Quantitative Volatility Trading

Daniel Bloch

Quant Finance

15/16th of February 2016

Introduction

- We introduce option pricing and discuss the BS framework.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.
- We introduce financial risk and present the Greeks.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.
- We introduce financial risk and present the Greeks.
- We discuss the importance of asset returns.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.
- We introduce financial risk and present the Greeks.
- We discuss the importance of asset returns.
- We discuss the importance of smile dynamics.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.
- We introduce financial risk and present the Greeks.
- We discuss the importance of asset returns.
- We discuss the importance of smile dynamics.
- We present market anomalies.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.
- We introduce financial risk and present the Greeks.
- We discuss the importance of asset returns.
- We discuss the importance of smile dynamics.
- We present market anomalies.
- We consider the choice of a volatility model and discuss the difficulty of generating arbitrage-free smiles.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.
- We introduce financial risk and present the Greeks.
- We discuss the importance of asset returns.
- We discuss the importance of smile dynamics.
- We present market anomalies.
- We consider the choice of a volatility model and discuss the difficulty of generating arbitrage-free smiles.
- We present a globally arbitrage-free parametric model, the MixVol model.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.
- We introduce financial risk and present the Greeks.
- We discuss the importance of asset returns.
- We discuss the importance of smile dynamics.
- We present market anomalies.
- We consider the choice of a volatility model and discuss the difficulty of generating arbitrage-free smiles.
- We present a globally arbitrage-free parametric model, the MixVol model.
- We discuss the model parameters, the shape of the smiles produced, and its associated risk management.

Introduction

- We introduce option pricing and discuss the BS framework.
- We present the implied volatility and define the no-arbitrage.
- We introduce financial risk and present the Greeks.
- We discuss the importance of asset returns.
- We discuss the importance of smile dynamics.
- We present market anomalies.
- We consider the choice of a volatility model and discuss the difficulty of generating arbitrage-free smiles.
- We present a globally arbitrage-free parametric model, the MixVol model.
- We discuss the model parameters, the shape of the smiles produced, and its associated risk management.
- We briefly introduce variance swaps and basket options in view of analysing Dispersion Trading.

The need for derivative products

- Uncertainty related to traditional economic variables, imposing costs and risks on the society, must be hedged away.
- Thus, derivative products developed.
- Both parties entering the deal must agree on the price of the contingent claim.
- The decisions are made contingent on the price behaviour of the underlying securities.
- The uncertainty is based on future trajectories of the risky asset seen as possible scenarios.
- Thus, we need to understand the properties of market returns.

Pricing in a probabilistic approach

- We consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathcal{F}_t$ is a right continuous filtration including all $\mathbb{P}$ negligible sets in $\mathcal{F}$.
- Using the concept of absence of arbitrage opportunities (AAO), asset returns must reflect the fact that the riskier the asset the higher the returns, resulting in an instantaneous drift higher than the risk-free rate.
- In an arbitrage-free market, the assumption of linearity of prices leads to the existence of a risk-neutral probability measure $\mathbb{Q}$ equivalent to the historical measure $\mathbb{P}$.
- Since our underlying is exchanged in a market, its discounted price needs to be a martingale under the risk-neutral measure $\mathbb{Q}$.

Pricing in a probabilistic approach Ctd

- We let $\mathcal{Q}$ be the set of coexistent equivalent measures $\mathbb{Q}$. In a complete market the risk-neutral probability is unique, but it is not the case in an incomplete market where it must be specified.
- For simplicity of exposition, given a market price of risk λ, we will assume that there exist an equivalent martingale measure $\mathbb{Q}^\lambda$ denoted by $\mathbb{Q}$.
- The price of a European contingent claim $C(t,x)$ on $]0,T]\times]0,+\infty[$ under the risk-neutral measure $\mathbb{Q}$ is

$$C(t,X_t) = E^Q[e^{-\int_t^T r_s\,ds} h(X_T)|\mathcal{F}_t] \tag{0.1}$$

where X_T is a $\mathbb{Q}$-random variable and h is a sufficiently smooth payoff function.

The replication portfolio: complete market

- We let $(X(t))_{t\in[0,T]}$ be a continuous semimartingale on the horizon $[0,T]$, representing the price process of a risky asset.
- A contingent claim is given by $H = h(X(T))$ for some function h, where H is a random variable such that $H \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.
- $M(t)$ is the risk-free money account, $\alpha(t)$ is the amount of asset held at time t, and $\beta(t)$ is the money account held at time t.
- The value of a portfolio at time t satifies

$$V(t) = \alpha(t)X(t) + \beta(t)M(t) \ , \ 0 \leq t \leq T$$

- Set $M(t) = 1$ for all $0 \leq t \leq T$. The trading strategy (α, β) is admissible, such that the value process $V(t)$ is square-integrable and have right-continuous paths defined by

$$V(t) = V_0 + \int_0^t \alpha(s) dX(s)$$

The replication portfolio: complete market Ctd

- For $\mathbb{Q}$-almost surely, every contingent claim H is attainable and admits the representation

$$V(T) = H = V_0 + \int_0^T \alpha(s) dX(s)$$

where $V_0 = E^Q[H]$.

- The strategy is self-financing, meaning the cost of the portfolio is a constant

$$V(t) - \int_0^t \alpha(s) dX(s) = V_0$$

where V_0 is a perfect hedge.

The assumption of lognormality

- Assuming the returns between two periods are measured by the difference between the logarithms of the asset prices, and that returns are modelled with a Brownian motion with volatility σ and a drift $\mu - \frac{1}{2}\sigma^2$, then for $\{S_t; t \in [0, T]\}$ the returns $\log S_t - \log S_s$ are normally distributed with mean $(\mu - \frac{1}{2}\sigma^2)(t - s)$ and variance $\sigma^2(t - s)$.
- Given $0 < t_1 < .. < t_n$, then the relative prices $\{\frac{S_{t_{i+1}}}{S_i}; 0 \leq i \leq n-1\}$ are independent and follow the same law. That is, there exists a Brownian motion $\widehat{W}$ such that

$$S_t = f(t, \widehat{W}_t) = xe^{\mu t + \sigma \widehat{W}_t - \frac{1}{2}\sigma^2 t}$$

with initial condition $S_0 = x$.

- Applying Ito's lemma, we get the dynamics of the stock price as

$$\frac{dS_t}{S_t} = \mu dt + \sigma d\widehat{W}_t$$

- For $\sigma \neq 0$, the market compares the rate of return per unit of time μ to the return a risk-free asset, so that the market reference is $\mu - r$.
- We let λ be the market price of risk assigned to the noise $\widehat{W}$ as

$$\lambda = \frac{\mu - r}{\sigma}$$

- We can then rewrite the dynamics of the stock price as

$$\frac{dS_t}{S_t} = rdt + \sigma\big(d\widehat{W}_t + \lambda dt\big)$$

and the market price of risk is not specific to the asset price S_t, but to the noise $\widehat{W}_t$.

The Black-Scholes framework

- Black et al. **?** derived a theoretical valuation formula. To do so, they assumed the following ideal conditions in the market

1. The short-term interest rate is known and constant through time.
2. The stock price follows a random walk in continuous time with a variance rate proportional to the square of the stock price.
3. The stock pays no dividends or other distributions.
4. The option is European.
5. There are no transaction costs in buying or selling the stock or the option.
6. It is possible to borrow any fraction of the price of a security to buy it or to hold it, at the short-term interest rate.
7. There are no penalties to short selling.

- Following that argument, BS defined the price of a derivative option as the price of its hedge.

The Black-Scholes framework Ctd

- A direct application of the above assumption is that we can express the Black-Scholes price dynamics as

$$dC_{BS}(t,S_t) = r_t C_{BS}(t,S_t)dt + \pi_t \sigma_t (d\widehat{W}_t + \lambda_t dt)$$

where $\pi_t = \alpha(t)S_t$.

- Using Girsanov's theorem, we can re-express that price dynamics under the risk-neutral measure $\mathbb{Q}$ by applying the change of measure

$$W_t = \widehat{W}_t + \int_0^t \lambda_s ds$$

where W_t is a $\mathbb{Q}$-Brownian motion.

- Thus, the Black-Scholes price at all future time is a solution to the following stochastic differential equation with terminal condition

$$\begin{aligned} dC_{BS}(t,S_t) &= r_t C_{BS}(t,S_t)dt + \pi_t \sigma_t dW_t \\ C_{BS}(T,S_T) &= h(S_T) \end{aligned} \qquad (0.2)$$

- The option price depends neither on the return μ of the risky asset nor the market price of risk λ, but the risk due to the variations of the risky asset still impacts the option price via the volatility σ.

The Black-Scholes formula

- The price of a call option seen at time t with strike K and maturity T is

$$
\begin{aligned}
C_{BS}(t,x,K,T) &= xe^{-q(T-t)}N(d_1(T-t,F(t,T),K)) \\
&- Ke^{-r(T-t)}N(d_2(T-t,F(t,T),K))
\end{aligned}
$$

  where $F(t,T) = xe^{(r-q)(T-t)}$ is the forward price.
- Further, we have

$$
d_2(t,x,y) = \frac{1}{\sigma\sqrt{t}}\log\frac{x}{y} - \frac{1}{2}\sigma\sqrt{t} \text{ and } d_1(t,x,y) = d_2(t,x,y) + \sigma\sqrt{t}
$$

The Black-Scholes formula Ctd

- We let $P(t,T)$ be the discount factor, $Re(t,T)$ be the repo factor and we define

$$\eta = \frac{K}{F(t,T)} = \frac{KP(t,T)}{xRe(t,T)}$$

to be the forward moneyness of the option.

- Defining the total variance (TV) as $\omega(t) = \sigma^2 t$, and writing $K = \eta F(t,T)$, the call price becomes

$$\frac{C_{TV}(t,x,K,T)\big|_{K=\eta F(t,T)}}{xRe(t,T)} = \big(N(d_1(\eta,\omega(T-t)) - \eta N(d_2(\eta,\omega(T-t)))\big)$$

where

$$d_2(\eta,\omega(t)) = -\frac{1}{\sqrt{\omega(t)}}\log\eta - \frac{1}{2}\sqrt{\omega(t)}$$

and $d_1(\eta,\omega(t)) = d_2(\eta,\omega(t)) + \sqrt{\omega(t)}$.

Some properties

- When the spot price $S_t = x$ is close to the ATM forward strike, $K \approx xe^{(r-q)(T-t)}$, the call price can be approximated with

$$\frac{C_{BS}(t,x,K,T)}{xRe(t,T)} \approx \big(\frac{1}{2} + \frac{1}{5}\sigma\sqrt{T-t}\big) - Ke^{-r(T-t)}\big(\frac{1}{2} - \frac{1}{5}\sigma\sqrt{T-t}\big)$$

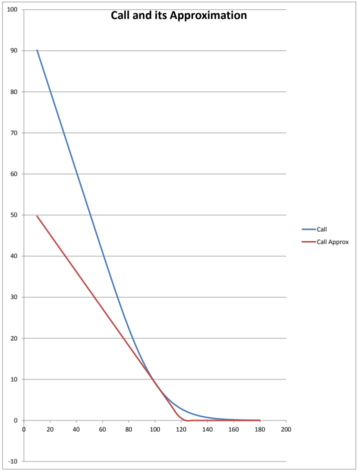
  which is linear in the spot price and the volatility.
- When the spot price $S_t = x$ is exactly ATM forward $K = F(t,T)$, the call price can be approximated with

$$\frac{C_{TV}(t,x,K,T)\big|_{K=F(t,T)}}{xRe(t,T)} \approx 0.4xe^{-q(T-t)}\sqrt{\omega(T-t)}$$

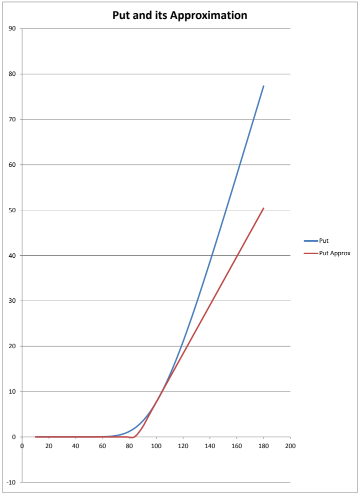
  which is linear in the spot price and the square root of the total variance.
- In the Black-Scholes model option prices are homogeneous functions

$$C_{BS}(t,\lambda x,\lambda K,T) = \lambda C_{BS}(t,x,K,T)$$

Linear approximation: Call



Linear approximation: Put



Defining the implied volatility

Definition

A call price surface parameterised by s is a function

$$C : [0, \infty) \times [0, \infty) \to \mathbb{R}$$
$$(K, T) \to C(K, T)$$

along with a real number $s > 0$.

- The implied volatility (IV) is a mapping from time, spot prices, strike prices and expiry days to $\mathbb{R}^+$

$$\Sigma : (t, S_t, K, T) \to \Sigma(t, S_t; K, T)$$

- Given the option price $C(t, S_t, K, T)$ at time t for a strike K and a maturity T, the market implied volatility $\Sigma(t, S_t; K, T)$ satisfies

$$C(t, S_t, K, T) = C_{BS}(t, S_t, K, T; \Sigma(K, T))$$

where $C_{BS}(t, S_t, K, T; \sigma)$ is the BS formula.

Defining the implied volatility Ctd

- The implied volatility is obtained by inverting the Black-Scholes formula $C_{BS}^{-1}(C(t, S_t, K, T); K, T)$.
- We refer to the two-dimensional map

$$(K, T) \to \Sigma(K, T)$$

  as the implied volatility surface.
- There is no closed-form solution for the IV, and Gerhol **?** showed that the BS-model did not belong to a class of functions for which solutions could easily be found.
- Iterative algorithms (Newton-Raphson method along the vega) have many shortages such as explosion when the gradient tends to zero.
- Researchers developed formulas to find an approximated implied volatility.

Graph of IVS

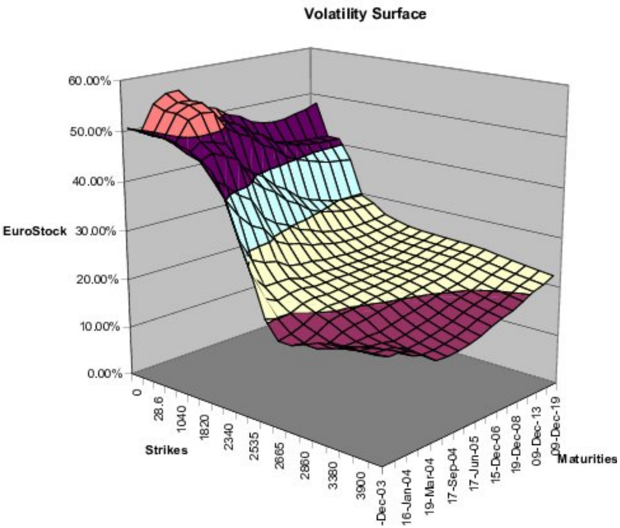


Figure: EuroStock Volatility Surface on the $8/12/2003$.

Arbitrage-free conditions: The call price surface

Theorem (necessary and sufficient conditions)

Let $s > 0$ *be a constant,* T *be the maturity, and let* $C : (0, \infty) \times [0, \infty) \to \mathbb{R}$ *satisfy the following conditions:*

(A1) *(Convexity in K)*
$C(\cdot, T)$ *is a convex function,* $\forall T \geq 0$

(A2) *(Monotonicity in* T*)*
$C(K, \cdot)$ *is non-decreasing,* $\forall K \geq 0$

(A3) *(Large Strike Limit)*
$\lim_{K \to \infty} C(K, T) = 0$, $\forall T \geq 0$

(A4) *(Bounds)*
$(s - e^{-rT} K)^+ \leq C(K, T) \leq s$, $\forall K > 0, T \geq 0$

(A5) *(Expiry Value)*
$C(K, 0) = (s - K)^+$, $\forall K > 0$

Then, there exists a non-negative Markov martingale X*, with the property that*

$$C(K, T) = \mathbb{E}[e^{-rT}(X_T - K)^+] \text{ , } \forall K \text{ and } \forall T \geq 0$$

Arbitrage-free conditions: The volatility surface

- To infer static arbitrage from implied volatility surface, one must first establish necessary and sufficient conditions on the call price surface for it to be free of static arbitrage, and then translate these conditions into conditions on the implied volatility surface.
- We are going to state the theorem given by Roper **?**. We set $x = \ln\left(\frac{K}{F(t,T)}\right)$ and we define the time scaled implied volatility as

$$\begin{aligned}\Xi : \mathbb{R} \quad \times \quad & [0,\infty) \to [0,\infty] \\ & (x, T) \to \sqrt{T-t}\Sigma(F(t,T)e^x, T)\end{aligned}$$

where $F(t,T)$ is the forward price. That is, $\Xi(x,T) = \sqrt{\omega(\overline{\eta},T)}$.

Arbitrage-free conditions: The volatility surface Ctd

Theorem

Let $F(t,T) > 0$*,* T *be the maturity, and let* $\Xi : \mathbb{R} \times [0,\infty) \to \mathbb{R}$*. Let* Σ *satisfies the following conditions:*

(IV1) *(Smoothness)*
$\forall T > 0$, $\Xi(\bullet, T)$ *is twice differentiable.*

(IV2) *(Positivity)*
$\forall x \in \mathbb{R}$ *and* $T > 0$, $\Xi(x,T) > 0$.

(IV3) *(Durrleman Condition)*
$\forall x \in \mathbb{R}$ *and* $T > 0$,

$$\left(1 - \frac{x\partial_x \Xi}{\Xi}\right)^2 - \frac{1}{4}\Xi^2(\partial_x \Xi)^2 + \Xi \partial_{xx}\Xi \geq 0$$

(IV4) *(Monotonicity in T)*
$\forall x \in \mathbb{R}$, $\Xi(x, \bullet)$ *is non-decreasing.*

Arbitrage-free conditions: The volatility surface Ctd

Theorem (Ctd)

(IV5) *(Large-Moneyness Behaviour)*
$\forall T > 0$, $\lim_{x \to \infty} d_1(x, \Xi(x, T)) = -\infty$.

(IV6) *(Value at Maturity)*
$\forall x \in \mathbb{R}$, $\Xi(x, 0) = 0$.

Then the call price surface parameterised by $F(t, T)$ is free from static arbitrage. In particular, there exists a non-negative Markov martingale X, with the property that

$$C(K, T) = \mathbb{E}[e^{-rT}(X_T - K)^+ | X_0 = F(0, T)] \ , \forall K \text{ and } \forall T \geq 0$$

Defining no-arbitrage on the volatility surface

- Other asymptotic behaviours of d_1 and d_2 hold in great generality. For instance, the Small-Moneyness Behaviour (SMB) condition states

$$\lim_{\overline{\eta} \to -\infty} d_2(\overline{\eta}, \omega(\overline{\eta}, t)) = \infty$$

- or, equivalently

$$\lim_{K \to 0} \partial_K C(K, t) = -P(0, t)$$

- These conditions were then recasted in terms of implied volatility providing a complete characterisation of an IVS free from static arbitrage.

Definition

Arbitrage-Free Implied Volatility Surface
Assuming deterministic interest and repo rates, an implied volatility surface is free from static arbitrage if and only if the following conditions are satisfied

(A1) *(Monotonicity of total variance)*
$\forall x$ *and* $T > 0$, $\partial_T \omega(x, T) \geq 0$

(A2) *(Convexity in K)*
$\forall T > 0$, $\partial_{KK} C(K, T) \geq 0$

(A3) *(Large-Moneyness Behaviour)*
$\lim_{\overline{\eta} \to \infty} d_1(\overline{\eta}, \omega(\overline{\eta}, t)) = -\infty$ *and* $\lim_{\overline{\eta} \to \infty} d_2(\overline{\eta}, \omega(\overline{\eta})) = -\infty$

(A4) *(Small-Moneyness Behaviour)*
$\lim_{\overline{\eta} \to -\infty} d_2(\overline{\eta}, \omega(\overline{\eta}, t)) = \infty$ *and* $\lim_{\overline{\eta} \to -\infty} d_1(\overline{\eta}, \omega(\overline{\eta})) = \infty$

Pricing principles

- Options provide leverage and give the ability to take a view on volatility as well as equity direction.
- The nature of options makes it possible to create a considerable number of speculative trading strategies.
- However, parsimonious information on option prices is available in time and space, and can only be accounted for with the No-Dominance law stating that valuation obeys to a trivial monotonicity rule.

Proposition

No dominance principle

Let X be the gain from a portfolio strategy with initial cost x. If $X \geq 0$ in every state of the world, then $x \geq 0$.

Pricing principles Ctd

The dominance principle is based on the following fundamental rules:

(D1) The law of one price: the same asset must trade at the same price on all markets.

(D2) Two assets with identical cash flows in the future must have the same price to start with.

(D3) An asset with a known future price must trade today at the price discounted at the risk free rate.

As an example of financial instruments following these rules we have:

1. Futures: they must satisfy rule number $(D3)$, any deviation from this equality leads to arbitrage.
2. Derivaties: buy and sell the same asset on two different markets, namely the spot market and the derivative market. Hence, they must satisfy rule number $(D1)$.

The Digital Bond and Digital Stock

- We can calculate the prices of some special contingent claims via the derivation of the call price $C(t, S_t, T, K)$ as follow

$$\begin{aligned} \text{Digital} &= -\frac{\partial C(t, S_t, T, K)}{\partial K} = P(t, T) E_t[I_{\{S_T \geq K\}}] \\ \text{Density} &= \frac{\partial^2 C(t, S_t, T, K)}{\partial K^2} = P(t, T) E_t[\delta(S_T - K)] \end{aligned}$$

- As a result, the price of a call option can be expressed in terms of those quantities as

$$\frac{C(t, S_t, T, K)}{P(t, T)} = E_t[S_T I_{\{S_T \geq K\}}] - K E_t[I_{\{S_T \geq K\}}] = \Delta_S + K \Delta_K$$

where Δ_K is the probability that the stock price end up higher or equal to the strike price at maturity, and Δ_S is a modified probability that the stock price end up higher or equal to the strike price at maturity.

The Digital Bond and Digital Stock Ctd

- the Digital Bond $D_B(S, t, T; \xi)$ is the value at time t of receiving one dollar at the maturity T if and only if a probabilist event ξ occurs.
- the Digital Stock $D_S(S, t, T; \xi)$ is the value at time t of receiving one share of the stock at the maturity T if and only if a probabilist event ξ occurs.
- The pricing of other European derivatives with piecewise linear and path-independent payoffs only requires valuing Digital Bond and Digital Share with event $\xi = \{L < S_T < H\}$ for some constants L and H.
- In the special case where $\xi = \{S_T > K\}$, we have $D_B(S, t, T; \xi) = P(t, T)\Delta_K$ and $D_S(S, t, T; \xi) = P(t, T)\Delta_S$. Geman et al. **?** showed that $\Delta_K = P(S_T \geq K)$ and $\Delta_S = P^S(S_T \geq K)$.

The digital option

- a Digital option $D(K, T) = P(t, T)\Delta_K$ pays \$1 when the stock price S_T is greater than the strike K, and zero otherwise. From a static replication argument, its price is given by

$$D(K, T) = \lim_{\Delta K \to 0} \frac{C(K, T) - C(K + \Delta K, T)}{\Delta K} = -\frac{\partial}{\partial K} C(K, T)$$

- Given $C(K, T) = C_{BS}(K, T; \Sigma_{BS}(K, T))$ with $\Sigma_{BS}(K, T)$, the Digital option becomes

$$D(K, T) = -\frac{\partial}{\partial K} C_{BS}(K, T; \Sigma_{BS}) - \frac{\partial}{\partial \Sigma} C_{BS}(K, T; \Sigma(K, T)) \frac{\partial}{\partial K} \Sigma(K, T)$$

- When $r = q = 0$, $T = 1$, $S_0 = 100$ and $K = 100$, then $\eta = 1$ and we get $d_2 = -\frac{1}{2}\Sigma_{BS}\sqrt{T}$. For a skew of 2.5% per 10% change in the strike and ATM volatility $\Sigma_{ATM} = 25\%$, the Digital option becomes

$$\begin{aligned} D(100, 1) &= N[-\frac{\Sigma_{ATM}}{2}] - S_0 n[\frac{\Sigma_{ATM}}{2}] \frac{-0.025}{0.1 S_0} \\ &\approx 0.45 + 0.25 \times 0.4 = 0.55 \end{aligned}$$

The butterfly option

- We consider a butterfly strategy at K where we are long a call option with strike $K - \Delta K$, long a call option with strike $K + \Delta K$, and short two call options with strike K. The value of the butterfly is

$$\begin{aligned} B(t_0, K, T) &= C(K - \Delta K, T) - 2C(K, T) + C(K + \Delta K, T) \\ &\approx P(t_0, T)\phi(t_0; K, T)(\Delta K)^2 \end{aligned}$$

  where $\phi(t_0; K, T)$ is the PDF of S_T evaluated at strike K.
- As a result, we have

$$\phi(t_0; K, T) \approx \frac{1}{P(t_0, T)} \frac{C(K - \Delta K, T) - 2C(K, T) + C(K + \Delta K, T)}{(\Delta K)^2}$$

  where $A_{t_0}(K, T) = P(t_0, T)\phi(t_0; K, T)$ is the price at time t_0 of a security paying $\$1$ at state K and future time T (Arrow-Debreu price).
- Letting $\Delta K \to 0$ to be in continuous time, the density becomes

$$\phi(t_0; T, K) = \frac{1}{P(t_0, T)} \frac{\partial^2}{\partial K^2} C(K, T)$$

The straddle

- A long straddle involves going long both a call option and a put option on the same stock with strike K, and satisfies

$$S_{BS}(t,x,K,T) = C_{BS}(t,x,K,T) + P_{BS}(t,x,K,T)$$

- Given the definition of the BS-formula, we can write the price of a straddle as

$$\begin{aligned}\frac{S_{BS}(t,x,K,T)}{xRe(t,T)} &= \big(2N(d_1(T-t,F(t,T),K)) - 1\big) \\ &- Ke^{-r(T-t)}\big(2N(d_2(T-t,F(t,T),K)) - 1\big)\end{aligned}$$

- We can compute the limit case of a long straddle when the volatility tends to zero as

$$\lim_{\sigma\to 0} S_{BS}(t,x,K,T) = |xe^{-q(T-t)} - Ke^{-r(T-t)}|$$

- When the volatility tends to infinity, we get

$$\lim_{\sigma\to\infty} S_{BS}(t,x,K,T) = xe^{-q(T-t)} + Ke^{-r(T-t)}$$

The strangle

- A strangle option spread is similar to a straddle but different strikes are used. The investor use put options having a strike below the call options, and vice versa. Assuming a long strangle with strikes $K_1 < K_2$, we get

$$\widehat{S}_{BS}(t,x,K_1,K_2,T) = P_{BS}(t,x,K_1,T) + C_{BS}(t,x,K_2,T)$$

- We can compute the limit case of a long strangle when the volatility tends to zero as

$$\lim_{\sigma\to 0}\widehat{S}_{BS}(t,x,K_1,K_2,T) = \begin{cases} xRe(t,T) - K_2P(t,T) \text{ if } \frac{xRe(t,T)}{P(t,T)} > K_2 \\ K_1P(t,T) - xRe(t,T) \text{ if } \frac{xRe(t,T)}{P(t,T)} < K_1 \\ 0 \text{ if } K_1P(t,T) \leq \frac{xRe(t,T)}{P(t,T)} \leq K_2 \end{cases}$$

- When the volatility tends to infinity

$$\lim_{\sigma\to\infty}\widehat{S}_{BS}(t,x,K_1,K_2,T) = K_1e^{-r(T-t)} + xe^{-q(T-t)}$$

The call-spread option

- A bull call spread option is a vertical spread made of two calls with the same expiration but different strikes where the strike price of the short call is higher than the strike of the long call.
- We consider a bull call spread with strikes $K_1 < K_2$ given by

$$CS_{BS}(t, x, K_1, K_2, T) = C_{BS}(t, x, K_1, T) - C_{BS}(t, x, K_2, T)$$

- we can compute the limit cases of the call spread as follow

$$\lim_{\sigma \to 0} CS_{BS}(t, x, K_1, K_2, T) = \begin{cases} (K_2 - K_1)e^{-r(T-t)} \text{ if } \frac{xRe(t,T)}{P(t,T)} \geq K_1 \\ \text{ and } \frac{xP(t,T)}{Re(t,T)} \geq K_2 \\ 0 \text{ if } \frac{xRe(t,T)}{P(t,T)} < K_1 \text{ and } \frac{xRe(t,T)}{P(t,T)} < K_2 \end{cases}$$

- and

$$\lim_{\sigma \to \infty} CS_{BS}(t, x, K_1, K_2, T) = xe^{-q(T-t)} - xe^{-q(T-t)} = 0$$

Financial risks

- Hedging an option is the practice of taking offsetting positions in one market to balance against adverse movements in the value of our positions in another market.
- The factors affecting option prices are the spot price S, the strike price K, the time to maturity T, the volatility of the underlying σ, and the risk-free interest rate r.
- Thus, when risk managing a book of options one should consider the sensitivity of option prices to these factors.
- Volatility is the only unkmown in the BS formula.
- Volatility is the key parameter for traders, allowing them to compute the necessary sensitivities, or Greeks, of their portfolio.

The volatility risk

- As a proxy for accouting for stochastic volatility, traders use one BS-model for every pair (K, T), generating the IV surface.
- Traders must therefore consider the sensitivity of option prices to changes in the IV surface.
- When the volatility σ_t changes with a certain volatility of its own, the sensitivity of the option price with respect to the volatility is called Vega.
- Theoretically, the BS-formula does not apply to stochastic volatility, although the explicit BS-vega calculation is a useful approximation to the actual vega over a small time horizon.

The Greeks

- Given the spot price $(S_t^x)_{t\geq 0}$, the Black-Scholes formula is a function

$$v(t, x, r, q; T, K; \sigma)$$

- A call option with maturity T and strike K can be hedged with the quantity of the risky asset

$$\Delta(t, S_t) = \partial_x C_{BS}(t, S_t, K, T) = e^{-q(T-t)} N(d_1(T-t, F(t,T), K))$$

- Some useful Greeks are

$$\begin{aligned}
\partial_{xx} C &= \Gamma = \frac{e^{-q(T-t)}}{x\sigma\sqrt{T-t}} n(d_1) > 0 \\
\partial_\sigma C &= Vega = xe^{-q(T-t)}\sqrt{T-t}\, n(d_1) \\
\partial_t C &= \Theta = -\frac{xRe(t,T)\sigma}{2\sqrt{T-t}} n(d_1) + qxRe(t,T)N(d_1) \\
&- rKP(t,T)N(d_2)
\end{aligned}$$

Interpretation

- The dynamics of the stock price with constant volatility in discrete time is

$$\Delta S_t = \mu S_t \Delta t + \sigma S_t y \sqrt{\Delta t}$$

- We let $S^+ = S + \alpha$ where $\alpha = \epsilon S$ with $\epsilon << 1$. The Delta with a one-sided finite difference is

$$\Delta \approx \frac{C(S^+) - C(S)}{\alpha}$$

- The modified Delta is $\widehat{\Delta} = \alpha \Delta$, such that

$$C(S^+) \approx \widehat{\Delta} + C(S)$$

- Similarly, we can rewrite the Gamma in terms of the Delta as

$$\Gamma \approx \frac{\Delta(S^+) - \Delta(S)}{\alpha}$$

and deduce that, for $\widehat{\Gamma} = \alpha \Gamma$, we get

$$\Delta(S^+) \approx \widehat{\Gamma} + \Delta(S)$$

Interpretation Ctd

- We let $\sigma^+ = \sigma + \alpha$ where $\alpha = \epsilon\sigma$ with $\epsilon << 1$. Using one-sided finite difference, we can approximate the Vega as

$$Vega \approx \frac{C(\sigma^+) - C(\sigma)}{\alpha}$$

- Market practice is to set $\alpha = 1\%$ and obtain the modified Vega as $\widehat{V}ega = \alpha\, Vega$, such that

$$C(\sigma^+) \approx \widehat{V}ega + C(\sigma)$$

- To express time decay over a one day period, we simply multiply the Theta with a fraction of time, getting

$$\widehat{\Theta} = \Theta \times \Delta t$$

where $\Delta t = \frac{1}{365}$.

Interpretation Ctd

- When risk managing options, traders need to estimate their risk over a short time period.
- Assuming σ is a stochastic volatility and setting $V = V(t, S, \sigma)$, we apply a Taylor expansion on ΔV getting the variation

$$\begin{aligned}\Delta V_t &= \frac{\partial V}{\partial t}\Delta t + \frac{\partial V}{\partial S}\Delta S_t + \frac{\partial V}{\partial \sigma}\Delta\sigma \\ &+ \frac{1}{2}\big[\frac{\partial^2 V}{\partial S^2}(\Delta S_t)^2 + \frac{\partial^2 V}{\partial \sigma^2}(\Delta\sigma)^2 + 2\frac{\partial^2 V}{\partial S \partial \sigma}(\Delta S_t)(\Delta\sigma)\big] + ...\end{aligned}$$

where $\Delta V_t = V(t + \Delta t, S + \Delta S, \sigma + \Delta\sigma) - V(t, S, \sigma)$.
- Using our previous notation, we can rewrite the portfolio variation as

$$\Delta V_t = \widehat{\Theta} + \widehat{\Delta} + \widehat{V}ega + \frac{1}{2}\widehat{\Gamma}\Delta S_t + \frac{1}{2}Volga(\Delta\sigma)^2 + Vanna(\Delta S_t)(\Delta\sigma) + ...$$

The summation of independent random variables

- The uncertainty affecting the underlying is modelled by considering future trajectories of the risky asset seen as possible scenarios. Consequently, it is important to understand the properties of the market returns in order to devise their dynamics.
- J. Bernoulli **?** considered the sequence of normalised sums $\frac{S_n}{n}$, $S_n = \sum_{i=1}^{n} X_i$, where independent random variables X_i take the value 1 with probability p and the value 0 with probability $1 - p$.
- Bernoulli's therorm states that for any arbitrary small but fixed $\epsilon > 0$ we have

$$P(\left|\frac{1}{n}S_n - p\right| > \epsilon) \to 0 \text{ , } n \to \infty$$

- It is the Bernoulli's form of the law of large numbers.

The Central Limit Theory

- We consider a sequence of i.i.d. random variables $X_1, X_2, ...$ Assuming that the mathematical expectation $a = E[X_i]$ and the variance $\sigma^2 = Var(X_i)$ of these variables are finite, we construct the corresponding sequence of the normalised sums $Z_1, Z_2, ...$

$$Z_n = \frac{S_n - na}{\sigma\sqrt{n}}$$

- Then for any $x_1 < x_2$, the Central Limit Theory (CLT) states

$$P(x_1 < Z_n < x_2) \Rightarrow \int_{x_1}^{x_2} p^G(x)dx \ , \ n \to \infty$$

where

$$p^G(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

is the density of the standard normal law.

Towards the generalised CLT

- Considering the i.i.d, sequence of random variables $X_1, X_2, \ldots$ witout any preliminary assumptions about their distribution, and using the real-valued constants $a_1, a_2, \ldots$ and positive constants $b_1, b_2, \ldots$, we introduce

$$Z_n = \frac{S_n - a_n}{b_n}$$

- We want to find the constants a_n and b_n in such a way that the distribution functions of Z_n weakly converge to some limit distribution function $G(x)$, that is,

$$P(Z_n < x) \Rightarrow G(x) \ , \ n \to \infty$$

  for any x which is a continuity point of the function $G(x)$.
- We will see that the stable law is a class of distributions playing the role of the limiting law.

Power-law distributions

- In the statistical extreme value theory (EVT), the extremes and the tail regions of a sample of i.i.d. random variables converge in distribution to one of only three types of limiting laws
    1. Exponential decay : it decreases at a rate proportional to its current value. That is, $X_t = X_0 e^{-\lambda t}$ where λ is the exponential decay constant and $\tau = \frac{1}{\lambda}$ is the mean lifetime.
    2. Power-law decay : one quantity varies as a power of another. Power laws are scale-invariant, that is, for $f(x) = ax^{-\alpha}$ we get $f(cx) = c^{-\alpha} f(x) \propto f(x)$. A power law $x^{-\alpha}$ has a well-defined mean over $x \in [1, \infty]$ only if $\alpha > 2$ and it has a finite variance only if $\alpha > 3$.
    3. The behaviour of distributions with finite endpoint of their support.
- A power-law probability distribution is a distribution whose density function has the form

$$p(x) \propto L(x) x^{-\alpha}$$

where $\alpha > 1$ and $L(x)$ is a slowly varying function at infinity.

Power-law distributions Ctd

- Heavy-tailed distributions are probability distributions whose tails are not exponentially bounded. That is, they have heavier tails than the exponential distribution.
- The distribution of a random variable X with distribution function F is said to have a heavy right tail if

$$\lim_{x \to \infty} e^{\lambda x} P(X > x) = \infty \text{ for all } \lambda > 0$$

- A fat-tailed distribution is a probability distribution exhibiting large skewness or kurtosis relative to the normal distribution. Some fat-tailed distributions have power law decay in the tail of the distribution, but do not necessarily follow a power law everywhere. The distribution of a random variable X is said to have a fat tail if

$$P(X > x) \sim x^{-\alpha} \text{ as } x \to \infty \text{ , } \alpha > 0$$

  That is, if X has a probability density function $f_X(x)$, then

$$f_X(x) \sim x^{-(1+\alpha)} \text{ as } x \to \infty \text{ , } \alpha > 0$$

Self-similarity

- A stochastic process $Y = (Y(t), t \geq 0)$ is called self-similar if there is H such that for all $c > 0$ we have

$$(Y(ct), t \geq 0) = (c^H Y(t), t \geq 0) \tag{0.3}$$

where H is the exponent of self-similarity, or scaling exponent.

- A self-similar process can not be stationary, but its increments can if they are stationary. If $X_i = Y(i) - Y(i-1)$, $i = 1, 2, ...$ is the increment process of Y, then the partial sum process S_n satisfies

$$S_n = Y(n) - Y(0) = n^H (Y(1) - Y(0)) = n^H S_1$$

- Self-similar objects with parameters N and s are described by a power law such as

$$N = s^D$$

where $D = \frac{\ln(N)}{\ln(s)}$ is the dimension of the scaling law, known as the Hausdorff dimension.

The stable distributions

- Stable distributions are determined by characteristic function with the properties of being stable under addition, self-similar, and capable of producing high peaks at the mean and fat tails.
- L-stable distributions are represented by $S(\alpha, \beta, c, \delta; k)$ where $\alpha \in (0, 2]$ is the characteristic exponent, $\beta \in [-1, 1]$ is the skewness parameter, $\delta \in \mathbb{R}$ is the location parameter and $c > 0$ is the scale parameter.
- Some special case are the Gaussian distribution ($\alpha = 2$), the Levy distribution ($\alpha = \frac{1}{2}$, $\beta = 1$) and the Cauchy distribution ($\alpha = 1$, $\beta = 0$).
- The stability parameter α define the existence, or not, of the variance.

The stable distributions Ctd

- A normal random variable is unbounded, and can take arbitrarily large (absolute) values, but can not take infinite values in the sense that

$$\lim_{x \to \infty} \int_x^\infty p(x^{'}) dx^{'} = \lim_{x \to -\infty} \int_{-\infty}^x p(x^{'}) dx^{'} = 0$$

- Thus, a random variable distributed by the normal law (or any other stable) law takes finite values with probability one.
- Any non-degenerate stable distribution has a smooth (infinitely differentiable) density function, but for $\alpha < 2$ the density has an asymptotic behaviour of a heavy-tail distribution, leading to infinite variance.
- The meaning of infinite variance can be expressed as follow

$$\int_x^\infty p(x^{'}) dx^{'} + \int_{-\infty}^{-x} p(x^{'}) dx^{'} \propto x^{-\alpha} \text{ , } x \to \infty \text{ , } 0 < \alpha < 2$$

- Hence, the existence of finite variance of the normal law is connected with just a faster decrease of tails as compared with the others.

The stable distributions Ctd

- One consequence of heavy tails is that not all moments exits.
- The first two moments are not generally useful for heavy tailed distributions since the integral expressions for these expectations may diverge. Instead, we can use fractional absolute moments $E[|X|^p] = \int_{-\infty}^{\infty} |x|^p f(x)dx$, where p is any real number.

One can show that for $\alpha < 2$,

- $E[|X|^p]$ is finite for $0 < p < \alpha$, and
- $E[|X|^p] = +\infty$ for $p \geq \alpha$.

Thus,

- when $\alpha \leq 1$, then $E[|X|] = +\infty$ and the mean of X is undefined.
- when $1 < \alpha \leq 2$, then $E[|X|] < \infty$ and the mean of X is defined.
- when $0 < \alpha < 2$, then $E[|X|^2] = E[X^2] = +\infty$ and stable distributions do not have finite second moments (variances).

The stable distributions Ctd

Definition

The random variable X with values in $\mathbb{R}$ has a stable distribution if the following condition holds:
If $n \in \mathbb{N}_+$ and $X_1, ..., X_n$ is a sequence of independent variables, each with the same distribution as X, then $S_n = \sum_{i=1}^n X_i$ has the same distribution as $a_n + b_n X$ for some $a_n \in \mathbb{R}$ and $b_n \in (0, \infty)$. X is strictly stable if and anly if $a_n = 0$ for all n.

- In the general case of strictly stable random variables, we get

$$S_n = \sum_{i=1}^n X_i = n^{\frac{1}{\alpha}} X$$

- One can show that the only possible choice for the scaling constants is $b_n = n^{\frac{1}{\alpha}}$ for some $\alpha \in (0, 2]$.

Remark

In the case where the summands in the Definition do not all have the same α, the sum will not be stable.

Dependencies between variables

Definition

Let $\underline{X}$ and $\underline{Y}$ be vector observations of the variables X and Y. Then

1. *$\underline{X}$ and $\underline{Y}$ are linearly independent iff there exists no constant a such that $a\underline{X} - \underline{Y} = 0$.*
2. *$\underline{X}$ and $\underline{Y}$ are orthogonal iff $(\underline{X})^\top \underline{Y} = 0$.*
3. *$\underline{X}$ and $\underline{Y}$ are uncorrelated iff $(\underline{X} - \overline{X}I)^\top(\underline{Y} - \overline{Y}I) = 0$, where $\overline{X}$ and $\overline{Y}$ are the means of $\underline{X}$ and $\underline{Y}$, respectively, and I is a vector of ones.*

For linearly independent variables we get the following situations

1. two variables that are perpendicular can become oblique once centered: orthogonal but not uncorrelated.
2. two variables not perpendicular can become perpendicular once centered: uncorrelated but not orthogonal.
3. two variables may be both orthogonal and uncorrelated.

Dependencies between variables Ctd

- It is important to note that correlation is only a measure of linear dependence, but it does not necessarily imply something about other kinds of dependence.
- We let X be normally distributed with $E[X] = 0$ and we let $Y = X^2$ be a quadratic function of X. Thus, X and Y are dependent (quadratic dependence), but $Cov(X, Y) = E[XY] = E[X^3] = 0$. Therefore, X and Y are uncorrelated, but not independent (situation (2)).
- X does not have to be normally distributed, and any density function symmetric about 0 and for which $\int |x|^3 dP$ exists, will do.

Long range dependence (LRD)

- LRD is related to memory in a stochastic process.
- Hurst measured how a reservoir level fluctuated around its average level over time, and found that the range of the fluctuation would change, depending on the length of time used for measurement.
- If the series were random, the range would increase with the square root of time

$$R = T^{\frac{1}{2}}$$

- To standardise the measure over time, he created a dimensionless ratio by dividing the range by the standard deviation, S, of the observation, obtaining the Rescaled Range

$$\frac{R}{S} = k \times T^H$$

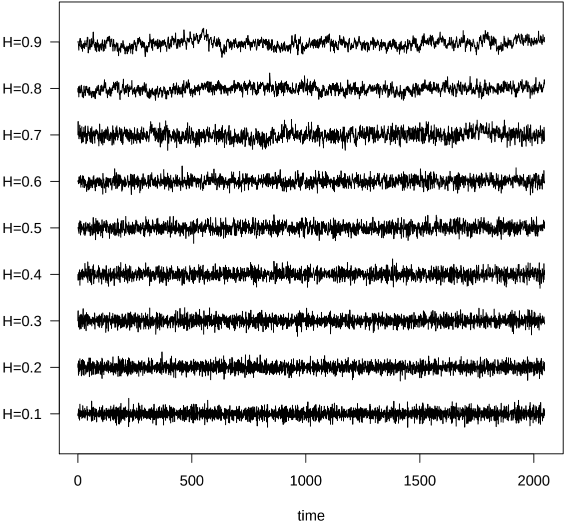
- He found that most natural phenomena follow a biased random walk, that is, a trend with noise which could be measured by how the rescaled range scales with time, or, how high the exponent H is above $\frac{1}{2}$.
- This phenomenon became known as the Hurst phenomenon.

Long range dependence Ctd

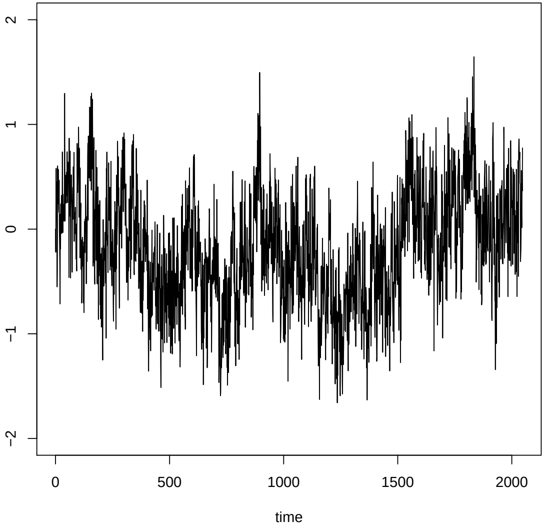
The value of the Hurst exponent varies between 0 and 1, with

- $H = \frac{1}{2}$ implying a random walk, or an independent process.
- For $0 \leq H < \frac{1}{2}$ we have anti-persistence (or ergodicity) where the process covers less distance than a random walk (mean reverting).
- For $\frac{1}{2} < H \leq 1$ we have persistence (or trend-reinforcing) where the process covers more distance than a random walk (long memory effects).

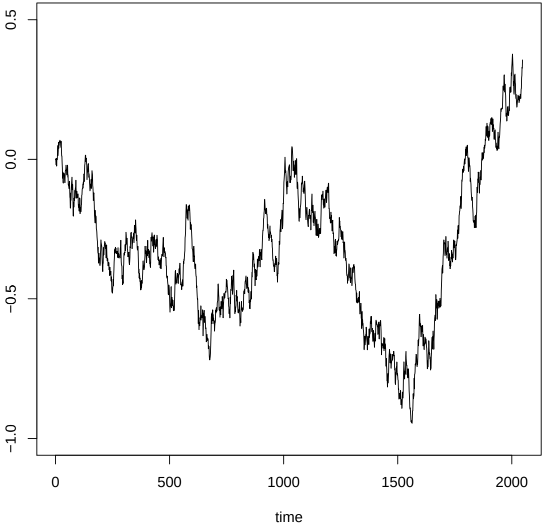
Long range dependence Ctd



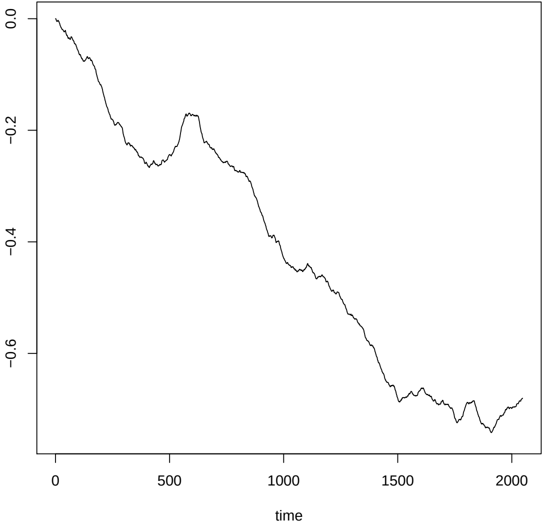
Long range dependence Ctd



Long range dependence Ctd



Long range dependence Ctd



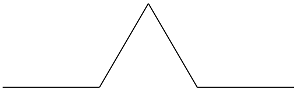
Long range dependence Ctd

- The fractal dimension D of a time series measures how jagged the time series is.
- Mandelbrot showed that the Hurst Exponent, H, is related to the fractal dimension D for a self-similar surface in n-dimensional space by the relation

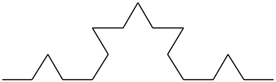
$$D = n + 1 - H$$

Koch Curve: Initial line

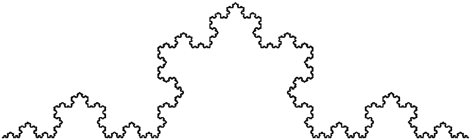
Koch Curve: Second step



Koch Curve: Third step



Koch Curve: Few iterations



Long range dependence Ctd

- In general LRD are based on 2nd-order properties of a stochastic process, since correlations are easy to understand and to estimate.

Definition (Time domain)

The weakly stationary time series $X(t)$ is said to be long-range dependent if $\sum_{k=-\infty}^{\infty} \rho_k$ diverges, where ρ_k is the autocorrelation function (ACF) of $X(t)$.

Definition (Frequency domain)

The weakly stationary time series $X(t)$ is said to be long-range dependent if its spectral density obeys $f(\lambda) \sim C_f |\lambda|^{-\beta}$ as $\lambda \to 0$, for some $C_f > 0$ and some real $\beta \in [0, 1]$.

1. in the time domain, LRD occurs at a high degree of correlation between distantly separated data points.
2. in the frequency domain, LRD occurs at a significant level of power at frequencies near zero.

Long range dependence: scaling analysis

- Since self-similarity can have very different origins, we need a general test for self-similarity.
- As a consequence of self-similarity we can set $c = \frac{1}{t}$ and get

$$X_t = t^H X_1 \text{ , } \forall t > 0$$

so that the distribution of X_t is completely determined by the distribution of X_1

$$F_t(x) = P(t^H X_1 \leq x) = F_1(\frac{x}{t^H})$$

- Differentiating the above equation, we get the density ρ_t of F_t as

$$\rho_t(x) = \frac{1}{t^H} \rho_1(\frac{x}{t^H})$$

and the kth moment is

$$E[|X_t|^k] = t^{kH} E[|X_1|^k]$$

- The scaling property of the density implies that $\widehat{\rho}_{n\Delta t}(x)$ and $\frac{1}{n^H} \widehat{\rho}_{\Delta t}(\frac{x}{n^H})$ should coincide.

The difficulties of measuring empirically LRD

- Statistical approaches are based on the moment properties of stochastic processes and must be restricted to second-order stationary processes.
- Fractal properties are only verified in the infinitesimal limit.
- Estimators of Hurst exponent works for very long, or infinite, time series
- In the time domain, at high lags only a few samples are available
- In the frequency domain, at frequencies near zero, measurement errors are largest.
- Statistical analysis applies to processes with stationary increments
- If the time series possess the two features of heavy tails and long range dependence, then many standard statistical tests will fail to work.
- Estimators are vulnerable to trends in data, periodicity, large fluctuations, etc.

Introducing multifractality

- Many time series do not exhibit a simple monofractal scaling behaviour described with a single scaling exponent.
- In complex system, such different scaling behaviour can be observed for many interwoven fractal subsets of the time series, in which case a multitude of scaling exponents is required for a full description of the scaling behaviour.
- A multifractal analysis must be applied. For instance,
  1. scaling analysis focus on global properties, such as moments and autocovariance.
  2. multifractal analysis adopt a more local viewpoint and examine the regularity of realised paths around a given instant.

Scaling analysis

- Scaling analysis investigate the multiscaling properties of the self-affine function $f(X)$ by calculating the qth order height-height correlation function. It exhibits a nontrivial multiscaling behaviour if

$$C_q(\Delta t) \sim (\Delta t)^{qH(q)}$$

  with $H(q)$ changing continuously with q.
- As a result, multiscaling in empirical data is typically identified by differences in the scaling behaviour of different (absolute) moments

$$E[|X(t, \Delta t)|^q] = c(q)(\Delta t)^{qH(q)} = c(q)(\Delta t)^{\tau(q)}$$

  where $c(q)$ and $\tau(q)$ are deterministic functions of the order of the moment q.

Multifractal analysis

- The multifractal analysis (MFA) studies how the (pointwise) local regularity of X fluctuates in time (or space).
- Local Holder regularity describes the regularity of sample paths of stochastic processes by means of a local comparison against a power law function and is therefore closely related to scaling in the limit of small scales.
- The exponent of this power law, $h(t)$, is called the (local) Holder exponent and depends on both time and the sample path of X.
- Researchers focused on a global description of the regularity of the function of f in form of multifractal spectrum (also called the singularity spectrum) $D(h)$ reflecting the size of the set of points for which the Holder exponent takes a certain value h.
- The main idea being that the relative frequency of the local exponents can be represented by a renormalised density called the multifractal spectrum.

Multifractal analysis Ctd

- Financial time series have non-stationary scaling properties (stochastic Hurst exponent) with characteristics of abrupt changes in the fractal structure that can be related to the theory of outliers.
- Outliers in time series have an inherently isolated and local character, with erratic behaviour (spikes), that can be detected and localised in time with the help of the effective local Holder exponents (ELHE).
- Even though the width of the multifractal spectra is capable of indicating the presence of large shocks, the oscillating nature of the local Holder exponent characterise the continuously changing dynamics of the response time distribution.
- Thus, the local Hurst exponent gives valuable indications on extreme values (upward and downward) in data series.

Building a theory of option pricing

- Bachelier: returns are normally distributed.
- Fama: the price changes are independent and may be a random walk (EMH).
- Samuelson: the market is efficient if it follows a martingale process.
- Market: focused on thick tails and long-memory volatility persistence, but did not account for the evidence of long memory in raw returns.
- Stylised facts on volatility: volatility clustering, asymmetry and mean reversion.
- Continuous time: jump-diffusion models. stochastic volatility + independent jumps.

Jump-diffusion models

Build a JD model differing from a BS-model only in the distribution of shocks. This is achieved by adding

- jumps to the process, where jumps means significant unexpected discontinuous changes in prices.
- stochastic volatility, meaning volatility changing at random over time.

Example

- First draw a random variable Y whose outcomes are 1 and 0 with respective probabilities p and $1-p$.
- Then, independently draw a standard normal random variable Z, $Z \sim N(0,1)$.
- Mixture model: $X = \alpha Z I_{\{Y=1\}} + \beta Z I_{\{Y=0\}}$.
- The standard deviation α and β must be chosen such that the variance of X is 1.

$$Var(X) = E[X^2] = p\alpha^2 + (1-p)\beta^2 = 1$$

Jump-diffusion models Ctd

- We can then deduce β as

$$\beta = \sqrt{\frac{1 - p\alpha^2}{1 - p}}$$

- We are left with choosing p and α to satisfy either a given kurtosis or a 0.99 critical value. The kurtosis of X is given by

$$E[X^4] = 3\big(p\alpha^4 + (1-p)\beta^4\big)$$

The regimes of volatility

Rules of thumb: provides an indication of some possible behaviour for the smile.

1. sticky delta : the strike of the option is rescaled according to how the current spot evolved with respect to the spot at inception
   $\Sigma(t, S_t; K, T) = \Sigma_0^{obs}(K\frac{S_0}{S_t}, T - t)$
2. absolute floating : the future implied volatility is obtained from the original smile surface by simply reducing time to maturity and linearly offsetting the strike by how much the spot has moved
   $\Sigma(t, S_t; K, T) = \Sigma_0^{obs}(K + (S_0 - S_t), T - t)$
3. absolute sticky (sticky strike) : any dependence on the current spot level or calendar time is ignored
   $\Sigma(t, S_t, K, T) = \Sigma_0^{obs}(K, T - t)$

The regimes of volatility Ctd

Example: the linear smile (only true when the strike is near the money)

$$\Sigma(t_0, S_0; K, T) = a - b(K - S_0) \ , \ b > 0$$

1. sticky moneyness : The dynamics of option prices derive entirely from the moneyness

$$\Sigma(t, S_t; K, T) = a - b(\frac{K}{S_t} - 1)S_0 \ , \ b > 0$$

2. sticky implied tree : In the implied tree, local volatility increases as the underlying decreases. We get the linear approximation to the skew

$$\Sigma(t, S_t; K, T) = a - b(K + S_t - 2S_0) \ , \ b > 0$$

so that the implied volatilities decrease as K or S increases.

Space homogeneity versus space inhomogeneity

At the two extremes

1. The smile is expressed in terms of moneyness (sticky delta rule). We get space homogeneous models like the Heston model. The delta is greater than the BS-delta, meaning that the dynamics of the IV follow that of the underlying asset.
2. The smile is expressed in terms of absolute value for the strike and spot prices (sticky implied tree rule). We get space inhomogeneous models like some LV models. The delta is smaller than the BS-delta, meaning that the dynamics of the IVS is inversly related to the underlying asset.

- In between, practitioners developed a range of universal models: the diffusion term is a mix of stochastic volatility and local volatility.

The implied Greeks

- Different option pricing models imply different smile dynamics leading to different Greeks.
- Assuming that the volatility $\Sigma = \Sigma(S_t)$ depends on the spot price, the delta becomes

$$\Delta(t, S_t) = \partial_x C_{BS}(t, S_t, K, T; \Sigma) + \partial_\Sigma C_{BS}(t, S_t, K, T; \Sigma) \frac{\partial \Sigma(t, S_t, K, T)}{\partial x}$$

Remark

The new delta will be either bigger or smaller than the BS delta, depending on the shape of the IV with repect to the spot price. If $\frac{\partial \Sigma}{\partial x} > 0$, the new delta will be bigger, but If $\frac{\partial \Sigma}{\partial x} < 0$, the new delta will be smaller.

The implied Greeks Ctd

- Differentiating the delta one more time with respect to the spot price, we get the gamma expressed in terms of the shadow gamma as follow

$$
\begin{aligned}
\Gamma(t,S_t) &= \partial_{xx} C_{BS}(t,S_t,K,T) + 2\partial_{x\Sigma} C_{BS}(t,S_t,K,T;\Sigma)\frac{\partial \Sigma}{\partial x} \\
+ &\quad \partial_{\Sigma\Sigma} C_{BS}(t,S_t,K,T;\Sigma)\big(\frac{\partial \Sigma}{\partial x}\big)^2 + \partial_{\Sigma} C_{BS}(t,S_t,K,T;\Sigma)\frac{\partial^2 \Sigma}{\partial x^2}
\end{aligned}
$$

- The gamma depends on the second derivative of the implied volatility with respect to the spot price.

Characterising the implied Greeks

- We need a model to describe how the IVS is modified when the underlying asset moves.
- Given the regimes of volatility we can relate these variations to the Skew and Curvature of the surface.
- The first derivative becomes

$$\frac{\partial \Sigma(t,S_t;K,T)}{\partial S_t} = k_S \frac{\partial \Sigma(t,S_t;K,T)}{\partial K}$$

$$k_S = \begin{cases} 0 \text{ if sticky strike} \\ 1 \text{ if sticky implied tree} \\ -\frac{K}{S_t} \text{ if sticky delta} \\ \frac{K}{S_t} \text{ if minimum variance} \end{cases}$$

- The second derivative becomes

$$\frac{\partial^2 \Sigma(t, S_t; K, T)}{\partial S_t^2} \approx k_C \frac{\partial^2 \Sigma(t, S_t; K, T)}{\partial K^2}$$

where

$$k_C = \begin{cases} 0 \text{ if sticky strike} \\ 1 \text{ if sticky implied tree} \\ (\frac{K}{S_t})^2 \text{ if sticky delta and minimum variance} \end{cases}$$

Dynamic delta hedging

- We let $v_\gamma(t, S_t)$ be the market price with misspecified volatility γ function of the current stock price. $P(t)$ for $t \in [0, T]$ is the option price and σ is the true volatility.
- We short the option P and hedge it by being long a quantity Δ of the stock and some cash. The Tracking Error satisfies the process

$$e(t) = \pi_\Delta(t) - P(t)$$

- After some calculation, the Tracking Error becomes

$$e_\gamma(t) = \frac{1}{2} M(t) \int_0^t \frac{1}{M(u)} [\gamma^2(u, S(u)) - \sigma^2(u)] S^2(u) \frac{\partial^2 v_\gamma}{\partial x^2}(u, S(u)) du$$

where $M(t)$ is the money market.

Dynamic delta hedging Ctd

- The expected PnL is maximum when the stock price is close to the money at maturity, where Gamma is largest, and when the instantaneous volatility is large at maturity.
- However, the PnL depends on the path taken by the underlying asset.
- Two trajectories with the same average volatility ending at-the-money can lead to two very different PnL.
- In practice the difficulty comes from the realisation (or not) of the delta hedging. The volatility is realised through delta hedging. If the underlying jumps the volatility can not be realised.

Dynamic delta hedging Ctd

Theorem

Given some assumptions, if

$$\sigma(t) \leq \gamma(t, S(t))$$

for Lebesgue-almost all $t \in [0, T]$*, almost surely, then* $(P_\gamma, \Delta_\gamma)$ *is a superstrategy and*

$$\pi_{\Delta_\gamma}(T) \geq h(S(T))$$

Inversly, if $\sigma(t) \geq \gamma(t, S(t))$*, then* $(P_\gamma, \Delta_\gamma)$ *is a substrategy.*

- Successful hedging is possible even under significant model error.
- It depends on the relationship between the misspecified volatility $\gamma(t, S(t))$ and the true volatility $\sigma(t)$.

Discrete dividends

- Rewrite the spot price with discrete dividends in terms of a process Y with no discrete dividends as follow

$$S_t = a(t)\,Y_t + b(t)$$

- The price of a call option on S is

$$C_S(t;K,T) = P(t,T)a(t,T)E_t[(Y_T - k)^+] = P(t,T)a(t,T)C_Y(t;k,T)$$

where where $k = \frac{K-b(t,T)}{a(t,T)}$.

- The capitalisation factor from time t until time T is

$$C(t,T) = \frac{Re(t,T)}{P(t,T)} = e^{\int_t^T \mu_s\,ds}$$

- We let D_t be the dividends paid till time t given by

$$D_t = \sum_{i=0}^{\infty} \mathcal{H}(t - t_{d_i})\,d_i\,C(t_{d_i},t)$$

Discrete dividends Ctd

- The dividends paid from time t till time T and capitalised at maturity T are given by

$$D(t,T) = D_T - C(t,T)D_t$$

- The stock price dynamics in the spot model and under the risk-neutral measure $\mathbb{Q}$ become

$$\begin{aligned} S_t &= C(0,t)Y_t - D_t \\ dY_t &= Y_t \sigma_{Y,t} dW_t \\ Y_0 &= S_0 \end{aligned}$$

with $a(t) = C(0,t)$, $b(t) = -D_t$.

Discrete dividends Ctd

- We let $Z_t = C(0,t)\,Y_t$ be the asset price S_t plus the forward value of all dividends paid from inception up to time t. The dynamics are

$$dZ_t = \mu_t Z_t + \sigma_{Z,t} Z_t dW_t$$

- It is popular because by adding the already paid dividends to the strike of the option one can get a closed-form solution for European option prices

$$C_S(K,T) = P(t,T)E_t^Q[(S_T - K)^+] = P(t,T)E_t^Q[(Z_T - K^{'})^+]$$

where $K^{'} = K + D(t,T)$ such that $a(t,T) = 1$ and $b(t,T) = -D(t,T)$.

Discrete dividends Ctd

- In the spot model the forward price is given by

$$F(t,T) = S_t C(t,T) - D(t,T) = S_t \frac{Re(t,T)}{P(t,T)} - \sum_{t_i \in [t,T]} d_i \frac{Re(t_i,T)}{P(t_i,T)}$$

- The put-call parity only holds for European options and not for American options. It is expressed as

$$\overline{C}(t, S_t, K, T) - \overline{P}(t, S_t, K, T) = S_t \frac{Re(t,T)}{P(t,T)} - D(t,T) - K$$

Inferring the implied yield from market prices

- We "mark-to-market" the repo curve by comparing the synthetic forward price implied by the put-call parity relationship with the theoretical forward.
- Given deterministic rates, the forward becomes

$$F(t,T_N) = F(t,T_{N-1})\frac{Re(T_{N-1},T_N)}{P(T_{N-1},T_N)} - \sum_{t_{d_i}\in[T_{N-1},T_N]} d_i \frac{Re(t_{d_i},T_N)}{P(t_{d_i},T_N)}$$

- The synthetic forward price satisfies

$$F^s(t,T) = \overline{C}(t,S_t,K,T) - \overline{P}(t,S_t,K,T) + K$$

- We assume piecwise constant repo rate

$$q(T_i) = \frac{1}{T_i}\sum_{j=1}^{i}(T_j - T_{j-1})q_{T_{j-1},T_j}$$

and find $q_i = q(T_i)$ such that

$$G(t,T_i,q_i) = F^s(t,T_i) - F(t,T_i,q_i) = 0 \quad \text{for} \quad i = 1,...,N$$

The parametric models

- Estimate each smile independently with some nonlinear function.
- Then, the IV surface is reconstructed by interpolating total variances along the forward moneyness.

Examples

- The polynomial models: Cubic model

$$\Sigma(t, S_t; K, T) = a_0 + a_1\overline{\eta} + a_2(\overline{\eta})^2 + a_3(\overline{\eta})^3 + a_4\tau + a_5\tau^2$$

  where $\overline{\eta} = \ln \frac{K}{S}$.
- The stochastic volatility models: asymptotic expansion of a stochastic volatility model (SABR, SVI).

The difficulty of generating arbitrage-free smiles

- Roper investigated three models: SVI, SABR, Quadratic parametrisation, and showed that they still admit arbitrage under certain parameter classifications.
- Setting $x = \ln\left(\frac{K}{F(t,T)}\right)$ and given the time scaled IV, Ξ, the parametrisation of the SVI follows

$$\Xi^2_{SVI}(x,\tau) = a + b\Big(\rho(x-m) + \sqrt{(x-m)^2+\sigma^2}\Big)$$

with parameters $a = 0.04$, $b = 0.8$, $\sigma = 0.1$, $\rho = -0.4$ and $m = 0$.

- The parametrisation of the SABR model is

$$\Xi_{SABR}(x,1) = \frac{k|x|}{\ln\left(k|f(x)| + \sqrt{1+k^2f^2(x)}\right)}$$

where

$$f(x) = \frac{1-e^{-\beta x}}{\sigma_0\beta}$$

with parameters $\sigma_0 = 0.2$, $\beta = -4.0$ and $k = 0.5$.

The difficulty of generating arbitrage-free smiles Ctd

- The Quadratic parametrisation is given by

$$\Xi_Q(x,1) = 0.16 - 0.34x + 4.45x^2$$

- In all cases, the Durrleman Condition is violated and the resulting parametrisation is not arbitrage free.
- The solution is to define a globally arbitrage-free parametric model in time and space.
- It requires complex calibration algorithms and the use of sophisticated optimisation engine to avoid local minima.

The parametric MixVol model

We consider a sum of shifted log-normal distributions, that is, using the Black-Scholes formula with shifted strike as an interpolation function.

- The market price $C_M(K,t)$ is estimated at time $t_0 = 0$ by the weighted sum

$$C_M(t_0, S_0; K, t) = \sum_{i=1}^{n} a_i(t) Call_{BS}(t_0, S_0, R_t, P_t, \overline{K}(K,t), t, \Sigma_i(t))$$

  where $a_i(t)$ for $i = 1, .., n$ are the weights, $\overline{K}(K,t) = K^{'}(K,t)(1 + \mu_i(t))$ with $K^{'}(K,t) = K + D_t$.
- The time function $t \to \Sigma_i(t)$ satisfies

$$\Sigma_i^2(t) = \frac{1}{t} \int_0^t \sigma^2(s) ds$$

  where $\sigma(t)$ is the instantaneous volatility of the underlying process.

The parametric MixVol model Ctd

- Given the function

$$f(t,x) = 1 - \frac{2}{1 + (1 + \frac{t}{x})^2}$$

we make the weight $a_i(t)$ proportional to $\frac{a_i^0}{f(t,\beta_i)}$ for some constant $a_i^0 > 0$, getting the representation

$$\mu_i(t) = \mu_i^0 f(t,\beta_i) \text{ and } a_i(t) = \frac{a_i^0}{f(t,\beta_i) \times norm}$$

where $norm = \sum_{i=1}^n \frac{a_i^0}{f(t,\beta_i)}$.

The parametric MixVol model Ctd

With seperable functions of time, we can prove that the no-free lunch constraints simplify.

Theorem

No-Arbitrage Constraints in the MixVol Model
We let the price of a European call option be given above and assume seperable functions of time $a_i(t)$, $\mu_i(t)$. Then the resulting implied volatility surface is free from static arbitrage provided the following parameter restrictions hold:

$$\begin{aligned} a_i^0 &\geq 0 \\ \sum_{i=1}^{n} a_i(t) &= 1 \\ \sum_{i=1}^{n} a_i^0 \mu_i^0 &= 0 \\ \mu_i^0 &\geq -1 \end{aligned} \tag{0.4}$$

The parametric MixVol model: the term structure

- The no-arbitrage condition with respect to time holds iff the total variance $\omega(K,t) = \Sigma^2_{imp}(K,t)t$ is an increasing function of time t

$$\partial_t \omega_i(t) \geq 0 \text{ , } \forall i$$

or equivalently

$$\Sigma_i(t) \geq -2t\partial_t \Sigma_i(t) \text{ , } \forall i$$

- To make sure that the IVS flattens for infinitely large expiries, we impose the limit behaviour

$$\lim_{t \to \infty} \Sigma_i(t) = d_i$$

The parametric MixVol model: the term structure Ctd

- To get a general volatility function capable of generating both an upward hump or a downward one, we propose the volatility function

$$\Sigma_i(t) = \big(a_i + b_i \ln(1 + e_i t)\big) e^{-c_i t} + d_i$$

where $c_i > 0$, $d_i > 0$ and $a_i \in \mathbb{R}$, $b_i \in \mathbb{R}$ and $e_i \in] -\frac{1}{t}, \infty[$.

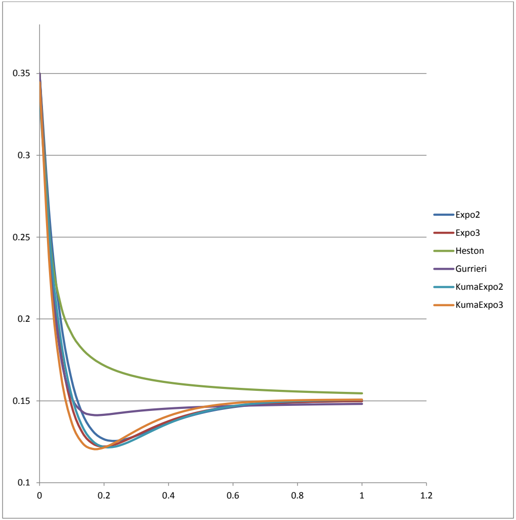
- The derivative of the function with respect to time t is

$$\Sigma_i^{'}(t) = \Big(- a_i c_i + b_i \frac{e_i}{1 + e_i t} - b_i c_i \ln(1 + e_i t)\Big) e^{-c_i t}$$

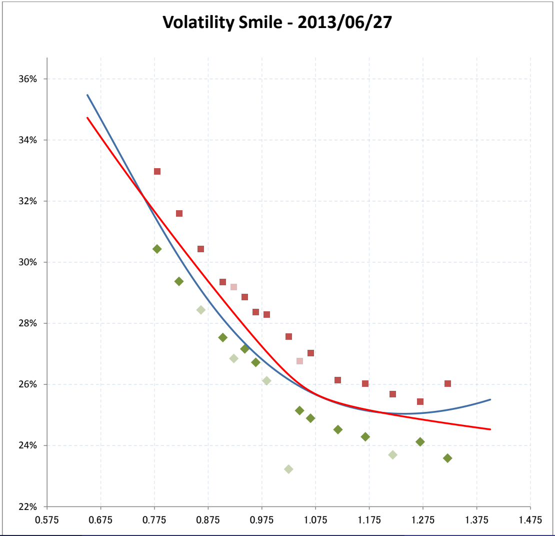
and the no-arbitrage constraint must satisfy

$$e^{-c_i t}\Big[(1 - 2tc_i)\big(a_i + b_i \ln(1 + e_i t)\big) + 2tb_i \frac{e_i}{1 + e_i t}\Big] + d_i \geq 0$$

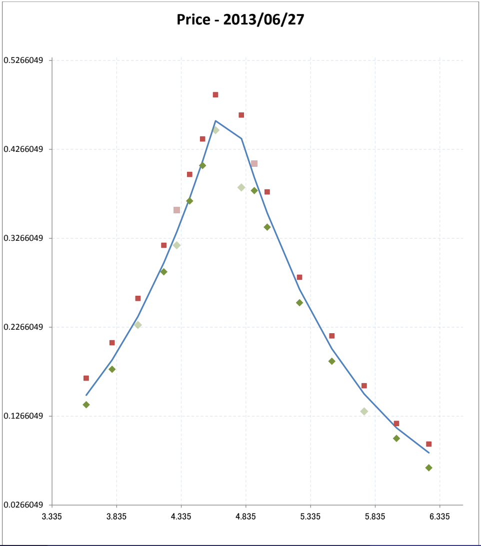
Example of term structure



Example of smile: volatility



Example of smile: price



Digital contracts in the MixVol model

- We consider the modified strike $K^{'}(K,t) = K + D_t$ with event $\xi = \{Z_t > K^{'}\}$ when pricing a Digital contracts.
- We set the interest rate to zero and multiply the strike with the discount factor, getting $\tilde{K}(K,t) = P(t_0,t)\overline{K}(K,t)$.
- Let $D(S,t_0,t;\xi_i)$ be the BS digital option, the Digital Bond satisfies

$$D_M(S,t_0,t;\xi) = \frac{1}{norm}\sum_{i=1}^{n}\overline{a}_i(t)(1+\mu_i(t))D(S,t_0,t;\xi_i)$$

- The Digital Share is

$$S_M(S,t_0,t;\xi) = xRe(t_0,t)\frac{1}{norm}\sum_{i=1}^{n}\overline{a}_i(t)N(d_1^i(t-t_0,x,\tilde{K}(K,t))$$

Computing the Skew and Curvature analytically

- The model Digital Bond must equate the market digital price. Thus, the Skew satisfies

$$Skew(K,t) = \frac{1}{Vega(K,t;\Sigma_{BS}(K,t))}\left[D_{BS}(S,t_0,t;\xi) - D_M(S,t_0,t;\xi)\right]$$

- The parametric density must equate the market density. Thus, the Curvature satisfies

$$\begin{aligned}\partial_{KK}\Sigma(K,t) &= \frac{\partial_{KK}C_M(t_0,S_0;K,t)}{Vega(K,t;\Sigma_{BS}(K,t))} \\ - & \frac{1}{K^2\Sigma(K,t)\tau}\left[1 + 2Kd_1\sqrt{\tau}Skew(K,t) + K^2 d_1 d_2 \tau (Skew(K,t))^2\right]\end{aligned}$$

where $\tau = (t - t_0)$.

Controlling the Greeks

- The implied Greeks are functions of the first derivative of the IV with respect to the spot price $\partial_x \Sigma$, and the second derivative of the IV with respect to the spot price $\partial_{xx} \Sigma$.
- These variations can be related to the Skew and Curvature of the IVS.
- After calibration, we infer the BS volatility $\Sigma_{BS}(K, T)$ for the pair (K, T) from the model price $C_M(K, T)$, and we compute the implied delta at (t, S_t) as

$$\partial_x C_M(K, T) = \partial_x C_{BS}(K, T; \Sigma) + \partial_\Sigma C_{BS}(K, T; \Sigma) k_S \frac{\partial \Sigma}{\partial K}$$

Controlling the Greeks Ctd

- The implied gamma at (t, S_t) is

$$\begin{aligned}\partial_{xx} C_M(K,T) &= \partial_{xx} C_{BS}(K,T;\Sigma) + 2\partial_{x\Sigma} C_{BS}(K,T;\Sigma) k_S \frac{\partial \Sigma}{\partial K} \\ &+ \partial_{\Sigma\Sigma} C_{BS}(K,T;\Sigma)\big(k_S \frac{\partial \Sigma}{\partial K}\big)^2 + \partial_{\Sigma} C_{BS}(K,T;\Sigma) k_C \frac{\partial^2 \Sigma}{\partial K^2}\end{aligned}$$

where k_S and k_C satisfy

$$k_S = \begin{cases} 0 \text{ if sticky strike} \\ 1 \text{ if sticky implied tree} \\ -\frac{K}{S_t} \text{ if sticky delta} \\ \frac{K}{S_t} \text{ if min var (MV)} \end{cases} \quad k_C = \begin{cases} 0 \text{ if sticky strike} \\ 1 \text{ if sticky implied tree} \\ (\frac{K}{S_t})^2 \text{ if sticky delta and MV} \end{cases}$$

Scenarios analysis

- We consider a range of plausible scenarios for the dynamics of the implied volatility surface.
- We bump each function's curve $f_i(t)$ inside the shift function $\mu_i(t)$ with a constant, getting for (t_0, S_0)

$$C_M(K,t;\epsilon) = \frac{1}{norm} \sum_{i=1}^{n} \overline{a}_i(t) Call_{BS}(K^{'}(K,t)(1 + \mu_i^0(f_i(t) + \epsilon)), t, \Sigma_i(t))$$

- Multiplying each parameter μ_i^0 with the same constant $(1 + \epsilon)$, rather than adding a constant, will preserve the no-arbitrage constraints. The shifted option price becomes

$$C_M(K,t;\epsilon) = \frac{1}{norm} \sum_{i=1}^{n} \overline{a}_i(t) Call_{BS}(K^{'}(K,t)(1 + (1 + \epsilon)\mu_i^0 f_i(t)), t, \Sigma_i(t))$$

Scenarios analysis Ctd

- We can shift each BS function $\Sigma_i(t)$ with a positive constant to generate a parallel shift type of movements of the IVS around the money.

$$C_M(K,t;\epsilon) = \frac{1}{norm} \sum_{i=1}^{n} \overline{a}_i(t) Call_{BS}(\overline{K}(K,t), t, \Sigma_i(t) + \epsilon)$$

where $\widehat{\Sigma}_i(t) = \Sigma_i(t) + \epsilon$.

- Combining all these bumps together, we obtain a no-arbitrage deformation of the IVS generated by the shifted option price

$$\begin{aligned} & C_M(K,t;\epsilon_\mu,\epsilon_\beta,\epsilon_\sigma) \\ = \quad & \frac{1}{norm} \sum_{i=1}^{n} \overline{a}_i(t) Call_{BS}(K^{'}(1 + (1+\epsilon_\mu)\mu_i^0 f(t,\beta_i(1+\epsilon_\beta))), t, \widehat{\Sigma}_i(t)) \end{aligned}$$

where $\widehat{\Sigma}_i(t) = \Sigma_i(t) + \epsilon_\sigma$.

Scenarios analysis Ctd

- For risk management purposes, we can therefore compute the modified vega as

$$\widehat{V}ega = C_M(K,t;\epsilon_\mu,\epsilon_\beta,\epsilon_\sigma) - C_M(K,t)$$

The variance swap

- The variance swap is a tool for trading and managing volatility risk.
- It gives investors exposure to variance without directly exposing themselves to movements in the underlying since it does not require delta hedging.
- The payoff at maturity of the variance swap is the difference between the realised variance of the underlying stock price and a fixed strike multiplied by a notional.
- The price of the variance swap is a forward contract on the realised annualised variance of the stock price.

The variance swap Ctd

- Given the period $[0, T]$ with business days $0 = t_0 < ... < T_n = T$, its payoff at time T is

$$U\Big(A \times \Big[\frac{1}{n}\sum_{i=1}^{n}(\log\frac{S_i}{S_{i-1}})^2 - (\frac{1}{n}\log\frac{S_n}{S_0})^2\Big] - K_{var}\Big)$$

where $(S_i)_{i=1,..,n}$ is the stock price at time t_i, $i = 1, .., n$, U is the notional amount of the swap, A the annualisation factor and K_{var} is the strike price.

The variance swap Ctd

- From the definition of the quadratic variation of a jump-diffusion process, we can deduce that when the number of sampling dates n tends to infinity the price of the variance swap converge to the quadratic variation of the logarithm of the stock price denoted by L_t.

Proposition

If we take the limit and let n tends to infinity, then the realised annualised variance $\overline{V}_n$ is approximated by the quadratic variation of the stock price, that is,

$$\overline{V}_T = \frac{1}{T}[L, L]_T = \frac{1}{T}\Big(\int_0^T \sigma(\omega, s)^2 ds + \int_0^T d\widehat{J}_L(s)\Big)$$

where $\widehat{J}_L(t) = \sum_{s \in [0,t]}^{\Delta L_s \neq 0} |\Delta L_s|^2 = \sum_j^{N_t} |Z_L(j)|^2$ *is a pure jump process with* $E[\widehat{J}_L(t)] = Var[J_L(t)] = \int_{[0,t]} \lambda \int_{\mathbb{R}} z^2 F_L(dz) ds$.

The variance swap Ctd

- From the absence of arbitrage opportunities, variance swap prices are martingales under the risk-neutral measure.

Definition

We let the price of the variance swap contract $(V_A(t,T))_{t\geq 0}$*, seen at time* t *with maturity* T*, satisfies under the risk-neutral measure* $\mathbb{Q}$

$$V_A(t,T) = E^Q[e^{-\int_t^T r_s ds}\Big(\int_0^T \sigma^2(\omega,s)ds + \int_0^T d\widehat{J}_L(s)\Big)|\mathcal{F}_t]$$

where the payoff of the contingent claim is the quadratic variation of a Lévy process.

The logarithm contract

- In a pure diffusion process we can link the price of the variance swap to the price of the logarithm contract.
- We denote $\overline{W}_T$ the total variance over the period $[0, T]$ and get

$$\overline{W}_T = \int_0^T |\sigma(t,\omega)|^2 dt = 2\Big(\int_0^T \frac{dS_t}{S_t} - \log \frac{S_T}{F_T(0)} + \log \frac{S_{t_0}}{F_T(0)}\Big)$$

- Taking the expectation of the total variance we get

$$\frac{E_{t_0}[\int_0^T |\sigma(t,\omega)|^2 dt]}{2} = \int_0^{F_T(0)} \frac{1}{K^2}\overline{P}(T,K)dK + \int_{F_T(0)}^{\infty} \frac{1}{K^2}\overline{C}(T,K)dK$$

The multi-underlying processes

- We let $(X_t)_{t\geq 0}$ be a Markov process taking values in $\mathbb{R}$ to be the logarithm of the underlying price of a contingent claim with dynamics under the risk-neutral measure $\mathbb{Q}$ being

$$dX_t = \mu_t dt + \sigma dW(t)$$

- For any two processes $X_{1,t}$ and $X_{2,t}$, we get $< dW_1, dW_2 >_t = \rho_{1,2} dt$ where $\rho_{1,2}$ is the instantaneous correlation between the two assets.
- Assuming that the correlation matrix is symmetric and positive definite, we use Cholesky decomposition to project the correlated Brownian motions W into a basis of independent orthogonal Brownian motions B.
- The dynamics of the one-dimensional Markov process become

$$dX_t = \mu_t dt + \sigma_t dB(t) = \mu_t dt + < \sigma_t, dB(t) >$$

where $\mu_t = r_t - \frac{1}{2}\sigma_t \sigma_t^\top - q_t$. The Brownian motion $B(t) = \big(B_1(t), ..., B_n(t)\big)^\top$ is a column vector of dimension $(n, 1)$ and σ_t is a matrix of dimension $(1, n)$ with element $\sigma_i(t)$.

The multi-underlying processes Ctd

- We can re-express the above equation in terms of the volatility $vol_S(t)$ and a unidimensional Brownian motion $Z(t)$ as

$$< \sigma_t, dB(t) > \; = vol_S(t) dZ(t)$$

- The link between the vector of local volatilities and the underlying volatility is

$$V(t) = vol_S^2(t) = \sum_{j=1}^{n} \sigma_j^2(t) = ||\sigma_t||^2 = \sigma_t \sigma_t^\top$$

- The average variance in the range $[t, T]$ is given by

$$I^2(t, T) = \frac{1}{T-t} \int_t^T V(s) ds$$

The multi-underlying processes Ctd

- The stochastic differential equation (SDE) of the process becomes

$$dX_t = \mu_t dt + \sqrt{V(t)} dZ(t)$$

with

$$dZ(t) = \frac{\sum_{j=1}^{n} \sigma_j dB_j(t)}{\sqrt{\sum_{j=1}^{n} \sigma_j^2(t)}}$$

- The instantaneous covariance between different underlyings $X_{1,t}$ and $X_{2,t}$ is

$$Cov(dX_{1,t}, dX_{2,t}) = < \sigma_t^1, \sigma_t^2 > dt = \sigma_t^1 (\sigma_t^2)^\top dt$$

since $X_{i,t} = \log S_{i,t}$ for $i = 1, .., n$.

- In the case where we have not yet decorrelated the Brownian motions, the covariance between two stocks $S_{1,t}$ and $S_{2,t}$ is

$$Cov(\frac{dS_{1,t}}{S_{1,t}}, \frac{dS_{2,t}}{S_{2,t}}) = < \sigma_t^1, \sigma_t^2 > dt = \rho_{12}(t) vol_{S_1}(t) vol_{S_2}(t) dt$$

The basket in the BS-world

- In a multidimensional BS model with N stocks $S_1, .., S_N$, the risk-neutral dynamics of the forward prices are given by

$$\frac{dF_i(t,T)}{F_i(t,T)} = \sum_{j=1}^{n} \sigma_{ij}(t) dB_j(t) \ , \ i = 1, ..., N$$

  where B_j for $j = 1, .., n$ are independent standard Brownian motions.
- Given the forward basket price $B(t,T) = \sum_{i=1}^{N} \omega_i F_i(t,T)$, with weights $\omega_i \geq 0$, its dynamics satisfy the SDE

$$\frac{dB(t,T)}{B(t,T)} = \sum_{i=1}^{N} \widehat{\omega}_i(t) \frac{dF_i(t,T)}{F_i(t,T)}$$

  where

$$\widehat{\omega}_i(t) = \frac{\omega_i F_i(t,T)}{\sum_{i=1}^{N} \omega_i F_i(t,T)} = \frac{\omega_i F_i(t,T)}{\omega^\top F(t,T)} \ , \ i = 1, ..., N$$

  with $0 \leq \widehat{\omega}_i(t) \leq 1$ and $\sum_{i=1}^{N} \widehat{\omega}_i(t) = 1$.

The basket in the BS-world Ctd

- We can compute the instantaneous variance of the basket as

$$V_B(t) = Var(\frac{dB(t,T)}{B(t,T)}) = Var(\sum_{i=1}^{N} \widehat{\omega}_i(t) \frac{dF_i(t,T)}{F_i(t,T)})$$

- From independence of the Brownian motions, it simplifies to

$$V_B(t) = \sum_{i=1}^{N} Var(\widehat{\omega}_i(t) \frac{dF_i(t,T)}{F_i(t,T)}) = \sum_{i=1}^{N} ||\widehat{\omega}_i(t) \sigma_t^i||^2$$

- In the case where we have not projected the Brownian motions W in a basis of independent Brownian motions B, we get

$$V_B(t) = \sum_{i,j=1}^{N} \widehat{\omega}_i(t) \widehat{\omega}_j(t) Cov(\frac{dS_i(t)}{S_i(t)}, \frac{dS_j(t)}{S_j(t)})$$

and the instantaneous variance of the basket becomes

$$V_B(t) = \sum_{i,j=1}^{N} \widehat{\omega}_i(t) \widehat{\omega}_j(t) \rho_{ij}(t) \sigma_i(t) \sigma_j(t) dt$$

Pricing basket option

- The price of the basket option at time $t = 0$ for strike K and maturity T, under the risk-neutral measure $\mathbb{Q}$, satisfies

$$\begin{aligned} BC(N, T, K) &= P(0, T) E^Q[(B(T) - K)^+] \\ &= P(0, T) E^Q[\big(\sum_{i=1}^N \omega_i S_i(T) - K\big)^+] \end{aligned}$$

- By sublinearity of the maximum, we get the following relation on the intrinsic value

$$\max\left(B(T) - K, 0\right) \leq \sum_{i=1}^N \omega_i \max\left(S_i(T) - K, 0\right)$$

- By the dominance principle we see that a basket option is cheaper than the corresponding portfolio of plain vanilla options.

Pricing basket option Ctd

- If we assume that the i-th stock price follows the dynamics of an Affine model, and apply Ito's Lemma to the forward basket price, then its dynamics are no-longer affine.
- There is no explicit analytical expression available for the distribution of the weighted sum of the assets.
- Thus, using the BS formula on a collection of underlying stocks will not produce a closed-form solution to the price of a basket option.
- Determining the price of a basket option is not a trivial problem. Upper and lower bounds can be found.

The problem of portfolio selection

- A direct consequence of the EMH is that the most important concepts in theoretical and empirical finance developed around the assumption that asset returns follow a normal distribution.
- In the problem of portfolio selection, the mean-variance approach is a simple trade-off between return and uncertainty.
- We assume zero transaction costs and portfolios with prices V_t taking values in $\mathbb{R}$ with dynamics given by

$$\frac{dV_t}{V_t} = \mu dt + \sigma_V dW_t$$

- Markowitz postulated that an investor should maximise the expected portfolio return μ, while minimising portfolio variance of return σ_V^2.
- It follows from the relation between the variance of the return of the portfolio σ_V^2 and the variance of return of its constituent securities σ_j^2 for $j = 1, 2, ..., N$ given by

$$\sigma_V^2 = \sum_j w_j^2 \sigma_j^2 + \sum_j \sum_{k \neq j} w_j w_k \rho_{jk} \sigma_j \sigma_k$$

The problem of portfolio selection Ctd

- One consequence of the CAPM is that statistical approaches must be restricted to second-oder stationary processes.
- Using volatilities and correlations calculated on any desired historical time interval for price data and different forecast times, we can then compute the risk of the portfolio.
- Rather than using volatilities and correlations statistically forecasted, it has been suggested to use implied volatility from option prices because they are biased estimators of the future volatility.
- Defining the Implied Correlation, a new type of correlation products developped, called Dispersion Trading.

Dispersion trading: the implied correlation

- Given an $\mathbb{R}$ random variable X, a portfolio is defined as a weighted sum of random variables. Markowitz's portfolio theory is derived from the variance of that sum.

Remark

The portfolio variance σ_I^2 is not to be confused with the instantaneous basket variance $V_B(t)$. The former is derived from a probability formula and has constant weights, while the latter is derived from stochastic calculs within infinitesimal time and has stochastic weights.

Remark

From the moment we are plugging implied volatility to both the LHS and RHS of the portfolio variance, we are trying to relate an option on a basket with a sum of options on its component. We enter the option pricing theory (OPT), where the no-dominance principle apply, and this relationship is no-longer valid.

Dispersion trading: the implied correlation Ctd

- Assuming σ_i and σ_j for the ith and jth component of an index I are implied volatilities, and letting $\rho = \rho_{ij}$ for $i \neq j$, practitioners defined the Implied Correlation (IC) as

$$\rho_{imp} = \frac{\sigma_I^2 - \sum_{i=1}^N w_i^2 \sigma_i^2}{2 \sum_{i=1}^N \sum_{j>i} w_i w_j \sigma_i \sigma_j}$$

- IC attempts at defining the correlation level between the actual implied volatility of the index and the implied volatilities of its stock components.
- Under some reasonable conditions, one can show that the term $\sum_{i=1}^N w_i^2 \sigma_i^2$ is close to zero. A good proxy for the IC is

$$\rho_{imp}^* = \frac{\sigma_I^2}{\left(\sum_{i=1}^N w_i \sigma_i\right)^2}$$

Dispersion trading: the implied correlation Ctd

- Given the average implied volatility of the components $\sigma_{Avg} = \sum_{i=1}^{N} w_i \sigma_i$ we can approximate the index volatility as follow

$$\sigma_I \approx \sqrt{\rho^*_{imp}} \sigma_{Avg}$$

for $\rho^*_{imp} > 0.15$ and $N > 20$.

- The Realised Correlation (RC) was proposed as an alternative to the implied correlation. Denoted $\widehat{\rho}$, it is defined as

$$\widehat{\rho} = \frac{\sum_{i \neq j}^{N} w_i w_j \rho_{ij}}{\sum_{i \neq j}^{N} w_i w_j}$$

- It can be approximated as

$$\widehat{\rho} = \frac{\widehat{\sigma}_I^2}{\left(\sum_{i=1}^{N} w_i \widehat{\sigma}_i \right)^2} = \frac{\widehat{\sigma}_I^2}{(\widehat{\sigma}_{Avg})^2}$$

where $(\widehat{\sigma}_I, \widehat{\sigma}_1, ..., \widehat{\sigma}_N)$ represent realised volatilities.

Dispersion trading

- Dispersion trading (DT) consists in selling the index option and buying options on the index components.
- Since all trades are delta-neutral, dispersion trading
    - sells index volatility and buys volatility on the index components
- It is a hedged strategy designed at taking advantage of relative value differences in implied volatilities between an index and a basket of component stocks.
- There are multiple instruments available to implement dispersion trading.
    - When using vanilla options it is necessary to delta-hedge each option.
    - Straddles and strangles have volatility exposure with limited delta.
- A DS that buys index straddles/strangles and sells straddle/strangle positions on individual components is hedged against large market movement and has low volatility risk.

Dispersion trading: Defining the problem

- We need to properly defining the relationship between an option on a basket and a weighted sum of options on its components.
- Setting the strike on the index option as

$$K = \sum_{i=1}^{N} \omega_i K_i$$

and applying Jensen's inequality, we get the following relation on the intrinsic value

$$\max\left(B(T,T) - K, 0\right) \leq \sum_{i=1}^{N} \omega_i \max\left(F_i(T,T) - K_i, 0\right)$$

- Taking the expectation of the payoff under the RN-measure and discounting, we can rewrite this relation in terms of prices as

$$C_B(t, B(t,T), K, T) \leq \sum_{i=1}^{N} \omega_i C_i(t, F_i(t,T), K_i, T)$$

Approximating the IV of a basket option

- Expressing the IV $\Sigma_B(K,T)$ of an option on a basket in terms of IV $\Sigma_{F_i}(K_i,T)$ of its components is not an easy task.
- As an example, we consider the model

$$C(t,S_t;K,T;I(t,T)) = C_{BS}(t,S_t,K,T;I_0(t,T)) + \alpha(K,T)$$

where $I_0^2(t,T) = E_t[I^2(t,T)]$ is the expected mean variance and $\alpha(K,T)$ produces the skew.

- We assume that the market price could be decomposed into the linear combination

$$C(t,x,K,T) = \sum_{i=1}^{n} a_i \big(C_{BS}(t,x,K,T;\Sigma_i(K,T)) + \alpha_i(K,T) \big)$$

where the weights $a_i > 0$ and such that $\sum_{i=1}^{n} a_i = 1$.

- In the special case where the strike is ATM-forward, the approximate implied volatility simplifies to

$$\Sigma(K,T)\big|_{K=F(T)} \approx \overline{\sigma} + \frac{1}{vega(\overline{\sigma})}\big(\sum_{i=1}^{n} a_i \alpha_i(K,T)\big)$$

  where $\overline{\sigma} = \sum_{i=1}^{n} a_i \sigma_i$.
- Going further and setting $\alpha_i(K,T) = 0$, $\forall i$, we get

$$\Sigma(K,T)\big|_{K=F(T)} \approx \overline{\sigma}$$

  which gives an idea of the over-simplification made by pratitioners for using Markowitz's portfolio theory.

A measure of dispersion

- A simple way of understanding the properities of linear combination of European option prices in the BS world is to assume options are ATM forward and to linearise them.
- Given $K = B(t, T)$ and $K_i = F_i(t, T)$, the implied volatilities no-longer depend on strikes, and the price relation above simplifies to

$$B(t, T)e^{-q(T-t)}\sigma_B\sqrt{T-t} \leq \sum_{i=1}^{N}\omega_i F_i(t, T)e^{-q_i(T-t)}\sigma_{F,i}\sqrt{T-t}$$

  where σ_B is the ATM IV of an option on a basket and $\sigma_{F,i}$ is the ATM volatility of its ith component.
- We can rewrite the relation as

$$\sigma_B \leq \sum_{i=1}^{N}\widehat{\omega}_i(t)\sigma_{F,i}$$

A measure of dispersion Ctd

- We can approximate the stochastic weights $\widehat{\omega}_i(t)$ with their time zero values, getting $\overline{\omega}_i$, $i = 1, .., N$.
- We can then approximate the no-arbitrage relation for ATM implied volatility of an option on a basket as

$$\sigma_B \leq \sum_{i=1}^{N} \overline{\omega}_i \sigma_{F,i} = \sigma_{WB}$$

  where $\sigma_{WB} = \sum_{i=1}^{N} \overline{\omega}_i \sigma_{F,i}$ is the approximated average implied volatility of the basket components.
- Thus, we define an ATM indicator of arbitrage between an option on a basket and a weighted sum of options on its components as follow

$$M_\rho = \frac{\sigma_B^2}{\sigma_{WB}^2} \leq 1$$

A measure of dispersion Ctd

- If M_ρ is close to 1, then the implied volatility of a basket option, σ_B, is very expensive, and it is likely to move down.
- If this indicator is greater than one, there is an arbitrage and we should sell the option on the basket.
- This indicator does not represent the market's expectation of the future realised correlation.

Remark

One can only relate Markowitz's portfolio theory to option pricing theory when the options have been linearised and the stochastic weights have been frozen.

Pricing dispersion

- When dealing with DT, we must be concerned by the ability of controlling the risks involved.
- As the basket components move, we must be able to price each individual option and compute its Greeks.
- We use the MixVol model to relate single option prices with different strikes, and compute their Greeks.
- The super-replicating portfolio $\widehat{C}_B(t_0, K, t)$ is given by

$$\widehat{C}_B(t_0, K, t) = \sum_{i=1}^{N} \omega_i \Big(\sum_{j=1}^{n} a_{ij}(t) Call_{BS}(t_0, \overline{K}(K, t), t, \Sigma_{ij}(t)) \Big)$$

  where $a_{ij}(t)$ for $j = 1, .., n$ are the weights for the ith underlying, $\Sigma_{ij}(t)$ is the volatility for the ith underlying.
- We calibrate independently the MixVol model's parameters $a_{ij}(t)$, μ_{ij}, $\Sigma_{ij}(t)$ to the liquid market prices of each underlying $F_i(0)$ for $i = 1, .., N$.

Pricing dispersion Ctd

- Assuming some dynamics for the underlying stock price, we can dynamically manage the evolution of the replicating portfolio.
- In practice we do not know the true volatility of the process, but only a misspecified one.
- Hence, when risk managing an option we are concerned with the PnL to be made when hedging options that are mispriced by the market.
- One can show that if the misspecified volatility dominates (or is dominated by) the true volatility, then the self-financing value of the misspecified hedging portfolio exceeds (or is dominated by) the payoff of the contingent claim at expiration.

bibliography I