Smooth surface generator

Arnaud RIVOIRA

January 25, 2019

1 Objective

The objective is to generate instances of smooth surfaces for the local volatility model.

The challenges are:

- Being able to produce a large choice of shapes with various types of curvatures (concave and convex profiles, saddle points, etc.)
- Avoiding the machine training to be polluted by too weird fluctuating surfaces

2 Framework description

It is going to be considered that the surface is represented by two coordinates (x, y) corresponding respectively to the rescaled log-moneyness $X_t = \log(S_t/S_0)/\log(S_{\max}/S_0)$ and $y = \sqrt{T/T_{\max}}$.

The main rationale for that choice is that the coordinates x and y are expected to be of the same nature, which is not the case for strike and maturity. In particular, for short maturities, any diffusive model tends to behave like a heat kernel which admits as invariant the combined rescaling $(x, t) \mapsto (\lambda x, \lambda^2 t)$ i.e. (meaning that space quantities translate naturally in terms of the square root of time quantities).

To make sure that the volatility never zeroes, we assume it is the exponential of some bivariate function P

$$\sigma(x, y) = e^{\gamma P(x - c_X, y - c_Y)} \tag{1}$$

where P is some multivariate polynomial:

$$P(x, y) = a^0 + a_0^1 x + a_1^1 y + a_0^2 \frac{x^2}{2} + a_1^2 \frac{2xy}{2} + a_2^2 \frac{y^2}{2} + \cdots \tag{2}$$

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{a_k^n}{k!(n-k)!} y^k x^{n-k} \tag{3}$$

and c_X and c_Y are two uniform random variables respectively in $[-1, 1]$ and $[0, 1]$.

The purpose of c_X and c_Y is to average the surface over the expected support of x and y to decrease the dependence of the probability distribution of the surface to the reference point of the Taylor expansion.

From Appendix.4.1, the coefficients of the expansion are related to the derivatives of P as follows:

$$a_k^n = \frac{\partial^n P(0,0)}{\partial y^k \partial x^{n-k}} \tag{4}$$

To insure that the generated surfaces are smooth, the derivatives must decay fast enough to zero. As changing the scale on the x and y is exactly equivalent to applying an exponential damping of the coefficients, we are going to add an extra-parameter α to insure that the generated surface remains close to a polynomial function outside the expected range of values for both x and y. Hence, a_k^n are going to be drawn as independent Gaussian variables of zero mean and standard deviation $\sigma_a(k,n) = e^{-\alpha n^2}$, where α controls the growth of the surface outside of the target unit box.

The benefit of using a non local basis such as that of the multivariate polynomials is precisely that the coefficients can be drawn independently without losing the smoothness property. That would, for instance, not have been the case drawing directly the samples $\sigma(x_i, y_j)$. In that case indeed, we would have had to introduce a strong correlation structure between close samples, making it somehow more involved to generate.

To determine the overall level of dispersion γ of the polynomial coefficients, one can consider the level of relative deviation of P at the origin:

$$\rho^2 = \frac{\mathbb{E}\left\{P(0,0)^2\right\} - \Big(\mathbb{E}\left\{P(0,0)\right\}\Big)^2}{\Big(\mathbb{E}\left\{P(0,0)\right\}\Big)^2} \tag{5}$$

$$= \frac{e^{2\gamma^2}}{e^{\gamma^2}} - 1 = e^{\gamma^2} - 1 \tag{6}$$

Hence:

$$\rho = \sqrt{e^{\gamma^2} - 1} \tag{7}$$

For $\gamma = 1/2$, the level of relative deviation is expected to be of circa 50%, which looks rather reasonable.

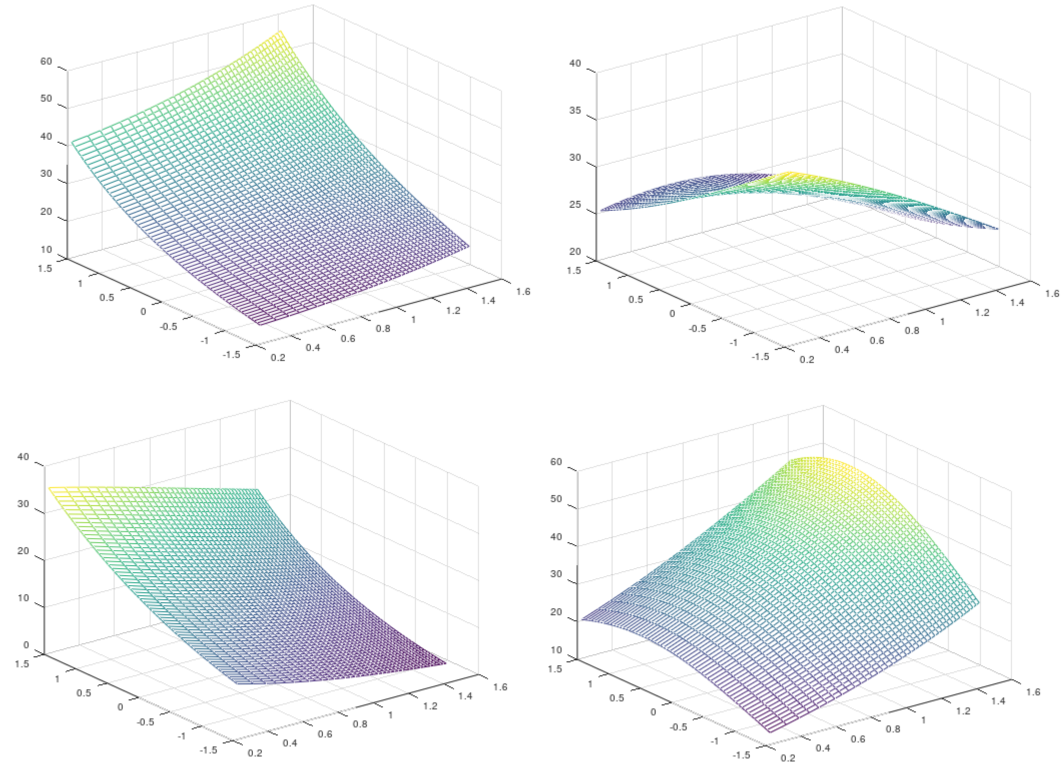
In real life, of course, the series is going to truncated. The maximal degree N must be related to the decay factor α by $e^{-\alpha N} = \epsilon$, where ϵ is the numerical tolerance of the truncation.

3 Examples of generated surfaces

Please find thereafter the kind of surfaces we obtain with the following parameters:

- $\alpha = 1/2$
- $\gamma = 1/2$
- Maximal degree for polynomial : 10

The values have all been scaled by a constant factor 25 corresponding to the average level of volatility in $\%/\sqrt{\text{year}}$.



4 Appendix

4.1 Multivariate Taylor

Taylor expansion reads:

$$P(a,b) = \left(e^{b\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}}P\right)(0,0) \tag{8}$$

Hence:

$$\begin{aligned}
e^{a\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}} &= \sum_{n=0}^{\infty}\frac{1}{n!}\left(a\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}\right)^{n} && (9)\\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}\left(a\frac{\partial}{\partial x}\right)^{k}\left(b\frac{\partial}{\partial y}\right)^{n-k} && (10)\\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{a^k}{k!}\frac{b^{n-k}}{(n-k)!}\frac{\partial^k}{\partial x^k}\frac{\partial^{n-k}}{\partial y^{n-k}} && (11)
\end{aligned}$$

Leading to:

$$P(a,b) = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{a^k}{k!}\frac{b^{n-k}}{(n-k)!}\frac{\partial^n P(0,0)}{\partial x^k\partial y^{n-k}} \tag{12}$$