Fast Lane beyond Black Scholes

Backward Kolmogorov equation

The derivative pricing framework consists of four steps : (1) model definition using SDE, which describes how underlying dynamic governed by model paramters, (2) derivation of derivative price as PDE, which removes all random terms by dynamic hedging and arbitrage free pricing, coming up with a formula that relates all greeks, (3) analytic solution of PDE for plain vanilla claim, resulting in a closed form plain vanilla price, which is very useful for model calibration (parameter estimation) using prevailing market data, and (4) numerical methodologies for pricing exotic derivatives (such as trees and Monte Carlo) using parameters obtained in step 3 for OTC business. Thus there are huge differences between SDE and PDE, the former describes how random variables change with risk factors (sources of randomness, i.e. dz_1 dz_2), while the latter is an equation relating different greeks. In short, SDE is underlying whereas PDE is about derivatives.

PDE of derivatives price can be derived via (1) creating a risk free portfolio using dynamic hedging and (2) making it arbitrage free by allowing it to grow at the prevailing risk free rate. We can go through these two steps for different models, such as Black Scholes Heston and SABR etc, however it is too cumbersome. On the contrary, Dr Yan proposed to do it in one line using the fundalmental theorem of asset pricing, this is called ***backward Kolmogorov equation***. Given the following SDEs with 2 risk factors :

$$dx_t = \alpha_{1t}dt + \beta_{1t}dz_{1t} \qquad \textit{SDE of } 1^{st} \textit{ risk factor (for example : underlying)}$$
$$dy_t = \alpha_{2t}dt + \beta_{2t}dz_{2t} \qquad \textit{SDE of } 2^{nd} \textit{ risk factor (for example : volatility, in that case } \beta_{2t} \textit{ is called vol-of-vol)}$$

Given contingent claim having risk free growth rate on *LHS*, and the same contingent claim having *P&L* breakdown into Greeks on *RHS*, weights in the linear combination of Greeks are the drift and diffusion terms in the *SDE*s. *LHS* and *RHS* should be equivalent, as they are simply different perspectives of the same thing. Let *f* be a contingent claim :

$$rf_t = \partial_t f + \alpha_{1t}(\partial_x f) + \alpha_{2t}(\partial_y f) + \frac{1}{2}\beta_{1t}^2(\partial_{xx} f) + \rho\beta_{1t}\beta_{2t}(\partial_{xy} f) + \frac{1}{2}\beta_{2t}^2(\partial_{yy} f) \qquad \textit{where } \rho \textit{ is correlation between risks}$$

$$\Rightarrow \quad \alpha_{1t} = \textit{drift term of risk 1, weight for 1st order derivatives wrt risk 1}$$
$$\alpha_{2t} = \textit{drift term of risk 2, weight for 1st order derivatives wrt risk 2}$$
$$\beta_{1t} = \textit{diffusion term of risk 1, weight for 2nd order derivatives,}$$
$$\beta_{2t} = \textit{diffusion term of risk 2, weight for 2nd order derivatives}$$

Only diffusion coefficients exist in second order derivatives, since $dz_{1t}\, dz_{1t} = dt$, $dz_{1t}\, dz_{2t} = \rho dt$ and $dz_{2t}\, dz_{2t} = dt$, while drift coefficients do not exist in second order derivatives as $dt\, dt = 0$. In particular, when x is the random process of ***underlying*** and if y is the random process of ***underlying volatility square***, then we have :

$$rf_t = timedecay + \alpha_{1t}delta + \alpha_{2t}vega + \frac{1}{2}\beta_{1t}^2 gamma + \rho\beta_{1t}\beta_{2t}vanna + \frac{1}{2}\beta_{2t}^2 vomma \qquad \textit{Don't forget 1/2, and } \rho.$$

where *vega* is the derivative of claim with respect to second risk factor, i.e. ***volatility square*** (not volatility), and *vomma* means *volga*.

Example 1 : Black Scholes model

$$dS_t = (r-q)S_t dt + \sigma S_t dz_t \quad \Rightarrow \quad rf_t = \partial_t f + (r-q)S_t\partial_s f + \frac{1}{2}(\sigma S_t)^2\partial_{ss} f \qquad \textit{no vega, vanna nor volga}$$

Example 2 : Heston model

$$dS_t = (r-q)S_t dt + \sqrt{v_t}S_t dz_{1t} \qquad \textit{SDE of underlying}$$
$$dv_t = \kappa(\vartheta - v_t)dt + \sigma\sqrt{v_t}dz_{2t} \qquad \textit{SDE of volatility square}$$
$$dz_{1t}dz_{2t} = \rho dt \quad \Rightarrow \quad rf_t = \partial_t f + (r-q)S_t\partial_s f + \kappa(\vartheta - v_t)\partial_v f + \frac{1}{2}(\sqrt{v_t}S_t)^2\partial_{ss} f + \rho\sigma S_t v_t \partial_{sv} f + \frac{1}{2}(\sigma\sqrt{v_t})^2\partial_{vv} f$$

Example 3 : SABR model ***This result is not verified.***

$$dF_t = \sigma_t F_t^{\beta} dz_{1t} \qquad \textit{SDE of forward, hence no drift term, besides the power offers skew}$$
$$d\sigma_t = \alpha\sigma_t dz_{2t} \qquad \textit{SDE of volatility, which is log normal without drift}$$
$$dz_{1t}dz_{2t} = \rho dt \quad \Rightarrow \quad rf_t = \partial_t f + \frac{1}{2}(\sigma_t F_t^{\beta})^2\partial_{FF} f + \rho\alpha\sigma_t\sigma_t F_t^{\beta}\partial_{F\sigma} f + \frac{1}{2}(\alpha\sigma_t)^2\partial_{\sigma\sigma} f$$

Fundalmental theorem of asset pricing

Fundalmental theorem of asset pricing states that in a complete market (every claim has a price for all market states), an instrument is arbitrage free iff there exists a measure such that the instrument is martingale, i.e. the *SDE* of the claim does not have a drift term.

$$\begin{aligned} dx_t &= \beta_t dz_t \\ E[x_{t+\Delta t} \mid I_t] &= x_t + \beta_t E[\Delta z_t] \\ &= x_t + \beta_t \Delta t E[\varepsilon] \\ &= x_t \end{aligned}$$

Dr Yan once asked, can we blindly assign call option price as a sine function of underlying? Furthermore, its delta must be a cosine function, hence we can always dynamically hedge using this pricing formula. Is this claim correct? No although we can do hedging with sine and cosine, this pricing formula is obviously not a martingale under common models for underlying S_t, i.e.

$$E[\sin(S_{t+\Delta t}) \mid S_t] \neq \sin(S_t) \qquad \textit{for Black Scholes } S_t \textit{ nor Heston } S_t$$

Derive risk neutral measure for Black Scholes

Let's find the new measure. We can show that *PV* of contingent claim becomes martingale if we replace the underlying drift by risk free rate. In other words, by changing the physical measure into a new measure (which we call the risk neutral measure), the drift μ will become risk free r, whereas *PV* of claim becomes martingale. Why call the new measure risk neutral measure? That's because it requests a risk free rate in the presence of volatility, hence it is neutral in terms of risk preference. Here is the proof for BS model :

$$\begin{aligned} d(e^{-r(T-t)} f_t) &= re^{-r(T-t)} f_t dt + e^{-r(T-t)} \partial_t f_t dt + e^{-r(T-t)} \partial_s f_t dS_t + \frac{1}{2} e^{-r(T-t)} \partial_{ss} f_t (dS_t)^2 && \textit{using Ito's lemma} \\ &= re^{-r(T-t)} f_t dt + e^{-r(T-t)} \partial_t f_t dt + e^{-r(T-t)} \partial_s f_t (\mu S_t dt + \sigma S_t dz_t) + \frac{1}{2} e^{-r(T-t)} \partial_{ss} f_t (\mu S_t dt + \sigma S_t dz_t)^2 \\ &= e^{-r(T-t)} (\underbrace{r f_t + \partial_t f_t + \mu S_t \partial_s f_t + \frac{1}{2} (\sigma S_t)^2 \partial_{ss} f_t}_{=0 \quad if \quad \mu = r}) dt + e^{-r(T-t)} \sigma S_t \partial_s f dz_t && \textit{using Black Scholes PDE} \\ &= e^{-r(T-t)} \sigma S_t \partial_s f dz_t \end{aligned}$$

Derivation of risk neutral measure for different models should be done separately. Heston is left as an exercise.

Different solutions for Black Scholes

Black Scholes can be solved in the following ways :

- solving BSPDE analytically
- solving BSPDE with Green function of BS differential operator
- derive characteristic function from BSPDE and go for Fourier approach
- proving martingale with BSPDE and perform risk neutral pricing

How about Heston? Fourier approach method and Green function approach are applicable to Heston PDE too.

Dr Yan's view about Fourier method

According to Dr Yan's view, for whatever models, Black Scholes or Heston, Fourier method works like this :

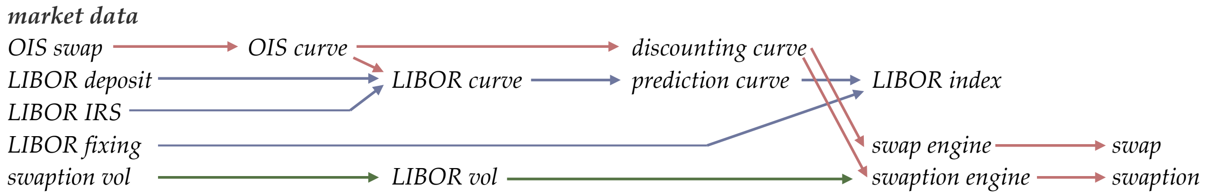
$$\begin{aligned} f_t(K) &= \int pdf(S_T)\, payoff(S_T, K) dS_T && \textit{pricing a batch of payoff with different strikes} \\ &= pdf(S_T) \otimes payoff(S_T, K) \\ &= FT^{-1}[FT[pdf(S_T)] \times FT[payoff(S_T, K)]] && \textit{convolution becomes multiplication} \end{aligned}$$

The Fourier transform of $pdf(S_T)$ can be found from BSPDE or Heston PDE. Is the above Fourier method the same as Carr Madan or Lewis? Yes this is Lewis approach.

Prediction vs discounting = index vs engine

In quant programming, underlying is modeled an index, which is composed of a historical set of fixing and future prediction using a stochastic model for nonlinear product (like Black Scholes for equity, Heston for FX rate, Hull White for IR rate) or simply combo of term structures, forming a prediction curve for linear product (like LIBOR + OIS for IR swap and foreign + domestic curve for FX swap). With prediction model or prediction curve, we can forecast future cashflows, which give PV when discounted, pricing tasks are done inside pricing engines, which should contain the risk free discounting curve, such as OIS curve.

Linear products are instruments that capture (i.e. having price depending on) drift term of underlying whereas non linear products are those which capture both the drift term and diffusion term. For example, here is the DAG for IR world :



Log normal vs normal
For a stochastic process, the drift term determines growth and the diffusion term determines distribution model.

$$
\begin{array}{lcll}
dx_t & = & \alpha_t x_t dt + \beta_t x_t dz_t & \textit{exponential growth with log normal model} \\
dx_t & = & \alpha_t x_t dt + \beta_t dz_t & \textit{exponential growth with normal model} \\
dx_t & = & \alpha_t dt + \beta_t x_t dz_t & \textit{linear growth with log normal model} \\
dx_t & = & \alpha_t dt + \beta_t dz_t & \textit{linear growth with normal model}
\end{array}
$$

Examples of log-normal model include Black Scholes, Heston, SABR with $\beta=1$, while examples of normal model include Hull White, Vasicek, Ornstein Uhlenbeck, SABR with $\beta=0$.

Can we replace standard normal in BS with other distributions?
When we model underlying Black Scholes, we are assuming that the random walk of the underlying follows Brownian motion. Yet, we never calibrate the Black Scholes volatility using historical statistics of the underlying, instead we derive an analytic solution for vanilla option, and imply the Black Scholes volatility using market data of option. The reason for doing so is that BS volatility is not constant in time, statistics represents the history, whereas market traded price represents expectation of the future, hence we prefer the latter. Yet from the observations of real market data, we found that there is skew in volatility smile, implying that Black Scholes is not sufficient to model the underlying. To solve the problem, we cannot replace $dz_t \sim \varepsilon(0,1)$ with other distributions, as stochastic calculus makes good uses of the nice properties of Gaussian, such as sum of Gaussians remains Gaussian. If unit normal is replaced by something else, Itos lemma does not hold. Therefore practitioners prefer to use stochastic volatility, i.e. modelling volatility with a separate *SDE*, forming a 2 factors model.

Physical intepretations from different PDEs
What does different *PDEs* imply? Suppose f is any single contingent claim or any dynamically hedged portfolio :

$$
\begin{array}{lcl}
rf_{BS,t} & = & \partial_t f_{BS} + (r-q)S_t\partial_s f_{BS} + \frac{1}{2}(\sigma S_t)^2\partial_{ss} f_{BS} \\
rf_{HT,t} & = & \partial_t f_{HT} + (r-q)S_t\partial_s f_{HT} + \kappa(\vartheta - v_t)\partial_v f_{HT} + \frac{1}{2}(\sqrt{v_t}S_t)^2\partial_{ss} f_{HT} + \rho\sigma S_t v_t\partial_{sv} f_{HT} + \frac{1}{2}(\sigma\sqrt{v_t})^2\partial_{vv} f_{HT}
\end{array}
$$

If the portfolio is perfectly hedged, i.e. having zero delta, then in Black Scholes world, time decay is related to gamma as :

$$
rf_{BS,t} = \underbrace{\left[\partial_t + \frac{1}{2}(\sigma S_t)^2\partial_{ss}\right]}_{operator} f_{BS,t} \qquad \textit{equation 1}
$$

However, this relation does not apply for Heston model, as there are vega, vanna and volga in Heston *PDE* :

$$
rf_{HT,t} \neq \underbrace{\left[\partial_t + \frac{1}{2}(\sigma S_t)^2\partial_{ss}\right]}_{operator} f_{HT,t} \qquad \textit{equation 2}
$$

In reality, market data tells us that *equation 1* is not true, implying that we need to handle volatility risk by adding volatility related terms into the *PDE*, like Heston. Please note that Dupire handles instantaneous volatility only, it does not consider volatility risk.

Linear vs nonlinear products

According to the fundalmental theorem of asset pricing, security price can be found by risk neutral pricing, i.e. risk free discounted expectation of future payoff under risk neutral measure :

$$price_t = \sum_n DF(T_n) E_Q[f_n(X_{T_n})] \qquad \textit{where } X_t \textit{ is the underlying random variable, } f_n \textit{ is the } n^{th} \textit{ payoff at time } T_n$$

A linear product is contract with payoffs being linear to the underlying random variable, thus expectation of payoff can capture the first order moment of underlying random variable only. Typical example is forward contract.

$$\begin{array}{lll} & f(x) & = & ax+b \\ \Rightarrow & E[f(X_T)] & = & aE[X_T]+b \end{array}$$

Non linear product is contract with payoffs being nonlinear to the underlying random variable, thus expectation of payoff captures higher order moments of underlying random variable, including volatility, skew and even kurtosis. The main reason is expectation cannot be moved inside a nonlinear function. Typical examples include various types of option.

$$\begin{array}{llll} & E[f(X_T)] & \neq & f(E[X_T]) \\ \text{if} & f(x) & = & x^2 \\ \text{if} & f(x) & = & (x-const)^+ \end{array}$$

Why Black Scholes? Can we model underlying sine?

Can we simply model underlying with deterministic sine function? We can find its Greeks anyway (delta is cosine in this case) and hence it can be perfectly hedged. What is the problem with that? Now recall that when we derive the fundamental theorem of asset pricing, there are two main steps, constructing a risk free portfolio (hedging) and ensuring arbitrage free, which leads to martingale property (also known as risk neutral measure). The sine model fulfills hedging only (we can always hedge claims derived from sine underlying by longing cosine shares of underlying), however, the sine model does not fulfill martingale requirement. In continuous context, a martingale process must have zero drift term. Therefore we cannot win a Nobel price with the sine model.