Extended Local Volatility Modeling

Steven E. Shreve
Carnegie Mellon University

Joint work with
Gerard Brunick
University of California
Santa Barbara

To appear in
Annals of Applied Probability

School of Mathematics
University of Minnesota
March 29, 2013

Outline

- Markov projection (Krylov/Gyöngy)
- First application: Index options
- Second application: Mixture models
- Extension of Markov projection
- Main result

Markov projection

Theorem (Gyöngy (1986); see also Krylov (1984))

Suppose

$$dX(t) = \mu_t \, dt + \sigma_t \, dW(t),$$

where μ_t *and* σ_t *are adapted random processes and* $W(t)$ *is a Brownian motion. Define (nonrandom) functions*

$$\widehat{\mu}(t,x) = \mathbb{E}\big[\mu_t \big| X(t) = x\big], \quad \widehat{\sigma}(t,x) = \Big(\mathbb{E}\big[\sigma_t^2 \big| X(t) = x\big]\Big)^{1/2}.$$

Then there exists a solution of the stochastic differential equation

$$dY(t) = \widehat{\mu}\big(t, Y(t)\big) \, dt + \widehat{\sigma}\big(t, Y(t)\big) \, dW(t)$$

with initial condition $Y(0) = X(0)$ *such that at each fixed time* t*,* $Y(t) \stackrel{\mathcal{D}}{=} X(t)$.

Comments on Gyöngy's Theorem

- When the stochastic differential equation for Y has a unique solution, this solution is Markov. Y is called the Markov projection of X.
- Within the context of finance, $\widehat{\sigma}(t, x)$ is the local volatility surface of Dupire.
- Because $Y(t) \stackrel{\mathcal{D}}{=} X(t)$ for each fixed t, European calls on X have the same prices as the corresponding calls on Y.
- Path-dependent options on X do not have the same prices as the corresponding path-dependent options on Y.
- Gyöngy's Markov projection theorem requires that σ_t is uniformly bounded away from zero and from above, and hence $\widehat{\sigma}(t, x)$ is unformly bounded away from zero and from above.
- Within the context of finance, Gyöngy does not apply to many stochastic volatility models, e.g., Heston.

First application of Markov projection: Index options

- Individual stocks: $dS_i(t) = \sigma_i S_i(t)\, dW_i(t)$, $i = 1, \ldots, n$
- Index: $S(t) = \sum_{i=1}^n w_i S_i(t)$
- Evolution of index: $dS(t) = \sum_{i=1}^n w_i \sigma_i S_i(t)\, dW_i(t)$
- Markov projection:
  $\widehat{\sigma}(t,s) = \left(\mathbb{E}\left[\sum_{i,j=1}^n w_i w_j \rho_{ij} \sigma_i \sigma_j S_i(t) S_j(t) \middle| S(t) = s\right]\right)^{1/2}$
- Local volatility model for index:

$$dS^{\ell v}(t) = \widehat{\sigma}\big(t, S^{\ell v}(t)\big)\, dW(t)$$

- European calls on S have the same prices as European calls on $S^{\ell v}$.
- See Piterbarg, "Markovian projection for volatility calibration," *Risk* 20(4), 2007, for an approximation of the local volatility surface.

Second application of Markov projection: Mixture models

The Black-Scholes model assumes that the underlying asset price has a log-normal distribution under a risk-neutral probability measure at the option expiration date T. It has been proposed to instead assume that the distribution is a mixture of log-normals.

- Empirical reason: Mixture of two log-normals fits the volatility smile.
- Computational reason: The mixture of log-normals gives prices and Greeks that are mixtures of Black-Scholes prices and Greeks.

A mixture of models is not a model.

A model describes the evolution of the underlying asset price, not just its risk-neutral distribution at the final time T.

- If we don't model the evolution, we cannot build successful trading strategies. Trading strategies generate profits and losses over time.
- If we don't model the evolution, we cannot price path-dependent options. The price of a path-dependent option depends on the joint distribution of the underlying asset at multiple time points.

A naive mixture model

Assume

$$0 < v_1 < v_2.$$

Consider a "model" with

$$dS_t = \sigma_0 S_t \, dW_t, \quad 0 \le t \le T,$$

where

$$\mathbb{P}\{\sigma_0^2 = v_1\} = \frac{1}{2}, \quad \mathbb{P}\{\sigma_0^2 = v_2\} = \frac{1}{2}.$$

We set the value of σ_0 at time zero, and then the risk-neutral distribution of $S(T)$ is a mixture of log-normals with volatilities $\sqrt{v_1}$ and $\sqrt{v_2}$.

Immediately after time zero, we can determine σ_0 from the observed returns, and we no longer have a mixture model.

A less naive mixture model

At time 0 choose a volatility σ_0 with

$$\mathbb{P}\{\sigma_0^2 = v_1\} = \frac{1}{2}, \quad \mathbb{P}\{\sigma_0^2 = v_2\} = \frac{1}{2}.$$

Use this volatility throughout to obtain a process S.

Choose a time-partition

$$\Pi : 0 = T_0 < T_1 < T_2 < \cdots < T_n = T.$$

At each time T_i, compute

$$p_1(T_i, s) = \mathbb{P}\{\sigma_0^2 = v_1 | S_{T_i} = s\}, \quad p_2(T_i, s) = \mathbb{P}\{\sigma_0^2 = v_2 | S_{T_i} = s\}.$$

Construct a second process S^Π recursively. On $[T_0, T_1)$ use the volatility σ_0 chosen above. At each subsequent time T_i, redraw the volatility according to

$$\mathbb{P}\{(\sigma^\Pi_{T_i})^2 = v_1\} = p_1(T_i, S^\Pi_{T_i}), \quad \mathbb{P}\{(\sigma^\Pi_{T_i})^2 = v_2\} = p_2(T_i, S^\Pi_{T_i}),$$

and use it on $[T_i, T_{i+1})$.

Relationship between S and S^{Π}

- We set $\sigma_0^{\Pi} = \sigma_0$.
- We use volatility σ_0 to generate S_t, $0 \le t \le T_1$.
- We use volatility σ_0^{Π} to generate S_t^{Π}, $0 \le t \le T_1$.
- Therefore, $S_t = S_t^{\Pi}$ for $0 \le t \le T_1$.
- At time T_1, we choose a new $\sigma_{T_1}^{\Pi}$ so that $(S_{T_1}, \sigma_0) \overset{\mathcal{D}}{=} (S_{T_1}^{\Pi}, \sigma_{T_1}^{\Pi})$.
- We use volatility $\sigma_{T_1}^{\Pi}$ to continue S_t^{Π}, $\quad T_1 \le t \le T_2$.
- Therefore, $(S_t, \sigma_0) \overset{\mathcal{D}}{=} (S_t^{\Pi}, \sigma_t^{\Pi})$, $T_1 \le t \le T_2$.
- At time T_2, we choose a new $\sigma_{T_2}^{\Pi}$ so that $(S_{T_2}, \sigma_0) \overset{\mathcal{D}}{=} (S_{T_2}^{\Pi}, \sigma_{T_2}^{\Pi})$.
- We use volatility $\sigma_{T_2}^{\Pi}$ to continue S_t^{Π}, $T_2 \le t \le T_3$.
- Therefore, $(S_t, \sigma_0) \overset{\mathcal{D}}{=} (S_t^{\Pi}, \sigma_t^{\Pi})$, $T_2 \le t \le T_3$.

Properties of S^Π

- For each fixed t, $S_t \stackrel{\mathcal{D}}{=} S_t^\Pi$, and so . . .
- European calls on S have the same prices as European calls on S^Π.
- S^Π has piecewise constant volatility.
- Immediately after each T_i, observation of S^Π reveals the volatility being used on $[T_i, T_{i+1})$, but not the volatilities that will be used after time T_{i+1}.

A local volatility model as the limit

Recall

$$0 = T_0 < T_1 < T_2 < \cdots < T_n = T.$$

Let $n \to \infty$ so that $\max_i |T_{i+1} - T_i| \to 0$.

S^Π converges to a process $S^{\ell v}$ satisfying

$$dS_t^{\ell v} = \widehat{\sigma}(t, S_t^{\ell v}) S_t^{\ell v} dW_t, \quad 0 \le t \le T,$$

where

$$\widehat{\sigma}(t,s) = \left(\mathbb{E}\left[\sigma_0^2 \middle| S_t = s\right]\right)^{1/2} = \left(\frac{v_1 \pi_1(t,s) + v_2 \pi_2(t,s)}{\pi_1(t,s) + \pi_2(t,s)}\right)^{1/2}$$

and $\pi_i(t,s)$ is the log-normal distribution corresponding to time t and volatility $\sqrt{v_i}$.

Comments

- This formula for the local volatility surface for a mixture model was worked out by Brigo and Mercurio, “A mixed-up smile,” *Risk*, September 2000.
- The formula is a special case of Gyöngy’s Markov projection theorem.
- The derivation of the formula given here provides a new proof technique that allows us to remove Gyöngy’s conditions that the diffusion coefficient must be uniformly bounded away from zero and bounded from above.

Extensions of Markov projection

Corollary (of the Main Result)

Assume

$$dS_t = \sigma_t S_t dW_t, \quad 0 \le t \le T,$$

where σ_t *is a random process. Define*

$$M_t \triangleq \max_{0 \le u \le t} S_u.$$

Then there exists a "local volatility" function $\widehat{\sigma}(t, s, m)$ *and there exists a solution of the stochastic differential equation*

$$dS_t^{\ell v} = \widehat{\sigma}(t, S_t^{\ell v}, M_t^{\ell v}) S_t^{\ell v} \, dW_t,$$

where

$$M_t^{\ell v} \triangleq \max_{0 \le u \le t} S_u^{\ell v},$$

such that for each fixed $t \ge 0$, *we have* $(S_t^{\ell v}, M_t^{\ell v}) \stackrel{\mathcal{D}}{=} (S_t, M_t)$.

Comments on extended Markov projection

- The extended local volatility surface is

$$\widehat{\sigma}(t, s, m) = \left(\mathbb{E}\left[\sigma_t^2 \middle| S_t = s, M_t = m\right]\right)^{1/2}.$$

- There are versions of the corollary for the minimum stock price, the average stock price, and combinations of these, so ....
- the corollary applies to path-dependent options whose payoff depends on the maximium, minimum or average stock price, e.g., barrier and Asian options.
- The proof follows the approach outlined for the mixture model.
- Disclaimer: We need to assume that σ_t is bounded away from zero in order to guarantee that the local volatility stochastic differential equation has a unique solution.

Example of nonuniqueness

Let $X_0 = 0$ and $dX(t) = \sigma_t \, dW_t$, where

$$\sigma_t = I_{(1,\infty)}(t) I_{\{W_1 > 0\}}.$$

The solution is

$$X_t = I_{(1,\infty)}(t) I_{\{W_1 > 0\}} (W_t - W_1).$$

We have $\widehat{\sigma}(t,x) = 0$ for $0 \le t \le 1$, and for $t > 1$,

$$\widehat{\sigma}^2(t,x) = \mathbb{E}[\sigma_t^2 | X_t = x] = \begin{cases} 1, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Both $X_t^{(1)} \equiv 0$ and

$$X^{(2)}(t) = I_{(1,\infty)}(t)(W_t - W_1)$$

are solutions of $dX_t^{\ell v} = \widehat{\sigma}(t, X_t^{\ell v}) dW_t$. The solution we want is $X^{(1)}$ with probability $\frac{1}{2}$ and $X^{(2)}$ with probability $\frac{1}{2}$.

The Main Result

Let C^d denote the space of continuous functions from $[0, \infty)$ to $\mathbb{R}^d$.

Define three operators mapping $C^d \times [0, \infty)$ to C^d:

- Shift operator: $\Theta(x, t) \triangleq x(t + \cdot)$,
- Stopping operator: $\nabla(x, t) \triangleq x(t \wedge \cdot)$,
- Difference operator: $\Delta(x, t) \triangleq x(t + \cdot) - x(t)$.

We say $\Phi\colon C^d \to C^d$ is an updating function if

- Initiation: $\Phi_0(x) = x(0)$,
- Non-anticipativity: $\nabla(\Phi(x), t) = \nabla\Big(\Phi\big(\nabla(x, t)\big), t\Big)$,
- "Markov" property: $\Theta\big(\Phi(x), t\big) = \Phi\big(\Phi_t(x) + \Delta(x, t)\big)$.

Theorem

Given

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s, \quad t \geq 0,$$

where $\mathbb{E} \int_0^t \left(\|\mu_s\| + \|\sigma_s \sigma_s^T\| \right) ds < \infty$ *for all* $t \geq 0$. *Let* $Z = \Phi(X)$. *For Lebesgue-almost-every* t, *there are versions*

$$\widehat{\mu}(t, Z_t) = \mathbb{E}[\mu_t | Z_t], \quad \widehat{\sigma}(t, Z_t) \widehat{\sigma}^T(t, Z_t) = \mathbb{E}[\sigma_t \sigma_t^T | Z_t],$$

and a weak solution

$$\begin{aligned} \widehat{X}_t &= \widehat{X}_0 + \int_0^t \widehat{\mu}(s, \widehat{Z}_s) \, ds + \int_0^t \widehat{\sigma}(s, \widehat{Z}_s) \, dW_s, \\ \widehat{Z} &= \Phi(\widehat{X}), \end{aligned}$$

such that $\widehat{Z}_t \stackrel{\mathcal{D}}{=} Z_t$ *for every* $t \geq 0$.