深圳大学研究生课程：模式识别理论与方法

课程作业实验报告

实验名称：模式类实验数据的生成

实验编号：Proj01-01

签　　名：

姓　　名：夏荣杰

学　　号：2170269107

截止提交日期：2018 年 3 月 23 日

摘要：本次实验的主要内容是针对模式识别研究中用到的随机数和随机向量的生成，以及计算随机数的概率密度值，进行 MATLAB 的仿真实验。产生低维或高维的随机数据是其他模式实验的基础。实验一主要是生成一维、二维和三维区间均匀分布的随机变量，并利用直方图或者二维、三维散点图进行直观的显示；实验二主要是生成服从不同均值矢量和协方差矩阵的高斯分布的一维、二维和三维随机矢量，并画出直方图和二维、三维散点图进行直观显示；实验三主要是利用 mvnpdf 函数计算概率密度函数值，并绘制二维高斯随机矢量的概率密度函数值的三维曲面图。本次实验涉及输入数据生成和数据观测，为以后的研究提供条件。

一、 背景技术 或 基本原理

在模式识别的研究中，产生模拟随机数据时其他模式分类实验的基础，从低维数据到高维数据，都可以用 MATLAB 提供的函数产生。

在 Matlab 中提供了很多产生随机数和随机向量的函数，以及计算随机函数的概率密度值的函数。下面是几个较常用的函数：

rand()生成(0,1)区间的均匀分布随机数；

randn()生成高斯分布随机数；

mvnrnd()生成多元高斯分布的随机向量矩阵；

mvnpdf()计算多元高斯分布的概率密度函数值；

下面是本实验中用到的绘图函数：

hist(data, xvalues)：直方图统计，xvalues 为直方图统计的范围；

scatter(x, y)：二维画散点图，x，y 为散点对应的坐标矩阵；

scatter3(x, y, z)：三维画散点图，x，y，z 为散点对应的坐标矩阵；

meshgrid(x, y)：以 x、y 向量为基准，产生在 x-y 平面的各栅格点坐标值的矩阵；

mesh(x, y, p): 以 x、y 为自变量，p 为因变量画三维网线图。

二、 实验方法 或 算法流程步骤

本实验用到的方法和步骤如下：

实验 1：产生均匀分布的随机数据主要是使用 rand()函数，该函数产生的是一维区间(0,1)上的“均匀随机数”，对该函数加以变形修改拓展到我们所需要的区间和维数。

① 在一维区间[a, b]中，生成 n 个均匀分布的随机数，可以利用下面的公式

$$r = a + (b - a) .* rand(n, 1)$$

② 一维区间内的随机点可以拓展到多维，例如在二维区间[a, b] × [c, d]中生成随机点，可以利用下面的公式：

$$[x, y] = [a + (b - a) .* rand(n, 1), c + (d - c) .* rand(n, 1)]$$

实验 2：生成一元特定均值和方差的高斯分布的随机数，主要是使用 randn()函数，该函数主要是生成均值为 0，方差为 1 的“标准高斯分布随机数”,对其变形修改成我们需要的特定均值和方差。

① 生成 n 个均值为 u，方差为σ^2的高斯分布的随机数，可以利用下面的公式:

$$r = u + sqrt(\sigma^2) * randn(n, 1)$$

② 对于多元高斯分布的随机向量矩阵，MATLAB 中有提供专门的函数 mvnrnd()，指定参数均值矢量和协方差矩阵即可生成。

实验 3：先确定任意的均值矢量和协方差矩阵，然后再利用 meshgrid()函数生成一个二维网格，利用 mvnpdf()计算出概率密度函数值，最后用 mesh()函数画出三维曲线图。

三、 实验结果

1. 实验 1 结果如下：

1）在一维区间[10, 70]中，生成 1000 个均匀分布的随机数，然后统计并绘制这些数的

直方图，如图 1-1；

2）在二维区间[1, 5]□ [20, 30]中，生成 5000 个均匀分布的二维随机点，并绘制出它们的二维散点图，如图 1-2；

3）在三维区间[10, 50]□ [30, 60]□ [10, 15]中，生成 10000 个均匀分布的三维随机点量，并绘制出它们的三维散点图，如图 1-3。

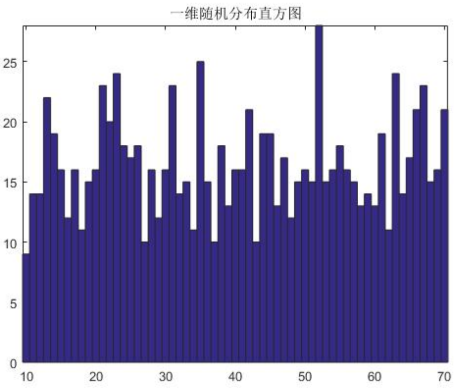

图 1-1

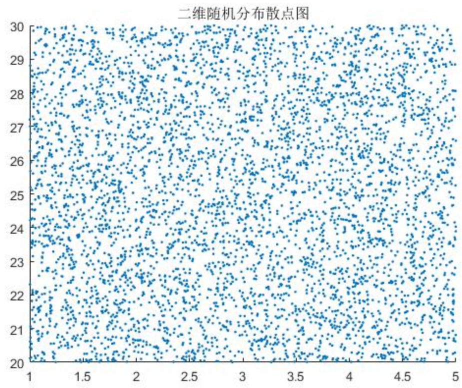

图 1-2

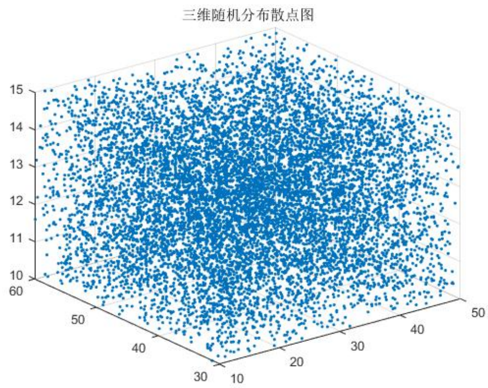

图 1-3

2. 实验 2 结果如下：

1）生成两组各 1000 个具有不同均值（μ1=0; μ1=5）和方差（σ1 =5; σ2 =10）的一维高斯分布的随机数，然后统计并绘制这些点的直方图，如图 2-1；

2）生成三组各 1000 个具有不同均值矢量（μ1=[3, 3]; μ2=[5, 5]; μ3=[5, 5]）和协方差矩阵（σ1=[1 0; 0 1]; σ2=[2 0; 0 2]; σ3=[3 0; 0 3]）的二维随机矢量，并绘制出它们的二维散点图，如图 2-2；

3）生成五组各 1000 个具有不同均值矢量（μ1 = [1, 1, 1]; μ2 = [3, 5, 5]; μ3 = [10, 5, 10]; μ4= [3, 10, 15]; μ5= [10, 10, 15]）和协方差矩阵（σ1 = [1 0 0; 0 1 0; 0 0 1]; σ2 = [2 0 0; 0 5 0; 0 0 2]; σ3 = [3 0 0; 0 3 0; 0 0 3]; σ4 = [2 0 0; 0 10 0; 0 0 4]; σ5 = [1 0 0; 0 5 0; 0 0 10]）的三维随机矢量，并绘制出它们的三维散点图，如图 2-3。

4）将 3）产生的五组各 1000 个具有不同均值矢量和协方差矩阵的三维随机矢量在二维平面上进行投影，画出子分量集合的二维散点图，如图 2-4（在 x-y 平面投影），图 2-5（在 x-z 平面投影）和图 2-6（在 y-z 平面投影）。

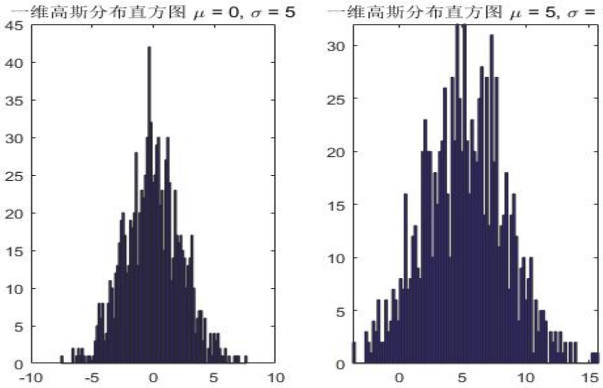

图 2-1

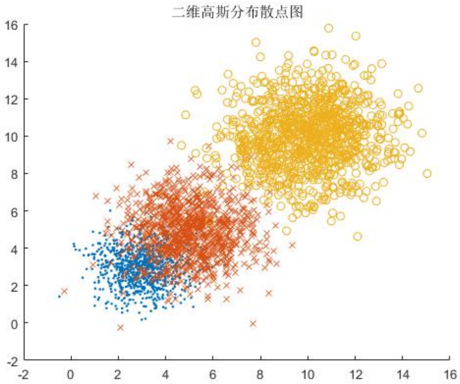

图 2-2

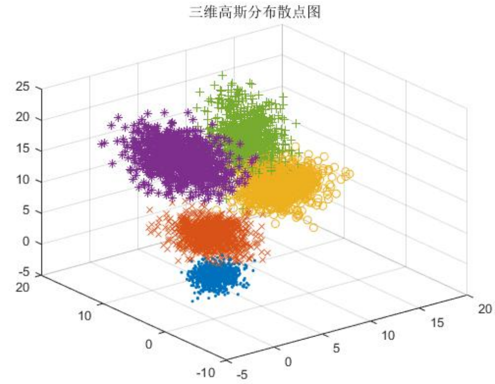

图 2-3

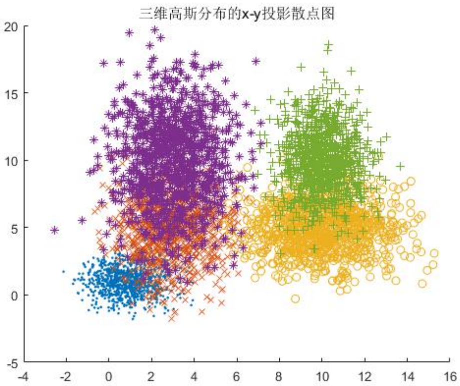

图 2-4

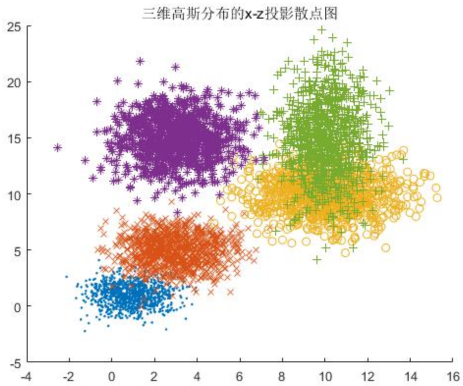

图 2-5

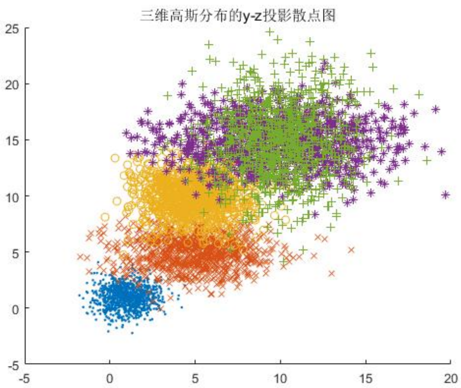

图 2-6

3. 实验 3 结果如下：

一个二维随机矢量（均值矢量μ = [0, 0]和协方差矩阵σ= [5 0; 0 5]）的概率密度函数值的三维曲面图，如图 3-1，纵坐标为概率密度函数值。

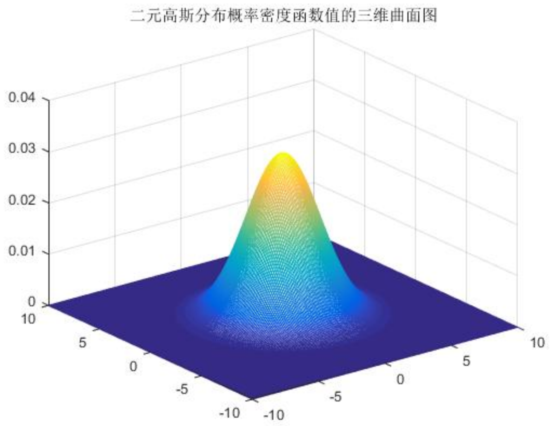

图 3-1

四、 讨论与分析

实验 1：使用函数 rand()能够生成不同数据长度和不同数据维度的随机数据。

如图 1-1，生成的一维随机数均匀分布在区间[10, 70]中；图 1-2，生成的二维随机数均匀分布在二维区间[1, 5]□ [20, 30]中；图 1-2，生成的三维随机数据均匀分布在三维区间[10, 50]□ [30, 60]□ [10, 15]中。

实验 2：使用函数 randn()能够生成服从高斯分布的一维高斯分布随机数的方法，以及使用函数 mvnrnd()能够生成二维、三维服从高斯分布的随机矢量。

如图 2-1，生成两组一维随机数服从高斯分布，由图可知，均值分别为 0 和 5，且第二组数据的方差比第一组大，与实验设置一致。图 2-2 是三组二维高斯随机矢量的二维散点图，由图可知它们具有不同的均值矢量和协方差矩阵。图 2-3 是五组三维高斯随机矢量的三维散点图，由图可知它们具有不同的均值矢量和协方差矩阵，而且它在不同的二维分量平面上的投影各不相同。

实验 3：使用函数 mvnpdf()可以求取多维随机向量的概率密度函数值，使用 meshgrid 能够绘制三维曲面图，使数据可视化。

图 3-1 概率密度函数值的三维曲面图服从高斯分布，由此可知，生成的二维随机矢量服从高斯分布，且由图可知均值为[0, 0]。

附录.

实验 1：

```
%%----------------Proj01-01：模式类实验数据的生成--------------%%
%%----------------Proj01-01-exp1----------------%%

clear; clc;

%%在一维区间[10,70]中，生成1000个均匀分布的随机数，然后统计并绘制这些数的直方图；
r = 10 + (70 - 10).* rand(1000, 1); % Generate values from the uniform distribution on the interval [a, b]. r = a + (b-a).*rand(100,1);
figure(1);
hist(r, 10: 70);
axis tight;
title('一维随机分布直方图');

%%在二维区间[1,5]*[20,30]中，生成5000个均匀分布的二维随机点，并绘制出它们的二维散点图；
x = 1 + (5 - 1).* rand(5000, 1);
y = 20 + (30 - 20).* rand(5000, 1);
figure(2);
scatter(x, y, '.', 'p');
title('二维随机分布散点图');

%%在三维区间[10,50]*[30,60]*[10,15]中，生成10000个均匀分布的三维随机点量，并绘制出它们的三维散点图。
x = 10 + (50 - 10).* rand(10000, 1);
y = 30 + (60 - 30).* rand(10000, 1);
z = 10 + (15 - 10).* rand(10000, 1);
figure(3);
scatter3(x, y, z, '.', 'p');
title('三维随机分布散点图');
```

实验 2：

```
%%----------------Proj01-01：模式类实验数据的生成--------------%%
%%----------------Proj01-01-exp2----------------%%

clear; clc;
n = 1000;
%%生成两组各1000个具有不同均值和方差的一维高斯分布的随机数，然后统计并绘制这些点的直方图；

u1 = 0; u2 = 5;%均值
sigma1 = 5;  sigma2 = 10;%方差
```

```
r1 = u1 + sqrt(sigma1) * randn(n, 1);
r2 = u2 + sqrt(sigma2) * randn(n, 1);
figure(1);
subplot(1, 2, 1); hist(r1, 100); title('一维高斯分布直方图 \mu = 0, \sigma = 5');
subplot(1, 2, 2); hist(r2, 100); title('一维高斯分布直方图 \mu = 5, \sigma = 10');
axis tight;

%%生成三组各 1000 个具有不同均值矢量和协方差矩阵的二维随机矢量，并绘制出它们的二维散点图；
m1 = [3, 3]; m2 = [5, 5]; m3 = [10, 10];%均值
a1 = [1 0; 0 1]; a2 = [2 0; 0 2]; a3 = [3 0; 0 3];%协方差矩阵
g1 = mvnrnd(m1, a1, n);
g2 = mvnrnd(m2, a2, n);
g3 = mvnrnd(m3, a3, n);
figure(2);
scatter(g1(:, 1), g1(:, 2), '.');
hold on; scatter(g2(:, 1), g2(:, 2), 'x');
scatter(g3(:, 1), g3(:, 2), 'o');
title('二维高斯分布散点图');

%%生成五组各 1000 个具有不同均值矢量和协方差矩阵的三维随机矢量，并绘制出它们的三维散点图。
k1 = [1, 1, 1]; k2 = [3, 5, 5]; k3 = [10, 5, 10]; k4= [3, 10, 15]; k5= [10, 10, 15];%均值
b1 = [1 0 0; 0 1 0; 0 0 1]; b2 = [2 0 0; 0 5 0; 0 0 2]; b3 = [3 0 0; 0 3 0; 0 0 3];
b4 = [2 0 0; 0 10 0; 0 0 4]; b5 = [1 0 0; 0 5 0; 0 0 10];%协方差矩阵
h1 = mvnrnd(k1, b1, n);
h2 = mvnrnd(k2, b2, n);
h3 = mvnrnd(k3, b3, n);
h4 = mvnrnd(k4, b4, n);
h5 = mvnrnd(k5, b5, n);
figure(3);
scatter3(h1(:, 1), h1(:, 2), h1(:, 3), '.');
hold on; scatter3(h2(:, 1), h2(:, 2), h2(:, 3), 'x');
scatter3(h3(:, 1), h3(:, 2), h3(:, 3), 'o');
scatter3(h4(:, 1), h4(:, 2), h4(:, 3), '*');
scatter3(h5(:, 1), h5(:, 2), h5(:, 3), '+');
title('三维高斯分布散点图');

%%进一步，绘制上述三维随机矢量数据集合的二维投影散点图。
figure(4);%三维随机矢量数据集合在第 1 和第 2 两个维度的二维投影散点图
```

```
scatter(h1(:, 1), h1(:, 2), '.');
hold on; scatter(h2(:, 1), h2(:, 2), 'x');
scatter(h3(:, 1), h3(:, 2), 'o');
scatter(h4(:, 1), h4(:, 2), '*');
scatter(h5(:, 1), h5(:, 2), '+');
title('三维高斯分布的 x-y 投影散点图');

figure(5);%三维随机矢量数据集合在第 1 和第 3 两个维度的二维投影散点图
scatter(h1(:, 1), h1(:, 3), '.');
hold on; scatter(h2(:, 1), h2(:, 3), 'x');
scatter(h3(:, 1), h3(:, 3), 'o');
scatter(h4(:, 1), h4(:, 3), '*');
scatter(h5(:, 1), h5(:, 3), '+');
title('三维高斯分布的 x-z 投影散点图');

figure(6);%三维随机矢量数据集合在第 2 和第 3 两个维度的二维投影散点图
scatter(h1(:, 2), h1(:, 3), '.');
hold on; scatter(h2(:, 2), h2(:, 3), 'x');
scatter(h3(:, 2), h3(:, 3), 'o');
scatter(h4(:, 2), h4(:, 3), '*');
scatter(h5(:, 2), h5(:, 3), '+');
title('三维高斯分布的 y-z 投影散点图');
```

实验 3：

```
%%----------------Proj01-01：模式类实验数据的生成--------------%%
%%----------------Proj01-01-exp3----------------%%

clc; clear;
%%确定一个二维的均值矢量和协方差矩阵，然后利用 matlab 中的 meshgrid 函数生成一个二维网格;
u = [0, 0];%均值
sigma = [5 0; 0 5];%协方差矩阵
[X, Y]  = meshgrid(-10: 0.1: 10, -10: 0.1: 10);

%%利用 mvnpdf 函数计算在每个网格点上的概率密度函数值，并绘制出这些函数值的三维曲面图。
p = mvnpdf([X(:) Y(:)], u, sigma);
p = reshape(p, size(X));%将行向量转换为矩阵向量
mesh(X, Y, p);
title('二元高斯分布概率密度函数值的三维曲面图');
```