

Estimating Smooth and Convex Functions

Eunji Lim¹, & Kihwan Kim²

¹ School of Business, Decision Sciences and Marketing, Adelphi University, Garden City, NY 11530, USA

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

Correspondence: Eunji Lim, Decision Sciences and Marketing, School of Business, Adelphi University, Garden City, NY 11530, USA. E-mail: elim@adelphi.edu

Received: June 17, 2020 Accepted: July 21, 2020 Online Published: July 28, 2020

doi:10.5539/ijsp.v9n5p40 URL: https://doi.org/10.5539/ijsp.v9n5p40

Abstract

We propose a new method for estimating an unknown regression function $f_* : [\alpha, \beta] \rightarrow \mathbb{R}$ from a dataset $(X_1, Y_1), \dots, (X_n, Y_n)$ when the only information available on f_* is the fact that f_* is convex and twice differentiable. In the proposed method, we fit a convex function to the dataset that minimizes the sum of the roughness of the fitted function and the average squared differences between the fitted function and f_* . We prove that the proposed estimator can be computed by solving a convex quadratic programming problem with linear constraints. Numerical results illustrate the superior performance of the proposed estimator compared to existing methods when i) f_* is the price of a stock option as a function of the strike price and ii) f_* is the steady-state mean waiting time of a customer in a single server queue.

Keywords: convexity regression, penalizes least squares, nonparametric regression, constrained smoothing, convex smoothing, smoothing splines

1. Introduction

In this paper, we propose a new method for estimating an unknown regression function $f_* : [\alpha, \beta] \rightarrow \mathbb{R}$ from a dataset $(X_1, Y_1), \dots, (X_n, Y_n)$, satisfying $\alpha < X_1 < \dots < X_n < \beta$ and $Y_i = f_*(X_i) + \epsilon_i$ for $1 \leq i \leq n$ and some error terms $(\epsilon_i : 1 \leq i \leq n)$, when the only information available on f_* is the fact that f_* is convex and twice differentiable.

In many applications, the unknown regression function f_* is not assumed to have any particular functional form, but is known to have certain shape characteristics, such as convexity, twice differentiability, or both. For example, the price of a call option is known to be a twice differentiable convex function as a function of the option's strike price; see Section 2 of Ait-Sahalia and Duarte (2003) for details. In such a setting, incorporating both shape restrictions, convexity and twice differentiability, on the estimator of f_* seems to be a natural step to follow.

When f_* is only known to be convex, an intuitive way to estimate f_* is to fit a convex function f minimizing the average of squared errors $\sum_{i=1}^n (Y_i - f(X_i))^2/n$. This estimator can be found by solving the following problem:

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 \\ \text{Subject to} \quad & f \text{ is convex,} \end{aligned} \quad (1)$$

which is equivalent to the following finite-dimensional quadratic programming problem in the decision variables $f(X_1), \dots, f(X_n)$:

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 \\ \text{Subject to} \quad & (f(X_2) - f(X_1))/(X_2 - X_1) \leq \dots \leq (f(X_n) - f(X_{n-1}))/ (X_n - X_{n-1}); \end{aligned} \quad (2)$$

see Hildreth (1954) for details. The solution to (2) is referred to as the “convex regression estimator,” and is typically not smooth, but piecewise linear.

On the other hand, when f_* is only known to be smooth enough to be twice differentiable, one can notice that the roughness of f_* can be measured by $\int_{\alpha}^{\beta} \{f_*^{(2)}(x)\}^2 dx$, where $f_*^{(2)}$ denotes the second derivative of f_* . We can then estimate f_* by finding the solution to

$$\text{Minimize}_{f \in \mathcal{F}} \quad \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 + \lambda \int_{\alpha}^{\beta} \{f^{(2)}(x)\}^2 dx \quad (3)$$

for some positive constant λ , where

$$\mathcal{F} = \{f \text{ is differentiable on } [a, b] \text{ and has absolutely continuous first derivative}\}.$$

It should be noted that $f \in \mathcal{F}$ if and only if f is differentiable everywhere on $[\alpha, \beta]$ with derivative $f^{(1)}$ and there is an integrable function $f^{(2)}$ such that $\int_{\alpha}^x f^{(2)}(t)dt = f^{(1)}(x) - f^{(1)}(\alpha)$ for all $x \in [\alpha, \beta]$. The first term in (3) is called the goodness-of-fit term, and measures how close the fitted function f is to the dataset $(X_1, Y_1), \dots, (X_n, Y_n)$, while the second term in (3) is called the roughness term, and measures how “wiggly” f is. There is a trade-off between the goodness-of-fit and roughness terms because as the function f gets close to the data points, its roughness increases. The constant λ appearing in (3) is called the smoothing constant, and makes a balance between the goodness-of-fit and roughness terms. The solution to (3) is called the penalized least squares estimator of f_* and can be computed by solving a system of linear equations; see, for example, Theorem 2.4 on page 19 of Green & Silverman (1994).

The convex regression estimator and the penalized least square estimator only incorporate one of the shape characteristics of f_* ; the former incorporates convexity only, while the latter incorporates twice differentiability only. In many settings, f_* is known to be both convex and twice differentiable, and hence, one wishes to incorporate both convexity and twice differentiability into an estimator of f_* . With this goal in mind, our proposed estimator is motivated as follows. Since f_* is twice differentiable, the solution to (3) provides a twice differentiable estimator of f_* , which is not necessarily convex. We notice that the solution \hat{f}_n to (3) exists uniquely and has the following properties:

- P1. \hat{f}_n is twice continuously differentiable,
- P2. \hat{f}_n is a polynomial of degree 3 on $[X_{i-1}, X_i]$ for $2 \leq i \leq n$, and
- P3. \hat{f}_n is a polynomial of degree 1 on $[\alpha, X_1]$ and $[X_n, \beta]$;

see, for example, Green & Silverman (1994) or Exercise 5.7 of Hastie et al. (2009). A function satisfying P1, P2, and P3 is called a natural spline of degree 3. In particular, \hat{f}_n can be represented as

$$\hat{f}_n(x) = a_0 + a_1x + \sum_{i=1}^n b_i(x - X_i)_+^3$$

for some constants $a_0, a_1, b_1, \dots, b_n \in \mathbb{R}$ and $x \in [\alpha, \beta]$, where $x_+ \triangleq \max(0, x)$ for $x \in \mathbb{R}$; see, for example, page 3 of Greville (1969).

Inspired by this, we consider the following variation of (3):

$$\text{Minimize}_{f \in \mathcal{G}} \quad \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 + \lambda \int_{\alpha}^{\beta} \{f^{(2)}(x)\}^2 dx, \tag{4}$$

where

$$\mathcal{G} = \{g : [\alpha, \beta] \rightarrow \mathbb{R} \text{ such that } g(x) = a_0 + a_1x + \sum_{i=1}^n b_i(x - X_i)_+^3 \text{ for some } a_0, a_1, b_1, \dots, b_n \in \mathbb{R}\}.$$

Additionally, when f_* is known to be convex, i.e., $f^{(2)}(x) \geq 0$ for all $x \in [\alpha, \beta]$, we can incorporate the convexity requirement by modifying (4) as follows:

$$\begin{aligned} \text{Minimize}_{f \in \mathcal{G}} \quad & \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 + \lambda \int_{\alpha}^{\beta} \{f^{(2)}(x)\}^2 dx \\ \text{Subject to} \quad & f^{(2)}(x) \geq 0 \text{ for all } x \in [\alpha, \beta]. \end{aligned} \tag{5}$$

Problem (5) appears to have an infinite number of constraints. However, one can realize that for any $f \in \mathcal{G}$, $f^{(2)}$ is linear over $[X_{i-1}, X_i]$ for $i = 2, \dots, n$, $[a, X_1]$, and $[X_n, b]$. Hence, the condition $f^{(2)}(x) \geq 0$ for $x \in [\alpha, \beta]$ can be replaced by $f^{(2)}(\alpha) \geq 0$, $f^{(2)}(X_i) \geq 0$ for $1 \leq i \leq n$, and $f^{(2)}(\beta) \geq 0$. We can therefore suggest the following formulation, which is equivalent to (5):

$$\begin{aligned} \text{Minimize}_{f \in \mathcal{G}} \quad & \frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 + \lambda \int_{\alpha}^{\beta} \{f^{(2)}(x)\}^2 dx \\ \text{Subject to} \quad & f^{(2)}(\alpha) \geq 0 \\ & f^{(2)}(X_i) \geq 0, \quad 1 \leq i \leq n \\ & f^{(2)}(\beta) \geq 0 \end{aligned} \tag{6}$$

Proposition 1 in Section 2 establishes that Problem (6) can be converted to a finite-dimensional convex quadratic programming problem. The solution \hat{g}_n to Problem (6) is our proposed estimator of f_* . Figure 1 displays instances of the convex regression estimator, the penalized least squares estimator, and the proposed estimator when $f_*(x) = x^2$ for $x \in [0, 1]$ and $Y_i = f(X_i) + \epsilon_i$ for $1 \leq i \leq n$, where the ϵ_i 's are independent and identically distributed (iid) normal random variables with a mean of 0 and a standard deviation of 0.5. Figure 1 shows that the convex regression estimator is not smooth, the penalized least squares estimator is not convex, but the proposed estimator is both convex and smooth.

Proposition 1 in Section 2 ensures that the proposed estimator \hat{g}_n exists always and can be found by solving a convex quadratic programming problem with linear constraints. There exist numerous algorithms that can solve a convex quadratic programming problem in polynomial time; see, for example, Ye & Tse (1984). Thus, Proposition 1 provides a framework that enables us to compute the proposed estimator with numerical efficiency. The numerical results in Section 3 demonstrate that \hat{g}_n converges to f_* as $n \rightarrow \infty$, thereby justifying the consistency of the proposed estimator numerically. They also indicate that \hat{g}_n shows superior performance compared to the convex regression estimator and the penalized least squares estimator.

The problem of enforcing the smoothness condition to the shape constrained estimators has received a considerable amount of attention in the literature. However, most work in the literature has focused on incorporating monotonicity and smoothness conditions (Mammen (1991), Hall and Huang (2001), Mammen et al. (2001), Du et al. (2013), Li et al. (2017), Sysoev and Burdakov (2019), and Yagi et al. (2020)). When incorporating convexity and smoothness conditions, most work has proposed methods for the case where f_* is both convex *and* monotone, and these methods cannot be easily generalized to the case where f_* is convex and not necessarily monotone, because they used basis functions that are both monotone and convex (Dole (1999) and Meyer (2008)). A number of researchers considered the case where f_* is convex and not necessarily monotone, but their methods are two-step procedures, in which the first step computes the smoothed version of the Y_i 's using nonparametric smoothing methods, and the second step finds the shape restricted estimator of the smoothed Y_i values; see, for example, Mammen and Thomas-agnan (1999). Ait-Sahalia and Duarte (2003) also consider a two-step procedure, in which the first step computes a shape restricted estimator, and the second step smoothes the shape restricted estimator using local polynomial regression. In our proposed method, we propose an estimator for a smooth convex function f_* that is not necessarily monotone. Our method uses a single-step, in which convexity and smoothness are incorporated simultaneously. In the numerical experiments conducted in Section 3, we compared our proposed estimator to an estimator computed from a two-step procedure, and observed that the proposed estimator shows superior performance. To our knowledge, this paper is the first to incorporate the convexity condition directly into the framework of the penalized least squares estimator in order to estimate a smooth convex function.

The rest of this paper is organized as follows. Section 2 establishes that the proposed estimator can be computed by solving a convex quadratic programming problem. The numerical performance of the proposed estimator is examined in Section 3 in three settings: i) f_* is a quadratic function, ii) f_* is the price of a European call option as a function of the strike price, and iii) f_* is the steady-state mean waiting time of a customer in a single server queue as a function of the service rate. Concluding remarks are included in Section 4.

2. Proposed Estimator

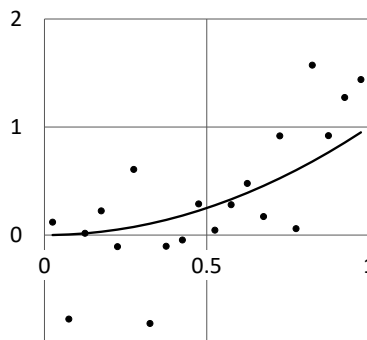
In this section, we establish Proposition 1, which proves that the solution to (6) exists and can be found by solving a convex quadratic programming problem. In Proposition 1, we view vectors as columns, and write x^T to denote the transpose of a vector x .

Proposition 1 *Suppose $\alpha < X_1 < \dots < X_n < \beta$. The solution \hat{g}_n to Problem (6) exists and can be represented as*

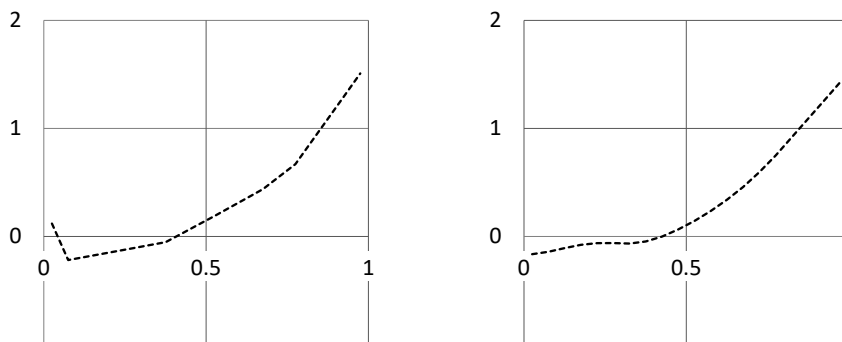
$$\hat{g}_n(x) = \hat{a}_0 + \hat{a}_1 x + \sum_{i=1}^n \hat{b}_i (x - X_i)_+^3,$$

where (\hat{a}_0, \hat{a}_1) , $(\hat{b}_1, \dots, \hat{b}_n)$, and $(\hat{g}_n(X_1), \dots, \hat{g}_n(X_n))$ is the solution to the following convex quadratic programming problem with the decision variables $a = (a_0, a_1) \in \mathbb{R}^2$, $b = (b_1, \dots, b_n) \in \mathbb{R}^n$, and $y = (y_1, \dots, y_n) \in \mathbb{R}^n$:

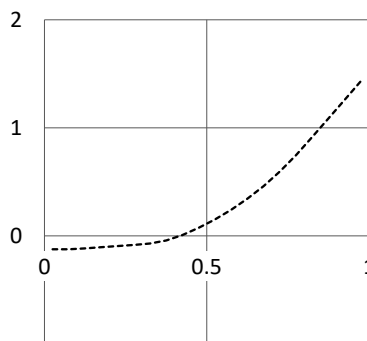
$$\begin{aligned} \text{Minimize} \quad & \frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2 + \lambda b^T Q b & (7) \\ \text{Subject to} \quad & A \begin{bmatrix} a \\ b \end{bmatrix} = y \\ & Bb \geq 0_n \\ & c^T b \geq 0, \end{aligned}$$



(a) f_* and the Y_i 's



(b) Convex regression estimator (left) and penalized least squares estimator (right)



(c) Proposed estimator

Figure 1. The solid line in (a) is $f_*(x) = x^2$ for $x \in [0, 1]$ and the dots in (a) are $Y_i = f_*(X_i) + N(0, 0.5^2)$ for $1 \leq i \leq 20$. The dashed line on the left side of (b) is the convex regression estimator. The dashed line on the right side of (b) is the penalized least squares estimator when $\lambda = 0.0001$. The dashed line in (c) is the proposed estimator, i.e., the solution to (6), when $\lambda = 0.0001$

where $0_n \in \mathbb{R}^n$ is the vector of zeros in all entries, $Q = (Q_{ij} : 1 \leq i, j \leq n)$,

$$Q_{ij} = Q_{ji} = 36(\beta^2/3 - (X_i + X_j)\beta^2/2 + X_i X_j \beta) - 36(X_j^3/3 - (X_i + X_j)X_j^2/2 + X_i X_j^2)$$

for $i \leq j$,

$$A = \begin{bmatrix} 1 & X_1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & X_2 & (X_2 - X_1)_+^3 & 0 & 0 & \cdots & 0 & 0 \\ 1 & X_3 & (X_3 - X_1)_+^3 & (X_3 - X_2)_+^3 & 0 & \cdot & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & X_n & (X_n - X_1)_+^3 & (X_n - X_2)_+^3 & (X_n - X_3)_+^3 & \cdots & (X_n - X_{n-1})_+^3 & 0 \end{bmatrix},$$

$B = (B_{ij} : 1 \leq i, j \leq n)$,

$$B_{ij} = \begin{cases} 6|X_i - X_j|, & \text{if } i \leq j \\ 0, & \text{otherwise,} \end{cases}$$

and $c = (6|\beta - X_i| : 1 \leq i \leq n)$.

Proof. For any $f(x) = a_0 + a_1x + \sum_{i=1}^n b_i(x - X_i)_+^3 \in \mathcal{G}$, let $y_i = f(X_i)$ for $1 \leq i \leq n$ and $\varphi_i : [\alpha, \beta] \rightarrow \mathbb{R}$ be defined by $\varphi_i(x) = (x - X_i)_+^3$ for $x \in [\alpha, \beta]$ and $1 \leq i \leq n$. Then, we have $\varphi_i^{(2)}(x) = 6(x - X_i)_+$ and $f^{(2)}(x) = \sum_{i=1}^n b_i \varphi_i^{(2)}(x)$. Also, we have

$$\int_{\alpha}^{\beta} \{f^{(2)}(x)\}^2 dx = \sum_{i=1}^n \sum_{j=1}^n b_i b_j \int_{\alpha}^{\beta} \varphi_i^{(2)}(x) \varphi_j^{(2)}(x) dx.$$

Thus, one can easily check

$$\frac{1}{n} \sum_{i=1}^n (Y_i - f(X_i))^2 + \lambda \int_{\alpha}^{\beta} \{f^{(2)}(x)\}^2 dx = \frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2 + \lambda b^T Q b,$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} f(X_1) \\ \vdots \\ f(X_n) \end{bmatrix} = A \begin{bmatrix} a \\ b \end{bmatrix},$$

$f^{(2)}(\alpha) = 0$, $f^{(2)}(\beta) = c^T b$, and

$$\begin{bmatrix} f^{(2)}(X_1) \\ \vdots \\ f^{(2)}(X_n) \end{bmatrix} = B b.$$

Hence, $\hat{g}_n(x) = \hat{a}_0 + \hat{a}_1x + \sum_{i=1}^n \hat{b}_i(x - X_i)_+^3$ is a solution to Problem (6) if and only if (\hat{a}_0, \hat{a}_1) , $(\hat{b}_1, \dots, \hat{b}_n)$, and $(\hat{g}_n(X_1), \dots, \hat{g}_n(X_n))$ is a solution to Problem (7).

It remains to show that Problem (7) is a convex quadratic programming problem and that there exists a solution to Problem (7). We will show that

$$\frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2 + \lambda b^T Q b$$

is convex and coersive in (a, b, y) . First, note that Q is positive definite because

$$b^T Q b = \int_{\alpha}^{\beta} \left(\sum_{i=1}^n b_i \varphi_i^{(2)}(x) \right)^2 dx \geq 0$$

for all $b \in \mathbb{R}^n$ and $b^T Q b = 0$ if and only if $b = 0_n$. Hence, $b^T Q b$ is convex in b and

$$\frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2 + \lambda b^T Q b$$

is convex in (a, b, y) .

On the other hand, the positive definiteness of Q implies the coersiveness of $b^T Q b$ in b . So,

$$\frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2 + \lambda b^T Q b$$

is coersive in (b, y) . The constraints $f(X_1) = y_1$ and $f(X_2) = y_2$ imply

$$\begin{aligned} a_0 + a_1 X_1 &= y_1 \\ a_0 + a_1 X_2 + (X_2 - X_1)_+^3 b_1 &= y_2, \end{aligned}$$

and hence,

$$\begin{aligned} a_0 + a_1 X_1 &= y_1 \\ a_1(X_2 - X_1) + (X_2 - X_1)_+^3 b_1 &= y_2, \end{aligned}$$

follow. Thus, $a_1^2 \rightarrow \infty$ implies $b_1^2 + y_2^2 \rightarrow \infty$, and $a_0^2 \rightarrow \infty$ implies $a_1^2 + y_1^2 \rightarrow \infty$. Therefore,

$$\frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2 + \lambda b^T Q b$$

is coersive in (a, b, y) .

Since $\frac{1}{n} \sum_{i=1}^n (Y_i - y_i)^2 + \lambda b^T Q b$ is coersive and continuous in (a, b, y) , and the feasible region of Problem (7) is non-empty ($a = (0, 0), b = 0_n, y = 0_n$ is a feasible solution) and closed, the solution to Problem (7) always exists (see, for example, Proposition 7.3.1 and Theorem 7.3.7 in pp 216 and 217 of Kurdila and Zabaranin (2005)), which completes the proof of Proposition 1. \square

3. Numerical Results

In this section, we examine the numerical performance of the proposed estimator in three settings: i) $f_* : [0, 1] \rightarrow \mathbb{R}$ is given by $f_*(x) = x^2$ for $x \in [0, 1]$, ii) $f_*(x)$ is the price of a European call option when the strike price is $x \in [0, 1]$, and iii) $f_*(x)$ is the steady-state mean waiting time of a customer in a single server queue when the service rate is $x \in [1.2, 1.7]$. In each of these settings, we compute the proposed estimator along with three competing estimators, i.e., the convex regression estimator, the penalized least squares estimator \hat{f}_n , and an estimator \hat{h}_n that combines convex regression and penalized least squares regression in a two-step procedure. To compute \hat{h}_n , we take the penalized least squares estimator \hat{f}_n , which is a smooth estimator of f_* , and compute its convex regression estimator by solving (2) with the Y_i 's replaced by the $\hat{f}_n(X_i)$'s. The proposed estimator, the convex regression estimator, and the penalized least squares estimator are computed by solving (7), (2), and (3). When solving the quadratic programming formulations (7) and (2), we use CVX (Grant and Boyd (2014)), a software package designed to solve convex programs. When solving (3) and (7), we choose λ using the cross validation method. In particular, to find λ in (3), we define $f_\lambda^{[k]}$ by the minimizer of

$$\frac{1}{n} \sum_{i=1, i \neq k}^n (Y_i - f(X_i))^2 + \lambda \int_\alpha^\beta \{f^{(2)}(x)\}^2 dx$$

over $f \in \mathcal{F}$. The cross validation function $V(\lambda)$ is defined by

$$V(\lambda) = \frac{1}{n} \sum_{k=1}^n (Y_k - f_\lambda^{[k]}(X_k))^2,$$

and λ_* is chosen as the minimizer of $V(\lambda)$ over $\lambda \in \{10^{-6}, 10^{-4}, 10^{-2}, 1, 10^2, 10^4, 10^6\}$; see p. 47 of Wabha (1990) for details. $\lambda = \lambda_*$ is used in (3) when computing the penalized least squares estimator. The smoothing parameter λ in (7) for the proposed estimator is chosen in a similar fashion.

Numerical results in Sections 3.1, 3.2, and 3.3 indicate that the proposed estimator converges to the true function f_* as $n \rightarrow \infty$ and that the proposed estimator displays superior performance compared to other competing estimators.

We conducted all simulations using a 64-bit computer with an Intel(R) Core(TM) i7-6700K CPU at 4 GHz and a RAM of 32 GB, and programmed all simulations in MATLAB R2018a.

3.1 Stylized Model

We consider the case where $f_* : [0, 1] \rightarrow \mathbb{R}$ is defined by $f_*(x) = x^2$ for $x \in [0, 1]$, $X_i = i/n - 1/(2n)$ for $1 \leq i \leq n$, and $Y_i = f_*(X_i) + \epsilon_i$ for $1 \leq i \leq n$, where the ϵ_i 's are iid normal random variables with a mean of 0 and a standard deviation of 0.15.

We generated $((X_i, Y_i) : 1 \leq i \leq n)$, and computed the proposed estimator, the convex regression estimator, the penalized least squares estimator, and \hat{h}_n . To measure the accuracy of the proposed estimator, we computed the empirical integrated

mean square error (EIMSE) between the true function f_* and the proposed estimator \hat{g}_n as follows:

$$\frac{1}{n} \sum_{i=1}^n (\hat{g}_n(X_i) - f_*(X_i))^2.$$

The EIMSE values for other competing estimators are computed in a similar fashion. We repeated this procedure 100 times independently, generating 100 copies of the EIMSE value for each of the estimators. Using these 100 copies, we then computed the 95% confidence interval of the EIMSE value. The 95% confidence intervals are reported in Table 1 for a variety of n values. The proposed estimator produces lower EIMSE values than other competing estimators.

Table 1. The 95% confidence intervals of the EIMSE values when $f_*(x) = x^2$ for $x \in [0, 1]$

| 95% confidence interval of the EIMSE value for | | | | |
|--|--------------------|-------------------|-------------------------|---|
| n | Proposed estimator | Convex regression | Penalized least squares | Penalized least squares + Convex regression |
| 10 | 0.0060 ± 0.0010 | 0.0088 ± 0.0013 | 0.0069 ± 0.0011 | 0.0066 ± 0.0011 |
| 20 | 0.0034 ± 0.0004 | 0.0047 ± 0.0005 | 0.0038 ± 0.0005 | 0.0036 ± 0.0004 |
| 30 | 0.0022 ± 0.0003 | 0.0032 ± 0.0004 | 0.0025 ± 0.0003 | 0.0023 ± 0.0003 |
| 40 | 0.0020 ± 0.0003 | 0.0030 ± 0.0003 | 0.0022 ± 0.0003 | 0.0021 ± 0.0003 |
| 50 | 0.0014 ± 0.0002 | 0.0024 ± 0.0003 | 0.0016 ± 0.0003 | 0.0015 ± 0.0003 |
| 60 | 0.0012 ± 0.0002 | 0.0022 ± 0.0003 | 0.0013 ± 0.0002 | 0.0013 ± 0.0002 |

3.2 Call Option

We consider the case where $f_*(x)$ is the price of a European call option with the strike price $x \in (0, 1)$ when the current price of the underlying stock is 0.5, the risk-free annual interest rate is 0.05, the stock price volatility is 0.3, and the time to maturity of the option is 1 year. We set $X_i = i/n - 1/(2n)$ for $1 \leq i \leq n$. For each $i \in \{1, \dots, n\}$, we assumed that the strike price of the option is X_i , generated a sample path $(S_t : 0 \leq t \leq 1)$ of a geometric Brownian motion up to time 1 with a drift of 0.05 and a volatility of 0.3, and computed the price of the option Y_i by computing $\exp(-r) \max(0, S_1 - X_i)$, where S_1 is the value of the geometric Brownian motion we generated at time 1.

Using $((X_i, Y_i) : 1 \leq i \leq n)$, we computed the proposed estimator, the convex regression estimator, the penalized least squares estimator, \hat{h}_n , and their EIMSE values. The 95% confidence intervals of these EIMSE values, based on 100 iid replications, are reported in Table 2 for a variety of n values. The proposed estimator produces lower EIMSE values than other competing estimators.

Table 2. The 95% confidence intervals of the EIMSE values when $f_*(x)$ is the price of a European call option with the strike price $x \in (0, 1)$

| 95% confidence interval of the EIMSE value for | | | | |
|--|--------------------|-------------------|-------------------------|---|
| n | Proposed estimator | Convex regression | Penalized least squares | Penalized least squares + Convex regression |
| 10 | 0.0034 ± 0.0008 | 0.0044 ± 0.0009 | 0.0039 ± 0.0008 | 0.0038 ± 0.0008 |
| 20 | 0.0018 ± 0.0004 | 0.0027 ± 0.0006 | 0.0020 ± 0.0004 | 0.0019 ± 0.0004 |
| 30 | 0.0015 ± 0.0003 | 0.0021 ± 0.0004 | 0.0018 ± 0.0003 | 0.0016 ± 0.0003 |
| 40 | 0.0010 ± 0.0002 | 0.0017 ± 0.0004 | 0.0011 ± 0.0002 | 0.0011 ± 0.0002 |
| 50 | 0.0008 ± 0.0001 | 0.0013 ± 0.0002 | 0.0009 ± 0.0001 | 0.0008 ± 0.0001 |
| 60 | 0.0007 ± 0.0001 | 0.0011 ± 0.0002 | 0.0008 ± 0.0001 | 0.0007 ± 0.0001 |

3.3 M/M/1 Queue

We consider the case where $f_*(x)$ is the steady-state mean waiting time of a customer in a single server first come/first serve queue with infinite capacity buffer, unit arrival rate, and a service rate of $x \in [1.2, 1.7]$ when the service times are iid exponential random variables, and the interarrival times are iid exponential random variables. We set $X_i = 1.2 + i/(2n) - 1/(4n)$ for $1 \leq i \leq n$. For each $i \in \{1, \dots, n\}$, we computed the average of the waiting times of the first 500 customers

arriving at the queue when the service rate is X_i and the queue is initialized empty and idle. We repeated this procedure 10 times independently, generated 10 averages, and set Y_i equal to the average of the 10 averages.

Using $((X_i, Y_i) : 1 \leq i \leq n)$, we computed the proposed estimator, the convex regression estimator, the penalized least squares estimator, \hat{h}_n , and their EIMSE values. The 95% confidence intervals of these EIMSE values, based on 200 iid replications, are reported in Table 3 for a variety of n values. The proposed estimator produces lower EIMSE values than other competing estimators.

Table 3. The 95% confidence intervals of the EIMSE values when $f_*(x)$ is the steady-state mean waiting time of a customer in an M/M/1 queue

| n | 95% confidence interval of the EIMSE value for | | | |
|-----|--|-------------------|-------------------------|---|
| | Proposed estimator | Convex regression | Penalized least squares | Penalized least squares + Convex regression |
| 10 | 0.0350 ± 0.0052 | 0.0472 ± 0.0111 | 0.0405 ± 0.0054 | 0.0387 ± 0.0054 |
| 20 | 0.0220 ± 0.0028 | 0.0286 ± 0.0044 | 0.0250 ± 0.0030 | 0.0234 ± 0.0029 |
| 30 | 0.0190 ± 0.0022 | 0.0253 ± 0.0028 | 0.0215 ± 0.0023 | 0.0200 ± 0.0023 |
| 40 | 0.0160 ± 0.0016 | 0.0226 ± 0.0032 | 0.0180 ± 0.0017 | 0.0169 ± 0.0016 |
| 50 | 0.0148 ± 0.0014 | 0.0199 ± 0.0018 | 0.0162 ± 0.0015 | 0.0153 ± 0.0014 |
| 60 | 0.0134 ± 0.0013 | 0.0168 ± 0.0015 | 0.0147 ± 0.0014 | 0.0138 ± 0.0013 |

4. Conclusions

In this paper, we proposed a new method for estimating a smooth and convex regression function. The proposed estimator can be easily computed by solving a convex quadratic programming problem and shows superior numerical performance compared to existing estimators. Possible future research topics include establishing the consistency of the proposed estimator as $n \rightarrow \infty$ and incorporating other shape restrictions, such as monotonicity, into the proposed estimator.

References

Ait-Sahalia, Y., & Duarte, J. (2003). Nonparametric option pricing under shape restrictions. *J. Econometrics*, 116, 9-47. [https://doi.org/10.1016/S0304-4076\(03\)00102-7](https://doi.org/10.1016/S0304-4076(03)00102-7)

Dole, D. (1999). CoSmo: A constrained scatterplot smoother for estimating convex, monotonic transformations. *J. Bus. Econom. Statist.*, 17, 444-455. <https://doi.org/10.1080/07350015.1999.10524832>

Du, P., Parmeter, C. F., & Racine, J. S. (2013). Nonparametric kernel regression with multiple predictors and multiple shape constraints. *Statist. Sinica*, 23, 1347-1371.

Grant, M., & Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.1. Retrieved May 2020, from <http://cvxr.com/cvx>

Green, P., & Silverman, B. (1994). *Nonparametric regression and generalized linear models: A roughness penalty approach*. Boca Raton: Chapman & Hall/CRC Press. <https://doi.org/10.1007/978-1-4899-4473-3>

Greville, T. N. E. (1969). *Theory and applications of spline functions*. New York: Academic Press.

Hall, P., & Huang, L-S. (2001). Nonparametric kernel regression subject to monotonicity constraints. *Ann. Statist.*, 29, 624-647. <https://doi.org/10.1214/aos/1009210683>

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference and prediction*. New York: Springer. <https://doi.org/10.1007/978-0-387-84858-7>

Hildreth, C. (1954). Point Estimates of Ordinates of Concave Functions. *J. Amer. Statist. Assoc.*, 49, 598-619. <https://doi.org/10.1080/01621459.1954.10483523>

Kurdila, A. J., & Zabrankin, M. (2005). *Convex functional analysis (Systems & control: Foundations & applications)* Switzerland: Birkhäuser.

Li, Z., Liu, G., & Li, Q. (2017). Nonparametric Knn estimation with monotone constraints. *Econometric Rev.*, 36, 988-1006. <https://doi.org/10.1080/07474938.2017.1307904>

Mammen, E. (1991). Estimating a smooth monotone regression function. *Ann. Statist.*, 19, 724-740. <https://doi.org/10.1214/aos/1176348117>

- Mammen, E., Marron, J. S., Turlach, B. A., & Wand, M. P. (2001). A general projection framework for constrained smoothing. *Statist. Sci.*, *16*, 232-248. <https://doi.org/10.1214/ss/1009213727>
- Mammen, E., & Thomas-Agnan, C. (1999). Smoothing splines and shape restrictions. *Scand. J. Stat.*, *26*, 239-252. <https://doi.org/10.1111/1467-9469.00147>
- Meyer, M. C. (2008). Inference using shape-restricted regression splines. *Ann. Statist.*, *2*, 1013-1033. <https://doi.org/10.1214/08-AOAS167>
- Sysoev, O., & Burdakov, O. (2019). A smoothed monotonic regression via L_2 regularization. *Knowledge and Information Systems*, *59*, 197-218. <https://doi.org/10.1007/s10115-018-1201-2>
- Yacine, Ä-S., & Duarte, J. (2003). Nonparametric option pricing under shape restrictions. *Journal of Economics*, *116*, 9-47. [https://doi.org/10.1016/S0304-4076\(03\)00102-7](https://doi.org/10.1016/S0304-4076(03)00102-7)
- Yagi, D., Chen, Y., Johnson, A. L., & Kuosmanen, T. (2020). Shape constrained kernel-weighted least squares: Estimating production functions for Chilean manufacturing industries. *J. Bus. Econom. Statist.*, *38*, 43-54. <https://doi.org/10.1080/07350015.2018.1431128>
- Ye, Y., & Tse, E. (1984). An extension of Karmarkar's projective algorithm for convex quadratic programming. *Math. Program.*, *44*, 157-179. <https://doi.org/10.1007/BF01587086>

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/>).