数理解析研究所講究録
第 1981 巻 2016 年 45-50

非負行列分解に対する前処理付き逐次射影法のノイズ頑強性

水谷 友彦 *
東京工業大学 経営工学専攻

Tomohiko Mizutani
Department of Industrial Engineering and Management
Tokyo Institute of Technology

概 要

本稿では [8] で得られた結果を紹介する．逐次射影法は分離可能性を仮定した下での非負行列分解を求めるための有効な数値手法であることが知られている．この行列分解は文章データのクラスタリングやトピック抽出などいくつかの応用を持つ．このような応用を考える際，手法はノイズに対して頑強であることが望まれる．Gillis-Vavasis は逐次射影法はノイズに対して頑強であることを理論的に示した．更に，前処理行列を利用するとその頑強性が向上することを示した．しかし，彼らは次元 d と分離ランク r が一致するという条件の下で，前処理付き逐次射影法のノイズ頑強性を評価している．応用を考える上ではこの条件は不自然である．本研究では $d=r$ という条件を課さずに前処理付き逐次射影法のノイズ頑強性を評価する．

1 はじめに

逐次射影法は分離可能性を仮定した下での非負行列分解を求めるための数値手法である．本稿ではそのノイズ頑強性について考える．以降では非負行列分解を NMF と記述することにする．実行列でその要素が全て非負であるような行列を**非負行列**と呼び，大きさが $d \times m$ の非負行列の集合を $\mathbb{R}_+^{d\times m}$ と書くことにする．$\boldsymbol{A} \in \mathbb{R}_+^{d\times m}$ が次のように分解できるとき，$\boldsymbol{A}$ は**分離可能な行列**であると定める．

$$\boldsymbol{A} = \boldsymbol{F}\boldsymbol{W} \text{ for } \boldsymbol{F} \in \mathbb{R}_+^{d\times r} \text{ and } \boldsymbol{W} = (\boldsymbol{I}, \boldsymbol{K})\boldsymbol{\Pi} \in \mathbb{R}_+^{r\times m} \tag{1}$$

ここで，$\boldsymbol{I}$ は $r \times r$ 単位行列，$\boldsymbol{K}$ は $r \times (m-r)$ 非負行列，$\boldsymbol{\Pi}$ は $m \times m$ 置換行列を表している．特に，$\boldsymbol{F}$ を**基底行列**，その列ベクトル $\boldsymbol{f}_i$ を**基底**，d を**次元**，r を**分解ランク**と呼ぶことにする．定義から，基底 $\boldsymbol{f}_1, \dots, \boldsymbol{f}_r$ は $\boldsymbol{A}$ の列ベクトルの集合の中に含まれていることが分かる．分離可能性を仮定した NMF とは次のような問題である．

問題 1 (Separable NMF). 分離可能な行列 $\boldsymbol{A}$ から基底行列 $\boldsymbol{F}$ に対応する $\boldsymbol{A}$ の列ベクトルを見つけよ．

これは分離可能な行列 $\boldsymbol{A}$ が与えられたとき $\boldsymbol{A}(\mathcal{I}) = \boldsymbol{F}$ となる添字集合 $\mathcal{I}$ を見つけるという問題と言い換えることができる．ここで $\boldsymbol{A}(\mathcal{I})$ とは $\boldsymbol{A}$ の部分行列で列ベクトル $\boldsymbol{a}_i$ の添字 i が $\mathcal{I}$ に含まれるもので構成される行列である．つまり，$\boldsymbol{A}(\mathcal{I}) = (\boldsymbol{a}_i : i \in \mathcal{I})$ である．問題 1 を Separable NMF と記述することにする．Separable NMF は NMF の特殊な場合となっている．NMF とは $\boldsymbol{A} \in \mathbb{R}_+^{d\times m}$ と整数 r が与えられたとき，$||\boldsymbol{A} - \boldsymbol{F}\boldsymbol{W}||_2$ が最小となるなるような $\boldsymbol{F} \in \mathbb{R}_+^{d\times r}$ と $\boldsymbol{W} \in \mathbb{R}_+^{r\times m}$ を求めるという問題として定式化できる．この最小化問題は NP 困難であることが知られている [9]．この困難さに対して，Arora らは [3] の中で NMF に $\boldsymbol{A}$ が分離可能な行列であるという仮定を置くと扱いやすい問題となることを指摘した．その問題が Separable NMF に対応する．Separable

*連絡先: mizutani.t.ab@m.titech.ac.jp

NMF は文章データのクラスタリングやトピック抽出 [4, 2, 7]，ハイパースペクトラル画像からの特徴画像の抽出 [5, 6] に応用できることが知られている．

Separable NMF な幾何的な解釈を持つ．$\boldsymbol{D}_1, \boldsymbol{D}_2$ を正則な対角行列とすると，分離可能な行列 $\boldsymbol{A}$ は $\boldsymbol{A} = \boldsymbol{F}\boldsymbol{W} \Leftrightarrow \boldsymbol{A}\boldsymbol{D}_1 = \boldsymbol{F}\boldsymbol{D}_2\boldsymbol{D}_2^{-1}\boldsymbol{W}\boldsymbol{D}_1$ となる．このことから一般性を失わずに $\boldsymbol{W}$ の列ベクトル $\boldsymbol{w}_i$ の 1 ノルムは 1 であると仮定できる．つまり，$\boldsymbol{A}$ の列ベクトル $\boldsymbol{a}_i$ は基底 $\boldsymbol{f}_1, \dots, \boldsymbol{f}_r$ の凸結合で与えられることを意味する．$\boldsymbol{f}_1, \dots, \boldsymbol{f}_r$ は線形独立であると仮定しよう．すると，$\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ の凸包は d 次元空間上の $(r-1)$ 次元単体となり，その r 個の頂点が $\boldsymbol{f}_1, \dots, \boldsymbol{f}_r$ に対応する．したがって，Separable NMF は $\boldsymbol{A}$ の列ベクトル $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ の凸包の頂点を全て見つける問題と解釈できる．

これまでに Sepearable NMF に対して幾つかのアルゴリズムが提案されている．Separable NMF は扱いやすい問題で多項式時間で解くことができる．具体的には入力として (1) の $\boldsymbol{A}$ とその分解ランク r が与えられたとき，$\boldsymbol{F} = \boldsymbol{A}(\mathcal{I})$ となる $\mathcal{I}$ を出力するような多項式時間アルゴリズムが幾つか存在する．Separable NMF の実問題への応用を考える．この場合，分離可能な行列はノイズを含むという状況を考える必要がある．(1) の $\boldsymbol{A}$ と $\boldsymbol{N} \in \mathbb{R}^{d \times m}$ に対して，

$$\widetilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{N} \in \mathbb{R}^{d \times m} \tag{2}$$

を**ノイズを含む分離可能な行列**であると定める．また $\boldsymbol{N}$ を**ノイズ行列**と呼ぶことにする．Separable NMF に対するアルゴリズムを考える．(2) の $\widetilde{\boldsymbol{A}}$ と分解ランク r が入力されたとき $\boldsymbol{F}$ と $\widetilde{\boldsymbol{A}}(\mathcal{I})$ の差が大きくならないような $\mathcal{I}$ を出力するようなアルゴリズムを**ノイズに対して頑強**であるという．

逐次射影法は計量化学の文脈で提案された手法である [1]．Gillis-Vavasis は [5] において Separable NMF に対して逐次射影法は $\boldsymbol{F} = \boldsymbol{A}(\mathcal{I})$ となる添字集合 $\mathcal{I}$ を出力することを示した．さらにノイズに対して頑強であることを示した．彼らの理論解析は何らかの手段で $\boldsymbol{F}$ の条件数を小さくすることができると，逐次射影法のノイズ頑強性が向上することを示唆している．そこで彼らは [6] の中で前処理行列を利用することで，逐次射影法のノイズ頑強性が向上するということを理論的に解析した．しかし，その解析では (1) の $\boldsymbol{A}$ において次元 d と分解ランク r が一致するという条件の下で行われている．この条件は Separable NMF の実問題への応用を考えた場合，不自然である．例えば Separable NMF を文章データからのトピック抽出に応用することを考えよう．この場合，d は文章の個数，r はトピックの個数に対応する．したがって，d は r より大きいとするのが自然である．本研究では $d = r$ という条件を仮定せずに前処理付き逐次射影法のノイズ頑強性を理論解析を行った．その結果を定理 2 にまとめる．

2 逐次射影法のノイズ頑強性

逐次射影法のアルゴリズムは Separable NMF の幾何的な解釈に基づいて設計されている．$\boldsymbol{A} \in \mathbb{R}^{d \times m}$ と正数 r が入力されたとき逐次射影法は以下のように動作する．$\boldsymbol{A}$ の列ベクトル $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ の中からベクトルの長さが最大となっている列ベクトル $\boldsymbol{a}_{i^*}$ を選択する．次に $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ を $\boldsymbol{a}_{i^*}$ の直交補空間に射影する．この操作を r 本の列ベクトルが選択されるまで繰り返す．逐次射影法の各ステップをアルゴリズム 1 にまとめる．

多面体の内部の点と頂点から構成される集合 $\mathcal{S}$ と凸関数 f を考える．このとき，$\mathcal{S}$ の要素の中で f の関数値が最大となるものは頂点に対応することが知られている．この事実を踏まえると，$\boldsymbol{A}$ が分離可能な行列な場合，逐次射影法の出力は基底行列 $\boldsymbol{F}$ を与えるということは直感的には理解できる．次のような仮定を用意する．

仮定 1. $\boldsymbol{A} \in \mathbb{R}^{d \times m}$ は $\boldsymbol{A} = \boldsymbol{F}\boldsymbol{W}$ と分解できる．$\boldsymbol{F}$ と $\boldsymbol{W}$ はそれぞれ $\boldsymbol{F} \in \mathbb{R}^{d \times r}$ と $\boldsymbol{W} = (\boldsymbol{I}, \boldsymbol{K})\boldsymbol{\Pi} \in \mathbb{R}_+^{r \times m}$ であり，$\boldsymbol{I}$，$\boldsymbol{K}$，$\boldsymbol{\Pi}$ は *(1)* におけるものと同じである．$\boldsymbol{F}$ と $\boldsymbol{K}$ の列ベクトル $\boldsymbol{k}_i$ は次の条件を満たす．

Algorithm 1 逐次射影法

入力: $\boldsymbol{A} \in \mathbb{R}^{d \times m}$ と正数 r
出力: 添字集合 $\mathcal{I}$

1: 行列 $\boldsymbol{S}$ と添字集合 $\mathcal{I}$ を用意し，$\boldsymbol{S} \leftarrow \boldsymbol{A}$，$\mathcal{I} \leftarrow \emptyset$ と初期化する．

2: $\boldsymbol{S}$ の列ベクトル $\boldsymbol{s}_i$ に対して，$i^* = \arg\max_{i=1,\dots,m} ||\boldsymbol{s}_i||_2^2$ となる添字 i^* を見つける．

3: ベクトル $\boldsymbol{t}$ を用意して $\boldsymbol{t} \leftarrow \boldsymbol{s}_{i^*}$ と設定する．そして，行列 $\boldsymbol{S}$ と添字集合 $\mathcal{I}$ を次のように更新する．

$$\begin{aligned} \boldsymbol{S} &\leftarrow \left(\boldsymbol{I} - \frac{\boldsymbol{t}\boldsymbol{t}^\top}{||\boldsymbol{t}||_2^2}\right)\boldsymbol{S} \\ \mathcal{I} &\leftarrow \mathcal{I} \cup \{i^*\} \end{aligned}$$

4: もし $|\mathcal{I}| < r$ ならばステップ 2 に戻り，そうでなければ $\mathcal{I}$ を出力する．

(a) $rank(\boldsymbol{F}) = r$

(b) $\boldsymbol{e}^\top \boldsymbol{k}_i \le 1$ *for* $i = 1, \dots, m - r$

ここで $\boldsymbol{e}$ は全ての要素が *1* となっている d 次元ベクトルを表している．

仮定 1 における $\boldsymbol{A}$ の分解は厳密には (1) における分解と一致していない．つまり，(1) における $\boldsymbol{F}$ は非負行列となる必要があるが，一方で，仮定 1 における $\boldsymbol{F}$ は必ずしも非負行列である必要はない．1 節で言及したように，$\boldsymbol{D}_1, \boldsymbol{D}_2$ を正則な対角行列とすると $\boldsymbol{A} = \boldsymbol{F}\boldsymbol{W} \Leftrightarrow \boldsymbol{A}\boldsymbol{D}_1 = \boldsymbol{F}\boldsymbol{D}_2\boldsymbol{D}_2^{-1}\boldsymbol{W}\boldsymbol{D}_1$ となる．したがって，仮定 1(b) は一般性を失わずに仮定できる．Gillis-Vavasis[5] は逐次射影法の入力行列 $\boldsymbol{A}$ が仮定 1 を満たすとき出力 $\mathcal{I}$ は $\boldsymbol{F} = \boldsymbol{A}(\mathcal{I})$ となることを示し，さらに，ノイズに対して頑強であることを示した．逐次射影法のノイズ頑強性に関する彼らの結果は次のように記述される．

定理 1 ([5] の定理 3)**.** $\boldsymbol{A} \in \mathbb{R}^{d \times m}$ と $\boldsymbol{N} \in \mathbb{R}^{d \times m}$ に対して $\widetilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{N}$ とする．$r \ge 2$ で $\boldsymbol{A}$ は仮定 *1* を満たすとする．もし $\boldsymbol{N}$ の列ベクトル $\boldsymbol{n}_i$ は $||\boldsymbol{n}_i||_2 \le \epsilon,\ i = 1, \dots, m$ で ϵ は

$$\epsilon < \min\left(\frac{1}{2\sqrt{r-1}}, \frac{1}{4}\right)\frac{\sigma_{min}(\boldsymbol{F})}{1 + 80\kappa(\boldsymbol{F})^2}$$

となるならば，アルゴリズム *1* に $(\widetilde{\boldsymbol{A}}, r)$ を入力すると出力 $\mathcal{I}$ はその要素の順番を適切に並べ替えると

$$||\widetilde{\boldsymbol{a}}_{\mathcal{I}(j)} - \boldsymbol{f}_j||_2 \le (1 + 80\kappa(\boldsymbol{F})^2)\epsilon,\ j = 1, \dots, r$$

を満たす．

$\mathcal{I}(j)$ は添字集合 $\mathcal{I}$ の要素をある順番で並べたときの j 番目の要素を表している．$\boldsymbol{f}_j$ は $\boldsymbol{F}$ の j 番目の列ベクトルを表している．$\sigma_{\min}(\boldsymbol{F})$ は $\boldsymbol{F}$ の最小特異値，$\kappa(\boldsymbol{F})$ は $\boldsymbol{F}$ の条件数を表している．

3 前処理付き逐次射影法のノイズ頑強性

定理 1 は逐次射影法のノイズ頑強性を更に改善するための示唆を与えている．(2) の $\widetilde{\boldsymbol{A}}$ を考える．このとき，$\widetilde{\boldsymbol{A}}$ における $\boldsymbol{F}$ に前処理行列 $\boldsymbol{Q}$ を施すことで $\boldsymbol{F}$ の条件数を 1 に近づけることができると，頑強性を保証できるノイズ許容範囲 ϵ を拡大し，かつ基底誤差 $||\widetilde{\boldsymbol{a}}_{\mathcal{I}(j)} - \boldsymbol{f}_j||$ に掛かっている係数を小

さくすることができるかもしれない．$Q \in \mathbb{R}^{d\times d}$ を正則行列とする．すると，$Q\widetilde{A} = QFW + QN$ となる．つまり，$Q\widetilde{A}$ はノイズを含む分離可能な行列として見ることができ，QF はその分離可能な行列，QN はノイズ行列に対応する．A が仮定 1 を満たすとき，Q は正則なので QF もこの仮定を満たすことが分かる．Q として F の条件数をできるだけ小さくするようなものを選択し，逐次射影法に $\widetilde{A}$ の代わりに $Q\widetilde{A}$ を入力すると，アルゴリズムのノイズ頑強性が向上することが期待できる．ただし，その新たなノイズ行列 QN は $||QN|| \le ||Q||||N||$ となるのでそのサイズは拡大するかもしれない．

3.1 $d = r$ の場合

Gilli-Vavasis は [6] の中で上述のような前処理行列 Q を構築するための手順を示している．簡単のためにアルゴリズム 1 の入力行列は分離可能な行列でノイズを含まないとする．つまり，入力行列は (1) の A とする．A は仮定 1 と $d = r$ を満たすと仮定する．この仮定の下では A の大きさは $r \times m$ で，$r \times r$ 基底行列 F を含むことになる．A の列ベクトル $a_1, \dots, a_m$ を集めて集合 $\mathcal{S} = \{a_1, \dots, a_m\}$ を作成し，次のような最適化問題を考える．

$$\begin{array}{lll} \mathsf{P}(\mathcal{S}): & \text{minimize} & -\log\det(L), \\ & \text{subject to} & a^\top L a \le 1 \text{ for all } a \in \mathcal{S}, \\ & & L \succ 0. \end{array}$$

$L \succ 0$ は L が正定値対称行列であることを表している．この決定変数は L である．$\mathsf{P}(\mathcal{S})$ は $\mathcal{S}$ の要素に対する体積最小閉包楕円を求めるための問題として知られている．また，これは凸計画問題となっており，効率的なアルゴリズムが存在する．$\mathsf{P}(\mathcal{S})$ の最適解 L^* は $L^* = (FF^\top)^{-1}$ となることが知られている [7, 6]．したがって，$\kappa((L^*)^{1/2}F)^2 = \kappa(F^\top L^* F) = \kappa(I) = 1$ となるので，$(L^*)^{1/2}$ は F の条件数を小さくするための前処理行列として利用できる．

次にアルゴリズム 1 の入力行列は分離可能な行列でノイズを含むとする．つまり，入力行列は (2) の $\widetilde{A}$ とする．$\widetilde{A}$ の列ベクトル $\widetilde{a}_1, \dots, \widetilde{a}_m$ を用いて $\mathcal{S} = \{\widetilde{a}_1, \dots, \widetilde{a}_m\}$ を作成する．この $P(\mathcal{S})$ の最適解 L^* は $(FF^\top)^{-1}$ とは一致しない．しかし，ノイズ行列 N のサイズが小さければ，L^* と $(FF^\top)^{-1}$ の差は小さいことが期待できる．このような場合，$(L^*)^{1/2}$ は F の条件数を小さくするため前処理行列として利用できるだろう．実際，Gillis-Vavasis は定理 2.9 においてこの前処理行列は逐次射影法のノイズ頑強性を向上させることを示している．彼らの結果の概要は以下のように記述できる．$A \in \mathbb{R}^{d\times m}$ と $N \in \mathbb{R}^{d\times m}$ に対して $\widetilde{A} = A + N$ とする．A は仮定 1 と $d = r$ を満たすと仮定する．上記のような $(L^*)^{1/2}$ に対してアルゴリズム 1 に $(L^*)^{1/2}\widetilde{A}$ と r を入力すると，N のサイズがある値より小さいとき，その出力 $\mathcal{I}$ に対して F と $\widetilde{A}(\mathcal{I})$ の差はある値より小さくなる．

3.2 $d \neq r$ の場合

Gillis-Vavasis[6] では $d \neq r$ における前処理行列の構築手順を示している．(2) の $\widetilde{A}$ を考える．彼らは，まず特異値分解 (SVD) を利用して $\widetilde{A}$ の次元 d を r まで削減し，その後，3.1 節で説明した手順を適用することを提案している．$\widetilde{A}$ の SVD を計算し $\widetilde{A} = U\Sigma V^\top$ と分解する．Σ は $d \times m$ 対角行列で，その対角要素 $\sigma_1, \dots, \sigma_d$ は非負となっている．U と V の列の順番を適切に選ぶことで常に $\sigma_1 \ge \cdots \ge \sigma_d \ge 0$ とすることできる．$\sigma_1, \dots, \sigma_r$ と V の列ベクトル $v_1, \dots, v_r$ に対して

$$P = \mathrm{diag}(\sigma_1, \dots, \sigma_r) V^r \in \mathbb{R}^{r\times m}$$

を構築する．ここで $V^r = (v_1, \dots, v_r)$ である．この P に対して 3.1 節で説明した手順を適用する．P の列ベクトル $p_1, \dots, p_m$ に対して $\mathcal{S} = \{p_1, \dots, p_m\}$ を作成し，$\mathsf{P}(\mathcal{S})$ の最適解 L^* を求める．そして，$(L^*)^{1/2}P$ に対して逐次射影法を適用する．アルゴリズム 2 に前処理付き逐次射影法の各ステップをまとめる．

Algorithm 2 前処理付き逐次射影法

入力: $\boldsymbol{A} \in \mathbb{R}^{d \times m}$ と正数 r
出力: 添字集合 $\mathcal{I}$

1: $\boldsymbol{A}$ の SVD を計算し $\boldsymbol{A} = \boldsymbol{U\Sigma V}^\top$ と分解する．$\boldsymbol{\Sigma}$ の対角要素の中で大きいものから順番に $\sigma_1, \dots, \sigma_r$ を選択する．また，$\boldsymbol{V}$ の列ベクトル $\boldsymbol{v}_1, \dots, \boldsymbol{v}_r$ に対して $\boldsymbol{V}^r = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_r)$ とする．これらから $\boldsymbol{P} = \mathrm{diag}(\sigma_1, \dots, \sigma_r)(\boldsymbol{V}^r)^\top$ を構築する．

2: $\boldsymbol{P}$ の列ベクトル $\boldsymbol{p}_1, \dots, \boldsymbol{p}_m$ から集合 $\mathcal{S} = \{\boldsymbol{p}_1, \dots, \boldsymbol{p}_m\}$ を作成し，$\mathsf{P}(\mathcal{S})$ の最適解 $\boldsymbol{L}^*$ を計算する．

3: 入力 $((\boldsymbol{L}^*)^{1/2}\boldsymbol{P}, r)$ に対してアルゴリズム 1 を実行し出力 $\mathcal{I}$ を求める．

3.3 本研究の結果

アルゴリズム 2 は [6] の中のアルゴリズム 2 とほぼ同じである．その論文ではアルゴリズム 2 はノイズに対して頑強であるという実験結果を報告しているが，理論的な解析は行われていない．本研究ではその解析を行った．以下が結果である．

定理 2. $\boldsymbol{A} \in \mathbb{R}^{d \times m}$ と $\boldsymbol{N} \in \mathbb{R}^{d \times m}$ に対して $\widetilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{N}$ とする．$r \ge 2$ で $\boldsymbol{A}$ は仮定 *1* を満たすとする．もし $||\boldsymbol{N}||_2 = \epsilon$ で ϵ は

$$\epsilon \le \frac{\sigma_{min}(\boldsymbol{F})}{1225\sqrt{r}},$$

となるならば，アルゴリズム *1* に $(\widetilde{\boldsymbol{A}}, r)$ を入力すると出力 $\mathcal{I}$ はその要素の順番を適切に並べ替えると

$$||\widetilde{\boldsymbol{a}}_{\mathcal{I}(j)} - \boldsymbol{f}_j||_2 \le (432\kappa(\boldsymbol{F}) + 4)\epsilon, \ j = 1, \dots, r$$

となる．

証明は [8] の 3 節にまとめてある．Gillis-Vavasis の [6] における定理 2.9 では $d = r$ という条件の下で前処理付き逐次射影法のノイズ頑強性を評価してるが，本結果は $d = r$ という条件を課さずにそれを評価している．1 節で述べたように，文章データからのトピック抽出など実際の応用から生じる Separable NMF では d と r が近いということは稀で，d と r は異なるとするのが自然である．したがって，本結果は応用から生じる Separable NMF に対する前処理付き逐次射影法のノイズ頑強性を考察する上でなんらかの助けになることが期待できる．

Gillis-Vavasis による定理 1 では，許容できるノイズ範囲を $||\boldsymbol{n}_i||_2$ で評価している．それにならって定理 2 のノイズ許容範囲を $||\boldsymbol{n}_i||_2$ で評価してみよう．一般に $d \times m$ 行列 $\boldsymbol{N}$ は $||\boldsymbol{N}||_2 \le \sqrt{m}\max_{i=1,\dots,m}||\boldsymbol{n}_i||_2$ となることを考慮すると以下の結果が得られる．

系 1. $\widetilde{\boldsymbol{A}}$ は定理 *2* と同じものとする．もし $||\boldsymbol{n}_i||_2 \le \epsilon$ で ϵ は

$$\epsilon \le \frac{\sigma_{min}(\boldsymbol{F})}{1225\sqrt{rm}},$$

となるならば，アルゴリズム *1* に $(\widetilde{\boldsymbol{A}}, r)$ を入力すると出力 $\mathcal{I}$ はその要素の順番を適切に並べ替えると

$$||\widetilde{\boldsymbol{a}}_{\mathcal{I}(j)} - \boldsymbol{f}_j||_2 \le (432\kappa(\boldsymbol{F}) + 4)\epsilon, \ j = 1, \dots, r$$

となる．

定理 2 と比較すると，許容できるノイズ範囲の評価において $1/\sqrt{m}$ という項が出現することが分かる．応用から生じる Separable NMF では m はデータの個数に対応し，その個数は多くなることがある．そのような場合，定理 2 でノイズに対する頑強性を保証できるノイズサイズの許容範囲は狭くなる．

参考文献

[1] U. M. C. Araújo, B. T. C. Saldanha, R. K. H. Galvão, T. Yoneyama, H. C. Chame, and V. Visani. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems*, 57(2):65–73, 2001.

[2] S. Arora, R. Ge, Y. Halpern, D. Mimno, and A. Moitra. A practical algorithm for topic modeling with provable guarantees. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, 2013.

[3] S. Arora, R. Ge, R. Kannan, and A. Moitra. Computing a nonnegative matrix factorization – Provably. In *Proceedings of the 44th symposium on Theory of Computing (STOC)*, pages 145–162, 2012.

[4] S. Arora, R. Ge, and A. Moitra. Learning topic models – Going beyond SVD. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1–10, 2012.

[5] N. Gillis and S. A. Vavasis. Fast and robust recursive algorithms for separable nonnegative matrix factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(4):698–714, 2014.

[6] N. Gillis and S. A. Vavasis. Semidefinite programming based preconditioning for more robust near-separable nonnegative matrix factorization. *SIAM Journal on Optimization*, 25(1):677–698, 2015.

[7] T. Mizutani. Ellipsoidal rounding for nonnegative matrix factorization under noisy separability. *Journal of Machine Learning Research*, 15:1011–1039, 2014.

[8] T. Mizutani. Robustness analysis of preconditioned successive projection algorithm for general form of separable nmf problem. arXiv:1506.08387, 2015.

[9] S. A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal of Optimization*, 20(3):1364–1377, 2009.