数理アルゴリズム　演習課題 9　提出日：2017 年 12 月 14 日
201611350　江畑拓哉

1　課題 1

1.1　$\boldsymbol{x}$ が最小二乗問題 $\min_{\boldsymbol{x}\in\mathbb{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ の解であるとき、かつそのときに限り正規方程式 $A^TA\boldsymbol{x}=A^T\boldsymbol{b}$ の解であることを示せ。

$$\begin{aligned}\|A\boldsymbol{x}-\boldsymbol{b}\|_2 &= (A\boldsymbol{x}-\boldsymbol{b})^T(A\boldsymbol{x}-\boldsymbol{b})\\ &= (\boldsymbol{x}^TA^T-\boldsymbol{b}^T)(A\boldsymbol{x}-\boldsymbol{b})\\ &= \boldsymbol{x}^TA^TA\boldsymbol{x}-2\boldsymbol{x}^TA^T\boldsymbol{b}+\boldsymbol{b}^T\boldsymbol{b}\\ \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}\boldsymbol{x}^TA^TA\boldsymbol{x}-2\boldsymbol{x}^TA^T\boldsymbol{b}+\boldsymbol{b}^T\boldsymbol{b} &= 2A^TA\boldsymbol{x}-2A^T\boldsymbol{b}\end{aligned}$$

よって $\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ を最小化するためには、 $A^TA\boldsymbol{x}=A^T\boldsymbol{b}$ であることが必要であると言える。

1.2　行列 $A\in\mathbb{R}^{m\times n}, rank\ A=n$ の特異値分解は

$$\begin{gathered}A=U\Sigma V^T,\ U\in\mathbb{R}^{m\times n},\ \Sigma\in\mathbb{R}^{n\times n},\ V\in\mathbb{R}^{n\times n},\\ U^TU=I,\ VV^T=V^TV=I\end{gathered}$$

と表せる。このとき、最小二乗問題 $\min_{\boldsymbol{x}\in\mathbb{R}^n}\|\boldsymbol{b}-A\boldsymbol{x}\|$ の解は $\boldsymbol{x}=V\Sigma^{-1}U^T\boldsymbol{b}$ であることを示せ。

(1-1) より、

$$\begin{aligned}A^TA\boldsymbol{x} &= A^T\boldsymbol{b}\\ &\Leftrightarrow \boldsymbol{x}=(A^TA)^{-1}A^T\boldsymbol{b}\\ &\Leftrightarrow \boldsymbol{x}=(V\Sigma U^TU\Sigma V^T)^{-1}A^T\boldsymbol{b}\\ &\Leftrightarrow \boldsymbol{x}=(V\Sigma^2V^T)^{-1}A^T\boldsymbol{b}\\ &\Leftrightarrow \boldsymbol{x}=V\Sigma^{-2}V^TV\Sigma U^T\boldsymbol{b}\\ &\Leftrightarrow \boldsymbol{x}=V\Sigma^{-1}U^T\boldsymbol{b}\end{aligned}$$

1.3　行列 A , ベクトル $\boldsymbol{b}$ を

$$A=\begin{pmatrix}1&1&1\\1&2&2\\1&2&3\\1&2&3\end{pmatrix} b=\begin{pmatrix}4\\5\\6\\7\end{pmatrix} \tag{1}$$

とし、最小二乗問題 $\min_{\boldsymbol{x}}\|\boldsymbol{b}-A\boldsymbol{x}\|_2$ を特異値分解を用いて解く Scilab プログラムを作成し、解を求めよ。

行列の宣言

```
A = [1, 1, 1; 1, 2, 2; 1, 2, 3; 1, 2, 3];
b = [4; 5; 6; 7];
```

最小二乗問題を解く関数

```
function [x] = solveLSP(U, S, V, b)
  x =  V * inv(S) * U' * b
endfunction
```

実行

```
[U, S, V] = svd(A, "e")
x = solveLSP(U, S, V, b)
```

解

```
  -->    x = solveLSP(U, S, V, b)
x  =


  3.
 -0.5
  1.5
```

2　課題 2

手書き数字の画像データセットである MNIST を用いて手書き数字の画像認識を行う。MNIST の行列データは manaba にアップロードされているため、ダウンロードして用いること。ファイルから得られる行列 X は画像データであり、各列に 1 枚の画像が格納されている。Y は各画像が属するグループを示している。 test_X, test_Y はそれぞれテストデータと各画像が属するグループを示している。今回の課題では訓練データとテストデータを 10,000 個ずつ用いる。　manaba にアップロードされている showMNISTimg.sci を使用すると行列 X の先頭から i 番目の列に格納されている訓練データを画像として表示することが出来る。

数理アルゴリズム　演習課題 9　提出日：2017 年 12 月 14 日
201611350　江畑拓哉

2.1　行列 $X \in \mathbb{R}^{m\times n}, Y \in \mathbb{R}^{10\times n}$ に対し、最小二乗問題

$$\min_{W\in\mathbb{R}^{10\times m}} \|Y - WX\|_F$$

の特異値分解を用いて解く Scilab プログラムを作成せよ。行列 X の特異値を $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$ としたとき、k を $\sigma_k/\sigma_1 \geq 10^{-14}$ を満たすような最大の整数とする。k 個の特異値とそれぞれの特異値に対応する k 本の特異ベクトルを用いて行列 X の低ランク近似を行うこと。ここでは $m = 784, n = 10,000$ である。

低ランク近似した X を得るための関数

```
function [X_k] = getX_k(U, S, V, k, X)
  X_k = zeros(size(X)(1), size(X)(2));
  for i=1:k,
   X_k = X_k + S(i, i) * U(1:$, i) * V(1:$, i)';
   if S(i, i) < S(1, 1) * 10e-14 then break; end,
  end
endfunction
```

ブロック最小二乗問題を解くための関数

```
function [W] = solveBLSP(U, i_S, V, Y)
  W = Y * V * i_S * U';
endfunction
```

対角行列の逆行列を求める関数

```
function [S_] = inv_(S)
  S_ = zeros(size(S)(1), size(S)(2))
  for i=1:rank(S),
    S_(i, i) =  1/S(i, i)
  end
endfunction
```

実行

```
[U, S, V] = svd(X);
k = 784;
```

```
X_k = getX_k(U, S, V, k, X);
[U, S, V] = svd(X_k, "e");
W = solveBLSP(U, inv_(S), V, Y);
```

2.2　得られた行列 W を用いてテストデータの画像がどのグループに属するか判定することができる。**manaba** にアップロードされている **check.sci** を用いて **(2-1)** で得られた **W** に対して処理判定を行い、画像認識の正答率を求めよ。**check(W, test_X, test_Y)** と実行すると正答率を求めることができる。

```
--> check(W, test_X, test_Y)
 ans  =

    0.8305
```

2.3　**(2-2)** では訓練データを **10,000** 個用いて画像認識を行ったが、訓練データ数を帰ることで画像認識の正答率は変化する。訓練データ数を行列の先端の列から $1,000, 2,000, \ldots, 10,000$ 個としたときの画像認識の正答率を求める **Scilab** プログラムを作成し、それぞれの訓練データ数における正答率をグラフに描画せよ。

```
c = zeros(0);
for i=1:10,
  [U, S, V] = svd(X(1:$, 1:1000*i));
  k = 784;
  X_k = getX_k(U, S, V, k, X(1:$, 1:1000*i));
  [U, S, V] = svd(X_k, "e");
  W = solveBLSP(U, inv_(S), V, Y(1:$, 1:1000*i));
  c(i) = check(W, test_X, test_Y);
end
plot(c, linspace(1000, 10000, 10))
```

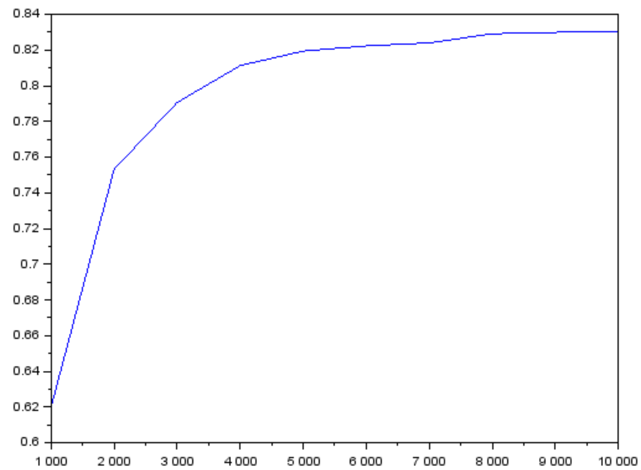

Figure 1: 訓練データ数と正答率のグラフ

2.4 行列 $\mathbf{X}$ の低ランク近似を行う際の整数 $\mathbf{k}$ を $\sigma_k/\sigma_1 \geq \delta$ を満たすような最大の整数とする。このとき、閾値 δ を $0.005, 0.01, 0.015, \ldots, 0.1$ と変えたときの画像認識の正答率を求める Scilab プログラムを作成し、各しきい値に対する正答率をグラフに描画せよ。ただし、訓練データは 10,000 個用いること。

閾値を変更できるようにした低ランク近似の関数

```
function [X_k] = getX_k_d(U, S, V, k, X, delta)
  X_k = zeros(size(X)(1), size(X)(2));
  for i=1:k,
   X_k = X_k + S(i, i) * U(1:$, i) * V(1:$, i)';
   if S(i, i) < S(1, 1) * delta then break; end,
  end
endfunction
#+end_src scilab


#+begin_src scilab
c = zeros(0);
deltas = linspace(0.005, 0.1, 20)
```

```
for i=1:20,
  [U, S, V] = svd(X);
  k = 784;
  X_k = getX_k_d(U, S, V, k, X, deltas(i));
  [U, S, V] = svd(X_k, "e");
  W = solveBLSP(U, inv_(S), V, Y);
  c(i) = check(W, test_X, test_Y);
end
plot(deltas, c)
```

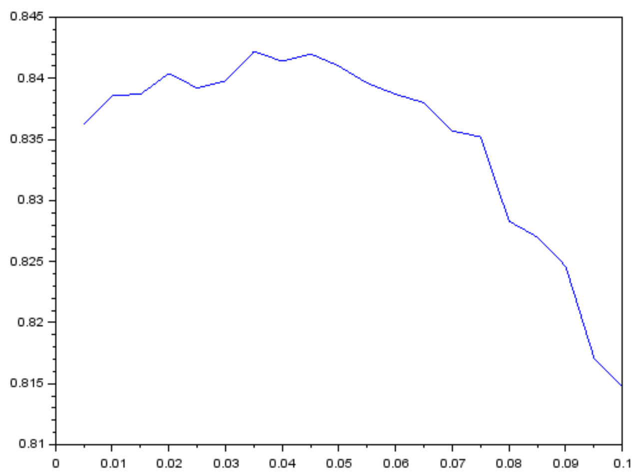

Figure 2: 閾値と正答率のグラフ