Vol.2017-AVM-98 No.13
2017/10/13

# 正定値対称行列における次元削減と辞書学習の統合学習法とその応用

笠井 裕之[†1,a)]

**概要**：近年，辞書学習とスパース表現係数を用いた処理手法は，信号処理，画像処理，コンピュータビジョン，パターン認識，機械学習，統計学習等の様々な分野で大きな注目を集めている．しかしながら，対象信号の高次元化に伴い，計算処理量が増大するだけでなく，潜在的構造に寄与する信号が高次元空間の中に埋没し，辞書学習やスパースコーディング処理における性能劣化を引き起こすという問題が指摘されている．本問題に対して次元削減手法は一つの解決手法であるが，高次元データに対して次元削減を行いその後辞書学習や識別処理を行う独立した処理が一般的であるため，後続の処理に対して最適な次元削減と辞書学習が行われないという問題がある．一方，データ処理で対象となる信号がユークリッド空間に属さず多様体上に分布する場合がある．正定値対称行列により表現されるデータはその一つであるが，近年，機械学習やコンピュータビジョンの様々なタスクにおいて，当該データを用いた処理手法が高い性能を示す結果が多数報告されている．そこで本稿では，正定値対称行列から構成されるリーマン多様体上に位置する信号を対象として，クラス識別タスクにおける性能向上を目指し，次元削減と辞書学習の統合学習法について提案する．

# Joint dimensionality reduction and dictionary learning on symmetric positive definite matrices

Hiroyuki Kasai[†1,a)]

***Abstract:*** Sparse representation based classification with dictionary learning from the training data has lead to promising results in many computer vision task. Dimensionality reduction also makes the computational cost lower and highlights the buried discriminative feature of high-dimensional data. However, data of interest are sometimes not characterized by simple vector features, and often originate from a manifold. One representative of them is symmetric positive definite (SPD) matrices that have been widely used in many computer vision tasks. In this paper, we propose a Riemannian joint dimensionality reduction and dictionary learning on SPD matrices for classification tasks. The joint learning considers the interaction between dimensionality reduction and dictionary learning procedures by connecting them into a unified framework.

## 1. はじめに

辞書学習とスパース表現係数を用いた信号処理手法は，機械学習やコンピュータビジョン分野等の様々なタスクにおいて非常に高い性能を示すことが報告されている [1], [2]．しかしながら，対象信号の高次元化に伴い，計算処理量が増大するだけでなく，データの潜在的構造に寄与する重要信号が高次元空間の中に埋没し，辞書学習やスパースコーディング処理における性能劣化を引き起こすという問題が指摘されている．本問題に対して次元削減手法は一つの解決手法であるが，高次元データに対して次元削減を行いその後辞書学習や識別処理を行う独立した処理が一般的であるため，後続のクラス識別などの処理に対して最適な次元削減と辞書学習が行われないという問題がある [3]．そこで Feng らは，次元削減のための射影行列とクラス識別性能を向上する辞書学習を統合的に行う手法を提案している [4]．

†1 現在，電気通信大学 大学院情報理工学研究科
Presently with Graduate School of Informatics and Engineering, The University of Electro-Communications
a) kasai@is.uec.ac.jp

ここでは，フィッシャー線形判別のような制約条件を係数行列に課すことでクラス識別性能の向上を実現している．他にも同様な研究として，Yang らは射影行列とクラス特有辞書を同時に学習する手法を提案し [5]，Li らは非負射影行列と辞書学習の統合学習手法を [6]，Foroughi らは係数行列だけでなく射影行列にもクラス識別性能向上を実現する制約を与える手法を提案している [7]．

対象信号がユークリッド空間に属さず多様体上に分布する問題が様々な基礎・応用問題で頻繁に現れる．“正定値対称行列” により表現されるデータはその一つであるが，近年，機械学習やコンピュータビジョンの様々なタスクにおいて，当該表現を用いた処理アルゴリズムが高い性能を示す報告が多数なされている．例えば， 領域分散共分散行列 (RCM: region covariance matrices) による識別器は，テクスチャ認識や顔認識などの応用で高い認識率を示している [8], [9]．一方で，例えば二つの正定値対称行列間の距離を評価する際，フロベニウスノルムの使用により劣悪な性能劣化を引き起こす事象がいくつか報告されている．これは，正定値対称行列がベクトル空間でなく多様体空間を成すため，その幾何的構造を考慮しない距離を使用することに起因する．

本稿では，正定値対称行列から構成されるリーマン多様体上で表現される信号を対象として，クラス識別処理における性能向上を目指し，次元削減と辞書学習の統合学習法について検討する．以後，$a \in \mathbb{R}$ はスカラー，$\boldsymbol{a} \in \mathbb{R}^d$ は $d$ 次元ベクトル，$\mathbf{A} \in \mathbb{R}^{m\times n}$ は $m \times n$ 次元行列，$\boldsymbol{\mathcal{A}} \in \mathbb{R}^{i\times j\times k}$ は 3 次多次元行列（テンソル）を表すものとする．

## 2. 正定値対称行列多様体とその幾何

$d \times d$ 正定値対称行列が作る空間はベクトル空間ではない．これは正定値対称行列に負のスカラーを掛けることで容易に確認できる．正定値対称行列の集合は “正定値実対称行列錐” と呼ばれ “多様体” として定義される．本稿では，$d \times d$ 正定値対称行列集合全体がつくる多様体を，特別に “正定値対称行列多様体” と呼び $\mathcal{S}^d_{++}$ と表記することとする．さて，多様体 $\mathcal{M}$ 上の各点 $p \in \mathcal{M}$ における接空間 $T_p\mathcal{M}$ に内積 $\langle \cdot, \cdot \rangle_p : T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$ を与えることで “距離” の概念を得ることができ，この内積を “リーマン計量” と呼ぶ．この時，正定値対称行列多様体を扱うアルゴリズムは，導入されたリーマン計量により定まる幾何的構造を考慮する必要がある．

正定値対称行列多様体 $\mathcal{S}^d_{++}$ の最も一般的なリーマン計量として Affine-invariant Riemannian Metric (AIRM) が提案されている [10], [11]．具体的には，多様体上の点 $\mathbf{P} \in \mathcal{S}^d_{++}$ において $\langle \mathbf{V}, \mathbf{W} \rangle_{\mathbf{P}} := \langle \mathbf{P}^{-1/2}\mathbf{V}\mathbf{P}^{-1/2}, \mathbf{P}^{-1/2}\mathbf{W}\mathbf{P}^{-1/2} \rangle_{\mathbf{P}} = \mathrm{Tr}(\mathbf{P}^{-1}\mathbf{V}\mathbf{P}^{-1}\mathbf{W})$，として定義される．ここで $\mathbf{V}, \mathbf{W} \in T_{\mathbf{P}}\mathcal{S}^d_{++}$ である．本計量の特徴は，$d$ 次一般線型群 $\mathrm{GL}(d)$ による作用により値が変化しない点にある．事実，$[\mathbf{X} \to \mathbf{M}\mathbf{X}\mathbf{M}^T, \mathbf{X} \in \mathcal{S}^d_{++}, ^\forall \mathbf{M} \in \mathrm{GL}(d)]$ である．さて，リーマン計量を与えることにより多様体上の異なる二点間の測地線距離が与えられる．特に AIRM を備える正定値対称行列多様体では，いかなる異なる二点間の測地線においてもそれが唯一に定まり最短のパスを与えることが知られている [11]．よって，AIRM により定まる測地線距離 $d_A : \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \to [0, \infty]$ は $d^2_A(\mathbf{A}, \mathbf{B}) := \mathrm{Log}||\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}||^2_F = \sum_{i=1}^d \ln \lambda_i(\mathbf{A}, \mathbf{B})$ により与えられる．ここで $\mathbf{A}, \mathbf{B} \in \mathcal{S}^n_{++}$, $||\cdot||_F$ はフロベニウスノルム， $\mathrm{Log}(\cdot)$ は行列対数， $\lambda_i(\mathbf{A}, \mathbf{B})$ は一般化固有値問題の $i$ 番目固有値である．尚，接空間 $T_{\mathbf{X}}\mathcal{S}^d_{++}$ は対称行列の集合から構成されることに注意されたい．

## 3. 従来手法

本節では，正定値対称行列多様体 $\mathcal{S}^d_{++}$ に特化し，$\mathcal{S}^d_{++}$ 上での辞書学習法 [12] と次元削減法 [13] について概説する．

$N$ 個の $d \times d$ 正定値対称行列により構成される訓練データを $\boldsymbol{\mathcal{X}}_{trn} = \{\mathbf{X}_1, \ldots, \mathbf{X}_N\} \in \mathbb{R}^{d\times d\times N}$ と定義する．$\mathbf{X}_j \in \mathcal{S}^d_{++}$ は構成する正定値対称行列を示す．ここで，$\boldsymbol{\mathcal{X}}_{trn}$ の各スライスが正定値対称行列から成るテンソル構造を有することに注意されたい．一方，同様に $\boldsymbol{\mathcal{D}} = \{\mathbf{D}_1, \ldots, \mathbf{D}_n\} \in \mathbb{R}^{d\times d\times n}$ は辞書テンソルを示し，$\mathbf{D}_k \in \mathcal{S}^d_{++}$ はその基底行列を表す．さらに，スパース係数ベクトルを $\boldsymbol{a}_j \in \mathbb{R}^n (j = 1, \ldots, N)$ と定義し，係数行列 $\mathbf{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N] \in \mathbb{R}^{n\times N}$ を形成するものと定義する．さて，辞書学習は $\boldsymbol{\mathcal{X}}_{trn}$ をより良く近似表現することが可能な辞書テンソル $\boldsymbol{\mathcal{D}}$ と係数ベクトル $\boldsymbol{a}_j$ を求めることであるが，ここで最も注意すべきは，要素 $\mathbf{X}_j$ を辞書テンソル $\boldsymbol{\mathcal{D}}$ と係数 $a_{j,i} (i = 1, \ldots, n)$ の錐結合により表現する点，つまり $a_{j,i}$ が非負値である点である．よって，この錐結合を $\mathbf{X}_j \approx \boldsymbol{\mathcal{D}}\boldsymbol{a}_j := \sum_{i=1}^n a_{j,i}\mathbf{D}_i (\boldsymbol{a}_j \geq 0)$ として表現する．以上から，正定値対称行列多様体における辞書学習の問題は $\min_{\boldsymbol{\mathcal{D}}, \mathbf{A} \in \mathbb{R}^{n\times N}_+} \frac{1}{2} \sum_{j=1}^N d^2(\mathbf{X}_j, \boldsymbol{\mathcal{D}}\boldsymbol{a}_j) + R_a(\boldsymbol{a}_j) + R_D(\boldsymbol{\mathcal{D}})$ として定義できる．ここで $d^2(\mathbf{A}, \mathbf{B})$ は，二点 $\mathbf{A}, \mathbf{B} \in \mathcal{S}^d_{++}$ 間の測地線距離を示す．さらに $R_a(\boldsymbol{a}_j)$ と $R_D(\boldsymbol{\mathcal{D}})$ は，それぞれ係数ベクトル $\boldsymbol{a}_j$ と辞書 $\boldsymbol{\mathcal{D}}$ に対する正則化項であり，前者は $\boldsymbol{a}_j$ のスパース性を誘導し $R_a(\boldsymbol{a}_j) = \lambda||\boldsymbol{a}_j||_1$ 等を使用する．$\lambda > 0$ は重み係数である．ここで，前記問題を最適化するには，辞書学習問題とスパースコーディング問題をそれぞれ交互に解く交互最適化法を使用する．

一方，訓練データである高次元正定値対称行列を，クラス識別能力を向上する低次元正定値対称行列へ埋め込む次元削減手法が提案されている [13]．$\mathbf{X} \in \mathcal{S}^m_{++}$ を次元削減対象の行列とすると，提案手法では射影行列 $\mathbf{U} \in \mathbb{R}^{m\times d} (m > d)$ を学習する．具体的には $f : \mathcal{S}^m_{++} \times \mathbf{U} \in \mathbb{R}^{m\times d} \to \mathcal{S}^d_{++}$ である関数 $f(\mathbf{X}, \mathbf{U}) = \mathbf{U}^T\mathbf{X}\mathbf{U}$ を考える．ここで $\mathbf{X}$ が正定値対称行列であり $\mathbf{U}$ が完全階数行列である場合，射影さ

れた $\mathbf{U}^T\mathbf{X}\mathbf{U}$ は正定値対称行列となり $\mathbf{U}^T\mathbf{X}\mathbf{U} \succ 0$ である．よって $\mathbf{U}$ の完全階数性を保証するため，新たに正規直交制約を付与する．この時，$\mathbf{U}$ は $m \times d$ の正規直交行列から成る集合 $\{\mathbf{U} \in \mathbb{R}^{m\times d} | \mathbf{U}^T\mathbf{U} = \mathbf{I}_d\}$ のシュティーフェル多様体 $\mathrm{St}(d, m)$ となる．

## 4. 正定値対称行列多様体における次元削減と辞書学習の統合学習法

本節では，次元削減と辞書学習の統合学習法を提案する．特に，**3.** で導入された表記に対して “クラス” を考慮して再定義する．$\boldsymbol{\mathcal{X}}_{trn}$ を $\boldsymbol{\mathcal{X}}_{trn} = \{\boldsymbol{\mathcal{X}}_1, \ldots, \boldsymbol{\mathcal{X}}_K\} \in \mathbb{R}^{m\times m\times N}$ と定義する．$\boldsymbol{\mathcal{X}}_k$ は $k$ 番目クラスの訓練データ群であり，$K$ はクラス数である．$\boldsymbol{\mathcal{X}}_k$ は，さらに訓練要素データから構成され $\boldsymbol{\mathcal{X}}_k = \{\mathbf{X}_{k,1}, \ldots, \mathbf{X}_{k,i}, \ldots, \mathbf{X}_{k,I_k}\}$ であり，$I_k$ は $k$ 番クラス内の $i$ 番目行列データ，$\sum_{k=1}^{K} I_k = N$ である．次に，辞書全体を $\boldsymbol{\mathcal{D}} = \{\boldsymbol{\mathcal{D}}_1, \ldots, \boldsymbol{\mathcal{D}}_K\}$ と定義し，$\boldsymbol{\mathcal{D}}_k$ を $k$ 番クラスのサブクラス辞書と定義する．$k$ 番クラスのサブクラス辞書 $\boldsymbol{\mathcal{D}}_k$ は, 訓練データと同様に $\boldsymbol{\mathcal{D}}_k = \{\mathbf{D}_{k,1}, \ldots, \mathbf{D}_{k,h}, \ldots, \mathbf{D}_{k,H_k}\}$ のように構成される．ここで，$H_k$ は $k$ 番クラスのサブ辞書 $\boldsymbol{\mathcal{D}}_k$ を構成する $d \times d$ 正定値対称行列 $\mathbf{D}_{k,h} \in \mathcal{S}^d_{++}$ の個数であり，$\sum_{k=1}^{K} H_k = n$ となる．

以上を元に，本提案手法では辞書テンソル $\boldsymbol{\mathcal{D}}$ と射影行列 $\mathbf{U}$ を同時に学習する．提案モデルにおける学習パラメータは $(\mathbf{U}, \boldsymbol{\mathcal{D}}) \in \mathcal{N}$ と $\mathbf{A} \in \mathbb{R}^{n\times N}$ であり，$\mathcal{N}$ を $\{\mathrm{St}(d, m) \times \prod_{i=1}^{n} \mathcal{S}^d_{++}\}$ で表される直積多様体として新たに定義する．特に，クラス識別性能の向上を考慮した誤差項及び正則化項を考慮することで，最終的な最適化問題を以下に定義する．

$$
\begin{aligned}
\min_{(\mathbf{U},\boldsymbol{\mathcal{D}})\in\mathcal{N}, \mathbf{A}\in\mathbb{R}^{n\times N}_{+}} \quad & J(\mathbf{U}, \boldsymbol{\mathcal{D}}, \mathbf{A}) + \lambda_a J_a(\mathbf{A}) + \lambda_u J_u(\mathbf{U}) \\
& + \lambda_d R_d(\boldsymbol{\mathcal{D}}) + \lambda_s R_s(\mathbf{A}).
\end{aligned}
$$

ここで，$J(\mathbf{U}, \boldsymbol{\mathcal{D}}, \mathbf{A})$ は識別性能を考慮した再現誤差項，$J_a(\mathbf{A})$ は係数に対する重み項，$J_u(\mathbf{U})$ は射影行列に対する重み項 $R_d(\boldsymbol{\mathcal{D}})$ は辞書に対する正則化項, $R_a(\mathbf{A})$ は係数に対するスパース性誘導の正則化項である．$\lambda_a, \lambda_u, \lambda_d, \lambda_s$ はそれぞれに対する正のスカラー係数である．ここで，$J(\mathbf{U}, \boldsymbol{\mathcal{D}}, \mathbf{A})$ は，全体クラス $\boldsymbol{\mathcal{D}}$ と $\mathbf{A}$ だけでなく，各クラスの $\boldsymbol{\mathcal{D}}_k$ と $\mathbf{A}_k$ に着目した誤差項を考慮する．また $J_a(\mathbf{A})$ と $J_u(\mathbf{U})$ は，クラス識別性能向上を考慮してクラス内近接グラフを考慮した重みを与える．

本問題の最適化にあたっては，辞書テンソル $\boldsymbol{\mathcal{D}}$ と射影行列 $\mathbf{U}$ の学習と，係数行列 $\mathbf{A}$ の学習をそれぞれサブ問題として分離し，交互最適化法により最適化を行う．具合的には，前者については，モデルパラメータ空間 $\mathcal{N}$ が直積多様体であることから “リーマン多様体上での最適化問題” として最適化を行う．詳細については [14], [15] に詳しい．一方，後者については凸空間への射影勾配法により解く．

## 5. まとめ

本稿では，正定値対称行列から構成される正定値対称行列から構成されるリーマン多様体上に位置する信号を対象として，クラス識別タスクにおける性能向上を目指し，次元削減と辞書学習の統合学習法について提案した．いくつかの評価結果は発表時に示したい．

### 参考文献

[1] Elad, M.: *Sparse and Redundant Representations - From Theory to Applications in Signal and Image Processing*, Springer (2010).
[2] 笠井裕之：スパースコーディングの研究動向，情報処理学会研究報告オーディオビジュアル複合情報処理，Vol. 2014-AVM-84, No. 8, pp. 1–10 (2014).
[3] Nguyen, H. V., PatelNasser, V. M. and Nasrabadi, N.M.and Chellappa, R.: Sparse Embedding: A Framework for Sparsity Promoting Dimensionality Reduction, *ECCV2012*, pp. 414–427 (2012).
[4] Feng, Z., Yang, M. Zhang, L., Liu, Y. and Zhang, D.: Joint discriminative dimensionality reduction and dictionary learning for face recognition, *Pattern Recognition*, Vol. 46, No. 8, pp. 2134–2143 (2013).
[5] Yang, B. Q., Gu, C.-C., Wu, K.-J., Zhang, T. and Guan, X.-P.: Simultaneous dimensionality reduction and dictionary learning for sparse representation based classification, *Multimedia Tools and Applications*, Vol. 76, No. 6, pp. pp 8969–8990 (2016).
[6] Liu, W., Yu, Z., Wen, Y., Lin, R. and Yang, M.: Jointly learning non-negative projection and dictionary with discriminative graph constraints for classification, *BMVC* (2016).
[7] Foroughi, H., Ray, N. and Zhang, H.: Object classification with joint projection and low-rank dictionary learning, *arXiv preprint arXiv:1612.01594* (2016).
[8] Pang, Y., Yuan, Y. and Li, X.: Gabor-based Region covariance matrices for face recognition, *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 18, No. 7, pp. 989–993 (2008).
[9] Tuzel, O., Porikli, F. and Meer, P.: Region covariance: a fast descriptor for detection and classification, *ECCV2006* (2006).
[10] Pennec, X., Fillard, P. and Ayache, N.: A Riemannian Framework for Tensor Computing, *Int. Jornal of Computer Vision*, Vol. 66, No. 1, pp. 41–66 (2006).
[11] Bhatia, R.: *Positive definite matrices*, Princeton series in applied mathematics, Princeton University Press (2007).
[12] Cherian, A. and Sra, S.: Riemannian Dictionary Learning and Sparse Coding for Positive Definite Matrices, *IEEE Trans. Neural Netw. Learn. Syst.* (2016).
[13] Harandi, M., Salzmann, M. and Richard, H.: Dimensionality Reduction on SPD Manifolds: The Emergence of Geometry-Aware Methods, *IEEE Trans. Pattern Anal. Mach. Intell.* (2017).
[14] Absil, P.-A., Mahony, R. and Sepulchre, R.: *Optimization Algorithms on Matrix Manifolds*, Princeton University Press (2008).
[15] 佐藤寛之，笠井裕之：リーマン多様体上の最適化の基本と最新動向（解説記事），システム制御情報学会誌「システム／制御／情報」(掲載予定).

 