Mathematics for New Technologies in Finance

Solution sheet 9

Exercise 9.1 (Local volatility and local stochastic volatility models) Assume to be working under the pricing measure.

(a) In Exercise 7.4, we have introduced the class of local volatility model

$$\mathrm{d}S_t = \sigma(t, S_t)S_t \mathrm{d}W_t,$$

and proved Dupire's formula. Is it true that a local volatility model using σ_{Dupire} can be perfectly calibrated to all the European call options in the market? Is the same true for the classical Black-Scholes model?

(b) Consider now the stochastic local volatility model

$$\mathrm{d}S_t = \alpha_t l(t, S_t)S_t \mathrm{d}W_t,$$

which includes a local volatility component in a stochastic volatility model. The process α is typically a diffusion process correlated with S, whose dynamics varies from model to model. Prove that, if such a model is required to perfectly calibrate European call options, it must satisfy

$$l(t, x) = \frac{\sigma_{\text{Dupire}}(t, x)}{\sqrt{\mathbb{E}[\alpha_t^2 | S_t = x]}}, \quad (t, x) \in [0, +\infty)^2. \tag{1}$$

Hint: Have a look at Gyöngy's mimicking Theorem [1]

Solution 9.1

(a) Dupire's formula ensures that we can build a Markovian model which can perfectly calibrate any European call option, in the sense that, if one chooses as $\sigma(t, x)$ the map σ_{Dupire} derived from Dupire's formula and defines the local volatility model

$$\mathrm{d}S_t^{\text{Dup}} = \sigma_{\text{Dupire}}(t, S_t^{\text{Dup}})S_t^{\text{Dup}} \mathrm{d}W_t,$$

it holds that

$$c(T, K) = \mathbb{E}[(S_T^{\text{Dup}} - K)^+], \quad \forall (T, K) \in [0, +\infty)^2, \tag{2}$$

where $c(T, K)$ is the price of an European call from the market. This is a direct consequence of Breeden-Litzenberger formula (see Exercise 7.3): the knowledge call option prices provides us exactly with the knowledge of the marginal distribution of any model S_t which has to be perfectly calibrated, for any $t \geq 0$. Hence, S^{Dup} is the only local volatility model perfectly calibrated to all the European call options on the market. It should be noted that this does not imply that S^{Dup} is, in general, the only model able to perfectly calibrate all the European calls. Indeed, the knowledge of the marginals does not imply the knowledge of the full distribution of S. Hence, there could exists other models, non Markovian, with the exact same marginals, and hence perfectly able to replicate call options as in 2.
Classical Black-Scholes model can not replicate all European calls: a simple way to motivate this is to notice that BS is a special case of local volatility dynamics, where however σ is a constant and does not (in real-world scenarios) coincide with σ_{Dupire}. Broadly speaking, $\sigma(t, x)$ needs to be a whole function, not a fixed constant, in order to have the required calibrating ability.

(b) We have highlighted above that any model which aims to perfectly calibrate any European call, in the sense of (2), must have the same marginal distribution of S_t^{Dup}, for any $t \geq 0$. The question on the existence of a Markovian diffusion process with the same marginals of a given process S was solved by Gyöngy's mimicking Theorem, [1]: the volatility $A(t,x)$ of the Markovian mimicking process has to satisfy

$$A(t,x) = \sqrt{\mathbb{E}[\alpha_t^2 l(t,S_t)^2 S_t^2 | S_t = x]}, \quad (t,x) \in [0,+\infty)^2.$$

We have already highlighted that the only Markovian model able to perfectly calibrate on European calls is S^{Dup}. Hence, it must hold that $A(t,x) = x\sigma_{\mathrm{Dupire}}(t,x)$, for any $(t,x) \in [0,+\infty)^2$. The theorem prescribes therefore that the following, additional condition must hold for S.

$$x^2\sigma^2_{\mathrm{Dupire}}(t,x) = \mathbb{E}[\alpha_t^2 l(t,S_t)^2 S_t^2 | S_t = x] = l(t,x)^2 x^2 \mathbb{E}[\alpha_t^2 | S_t = x], \quad \forall (t,x) \in [0,+\infty)^2,$$

which directly implies 1. In essence, this relationship tells us that one has "only" to project the volatility process α on the information generated by the current value S_t of the asset in order to obtain perfect calibration. Finally, it should be noted that the newly generated SDE

$$\mathrm{d}S_t = \frac{\alpha_t}{\sqrt{\mathbb{E}[\alpha_t^2 | S_t]}} \sigma_{\mathrm{Dupire}}(t,S_t) S_t \mathrm{d}W_t, t \in [0,T],$$

is inherently hard, being a non-linear SDE with McKean–Vlasov dynamics. Even ensuring the existence of a unique solution to such an equation is a very hard task, for general volatility α.

Exercise 9.2 (Calibrating local stochastic volatility models via NNs) In Exercise 1 we have presented the theoretical material required to calibrate a LSV model to the price of European call options. One should however notice that computing $l = l(t,x)$ is not as straightforward as in a local volatility model: one has to use particle methods to deal with the derived McKean-Vlasov SDE. A different approach, inspired by the deep-hedging methodology, is outlined in lecture notebook 6. Taking inspiration from that, write a code which calibrates a LSV model to the price of European calls via neural networks.

Solution 9.2 See solution notebook 1.

Exercise 9.3 (Stratonovich-Taylor expansion) Consider the scalar stochastic differential equation

$$X_t = X_0 + \sum_{i=0}^{d} \int_0^t V_i(X_s) \circ \mathrm{d}W_s^i, \tag{3}$$

where by notation $\mathrm{d}W_t^0 = \mathrm{d}t$, $\{W^1, \dots, W^d\}$ are independent Brownian motions, the integrals are taken in the Stratonovich sense and $\{V^0, \dots, V^d\}$ are $\mathcal{C}^2$.

(a) Consider a scalar, $\mathcal{C}^2$ function f. Derive from Ito's formula that

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \circ \mathrm{d}X_s := f(X_0) + \sum_{i=0}^{d} \int_0^t f'(X_s) V_i(X_s) \circ \mathrm{d}W_s^i.$$

(b) Assume that f is $\mathcal{C}^\infty$. Define the transport operator $V_i f(x) := f'(x)V_i(x)$, for any $x \in \mathbb{R}$. Define $\mathbf{W} = (W^0, W^1, \dots, W^d)$. Prove more in general that for any $M \in \mathbb{N}$,

$$f(X_t) = f(X_0) + \sum_{k=1}^{M} \sum_{\substack{w \in W \\ |w| = k}} V_{i_1} \dots V_{i_k} f(X_0) \mathbf{Sig}^w_{[0,t]}(\mathbf{W}) + R_{M+1}(t,f), \tag{4}$$

where R_{M+1} is a residual term of order $M+1$ (dependent on f and t), $w = (i_1, \dots, i_k)$ and $W = \{0, \dots, d\}^k$.

(c) Consider the explicit dynamics

$$Y_t = Y_0 + \int_0^t \circ \mathrm{d}B_s,$$

and $f(y) = y^p$, $p \in \mathbb{N}$. Compute explicitly $f(Y_t)$ at the second order and provide the general form.

(d) Do as in the previous point, with

$$Y_t = Y_0 + \int_0^t aY_s \mathrm{d}s + \int_0^t bY_s \circ \mathrm{d}B_s,$$

where $\{a, b\} \in \mathbb{R}^2$ are constants.

Solution 9.3

(a) Let h be a $\mathcal{C}^2$ function; we recall first that Ito and Stratonovich integrals can be related to each other as follows:

$$\int_0^t h(X_s) \circ \mathrm{d}B_s = \int_0^t h(X_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t h'(X_s)h(X_s)\mathrm{d}s, \tag{5}$$

where the integral on the right-hand side is taken in Ito's sense. Hence, we can rewrite

$$X_t = X_0 + \int_0^t V_0(X_s)\mathrm{d}s + \sum_{i=1}^d \int_0^t V_i(X_s)\mathrm{d}W_s^i + \frac{1}{2}\sum_{i=1}^d \int_0^t V_i'(X_s)V_i(X_s)\mathrm{d}s,$$

Applying Ito's formula, we obtain

$$\begin{aligned} f(X_t) =& f(X_0) + \int_0^t f'(X_s)V_0(X_s)\mathrm{d}s + \frac{1}{2}\sum_{i=1}^d \int_0^t f'(X_s)V_i'(X_s)V_i(X_s)\mathrm{d}s \\ &+ \sum_{i=1}^d \int_0^t f'(X_s)V_i(X_s)\mathrm{d}W_s^i + \frac{1}{2}\sum_{i=1}^d \int_0^t f''(X_s)(V_i)^2(X_s)\mathrm{d}s. \end{aligned}$$

One has simply to recognize that the relationship between Ito and Stratonovich integral implies that, for any $i \in \{1, \dots, d\}$,

$$\int_0^t f'(X_s)V_i(X_s) \circ \mathrm{d}W_s^i = \int_0^t f'(X_s)V_i(X_s)\mathrm{d}W_s^i + \frac{1}{2}\int_0^t [f'(X_s)V_i'(X_s) + f''(X_s)V_i(X_s)]V_i(X_s)\mathrm{d}s$$

to conclude. Let us now comment briefly on the differences between Ito and Stratonovich integrals.

- Stratonovich integral, contrary to Ito, offers a Taylor-like expansion, which does not require correction terms. This property is actually true as long as the integrand does not have jumps. For instance, in case of a generic Lévy process with jumps, even Stratonovich integral requires some correction terms in a similar expansion.
- Stratonovich integral does not preserve the martingale (or local martingale) property, even when the integrator is a Brownian motion, as we immediately see from (5). This is due to the fact that, in its infinitesimal derivation, Stratonovich integral is forward looking, in the sense that integral at time t requires looking at time $t + \epsilon$, for $\epsilon > 0$ small. Such a property, which is very relevant for simulations, is not indicate for financial applications. Notice that in the limit, this forward-looking property is not present, as we see from (5). For this reason, Ito's integral, whose integral does not look in the future, is much more commonly used in Financial Mathematics.

- Stratonovich integral requires more integrability in the integrand, compare to Ito, where only predictability is needed. In this sense, Ito's integration theory is applicable to a wider class of processes.

(b) This is a direct consequence of the previous result, simply iterated. We have first

$$f(X_t) = f(X_0) + \sum_{i=0}^{d} \int_0^t V_i f(X_s) \circ \mathrm{d}W_s^i.$$

One can now expand the term $V_i f(X_s) = V_i f(X_0) + \sum_{j=0}^d \int_0^s V_j V_i f(X_u) \circ \mathrm{d}W_u^j$, leading to

$$f(X_t) = f(X_0) + \sum_{i=0}^{d} V_i f(X_0) \int_0^t \circ \mathrm{d}W_s^i + R_2(t,f),$$

where $R_2(t,f) := \sum_{i=0}^d \sum_{j=0}^d \int_0^t \int_0^s V_j V_i f(X_u) \circ \mathrm{d}W_u^j \circ \mathrm{d}W_s^i$. Iteratively,

$$f(X_t) = f(X_0) + \sum_{k=1}^{M} \sum_{\substack{w \in W \\ |w|=k}} V_{i_1} \dots V_{i_k} f(X_0) \int_0^t \dots \int_0^{t_2} \circ \mathrm{d}W_{t_1}^{i_1} \dots \circ \mathrm{d}W_{t_k}^{i_k} + R_{M+1}(t,f),$$

$0 \le t_1 \le \ldots \le t_l \le t$, which is exactly the required equation.

(c) By point (a),

$$Y_t^p = Y_0^p + p \int_0^t Y_s^{p-1} \circ \mathrm{d}B_s = Y_0^p + pY_0^{p-1} \int_0^t \circ \mathrm{d}B_s + p(p-1) \int_0^t \int_0^s Y_u^{p-2} \circ \mathrm{d}B_u \circ \mathrm{d}B_s$$

And, in general, if we assume $p \ge M$,

$$Y_t^p = Y_0^p + \sum_{k=1}^{M} \frac{p!}{(p-k)!} Y_0^{p-k} \mathbf{Sig}_{[0,t]}^{(1)^{\otimes k}}(B) + R_{M+1}(t,f). \tag{6}$$

Notice that in this case $W = \{1\}^d$, and hence for each length k, there is only one possible word, which we indicate as $(1)^{\otimes k}$. As a comment, let us highlight that this formula was indeed expected: by (5), $\int_0^t \circ \mathrm{d}B_s = B_t$. Hence, we can actually compute

$$Y_t^p = (Y_0 + B_t)^p = Y_0^p + \sum_{k=1}^{M} \frac{p!}{(p-k)!\,k!} Y_0^{p-k} B_t^k.$$

One exactly recovers (6) since it is possible to iteratively prove that

$$\mathbf{Sig}_{[0,t]}^{(1)^{\otimes k}}(B) = \frac{B_t^k}{k!},$$

when the integral is taken in Stratonovich sense.

(d) Again by point (a),

$$\begin{aligned}
Y_t^p =& Y_0^p + \int_0^t apY_s^p \mathrm{d}s + \int_0^t bpY_s^p \circ \mathrm{d}B_s \\
=& Y_0^p + apY_0^p \int_0^t \mathrm{d}s + bpY_0^p \int_0^t \circ \mathrm{d}B_s \\
&+ (ap)^2 \int_0^t \int_0^s Y_u^p \mathrm{d}u\mathrm{d}s + abp^2 \int_0^t \int_0^s Y_u^p \circ \mathrm{d}B_u \mathrm{d}s \\
&+ abp^2 \int_0^t \int_0^s Y_u^p \mathrm{d}u \circ \mathrm{d}B_s + (bp)^2 \int_0^t \int_0^s Y_u^p \circ \mathrm{d}B_u \circ \mathrm{d}B_s.
\end{aligned}$$

More generally, we can write

$$Y_t^p = Y_0^p + \sum_{k=1}^{M} \sum_{\substack{w \in W \\ |w|=k}} p^k a^{\#0(w)} b^{k-\#0(w)} Y_0^p \mathbf{Sig}^w_{[0,t]}(\mathbf{W}) + R_{M+1}(t,f).$$

In this case, $W = \{0,1\}^d$, and $\#0(w)$ counts the number of zeros in the word w.

Exercise 9.4 (Signatures and reservoirs computing)

(a) Consider a process X with dynamics (3). The same assumptions of Exercise 9.3 are made. Prove that

$$X_t = \langle \mathbf{R}, \mathbf{Sig}_{[0,t]}(\mathbf{W})\rangle X_0, \tag{7}$$

and express the readout $\mathbf{R}$ in terms of $(V_i)_{i=0}^d$.

(b) Relate (7) with reservoirs computing.

Solution 9.4

(a) This is an immediate consequence of equation (4). We simply set $f = \mathrm{id}$ and send $M \to \infty$, to obtain

$$X_t = X_0 + \sum_{k=1}^{\infty} \sum_{\substack{w \in W \\ |w|=k}} V_{i_1} \dots V_{i_k} X_0 \mathbf{Sig}^w_{[0,t]}(\mathbf{W}),$$

which is exactly the requested equation, written in a more explicit form.

(b) The solution of any stochastic differential equation can be express via an universal, dynamical reservoir, namely $\mathbf{Sig}_{[0,t]}(\mathbf{W})$, dependent only on the driving path $\mathbf{W}$ (but not on X itself) and a linear readout $\mathbf{R}$. The latter depends on X_0 and on the vector field $(V_i)_{i=0}^d$. This is very similar to classical reservoir computing, where one fixes the reservoir (usually a large, random recurrent neural network) and then only train a single linear layer that reads off the quantities of interest. The reservoir $\mathbf{Sig}_{[0,t]}(\mathbf{W})$ is also "universal", in the sense that changing the readout $\mathbf{R}$ (or equivalently the vector field $(V_i)_{i=0}^d$), one can recover the dynamics of any SDE X driven by $\mathbf{W}$.

Exercise 9.5 (Indifference pricing via deep hedging) We aim to compute the indifference price of an European call option in Heston model. There are at least 2 options: (a) Montecarlo runs, (b) deep-hedging approach. Produce a code comparing the two approaches.

Solution 9.5 See solution notebook 2.

References

[1] István Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an ito differential. *Probability Theory and Related Fields*, 71:501–516, 1986.