Chapter 4

Generating Brownian Motions and Other Processes in Finance

Why Brownian Motion?

In finance, the stock prices are often assumed to follow (under the risk-neutral measure)

$$dS_t = rS_t dt + \sigma S_t dB_t \qquad (*)$$

where r --- risk-free interest rate

σ --- volatility

B_t --- the standard BM.

The equation (*) has the analytical solution:

$$S_t = S_0 \exp((r - \sigma^2/2)t + \sigma B_t).$$

(Geometric Brownian motion, GBM)

So if we can simulate BM, we can simulate stock prices.

History of Brownian Motion (BM)

- BM is named after the Scottish botanist Brown who first described the random motions of pollen grains suspended in water (1827).

- The mathematics of this process were formalized by Bachelier in an option pricing context (1900) and by Einstein (1905).

- A mathematically concise definition and a rigorous theory of BM, was developed by Wiener (1918), thus BM is also known as **Wiener process**.

1. Brownian Motion

Definition of a standard BM:

$(1)\, B_0 = 0$ **or** $P\{\omega \mid B_0(\omega) = 0\} = 1;$

(2) **For** $0 < t_1 < \cdots < t_n, B_{t_1} - B_0, \cdots, B_{t_n} - B_{t_{n-1}}$ **are independent**

$(3)\, B_{t+s} - B_s \sim N(0, t);$

$(4)\, B_t$ **is a continuous function of** t **on** $[0, T]$ **(with prob.** 1)

A BM X(t) with drift μ and diffusion coefficient σ^2:

if $\dfrac{X(t) - \mu t}{\sigma}$ is a standard BM.

1. Brownian Motion

- **BM plays a central role in finance.**
- **In many applications it is required to simulate BM, i.e., to simulate the values**

$$(B(t_1), B(t_2), \dots, B(t_n))^T$$

at a discrete set of points $0 < t_1 < \dots < t_n$.

The mean and covariance matrix of $(B(t_1),\ldots,B(t_n))^T$

Since B(t) ~ N(0, t), we have E(B(t))=0 for all t>0. Thus the **mean** is identically 0.

For 0 < s < t < T, we have

$$\mathbf{Cov}[B(s),B(t)] = \mathbf{Cov}(B(s),B(s)+(B(t)-B(s)))$$
$$= \mathbf{Cov}[B(s),B(s)] + \mathbf{Cov}[\underline{B(s)},\underline{B(t)-B(s)}] = s+0 = s.$$

independent

Thus the **covariance matrix** is C with

$$\boldsymbol{C}_{ij} = \min(\boldsymbol{t}_i,\boldsymbol{t}_j),\ \textbf{i.e.,}\ \boldsymbol{C} = \begin{pmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_1 & \cdots & \boldsymbol{t}_1 \\ \boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_n \end{pmatrix}.$$

Classical Constructions of BM

- Random Walk Construction (Standard Construction)
- Brownian Bridge Construction
- Principal Component Analysis Construction

2. Random Walk Construction (Standard Construction)

Because BM has **independent normally distributed increments**, thus we can generate BM sequentially

$$B(0) = 0;$$

$$B(t_{i+1}) = B(t_i) + \sqrt{t_{i+1} - t_i}\, Z_{i+1}, \quad i = 1, \ldots, n-1,$$

where $Z_1, \ldots, Z_n$ are iid $N(0, 1)$.

Note: The method **is exact** in the sense that the joint distribution of $(B(t_1), \ldots, B(t_n))^T$ **coincides** with the joint distribution of the true BM at $0 < t_1 < \ldots < t_n$.

Random Walk Construction

$$B(0) = 0;$$

$$B(t_1) = \sqrt{t_1}\, Z_1\ ,$$

$$B(t_2) = B(t_1) + \sqrt{t_2 - t_1}\, Z_2 \ = \sqrt{t_1}\, Z_1 \ + \sqrt{t_2 - t_1}\, Z_2$$

$$B(t_3) = B(t_2) + \sqrt{t_3 - t_2}\, Z_3 \ = \sqrt{t_1}\, Z_1 \ + \sqrt{t_2 - t_1}\, Z_2 \ + \sqrt{t_3 - t_2}\, Z_3$$

……

where $Z_1, \ldots, Z_n$ are iid N(0, 1). **In matrix form:**

$$\begin{pmatrix} B(t_1) \\ B(t_2) \\ \vdots \\ B(t_n) \end{pmatrix} = \begin{pmatrix} \sqrt{t_1} & 0 & \cdots & 0 \\ \sqrt{t_1} & \sqrt{t_2 - t_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{t_1} & \sqrt{t_2 - t_1} & \cdots & \sqrt{t_n - t_{n-1}} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{pmatrix}.$$

But this is slower than the one in previous page

Other Constructions?

What is the exact joint distribution of $(B(t_1),\ldots,B(t_n))^T$?

According to the definition of BM, $B_{t_1}-B_0,\cdots,B_{t_n}-B_{t_{n-1}}$ are independent and each has normal distribution, thus their joint distribution is normal.

So the joint distribution of $(B(t_1),\ldots,B(t_n))^T$ is **normal**, since it can be obtained by a linear transformation from

$$B_{t_1}-B_0, B_{t_2}-B_{t_1},\cdots,B_{t_n}-B_{t_{n-1}}.$$

We have $(B(t_1),\ldots,B(t_n))^T \sim N(0,C).$

Other Constructions?

Since

$$(B(t_1),\ldots,B(t_n))^T \sim N(0,C),$$

we may simulate it by using the method for normal variate:

$$(B(t_1),\ldots,B(t_n))^T = AZ,$$

$$\textbf{with } AA^T = C \textbf{ and } Z \sim N(0,I),$$

where A is called the **generating matrix of BM**.

Other Constructions?

The **random walk construction** is equivalent to using the **Cholesky factorization** of C:

$$A = \begin{pmatrix} \sqrt{t_1} & 0 & \cdots & 0 \\ \sqrt{t_1} & \sqrt{t_2 - t_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{t_1} & \sqrt{t_2 - t_1} & \cdots & \sqrt{t_n - t_{n-1}} \end{pmatrix}.$$

Other Constructions?

- The only requirement of generating BM is that the vector $(B(t_1),\ldots,B(t_n))^T$ has the RIGHT distribution:

$$(B(t_1),\ldots,B(t_n))^T \sim N(0,C).$$

- Any matrix A satisfies $AA^T = C$ can be used:

$$(B(t_1),\ldots,B(t_n))^T = AZ,$$

$$\textbf{with } AA^T = C \textbf{ and } Z \sim N(0,I).$$

Remark:

- Different generating matrices are equivalent in MC setting.
- But they may have quite different effect in **quasi-Monte Carlo** (**why?**).
- Thus it is important to choose good generating matrix.

- **Important question:**

  How best to choose the generating matrix?

3. Brownian Bridge Construction

We may generate the values of BM

$$B(t_1),\ldots,B(t_n)$$

in any order, provided that at each step we sample from the correct conditional distribution, given the values already generated.

We only need to guarantee that the vector

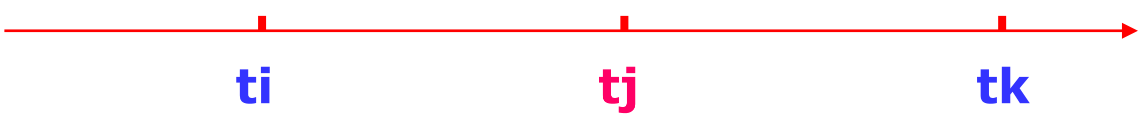
$$(B(t_1),\ldots,B(t_n))^T$$

has the **RIGHT distribution N(0, C)**.

Theorem:

Given the past value $B(t_i) = b_i$ and a future value $B(t_k) = b_k$, then $B(t_j)$ (with $i < j < k$) conditional on these two values is a normal variable $N(\mu, \sigma^2)$, where

$$\mu = \frac{t_k - t_j}{t_k - t_i} b_i + \frac{t_j - t_i}{t_k - t_i} b_k, \quad \sigma^2 = \frac{(t_k - t_j)(t_j - t_i)}{t_k - t_i}.$$



Lemma: (Conditioning Formula)

Suppose the vector (X, Y) **(X, Y themselves may be vectors)**

is multivaria te normal

$$N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\right),$$

and assumes that Σ_{22}^{-1} **exists, then**

$$\mathbf{X} \mid \mathbf{Y} = \mathbf{y} \sim N\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(y - \mu_2),\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right).$$

(seeGlasserman P.65)

Linear Transformation Property

Any linear transformation of a normal vector is again normal :

$$X \sim N(\mu, \Sigma) \Rightarrow AX \sim N(A\mu, A\Sigma A^T),$$

for any d-vector μ, and $d \times d$ matrix Σ, and any $k \times d$ matrix A, for any k.

(see Glasserman P.65)

Proof of the Theorem (outline):

$$\begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} t_1 & t_1 & t_1 \\ t_1 & t_2 & t_2 \\ t_1 & t_2 & t_3 \end{pmatrix}\right)$$

Natural order

$$\Rightarrow \begin{pmatrix} B_2 \\ B_1 \\ B_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} t_2 & t_1 & t_2 \\ t_1 & t_1 & t_1 \\ t_2 & t_1 & t_3 \end{pmatrix}\right)$$

Changed order:
后者=P12*前者,
后者协方差
= P12 *前者协方差*P12^T

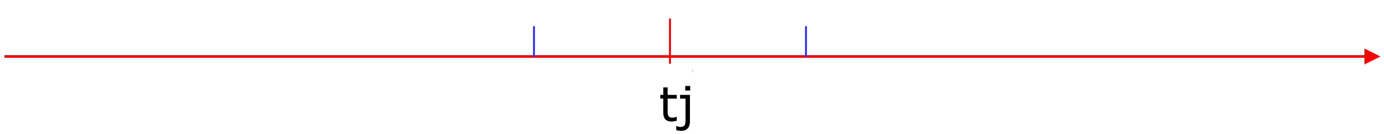
By conditioni ng formula,

$$B_2 \mid B_1 = b_1, B_3 = b_3 \sim N\left(\frac{t_3 - t_2}{t_3 - t_1} b_1 + \frac{t_2 - t_1}{t_3 - t_1} b_3,\ \frac{(t_3 - t_2)(t_2 - t_1)}{t_3 - t_1}\right).$$

3. Brownian Bridge Construction

Suppose the elements of $(B(t_1),\ldots,B(t_d))$ need to be generated in the order $(B(t_{\pi(1)}),\ldots,B(t_{\pi(d)}))$ for some permutation π of d elements.

In computing $B(t_{\pi(j)})$ we need to take into account the previously computed elemnts, and at most two of those are of relevent: the one next to $\pi(j)$ on the left and the one next to $\pi(j)$ on the right.



$$N\left(\frac{t_k - t_j}{t_k - t_i} b_i + \frac{t_j - t_i}{t_k - t_i} b_k,\ \frac{(t_k - t_j)(t_j - t_i)}{t_k - t_i}\right)$$

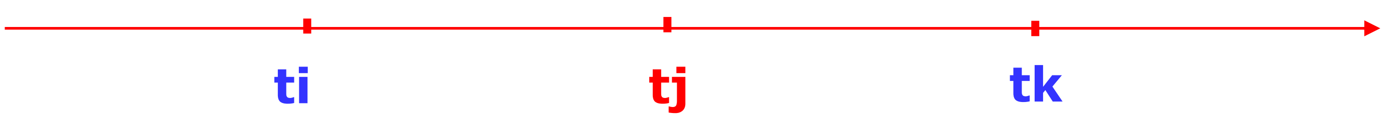
3. Brownian Bridge Construction

For **equally time interval**, given the past value B(t_i) and a future value B(t_k), the value of B(t_j) (with i<j<k) can be generated as

$$B_{t_j} = (1-\rho)B_{t_i} + \rho B_{t_k} + \sqrt{\rho(1-\rho)(k-i)\Delta t}\, Z,$$

$$Z \sim N(0,1),\quad \rho = (j-i)/(k-i),\quad t_i < t_j < t_k$$

The Brownian bridge formula provides new ways to generate BM.



(for n=power of 2)

We first generate the end point B_T, then the middle point $B_{T/2}$, …

$$
\begin{aligned}
B_T &= \sqrt{T}\, z_1; \\
B_{T/2} &= \frac{1}{2}(B_0 + B_T) + \sqrt{\frac{T}{4}}\, z_2 = \frac{\sqrt{T}}{2}\, z_1 + \frac{\sqrt{T}}{2}\, z_2; \\
B_{T/4} &= \frac{1}{2}(B_0 + B_{T/2}) + \sqrt{\frac{T}{8}}\, z_3 = \frac{\sqrt{T}}{4}\, z_1 + \frac{\sqrt{T}}{4}\, z_2 + \sqrt{\frac{T}{8}}\, z_3; \\
B_{3T/4} &= \frac{1}{2}(B_{T/2} + B_T) + \sqrt{\frac{T}{8}}\, z_4 = \frac{3\sqrt{T}}{4}\, z_1 + \frac{\sqrt{T}}{4}\, z_2 + \sqrt{\frac{T}{8}}\, z_4; \\
&\vdots
\end{aligned}
$$

$$
\begin{aligned}
B_T &= \sqrt{T}\, z_1; \\
B_{T/2} &= \frac{1}{2}(B_0 + B_T) + \sqrt{\frac{T}{4}}\, z_2 = \frac{\sqrt{T}}{2}\, z_1 + \frac{\sqrt{T}}{2}\, z_2;
\end{aligned}
$$

$$
\begin{aligned}
B_{T/4} &= \frac{1}{2}(B_0 + B_{T/2}) + \sqrt{\frac{T}{8}}\, z_3 = \frac{\sqrt{T}}{4}\, z_1 + \frac{\sqrt{T}}{4}\, z_2 + \sqrt{\frac{T}{8}}\, z_3; \\
B_{3T/4} &= \frac{1}{2}(B_{T/2} + B_T) + \sqrt{\frac{T}{8}}\, z_4 = \frac{3\sqrt{T}}{4}\, z_1 + \frac{\sqrt{T}}{4}\, z_2 + \sqrt{\frac{T}{8}}\, z_4; \\
&\vdots
\end{aligned}
$$

Question:

What is the generating matrix A corresponding to BB?

Natural order

$$
\begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{pmatrix} = A \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}.
$$

3. Brownian Bridge Construction

(for n= 2^M)

$$\text{Step 1. } B_0 = 0, h = T, B_T = \sqrt{T}\, Z, \;\; Z \sim N(0,1);$$

$$\text{Step 2: For } k = 1 \text{ to } M, \text{ do}$$

$$\text{(a) Set } h = h/2;$$

$$\text{(b) For } j = 1 \text{ to } 2^{k-1}, \text{ do}$$

$$\text{i. Simulate a new } Z \sim N(0,1),$$

$$\text{ii. } B_{(2j-1)h} = \frac{1}{2}\left(B_{2(j-1)h} + B_{2jh}\right) + \sqrt{h/2}\, Z.$$

Remarks

- The random walk and BB algorithms are *exact*, in that the $\{B_i\}$ are drawn exactly according to their respective distributions.
- Nevertheless, these algorithms returns only a discrete skeleton of the true continuous process.
- To obtain a continuous path approximation to the exact path of the Wiener process, one could use linear interpolation on the points obtained.
- It is possible to adaptively refine the path by using a Brownian bridge method.

3. Brownian Bridge Construction

- In the BB construction:
  - A single normal variable Z_1 determine the end point of the path;
  - Conditional on the end point, a second normal point Z_2 determines the midpoint of the path,
  - and so on.

- Thus much of ultimate shape of BM is determined or explained by the values of **just a few normals**.

3. Brownian Bridge Construction

Questions:

- What is the "optimal" order to construct BM?
- In what sense?
- Is there a construction (among all constructions) that maximizes the variability of the path explained by $Z_1, \ldots, Z_k$ for all $k=1, \ldots, n$?

4. Principal Component Analysis Construction

What is PCA construction?

PCA construction takes the generating matrix as

$$A^{PCA} = V D^{1/2} = (\sqrt{\lambda_1}\, v_1, \cdots, \sqrt{\lambda_n}\, v_n)$$

where

$$D = \mathbf{diag}(\lambda_1, \cdots \lambda_n),$$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ **are the eigenvalues of C** ,

$V = (v_1, v_2, \cdots, v_n)$,

v_i **is the eigenvectors of unit length of** λ_i .

4. PCA Construction

In Matlab:

EIG Eigenvalues and eigenvectors:

- **E = EIG(A) is a vector containing the eigenvalues of a square matrix A.**

- **[V,D] = EIG(X) produces a diagonal matrix D of eigenvalues and a full matrix V whose columns are the corresponding eigenvectors so that**

  A*V = V*D.

4. PCA Construction

What is PCA construction?

PCA construction takes the generating matrix as

$$A^{PCA} = V\,D^{1/2} = (\sqrt{\lambda_1}\,v_1, \cdots, \sqrt{\lambda_n}\,v_n)$$

where

$$D = \mathbf{diag}(\lambda_1, \cdots \lambda_n),$$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ **are the eigenvalues of C**,

$$V = (v_1, v_2, \cdots, v_n),$$

v_i **is the eigenvectors of unit length of** λ_i **.**

4. PCA Construction

Is PCA construction a right construction?

Since the covariance matrix C is real symmetric and positive definite (or semidefinite),

it has n positive (or nonnegative) eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

and has associated orthonormal set of eigenvectors

$$v_1, \cdots, v_n$$

satisfying $Cv_i = \lambda_i v_i, i = 1, \ldots, n,$

$$v_i^T v_i = 1, v_i^T v_j = 0.$$

4. PCA Construction

From $Cv_i = \lambda_i v_i, \ i = 1, \ldots, n,$

$$v_i^T v_i = 1, \ v_i^T v_j = 0,$$

it follows that

$$CV = VD$$

$$\Rightarrow C = VDV^T = (VD^{1/2})(VD^{1/2})^T = A^{PCA}\left(A^{PCA}\right)^T,$$

implying that **PCA is a valid (right) construction.**

Note:

$V = (v_1, \cdots, v_n)$ is an orthogonal matrix: $VV^T = I.$

4. PCA Construction

What is the benefit of using PCA construction?

- **PCA provides an "optimal" lower-dim approximation to a random vector.**

- **PCA maximizes the variability of the path explained by $Z_1, \ldots, Z_k$ for all $k=1, \ldots, n$.**

Connection with Principal Component in statistics

Let $Y := (B_{t_1}, \cdots, B_{t_n})^T = A^{PCA} Z.$

We have

$$Z = (A^{PCA})^{-1} Y = (VD^{1/2})^{-1} Y = D^{-1/2} V^{-1} Y = D^{-1/2} V^T Y = \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} v_1^T \\ \vdots \\ \frac{1}{\sqrt{\lambda_n}} v_n^T \end{pmatrix} Y,$$

$$\Rightarrow Z_j = \frac{1}{\sqrt{\lambda_j}} v_j^T Y \qquad \textbf{(a linear combination of } Y\textbf{)}$$

Z_j is the (unit variance) **j-th principal component** of Y (as defined in statistical literature). Why?

It turns out that Zj is optimal in a sense defined below.

4. PCA Construction

Let *Y* be a random vector with covariance matrix C. Suppose we want to find the best linear combination $b^T Y$ that best captures the variability of the component of Y (i.e., has the largest variance among all b with ||b||=1):

$$\max_{\|b\|=1} b^T C\, b$$

This problem is solved by $b = v_1$.

The max. value = $v_1^T C v_1 = v_1^T \lambda_1 v_1 = \lambda_1$.

The linear combination of $v_1^T Y$ is called the first **principal component of Y.** The unit variance version is $\frac{1}{\sqrt{\lambda_1}} v_1^T Y$

To solve $$\max_{b\in R^n} b^T C\, b$$

$$s.t. \| b \| = 1$$

We introduce Lagrange multiplier λ

$$\frac{\partial}{\partial b_i}[b^T C b + \lambda(1 - b^T b)] = 0$$

$$\Rightarrow C\, b = \lambda\, b$$

$\Rightarrow \lambda$ **is an eigenvalue of C and b is the eigenvector.**

Since $b^T C b = b^T \lambda b = \lambda b^T b = \lambda$, **thus**

the solution is the first eigenvalue λ_1 ,

and the corresponding eigenvector is v_1.

Choose the largest

4. PCA Construction

- The problem of finding **the next best** vector orthogonal to the first one reduces to

$$\max_{||b||=1,\, b^T v_1 = 0} b^T C\, b$$

This is solved by $b = v_2$. The max. value = λ_2

- **Definition:**

$v_2^T Y$ is called the **second PC of Y.**

The unit variance version is $\dfrac{1}{\sqrt{\lambda_2}} v_2^T Y$.

To solve $$\max_{b \in R^n} b^T C\, b$$

$$s.t. \|b\| = 1, v_1^T b = 0$$

We introduce Lagrange multiplier λ, η

$$\frac{\partial}{\partial b_i}[b^T C b + \lambda(1 - b^T b) + \eta v_1^T b] = 0$$

$$\Rightarrow C\, b - \lambda\, b - \eta v_1 = 0 \qquad (*)$$

Multiply both sides on the left by v_1^T :

$$v_1^T C\, b - \lambda\, v_1^T b - \eta v_1^T v_1 = 0$$

(the first two terms are zeros : $v_1^T C\, b = \lambda_1 v_1^T b = 0$)

$$\Rightarrow \eta v_1^T v_1 = 0 \Rightarrow \eta = 0$$

Again, (*) implies that λ is an eigenvalue of C and b is the corresponding eigenvector....

Choose the 2nd largest

4. PCA Construction

- Proceeding inductively in this way, we obtain

$$\lambda_{k+1} = \max_{\|b\|=1, b^T v_j = 0, j=1,\dots,k} b^T C\, b$$

The maximaizer of the right hand side is the eivenvector v_{k+1} corresponding to λ_{k+1}.

- **Definition:**

$v_j^T Y$ is called the **j-th PC of Y**

The unit variance version is $\dfrac{1}{\sqrt{\lambda_j}} v_j^T Y$.

An alternative understanding of the optimality of PCA

Consider

$$Y = AZ = (a_1,\ldots,a_n)\begin{pmatrix} Z_1 \\ \vdots \\ Z_n \end{pmatrix}$$

$$= a_1 Z_1 + \cdots + a_n Z_n,\ \ Z = (Z_{1,} \cdots, Z_n)^T \sim N(0, I)$$

where a_i is the i-th column of generating matrix A with $AA^T = C.$

Can we use several normals to approximate well the vector Y ?

An alternative understanding of the optimality of PCA

The approximation error

$$E\left[\| Y - \sum_{i=1}^{k} a_i Z_i \|^2 \right] = \| a_{k+1} \|^2 + \cdots + \| a_n \|^2$$

is minimized for all k =1, …, n for PCA construction, i.e., with

$a_i = \sqrt{\lambda_i}\, v_i$ (i-th column of the generating matrix in PCA)

$Z_i = \frac{1}{\sqrt{\lambda_i}} v_i^T Y$ (the normalized i-th PC of Y)

Thus the PCs provide an optimal lower-dimensional approximation to a random vector.

Let *Y* be a RV with covariance matrix C

$$\mathrm{Tr}\,(\mathrm{C}) = \sum_{i=1}^{n} \mathrm{C_{ii}} = \sum_{i=1}^{n} \mathrm{Var}(\mathrm{Y_i}).$$

On the other hand, $\mathrm{Tr}\,(\mathrm{C}) = \lambda_1 + \cdots + \lambda_n$

$$\Rightarrow \lambda_1 + \cdots + \lambda_n = \sum_{i=1}^{n} \mathrm{Var}(\mathrm{Y_i})$$

An important goal of PCA is to determine if the first few PCs

can account for most of the overall variance $\sum_{i=1}^{n} \mathrm{Var}(\mathrm{Y_i})$.

This account to determine whether

$$(\lambda_1 + \cdots + \lambda_k)/(\lambda_1 + \cdots + \lambda_n)$$

is near 1 for some small k.

4. PCA Construction

Given a generating matrix A, consider the expression (a_j is the j-th column of A)

$$Y = AZ = a_1 Z_1 + \cdots + a_n Z_n, \; Z = (Z_1, \cdots, Z_n)^T \sim N(0, I)$$

Definition: The **variance explained** by the first k normals is defined as

$$\frac{\| a_1 \|^2 + \cdots + \| a_k \|^2}{\| a_1 \|^2 + \cdots + \| a_n \|^2} = \frac{\| a_1 \|^2 + \cdots + \| a_k \|^2}{\operatorname{tr}(C)}$$

4. PCA Construction

In PCA construction,

$$A^{PCA} = V D^{1/2} = (\sqrt{\lambda_1}\, v_1, \cdots, \sqrt{\lambda_n}\, v_n)$$

the **variance explained** by the first k normals is

$$\frac{\lambda_1 + \cdots + \lambda_k}{\lambda_1 + \cdots + \lambda_k + \cdots + \lambda_n}$$

The first PC is chosen to explain as much variance as possible.

In general, the eigenvalues decay rapidly, thus the first few normals capture most of the variability.

The eigenvalues for 16-D matrix C (of BM with T=1)

1. 6.901383755994658e+000
2. 7.714710949798159e-001
3. 2.811147944365275e-001
4. 1.460642768467739e-001
5. 9.054316061229008e-002
6. 6.249999999999987e-002
7. 4.643856778472944e-002
8. 3.643528114335882e-002
9. 2.983060277002503e-002
10. 2.528810361150039e-002
11. 2.207833018085853e-002
12. 1.977776864282738e-002
13. 1.812924691929490e-002
14. 1.697213308367630e-002
15. 1.620541363759517e-002
16. 1.576746935606724e-002

The eigenvalues for matrix C (analytical formula for equal time interval)

For the covariance matrix $\mathbf{V}$ in (5), it is possible to calculate its eigenvalues and the corresponding eigenvectors. The inverse of $\mathbf{V}$ is given by

$$\mathbf{V}^{-1} = \Delta t^{-1} \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

The eigenvalues of $\mathbf{V}^{-1}$ are

$$\beta_j := 4\Delta t^{-1} \sin^2 h_j \quad \text{with } h_j := \frac{(2j-1)\pi}{2(2d+1)}, \quad j = 1, \ldots, d.$$

The unit-length eigenvector corresponding to the eigenvalue β_j is

$$\frac{2}{\sqrt{2d+1}} (\sin(2h_j), \sin(4h_j), \ldots, \sin(2dh_j))^T. \tag{15}$$

Therefore, the eigenvalues of $\mathbf{V}$ are (in a decreasing order)

$$\lambda_j = \frac{1}{\beta_j} = \frac{\Delta t}{4 \sin^2 h_j}, \quad j = 1, \ldots, d.$$

Note that the largest eigenvalue is $O(d)$, while the smallest one is $O(d^{-1})$. The unit-length eigenvector corresponding to λ_j is given by (15). Thus, the sum of the jth column of

See: Wang (Operations Research, 2006)

5. Comparison of Different Constructions

Note: Each construction corresponds to a method of decomposing the covariance matrix C.

- Which construction should be used in practice?

  In MC setting, all constructions are equivalent (in probabilistic sense).

  However, in Quasi-Monte Carlo, their efficiency can be quite different!

- See: Wang & Fang (J. Complexity, 2003)
  Sobol' (Monte Carlo Methods Appl., 2005)
  Wang (Operations Research, 2006)

5. Comparison of Different Constructions

- **Cumulative explained variability** from the first five dimensions, based on **n = 64**

Construction	Cumulative explained variability (%)				
standard	3.1	6.1	9.1	12.0	14.9
Brownian Bridge	67.2	83.6	87.7	91.8	92.9
PCA	81.6	90.1	93.3	95.0	96.0

Standard method is sensitive with respect to the dimension n, whereas BB and PCA are robust.

5. Comparison of Different Constructions

- **Cumulative explained variability** from the first five dimensions, based on **n = 256**

Construction	Cumulative explained variability (%)				
standard	0.8	1.6	2.3	3.1	3.9
Brownian Bridge	66.8	83.4	87.6	91.7	92.7
PCA	81.1	90.1	93.3	95.0	96.0

Standard method is sensitive with respect to the dimension n, whereas BB and PCA are robust.

Remarks

- The superiority of BB and PCA or other methods (in the sense of explained variability) does not imply their superiority for practical problems.
- Indeed, no method provides a consistent advantage in QMC. A method works well for one problem can work badly for another.

- See Wang & Sloan (Operations Research, 2011); Wang & Tan (Management Science, 2013).
- **Big Problem:**

  How to find a good or optimal method in practice?

5. Comparison of Different Constructions

- **Comparisons on Computational cost:**

- Random walk: O(n);
- BB: O(n);
- PCA: O(n^3);

So in MC, just use random walk construction.

But in QMC, much more care is needed, since they can affect

- the effective dimension (the number of important variables)
- the smoothness property of the function.

Both are crucial for QMC.

6. Connection of Different Constructions

Theorem: Let A be a dxd matrix satisfying $\mathbf{AA}^{\mathrm{T}} = \mathbf{C}$.
Then $\mathbf{BB}^{\mathrm{T}} = \mathbf{C}$, if and only if B can be written as **B=AU** for some orthogonal matrix **U** ($\mathbf{UU}^{\mathrm{T}} = \mathbf{I}$).

Proof. If B=AU, and U is an orthogonal matrix, then

$$\mathbf{BB}^{\mathrm{T}} = \mathbf{AUU}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \mathbf{AA}^{\mathrm{T}} = \mathbf{C}.$$

On the other hand, suppose $\mathbf{BB}^{\mathrm{T}} = \mathbf{C}$.
Let $\mathbf{U} = \mathbf{A}^{-1}\mathbf{B}$. Then **B=AU**, and from

$$\mathbf{AA}^{\mathrm{T}} = \mathbf{C} = \mathbf{BB}^{\mathrm{T}} = \mathbf{A}\underline{\mathbf{UU}^{\mathrm{T}}}\mathbf{A}^{\mathrm{T}},$$

we have $\mathbf{UU}^{\mathrm{T}} = \mathbf{I}$.
(We may directly verify that $\mathbf{UU}^{\mathrm{T}} = \mathbf{I}$ using $\mathrm{U} = \mathrm{A}^{-1}\mathrm{B}$)

6. Connection of Different Constructions

To find a good or optimal construction:

- Fix an initial matrix A, such that AA^T=C;
- Find a good or optimal orthogonal matrix, such that UU^T =I;
- Let B= AU.

7. Multiple Dimensions

- The preceding discussion concerns the construction of a single, scalar Brownian motion.
- Suppose now that we have to generate a d-dim Brownian motion with correlation matrix between the different components.
- What do we do?

7. Multiple Dimensions

We call a process $W(t) = (W_1(t),\dots,W_d(t))^T, 0 \le t \le T$ a **standard d-dimensional BM**, if

(1) $W(0) = 0$;

(2) For $0 < t_1 < \cdots < t_n, W(t_1) - W(0), \cdots, W(t_n) - W(t_{n-1})$ are independent;

(3) $W(t) - W(s) \sim N(0, (t-s)\, I_d)$;

(4) $W(t)$ has continuous sample paths.

Remark:

Each component is a standard one-dim BM, and Wi and Wj are independent for $i \neq j$.

7. Multiple Dimensions

Let $\boldsymbol{\mu} \in \boldsymbol{R}^d$, $\boldsymbol{\Sigma}$ is a d x d matrix.

A process X(t) is called

a BM with drift μ and covariance Σ,

if X has continuous sample paths and independent increments with

$$X(t) - X(s) \sim N((t-s)\mu, (t-s)\Sigma).$$

It is denoted as $X \sim \mathbf{BM}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

7. Multiple Dimensions

When $W \sim BM(0,I)$ on R^k, and B is a $d \times k$ matrix satisfying $BB^T = \Sigma$, then the process

$$X(t) = \mu t + B\, W(t) \sim BM(\mu, \Sigma).$$

This process solves the SDE

$$dX(t) = \mu\, dt + B\, dW(t).$$

7. Multiple Dimensions

BM with deterministic and time-varying $\mu(t)$ and $\Sigma(t)$: Solution to the SDE

$$dX(t) = \mu(t)dt + B(t)dW(t),$$

where $B(t)B^T(t) = \Sigma(t)$.

The process X(t) has continuous sample paths, independent increaments with

$$X(t) - X(s) \sim N\left(\int_s^t \mu(u)du,\ \int_s^t \Sigma(u)du\right)$$

Construction of BM(0, I)

- We want to construct

$$W_1(t_1), W_1(t_2), \ldots, W_1(t_n)$$

$$W_2(t_1), W_2(t_2), \ldots, W_2(t_n)$$

$$\cdots\cdots$$

$$W_d(t_1), W_d(t_2), \ldots, W_d(t_n)$$

- **Method:**

For each dimension W_j, we may use any method for one-dim BM (random walk, BB, or PCA).

7. Multiple Dimensions

- This leads to the dimension-by-dimension construction for BM(0, I):

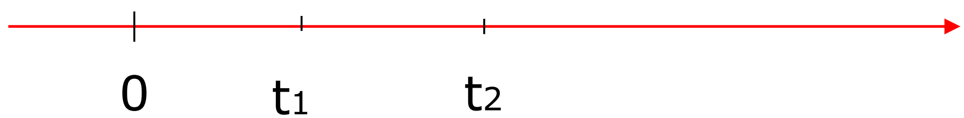
  For each dimension W_j, $j=1, \ldots, d$,
  we may use any of the construction for one-dim BM (random walk, BB, PCA).

- An alternative is the "time-by-time vector construction".

Time-by-time construction
(vector version of random walk construction)

Because d-dim Brownian motion BM(0, I) has **independent normally distributed increments**, thus we can generate BM as (time-by-time)

$$W(0) = 0;$$



$$W(t_{i+1}) = W(t_i) + \sqrt{t_{i+1} - t_i}\; Z_{i+1}, \quad i = 1, \ldots, n-1,$$

where $Z_1, \ldots, Z_n$ are iid **N(0, I)** vector.

- **Time-by-time BB construction** is also possible.

7. Multiple Dimensions

- **Construction of** $X \sim \mathrm{BM}(\mu, \Sigma)$

  We want to construct

$$\mathbf{X_1(t_1),\ X_1(t_2),\ldots,\ X_1(t_n)}$$

$$\mathbf{X_2(t_1),\ X_2(t_2),\ldots,\ X_2(t_n)}$$

$$\mathbf{\cdots\cdots}$$

$$\mathbf{X_d(t_1),\ X_d(t_2),\ldots,\ X_d(t_n).}$$

Method: Construction of $X \sim \mathrm{BM}(\mu, \Sigma)$

- **Step 1: Construct $\mathbf{W} \sim \mathbf{BM}(0, I)$ (using RW, BB or PCA):**

$$\mathbf{W_1(t_1), W_1(t_2), \ldots, W_1(t_n)}$$
$$\mathbf{W_2(t_1), W_2(t_2), \ldots, W_2(t_n)}$$
$$\cdots\cdots$$
$$\mathbf{W_d(t_1), W_d(t_2), \ldots, W_d(t_n)}$$

- **Step 2:** Choose a decomposition for $\Sigma = BB^T$ (Cholesky or PCA decomposition), and for each time step t_j, use the transformation

$$X(t) = \mu t + BW(t).$$

7. Multiple Dimensions

In Step 1, we may use RW, BB or PCA construction. This corresponds to a method of decompose matrix C:

$$C = \begin{pmatrix} t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & \cdots & t_2 \\ \vdots & \vdots & \ddots & \vdots \\ t_1 & t_2 & \cdots & t_d \end{pmatrix}.$$

In Step 2, we may use Cholesky or PCA decomposition for the matrix Σ .

There are 3X2=6 methods to generate multi-dim BM.

- **Other methods? Better methods?**

Other methods?

- Yes! Many! An important one is "One step PCA".
  Consider a long vector (for the case $\mu = 0$)

$$\boldsymbol{Y} := (\underline{\boldsymbol{X}_1(\boldsymbol{t}_1),\ldots,\boldsymbol{X}_d(\boldsymbol{t}_1)},\ \underline{\boldsymbol{X}_1(\boldsymbol{t}_2),\ldots,\boldsymbol{X}_d(\boldsymbol{t}_2)},\cdots,\boldsymbol{X}_1(\boldsymbol{t}_n),\ldots,\boldsymbol{X}_d(\boldsymbol{t}_n))^T$$

Time-by-time arrangement

- Its covariance matrix is the **Kronecker product**:

$$(\boldsymbol{C} \otimes \Sigma) = \begin{pmatrix} \boldsymbol{C}_{11}\Sigma & \boldsymbol{C}_{12}\Sigma & \cdots & \boldsymbol{C}_{1\boldsymbol{n}}\Sigma \\ \boldsymbol{C}_{21}\Sigma & \boldsymbol{C}_{22}\Sigma & \cdots & \boldsymbol{C}_{2\boldsymbol{n}}\Sigma \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{C}_{\boldsymbol{n}1}\Sigma & \boldsymbol{C}_{\boldsymbol{n}2}\Sigma & \cdots & \boldsymbol{C}_{\boldsymbol{nn}}\Sigma \end{pmatrix}.$$

This follows from $\mathbf{Cov}\,(X_i(s), X_j(t)) = \min(s,t)\boldsymbol{\Sigma}_{ij}.$

Note:

$$X(s) := (X_1(s), \ldots, X_d(s))^T \sim N(s\boldsymbol{\mu}, s\boldsymbol{\Sigma}).$$

For s < t,

$$\textbf{Cov}\,(X_i(s), X_j(t))$$

$$= \textbf{Cov}\,(X_i(s), X_j(s) + X_j(t) - X_j(s))$$

$$= \textbf{Cov}\,(\underline{X_i(s), X_j(s)}) + \textbf{Cov}\,(X_i(s), X_j(t) - X_j(s))$$

$$= s\boldsymbol{\Sigma}_{ij}$$

$\boxed{\textbf{independent}}$

7. Multiple Dimensions

- **We can generate Y as:**

$$Y = LZ, \text{where } \mathrm{LL}^{\mathrm{T}} = (C \otimes \Sigma), Z \sim N(0, I_{nd}).$$

- **Decompose the matrix** $\left(C \otimes \Sigma\right)$ **by PCA:**

$$\left(C \otimes \Sigma\right) = DD^T$$

- **Then the "long vector"**

$$\boldsymbol{Y} := (\boldsymbol{X}_1(\boldsymbol{t}_1), \ldots, \boldsymbol{X}_{\boldsymbol{d}}(\boldsymbol{t}_1), \boldsymbol{X}_1(\boldsymbol{t}_2), \ldots, \boldsymbol{X}_{\boldsymbol{d}}(\boldsymbol{t}_2), \cdots, \boldsymbol{X}_1(\boldsymbol{t}_{\boldsymbol{n}}), \ldots, \boldsymbol{X}_{\boldsymbol{d}}(\boldsymbol{t}_{\boldsymbol{n}}))^{\boldsymbol{T}}$$

can be generated as

$$Y = DZ, Z \sim N(0, I_{nd}).$$

- **For** $\mu \neq 0,$ **use**

$$X(t) = \mu t + \overline{X}, \quad \overline{X} \sim BM(0, \boldsymbol{\Sigma}).$$

7. Multiple Dimensions

- **An alternative way:**

  Instead of

  Time-by-time arrangement

$$Y := (X_1(t_1),\ldots,X_d(t_1),X_1(t_2),\ldots,X_d(t_2),\cdots,X_1(t_n),\ldots,X_d(t_n))^T$$

we may consider

$$\overline{Y} := (\underline{X_1(t_1),\ldots,X_1(t_n)},\underline{X_2(t_1),\ldots,X_2(t_n)},\cdots,\underline{X_d(t_1),\ldots,X_d(t_n)})^T$$

dimension-by-dimension arrangement

- **How to generate $\overline{Y}$? Please try!**

8. Generating Geometric BM (GBM)

Suppose that

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

where W_t is the standard BM, then S_t follows GBM.

Based on Ito Lemma, we have

$$S_t = S_0 \exp((\mu - \sigma^2 / 2)t + \sigma W_t).$$

Thus to generate GBM, we only need to generate BM using the method we studied.

Unfortunately, this is not true for most models.

8. Generating GBM

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

An alternative generation of GBM is based on the discretization

$$\frac{S(t+\Delta t) - S_t}{S_t} = \mu \Delta t + \sigma \Delta W_t$$

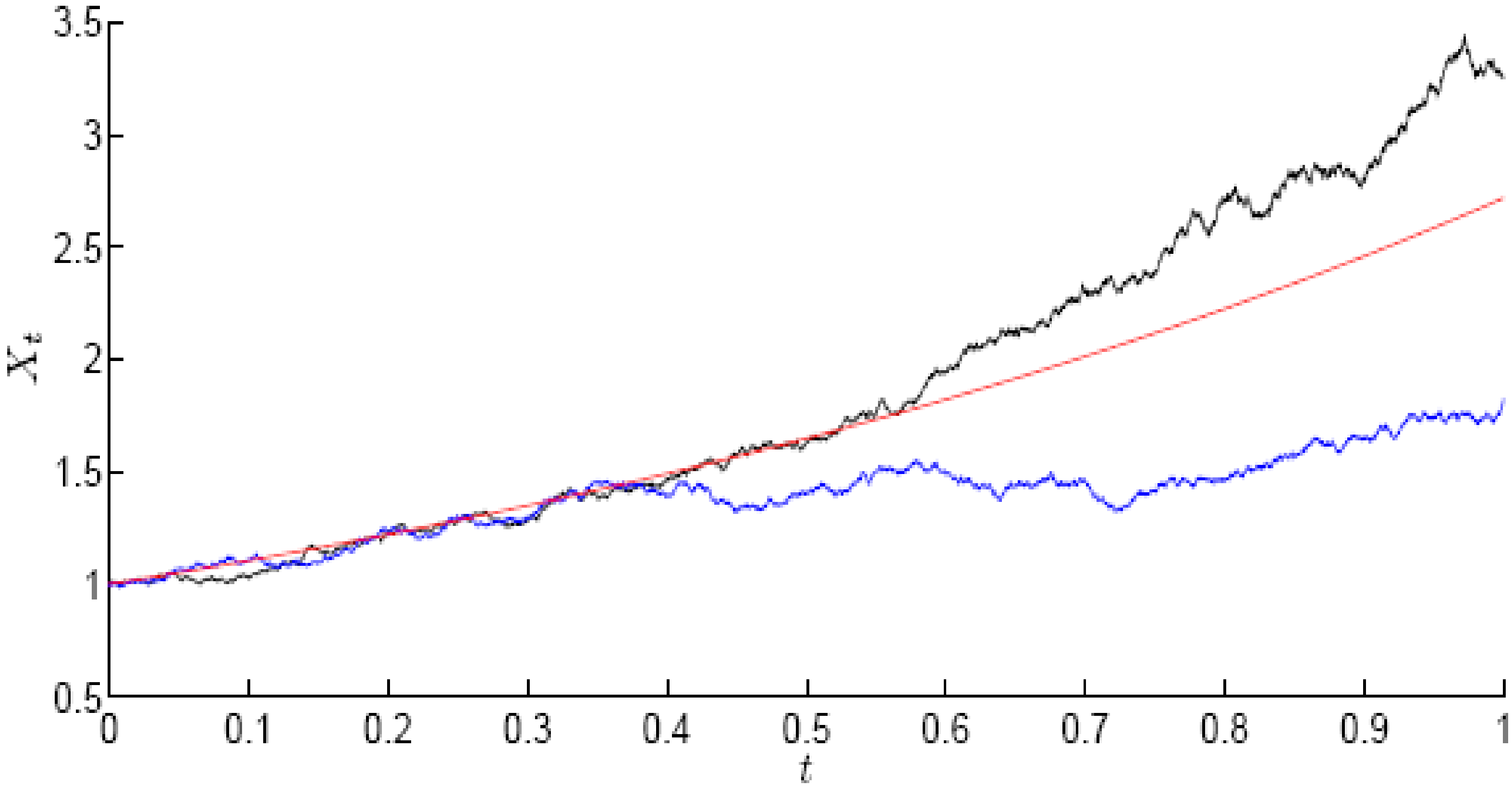
or $$S(t+\Delta t) = S_t + S_t(\mu \Delta t + \sigma \sqrt{\Delta t} Z), Z \sim N(0,1).$$

But this will lead to discretization error, thus it is not usually used.

When the model is more complicated, one has to use discretization method (later).

8. Generating GBM

Two realizations of GBM:



8. Generating GBM--- multiple dimension

Suppose that

$$dS_j(t) = \mu_j S_j(t)dt + \sigma_j S_j(t)dX_j(t),\ j = 1,\dots,d,$$

where each X_j is a standard BM, and X_i and X_j have correlation ρ_{ij}. Then

$$(X_1(t),\dots,X_d(t)) \sim BM(0,\Sigma)\ \textbf{with}\ \Sigma_{ij} = \rho_{ij}.$$

Based on Ito Lemma, we have

$$S_j(t) = S_j(0)\exp((\mu_j - \sigma_j^2/2)t + \underline{\sigma_j X_j(t)}).$$

Thus to generate multiple GBM, we only need to generate multi-dimensional BM $X(t)=(X_i(t),\dots, X_j(t))$,

9. Applications: Option Pricing

- **Risk-neutral valuation principle**:

  Assume the payoff function of an option is (path-dependent)

$$g(S_{t_1},\cdots,S_{t_n}).$$

Then the value of the option at t=0 is:

$$\mathbf{Price} = \mathbf{e^{-rT}E_Q[g(s_{t_1},\cdots,s_{t_n})]}.$$

Option price is a Mathematical Expectation.

9. Applications: Option Pricing

When MC is used to value European stock options, this involves the following steps:

1. **Simulate 1 path for the stock price in a risk neutral world, say, according to**

$$dS_t = rS_t dt + \sigma S_t dW_t.$$

2. Calculate the payoff from the stock option;
3. Repeat steps 1 and 2 many times to get many sample payoff;
4. Calculate mean payoff;
5. Discount mean payoff at risk free rate to get an estimate of the value of the option.

9. Applications: Option Pricing

The MC approximation for the option price is:

$$\mathbf{Price} \approx e^{-rT} \frac{1}{N} \sum_{k=1}^{N} g(S_{t_1}^k, \ldots, S_{t_n}^k) =: Q_N$$

$$\textbf{Error : an unbasied estimate for standard error of } \underline{Q_N} \textbf{ is}$$

$$\bar{\sigma} = \sqrt{\frac{1}{\underline{N(N-1)}} \sum_{k=1}^{N} \left[e^{-rT} g(S_{t_1}^k, \ldots, S_{t_n}^k) - Q_N \right]^2}.$$

$$\textbf{The empirical 95\% confidence interval :}$$

$$[Q_N - 1.96\,\bar{\sigma}, Q_N + 1.96\,\bar{\sigma}].$$

This is a practical tool for path-dependent options.

Summary

Step 1 : Simulating the path of the underlying $S^{(k)}$ by MC or QMC according to the dynamics of the underlying ,

Step 2 : Calculating the payoff $g(S^{(k)})$;

Step 3 : Repeating Steps 1 and 2 many times (N), calculate the mean and then discount :

$$\text{Value} \approx e^{-rT}\frac{1}{N}\sum_{k=1}^{N} g(S^{(k)}).$$

It is customary to report both the estimate and the standard error.

Example: Asian Options

- Asian options payoffs depend on the average of the underlying asset prices during the option life.
- Asian options are popular in financial industry because Asian options cost less than their vanilla counterparts and are less sensitive to the change in underlying asset prices.
- The common forms of averaging in option contracts can either be geometric average or arithmetic average of the underlying variables.

Asian Options: Asian Options

Asian option : discrete monitoring .

Floating strike Asian call : Payoff $= \max(S(T) - \bar{S}, 0)$

Fixed strike Asian call : Payoff $= \max(\bar{S} - K, 0)$

Arithmetic average option : $\bar{S} = \frac{1}{n}\sum_{k=1}^{n} S(t_k);$

Geometric average option : $\bar{S} = \left(\prod_{k=1}^{n} S(t_k)\right)^{1/n};$

The latter is analytically tractable.

Example: Lookback Options

The payoffs of lookback options depend on the maximum or minimum stock price during the life of the option.

$$\text{Floating strike call}: \text{Payoff} = S(t_n) - \min_{k=1,\dots,n} S(t_k).$$

$$\text{Floating strike put}: \text{Payoff} = \max_{k=1,\dots,n} S(t_k) - S(t_n).$$

$$\text{Fixed strike call}: \text{Payoff} = (\max_{k=1,\dots,n} S(t_k) - K)^+.$$

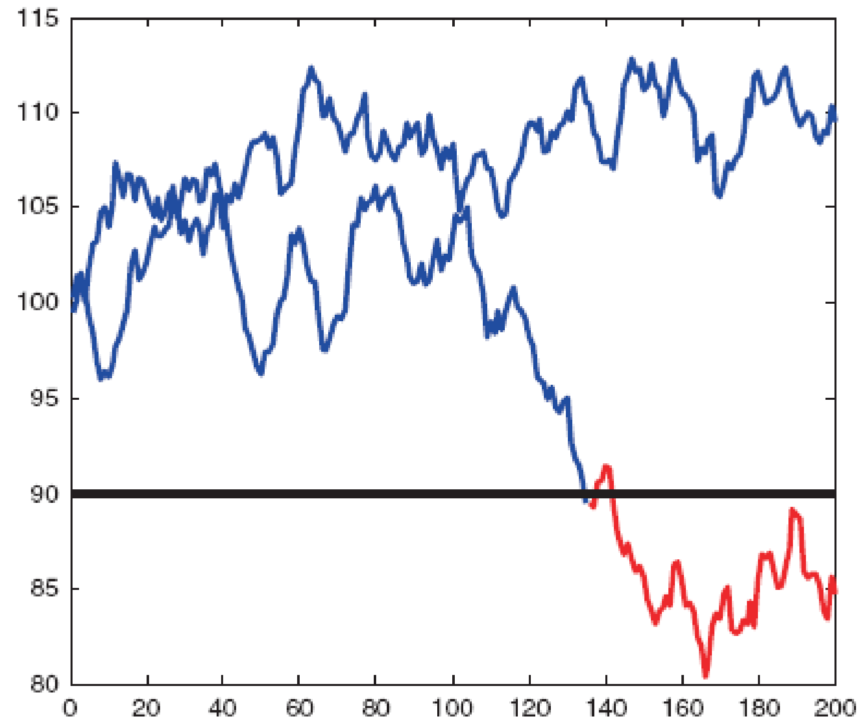
$$\text{Fixed strike put}: \text{Payoff} = (K - \min_{k=1,\dots,n} S(t_k))^+.$$

Example: Barrier Options

- Barrier options are options where the payoff depends on whether the underlying asset's price reaches a certain level during a certain period of time.

- EX:

  Two possible paths of the asset price.
  When the price hits the barrier (lower path), the option expires as worthless.



Example: Barrier Options

--- Knock-in option: Option comes into existence only if stock price hits barrier before option maturity.

--- Knock-out option: Option dies if stock price hits barrier before option maturity.

--- 'Up' options : Stock price must hit barrier from below (the barrier is above the initial asset value) (in or out).

--- 'Down' options: Stock price must hit barrier from above (the barrier is below the initial asset value) (in or out).

--- Option may be a put or a call

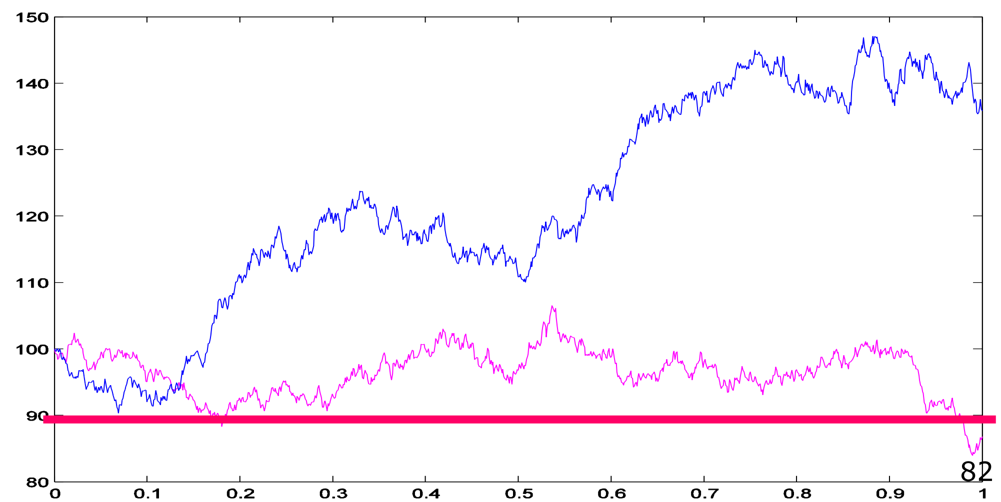
- Eight possible combinations

$$\begin{Bmatrix} \textbf{in} \\ \textbf{out} \end{Bmatrix} \textbf{ or } \begin{Bmatrix} \textbf{up} \\ \textbf{down} \end{Bmatrix} \textbf{ or } \begin{Bmatrix} \textbf{call} \\ \textbf{put} \end{Bmatrix}$$

Example: Barrier Options

- Down-and-out (call or put): ceases to exist if the asset price reaches a certain barrier level H. The barrier level is below the initial asset price.
- Down-and-in (call or put): comes into existence only if the asset price reaches the barrier level H.
- Clearly, $c_{do} + c_{di} = c$, $p_{do} + p_{di} = p$

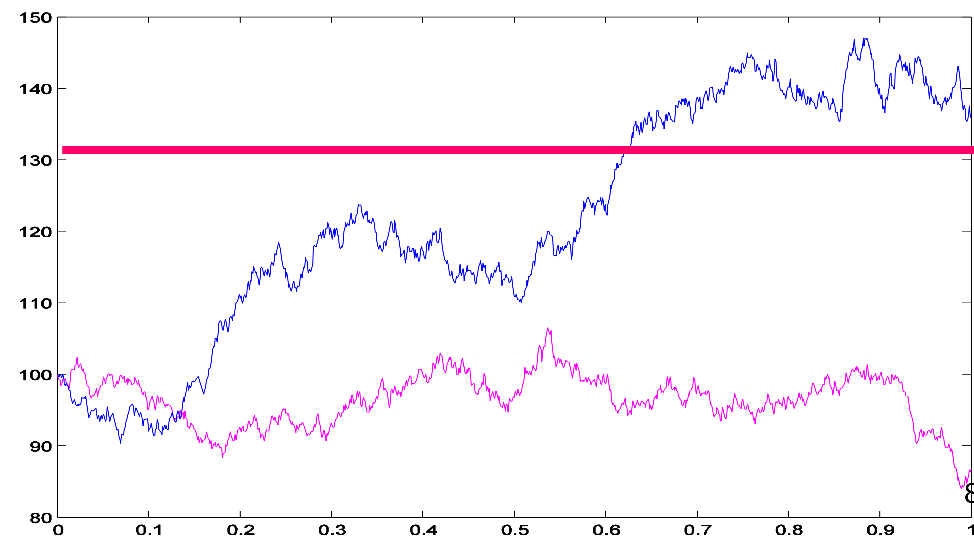
In + Out = Vanilla



Example: Barrier Options

- Up-and-out (call or put): ceases to exist if the asset price reaches a certain barrier level H, that is, higher than the initial asset price.
- Up-and-in (call or put): comes into existence only if the asset price reaches the barrier level H.
- Clearly, $c_{uo} + c_{ui} = c, \qquad p_{uo} + p_{ui} = p$

In + Out = Vanilla



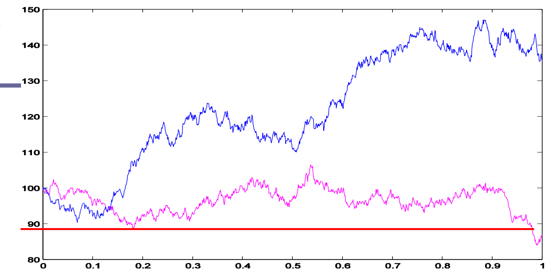
Remarks

- There are single barrier and double barrier options.
- A single barrier option has one barrier that may be either greater than or less than the strike price.
- A double barrier option (corridor option) has barriers on either side of the strike.

Why would we ever buy options with a barrier on it?

- Because it is cheaper and we have a specific view about the path that spot will take over the lifetime of the structure.

Example: Barrier Options



Payoffs:

Down-and-out call:

$$\text{Payoff}: \max(S(T)-K,\ 0)\cdot I\{\tau(b) > T\}$$

where $\tau(b) = \inf\{t_i : S(t_i) < b\}$

is the first time in $\{t_1,\ldots,t_n\}$ the price of the underlying asset drops below b.

Down-and-in call:

$$\text{Payoff}: \max(S(T)-K,\ 0)\cdot I\{\tau(b) \le T\}$$

Ex. Multi-Asset Options (basket options)

- The payoff of a basket option is the weighted average of two or more underlying assets --- a "basket" of assets.
- The complication in evaluating a basket option is that the underlying assets are almost always correlated. Fortunately, correlation is no problem for MC method.

$$dS_j(t) = \mu_j S_j(t)dt + \sigma_j S_j(t)dX_j(t),\ j = 1,\dots,d,$$

$$(X_1(t),\dots,X_d(t)) \sim BM(0,\Sigma) \text{ with } \Sigma_{ij} = \rho_{ij}.$$

$$S_j(t) = S_j(0)\exp((\mu_j - \sigma_j^2/2)t + \sigma_j X_j(t)).$$

Example: Multi-Asset Options

Basketoption :

$$\text{Payoff} : \max(\bar{S} - K, \ 0), \quad \text{where} \quad \bar{S} = \sum_{k=1}^{d} w_k S_k(T);$$

Spread option :

$$\text{Payoff} = ([S_1(T) - S_2(T)] - K)^+;$$

Outperformance option :

$$\text{Payoff} = (\max\{w_1 S_1(T), \ldots, w_d S_d(T)\} - K)^+;$$

Barrier option :

$$\text{Payoff} = (K - S_1(T))^+ \cdot I\{\min_{i=1,\ldots,n} S_2(t_i) < b\};$$

Example: Basket Asian Options

Basket Asian option : discrete monitoring .

Payoff : $\max(\bar{S} - K, \ 0)$

Arithmetic average option : $\bar{S} = \frac{1}{nd} \sum_{i=1}^{n} \sum_{j=1}^{d} S_j(t_i);$

Geometric average option : $\bar{S} = \left(\prod_{i=1}^{n} \prod_{j=1}^{d} S_j(t_i) \right)^{1/nd};$

Remark:

A single asset at multiple dates $:(S_{t1},...,S_{tn}):$

$$(U_1,...,U_n) \Rightarrow (Z_1,...,Z_n) \Rightarrow (B_{t1},...,B_{tn}) \Rightarrow (S_{t1},...,S_{tn})$$

Multiple assets at one date $:(S_T^1,...,S_T^d):$

$$(U_1,...,U_d) \Rightarrow (Z_1,...,Z_d) \Rightarrow (B_T^1,...,B_T^d) \Rightarrow (S_T^1,...,S_T^d)$$

Multiple assets at multiple dates $:(S_{t1}^1,...,S_{t1}^d,...S_{tn}^1,...,S_{tn}^d):$

$$(U_1,...,U_{nd}) \Rightarrow (Z_1,...,Z_{nd}) \Rightarrow \begin{pmatrix} B_{t1}^1,...,B_{tn}^1 \\ ... \\ B_{t1}^d,...,B_{tn}^d \end{pmatrix} \Rightarrow \begin{pmatrix} S_{t1}^1,...,S_{tn}^1 \\ ... \\ S_{t1}^d,...,S_{tn}^d \end{pmatrix}$$

10. Simulating Jump-diffusion model of Merton

- The continuous-time stochastic for stock price modelling have continuous paths. Hence, if we monitor such a process closely, we cannot be caught by surprise by an exceptionally big move.
- However, there are many reasons such as catastrophes or surprising news (unexpected political changes, economic scandal, pandemic and so forth) that can make a real-life process jump.
- Due to discreteness of measurement, it can also be argued that processes modelling the real world should allow for discontinuities in their paths.

10. Simulating Jump-diffusion model of Merton

- Financial models for which randomness is input **through BM** are such that the price paths are almost surely **continuous**.
- Sometimes, prices move in an abrupt way that cannot be captured by a continuous model.
  It is of interest to study models that **allow jumps** to occur in the price paths.
- **Examples:**
  - The jump-diffusion model of Merton
  - Variance Gamma model of Madan
  - Normal Inverse Guassian model

Jump-diffusion model of Merton

$$\mathrm{dlog}\ S = \mu\, dt + \sigma\, dB(t) + Q dN(t),$$

S --- the stock price;

B(t) --- standard BM;

N(t) --- Poission process with an intensity λ;

Q --- Normal with mean k and variance s^2;

B, N and Q are independent at each time t.

Discrete Approximation

$$\Delta \log S = \mu \Delta t + \sigma \Delta B + Q \Delta N.$$

when Δt **is small,** ΔN **is**

either 1, with prob. $\lambda \Delta t$ **; or 0, with prob.** $1 - \lambda \Delta t$ **:**

$$\Delta N \approx \begin{cases} 1, & \textbf{with prob. } \lambda \Delta t, \\ 0, & \textbf{with prob. } 1 - \lambda \Delta t. \end{cases}$$

Define $X_i = \log S_i - \log S_{i-1}$. **Then**

$$X_i \mid \Delta N = 0 \sim N(\mu \Delta t, \sigma^2 \Delta t);$$

$$X_i \mid \Delta N = 1 \sim N(\mu \Delta t + k, \sigma^2 \Delta t + s^2).$$

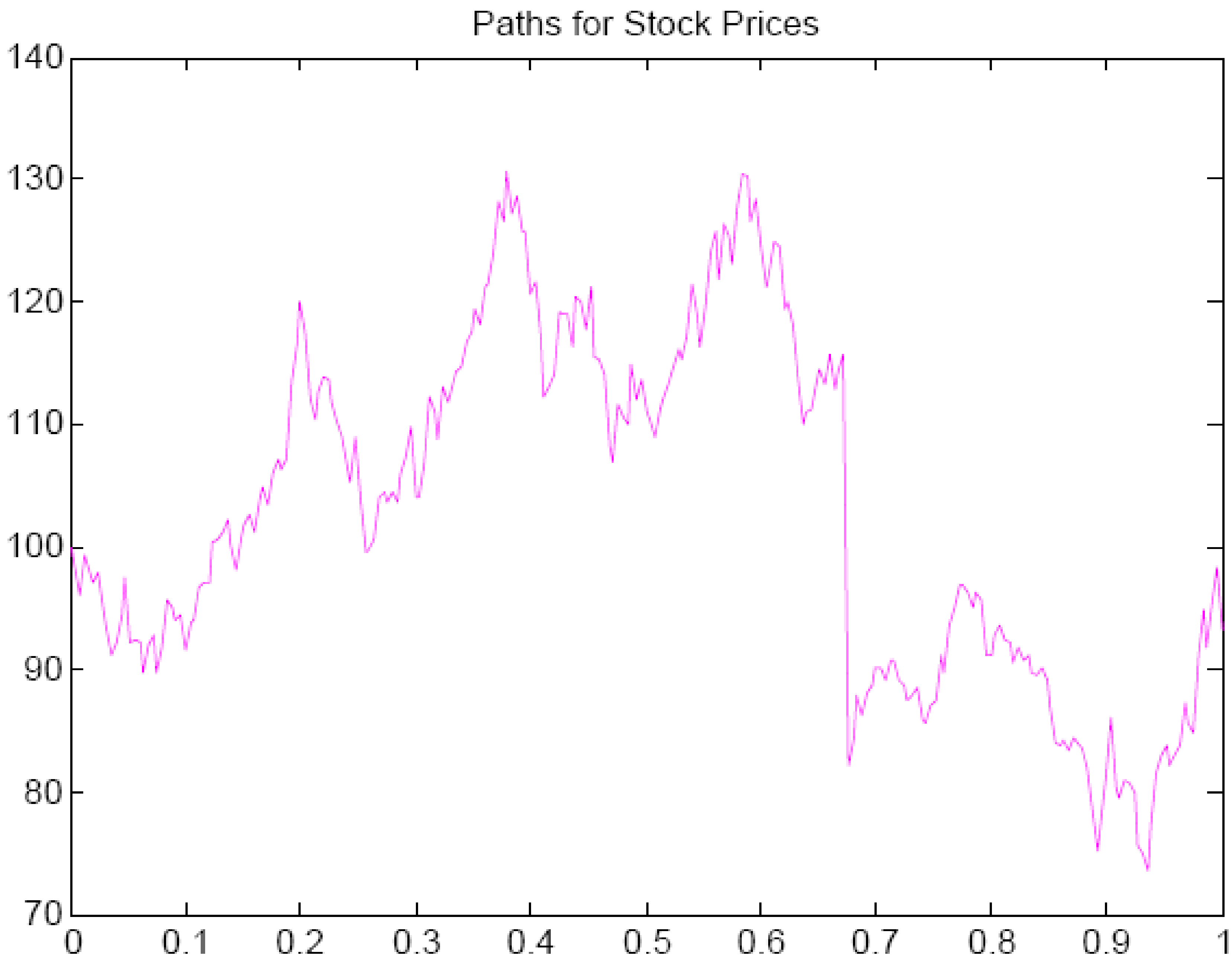
Algorithm

1. Generate $U \sim U(0,1),\ Z \sim N(0,1)$ **independently;**

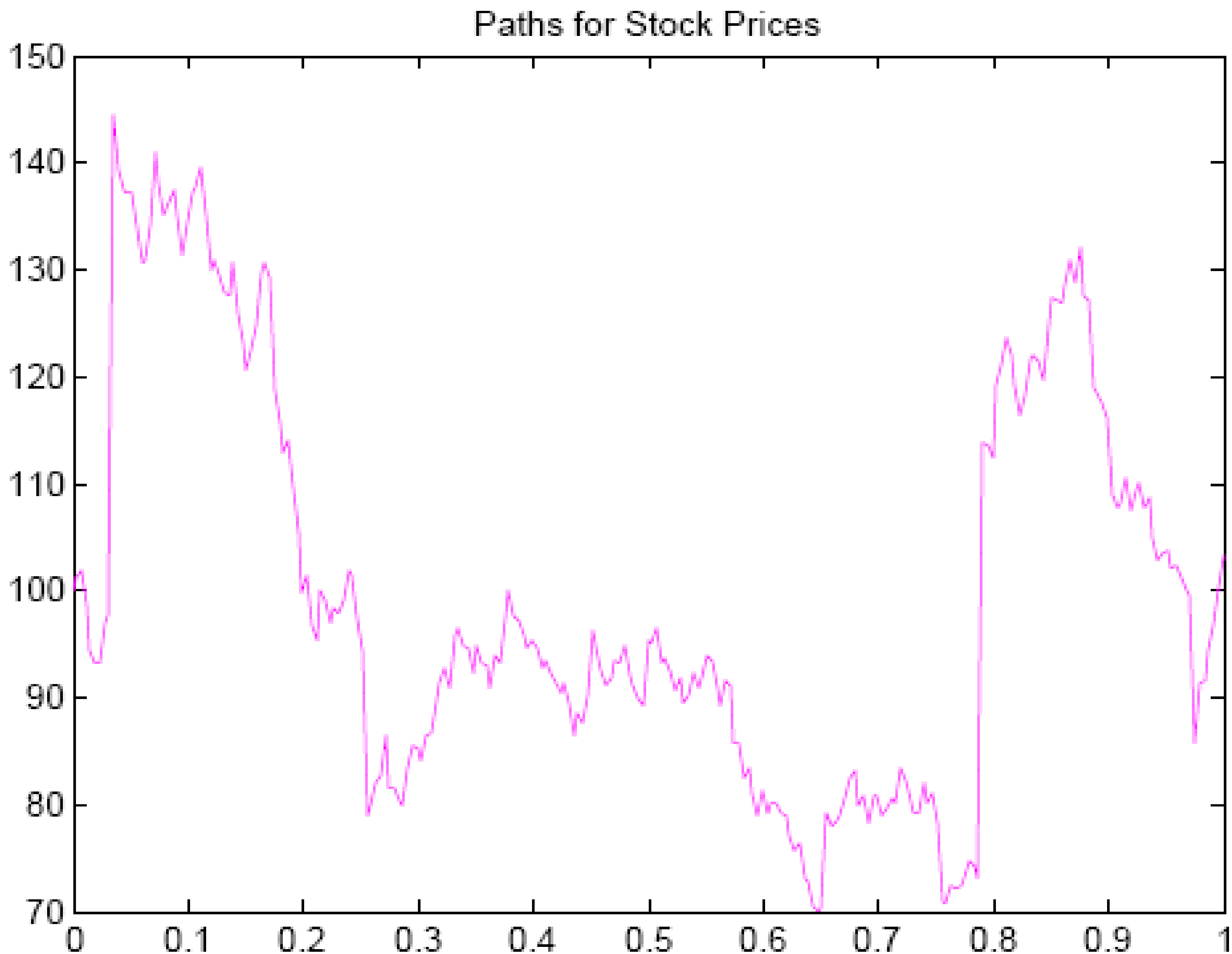
2. Let $X_i = \begin{cases} \mu\Delta t + \sigma\sqrt{\Delta t}\ Z; & \textbf{if } U > \lambda\Delta t; \\ \mu\Delta t + k + \sqrt{\sigma^2\Delta t + s^2}\ Z, & \textbf{if } U \le \lambda\Delta t. \end{cases}$

jump

From $\{X_i\} \Rightarrow \{\log S_i\} \Rightarrow \{S_i\}$ **(one path).**



Paths for Stock Prices
140
130
120
110
100
90
80
70
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Paths for Stock Prices
150
140
130
120
110
100
90
80
70
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Jump-diffusion model of Merton

- The simulation method above has discretization error, depending on the time step size.
- **Exact simulation** (without discretization) for the jump-diffusion model is possible and now is widely used in practice.

Jump-diffusion model of Merton
(with the same assumptions as Merton)

$$\frac{dS(t)}{S(t-)} = \mu\, dt + \sigma\, dB(t) + dJ(t),$$

$B(t)$ - standard BM; $S(t)$ - stock price;

$S(t-)$ - stock price just before a potential jump

$$S(t-) = \lim_{u \to t-0} S(u).$$

$$J(t) = \sum_{j=1}^{N(t)} (Y_j - 1) \text{ - compound Poission process, } Y_j \ldots \text{iid}.$$

$$Y_j \sim LN(a, b^2), \text{ i.e., } \log Y_j \sim N(a, b^2);$$

$N(t)$ - Poission process with an intensity λ;

$B(t), N(t),$ and $\{Y_1, Y_2, \ldots\}$ are independent at each time t.

The jump times and the jump size

$$\frac{dS(t)}{S(t-)} = \mu\, dt + \sigma\, dB(t) + dJ(t), \qquad J(t) = \sum_{j=1}^{N(t)} (Y_j - 1)$$

Let: Jump times $0 < \tau_1 < \tau_2 < \cdots$

$N(t) = \sup\{n : \tau_n \le t\}$: number of total jumps in $[0, t]$;

The jump in S at τ_j is

$$S(\tau_j) - S(\tau_j -) = S(\tau_j -)[J(\tau_j) - J(\tau_j -)] = S(\tau_j -)(Y_j - 1).$$

Hence $S(\tau_j) = S(\tau_j -) Y_j$.

Y_j -1 is the relative size of jump

So Y_j is the ratio of the assetprice before and after a jump :

The jumps are multiplica tive.

We have $\log S(\tau_j) = \log S(\tau_j -) + \log Y_j$.

The analytical solution

$$\frac{dS(t)}{S(t-)} = \mu\, dt + \sigma\, dB(t) + dJ(t), \quad \text{where} \quad J(t) = \sum_{j=1}^{N(t)} (Y_j - 1).$$

The solution is

$$S(t) = S(0) \exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)] \cdot \prod_{j=1}^{N(t)} Y_j.$$

Conditional on $N(t) = n$, $S(t)$ has a lognormal distribution (denoted by $F_{n,t}$). Thus

Poission mixture of lognormal

$$P(S(t) \le x) = \sum_{n=0}^{\infty} F_{n,t}(x) \cdot \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

LAW OF TOTAL PROBABILITY

Page 137

Based on this, Merton obtained the price of European option.

Remark: Risk-Neutral Measure

$$S(t) = S(0)\exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)] \cdot \prod_{j=1}^{N(t)} Y_j =: S^*(t) \cdot J(t)$$

$$\Rightarrow E[J(t)] = \exp(\lambda m t), \quad m = E(Y_j) - 1.$$

Next next page

Since $E[S^*(t)] = S^*(0)e^{\mu t}$

$$\Rightarrow E[S(t)] = E[S^*(t)J(t)] = E[S^*(t)]\,E[J(t)] \quad \text{(by independence)}$$

$$= S^*(0)\exp[\mu t + \lambda m t]$$

In risk-neutral world : $S(t)\,e^{-rt}$ is a martingale :

$$E[S(t)] = S(0)\,e^{rt}.$$

$\Rightarrow$ Risk-neutral prob. will result when μ is given by

$$\mu t + \lambda m t = rt \quad \Rightarrow \mu = \mu^* = r - \lambda m.$$

If lamda =0, then this reduces to the GBM case

Jump-diffusion model of Merton

With $\mu = r - \lambda m$, we have

$$\frac{dS(t)}{S(t-)} = \mu\, dt + \sigma\, dB(t) + dJ(t),$$

$$= r\, dt + \sigma\, dB(t) + \underline{[dJ(t) - \lambda m dt]}.$$

The last two terms on the right are martingales, and the net growth rate in S(t) is indeed r.

$$\text{Let } J(t)=\prod_{i=1}^{N(t)} Y_i. \text{ Then } E[J(t)]=\exp(\lambda m t),\ m=E(Y_i)-1.$$

Proof. Since $E[J(t) \mid N(t)=n]$

$$=E[\prod_{i=1}^{N(t)} Y_i \mid N(t)=n]=E[\prod_{i=1}^{n} Y_i \mid N(t)=n]$$

By independence of Yi and N(t)

$$=E[\prod_{i=1}^{n} Y_i]=\prod_{i=1}^{n} E[Y_i]=(m+1)^n.$$

$$E[J(t)]=\sum_{n=0}^{\infty} E[J(t) \mid N(t)=n]\cdot P(N(t)=n)$$

$$=\sum_{n=0}^{\infty} (m+1)^n \cdot e^{-\lambda t}(\lambda t)^n/n!$$

$$=e^{-\lambda t}\sum_{n=0}^{\infty} ((m+1)\lambda t)^n/n! = \exp(\lambda m t)$$

Simulating at fixed dates

Similating at fixed dates $0 = t_0 < t_1 < t_2 < \cdots$

$$S(t) = S(0)\exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)] \cdot \prod_{j=1}^{N(t)} Y_j.$$

$\Rightarrow$

$$S(t_{i+1}) = S(t_i)\exp[(\mu - \frac{1}{2}\sigma^2)(t_{i+1} - t_i) + \sigma(B(t_{i+1}) - B(t_i))] \cdot \prod_{j=N(t_i)+1}^{N(t_{i+1})} Y_j.$$

Set $X(t) = \log S(t)$. We have

$$X(t_{i+1}) = X(t_i) + (\mu - \frac{1}{2}\sigma^2)(t_{i+1} - t_i) + \sigma[B(t_{i+1}) - B(t_i)] + \sum_{j=N(t_i)+1}^{N(t_{i+1})} \log Y_j.$$

Similation from t_i to t_{i+1}: Based on

$$X(t_{i+1}) = X(t_i) + (\mu - \frac{1}{2}\sigma^2)(t_{i+1} - t_i) + \sigma[B(t_{i+1}) - B(t_i)] + \underline{\sum_{j=N(t_i)+1}^{N(t_{i+1})} \log Y_j}.$$

M

1. Generate $Z \sim N(0,1)$;
2. Generate $N \sim \text{Poission}(\lambda(t_{i+1} - t_i))$; (number of jumps) P.128.

   if $N = 0$, set $M = 0$ and goto step 4.
3. Generate $Z_2 \sim N(0,1)$ and set $M = aN + b\sqrt{N}Z_2$;
4. Set

$$X(t_{i+1}) = X(t_i) + (\mu - \frac{1}{2}\sigma^2)(t_{i+1} - t_i) + \sigma\sqrt{t_{i+1} - t_i}\, Z + M.$$

Note: since $\log Y_j \overset{iid}{\sim} N(a, b^2)$, then $\underline{\sum_{j=1}^{n} \log Y_j \sim N(aN, b^2 N)}$.

Simulating jump times

- In simulating jump-diffusion processes, the easiest way to incorporate the jumps is to **firstly simulate the times when the jumps occur.**
- Once one has the jump times and the jump size distribution, the cumulative size of the jump component for each jump time can be determined, and then this is simply added to the diffusion component.

Simulating jump times $\tau_1, \tau_2, \ldots$

From one jump to the next, stock price S(t) evolves like ordinary GBM. We have

$$\mathbf{S(\tau_{i+1}-) = S(\tau_i) + (\mu - \frac{1}{2}\sigma^2)(\tau_{i+1} - \tau_i) + \sigma[B(\tau_{i+1}) - B(\tau_i)].}$$

and $\mathbf{S(\tau_{i+1}) = S(\tau_{i+1}-)\, \underline{Y}_{i+1}.}$

Simulating jump times $\tau_1, \tau_2, \ldots$

1. Generate R_{i+1} from EXP (λ), with mean $1/\lambda$;
2. Generate $Z_{i+1} \sim N(0,1)$;
3. Generate $\log Y_{i+1} \sim N(a, b^2)$;
4. Set $\tau_{i+1} = \tau_i + R_{i+1}$, and

$$X(\tau_{i+1}) = X(\tau_i) + (\mu - \frac{1}{2}\sigma^2) R_{i+1} + \sigma \sqrt{R_{i+1}}\, Z_{i+1} + \log Y_{i+1}.$$

$$R_{i+1} = -\frac{1}{\lambda}\log(U), U \sim U(0,1).$$

Interarrival Time

Note:

The interarrival time $\tau_{i+1} - \tau_i$ is exponentially distributed:

$$P(\tau_{i+1} - \tau_i \le t) = 1 - e^{-\lambda t}.$$

See next page

Remark: Poisson分布与指数分布的关系

设某股票在任意长为 t 的时段[0, t]内发生跳跃的次数 $N(t) \sim P(\lambda t)$. 则相邻两次跳跃之间的时间间隔 $T \sim \mathrm{Exp}(\lambda)$.

证：由$N(t) \sim P(\lambda t)$,有$P(N(t)=k) = \dfrac{(\lambda t)^k}{k!} e^{-\lambda t}, k = 0,1,\ldots$

以0记上一次跳跃时刻，T为下一次跳跃时刻(相邻两次跳跃间隔时间),

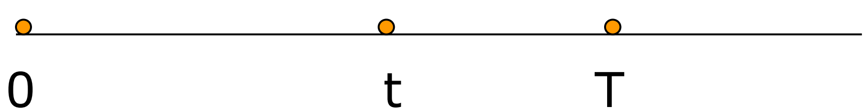
则$\{T > t\}$等价于$\{N(t)=0\}$,即$[0,t]$内没有发生跳跃。

当$t<0$时, $F_T(t) = P(T \le t) = 0$;

当$t \ge 0$时, $F_T(t) = P(T \le t) = 1 - P(T > t)$

$$= 1 - P(N(t)=0) = 1 - e^{-\lambda t}.$$

$$\Rightarrow T \sim \mathbf{Exp}(\lambda).$$



Combined Approach

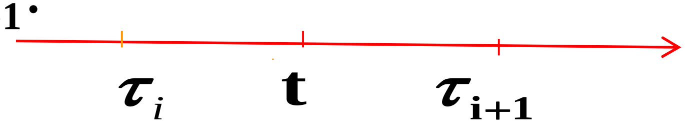
We fix a date t in advance that we would like to include among the simulated dates. Suppose $\tau_i < t < \tau_{i+1}$. Then from τ_i to the time t, and from t to τ_{i+1}-, the stock price S(t) evolves like ordinary GBM :

$$S(t) = S(\tau_i) + (\mu - \frac{1}{2}\sigma^2)(t - \tau_i) + \sigma[B(t) - B(\tau_i)],$$

$$S(\tau_{i+1}-) = S(t) + (\mu - \frac{1}{2}\sigma^2)(\tau_{i+1} - t) + \sigma[B(\tau_{i+1}) - B(t)],$$

and at τ_{i+1}, we have a jump

$$S(\tau_{i+1}) = S(\tau_{i+1}-)\, Y_{i+1}.$$



11. Generating Other Processes

- Stochastic Volatility model;
- Time-changed Brownian motion;
- Variance Gamma model of Madan;
- Normal Inverse Guassian model;
- Levy process.

(1) Stochastic Volatility model

This model replaces the constant volatility in the lognormal model by a stochastic volatility:

Square root diffusion process

$$dS(t) = rS(t)dt + \underline{\boldsymbol{\sigma}(t)} S(t) dW_1(t),$$

$$d\boldsymbol{\sigma}^2(t) = \boldsymbol{\kappa}(\boldsymbol{\theta} - \boldsymbol{\sigma}^2(t))dt + \underline{\boldsymbol{\xi}}\, \boldsymbol{\sigma}(t) dW_2(t),$$

$$\textbf{where} \begin{pmatrix} \mathbf{W_1(t)} \\ \mathbf{W_2(t)} \end{pmatrix} \sim \mathbf{BM}(0, \Sigma), \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

$$\textbf{Each } \mathbf{W_i(t)} \textbf{ is standard BM} : \mathbf{E(W_i(t)) = 0, D(W_i(t)) = t.}$$

$$\mathbf{W_1(t)} \textbf{ and } \mathbf{W_2(t)} \textbf{ have correlation } \rho : \mathbf{Cov(W_1(t), W_2(t))} = \rho t.$$

ξ **is the volatility of the volatility process.**

(1) Stochastic Volatility model

$$dS(t) = rS(t)dt + \sigma(t)S(t)dW_1(t),$$

$$d\sigma^2(t) = \kappa(\theta - \sigma^2(t))dt + \xi\sigma(t)dW_2(t),$$

$$\textbf{For} \quad \begin{pmatrix} \mathbf{W_1(t)} \\ \mathbf{W_2(t)} \end{pmatrix} \sim \mathbf{BM}(\mathbf{0}, \Sigma),\ \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

$$\Rightarrow \begin{pmatrix} \mathbf{W_1(t)} \\ \mathbf{W_2(t)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} \mathbf{B_1(t)} \\ \mathbf{B_2(t)} \end{pmatrix},$$

$\mathbf{B_1(t)}$ and $\mathbf{B_2(t)}$ are two independent BM.

Thus the model can be written as

$$dS(t) = rS(t)t + \sigma(t)S(t)dB_1(t),$$

$$d\sigma^2(t) = \kappa(\theta - \sigma^2(t))dt + \xi\sigma(t)\,[\rho dB_1(t) + \sqrt{1-\rho^2}\,dB_2(t)],$$

(1) Stochastic Volatility model

- A closed-form expression can be derived for the price of a plain call option under the stochastic volatility model.
- But for more complicated options we might need to use Monte Carlo method and a discretization of the process.
- Next is the method of Euler discretization. Exact simulation of the model is possible and was extensively studied.

Discretized Version

$$\Delta S(t) = rS(t)\Delta t + \sigma(t)S(t)\Delta W_1(t),$$

$$\Delta\sigma^2(t) = \kappa(\theta - \sigma^2(t))\Delta t + \xi\sigma(t)\Delta W_2(t),$$

Since :

$$\begin{pmatrix} \Delta W_1(t) \\ \Delta W_2(t) \end{pmatrix} \sim N(0, \Delta t\ \Sigma),\ AA^T = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \textbf{ where } A = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix},$$

we have $\begin{pmatrix} \Delta W_1(t) \\ \Delta W_2(t) \end{pmatrix} = \sqrt{\Delta t} \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},\ Z_1, Z_2 \overset{iid}{\sim} N(0,1)$

$$S(t+\Delta t) = S(t) + rS(t)\Delta t + \sigma(t)S(t)[\sqrt{\Delta t}\ Z_1],$$

$$\sigma^2(t+\Delta t) = \sigma^2(t) + \kappa(\theta - \sigma^2(t))\Delta t + \xi\sigma(t)\sqrt{\Delta t}[\rho Z_1 + \sqrt{1-\rho^2}\ Z_2].$$

(2) Time-Changed BM

- If one looks at time series of stock price returns, then one often observes phases of high frequency price changes followed by phases where the intensity of price changes is comparably low. This is referred to as **volatility clustering.**
- To model this, one introduces **a different clock** where time moves in velocity proportional to the trading activity. When there are a lot of trades, the internal time of the process runs faster. It runs slower if nearly nothing happens.
- In such a way it is hoped to care for the main weakness of the BM-based BS model which moves in a too uniform way over time.

(2) Time-Changed BM

An arithmetic BM with drift θ and volatility σ:

$$\theta t + \sigma W_t,$$

where $\{W_t\}_{t\geq 0}$ is a standard BM.

Consider a subordinator $T = (G_t)_{t\geq 0}$, indenpendent of W_t.

Substituting $t = G_t$ in the arithmetic BM, we have a new process $\{X_t\}_{t\geq 0}$

$$X_t = \theta G_t + \sigma W_{G_t}$$

which is time-changed BM. Using it we may define various processes.

Remarks: Subordinator

- If a pure jump process $T=(T_t)$ is non-decreasing, then it is referred to as the subordinator or intrinsic time process.
- With this random time change, a BM can again be used as the basic building block for modelling the uncertainty of the future stock price evolution.
- A simple recipe for generating a new process is to use a subordinator as a model for the evolution of time and plugging it into a BM.
- Poisson, Gamma, and inverse Gaussian processes are non-decreasing and then they are subordinator.

(2) Time-Changed BM

A large part of modern finance has been concerned with modelling the evolution of return processes over time.

By subordination, it is possible to capture empirically observed anomalies that contradict the classical log-normallity assumption for asset prices.

(2) Time-Changed BM: Simulation

Simulating Time-changed BM

$$X_t = \theta G_t + \sigma W_{G_t}$$

1. Fixed a time grid $t_1, \dots, t_n$ and $G_0 = 0$;
2. Simulte increments of the subordinator

$$\Delta G_i = G_{ti} - G_{ti-1}.$$

3. Simulate iid standard normal $Z_1, \dots, Z_n.$
4. Calculate increments

$$\Delta X_i = \theta \underline{\Delta G_i} + \sigma \sqrt{\underline{\Delta G_i}}\, Z_i.$$

当成时间间隔

(3) Variance Gamma Model

A Variance Gamma process is a stochastic process $\{X_t\}_{t\geq 0}$ defined by

$$X_t = \theta G_t + \sigma W_{G_t}$$

Gamma process is the subordinator

where $\{W_t\}_{t\geq 0}$ is a standard BM, $\{G_t\}_{t\geq 0}$ is a Gamma process with some parameters.

VG is a time-changed BM, where the subordinator is a Gamma process. It is obtained on evaluating BM at a time given by Gamma process.

Remarks

- $X(t)$ can be viewed as the result of applying a random time-change to an ordinary BM.
- The deterministic time argument t has been replaced by the random time $G(t)$, which is the conditional variance of $W(G(t))$ given $G(t)$.
- This explains the name Variance Gamma.

Remark: Gamma Distribution

Gamma distribution Gamma (a, b)

See Glasserman (P. 125) for Gamma distribution

$$f(y) = \frac{1}{\Gamma(a) b^a} y^{a-1} e^{-y/b}, \; y \geq 0.$$

It has mean ab and variance ab^2.

Gamma distribution is infinitely divisible:

Let $Y_1, \ldots, Y_n \overset{iid}{\sim} \text{Gamma}(a/n,\ b)$, then $Y_1 + \cdots + Y_n \sim \text{Gamma}(a,\ b)$.

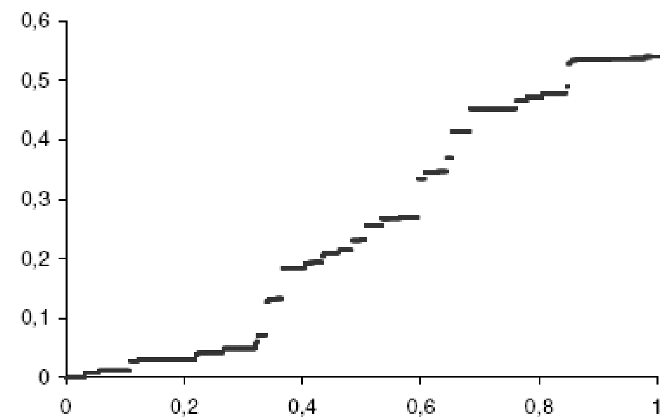
Gamma process with parameters $a, b > 0$:

for fixed time $s < t$, the increment

$$G(t) - G(s) \sim \text{Gamma}(a(t-s), b)$$

and the increments are independent.

A simulated path of Gamma process



(3) Variance Gamma Model

For a Variance Gamma process

$$X_t = \theta G_t + \sigma W_{G_t},$$

conditional on the increment $G(t_{i+1}) - G(t_i)$, the increment $X(t_{i+1}) - X(t_i)$ has normal distribution with mean $\theta[G(t_{i+1}) - G(t_i)]$ and variance $\sigma^2[G(t_{i+1}) - G(t_i)]$.

We may simulate variance gamma process as follows:

1. Generate $Y \sim \mathrm{Gamma}((t_{i+1} - t_i)/b, b)$

for a=1/b

2. Generate $Z \sim N(0,1)$,

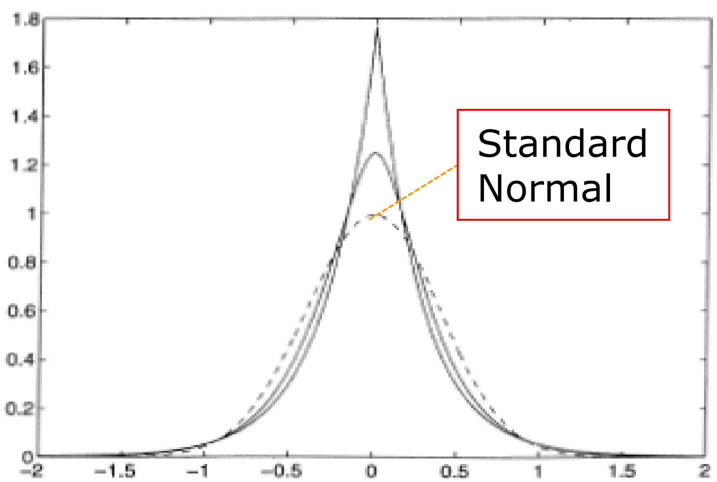
3. Set $X(t_{i+1}) = X(t_i) + \theta \underline{Y} + \sigma \sqrt{\underline{Y}}\, Z.$

当成时间间隔

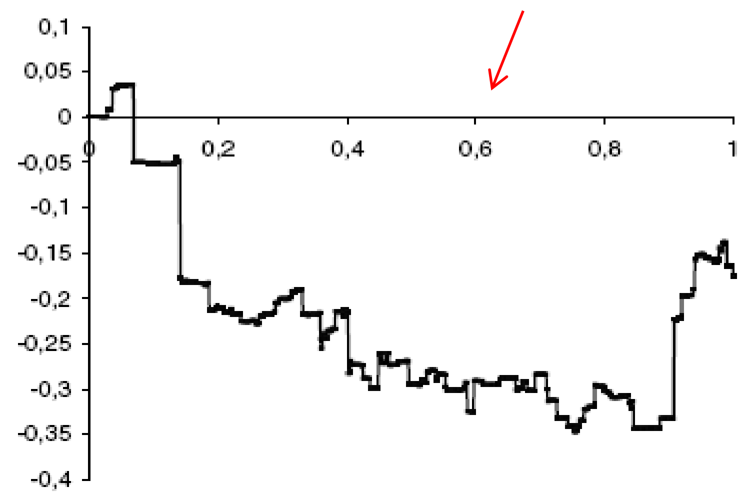
(3) Variance Gamma Model

- As gamma process is strictly increasing by jump, VG process only changes by jumps.
- Next figure (right) shows a simulated path of VG process. We have drawn the paths as a continuous function for better visualization. Actually, a VG process is a pure jump process.

Variance Gamma densities



X(t)=W(G(t))



(3) Variance Gamma Model

The Variance Gamma model for the stock price is :

$$S(t) = S(0)\exp(X(t)),$$

with $X(t) = W(G(t))$,

The deterministic time argument t has been replaced by the random time G(t).

where $W(\cdot)$ - BM with drift μ and diffusion coefficient σ;

$G(\cdot)$ - Gamma process with parameters $a, b > 0$:

for fixed time $s < t$, the increment

$$G(t) - G(s) \sim \mathrm{Gamma}((t-s)/b, b) \qquad (a = 1/b)$$

and these increments are independent.

(3) Variance Gamma Model

- We have two possibilities for sampling VG paths. The MC method for option pricing in VG model is straightforward:

1. Sample a sufficiently big number of stock price paths of the VG process to obtain VG stock price paths under the risk-neutral measure;
2. Calculate the corresponding option payoffs;
3. Estimate the option price by the discounted average over these payoffs.

- Note: For path-independent options it is enough to generate the final value of the VG process and not the whole path.

(4) Normal Inverse Gaussian Process

A Normal Inverse Gaussian (NIG) process is a stochastic process $\{X_t\}_{t\geq 0}$ defined by

$$X_t = \theta G_t + \sigma W_{G_t}$$

where $\{G_t\}_{t\geq 0}$ is a Inverse Gaussian (IG) process with some parameters.

NIG process is a time-changed BM, where the subordinator is a IG process. It is obtained on evaluating BM at a time given by IG process.

(4) Normal Inverse Gaussian Process

- A property that makes the *NIG*-distribution suitable for a log-return model is the fact that it is much more flexible than the normal distribution.
- Also, it can produce a higher peak in the centre and at the same time more heavy tails than the normal distribution while having the same mean and variance.

(5) Levy Process

A stochastic process $X=(X_t)$ is called a Levy process if the following five conditions are satisfied:

1. $X_0 = 0$ a.s.
2. X has independent increments.
3. X has stationary increments.

The distribution of $X_{t+h} - X_t$ does not depend on t, it has the same distribution as X_h.

4. X is stochastically continuous

$$\lim_{s \to t} P\{|X_s - X_t| > \varepsilon\} = 0 \; (\forall t \geq 0, \forall \varepsilon > 0)$$

The probability of seeing a jump at t is zero.

5. X is right continuous and has left limits (*cadlag*).

(5) Levy Process

- The condition 4 does not imply that the sample paths are continuous.
- It is merely a formalization of the idea that for a given time t, the probability of seeing a jump is zero.
- In other words, the jumps are not predictable - they occur at random times. These exclude processes which have jumps at fixed times.

(5) Levy Process

- It is straightforward to verify that BM satisfies the above conditions, making it an example of a Levy process.
- The class of Levy processes is bigger than BM though and includes Poisson processes and Compound Poisson processes among many others.
- Gamma process and NIG are also Levy process.

(5) Levy Process

- Besides Wiener processes there are several known Levy processes. The simplest is pure drift, $X_t = m\ t$.
- This and the Wiener process are the only two that are continuous, all others have jumps.
- The simplest non-continuous Levy process is the *Poisson process.*
- Using subordinator instead of physical time for BM, we obtained various pure jump processes.
- All these processes are included in the class of Levy process.

(5) Levy Process

Infinitely divisibility:

A random variable Y (more precisely, its distribution) is said to be infinitely divisible, if for each n=2, 3, …, there are i.i.d. RVs $Y_1, \ldots, Y_n$, such that $Y_1 + \ldots + Y_n$ has the distribution of Y.

- Y can be ''divided'' into n i.i.d. parts.
- infinitely divisible distributions can be split into arbitrarily many identically distributed `bits'.

(5) Levy Process

Infinitely divisibility:

- The most common infinitely divisible laws are: Gaussian, Gamma, and Poisson distributions.
- A RV having any of those distributions can be decomposed into a sum of n iid parts having the same distribution but with modified parameters.
- For example, if X~N(mu, sigma^2), then one can write that $X= X_1+\ldots+X_n$, where $X_1, \ldots, X_n$ are iid with law N(mu/n, sigma^2/n).
- The uniform law on an interval is not infinitely divisible.

(5) Levy Process

- We must be able to:
  - (1) divide the fundamental interval, $[0, T]$ into arbitrary subintervals $\Delta t = T/n$,
  - (2) simulate identical and independent random increments ΔX_i on each subinterval,
  - (3) add the increments together, and get the same result statistically, that is in terms of probability density, as for any other subdivision.
- Such a process is said to be ***infinitely divisible*.**

Infinitely divisibility and Levy process

Proposition

- Let X_t be a Levy process. Then for every t, X_t has infinitely divisible distribution.
- Conversely, if F is a infinitely divisible distribution, then there exists a Levy process X_t such that the distribution of X_1 is given by F.

- The result implies a connection between
  - **Levy processes and**
  - **Infinitely divisible distributions**.

(5) Levy Process

- Using Lévy process instead of BM to model log returns has recently become popular.
- Motivations:
- First: The GBM BS model does not appear to describe reality well. Log returns appear to have a clearly nonnormal distribution, and the implied volatility does not appear to be a constant function of the strike price, but exhibits a skew or a smile.
- Lévy process widens the class of possible marginals to the class of infinitely divisible distributions, which is much more flexible than the class of normals.

(5) Levy Process

- Second: Lévy processes allow for sample paths with jumps.

12. Case Studies: Vasicek Model

Mean Reversion

- The GBM provides the foundation for modeling the dynamics for asset prices of many different securities, including stock prices.
- However, in some cases it is not justified to assume that the asset price evolve with a particular drift, or can deviate arbitrarily far from some kind of a representative value.
- Interest rate, exchange rate, and the prices of some commodities are examples for which GBM does not provide a good representation over the long time.

12. Case Studies: Vasicek Model

- For example, if the price of copper becomes high, copper mines would increase production in order to maximize profits. This would increase the supply of copper in the market, therefore decreasing its price to some equilibrium level.
- Moreover, the price of copper becomes high, consumers may look for substitutes, which would reduce the price to its equilibrium level.
- There is some kind of a long-term average level to which they return after deviating up or down. This behavior is known as Mean Reversion.

11. Case Studies: Vasicek Model

See Wang (Operations Research, 2006)

4. Bond Valuation According to Term Structure Models

Consider the problem of finding the fair price of a T-year zero-coupon bond with a face price of one (see Ninomiya and Tezuka 1996 and Morokoff 1998). Assume that the interest rate follows the Vasicek model (Vasicek 1977)

$$dr = a(b-r)dt + \sigma dB, \quad r(0) = r_0,$$

where a and b are constants (the reversion speed and reversion level, respectively), σ is volatility, and dB is the standard Brownian motion. According to the arbitrage pricing theory (see Hull 2002), the price of a discount bond at time 0 with maturity T is

$$P(T) = \mathbb{E}\left(\exp\left(-\int_0^T r(s)\,ds\right)\right) \tag{28}$$

under the risk-neutral measure. This problem can be solved analytically (Vasicek 1977). For the purpose of illustrating the impact of dimension reduction techniques, this example is appropriate.

If the interest rate deviates from its long-run average, the process will pushed back to equilibrium at a speed a.

12. Case Studies: Vasicek Model

Vasicek Model: Mean reversion

$$dr(t) = a\,(b-r)\,dt + \text{sigma}\,dB$$

- b is the long-run mean of the short rate process and a is interpreted as the rate of mean reversion.
- When $r(t) > b$, the drift term will be negative and push r down towards b.
- When $r(t) < b$, the drift term will be positive and push r up towards b.

Vasicek Model: Discretization

A standard simulation procedure is to discretize the time and to approximate the Vasicek model by the basic Euler approximation

$$r_j - r_{j-1} = a(b - r_{j-1})\Delta t + \sigma\sqrt{\Delta t}\, z_j,$$
$$z_j \sim N(0,1), \quad j = 1, \ldots, d, \quad (29)$$

where $\Delta t = T/(d+1)$ and r_j is the interest rate at $t_j = j\Delta t$. The discrete-time version of (28) is

$$P_{d+1}(T) = \mathbb{E}\left(\exp\left(-\Delta t \sum_{j=0}^{d} r_j\right)\right). \quad (30)$$

The resulting integral is d-dimensional. To make the time discretization error small enough, the dimension can be as large as thousands (Ninomiya and Tezuka 1996). The

Vasicek Model: Discretization

of dimension reduction techniques. Note that (29) can be written as

$$r_j - b = (r_{j-1} - b)\beta + \sigma\sqrt{\Delta t}z_j, \qquad j = 1, \ldots, d,$$

where $\beta = 1 - a\Delta t$. This leads in a matrix notation to

$$(r_1, \ldots, r_d)^T = \sigma A(z_1, \ldots, z_d)^T + (H_1, \ldots, H_d)^T, \tag{31}$$

where

$$A = (a_{kj}) = \sqrt{\Delta t}\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \beta & 1 & 0 & \cdots & 0 \\ \beta^2 & \beta & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta^{d-1} & \beta^{d-2} & \beta^{d-3} & \cdots & 1 \end{pmatrix} \tag{32}$$

and $H_j = b + (r_0 - b)\beta^j$. If $a = 0$, then $\beta = 1$, and the

Vasicek Model: Discretization

Other methods to generate the paths of the interest rates are possible. Let

$$\mathbf{V} := AA^T$$

be the covariance matrix of the random vector $(r_1, \ldots, r_d)^T$, where A is given by (32). For an arbitrary matrix $\tilde{A}$ satisfying $\tilde{A}\tilde{A}^T = \mathbf{V}$, let the interest rates be generated by

$$(r_1, \ldots, r_d)^T = \sigma \tilde{A}(z_1, \ldots, z_d)^T + (H_1, \ldots, H_d)^T. \tag{35}$$

Vasicek Model: PCA path simulation

generation (35) is more suited to QMC. The ideas of BB and PCA remain the same as in the previous section. The matrix corresponding to PCA is

$$A^{\mathrm{PCA}} = WD^{1/2}, \tag{36}$$

with D being a diagonal matrix of eigenvalues of $\mathbf{V}$ arranged in nonincreasing order and the columns of W being the corresponding eigenvectors of unit length.

Question:

How to generate the interest rate paths by BB?

12. Case Studies: Cox-Ingersoll-Ross Model

See Wang (working paper, 2012)

Our first example is from Glasserman et al. (1999) about the interest rate model of Cox, Ingersoll and Ross (1985). The risk-neutral process for the interest rate r_t in CIR model is

$$dr_t = \kappa(a - r_t)\,dt + \sigma\sqrt{r_t}\,dW_t, \tag{21}$$

where κ, a, σ are constants and W_t is the standard Brownian motion. Our purpose is to calculate the fair price of a discount bond at time 0 with maturity T, which is given by the expectation

$$\mathbb{E}\left(\exp\left(-\int_0^T r_t\,dt\right)\right), \tag{22}$$

Square root diffusion process

There is no explicit solution to the SDE

12. Case Studies: Cox-Ingersoll-Ross Model

- The difference between the CIR model and the Vasicek model is that the volatility in the CIR model is proportional to the square root of the short rate rather than being constant.
- Because of this fact, the short rate can never be negative. The reason is that the volatility (i.e. sigma sqrt(r(t))) is very small if r(t) is near zero, so the drift will dominate the change in r(t), pushing it upwards towards the level a. This prevents r(t) from taking negative values.

12. Case Studies: Cox-Ingersoll-Ross Model

A simple Euler discretization of the CIR model

$$r(t_{i+1}) = r(t_i) + \kappa(a - r(t_i))[t_{i+1} - t_i] + \sigma\sqrt{r(t_i)^+}\sqrt{t_{i+1} - t_i}\, Z_{i+1},$$

$$\text{where } Z_i \overset{\text{iid}}{\sim} N(0,1).$$

Time-dependent coefficient a(t) is possible.

It is possible to develop BB-like and PCA-like path simulation methods for the CIR model.

Please try!

Assignment 2 (Dead Line: March 31, 2020)

1. Prove the Brownian bridge theorem.

2. Prove that $(B(t_1),\ldots,B(t_n))^T \sim N(0,C)$, where $B(t_1),\ldots,B(t_n)$ are the values of BM at times t_1, …, t_n, and the entries of C are given by

$$c_{i,j} = \min(t_i, t_j).$$

3. **Prove that if** $AA^T = C$, **then** $BB^T = C$, **if and only if** B **can be written as** $B = AU$ **for some orthogonal matrix** U.

Assignment 2 (Dead Line: March 31, 2020)

4. Find the generating matrix A (analytically or numerically) corresponding to the Brownian bridge construction, i.e., the matrix A in

$$(\boldsymbol{B}(\boldsymbol{t}_1),\ldots,\boldsymbol{B}(\boldsymbol{t}_n))^T = \boldsymbol{AZ}, \textbf{ with } \boldsymbol{AA}^T = \boldsymbol{C} \textbf{ and } \boldsymbol{Z} \sim N(0,\boldsymbol{I}),$$

with the values $B(t_i)$ generated by BB.

5. Write programs to generate BM for n being a power of 2, and being not a power of 2, respectively.

Assignment 2 (Dead Line: March 31, 2020)

6. **Derive analytical formulas for the cumulative explained variability in the random walk construction, PCA (or even in BB construction).**

7. **Compare the cumulative explained variability in random walk construction, BB and PCA constructions in dimension 16, 64 and 256.**

8. **Suppose** $\mathrm{C} = \mathrm{AA}^{\mathrm{T}}, \Sigma = \mathrm{BB}^{\mathrm{T}}$. **Prove that**

$$(\mathrm{C} \otimes \Sigma) = (\mathrm{A} \otimes \mathrm{B})(\mathrm{A} \otimes \mathrm{B})^{\mathrm{T}}.$$

Assignment 2 (Dead Line: March 31, 2020)

9. Let $X \sim \mathrm{BM}(\mu, \Sigma)$ and

$$\boldsymbol{Y} := (\boldsymbol{X}_1(\boldsymbol{t}_1), \ldots, \boldsymbol{X}_1(\boldsymbol{t}_n), \boldsymbol{X}_2(\boldsymbol{t}_1), \ldots, \boldsymbol{X}_2(\boldsymbol{t}_n), \cdots, \boldsymbol{X}_d(\boldsymbol{t}_1), \ldots, \boldsymbol{X}_d(\boldsymbol{t}_n))^T.$$

Prove that the covariance matrix of Y is $(\Sigma \otimes C)$.

10. How to obtain the eigenvalues and eigenvector of the matrix $(C \otimes \Sigma)$ from these of matrix C and matrix $\sum$.

11*. What is the difference in generating multi-dimensional BM between "one-step PCA" and "two-step-PCA" (i.e., use PCA for both matrix C and matrix Σ)?

Projects

Using MC simulation to price European or exotic options (e.g. Asian options, barrier options, lookback options, interest rate derivatives, etc.) or other derivatives under one or several of the following models:

(1) Black-Scholes model;

(2) The jump-diffusion model;

(3) Stochastic volatility model;

(4) Variance Gamma model;

(5) Normal inverse Gaussian model;

(6) Some other models (say, interest rate models).

The End of Chapter 4