Markov Chain Monte-Carlo Method

Author: Xiaohua_Ye

1. MCMC 算法简介

MCMC 算法的核心是通过生成一个平稳分布为$p(x)$的 Markov chain 来生成遵循某一特定分布的随机数序列。其核心是构建满足细致平稳条件的不可约、非周期马氏链。马氏链细致平稳条件为：

$$p(i)P_{ij} = p(j)P_{ji}$$

Metropolis 算法是 MCMC 算法中较为基础的算法，可以将任意对称分布转化为其他分布。而后 Hastings 对 Metropolis 算法进行了一些改进，扩大了提议分布（proposal distribution）的来源，使得非对称分布也可以成为提议分布。在生成遵循高维分布的随机向量时，有三种不同的 M - H 算法的变式可供人们选择。Blockwise M - H 算法通过从与目标分布相同维度的提议分布中抽取随机向量对状态进行更新，Componentwise M - H 算法则依次从一维提议分布中抽取随机数对随机向量某一个维度分量进行更新，这两种 M - H 算法的高维版本虽然只要求我们知道联合分布即可，但是在采样过程中可能存在大量的拒绝状态转移的情形，进而降低了算法效率，使其收敛速度变慢，这种情况随着维度的增加会变的越发明显。Gibbs Sampling 算法则解决了这一问题，通过将条件分布作为提议分布的方式，Gibbs Sampling 确保每一次都可以进行状态转移，使得算法有了更快的收敛速度，然而相应的，Gibbs Sampling 要求我们能够生成服从条件分布的随机数。理论上 MCMC 算法可以生成任意分布的随机数，在蒙特卡洛积分、贝叶斯估计、随机数生成等领域都有着巨大的作用。

2. 常见的 MCMC 算法及其实例

Algorithm 1. Metropolis 算法

1. 初始化 $t = 1$
2. 初始化 $X_0 = x_0$
3. 重复下述步骤直到$t = T$：

   $t = t + 1$

   从任意已知对称分布 Q 中选取 $x^* \sim Q(x \mid X_{t-1})$

   计算接受率 $\alpha = \min(1, \frac{p(x^*)}{p(X_{t-1})})$

   选取 $u \sim U(0,\ 1)$

   如果 $u \le \alpha$ 则 $X_t = x^*$ 否则 $X_t = X_{t-1}$

eg. 1. Metropolis 算法实例：利用正态分布生成标准柯西分布

注：柯西分布概率密度函数

$$pCauchy(x) = \frac{1}{\pi\alpha\left(1+\left(\frac{x-\theta}{\alpha}\right)^2\right)}$$

对于标准柯西分布而言，我们有 $\theta = 0$，$\alpha = 1$.

核心伪代码：

$X_0 = 0$

for i from 1 to T begin:

  $x^* = X_{i-1} + rndfromNormal(0,1)$ //即$x^* \sim N(X_{i-1},1)$

  $accept = pCauchy(x^*)/pCauchy(X_{i-1})$

  if $rnd(1) \le accept$ then begin:

    $X_i = x^*$ //以概率$accept$接受转移状态

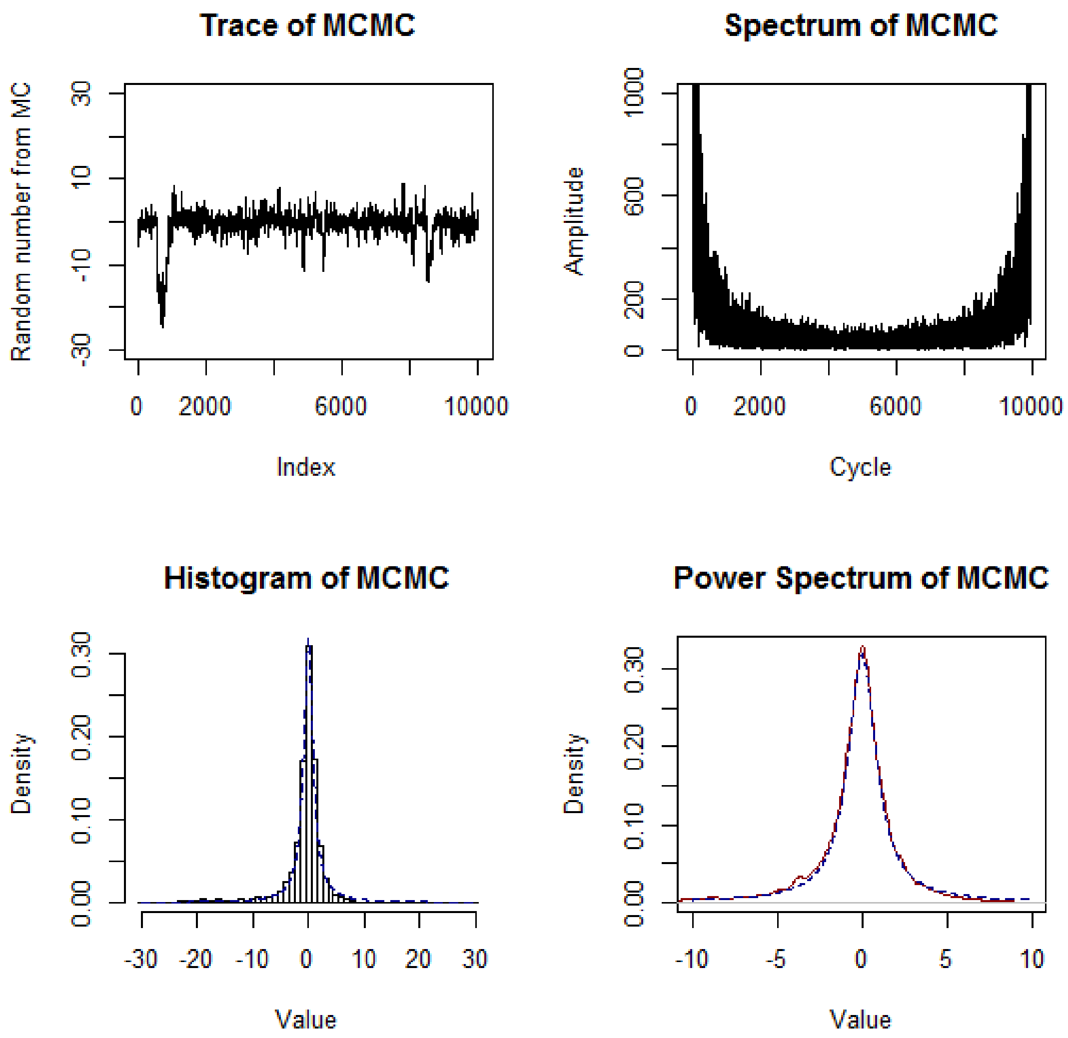
  else:

    $X_i = X_{i-1}$ //以概率$1 - accept$拒绝转移

  end if

end for

算法运行结果：



Algorithm 2. Metropolis - Hastings 算法
1. 初始化 $t = 1$
2. 初始化 $X_0 = x_0$
3. 重复下述步骤直到$t = T$：
$t = t + 1$
从任意已知 $p.d.f.$ 为 q 的分布 Q 中选取 $x^* \sim Q(x \mid X_{t-1})$
计算接受率 $\alpha = \min(1, \frac{p(x^*)}{p(X_{t-1})} \frac{q(X_{t-1} \mid x^*)}{q(x^* \mid X_{t-1})})$
选取 $u \sim U(0,\ 1)$
如果 $u \leq \alpha$ 则 $X_t = x^*$ 否则 $X_t = X_{t-1}$

eg. 2. Metropolis - Hastings 算法实例：利用卡方分布生成服从参数 1 的指数分布

注：服从参数λ的指数分布的概率密度函数

$$pExp(x) = \lambda \exp(-\lambda x)$$

核心伪代码：

$X_0 = 1$

for i from 1 to T begin:

$x^* = rndfromchisq(df = X_{i-1})$

$accept = \frac{pExp(x^*)}{pExp(X_{i-1})} * \frac{pChisq(X_{i-1}|df=x^*)}{pChisq(x^*|df=X_{i-1})}$ //$pChisq$即卡方分布概率密度

if $rnd(1) \le accept$ then begin:

$X_i = x^*$ //以概率$accept$接受转移状态

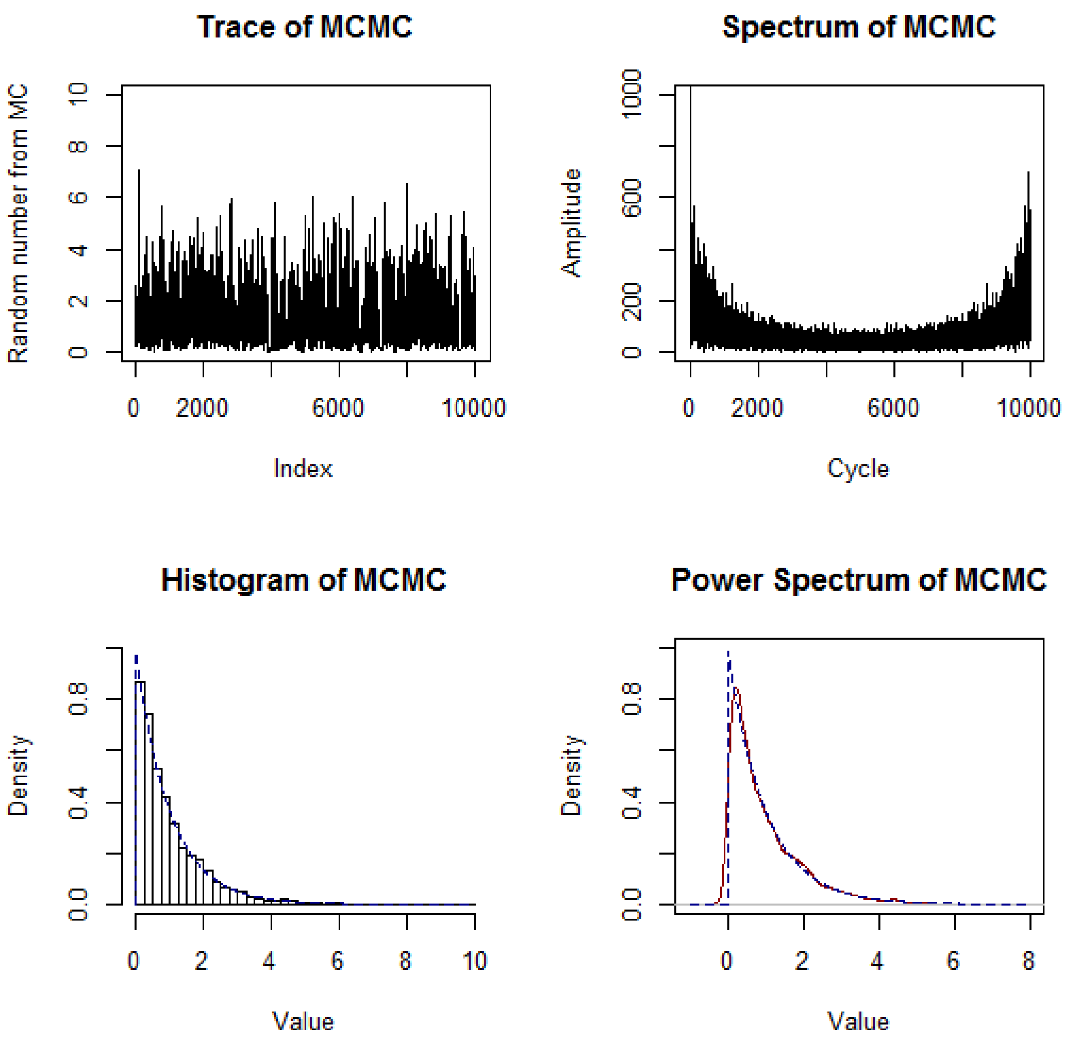
else:

$X_i = X_{i-1}$ //以概率$1 - accept$拒绝转移

end if

end for

算法运行结果：



Algorithm 3. Blockwise Metropolis - Hastings 算法
1. 初始化 $t = 1$

2. 初始化 $\boldsymbol{X_0} = (x_{01}, x_{02}, x_{03}, \dots, x_{0n})$
3. 重复下述步骤直到$t = T$：

   $t = t + 1$

   从任意已知 $p.d.f.$ 为 q 的 n 维分布 Q 中选取向量 $\boldsymbol{X}^* \sim Q(\boldsymbol{X} \mid \boldsymbol{X_{t-1}})$

   计算接受率 $\alpha = \min(1, \frac{p(\boldsymbol{X}^*)}{p(\boldsymbol{X_{t-1}})} \frac{q(\boldsymbol{X_{t-1}} \mid \boldsymbol{X}^*)}{q(\boldsymbol{X}^* \mid \boldsymbol{X_{t-1}})})$

   选取 $u \sim U(0,\ 1)$

   如果 $u \le \alpha$ 则 $\boldsymbol{X_t} = \boldsymbol{X}^*$ 否则 $\boldsymbol{X_t} = \boldsymbol{X_{t-1}}$

eg. 3. Blockwise M - H 算法实例：利用二维均匀分布生成二维正态分布

注：二维正态分布的概率密度函数

$$p2DimNorm(x,y) = \frac{exp\left(-\frac{(x-\mu_1)^2}{2(1-\rho^2)\sigma_1{}^2} - \frac{(y-\mu_2)^2}{2(1-\rho^2)\sigma_2{}^2} + \frac{\rho(x-\mu_1)(y-\mu_2)}{(1-\rho^2)\sigma_1\sigma_2}\right)}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

本例中的二维正态分布参数为：$\mu_1 = 3, \mu_2 = 0, \sigma_1 = 2, \sigma_2 = 1, \rho = 0.7$

核心伪代码：

$\boldsymbol{X_0} = (0, 0)$

for i from 1 to T begin:

  $\boldsymbol{X}^* = \boldsymbol{X_{i-1}} + 2 * rnd(2) - 1$ //即$\boldsymbol{X}^* \sim U(X_{i-1} - 1, X_{i-1} + 1)$

  $accept = \frac{p2DimNorm(\boldsymbol{X}^*)}{p2DimNorm(\boldsymbol{X_{i-1}})}$ //均匀分布对称，因此只有一项

  if $rnd(1) \le accept$ then begin:

    $\boldsymbol{X_i} = \boldsymbol{X}^*$ //以概率$accept$接受转移状态

  else:

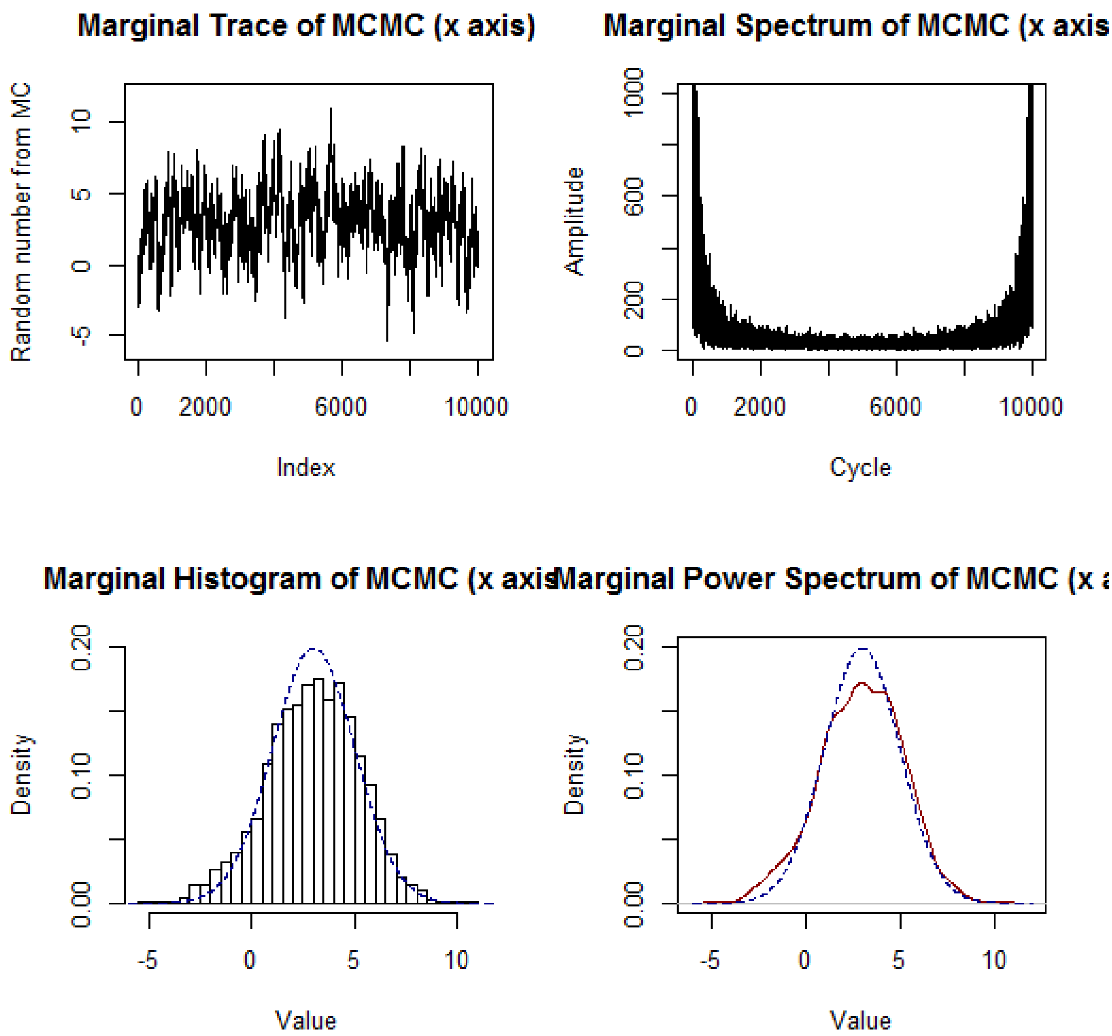
    $\boldsymbol{X_i} = \boldsymbol{X_{i-1}}$ //以概率$1 - accept$拒绝转移

  end if

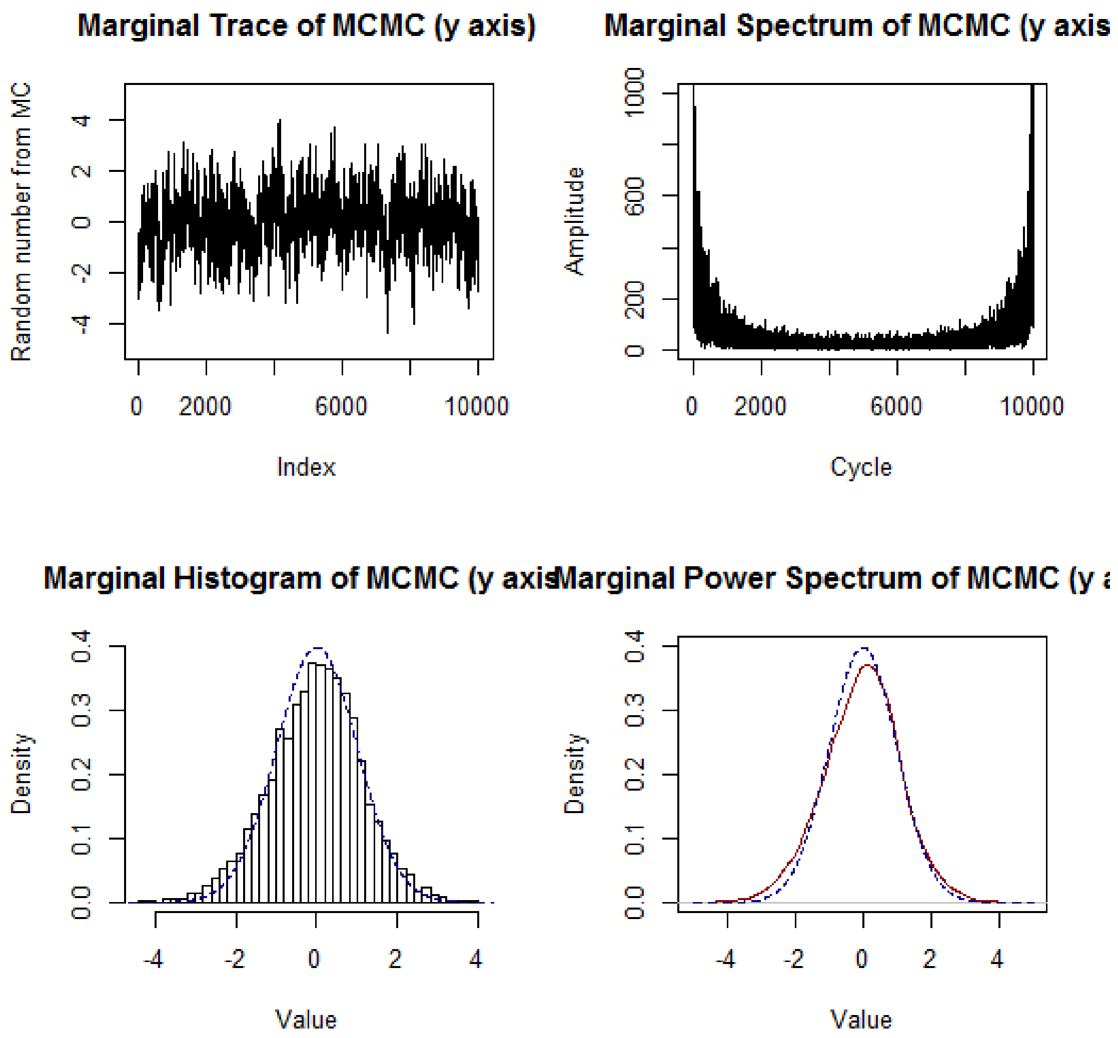
end for

算法运行结果：

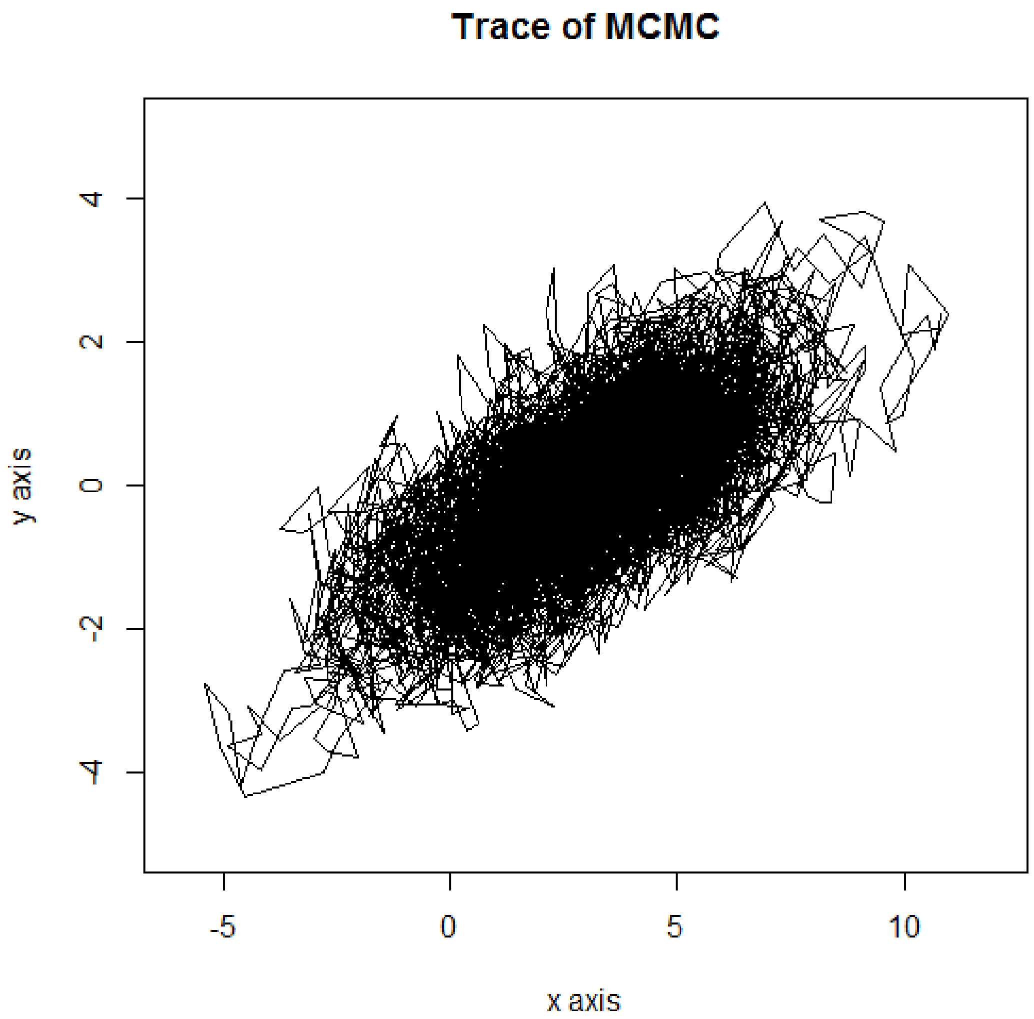
x 轴边际分布：


Marginal Trace of MCMC (x axis)
Random number from MC
-5
0
5
10
0
2000
6000
10000
Index
Marginal Spectrum of MCMC (x axis
Amplitude
0
200
600
1000
0
2000
6000
10000
Cycle
Marginal Histogram of MCMC (x axis
Density
0.00
0.10
0.20
-5
0
5
10
Value
Marginal Power Spectrum of MCMC (x a
Density
0.00
0.10
0.20
-5
0
5
10
Value

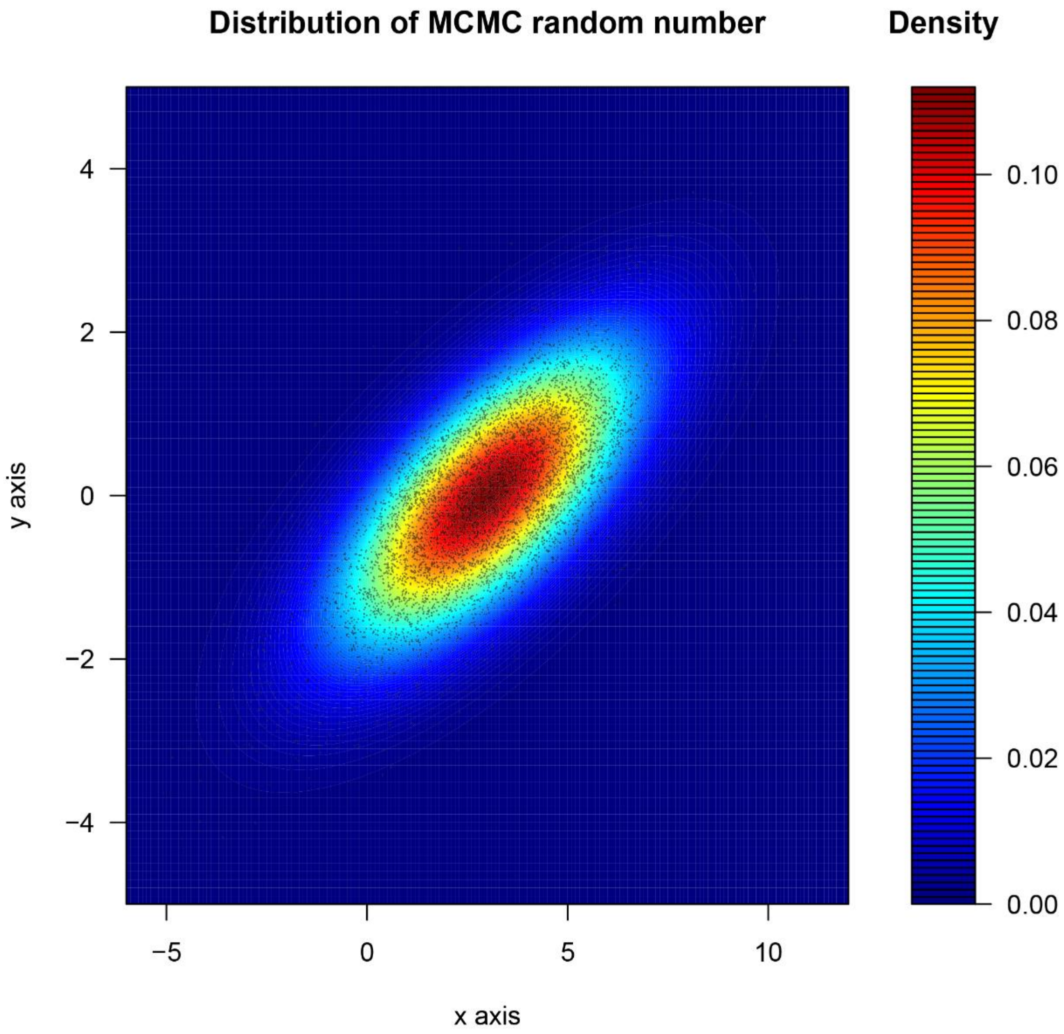
y 轴边际分布：


Marginal Trace of MCMC (y axis)
Random number from MC
Index
Marginal Spectrum of MCMC (y axis
Amplitude
Cycle
Marginal Histogram of MCMC (y axis
Density
Value
Marginal Power Spectrum of MCMC (y a
Density
Value

轨迹图：



联合分布投影图：



Algorithm 4. Componentwise Metropolis - Hastings 算法

1. 初始化 $t=1$
2. 初始化 $\boldsymbol{X_0}=(x_{01},x_{02},x_{03},\dots,x_{0n})$
3. 重复下述步骤直到$t=T$：

   $t=t+1$

   对每一维：$i=1,2,3,\dots,n$

   从任意已知 $p.d.f.$ 为 q 的一维分布 Q 中选取 $x_i^*\sim Q(x\mid x_{t-1,i})$

   令 $\boldsymbol{X}=(x_{t-1,1},x_{t-1,2},\dots,x_i^*,\dots,x_{t-1,n})$

   计算接受率 $\alpha=\min(1,\ \frac{p(\boldsymbol{X})}{p(\boldsymbol{X_{t-1}})}\frac{q(x_{t-1,i}\mid x_i^*)}{q(x_i^*\mid x_{t-1,i})})$

   选取 $u\sim U(0,\ 1)$

   如果 $u\le\alpha$ 则 $x_{ti}=x_i^*$ 否则 $x_{ti}=x_{t-1,i}$

eg. 4. Componentwise M - H 算法实例：利用一维均匀分布生成二维正态分布

注：二维正态分布的概率密度函数

$$p2DimNorm(x,y)=\frac{exp\left(-\frac{(x-\mu_1)^2}{2(1-\rho^2){\sigma_1}^2}-\frac{(y-\mu_2)^2}{2(1-\rho^2){\sigma_2}^2}+\frac{\rho(x-\mu_1)(y-\mu_2)}{(1-\rho^2)\sigma_1\sigma_2}\right)}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

本例中的二维正态分布参数为：$\mu_1=3,\mu_2=0,\sigma_1=2,\sigma_2=1,\rho=0.7$

核心伪代码：

$\boldsymbol{X_0}=(0,0)$

for i from 1 to T begin:

for j form 1 to 2 begin:

$x_j^*=x_{i-1,j}+2*rnd(1)-1$ //即$x_j^*\sim U\left(x_{i-1,j}-1,x_{i-1,j}+1\right)$

if $j=1$ then $\boldsymbol{X_j^*}=\left(x_j^*,x_{i-1,2}\right)$ else $\boldsymbol{X_j^*}=\left(x_{i-1,1},x_j^*\right)$

$accept_1=\frac{p2DimNorm(\boldsymbol{X_1^*})}{p2DimNorm(\boldsymbol{X_{i-1}})}$ //均匀分布对称，因此只有一项

if $rnd(1)\le accept_1$ then begin:

$x_{ij}=x_j^*$ //以概率$accept$接受转移状态

else:

$x_{ij}=x_{i-1,j}$ //以概率$1-accept$拒绝转移

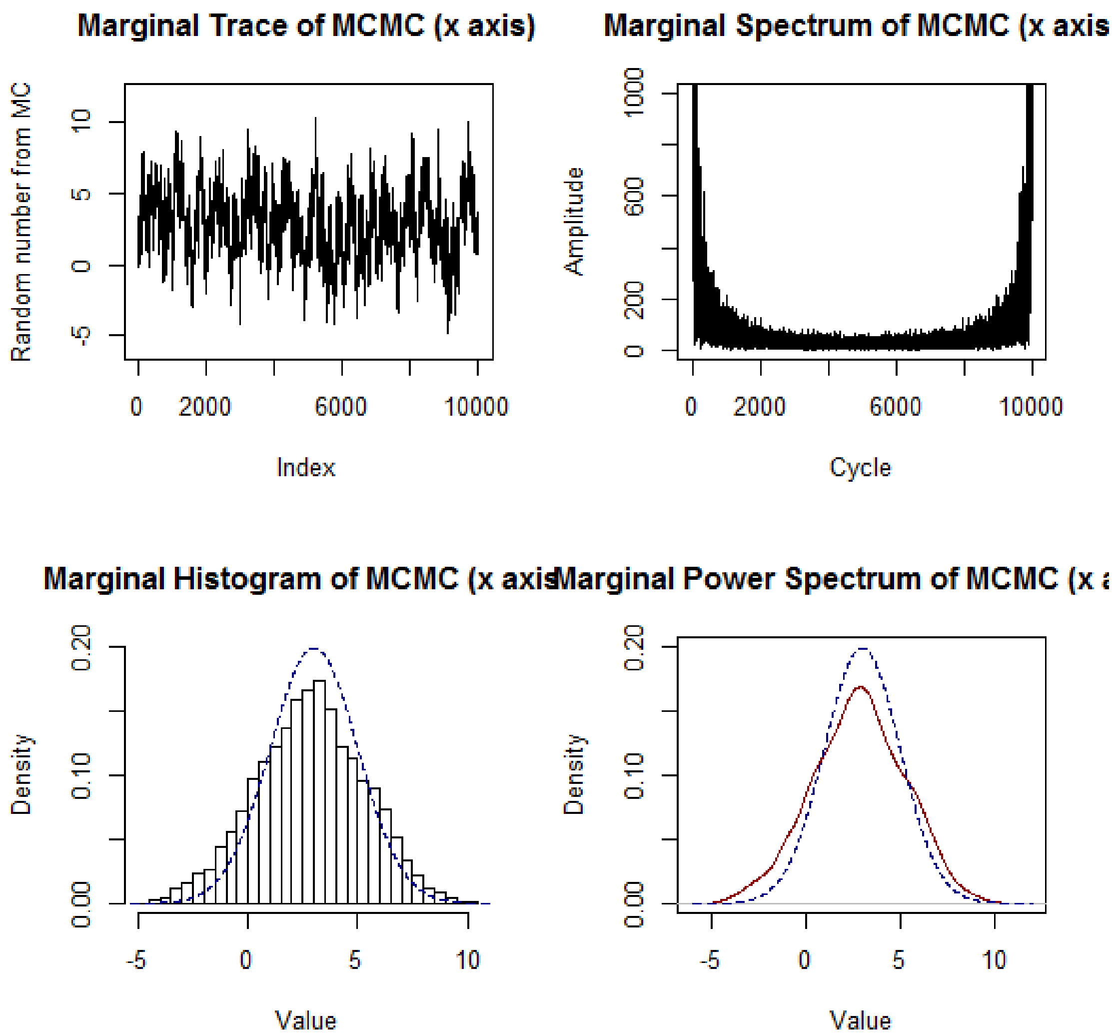
end if

end for j

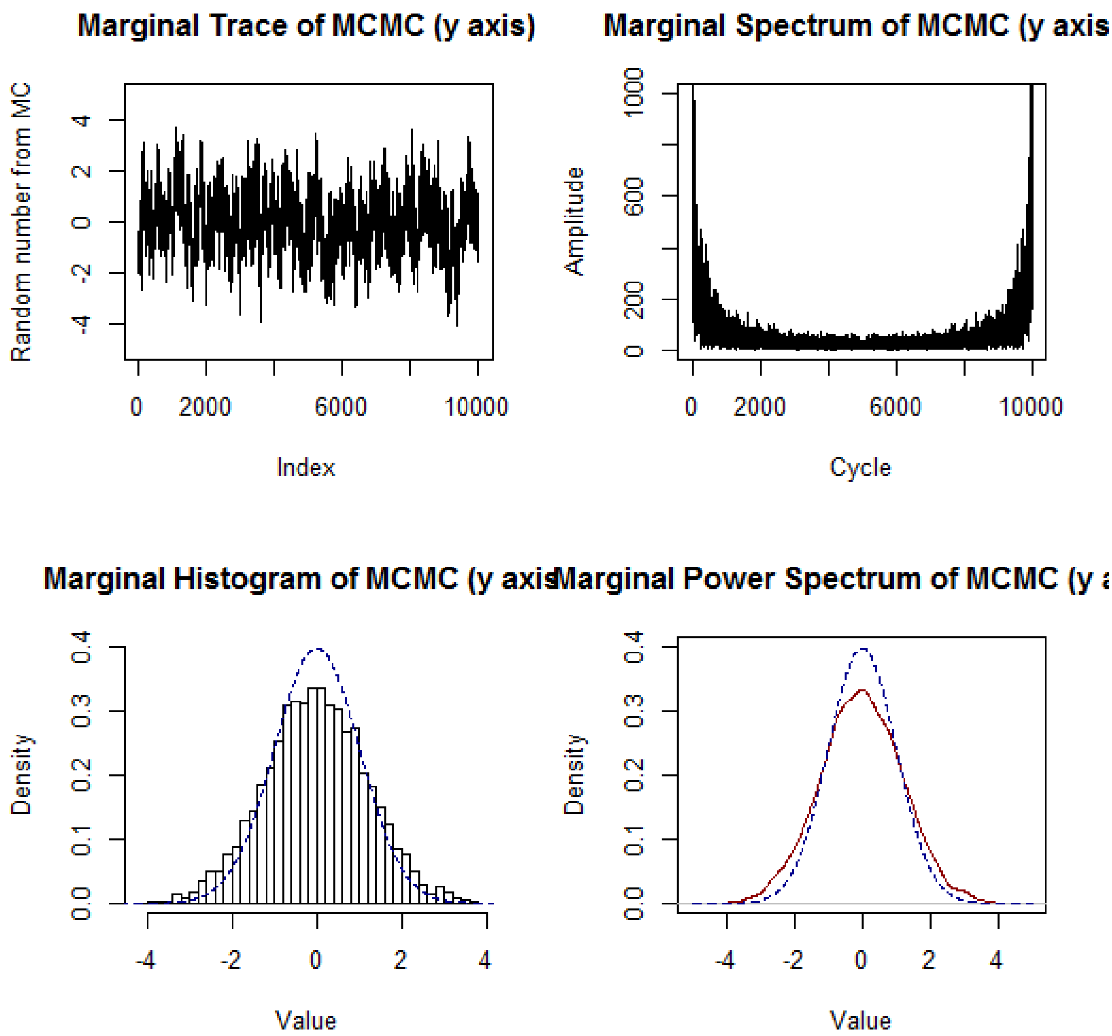
end for i

算法运行结果：

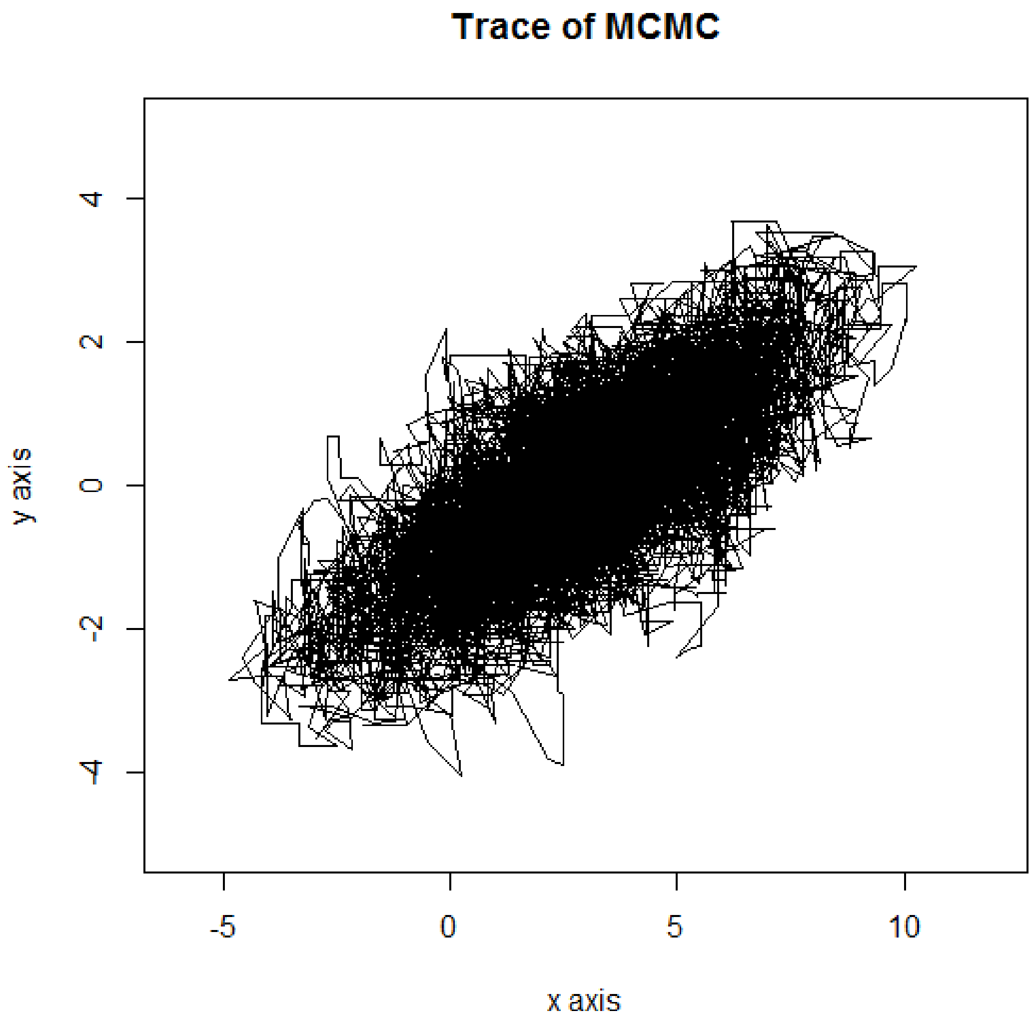
x 轴边际分布：


Marginal Trace of MCMC (x axis)
Random number from MC
Index
Marginal Spectrum of MCMC (x axis
Amplitude
Cycle
Marginal Histogram of MCMC (x axis
Density
Value
Marginal Power Spectrum of MCMC (x a
Density
Value

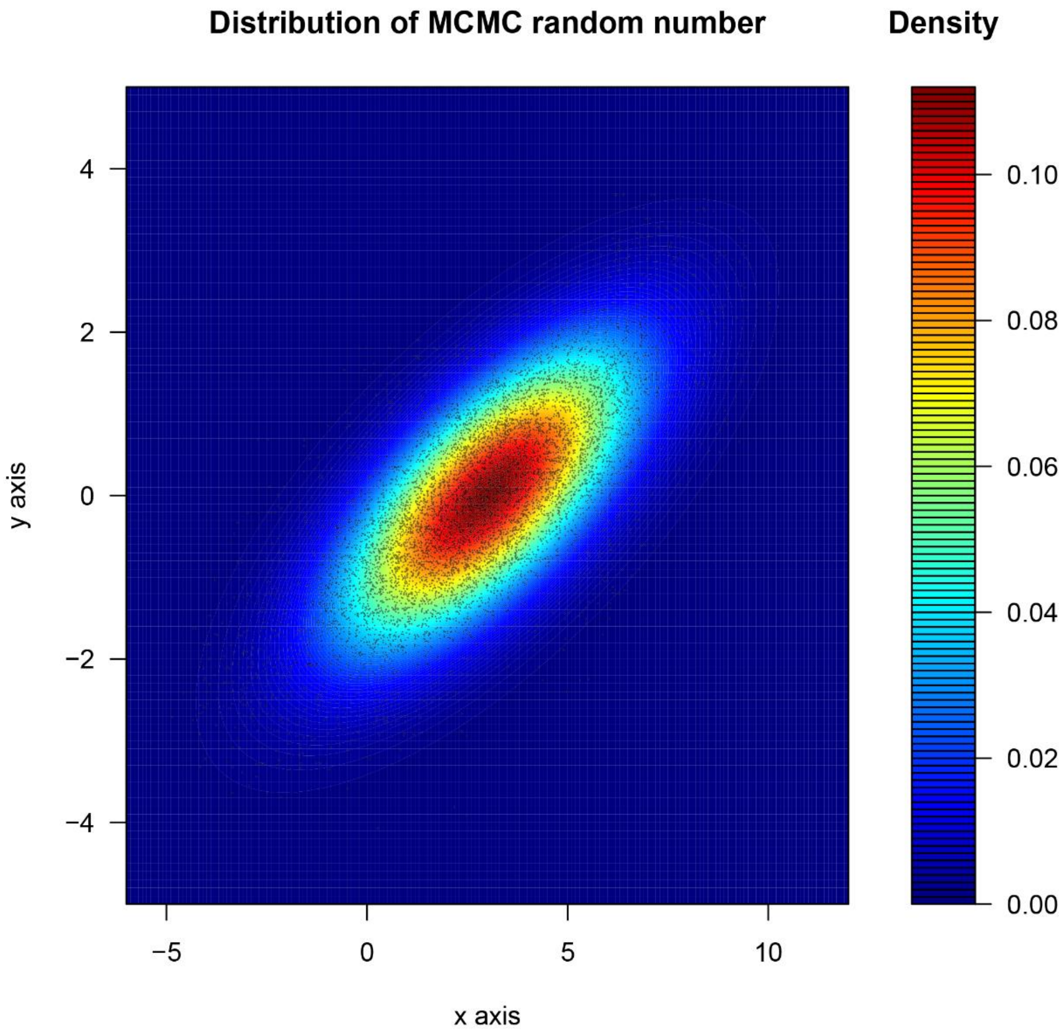
y 轴边际分布：


Marginal Trace of MCMC (y axis)
Random number from MC
-4 -2 0 2 4
0 2000 6000 10000
Index
Marginal Spectrum of MCMC (y axis
Amplitude
0 200 600 1000
0 2000 6000 10000
Cycle
Marginal Histogram of MCMC (y axis
Density
0.0 0.1 0.2 0.3 0.4
-4 -2 0 2 4
Value
Marginal Power Spectrum of MCMC (y a
Density
0.0 0.1 0.2 0.3 0.4
-4 -2 0 2 4
Value

轨迹图：



联合分布投影图：



Algorithm 5. Gibbs Sampling 算法
1. 初始化 $t=1$
2. 初始化 $\boldsymbol{X_0}=(x_{01},x_{02},x_{03},\dots,x_{0n})$
3. 重复下述步骤直到$t=T$：
$t=t+1$
对每一维：$i=1,2,3,\dots,n$
选取 $x_{t1}\sim p(x\mid x_{t-1,2},x_{t-1,3},\dots,x_{t-1,n})$
$x_{t2}\sim p(x\mid x_{t1},x_{t-1,3},\dots,x_{t-1,n})$
……
$x_{ti}\sim p(x\mid x_{t1},x_{t2},\dots,x_{t,i-1},x_{t-1,i+1},\dots,x_{t-1,n})$
……
$x_{tn}\sim p(x\mid x_{t1},x_{t2},\dots,x_{t,n-1})$

eg. 5. Gibbs Sampling 算法实例：生成二维正态分布

注：二维正态分布的概率密度函数

$$p2DimNorm(x,y)=\frac{exp\left(-\frac{(x-\mu_1)^2}{2(1-\rho^2){\sigma_1}^2}-\frac{(y-\mu_2)^2}{2(1-\rho^2){\sigma_2}^2}+\frac{\rho(x-\mu_1)(y-\mu_2)}{(1-\rho^2)\sigma_1\sigma_2}\right)}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

本例中的二维正态分布参数为：$\mu_1=3,\mu_2=0,\sigma_1=2,\sigma_2=1,\rho=0.7$

其条件分布依然为正态分布，且有

$$condNorm(x\mid y=y_0)\sim N\left(\rho\frac{\sigma_1}{\sigma_2}(y_0-\mu_2)+\mu_1,\sigma_1\sqrt{1-\rho^2}\right)$$

$$condNorm(y\mid x=x_0)\sim N\left(\rho\frac{\sigma_2}{\sigma_1}(x_0-\mu_1)+\mu_2,\sigma_2\sqrt{1-\rho^2}\right)$$

核心伪代码：

$\boldsymbol{X_0}=(0,0)$

for i from 1 to T begin:

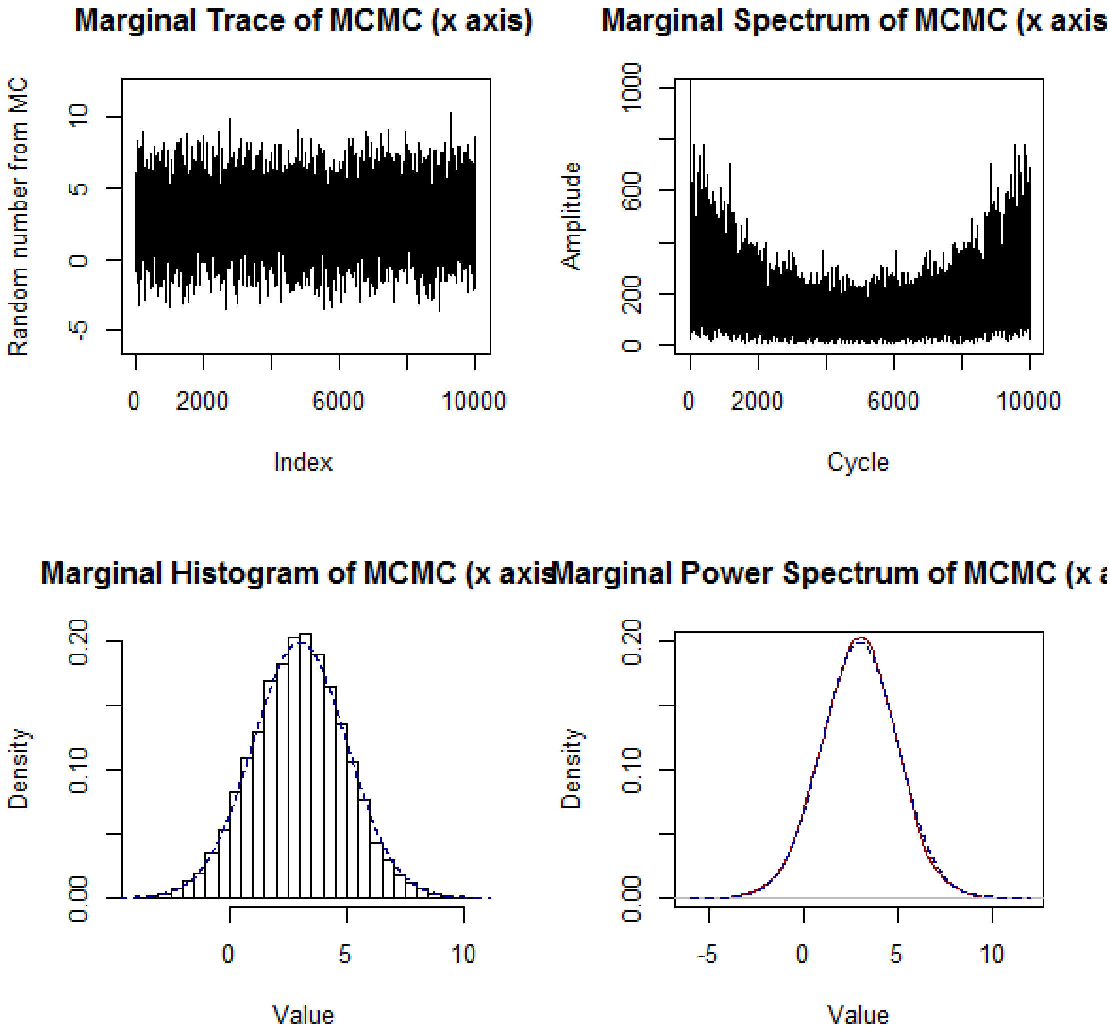
$$x_{t2}=rndfromNorm\left(\rho\frac{\sigma_2}{\sigma_1}\left(x_{t-1,1}-\mu_1\right)+\mu_2,\sigma_2\sqrt{1-\rho^2}\right)$$

$$x_{t1}=rndfromNorm\left(\rho\frac{\sigma_1}{\sigma_2}\left(x_{t2}-\mu_2\right)+\mu_1,\sigma_1\sqrt{1-\rho^2}\right)$$

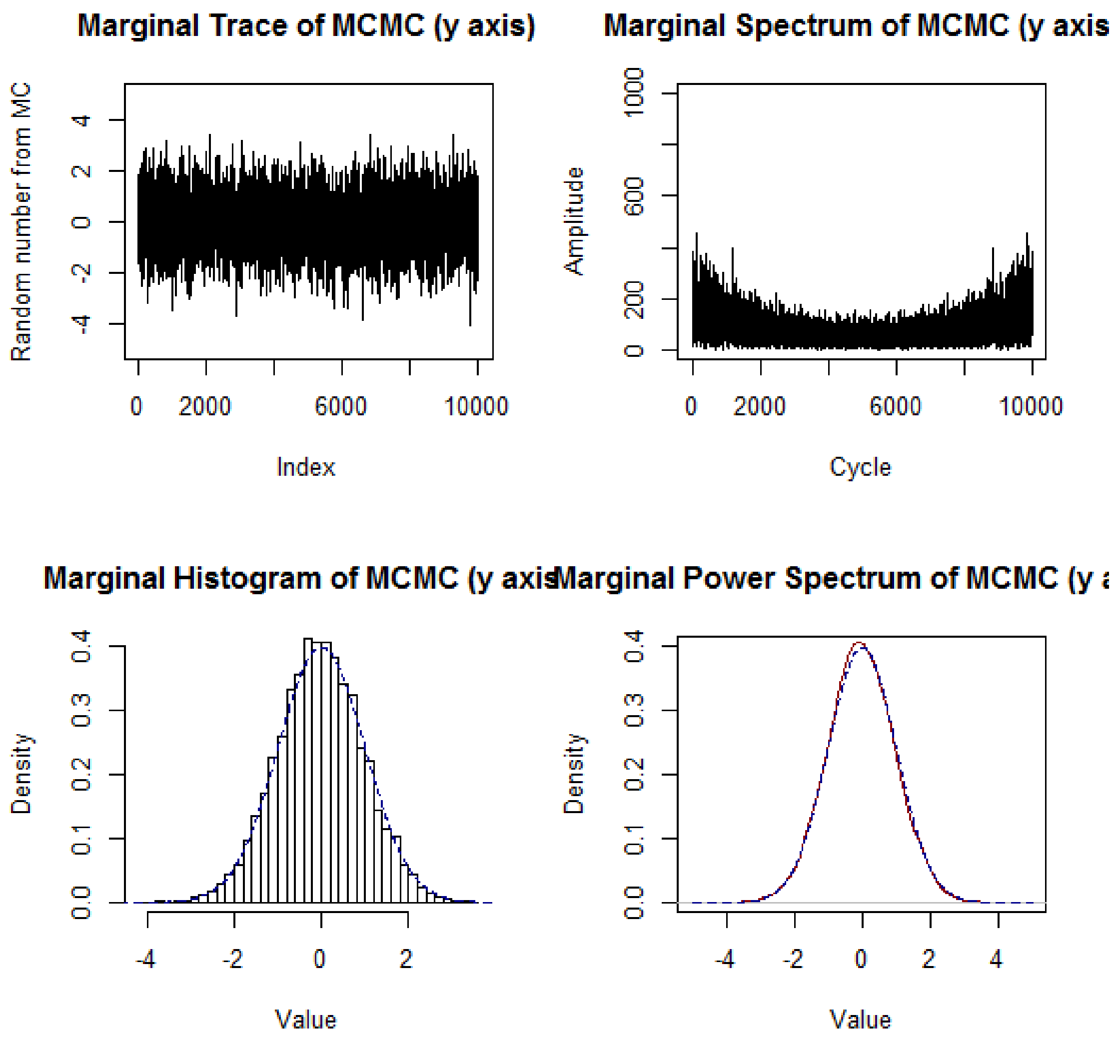
end for i

算法运行结果：

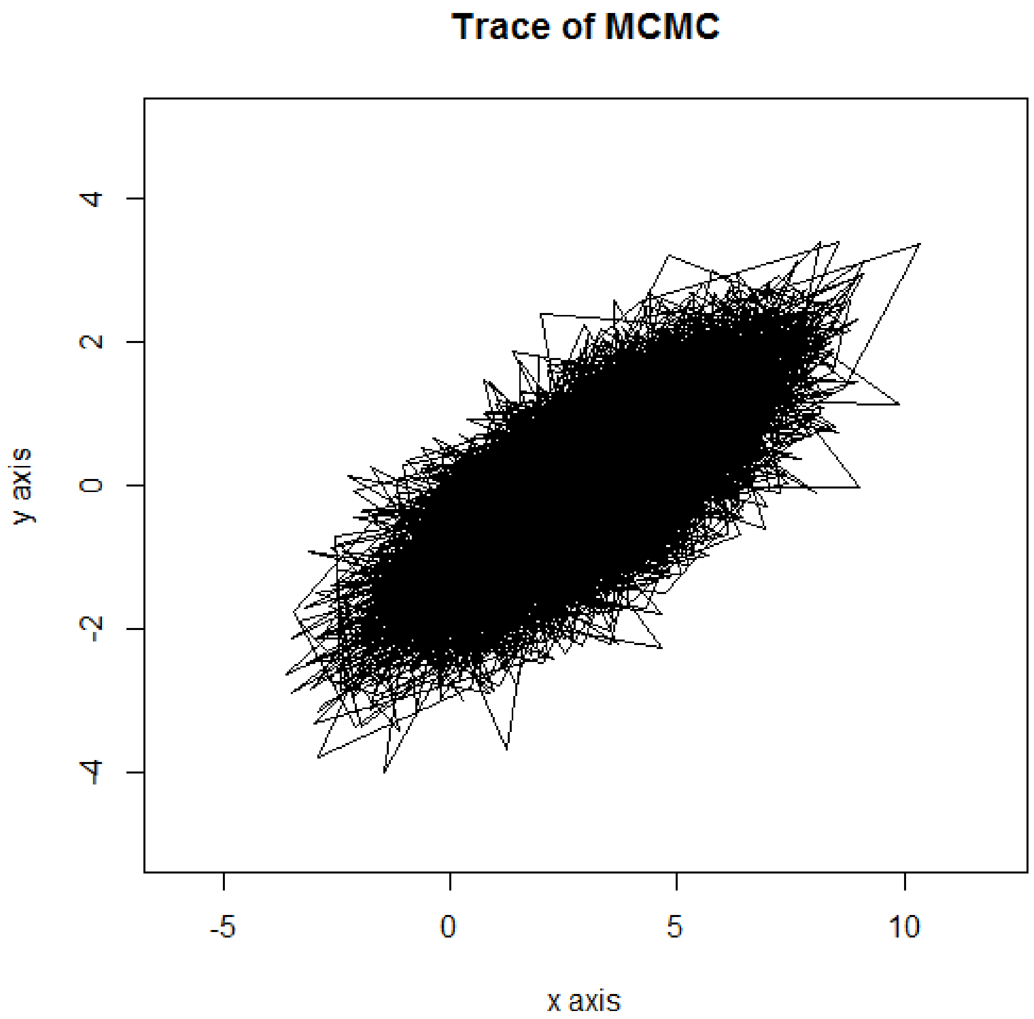
x 轴边际分布：


Marginal Trace of MCMC (x axis)
Random number from MC
Index
Marginal Spectrum of MCMC (x axis
Amplitude
Cycle
Marginal Histogram of MCMC (x axis
Density
Value
Marginal Power Spectrum of MCMC (x a
Density
Value

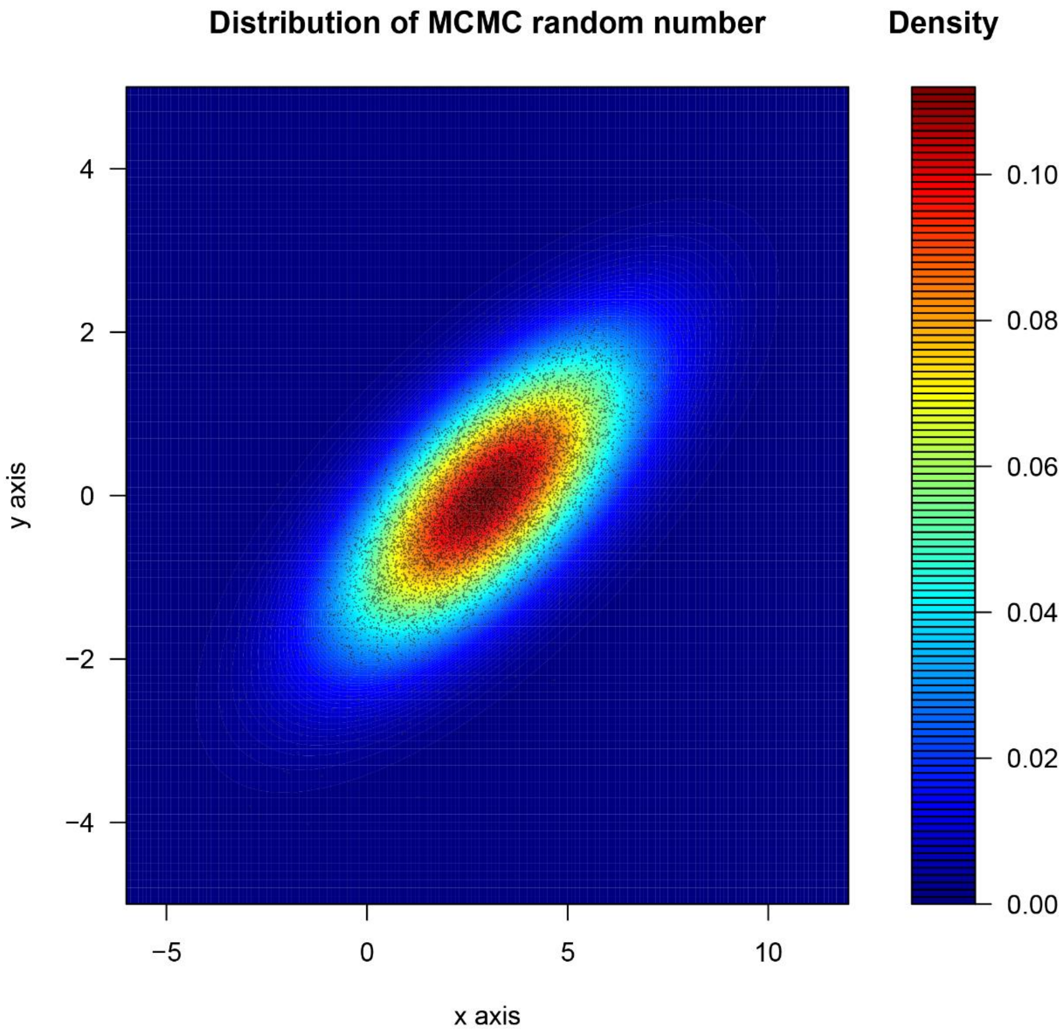
y 轴边际分布：


Marginal Trace of MCMC (y axis)
Random number from MC
4
2
0
-2
-4
0
2000
6000
10000
Index
Marginal Spectrum of MCMC (y axis
Amplitude
1000
600
200
0
0
2000
6000
10000
Cycle
Marginal Histogram of MCMC (y axis
Density
0.4
0.3
0.2
0.1
0.0
-4
-2
0
2
Value
Marginal Power Spectrum of MCMC (y a
Density
0.4
0.3
0.2
0.1
0.0
-4
-2
0
2
4
Value

轨迹图：



联合分布投影图：


Distribution of MCMC random number
Density
4
2
0
−2
−4
y axis
−5
0
5
10
x axis
0.10
0.08
0.06
0.04
0.02
0.00