ullint5b.tex pm Wed 27.2.2019

10. Monte-Carlo pricing of swaptions with LMM

Recall our study of swaptions above (V.4). In the notation used there,

$$\begin{aligned} E_B[\frac{B(0)}{B(T_i)}(S_{ik}(T_i)-K)_+ \sum_{j=i+1}^{k} \tau_j P(T_i,T_j)] &= E_i[\frac{P(0,T_i)}{P(T_i,T_i)}(S_{ik}(T_i)-K)_+ \sum_{j=i+1}^{k} \tau_j P(T_i,T_j)] \\ &= [P(0,T_i)E_i[(S_{ik}(T_i)-K)_+ \sum_{j=i+1}^{k} \tau_j P(T_i,T_j)]. \end{aligned}$$

Since

$$S_{ik}(T_i) = \frac{1-\prod_{j=i+1}^{k} 1/(1+\tau_j F_j(T_i))}{\sum_{j=i+1}^{k} \tau_j \prod_{\ell=i+1}^{j} \tau_\ell 1/(1+\tau_j F_j(T_i))},$$

this depends on the joint distribution under $\mathbb{Q}_i$ of

$$(F_{i+1}(T_i), \cdots, F_k(T_i)).$$

Recall the dynamics of forward rates under $\mathbb{Q}_i$:

$$dF_k(t) = \sigma_k(t)F_k(t)\sum_{j=i+1}^{k} \frac{\rho_{k,j}\tau_j F_j}{1+\tau_j F_j(t)}dt + \sigma_k(t)F_k(t)dZ_k.$$

By Itô's formula,

$$d\log F_k = dF_k/F_k - \frac{1}{2}(dF_k)^2/F_k^2.$$

So as by above

$$dF_k(t)/F_k(t) = \sigma_k(t)\sum_{j=i+1}^{k} \frac{\rho_{k,j}\tau_j F_j}{1+\tau_j F_j(t)}dt + \sigma_k(t)dZ_k, \qquad (dF_k(t))^2 = \sigma_k(t)^2 dt,$$

this gives

$$d\log F_k(t) = \sigma_k(t)\sum_{j=i+1}^{k} \frac{\rho_{k,j}\tau_j F_j}{1+\tau_j F_j(t)}dt - \frac{1}{2}\sigma_k(t)^2 dt + \sigma_k(t)dZ_k.$$

Numerical solution of the SDE.

This SDE, like most SDEs, must be solved numerically (the two main exceptions are, fortunately, the two we meet earliest – OU and GBM). There is a whole field on numerical solutions of SDEs – a combination of numerical solution of ODEs and PDEs within Numerical Analysis, and SDEs within Probability. The standard work, for reference, is

[KP] Peter E. KLOEDEN and Eckhard PLATEN, *Numerical solutions of stochastic differential equations*, Springer, 1992.

There, one will find the *Milstein scheme* [KP, 10.3], which is a discretisation. The Milstein scheme for $\log F_k$ here gives

$$\log F_k^{\Delta t}(t+\Delta t) = \log F_k^{\Delta t}(t) + \sigma_k(t) \sum_{j=i+1}^{k} \frac{\rho_{k,j}\tau_j F_j^{\Delta t}}{1+\tau_j F_j^{\Delta t}(t)} \Delta t - \frac{1}{2}\sigma_k(t)^2 dt + \sigma_k(t)(Z_k(t+\Delta t) - Z_k(t)).$$

This leads to an approximate solution with strong convergence properties:

$$E_i[|\log F_k^{\Delta t}(T_i) - \log F_k(T_i)|] \leq C(T_i)\Delta t \qquad \forall \Delta t \leq \delta_0,$$

where $C(T_i) > 0$ is a constant. Now $(Z_k(t+\Delta t) - Z_k(t))$ is *Gaussian* and *known* (that is, its mean and variance, and so being Gaussian, its distribution, are known). So this is easy to *simulate*, by Monte Carlo (MC), below.

Monte-Carlo estimation

Assume we need to value a payoff $\Pi(T)$ depending on the realisation of different forward LIBOR rates

$$F(t) = (F_{i+1}(t), \cdots, F_k(t))^T$$

in a time-interval $t \in [0,T]$, where typically $T \leq T_i$. We have seen a particular case of $\Pi(T) = \Pi(T_i)$ as the swaption payoff. The simulation scheme above for the rates entering the payoff provides us with the Fs needed to form scenarios of $\Pi(T)$. Denote by a superscript the scenario (or path) under which a quantity is considered, and the number of paths by n_p. Then the Monte-Carlo simulated price of our payoff is

$$E[D(0,T)\Pi(T)] = P(0,T)E_T[\Pi(T)] = P(0,T).\frac{1}{n_p}\sum_{j=1}^{n_p} \Pi_j(T),$$

where the forward rates F_j entering $\Pi_j(T)$ have been simulated under the T-forward measure. We omit the T-arguments to save notation: all distributions, means, variances etc. are under the T-forward measure. However, the argument is general and extends to any other measure.

We turn to estimating the error of this MC estimate. Take a sequence Π_j of independent and identically distributed (iid) random variables, with the distribution of Π. By the Central Limit Theorem, writing SD for the standard deviation (square root of the variance), we have the convergence in distribution

$$\frac{\sum_{j=1}^{n_p}(\Pi_j - E[\Pi])}{\sqrt{n_p}SD(\Pi)} \to \Phi = N(0,1)$$

as n_p increases. So for large n_p, we have approximately

$$\frac{1}{n_p}\sum_{j=1}^{n_p}\Pi_j - E[\Pi] \sim \frac{SD(\Pi)}{\sqrt{n_p}}N(0,1),$$

in an obvious notation. So

$$\begin{aligned}\mathbb{Q}_T\Big(|\frac{1}{n_p}\sum_{j=1}^{n_p}\Pi_j - E[\Pi]| < \epsilon\Big) &= \mathbb{Q}_T(|N(0,1)| < \epsilon\sqrt{n_p}/SD(\Pi)) \\ &= 2\Phi(\epsilon\sqrt{n_p}/SD(\Pi)) - 1,\end{aligned}$$

where as usual Φ denotes the distribution function $N(0,1)$ of the standard Gaussian random variable. All this is familiar from a first course on Statistics, in particular, *confidence intervals*. From tables of Φ,

$$2\Phi(z) - 1 = 0.98 \Leftrightarrow \Phi(z) = 0.99 \Leftrightarrow z \sim 2.33.$$

So choosing

$$\epsilon = 2.33SD(\Pi)/\sqrt{n_p},$$

$E[\Pi]$ lies in the 'window' between the bounds

$$\frac{1}{n_p}\sum_{j=1}^{n_p}\Pi_j \pm 2.33SD(\Pi)/\sqrt{n_p}$$

with probability 98 %, giving a 98% confidence interval. Similarly for other confidence levels you may choose to use.

Here, the width of the window – the accuracy of our estimate – shrinks by only the *square root* of the 'sample size' (number of paths, n_p). Worse, we do not know $SD(\Pi)$, or its square, $var(\Pi)$. This is a *population variance*, an unknown parameter. We have to estimate it from the data. We can do this by using the *sample variance* $S^2(\Pi)$ to approximate it. See e.g. SMF, Ch. I.

One can do better by using *control variates*. But we must refer for this elsewhere, to a work on Simulation, or MATL484.

11. Analytical pricing of swaptions with LMM

There is an alternative method – an analytical approximation – to compute LMM swaption prices which avoids the need for Monte-Carlo. See [BDB] A. BRACE, T. DUN and G. BARTON, Towards a central interest-rate model. *Handbook in Mathematical Finance: Topics in Option Pricing, Interest Rates and Risk Management*, CUP, 2001 (Working Paper, 1998).

Recall the swap model SMM leading to Black's swaption formula:

$$dS_{ik}(t) = \sigma_{ik}(t)S_{ik}(t)dW_{ik}(t), \qquad \mathbb{Q}_{ik}.$$

A crucial role is played by the Black swap volatility component

$$\begin{aligned}\int_0^T \sigma_{ik}(t)^2 dt &= \int_0^T \sigma_{ik}(t)dW_{ik}(t).\sigma_{ik}(t)dW_{ik}(t) \\ &= \int_0^T (d\log S_{ik}(t)).(d\log S_{ik}(t)).\end{aligned}$$

We computed an analogous approximate quantity in the LMM:

$$S_{ik}(t) = \sum_{j=i+1}^{k} w_j(t)F_j(t),$$

where

$$\begin{aligned}w_j(t) &= w_j(F_{i+1}(t), \cdots, F_k(t)) \\ &= \frac{\tau_j \prod_{\ell=i+1}^{j} 1/(1+\tau_\ell F_\ell(t))}{\sum_{j=i+1}^{k} \tau_j \prod_{\ell=i+1}^{j} 1/(1+\tau_\ell F_\ell(t))}.\end{aligned}$$

Freezing

To reduce the dimensionality from uncountably infinite (as with functions of time t) to finite (as with matrix elements), we need to *approximate*. Often, it will be useful to *freeze* functions of t to their value at the beginning of the relevant time-interval. For the $S_{ik}(t)$, *freeze* the ws at time 0:

$$S_{ik}(t) = \sum_{j=i+1}^{k} w_j(t)F_j(t) \sim S_{ik}(t) = \sum_{j=i+1}^{k} w_j(0)F_j(t)$$

– variability of the ws is much smaller than variability of the Fs. This gives

$$dS_{ik} \sim \sum_{j=i+1}^{k} w_j(0)dF_j = (\cdots)dt + \sum_{j=i+1}^{k} w_j(0)\sigma_j(t)F_j(t)dZ_j(t),$$

under any of the forward measures. So using $dZ_i(t)dZ_j(t) = \rho_{ij}dt$, this gives

$$(dS_{ik}(t))^2 \sim \sum_{j,\ell=i+1}^{k} w_j(0)w_\ell(0)\sigma_j(t)\sigma_\ell(t)F_j(t)F_\ell(t)\rho_{i,j}dt.$$

As in V.10,

$$\begin{aligned}(d\log S_{ik}(t))^2 &= (dS_{ik}(t)/S_{ik}(t))^2 \\ &\sim \frac{\sum_{j,\ell=i+1}^{k} w_j(0)w_\ell(0)\sigma_j(t)\sigma_\ell(t)F_j(t)F_\ell(t)\rho_{i,j}}{S_{ik}(t)^2}dt.\end{aligned}$$

This has the interpretation of the percentage quadratic covariation. Introduce a further approximation by *freezing* all the Fs to 0, as with the ws above. This gives

$$(d\log S_{ik}(t))^2 \sim \sum_{j,\ell=i+1}^{k} \frac{w_j(0)w_\ell(0)F_j(0)F_\ell(0)\rho_{i,j}}{S_{ik}(0)^2}.\sigma_j(t)\sigma_\ell(t)dt.$$

Now take the time-average of this: we obtain

Proposition (Rebonato's formula). The time-averaged percentage variance of S is given aproximately by

$$\begin{aligned}(v_{ik}^{LMM})^2 &= \frac{1}{T}\int_0^T (d\log S_{ik}(t))^2 \\ &\sim \sum_{j,\ell=i+1}^{k} \frac{w_j(0)w_\ell(0)F_j(0)F_\ell(0)\rho_{i,j}}{S_{ik}(0)^2}.\int_0^T \sigma_j(t)\sigma_\ell(t)dt.\end{aligned}$$

Here v_{ik}^{LMM} can be used as a proxy for the Black volatility $v_{ik}(T_i)$. So, making this replacement, we can use Black's swaption formula to price swaptions *analytically* with LMM, to this degree of approximation. This works well in practice, as pointed out in e.g. [BM] and the paper [BDB].

There is an alternative, with similar accuracy in practice, due to Hull and White (1999). They differentiate the $S_{ik}(t)$ without freezing the ws. This gives an *algebraic* pricing formula, which is very quick to implement.

Instantaneous correlations: Inputs or outputs?

For swaptions, we want to match market prices (observed in the market) to model prices (which depend on (σ, ρ)). Should we infer ρ itself from swaption market quotes, or should we estimate ρ exogenously and impose it, leaving the calibration only to σ? Are the parameters in ρ inputs or outputs to the calibration?

Inputs?

We might consider a time series of past interest-rate curve data, which are observed under the real-world (objective) probability measure. This would allow us, through interpolation, to obtain a corresponding time series for the particular forward LIBOR rates being modelled in our LIBOR model. These series would be observed under the objective measure, $\mathbb{P}$. Thanks to the Girsanov theorem, this is not a problem, since instantaneous correlations are the same under $\mathbb{P}$ and $\mathbb{Q}$: considered as instantaneous covariances between driving Brownian motions in forward-rate dynamics, they do not depend on the probability measure. So, by using historical estimation (under $\mathbb{P}$), we obtain a historical estimation of the instantaneous correlation matrix ρ. This ρ, or a stylised version of it, can be considered as a given ρ for our LIBOR model, and the remaining free parameters σ are to be used to calibrate market derivatives data. In this case, calibration consists of finding the σs such that the model (caps and) swaptions prices match the corresponding market prices. In this matching procedure – often done by Optimisation (a subject in its own right!) – ρ is fixed from the start as the historical estimate found above, and we play on the volatility parameters σ to achieve our matching.

Outputs?

This second possibility considers instantaneous correlations as parameters to be fitted, like those in σ above.

Which of these two methods is preferable? See later. For now, we try to find a decent historical estimate of ρ, in case we opt for the 'inputs' approach.

12. Instantaneous correlations as Inputs: The historical matrix

It turns out (see e.g. Jäckel and Rebonato (2000)) that European swaptions are relatively insensitive to instantaneous (rather than terminal) correlations. So we may impose a good exogenously-derived instantaneous correlation matrix, if we have one, and then use volatilities to calibrate swaptions:
(a) Smoothing the rough historically estimated matrix through a parsimonious 'pivot' form (below) enjoying desirable properties may guarantee a smooth and regular behaviour of terminal correlations, and slightly more regular σ when calibrating. See Rebonato and Jäckel (1999), who propose to fit a parametric form onto the estimate.
(b) The chosen parametric forms may have particularly interesting properties typical of forward-rate correlations.
(c) Such pivot forms depend on a small number of parameters. It is always desirable to work with as few parameters as possible! Incorporating personal views, or recent changes in the market, is also easier with pivot forms.

To find our reduced-rank pivot historical-correlation matrix:
1. A market historical correlation matrix is estimated.
2. A parsimonious parametric form is chosen, and the parameters in it are estimated from the historical estimate above.
3. An angles form of the desired rank is fitted to the resulting parsimonious matrix (Rebonato's angles: see the end of V.9 above).

Historical estimation

In estimating correlations, we take into account the particular nature of LMM forward rates, characterised by a fixed *maturity*, in contrast to market quotations, where a fixed *time-to-maturity* is usually considered as time passes. We observe from the market, at different times t,

$$P(t, t + Z), P(t + 1, t + 1 + Z), \cdots, P(t + n, t + n + Z),$$

where Z ranges in a standard set of times-to-maturity. We need instead

$$P(t, T), P(t + 1, T), \cdots, P(t + n, T),$$

for the maturities T included in the tenor structure of the chosen LMM. Accordingly, a log-interpolation between discount factors is carried out and only one year of data is used (e.g., 1 Feb 2001 - 1 Feb 2002), since the first forward rate in the family expires one year after the starting date. From these daily quotations of notional ZCBs, whose maturities range from 1y to 20y from today, we extract daily log-returns of the annual forward rates involved in the model. Starting from the usual Gaussian approximation,

$$(\log(F_1(t+\Delta t)/F_1(t)), \cdots, \log(F_{19}(t+\Delta t)/F_{19}(t))) \sim N(\mu, V),$$

where the time-step $\Delta t = 1d$, our estimates of the parameters μ (population mean vector) and V (population covariance matrix) are the sample mean and sample covariance for the Gaussian variables (these are the maximum-likelihood estimators (MLEs): SMF, IV.5):

$$\hat{\mu}_i = \frac{1}{m}\sum_{k=0}^{m-1} \log(F_i(t_{k+1})/F_i(t_k)),$$

$$\hat{V}_{ij} = \frac{1}{m}\sum_{k=0}^{m-1}(\log(F_i(t_{k+1})/F_i(t_k)) - \hat{\mu}_i)(\log(F_j(t_{k+1})/F_j(t_k)) - \hat{\mu}_j),$$

where m is the number of observed log-returns for each rate. So our estimate of the general correlation coefficient ρ_{ij} is

$$\hat{\rho}_{ij} = \hat{V}_{ij}/\sqrt{\hat{V}_{ii}\hat{V}_{jj}}.$$

Principal Components Analysis (PCA: SMF III.5) reveals that, for typical data, 7 factors are required to explain 90 % of the overall variability.

Pivot matrices

We focus on an example seen earlier, ($SC3$) (Rebonato exponential). The classic approach is to fit by minimising some loss function of the difference between the two matrices – an optimisation problem.

Morini (2002) proposes instead to invert directly the functional structure of the parametric forms. We shall show how to do this for our chosen example below. Parameters are expressed as functions of key elements of the target historical matrix so as to reproduce these exactly. We call such key elements *pivot points*, or just *pivots*, of the historical matrix, and the resulting parametric matrices *pivot matrices*. The pivot approach:
(a) does not need any optimisation;
(b) with well-chosen pivots, it typically gives a matrix with the same monotonicity and positivity properties as the original one;

(c) parameters have a clear, intuitive meaning, as they are expressed in terms of correlation entries chosen as they are considered particularly significant. This allows us to alter the matrix easily by playing with the parameters in a controlled way, as might be needed in market practice.
(d) It avoids the irregularities and outliers typical of historical estimates.
(e) In our examples, the fitting error with pivots is nearly optimal.

Pivots must be chosen carefully. We consider Rebonato's exponential form:

$$\rho_{i,j} = \rho_\infty + (1 - \rho_\infty) \exp\{-|i-j|(\beta - \alpha(\max(i,j) - 1))\}.$$

Morini (2003) shows that here, the parameters satisfy the following equations:

$$\left(\frac{\rho_{1M} - \rho_\infty}{1 - \rho_\infty}\right) = \left(\frac{\rho_{M-1,M} - \rho_\infty}{1 - \rho_\infty}\right)^{M-1}$$

for ρ_∞; and for α, β,

$$\alpha = \frac{1}{2-M} \log\left(\frac{\rho_{12} - \rho_\infty}{\rho_{M-1,M} - \rho_\infty}\right),$$

$$\beta = \alpha - \log\left(\frac{\rho_{12} - \rho_\infty}{1 - \rho_\infty}\right).$$

He also considers the Schoenmaker-Coffey model ($SC3$). The pivot method applies here too. The two models are compared, for various data sets. Which is better depends on the loss function (optimisation criterion) – and indeed, on the data set, and the purpose for which the calibration is being done.

Cascade calibration

We mention briefly an alternative method, *cascade calibration*. This method also relies on an external instantaneous correlation matrix ρ, that can be estimated historically, as above. For a parametric model, such as Rebonato's exponential model above, one can fit the ρ-parameters to the historic matrix, using when applicable rank reduction by eigenvalue-zeroing, or optimisation by Rebonato's angles etc. Using the resulting historically-based ρ-matrix, one can then use the σ-parameters to fit the swaption market. Cascade calibration is a very fast and accurate calibration procedure. For a detailed study, with numerical examples, see [BM, 7.4, 7.6].

13. Smile: volatility modelling; Breeden-Litzenberger and Dupire formulae

Here we follow [BM, Part IV, Ch. 9 - 12]. Also useful is Gatheral [G].

Recall (II.3) that a caplet is like an *option* – indeed, it *is* an option, on an interest rate rather than a stock.

Recall also (MATL480) that in the Black-Scholes theory, the volatility σ is constant. Now volatility – the unknown parameter in the Black-Scholes formula, which has to be inferred from option prices in the market – is so important that it has been studied intensively. It has been observed to be *by no means constant.* A graph of (implied) volatility σ against stock-price S, or strike K, is not flat, but typically turns up at the sides, producing the 'happy face' (look at :) sideways, with the : as eyes!) of a smile. Hence the term smile for volatility modelling, variation, dynamics etc.

The situation is similar here. Recall *Black's caplet formula* (V.2):

$$Cpl(0, T_1, T_2, K) = P(0, T_2)\tau[F_2(0)\Phi(d_1) - K\Phi(d_2)],$$

where

$$d_1, d_2 = \frac{\log(F_2(0)/K) \pm \frac{1}{2}T_1 v_1(T_1)^2}{\sqrt{T_1}v_1(T_1)}.$$

We re-write this as

$$Cpl(0, T_1, T_2, K) = P(0, T_2)\tau Bl(K, F_2(0), v_2(T_1)).$$

Suppose now that we have two different strikes K_1, K_2, on two otherwise similar caplets. Can one find a *single* volatility, $v(T_1)$ say, so that *both* of

$$Cpl(0, T_1, T_2, K_1) = P(0, T_2)\tau Bl(K_1, F_2(0), v(T_1)),$$

$$Cpl(0, T_1, T_2, K_2) = P(0, T_2)\tau Bl(K_2, F_2(0), v(T_1))$$

hold? The answer is **no**. What we need are two *different* volatilities, $v(T_1, K_1)$, $v(T_1, K_2)$. That is, each caplet market price requires its *own* Black volatility

$$v^{Mkt}(T_1, K),$$

depending on the caplet strike K. So the market uses Black's formula simply as a metric to express caplet prices as volatilities.

Volatility and its measurement

As in MATL480, volatility is both vitally important (it appears explicitly in the key formulae – Black-Scholes, Black caplet, Slack swaption etc.), and intangible directly. Volatility is really about market *sentiment* – how the people involved in financial markets are feeling (particularly, how they react to changing circumstances, especially unfavourable ones). But, we can't get at this directly: it is a matter of group (and individual) psychology. So we have to rely on what we *can* get at directly: past price data (leading to *historic volatility*), and – more relevant in times of change – present price data (of options – on stock in MATL480, on interest rates here), leading to *implied volatility*.

The overall market volatility is more important and informative than volatility on individual products. So, there is a demand for some measurement of it. This led to the CBOE introducing the VIX (volatility index). This allows traders to follow overall market sentiment over time, rather as in ordinary life we look at the temperature. Indeed, there is an analogy between the two: one speaks of markets over-heating, using the language of temperature. It also allows traders to trade in options on VIX – in effect, to bet on where market sentiment is going to go, and back their judgement with money. It is arguable whether this is a good thing: some would say that there are too many exotic financial derivatives, allowing traders (or speculators) to make bets with other people's money. One could also argue that trading on VIX contributes to market instability. We note in passing the emergence of XIV (a sort of "opposite to VIX").

Be that as it may, the details of how VIX is calculated (which change over time) are of interest in that they illustrate the 'state of the art' on volatility measuring.

We will not pursue this further: although interesting, these matters are rather specialied.