Unsupervised learning of asymmetric high-order autoregressive stochastic volatility model

Ivan Gorynin, Emmanuel Monfrini and Wojciech Pieczynski

SAMOVAR, Telecom SudParis, CNRS, Université Paris-Saclay, 9, rue Charles Fourier, Evry, France.

Context

- (X, Y) : asymmetric p-order autoregressive stochastic volatility model (A-ARSV(p), [1]) : (L) hidden scalar, (Y) observed:

$$L_{n+1} = \sum_{j=1}^{p} \varphi_j L_{n+1-j} + \psi W_n + \sigma V_{n+1};$$

$$Y_n = W_n \exp\left(\frac{L_n}{2} + \frac{\mu}{2}\right).$$

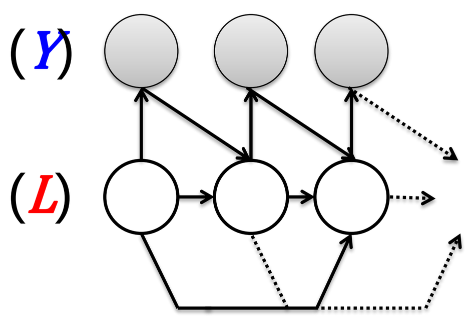

Fig. 1 Oriented dependency graph of A-ARSV(2).

Y is known as a log-return process, L is log-volatility. ψ, usually assumed negative, parameterizes the asymmetric volatility, according to which the returns and conditional volatility are negatively correlated. (V_n, W_n) are independent $\mathcal{N}(\mathbf{0}, \mathbf{I}_2)$-distributed.

- Contribution: present an approach for estimating the order p of A-ARSV model from an observed sequence $Y_{1..N}$, and related parameters $\theta = (\varphi_1, \ldots, \varphi_p, \psi, \sigma, \mu)$.

Approach proposed

- Consider an augmented state vector X and write A-ARSV(p) as a hidden Markov model $X_{n+1} = AX_n + BU_{n+1}$, $Y_n = h(X_n)$,

$$X_n = \begin{bmatrix} L_{n-p+1} \\ L_{n-p+2} \\ \ldots \\ L_n \\ W_n \end{bmatrix}, A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \varphi_p & \varphi_{p-1} & \varphi_{p-2} & \ldots & \varphi_1 & \psi \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ \sigma & 0 \\ 0 & 1 \end{bmatrix}, U_n = \begin{bmatrix} V_n \\ W_n \end{bmatrix}.$$

- Approximate the filtering $p_\theta(\mathbf{x}_n|\mathbf{y}_{1..n})$ and predictive $p_\theta(\mathbf{x}_{n+1}|\mathbf{y}_{1..n})$ distributions by normal densities. The parameters of these densities can be computed from θ and $Y_{1..N}$ in the Bayesian estimation framework by using a one-dimensional Gauss-Hermite quadrature rule (see, e.g., [2]) recursively. Indeed, Y_n depends only on 2 components in X_n, thus

$$\log p_\theta(y_{1\ldots N}) = \log p_\theta(y_1) + \sum_{n=1}^{N-1} \log p_\theta(y_{n+1}|y_{1..n}) = \log \int p_\theta(y_1|x_1) p_\theta(x_1) dx_1 + \sum_{n=1}^{N-1} \log \int p_\theta(y_{n+1}|x_{n+1}) p_\theta(x_{n+1}|y_{1..n}) dx_{n+1}$$

can be approximated by using a one-dimensional Gauss-Hermite quadrature too.

- Parameter estimate θ^* is obtained by maximizing $p_\theta(\mathbf{y}_{1..N})$ w.r.t. θ by numerical optimization techniques. Finally, for each candidate order p, we compute $\theta^*(p)$ and the Bayesian information criterion (BIC, [3]) defined as BIC(p) = $-2\log p(\mathbf{y}_{1..N}|\theta^*(p)) + (p+3)\log N$ in the case of the A-ARSV(p) model. The best-suited order p is chosen as that which minimizes BIC(p).

Results and discussion

- Evidence from the BIC plots of ARSV(p) and A-ARSV(p) models for U.S. stock indices suggests that $p=2$ would be sufficient in most cases. It was frequent to observe BIC(A-ARSV(2))<BIC(ARSV(2))< BIC(ARSV(1))< BIC(A-ARSV(1)) – see e.g. Fig. 2. Considering a higher order of autoregression may allow to better highlight the asymmetric volatility phenomenon.
- Our method can be generalized to take into account any non-linear/non-Gaussian conditional distribution of Y in the case where the hidden process is scalar autoregressive. The Gauss-Hermite quadrature rule is particularly efficient in the one-dimensional case, what makes our estimation procedure preferable to the simulation-based alternatives.

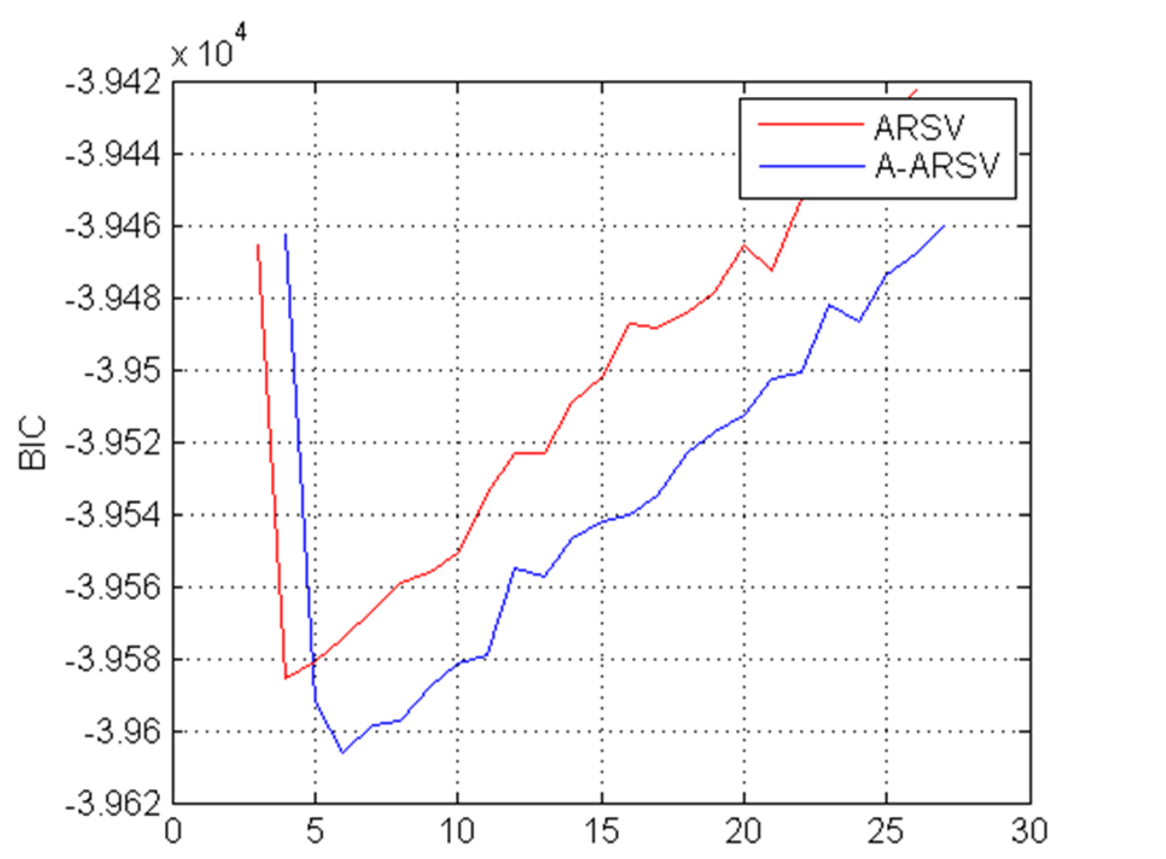

Fig. 2 BIC values of the ARSV(p), A-ARSV(p) models estimated from the EEM index (lower is better).

References

[1] M. Centedo and R. Salido, "Estimation of asymmetric stochastic volatility models for stock-exchange index returns," *International advances in economic research*, vol. 15, no. 3, pp. 70–87, 2003.
[2] H. Salzer, H. Zucker and R. Capuano, *Table of the zeros and weight factors of the first twenty Hermite polynomials*, US Government Printing Office, 1952.
[3] G. Schwarz, "Estimating the dimension of a model", *Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.

Contact Ivan.Gorynin@telecom-sudparis.eu - http://citi.telecom-sudparis.eu/fr/
Emmanuel.Monfrini@telecom-sudparis.eu