Bayesian Inference for Stochastic Volatility in Financial Markets: A Case Study on the S&P 500

Group 4: Zhiyuan Jia, Abdumalik Abdukayumov, Baran Sen

Department of Economics & Computer Science, Duke University

April 10, 2024

Overview

1. Introduction
2. Estimation: the theory
3. An Application: S&P 500 Returns
4. Conclusion and references

Motivation

- To predict the returns of a financial asset, we want to understand the evolution behind it
- Finance Model: there is an unobservable variable that drove the evolution: Volatility
- Our project: forecast daily S&P 500 closing returns

Motivation

- To predict the returns of a financial asset, we want to understand the evolution behind it
- Finance Model: there is an unobservable variable that drove the evolution: Volatility
- Our project: forecast daily S&P 500 closing returns
- Takeaway of presentation:
  1. A brief introduction to a canonical SV Model (State space model)
  2. Estimation technique in the model

Equation 1: observable variable

Formally,

- Start with a list of observable data (the return of assets at different time)

$$y_1, y_2, \ldots, y_t, \ldots, y_T.$$

[1] observable noise

Equation 1: observable variable

Formally,

- Start with a list of observable data (the return of assets at different time)

$$y_1, y_2, \dots, y_t, \dots, y_T.$$

- Finance Model:

$$y_t = \exp\left(\frac{h_t}{2}\right)\varepsilon_t \tag{1}$$

for h_t as the unobservable volatility, ε_t is i.i.d.[1] $N(0,1)$

[1] observable noise

Equation 2: volatility process

Formally,

- volatility is not constant but evolves over time

$$h_{t+1} = \mu + \phi\,(h_t - \mu) + \eta_t \tag{2}$$

- μ is the long time mean to which volatility reverts over time
- ϕ is the persistence parameter
- $\eta_t \sim N(0, \sigma_\eta^2)$ i.i.d, the dynamic noise
- initial condition $h_1 \sim N\left(\mu, \frac{\sigma^2}{1-\phi^2}\right)$

Estimation

$$y_t = \exp\left(\frac{h_t}{2}\right)\varepsilon_t$$

$$h_{t+1} = \mu + \phi\left(h_t - \mu\right) + \eta_t$$

$$\varepsilon_t \sim N(0,1), \eta_t \sim N(0,\sigma_\eta^2), h_1 \sim N\left(\mu, \frac{\sigma^2}{1-\phi^2}\right)$$

Given the model and data $y^T := \{y_1, y_2, ...y_T\}$, our project

Estimation

$$y_t = \exp\left(\frac{h_t}{2}\right)\varepsilon_t$$

$$h_{t+1} = \mu + \phi\,(h_t - \mu) + \eta_t$$

$$\varepsilon_t \sim N(0,1), \eta_t \sim N(0,\sigma_\eta^2), h_1 \sim N\left(\mu, \frac{\sigma^2}{1-\phi^2}\right)$$

Given the model and data $y^T := \{y_1, y_2, ...y_T\}$, our project

- estimate the parameters $\{\mu, \phi, \sigma_\eta^2\}$ in the spirit of Bayesian

Estimation

$$y_t = \exp\left(\frac{h_t}{2}\right)\varepsilon_t$$

$$h_{t+1} = \mu + \phi\,(h_t - \mu) + \eta_t$$

$$\varepsilon_t \sim N(0,1), \eta_t \sim N(0,\sigma_\eta^2), h_1 \sim N\left(\mu, \frac{\sigma^2}{1-\phi^2}\right)$$

Given the model and data $y^T := \{y_1, y_2, ...y_T\}$, our project

- estimate the parameters $\{\mu, \phi, \sigma_\eta^2\}$ in the spirit of Bayesian
- estimate the unobservable variable $h^T := \{h_1, h_2, ...h_T\}$

General idea

- Derive full conditionals and use MCMC (Gibbs Sampling)
  1. Derive $\sigma_\eta^2 \mid y, h, \phi, \mu$
  2. Derive $\mu \mid y, h, \phi, \sigma_\eta^2$
  3. Derive $\phi \mid y, h, \phi, \mu, \sigma_\eta^2$

General idea

- Derive full conditionals and use MCMC (Gibbs Sampling)
  1. Derive $\sigma_\eta^2 \mid y, h, \phi, \mu$
  2. Derive $\mu \mid y, h, \phi, \sigma_\eta^2$
  3. Derive $\phi \mid y, h, \phi, \mu, \sigma_\eta^2$
- Untrivial part: posterior of h^T
  - Alternative: derive full conditionals for each h_t seperately
  - Kim (1998) developed a more efficient method

Preliminary: Kalman Filter

- Consider a linear system with Gaussian noise

$$\begin{aligned} y_t &= x_t + v_t \\ x_t &= \theta x_{t-1} + w_t \end{aligned} \tag{3}$$

where $w_t \sim \mathcal{N}\left(0, \tau^2\right)$ and $v_t \sim \mathcal{N}\left(0, \sigma^2\right)$

Preliminary: Kalman Filter

- Consider a linear system with Gaussian noise

$$\begin{aligned} y_t &= x_t + v_t \\ x_t &= \theta x_{t-1} + w_t \end{aligned} \tag{3}$$

where $w_t \sim \mathcal{N}\left(0, \tau^2\right)$ and $v_t \sim \mathcal{N}\left(0, \sigma^2\right)$

- Given θ, τ^2, σ^2, and data $\{y_t\}_{t=1}^T$, there is a simulator for $\{x_t\}_{t=1}^T$: Kalman filtering and smoothing (omit detail here)

Generate all h^T at the same time

- Taking log over (1) to get a linear model[2]

$$y_t^* = h_t + \varepsilon_t^* \tag{4}$$

[2] where $y_t^* = log(y_t^2), \varepsilon_t^* = log(\varepsilon_t^2)$, and so $E[\varepsilon_t^*] = 1.27036$ and $Var(\varepsilon_t^*) = 4.93$.

Generate all h^T at the same time

- Taking log over (1) to get a linear model[2]

$$y_t^* = h_t + \varepsilon_t^* \tag{4}$$

- ε_t^* does not follow normal distribution anymore

[2] where $y_t^* = log(y_t^2), \varepsilon_t^* = log(\varepsilon_t^2)$, and so $E[\varepsilon_t^*] = 1.27036$ and $Var(\varepsilon_t^*) = 4.93$.

Generate all h^T at the same time

- Taking log over (1) to get a linear model[2]

$$y_t^* = h_t + \varepsilon_t^* \tag{4}$$

- ε_t^* does not follow normal distribution anymore
- Solution: using Gaussian mixture model
  - Approximate the exact likelihood
  - Why we want to do so?

[2] where $y_t^* = log(y_t^2), \varepsilon_t^* = log(\varepsilon_t^2)$, and so $E[\varepsilon_t^*] = 1.27036$ and $Var(\varepsilon_t^*) = 4.93$.

Approximation

Now, using the Gaussian mixture model to approximate the distribution of ε_t^*, namely,

$$f\left(\varepsilon_t^*\right)=\sum_{i=1}^{7} q_i f_{normal}\left(\varepsilon_t^* \mid m_i-1 \cdot 2704, v_i^2\right) \tag{5}$$

$s_t = i$	q_i	m_i	v_i^2
1	0.04395	2.77786	0.16735
2	0.24566	1.79518	0.34023
3	0.34001	0.61942	0.64009
4	0.25750	-1.08819	1.26261
5	0.10556	-3.97281	2.61369
6	0.00002	-8.56686	5.17950
7	0.00730	-10.12999	5.79596

Sampling from mixture

- Naturally we could use a two-step sampling to sample from the mixture of normal distribution. namely,

$$p(h_t|s_t, \{y_t\}_{t=1}^T, \mu, \phi, \sigma_\eta^2) * p(s_t|\{y_t\}_{t=1}^T, \mu, \phi, \sigma_\eta^2)$$

- **sample** $s_t|\{y_t\}_{t=1}^T, \phi, \mu, \sigma_\eta^2$

$$\Pr(s_t = i \mid y_t^*, h_t) \propto q_i f_{Normal}\left(y_t^* \mid h_t + m_i - 1 \cdot 2704, v_i^2\right)$$

and notice in full conditionals, the right-hand side is known.

Sampling from mixture (continue)

- **sample** $h_t | s_t, \{y_t\}_{t=1}^{T}, \phi, \mu, \sigma_\eta^2$
  - We want to solve $\{h_t\}_{t=1}^{T}$ in the system

$$
\begin{aligned}
y_t^* &= h_t + \varepsilon_t^* | s_t \\
h_t &= \mu + \phi\left(h_{t-1} - \mu\right) + \eta_{t-1}
\end{aligned}
$$

  - Normal noise, there is a standard method to deal with: Kalman filter (see detail in appendix)
  - Can use package/functions: KalmanLike(), KalmanRun(), and KalmanSmooth()

Full conditionals of σ_η^2, μ, ϕ

using the regression method discussed in Lecture 15, and equation (2) structure, for σ_η^2, given an inverse gamma prior,

$$\sigma_\eta^2 \sim IG(\frac{v_0}{2}, \frac{v_0\sigma_0^2}{2}) \tag{6}$$

$$\begin{gathered} \sigma_\eta^2 \mid y, h, \phi, \mu \sim IG\left\{\frac{n+v_0}{2}, \frac{v_0\sigma_0^2 + SSR}{2}\right\} \\ SSR = (h_1-\mu)^2\left(1-\phi^2\right) + \sum_{t=1}^{n-1}\left((h_{t+1}-\mu) - \phi\left(h_t-\mu\right)\right)^2 \end{gathered} \tag{7}$$

Full conditionals of σ_η^2, μ, ϕ

For μ, given a diffuse prior, [3]

$$
\begin{aligned}
&\mu \mid h, \phi, \sigma_\eta^2 \sim \mathcal{N}\left(\hat{\mu}, \sigma_\mu^2\right) \\
&,\hat{\mu} = \sigma_\mu^2 \left\{ \frac{(1-\phi^2)}{\sigma_\eta^2} h_1 + \frac{(1-\phi)}{\sigma_\eta^2} \sum_{t=1}^{n-1} (h_{t+1} - \phi h_t) \right\} \\
&\sigma_\mu^2 = \sigma_\eta^2 \left\{ (n-1)(1-\phi)^2 + \left(1-\phi^2\right) \right\}^{-1}.
\end{aligned} \tag{8}
$$

Similarly, we can derive a posterior for ϕ, given different prior.

[3] in practice, can use a slightly informative prior

MCMC

1. Initialize s, ϕ, σ_η^2 and μ.
2. Sample h from $h \mid y^*, s, \phi, \sigma_\eta^2, \mu$.
3. Sample s from $s \mid y^*, h$.
4. Update ϕ, σ_η^2, μ
5. Goto 2.

S&P 500

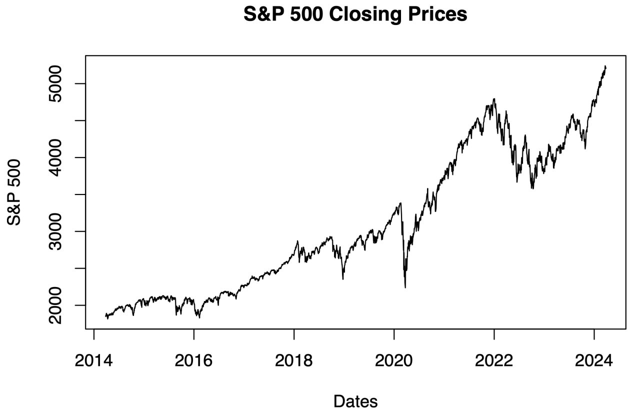

Figure: Closing Prices of SP 500

Log S&P 500 Returns

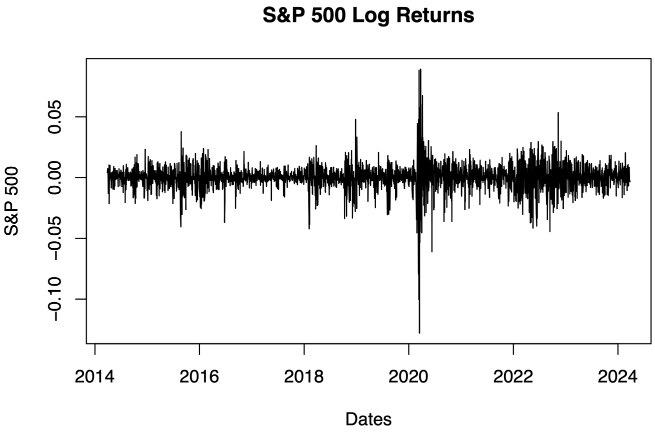

Figure: Log Returns of SP 500

Estimated Volatility

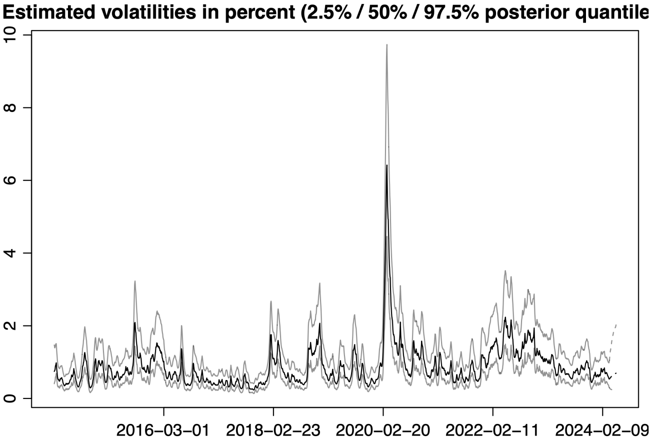

Figure: Volatility Plot

Trace Plots of Model Parameters

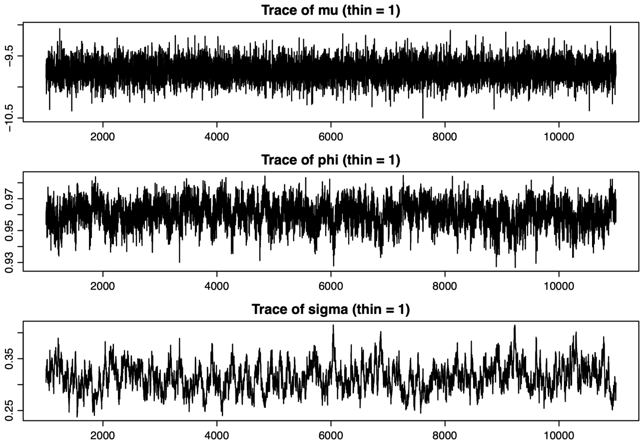

Figure: Model Parameters

Density Plots of Model Parameters

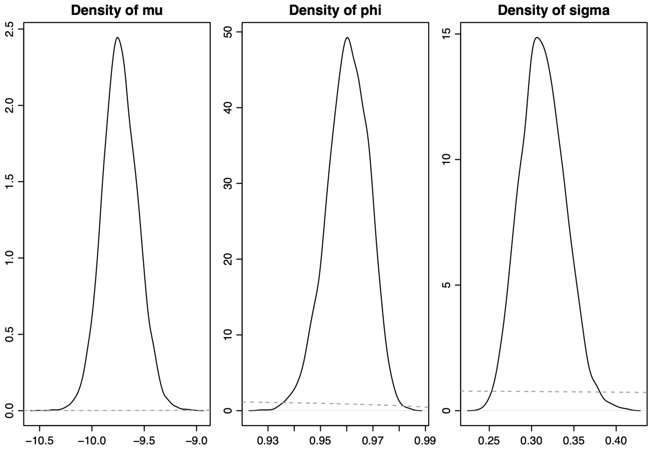

Figure: Model Parameter Densities

Forecasting

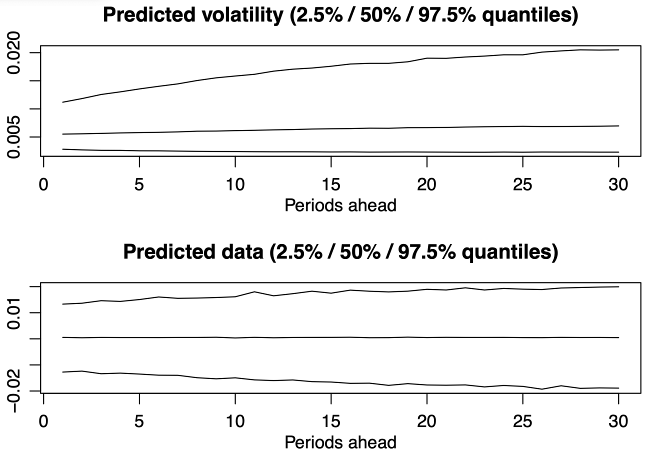

Figure: 30 - Day Forecast

Takeaways

- SV/ State-space Model could be used to model financial data, which consider the randomness in the volatility
- Estimation of the volatility: using Gaussian Mixture model to approximate $log\chi^2$
- Limitations : incorporate unexpected shock in the economy (e.g. Covid)
- Future work: we can build the model with more complex specifications (SVt, include AR terms)

References

- Greenberg, E. (2012). Introduction to Bayesian econometrics. Cambridge University Press.
- Hosszejni D, Kastner G (2021). "Modeling Univariate and Multivariate Stochastic Volatility in R with stochvol and factorstochvol." *Journal of Statistical Software*, **100**(12), 1–34. doi:10.18637/jss.v100.i12 .
- Kim, S., Shephard, N., Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. The review of economic studies, 65(3), 361-393.

Appendix

Rudolf Kalman developed a method to do so:

1. Initialize λ_0^0, P_0^0
2. Calculate $\lambda_1^1, P_1^1, \lambda_2^2, P_2^2, \lambda_3^3, P_3^3 \ldots, \lambda_T^T, P_T^T$ using the following update:
   1. update

$$\lambda_t^{t-1} = \theta \lambda_{t-1}^{t-1}$$
$$P_t^{t-1} = \theta^2 P_{t-1}^{t-1} + \tau^2$$

   2. calculate $K_t = \frac{P_t^{t-1}}{P_t^{t-1} + \sigma^2}$ and update

$$\lambda_t^t = \lambda_t^{t-1} + K_t \left(y_t - \lambda_t^{t-1} \right)$$
$$P_t^t = (I - K_t) P_t^{t-1}$$

3. sample x_t from $\mathcal{N}(\lambda_t^t, P_t^t)$

Thanks for your attention