INVERSE PROBLEMS FOR DERIVATIVE PRICING IN REGIME SWITCHING VOLATILITY MODELS (HIDDEN MARKOV MODEL FOR VOLATILITY)

RAYMOND BRUMMELHUIS, SERGE-ANDRÉ MASSON

1. Introduction

For local volatility models of option pricing (that is, models for which the instantaneous volatility is a function of the underlying asset and of time only) it has been known since Dupire [**?**] that knowledge, at some given date, of all European call options prices for all possible maturities and strikes allows one to reconstruct the local volatility function. It is natural to ask whether a similar result can be true for more general stochastic volatility models in which the volatility process has an independent component from the stochastic process which drives the stock price changes (or, if one prefers and in a Brownian motion setting, models for which volatility changes are not perfectly correlated with price changes). More generally, anticipating a possibly negative answer to this question, one can ask to what extend European call prices for all possible maturities and strikes determine the (parameters of the) volatility process. In this paper we examine this question for models in which the volatility is driven by a not directly observable finite state continuous-time Markov chain.

2. The model

⇒ Literature-review: Elliott and co-authors, others (?) - à faire

2.1. **The Regime-Switching Stochastic Vol or RSSV model.** Let $(X_t)_{t\geq 0}$ be a continuous time Markov chain with finite state space $\{1, \dots, N\}$, and risk-neutral transition probability rates q_{ij} defined by

$$\mathbb{P}(X_{t+dt} = j | X_t = i) = \delta_{ij} + q_{ij} dt \tag{1}$$

where $\mathbb{P}$ will denote the risk-neutral probability selected by the market for pricing traded assets.

Commentaires:

- Une façon plus propre de formuler ceci: $\mathbb{P}(X_{t+h} = j | X_t = i) = \delta_{ij} + q_{ij} h + o(h)$, $h \to 0$
- Attention: changement de notation par rapport à celle utilisée auparavant: q_{ij} en lieu de a_{ji}
- Faire une rémarque sur les différentes probabilités risque-neutres? Le modèle (2) avec un drift ν (plus un actif sans risque - compte en banque) n'est pas complèt.

We consider options written on a frictionlessly traded asset whose risk-neutral price dynamics is given by

$$dS_t = rS_t dt + \sigma(X_t) S_t dW_t, \tag{2}$$

where r is the constant risk-free rate and $\sigma : \{1, \dots, N\} \to \mathbb{R}$ is some given function. We suppose that the Brownian motion $(W_t)_{t\geq 0}$ is independent of the Markov chain $(X_t)_{t\geq 0}$. We note that the process $(S_t)_{t\geq 0}$ has a.s. continuous trajectories (so we don't have to write S_{t-} instead of S_t): the jumps are in the, not directly observable, instantaneous volatilities.

Remark 2.1. We could also work directly with the Markov chain $\sigma_t := \sigma(X_t)$, whose state space is $\{\sigma_1, \dots, \sigma_N\}$ where $\sigma_j = \sigma(j)\}$ and (infinitessimal) transition probabilities

$$\mathbb{P}(\sigma_{t+dt} = \sigma_j | \sigma_t = \sigma_k) = \delta_{ij} + q_{ij}dt.$$

One advantage of using a hidden Markov process X_t defined on an abstract state space is that the model then naturally extends to include for example stochastic risk-free interest rates, by specifying a second function $r : E \to \mathbb{R}$ and specifying price dynamics by

$$dS_t = r(X_t)S_t dt + \sigma(X_t)S_t dW_t.$$

Similarly, one could add state-dependend dividend rates. As before, sample paths of $(S_t)_{t\geq 0}$ are a.s. continuous: the jumps are in the (Ito-) derivatives of S_t. For now we will take r constant, and concentrate on the regime-switching stochastic volatilty model (2).

2.2. **Derivative pricing in RSSV models.** As explained in the introduction, we are interested in the inverse problem of determining the model parameters σ_i, q_{ij} from observed European call option prices for all strikes and maturities. To that effect, we start by reviewing European option pricing in our model. The problem of option pricing in regime-switching models (not only models for stochastic volatility but also for example interest rate en credit risk models) has drawn a lot attention in the mathematical finance literature, notably in papers by Robert Elliott and his co-authors: see [add string of references papers by Elliott et al] and also [other papers?]. The majority of these are concerned with the direct problem of computing option prices for a given set of parameters, though some papers also examine calibration issues: see for example Xi, Rodrigo and Mamo [1] We will review the PDE approach to pricing, which for the model (2) amounts to solving a system of PDEs for the option prices in the different Markov-chain states. We then, following an idea of [1] derive a Dupire-type equation for call-option prices as function of strike and maturity.

By risk-neutral pricing, a European derivative written on the asset S_t and paying off an amount of $F(S_T)$ at its maturity T will have a time-t price given by the discounted risk-neutral expectation

$$V_t = \mathbb{E}\left(e^{-r(T-t)}F(S_T)|\mathcal{F}_t^{S,X}\right) \tag{3}$$

where $\mathcal{F}_t^{S,X}$ is the filtration generated by the process $(S_t, X_t)_{t\geq 0}$. Since the latter is Markov, we have that $V_t = V(S_t, X_t, t)$, where

$$V(S, i, t) = \mathbb{E}\left(e^{-r(T-t)}F(S_T)|S_t = S, X_t = i\right), i = 1, \dots, N, \tag{4}$$

remembering that $X_t \in \{1, \dots, N\}$. It will be convenient to collect these N functions into a column vector

$$V(S, t) := (V(S, 1, t), \dots, V(S, N, t)^T$$

^{T} standing for "transpose", and we will consequently write $V_i(S, t)$ for $V(S, i, t)$.

The prices $V_i(S, t)$ satisfy a system of PDEs.

Theorem 2.2. *Suppose that $f = f(S, X, t)$ is a $C^{2,1}$-function*[1] *on $\mathbb{R} \times \{1, \dots, N\} \times \mathbb{R}$. Then*

$$\mathbb{E}\left(df(S_t, X_t, t)|S_t = S, X_t = i\right) \tag{5}$$
$$= \; = \left(\partial_t f(S,i,t) + rS\partial_S f(S,i,t) + \frac{1}{2}\sigma(i)^2 S^2 \partial_S^2 f(S,i,t) + \sum_{j=1}^{N} q_{ij} f(S,j,t)\right) dt.$$

Proof. Conditioning first on the Markov chain at $t + dt$ and using that the Markov chain is, by assumption, independent of the Brownian motion, we have

$$\begin{aligned}
\mathbb{E}(f(S_{t+dt}, X_{t+dt}, t)\,|S_t = S, X_t = i) &= \mathbb{E}\left(\mathbb{E}(f(S_{t+dt}, X_{t+dt}, t)\,|X_{t+dt})|S_t = S, X_t = i\right) \\
&= \mathbb{E}\left(\sum_{j=1}^{N} f(S_{t+dt}, j, t)(\delta_{ij} + q_{ij}dt)\,|S_t = S\right) \\
&= \sum_{j=1}^{N} \mathbb{E}\left(f(S + rSdt + \sigma_j S dW_t, j)\right)(\delta_{ij} + q_{ij}dt), \qquad (6)
\end{aligned}$$

since to order dt there can be at most on jump in $[t, t + dt]$. By Ito's lemma,

$$\mathbb{E}\left(f(S + rSdt + \sigma_j S dW_t, j, t)\right) = f(S, j, t) + \left(\partial_t f + rS\partial_S f + \frac{1}{2}\sigma_j^2 \partial_S^2 f\right) dt,$$

with f's derivatives all evaluated in (S, j, t). Substituting this into (6) and using that $(dt)^2 = 0$, we see that only the terms $f(S, j, t)q_{ij}dt$ and $f(S, i, t) + \left(\partial_t f + rS\partial_S f + \frac{1}{2}\sigma_j^2 \partial_S^2 f\right) dt$ remain which, after subtracting $f(S, i, t)$, proves (5). $\square$

Assuming we would know that the $V_i(S, t)$ are $C^{2,1}$ as a function of S and t, the fact that $e^{-rt}V(S_t, X_t, t)$ is a martingale, and therefore has drift 0, and theorem 2.2 applied to $e^{-rt}V_i(S, t)$ immediately implies that they must satisfy the system of PDEs

$$\partial_t V_i + \frac{1}{2}\sigma_i(S,t)^2 S^2 \partial_S^2 V_i + rS\partial_S C + \sum_i q_{ij} V_j = rV_i, \quad t < T, \tag{7}$$

with final condition $V(S, T) = F(S)$. It is possible to prove directly from (4) that the V_i's are $C^{2,1}$: see for example [?]. Alternatively, one can use the theory of linear PDEs: the system (7) with the final condition F has a unique smooth solution - reference? Friedman's book on parabolic PDE? - . By theorem 2.2, $e^{-rt}V(S_t, X_t, dt)$ is a local martingale. If F is for example bounded, then so is the solution $(V_i)_i$, which implies that the local martingale is a martingale, so that

$$e^{-rt}V(S_t, i, t) = \mathbb{E}\left(e^{-rT}F(S_T)|S_t, X_t\right),$$

and $V(S_t, X_t, t)$ is the price of the derivative.

Deux remarques:

- dernier argument à re-vérifier et à généraliser pour un call (dont le pay-off n'est pas borné)
- Le "payoff" F peut en principe être vectoriel, c'est à dire, dépendant de l'état de la chaine de Markov à T, mais admettre de tels pay-off vectoriel impliquérait que les états de la chaine de Markov sont observables, ce qui n'est pas le cas pour notre modèle, puisqu'on peut pas observer la volatilité instantannée $\sigma(X_T)$ à T

[1] two times continuously differentiable with respect to S, once with respect to t

We now specialize to European call options with (state-independent) pay-off $F(S_T) = \max(S_T - K, 0)$. We will denote the value of the call by $C(S, X, t; K, T)$ ($X \in \{1, \ldots, N\}$ and also as a column vector $C(S, t; K, T) = (C(S, 1, t; K, T), \ldots, C(S, N, t; K, T))^T$, where $C_i(S, t; K, T) = C(S, i, t; K, T)$. It will satisfy the system

$$\partial_t C + \frac{1}{2}\Sigma^2 S^2 \partial_S^2 C + rS\partial_S C + QC = rC, \tag{8}$$

where

$$\Sigma^2 = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_N^2 \end{pmatrix} \tag{9}$$

is the diagonal matrix of the state-dependend volatilities, and $Q = (q_{ij})_{1 \le i,j \le N}$ is the matrix of transition probability rates of the Markov chain. The matrix Q is row-stochastic: if $\mathbf{1} := (1, 1, \ldots, 1)^T$, then

$$Q\mathbf{1} = 0 \tag{10}$$

Since the model is time-homogeneous, we can write $C(S, t; K, T) = C(S, K, T - t)$ (with a slight abuse of notation).

The inverse problem we are interested in then is the following:

Question 1. *Suppose that at a given time t_0 we observe all call prices $C_i(S_0, t_0; K, T)$ for arbitrary strike $K > 0$ and maturity $T > t_0$, where i is the state of the Markov chain in which we are in at time t_0. How much of the model parameters N (the number of Markov states), σ_i (the state-dependent volatilities) and q_{ij} (the transition probability rates) can we reconstruct ?*

We have a total of $1 + n + n^2 - n = n^2 + 1$ parameters and a continuum of observed prices (in our idealized set-up), so the problem seems at first sight over-determined.

Question: is $2N - 2 =$ maximal size for which the matrix $\left((A^j v,^T A^k w)\right)_{j,k}$ is of full rang? Here $A = z\Sigma + Q$

2.3. Dupire's equation.

Theorem 2.3. *(Xi, Rodrigo and Mamon [1]) Fix $S = S_0, t = t_0$. Then as a function of (K, T), the vector of prices $C(S, t; K, S)$ satisfies the system of PDEs*

$$\partial_T C = \frac{1}{2}\Sigma^2 K^2 \partial_K^2 C - rK\partial_K C + QC, \quad T > 0, \tag{11}$$

with initial value $C(S_0, t_0, K, 0) = \max(S_0 - K, 0)\mathbf{1}$.

Proof. Xi *et al.* [1] first observe that C is homogeneous of order 1 in (K, S) by showing that $C(\lambda S, t; \lambda K, T)$ and $\lambda C(S, t; K, T)$ both satisfy the $N \times N$-system (8), since the (matrix-)coefficients of this system are constant. They both have the the same final value value $\lambda \max(S - K, 0)$ at T, and are therefore identical. The Euler relation

$$S\partial_S C + K\partial_K C = C$$

then allows to express derivatives with respect to S in terms of derivatives with respect to K, and (11) follows from (8). Alternatively, one can use the relation that $C(S, K, T - t_0) = S\, C(1, K/S, T - t_0)$ to derive (11). ∎

Interrogation: la question se pose si c'est vraiement nécessaire, pour notre problème inverse, d'utiliser une équation de Dupire, dans le sens qu'on peut déduire, pour ce modèle, une formule explicite pour la transformation de Fourier du prix en résolvant (8) dans l'espace de Fourier (après passage au prix logarithmique $x = \log S/S_0$) comme on le fait pour Dupire ci-bas, formule qu'on peut ensuite manipuler en tant que fonction de K ou de $\log(K/S_0)$ (S_0 étant le prix du sous-jacent au moment de l'observation) et de T. On peut peut-être pour ce modèle spécifique, en quelque sorte "court-circuiter" Dupire? À suivre.

Passing to log-coordinates $x = \log K$ and letting $c(x,T) := C(S_0, e^x, T)$ (suppressing the S_0-dependence from the notions and taking welog $t_0 = 0$) we find that

$$\partial_T c = \frac{1}{2}\Sigma^2 \partial_x^2 c - \left(\frac{1}{2}\Sigma^2 + r\right)\partial_x c + Qc, \quad T > 0 \tag{12}$$

with initial condition $c(x,0) = c_0(x) := \max(S_0 - e^x, 0)\mathbf{1}$. It is natural to solve this using the Fourier transform: if we take $r = 0$ to simplify, and if $\widehat{c}(\xi,T)$ is the Fourier transform with respect to the x-variable of the (vector-valued) call price function c, then

$$\partial_T \widehat{c} = -\left(\frac{1}{2}(\xi^2 + i\xi)\Sigma^2 - Q\right)\widehat{c},$$

with initial condition $\widehat{c}(\xi,0)$. If the initial condition would have been an integrable function, the solution is

$$c(\xi,T) = e^{-T(\frac{1}{2}(\xi^2+i\xi)\Sigma^2 - Q)}\widehat{c}(\xi,0)\mathbf{1}, \tag{13}$$

where the exponential is a matrix exponential. In our case, $c(x,0)$ is only a bounded function and thus a tempered distribution, as is its Fourier transform. To show that (13) defines a tempered distribution we have to check that $\xi \to \exp(-T(\frac{1}{2}(\xi^2 - i\xi)\Sigma^2 - Q))$ belongs to the Schwarz-class of rapidly decreasing functions. While this is not in doubt, proving it is slightly technical since we are dealing with the exponential of a sum of two non-commuting matrices, and there are no simple opper bounds we are aware off of for example $||e^{A+B}||$ en termes of $||e^A||$ and $||e^B||$ when A and B are non-commuting matrices.

Lemma 2.4. *- lemme technique: peut être sauté en première lecture -* $\xi \in \exp(-T(\frac{1}{2}(\xi^2 + i\xi)\Sigma^2 - Q))$ *is a rapidly decreasing function of* ξ *(with values in the space of* N*-dimensional matrices), and (13) is therefore well-defined as a tempered distribution, for ant tempered distribution* $c(x,0)$.

Proof. We will exploit the fact that $P(t) := e^{tQ}$ is a row-stochastic non-negative matrix, since

$$P_{ij}(t) = \mathbb{P}(X_t = j | X_0 = i),$$

and therefore $\sum_j P_{ij}(t) = 1$. If $||v||_\infty := \max_i |v_i|$ is the sup-norm on $\mathbb{C}^N$, then any non-negative row-stochastic matrix P is a contraction with respect to this norm:

$$||Pv||_\infty \leq ||v||_\infty,$$

as is easily checked[2].

Next, we recall Lie's formula (reference [?]):

$$e^{A+B} = \lim_{n\to\infty}\left(e^{A/n}e^{B/n}\right)^n,$$

which implies that $||e^{A+B}|| \leq \lim_{n\to\infty} ||e^{A/n}||^n\,||e^{B/n}||^n$ for any matrix-norm $||A||$, and in particular for $||A||_\infty = \sup_{||v||_\infty=1}||Av||_\infty$. Applying this with $A = -T(\xi^2 + i\xi)\Sigma^2$ and $B = TQ$ and using

[2]Since $P_{ij} \geq 0$, $||Pv||_\infty = \max_i |\sum_j P_{ij}v_j| \leq \max_i \sum_j P_{ij}||v||_\infty = ||v||_\infty$, since $\sum_j P_{ij} = 1$.

that $||e^{TQ/n}||_\infty \leq 1$ (in fact, equal to 1, since if P is a stochastic matrix, $P\mathbf{1} = \mathbf{1}$, which shows that $\sup_{||v||_\infty=1} ||Pv||_\infty = 1$), we find that

$$\left|\left| e^{-T(\xi^2+i\xi)\Sigma^2+TQ} \right|\right|_\infty \leq \lim_{n\to\infty} \left|\left| e^{-\frac{T}{n}(\xi^2+i\xi)\Sigma^2} \right|\right|_\infty^n$$

If $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \Lambda_N)$ is diagonal, then $||\Lambda||_\infty \leq \max_i |\lambda_i|$. Applying this to $\Lambda = \exp(-\frac{T}{2n}(\xi^2 + i\xi)\Sigma^2)$ with $\lambda_j = e^{-(T/2n)(\xi^2+i\xi)\sigma_j^2}$, we see that the right hand side equals

$$\left|\left| e^{-T(\xi^2+i\xi)\Sigma^2+TQ} \right|\right|_\infty \leq \lim_{n\to\infty} \left(e^{-\frac{T}{2n}\xi^2 \min_j \sigma_j^2} \right)^n = e^{-T\xi^2(\min_j \sigma_j^2)/2},$$

which is rapidly decreasing in ξ.

We next examine the derivatives with respect to $\xi \in \mathbb{R}$ of $e^{-T(\xi^2+i\xi)\Sigma^2+TQ}$. Again, there is no closed formula for the derivative with respect to ξ, since the matrices Σ^2 and Q do not commute. We will use the following formula [Wilcox, R. M., Exponential operators and parameter differentiation in quantum physics, J. Math. Phys (1967)]: if $A(\xi)$ is a C^1 matrix-valued function on $\mathbb{R}$, then

$$\frac{d}{d\xi} e^{tA(\xi)} = \int_0^t e^{(t-s)A(\xi)} A'(\xi) e^{sA(\xi)} ds.$$

Applying this with $t = 1$ to $A(\xi) = -\frac{1}{2}T((\xi^2 + i\xi)\Sigma^2 - Q)$ and using our estimate above for the norm of $e^{A(\xi)}$ we find

$$\left|\left| \frac{d}{d\xi} e^{A(\xi)} \right|\right|_\infty \leq C(|\xi| + 1) \int_0^1 ||e^{-sA(\xi)}||_\infty ||e^{(1-s)A(\xi)}||_\infty ds \leq C(|\xi| + 1) e^{-T\xi^2(\min_j \sigma_j^2)/2},$$

with $C = T||\Sigma^2||_\infty = T \max_j \sigma_j^2$, and where the ℓ^∞-norm can of course be replaced by any other matrix norm. Higher order derivatives cab be treated by iterating Wilcox's formula, e.g.

$$\begin{aligned} \frac{d^2}{d\xi^2} e^{tA(\xi)} &= \int_0^t e^{(t-s)A} A''(\xi) e^{sA} ds + \int_0^t \left(\int_0^{t-s} e^{(t-s-u)A} A'(\xi) e^{uA} du \right) A'(\xi) e^{sA} ds \\ &\quad + \left(\int_0^t e^{(t-s)A} A'(\xi) \int_0^s e^{(s-u)A} A'(\xi) e^{uA} du \right) ds; \end{aligned}$$

□

The fact that Σ^2 and Q will never commute, except in trivial cases[3] will be the cause of most of the technical problems in this paper, and makes the direct and inverse problem of option pricing in our hidden Markov model interesting and non-trivial, even in the simplest case of a two-state Markov chain.

The (distributional) Fourier transform of $c(x, 0) = \max(S_0 - e^x, 0)$ can be computed explicitly, and is equal to

$$\widehat{c}(\xi, 0) = iS_0^{1-i\xi} \left(\frac{1}{\xi + i0} - \frac{1}{\xi + i} \right), \tag{14}$$

where $(\xi + i0)^{-1} := \lim_{\varepsilon\to 0+}(\xi + \varepsilon)^{-1} = \mathrm{pv}(1/\xi) - i\pi\delta_0(\xi)$.

[3]if all σ_i^2 are distinct, then $[\Sigma^2, Q] = 0$ with $Q \neq 0$ implies that Q is a permutation matrix (au moins, je crois). The row sums of a permutation matrix are all equal to 1, so Q cannot be the generator of a Markov chain then. If for example $\sigma_1^2 = \sigma_2^2$, then Σ^2 can commute with non-zero generator matrices Q whose non-zero elements correspond to transitions between states 1 and 2, but these will then have no effect on the volatility S_t

Détails du calcul: $c_0(x) := \max(S_0 - e^x, 0)$ n'est pas intégrable, mais $c_\varepsilon(x) := e^{\varepsilon x}\max(S_0 - e^x, 0)$ l'est, pour tout $\varepsilon > 0$, et $c_\varepsilon \to c_0$ comme distributions temperées, et donc $\widehat{c}_\varepsilon \to \widehat{c}$. Or,

$$\begin{aligned}
\widehat{c}_\varepsilon(\xi) &= \int_{\mathbb{R}} c_0(x) e^{\varepsilon x - ix\xi} dx \\
&= \int_{-\infty}^{\log S_0} \left(S_0 e^{(\varepsilon - i\xi)x} - e^{(\varepsilon+1-i\xi)x} \right) dx \\
&= S_0 \cdot \frac{e^{(\varepsilon - i\xi)\log S_0}}{\varepsilon - i\xi} - \frac{e^{(\varepsilon+1-i\xi)\log S_0}}{\varepsilon + 1 - i\xi} \\
&= iS_0^{1+\varepsilon - i\xi} \left(\frac{1}{\xi + i\varepsilon} - \frac{1}{\xi + i(1+\varepsilon)} \right) \\
&\to S_0^{1-i\xi} \left(\frac{1}{\xi + i0} - \frac{1}{\xi + i} \right).
\end{aligned}$$

Note that away from $\xi = 0$, $\widehat{c}(\xi, 0)$ can be identified with a non-vanishing locally integrable function. It follows that if we would know all call prices $C_1(S_0, 0; K, T) = (C(S_0, 0; K, T), e_1)$, for all positive K and T, assuming without essential loss of generality that at the time of observation $t = 0$ we are in the hidden Markov state 1, then we would know the function $(c(x, T), e_1)$ and therefore its Fourier transform $(\widehat{c}(\xi, T), e_1)$ given by (13). Since $\widehat{c}(\xi, 0) \neq 0$ for all $\xi \neq 0$, this implies that we would know the function

$$\left(e^{-T(\zeta\Sigma^2 - Q)}\mathbf{1}, e_1 \right), \quad (\xi, T) \in \mathbb{R} \times \mathbb{R}_{>0}, \tag{15}$$

where we put $\zeta = \zeta(\xi) := \xi^2 + i\xi$, to simplify notations. In particular, evaluating the derivatives ∂_T^k at $T = 0$, we would know

$$\left((-\zeta\Sigma^2 + Q)^k \mathbf{1}, e_1 \right), \quad k = 0, 1, 2, \ldots \tag{16}$$

and the inverse problem we study becomes

Question 2. *How much of the matrices Σ^2 and Q can one reconstruct from knowledge of (16) (under appropriate conitions on Σ^2 and on Q)?*

We note in passing that since (16) are polynomials in ζ, if we know them for all ζ of the form $\zeta = \xi^2 - i\xi$, we know them for all $\zeta \in \mathbb{C}$ (in fact, we only need to know their values on $k+1$ different points).

2.4. Relation between (16) and observed option prices.

We can take $S_0 = 1$ without essential loss of generality. We first note that

$$(\xi^2 + i\xi)\widehat{c}(\xi, 0) = \xi(\xi + i) \cdot i \left(\frac{1}{\xi + i0} - \frac{1}{\xi + i} \right) = i((\xi + i) - \xi) = -1,$$

(which is equivalent to $(\partial_x^2 - \partial_x)c(x, 0) = \delta_0$). Therefore

$$(\xi^2 + i\xi)\widehat{c}(\xi, T) = -e^{T(-\frac{1}{2}\zeta\Sigma^2 + Q)}\mathbf{1},$$

where $\zeta := \zeta(\xi) := \xi^2 + i\xi$, and

$$(-\zeta\Sigma^2 + Q)^k \mathbf{1} = -(\xi^2 + i\xi)\frac{\partial^k}{\partial T^k}\widehat{c}(\xi, T)|_{T=0},$$

Now

$$(-\zeta\Sigma^2+Q)^k=\sum_{j=0}^{k}P_j^{(k)}(\Sigma^2,Q)(-1)^j\zeta^j,$$

where $P_j^{(k)}(\Sigma^2,Q)$ s a polynomial in the nonommuting (!) variables Σ^2 and Q: for example, $P_0^{(k)}(\Sigma^2,Q)=Q^k$, while

$$P_1^{(k)}(\Sigma^2,Q)=Q^{k-1}\Sigma^2+Q^{k-2}\Sigma^2Q+\cdots+\Sigma^2Q^{k-1}.$$

Note that since $Q\mathbf{1}=0$,

$$\left(P_1^{(k)}(\Sigma^2,Q)\mathbf{1},e_1\right)=(Q^{k-1}\Sigma^2\mathbf{1},e_1),$$

that is, only the first term survives when looking at the for us relevant matrix element.

Remembering that $\zeta=\xi^2+i\xi$, we therefore have

$$\begin{aligned}P_j^{(k)}(\Sigma^2,Q) &= \frac{(-1)^j}{(2j)!}\partial_\xi^{2j}(-\zeta\Sigma^2+Q)^k|_{\xi=0}\\ &= -\frac{(-1)^j}{(2j)!}\partial_\xi^{2j}\partial_T^k\left((\xi^2+i\xi)\widehat{c}(\xi,T)\right)|_{\xi=0,T=0}\end{aligned}$$

Now

$$-\partial_\xi^{2j}(\xi^2+i\xi)\widehat{c}(\xi,T)=\mathcal{F}_{x\to\xi}\left((ix)^{2j}(\partial_x^2-\partial_x)c(x,T)\right),$$

so that we find that

$$p_{k,j}:=\left(P_j^{(k)}(\Sigma^2,Q)\mathbf{1},e_1\right)=\int_{\mathbb{R}}x^{2j}(\partial_x^2-\partial_x)\partial_T^kc_1(x,T)\big|_{T=0}dx \tag{17}$$

Transforming variables back to $K=e^x$, we can also write this as

$$p_{k,j}=\int_0^\infty(\log K)^{2j}\left(K^2\partial_K^2\partial_T^kC\right)(1;0,K,T)\,\frac{dK}{K}\quad(?) \tag{18}$$

which can obviously be determined from observed option prices.

3. **Inverse option pricing problem for a two-state Markov chain**

In this subsection we show that if $N=2$, then (17) with $k=1,2,3$ uniquely determine the σ_k^2 and the q_{ij}. First of all, since the row-sums of Q are 0 ($Q\mathbf{1}=0$), we have

$$Q=\begin{pmatrix}-q_{12} & q_{12}\\ q_{21} & -q_{21}\end{pmatrix}$$

so Q is determined by the two parameters q_{12} and q_{21} which, together with the two volatilies (squared) σ_1^2 and σ_2^2 makes for a total of 4 parameters to be determined. Next we look at $(-\zeta\Sigma^2+Q)^k$ which we expand for $k=1,2$ and 3: we put $\Sigma^2=V$, to simplify the appearance of the formulas and avoid confusion with the powers of Σ which occur (which are forcibly even). For $k=1$ there is nothing to do, while for the other k's we find

$$\begin{aligned}(-\zeta V+Q)^2 &= \zeta^2V^2-(VQ+QV)\zeta+Q^2\\ (-\zeta V+Q)^3 &= -\zeta^3V^3+\zeta^2(V^2Q+VQV+QV^2)\\ &\quad-\zeta(Q^2V+QVQ+VQ^2)+Q^3\end{aligned}$$

When applying this to the vector $\mathbf{1}$, all terms starting with a Q on the left will be 0, so (17) will give us

$$p_{1,1} = (V\mathbf{1}, e_1) = \sigma_1^2,\ p_{2,2} = (V^2\mathbf{1}, e_1) = \sigma_1^4,\ p_{3,3} = (V^2\mathbf{1}, e_1) = \sigma_1^6,$$

which are all dependent. We already observed that $p_{1,0} = (Q\mathbf{1}, e_1) = 0$, and similarly for the other $p_{k,0}$. Next,

$$p_{2,1} = (QV\mathbf{1}, e_1),\ p_{3,2} = ((VQV + QV^2)\mathbf{1}, e_1) = \sigma_1^2 (QV\mathbf{1}, e_1) + (QV^2\mathbf{1}, e_1), \tag{19}$$

so that the last equation translates into

$$(QV^2\mathbf{1}, e_1) = p_{3,2} - \sigma_1^2 p_{2,1} = p_{3,2} - p_{1,1}p_{2,1}, \tag{20}$$

while finally

$$p_{3,1} = (Q^2V\mathbf{1}, e_1). \tag{21}$$

Computing $QV\mathbf{1} = Q(\sigma_1^2, \sigma_2^2)^t$ and $QV^2\mathbf{1}$, the first equation of (19) and (20) give

$$q_{12}\,(\sigma_2^2 - \sigma_1^2) = p_{2,1}$$
$$q_{12}\,(\sigma_2^4 - \sigma_1^2) = p_{3,2} - p_{1,1}p_{2,1}.$$

Dividing the second equation by the first, we find $\sigma_2^2 + \sigma_1^2 = (p_{3,2} - p_{1,1}p_{2,1})/p_{2,1} = (p_{3,2}/p_{2,1}) - p_{1,1}$, so that

$$\sigma_2^2 = \frac{p_{3,2}}{p_{2,1}} - 2p_{1,1}.$$

The first equation then yields q_{12}:

$$q_{12} = \frac{p_{2,1}}{\sigma_2^2 - \sigma_1^2} = \frac{(p_{2,1})^2}{p_{3,2} - 3p_{1,1}p_{2,1}}.$$

Finally (21) translates into

$$(\sigma_1^2 - \sigma_2^2)(q_{12}^2 + q_{12}q_{21}) = p_{3,1},$$

with solution

$$q_{21} = \frac{p_{3,1} + (\sigma_2^2 - \sigma_1^2)q_{12}^2}{q_{12}(\sigma_1^2 - \sigma_2^2)} = \frac{p_{3,1}}{q_{12}(\sigma_1^2 - \sigma_2^2)} - q_{12} = -\frac{p_{3,1}}{p_{2,1}} - q_{12}.$$

Calculs à vérifier encore; formuler tout ceci comme théorème;

Theorem 3.1. *Suppose all call-prices $C(S_0 = 1, 0; K, T)$ are known. If*

$$p_{k,j} = \int_0^\infty (\log K)^{2j} \left(K^2 \partial_K^2 \partial_T^k C\right)(1; 0, K, T)\, \frac{dK}{K}, \tag{22}$$

then if $N = 2$, the model parameters are given by ...

References

[1] X. Xi, M. Rodrigo and R.S. Mamon, 2012, Parameter estimation of a regime-switching model using an inverse Stieltjes moment approach?, In: *Stochastic Processes, Finance and Control (Festschrift in Honour of Robert Elliott's 70th Birthday)*, Advances in Statistics, Probability and Actuarial Science, Volume I , (eds.: Cohen, S., Madan, D., Siu, T. and Yang, H.), World Scientific, 549-569