dB, and $\mu = 0.55$.

Fig. 5.6 はノイズにレイリー分布を用いた時の、SNR $r = 0.88$ dB の条件下の各欠損率 α 及び基底ベクトル数 K における L_1 -RNMF の補完実験結果である。この時、重み係数 $\nu = 1.44$ である。Fig. 5.6 より、実験を行った全ての範囲について、 L_1 -RNMF の補完精度が KLNMF を上回っていることがわかる。総じて、補完精度は欠損率 α が低いほど改善量が大きく、同じ欠損率 α について基底ベクトル数 K の値によらず同程度の性能であることがわかる。

Fig. 5.7 はノイズにレイリー分布を用いた時の、SNR $r = 5.99$ dB の条件下の各欠損率 α 及び基底ベクトル数 K における L_1 -RNMF の補完実験結果である。この時、重み係数 $\nu = 0.89$ である。Fig. 5.7 より、実験を行ったほぼ全ての範囲について、 L_1 -RNMF の補完精度が KLNMF を上回っていることがわかる。総じて SNR $r = 0.88$ dB の条件時と同様に、補完精度は欠損率 α が低いほど改善量が大きく、同じ欠損率 α について基底ベクトル数 K の値によらず同程度の性能であることがわかる。但し、欠損率 $\alpha = 80\%$ に近づくと、改善がほとんど無いか、KLNMF の方が補完精度が良いところが現れ、SNR $r = 0.88$ dB の条件時よりは改善が減少している。

Fig. 5.8 はノイズにレイリー分布を用いた時の、SNR $r = 21.54$ dB の条件下の各欠損率 α 及び基底ベクトル数 K における L_1 -RNMF の補完実験結果である。この時、重み係数 $\nu = 0.08$ である。Fig. 5.8 より、欠損率 α がおよそ 50% を下回る範囲について、 L_1 -RNMF の補完精度が KLNMF を上回っていることがわかる。総じて SNR $r = 0.88$ dB 及び SNR $r = 5.99$ dB の条件下同様、補完精度は欠損率 α が低いほど改善量が大きく、同じ欠損率 α について基底

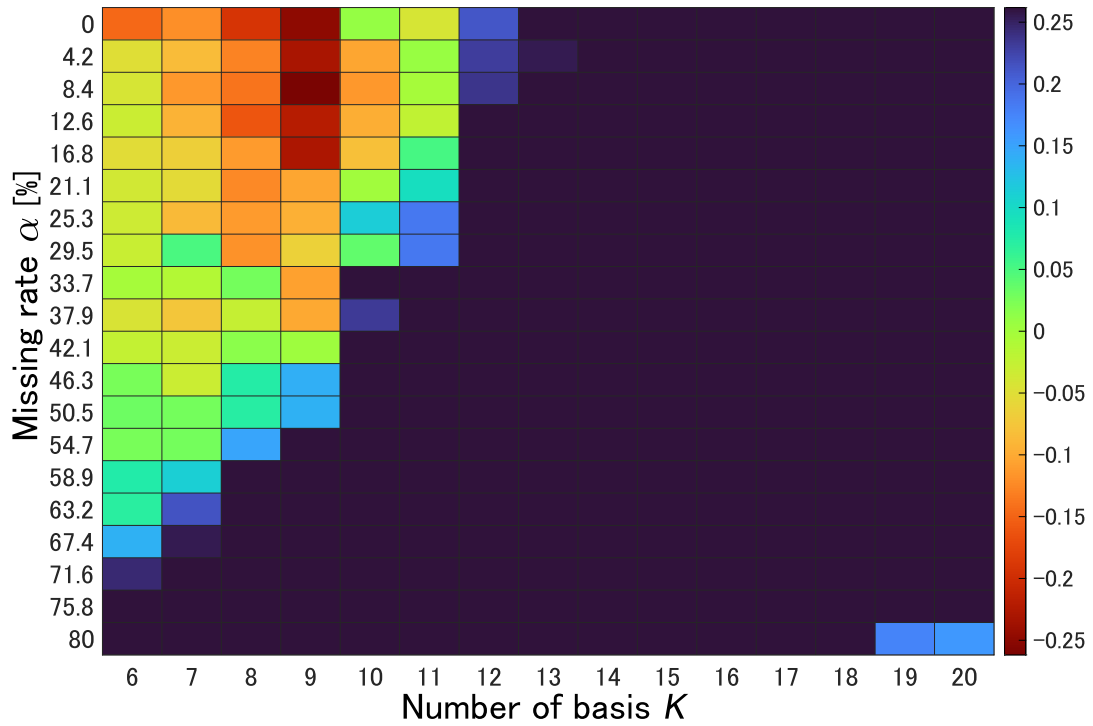


Fig. 5.5. Improvement value obtained by CSRNMf: noise type is Rayleigh dist., $r = 21.54$ dB, and $\mu = 0.55$.

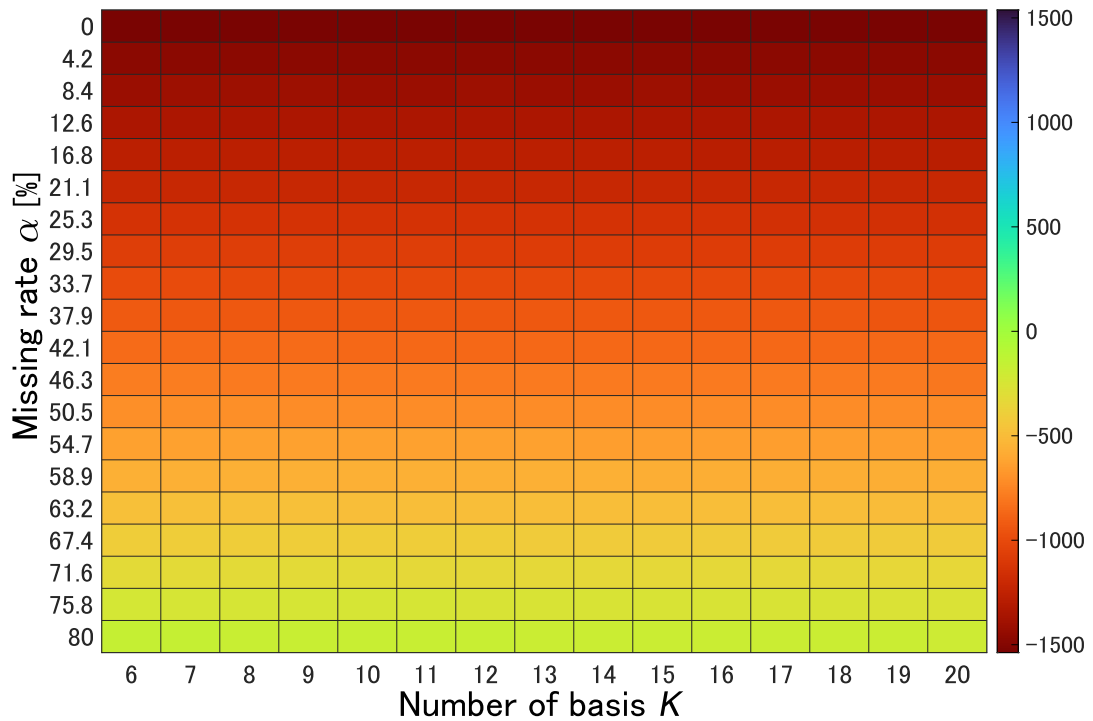


Fig. 5.6. Improvement value obtained by L_1 -RNMF: noise type is Rayleigh dist., $r = 0.88$ dB, and $\nu = 1.44$.

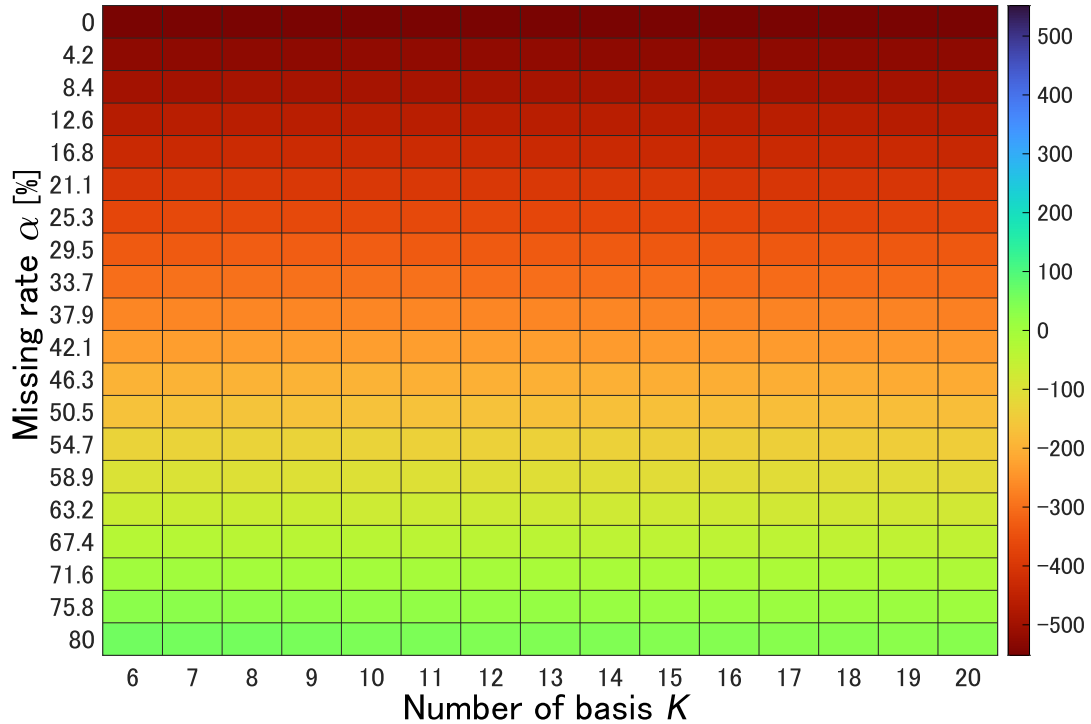


Fig. 5.7. Improvement value obtained by L_1 -RNMF: noise type is Rayleigh dist., $r = 5.99$ dB, and $\nu = 0.89$.

ベクトル数 K の値によらず同程度の性能であることがわかる。一方で、欠損率 α が 50% を超えると KLNMF の補完精度を下回りはじめ、SNR $r = 5.99$ dB の条件下より広い範囲で改善が得られなかった。

以上より、 L_1 -RNMF はノイズの大小に関係なく、欠損率 α が小さい時に基底ベクトル数 K によらず大きな改善を示し、欠損率 α が大きくなるにつれ改善量が小さくなっていった。即ち、ノイズの影響を受けずに補完精度を向上させており、基底ベクトル数 K によって改善量が変化しないことから基底ベクトル数 K の誤差に対する特性は KLNMF による行列補完と同様であると言える。これより、 L_1 ノルム正則化項はノイズに対する頑健性を向上させており、行列補完の精度を底上げする正則化項であると言える。

CSRNMF と L_1 -RNMF の補完精度を比較すると、いずれの条件においても L_1 -RNMF の補完精度が上回ることがわかる。但し、CSRNMF は改善量は小さい一方で、SNR $r = 0.88$ dB のようにノイズが特に大きい時に欠損率 α が大きい範囲で良い性能を発揮しており、欠損率 α が小さい範囲を得意とする L_1 -RNMF とは異なる特徴を持つことがわかる。それぞれの正則化付き NMF の実験結果の特徴をまとめる。

いずれの正則化付き NMF に共通することとして、ノイズが大きい時ほど改善量が大きく、ノイズが小さくなると改善量も減少していた。これは、Fig. 5.9 に示す結果より説明できる。Fig. 5.9 は各 SNR の条件における、KLNMF による補完行列と正解行列との一般化 KL ダイバージェンスにおける距離の推移である。ノイズにはレイリー分布を用いている。Fig. 5.9 よ

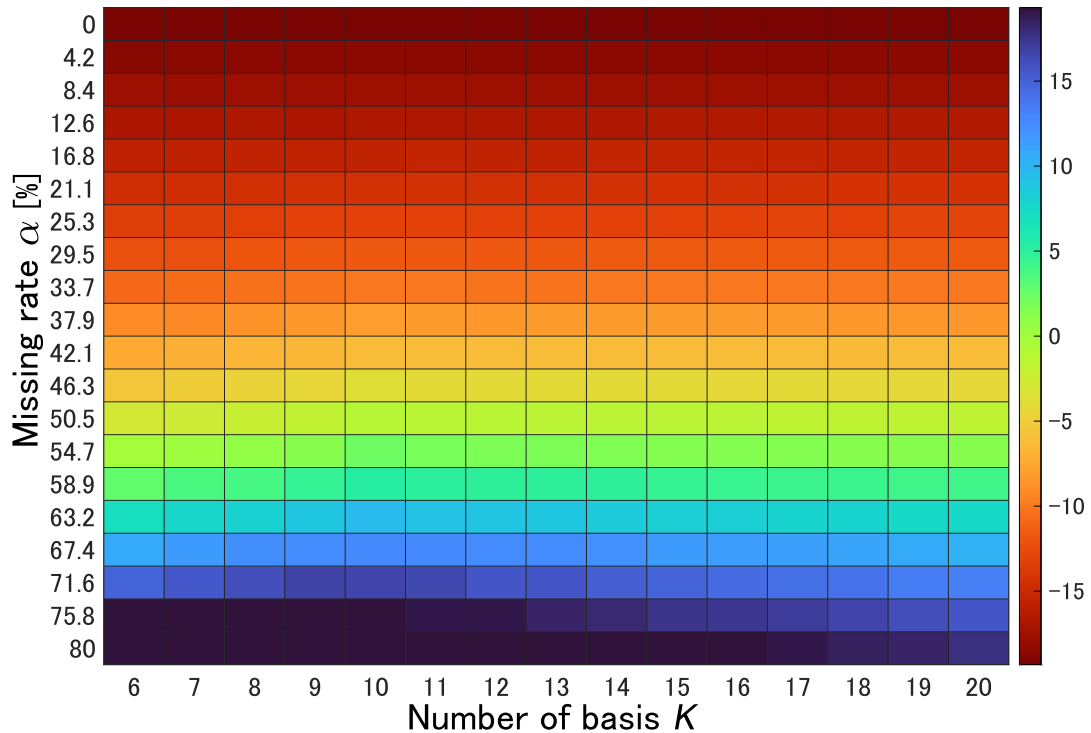


Fig. 5.8. Improvement value obtained by L_1 -RNMF: noise type is Rayleigh dist., $r = 21.54$ dB, and $\nu = 0.08$.

り、ノイズが小さくなるにつれ、KLNMFによる行列補完の精度が増す。つまり、ノイズが小さくなると行列補完の精度が増すことで一般化 KL ダイバージェンスの値が小さくなり、改善の余地が減少する。このため、改善があった場合もその値が小さくなっている。また、いずれの正則化付き NMF においても、ノイズが大きい時には重み係数 μ 及び ν が大きいほどよい性能を示し、ノイズが小さい時には重み係数の値が小さい時によい性能を示している。これは、正則化の影響を大きくすることで、Fig. 4.3 及び Fig. 4.4 の基底行列または係数行列のような推定する傾向が強くなる。このような基底行列または係数行列は、観測行列の中に含まれる「低ランク性に寄与しない成分」を無視するような推定となっている。そのため、観測行列に含まれるノイズの影響をあまり受けずに正解行列の低ランク構造を抽出することができ、結果としてノイズに対する頑健性が向上しているといえる。

5.4 本章のまとめ

本章では、人工欠損行列を用いた補完実験により、それぞれの正則化付き NMF の行列補完の精度について検証した。いずれの正則化付き NMF においても改善が得られた一方で、補完対象がノイズ行列であるため、実用性が検証可能な実験条件であるとは言えない。そこで、次章では音響信号の振幅スペクトログラムに欠損を与え、これの補完実験によって補完性能の検証を行う。

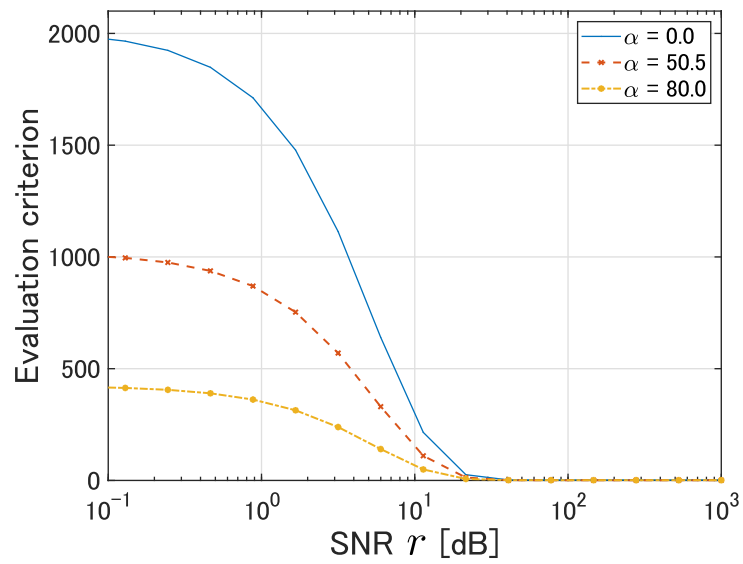


Fig. 5.9. Evaluation criterion for each SNR conditions: : noise type is Rayleigh dist. and $K = 10$.

第 6 章

欠損音声スペクトログラムの補完 実験

6.1 はじめに

本章では、音源分離等の処理を通して音響信号に発生しうる欠損値の補完について、正則化付き NMF の補完精度の評価を行う。6.2 節では、NMF によって音声信号を扱う上で欠かせない STFT について説明する。また、本実験における欠損原因となる時間周波数マスクについて説明する。6.3 節では実験の詳細と各種条件について説明する。6.4 節では実験結果を示し、各正則化付き NMF の補完性能と特徴について述べる。6.5 節で本章についてまとめる。

6.2 実験背景

本実験では、音響信号の振幅スペクトログラムに発生する欠損について補完実験を行う。本節では、実験において、音声信号を非負値行列（振幅スペクトログラム）に変換するために用いる STFT と本実験が仮定する欠損発生メカニズムである時間周波数マスクについて説明する。

6.2.1 STFT

観測された音響信号を非負行列として表現する手法として、STFT がある。STFT とは Fig. 6.1 に示すように、一次元の時間信号を二次元の時間周波数信号に変換する処理である。STFT の分析窓関数の長さ及びシフト長をそれぞれ Q 及び τ としたとき、時間領域の信号 $z(p)$ ($p = 1, 2, \dots, P$ は離散時間サンプルのインデックスを表す) の j 番目の短時間区間（時間フレーム）の信号は次式で表される。

$$\begin{aligned} \mathbf{z}^{(j)} &= [z((j-1)\tau+1), z((j-1)\tau+2), \dots, z((j-1)\tau+Q)]^T \\ &= [z^{(j)}(1), z^{(j)}(2), \dots, z^{(j)}(q), \dots, z^{(j)}(Q)]^T \in \mathbb{R}^Q \end{aligned} \quad (6.1)$$

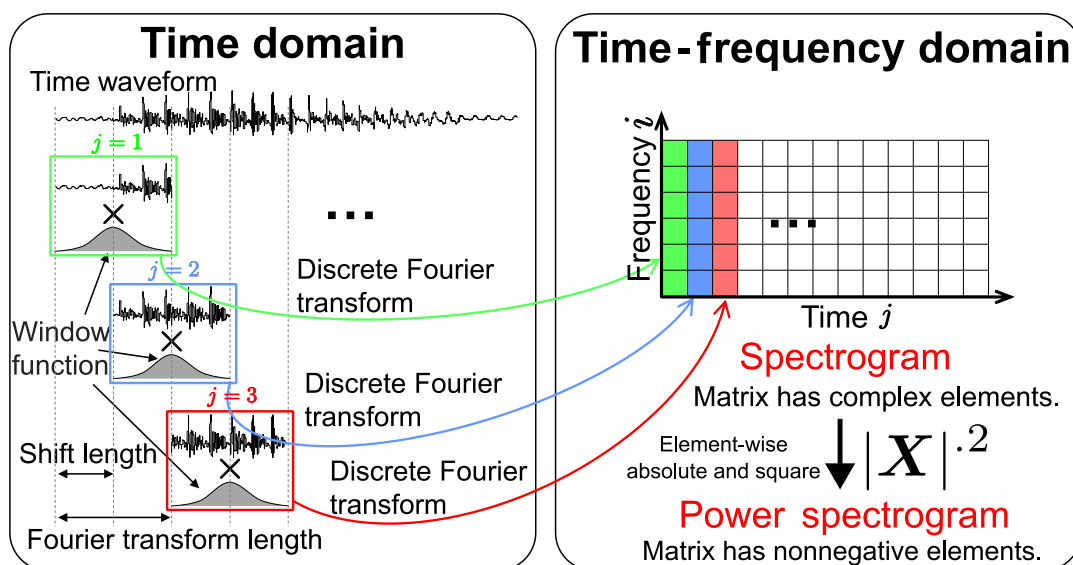


Fig. 6.1. Process of STFT.

ここで、 $j = 1, 2, \dots, J$ 及び $q = 1, 2, \dots, Q$ は、それぞれ時間フレーム及び時間フレーム内のサンプルのインデクスを示す。また、時間フレーム数 J は次式によって与えられる。

$$J = \frac{P}{\tau} \quad (6.2)$$

但し、信号長 L はセグメント数 J が整数となるように各時間フレームの信号の両端にゼロを挿入する処理（ゼロパディング）が施されている。式 (6.1) で定義される時間フレームの信号を全ての j についてまとめた全時間フレームの信号を $\mathbf{z} = [\mathbf{z}^{(1)} \ \mathbf{z}^{(2)} \ \dots \ \mathbf{z}^{(J)}] \in \mathbb{R}^{Q \times J}$ と表記すると、STFT の処理は次式のように表される。

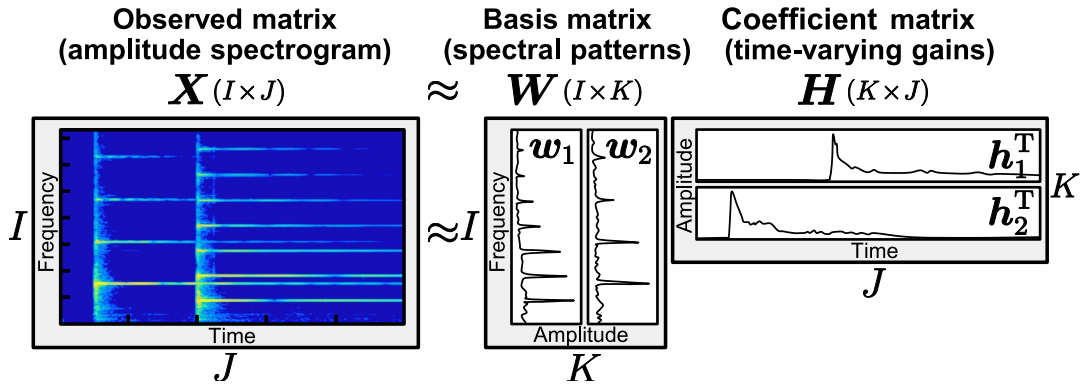
$$\mathbf{Z} = \text{STFT}_{\omega}(\mathbf{z}) \in \mathbb{C}^{I \times J} \quad (6.3)$$

ここで、 $\omega = [\omega(1), \omega(2), \dots, \omega(Q)]^T \in \mathbb{R}^Q$ は STFT で用いる窓関数である。スペクトログラム \mathbf{Z} の (i, j) 番目の要素は次式で表される。

$$z_{ij} = \sum_q \omega(q) z^{(j)}(q) \exp \left\{ \frac{-i2\pi(q-1)(i-1)}{F} \right\} \quad (6.4)$$

ここで F は $\lfloor \frac{F}{2} \rfloor + 1 = I$ を満たす整数 ($\lfloor \cdot \rfloor$ は床関数)、 $i = 1, 2, \dots, I$ は周波数ビンのインデクス、 i は虚数単位を示している。このように、時間領域の信号に対して一定幅の短時間ごとに分析窓関数を乗じて離散フーリエ変換を行うことで、横軸が時間、縦軸が周波数のスペクトログラムと呼ばれる複素行列 \mathbf{Z} で表すことができる。

音響信号に NMF を適用する場合、時間信号に STFT を適用して得られるスペクトログラム（複素行列）の振幅値（振幅スペクトログラム）やパワー値（パワースペクトログラム）を非負値観測行列 $\mathbf{X} \in \mathbb{R}_{\geq 0}^{I \times J}$ として扱うことが一般的である。振幅スペクトログラム \mathbf{X} を NMF で分解した様子が Fig. 6.2 である。振幅スペクトログラム \mathbf{X} は NMF を用いることで基底行

Fig. 6.2. NMF for audio signals, where $K = 2$.

列 \mathbf{W} と係数行列 \mathbf{H} の行列積で近似されている。Fig. 6.2 において、基底行列 \mathbf{W} は2つの基底ベクトルから成る ($K = 2$ に設定されている)。即ち、観測行列 \mathbf{X} のスペクトログラムを2本のスペクトルパターン (基底ベクトル) で近似的に表現している。同様に、係数行列 \mathbf{H} は、 \mathbf{W} 中の2本の基底ベクトルがどの時間にどの程度の強さで生起するかを表す時間的強度変化を行ベクトルとして含んでいる。

このように、音響信号であっても、STFTによってスペクトログラムを得ることでNMFを適用することができる。NMFで抽出される非負構成要素は解釈が容易であり、音響信号中のスペクトルパターンを教師なし学習できるため、音楽信号解析 [18] や音源分離 [7, 10] 等に頻繁に適用される。本実験では、単に欠損のある音響信号をNMFに基づく行列補完の観測行列として適用するため、STFTによってパワースペクトログラムに変換する。

6.2.2 時間周波数マスク

次に、時間周波数マスクについて説明する。時間周波数マスクとは、観測信号のスペクトログラムのある時間周波数要素に対して、目的の分離信号の成分が存在しているかどうかを表す二次元行列である。ソフトマスクであれば、時間周波数マスクは0から1までの値で構成され、バイナリマスク行列であれば、0又は1の値で構成される。この概要をFig. 6.3に示す。図のような赤、緑、青の音源から成る混合信号から赤の音源成分のみを取り出したい状況を仮定する。この時、赤の音源の部分を1それ以外を0とするようなバイナリの時間周波数マスクを構成し、混合信号のスペクトログラムと時間周波数マスクを要素毎に乗算することで、赤の音源とその他の音源を分離することが可能である。

このような時間周波数マスクはそのまま音源分離に適用され、いかに目的音源のみを通すマスクを作成できるかが検討されている (例えば, [19, 20, 21] 等)。適切な時間周波数マスクを生成できれば、簡単に音源分離可能な一方で、特にバイナリ時間周波数マスクによって分離された音源は、スペクトログラムの中に0を含むことになる。即ち、元の目的音源に対し欠損した状態であり、これは、人間の聴覚には歪みとして現れる。

本実験では、バイナリ時間周波数マスクを適用して音源分離された音響信号に対し、歪みを

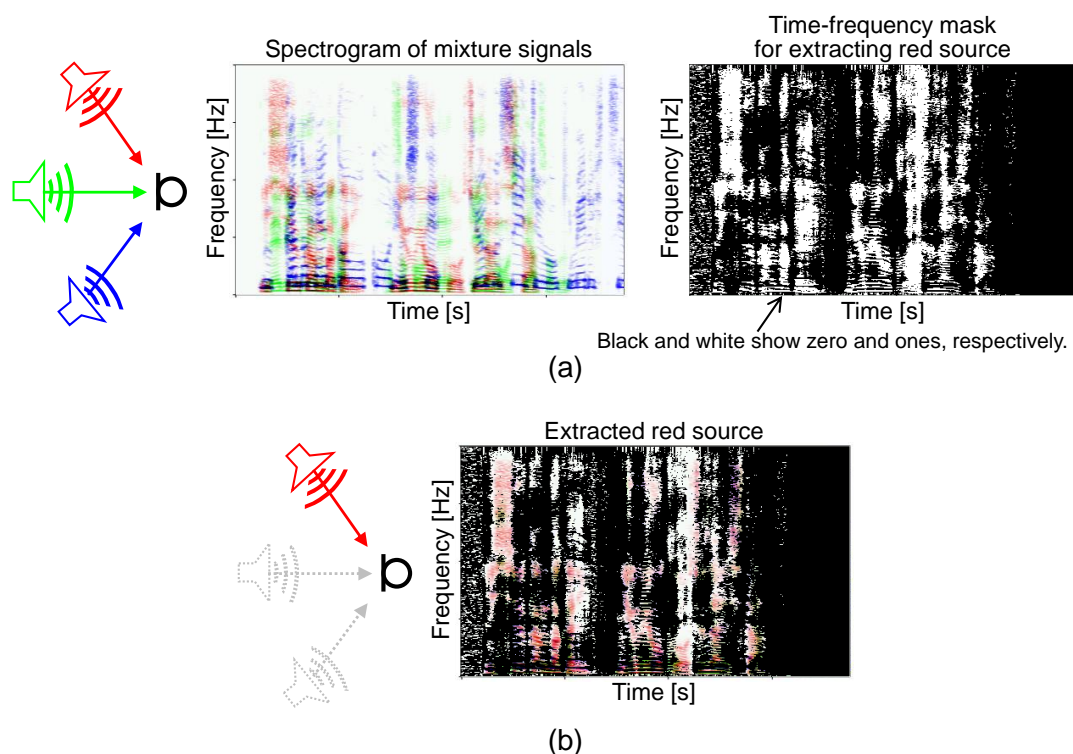


Fig. 6.3. Overview of time-frequency mask: (a) mixture signal and time-frequency binary mask for red source and (b) extracted red source by time-frequency binary masking.

軽減することを目的として提案手法で欠損補完を施すことを想定する。バイナリ時間周波数マスクは欠損のある音源を作成するための、完全に分離された（混合前の）音源信号のパワースペクトログラムから生成する。これは即ち、前段のバイナリ時間周波数マスクに基づく音源分離アルゴリズムが理想的な性能を達成した場合において、提案手法による欠損補完が分離音の品質をどの程度改善させることができるか、ということと比較・検証することに相当する。

6.3 実験方法

本実験では、音源に対しバイナリ時間周波数マスクをかけることで欠損を生成し、これを観測行列とした補完実験を行う。音源には SiSEC2011 [22] データセットから楽曲 bearlin-roads のアコースティックギター音源 (Gt.) 及びピアノ音源 (Pf.) の2つの音源を用いる。Table 6.1 に使用音源を示す。SiSEC2011 はプロ奏者の音楽信号が収録されたデータセットであり、bearlin-roads には音源分離への活用を想定してトレーニングセットとテストセットの2つのデータセットが用意されている。本実験において学習は不要のため、トレーニングセットのみを用いる。バイナリ時間周波数マスクは、各音源のパワースペクトログラムの各要素の値から生成する。

今、アコースティックギター音源のパワースペクトログラム $\mathbf{Z}^{(\text{Gt})} \in \mathbb{R}_{\geq 0}^{I \times J}$ 及びピアノ音源

Table 6.1. Source signals

| Dataset name | Instrument type |
|---|-----------------|
| dev1__bearlin-roads__snip_85_99__acoustic_guit_main.wav | Acoustic guitar |
| dev1__bearlin-roads__snip_85_99__piano.wav | Piano |

Table 6.2. STFT conditions

| | |
|-------------------------|-----------------|
| Audio length | 14.0 s |
| Sampling rate | 44.1 kHz |
| Window function in STFT | Blackman window |
| Window length in STFT | 92.9 ms |
| Shift length in STFT | 46.4 ms |

のパワースペクトログラム $\mathbf{Z}^{(\text{Pf})} \in \mathbb{R}_{\geq 0}^{I \times J}$ が与えられた時、ピアノ音源のバイナリ時間周波数マスク $\mathbf{M}^{(\text{Gt})} \in \{0, 1\}^{I \times J}$ 及びギター音源のバイナリ時間周波数マスク $\mathbf{X}^{(\text{Pf})} \in \{1, 0\}^{I \times J}$ は次式によって得る.

$$m_{ij}^{(\text{Gt})} = \begin{cases} 1 & \text{if } z_{ij}^{(\text{Gt})} \geq z_{ij}^{(\text{Pf})} \\ 0 & \text{otherwise} \end{cases} \quad \forall i, j \quad (6.5)$$

$$m_{ij}^{(\text{Pf})} = \begin{cases} 1 & \text{if } m_{ij}^{(\text{Gt})} = 0 \\ 0 & \text{otherwise} \end{cases} \quad \forall i, j \quad (6.6)$$

ここで, $z_{ij}^{(\text{Gt})}$, $z_{ij}^{(\text{Pf})}$, $m_{ij}^{(\text{Gt})}$, 及び $m_{ij}^{(\text{Pf})}$ はそれぞれ $\mathbf{Z}^{(\text{Gt})}$, $\mathbf{Z}^{(\text{Pf})}$, $\mathbf{M}^{(\text{Gt})}$, 及び $\mathbf{M}^{(\text{Pf})}$ の要素である. 即ち, 式 (6.5) 及び式 (6.6) は, 2つの観測信号のパワースペクトログラムの各要素について大小比較し, 大きい要素を持つ観測信号のバイナリ時間周波数マスクを1, もう一方を0として生成する. 完全に分離された信号のスペクトログラム $\mathbf{Z}^{(\text{Gt})}$ 及び $\mathbf{Z}^{(\text{Pf})}$ を利用しているので, 実際の音源分離では作ることができない理想的なバイナリ時間周波数マスクである.

欠損を含む観測行列は, 観測信号のパワースペクトログラムにバイナリ時間周波数を掛け合わせるにより生成する. アコースティックギター音源の観測行列を $\mathbf{X}^{(\text{Gt})} \in \mathbb{R}_{\geq 0}^{I \times J}$, ピアノ音源の観測行列を $\mathbf{X}^{(\text{Pf})} \in \mathbb{R}_{\geq 0}^{I \times J}$ とすると, 次式より得られる.

$$\mathbf{X}^{(\text{Gt})} = \mathbf{M}^{(\text{Gt})} \odot \mathbf{Z}^{(\text{Gt})} \quad (6.7)$$

$$\mathbf{X}^{(\text{Pf})} = \mathbf{M}^{(\text{Pf})} \odot \mathbf{Z}^{(\text{Pf})} \quad (6.8)$$

以上より, 観測行列 $\mathbf{X}^{(\text{Gt})}$ 及び $\mathbf{X}^{(\text{Pf})}$ を得る. Fig. 6.4 に使用音源の時間信号とパワースペクトログラムを, Fig. 6.5 にバイナリ時間周波数マスクと観測行列をそれぞれ示す. また, この時の STFT の条件を Table 6.2 に示す. 本実験では, 観測行列 $\mathbf{X}^{(\text{Gt})}$ 及び $\mathbf{X}^{(\text{Pf})}$

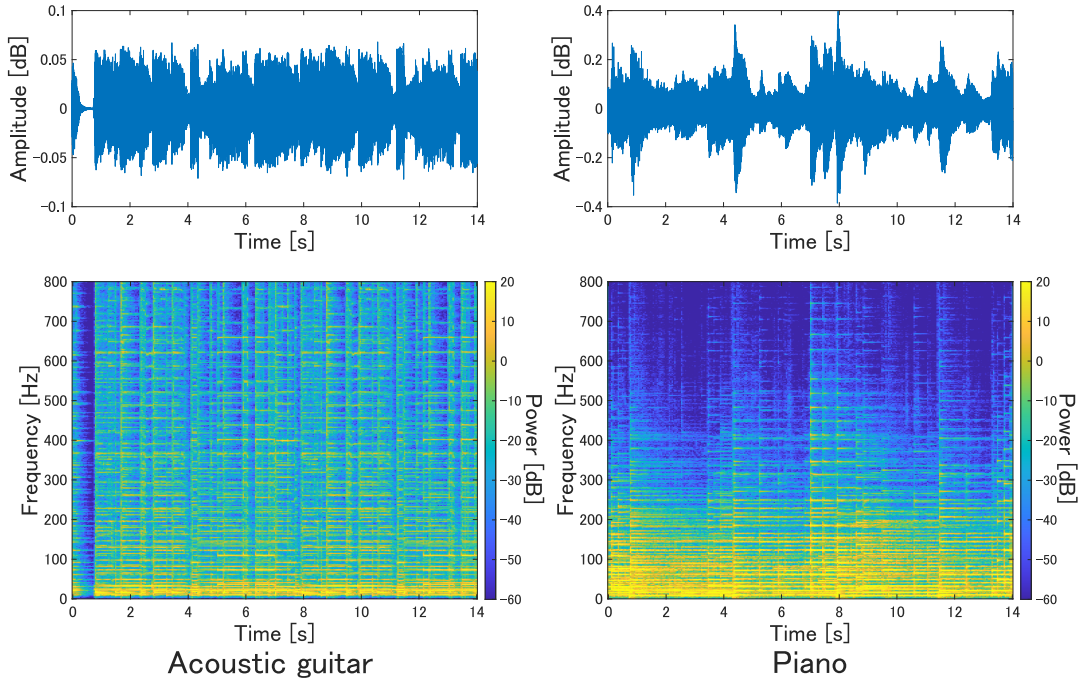


Fig. 6.4. Source signals and their power spectrograms $Z^{(\text{Gt})}$ and $Z^{(\text{Pf})}$ used in experiments.

を KLNMF 及び 2 つの正則化付き NMF でそれぞれ補完を行い、性能を評価する。バイナリ時間周波数マスクによって得られる分離音源は波長の一部に 0 を含むことで歪みを有する。行列補完により適切に欠損値が埋められた場合、このような歪みは減少すると考えられる。また、補完時に目的音源のパターンを強く抽出することができれば、非目的音源の干渉を抑えることができる。従って、音源分離分野においてよく用いられる指標である信号対歪み比 (source-to-distortion ratio: SDR), 信号対干渉比 (source-to-interference ratio: SIR), 及び信号群対歪み比 (sources-to-artifacts ratio: SAR) [23] の改善量によって性能を評価する。これらは分離信号の品質を測る指標であり, SIR は値が高いほど非目的音源の干渉が小さいことを, SAR は値が高いほど信号に残留する人工歪みが小さいことをそれぞれ示す。また, SDR は SIR と SAR の両指標を加味した総合的な音源分離性能を評価する指標である。

今, 目的音源 $s(p)$ における推定信号 $\hat{s}(p)$ は以下の式に示すように分解することができる。

$$\hat{s}(p) = s_{\text{target}}(p) + s_{\text{interf}}(p) + e_{\text{artif}}(p) \quad (6.9)$$

ここで, $s_{\text{target}}(p)$, $s_{\text{interf}}(p)$, 及び $e_{\text{artif}}(p)$ はそれぞれ推定信号 $\hat{s}(p)$ 中の目的音源成分, 残留した非目的音源成分, 及び音源分離によって生じた人工的な歪み成分を表す。SDR は次式で表される。

$$\text{SAR} = 10 \log_{10} \frac{\sum_p |s_{\text{target}}(p)|^2}{\sum_p |s_{\text{interf}}(p) + e_{\text{artif}}(p)|^2} \text{ [dB]} \quad (6.10)$$

従って, 高い SDR 値を達成するには $s_{\text{interf}}(p)$ 及び $e_{\text{artif}}(p)$ が少なく, $s_{\text{target}}(p)$ が高精度に

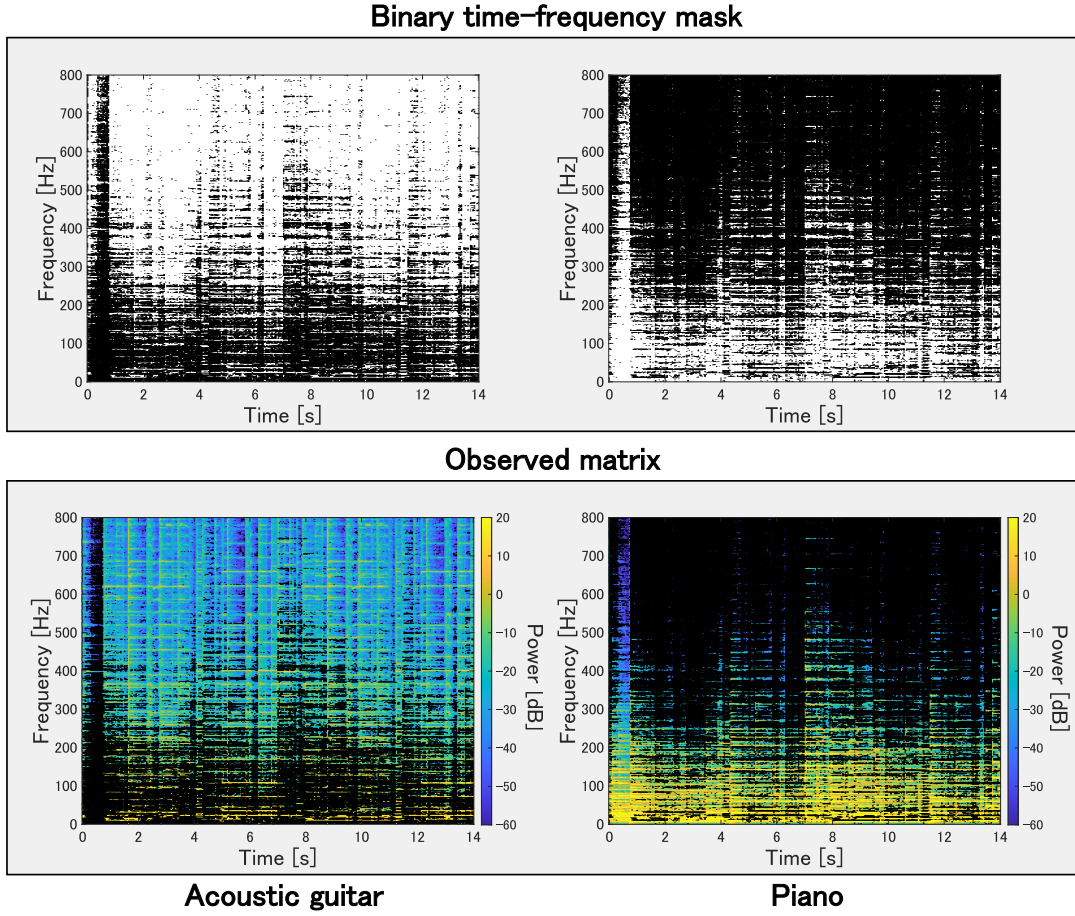


Fig. 6.5. Binary time-frequency masks and separated power spectrograms, which are used as observed matrix in matrix completion experiment. Black elements in observed matrix show missing values.

推定されている必要がある。SIR は、 $s_{\text{target}}(p)$ 及び $s_{\text{interf}}(p)$ を用いて次式で表される。

$$\text{SIR} = 10 \log_{10} \frac{\sum_p |s_{\text{target}}(p)|^2}{\sum_p |s_{\text{interf}}(p)|^2} \text{ [dB]} \quad (6.11)$$

従って、高い SIR 値を達成するには $s_{\text{interf}}(p)$ をより小さくする必要がある。SAR は、 $s_{\text{target}}(p)$ 、 $s_{\text{interf}}(p)$ 、及び $e_{\text{artif}}(p)$ を用いて次式で表される。

$$\text{SAR} = 10 \log_{10} \frac{\sum_p |s_{\text{target}}(p) + s_{\text{interf}}(p)|^2}{\sum_l |e_{\text{artif}}(p)|^2} \text{ [dB]} \quad (6.12)$$

従って、高い SAR 値を達成するには目的音源及び非目的音源によって構成される分離音源の大きさに対し、歪み成分の大きさをより小さくする必要がある。以上より、これら3つの指標を確認することで補完性能について評価することができる。

実験手順と各処理で使用するパラメータを Fig. 6.6 にまとめる。本実験では、使用音源をパワースペクトログラムに変換した後、バイナリ時間周波数マスクによって観測行列 $\mathbf{X}^{(\text{Gt})}$ 及

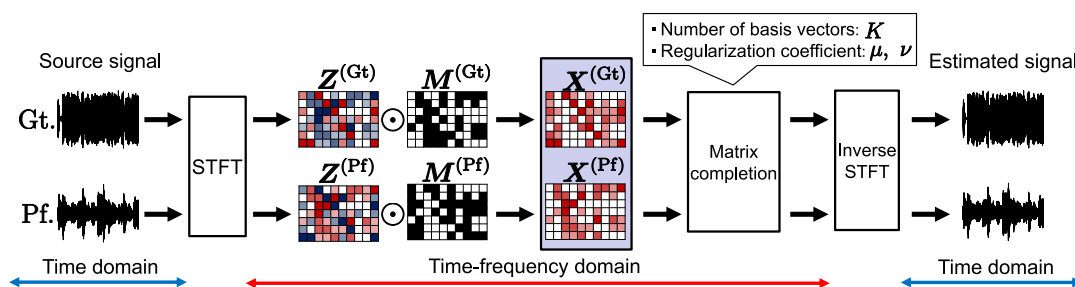


Fig. 6.6. Experimental conditions changed in evaluation.

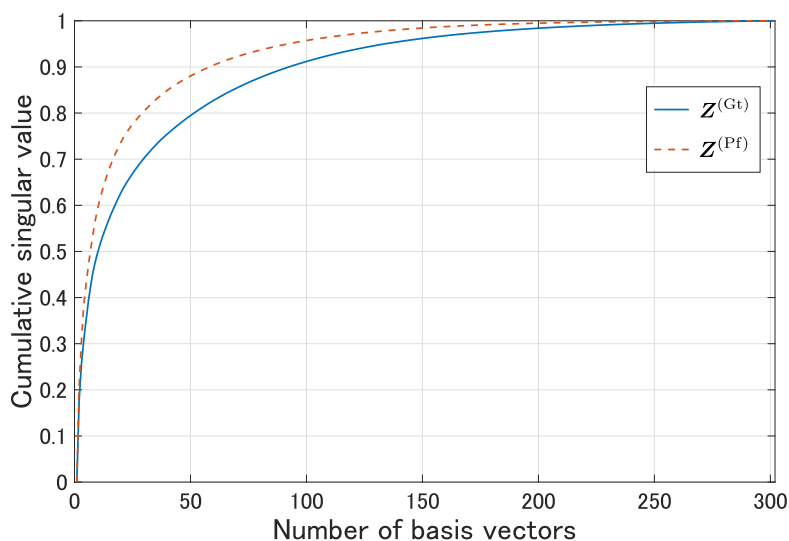


Fig. 6.7. Cumulative singular values of power spectrogram of acoustic guitar and piano sources.

び $\mathbf{X}^{(Pf)}$ を生成する。観測行列に対し各設定値の NMF を適用し行列補完を行い、時間領域に変換した補完信号と観測信号を比較し、補完精度を評価する。

NMF の基底ベクトル数 K の設定値候補を算出するために、 $\mathbf{Z}^{(Gt)}$ 及び $\mathbf{Z}^{(Pf)}$ がどの程度のランクで良く近似できるかを確認する。具体的には、累積特異値が 80% 程度に到達するのに必要な特異ベクトルの個数を確認し、その周辺の値を NMF の基底ベクトル数 K の設定値候補とする。実際に $\mathbf{Z}^{(Gt)}$ 及び $\mathbf{Z}^{(Pf)}$ の特異値分解により累積特異値を求めた結果を Fig. 6.7 に示す。これより、いずれの音源においてもランク 50 程度で累積特異値の割合が全体の 80% を超えることがわかる。また、観測行列は 2049 行 302 列の行列であり、NMF は低ランク近似であることから、基底ベクトル数 K の設定値は 4 から 100 までとした。正則化付き NMF の重み係数 μ 及び ν は 10^{-3} から 10^5 まで対数スケールで 40 点推移させた。各実験条件において NMF の基底行列及び係数行列の初期化に用いる乱数シードを 10 種類用意し、その平均を評価値とした。

6.4 実験結果

Figs. 6.8–6.10 に、基底ベクトル数 K 毎の各 NMF による補完信号の SDR 改善量 (SDR_i), SIR 改善量 (SIR_i), 及び SAR 改善量 (SAR_i) の推移をそれぞれ示す。これらの改善量は、バイナリ時間周波数マスクによって欠損したパワースペクトログラムをそのまま（補完せずに）時間領域へ戻した信号の SDR 値, SIR 値及び SAR 値との差であり、値が高いほど行列補完処理によって品質が改善されたことを示す。

Fig. 6.8 より、両音源において CSRNMF 及び KLNMF は SDR が無補完信号に対し低下しており、分離信号として劣化したことがわかる。KLNMF と CSRNMF を比較すると、基底ベクトル数 K 値によって前後するものの同程度の値をとっているといえる。一方で、 L_1 -RNMF は SDR_i が常にほぼ 0 となっており、僅かな改善もみられる。また、KLNMF や CSRNMF と異なり、基底ベクトル数 K の変化に対して一定の値をとっている。

Fig. 6.9 より、アコースティックギター音源において SIR は、KLNMF 及び CSRNMF 共に無補完信号に対して劣化はあるものの、KLNMF が平均的に 25 dB 程度低下したのに対し CSRNMF は 8 dB 程度の低下に抑えられている。一方で、ピアノ音源については CSRNMF が無補完信号に対して広い基底ベクトル数 K 値において改善を示しており、最大で 10 dB 程度まで改善した。また、 L_1 -RNMF は SDR 同様ほぼ 0 の値を推移しており、ピアノ音源において極僅かな改善量が見られる。

Fig. 6.10 より、アコースティックギター音源において SAR は、KLNMF 及び CSRNMF 共に改善が見られず、基底ベクトル数 K 値が大きくなるにつれていずれの両手法とも減少していることがわかる。 L_1 -RNMF は基底ベクトル数 K が低い時に僅かな改善を示し、基底ベクトル数 K が大きくなるにつれ、KLNMF 及び CSRNMF 同様に SAR が低下していることがわかる。一方で、ピアノ音源においては、KLNMF 及び CSRNMF が低下したのに対し L_1 -RNMF は広い基底ベクトル数 K 値の範囲の多くの点で 1 dB 程度の改善を示した。

以上より、 L_1 -RNMF は SDR, SIR 及び SAR において、音源によらず 0 に近い値をとっており、一部において僅かな改善が見られた。 L_1 -RNMF において特徴的なのは、ピアノ音源における SAR が 1 dB 程度上昇したことである、ピアノ音源においてのみこのような結果を示したことから、本実験におけるアコースティックギター音源とピアノ音源の特徴の違いが、 L_1 -RNMF の性能に影響したと考えられる。Fig. 6.4 のスペクトログラムにおいて特に強い振幅を持つことを示す黄色い要素が、Fig. 6.5 においてアコースティックギター音源は多く欠損したのに対し、ピアノ音源は保持していることがわかる。これより、ピアノ音源のスペクトログラムは重要な特徴を残した行列であるといえる。このような行列のため、欠損値をピアノ音源に矛盾なく補完することができ、結果として SAR が改善したと考えられる。CSRNMF は SDR 及び SAR が、多くの条件において KLNMF と同等か低い結果となった。一方で、SIR においては KLNMF に対し明らかに性能が向上しており、ピアノ音源においては無補完信号に対し 10 dB 程度の改善を示した。このような結果の理由として、 L_1 -RNMF 同様ピアノ音

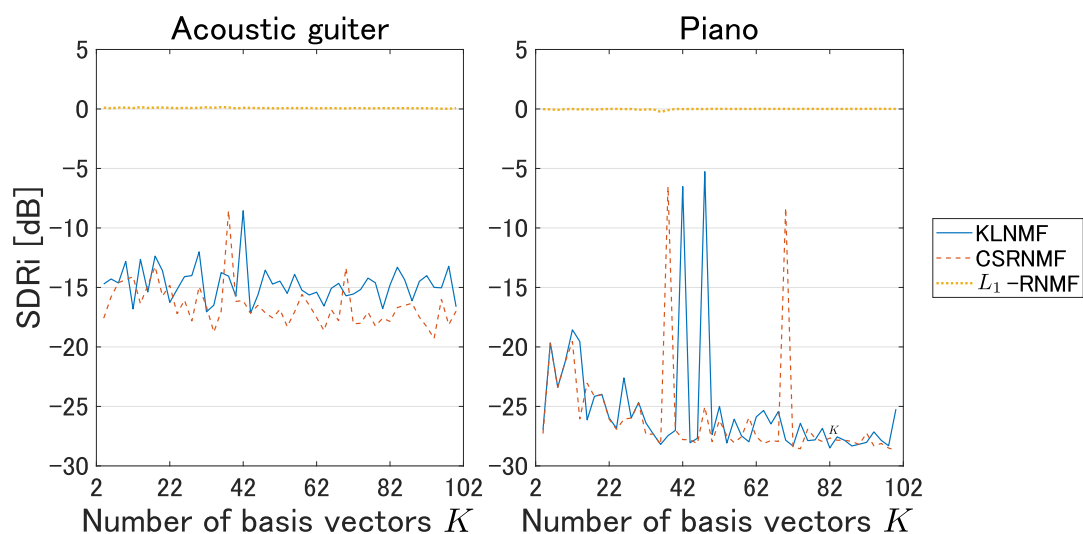


Fig. 6.8. Average SDRi obtained by matrix completion based on KLNMF, CSRNMF, and L_1 -RNMF when $\mu = 0.001$, and $\nu = 42.813$.

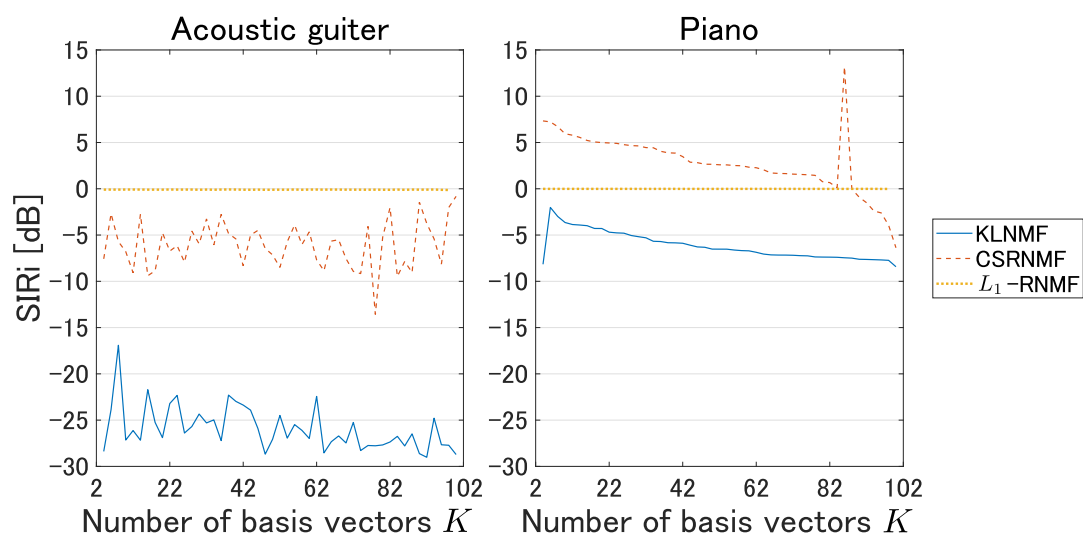


Fig. 6.9. Average SIRi obtained by matrix completion based on KLNMF, CSRNMF, and L_1 -RNMF when $\mu = 61585$, and $\nu = 100000$.

源のスペクトログラムが重要な特徴を保持しているためであると考えられる。無補完信号における干渉信号とはバイナリ時間周波数マスクによる0の信号である。CSRNMFによってこれを補完するにあたり、より元の信号の特徴を含んだピアノ音源において性能が向上したと考えられる。また、SDRが全体的に低下している要因として、バイナリ時間周波数マスクによる音源分離が理想的状態が大きい。欠損値を補完する値が元音源と異なるとき、その誤差が人工的歪みや非目的音源の干渉となってしまう、音源分離の総合的指標となるSDRは改善しにくいと言える。

最後に、各手法について、SDR、SIR及びSARから総合的に評価する。Fig. 6.11に、基底

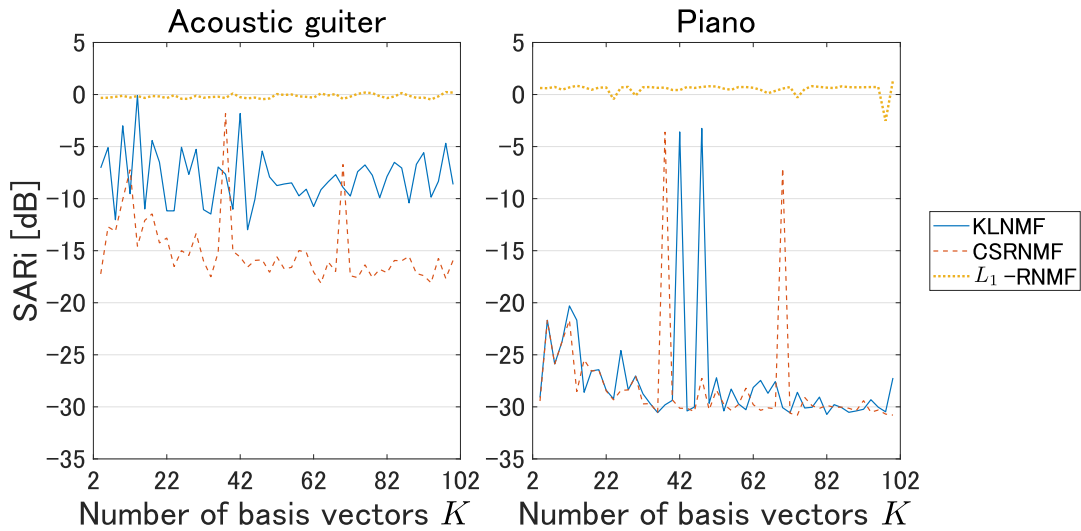


Fig. 6.10. Average SARI obtained by matrix completion based on KLNMF, CSRNMF, and L_1 -RNMF when $\mu = 100000$, and $\nu = 3.793$.

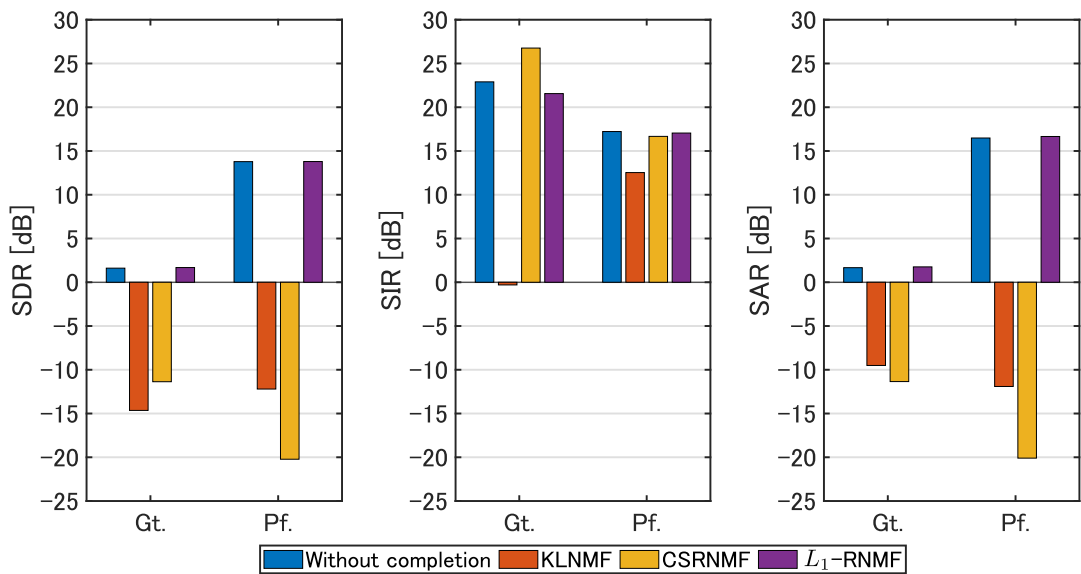


Fig. 6.11. Average SDR, SIR, and SAR values when $K = 30$, $\mu = 6.159$, and $\nu = 297.635$.

ベクトル数 K が 30 の時における無補完信号を含めた各手法の最大値を示す。また、その他の基底ベクトル数 K における各手法の最大値を付録 C に示す。Fig. 6.11 において、CSRNMF 及び L_1 -RNMF の重み係数はそれぞれ $\mu = 6.159$, $\nu = 297.635$ である。また、この時の各値を Table 6.3 に示す。これより、KLNMF は全ての指標において無補完信号から低下する一方、CSRNMF は SIR が、 L_1 -RNMF は SDR 及び SAR の改善が見られ、補完手法として性能向上が確認できる。総合的な音源補完の性能としては、 L_1 -RNMF が最も安定していると言え、SIR を最も重視する場合には CSRNMF が有効であると考えられる。いずれの場合においても、正則化項による効果が確認でき、補完性能の向上が得られた。

Table 6.3. Average SDR, SAR, and SIR values when $K = 30$, $\mu = 6.159$, and $\nu = 297.635$.

| Method | Acoustic guitar | | | Piano | | |
|--------------------|-----------------|---------------|--------------|---------------|---------------|---------------|
| | SDR [dB] | SIR [dB] | SAR [dB] | SDR [dB] | SIR [dB] | SAR [dB] |
| Without completion | 1.612 | 22.900 | 1.667 | 13.784 | 17.224 | 16.485 |
| KLNMF | -14.645 | -0.305 | -9.508 | -12.202 | 12.534 | -11.901 |
| CSRNMF | -11.366 | 26.762 | -11.349 | -20.223 | 16.670 | -20.098 |
| L_1 -RNMF | 1.682 | 21.559 | 1.757 | 13.796 | 17.052 | 16.658 |

6.5 本章のまとめ

本章では、実際の音源を用いた補完実験を行い、それぞれの手法における性能を評価した。SCRNMF による行列補完は SDR が改善しなかったものの、SIR は大きく改善する場合が確認できた。 L_1 -RNMF による行列補完は SDR, SIR 及び SAR において僅かな改善が見られ、KLNMF 及び CSRNMF に対し最も高い補完性能が得られた。これより、正則化項が働くことで補完性能に一定の改善を得られることが確認できた。

第7章

結言

本論文では、あらゆるデータにおいて発生する欠損値を低コストかつ高精度に補完するため、2つの正則化付き NMF を検討した。コサイン類似度正則化は NMF に付与することにより、基底行列に含まれる各基底ベクトルを相互に直交させることが確認できた。 L_1 ノルム正則化は NMF に付与することにより、係数行列の要素がスパースになるよう促すことが確認できた。人工欠損行列を用いた実験において、CSRNMF が特にノイズの大きい条件について従来の KLNMF を用いた手法より高精度な補完性能を発揮した。また、 L_1 -RNMF は多くの条件において、KLNMF に対し補完性能が向上しており、基底ベクトル数 K のモデル誤差及びノイズに対する頑健性が確認できた。データセットの音源を用いた実験においては、 L_1 -RNMF が安定した補完性能を見せたものの劇的な音源の改善は見られなかった。また、CSRNMF は特定の条件において SIR のみ 10 dB 程度の大きな改善が確認できたが、これに伴い SDR 及び SAR は低下した。

本論文では、正則化を NMF に適用することで行列補完性能を向上させることができることを確認した。一方で、正則化項の影響度を定める重み係数によってその性能は大きく変化し、各実験において条件ごとに最適な値を示す重み係数は異なっていた。現状、最適な値は網羅的に重み係数値を変更して調べる必要があり、多大な時間を要する。この問題を解決するため、効率的な最適な重み係数のチューニング方法の検討が必要となる。

謝辞

本論文は、香川高等専門学校電気情報工学科北村研究室にて行われた研究に基づくものです。

初めに、本研究を進めるにあたり、ご多忙のところ多大なご指導を賜った、北村大地講師に深く感謝いたします。北村大地講師にはこの1年間で、本研究分野に関する基礎的な知識をはじめ、資料作成、実験結果の考察など細部に至るまで熱心にご指導下さいました。また、自分の進路についても非常に丁寧にサポート頂き、本当にありがとうございました。

本論文の副査を担当頂いた柿元健准教授には、論文の構成や記述に関して大変有益な助言を頂き、大変お世話になりました。ここに厚く御礼申し上げます。

北村研究室の先輩である川口翔也氏、溝渕悠朔氏、村田佳斗氏、綾野翔馬氏にはソフトウェアの使い方やプログラミングの記法、実験方法の助言など、研究に必要な多くのことを支援頂きました。研究について高い技術と熱意を持ちながら、研究室内をユーモアあふれる空間にして下さったことで、臆することなく様々な学びを得ることができました。また、綾野翔馬氏には本論文執筆において様々な助言を頂き、自身の研究についてより理解を深めることができました。心より感謝致します。

北村研究室の同期である加藤大輝氏、鈴木慶氏、松本愛花氏には、日頃のディスカッションをはじめ、研究室生活を楽しいものにして頂きました。特に、加藤大輝氏とは行動を共にすることも多く、その中で多くのインスピレーションを与えていただきました、モチベーションの高い同期に囲まれたこの環境だからこと、研究を進めることができました。

最後に、実家を離れ香川高専へ進学することを快く後押ししてくれた家族には感謝の念に堪えません。本科5年間においても、安心して寮生活を送れるよう暖かく見守って下さり、本当にありがとうございました。重ねて、本研究に協力して下さいました皆様には心より感謝申し上げます。

参考文献

- [1] L. Ren, T. Wang, A. S. Seklouli, H. Zhang, and A. Bouras, “A review on missing values for main challenges and methods,” *Information Systems*, vol. 119, 2023.
- [2] T. Emmanuel, T. Maupong, D. Mpoeleng, T. Semong, B. Mphago, and O. Tabona, “A survey on missing data in machine learning,” *Journal of Big Data*, vol. 8, 2021.
- [3] D. D. Lee and H. S. Seung, “Learning the parts of objects by non-negative matrix factorization,” *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [4] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *Proc. Neural Info. Process. Syst.*, pp. 556–562, 2000.
- [5] T. Kujiraoka, S. Fumiaki, and I. Syohei, “Extraction of customer satisfaction topics regarding product delivery using non-negative matrix factorization,” in *Proc. IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 225–229, 2017.
- [6] D. Kitamura, H. Saruwatari, H. Kameoka, Y. Takahashi, K. Kondo, and S. Nakamura, “Multichannel signal separation combining directional clustering and nonnegative matrix factorization with spectrogram restoration,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 4, pp. 654–669, 2015.
- [7] 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.
- [8] D. R. Hunter and K. Lange, “Quantile regression via an MM algorithm,” *J. Comput. Graph. Stat.*, vol. 9, no. 1, pp. 60–77, 2000.
- [9] P. Smaragdis, B. Raj, and M. Shashanka, “Supervised and semi-supervised separation of sounds from single-channel mixtures,” in *Proc. LVA/ICA*, pp. 414–421, 2007.
- [10] D. Kitamura, H. Saruwatari, K. Yagi, K. Shikano, Y. Takahashi, and K. Kondo, “Music signal separation based on supervised nonnegative matrix factorization with orthogonality and maximum-divergence penalties,” *IEICE Trans. Fundamentals*, vol. E97–A, no. 5, pp. 1113–1118, 2014.
- [11] Y. Iwase and D. Kitamura, “Supervised audio source separation based on nonnegative matrix factorization with cosine similarity penalty,” *IEICE Trans. Fundamentals*, vol. E105–A, no. 6, pp. 906–913, 2022.

- [12] P. O. Hoyer, “Non-negative matrix factorization with sparseness constraints,” *The Journal of Machine Learning Research* 5, pp. 1457–1469, 2004.
- [13] W. Liu, N. Zheng, and X. Lu, “Non-negative matrix factorization for visual coding,” in *Proc. ICASSP*, pp. III–293, 2003.
- [14] J. Le Roux, F. Weninger and J. R. Hershey, “Sparse NMF - half-baked or well done?,” *Mitsubishi Electric Research Lab. Technical Report*, TR2015–023, 2015.
- [15] R. Tibshirani, “Regression shrinkage and selection via the Lasso,” *J. Royal Statist. Soc. Ser. B* 58, pp. 267–288, 1996.
- [16] M. A. Rasmussen, and R. Bro, “A tutorial on the Lasso approach to sparse modeling,” *Chemometrics and Intelligent Laboratory Systems*, vol. 119, pp. 21–31, 2012.
- [17] A. Marmin, J. Goulart, and C. Févotte, “Majorization-minimization for Sparse Non-negative Matrix Factorization with the β -divergence.” *IEEE Trans. Signal Process.*, pp. 1435–1447, 2023.
- [18] C. Févotte, N. Bertin, and J.-L. Durrieu, “Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis,” *Neural Comput.*, vol. 21, no. 3, pp. 793–830, 2009.
- [19] H. Sawada, S. Araki, and S. Makino, “Underdetermined convolutive blind source separation via frequency bin-wise clustering and permutation alignment,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 3, pp. 516–527, 2011.
- [20] Y. Luo and N. Mesgarani, “Conv-TasNet: Surpassing ideal time–frequency magnitude masking for speech separation,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 27, no. 8, pp. 1256–1266, 2019.
- [21] S. Oyabu, D. Kitamura, and K. Yatabe, “Linear multichannel blind source separation based on time-frequency mask obtained by harmonic/percussive sound separation,” in *Proc. ICASSP*, pp. 201–205, 2021.
- [22] S. Araki, F. Nesta, E. Vincent, Z. Koldovský, G. Nolte, A. Ziehe, and A. Benichoux, “The 2011 signal separation evaluation campaign (SiSEC2011) - audio source separation -,” in *Proc. LVA/ICA*, pp. 414–422, 2012.
- [23] E. Vincent, R. Gribonval, and C. Févotte, “Performance measurement in blind audio source separation,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 4, pp. 1462–1469, 2006.

付録 A

補助関数法で利用される不等式

A.1 Jensen の不等式

補題 A.1. (Jensen の不等式) $\alpha_i > 0$ を, $\sum_i \alpha_i = 1$ を満たす補助変数とする. 関数 $f(x)$ が凸関数であるとき, x_i ($i = 1, \dots, I$) に対して以下の不等式が成立する.

$$f\left(\sum_{i=1}^I \alpha_i x_i\right) \leq \sum_{i=1}^I \alpha_i f(x_i) \quad (\text{A.1})$$

$f(x)$ が狭義凸関数であるとき, 不等式中の等号が成立するための条件は $x_1 = \dots = x_i = \dots = x_I$ である.

付録 B

人工欠損行列の補完実験結果

本付録では、5章で行った人工欠損行列の補完実験の内、本文で示した条件以外の実験結果を掲載する。Figs. B.1–B.11 にノイズに一樣分布を用いた時の各 SNR 値における実験結果を示す。また、Figs. B.12–B.19 にノイズにレイリー分布を用いた時の各 SNR 値における実験結果を示す。

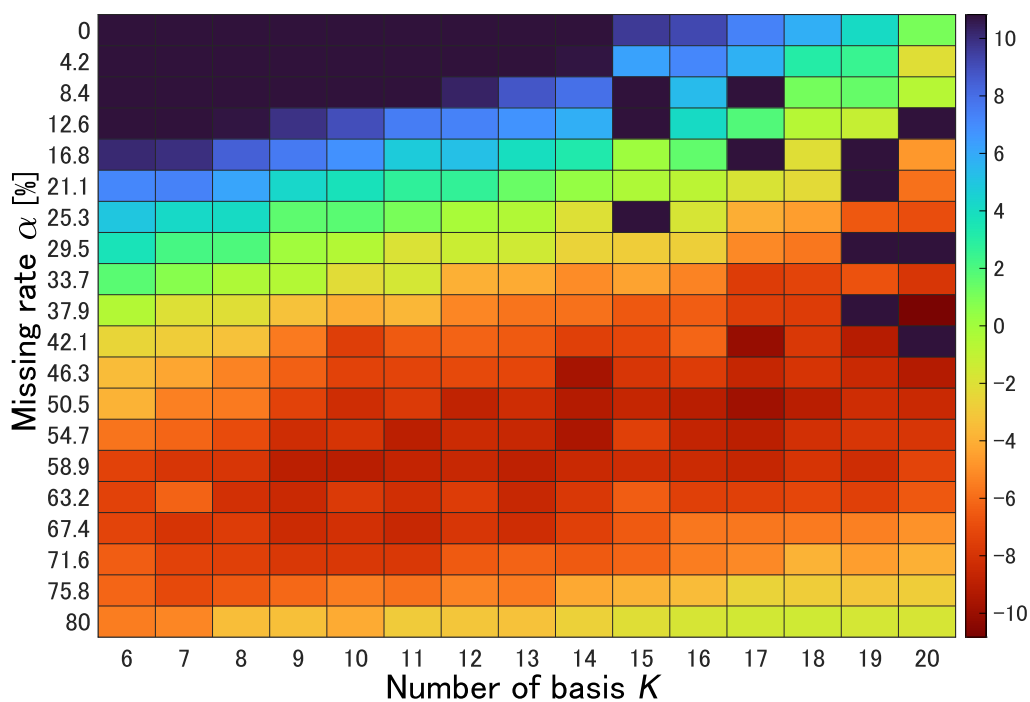


Fig. B.1. Improvement value obtained by CSRNMF: noise type is uniform dist., $r = 0$ dB, and $\mu = 1274.28$. The smaller the value and the redder the color, the better the performance.

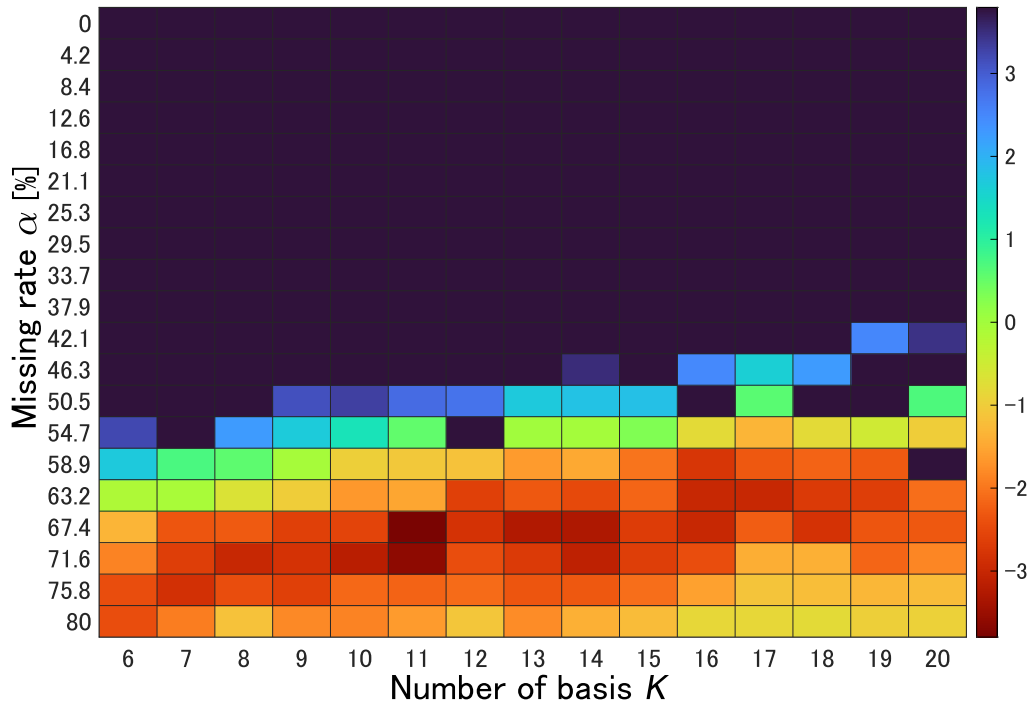


Fig. B.2. Improvement value obtained by CSRNMF: noise type is uniform dist., $r = 3.16$ dB, and $\mu = 5455.59$.

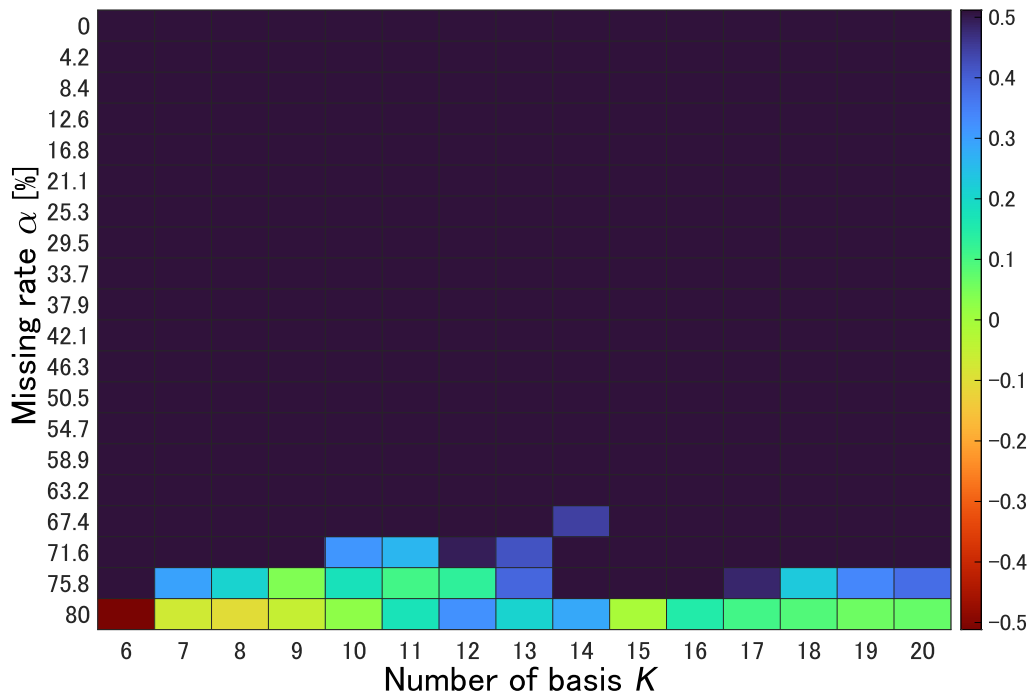


Fig. B.3. Improvement value obtained by CSRNMF: noise type is uniform dist., $r = 5.99$ dB, and $\mu = 6.16$.

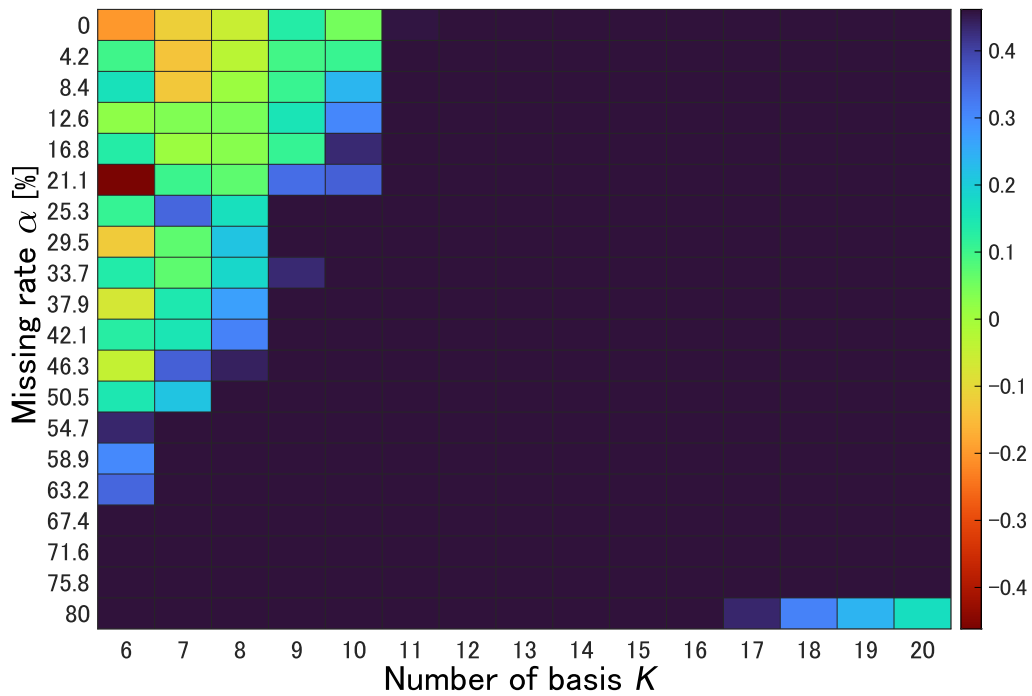


Fig. B.4. Improvement value obtained by CSRNMf: noise type is uniform dist., $r = 21.54$ dB, and $\mu = 0.88$.

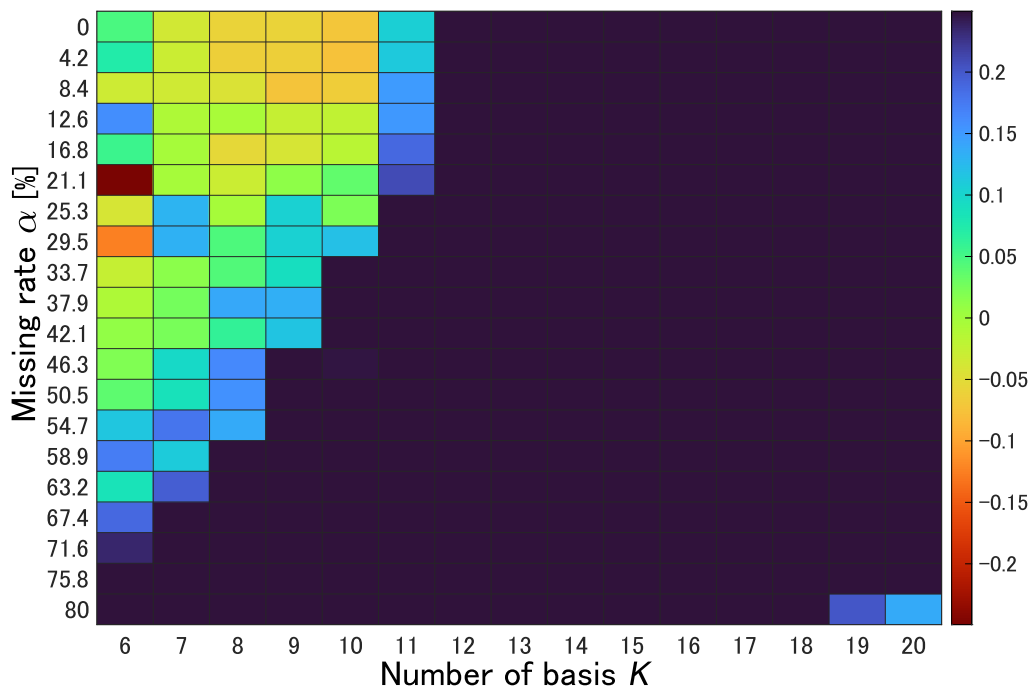


Fig. B.5. Improvement value obtained by CSRNMf: noise type is uniform dist., $r = 40.84$ dB, and $\mu = 0.55$.

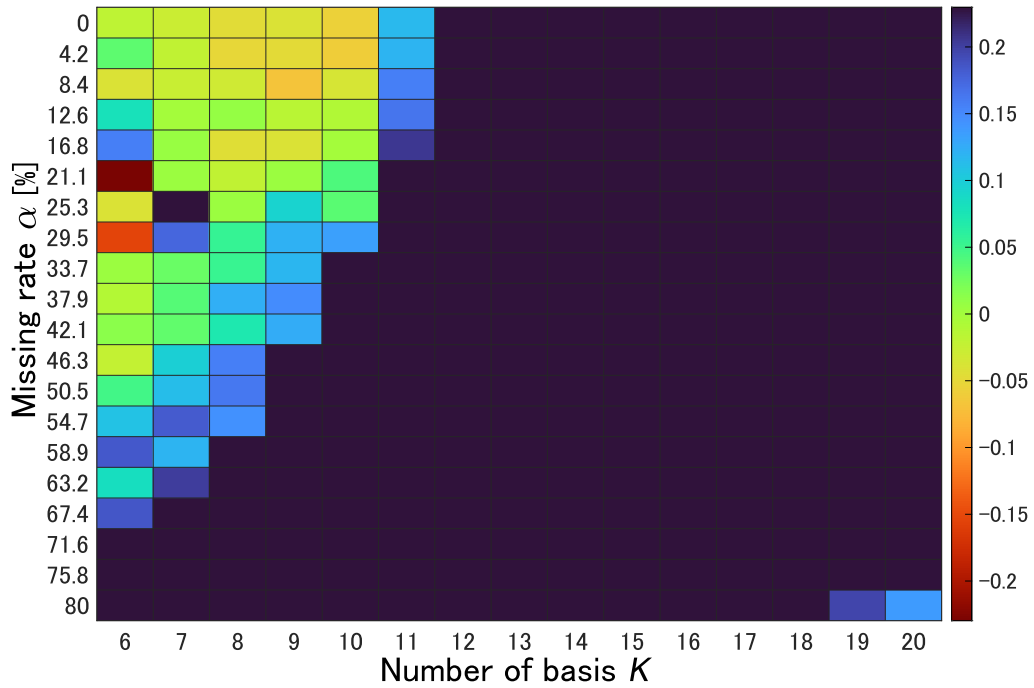


Fig. B.6. Improvement value obtained by CSRNMF: noise type is uniform dist., $r = 77.43$ dB, and $\mu = 0.55$.

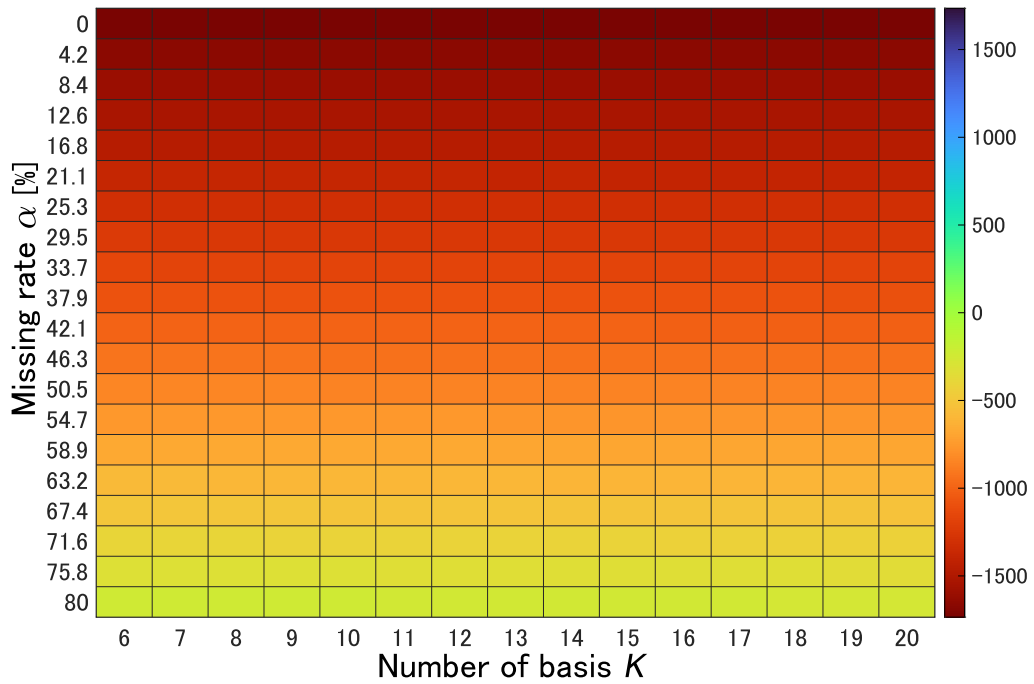


Fig. B.7. Improvement value obtained by L1RNMF: noise type is uniform dist., $r = 0$ dB, and $\nu = 1.44$.

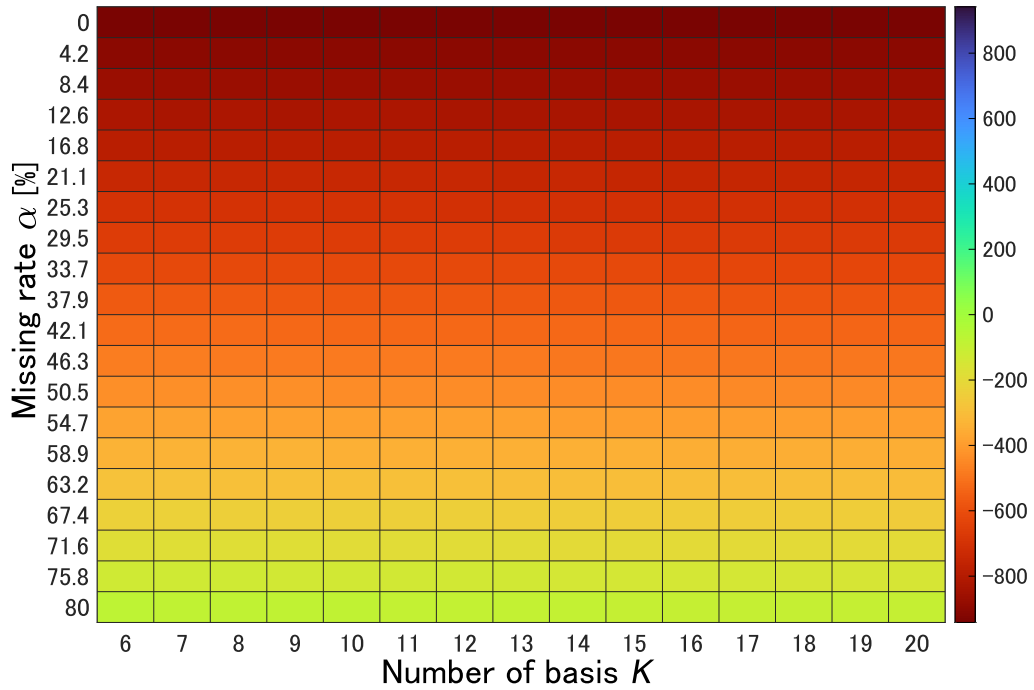


Fig. B.8. Improvement value obtained by L_1 RNMF: noise type is uniform dist., $r = 3.16$ dB, and $\nu = 0.89$.

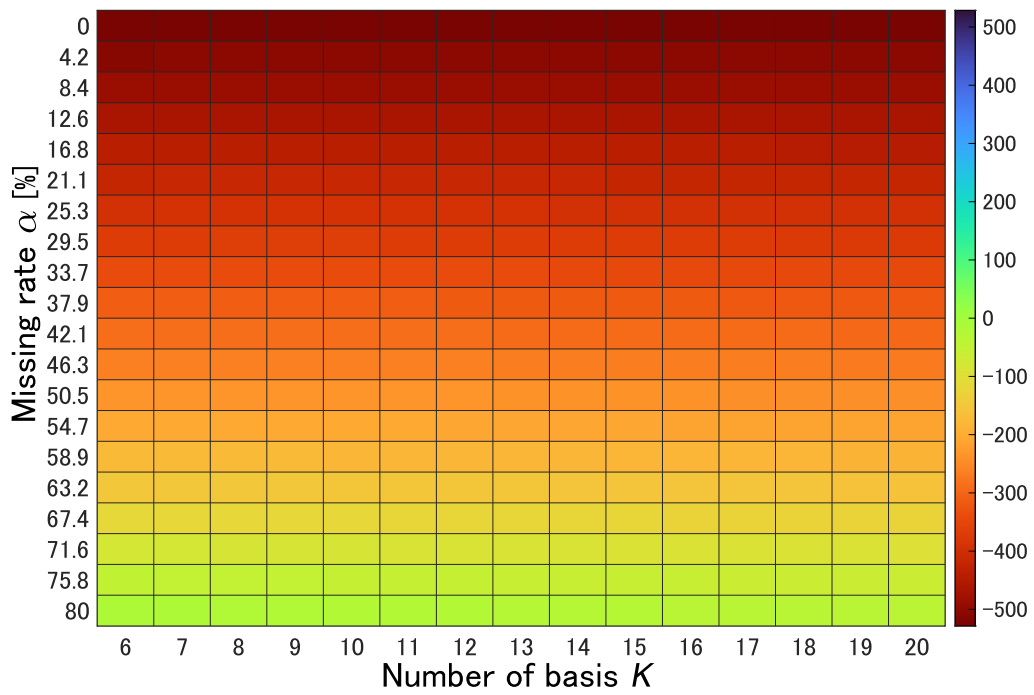


Fig. B.9. Improvement value obtained by L_1 -RNMF: noise type is uniform dist., $r = 5.99$ dB, and $\nu = 0.55$.

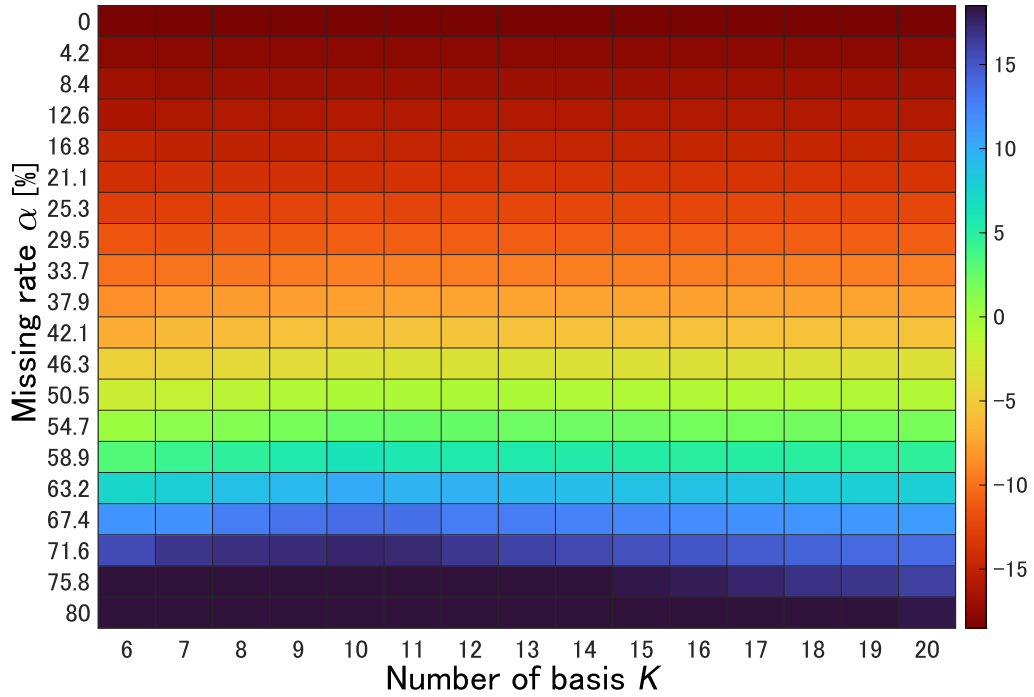


Fig. B.10. Improvement value obtained by L_1 -RNMF: noise type is uniform dist., $r = 21.54$ dB, and $\nu = 0.08$.

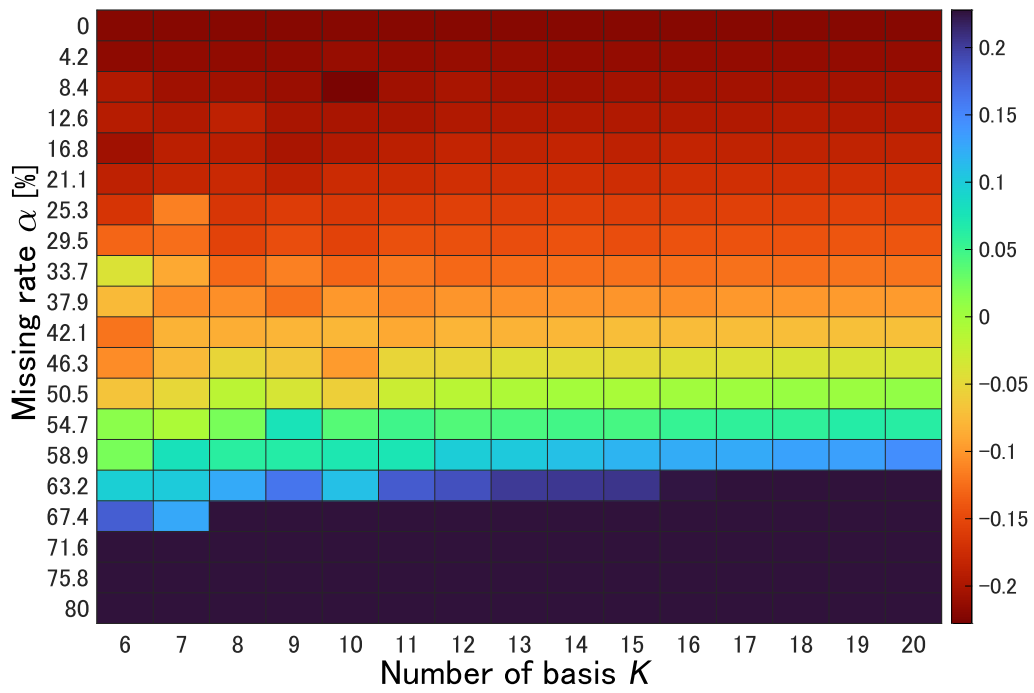


Fig. B.11. Improvement value obtained by L_1 -RNMF: noise type is uniform dist., $r = 40.84$ dB, and $\nu = 0.01$.

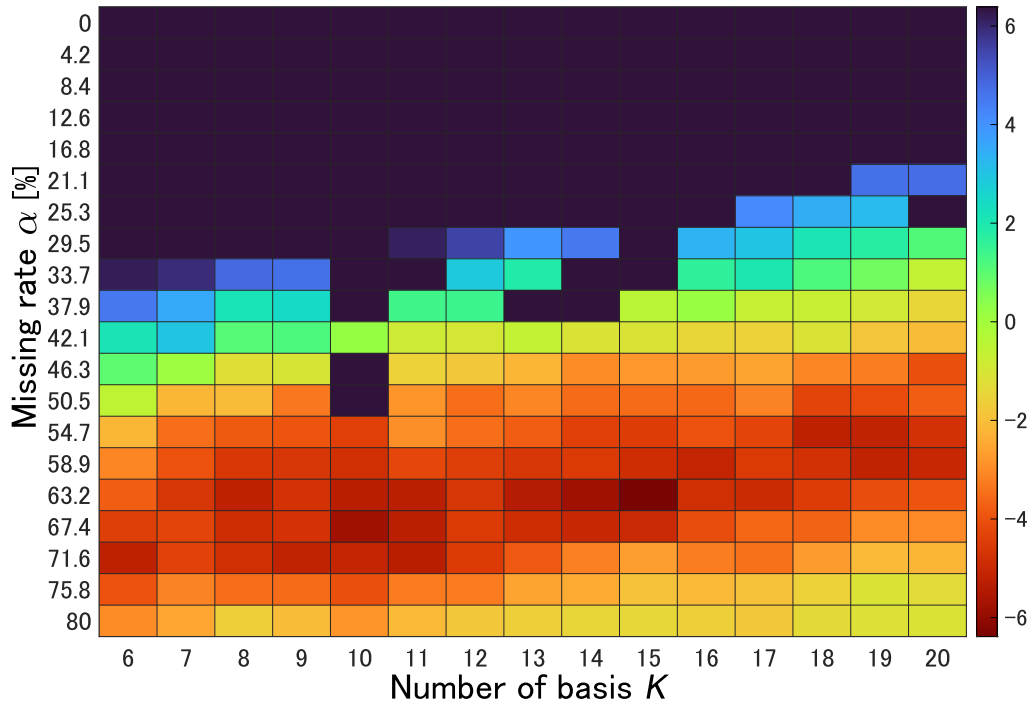


Fig. B.12. Improvement value obtained by CSRNMF: noise type is Rayleigh dist., $r = 0$ dB, and $\mu = 61584.82$.

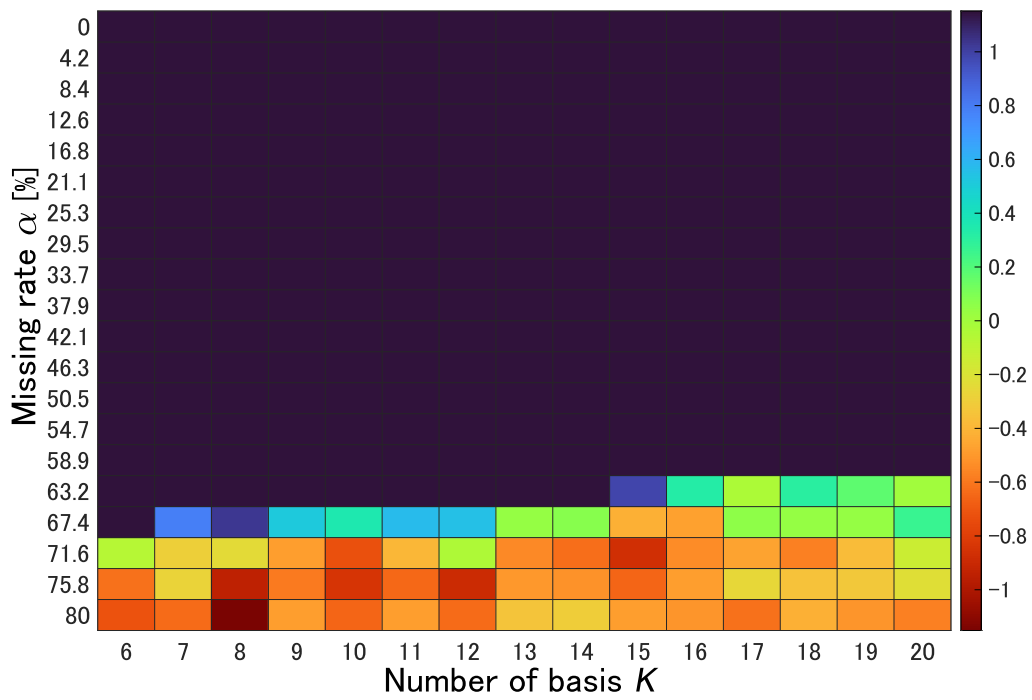


Fig. B.13. Improvement value obtained by CSRNMF: noise type is Rayleigh dist., $r = 3.16$ dB, and $\mu = 784.76$.

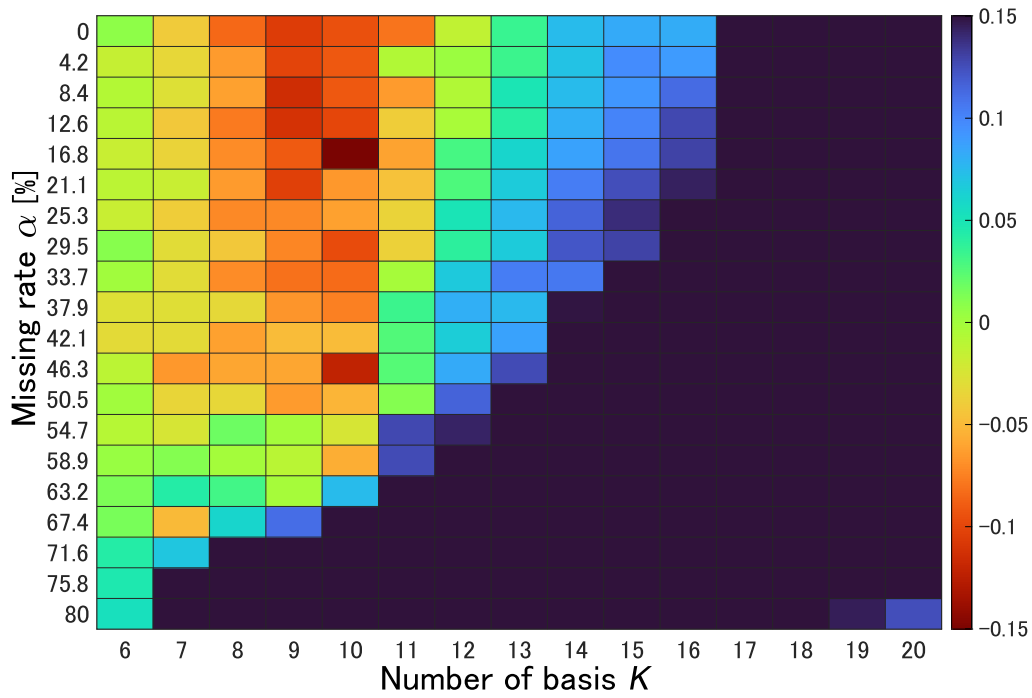


Fig. B.14. Improvement value obtained by CSRNMf: noise type is Rayleigh dist., $r = 40.84$ dB, and $\mu = 0.21$.

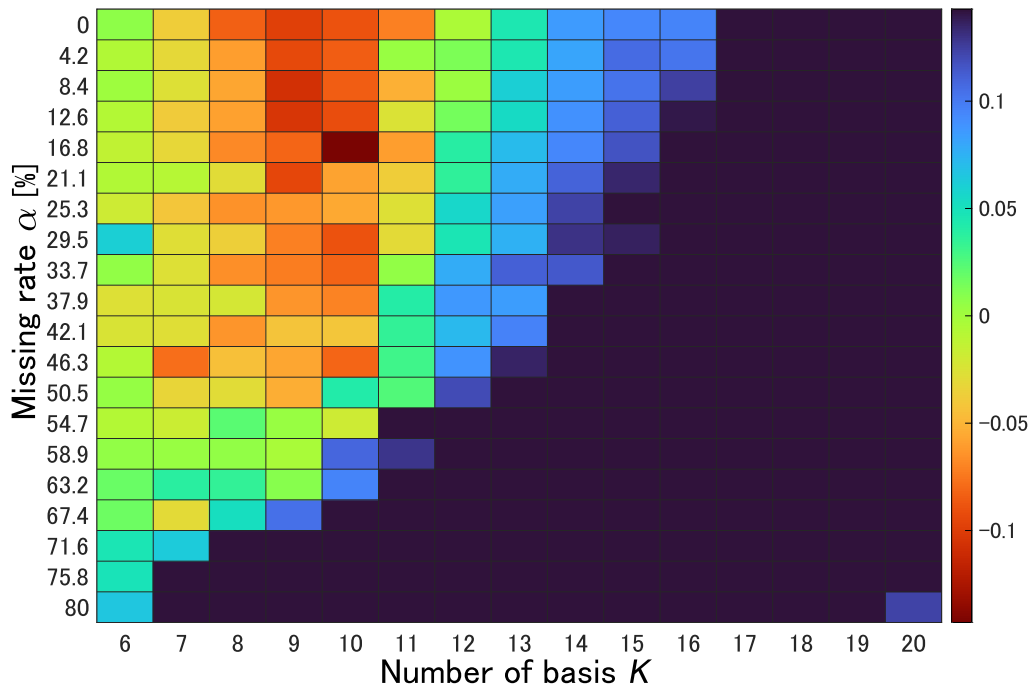


Fig. B.15. Improvement value obtained by CSRNMf: noise type is Rayleigh dist., $r = 77.43$ dB, and $\mu = 0.21$.

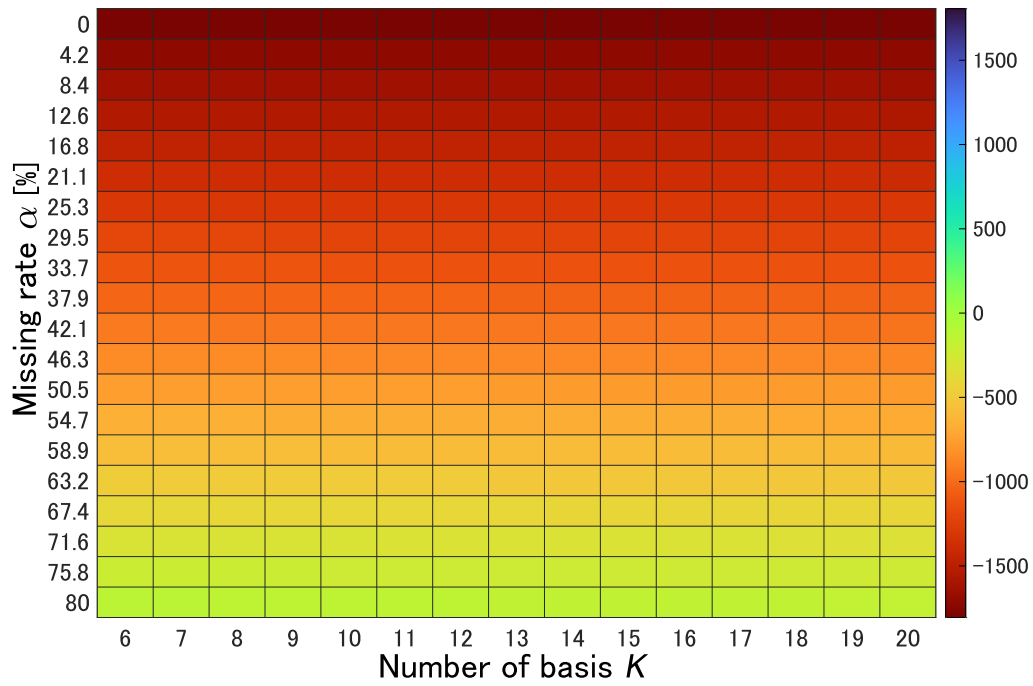


Fig. B.16. Improvement value obtained by L_1 -RNMF: noise type is Rayleigh dist., $r = 0$ dB, and $\nu = 2.34$.

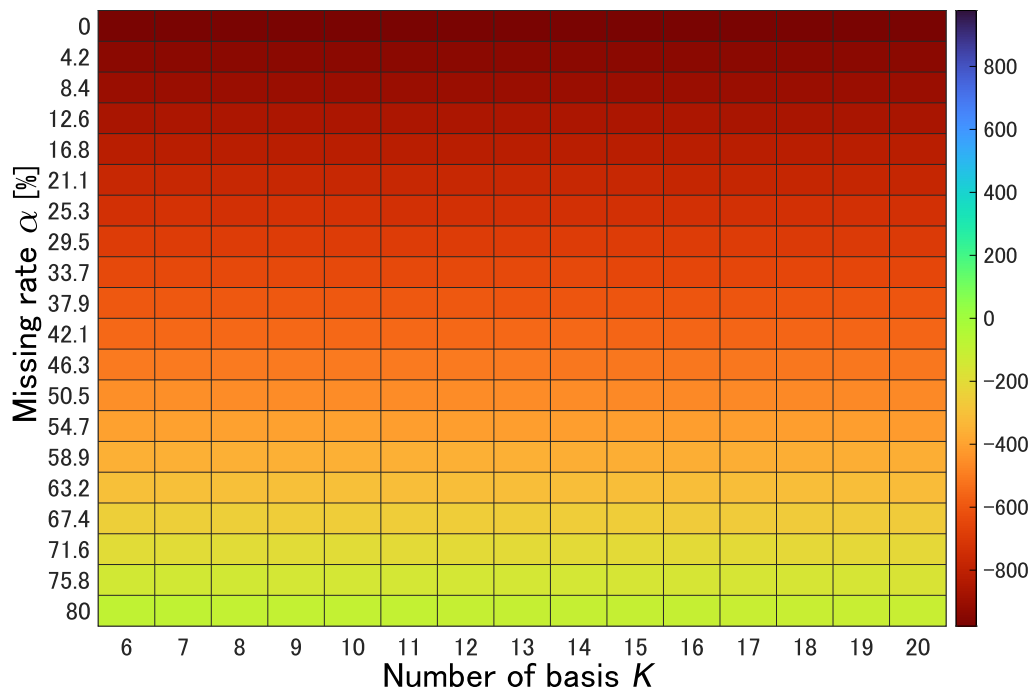


Fig. B.17. Improvement value obtained by L_1 -RNMF: noise type is Rayleigh dist., $r = 3.16$ dB, and $\nu = 0.89$.

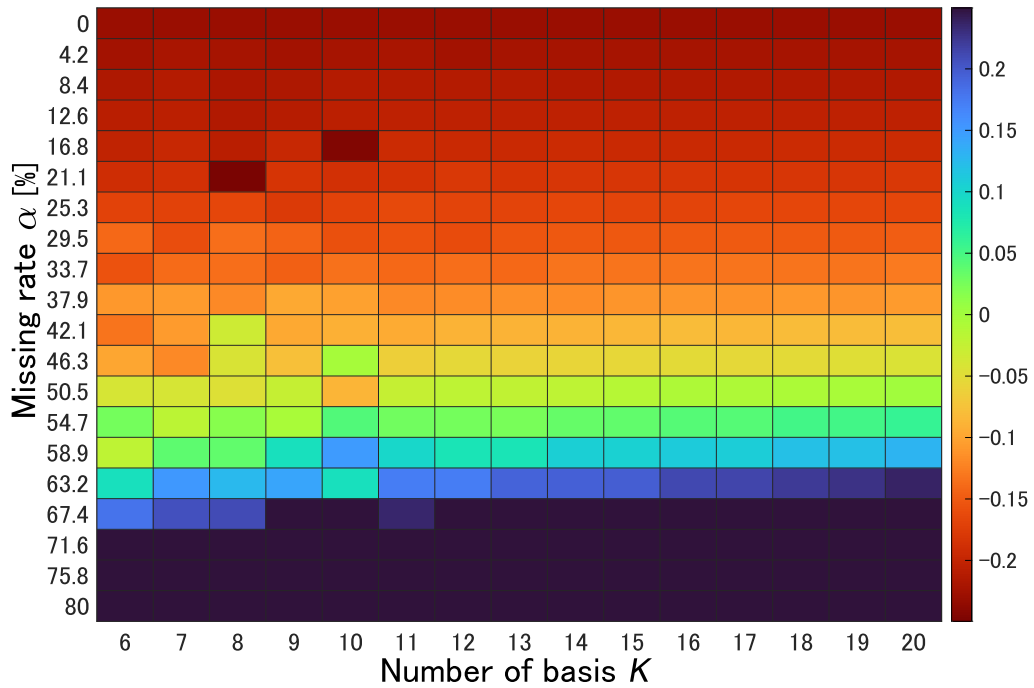


Fig. B.18. Improvement value obtained by L_1 -RNMF: noise type is Rayleigh dist., $r = 40.84$ dB, and $\nu = 0.01$.

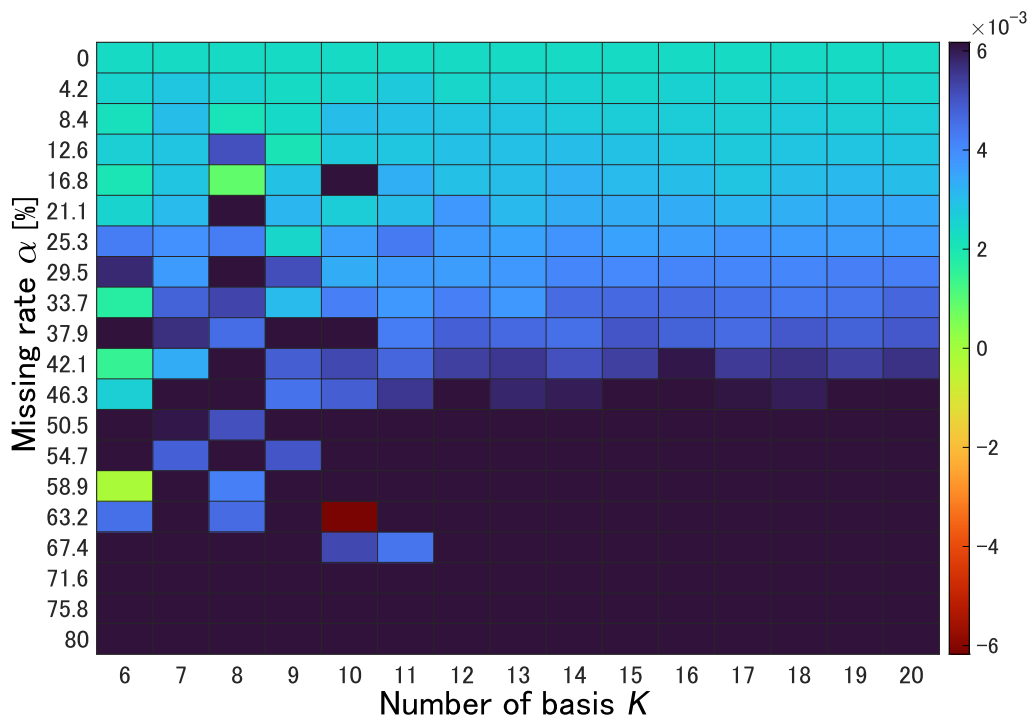


Fig. B.19. Improvement value obtained by L_1 -RNMF: noise type is Rayleigh dist., $r = 77.43$ dB, and $\nu = 0.01$.

付録 C

欠損音声スペクトログラムの補完実験結果

本付録では、6章で行った欠損音声スペクトログラムの補完実験の内、本文で示した条件以外の実験結果を掲載する。Figs. C.1–C.17に、各基底ベクトル数 K 値における各手法の最大値を示す。

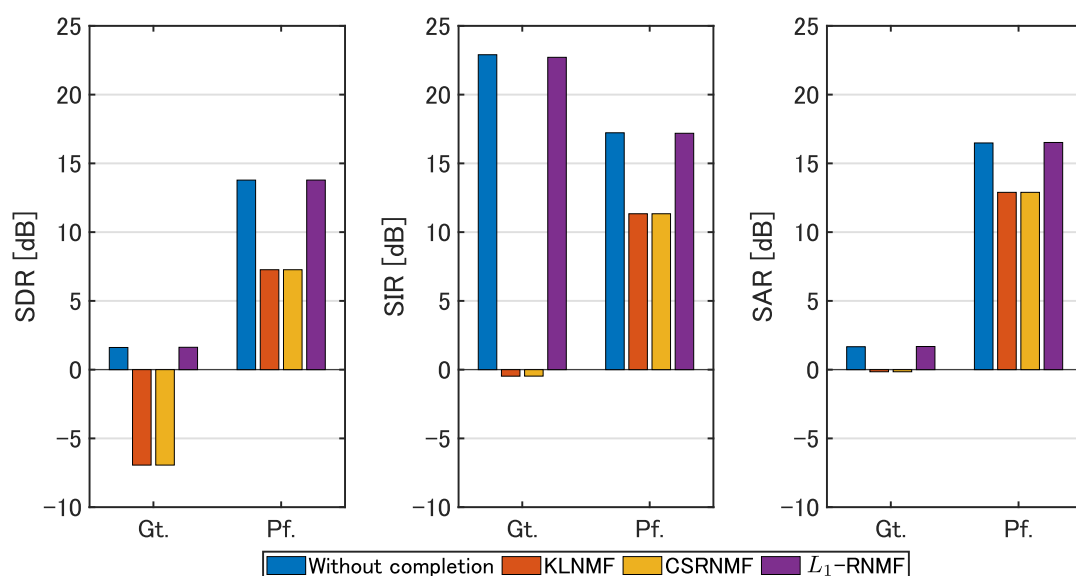


Fig. C.1. Average SDR, SIR, and SAR values when $K = 4$, $\mu = 0.001$, and $\nu = 23357.214$.

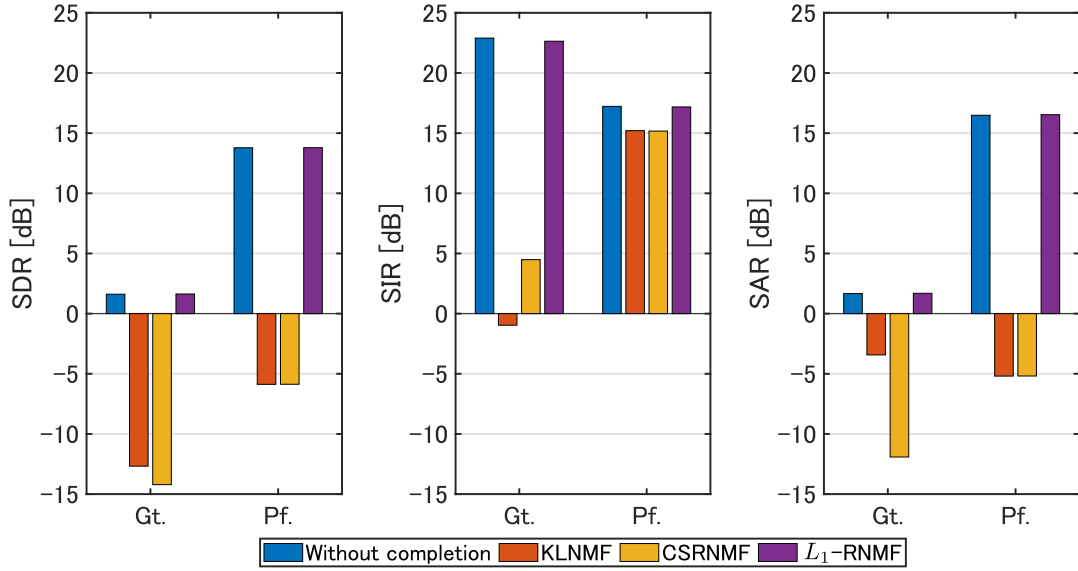


Fig. C.2. Average SDR, SIR, and SAR values when $K = 10$, $\mu = 0.004$, and $\nu = 8858.668$.

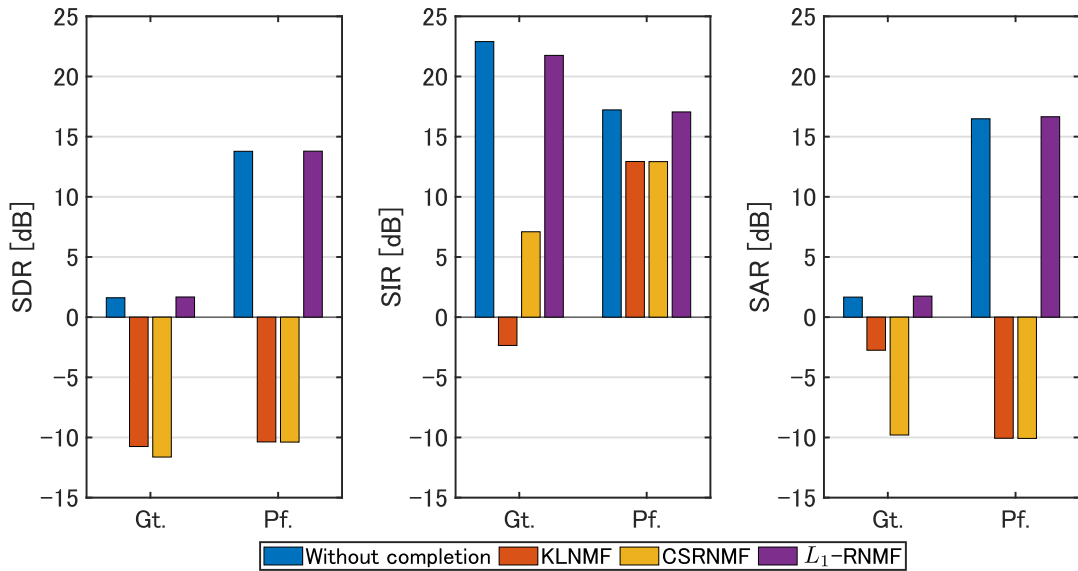


Fig. C.3. Average SDR, SIR, and SAR values when $K = 16$, $\mu = 0.001$, and $\nu = 483.293$.

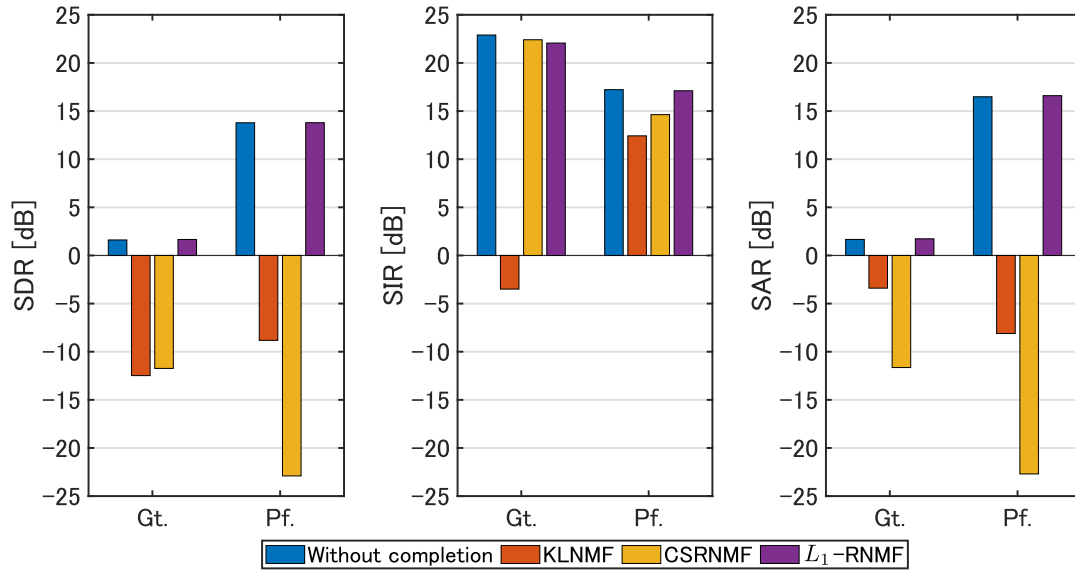


Fig. C.4. Average SDR, SIR, and SAR values when $K = 22$, $\mu = 10$, and $\nu = 784.760$.

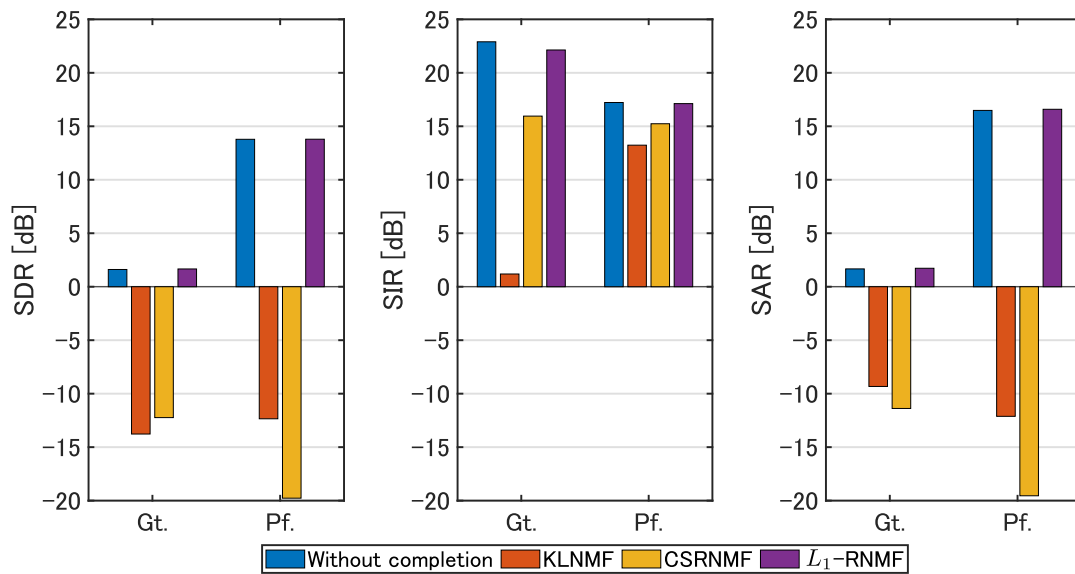


Fig. C.5. Average SDR, SIR, and SAR values when $K = 28$, $\mu = 3.793$, and $\nu = 784.760$.

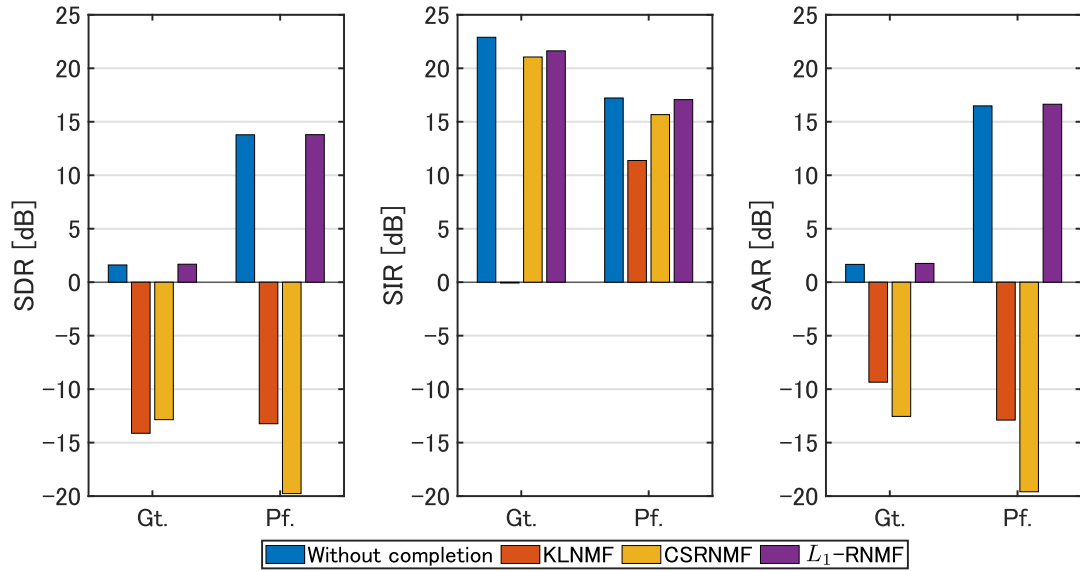


Fig. C.6. Average SDR, SIR, and SAR values when $K = 28$, $\mu = 2.335$, and $\nu = 297.635$.

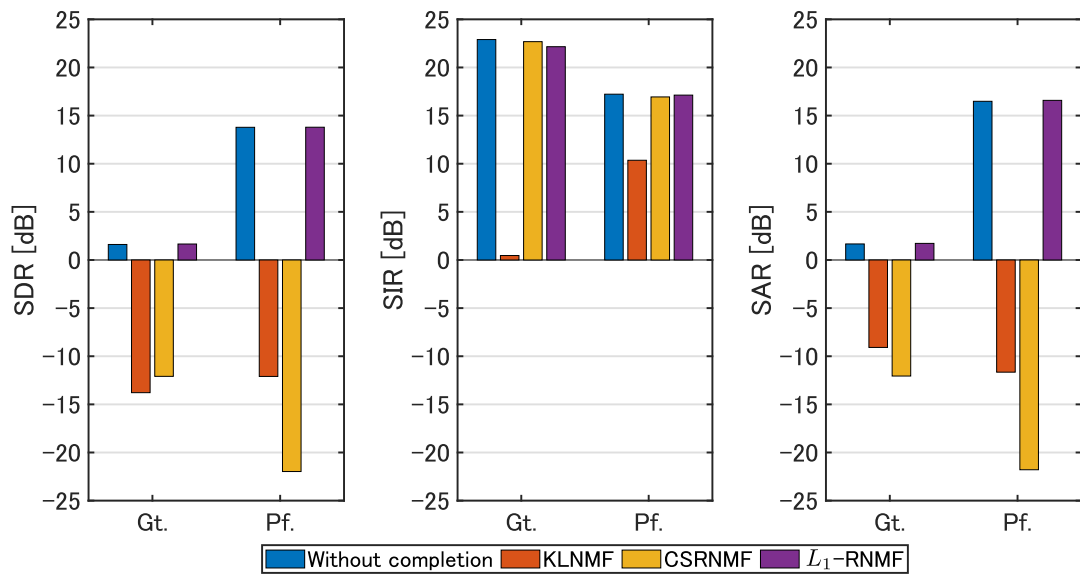


Fig. C.7. Average SDR, SIR, and SAR values when $K = 40$, $\mu = 3.703$, and $\nu = 784.760$.

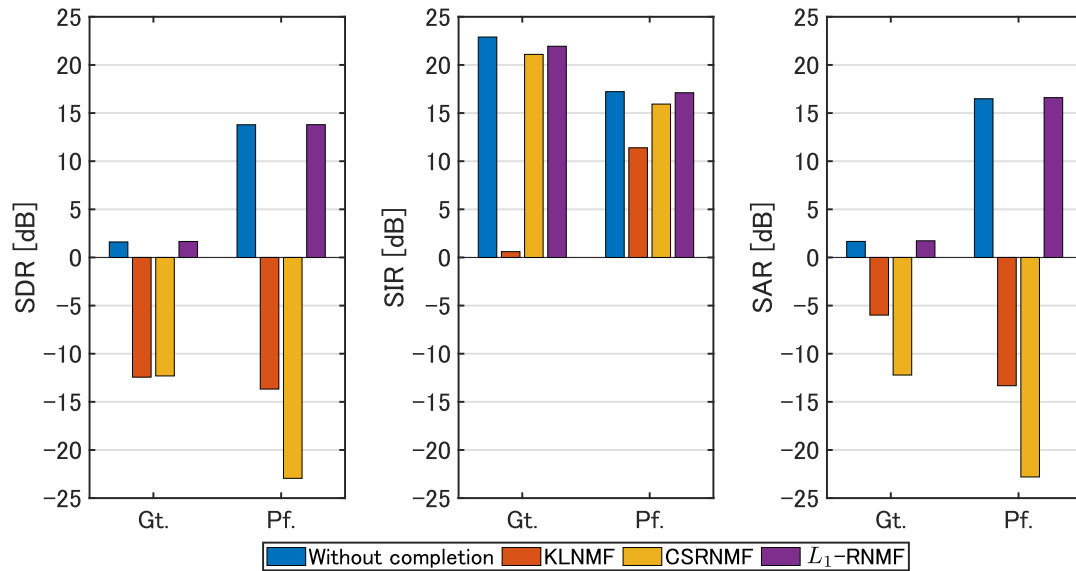


Fig. C.8. Average SDR, SIR, and SAR values when $K = 46$, $\mu = 3.793$, and $\nu = 483.293$.

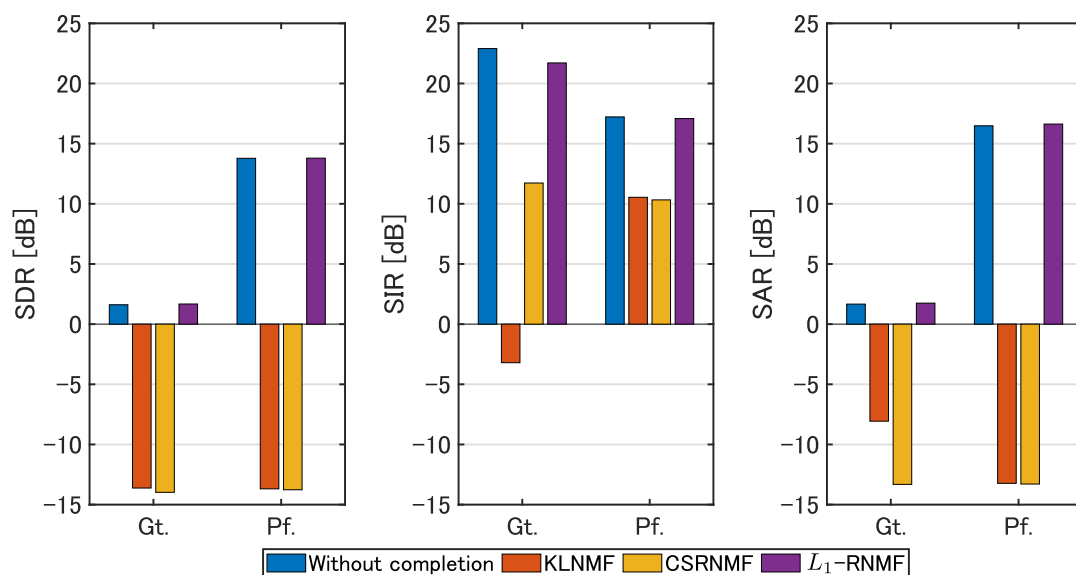


Fig. C.9. Average SDR, SIR, and SAR values when $K = 52$, $\mu = 0.001$, and $\nu = 297.635$.

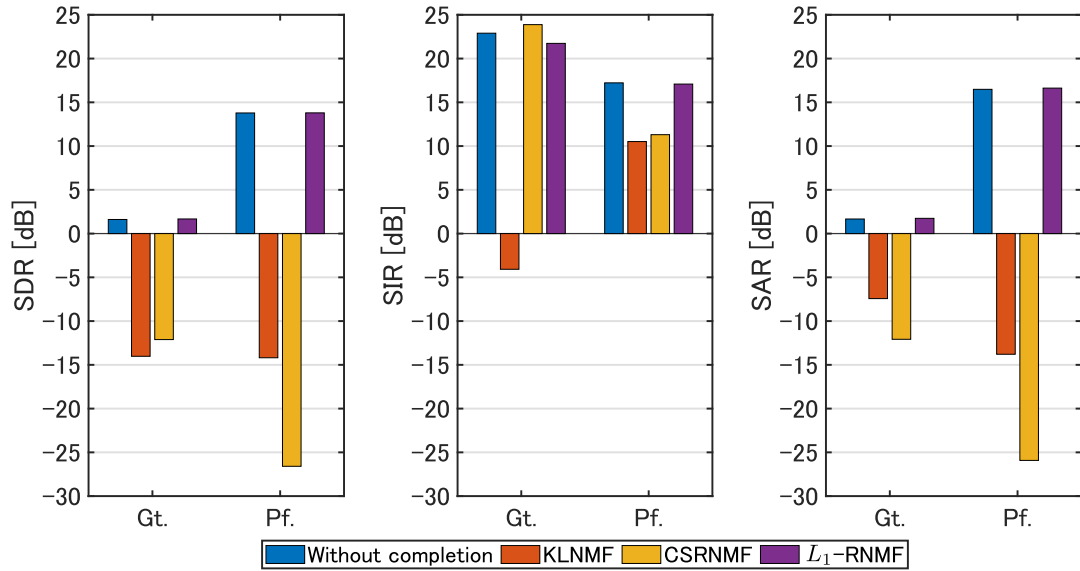


Fig. C.10. Average SDR, SIR, and SAR values when $K = 58$, $\mu = 1.438$, and $\nu = 297.635$.

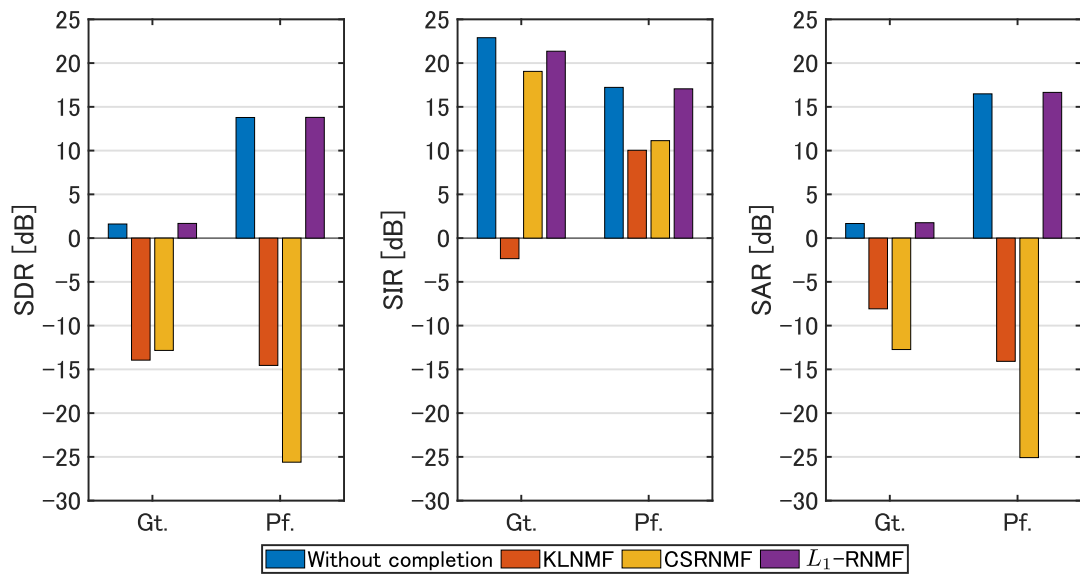


Fig. C.11. Average SDR, SIR, and SAR values when $K = 64$, $\mu = 2.336$, and $\nu = 183.298$.

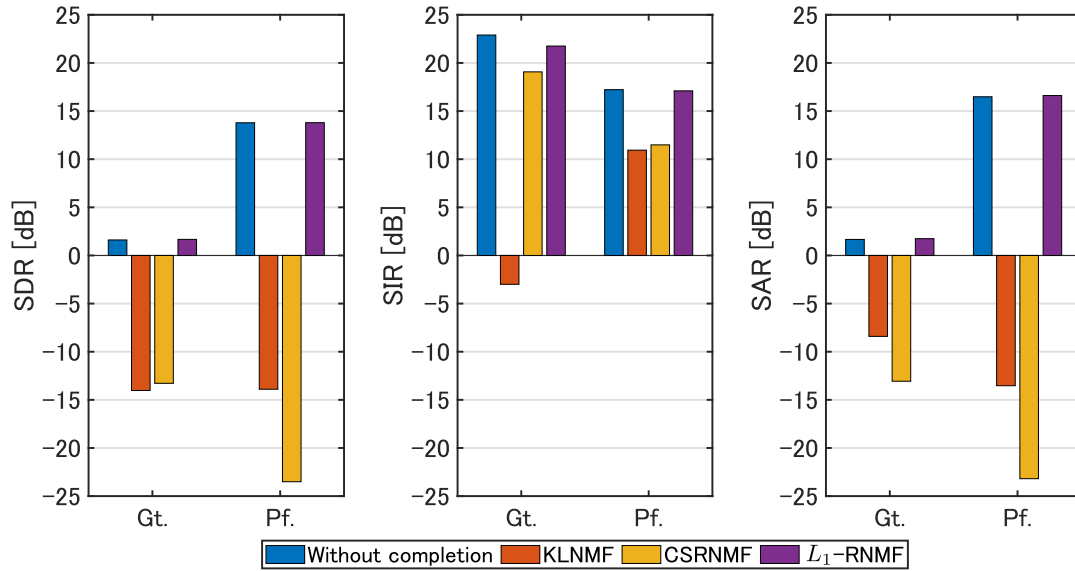


Fig. C.12. Average SDR, SIR, and SAR values when $K = 70$, $\mu = 0.336$, and $\nu = 297.635$.

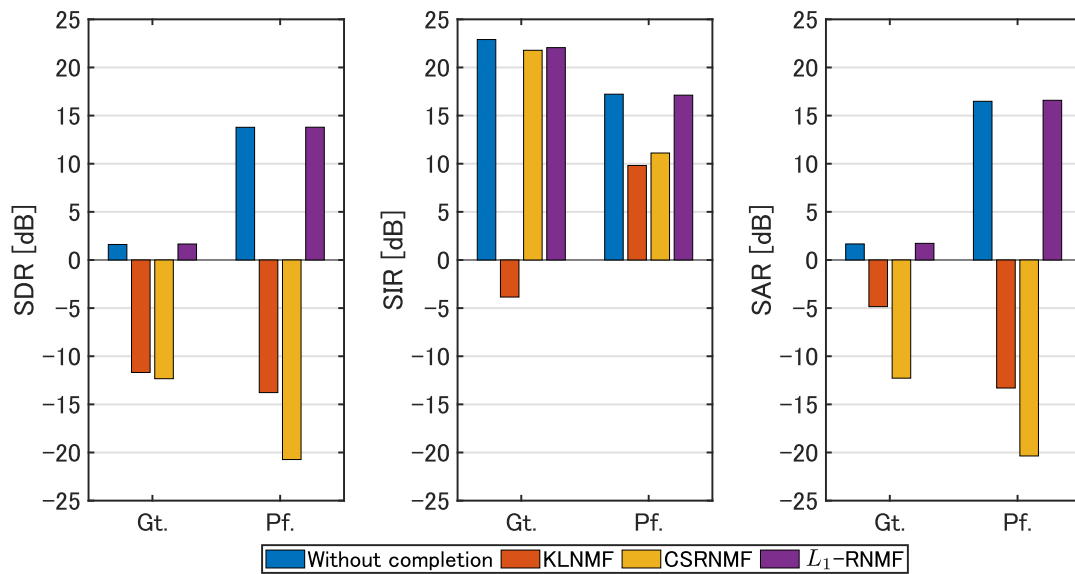


Fig. C.13. Average SDR, SIR, and SAR values when $K = 76$, $\mu = 0.207$, and $\nu = 483.293$.

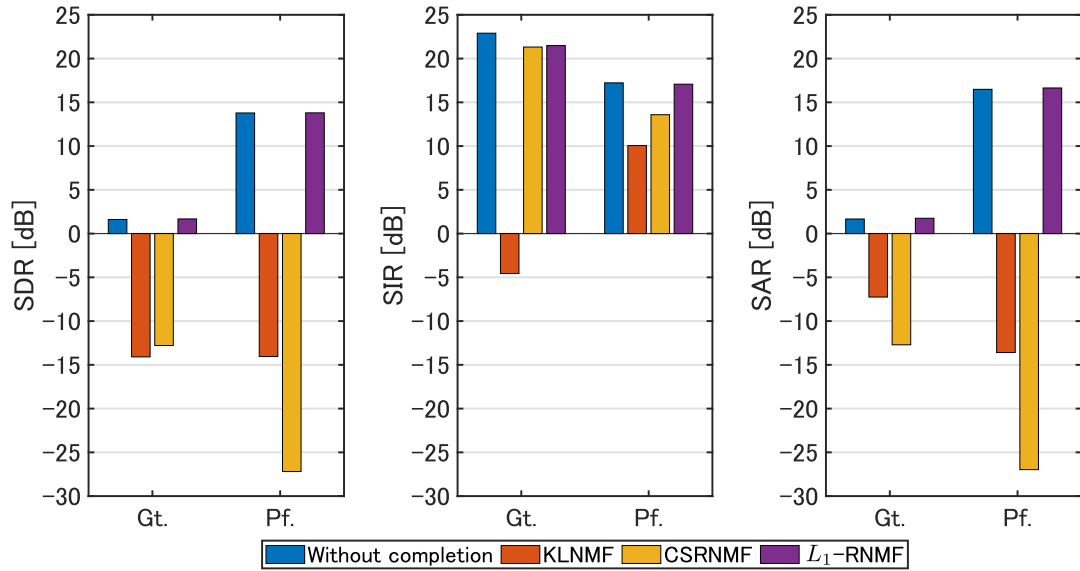


Fig. C.14. Average SDR, SIR, and SAR values when $K = 82$, $\mu = 0.546$, and $\nu = 183.298$.

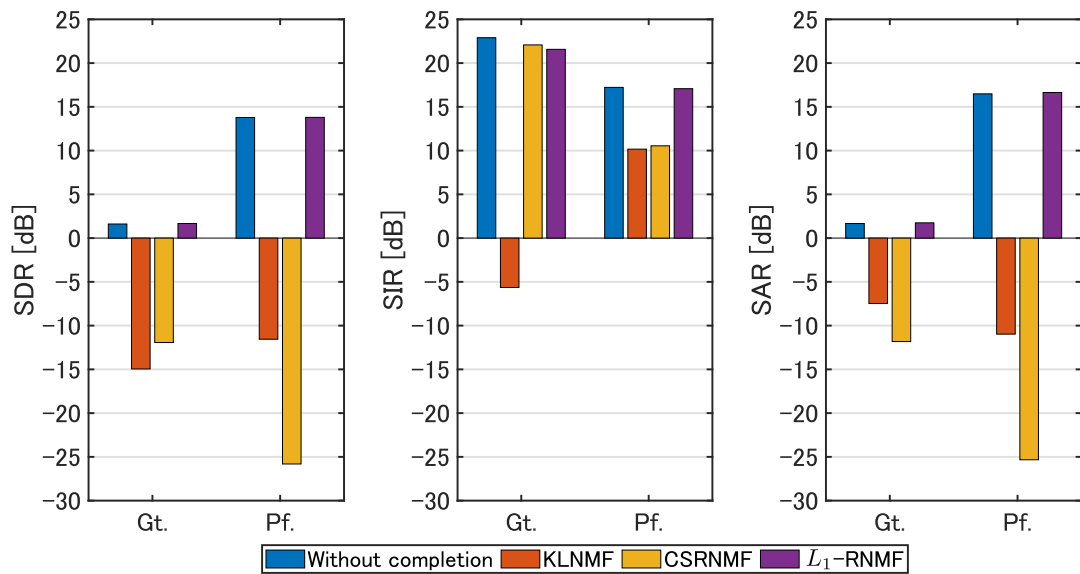


Fig. C.15. Average SDR, SIR, and SAR values when $K = 88$, $\mu = 0.546$, and $\nu = 183.298$.

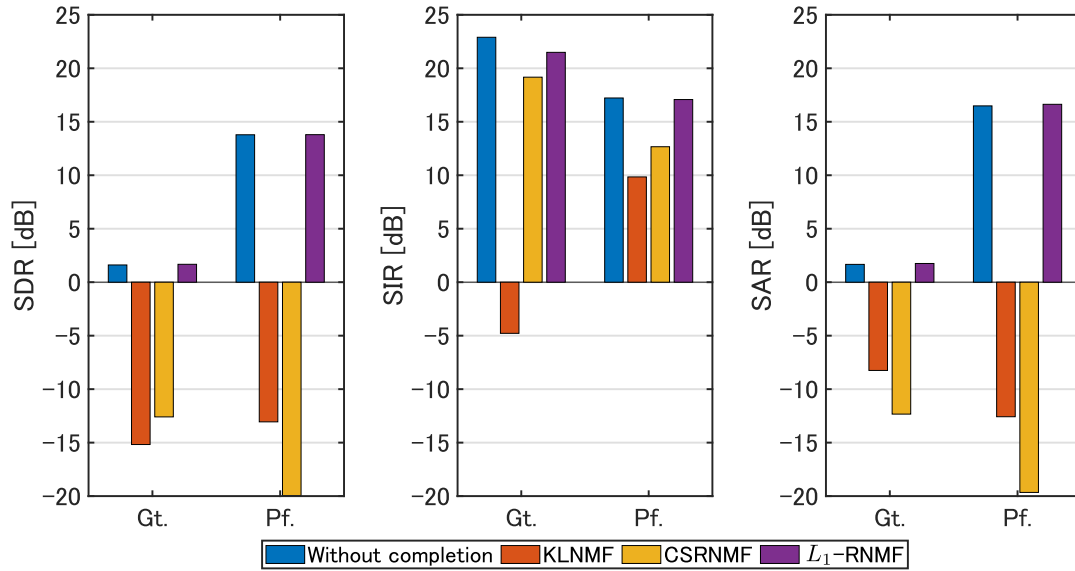


Fig. C.16. Average SDR, SIR, and SAR values when $K = 94$, $\mu = 0.207$, and $\nu = 183.298$.

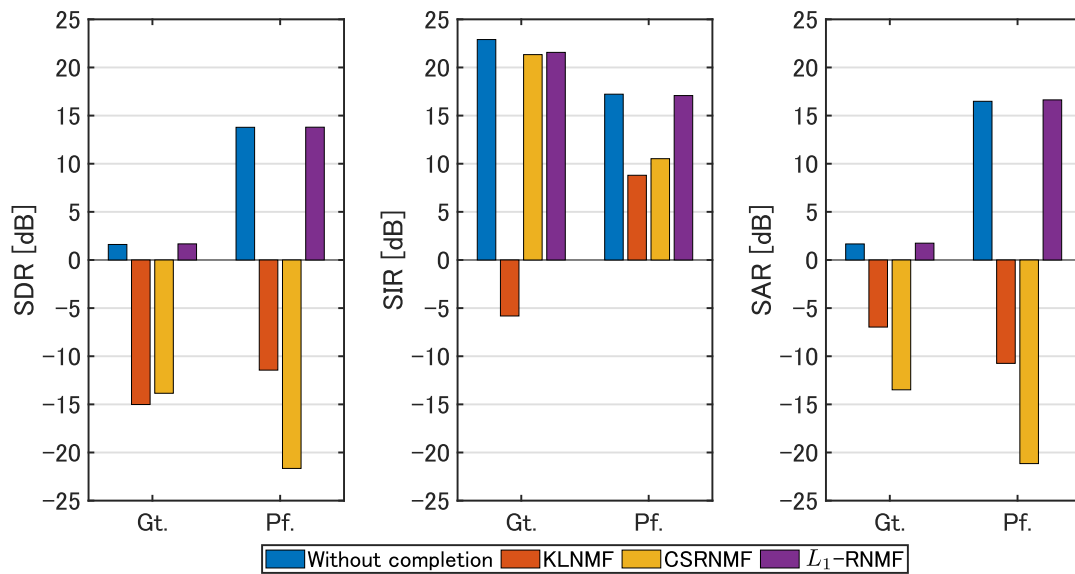


Fig. C.17. Average SDR, SIR, and SAR values when $K = 100$, $\mu = 0.127$, and $\nu = 183.298$.