Communications of the Korean Statistical Society Vol. 15, No. 5, 2008, pp. 675–685

Sequential Shape Modification for Monotone Convex Function: $\mathcal{L}_2$ Monotonization and Uniform Convexification

Johan Lim[1), Sungim Lee[2)

Abstract

This paper studies two sequential procedures to estimate a monotone convex function using $\mathcal{L}_2$ monotonization and uniform convexification; one, denoted by FMSC, monotonizes the data first and then, convexifies the monotone estimate; the other, denoted by FCSM, first convexifies the data and then monotonizes the convex estimate. We show that two shape modifiers are not commutable and so does FMSC and FCSM. We compare them numerically in uniform error(UE) and integrated mean squared error(IMSE). The results show that FMSC has smaller uniform error(UE) and integrated mean squared error(IMSE) than those of FCSC.

Keywords: Commutability, $\mathcal{L}_2$ monotonization, monotone convex function, sequential estimation, uniform convexification.

1. Introduction

Suppose the observations are from the regression model

$$Y_i = f(X_i) + \sigma\epsilon_i, \tag{1.1}$$

where f is a monotone convex function from an interval $\Omega \subseteq \mathbb{R}$ into $\mathbb{R}$, ϵ_i are independent and identically distributed(i.i.d.) random variables with zero mean and unit variance and $\sigma > 0$. This paper concerns the problem of estimating of f, using the samples $(x_1, y_1), \ldots, (x_n, y_n)$ from this model.

This paper studies sequential procedures to estimate the monotone convex function based on two shape modifiers $\mathcal{L}_2$ monotonization and uniform convexification, which are commonly used for the estimation of monotone convex function. For a given sample $\big\{(x_i, y_i)\big\}_{i=1}^n$, the $\mathcal{L}_2$ monotonization is defined as the solution to the infinite-dimensional least square problem

$$\text{minimize} \sum_{i=1}^n \big\{y_i - f(x_i)\big\}^2 \quad \text{subject to} \quad f \in \mathcal{F}_m, \tag{1.2}$$

1) Associate Professor, Department of Statistics, Seoul National University, San 56-1, Shinlim-Dong, Gwanak-Gu, Seoul 151-747, Korea. Correspondence: johanlim@snu.ac.kr
2) Assistant Professor, Department of Statistics, Dankook University, 126, Jukjeon-dong, Suji-gu, Yongin-si, Gyeonggi-do 448-701, Korea.

in which $\mathcal{F}_m$ is the set of all monotone functions. On the other hand, the uniform convexification is defined as the solution to the optimization problem

$$\text{minimize} \max_{i=1,\ldots,n} \left| y_i - f(x_i) \right| \quad \text{subject to } f \in \mathcal{F}_{\mathrm{c}},$$

in which $\mathcal{F}_{\mathrm{c}}$ is the set of all convex functions.

The sequential application of the above two shape modifiers are recently discussed in Lee *et al.* (2008). They study the procedure which first monotonizes the raw data and then convexifies the monotonized outcomes of the first stage. In below, we denote this $\mathrm{FMSC} - r$ which implies the application of FMSC to raw data. They prove the global convergence rate of the isotonic(monootone) regression estimate and show the estimate by the given sequential shape modification converges at least as fast as the computed global convergence rate of the isotonic regression.

In modifying shape sequentially, a very next question we have is the application order of two shape modifiers. To be specific, we could consider two types of procedures according to the application order of two shape modifiers; first monotonization and second convexification(FMSC) which is considered in Lee *et al.* (2008), or first convexification and second monotonization(FCSM). We do question that whether the estimates by FMSC and FCSM are equal to each other or not. Further, if they are not equal, which is better in estimation error such as uniform error(UE) or integrated mean squared error(IMSE); UE and IMSE are defined more specifically later.

To answer the questions, in this paper, we show that two shape modifiers are not commutable to each other and hence, the estimates by FMSC and FCSM are not equal to each other in general. We further compare them numerically in estimating a monotone convex function. We apply FMSC and FCSM to estimating raw data and modifying the shape of kernel smooth estimates. For each of these four trials, we use the following notations:

- FMSC: The sequential shape modifier that first monotonizes and second convexifixes.
- FCSM: The sequential shape modifier that first convexifies and second monotonizes.
- $\mathrm{FMSC} - r$: The procedure to apply FMSC to raw data.
- $\mathrm{FCSM} - r$: The procedure to apply FCSM to raw data.
- $\mathrm{FMSC} - s$: The procedure to apply FMSC to smooth estimate.
- $\mathrm{FCSM} - s$: The procedure to apply FCSM to smooth estimate.

The remainder of the paper is as follows. We shortly review two shape modifiers in Section 2. In Section 3, we study the commutability between two shape modifiers of interest. Section 4 numerically compare FMSC and FCSM in terms of UE and IMSE. Section 5 applies the methods to analyzing a real example. Section 6 concludes the paper.

2. Monotonization and Convexification

To lay a groundwork for the addressed sequential estimators, we study some results on $\mathcal{L}_2$ monotonization and uniform convexification.

2.1. $\mathcal{L}_2$ approximation with monotone functions

Isotonic(monotone) regression which solves (1.2) dates back to 1950s (Brunk, 1955; Brunk, 1958) and excellent exposition of the issue appears in Barlow (1972) and later on in Robertson *et al.* (1988). Here we briefly review some basic results of isotonic regression.

The solution to the minimization problem (1.2) is piecewise linear and is given as the slope of the greatest convex minorant of the cumulative sum diagram by the points $\big(x_k,\ \sum_{i=1}^{k} y_i\big)$, $k = 1, 2, \ldots, n$. The slope is characterized as

$$\widehat{f}_n^{\mathrm{IS}}(x_k) = \max_{i \le k} \min_{j \ge k} P\big(i,\, j\big) = \min_{j \le k} \max_{i \ge k} P\big(i,\, j\big) = \min_{j \ge k} \max_{i \le k} P\big(i,\, j\big), \tag{2.1}$$

where

$$P\big(i,\, j\big) = \frac{1}{j - i + 1} \sum_{k=i}^{j} y_k.$$

The pooled adjacent violators algorithm can also characterize the solution which iteratively modifies the adjacent violators by pooling them. In the algorithms above, $\widehat{f}_n^{\mathrm{IS}}$ can be completely characterized with $O(n)$ operations; see Barlow (1972).

2.2. Uniform approximation with convex functions

We review some results on the constrained uniform approximation problem for a function f from a closed interval Ω into $\mathbb{R}$:

$$\text{minimize } \sup_{x \in \Omega} \big| f(x) - \tilde{f}(x) \big| \quad \text{subject to } \tilde{f} \in \mathcal{F}_{\mathrm{c}}. \tag{2.2}$$

This is an infinite-dimensional optimization problem, but we can find an analytic solution, as will be seen soon. We refer the reader to Kim and Lim (2006) and Lee *et al.* (2008) for a more extensive discussion.

For the moment, we assume that $f : \Omega \to \mathbb{R}$ is piecewise linear and continuous with n break points $x_1, \ldots, x_n$. Then, the convex envelope $f_{\mathbf{env}}$ is defined by the *unique* piecewise linear and continuous function that satisfies the interpolation condition:

$$f_{\mathbf{env}}(u_i) = z_i, \quad i = 1, \ldots, r,$$

where $\{(u_i, z_i) : i = 1, \ldots, r\}$ is the set of the vertices of the lower convex hull of the break points of f. Then, the function

$$h = f_{\mathbf{env}} + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2} = f_{\mathbf{env}} + \frac{\max_{i=1,\ldots,r} \{y_i - f_{\mathbf{env}}(x_i)\}}{2}$$

is a solution to the constrained uniform approximation problem (2.2) (See Kim and Lim (2006) for the proof). Note that the break points of this solution and hence the piecewise linear solution itself, can be characterized in $O(n \log n)$ operations by applying the convex hull algorithm to the break points of f.

When used as an approximation to the function f, the solution h tends to over-approximate around the boundaries of the end points. To solve this difficulty, Kim and Lim (2006) suggest a modification of the function $f_{\mathbf{env}} + \|f - f_{\mathbf{env}}\|_{\infty,\Omega}/2$ to reduce the overestimation error around the boundaries of the end points.

- Boundary Correction Algorithm

1. Find the lower envelope $f_{\mathbf{env}}$ of f.
2. Find the points x_l and x_r which satisfies

$$\begin{aligned} f_{\mathbf{env}}(x_l) + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2} &= f(x_l), \\ f_{\mathbf{env}}(x) + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2} &> f(x), \quad x \le x_l, \\ f_{\mathbf{env}}(x_r) + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2} &= f(x_r), \\ f_{\mathbf{env}}(x) + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2} &> f(x), \quad x \ge x_r. \end{aligned}$$

3. Of the break points of $f_{\mathbf{env}}$, find the left adjacent point $\tilde{x}_l$ of x_l and the right adjacent point of $\tilde{x}_r$.
4. Define the function $f^\star : \Omega \to \mathbb{R}$ by

$$f^\star(x) = \begin{cases} \max\{s_l(x),\, f_{\mathbf{env}}(x)\}, & \text{if } x \le x_l, \\ f_{\mathbf{env}}(x) + \dfrac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2}, & \text{if } x_l \le x \le x_r, \\ \max\{s_r(x),\, f_{\mathbf{env}}(x)\}, & \text{if } x \ge x_r, \end{cases}$$

in which s_l is the affine function that passes through the two points

$$\left(x_l,\, f_{\mathbf{env}}(x_l) + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2}\right), \qquad \left(\tilde{x}_l,\, f_{\mathbf{env}}(\tilde{x}_l) + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2}\right)$$

and s_r is the affine function that passes through the two points

$$\left(x_r,\, f_{\mathbf{env}}(x_r) + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2}\right), \qquad \left(\tilde{x}_r,\, f_{\mathbf{env}}(\tilde{x}_r) + \frac{\|f - f_{\mathbf{env}}\|_{\infty,\Omega}}{2}\right).$$

The function $f^\star$ generated by the algorithm above is piecewise linear and hence is completely characterized by its break points that can be computed by the convex hull algorithm in $O(n \log n)$ operations. In what follows, the (uniform) convexification of a function f means applying the boundary correction algorithm to f.

3. Commutability

In this section, we prove that FMSC and FCSM are not equal to each other, unless the data points are monotone or convex.

We first introduce some definitions and notations which will be used in the remainder of the paper.

Let the observations $(x_i,\, y_i)$ for $i = 1, 2, \ldots, n$, be independently from the model (1.1). Let $\mathrm{Y}_n = (y_1, y_2, \ldots, y_n)$ and $\mathrm{X}_n = (x_1, x_2, \ldots, x_n)$ with $x_1 \le x_2 \le \cdots \le x_n$. By the monotonicity of Y_n on X_n, we imply that $(x_i - x_j)(y_i - y_j) \ge 0$ for any i and j, *i.e.*, the piecewise linear function that connects the points $\{(x_i, y_i)\}$ is monotone. Similarly, we say that the observations are *convex* on X_n if

$$(y_i - y_j)(x_k - x_j) \geq (y_k - y_j)(x_i - x_j),$$

whenever $x_i < x_j < x_k$, *i.e.*, the piecewise linear function, connecting the observed data points, is convex.

The following two lemmas from Lee *et al.* (2008) show that for any $p \geq 1$, the $\mathcal{L}_p$ convexification(or monotonization) of a monotone(or convex) function is still monotone(or convex).

Lemma 3.1 Let $\mathcal{F}_{\mathrm{c}}(X_n)$ be the set of functions from $\mathbb{R}$ to $\mathbb{R}$ with the convex property on X_n and let $\widehat{f}_c^*$ be the solution to the optimization problem

$$\text{minimize} \sum_{i=1}^{n} \left|y_i - f_c(x_i)\right|^p \quad \text{subject to } f_c \in \mathcal{F}_{\mathrm{c}}(\mathrm{X}_n), \tag{3.1}$$

where $1 \leq p \leq \infty$. Suppose that Y_n is monotone on X_n. Then, $\widehat{f}_c^*$ is monotone on X_n.

Lemma 3.2 Let $\mathcal{F}_{\mathrm{m}}$ be the set of functions from $\mathbb{R}$ to $\mathbb{R}$ with the monotone property on X_n and let $\widehat{f}_m^*$ be the solution to the optimization problem

$$\text{minimize} \sum_{i=1}^{n} \left|y_i - f_m(x_i)\right|^p \quad \text{subject to } f_m \in \mathcal{F}_{\mathrm{m}}(\mathrm{X}_n), \tag{3.2}$$

where $1 \leq p \leq \infty$. Then, $\widehat{f}_m^*$ is convex on X_n, if Y_n is convex on X_n.

The preceding two lemmas provide a sufficient condition when both FCSM and FMSC yields the same monotone convex function. However this is not true in general.

Theorem 3.1 Both FCSM and FMSC applied to the data points $(\mathrm{X}_n, \mathrm{Y}_n)$ yield the same monotone convex function if Y_n is monotone or convex on X_n. However, the estimate by FCSM is not equal to that by FMSC in general.

4. Comparison

We carry out a simulation study to compare FMSC and FCSM. In the simulation study, the design points $x_1, \ldots, x_n$ are taken as $x_i = i/n$ and the response data y_i are generated from the model (1.1). We consider three true functions: $f(x) = 1$ (neither strictly monotone nor strictly convex), $f(x) = 2x$ (strictly monotone but not strictly convex), $f(x) = 2x^2$ (strictly monotone and convex) on $\Omega = [0, 1]$. Two error distributions with mean 0 and variance σ^2 are considered for ϵ_i: the Gaussian and the double exponential distribution. The double exponential distribution has the form $f(x) = (1/2\sigma)\exp(-|x|/\sigma)$. For each pair of the true function and the error distribution, we generate 100 data sets and apply the proposed sequential methods to estimating the raw data, FMSC $- r$ and FCSM $- r$. We also apply FMSC and FCSM to modifying the shape of kernel estimates, FMSC $- s$ and FCSM $- s$. The kernel estimates are computed using Gaussian kernel and

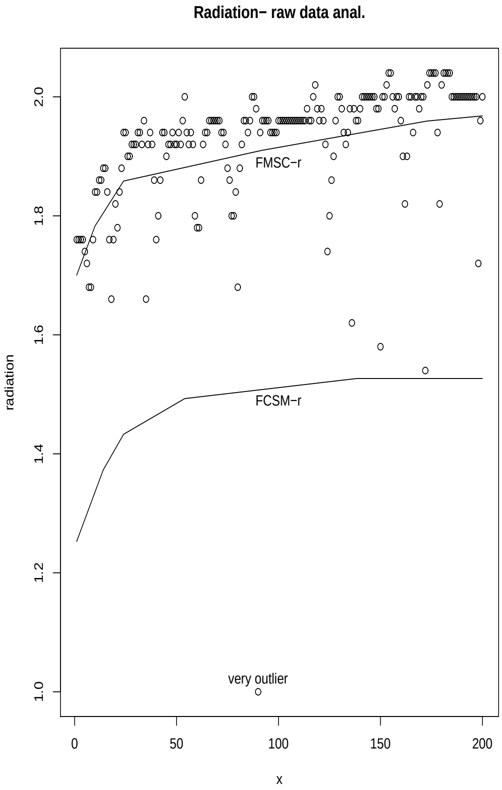
Table 4.1: Comparison between FMSC and FCSM based on 100 data sets of size $n = 50$. In the table, “Gauss” and “DE” implies Gaussian error distribution and double exponential error distribution, respectively. The numbers in the parenthesis are the SEs. Here, $f_1(x) = 1$, $f_2(x) = 2x$ and $f_3(x) = 2x^2$

		$\text{FMSC}-r$		$\text{FCSM}-r$		$\text{FMSC}-s$		$\text{FCSM}-s$	
G	σ	UE	IMSE	UE	IMSE	UE	IMSE	UE	IMSE
f_1	0.1	0.0554	0.0024	0.2388	0.0196	0.0662	0.0027	0.0825	0.0034
		(0.0078)	(0.0024)	(0.0130)	(0.0054)	(0.0057)	(0.0004)	(0.0073)	(0.0005)
	0.3	0.1128	0.0093	0.5027	0.0796	0.1198	0.0088	0.1572	0.0108
		(0.0102)	(0.0033)	(0.0168)	(0.0075)	(0.0087)	(0.0010)	(0.0146)	(0.0013)
	0.5	0.1578	0.0153	0.6257	0.1128	0.1499	0.0138	0.1921	0.0179
		(0.0030)	(0.0009)	(0.0030)	(0.0009)	(0.0125)	(0.0018)	(0.0166)	(0.0027)
f_2	0.1	0.2656	0.0159	1.1740	0.3122	0.2063	0.0087	0.2171	0.0091
		(0.0120)	(0.0040)	(0.0143)	(0.0091)	(0.0098)	(0.0007)	(0.0097)	(0.0007)
	0.3	0.4077	0.0534	1.3326	0.3510	0.2718	0.0204	0.2874	0.0211
		(0.0143)	(0.0075)	(0.0183)	(0.0122)	(0.0107)	(0.0017)	(0.0110)	(0.0017)
	0.5	0.4592	0.0716	1.4781	0.3920	0.3058	0.0272	0.3268	0.0285
		(0.0157)	(0.0091)	(0.0205)	(0.0155)	(0.0157)	(0.0024)	(0.0172)	(0.0025)
f_3	0.1	0.2688	0.0284	1.0236	0.3465	0.1744	0.0098	0.1829	0.0098
		(0.0110)	(0.0058)	(0.0122)	(0.0097)	(0.0059)	(0.0009)	(0.0057)	(0.0008)
	0.3	0.4190	0.0611	1.1609	0.3543	0.2491	0.0240	0.2564	0.0218
		(0.0122)	(0.0068)	(0.0177)	(0.0117)	(0.0101)	(0.0023)	(0.0102)	(0.0020)
	0.5	0.4641	0.0798	1.3257	0.3495	0.2913	0.0285	0.2928	0.0261
		(0.0146)	(0.0085)	(0.0205)	(0.0135)	(0.0116)	(0.0028)	(0.0111)	(0.0024)

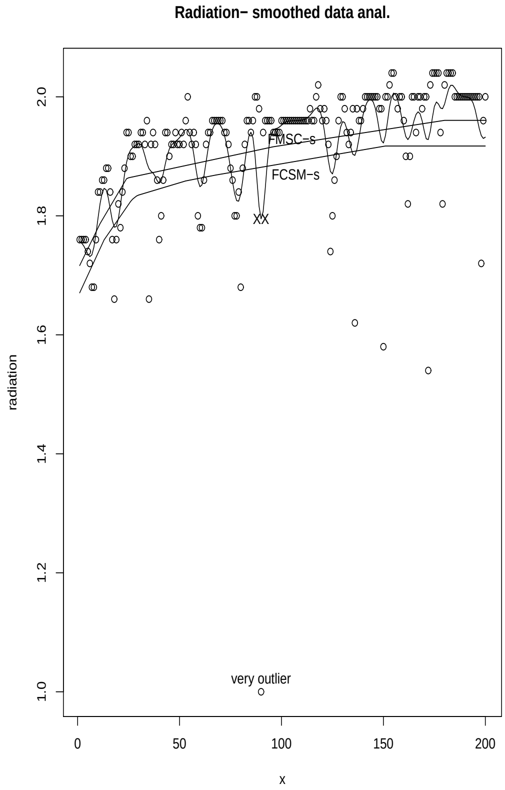
		$\text{FMSC}-r$		$\text{FCSM}-r$		$\text{FMSC}-s$		$\text{FCSM}-s$	
DE	σ	UE	IMSE	UE	IMSE	UE	IMSE	UE	IMSE
f_1	0.1	0.0768	0.0037	0.3892	0.0708	0.0553	0.0020	0.0792	0.0032
		(0.0077)	(0.0020)	(0.0135)	(0.0067)	(0.0039)	(0.0002)	(0.0065)	(0.0005)
	0.3	0.1232	0.0091	0.6424	0.1651	0.1203	0.0092	0.1516	0.0129
		(0.0103)	(0.0038)	(0.0189)	(0.0126)	(0.0097)	(0.0012)	(0.0124)	(0.0020)
	0.5	0.1429	0.0153	0.8773	0.3006	0.1515	0.0139	0.2018	0.0193
		(0.0115)	(0.0045)	(0.0208)	(0.0162)	(0.0122)	(0.0016)	(0.0134)	(0.0022)
f_2	0.1	0.2605	0.0186	1.2588	0.3363	0.1851	0.0072	0.1969	0.0078
		(0.0121)	(0.0045)	(0.0168)	(0.0111)	(0.0078)	(0.0005)	(0.0073)	(0.0006)
	0.3	0.3997	0.0529	1.5298	0.4490	0.2957	0.0218	0.2998	0.0211
		(0.0163)	(0.0091)	(0.0217)	(0.0168)	(0.0140)	(0.0021)	(0.0131)	(0.0018)
	0.5	0.4534	0.0778	1.6943	0.5584	0.3401	0.0308	0.3524	0.0327
		(0.0153)	(0.0083)	(0.0237)	(0.0188)	(0.0145)	(0.0026)	(0.0144)	(0.0026)
f_3	0.1	0.3121	0.0423	1.1069	0.3787	0.1837	0.0103	0.1976	0.0107
		(0.0114)	(0.0056)	(0.0168)	(0.0117)	(0.0081)	(0.0009)	(0.0073)	(0.0009)
	0.3	0.4280	0.0758	1.4093	0.4210	0.2575	0.0231	0.2632	0.0220
		(0.0144)	(0.0081)	(0.0212)	(0.0160)	(0.0098)	(0.0022)	(0.0097)	(0.0018)
	0.5	0.4487	0.0712	1.5829	0.5798	0.2972	0.0336	0.3134	0.0390
		(0.0158)	(0.0092)	(0.0245)	(0.0196)	(0.0135)	(0.0028)	(0.0143)	(0.0045)

optimally chosen bandwidth by leave-one-out cross-validation. In each of 100 data sets, we compute the uniform error, $\max_{i=1}^n |\widehat{f}_n(x_i) - f(x_i)|$ and the integrated mean squared error(IMSE), $(1/n)\sum_{i=1}^n |\widehat{f}_n(x_i) - f(x_i)|^2$.

Table 4.1 summarizes the simulation results for $n = 50$ for three noise level $\sigma = 0.1, 0.2$ and 0.3. We also try $n = 20$ and $n = 100$ but do not report here, since the results are similar. The Table 4.1 shows that, in all cases, FMSC is superior to FCSM for both



(a) Raw radiation data.



(b) Smoothed radiation data.

Figure 5.1

estimating the raw data and modifying the shape of smooth nonparametric function estimate. Further, since the true functions are smooth, $\mathrm{FMSC}-s$ and $\mathrm{FCSM}-s$ has smaller errors in overall than the $\mathrm{FMSC}-r$ and $\mathrm{FCSM}-r$, respectively.

5. Example: Radiation Data Analysis

In this section, we illustrate the proposed procedures by analyzing the radiation data reported in Davies and Gather (1993). The data set is available from `http://lib.stat.cmu.edu/datasets/`. The data has 2001 observations of radiation level taken from a balloon about 30 kilometers above the surface of the earth. There is a non-decreasing concave trend in the data, that is degrading increment of the radiation with height. As stated in Davies and Gather (1991), there are multiple outliers that caused by the fact that the balloon slowly rotates, causing the ropes from which the measuring instrument is suspended to cut off the direct radiation from the sun. There outlier will make much of the difference between FMSC and FCSM. In below, we use the first 200 sequence of the entire data set.

We first apply the FMSC and FCSM to raw data and plot the results in Figure 5.1 (a). The 90^{th} observation is very outlier. The uniform convexification is known to be more sensitive to outliers than $\mathcal{L}_2$ monotonization. Thus, if we apply the uniform convexification first, then the first stage estimate is dragged down to the 90^{th} observation and in consequence, the FCSM $- r$ is too. The FMSC $- r$ is also influenced by the very outlier and is slightly smaller than main body of the data. However, it is less sensitive than the FCSM $- r$.

We then apply the two sequential shape modifier to smoothed data. The smoothing procedure is exactly same with that in numerical study in previous section. We plot the results in in Figure 5.1 (b). The very outlier still gives an influence to the smoothed estimates. It drag down the estimates(please see "XX" in the figure) and results in downward bias of both estimates. However, again, the bias of FMSC $- s$ is smaller than that of FCSM $- s$.

6. Conclusion

This paper studies two shape modifiers, $\mathcal{L}_2$ monotonization and uniform convexification. They provide two sequential shape modifiers to estimate a monotone convex function according the the order of application. We study commutability of two shape modifiers and show numerically FMSC provides better results that the other, FCSM, in UE and IMSE.

Acknowledgement

We are grateful to the referee and the associate editor for the suggestions which improved the quality of the paper. Johan Lim's research was supported by Research Settlement Fund for the new faculty of College of Natural Science, Seoul National University.

Appendix

Proof of Theorem 1

Sufficient condition for FMSC=FCSM

To see this, let $\widehat{f}_{\mathrm{FCSM}}$ and $\widehat{f}_{\mathrm{FMSC}}$ be the estimate obtained via FCSM and the one via FMSC, respectively. Suppose Y_n is monotone on X_n. Then, we can see from Lemma 2 that the solution $\widehat{f}_c^*$ to the problem

$$\text{minimize} \sum_{i=1}^{n} \left|f_c(x_i) - y_i\right|^p \quad \text{subject to } f_c \in \mathcal{F}_{\mathrm{c}}$$

is monotone. Thus, the monotonization of the first stage convex estimate, $\widehat{f}_c^*$ is $\widehat{f}_c^*$ itself, *i.e.*, $\widehat{f}_{\mathrm{FCSM}} = \widehat{f}_c^*$. On the other hand, since Y_n is monotone, the first stage monotone estimate, $\widehat{f}_m^*$ is $\mathrm{Y}_n = (y_1, y_2, \ldots, y_n)$ itself. Thus, the estimate by FMSC is given by the solution to (A.1), which is equal to $\widehat{f}_c^*$. In other words, $\widehat{f}_{\mathrm{FMSC}} = \widehat{f}_c^*$. Hence, $\widehat{f}_{\mathrm{FCSM}} = \widehat{f}_{\mathrm{FMSC}}$. Through similar arguments, we can show that if Y_n is convex on X_n, then both FCSM and FMSC applied to $(\mathrm{X}_n, \mathrm{Y}_n)$ lead to the same result.

FCSM is not equal to FMSC in general

In this section, by giving a counter example, we show that FCSM and FMSC applied to the data points $(\mathrm{X}_n,\ \mathrm{Y}_n)$ do not always provide same answer.

Consider the following data points

$$\big\{(x_1,\ y_1), \ldots, (x_5,\ y_5)\big\} = \big\{(0,\ 0),\ (1,\ -2),\ (2,\ 0),\ (3,\ -100),\ (4,\ 0)\big\}.$$

We first solve the monotone regression (1.2) for these data points. A solution $\widehat{f}_{\mathrm{m}}^*$ on $\Omega = [\,0,\ 4\,]$ is piecewise linear with break points at 0, 1, 2, 3, 4. At the break points, the function $\widehat{f}_{\mathrm{m}}^*$ takes it values as

$$\widehat{f}_{\mathrm{m}}^*(x) = \begin{cases} -\dfrac{51}{2}, & \text{if } x = 0, \ldots, 3, \\ \ \ 0\ , & \text{if } x = 4. \end{cases}$$

Note that the piecewise function $\widehat{f}_m^*(x)$ is convex on $\Omega = [\,0,\ 4\,]$. Thus, $\widehat{f}_{\mathrm{FMSC}}(x)$ is identical to $\widehat{f}_m^*(x)$ and unique. To find $\widehat{f}_{\mathrm{FCSM}}$, we find $\widehat{f}_c^*$ by solving the uniform approximation problem

$$\text{minimize} \max_{i=1,\ldots,5} \big|\ y_i - f(x_i)\big| \quad \text{subject to } f \in \mathcal{F}_{\mathrm{c}}.$$

Note that the unique best lower convex approximation to Y_n, say $\underline{f}_c^*$, is piecewise linear with break points only at $0, 1, 2, 3, 4$ and take its values at the points as

$$\underline{f}_c^*(x) = \begin{cases} \ \ 0\ , & \text{if } x = 0,\ 4, \\ -\dfrac{100}{3}, & \text{if } x = 1, \\ -\dfrac{200}{3}, & \text{if } x = 2, \\ -100\,, & \text{if } x = 3. \end{cases}$$

From the results in Section 2.2, the uniform error by a best uniform approximation is 100/3, a half of that by $\underline{f}_c^*(x)$.

Let $\widehat{f}_c^*(x)$ be a function in the set of best convex approximations to Y_n(on X_n), $\mathcal{C}(\mathrm{Y}_n, \mathcal{F}_c)$. Then, $-100/3 \le \widehat{f}_c^*(0),\ \widehat{f}_c^*(2) \le 100/3$ and $-400/3 \le \widehat{f}_c^*(3) \le -200/3$, since the uniform error is 100/3. Also, note from the convexity of $\widehat{f}_c^*(x)$ that

$$\widehat{f}_c^*(0) \ge \widehat{f}_c^*(3) + \Big\{\widehat{f}_c^*(3) - \widehat{f}_c^*(2)\Big\}\Big\{0 - \widehat{f}_c^*(2)\Big\}.$$

Here the minimum value of the right-hand side is 100/3 at $\widehat{f}_c^*(2) = -100/3$ and $\widehat{f}_c^*(3) = -200/3$. Hence, $\widehat{f}_c^*(0) = 100/3$ is the only possible value for the uniform error to be

100/3. In a similar way, we can show that $\widehat{f}_c^*(1) = 0$ and $\mathcal{C}(\mathrm{Y}_n, \mathcal{F}_c)$ consists of functions of the form

$$\widehat{f}_c^*(x) = \begin{cases} \dfrac{100}{3}, & \text{if } x = 0, \\ 0\ , & \text{if } x = 1, \\ -\dfrac{100}{3}, & \text{if } x = 2, \\ -\dfrac{200}{3}, & \text{if } x = 3, \\ c\ , & \text{if } x = 4, \end{cases}$$

for any $c \in [-100/3, 100/3]$. For each $\widehat{f} \in \mathcal{C}(\mathrm{Y}_n, \mathcal{F}_c)$, we can solve the associated projection problem to the class of monotone functions $\mathcal{F}_{\mathrm{m}}$ and find a monotone convex function. Let $\mathcal{C}(\mathcal{M}(\mathrm{Y}_n, \mathcal{F}_c))$ be the set of resulting monotone convex functions. Using the complete characterization (2.1) of the solution to the isotonization problem (1.2), we can see that any piecewise linear function $\widehat{f}_{\mathrm{FCSM}} \in \mathcal{C}(\mathcal{M}(\mathrm{Y}_n, \mathcal{F}_c))$ is of the form

$$\widehat{f}_{\mathrm{FCSM}}(x) = \begin{cases} -\dfrac{100}{3} + \dfrac{c}{3}, & \text{if } x = 0, \\ -\dfrac{100}{3} + \dfrac{c}{3}, & \text{if } x = 1, \\ -\dfrac{100}{3} + \dfrac{c}{3}, & \text{if } x = 2, \\ -\dfrac{100}{3} + \dfrac{c}{3} \left(\text{or } \dfrac{c}{2}\right), & \text{if } x = 3 \ \text{ and } \ c, \leq 0 \text{ (or } c \geq 0), \\ c, & \text{if } x = 4. \end{cases}$$

Now, it is evident that

$$\underline{f}_m^* \notin \mathcal{C}\big(\mathcal{M}\left(\mathrm{Y}_n, \mathcal{F}_c\right)\big).$$

This counterexample shows that two sequential shape modification procedures, FCSM and FMSC do not always yield the same result.

References

Barlow, R. E. (1972). *Statistical Inference under Order Restrictions,* John Wiley & Sons, New York.

Brunk, H. D. (1955). Maximum likelihood estimates of monotone parameters, *The Annals of the Mathematical Statistics,* **26**, 607–616.

Brunk, H. D. (1958). On the estimation of parameters restricted by inequalities, *The Annals of the Mathematical Statistics,* **29**, 437–454.

Davis, L. and Gather, U. (1993). The identification of multiple outliers, *Journal of the American Statistical Association,* **88**, 782–801.

Kim, S. J. and Lim, J. (2006). Uniform approximation and estimation of shape restricted functions, Preprint available upon request.

Lee, S., Lim, J., Kim, S. J. and Joo, Y. (2008). Estimating monotone convex functions via sequential shape modification, To appear in *Journal of Statistical Compuation and Simulation*.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference,* John Wiley & Sons, New York.

[Received June 2008, Accepted July 2008]