Chapter 3

Calibrating the model

In this chapter we try to calibrate the model to the market datas, that is to the assets traded in the market. Tipically market datas are the prices of some coupon bonds, from which one can derive prices of zero-coupon bonds with typical maturities $T = j \cdot 3$ months, $j = 1, \ldots, N$ (for example $N = 40$, that is 10 years), and prices of some interest rate derivatives, such as futures or options on some bonds, caps, floors and swaptions.

3.1 An "implied" approach

Now we try to find the diffusion function τ implicitly from the prices of the caps in the market. We will find some general finite dimensional constraints for τ, which it must satisfy in order to price the caps.

A **cap** is the sum of several caplets, each one having a different maturity $T_i = i \cdot \theta, i = 1, \ldots, N$, where θ usually is equal to 0.25 or 0.5 (three or six months), and the same exercise rate.

A **caplet** is a claim giving the right to exchange the linear rate $R_L(T_{i-1}, \theta)$, defined by

$$1 + \theta R_L(T_{i-1}, \theta) = \frac{1}{B(T_{i-1}, T_i)}$$

with the fixed rate R^f, multiplied by θ, in the maturity T_i; hence the caplet gives the payoff

$$\left(\theta R_L(T_{i-1}, \theta) - \theta R^f\right)^+$$

at time T_i, so it has the price:

$$\text{Caplet}_i(t) = \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_i} r_u \ du}\left(\theta R_L(T_{i-1}, \theta) - \theta R^f\right)^+\right]$$

where we indicate with $\mathbb{E}^t_{\mathbb{Q}}$ the conditional expectation with respect to the probability $\mathbb{Q}$ and the σ-algebra $\mathcal{F}_t$. We notice that the caplet will be exercised if and only if it gives a positive payoff, that is if and only if we are in the so called **exercise set**

$$\mathcal{E}_i = \left\{\frac{B(T_{i-1}, T_i)}{B(T_i, T_i)} \leq \frac{1}{1 + \theta R^f}\right\}$$

We also notice that the set $\mathcal{E}_i$ is $\mathcal{F}_{T_{i-1}}$-measurable, so one can know in T_{i-1} whether to exercise the cap in T_i or not. Now, introducing the forward neutral probabilities $\mathbb{Q}_{T_{i-1}}$ and $\mathbb{Q}_{T_i}$ (see [7]), we calculate explicitly the price of a caplet:

$$\begin{aligned}
\mathrm{Caplet}_i(t) &= \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_i} r_u\, du}\left(\frac{1}{B(T_{i-1},T_i)} - (1+\theta R^f)\right)^+\right] = \\
&= \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_i} r_u\, du}\frac{1}{B(T_{i-1},T_i)}\mathbf{1}_{\mathcal{E}_i} - (1+\theta R^f)\mathbf{1}_{\mathcal{E}_i}\right] = \\
&= \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_{i-1}} r_u\, du}\mathbb{E}^{T_{i-1}}_{\mathbb{Q}}\left[e^{\int_{T_{i-1}}^{T_i} r_u\, du}\right]\frac{1}{B(T_{i-1},T_i)}\mathbf{1}_{\mathcal{E}_i}\right] - \\
&\quad -(1+\theta R^f)\mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_i} r_u\, du}\mathbf{1}_{\mathcal{E}_i}\right] = \\
&= \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_{i-1}} r_u\, du}\mathbf{1}_{\mathcal{E}_i}\right] - (1+\theta R^f)B(t,T_i)\mathbb{Q}^t_{T_i}(\mathcal{E}_i) = \\
&= B(t,T_{i-1})\mathbb{Q}^t_{T_{i-1}}(\mathcal{E}_i) - (1+\theta R^f)B(t,T_i)\mathbb{Q}^t_{T_i}(\mathcal{E}_i)
\end{aligned}$$

Now we suppose that

$$\frac{dB(t,T)}{B(t,T)} = r_t\, dt + \Gamma(t,T)\, dW_t$$

for all $T = T_i$, where W is a d-dimensional (infinite dimensional?) Brownian motion under $\mathbb{Q}$, and Γ is deterministic; this is equivalent to suppose that

$$dr_t(x) = (Ar_t(x) + c(t,x))\, dt + \tau(t,x)\, dW_t$$

with τ deterministic function of t and x such that

$$\tau(t,x) = -\frac{\partial}{\partial T}\Gamma(t,t+x), \quad \text{that is} \quad \Gamma(t,T) = -\int_0^{T-t}\tau(t,u)\, du \tag{3.1}$$

Then the price of a caplet is given by:

$$\mathrm{Caplet}_i(t) = B(t,T_{i-1})N(d_{0,i}) - B(t,T_i)(1+\theta R^f)N(d_{1,i})$$

where

$$d_{0,i} = \frac{1}{\sigma_{t,T_i}}\log\left(\frac{B(t,T_{i-1})}{(1+\theta R^f)B(t,T_i)}\right) + \frac{1}{2}\sigma_{t,T_i}, \quad d_{1,i} = d_{0,i} - \sigma_{t,T_i}$$

and

$$\sigma^2_{t,T_i} = \int_t^{T_{i-1}} \|\Gamma(s,T_i) - \Gamma(s,T_{i-1})\|^2\, ds$$

where $\|\cdot\|$ is the norm on $\mathbb{R}^d$.

We notice that this formula is very similar to the Black-Scholes formula; in particular, $\mathrm{Caplet}_i(t)$ it is an increasing function of σ_{t,T_i}. This means that if one knows the price of a caplet and all the variables in the price, except σ_{t,T_i}, one can derive it inverting the function.

The problem is that in the market caplets are not priced, but one can read only the price of caps, which are given by

$$\text{Cap}(t) = \sum_{i=1}^{N} \text{Caplet}_i(t)$$

so it is not well clear how to derive the σ_{t,T_i} for the single caplet. One method could be to ricave a volatility for each cap priced in the market, and then to make a linear interpolation for the volatility of each caplet; this method is presented in [3] as being the market practise.

Once that we have the quantities σ_{t,T_i}, we obtain a constraint for the function τ:

$$\frac{\partial}{\partial t}\sigma^2_{t,T_i} = -\|\Gamma(t,T_i) - \Gamma(t,T_{i-1})\|^2 = -\left\| \int_{T_{i-1}-t}^{T_i-t} \tau(t,u)\ du \right\|^2$$

however, these are only finite dimensional constraints, and it is still to be specified:

- the "dimension" of the model, that is the number d of Brownian motions to use
- the shape of the functions τ: for example, in [3] the authors use cubic splines, but say that the market practise is linear interpolation.

3.2 Two "historical" approaches

Now we present two historical approaches, each one based on the observation of the prices of zero-coupon bonds, which we suppose to have dynamics

$$\frac{dB(t,T)}{B(t,T)} = r_t(0)\ dt + \Gamma(t,T)\ dW_t$$

To this purpose we define $R(t,T)$ and $\tilde{R}(t,T)$ in this way:

$$B(t,T) = \exp\left(-\int_0^{T-t} r_t(u)\ du\right) = \exp\left(-(T-t)R(t,T)\right) = \exp\left(\tilde{R}(t,T)\right)$$

then, since $\tilde{R}(t,T) = \log B(t,T)$, we have:

$$d\tilde{R}(t,T) = (r_t(0) + 1/2\Gamma^2(t,T))\ dt + \Gamma(t,T)\ dW_t$$

and since $R(t,T) = \frac{\tilde{R}(t,T)}{T-t}$, we have:

$$\begin{aligned} dR(t,T) &= -\frac{\tilde{R}(t,T)}{(T-t)^2}\ dt + \frac{d\tilde{R}(t,T)}{T-t}\frac{d\langle\tilde{R}(\cdot,T)\rangle_t}{2(T-t)} = \\ &= -\frac{R(t,T)}{T-t}\ dt + \frac{r_t(0) + 1/2\Gamma^2(t,T)}{T-t}\ dt + \frac{\Gamma(t,T)}{T-t}\ dW_t \end{aligned}$$

3.2.1 With less datas

In this approach we suppose to have datas in the form $B(t,T)$, $T = j{\cdot}3$ months, $j = 1, \ldots, N$ (for example $N = 40$, that is 10 years), $t = i$ days, $i = 1, \ldots, 90$. We can get the rates in this way:

$$B(t,T) = \exp\left(-\int_t^T f(t,u)\ du\right) = \exp\left(-\int_0^{T-t} r(t,u)\ du\right)$$

$$f(t,T) = -\frac{\partial}{\partial T}\log B(t,T), \qquad r(t,x) = -\frac{\partial}{\partial T}\log B(t,t+x)$$

(we can get only approximate values)

If we put $x = T - t$, then for each day t the values of $(B(t,t+x))_x$ have a 1 day shift to the left, so if we have 90 samples we have a discrete scheme with $90 \cdot N$ values. If we interpolate each curve (which way? linear? spline? exponential?), then we could have 90 samples of forward rates' curves. Then we can use Karhunen-Loewe method to find a "good" basis, and try to remake the interpolation using Karhunen-Loewe basis...

3.2.2 With more datas

In this approach we suppose to have datas in the form $B(t,T_i)$, $T_i = i{\cdot}3$ months, $i = 1, \ldots, N$ (for example $N = 40$, that is 10 years), and datas in t with a difference in t very near (for example, if we have tick by tick datas: this datas are taken in each transaction made in the stock market, and tipically the difference between two t's near is 5-10 seconds). From these datas, we can get the volatility Γ of the bonds

$$dB(t,T) = B(t,T)(r_t\ dt + \Gamma(t,T)\ dW_t)$$

where Γ is a d-dimensional vector.

A possible way is the following: we fix the j-th day, and let t vary only into the j-th day. We suppose that $\Gamma(t,T_i) \simeq \Gamma(T_i)$ for $t \in [i \text{ day}, i{+}1 \text{ day}]$, and we pose $X_t^i = \log B(t,T_i)$, so we have

$$X_t^i = X_0^i + (r_t - 1/2\Gamma^2(T_i))t + \Gamma(T_i)W_t$$

Now we discretize $[i \text{ day}, i+1 \text{ day}]$ with a step $h = 1 \text{ day}/n$, and we pose:

$$\hat{\sigma}_n^{lm} = \sum_{i=0}^{n-1}(X_{(i+1)h}^l - X_{ih}^l)(X_{(i+1)h}^m - X_{ih}^m)$$

Then we have:

$$\begin{aligned}
\hat{\sigma}_n^{lm} &= \sum_{i=0}^{n-1}\big((r_t - 1/2\Gamma^2(T_l))h + \Gamma(T_l)(W_{(i+1)h} - W_{ih})\big) \cdot \\
&\quad \cdot \big((r_t - 1/2\Gamma^2(T_m))h + \Gamma(T_m)(W_{(i+1)h} - W_{ih})\big) = \\
&= (r_t - 1/2\Gamma^2(T_l))(r_t - 1/2\Gamma^2(T_m))\sum_{i=0}^{n-1} h^2 + \Gamma(T_l)\Gamma(T_m)\sum_{i=0}^{n-1}(W_{(i+1)h} - W_{ih})^2 +
\end{aligned}$$

$$+\left((r_t - 1/2\Gamma^2(T_l))\Gamma(T_m) + (r_t - 1/2\Gamma^2(T_m))\Gamma(T_l)\right) h \sum_{i=0}^{n-1}(W_{(i+1)h} - W_{ih}) =$$

$$= (r_t - 1/2\Gamma^2(T_l))(r_t - 1/2\Gamma^2(T_m))\frac{t^2}{n} + \Gamma(T_l)\Gamma(T_m)\sum_{i=0}^{n-1}(W_{(i+1)h} - W_{ih})^2 +$$

$$+(r_t - 1/2\Gamma(T_l)\Gamma(T_m))(\Gamma(T_l) + \Gamma(T_m))\frac{t}{n}W_t$$

As n tends to infinity, the first and the third addend tend to zero, while the second tends to $\Gamma(T_l)\Gamma(T_m)t$; so the estimator $\hat{\sigma}_n = (\hat{\sigma}_n^{lm})_{lm}$ tends to $\Gamma \cdot \Gamma^t$, where Γ is the matrix having as rows $\Gamma(T_l), l = 1, \dots, d$. We can then get Γ making the square root of the resulting matrix.

Once we have $\Gamma(t, T)$, we get:

$$\tau(t, x) = -\frac{\partial}{\partial T}\Gamma(t, t + x)$$

(as before, we can get only approximate values)

3.3 Karhunen–Loeve expansion

In this part we follow the approach of [12]; the Karhunen–Loeve expansion is a tool that is classically used in fluidodynamics giving nice numerical results, and we believe that it succeeds in treating well also our problem.

We suppose we have a random variable u with values in a function space H, dual of another space K.

Example 3.1 *In our case, we could take* $H = K = L^2(\mathbb{R}^+), H^1(\mathbb{R}^+), L^2([0, T])$ *or* $H^1([0, T])$.

From now on, we suppose that $\mathbb{E}[u] = 0$ (we can always reduce the problem to this case by defining $u' = u - \mathbb{E}[u]$). We suppose then that the bilinear function defined by the variance R of u from H^2 into $\mathbb{R}^+$

$$R(k, k) = \mathbb{E}[\langle u, k\rangle^2] \quad \forall k \in K$$

is continuous; then, for the **kernel representation theorem**, we can find a function $R(t, t')$ such that

$$R(k, k) = \int R(t, t')k(t)k(t')\ dt\ dt'$$

then we can interpret R as an operator going from K to H.

3.3.1 The representation problem

Now we want to maximize the expression

$$\lambda = \frac{\mathbb{E}[\langle u, k\rangle^2]}{\langle k, k\rangle} = \frac{R(k, k)}{\langle k, k\rangle}$$

in $k \in K \cap H$. From a standard procedure by calculus of variations, the extremal points are the k such that

$$R(k,k') = \lambda\langle k,k'\rangle \quad \forall k' \in K \cap H$$

that is, the extremal k are the eigenvectors of R:

$$Rk = \lambda k$$

3.3.2 The "finite energy" case: Karhunen–Loeve expansion

Now we suppose that $(f_n)_n$ is a complete (non necessary orthonormal) system in K (in L^2), decreasing exponentially to infinity, and that

$$\sum_{n=1}^{\infty} \mathbb{E}[\langle u, f_n\rangle^2] < +\infty$$

this is equivalent to $\sum_{n=1}^{\infty}\langle u, f_n\rangle^2 < +\infty$ $\mathbb{P}$-a.s.; then $u \in L^2$ $\mathbb{P}$-a.s.; besides, R is symmetric and nonnegative-definite, so it exists a complete orthonormal set $(k_n)_n$ such that $Rk_n = \lambda_n k_n$, with $\lambda_n \geq 0$.

3.3.3 Calculation of the eigenvalues of R

Now we present an iterative method to find the first eigenvalue λ_1 of R and the corresponding unitary eigenvector k_1.

We fix $k^0 \in K$, and we define:

$$k^{n+1} = Rk^n$$

then we have that:

$$\frac{k^{n+1}}{k^n} \to \lambda_1, \qquad \frac{k^n}{\lambda_1^n} \to \langle k_1, k^0\rangle k_1$$

In order to find the other eigenvalues and eigenvectors, we can apply the method to the operator $R_1 = R - \lambda k_1 \otimes k_1$.

3.3.4 Rate of convergence of the series

We search for a number N that would be in some way an index of how much of the terms of the series are "significative"; we define

$$N = \frac{\sum_{n=1}^{\infty} \lambda_n}{\lambda_1}$$

this means that *at least* N terms of the series are needed in order that the sum is significative. If we suppose that L is large enough to include in $[0,L]^2$ "most" of the support of the function $R(\cdot,\cdot)$, then we have:

$$\lambda_1 \geq \frac{1}{L}\int_0^L \int_0^L R(t,t')\ dt\ dt'$$

and we have the estimation:

$$N \leq L \frac{\int_0^\infty R(t,t)\ dt}{\int_0^L \int_0^L R(t,t')\ dt\ dt'}$$

3.3.5 Application: the Musiela model

Now we want to apply this expansion to the process $(r_t)_t$ in this way: we suppose that $(r_t)_t$ is a stationary solution of the Musiela equation:

$$\begin{cases} dr_t(x) = (Ar_t(x) + c_t(x))\ dt + \tau_t(x)\ dW_t \\ r_0(\cdot) \in L^2(\Omega, \mathcal{F}, \mathbb{P}; H) \end{cases}$$

where τ is a priori an element of H^∞ such that $\sum_{n=1}^\infty \|\tau_n\|^2 < +\infty$. Then the marginal law of r_t is

$$\mathbb{P}(r_t) = N(b^*(\cdot) + b_0, Q_\infty) \quad \forall t \geq 0$$

with $b_0 \in \mathbb{R}$ and

$$Q_\infty = \int_0^\infty \sum_{n=1}^\infty S_u\tau_n \otimes S_u\tau_n\ du, \qquad b^*(\cdot) = -\int_0^x c(u)\ du$$

Since the process is stationary, in order to know its law it is sufficient to know Q_∞; in fact, once we know Q_∞, we are able to get the τ_n and to get c from them.

Here there is a hidden detail: in fact we always studied the Musiela model under the risk-neutral probability $\mathbb{Q}$, while we can obtain from law of the time series an estimation of the hystorical probability $\mathbb{P}$. This doesn't create problems, because since $\mathbb{Q}$ and $\mathbb{P}$ are equivalent, the Gaussian random variable we are studying has the same covariance both under $\mathbb{P}$ as under $\mathbb{Q}$, so we can proceed without troubles.

Now we suppose to have $\bar{N}$ observations $r_{t_i}(x)$, $i = 1, \ldots, \bar{N}$ of the forward curve for $x \in [0, T]$. Since the process is stationary, we can consider a realization of r_t consisting of $\bar{N}$ observations r_{t_i}, $i = 1, \ldots, \bar{N}$ as $\bar{N}$ realizations of the random variable r_t for a given $t \geq 0$.

From now on, we indicate $r_i = r_{t_i}$ for brevity. Then we can approximate:

$$\mathbb{E}[r_t] \simeq \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} r_i, \quad \text{Cov}\ (r_t, r_t) \simeq \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} r_i \otimes r_i$$

In order to implement the algorithm, we follow these procedure:

1st step we center r, passing to $r'_t = r_t - \mathbb{E}[r_t]$:

$$r'_j = r_j - \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} r_i$$

then $r'_t \circ N(0, Q_\infty)$, that is $R = Q_\infty$.

2nd step we apply the algorithm: we fix an initial k^0 and we define:

$$k^{n+1} = \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} \langle r_i, k^n \rangle r_i$$

and we find λ_1 and k_1.

3rd step we define R_1:

$$R_1 = \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} r_i \otimes r_i - \lambda_1 k_1 \otimes k_1$$

and we come back to the 2nd step (we can stop to the suitable N).

Once we obtain the $(k_n)_n$, we can obtain the $(\tau_n)_n$ in this way: we have two expressions of the covariance Q_∞:

$$Q_\infty = \sum_{n=1}^{\infty} \int_0^\infty S_u \tau_n \otimes S_u \tau_n \; du = \sum_{n=1}^{\infty} \lambda_n k_n \otimes k_n$$

each term in the two sums corresponding to a different component of the Brownian motion W. Since the components of W are independent, then we have:

$$\int_0^\infty S_u \tau_n \otimes S_u \tau_n \; du = \lambda_n k_n \otimes k_n \quad \forall n \in \mathbb{N}$$

so we have:

$$\int_0^\infty \tau_n(x+u) \; du = \sqrt{\lambda_n} k_n(x) \quad \forall x \in \mathbb{R}^+, n \in \mathbb{N}$$

Deriving under the integral sign, we obtain

$$\sqrt{\lambda_n} k_n'(x) = \int_0^\infty \tau_n'(x+u) \; du = [\tau_n(x+u)]_0^{+\infty} = -\tau_n(x) \quad \forall x \in \mathbb{R}^+, n \in \mathbb{N}$$

because $\lim_{x \to +\infty} \tau_n(x) = 0$; so we have that

$$\tau_n = -\sqrt{\lambda_n} k_n'$$