Volatility Cube Construction

Arthur PHAM[*], El Hadj FALL, Achim SCHNIEWIND

March 14, 2017

Abstract

Adjusting for the volatility skew is seen as essential today to accurately price interest rate derivatives such as swaptions. However, market data for swaption skews is neither transparently nor freely available across a wide range of currencies. We will derive swaption volatility cubes (by expiry, underlying tenor, and strike) from the more generally available cap-floor volatility surfaces.

To interpolate accurately on the cube, the SABR model has been implemented. The SABR model is a well established model in the market, which is highly tractable, enabling frequent updates of the derived swaption volatility cube.

In this document we explain in detail the calculation methodology used to derive swaption volatility smiles from quoted ATM swaption volatilities and cap volatility surfaces.

[*]Address: Thomson Reuters, 3 Times Square, 10036 New York NY, USA. Email: arthur.pham@thomsonreuters.com

1 Introduction

From the cap smile, we strip for each expiry the caplet smile for quoted strikes. This smile is applied (in ways which will be described below) to the ATM swaption volatilities with same option expiry. The resulting curve is supposed to be the swaption smile. The hypothesis in that approach is that the caplet smile and all swaptions smiles with the same expiry (regardless of the underlying swap tenor) have a similar shape. The main limitation of this hypothesis is that it makes strong implicit assumptions about the structure of correlations between forward rates. However, we do not have sufficient information to imply that correlation structure from market rates. Hence we calibrate the SABR model for swaptions with the help of caplets incorporating the ATM swaption vol information.

To obtain swaption volatilities for non standard expiries and tenors, we perform a further interpolation in the expiry/tenor dimension.

Available data

As opposed to caps/floors, swaption vols are only readily observable for the ATM case. Summarizing the available data are:

- libor rates, swap rates
- ATM/ITM/OTM caps/floors vols for specific maturities,
- ATM swaption vols for specific expiry/tenor. Recall that a caplet can be seen as a one period swaption.

Algorithm

The vol cube construction procedure consists of the following main steps to be described in the subsequent chapters:

1. Strip the caplet vols from the quoted cap vols.
2. As we don't have OTM/ITM swaption quotation, we estimate the swaption smile by lifting the caplet smile to the swaptions.
3. For each point in the expiry/tenor space, we will calibrate the SABR model in strike dimension and calculate implied volatilities for each requested strike. To cover non standard expiries and tenors, we perform a linear interpolation of the obtained parameters in the expiry/tenor dimension.

2 Stripping Caplet Volatilities From Cap Volatilities

Cap volatilities are typically quoted by expiry and strike as in the following example :

	ATM strike	ATM vol	0.0025	0.005	0.0075	0.01	0.015	0.02	0.025	0.03	0.04
1Y	0.004	0.4707	0.429	0.497	0.539	0.561	0.591	0.608	0.62	0.628	0.639
2Y	0.0046	0.5435	0.511	0.551	0.586	0.607	0.631	0.645	0.654	0.66	0.668
3Y	0.0055	0.5873	0.563	0.582	0.608	0.628	0.649	0.662	0.67	0.676	0.683
4Y	0.0065	0.587	0.574	0.578	0.594	0.611	0.629	0.639	0.646	0.651	0.658
5Y	0.0076	0.5593	0.564	0.553	0.559	0.573	0.589	0.598	0.605	0.611	0.618
6Y	0.0089	0.519	0.544	0.52	0.514	0.524	0.536	0.546	0.553	0.559	0.567
7Y	0.0103	0.4775	0.525	0.491	0.474	0.477	0.483	0.492	0.5	0.506	0.515
8Y	0.0116	0.4392	0.506	0.467	0.443	0.439	0.44	0.447	0.454	0.46	0.47
9Y	0.0129	0.4052	0.49	0.447	0.419	0.409	0.404	0.41	0.417	0.424	0.435
10Y	0.0141	0.3766	0.476	0.431	0.399	0.386	0.376	0.381	0.388	0.395	0.407

We will now describe several caplet stripping methods that can be used.

2.1 Piece-wise Constant Caplet Volatilities

This method is guaranteed to work if cap data are consistent, i.e. if all forward cap premiums are above their intrinsic values. It only requires one dimensional solving by implying forward cap volatilities. If this method fails, the data should be questioned.

Let $\{T_i, 1 \leq i \leq N\}$ be the cap maturities, with associated volatilities $\{\Sigma(i,K), 1 \leq i \leq N\}$ for a given strike, and $\{\sigma(i,K), 1 \leq i \leq N\}$ the caplet volatilities we are trying to compute. Let's denote $Cap(T_0, T_i, K, \Sigma_i)$ the cap premium for strike K, maturity T_i, and volatility Σ_i.

By definition and quotation rules, we have :

$$Cap(T_0, T_i, K, \Sigma_i) = \sum_{j=1}^{N_i} Caplet(t_{j-1}, t_j, K, \Sigma_i)$$

where $\{t_j, 1 \leq j \leq N_i\}$ is the caplet term, and we define $t_0 = T_0$.

Caps are usually $1Y$ spaced while caplets are typically $3M$ or $6M$ based (depending on currency).

Caplet prices are computed from the Black&Scholes formula (lognormal) :

$$Caplet(t_{j-1}, t_j, K, \sigma) = B(t_0, t_j) YF(t_{j-1}, t_j) \left[FN(d_1) - KN(d_2) \right]$$

where :

$$\begin{aligned}
F = F(t_0, t_{j-1}, t_j) &: \text{the forward Libor rate} \\
B(t_0, t_j) &: \text{the Zero-Coupon value (discount factor)} \\
YF(t_{j-1}, t_j) &: \text{the year fraction between } t_{j-1} \text{ and } t_j \\
d_1 &= \frac{\ln(\frac{F}{K})}{\sigma\sqrt{YF(t_0, t_j)}} + \frac{\sigma^B \sqrt{YF(t_0, t_j)}}{2} \\
d_2 &= d_1 - \sigma^B \sqrt{YF(t_0, t_j)}
\end{aligned}$$

Another model used to price caplets is the Bachelier model (normal) :

$$Caplet(t_{j-1}, t_j, K, \sigma) = B(t_0, t_j) YF(t_{j-1}, t_j) \sigma^N \sqrt{YF(t_0, t_j)} \left[dN(d) + n(d) \right]$$

where :

$$\begin{aligned} F = F(t_0, t_{j-1}, t_j) &: \text{the forward Libor rate} \\ B(t_0, t_j) &: \text{the Zero-Coupon value (discount factor)} \\ YF(t_{j-1}, t_j) &: \text{the year fraction between } t_{j-1} \text{ and } t_j \\ d &= \frac{F - K}{\sigma^N \sqrt{YF(t_0, t_j)}} \end{aligned}$$

Now let's describe the main steps of the stripping algorithm :
From the cap prices we can compute forward cap prices:

$$FwdCap(T_{i-1}, T_i, K, \Sigma_i^{fwd}) = Cap(T_0, T_i, K, \Sigma_i) - Cap(T_0, T_{i-1}, K, \Sigma_{i-1})$$

In order to find the piece-wise constant caplet vols, we use the so obtained FwdCap in the following equation

$$FwdCap(T_{i-1}, T_i, K, \Sigma_i^{fwd}) = \sum_{j=N_{i-1}+1}^{N_i} Caplet(t_{j-1}, t_j, K, \sigma_i)$$

and solve for the σ_i. This way we calculate caplet volatilities for all expiries across all strikes.[1]

We also take into account the ATM caps by using a similar solving procedure and interpolating missing points.

2.2 Linearly Interpolated Caplet Volatilities

For a given strike, we build a piece-wise linear caplet volatility curve (with respect to maturity date) such that we reprice all given caps with strike K.

Consider the first cap $Cap(T_0, T_1, K, \Sigma_1)$. Once we have specified the starting value (i.e. the value of the first caplet volatility), the linear caplet volatility curve until the cap term is completely determined by the the first cap value which we are aiming to match. Having specified the caplet vol curve until the term of the first cap, we can proceed with the second cap. The first caplet volatilities (until maturity of the first cap) are already implied from the first cap. The following volatilities until the term of the second cap are again completely determined by the value of the second cap, given the linear shape of the vol curve. For the third and all subsequent caps we apply the same procedure.

The above procedure leaves the first caplet volatility as a degree of freedom. One possible method is to specify it such that the caplet vols are flat until the first cap term.

[1] The caplet dates for caps of different tenors may in fact differ for a particular (overlapping) period. We account for that by some amendments of the described procedure, but the principle remains the same.

3 Applying the Caplet Smile to Swaptions

Having obtained caplet volatilities across expiries and strikes, we now apply these to swaptions. There are three different methods available for transferring the caplet information to swaptions [4]. For any swaption term expiring at T_i, we build strikes K^S and volatilities σ^S for swaptions from K^C and σ^C (caplets). Then this lifted swaption smile K^S, σ^S is used as an input to the shifted SABR calibration.

3.1 Strike adjustment

3.1.1 Absolute

$$K^S = K^C + K^S_{ATM} - K^C_{ATM}$$

3.1.2 Relative

$$K^S = K^C * \frac{K^S_{ATM}}{K^C_{ATM}}$$

3.2 Vol adjustment

3.2.1 Relative Rescaling

$$\sigma^S(K^S) = \sigma^C(K^C) * \frac{\sigma^S_{ATM}}{\sigma^C_{ATM}}$$

3.2.2 Parallel Shift

$$\sigma^S(K^S) = \sigma^C(K^C) + \sigma^S_{ATM} - \sigma^C_{ATM}$$

4 Interpolation in Strike Dimension : Static SABR Model

4.1 Model

The shifted SABR model can be described by the following equations [2]:

$$\begin{aligned}
dF_t &= \alpha_t (F_t + \epsilon)^\beta dW^1_t \\
d\alpha_t &= \nu \alpha_t dW^2_t \\
dW^1_t dW^2_t &= \rho dt \\
F_0 &= F_{ATM} \\
\alpha_0 &= \alpha_{ATM}
\end{aligned}$$

- dW_t^1 and dW_t^2 are two correlated one dimensional Brownian motions
- F_{ATM} is the value of the at-the-money forward at time zero
- α_{ATM} is the value of α_t at time zero such that the at-the-money option price is recovered
- ϵ is a constant deterministic shift allowing negative rates ($b = 0$ for the classic SABR model)

4.2 Approximate Formula

Using singular perturbation techniques, it can be shown that the price of a vanilla option under the SABR model is given by the Black formula, using the following volatility (lognormal vol σ^B and normal vol σ^N):

4.2.1 Original Hagan formulas

$$\begin{aligned}
\sigma^B(K, F_0) &= \frac{\alpha\left(1+\left(\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(F_0K)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha}{(F_0K)^{\frac{1-\beta}{2}}}+\frac{2-3\rho^2}{24}\nu^2\right)\tau\right)}{(F_0K)^{\frac{1-\beta}{2}}\left[1+\frac{(1-\beta)^2}{24}\ln^2\frac{F_0}{K}+\frac{(1-\beta)^4}{1920}\ln^4\frac{F_0}{K}\right]}\frac{z}{\chi(z)} \\
z &= \frac{\nu}{\alpha}(F_0K)^{\frac{1-\beta}{2}}\ln\frac{F_0}{K} \\
\chi(z) &= \ln\left(\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right)
\end{aligned}$$

$$\begin{aligned}
\sigma^N(K, F_0) &= \alpha(F_0K)^{\frac{\beta}{2}}\frac{1+\frac{1}{24}\ln^2\frac{F_0}{K}+\frac{1}{1920}\ln^4\frac{F_0}{K}}{1+\frac{(1-\beta)^2}{24}\ln^2\frac{F_0}{K}+\frac{(1-\beta)^4}{1920}\ln^4\frac{F_0}{K}}\frac{z}{\chi(z)} * \\
&\quad \left(1+\left(\frac{-\beta(2-\beta)\alpha^2}{24(F_0K)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha}{(F_0K)^{\frac{1-\beta}{2}}}+\frac{2-3\rho^2}{24}\nu^2\right)\tau\right) \\
z &= \frac{\nu}{\alpha}(F_0K)^{\frac{1-\beta}{2}}\ln\frac{F_0}{K} \\
\chi(z) &= \ln\left(\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right)
\end{aligned}$$

4.2.2 Hagan-Obloj formulas

We take into account the remarks from [3] and modify it for the shifted SABR [1]:

$$\sigma^B(K,F_0) = \frac{1}{x(\zeta(K))}\ln\left(\frac{F_0+\epsilon}{K+\epsilon}\right) *$$

$$\left[1+\left(g(K)+\frac{1}{4}\rho\nu\alpha\beta(F_0+\epsilon)^{\frac{\beta-1}{2}}(K+\epsilon)^{\frac{\beta-1}{2}}+\frac{1}{24}(2-3\rho^2)\nu^2\right)\tau\right]$$

$$g(K) = \frac{1}{24}(\beta-1)^2(F_0+\epsilon)^{\beta-1}(K+\epsilon)^{\beta-1}\alpha^2$$

$$\zeta(z) = \frac{\nu}{\alpha(1-\beta)}\left((F_0+\epsilon)^{1-\beta}-(K+\epsilon)^{1-\beta}\right)$$

$$x(\zeta) = \frac{1}{\nu}\ln\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2}-\rho+\zeta}{1-\rho}\right)$$

$$\sigma^N(K,F_0) = \frac{F_0-K}{x(\zeta(K))} *$$

$$\left[1+\left(g(K)+\frac{1}{4}\rho\nu\alpha\beta(F_0+\epsilon)^{\frac{\beta-1}{2}}(K+\epsilon)^{\frac{\beta-1}{2}}+\frac{1}{24}(2-3\rho^2)\nu^2\right)\tau\right]$$

$$g(K) = \frac{1}{24}(\beta^2-2\beta)(F_0+\epsilon)^{\beta-1}(K+\epsilon)^{\beta-1}\alpha^2$$

$$\zeta(z) = \frac{\nu}{\alpha(1-\beta)}\left((F_0+\epsilon)^{1-\beta}-(K+\epsilon)^{1-\beta}\right)$$

$$x(\zeta) = \frac{1}{\nu}\ln\left(\frac{\sqrt{1-2\rho\zeta+\zeta^2}-\rho+\zeta}{1-\rho}\right)$$

4.3 Calibration of the Model

The SABR model has β, F_0, σ_{atm} as exogenous inputs, leaving only ρ and ν to be calibrated on a set of market prices (see also [5]).

4.3.1 β parameter

This parameter has to be chosen beforehand as otherwise it forms an over-parametrization of the model. It may be adjusted dependent on the market; by default we will set $\beta = 0.5$.

Statistical procedures (log-log regression) are suggested in [2] to find a reliable and stable value for β.

4.3.2 α parameter

In the ATM case, if $K = F_0$, then $\frac{z}{\chi(z)} = 1$ [2] and

[2] If $z = 0$, we must take care that we do not divide by 0 and rather use the limit $\frac{z}{\chi(z)} = 1$

$$\sigma^B_{ATM} = \alpha(F_0+\epsilon)^{\beta-1}\left(1+\left(\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(F_0+\epsilon)^{2-2\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha}{(F_0+\epsilon)^{1-\beta}}+\frac{2-3\rho^2}{24}\nu^2\right)\tau\right) \quad (1)$$

$$\sigma^N_{ATM} = \alpha(F_0+\epsilon)^{\beta}\left(1+\left(\frac{(\beta^2-2\beta)}{24}\frac{\alpha^2}{(F_0+\epsilon)^{2-2\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha}{(F_0+\epsilon)^{1-\beta}}+\frac{2-3\rho^2}{24}\nu^2\right)\tau\right) \quad (2)$$

As we want to preserve σ^B_{ATM} or σ^N_{ATM}, knowing ρ, ν, it can be shown that α is the root of the cubic :

$$\frac{(1-\beta)^2\tau}{24(F_0+\epsilon)^{2-2\beta}}\alpha^3+\frac{\rho\beta\nu\tau}{4(F_0+\epsilon)^{1-\beta}}\alpha^2+(1+\frac{2-3\rho^2}{24}\nu^2\tau)\alpha-\sigma^B_{ATM}(F_0+\epsilon)^{1-\beta}=0$$

$$\frac{(\beta^2-2\beta)\tau}{24(F_0+\epsilon)^{2-2\beta}}\alpha^3+\frac{\rho\beta\nu\tau}{4(F_0+\epsilon)^{1-\beta}}\alpha^2+(1+\frac{2-3\rho^2}{24}\nu^2\tau)\alpha-\sigma^N_{ATM}(F_0+\epsilon)^{-\beta}=0$$

We use the Tartaglia algorithm to solve the real root that is of order 1.

4.3.3 ρ and ν parameters

We minimize the following error function using a minimization algorithm (conjugate gradient with automatic differentiation, Powell or Nelder-Mead simplex) :

$$\min_{\rho,\nu}\sum_i^n L(\sigma_i^{model}-\sigma_i^{market})$$

subject to the constraint to match the σ_{ATM} with formula (1) or (2), using a metric $L(x) = abs(x)$ or $L(x) = x^2$.

We also add a penalty to constraint the range of a parameters. For example, the correlation has to be between -1 and 1.

4.4 Converting volatility

In general, we calibrate the SABR using the volatility type of the caps and swaption. So if we use normal vols in input, we calibrate the SABR using those normal, compute the normal vols in output using the SABR approximation formula.

If we wanted to compute normal vols from lognormal vol (in case the caps and swaptions taken in input of the volcube are lognormal), we use this formula :

$$\sigma_N = \sigma_B\sqrt{F_0K}\frac{1+{}^1\!/_{24}\ln^2 {}^{F_0}\!/_K+{}^1\!/_{1920}\ln^4 {}^{F_0}\!/_K}{1+{}^1\!/_{24}(1-{}^1\!/_{120}\ln^2 {}^{F_0}\!/_K)\sigma_B^2\tau+{}^1\!/_{5760}\sigma_B^4\tau}$$

If we wanted to compute lognormal vols from that, we would solve the previous formula using a 1D numerical solver like Newton.

References

[1] Fabien Le Floc'h and Gary Kennedy. Explicit sabr calibration through simple expansions. *SSRN*, 2014.

[2] Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing smile risk. *Wilmott Magazine*, m:84–108, 2002.

[3] Jan Obloj. Fine-tune your smile correction to hagan et al. *arXiv*, 2008.

[4] Harvey Stein. Valuation of Exotic Interest Rate Derivatives-Bermudans and Range Accruals. *Quantitative Finance*, (December 2007):1–75, 2007.

[5] Graeme West. Calibration of the SABR model in illiquid markets. *Applied Mathematical Finance*, 12(4):371–385, 2005.