British Journal of Mathematics & Computer Science
16(6): 1-6, 2016, Article no.BJMCS.25967
ISSN: 2231-0851

SCIENCEDOMAIN *international*
www.sciencedomain.org

Statistical Arbitrage Opportunities Using Bellman Equation

Reza Habibi[1*] and Hamed Habibi[2]

[1]Iran Banking Institute, Central Bank of Iran, Tehran, Iran.
[2]Faculty of Science and Engineering, Curtin University, Perth, Australia.

Authors' contributions

This work was carried out in collaboration between both authors. Authors designed the study, wrote the protocol and supervised the work. Authors performed the statistical analysis. Authors wrote the first draft of the manuscript. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMCS/2016/25967
Editor(s):
(1) H. M. Srivastava, Department of Mathematics and Statistics, University of Victoria, Canada.
Reviewers:
(1) Stephen Akandwanaho, University of KwaZulu-Natal, South Africa.
(2) Alexandre Gonçalves Pinheiro, Ceará State University, Brazil.
Complete Peer review History: http://sciencedomain.org/review-history/14717

Original Research Article

Received: 28th March 2016
Accepted: 4th May 2016
Published: 21st May 2016

Abstract

This paper uses the dynamic programming to detect the optimal statistical arbitrage opportunities in a market including a bond and a stock. First, it is assumed that the growth rates of stock are independent random variables and Bellman equation is derived for probability of gain of a portfolio containing a long position in stock and short position in bond. The Bellman equation is derived and its approximations are studied. Then, using the simulation, the performance of method in correlated growth rates cases is proposed. Conclusions are also given.

Keywords: Bellman equation; bond; dynamic programming; probability of statistical arbitrage; stock.

1 Introduction

This paper deals with construction of statistical arbitrage strategy to obtain possible profit with a high probability in future with zero investment at the current time. [1] investigated the relationship of market efficiency and statistical arbitrage. A comprehensive review in statistical arbitrage may be found in [2]. Authors [3] applied the statistical arbitrage detection techniques in U.S. equities markets. In foreign

**Corresponding author: E-mail: habibi1356@gmail.com*

exchange markets, there are many basic shortest path algorithms to detect statistical arbitrage, see [4]. There are many basic strategies to this end, like market neutral strategies, pairs trading, high frequency trading, co-integration trading and mean revert trading, see [5] and references therein. The dynamic programming is a method for optimal control like (i) maximum principle and (ii) calculus of variation which may be used to exploit the statistical arbitrages, see [6,7].

Here, it is assumed that B_t is a bond with riskless rate r_f growth, that is $B_t = B_0(1+r_f)^t$ and S_t is a stock with growth rate $r_t, t \geq 1$, i.e.,

$$S_t = S_0 \prod_{i=1}^{t} (1 + r_i).$$

In this paper, is it assumed that r_i's are independent random variables and for each $t \geq 1$, $E(r_t) > r_f$. The value of a portfolio with long position in one unit of stock and α_t unit of bond in short position is $\pi_t = S_t - \alpha_t B_t$. Suppose that $\pi_0 = 0$, then $\alpha_0 = \frac{S_0}{B_0}$. For $t \geq 1$, let $\gamma_t = \ln\left(\frac{\alpha_t}{\alpha_0}\right)$ and p_t be the gain probability that is $p_t = P(\pi_t > 0)$. Then,

$$p_t = P(S_t - \alpha_t B_t > 0) = P\left(\frac{S_t}{B_t} > \alpha_t\right) =$$

$$P\left(\sum_{i=1}^{t} \left[\ln(1+r_i) - \ln\left(1+r_f\right)\right] > \gamma_t\right) \approx P\left(\sum_{i=1}^{t} \left(r_i - r_f\right) > \gamma_t\right).$$

The last approximation is true because $\ln(1+x) \approx x$ for small x. Let F_t be the distribution function of $\sum_{i=1}^{t}\left(r_i - r_f\right)$. Then, $p_t = 1 - F_t(\gamma_t)$. Notice that, $p_t = p_{t-1} - F_t(\gamma_t) + F_{t-1}(\gamma_{t-1})$. One can see that

$$F_t(x) = E\left\{F_{t-1}\left(x + r_f - r_t\right)\right\}.$$

Assuming, $F_{t-1}(\gamma_t) - F_{t-1}(\gamma_{t-1}) \approx 0$, then,

$$p_t = p_{t-1} + E\{F_{t-1}(\gamma_t) + F_{t-1}\left(\gamma_t + r_f - r_t\right)\}.$$

Therefore, to maximize the gain probability with respect to γ_t, it is enough to solve the following equation

$$p_t = \max_{\gamma_t} \left(p_{t-1} + E\{F_{t-1}(\gamma_t) + F_{t-1}\left(\gamma_t + r_f - r_t\right)\}\right).$$

The last equation defines a dynamic programming framework for p_t. As soon as γ_t is determined, then $\alpha_t = \alpha_0 e^{\gamma_t}$ is derived. The rest of paper is organized as follows. In section 2, assuming normal distribution for r_t's the p_t, γ_t and α_t are derived. The exact and approximated solutions are derived. Section 3 examines the correlated r_t's. A real data set is analyzed.

2 Normal r_t's

In this section, first, the exact dynamic programming is presented. Then, the approximated solution is proposed. To this end, suppose that r_t's have normal distribution with mean μ_t and variance σ_t^2.

2.1 Exact solution

The F_t is distribution function of normal distribution with mean $\sum_{i=1}^{t}(\mu_i - r_f)$ and variance $\sum_{i=1}^{t}\sigma_i^2$. For two independent normally distributed random variables X and Y, it is true that $E(F_X(Y)) = P(X \le Y)$, thus,

$$E\{F_{t-1}(\gamma_t + r_f - r_t)\} = F_t(\gamma_t),$$

then,

$$p_t = \max_{\gamma_t}(p_{t-1} + F_{t-1}(\gamma_t) + F_t(\gamma_t)).$$

Example 1. Let $r_f = 0.05$ and $\mu_i = a\mu_{i-1}$ for $|a| < 1$ and $\sigma_t^2 = 0.16$. Notice that $p_0 = 0$. Let $\alpha_0 = 1$. The following Table gives the values of p_t, γ_t and α_t for $a = 0.2$.

Table 1. The values of p_t, γ_t and α_t

t	1	2	3	4	5	6	7	8	9	10
p(t)	0.172	0.22	0.234	0.239	0.351	0.36	0.367	0.373	0.378	0.383
α(t)	0.27	0.48	0.144	8.891	0.105	0.079	0.089	0.081	0.074	0.081
γ(t)	1.305	0.73	1.94	-2.185	2.26	2.54	2.425	2.515	2.61	2.5

2.2 Approximations

The following three propositions considers different approximations. To study the approximated solution, notice that

$$p_t - p_{t-1} = F_{t-1}(\gamma_{t-1}) - F_t(\gamma_t).$$

Proposition 1. For special case for $\gamma_t = -bt$, for some $b > 0$.

(a) The optimum b is

$$b = \frac{tF'_{t-1}(0) - (t-1)F'_t(0)}{t^2F''_{t-1}(0) + (t-1)^2F''_t(0)}.$$

(b) For fixed t, the optimum b is given by Newton-Raphson recursive formulae given by

$$b_{k+1} = b_k - \frac{g(b_k)}{g'(b_k)},$$

where $g(x) = tF'_t(-tx) - (t-1)F'_{t-1}(-(t-1)x)$.

Proof. Notice that

$$p_t - p_{t-1} = \max_b \{F_{t-1}(0) - btF'_{t-1}(0) + \frac{b^2t^2F''_{t-1}(0)}{2} -$$
$$F_t(0) + b(t-1)F'_t(0) + \frac{b^2(t-1)^2F''_t(0)}{2}\}.$$

This completes the proof of part (a). For part (b), notice that the optimum b, is obtained by letting the derivative of $F_{t-1}(\gamma_{t-1}) - F_t(\gamma_t)$ with respect to b equal to zero.

Proposition 2. If $\gamma_t - \gamma_{t-1} = hd_t$, then

$$p_t - p_{t-1} = F_{t-1}(\gamma_t) - F_t(\gamma_t) - hd_t F'_{t-1}(\gamma_t) + \frac{h^2 d_t^2}{2} F''_{t-1}(\gamma_t).$$

Proof. Use the straightforward application of Taylor series completes the proof. The following proposition studies the Monte Carlo approximation.

Proposition 3. The Monte Carlo approximation for $p_t - p_{t-1}$ is given by

$$\frac{1}{M}\sum_{i=1}^{M}(F_{t-1}(\gamma_t) - F_{t-1}(\gamma_t + r_f - r_t^i),$$

where $r_t^i, i = 1,2,\dots,M$ is a Monte Carlo sample from r_t.

3 Correlated r_t's

In financial applications, usually, it is not true to assume that r_t's are independent. In this section, it is assumed that they follow a first order autoregressive process with GARCH(1,1) errors, a rich practical model which works for many types of financial data sets, that is

$$r_t - r_f = \alpha + \beta(r_{t-1} - r_f) + \sigma_t z_t,$$

where z_t's are iid standard normal random variables and σ_t constitutes a GARCH(1,1) series. In the rest of this section, the existence the statistical arbitrage in stock of Intel corporation, a multinational technology company, is surveyed. The daily stock price are collected for period of study 20 Feb 2015 to 18 Feb 2016, including 250 log-returns. They are taken from Google-finance website[1]. Time series plot of return series is given by Fig. 1. The x-axis is time and y-axis is the logarithmic return.

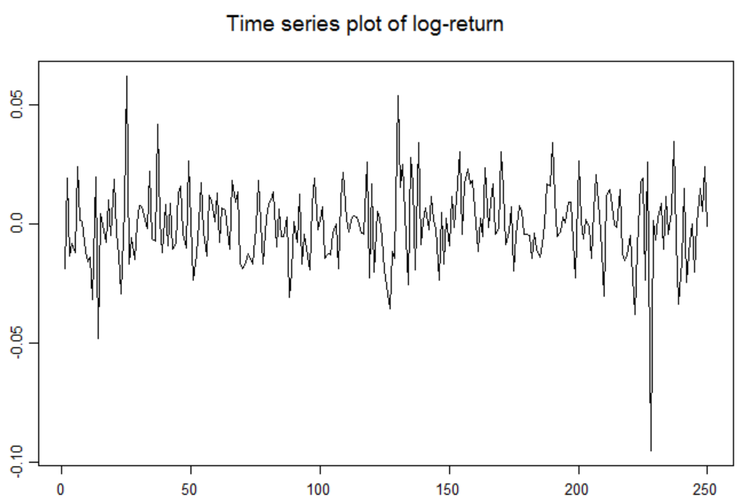

Fig. 1. Time series plot of log-returns

[1] http://www.google.com/finance?cid=284784

According to the Proposition 1 part (a), for $\gamma_t = -bt$, the optimum b is

$$b = \frac{tF'_{t-1}(0) - (t-1)F'_t(0)}{t^2F''_{t-1}(0) + (t-1)^2F''_t(0)}.$$

The quantities of $F'_{t-1}(0)$, $F'_t(0)$, $F''_{t-1}(0)$ and $F''_t(0)$ are derived using the Monte Carlo approximation and numerical differentiating. The plot of p_t is given as follows in Fig. 2. The x-axis is time and y-axis is the probability of statistical arbitrage.

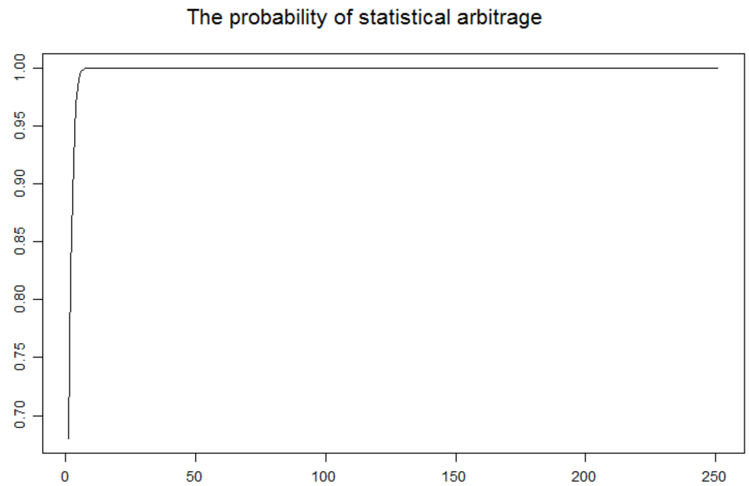

Fig. 2. The probability of statistical arbitrage over time

4 Conclusions

The application of Bellman equation is studied to fine the maximum probability of achieving the statistical arbitrage opportunities. Two cases of independent and correlated returns are studied and in both cases high probability of statistical arbitrage is derived.

Acknowledgement

The author thanks Editorial Board of Journal and two Referees for suggesting several comments to improve the paper.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Jarrow R, Teo M, Tse YK, Warachka M. Statistical arbitrage and market efficiency: Enhanced theory, robust tests and further applications. Technical reports. Johnson Graduate School of Management, Cornell University. USA; 2005.

[2] Pole A. Statistical arbitrage: Algorithmic trading insights and techniques. Wiley; 2007.

[3] Avellaneda M, Lee JH. Statistical arbitrage in the U.S. equities market. Technical reports. Finance Concepts SARL. Paris. France; 2008.

[4] Goldberg AV. Basic shortest path algorithms. Microsoft research. DIKU summer school on shortest path. USA; 2008.

[5] Aldridge I. High frequency trading: A practical guide to algorithmic strategies and trading systems. Wiley; 2009.

[6] Intriligator MD. Mathematical optimization and economic theory. SIAM; 2002.

[7] Zisserman A. Lecture C25 on optimization. Hilary Term; 2013.
Available: http:// www. robots. ox. ac. uk/~az/lectures/opt

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/14717