Part 4 MC and multivariate distribution

Dr. Nguyen Quang Huy

Feb 18, 2021

MC simulation and multivariate distribution

To simulate $(X_1, X_2, \cdots, X_n)$ with distribution function $F(x_1, x_2, \cdots, x_n)$:

1. Calculate $F_1(x_1)$ the distribution function of X_1; simulate $U_1 \sim \mathcal{U}(0,1)$ and return $X_1 = F_1^{\leftarrow}(U_1)$.
2. Calculate $F_{2|1}(x_2|x_1)$, the conditional distribution function of X_2 given X_1; simulate $U_2 \sim \mathcal{U}(0,1)$ and return $X_2 = F_{2|1}^{\leftarrow}(U_2)$.
3. Calculate $F_{3|2,1}(x_3|x_2, x_1)$, the conditional distribution function of X_3 given X_2 and X_1; simulate $U_3 \sim \mathcal{U}(0,1)$ and return $X_3 = F_{3|2,1}^{\leftarrow}(U_3)$.

...

n. Calculate $F_{n|n-1,\ldots,1}(x_n|x_{n-1}, \ldots, x_1)$, the conditional distribution function of X_n given $X_{n-1}, \ldots, X_1$; simulate $U_n \sim \mathcal{U}(0,1)$ and return $X_n = F_{n|(n-1),\ldots,1}^{\leftarrow}(U_n)$.

MC simulation and multivariate distribution

Example: Given (X, Y) with multivariate distribution function $F(x, y) : (0, \infty)^2 \to [0, 1]$ and

$$F(x, y) = 1 - \frac{1}{x+1} - \frac{1}{y+1} + \frac{1}{1+x+y}$$

1. Calculate the distribution F_X of X and F_Y of Y?
2. X and Y are independent?
3. Calculate the conditional distribution $F_{Y|X}(y|x)$

$$F_{Y|X}(y|x) = \mathbb{P}(Y \leq y | X = x) = \frac{\partial F(x, y)/\partial x}{f_X(x)}$$

4. Simulate X based on F_X then simulate Y based on $F_{Y|X}(y|x)$

MC simulation and multivariate distribution

1. Distribution F_X of X and F_Y of Y

MC simulation and multivariate distribution

1. Distribution F_X of X and F_Y of Y

$$\begin{aligned} F_X(x) &= F(x,\infty) = 1 - \frac{1}{x+1} - \frac{1}{1+\infty} + \frac{1}{1+x+\infty} \\ &= 1 - \frac{1}{x+1} \end{aligned}$$

Similarly,

$$F_Y(y) = 1 - \frac{1}{y+1}$$

Density functions of X and Y

$$\begin{aligned} f_X(x) &= \frac{1}{(x+1)^2} \\ f_Y(y) &= \frac{1}{(y+1)^2} \end{aligned}$$

MC simulation and multivariate distribution

2. X and Y are independent

MC simulation and multivariate distribution

2. X and Y are independent if and only if

$$\begin{aligned} F(x,y) &= \mathbb{P}(X \le x, Y \le y) \\ &= \mathbb{P}(X \le x)\mathbb{P}(Y \le y) \\ &= F_X(x)F_Y(y) \quad \forall x, y \end{aligned}$$

However

$$F(x,y) = 1 - \frac{1}{x+1} - \frac{1}{y+1} + \frac{1}{1+x+y}$$
$$F_X(x)F_Y(y) = \left(1 - \frac{1}{x+1}\right)\left(1 - \frac{1}{y+1}\right)$$

$\rightarrow$ X and Y are NOT independent.

MC simulation and multivariate distribution

3. Calculate the conditional distribution $F_{Y|X}(y|x)$

$$F_{Y|X}(y|x) = \frac{\partial F(x,y)/\partial x}{f_X(x)}$$

We have

$$\frac{\partial F(x,y)}{\partial x} = \frac{1}{(1+x)^2} - \frac{1}{(1+x+y)^2}$$

then

$$\begin{aligned} F_{Y|X}(y|x) &= \left(\frac{1}{(1+x)^2} - \frac{1}{(1+x+y)^2}\right)(1+x)^2 \\ &= 1 - \frac{(1+x)^2}{(1+x+y)^2} \end{aligned}$$

MC simulation and multivariate distribution

4. Calculate the inverse function of $F_X(x)$ and $F_{Y|X}(y|x)$ (a function of y given x)

MC simulation and multivariate distribution

4. Calculate the inverse function of $F_X(x)$ and $F_{Y|X}(y|x)$ (a function of y given x)

$$
\begin{aligned}
F_X^{\leftarrow}(t) &= \frac{1}{1-t} - 1 \\
F_{Y|X}^{\leftarrow}(t|x) &= (1+x)\left(\frac{1}{\sqrt{1-t}} - 1\right)
\end{aligned}
$$

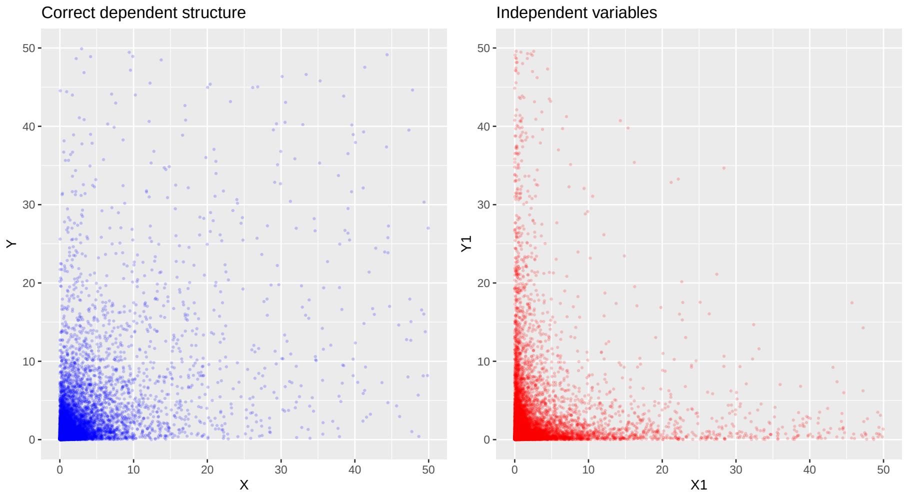
To simulate (X, Y)

- Simulate $U_1 \sim \mathcal{U}(0,1)$, return $X = \dfrac{1}{1-U_1} - 1$
- Simulate $U_2 \sim \mathcal{U}(0,1)$, return $Y = (1+X_1)\left(\dfrac{1}{\sqrt{1-U_2}} - 1\right)$

MC simulation and multivariate distribution

```
N<-10^3
u1<-runif(N,0,1)
u2<-runif(N,0,1)
X<-(1)/(1-u1)-1
Y<-(1+X)*(1/(1-u2)^0.5-1)
# INDEPENDENT X1 and Y1
X1<-(1)/(1-u1)-1
Y1<-(1)/(1-u2)-1
dat<-data.frame("X"=X,"Y"=Y,"X1"=X1,"Y1"=Y1)
p1<-ggplot(dat,aes(X,Y))+geom_point(cex=0.5,alpha=0.5,color="b
  ggtitle("Correct dependent structure")
p2<-ggplot(dat,aes(X1,Y1))+geom_point(cex=0.5,alpha=0.5,color=
  ggtitle("Independent variable")
grid.arrange(p1, p2, ncol=2)
```

MC simulation and multivariate distribution



MC simulation and multivariate distribution

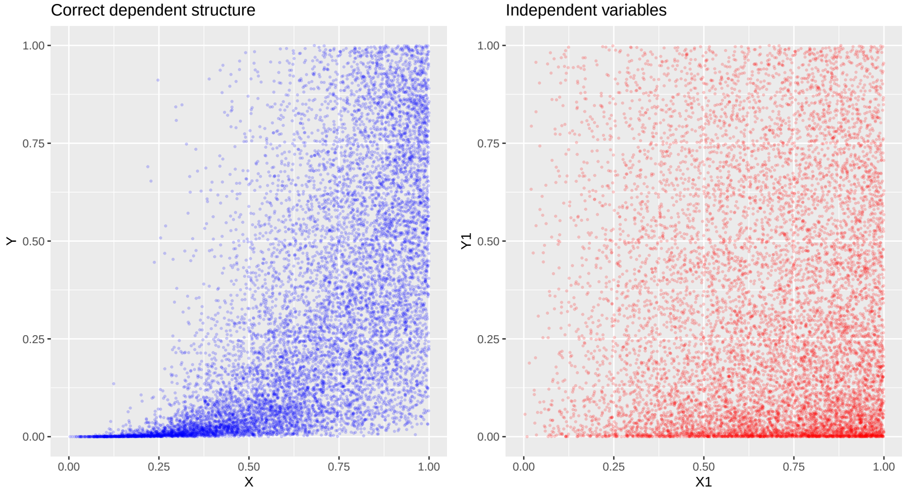
Given (X,Y) with multivariate distribution function $F(x, y) : (0, 1)^2 \to [0, 1]$ and

$$F(x, y) = \left(x^{-2\theta} + y^{-\theta/2} - 1\right)^{-1/\theta}$$

with $\theta > 0$

1. Calculate $F_X(x)$, $F_{Y|X}(y|x)$
2. Calculate $F_X^{\leftarrow}(x)$, $F_{Y|X}^{\leftarrow}(y|x)$
3. With $\theta = ...$, simulate (X, Y)
4. Calculate $\mathbb{P}(X + Y > 0.8)$ with $\theta = ...$

MC simulation and multivariate distribution



MC simulation and multivariate normal distribution

- In some cases, the simulation of dependent random variables does not need the calculation of conditional distribution, multivariate normal random variable is an example.
- Suppose that $\mathbf{N} = (N_1, N_2, \cdots, N_k)^{'}$ is a k-dimensional normal random vector with mean μ and covariance matrix Σ

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \cdots \\ \mu_k \end{pmatrix} ; \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix}$$

- We can prove that $\mathbf{N} = \mu + \mathbf{CZ}$ where $\mathbf{Z} = (Z_1, Z_2, \cdots, Z_k)^{'}$ is the vector of k i.i.d standard normal r.v and matrix $\mathbf{C}$ is the Cholesky decomposition of Σ: $\Sigma = \mathbf{CC}^{'}$

MC simulation and multivariate normal distribution

Given $\mathbf{N} \sim \mathcal{N}(\mu, \Sigma)$ where

$$\mu = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \; ; \Sigma = \begin{pmatrix} 4 & 4 \\ 4 & 9 \end{pmatrix}$$

- Find a lower triangular matrix $\mathbf{C}$ such that $\Sigma = \mathbf{C}\mathbf{C}'$
- Using stochastic representation $\mathbf{N} = \mu + \mathbf{CZ}$ to simulate $\mathbf{N}$.

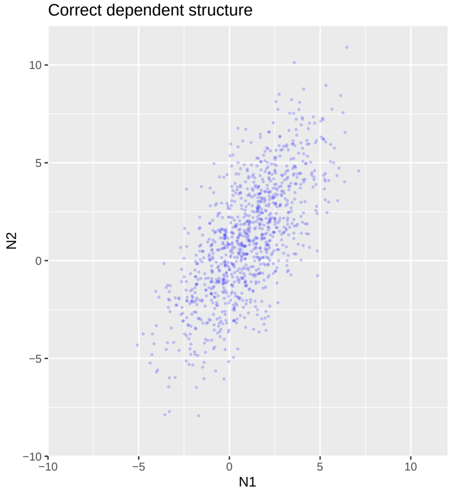
MC simulation and multivariate normal distribution

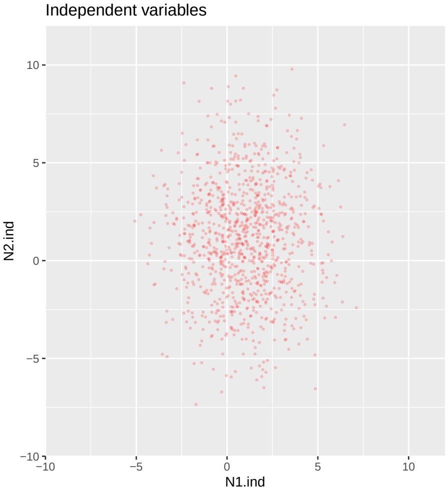
```
n<-1000
Z<-matrix(rnorm(2*n,0,1),2,n)
C<-matrix(c(2,2,0,5^0.5),2,2)
mu<-matrix(rep(c(1,1),n),2,n)
N<-mu + C %*% Z

# INDEPENDENT X1 and Y1
C1<-matrix(c(2,0,0,3),2,2)
N1<-mu+C1 %*% Z

dat<-data.frame("N1"=N[1,],"N2"=N[2,],"N1.ind"=N1[1,],"N2.ind"
p1<-ggplot(dat,aes(N1,N2))+geom_point(cex=0.5,alpha=0.2,color=
  ggtitle("Correct dependent structure")+xlim(c(-9,11))+ylim(c
p2<-ggplot(dat,aes(N1.ind,N2.ind))+geom_point(cex=0.5,alpha=0.
  ggtitle("Independent variables")+xlim(c(-9,11))+ylim(c(-9,11
grid.arrange(p1, p2, ncol=2)
```

MC simulation and multivariate normal distribution





MC simulation and multivariate normal distribution

- The function *chol(M)* return the **transpose** of the cholesky decomposition matrix M.

```
Sigma<-matrix(c(4,4,4,9),2,2)
t(chol(Sigma))
```

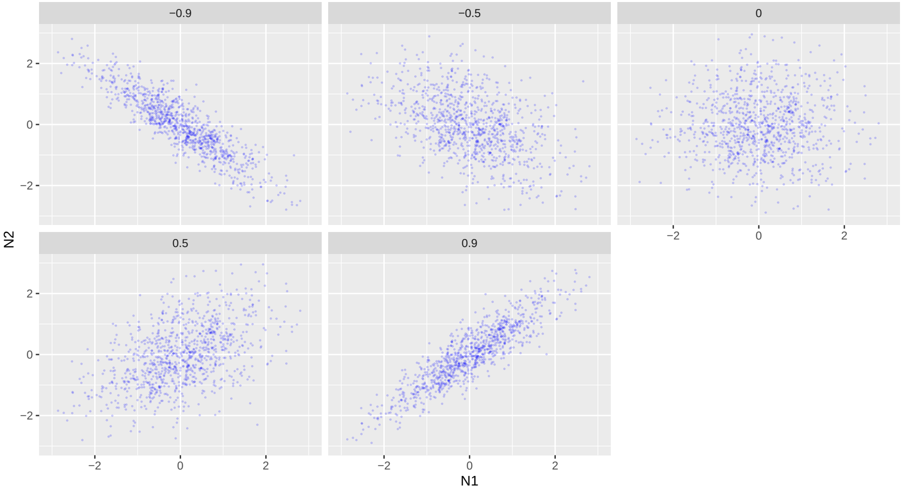
```
##      [,1]     [,2]
## [1,]    2 0.000000
## [2,]    2 2.236068
```

- Simulate $N = (N_1, N_2)$ where N_1 and N_2 are standard normal r.v and correlation between N_1 and N_2 is ρ; $\rho = -0.9, -0.5, 0, 0.5, 0.9$ respectively

MC simulation and multivariate normal distribution

```
rho<-c(-0.9,-0.5,0,0.5,0.9)
n<-1000
N<-matrix(0,2,5*n)
Z<-matrix(rnorm(2*n,0,1),2,n)
for (i in 1:5){
  C<-matrix(c(1,rho[i],rho[i],1),2,2)
  N[,((i-1)*n+1):(i*n)]<-t(chol(C)) %*% Z
}
dat<-data.frame("N1"=N[1,],"N2"=N[2,],
                gr=(((1:(5*n))-1) %/% n + 1))
dat<-mutate(dat,cor=as.factor(rho[gr]))
dat%>%ggplot(aes(N1,N2))+
  geom_point(cex=0.2,alpha=0.2,col="blue")+
  xlim(c(-3,3))+
  ylim(c(-3,3))+facet_wrap(~cor)
```

MC simulation and multivariate normal distribution



Bivariate Brownian motion

A 2-dimensional Brownian motion $B_t = (B_t^{(1)}, B_t^{(2)})$ is defined as follows:

1. $B_0 = (0, 0)$
2. $(B_{t_2}^{(i)} - B_{t_1}^{(i)})$ and $(B_{s_2}^{(j)} - B_{s_1}^{(j)})$ are independent for all $1 \leq i, j \leq 2$ and $s_1 \leq s2 \leq t_1 \leq t_2$
3. $(B_{t_2}^{(1)} - B_{t_1}^{(1)}, B_{t_2}^{(2)} - B_{t_1}^{(2)})$ is a 2-dimensional normal random vector with mean μ and covariance matrix Σ

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \; ; \Sigma = \begin{pmatrix} (t_2 - t_1) & \rho(t_2 - t_1) \\ \rho(t_2 - t_1) & (t_2 - t_1) \end{pmatrix}$$

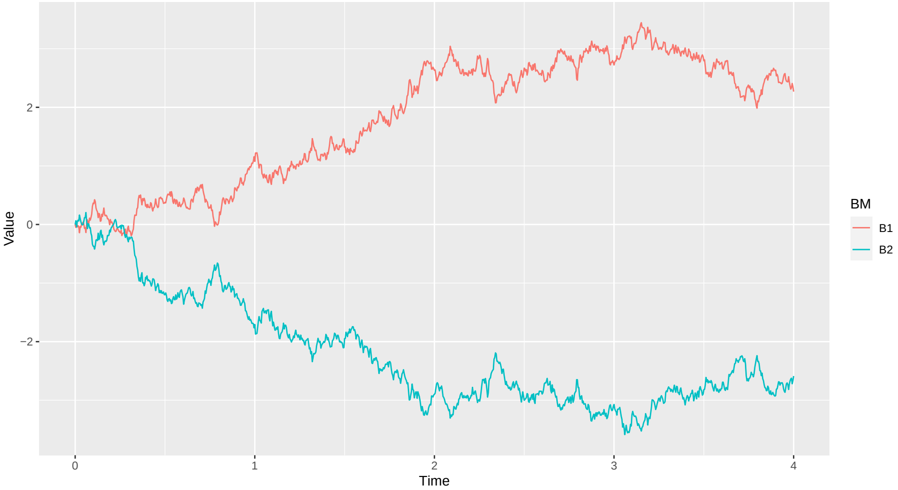
where ρ is the correlation between these Brownian motions.

Bivariate Brownian motion

Simulate a 2-dimensional Brownian motion from 0 to T with $T = 4$, $n = 1000$ and $\rho = -0.9, 0, 0.9$, respectively

```
T<-4
n<-1000
delta<-T/n
d<-2
rho<-0.9
Bt<-matrix(0,d,(n+1))
C<-t(chol(matrix(c(delta,rho*delta,rho*delta,delta),d,d)))
for (i in 2:(n+1)){
  dBt<-C%*%rnorm(d,0,1)
  Bt[,i]<-Bt[,(i-1)]+dBt
}
dat<-data.frame("B1"=Bt[1,],"B2"=Bt[2,],Time=seq(0,T,T/n))
dat<-gather(dat,BM,Value,B1:paste0("B",d))
dat%>%ggplot(aes(Time,Value,group=BM,col=BM))+geom_line()
```

Bivariate Brownian motion

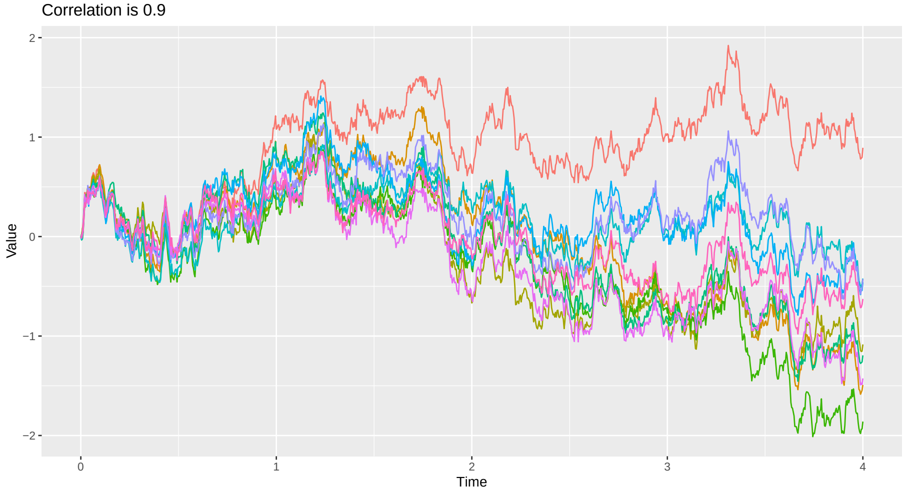


Multivariate Brownian motion

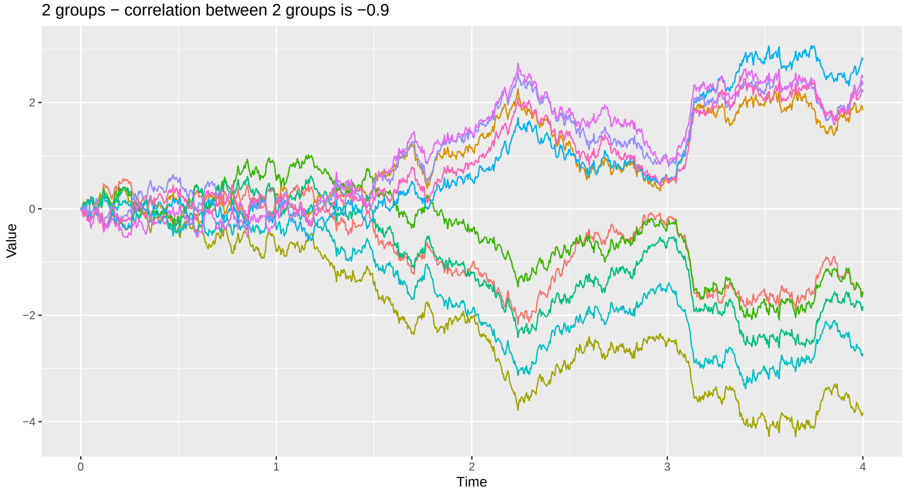
Simulate a d-dimensional Brownian motion from 0 to T with d = 10, T = 4, n = 1000 and $\rho_{ij} = 0.9$ for all $1 \leq i, j \leq k$ and $i \neq j$

```
d<-10
rho<-0.9
Bt<-matrix(0,d,(n+1))
Sigma<-matrix(rho,d,d)
Sigma<-delta*(Sigma + diag(1-rho,d,d))
C<-t(chol(Sigma))
for (i in 2:(n+1)){
  dBt<-C%*%rnorm(d,0,1)
  Bt[,i]<-Bt[,(i-1)]+dBt
}
dat<-as.data.frame(t(Bt))
names(dat)<-paste0("B",1:d)
dat<-dat%>%mutate(Time=seq(0,T,T/n))%>%
  gather(dat,BM,Value,B1:paste0("B",d))
```

Multivariate Brownian motion



Multivariate Brownian motion



Application of multivariate Brownian motion

In one factor (One Brownian motion) short rate models (Vasicek, CIR) the continuously compounded spot rates $R(t, T)$ is a function of r_t

$$\begin{aligned} P(t, T) &= A(t, T) \times exp(-B(t, T) r_t) \\ R(t, T) &= \frac{-\,log(P(t, T))}{(T - t)} \\ &= a(t, T) + r_t b(t, T) \end{aligned}$$

where $a(t, T)$ and $b(t, T)$ are deterministic functions. With $T_1 < T_2$

$$\begin{aligned} cor(R(t, T_1), R(t, T_2)) &= cor(a(t, T_1) + r_t b(t, T_1), a(t, T_2) + r_t b(t, T_2)) \\ &= 1 \quad \forall T_1 < T_2 \end{aligned}$$

$\rightarrow$ the price of 1-month ZC bond and the price of 30-years ZC bond are perfectly correlated!!!

Application of multivariate Brownian motion

- The two-factor version of Vasicek short rate model (G2 model):

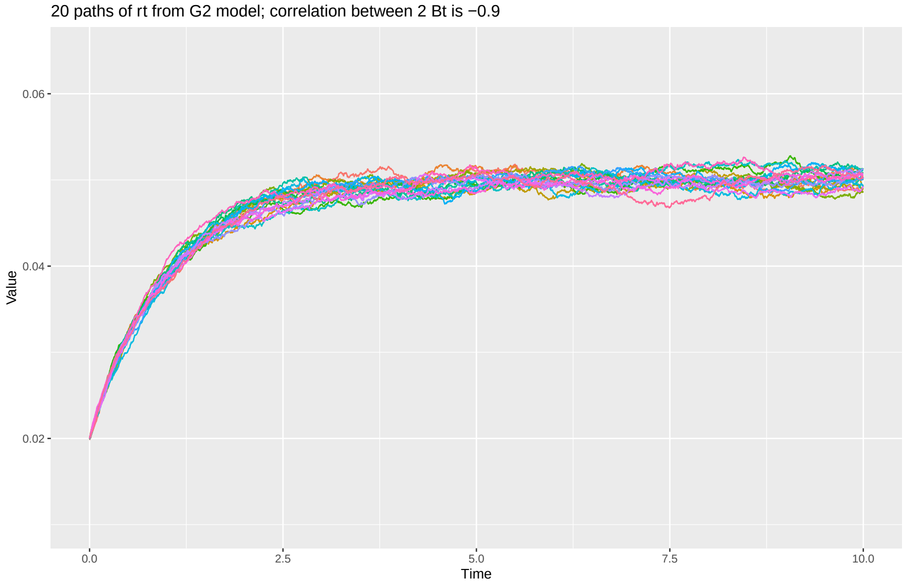
$$\begin{aligned} r_t &= x_t + y_t \\ x_t &= \kappa_x(\theta_x - x_t)dt + \sigma_x dB_t^{(1)} \\ y_t &= \kappa_y(\theta_y - y_t)dt + \sigma_y dB_t^{(2)} \end{aligned}$$

  where $(B_t^{(1)}, B_t^{(2)})$ is a 2-dimensional Brownian motion with correlation ρ

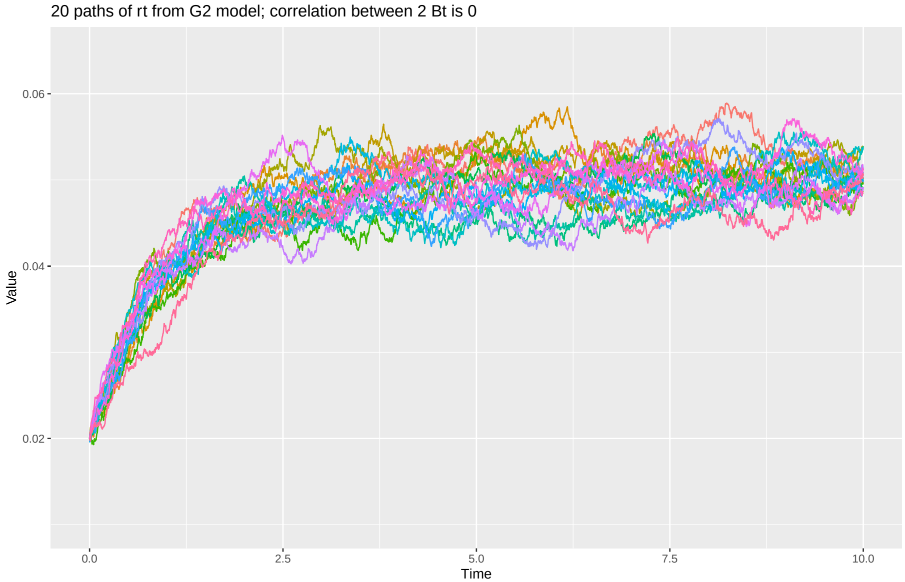
- Simulate r_t from G2 model with $\rho = -0.9, 0, 0.9,\ T = 10,\ n = 2500$ and

$$\begin{aligned} x_0 = 1\%, \kappa_x = 1, \theta_x = 2.5\%, \sigma_x = 0.25\% \\ y_0 = 1\%, \kappa_y = 1, \theta_y = 2.5\%, \sigma_y = 0.25\% \end{aligned}$$

Application of multivariate Brownian motion



Application of multivariate Brownian motion



Application of multivariate Brownian motion