

Estimation of Multivariate Smooth Functions via Convex Programs

Eunji Lim¹ & Fanny Gonzalez²

¹ College of Business and Public Administration, Kean University, Union, NJ 07083, USA

² College of Business and Public Administration, Kean University, Union, NJ 07083, USA

Correspondence: Eunji Lim, College of Business and Public Administration, Kean University, Union, NJ 07083, USA.
Tel: 1-908-737-4709. E-mail: elim@kean.edu

Received: January 13, 2017 Accepted: March 3, 2017 Online Published: March 14, 2017

doi:10.5539/ijsp.v6n3p1 URL: <https://doi.org/10.5539/ijsp.v6n3p1>

Abstract

A new method for estimating an unknown, multivariate function from noisy data is proposed in the case where the unknown function is assumed to be smooth. The proposed method finds the minimizer of the smoothness of a function while imposing an upper bound on the sum of squared errors between the function and the data set. The proposed estimator is designed to be numerically sound by eliminating the dependency on any artificially plugged-in parameters that traditional methods use and by tackling the ill-conditioned numerical settings that traditional methods suffer from. Hence, it is expected to perform better than existing estimators numerically. We prove the existence of the proposed estimator and show that we can compute the proposed estimator through a convex program. Empirical studies illustrate that the proposed method is effectively applied to the problem of estimating the average payoff of a stock option that is contingent on two different stocks.

Keywords: nonparametric regression, multivariate regression, convex programming

1. Introduction

We look at the problem of estimating an unknown function $f_* : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}$, where \mathbb{R}^d is the d -dimensional Euclidean space. We presume that we observe r noisy measurements Y_{i1}, \dots, Y_{ir} of f_* at each $t_i = (x_1(i), \dots, x_d(i)) \in \Omega$ and that

$$Y_{ij} = f_*(t_i) + \epsilon_{ij}$$

for $1 \leq i \leq n$ and $1 \leq j \leq r$, where $\epsilon_{i1}, \dots, \epsilon_{ir}$ are independent and identically distributed (iid) random variables with a mean of 0 and a variance of $\sigma_i^2 < \infty$.

In many applications, f_* is assumed to be smooth. We thus assume that f_* has partial derivatives of total order m that are square integrable over Ω . In such a case, one naturally tries to fit a function f to data by minimizing the smoothness of the fitted function f , which can be measured by

$$J_m^d(f) = \sum_{\substack{\alpha_1 + \dots + \alpha_d = m, \\ \alpha_1, \dots, \alpha_d \text{ are non-negative integers}}} \int_{\Omega} \dots \int \left(\frac{\partial^m f}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right)^2 dx_1 \dots dx_d$$

over $f \in \mathcal{X}$, where

$$\mathcal{X} = \left\{ g : \Omega \rightarrow \mathbb{R} \mid g \text{ is an analytic function and } \int_{\Omega} \dots \int \left(\frac{\partial^m g}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right)^2 dx_1 \dots dx_d < \infty \right. \\ \left. \text{for all nonnegative integers } \alpha_1, \dots, \alpha_d \text{ satisfying } \alpha_1 + \dots + \alpha_d = m \right\}.$$

While minimizing the smoothness J_m^d of f over \mathcal{X} , we also want to ensure that the fitted function f is close to the data points. The closeness between f and the data points can be measured by

$$I(f) = \frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - f(t_i))^2,$$

where $\bar{Y}_i = \sum_{j=1}^r Y_{ij}/r$ for $1 \leq i \leq n$. An upper bound on $I(f)$ can be obtained by noticing that

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - f_*(t_i))^2 &= \frac{1}{n} \sum_{i=1}^n \bar{\epsilon}_i^2 \quad \left(\bar{\epsilon}_i \triangleq \sum_{j=1}^r \epsilon_{ij}/r \text{ for } 1 \leq i \leq n \right) \\ &\approx \frac{1}{nr} \sum_{i=1}^n (N(0, \sigma_i^2))^2 \text{ for large } r \text{ by the weak law of large numbers} \\ &\approx \frac{1}{nr} \sum_{i=1}^n \sigma_i^2 \text{ for large } n \text{ by the strong law of large numbers,} \end{aligned}$$

where $N(0, \sigma^2)$ is normally distributed with a mean of 0 and a variance of σ^2 . The notation “ \approx ” expresses equality in some appropriate sense.

Since σ_i^2 can be estimated by the sample variance $\sum_{j=1}^r (Y_{ij} - \bar{Y}_i)^2 / (r - 1)$ for $1 \leq i \leq n$, an upper bound S^2 on $I(f)$ can be obtained by

$$S^2 = \frac{1}{nr} \sum_{i=1}^n \sum_{j=1}^r (Y_{ij} - \bar{Y}_i)^2 / (r - 1). \quad (1)$$

So, we propose \hat{f}_n , the solution to the following minimization problem, as our estimate of f_* :

$$\begin{aligned} \text{Problem (A):} \quad & \text{Minimize}_{f \in \mathcal{X}} \sum_{\substack{\alpha_1 + \dots + \alpha_d = m, \\ \alpha_1, \dots, \alpha_d \text{ are non-negative integers}}} \int_{\Omega} \dots \int \left(\frac{\partial^m f}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right)^2 dx_1 \dots dx_d \\ \text{Subject to} \quad & \frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - f(t_i))^2 \leq S^2. \end{aligned}$$

Problem (A) can serve as an alternative approach to the following traditional problem:

$$\text{Problem (B):} \quad \text{Minimize}_{f \in \mathcal{X}} \frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - f(t_i))^2 + \lambda J_m^d(f)$$

for some $\lambda > 0$, which was studied extensively in the literature by Duchon (1977), Meinguet (1979), and Cox (1984) among others. Despite its rich theory established by numerous researchers, the performance of the solution to Problem (B) is dependent on the parameter λ , and the procedure of finding the right value of λ is computationally burdensome. Furthermore, the solution to Problem (B) can be computed by solving a system of linear equations, which involves an ill-conditioned matrix whose diagonal elements are all zero (page 145 of Dierckx, 1993). To tackle this ill-conditioned setting, additional numerical procedures must be performed.

On the contrary, the numerical performance of the solution to Problem (A) does not depend on any artificially plugged-in parameters since one can obtain a good estimate of S^2 via Equation (1). Moreover, we prove that Problem (A) always has a solution and that the solution can be found by using a convex program. There are many efficient algorithms that can solve a convex program with guaranteed convergence, so our convex program formulation makes it possible to utilize such algorithms to find the solution to Problem (A) with guaranteed convergence. The numerical results in Section 3 illustrate that the solution to Problem (A) is computed successfully within a few seconds.

There are many applications where the independent variable of a regression function lies in the multi-dimensional space. Even though Problem (A) was mentioned as one of the possible solutions to estimation of a multivariate smooth function in Equations (8.2) and (8.3) on page 139 of Dierckx (1993), the existence of the solution to Problem (A) has not been proved yet and a numerical procedure with guaranteed convergence has not been proposed so far. This paper provides these missing foundations by proving the existence of the solution to Problem (A) and by providing numerical procedures with guaranteed convergence based on convex programs.

This paper is organized as follows. In Section 2, we prove the existence of the solution to Problem (A), and provide the convex program formulation of Problem (A). In Section 3, we present numerical results to illustrate that our proposed estimator successfully recovers the underlying true function as n increases to infinity. Concluding remarks are included in Section 4.

2. Proposed Formulation

In this section, we prove that Problem (A) always has a solution in Proposition 1. Proposition 1 also describes how we can solve Problem (A) with a convex program.

To describe the details of Proposition 1, we need some preliminaries. Let \mathcal{P} be the set of all polynomials over Ω of the total degree less than m . It should be noted that \mathcal{P} is the vector space of the dimension

$$K = \binom{m+d-1}{d}$$

spanned by the monomials of the total degree less than m . For example, when $d = 2$, $\Omega = [0, 1]^2$, and $m = 3$, \mathcal{P} is spanned by $\varphi_1, \dots, \varphi_6$, which are defined by

$$\varphi_1(x_1, x_2) = 1, \varphi_2(x_1, x_2) = x_1, \varphi_3(x_1, x_2) = x_2, \varphi_4(x_1, x_2) = x_1^2, \varphi_5(x_1, x_2) = x_2^2, \varphi_6(x_1, x_2) = x_1 x_2$$

for all $(x_1, x_2) \in \Omega$. We will denote the K monomials of the total degree less than m by $\varphi_1, \dots, \varphi_K$.

A set $\{u_1, \dots, u_K\}$ of distinct points of Ω is called unisolvent for \mathcal{P} if there exists a unique $p \in \mathcal{P}$ satisfying $p(u_i) = z_i$ for $1 \leq i \leq K$ and for any real numbers z_1, \dots, z_K . It should be noted that $\{u_1, \dots, u_K\}$ is unisolvent for \mathcal{P} if and only if the matrix $M = \{M_{ij} : 1 \leq i, j \leq K\}$, whose (i, j) th entry is $\varphi_i(u_j)$, is invertible.

For $t = (x_1, \dots, x_d) \in \mathbb{R}^d$, $\|t\|$ is defined by $(x_1^2 + \dots + x_d^2)^{1/2}$.

We are ready to present Proposition 1.

Proposition 1 Assume that $2m > d$ and that $\{t_1, \dots, t_n\}$ contains a subset $\{u_1, \dots, u_K\}$ which is unisolvent for \mathcal{P} . Then there exists a solution \hat{f}_n to Problem (A). Furthermore,

(a) $\hat{f}_n : \Omega \rightarrow \mathbb{R}$ can be represented as

$$\hat{f}_n(t) = \hat{a}_1 \varphi_1(t) + \dots + \hat{a}_K \varphi_K(t) + \sum_{i=1}^n \hat{a}_{K+i} R(t, t_i)$$

for $t \in \Omega$, where

$$R(t, t_i) = \begin{cases} \|t - t_i\|^{2m-d} \ln \|t - t_i\|, & \text{for } t \neq t_i \\ 0, & \text{for } t = t_i \end{cases}$$

for $t \in \mathbb{R}^2$ if $2m - d$ is an even integer and $R(t, t_i) = \|t - t_i\|^{2m-d}$ for $t \in \mathbb{R}^d$ if $2m - d$ is an odd integer.

(b) $\hat{a}_1, \dots, \hat{a}_{K+n}, \hat{y}_1, \dots, \hat{y}_n$ is the solution to the following program with the decision variables $a_1, \dots, a_{K+n}, y_1, \dots, y_n \in \mathbb{R}$:

$$\begin{aligned} & \text{Minimize} && \sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} a_i a_j \\ & \text{Subject to} && \frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - y_i)^2 \leq S^2 \\ & && \sum_{i=1}^{K+n} a_i \varphi_i(t_j) = y_j, \quad j = 1, \dots, n, \end{aligned} \tag{2}$$

where

$$C_{ij} = \sum_{\substack{\alpha_1 + \dots + \alpha_d = m, \\ \alpha_1, \dots, \alpha_d \text{ are non-negative integers}}} \int_{\Omega} \dots \int \left(\frac{\partial^m \varphi_i}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right) \left(\frac{\partial^m \varphi_j}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \right) dx_1 \dots dx_d$$

for $1 \leq i, j \leq K + n$ and $\varphi_{K+i}(t) = R(t, t_i)$ for $t \in \Omega$ and $1 \leq i \leq n$.

(c) Problem (2) is a convex program.

Proof. Let

$$C = \left\{ (y_1, \dots, y_n) : \sum_{i=1}^n (\bar{Y}_i - y_i) / n \leq S^2 \right\}.$$

It should be noted that C is a nonempty, closed, and bounded subset of \mathbb{R}^n .

By Talmi and Gilat (1977), for any (y_1, \dots, y_n) in \mathbb{R}^n , there exists a unique $f \in \mathcal{X}$ such that

1. $f(t) = \sum_{i=1}^{K+n} a_i \varphi_i(t)$ for all $t \in \Omega$ and for some a_1, \dots, a_{K+n} ,
2. $f(t_i) = y_i$ for $1 \leq i \leq n$, and
3. $J_m^d(f) \leq J_m^d(g)$ for any $g \in \mathcal{X}$ satisfying $g(t_i) = y_i$ for $1 \leq i \leq n$.

The a_i 's in 1 are uniquely determined by the following linear system with $K+n$ variables and $K+n$ linear equations:

$$\begin{aligned} \sum_{i=1}^{K+n} a_i \varphi_i(t_j) &= y_j, & 1 \leq j \leq n \\ \sum_{j=1}^n a_{K+j} \varphi_i(t_j) &= 0, & 1 \leq i \leq K. \end{aligned} \quad (3)$$

By the uniqueness of the a_i 's, the linear system (3) is nonsingular, and hence, the mapping from (y_1, \dots, y_n) to $\sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} a_i a_j$ is continuous. Hence, there exists $(\hat{y}_1, \dots, \hat{y}_n) \in C$ that minimizes $\sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} a_i a_j$ over C .

Let $\hat{a}_1, \dots, \hat{a}_{K+n}$ be the solution to (3) with (y_1, \dots, y_n) replaced with $\hat{y}_1, \dots, \hat{y}_n$. It remains to show that $\hat{f}_n : \Omega \rightarrow \mathbb{R}$ given by $\hat{f}_n(t) = \sum_{i=1}^{K+n} \hat{a}_i \varphi_i(t)$ for $t \in \Omega$ minimizes Problem (A). Let g be any function in \mathcal{X} satisfying $\sum_{i=1}^n (\bar{Y}_i - g(t_i))^2 / n \leq S^2$. By Talmi and Gilat (1977), there exists a function $\bar{g} \in \mathcal{X}$ and $\bar{a}_1, \dots, \bar{a}_{K+n}$ that satisfy 1, 2, and 3 with y_i replaced with $g(t_i)$ for $1 \leq i \leq n$. It should be noted that $J_m^d(\bar{g}) \leq J_m^d(g)$. Since $(g(t_1), \dots, g(t_n)) \in C$ and $(\hat{y}_1, \dots, \hat{y}_n)$ is the minimizer of $\sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} a_i a_j$,

$$J_m^d(\hat{f}_n) = \sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} \hat{a}_i \hat{a}_j \leq \sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} \bar{a}_i \bar{a}_j = J_m^d(\bar{g}) \leq J_m^d(g),$$

resulting in $J_m^d(\hat{f}_n) \leq J_m^d(g)$. Hence, \hat{f}_n is a minimizer of Problem (A), proving (a).

To prove (b), we first notice that $\hat{a}_1, \dots, \hat{a}_{K+n}, \hat{y}_1, \dots, \hat{y}_n$ is a feasible solution to (2). Let $\tilde{a}_1, \dots, \tilde{a}_{K+n}, \underline{y}_1, \dots, \underline{y}_n$ be a feasible solution to (2). By Talmi and Gilat (1977), there exists a unique $\underline{f} \in \mathcal{X}$ such that

- i. $\underline{f}(t) = \sum_{i=1}^{K+n} \underline{a}_i \varphi_i(t)$ for all $t \in \Omega$ and for some $\underline{a}_1, \dots, \underline{a}_{K+n}$,
- ii. $\underline{f}(t_i) = \underline{y}_i$ for $1 \leq i \leq n$, and
- iii. $J_m^d(\underline{f}) \leq J_m^d(g)$ for any $g \in \mathcal{X}$ satisfying $g(t_i) = \underline{y}_i$ for $1 \leq i \leq n$.

Since $\tilde{f} : \Omega \rightarrow \mathbb{R}$ defined by $\tilde{f}(t) = \sum_{i=1}^{K+n} \tilde{a}_i \varphi_i(t)$ for $t \in \Omega$ satisfies $\tilde{f}(t_i) = \underline{y}_i$ for $1 \leq i \leq n$, it follows that

$$J_m^d(\underline{f}) \leq J_m^d(\tilde{f}). \quad (4)$$

Since $\hat{y}_1, \dots, \hat{y}_n$ is the minimizer of $\sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} a_i a_j$,

$$J_m^d(\hat{f}) = \sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} \hat{a}_i \hat{a}_j \leq J_m^d(\underline{f}). \quad (5)$$

A combination of (4) and (5) yields

$$\sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} \hat{a}_i \hat{a}_j \leq \sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} \tilde{a}_i \tilde{a}_j,$$

concluding that $\hat{a}_1, \dots, \hat{a}_{K+n}, \hat{y}_1, \dots, \hat{y}_n$ is a minimizer of (2).

To prove (c), we note that

$$\sum_{i=1}^{K+n} \sum_{j=1}^{K+n} C_{ij} a_i a_j = \sum_{\substack{\alpha_1 + \dots + \alpha_d = m, \\ \alpha_1, \dots, \alpha_d \text{ are non-negative integers}}} \int_{\Omega} \dots \int \left(\frac{\partial^m \varphi_1}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} a_1 + \dots + \frac{\partial^m \varphi_{K+n}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} a_{K+n} \right)^2 dx_1 \dots dx_d \geq 0$$

for all $a_1, \dots, a_{K+n} \in \mathbb{R}$, and hence, Problem (2) is a quadratically constrained quadratic program. Thus, it follows that Problem (2) is a convex program. \square

3. Numerical Results

We now examine how the proposed estimator \hat{f}_n can be calculated in two numerical settings.

In the first setting, we assume that the true function $f_* : [1, 3]^3 \rightarrow \mathbb{R}$ is given by $f_*(x_1, x_2, x_3) = x_1^{1/2} + x_2^{1/2} + x_3^{1/2}$ for $(x_1, x_2, x_3) \in [0, 1]^3$ and that we can observe a noisy measurement $f_*(x_1, x_2, x_3) + \epsilon$ of $f_*(x_1, x_2, x_3)$, where ϵ is normally distributed with a mean of 0 and a variance of 4.

In the second setting, we look at the case where the true function $f_*(x_1, x_2)$ is the average payoff of a stock option that is contingent on two different stocks, where x_1 and x_2 denote the prices of the underlying stocks.

In both settings, we calculate the mean square error (MSE) of the proposed estimator \hat{f}_n , defined by

$$\frac{1}{n} \sum_{i=1}^n (\hat{f}_n(t_i) - f_*(t_i))^2,$$

and compare it to the MSE of the crude estimator $\bar{Y}_1, \dots, \bar{Y}_n$, defined by

$$\frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - f_*(t_i))^2.$$

The numerical results show that the MSE of the proposed estimator converges to 0 as n increases to infinity while the MSE of the crude estimator converges to a positive constant. This phenomenon empirically shows that our proposed estimator converges to the underlying function as n increases to infinity in the L_2 sense.

We conducted all simulations using a 64-bit computer with an Intel(R) Core TM i7-6700K CPU at 4GHz and a 32 GB RAM. We also programmed all simulations in MATLAB R2016a.

3.1 A Simple Case

We assume that $f_* : [1, 3]^3 \rightarrow \mathbb{R}$ is given by $f_*(x_1, x_2, x_3) = x_1^{1/2} + x_2^{1/2} + x_3^{1/2}$ for $(x_1, x_2, x_3) \in [1, 3]^3$, $\{t_1, \dots, t_n\}$ is given by

$$\left\{ (1 + 2u/n^{1/3} - 1/n^{1/3}, 1 + 2v/n^{1/3} - 1/n^{1/3}, 1 + 2w/n^{1/3} - 1/n^{1/3}) : 1 \leq u, v, w \leq n^{1/3} \right\},$$

$Y_{ij} = f(t_i) + \epsilon_{ij}$ for $1 \leq i \leq n$ and $1 \leq j \leq r$ with $r = 1$, and the ϵ_{ij} 's are iid normal random variables with a mean of 0 and a variance of 4.

We assume that it is known a priori that f_* has integrable partial derivatives of a total degree of 2. Then, our proposed estimator \hat{f}_n can be computed as

$$\hat{f}_n(x_1, x_2, x_3) = \hat{a}_1 \varphi_1(x_1, x_2, x_3) + \hat{a}_2 \varphi_2(x_1, x_2, x_3) + \hat{a}_3 \varphi_3(x_1, x_2, x_3) + \hat{a}_4 \varphi_4(x_1, x_2, x_3) + \sum_{i=1}^n \hat{a}_{i+4} \varphi_{i+4}(x_1, x_2, x_3)$$

for $(x_1, x_2, x_3) \in [1, 3]^3$, where

$$\begin{aligned} \varphi_1(x_1, x_2, x_3) &= 1, \varphi_2(x_1, x_2, x_3) = x_1, \varphi_3(x_1, x_2, x_3) = x_2, \varphi_4(x_1, x_2, x_3) = x_3, \\ \varphi_{i+4}(x_1, x_2, x_3) &= \left((x_1 - x_1(i))^2 + (x_2 - x_2(i))^2 + (x_3 - x_3(i))^2 \right)^{1/2} \end{aligned}$$

for $1 \leq i \leq n$ and $(x_1, x_2, x_3) \in [1, 3]^3$, and $(\hat{a}_1, \dots, \hat{a}_{n+4})$ is computed from the following program:

$$\begin{aligned} &\text{Minimize} \quad \sum_{i=1}^{n+4} \sum_{j=1}^{n+4} a_i a_j \\ &\quad \times \left(\int_1^3 \int_1^3 \int_1^3 \left(\frac{\partial^2 \varphi_i}{\partial x_1^2} \frac{\partial^2 \varphi_j}{\partial x_1^2} + \frac{\partial^2 \varphi_i}{\partial x_2^2} \frac{\partial^2 \varphi_j}{\partial x_2^2} + \frac{\partial^2 \varphi_i}{\partial x_3^2} \frac{\partial^2 \varphi_j}{\partial x_3^2} + \frac{\partial^2 \varphi_i}{\partial x_1 \partial x_2} \frac{\partial^2 \varphi_j}{\partial x_1 \partial x_2} + \frac{\partial^2 \varphi_i}{\partial x_2 \partial x_3} \frac{\partial^2 \varphi_j}{\partial x_2 \partial x_3} + \frac{\partial^2 \varphi_i}{\partial x_3 \partial x_1} \frac{\partial^2 \varphi_j}{\partial x_3 \partial x_1} \right) dx_1 dx_2 dx_3 \right) \\ &\text{Subject to} \quad \frac{1}{n} \sum_{i=1}^n \left(\bar{Y}_i - \sum_{j=1}^{n+4} a_j \varphi_j(t_i) \right)^2 \leq \frac{1}{nr} \sum_{i=1}^n \sigma_i^2, \end{aligned} \tag{6}$$

with $\sigma_i^2 = 4$ for $1 \leq i \leq n$. The triple integration appearing in the objective function of Problem (6) is computed in MATLAB using MATLAB functions “triplequad” and “integral3”. Problem (A) is then solved using CVX, developed by Grant and Boyd (2014).

The MSE of \hat{f}_n is calculated by

$$\frac{1}{n} \sum_{i=1}^n (\hat{f}_n(t_i) - f_*(t_i))^2$$

and is compared to the MSE of the crude estimator $\bar{Y}_1, \dots, \bar{Y}_n$,

$$\frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - f_*(t_i))^2.$$

We calculate the 95% confidence intervals of the MSE using 400 iid replications, and report them in Table 1. We also measure the amounts of time spent while calculating the proposed estimator, and compute their averages using 400 iid replications. We report these values in Table 1. We observe that the MSEs of the proposed estimator are less than those of the crude estimator. The MSE of the proposed estimator converges to 0 as n increases. We can thus deduce that the proposed estimator converges to the true function as n increases to infinity.

Table 1. The 95% confidence intervals of the MSE and the average amount of time spent when $f_*(x_1, x_2, x_3) = x_1^{1/2} + x_2^{1/2} + x_3^{1/2}$ and $r = 1$.

n	MSE of the proposed estimator	MSE of the crude estimator	Time (second)
8	2.0674 ± 0.1376	4.0770 ± 0.1895	0.07
27	0.5789 ± 0.0373	3.9562 ± 0.1021	0.08
64	0.2703 ± 0.0174	4.0012 ± 0.0682	0.09
125	0.1474 ± 0.0102	4.0070 ± 0.0509	0.12
216	0.0884 ± 0.0057	3.9841 ± 0.0364	0.24

We next repeat the above experiment with $r = 30$ and with the right hand side of the constraint of Problem (6) replaced by $\frac{1}{nr} \sum_{i=1}^n \sum_{j=1}^r (Y_{ij} - \bar{Y}_i)^2 / (r - 1)$. Table 2 reports the 95% confidence intervals of the MSE and the average amounts of time spent to compute the proposed estimator, using 400 iid replications.

Table 2. The 95% confidence intervals of the MSE and the average amount of time spent when $f_*(x_1, x_2, x_3) = x_1^{1/2} + x_2^{1/2} + x_3^{1/2}$ and $r = 30$.

n	MSE of the proposed estimator	MSE of the crude estimator	Time (second)
8	0.0694 ± 0.0056	0.1387 ± 0.0078	0.53
27	0.0206 ± 0.0016	0.1334 ± 0.0039	0.55
64	0.0093 ± 0.0007	0.1332 ± 0.0027	0.55
125	0.0052 ± 0.0004	0.1325 ± 0.0020	0.60
216	0.0033 ± 0.0002	0.1333 ± 0.0015	0.71

3.2 Estimation of “Gamma” of a Stock Option

We consider the case where $f_* : [0.5, 1.5]^2 \rightarrow \mathbb{R}$ is given so that $f_*(x_1, x_2)$ is the average payoff of a stock option that is contingent on the stock price x_1 of Stock 1 and the stock price x_2 of Stock 2. The payoff structure of this stock option is given as follows. The option is active between time 0 and time $T = 365$. The prices of Stock 1 and Stock 2 at time $t \in [0, T]$ are denoted by S_t^1 and S_t^2 , respectively. There are 6 time instances, say d_1, \dots, d_6 , when the option expires early. In other words, on day d_i , the option expires and generates a payoff of $\$r_i$ if both $S_{d_i}^1/S_0^1$ and $S_{d_i}^2/S_0^2$ are above a certain price b_i for $1 \leq i \leq 6$. If the option does not expire before time $T = 365$, the option expires at time $T = 365$ as follows. If the option does not expire before time $T = 365$, and both S_T^1/S_0^1 and S_T^2/S_0^2 stay above the price b over the entire lifetime $t \in [0, T]$ of the option, then the option expires at time $T = 365$ and generates a payoff of \$1. If the option does not expire before time $T = 365$, and either S_t^1/S_0^1 or S_t^2/S_0^2 drops below b at any time instance t between 0 and $T = 365$, the option expires at time $T = 365$ and generates a payoff equal to the minimum of S_T^1/S_0^1 and S_T^2/S_0^2 .

We let

$$\{t_1, \dots, t_n\} = \left\{ \left(0.5 + u/n^{1/2} - 1/(2n^{1/2}), 0.5 + v/n^{1/2} - 1/(2n^{1/2}) \right) : 1 \leq u, v \leq n^{1/2} \right\}.$$

We assume that the current time is 60 days before $T = 365$. For each $t_i = (x_1(i), x_2(i))$, we generate two independent sample paths of a geometric Brownian motion as the trajectories of the prices of Stock 1 and Stock 2 between the current time and $T = 365$ with the current stock prices of Stock 1 and Stock 2 equal to $x_1(i)$ and $x_2(i)$, respectively. We then compute the corresponding payoff of the stock option. The following parameters are used to compute the payoff of the stock option: $d_1 = 61, d_2 = 122, d_3 = 182, d_4 = 243, d_5 = 304, d_6 = 365, b_1 = 0.90, b_2 = 0.90, b_3 = 0.85, b_4 = 0.85, b_5 = 0.80, b_6 = 0.80, r_1 = 1.05, r_2 = 1.10, r_3 = 1.15, r_4 = 1.20, r_5 = 1.25, r_6 = 1.30$, and $b = 0.70$. The volatility per year is 30% and the risk-free interest rate per year is 5%. The prices of Stock 1 and Stock 2 at time 0 are assumed to be \$1.

We set $r = 5$, so 5 iid replications Y_{i1}, \dots, Y_{i5} are generated at each $(x_1(i), x_2(i))$ and $\bar{Y}_i = \sum_{j=1}^r Y_{ij}/r$ is computed for $1 \leq i \leq n$.

The second derivative, often called the “gamma” of an option price is of particular interest to portfolio managers because it determines how many shares of the option must be sold or bought to stay in the delta-hedged position. Moreover, it is often assumed that the gamma is a smooth function of underlying stock prices because a smooth gamma curve suggests consistent buying or selling strategies over a domain of possible stock price. Hence, it is natural for us to assume that f_* has the smooth derivative. Since we can compute the smoothness of the second derivative of f_* using J_m^d with $m = 4$, our proposed estimator is computed as

$$\hat{f}_n(x_1, x_2) = \hat{a}_1 \varphi_1(x_1, x_2) + \dots + \hat{a}_{10} \varphi_{10}(x_1, x_2) + \sum_{i=1}^n \hat{a}_{i+10} \varphi_{i+10}(x_1, x_2)$$

for $(x_1, x_2) \in [0.5, 1.5]^2$, where

$$\begin{aligned} \varphi_1(x_1, x_2) &= 1, \varphi_2(x_1, x_2) = x_1, \varphi_3(x_1, x_2) = x_2, \varphi_4(x_1, x_2) = x_1^2, \varphi_5(x_1, x_2) = x_1 x_2, \varphi_6(x_1, x_2) = x_2^2, \\ \varphi_7(x_1, x_2) &= x_1^3, \varphi_8(x_1, x_2) = x_1^2 x_2, \varphi_9(x_1, x_2) = x_1 x_2^2, \varphi_{10}(x_1, x_2) = x_2^3, \\ \varphi_{i+10}(x_1, x_2) &= \left((x_1 - x_1(i))^2 + (x_2 - x_2(i))^2 \right)^3 \ln \left((x_1 - x_1(i))^2 + (x_2 - x_2(i))^2 \right)^{1/2} \end{aligned}$$

for $1 \leq i \leq n$ and $(x_1, x_2) \in [0.5, 1.5]^2$, and $(\hat{a}_1, \dots, \hat{a}_{n+10})$ is calculated from the following program:

$$\begin{aligned} \text{Minimize} \quad & \sum_{i=1}^{n+10} \sum_{j=1}^{n+10} a_i a_j \\ & \times \left(\int_{0.5}^{1.5} \int_{0.5}^{1.5} \left(\frac{\partial^4 \varphi_i}{\partial x_1^4} \frac{\partial^4 \varphi_j}{\partial x_1^4} + \frac{\partial^4 \varphi_i}{\partial x_1^3 \partial x_2} \frac{\partial^4 \varphi_j}{\partial x_1^3 \partial x_2} + \frac{\partial^4 \varphi_i}{\partial x_1^2 \partial x_2^2} \frac{\partial^4 \varphi_j}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 \varphi_i}{\partial x_1 \partial x_2^3} \frac{\partial^4 \varphi_j}{\partial x_1 \partial x_2^3} + \frac{\partial^4 \varphi_i}{\partial x_2^4} \frac{\partial^4 \varphi_j}{\partial x_2^4} \right) dx_1 dx_2 \right) \\ \text{Subject to} \quad & \frac{1}{n} \sum_{i=1}^n \left(\bar{Y}_i - \sum_{j=1}^{n+10} a_j \varphi_j(t_i) \right)^2 \leq \frac{1}{nr} \sum_{i=1}^n \sum_{j=1}^r (Y_{ij} - \bar{Y}_i)^2 / (r-1). \end{aligned} \tag{7}$$

Problem (7) is solved with CVX. The double integration appearing in the objective function Problem (7) is computed numerically in MATLAB.

The MSE of the proposed estimator \hat{f}_n is calculated by

$$\frac{1}{n} \sum_{i=1}^n (\hat{f}_n(t_i) - f_*(t_i))^2$$

and is compared to the MSE of the crude estimator $\bar{Y}_1, \dots, \bar{Y}_n$

$$\frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - f_*(t_i))^2.$$

We calculate the 95% confidence intervals of the MSE using 400 iid replications, and report them in Table 3. We also measure the amounts of time spent while calculating the proposed estimator, and compute their averages using 400 iid replications. We report these values in Table 3. We observe that the MSEs of the proposed estimator are less than those of the crude estimator. The MSE of the proposed estimator converges to 0 as n increases. We can thus deduce that the proposed estimator converges to the true function as n increases to infinity.

Table 3. The 95% confidence intervals of the MSE and the average amount of time spent when f_* is the average payoff of a stock option contingent on two types of stocks.

n	MSE of the proposed estimator	MSE of the crude estimator	Average amount of time spent for the computation of the proposed estimator (second)
36	0.0025 ± 0.0001	0.0044 ± 0.0002	0.53
64	0.0022 ± 0.0001	0.0041 ± 0.0001	0.54
100	0.0015 ± 0.0001	0.0042 ± 0.0001	0.56
144	0.0013 ± 0.0001	0.0042 ± 0.0001	0.63
196	0.0010 ± 0.0000	0.0041 ± 0.0001	0.73
256	0.0008 ± 0.0000	0.0042 ± 0.0001	0.81

4. Concluding Remarks

A new formulation for estimating a multivariate smooth function is proposed. The solution to the proposed formulation can be easily obtained by solving a convex program. It does not rely on any artificially plugged-in parameters and depends on the observed data set only. Hence, the proposed formulation tackles the computational inefficiency of the existing methods, which suffer from the dependency on the smoothing parameter and the ill-conditioned numerical settings.

Numerical results show that the proposed method is applied to 2-dimensional and 3-dimensional problems successfully. They also suggest that the proposed estimator converges to the true function as the number of data points increases to infinity. Thus, a promising research topic for the future is to prove the consistency of the proposed estimator.

References

- Cox, D. (1984). Multivariate smoothing spline functions. *SIAM J. Numer. Anal.*, 21, 789-813.
<https://doi.org/10.1137/0721053>
- Dierckx, P. (1993). *Curve and surface fitting with splines*. New York, NY: Oxford University Press Inc.
<https://doi.org/10.5860/choice.31-2162>
- Duchon, J. (1977). Splines minimizing rotation-invariant semi-norms in Sobolev spaces. *Constructive Theory of Functions of Several Variables. Lecture Notes in Mathematics*, 571. Berlin: Springer-Verlag. 85–100.
<https://doi.org/10.1007/BFb0086566>
- Grant, M., & Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.1. Retrieved January 2017, from <http://cvxr.com/cvx>
- Meinguet, J. (1979). Multivariate interpolation at arbitrary points made simple. *Journal of Applied Mathematics and Physics*, 30, 292–304. <https://doi.org/10.1007/BF01601941>
- Talmi, A., & Gilat, G. (1977). Method for smooth approximation of data. *J. Comput. Phys.*, 23, 93–123.
[https://doi.org/10.1016/0021-9991\(77\)90115-2](https://doi.org/10.1016/0021-9991(77)90115-2)

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).