SDE course project

Stochastic Control Approach to Pair Trading

Jiahui Shui
Finance(Finance and Artificial Intelligence)
School of Finance
Southwestern University of Finance and Economics,benshui@smail.swufe.edu.cn

Abstract: In this paper, we propose a stochastic optimal control approach to pair trading. The constant volatility case is being considered at first, with an analytical solution. Then we extend this model into time varying volatility case, solving by numerical methodA. Finally, we use some numerical experiments and apply our model to the real financial market. We found that our strategy is better than holding the stocks.

Key words: stochastic control;pair trading;cointegration;finite difference;GMM

1. Introduction

Since the birth of pair trading, it has been widly used in many Hedge Funds. Consider two co-integrated and correlated stocks which trade at some spread. For a given period, we need to maximise the agents terminal utility of wealth subject to budget constraints.

In this paper, we follow the steps of [4]. First, we introduce the constant volatility case. By using some transformations, we are able to get a linear PDE instead a nonlinear one. Furthermore, we successfully find out the analytical solution to this problem with CRRA utility function.

After that, we consider the time-varying volatility case. For simplicity, we choose Constant Elasticity of Variance model. Unfortunately, it is impossible to get an analytical solution in this case.

Finally, we use some numerical results and apply our model into the real market to illustrate the efficience of pair trading strategy. We first compare the analytical solution in constant volatility case with holding stocks or cash. For timevarying case, we use numerical method to calculate the optimal threshold. The results show that pair trading strategy has great priority than other strategies.

2. Constant Volatility Case

In this section, we figure out the analytical solution of the stochastic optimal control under constant volatility case.

Granger's representation theorem enables the cointegrated vector time series to be expresssed as an error correction model. In discrete time, an error correction dynamic for the n-component asset price time series with $k(1 \leq k \leq n)$ cointegrating factors is defined as follows

$$
\begin{aligned}
\ln S_{i,t} - \ln S_{i,t-1} &= \mu_i + \sum_{j=1}^{k} \delta_{ij} z_{j,t-1} + \sigma_{i,t}\varepsilon_{i,t} \\
z_{j,t} &= a_j + b_j t + \sum_{i=1}^{n} c_{ij} \ln S_{i,t}
\end{aligned} \tag{1}
$$

We use the model derived in [3] as the diffusion limit of a discrete-time model. We fix the time horizon $T > 0$. With only two assets and the volatility is constant, the cointegrated asset prices $S_t^{(1)}$ and $S_t^{(2)}$ satisfy the following stochastic differential equations.

$$
\mathrm{d}\ln S_t^{(1)} = \left(\mu_1 - \frac{\sigma_1^2}{2} + \delta_1 z_t\right)\mathrm{d}t + \sigma_1 \mathrm{d}W_t^{(1)} \tag{2}
$$

$$
\mathrm{d}\ln S_t^{(2)} = \left(\mu_2 - \frac{\sigma_2^2}{2} + \delta_2 z_t\right)\mathrm{d}t + \sigma_2 \mathrm{d}W_t^{(2)} \tag{3}
$$

where $\mu_1, \mu_2, \sigma_1, \sigma_2, r$ are constants. And $W_t^{(1)}, W_t^{(2)}$ are Brownian Motions with correlation $\rho \in (-1, 1)$. The cointegrated vector z_t is defined as following:

$$
z_t = a + bt + \ln S_t^{(1)} + \beta \ln S_t^{(2)} \tag{4}
$$

Also, we suppose that there is a risk-free asset in the market with interest $r > 0$. The price of this asset is denoted by $S_t^{(0)}$, and follows

$$
\mathrm{d}S_t^{(0)} = r S_t^{(0)} \mathrm{d}t \tag{5}
$$

Therefore, by Itô lemma, we have that

$$
\begin{aligned}
\mathrm{d}z_t &= \left(b + \mu_1 - \frac{\sigma_1^2}{2} + \beta\mu_2 - \beta\frac{\sigma_2^2}{2} + \delta_1 z_t + \beta\delta_2 z_t\right)\mathrm{d}t + \sigma_1 \mathrm{d}W_t^{(1)} + \sigma_2 \mathrm{d}W_t^{(2)} \\
&= \alpha(\eta - z_t)\mathrm{d}t + \sigma_\beta \mathrm{d}W_t
\end{aligned} \tag{6}
$$

where we denote $\alpha = -\delta_1 - \beta\delta_2, \sigma_\beta = \sqrt{\sigma_1^2 + \beta^2\sigma_2^2}$, and

$$
\eta = \frac{1}{\alpha}\left(b + \mu_1 - \frac{\sigma_1^2}{2} + \beta\left(\mu_2 - \frac{\sigma_2^2}{2}\right)\right)
$$

Now we consider self-financing condition, $\pi_t^{(1)}$ and $\pi_t^{(2)}$ denote the fraction of wealth invested on two stocks respectively at time t. Therefore, we get the wealth dynamic as

$$
\mathrm{d}X_t = \pi_t^{(1)} X_t \frac{\mathrm{d}S_t^{(1)}}{S_t^{(1)}} + \pi_t^{(2)} X_t \frac{\mathrm{d}S_t^{(2)}}{S_t^{(2)}} + (1 - \pi_t^{(1)} - \pi_t^{(2)}) X_t \frac{\mathrm{d}S_t^{(0)}}{S_t^{(0)}} \tag{7}
$$

Our goal is to maximize the utility at time T. Define the optimal value function

$$
u(t, v, x, y) = \sup_{\pi_1, \pi_2} \mathbb{E}[U(V_T^{t,v,x,y,\pi_1,\pi_2})] \tag{8}
$$

We focus on the CRRA utility function

$$U(x)=\frac{x^{\gamma}}{\gamma}, \quad \gamma\in(0,1) \tag{9}$$

The HJB equation of above stochastic control problem can be written as

$$\begin{aligned}
u_t+\sup_{\pi_1,\pi_2}\Bigg\{&[\pi_1(\mu_1+\delta_1 z)+\pi_2(\mu_2+\delta_2 z)+r(1-\pi_1-\pi_2)]vu_v\\
&+\left(\mu_1-\frac{1}{2}\sigma_1^2+\delta_1 z\right)u_x+\left(\mu_2-\frac{1}{2}\sigma_2^2+\delta_2 z\right)u_y\\
&+(\pi_1\sigma_1^2 v+\pi_2\rho\sigma_1\sigma_2)u_{vx}+(\pi_2\sigma_2^2 v+\pi_1\rho\sigma_1\sigma_2)u_{vy}\\
&+\frac{1}{2}\left(\pi_1^2\sigma_1^2+\pi_2^2\sigma_2^2+2\pi_1\pi_2\rho\sigma_1\sigma_2\right)v^2 u_{vv}\\
&+\frac{1}{2}\sigma_1^2 u_{xx}+\frac{1}{2}\sigma_2^2 u_{yy}+\rho\sigma_1\sigma_2 u_{xy}\Bigg\}=0
\end{aligned} \tag{10}$$

To get the analytical solution of this PDE, we suppose that the value funciton can be separated as following

$$u(t,v,x,y)=\frac{1}{\gamma}v^{\gamma}h(t,z) \tag{11}$$

Note that $h(t,z)$ is a function of t,z instead of t,x,y. Then we can derive the PDE for $h(t,z)$

$$\begin{aligned}
h_t+\sup_{\pi_1,\pi_2}\Bigg\{&\Bigg[\frac{1}{2}\left(\pi_1^2\sigma_1^2+\pi_2^2\sigma_2^2+2\pi_1\pi_2\rho\sigma_1\sigma_2\right)\gamma(\gamma-1)\\
&+\pi_1(\mu_1+\delta_1 z)+\pi_2(\mu_2+\delta_2 z)+r(1-\pi_1-\pi_2)\Bigg]\gamma h\\
&+\Bigg[b+\left(\mu_1-\frac{1}{2}\sigma_1^2+\delta_1 z\right)+\beta\left(\mu_2-\frac{1}{2}\sigma_2^2+\delta_2 z\right)\\
&+\gamma\left(\pi_1\sigma_1^2+\pi_2\sigma_2^2\beta++\pi_2\rho\sigma_1\sigma_2+\pi_1\beta\rho\sigma_1\sigma_2\right)\Bigg]h_z\\
&+\frac{1}{2}\left(\sigma_1^2+\sigma_2^2\beta^2+2\beta\rho\sigma_1\sigma_2\right)h_{zz}\Bigg\}=0
\end{aligned} \tag{12}$$

Taking derivatives with respect π and π_2, we can get the following linear equation

$$\begin{cases}
\pi_1\sigma_1+\pi_2\rho\sigma_1\sigma_2=\dfrac{\mu_1-r+\delta_1 z}{1-\gamma}+\dfrac{\sigma_1^2+\beta\rho\sigma_1\sigma_2}{1-\gamma}\dfrac{h_z}{h}\\
\pi_1\rho\sigma_1\sigma_2+\pi_2\sigma_2^2=\dfrac{\mu_2-r+\delta_2 z}{1-\gamma}+\dfrac{\beta\sigma_2^2+\rho\sigma_1\sigma_2}{1-\gamma}\dfrac{h_z}{h}
\end{cases} \tag{13}$$

Note that here $\rho\neq\pm1$, otherwise there will not have any solution to π_1^* and π_2^*. Now the optimal controls given in terms of h and its derivatives are

$$\begin{aligned}
\pi_1^*&=\frac{\mu_1-r+\delta_1 z}{\sigma_1^2(1-\gamma)(1-\rho^2)}-\rho\frac{\mu_2-r+\delta_2 z}{\sigma_1\sigma_2(1-\gamma)(1-\rho^2)}+\frac{h_z}{(1-\gamma)h}\\
\pi_2^*&=\frac{\mu_2-r+\delta_2 z}{\sigma_2^2(1-\gamma)(1-\rho^2)}-\rho\frac{\mu_1-r+\delta_1 z}{\sigma_1\sigma_2(1-\gamma)(1-\rho^2)}+\frac{\beta h_z}{(1-\gamma)h}
\end{aligned} \tag{14}$$

Substituting the optimal controls (12) into (14) yields

$$
\begin{aligned}
& h_t-\left\{\frac{\gamma}{2(\gamma-1)\left(1-\rho^2\right)}\left[\frac{\left(\mu_1-r+\delta_1 z\right)^2}{\sigma_1^2}+\frac{\left(\mu_2-r+\delta_2 z\right)^2}{\sigma_2^2}\right.\right. \\
& \left.\left.+\ -2 \frac{\rho\left(\mu_1-r+\delta_1 z\right)\left(\mu_2-r+\delta_2 z\right)}{\sigma_1 \sigma_2}\right]-r \gamma\right\} h \\
& -\left\{\frac{1}{\gamma-1}\left[\left(\mu_1+\delta_1 z\right)+\beta\left(\mu_2+\delta_2 z\right)-r \gamma(1+\beta)\right]-b\right. \\
& \left.+\frac{1}{2}\left(\sigma_1^2-\beta \sigma_2^2\right)+\frac{1}{2}\left(\sigma_1^2-\beta \sigma_2^2\right)\right\} h_z+\frac{1}{2}\left(\sigma_1^2+\beta^2 \sigma_2^2\right. \\
& \left.+\ 2 \beta \rho \sigma_1 \sigma_2\right)\left(h_{zz}-\frac{\gamma}{\gamma-1} \frac{h_z^2}{h}\right)=0
\end{aligned} \tag{15}
$$

with the terminal condition $h(T,z)=1$. To reduce the nonlinear term in (15), we can do the following trasnformation with an unknown function ϕ

$$
h=\frac{1}{1-\gamma}\phi^{1-\gamma} \tag{16}
$$

Then we get the linear PDE of ϕ

$$
\begin{aligned}
& \phi_t+\left\{\frac{\gamma}{2(\gamma-1)^2\left(1-\rho^2\right)}\left[\frac{\left(\mu_1+\delta_1 z-r\right)^2}{\sigma_1^2}+\frac{\left(\mu_2+\delta_2 z-r\right)^2}{\sigma_2^2}\right.\right. \\
& \left.\left.-2 \frac{\rho\left(\mu_1-r+\delta_1 z\right)\left(\mu_2-r+\delta_2 z\right)}{\sigma_1 \sigma_2}\right]+\frac{r \gamma}{1-\gamma}\right\} \phi \\
& -\left\{\frac{1}{\gamma-1}\left[\left(\mu_1+\delta_1 z\right)+\beta\left(\mu_2+\delta_2 z\right)\right]-+\frac{r \gamma(1+\beta)}{\gamma-1}-b\right. \\
& \left.+\frac{1}{2}\left(\sigma_1^2+\beta \sigma_2^2\right)\right\} \phi_z+\frac{1}{2}\left(\sigma_1^2+\beta^2 \sigma_2^2+2 \beta \rho \sigma_1 \sigma_2\right) \phi_{zz}=0
\end{aligned} \tag{17}
$$

Note that (17) is an Euler type equation. To solve this equation, we can try

$$
\phi(t,z)=\exp\{f_0(t)+f_1(t)z+f_2(t)z^2\} \tag{18}
$$

Substituting (18) into (17), we can obtain the following equation

$$
A(t)z^2+B(t)z+C(t)=0 \tag{19}
$$

where A,B,C are the functions of t. Therefore, $A=B=C=0$ can give us three ordinanry differential equations (system). However, we just need the first and the second equation to get the stochastic optimal control since we can solve $f_1(t)$ and $f_2(t)$ without the third eqaution. For simplicity, we define the following constants

$$
\begin{aligned}
c_1 &= \sigma_1^2+\beta^2\sigma_2^2+2\beta\rho\sigma_1\sigma_2>0, c_2=\frac{\alpha}{2(1-\gamma)c_1}>0,\\
c_0 &= \frac{\alpha^2}{2(1-\gamma)^2c_1}-\frac{\gamma}{2(1-\gamma)^2\left(1-\rho^2\right)}\left(\frac{\delta_1^2}{\sigma_1^2}+\frac{\delta_2^2}{\sigma_2^2}-2\rho\frac{\delta_1\delta_2}{\sigma_1\sigma_2}\right)\\
c_3 &= 2b+\frac{2r\gamma(1+\beta)}{\gamma-1}-\frac{1}{\gamma-1}\left[2\left(\mu_1+\beta\mu_2\right)\right]-\left(\sigma_1^2+\beta\sigma_2^2\right)\\
c_4 &= \frac{\gamma}{2(\gamma-1)^2\left(1-\rho^2\right)}\left[\frac{\left(\mu_1-r\right)\delta_1}{\sigma_1^2}+\frac{\left(\mu_2-r\right)\delta_2}{\sigma_2^2}-2\rho\left(\frac{\delta_1\left(\mu_2-r\right)}{\sigma_1\sigma_2}+\frac{\delta_2\left(\mu_1-r\right)}{\sigma_1\sigma_2}\right)\right]
\end{aligned}
$$

Then the first and the second ordinary differential equations are

$$
f_2'(t)+2c_1(f_2(t)-c_2)^2-c_0=0 \tag{20}
$$

$$
f_1'(t)+[-2c_1c_2+2c_1f_2(t)]f_1(t)+c_3f_2(t)+c_4=0 \tag{21}
$$

coupled with the terminal condition $f_1(T)=f_2(T)=0$.

The solution of f_2 depends on the sign of c_0. From [1], when $c_0>0$, we have a particular solution to (20) this Riccati equation

$$
f_2(t)=c_2\left(1-\frac{c_0}{2c_1c_2^2}\right)\frac{\sinh\left(\sqrt{2c_1c_0}(T-t)\right)}{\sinh\left(\sqrt{2c_1c_0}(T-t)\right)+\frac{1}{c_2}\sqrt{\frac{c_0}{2c_1}}\cosh\left(\sqrt{2c_1c_0}(T-t)\right)} \tag{22}
$$

It is easy to varify that the (22) is indeed the solution to (20). Numerical methods also show that it is the solution.

With the solution of f_2, we can easily get the solution to f_1, which is given by

$$
f_1(t)=\int_t^T[c_3f_2(s)+c_4]\exp\left\{2c_1\int_t^s(f_2(u)-c_2)\mathrm{d}u\right\}\mathrm{d}s \tag{23}
$$

Finally, we get the stochastic optimal controls of analytical form

$$
\begin{aligned}
\pi_1^* &= \frac{\mu_1-r+\delta_1z}{\sigma_1^2(1-\gamma)\left(1-\rho^2\right)}-\rho\frac{\mu_2-r+\delta_2z}{\sigma_1\sigma_2(1-\gamma)\left(1-\rho^2\right)}+2f_2(t)z+f_1(t)\\
\pi_2^* &= \frac{\mu_2-r+\delta_2z}{\sigma_2^2(1-\gamma)\left(1-\rho^2\right)}-\rho\frac{\mu_1-r+\delta_1z}{\sigma_1\sigma_2(1-\gamma)\left(1-\rho^2\right)}+\beta\left(2f_2(t)z+f_1(t)\right)
\end{aligned} \tag{24}
$$

3. Time-varying Volatility Case

In this section, we do not suppose that the volatility are constants. A simple choice of time-varying volatility model is constant elasticity of variance, proposed by [2],

$$
\sigma_1(t,x)=\sigma_1e^{\theta_1x},\quad \sigma_2(t,y)=\sigma_2e^{\theta_2y}
$$

where $\theta_1, \theta_2 \in (-1, 0)$. This model is not very realistic, however, at least it reflects the negative relationship between the stock price and the volatility observed in the real stock market.

Our derivation remains under the time-varying volatility model, except that σ_1 and in the equations should change into $\sigma_1(t,x)$ and $\sigma_2(t,y)$ respectively. The HJB equation of the time-varying volatility case is the same as (10). Then we use the separation of variable

$$u(t,v,x,y) = \frac{1}{\gamma} v^\gamma g(t,x,y) \tag{25}$$

Then g satisfies the PDE:

$$\begin{aligned} & g_t + \sup_{\pi_1,\pi_2} \Big\{ \left[\pi_1 \left(\mu_1 + \delta_1 z\right) + \pi_2 \left(\mu_2 + \delta_2 z\right) + r\left(1 - \pi_1 - \pi_2\right)\right] \gamma g \\ & + \frac{1}{2} \left(\pi_1^2 \sigma_1^2(t,x) + \pi_2^2 \sigma_2^2(t,y) + 2\pi_1\pi_2\rho\sigma_1(t,x)\sigma_2(t,y)\right) \gamma(\gamma-1) g \\ & + \left[\mu_1 - \frac{1}{2}\sigma_1^2(t,x) + \delta_1 z + \gamma\left(\pi_1\sigma_1^2(t,x) + \pi_2\rho\sigma_1(t,x)\sigma_2(t,y)\right)\right] g_x \\ & + \left[\mu_2 - \frac{1}{2}\sigma_2^2(t,y) + \delta_2 z + \gamma\left(\pi_2\sigma_2^2(t,y) + \pi_1\rho\sigma_1(t,x)\sigma_2(t,y)\right)\right] g_y \\ & + \frac{1}{2}\sigma_1^2(t,x) g_{xx} + \frac{1}{2}\sigma_2^2(t,y) g_{yy} + \rho\sigma_1(t,x)\sigma_2(t,y) g_{xy} \Big\} \\ & = 0 \end{aligned} \tag{26}$$

for all $0 \le t < T$ coupled with $g(T,x,y) = 1$ for all x, y. Then the optimal controls involving g and its derivatives are given by

$$\begin{aligned} \pi_1^* &= \frac{\mu_1 - r + \delta_1 z}{\sigma_1^2(t,x)(1-\gamma)\left(1-\rho^2\right)} - \rho \frac{\mu_2 - r + \delta_2 z}{\sigma_1(t,x)\sigma_2(t,y)(1-\gamma)\left(1-\rho^2\right)} + \frac{g_x}{(1-\gamma)g}, \\ \pi_2^* &= \frac{\left(\mu_2 - r + \delta_2 z\right)}{\sigma_2^2(t,y)(1-\gamma)\left(1-\rho^2\right)} - \rho \frac{\mu_1 - r + \delta_1 z}{\sigma_1(t,x)\sigma_2(t,y)(1-\gamma)\left(1-\rho^2\right)} + \frac{g_y}{(1-\gamma)g} \end{aligned} \tag{27}$$

Substituting them back to the PDE

$$\begin{aligned} g_t - & \Bigg\{ \frac{\gamma}{2(\gamma-1)\left(1-\rho^2\right)} \left[\frac{\left(\mu_1 + \delta_1 z - r\right)^2}{\sigma_1^2(t,x)} + \frac{\left(\mu_2 + \delta_2 z - r\right)^2}{\sigma_2^2(t,y)} \right. \\ & \left. - 2\rho \frac{\left(\mu_1 - r + \delta_1 z\right)\left(\mu_2 - r + \delta_2 z\right)}{\sigma_1 \sigma_2} \right] + r\gamma \Bigg\} g \\ & - \frac{1}{\gamma - 1}\left[\mu_1 + \delta_1 z - r\gamma + \frac{1}{2}\sigma_1^2(t,x)\right] g_x \\ & - \frac{1}{\gamma - 1}\left[\mu_2 + \delta_2 z - r\gamma + \frac{1}{2}\sigma_2^2(t,y)\right] g_y \\ & - \frac{\gamma}{2(\gamma-1)} \left(\sigma_1^2(t,x) \frac{g_x^2}{g} + \sigma_2^2(t,y) \frac{g_y^2}{g} \right. \\ & \left. + \rho\sigma_1(t,x)\sigma_2(t,y) \frac{g_x g_y}{g} \right) + \frac{1}{2}\sigma_1^2(t,x) g_{xx} + \frac{1}{2}\sigma_2^2(t,y) g_{yy} \\ & + \rho\sigma_1(t,x)\sigma_2(t,y) g_{xy} = 0 \end{aligned} \tag{28}$$

To get rid of the nonlinear term in this PDE, we use the same technique as in previous section. Do the transform

$$g(t,x,y) = \frac{1}{1-\gamma}\phi^{1-\gamma}(\tau,x,y) \tag{29}$$

where $\tau = T - t$. Then we get the PDE of ϕ

$$\begin{aligned}
\phi_\tau - &\left\{ \frac{\gamma}{2(\gamma-1)^2\,(1-\rho^2)} \left[\frac{(\mu_1+\delta_1 z - r)^2}{\sigma_1^2(t,x)} + \frac{(\mu_2+\delta_2 z - r)^2}{\sigma_2^2(t,y)} \right.\right. \\
&\left.\left. -2\rho \frac{(\mu_1 - r + \delta_1 z)\,(\mu_2 - r + \delta_2 z)}{\sigma_1(t,x)\sigma_2(t,y)} \right] + \frac{r\gamma}{1-\gamma} \right\} \phi \\
&+ \left[\frac{\mu_1 + \delta_1 z - r\gamma}{\gamma - 1} + \frac{1}{2}\sigma_1^2(t,x) \right] \phi_x \\
&+ \left[\frac{\mu_2 + \delta_2 z - r\gamma}{\gamma - 1} + \frac{1}{2}\sigma_2^2(t,y) \right] \phi_y - \frac{1}{2}\sigma_1^2(t,x)\phi_{xx} \\
&- \frac{1}{2}\sigma_2^2(t,y)\phi_{yy} - \rho\sigma_1(t,x)\sigma_2(t,y)\phi_{xy} = 0
\end{aligned} \tag{30}$$

coupled with the initial condition $\phi(0,x,y) = (1-\gamma)^{\frac{1}{1-\gamma}}$. Finally, we get the optimal controls involving ϕ and its derivatives

$$\begin{aligned}
\pi_1^* &= \frac{\mu_1 - r + \delta_1 z}{\sigma_1^2(t,x)(1-\gamma)\,(1-\rho^2)} - \rho \frac{\mu_2 - r + \delta_2 z}{\sigma_1(t,x)\sigma_2(t,y)(1-\gamma)\,(1-\rho^2)} + \frac{\phi_x}{\phi} \\
\pi_2^* &= \frac{\mu_2 - r + \delta_2 z}{\sigma_2^2(t,y)(1-\gamma)\,(1-\rho^2)} - \rho \frac{\mu_1 - r + \delta_1 z}{\sigma_1(t,x)\sigma_2(t,y)(1-\gamma)\,(1-\rho^2)} + \frac{\phi_y}{\phi}
\end{aligned} \tag{31}$$

We can see that the first two components in both controls are the same as in the constant volatility case except for the volatility term are time-varying now. It is impossible to get the analytical solution in time-varying case, we will use numerical method to solve this problem later.

4. Numerical Result

In this section, we are going to do some numerical experiments to illustrate our result better.

4.1. Constant Volatility Case

From previous sections, we know that when the volatility is constant, then the analytical solution exists. Parameters we used in this part are shown below

$$\begin{aligned}
&r = 0.01, \mu_1 = 0.2, \mu_2 = 0.08, \sigma_1 = 0.4, \sigma_2 = 0.45 \\
&\beta = -0.6, a = -0.01, b = -0.01, \delta_1 = -0.1, \delta_2 = 0.1 \\
&\theta_1 = -0.2, \theta_2 = -0.15, \gamma = 0.1, \rho = 0.5, S_0^{(1)} = 12.18 \\
&S_0^{(2)} = 20.09
\end{aligned}$$

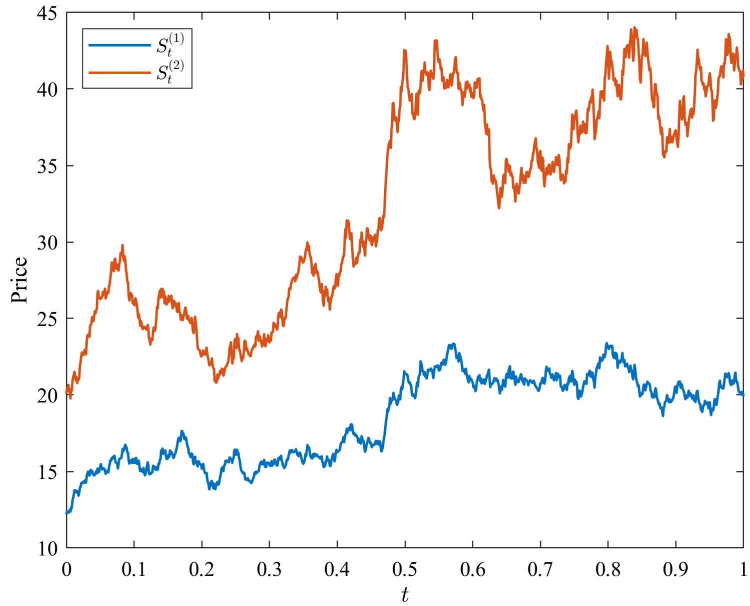

Figure 1 The stock prices in constant volatility case

The analytical solution of $f_1(t)$ and $f_2(t)$ is shown as below. From this figure, we can see that $f_1(t)$ is increasing and $f_2(t)$ is decreasing.

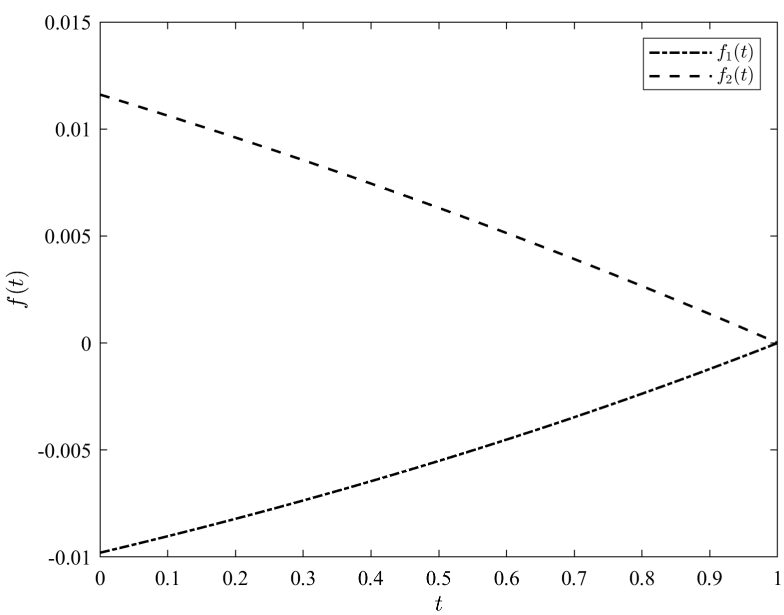

Figure 2 The analytical solution of $f_1(t)$ and $f_2(t)$

Then we can get the optimal controls $\pi_t^{(1)}$ and $\pi_t^{(2)}$

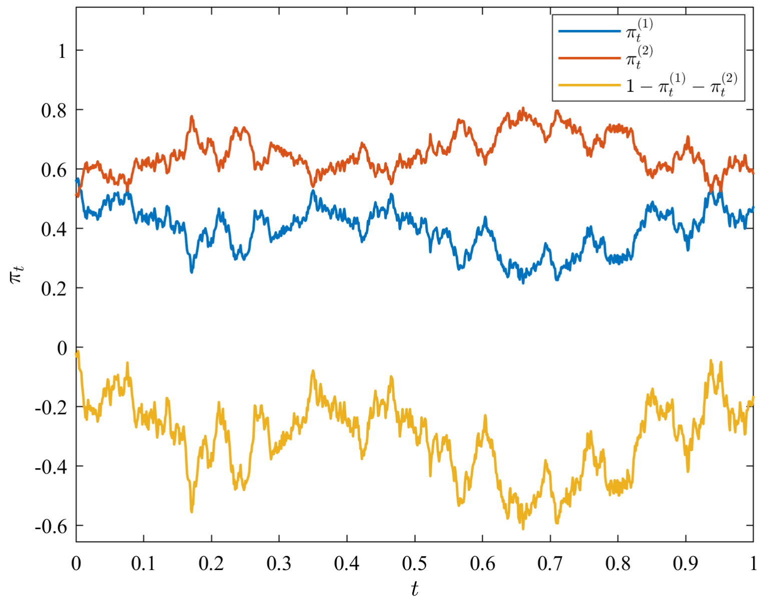

Figure 3 Optimal controls in constant volatility case

Now we calculate the total wealth by four different strategies

1. Keep cash
2. Invest all the money to Stock 1
3. Invest all the money to Stock 2
4. Using Pair Trading Strategy

The dynamics of total wealth using these four strategies are shown in fig 4. We can see that pair trading strategy slightly beats other strategies.

4.2. Time Varying Volatility Case

Since the equation (30) can not be solved analytically, we try to use numerical method to handle it. We set up the parameters the same as the constant volatility case, except for

$$\theta_1 = -0.2, \quad \theta_2 = -0.15$$

First we plot the dynamics of the stock prices

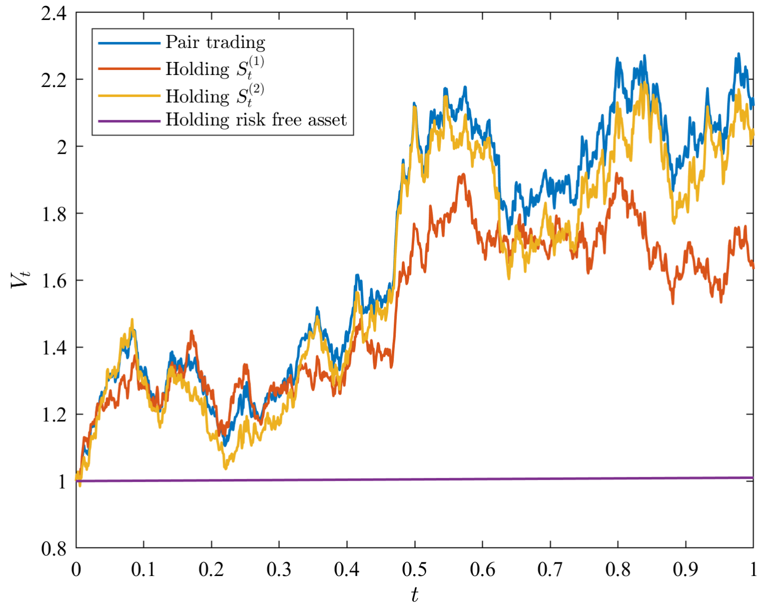

Figure 4 **The dynamics of total wealth using different strategies in constant volatility case**

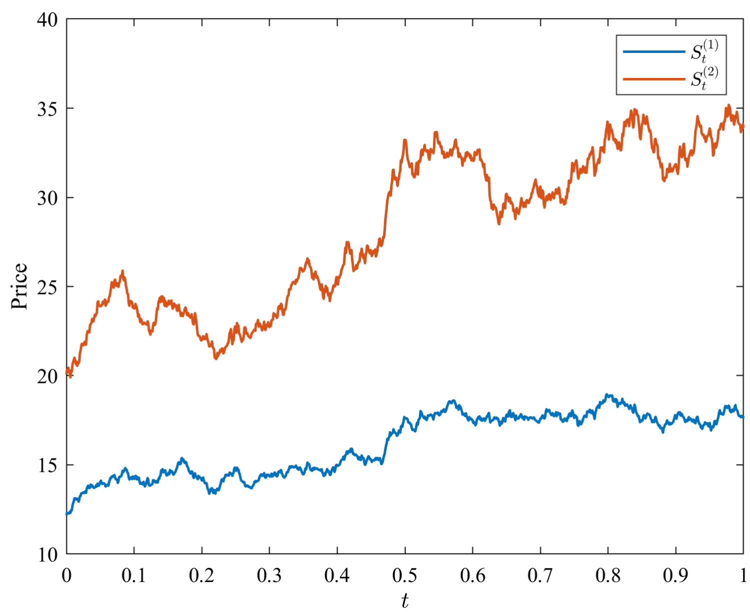

Figure 5 **The stock prices in time-varying volatility case**

Then we need to use finite difference method to solve the (30). Details are in the Appendix. We show the function ϕ in the fig 6. After that, the optimal controls can be easily calculated, which are shown in fig 7

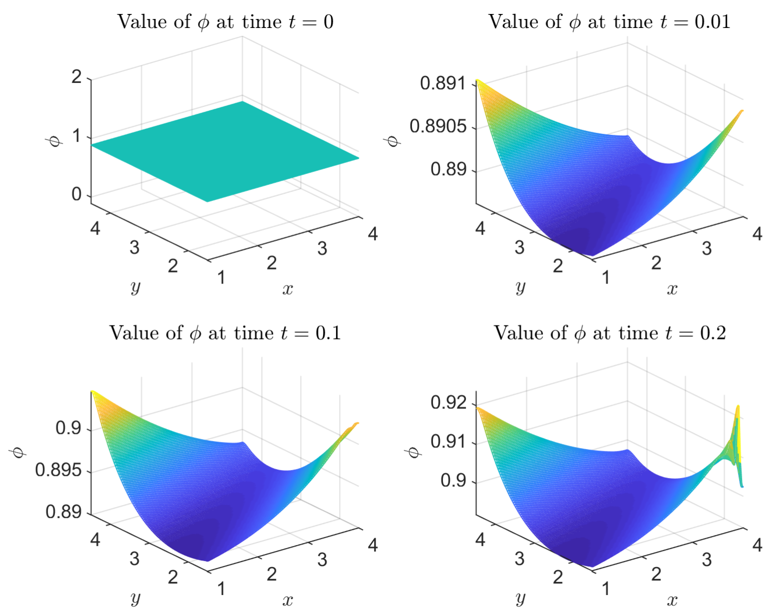

Figure 6 **The value of function ϕ at** $t = 0.001, 0.01, 0.1, 0.2$

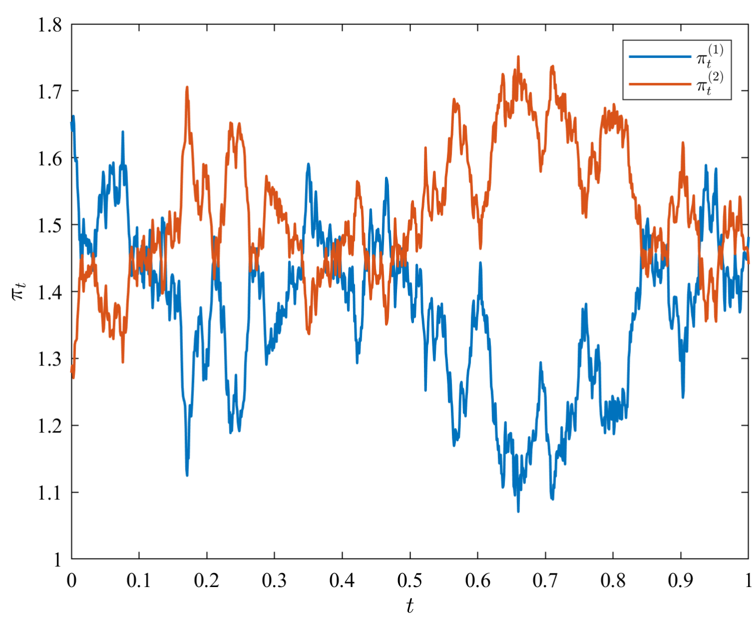

Figure 7 **Optimal controls in timevarying volatility case**

The total wealth under different strategy are also shown in the figure

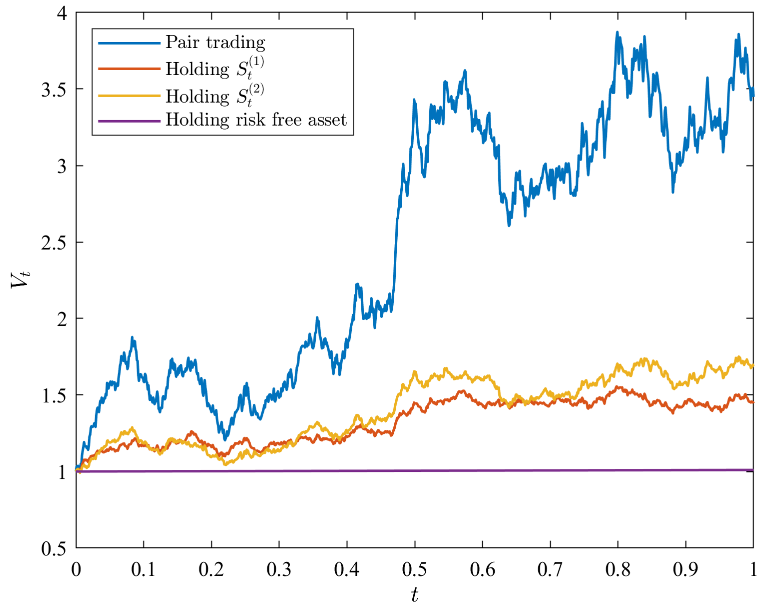

Figure 8 **The dynamics of total wealth using different strategies in time varying volatility case**

5. Application in Real Market

Now, let us use Generalized Method of Moments to estimate the following parameters:

$$\mu_1, \sigma_1, \delta_1, \theta_1, \mu_2, \sigma_2, \delta_2, \theta_2, \rho, a, b, \beta$$

We use Apple and Microsoft stock prices as our example.

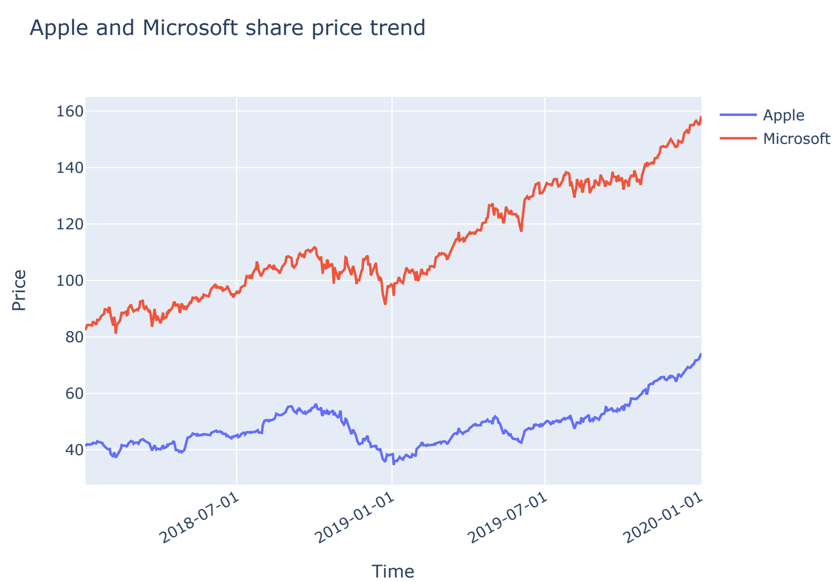

Figure 9 **Price of Apple and Microsoft**

The GMM scheme is shown in Appendix. The GMM in just-identified case yields the following result :

$$\mu_1 = 0.6324, \sigma_1 = 0.2438, \delta_1 = -0.00096, \theta_1 = -0.00553$$

$$\mu_2 = 0.8360, \sigma_2 = 0.3401, \delta_2 = -0.0013, \theta_2 = -0.0549$$

$$\rho = -0.0439, a = -113.6311, b = -174.3466, \beta = 171.8160$$

Solving the model with parameters above, we get $\pi_t^{(1)}$ and $\pi_t^{(2)}$. Finally, we plot the dynamics of total wealth. In fig 11, we can see that the computed strategy generate a profit of roughly 500%, however, it was beated by holding strategy before $t = 1.8$. The reason may be that our estimation is not that accurate, then the strategy may not be that efficient.

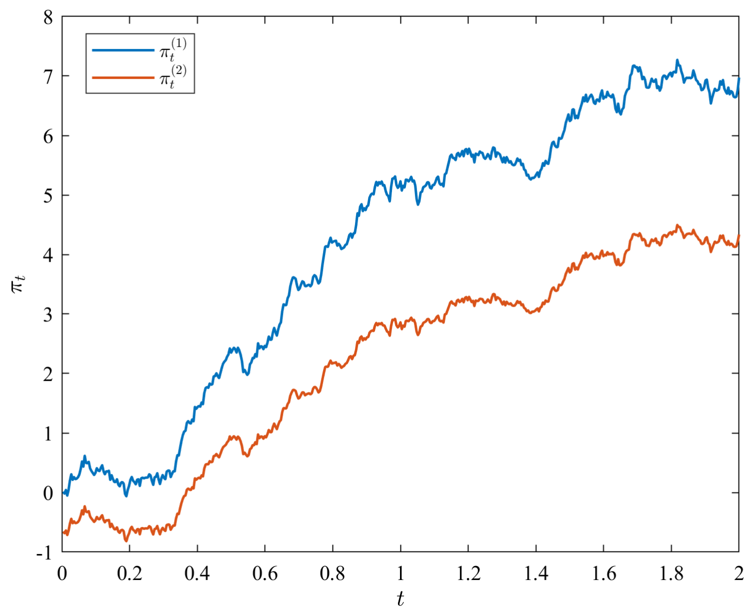

Figure 10 The propotion of total wealth invested to Apple($\pi_t^{(1)}$) and Microsoft($\pi_t^{(2)}$)

6. Conclusion

In this article, we derived pair trading strategy from stochastic optimal control perspect. First, we successfully figure out the analytical solution in the constant volatility case. Then we extend our model into the time varying volatility case, however, in this situation, it is impossible to derive an analytical solution.

After setting up those two models. We use numerical experiments to have a taste of how powerful this strategy is in ”not real world”. However, when we apply our method into the real market, we see its shortage. First, we have cash in short, which means we have to borrow money from banks.

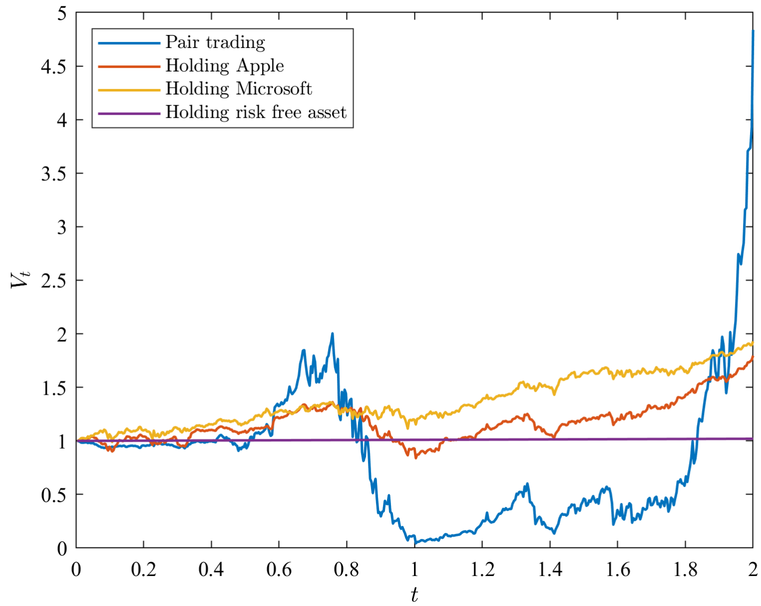

Figure 11 **The dynamic of total wealth in real market**

This may bring tons of risk to become bankrupt. Second, transaction cost is not being considered, our strategy may not that powerful in the real world. There are two ways I want to improve this model. One is using GARCH model for time varying volatility. The other one is adding transaction costs into the model.

All models are wrong, but at least they can depict parts of the real world.

APPENDIX

Appendix A. Finite Difference Scheme

We first discretise x, y, z as

$$x_i = x_{min} + i\Delta x, \quad y_j = y_{min} + j\Delta y$$
$$z_{i,j}^k = a + b(T - k\Delta t) + x_i + \beta y_j$$

where $(i,j) \in \{0,1,\cdots,I\} \times \{0,1,\cdots,J\}$, with $I\Delta x = x_{max} - x_{min}$ and $J\Delta y = y_{max} - y_{min}$. Also, $k \in \{0,1,\cdots,K\}$ with $K\Delta t = T$.

Denote $\phi_{i,j}^k := \phi(t_k, x_i, y_j)$, then we can use following to approximate the derivatives.

$$\phi_x \approx \frac{\phi_{i+1,j}^k - \phi_{i,j}^k}{\Delta x}$$
$$\phi_{xx} \approx \frac{\phi_{i-1,j}^k - 2\phi_{i,j}^k + \phi_{i+1,j}^k}{\Delta x^2}$$
$$\phi_t \approx \frac{\phi_{i,j}^{k+1} - \phi_{i,j}^k}{\Delta t}$$

The mixture term is depend on the sign of ρ. If $\rho \geq 0$, we have

$$\phi_{xy} \approx \frac{2\phi_{i,j}^k + \phi_{i+1,j+1}^k + \phi_{i-1,j-1}^k - \phi_{i+1,j}^k - \phi_{i-1,j}^k - \phi_{i,j+1}^k - \phi_{i,j-1}^k}{2\Delta x \Delta y}$$

If $\rho < 0$, we use

$$\phi_{xy} \approx \frac{-2\phi_{i,j}^k - \phi_{i+1,j-1}^k - \phi_{i-1,j+1}^k + \phi_{i+1,j}^k + \phi_{i-1,j}^k + \phi_{i,j+1}^k + \phi_{i,j-1}^k}{2\Delta x \Delta y}$$

Substitute them into (30) we can get the numerical solution.

Appendix B. Generalized Method of Moments Scheme

First, we discretize the S_t.

$$\ln S_{t+1}^{(i)} - \ln S_t^{(i)} = \left(\mu_i - \frac{1}{2}\sigma_i^2 e^{2\theta_i \ln S_t^{(i)}} + \delta_i \left(a + bt + \ln S_t^{(1)} + \beta \ln S_t^{(2)}\right)\right)\Delta t + \epsilon_{t+1}^i \tag{32}$$

where ϵ_t^i satisfies

$$\begin{aligned} &\mathbb{E}\left(\epsilon_{t+1}^i\right) = 0 \\ &\mathbb{E}\left[\left(\epsilon_{t+1}^i\right)^2\right] = \sigma_i^2 e^{2\theta_i \ln S_t^i}\Delta t, \quad \text{for } i = 1,2, \\ &\mathbb{E}\left[\epsilon_{t+1}^1 \epsilon_{t+1}^2\right] = \rho\sigma_1 e^{\theta_1 \ln S_t^1}\sigma_2 e^{\theta_2 \ln S_t^2}\Delta t \end{aligned} \tag{33}$$

Then we have that

$$\epsilon_{t+1}^i = \ln S_{t+1}^i - \ln S_t^i - \left(\mu_i - \frac{1}{2}\sigma_i^2 e^{2\theta_i \ln S_t^i} + \delta_i\left(a + bt + \ln S_t^1 + \beta \ln S_t^2\right)\right)\Delta t \tag{34}$$

Now we can define the vector $f_t(\lambda)$ of moment functions

$$f_t(\lambda) = \begin{bmatrix} \epsilon_{t+1}^1 \\ \epsilon_{t+1}^2 \\ \left(\epsilon_{t+1}^1\right)^2 - \sigma_1^2 e^{2\theta_1 \ln S_t^1}\Delta t \\ \left(\epsilon_{t+1}^2\right)^2 - \sigma_2^2 e^{2\theta_2 \ln S_t^2}\Delta t \\ \epsilon_{t+1}^1\epsilon_{t+1}^2 - \sigma_1 e^{\theta_1 \ln S_t^1}\sigma_2 e^{\theta_2 \ln S_t^2}\rho\Delta t \end{bmatrix} \otimes \begin{bmatrix} 1 \\ t \\ \ln S_t^1 \\ \ln S_t^2 \end{bmatrix} \tag{35}$$

where $\otimes$ denotes the Kronecker product. We have that $\mathbb{E}[f_t(\lambda)] = 0$. Then we get sample average

$$g(\lambda) = \frac{1}{T}\sum_{t=1}^{T} f_t(\lambda)$$

The goal is to minimize

$$J(\lambda) = g(\lambda)^T W g(\lambda)$$

where W is positive-definite weighting matrix. And our estimator is

$$\hat{\lambda} = \arg\min_{\lambda \in \Theta} J(\lambda)$$

References

[1] Fred Espen Benth and Kenneth Hvistendahl Karlsen. A note on merton's portfolio selection problem for the schwartz mean-reversion model. *Stochastic analysis and applications*, 23(4):687–704, 2005.

[2] John C Cox and Stephen A Ross. The valuation of options for alternative stochastic processes. *Journal of financial economics*, 3(1-2):145–166, 1976.

[3] Jin-Chuan Duan and Stanley R Pliska. Option valuation with co-integrated asset prices. *Journal of Economic Dynamics and Control*, 28(4):727–754, 2004.

[4] Thomas Nanfeng Li and Agnès Tourin. Optimal pairs trading with time-varying volatility. *International Journal of Financial Engineering*, 3(03):1650023, 2016.