Parameter Estimation for a Stochastic Volatility Model with Additive and Multiplicative Noise

Ibukun Amusan

Professor: Dr. Ewald

June 11, 2012

Contents

Chapter 1

Introduction

The pricing of derivatives is a problem that people have been examining over the years. The model by Black and Scholes[**?**] provided a framework in which it is possible to solve the problem. Their model formed a basis for the pricing of derivatives from the simplest to the most exotic. However, the assumptions of the model placed some limitations on its use. These limitations on the model have prompted and still continue to prompt various attempts and corrections that are aimed at finding models that are more accurate with respect to prices that are obtained in markets.

One of the assumptions underlying the model is that of constant volatility. The graph of the implied volatility of the model against the strike price produces a smile curve, which shows that volatility is not constant. Many models have been proposed to correct the Black-Scholes model by making volatility non-constant. Stochastic volatility models are models whereby the volatility is assumed to be a stochastic process with its own dynamics. A property that is particularly desirable for such models is mean-reversion.

An Ornstein-Uhlenbeck (OU) process is an example of a process with the mean-reverting property and some stochastic volatility models assume that volatility has the dynamics of an OU process. Stochastic volatility models are meant to account for the skewness observed in the implied volatility curve. The stationary distribution of an OU process is Gaussian, which has constant skewness and kurtosis, regardless of its parameters. However, the skewness and kurtosis of historical volatilities are not constant. This flaw can be improved on by adding another source of noise to the model to give another stochastic process with additive and multiplicative noise.

The skewness and kurtosis of the stochastic process with additive and multiplicative noise has been found to be not constant. Also, results obtained from simulation show that the graph of kurtosis against skewness of paths followed by the process with additive and multiplicative noise has a similar shape to the graph of kurtosis versus skewness of the log of historical volatilities.

For parameter estimation, we first note that the Euler discretization used for simulation provide a linear relationship between successive pairs of data. Least squares regression method is used to find the parameters for the linear relationship, and these parameters are then used to find the parameters for the model. The first two parameters are found directly from the least square parameters and the remaining three parameters are estimated by further using the maximum likelihood method on the residuals.

Chapter 2

Background for Stochastic Volatility Models

Stochastic volatility models for option pricing were developed to try to correct the unrealistic assumption of constant volatility of the Black-Scholes model. The assumption of constant volatility is particularly shown to be untrue by the implied volatility of the model.

The implied volatility is the volatility that makes the price from the Black-Scholes formula equal to the market price of the option. Different strike prices and maturities yield different implied volatilities. This is clearly seen from the skewness and so called smile curve which is obtained by plotting the implied volatilities against the strike prices of options on a particular underlying asset.

Let $C_{BS}(S_t, t; K, T; \sigma)$ be the price obtained from the Black-Scholes formula for a European call option on an underlying asset S_t at time t with strike K, maturity T and constant volatility σ. Let C^{obs} be the observed market value for the call option. The constant volatility that makes the Black-Scholes call value equal to the market price is $\sigma = I$, where I is called the implied volatility. So we have

$$C_{BS}(S_t, t; K, T; I) = C^{obs}. \tag{2.1}$$

Several models have been proposed to handle assets with non-constant volatility as evidenced by the differing implied volatilities obtained using the Black-Scholes formula for options on the same asset. One such general class of models, referred to as local volatility models, uses deterministic functions of the asset price and time for the volatility. Local volatility models provide improvements on the Black-Scholes model but they are still not good enough as volatility is in no way deterministic.

Another class of models, known as stochastic volatility models, treats volatility as a stochastic process. Here, volatility is taken to be a random variable having its own stochastic differential equation. A popular feature of this class of model is mean-reversion. Although there are other models, the mean-reverting property is particularly desirable because it can help to keep the growth of the volatility in check.

An Ornstein-Uhlenbeck (OU) process is an example of a mean-reverting process that is used by some stochastic volatility models. The OU process satisfies the stochastic differential equation

$$dY_t = \kappa(\theta - Y_t)dt + \alpha dW_t^1 \tag{2.2}$$

where κ, θ, α, μ are constants, and W_t^1 is a Wiener process. The OU process has a Gaussian distribution. Its stationary distribution is Gaussian with mean θ and variance $\frac{\alpha^2}{2\kappa}$.

The problem with using the OU process for volatility is that the observed volatility does not really follow a Gaussian distribution. Figure 2.1 and figure 2.2 are plots of the skewness against kurtosis of historical volatilities and log of historical volatilities respectively. The historical data was obtained from the Chicago Board Options Exchange (CBOE) website. The graphs show that the skewness and kurtosis of volatility are not constants. These call into question the suitability of the OU process for stochastic volatility models.

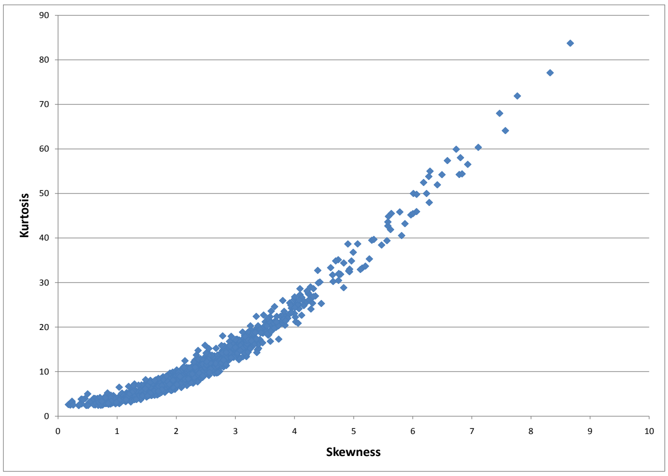

Figure 2.1: Skewness vs kurtosis for historical volatilities of stock returns. Data obtained from CBOE.

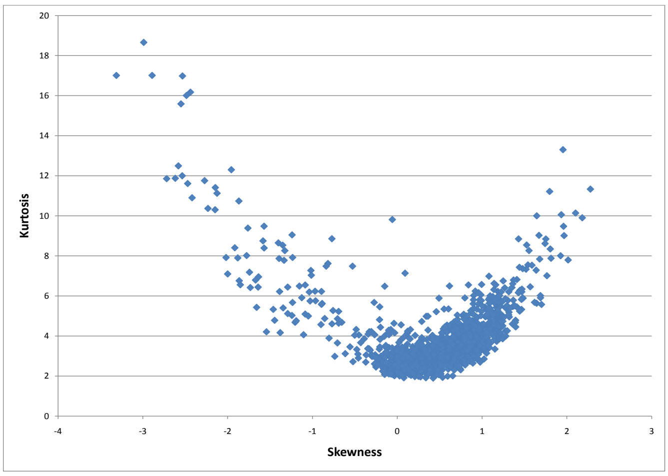

Figure 2.2: Skewness vs kurtosis for log of historical volatilities of stock returns. Data obtained from CBOE.

Chapter 3

Coupled Additive-Multiplicative Noise Model

The OU process used by some stochastic volatility models has one source of noise, which is additive. Another type of stochastic model that will now be considered has an extra source of noise, the multiplicative noise, in addition to the additive noise of the OU process.

Consider the stochastic model for a process Y_t:

$$dY_t = \kappa(\theta - Y_t)dt + \alpha dW_t^1 + \beta Y_t dW_t^2 \tag{3.1}$$

where W_t^1 and W_t^2 are correlated Wiener processes. αdW_t^1 is the additive noise and $\beta V_t dW_t^2$ is the multiplicative noise part.

Let

$$\begin{aligned} W_t^1 &= \rho Z_t^2 + \sqrt{1-\rho^2} Z_t^1 \\ W_t^2 &= Z_t^2 \end{aligned}$$

Then

$$dY_t = \kappa(\theta - Y_t)dt + \alpha\sqrt{1-\rho^2} dZ_t^1 + (\alpha\rho + \beta Y_t) dZ_t^2$$

where Z_t^1 and Z_t^2 are independent Wiener processes.

Let $\mathbb{E}[Y_t] = \mu_1$, $\mathbb{E}[Y_t^2] = \mu_2$, $\mathbb{E}[Y_t^3] = \mu_3$, ... For the process Y_t,

$$d\mu_1 = \kappa(\theta - \mu_1)dt. \tag{3.2}$$

Solving for μ_1 gives

$$\theta - \mu_1 = C_1 e^{-\kappa t}, \tag{3.3}$$

C_1 a constant, and this further implies that

$$\mu_1 = \theta(1 - e^{-\kappa t}) + Y_0 e^{-\kappa t} \tag{3.4}$$

where Y_0 is the initial value of Y_t. As $t \to \infty$, $\mu_1 \to \theta$. And so, θ is the long-run mean of the process Y_t.

For Y_t^2, the stochastic differential equation is given by

$$dY_t^2 = 2Y_t dY_t + (dY_t)^2. \tag{3.5}$$

Substituting Y_t and simplifying gives

$$\begin{aligned} dY_t^2 &= [(\beta^2 - 2\kappa)Y_t^2 + 2(\kappa\theta + \alpha\beta\rho)Y_t + \alpha^2]dt \\ &\quad + 2\alpha\sqrt{1-\rho^2}Y_t dZ_t^1 + 2(\alpha\rho + \beta Y_t)Y_t dZ_t^2. \end{aligned} \tag{3.6}$$

Taking the expectation then yields

$$d\mu_2 = [(\beta^2 - 2\kappa)\mu_2 + 2(\kappa\theta + \alpha\beta\rho)\mu_1 + \alpha^2]dt. \tag{3.7}$$

To have finite moments as $t \to \infty$, the condition $\beta^2 - 2\kappa < 0$ must hold.

Doing similar calculations for μ_3, μ_4, μ_5 and μ_6, we obtain

$$d\mu_3 = [3(\beta^2 - \kappa)\mu_3 + 3(\kappa\theta + 2\alpha\beta\rho)\mu_2 + 3\alpha^2\mu_1]dt, \tag{3.8}$$

$$d\mu_4 = [2(3\beta^2 - 2\kappa)\mu_4 + 4(\kappa\theta + 3\alpha\beta\rho)\mu_3 + 6\alpha^2\mu_2]dt, \tag{3.9}$$

$$d\mu_5 = [5(2\beta^2 - \kappa)\mu_5 + 5(\kappa\theta + 4\alpha\beta\rho)\mu_4 + 10\alpha^2\mu_3]dt, \tag{3.10}$$

and

$$d\mu_6 = [3(5\beta^2 - 2\kappa)\mu_6 + 6(\kappa\theta + 5\alpha\beta\rho)\mu_5 + 15\alpha^2\mu_4]dt. \tag{3.11}$$

To have finite moments as $t \to \infty$, the following conditions must hold:

$$\text{For } \mu_2: \quad \beta^2 < 2\kappa \quad \text{or} \quad \beta^2 < \frac{2}{1}\kappa. \tag{3.12}$$

$$\text{For } \mu_3: \quad \beta^2 < \kappa \quad \text{or} \quad \beta^2 < \frac{2}{2}\kappa. \tag{3.13}$$

$$\text{For } \mu_4: \quad 3\beta^2 < 2\kappa \quad \text{or} \quad \beta^2 < \frac{2}{3}\kappa. \tag{3.14}$$

$$\text{For } \mu_5: \quad 2\beta^2 < \kappa \quad \text{or} \quad \beta^2 < \frac{2}{4}\kappa. \tag{3.15}$$

$$\text{For } \mu_6: \quad 5\beta^2 < 2\kappa \quad \text{or} \quad \beta^2 < \frac{2}{5}\kappa. \tag{3.16}$$

Generally, for μ_i, $i \geq 2$,

$$\beta^2 < \frac{2}{i-1}\kappa. \tag{3.17}$$

Solving the differential equations (3.7) to (3.9) for μ_2, μ_3 and μ_4 gives

$$\mu_2 = C_2 e^{(\beta^2-2\kappa)t} + \frac{2C_1(\kappa\theta + \alpha\beta\rho)}{\beta^2 - \kappa} e^{-\kappa t} - \frac{2(\kappa\theta + \alpha\beta\rho)\theta + \alpha^2}{\beta^2 - 2\kappa} \tag{3.18}$$

$$\mu_3 = C_3 e^{3(\beta^2-\kappa)t} - \frac{3C_2(\kappa\theta + 2\alpha\beta\rho)}{2\beta^2 - \kappa} e^{(\beta^2-2\kappa)t} - \frac{3C_1}{3\beta^2 - 2\kappa} \Upsilon e^{-\kappa t} - \frac{1}{\beta^2 - \kappa}\Psi \tag{3.19}$$

$$\begin{aligned}
\mu_4 = {} & C_4 e^{(6\beta^2-4\kappa)t} - \frac{4C_3(\kappa\theta + 3\alpha\beta\rho)}{3\beta^2 - \kappa} e^{3(\beta^2-\kappa)t} \\
& - \frac{C_2}{5\beta^2 - 2\kappa}\left(6\alpha^2 - \frac{12(\kappa\theta + 2\alpha\beta\rho)(\kappa\theta + 3\alpha\beta\rho)}{2\beta^2 - \kappa}\right) e^{(\beta^2-2\kappa)t} \\
& - \frac{C_1}{6\beta^2 - 3\kappa}\left(\frac{12\alpha^2(\kappa\theta + \alpha\beta\rho)}{\beta^2 - \kappa} - \frac{12(\kappa\theta + 3\alpha\beta\rho)}{3\beta^2 - 2\kappa}\Upsilon\right) e^{-\kappa t} \\
& + \frac{1}{6\beta^2 - 4\kappa}\left(\frac{6\alpha^2(2(\kappa\theta + \alpha\beta\rho)\theta + \alpha^2)}{\beta^2 - 2\kappa} + \frac{4(\kappa\theta + 3\alpha\beta\rho)}{\beta^2 - \kappa}\Psi\right)
\end{aligned} \tag{3.20}$$

where

$$\Upsilon = \left(\frac{2(\kappa\theta + \alpha\beta\rho)(\kappa\theta + 2\alpha\beta\rho)}{\beta^2 - \kappa} - \alpha^2 \right) \tag{3.21}$$

and

$$\Psi = \left(\alpha^2\theta - (\kappa\theta + 2\alpha\beta\rho) \left(\frac{2(\kappa\theta + \alpha\beta\rho)\theta + \alpha^2}{\beta^2 - 2\kappa} \right) \right). \tag{3.22}$$

$C_1 = \theta - Y_0$. C_2, C_3 and C_4 are constants that can be obtained by setting $t = 0$ in the above solutions for μ_2, μ_3 and μ_4.

Let $\mu_i' = \lim_{t\to\infty} \mu_i$. $\mu_1' = \theta$ from equation (3.4). Suppose β and κ satisfy the inequality in (3.17). Then,

$$\begin{aligned}
\mu_2' &= -\frac{2(\kappa\theta + \alpha\beta\rho)\theta + \alpha^2}{\beta^2 - 2\kappa} && (3.23)\\
&= -\frac{1}{\beta^2 - 2\kappa}\left[2(\kappa\theta + \alpha\beta\rho)\mu_1' + \alpha^2\right] && (3.24)\\
\mu_3' &= -\frac{1}{\beta^2 - \kappa}\left[(\kappa\theta + 2\alpha\beta\rho)\mu_2' + \alpha^2\mu_1'\right] && (3.25)\\
\mu_4' &= -\frac{1}{6\beta^2 - 4\kappa}\left[4(\kappa\theta + 3\alpha\beta\rho)\mu_3' + 6\alpha^2\mu_2'\right]. && (3.26)
\end{aligned}$$

Let $\hat{\mu}$, $\hat{\sigma}^2$, $\hat{\gamma}$ and $\hat{K}$ be the stationary mean, variance, skewness and kurtosis of Y_t respectively. Then, using the definition of standardized moments of random variables, we get

$$\begin{aligned}
\hat{\mu} &= \theta && (3.27)\\
\hat{\sigma}^2 &= -\frac{(\theta\beta + \alpha\rho)^2 + \alpha^2(1 - \rho^2)}{\beta^2 - 2\kappa} && (3.28)\\
\hat{\gamma} &= \frac{2\alpha\beta\rho(\alpha^2 + 3\theta^2\beta^2) + 2\theta\beta^2(\alpha^2 + 2\alpha^2\rho^2 + \theta^2\beta^2)}{(\beta^2 - 2\kappa)(\beta^2 - \kappa)\hat{\sigma}^3} && (3.29)\\
\hat{K} &= \frac{\hat{A}}{(6\beta^2 - 4\kappa)(\beta^2 - 2\kappa)(\beta^2 - \kappa)\hat{\sigma}^4} && (3.30)
\end{aligned}$$

where

$$\begin{aligned}\hat{A} &= 6\alpha^4(\beta^2-\kappa)-6\theta^4\beta^4(3\beta^2+\kappa)-12\beta^2\alpha^2\theta^2(\beta^2+\kappa)\\ &\quad -24\alpha^2\beta^2\rho^2(\alpha^2+\kappa\theta^2+4\beta^2\theta^2)\\ &\quad -24\alpha\rho\theta\beta(\alpha^2\beta^2+\kappa\theta^2\beta^2+\alpha^2\kappa+3\theta^2\beta^4+2\alpha^2\beta^2\rho^2). \qquad (3.31)\end{aligned}$$

It can be observed from the above stationary moments for the stochastic model with additive and multiplicative noise that the stationary variance is always positive because of the assumption that $\beta^2 - 2\kappa < 0$. The stationary skewness is not a constant and it can be either positive or negative, depending on the values of the parameters. If the first term $(2\alpha\beta\rho(\alpha^2 + 3\theta^2\beta^2))$ in the numerator of $\hat{\gamma}$ is negative and its absolute value is greater than the second term $(2\theta\beta^2(\alpha^2+2\alpha^2\rho^2+\theta^2\beta^2))$ in the numerator, then the stationary skewness will be negative; otherwise, the stationary skewness will be positive. The stationary kurtosis, $\hat{K}$, depends on κ, θ, α, β and ρ, and is therefore also not a constant.

The fact that the skewness and kurtosis are not constant for different parameters is enough evidence that the random variable Y_t is not Gaussian, since Gaussian distribution has constant skewness and kurtosis. This is in contrast to the OU process without multiplicative noise, which has a Gaussian distribution. Thus, the addition of the multiplicative noise to the OU process results in another process with a non-Gaussian distribution.

Monte carlo simulation of the paths followed by the process Y_t in equation (3.1) suggests that the stochastic equation may be used to model the log of volatility of stock prices. Figure 3.1 is the plot of the skewness against kurtosis of some simulated paths of Y_t and it shows that the skewness and kurtosis are not constant for different parameters of the model. The graph has a similar shape to the graph of the log of volatility for historical data in figure 2.2 and both graphs can be fitted on top of a parabola.

Stochastic models of this kind have been used for modeling in other fields. In meteorology and oceanography, Sardeshmukh and Sura[**?**, **?**] have investigated and proposed the suitability of using the stochastic model with additive and multiplicative noise for daily sea surface temperature variations. They showed that the scattered plot of the kurtosis against the skewness of some

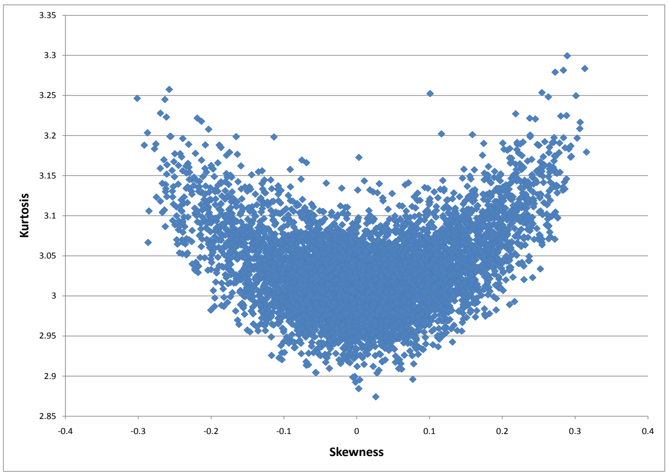

Figure 3.1: Skewness vs kurtosis for the simulated paths of process Y_t of equation (3.1).

observed data fits above a parabola. From their analysis of the moments of a process with the stochastic equation with multiplicative noise, they found that the kurtosis is greater than or equal three-halves of the squared-skewness ($kurt \geq (3/2)skew^2$), which gives a lower limit for the graph of the observed data. The graph of the kurtosis and skewness of the log of volatilities of stocks in figure 2.2 is similar to the graph obtained by Sardeshmukh and Sura[**?**] for the variations in sea surface temperature. It is this perceived similarity of the stochastic nature of ocean temperature variations and stock volatilities that informed this study.

In finance, the steady state probability density function (PDF) has been obtained by using the Fokker-Planck equation for the model. Anteneodo and Riera[**?**] considered the case where the Wiener processes, W_t^1 and W_t^2, are not correlated. The PDF obtained provided a good fit to the empirical PDF. Cheng *et al*[**?**] examined the PDF for when the Wiener processes are correlated.

Chapter 4

Parameter Estimation

The coupled additive-multiplicative noise model for stochastic volatility is

$$dY_t = \kappa(\theta - Y_t)dt + \alpha dW_t^1 + \beta Y_t dW_t^2$$

where W_t^1 and W_t^2 are correlated Wiener processes with correlation coefficient ρ.

The stochastic equation can also be written as

$$dY_t = \kappa(\theta - Y_t)dt + \alpha\sqrt{1-\rho^2}dZ_t^1 + (\alpha\rho + \beta Y_t)dZ_t^2$$

where Z_t^1 and Z_t^2 are independent Wiener processes.

In order to simulate paths for Y_t, some form of discretization is needed. Discretizing the model using Euler method gives

$$\begin{aligned} Y_j = Y_{j-1} + \kappa(\theta - Y_{j-1})\delta t + \alpha\sqrt{(1-\rho^2)\delta t}z_1 \\ +(\alpha\rho + \beta Y_{j-1})\sqrt{\delta t}z_2 \end{aligned}$$

where z_1 and z_2 are independent standard random normal.
The terms involving z_1 and z_2 can be combined into one term as follows:

$$Y_j = Y_{j-1}(1 - \kappa\delta t) + \kappa\theta\delta t + \sqrt{\zeta_{j-1}\delta t}z$$

where

$$\zeta_{j-1} = \alpha^2 + 2\alpha\beta\rho Y_{j-1} + \beta^2 Y_{j-1}^2$$

and z is standard normal.

There are five parameters to be estimated: κ, θ,α,β, and ρ. κ and θ are estimated using least squares regression method, and maximum likelihood estimation (MLE) method is further used for the remaining three parameters.

4.1 Least Squares Regression

The last equation shows that there is a linear relationship between Y_{j-1} and Y_j. The Euler Discretization can be written in the form

$$Y_j = Y_{j-1}(a) + b + \sqrt{\delta t}\epsilon_{j-1}$$

where

$$a = 1 - \kappa\delta t$$

$$b = \kappa\theta\delta t$$

and

$$\epsilon_{j-1} = \sqrt{\zeta_{j-1}}z$$

Fitting a line using least squares, we get a, b and ϵ_{j-1}.
The conditional distribution of ϵ_{j-1} given Y_{j-1} is normal with mean zero and variance ζ_{j-1}.

4.1.1 Estimation of κ and θ

Now, $a = 1 - \kappa\delta t$ implies

$$\kappa = \frac{1-a}{\delta t}$$

and $b = \kappa\theta\delta t$ implies

$$\theta = \frac{b}{1-a}.$$

So given the values of a and b, we can estimate κ and θ.

4.2 Maximum Likelihood Estimation of α, β, ρ

Since we now know a and b, we can define the new variable ϵ_{j-1} as follows:

$$\frac{Y_j - (aY_{j-1} + b)}{\sqrt{\delta t}} = \epsilon_{j-1}(\alpha, \beta, \rho; Y_{j-1}) \sim \mathcal{N}(0, \zeta_{j-1})$$

where

$$\zeta_{j-1} = \alpha^2 + 2\alpha\beta\rho Y_{j-1} + \beta^2 Y_{j-1}^2$$

The conditional probability density function of ϵ_{j-1} given Y_{j-1} is

$$f(\epsilon_{j-1} \mid Y_{j-1}) = \frac{1}{\sqrt{2\pi\zeta_{j-1}}} \exp\{-\frac{\epsilon_{j-1}^2}{2\zeta_{j-1}}\}.$$

Using their conditional density, the log-likelihood function of the ϵ_{j-1}'s is

$$\mathcal{L}(\alpha, \beta, \rho) = -\frac{1}{2} \sum (\log 2\pi + \log \zeta_{j-1} + \frac{\epsilon_{j-1}^2}{\zeta_{j-1}}).$$

The maximum of the log-likelihood will give the values for the estimates of the three parameters. Finding the place at which the above log-likelihood function is maximized is equivalent to minimizing the function

$$M(\alpha, \beta, \rho) = \sum (\log \zeta_{j-1} + \frac{\epsilon_{j-1}^2}{\zeta_{j-1}})$$

over α, β and ρ.

The minimization problem can be solved by using some method of optimization. For instance, the problem can be solved by Excel Solver, which uses the "Generalized Reduced Gradient (GRG2) Algorithm" developed by Leon Lasdon and Allan Waren. However, since the minimization is with respect to three variables, optimizing M directly may be computationally costly.

One method for finding the maximum likelihood estimates is to equate the gradient of M to zero and then solve the resulting three equations simultaneously. The values of α, β, ρ for which $\nabla M = 0$, where $\nabla = (\frac{\partial}{\partial\alpha}, \frac{\partial}{\partial\beta}, \frac{\partial}{\partial\rho})$ are the maximum likelihood estimates.

For

$$M(\alpha, \beta, \rho) = \sum (\log \zeta_{j-1} + \frac{\epsilon_{j-1}^2}{\zeta_{j-1}})$$

and

$$\zeta_{j-1} = \alpha^2 + 2\alpha\beta\rho Y_{j-1} + \beta^2 Y_{j-1}^2,$$

$$
\begin{aligned}
\frac{\partial M}{\partial \alpha} &= \sum \left\{ \frac{2\alpha + 2\beta\rho Y_{j-1}}{\zeta_{j-1}} - \frac{(2\alpha + 2\beta\rho Y_{j-1})\epsilon_{j-1}^2}{\zeta_{j-1}^2} \right\} = 0 \\
\frac{\partial M}{\partial \beta} &= \sum \left\{ \frac{2\alpha\rho Y_{j-1} + 2\beta Y_{j-1}^2}{\zeta_{j-1}} - \frac{(2\alpha\rho + 2\beta Y_{j-1}^2)\epsilon_{j-1}^2}{\zeta_{j-1}^2} \right\} = 0 \\
\frac{\partial M}{\partial \rho} &= \sum \left\{ \frac{2\alpha\beta Y_{j-1}}{\zeta_{j-1}} - \frac{(2\alpha\beta Y_{j-1})\epsilon_{j-1}^2}{\zeta_{j-1}^2} \right\} = 0.
\end{aligned}
$$

The equation $\frac{\partial M}{\partial \rho} = 0$ simplifies to

$$\sum \left\{ \frac{Y_{j-1}}{\zeta_{j-1}} - \frac{Y_{j-1}\epsilon_{j-1}^2}{\zeta_{j-1}^2} \right\} = 0$$

which implies that

$$\sum \frac{Y_{j-1}}{\zeta_{j-1}} = \sum \frac{Y_{j-1}\epsilon_{j-1}^2}{\zeta_{j-1}^2}.$$

Also, $\frac{\partial M}{\partial \alpha} = 0$ can be factorized as

$$2\alpha \sum \left\{ \frac{1}{\zeta_{j-1}} - \frac{\epsilon_{j-1}^2}{\zeta_{j-1}^2} \right\} + 2\beta\rho \sum \left\{ \frac{Y_{j-1}}{\zeta_{j-1}} - \frac{Y_{j-1}\epsilon_{j-1}^2}{\zeta_{j-1}^2} \right\} = 0$$

and, upon substituting the result from $\frac{\partial M}{\partial \rho} = 0$, gives

$$\sum \frac{1}{\zeta_{j-1}} = \sum \frac{\epsilon_{j-1}^2}{\zeta_{j-1}^2}.$$

Similarly, $\frac{\partial M}{\partial \beta} = 0$ can be factorized thus:

$$2\alpha\rho \sum \left\{ \frac{Y_{j-1}}{\zeta_{j-1}} - \frac{Y_{j-1}\epsilon_{j-1}^2}{\zeta_{j-1}^2} \right\} + 2\beta \sum \left\{ \frac{Y_{j-1}^2}{\zeta_{j-1}} - \frac{Y_{j-1}^2\epsilon_{j-1}^2}{\zeta_{j-1}^2} \right\} = 0.$$

Substituting the result from $\frac{\partial M}{\partial \rho} = 0$ gives

$$\sum \frac{Y_{j-1}^2}{\zeta_{j-1}} = \sum \frac{Y_{j-1}^2\epsilon_{j-1}^2}{\zeta_{j-1}^2}.$$

So we need to solve the following system of three equations for α, β and ρ :

$$\begin{aligned}
\sum \frac{1}{\zeta_{j-1}} &= \sum \frac{\epsilon_{j-1}^2}{\zeta_{j-1}^2} \\
\sum \frac{Y_{j-1}}{\zeta_{j-1}} &= \sum \frac{Y_{j-1}\epsilon_{j-1}^2}{\zeta_{j-1}^2} \\
\sum \frac{Y_{j-1}^2}{\zeta_{j-1}} &= \sum \frac{Y_{j-1}^2\epsilon_{j-1}^2}{\zeta_{j-1}^2}.
\end{aligned}$$

This system can be simplified further by rewritting ζ_{j-1} :

$$\zeta_{j-1} = \alpha^2 + 2\alpha\beta\rho Y_{j-1} + \beta^2 Y_{j-1}^2.$$

Completing the square in Y_{j-1} gives

$$\zeta_{j-1} = \alpha^2 - \alpha^2\rho^2 + (\alpha\rho + \beta Y_{j-1})^2.$$

and after factoring out $\alpha^2 - \alpha^2\rho^2 = \alpha^2(1-\rho^2)$,

$$\zeta_{j-1} = \alpha^2(1-\rho^2)\left[1 + \left(\frac{\rho}{\sqrt{1-\rho^2}} + \frac{\beta Y_{j-1}}{\alpha\sqrt{1-\rho^2}}\right)\right].$$

Now, let

$$A = \alpha^2(1-\rho^2), B = \frac{\rho}{\sqrt{1-\rho^2}}, C = \frac{\beta Y_{j-1}}{\alpha\sqrt{1-\rho^2}}.$$

Then $\zeta_{j-1} = A\left[1 + (B + CY_{j-1})^2\right]$. That is, ζ_{j-1} is transformed from being a function of α, β, ρ to a function of A, B, C.

Furthermore, let

$$\eta_{j-1} = 1 + (B + CY_{j-1})^2$$

so that

$$\zeta_{j-1} = A\eta_{j-1}.$$

Substituting this into the system of equations gives

$$\begin{aligned}
\sum \frac{1}{A\eta_{j-1}} &= \sum \frac{\epsilon_{j-1}^2}{A^2\eta_{j-1}^2} \\
\sum \frac{Y_{j-1}}{A\eta_{j-1}} &= \sum \frac{Y_{j-1}\epsilon_{j-1}^2}{A^2\eta_{j-1}^2} \\
\sum \frac{Y_{j-1}^2}{A\eta_{j-1}} &= \sum \frac{Y_{j-1}^2\epsilon_{j-1}^2}{A^2\eta_{j-1}^2}.
\end{aligned}$$

Since A can be pulled out of the summation, the system reduces to

$$\begin{aligned}
\sum \frac{1}{\eta_{j-1}} &= \frac{1}{A} \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}^2} \\
\sum \frac{Y_{j-1}}{\eta_{j-1}} &= \frac{1}{A} \sum \frac{Y_{j-1}\epsilon_{j-1}^2}{\eta_{j-1}^2} \\
\sum \frac{Y_{j-1}^2}{\eta_{j-1}} &= \frac{1}{A} \sum \frac{Y_{j-1}^2\epsilon_{j-1}^2}{\eta_{j-1}^2},
\end{aligned}$$

which can also be written as

$$A = \frac{\sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}^2}}{\sum \frac{1}{\eta_{j-1}}}, A = \frac{\sum \frac{Y_{j-1}\epsilon_{j-1}^2}{\eta_{j-1}^2}}{\sum \frac{Y_{j-1}}{\eta_{j-1}}}, A = \frac{\sum \frac{Y_{j-1}^2\epsilon_{j-1}^2}{\eta_{j-1}^2}}{\sum \frac{Y_{j-1}^2}{\eta_{j-1}}}.$$

To get rid of the fraction of summations, the log of A can be taken to obtain

$$\begin{aligned}
\log A &= \log \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}^2} - \log \sum \frac{1}{\eta_{j-1}} \\
\log A &= \log \sum \frac{Y_{j-1}\epsilon_{j-1}^2}{\eta_{j-1}^2} - \log \sum \frac{Y_{j-1}}{\eta_{j-1}} \\
\log A &= \log \sum \frac{Y_{j-1}^2\epsilon_{j-1}^2}{\eta_{j-1}^2} - \log \sum \frac{Y_{j-1}^2}{\eta_{j-1}}.
\end{aligned}$$

This shows that the variable A can be expressed as a function of the two other variables B and C of η_{j-1}. The system of three unknowns (A, B, C) can now be changed to a system of two equations and two unknowns (B, C) by equating any two of the expressions for A.

The system of two equations in the two unknowns, B and C, is a simpler problem to solve. After obtaining the values for B and C, any of the expressions for A can then be used to find the value of A.

Also, the expressions for A or the equations in B and C can be visualized by plotting their 3-dimensional graphs. The ranges where B and C lie can be seen from the graphs and this can help in restricting the interval for searching for the values of B and C.

Moreover, a simpler expression can be obtained for A by considering the function to minimize as a function of A, B and C. That is, for

$$M = \sum \left(\log \zeta_{j-1} + \frac{\epsilon_{j-1}^2}{\zeta_{j-1}} \right),$$

if $\zeta_{j-1} = A\eta_{j-1}$, then

$$M(A, B, C) = \sum \left\{ \log A + \log \eta_{j-1} + \frac{\epsilon_{j-1}^2}{A\eta_{j-1}} \right\}.$$

Looking at the first order condition with respect to A, the partial derivative of M with respect to A is equated to zero, which gives

$$\sum \left(\frac{1}{A} - \frac{\epsilon_{j-1}^2}{A^2\eta_{j-1}} \right) = 0.$$

Like before, the A's can be factored out of the summation to get

$$\frac{1}{A} \sum 1 = \frac{1}{A^2} \sum \frac{\epsilon_{j-1}^2}{A^2\eta_{j-1}},$$

which simplifies to $AN = \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}}$.

Therefore, the minimum value of M is obtained when

$$A = \frac{1}{N} \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}},$$

or equivalently when

$$\log A = \log \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}} - \log N.$$

This expression for A, which is also a function of B and C, but simpler than the other three expressions previously obtained, can be used in conjunction with any two of the previous three to get two equations in B and C.

So another system to solve for B and C could be comprised of any two equations from

$$\begin{aligned}
\log \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}} - \log N &= \log \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}^2} - \log \sum \frac{1}{\eta_{j-1}} \\
\log \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}} - \log N &= \log \sum \frac{Y_{j-1}\epsilon_{j-1}^2}{\eta_{j-1}^2} - \log \sum \frac{Y_{j-1}}{\eta_{j-1}} \\
\log \sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}} - \log N &= \log \sum \frac{Y_{j-1}^2\epsilon_{j-1}^2}{\eta_{j-1}^2} - \log \sum \frac{Y_{j-1}^2}{\eta_{j-1}}.
\end{aligned}$$

Also, the minimization function M can be changed to a function of two variables B and C by substituting the optimal value of A into the function as follows:

$$M(A,B,C) = \sum \left\{ \log A + \log \eta_{j-1} + \frac{\epsilon_{j-1}^2}{A\eta_{j-1}} \right\}.$$

Substituting $A = \frac{1}{N}\sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}}$ gives

$$M = \sum \left\{ \log \left(\sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}} \right) - \log N + \log \eta_{j-1} + \frac{\epsilon_{j-1}^2}{\eta_{j-1}} \frac{N}{\sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}}} \right\}.$$

This implies that

$$M = N \log \left(\sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}} \right) - N \log N + \sum \log \eta_{j-1} + N.$$

Since $-N \log N + N$ is a constant, it can be removed from the function to give

$$M_1(B,C) = N \log \left(\sum \frac{\epsilon_{j-1}^2}{\eta_{j-1}} \right) + \sum \log \eta_{j-1},$$

which is a function of two variables B and C. The minimum value of M_1 can then be found using some optimization method. The graph of M_1 can be plotted against B and C to see the behavior of the function.

Once the values for A, B and C have been obtained, they can be transformed back to get α, β and ρ. So, solving

$$A = \alpha^2(1-\rho^2), B = \frac{\rho}{\sqrt{1-\rho^2}}, C = \frac{\beta Y_{j-1}}{\alpha\sqrt{1-\rho^2}}$$

for α, β and ρ gives

$$\alpha = \sqrt{A(1+B^2)}, \beta = C\sqrt{A}, \rho = \frac{B}{\sqrt{1+B^2}},$$

which are the remaining three parameter estimates.

4.3 Simulation results and Application

The parameters used for simulations are $\kappa = 4$, $\theta = 1$,$\alpha = 0.5$,$\beta = 0.6$,$\rho = 0.7$,$\delta t = \frac{1}{12}$.

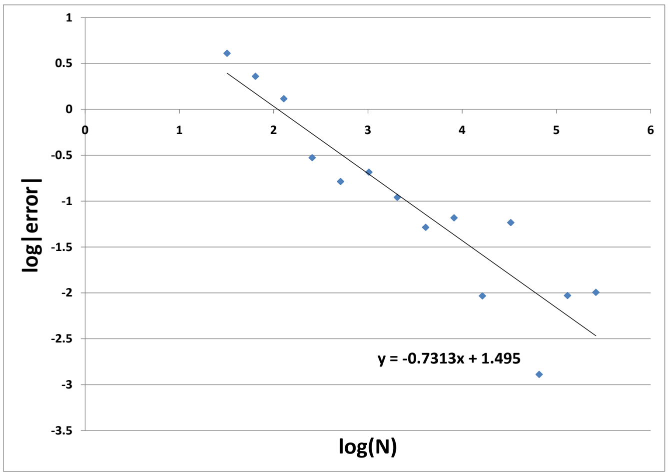

Figure 4.1: Log of error ($|Exact - Estimate|$) of κ against log of data size (N)

Sample paths were generated using these parameters. N is the amount of data generated for each sample path.
The described parameter estimation method is used to estimate the five parameters for the generated data.

The graphs below show that, as the data size (N) gets bigger, the error (difference between estimated and exact parameter values) gets smaller for all five parameters.

The same method can be used to estimate the parameters of the CAM stochastic volatility model for different stocks by treating volatility of stock returns as time series.

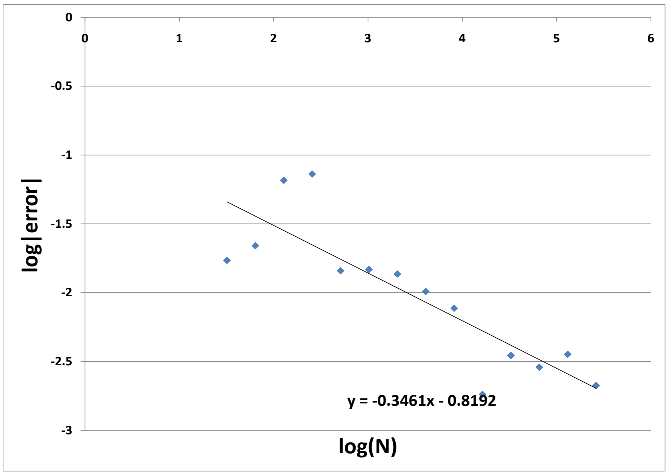

Figure 4.2: Log of error ($|Exact - Estimate|$) of θ against log of data size (N)

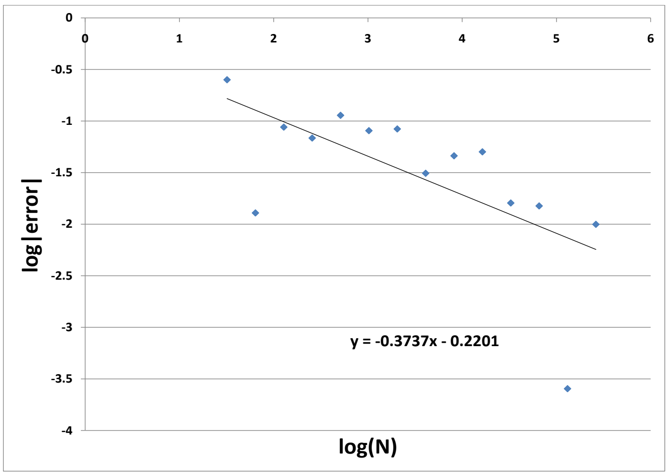

Figure 4.3: Log of error ($|Exact - Estimate|$) of α against log of data size (N)

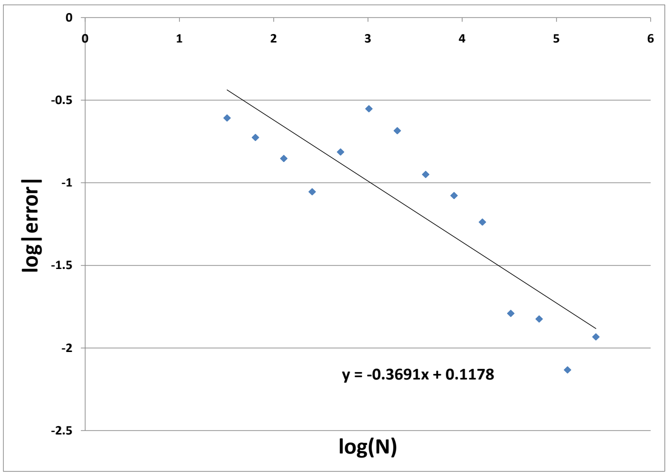

Figure 4.4: Log of error ($|Exact - Estimate|$) of β against log of data size (N)

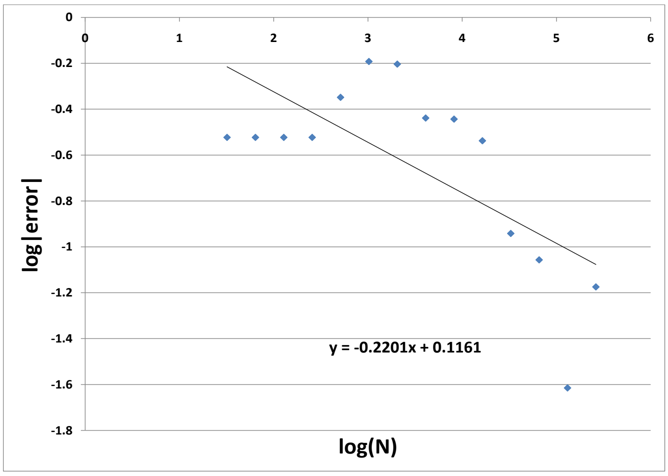

Figure 4.5: Log of error ($|Exact - Estimate|$) of ρ against log of data size (N)

Chapter 5

Conclusion

We have considered a stochastic volatility model that has additive and multiplicative noise. The model has five parameters and we have been able to estimate all of them. The first two parameters were found using the method of least squares on successive observation pairs. Then the remaining three parameters were estimated by further using the maximum likelihood method on the least squares residuals.

In the process of estimation of the last three parameters using the maximum likelihood method, we initially had three equations in three unknowns. However, by doing a change of variable, one of the three new variables could be written in terms of the other two variables such that, upon substituting into the optimization problem, we then got a function in two variables instead of the original three we had.

The function of two variables could now be viewed in a 3-dimensional graph. This would allow one to have a better guess of the point where the function is optimized, which could further help in speeding up the process of optimizing the function and thus obtaining quicker the parameters of the stochastic volatility model.