

修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 情報理工 学研究科 情報・ネットワーク工学 専攻 博士前期課程		
氏 名	藤田 隼輔	学籍番号	1631127
論 文 題 目	ロバスト性及び幾何的性質を考慮した 直交制約付き非負値行列分解に関する検討		
<p>要 旨</p> <p>情報化社会の現代，インターネットを利用して映像，画像，文書等様々なデータや信号が取得される。取得された信号は様々な要因により，ノイズ，外れ値，データ欠損が発生する場合がある。そのような信号やデータを効率よく処理するためには，高いロバスト性を持ち，さらに精度の高いクラスタリングが必要とされる。取得された信号は，行列で表現することが可能であり，その要素を非負値と仮定しても多くの場合不都合を生じないものである。非負値の値を持つ行列を解析する一手法として，Nonnegative Matrix Factorization(NMF)がある。この手法は，NMF は 1990 年代に画像処理の技術の一旦として研究されてきた手法であり，2001 年に乗法的更新則によって実現するアルゴリズムが発表され，現在もその応用の可能性から研究されている手法である。応用の一例としてクラスタリングがある。拡張手法に Orthogonal Nonnegatively Penalized Matrix Factorization (ONP-MF) がある。ONP-MF はロバスト性に乏しく，幾何的な情報が欠如している課題があり，クラスタリングとして応用した時の精度の低下の要因となっている。そこで，課題解決のため，ONP-MF を $L_{2,1}$ ノルムと局所的不変ペナルティ項によって拡張した Robust-Locally-invariant-ONP-MF を提案した。提案手法のロバスト性の確認及び，局所的な不変ペナルティ項による精度への影響を確認する実験を行なった。ノイズ密度を変化させた評価結果より，多手法より高いロバスト性が確認できた。また，幾何学的情報の付与によるクラスタリング精度が向上を確認した。今回の実験では，対象データと結果値が少ないため，今後として対象とするデータを増やすこと，パラメータの繊細化，動作時間等によつての評価を行い拡張による有効性を確認する必要がある。</p>			

平成 29 年度 修士論文

ロバスト性及び幾何的性質を考慮した
直交制約付き非負値行列分解に関する検討

電気通信大学大学院
情報・ネットワーク工学専攻
電子情報学プログラム

学籍番号：1631127

藤田 隼輔

指導教員

笠井 裕之 准教授

森田 啓義 教授

平成 30 年 1 月 29 日

目次

第1章 序論	5
1.1 背景	5
1.2 目的	5
1.3 問題点	6
1.4 構成	6
第2章 関連研究	7
2.1 概要の説明	7
2.1.1 式の表記	7
2.1.2 フロベニウスノルム	7
2.1.3 $L_{2,1}$ ノルム	7
2.1.4 クラスタリングの次元縮約	8
2.1.5 特異値分解 (SVD)	8
2.1.6 勾配法	8
2.1.7 直線探索	9
2.1.8 ペナルティ法	9
2.1.9 拡張ラグランジュ法	9
2.1.10 双対定理	10
2.1.11 極分解	10
2.2 非負値行列分解 (Nonnegative Matrix Factorization:NMF)	11
2.2.1 乗法的更新則	11
2.2.2 クラスタリングへの応用法	12
2.3 NMFの拡張方式	13
2.3.1 Orthogonal Nonnegative Matrix Factorization(ONMF)	15
2.3.2 Robust Nonnegative Matrix Factorization(RNMF)	16
2.3.3 Graph Regularized NMF(GNMF)	16

2.3.4	Orthogonal Nonnegatively Penalized Matrix Factorization(ONP-MF)	18
2.4	手法の比較	21
第3章	提案手法	22
3.1	Robust Locally ONP-MF(Rbust-Li-ONP-MF)	22
3.2	分解行列とラグランジュ乗数の更新	23
3.2.1	\mathbf{U} 更新	23
3.2.2	\mathbf{V} 更新	23
3.2.3	Λ 更新	24
第4章	評価実験	25
4.1	対象データセット	25
4.2	実験条件	25
4.3	評価指標	27
第5章	実験結果	29
第6章	結論	39

目 次

2.1	NMF クラスタリング	12
2.2	関連度例	13
2.3	NMF モデルとアルゴリズムのカテゴリ [1]	15
5.1	精度評価図 ($\mu=0.01, p=5, \text{noise}=0.05$)	36
5.2	精度評価図 ($\mu=0.01, p=10, \text{noise}=0.1$)	36
5.3	精度評価図 ($\mu=0.01, p=5, \text{noise}=0.05$)	37
5.4	精度評価図 ($\mu=0.1, p=10, \text{noise}=0.05$)	37
5.5	精度評価図 ($\mu=0.1, p=20, \text{noise}=0.05$)	38

表 目 次

2.1	主問題と双対問題	10
2.2	手法の比較	21
4.1	実験データセット	25
5.1	クラスタリング実験結果 ($p\text{-num}=5$, noise level=0.05)	30
5.2	クラスタリング実験結果 ($p\text{-num}=10$, noise level=0.05)	30
5.3	クラスタリング実験結果 ($p\text{-num}=15$, noise level=0.05)	31
5.4	クラスタリング実験結果 ($p\text{-num}=20$, noise level=0.05)	31
5.5	クラスタリング実験結果 ($p\text{-num}=5$, noise level=0.075)	32
5.6	クラスタリング実験結果 ($p\text{-num}=10$, noise level=0.075)	32
5.7	クラスタリング実験結果 ($p\text{-num}=15$, noise level=0.075)	33
5.8	クラスタリング実験結果 ($p\text{-num}=20$, noise level=0.075)	33
5.9	クラスタリング実験結果 ($p\text{-num}=5$, noise level=0.1)	34
5.10	クラスタリング実験結果 ($p\text{-num}=10$, noise level=0.1)	34
5.11	クラスタリング実験結果 ($p\text{-num}=15$, noise level=0.1)	35
5.12	クラスタリング実験結果 ($p\text{-num}=20$, noise level=0.1)	35

第1章 序論

1.1 背景

情報化社会の現代，インターネットを利用して映像，画像，文書等様々なデータや信号が取得される．取得された信号は様々な要因により，ノイズ，外れ値，データ欠損が発生する場合がある．そのような信号やデータを効率よく処理するためには，高いロバスト性を持ち，さらに精度の高いクラスタリングが必要とされる．

1.2 目的

取得された信号は，行列で表現することが可能であり，その要素を非負値と仮定しても多くの場合不都合を生じないものである．非負値の値を持つ行列を解析する一手法として，非負値行列因子分解 [Nonnegative Matrix Factorization (以後 NMF と記述する)] がある [1]．この手法は，NMF は 1990 年代に画像処理の技術の一旦として研究されてきた手法であり，2001 年 [2] に乗法的更新則によって実現するアルゴリズムが発表され，現在もその応用の可能性から研究されている手法である．応用の一例としてクラスタリングがあり，様々な制約等を与えることで，それに応じた性能を得ることができる手法である．また，NMF のクラスタリングは，次元縮約 [3] を利用したものであるため，大規模なデータの処理も行うことが可能である．本研究では，NMF の性質を利用し，各々の問題点の解決のための関連手法の特徴を取り入れた Robust-Li-ONP-MF を提案し，その特徴を示すかの検討である．検討の結果，高い分類性能を示すことができれば，ノイズや外れ値が含まれる大規模なデータや信号に関しても，効率よく処理することが可能となると考えられる．

1.3 問題点

NMFは直交制約を付加した Orthogonal NMF (ONMF) がある。直交制約を付加することでk-meansと同等のクラスタリング性能を持つようになる。これは、非負値制約を満たしながら漸近的に直交制約を満たすように更新を行うものである。

ONMFは非負値を満たす更新によって局所解に収束する場合があること、収束値が初期値の影響を受けるという問題点がある。それらを解決する手法として、Orthogonal Nonnegative Penalized Matrix Factorization (ONP-MF) がある。ONP-MFはONMFの目的関数を拡張ラグランジュ法によって定義されたものであり、直交制約を満たしながら漸近的に非負値を満たすものである。また、初期値を乱数生成ではなく、Singular Value Decomposition (SVD) によって一意に設定するものである。

ONP-MFの問題点として、ロバスト性が乏しい、つまり、ノイズや外れ値データの分類精度が低いこと、幾何学的な情報の欠如という問題点がある。これらの問題点はクラスタリングをする際に精度の低下をもたらす原因となる。

そこで、以上に述べた問題解決のため、本研究ではONP-MFの拡張を行う。拡張の内容として、ONP-MFにロバスト性を持たせること、データの幾何学的な構造を保存を行う。

1.4 構成

本稿の構成を以下に示す。

2章では、関連研究について紹介する。3章では、関連研究の特徴を取り入れた提案手法について説明する。4章では、提案手法についての精度評価実験について述べる。最後に5章でまとめについて述べる。

第2章 関連研究

2.1 概要の説明

はじめに関連研究を説明する上での，表記や語句等に概要について述べる．

2.1.1 式の表記

本稿では，行列を \mathbf{A} のように太字の大文字表記をする．また，ベクトルの表記は \mathbf{a} のように太字の小文字表記をする． \odot はアダマール積， \div は elementwise division を表す． $\mathbf{A} \geq 0$ は行列が非負値であることを示している．subject to は制約条件， $\mathbf{A} \in \mathbb{R}^{m \times k}$ は行列 \mathbf{A} が m 行 k 列であることを示している．

2.1.2 フロベニウスノルム

行列の大きさを表す量の一つである． $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n]$ 場合，フロベニウスノルムは式 (2.1) のように定義される．

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i,j} a_{i,j}^2} \quad (2.1)$$

また，フロベニウスノルムには，式 (2.2) のように，トレース行列として表現することができる．

$$\|\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}\mathbf{A}^T) = \text{tr}(\mathbf{A}^T\mathbf{A}) \quad (2.2)$$

2.1.3 $L_{2,1}$ ノルム

$L_{2,1}$ ノルムはサブベクトルごとに L_2 ノルムを計算し，その総和をとったものである．フロベニウスノルムと違い，二乗誤差を行わないノルム

である。式 (2.3) のように定義される。NMFをこのノルムを用いて目的関数を定義することで、ロバスト性を持つNMFが提案されている。2.3.2で取り扱う。

$$\|\mathbf{A}\|_{2.1} = \sum_{i=1}^n \|a_i\| \quad (2.3)$$

2.1.4 クラスタリングの次元縮約

クラスタリングには以下のような課題が存在する。

- 高次元になることでベクトル間距離が離れていく
- 本質的にクラスタリングは距離の近いものをまとめる
- 高次元になるほど妥当な結果を得ることが難しくなる

次元縮約 [3] を用いることで以下のような課題を解決することが可能である。

- 対処法として次元縮約を行う
- 高次元のベクトル間の位置関係をできるだけ保存
- より低次元のベクトルに変換する処理

2.1.5 特異値分解 (SVD)

特異値分解 [Singular Value Decomposition (SVD)] [4] は $\mathbf{A} \in \mathbb{R}^{m \times n}$ の任意の行列に対して、 $\mathbb{R}^{m \times m}$ 、 $\mathbb{R}^{n \times n}$ の直交行列と $m \times n$ の対角行列内積で表現される。式 (2.4) で表現される。

$$\mathbf{A} = \mathbf{U} \times \mathbf{S} \times \mathbf{V} \quad (2.4)$$

\mathbf{S} の非対角成分は0であり、対角成分は、降順となっている。

2.1.6 勾配法

勾配法 [9] は目的関数を適当な初期値から目的関数をより小さくする方向へ繰り返し移動することで最終的に局所的最適解へ行きつくという考え方による手法の一つである。目的関数は式 (2.5) で表される。

- f :目的関数, x :初期解, 次解: \hat{x} , α_0 :初期のステップ幅

$$\hat{x} = \hat{x} - \alpha_0 \nabla f(x) \quad (2.5)$$

- ステップ幅は $f(\hat{x}) < f(x)$ となるように選ぶ
- x^t が1次の必要条件を満たすまで繰り返し行い最適解を求める

2.1.7 直線探索

直線探索[9]は, 連続最適化問題に置いて, 目的関数を求めるための手法の一つである. 現在の近似解 x から降下する方向 d だけ移動した点 $x + \alpha d$ が最小となるような α を求める手法である.

2.1.8 ペナルティ法

ペナルティ法[9]は, 制約条件を破った時に目的関数値にペナルティを加えることによって, 制約条件を目的関数に取り込み, 制約なし問題として解く手法である. 制約条件を破った場合には, その度合いに応じてペナルティを課すような値を与える.

2.1.9 拡張ラグランジュ法

拡張ラグランジュ法[10]は制約ありの最適化問題の局所最適解を数値的に求める計算アルゴリズムの一つである. ラグランジュ未定乗数法とペナルティ法を組み合わせたものである. 式 (2.6) で表現される.

- λ :ラグランジュ乗数
- ρ :二次ペナルティパラメータ
- $\rho > 0$

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x) + \frac{\rho}{2} \sum_{j=1}^m \{g_j(x)\}^2 \quad (2.6)$$

2.1.10 双対定理

定式化された線形計画問題には，表裏の関係にある線形計画問題が存在し，表に当たる問題を主問題，裏にあたる問題を双対問題[9]という．主問題が最適解を持つ場合，双対問題も最適解を持ち，主問題における最小値 (z) と双対問題における最大値 (w) は等しくなる．また，双対問題のそのまた双対問題は主問題と等しいという性質を持つ定理．

表 2.1: 主問題と双対問題

目的関数	制約条件
主問題	
$z = c^T x$ を最小化	$Ax = b, x \geq 0$
双対問題	
$w = b^T y$ を最大化	$A^T y \leq c$

2.1.11 極分解

極分解[11]は任意の実正方行列 M は直交行列 U と半正定値行列 P に分解する手法である．それぞれの行列について以下に示す．

$$M = UP \quad (2.7)$$

- 実正方行列
実数を成分にとる正方行列
- 半正定値行列
 $n \times n$ の実対象行列 M が以下の条件の一つを満たす行列
 - i 全ての n 次元実ベクトル x に対して二次形式 $x^T M x$ が非負
 - ii M の固有値は全て非負
 - iii ある実正方行列 X が存在し， $M = X^T X$ と表すことができる
 - iv M の主小行列式が全て非負

2.2 非負値行列分解 (Nonnegative Matrix Factorization:NMF)

NMF[2] は非負信号から構成される入力行列 $\mathbf{X} \in \mathbb{R}^{m \times n}$ を, ランク k の2つの非負の行列 $\mathbf{U} \in \mathbb{R}^{m \times k}$, $\mathbf{V} \in \mathbb{R}^{k \times n}$ に分解する問題である. \mathbf{U} は辞書行列, \mathbf{V} は係数行列と呼ばれる. NMF をクラスタリングに用いる場合は係数行列 \mathbf{V} を用いて行う. 目的関数を式 (2.8) に示す.

- $\mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{U} \in \mathbb{R}^{m \times k}, \mathbf{V} \in \mathbb{R}^{k \times n}$

$$\begin{aligned} \min_{\mathbf{U} \in \mathbb{R}^{m \times k}, \mathbf{V} \in \mathbb{R}^{k \times n}} \quad & \|\mathbf{X} - \mathbf{UV}\|_F^2 \\ \text{subject to} \quad & \mathbf{U} \geq 0, \mathbf{V} \geq 0 \end{aligned} \quad (2.8)$$

NMF は非負値という制約以外に制約を与えたり, 緩和させることで, それに応じた能力を持たせることができる手法である [12]. 1.3 で述べた問題解決のため, 幾らかの関連研究の特徴に着目した2章ではこれについて述べる.

2.2.1 乗法的更新則

値の収束法として, NIPS2001 において, 非負制約を満たしながら乗法的更新則 (Multiplicative Update) により実現するアルゴリズムが Lee らにより提案されている. NMF の \mathbf{U} 及び \mathbf{V} の乗法的更新則を式 (2.9), (2.10) に示す.

- $\mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{U} \in \mathbb{R}^{m \times k}, \mathbf{V} \in \mathbb{R}^{k \times n}$

- \div は elementwise division

- \odot はアダマール積

$$\mathbf{U} \leftarrow \mathbf{U} \odot \frac{\mathbf{XV}^T}{\mathbf{UVV}^T} \quad (2.9)$$

$$\mathbf{V} \leftarrow \mathbf{V} \odot \frac{\mathbf{U}^T \mathbf{X}}{\mathbf{U}^T \mathbf{UV}} \quad (2.10)$$

式 (2.9), (2.10) を条件を満たすまで繰り返し反復させることで, 最適な値に収束させる.

2.2.2 クラスタリングへの応用法

NMFは次元縮約を利用したクラスタリング手法であり，係数行列 V を用いる[3]．辞書行列 V の i 番目の行ベクトルが元のデータ X の i 番目の列ベクトルを縮約した結果となっている．これは，縮約後の各々の軸がクラスタとなるトピックを表現している．つまり，辞書行列 V の k 列目の要素の値が， k 番目のトピックとの関連度の大きさを表している．よって， i 番目のトピックのクラスタ番号は式(2.11)のように定義される[13]．
 このように，NMFは次元縮約を用いたクラスタリングであるため，大規模なデータのクラスタリングも可能な手法である．イメージを図2.1, 2.2に示す．

$$\arg \max_h v_{ih} \quad (2.11)$$

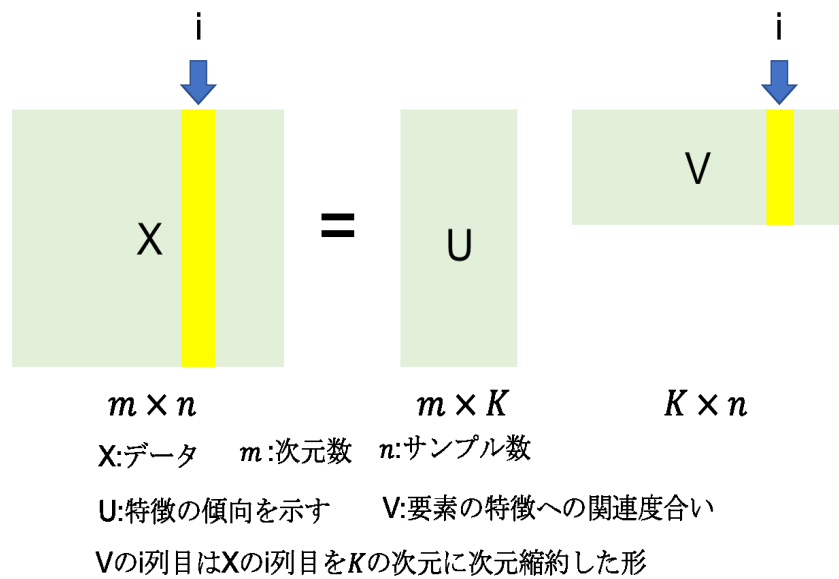


図 2.1: NMF クラスタリング

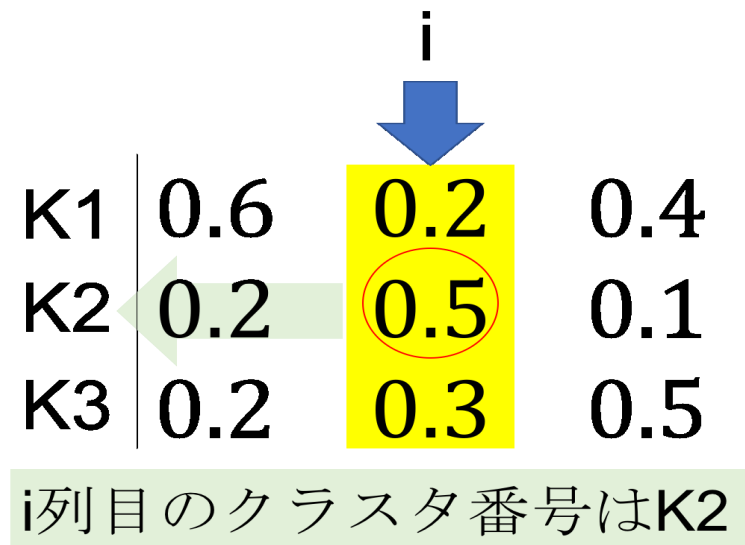


図 2.2: 関連度例

2.3 NMF の拡張方式

NMF は様々な応用研究が行われている。カテゴリの図を図 2.3 に示す。現在存在する NMF は以下の 4 つのカテゴリに分けることができる [1]。

- Basic NMF (BNMF) : 非負値制約のみの NMF
- Constrained NMF (CNMF) : 正則化のためにいくつかの制約のある NMF
- Structured NMF (SNMF) : 通常の因子分解を変更した NMF
- Generalized NMF (GNMF) : 広義的に従来のデータ型や因子モデルを打ち破っている NMF

更に 4 つのカテゴリからいくつかのサブクラスにも分類される。CNMF は以下の 4 つのサブクラスに分類される。

- Sparse NMF (SPNMF) : スパース制約を課した NMF
- Orthogonal NMF (ONMF) : 直交制約を課した NMF
- Discriminant NMF (DNMF) : 分類と判別のための情報を含んだ NMF
- NMF on manifold (MNMF) : 局所的にトポロジカルな特性を保っている NMF

SNMF は以下の3つのサブクラスに分類される.

- Weighed NMF (WNMF) :各成分の相対的重要度の違いによって重み付けする NMF
- Conventional NMF (CVNMF) :時間-周波数領域での因子分解を考慮する NMF
- Nonnegative Matrix Tri-Factorization (NNTF) :データ行列を3つに因子分解する NMF

GNMF は以下の4つのサブクラスに分類される.

- Semi- NMF:特定の因子行列に制約を緩和した NMF
- Nonnegative Tensor Factorization (NTF) :行列形式のデータを高次元のテンソルに一般化する手法
- Nonnegative Matrix-Set Factorization (NMSF) :データセットを行列から行列のセットに拡張する NMF
- NMF on manifold (KNMF) :非線形モデルの NMF

ここで挙げた手法のうち, 本稿で取り上げた問題点解決のため, いくつかの手法について説明する.

図 2.3 は [1] を参考に描いた図である.

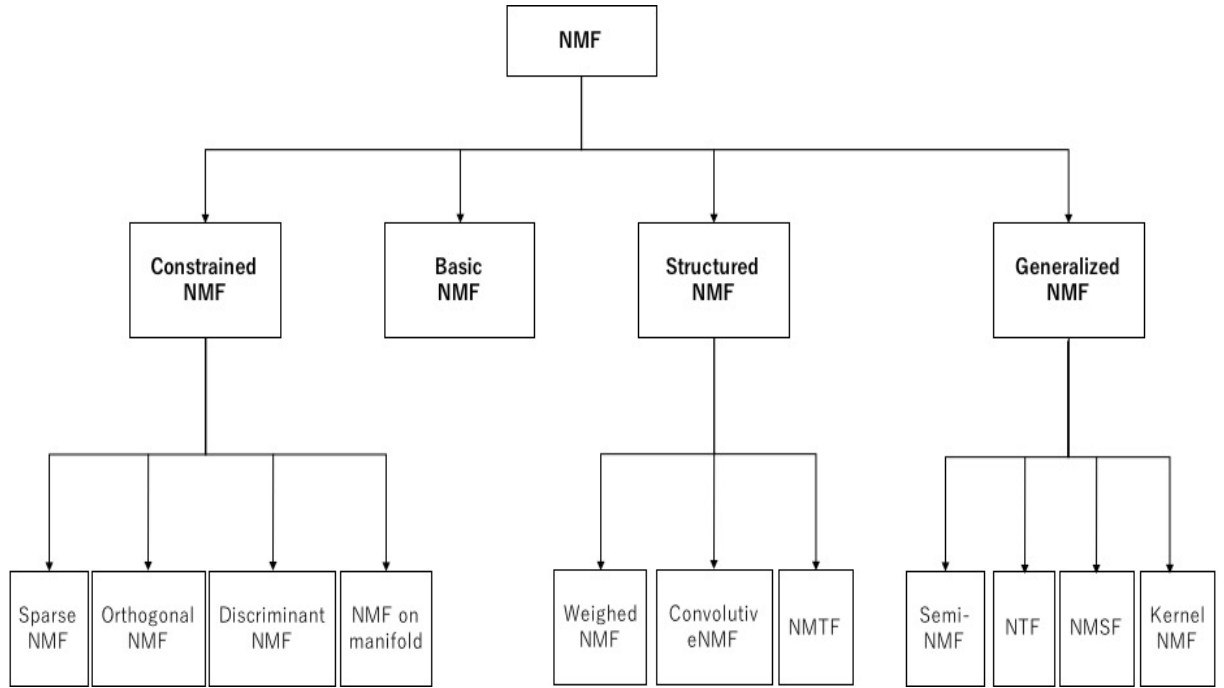


図 2.3: NMF モデルとアルゴリズムのカテゴリ [1]

2.3.1 Orthogonal Nonnegative Matrix Factorization(ONMF)

ONMF[14] は NMF の V に直行制約を付与したものである. NMF より優れた精度結果が得られている手法である. NMF と同様に, 非負値を満たしながら乗法的更新則によって実現される. 目的関数を式 (2.12) に示す. k は近似されるランクである.

- $\mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{U} \in \mathbb{R}^{m \times k}, \mathbf{V} \in \mathbb{R}^{k \times n}$

$$\begin{aligned} \min_{\mathbf{U} \in \mathbb{R}^{m \times k}, \mathbf{V} \in \mathbb{R}^{k \times n}} \quad & \|\mathbf{X} - \mathbf{UV}\|_F^2 \\ \text{subject to} \quad & \mathbf{U} \geq 0, \mathbf{V} \geq 0, \mathbf{V}\mathbf{V}^T = \mathbf{I}_k \end{aligned} \quad (2.12)$$

係数行列 \mathbf{V} に直交制約を与えることで, 必然的に各列に一つだけ非負値要素を有するスパースな行列となる. スパース性がさらに高まるため, 帰属度が集約し, 統計的独立性が高まるため, 行列分解の一意性を向上させる. また, 副次的効果として, 辞書行列 \mathbf{U} の持つ特徴の情報が増える [15]. このような制約は, 辞書行列や係数行列の過学習を防ぐ役割も果たす. 制約を敢えて両辺に与えないことで, 用途に応じた柔軟なクラスタリングが可能となる. そこで, クラスタリングで用いる \mathbf{V} にのみ制約を与えている.

2.3.2 Robust Nonnegative Matrix Factorization(RNMF)

RNMF[16] は高次元かつノイズを扱うことができるNMFである．目的関数を式 (2.13) に示す．

- $\mathbf{X} \in \mathbb{R}^{m \times n}$, $\mathbf{U} \in \mathbb{R}^{m \times k}$, $\mathbf{V} \in \mathbb{R}^{k \times n}$
- $\|\mathbf{A}\|_{2,1} = \sum_{i=1}^n \|a_i\|$ と定義される．

$$\begin{aligned} \min_{\mathbf{U} \in \mathbb{R}^{m \times k}, \mathbf{V} \in \mathbb{R}^{k \times n}} \quad & \|\mathbf{X} - \mathbf{UV}\|_{2,1} \\ \text{subject to} \quad & \mathbf{U} \geq 0, \mathbf{V} \geq 0 \end{aligned} \quad (2.13)$$

フロベニウスノルムは二乗誤差であるために、大きく定まってしまうが、L2,1ノルムを用いることで、その問題を解決することが可能である．そのため、NMFに比べてロバスト性能を持つ [6]．外れ値、ノイズ、データ欠損に影響を受けてしまう問題の解決を図ることが可能な手法である．

2.3.3 Graph Regularized NMF(GNMF)

GNMF [8] は局所的不変性に着目したNMFである．いわば、データ空間の幾何学関係を考慮したNMFである．元の信号の構造は射影後にも保存されるということを、隣接行列と任意の2つのデータ点間の距離関係によって導出した局所的な不変ペナルティ項 ($\text{Tr}(\mathbf{V}\mathbf{L}\mathbf{V}^T)$) で実現している [17]．目的関数を式 (2.14) に示す． \mathbf{W} は隣接行列、 μ は正則化パラメータを表している．隣接行列は元の特徴空間における幾何学的関係の類似性をグラフ表現した行列である．

- $\mathbf{X} \in \mathbb{R}^{m \times n}$, $\mathbf{U} \in \mathbb{R}^{m \times k}$, $\mathbf{V} \in \mathbb{R}^{k \times n}$
- $\mathbf{D}\mathbf{b} = \sum_l \mathbf{W}_{jl}$, $\mathbf{L} = \mathbf{D}\mathbf{b} - \mathbf{W}$, $\mu \geq 0$

$$\begin{aligned} \min_{\mathbf{U} \in \mathbb{R}^{m \times k}, \mathbf{V} \in \mathbb{R}^{k \times n}} \quad & \|\mathbf{X} - \mathbf{UV}\|_F^2 + \mu \text{Tr}(\mathbf{V}\mathbf{L}\mathbf{V}^T) \\ \text{subject to} \quad & \mathbf{U} \geq 0, \mathbf{V} \geq 0 \end{aligned} \quad (2.14)$$

この手法は入力信号空間で近いデータサンプル ($\mathbf{x}_j, \mathbf{x}_l$) は、それ自体に内在する幾何空間において近い距離であるはず．つまり、新しく求める辞書行列 \mathbf{U} により表現される空間において、それらのデータサンプルは

お互いに近い距離に位置するはずであるということ，いわば局所性に着目している．そのために，入力信号空間における2つのデータサンプル $(\mathbf{x}_j, \mathbf{x}_l)$ の距離関係と，求める係数行列 \mathbf{V} 内にそれぞれ対応する係数ベクトル $(\mathbf{z}_j, \mathbf{z}_l)$ との関係に着目し，近くに位置するデータサンプル同士が，異なる係数ベクトルで表現されることに対するペナルティ項を，NMFの目的関数で考慮したものである．ペナルティ項は $\text{Tr}(\mathbf{V}\mathbf{L}\mathbf{V}^T)$ で表される．ペナルティ項について2.3.3に示す．

ペナルティ項の導出

ペナルティ項を考える際，信号の特徴空間における幾何的關係を行列で表現したい．そのために，元の特徴空間における幾何的關係の類似性をグラフに表現した行列である隣接行列とデータ間距離を用いる[18]．

- 隣接行列 \mathbf{W} について

隣接行列は，元の特徴空間における幾何的關係の類似性をグラフに表現した行列である．

$$\mathbf{W}_{jl} = w_{jl} = \begin{cases} \exp\left(-\frac{\|\mathbf{x}_j - \mathbf{x}_l\|}{2\sigma^2}\right) & \mathbf{x}_j \text{ と } \mathbf{x}_l \text{ が隣接している場合} \\ 0 & \text{otherwise} \end{cases} \quad (2.15)$$

\mathbf{W}_{jl} は2つのデータ点 $(\mathbf{x}_j, \mathbf{x}_l)$ 間のデータ間の距離関係を表現している．

各サンプルについて最も類似している K 個に重みを与え，他のサンプルには重みを与えない．これによって重み計算がよりシンプルに考えることができる．

- 係数ベクトル間距離 $d(\cdot, \cdot)$ の定義

係数行列表現を $\mathbf{V} = [\mathbf{z}_1, \dots, \mathbf{z}_j, \dots, \mathbf{z}_n]$ と定義する．新しい基底による2つのデータ点 $(\mathbf{x}_j, \mathbf{x}_l)$ に対応する表現係数ベクトル $(\mathbf{z}_j, \mathbf{z}_l)$ 間の距離 $d(\mathbf{z}_j, \mathbf{z}_l)$ を式 (2.16) に定義する．

$$d(\mathbf{z}_j, \mathbf{z}_l) = \|\mathbf{z}_j - \mathbf{z}_l\|^2 \quad (2.16)$$

- ペナルティ項の考え方

入力信号空間における2つのデータサンプル $(\mathbf{x}_j, \mathbf{x}_l)$ が近い場合 \mathbf{W}_{jl} の値が増加するようなペナルティ項を考える．また，係数ベクト

ル間距離が大きい場合においても $d(z_j, z_l)$ が増加するようなペナルティ項を考える。 \mathbf{W}_{jl} 及び $d(z_j, z_l)$ の値が増加することで、罰則効果も増加するように設定する。

- 局所的不変ペナルティ項を定義

上記のペナルティ項の考え方より、隣接行列 \mathbf{W} と係数ベクトル間距離 $d(z_j, z_l)$ から局所的不変ペナルティ項を定義する。

$$\frac{1}{2} \sum_{j,l=1}^N \|z_j - z_l\|^2 \mathbf{W}_{jl} \quad (2.17)$$

- 局所的不変ペナルティ項の導出

式 (2.17) を変形し、局所的不変ペナルティ項を導出する。ここで $\mathbf{Db} = \sum_l \mathbf{W}_{jl}$, $\mathbf{L} = \mathbf{Db} - \mathbf{W}$ と定義

$$\begin{aligned} \frac{1}{2} \sum_{j,l=1}^N \|z_j - z_l\|^2 \mathbf{W}_{jl} &= \frac{1}{2} \sum_{j,l=1}^N (z_j^T z_j - 2z_j^T z_l + z_l^T z_l) \mathbf{W}_{jl} \\ &= \frac{1}{2} \sum_{j,l=1}^N z_j^T z_j \mathbf{W}_{jl} - \sum_{j,l=1}^N z_j^T z_l \mathbf{W}_{jl} + \frac{1}{2} \sum_{j,l=1}^N z_l^T z_l \mathbf{W}_{jl} \\ &= \sum_{j=1}^N z_j^T z_j \mathbf{Db}_{jj} - \sum_{j,l=1}^N z_j^T z_l \mathbf{W}_{jl} \\ &= \text{Tr}(\mathbf{VDbV}^T) - \text{Tr}(\mathbf{VWV}^T) \\ &= \text{Tr}(\mathbf{VLV}^T) \end{aligned}$$

このペナルティ項は提案手法にも用いられている。ペナルティ項は、クラスタリングの際、任意の隣接距離 N と正則化パラメータ p を設定する必要があり、その最適値は実験測で得ることができる。本稿においても、任意の隣接距離 N と正則化パラメータ p を設定した実験を行なっている。

2.3.4 Orthogonal Nonnegatively Penalized Matrix Factorization(ONP-MF)

Orthogonal Nonnegatively Penalized Matrix Factorization(以後 ONP-MF と記述する) [19] は非負値を罰則ベースに拡張ラグランジュ法 [10] によって最適化目的関数を定義した NMF である。拡張ラグランジュ法は、制約のある

最適化問題の局所最適解を数値的に求めるアルゴリズムである。NMF等とは違い、直交制約を保証したパラメータ更新式を行いながら、直交制約のペナルティにより、漸近的に非負値制約を実現するものである。また、ロバスト性も持つ。初期値をSVDによって与えているため、ランダムで初期値を求めるNMFと比較して、初期値依存性を低下させることが可能である。目的関数を式 (2.18) に示す。

- $\mathbf{X} \in \mathbb{R}^{m \times n}$, $\mathbf{U} \in \mathbb{R}^{m \times k}$, $\mathbf{V} \in \mathbb{R}^{k \times n}$
- $\Lambda \in \mathbb{R}^{k \times n}$, $\rho \geq 0$ (二次のペナルティパラメータ)

$$\begin{aligned}
& \min_{\mathbf{U}, \mathbf{V}, \Lambda} \quad \mathcal{L}_\rho(\mathbf{U}, \mathbf{V}, \Lambda) \\
& \text{where} \quad \mathcal{L}_\rho(\mathbf{U}, \mathbf{V}, \Lambda) = \frac{1}{2} \|\mathbf{X} - \mathbf{UV}\|_F^2 \\
& \quad \quad \quad + \langle \Lambda, -\mathbf{V} \rangle + \frac{\rho}{2} \|\min(\mathbf{V}, 0)\|_F^2 \\
& \text{subject to} \quad \mathbf{U} \geq 0, \mathbf{V}\mathbf{V}^T = \mathbf{I}_k
\end{aligned} \tag{2.18}$$

この手法は乗法的更新則ではなく、勾配法等を用いて収束を行う。値の更新法について2.3.4に示す。

値の更新

ONP-MFでは、辞書行列 \mathbf{U} 、係数行列 \mathbf{V} 、及びラグランジュ乗数 Λ を条件を満たすまで反復することで、値を収束させる。なお、双対定理から、解 (\mathbf{U}, \mathbf{V}) は双対問題の解となることに注意して求める必要がある。

- \mathbf{U} の更新
 \mathbf{V} と Λ を固定し、以下の非負値制約の付いた最小二乗問題を解く。
ここで、 \mathbf{X} , \mathbf{V} , givenな値である。

$$\mathbf{U} \leftarrow \arg \min_{\mathbf{X} \in \mathbb{R}^{m \times n}} \|\mathbf{X} - \mathbf{YV}\|_F^2 \tag{2.19}$$

- \mathbf{V} の更新
勾配射影法 [20] を用いて更新を行なっている。直交制約を保証しながら \mathbf{V} を更新を行う。入力 \mathbf{V} に対して、最急勾配降下法及び直線探索 [19] によって更新し、直交制約保証のため、 $\mathbf{V}\mathbf{V}^T = \mathbf{I}_k$ となる空間

への射影を行い，収束条件が満たされるまで繰り返し計算を行う．
ステップサイズ β の導出に直線探索を用いる．収束条件が満たされるまで，(A)，(B)を繰り返す．

(A) 最急勾配降下法及び直線探索

まず \mathbf{V} を最急降下法によって導出する．その際，ステップサイズを求めるために直線探索を用いる．

・ \mathbf{Z} : 最急降下法によって \mathbf{V} を収束させた値

$$\begin{aligned}\mathbf{Z} &= \mathbf{V} - \beta \frac{\partial \mathcal{L}_\rho}{\partial \mathbf{V}} \\ &= \mathbf{V} - \beta (\mathbf{U}^T \mathbf{X} - \mathbf{U}^T \mathbf{U} \mathbf{V} + \Lambda + \rho \max(-\mathbf{V}, 0))\end{aligned}\quad (2.20)$$

(B) \mathbf{Z} の $\mathbf{V}\mathbf{V}^T = \mathbf{I}_k$ 空間への射影

(A) で求めた \mathbf{Z} を， $\mathbf{V}\mathbf{V}^T = \mathbf{I}_k$ を満たす空間へ射影する． \mathbf{Z} に極分解[11]を用いることで，式(2.21)，(2.22)で表されるように，正規直交基底を取り出すことで，それを達成している．

・ $\hat{\mathbf{V}} : \mathbf{Z}\mathbf{Z}^T = \mathbf{I}_k$ を満たす \mathbf{V} (\mathbf{V} を更新したもの)

$$\hat{\mathbf{V}} = \arg \min_{\mathbf{Y}} \|\mathbf{Z} - \mathbf{Y}\|_F^2 \quad \text{subject to} \quad \mathbf{Y}\mathbf{Y}^T = \mathbf{I}_k \quad (2.21)$$

$$\hat{\mathbf{V}} = \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1/2} \quad (2.22)$$

● ラグランジュ乗数 Λ の更新

関数 $\Lambda \mapsto \max_{\Lambda \geq 0} L_\rho(\mathbf{U}, \mathbf{V}, \Lambda)$ の Λ に対する勾配は $-\mathbf{V}$ であること，また， Λ が非負値であることに注意して，式(2.23)のように更新する．ここで， α はステップサイズである．

$$\Lambda \leftarrow \max(0, \Lambda - \alpha \mathbf{V}) \quad (2.23)$$

ステップサイズ α は， $\alpha = \alpha_0/t$ (t :更新回数)で設定している．

2.4 手法の比較

これまで述べた手法の特徴を表2.2に示す．それぞれの特徴として，非負値制約，直行制約，幾何的制約，L2,1ノルムが挙げられる．ONMF, GNMFの制約は精度向上に繋がり，RNMFとONP-MFはロバスト性能を持つため，それらの制約をNMFに与えることで，ノイズデータ等に対してのクラスタリング精度が向上するのではないかと考えられる．提案手法についての詳細は3章で述べる．

表 2.2: 手法の比較

方式	非負値制約	直交制約	幾何的制約	L2,1ノルム	初期値
NMF	Direct (MU)				random
ONMF	Direct (MU)	✓			random
RNMF	Direct (MU)			✓	random
GNMF	Direct (MU)		✓		random
ONP-MF	Lagrangian	Direct			SVD
提案手法	Lagrangian	Direct (Projection)	✓	✓	SVD

- Direct: 直接的 (強制的) に制御
- MU: Multiplicative Update (乗法的更新則)
- Projection: 射影手法による実現
- Lagrangian: ラグランジュ関数による実現

第3章 提案手法

本研究は、1章で述べたNMFの問題各々に関して、その解決手法の一例が、2章で述べた関連研究である。では、それらの問題を総じて解決するためには、その特徴をNMFに取り入れ、組み合わせることで改善できないかと考えた。そこで、実際に取り入れ、組み合わせたNMFは、ノイズデータに対してクラスタリングを行なった際、従来のロバスト性を持つ手法より質が改善されているかという検討を行ったものである。

3.1 Robust Locally ONP-MF(Rbust-Li-ONP-MF)

目的関数を式 (3.1) に示す。

- $\mathbf{X} \in \mathbb{R}^{m \times n}$, $\mathbf{U} \in \mathbb{R}^{m \times k}$, $\mathbf{V} \in \mathbb{R}^{k \times n}$
- $\Lambda \in \mathbb{R}^{k \times n}$, $\rho \geq 0$ (二次のペナルティパラメータ)
- $\mathbf{D}_{i,i} = \frac{1}{\|\mathbf{x}_i - \mathbf{U}\mathbf{v}_i\|}$
- $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n]$, $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n]$ と定義する。

$$\begin{aligned}
 \min_{\mathbf{U}, \mathbf{V}, \Lambda} \quad & \mathcal{L}_{\rho, \mu}(\mathbf{U}, \mathbf{V}, \Lambda) \\
 \text{where} \quad & \mathcal{L}_{\rho, \mu}(\mathbf{U}, \mathbf{V}, \Lambda) = \frac{1}{2} \|\mathbf{X} - \mathbf{U}\mathbf{V}\|_{2,1} + \langle \Lambda, -\mathbf{V} \rangle \\
 & + \frac{\rho}{2} \|\min(\mathbf{V}, 0)\|_F^2 + \mu \text{Tr}(\mathbf{V}\mathbf{L}\mathbf{V}^T) \\
 \text{subject to} \quad & \mathbf{U} \geq 0, \mathbf{V}\mathbf{V}^T = \mathbf{I}_k
 \end{aligned} \tag{3.1}$$

\mathbf{U} , \mathbf{V} , Λ を反復的に更新することで、式 (3.1) が最小となる値を求める。2.3.4の式を軸に関連研究で述べた手法の特徴を組み合わせている。そのため、値の収束のための更新式は2.3.4で述べたものと似通ったものである。なお、双対定理から、解 (\mathbf{U}, \mathbf{V}) は双対問題の解となることに注意して求める必要のも同様の理由である。

3.2 分解行列とラグランジュ乗数の更新

3.2.1 U更新

Uは勾配射影法[20]によって更新を行う．勾配法を用いて値の収束を行う．

1. \mathbf{A} は最急降下法で求める．ステップサイズ γ は直線探索[19]によって求める．

$$\begin{aligned}\mathbf{A} &= \mathbf{U} - \gamma \frac{\partial \mathcal{L}_{\rho, \mu}}{\partial \mathbf{U}} \\ &= \mathbf{U} - \gamma (\mathbf{U} \mathbf{V} \mathbf{D} \mathbf{V}^T - \mathbf{X} \mathbf{D} \mathbf{V}^T)\end{aligned}\quad (3.2)$$

2. \mathbf{A} を $\mathbf{A} \geq 0$ を満たす空間へ射影する．

$$\hat{\mathbf{U}}_{i,j} = \begin{cases} \mathbf{A}_{i,j} & (\mathbf{A}_{i,j} > 0) \\ 0 & (\mathbf{A}_{i,j} \leq 0) \end{cases}$$

3.2.2 V更新

2.3.4と同様に，勾配射影法[20]によって更新を行う．入力 \mathbf{V} を，最急降下法により更新し，直交制約保証のための $\mathbf{V} \mathbf{V}^T = \mathbf{I}_k$ となる空間への射影を行い，収束条件が満たされるまで繰り返し計算を行う．ステップサイズ β は直線探索[19]によって求める．

1. 最急降下法を用いて最適値 \mathbf{Z} を求める．

$$\begin{aligned}\mathbf{Z} &= \mathbf{V} - \beta \frac{\partial \mathcal{L}_{\rho}}{\partial \mathbf{V}} \\ &= \mathbf{V} - \beta (\mathbf{U}^T \mathbf{U} \mathbf{V} \mathbf{D} - \mathbf{U}^T \mathbf{X} \mathbf{D} + \Lambda + \rho \max(-\mathbf{V}, 0))\end{aligned}\quad (3.4)$$

2. \mathbf{Z} を $\mathbf{V} \mathbf{V}^T = \mathbf{I}_k$ 空間への射影 (\mathbf{Z} , given, $\hat{\mathbf{V}}$ は極分解により導出) する．

$$\hat{\mathbf{V}} = \arg \min_{\mathbf{Y}} \|\mathbf{Z} - \mathbf{Y}\|_F^2 \quad \text{subject to} \quad \mathbf{Y} \mathbf{Y}^T = \mathbf{I}_k \quad (3.5)$$

入力信号 \mathbf{Z} の正規直交基底を取り出す極分解[11]により解 $\hat{\mathbf{V}}$ を計算する．

$$\hat{\mathbf{V}} = \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1/2} \quad (3.6)$$

3.2.3 Λ 更新

ラグランジュ関数 Λ も勾配法を用いて求める．2.3.4と同様に，関数 $\Lambda \mapsto \max_{\Lambda \geq 0} L_\rho(\mathbf{U}, \mathbf{V}, \Lambda)$ に対する勾配は $-\mathbf{V}$ であること，非負値であることに注意して，以下の式で更新する．ステップサイズ α は， $\alpha = \alpha_0/t$ で設定する． t は更新回数である．

$$\Lambda \leftarrow \max(0, \Lambda - \alpha \mathbf{V}) \tag{3.7}$$

第4章 評価実験

1.3で問題解決のため、関連研究で述べた手法の特徴を取り入れた提案法を考えた。従来のものと比較して、ロバスト性及びクラスタリング精度が向上するものになると考えられる。そこで、ノイズを混入したデータに対してのクラスタリングを行い、従来手法との比較実験を行なった。その際以下の2つの実験を行った。

- ノイズ密度の変化
ノイズの密度が変化した場合の性能への影響の確認のため
- 正則化パラメータ μ や隣接距離 p の値の変化
正則化パラメータ μ や隣接距離 p の性能へ与える影響確認のため

4.1 対象データセット

対象データセットは表4.1のものを用いた。

表 4.1: 実験データセット

dataset	type	Instances(size)	Features(dimentionality)	Number of clusses
COIL20	Image	1440	1024	20

- COIL20:20 種類の物体を5度ずつ角度を変えて撮影した画像データセット

4.2 実験条件

本稿の提案手法は、2.3.3と同様に、正則化パラメータ μ 及び隣接距離 p の値によって精度が変化する。そのため、値を変えることによって、どのようにクラスタリング精度が変化するのに着目した実験を行なった。

ロバスト性を確認及びノイズのクラスタリング精度への影響の確認のため、ノイズデータを使用した。本実験で行なったものは以下の3つの実験である。

- 1 正則化パラメータを変えた時のクラスタリング精度
パラメータは{0.01, 0.1, 1}の5つを設定した。
- 2 隣接距離の値を変化させた時のクラスタリング精度
パラメータは{5, 10, 15, 20}の4つを設定した。
- 3 ノイズの精度への影響{0.05, 0.075, 0.1}のレベルのノイズデータを使用して実験を行なった。

比較手法を以下に示す。

- k-means[7]
- NMF 2.2 参照
- GNMF 2.3.3 参照
- RNMF 2.3.2 参照
- ONP-MF 2.3.4 参照

実験のその他条件を以下に示す。

- iteration number:1000
- 提案手法及び比較手法各々5回試行
- 4.3で説明する評価指標の5回平均及びその標準偏差を結果としている
- 初期値は全て同様の値で、SVDの左特異値と右特異値を設定している[21]
- $\alpha_0=1.01$, $\rho=0.01$ と設定

4.3 評価指標

結果の評価には, *purity*, *Accuracy*, *NMI*を指標として用いた. これらの値は大きな値である程良いクラスタリング結果であることを示す指標である.

purity (純度) [22] は式 (4.1) で定義される指標である. この指標は, 正解クラスタのデータを含む割合を示すものである.

- N : データ数, K : 生成クラスタ数
- $x_{i,j}$: 生成クラスタの i 番目に置いて, j 番目の正しいクラスタに属するデータ数

$$purity = \frac{1}{N} \sum_{i=1}^K \max_j(x_{i,j}) \quad (4.1)$$

Accuracy[23] (正確性) は式 (4.2) で定義される指標である. この指標はクラスタの正確性を示すものである.

- r_i : クラスタラベル, d_i : 正解ラベル, N : データ数
- $\delta(x, y)$: if $x = y$ を示す関数, $map(r_i)$: それぞれの正解ラベルとクラスタラベルを対応させる関数

$$Accuracy = \max \frac{\sum_{i=1}^N \delta(map(r_i), d_i)}{N} \quad (4.2)$$

NMI (正規化相互情報量) [24] は式 (4.3) で定義される指標である. この指標はクラスタの類似性を示すものである.

- K : 生成クラスタ集合, T : 正解クラスタの集合
- MI : 相互情報量, H : エントロピー

$$NMI = \frac{MI(K, T)}{\max(H(K), H(T))} \quad (4.3)$$

$H(K)$, $H(T)$ は式 (4.4) で表記される.

- N : 全データ数
- $P(K_i) = \frac{|K_i|}{N}$, $|K_i|$: 生成されたクラスタ K_i に含まれるデータ数

$$H(K) = - \sum_{\substack{i \\ i \in \{1, 2, \dots, k\}}}^k P(K_i) \log P(K_i)$$

また, $MI(K, T)$ は式 (4.4) で定義される.

$$MI(K, T) = H(T) + H(K) - H(K, T) \tag{4.4}$$

第5章 実験結果

精度評価結果を表5.1から表5.12に示す. また, パラメータの組み合わせの違う精度評価図を図5.1から図5.5に示す. 表中の値は評価値の5回平均の値を, \pm は標準偏差を示す. 表は μ は正則化パラメータを, p は隣接距離を示す. また, 正則化パラメータ μ や隣接距離 p の値が影響の無いk-means, NMF, RNMF, ONP-MFは各々の値が変化しても, 評価値等に影響は及ぼさない手法である. また, 最も高い精度を示した結果を太字, 時点で高い精度を示した結果に下線で表す.

精度のみ着目すると, 正則化パラメータ等によって最高精度を示す組み合わせは違うが, 表5.1から表5.12どれも従来手法より高い精度を示す結果となった. これは, ロバスト性と幾何的な情報量を取り入れたことによる精度の向上であると考えられる.

また, 表5.7と表5.11に着目すると, ノイズの密度が増しているにもかかわらず, 精度が増していることが確認できる. これは, データによって, 最適な正則化パラメータと隣接距離の組み合わせによることが原因と考えられる. つまり, 幾何的な構造の情報によって精度が向上していることが確認できる.

よって, ロバスト性及び幾何的な構造の情報の付与によってONP-MFの問題を低減し, 精度の向上につなげることができた結果となった.

表 5.1: クラスタリング実験結果 (p -num=5, noise level=0.05)

method	μ -num	purity	Accuracy	NMI
k-means	—	52.6	47.8	52.6
NMF	—	54.7 \pm 0.9	51.0 \pm 1.2	59.7 \pm 0.6
Robust NMF	—	17.8 \pm 0.3	17.3 \pm 0.2	17.4 \pm 0.1
GNMF	0.01	12.9 \pm 0.08	11.9 \pm 0.1	9.5 \pm 0.1
	0.1	11.0 \pm 0.3	9.7 \pm 0.3	6.7 \pm 0.4
	1	11.1 \pm 0.2	10.3 \pm 0.2	8.4 \pm 0.1
ONP-MF	—	71.5 \pm 1.1	68.5 \pm 1.0	<u>78.5</u> \pm 1.1
Robust-Li-ONP	0.01	69.6 \pm 1.6	68.0 \pm 1.6	77.5 \pm 0.6
	0.1	<u>72.7</u> \pm 0.9	<u>70.7</u> \pm 1.6	78.4 \pm 0.8
	1	72.9 \pm 1.0	71.1 \pm 1.9	78.8 \pm 0.5

表 5.2: クラスタリング実験結果 (p -num=10, noise level=0.05)

method	μ -num	purity	Accuracy	NMI
k-means	—	52.6	47.8	52.6
NMF	—	54.7 \pm 0.9	51.0 \pm 1.2	59.7 \pm 0.6
Robust NMF	—	17.8 \pm 0.3	17.3 \pm 0.2	17.4 \pm 0.1
GNMF	0.01	15.4 \pm 0.7	13.8 \pm 0.7	13.8 \pm 0.9
	0.1	13.5 \pm 0.08	12.6 \pm 0.1	11.5 \pm 0.05
	1	12.0 \pm 0.03	11.2 \pm 0.03	7.9 \pm 0.06
ONP-MF	—	<u>0.72</u> \pm 0.1	<u>0.69</u> \pm 0.09	<u>0.79</u> \pm 0.01
Robust-Li-ONP	0.01	0.71 \pm 0.1	0.68 \pm 0.1	0.78 \pm 0.07
	0.1	0.72 \pm 0.2	0.71 \pm 0.3	0.79 \pm 0.1
	1	0.71 \pm 0.1	0.68 \pm 2.3	0.78 \pm 1.2

表 5.3: クラスタリング実験結果 (p -num=15, noise level=0.05)

method	μ -num	purity	Accuracy	NMI
k-means	—	52.6	47.8	52.6
NMF	—	54.7 \pm 0.9	51.0 \pm 1.2	59.7 \pm 0.6
Robust NMF	—	17.8 \pm 0.3	17.3 \pm 0.2	17.4 \pm 0.1
GNMF	0.01	12.3 \pm 0.3	11.7 \pm 0.5	9.4 \pm 0.2
	0.1	9.2 \pm 0.1	8.3 \pm 0.1	5.2 \pm 0.1
	1	12.9 \pm 1.1	12.5 \pm 1.1	8.6 \pm 0.9
ONP-MF	—	<u>71.5</u> \pm 1.1	<u>68.5</u> \pm 1.0	<u>78.5</u> \pm 1.1
Robust-Li-ONP	0.01	70.7 \pm 2.2	68.3 \pm 2.6	77.4 \pm 1.5
	0.1	71.3 \pm 1.6	67.9 \pm 2.3	78.2 \pm 0.8
	1	72.8 \pm 1.9	70.8 \pm 2.4	79.0 \pm 1.3

表 5.4: クラスタリング実験結果 (p -num=20, noise level=0.05)

method	μ -num	purity	Accuracy	NMI
k-means	—	52.6	47.8	52.6
NMF	—	54.7 \pm 0.9	51.0 \pm 1.2	59.7 \pm 0.6
Robust NMF	—	17.8 \pm 0.3	17.3 \pm 0.2	17.4 \pm 0.1
GNMF	0.01	8.7 \pm 0.5	8.0 \pm 0.5	4.4 \pm 0.5
	0.1	9.5 \pm 0.6	9.0 \pm 0.5	6.3 \pm 0.9
	1	10.3 \pm 1.3	9.7 \pm 1.2	6.1 \pm 1.4
ONP-MF	—	<u>71.5</u> \pm 1.1	68.5 \pm 0.9	78.5 \pm 1.1
Robust-Li-ONP	0.01	71.1 \pm 2.5	<u>69.3</u> \pm 3.8	78.0 \pm 1.2
	0.1	72.4 \pm 2.1	70.3 \pm 3.1	78.4 \pm 1.0
	1	70.9 \pm 1.7	69.2 \pm 1.9	78.4 \pm 0.7

表 5.5: クラスタリング実験結果 (p -num=5, noise level=0.075)

method	μ -num	purity	Accuracy	NMI
k-means	—	51.9	49.0	51.9
NMF	—	53.6 \pm 0.7	50.8 \pm 1.1	59.8 \pm 0.5
Robust NMF	—	16.4 \pm 0.5	16.2 \pm 0.5	15.2 \pm 0.8
GNMF	0.01	11.4 \pm 0.4	10.3 \pm 0.4	8.6 \pm 0.7
	0.1	7.7 \pm 0.5	7.1 \pm 0.5	3.7 \pm 0.7
	1	10.8 \pm 0.6	9.2 \pm 0.6	7.4 \pm 0.6
ONP-MF	—	70.7\pm1.7	68.8\pm2.4	77.5 \pm 1.0
Robust-Li-ONP	0.01	69.5 \pm 1.6	68.0 \pm 2.6	77.7\pm0.7
	0.1	69.3 \pm 1.6	66.5 \pm 1.2	77.7\pm0.7
	1	<u>70.5\pm2.7</u>	<u>68.4\pm3.2</u>	77.7\pm1.3

表 5.6: クラスタリング実験結果 (p -num=10, noise level=0.075)

method	μ -num	purity	Accuracy	NMI
k-means	—	51.9	49.0	51.9
NMF	—	53.6 \pm 0.7	50.8 \pm 1.1	59.8 \pm 0.5
Robust NMF	—	16.4 \pm 0.5	16.2 \pm 0.5	15.2 \pm 0.8
GNMF	0.01	14.6 \pm 0.2	14.1 \pm 0.2	13.8 \pm 0.2
	0.1	7.2 \pm 0.2	6.5 \pm 0.8	2.8 \pm 0.2
	1	9.3 \pm 0.07	8.5 \pm 0.05	5.1 \pm 0.07
ONP-MF	—	70.7 \pm 1.7	68.8 \pm 2.4	77.5 \pm 1.0
Robust-Li-ONP	0.01	70.2 \pm 1.9	67.5 \pm 1.6	78.4\pm0.5
	0.1	<u>70.8\pm1.4</u>	69.9\pm2.2	<u>78.3\pm0.7</u>
	1	70.9\pm2.6	<u>69.3\pm2.1</u>	78.0 \pm 0.6

表 5.7: クラスタリング実験結果 (p -num=15, noise level=0.075)

method	μ -num	purity	Accuracy	NMI
k-means	—	51.9	49.0	51.9
NMF	—	53.6 \pm 0.7	50.8 \pm 1.1	59.8 \pm 0.5
Robust NMF	—	16.4 \pm 0.5	16.2 \pm 0.5	15.2 \pm 0.8
GNMF	0.01	9.2 \pm 0.4	8.5 \pm 0.3	5.7 \pm 0.7
	0.1	10.8 \pm 0.4	9.9 \pm 0.4	7.1 \pm 0.5
	1	9.0 \pm 0.6	8.2 \pm 0.5	4.3 \pm 0.5
ONP-MF	—	70.7 \pm 1.7	68.8 \pm 2.4	77.5 \pm 1.0
Robust-Li-ONP	0.01	<u>71.8</u> \pm 1.7	69.9 \pm 1.9	78.7 \pm 0.9
	0.1	72.0 \pm 1.6	<u>69.6</u> \pm 1.5	78.7 \pm 0.8
	1	70.4 \pm 1.2	68.5 \pm 1.4	77.1 \pm 0.9

表 5.8: クラスタリング実験結果 (p -num=20, noise level=0.075)

method	μ -num	purity	Accuracy	NMI
k-means	—	51.9	49.0	51.9
NMF	—	53.6 \pm 0.7	50.8 \pm 1.1	59.8 \pm 0.5
Robust NMF	—	16.4 \pm 0.5	16.2 \pm 0.5	15.2 \pm 0.8
GNMF	0.01	8.9 \pm 0.04	8.4 \pm 0.07	4.8 \pm 0.09
	0.1	14.3 \pm 0.04	13.4 \pm 0.04	11.2 \pm 0.01
	1	12.0 \pm 0.03	11.4 \pm 0.03	06.5 \pm 0.2
ONP-MF	—	70.7 \pm 1.7	68.8 \pm 2.4	77.5 \pm 1.0
Robust-Li-ONP	0.01	<u>72.1</u> \pm 1.8	<u>69.6</u> \pm 2.6	<u>78.5</u> \pm 0.8
	0.1	70.8 \pm 2.4	68.7 \pm 2.8	78.3 \pm 1.0
	1	73.0 \pm 0.4	72.1 \pm 0.6	79.5 \pm 0.3

表 5.9: クラスタリング実験結果 (p -num=5, noise level=0.1)

method	μ -num	purity	Accuracy	NMI
k-means	—	35.8	32.9	52.7
NMF	—	52.5 \pm 1.1	49.9 \pm 1.1	59.2 \pm 0.8
Robust NMF	—	11.1 \pm 0.6	10.8 \pm 0.6	7.7 \pm 0.4
GNMF	0.01	13.8 \pm 0.5	13.4 \pm 0.5	9.2 \pm 1.3
	0.1	10.2 \pm 1.0	9.6 \pm 0.7	7.1 \pm 1.6
	1	8.4 \pm 0.9	7.8 \pm 0.8	4.7 \pm 1.2
ONP-MF	—	70.2 \pm 1.9	68.3 \pm 1.9	77.7 \pm 0.8
Robust-Li-ONP	0.01	71.9\pm1.9	<u>70.0\pm2.6</u>	78.4\pm1.3
	0.1	71.5 \pm 1.1	69.7 \pm 1.7	<u>78.3\pm0.9</u>
	1	71.9\pm1.3	70.3\pm1.6	<u>78.3\pm0.9</u>

表 5.10: クラスタリング実験結果 (p -num=10, noise level=0.1)

method	μ -num	purity	Accuracy	NMI
k-means	—	35.8	32.9	52.7
NMF	—	52.5 \pm 1.1	49.9 \pm 1.1	59.2 \pm 0.8
Robust NMF	—	11.1 \pm 0.6	10.8 \pm 0.6	7.7 \pm 0.4
GNMF	0.01	9.3 \pm 0.4	7.9 \pm 0.4	5.2 \pm 0.4
	0.1	12.6 \pm 0.1	11.5 \pm 0.1	10.0 \pm 0.4
	1	11.6 \pm 0.06	11.1 \pm 0.06	8.0 \pm 0.1
ONP-MF	—	70.2 \pm 1.9	68.3 \pm 1.9	77.7 \pm 0.8
Robust-Li-ONP	0.01	70.6 \pm 2.1	68.6\pm2.5	<u>77.9\pm1.3</u>
	0.1	70.8\pm1.4	<u>68.4\pm1.1</u>	78.0\pm0.8
	1	<u>70.7\pm1.8</u>	68.0 \pm 0.3	<u>77.9\pm0.4</u>

表 5.11: クラスタリング実験結果 (p -num=15, noise level=0.1)

method	μ -num	purity	Accuracy	NMI
k-means	—	35.8	32.9	52.7
NMF	—	52.5 \pm 1.1	49.9 \pm 1.1	59.2 \pm 0.8
Robust NMF	—	11.1 \pm 0.6	10.8 \pm 0.6	7.7 \pm 0.4
GNMF	0.01	13.4 \pm 0.4	11.8 \pm 0.5	10.2 \pm 0.5
	0.1	10.3 \pm 0.2	9.2 \pm 0.2	6.6 \pm 0.2
	1	11.7 \pm 0.2	10.8 \pm 0.2	7.0 \pm 0.1
ONP-MF	—	70.2 \pm 1.9	68.3 \pm 1.9	77.7 \pm 0.8
Robust-Li-ONP	0.01	72.2 \pm 1.4	70.0 \pm 2.2	78.7\pm0.7
	0.1	72.6\pm1.8	70.4\pm3.3	78.7\pm0.7
	1	70.0 \pm 2.2	68.1 \pm 2.5	77.5 \pm 0.8

表 5.12: クラスタリング実験結果 (p -num=20, noise level=0.1)

method	μ -num	purity	Accuracy	NMI
k-means	—	35.8	32.9	52.7
NMF	—	52.5 \pm 1.1	49.9 \pm 1.1	59.2 \pm 0.8
Robust NMF	—	11.1 \pm 0.6	10.8 \pm 0.6	7.7 \pm 0.4
GNMF	0.01	13.1 \pm 0.6	12.3 \pm 0.5	14.1 \pm 0.4
	0.1	10.9 \pm 0.3	10.5 \pm 0.3	7.1 \pm 0.6
	1	9.6 \pm 1.4	9.3 \pm 1.4	5.3 \pm 1.3
ONP-MF	—	70.2 \pm 1.9	68.3 \pm 1.9	77.7 \pm 0.8
Robust-Li-ONP	0.01	71.4 \pm 2.1	69.3 \pm 2.2	78.1 \pm 0.9
	0.1	72.7\pm2.0	71.5\pm2.3	78.9\pm0.9
	1	72.2 \pm 2.1	70.6 \pm 3.1	78.6 \pm 1.2

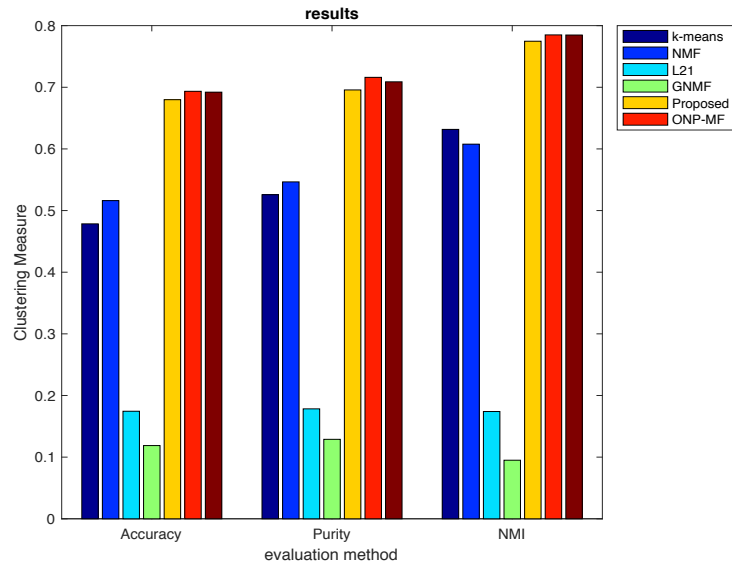


図 5.1: 精度評価図 ($\mu=0.01, p=5, \text{noise}=0.05$)

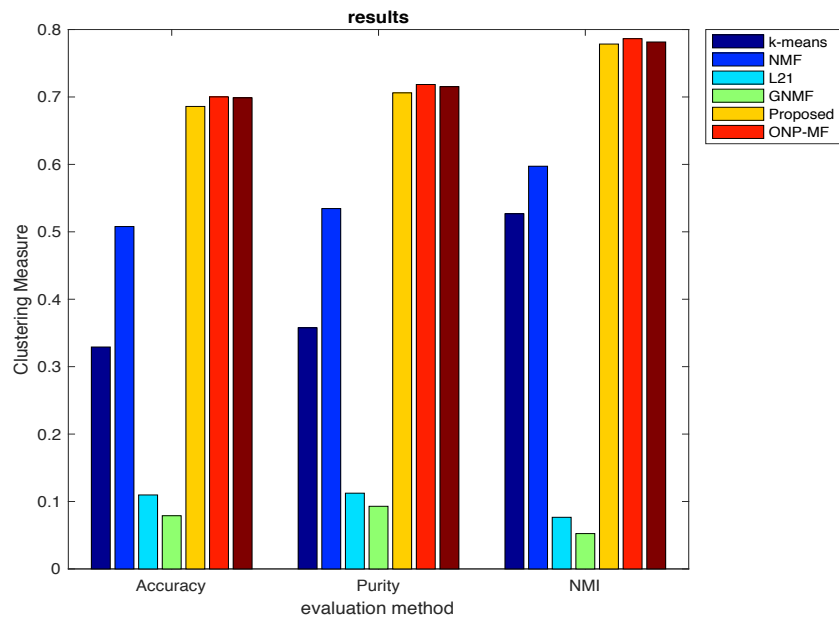


図 5.2: 精度評価図 ($\mu=0.01, p=10, \text{noise}=0.1$)

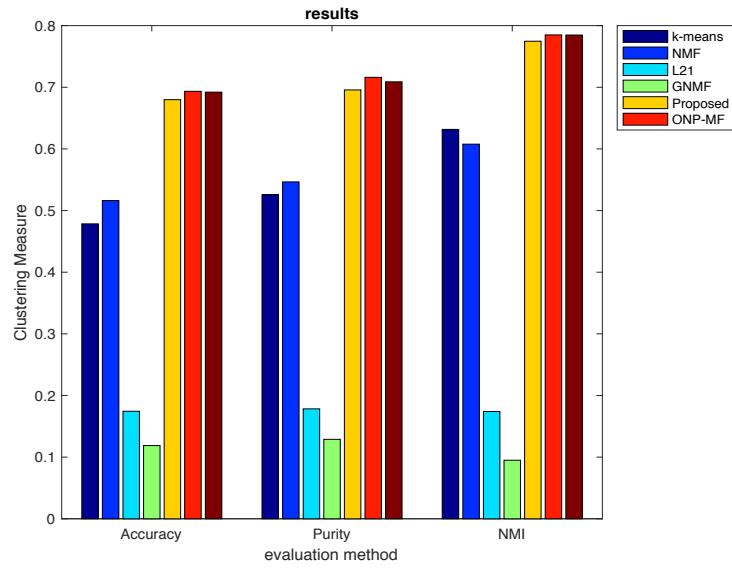


図 5.3: 精度評価図 ($\mu=0.01, p=5, \text{noise}=0.05$)

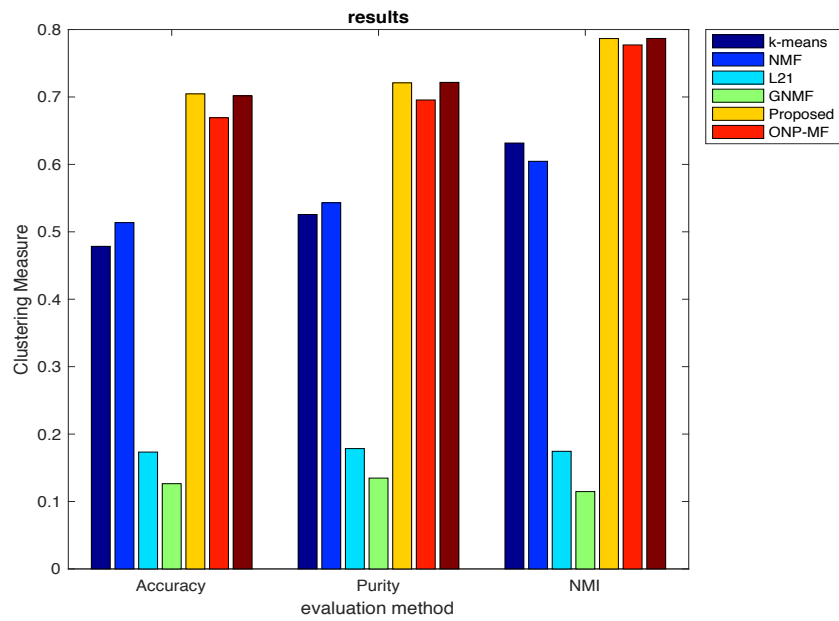


図 5.4: 精度評価図 ($\mu=0.1, p=10, \text{noise}=0.05$)

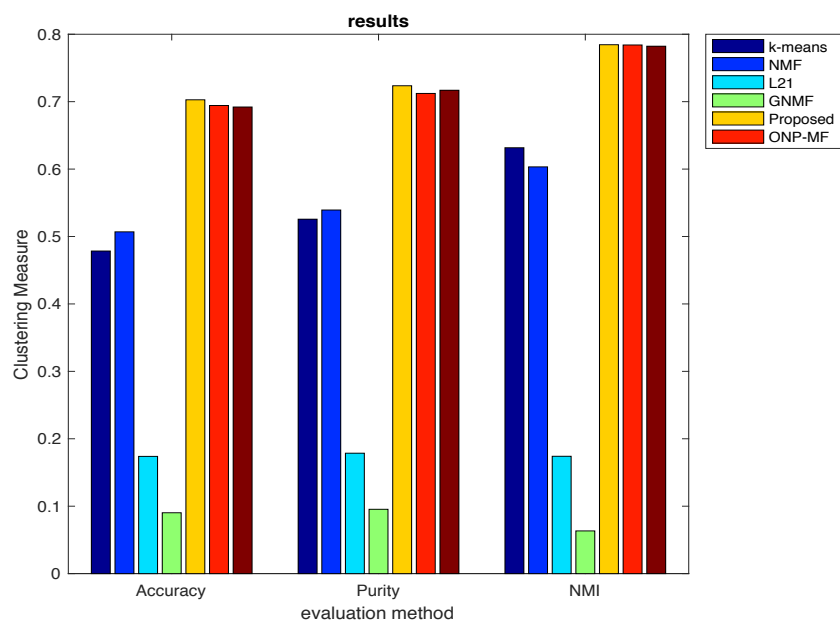


図 5.5: 精度評価図 ($\mu=0.1, p=20, \text{noise}=0.05$)

第6章 結論

本研究では，NMFを拡張したONP-MFの問題解決のため，ロバスト性と幾何学的情報を手法に付加したRobust-Li-ONP-MFの目的関数及び値の更新についての提案を行った．

結果として，多手法と比較して，僅かだが高いロバスト性が確認できた．また，評価指標から，幾何学的構造による精度の向上が確認することができた．

今後として，多種のデータに関しても同様の実験を行い，どのデータに関しても同様の性能が得られるかの確認を行う必要がある．また，任意のパラメータに関してもより細かく設定し，適切なパラメータはどの範囲に集約するか等の検討を行う．

謝辞

本研究を実施するにあたって，ご指導をいただきました笠井裕之先生に心から感謝申し上げます．また，ゼミや講座内発表等でご助言，ご指導をいただきました，森田啓義先生，眞田亜紀子先生，応用ネットワーキング講座の様々な面におきまして，助けていただきました片桐さんに心から感謝申し上げます．そして，有意義な研究活動を行うことができたのは共に支え合い，助け合うことができた同講座内の仲間がいたからです．彼らにも，感謝の意を示します．ありがとうございました．

References

- [1] Wang Yu-Xiong and Zhang Yu-Jin. Nonnegative matrix factorization: A comprehensive review. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, 25(6), June 2013.
- [2] Daniel D Lee and H. Sebastian Seung. Algorithms for non-negative matrix factorization. *Advances in neural information processing systems (NIPS)*, pages 556–562, 2001.
- [3] Zhang Zhong-Yuan. Nonnegative matrix factorization: Models, algorithms and applications. *Data Mining: Foundations and Intelligent Paradigms*, 24:99–134, 2012.
- [4] C. Boutsidis and E. Gallopoulos. Svd based initialization: A head start for non-negative matrix factorization. *Pattern Recogn.*, 41(4):1350–1362, April 2008.
- [5] *Robust Capped Norm Nonnegative Matrix Factorization*, October 2015.
- [6] Kong Deguang, Ding Chris, and Huang Heng. Robust nonnegative matrix factorization using l21-norm. *Association for Computing Machinery(ACM)*, pages 673–682, 2011.
- [7] Seungjin Choi. Algorithms for orthogonal nonnegative matrix factorization. In *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence)*. IEEE, June 2008.
- [8] Cai Deng, He Xiaofei, Han Jiawei, and S. Huang Thomas. Graph regularized nonnegative matrix factorization for data representation. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 33(8):Nonnegative matrix factorization, graph Laplacian, manifold regularization, clustering, August August 2011.
- [9] 尾形 わかは. 数理線形計画法. オーム社, November 2010.

- [10] Nocedal Jorge and J. Wright Stephen. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag New York, 2006.
- [11] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, NY, USA, 2nd edition, 2012.
- [12] M. Chu, F. Diele, R. Plemmons, and S. Rangni. Optimality, computation, and interpretation of nonnegative matrix factorizations. *JOURNAL ON MATRIX ANALYSIS*, October 2004.
- [13] N. Del Buono and G. Pio. Non-negative matrix tri-factorization for co-clustering: An analysis of the block matrix. *nformation Sciences*, 301(C):13–26, April 2015.
- [14] Ding Chris, Li Tao, Peng Wei, and Park Haesun. Orthogonal nonnegative matrix tri-factorizations for clustering. *In SIGKDD*, pages 126–135, August 2006.
- [15] 亀岡 弘和. 非負値行列因子分解. In *the Society of Instrument and Control Engineers*, volume 51, pages 835–844. コロナ社, 9 2012.
- [16] Yang Shizhun, HouChangshui Chenping, and Wu ZhangYi. Robust non-negative matrix factorization via joint sparse and graph regularization. *Neural Computing and Applications*, 23:541–559, August 2013.
- [17] Li Yeqing, Huang Junzhou, and Liu Wei. Scale sequential spectral clustering. In *AAAI’16 Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pages 1809–1815, February 2016.
- [18] Mikhail Belkin and Niyogi Partha. Laplacian eigenmaps and spectral techniques for embedding and clustering. In T. G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems 14*, pages 585–591. MIT Press, 2002.
- [19] Pompili Filippo, Gillis Nicolas, Absil P.-A, and Glineur Francois. Two algorithms for orthogonal nonnegative matrix factorization with application to clustering. *Neurocomputing*, (141):15–25, April 2014.

- [20] Wang Enyuan, Zhu Zhen, Chenc Wei, and Xiaod Pengcheng. Manifold nmf with l21 norm for clustering. *Neurocomputing*, 2017.
- [21] Pompili Filippo, Gillis Nicolas, Absil P.-A, and Glineur Francois. ONP-MF: An orthogonal nonnegative matrix factrization algorithm with application to clustering. In *Computational Intelligence and Machine Learning (ESANN 2013)*, April 2013.
- [22] Peng Chong, Kang Zhao, Hu Yunhong, Cheng Jie, and Cheng Qiang. Robust graph regularized nonnegative matrix factorization for clustering. *ACM Transactions on Knowledge Discovery from Data*, 11(33), April 2017.
- [23] Ma Huifang, Zhao Weizhong, Tan Qing, and Shi Zhongzhi. Orthogonal non-negative matrix tri-factorization for semi-supervised document co-clustering. *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 189–200, 2010.
- [24] Cheng Hao, A.Hua Kien, and Vu Khanh. Constrained locally weighted clustering. *Constrained Locally Weighted Clustering*, Volume 1:Pages 90–101, August 2008.