Stochastic Models in Energy Markets

Matteo Gardini

Helsinki, 26 February 2025

Who am I?

My name is Matteo, I work as Quantitative Analyst at Eni Plenitude, Italy.

`matteo.gardini88@gmail.com`

Table of Contents

Bibliography and references

Key concepts are contained in:

- Brigo et al. [2007]: gives a very useful overview of the topic and provides some *ready to use* MATLAB code. It deals with **historical calibration** and **processes simulation**.
- Cont and Tankov [2003, Chapter 13]: it deals with **risk-neutral** calibration and provides a rich overview of Lévy processes applied to the mathematical finance.
- Oosterlee and Grzelak [2019]: a huge book with theory and a lot of code in MATLAB and Python.

Table of Contents

Working path

Working path

- Analyze the market and try to find the model which better describe the phenomena.
- Once you have selected the model, $\mathscr{M}(\Theta)$, try to fit the parameters Θ.
  - **Historical calibration**: observe a time series and use it to get Θ.
  - **Risk Neutral calibration**: look at the derivative market and use it to properly fit $\mathscr{M}(\Theta)$.
- Apply your model to price derivatives, compute risk metrics, simulate path and so on.

Table of Contents

Calibration

Once you defined a model $\mathscr{M}(\Theta)$, you have to **calibrate** it: this means you have to fit the parameters according to some real data.

This is a very important and delicate topic. Roughly speaking, calibrating a model leads to **solve a optimization problems**. Commonly used calibration techniques are:

- Log-Likelihood estimation.
- Least-Squares.
- Generalized Method of Moments.
- Genetic Algorithms.
- Kalman Filter.

Model Calibration

Calibration is considered an hard-topic for several reasons:

- The objective function to minimize may be highly non-linear with multiple local minima, making it unlikely to reach the global minimum.
- Inverse problems are ill-posed, meaning the solution may not be unique.
- Advanced optimization algorithms are necessary, most of which require computing the gradient of the objective function—a numerically challenging task.

Table of Contents

Historical Calibration - Maximum Likelihood Estimation

The idea behind historical calibration is simple: *given a set of observed data try to find a set of parameters that are "compatible" with the observed data in some sense.* Usually, a **Maximum Likelihood Estimation** (MLE) technique is used.

Definition (Likelihood)

Given a parameterized family of probability density functions $f(x|\Theta)$

$$x \longmapsto f(x|\Theta)$$

where Θ is a parameter, the **likelihood function** is

$$\Theta \longmapsto f(x|\Theta)$$

written $\mathscr{L}(\Theta) = f(x|\Theta)$ where x is the outcome of the experiment.

Observe that the likelihood is not a probability density function.

Consider the normal distribution and a given data-set $\boldsymbol{x} = (x_1, \ldots, x_n)$ of independent realizations of an experiment, $\Theta = (\mu, \sigma)$ and the *pdf* given by:

$$f(x|\Theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

The likelihood is $f(x|\Theta)$ regarded as a function of Θ. If $\{x_i\}_{i\in\mathbb{N}}$ are i.i.d., then the likelihood of the random vector $\boldsymbol{x}$ is given by:

$$\mathscr{L}(\Theta) := f(\boldsymbol{x}|\Theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}.$$

Exercize: find the value of μ and σ that maximize the likelihood $\mathscr{L}(\Theta)$ (Solution in Brigo et al. [2007]).

How to use this concept to calibrate a model?

Maximum Likelihood Estimation (intuitively)

Find the set of parameters Θ such that, under the assumed statistical model $\mathscr{M}(\Theta)$, the observed data-set is the most probable. The point in the parameter space that maximizes the likelihood function is called the *maximum likelihood estimate.* In order to do so, maximize the likelihood $\mathscr{L}(\Theta)$ with respect to Θ.

Maximum Likelihood Estimation (formally)

Let $\mathscr{M}(\Theta)$ be a model and let Θ be its set of parameters. Let $\mathscr{L}(\Theta)$ be the likelihood of observing a particular data sample. The *MLE* $\hat{\Theta}$ is such that:

$$\hat{\Theta} = \arg\max_{\Theta} \mathscr{L}(\Theta)$$

Remark

In order to successfully apply this method, the likelihood function must be available is some "nice" form (as in the previous example). Under some model assumptions the likelihood assumes a simple form.

Historical Calibration - Maximum Likelihood for Lévy processes

How to find the likelihood for Lévy processes (see Brigo et al. [2007])?

Let $Y = \{Y(t); t \geq 0\}$ be a Lévy process: we model the risky asset process $S = \{S(t); t \geq 0\}$ as:

$$S(t) = S(0)e^{Y(t)}, \quad t \geq 0, \quad S(0) = S_0.$$

Assume to have a given data sample $x = (x_1, x_2, \ldots, x_n)$[1]. In our example we suppose that the realization of the process $Y(t)$ are evenly spaced on a time grid $t_0, t_1, \ldots, t_n$ where:

$$\Delta t = t_{i+1} - t_i.$$

[1] We state with x_i the realization of the increment of the process Y at time t, i.e. $x_i = X(t_i) := Y(t_i) - Y(t_{i-1}) = \log S(t_i) - \log S(t_{i-1})$.

Let be:

$$X\left(t_i\right) := \log S\left(t_i\right) - \log S\left(t_{i-1}\right).$$

By definition the likelihood function is:

$$\mathscr{L}\left(\Theta\right) := f_{X(t_0),\dots,X(t_n);\Theta}.$$

Since X is a Markov process then it simplifies:

$$\begin{aligned}\mathscr{L}\left(\Theta\right) &= f_{X(t_0),\dots,X(t_n);\Theta}\\ &= f_{X(t_n)|X(t_{n-1});\Theta} \cdot f_{X(t_{n-1})|X(t_{n-2});\Theta} \cdot \dots \cdot f_{X(t_0);\Theta}\end{aligned}$$

and, since the increments of Y are i.i.d.

$$f_{X(t_i)|X(t_{i-1});\Theta} = f_{X(t_i);\Theta},$$

and eventually:

$$\mathscr{L}\left(\Theta\right) = \prod_{i=1}^{n} f_{X(t_i);\Theta}.$$

Substituting the observed data set $x = (x_1, \dots, x_n)$, we get:

$$\mathscr{L}\left(\Theta\right) = \prod_{i=1}^{n} f_{\Theta}\left(x_i\right).$$

Historical Calibration - Maximum Likelihood for Lévy processes II

Since the product of very small numbers is hard to maximize, it is better to compute:

$$\mathscr{L}^{*}(\Theta) = \log \mathscr{L}(\Theta) = \sum_{i=1}^{n} \log f_{\Theta}(x_i)$$

Historical Calibration - Maximum Likelihood for Lévy processes

The set of parameters $\hat{\Theta}$ which maximizes the likelihood is obtained by solving:

$$\hat{\Theta} = \arg\max_{\Theta} \mathscr{L}^{*}(\Theta).$$

Such maximization can be solved analytically only in rare cases, such as the Black-Scholes model (Brigo et al. [2007]). More commonly, numerical techniques are required. Tools like *MATLAB*, *Python*, and *R* offer numerous ready-to-use numerical algorithms.

How to apply MLE techniques to the **Variance Gamma** process?

In the Variance Gamma model the log-prices of a given risky asset S follow the dynamic:

$$d\log S(t) = \theta\, dg(t) + \sigma\, dW\big(g(t)\big), \quad S(0) = S_0,\ t \geq 0.$$

where $\theta \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, $g = \{g(t); t \geq 0\}$ is a Gamma process $g(t) \sim \Gamma(\frac{t}{\nu}, \nu)$, $\nu \in \mathbb{R}^+$ and W is a standard Brownian motion independent of g. In this case the vector parameter we want to estimate is $\Theta = (\theta, \sigma, \nu)$.

In order to write the likelihood we must have a form for $f_\Theta(x_i)$ which is given by (see Brigo et al. [2007]):

$$f_\Theta(x_i) = \frac{2e^{\frac{\theta x}{\sigma^2}}}{\sigma\sqrt{2\pi}\nu^{\Delta t/\nu}\Gamma(\frac{1}{\nu})}\left(\frac{|x|}{\sqrt{\frac{2\sigma^2}{\nu}+\theta^2}}\right)^{\frac{\Delta t}{\nu}-\frac{1}{2}} K_{\frac{\Delta t}{\nu}-\frac{1}{2}}\left(\frac{|s|\sqrt{\frac{2\sigma^2}{\nu}+\theta^2}}{\sigma^2}\right) \tag{1}$$

where $\Gamma(x)$ is the Gamma function and $K_\eta(\cdot)$ is the modified Bessel function of the Third kind.

Note: Gamma and Bessel functions can be computed numerically in a very efficient way.

Note: Equation (1) seems to be an act of faith! See Brigo et al. [2007] for a brief explanation.

Exercise: derive the expression in (1).

Now we can get the desired $\hat{\Theta} = \left(\hat{\theta}, \hat{\sigma}, \hat{v}\right)$ by numerically computing:

$$\hat{\Theta} = \arg\max_{\Theta} \mathscr{L}^* \left(\Theta\right),$$

where

$$\mathscr{L}^* \left(\Theta\right) = \sum_{i=1}^{n} \log f_{\Theta}\left(x_i\right)$$

Since $\mathscr{L}\left(\Theta\right)$ is not linear, local maxima may be present. Most optimization algorithms require a starting point, Θ_0, and the optimal value of $\mathscr{L}\left(\Theta\right)$ may depend on the chosen Θ_0. Is there a systematic way to select a good starting point instead of choosing it randomly?

Exercise: Run the code and try changing the starting point Θ_0 of the minimization algorithm. What do you observe?

We can use the **Generalized Method of Moments** to pick up the starting point Θ_0.

Remember that if X is a Variance Gamma process with parameters θ, σ, ν, the Moment Generating Function $M_X(u), u \in \mathbb{R}$ of the increment over Δt is given by:

$$M_X(u) = \left(1 - \theta\nu u - \frac{1}{2}\nu\sigma^2 u^2\right)^{-\frac{\Delta t}{\nu}}.$$

which can be used to compute the first four central moments $\left(\mathbb{E}\left[X^n\right] = M_X^n(0)\right)$:

$$\begin{aligned}
\mathbb{E}[X] &= \theta\Delta t \\
\mathbb{E}[(X - \mathbb{E}[X])^2] &= (\nu\theta^2 + \sigma^2)\Delta t \\
\mathbb{E}[(X - \mathbb{E}[X])^3] &= (2\theta^3\nu^2 + 3\sigma^2\nu\theta)\Delta t \\
\mathbb{E}[(X - \mathbb{E}[X])^4] &= (3\nu\sigma^2 + 12\theta^2\sigma^2\nu^2 + 6\theta^4\nu^2)\Delta t + (3\sigma^4 + 6\theta^2\sigma^2\nu + 3\theta^4\nu^2)\Delta t^2.
\end{aligned}$$

We recall the definition of variance V, skewness S and kurtosis K:

$$M = \mathbb{E}[X] \qquad V = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

$$S = \frac{\mathbb{E}[(X - \mathbb{E}[X])^3]}{(\mathbb{E}[(X - \mathbb{E}[X])^2])^{\frac{3}{2}}} \qquad K = \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{(\mathbb{E}[(X - \mathbb{E}[X])^2])^2}$$

For each of these quantities estimators exist, and hence estimate can be computed from the dataset $x = (x_i, \dots, x_n)^2$ and compared with the theoretical ones obtained from previous relations.
For example

$$\hat{M} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \hat{V} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

$$\hat{S} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^3}{\hat{V}^{\frac{3}{2}}} \qquad \hat{K} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^4}{\hat{V}^2}$$

Now match M with $\hat{M}$, V with $\hat{V}$ and so on ...

If we assume that θ is "small" (as typically is in the applications) we get the following set for the starting point Θ_0 of our optimization algorithm:

$$\sigma_0 = \sqrt{\frac{V}{\Delta t}}, \quad \nu_0 = \left(\frac{K}{3} - 1\right)\Delta t, \quad \theta_0 = \frac{S\sigma\sqrt{\Delta t}}{3\nu}.$$

Maximum Likelihood Estimator starting from Θ_0

Now we can run the optimization algorithm and maximizing $\mathscr{L}(\Theta)$ with respect to Θ, choosing as starting point $\Theta_0 = (\theta_0, \sigma_0, \nu_0)$ and obtain the desired value $\hat{\Theta}$ such that:

$$\hat{\Theta} = \arg\max_{\Theta} \mathscr{L}^* (\Theta),$$

Generalized Method of Moments (GMM)

Roughly speaking, when selecting the starting point Θ_0, we choose parameters θ, σ, and ν to match the moments. This approach, known as the **Generalized Method of Moments**, is an alternative technique for calibrating our model.

[2] $x_i = \log S(t_i) - \log S(t_{i-1})$

Table of Contents

Risk-Neutral Calibration - Intuition

Call option prices (and those of other derivative contracts in general) reflect the market's "expectation of future behavior". Historical prices, on the other hand, represent past data and, therefore, "do not provide information about the future".

To price a derivative, you must use information about the future. Furthermore, the price of the derivative you're valuing must be consistent with the prices of other quoted market products. In other words, consistency means that **arbitrage opportunities must be avoided**.

eex Markets Market Data Trading Resources

Option Expiry 2025-12-11

Call					
Strike	Last Price	Last Volume	Settlement Price	Total Volume	Open Interest
51.00	-	0	16,322	0	0
52.00	-	0	15,572	0	0
53.00	-	0	14,842	0	0
54.00	-	0	14,134	0	0
55.00	-	0	13,448	0	0
56.00	-	0	12,785	0	0

Figure 1: Market snapshot: call option prices on Power futures German calendar.

We denote by C_i the market price of a call option.

In a Risk-neutral world the risky-asset is modeled as

$$S(t) = S(0)e^{\omega t + rt + Y(t)}$$

where $Y = \{Y(t); t \geq 0\}$ is a Lévy process, r is the risk-free rate and ω is a parameter that must be chosen such that the discounted prices are martingales. Given the characteristic function of Y:

$$\phi_Y(u) = \mathbb{E}[e^{iuY}]$$

where $i = \sqrt{-1}$, a simple way to get the martingale condition is to chose ω such that:

$$\phi_Y(-i) = -\omega t. \tag{2}$$

In the Variance Gamma condition (2) leads to:

$$\omega = \frac{1}{\nu} \log\left(1 - \frac{\sigma^2 \nu}{2} - \theta \nu\right).$$

Risk Neutral Calibration - Idea

Let:

- C_i for $i = 1, \ldots, n$ n European call option prices observed in the market with underlying asset S.
- Θ represent the vector of unknown parameters.
- $C_i^{\Theta}(K,T)$ denotes the price output from the chosen market model.

The optimal $\hat{\Theta}$ can be found by solving:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{i=1}^{n} \left(C_i^{\Theta}(K,T) - C_i \right)^2, \tag{3}$$

Intuitive explanation

This time, we are looking for the set of parameters Θ that minimizes the distance between the model output prices C_i^{Θ} and the market prices C_i.

Since want to solve:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{i=1}^{n} \left(C_i^{\Theta}(K,T) - C_i \right)^2 . \tag{4}$$

we need a method to compute $C_i^{\Theta}(K,T)$ for the given model $\mathscr{M}(\Theta)$.

Typically, fast methods for computing vanilla option prices (such as call options) are available for many models.

Since we must use a minimization algorithm to solve the problem above, and optimization algorithms are generally iterative, the method for computing $C_i^{\Theta}(K,T)$ must be as fast as possible.

What are the commonly available methods to compute $C_i^{\Theta}(K,T)$?

- Analytical formula ✓.
- Monte Carlo Methods ✗.
- Partial Differential Equation methods ✗.
- Fourier Methods: FFT, Lewis, Convolution, COS and so on ✓.

Analytical or semi-analytical formulas are not always available, but Fourier Methods can be used for almost all Lévy models.

Conclusion

By minimizing with respect to Θ:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{i=1}^{n} \left(C_i^{\Theta}(K,T) - C_i \right)^2,$$

we get $\hat{\Theta}$ and hence the model $\mathscr{M}(\Theta)$ is calibrated on the quoted derivatives products.

For the variance Gamma model a semi-analytic closed formula for Call options is available and it is given by:

$$C^{\Theta}(K,T) = \int_0^{\infty} C(g) \frac{\frac{1}{\nu}^{\frac{t}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)} g^{\frac{t}{\nu}-1} e^{\frac{-g}{\nu}} \tag{5}$$

where $C(g)$ is the price of a Call within the Black-Scholes model. The integral in Equation (5) may appear intimidating, but it can actually be computed very efficiently using Bessel functions and the degenerate hypergeometric function, as demonstrated by Madan and Seneta [1990].

Obs: In the following numerical experiment we do not use the explicit formula but the FFT method to compute the call option price $C^{\Theta}(K,T)$. This because FFT method can be applied to other Lévy models and hence it is a widely used method!

Table of Contents

Numerical Application - Variance Gamma

We apply all methods analyzed before to a real case. We focus on Power Germany Future prices. The outline is the following:

1. We check that Lévy modeling is suitable.
2. We Calibrate the Variance Gamma model on historical future prices (use **MLE**).
3. We Calibrate the Variance Gamma model on quoted Call options written on power Germany future prices (use **LS**).
4. We compare the obtained results.

We assume a risk-free rate of $r = 0.01$.

A power futures contract is a standardized financial contract that obligates the buyer to purchase, and the seller to deliver, a specified amount of electricity at a future date, at a predetermined price. These contracts are used by participants in the energy markets (such as utilities, producers, traders, and hedgers) to manage price risk and speculate on future energy prices.

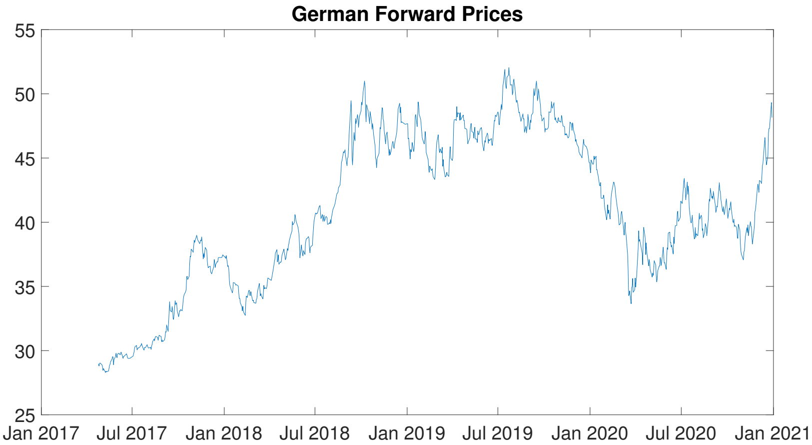

Figure 2: Historical German Power Future quotations.

First of all we have to check that increments (i.e. log-returns) of n log-price process realizations are independent. One possible way is to plot the so called *Autocorrelation function* of lag k defined as:

$$ACF(k) = \frac{1}{(n-k)\hat{v}} \sum_{i=1}^{n-k} (x_i - \hat{m})(x_i + k - \hat{m}), \quad k = 1, 2, \ldots$$

where $\hat{m}$ and $\hat{v}$ are the sample mean and variance of the series, respectively. $ACF(k)$ gives an estimate of the correlation between $X(t_i)$ and $X(t_{i+k})$. If we want to use Lévy modeling framework we must observe a low level of increments correlation[3].

[3] Use the function *autocorr* in *MATLAB*.

Numerical Application - Independence of increments

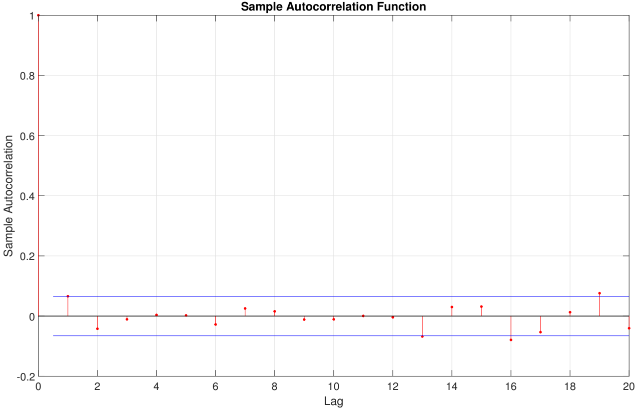

Figure 3: We see that no auto-regressive component in log-returns seems to be present. A Lévy process should be okay for modeling purposes.

Now we have to choose the "right" Lévy process to model the log-return process. To check the normality of log-returns we can use the QQ-plot (use *qqplot* function in *MATLAB*.)

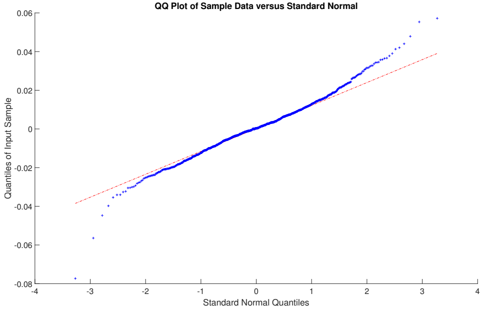

Figure 4: The QQ-plot shows that the historical quantiles in the tail of the distribution are significantly larger compared to the normal distribution.

Using the Black-Scholes model might be too restrictive, as log-returns are not normally distributed. Therefore, we use the Variance-Gamma process instead.

θ	σ	ν
0.1873	0.1917	0.002

Table 1: Historical calibration of the Variance Gamma model using the MLE estimator.

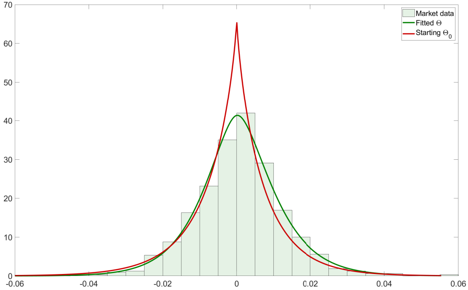

Figure 5: Fitted distribution compared to the historical one.

Numerical Application - Risk-Neutral Calibration

θ	σ	ν
0.0900	0.2220	0.2231

Table 2: Risk-Neutral calibration of the Variance Gamma model using Least-Squares.

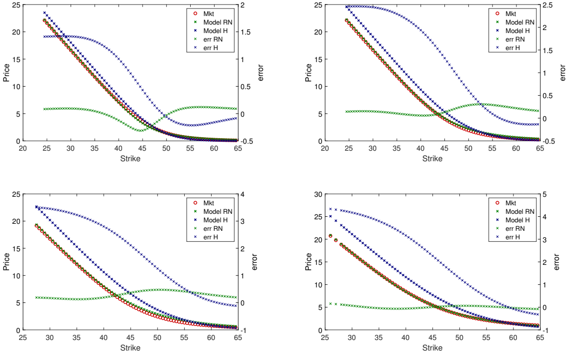

Figure 6: Market Repricing for different maturities: observe that historical calibration and Risk-Neutral calibration leads to different parameters and hence to different derivative pricing.

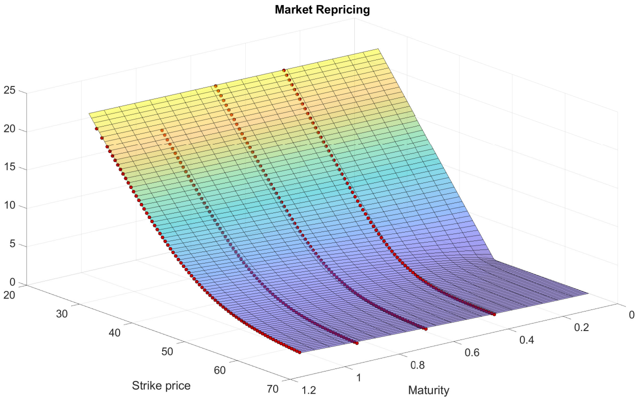

Figure 7: Market repricing using Risk-Neutral parameters.

Important remark:

"Never" use historical calibration for derivative pricing! The obtained prices are not coherent with the ones observed in the market!

Obs: Sometimes the derivative market lacks liquidity, making it reasonable to use parameters from historical calibration for derivative pricing.

Definition (Implied volatility)

In financial mathematics, the implied volatility $\hat{\sigma}$ of a Call option contract C_{mkt} is that value of the volatility of the underlying instrument which, when input in an option pricing model (such as Black–Scholes), will return a theoretical value equal to the current market price of said option. In formula the value $\hat{\sigma}$ such that:

$$C_{mkt} = C(t, S(t), \hat{\sigma}, r, K)$$

Black-Scholes model assumes that the volatility is constant. But, if you retrieve the implied volatility from quoted option prices varying the strike price you get something strange...

Numerical Application - Volatility smiles

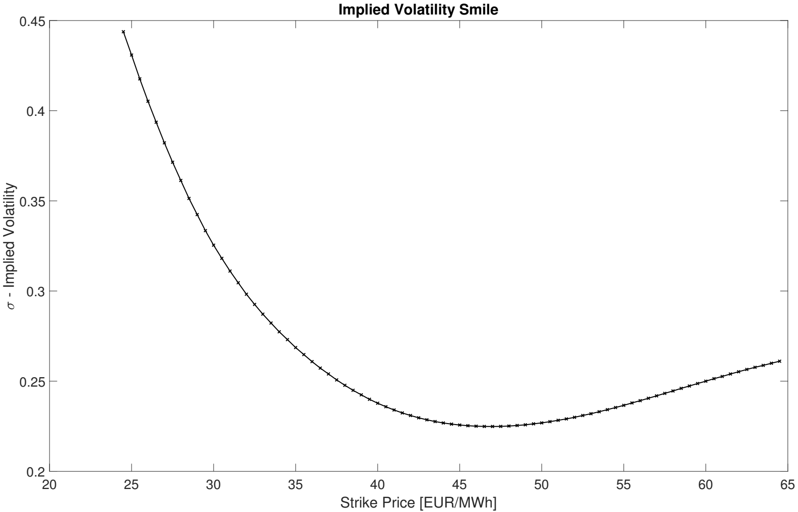

Figure 8: Volatility smile varying the strike price K, same maturity T.

Numerical Application - Volatility surface

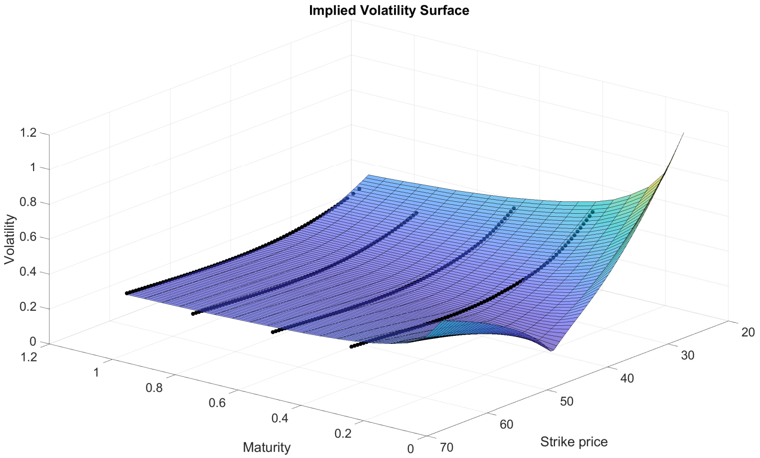

Figure 9: Volatility surface generated by the Variance Gamma model. It is able to fit the implied market volatility. Unlike the Black-Scholes model the volatility is no more constant.

Conclusions

- **Risk-Neutral calibration of Variance Gamma process is easy** since a semi-analytic option pricing formula for Vanilla option is available.
- Variance Gamma distribution fits the market log-returns better than the normal distribution.
- Variance Gamma can replicate **volatility smiles**.

Bibliography I

D. Brigo, A. Dalessandro, Neugebauer M., and F. Triki. A Stochastic Process Toolkit for Risk Management. Technical report, 2007.

R. Cont and P. Tankov. *Financial Modelling with Jump Processes.* Chapman and Hall, 2003.

D. B. Madan and E. Seneta. The Variance Gamma (V.G.) Model for Share Market Returns. *The Journal of Business*, 63(4):511–24, 1990.

C. W. Oosterlee and L. A. Grzelak. *Mathematical Modeling and Computation in Finance.* World Scientific, 2019.