Ex 8.1

In `BayesianPriorElicitation.m`, the whitening process is completed by choosing `xn` from the standard normal distribution (line 27) and then taking `x = gamma * (L'\xn)` (line 30). This is equivalent to the step mentioned in the lecture notes, where $m_i = \overline{m} + R Z_i$.

Ex 8.2

Rejection Sampling ($a > 1$):

The results for the rejection sampling are shown below with the scaled target density. It can be seen that the shape of the histogram resembles the exact density shape.
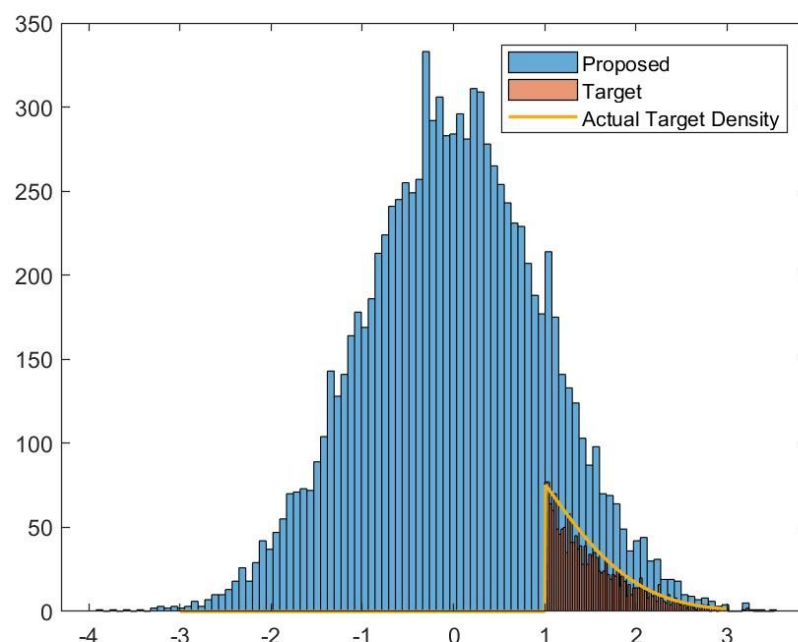
If the value of 'a' keeps on increasing, fewer samples are accepted as shown below. The target samples then do not truly follow the target distribution.
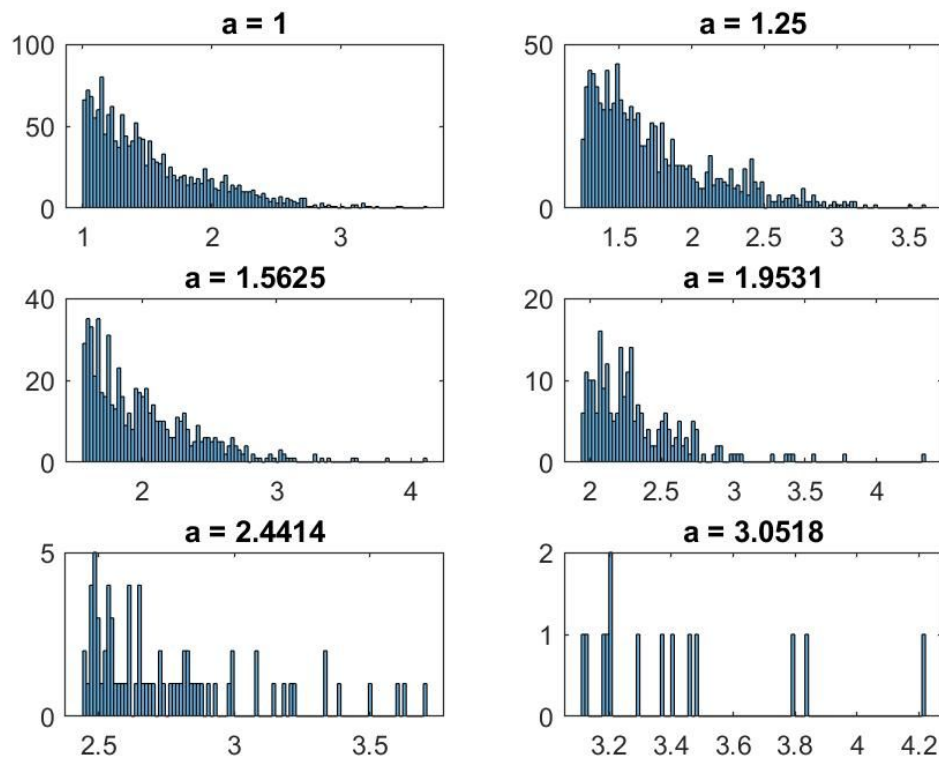
Code:

```
% target density, p(m)  = g(m)/C*exp(-m^2/2)
% proposal density, q(m)  = 1/sqrt(2*pi)*exp(-m^2/2)

clear all
close all

%% Generate Samples

n = 10000;
proposedSamples = randn(n,1);

a = 1;
g = double(proposedSamples>a);

%% Accept/Reject

alpha = g;
acceptor = rand(n,1)>1-alpha;
targetSamples = proposedSamples(acceptor);

%% Plots

figure(1)
m = -3:0.01:3;
g = double(m>a);
target = g.*exp(-m.^2/2);
histogram(proposedSamples,100)
hold on
histogram(targetSamples,100)
plot(m,50*sqrt(2*pi)*target,'LineWidth',1.5)
legend('Proposed','Target','Actual Target Density')
```

```
%% Increase 'a'

a = 1;
figure(2)
for i = 1:6
    proposedSamples = randn(n,1);
    g = double(proposedSamples>a);
    alpha = g;
    acceptor = rand(n,1)>1-alpha;
    targetSamples = proposedSamples(acceptor);
    subplot(3,2,i)
    histogram(targetSamples,100)
    title(['a = ' num2str(a)])
    a = 1.25*a;
end
```

Ex 8.3

We empirically find the value of D which suits the problem. If we take the value of $D = 1.7\sqrt{2\pi}C$, the contour map for the value of $(Dq(m) - C\pi(m))/C$ is shown below. It can be seen that this is positive everywhere, which implies that assumption (8.5) holds.
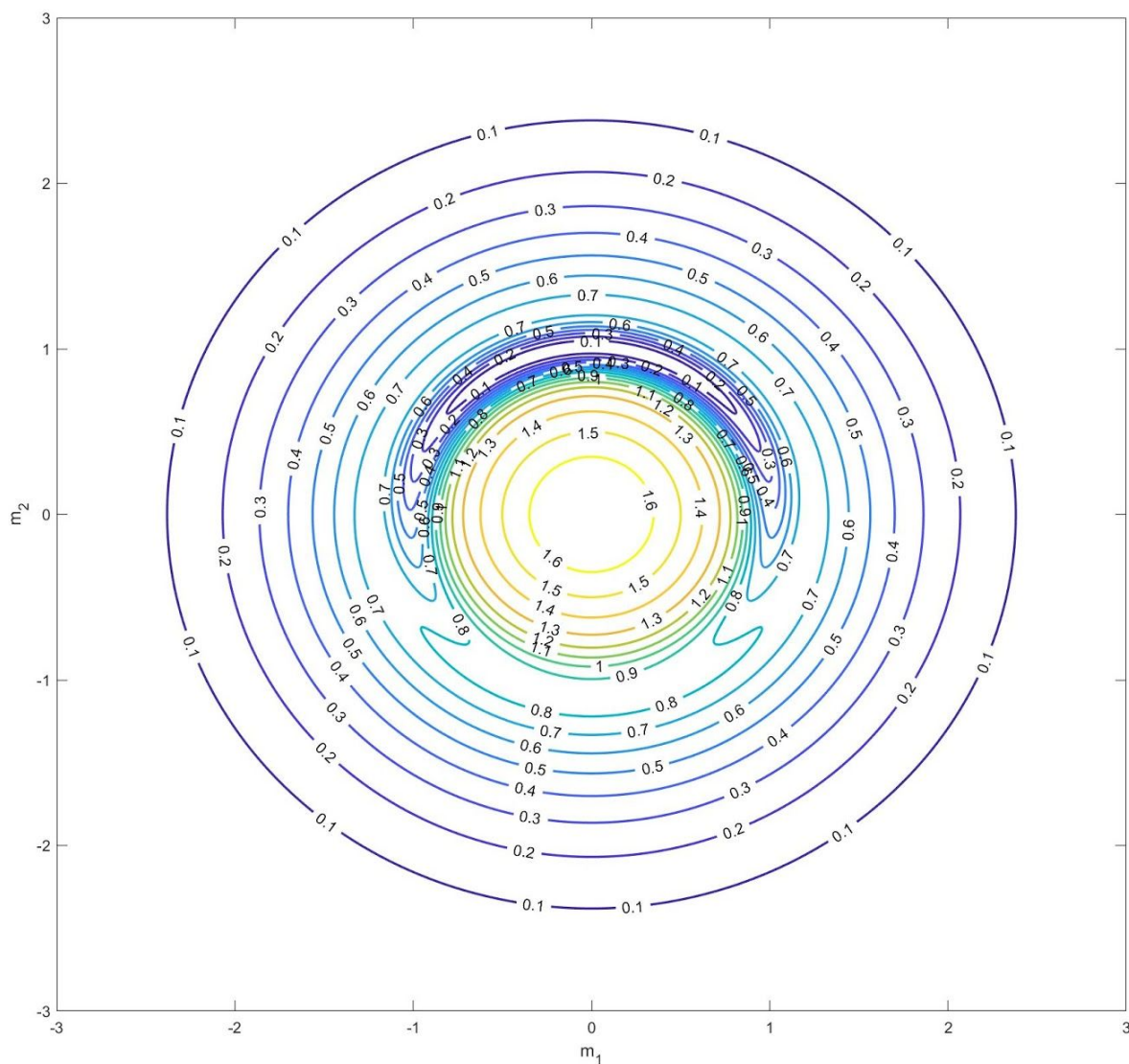
The sampling distribution (horseshoe) and the sampled points are shown below. It can be seen that the points follow the true distribution.
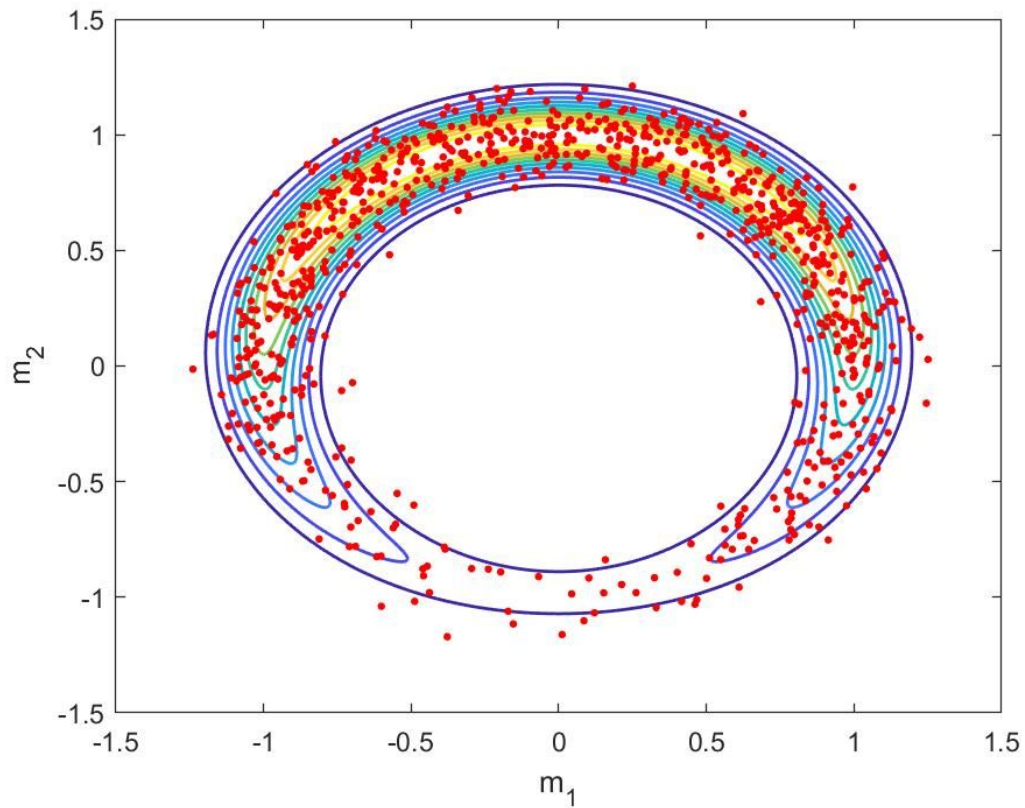
Code:

```
clear all

%% Contours

m1 = linspace(-3,3,1000);
m2 = linspace(-3,3,1000);
[M1,M2] = meshgrid(m1,m2);
sigma = 0.1;
delta = 1;
Z = exp(-1/(2*sigma^2)*((M1.^2 + M2.^2).^0.5 - 1).^2 - 1/(2*delta^2)*(M2 - 1).^2);
ZBar = 1.7*exp(-1/(2)*(M1.^2 + M2.^2));
ZTilde = ZBar - Z;
figure(1)
contour(M1,M2,ZTilde,[-0.001,0:0.1:3],'LineWidth',1.5,'ShowText','on');
xlabel('m_1')
ylabel('m_2')

%% Generate Samples

n = 10000;
m1Proposed = randn(n,1);
m2Proposed = randn(n,1);

%% Accept/Reject

alpha    =    exp(-1/(2*sigma^2)*((m1Proposed.^2    +    m2Proposed.^2).^0.5    -    1).^2    -
1/(2*delta^2)*(m2Proposed - 1).^2)./(1.7*exp(-1/(2)*(m1Proposed.^2 + m2Proposed.^2)));
acceptor = rand(n,1)>1-alpha;
m1Target = m1Proposed(acceptor);
m2Target = m2Proposed(acceptor);

figure(2)
contour(M1,M2,Z,10,'LineWidth',1.2);
hold on
```
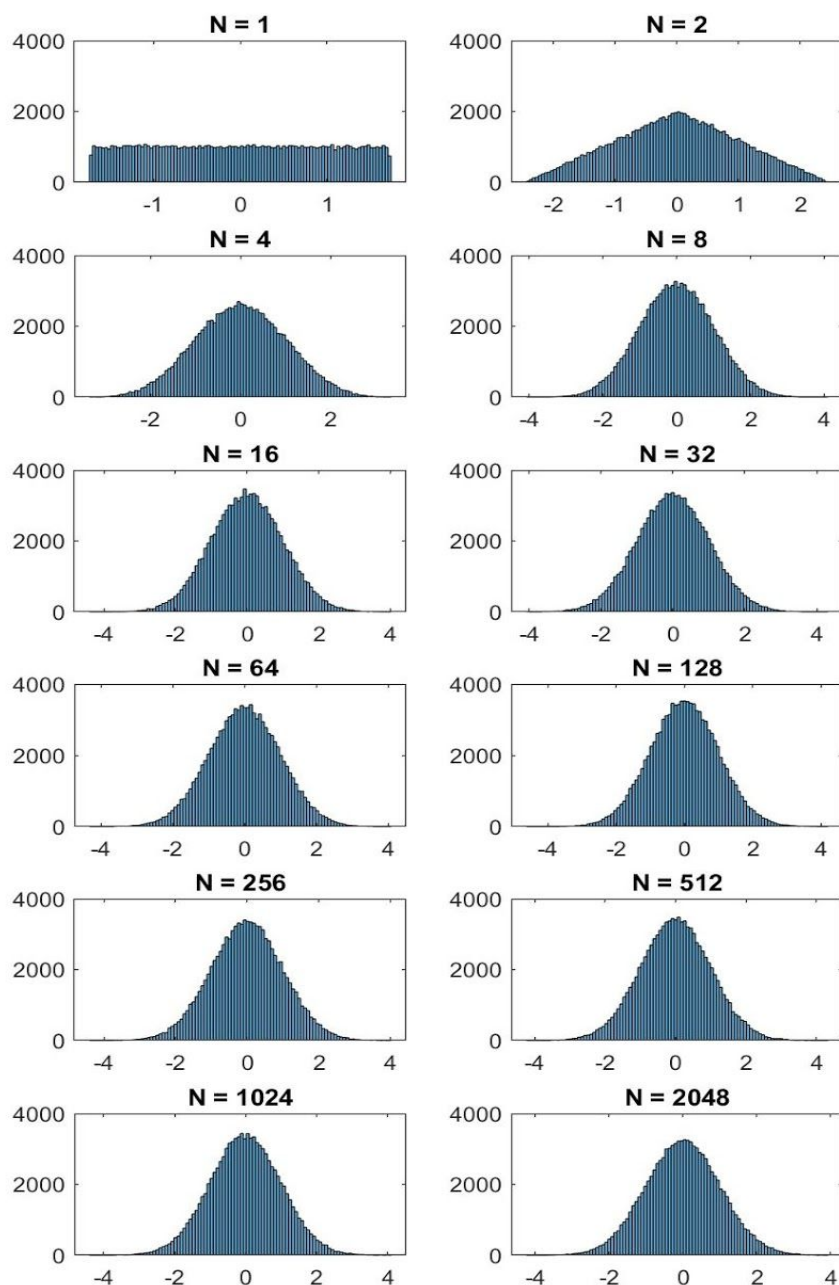
```
scatter(m1Target,m2Target,8,'r','filled');
xlabel('m_1')
ylabel('m_2')
xlim([-1.5,1.5])
ylim([-1.5,1.5])
```

Ex 9.1

The histogram of Z_N is shown with increasing N.

It can be seen that the distribution of Z_N converges to the standard normal distribution. The estimated mean and standard deviation of Z_N for the 12 cases shown above are shown below. They converge to 0 and 1 respectively.

```
Mean, Z  =  -0.0036128       0.0042228       0.0012264  -0.00042016     -0.0031531       0.0011771  -5.9222e-05
-0.00069034 -0.00094028    2.6691e-05   -0.0036722      0.0048895
Sigma,  Z  =  0.99877         0.99914         0.99758         0.99791         0.99917         1.0035          0.99754
0.99838       0.99989         0.9968          0.9984          1.0045
```

The value of S_N with increasing N is also shown below. It is seen that the value converges to 0.5.
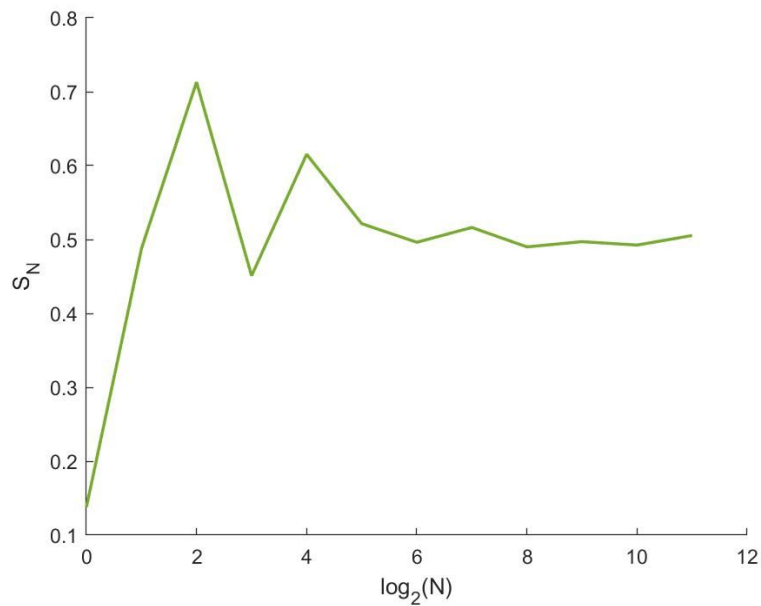
Code:

```
clear all
close all

N = 1;
samples = 100000;
mBar = 0.5;
sigma = sqrt(1/12);

%% Z Histogram

figure(1);
for i = 1:12
    subplot(6,2,i)
    m = rand(N,samples);
    Z = 1/(sigma*sqrt(N))*sum(m,1) - mBar/sigma*sqrt(N);
    histogram(Z,100)
    ylim([0 4000])
    title(['N = ' num2str(N)])
    ZMean(i) = mean(Z);
    ZStd(i) = std(Z);
    SN(i) = mean(m(:,1));
    num(i) = N;
    N = N*2;
end

%% Convergence of SN

figure(2)
hold on
for i = 1:5
    plot(log2(num),SN,'LineWidth',1.5)
end
xlabel('log_2(N)')
ylabel('S_N')

%% Display Mean, Std. Dev. of Z
```

```
disp(['Mean, Z = ' num2str(ZMean)])
disp(['Sigma, Z = ' num2str(ZStd)])
```

Ex 9.2

Ex. 9.2

$$\mathbb{E}\left[(Z_N - 0)^2\right] \underset{\text{def.}}{=} \mathbb{E}\left[\left(\frac{1}{\sigma\sqrt{N}}\sum m_i - \frac{\bar{m}\sqrt{N}}{\sigma}\right)^2\right]$$

$$Z_N$$

$$\|$$

$$\mathbb{E}\left[\frac{N}{\sigma^2}\left(\frac{\sum m_i}{N} - \bar{m}\right)^2\right]$$

$\|$ def. S_N

$$\frac{N}{\sigma^2}\,\mathbb{E}\left[(S_N - \bar{m})^2\right]$$

$\|$ $L_2(\mathbb{P}) := \left\{F(w): \int F(w)^2 \, d\mathbb{P} < \infty\right\}$

$$\frac{N}{\sigma^2}\,\|S_N - \bar{m}\|^2_{L^2(\mathbb{P})}$$

$$\Downarrow$$

$$\frac{N}{\sigma^2}\,\|S_N - \bar{m}\|^2_{L^2(\mathbb{P})} \simeq 1$$

$$\|S_N - \bar{m}\|_{L^2(\mathbb{P})} \simeq \frac{\sigma}{\sqrt{N}}$$

or $$\|S_N - \bar{m}\|^2_{L^2(\mathbb{P})} \simeq \frac{\sigma^2}{N}$$

9.3

Suppose the density of m is $\pi(m)$ and $z = g(m)$
is differentially invertible.
What is the density of $g(m)$?

Let $G(z)$ be the CDF of $g(m)$.

$$= P(Z \le z)$$
$$= P(g(m) \le z)$$
$$= P(m \le g^{-1}(z)) = \int_{-\infty}^{g^{-1}(z)} \pi(m)\, dm$$

So the density of $g(m)$ is $\dfrac{d\, G(z)}{dz}$

chain rule:
+ fundamental theorem
of calculus.

$$= \pi(g^{-1}(z)) \underbrace{\frac{dg^{-1}(z)}{dz}}_{} \, dz$$

$$\pi_z(z)$$

9.4 Define $z = g(m) \in T$. $\quad E[G(m)] = \int_T z\, \pi_z(z)\, dz$

Derive $E[G(m)] = \int_S g(m)\pi(m)\, dm$

$g(m)$ diff invertible $\Rightarrow$ $g(m)$ ~~???~~ bijective.

$\forall z \in T \quad g(m) = z \quad \Rightarrow \quad$ ~~????~~ $g(S) = T$
$$g^{-1}(T) = S.$$

$$\int_T z\, \pi_z(z)\, dz \qquad \text{take } z = g(m) \qquad m = g^{-1}(z)$$
$$dz = \frac{dg(m)}{dm}\, dm \qquad dm = \frac{dg^{-1}(z)}{dz}\, dz$$

$$= \int_{g^{-1}(T)} g(m)\, \pi(g^{-1}(g(m)))\, \frac{dg^{-1}(g(m))}{dz} \cdot \frac{dm}{dz}\, dz$$

$$= \int_S g(m)\, \pi(m)\, dm$$

9.5

9.5 Derive an unbiased estimator for Γ.

1. Show $E[(\hat{m}-\bar{m})(\hat{m}-\bar{m})^T] = \frac{\Gamma}{N}$

2. Start from $J = \sum_{i=1}^{N}(m_i - \hat{m})(m_i - \hat{m})^T$, show $E[J] = (N-1)\Gamma$

$$J = \sum_{i=1}^{N}(m_i - \hat{m})(m_i - \hat{m})^T = \sum (m_i - \bar{m} + \bar{m} - \hat{m})(m_i - \bar{m} + \bar{m} - \hat{m})^T$$

$$= \sum (m_i - \bar{m})(m_i - \bar{m})^T + \sum(\bar{m} - \hat{m})(\bar{m} - \hat{m})^T + \sum (\bar{m} - \hat{m})(m_i - \bar{m})^T + \sum(\bar{m} - \hat{m})$$

expectation over each sum.

$$= N\Gamma$$

$\left(\hat{m} = \sum m_i / N \right)$

$N E[(\hat{m} - \bar{m})(\hat{m} - \bar{m})^T]$

$N\left(\frac{-\Gamma}{N} \right) = -\Gamma$ (kwingely) $-\Gamma$ (by symmetry)

$+ N\Gamma$

$\frac{}{N}$

$$= N\Gamma - \Gamma - \Gamma + \Gamma = (N-1)\Gamma$$

Then $E\left[\dfrac{J}{N-1} \right] = \Gamma = \dfrac{1}{N-1}\sum_{i=1}^{N}(m_i - \hat{m})(m_i - \hat{m})^T$.

Now all that remains to be shown is (1).

$$(\hat{m} - \bar{m})(\hat{m} - \bar{m})^T = \left(\frac{1}{N}\sum_i m_i - \bar{m} \right)\left(\frac{1}{N}\sum_i m_i - \bar{m} \right)^T$$

$$= \frac{1}{N^2}\sum_i\sum_i (m_i - \bar{m})(m_i - \bar{m})^T$$

$$= \frac{N\Gamma}{N^2} = \frac{\Gamma}{N}.$$

10.1

- Use Algorithm 1 to reproduce figure 10.2
- Compare N=1500 with the stationary distribution for different starting positions

Code:

```
P = [
      0, 0.5, 0, 0.5, 0;
      1, 0, 0, 0, 0;
      0, 0, 0, 0.5, 0.5;
      1/3, 1/3, 0, 0, 1/3;
      0, 0, 1/2, 1/2, 0
      ];
nTrials = 1500;
n = size(P, 1);
```

```matlab
m = zeros(nTrials, 1);
for start=1:n
    x = rand(nTrials, 1);
    m(1) = linecheck(x(1), P(start, :));
    for trials=2:nTrials
        m(trials) = linecheck(x(trials), P(m(trials-1), :));
    end
    sv = zeros(1, n);
    sv(start) = 1;
    pi_inf = sv*P^1500;
    [a,d]=hist(m,unique(m));
    pi_n = zeros(1, n);
    for idx=d
        pi_n(d) = a(d)/nTrials;
    end
    figure()
    bar(1:n, [pi_inf', pi_n'])
    stitle = sprintf('Starting at State %d', start);
    title(stitle)
    legend('Invariant Dist', 'Frequency')
    xlabel('State')
    ylabel('Frequency')
end
bal = 1;
for i=1:size(P, 1)
    for j=1:size(P, 2)
        if(norm(pi_inf(i)*P(i, j) - pi_inf(j)*P(j, i)) > 1e-3)
            bal = 0
            break
        end
    end
    if(~bal)
        break
    end
end
if bal
    fprintf("This satisfies detailed balance\n")
else
    fprintf("This doesn't sastify detailed balance\n")
end


function m = linecheck(x, v)
    b = 0;
    for idx=1:length(v)
        b = b + v(idx);
        if(x < b)
            break;
        end
    end
    m = idx;
end
```
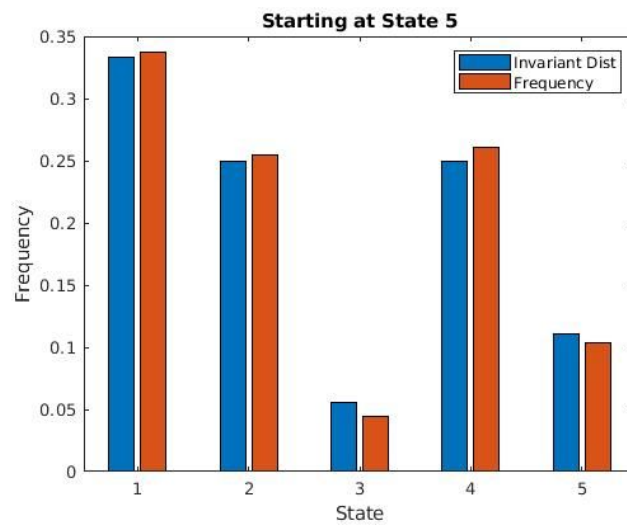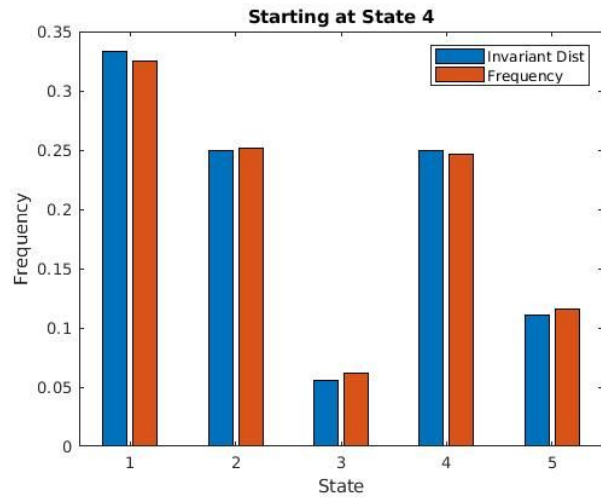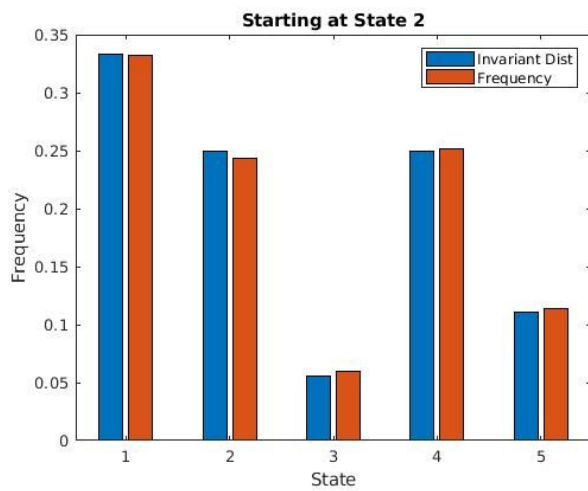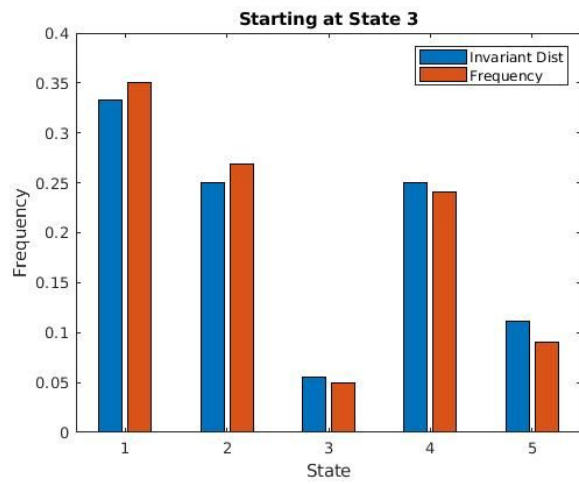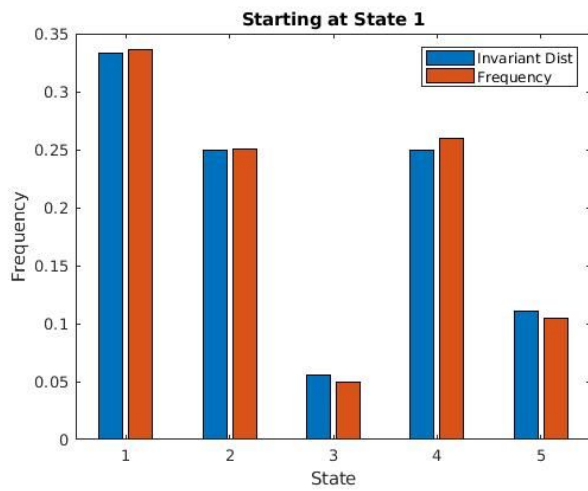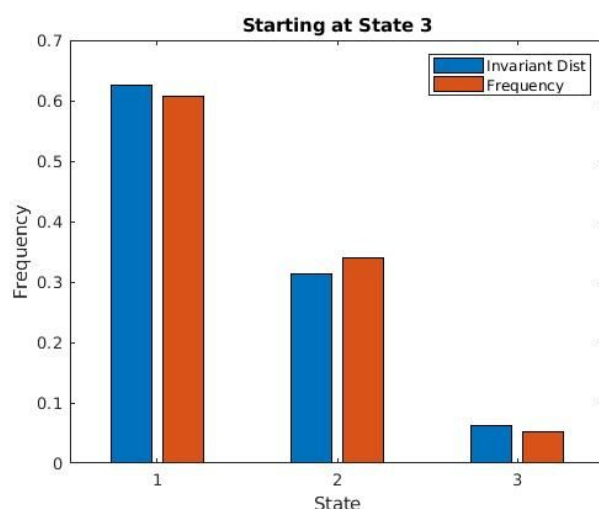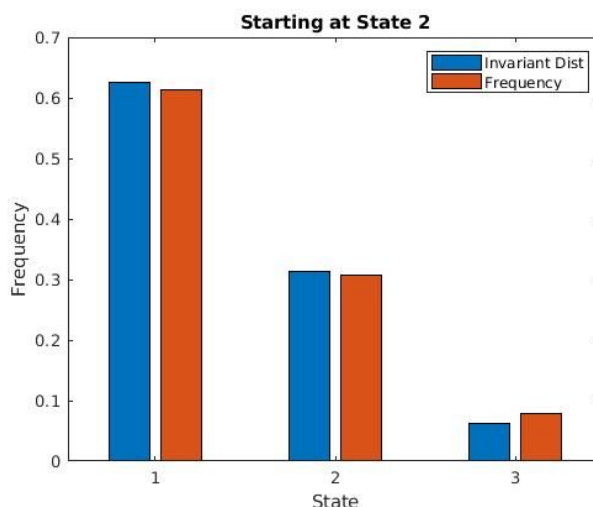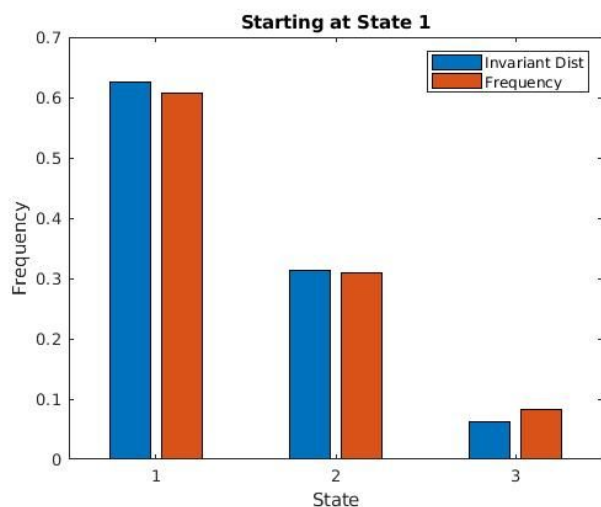
10.2

- **Same as 10.1 but with a different system given by:**

```
P = [
    0.9, 0.075, 0.025;
    0.15, 0.8, 0.05;
    0.25, 0.25, 0.5];
```

Starting at State 1

Starting at State 2

Starting at State 3

The system has a 6.25% chance of being in the recession (state 3) at the stationary distribution.

10.3

- What is the discrete version of detailed balance?

$\pi_i P_{i,j} = \pi_j P_{j,i}$

- Is 10.1 reversible? No.
- Is 1.2 reversible? Yes.

Checks for reversibility are in the code above.

10.4

- What is the probability of staying at m_k ?

In Metropolis-Hastings this is the same as the acceptance probability, given a proposal point p we stay at m_k with probability:

$$1 - \alpha(m_k, p) = 1 - \min\left\{\frac{\pi(p)q(p, m_k)}{\pi(m_k)q(m_k, p)}, 1\right\}$$

Since p depends on m_k we cannot simply this more without knowing the explicit form of the proposal density function.

The probability of staying at m_k then is this probability marginalized over the space of all proposal samples p.

$$1 - \int_p \alpha(m_k, p) \, q(p \mid m_k)$$
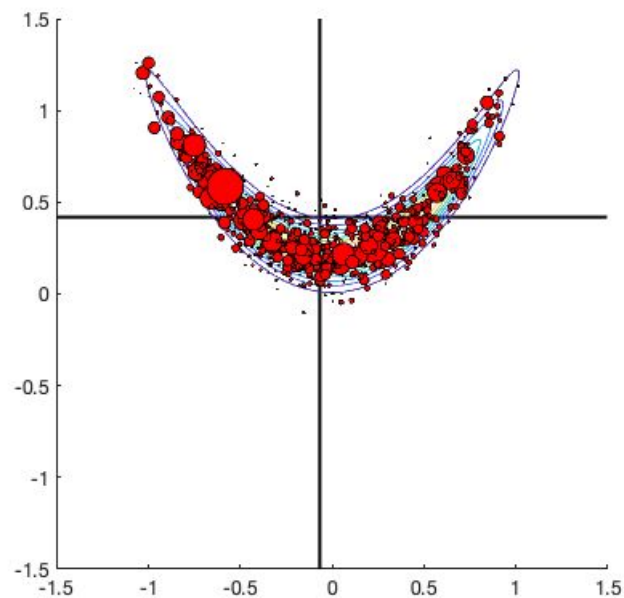
Ex 11.1

(a) .

Change the corresponding part of BayesianMCMC.m to
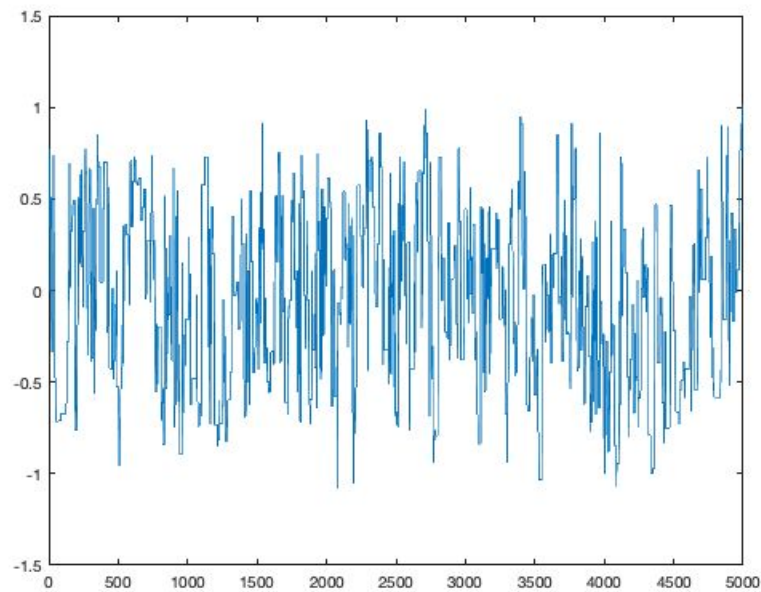
```
% number of samples

N = 5000;


% target density

sigma = 0.1; delta = 1; y = [-0.2, 0.1];

h = @(m) [m(1)^2-m(2), m(2)/5];

Pi = @(m) exp(-0.5/sigma^2*sum((y-h(m)).^2) - 0.5/delta^2* ...

            sum(m.^2));


% proposal variance, and proposal density

gamma = 0.5;

q = @(m,p) 1./sqrt(2*pi*gamma^2)*exp(-0.5/gamma^2*sum((m-p).^2));
```
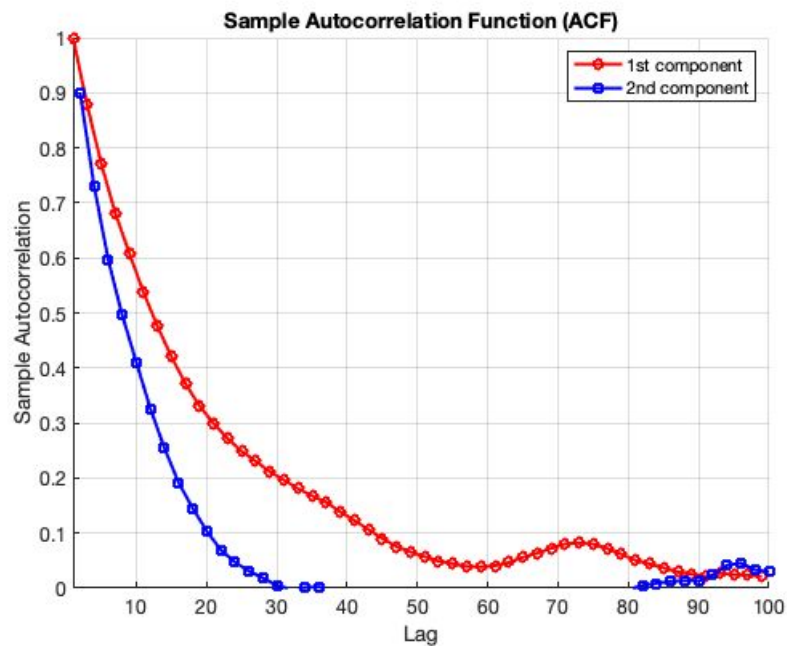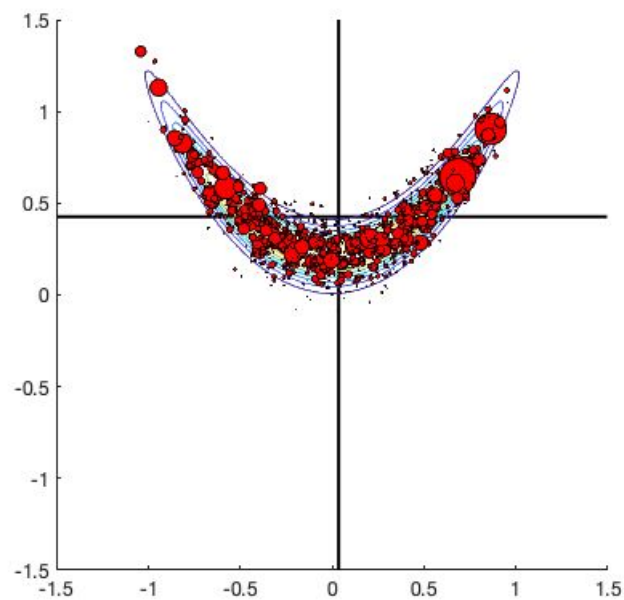

(a) scatter plot

(b) *Trace plots of the first component of m.*

(a) *Autocorrelation function plot for both components of m.*

Figure 1. Plots for Ex. 11.1 - 1.

(b) The γ should be 0.364 so that the corresponding to average acceptance probability is about 0.234.

Change the corresponding part of BayesianMCMC.m to

```
% number of samples

N = 5000;


% target density

sigma = 0.1; delta = 1; y = [-0.2, 0.1];
```

```
h = @(m) [m(1)^2-m(2), m(2)/5];

Pi = @(m) exp(-0.5/sigma^2*sum((y-h(m)).^2) - 0.5/delta^2* ...

                sum(m.^2));


% proposal variance, and proposal density

gamma = 0.364;

q = @(m,p) 1./sqrt(2*pi*gamma^2)*exp(-0.5/gamma^2*sum((m-p).^2));
```
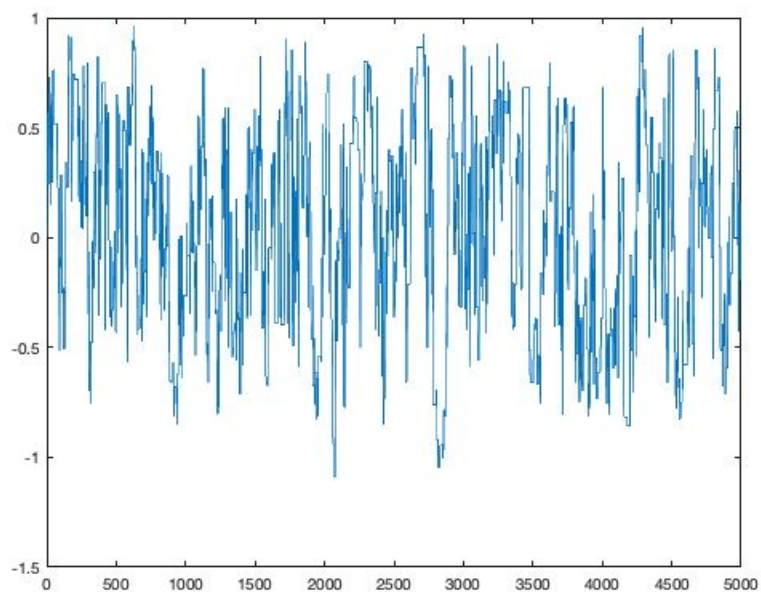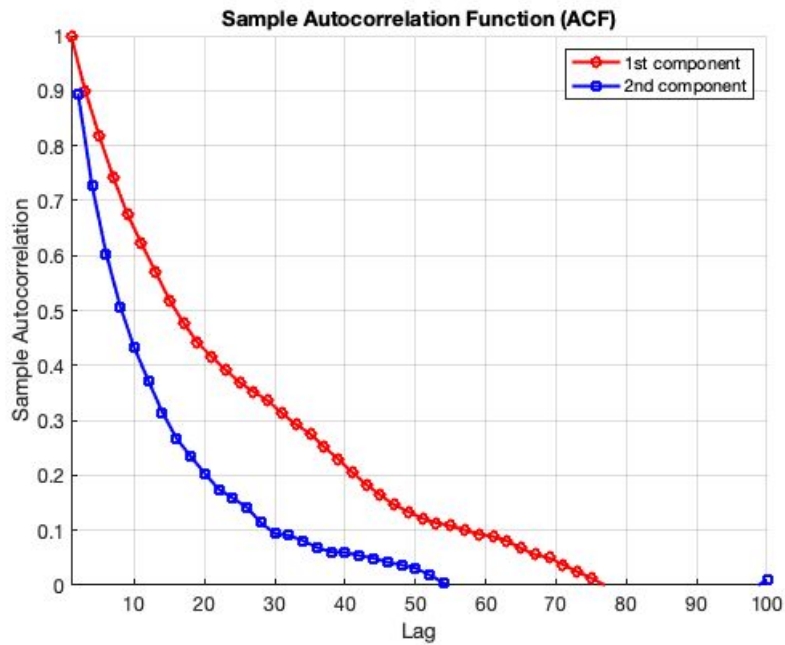

(a) scatter plot

(b) *Trace plots of the first component of m.*

(a) *Autocorrelation function plot for both components of m.*

Figure 1. Plots for Ex. 11.1 - 2.
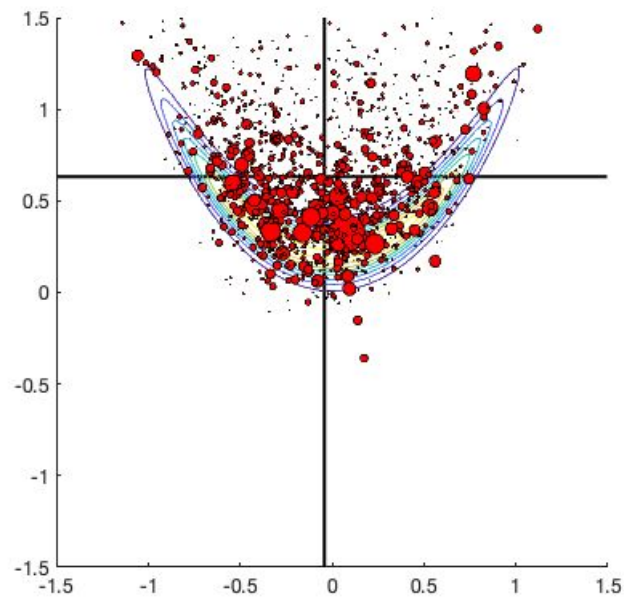
Ex 11.2

(1) $H(m_k) = (\nabla^T h(m_k)) \, \nabla h(m_k)$

(2)

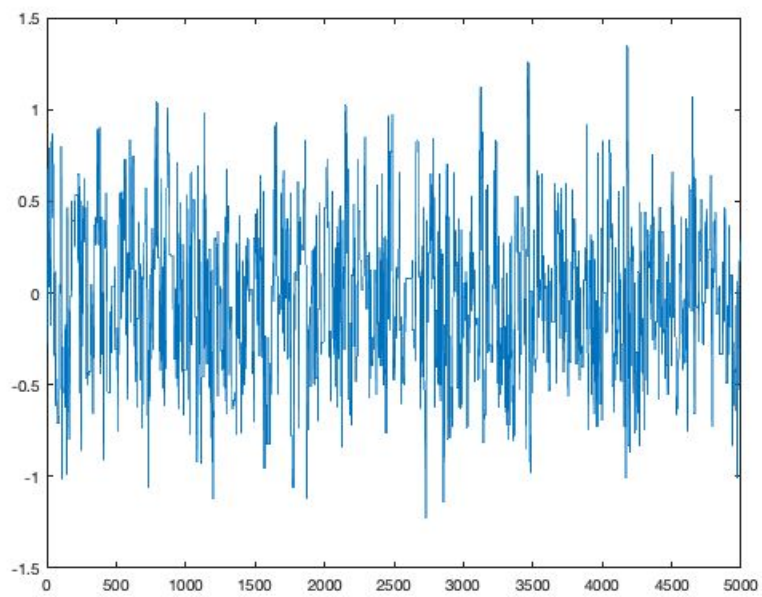Change the corresponding part of BayesianMCMC.m to

```
% number of samples

N = 5000;


% target density

sigma = 0.1; delta = 1; y = [-0.2, 0.1];

h = @(m) [m(1)^2-m(2), m(2)/5];

Pi = @(m) exp(-0.5/sigma^2*sum((y-h(m)).^2) - 0.5/delta^2* ...

            sum(m.^2));


% proposal variance, and proposal density

gamma = 0.5;

gh = @(m) [2*m(1) -1; 0 0.2];

q = @(m,p) 1./sqrt(2*pi*sigma^2)*exp(-0.5/sigma^2*sum((gh(m)*(p-m)).^2));
```
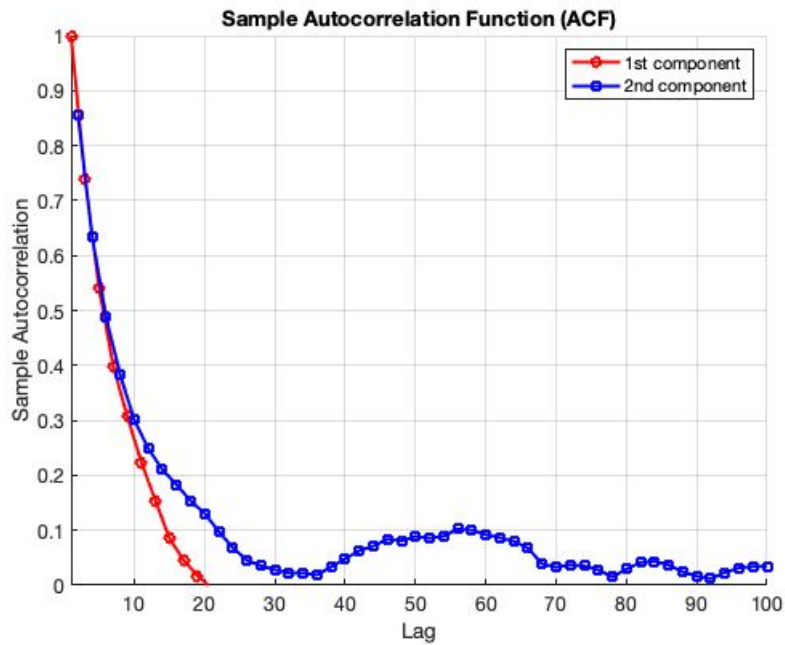
(a) scatter plot

(b) *Trace plots of the first component of m.*

(a) *Autocorrelation function plot for both components of m.*

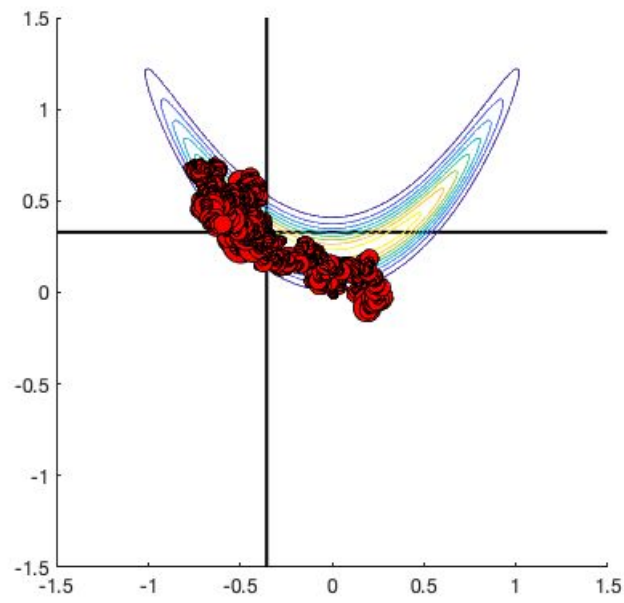Figure 3. Plots for Ex. 11.2.

Ex 11.3

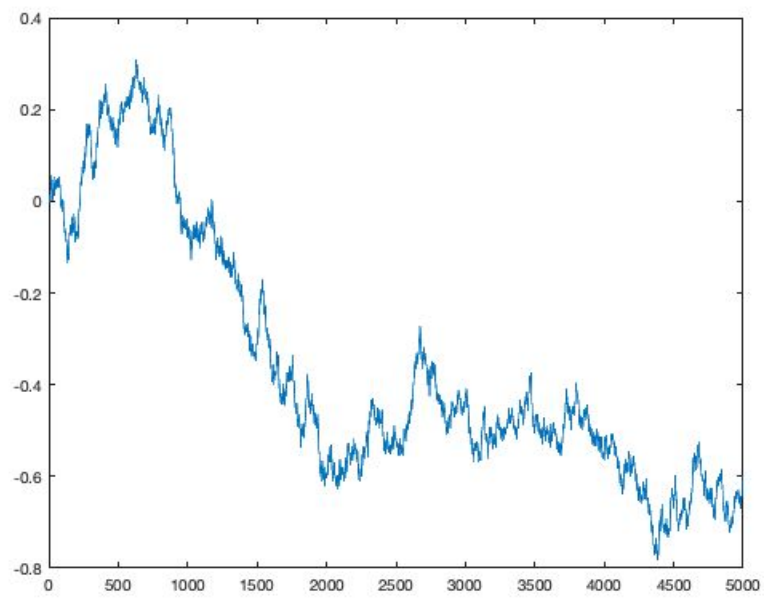Change the corresponding part of BayesianMCMC.m to

```
% number of samples

N = 5000;

% target density

sigma = 0.1; delta = 1; y = [-0.2, 0.1];

h = @(m) [m(1)^2-m(2), m(2)/5];

Pi = @(m) exp(-0.5/sigma^2*sum((y-h(m)).^2) - 0.5/delta^2* ...

             sum(m.^2));


% proposal variance, and proposal density

gamma = 0.0087;

C = [0.1896 0.0095; 0.0095 0.0463];

q = @(m,p) 1./sqrt(4*pi^2*gamma)*exp(-0.5*(p-m)'*C^(-1)*(p-m));
```
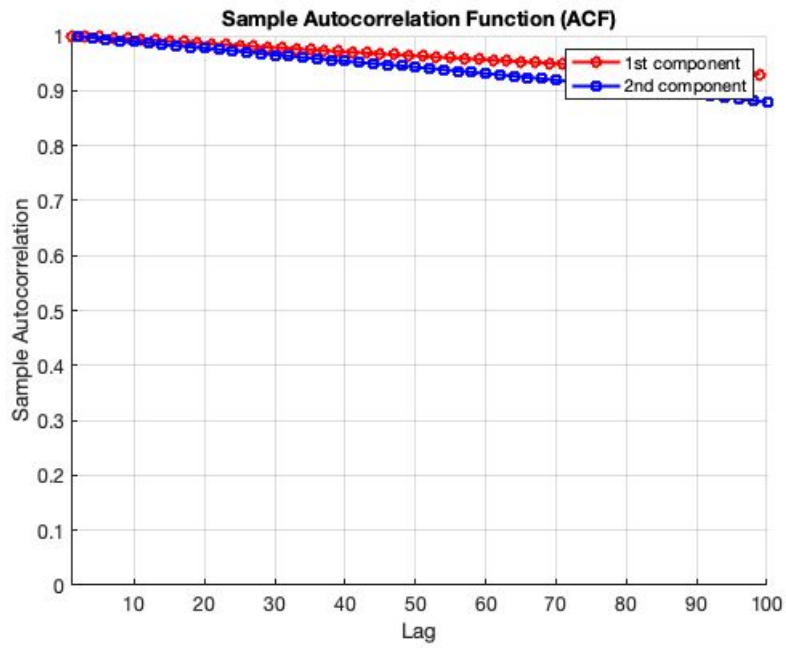
(a) scatter plot

(b) *Trace plots of the first component of m*.

(a) *Autocorrelation function plot for both components of m*.

Figure 4. Plots for Ex. 11.3.

8.2: the conclusion is not true. Since most of samples are rejected you need to increase the number of samples in order to get good result. Also, you don't attempt to explain why the rejection is high (-2)
11.2 and 11.3; you need to tune the sigma to get optimal convergence rate. it is supposed to be better than random walk.