ASSET PRICES WITH REGIME SWITCHING VARIANCE GAMMA DYNAMICS

Robert Elliott

Joint work with
Andrew Royal , Haskayne School of Business, University of Calgary

Abstract

In a recent article Elliott and Osakwe discussed option pricing when the price process has dynamics described by a regime switching Lévy process. The regime switching is determined by an observable Markov chain. In this paper a related framework is considered but the regime switching chain is not observed directly. Its state and dynamics can only be estimated using some new filters. The results are tested empirically for option prices using S&P data.

The state of the economy or market will be modelled by a process, X, which is a continuous time finite state Markov chain, taking values in the set of N canonical basis vectors. More specifically, the stochastic process $X(\omega, t) : \Omega \times [0, \infty) \to \mathcal{L}$, where the set $\mathcal{L}$ is defined as

$$\mathcal{L} \equiv \{\vec{e}_1, \ldots, \vec{e}_N\} \subset \Re^N,$$

and $\vec{e}_i \in \Re^N$ has 1 in the i^{th} position and zero elsewhere. As usual, we shall write $X_t(\omega) = X(\omega, t)$, depending on context. Write $p_t^i = \bar{P}(X_t = \vec{e}_i \mid X_0)$ and $p_t = (p_t^1, p_t^2, \ldots, p_t^N)'$.

The evolution of the chain is usually described in terms of its rate matrix, (or Q matrix), A so that

$$\frac{dp_t}{dt} = Ap_t.$$

Then X has semi-martingale decomposition, (see [5]),

$$X_t(\omega) = X_0(\omega) + \int_0^t AX_s(\omega)ds + V_t.$$

Here V_t is a $(\bar{P}, \mathbf{G})$ - martingale with values in $\Re^N$, which by definition is independent of Y.

The observation process will be a map $Y(\omega, t) : \Omega \times [0, \infty) \to \Re$ which we shall suppose is a variance gamma, VG, process. This can be represented in a number of equivalent ways.

The first representation of the VG process is as a time changed Brownian motion with drift. Write $Z_t(\omega) = \theta t + \sigma B_t(\omega)$, where B_t is a Brownian motion. Here the drift is θt and the instantaneous variance σ^2. We shall write $G_t^\nu(\omega)$ for a Gamma process independent of the Brownian motion. That is,

$$G_{t+h}^\nu - G_t^\nu \overset{law}{\equiv} \gamma\left(\frac{h}{\nu}, \frac{1}{\nu}\right),$$

where $\gamma\left(\frac{h}{\nu}, \frac{1}{\nu}\right)$ is a gamma variable with mean h and variance νh. We write Y as a Brownian motion subordinated by a gamma process, i.e.,

$$Y_t \overset{law}{=} Z_{G_t^\nu}.$$

This representation has a natural economic interpretation: "packets" of information affecting security prices arrive at the market randomly according to a gamma process. At those times, the size of the returns will be given by the drifted Brownian motion, Z evaluated at the new time. The gamma process is discontinuous, and consequently the composition of the VG process is too.

The first representation makes it easy to write down the characteristic function by using a conditional expectations argument. We have

$$\begin{aligned} E\left[e^{iuY_t}\right] &= E\left[E\left[\exp\left(iuZ_{G_t^\nu}\right)|G_t^\nu\right]\right] \\ &= E\left[\exp\left(\left(iu\theta - \frac{\sigma^2 u^2}{2}\right)G_t^\nu\right)\right] \\ &= \left(1 - iu\theta\nu + \frac{\sigma^2 u^2 \nu}{2}\right)^{-\frac{t}{v}}. \end{aligned} \tag{1}$$

A second representation of a VG process is via a decomposition of the characteristic function in (1).
Write

$$\begin{aligned} \frac{1}{C} &= \nu, \\ \frac{1}{G} &= \sqrt{\frac{\theta^2\nu^2}{4} + \frac{\sigma^2\nu}{2}} - \frac{\theta\nu}{2}, \\ \frac{1}{M} &= \sqrt{\frac{\theta^2\nu^2}{4} + \frac{\sigma^2\nu}{2}} + \frac{\theta\nu}{2}. \end{aligned} \tag{2}$$

For ease of computation, we note that

$$\begin{aligned} \frac{1}{GM} &= \frac{\sigma^2\nu}{2}, \\ \text{and } \frac{1}{M} - \frac{1}{G} &= \theta\nu. \end{aligned}$$

Consequently, we may factor equation (1) as

$$\left(1 - iu\theta\nu + \frac{\sigma^2 u^2 \nu}{2}\right)^{-\frac{t}{v}} = \left(1 - \frac{iu}{M}\right)^{-Ct} \left(1 + \frac{iu}{G}\right)^{-Ct}. \qquad (3)$$

The importance of this factorisation is that we can interpret the VG process as the difference of two independent Gamma processes.

In the context of Lévy processes, the characteristic function takes on extra special significance through the following theorem.
Theorem 2.1. (Lévy-Khintchine Formula) If Y is a square integrable Lévy process, then its characteristic function may be written in the following way

$$E[e^{iuY_t}] \quad = \quad \exp\left(t\phi(u)\right)$$

where the unit log characteristic function is written

$$\phi(u) \quad = iub - \tfrac{1}{2}\sigma^2 u^2 + \int_{\Re} \left(e^{iux} - 1 - iuxI_{\{|x|\leq 1\}}\right)\nu(dx).$$

Here $\nu(dx)$ is a sigma-finite measure which satisfies $\nu(\{0\}) = 0$ and the integrability condition

$$\int_{\Re} \left(1 \wedge x^2\right)\nu(dx) < \infty$$

We also have $\sigma > 0$ and b are real numbers.

We have the following representation for the variance gamma process:

$$
\begin{aligned}
& \left(1-\frac{iu}{M}\right)^{-Ct}\left(1+\frac{iu}{G}\right)^{-Ct} \\
= \;& \exp\left(Ct\int_0^\infty (e^{iux}-1)\frac{e^{-Mx}}{x}dx\right)\exp\left(Ct\int_0^\infty (e^{-iux}-1)\frac{e^{-Gx}}{x}dx\right) \\
= \;& \exp\left(Ct\int_0^\infty (e^{iux}-1)\frac{e^{-Mx}}{x}dx\right)\exp\left(Ct\int_{-\infty}^0 (e^{iux}-1)\frac{e^{-G|x|}}{|x|}dx\right) \\
= \;& \exp\left(t\int_{\Re}\left(e^{iux}-1\right)\nu(dx)\right) \\
= \;& \exp\left(tiub+t\int_{\Re}\left(e^{iux}-1-iuxI_{\{|x|\le 1\}}\right)\nu(dx)\right)
\end{aligned}
$$

Here we have written $\nu(dy) = k(y)dy$ as the Lévy measure where k is given by

$$k(y) = \frac{C\exp(-My)}{y} I_{\{y>0\}} + \frac{C\exp(-G|y|)}{|y|} I_{\{y<0\}}. \qquad (4)$$

and the drift, b, is given by

$$\begin{aligned} b &= -C\int_{-1}^{0} e^{-G|x|}dx + C\int_{0}^{1} e^{-Mx}dx \\ &= \frac{C}{M}\left(1-e^{-M}\right) - \frac{C}{G}\left(1-e^{-G}\right). \end{aligned}$$

Thus, the Lévy triple is given by $(b, 0, \nu)$.

Definition (Random Measure) A random measure on $\Re \times \Re_+$ is a family $\mu = \{\mu(dy, dt; \omega); \omega \in \Omega\}$ of non-negative measures on the measure space $(\Re \times \Re_+, \mathbf{B}(\Re) \otimes \mathbf{B}(\Re_+))$ which satisfy $\mu(\Re \times \{0\}; \omega) = 0$. First, (dropping the ω subscript), the random (jump) measure induced by Y is defined as follows:

$$\begin{aligned} \mu(A, (0, t]) &\triangleq \sum_{0<s\leq t} I_{\{\Delta Y_s \in A, \Delta Y_s \neq 0\}} \\ &= \sum_{s\in(0,t]} I_{\{\Delta Y_s \in A\}} \delta_{(s,\Delta Y_s)}((0, t] \times A). \end{aligned}$$

Here, $A \subset \Re$ is any Borel set which does not contain an open interval surrounding the origin, thus ensuring that μ is finite. As usual, δ_a is the Dirac measure for the point $a \in \Re_+ \times \Re$.

Remark 2.6 The quantity $\mu(A,(0,t])$ is integer valued and can be intuitively thought of as the number of jumps of "size" A that the process Y makes in the time interval $[0,t]$. The integral of a suitable function $f : \Re \times \Re_+ \times \Omega \to \Re$ with respect to this jump measure is defined by

$$
\begin{aligned}
(f * \mu)_t &= \int_{(0,t]} \int_{\Re} f(y, s-)\mu(dy, ds) \\
&= \sum_{0<s\le t} f(\Delta Y_s, s I_{\{\Delta Y_s \neq 0\}}).
\end{aligned}
$$

Next, we shall define the predictable compensator, ν^P of this quantity as a random measure such that $E[(f * \nu^P)_t] = E[(f * \mu)_t]$ for all $t \geq 0$ and all f such that f is a left continuous process (i.e. f is predictable).

Then

$$\begin{aligned}
&e^{iuX_t} \\
= \;& 1 + \sum_{0<s\leq t} e^{iuX_s} - e^{iuX_{s-}} \\
= \;& 1 + \sum_{0<s\leq t} e^{iuX_{s-}} \left(e^{iu\Delta X_s} - 1 \right) \\
= \;& 1 + \int_0^t e^{iuX_{s-}} \int_{\Re} \left(e^{iux} - 1 \right) \mu(dx; ds) \\
= \;& 1 + \int_0^t e^{iuX_{s-}} \int_{\Re} \left(e^{iux} - 1 \right) \left(\mu(dx; ds) - \nu^P(dx; ds) \right) \\
+ \;& \int_0^t e^{iuX_{s-}} \int_{\Re} \left(e^{iux} - 1 \right) \nu^P(dx; ds) \\
= \;& 1 + \text{martingale} + \int_0^t e^{iuX_{s-}} \int_{\Re} \left(e^{iux} - 1 \right) \nu^P(dx; ds).
\end{aligned}$$

Taking expectations and solving, we observe that

$$\begin{aligned} E_P[e^{iuX_t}] &= 1 + \int_0^t E_P[e^{iuX_{s-}}] \int_{\Re} \left(e^{iux} - 1\right) \nu(dx, ds) \\ &= \exp\left(\int_0^t \int_{\Re} (e^{iux} - 1)\nu(dx; ds)\right). \end{aligned}$$

Equating this with the Lévy-Kintchine formulae, we see that the compensator measure $\nu(dx; ds)$ is the same as the Lévy measure. That is, the predictable compensator of the VG process Y is given by

$$\begin{aligned}\nu^P(dy; ds) &= k(y)dyds \\ &= \left(\frac{C\,\mathsf{exp}(-My)}{y} I_{\{y>0\}} + \frac{C\,\mathsf{exp}(-G|y|)}{|y|} I_{\{y<0\}}\right) dyds.\end{aligned} \tag{5}$$

VG Processes under absolutely continuous changes of measures.

At this point it is useful to introduce a truncated version of the variance gamma process such that we can ensure that the change of measure works as we wish. Instead of considering Y, we consider the tails of this process, namely, for any $\epsilon > 0$, write

$$Y_t^{\epsilon} = Y_0 + \sum_{0<s\leq t} \Delta Y_s I_{\{|\Delta Y_s| > \epsilon\}}.$$

For $A \in \mathbf{B}(\Re)$ the integer valued random measure is then defined by $\mu^{\epsilon} : \Re \times \Re_{+} \times \Omega \rightarrow \mathbf{N}$ by

$$\mu^{\epsilon}\left(A \times (0,t]\right)(\omega) = \sum_{s \in (0,t]} I_{\{\Delta Y_s^{\epsilon}(\omega) \in A\}} \delta_{(s, \Delta Y_s^{\epsilon})}((0,t] \times A).$$

The truncated VG process makes this sum almost surely finite because infinite jump activity of size less than ϵ is ignored.

Definition 3.1 (Local Absolutely Continuous Measures) P is locally absolutely continuous with respect to P', (written $P \overset{\text{loc}}{\ll} P'$), if $P|_{\mathcal{F}_t} \ll P'|_{\mathcal{F}_t}$ for all $t \geq 0$.

The importance of this truncation is that the integrability conditions of the following theorem hold.

Theorem 3.2 If, under P_i, the Lévy triple is given by $(b_i, 0, \nu_i)$ for $i = 1, 2$, then $P_1 \overset{\text{loc}}{\ll} P_2$ provided that

1. $\int_{-1}^{1} |xk_1(x) - xk_2(x)| dx < \infty$,
2. $b_1 = b_2 + \int_{-1}^{1} x(k_1(x) - k_2(x))dx$, and
3. $\int_{\Re} \left(\sqrt{k(x)} - \sqrt{k'(x)} \right)^2 dx < \infty$.

proof: See Jacod and Shiryaev, theorem IV.4.39(c).

The Reference Measure

For the rest of the talk, we shall write Y, (resp. μ,k), for the truncated VG process, Y^ϵ, (resp. for the jump measure μ^ϵ, for the compensator measure k^ϵ), to avoid the more cumbersome notation.

Suppose under $\bar{P}$, Y is a truncated VG process with parameters $C = 1, G = M = \sqrt{2}$. For $j = 1, \ldots, m$, $C_j, M_j, G_j \in \Re_+$, consider a *likelihood function* $L_j^\epsilon(y)$ defined by

$$
\begin{aligned}
L_j^\epsilon(y) &= C_j \exp\big(-(M_j - \sqrt{2})y\big) I_{\{y>\epsilon\}} \\
&+ C_j \exp\big(-(G_j - \sqrt{2})|y|\big) I_{\{y<-\epsilon\}} + I_{\{|y|\leq\epsilon\}} \\
&= \frac{k_j(y)}{k(y)} I_{\{|y|>\epsilon\}} + I_{\{|y|\leq\epsilon\}},
\end{aligned}
$$

where $k(y)$, (resp. $k_j(y) I_{|y|>\epsilon}$) is the Lévy measure of a VG process with parameters $C = 1, G = M = \sqrt{2}$, (resp. a VG process with parameters C_j, M_j, G_j).

Consider two processes on $(\Omega, \mathcal{F}, \mathbf{G}, \bar{P})$ defined by

$$\bar{U}_t = \int_{(0,t]} \sum_{j=1}^{N} < X_{s-}, \vec{e}_j > \int_{\Re} (L_j^{\epsilon}(y) - 1)\{\mu(dy; ds) - k(y)dyds\},$$

$$\bar{\Lambda}_t = 1 + \int_{(0,t]} \bar{\Lambda}_{s-} d\bar{U}_s.$$

Then $\bar{\Lambda}$ is given by the Doléans-Dade exponential, (see Jacod and Shiryaev, [7]),

$$
\begin{aligned}
\bar{\Lambda}_t &= E_t(\bar{U}) \\
&= e^{\bar{U}_t} \prod_{0<s\leq t} (1+\Delta \bar{U}_s) e^{-\Delta \bar{U}_s} \\
&= \exp\Bigg(\int_{(0,t]} \sum_{j=1}^{N} < X_{s-}, \vec{e}_j > \int_{\Re} \log(L_j^{\epsilon}(y)) \mu(dy; ds) \\
&- \int_{(0,t]} \sum_{j=1}^{N} < X_{s-}, \vec{e}_j > \int_{\Re} (L_j^{\epsilon}(y) - 1) k(y) dy ds \Bigg).
\end{aligned}
$$

We now consider the change of measure

$$\left.\frac{dP}{d\bar{P}}\right|_{\mathcal{G}_t} \equiv \bar{\Lambda}_t, \quad t \geq 0.$$

We require that $P \sim \bar{P}$, which forced the truncation of our Lévy measure from $k(y)$ to $k(y)I_{\{|y|>\epsilon\}}$. Under the measure P, (which will be referred to as the *historical measure*)

$$Y_t^\epsilon - \sum_{j=1}^{N} \int_{(0,t]} < X_{s-}, \vec{e}_j > \int_{\Re} y k_j(y) I_{\{|y|>\epsilon\}} dy ds$$

is a martingale. In other words, under P the process Y^ϵ has predictable compensator defined by the measure

$$\sum_{j=1}^{N} < X_{s-}, \vec{e}_j > k_j(y) I_{\{|y|>\epsilon\}} dy ds.$$

We wish to calculate quantities such as $E\big[H_t X_t | \mathcal{Y}_t\big]$, where H_t is a $\mathcal{G}_t$ measurable process. More specifically, we want to find estimates of processes of the form

$$H_t = H_0 + \int_0^t \alpha_s ds + \int_0^t \beta_s' dV_s + \int_0^t \int_{\Re} \xi_s \mu(dy; ds).$$

Here α, ξ and β, are predictable square integrable processes, with $\beta \in \Re^N$.

Lemma 4.2 (Conditional Bayes' Formula)

$$E\big[H_t X_t | \mathcal{Y}_t\big] \quad = \quad \frac{\bar{E}\big[\bar{\Lambda}_t H_t X_t | \mathcal{Y}_t\big]}{\bar{E}\big[\bar{\Lambda}_t | \mathcal{Y}_t\big]}.$$

proof: See Elliott et al, Theorem 3.2, page 23.
□

Write $q_t(HX) = \bar{E}\big[\bar{\Lambda}_t H_t X_t | \mathcal{Y}_t\big]$. We use Lemma 4.2 to calculate the linear Zakai equation described in the following theorem.

Theorem 4.3 (Linear Zakai Equation)

$$\begin{aligned}
& q_t(HX) \\
= \; & q_0(HX) \\
+ \; & \int_0^t \Big\{ q_s(HAX + \alpha X) + \sum_{i,j=1}^N a_{ji} < q_s\big((\beta^j - \beta^i)X\big), \vec{e}_i > (\vec{e}_j - \vec{e}_i) \Big\} \\
+ \; & \int_0^t \int_{\Re} \sum_{j=1}^N < q_{s-}(\xi_s(y)X), \vec{e}_j > L_j^{\epsilon}(y) \mu(dy; ds) \vec{e}_j \\
+ \; & \int_0^t \sum_{j=1}^N < q_{s-}(HX), \vec{e}_j > \int_{\Re} (L_j^{\epsilon}(y) - 1)\{\mu(dy; ds) - k(y) dy ds\} \vec{e}_j .
\end{aligned}$$

Parameter Estimation

Here we derive an EM algorithm for parameter updating. The parameter space for our model is given by

$$\Theta = \Big\{(a_{ij})_{i,j=1}^{N,N}, (C_j)_{j=1}^{N}, (G_j)_{j=1}^{N}, (M_j)_{j=1}^{N};$$

$$C_j, G_j, M_j, a_{ij} \geq 0,\ i \neq j,\ i, j = 1, \ldots, N, \text{and } \sum_{j=1}^{N} a_{ij} = 0,$$

$$i = 1, \ldots, N\Big\}.$$

Consider a fixed $\theta \in \Theta$. Associated with this parameter, we have a probability P_θ under which the processes X has rate matrix A, and Y is a VG process with parameters C_j, G_j, M_j when X is in state j. We wish to estimate a better $\hat{\theta} \in \Theta$. For this we use the EM algorithm: given observations $\{y_t\}_{0 \le t \le T}$ choose

$$\hat{\theta} \in argmax_{\psi \in \Theta} E_\theta \big[\ln \frac{dP_\psi}{dP_\theta} \mid \mathcal{Y}_T \big].$$

We then define $P_{\hat{\theta}}$ so that our model has parameter $\hat{\theta}$.

Firstly, we endeavor to find the Radon-Nikodym derivative which changes the drift of the state process from $\int_0^t AX_s ds$ to $\int_0^t \hat{A} X_s ds$. Consider the counting process, J_t^{ij}, $i \neq j$, which counts the number of jumps from state i to j in the interval $(0, t]$. This has representation

$$
\begin{aligned}
J_t^{ij} &= \int_0^t < X_{s-}, \vec{e}_i >< dX_s, \vec{e}_j > \qquad (6) \\
&= \int_0^t < X_{s-}, \vec{e}_i >< AX_s ds, \vec{e}_j > + \int_0^t < X_{s-}, \vec{e}_i >< dV_s, \vec{e}_j > \\
&= \int_0^t a_{ji} < X_s, \vec{e}_i > ds + V_t^{ij} \\
&= a_{ji} O_t^i + V_t^{ij}.
\end{aligned}
$$

Here, O_t^i is the occupation time of the state process in state i up to time t; also, the third equality holds because the set $\{s \in [0,t]; X_{s-} \neq X_s\}$ is a.s., finite, and has Lebesgue measure zero.
Write

$$\begin{aligned}\hat{\Lambda}_t^{ij} &= E_t\left(\left(\frac{\hat{a}_{ji}}{a_{ji}} - 1\right)V^{ij}\right) \\ &= \left(\frac{\hat{a}_{ji}}{a_{ji}}\right)^{J_t^{ij}} \exp(-(\hat{a}_{ji} - a_{ji})O_t^i).\end{aligned}$$

Then the product $\prod_{i \neq j} \hat{\Lambda}_t^{ij}$ defines the required Radon-Nikodym derivative. Write k, (resp. $\hat{k}$) for the Lévy measure with parameters $\{C_j, G_j, M_j\}$, (resp. $\{\hat{C}_j, \hat{G}_j, \hat{M}_j\}$), $\hat{L}_j^\epsilon(y) = \frac{\hat{k}(y)}{k(y)} I_{\{|y|>\epsilon\}} + I_{\{|y|\leq\epsilon\}}$, and

$$\begin{aligned} \hat{U}_t &= \int_{(0,t]} \sum_{j=1}^{N} < X_{s-}, \vec{e}_j > \int_{\Re} \big(\hat{L}_j^\epsilon(y) - 1\big)\{\mu(dy; ds) - k(y)dyds\}, \\ \hat{\Lambda}_t &= \mathcal{E}_t\big(\hat{U}\big) \prod_{i \neq j} \hat{\Lambda}_t^{ij}. \end{aligned}$$

We now define the measure $\hat{P}$ by

$$\frac{d\hat{P}}{dP} \mid_{\mathcal{G}} = \hat{\Lambda}_t.$$

Then, under $\hat{P}$, the observation process has parameter $\hat{\theta}$. Now

$$\begin{aligned}
\ln \hat{\Lambda}_T &= \ln E_T(\hat{U}) + \sum_{i \neq j} \ln \hat{\Lambda}_T^{ij} \\
&= -\big((\hat{L}^\epsilon - 1) * k\big)_T + \big((\ln \hat{L}^\epsilon) * \mu\big)_T \\
&+ \sum_{i \neq j} \left(J_T^{ij} \ln \big(\frac{\hat{a}_{ij}}{a_{ij}}\big) - (\hat{a}_{ij} - a_{ij}) O_T^i \right).
\end{aligned}$$

Thus the conditional log likelihood function looks like

$$\begin{aligned}
\ell(\hat{\theta},\theta) &= E_\theta\big[\ln\hat{\Lambda}_T \mid \mathcal{Y}_T\big] \\
&+ E_\theta\Bigg[\sum_{i\neq j}\bigg(J_T^{ij}\ln\big(\frac{\hat{a}_{ij}}{a_{ij}}\big) - (\hat{a}_{ij}-a_{ij})O_T^i\bigg) \mid \mathcal{Y}_T\Bigg] \\
&= E_\theta\Bigg[-\bigg((\hat{L}^\epsilon - 1) * k\bigg)_T + \bigg((\ln\hat{L}^\epsilon) * \mu\bigg)_T \mid \mathcal{Y}_T\Bigg] \\
&+ \sum_{i\neq j}\bigg(\mathcal{J}_T^{ij}\ln\bigg(\frac{\hat{a}_{ij}}{a_{ij}}\bigg) - (\hat{a}_{ij}-a_{ij})\mathcal{O}_T^i\bigg).
\end{aligned}$$

Here, $\mathcal{J}_T^{ij} = E_\theta[J_T^{ij} \mid \mathcal{Y}_T]$ and $\mathcal{O}_T^i = E_\theta[O_T^i \mid \mathcal{Y}_T]$. We also introduce the statistics

$$P_T^j = \int_0^T < X_{s-}, \vec{e}_j > \int_\epsilon^\infty y\mu(dy; ds) \quad (7)$$

$$Q_T^j = \int_0^T < X_{s-}, \vec{e}_j > \int_{-\infty}^{-\epsilon} |y|\mu(dy; ds), \quad (8)$$

$$R_T^j = \int_0^T < X_{s-}, \vec{e}_j > \mu(|y| > \epsilon; ds) \quad (9)$$

and the corresponding estimates $\mathcal{P}_T^j = E[P_T^j|\mathcal{Y}_T]$, $\mathcal{Q}_T^j = E[Q_T^j|\mathcal{Y}_T]$, and $\mathcal{R}_T^j = E[R_T^j|\mathcal{Y}_T]$.

Taking derivatives with respect to the parameter $\hat{\theta}$, we get the following result.

Theorem (Parameter Updates) The parameter updates given the observations, $\mathcal{Y}_T$ are given by

$$\hat{a}_{ji} \quad = \quad \frac{\mathcal{J}_T^{ij}}{\mathcal{O}_T^i}, \quad , \forall i \neq j \tag{10}$$

$$\hat{C}_j \quad = \quad \frac{\mathcal{R}_T^j}{\big(E_1(\hat{G}_j \epsilon) + E_1(\hat{M}_j \epsilon)\big) \mathcal{O}_T^j}, \tag{11}$$

$$\hat{M}_j \quad = \quad \frac{\hat{C}_j \mathcal{O}_T^j}{\mathcal{P}_T^j}, \text{ and} \tag{12}$$

$$\hat{G}_j \quad = \tfrac{\hat{C}_j \mathcal{O}_T^j}{\mathcal{Q}_T^j} \tag{13}$$

Here $E_1(x) = \int_x^\infty \frac{\exp(-s)}{s} ds$ is the exponential integral function described in Abramowitz and Stegun, [1], page 228.

The linear Zakai equation (4.1) can be transformed into a simpler stochastic differential than the one derived, and when the process $\xi = 0$ the resulting equation becomes a pathwise ordinary differential equation. For $\ell = 1, \dots, N$, write

$$\begin{aligned} q_t(HX)^\ell &= < q_t(HX), \vec{e}_\ell >, \\ U_t^\ell &= \int_{(0,t]} \int_{\Re} \Big(\frac{1}{L_\ell^\epsilon(y)} - 1\Big)\{\mu(dy; ds) - k_\ell(y)dyds\}, \\ \lambda_t^\ell &= 1 + \int_{(0,t]} \lambda_{s-}^\ell dU_s^\ell \\ &= \mathsf{E}_t(U^\ell). \end{aligned}$$

Write $\Gamma_t = diag(\lambda_t^\ell)$ and $\overline{q}_t = \Gamma_t q_t$. Then, we have the following result.

Theorem (Robust Filters.)

$$
\begin{aligned}
\bar{q}_t(HX) &= \bar{q}_0(HX) + \int_0^t \big\{\bar{q}_s(HAX + \alpha X) \\
&+ \sum_{i,j=1}^{N} \Gamma_s a_{ji}(\vec{e}_j - \vec{e}_i)\vec{e}_i^{\,\prime}\Gamma_s^{-1}\bar{q}_s((\beta^j - \beta^i)X)\big\} ds \\
&+ \int_0^t \int_{\Re} \bar{q}_{s-}(\xi(y)X)\mu(dy; ds).
\end{aligned}
$$

□

This section calculates the recursive equations for the parameter updates. As we are using VG dynamics some of the filters are new and have a different form. Set $\alpha = \xi = 0, \beta = 0 \in \Re^N$, and $H_t = H_0 = 1$. Then we obtain the following ordinary differential equation:

$$\begin{aligned} \bar{q}_t(X)^\ell &= \bar{q}_0(X)^\ell + \int_0^t \bar{q}_s(AX)^\ell ds \\ &= \bar{q}_0(X)^\ell + \int_0^t \vec{e}_\ell' \Gamma_s A \Gamma_s^{-1} \bar{q}_s(X) ds. \end{aligned}$$

Consequently,

$$\bar{q}_t(X) \quad = \quad \bar{q}_0(X) + \int_0^t \Gamma_s A \Gamma_s^{-1} \bar{q}_s(X) ds$$

To calculate the normalised probability, we need only calculate

$$E_\theta[X \mid \mathcal{Y}_T] = \frac{\bar{q}_t(X)}{\bar{q}_t(X)' 1_N}.$$

Occupation Time of the State Process.

Write $H_t = O_t^j$, for the occupation time of the hidden Markov chain in state j up to time t. Write $\mathcal{O}_t^j = E_\theta[O_t^j \mid \mathcal{Y}_t]$ for the corresponding estimate. Also, set $H_0 = 0$, and $\alpha_s = < X_{s-}, \vec{e}_j >$, $\beta = 0 \in \Re^N$, $\xi = 0$. Then

$$\begin{aligned}\bar{q}_t(O^j X)^\ell &= \int_0^t \bar{q}_s(O^j AX + < X_-, \vec{e}_j > X)^\ell ds \\ &= \int_0^t \vec{e}_\ell' \Gamma_s A \Gamma_s^{-1} \bar{q}_s(O^j X) + \vec{e}_\ell' \vec{e}_j \vec{e}_j' \bar{q}_s(O^j X) ds\end{aligned}$$

Consequently,

$$\bar{q}_t(O^j X) = \int_0^t \Gamma_s A \Gamma_s^{-1} \bar{q}_s(O^j X) + \vec{e}_j \vec{e}_j' \bar{q}_s(X) ds.$$

Cumulative Size of Positive Jumps in the Observation Process.

Number of Jumps From State i to State j.

$$\bar{q}_t(J^{ij}X) = \int_0^t \Gamma_s A \Gamma_s^{-1} \bar{q}_s(J^{ij}X) ds + \int_0^t \Gamma_s a_{ji} \vec{e}_j \vec{e}_i' \Gamma_s^{-1} \bar{q}_s(X) ds.$$

$$\bar{q}_t(P^j X) = \int_0^t \Gamma_s A \Gamma_s^{-1} \bar{q}_s(P^j X) ds + \int_0^t \bar{q}_{s-}(X) \int_\epsilon^\infty y \mu(dy; ds) \vec{e}_j.$$

Cumulative Size of Negative Jumps in the Observation Process.

$$\bar{q}_t(Q^j X) = \int_0^t \Gamma_s A \Gamma_s^{-1} \bar{q}_s(Q^j X) ds + \int_0^t \bar{q}_{s-}(X) \int_{-\infty}^{-\epsilon} y \mu(dy; ds) \vec{e}_j.$$

Number of Jumps in the Observation Process.

$$\bar{q}_t(R^j X) = \int_0^t \Gamma_s A \Gamma_s^{-1} \bar{q}_s(R^j X) ds + \int_0^t \bar{q}_{s-}(X) \mu(|y| \geq \epsilon; ds) \vec{e}_j.$$

Discretisations

State Estimates.

$$\begin{aligned} q_0(X) &= \vec{\pi} \in \Re^N, \\ q_{t_{k+1}}(X) &\approx \tilde{B}_k \tilde{A} q_{t_k}(X). \end{aligned}$$

Occupation Time.

$$\begin{aligned} q_0(O^j X) &= \vec{0} \in \Re^N, \\ q_{t_{k+1}}(O^j X) &\approx \tilde{B}_k \tilde{A} q_{t_k}(O^j X) + \tau_{k+1} \tilde{B}_k \vec{e}_j \vec{e}_j' q_{t_k}(X). \end{aligned}$$

Number of Jumps From State i to State j, $i \neq j$.

$$\begin{aligned} q_0(J^{ij} X) &= \vec{0} \in \Re^N, \\ q_{t_{k+1}}(J^{ij} X) &\approx \tilde{B}_k \tilde{A} q_{t_k}(J^{ij} X) + \tau_{k+1} a_{ji} \tilde{B}_k \vec{e}_j \vec{e}_i' q_{t_k}(X). \end{aligned}$$

Cumulative Size of Positive Jumps in the Observation Process.

$$q_0(P^j X) \quad = \vec{0} \in \Re^N,$$

$$\begin{aligned} & q_{t_{k+1}}(P^j X) \\ \approx \quad & \tilde{B}_k \tilde{A} q_{t_k}(P^j X) + \tilde{B}_k \vec{e}_j \vec{e}_j' q_{t_k}(X) \times \sum_{t_k < s \leq t_{k+1}} \Delta Y_s I_{\{\Delta Y_s > \epsilon\}}. \end{aligned}$$

Cumulative Size of Negative Jumps in the Observation Process.

$$q_0(Q^j X) \quad = \vec{0} \in \Re^N,$$

$$\begin{aligned} & q_{t_{k+1}}(Q^j X) \\ \approx \quad & \tilde{B}_k \tilde{A} q_{t_k}(Q^j X) + \tilde{B}_k \vec{e}_j \vec{e}_j' q_{t_k}(X) \times \sum_{t_k < s \leq t_{k+1}} \Delta Y_s I_{\{\Delta Y_s < -\epsilon\}}. \end{aligned}$$

Number of Jumps in the Observation Process.

$$q_0(R^j X) \quad = \vec{0} \in \Re^N,$$

$$q_{t_{k+1}}(R^j X) \quad \approx \tilde{B}_k \tilde{A} q_{t_k}(R^j X) + \tilde{B}_k \vec{e}_j \vec{e}_j' q_{t_k}(X) \times \sum_{t_k < s \leq t_{k+1}} I_{\{|\Delta Y_s| \geq \epsilon\}}.$$

Simulation

The simulation we use the discretisation where the $t_k = T_k^\epsilon$ are in fact the jump times of the observation process, Y. In this case, we have $\sum_{t_k < s \le t_{k+1}} \Delta Y_s I_{\{\Delta Y_s > \epsilon\}} = \Delta Y_{t_{k+1}} I_{\{\Delta Y_{t_{k+1}} > \epsilon\}}$ and $\sum_{t_k < s \le t_{k+1}} I_{\{|\Delta Y_s| \ge \epsilon\}} = I_{\{|\Delta Y_{t_{k+1|}} \ge \epsilon\}}$. We can then calculate the recursive equations given above, whereupon we substitute these 'structural equations' into the new parameter estimates calculated in equations (10). We draw the observations from a simulated process with parameters given by

$$A = \begin{bmatrix} -0.025 & 0.015 \\ 0.025 & -0.015 \end{bmatrix}, \quad \theta = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}, \quad \sigma = \begin{bmatrix} 0.05 \\ 0.15 \end{bmatrix},$$

and

$$\nu = \begin{bmatrix} 5.0 \\ 3.0 \end{bmatrix}.$$

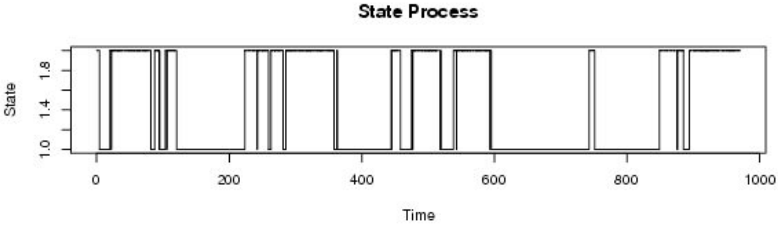
We then use an initial guess of the EM Algorithm given by

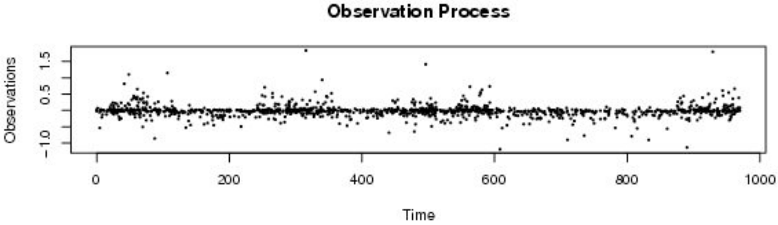
$$A^{(0)} = \begin{bmatrix} -0.2 & 0.1 \\ 0.2 & -0.1 \end{bmatrix}, \quad \theta^{(0)} = \begin{bmatrix} -0.05 \\ 0.25 \end{bmatrix}, \quad \sigma^{(0)} = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix},$$

and

$$\nu^{(0)} = \begin{bmatrix} 1.0 \\ 0.5 \end{bmatrix}.$$

The results for the state estimation are shown in figure 8.1, while the convergence of the parameters is shown in figure 8.2.


State Process
State
1.0
1.4
1.8
0
200
400
600
800
1000
Time


Observation Process
Observations
−1.0
0.5
1.5
0
200
400
600
800
1000
Time

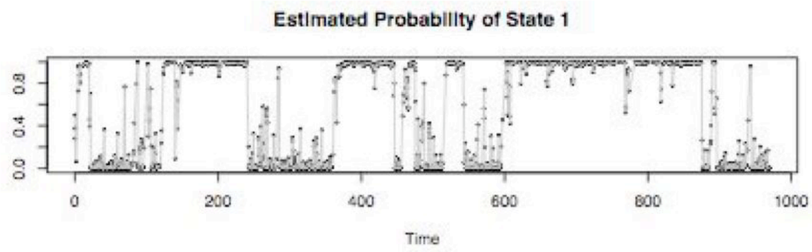


Figure: 8.2 State estimates from filtering the VG process. The upper pane shows the (hidden) state process, X_t. The middle pane indicates observations, Y_t, while the dotted grey line in the bottom pane indicates the estimated state process, $P^{(50)}(X_t = \vec{e}_1|_t) = \frac{<q_t(X_t), \vec{e}_1>}{<q_t(X_t), 1_2>}$. Note, that the estimates of the probability are using the estimated parameters after the 50^{th} pass of the EM Algorithm.

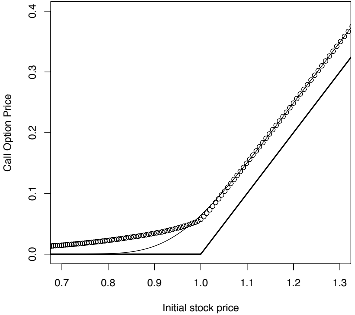


Figure: A comparison of the option prices under different assumptions about the underlying process. The dotted line is the option price when the underlying has VG regime switching returns. The thin solid line is the same but when the underlying returns have a normal distribution. Both processes have the same variance.

We compare the estimated call option price with the actual call option price for the date 29th October, 2005 and using daily data from the S&P500 found between 20th January, 1986 and the 20th January 2006. We shall use the full array of strikes that are traded on the underlying. Using the estimation scheme, we find the estimated point estimates for a two state model are given by $(A^{(50)}, \theta^{(50)}, \sigma^{(50)}, \nu^{(50)})$ where

$$A^{(50)} = \begin{bmatrix} -0.2582563 & 0.3234474 \\ 0.2582563 & -0.3234474 \end{bmatrix},$$

$$\theta^{(50)} = \begin{bmatrix} 0.14348039 \\ 0.06117406 \end{bmatrix},$$

$$\nu^{(50)} = \begin{bmatrix} 0.04702351 \\ 0.05296345 \end{bmatrix},$$

$$\text{and} \quad \sigma^{(50)} = \begin{bmatrix} 0.1992016 \\ 0.3913205 \end{bmatrix}.$$

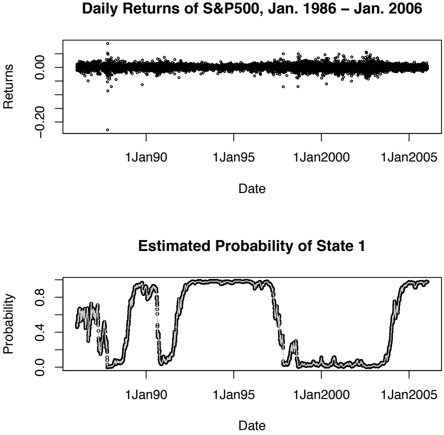


Figure: Daily Observations of the S&P500 (top plot) with filtered state estimates (bottom plot).

A diagram of the working filter for X is shown in figure (9.2). The filter clearly shows the switching occurring between two main regimes. State 1 appears to be prominent during the mid 1990s and also in the last year, while state 2 is much more in evidence during times of high volatility and market crashes. Note that our parameter estimates corresponds to a fairly reasonable (approximate) conditional yearly standard deviation of

$$\sqrt{(\theta^{(50)})^2\nu^{(50)} + (\sigma^{(50)})^2} = \begin{bmatrix} 0.2016168 \\ 0.3915737 \end{bmatrix}.$$

This shows that we have managed to decompose the time series into two states: one with high volatility (the second state), and one with low volatility. This is only a guide for yearly volatility as observations were not taken at regular intervals.

Pricing of options using this approach now becomes easy: we use these estimates given, and substitute them into our FFT framework. We compare our estimates with the actual price for options struck on the S&P500 on the date 29th October 2005.

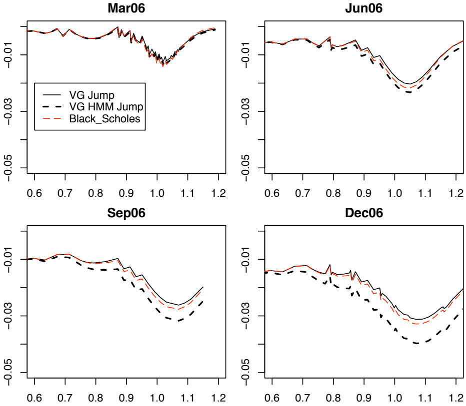


Figure: Absolute call option pricing errors for three different option pricing models.