Proceedings of the 69th Annual Conference of the Institute of Systems, Control and Information Engineers (ISCIE), Kobe, May 25-27, 2025

13C2

Nonlinear Parameter Estimation for the Black-Scholes Process

Tomoji TAKATSU (TOKEN C. E. E. Consultants Co.,Ltd.)

Abstract In this study, we derive a nonlinear estimation mechanism for the time-invariant parameters of the Black-Scholes process by representing them as system state variables. The efficacy of this approach is validated through empirical analysis using public available stock data.

1 Introduction

Parameter estimation for stochastic differential equations is an important research area with numerous applications. This field encompasses not only linear systems but also nonlinear ones. In this paper, we focus on the Black-Scholes process (BSP) because stock data is readily available from various websites. The BSP model has been extensively studied, particularly in the context of modeling volatility changes. Several models, such as the CEV, SABR, and GARCH models, are based on the BSP, underscoring the significance of volatility modeling [1]. Rather than attempting to improve the structure of these existing models, we concentrate on developing more effective methods for estimating their constant parameters. Thus, our objective is to establish a parameter estimation method under the assumption that these parameters are constant.

2 Drift Parameter Estimation

The drift and volatility parameters estimations are developed within the framework of Kalman Filter. The Black-Scholes Process, which is the model of stock prices and ETF, is as follows[2],

$$dS(t) = \mu S(t)dt + \sigma S(t)dw(t), \;\; S(0) = S_0, \qquad (1)$$

where $t \in [0, \infty)$, $S(t) (> 0)$ is a series of stock prices or price of ETF, $w(t)$ is standard Browian Motion Process, μ $(0 < |\mu| < \infty)$ is a constant drift parameter and σ $(0 < |\sigma| < \infty)$ is a constant volatility parameter.

We set $x(t) = [\mu \; \sigma]^T$ and the observation and estimation mechanism as follows,

$$dy(t) = h_1(t)x(t)dt + h_2(t)x(t)dw(t) \qquad (2)$$

$$d\hat{x}(t|t) = \tilde{K}(t)\left[dy(t) - h_1(t)\hat{x}(t|t)dt\right], \qquad (3)$$

$$\hat{x}(0|0) = \hat{x}_0,$$

where $y(t) = S(t)$ is observation data, $dx(t)/dt = 0$, $h_1(t) = [S(t) \; 0]$, $h_2(t) = [0 \; S(t)]$, $\hat{x}(t|t) = \mathcal{E}\{x(t)|\mathcal{Y}_t\} = [\hat{\mu}(t|t) \; \hat{\sigma}(t|t)]^T$, $\tilde{K}(t) = [\tilde{K}_1(t) \; \tilde{K}_2(t)]^T$ is an estimation gain vector and where $\mathcal{Y}_t = \{y(\tau), 0 \le \tau \le t\}$.

(3) is equivalent to the following as

$$d\hat{\mu}(t|t) = \tilde{K}_1(t)\left[dy(t) - S(t)\hat{\mu}(t|t)dt\right], \qquad (4)$$

$$d\hat{\sigma}(t|t) = \tilde{K}_2(t)\left[dy(t) - S(t)\hat{\mu}(t|t)dt\right]. \qquad (5)$$

Since $K(\bullet)$ is independent of μ and σ, $\hat{\sigma}(\bullet|\bullet)$ is not updated by (5). Therefore, our strategy must be modified to estimate the unknown parameter σ. This is because, by analogy with the Kalman filter, $\tilde{K}(\bullet)$ is independent of both μ and σ. The novle strategy is as follows. Initially, we construct an estimation mechanism for μ, assuming σ is given. Subsequently, we develop another estimation mechanism for σ, and the estimated σ is used in the estimation mechanism for μ.

Assuming σ is given, we set the following observation and estimation mechanisms,

$$dy(t) = S(t)\mu dt + S(t)\sigma dw(t), \qquad (6)$$

$$d\hat{\mu}(t|t) = K(t)\left[dy(t) - S(t)\hat{\mu}(t|t)dt\right], \qquad (7)$$

$$\hat{\mu}(0|0) = \hat{\mu}_0,$$

where $dy(t) \, (= dS(t))$ is observation data, $K(t)$ is estimation gain and $\hat{\mu}(t|t) = \mathcal{E}\{\mu|\,\mathcal{Y}_t\}$, and where $\mathcal{Y}_t = \{y(\tau), 0 \le \tau \le t\}$.

Therefore, we have formulated the parameter estimation problem as a state estimation problem. In order to have the estimation gain in (7), we set the estimation error covariance $P(t|t)$ as follows,

$$P(t|t) := \mathcal{E}\left\{[\mu - \hat{\mu}(t|t)]^2 \middle| \mathcal{Y}_t\right\}. \qquad (8)$$

Assuming the parameter estimation is optimal in some respect, then the $dP(t|t) = P(t+dt|t+dt) - P(t|t)$ is obtained as follows with $\mathcal{E}\{\bullet|\mathcal{Y}_{t+dt}\} \cong \mathcal{E}\{\bullet|\mathcal{Y}_t\}$, $\mathcal{E}\{dw(t)|\mathcal{Y}_t\} = 0$, $\mathcal{E}\{dw^2(t)|\mathcal{Y}_t\} = dt$, $d\mu/dt = 0$ and $\hat{\mu}(t+dt|t+dt) \cong \hat{\mu}(t|t) + d\hat{\mu}(t|t)$,

$$dP(t|t) = -2K(t)S(t)P(t|t)dt + K^2(t)S^2(t)\sigma^2 dt + o(dt^{\frac{3}{2}}). \qquad (9)$$

From (9), we have following differential equation as

$$\frac{dP(t|t)}{dt} = -2K(t)S(t)P(t|t) + K^2(t)S^2(t)\sigma^2. \qquad (10)$$

The estimation gain $K(t)$ is designed to achieve minimum the estimation error variance $P(t|t)$ satisfied (10). Next, we set the cost functional $V[t, P(t|t)]$ as following scalar function,

$$V[t, P(t|t)] := \min_{\substack{K(\tau)\\0\le\tau\le t}} P(t|t). \qquad (11)$$

Substituting (10) into (11), we have

$$V[t, P(t|t)] = \min_{\substack{K(\tau)\\0\le\tau\le t}} \left\{P(0|0) + \int_0^t \frac{dP(\tau|\tau)}{d\tau}d\tau\right\}$$

$$= P(0|0) + \min_{\substack{K(\tau)\\0\le\tau\le t}} \left\{ \int_0^t \left[-2K(\tau)S(\tau)P(\tau|\tau) + K^2(\tau)S^2(\tau)\sigma^2\right] d\tau \right\}. \quad (12)$$

While the dynamic programming is a common method to solve the second term on the right-hand side of (12), we employ the calculus of variations in this study. Therefore, we proceed to solve the following equation.

$$\frac{\partial}{\partial \epsilon} \int_0^t \{-2[K(\tau)+\epsilon Q(\tau)]S(\tau)P(\tau|\tau) + [K(\tau)+\epsilon Q(\tau)]^2(\tau)S^2(\tau)\sigma^2\} d\tau \Big|_{\epsilon=0} = 0, \quad (13)$$

where $Q(\tau)$ is in the same mathematical class of the optimal $K(\tau)$ with $Q(0) = Q(t) = 0$ [3].

From (13), we derive the following identity with respect to Q.

$$\int_0^t Q(\tau)\left\{K(\tau)S^2(\tau)\sigma^2 - S(\tau)P(\tau|\tau)\right\} d\tau = 0. \quad (14)$$

Consequently, we obtain a relation between $K(\tau)$ and $P(\tau|\tau)$, which is given by

$$K(\tau) = \frac{P(\tau|\tau)}{S(\tau)\sigma^2} \ \ (0 \le \tau \le t). \quad (15)$$

From (10) and (15), the following differential equation is hold,

$$\frac{dP(t|t)}{dt} = -\frac{P^2(t|t)}{\sigma^2}, \ \ P(0|0) = P_0. \quad (16)$$

The solotion of (16) is as follows,

$$P(t|t) = \frac{P_0\sigma^2}{P_0 t + \sigma^2}. \quad (17)$$

The convergence of $\hat{\mu}(t|t)$, which is $\lim_{t\to\infty} P(t|t) = 0$, is shown as follows,

$$\lim_{t\to\infty} P(t|t) = \lim_{t\to\infty} \frac{P_0\sigma^2}{P_0 t + \sigma^2} = 0. \quad (18)$$

Assuming $\lim_{t\to\infty} \mathcal{E}\{\bullet|\mathcal{Y}_t\} = \mathcal{E}\{\bullet\}$, from (8), (18) can be expressed as follows,

$$\underset{t\to\infty}{\mathrm{l.i.m.}}\ \hat{\mu}(t|t) = \mu. \quad (19)$$

Substituting (15) and (17) into (7), we consequently obtain the following estimation,

$$d\hat{\mu}(t|t) = \frac{P_0}{S(t)\left(P_0 t + \sigma^2\right)} \left[dy(t) - S(t)\hat{\mu}(t|t)dt\right]. \quad (20)$$

The estimation mechanism for μ is derived assuming σ is given. Since σ is unknown, we must use estimated σ which is not always true value of σ. Consequently, we assess the impact on the estimate of μ when employing a value of σ that deviates from its true value. Taking the partial derivative of both sides of equation (17) with respect to σ^2, we obtain the following equation

$$\frac{\partial P(t|t)}{\partial \sigma^2} = \frac{P_0^2 t}{\left(P_0 t + \sigma^2\right)^2}. \quad (21)$$

From (21), we have

$$\lim_{t\to\infty} \frac{\partial P(t|t)}{\partial \sigma^2} = \lim_{t\to\infty} \frac{P_0^2 t}{\left(P_0 t + \sigma^2\right)^2} = 0 \quad (22)$$

$$\frac{\partial P(t|t)}{\partial \sigma^2}\Big|_{t=0} = \frac{P_0^2 t}{\left(P_0 t + \sigma^2\right)^2}\Big|_{t=0} = 0. \quad (23)$$

Taking the partial derivative of both sides of equation (17) with respect to t, the following equation holds,

$$\frac{\partial}{\partial t}\left(\frac{\partial P(t|t)}{\partial \sigma^2}\right) = \frac{P_0^2\sigma^2 - P_0^3 t}{\left(P_0 t + \sigma^2\right)^3}. \quad (24)$$

From above equations, (21) takes the maximam value of $P_0/4\sigma^2$ when $t = \sigma^2/P_0$.

We have shown that σ in equation (20) can be replaced with some suitable approximation.

3 Volatility Parameter Estimation

The volatility parameter estimation is developed separate from Kalman Filter. Applying Itô formula (1), we derive the following expressions.

$$\mathcal{E}\left\{\left(d\left[\ln S(t)\right]\right)^2\right\} = \mathcal{E}\left\{\left[\left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dw(t)\right]^2\right\} = \sigma^2 dt + o(dt^{\frac{3}{2}}). \quad (25)$$

As $dt \to 0$, the following equation holds.

$$\sigma^2 dt = \mathcal{E}\left\{\left(d\left[\ln S(t)\right]\right)^2\right\}. \quad (26)$$

Now we define the estimated σ^2 as follows,

$$\hat{\sigma}^2(t|t) := \mathcal{E}\left\{\sigma^2|\mathcal{Y}_t\right\}. \quad (27)$$

Applying the operator $\mathcal{E}\{\bullet|\mathcal{Y}_t\}$ to both sides of (26), we obtain the following equations with the property of $\mathcal{E}\{\mathcal{E}\{\bullet\}|\mathcal{Y}_t\} = \mathcal{E}\{\bullet|\mathcal{Y}_t\}$.

$$\hat{\sigma}^2(t|t)dt = \mathcal{E}\left\{\left(d\left[\ln S(t)\right]\right)^2|\mathcal{Y}_t\right\} \quad (28)$$

$$\begin{aligned} d\hat{\sigma}^2(t|t)dt &:= \hat{\sigma}^2(t+dt|t+dt)dt - \hat{\sigma}^2(t|t)dt \\ &= \mathcal{E}\left\{\left(d\left[\ln S(t+dt)\right]\right)^2|\mathcal{Y}_{t+dt}\right\} \\ &\quad -\mathcal{E}\left\{\left(d\left[\ln S(t)\right]\right)^2|\mathcal{Y}_t\right\}. \end{aligned} \quad (29)$$

In (28), replacing the expectation operation with the arithmetic mean, we obtain the following equation.

$$\hat{\sigma}^2(t|t)dt \cong \hat{\sigma}^2(t_N|t_N)\Delta t$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left[\ln S(t_i) - \ln S(t_{i-1})\right]^2, \quad (30)$$

where $t_0 < t_1 < t_2 < \cdots < t_N = N\Delta t = t$, and where $t_j - t_{j-1} = \Delta t$ $(j = 1, 2, \cdots, N)$.

From (29) and (30), we obtain the following estimation mechanism with $y(t) = S(t)$ and $d[\ln y(t)] = \ln y(t+dt) - \ln y(t)$.

$$\begin{aligned} d\hat{\sigma}^2(t|t) &\cong \frac{\sum_{i=1}^{N+1}\left[\ln S(t_i) - \ln S(t_{i-1})\right]^2}{(N+1)\Delta t} \\ &\quad - \frac{\sum_{i=1}^{N}\left[\ln S(t_i) - \ln S(t_{i-1})\right]^2}{N\Delta t} \\ &\cong \frac{\left[\ln y(t_{N+1}) - \ln y(t_N)\right]^2}{(N+1)\Delta t} - \frac{\hat{\sigma}^2(t_N|t_N)\Delta t}{(N+1)\Delta t} \\ &\cong \frac{1}{t+\Delta t}\left\{(\Delta[\ln y(t)])^2 - \hat{\sigma}^2(t|t)\Delta t\right\} \quad (31) \\ &\quad \hat{\sigma}^2(0|0) = \hat{\sigma}_0^2. \end{aligned}$$

Moreover, from eq. (31), we have discrete equation:

$$\begin{aligned} &\hat{\sigma}^2(t_{N+1}|t_{N+1}) \\ &\quad = \frac{N\hat{\sigma}^2(t_N|t_N)\Delta t + \left[\ln y(t_{N+1}) - \ln y(t_N)\right]^2}{(N+1)\Delta t} \\ &\quad = \frac{\left[\ln y(t_{N+1}) - \ln y(t_N)\right]^2}{(N+1)\Delta t} + \frac{N}{N+1}\hat{\sigma}^2(t_N|t_N). \quad (32) \end{aligned}$$

As $\hat{\sigma}^2(t|t) < \infty$ comes from $|\sigma| < \infty$, (31) implies that $\lim_{t\to\infty} d\hat{\sigma}^2(t|t) = 0$ i.e. $\hat{\sigma}^2(t|t)$ is the Cauchy sequence. Assuming $\lim_{t\to\infty} \mathcal{E}\{\bullet|\mathcal{Y}_t\} = \mathcal{E}\{\bullet\}$, from (26) and (28), we have the following equation

$$\lim_{t\to\infty} \hat{\sigma}^2(t|t) = \sigma^2. \quad (33)$$

The discrete representation of eq. (20) with $\hat{\sigma}(\bullet|\bullet)$ is as follows:

$$\begin{aligned} \hat{\mu}(t_{N+1}|t_{N+1}) &= \left(1 - \frac{P_0\Delta t}{P_0 t_N + \hat{\sigma}^2(t_N|t_N)}\right)\hat{\mu}(t_N|t_N) \\ &\quad + \frac{P_0\left[y(t_{N+1}) - y(t_N)\right]}{\left[P_0 t_N + \hat{\sigma}^2(t_N|t_N)\right] y(t_N)}. \quad (34) \end{aligned}$$

Equation (34) also shows the convergence of $\hat{\mu}(t_N|t_N)$.

Although equations (20) and (31) have different forms of observation data, they have the same form as an estimation mechanism. Therefore, we attempt to reconstruct them into a single estimation mechanism. Setting $\hat{\theta}(t|t) = \left[\hat{\mu}(t|t)\ \hat{\sigma}^2(t|t)\right]^T$, we formally reconstruct the estimation mechanisms as follows:

$$d\hat{\theta}(t|t) = K_\theta(t, \hat{\theta}(t|t))\left\{dz(t) - H(t)\hat{\theta}(t|t)dt\right\} \quad (35)$$

$$\hat{\theta}(0|0) := \left[\hat{\mu}_0\ \hat{\sigma}_0^2\right]^T \quad (0 < t < \infty)$$

where $K_\theta(t, \hat{\theta}(t|t)) = \mathrm{diag}\left\{\frac{P_0}{S(t)[P_0 t + \hat{\sigma}^2(t|t)]}, \frac{1}{t}\right\}$, $dz(t) = \left[dy(t)\ \ (d[\ln y(t)])^2\right]^T$ and $H(t) = \mathrm{diag}\{S(t), 1\}$.

The convergence of $\hat{\theta}(t|t)$ is guaranteed by equations (18) and (33).

4 Simulation Experiments

We used weekly data, because the data sampling interval is consistent. The data sampling interval dt is $7/365 \cong 0.0192$[year]. Since the S&P500 market index has easily accessible data, we use them. On March 4, 1957, the S&P 500 Index has begun operating in its current form.

It is widely reported that the expected return and volatility of the S&P500 are approximately 0.1 and 0.2, respectively. Consequently, in (1), we set the true values to $\mu = 0.1$ and $\sigma = 0.2$ in order to generate stock price data. We calculated the approximate values based on data provided in [4]. The initial values are set $S(0) = 44$[\$], which is the price of the S&P500 on Mar. 4, 1957, $\hat{\mu}_0 = 0.1$, $\hat{\sigma}(0|0) = 0.2$ and $P_0 = 0.1$. We set the simulation period to 70 years, corresponding to the historical timeframe of the S&P 500. The simulation results are illustrated from Figures 1 to 4.

Subsequently, we estimate the drift parameter μ and the volatility parameter σ using actual stock prices. These data were obtained from a publicly available website [4]. March 4, 1957, is set as $t = 0$.

The estimation results are illustrated from Figures 5 to 8. Since the Black-Scholes process does not fully model stock price movements, parameter estimates derived from real data fail to converge to specific values.

5 Conclusing Remarks

We derived two distinct estimation mechanisms for the drift and the volatility parameters in the Black-Scholes process. Furthermore, we successfully integrated the two estimation mechanisms into a single framework. As a result, the estimated parameter vector has been changed from $[\mu\ \sigma]^T$ to $[\mu\ \sigma^2]^T$, and the observation data has been changed from $dy(t)$ to $[dy(t)\ \ (d[\ln y(t)])^2]^T$. We consuquently derive a nonlinear estimation of the parameter $[\mu\ \sigma^2]^T$, instead of $[\mu\ \sigma]^T$, for the Black-Scholes Process.

References

[1] Stochastic volatility; https://en.wikipedia.org/wiki/Stochastic_volatility, Accessed December 1 2024.

[2] A. Ohsumi: *Introduction to Stochastic Systems*, Asakura Publ., Tokyo, 2002 (in Japanese).

[3] Andrew Hodges: Calculus of Variations, Lecture Notes for Trinity Term, 2016, https://www.maths.ox.ac.uk/system/files/attachments/CV2016_0.pdf, Accessed December 1, 2024.

[4] https://stooq.com/, Accessed December 1 2024.

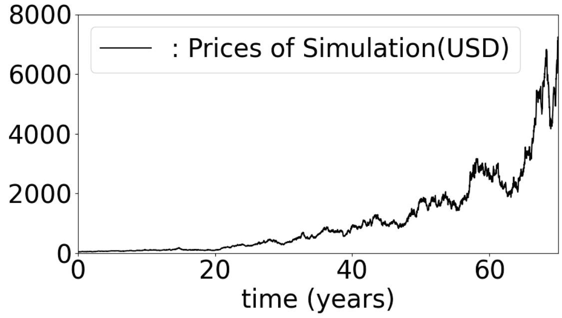

Figure 1: Generated Stock Prices.

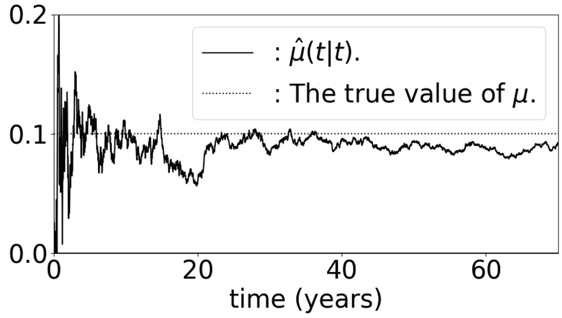

Figure 2: Estimation of μ.

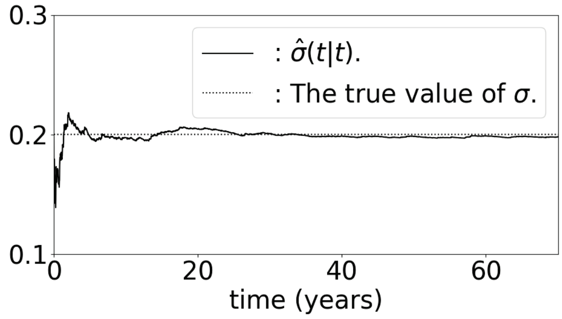

Figure 3: Estimation of σ.

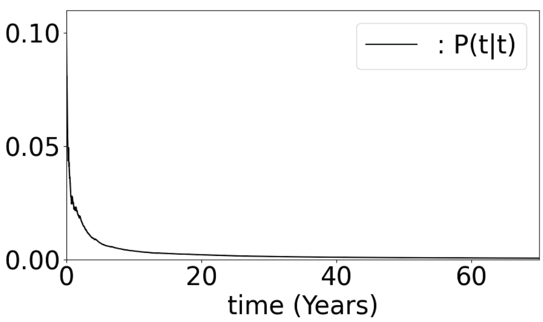

Figure 4: Estimation Error Variance of $\hat{\mu}$.

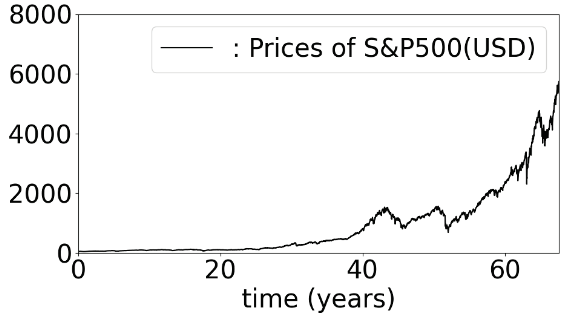

Figure 5: Stock Prices of S&P 500.

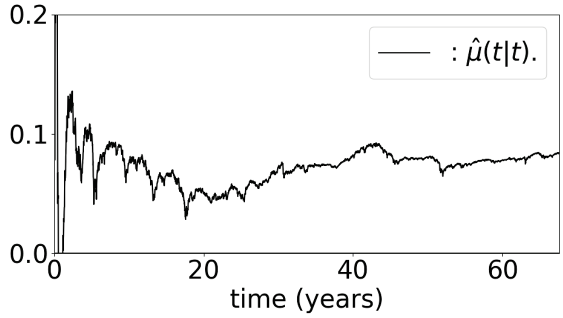

Figure 6: Estimation of μ (S&P500).

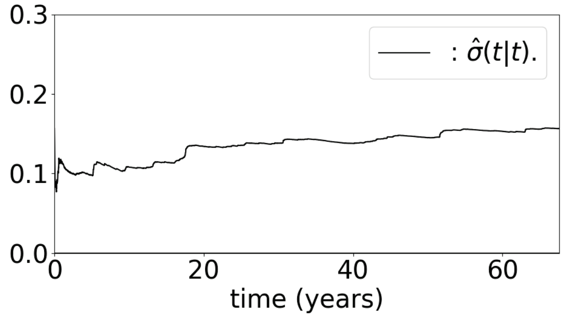

Figure 7: Estimation of σ (S&P500).

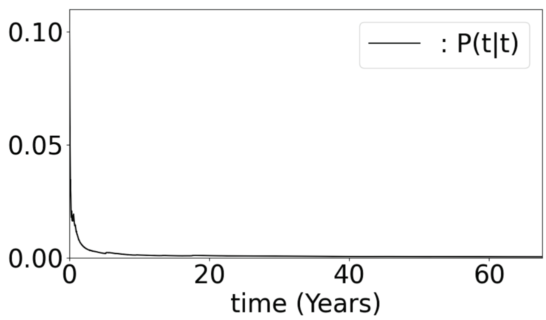

Figure 8: Estimation Error Variance of $\hat{\mu}$ (S&P500).