

2024 年度 応用プログラミング実験

第 4 回 マルチメディア情報検索 実験レポート

学修番号: 22140026
谷 知拓 - Tomohiro Tani^{*1}

レポート提出日 : 2024-05-09

^{*1} 東京都立大学 システムデザイン学部 情報科学科
mail@taniii.com

1 はじめに

本書では、『応用プログラミング実験』第4回の実施報告を行う。
部分空間法を用いて手書きの数字の認識を行う。

2 実験環境

実験環境は以下の通りである。

- OS^{*2}: macOS Ventura 13.4.1
- CPU^{*3}: Apple M2 arm64^{*4}
- メインメモリ・ビデオメモリ共通: 16GB ユニファイドメモリ^{*4}
- 実行環境 (matlab): matlab R2023b (23.2.0.2380103) 64-bit (maca64)

3 実験の概要

本実験では、部分空間法を用いて手書きの数字の認識を行う。

具体的には、部分空間の次元を $r = 1, \dots, 20$ と変化させたときの、認識率について調べる。

また、部分空間法を用いて手書きの数字の認識を行う際に、どのような手法を用いるか、また、その手法を用いることでどのような効果があるかを考察する。

4 部分空間法の原理

部分空間法は、パターン認識において広く用いられる手法であり、高次元データの持つ構造をより低次元の空間に射影し、その射影された空間でのデータの特性を利用して識別を行う手法である。

この手法を用いて、データの次元を削減することで、様々なパターン変動 (認識対象のパターンが様々な要因により変化すること) を、小さい辞書で表現することができ、高速かつ高精度な認識が可能となる。

部分空間法を用いたパターン認識の基本的な流れは、まず学習段階において、クラスごとに、エントロピーが最小化されるような部分空間 (より低次元な空間) を見つける。

そして、識別段階において、データをその部分空間に射影し、射影されたデータの特性を利用して識別を行う。

以下では、具体的な手法について説明する。

4.1 学習

学習段階においては、クラスごとに、エントロピーが最小化されるような部分空間を見つける。

式1で表されるエントロピーが最小となるような、部分空間軸を求める。

^{*2} Operating System

^{*3} Central Processing Unit

^{*4} <https://www.apple.com/jp/macbook-air-13-and-15-m2/specs/>

$$\min_{\mathbf{u}_i} - \sum_{i=1}^d p(\mathbf{u}_i) \log p(\mathbf{u}_i) \quad (1)$$

ただし、 $p(\mathbf{u}_i)$ は、部分空間軸 \mathbf{u}_i におけるデータの分布を表す確率密度関数である。

ここで、データから最適な部分空間軸 (部分空間の基底ベクトル) を求めるために、主成分分析 (PCA) や線形判別分析 (LDA) を用いる方法がある。

主成分分析において、最も分散が大きい方向を求める場合や、線形判別分析において、クラス間分散を最大化しつつクラス内分散を最小化する方向を見つける場合に、**制約付き最適化問題** を解く必要があり、これに **ラグランジュの未定乗数法** を用いることがある。

ラグランジュの未定乗数法は、以下のように解くことができる。
仮に、以下のような制約付き最適化問題を考える。

制約条件 $g(\mathbf{x}) = 0$ のもとで関数 $f(\mathbf{x})$ の極地を求めよ

まず、ラグランジュ関数を以下のように定義する。

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x}) \quad (2)$$

このラグランジュ関数を最小化するために、以下のような方程式を解く。

$$\frac{\partial}{\partial \mathbf{x}} L = \nabla f - \lambda \nabla g = \mathbf{0}, \quad \frac{\partial L}{\partial \lambda} = 0 \quad (3)$$

この方程式を解くことで、制約条件のもとでの関数の極値を求めることができる。

部分空間法では、 $\mathbf{u}^\top \mathbf{u} = 1$ を制約条件として、 $\mathbf{u}^\top \mathbf{A} \mathbf{u}$ の最大値を求める問題を解く必要があることが頻出である。

$$\begin{aligned} \max_{\mathbf{u}} \quad & \mathbf{u}^\top \mathbf{A} \mathbf{u} \\ \text{s.t.} \quad & \mathbf{u}^\top \mathbf{u} - 1 = 0 \end{aligned} \quad (4)$$

ただし、 \mathbf{A} は半正定値対称行列である。

このような問題を解くために、ラグランジュの未定乗数法を用いることがある。
ラグランジュ乗数 λ を導入し、ラグランジュ関数を以下のように定義する。

$$L(\mathbf{u}, \lambda) = \mathbf{u}^\top \mathbf{A} \mathbf{u} + \lambda (\mathbf{u}^\top \mathbf{u} - 1) \quad (5)$$

制約条件のもとで、関数が極値を取る点は次式を満たす。

$$\frac{\partial L}{\partial \mathbf{u}} = \mathbf{0}, \quad \frac{\partial L}{\partial \lambda} = 0 \quad (6)$$

$\frac{\partial L}{\partial \mathbf{u}} = 2\mathbf{A}\mathbf{u} - 2\lambda\mathbf{u} = \mathbf{0}$ より、 $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$ が成り立つ。

4.2 識別

続いて、識別段階においては、データを学習によって得られた部分空間に射影し、射影されたデータの特性を利用して、クラス分けを行う。

識別においては、以下のような問題設定を考える。

1. 総数 n 個の d 次元訓練パターン x_i が与えられており、それらが C 個のクラスのいずれか 1 つに分類されている。
2. クラス j に属する訓練パターンを良く近似することのできる部分空間を張る正規直交ベクトルの行列を \mathbf{U}_j とする。

このとき、未知パターン x が与えられたとき、属するクラスを求めるための、部分空間法の識別則は以下のように定義できる。

$$\max_{j=1,\dots,C} \{\|\mathbf{U}_j^T \mathbf{x}\|\} = \|\mathbf{U}_k^T \mathbf{x}\| \Rightarrow \mathbf{x} \in \text{class } k \quad (7)$$

これは、未知パターン x を各クラスの部分空間に射影し、射影されたデータのノルムが最大となるクラスに分類するということを意味している。

部分空間法よりシンプルな識別則として、最小距離法が考えられるが、これは、 $d_j^2 = \|\mathbf{x}\|^2 - \|\mathbf{U}_j^T \mathbf{x}\|^2$ が最小となるクラスに分類するというものである。

ただし、 d_j は、未知パターン x とクラス j の部分空間との距離を表す。

ここで、 $\|\mathbf{x}\|^2$ は、クラスに関わらず共通であるため、 $\|\mathbf{U}_j^T \mathbf{x}\|^2$ が最大となるクラスを探せば良い。したがって、最小距離法による識別則と、部分空間法による識別則は、矛盾しない。

このようにして、部分空間法を用いて、未知パターンをクラス分けすることができる。

5 実験結果

実験結果は以下の通りである。

図 1 は、次元 $r = 1, \dots, 20$ と認識率の関係をグラフにしたものである。

また、表 1 は、次元 $r = 1, 5, 10, 15, 19, 20$ における認識率を示したものである。なお、 $r = 19$ において、認識率が最大となった。

6 考察

実験結果より、 $r = 1, \dots, 19$ の範囲において、次元 r を増やすことで、認識率が向上することがわかった。

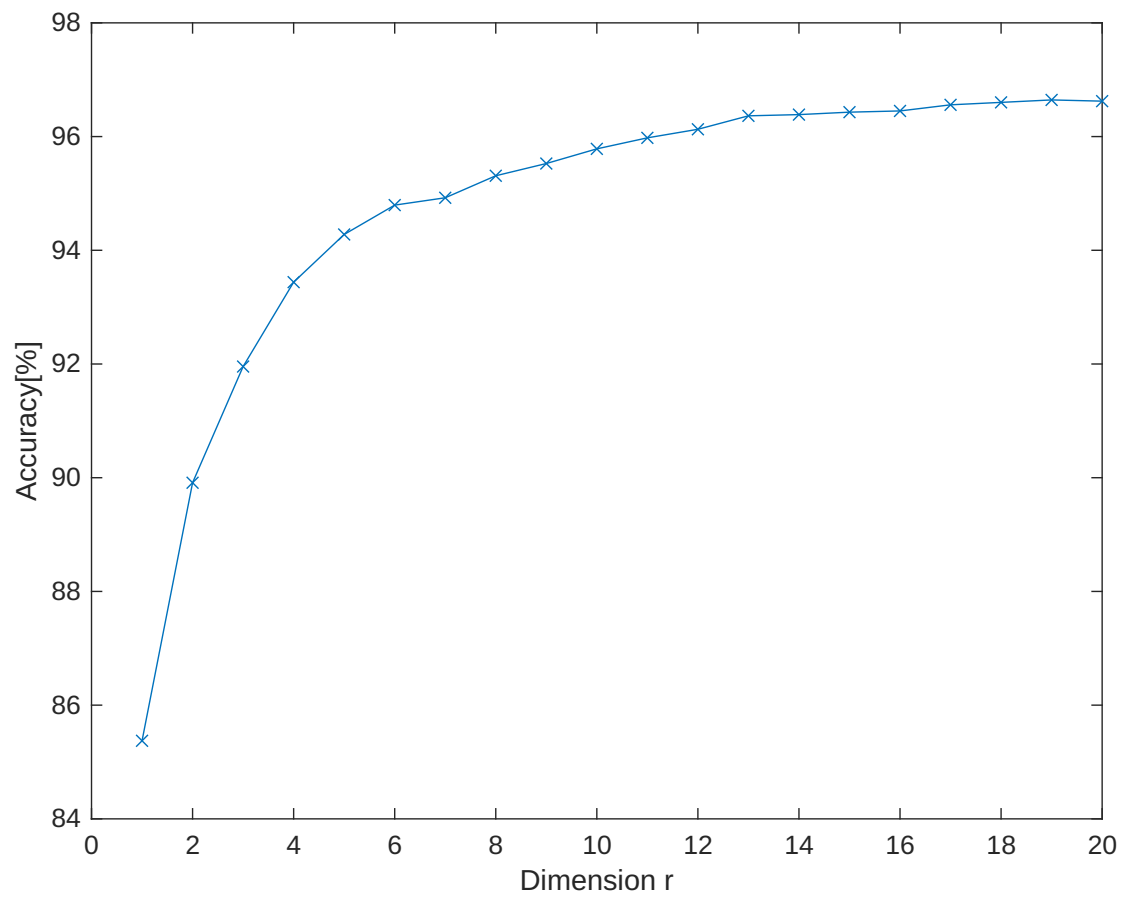
これは、部分空間法において、次元 r を増やすことで、データの持つ構造をより正確に表現できるためであると考えられる。

一方で、次元が増えるのに伴い、認識率の上昇はゆるやかになっている。また、 $r = 19$ と $r = 20$ を比較すると、 $r = 20$ の方が認識率が低いことがわかる。

これは、部分空間の次元 r を増やしすぎることにより、過学習が発生しているものと推察できる。

過学習は、次元 r が増えることにより、学習データの構造を過剰に反映してしまい、未知データに対する汎化性能が低下することにより生じるものである。

したがって、部分空間法においては、次元 r を適切に設定することが重要である。

図 1 次元 ($r = 1, \dots, 20$) と認識率の関係表 1 次元 $r = 1, 5, 10, 15, 19, 20$ における認識率

次元 r	認識率
1	85.3732
5	94.2783
10	95.7840
15	96.4293
19	96.6444
20	96.6229