Estimating the expected return and volatility of Microsoft stock using adaptive MCMC in PyMC

Vili Niemelä

March 6, 2025

Abstract

The expected daily return and volatility of Microsoft one-year stock data were estimated with the NUTS sampler in the PyMC Python package. The stock price was assumed to follow geometric Brownian motion with constants μ and σ. The estimated expected daily return was 0.3% and the daily volatility 1.3%. The NUTS sampler in the package worked well in estimating the parameters.

1 Introduction

In this assignment, the expected return and volatility of Microsoft stock were estimated using the NUTS sampler, introduced in the article (Hoffman et al. (2014)). The stock price was assumed to follow a geometric Brownian motion (GBM), which is a stochastic differential equation. To solve the GBM and formulate the equation for the daily returns, Itô's lemma was used (for full theory of Itô's stochastic calculus see Karatzas and Shreve (1991)). For the implementation, the Python PyMC package was utilized. The data from which the expected return and volatility were estimated is the one-year Microsoft stock daily closing price data from which the daily log returns were calculated using the equation (5). The one year daily closing prices and log returns can be seen in Figure (1).

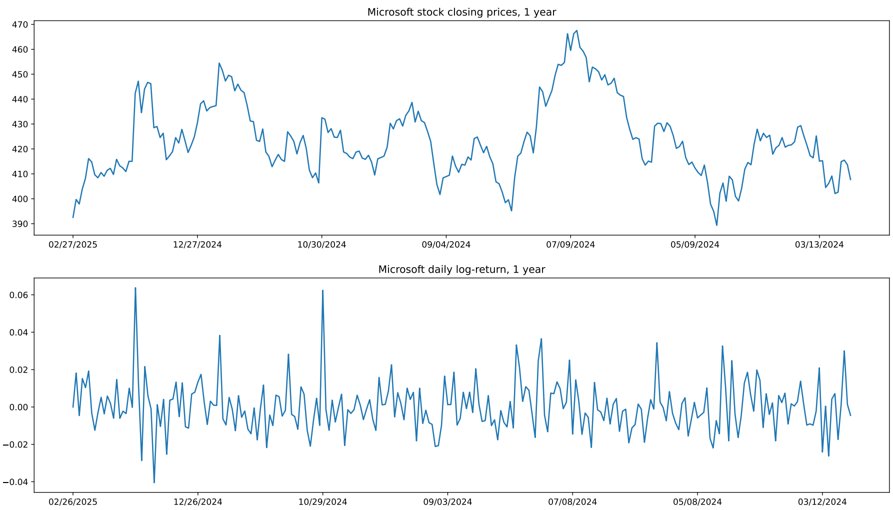

Figure 1: Microsoft stock price and daily log-return, one year.

2 Model

Geometric Brownian motion can be expressed as follows

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \tag{1}$$

where $\mu \in \mathbb{R}$ is expected return, $S_t \in \mathbb{R}$ stock price and $\sigma \in \mathbb{R}$ volatility of the stock. $B_t \in \mathbb{R}$ is a Brownian motion term. In this assignment the target parameters to be estimated are μ and σ.

GBM can be solved using the Itô's lemma. For solving the equation we denote $f(S_t) = \log(S_t)$ and take the Taylor expansion of df

$$df = f(s+ds) - f(s) = \frac{df}{ds}ds + \frac{1}{2}\frac{d^2 f}{ds^2}(ds)^2 + ... \tag{2}$$

By substituting $ds = dS_t = \mu S_t dt + \sigma S_t dB_t$ to (2) we get

$$\begin{aligned} df &= \frac{df}{ds}(\sigma S_t dB_t + \mu S_t dt) + \frac{1}{2}\frac{d^2 f}{ds^2}(\sigma^2 S_t^2 (dB_t)^2 + \mu S_t^2 (dt)^2 + 2\sigma S_t dB_t \mu S_t dt) + ... \\ &\approx f'(S_t)dS_t + \frac{1}{2}f''(S_t)(dS_t)^2. \end{aligned}$$

Notice that $(dt)^2, dtdB_t \approx 0$ and $(dB_t)^2 = dt$. When setting $(dt)^2, dtdB_t$ to zero we get

$$\begin{aligned} df &= \frac{1}{S_t}(\sigma S_t dB_t + \mu S_t dt) + \frac{1}{2}(-S_t^{-2})\sigma^2 S_t^2 dt \\ &= \sigma dB_t + (\mu - \frac{\sigma^2}{2})dt. \end{aligned}$$

By integrating both sides

$$\int_0^t d(\log(S_t)) = \int_0^t (\mu - \frac{\sigma^2}{2})dt + \int_0^t \sigma dB_t$$

we obtain

$$\log(S_t) - \log(S_0) = (\mu - \frac{\sigma^2}{2})t + \sigma(B_t - B_0).$$

Since B_t is Brownian motion, $B_0 = 0$ and we get

$$\log\left(\frac{S_t}{S_0}\right) = (\mu - \frac{\sigma^2}{2})t + \sigma B_t. \tag{3}$$

Finally from the equation (3) it can be noted that

$$\log\left(\frac{S_t}{S_0}\right) \sim \mathcal{N}((\mu - \frac{\sigma^2}{2})t, \sigma^2 t). \tag{4}$$

Now we can denote the returns at time point t_i as

$$Y_{t_i} = \log\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right) \sim \mathcal{N}((\mu - \frac{\sigma^2}{2})dt, \sigma^2 dt), \tag{5}$$

where $dt = t_i - t_{i-1}$ is a time step of the time discretization. For simpler denotations, we define $\tilde{\mu} = (\mu - \frac{\sigma^2}{2})dt$ and $\tilde{\sigma} = \sigma^2 dt$. By using $\tilde{\mu}$ and $\tilde{\sigma}$ we can express the likelihood as follows

$$p(\mathbf{Y}|\mu,\sigma) = \frac{1}{\sqrt{(2\pi)^n|\tilde{\sigma}^2 I|}} \exp\left(-\frac{1}{2}(\mathbf{Y}-\tilde{\mu})^T(\tilde{\sigma}^2 I)^{-1}(\mathbf{Y}-\tilde{\mu})\right), \tag{6}$$

where bolded symbols denotes vectors, for example $\mathbf{Y}$ denotes the return vector calculated from the stock data. In this assignment dt is set to one since the interval between the data points is one day and the one-day expected return and volatility are estimated.

The prior used for the parameter μ is the standard Gaussian with the mean zero and the standard deviation of 3, which can be expressed as

$$p(\mu) = \frac{1}{\sqrt{18\pi}} \exp\left(-\frac{\mu^2}{18}\right). \tag{7}$$

The prior for the parameter σ is the half-normal distribution with the scale parameter five. It can be expressed as follows

$$p(\sigma) = \begin{cases} \frac{\sqrt{2}}{5\sqrt{\pi}} \exp\left(-\frac{\sigma^2}{50}\right), & \sigma > 0 \\ 0, & \sigma \leq 0 \end{cases} \tag{8}$$

Now the posterior can be constructed from the priors and the likelihood

$$p(\mu,\sigma|\mathbf{Y}) \propto p(\mathbf{Y}|\mu,\sigma)p(\mu)p(\sigma). \tag{9}$$

3 Methods

To sample from the posterior given in the equation (9) the No-U-Turn (NUTS) sampler is used. NUTS sampler is an extension to the Hamiltonian Monte Carlo (HMC) sampler that transforms the problem of sampling from the target distribution into the simulation problem of Hamiltonian dynamics. In HMC user has to specify two parameters ϵ and L which correspond to a step size and a number of steps to run the Hamiltonian simulation. Since the poor choice of ϵ and L results in a decrease in the HMC efficiency, the NUTS sampler eliminates this problem by automatically tuning the parameters. (Hoffman et al. (2014))

4 Numerical Examples

In Figure (2) traces and marginal densities of the parameters μ and σ are presented. In Figure (3) autocorrelation functions of the parameters can be seen. The erdogic means of the traces are for μ 0.00031 and for σ 0.013. The corresponding effective sample sizes for the parameters are 1063 and 1136.

5 Conclusions

Based on estimates, the expected daily return for the Microsoft stock has been 0.3% and the daily volatility 1.3% during the last year. Based on Figures (2) and (3), the traces of the parameter estimates are converged and do not have a significant autocorrelation. Furthermore,

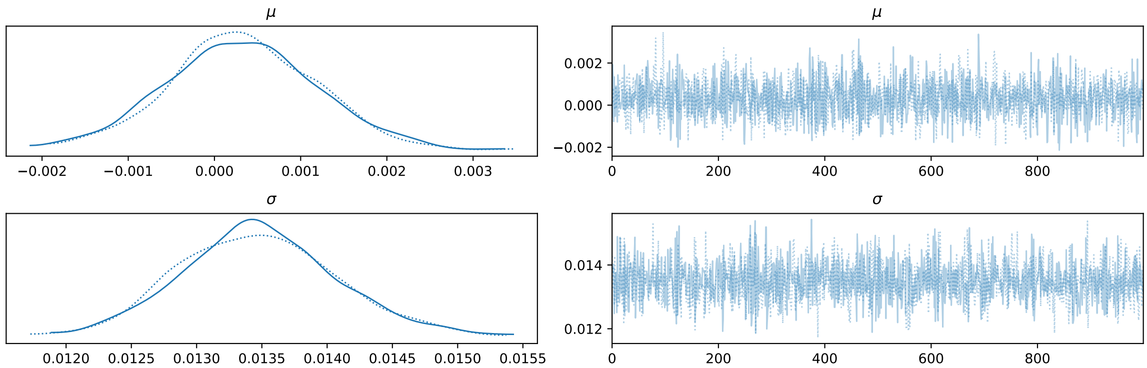

Figure 2: Traces and marginal densities of the parameters μ and σ.

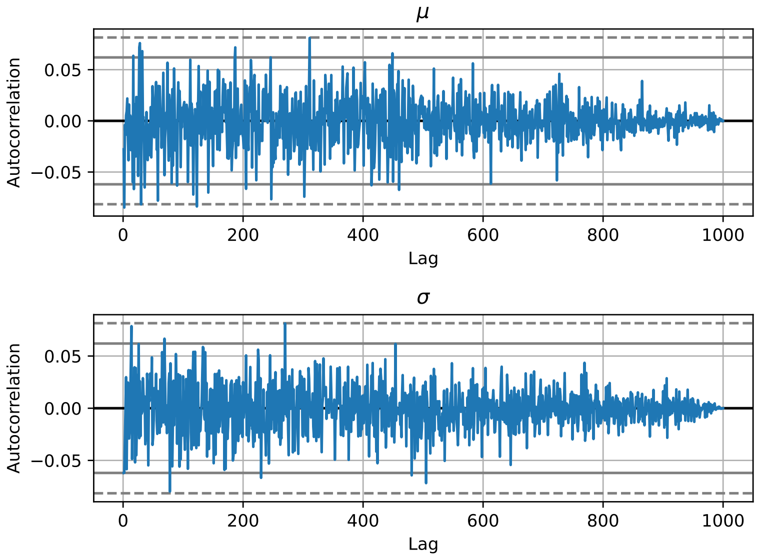

Figure 3: Autocorrelations of the μ and σ traces.

effective sample sizes are close to the chain lengths (1000). These results show that the NUTS sampler in the PyMC package worked well in estimating the expected return and volatility of the stock data. The package was also easy and fast to use thanks to good documentation. As the assumption of constant volatility is actually not correct, the topic for a possible future study relating to this study would be to estimate the time-varying volatility of Microsoft stock, for example, with the Kalman Filter algorithm.

References

Hoffman, M. D., Gelman, A., et al. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623.

Karatzas, I. and Shreve, S. (1991). *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media.