Volatility Estimation via Chaos Expansions

Alireza Javaheri
Ecole des Mines de Paris, France

1 Introduction

The question of estimation of volatility has been studied extensively in the past two decades. As discussed for instance in (Javaheri, Lautier and Galli 2003) one possible approach would be the use of nonlinear filtering. In that article techniques such as the extended Kalman filter (EKF), the unscented Kalman filter (UKF) and simulation based particle filtering (PF) were applied.

On the other hand, the concept of Wiener chaos expansions (WCE) has been recently introduced in finance in the context of Malliavin calculus (Oksendal 1997), option hedging (Lacoste 1996) as well as unified framework for interest rate and foreign exchange modeling (Hughston and Rafailidis 2005). Interestingly WCE has a natural application to the problem of nonlinear filtering as pointed by many such as (Lototsky, Mikulevicius and Rozovskii 1997).

In this paper this idea is used to present a series expansion allowing us to find the optimal estimation of the hidden volatility.

2 The Estimation Problem

We suppose we have an unobservable variance v_t which could correspond either the instantaneous volatility in a stochastic volatility framework, or the implied volatility in a stochastic implied volatility model.

We assume that the initial volatility has a known distribution

$$v_0 \sim q$$

and follows the stochastic differential equation (SDE)

$$dv_t = f(t, v_t)dt + \sigma(t, v_t)dz_t \tag{1}$$

where dz_t is a Brownian motion (BM) and $f(t, v)$ and $\sigma(t, v)$ are known deterministic functions. For instance for a square root stochastic volatility model (with zero correlation[1]) we would have under the risk neutral measure $f(t, v) = \kappa(\theta - v_t)$ and $\sigma(t, v) = \sigma\sqrt{v_t}$ where the parameters κ, θ and σ are the usual speed of mean reversion, long term variance and volatility of volatility.

The observation y_t could be for instance the option price

$$dy_t = h(v_t)dt + dw_t \tag{2}$$

where dw_t is a BM uncorrelated with dz_t and $h(.)$ is a (usually nonlinear) deterministic function. For example, for a stochastic implied volatility model $h(.)$ will be the Black Scholes function, and for the square root stochastic volatility model $h(.)$ will be the Heston closed form model. As for dw_t it could be considered as the "market noise".

In a discrete observation framework, having $y_n = h(v_n) + B_n$, it is known (Kushner and Budhiraja A. S. 2000) that the optimal conditional density ρ_n of v_n given all the observations y_k for $0 \le k \le n$ follows the iterative Bayes equation

$$\rho_n(v) = c_n.\mathcal{L}_{n-1,n}(\rho_{n-1})(v).\exp\left\{-(y_n - h(v))^2/2\right\}$$

where c_n is a normalizing constant and $\mathcal{L}_{n-1,n}(p)$ is the Fokker-Plank solution on $[t_{n-1}, t_n]$. It is interesting to note that setting $\exp\left\{-(y_n - h(v))^2/2\right\}$ and repeating this operation many times, we shall have

$$\rho_n(v) \propto \eta_n(v).\mathcal{L}_{n-1,n}\left(\eta_{n-1}(v)\mathcal{L}_{n-2,n-1}\left(\ldots\eta_1(v).\mathcal{L}_{0,1}q(v)\right)\right)$$

which is a form of chaotic expansion. Therefore the concept of chaos is introduced quite naturally in this manner.

In the continuous framework the optimal conditional density in $\pi_t(g) = E[g(v_t)|\mathcal{Y}_t]$, where $g(.)$ is any deterministic function (for instance identity $g(v_t) = v_t$) and $\mathcal{Y}_t$ all the information between times 0 and t, follows a stochastic partial derivatives equation (SPDE) commonly referred to as the Kushner-Stratonovich (KS) equation

$$\pi_t(g) = \pi_0(g) + \int_0^t \pi_s(\mathcal{L}_s g)ds + \int_0^t [\pi_s(gh) - \pi_s(g)\pi_s(h)](dy_s - \pi_s(h)ds)$$

with again $\mathcal{L}$ the Fokker-Planck operator associated to the volatility SDE

$$\mathcal{L}(p) = -\frac{\partial(p(v)f(v))}{\partial v} + \frac{1}{2}\frac{\partial^2(p(v)\sigma^2(v))}{\partial v^2} \quad (3)$$

By performing a change of measure we can significantly simplify the KS equation. Indeed, posing

$$\lambda_t = \exp\left(-\int_0^t h(v_s)dw_s - \frac{1}{2}\int_0^t h^2(v_s)ds\right)$$

and making the change of measure $dP^*/dP = \lambda_t$ where dy_t becomes a BM uncorrelated to dv_t, we shall have from the conditional version of the Bayes theorem (Elliott, Aggoun and Moore 1995)

$$\pi_t(g) = E[g(v_t)|\mathcal{Y}_t] = \frac{E^*[g(v_t)\lambda_t^{-1}|\mathcal{Y}_t]}{E^*[\lambda_t^{-1}|\mathcal{Y}_t]}$$

calling $u(t, v)$ the unnormalized density, we will therefore get

$$E[g(v_t)|\mathcal{Y}_t] = \frac{\int_R g(v).u(t,v)dv}{\int_R u(t,v)dv} \quad (4)$$

and this new density $u(t, v)$ follows a simpler SPDE called the Zakai equation

$$du(t,v) = \mathcal{L}(u)dt + h(v).u(t,v).dy_t \quad (5)$$

3 The Chaos Expansion

One way to solve for $u(t, v)$ would be to use a WCE as follows. Taking an orthonormal functional basis $(m_k(t))_{k\geq 1}$ we can introduce the complete orthonormal system

$$\xi_\alpha(y) = \prod_k \frac{1}{\sqrt{\alpha_k!}} H_{\alpha_k}\left(\int_0^t m_k(s)dy_s\right)$$

where $\alpha = (\alpha_k)_{k\geq 1}$ is a multi-index of integers such that the length $|\alpha| = \sum_{k\geq 1}\alpha_k$ remains finite and $H_n(x)$ s the Hermite polynomial of order n.

A form of WCE (referred to as the "second" form or the Cameron-Martin version) states that we can write

$$u(t,v) = \sum_{k\geq 0}\sum_{|\alpha|=k} \frac{1}{\sqrt{\alpha!}}\xi_\alpha(y).\phi_\alpha(t,v) \quad (6)$$

where $\phi_n(t, v)$ are *deterministic* coefficients to be determined.

It is fundamental to note that this expansion (much like Fourier or Karhunen-Loeve expansions) *separates* the SPDE in two pieces: a deterministic piece represented by the functions $\phi_n(t, v)$ which can be solved just by knowing the volatility dynamics and the nature of the nonlinear observation function $h(v)$, and on the other hand, a stochastic piece putting all the randomness of dy_s in the Hermite polynomials. Hence the computation of the deterministic functions could be carried out separately and could be viewed as a calibration process.

As shown in (Lototsky, Mikulevicius and Rozovskii 1997) the above coefficients could be obtained by solving (deterministic) PDE's with usual techniques. The PDE's could be written as

$$\frac{\partial\phi_\alpha(t,v)}{\partial t} = \mathcal{L}(\phi_\alpha)(t,v) + \sum_k \alpha_k.m_k(t).h(v).\phi_{\alpha^*(k)}(t,v) \quad (7)$$

where $\alpha^*(k) = (\alpha_1, \ldots, \alpha_{k-1}, \alpha_k - 1, \alpha_{k+1}, \ldots)$ is another multi-index. A quick justification for the above is the following:

It can be verified that the above polynomial $\xi_\alpha(y)$ verifies

$$\xi_\alpha(y) = \frac{1}{\sqrt{\alpha!}}\frac{\partial^\alpha}{\partial z^\alpha}P_t(z)|_{z=0}$$

where $\frac{\partial^\alpha}{\partial z^\alpha}$ is the derivative taken upon all indexes α_k's present in α and

$$P_s(z) = \exp\left\{\int_0^s \sum_{k\geq 1} m_k(\tau)z_k.dy(\tau) - \frac{1}{2}\int_0^s \sum_{k\geq 1}(m_k(\tau)z_k)^2 d\tau\right\}$$

with (z_k) a sequence of real variables.

Clearly

$$\frac{dP_s(z)}{P_s(z)} = \sum_{k\geq 1} m_k(s)z_k dy(s)$$

and $u(t, v)$ satisfying the Zakai equation (5), by Ito lemma we shall have

$$d(u(s,v)P_s(z)) = \mathcal{L}u(s,v)P_s(z)ds + h(v)\sum_{k\geq 1} m_k(s)z_k u(s,v)P_s(z)ds + h(v)u(s,v)P_s(z)dy(s) + u(s,v)\sum_{k\geq 1} m_k(s)z_k P_s(z)dy(s)$$

and taking expectations $E^*[.]$ and setting $\phi(s, v, z) = E^*[u(s, v).P_s(z)]$ we will obtain the deterministic PDE[2]

$$\frac{\partial\phi}{\partial s} = \mathcal{L}\phi + \sum_{k\geq 1} m_k(s)z_k h(v)\phi$$

On the other hand by definition of the Cameron-Martin decomposition, the coefficient

$$\phi_\alpha(s,v) = \sqrt{\alpha!}E^*[u(s,v).\xi_\alpha(y)]$$

therefore using the previous identification of $\xi_\alpha(y)$

$$\phi_\alpha(s,v) = \frac{\partial^\alpha}{\partial z^\alpha}E^*[u(s,v)P_s(z)]|_{z=0}$$

applying the operator $\frac{\partial^\alpha}{\partial z^\alpha}$ to the previous PDE

$$\frac{\partial\phi_\alpha}{\partial s} = \mathcal{L}\phi_\alpha + \sum_{k\geq 1} m_k(s)h(v)\frac{\partial^\alpha}{\partial z^\alpha}[z_k\phi(s,v,z)]|_{z=0}$$

(QED)

Practically we can choose a length $|\alpha| < N$ and an order (greatest non-zero index) $d(\alpha) < n$ and subdivide the time interval over sub-intervals of a length Δ as $\{[t_{i-1}, t_i]\}_{i\geq 1}$ and solve

$$\frac{\partial \phi^i_\alpha(t,v)}{\partial t} = \mathcal{L}(\phi^i_\alpha)(t,v) + \sum_k \alpha_k.m_k(t).h(v).\phi^i_{\alpha^*(k)}(t,v)$$

over each $[t_{i-1}, t_i]$ with the boundary condition $\phi^i_\alpha = u(t_{i-1}, v)$ if $|\alpha| = 0$ and the null function otherwise.

A convenient choice for the orthonormal basis $(m^i_k(s))_{k\geq 1}$ would be setting $m^i_k(s) = m_k(s - t_{i-1})$ and

$$m_1(s) = 1/\sqrt{\Delta}$$
$$m_k(s) = \sqrt{2/\Delta}\cos\left(\frac{\pi(k-1)s}{\Delta}\right) \tag{8}$$

for $k > 1$ within $[0, \Delta]$ and null outside.

With this choice, we will have

$$\int m_1(s)dy(s) \approx \frac{y(t_i) - y(t_{i-1})}{\sqrt{\Delta}}$$

and for $k > 1$ taking over $[t_{i-1}, t_i]$

$$y_s \approx y_{i-1} + \frac{y_i - y_{i-1}}{\Delta}s$$

with $0 \leq s \leq \Delta$ we will get

$$\int m_k(s)dy(s) \approx \frac{\sqrt{2\Delta}}{\pi^2(k-1)^2}(y_i - y_{i-1})\left((-1)^{k-1} - 1\right)$$

3.1 Chaos Expansion of Order One

We can then take simple cases such as $N = n = 1$ and

$$\frac{\partial \phi^i_0(t,v)}{\partial t} = \mathcal{L}(\phi^i_0)(t,v)$$
$$\frac{\partial \phi^i_1(t,v)}{\partial t} = \mathcal{L}(\phi^i_1)(t,v) + m_1(t).h(v).\phi^i_0(t,v)$$

with boundary condition $\phi^i_0(0,v) = u(t_{i-1}, v)$ and $\phi^i_1(0,v) = 0$ and of course $u(t_0, v) = q(v)$

As for the orthonormal set we take

$$m_1(s) = 1/\sqrt{\Delta}$$

Having ϕ_0 and then ϕ_1 over the time interval $[t_{i-1}, t_i]$ then

$$u(t_i, v) \approx \phi_0(t,v) + \phi_1(t,v)\xi_1 \tag{9}$$

where

$$\xi_1 = \int m_1(s)dy(s) = \frac{y(t_i) - y(t_{i-1})}{\sqrt{\Delta}}$$

Note that this would be equivalent to applying an Euler discretization scheme to the Zakai equation.

3.1.1 Solving the PDE's

The PDE's could be solved for example via a finite difference (FD) technique. Taking a time grid $t_i = t_0 + i\Delta$ and a space grid $(v^j)_{1<j<M}$ with an interval δ we can write

$$\begin{aligned} u^j_i = u^j_{i-1}.[1 - \frac{\Delta}{\delta^2}\sigma(t_{i-1}, v^j_{i-1}) + h(v^j_{i-1}).(y_i - y_{i-1})] \\ + u^{j+1}_{i-1}.[-\frac{\Delta}{2\delta}f(t_{i-1}, v^{j+1}_{i-1}) + \frac{\Delta}{2\delta^2}\sigma(t_{i-1}, v^{j+1}_{i-1})] \\ + u^{j-1}_{i-1}.[\frac{\Delta}{2\delta}f(t_{i-1}, v^{j-1}_{i-1}) + \frac{\Delta}{2\delta^2}\sigma(t_{i-1}, v^{j-1}_{i-1})] \end{aligned}$$

this is a forward equation with the initial condition $u^j_0 = q^j$.

It is fundamental to note that this FD scheme can be separated in two pieces

$$\begin{aligned} u0^j_i = u^j_{i-1}.\left[1 - \frac{\Delta}{\delta^2}\sigma\left(t_{i-1}, v^j_{i-1}\right)\right] \\ + u^{j+1}_{i-1}.\left[-\frac{\Delta}{2\delta}f\left(t_{i-1}, v^{j+1}_{i-1}\right) + \frac{\Delta}{2\delta^2}\sigma\left(t_{i-1}, v^{j+1}_{i-1}\right)\right] \\ + u^{j-1}_{i-1}.\left[\frac{\Delta}{2\delta}f\left(t_{i-1}, v^{j-1}_{i-1}\right) + \frac{\Delta}{2\delta^2}\sigma\left(t_{i-1}, v^{j-1}_{i-1}\right)\right] \end{aligned}$$

which corresponds to Chaos of order zero, or the Fokker-Planck equation, and

$$u1^j_i = \left[u^j_{i-1}.h(v^j_{i-1}).\sqrt{\Delta}\right].(y_i - y_{i-1})/\sqrt{\Delta}$$

which adds the first degree supplement, and of course

$$u^j_i = u0^j_i + u1^j_i$$

this is because $\phi^i_1(0,v) = 0$ which is also true for all higher order ϕ's. Therefore the actual Fokker-Planck FD happens only for ϕ_0 and then we can add the higher order terms.

Having obtained u^j_i in this manner, the estimation of the variance v_i will be

$$\hat{v}_i \approx \frac{\sum_{j=1}^M u^j_i v^j_i}{\sum_{j=1}^M u^j_i}$$

3.2 Chaos Expansion of Order Two

We take $|\alpha| \leq N = 2$ and $d(\alpha) \leq n = 2$ which means in addition to the above functions we will have ϕ_2 as well as $\phi_{1,1}, \phi_{2,2}$ and $\phi_{1,2}$

$$\frac{\partial \phi^i_2(t,v)}{\partial t} = \mathcal{L}(\phi^i_2)(t,v) + m_2(t).h(v).\phi^i_0(t,v)$$

as well as

$$\frac{\partial \phi^i_{1,1}(t,v)}{\partial t} = \mathcal{L}(\phi^i_{1,1})(t,v) + 2m_1(t).h(v).\phi^i_1(t,v)$$

where $\alpha_1 = 2$ and $\alpha_2 = 0$

$$\frac{\partial \phi^i_{2,2}(t,v)}{\partial t} = \mathcal{L}(\phi^i_{2,2})(t,v) + 2m_2(t).h(v).\phi^i_2(t,v)$$

where $\alpha_1 = 0$ and $\alpha_2 = 2$

$$\frac{\partial \phi^i_{1,2}(t,v)}{\partial t} = \mathcal{L}(\phi^i_{1,2})(t,v) + m_2(t).h(v).\phi^i_1(t,v) + m_1(t).h(v).\phi^i_2(t,v)$$

where $\alpha_1 = 1$ and $\alpha_2 = 1$.

This therefore involves solving six PDE's instead of two, and[3]

$$\xi_2 = \int m_2(s)dy(s)$$

$$\xi_{1,1} = \frac{1}{\sqrt{2}}\left[\left(\int m_1(s)dy(s)\right)^2 - 1\right]$$

$$\xi_{2,2} = \frac{1}{\sqrt{2}}\left[\left(\int m_2(s)dy(s)\right)^2 - 1\right]$$

$$\xi_{1,2} = \xi_{2,1} = \left(\int m_1(s)dy(s)\right)\left(\int m_2(s)dy(s)\right)$$

and

$$u(t,v) \approx \phi_0(t,v) + \phi_1(t,v)\xi_1 + \phi_2(t,v)\xi_2 + \phi_{1,1}(t,v)\xi_{1,1} + 2\phi_{1,2}(t,v)\xi_{1,2} + \phi_{2,2}(t,v)\xi_{2,2} \quad (10)$$

which completes the algorithm. The numeric solving via FD could be carried out similarly to the first degree expansion.

4 Future Work

In order to test the performance of the above tools, we could simulate a time series of (theoretically unknown) instantaneous variances v_i as well as for instance three-months at-the-money option prices $y_i = C(v_i) + w_i$ with $C(v)$ the Heston closed form function for a given set of parameters (κ, θ, σ) and zero stock volatility correlation. Given that we will have generated this series artificially, we will have access to the unobservable v_i and we will therefore be able to estimate the error $e_i = |v_i - \hat{v}_i|$. Hence we could compare the performance of these tools to the commonly used EKF and UKF.

Preliminary tests indicate that the WCE of order one is comparable to EKF, however the higher order WCEs outperform both EKF and UKF. This is not surprising since as mentioned for instance in (Ito and Xiong 2000) an Euler discretization of the Zakai equation and the application of the usual discrete filter (Javaheri, Lautier and Galli 2003)

$$p(v_i|y_{1:i-1}) = \int p(v_i|v_{i-1})p(v_{i-1}|y_{1:i-1})dv_{i-1}$$

$$p(v_i|y_{1:i}) = \frac{p(y_i|v_i)p(v_i|y_{1:i-1})}{p(y_i|y_{1:i-1})}$$

would precisely provide a Gaussian filter such as Kalman's. Furthermore, as we saw this corresponds to a first order chaos expansion. A more accurate measure of the performance of these chaos based algorithms is left to a future publication in the interest of brevity.

FOOTNOTES & REFERENCES

1. We could introduce the correlation by using techniques such as the ones discussed in (Javaheri, Lautier and Galli 2003).
2. Needless to say, the operator $\mathcal{L}$ only affects the variable v.
3. Note that $\phi_{1,2} = \phi_{2,1}$

■ Elliott R., Aggoun L., Moore J.B. (1995) "Hidden Markov Models: Estimation and Control" *Springer-Verlag*

■ Hughston L.P., Rafailidis A. (2005) "A Chaotic Approach to Interest Rate Modeling" *Finance and Stochastics*, Volume 9, Number 1

■ Ito K., Xiong K. (2000) "Gaussian Filters for Nonlinear Filtering Problems" *IEEE Transactions on Automatic Control*, Volume 45, Number 5

■ Javaheri A., Lautier D., Galli A. (2003) "Filtering in Finance" *Wilmott*, May 2003 Issue, Number 5

■ Kushner H. J., Budhiraja A. S. (2000) "A Nonlinear Filtering Algorithm based on an Approximation of the Conditional Distribution" *IEEE Transactions on Automatic Control*, Volume 45, Number 3

■ Lacoste V. (1996) "Wiener Chaos: A New Approach to Option Hedging" *Mathematical Finance*, Volume 6, Number 2

■ Lototsky S., Mikulevicius R., Rozovskii B. (1997) "Nonlinear Filtering Revisited: A Spectral Approach" *SIAM Journal of Control and Optimization*, Volume 35, Number 2

■ Oksendal B. (1997) "An Introduction to Malliavin Calculus with Applications to Economics" *Department of Mathematics, University of Oslo*