修士論文
主成分およびマイナー成分抽出に対する正規直交化ペナルティ法

北陸先端科学技術大学院大学
知識科学研究科 知識システム基礎学専攻
950081 松久保 潤

平成 13 年 2 月

審査委員 ： 林　幸雄　助教授 (主査)
櫻井　彰人　教授
中森　義輝　教授
本多　卓也　教授

Copyright © 2001 by Jun Matsukubo

目 次

1 はじめに **1**
1.1 背景と目的 1
1.2 論文の構成 3

2 従来の手法 **5**
2.1 Oja の部分空間抽出アルゴリズム 5
2.2 Oja の主成分抽出アルゴリズム 8
2.3 Wang の双勾配アプローチ 10
2.4 Chen の部分空間アルゴリズム 12

3 提案手法 **13**

4 数値実験 **16**
4.1 収束特性と利用可能なパラメータの範囲 16
4.1.1 主成分抽出に対する実験結果 18
4.1.2 マイナー成分抽出に対する実験結果 24
4.1.3 行列 D に対する比較 29
4.2 グラフの隣接行列に対する適用例 55

5 おわりに **62**

A 定理 1 の証明 **63**

B 定理 2 の証明 **70**

1　はじめに

1.1　背景と目的

多変量解析の手段の一つである主成分分析 (Principal Component Anarysis; PCA) は，信号処理および画像処理におけるデータ圧縮，フィルタリング，等の多くの応用がある．PCA がデータ共分散行列の固有値の大きいものに対応する固有ベクトルから順に求めるのに対して，固有値の小さなものに対応する固有ベクトルから順に求めるマイナー成分分析 (Minor Component Anarysis; MCA) がある．マイナー成分分析は信号処理および画像処理における雑音除去や振動解析における低振動数成分の抽出等の応用に対して近年特に注目されている．PCA と MCA は大規模な行列を対象とすることが多い．大規模な疎行列の固有値問題を解くために，数値解析の分野において Jacobi 法，QR 法，ベキ乗法等の反復解法が提案されている．

一方，ニューラル・ネットワーク等の力学系のダイナミクスに基づいた手法が Matsuoka and Kawamoto[94] と，Oja[92a][92b]，Sangar[89] ，Luo[97]，Douglas[98]，Wang and Kalhunen[96]，Chen[97] 等により提案されている．Matsuoka らのアプローチはネットワーク構造や結合重みの値のパラメータを変更することにより主成分抽出，直交射影演算，マイナー成分抽出という三つの異なる機能を実現している．これに対して，Matsuoka ら以外の手法は力学系の方程式において正負の符号の違いのみで主成分とマイナー成分を抽出するため，より簡便で，かつ双勾配や正規直交化など，数学的な意味においても明確である．ただし，いずれも Oja[92a] の部分空間抽出アルゴリズムを基礎としている点においては共通している．

アルゴリズム名	アルゴリズム	特徴
Oja[92a]	$W_{k+1} = W_k + \alpha_k(CW_k - W_k W_k^T C W_k)$	数値的に不安定
Chen[97]	$W_{k+1} = W_k \pm \alpha_k(CW_k - W_k W_k^T C W_k)$ $+\gamma W_k(I - W^T W)$	正規直交化を強制する ペナルティ項
Douglas[98]	$W_{k+1} = W_k - \alpha_k C W_k (W_k^T W_k)(W_k^T W_k)$ $+W_k W_k^T C W_k$	行列の積

表 1: 数値安定化の手法

アルゴリズム名	アルゴリズム	特徴
Sanger[89]	$W_{k+1} = W_k + \alpha_k(CW_k - \mathrm{LT}[W_k W_k^T C W_k])$	下三角成分の抽出
Oja[92b]	$W_{k+1} = W_k + \alpha_k(CW_k - W_k D^{-1} W_k^T C W_k D)$	行列の積
Wang[96]	$W_{k+1} = W_k \pm \alpha_k C W_k +$ $\gamma_k W_k \times \mathrm{UT}[I - W_k^T W_k]$	上三角成分の抽出 ($\gamma_k \to \infty$ が条件)

表 2: アルゴリズムの非対称化の手法

さて，アルゴリズムの説明の前にいくつか記号を定義しておく．各記号の説明は以下の通りである．

W = $[w_1 ... w_N] \in \mathbf{R}^{M \times N}$: 解ベクトルからなる行列
x = $[x_1 ... x_M] \in \mathbf{R}^M$: データからなるベクトル
C = $E\{x^T x\}$: データ共分散行列
D = $\mathrm{diag}[\theta_1, ..., theta_N]$: 対角行列
$\alpha, \gamma, \theta_i, (i = 1, ..., N) \in \mathbf{R}$: アルゴリズムのパラメータ
$\mathrm{upper}[A]$: 行列 A の上三角成分

まず，表 1 に最も基本となる Oja の部分空間アルゴリズムと数値安定化を施したアルゴリズムを示す．Oja の部分空間アルゴリズム [92a] は，解ベクトルに対する $W^T W = I$ という条件の下で主成分が張る部分空間の基底を抽出する．理論上はパラメータの符号を負にすることで，マイナー成分が張る部分空間を抽出することが可能であるが，実際には数値的不安定性のために困難である．この問題を改善するための数値安定化手法が Chen[97]，Douglas ら [98] によって提案されている．Chen[97] のアルゴリズムはペナルティ項により数値的不安定性を抑制している．一

方，Douglas ら [98] のアルゴリズムは $(W_k^T W_k)(W_k^T W_k)$ の積により数値的不安定性を抑制している．

以上，表 1 にある数値安定化の手法は $W^T W = I$ を条件としているアルゴリズムであり，W の各列ベクトル $w_n, (n = 1, ..., N)$ が行列 C の固有ベクトルに収束することが保証されていない．

次に，部分空間ではなく，各々の固有ベクトルを抽出する手法を表 2 に示す．これらの手法のうち，Sanger[89] と Wang ら [96] はアルゴリズム中で三角成分を抽出することで，Gram-Schmidt の直交化に対応するような非対称化を行っている．一方，Oja の主成分抽出アルゴリズム [92b] は，行列 D の積が Gram-Schmidt の直交化に対応している．Oja[92a] と同様に，この手法は理論上はパラメータの符号を負にすることでマイナー成分を抽出することが可能であるが，実際には数値的不安定性のために困難である．

従って，マイナー成分抽出のための数値安定化と固有ベクトルを抽出するための非対称化は $W^T W = I$ の条件のために，相反する性質をもっているため，単純に組み合わせることができない．

本論文では，Oja，Wang，Chen のアプローチに基づき正負の符号の違いのみで主成分とマイナー成分の固有ベクトルを同様に抽出するとともに，数値的に安定なアルゴリズムを提案する．また提案するアルゴリズムに対して主成分およびマイナー成分の抽出に利用可能なパラメータの範囲を数値的に示すことを目的とする．

1.2 論文の構成

次節から，以下のことについて述べる．

第 2 章では，本研究の基礎となっている Oja の部分空間アルゴリズム，Oja の主成分抽出アルゴリズム，Wang らの双勾配系アプローチ，Chen の部分空間アルゴリズムの特徴について順に説明する．

第 3 章では，修正 Oja アルゴリズムを提案する．このアルゴリズムは，パラメータの符号の正負によって主成分およびマイナー成分固有ベクトルの抽出を選択することができる．また，この章で平衡点近傍での収束特性と利用可能なパラメータ範囲に関する定理を与える．

第 4 章では，提案されるアルゴリズムの収束特性を検証するために行った数値実験の結果を示す．また，これまで類似したアルゴリズムに対して利用可能なパラメータの範囲を系統的に調べられた研究は少ない．し

かしながら，本研究では提案するアルゴリズムに対して数値実験によってその範囲を検証した．

さらに，一つの適用例として，Zipf の法則に従うリンク構造をもつネットワークの隣接行列に対して，提案したアルゴリズムを用いて，主成分とマイナー成分のそれぞれに対するリンク構造の関係について議論する．

最後に第 5 章でまとめと今後の課題を述べる．

2　従来の手法

この章では，2.1 節で Oja の部分空間アルゴリズム [92a] について説明する．この手法は

1. 主成分およびマイナー成分の固有ベクトルへ常微分方程式の解が収束することが保証されていない

2. マイナー成分が張る部分空間の抽出は数値的に不安定である

という問題点をもっている．

2.2 節で，1．の問題点を改善する手法として，Oja により主成分固有ベクトルを抽出できるように改良された主成分抽出アルゴリズム [92b] について述べる．

2.3 節で，2．の問題点の解決法として解の正規直交化を強制するペナルティ項を用いた Wang らの双勾配系アルゴリズム [96]，主成分およびマイナー成分が張る部分空間の基底を抽出する Chen のアルゴリズム [97] について述べる．

2.1　Oja の部分空間抽出アルゴリズム

まず，最も基本となっている Oja の部分空間抽出アルゴリズムについて説明する．

与えられたデータからなるベクトル $x = [x_1, ..., x_K]^T \in \mathbf{R}^K$ と，変数ベクトル $w_n, (n = 1, ..., N) \in \mathbf{R}^K$ からなる $(K \times N)$ 行列 $W = [w_1 ... w_N] \in \mathbf{R}^{K \times N}$ について，

$$\begin{aligned} \Phi_1(W) &\stackrel{\text{def}}{=} \mathrm{trace}(x^T W)^T (x^T W) && (1) \\ &= \sum_{n=1}^{N} \Big\{ (x^T w_n)^T (x^T w_n) \Big\} && (2) \end{aligned}$$

を最大化することを考える．x の共分散行列を $C = E\{xx^T\}$ と表すと，式 (1) は

$$\begin{aligned} J_1(W) &\stackrel{\text{def}}{=} \mathrm{trace}(W^T C W) && (3) \\ &= \sum_{i=1}^{N} (w_i^T C w_i) && (4) \end{aligned}$$

となる．ここで，E は入力データに対する期待値を表す．式 (3) を W について微分すると，

$$\frac{\partial}{\partial W}J_1(W) = CW \tag{5}$$

となる．

まず，$N = 1$ のとき，式 (4) の $J_1(w)$ を最大化する $w \in \mathbf{R}^K$ が行列 C の第一主成分となることを示す．

まず，実対称行列 C に対して，Reyleigh 商

$$R(w) \stackrel{\text{def}}{=} \frac{w^T C w}{w^T w} \tag{6}$$

の最大値は行列 C の最大の固有値 λ_1 となり，$R(w)$ を最大化する w は λ_1 に対応する固有ベクトル，すなわち行列 C の第一主成分である．ここで，$w^T w = 1$ であるなら，式 (6) は

$$R(w) = w^T C w$$

となる．しかしながら，$N > 1$ の場合，すなわち $W = [w_1 ... w_N]$ の場合にそれぞれの $w_n (n = 1, ..., N)$ に対して，$R(w_i)$ の最大化を試みるとすべての w_i が第一主成分に向かってしまうので，各 w_i の直交化が必要となる．

Oja は $w_n^T w_n = 1 (n = 1, ..., N)$ という条件の下に式 (5) を用いて，$J_1(W)$ を最大化する勾配系

$$\tilde{W}_{k+1} = W_k + \alpha_k C W_k \tag{7}$$

$$W_{k+1} = \tilde{W}_{k+1} S_{k+1}^{-1} \tag{8}$$

を提案した．ここで，α_k はスカラーのパラメータであり，S_{k+1}^{-1} は $\tilde{W}_{k+1}$ のそれぞれの列ベクトルを正規直交化する行列である．

Oja は S_{k+1} を対称に選ぶことで $\tilde{W}_{k+1}$ を正規直交化する手法を提案した．

このとき，S_{k+1} は

$$S_{k+1} \stackrel{\text{def}}{=} \{\tilde{W}_{k+1}^T \tilde{W}_{k+1}\} \tag{9}$$

のように選ばれる．

ここで，α_k が十分小さく，

$$W_k^T W_k = I \qquad (I : \text{単位行列})$$

が成立すると仮定すると，S_{k+1}^{-1} は以下のように展開される．

$$\begin{aligned} S_{k+1}^{-1} &= (\tilde{W}_{k+1}^T \tilde{W}_{k+1})^{-\frac{1}{2}} \\ &= [(W_{k-1}^T + \alpha_k W_{k-1}^T C) \times (W_{k-1} + \alpha_k C W_{k-1})]^{-\frac{1}{2}} \\ &= [W_{k-1}^T \cdot W_{k-1} + W_{k-1}^T \cdot \alpha_k C W_{k-1} \\ &\qquad +\alpha_k W_{k-1}^T C \cdot W_{k-1} + O(\alpha_k^2)]^{-\frac{1}{2}} \\ &= [I + 2\alpha_k W_{k-1}^T C W_{k-1} + O(\alpha_k^2)]^{-\frac{1}{2}} \end{aligned}$$

有理化すると，

$$\begin{aligned} S_{k+1}^{-1} &= [I + O(\alpha_k^2)]^{-\frac{1}{2}} [I - 2\alpha_k W_{k-1}^T C W_{k-1} + O(\alpha_k^2)]^{\frac{1}{2}} \\ &= [(I - \alpha_k W_{k-1}^T C W_{k-1} + O(\alpha_k^2))^2]^{-\frac{1}{2}} \\ &= I - \alpha_k W_{k-1}^T C W_{k-1} + O(\alpha_k^2) \end{aligned} \tag{10}$$

となる．式 (10) を用いて，式 (7) と式 (8) を整理すると，

$$W_{k+1} = W_k + \alpha_k (C W_k - W_k W_k^T C W_k) \tag{11}$$

となる．

次に，このアルゴリズムの時間 t に関する常微分方程式の形式

$$\frac{dW}{dt} = CW - WW^T CW \tag{12}$$

が $W(0)^T W(0) = I$ の場合に，$W(t)^T W(t) = I (k = 1, 2, ...)$ を満すことを示す．

$$\begin{aligned} \frac{dW^T W}{dt} &= \frac{dW^T}{dt} W + W^T \frac{dW}{dt} \\ &= (W^T C - W^T C W W^T) W + W^T (CW - WW^T CW) \\ &= (W^T CW - W^T CWW^T W) + (W^T CW - W^T WW^T CW) \\ &= (W^T CW - W^T CW) + (W^T CW - W^T CW) \\ &= \mathbf{0} \end{aligned}$$

ここで，$\mathbf{0}$ は零行列である．

このアルゴリズムには以下の二つの問題点がある．第 1 番目は，$J_1(W)$ の最大化のみを目的としているため，W の各列ベクトル w_n が主成分である各々の固有ベクトルに収束することが保証されてない点である．

この点に関して，Oja は条件 $W^TW = I$ を崩すようなアルゴリズムの非対称化が必要であると述べている．非対称化の手法は Gram-Schmidt の正規直交化に対応する Sanger[89]，Wang ら [96] による三角成分の抽出や，Oja[92b] による対角行列の積による手法が提案されている．

第 2 番目は，式 (11) は理論的には，$J_1(W)$ を最小化するようにパラメータ α の符号を負に変更した場合の常微分方程式

$$\frac{dW}{dt} = -CW + WW^TCW$$

においても条件 $W^TW = I$ を満たすので，マイナー成分が張る部分空間の基底を抽出することが可能であるが，実際には数値的不安定性により，困難である．Chen[97]，Douglas ら [98] によって，Oja のアルゴリズム (11) の数値安定化手法が提案されているが，それらの安定化手法では $W^TW = I$ が成立することが条件となっている．そのため，マイナー成分が張る部分空間の基底を抽出することができるが，$W^TW = I$ の条件を崩す非対称化の操作と組み合わせてマイナー成分である固有ベクトルを抽出することは困難である．

2.2 Oja の主成分抽出アルゴリズム

Oja の部分空間抽出アルゴリズムに，非対称化の操作を施して主成分固有ベクトルを抽出するアルゴリズムを説明する．

Oja[92b] は (8) において，S_{k+1} を

$$S^*_{k+1} = \{D^{-1}\tilde{W}^T_{k+1}\tilde{W}_{k+1}D\} \tag{13}$$

のように選ぶことで，行列 D によるアルゴリズムの非対称化により，主成分が張る部分空間の基底を抽出するアルゴリズムを提案している．

ただし，D は

$$D = \mathrm{diag}[\theta_1, ..., \theta_N]$$
$$\theta_n \in \mathbf{R}, (n = 1, ..., N)$$

と表される対角行列である．ここで，α_k が十分小さく，

$$W^T_kW_k = I \qquad (I : \text{単位行列})$$

が成立すると仮定すると，S_{k+1}^{*-1} は以下のように展開される．

$$\begin{aligned}
S_{k+1}^{*-1} &= (D^{-1}\tilde{W}_{k+1}^T\tilde{W}_{k+1}D) \\
&= [D^{-1}(W_{k-1}^T + \alpha_k W_{k-1}^T C) \times (W_{k-1} + \alpha_k C W_{k-1})D]^{-\frac{1}{2}} \\
&= [(D^{-1}W_{k-1}^T + \alpha_k D^{-1}W_{k-1}^T C) \times (W_{k-1}D + \alpha_k C W_{k-1}D)]^{-\frac{1}{2}} \\
&= [D^{-1}W_{k-1}^T \cdot W_{k-1}D + D^{-1}W_{k-1}^T \cdot \alpha_k C W_{k-1}D \\
&\qquad +\alpha_k D^{-1}W_{k-1}^T C \cdot W_{k-1}D + O(\alpha_k^2)]^{-\frac{1}{2}} \\
&= [I + 2\alpha_k D^{-1}W_{k-1}^T C W_{k-1}D + O(\alpha_k^2)]^{-\frac{1}{2}}
\end{aligned}$$

有理化すると，

$$\begin{aligned}
S_{k+1}^{*-1} &= [I + O(\alpha_k^2)]^{-\frac{1}{2}}[I - 2\alpha_k D^{-1}W_{k-1}^T C W_{k-1}D + O(\alpha_k^2)]^{\frac{1}{2}} \\
&= [(I - \alpha_k D^{-1}W_{k-1}^T C W_{k-1}D + O(\alpha_k^2))^2]^{-\frac{1}{2}} \\
&= I - \alpha_k D^{-1}W_{k-1}^T C W_{k-1}D + O(\alpha_k^2) \qquad (14)
\end{aligned}$$

となる．式 (14) を用いて，式 (7) と式 (8) を整理すると，

$$W_{k+1} = W_k + \alpha_k(CW_k - W_k D^{-1}W_k^T C W_k D) \qquad (15)$$

となる．行列 D のパラメータが

$$0 < \theta_1 < \theta_2 < ... < \theta_N \qquad (16)$$

の場合に，

$$W_k \in \mathbf{R}^{K\times N}$$

は，

$$V = [v_1 ... v_N]$$

に収束する．ただし，$v_n, (n = 1, ..., N)$ は C の固有値

$$\lambda_1 > \lambda_2 > ... > \lambda_N \geq \lambda_{N+1} \geq ... \geq \lambda_K \geq 0, (K > N)$$

のそれぞれに対応する固有ベクトルである．

次に，このアルゴリズムの時間 t に関する常微分方程式

$$\frac{dW}{dt} = CW - WD^{-1}W^T C W D \qquad (17)$$

が部分空間抽出アルゴリズム (11) とは異なり，$W(0)^T W(0) = I$ であっても，$W(t)^T W(t) = I(k = 1, 2, ...)$ を満たさないことを示す．

$$\begin{aligned}
\frac{dW^T W}{dt} &= (W^T C - DW^T CWD^{-1}W^T)W \\
&\qquad +W^T(CW - WD^{-1}W^T CDW) \\
&= (W^T CW - DW^T CWD^{-1}W^T W) \\
&\qquad +(W^T CW - W^T WD^{-1}W^T CWD) \\
&\neq \mathbf{0}
\end{aligned}$$

しかしながら，W の各列ベクトル $w_n, (n = 1, ..., N)$ に対して $w_n^T w_n = 1$ であるなら，次の等式が成り立つ．

$$\begin{aligned}
\frac{dw_n^T w_n}{dt} &= (w_n^T C - \theta_n w_n^T Cw_n \theta_n^{-1} w_n^T)w_n \\
&\qquad +w_n^T(Cw_n - w_n \theta_n^{-1} w_n^T Cw_n \theta_n) \\
&= (w_n^T Cw_n - w_n^T Cw_n w_n^T w_n) + (w_n^T Cw_n - w_n^T w_n w_n^T Cw_n) \\
&= (w_n^T Cw_n - w_n^T Cw_n) + (w_n^T Cw_n - w_n^T Cw_n) \\
&= 0
\end{aligned} \tag{18}$$

このアルゴリズムも部分空間抽出アルゴリズムと同様にパラメータ α の符号を負に変更した場合の常微分方程式の形式

$$\frac{dW}{dt} = -CW + WD^{-1}W^T CWD$$

においても条件 $w_n^T w_n = 1$ を満たすので，理論的にはマイナー成分を抽出することが可能であるが，実際には数値的不安定性により困難である．また，$W^T W = I$ の条件が崩れているので，この条件を用いて提案されている数値的安定化の手法 (Douglas[98]) と組み合わせることも困難である．

2.3 Wang の双勾配アプローチ

$J_1(W)$ を最小化してマイナー成分が張る部分空間の基底およびマイナー成分固有ベクトルを抽出する場合に Oja のアルゴリズム (11) および (15) で数値的に不安定になると，前提である $W^T W = I$ という条件が崩れる．

Wang ら [96] はこれを抑制するために，次のペナルティ関数

$$J_2(W) = \frac{1}{2}\mathrm{trace}(I - W^T W)^2 \tag{19}$$

$$= \frac{1}{2}\sum_{n=1}^{N}(1 - w_n^T w_n) \tag{20}$$

を最小化する操作を組み合わせて，マイナー成分が張る部分空間の基底およびマイナー成分の固有ベクトルを抽出する次のようなアプローチを提案している．

まず，$J_2(W)$ の勾配は

$$\frac{\partial J_2(W)}{\partial W} = -W(I - W^T W) \tag{21}$$

となる．Wang らのアプローチは式 (21) を用いて，

$$W_{k+1} = W_k + \alpha_k \frac{\partial J_1}{\partial W}|_{W=W_k} + \gamma_k \frac{\partial J_2}{\partial W}|_{W=W_k} \tag{22}$$

のように表される．ここで，$J_2(W)$ の勾配は W_k の正規直交化を強制するペナルティ項としての機能を担う．また γ_k はペナルティ項へのゲイン・パラメータである．

このアルゴリズムはパラメータ α_k の符号が正のとき主成分が張る部分空間の基底，α_k の符号が負のとき，マイナー成分が張る部分空間の基底を抽出する．

しかしながら，ペナルティ項として式 (21) を用いると，アルゴリズムの非対称化が行われていないので，W_k の各列ベクトルは個々の固有ベクトルに収束するとは限らない．

そこで，Wang らは $J_2(W)$ を次のように非対称化している．

$$\frac{\partial J_2'(W)}{\partial W} = -W \times \mathrm{UT}[I - W^T W] \tag{23}$$

ここで，$\mathrm{UT}[A]$ は行列 A の上三角成分をそのままにして，それ以外の要素をすべて 0 にする操作を表している．

式 (21) を式 (23) に置き換え，パラメータ α_k と γ_k が十分小さいと仮定することで，Wang らのアルゴリズムは次のように表される．

$$W_{k+1} = W_k + \alpha_k C W_k + \gamma_k W_k \times \mathrm{UT}[1 - W_k^T W_k] \tag{24}$$

このアルゴリズムはパラメータ α_k の符号が正にとられるとき，主成分固有ベクトルを，負にとられるとき，マイナー成分固有ベクトルを抽出する．

しかしながら，$k \to \infty$ のとき，$\gamma_k \to \infty$ となるように γ_k を変更しなければ，W が正規化されない．また，γ_k を適当に変更しなければ，アルゴリズムが不安定になることがあるが，適当にパラメータを変更することは容易ではない．

2.4 Chen の部分空間アルゴリズム

Chen によって提案された部分空間抽出アルゴリズムについて説明する．
Chen のアルゴリズムは，式 (3) と式 (19) を用いて，

$$\begin{aligned} \tilde{W}_{k+1} &= W_k + \alpha_k \frac{\partial J_1(W)}{\partial W}|_{W=W_k} \\ &= W_{k-1} + \alpha_k C W_{k-1} \qquad (25) \\ W_{k+1} &= \tilde{W}_{k+1} S_k^{-1} + \gamma_k \frac{\partial J_2(W)}{\partial W}|_{W=W_k} \\ &= \tilde{W}_{k+1} S_k^{-1} + \gamma_k W_k (I - W_k^T W_k) \qquad (26) \end{aligned}$$

のようにして導出される．ここで，S_k^{-1} は Oja の部分空間抽出アルゴリズム (11) と同様に

$$S_k = (\tilde{W}_k^T \tilde{W}_k)^{\frac{1}{2}}$$

と選ばれている．パラメータ γ は Wang らのアルゴリズム (24) と同様に正規直交化を強制するペナルティ項に対するゲイン・パラメータであるが，式 (24) とは異なり，一定とされる．

式 (25) と式 (26) においてパラメータ α_k が十分小さいと仮定することで，Chen のアルゴリズムは次の形式で表される．

$$\begin{aligned} W_{k+1} &= W_k + \alpha_k [C W_k - W_k W_k^T C W_k^T \\ &\quad + \gamma (I - W_k W_k^T) W_k] \qquad (27) \end{aligned}$$

このアルゴリズムは α_k の符号が正で，かつ $\gamma > 0$ の場合に主成分が張る部分空間の基底を，α_k の符号が負で，かつマイナー成分に対応する固有値が $\mu_1 < ... < \mu_N$ となっている場合に $\gamma > \mu_N$ とすることで，マイナー成分が張る部分空間の基底を抽出する．

しかしながら，Oja の部分空間抽出アルゴリズム (11) と同様に，各々の固有ベクトルが得られる保証はない．

3　提案手法

前章までで，従来提案されているアルゴリズムについて説明してきた．これらの手法には

1. マイナー成分の抽出における数値安定化のために正規直交行列の空間上の力学系を考えている．
2. 各々の固有ベクトルを抽出するためには，Gram-Schmidt の正規直交化に対応するようなアルゴリズムの非対称化が必要である．ただし，上記の手法 (W^TW の積による正規直交化の強化) では非対称化によって $W^TW = I$ が崩れる．

という相反する条件が必要であった．そこで，主成分およびマイナー成分を抽出するための非対称化を行いつつ，ペナルティ項を付加することで正規直交化を行う手法を提案する．

提案する修正 Oja アルゴリズムは，

$$\begin{aligned} W_{k+1} &= W_k \pm \alpha_k(CW_k - W_k D_{-1} W_k^T C W_k^T D) \\ &\quad +(\alpha_k\gamma)W_k(I - W_k^T W_k) \end{aligned} \tag{28}$$

と表される．ここで，行列 D は Oja の主成分抽出アルゴリズムと同様に

$$\begin{aligned} &D = \mathrm{diag}(\theta_1, ..., \theta_N), \\ &0 < \theta_1 < ... < \theta_N \end{aligned}$$

と定められ，行列 W の各列ベクトル $w_n, (n = 1, ..., N)$ が行列 C の固有ベクトルに収束することが保証される．また，右辺第三項は Chen のアルゴリズム (27) と同じく，正規直交化を強制するペナルティ項として機能する．γ はペナルティ項へのゲインである．右辺第三項に与えられる係数が $(\alpha_k\gamma)$ となっているのは，後の微分方程式 (29) との対応のためである．

式 (28) の右辺第二項にあるパラメータ α_k の符号の正負がそれぞれ主成分およびマイナー成分の抽出に対応している．また，式 (28) の右辺第三項のパラメータ γ に関する説明は後述する．

このアルゴリズムによって主成分を抽出する場合に対して，以下の定理が成立する．

【定理 1 】

K 個の主成分である固有ベクトルを列ベクトルとした行列

$$\tilde{V} = [v_1 v_2 ... v_K], ||v_i||^2 = 1, i = 1, ..., K$$

と，それぞれに対応する固有値

$$\lambda_1, \lambda_2, ..., \lambda_K, \lambda_1 > \lambda_2 > ... > \lambda_N \geq \lambda_{N+1} \geq ... \geq \lambda_K \geq 0, (K > N)$$

をもつ対角行列

$$\tilde{\Lambda} = diag(\lambda_1, \lambda_2, ..., \lambda_K)$$

で表される正定値対称行列

$$C = \tilde{V}\tilde{\Lambda}\tilde{V}^T \in \mathbf{R}^{K \times K}$$

と対角行列

$$D = diag(\theta_1, ..., \theta_N) \in \mathbf{R}^{N \times N}$$

で定義された

$$\frac{dW}{dt} \equiv CW - WD^{-1}W^T CWD + \gamma W(I - W^T W) \tag{29}$$

の解

$$W \in \mathbf{R}^{K \times N}$$

は，

$$V = [v_1 v_2 ... v_N]$$

で与えられ，

$$0 < \theta_1 < \theta_2 < ... < \theta_N \tag{30}$$

かつ

$$\gamma > 0 \tag{31}$$

の場合に，漸近的に安定である．

また，マイナー成分を抽出する場合に対しては以下の定理が成立する．

【定理 2 】

k 個のマイナー成分である固有ベクトルを列ベクトルとした行列

$$\tilde{V} = [v_1 v_2 ... v_K], ||v_i||^2 = 1, i = 1, ..., K$$

とその固有値

$$\mu_1, \mu_2, ..., \mu_K, 0 < \mu_1 < \mu_2 < ... < \mu_N \leq \mu_{N+1} \leq ... \leq \mu_K (K > N)$$

をもつ対角行列

$$\tilde{M} = diag(\mu_1, \mu_2, ..., \mu_K)$$

で表される正定値対称行列

$$C = \tilde{V}\tilde{M}\tilde{V}^T \in \mathbf{R}^{K \times K}$$

と対角行列

$$D = diag(\theta_1, ..., \theta_N) \in \mathbf{R}^{N \times N}$$

で定義された

$$\frac{dW}{dt} \equiv -CW + WD^{-1}W^T CWD + \gamma W(I - W^T W) \tag{32}$$

の解

$$W \in \mathbf{R}^{K \times N}$$

は，

$$V = [v_1 v_2 ... v_N]$$

で与えられ，

$$0 < \theta_1 < \theta_2 < ... < \theta_N \tag{33}$$

かつ

$$\gamma > \mu_N \tag{34}$$

の場合に，漸近的に安定である．

以上の 2 つの定理の証明は付録で行われている．また次節では，このアルゴリズムの大域的な収束特性と利用可能なパラメータの範囲について調べるために行った数値実験の結果を示す．

4　数値実験

提案するアルゴリズムの収束特性と利用可能なパラメータの範囲を検証するために，以下に数値実験を行った．

4.1　収束特性と利用可能なパラメータの範囲

まず，以下の (5×5) の実対称行列

$$C = \begin{bmatrix} 1.5788 & 0.5600 & 0.3143 & 0.1957 & 0.0916 \\ 0.5600 & 2.0936 & 0.1623 & 0.3555 & -0.2863 \\ 0.3143 & 0.1623 & 2.1498 & 0.0467 & 0.4530 \\ 0.1957 & 0.3555 & 0.0467 & 1.7025 & 0.7877 \\ 0.0916 & -0.2763 & 0.4530 & 0.7877 & 1.9752 \end{bmatrix} \tag{35}$$

に対して，主成分とマイナー成分を求める数値実験を行った．この行列は Chen[97] の数値実験例で用いられている．また，他のデータに対しても同様の特性を示した．

この行列がもつ固有値は

$$\begin{aligned} \lambda_1 &= 2.998 \\ \lambda_2 &= 2.504 \\ \lambda_3 &= 1.999 \\ \lambda_4 &= 1.200 \\ \lambda_5 &= 0.7976 \end{aligned}$$

となっている．これらは，LAPACK[95] に含まれる固有値問題を解くためのルーチンを用いて予め求めた固有値である．

また計算途中，解ベクトルを評価するために，指標として，以下の最大値ノルム

$$\rho_k = \max|CW - W\Lambda| \tag{36}$$

$$\eta_k = \max|I - W^T W| \tag{37}$$

を用いた．ただし，$\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_N)$，$|A|$ は行列 A の各要素の絶対値をとった行列である．ρ_k と η_k はそれぞれ各ステップ k における解ベクトル w の固有値の近似と正規直交化の度合いを表している．

まず，ペナルティ項のパラメータ γ による主成分およびマイナー成分の抽出における大域的な収束特性を調べる数値実験の結果を示す．ここでは，パラメータを

$$\begin{aligned} \alpha &= 0.05 \\ D &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \end{aligned} \tag{38}$$

で一定，W の列数は $N = 3$ とした．

4.1.1 主成分抽出に対する実験結果

この節では以下の場合について数値実験を行った．

(P_1)	$\gamma = 0$	図 1
(P_2)	$\gamma = 3$	図 2
(P_3)	$\gamma = 23$	図 3
(P_4)	$\gamma = 30$	図 4

(P_1) 主成分の抽出に対して，$\gamma = 0$ で一定とした場合の固有ベクトルの近似の度合を表すノルム ρ_k と正規直交化の度合を表すノルム η_k のステップ毎の変化を示す．この場合は $\gamma = 0$ であることから，Oja の主成分抽出アルゴリズム (15) に他ならない．

ρ_k と η_k がともに 0 に収束していることから，解ベクトルが各主成分に収束していることが分かる．

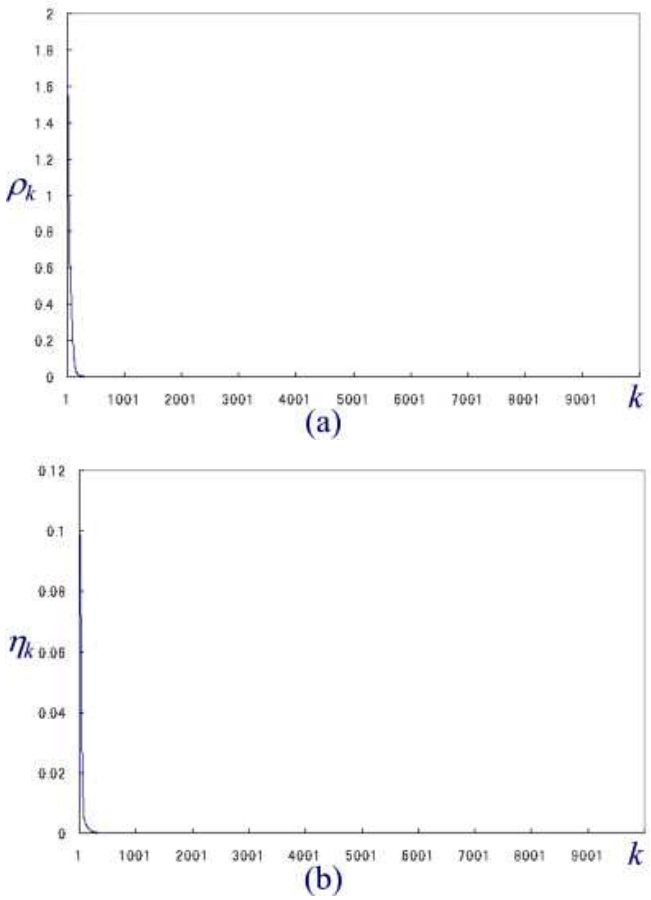

図 1: (P_1)PCA **の結果** $1(\alpha = 0.05, \gamma = 0)$

(P_2) 主成分の抽出に対して，本提案手法 (28) において $\gamma = 3$ で一定とした場合のノルム ρ_k と η_k のステップ毎の変化を示す．

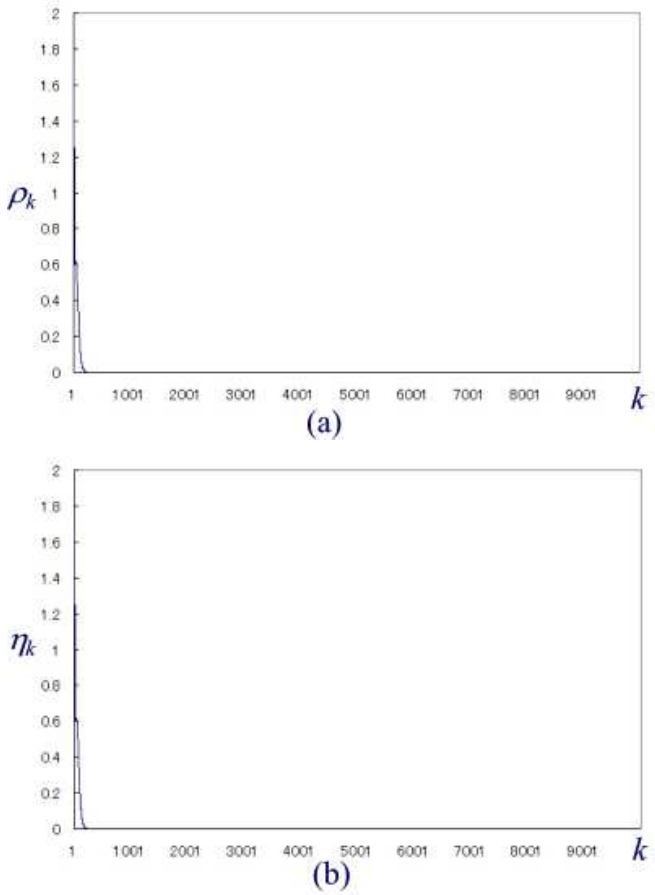

図 2: (P_2)PCA の結果 $2(\alpha = 0.05, \gamma = 3)$

この場合も (P_1) と同様に，ρ_k ，η_k ともに 0 に収束していることから，図 $2(\mathrm{a})(\mathrm{b})$ は解ベクトルが各主成分に収束していることを示している．また図 $1(\mathrm{a})(\mathrm{b})$ と図 $2(\mathrm{a})(\mathrm{b})$ を比較して，(P_1) と (P_2) の場合，正規直交化ペナルティ項が主成分の抽出において大域的な収束特性に影響を与えないことが分かる．

(P_3) 主成分の抽出に対して，本提案手法 (28) においてさらにペナルティ項の影響を大きくするために $\gamma = 23$ で一定とした場合のノルム ρ_k と η_k のステップ毎の変化を示す．

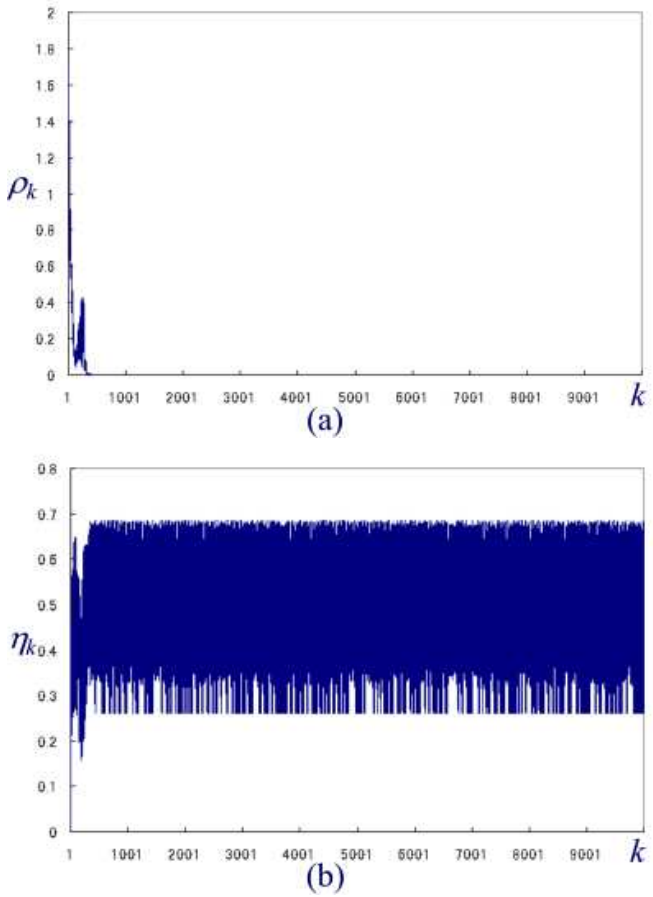

図 3: (P_3)PCA の結果 $3(\alpha = 0.05, \gamma = 23)$

この場合では，ρ_k は 0 に収束しているように見えるが，η_k は振動している．Reyleigh 商 (6) によって求めた $w_n, (n = 1, 2, 3)$ のそれぞれに対応する近似固有値は

$$\begin{aligned} \lambda_1^* &= 1.5788, \\ \lambda_2^* &= 2.0936, \\ \lambda_3^* &= 2.1498 \end{aligned}$$

となっていて，16 ページに示した正しい固有値に対応する成分の抽出が
正確に行われていない．

(P_4) 主成分の抽出に対して，本提案手法 (28) において $\gamma = 30$ で一定とした場合のノルム ρ_k と η_k のステップ毎の変化を示す．

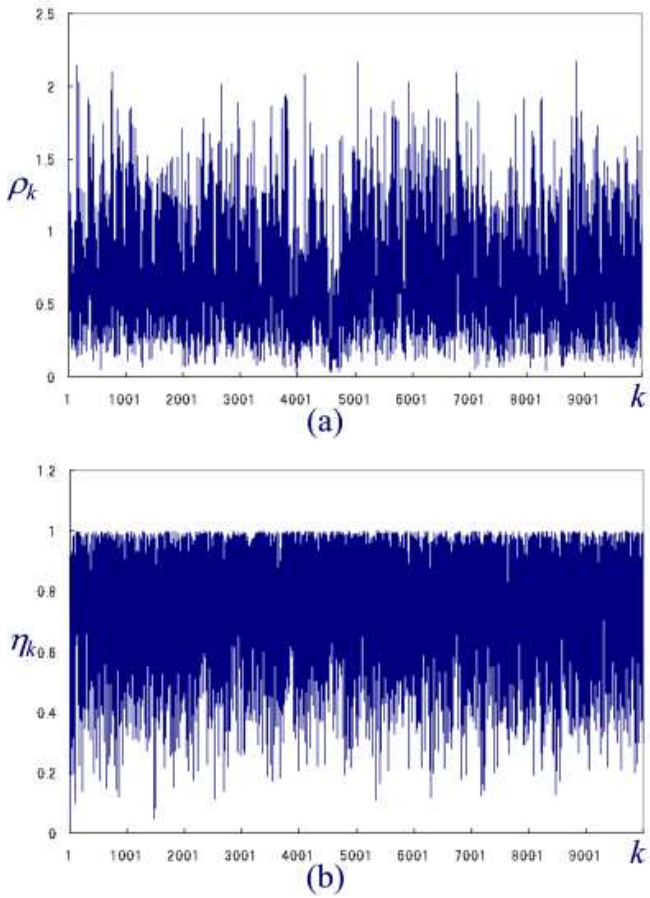

図 4: (P_4)PCA の結果 $4(\alpha = 0.05, \gamma = 30)$

この場合は ρ_k，η_k ともに振動していて，成分の抽出が行われていない．

(P_3) の結果と併せて，ペナルティ項の影響を過剰に大きくした場合に，アルゴリズムが不安定になることが分かる．

4.1.2　マイナー成分抽出に対する実験結果

この節では以下の場合について数値実験を行った．

(M_1)	$\gamma = 5$	図 5
(M_2)	$\gamma = 1.025$	図 6
(M_3)	$\gamma = 23$	図 7
(M_4)	$\gamma = 40$	図 8

(M_1) マイナー成分の抽出に対して，本提案手法において $\gamma = 5$ で一定とした場合のノルム ρ_k と η_k のステップ毎の変化を示す．

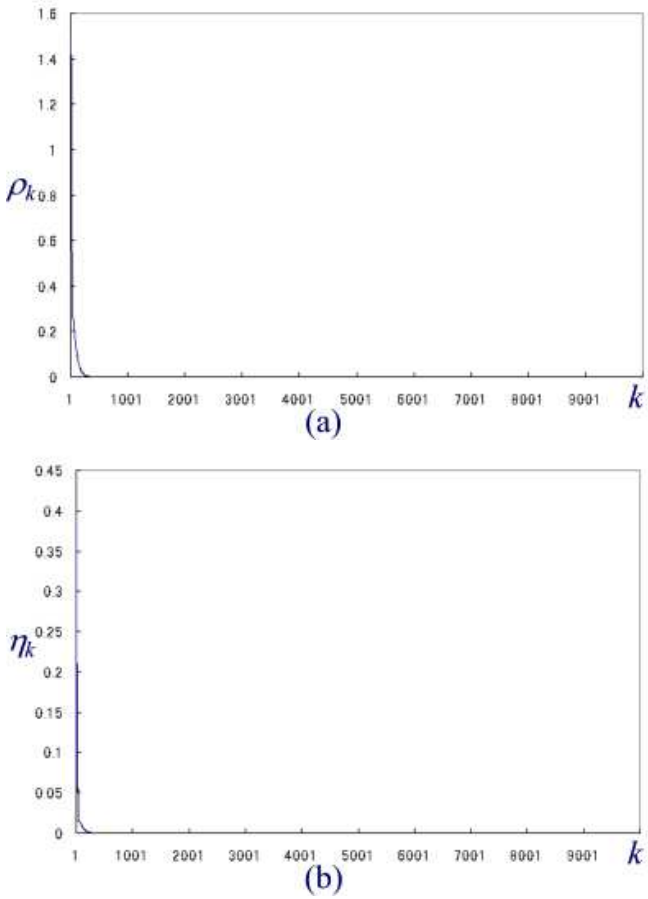

図 5: (M_1)MCA の結果 $1(\alpha = 0.05, \gamma = 5)$

主成分の場合と同様に，ρ_k ，η_k ともに 0 に収束していることが，$w_n, (n = 1, 2, 3)$ のそれぞれが各マイナー成分に収束していることを示している．また，図 1，図 2 と図 5 を比較して，大域的な収束特性は主成分とマイナー成分どちらの抽出においても同様であることが分かる．

(M_2) マイナー成分の抽出に対して，本提案手法において $\gamma = 1.025 < \lambda_3$ で一定とした場合のノルム ρ_k と η_k のステップ毎の変化を示す．このとき，γ は定理 2 の条件 (34) を満たしていない．

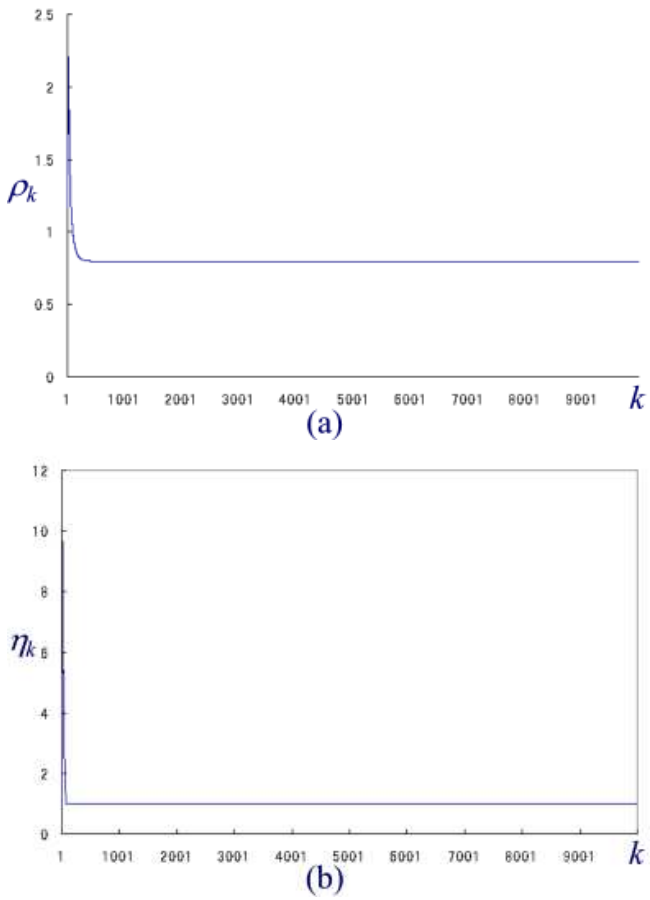

図 6: (M_2)MCA の結果 $2(\alpha = 0.05, \gamma = 1.025)$

この場合，ρ_k, η_k がともに 0 に収束していない．レイレー商を用いて求めた近似固有値が

$$\begin{aligned} \lambda_5^* &= -3.3595 \times 10^{-47}, \\ \lambda_4^* &= 3.5919 \times 10^{-107}, \\ \lambda_3^* &= -3.3595 \times 10^{-47} \end{aligned}$$

となっていることから，16 ページに示した成分の抽出は行われていない．

(M_3) マイナー成分の抽出に対して，本提案手法において $\gamma = 23 > \lambda_3$ で一定としてペナルティ項の影響を大きくした場合のノルム ρ_k と η_k のステップ毎の変化を示す．

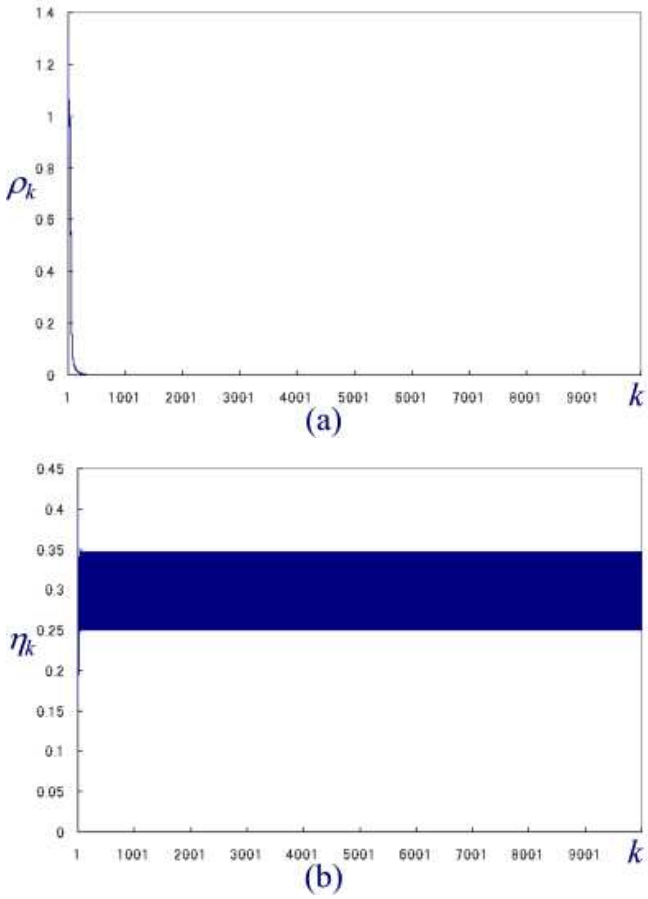

図 7: (M_3)MCA の結果 $3(\alpha = 0.05, \gamma = 23)$

この場合は，ρ_k は 0 に収束しているが，η_k は振動している．Reyleigh 商 (6) による近似固有値は

$$\begin{aligned} \lambda_5^* &= 1.5788, \\ \lambda_4^* &= 2.0936, \\ \lambda_3^* &= 2.1498 \end{aligned}$$

となっていて，成分の抽出は行われていない．

(M_4) マイナー成分の抽出に対して，本提案手法において $\gamma = 40$ で一定とした場合のノルム ρ_k と η_k のステップ毎の変化を示す．

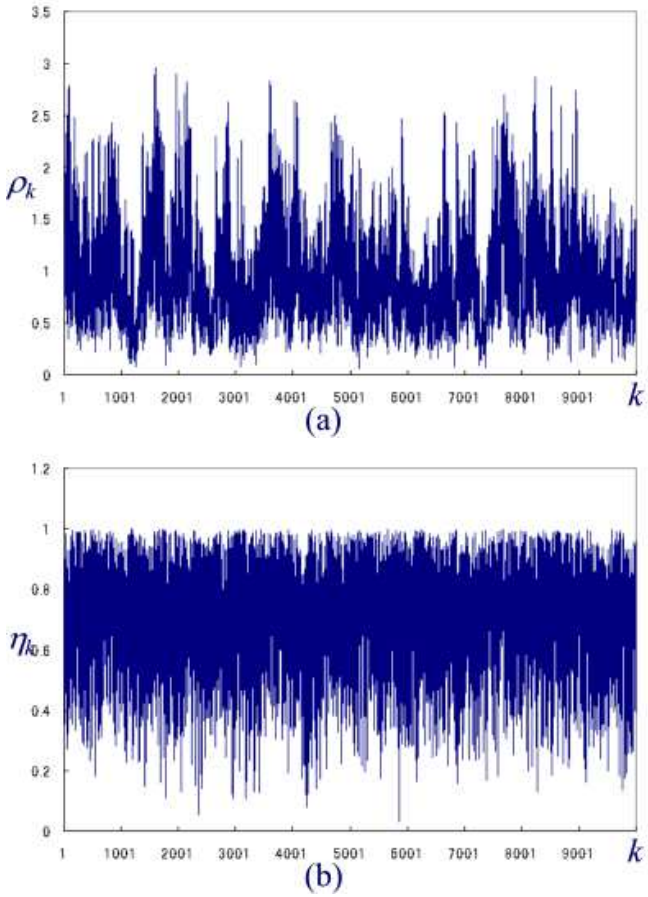

図 8: (M_4)MCA の結果 $4(\alpha = 0.05, \gamma = 40)$

この場合は，ρ_k，η_k ともに振動していて，0 に収束していないことから，成分が抽出されていない．また，(M_3) の結果と併せて，主成分の抽出の場合と同じく，ペナルティ項の影響を過剰に大きくすると，アルゴリズムが不安定になる．

4.1.3　行列 D に対する比較

4.1.1 と 4.1.2 では，PCA と MCA におけるペナルティ項の影響について示してきた．

続いて，式 (35) の対称行列 C の PCA，MCA に対して，利用可能なパラメータの範囲が下に示す行列 D によってどのように変化するのかを調べた．

$$(D_1) \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$(D_2) \qquad D = \begin{bmatrix} 0.09 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(D_3) \qquad D = \begin{bmatrix} 0.81 & 0 & 0 \\ 0 & 0.9 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(D_4) \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

ここで利用可能なパラメータとは ρ_k と η_k が 10000 ステップ以内に 100 回以上連続で 10^{-3} となるようなパラメータを示している．ここでは，求める成分の数を $N = 3$ としている．

また，行列 D のそれぞれの場合に対して，行列 C の固有ベクトルからなる行列 $V = [v_1 ... v_5]$ とそれぞれの解ベクトル $w_1, w_2 . w_3$ の積 $(Vw_1), (Vw_2), (Vw_3)$ のステップ毎の変化のグラフを図 11，12，15，16，19，20，23，24 に示す．

パラメータ α と γ は主成分抽出では $\alpha = 0.05, \gamma = 2$，マイナー成分抽出では $\alpha = 0.05$，$\gamma = 10$ で一定としている．

(D_1) D を

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

とした場合の結果を示す．図中の色が付いている部分が利用可能なパラメータの範囲を示している．

図 9 と図 10 の各点 (P_1)，(P_2)，(P_3)，(P_4)，(M_1)，(M_2)，(M_3)，(M_4) は $4.1.1$ と $4.1.2$ で示した結果図 1，2，3，4，5，6，7，8 を求めたときのパラメータに対応している．

図 10 における利用可能なパラメータの γ の下限は定理 2 の γ に対する条件 (34) の下限である $\lambda_3 = 1.999$ とほぼ一致している．

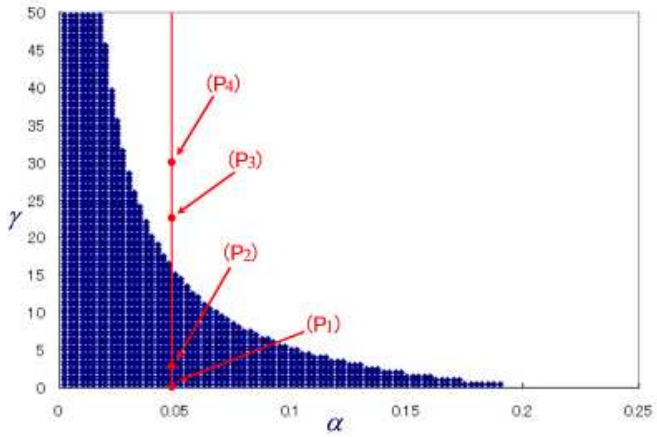

図 9: PCA のパラメータ・マップ 1

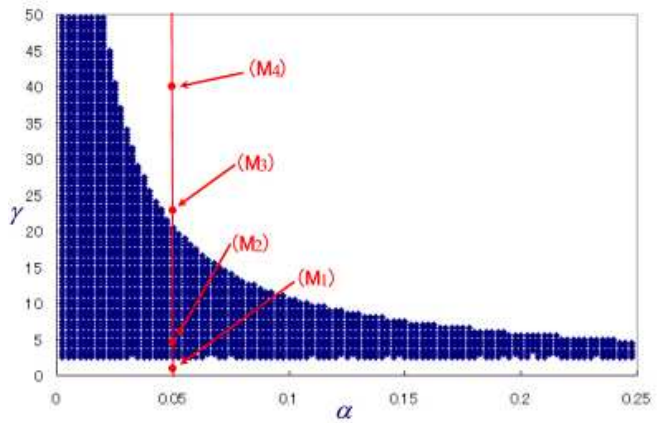

図 10: MCA のパラメータ・マップ 1

下に示す 3 つの図は主成分の抽出における C の固有ベクトルからなる行列 $V = [v_m, (m = 1, 2, 3, 4, 5)]$ と各々の解ベクトル $w_n, (n = 1, 2, 3)$ の積 $V \cdot w_n, (n = 1, 2, 3)$ のステップ毎の変化を示している．グラフから，1000 回程度で収束していることが分かる．

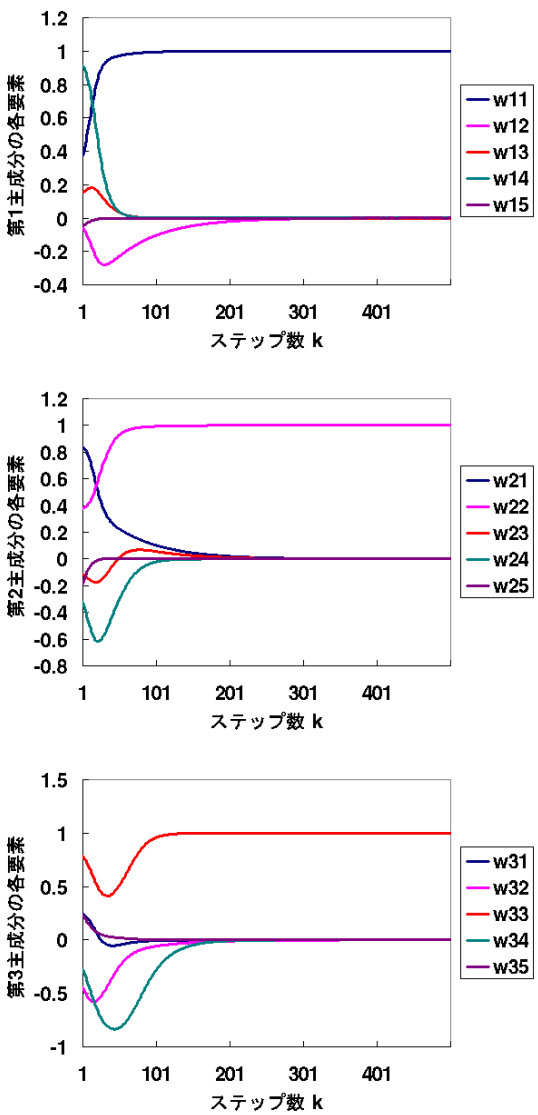

図 11: 主成分抽出における各成分のステップ毎の変化 1

以下も図 11 と同様にマイナー成分の抽出における $V \cdot w_n, (n = 1, 2, 3)$ のステップ毎の変化を示している．各要素の収束が図 11 の主成分の場合よりも遅くなっている．

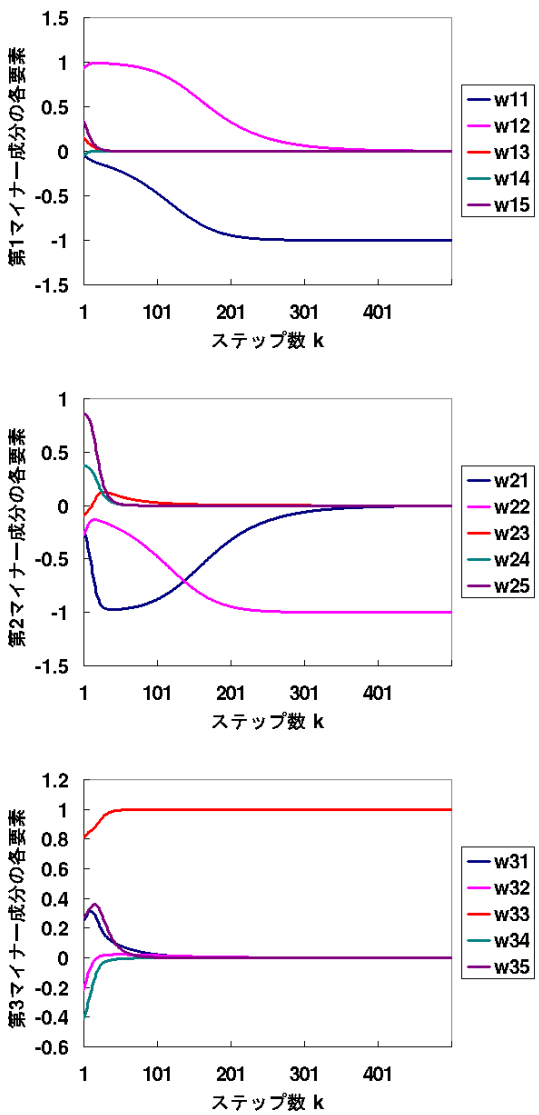

図 12: マイナー成分抽出における各成分のステップ毎の変化 1

(D_2) D を

$$D = \begin{bmatrix} 0.09 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

とした場合の結果を示す．

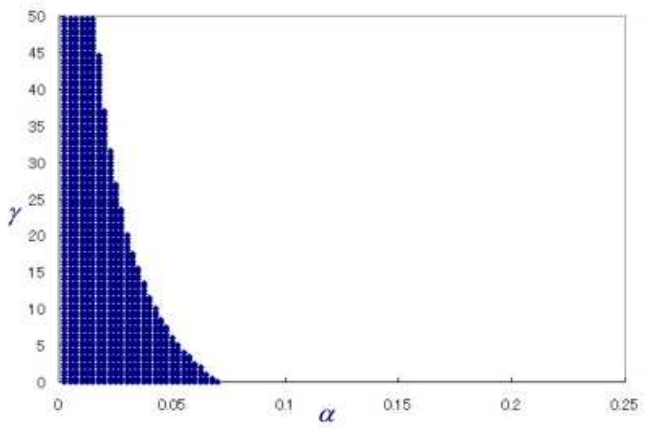

図 13: PCA のパラメータ・マップ 2

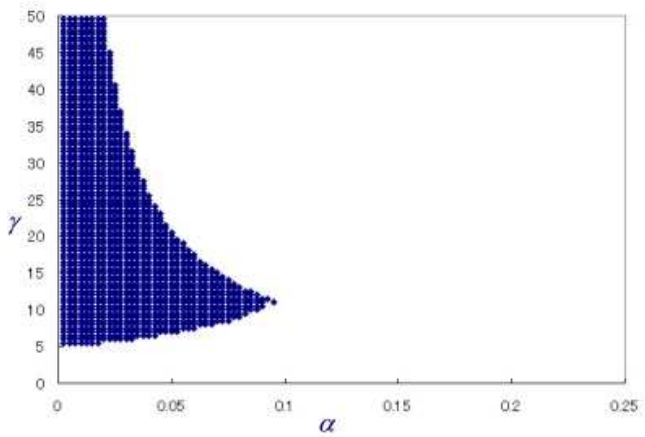

図 14: MCA のパラメータ・マップ 2

(D_1) の場合と比較して，

- 利用可能なパラメータの α の上限が主成分の抽出に対して 0.19 から 0.07 程度にマイナー成分の抽出に対して 0.25 以上から 0.1 程度に低くなっている．
- MCA の γ の下限が 2.0 から 5.0 程度まで高くなっている．また，このことは定理 2 の γ に対する条件 (34) : $\gamma > \lambda_3 = 1.999$ が平衡点での必要条件に他ならないことを反映している．すなわち，実際にはより大きな値でないと収束しない．

以上の 2 点の変化が見られた．

以下の 3 つの図は主成分の抽出における C の固有ベクトルからなる行列 $V = [v_m, (m = 1, 2, 3, 4, 5)]$ と各々の解ベクトル $w_n, (n = 1, 2, 3)$ の積 $V \cdot w_n, (n = 1, 2, 3)$ のステップ毎の変化を示している．グラフから，500 回程度で収束していることが分かる．

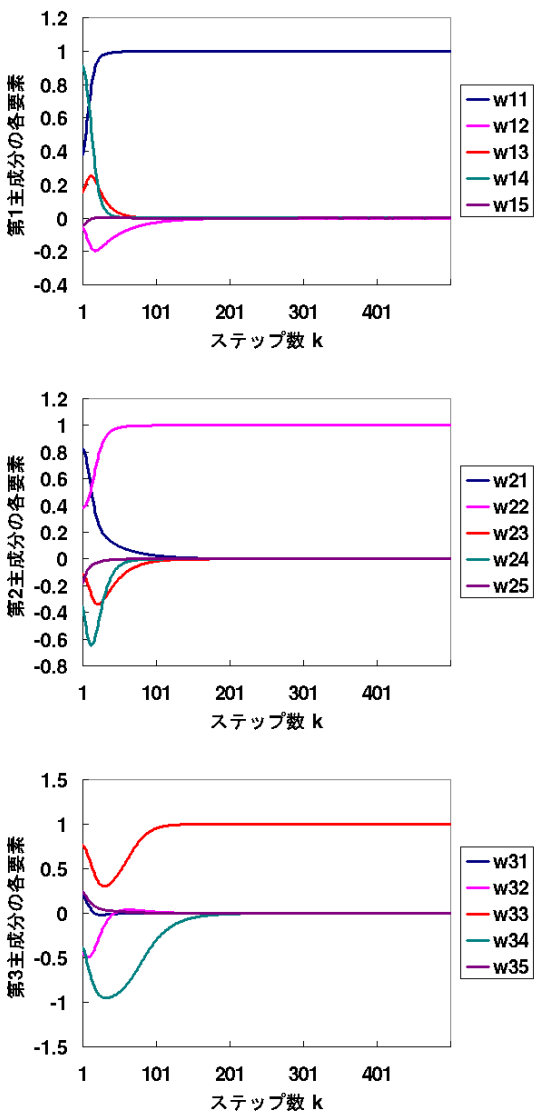

図 15: 主成分抽出における各成分のステップ毎の変化 2

下に示す 3 つの図はマイナー成分の抽出における C の固有ベクトルからなる行列 $V = [v_m, (m = 1, 2, 3, 4, 5)]$ と各々の解ベクトル $w_n, (n = 1, 2, 3)$ の積 $V \cdot w_n, (n = 1, 2, 3)$ のステップ毎の変化を示している．主成分の場合と同様に，500 回程度で収束していることが分かる．

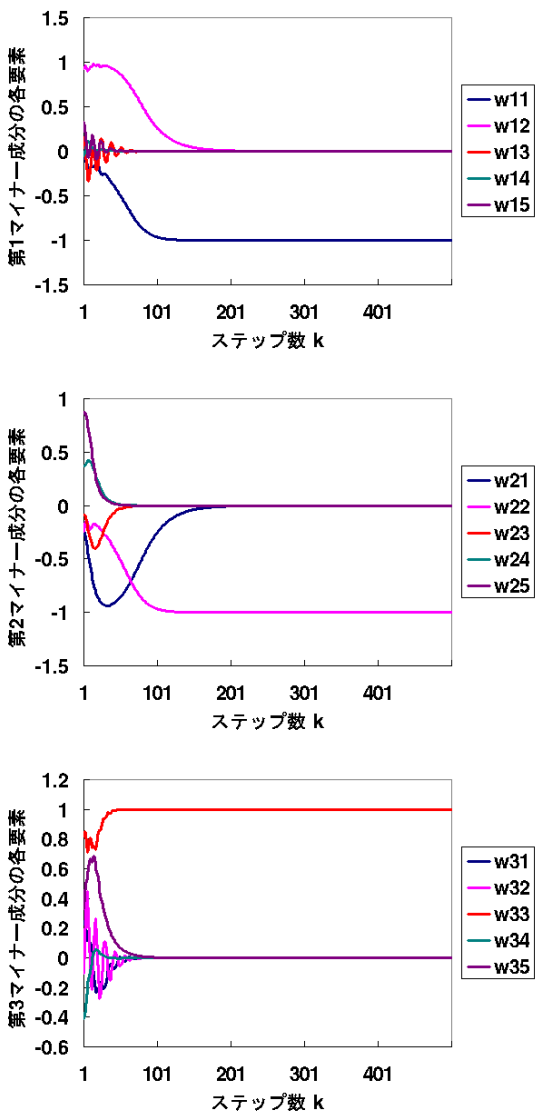

図 16: マイナー成分抽出における各成分のステップ毎の変化 2

(D_3) D を

$$D = \begin{bmatrix} 0.81 & 0 & 0 \\ 0 & 0.9 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

とした場合の結果を示す．

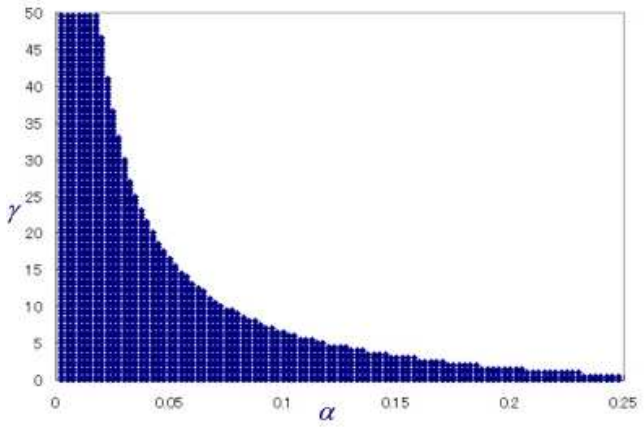

図 17: PCA のパラメータ・マップ 3

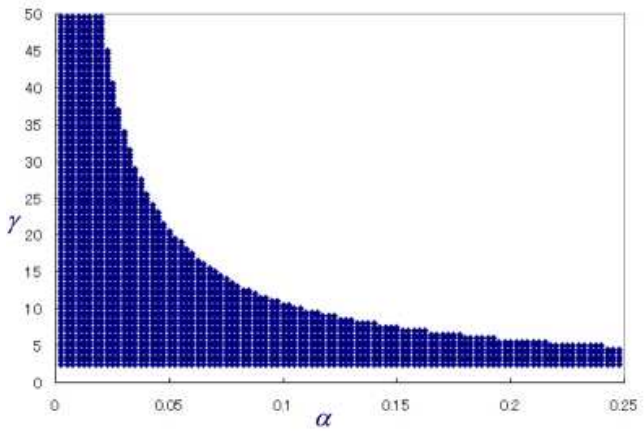

図 18: MCA のパラメータ・マップ 3

(D_1) と同様に，図 17 の γ の下限は定理 2 の γ に関する条件の下限 $\lambda_3 =$

1.999 と一致している．また，(D_1) と比較して利用可能なパラメータの α に付いての上限が 0.25 以上にまで高くなっている．

下に示す 3 つの図は主成分の抽出における C の固有ベクトルからなる行列 $V = [v_m, (m = 1, 2, 3, 4, 5)]$ と各々の解ベクトル $w_n, (n = 1, 2, 3)$ の積 $V \cdot w_n, (n = 1, 2, 3)$ のステップ毎の変化を示している．(D_1)，(D_2) とは異なり各要素が収束するまでステップ数が 3000 回程度掛かっている．

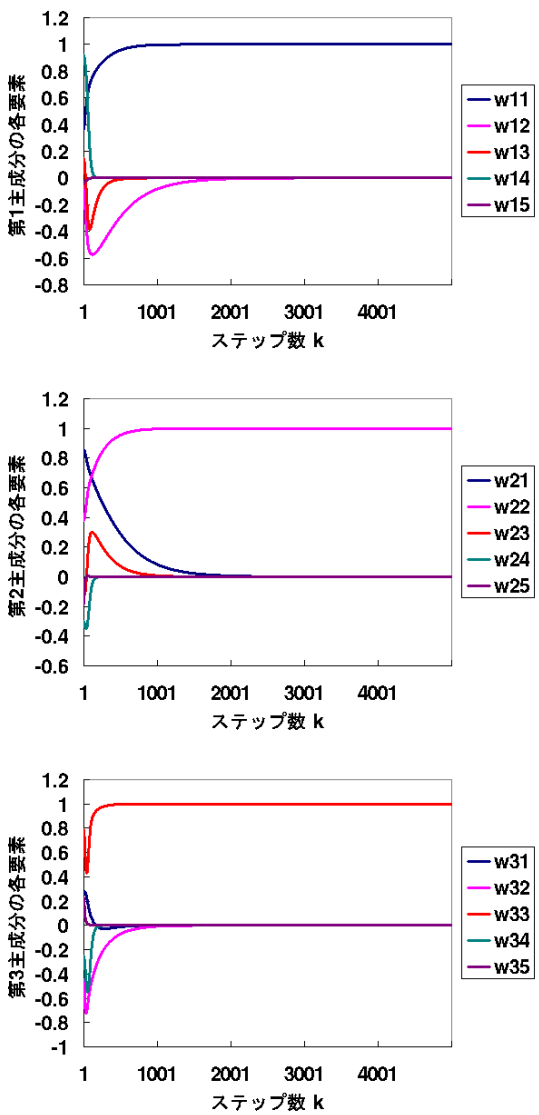

図 19: 主成分抽出における各成分のステップ毎の変化 3

下の 3 つの図は主成分の抽出における C の固有ベクトルからなる行列 $V = [v_m, (m = 1, 2, 3, 4, 5)]$ と各々の解ベクトル $w_n, (n = 1, 2, 3)$ の積 $V \cdot w_n, (n = 1, 2, 3)$ のステップ毎の変化を示している．図 19 の場合と同様に各要素が収束するまでステップ数が 3000 回程度掛かっている．図 19 の結果と併せて，行列 D の要素の比が 1 に近くなると収束が遅くなる傾向があることが分かる．

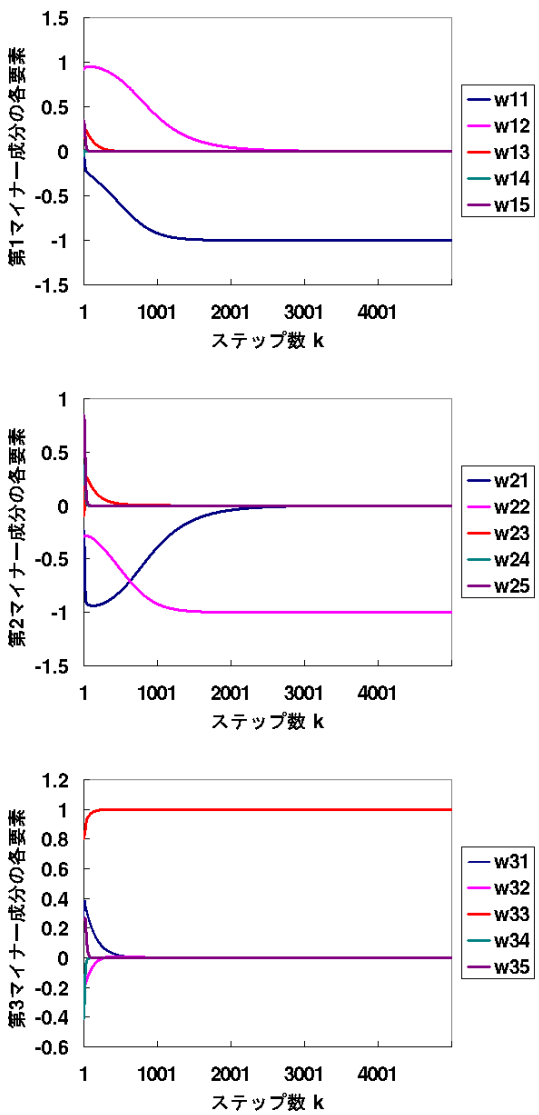

図 20: マイナー成分抽出における各成分のステップ毎の変化 3

(D_4) D を

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

とした場合の結果を示す．

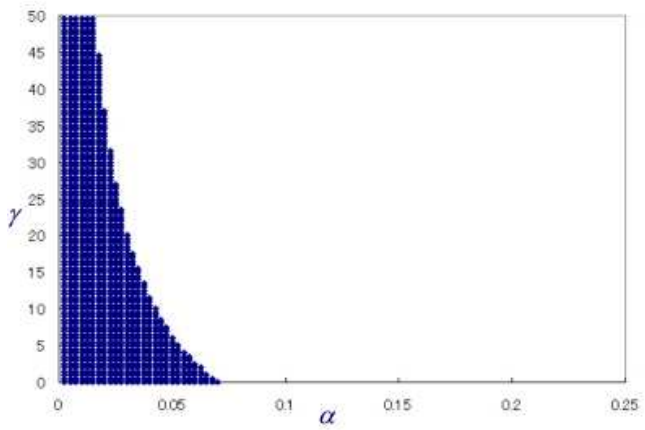

図 21: PCA のパラメータ・マップ 4

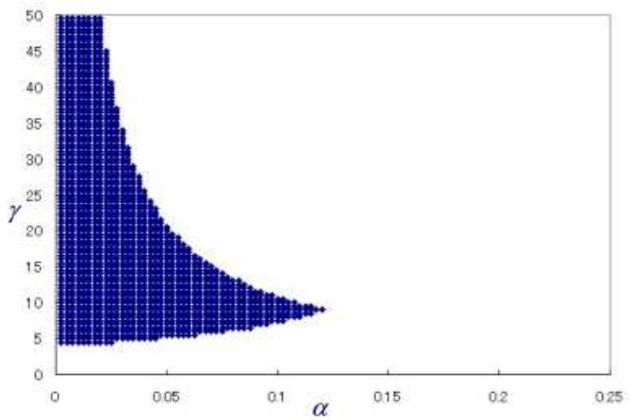

図 22: MCA のパラメータ・マップ 4

(D_2) の場合と同様に (D_1)，(D_3) の場合と比較して，

- 利用可能なパラメータの α の上限が主成分の抽出に対して 0.07 程度までマイナー成分の抽出に対して 0.25 以上から 0.12 程度まで低くなっている．
- MCA の γ の下限が 2.0 から 5.0 程度まで高くなっている．また，このことは定理 2 の γ に対する条件 (34) : $\gamma > \lambda_3 = 1.999$ が平衡点での必要条件に他ならないことを反映している．すなわち，実際にはより大きな値でないと収束しない．

以上の 2 点の変化が見られた．

以下の 3 つの図は主成分の抽出における C の固有ベクトルからなる行列 $V = [v_m, (m = 1, 2, 3, 4, 5)]$ と各々の解ベクトル $w_n, (n = 1, 2, 3)$ の積 $V \cdot w_n, (n = 1, 2, 3)$ のステップ毎の変化を示している．グラフから，各要素が 500 回程度で収束していることが分かる．

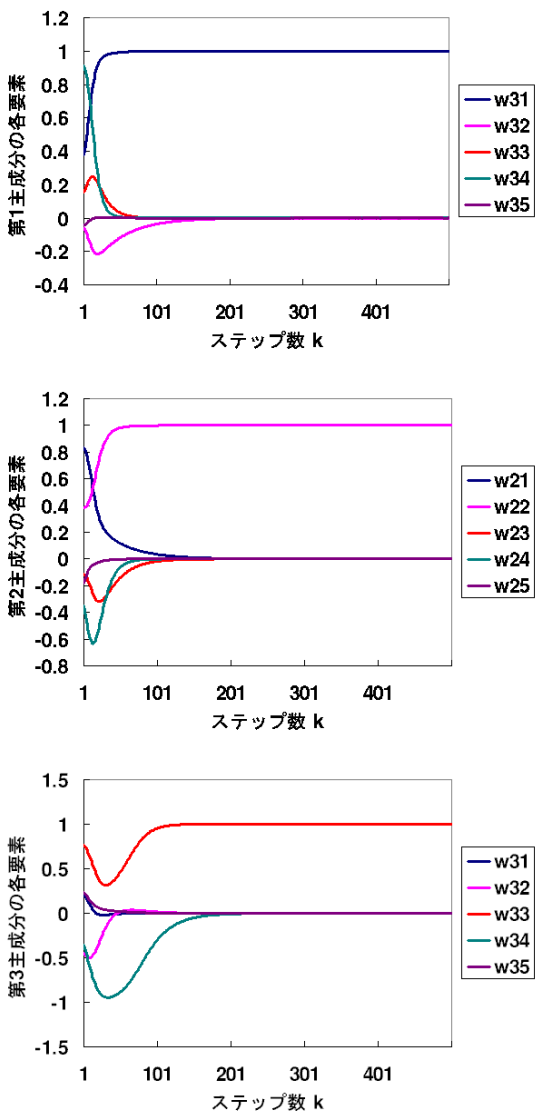

図 23: 主成分抽出における各成分のステップ毎の変化 4

以下の 3 つの図はマイナー成分の抽出における C の固有ベクトルからなる行列 $V = [v_m, (m = 1, 2, 3, 4, 5)]$ と各々の解ベクトル $w_n, (n = 1, 2, 3)$ の積 $V \cdot w_n, (n = 1, 2, 3)$ のステップ毎の変化を示している．図 23 の場合と同様に，500 回程度で収束していることが分かる．

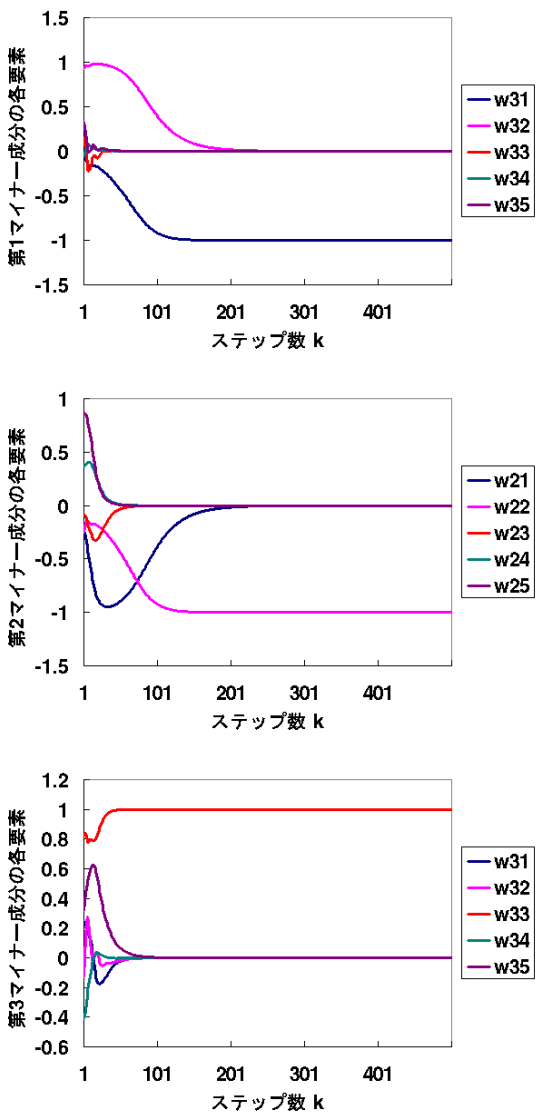

図 24: マイナー成分抽出における各成分のステップ毎の変化 4

以上の結果より，

- (D_2)，(D_4) から D の要素の比が 1 から離れる程，収束が速くなるが，利用可能なパラメータの範囲が狭くなる．
- (D_3) から，D の要素の比が 1 に近い方が，利用可能なパラメータの範囲が広い．しかし，解の収束が遅くなる．

ということが分かった．

4.2 グラフの隣接行列に対する適用例

本節では，本提案手法を Kleinberg[98] で行われている Web ページのリンク構造の解析に対して適用した結果を示す．

Kleinberg は Web ページのリンク構造を解析するために，それぞれの Web ページに対して参照に関する hub 重みと被参照に関する authority 重みを計算する次のようなアルゴリズムを提案している．

$x(p)$ をページ p の authorty 重み，$y(p)$ をページ p の hub 重み，行列 A を $(M \times M)$ のネットワークのリンク構造を表す隣接行列であるとする．

各 Web ページの hab 重み $x(p), (p = 1, ..., M)$ と authority 重み $y(p), (p = 1, ..., M)$ からなるベクトルを $\mathbf{x} = [x(1)...x(M)]^T$，$\mathbf{y} = [y(1)...y(M)]^T$ として，Kleinberg のアルゴリズムを構成する基本部分は次のように表される．

$$\mathbf{x} \leftarrow \mathbf{A^T y} \leftarrow \mathbf{A^T A x}$$
$$\mathbf{y} \leftarrow \mathbf{A x} \leftarrow \mathbf{A A^T y}$$

これは，$\mathbf{A}^T\mathbf{A}$ と $\mathbf{A}\mathbf{A}^T$ の第 1 主成分を求めるベキ乗法に他ならない．

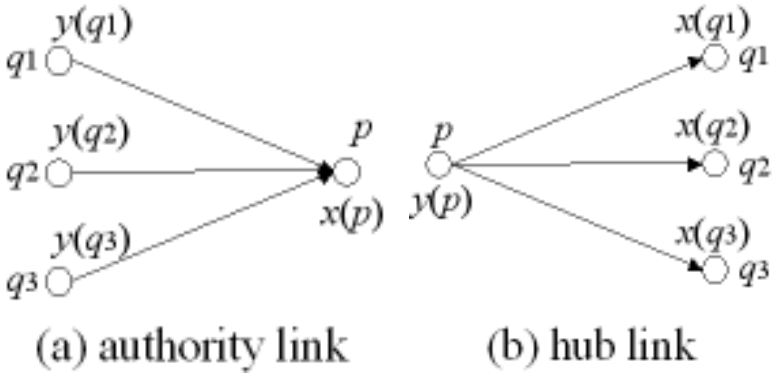

図 25: authority link と hub link

Kleinberg[98] は実際の Web ページのリンク構造を解析している．一方で，Web ページのリンク構造が Zipf の法則に従う傾向をもつことが明らかにされている ([99])．そこで，この例では in-degree(authority link の数) および out-degree(hub link の数) に対するノード (Web ページに対応) の

ただし，$||\mathbf{x}|| = 1$，$||\mathbf{y}|| = 1$ とする．

個数が Zipf の法則に従う人工的なネットワークに対して提案手法を適用した．ネットワークは Kumar ら [99] の (α, β) モデルにより構成した．

ここでは，ネットワークのリンク構造を表している隣接行列 A に対して修正 Oja アルゴリズムを用いて第 1，第 2，第 3 主成分と第 1，第 2，第 3 マイナー成分を抽出した．また，隣接行列のサイズは (550×550) となっている．

まず，(α, β) モデルによって生成したネットワーク内に存在する各ノードの in-degree および out-degree に対するノードの数の分布を示す．

以下で，

- (a) authority link の解析
- (b) hub link の解析

の 2 つの問題について，本提案手法を適用した結果を示す．

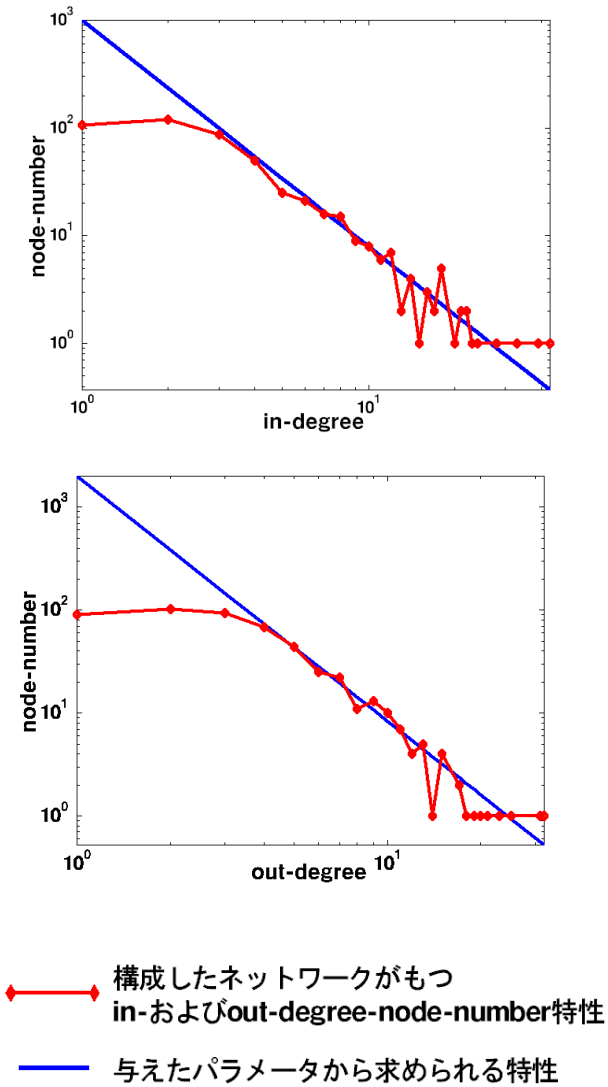

図 26: authrity link および hub link に対するノードの個数の分布

(a) authorty link に関する主成分およびマイナー成分抽出の結果を図 27 に示す．

10 万回のステップ数で求めた解ベクトルを用いて Reyleigh 商 (6) により計算した近似固有値は $\lambda_1 = 100.9, \lambda_2 = 46.42, \lambda_3 = 41.41, \mu_1 = 0.0017, \mu_2 = 0.0021, \mu_3 = 0.0025$ となった．

図 27 から，

- 主成分については大きな in-degree をもつノードに対して主成分固有ベクトルの大きな値をもつ要素が対応している．これは，Kleinberg の結果とほぼ一致する．
- マイナー成分については主成分とは逆に小さな in-degree(0 を除く) をもつノードに対してマイナー成分固有ベクトルの大きな値をもつ要素が対応している場合が多い．

という結果が得られた．

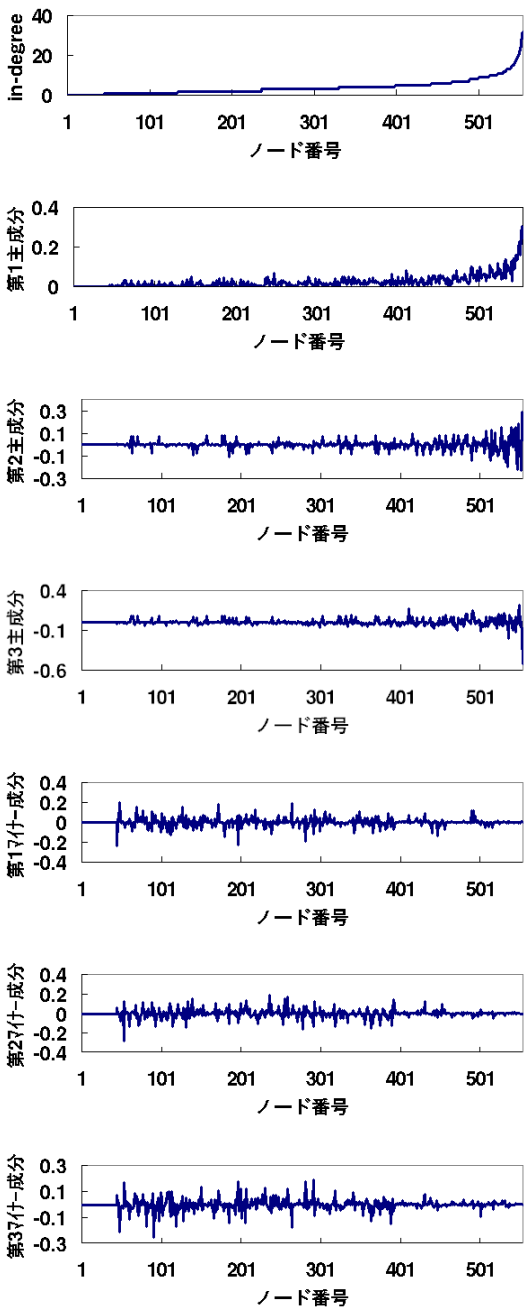

図 27: in-degree と authority link の主成分とマイナー成分

(b) hub link に関する主成分およびマイナー成分抽出の結果を図 28 に示す．

このとき，Reylegh 商により求めた近似固有値は $\lambda_1 = 100.9, \lambda_2 = 46.42, \lambda_3 = 41.41, \mu_1 = 1.7 \times 10^{-5}, \mu_2 = 3.3 \times 10^{-5}, \mu_3 = 6.0 \times 10^{-6}$ となっている．

図 28 から，

- 主成分については authority link と同様に大きな out-degree をもつノードに対して主成分固有ベクトルの大きな値をもつ要素が対応している．これは，Kleinberg の結果とほぼ一致する．
- マイナー成分については主成分とは逆に小さな in-degree をもつノードに対してマイナー成分固有ベクトルの大きな値をもつ要素が対応している場合が多い．

という結果が得られた．

authority link と hub link のマイナー成分の関係は過去に調べられていないので更に考察が必要である．また，hub link のマイナー成分については，in-degree の場合と異なり，out-degree が 0 であるノードに対しても大きな値を持つ要素が存在している．hub link に関する MCA の 10 万ステップ終了時の誤差ノルム ρ_k と η_k の値はそれぞれ，$\rho_k = 2.0 \times 10^{-5}$，$\eta_k = 1.0 \times 10^{-6}$ となっていた．ρ_k の計算には真の固有値の代わりにステップ毎に $w_n, (n = 1, 2, 3)$ を用いて Reyleigh 商により求めた近似固有値を用いた．しかしながら，固有値が昇順に求められていないことから，誤差の大きな結果が得られていると考えられる．この原因は，MCA の対象となった行列のマイナー成分に対応する固有値がお互いに近いことと，問題の行列が提案手法の反復計算において，数値誤差が大きくなりやすい構造をもっているためであると推測される．

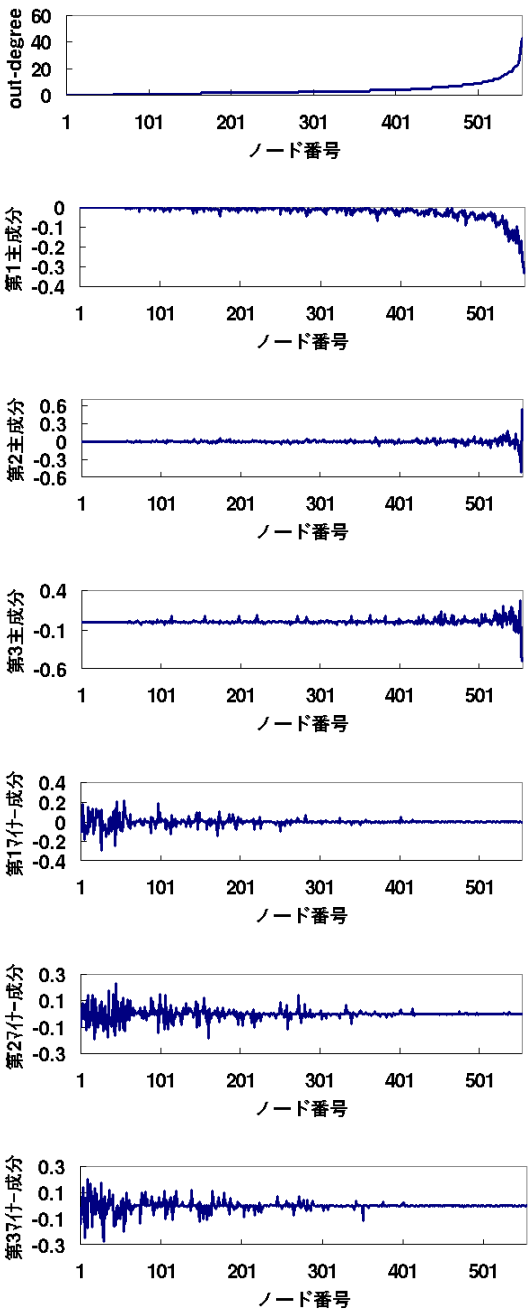

図 28: out-degree と hub link の主成分とマイナー成分

5　おわりに

本研究では，正負の符号の違いのみで主成分とマイナー成分を抽出できるとともに，数知的に安定なアルゴリズムを提案した．

従来の手法では，

1. マイナー成分の抽出における数値安定化のために正規直交行列の空間上の力学系を考えている．

2. 各々の固有ベクトルを抽出するためには，Gram-Schmidt の正規直交化に対応するようなアルゴリズムの非対称化が必要である．ただし，上記の手法 (W^TW の積による正規直交化の強化) では非対称化によって $W^TW = I$ が崩れる．

という相反する条件が必要であったが，本研究で提案したアルゴリズムは Gram-Schmidt の直交化法に対応する非対称化と，数値安定化のためのペナルティ項を組み合わせることで上の問題を回避している．

また，既存の主成分およびマイナー成分を抽出するアルゴリズムでは利用可能なパラメータの範囲を系統的に調べられてないので，平衡点近傍での漸近解析と数値実験により利用可能なパラメータの範囲を調べた．

漸近解析ではマイナー成分の抽出においてはペナルティ項に与えるゲイン γ が少なくとも，抽出されるマイナー成分に対応する固有値よりも大きくないと解がマイナー成分固有ベクトルに収束しないという結果が得られた．

数値実験では，行列 D が定理の条件を満たしていたとしても，パラメータ $\theta_n, (n = 1, ..., N)$ の比により利用可能なパラメータの範囲が変化し，θ_n の比が大きい場合には，その範囲が狭くなった．これは漸近解析の結果が大域的には十分条件になっていないことを示すものである．

さらに，より現実的な規模の適用例としてネットワークのリンク構造解析に対して，提案したアルゴリズムを用いた．

得られた主成分は接続リンク数が多いノードに対して大きな値をもつ要素が対応しており，Kleinberg とほぼ一致した結果が得られた．

一方，マイナー成分は接続リンク数が少ないノードに対して，大きな値をもつ要素が対応しているという傾向が見られた．しかしながら，各マイナー成分とリンク構造の関係については明確なものではなく，更に考察が必要である．

A　定理 1 の証明

【定理 1 】

K 個の主成分である固有ベクトルを列ベクトルとした行列

$$\tilde{V} = [v_1 v_2 ... v_K], ||v_i||^2 = 1, i = 1, ..., K$$

と，それぞれに対応する固有値

$$\lambda_1, \lambda_2, ..., \lambda_K, \lambda_1 > \lambda_2 > ... > \lambda_N \geq \lambda_{N+1} \geq ... \geq \lambda_K \geq 0, (K > N)$$

をもつ対角行列

$$\tilde{\Lambda} = diag(\lambda_1, \lambda_2, ..., \lambda_K)$$

で表される正定値対称行列

$$C = \tilde{V}\tilde{\Lambda}\tilde{V}^T \in \mathbf{R}^{K \times K}$$

と対角行列

$$D = diag(\theta_1, ..., \theta_N) \in \mathbf{R}^{N \times N}$$

で定義された

$$\frac{dZ}{dt} \equiv CZ - ZD^{-1}Z^T CZD + \gamma Z(I - Z^T Z) \tag{39}$$

の解

$$Z \in \mathbf{R}^{K \times N}$$

は，

$$V = [v_1 v_2 ... v_N]$$

で与えられ，

$$0 < \theta_1 < \theta_2 < ... < \theta_N \tag{40}$$

かつ

$$\gamma > 0 \tag{41}$$

の場合に，漸近的に安定である．

【証明】

$$V = [v_1 ... v_N]$$

に対応する固有値からなる対角行列

$$\Lambda = diag(\lambda_1, ..., \lambda_N)$$

を考える．そのとき，V は直交行列になるので，

$$CV = V\Lambda \tag{42}$$
$$V^T CV = \Lambda \tag{43}$$

となる．

ここで，式 (39) の右辺に，この定数行列 V を代入すると，

$$\begin{aligned}
(\text{右辺}) &= CV - VD^{-1}(V^T CV)D \\
&= V\Lambda - VD^{-1}\Lambda D \\
&= V\Lambda - V\Lambda D^{-1}D \\
&= V\Lambda - V\Lambda \\
&= 0
\end{aligned} \tag{44}$$

次に，$\bar{Z}$ の漸近安定性を証明するために，平衡点 $\bar{Z}$ からの微小な摂動 $E(t) = [e_1(t)...e_n(t)]$ を考える．

$$\begin{aligned}
Z(t) &= \bar{Z} + E(t) \\
&= V + E(t)
\end{aligned}$$

を式 (39) に代入すると，$E(t)$ に対して次の式を得る．

$$\begin{aligned}
\frac{dE}{dt} &= \frac{dZ}{dt}|_{Z=V+E} \\
&= C(V+E) \\
&\qquad -(V+E)D^{-1}(V+E)^T C(V+E)D \\
&\qquad +\gamma(V+E)[I-(V+E)^T(V+E)] \\
&= [CV + CE \\
&\quad -VD^{-1}V^T CVD - VD^{-1}V^T CED \\
&\qquad -VD^{-1}E^T CVD - ED^{-1}V^T CVD] \\
&\quad -\gamma[V^T VE - VE^T V] + O\Big(||E||^2\Big)
\end{aligned}$$

ここで，$O(||E||^2)$ は E の要素の 2 次以上の項を示している．$||E||^2$ が十分に小さいと仮定すると，式 (42) と式 (43) を用いて次の式が得られる．

$$\begin{aligned}
\frac{dE}{dt} &= CV + CE \\
&\quad -VD^{-1}V^TCVD - VD^{-1}V^TCED \\
&\qquad -VD^{-1}E^TCVD - ED^{-1}V^TCVD \\
&\qquad -\gamma E - \gamma VE^TV \\
&= V\Lambda + CE - VD^{-1}\Lambda D - VD^{-1}\Lambda V^TED \\
&\qquad -VD^{-1}E^TV\Lambda D - ED^{-1}\Lambda D \\
&\qquad -\gamma E - \gamma VE^TV \\
&= V\Lambda + CE - V\Lambda - VD^{-1}\Lambda V^TED - VD^{-1}E^TV\Lambda D - E\Lambda \\
&\qquad -\gamma E - \gamma VE^TV \\
&= (C-\gamma I)E - E\Lambda - VD^{-1}\Lambda V^TED \\
&\qquad -VD^{-1}E^TV\Lambda D - \gamma VE^TV \qquad (45)
\end{aligned}$$

E の n 番目の列ベクトル e_n に対して、式 (45) は

$$\begin{aligned}
\frac{de_n}{dt} &= [C-(\gamma+\lambda_n)I]e_n - \left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\lambda_m\Big)v_mv_m^T\right)e_n \\
&\qquad -\left(\sum_{m=1}^{N}\Big(\frac{\lambda_n}{\theta_m}\theta_n - \gamma\Big)v_me_m^T\right)v_n \qquad (46)
\end{aligned}$$

となる．

(46) の両辺に，左から固有ベクトル $v_r, r = 1, ..., K$ を掛け，以下の三つの場合に分けて議論する．

(i) まず，$r > N$ とする．このとき，実対称行列 C の固有ベクトル v_r の正規直交性により次のようになる．

$$\begin{aligned}
\frac{d}{dt}(v_r^Te_n) &= v_r^T[C-(\lambda_n+\gamma)I]e_n - v_r^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\lambda_m\Big)v_mv_m^T\right)e_n \\
&\qquad -v_r^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\lambda_n - \gamma\Big)v_me_m^T\right)v_n \qquad (47) \\
&= \Big(\lambda_r - \lambda_n - \gamma\Big)v_r^Te_n \qquad (48)
\end{aligned}$$

したがって，$v_r^Te_n$ は $\lambda_r - \lambda_n - \gamma < 0$ のとき漸近的に 0 に近づく．

この結果は条件 (40) に依らず，そして θ_n がすべて等しい場合においても正しい．

(ii) つぎに，$r = n, r, n \leq N$ とすると，

$$\begin{aligned}
\frac{d}{dt}(v_n^T e_n) &= v_n^T[C - (\lambda_n + \gamma)I]e_n - v_n^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\lambda_m\Big)v_m v_m^T\right)e_n \\
&\qquad -v_n^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\lambda_n - \gamma\Big)v_m e_m^T\right)v_n \\
&= \Big(\lambda_n - \lambda_n - \gamma\Big)(v_n^T e_n) - \lambda_n(v_n^T e_n) - \left(\frac{\lambda_n\theta_n}{\theta_n} - \gamma\right)(e_n^T v_n) \\
&= -(\lambda_n + \lambda_n)(v_n^T e_n) \\
&= -2\lambda_n(v_n^T e_n) \qquad (49)
\end{aligned}$$

$\lambda_n > 0$ より，この式は漸近的に 0 に近づく．

(iii) 最後に，$r \neq n, (r, n \leq N)$ とする．一般性を失うこと無く，$r < n$ と仮定することができる．式 (47) より，次の式を得る．

$$\begin{aligned}
\frac{d}{dt}(v_r^T e_n) &= v_r^T[C - (\lambda_n + \gamma)I]e_n - v_r^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\lambda_m\Big)v_m v_m^T\right)e_n \\
&\qquad -v_r^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\lambda_n - \gamma\Big)v_m e_m^T\right)v_n \\
&= \left[\Big(1 - \frac{\theta_n}{\theta_r}\Big)\lambda_r - \lambda_n - \gamma\right](v_r^T e_n) \\
&\qquad -\Big(\frac{\theta_n}{\theta_r}\lambda_n - \gamma\Big)(v_n^T e_r) \qquad (50)
\end{aligned}$$

また，上式の r と n を置換すると，次のようになる．

$$\frac{d}{dt}(v_n^T e_r) = \left[\Big(1 - \frac{\theta_r}{\theta_n}\Big)\lambda_n - \lambda_r - \gamma\right](v_n^T e_r) - \Big(\frac{\theta_r}{\theta_n}\lambda_r - \gamma\Big)(v_r^T e_n) \qquad (51)$$

更に，式 (50) より，

$$\begin{aligned}
(v_n^T e_r) &= -\Big(\frac{\theta_n}{\theta_r}\lambda_n - \gamma\Big)^{-1} \\
&\qquad \times\left[\frac{d}{dt}(v_r^T e_n) - \Big\{\Big(1 - \frac{\theta_n}{\theta_r}\Big)\lambda_r - \lambda_n - \gamma\Big\}(v_r^T e_n)\right] \qquad (52)
\end{aligned}$$

この結論も式 (16) とは独立である．そして，すべての θ_n が等しい，または正値をとる場合においても成り立つ．

式 (52) を式 (51) の右辺に代入すると，

$$\begin{aligned}\frac{d}{dt}(v_n^T e_r) &= -\Big(\frac{\theta_n}{\theta_r}\lambda_n-\gamma\Big)^{-1}\Big[\Big(1-\frac{\theta_r}{\theta_n}\Big)\lambda_n-\lambda_r-\gamma\Big]\\ &\qquad\times\Big[\frac{d}{dt}(v_r^T e_n)-\Big\{\Big(1-\frac{\theta_n}{\theta_r}\Big)\lambda_r-\lambda_n-\gamma\Big\}(v_r^T e_n)\Big]\\ &\quad -\Big(\frac{\theta_r}{\theta_n}\lambda_r-\gamma\Big)(v_r^T e_n) \qquad (53)\end{aligned}$$

となる．

$(v_r^T e_n)$ のみに関する方程式を得るために，式 (50) を時間 t で微分して，

$$\frac{d^2}{dt^2}(v_r^T e_n)=\Big[\Big(1-\frac{\theta_n}{\theta_r}\Big)\lambda_r-\lambda_n-\gamma\Big]\frac{d}{dt}(v_r^T e_n)-\Big(\frac{\theta_n}{\theta_r}\lambda_n-\gamma\Big)\frac{d}{dt}(v_n^T e_r)\,(54)$$

を得る．上式の右辺に，式 (53) を代入すると，次式を得る．

$$\begin{aligned}\frac{d^2}{dt^2}(v_r^T e_n) &= \Big[\Big(1-\frac{\theta_n}{\theta_r}\Big)\lambda_r-\lambda_n-\gamma\Big]\frac{d}{dt}(v_r^T e_n)+\Big[\Big(1-\frac{\theta_r}{\theta_n}\Big)\lambda_n-\lambda_r-\gamma\Big]\frac{d}{dt}(v_r^T e_n)\\ &\quad -\Big[\Big\{\Big(1-\frac{\theta_n}{\theta_r}\Big)\lambda_r-\lambda_n-\gamma\Big\}\Big\{\Big(1-\frac{\theta_r}{\theta_n}\Big)\lambda_n-\lambda_r-\gamma\Big\}\\ &\qquad -\Big(\frac{\theta_n}{\theta_r}\lambda_n-\gamma\Big)\Big(\frac{\theta_r}{\theta_n}\lambda_r-\gamma\Big)\Big](v_r^T e_n) \qquad (55)\end{aligned}$$

ここで，

$$\beta\equiv\frac{\theta_n}{\theta_r},\qquad \epsilon\equiv\frac{\lambda_n}{\lambda_r}$$

とすると，式 (55) は

$$\begin{aligned}\frac{d^2}{dt^2}(v_r^T e_n) &= \Big[(1-\beta)\lambda_r-\epsilon\lambda_r-\gamma\Big]\frac{d}{dt}(v_r^T e_n)+\Big[(1-\beta^{-1})\epsilon\lambda_r-\lambda_r-\gamma\Big]\frac{d}{dt}(v_r^T e_n)\\ &\quad -\Big[\Big\{(1-\beta)\lambda_r-\epsilon\lambda_r-\gamma\Big\}\Big\{(1-\beta^{-1})\epsilon\lambda_r-\lambda_r-\gamma\Big\}\\ &\qquad -(\beta\epsilon\lambda_r-\gamma)(\beta^{-1}\lambda_r-\gamma)\Big](v_r^T e_n)\\ &= \Big[(1-\beta)\lambda_r-\epsilon\lambda_r-\gamma+\{(1-\beta^{-1})\epsilon-1\}\lambda_r-\gamma\Big]\frac{d}{dt}(v_r^T e_n)\\ &\quad -\Big[\Big\{(1-\beta)\lambda_r-\epsilon\lambda_r-\gamma\Big\}\Big\{(1-\beta^{-1})\epsilon\lambda_r-\lambda_r-\gamma\Big\}\\ &\qquad -(\beta\epsilon\lambda_r-\gamma)(\beta^{-1}\lambda_r-\gamma)\Big](v_r^T e_n)\\ &= -\Big[(\beta+\beta^{-1}\epsilon)\lambda_r+2\gamma\Big]\frac{d}{dt}(v_r^T e_n)\\ &\quad -\Big[\Big\{(\beta^{-1}-1)\epsilon^2-(\beta-3+\beta^{-1})\epsilon+(\beta-1)\Big\}\\ &\qquad +(\beta+\beta^{-1}\epsilon)\lambda_r\gamma+\gamma^2\Big](v_r^T e_n)\end{aligned}$$

となる．これは定数係数 2 階線形微分方程式である．この解は二つの特性根が負の実部をもつ場合，またその場合に限り 0 に収束する．上式の特性方程式の解を s として求める．

$$\begin{aligned} s &= (2\beta)^{-1}[-2\beta\gamma - \lambda_r\beta^2 - \lambda_n \\ &\quad \pm\{4\beta^2\gamma^2 - 4(\beta^3\lambda_n + \beta\lambda_r)\gamma \\ &\quad +\lambda_r^2\beta^4 + 4\lambda_r(\lambda_n - \lambda_r)\beta^3 \\ &\quad +(4\lambda_r^2 - 6\lambda_r\lambda_n + 4\lambda_n^2)\beta^2 \\ &\quad +4\lambda_n(\lambda_r - \lambda_n)\beta + \lambda_n^2\}^{-\frac{1}{2}}] \end{aligned} \tag{56}$$

となる．よって，式 (55) が 0 に収束するためには，

$$\begin{aligned} s_1 &= (2\beta)^{-1}[-2\beta\gamma - \lambda_r\beta^2 - \lambda_n \\ &\quad +\{4\beta^2\gamma^2 - 4(\beta^3\lambda_n + \beta\lambda_r)\gamma \\ &\quad +\lambda_r^2\beta^4 + 4\lambda_r(\lambda_n - \lambda_r)\beta^3 \\ &\quad +(4\lambda_r^2 - 6\lambda_r\lambda_n + 4\lambda_n^2)\beta^2 \\ &\quad +4\lambda_n(\lambda_r - \lambda_n)\beta + \lambda_n^2\}^{-\frac{1}{2}}] \\ &< 0 \end{aligned} \tag{57}$$

が条件となる．上の不等式を γ について Maple を用いて解くと，

$$\gamma > \frac{(\epsilon-1)(\beta-1)(\beta+\epsilon)}{(\epsilon+1)(\beta^2+1)}\lambda_r \equiv I \tag{58}$$

となる．ここで，

$$\begin{aligned} &\lambda_1 > \lambda_2 > ... > \lambda_N \geq \lambda_{N+1} = ... = \lambda_K = 0, \\ &0 < \theta_1 < \theta_2 < ... < \theta_N \end{aligned}$$

であることから，

$$0 < \epsilon < 1 \text{ のとき} \beta > 1 \text{，また} \epsilon > 1 \text{ のとき } 0 < \beta < 1 \tag{59}$$

なので，

$$(\epsilon - 1)(\beta - 1) < 0$$

となる．また，$(\beta^2+1)>0$ であり，

$$\begin{aligned}\frac{\beta+\epsilon}{\epsilon+1} &= \frac{\beta+\epsilon+1-1}{\epsilon+1}\\ &= \frac{\beta-1}{\epsilon+1}+1\\ &> 0\end{aligned}$$

となる．ゆえに $I<0$ となるので，$\gamma>0$ であれば式 (58) を満足する．

【証明終】

B　定理2の証明

【定理 2 】

k 個のマイナー成分である固有ベクトルを列ベクトルとした行列

$$\tilde{V} = [v_1 v_2 ... v_K], ||v_i||^2 = 1, i = 1, ..., K$$

とその固有値

$$\mu_1, \mu_2, ..., \mu_K, 0 < \mu_1 < \mu_2 < ... < \mu_N \leq \mu_{N+1} \leq ... \leq \mu_K (K > N)$$

をもつ対角行列

$$\tilde{M} = diag(\mu_1, \mu_2, ..., \mu_K)$$

で表される正定値対称行列

$$C = \tilde{V}\tilde{M}\tilde{V}^T \in \mathbf{R}^{K \times K}$$

と対角行列

$$D = diag(\theta_1, ..., \theta_N) \in \mathbf{R}^{N \times N}$$

で定義された

$$\frac{dZ}{dt} \equiv -CZ + ZD^{-1}Z^T CZD + \gamma Z(I - Z^T Z) \tag{60}$$

の解

$$Z \in \mathbf{R}^{K \times N}$$

は，

$$V = [v_1 v_2 ... v_N]$$

で与えられ，

$$0 < \theta_1 < \theta_2 < ... < \theta_N \tag{61}$$

かつ

$$\gamma > \mu_N \tag{62}$$

の場合に，漸近的に安定である．

【証明】

$$V = [v_1 ... v_N]$$

に対応する固有値からなる対角行列

$$M = diag(\mu_1, ..., \mu_N)$$

を考える．そのとき，V は直交行列になるので，

$$CV = VM \tag{63}$$
$$V^T CV = M \tag{64}$$

となる．

ここで，式 (60) の右辺に，この定数行列 V を代入すると，

$$\begin{aligned}(\text{右辺}) &= -CV + VD^{-1}(V^T CV)D \\ &= -VM + VD^{-1}MD \\ &= -VM + VMD^{-1}D \\ &= -VM + VM \\ &= 0 \end{aligned} \tag{65}$$

次に，$\bar{Z}$ の漸近安定性を証明するために，平衡点 $\bar{Z}$ からの微小な摂動 $E(t) = [e_1(t) ... e_n(t)]$ を考える．

$$\begin{aligned} Z(t) &= \bar{Z} + E(t) \\ &= V + E(t) \end{aligned}$$

を式 (60) に代入すると，$E(t)$ に対して次の式を得る．

$$\begin{aligned} \frac{dE}{dt} &= \frac{dZ}{dt}|_{Z=V+E} \\ &= -C(V+E) \\ &\qquad +(V+E)D^{-1}(V+E)^T C(V+E)D \\ &\qquad +\gamma(V+E)[I-(V+E)^T(V+E)] \\ &= [-CV - CE \\ &\quad +VD^{-1}V^T CVD + VD^{-1}V^T CED \\ &\qquad +VD^{-1}E^T CVD + ED^{-1}V^T CVD] \\ &\quad -\gamma[V^T VE - VE^T V] + O\Big(||E||^2\Big) \end{aligned}$$

ここで，$O(||E||^2)$ は E の要素の 2 次以上の項を示している．$||E||^2$ が十分に小さいと仮定すると，式 (63) と式 (64) を用いて次の式が得られる．

$$\begin{aligned}
\frac{dE}{dt} &= -CV - CE \\
&\quad +VD^{-1}V^TCVD + VD^{-1}V^TCED \\
&\qquad +VD^{-1}E^TCVD + ED^{-1}V^TCVD \\
&\qquad -\gamma E - \gamma VE^TV \\
&= -VM - CE + VD^{-1}MD + VD^{-1}MV^TED \\
&\qquad +VD^{-1}E^TVMD + ED^{-1}MD \\
&\qquad -\gamma E - \gamma VE^TV \\
&= -VM - CE - VM + VD^{-1}MV^TED + VD^{-1}E^TVMD + EM \\
&\qquad -\gamma E - \gamma VE^TV \\
&= -(C+\gamma I)E + EM + VD^{-1}MV^TED \\
&\qquad +VD^{-1}E^TVMD - \gamma VE^TV \qquad (66)
\end{aligned}$$

E の n 番目の列ベクトル e_n に対して、式 (66) は

$$\begin{aligned}
\frac{de_n}{dt} &= -[C + (\gamma - \mu_n)I]e_n + \left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\mu_m\Big)v_mv_m^T\right)e_n \\
&\qquad +\left(\sum_{m=1}^{N}\Big(\frac{\mu_n}{\theta_m}\theta_n - \gamma\Big)v_me_m^T\right)v_n \qquad (67)
\end{aligned}$$

となる．

(67) の両辺に，左から固有ベクトル $v_r, r = 1, ..., K$ を掛け，以下の三つの場合に分けて議論する．

(i) まず，$r > N$ とする．このとき，実対称行列 C の固有ベクトル v_r の正規直交性により次のようになる．

$$\begin{aligned}
\frac{d}{dt}(v_r^Te_n) &= -v_r^T[C + (\gamma - \mu_n)I]e_n + v_r^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\mu_m\Big)v_mv_m^T\right)e_n \\
&\qquad +v_r^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\mu_n - \gamma\Big)v_me_m^T\right)v_n \qquad (68) \\
&= -\Big(\mu_r - \mu_n + \gamma\Big)v_r^Te_n \qquad (69)
\end{aligned}$$

したがって，$\mu_n < \mu_r, \gamma > 0$ により $\mu_r - \mu_n - \gamma < 0$ となり、$v_r^Te_n$ は漸近的に 0 に近づく．

この結果は条件 (16) に依らず，そして θ_n がすべて等しい場合においても正しい．

(ii) つぎに，$r = n, (r, n \leq N)$ とすると，

$$\begin{aligned}
\frac{d}{dt}(v_n^T e_n) &= -v_n^T[C + (\gamma - \mu_n)I]e_n + v_n^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\mu_m\Big)v_m v_m^T\right)e_n \\
&\qquad + v_n^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\mu_n - \gamma\Big)v_m e_m^T\right)v_n \\
&= \Big\{-(\mu_n - \mu_n + \gamma) + \lambda_n + \lambda_n - \gamma\Big\}(v_n^T e_n) \\
&= 2(\mu_n - \gamma)(v_n^T e_n) \qquad (70)
\end{aligned}$$

よって，仮定により $\mu_n - \gamma < 0$ なので，この式は漸近的に 0 に近づく．

(iii) 最後に，$r \neq n$ とする．一般性を失うこと無く，$r < n$ と仮定することができる．式 (68) より，次の式を得る．

$$\begin{aligned}
\frac{d}{dt}(v_r^T e_n) &= -v_r^T[C + (\gamma - \mu_n)I]e_n + v_r^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\mu_m\Big)v_m v_m^T\right)e_n \\
&\qquad + v_r^T\left(\sum_{m=1}^{N}\Big(\frac{\theta_n}{\theta_m}\mu_n - \gamma\Big)v_m e_m^T\right)v_n \\
&= \left[\Big(\frac{\theta_n}{\theta_r} - 1\Big)\mu_r + \mu_n - \gamma\right](v_r^T e_n) \\
&\qquad + \Big(\frac{\theta_n}{\theta_r}\mu_n - \gamma\Big)(v_n^T e_r) \qquad (71)
\end{aligned}$$

また，上式の r と n を置換すると，次のようになる．

$$\frac{d}{dt}(v_n^T e_r) = \left[\Big(\frac{\theta_r}{\theta_n} - 1\Big)\mu_n + \mu_r - \gamma\right](v_n^T e_r) - \Big(\frac{\theta_r}{\theta_n}\mu_r - \gamma\Big)(v_r^T e_n) \qquad (72)$$

更に，式 (71) より，

$$(v_n^T e_r) = \Big(\frac{\theta_n}{\theta_r}\mu_n - \gamma\Big)^{-1}\left[\frac{d}{dt}(v_r^T e_n) - \Big\{\Big(\frac{\theta_n}{\theta_r} - 1\Big)\mu_r + \mu_n - \gamma\Big\}(v_r^T e_n)\right] \qquad (73)$$

式 (73) を式 (72) の右辺に代入すると，

$$\begin{aligned}
\frac{d}{dt}(v_n^T e_r) &= \Big(\frac{\theta_n}{\theta_r}\mu_n - \gamma\Big)^{-1}\left[\Big(\frac{\theta_r}{\theta_n} - 1\Big)\mu_n + \mu_r - \gamma\right] \\
&\qquad \times\left[\frac{d}{dt}(v_r^T e_n) - \Big\{\Big(\frac{\theta_n}{\theta_r} - 1\Big)\mu_r + \mu_n - \gamma\Big\}(v_r^T e_n)\right] \\
&\quad + \Big(\frac{\theta_r}{\theta_n}\mu_r - \gamma\Big)(v_r^T e_n) \qquad (74)
\end{aligned}$$

この結論も式 (16) とは独立である．そして，すべての θ_n が等しい，または正値をとる場合においても成り立つ．

となる．

$(v_r^T e_n)$ のみに関する方程式を得るために，式 (71) を時間 t で微分して，

$$\begin{aligned}\frac{d^2}{dt^2}(v_r^T e_n) &= \Big[\Big(\frac{\theta_n}{\theta_r}-1\Big)\mu_r+\mu_n-\gamma\Big]\frac{d}{dt}(v_r^T e_n)\\ &\quad+\Big(\frac{\theta_n}{\theta_r}\mu_n-\gamma\Big)\frac{d}{dt}(v_n^T e_r) \qquad (75)\end{aligned}$$

を得る．上式の右辺に，式 (74) を代入すると，次式を得る．

$$\begin{aligned}\frac{d^2}{dt^2}(v_r^T e_n) &= \Big[\Big(\frac{\theta_n}{\theta_r}-1\Big)\mu_r+\mu_n-\gamma+\Big(\frac{\theta_r}{\theta_n}-1\Big)\mu_n+\mu_r-\gamma\Big]\frac{d}{dt}(v_r^T e_n)\\ &\quad-\Big[\Big\{\Big(\frac{\theta_n}{\theta_r}-1\Big)\mu_r+\mu_n-\gamma\Big\}\Big\{\Big(\frac{\theta_r}{\theta_n}-1\Big)\mu_n+\mu_r-\gamma\Big\}\\ &\quad-\Big(\frac{\theta_n}{\theta_r}\mu_n-\gamma\Big)\Big(\frac{\theta_r}{\theta_n}\mu_r-\gamma\Big)\Big](v_r^T e_n) \qquad (76)\end{aligned}$$

ここで，

$$\beta\equiv\frac{\theta_n}{\theta_r},\qquad \epsilon\equiv\frac{\mu_n}{\mu_r}$$

とすると，式 (76) は

$$\begin{aligned}\frac{d^2}{dt^2}(v_r^T e_n) &= \Big[(\beta-1)\mu_r+\epsilon\mu_r-\gamma+(\beta^{-1}-1)\epsilon\mu_r+\mu_r-\gamma\Big]\frac{d}{dt}(v_r^T e_n)\\ &\quad-\Big[\Big\{(\beta-1)\mu_r+\epsilon\mu_r-\gamma\Big\}\Big\{(\beta^{-1}-1)\epsilon\mu_r+\mu_r-\gamma\Big\}\\ &\quad-(\beta\epsilon\mu_r-\gamma)(\beta^{-1}\mu_r-\gamma)\Big](v_r^T e_n)\\ &= \Big[(\beta+\beta^{-1}\epsilon)\lambda_r-2\gamma\Big]\frac{d}{dt}(v_r^T e_n)\\ &\quad-\Big[\Big\{(\beta^{-1}-1)\epsilon^2-(\beta-2+\beta^{-1})\epsilon-1\Big\}\mu_r^2\\ &\quad+\Big\{(\beta-\beta^{-1})\epsilon+\beta^{-1}-\beta\Big\}\mu\gamma\Big](v_r^T e_n) \qquad (77)\end{aligned}$$

となる．これは定数係数 2 階線形微分方程式である．この解は二つの特性根が負の実部をもつ場合，またその場合に限り 0 に収束する．上式の特性方程式の解を s として求める．

$$\begin{aligned}s &= (2\beta)^{-1}[-2\beta\gamma+(\beta^2+1)\mu_r\\ &\quad\pm\{4\beta^2\gamma^2-4(\beta^3\epsilon+\beta)\mu_r\gamma\\ &\quad+\mu_r^2(\beta^4+4(\epsilon-1)\beta^3+\\ &\quad(4-6\epsilon+4\epsilon^2)\beta^2+4\epsilon(1-\epsilon)\beta+\epsilon^2)\}^{-\frac{1}{2}}] \qquad (78)\end{aligned}$$

となる．よって，式 (76) が 0 に収束するためには，

$$\begin{aligned} s_1 &= (2\beta)^{-1}[-2\beta\gamma + (\beta^2+1)\mu_r \\ &\quad +\{4\beta^2\gamma^2 - 4(\beta^3\epsilon+\beta)\mu_r\gamma \\ &\quad +\mu_r^2(\beta^4 + 4(\epsilon-1)\beta^3 + \\ &\quad (4-6\epsilon+4\epsilon^2)\beta^2 + 4\epsilon(1-\epsilon)\beta + \epsilon^2)\}^{-\frac{1}{2}}] \\ &< 0 \end{aligned} \tag{79}$$

が条件となる．上の不等式を γ について解くと，

$$\gamma > \frac{\beta+\epsilon}{\beta+1}\mu_r \equiv I \tag{80}$$

となる．仮定，

$$\begin{aligned} &\mu_1 < \mu_2 < ... < \mu_N, \\ &0 < \theta_1 < \theta_2 < ... < \theta_N \end{aligned}$$

より，

$$0 < \epsilon < 1 \text{ のとき } 0 < \beta < 1 \text{，また} \epsilon > 1 \text{ のとき} \beta > 1 \tag{81}$$

となる．

(a) $0 < \epsilon < 1, 0 < \beta < 1$ の場合

$$\beta + \epsilon < \beta + 1$$

なので，

$$I = \frac{\beta+\epsilon}{\beta+1}\mu_r < \mu_r$$

となる．

(b) $\epsilon > 1, \beta > 1$ の場合

$$\frac{\beta+\epsilon}{\beta+1} < \frac{\beta\epsilon+\epsilon}{\beta+1} = \epsilon$$

なので，

$$I = \frac{\beta+\epsilon}{\beta+1}\mu_r < \epsilon\mu_r \leq \mu_N$$

となる．ゆえに，$\gamma > \mu_N$ であれば式 (80) を満足する．

【証明終】

参考文献

[1] Oja, E., Ogawa, H., & Wangviwattana, J.(1992a) "Principal Component Analysis by Homogeneous Neural Networks, Part II: Analysis an d Extentions of the Learning Algorithm" IEICE TRANS. INF. & SYST., Vol.E75-D, pp.376-381, No.3

[2] Oja, E.(1992b) "Principal Components, Minor Components, andd Linear Neural Networks" Neural Networks, Vol.5, pp.927-935

[3] Wang, L., & Kalhunen. J.(1996) "A Unified Neural Bigradient Algorithm for Robust PCA and MCA" International Journal of Neural Systems, Vol.7, No.1, pp.53-67

[4] Chen, T.(1997) "Modified Oja's Algorithms for Principal Subspace and Minor Subspace Extraction" Neural Processing Letters, Vol.5, pp.105-110

[5] Sanger, T(1989) "Optional Unspervised Learning in a Single-Layer Linear Feedforward Neural Network" Neural Networks, Vol.2, pp.459-473

[6] Douglas, S., Kung, S., & Amari, S.(1998) "A Self-Stabilized Minor Subspace Rule" IFFF Signal Processing Letters, Vol.5, No.12, pp.328-331

[7] Luo, L., Unbehauen, R., & Cichocki, A.(1997) "A Minor Component Analysis Algorithm" Neural Networks, Vol.10, No.2, pp.291-297

[8] Kumar, R., Raghaven, P., Rajagopalan, S., & Tomkins, A.(1999) "Extracting large-scale knowledge bases from the web" IEEE Int. Conf. on VLDB

[9] Kleinberg, J(1998) "Authoritative Sources in a Hyperlinked Environment" Proc.of the 9th ACM-SIAM Symp. on Doc. Algo, pp.668-677

[10] Anderson, E., et al , (小国 力 訳)(1995) "LAPACK 利用の手引" 丸善