KRIGING FOR IMPLIED VOLATILITY SURFACE

Areski Cousin, Djibril Gueye

KRIGING FOR IMPLIED VOLATILITY SURFACE

Cousin Areski
Institut de Recherche Mathématiques Avancées
Université de Strasbourg
a.cousin@unistra.fr

Gueye DJibril
Institut de Recherche Mathématiques Avancées
Université de Strasbourg
djibril.gueye@unistra.fr

ABSTRACT

Implied volatility surface is of crucial interest for risk management and exotic option pricing models. Its construction is usually carried out in accordance with the arbitrage-free principle. This condition leads to shape restrictions on the option prices such as monotonicity with respect to maturities and convexity with respect to strike prices. In this paper, we propose a new arbitrage-free construction method that extends classical spline techniques by additionally allowing for quantification of uncertainty. The proposed method extends the constrained kriging techniques developed in [MB16] and [CMR16] to the context of volatility surface construction. Assuming a Gaussian process prior, the posterior price surface becomes a truncated Gaussian field given shape constraints and market observations. Prices of illiquid instruments can also be incorporated when considered as noisy observations. Starting from a suitable finite-dimensional approximation of the Gaussian process prior, the no-arbitrage condition on the entire input domain is characterized by a finite number of linear inequality constraints. We define the most likely response surface and the most-likely noise values as the solution of a quadratic optimization problem. We use Hamiltonian Monte Carlo technics to simulate the posterior truncated Gaussian surface and build pointwise confidence bands. The Gaussian process hyper-parameters are estimated using maximum likelihood. The method is illustrated on Euro Stoxx 50 option prices by building no-arbitrage volatility surfaces and their corresponding confidence bands.

1 Introduction

Volatility surfaces are important building blocks of risk management systems designed for pricing and hedging of financial derivatives or optional guarantees under market-consistent asset models. For a given underlying asset, the option value depends on the characteristics of the contract, i.e., the date of maturity and the strike price for equity options. However, the market value of such options is typically available (or reliable) for a limited number of standard characteristics (maturity, strike price) whereas one may require the whole price surface or the price for some unobserved characteristics for some applications. In addition, the construction of these surfaces is usually made in accordance with the principle of no arbitrage opportunity, i.e, the constructed options prices must be exempted financial arbitrage. Constructing no-arbitrage price surface can be cast into a functional learning problem with shape constraints and possibly noisy observations.

Once the price surface is interpolated, we can use the inversion formula of Black and Scholes for obtaining the implied volatility surface. Besides it is possible to construct the local volatility surface using the implied volatilities rather than prices. This consists in extracting the local volatility surface from the implied volatility one using the formula (1.10) of [Gat11]. These approaches that use the observed prices (or implied volatility for constructing local volatility) in the volatility surfaces construction are known as indirect.

Beyond these direct approches, there exist some ones which parametrize directly the unobserved volatility through a function in which its parameters can be estimated by minimizing a suitable objective function obtained for calibrating that function to market data. These include among others the stochastic volatility inspired (SVI) which parametrize the implied volatility smile and its extension (SSVI) for the implied volatility surface (see [GJ14; Gat04]).

The most known techniques in the context of volatility surface construction using indirect approaches have been based on constrained splines, see for example, [Fen09; Hom11; Lau11; WYQ04] and so forth.
Some techniques in deep learning have recently been developed in order to fill the gap in the usual models by recording better performances of learning arbitrage-free price surface. The work of [Dug+09] is the starting point for incorporating no-arbitrage constraints in neural networks (NN) for learning option prices. [Dug+09] used the so-called hard constraints approach through a special neural network architecture. This approach is different to the one called soft constraints approach which consists in penalizing some variables, in the objective function, that do not verify the constraints (see, e.g., [Itk19]).
The main difficulty of the approach of [Dug+09] relies on the fact that incorporating inequality constraints in the neural networks reduces its flexibility in practice and presents highly non linear transformation for option pricing computation at each iteration as shown in [ATV19].
For constructing the local volatility surface, [ATV19] proposed to directly learn the implied volatility surface instead of prices surfaces by using the soft constraint approach.
[CCD20] proposed simple neural network architectures for both arbitrage-free option prices and implied volatility. They show that hard constraints reduce the power of the network and that soft constraints provide best accurate prices and implied volatilities.
The main limit of these approaches lies on the fact that they do not allow to quantify uncertainty in the estimation of the variables of interest.

The aim of this part is to adopt the constrained kriging techniques developed in [MB17] for imposing hard constraints and quantifying the associated uncertainty in the context of implied volatility surface construction. We use the indirect approaches that consist in learning prices through kriging and then use the inversion formula for obtaining the implied volatility surface. There is no need to extend the approach of [MB17] for more general inequality constraints as it is done in [LL+18], since we limit ourselves to the convexity and monotonicity constraints.
The idea is to assume that the distribution of the underlying process is a realization of a truncated Gaussian process, verifying a number of constraints. As for the rate curves (see [CMR16]), the compatibility constraints with the observations (price compatibility of European options) is translated into a linear relationship on the kriging process. The arbitrage-free hypothesis on option prices is equivalent to the existence of a martingale measure: the achievements of the kriging process must correspond to marginal distributions of martingales. According to Theorem 2.6 of [BJ+16], it is then necessary to build a family of marginal distributions from a bivariate Gaussian process which is conditionally:

- increasing with respect to the direction of maturities
- convex with respect to the strike prices
- solution of a linear relationship on certain maturities and the strike prices.

Therefore, the first two constraints correspond to the inequality constraints that must be taken into account by the kriging model. The third constraint makes the kriging model compatible with the observed European options with different maturities and exercise prices.

Kriging, known as Gaussian process (GPs) regression, is a spatial interpolation approach and its techniques have been developed in geostatistics for estimating the distribution of mineral resource in the ground given the relatively small set of boreholes (see, e.g., [Mat63; Cre90; KM82]). It gains popularity with the works of [WR06] and has been adapted into various areas such as hydrology, meteorology, epidemiology and among others. Recent works in kriging have been developed in quantitative finance. For instance, [SEG12] used kriging for learning model parameters when calibrating the Vasicek interest rate model under the risk neutral measure. They made a Gaussian prior on the zero coupon bond log prices and learnt model parameters by maximizing the log likelihood of the training data given the parameters. [Lud18] improves the Monte Carlo Least square method for the valuation of Bermuda option by using kriging for the regression step which consists in approximating the expected value from continuation. [DS+18] show the speed up of kriging comparing to monte carlo method for pricing options and approximating implied volatility. However they do not take into account the arbitrage-free conditions in their approximation. In derivative portfolio modeling, [DC18] used a multi-response kriging in portfolio valuation through computation of credit valuation adjustment (CVA). They show that kriging provides fast approximation of portfolio value. [Gon+19] explore the use of kriging in finance and show that this latter is a powerful tool for fitting the yield curve. However their interpolation does not take in consideration the arbitrage-free conditions too.

A first attempt at using kriging for local volatililty is done by [TR19] who place a Gaussian prior directly on the local volatility surface (to guarantee the positivity of the local volatility, they assign a positive function on the prior). Such an approach leads to a nonlinear least squares training loss function, as it involves the nonlinear transformation of the local volatility into the corresponding vanilla option prices. Such a loss function is not obviously amenable to gradient descent (stochastic or not), so the authors resort to a MCMC optimization. Moreover, the performance of their GPs is measured by in-sample RMSE and does not seem to be assessed for overfitting. Furthermore, they do not benchmark their approach against alternatives.
Actually, adding that positivity constraint and several others (such as monotonicity, convexity, boundedness) in kriging is not an easy task and constitutes a big challenge. Several approaches exist. The frameworks presented in [DVM12; Gol+15; RV10; WB16], among many others, are based on the fact that the constraints are only satisfied on a virtual observation locations and does not guarantee the constraints in the entire domain. For instance, for incorporating increasing monotonicity constraint in kriging, [WB16] proposed to select the locations with large probabilities of having negative derivatives as the virtual locations set in which they enforce the constraints. The approach of [Agr19] extended that of [WB16] for multiple constraints through an efficient method for sampling the posterior process based on the derivation of the posterior of the constrained Gaussian process using a linear operator.

Adding some inequality constraints in the quantities of interest leads to the fact that the posterior process is no more Gaussian and these constraints are usually infinite-dimensional. This motivated [MB14] to use the finite dimensional approximation of Gaussian processes for which the the inequality constraints are easy to check. They developed a sampling scheme called Rejection Sampling from the Mode (RSM) for sampling from truncated multivariate Gaussian distribution which is needed for estimating the posterior process. Their approach guarantees the constraints in the whole domain. [CMR16] show the extensions of classical spline interpolation by constrained kriging techniques developed in [MB14] to ensure non-arbitrable and error-controlled yield-curve and CDS curve interpolation. [LL+18] extended the framework of [MB14] for more general inequality constraints and proposed to use the Hamiltonian Monte Carlo (HMC) of [PP14] which is more efficient than the RSM of [MB14] for sampling from truncated multivariate Gaussian distribution.

Unlike spline interpolation and NN approaches, kriging makes it possible to quantify the uncertainty in the estimation of the variables of interest. It is possible for example, to obtain confidence intervals at the points of the surface where the observations are unavailable or considered unsuitable.

The main contributions of the paper is to show that kriging is a suitable tool for constructing option prices and quantifying uncertainty in the presence of noisy observation, and for computing the associated implied volatility.

The outline of the paper is organized as follows. In section 2, we recall some well known facts concerning arbitrage-free conditions. Section 3 emphasizes on constrained kriging and Section 4 is devoted to the numerical illustrations. We give, in Section 5, an extension of kriging in presence of interest rate and dividends.

2 Option pricing in no-arbitrage models

The goal of this section is to recall the notion of implied volatility surface and explain the construction problem of such surface under the no-arbitrage assumptions. For this purpose, we discuss it through the standard model of Black-Scholes. More details can be founded in [JYC09, subsection 2.3], [Shr04, subsection 4.5], [Kar97, section 1.2].

2.1 Black-Scholes model

We consider an European option on a given asset S, with maturity T and strike price X. By denoting Φ its payoff function, the value of the option at time T is given, in term of Call option, by $\Phi(S_T) = (S_T - X)^+$, where $(S_T - X)^+ = \max(S_T - X; 0)$. While, in term of put option, its value at time T is $\Phi(S_T) = \max(X - S_T; 0)$.
In [BS73] model, the risky asset price $S = (S)_{t\geq 0}$ defined on the filtered probability space $(\Omega, \mathcal{G}, \mathbb{P}, \mathbb{F})$ is assumed to follow a Geometric Brownian Motion, i.e., its dynamics are of the form :

$$dS_t = \mu S_t dt + \sigma S_t dB_t,\ \forall t \geq 0 \tag{1}$$

where $B = (B_t)_{t\geq 0}$ is a Brownian Motion defined on $(\Omega, \mathcal{G}, \mathbb{P})$, μ represents the drift of the returns of S, σ its instantaneous volatility and $S_0 > 0$ the initial stock price. The interest rate r is assumed to be constant and positive so that the discounted stock price $\hat{S}$ has the following dynamics under the probability $\mathbb{P}$:

$$d\hat{S}_t = (\mu - r)\hat{S}_t dt + \sigma \hat{S}_t dB_t,\ \forall t \geq 0\,.$$

One can show from the Girsanov theorem that there exists a (unique) measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, given by $\mathbb{Q}|_{\mathcal{F}_t} = e^{-\theta B_t - \frac{1}{2}\theta^2 t}\mathbb{P}|_{\mathcal{F}_t}$, with $\theta = \frac{\mu - r}{\sigma}$ [JYC09, Proposition 2.3.1.1] and under which $\hat{S}$ is an $\mathbb{F}$-martingale. The fundamental theorem of asset pricing provides that this existence of such a $\mathbb{Q}$ is equivalent to the absence of arbitrage opportunity of this market model. Therefore, under the arbitrage-free conditions, the value at time t, of the European option is the expected present value under the risk neutral measure $\mathbb{Q}$ of future payoffs, which has the following expression:

$$V_t(S_t, t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}\Phi(S_T)|\mathcal{F}_t\right]. \tag{2}$$

The computation of the expression (2) leads to the famous Black-Scholes formula.
For instance, the value of a put option on S, denoted by P_t^{BS} is given by

$$P_t^{BS}(S_t, T, X, \sigma, r) = Xe^{-r(T-t)}N(-d_2) - S_t N(-d_1), \tag{3}$$

where

$$d_1 = \frac{\log(S_t/X) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$
$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

and N is the cumulative normal density function.

2.2 Implied volatility surface

The equality (3) links the European put option price to the underlying price S, the interest rate r, the volatility of the stock σ, the strike price X and the maturity T. All these parameters except the volatility σ may be directly observed in the market. Actually, the option price is a function of the volatility of the underlying asset and then, inverting it leads to the so called implied volatility, which is of crucial interest for risk management and exotic option pricing models. This means that for a given observed market option price P_t^{Market}, quoted at the date t, with corresponding maturity T and Strike X, the implied volatility is the solution value $\sigma_t^{IV}(T,X)$ of

$$P_t^{BS}(S_t, T, X, \sigma_t^{IV}(T,X), r) = P_t^{Market}. \quad (4)$$

Accordingly, our interpolation problem consists in computing at any time t, the whole surface $(T,X) \to P_t^{BS}(T,X)$ of option prices which allows us to derive the implied volatility surface $(T,X) \to \sigma_t^{IV}(T,X)$ by using the inversion problem. However, this construction must take into account some characteristics such as :

- **Incomplete information** : the option price surface is only known or can only be estimated for few input locations (the observed couples Strike-maturity).
- **Noisy measurement** : observed prices may not be fully reliable (ex : price of illiquid instruments).
- **Smoothness constraints** : the price surface should be differentiable (important for deriving the local volatility surface).
- **Shape constraints** : the price surface should not allow to generate arbitrage.

In this regard, we propose to use the kriging techniques for such construction. We discuss our methodology that requires the transition from classical kriging, which does not allow to obtain arbitrage-free option price surfaces, to the constrained kriging which does allow to get free-arbitrage option price surfaces. We present these techniques in the next section.

In what follows, we only consider Put option, Call option price can be recovered using the call-put parity. Before going further, we recall how no-arbitrage conditions translate into shape constraints on the put option price.

Proposition 1 *We place ourselves at a fixed date of evolution t_0 and we consider S_0 the value of the underlying at t_0. The put price surface $(T,X) \to P(T,X)$ is free of static arbitrage if and only if*

(i) $X \to P(T,X)$ *is a convex function such that* $P(T,0) = 0$ *and* $\frac{\partial P}{\partial X}(T,0) = 0$*, for any* $T \geq 0$.

(ii) $T \to P(T,X)$ *is a non-decreasing function, for any* $X \geq 0$.

(iii) $\lim\limits_{X\to\infty} P(T,X) = X - S_0$.

(iv) $P(0,X) = (X - S_0)^+$.

PROOF: This follows by using the call-put parity formula in [Rop10, Theorem 2.1].

3 Kriging for learning arbitrage-free put option price surfaces

In this suebsection, we present the classical kriging techniques which allow to construct option price surface by only incorporating the three first characteristics mentioned above (i.e, incomplete information, noisy measurement and smoothness constraints).

Given the input domain $\mathcal{D}$ in time and space, we aim at constructing, at a given quotation date, put price surface

$$P : \mathcal{D} \to \mathbb{R}^+$$
$$(T,X) \mapsto P(T,X)$$

satisfying arbitrage-free conditions given in Proposition 1 and given n noisy observations $\boldsymbol{y} = [y_1, ..., y_n]^\top$ of function P at input points $\mathbf{X} = [\mathbf{X}_1, ..., \mathbf{X}_n]$ where $\mathbf{X}_i = (T_i, X_i)$, $i = 1, \dots, n$. Then, this construction should be compatible with market fit condition

$$\boldsymbol{y} = P(\mathbf{X}) + \boldsymbol{\varepsilon} \quad (1)$$

where $P(\mathbf{X}) := [P(\mathbf{X}_1), ..., P(\mathbf{X}_n)]^\top$. The additive noise term $\boldsymbol{\varepsilon} = [\varepsilon_1, ..., \varepsilon_n]^T$ is assumed to be a zero-mean Gaussian vector, independent from P, and with an homoscedastic covariance matrix given as $\varsigma^2 I_n$, where I_n is the identity matrix of size n.

Comments 2 *As mentioned in [CMR16], this framework which takes account of the presence of error noise being quite considerable inasmuch as it allows to construct implied volatility surface in presence of illiquid options. Accordingly a best way to incorporate some noises in the response variable is to investigate the main sources of these noises. In [Hen03] three kinds of sources of measurement errors have been referred in the option prices such as the finite quote precision which is based on the tick price that represents the minimum increment between bid and ask prices of an asset in the trading system which orients the prices movement in a discrete setting. However the real market prices move in a continuous way and this discrete increment should be a source of noisy observations. In addition to the finite quote precision, Hentschel invoques the non-synchronous prices and the bid-ask spread which is the difference between the ask (which represents the supply for a particular asset) and the bid (the demand for asset). The higher the bid-ask spread of an option, the more liquid this option becomes.*

Two approaches for choosing our observations and incorporating noises might be taken into account, the first one consists in considering the realizations of the Gaussian process to be the mid-point prices and assuming the noise variance term ς to be proportional to the magnitudes of the bid-ask spreads. The second one consists in considering both the bid and ask prices as two independent realizations of the Gaussian process and estimate ς by using an appropriate method of parameter estimation. This last approach allows to obtain a option price surface which lies between the free-arbitrage bid and ask surfaces.

3.1 Classical GPs regression or kriging

We consider a zero-mean Gaussian process prior on the mapping $P = P(T, X)_{(T,X)\in\mathcal{D}}$ with covariance function (kernel function) $\mathcal{K}$. Then, the output vector $P(\mathbf{X})$ has a normal distribution with zero mean and covariance matrix $\boldsymbol{C}$ with components $\mathrm{cov}(P(T_i, X_i), P(T_j, X_j)) = \mathcal{K}((T_i, X_i), (T_j, X_j))$. We consider a 2-dimensional isotropic covariance kernel given as a tensor product, i.e., for $\mathbf{x} = (T, X)$ and $\mathbf{x}' = (T', X')$ two elements of $\mathcal{D}$,

$$\mathcal{K}(\mathbf{x}, \mathbf{x}') = \sigma^2 R_T(T - T', \theta_T) R_X(X - X', \theta_X)$$

where $\theta = (\theta_T, \theta_X)$ and σ^2 correspond to the length scale and the variance hyper-parameters of the kernel function $\mathcal{K}$ and the functions R_T and R_X are kernel correlation functions.
It is well known that the conditional process $P \mid \boldsymbol{y} = P(\mathbf{X}) + \boldsymbol{\varepsilon}$ is Gaussian with mean function η and covariance function $\mathcal{K}^\star$ given respectively by (see [WR06]):

$$\eta(\mathbf{x}) = \mathbf{c}(\mathbf{x})^\top (\boldsymbol{C} + \varsigma^2 I_n)^{-1} \boldsymbol{y}, \quad \mathbf{x} = (T, X) \in \mathcal{D} \tag{2}$$

$$\mathcal{K}^\star(\mathbf{x}, \mathbf{x}') = \mathcal{K}(\mathbf{x}, \mathbf{x}') - \mathbf{c}(\mathbf{x})^\top (\boldsymbol{C} + \varsigma^2 I_n)^{-1} \mathbf{c}(\mathbf{x}'), \quad \mathbf{x}, \mathbf{x}' \in \mathcal{D} \tag{3}$$

where $\mathbf{c}(\mathbf{x}) = [\mathcal{K}(\mathbf{x}, (T_1, X_1)), ..., \mathcal{K}(\mathbf{x}, (T_n, X_n))]^\top$.

Without considering arbitrage-free conditions as described in Proposition 1, estimation of the price function P under this framework is known as classical GPs regression or classical kriging. In this setting, prediction and uncertainty quantification is made using the conditional distribution $P \mid \boldsymbol{y} = P(\mathbf{X}) + \boldsymbol{\varepsilon}$. The Best Linear Unbiased Estimator (BLUE) of P is given as the kriging mean function (2). The conditional covariance function $\mathcal{K}^*$ can be used to obtain confidence bands around the predicted price surface. The hyper-parameters of the kernel function $\mathcal{K}$ as well as the variance of the noise can be estimated using the maximum likelihood estimator (MLE) (see, e.g., [BLLL18]).

3.2 Imposing the no-arbitrage conditions

Conditionally to the market fit condition (1) and conditions (iii) and (iv) of Proposition 1, P is still Gaussian. However, conditionally to the inequality constraints (i.e, the monotonicity of the put price with respect to the maturities direction and its convexity with respect to the strike prices) by means (i) and (ii) of Proposition 1, the process P is no longer Gaussian and this issue obviously run across the difficulties of simulating the posterior process in the sense that the range of constraint check points is usually infinite-dimensional in the simulation. We adopt the solution of [CMR16] that consists in constructing a finite-dimensional approximation P^N of the Gaussian prior P for which the constraints can be checked in the entire domain $\mathcal{D}$ with a finite number of checks.
We first consider a discretized version of the input space $\mathcal{D}$ as a $N = (N_T + 1) \times (N_X + 1)$ regular grid with knots (u_i, v_j), $i = 1, ..., N_T$, $j = 1, ..., N_X$ with $u_i = i\delta_T$ and $v_j = j\delta_X$, where $\delta_T = \frac{1}{N_T}$ and $\delta_X = \frac{1}{N_X}$. For each knot (u_i, v_j), we introduce the hat basis function defined as the following tensor product

$$\phi_{i,j}(T, X) := \max(1 - \frac{|T - u_i|}{\delta_T}, 0) \max(1 - \frac{|X - v_j|}{\delta_X}, 0) \ .$$

Then, the process P is approximated on $\mathcal{D}$ by the process P^N given by

$$P^N(T, X) = \sum_{i=0}^{N_T} \sum_{j=0}^{N_X} P(u_i, v_j)\phi_{i,j}(T, X), \quad \text{for all } (T, X) \in \mathcal{D} \tag{4}$$

which is a piecewise linear interpolation of P at knots $(u_i, v_j)_{i,j}$. If we denote $\xi_{i,j} = P(u_i, v_j)$, for $i = 1, ..., N_T$, $j = 1, ..., N_X$, then $\boldsymbol{\xi} = [\xi_{0,0}, ..., \xi_{i,j} ..., \xi_{N_T,N_X}]^\top$ is a zero-mean Gaussian vector with $N \times N$ covariance matrix Γ^N such that $\Gamma^h_{\imath_1,\imath_2} = \mathcal{K}((u_{i_1}, v_{j_1}), (u_{i_2}, v_{j_2}))$, for any two grid index pairs (i_1, j_1) and (i_2, j_2) corresponding to global indices $\imath_1$ and $\imath_2$ respectively and $\boldsymbol{\phi}(T, X)$ a vector of size N given by

$$\boldsymbol{\phi}(T, X) = [\phi_{0,0}(T, X), ..., \phi_{i,j}(T, X), ..., \phi_{N_T,N_X}(T, X)].$$

The equality (4) can be written in the following matrix form

$$P^N(T, X) = \boldsymbol{\phi}(T, X) \cdot \boldsymbol{\xi}$$

so that when denoting $\boldsymbol{\Phi}(\mathbf{X})$ the $n \times N$ matrix of basis function in which, each row l corresponds to the vector $\boldsymbol{\phi}(T_l, X_l)$, one has $P^N(\mathbf{X}) = \boldsymbol{\Phi}(\mathbf{X}) \cdot \boldsymbol{\xi}$, with $P^N(\mathbf{X}) := [P^N(\mathbf{X}_1), ..., P^N(\mathbf{X}_n)]^\top$.

In what follows we use the shorthand $\boldsymbol{\Phi}(\mathbf{X}) = \boldsymbol{\Phi}$.

Proposition 3 *(see [Maa17])*
The following statements hold.

- *The finite-dimensional process P^N uniformly converges to P on $\mathcal{D}$ as $N_X \to \infty$ and $N_T \to \infty$, almost surely,*
- *$P^N(T, X)$ is non-decreasing function of T if and only if $\xi_{i+1,j} \geq \xi_{i,j}$,*
- *$P^N(T, X)$ is a convex function of X if and only if $\xi_{i,j+2} - \xi_{i,j+1} \geq \xi_{i,j+1} - \xi_{i,j}$.*

Given the first statement of Proposition 3, by denoting $\mathcal{M}$ the convex set of inequality constraints, i.e., $\mathcal{M}$ is the set of 2-d continuous functions which are no-decreasing w.r.t. T and convex w.r.t. the X, our construction problem (we denote it $(\mathcal{P})$) consists in finding the conditional distribution of P^N given

$$\begin{cases} \boldsymbol{y} = P^N(\mathbf{X}) + \boldsymbol{\varepsilon} \\ P^N \in \mathcal{M}. \end{cases}$$

The last two statements of Proposition 3 justify the choice of the hat basis functions which is due to the fact that P^N satisfies the inequality constraints in the entire domain $\mathcal{D}$ when it satisfies these constraints at the knots (see [MB14]), i.e., $P^N \in \mathcal{M}$ if and only if $\boldsymbol{\xi} \in \mathcal{C}_{ineq}$ where $\mathcal{C}_{ineq}$ is a set of linear inequality constraints on $\boldsymbol{\xi}$ as given by the two last points of Proposition 3. In this line of thinking, our construction problem $(\mathcal{P})$ is equivalent to estimate $\boldsymbol{\xi}$ restricted to

$$\begin{cases} \boldsymbol{y} = \boldsymbol{\Phi} \cdot \boldsymbol{\xi} + \boldsymbol{\varepsilon} \\ \boldsymbol{\xi} \in \mathcal{C}_{ineq}. \end{cases}$$

3.3 Hyper-parameter learning

Hyper-parameters, namely the length scale θ and the variance parameter σ, of the kernel function $\mathcal{K}$ as well as the noise parameter ς can be either specified or estimated. By denoting λ the set of these parameters (i.e., $\lambda = [\theta, \sigma, \varsigma]^\top$), we propose to maximize the marginal log likelihood $\mathcal{L}(\lambda)$ for the process P^N w.r.t. λ for parameters learning.
Under the finite dimensional approximation, the marginal log likelihood can be expressed as (see, e.g., [WR06]):

$$\mathcal{L}(\lambda) = -\frac{1}{2}\boldsymbol{y}^\top(\boldsymbol{\Phi}\Gamma^N\boldsymbol{\Phi}^\top + \varsigma^2 I_n)^{-1}\boldsymbol{y} - \frac{1}{2}\log(\det(\boldsymbol{\Phi}\Gamma^N\boldsymbol{\Phi}^\top + \varsigma^2 I_n)) - \frac{n}{2}\log(2\pi). \tag{5}$$

3.4 The most probable response surface and measurement noises

The MAP of P^N is given by

$$\mathbf{m}_{P^N}(T, X) := \sum_{i=0}^{N_X}\sum_{j=0}^{N_T} \boldsymbol{\nu}_{\boldsymbol{\xi}}^{i,j}\phi_{i,j}(T, X) \tag{6}$$

where $\boldsymbol{\nu}_{\boldsymbol{\xi}} = \left(\boldsymbol{\nu}_{\boldsymbol{\xi}}^{(0,0)}, \dots \boldsymbol{\nu}_{\boldsymbol{\xi}}^{(i,j)}, \cdots, \boldsymbol{\nu}_{\boldsymbol{\xi}}^{(N_x, N_t)}\right)^\top$ is the MAP of the Gaussian coefficients $\boldsymbol{\xi}$. The MAP $\boldsymbol{\nu}_{\boldsymbol{\xi}}$ of $\boldsymbol{\xi}$ is paramount in the sampling of $\boldsymbol{\xi}$ since it satisfies the inequality constraints hence it can be considered as the initial vector when using the sampling algorithm of HMC (see [LL+18]). This can be obtained directly by maximizing the density function of the conditional Gaussian vector $\boldsymbol{\xi}|\boldsymbol{\xi} \in C_{ineq}$ restricted to C_{ineq} which implies that

$$\boldsymbol{\nu}_{\boldsymbol{\xi}} = \arg\min_{\boldsymbol{\vartheta} \in C_{ineq}} (\boldsymbol{\vartheta}^\top \Sigma_{\text{cond}}^{-1}\boldsymbol{\vartheta} - \boldsymbol{\vartheta}^\top\Sigma_{\text{cond}}^{-1}\mu_{\text{cond}}) \tag{7}$$

where μ_{cond} and Σ_{cond} are respectively defined by (10) and (11). This is a quadratic optimization problem and it is equivalent to

$$\boldsymbol{\nu}_{\boldsymbol{\xi}} := \arg\min_{\boldsymbol{\Phi}\cdot\boldsymbol{\vartheta}+\boldsymbol{e}=\boldsymbol{Y},\, \boldsymbol{\vartheta}\in\mathcal{C}_{ineq}} (\boldsymbol{\vartheta}^\top(\Gamma^N)^{-1}\boldsymbol{\vartheta}), \tag{8}$$

with Γ^N the covariance matrix of $\boldsymbol{\xi}$.
In order to identify the locations x of the largest noises and their values, we compute the joint MAP $(\boldsymbol{\nu}_{\boldsymbol{\xi}}, \boldsymbol{\nu}_{\boldsymbol{\varepsilon}})$ of the truncated gaussian vector $\boldsymbol{\xi}$ and the MAP of the Gaussian noise vector $\boldsymbol{\varepsilon}$. This can be defined as solution of

$$\max_{\boldsymbol{\vartheta},\boldsymbol{e}} \mathbb{P}\left(\boldsymbol{\xi} \in [\boldsymbol{\vartheta}, \boldsymbol{\vartheta} + d\boldsymbol{\vartheta}], \boldsymbol{\varepsilon} \in [\boldsymbol{e}, \boldsymbol{e} + d\boldsymbol{e}] \mid \boldsymbol{\Phi}\cdot\boldsymbol{\xi} + \boldsymbol{\varepsilon} = \tilde{\boldsymbol{v}},\, \boldsymbol{\xi} \in \mathcal{C}_{ineq}\right).$$

As $(\boldsymbol{\xi}, \boldsymbol{\nu}_{\boldsymbol{\varepsilon}})$ is Gaussian centered with block-diagonal covariance matrix with blocks Γ^N and $\varsigma^2 I_n$, this implies that the mode $(\boldsymbol{\nu}_{\boldsymbol{\xi}}, \boldsymbol{\nu}_{\boldsymbol{\varepsilon}})$ is a solution to the following quadratic problem

$$\min_{\boldsymbol{\Phi}\cdot\boldsymbol{\vartheta}+\boldsymbol{e}=\tilde{\boldsymbol{v}},\, \boldsymbol{\vartheta}\in\mathcal{C}_{ineq}} \left(\boldsymbol{\vartheta}^\top(\Gamma^N)^{-1}\boldsymbol{\vartheta} + \boldsymbol{e}^\top(\varsigma^2 I_n)^{-1}\boldsymbol{e}\right). \tag{9}$$

As a consequence, we define the most probable measurement noises to be $\boldsymbol{\nu}_{\boldsymbol{\varepsilon}}$ and the most probable response surface $\mathbf{m}_{P^N}$ given by $\mathbf{m}_{P^N} := \boldsymbol{\Phi}\cdot\boldsymbol{\nu}_{\boldsymbol{\xi}}$. Distance to the data can be an effect of arbitrage opportunities within the data and/or the misspecification / lack of expressiveness of the kernel.

3.5 Sampling finite-dimensional GPs with shape constraints

As we mentioned in Subsection 3.2, the construction of the put price surface consists in sampling $\boldsymbol{\xi}$ truncated on C_{ineq}. Known that the distribution of $\boldsymbol{\xi}$ given $\boldsymbol{y} = \boldsymbol{\Phi}\cdot\boldsymbol{\xi} + \boldsymbol{\varepsilon}$ is multivariate Normal $\mathcal{N}(\boldsymbol{\mu}_{cond}, \Sigma_{cond})$ (see [WR06]) where

$$\mu_{\text{cond}} = \Gamma^N\boldsymbol{\Phi}^\top(\boldsymbol{\Phi}\Gamma^N\boldsymbol{\Phi}^\top + \varsigma^2 I_n)^{-1}\boldsymbol{y} \tag{10}$$

and

$$\Sigma_{\text{cond}} = \Gamma^N \mathbf{\Phi}^\top (\mathbf{\Phi}\Gamma^N \mathbf{\Phi}^\top + \varsigma^2 I_n)^{-1} \mathbf{\Phi}\Gamma^N. \tag{11}$$

Hence we are face to a problem of sampling from truncated multivariate Gaussian distribution, which we do by Hamiltonian Monte Carlo (see [LL+18]), using the maximum a posteriori probability estimate (MAP) of $\boldsymbol{\xi}$ as initial vector (which must verify the constraints) in the algorithm.

4 Numerical illustrations

The aim of this empirical investigation is to illustrate the construction methods introduced in the previous sections using real financial data. In particular, we construct the put prices surface from both classical and constrained kriging in such a way to use it for deriving the implied volatility surface from the inversion techniques. Our study turns out to be very interesting since the results would certainly be useful for no iliquid options pricing. It would also be possible through this study to understand the option prices consistency under bid ask spreads. In addition, a useful case study could be to investigate for instance, the liquidity measure which is one of the most important tools of a financial markets and characterizes the ability of market makers to execute trades at the determined market prices and for a large volume without affecting the stock price.
The illustrations are carried out on the Euro Stoxx 50. We present an approach which consists in constructing the put prices surface through the mid price by estimating the standard deviation of the noise by MLE.

Remark 4 *One can also construct the whole put prices surface through the mid price by defining the standard deviation of the noise as the difference between the ask and mid prices. Another approach consists in considering both the bid and ask put prices as independent responses of a zero-mean Gaussian process Y and constructing a put prices surface by estimating the attributed noise variance. As such, we can study the behavior of this surface relatively to the surfaces constructed from the bid put prices and the ask put prices.*

We observe at a particular market quotes t_0 as of January 10, 2019, a series of put option prices $f(T_i, X_i) = P(T_i, X_i)$ for different characteristics (T_i, X_i), $i = 1, \ldots, n = 1232$ with T_i the maturities in which, its range goes from January 18, 2019 to December 17, 2021 and X_i the strike prices which range from 250 to 4000. The spot price is equal to $S_{t_0} = 3070.24$. These observations are represented in Figure 1. Our goal is to construct the whole surface of the put option prices (using both classical and constrained kriging) that we use for computing the implied volatility surface by the inversion techniques.

To this end, we consider our input data which contains $n_T = 15$ maturities and $n_X = 88$ strike prices and which does not necessarily contain gridded data (as we may observe in Figure 1).

We randomly choose 5% of these data as training set. Thus the unknown function is evaluated at only 5% of the input data observed. After scaling the input space $\mathcal{D}$ into $[0,1] \times [0,1]$, we choose $N = 600$ basis functions given by $N = N_T \times N_X$ where $N_T = 30$ represents the number of nodes related to the time-to-maturities in $[0,1]$ and $N_X = 20$ the number of nodes linked to the strike prices in $[0,1]$. We consider a two dimensional Gaussian covariance function defined as

$$\mathcal{K}(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\left(-\frac{(X - X')^2}{\theta_X^2} - \frac{(T - T')^2}{\theta_T^2}\right)$$

where the components of the two vectors $\mathbf{x} = (T, X)$ and $\mathbf{x}' = (T', X')$ represent respectively the vectors of strike prices and time to maturities.

The GP hyperparameters $\theta_X, \theta_T, \sigma$ and the standard deviation of the noise ς are estimated using the MLE. Figure 2 represents their stability relatively to the chosen number of basis functions. Here, we study the convergence of the

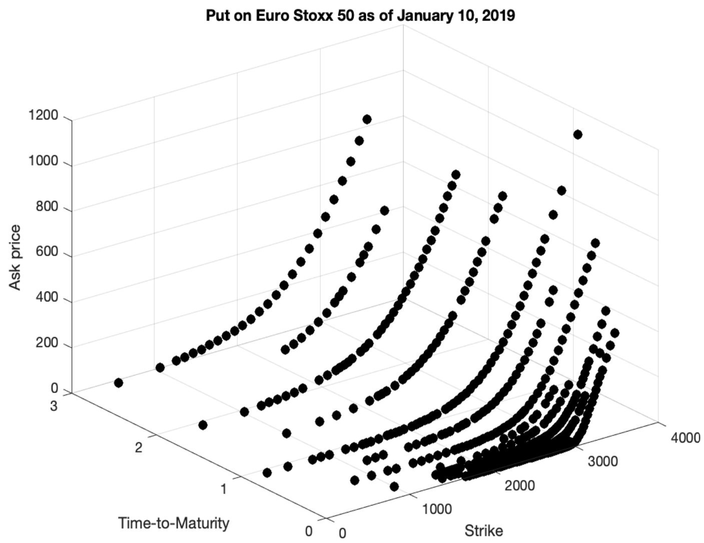

Figure 1: Import input observed data

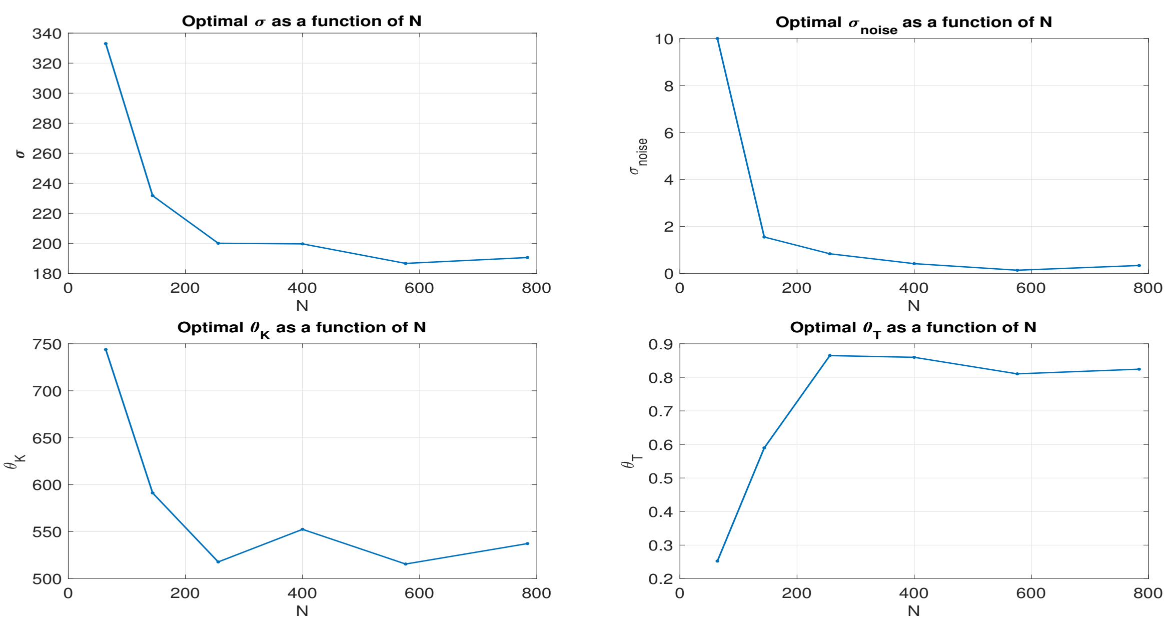

Figure 2: Convergence of optimal parameter as a function of N (number of basis functions).

parameters by increasing the number of basis function. It is clear that this convergence is achieved among from 250 basis functions. In short, parameters convergence is reached when the number of basis functions tends to infinity.

Then, as explained in Subsection 3.4, quadratic programming is used to find the most probable response surface and measurement noises, while enforcing the constraints, using the `interior-point-convex` iterative algorithm with a tolerance of 1×10^{-12}.

In Figure 3, we represent the most probable response surface (left) and measurement noises (right). It is clear that the surface of put prices respects the arbitrage-free conditions. The scattered noisy points show that low maturity options are associated with a significant error and therefore a high distance from the most probable non-arbitrable surface.

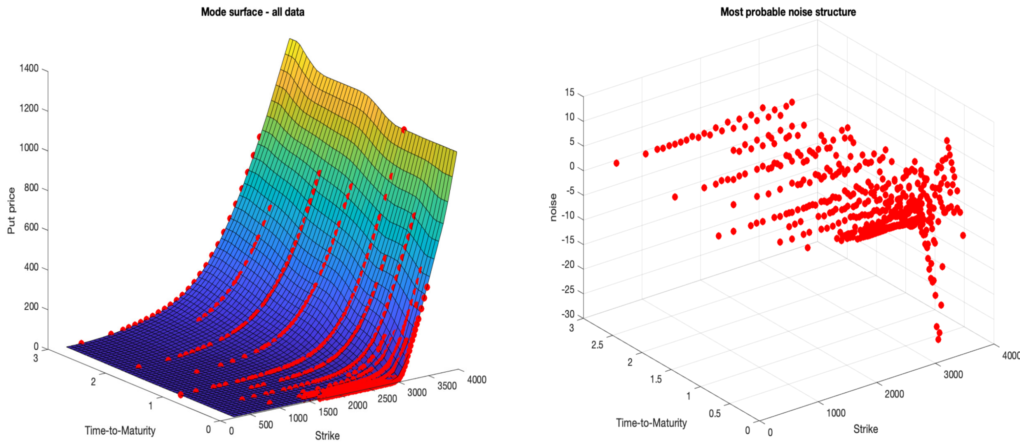

Figure 3: Most probable surface (left) vs most probable noise values (right).

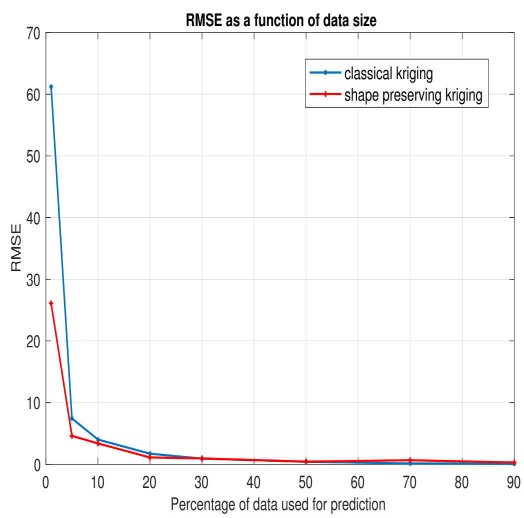

Figure 4: Mode estimator - prediction accuracy

Figure 4, illustrates the prediction accuracy of the mode estimator (MAP) which has been carried out by the following way:

- We first construct a series of randomly chosen data subsets with increasing number of points,
- We apply classical kriging and shape-preserving kriging on these subsets by computing the two mode estimators,
- For each data size, we compute the Root Mean Square Error (RMSE) with respect to the original data set.

The results show that for a small percent of data size (less than 30 %), the MAP estimated from the constrained kriging outperforms the one estimated from the unconstrained kriging. However, when more than 30% of data points are used for training, the two estimated MAP present a similar accuracy with their RMSE which tend to zero.

We now generate 5000 paths from both the classical and the constrained kriging using the Exact Hamiltonian Monte Carlo algorithm (see [PP14]). Then we compute the whole implied volatility surface from the inversion of the Black and Scholes formula as described in Subsection 2.1 with the spot price S_{t_0} and interest rate $r = 1\%$. Figure 5 presents comparison results on the MAP estimator between classical and constrained kriging. One can see that the put prices surface constructed from constrained kriging (in left) clearly verifies the no-arbitrage constraints while the one constructed using classical kriging does not fulfill these constraints. Also Figure 6 represents the obtained implied volatility surface which shows that the constrained kriging allows to obtain some surfaces which are more smooth than the ones obtained in classical kriging. In Figure 7 which shows 5% and 95% estimated quantiles of the fitted GPs, one can see a large confidence interval in the case of classical kriging where there is no observation while this confidence interval becomes more restricted in the constrained kriging, due to the knowledge of the no-arbitrage conditions.

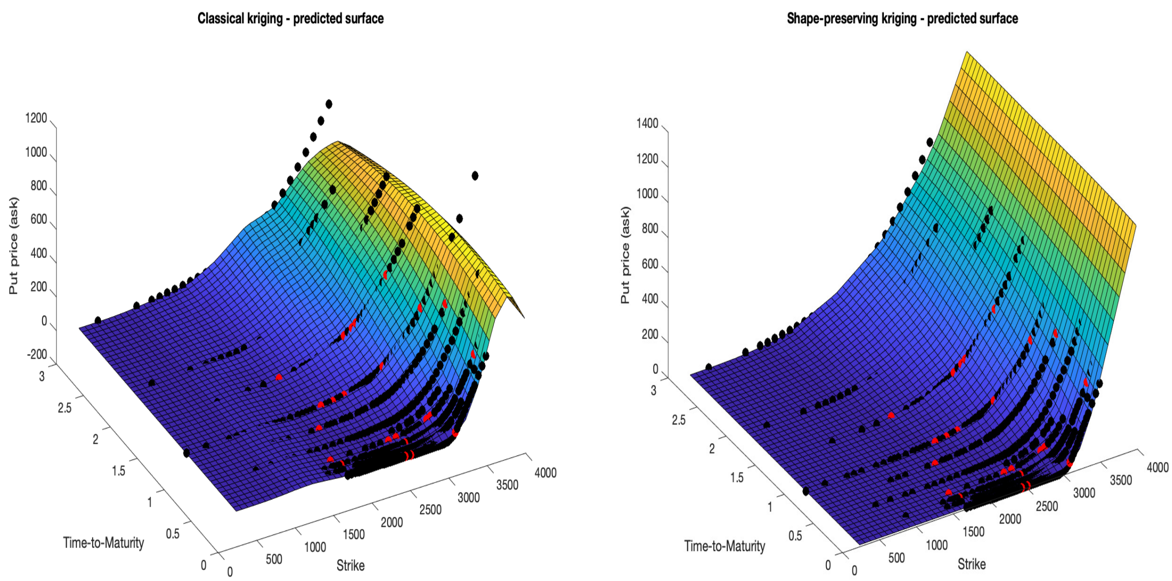

Figure 5: Put prices surface constructed from classical kriging (left) vs put prices surface constructed from constrained kriging (right). The red points represent the 5% of the data used for training.

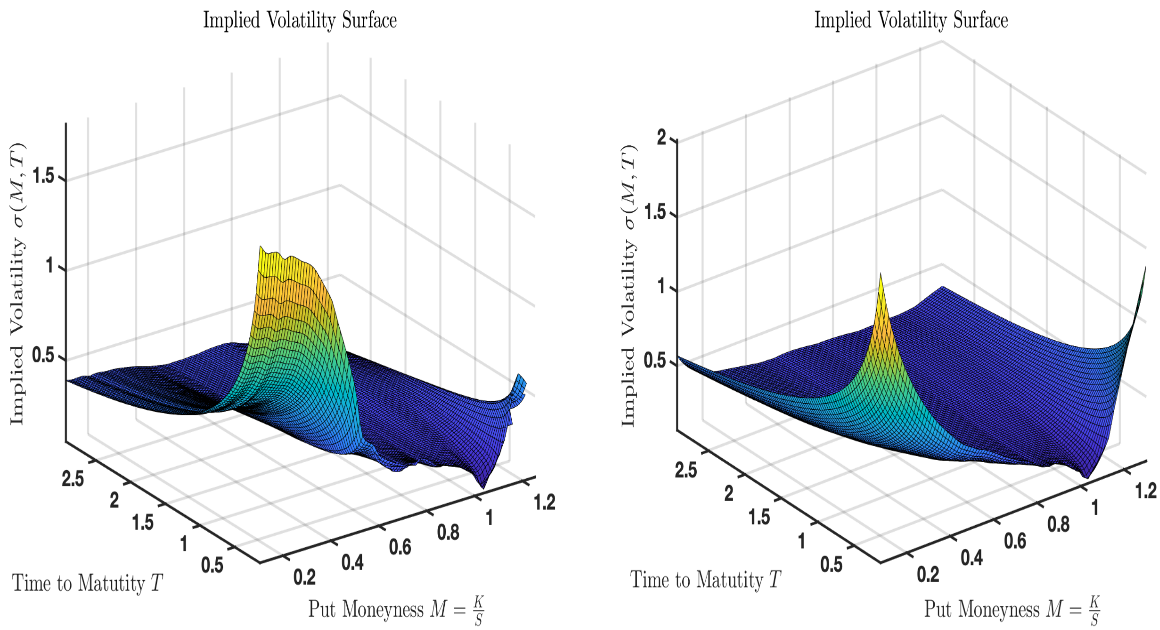

Figure 6: Implied volatility surface obtained from classical kriging (left) vs Implied volatility surface obtained from constrained kriging (right). Implied volatility surface obtained from constrained kriging is more smoother.

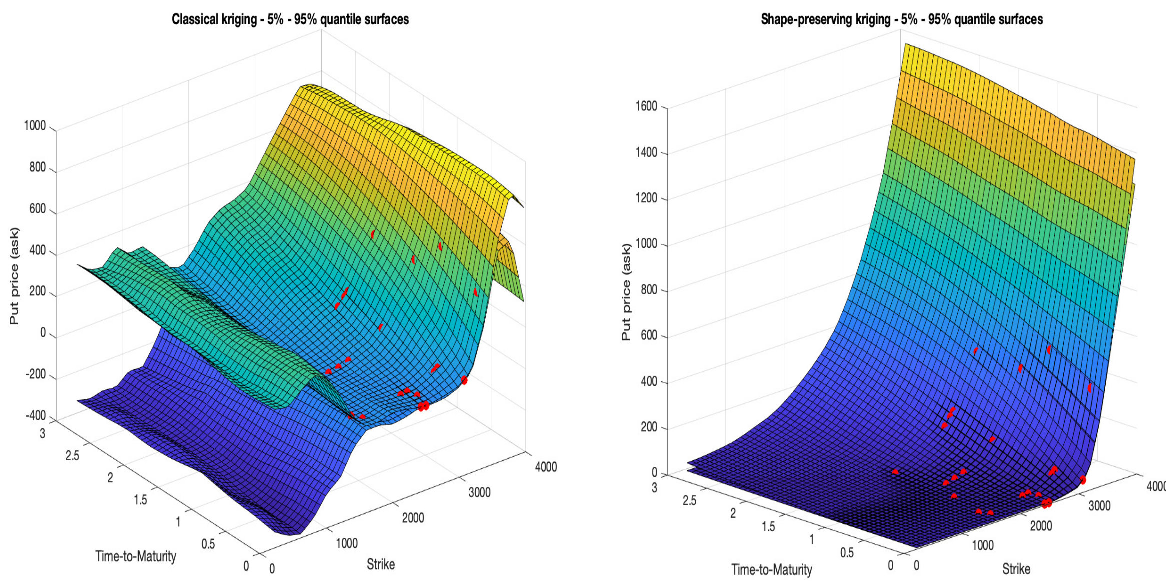

Figure 7: 5% and 95% estimated pointwise quantiles of the fitted GPs with classical kriging (left) vs constrained kriging (right).

In Figure 8, we present a 5% and 95% estimated pointwise quantiles of the constrained fitted Gaussian process with extrapolation in the time-to-maturities direction by adding two years. Unsurprisingly, this leads to an increase of the confidence interval due to the fact that no price is observed for maturities greater than two years.

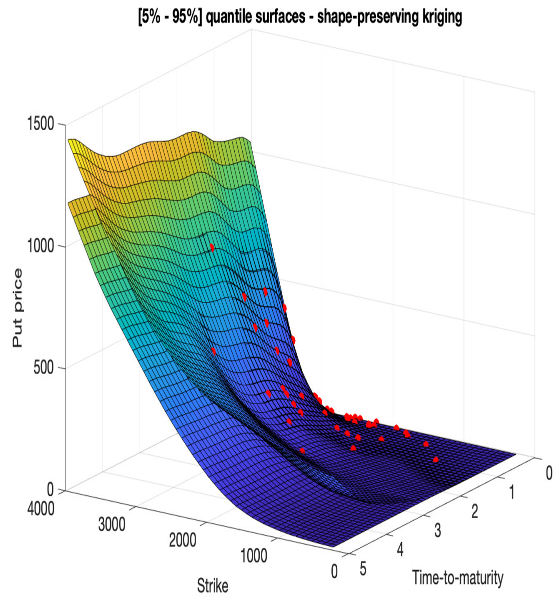

Figure 8: 5% and 95% estimated pointwise quantiles of the constrained fitted GPs with extrapolation in the time-to-maturities direction (adding 2 years).

5 Extentions of Kriging techniques in the case with interest rate and dividend

Given a stock S we consider a European vanilla put option on S, with maturity T and strike price X. We consider the following risk neutral dynamics of S

$$dS_t = S_t[(r(t) - q(t))dt + \sigma(t, S_t)dB_t],\ \forall t \geq 0 \tag{1}$$

where $B = (B_t)_{t\geq 0}$ is a Brownian Motion, $r(t)$ represents a deterministic short interest rate term structure of the corresponding economy and $q(t)$ a deterministic continuous-dividend-yields on S and $\sigma(t, S_t)$ the instantaneous volatility function of S and t. We denote respectively by $D_t = \exp\left(-\int_0^t r(s)ds\right)$ and $H_t = \exp\left(-\int_0^t q(s)ds\right)$ the discount and dividend factors.

Without any restriction, we only pay particular attention to the put prices constraints since the the call prices constraints can be derived from the Call-Put parity. As such, let us denote by $P(T, X)$ and $C(T, X)$, the prices at time $t = 0$ of, respectively, the European put and the European Call options on the underlying asset S with maturity T and strike price X. We recall the Call-Put parity which can be expressed as

$$C(T, X) - P(T, X) = S_0 H(T) - X D(T) \tag{2}$$

Once, we know $C(T, X) - P(T, X)$ for $T = T_1, ..., T_j$, with $j \in \mathbb{N}$, we can regress $C(T, X) - P(T, X)$ on the direction of X, for $X = X_1, ..., X_{n_j}$, with $n_j \in \mathbb{N}$ in order to obtain an approximation of $H(T_j)$ and $D(T_j)$. One further constructs $H(T)$ and $D(T)$, for any T by interpolation.

Definition 5 *Under the arbitrage-free conditions, the transformed process $M = (M_t)_{t\geq 0}$, given by*

$$M_t = \exp\left(-\int_0^t (r(s) - q(s))ds\right) S_t,$$

is a $(\mathbb{Q}, \mathbb{F})$-martingale (see [Shr04, Subsection 5.5.1]) and the price at time $t = 0$ of the Put option is given by

$$P(T, X) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^T r(s)ds\right)(X - S_T)^+\right]. \tag{3}$$

The equality (3) can be expressed as

$$P(T,X) = H_T \mathbb{E}^{\mathbb{Q}}\left[\left(\frac{D_T}{H_T}X - M_T\right)^+\right] \tag{4}$$

with

$$M_T = \frac{D_T}{H_T} S_T.$$

By setting

$$k = \frac{D_T}{H_T} X, \tag{5}$$

one has, from (4),

$$P(T,X) = H_T \widehat{P}(T,k) \tag{6}$$

where

$$\widehat{P}(T,k) = \mathbb{E}^{\mathbb{Q}}\left[(k - M_T)^+\right]. \tag{7}$$

Proposition 6 *In presence of dividend and interest rate, a collection of Put prices $P(T,X)_{T\geq 0, X\geq 0}$ is arbitrage-free if and only if the process $\widehat{P}(T,k)_{T,k}$ given by (7) verifies the following conditions*

1. *$\widehat{P}(T,.)$ is a convexe function, for any $T \geq 0$,*
2. *$\widehat{P}(.,k)$ is an non-decreasing function, for any $k \geq 0$,*
3. *$\widehat{P}(T,0) = 0, \forall\, T \geq 0$,*
4. *$\frac{\partial}{\partial k}\widehat{P}(T,0) = 0, \forall\, T \geq 0$,*
5. *$\lim\limits_{k \longrightarrow \infty} \frac{\partial}{\partial k}\widehat{P}(T,k) = k - M_0$,*
6. *$\widehat{P}(0,k) = (k - M_0)^+,\ \forall k \geq 0$.*

Accordingly, learning the put price function $P(\cdot,\cdot)$ is equivalent to learning the reduced price$\widehat{P}$ given the observed put prices and Proposition 6 and applying (6) where the q can be extracted from the forward contracts on the stock S.

6 Conclusion

In this paper we have adopted the constrained kriging techniques in the aim of constructing implied volatility surface. We have shown that constrained kriging techniques allow to estimate the option prices in a market where only noisy prices are available. Our approach incorporates all the characteristics of surface construction problem, such as incomplete information, indirect observation, noisy measurement and shape constraints. The most likely measurement noises have been computed. We have shown, through a comparison study, that constrained kriging is more adapted for implied volatility surface construction than the classical one.

References

[ATV19] Damien Ackerer, Natasa Tagasovska, and Thibault Vatter. "Deep smoothing of the implied volatility surface". In: *Available at SSRN 3402942* (2019).

[Agr19] Christian Agrell. "Gaussian processes with linear operator inequality constraints". In: *arXiv preprint arXiv:1901.03134* (2019).

[BLLL18] François Bachoc, Agnès Lagnoux, and Andrés F López-Lopera. "Maximum likelihood estimation for Gaussian processes under inequality constraints". In: *arXiv preprint arXiv:1804.03378* (2018).

[BJ+16] Mathias Beiglböck, Nicolas Juillet, et al. "On a problem of optimal transport under marginal martingale constraints". In: *The Annals of Probability* 44.1 (2016), pp. 42–106.

[BS73] Fischer Black and Myron Scholes. "The pricing of options and corporate liabilities". In: *Journal of political economy* 81.3 (1973), pp. 637–654.

[CCD20] Marc Chataigner, Stéphane Crépey, and Matthew Dixon. "Deep Local Volatility". In: *Risks* 8.3 (2020), p. 82. URL: http://dx.doi.org/10.3390/risks80300.

[CMR16] Areski Cousin, Hassan Maatouk, and Didier Rullière. "Kriging of financial term-structures". In: *European J. Oper. Res.* 255.2 (2016), pp. 631–648. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2016.05.057. URL: https://doi.org/10.1016/j.ejor.2016.05.057.

[Cre90] Noel Cressie. "The origins of kriging". In: *Math. Geol.* 22.3 (1990), pp. 239–252. ISSN: 0882-8121. DOI: 10.1007/BF00889887. URL: https://doi.org/10.1007/BF00889887.

[DVM12] Sébastien Da Veiga and Amandine Marrel. "Gaussian process modeling with inequality constraints". In: *Annales de la Faculté des sciences de Toulouse: Mathématiques*. Vol. 21. 3. 2012, pp. 529–555.

[DS+18] Jan De Spiegeleer, Dilip B Madan, Sofie Reyners, and Wim Schoutens. "Machine learning for quantitative finance: fast derivative pricing, hedging and fitting". In: *Quantitative Finance* 18.10 (2018), pp. 1635–1643.

[DC18] Matthew F Dixon and Stéphane Crépey. "Multivariate Gaussian Process Regression for Derivative Portfolio Modeling: Application to CVA". In: (2018).

[Dug+09] Charles Dugas, Yoshua Bengio, François Bélisle, Claude Nadeau, and René Garcia. "Incorporating functional knowledge in neural networks". In: *Journal of Machine Learning Research* 10.Jun (2009), pp. 1239–1262.

[Fen09] Matthias R Fengler. "Arbitrage-free smoothing of the implied volatility surface". In: *Quantitative Finance* 9.4 (2009), pp. 417–428.

[Gat04] Jim Gatheral. "A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives". In: *Presentation at Global Derivatives & Risk Management, Madrid* (2004), p. 0.

[Gat11] Jim Gatheral. *The volatility surface: a practitioner's guide*. Vol. 357. John Wiley & Sons, 2011.

[GJ14] Jim Gatheral and Antoine Jacquier. "Arbitrage-free SVI volatility surfaces". In: *Quantitative Finance* 14.1 (2014), pp. 59–71.

[Gol+15] Shirin Golchi, Derek R Bingham, Hugh Chipman, and David A Campbell. "Monotone emulation of computer experiments". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 370–392.

[Gon+19] Joan Gonzalvez, Edmond Lezmi, Thierry Roncalli, and Jiali Xu. "Financial applications of Gaussian processes and Bayesian optimization". In: *arXiv preprint arXiv:1903.04841* (2019).

[Hen03] Ludger Hentschel. "Errors in implied volatility estimation". In: *Journal of Financial and Quantitative analysis* 38.4 (2003), pp. 779–810.

[Hom11] Cristian Homescu. "Implied volatility surface: Construction methodologies and characteristics". In: *Available at SSRN 1882567* (2011).

[Itk19] Andrey Itkin. "Deep learning calibration of option pricing models: some pitfalls and solutions". In: *arXiv preprint arXiv:1906.03507* (2019).

[JYC09] Monique Jeanblanc, Marc Yor, and Marc Chesney. *Mathematical methods for financial markets*. Springer Science & Business Media, 2009.

[Kar97] Ioannis Karatzas. *Lectures on the Mathematics of Finance*. Vol. 8. American Mathematical Soc., 1997.

[KM82] Daniel G. Krige and Eduardo J. Magri. "Geostatistical case studies of the advantages of lognormal-de Wijsian Kriging with mean for a base metal mine and a gold mine". In: *J. Internat. Assoc. Math. Geol.* 14.6 (1982), pp. 547–555. ISSN: 0020-5958. DOI: 10.1007/BF01033878. URL: https://doi.org/10.1007/BF01033878.

[Lau11] Márcio Poletti Laurini. “Imposing no-arbitrage conditions in implied volatilities using constrained smoothing splines”. In: *Applied Stochastic Models in Business and Industry* 27.6 (2011), pp. 649–659.

[LL+18] Andrés F López-Lopera, François Bachoc, Nicolas Durrande, and Olivier Roustant. “Finite-dimensional Gaussian approximation with linear inequality constraints”. In: *SIAM/ASA Journal on Uncertainty Quantification* 6.3 (2018), pp. 1224–1255.

[Lud18] Michael Ludkovski. “Kriging metamodels and experimental design for Bermudan option pricing”. In: (2018).

[Maa17] Hassan Maatouk. “Finite-dimensional approximation of Gaussian processes with inequality constraints and errors measurements”. In: (2017).

[MB14] Hassan Maatouk and Xavier Bay. *Gaussian Process Emulators for Computer Experiments with Inequality Constraints. hal-01096751*. 2014.

[MB16] Hassan Maatouk and Xavier Bay. “A new rejection sampling method for truncated multivariate Gaussian random variables restricted to convex sets”. In: *Monte carlo and quasi-monte carlo methods*. Springer, 2016, pp. 521–530.

[MB17] Hassan Maatouk and Xavier Bay. “Gaussian process emulators for computer experiments with inequality constraints”. In: *Mathematical Geosciences* 49.5 (2017), pp. 557–582.

[Mat63] Georges Matheron. “Principles of geostatistics”. In: *Economic geology* 58.8 (1963), pp. 1246–1266.

[PP14] Ari Pakman and Liam Paninski. “Exact hamiltonian monte carlo for truncated multivariate gaussians”. In: *Journal of Computational and Graphical Statistics* 23.2 (2014), pp. 518–542.

[RV10] Jaakko Riihimäki and Aki Vehtari. “Gaussian processes with monotonicity information”. In: *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. 2010, pp. 645–652.

[Rop10] Michael Roper. “Arbitrage free implied volatility surfaces”. In: *preprint* (2010).

[Shr04] Steven E Shreve. *Stochastic calculus for finance II: Continuous-time models*. Vol. 11. Springer Science & Business Media, 2004.

[SEG12] J Beleza Sousa, Manuel L Esquível, and Raquel M Gaspar. “Machine learning Vasicek model calibration with Gaussian processes”. In: *Communications in Statistics-Simulation and Computation* 41.6 (2012), pp. 776–786.

[TR19] Martin Tegnér and Stephen Roberts. “A probabilistic approach to nonparametric local volatility”. In: *arXiv preprint arXiv:1901.06021* (2019).

[WB16] Xiaojing Wang and James O Berger. “Estimating shape constrained functions using Gaussian processes”. In: *SIAM/ASA Journal on Uncertainty Quantification* 4.1 (2016), pp. 1–25.

[WYQ04] Y Wang, H Yin, and L Qi. “No-arbitrage interpolation of the option price function and its reformulation”. In: *Journal of Optimization Theory and Applications* 120.3 (2004), pp. 627–649.

[WR06] Christopher K Williams and Carl Edward Rasmussen. “Gaussian processes for machine learning”. In: *the MIT Press* 2.3 (2006), p. 4.