

第 4 章

第 3 章 特徴抽出 演習問題

準備

1. 平均値を求める (mean 関数)
2. 分散を求める (variance 関数)
3. 標準偏差を求める (stdev 関数)
4. 行列に対し、列ごとに平均値・分散を求める
5. 行列に対し、共分散行列を求める (cov 関数)
6. 主成分分析を行う (pca 関数)
7. 複数のグラフの表示場所を設定する (subplot 関数)
8. グラフを表示する (plot2d 関数)
9. CSV ファイルを読み込む (csvRead 関数)

```
-->a = [5 8 7 2 3];

-->mean(a)
ans =
    5.

-->variance(a)
ans =
    6.5

-->stdev(a)
ans =
    2.5495098

-->M = [5 8 7 2 3; 5 4 6 8 9]'
M =
    5.    5.
    8.    4.
    7.    6.
    2.    8.
    3.    9.

-->mean(M, 'r')
ans =
    5.    6.4

-->variance(M, 'r')
```

```

ans =
    6.5    4.3

-->cov(M)
ans =
    6.5    -4.5
   -4.5    4.3

-->[l f c] = pca(M)
c =                                     // 主成分
   -0.4773960    0.4773960
   -1.6504435    -0.0136571
   -0.6910991    -0.4183013
    1.3776458    0.2864548
    1.4412928    -0.3318924
f =                                     // 固有ベクトル
   -0.7071068    -0.7071068
    0.7071068    -0.7071068
l =                                     // 固有値・全体に対する比
    1.8511804    0.9255902
    0.1488196    0.0744098

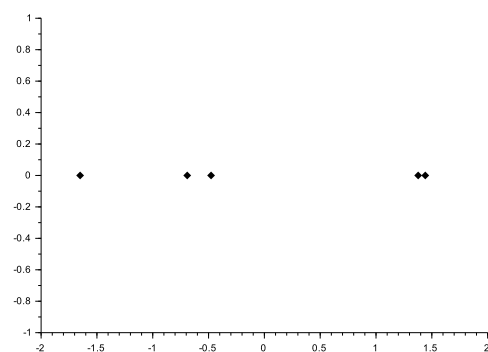
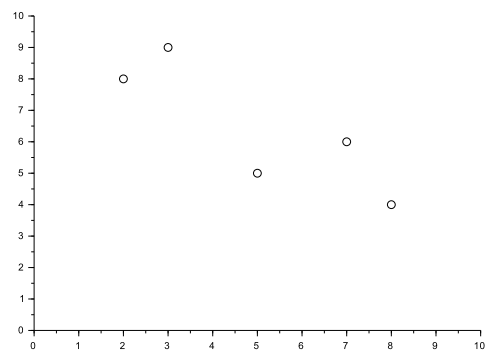
-->subplot(2, 1, 1)

-->plot2d(M(:,1), M(:,2), style=-9, rect=[0,0,10,10])

-->subplot(2, 1, 2)

-->plot2d(c(:,1)', zeros(1,length(c(:,1)))), style=-4, rect=[-2,-1,2,1])

--> csvRead('iris.csv');
```



実践演習 4-1

ソースコード 3.1 の空欄を埋め、データの標準化と主成分分析を行う Scilab のコードを完成させよ。

```
clear;

x=[3 2; 3 4; 5 4; 5 6];
[n d] = (ア) (x);
// 元データの表示
subplot(2,1,1);
plot2d(x(:,1), x(:,2), style=-9, rect=[0,0,8,8])

// 標準化
m = (イ) (x, 'r');
s = (ウ) (x, 'r');
normX = (x - repmat( (エ) , [n,1])) ./ repmat( (オ) , [n,1]);

// 標準化後のデータの表示
subplot(2,1,2);
plot2d(normX(:,1), normX(:,2), style=-10, rect=[-2,-2,2,2])

// 主成分分析
[lambda, facpr, comprinc] = (カ) (normX);
z = normX * facpr(:,1);
disp(z)
```

実践演習 4-2

アヤメの分類問題のサンプルデータ iris.csv を読み込み、4次元のデータを2次元で可視化せよ。