Theses of the PhD. Dissertation

Zsolt Nika

Supervisors: Dr. Miklós Rásonyi and Dr. Péter Szolgay

Optimal Investments for Price Processes with Long Memory

Roska Tamás Doctoral School of Sciences and Technology
Pazmány Péter Catholic University

Budapest, Hungary
2020

Contents

Introduction

In financial investments, the investor decides which securities (e.g. stocks or bonds) to buy. Naturally, the investor's goal is to maximize their wealth at a given point in the future. However, I can have no prior knowlege of the exact result or outcome of a given investment; this uncertainty is termed risk, which different investors take into account in differing ways. The investor's preference, that is, to what extent risk influences their decision, is called *risk aversion*.

Since the investor has to make decisions based on available information, such as the price of a bond, the control theory of stochastical systems are used to handle the theory of investments. These decisions determine which part of the available, disposable wealth is invested into financial assets. The price takes on the role of independent variable, whereas the objective function depends on the risk senstivity that the investor wishes to optimize.

Methodology and basic concepts of the research

This research contains numerous simulation results. In these instances, I used the Monte-Carlo method for the price process of the stocks corresponding to the given dynamic system. The research necessitated the use of numerical integrals and the programming of algorithms that I had defined. The source code was written in R, while the studies were carried out on a personal computer (with circa 8 GB RAM and a 2.6 Ghz dual-core processor)

The theses will discuss two types of investments chosen based on the investor's goal during investment. This choice is essential because the degree of risk differentiates investments with identical expected values. For instance, in a game of chance, if a player has a 50% chance of winning 100 dollars and a 50% chance of winning 0 dollars, the expected value of the win is 50 dollars. However, investors consider investments more favorable when the winnings are for certain 50 dollars. This is because there is no risk - therefore, investors "punish" the presence of risk.

The risk aversion, that is, the value of the investor's risk (the amount of expected profit they are willing to forgo in order to avoid a given risk) varies from investor to investor. Investors with greater capital are usually greater risk-takers, while those with less capital are more risk-averse (for example, to avoid bankruptcy). These two criteria (risk sensitivity, or in the case of greater wealth, greater risk-taking) are described by several types of objective functions. Of these, the most well-known is the *log-optimal* portfolio, in which the the investor maximizes the expected value of the weatlh's logarithm. If the investor does not take risk into account, the investor is *risk-neutral*. In this case, I do not differentiate between the two cases in the above example.

In addition to the objective function, the decision function must also be given. In both cases I assume that the investor is *self-financing*; that is, in every moment they can only use their present wealth. For example, they cannot provide or receive loans.

Financial assets may be grouped according to whether their future price is a deterministic or a stochastic variable. The former case is termed a *risk-free* asset (e.g. bonds) and the latter case is termed a *risky* asset. For simplicity, I assume that during trading, one of each type of asset is at the investor's disposal. This restriction is not a strict one and instead serves to simplify the notation, as almost every statement would be true even in the case of several stocks following simple modifications.

Another common assumption is that the objective function does not depend on the value of the portfolio at a given time but rather on its value at infinity. These are called *long term* investments. Its purpose is to investigate long term behavior of the market rather than its fluctuation in time. That is to stay, its average, characteristic behavior in time is investigated.

In the following subsections, I will present those basic terms that will be fundamental to understading the thesees.

Log-optimal portfolios

Let us presume a discrete trading system with a risk-free asset whose price at time t is denoted by B_t and a risky asset whose price is S_t without a transaction fee in a liquid market. The wealth of the trader at time t is as follows:

$$W_t = W_{t-1}(1-\pi_t)B_{t+1}/B_t + W_{t-1}\pi_t S_{t+1}/S_t,$$

where the investor's decision is the $\pi_t \in [0,1]$ trading strategy, i.e. what percent of their wealth should be invested in stocks. (As an example, the $\pi_t \notin [0,1]$ would correspond to the loan that I excluded.) A general assumption is that the bond price evolves in time deterministically with a fixed (constant) interest rate $r \geq 0$, namely, $B_t = B_{t-1}(1+r)$, $\forall t$. Moreover, I assume that $r = 0$, in order to exclude the effect of interest rate. This does not result in an actual limitation to the model and allows for clearer notation. For statistical reasons, it is convenient to describe stock dynamics by its *log-return* $H_t := \log(S_t/S_{t-1})$. In this way, the above equation may be rewritten as

$$W_t/W_{t-1} = (1-\pi_t) + \pi_t \exp(H_t)$$

following division by W_{t-1}.

At time $t = 0$, the investor possesses an initial capital of w_0. The investor's objective is maximizing the logarithmic utility function by the long-term trading of their entire wealth, namely,

$$\liminf_{t\to\infty} \frac{1}{t}\mathbb{E}[\log(W_t)] \to \max.$$

(If the limit clearly exists and is finite, I denote it as G^*, and the investor's wealth increases exponentially at an average rate of at least G^*, that is: $\mathbb{E}[W_t] \gtrapprox W_0 \exp(G^* t)$.)

The investor's wealth at time t may also be expressed in another way: $W_t = w_0 \prod_{j=1}^{t} W_j/W_{j-1}$ for an initial capital $w_0 > 0$. Using one of the basic properties of the conditional expected value, the tower rule, I find that the optimal strategy may be found by maximizing the conditional expected value of the portfolio value log increments ($\log(W_t/W_{t-1})$) at every time t:

$$\max_{\pi_t \in [0,1]} \mathbb{E}[\log((1-\pi_t) + \pi_t \exp(H_t))|\mathcal{F}_{t-1}]. \tag{1}$$

Provided that there is a strategy π_t^* which optimizes the expression at every t, it must be shown that the objective function thus obtained is indeed convergent. However, it must be noted that this process– where at each point in time only the increment of the next point in time must be optimized– can solely be utilised in the case of log optimal portfolios; with other cases of utility functions, it cannot be used.

Results thus far may be classified into three groups. On the assumption that the H_t process is Markovian, the problem may be solved using dynamic programming. When only assuming ergodicity in a more general case, the convergence of the objective function is proven, albeit the procedure to find the optimal strategy is not given. Furthermore, it is difficult in many cases to prove ergodicity in financially relevant models. The third method is the set of learning algorithms for large data sets. However, here the large amount of data provide a strong restriction. Additionally, the results cannot be interpreted as these methods are generally nonparametric. The model and the investment cannot be known or interpreted.

In my research, I found a general method of constructing the optimal strategy for more complex models. I numerically examine and prove the convergence of the objective function. Further, I also introduce an effective approximative strategy and define a threshold strategy whose optimum can be found by a learning algorithm. This learning algorithm does not assume the availabilitiy of a large amount of data during the investment, nor does it require knowledge of the type of process (within reasonable limits).

Considering transaction fees in the case of risk-neutral preference

When trading continuously in time, and allowing for a transaction fee proportional to the rate of trading, let X_t^0 be the wealth stored in risk free assets, while X_t^1 is the number of stocks (risky assets); let the price process of the stocks be proportional to the fractional Brownian motion: $S_t = \sigma B_t^H$ (that is, a negative price is possible price process has long memory). Trading takes place in the interval $[0, T]$ and its goal is to maximize the expected value of cash at time T, provided that a maximum exists. It is the investor's decision how "fast" they trade at a given point in time t, therefore I denote this as ϕ_t. In this way, the number of stocks is $X_t^1(\phi) = X_0^1 + \int_0^t \phi_u du$.

The value of cash at time t is the following:

$$X_t^0(\phi) = X_0^0 - \int_0^t \phi_u S_u du - \int_0^t \lambda |\phi_u|^\alpha du.$$

The initial conditions of X_0^0 and X_0^1 can be chosen freely, while the parameters λ and α are characteristics of the market.

Because it is established that an optimal strategy exists, ϕ_t^* is a speed of trading that reaches the optimal objective function $\sup_\phi \mathbb{E}[X_T(\phi)] =: u^*(T)$. However, finding this strategy for a finite T is a lost cause, although $T \to \infty$ may arbitrarily approach the optimal strategy.

It is an established result that for $T \to \infty$ the growth of the optimal objective function $u^*(T)$ is

$$\limsup_{T\to\infty} \frac{u^*(T)}{T^{H(1+1/(\alpha-1))+1}} < \infty.$$

That is, at the time of expiry T, the expected value of the investment grows according to a power law $u^*(T) \approx T^{H(1+1/(\alpha-1))+1}$.

Let κ be the trading "intensity," which denotes that the speed of trading is proportional to κ-th power of the absolute value of the price process.

The upper bound of the intensity is that $\kappa < 1/(\alpha - 1)$. The strategy

$$\phi_t(T, \kappa) = \begin{cases} \text{sgn}(S_t(H - 1/2))|S_t|^\kappa, & t \in [0, T/2), \\ -\frac{1}{T/2} \int_0^{T/2} \phi_s \, ds, & t \in [T/2, T], \end{cases}$$

allows the utility function grows such that $u(T) \approx T^{H(1+\kappa)+1}$. For $\kappa = 1/(\alpha-1)$ the growth would be optimal but this is not provided for by theorem.

The question regarding this investment is what can be known about the strategy that approaches optimality. Because the investment does not take into account risk, it is uncertain if the investment's Sharpe ratio is meaningful. The Sharpe ratio is the ratio of the expected profit and the risk; it is vital in the determination of how realistic a theoretical investment model is. Furthermore, the way that the investment depends on the financial parameters is, in theory, only numerically verifiable (the market parameters: H Hurst parameter, α price impact, λ volatility; the parameters of the investment: κ intensity). Aside from this, it is also uncertain whether the optimal growth rate $\kappa = 1/(\alpha-1)$ is attainable, despite a lack of theoretical basis.

Research Goals

Studies on investment theory are lacking for parametric models that take long memory into consideration for stock prices. In my PhD dissertation, I examine constructive trading strategies where it is both proven the solution exists and where the optimal strategy can be found numerically. Due to my examination of parametric models, I was able to examine how the dynamic parameters that desribe the price processes affect the optimal decision based on numerical simulations. In the course of this disseration, my goals include the following:

- Creation of a constructive trading strategy for investors of differing risk sensitivites in which the stock prices have long memory

- Defining a general class of models and specific price process dynamics which are parametric and have long memory. Furthermore, long memory is taken into account with only one parameter

- Developing an approximative strategy whose computational cost is significantly more favorable than the log-optimal strategy

- Examining log-optimal and approximative strategies with numerical simulations

- Creation of learning algorithms that adapt to the dual expectation of financial data arriving both continually and quickly. Due to this, the method must also be applicable on a small data set and have a low computational cost.

Theses

Log-optimal trading in Conditionally Gaussian processes

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{Z}}, \mathbb{P})$ be a probability space, where the σ-algebra $\mathcal{F}_t$ is generated by $\{\varepsilon_j\}_{j\leq t}$ and $\{\eta_j\}_{j\leq t}$ where $\{\varepsilon_j\}_{j\in\mathbb{Z}}$ and $\{\eta_j\}_{j\in\mathbb{N}}$ are two i.i.d. sequences that are independent. I assume that ε_j, $j \in \mathbb{Z}$ are standard Gaussian.

The log-return process is called **Conditionally Gaussian** is it evolves in time

$$H_t = F(Z_{t-1}, Y_t, \varepsilon_t, \eta_t), \tag{2}$$

where F is a measurable function, Y_t is a stationary Gaussian process such that Y_t is $\sigma(\varepsilon_j,\ j \leq t)$-measurable.

Here Z_{t-1} contains the information about the past values of the log-return, so it is assumed that $Z_{t-1} = j(H_{t-1}, H_{t-2}, \dots)$ for some measurable $j : \mathbb{R}^{\mathbb{N}} \to \mathbb{B}$ for some Banach space $\mathbb{B}$.

Corresponding publication: [1].

Thesis One: Provided that the dynamics of the price process $F(\cdot)$ meets a given ergodic-like criteria, then a strategy (π_t^*) can be constructed that is a function of the information available at time $(t-1)$. The investment based on this strategy converges with probability one to the log-optimal investment in the long run. Therefore, the following is also true:

$$\liminf_{T\to\infty} \frac{1}{T}\mathbb{E}[\log(W_T^{\pi^*})] = \max_{\pi} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}[\log(W_T)].$$

Using the function $\hat{\pi}(z, \nu, \kappa)$, the optimal strategy may be constructed:

$$\pi_t^* = \hat{\pi}(Z_{t-1}, \nu_{t-1}, \kappa_{t-1}),$$

where $\nu_{t-1} := \mathbb{E}[Y_t|\mathcal{F}_{t-1}]$ and $\kappa_t := Var[Y_t|\mathcal{F}_{t-1}]$. The function $\hat{\pi}$ can be determined in advance for any three (z, ν, κ) by applying two-dimensional numerical integration.

Comment The importance of the above statement lies in the fact that Conditionally Gaussian-type processes are compatible with the majority of price process dynamics used in practice, assuming that the noise is of Gaussian distribution. This provides a general method for the construction of the optimal strategy.

Price Process Models with Long Memory

The *Discrete Gaussian Stochastic Volatility* (DGSV), a novel proposal for the dynamics of the log-return (H_t, $t \in \mathbb{N}$), is defined in a similar way to the popular model *rough Volatility*:

$$H_t = \mu + \alpha H_{t-1} + \sigma e^{Y_t} \left(\rho \varepsilon_t + \sqrt{1-\rho^2} \eta_t \right);$$

$$Y_t = \sum_{j=0}^{\infty} \beta_j \varepsilon_{t-j}, \qquad \mu, \sigma, \beta_j \in \mathbb{R}, \quad \alpha, \rho \in (-1, 1).$$

The *Bilinear ARCH* model, the long-memory version of ARCH (R. F. Engle and C. Granger's Nobel prize-winning model), is defined as follows:

$$H_t = c_0 + c_1 H_{t-1} + \eta_t \sigma_t;$$

$$\sigma_t = a + \sum_{j=1}^{\infty} \beta_j H_{t-j}, \qquad a, c_0, c_1, \beta_j \in \mathbb{R}.$$

The variables ε_t and η_t are independent white noise sequences with Gaussian distribution.

Corresponding publication: [1].

Thesis Two: The following statements may be made about Discrete Gaussian Stochastic Volatility and Bilinear Arch models. The models are compared based on the examination of 1,250 price processes of stocks on the New York Stock Exchange between January 1, 2010, and December 31, 2016.

1. The parameters of both models may be chosen so that the first four moments of the generated data reflect reality. In addition, both models are capable of describing the so-called "leverage effect", particularly the DGSV model, in which this is controlled by the ρ parameter.

2. Several research results, in addition to my own results, confirm that long memory appears in volatility rather than in the drift of the log-return. Specifically, while log-returns are not correlated in time, their absolute values are (therefore $\tau \to Cov(|H_t|, |H_{t+\tau}|)$ decays by a power law). Both models are capable of describing this, provided that $\beta_j = b_0(1+j)^{-b}$, $b > 0.5, b_0 > 0$, $j \in \mathbb{N}$.

Log-optimal Investments with the DGSV and BARCH models

Corresponding publication: [1].

Thesis Three: Both the DGSV and the Bilinear ARCH meet ergodic-like criteria which guarantee the log-optimal investment for each realization (with probability one). Results based on numerical simulations:

1. The steady state is achieved in 1,000-15,000 time steps (which represents 4-6 years of daily trading), while Thesis One only proves the limit case $t \to \infty$. The speed of the convergence decreases as the strength of the memory increases.

2. The optimal solution is a decreasing linear function of the strength of the memory (b_0) when the variance of the log-volatility is kept constant.

3. The rate of the memory's decay (b) does not effect the optimal solution.

4. The optimal strategy is "extreme" in the sense that for each realization in the majority of the time, it takes only two values, either 0 or 1 (the value it takes changes in time). This suggests that the range of strategies may be reduced from $[0, 1]$ to $\{0, 1\}$ in practice.

Approximative and Threshold Strategies

The threshold strategy is defined as follows:

$$\mathbb{E}[\log((1-\pi)(1+r)+\pi\exp(H_t))|\mathcal{F}_{t-1}] \to \max$$
$$\pi := \mathbb{1}_{\{f(H_{t-1})>0\}} = \mathbb{1}_{\{H_{t-1}>\theta\}}.$$

If the σ-algebra $\mathcal{F}_{t-1}$ does not provide more information on the investment than the function $f()$, then the two strategies result in the same outcome.

Corresponding publication has not peered yet, results are available at arXiv 1907.02457.

Thesis Four The optimal threshold strategy may be searched for using the Kiefer-Wolfowitz algorithm, where Isearch for the $g(\theta) := \mathbb{E}[\log((1-\pi_t)+\pi_t\exp(H_t))]$ function's maximum when using the $\pi_t := \mathbb{1}_{\{H_{t-1}>\theta\}}$ strategy. The alorithm is below:

$$\theta_{t+1} = \theta_t + a_t\frac{H_t\mathbb{1}_{\{H_{t-1}\in[\theta_t\pm c_t]\}}}{c_t}, \quad a_t = t^{-1}, c_t = t^{-1/3}.$$

With this method, the investment may be optimized without conducting statistical examinations beforehand on the price process of stocks. Following the completion of numerical analyses, I found several possible pitfalls for which I propose solutions.

1. The algorithm's computational cost is cheap; it contains only multiplication, addition, and conditional expressions. The incoming data is processed one at a time without the use of "big data."
2. A vizsgált esetekben a $g(\theta)$ függvénynek pontosan egy maximuma van. , the g fuction has exactly one maximum
3. Although in theory it cannot be confirmed, numerical simulations show that the θ_t series converges in mean square to the optimal value.

Comment One Numerical results show that the linear approximation strategy is the same as the log-optimal strategy in the majority of cases, although its construction does not necessitate numerical integration, therefore the strategy can easily be constructed for many stocks.

Comment 2 The threshold strategy is similar to the approximative strategy but does not assume knowledge of $\mathcal{F}_{t-1}$ which is not available to the investor when using real data. The statement can be generalised information other than only H_{t-1}, such as external market information which effects the stock price. The algorithm in Thesis Four searches for the optimal investment during the continuous infux of the data stream (in general cases, it is impossible to prove that there exists only one maximum point).

Trading with Fractional Brownian motion

The definition of the Sharpe ratio is as follows: $SR := \mathbb{E}[X_T]/\mathbb{D}(X_T)$.

Corresponding publication: [2].

Thesis Five:

1. The Sharpe ratio is bounded with the speed of trading $\phi_t(T, \kappa)$.
2. Based on numerical results, the Sharpe ratio grows with the strengthening of the memory (that is, $|H - 0.5|$ is large). The Sharpe ratio of the trading is a constant function of intensity κ when memory is weak and a decreasing function when the memory is strong.
3. From the previous two points, it follows that the maximization of the objective function and the maximization of the Sharpe ratio are mutually exclusive expectations: during high-intensity trading ($\kappa \approx 1/(\alpha - 1)$), the expected value is nearly optimal but its Sharpe ratio is minimal; at weak-intensity trading, the opposite holds true.
4. The suggested trading strategy is applicable even in the case $\kappa = 1/(\alpha - 1)$ and numerical experiments show that it attains the optimal growth rate (though this has not been theoretically verified).

Publications

[1] Nika, Zsolt, and Rásonyi, Miklós (2018). "Log-Optimal Portfolios with Memory Effect." *Applied Mathematical Finance (Taylor & Francis)*, 25(5-6), 557–585.

[2] Guasoni, Paolo and Nika, Zsolt and Rásonyi, Miklós (2019). "Trading Fractional Brownian Motion." *SIAM Journal on Financial Mathematics*, 10(3), 769–789.