論　文

標本共分散行列の固有ベクトルを用いた真のマハラノビス距離の推定法

岩村　雅一[†]　　大町真一郎[†]　　阿曽　弘具[†]

A Method to Estimate the True Mahalanobis Distance from Eigenvectors of Sample Covariance Matrix

Masakazu IWAMURA[†], Shinichiro OMACHI[†], and Hirotomo ASO[†]

あらまし　パターン認識で用いられる統計的識別関数は，パターンの確率分布が正しく与えられたときに最適になる．パターンの確率分布は一般に未知であるため，正規分布を仮定して，その平均ベクトルと共分散行列は学習サンプルからの推定値を用いることが多い．しかし，推定に用いる学習サンプルが不足すると推定値が誤差を含み，認識性能が低下することが知られている．平均ベクトルや共分散行列の固有値の推定誤差についてはよく調査され，認識性能の低下を回避する方法が提案されている．しかし，共分散行列の固有ベクトルの推定誤差についてはこれまでほとんど考慮されてこなかった．本論文では，学習サンプルが十分用意できない場合においても高精度な認識を行うことを目的とし，共分散行列の固有ベクトルに誤差がある場合でも，2 次識別関数の主要部分であるマハラノビス距離を正しく推定する手法を提案する．提案手法を文字認識に適用した結果，パラメータの推定に用いる学習サンプルが少ない場合でもマハラノビス距離を正しく推定することができ，認識性能の改善が見られた．

キーワード　パターン認識，識別関数，固有ベクトル，推定誤差，マハラノビス距離

1. まえがき

ベイズの定理から導かれる統計的識別関数は，パターンの確率分布が正しく与えられたときに，誤認識によって生じる損失の期待値を最小に留める最適性を持つ．しかし，実際のパターン認識ではパターンの確率分布は未知で，パターンの確率分布として正規分布を仮定し，分布パラメータを学習用のサンプルから推定することが多い．このとき，用いる学習サンプルが特徴量の次元数に比べて不足すると，パラメータの推定誤差のために認識性能が低下することが知られている．

パターンの確率分布を多次元正規分布と仮定した識別関数として，2 次識別関数やマハラノビス距離が知られている．多次元正規分布の分布パラメータである平均ベクトルと共分散行列を学習サンプルから推定したとき，推定値である標本平均ベクトルと標本共分散行列は誤差を含む．竹下ら [1] は，2 次識別関数の値を実質的に決定するマハラノビス距離が平均ベクトルと共分散行列のそれぞれの推定誤差に対してどのような影響を受けるのかを理論的に調査した．その結果，学習サンプル数が 10 を超えるときは，平均ベクトルの推定誤差の影響は共分散行列の推定誤差に比べてほとんど無視できることを導いた．従って，2 次識別関数やマハラノビス距離では共分散行列の推定誤差が認識精度を低下させているといえる．

共分散行列の推定誤差については，共分散行列が固有値展開により，その固有値と固有ベクトルに分解できることから，固有値の推定誤差と固有ベクトルの推定誤差に分けて考えられてきた．このうち固有値の推定誤差については，特に小さな固有値の推定誤差が認識性能に悪影響を及ぼすことが知られ，信頼性の低い推定値を用いない手法が提案されている．このような手法には，木村らの修正 2 次識別関数 (MQDF) [2]，加藤らの改良型マハラノビス距離 (MMD) [3]，孫らの簡素化マハラノビス距離 (SMD) [4]，Friedman の Regularized Discriminant Analysis [5] などがある．

[†] 東北大学大学院工学研究科，仙台市
Graduate School of Engineering, Tohoku University, Sendai-shi, 980–8579 Japan

それに対して，信頼性の低い固有値を補正して用いる手法も提案されている．分布のパラメータを補正する手法は，信頼性の低いパラメータを用いない手法に比べて難しい反面，補正が正しく行われればより高い認識性能が得られる可能性がある．また，補正法を導く理論自体がパターン認識に与える貢献も大きい．James と Stein [6] は，従来の標本共分散行列は非許容的であり，より良い推定量が存在することを指摘し，標本共分散行列を改良する推定量 (James-Stein 推定量) を導いた．Takeshita ら [7] は実験によって James-Stein 推定量がパターン認識に有効であることを確認した．酒井ら [8] は標本共分散行列の固有値の平均的な偏りから標本固有値を補正する手法 (RQDF) を提案し，良好な認識結果を得ている．

一方，固有ベクトルの推定誤差が認識性能に及ぼす影響についてはこれまであまり考慮されてこなかった．酒井ら [9] は，標本固有ベクトルの揺らぎのために，パターン認識で推定すべき共分散行列は真の共分散行列ではなく，真の共分散行列をぼかした行列 (ぼけ共分散行列) であると主張し，固有値を二回補正する手法 (WRQDF) を提案している．しかし，何故ぼけ共分散行列を用いるべきかについての議論は行われていない．また，二回目の補正では標本固有値も積分変数としているため，ぼけ共分散行列が標本固有値が持つ情報を十分に反映していない．

本論文では共分散行列の固有ベクトルの推定誤差を考慮することで，学習サンプルが十分用意できない場合においても高精度な認識を行うことを目的とし，マハラノビス距離を正しく推定するために固有値を修正する手法を提案する．2. では，固有ベクトルの推定誤差によってマハラノビス距離に推定誤差が生じ，この誤差は固有値の誤差と解釈することができることを示す．そして，固有値を修正することにより，辞書の作成に用いるサンプルが少ないときでも，マハラノビス距離を正しく推定する手法を提案する．3. では，提案手法を文字認識に適用し，マハラノビス距離が正しく推定されること，認識性能が改善されることを示す．

なお，本論文では確率変数と定数を区別するため，$\acute{A}$ や $\acute{\boldsymbol{A}}$ のように，確率変数には ´ を付ける．

2. 真のマハラノビス距離の推定法

2.1 誤差を含む固有ベクトルを用いたマハラノビス距離の推定

真の共分散行列の第 i 固有値，固有ベクトルを λ_i，$\boldsymbol{\phi}_i$ とすると，真の固有値，固有ベクトルからなる行列は

$$\boldsymbol{\Lambda} = \mathrm{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_d) \tag{1}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 & \cdots & \boldsymbol{\phi}_d \end{bmatrix} \tag{2}$$

とおける．固有値は降順に並んでいるとする．真の平均ベクトルを $\boldsymbol{\mu}$，未知入力ベクトルを $\boldsymbol{x}$ とすれば，マハラノビス距離 $d(\boldsymbol{x})$ は次式で表される．

$$d(\boldsymbol{x}) = (\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Phi} \boldsymbol{\Lambda}^{-1} \boldsymbol{\Phi}^{\mathrm{T}} (\boldsymbol{x}-\boldsymbol{\mu}) \tag{3}$$

一方，標本共分散行列の第 i 固有ベクトルを $\hat{\boldsymbol{\phi}}_i$ とし，標本共分散行列の固有ベクトルからなる行列を

$$\hat{\boldsymbol{\Phi}} = \begin{bmatrix} \hat{\boldsymbol{\phi}}_1 & \hat{\boldsymbol{\phi}}_2 & \cdots & \hat{\boldsymbol{\phi}}_d \end{bmatrix} \tag{4}$$

とおく．固有ベクトルとして，推定誤差を含む標本固有ベクトルを用いたときのマハラノビス距離は次式で表される．

$$\hat{d}(\boldsymbol{x}) = (\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}} \hat{\boldsymbol{\Phi}} \boldsymbol{\Lambda}^{-1} \hat{\boldsymbol{\Phi}}^{\mathrm{T}} (\boldsymbol{x}-\boldsymbol{\mu}) \tag{5}$$

いま，正規直交基底である固有ベクトルが作る座標系を考える．このように考えると，真の固有ベクトルが作る座標系と標本固有ベクトルが作る座標系のずれ (回転) が固有ベクトルの推定誤差であると考えられる．そこで，第 j 標本固有ベクトルが真の第 i 固有ベクトル方向へずれた量を ij 成分に持つ行列 $\hat{\boldsymbol{\Psi}}$ を，固有ベクトルの推定誤差を表す行列として次のように定義する．

$$\hat{\boldsymbol{\Psi}} \equiv \boldsymbol{\Phi}^{\mathrm{T}} \hat{\boldsymbol{\Phi}} \tag{6}$$

$\boldsymbol{\Phi}$，$\hat{\boldsymbol{\Phi}}$ は正規直交基底 (回転行列) であるので，$\hat{\boldsymbol{\Psi}}$ も正規直交基底である．(6) 式から

$$\hat{\boldsymbol{\Phi}} = \boldsymbol{\Phi} \hat{\boldsymbol{\Psi}} \tag{7}$$

と表せるので，(7) 式を (5) 式に代入する．

$$\hat{d}(\boldsymbol{x}) = (\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Phi} \hat{\boldsymbol{\Psi}} \boldsymbol{\Lambda}^{-1} \hat{\boldsymbol{\Psi}}^{\mathrm{T}} \boldsymbol{\Phi}^{\mathrm{T}} (\boldsymbol{x}-\boldsymbol{\mu}) \tag{8}$$

$$= (\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Phi} \left(\hat{\boldsymbol{\Psi}} \boldsymbol{\Lambda} \hat{\boldsymbol{\Psi}}^{\mathrm{T}} \right)^{-1} \boldsymbol{\Phi}^{\mathrm{T}} (\boldsymbol{x}-\boldsymbol{\mu}) \tag{9}$$

(9) 式を (3) 式と比較すれば，(9) 式の $\left(\hat{\boldsymbol{\Psi}} \boldsymbol{\Lambda} \hat{\boldsymbol{\Psi}}^{\mathrm{T}} \right)^{-1}$ と (3) 式の固有値 $\boldsymbol{\Lambda}^{-1}$ は明らかに異なり，標本固有ベクトルを用いた $\hat{d}(\boldsymbol{x})$ は真のマハラノビス距離 $d(\boldsymbol{x})$

との誤差を含むことがわかる．そして，その誤差は $\left(\hat{\boldsymbol{\Psi}}\boldsymbol{\Lambda}\hat{\boldsymbol{\Psi}}^{\mathrm{T}}\right)^{-1}$ の非対角成分を無視できるなら，固有ベクトルの推定誤差によって生じた固有値の誤差とみなすことができる．そこで，(5) 式で $\boldsymbol{\Lambda}$ の代わりに適当な対角行列 $\tilde{\boldsymbol{\Lambda}}$ を用いることを考える．すると，(9) 式の固有値に相当する部分は $\left(\hat{\boldsymbol{\Psi}}\tilde{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Psi}}^{\mathrm{T}}\right)^{-1}$ となる．これを $\boldsymbol{\Lambda}^{-1}$ に近付けることにより，固有ベクトルが誤差を含む場合でもマハラノビス距離を正しく推定できることになる．すなわち，

$$\boldsymbol{\Lambda}^{-1} \approx \left(\hat{\boldsymbol{\Psi}}\tilde{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Psi}}^{\mathrm{T}}\right)^{-1} \tag{10}$$

となる $\tilde{\boldsymbol{\Lambda}}$ を選ぶことを考える．(10) 式から，

$$\tilde{\boldsymbol{\Lambda}} \approx \hat{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\Psi}} \tag{11}$$

が成立するので，

$$\tilde{\boldsymbol{\Lambda}} = \mathrm{D}\left(\hat{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\Psi}}\right) \tag{12}$$

とおく．ここで $\mathrm{D}(\cdot)$ は行列の対角成分を返す関数である．$\hat{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\Psi}}$ の非対角成分を無視することの妥当性は 3.1 の実験で確認する．(12) 式を (5) 式の $\boldsymbol{\Lambda}$ の代わりに用いることにより，次式が得られる．

$$\tilde{d}(\boldsymbol{x}) = (\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\hat{\boldsymbol{\Phi}}\tilde{\boldsymbol{\Lambda}}^{-1}\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{\mu}) \tag{13}$$

(13) 式は，誤差を含んだ固有ベクトルを用いて真のマハラノビス距離の推定値を与えると考えられる．

2.2 固有ベクトルの推定誤差を補償する固有値

(12) 式では，$\tilde{\boldsymbol{\Lambda}}$ が真の固有値 $\boldsymbol{\Lambda}$ と固有ベクトルの推定誤差 $\hat{\boldsymbol{\Psi}}$ で定義されている．固有ベクトルの推定誤差 $\hat{\boldsymbol{\Psi}}$ は (6) 式で真の固有ベクトル $\boldsymbol{\Phi}$ を用いて定義されている．いま，真の固有ベクトルが未知である場合を考えているので，$\hat{\boldsymbol{\Psi}}$ は観測できない．しかし，付録で示すように，固有ベクトルの推定誤差の確率分布は特徴量の次元数，学習サンプル数，真の固有値，標本固有値のみから決まり，真の固有ベクトルとは無関係である．従って，真の固有ベクトルが未知であっても，固有ベクトルの推定誤差の期待値であれば観測可能である．そこで，固有ベクトルの推定誤差を表す確率変数 $\acute{\boldsymbol{\Psi}}$ を考える．(12) 式で定めた $\tilde{\boldsymbol{\Lambda}}$ も確率変数となり，$\acute{\boldsymbol{\Lambda}}$ と記すと，(12) 式は次のようになる．

$$\acute{\boldsymbol{\Lambda}} = \mathrm{D}\left(\acute{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Lambda}\acute{\boldsymbol{\Psi}}\right) \tag{14}$$

(14) 式の両辺の標本固有値 $\hat{\boldsymbol{\Lambda}}$ に関する条件付き期待値をとると，次のようになる．

$$\begin{aligned}\tilde{\bar{\boldsymbol{\Lambda}}} = \mathrm{E}\left[\acute{\boldsymbol{\Lambda}}\middle|\hat{\boldsymbol{\Lambda}}\right] &= \mathrm{E}\left[\mathrm{D}\left(\acute{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Lambda}\acute{\boldsymbol{\Psi}}\right)\middle|\hat{\boldsymbol{\Lambda}}\right] \\ &= \mathrm{D}\left(\mathrm{E}\left[\acute{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Lambda}\acute{\boldsymbol{\Psi}}\middle|\hat{\boldsymbol{\Lambda}}\right]\right)\end{aligned} \tag{15}$$

$\tilde{\bar{\boldsymbol{\Lambda}}}$ は固有ベクトルの推定誤差の補償する固有値 $\tilde{\bar{\lambda}}_i$ からなる行列

$$\tilde{\bar{\boldsymbol{\Lambda}}} = \mathrm{diag}\left(\tilde{\bar{\lambda}}_1, \tilde{\bar{\lambda}}_2, \ldots, \tilde{\bar{\lambda}}_d\right) \tag{16}$$

である．(15) 式の第 i 対角要素を取り出すと，

$$\tilde{\bar{\lambda}}_i = \mathrm{E}\left[\sum_{j=1}^{d}\left\{\acute{\psi}_{ji}\right\}^2\lambda_j\middle|\hat{\boldsymbol{\Lambda}}\right] \tag{17}$$

$$= \sum_{j=1}^{d}\mathrm{E}\left[\left\{\acute{\psi}_{ji}\right\}^2\middle|\hat{\boldsymbol{\Lambda}}\right]\lambda_j \tag{18}$$

となる．さらに

$$\left\{\tilde{\psi}_{ji}\right\}^2 = \mathrm{E}\left[\left\{\acute{\psi}_{ji}\right\}^2\middle|\hat{\boldsymbol{\Lambda}}\right] \tag{19}$$

とおくと，$\tilde{\bar{\lambda}}_i$ は次式で得られる．

$$\tilde{\bar{\lambda}}_i = \sum_{j=1}^{d}\left\{\tilde{\psi}_{ji}\right\}^2\lambda_j \tag{20}$$

2.3 (20) 式の計算法

本節では (20) 式の計算法を与える．(19) 式の標本固有値 $\hat{\boldsymbol{\Lambda}}$ に関する条件付き期待値を一般化して，$\acute{\boldsymbol{\Psi}}$ から値が決まる関数 $f(\acute{\boldsymbol{\Psi}})$ に関する条件付き期待値を考える．この積分表現は次のようになる．

$$\mathrm{E}\left[f(\acute{\boldsymbol{\Psi}})\middle|\hat{\boldsymbol{\Lambda}}\right] = \int_{\acute{\boldsymbol{\Psi}}} f(\acute{\boldsymbol{\Psi}})\mathrm{P}(\acute{\boldsymbol{\Psi}}|\hat{\boldsymbol{\Lambda}})d\acute{\boldsymbol{\Psi}} \tag{21}$$

$\mathrm{P}(\acute{\boldsymbol{\Psi}}|\hat{\boldsymbol{\Lambda}})$ は固有ベクトルの推定誤差 $\acute{\boldsymbol{\Psi}}$ の確率密度関数である．高次元の場合，この期待値を厳密に求めるのは困難である．そこで，数値的に求めることを考える．モンテカルロ法を用いて計算する場合，正規分布に従うサンプルから標本固有値が $\hat{\boldsymbol{\Lambda}}$ であるときの固有ベクトルを作成して，期待値を計算する必要がある．しかし，これは実現不可能である．本論文では (21) 式を条件付きモンテカルロ法 [10] を用いて計算する．$f(\acute{\boldsymbol{\Psi}}) = \left\{\acute{\psi}_{ji}\right\}^2$ の場合を考えると，(19) 式の $\left\{\tilde{\psi}_{ji}\right\}^2$ が (21) 式で求まるので，(20) 式の $\tilde{\bar{\lambda}}_i$ が計算できる．l 条件付きモンテカルロ法を実現するために，(21) 式

の右辺を変形する．確率変数 $\acute{\mathbf{\Sigma}}$ と $\acute{\mathbf{\Lambda}}$ をそれぞれ

$$\acute{\mathbf{\Sigma}} = \acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}} \tag{22}$$

を満たす対称行列と対角行列とする．固有ベクトルの推定誤差の分布は真の固有ベクトルに依らないので，真の固有ベクトルが単位行列の場合 ($\mathbf{\Phi} = \boldsymbol{I}$) を考えても一般性を失わない．このとき，(6) 式から標本固有ベクトルと固有ベクトルの推定誤差が一致する ($\acute{\mathbf{\Phi}} = \acute{\mathbf{\Psi}}$) ので，$\acute{\mathbf{\Sigma}}$ の分布は Wishart 分布 (付録参照) で与えられる．(22) 式の関係から，次が成立する．

$$\mathrm{P}(\acute{\mathbf{\Sigma}}) = \mathrm{P}(\acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}}) = \mathrm{P}(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}})J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}}) \tag{23}$$

ただし，$J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}})$ は変換のヤコビアンで，

$$J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}}) = \frac{d\acute{\mathbf{\Psi}}d\acute{\mathbf{\Lambda}}}{d\acute{\mathbf{\Sigma}}} \tag{24}$$

である．$\hat{\mathbf{\Lambda}}$ は確率変数 $\acute{\mathbf{\Lambda}}$ の実現値であるから，

$$\mathrm{P}(\acute{\mathbf{\Psi}}\hat{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}}) = \mathrm{P}(\acute{\mathbf{\Psi}}, \hat{\mathbf{\Lambda}})J(\acute{\mathbf{\Psi}}, \hat{\mathbf{\Lambda}}) \tag{25}$$

が成り立つ．さらに，任意の関数 $g(\acute{\mathbf{\Lambda}})$ を定義し，

$$G = \int_{\acute{\mathbf{\Lambda}}} g(\acute{\mathbf{\Lambda}})d\acute{\mathbf{\Lambda}} \tag{26}$$

とおく．以上の準備のもとで，次式が成り立つ．

$$\begin{aligned}
&\int_{\acute{\mathbf{\Psi}}} f(\acute{\mathbf{\Psi}})\mathrm{P}(\acute{\mathbf{\Psi}}|\hat{\mathbf{\Lambda}})d\acute{\mathbf{\Psi}} \\
&= \int_{\acute{\mathbf{\Psi}}} f(\acute{\mathbf{\Psi}})\frac{\mathrm{P}(\acute{\mathbf{\Psi}}|\hat{\mathbf{\Lambda}})}{G}\left[\int_{\acute{\mathbf{\Lambda}}} g(\acute{\mathbf{\Lambda}})d\acute{\mathbf{\Lambda}}\right]d\acute{\mathbf{\Psi}} \\
&= \int_{\acute{\mathbf{\Psi}}\times\acute{\mathbf{\Lambda}}} f(\acute{\mathbf{\Psi}})\frac{\mathrm{P}(\acute{\mathbf{\Psi}}, \hat{\mathbf{\Lambda}})}{\mathrm{P}(\hat{\mathbf{\Lambda}})}\frac{g(\acute{\mathbf{\Lambda}})}{G}d\acute{\mathbf{\Psi}}d\acute{\mathbf{\Lambda}} \\
&= \frac{1}{\mathrm{P}(\hat{\mathbf{\Lambda}})}\int_{\acute{\mathbf{\Sigma}}} f(\acute{\mathbf{\Psi}})\frac{\mathrm{P}(\acute{\mathbf{\Psi}}\hat{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}})}{J(\acute{\mathbf{\Psi}}, \hat{\mathbf{\Lambda}})}\frac{g(\acute{\mathbf{\Lambda}})}{G}J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}})d\acute{\mathbf{\Sigma}} \\
&= \frac{1}{\mathrm{P}(\hat{\mathbf{\Lambda}})}\int_{\acute{\mathbf{\Sigma}}} f(\acute{\mathbf{\Psi}})w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}})\mathrm{P}(\acute{\mathbf{\Sigma}})d\acute{\mathbf{\Sigma}}
\end{aligned} \tag{27}$$

ただし，

$$w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}}) = \frac{\mathrm{P}(\acute{\mathbf{\Psi}}\hat{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}})}{J(\acute{\mathbf{\Psi}}, \hat{\mathbf{\Lambda}})}\frac{J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}})}{\mathrm{P}(\acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}})}\frac{g(\acute{\mathbf{\Lambda}})}{G} \tag{28}$$

である．(27) 式は，確率密度 $\mathrm{P}(\acute{\mathbf{\Psi}}|\hat{\mathbf{\Lambda}})$ を持つ固有ベクトルを作成して $f(\acute{\mathbf{\Psi}})$ の期待値を取ることが，確率密度 $\mathrm{P}(\acute{\mathbf{\Sigma}})$ を持つ共分散行列を作成して $f(\acute{\mathbf{\Psi}})w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}})\frac{1}{\mathrm{P}(\hat{\mathbf{\Lambda}})}$ の期待値を取ることに等しいことを意味する．

$w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}})$ は付録に示す (A·9)，(A·11) 式を (28) 式に代入して変形することにより，

$$\begin{aligned}
w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}}) = &\frac{|\hat{\mathbf{\Lambda}}|^{\frac{1}{2}(n-p-2)}\prod_{i<j}^{d}(\hat{\lambda}_i - \hat{\lambda}_j)}{|\acute{\mathbf{\Lambda}}|^{\frac{1}{2}(n-p-2)}\prod_{i<j}^{d}(\acute{\lambda}_i - \acute{\lambda}_j)} \\
&\cdot \frac{\exp\left(-\frac{n-1}{2}\operatorname{tr}\mathbf{\Lambda}^{-1}\acute{\mathbf{\Psi}}\hat{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}}\right)}{\exp\left(-\frac{n-1}{2}\operatorname{tr}\mathbf{\Lambda}^{-1}\acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}}\right)}\frac{g(\acute{\mathbf{\Lambda}})}{G}
\end{aligned} \tag{29}$$

となる．$\mathrm{P}(\hat{\mathbf{\Lambda}})$ は解析的に得られるが，特に高次元の場合，数値計算を行うのは困難である．しかし，(27) 式で $f(\acute{\mathbf{\Psi}}) = 1$ の場合を考えると，確率の公理より，

$$\frac{1}{\mathrm{P}(\hat{\mathbf{\Lambda}})}\int_{\acute{\mathbf{\Sigma}}} w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}})\mathrm{P}(\acute{\mathbf{\Sigma}})d\acute{\mathbf{\Sigma}} = 1 \tag{30}$$

が成り立つので，$\mathrm{P}(\hat{\mathbf{\Lambda}})$ は

$$\mathrm{P}(\hat{\mathbf{\Lambda}}) = \int_{\acute{\mathbf{\Sigma}}} w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}})\mathrm{P}(\acute{\mathbf{\Sigma}})d\acute{\mathbf{\Sigma}} \tag{31}$$

で与えられる．

ここで，$f(\acute{\mathbf{\Psi}}) = \left\{\acute{\psi}_{ji}\right\}^2$ の場合を考えると，(27) 式は (19) 式の右辺に等しいので，(27) 式の $\mathrm{P}(\hat{\mathbf{\Lambda}})$ に (31) 式を代入すると，(19) 式は次式で表される．

$$\left\{\tilde{\psi}_{ji}\right\}^2 = \frac{\int_{\acute{\mathbf{\Sigma}}}\left\{\acute{\psi}_{ji}\right\}^2 w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}})\mathrm{P}(\acute{\mathbf{\Sigma}})d\acute{\mathbf{\Sigma}}}{\int_{\acute{\mathbf{\Sigma}}} w_0(\acute{\mathbf{\Sigma}}; \hat{\mathbf{\Lambda}})\mathrm{P}(\acute{\mathbf{\Sigma}})d\acute{\mathbf{\Sigma}}} \tag{32}$$

(32) 式の分子，分母の積分を Wishart 分布に従って生成した $\acute{\mathbf{\Sigma}}$ の t 個のサンプルによる算術平均として表すと，次式が得られる．

$$\left\{\tilde{\Psi}_{ji}\right\}^2 = \frac{\sum_{k=1}^{t}\left\{\acute{\Psi}_{k,ji}\right\}^2 w_0(\acute{\mathbf{\Psi}}_k\acute{\mathbf{\Lambda}}_k\acute{\mathbf{\Psi}}_k^{\mathrm{T}}; \hat{\mathbf{\Lambda}})}{\sum_{k=1}^{t} w_0(\acute{\mathbf{\Psi}}_k\acute{\mathbf{\Lambda}}_k\acute{\mathbf{\Psi}}_k^{\mathrm{T}}; \hat{\mathbf{\Lambda}})} \tag{33}$$

以上の議論から，条件付きモンテカルロ法を用いて (20) 式の $\tilde{\tilde{\lambda}}_i$ を推定する手順を Algorithm 1 に示す．n は実際のパターン認識において学習に使用できるサンプル数，$\check{\mathbf{\Lambda}}$ は真の固有値，もしくは既存の固有値補正法で補正された標本固有値である．本論文では，$g(\acute{\mathbf{\Lambda}}_k) = \frac{1}{t}$ とした．

Algorithm 1 $\tilde{\tilde{\lambda}}_i$ の推定

1: 乱数を用いて正規分布 $\mathrm{N}(\mathbf{0}, \check{\mathbf{\Lambda}})$ に従う nt 個の人工サンプル $X_1, \ldots, X_{nt}$ を作成する．
2: **for** $k = 1$ to t **do**
3: n 個の人工サンプル $X_{n(k-1)+1}, \ldots, X_{nk}$ から標本共分散行列 $\acute{\mathbf{\Sigma}}_k$ を求め，標本固有値 $\acute{\mathbf{\Lambda}}_k$ と固有ベクトルの推定誤差 $\acute{\mathbf{\Psi}}_k$ を得る．
4: **end for**
5: (33) 式を用いて $\left\{\tilde{\Psi}_{ji}\right\}^2$ を求める．
6: (20) 式で $\tilde{\tilde{\lambda}}_i$ を得る．

2.4 標本固有ベクトルを用いて真のマハラノビス距離を推定する手順

実際のパターン認識において学習サンプルが n 個あるとき，固有値を以下の手順で修正することにより，標本固有ベクトルを用いて真のマハラノビス距離を推定することができる．

（1）辞書の作成に使用可能な n 個のサンプルから標本平均ベクトルと標本共分散行列を求め，標本固有値 $\hat{\mathbf{\Lambda}}$ と標本固有ベクトル $\hat{\mathbf{\Phi}}$ を得る．

（2）標本固有値を固有値の推定誤差を補正する固有値補正法 (例えば酒井らの手法 [8]，James-Stein 推定量 [6] など) を適用して，標本固有値を補正した $\check{\mathbf{\Lambda}}$ を得る．

（3）Algorithm 1 を用いて $\tilde{\tilde{\lambda}}_i$ を推定する．

（4）$\tilde{\tilde{\lambda}}_i$ を固有値として，標本固有ベクトル $\hat{\mathbf{\Phi}}$ と共に認識に用いる．

3. 提案手法の性能評価

3.1 マハラノビス距離の推定値

標本固有ベクトルと提案手法で修正した固有値を用いたときに，マハラノビス距離が正しく推定できることを実験により確認する．

まず，3 種類のマハラノビス距離 t_{ij}，e_{ij}，p_{ij} を定義する．t_{ij} をクラス i の平均ベクトルとクラス i に属する j 番目の未知入力サンプルとの真のマハラノビス距離 $d(\boldsymbol{x})$((3) 式) とする．e_{ij} を真の平均ベクトル，真の固有値，標本固有ベクトルを用いて求めたマハラノビス距離 $\hat{d}(\boldsymbol{x})$ ((5) 式) とする．p_{ij} を真の平均ベクトル，提案手法で求めた固有値，標本固有ベクトルを用いて求めたマハラノビス距離 $\tilde{d}(\boldsymbol{x})$ ((13) 式) とする．t_{ij} に対する e_{ij} の比の平均値 r_{e} と，t_{ij} に対する p_{ij} の比の平均値 r_{p} を次のように定義する．

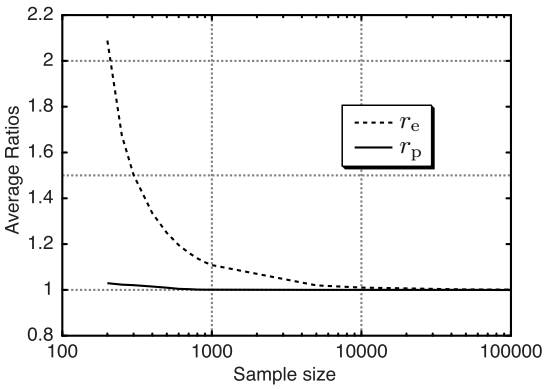


図 1 マハラノビス距離の比の平均値
Fig. 1 The mean ratios of Mahalanobis distances.

$$r_{\mathrm{e}} = \frac{1}{cs}\sum_{i}^{c}\sum_{j}^{s}\frac{e_{ij}}{t_{ij}} \tag{34}$$

$$r_{\mathrm{p}} = \frac{1}{cs}\sum_{i}^{c}\sum_{j}^{s}\frac{p_{ij}}{t_{ij}} \tag{35}$$

ここで c はクラス数，s はクラス毎のテストサンプル数である．r_{e}，r_{p} が 1 に近い程，マハラノビス距離が正しく推定されているといえる．本実験では r_{e} と r_{p} を比較する．

実験では，実際の文字画像データから推定した分布パラメータと同じパラメータを持つ正規分布に従う人工特徴量を用いる．文字画像は NIST Special Database 19 [11] の数字サンプルで，64×64 の大きさに非線形正規化 [12] した後，196 次元の方向線素特徴量 [13] を抽出する．クラス i の平均ベクトル $\boldsymbol{\mu}_i$ と共分散行列 $\mathbf{\Sigma}_i$ をクラス毎に文字画像の特徴量 36,000 文字分から推定する．人工サンプルはこの推定値を真の値として，正規分布 $\mathrm{N}(\boldsymbol{\mu}_i, \mathbf{\Sigma}_i)$ に従うように乱数を用いて作成する．

辞書の作成に用いる学習サンプル数を 200 個から 100,000 個まで変化させ，テストサンプルは 1,000 個ずつ用いる．Algorithm 1 のパラメータ t を 10,000 とした．

t_{ij} に対する e_{ij} の比の平均値 r_{e} と，t_{ij} に対する p_{ij} の比の平均値 r_{p} を図 1 に示す．r_{e} では辞書の推定に用いる学習サンプル数が少ない程，マハラノビス距離の推定値が大きく偏っているが，r_{p} では学習サンプル数に依らずほぼ正しく推定できている．これにより，学習サンプル数が少なく，固有ベクトルが推定誤差を含む場合でも，提案手法を用いることでマハラノ

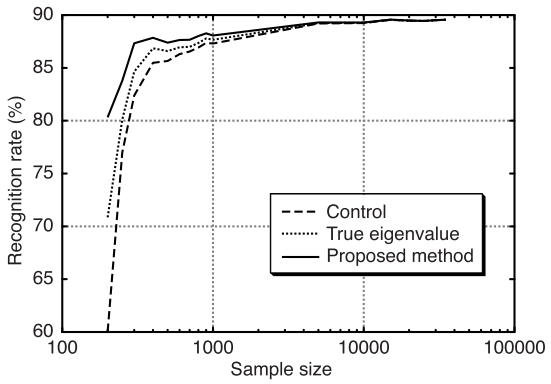


図 2　認識実験の結果 (文字画像)
Fig. 2　The recognition rates of the recognition experiments on feature vectors of actual character samples.

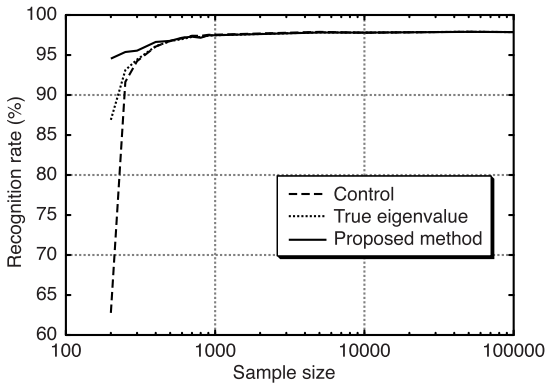


図 3　認識実験の結果 (人工サンプル)
Fig. 3　The recognition rates of the recognition experiments on artificial feature vectors.

ビス距離が正しく推定できることを確認できた．この結果は，2. で $\hat{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\Psi}}$ の非対角成分を無視したことの妥当性も示している．

3.2　認識性能の確認

提案手法を統計的識別手法に適用したときの認識性能を確認するために，識別関数としてマハラノビス距離を用いた文字認識実験を行う．実験は固有値だけが異なる 3 種類の辞書を作成し，それぞれの辞書を用いたときの認識率を比較する．3 種類の辞書は，固有値として標本固有値を用いる “Control”，真の固有値を用いる “True eigenvalue”($\hat{d}(\boldsymbol{x})$)，提案手法で修正した固有値を用いる “Proposed method”($\tilde{d}(\boldsymbol{x})$) である．全ての辞書で共通の標本平均ベクトル，標本固有ベクトルを用いる．

サンプルは 3. 1 で作成した文字画像の特徴量と正規分布に従う人工特徴量を用いる．辞書の作成に用いる学習サンプル数を文字画像の特徴量では 200 個から 35,000 個まで，人工サンプルでは 200 個から 100,000 個まで変化させる．テストサンプルは 1,000 個用いる．実際の文字画像では真の固有値が得られないため，酒井らの手法 [8] で補正した固有値を真の固有値の代わりに用いる．Algorithm 1 のパラメータ t を 10,000 とした．

実際の文字画像を用いた認識実験の結果を図 2 に，人工サンプルを用いたときの結果を図 3 に示す．実際の文字画像，人工サンプル共に，“Proposed method” の認識率は全ての学習サンプル数で “True eigenvalue” より高く．提案手法を適用した識別関数は認識に有効であることが確認できた．“True eigenvalue” と “Proposed method” の認識率の差を学習サンプル数との関係を考慮して見てみると，学習サンプル数が多いときには認識率にほとんど差はないが，少なくなるほど認識率の差は大きくなる．これは学習サンプル数が多いときには，補正されるべき推定誤差がもともと小さく，少ないときには認識性能に影響する固有ベクトルの推定誤差が大きいためであると考えられる．また，真の固有値を用いた “True eigenvalue” (文字画像の特徴量では補正した標本固有値) は，標本固有値をそのまま用いた “Control” よりも高い認識率が得られた．

4. むすび

本論文ではまず，共分散行列の固有ベクトルの推定誤差によってマハラノビス距離に推定誤差が生じ，この誤差は固有値の誤差と解釈できることを示した．そして，辞書の作成に用いるサンプルが少ないときでも，固有ベクトルの推定誤差を補償するように固有値を修正することで，マハラノビス距離を正しく推定する手法を提案した．提案手法を文字認識に適用し，マハラノビス距離が正しく推定されること，認識性能が改善されることを確認した．マハラノビス距離の推定値を改善する効果はパラメータの推定に用いる学習サンプルが少ないほど高く，本手法は学習サンプルを十分用意できない場合において特に有効である．

文　献

[1] 竹下鉄夫，鳥脇純一郎，“マハラノビス距離の推定値の確率分布に関する一性質，” 信学論 (D-II)，vol.J76-D-II, no.1，pp.59–64，Jan.，1993．

[2] F. Kimura, K. Takashina, S. Tsuruoka and

Y. Miyake, "Modified quadratic discriminant functions and the application to chinese character recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.9, no.1, pp.149–153, Jan., 1987.
[3] 加藤寧，安倍正人，根元義章，"改良型マハラノビス距離を用いた高精度な手書き文字認識，" 信学論 (D-II)，vol.J79-D-II, no.1，pp.45–52，Jan.，1996．
[4] F. Sun, S. Omachi and H. Aso, "Precise selection of candidates for handwritten character recognition using feature regions," IEICE Trans. Inf. & Syst., vol.E79-D, no.5, pp.510–515, May, 1996.
[5] J. H. Friedman, "Regularized discriminant analysis," Journal of American Statistical Association, vol.84, no.405, pp.165–175, 1989.
[6] W. James and C. Stein, Estimation with quadratic loss, Proc. 4th Berkeley Symp. on Math. Statist. and Prob., pp.361–379, 1961.
[7] T. Takeshita and J. Toriwaki, "Experimental study of performance of pattern classifiers and the size of design samples," Pattern Recognition Letters, vol.16, no.3, pp.307–312, Mar., 1995.
[8] 酒井充，米田政明，長谷博行，丸山博，直江美知子，"固有値の偏り補正に基づく 2 次識別関数，" 信学論 (D-II)，vol.J82-D-II, no.4，pp.631–640，Apr.，1999．
[9] 酒井充，米田政明，長谷博行，丸山博，"固有値を二重に補正した共分散行列を用いた 2 次識別関数，" 信学技報 PRMU99–30，pp.1–8，June，1999．
[10] 津田孝夫，モンテカルロ法とシミュレーション—電子計算機の確率論的応用—，培風館，東京，三訂版，1995．
[11] P. J. Grother, Nist special database 19 handprinted forms and characters database, Technical report, National Institute of Standards and Technology, Mar., 1995.
[12] 山田博三，斉藤泰一，山本和彦，"線密度イコライゼーション—相関法のための非線形正規化法，" 信学論 (D)，vol.J67-D, no.11，pp.1379–1383，Nov.，1984．
[13] 孫寧，田原透，阿曽弘具，木村正行，"方向線素特徴量を用いた高精度文字認識，" 信学論 (D-II)，vol.J74-D-II, no.3，pp.330–339，Mar.，1991．
[14] R. J. Muirhead, Aspects of Multivariate Statistical Theory, John Wiley & Sons, Inc., New York, 1982.
[15] 竹村彰通，多変量推測統計の基礎，共立出版株式会社，東京，1991．

付　　録

固有ベクトルの推定誤差の確率密度関数

(i) Wishart 分布

正規分布 $\mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$ に従う各々独立な n 個の d 次元確率ベクトル $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ の平方和積和行列 $\boldsymbol{W}$ を

$$\boldsymbol{W} = \sum_{t=1}^{n} \boldsymbol{X}_t \boldsymbol{X}_t^{\mathrm{T}} \tag{A·1}$$

とおくと，$\boldsymbol{\Sigma}$ が与えられたときの $\boldsymbol{W}$ の分布 $\mathrm{W}_d(n, \boldsymbol{\Sigma})$ は Wishart 分布として知られている [14], [15]．その確率密度関数は次式で与えられる．

$$\mathrm{P}(\boldsymbol{W}|\boldsymbol{\Sigma}) = \frac{v(d,n)|\boldsymbol{W}|^{\frac{1}{2}(n-d-1)}}{|\boldsymbol{\Sigma}|^{\frac{1}{2}n}} \exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\right) \tag{A·2}$$

ここで $v(d,n)$ は

$$v(d,n) = \frac{1}{2^{\frac{1}{2}nd}\pi^{\frac{1}{4}d(d-1)}\prod_{i=1}^{d}\boldsymbol{\Gamma}\left[\frac{1}{2}(n+1-i)\right]} \tag{A·3}$$

である．

(ii)　標本共分散行列の確率密度関数

標本共分散行列 $\acute{\boldsymbol{\Sigma}}$，標本平均ベクトル $\acute{\boldsymbol{\mu}}$ を

$$\acute{\boldsymbol{\Sigma}} = \frac{1}{n-1}\sum_{t=1}^{n}(\boldsymbol{X}_t - \acute{\boldsymbol{\mu}})(\boldsymbol{X}_t - \acute{\boldsymbol{\mu}})^{\mathrm{T}} \tag{A·4}$$

$$\acute{\boldsymbol{\mu}} = \frac{1}{n}\sum_{t=1}^{n}\boldsymbol{X}_t \tag{A·5}$$

とおくと，$\acute{\boldsymbol{\Sigma}}$ の分布は $\mathrm{W}_d(n-1, \frac{1}{n-1}\boldsymbol{\Sigma})$ で与えられ，$\acute{\boldsymbol{\Sigma}}$ の確率密度関数は

$$\mathrm{P}(\acute{\boldsymbol{\Sigma}}|\boldsymbol{\Sigma}) = (n-1)^{\frac{1}{2}(n-1)d}v(d,n-1) \cdot \frac{|\acute{\boldsymbol{\Sigma}}|^{\frac{1}{2}(n-d-2)}\exp\left(-\frac{n-1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\acute{\boldsymbol{\Sigma}}\right)}{|\boldsymbol{\Sigma}|^{\frac{1}{2}(n-1)}} \tag{A·6}$$

となる [14], [15]．

(iii)　固有値，固有ベクトルの推定誤差の同時確率密度関数

共分散行列 $\boldsymbol{\Sigma}$，標本共分散行列 $\acute{\boldsymbol{\Sigma}}$ は固有値展開により，

$$\boldsymbol{\Sigma} = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{\mathrm{T}} \tag{A·7}$$

$$\acute{\boldsymbol{\Sigma}} = \acute{\boldsymbol{\Phi}}\acute{\boldsymbol{\Lambda}}\acute{\boldsymbol{\Phi}}^{\mathrm{T}} = \boldsymbol{\Phi}\acute{\boldsymbol{\Psi}}\acute{\boldsymbol{\Lambda}}\acute{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}} \tag{A·8}$$

となる．(A·6) 式は (A·7)，(A·8) を用いて次式のように書き直すことができる．

$$\mathrm{P}(\boldsymbol{\Phi}\acute{\boldsymbol{\Psi}}\acute{\boldsymbol{\Lambda}}\acute{\boldsymbol{\Psi}}^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}}|\boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{\mathrm{T}}) = (n-1)^{\frac{1}{2}(n-1)d}v(d,n-1) \cdot \frac{|\acute{\boldsymbol{\Lambda}}|^{\frac{1}{2}(n-d-2)}\exp\left(-\frac{n-1}{2}\operatorname{tr}\boldsymbol{\Lambda}^{-1}\acute{\boldsymbol{\Psi}}\acute{\boldsymbol{\Lambda}}\acute{\boldsymbol{\Psi}}^{\mathrm{T}}\right)}{|\boldsymbol{\Lambda}|^{\frac{1}{2}(n-1)}} \tag{A·9}$$

(A·9) 式は $\mathbf{\Phi}$ に依存しないので，$\mathrm{P}(\acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}}|\mathbf{\Lambda})$ となり，$\mathrm{P}(\acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}})$ と書く．この $\mathrm{P}(\acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}})$ は

$$\mathrm{P}(\acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}}) = \mathrm{P}(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}}) J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}}) \qquad \text{(A·10)}$$

と変形できる．ヤコビアン $J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}})$ は (A·9) 式より，

$$J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}}) = \frac{\prod_{i=1}^{d} \mathbf{\Gamma}\left[\frac{1}{2}(d+1-i)\right]}{2^d \pi^{\frac{1}{4}d(d+1)}} \frac{1}{\prod_{i<j}^{d} (\acute{\lambda}_i - \acute{\lambda}_j)}. \qquad \text{(A·11)}$$

である [14], [15]．

(iv) 固有ベクトルの推定誤差の確率密度関数

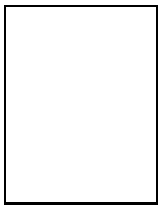
固有ベクトルの推定誤差の確率密度関数 $\mathrm{P}(\acute{\mathbf{\Psi}}|\acute{\mathbf{\Lambda}})$ は (A·10) 式を用いて，

$$\mathrm{P}(\acute{\mathbf{\Psi}}|\acute{\mathbf{\Lambda}}) = \frac{\mathrm{P}(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}})}{\mathrm{P}(\acute{\mathbf{\Lambda}})} \qquad \text{(A·12)}$$

$$= \frac{\mathrm{P}(\acute{\mathbf{\Psi}}\acute{\mathbf{\Lambda}}\acute{\mathbf{\Psi}}^{\mathrm{T}})}{\mathrm{P}(\acute{\mathbf{\Lambda}}) J(\acute{\mathbf{\Psi}}, \acute{\mathbf{\Lambda}})} \qquad \text{(A·13)}$$

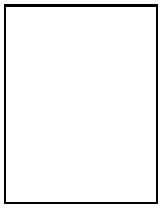
となる．(A·10)，(A·11) 式から，$\mathrm{P}(\acute{\mathbf{\Psi}}|\acute{\mathbf{\Lambda}})$ は標本固有値 $\acute{\mathbf{\Lambda}}$ が与えられたとき，次元数 d，サンプル数 n，真の固有値 $\mathbf{\Lambda}$ のみから決まり，真の固有ベクトル $\mathbf{\Phi}$ に依存していないことがわかる．

（平成 x 年 xx 月 xx 日受付）



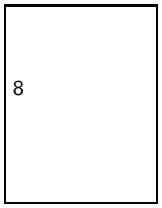
岩村　雅一　（学生員）

平 10 東北大・工・通信卒．平 12 同大大学院博士課程前期了．現在，博士課程後期在学中．パターン認識に関する研究に従事．



大町真一郎　（正員）

昭 63 東北大・工・情報卒．平 5 同大大学院博士課程了．同年同大情報処理教育センター助手．平 8 同大工学部助手．平 11 同大大学院工学研究科助教授．現在に至る．博士（工学）．その間，平 12 ~ 13 米国ブラウン大学客員助教授．パターン認識，コンピュータビジョン，並列処理，文字認識システムの開発などの研究に従事．IEEE，情報処理学会，人工知能学会，Pattern Recognition Society 各会員．



阿曽　弘具　（正員）

昭 43 東北大・工・電気卒．昭 49 同大大学院博士課程了．昭 48 東北大・工・助手，昭 54 名大・工・講師．昭 57 同助教授，昭 61 東北大・工・助教授を経て，平 3 同教授．現在，同大学院工学研究科教授．工博．その間，学習オートマトン，セル構造オートマトン，並行処理理論，シストリックアルゴリズム設計論，文字認識，音声認識，ニューラルネットワークなどの研究に従事．平成 3 年度本学会業績賞受賞．IEEE, ACM, EATCS，情報処理学会，人工知能学会，日本認知科学会，LA 各会員．