Option-based Maximum Likelihood Estimation in Stochastic Volatility Models

Stefan Häggesten, Jimmy Olsson

Center of Mathematical Sciences, Lund University, Box 118 SE-22100 Lund, SWEDEN
e-mail: jimmy@maths.lth.se, stefan.hagnesten@gmail.com

Abstract: In this note we apply the particle-based iterated filtering algorithm proposed by Ionides et al. (2009) to the problem of calibration of stochastic volatility models. We assume that the underlying asset follows the Heston stochastic volatility dynamics and consider a hidden Markov model formulation of the observed option prices where the volatility of the underlying asset is treated as a latent signal. In this setting, robust approximations of the maximum likelihood estimator (MLE) are obtained by introducing a time varying parameter process following a random walk dynamics with decreasing size of the increments. A strength of this approach is that a consistent estimate of the MLE can be formed without having access to a closed-form expression of the Markovian transition density of the latent process. The technique is demonstrated on simulated data.

Keywords and phrases: Heston model, iterated filtering, maximum likelihood, option pricing, particle filter, sequential Monte Carlo methods.

1. Introduction

Ever since the historic paper by Bachelier (1900), the field of mathematical finance has developed extensively. The theory of derivative pricing is one of its main subareas, presenting a vast amount of non-trivial problems to be solved. Whatever application, efficient and precise calibration of underlying model parameters is crucial, and in this note we thus explore new techniques for estimating parameters in financial hidden Markov models (HMM) (see e.g. Cappé et al., 2005, for an extensive treatment) via observed option prices.

In this extended abstract we let the dynamics of the underlying asset process \( S_t \) be governed by the Heston model (Heston, 1993). More specifically, the dynamics of the asset under the risk neutral measure \( Q \) is described by the system

\[
\begin{align*}
    dS_t &= rS_t dt + \sqrt{V_t} S_t dW^S_t, \\
    dV_t &= \kappa (\xi - V_t) dt + \sigma_V \sqrt{V_t} dW^V_t,
\end{align*}
\]

of stochastic differential equations, where \( \kappa > 0 \) is the mean reversion rate, \( \xi > 0 \) is the mean reversion level, and \( (W^S_t)_{t \geq 0} \) and \( (W^V_t)_{t \geq 0} \) are standard Brownian motions such that \( \text{Cov}(dW^S_t, dW^V_t) = \rho dt \), with \(-1 \leq \rho \leq 1\). In addition, \( \sigma_V > 0 \) denotes the volatility of the volatility \( V_t \). We denote by \( \theta \equiv (\kappa, \xi, \sigma_V, \rho) \) the full parameter vector and by \( C^\theta_t(T, K, s, v) \) the theoretical price of a European call option at time \( t \), written on the underlying asset value \( s \) and volatility \( v \), with strike level \( K \) and time \( T \) to maturity. With this notation we assume that the sequence \( (C^\theta_k)_{k \in \mathbb{N}} \) of market prices consists of noisy observations of theoretical prices at a grid \( (t_k)_{k \in \mathbb{N}} \) of discrete time points according to

\[
C^\epsilon_k = C^\theta_{t_k}(T, K, S_k, V_k) + \sigma_V \epsilon_k,
\]

where \( (\epsilon_k)_{k \in \mathbb{N}} \) is a sequence of independent standard Gaussian variables and \( S_k \equiv S_{t_k} \) and \( V_k \equiv V_{t_k} \) for \( k \in \mathbb{N} \). The additional noise term enables us to capture the bid-ask...
uncertainty of the price process. The conditional distribution of \( C^*_k \) given \( (V_k) \) (and \( S_k \)) is referred to as the likelihood of the state and will be denoted by \( G_\theta \). In this setting, assuming that the underlying asset price is known while the volatility process is unknown at any time instant, the bivariate process \((V_k, C^*_k)_{k \in \mathbb{N}}\) form a time inhomogeneous HMM, and we denote by \((Q_{\theta,k})_{k \in \mathbb{N}}\) the Markov transition kernels of the volatility process conditional on the asset price process \( S \), i.e., \( Q_{\theta,k}(V_k, A) = \mathbb{P}_\theta(V_{k+1} \in A|V_k, \mathcal{F}_{k+1}) \) with \( (\mathcal{F}_t)_{t \geq 0} \) being the filtration generated by \( S \). Unfortunately, the transition kernels lack closed-form expressions. Our basic problem is to form a consistent estimate of \( \theta \) with \( \mathbb{E}_\theta[n] \) being the filtration generated by \( S \).

2. Parameter estimation in the Heston model using iterated filtering

Each iteration of the algorithm under consideration comprises two main operations: Firstly, an approximation of the likelihood gradient (or score function), evaluated at a given crude estimate of the maximum likelihood estimator (MLE), is formed by solving, using a particle filter, an optimal filtering problem for the extended HMM obtained when imposing a random walk dynamics on the unknown parameter. Secondly, an improved estimate of the MLE is obtained by means of a standard stochastic approximation scheme. It has been successfully applied to state-of-the-art HMM inference problems within biology and environmental science (see Ionides et al., 2006; King et al., 2008; He et al., 2009; Bretó et al., 2009), but has not, at far as known to the authors, been used for inference in financial models.

In order to describe the method more specifically, let \( \ell_n(\theta) \) the log-likelihood function of the observed option prices \((C^*_k)_{k=0}^n\); in order to find a zero of the score function \( \nabla_\theta \ell_n(\theta) \) a first obvious approach would be to apply the iterative Robbins-Monro scheme

\[
\theta_{\ell+1} = \theta_\ell + \gamma_{\ell+1} \nabla_\theta \ell_n(\theta_\ell), \quad \ell = 0, 1, \ldots
\]

for some crude initial estimate \( \theta_0 \) on the parameter. However, since the score is not available on closed-form we are forced to replace \( \nabla_\theta \ell_n(\theta_\ell) \) by some suitable approximation. Following Ionides et al. (2006), such a pointwise approximation may be obtained by imposing, in a Bayesian manner, a random dynamics on the parameter as well. More precisely, we consider the extended HMM \( (V_k, \theta_k, C^*_k)_{k \in \mathbb{N}} \) with signal transition kernel, observation transition kernel, and initial distribution given by, respectively,

\[
Q^*_{\sigma,k}([v, \theta], \cdot) \equiv Q_{\theta,k}(v, \cdot) \otimes P_{\theta}(\cdot), \\
G^*([v, \theta], \cdot) \equiv G_\theta(v, \cdot), \\
\chi^*_{\sigma,\theta} \equiv \chi \otimes P_{\sigma}(\cdot),
\]

with \( \otimes \) denoting product measures and \( \overline{\theta} \) being a given parameter value. In the equations above \( P_{\sigma} \) is a Markovian transition kernel, indexed by some parameter \( \sigma \), for the parameter process \( \theta_k \) \( k \in \mathbb{N} \) and \( \varsigma(\cdot) \) is a function of \( \sigma \); in this way, \( \sigma \) and the initial parameter vector \( \theta \) determines the full dynamics of the extended model. From the definition of \( Q^*_{\sigma,k} \) it is clear that the parameter and volatility processes evolve independently. In this note we let \( P_{\sigma} \) describe a simple random walk where all components of the parameter vector move independently of each other with Gaussian-distributed increments with mean zero and standard deviation proportional to \( \sigma \). Assume that \( \varsigma \) is such that \( \varsigma(\sigma)/\sigma \to 0 \) as \( \sigma \to 0 \); then, it is clear that letting \( \sigma \to 0 \) reproduces the original HMM.
shown by Ionides et al. (2006), under rather mild assumptions,

\[
\lim_{\sigma \to 0} \frac{1}{n} \sum_{m=1}^{n} \nabla_{\sigma, \beta} \left[ \theta_m | C_{0:m-1}^* \right] \left( \mathbb{E}_{\sigma, \beta} [\theta_m | C_{0:m}^*] - \mathbb{E}_{\sigma, \beta} [\theta_m | C_{0:m-1}^*] \right) = \nabla_{\theta} \ell_n (\theta) \quad (3)
\]

where \( \mathbb{E}_{\sigma, \beta} \) and \( \nabla_{\sigma, \beta} \) denote expectation and variance in the extended model. The limit (3) transfers the problem of producing a pointwise approximation of the score function to the problem of solving a nonlinear optimal filtering (and prediction) problem in extended model (2), i.e., to compute the posterior distribution of \( \theta_m \) (or \( \theta_{m+1} \) in the case of prediction) conditional on the observed data \( C_{0:m}^* \) as well as the information \( \mathcal{F}_{t_m}^S \). These posterior distributions lack however closed-form expressions but may be approximated using sequential Monte Carlo (SMC) methods. SMC methods refer to a class of algorithms approximating a sequence of probability distributions, defined on a sequence of probability spaces, by updating recursively a set of random particles with associated nonnegative importance weights. In our setting, having at hand a such a weighted particle sample \( \{(v^i_m, \tilde{\theta}_m^i, \omega^i_m)\}_{i=1}^N \), where each \( v^i_m \) and \( \tilde{\theta}_m^i \) are random draws in the state space of the volatility process and in the parameter space, respectively, approximating the filter posterior distribution of \( (V_m, \theta_m) \) conditional on \( C_{0:m}^* \) and \( \mathcal{F}_{t_m} \), an updated sample \( \{(v^i_{m+1}, \tilde{\theta}_m^{i+1}, \omega^i_{m+1})\}_{i=1}^N \) approximating the corresponding posterior distribution at time \( m + 1 \) is obtained by (i) transforming, by resampling the particles multinomially according to the normalized importance weights, the original sample into a uniformly weighted sample \( \{(v^i_m, \tilde{\theta}_m^i, 1)\}_{i=1}^N \) and then (ii) simulating, conditionally independently, \( N \) new particles according to \( (v^i_{m+1}, \tilde{\theta}_m^{i+1}) \sim Q_{\sigma, k}(\{v^i_m, \tilde{\theta}_m^i\}, \cdot) \). The uniform sample \( \{(v^i_{m+1}, \tilde{\theta}_m^{i+1}, 1)\}_{i=1}^N \) now targets the predictive distribution of \( (V_{m+1}, \theta_{m+1}) \) given \( C_{0:m}^* \) and \( \mathcal{F}_{t_m} \). Finally, (iii) each particle \( (v^i_{m+1}, \tilde{\theta}_m^{i+1}) \) is assigned the importance weight \( \omega^i_{m+1} \defeq g^i(\{v^i_{m+1}, \tilde{\theta}_m^{i+1}, C_{m+1}^*\}) \). Now, standing at iteration \( \ell \) of (1), we run the particle filter under the extended model dynamics \( (\sigma_{\ell}, \theta_{\ell}) \) for obtaining samples \( \{(v^i_{m,\ell}, \tilde{\theta}_m^{i,\ell}, \omega_{m,\ell})\}_{i=1}^N \), \( 0 \leq m \leq n \), and approximate \( \nabla_{\theta} \ell_n (\theta_{\ell}) \) by the quantity

\[
\hat{s}_{\ell} \defeq \frac{N_{\ell} - 1}{N_{\ell}} \sum_{m=1}^{n} \sum_{j=1}^{N_{\ell}} \left( \tilde{\theta}_{m,\ell} - \hat{\theta}_{m-1,\ell} \right) \left( \tilde{\theta}_{m,\ell} - \hat{\theta}_{m-1,\ell} \right)^T
\]

and letting \( \theta_{\ell+1} = \theta_{\ell} + \gamma_{\ell} \hat{s}_{\ell} \). Ionides et al. (2009, Theorem 3) provide criteria on how to decrease the random walk standard deviation \( \sigma_{\ell} \) and the step size \( \gamma_{\ell} \) as well as how to increase the number of particles \( N_{\ell} \) with the iteration index \( \ell \) in order to obtain an algorithmic output \( (\theta_{\ell})_{\ell \in \mathbb{N}} \) that converges to the exact MLE with probability one. Note that the algorithm above presupposes only the possibility of simulating transitions according to \( Q_{\theta, k} \) for all \( k \), i.e. we do not need closed-form expression of the transition densities. If the asset price is sampled at high frequency, transitions can simulated with relatively high precision using the Euler scheme.

3. Simulation Study

In order to study the efficiency of the method, we simulated trajectories of the bivariate process \((S, V)\) over \( n = 100 \) time steps using the Euler scheme. Here the time steps correspond to daily measurements. The parameter values chosen (and suggested by Zhang and Shu, 2003) were \( \kappa = 2.75, \xi = 0.035, \sigma_V = 0.425, \) and \( \rho = -0.4644, \) and the HMM was initialized at values \((S_0, V_0) = (1000, 0.1)\). Moreover, 100 daily option prices—one at each time step—were simulated over 5 regularly spaced time to maturities, \( T \in [0.25, 1], \) and 20 strike levels with \( K_t \in [0.85 S_t, 1.15 S_t] \). When propagating the particles for the process \((S, V)\), an Euler discretization with 10 times higher resolution was used, yielding a dynamics of the particles that was close to that of the true system. Initial values of
the parameters were set to $\kappa_0 = 3.0$, $\xi_0 = 0.08$, $\sigma_{\epsilon,0} = 0.6$, and $\rho_0 = -0.7$. As evident from Figure 1., all parameters converge successfully to the true values, except for $\kappa$ which is not as informative on this scale. In the present formulation, the performance of the algorithm depends rather strongly on the size of the measurement noise standard deviation $\sigma_\epsilon$ (reflecting the bid-ask uncertainty). Modeling the bid-ask spread as 1, the size of the measurement noise deviation was set to $\sigma_\epsilon = 0.2$. This is motivated with that we expect at least 95% of the true option prices to be located within the spreads. A smaller size of $\sigma_\epsilon$ would decrease the performance of the algorithm due to the fact that the particles are propagated blindly, i.e. without taking into account any information about the subsequent option price, according to $Q_{k,\sigma}^*$, and with a small observation noise many particles are expected to be proposed in regions of small posterior probability (as measured by the likelihood $G^*$). This leads to degeneracy of the particle sample. A way to cope with this problem is to, within the framework of the auxiliary particle filter proposed by Pitt and Shephard (1999), resample the particles according to probabilities proportional to adjusted weights $(\omega_i^m \lambda_m^i)^{N_{i=1}}$, where $\lambda_m^i$ is a adjustment multiplier incorporating information about the subsequent option price. In the case of partially observed diffusion or Lévy processes, such adjustment multipliers could be designed using e.g. novel results obtained by Olsson and Ströjby (2009).

References

Bachelier, L. (1900). Théorie de la spéculation. Annales scientifiques de l’É.N.S., 3e série, pp. 21-86.


In Proceedings of the national academy of sciences of the USA, 103, pp. 18438-18443.


