Physica A 461 (2016) 788–794

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Extracting volatility signal using maximum a posteriori estimation

David Neto [1]

University of Geneva, Switzerland

HIGHLIGHTS

- We consider volatility estimation problem as a signal processing problem.
- A Laplace process is used as prior for the log-volatility to capture sharp jumps.
- Our empirical results show that our method outperforms RV measure to capture extremes.

ARTICLE INFO

Article history:
Received 22 November 2015
Available online 14 June 2016

Keywords:
Hidden Markov model
Volatility
Maximum a posteriori estimator
IDM algorithm
Laplace process
Realized volatility

ABSTRACT

This paper outlines a methodology to estimate a denoised volatility signal for foreign exchange rates using a hidden Markov model (HMM). For this purpose a maximum a posteriori (MAP) estimation is performed. A double exponential prior is used for the state variable (the log-volatility) in order to allow sharp jumps in realizations and then log-returns marginal distributions with heavy tails. We consider two routes to choose the regularization and we compare our MAP estimate to realized volatility measure for three exchange rates.

1. Introduction

Since volatility is naturally unobservable, its estimation represents a serious challenge for investors. As pointed out by Andersen et al. [1], direct volatility indicators, such as ex post squared, are contaminated by noise for which the variance can be very large relative to that of the signal [2]. On the other hand, the volatility estimation issue has mainly been addressed using parametric approaches like GARCH family models, stochastic volatility models, which are obviously exposed to the risk of misspecification. It is why this last decade has seen a growing interest for the realized volatility measure introduced by Andersen et al. [1], which only requires high-frequency data aggregation.

In this paper we propose an alternative in the spirit of the contribution of Neto, Sardy and Tseng [3], and Neto and Sardy [4]. However, a different direction is taken in this paper by relaxing the Taylor' [5] stochastic volatility parametrization in the hidden Markov model (henceforth HMM). The volatility is discussed here as in Ref. [6] who propose a non-stationary nonparametric volatility model where the volatility is controlled by a smooth nonparametric function depending on an observable volatility proxy (like the VIX index). However, in contrast with these authors, the volatility is considered in our context as a fully (non-stationary) hidden signal, which also moves smoothly, but occasionally exhibits some jumps allowed by a Double exponential prior over the log-volatility. Hence, the signal appears as locally stationary [7].

E-mail address: david.neto@unige.ch.

[1] The author was a Post-Doctoral Research Assistant at the Department of Economics, University of Geneva, when an early version of this paper was written.

http://dx.doi.org/10.1016/j.physa.2016.05.065
0378-4371/

Regarding the estimation issue, it is considered as a signal processing problem. Indeed, commonly used in image processing (for reconstruction, segmentation, or 3D vision), hidden Markov models and their related estimators (the maximum a posteriori (MAP) or the Markov random field estimators), have not had much echo in empirical finance, even so these techniques have interesting arguments in their favor in terms of implementation ease. Yet, one of the main advantages of these estimators probably resides on the fact that they use deterministic optimization algorithms instead of simulated estimation methods. However, the tricky aspect of these estimation procedures is that they requires regularization (*i.e.* a penalty for complexity). For exchange rates volatility issue, we adopted a total variation regularization which authorizes abrupt jumps but makes the MAP estimation quite challenging since the a posteriori likelihood function is no longer differentiable. Additionally, the choice of the regularization parameter can also be a sensitive issue. Two routes were performed and compared for the exchange rates: the universal thresholding approach and a cross validation procedure.

The outline of the paper is as follows. The HMM is presented in Section 2. Section 3 addresses the estimation issues. Section 4 is devoted to our empirical application to exchange rates. We compare our signal obtained from our model with the realized volatility measure. In particular we focus on the ability of both estimates to capture heavy tails in the marginal distributions. The Section 5 concludes.

2. Hidden Markov model for market volatility

Markov models are useful as prior models for state variables (*i.e.* unobservable variables) which have to be inferred from measurements. For the exchange rate volatility estimation issue, log returns, given by $R_t = \ln(S_t/S_{t-1})$, where S_t denotes the exchange spot rate on trading date t, are the measurements and the volatility, denoted σ_t, is the state variable which can also be considered as a time-varying parameter. The inference problem in which the posterior distribution of the possible state $\mathbf{X}$, given the corresponding set of observations, is given by the Bayes' formula:

$$\mathbb{P}(\mathbf{X} = \mathbf{x}|\mathbf{R} = \mathbf{r}) \propto \mathbb{P}(\mathbf{R} = \mathbf{r}|\mathbf{X} = \mathbf{x})\mathbb{P}(\mathbf{X} = \mathbf{x}), \tag{1}$$

where $\mathbf{R} = (R_1, \ldots, R_t, \ldots, R_T)'$, $\mathbf{X}$ are the log volatility, *i.e.* $X_t = \log(\sigma_t)$ and $\mathbf{X} = (X_1, \ldots, X_T)$, $\mathbb{P}(\mathbf{X} = \mathbf{x})$ is the prior over the log-volatility, $\mathbb{P}(\mathbf{R} = \mathbf{r}|\mathbf{X} = \mathbf{x})$ is the likelihood of the observations which measure the quality of the measurements. The hidden Markov model consists of describing the latter prior itself as a first-order Markov chain for which the joint density is the product of the conditional densities: $\mathbb{P}(\mathbf{X} = \mathbf{x}) = \prod_{t=2}^{T} \mathbb{P}(X_t = x_t|X_{t-1} = x_{t-1})$. Let us specify the prior conditional distribution as a centered Laplace process:

$$X_t|X_{t-1} \sim \mathscr{L}(X_{t-1}, \kappa), \tag{2}$$

with the scale parameter $\kappa \in \mathbb{R}_+^*$. The conditional distribution of the log returns is assumed to be Gaussian:

$$R_t|X_t \sim \mathscr{N}(0, \exp(2X_t)). \tag{3}$$

Considering such a prior (2) for the log volatilities relates our HMM to the variance Gamma (VG) model commonly used in finance. Indeed, this model has had attracted a particular attention since it was introduced in its complete formulation by Madan and Seneta [8]. The idea of the VG model is based on stochastic time changes defined in the early paper of Clark (1973). Hence, the time change is supposed to reflect the information arrival flow on the market which conditions the market activity (see also Ref. [13]). Technically it consists of using subordinated processes. For instance, the VG model describes the stock returns as a diffusion process where the Brownian motion is evaluated at random time distributed as a Gamma process. Such a process is also called a Laplace motion [9] and it is a pure jump compound process.[2] Hence, the prior (2) allows sharp changes in the realizations of the log volatilities. For small values of κ, the hidden process will exhibit some frequent jumps, whilst the trajectory will be smoother for large values of κ. From (1), the negative posterior log-likelihood of model (2)–(3) is given by:

$$\text{npl}(r_t, X_t, \theta) = \sum_{t=1}^{T} X_t + 0.5r_t^2 \exp(-2X_t) + \kappa \sum_{t=2}^{T} |\Delta X_t|, \tag{4}$$

where $\theta' = (\mathbf{X}', \kappa)$ and where κ is the tuning parameter which controls the smoothing of the signal. Following Han and Zhang [6], we also assume a smooth underlying signal for the volatility, however, in contrast with these authors, it is free of covariates.

3. Estimation method: the maximum a posteriori

The MAP method consists of estimating the most probable sequence $\mathbf{X}$ as the mode of the posterior distribution. In other words, the estimation of $\mathbf{X}$ merely consists of minimizing (4). By not considering the hyperparameter κ as a direct parameter

[2] Laplace motion can be written as a compound Poisson process with independent and random jumps.