Statistical inference of semi-Markov process and application in Finance

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Sanket Nandan

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2015

Supervisor: Dr. Anindya Goswami

This is to certify that this dissertation entitled Statistical inference of semi-Markov process and application in Finance towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents the research carried out by Sanket Nandan at Indian Institute of Science Education and Research under the supervision of Dr. Anindya Goswami, Assistant Professor, Dept. of Mathematics , during the academic year 2014-2015.

Dr. Anindya Goswami

Committee:

Dr. Anindya Goswami

Prof. M. K. Ghosh (IISc, Bangalore)

This thesis is dedicated to my beloved family and my grandma.

Declaration

I hereby declare that the matter embodied in the report entitled Statistical inference of semi-Markov process and application in Finance are the results of the investigations carried out by me at the Dept. of Mathematics, Name of the Institute, under the supervision of Dr. Anindya Goswami and the same has not been submitted elsewhere for any other degree.

Sanket Nandan

Acknowledgments

I would like to thank all the people who contributed and encouraged in some way to the work described in this thesis. First and foremost, I would like to thank my guide *Dr. Anindya Goswami*, for engaging me in new ideas and showing me several ways of assimilating mathematical ideas. Additionally, I would like to thank my committee member *Prof. Mrinal K. Ghosh* for the interest and support in my work.

I would also like to thank my parents, elder uncle, my cousin brothers, my not so studious friends and Prof. P. K. Panigrahi for endless guidance and support. Special thanks to Mayur Meena. And Pink Floyd.

Abstract

In the literature of derivative pricing, for a real life application one must know the values of the co-efficients which appear in the non-local parabolic PDE of the initial-boundary value problem that arises in the study of derivative pricing in a semi-Markov modulated market model. The co-efficients that are involved as parameters in the PDE should be estimated from the market data in a semi-Markov modulated GBM (geometric Brownian motion) model. There exists one functional parameter, known as hazard rate function and this current thesis studied the maximum likelihood estimation (MLE) of transition rate of a semi-Markov process, and thereby the MLE of the corresponding hazard rate for a given transition matrix. The study of the convergence of MLE of transition rate through asymptotic behaviour of the estimator has been done, that can be extended to the convergence of MLE of hazard rate for our market model with given transition matrix. In this thesis we did perform some numerical experiments to illustrate the convergence of MLE of hazard rate function. Finally, by looking towards the application in Finance, it has been shown that the solution of the approximated price equation of European call option (through maximum likelihood estimation of hazard rate) converges to the true solution.

Contents

Chapter 1

Introduction

Purpose of the project: This current project aims to survey recent development on certain problems in Mathematical Finance and contribute something new. Following the seminal work by Black, Scholes and Merton on pricing options in 1973, several alternative models are still being proposed in the literature and thereby new mathematical challenges are arising. In recent years researchers are interested to replace the constant market parameters by semi-Markov chains which evolve according to some prescribed transition rates[8]. It is also shown in the literature that by considering such kind of general model, many drawbacks of the original Black-Scholes-Merton model can be fixed. While fixing the deficits, it also retains the mathematical tractability. More precisely, the price function still solves a parabolic PDE which can be computed numerically. Despite that, for a real life application one must know the values of the coefficients which appear in the equation. Even for Black-Scholes-Merton model this is a challenge. For semi-Markov modulated GBM model, the PDE involves a few more parameters which should be estimated from the market data.

In this thesis, we put our concern on the parameter, called hazard rate function, which needs to be studied for the purpose of solving or drawing some inferences about the price function in a semi-Makov modulated GBM model. This current thesis studied the MLE of transition rate of a semi-Markov process, and thereby the MLE of the corresponding hazard rate for a given transition matrix. We also studied of the convergence of MLE of transition rate through asymptotic behaviour of the estimator, that can be extended to the convergence of MLE of hazard rate for our market model, given the transition matrix is known. We also performed some numerical experiments to illustrate the convergence of MLE of hazard rate function. Finally, it has been shown that the solution of the approximated price equation of European call option (through maximum likelihood estimation of hazard rate) converges to the true solution.

Techniques: This present thesis studied the techniques of statistical inference of semi-Markov processes. It did scrutiny the applicability of the nonparametric estimation for semi-Markov processes based on its hazard rate functions[16]. This work has required knowledge of real analysis, stochastic calculus, ordinary and partial differential equation, numerical analysis, quantitative finance, maximum likelihood estimation and FORTRAN95 (programming language).

Outcomes: We have studied the convergence of the estimated hazard rate through maximum likelihood estimation in a semi-Markov modulated market model. The estimated hazard rate has been compared with the theoretical hazard rate function for a typical semi-Markov modulated market, as shown in graph [see chapter 4].

As application towards mathematical Finance, the approximated price function is defined with the dependence of estimated hazard rate function as it appears in the price equation as a parametric co-efficient. Thereafter, we studied the convergence of HREB approximation of the European call option price function to the true value of the price. Later, we have indicated a numerical method to actually compute the approximated price function.

Chapter 2

Preliminaries

2.1 Markov semigroup and its generator

Definition 2.1.1. *(Semigroup) A semigroup of operators $\{S(t)\}_{t\geq 0}$ is a family of bounded linear operators on a Banach space V such that*
(i) $S_0 f = f \quad \forall f \in V$, and
(ii) $S_{t+s} = S_t \circ S_s \quad \forall t, s \geq 0$.

Given a semi-group of operators, $\{S_t\}_{t\geq 0}$, one may associate a generator in the following way.

Define,

$$\mathcal{D} := \{f \in V | \lim_{t \downarrow 0} \frac{S_t f - f}{t} \text{ exists}\}$$

and

$$\mathcal{A} f := \lim_{t \downarrow 0} \frac{S_t f - f}{t} \quad \forall f \in \mathcal{D}. \tag{2.1.1}$$

Then, $\mathcal{A}$ is called infinitesimal generator and $\mathcal{D}$ is the domain.

Given a Markov chain X_t we define,

$$S_t f(x) = E(f(X_t)|X_0 = x) \qquad \forall f \text{ continuous and bounded.} \tag{2.1.2}$$

Thus, one can always associate a semi-group like this.

Theorem 2.1.1. $\{S_t\}_{t\geq 0}$ is a semi-group.

Proof. We write,

$$\begin{aligned} S_{t+s}f(x) &= E(f(X_{t+s})|X_0 = x) \\ &= E(E[f(X_{t+s})|X_t]|X_0 = x) \\ &= E(S_s f(X_t)|X_0 = x) \\ &= S_t(S_s f)(x) \\ &= S_t \circ S_s f(x). \end{aligned}$$

And it is easy to check, $S_0 = \mathbb{I}$. Hence $\{S_t\}_{t\geq 0}$ is a semi-group. □

Let $\mathcal{L}$ be an operator such that,

$$f(X_t) = f(X_0) + \int_0^t \mathcal{L}f(X_s)ds + M_t^f \qquad \forall f \in \mathrm{Dom}(\mathcal{L}),$$

where M^f is a Martingale. We see,

$$\begin{aligned} (S_\delta - I)f &= E[f(X_\delta)|X_0 = x] - f(x) \\ &= E[f(X_\delta) - f(x)|X_0 = x] \\ &= E[\int_0^\delta \mathcal{L}f(X_s)ds + M_\delta|X_0 = x] \\ &= E[\int_0^\delta \mathcal{L}f(X_s)ds|X_0 = x]. \end{aligned}$$

So,

$$\lim_{\delta\mapsto 0} \frac{(S_\delta - I)f}{\delta} = \lim_{\delta\mapsto 0} \frac{1}{\delta} E[\int_0^\delta \mathcal{L}f(X_s)ds|X_0 = x].$$

If f is such that on the RHS the limit and expectation can change the order, we get

$$\mathcal{A}f(x) = \mathcal{L}f(x),$$

i.e., $\mathcal{L}$ is the generator.

2.2 Semi-Markov process and its augmentation

2.2.1 Semi-Markov process:

The process $\{X_t\}_{t\geq 0}$ is a semi-Markov process on the statespace $S = \{1, 2, 3, \ldots, k\}$ if the following holds:

(i) $\{X_{T_n}\}_{n\geq 0}$ is a Markov chain with the transition matrix (p_{ij}) having $p_{ii} = 0\ \forall i$, where $\{T_n\}$ are transition times with $T_0 = 0$ a.s.

(ii)$P(X_{T_n} = j, T_n - T_{n-1} \leq y | X_{T_{n-1}} = i, \ldots, X_0) = p_{ij}F(y|i)\ \ \forall j \neq i$ while $F(\cdot|i)$ is a c.d.f. for each i.

The two independent quantities for the description of a semi-Markov process can be interpreted as follows:

(i) p_{ij} gives the probability of going to state j in the next transition, given the current state is i,

$$P(X_{T_n} = j | X_{T_{n-1}} = i) = P(X_{T_n} = j, T_n - T_{n-1} < \infty | X_{T_{n-1}} = i) = \lim_{y\to\infty} p_{ij}F(y|i) = p_{ij}.$$

(ii) $F(\cdot|i)$ denotes the conditional distribution of holding time,

$$\begin{aligned} P(T_n - T_{n-1} \leq y | X_{T_{n-1}} = i) &= P(X_{T_n} \in S, T_n - T_{n-1} \leq y | X_{T_{n-1}} = i) \\ &= \sum_{j\in S} P(X_{T_n} = j, T_n - T_{n-1} \leq y | X_{T_{n-1}} = i) = F(y|i). \end{aligned}$$

2.2.2 Augmented Markov process (X_t, Y_t):

In our discussion we consider a particular class of semi-Markov processes which satisfy the following conditions,

(i) $F(y|i)$ is twice differentiable $\forall i$ and $f(y|i) := \frac{d}{dy}F(y|i)$ is bounded,

(ii) $f(y|i) > 0$ $\forall i$ & $y > 0$.

We embed S in $\mathbb{R}^k$ through $i \sim e_i \in \mathbb{R}^k$ and define for $y \in [0, \infty)$

$$\lambda_{ij}(y) := p_{ij}\frac{f(y|i)}{1 - F(y|i)}, \qquad \forall j \neq i \in S \tag{2.2.1}$$

and

$$\lambda_{ii}(y) := -\sum_{j \in S, j \neq i} \lambda_{ij}(y) \qquad \forall i \in S. \tag{2.2.2}$$

Define, $\Lambda_{ij}(y)$ as a left closed and right open interval of the real line for a particular $y \in \mathbb{R}^+ \forall j \neq i \in S$ with the intervals having lengths $\lambda_{ij}(y)$ and arranged consecutively according a lexicographic order starting from the origin. Define functions $h : S \times \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^k$ and $g : S \times \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^+$ as

$$h(i, y, z) := \sum_{j \in S, j \neq i} (j - i) 1_{\Lambda_{ij}(y)}(z)$$

and

$$g(i, y, z) := \sum_{j \in S, j \neq i} y 1_{\Lambda_{ij}(y)}(z).$$

Theorem 2.2.1. Let us consider a set of stochastic integral equations, given by,

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}} h(X_{u-}, Y_{u-}, z)\wp(du, dz) \tag{2.2.3}$$

$$Y_t = t - \int_0^t \int_{\mathbb{R}} g(X_{u-}, Y_{u-}, z)\wp(du, dz), \tag{2.2.4}$$

where $\wp$ is Poisson random measure with intensity as Lebesgue measure $dudz$ and

h and g are mentioned above. The equations (2.2.3) and (2.2.4) has a strong solution (X_t, Y_t), which is a Markov process and $\{X_t\}_{t\geq 0}$ is a semi-Markov process with transition matrix (p_{ij}) and holding time c.d.f. $F(\cdot|i)$.

Proof. The proof of existence of a strong solution is standard, and we omit. To show the Markovity of (X_t, Y_t), we need to check that the dependence of future (X_T, Y_T) of the processes on past is only with the present state (X_t, Y_t).

$$\begin{aligned} X_T &= X_0 + \int_0^T \int_{\mathbb{R}} h(X_{u-}, Y_{u-}, z)\wp(du, dz) \\ &= X_0 + \int_0^t \int_{\mathbb{R}} h(X_{u-}, Y_{u-}, z)\wp(du, dz) + \int_t^T \int_{\mathbb{R}} h(X_{u-}, Y_{u-}, z)\wp(du, dz) \\ &= X_t + \int_t^T \int_{\mathbb{R}} h(X_{u-}, Y_{u-}, z)\wp(du, dz). \end{aligned}$$

$$\begin{aligned} Y_T &= T - \int_0^T \int_{\mathbb{R}} g(X_{u-}, Y_{u-}, z)\wp(du, dz) \\ &= (T-t) + t - \int_0^t \int_{\mathbb{R}} g(X_{u-}, Y_{u-}, z)\wp(du, dz) + \int_t^T \int_{\mathbb{R}} g(X_{u-}, Y_{u-}, z)\wp(du, dz) \\ &= (T-t) + Y_t + \int_t^T \int_{\mathbb{R}} g(X_{u-}, Y_{u-}, z)\wp(du, dz). \end{aligned}$$

Thus, we prove that (X_t, Y_t) is a Markov process.

From the equations one can easily derive to get the following interpretations, which ensure the integrals to represent the augmented Markov process, viz.,

(1) Considering $\tau_n = T_n - T_{n-1}$ we will first prove $P(\tau_{n+1} \leq y|\mathcal{F}_{T_n}, X_{T_n} = i) = F(y|i)$.

From (2.2.1), our consideration of f to be the derivative of F and knowing

$\sum_{j\neq i\in S} p_{ij} = 1$, we get,

$$\begin{aligned}
\frac{dF(s|i)}{ds} = f(s|i) &= \sum_{j\in S} p_{ij} f(s|i) \\
&= (1 - F(s|i)) \sum_{j\neq i\in S} \lambda_{ij}(s).
\end{aligned}$$

Since $F(0|i) = 0$, thus,

$$\begin{aligned}
\int_0^y \frac{dF(s|i)}{1 - F(s|i)} &= \int_0^y \sum_{j\neq i\in S} \lambda_{ij}(s) ds \\
\Rightarrow -\ln(1 - F(y|i)) &= \int_0^y \sum_{j\neq i\in S} \lambda_{ij}(s) ds.
\end{aligned}$$

So,

$$\begin{aligned}
F(y|i) &= 1 - \exp(-\int_0^y \sum_{j\neq i\in S} \lambda_{ij}(s) ds) \\
&= 1 - P(\wp(\bigcup_{0<s\leq y} (\{T_n + s\} \times \bigcup_{j\neq i} \Lambda_{ij}(s))) = 0) \\
&= 1 - P(\text{no jump in } (T_n, T_n + y] | \mathcal{F}_{T_n}, X_{T_n} = i) \\
&= P(\tau_{n+1} \leq y | \mathcal{F}_{T_n}, X_{T_n} = i).
\end{aligned}$$

(2) Secondly we will prove $P(X_{T_{n+1}} = j | \mathcal{F}_{T_{n+1}-}, T_{n+1}) = p_{X_{T_n} j}$.

Owing to the property of Markovity of (X_t, Y_t), we see,

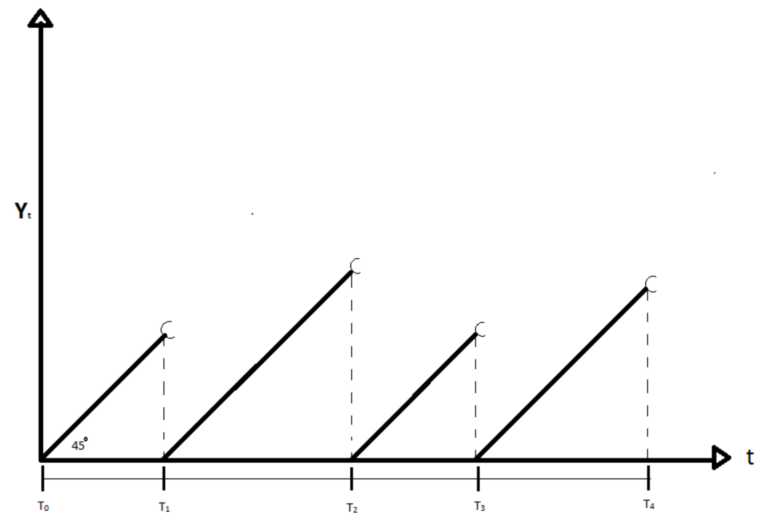
$$
\begin{aligned}
P(X_{T_{n+1}} = j|\mathcal{F}_{T_{n+1}-}, T_{n+1}) &= P(X_{T_{n+1}} = j|X_{T_{n+1}-} = X_{T_n}, Y_{T_{n+1}-} = T_{n+1} - T_n) \\
&= P(\int_{\mathbb{R}} h(X_{T_n}, T_{n+1} - T_n, z)\wp(\{T_{n+1}\} \times dz) = j - X_{T_n}| \\
&\qquad \int_{\mathbb{R}} h(X_{T_n}, T_{n+1} - T_n, z)\wp(\{T_{n+1}\} \times dz) \neq 0) \\
&= P(\wp(\{T_{n+1}\} \times \Lambda_{X_{T_n}j}(T_{n+1})) \neq 0|\wp(\{T_{n+1}\} \times \Lambda_{X_{T_n}l}(T_{n+1})) \neq 0 \\
&\qquad \text{for some } l) \\
&= \frac{|\Lambda_{X_{T_n}j}(T_{n+1})|}{|\bigcup_{l \neq X_{T_n}} \Lambda_{X_{T_n}l}(T_{n+1})|} \\
&\qquad [\ \wp \text{ has uniform distribution}] \\
&= \frac{\lambda_{X_{T_n}j}(T_{n+1})}{\sum_{l \neq X_{T_n}} \lambda_{X_{T_n}l}(T_{n+1})} \\
&= p_{X_{T_n}j}.
\end{aligned}
$$

Now, we proceed to look at the semi-Markov kernel and prove $P(X_{T_{n+1}} = j, \tau_{n+1} \leq y|\mathcal{F}_{T_n}) = p_{X_{T_n}j}F(y|X_{T_n})$.

$$
\begin{aligned}
P(X_{T_{n+1}} = j, \tau_{n+1} \leq y|\mathcal{F}_{T_n}) &= E[P(X_{T_{n+1}} = j, \tau_{n+1} \leq y|\mathcal{F}_{T_{n+1}-}, T_{n+1})|\mathcal{F}_{T_n}] \\
&= E[p_{X_{T_n}j}1_{\tau_{n+1} \leq y}|\mathcal{F}_{T_n}] \\
&= p_{X_{T_n}j}P(\tau_{n+1} \leq y|\mathcal{F}_{T_n}) \\
&= p_{X_{T_n}j}F(y|X_{T_n}). \quad \square
\end{aligned}
$$

In view of the previous theorem, given a semi-Markov process X_t, we can augment that with the holding time process Y_t to obtain a Markov process (X_t, Y_t) and we call that the augmented Markov process.

To illustrate the behaviour of the process Y_t, we plot a typical realization below.



2.3 Derivation of generator of augmented Markov process

Now our aim is to find the infinitesimal generator of the augmented Markov process on the domain of smooth functions.

Using Itô's formula for function of RCLL (right continuous with left limit) processes(Theorem II.31 [18]), we calculate the infinitesimal change of some function of the augmented Markov process $\phi(X_t, Y_t)$ as

$$\begin{aligned} d\phi(X_t,Y_t) &= \frac{\partial \phi}{\partial y}(X_t,Y_t)dY_t^c + \phi(X_t,Y_t) - \phi(X_{t^-},Y_{t^-}) \\ &= \frac{\partial \phi}{\partial y}(X_t,Y_t)dt + \int_{\mathbb{R}} [\phi(X_{t^-} + h(X_{t^-},Y_{t^-},z), Y_{t^-} - g(X_{t^-},Y_{t^-},z)) - \phi(X_{t^-},Y_{t^-})] \\ &\qquad (\hat{\wp}(dt,dz) + dtdz). \end{aligned}$$

Note that: $\hat{\wp}$ is compensated Poisson random measure with expectation zero. We denote the process obtained by integrating w.r.t. $\hat{\wp}$ as $\{M_t\}_{t\geq 0}$. Clearly M_t is a martingale and we obtain

$$\begin{aligned} d\phi(X_t, Y_t) &= \frac{\partial \phi}{\partial y}(X_t, Y_t)dt + \sum_{j \neq X_{t-}} [\phi(j,0) - \phi(X_{t^-}, Y_{t^-})]\lambda_{X_{t-}j}(Y_{t^-})dt + dM_t \\ &= \frac{\partial \phi}{\partial y}(X_t, Y_t)dt + \frac{f(Y_{t^-}|X_{t^-})}{1 - F(Y_{t^-}|X_{t^-})} \sum_{j \neq X_{t-}} p_{X_{t-}j}[\phi(j,0) - \phi(X_{t^-}, Y_{t^-})]dt + dM_t. \end{aligned}$$

Thus the infinitesimal generator $\mathcal{L}$ of the augmented Markov process, restricted in the class of $\mathcal{C}^1$ function (in y) is given by

$$\mathcal{L}\phi(i,y) = \frac{\partial \phi}{\partial y}(i,y) + \frac{f(y|i)}{1 - F(y|i)} \sum_{j \neq i} p_{ij}[\phi(j,0) - \phi(i,y)]. \tag{2.3.1}$$

2.4 Application in Finance

The operator, $\mathcal{L}$ obtained in appears in several differential equations in the field of mathematical finance[8, 12]. In the next chapter we provide a short description of one such equation which arises in derivative pricing.

Chapter 3

Price equation in a semi-Markov modulated market

In the literature of derivative pricing, we often encounter incomplete market, where to hedge for a contingent claim, no class of self-financing hedging strategy is adequate to replicate the given legitimate claim at maturity. Although, there are many approaches to deal with pricing problem through formulation of price. One of them is the local risk minimization approach mentioned by Föllmer and Sondermann[7]; Schweizer[20]; Schweizer[21]; Schweizer[22]; Schweizer[23]. It says that for hedging a claim, one must adopt a dynamic strategy through dynamic (varies with time) allocation to the assets and it must replicate the given claim through accumulation of additional cash flow in a continuous trading. Here, optimal hedging is done by minimizing the quadratic residual risk (QRR), a particular measure of the additional cash flow, under certain set of constraints. Föllmer and Schweizer[6] showed that existence of an optimal hedging is equivalent to the existence of Föllmer Schweizer decomposition of discounted claim in a particular form, for a arbitrage free market. For some typical market models, it is quite possible to derive and solve a system of differential equations (price function associated) for obtaining price and optimal hedging of a claim.

We consider regime switching market model by assuming the market parameters to vary according to a stochastic process with finite state-space, with time. We are more concerned in a semi-Markov regime switching market model, where the parameters follow a semi-Markov process,$\{X_t\}_{t\geq 0}$, where t is the time. This market model is called a semi-Markov modulated GBM (geometric Brownian motion) model. Hunt & Devolder[12] have talked about the advantages of using semi-Markov switching models over homogeneous Markov switching models as the memoryless Markovian property restricts the flexibility of semi-Markov process in exhibiting holding time duration dependence. For example, it is useful in tackling the duration dependent business cycles. Thus, it gives us the motivation to study this generalized market model.

The study on option pricing using Föllmer Schweizer decomposition, in a semi-Markov modulated GBM model, was done in Ghosh & Goswami[8], where they have shown the price function to satisfy a non-local system of parabolic PDEs, which arises while finding the optimal hedging strategy for a claim, a European call option with prescribed strike price and maturity time, in semi-Markov modulated market. Goswami, Patel & Shevgaonkar[9] have shown that the same price function also satisfies a Volterra integral equation of second kind and that non-local PDE is equivalent to the Volterra equation.

In the next section a brief description of locally risk minimizing hedging in general (incomplete) market, following Föllmer and Schweizer, is given.

3.1 Locally risk minimizing pricing

Consider, a market with two assets, $\{S_t\}_{t\geq 0}$ and $\{B_t\}_{t\geq 0}$ where S_t and B_t are continuous semi-martingales and B_t has finite variation. Then, an admissible strategy is defined by,

$$\pi = \{\pi_t = (\xi_t, \varepsilon_t), 0 \leq t \leq T\}, \text{ that satisfies (A1).}$$

The portfolio value at time t is given by,

$$V_t = \xi_t S_t + \varepsilon_t B_t. \tag{3.1.1}$$

Assumption (A1) (i)ξ_t is square integrable w.r.t. S_t, (ii) expectation of ε_t^2 is finite, (iii) $\exists a > 0$ s.t. $P(V_t \geq -a, t \in [0,T]) = 1$.

One can allow adding an instantaneous cash flow, ΔC_t with the return of investment at $t - \Delta$ in the portfolio value at time t, thus

$$V_t = \xi_{t-\Delta} S_t + \varepsilon_{t-\Delta} B_t + \Delta C_t. \tag{3.1.2}$$

From (3.1.1) and (3.1.2), we get the discrete equation,

$$V_t - V_{t-\Delta} = \xi_{t-\Delta}(S_t - S_{t-\Delta}) + \varepsilon_{t-\Delta}(B_t - B_{t-\Delta}) + \Delta C_t,$$

equivalently we got the SDE as

$$dV_t = \xi_t dS_t + \varepsilon_t dB_t + dC_t. \tag{3.1.3}$$

Definition 3.1.1. *Strategy, $\pi_t = (\xi_t, \varepsilon_t)$, is defined to be self financing if*

$$dV_t = \xi_t dS_t + \varepsilon_t dB_t, \quad \forall t \geq 0.$$

From the above definition, it's clear to see, for a self financing strategy π, the cost process $C_t(\pi) =$ initial cash flow $=$ constant.

There are market models, which are incomplete in the sense that there exists no class of self financing strategies to perfectly hedge for given contingent claim. These markets are known as incomplete markets. Here, one can get an optimal strategy through minimizing quadratic residual risk, a measure of cash flow, under certain constraint[6]. It is shown in [6] that for arbitrage free market, the existence of an optimal strategy to hedge an $\mathcal{F}_T$ measurable claim H is equivalent to that of Föllmer Schweizer decomposition of discounted claim, $H^* := B_T^{-1} H$ in the following form

$$H^* = H_0 + \int_0^T \xi_u^{H^*} dS_u^* + L_T^{H^*} \tag{3.1.4}$$

where $H_0 \in L^2(\Omega, \mathcal{F}_0, P), L^{H^*} = \{L_t^{H^*}\}_{0 \leq t \leq T}$ is a square integrable martingale, orthogonal to the martingale part of S_t and $\xi^{H^*} = \{\xi_t^{H^*}\}$ satisfies (A1)(i). Here, ξ^{H^*} in the decomposition constitutes the optimal strategy. So, the optimal strategy $\pi_t = (\xi_t, \varepsilon_t)$ is given by

$$\begin{aligned} \xi_t &:= \xi_t^{H^*}; \\ \varepsilon_t &:= V_t^* - \xi_t^{H^*} S_t^*; \end{aligned}$$

with $V_t^* := H_0 + \int_0^t \xi_u^{H^*} dS_u^* + L_t^{H^*}$,$S_t^* := B_t^{-1} S_t$ and $B_t V_t^*$ represents the locally risk minimizing price at t of the claim H.

3.2 Description of market model

Considering $(\Omega, \mathcal{F}, P)$ to be underlying complete probability space, let us model the hypothetical state (assumed to be observable) by $\{X_t\}_{t\geq 0}$, a semi-Markov process on finite state space $S = \{1, 2, \ldots, \theta\}$ with transition probabilities (p_{ij}) and conditional holding time distributions $F(\cdot|i)$. Also consider $\{T_n\}_{n\geq 0}$ to be consecutive jump/transition times with $T_0 = 0$.

We assume in a semi-Markov regime, that the interest rate,r_t evolves depending on the hypothetical state of the market. Here the market consists of one locally risk free money market and one stock as the risky asset. Let, $\{B_t\}_{t\geq 0}$ be the price of money market account at t, also called the bond, with $r_t = r(X_t)$ and $B_0 = 1$. Therefore,

$$B_t = \exp(\int_0^t r(X_u) du).$$

Consider the stock price process to be $\{S_t\}_{t\geq 0}$, governed by semi-Markov modulated GBM as

$$dS_t = S_t(\mu(X_t)dt + \sigma(X_t)dW_t), \qquad S_0 > 0 \tag{3.2.1}$$

where $\{W_t\}_{t\geq 0}$ is standard Wiener process independent of the semi-Markov process $\{X_t\}_{t\geq 0}$; the drift coefficient, μ is real valued and the volatility, $\sigma \in (0, \infty)$ on domain S. The solution of SDE(3.2.1) is an $\mathcal{F}_t$ semi-martingale with almost sure continuous paths, where $\mathcal{F}_t$ is a filtration of $\mathcal{F}$ with right continuous version of X_t and S_t generated filtration. Ghosh & Goswami[8] have shown that this market model does admit the existence of equivalent martingale measure, thus it is arbitrage free under admissible strategy. The stock price involves uncertainties arising because of the driving Brownian motion and the semi-Markov switch. Due to semi-Markov switching, the market is incomplete.

3.3 B-S-M type price equation of European call option

Consider a particular contingent claim, a European call option, on S_t with strike price K and maturity time T. Then the $\mathcal{F}_T$ measurable contingent claim H, is given by,

$$H = (S_T - K)^+.$$

In questing for an optimal hedging strategy for the claim in semi-Markov modulated market, we encounter the following system of differential equations[8, 9, 12],

$$\begin{aligned}
&\frac{\partial}{\partial t}\phi(t,s,i,y) + \frac{\partial}{\partial y}\phi(t,s,i,y) + r(i)s\frac{\partial}{\partial s}\phi(t,s,i,y) + \frac{1}{2}\sigma^2(i)s^2\frac{\partial^2}{\partial s^2}\phi(t,s,i,y) \\
&\frac{f(y|i)}{1-F(y|i)}\sum_{j\neq i} p_{ij}[\phi(t,s,j,0) - \phi(t,s,i,y)] = r(i)\phi(t,s,i,y),
\end{aligned} \tag{3.3.1}$$

defined on

$$\mathcal{D} := \{(t,s,i,y) \in (0,T) \times \mathbb{R}^+ \times S \times (0,T) | y \in (0,t)\}, \tag{3.3.2}$$

with boundary conditions

$$\begin{aligned}
&\phi(t,0,i,y) = 0, \quad \forall t \in [0,T], \\
&\phi(T,s,i,y) = (s-K)^+; \quad s \in \mathbb{R}^+; \quad 0 \le y \le T; \quad i = 1,2,...,\theta;
\end{aligned} \tag{3.3.3}$$

where $r(\cdot)$,$\sigma(\cdot)$,(p_{ij}),$F(\cdot|i)$ are mentioned in section (3.2) and $f(\cdot|i)$ is derivative of $F(\cdot|i)$. Goswami, Patel & Shevgaonkar[9] have shown that the solution gives locally risk minimizing price function of the European call option through establishing existence and uniqueness of solution of an equivalent Volterra integral equation of second kind.

Chapter 4

Estimation of Instantaneous transition rate function

4.1 Convergence of the MLE

Considering an augmented Markov process $(X, Y) = (X_t, Y_t)_{t\geq 0}$, defined on a complete probability space with $S = \{1, 2, ..., \theta\}$ such that $\theta < \infty$, as the state-space for the underlying Markov chain $(X_{T_n})_{n\geq 0}$. $(Y_{T_n})_{n\geq 0}$ are the sojourn times, defined by $Y_{T_0} = 0$ and $Y_{T_n-} = T_n - T_{n-1}$, in these states measured in $\overline{\mathbb{R}}^+$.

For simplicity, we define $X_n := X_{T_n}$ and $Y_n := Y_{T_n-}$. To define the augmented Markov process, the initial law has been defined as $P(X_0 = j) := p(j)$ and the following semi-Markov kernel is specified as,

$$P(X_{n+1} = j, Y_{n+1} \leq y | X_0, X_1, ..., X_n, Y_1, ..., Y_n) := Q_{X_n j}(y) \quad (a.s.) \tag{4.1.1}$$

$\forall y \in \overline{\mathbb{R}}^+$ and $1 \leq j \leq \theta$.

With the assumption, $Q_{ii}(y) = 0$, $\forall i \in S$,the conditional distribution function of the holding/sojourn time in state i is defined by,

$$F(y|i) := \sum_{j=1}^{\theta} Q_{ij}(y), \quad \forall y \in \overline{\mathbb{R}}^{+}.$$

Naturally we can see,

$$\begin{aligned} Q_{ij}(y) &= P(X_{n+1} = j, Y_{n+1} \leq y | X_0, X_1, ..., X_n = i, Y_1, ..., Y_n) \\ &= P(Y_{n+1} \leq y | X_{n+1} = j, X_n = i) P(X_{n+1} = j | X_n = i) \\ &= p_{ij} F(y|i, j). \end{aligned}$$

Now we assume that the underlying semi-Markov process has a special property which states that $F(y|i,j)$ does not depend on the j variable. Hence, $Q_{ij}(y) = p_{ij}F(y|i)$.

We define the instantaneous transition rate function, $\forall j \neq i \in S$,of a semi-Markov kernel by,

$$\lambda_{ij}(y) := \lim_{\Delta y \downarrow 0} \frac{P(X_{n+1} = j, y < Y_{n+1} \leq y + \Delta y | X_n = i, Y_{n+1} > y)}{\Delta y},$$

$$\lambda_{ii}(y) := - \sum_{j \in S, j \neq i} \lambda_{ij}(y) \quad \forall i \in S.$$

i.e.,

$$\lambda_{ij}(y) = \begin{cases} \frac{q_{ij}(y)}{1 - F(y|i)} & \text{if } p_{ij} > 0 \text{ and } F(y|i) < 1 \\ 0 & \text{otherwise,} \end{cases}$$

where $q_{ij}(.)$, defined by $q_{ij}(y) := p_{ij} f(y|i)$, is the density/derivative of $Q_{ij}(.)$, assuming $Q_{ij}(.)$ to be absolute continuous w.r.t. Lebesgue's measure. We also define,

$$\lambda_i(y) := |\lambda_{ii}(y)|,$$

$$\Lambda_i(y) := \sum_{j=1}^{\theta} \int_0^y \lambda_{ij}(u) du.$$

Consider a history of augmented Markov process censored at fixed time τ,

$$\mathcal{H}(\tau) = (X_0, X_1, ..., X_{N_\tau}, Y_1, Y_2, ..., Y_{N_\tau}, U_\tau),$$

where $U_\tau = \tau - T_{N_\tau}$ is the backward recurrence time, $(T_n)_{n\geq 0}$ are jump times and N_τ is the number of jumps before time τ.

The associated log-likelihood function is maximized to obtain the maximum likelihood estimator(MLE) of the transition rate function,$\lambda_{ij}(.)$. The likelihood function for $\mathcal{H}(\tau)$ is

$$L(\tau) = p(X_0)(1 - \sum_{l=1}^{\theta} Q_{X_{N_\tau}l}(U_\tau)) \prod_{l=0}^{N_\tau - 1} p_{X_l X_{l+1}} f(Y_{l+1}|X_l)$$

$$\implies p(X_0)^{-1}L(\tau) = exp(-\Lambda_{X_{N_\tau}}(U_\tau)) \prod_{l=0}^{N_\tau - 1} exp(-\Lambda_{X_l}(Y_{l+1}))\lambda_{X_l, X_{l+1}}(Y_{l+1})$$

Then we consider log-likelihood as

$$l(\tau) := \log p(X_0)^{-1}L(\tau) = \sum_{l=0}^{N_\tau - 1}(\log \lambda_{X_l, X_{l+1}}(Y_{l+1}) - \Lambda_{X_l}(Y_{l+1})) - \Lambda_{X_{N_\tau}}(U_\tau).$$

Approximate the transition rate $\lambda_{ij}(y)$ by piecewise constant function $\lambda^*_{ij}(y)$ defined by $\lambda^*_{ij}(y) = \lambda_{ij}(v_k) = \lambda_{ijk}$ for $y \in (v_k, v_{k+1}] = I_k$, where $(v_k)_{0\leq k\leq M-1}$ is a regular subdivision of $[0, \tau]$, i.e., $0 = v_0 < v_1 < v_2 < ... < v_{M-1} < v_M = \tau$ with step $\Delta_\tau = \frac{\tau}{M}$ and $M = [\tau^{1+\alpha}]$, where [.] is the Greatest integer or box function and $\alpha \in (0, 1)$,to get

$$\lambda^*_{ij}(y) = \sum_{k=0}^{M-1} \lambda_{ijk} 1_{(v_k, v_{k+1}]}(y). \tag{4.1.2}$$

Then taking the log-likelihood and (4.1.2) and doing some calculus one obtains,

$$l(\tau) = \sum_{i,j\in S} \sum_{k=0}^{M-1} (d_{ijk} \log \lambda_{ijk} - \lambda_{ijk} v_{ik}), \tag{4.1.3}$$

where v_{ik} is a function of sojourn time in state i on the interval of I_k such that $\{v_{ik}\}_k$ is a decreasing sequence while each member is a function of sojourn time in state i on the interval I_k, and d_{ijk} is the number of transitions from state i to state j for which the observed sojourn time in state i lies in I_k and for $N_\tau \geq 1$, they are respectively given by

$$v_{ik} = \sum_{l=0}^{N_\tau - 1} (Y_{l+1} \wedge v_{k+1} - v_k) 1_{\{i\}\times(v_k,\infty)}(X_l, Y_{l+1}) + (U_\tau \wedge v_{k+1} - v_k) 1_{\{i\}\times(v_k,\infty)}(X_{N_\tau}, U_\tau),$$

and

$$d_{ijk} = \sum_{l=0}^{N_\tau - 1} 1_{\{i\}\times\{j\}\times I_k}(X_l, X_{l+1}, Y_{l+1}).$$

So the MLE of λ_{ijk} is

$$\hat{\lambda}_{ijk} = \begin{cases} d_{ijk}/v_{ik} & \text{if } \; v_{ik} > 0; \\ 0 & \quad\text{otherwise.} \end{cases} \tag{4.1.4}$$

Therefore the estimator of $\lambda_{ij}(y)$ is given by,

$$\hat{\lambda}_{ij}(y, \tau) = \sum_{k=0}^{M-1} \hat{\lambda}_{ijk} 1_{(v_k, v_{k+1}]}(y). \tag{4.1.5}$$

Suppose, $N_i = N_i(\tau)$ and $N_{ij} = N_{ij}(\tau)$ be the number of occurrence of event $\{X_l = i\}$ for $0 < l < N_\tau + 1$ and number of occurrence of i to j transition in $[0, \tau]$, respectively. Let us denote N_τ as N. Then, we define the following estimators:

$$Q_{ij}^N(y) = \frac{1}{N_i} \sum_{l=1}^{N} 1_{\{i\}\times\{j\}\times[0,y]}(X_l, X_{l+1}, Y_{l+1}),$$

$$q_{ij}^N(y) = \frac{Q_{ij}^N(v_{k+1}) - Q_{ij}^N(v_k)}{\Delta_\tau} \quad \text{if } \; y \in (v_k, v_{k+1}],$$

$$\begin{aligned}
\hat{f}_{ij}(y,\tau) &= \frac{1}{\Delta_\tau N_{ij}} \sum_{l=1}^{N} \sum_k 1_{\{i\}\times\{j\}\times I_k}(X_l, X_{l+1}, Y_{l+1}).1_{I_k}(y) \\
&= \frac{1}{\Delta_\tau N_{ij}} \sum_{r=1}^{N_{ij}} \sum_k 1_{I_k}(Y_{l_r+1}) 1_{I_k}(y) \\
&= \frac{1}{\Delta_\tau N_{ij}} \sum_{r=1}^{N_{ij}} \zeta_{ijr},
\end{aligned}$$

where

$$\zeta_{ijr} := \sum_k 1_{I_k}(Y_{l_r+1}) 1_{I_k}(y) = \sum_{k^*} 1_{I^*_k}(Y_{l_r+1}).$$

It is clear to see, $q^N_{ij}(y) = \frac{N_{ij}}{N_i} \hat{f}_{ij}(y,\tau)$.

We also define $G^N_{ij}(.)$ as

$$G^N_{ij}(y) = \sum_{l=1}^{N} \sum_k \frac{(Y_{l+1} - v_k)}{\Delta_\tau N_i} 1_{\{i\}\times\{j\}\times I_k \cap (y,\infty)}(X_l, X_{l+1}, Y_{l+1}).1_{I_k}(y).$$

Now the estimator of hazard rate (4.1.5) can be written as

$$\hat{\lambda}_{ij}(y,\tau) = \sum_k \frac{q^N_{ij}(v_k)}{1 - \sum_{j=1}^{\theta} \{Q^N_{ij}(v_{k+1}) - G^N_{ij}(v_k)\} + [\frac{((U_\tau \wedge v_{k+1} - v_k))}{\Delta_\tau N_i}] 1_{\{i\}\times[v_k,\infty)}(X_{N_\tau}, U_\tau)}.1_{I_k}(y). \tag{4.1.6}$$

In order to study asymptotic property of hazard rate, we define the following quantities:

$$A_{ij}(y) = \sum_k \frac{q^N_{ij}(y)}{1 - \sum_{j=1}^{\theta} \{Q^N_{ij}(v_k)\} + [\frac{((U_\tau \wedge v_{k+1} - v_k))}{\Delta_\tau N_i}] 1_{\{i\}\times[v_k,\infty)}(X_{N_\tau}, U_\tau)}.1_{I_k}(y)$$

and

$$B_{ij}(y) = \sum_k \frac{q^N_{ij}(y)}{1 - \sum_{j=1}^{\theta} \{Q^N_{ij}(v_{k+1})\} + [\frac{((U_\tau \wedge v_{k+1} - v_k))}{\Delta_\tau N_i}] 1_{\{i\}\times[v_k,\infty)}(X_{N_\tau}, U_\tau)}.1_{I_k}(y).$$

Lemma 4.1.1. For any fixed $y \in [0, \tau]$, we get

$$A_{ij}(y) \leq \hat{\lambda}_{ij}(y, \tau) \leq B_{ij}(y).$$

Proof. The right inequality is obvious owing to the non-negativity of $G_{ij}^N(y)$. For the left one, we have for $y \in I_k$, $(Y_k - v_k)/\Delta_\tau \leq 1$, so

$$\begin{aligned}
1 - \sum_{j=1}^{\theta}\{Q_{ij}^N(v_{k+1}) - G_{ij}^N(v_k)\} &\leq 1 - \sum_{j=1}^{\theta}\{Q_{ij}^N(v_{k+1}) - \sum_{l=1}^{N_i} 1_{\{i\}\times\{j\}\times I_k}(X_l, X_{l+1}, Y_{l+1})/N_i\} \\
&= 1 - \sum_{j=1}^{\theta}\{Q_{ij}^N(v_{k+1}) - Q_{ij}^N(v_{k+1}) + Q_{ij}^N(v_k)\} \\
&= 1 - \sum_{j=1}^{\theta} Q_{ij}^N(v_k).
\end{aligned}$$

Theorem 4.1.1. Consider the transition rate function, $\lambda_{ij}(.)$, to be continuous.

The estimator of transition rate, $\hat{\lambda}_{ij}(., \tau)$, is uniformly strongly consistent for $\lambda_{ij}(.)$, for $\alpha \in (0, 1/2)$, in all compacts $[0, T]$,$T \in \mathbb{R}_+$, in the sense that

$$\max_{i,j} \sup_{y \in [0,T]} |\hat{\lambda}_{ij}(y, \tau) - \lambda_{ij}(y)| \to 0, \quad \text{almost surely, as} \quad \tau \to \infty.$$

Proof. From the above lemma 4.1.1., the sufficient condition for the theorem is that $A_{ij}(y)$ and $B_{ij}(y)$ converge in probability, as $\tau \to \infty$, to the same limit $\lambda_{ij}(y)$.

The proof is divided into two parts, firstly we prove that $q_{ij}^N(y)$ uniformly converges almost surely, as $\tau \to \infty$, to the limit $q_{ij}(y) = p_{ij} f(y|i, j)$.

We check,

$$\begin{aligned}
|q_{ij}^N(y) - q_{ij}(y)| &= |\frac{N_{ij}}{N_i}(\hat{f}_{ij}(y, \tau) - f(y|i, j)) + f(y|i, j)(\frac{N_{ij}}{N_i} - p_{ij})| \\
&\leq \frac{N_{ij}}{N_i}|\hat{f}_{ij}(y, \tau) - f(y|i, j)| + f(y|i, j)|\frac{N_{ij}}{N_i} - p_{ij}|.
\end{aligned}$$

From Billingsley[2],

$$|\frac{N_{ij}}{N_i} - p_{ij}| \to 0 \text{ as } \tau \to \infty \qquad \text{(a.s.)}.$$

It is clear to see the asymptotic unbiasedness of $\hat{f}_{ij}(y,\tau)$,

$$E[\hat{f}_{ij}(y,\tau)] - f(y|i,j) \to 0 \text{ as } \tau \to \infty.$$

Suppose, $0 < \alpha < \frac{1}{2}$. To prove the almost sure convergence of $\hat{f}_{ij}(y,\tau)$ to $f(y|i,j)$, it is sufficient to prove that

$$\hat{f}_{ij}(y;n) = \frac{1}{n\Delta}\sum_{r=1}^{n}\zeta_{ijr}$$

converges almost surely to $f(y|i,j)$ as $n \to \infty$, with $\Delta = n^{-\alpha}$.

We see,

$$\begin{aligned}|\hat{f}_{ij}(y,n) - E[\hat{f}_{ij}(y,n)]| &= |\frac{\hat{F}(v_{k+1};n|i,j) - \hat{F}_i(v_k;n|i,j)}{\Delta} - \frac{F(v_{k+1}|i,j) - F(v_k|i,j)}{\Delta}| \\ &\leq \frac{2}{\Delta}\sup_y |\hat{F}(y;n|i,j) - F(y|i,j)|,\end{aligned}$$

where $\hat{F}(y;n|i,j)$ is the associated (with f) c.d.f. estimator.

From Dvoretzky's et al., inequality[5], the measure has been shown to be bounded by some exponential function of n, as

$$\begin{aligned}P\{\sup_y |\hat{f}_{ij}(y;n) - E[\hat{f}_{ij}(y;n)]| > \epsilon\} &\leq P\{\sup_y |\hat{F}(y;n|i,j) - F(y|i,j)| > \epsilon\frac{\Delta}{2}\} \\ &< C\exp(-2(\frac{\epsilon n^{1/2-\alpha}}{2})^2) \\ &= C\exp(\frac{-\epsilon^2}{2}n^{1-2\alpha}).\end{aligned}$$

For $0 < \alpha < 1/2$,

$$\sum_{n=0}^{\infty} P\{\sup_y |\hat{f}_{ij}(y;n) - E[\hat{f}_{ij}(y;n)]| > \epsilon\} \quad < \quad C\sum_{n=0}^{\infty} \exp\left(\frac{-\epsilon^2}{2} n^{1-2\alpha}\right) < \infty,$$

so using Borel Cantelli's Lemma, then

$$P\{\lim_n \sup_y |\hat{f}_{ij}(y;n) - E[\hat{f}_{ij}(y;n)]| > \epsilon\} = 0$$

i.e.

$$P\Big(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{\sup_y |\hat{f}_{ij}(y;n) - E[\hat{f}_{ij}(y;n)]| > \epsilon\}\Big) = 0$$

$$\Longrightarrow \sup_y |\hat{f}_{ij}(y;n) - E[\hat{f}_{ij}(y;n)]| \to 0 \quad \text{almost surely, as} \;\; n \to \infty.$$

By theorem 2.1, p. 10, in Gut[11],

$$\sup_y |\hat{f}_{ij}(y,\tau) - E[\hat{f}_{ij}(y,\tau)]| \to 0 \quad \text{almost surely, as} \;\; \tau \to \infty.$$

There exists almost sure convergence of $\hat{f}_{ij}(y,\tau)$ to $f(y|i,j)$ as $\hat{f}_{ij}(y,\tau)$ shows asymptotic unbiasedness and convergence is uniform on any compact $[0,T]$, as $f(y|i,j)$ is continuous (from the assumption made for the theorem).

Therefore,

$$\max_{i,j} \sup_{y\in[0,T]} |q_{ij}^N(y) - q_{ij}(y)| \to 0 \qquad \text{(a.s.), as} \;\; \tau \to \infty.$$

Consider, $\hat{H}_i(y,\tau)$ be the estimator of $H_i(y) := F(y|i)[\because F(y|i,j) = F(y|i)$ as described in our model],

$$\hat{H}_i(y,\tau) = \sum_{j=1}^{\theta} Q_{ij}^N(y) - \frac{((U_\tau \wedge v_{k+1}) - v_k)}{\Delta_\tau N_i} 1_{\{i\}\times[v_k,\infty)}(X_{N_\tau}, U_\tau).$$

For $y \in I_k$, $A_{ij}(y)$ and $B_{ij}(y)$ can be written as

$$A_{ij}(y) = \frac{q_{ij}^N(y)}{1 - H_i(y)} \sum_{n=0}^{\infty} (\frac{\hat{H}_i(v_k, \tau) - H_i(y)}{1 - H_i(y)})^n,$$

$$B_{ij}(y) = \frac{q_{ij}^N(y)}{1 - H_i(y)} \sum_{n=0}^{\infty} (\frac{\hat{H}_i(v_{k+1}, \tau) - H_i(y)}{1 - H_i(y)})^n.$$

Now, as $\tau \to \infty$, $\hat{H}_i(y, \tau)$ converges uniformly to $H_i(y)$ on $\mathbb{R}^+$[15, 17, 10]. Therefore, $A_{ij}(y)$ and $B_{ij}(y)$ converge uniformly, as $\tau \to \infty$ with probability 1, to the same limit $\lambda_{ij}(y)$. □

Remark: Continuous time Markov process with finite state space is a special case of the semi-Markov process we consider here.

4.2 Numerical experiment

We did numerical experiment to see the comparison between maximum likelihood estimation of hazard rate function and the theoretical/predefined hazard rate of the semi-Markov process. We choose, the state-space to be $S = \{1, 2, 3\}$ and transition matrix

$$(p_{ij}) = \begin{pmatrix} 0.0 & 0.1 & 0.9 \\ 0.4 & 0.0 & 0.6 \\ 0.7 & 0.3 & 0.0 \end{pmatrix}.$$

,

for the semi-Markov process.

In order to have the semi-Markov process defined, we generate the holding times, Y_n by adding two identically independent random variables with distribution $\exp(1) = \Gamma(1, 1)$, thus

$$Y_n \sim \Gamma(2, 1).$$

So, the associated p.d.f. for $\{Y_n\}$ is $f(y;2,1) = \frac{ye^{-y}}{1^2.\Gamma(2)} = ye^{-y}$ and c.d.f. is

$$F(y;2,1) = \int_0^y ue^{-u}du = -ye^{-y} + 1 - e^{-y}.$$

Therefore, the theoretical hazard rate function for any $y \in [0,\tau]$, is given by,

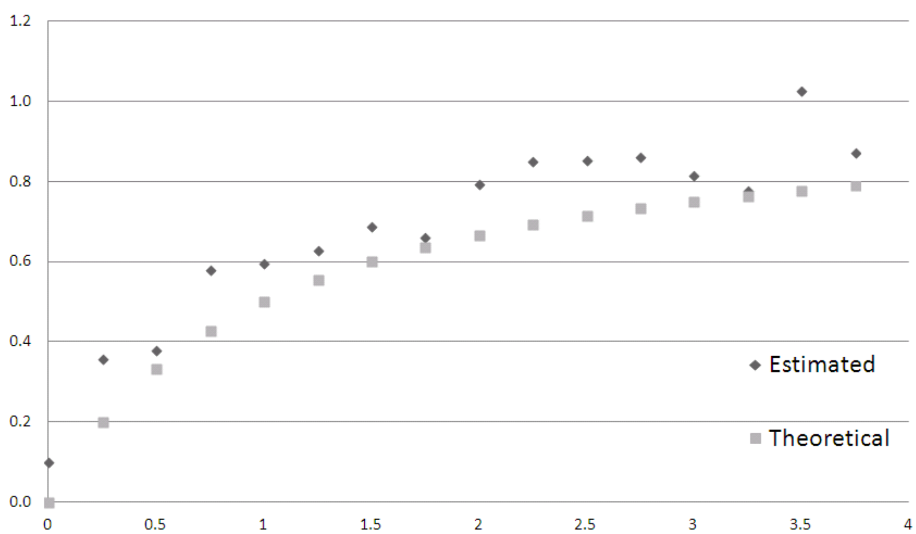
$$\lambda_i(y) = \lambda_{ij}(y)/p_{ij} = \frac{f(y|i)}{1-F(y|i)} = \frac{y}{y+1} \quad \forall i \in S \text{ and } i \neq j.$$

Estimated hazard rate(approximated to step function for regular intervals I_k) for $y \in [0,\tau]$,

$$\hat{\lambda}_i(y) = \hat{\lambda}_{ijk}/p_{ij} \quad \text{if } y \in I_k, \forall i \in S \text{ and } i \neq j.$$

The estimated hazard rate function (approximated step function), $\hat{\lambda}_i(.)$, for fixed $(i,j) = (1,3)$, in the interval of $y \in [0, T = 4]$,$T \in [0,\tau]$, is therefore generated, over the history of the semi-Markov process censored at $\tau = 3900$.

Comparison between the theoretical and estimated hazard rate function (y takes the values from X-axis):



Remark: The estimated hazard rate function has been estimated as a step function for each individual intervals $\{I_k\}_{k\geq 0}$ and then plotted at the left limit i.e.

v_k for each. Now, as the theoretical hazard rate is strictly increasing, the estimated hazard rate appears to be slightly overestimated in the graph than actual estimation.

Chapter 5

Approximation of price function

5.1 Convergence of approximation

Definition 5.1.1. *Let $\hat{\lambda}_i$ be an MLE of λ_i, the hazard rate, and ϕ be the solution of the problem* (3.3.1)-(3.3.3)*. If the function λ_i is replaced by $\hat{\lambda}_i$ in* (3.3.1)-(3.3.3)*, then the solution of the modified equation is called the HREB (hazard rate estimation based) approximation of ϕ with parameter $\hat{\lambda}_i$.*

Theorem 5.1.1. *Let, $\hat{\phi}^n$ be the HREB approximation of ϕ with parameter $\hat{\lambda}_i^n$, where $\hat{\lambda}_i^n$ is the MLE of λ_i using observation of the semi-Markov process of time length n with number of grid points $\lfloor n^{1+\alpha} \rfloor$, $0 < \alpha < \frac{1}{2}$.*

Then, $\hat{\phi}^n$ converges to ϕ point-wise, as $n \to \infty$.

Proof. The classical solution of (3.3.1)-(3.3.3) exists uniquely[8, 9]. So, ϕ is unique.

Replacing the functional parameter, hazard rate, $\lambda_i(y)$ by the MLE, $\hat{\lambda}_i^n(y)$, using observation of the semi-Markov process of time length n with number of grid points $\lfloor n^{1+\alpha} \rfloor$, for a fixed α s.t. $0 < \alpha < \frac{1}{2}$ in (3.3.1)-(3.3.3). So, $\hat{\phi}^n$ satisfies a modified (3.3.1)-(3.3.3) which also has a unique solution.

Now, we define the difference of the functions ϕ and its HREB estimator, $\hat{\phi}^n$ to be,

$$\psi^n(t, s, i, y) := \phi(t, s, i, y) - \hat{\phi}^n(t, s, i, y).$$

Now, by doing some algebra in (3.3.1)-(3.3.3), we see that ψ^n satisfies the following system of equations,

$$\frac{\partial}{\partial t}\psi(t,s,i,y) + r(i)s\frac{\partial}{\partial s}\psi(t,s,i,y) + \frac{1}{2}\sigma^2(i)s^2\frac{\partial^2}{\partial s^2}\psi(t,s,i,y) + \lambda_i(y)\sum_{j\neq i} p_{ij}\phi(t,s,j,0) - \lambda_i(y)\phi(t,s,i,y)$$
$$= r(i)\psi(t,s,i,y) + (\lambda_i(y) - \hat{\lambda}_i^n(y))\sum_{j\neq i} p_{ij}\hat{\phi}^n(t,s,j,0) - (\lambda_i(y) - \hat{\lambda}_i^n(y))\hat{\phi}^n(t,s,i,y),$$

defined on

$$\mathcal{D} := \{(t,s,i,y) \in (0,T)\times\mathbb{R}^+\times S\times(0,T) | y \in (0,t)\},$$

with boundary conditions

$$\psi(t,0,i,y) = 0, \quad \forall t \in [0,T],$$

$$\psi(T,s,i,y) = 0, \quad s \in \mathbb{R}^+; \quad 0 \le y \le T; \quad i = 1,2,\cdots,\theta;$$

where $r(\cdot)$,$\sigma(\cdot)$,(p_{ij}),$F(\cdot|i)$,$f(\cdot|i)$ are mentioned in section (3.2) and (3.3).

We rewrite the above system of equations as,

$$\frac{\partial}{\partial t}\psi(t,s,i,y) + \mathcal{L}\psi(t,s,i,y) = r(i)\psi(t,s,i,y) - f^n, \tag{5.1.1}$$

where

$$(\mathcal{L}\psi)(t,s,i,y) := \left[r(i)s\frac{\partial}{\partial s} + \frac{1}{2}\sigma^2(i)s^2\frac{\partial^2}{\partial s^2}\right]\psi(t,s,i,y) + \lambda_i(y)[\sum_{j=1}^{\theta} p_{ij}\psi(t,s,j,0) - \psi(t,s,i,y)],$$

and $\mathcal{L}$ is the infinitesimal generator of (S_t, X_t, Y_t) satisfying

$$dS_t = S_t(r(X_{t-})dt + \sigma(X_{t-})dW_t),$$

where X_t is a semi-Markov process with hazard rate $\lambda_i(y)$ and transition matrix (p_{ij}), and Y_t is the holding time process and

$$f^n(t,s,i,y) := (\hat{\lambda}_i^n(y) - \lambda_i(y))[\sum_{j\neq i} p_{ij}\hat{\phi}^n(t,s,j,0) - \hat{\phi}^n(t,s,i,y)].$$

Then, using Feynman-Kac formula,

$$\psi^n(t,s,i,y) = E[\int_t^T \exp\left(-\int_t^\tau r(X_u)du\right)\eta^n d\tau | S_t = s, X_t = i, Y_t = y], \qquad (5.1.2)$$

where,

$$\eta^n(\tau) = f^n(\tau, S_\tau, X_\tau, Y_\tau).$$

From theorem 4.1.1.[16], we have the uniform convergence of $\hat{\lambda}^n_{X_\tau}$ to λ_{X_τ} in $y \in [0,T]$,$T \in \mathbb{R}^+$, almost surely, for $0 < \alpha < \frac{1}{2}$. Or, in other words we get, for any given $\epsilon(> 0)$, $\exists N$ s.t. $P(N < \infty) = 1$ and for $n \geq N$,

$$|\lambda^n_{X_\tau}(y) - \lambda_{X_\tau}(y)| < \epsilon \quad \forall y \in [0,T]. \qquad (5.1.3)$$

Hence, it is clear that for fixed (τ,ω),

$$\eta^n(\tau)(\omega) \downarrow 0 \text{ point-wise as } n \to \infty.$$

For European call price function, $\hat{\phi}^n(\tau, S_\tau, j, 0)$, we see

$$0 \leq \hat{\phi}^n(\tau, S_\tau, j, 0) \leq S_\tau \quad \forall j,$$

then,

$$|\sum_{j \neq X_\tau} P_{X_\tau j}\hat{\phi}^n(\tau, S_\tau, j, 0) - \hat{\phi}^n(\tau, S_\tau, X_\tau, Y_\tau)| \leq S_\tau. \qquad (5.1.4)$$

Therefore, from (5.1.3) and (5.1.4), we get, $|\eta^n(\tau,\omega)| \leq \epsilon S_\tau \quad \forall n \geq N$ & $y \in [0,T]$.

Define, $g(\tau)(\omega) := \epsilon S_\tau(\omega)$.

To prove for Lebesgue integrability of g in any interval $[a,b] \in [0,T]$, we need,

$$E[\int_0^T \epsilon S_\tau d\tau] < \infty. \qquad (5.1.5)$$

By applying Tonelli's theorem for $S_\tau(> 0)$, we get,

$$E[\int_0^T \epsilon S_\tau d\tau] = \int_0^T \epsilon E[S_\tau] d\tau. \tag{5.1.6}$$

We know,

$$S_\tau = S_0 \exp[\int_0^\tau \{\mu(X_u) - \frac{1}{2}\sigma^2(X_u)\} du + \int_0^\tau \sigma(X_u) dW_u]. \tag{5.1.7}$$

Let us define the following variables as

$$c := \max_{i \in S} \{\mu(i) - \frac{1}{2}\sigma^2(i)\},$$

$$d := \max_{i \in S} \{\sigma^2(i)\}.$$

Clearly,

$$S_\tau \leq S_0 \exp(\int_0^\tau c du) \exp(\int_0^\tau \sigma(X_u) dW_u). \tag{5.1.8}$$

Hence, it is sufficient to prove that the R.H.S. has finite expectation. To this end we observe,

$$\begin{aligned}
\int_0^\tau \sigma(X_u) dW_u &= \sum_{n=1}^{\infty} \int_{T_{n-1}\wedge\tau}^{T_n\wedge\tau} \sigma(X_{T_{n-1}}) dW_u \\
&= \sum_{n=1}^{\infty} \int_{T_{n-1}\wedge\tau}^{T_n\wedge\tau} \sigma(X_{T_{n-1}})(W_{T_n\wedge\tau} - W_{T_{n-1}\wedge\tau})
\end{aligned}$$

Hence, the conditional distribution of $\int_0^\tau \sigma(X_u) dW_u$ given $\mathcal{F}_\tau^X$ is normal with mean zero and variance

$$\sum_{n=1}^{\infty} \sigma^2(X_{T_{n-1}})[T_n \wedge \tau - T_{n-1} \wedge \tau],$$

where $\mathcal{F}_\tau^X$ is the filtration of $\mathcal{F}$ generated by $X = \{X_t\}_{t\geq 0}$.

Now using the formula of variance of a Log normal random variable, we get,

$$\begin{aligned} E\left[\exp\left(\int_0^\tau \sigma(X_u)dW_u\right)\right] &= E\left[\exp\left(\frac{1}{2}[\sum_{n=1}^{\infty}\sigma^2(X_{T_{n-1}})[T_n\wedge\tau - T_{n-1}\wedge\tau]\right)\right] \\ &\le E\left[\exp\left(\frac{d}{2}\sum_{n=1}^{\infty}(T_n\wedge\tau - T_{n-1}\wedge\tau)\right)\right] \\ &= \exp\left(\frac{d}{2}\tau\right). \end{aligned}$$

Using above inequality we obtain from (5.1.8),

$$E(S_\tau) \le S_0 e^{c\tau} e^{\frac{d}{2}\tau} = S_0 e^{\left(c+\frac{d}{2}\right)\tau}.$$

Hence,

$$\begin{aligned} \int_0^T E(S_\tau)d\tau &= S_0\int_0^T e^{\left(c+\frac{d}{2}\right)\tau}d\tau \\ &= \frac{S_0}{(c+\frac{d}{2})}\left(e^{\left(c+\frac{d}{2}\right)T}-1\right) < \infty. \end{aligned}$$

Thus, from (5.1.5) and (5.1.6), $g(\tau)(\omega)$ is Lebesgue integrable and we have $|\eta^n| \le g$. Hence, using Dominated Convergence Theorem, we can see

$$\begin{aligned} \lim_n \psi^n(t,s,i,y) &= \lim_n E[\int_t^T \exp\left(-\int_t^\tau r(X_u)du\right)\eta^n d\tau | S_t = s, X_t = i, Y_t = y] \\ &= E[\int_t^T \exp\left(-\int_t^\tau r(X_u)du\right)\lim_n \eta^n d\tau | S_t = s, X_t = i, Y_t = y] \\ &= 0. \end{aligned}$$

So, $\psi^n \to 0$ as $n \to \infty$, point-wise. Therefore, $\hat{\phi}^n$ converges to ϕ point-wise, as $n \to \infty$. $\square$

5.2 Computation of price function

To compute the modified (replacing the hazard rate through its estimator) (3.3.1)-(3.3.3) numerically, the MLE $\hat{\lambda}_i(y, \tau)$ has to be interpolated. Since this functional parameter should satisfy differentiability, we adopt cubic spline as we know the existence,uniqueness and convergence of spline. As the PDE that involves European call option price function, can be computed numerically, but to do this one needs to fit a spline on the estimated hazard rate and then the nature of the approximated price function can be illustrated.

Appendices

Appendix A

Table1

Table A.1: The following is the table for the graph for estimated hazard rate along with theoretical, with given ($i = 1, j = 3$), generated [see chapter 4] through estimation of transition rate function[16].

I_k	$\hat{\lambda}_1(y)$	$\lambda_1(y)$
0	.100	0
0.25	.358	0.2
0.5	.380	0.333333333
0.75	.580	0.428571429
1	.596	0.5
1.25	.629	0.555555556
1.5	.688	0.6
1.75	.662	0.636363636
2	.795	0.666666667
2.25	.850	0.692307692
2.5	.853	0.714285714
2.75	.861	0.733333333
3	.815	0.75
3.25	.777	0.764705882
3.5	1.03	0.777777778
3.75	.871	0.789473684

Bibliography

[1] Basak, G., Ghosh, M., & Goswami, A., Risk minimizing option pricing for a class of exotic options in a Markov-modulated market. Stochastic Analysis and Applications, 29(2011), 259-281.

[2] Billingsley, P., Statistical inference for Markov process. The University of Chicago Press, 1961.

[3] Colvert, R.E. & Boardman, T., Estimation in the piece-wise constant hazard rate model, Comm. Statist.-Theor. Meth. A5 (1976), 1013-1029.

[4] Deshpande, A., Ghosh, M., Risk minimizing option pricing in a regime switching market (English summary), Stoch. Anal. Appl. 26 (2008), no. 2, 313–324.

[5] Dvoretzky, A., Kiefer, J. & Wolfowitz, J., Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator, Ann. Math. Statist. 27 (1956), 642-669.

[6] Föllmer, H. Schweizer, M., Hedging of contingent claims under incomplete information, Applied stochastic analysis (London, 1989), 389–414, Stochastics Monogr., 5, Gordon and Breach, New York, 1991.

[7] Föllmer, H., Sondermann, D., Hedging of non-redundant contingent claims in Contribution to Mathematical Economics (1986), 205-223.

[8] Ghosh, M., Goswami, A., Risk minimizing option pricing in a semi-Markov modulated market, SIAM J. Control Optim. 48 (2009), no. 3, 1519–1541.

[9] Goswami, A., Patel, J., Shevgaonkar, P., The optimal hedging in a semi-Markov modulated market, 2014. Arxive.

[10] Greenwood, P.E. & Wefelmeyer, W., Empirical estimators for semi-Markov processes, Math. Meth. of Satatist. 5(3) (1996), 299-315.

[11] Gut, A., Stopped Random Walks: Limits Theorems and Applications. Springer-Verlag, 1988.

[12] Hunt, J., Devolder, P., Semi-Markov regime switching interest rate models and minimal entropy measure, Physica A: Statistical Mechanics and its Applications 390, 15(2011), 3767-3781.

[13] Leunberger, D., Investment Science, Second Edition, Oxford University Press, 2013.

[14] Miller, R.G., Jr, Early failures in life testing, J. Amer. Statist. Assoc. 55 (1960), 491-502.

[15] Moore, E.H., Pyke, R., Estimation of the transition distributions of a Markov renewal process. Ann. Inst. Statist. Math., 20, 411– 424 (1968).

[16] Ouhbi, B. Limnios, N., Nonparametric estimation for semi-Markov processes based on its hazard rate functions. Statistical Inference for Stochastic Processes, 2 (2), 151–173 (1999).

[17] Ouhbi, B., Limnios, N., Nonparametric estimation for semi-Markov kernels with application to reliability analysis, Appl. Stochast. Mod. and Data Anal. 12(4) (1996), 209-220.

[18] Protter, P., Stochastic Integration and Differential Equations, Springer, 21 (2005).

[19] Pyke, R., Schaufele, R., Limit theorems for Markov renewal processes. Ann. Math. Statist., 35, 1746–1764 (1964).

[20] Schweizer, M., Risk-minimality and orthogonality of martingales, Stoch. Process. Appl. 30(1990), 123-131.

[21] Schweizer, M., Option hedging for semi-martingales, Stochastic Processes and their Applications (1991), 339-363.

[22] Schweizer, M., Martingale densities for general asset prices, J. Math. Econ. 21(1992), 363-378.

[23] Schweizer, M., A guided tour through quadratic hedging approaches(2001). In E.Jouini, J.Cvitanić, & M.Musiela (Eds.), Option pricing interest rates and risk management (pp. 538-574). Cambridge: Cambridge University Press.