Analysis and development of SABR framework

Marek Musiela

Sydney March 2016

Mathematical models in finance

Black Scholes model

- Developed to value equity options
- Soon after applied to the FX market. The issue was how to deal with the domestic and foreign interest rates
- Later on adapted to the pricing of caps and swaptions. The issue was how to deal with the fact that the model assumes constant interest rates

Evolution of models

- Pricing models are calibrated to liquid options and used to price more exotic products.
- Liquidity growth in options markets lead to the introduction of models with local volatility to represent the market smile
- Stochastic volatility was added to represent volatility dynamics and the skew.

Evolution of markets

- Model is introduced at one point in time
- It works well within the range of variables generated by the market at that time
- Market conditions change and the model may not be adequate anymore
- The process of a model change may be very complicated. It often leads to booking of substantial losses

Constant elasticity of variance or CEV model also known as Cox model

CEV or Cox model

- The asset price X follows

$$dX_t = \sigma X_t^{\beta} dW_t, \qquad t > 0, \qquad X_0 = x$$

$$\sigma > 0, \qquad 0 \le \beta \le 1$$

- Reduces to the Black Scholes model
- Log-normal volatility $\sigma X^{\beta-1}$ changes inversely with the price
- Does not reduce to the Bachelier model

CEV model between ½ and 1

- Pathwise existence and uniqueness hold
- For

$$\frac{1}{2} \leq \beta \leq 1, \qquad X_0 = x > 0$$

we have for all t>0

$$X_t > 0, \qquad X_\infty = 0$$

CEV model between 0 and 1/2

- Pathwise existence and uniqueness hold in the class of nonnegative solutions
- Define $\tau = \inf\{t > 0 : X_t = 0 \}$
- For $0 \le \beta < \frac{1}{2}, \qquad X_0 = x > 0$

  we have $P(\tau < \infty) = 1, \qquad X_t = 0, \qquad \forall t \ge \tau$

$$\{\tau > t\} = \{X_t > 0\}$$

- Of course week solutions exist up to the explosion time and are unique in law

First passage time to zero

- Distribution of the first time process X hits zero can be calculated analytically
- Indeed one can show that it is given by

$$P(\tau > t) = v\left(\frac{x}{t^{\gamma}}\right), \qquad \gamma = \frac{1}{2(1-\beta)}$$

$$\frac{1}{2(1-\beta)}\xi v'(\xi) + \frac{\sigma^2}{2}\xi^{2\beta} v''(\xi) = 0,$$

$$v(0) = 1, v(\infty) = 1$$

First passage time to zero

- It follow that

$$P(\tau > t) = v\left(\frac{x}{t^{\gamma}}\right), \qquad \gamma = \frac{1}{2(1-\beta)}$$

$$v(\xi) = \frac{1}{c}\int_0^{\xi} \exp\left(-\frac{\eta^{2-2\beta}}{2\sigma^2(1-\beta)^2}\right) d\eta,$$

$$c = \int_0^{\infty} \exp\left(-\frac{\eta^{2-2\beta}}{2\sigma^2(1-\beta)^2}\right) d\eta$$

Behaviour for large t

- As t goes to infinity we get

$$P(\tau > t) \approx \frac{x}{ct^{\gamma}}, \qquad \gamma = \frac{1}{2(1-\beta)}$$

One dimensional diffusion

- The asset price follows

$$dX_s = \sigma(X_s, s)dW_s, \qquad \sigma(x,s) = \sigma x^{\beta}$$

- The arbitrage free price is given by

$$u(x,t) = E\left(\varphi(X_T) \middle| X_t = x\right)$$

Maximum principle

- The function satisfies u

$$u_t + \frac{1}{2}\sigma^2 u_{xx} = 0$$

- Its second derivative v satisfies

$$v_t + \frac{1}{2}\left(\sigma^2 v\right)_{xx} = 0$$

- The maximum principle implies that v is positive and hence u is convex

CEV model - properties

- The CEV process is a nonnegative martingale which converges to zero
- It has finite moments of any order
- It converts convex option payoffs into convex functions of the underlying asset

Sigma alpha beta rho model
also know as SABR

Brief history of SABR

- I came across SABR model for the first time in the late 90’s (1997-1998?) while working as a consultant for Paribas
- Two teams worked on the development of the approximate formula for implied volatility, one in London and one in New York
- The approximate formula developed by Pat Hagan in Paribas New York in 1998 was better

Brief history of SABR - 2

- A trader left Paribas and soon after the formula become well known in the market
- Very rapidly SABR become the market standard for quoting cap and swaption volatilities
- Today it is also used in the FX and equity markets

SABR dynamics

- The SABR model is defined by the following equations

$$dX_t = Y_t X_t^{\beta} dW_t^1, X_0 = x \qquad 0 \le \beta \le 1, \qquad \alpha \ge 0$$

$$dY_t = \alpha Y_t dW_t^2, Y_0 = \sigma \qquad d\langle W^1, W^2 \rangle_t = \rho dt, \qquad |\rho| \le 1$$

- It reduces to standard CEV model when volatility of the volatility is zero
- Correlation between the Brownian motions determines many important properties of the model

SABR parameters

- All SABR parameters have important and intuitive roles to play
- In the first part of this talk I’m focussing on the role played by the correlation between the two Brownian motions
- In the second I propose and analyse suitable dependence structure between fractional Brownian motions used in the definition of the fractional SABR model

From CEV to SABR

- SABR model can be written as follows

$$dX = X^{\beta} dM, \qquad dM = Y dW^1, \qquad dY = \alpha Y dW^2$$

- One can also easily show that quadratic variation of M satisfies

$$\langle M \rangle_{\infty} = \infty \qquad for \qquad \alpha = 0 \qquad CEV$$
$$\langle M \rangle_{\infty} < \infty \qquad for \qquad \alpha > 0 \qquad SABR$$

- This is a fundamental reason of the differences between CEV and SABR models

SABR distributions

- The distributions of the random variables

$$\langle M \rangle_\infty = \int_0^\infty Y_t^2 dt, \qquad \xi = \inf\left(t > 0 : W_t = \frac{\sigma}{\alpha} \right)$$

coincide.

- The density is given by

$$\frac{\sigma}{\alpha}\left(2\pi x^3\right)^{-\frac{1}{2}} \exp\left(-\frac{\sigma^2}{2\alpha^2 x} \right)$$

Laplace transform

- Indeed

$$E\exp\left(-\frac{\lambda^2}{2}\int_0^t Y_s^2\,ds\right)$$

solves the equation

$$u_t - \frac{\alpha^2}{2}y^2u_{yy} + \frac{\lambda^2}{2}y^2u = 0, \qquad u(0,t) = u(y,0) = 1$$

and as t goes to infinity the Laplace transform converges to the bounded solution of the following equation

Equation

$$-\frac{\alpha^2}{2}u_{yy}+\frac{\lambda^2}{2}u=0, \qquad u(0)=1$$

- Solution to this equation is given by

$$u(y)=\exp\left(-\frac{\lambda}{\alpha}y\right)$$

- Coincides with the Laplace transform of the first hitting time

‘Lognormal’ case

- The ‘Lognormal’ model is defined by

$$dX_t = Y_t X_t \, dW_t^1, \qquad \beta = 1, \qquad \alpha > 0$$

$$dY_t = \alpha Y_t \, dW_t^2, \qquad d\left\langle W^1, W^2 \right\rangle_t = \rho dt, \qquad |\rho| \le 1$$

- X is positive continuous local martingale and hence it converges a.s. as t goes to infinity to an integrable random variable we call

$$X_\infty$$

Measure change

- Define a new measure

$$\hat{P}(A) = E\left(\exp\left(\int_0^T Y dW^1 - \frac{1}{2}\int_0^T Y^2 dt \right) I_A \right), \qquad A \in F_T$$

- The following process is a Brownian motion under the new measure

$$d\begin{pmatrix} \hat{W}^1 \\ \hat{W}^2 \end{pmatrix} = d\begin{pmatrix} W^1 \\ W^2 \end{pmatrix} - \begin{pmatrix} Y \\ \rho Y \end{pmatrix} dt$$

‘Lognormal’ case -2

- Assume that the correlation between the Brownian motions is strictly positive. Then

$$EX_t = x\hat{P}(\tau \geq t), \qquad t \geq 0$$
$$\tau = \inf\left(t > 0 : X_t \notin (0, \infty)\right)$$

- It turns out that X is a strict positive super-martingale under P

'Lognormal' case -3

- When the correlation between the two Brownian motions is negative or zero then the martingale X is uniformly integrable because

$$EX_{\infty} = x\hat{P}(\tau = \infty) = x,$$
$$\tau = \inf\left(t \geq 0 : X_t \notin (0,\infty)\right)$$

Zero correlation

- If the correlation between the two Brownian motions is zero then the process X is a continuous positive martingale
- We also have for all positive T

$$\sup_{0 \le t \le T} E X_t \left| \log X_t \right| < \infty$$

Negative correlation

- If the correlation is negative the process X is a positive continuous martingale and for any m>1

$$\sup_{0 \le t \le T} E X_t^m < \infty \qquad iff \qquad \rho \le -\sqrt{\frac{m-1}{m}}$$

Propagation of convexity

- The call price on X with maturity T and strike K is given by

$$E(X_T - K)^+ = E(xZ - K)^+, \qquad Z = \exp\left(\int_0^T Y dW^1 - \frac{1}{2}\int_0^T Y^2 dt\right)$$

- The function

$$x \to E(X_T - K)^+$$

is convex

Back to SABR

- Assume the following SABR parameters

$$0 \le \beta < 1, \qquad \alpha > 0 \qquad |\rho| \le 1$$

- The distribution of X at infinity is given by the distribution of the random variable

$$\xi_\tau, \qquad d\xi = \xi^\beta dW^1, \qquad \tau = \inf\left(t > 0 : \sigma + \alpha W_t^2 = 0\right)$$

Uniform integrability

- The expected value of X at infinity is equal to x if and only if the correlation is strictly less than 1, i.e.,

$$EX_{\infty} = x \qquad \Leftrightarrow \qquad \rho < 1$$

- It follows that the martingale X is uniformly integrable

Propagation of convexity

- SABR model coverts convex payoffs into convex functions of the underlying asset X if and only if the correlation between the two Brownian motions is negative or zero
- The proof is rather technical and uses PDE and probabilistic arguments

Idea of proof

- Use the following transformation of variables

$$\xi_t = X_t Y_t^{-\theta}, \qquad \theta = \frac{1}{1-\beta}, \qquad 0 \le \beta < 1$$

- Observe that

$$d\xi_t = \xi_t^{\beta} dW_t^1 - \alpha\theta\xi_t dW_t^2$$
$$+ \frac{\alpha^2}{2}\theta(\theta+1)\xi_t dt - \rho\alpha\theta\xi_t^{\beta} dt$$

Convexity

- Convexity of

$$E\varphi\left(X_t\right) \qquad in \qquad x$$

is equivalent to the convexity of

$$E\varphi\left(\xi_t Y_t^{\theta}\right) \qquad in \qquad \xi$$

SABR for interest rates

- For a long period of time interest rates were very low and this regime may continue for a while
- Volatility relative to the level of rates (normal) is high
- Market prices suggest that one should consider models that allow for negative interest rates

SABR calibration

- In order to calibrate SABR to the market prices one needs to work with a very low beta parameter
- In this regime the model hits zero in finite time and stays there
- The probability of this event is very high
- In this situation risk management breaks down

Other asset classes

- SABR was developed to price caps and swaptions, i.e., options on LIBOR and swap rates.
- More recently people applied SABR to price equity and FX options.
- This generated new challenges for the SABR framework and suggests a new class of models inspired by SABR.

Fractional SABR

Definition of SABR

- Volatility in the SABR model follows a lognormal martingale. Hence we have the following representation

$$dX_t = Y_t X_t^{\beta} dW_t^1, X_0 = x \qquad 0 \le \beta \le 1, \qquad \alpha \ge 0$$

$$Y_t = \sigma \exp\left(\alpha W_t^2 - \frac{1}{2}\alpha^2 t \right)$$

$$d\left\langle W^1, W^2 \right\rangle_t = \rho dt, \qquad |\rho| \le 1$$

New class of models

- This suggests a more general class of models of the type

$$dX_t = Y_t X_t^{\beta} dW_t \ , X_0 = x \qquad 0 \le \beta \le 1, \qquad \alpha \ge 0$$

$$Y_t = \sigma \exp\left(Z_t - \frac{1}{2} VarZ_t \right),$$

- The process (W,Z) is jointly Gaussian with zero mean. W is a Brownian Motion.
- How to choose (W,Z) process?

Beyond SABR

- Log-volatility behaves like a fractional Brownian Motion with Hurst exponent H of order 0.1 at any reasonable time scale (ref. Gatheral, Jaisson, Rosenbaum 2014)
- At-the-money volatility skew is well approximated by a power law function of time to expiry (ref. Gatheral 2014 and Bayer, Friz and Gatheral 2015)

Implied volatility

- Implied volatility of an option as a function of log-moneyness and time to expiration is denoted by

$$\sigma_{BS}(k,\tau)$$

- At-the-money volatility skew is given by

$$\psi(\tau)=\frac{\partial}{\partial k}\sigma_{BS}(k,\tau)\bigg|_{k=0}$$

ATM volatility skew

- Empirical evidence suggests that for a large range of time to expiry

$$\psi(\tau) = C\tau^{-\alpha}, \qquad 0 < \alpha < \frac{1}{2}$$

- Fukasawa 2011 shows that a model where log-volatility behaves like fractional Brownian Motion with Hurst exponent H generates ATM volatility skew of the form

$$\psi(\tau) = C\tau^{H-\frac{1}{2}}$$

Fractional Brownian Motion

- A mean-zero Gaussian process Z is called fractional Brownian Motion if

$$EZ_sZ_t = \frac{1}{2}\left(|s|^{2H} + |t|^{2H} - |t-s|^{2H}\right), \qquad , s,t > 0,$$

$$0 < H < 1$$

- Z is self similar and has stationary increments.
- Z is a Brownian Motion when H=1/2.

Small problem

- When H=1/2 does not reduce to the Brownian motion unless we assume additionally that fractional Brownian motion is a continuous zero mean Gaussian process with the covariance function given on the previous slide

Process (W,Z)

- The classical SABR model is defined as a solution of a stochastic differential equation driven by two correlated Brownian Motion.
- If we choose Z to be a fractional Brownian Motion we need to define the joint distribution of (W,Z) where W is a Brownian Motion.

Dependence structure

- As we have seen in the classical SABR framework the level of correlation between the two Brownian Motions determines many important properties of the model
- Different choices of dependence between W and Z will imply different model properties.

Molchan-Golosov definition

- One can define Z using a correlated with W Brownian Motion B and Molchan-Golosov formula (F- Gauss hypergeometric function)

$$Z_t = \int_0^t K_H(t,s)\,dB_s$$

$$K_H(t,s) = c(H)\,F\left(\frac{1}{2}-H, H-\frac{1}{2}, \frac{1}{2}+H, \frac{s-t}{s}\right)(t-s)^{H-\frac{1}{2}}$$

$$c(H) = \sqrt{\frac{2H\Gamma\left(\frac{3}{2}-H\right)}{\Gamma\left(H+\frac{1}{2}\right)\Gamma(2-2H)}}$$

Covariance of W and Z

- In this case it is easy to see that

$$EW_sZ_t = \delta \int_0^{s \wedge t} K_H(t,u)du, \qquad \delta = EW_1B_1$$

- It follows that the process (W,Z) is Gaussian with self-similar marginals
- There are other representations of a fractional BM in terms of a correlated BM

Mandelbrot-Van Ness

- Alternatively we may use the Mandelbrot-Van Ness formula

$$Z_t = \frac{1}{c_1(H)} \int_R f_t(u) dB_u$$

$$f_t(u) = \left((t-u)^+\right)^{H-\frac{1}{2}} - \left((-u)^+\right)^{H-\frac{1}{2}}$$

$$c_1(H) = \left(\int_0^\infty \left((1+u)^{H-\frac{1}{2}} - u^{H-\frac{1}{2}} \right)^2 du + \frac{1}{2H} \right)^{\frac{1}{2}}$$

Covariance of W and Z

- In this case we get the following covariance function

$$EW_sZ_t = \frac{\delta}{c_1(H)\left(\frac{1}{2}+H\right)} \begin{cases} t^{\frac{1}{2}+H} - (t-s)^{\frac{1}{2}+H} & 0 < s < t \\ t^{\frac{1}{2}+H} & 0 < t < s \end{cases}$$

- From the modelling perspective one would like to know which one to choose

Multivariate BM

- A multivariate Brownian Motion has the following properties
  - It has stationary increments
  - It has independent increments
  - It is continuous
- It follows that it is Gaussian process

Multivariate fBM

- A multivariate fractional Brownian Motion with multidimensional parameter H is a process which satisfies the following three properties (Pierre-Olivier Amblard, et al)
  - It is Gaussian
  - It is self-similar with parameter H
  - It has stationary increments
- Additionally we assume that it is continuous

Multivariate self-similarity

- A p-dimensional mean-zero Gaussian process Z is self-similar if there exists a vector

$$H=(H_1,\dots,H_p), \qquad 0<H_i<1, \qquad i=1,\dots,p$$

such that for each $\lambda>0$ the following processes have the same finite dimensional distributions:

$$(Z_1(\lambda t),\dots,Z_p(\lambda t)), \qquad t\geq 0$$

$$(\lambda^{H_1}Z_1(t),\dots,\lambda^{H_p}Z_p(t)), \qquad t\geq 0$$

Fractional SABR - fSABR

- Fractional SABR is defined by the following equations

$$dX_t = Y_t X_t^{\beta} dW_t \;, X_0 = x \qquad 0 \le \beta \le 1, \qquad \alpha \ge 0$$

$$Y_t = \sigma \exp\left(\alpha Z_t - \frac{1}{2}\alpha^2 t^{2H} \right), \qquad 0 < H < 1$$

- The process (W,Z) is mean-zero bivariate fractional Brownian Motion with parameter

$$\left(\frac{1}{2}, H \right)$$

Properties

- Note that when the process Z is a fractional Brownian Motion the process Y is not a martingale, like in the classical SABR case.
- However it still has a constant mean value.
- In general it is not a semimartingale hence it falls outside of the class of classical models of volatility used for option pricing.

Covariance functions

- Covariance functions of bivariate process (W,Z) are given by the following formulas

$$EW_sW_t = \min(s,t), \qquad s,t > 0$$

$$EZ_sZ_t = \frac{1}{2}\left(|s|^{2H} + |t|^{2H} - |t-s|^{2H}\right) \qquad s,t > 0$$

Case 0<s<t

- There exist constants $\rho \in [-1,1],\ \eta \in R$ such that

$$EW_s Z_t = \frac{1}{2}\left((\rho+\eta)|s|^{\frac{1}{2}+H} + (\rho-\eta)|t|^{\frac{1}{2}+H} - (\rho-\eta)|t-s|^{\frac{1}{2}+H} \right), \qquad 0 < s < t$$

where

$$EW_1 Z_1 = \rho$$

- Ref. Amblard, Coeurjolly, Lavancier, Philippe, Surgailis 2009, 2010, 2011

Case 0<t<s

- In this case we have

$$EW_sZ_t = \frac{1}{2}\left((\rho+\eta)|s|^{\frac{1}{2}+H} + (\rho-\eta)|t|^{\frac{1}{2}+H} - (\rho+\eta)|t-s|^{\frac{1}{2}+H} \right), \qquad 0<t<s$$

- The question is for what range of constants this is a covariance function.
- When H=1/2 (W,Z) are correlated Brownian Motions as in the classical SABR case.

Admissible range

- Define the matrix

$$\Sigma(s,t) = \begin{pmatrix} EW_s W_t & EW_s Z_t \\ EW_s Z_t & EZ_s Z_t \end{pmatrix}$$

with the explicit expressions given in the previous pages.

- The following result is proved in Amblard et al. 2014

Covariance matrix

- The matrix $\Sigma(s,t)$ is a covariance matrix if and only if

$$\frac{\Gamma\left(\frac{3}{2}+H\right)^2}{\Gamma(2H+1)\sin\pi H}\left(\rho^2\sin^2\left(\frac{\pi}{2}\left(\frac{1}{2}+H\right)\right)+\eta^2\cos^2\left(\frac{\pi}{2}\left(\frac{1}{2}+H\right)\right)\right)\leq 1$$

- The above inequality imposes constraints on the parameters ρ,η and H.

Case 0<t<s revisited

- Note that in this case by martingale property of W and joint self similarity of (W,Z) we get

$$EW_sZ_t = E\left(Z_t E\left(W_s \middle| F_t\right)\right) = EW_tZ_t$$

$$= Et^{\frac{1}{2}}W_1 t^H Z_1 = t^{\frac{1}{2}+H} EW_1Z_1 = \rho t^{\frac{1}{2}+H}$$

- It follows from the previous general expression that

$$\eta = -\rho$$

Covariance structure

- We have the following expressions for the cross covariance

$$EW_s Z_t = \rho t^{\frac{1}{2}+H}, \qquad 0 < t < s$$

$$EW_s Z_t = \rho t^{\frac{1}{2}+H} - \rho(t-s)^{\frac{1}{2}+H}, \qquad 0 < s < t$$

- We also have

$$EW_s W_t = \min(s,t), \qquad s,t > o$$

$$EZ_s Z_t = \frac{1}{2}\left(|s|^{2H} + |t|^{2H} - |t-s|^{2H}\right), \qquad s,t > 0$$

Correlation

- It follows that the correlation must satisfy

$$\rho^2 \le \frac{\Gamma(2H+1)\sin \pi H}{\Gamma\left(H+\frac{3}{2}\right)^2}$$

- Gatheral et al. demonstrate that log-volatility behaves as fractional BM with H of order 0.1. For H=0.1 correlation must satisfy

$$|\rho| \le 0.58$$

Mandelbrot – Van Ness

- It turns out that when we use Mandelbrot Van Ness formula to define process Z using correlated with W Brownian Motion B we define bivariate process (W,Z) which has the following properties
  - It is Gaussian
  - It is self-similar with vector parameter (1/2,H)
  - It has stationary increments
  - It is continuous

Moments – case $0 \le \beta < 1,$

- The process X is given by

$$dX = Y X^{\beta} dW \ , X_0 = x \qquad 0 \le \beta < 1, \qquad \alpha \ge 0$$

- It follows that

$$EX^m_{t \wedge \tau_n} \le a(t) + b\int_0^t EX^m_{s \wedge \tau_n} ds$$

$$\tau_n = \inf\{s \ge 0 : X_s \ge n\}$$

- Hence X is a martingale in $L^m, \qquad 1 < m < \infty$

Fractional 'lognormal case'

- Fractional 'lognormal' model is defined by

$$dX_t = Y_t X_t dW_t \ , X_0 = x \qquad \alpha \geq 0$$

$$Y_t = \sigma \exp\left(\alpha Z_t - \frac{1}{2}\alpha^2 t^{2H} \right), \qquad 0 < H < 1$$

- The process (W,Z) is mean-zero bivariate fractional Brownian Motion with parameter

$$\left(\frac{1}{2}, H \right)$$

Properties

- X is a continuous, positive local martingale and hence integrable supermartingale
- Y is a lognormal process with constant mean
- Y converges to zero when t goes to infinity
- Moments of Y are

$$EY_t^p = \sigma^p \exp\left(\frac{1}{2}\alpha^2 t^{2H} p(p-1)\right)$$

Measure change

- Define a new measure

$$\hat{P}(A) = E\left(\exp\left(\int_0^T Y dW - \frac{1}{2}\int_0^T Y^2 dt\right) I_A\right), \qquad A \in F_T$$

- The following process is a Brownian motion under the new measure

$$d\hat{W} = dW - Ydt$$

W and Z independent

- Under the new measure the process Z is a fractional Brownian Motion with the Hurst exponent H
- X is a martingale
- Moreover

$$EX_t\left|\log X_t\right| \le x\log x + \frac{1}{2}x\sigma^2\int_0^t \exp\left(\alpha^2 s^{2H}\right)ds$$

References

- Multivariate fractional Brownian Motion: Amblard, Coeurjoly, Lavancier, Philippe, Surgailis,…
- Fractional SABR: Bayer, Fritz, Gatheral, Jaisson, Rosenbaum, Fukasawa, Comte, Renault,…