3. Part I

a. μ and Σ from the first 10 data samples:

$$\mu = \begin{bmatrix} 0.8190 & -0.6271 \end{bmatrix}^T, \quad \Sigma = \begin{bmatrix} 0.7461 & -0.1474 \\ -0.1474 & 1.6047 \end{bmatrix}$$

b. μ and Σ from the first 100 data samples:

$$\mu = \begin{bmatrix} 0.9977 & -0.9725 \end{bmatrix}^T, \quad \Sigma = \begin{bmatrix} 2.2580 & 1.0856 \\ 1.0856 & 2.1439 \end{bmatrix}$$

c. μ and Σ from the first 1000 data samples:

$$\mu = \begin{bmatrix} 1.0222 & -0.9629 \end{bmatrix}^T, \quad \Sigma = \begin{bmatrix} 2.2118 & 1.1878 \\ 1.1878 & 2.0332 \end{bmatrix}$$

c. μ and Σ from the first 10000 data samples:

$$\mu = \begin{bmatrix} 0.9947 & -1.0027 \end{bmatrix}^T, \quad \Sigma = \begin{bmatrix} 1.9978 & 0.9643 \\ 0.9643 & 1.9237 \end{bmatrix}$$

e. Parameter estimation errors

Measure 1:

case	a	b	c	d
ε	2.0921	0.2863	0.3424	0.0509

Measure 2:

case	a	b	c	d
ε_μ	0.2931	0.0195	0.0306	0.0042
ε_Σ	1.1840	0.1645	0.1961	0.0292

In both *Measure 1* and *Measure 2* we notice that parameter estimation errors decrease as the number of data samples increase. Maximum likelihood estimation assumes that the parameter $\boldsymbol{\theta}$ is fixed then seeks to find the parameter value to maximize the probability of the training data being observed.

e. Plot of first 100 data samples and 2D contours of estimated Gaussian pdf

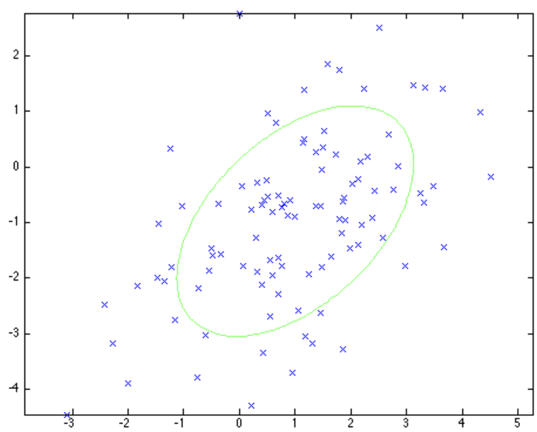

Figure 1: 100 data samples and estimated Gaussian pdf 2D contours

Part II

a. μ and Σ from the first 10 data samples:

$$\mu = \begin{bmatrix}1.8829 & -1.8135\end{bmatrix}^T, \quad \Sigma = \begin{bmatrix}5.6385 & -5.3104\\ -5.3104 & 5.3521\end{bmatrix}$$

b. μ and Σ from the first 100 data samples:

$$\mu = \begin{bmatrix}1.1741 & -1.2216\end{bmatrix}^T, \quad \Sigma = \begin{bmatrix}2.6753 & -2.5961\\ -2.5961 & 2.6913\end{bmatrix}$$

c. μ and Σ from the first 1000 data samples:

$$\mu = \begin{bmatrix}0.9539 & -0.9530\end{bmatrix}^T, \quad \Sigma = \begin{bmatrix}1.9939 & -1.9344\\ -1.9344 & 2.0528\end{bmatrix}$$

c. μ and Σ from the first 10000 data samples:

$$\mu = \begin{bmatrix}1.0023 & -1.0031\end{bmatrix}^T, \quad \Sigma = \begin{bmatrix}1.9659 & -1.8639\\ -1.8639 & 1.9582\end{bmatrix}$$

e. Parameter estimation errors

Measure 1:

case	a	b	c	d
ε	6.1596	1.2266	0.0126	0.0616

Measure 2:

case	a	b	c	d
ε_μ	0.8489	0.1993	0.0056	0.0028
ε_Σ	3.4881	0.6893	0.0057	0.0355

In both *Measure 1* and *Measure 2* we notice that parameter estimation errors decrease as the number of data samples increase. Maximum likelihood estimation assumes that the parameter $\boldsymbol{\theta}$ is fixed then seeks to find the parameter value to maximize the probability of the training data being observed.

e. Plot of first 100 data samples and 2D contours of estimated Gaussian pdf

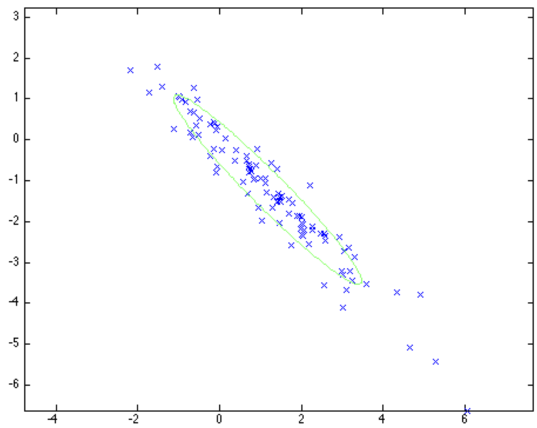

Figure 2: 100 data samples and estimated Gaussian pdf 2D contours

Appendix:

assignment_3.m

```
%
% CS7720 Spring 2015
% Introduction to Machine Learning and Pattern Recognition
% University of Missouri-Columbia
%
% Author: Chanmann Lim
% email: cl9p8@mail.missouri.edu
%
% Homework Assignment 3
% Problem 4
%
clc; clear; close all;

%%
% Problem 3. Part I
%
%   dataset - GDdataMLE1 dataset
%   m - true mean
%   P - true covariance

dataset = load('GDdataMLE1.txt');
m = [1; -1];
P = [2 1; 1 2];

problem_3_report;

%%
% Problem 3. Part II
%
%   dataset - GDdataMLE2 dataset
%   m - true mean
%   P - true covariance

dataset = load('GDdataMLE2.txt');
m = [1; -1];
P = [2 -1.9; -1.9 2];

problem_3_report;
```