

主専攻実験
数理モデリングとアルゴリズム

課題2：画像認識

1 はじめに

本実験では手書き文字をコンピュータに判別させる方法について学ぶ。コンピュータによる手書き文字の判別はパターン認識の分野において一般的な問題である。このような問題に対する解法には平均値を用いる方法や、主成分分析を用いる方法などがある。本実験では MATLAB を用いてノルムの計算や特異値分解を行い、手書き数字の画像認識を行う。画像は 16×16 のグレースケール画像を用いる。

2 実験課題 2-1：平均値を用いた画像認識

画像認識は、まずコンピュータにグループ分けされた画像（学習データ）を学習させる処理を行い、その後未知の画像データ（テストデータ）がどのグループに属するかを判定させる処理を行う。本課題では学習処理に学習データの平均値を用い、判定処理にベクトルのノルムを用いる。この方法によって、学習処理で得られるデータや画像認識の正答率がどのようになるのかを調べる。

2.1 画像データ

16×16 の画像データは以下の HP からダウンロードすることができる。

<http://www.cad.zju.edu.cn/home/dengcai/Data/MLData.html>

図1のような画像データが複数保存されている。

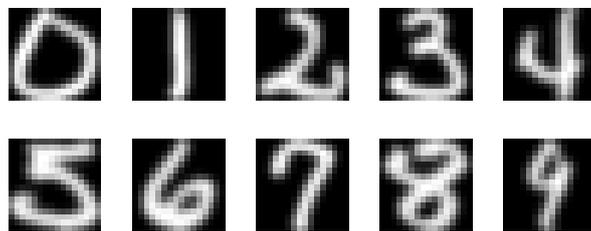


図 1: 手書き数字画像

テストデータと学習データは以下のプログラム loadUSPS.m で作成することができる。

ソースコード 1: 学習データとテストデータ作成プログラム

```
1 function [train_data,train_id,test_data,test_id] = loadUSPS();  
2  
3 train_data=[]; train_id=[]; test_data=[]; test_id=[];
```

```

4
5 try
6     load USPS.mat;
7 catch
8     disp('Error: the file usps_resampled.mat was not found.')
9     return
10 end
11
12 train_data_size=1707;
13 test_data_size=2007;
14
15 if test_data_size>2007
16     disp('Error: number of test data was too large.')
17     return;
18 end
19
20 train_data=fea(1:train_data_size,:);
21 train_id=gnd(1:train_data_size);
22
23 test_data=fea(7292:7292+test_data_size-1,:);
24 test_id=gnd(7292:7292+test_data_size-1);

```

実行する際は以下のようにする.

```
[train_data,train_id,test_data,test_id] = loadUSPS;
```

出力値について `train_data` は学習データであり、各行ごとに画像データが格納されている。`train_id` は `train_data` が属するグループを示している。`test_data`, `test_id` はテストデータと、その画像が属するグループを示している。このプログラムでは学習データを 1707 個、テストデータを 2007 個出力する。

2.2 学習処理

16×16 の画像データを 256 次元の画像ベクトルとして扱う。本課題では学習処理にグループごとの学習データの平均値を用いる。この学習処理によって得られる画像データがどのようになるのかを調べよ。以下の関数で画像データを描画することができる。

```

r_data=reshape(data,16,16);
image(r_data,'CDataMapping','scaled'); colormap gray;

```

`data` は 256 次元の画像データのベクトルである。関数 `reshape` は行列データの行サイズ、列サイズを変更する関数である。この関数によりベクトルのデータは 16×16 次元の行列データに変換している。関数 `image` は行列データの値を色として描画する関数である。

2.3 判定処理

学習処理によって得られるデータ \mathbf{x}_j ($j = 0, 1, \dots, 9$) とテストデータ \mathbf{y} との差のノルム計算

$$\|\mathbf{x}_j - \mathbf{y}\|_2$$

をする事で、テストデータがどのグループに近いのかを判別することができる。この判別処理による正答率を調べよ。

3 実験課題 2-2：主成分分析を用いた画像認識

前課題では学習処理に学習データの平均値を用い、判定処理にベクトルのノルムを用いた。本課題では学習処理に主成分分析を用い、判定処理に行列のノルムを用いる。この方法によって、学習処理で得られるデータや画像認識の正答率がどのようになるのかを調べる。前課題と同様 16×16 のグレースケール画像を用いて判別を行う。

3.1 主成分分析

主成分分析とは複数のデータを互いに相関の無いいくつかの特性値にまとめる手法である。例として2値のデータの主成分分析について説明する。図2は2値のデータの値グラフ化したものである。青点が各データを表す。これらのデータ群の特徴を反映している直線は各データのばらつきをよく評価している直線となるため、 L_1 がデータ群の特徴を一番反映している。このとき L_1 を第1主成分、 L_2 を第2主成分と呼ぶ。本課題ではこの主成分分析を用いて、学習処理を行う。主成分分析は行列の特異値分解を用いることで解析することができる。

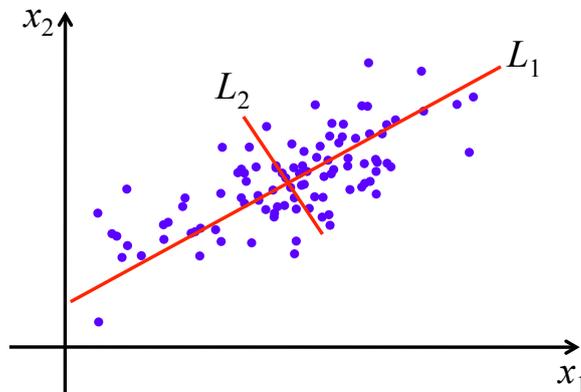


図 2: 2 値データの主成分分析

3.2 行列の特異値分解

A を $m \times n$ 行列とする。このとき行列 A は以下の行列に分解することができる。

$$A = U\Sigma V^T$$

ここで、 $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m]$ は $m \times m$ 直交行列、 V は $n \times n$ 直交行列、 Σ は以下の対角行列となる。

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{pmatrix}$$

$$\sigma_1 \geq \sigma_2 \geq \dots \sigma_n \geq 0$$

3.3 学習処理

16 × 16 の画像データを 256 次元の画像ベクトルとして扱う。本実験ではグループごとの学習データの主成分分析を行う。グループごとの学習データを結合させた 256 × n 行列を A とする。n は同グループの学習データの数である。この行列に対して特異値分解を行うことで、学習データの特性値を得ることができる。u_i は第 i 主成分の画像データを示しており、σ_i は第 i 主成分の寄与率を示している。この学習処理によって得られる画像データ u_i がどのようなのかを調べよ。

3.4 判定処理

学習処理によって得られるデータ $U_k^{(j)} = [\mathbf{u}_1^{(j)}, \mathbf{u}_2^{(j)}, \dots, \mathbf{u}_k^{(j)}]$ (k は任意の値, j はグループ番号) とテストデータ \mathbf{y} との差のノルム計算

$$\|(I - U_k^{(j)} U_k^{(j)T})\mathbf{y}\|_2$$

をする事で、テストデータがどのグループに近いのかを判別することができる。今回は $k = 10$ とし、この判別処理による正答率を調べよ。

4 実験課題 3-3：応用課題

前課題では画像認識に平均値を利用した方法や、主成分分析を利用した方法について述べた。これらの方法は学習データ数や、 $\mathbf{u}_k^{(j)}$ の数によって正答率が変化する。本実験ではそれらのデータを変えた際に正答率がどのように変化するのかを調べよ。また手書き数字を実際に作成し、正しく判別されるか実験せよ。