Linnæus University
Sweden

Degree project

Stock-Price Modeling by the Geometric Fractional Brownian Motion

A View towards the Chinese Financial Market

Author: Zijie Feng
Supervisor: Roger Pettersson
Examiner: Astrid Hilbert
Date: 2018-10-08
Level: Bachelor

Department Of Mathematics

Abstract

As an extension of the geometric Brownian motion, a geometric fractional Brownian motion (GFBM) is considered as a stock-price model. The modeled GFBM is compared with empirical Chinese stock prices. Comparisons are performed by considering logarithmic-return densities, autocovariance functions, spectral densities and trajectories. Since logarithmic-return densities of GFBM stock prices are Gaussian and empirical stock logarithmic-returns typically are far from Gaussian, a GFBM model may not be the most suitable stock price model.

Key words: geometric fractional Brownian motion, fractional Brownian motion, fractional Gaussian noise, Hurst exponent.

Contents

1 Preliminaries **4**
- 1.1 Stock price . . . 4
- 1.2 Logarithmic return . . . 4
- 1.3 Stationarity . . . 5
- 1.4 Gaussian distribution and Gaussian process . . . 5
- 1.5 Geometric Brownian motion model . . . 7
- 1.6 Geometric fractional Brownian motion model . . . 7

2 Simulation Steps and Theoretical Analysis **14**
- 2.1 Simulation of fractional Gaussian noise . . . 14
  - 2.1.1 The Cholesky method . . . 15
  - 2.1.2 The Davies and Harte method . . . 16
  - 2.1.3 Examples . . . 17
- 2.2 Simulated fractional Brownian motion . . . 22
- 2.3 Simulated stock prices . . . 23

3 Parameter Estimation **25**
- 3.1 The Hurst exponent . . . 25
  - 3.1.1 Rescaled range analysis (RS) . . . 25
  - 3.1.2 Periodogram method (PE) . . . 27
- 3.2 Examples . . . 28
- 3.3 Volatility and drift . . . 29
- 3.4 Examples . . . 29

4 Case Analysis **31**
- 4.1 Application 1 . . . 31
- 4.2 Application 2 . . . 35
- 4.3 Application 3 . . . 37
- 4.4 Summary . . . 39

5 Conclusion **40**
- 5.1 Result . . . 40
- 5.2 Further work . . . 40

6 Appendix **42**

Introduction

Brownian motion (BM) is a random type of particles' motion suspended in a fluid, discovered in 1827 and named after the British botanist Robert Brown. In 1900, the French mathematician Louis Bachelier introduced a BM model of stock prices as part of his PhD dissertation *Théorie de la spéculation*. Somewhat later, Albert Einstein, Marian Smoluchowski and many other scientists contributed a lot to the illustration of the principle of BM, as described by *Hänggi and Marchesoni* [1]. It was not until 1918 a precise definition of the BM was given by the American mathematician and philosopher Norbert Wiener.

Built on the randomness of the BM, it evolves into the geometric Brownian motion (GBM) model, an important model frequently used for stock prices. The most famous application is relevant to the Black-Scholes model (see Section 8.3 in *Capinski and Zastawniak* [2]), which is named after Fischer Black, Myron Scholes but also developed by Robert Cox Merton. Many other scholars have validated and applied the GBM model in stock markets at different countries and regions. Some detailed applications and accounts about the GBM model are given by many scholars such as *Reddy and Childers* [4], *Hu* [20], *Zhang et al.* [21] and *Gao* [22].

In financial mathematics, the application of the GBM model is a start for stock-price modeling. It may need to be modified since the increments of real stocks are not mutually independent as in the assumption of the GBM model defined in [3] and [15]. To improve the accuracy of stock-price modeling, *Mandelbrot and Van Ness* [5] introduced the definition of fractional Brownian motion or fractal Brownian motion (FBM) in 1968 generalizing the BM by considering the Hurst exponent. Therefore, a new stock-price model, the geometric fractional Brownian motion (GFBM) model was published as an extension of the GBM model.

The GFBM model is more general than the GBM model. It can explain more situations about the change of stock prices. With four-decade development, the GFBM model is thoroughly analyzed in the stock market (see for example [8], [12] and [19]). However, few journal articles, reports or market analysis on the applicability of the GFBM model can be found for the Chinese financial market. In China, most of the existing papers considering the GFBM model skip the basic analysis of GFBM stock-price modeling, but discuss stock derivatives (e.g. option pricing). This thesis will give a detailed introduction to stock-price modeling by the GFBM model and thereby analyze its applications to example

data from the Chinese financial market. To evaluate the GFBM stock-price model, simulation and parameter estimations for the GFBM is included.

This thesis is divided into six chapters. In chapter 1, we will give some preliminary knowledge and introduce the definition of the GFBM model as a generalization of the GBM model. In chapter 2, we will show how to simulate a series of stock prices step by step, and then analyze the difference among models with different Hurst exponents. Chapter 3 illustrates how to estimate the parameters in the GFBM model. In chapter 4, GFBM modeling to empirical data will be given. Chapter 5 is the conclusion part and chapter 6 provides the entire codes written by Zijie Feng and applied for MATLAB. All the references are located at the end of this thesis.

Acknowledgements

I would like to thank my supervisor Roger Pettersson for all his help and suggestions. His patience and guidance encourages me and has left me a precious and valuable study experience in my bachelor thesis. I would also like to thank Zhizheng Wang and Achref Lemjid for improving the formulations and usage of LaTeX, which hopefully makes sure the thesis is readable.

Chapter 1

Preliminaries

1.1 Stock price

Following *Capinski and Zastawniak* [2], assets in finance can for simplicity be divided into risk-free assets and risky assets. The former includes a bank deposit or a bond issued by the government or other financial institutions. These kinds of assets are given with fixed returns and their future values are known beforehand.

The risky assets can be gold, foreign currency and other virtual assets of which their future price is unknown currently. A stock is a typical risky asset held by different investors. Its price can represent the unit value of a stock in the stock market. Since return of a risky asset can be considered as random, the price of a stock have unpredictability in the sense that we can not tell for sure its future prices.

Since stock prices are unpredictable, there are many properties in the returns of stock prices which still attract many mathematicians and financiers. In GFBM modeling of stock prices, the stock prices at a small time scale and a large time scale share a common property of fractional behavior determined by a specific parameter, a Hurst exponent. For GFBM, such property is denoted as self-similarity. We will give a mathematical description of self-similarity in Section 1.5 based on *Mandelbrot and Van Ness* [5], *Niu and Liu* [13]) and *Embrechts and Maejima* [14].

1.2 Logarithmic return

For a risk-free asset with compounding interest payment, its value at time t is

$$P(t) = P_0(1 + \frac{r}{m})^{mt}, \quad t = 0, \frac{1}{m}, \frac{2}{m}, ...$$

where m represents a compounding frequency. The constant P_0 is the initial price and r is also a constant, a periodic compounding interest rate. If the compounding frequency increases to infinite, we get

$$P(t) = P_0 e^{rt}, \quad t > 0. \tag{1.1}$$

Then r is a continuously compounding interest rate, *Capinski and Zastawniak* [2]. Observe that for the risk-free asset with price (1.1) and time span Δt,

$$\ln \frac{P(t+\Delta t)}{P(t)} = r. \tag{1.2}$$

Since logarithm is additive and convenient in mathematics, we define and use logarithmic return or continuously compounding return in financial models based on equation (1.2).

Definition 1. *For a given asset with price $S(t)$ at time t, the logarithmic return of such asset during $[t, t+\Delta t]$ is*

$$R_{\Delta t}(t) := \ln \frac{S(t+\Delta t)}{S(t)}.$$

1.3 Stationarity

A stochastic process for which any joint probability distribution does not change over time is said to be a stationary process or a strictly stationary process. Wide-sense stationarity is a weaker form of stationarity.

Definition 2. *A stochastic process $\{X(t), t \in \mathbb{R}\}$ is said to be wide-sense stationary (WSS) if its mean is constant and the correlation function only depends on the time lag, i.e.*

$$\mathrm{E}[X(t)] = \mu,$$

for all t for some constant μ and

$$\mathrm{Cov}[X(t), X(t+\tau)] = r(\tau),$$

for all t and τ for some function $\tau \mapsto r(\tau)$.

1.4 Gaussian distribution and Gaussian process

From *Miller and Childers* [3] and *Zhou et al.* [15], the Gaussian distribution or normal distribution is frequently used in statistics.

Definition 3. *A continuous random variable X for which its probability distribution function is*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad x \in \mathbb{R}, \tag{1.3}$$

is said to be Gaussian distributed, denoted as $X \sim N(\mu, \sigma^2)$.

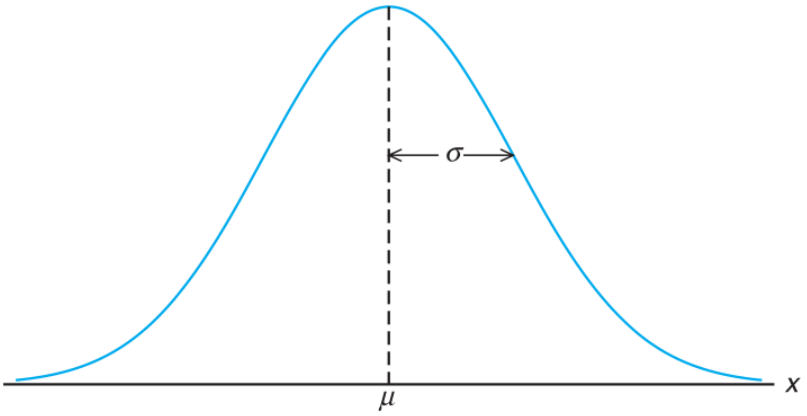

Figure 1.1: Normal curve

An illustration of the Gaussian density also known as a normal curve can be seen in Figure 1.1. Such probability distribution figure is symmetric and has maximal value at $x = \mu$. The standard deviation σ represents a mean deviation between the distribution and the mean. The Gaussian distribution has the property that if two Gaussian random variables are not correlated then they are also independent.

To define Gaussian processes, a definition of multivariate Gaussian distribution is necessary.

Definition 4. *A random vector* $\vec{X} = (X_1, ..., X_p)'$ *with mean vector* $\vec{\mu}$ *and covariance matrix* Σ *given by*

$$\vec{\mu} = \begin{pmatrix} \mathrm{E}[X_1] \\ \vdots \\ \mathrm{E}[X_p] \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \mathrm{Cov}[X_1, X_1] & \cdots & \mathrm{Cov}[X_1, X_p] \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}[X_p, X_1] & \cdots & \mathrm{Cov}[X_p, X_p] \end{pmatrix}.$$

is multivariate Gaussian distributed if its probability density function satisfies

$$f(\vec{x}) = \frac{1}{\sqrt{(2\pi)^p |\Sigma|^{\frac{1}{2}}}} \exp\left[-\frac{(\vec{x} - \vec{\mu})' \Sigma^{-1} (\vec{x} - \vec{\mu})}{2}\right], \quad \vec{x} \in \mathbb{R}^p.$$

Then we denote $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$.

A stochastic process $\{X(t), t \in \mathbb{R}\}$ is a collection of random variables. The following statement is a standard definition of a Gaussian process.

Definition 5. *If each finite collection* $\{X(t_1), ..., X(t_n)\}$ *of random variables of a stochastic process is multivariate Gaussian distributed, the stochastic process is said to be a Gaussian process.*

Here we can observe that any WSS Gaussian process is also strictly stationary.

1.5 Geometric Brownian motion model

Definition 6. *A stochastic process* $\{B(t), t \geq 0\}$ *is a Brownian motion (BM) if and only if it satisfies:*

(i) For any time points $0 \leq t_1 \leq t_2 \leq t_3 \leq t_4$*, the increments* $B(t_4) - B(t_3)$ *and* $B(t_2) - B(t_1)$ *are independent.*

(ii) Each increment is a zero-mean Gaussian random variable with variance equals the difference in time:

$$B(t) - B(s) \sim N(0, t - s).$$

(iii) $B(0) = 0$.

The classical BM is also known as the Wiener process in honor of Wiener's contribution. Notice that the BM is a non-stationary Gaussian process but its increments $B(t+h) - B(t)$ is a stationary Gaussian process keeping t fixed and varying $h \geq 0$. A geometric Brownian motion (GBM) is a stochastic process $\{S(t), t \geq 0\}$ such that:

$$S(t) = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma B(t)}, \quad t \geq 0. \tag{1.4}$$

where $\{B(t), t \geq 0\}$ is a BM. The GBM model is a widely used stock-price model, see for instance *Rostek and Schöbel* [12]. The constant S_0 means the initial stock price. The parameters μ and σ represent the drift and the volatility of the stock respectively.

Following Definition 1, the logarithmic return of a GBM modeled stock during time span Δt for the GBM model is

$$\begin{aligned} R_{\Delta t}(t) &= \ln \frac{S(t + \Delta t)}{S(t)} \\ &= (\mu - \frac{\sigma^2}{2})\Delta t + \sigma[B(t + \Delta t) - B(t)], \quad t \geq 0. \end{aligned}$$

Moreover, since $B(t+\Delta t) - B(t) \sim N(0, \Delta t)$, the distribution of the logarithmic return for the GBM can be presented as

$$R_{\Delta t}(t) \sim N((\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2 \Delta t).$$

1.6 Geometric fractional Brownian motion model

The fractional BM (FBM) is a generalization of the BM. Compared with the BM, the increments of the FBM need not to be independent by the definition by *Mandelbrot and Van Ness* [5].

Definition 7. *The FBM is a zero-mean Gaussian process* $\{B_H(t), t \geq 0\}$ *which starts at zero, with autocovariance function*

$$\mathrm{Cov}[B_H(t), B_H(t + \Delta t)] = \frac{1}{2}(|t|^{2H} + |t + \Delta t|^{2H} - |\Delta t|^{2H}), \tag{1.5}$$

where H *is a parameter such that* $H \in (0, 1)$.

The parameter H is called the Hurst exponent or Hurst index. The BM is a FBM with $H = 0.5$. The covariance between the FBM and its increments is

$$\begin{aligned}
&\mathrm{Cov}[B_H(t+\Delta t) - B_H(t), B_H(t)] \\
&= \mathrm{Cov}[B_H(t+\Delta t), B_H(t)]] - \mathrm{Cov}[B_H(t), B_H(t)] \\
&= \frac{1}{2}(|t+\Delta t|^{2H} + |t|^{2H} - |\Delta t|^{2H}) - \frac{1}{2}(2|t|^{2H}) \\
&= \frac{1}{2}(|t+\Delta t|^{2H} - |t|^{2H} - |\Delta t|^{2H}).
\end{aligned}$$

If for instance $H \in (0, 0.5)$, $x \mapsto x^{2H}$ is strictly concave for $x \geq 0$. So

$$\begin{aligned}
&|t|^{2H} + |\Delta t|^{2H} \\
&= |\frac{t}{t+\Delta t}(t+\Delta t) + \frac{\Delta t}{t+\Delta t} \cdot 0|^{2H} + |\frac{\Delta t}{t+\Delta t}(t+\Delta t) + \frac{t}{t+\Delta t} \cdot 0|^{2H} \\
&> \frac{t}{t+\Delta t}|t+\Delta t|^{2H} + \frac{\Delta t}{t+\Delta t} \cdot 0^{2H} + \frac{\Delta t}{t+\Delta t}|t+\Delta t|^{2H} + \frac{t}{t+\Delta t} \cdot 0^{2H} \\
&= |t+\Delta t|^{2H}.
\end{aligned}$$

Figure 1.2 is a graphical motivation that

$$\frac{\Delta t}{t+\Delta t} \cdot 0^{2H} + \frac{t}{t+\Delta t}|t+\Delta t|^{2H} < |t|^{2H}. \tag{1.6}$$

where the strictly convexity of $x \mapsto |x|^{2H}$ for $H \in (0, 0.5)$ was used. By the same reasoning,

$$\frac{\Delta t}{t+\Delta t}|t+\Delta t|^{2H} + \frac{t}{t+\Delta t} \cdot 0^{2H} < |t|^{2H}. \tag{1.7}$$

Combining (1.6) and (1.7) gives

$$|t+\Delta t|^{2H} < |t|^{2H} + |\Delta t|^{2H}.$$

Similarly for the case $H \in (0.5, 1)$ for which $x \mapsto |x|^{2H}$ is strictly convex.

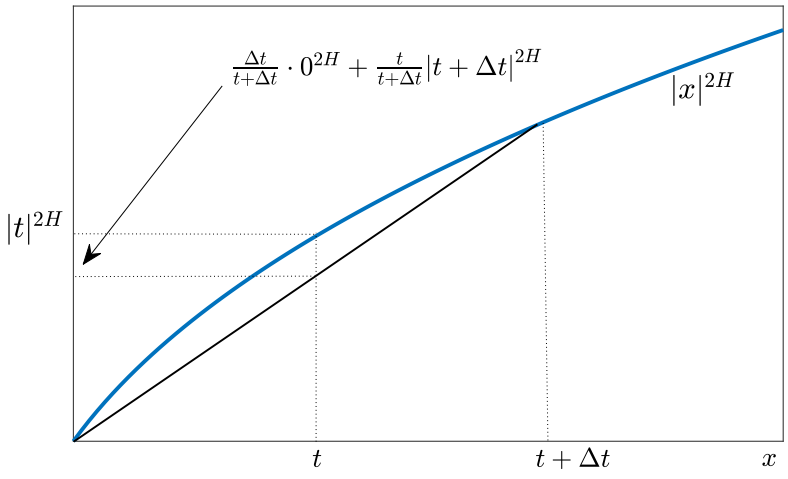

Figure 1.2: $\frac{\Delta t}{t+\Delta t} \cdot 0^{2H} + \frac{t}{t+\Delta t}|t+\Delta t|^{2H} < |t|^{2H}$, $H \in (0, 0.5)$.

We obtain:

(i) If $H \in (0, 0.5)$, the increments of FBM are negatively correlated, in particular for $\Delta t > 0$,

$$\mathrm{Cov}[B_H(t + \Delta t) - B_H(t), B_H(t)] < 0.$$

(ii) If $H = 0.5$, then $B_H(t)$ is a BM. The increments $B_H(t + \Delta t) - B_H(t)$ of FBM are thereby independent, in particular for $\Delta t > 0$,

$$\mathrm{Cov}[B_H(t + \Delta t) - B_H(t), B_H(t)] = 0.$$

(iii) If $H \in (0.5, 1)$, the increments of FBM are positively correlated, in particular for $\Delta t > 0$,

$$\mathrm{Cov}[B_H(t + \Delta t) - B_H(t), B_H(t)] > 0.$$

Now we will discuss self-similarity. We first need to be clear of the meaning of equality in distribution. Two random variables X and Y are said to be equal in distribution, denoted by $X \overset{\mathrm{d}}{=} Y$, if their cumulative distribution functions are the same, that is $\mathrm{P}(X \leq t) = \mathrm{P}(Y \leq t)$ for all t. Learning from *Niu and Liu* [13] and *Embrechts and Maejima* [14], self-similarity is defined as follows.

Definition 8. *A stochastic process $\{X(t), t \geq 0\}$ is said to be self-similar if for any $a > 0$, there exists $b > 0$ such that $X(at) \overset{d}{=} bX(t)$. Especially, if for arbitrary $a > 0$, $X(at) \overset{d}{=} a^H X(t)$ for an $H > 0$, not depending on a.*

The self-similarity is also known with the Hurst exponent. According to auto-covariance function (1.5), the variance of the FBM can be calculated as

$$\begin{aligned} \mathrm{Var}[B_H(t)] &= \mathrm{Cov}[B_H(t), B_H(t)] \\ &= \frac{1}{2}(|t|^{2H} + |t + 0|^{2H} - |0|^{2H}) \\ &= |t|^{2H}. \end{aligned}$$

Therefore, a FBM is a zero-mean Gaussian process with variance at time t equal to $|t|^{2H}$ which is denoted as $B(t) \sim N(0, |t|^{2H})$. According to the probability density function (1.3), its cumulative distribution is

$$P[B_H(t) \leq x] = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}|t|^H} \exp(-\frac{u^2}{2|t|^{2H}})du.$$

Hence,

$$\begin{aligned} P[B_H(at) \leq x] &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}|at|^H} \exp(-\frac{u^2}{2|at|^{2H}})du \\ &= \frac{1}{\sqrt{2\pi}|at|^H} \int_{-\infty}^{x} \exp\,(-\frac{u^2}{2|at|^{2H}})du \\ &= \frac{1/a^H}{\sqrt{2\pi}|t|^H} \int_{-\infty}^{x} \exp(-\frac{(u/a^H)^2}{2|t|^{2H}})du \\ &= \frac{1}{\sqrt{2\pi}|t|^H} \int_{-\infty}^{x/a^H} \exp(-\frac{s^2}{2|t|^{2H}})ds \\ &= P[B_H(t) \leq x/a^H] \\ &= P[a^H B_H(t) \leq x]. \end{aligned}$$

In the light of Definition 8 and the computations above, FBM is a Gaussian self-similar process, such that

$$B_H(at) \overset{\mathrm{d}}{=} a^H B_H(t).$$

This means that the FBM has a scale invariance which means that we can always find similar characteristics and structures when we observe different instances of FBM with different scales.

In financial mathematics, the GFBM model is a generalization of the GBM model (1.4), where $B(t)$ is replaced by $B_H(t)$. Thus, following Definition 7 and function (1.4), the GFBM model is

$$S_H(t) = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma B_H(t)}, \quad t \geq 0. \tag{1.8}$$

For the GFBM model, the logarithmic returns can be represented as

$$R_{H,\Delta t}(t) = \ln \frac{S_H(t+\Delta t)}{S_H(t)} = (\mu - \frac{\sigma^2}{2})\Delta t + \sigma X_{H,\Delta t}(t), \tag{1.9}$$

where

$$X_{H,\Delta t}(t) := B_H(t+\Delta t) - B_H(t), \quad t \geq 0, \tag{1.10}$$

is the increment of BM, called fractional Gaussian noise (FGN) during the time span Δt. Since $\mathrm{E}[B_H(t+\Delta t)] = \mathrm{E}[B_H(t)] = 0$, we have that $\mathrm{E}[X_H(t)] = 0$ and

$$\mathrm{E}[R_{H,\Delta t}(t)] = (\mu - \frac{\sigma^2}{2})\Delta t. \tag{1.11}$$

To get a step further, the variance of the FGN is

$$\begin{aligned}
\mathrm{Var}[X_{H,\Delta t}(t)] &= \mathrm{Var}[B_H(t+\Delta t) - B_H(t)] \\
&= \mathrm{Var}[B_H(t+\Delta t)] + \mathrm{Var}[B_H(t)] - 2\mathrm{Cov}[B_H(t), B_H(t+\Delta t)] \\
&= |t+\Delta t|^{2H} + |t|^{2H} - 2 \cdot \frac{1}{2}(|t|^{2H} + |t+\Delta t|^{2H} - |\Delta t|^{2H}) \\
&= |\Delta t|^{2H}.
\end{aligned}$$

The variance of the returns is then obtained as

$$\mathrm{Var}[R_{H,\Delta t}(t)] = \sigma^2 \mathrm{Var}[X_{H,\Delta t}(t)] = \sigma^2 |\Delta t|^{2H}. \tag{1.12}$$

Note that a zero-mean standardization of logarithmic returns in the GFBM model is a rescaled FGN:

$$\frac{R_{H,\Delta t}(t) - E[R_{H,\Delta t}(t)]}{\sqrt{\mathrm{Var}[R_{H,\Delta t}(t)]}} = \frac{1}{|\Delta t|^H} X_{H,\Delta t}(t). \tag{1.13}$$

In addition, we can deduce the autocovariance function of the FGN with lag τ from autocovariance function (1.5) and relation (1.10):

$$\begin{aligned}
&\mathrm{Cov}[X_{H,\Delta t}(t+\tau), X_{H,\Delta t}(t)] \\
&= \mathrm{E}[X_{H,\Delta t}(t+\tau), X_{H,\Delta t}(t)] - \mathrm{E}[X_{H,\Delta t}(t+\tau)]\mathrm{E}[X_{H,\Delta t}(t)] \\
&= \mathrm{E}[X_{H,\Delta t}(t+\tau), X_{H,\Delta t}(t)] - 0 \\
&= \mathrm{E}[(B_H(t+\Delta t+\tau) - B_H(t+\tau))(B_H(t+\Delta t) - B_H(t))] \\
&= \mathrm{E}[B_H(t+\Delta t+\tau)B_H(t+\Delta t)] - \mathrm{E}[B_H(t+\tau)B_H(t+\Delta t)] \\
&- \mathrm{E}[B_H(t+\Delta t+\tau)B_H(t) + \mathrm{E}[B_H(t+\tau)B_H(t)] \\
&= \frac{1}{2}(|t+\Delta t+\tau|^{2H} + |t+\Delta t|^{2H} - |\tau|^{2H}) \\
&- \frac{1}{2}(|t+\tau|^{2H} + |t+\Delta t|^{2H} - |\tau-\Delta t|^{2H}) \\
&- \frac{1}{2}(|t+\Delta t+\tau|^{2H} + |t|^{2H} - |\Delta t+\tau|^{2H}) \\
&+ \frac{1}{2}(|t+\tau|^{2H} + |t|^{2H} - |\tau|^{2H}) \\
&= \frac{1}{2}(|\tau-\Delta t|^{2H} - 2|\tau|^{2H} + |\tau+\Delta t|^{2H}) =: c_{H,\Delta t}(\tau) \qquad (1.14)
\end{aligned}$$

Notice that the autocovariance $\mathrm{Cov}[X_{H,\Delta t}(t+\tau), X_{H,\Delta t}(t)]$ does not depend on t, but only depends on the Hurst exponent, the time lag τ and the time span Δt for the FBM. Different from FBM, FGN for a fixed Δt is a stationary Gaussian process. The distribution of the FGN can be completely specified by its autocovariance. For $H \in (0.5, 1)$, by the strict convexity of $x \mapsto |x|^{2H}$,

$$\begin{aligned}
|\tau|^{2H} &= |\frac{1}{2}(\tau-\Delta t) + \frac{1}{2}(\tau+\Delta t)|^{2H} \\
&< \frac{1}{2}|\tau-\Delta t|^{2H} + \frac{1}{2}|\tau+\Delta t|^{2H}
\end{aligned}$$

giving $c_{H,\Delta t}(\tau) > 0$ for all τ. Figure 1.3 illustrates the strict convexity of $|\tau|^{2H}$ for $H \in (0.5, 1)$.

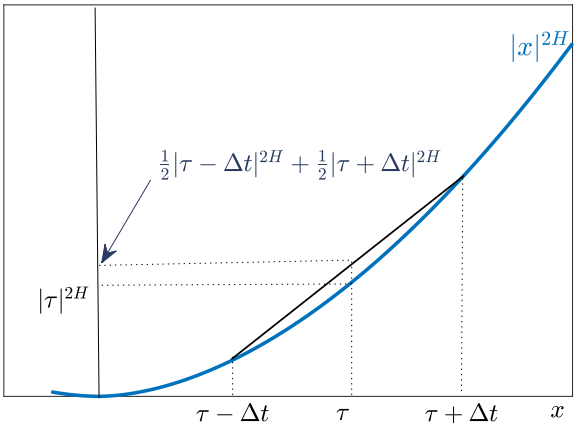

Figure 1.3: Graphical motivation of $|\tau|^{2H} < \frac{1}{2}|\tau-\Delta t|^{2H}+\frac{1}{2}|\tau+\Delta t|^{2H}$ for $H \in (0.5, 1)$.

For $H \in (0, 0.5)$, by the strict concavity of $x \mapsto |x|^{2H}$ for $x \geq 0$, for $\Delta t > 0$ and $\tau \geq \Delta t$,

$$\begin{aligned} |\tau|^{2H} &= |\frac{1}{2}(\tau-\Delta t)+\frac{1}{2}(\tau+\Delta t)|^{2H} \\ &> \frac{1}{2}|\tau-\Delta t|^{2H}+\frac{1}{2}|\tau+\Delta t|^{2H}, \end{aligned}$$

giving $c_{H,\Delta t}(\tau) < 0$ for all $\tau \geq \Delta t$. Figure 1.4 illustrates the strict concavity of $|\tau|^{2H}$ for $H \in (0, 0.5)$.

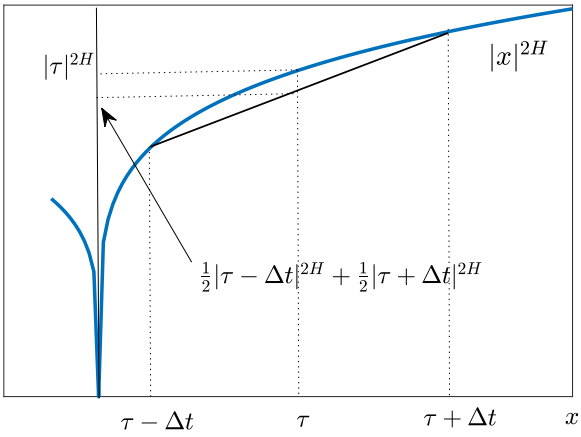

Figure 1.4: Graphical motivation of $|\tau|^{2H} > \frac{1}{2}|\tau-\Delta t|^{2H}+\frac{1}{2}|\tau+\Delta t|^{2H}$, $\tau \geq \Delta t$, for $H \in (0, 0.5)$.

We also have

$$c_{H,\Delta t}(0) = \frac{1}{2}(|\Delta t|^{2H}+|\Delta t|^{2H}) > 0.$$

By continuity, $c_{H,\Delta t}(\tau) > 0$ for τ close to zero, and $c_{H,\Delta t}(\tau) < 0$ for τ larger than some τ_0, $\tau_0 < \Delta t$. We get in particular that

$$c_{H,\Delta t}(0) = |\Delta t|^{2H} > 0.$$

To sum up,

1. for $H \in (0, 0.5)$: $\quad c_{H,\Delta t}(k\Delta t) < 0, \quad k = \pm 1, \pm 2, ...$
2. for $H = 0.5$: $\qquad c_{H,\Delta t}(k\Delta t) = 0, \quad k = \pm 1, \pm 2, ...$
3. for $H \in (0.5, 1)$: $\quad c_{H,\Delta t}(k\Delta t) > 0, \quad k = \pm 1, \pm 2, ...$

Here it was also used that autocovariance functions are symmetric.

Chapter 2

Simulation Steps and Theoretical Analysis

To evaluate a GFBM stock-price model it is natural to be able to simulate GFBM paths. Then empirical stock prices can be visually compared with simulated FBM paths. Therefore this section is devoted to simulation of the FBM from which by (1.8)-(1.10) we can simulate GFBM paths.

The procedure of stock simulation of a GFBM model can be divided into three different steps in the view of *Øksendal* [8] and *Dieker* [9]. The first step is to simulate a finite sequence of FGN based on a given Hurst exponent and the autocovariance function of FGN. The second step is to create a sequence of FBM by taking the cumulative sum of our simulated FGN. Finally, the simulation of stock prices can be obtained from the GFBM model (1.8) with the help of the created FBM.

Remark For convenience to understand and perform analysis, we only simulate a series of daily stock prices for which we assume $\Delta t = 1$.

2.1 Simulation of fractional Gaussian noise

For a stationary Gaussian process $\{X(t), t \geq 0\}$ evaluated at time points $t_1, t_2, ..., t_n$, the random vector $\vec{X} = (X(t_1), X(t_2), ..., X(t_n))'$ is an n-dimensional Gaussian variable with $n \times n$ covariance matrix C given by

$$c_{ij} = \text{Cov}[X(t_i), X(t_j)], \quad i, j = 1, .., n.$$

For the particular case when $\{X(t), t \geq 0\}$ is a FGN and $t_1 = 1, ..., t_n = n$, by (1.14) with $\Delta t = 1$,

$$\begin{aligned} c_{ij} &= c_{H,1}(i-j) \\ &= \frac{1}{2}(|i-j-1|^{2H} - 2|i-j|^{2H} + |i-j+1|^{2H}). \end{aligned}$$

So

$$C = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_1 & a_0 & \cdots & a_{n-3} & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-2} & a_{n-3} & \cdots & a_0 & a_1 \\ a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{pmatrix},$$

where

$$a_k = \frac{1}{2}(|k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H}).$$

To simulate $X(t_1), X(t_2), ..., X(t_n)$, we find a decomposition of the covariance matrix C such that $C = LL'$, where L is an $n \times n$-dimensional matrix and L' its transpose. After that, an n-row vector V is needed where all the entries are independent and identically distributed (i.i.d.) standard Gaussian random variables. Then the n-dimensional random vector LV is Gaussian distributed with n-dimensional mean vector

$$\mathrm{E}[LV] = L\mathrm{E}[V] = L\vec{0} = \vec{0},$$

where $\vec{0} = (0, 0, ..., 0)'$ and covariance matrix

$$\begin{aligned} \mathrm{Cov}[LV] &= L\mathrm{Cov}[V]L' \\ &= LIL' \\ &= LL' \\ &= C, \end{aligned}$$

where I denotes the $n \times n$ dimensional identity matrix. This means that LV has the same distribution as $\vec{X} = (X(t_1), X(t_2), ..., X(t_n))'$.

2.1.1 The Cholesky method

The Cholesky decomposition is an algorithm which can separate a positive-definite matrix into the product of a lower triangular matrix and its conjugate transpose. It means that L is a lower triangular matrix of the type

$$L = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix}.$$

The following codes are used to simulate a series of FGN by Cholesky decomposition in MATLAB:

```
% input the length of FGN and Hurst exponent
function [X]=cholesky(n,h)
    % n*1 vector with each random element following N(0,1)
    V=normrnd(0,1,[n,1]);
    % covariance matrix of random vector
    sigma=zeros(n);
    % time span is 1 day
```

```
        dt=1;
        % start at time 1 and end at time n
        for t=1:n
            for s=1:n
                ds=t-s;
                % the autocovariance function of FGN
                c=(1/2)*(abs(ds+dt).^(2*h)...
                    -2*abs(ds).^(2*h)+abs(ds-dt).^(2*h));
                sigma(t,s)=c;
            end
        end
        % Cholesky decomposition
        L=chol(sigma,'lower');
        % the random vector
        X=L*V;
    end
```

2.1.2 The Davies and Harte method

This method is designed for simulating a series of FGN with $2n = 2^m$ elements (m is a positive integer). We can firstly embed the $n \times n$ covariance matrix C in a $2n \times 2n$ circulant matrix Σ. Such a circulant matrix is defined as

$$\Sigma = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} & 0 & a_{n-1} & \cdots & a_2 & a_1 \\ a_1 & a_0 & \cdots & a_{n-2} & a_{n-1} & 0 & \cdots & a_3 & a_2 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ a_2 & a_3 & \cdots & a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 & a_1 \\ a_1 & a_2 & \cdots & 0 & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{pmatrix}.$$

Since C is a symmetric positive definite matrix and thereby Σ is a symmetric semi-positive definite matrix, a spectral decomposition $\Sigma = Q\Lambda Q'$ can be found, where Λ is a diagonal matrix with all eigenvalues of Σ, and Q is the corresponding unitary matrix. Afterwards, defining a matrix

$$S = Q\Lambda^{\frac{1}{2}}Q'$$

we obtain

$$\begin{aligned} SS' &= Q\Lambda^{\frac{1}{2}}Q'(Q\Lambda^{\frac{1}{2}}Q')' \\ &= Q\Lambda^{\frac{1}{2}}Q'Q\Lambda^{\frac{1}{2}}Q' \\ &= Q\Lambda^{\frac{1}{2}}I\Lambda^{\frac{1}{2}}Q' \\ &= Q\Lambda Q' \\ &= \Sigma. \end{aligned}$$

By creating a $2n$-row vector W where all the entries are i.i.d. standard normal random variables, we observe that

$$\mathrm{E}[SW] = S\mathrm{E}[W] = S\vec{0} = \vec{0},$$

where $\vec{0}$ is a $2n$-dimensional vector of zeros. Furthermore,

$$\begin{aligned} \mathrm{Cov}[SW] &= S\mathrm{Cov}[W]S' \\ &= SIS' \\ &= SS' \\ &= \Sigma. \end{aligned}$$

A sample of the $2n$-dimensional FGN is obtained by the SW. The Davies and Harte method can be motivated by the speed with which Λ and Q can be obtained using a Fast Fourier Transform (FFT), [9].

The codes below implements a simplified Davies and Harte MATLAB simulation not using FFT:

```
% input the length of FGN and Hurst exponent
function [X]=DH_method(N,h)
    % N=2n is a power of 2
    n=N/2;
    % the empty circulant matrix
    sigma=zeros(N);
    % the first row of circulant matrix
    temp=zeros(1,N);
    temp(1)=(1/2)*(abs(0+1).^(2*h)...        % c(0)
        -2*abs(0).^(2*h)+abs(0-1).^(2*h));
    temp(n+1)=0;
    for i=2:n
        ds=i-1;
        % autocovariance function of FGN
        c=(1/2)*(abs(ds+1).^(2*h)...
            -2*abs(ds).^(2*h)+abs(ds-1).^(2*h));
        temp(i)=c;
        temp(N+2-i)=c;
    end
    for i=1:N
        % the filled circulant covariance matrix
        sigma(i,:)=circshift(temp,[0,i-1]);
    end
    % N*1 vector with each random element ¬ N(0,1)
    V=normrnd(0,1,[N,1]);
    % unitary matrix and diagonal matrix
    [Q,A]=eig(sigma);
    S=Q*(A.^(1/2))*Q';
    % the random vector
    X=S*V;
end
```

2.1.3 Examples

By the FGN-simulation methods, it is easy to visualize the relation among models with different Hurst exponents and the corresponding FGNs. In the forthcoming description, all the figures are for simplicity produced by the Cholesky method.

Fractional Gaussian noise

Figure 2.1 includes three example paths of FGNs with 1000 elements and different Hurst exponents $H = 0.2$, $H = 0.5$ and $H = 0.8$.

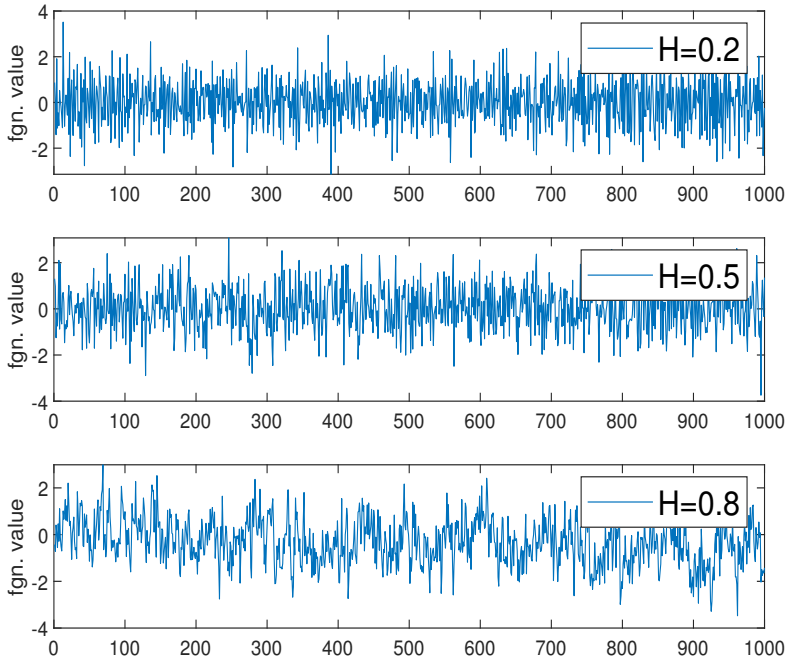

Figure 2.1: Simulated fractional Gaussian noises

The top path represents a simulated sequence of FGN with $H = 0.2$. It seems to be similar to the middle path, a simulated sequence of FGN with $H = 0.5$. However, they are actually different since the outcomes of the former path are negatively correlated. Such path flips up and down rapidly, which is in line with the discussion of the sign of $c_{H,\Delta t}(\tau)$. For FGN with $H = 0.5$, the outcomes are independent.

Different from the first two paths, for the bottom path which is a simulated sequence of FGN with $H = 0.8$, there obviously exists positive autocorrelation for adjacent time points and that is also in line with the discussion of the sign of $c_{H,\Delta t}(\tau)$.

Autocovariance

The Hurst exponent H is defined for $H \in (0, 1)$. To understand the qualitative behavior for H close to 0 and 1, we can consider the limits as $H \to 0$ and $H \to 1$. From the autocovariance function (1.14), if $H \to 0$, the autocovariances can be

operated as

$$\begin{aligned}
\lim_{H\to 0} c_{H,\Delta t}(0) &= \lim_{H\to 0} \frac{1}{2}(|1|^{2H} - 2|0|^{2H} + |-1|^{2H}) = 1, \\
\lim_{H\to 0} c_{H,\Delta t}(1) &= \lim_{H\to 0} \frac{1}{2}(|2|^{2H} - 2|1|^{2H} + |0|^{2H}) = -\frac{1}{2}, \\
\lim_{H\to 0} c_{H,\Delta t}(k) &= \lim_{H\to 0} \frac{1}{2}(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}) = 0, \quad k = 2, 3, ...
\end{aligned}$$

It means in particular that for H close to 0, adjacent outcomes are negatively correlated. Similarly, the results of $H \to 1$ are

$$\begin{aligned}
\lim_{H\to 1} c_{H,\Delta t}(k) &= \lim_{H\to 1} \frac{1}{2}(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}) \\
&= \frac{1}{2}[(k+1)^2 - 2k^2 + (k-1)^2] \\
&= 1, \quad k \geq 0.
\end{aligned}$$

It means that for H close to 1, the FGN is almost a perfectly correlated sequence for all outcomes: the process is constantly equal to zero. The case of $H = 0.5$ can be calculated using the similar operations

$$\begin{aligned}
c_{0.5,\Delta t}(0) &= \frac{1}{2}(|2| - 2|0| + |-1|) = 1, \\
c_{0.5,\Delta t}(k) &= \frac{1}{2}(|k+1| - 2|k| + |k-1|) = 0, \quad k = 1, 2, 3, ...
\end{aligned}$$

It means that for H equal to 0.5 the FGN is an i.i.d. zero-mean normal random variable sequence.

Now we discuss estimates for a FGN. We first start with a general discussion on estimates of mean and autocovariance functions for a general WSS process. For a sequence $(x(t_1), x(t_2), ..., x(t_n))'$ of a WSS time-series $\{X(t), t \geq 0\}$ with mean $E[X(t)] = \mu$ (μ is a number) and autocovariance function $c(\tau) = \text{Cov}[X(t), X(t+\tau)]$, we can estimate μ by

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x(t_i)$$

and the autocovariance function by

$$\hat{c}(\tau) = \begin{cases} \frac{1}{n-\tau}\sum\limits_{i=0}^{n-\tau-1}[x(t_i+\tau) - \hat{\mu}][x(t_i) - \hat{\mu}], & \tau = 0, ..., n-1, \\ \hat{c}(-\tau), & \tau = -1, ..., 1-n. \end{cases}$$

Figure 2.2 contains graphs of the autocovariances (1.14) for $H = 0.2$, $H = 0.5$ and $H = 0.8$ respectively and their estimates $\hat{c}_H$ for the simulated sequences illustrated in Figure 2.1.

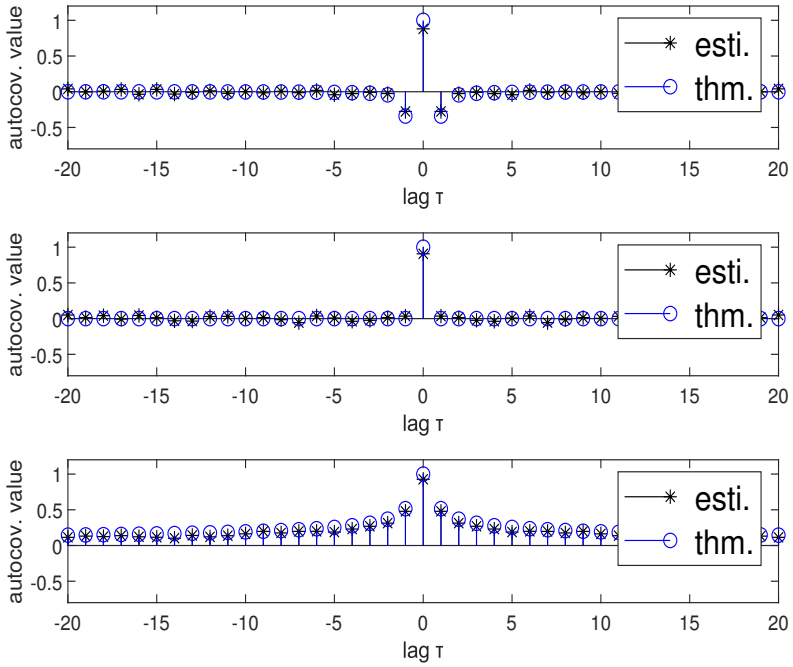

Figure 2.2: Autocovariances and their estimates

The stars represent the estimated autocovariances $\hat{c}_H$ from previous simulated FGNs. The circles represents the autocovariances of the theoretical FGNs. It is apparent that each estimated autocovariance is quite close to corresponding theoretical autocovariance with the same Hurst exponent.

In the first subfigure with $H = 0.2$, $\hat{c}_{0.2,\Delta t}(1)$ is very close to

$$c_{0.2,\Delta t}(1) = \frac{1}{2}(-2 \cdot 1 + 2^{0.4}) \approx -0.34,$$

but is approximately 0 at other non-zero lags. It means that there is a negative correlation between every adjacent elements and such a correlation is significantly smaller when $|\tau| > 1$.

The second subfigure with $H = 0.5$ confirms that the FGN is a sequence of uncorrelated random variables.

The final subfigure is very different from the previous two subfigures. There exists a strong positive correlation among all elements in the sequence. With the increase of lag τ, both the estimated and theoretical autocovariances are positive and decrease simultaneously.

Power spectral density

Besides the analysis of FGN in time domain, power spectral density or spectral density is another analysis method of FGN here in frequency domain. By the Wiener-Khinchin theorem (see Section 10.2 in [3]), the spectral density $S(f)$ and the autocovariance function $c_{H,\Delta t}(\tau)$ of a WSS stochastic process defined at discrete time points $t = 0, \pm 1, \pm 2, \pm 3, ...$ are Fourier transform pairs, with $S(f)$ given by

$$S(f) = \sum_{\tau=-\infty}^{\infty} e^{-i2\pi f\tau} c_{H,\Delta t}(\tau), \quad -\frac{1}{2} \leq f \leq \frac{1}{2}. \tag{2.1}$$

For a sequence $\vec{x} = (x(t_1), ..., x(t_n))'$ of a stationary time series $\{X(t), t = 0, \pm 1, \pm 2, ...\}$ the spectral density can be estimated by a periodogram

$$\hat{S}(f) = \frac{1}{n}\left|\sum_{\tau=0}^{n-1}[x(\tau) - \hat{\mu}]e^{-i2\pi f\tau}\right|^2, \quad -\frac{1}{2} \leq f \leq \frac{1}{2}. \tag{2.2}$$

Figure 2.3 includes spectral densities and periodograms for $H = 0.2$, $H = 0.5$ and $H = 0.8$ respectively. The spectral densities are based on equation (2.1) where the number of terms are 400 and autocovariances given by (1.14). For the periodograms, the simulated sequences illustrated in Figure 2.1 were used.

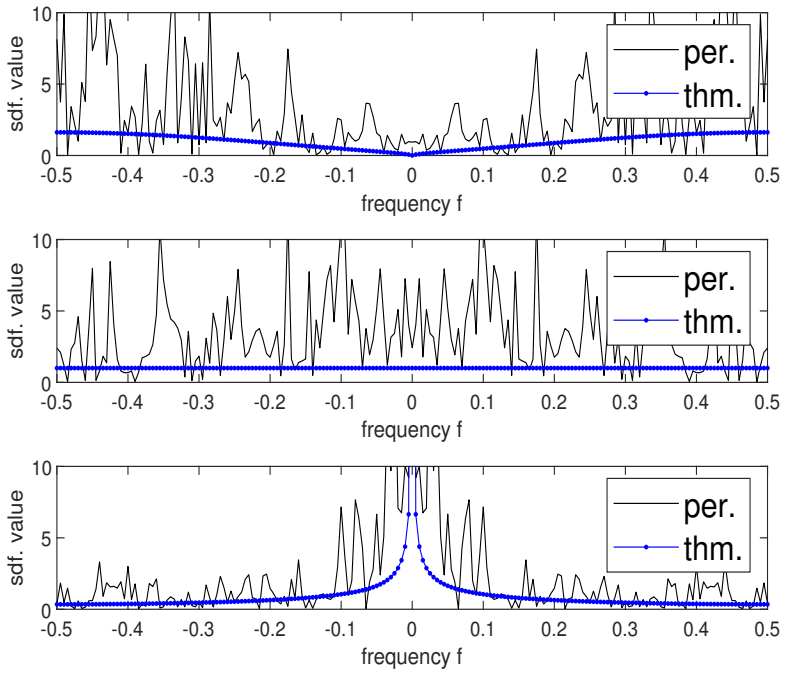

Figure 2.3: Spectral densities and periodograms

The solid paths represent the estimated spectral densities from simulated FGNs. Due to the randomness of simulated FGNs, periodograms are much rougher compared with the spectral densities from the theoretical FGNs (the dash-dot paths). However, the distributions of different frequencies are similar.

As before, the limiting cases $H \to 0$ and $H \to 1$ can show the extreme situations of the spectral densities of FGNs and illustrate the tendency.
For $H \to 0$,

$$\begin{aligned}
&\lim_{H\to 0} S_H(f) \\
&= e^{i2\pi f\cdot(-1)} \lim_{H\to 0} c_H(-1) + e^{i2\pi f\cdot 0} \lim_{H\to 0} c_H(0) + e^{i2\pi f\cdot 1} \lim_{H\to 0} c_H(1) \\
&= e^{-i2\pi f}(-\tfrac{1}{2}) + 1 + e^{i2\pi f}(-\tfrac{1}{2}) \\
&= 1 - \tfrac{1}{2}(e^{-i2\pi f} + e^{i2\pi f}) \\
&= 1 - \cos(2\pi f),
\end{aligned}$$

which is an increasing functions for $f \in [0, 0.5]$. It means that large frequencies dominate causing the processes to fluctuate quite wildly for $H \to 0$, in accordance with the top figure (with $H = 0.2$) in Figure 2.3.
For $H \to 1$,

$$\begin{aligned}
\lim_{H\to 1} S(f) &= \textstyle\sum_{\tau=-\infty}^{\infty} \lim_{H\to 1} c_H(\tau) \\
&= \textstyle\sum_{\tau=-\infty}^{\infty} e^{i2\pi f\cdot\tau} 1 \\
&= \delta(f) \\
&= \begin{cases} 1, & f = 0, \\ 0, & f \neq 0. \end{cases}
\end{aligned}$$

This means that the process has zero frequency, i.e., the process is constantly equal to zero as the bottom figure (with $H = 0.8$) in Figure 2.3 indicates.
For $H = 0.5$,

$$S_{0.5}(f) = e^{i2\pi f\cdot 0} c_H(0) = 1.$$

All frequencies are then equally weighted, the FGN is a Gaussian white noise which means that the underlying data is a sequence of independent standard normal variables here. The periodogram in the middle subfigure (with $H = 0.5$) of Figure 2.3 flips up and down around 1. It is similar to the above theoretical analysis.

2.2 Simulated fractional Brownian motion

According to relation (1.10) and $B_H(0) = 0$, for a given series of FGN, we can obtain a series of FBM at integer-time points by

$$B_H(t) = \sum_{k=1}^{t} X_H(k), \quad t = 1, 2, 3, ...n,$$

where $X_H(k) := X_{H,1}(k)$ is the FGN, and $B_H(t)$ is the FBM. Figure 2.4 contains three paths of FBMs generated from the FGNs in Figure 2.1.

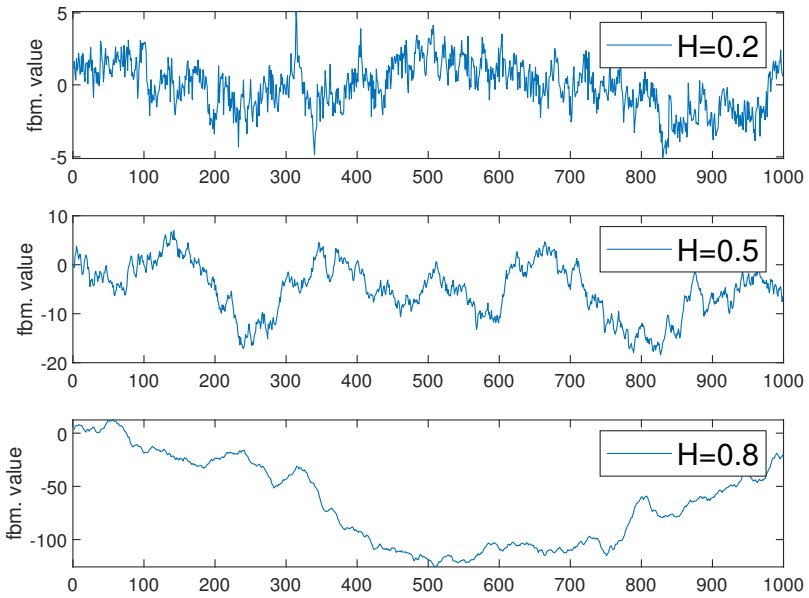

Figure 2.4: Simulated fractional Brownian motions

The first FBM ($H = 0.2$) has negatively correlated increments and it is much rough compared with the two other FBMs, especially compared with the FBM with $H = 0.8$. The range of a path is relatively small, here only from -5 to 5.

The bottom path, FBM with $H = 0.8$, nearly has highly dependent increments, which helps the path to get a dramatically big range, which here is from 20 to -130.

FBM with $H = 0.5$ looks like an intermediate of FBMs with $H = 0.2$ and with $H = 0.8$. It not only fluctuates frequently, but also has quite large range. It is a type of a one-dimensional random walk and moves up and down with the same probabilities.

2.3 Simulated stock prices

To simulate a series of stock prices, we use model (1.8). Assume for instance a drift $\mu = 0.0060$, a volatility $\sigma = 0.027$ and an initial stock price $S_0 = 10$. Figure 2.5 gives the output sequences based on those parameter values and the FBMs from Figure 2.4.

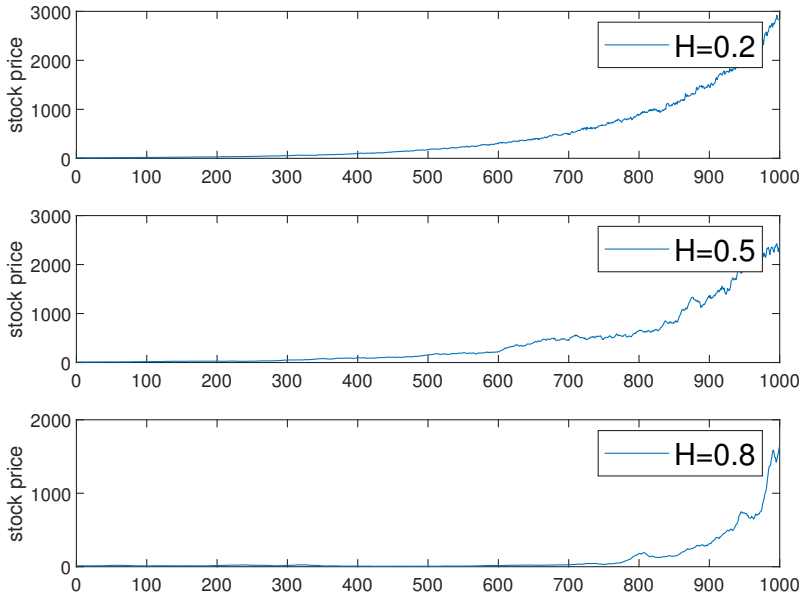

Figure 2.5: Simulated stock prices

Although the growth tendency of prices of the three example sequences of stocks in Figure 2.5 are similar, they are totally different with the development of n. Opposite to Figure 2.4, the simulated stock with $H = 0.2$ seems to at a large time scale grow more smoothly compared with the other stocks. That can be explained by the fact that the FGN for $H = 0.2$ fluctuates wildly around zero while its cumulative sum is more concentrated to zero than for $H = 0.5$ and $H = 0.8$ due to the negative correlations for $H = 0.2$. That makes the GFBM to be more exponential like with the randomness smoothed out.

On the contrary, the stock of $H = 0.8$ is an undulating path with relatively large periods of ups and downs. The cumulative sum of the FGN for $H = 0.8$ are at long periods increasing and decreasing due to the positive correlation of the FGN implying that the path looks to be more random at a large time scale.

Stock prices with $H = 0.5$ still seems like an intermediate between the first path and the third path. Its stock price seems to fluctuate more roughly than the stock price with $H = 0.2$ do at a large time scale, and have shorter periods of ups and downs than stock prices with $H = 0.8$.

Chapter 3

Parameter Estimation

In the GFBM model (1.8) there are parameters which need to be estimated on the basis of the given data. The parameters are the Hurst exponent H, volatility σ, and drift μ.

3.1 The Hurst exponent

At first, the Hurst exponent is necessary to be estimated. In this thesis, we will introduce two methods to estimate the Hurst exponent from a given sequence. In time domain, we use rescaled range analysis. In the frequency domain, we will use a periodogram method.

3.1.1 Rescaled range analysis (RS)

For a given time series X_1,X_2,...,X_N with a series of cumulative means

$$m_n = \frac{1}{n}\sum_{t=1}^{n} X_t, \quad n = 1, 2, ..., N,$$

we calculate the cumulative deviation series

$$Z_n = \sum_{t=1}^{n}[X_t - m_n], \quad n = 1, 2, ..., N,$$

and the range deviation series

$$R_n = \max(Z_1, Z_2, ..., Z_n) - \min(Z_1, Z_2, ..., Z_n), \quad n = 1, 2, ..., N.$$

Then we calculate the standard deviation series as

$$S_n = \sqrt{\frac{1}{n}\sum_{t=1}^{n}[X_t - m_n]^2}, \quad n = 1, 2, ..., N.$$

Define a sequence R_n/S_n called rescaled range deviation of a given time series. It is reasonable that rescaled range deviation would increase by n, since R_n will likely grow by n while S_n will likely converge to a number, at least if the time

series is WSS. Figure 3.1 shows an example for a FGN time series with $H = 0.5$ and $N = 1000$ points.

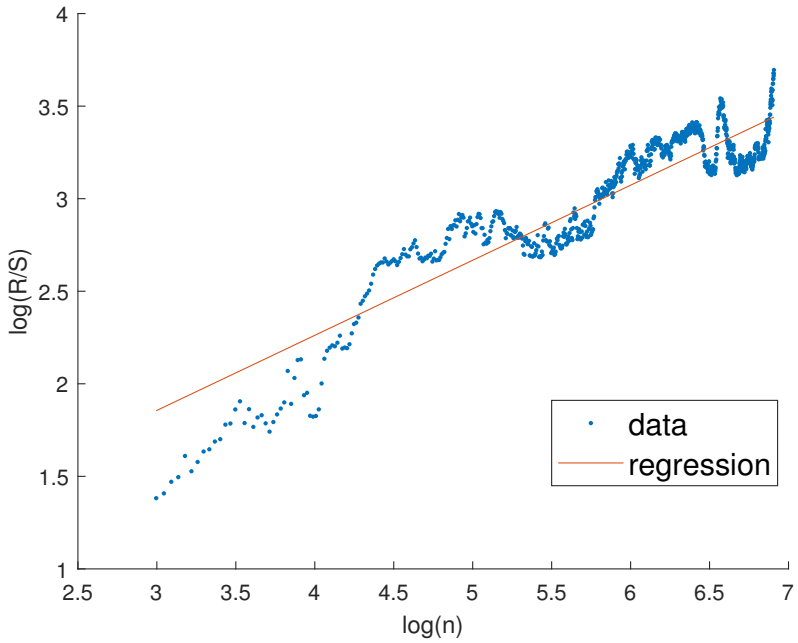

Figure 3.1: Linear regression

The Hurst exponent can represent the asymptotic behavior of the rescaled range as a function of the time span of a time series. For large n, such a relation can be represented as

$$\lim_{n\to\infty} n^{-H}\mathrm{E}[\frac{R_n}{S_n}] = C,$$

where C is a constant, [6]. It can be re-written for large n as

$$\mathrm{E}[\frac{R_n}{S_n}] \approx C \cdot n^H.$$

Thus, for a series of (R_n/S_n), we have

$$\ln(\frac{R_n}{S_n}) \approx \ln C + H \ln(n), \quad n = 1, 2, ...$$

The parameter H can be estimated by linear regression of $\ln(R_n/S_n)$ and $\ln(n)$, for which H is the slope.

The following codes achieve rescaled range analysis in MATLAB:

```
% input a series of FGN {X_t, t>0}
function [hurst]=RS(X)
    N=length(X);
    % skip the initial 20 points
    n=20;
    xvals=n: N;
    logx=log(xvals);
    yvals=zeros(1, length(xvals));
    for t=n: N
        tmpX=X(1:t);
        % deviation series with means
        Y=tmpX-mean(tmpX);
        % cumulative series
        Z=cumsum(Y);
        % range deviation series
        R=max(Z)-min(Z);
        % standard deviation series
        S=std(tmpX);
        yvals(t-(n-1))=R/S;
    end
    logy=log(yvals);
    p=polyfit(logx,logy,1);
    % Hurst exponent is the slope of linear-fit plot
    hurst=p(1);

    % the scatter figure of linear regression
    scatter(logx,logy,'.')
    hold on
    plot(logx, logx*hurst+p(2))
    xlabel('log(n)')
    ylabel('log(R/S)')
    legend('data','regression')
    hold off
end
```

3.1.2 Periodogram method (PE)

By *Liu et al.* [7], the spectral density of a FGN defined according to (1.10) with $\Delta t = 1$ can be rewritten as

$$\begin{aligned} S(f) &= \sum_{\tau=-\infty}^{\infty} e^{-i2\pi f\tau} c_H(\tau) \\ &= 4\sigma^2 C_H \sin^2(\pi f) \times \sum_{j=-\infty}^{\infty} \frac{1}{|f+j|^{2H+1}}, \quad -\frac{1}{2} \le f \le \frac{1}{2} \end{aligned} \tag{3.1}$$

where $C_H = \Gamma(2H+1)\sin(\pi H)/(2\pi)^{2H+1}$. Using Taylor expansion to equation (3.1), the corresponding spectral density is related to frequency by a power law:

$$S(f) \approx \sigma^2 C_H (2\pi)^2 |f|^{\gamma}, \quad -\frac{1}{2} \le f \le \frac{1}{2}, \tag{3.2}$$

where $\gamma \in (-1, 1)$ is the spectral exponent

$$\gamma = 1 - 2H.$$

By [7], the approximation $\hat{S}(f) \sim |f|^{\gamma}$ is valid for low frequences up to the size $N^{\frac{4}{5}}$, where N is the length of data. We can easily obtain an estimation of H by the slope of a linear regression from equation (3.2):

$$\ln \hat{S}(f) \approx \ln[\sigma^2 C_H (2\pi)^2] + \gamma \ln |f|.$$

Implemented codes of the periodogram method based on (2.2) in MATLAB are written below:

```
% Periodogram Method
function [hurst]=PE(X)
    N=length(X);
    % separate (-0.5,0.5) into 601 parts
    len=601;
    lags=linspace(-1/2,1/2, len);
    % estimated spectral density
    sdf=zeros(1,len);
    for f=1: len
        tau=0:N-1;
        xl=X-mean(X);
        xr=exp(-1i*2*pi*lags(f)*tau);
        cum=xr*xl;
        sdf(f)=(abs(cum)^2)/N;
    end
    mid=(len+1)/2;
    % during low-frequency part
    len=fix(mid^(4/5));
    x=lags(mid+1: mid+1+len);
    y=sdf(mid+1: mid+1+len);
    logx=log(x);
    logy=log(y);
    % spectral exponent
    gamma=polyfit(logx,logy,1);
    % Hurst exponent
    hurst=(1-gamma(1))/2;
end
```

3.2 Examples

We use both rescaled range analysis and the periodogram method to estimate the Hurst exponents of the simulated FGNs in Figure 2.1. The results are shown in Table 3.1.

H	Rescaled range analysis	Periodogram method
0.2	0.3438	0.2262
0.5	0.4655	0.4482
0.8	0.8586	0.7697

Table 3.1: Hurst exponent estimations

Notice that the rescaled range analysis and the periodogram method only provide rough values of Hurst parameters for our small samples of size $N = 1000$. Some methods with more accurate results can be found in [6] and [7].

3.3 Volatility and drift

When we have estimated the Hurst exponent of a series of logarithmic stock returns, the volatility needs to be estimated. After that, we can estimate the drift.

Assume we have a series of logarithmic returns $\{R_{H,\Delta t}(t_i), i = 1, ..., N\}$ with sample mean

$$\bar{R} = \frac{1}{N} \sum_{i=1}^{N} R_{H,\Delta t}(t_i), \tag{3.3}$$

and sample variance

$$s_R = \frac{1}{N-1} \sum_{i=1}^{N} (R_{H,\Delta t}(t_i) - \bar{R})^2. \tag{3.4}$$

The volatility and drift of stock returns $R_H(t)$ can be estimated from equation (1.11) and (1.12),

$$\hat{\sigma} := \sqrt{\frac{s_R}{|\Delta t|^{2H}}} \tag{3.5}$$

and

$$\hat{\mu} := \frac{\bar{R}}{\Delta t} + \frac{\hat{\sigma}^2}{2} \tag{3.6}$$

respectively, where here $\Delta t = t_2 - t_1 = ... = t_N - t_{N-1}$.

Here are the estimates of volatility and drift in MATLAB:

```
% volatility
function [volatility]=fgn_volatility(x,dt,h)
    v=var(x);    % sample variance
    volatility=sqrt(v/(dt.^(2*h)));
end
% drift
function drift=fgn_drift(x,dt,sigma)
    drift=mean(x)/dt+(sigma.^2)/2;
end
```

3.4 Examples

Figure 3.2 shows three sequences of returns from the simulated stocks in Figure 2.5.

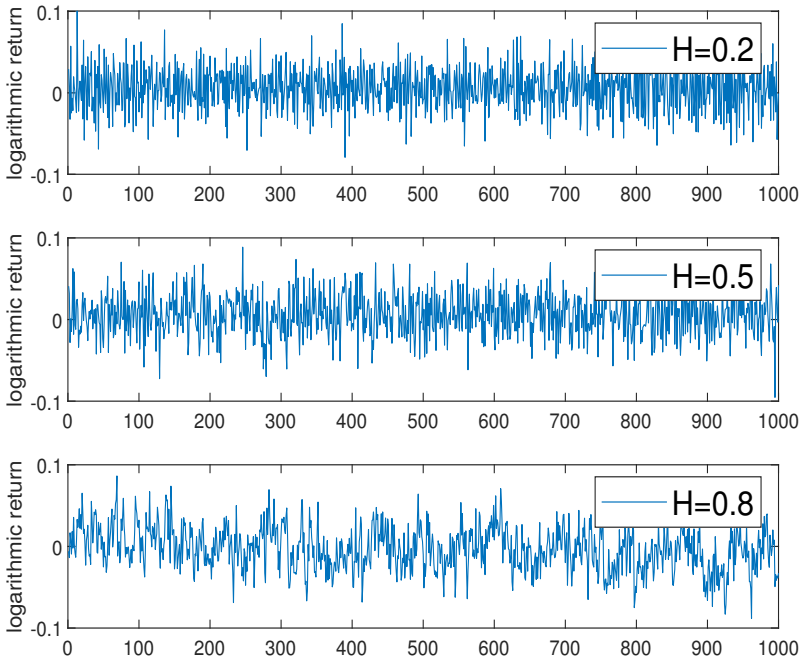

Figure 3.2: Returns of simulated stocks

Figure 3.2 is similar to Figure 2.1, because of the relation (1.13) between the FGN and the returns in the GFBM model. For each of the returns, μ and σ are estimated. The results are shown in the following table, here with known H, and given $\sigma = 0.027$ and $\mu = 0.0060$,

H	s_R	$\bar{R}$	$\hat{\sigma}$	$\hat{\mu}$
0.2	0.0272	0.0057	0.0272	0.0060
0.5	0.0261	0.0055	0.0261	0.0058
0.8	0.0274	0.0051	0.0274	0.0055

Table 3.2: Drift & volatility estimations

By Table 3.2, the estimated values $\hat{\sigma}$ and $\hat{\mu}$ approximate the given $\sigma = 0.027$ and $\mu = 0.0060$ respectively quite well, even though they are from models with different Hurst exponents.

Chapter 4

Case Analysis

Based on *stock.sohu.com* (a famous Internet company in China), we focus on three sequences of data. They are the closing prices of *Jiangling Motors Co., Ltd.* (JMC), *Shanxi Blue Flame Holding Company Limited* (SBFHC) published on Shanghai Stock Exchange, and *Shanghai Composite Index* (SZZS). The SZZS is an index which represents the tendency of all the stocks in the Shanghai Stock Exchange. We select these three sequences since they have the similar length of data from 2006 to 2018.

4.1 Application 1

We firstly estimate the Hurst exponents, volatilities and drifts of the three sequences of SZZS stock from 2006-7-1 to 2018-7-1 (12 years), 2010-7-1 to 2018-7-1 (8 years) and 2014-7-1 to 2018-7-1 (4 years) respectively. Then we simulate three corresponding paths with the estimated parameters and compare them with the original paths. Table 4.1 consists of the estimated Hurst exponents and Figure 4.1 to 4.3 show all the price paths, the x-axis represents time point t, the y-axis represents the stock prices. The sample sizes are 2971, 1942 and 976 respectively.

By rescaled range analysis, the results we obtained the indicate that for the returns of SZZS there exists a positive correlation during the recent 12 years, 8 years and 4 years respectively, since the estimated Hurst exponents of such three paths are larger than 0.5.

Different from rescaled range analysis, the periodogram method indicates that the logarithmic returns are positively correlated during 12 years and 4 years respectively, since then the estimated Hurst exponents are larger than 0.5, while the returns from 2010-7-1 to 2018-7-1 are indicated to be negatively correlated due to that the estimated Hurst exponents are smaller than 0.5.

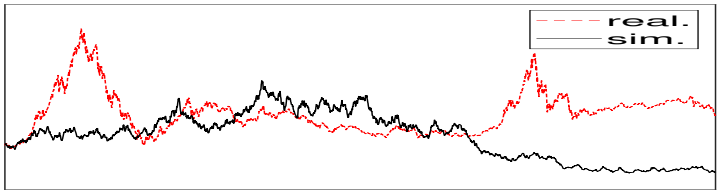

Figure 4.1: 12 years

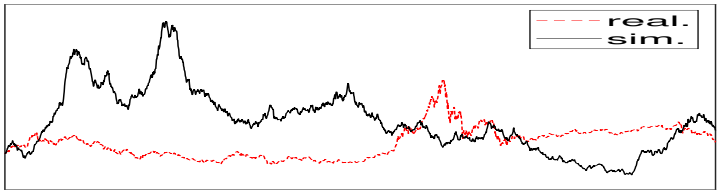

Figure 4.2: 8 years

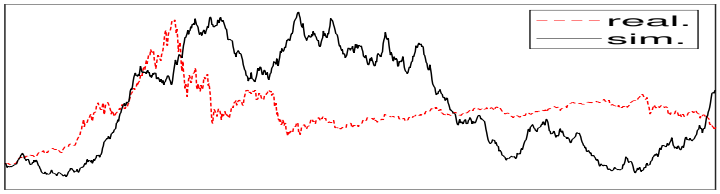

Figure 4.3: 4 years

	2006-7-1 to 2018-7-1	2010-7-1 to 2018-7-1	2014-7-1 to 2018-7-1
RS	0.6738	0.6013	0.6738
PE	0.6119	0.4755	0.6119

Table 4.1: Hurst exponents of SZZS

If the stock prices are GFBM, then the corresponding approximate FGN or standardized returns from (1.13),

$$\hat{X}_{H,\Delta t}(t) = \frac{(R_{H,\Delta t}(t) - \bar{R}_{H,\Delta t})|\Delta t|^H}{s_{R_{H,\Delta t}}} \tag{4.1}$$

with $\Delta t = 1$ are approximate standard Gaussian distributed. Figure 4.4 to 4.6 contain the theoretical autocovariances and estimated autocovariances by approxiamte FGN both from simulated and actual paths.

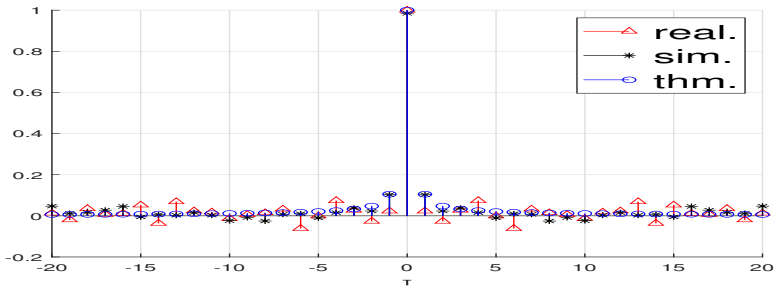

Figure 4.4: 12 years

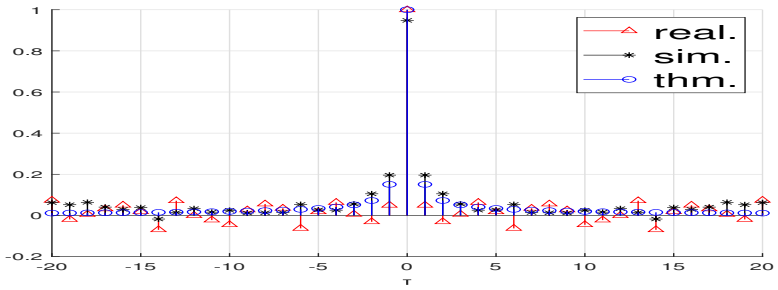

Figure 4.5: 8 years

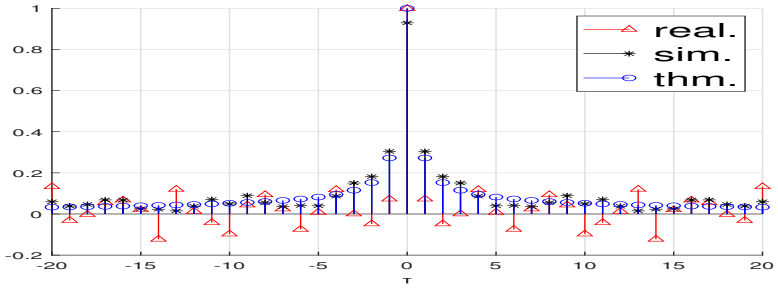

Figure 4.6: 4 years

The estimated autocovariances (black stars) from simulated paths and the theoretical autocovariances (blue circles) seem to follow the similar patterns that the FGNs are positively correlated. However, the autocovariances from real paths (red triangles) is a bit different from the others. Their autocovariances' values are located around 0, which means that some of the element pairs are positively correlated and some of them are negatively correlated. Such plots make the applicability of the GFBM model suspicious to our example data.

Figures 4.7 to 4.9 show the probability densities of standardized log returns for SZZS stocks. The densities for the real and simulated standardized returns are approximated by kernel density estimate. The dash-dotted lines represent the standard normal curves ($\mu = 0$, $\sigma = 1$).

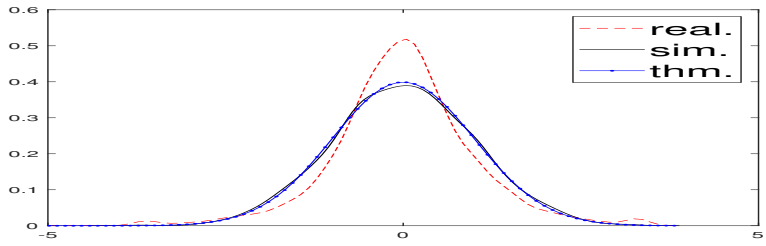

Figure 4.7: 12 years

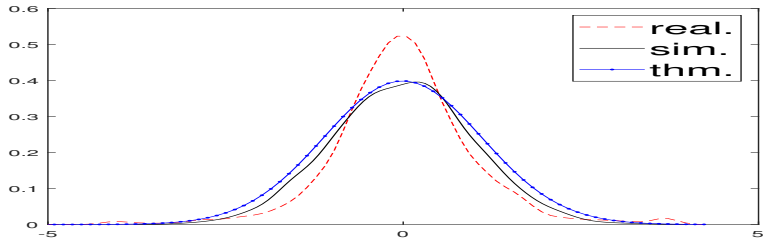

Figure 4.8: 8 years

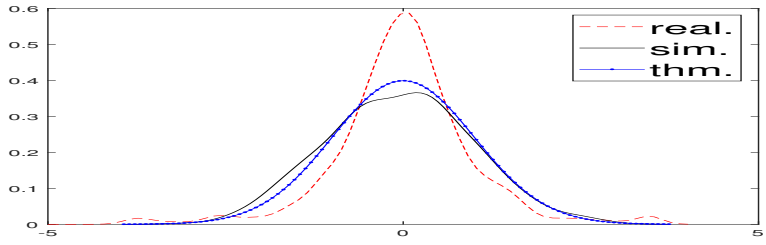

Figure 4.9: 4 years

We also use for simplicity the Jarque-Bera test to check whether the approximate FGNs of such actual and simulated stock prices are standard Gaussian distributed. In MATLAB (by the *jbtest*), the Gaussian hypothesis is clearly rejected at the 5% significance level. Note however that the Jarque-Bera test assumes that the outcomes in the sample are independent, which the standardized returns may not be. Therefore the conclusion of non-Gaussianity should be expressed by some small reservation.

Compared with the standard normal curves, all the FGNs from the simulated stocks seem to be Gaussian, but the corresponding FGN given by (4.1) for real stocks seem not be Gaussian. It is in line with the Gaussian hypothesis rejection.

4.2 Application 2

We simulate three paths for the stock prices of JMC based on the estimated parameters from the real prices. Figure 4.10 to 4.12 contain both simulated and real stock prices of JMC. The sample sizes are 2905, 1937 and 976 respectively.

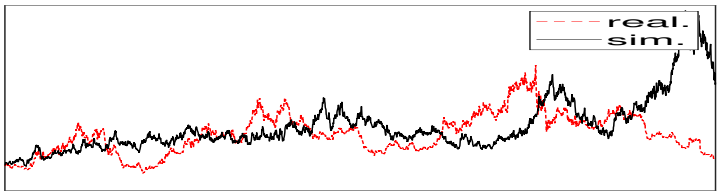

Figure 4.10: 12 years

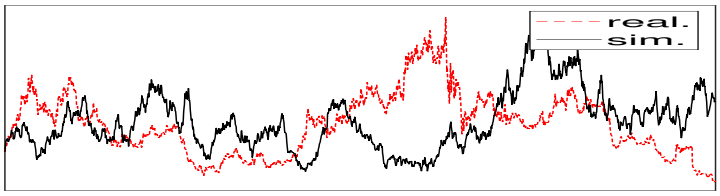

Figure 4.11: 8 years

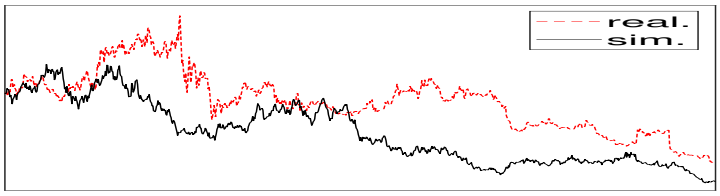

Figure 4.12: 4 years

	2006-7-1 to 2018-7-1	2010-7-1 to 2018-7-1	2014-7-1 to 2018-7-1
RS	0.4448	0.4928	0.4966
PE	0.5213	0.3696	0.4913

Table 4.2: Hurst exponents of JMC

According to Table 4.2, rescaled range analysis indicates that the returns of JMC are always negatively correlated, but the periodogram method indicates that such returns are positively correlated between 2006-7-1 and 2018-7-1.

Besides, the actual FGNs are still not Gaussian distributed by Jarque-Bera tests. Figure 4.13 to 4.15 are the probability densities of approximate real, simulated and theoretical FGNs.

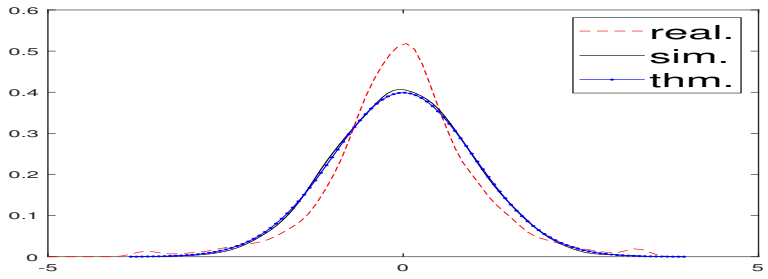

Figure 4.13: 12 years

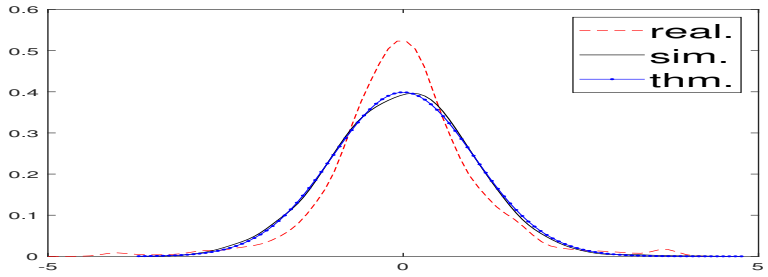

Figure 4.14: 8 years

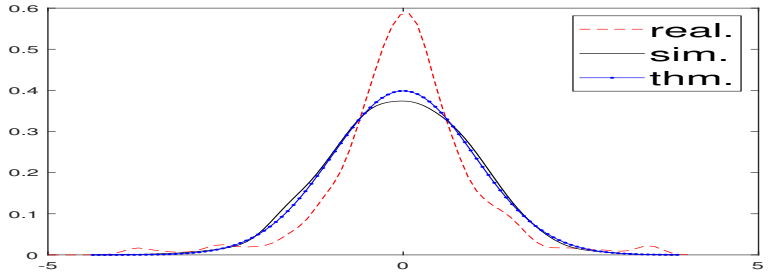

Figure 4.15: 4 years

4.3 Application 3

Figure 4.16 to 4.18 comprise both simulated and real stock prices of SBFHC while all the corresponding Hurst exponents are in Table 4.3. The sample sizes are 2673, 1724 and 771 respectively.

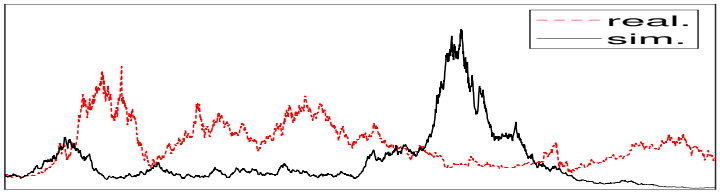

Figure 4.16: 12 years

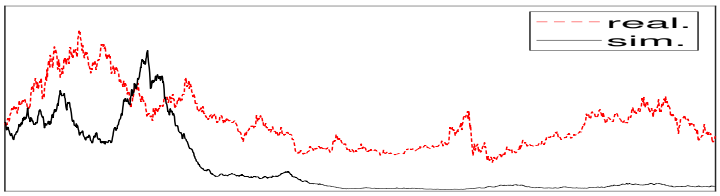

Figure 4.17: 8 years

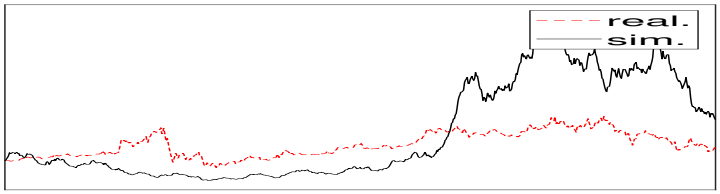

Figure 4.18: 4 years

	2006-7-1 to 2018-7-1	2010-7-1 to 2018-7-1	2014-7-1 to 2018-7-1
RS	0.5418	0.6379	0.6415
PE	0.5362	0.4646	0.4746

Table 4.3: Hurst exponents of SBFHC

The results of rescaled range analysis are similar to Application 1. For the three series there is a positive correlation of the returns of SBFHC, which is not true by the periodogram method. Once again, none of the actual FGNs of SBFHC are confirmed to be Gaussian by the Jarque-Bera test in line with Figure 4.19 to 4.15.

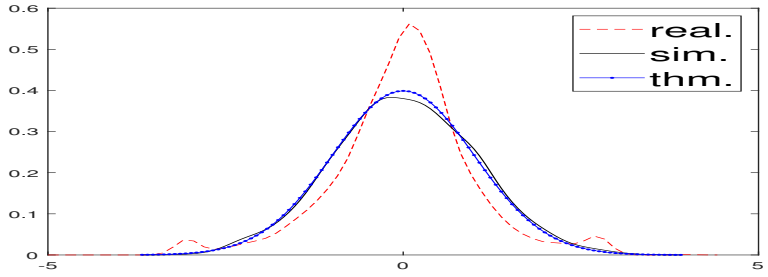

Figure 4.19: 12 years

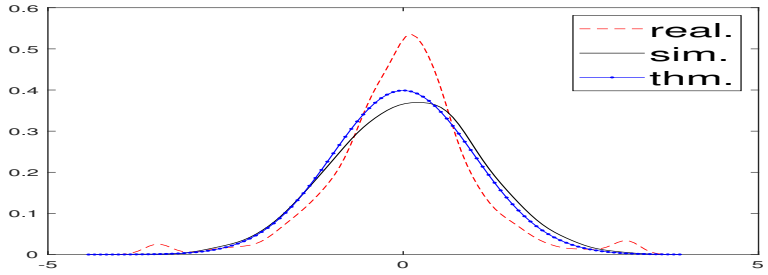

Figure 4.20: 8 years

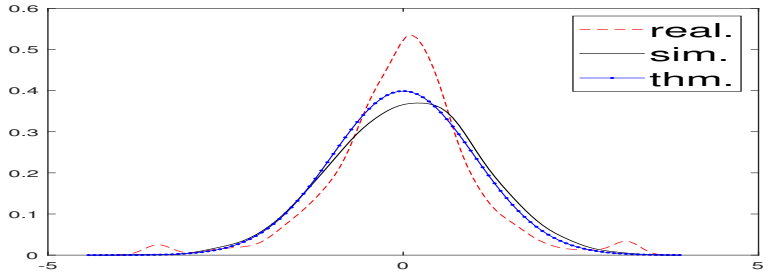

Figure 4.21: 4 years

4.4 Summary

Depending on the results we got, it is clear that the GFBM model is not a good choice for the Chinese data we select, because all the logarithmic returns are not Gaussian at the 5% significance level. This might also be one of the reasons why the results of rescaled range analysis and periodogram method are sometimes different.

Chapter 5

Conclusion

5.1 Result

To sum up, the GFBM model cannot describe stock prices well built on the data we select because their logarithmic returns are not fractional Gaussian distributed. Even if we could estimate Hurst exponents from the sample data, we cannot dare to tell that such returns are correlated.

However, in the Chinese financial market, the price limit of stocks cannot increase or decrease larger than 44% in the first day of company listing, and 10% on the next further days after the listing day. This rule could ensure the maintenance of the stock market. Moreover, it might control the correlation of stock prices. The GFBM model uses a simple but effective mathematical model to describe some potential features about stock prices. Understanding its specific characteristics of the GFBM model is a significant reference for financial mathematics.

5.2 Further work

To find and test a better mathematical model for stock-price modeling, since FBM is the only self-similar Gaussian process, it is reasonable and necessary to try some other self-similar stochastic process such as an α-stable Lévy process instead of the FBM. These processes are more complex but still convenient to handle. In that case, other structures are created by different processes depending on their special features.

Furthermore, finding better methods to estimate such parameters as Hurst exponent, drift and volatility, can also contribute a lot to the modification. Such modified methods can be applied not only to the GFBM model, but also in other potential models as well.

At last, data pre-processing is another important fact. For example, on the one hand, stock institutions do not record the stock prices at weekends and holidays. This will impact the continuity and correlation of data. On the other hand, inflation or deflation might be another fact influencing our model. Because of inflation or deflation, stock prices may not be self-similar. From the

point of macroscopic view, the stock price is growing, but its growth is unseen during a short time. Based on such situation, it could be preferable to apply the intrinsic values of stocks, not the empirical prices.

Based on the same features, the GFBM model could also contribute to such extra fields as calculation of surface growth (see [16]), computed tomography image edge detection (see [17]), optical flow determination (see [18]), etc. It is desirable to inherit the features of the GFBM model for the development of better models.

Chapter 6

Appendix

The MATLAB codes are divided into three parts. The first part is for the plots in chapter 2.

```
clear
clc
n=1000;
h=[0.2 0.5 0.8];                                % hurst exponents

%% FGN
fgn=zeros(n,3);
figure
for i=1:3
    fgn0=A.cholesky(n,h(i));                %fractional Gaussian noise
    fgn(:,i)=fgn0;
    subplot(3,1,i);
    plot(1:n, fgn0)
    ll=legend("H="+h(i));
    set(ll,'Fontsize',14);
        ylabel('fgn. value')
end

%% FBM
figure
fbm=zeros(n,3);
for i=1:3
    fbm0=cumsum(fgn(:,i));
    fbm(:,i)=fbm0;
    subplot(3,1,i);
    plot(1:n, fbm0)
    ll=legend("H="+h(i));
    set(ll,'Fontsize',14);
        ylabel('fbm. value')
end

%% Stock
mu=0.006;   sigma=0.027;   S0=10;
figure
S_sim=zeros(n,3);                   %stocks
S_sim(1,:)=S0;
for i=1:3
    f=fbm(:,i);
    for t=1:n
```

```
        S_sim(t+1,i)=S0*exp(mu*t-sigma*sigma*t/2+sigma*f(t));
    end
    subplot(3,1,i)
    plot(0:n, S_sim(:,i))
    ll=legend("H="+h(i));
    set(ll,'Fontsize',14);
        ylabel('stock price')
end

%% Return
rs=zeros(n,3);
figure
for i=1:3
    r0=diff(log(S_sim(:,i)));
    rs(:,i)=r0;
    subplot(3,1,i)
    plot(1:n, rs(:,i))
    ll=legend("H="+h(i));
    set(ll,'Fontsize',14);
    ylabel('logarithmic return')
end

%% parameters estimations
mu1=mean(rs);    sigma1=std(rs);
mu2=mu1+(sigma1.^2)/2;
h1=zeros(3,1);
h2=zeros(3,1);
for i=1:3
    h1(i)=A.RS(rs(:,i));
    h2(i)=A.PE(rs(:,i));
end

%% Covariance
cov_est=zeros(n*2-1,3);    %autocovariance by generalized formula
cov_thm=zeros(n*2-1,3);    %autocovariance by theorem
figure
for i=1:3
    [cov_est0]=A.ecov((fgn(:,i)));
    [cov_thm0,xval]=A.tcov(n,h(i));
    cov_thm(:,i)=cov_thm0';
    cov_est(:,i)=cov_est0';
    subplot(3,1,i);
    len=20;
    stem(xval(n-len:n+len),cov_est0(n-len:n+len));
    hold on
    stem(xval(n-len:n+len),cov_thm0(n-len:n+len));
    hold off
    ylim([-0.8 1.2])
    ylabel('autocov. value')
    xlabel('lag $\tau$');
    ll=legend('sim.','thm.');
    set(ll,'Fontsize',14);
end

%% Spectral density
figure
for i=1:3
    [sdf_thm,lags]=A.dftSDF(cov_thm(:,i)');
    %spectral density from cov_est
    [sdf_est]=A.dftSDF(cov_est(:,i)');
    % spectral density from cov_thm
    subplot(3,1,i);
```

```
    plot(lags,sdf_est,lags,sdf_thm)
    ll=legend('sim.','thm.');
    set(ll,'Fontsize',14);
    ylim([0,8])
    ylabel('sdf. value')
    xlabel('frequency f');
end
```

The second part is for application in chapter 4.

```
% raw GFBM
unnamed;
%% Stocks
S=flipud(unnamed);
nn=length(S);       % with 0
n=nn-1;             % without 0
R=diff(log(S));
X=(R-mean(R))/std(R);
h1=A.PE(X);                 % estimated hurst exponent
h=A.RS(X);
XX=A.cholesky(n,h);       % new FGN
BB=cumsum(XX);          % new FBM
dt=1;                                   % lag t
std0=A.fgn_volatility(R,dt,h);       % estimated volatility
mu0=A.fgn_drift(R,dt,std0);           % estimated drift
S0=S(1);
SS=zeros(nn,1); SS(1)=S0;
for t=1:n                                        % new stocks
    SS(t+1)=S0*exp(mu0*t-std0*std0*t/2+std0*BB(t));
end
figure
plot(0:n,S,"r--",0:n,SS,"k");
ll=legend("real.","sim.");
% title("Stock prices");
set(ll,'Fontsize',20);
set(gca,'xtick',[],'xticklabel',[])
set(gca,'ytick',[],'yticklabel',[])
grid on
xlim([0 n])
%% Probability density
figure
[f,xi1]=ksdensity(X);
[f1,xi2]=ksdensity(XX);
plot(xi1,f,"r--",xi2,f1,"k",xi2,normpdf(xi2),"b.-")
ll=legend('real.','sim.','thm.');
set(ll,'Fontsize',20);
xlim([-5,5]);
disp("Simulated stocks are not a Gaussion process: "+ jbtest(X));
%% Return
RR=diff(log(SS));
figure
plot(1:n,R,"r--",1:n,RR,"k")
grid on
ll=legend('real.','sim.');
% title("Log returns");
set(ll,'Fontsize',20);
xlim([1 n])
%% parameter estimation
std1=A.fgn_volatility(RR,dt,h);      % test volatility of S_sim
mu1=A.fgn_drift(RR,dt,std1);           % test drift of S_sim
```

```
%%  Autocovariance
[cov_real, lags]=A.ecov(X);
cov_sim=A.ecov(XX);
cov_thm=A.tcov(n,h);
figure
re=20;
hold on
stem(lags(n-re:n+re), cov_real(n-re:n+re),"r^-");
stem(lags(n-re:n+re), cov_sim(n-re:n+re),"k*-");
stem(lags(n-re:n+re), cov_thm(n-re:n+re),"bo-");
ll=legend('real.','sim.','thm.');
% title("Autocovariance");
grid on
xlabel('tau');
set(ll,'Fontsize',20);
hold off
%% Spectral density function
sdf_real = A.dftSDF(cov_real);
[sdf_sim, xval]= A.dftSDF(cov_sim);
[sdf_thm]=A.dftSDF(cov_thm);
figure
plot( xval, sdf_real,"r--", xval, sdf_sim,"k", xval, sdf_thm,"b.-")
ll=legend('real.','sim.','thm.');
% title("Spectral density");
set(ll,'Fontsize',20);
```

The final part is a tool class containing all functions.

```
classdef A
    methods(Static)
        %% Simulation methods of fractional Gaussian noise
    % input the length of FGN and Hurst exponent
        function [X]=cholesky(n,h)
        % n*1 vector with each random element¬N(0,1)
            V=normrnd(0,1,[n,1]);
            % covariance matrix of FGN
            sigma=zeros(n);
            dt=1;
            for t=1:n
                for s=1:n
                    ds=t-s;
                    c=(1/2)*(abs(ds+dt).^(2*h)...
                        -2*abs(ds).^(2*h)+abs(ds-dt).^(2*h));
                    sigma(t,s)=c;
                end
            end
            L=chol(sigma,'lower');
            X=L*V;
        end

    % input the length of FGN and Hurst exponent
        function [X]=DH_method(n,h)
            N=n/2;
            sigma=zeros(n);
            temp=zeros(1,n);
            temp(1)=(1/2)*(abs(0+1).^(2*h)...
                -2*abs(0).^(2*h)+abs(0-1).^(2*h));
            temp(N+1)=0;
            for i=2:N
                ds=i-1;
```

```
33                  c=(1/2)*(abs(ds+1).^(2*h)...
34                      -2*abs(ds).^(2*h)+abs(ds-1).^(2*h));
35                  temp(i)=c;
36                  temp(n+2-i)=c;
37              end
38              for i=1:n
39                  % create the circulant covariance matrix
40                  sigma(i,:)=circshift(temp,[0,i-1]);
41              end
42              V=normrnd(0,1,[n,1]);
43              [Q,A]=eig(sigma);
44              S=Q*(A.^(1/2))*Q';
45              X=S*V;
46          end
47
48          %% Autocovariance methods
49          function [c,lags]=ecov(X)      %real autocovariance
50              N=length(X);
51              lags=-N+1:N-1;
52              c=zeros(1,2*N-1);
53              mu=mean(X);
54              for m=0: N-1                  %lag start from 0
55                   xl=(X(m+1:N)-mu)';
56                  xr=X(1:N-m)-mu;
57                  cum=(xl*xr)/(N*var(X));
58                  c(N+m)=cum;
59                  c(N-m)=cum;
60              end
61          end
62
63          function [c,lags]=tcov(N,h)         %theoritical ...
                autocovariance
64              lags=-N+1:N-1;
65              c=zeros(1,2*N-1);
66              for i=1:2*N-1
67                  tau=i-N;
68                  c(i)=(1/(2))*( abs(tau+1).^(2*h)...
69                      -2*abs(tau).^(2*h)+abs(tau-1).^(2*h) );
70              end
71          end
72
73          %% Spectral density methods
74          function [sdf, lags] = dftSDF(c)       % computation SDF ...
                by autocovariances
75              R=length(c);
76              N=(R+1)/2;
77              len=201;
78              lags=linspace(-1/2, 1/2, len);
79              tau=-N+1:N-1;
80              sdf=zeros(1, len);
81              for f=1: len
82                  x=exp(-1i*2*pi*lags(f)*tau);
83                  cum=c*x';
84                  sdf(f)=abs(cum);
85              end
86          end
87
88          %% Estimation of parameters
89          % volatility
90          function [volatility]=fgn_volatility(x,dt,h)
91              v=var(x);                          % sample variance
92              volatility=sqrt(v/(dt.^(2*h)));
```

```
93          end
94
95          % drift
96          function drift=fgn_drift(x,dt,sigma)
97              drift=mean(x)/dt+(sigma.^2)/2;
98          end
99
100         % Rescaled Range Analysis
101         function [hurst]=RS(X)
102             N=length(X);
103             n=20;          % skip the initial 20 points
104             xvals=n: N;
105             logx=log(xvals);
106             yvals=zeros(1, length(xvals));
107             for t=n: N
108                 tmpX=X(1:t);
109                 Y=tmpX-mean(tmpX);
110                 Z=cumsum(Y);
111                 R=max(Z)-min(Z);
112                 S=std(tmpX);
113                 yvals(t-(n-1))=R/S;
114             end
115             logy=log(yvals);
116             p=polyfit(logx,logy,1);
117             % Hurst exponent is the slope of linear-fit plot
118             hurst=p(1);
119 %             scatter(logx,logy,'.')
120 %             hold on
121 %             plot(logx, logx*hurst+p(2))
122 %             xlabel('log(n)')
123 %             ylabel('log(R/S)')
124 %             legend('data','regression')
125 %               hold off
126         end
127
128         % Periodogram Method
129         function [hurst]=PE(X)
130             N=length(X);
131             len=601;        % separate (-0.5,0.5) into 601 parts
132             lags=linspace(-1/2,1/2, len);
133             sdf=zeros(1,len);
134             for f=1: len
135                 tau=0:N-1;
136                 xl=X-mean(X);
137                 xr=exp(-1i*2*pi*lags(f)*tau);
138                 cum=xr*xl;
139                 sdf(f)=(abs(cum)^2)/N;
140             end
141             mid=(len+1)/2;
142             len=fix(mid^(4/5));
143             x=lags(mid+1: mid+1+len);
144             y=sdf(mid+1: mid+1+len);
145             logx=log(x);
146             logy=log(y);
147             gamma=polyfit(logx,logy,1);
148             hurst=(1-gamma(1))/2;
149         end
150     end
151 end
```

Bibliography

[1] Hänggi, P., & Marchesoni, F. (2005). *Introduction: 100 years of Brownian motion.*

[2] Capinski, M., & Zastawniak, T. (2006). *Mathematics for finance: an introduction to financial engineering.* Springer.

[3] Miller, S., & Childers, D. (2012). *Probability and random processes: With applications to signal processing and communications.* Academic Press.

[4] Reddy, K., & Clinton, V. (2016). Simulating Stock Prices Using Geometric Brownian Motion: Evidence from Australian Companies. *Australasian Accounting, Business and Finance Journal, 10*(3), 23-47.

[5] Mandelbrot, B. B., & Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM review, 10*(4), 422-437.

[6] CHEN, J., TAN, X. H., & JIA, Z. (2006). Performance analysis of seven estimate algorithms about the Hurst coefficient [J]. *Journal of Computer Applications, 4*, 059.

[7] Liu, Y., Liu, Y., Wang, K., Jiang, T., & Yang, L. (2009). Modified periodogram method for estimating the Hurst exponent of fractional Gaussian noise. emphPhysical Review E, 80(6), 066207.

[8] Øksendal, B. (2003). Fractional Brownian motion in finance. *Preprint series. Pure mathematics http://urn. nb. no/URN: NBN: no-8076.*

[9] Dieker, T. (2004). Simulation of fractional Brownian motion. *MSc theses, University of Twente, Amsterdam, The Netherlands.*

[10] Härdle, W., & Simar, L. (2007). *Applied multivariate statistical analysis* (Vol. 22007, pp. 1051-8215). Berlin: Springer.

[11] Koutsoyiannis, D. (2002). The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrological Sciences Journal, 47*(4), 573-595.

[12] Rostek, S., & Schöbel, R. (2013). A note on the use of fractional Brownian motion for financial modeling. *Economic Modelling, 30*, 30-35.

[13] Niu, F. G., & Liu, W. Q. (2010). Relations of Fractional Brownian Motion and Hurst Exponent. *Journal of Shanxi University, 33*, 380-383.

[14] Embrechts, P., & Maejima, M. (2000). An introduction to the theory of self-similar stochastic processes. *International Journal of Modern Physics B, 14*(12n13), 1399-1420.

[15] Zhou, S., Shiqian, X., & Chengyi, P. (2008). *Probability theory and mathematical statistics*. Beijing: Higher Education Press.

[16] Barabási, A., & Stanley, H. (1995). *Fractal Concepts in Surface Growth*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511599798

[17] Dawei Qi, Lei Yu, & Xiaocheng Feng. (2008). A detection method on wood defects of CT image using multifractal spectrum based on fractal brownian motion. *Automation and Logistics, 2008. ICAL 2008. IEEE International Conference on*, 1539-1544.

[18] Ma, L. T., Wang, L., Chen, X., & Li, B. (2011). Determining optical flow of natural scene based on the fractal Brownian motion model. Procedia Engineering, 15(C), 5514-5518.

[19] Xiao, Y. Q., & Zou, J. Z. (2008). 分数布朗运动环境下的期权定价与测度变换 (The Option pricing and measure transformation in geometric Brownian motion). *Journal of Mathematics in Practice and Theory, 38*(20), 58-62.

[20] Hu, R. D.(1996). 布朗运动在股票期权定价模型中的应用 *(The application of Brownian motion in the option-pricing model)* (Doctoral dissertation).

[21] Zhang, J. H., Dai, L., & Liu, W. H. (2014). 股票价格行为关于几何布朗运动的模拟–基于中国上证综指的实证研究 (The simulation of stock prices by geometric Brownian motion: based on Shanghai Composite Index). *Money China*, (29), 31-33.

[22] Gao, L. (2010). 股票指数几何布朗运动模拟及实证分析 (The simulation and case analysis of stock prices by geometric Brownian motion). *Modern Economic Information*, (6X), 18-18.

Lnu.se

Linnæus University
Sweden

Faculty of Technology
SE-391 82 Kalmar | SE-351 95 Växjö
Phone +46 (0)772-28 80 00
teknik@lnu.se
Lnu.se/faculty-of-technology?l=en