STAT391, Lecture 9

Use and calibration of the LFM

A different representation of the forward rates

From (0.18), Lecture 8, we have for \( i = 1, \ldots, n, \)

\[
dF(t, T_{i-1}, T_i) = F(t, T_{i-1}, T_i)\sigma'(t, T_{i-1}, T_i)dW^{T_i}(t), \quad t \leq T_{i-1}, \tag{0.1}
\]

where \( W^{T_i} \) is a \( d \)-dimensional Brownian motion with respect to the forward measure \( P^{T_i} \). We also saw in Lecture 8 that the different Brownian motions were related as

\[
dW^{T_i}(t) = dW^{T_k}(t) - \sum_{j=i+1}^{k} \frac{r_j F(t, T_{j-1}, T_j)}{1 + r_j F(t, T_{j-1}, T_j)} \sigma(t, T_{j-1}, T_j) dt, \quad t \leq T_i, \tag{0.2}
\]

when \( i < k \) and

\[
dW^{T_i}(t) = dW^{T_k}(t) + \sum_{j=k+1}^{i} \frac{r_j F(t, T_{j-1}, T_j)}{1 + r_j F(t, T_{j-1}, T_j)} \sigma(t, T_{j-1}, T_j) dt, \quad t \leq T_k \tag{0.3}
\]

when \( k < i \).

The volatility \( \sigma(t, T_{i-1}, T_i) \) is assumed to be a \( d \)-dimensional deterministic function, and henceforth we assume that it can be written as

\[
\sigma(t, T_{i-1}, T_i) = \sigma_i(t) H_i, \quad i = 1, \ldots, n,
\]

where \( \sigma_i(t) \) is a scalar deterministic function and \( H_i \) is a constant \( d \)-vector. Define the constant \( n \times n \) matrix \( \rho \) by

\[
\rho_{ij} = H'_i H_j, \quad i, j = 1, \ldots, n.
\]

If we define the \( n \times d \) matrix \( H \) by letting \( H'_i = H_i \) be its \( i \)'th row then clearly

\[
\rho = HH'.
\]

Now for any \( n \)-vector \( \alpha \) the quadratic form

\[
\alpha' \rho \alpha = \alpha' HH' \alpha = (H' \alpha)'(H' \alpha) = |H' \alpha|^2 \geq 0,
\]

hence \( \rho \) is nonnegative definite, and it can be seen as a covariance matrix. Now for any \( c \neq 0, \sigma_i(t) H_i = (c \sigma_i(t))(c^{-1} H_i) \), so that there is a indeterminacy of scale in the choice of \( \sigma_i(t) \) and \( H_i \). This can be solved by letting \( |H_i|^2 = 1 \), i.e. by letting the diagonal elements of \( \rho \) all be equal to one. Then \( \rho \) will in fact be a correlation matrix, so that there exists a random vector \( X = (V_1, \ldots, V_n) \) (in fact lots of them), so that \( \text{Corr}(V_i, V_j) = \rho_{ij} \).
Now consider the $n$-dimensional process

$$Z^{T_i}(t) = HW^{T_i}(t),$$

so that the $j$'th component equals $Z_j^{T_i}(t) = H_j^i W^{T_i}(t)$. Obviously $Z_j^{T_i}$ is continuous, and

$$E^{T_i}[Z_j^{T_i}(t + s)|\mathcal{F}_i] = H_j^i E^{T_i}[W^{T_i}(t + s)|\mathcal{F}_i] = H_j^i W^{T_i}(t) = Z_j^{T_i}(t),$$

so $Z_j^{T_i}$ is a martingale. Furthermore

$$E^{T_i}[(Z_j^{T_i}(t + s) - Z_j^{T_i}(t))^2|\mathcal{F}_i] = E^{T_i}[(H_j^i(W^{T_i}(t + s) - W^{T_i}(t)))^2|\mathcal{F}_i]$$

$$= H_j^i E^{T_i}[(W^{T_i}(t + s) - W^{T_i}(t))(W^{T_i}(t + s) - W^{T_i}(t))'|\mathcal{F}_i] H_j^i$$

$$= H_j^i t I_d H_i = t H_j^i H_i = t,$$

where $I_d$ is the $d \times d$ identity matrix. It now follows from Theorem 0.1, Lecture 1, that each $Z_j^{T_i}$ is a Brownian motion under $P^{T_i}$. However, $Z^{T_i}$ is not an $n$-dimensional Brownian motion under $P^{T_i}$ since

$$E^{T_i}[dZ^{T_i}(t)(dZ^{T_i}(t))'|\mathcal{F}_i] = HH' dt = \rho dt,$$

so the components are correlated.

By (0.1) we have,

$$dF(t, T_{i-1}, T_i) = F(t, T_{i-1}, T_i) \sigma_i(t) H_i' dW^{T_i}(t) = F(t, T_{i-1}, T_i) \sigma_i(t) dZ_i^{T_i}(t).$$

By (0.1) and (0.3), for $i < k$,

$$dF(t, T_{i-1}, T_i) = F(t, T_{i-1}, T_i) \sigma_i(t) H_i' \left( dW^{T_k}(t) - \sum_{j=i+1}^{k} \frac{\tau_j F(t, T_{j-1}, T_j) \sigma_j(t) H_j dt}{1 + \tau_j F(t, T_{j-1}, T_j)} \right)$$

$$= -F(t, T_{i-1}, T_i) \sigma_i(t) \sum_{j=i+1}^{k} \frac{\rho_{ij} \tau_j F(t, T_{j-1}, T_j) \sigma_j(t)}{1 + \tau_j F(t, T_{j-1}, T_j)} dt$$

$$+ F(t, T_{i-1}, T_i) \sigma_i(t) dZ_i^{T_k}(t).$$

Similarly, for $i > k$, using (0.1) and (0.3),

$$dF(t, T_{i-1}, T_i) = F(t, T_{i-1}, T_i) \sigma_i(t) \sum_{j=k+1}^{i} \frac{\rho_{ij} \tau_j F(t, T_{j-1}, T_j) \sigma_j(t)}{1 + \tau_j F(t, T_{j-1}, T_j)} dt$$

$$+ F(t, T_{i-1}, T_i) \sigma_i(t) dZ_i^{T_k}(t).$$

(0.4)
To compute the value of caplets and hence of caps, according to the formula (0.19) of Lecture 8, the only quantity to be computed is

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

$$= \int_t^{T_{i-1}} |H_i \sigma_i(s)|^2 ds$$

$$= \int_t^{T_{i-1}} |H_i|^2 |\sigma_i(s)|^2 ds$$

$$= \int_t^{T_{i-1}} |\sigma_i(s)|^2 ds,$$

so once $\sigma_i^2$ is known, the caplet and cap pricing is solved. Note that even though the cap includes several time instants, the correlation matrix $\rho$ does not enter into its price.

**Monte Carlo pricing of swaptions with the LFM**

Another very important fixed income derivative is the swaption. We saw in Exercise 3, HW8, that pricing a swaption is easy if the swap rate $S_{T_n}^w(t)$ follows a lognormal model under the forward swap measure $P_{T_n}^f$. However, if we work with the LFM, the pricing of swaptions is much more complicated, and is often done using Monte Carlo simulations. To explain how this works, remember that the swaption payout at time $T_0$ equals

$$X = \left(\sum_{i=1}^{\infty} \tau_i P(T_0, T_i)(F(T_0, T_{i-1}, T_i) - R)\right)^+.$$

Using the forward measure $P^{T_0}$, we get

$$\Pi_X(t, T_0) = P(t, T_0)E^{T_0}\left[\left(\sum_{i=1}^{\infty} \tau_i P(T_0, T_i)(F(T_0, T_{i-1}, T_i) - R)\right)^+ \bigg| \mathcal{F}_t\right].$$

Now

$$1 + \tau_i F(T_0, T_{i-1}, T_i) = G(T_0, T_{i-1}, T_i) = \frac{P(T_0, T_{i-1})}{P(T_0, T_i)},$$

hence

$$P(T_0, T_i) = \frac{P(T_0, T_{i-1})}{1 + \tau_i F(T_0, T_{i-1}, T_i)}.$$

Recursion of this expression and using that $P(T_0, T_0) = 1$ gives,

$$P(T_0, T_i) = \prod_{j=1}^{i} \frac{1}{1 + \tau_j F(T_0, T_{j-1}, T_j)}.$$

This gives, setting $t = t_0$ to avoid notational confusion,

$$\Pi_X(t_0, T_0) = P(t_0, T_0)E^{T_0}\left[\left(\sum_{i=1}^{n} \tau_i \left(\prod_{j=1}^{i} \frac{1}{1 + \tau_j F(T_0, T_{j-1}, T_j)}\right) (F(T_0, T_{i-1}, T_i) - R)\right)^+ \bigg| \mathcal{F}_{t_0}\right].$$

3
Assume we can use the Monte Carlo method to simulate \( M \) independent realizations of

\[
\left( \sum_{i=1}^{n} \tau_i \left( \prod_{j=1}^{i} \frac{1}{1 + \tau_j F(T_0, T_{j-1}, T_j)} \right) (F(T_0, T_{i-1}, T_i) - R) \right)^+
\]

under \( P^{T_0} \), and given \( \mathcal{F}_0 \). Call these \( \hat{X}_1, \ldots, \hat{X}_M \). Our estimate of the swaption price is then

\[
\hat{\Pi}_X(t_0, T_0) = P(t_0, T_0) \frac{1}{M} \sum_{m=1}^{M} \hat{X}_m.
\]

If the simulations are done properly, \( E^{T_0}[\hat{X}_m|\mathcal{F}_0] = E^{T_0}[X_m|\mathcal{F}_0] \), giving

\[
E^{T_0}[\hat{\Pi}_X(t_0, T_0)|\mathcal{F}_0] = \Pi_X(t_0, T_0).
\]

Furthermore

\[
\text{Var}^{T_0}[\hat{\Pi}_X(t_0, T_0)|\mathcal{F}_0] = \frac{P^2(t_0, T_0)}{M} \text{Var}^{T_0}[\hat{X}|\mathcal{F}_0].
\]

We shall now address the question of how to get the \( \hat{X}_m \). By (0.4) we have setting \( k = 0 \),

\[
dF(t, T_{i-1}, T_i) = F(t, T_{i-1}, T_i) \mu_i(t)dt + F(t, T_{i-1}, T_i)\sigma_i(t)dZ_i^{T_0}(t),
\]

where

\[
\mu_i(t) = \sum_{j=1}^{i} \rho_{ij} \tau_j F(t, T_{j-1}, T_j) \sigma_j(t) \frac{1}{1 + \tau_j F(t, T_{j-1}, T_j)}.
\]

(0.5)

There is no closed form solution to this system of SDE, hence a discretization is necessary. To make it somewhat better suited for discretization, let \( F_L(t, T_{i-1}, T_i) = \log F(t, T_{i-1}, T_i) \). Itô's formula then gives

\[
dF_L(t, T_{i-1}, T_i) = (\mu_i(t) - \frac{1}{2}\sigma^2_i(t))dt + \sigma_i(t)dZ_i^{T_0}(t).
\]

Here the volatility term is nonstochastic, and it is well known that this property makes a simulation procedure more accurate. Therefore we simulate \( F_L(t, T_{i-1}, T_i) \) and then calculate \( F(t, T_{i-1}, T_i) = \exp\{F_L(t, T_{i-1}, T_i)\} \).

The simulation will be done simultaneously for \( i = 1, \ldots, n \). Divide the interval \([t_0, T_0]\) into \( K + 1 \) equidistant points so that \( h = (T_0 - t_0)/K \) is the distance between two points. Set \( t_k = t_0 + kh \), so in particular \( t_K = T_0 \). Initially \( F(t_0, T_{i-1}, T_i) \) is known for all \( i \), and recursively

\[
F_L(t_{k+1}, T_{i-1}, T_i) - F_L(t_k, T_{i-1}, T_i) = (\mu_i(t_k) - \frac{1}{2}\sigma^2_i(t_k))h
+\sigma_i(t_k)(Z_i^{T_0}(t_{k+1}) - Z_i^{T_0}(t_k)) + \text{Remainder}.
\]
For $K$ sufficiently large, or equivalently $h$ sufficiently small, the remainder can be ignored to give approximately

$$F_L(t_{k+1}, T_{i-1}, T_i) = F_L(t_k, T_{i-1}, T_i) + (\mu_i(t_k) - \frac{1}{2} \sigma_i^2(t_k))h + \sigma_i(t_k)(Z_{T_0}^{T_0}(t_{k+1}) - Z_{T_0}^{T_0}(t_k)).$$

From (0.5) it is seen that in order to compute $\mu_i(t_k)$, all the $F(t_k, T_{j-1}, T_j)$, $j \leq i$ are needed. This is why the simulation is done simultaneously for all $i = 1, \ldots, n$.

To do this, we must simulate the $n$-dimensional

$$Z_{T_0}^{T_0}(t_{k+1}) - Z_{T_0}^{T_0}(t_k) \sim \mathcal{N}(0, \rho h).$$

Here $\rho$ is an $n \times n$ matrix with rank equal to $d$. Therefore, it can be decomposed as $\rho = HH'$, where $H$ is an $n \times d$ matrix of rank $d$. It is assumed that $d \leq n$, which is the typical case. In fact $n$ may be larger than 10, while $d$ will often be smaller than 5. This decomposition can be obtained in several ways, it is not unique, but that is of no concern to us. Also it is the same for all time steps $t_k$, so it is in fact done before the recursion of the $F_L(t_k, T_{i-1}, T_i)$ starts.

With this decomposition at hand, to obtain the new $Z_{T_0}^{T_0}(t_{k+1}) - Z_{T_0}^{T_0}(t_k)$, just generate the independent standard normally distributed $U_1, \ldots, U_d$ and let $U = (U_1, \ldots, U_d)'$. Then set

$$\hat{Z} = HU \sqrt{h}.$$

This gives $E[\hat{Z}] = 0$ and $E[\hat{Z} \hat{Z}'] = H I_d H' h = \rho h$. Therefore the update becomes,

$$F_L(t_{k+1}, T_{i-1}, T_i) = F_L(t_k, T_{i-1}, T_i) + (\mu_i(t_k) - \frac{1}{2} \sigma_i^2(t_k))h + \sigma_i(t_k)\hat{Z}_i, \quad i = 1, \ldots, n.$$  

Continuing this process finally gives $F(t_K, T_{i-1}, T_i) = F(T_0, T_{i-1}, T_i)$ for all $i = 1, \ldots, n$.

Calibration

It is clear from the above that in addition to the initial forward rates $F(t_0, T_{i-1}, T_i)$, $i = 1, \ldots, n$, the price of the swaption depends on both the volatilities $\sigma_i(t)$ as well as the $n \times n$ correlation matrix $\rho$. We will now discuss how to choose these to fit the observed caplet and swaption volatilities. Remember that in the pricing of caps and caplets the correlation matrix $\rho$ does not enter, so it can be used entirely to fit the swaption prices. The volatilities $\sigma_i(t)$ can partly be used to calibrate the caplets, and partly to calibrate the swaptions.

Models for $\sigma_i(t)$.

We assume first that

$$\sigma_i(t) = \sigma_{i,a}(t), \quad t \leq T_i,$$
where as before, \( a(t) = \min \{ k : T_k \geq t \} \). Here the \( \sigma_{i,j} \) are constants, and it is further assumed that they are of the form

\[
\sigma_{i,j} = \Phi_i \psi_{i-j}, \quad j < i, \quad (0.6)
\]

where then the \( \Phi_i \) and the \( \psi_j \) are constants. In table form it looks like

\[
\begin{array}{|c|c|c|c|c|}
\hline
 t & (0, T_0) & (T_0, T_1) & (T_1, T_2) & \cdots & (T_{n-2}, T_{n-1}) \\
\hline
 \sigma_1(t) & \Phi_1 \psi_1 & \times & \times & \cdots & \times \\
 \sigma_2(t) & \Phi_2 \psi_2 & \Phi_2 \psi_1 & \times & \cdots & \times \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \times \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \times \\
 \sigma_n(t) & \Phi_n \psi_n & \Phi_n \psi_{n-1} & \Phi_n \psi_{n-2} & \cdots & \Phi_n \psi_1 \\
\hline
\end{array}
\]

We see that the \( \Phi_i \) are the same when the maturities are the same, while the \( \psi_j \) are the same when the time to maturities are the same.

A slightly different model is the following,

\[
\sigma_i(t) = \Phi_i((a(T_{i-1} - t) + d)e^{-b(T_{i-1} - t)} + c), \quad (0.7)
\]

for \( i = 1, \ldots, n \). Note that the same \( a, b, c \) and \( d \) are used for all \( i \), so if \( n > 4 \), which is typically the case, (0.6) has more parameters to calibrate than (0.7).

The definition of caplet volatility as given in Lecture 7 is

\[
\hat{\sigma}_{\text{cpl}}^2(T_{i-1}) = \frac{1}{T_i - T_{i-1}} \int_0^{T_i - T_{i-1}} (d \log F(t, T_{i-1}, T_i))^2 dt = \frac{1}{T_{i-1}} \int_0^{T_{i-1}} \sigma_i^2(t) dt.
\]

For the model (0.6) this gives with \( T_{i-1} = 0 \),

\[
\hat{\sigma}_{\text{cpl}}^2(T_{i-1}) = \frac{1}{T_{i-1}} \sum_{j=0}^{i-1} \int_{T_{j+1}}^{T_{j+2}} (\Phi_i \psi_{i-j})^2 dt = \frac{\Phi_i^2}{T_{i-1}} \sum_{j=0}^{i-1} \tau_j \psi_{i-j}^2, \quad (0.8)
\]

where as before \( \tau_i = \tau(T_{i-1}, T_i) \). We are here working with the LFM, so \( \hat{\sigma}_{\text{cpl}}^2(T_i) = \hat{\sigma}_{\text{cpl}}^2(T_i) \), and these are given by the market, hence they can also be written as \( \hat{\sigma}_{\text{Market}}^2(T_i) \). Therefore, (0.8) becomes

\[
\Phi_i^2 = \frac{T_{i-1} \hat{\sigma}_{\text{cpl}}^2(T_{i-1})}{\sum_{j=0}^{i-1} \tau_j \psi_{i-j}^2}. \quad (0.9)
\]

This gives the \( \Phi_i \) in terms of market quotations and the \( \psi_j \). The \( \psi_j \) as well as the correlation matrix \( \rho \) can now be used to calibrate the model to e.g. the swaption prices.
If we instead assume the volatility (0.7), then
\[
\hat{\sigma}^2_{cpl}(T_{i-1}) = \frac{\Phi^2_t}{T_{i-1}} \int_0^{T_{i-1}} \left((a(T_i - t) + d)e^{-b(t-t_i)} + c \right)^2 dt
\]
\[
= \frac{\Phi^2_t}{T_{i-1}} \int_0^{T_{i-1}} (at + d)e^{-bt} + c)^2 dt
\]
\[
= \frac{\Phi^2_t}{T_{i-1}} I(T_{i-1}; a, b, c, d),
\]
where \(I(T; a, b, c, d)\) is easily calculated, the boring details are left to the reader. Then as above we get
\[
\Phi^2_t = \frac{T_{i-1}\hat{\sigma}^2_{cpl}(T_{i-1})}{I(T_{i-1}; a, b, c, d)}. \tag{0.10}
\]
Again, the parameters \(a, b, c\) and \(d\) as well as the correlation matrix \(\rho\) can be used to calibrate the swaption prices.

There are of course lots of other ways to represent \(\sigma_i(t)\) than (0.6) and (0.7) above. The reason for choosing these is that they preserve the form of the volatility curve as time passes. To explain why, consider the graph of the volatility curve as seen at time \(t = 0\),
\[
\{(T_0, \hat{\sigma}(T_0)), (T_1, \hat{\sigma}(T_1)), \ldots, (T_{n-1}, \hat{\sigma}(T_{n-1}))\},
\]
where \(\hat{\sigma}(T_i) = \hat{\sigma}_{cpl}(T_i)\). The humped form means that for the first \(T_i\), the corresponding \(\hat{\sigma}(T_i)\) will increase, and after that it will decrease.

Now keep the maturity \(n\) fixed, but assume we move to \(t = T_0\). Then the graph of the remaining volatilities becomes
\[
\{(T_1, V(T_0, T_1)), (T_2, V(T_0, T_2)), \ldots, (T_{n-1}, V(T_0, T_{n-1}))\},
\]
where for \(i > 0\),
\[
V^2(T_0, T_{i-1}) = \frac{1}{\tau(T_0, T_{i-1})} \int_0^{T_{i-1}} \sigma_i^2(t) dt.
\]
If the original curve starting at \(t = 0\) had a hump with a maximum at \(T_3\) say, then the curve starting at \(t = T_0\) should be a time translation of the first one, having the same hump, but now with a maximum at \(T_3\) (assuming equidistant points).

Continuing to \(t = T_j\), the remaining graph is now
\[
\{(T_{j+1}, V(T_j, T_{j+1})), (T_{j+2}, V(T_j, T_{j+2})), \ldots, (T_{n-1}, V(T_j, T_{n-1}))\},
\]
where for \(k > 1\),
\[
V^2(T_j, T_{j+k-1}) = \frac{1}{\tau(T_j, T_{j+k-1})} \int_{T_j}^{T_{j+k-1}} \sigma_{j+k}^2(t) dt.
\]
With $\sigma_i(t)$ given by (0.6), we get for $t = T_j$ just as in (0.8),

$$V^2(T_j, T_{j+k-1}) = \frac{1}{\tau(T_j, T_{j+k-1})} \int_{T_j}^{T_{j+k-1}} \sigma_{j+k}^2(t) dt$$

$$= \frac{1}{\tau(T_j, T_{j+k-1})} \sum_{i=j+1}^{j+k-1} \tau_{i} \psi_{j+k-i}^2$$

By going from $t = T_j$ to $t = T_{j+1}$, the corresponding point on the graph will be $V(T_{j+1}, T_{j+k})$, and therefore we would like to have $V(T_{j+1}, T_{j+k}) \approx V(T_j, T_{j+k-1})$. But the same arguments as above gives

$$V^2(T_{j+1}, T_{j+k}) = \frac{\Phi_{j+k+1}^2}{\tau(T_{j+1}, T_{j+k})} \sum_{i=j+2}^{j+k} \tau_{i} \psi_{j+k-i+1}^2$$

Assume now that the $\{T_i\}$ are equidistant points so that $\tau_0 = \tau_1 = \cdots = \tau_n = \tau$, giving $\tau(T_j, T_{j+k}) = k\tau$. Then

$$V^2(T_j, T_{j+k-1}) = \frac{\Phi_{j+k}^2}{k-1} \sum_{i=j+1}^{j+k-1} \psi_{j+k-i}^2 = \frac{\Phi_{j+k}^2}{k-1} \sum_{i=1}^{k-1} \psi_{k-i}^2$$

$$V^2(T_{j+1}, T_{j+k}) = \frac{\Phi_{j+k+1}^2}{k-1} \sum_{i=j+2}^{j+k} \psi_{j+k-i+1}^2 = \frac{\Phi_{j+k+1}^2}{k-1} \sum_{i=1}^{k-1} \psi_{k-i}^2$$

We see that the only difference is that $\Phi_{j+k}^2$ becomes $\Phi_{j+k+1}^2$, so if the $\Phi_i$ are reasonable stable over time, the shape of the volatility curve will remain mostly the same. It will move from left to right, and for each increase in $j$, it is capped at the right tail.

Since the form of $\sigma_i(t)$ in (0.7) has much the same time structure as that in (0.6), this form will also preserve the volatility curve as time passes.

Consider the volatility matrix $\rho$. We saw above that it can be written as $\rho = HH'$ where $H$ is an $n \times d$ matrix. We shall assume that $H$ has full rank $d \leq n$. If $d = 1$, then since each element in $H$ has length 1, $H$ can be set equal to the $n \times 1$ matrix, i.e. the $n$-vector, consisting of ones.

For $d = 2$, $H$ must be of the form $H_i = (\cos \theta_i, \sin \theta_i)'$, since this gives $|H_i|^2 = \cos^2 \theta_i + \sin^2 \theta_i = 1$. Furthermore

$$\rho_{ij} = H_i' H_j = \cos \theta_i \cos \theta_j + \sin \theta_i \sin \theta_j = \cos(\theta_i - \theta_j).$$
Adjacent maturities are typically highly positively correlated, so \( \theta_i \) and \( \theta_{i+1} \) should lie fairly near each other.

For general \( d \) we can set for \( i = 1, \ldots, n, \)
\[
h_{i1} = \cos \theta_{i1},
\]
\[
h_{ij} = \cos \theta_{ij} \prod_{k=1}^{j-1} \sin \theta_{ik}, \quad j = 2, \ldots, d - 1,
\]
\[
h_{id} = \prod_{k=1}^{d-1} \sin \theta_{ik}.
\]

Note that
\[
|H_i|^2 = \sum_{j=1}^{d} h_{ij}^2
\]
\[
= \cos^2 \theta_{i1} + \sum_{j=1}^{d-1} \cos^2 \theta_{ij} \prod_{k=1}^{j-1} \sin^2 \theta_{ik} + \prod_{k=1}^{d-1} \sin^2 \theta_{ik}
\]
\[
= \cos^2 \theta_{i1} + \sum_{j=1}^{d-2} \cos^2 \theta_{ij} \prod_{k=1}^{j-1} \sin^2 \theta_{ik} + \cos^2 \theta_{i,d-1} \prod_{k=1}^{d-2} \sin^2 \theta_{ik} + \prod_{k=1}^{d-2} \sin^2 \theta_{ik}
\]
\[
= \cos^2 \theta_{i1} + \sum_{j=1}^{d-2} \cos^2 \theta_{ij} \prod_{k=1}^{j-1} \sin^2 \theta_{ik} + \prod_{k=1}^{d-2} \sin^2 \theta_{ik}
\]
\[
= \ldots
\]
\[
= \cos^2 \theta_{i1} + \sin^2 \theta_{i1} = 1,
\]
therefore this suggested form can be used as a correlation matrix.

For given \( d \), the above representation of \( \rho \) gives \( n \cdot (d - 1) \) different \( \theta_{ij} \) to calibrate, and if we use the model (0.6) for \( \theta_{ij} \), we have additionally \( n \) \( \psi_i \)'s, assuming the \( \Phi_i \)'s are calibrated according to the caplet volatilities as in (0.9). If we instead use (0.7), in addition to the \( n \cdot (d - 1) \) different \( \theta_{ij} \), we have \( a, b, c \) and \( d \) free to do the calibration, assuming the \( \Phi_i \)'s are calibrated according to (0.10).

All this remaining calibration will be done in accordance with a table of swaption volatilities. To see how, consider a swaption with time of delivery \( T_0 \) and lifetime \( T_n - T_0 \), which at time \( T_0 \) pays (cf. Exercise 3, HW 8),
\[
Y = (S_{\tau_n}(T_0) - R)^+ C_{\tau_n}(T_0),
\]
where
\[
S_{\tau_n}(t) = \frac{P(t, T_0) - P(t, T_n)}{C_{\tau_n}(t)},
\]
\[
C_{\tau_n}(t) = \sum_{i=1}^{n} \tau_i P(t, T_i).
\]
Assume the swap rate follows the LSM (Lognormal Forward-Rate Model),

\[ dS_{T_n}(t) = S_{T_n}(t)\sigma_{T_n}(t)dW_{T_n}(t), \]

with \( \sigma_{T_n}(t) \) deterministic. Then according to Exercise 3, HW 8, the Black formula for pricing the swap option at time \( t = 0 \) becomes

\[ \Pi_T(0, T_0) = C_{T_n}(0)S_{T_n}(0)N(d) - RC_{T_n}(0)N(d - b(0, T_0)), \]

where

\[ d = \log \left( \frac{S_{T_n}(0)}{K} \right) + \frac{1}{2}b^2(0, T_0) \]

\[ b^2(0, T_0) = \int_0^{T_0} |\sigma_{T_n}(s)|^2 ds. \]

Now let

\[ v_{T_n}^2(T_0) = \frac{1}{T_0} b^2(0, T_0) = \frac{1}{T_0} \int_0^{T_0} (d\log S_{T_n}(t))^2. \]

This is the formula used by the market to price swaptions, at least for at the money swaptions. Therefore, for given \( T_0 \) and \( T_n \), the market volatility \( v_{T_n}(T_0) \) is the volatility that gives the quoted price for an at the money swaption.

A swaption table is now an \( A \times B \) table, where the element in row \( a \) and column \( b \) is \( v_{T_{n_a}}(T_{n_b}) \), with \( T_{n_a} \) the delivery date and \( T_{n_b} \) the subsequent dates. Usually \( T_i - T_{i-1} \) is one year. In this case, if \( T_{n_a} = 3 \) and \( T_{n_b} = (4, 5, \ldots, 12) \), this means that the delivery of the swaption is in 3 years, and it will cover the subsequent 9 years. Furthermore, all the volatilities in this row of the table will have delivery date in 3 years, but with different lengths, while all the volatilities in this column will have lengths 9 years, but with different delivery dates.

In order to do this calibration fast enough, it is necessary to be able to calculate the swaption volatility implied by the LFM sufficiently efficient. An exact calculation is not feasible, but we shall here present Rebonato’s formula which seems to work well. We start be proving a useful representation of the swaption rate \( S_{T_n}(t) \). Note that

\[ P(t, T_0) - P(t, T_n) = \sum_{i=1}^{n} (P(t, T_{i-1}) - P(t, T_i)) \]

\[ = \sum_{i=1}^{n} P(t, T_i) \left( \frac{P(t, T_{i-1})}{P(t, T_i)} - 1 \right) \]

\[ = \sum_{i=1}^{n} P(t, T_i) \tau_i F(t, T_{i-1}, T_i). \]
Therefore
\[ S_{T_n}(t) = \frac{P(t, T_0) - P(t, T_n)}{\sum_{k=1}^{n} \tau_k P(t, T_k)} = \sum_{i=1}^{n} w_i(t) F(t, T_{i-1}, T_i), \]  
(0.11)
where
\[ w_i(t) = \frac{\tau_i P(t, T_i)}{\sum_{k=1}^{n} \tau_k P(t, T_k)}. \]

Note that \( \sum_{i=1}^{n} w_i(t) = 1 \) for all \( t \).

Since the \( w_i(t) \) vary far less than the \( F(t, T_{i-1}, T_i) \), the first approximation is to set them all equal to \( w_i(0) \), and then (0.11) becomes approximately
\[ S_{T_n}(t) \approx \sum_{i=1}^{n} w_i(0) F(t, T_{i-1}, T_i). \]

Taking the differential gives,
\[ dS_{T_n}(t) \approx c_t dt + \sum_{i=1}^{n} w_i(0) F(t, T_{i-1}, T_i) \sigma_i(t) dZ^T_0(t), \]

where, since we are only interested in volatilities, \( c_t \) is an uninteresting adapted process. Now Itô’s formula gives that
\[ (d \log S_{T_n}(t))^2 = \left( \frac{dS_{T_n}(t)}{S_{T_n}(t)} \right)^2, \]
since the difference in \( d \log S_{T_n}(t) \) and \( \frac{dS_{T_n}(t)}{S_{T_n}(t)} \) is only a \( dt \) term, and this disappears when taking the squares. Therefore, using that \( dZ^T_0(t) dZ^T_0(t) = \rho_{ij} dt \), we get
\[ (d \log S_{T_n}(t))^2 \approx \sum_{i,j=1}^{n} w_i(0) w_j(0) F(t, T_{i-1}, T_i) F(t, T_{j-1}, T_j) \sigma_i(t) \sigma_j(t) \rho_{ij} \frac{S^2_{T_n}(t)}{S^2_{T_n}(0)}. \]

As a second approximation, set \( F(t, T_{i-1}, T_i) = F(0, T_{i-1}, T_i), F(t, T_{j-1}, T_j) = F(0, T_{j-1}, T_j) \) and \( S_{T_n}(t) = S_{T_n}(0) \). Then
\[ (d \log S_{T_n}(t))^2 \approx \sum_{i,j=1}^{n} a_{ij} \rho_{ij} \sigma_i(t) \sigma_j(t) dt, \]
where
\[ a_{ij} = \frac{w_i(0) w_j(0) F(0, T_{i-1}, T_i) F(0, T_{j-1}, T_j)}{S^2_{T_n}(0)}, \quad i, j = 1, \ldots, n. \]

Therefore
\[ \frac{1}{T_0} \int_{T_0}^{T_0} (d \log S_{T_n}(t))^2 \approx \sum_{i,j=1}^{n} a_{ij} \rho_{ij} \frac{1}{T_0} \int_{T_0}^{T_0} \sigma_i(t) \sigma_j(t) dt. \]
Rebonato’s formula for the LFM swaption model volatility now becomes

\[(v_{T_n}^{\text{LFM}}(T_0))^2 \overset{\text{def}}{=} \sum_{i,j=1}^n a_{ij}\rho_{ij} \frac{1}{T_0} \int_0^{T_0} \sigma_i(t)\sigma_j(t)dt.\]

For a good calibration, this should be as near to the market priced \(v_{T_n}(T_0)\) as possible, simultaneously for all the market volatilities in the swaption table described above.

Letting \(\sigma_i(t)\) be given by (0.6), i.e. \(\sigma_i(t) = \Phi_i\psi_t\) when \(t \leq T_0\), the formula becomes

\[(v_{T_n}^{\text{LFM}}(T_0))^2 = \sum_{i,j=1}^n a_{ij}\rho_{ij}\Phi_i\Phi_j\psi_i\psi_j.\]

Then taking advantage of the already calibrated caplet volatilities (0.9), we get

\[(v_{T_n}^{\text{LFM}}(T_0))^2 = \sum_{i,j=1}^n a_{ij}\rho_{ij}\psi_i\psi_j \sqrt{\frac{T_{i-1}\tilde{\nu}_{\text{cpl}}^2(T_{i-1}) T_{j-1}\tilde{\nu}_{\text{cpl}}^2(T_{j-1})}{\sum_{k=0}^{i-1} t_k\psi_i^2 - \sum_{l=0}^{j-1} t_l\psi_j^2}}.\]

The \(\theta_{ij}\) that gives \(\rho\) as well as the \(\psi_t\)’s can now be found by minimizing e.g. the sum of squares

\[\sum_{a=1}^A \sum_{b=1}^B \left( v_{T_n_{a,b}}(T_0) - v_{T_n_{a,b}}^{\text{LFM}}(T_0) \right)^2.\]