International Journal of Statistics and Probability; Vol. 6, No. 5; September 2017
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Nonparametric Tests for Convexity/Monotonicity/Positivity of Multivariate Functions with Noisy Observations

Eunji Lim[1] & Elizabeth Tavarez[2]

[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. E-mail: elim@kean.edu

Received: June 7, 2017 Accepted: July 3, 2017 Online Published: July 21, 2017

doi:10.5539/ijsp.v6n5p18 URL: https://doi.org/10.5539/ijsp.v6n5p18

Abstract

We propose a new method of testing for a function's convexity, monotonicity, or positivity, based on some noisy observations of the function made over a finite set $\mathcal{T}$ of points in the domain, where the observations can be made multiple times at each point in $\mathcal{T}$. One of the traditional approaches to the test of a function's shape characteristic is to fit a convex, a monotone, or a positive function, depending on the shape characteristic we wish to test for, to the data set minimizing the sum of squared errors, and to compute the sum of squared differences (SSD) between the fit and the data set. While the traditional approach proceeds by observing the SSD as the number of points in $\mathcal{T}$ increases to infinity, we propose observing the SSD as r, the number of observations taken at each point in $\mathcal{T}$, increases to infinity. This new way of observing the asymptotic behavior of the SSD leads to a test procedure that does not require the estimation of any additional parameters, and hence, is easy to implement. The proposed test procedure is proven to achieve a prescribed power as $r \to \infty$. Numerical examples illustrate that the proposed test successfully detects the convexity/monotonicity/positivity of a function, as well as the non-convexity/non-monotonicity/non-positivity of a function.

Keywords: convexity detection, monotonicity detection, positivity detection, hypothesis test, convex regression, isotonic regression, nonparametric test, multivariate regression

1. Introduction

The goal of this paper is to develop a hypothesis test for determining whether an unknown function $f_* : \mathbb{R}^d \to \mathbb{R}$ is convex, monotone, or positive using some noisy measurements of f_* at points $X_1, \ldots, X_n \in \mathbb{R}^d$.

The convexity/monotonicity/positivity of a function has significant implications in many areas of applications. In economics, the concavity of the utility curve as a function of a person's income implies that the marginal utility diminishes as income increases (p. 31 of Keynes, 1935). In the context of statistical inference, the monotonicity of a function that one wishes to estimate from noisy data indicates that one can fit a monotone function to the data set to reconstruct the underlying function (Barlow et al., 1972).

In this paper, we assume that we can observe noisy measurements of $f_*(X_i)$ for $1 \leq i \leq n$, and these observations can be made multiple times independently of each other. Thus, we are able to obtain the data set $((X_i, Y_{ij}) : 1 \leq i \leq n, 1 \leq j \leq r)$, where

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

the X_i's are either $\mathbb{R}^d$-valued random vectors or deterministic points in $\mathbb{R}^d$, and the ϵ_{ij}'s are independent and identically distributed (iid) random variables with $\mathbb{E}(\epsilon_{ij} \mid X_1, \ldots, X_n) = 0$ and $\mathbb{E}(\epsilon_{ij}^2 \mid X_1, \ldots, X_n) = \sigma^2$ for some $\sigma < \infty$.

To determine whether the function f_* is convex/monotone/positive or not, we consider the following pairs of the null and alternative hypotheses:

Null Hypothesis	Alternative Hypothesis	
$H_c^0 : f_* \notin \mathcal{F}_c$	$H_c^a : f_* \in \mathcal{F}_c$	for a test of convexity,
$H_m^0 : f_* \notin \mathcal{F}_m$	$H_m^a : f_* \in \mathcal{F}_m$	for a test of monotonicity,
$H_p^0 : f_* \notin \mathcal{F}_p$	$H_p^a : f_* \in \mathcal{F}_p$	for a test of positivity,

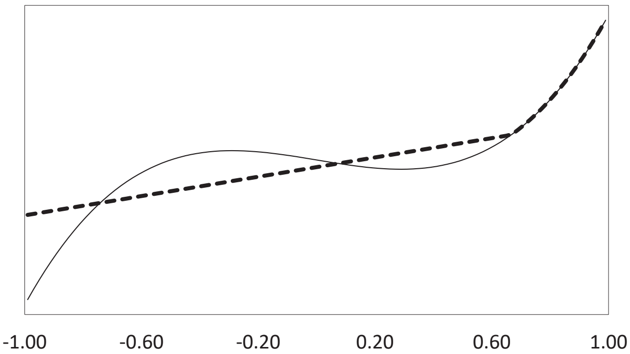

Figure 1. The solid line is f_* and the dashed line is the function to which $\hat{f}_c$ converges as $r \to \infty$.

where

$$\begin{aligned}
\mathcal{F}_c &= \{g : \mathbb{R}^d \to \mathbb{R} \mid g(tv + (1-t)w) \le tg(v) + (1-t)g(w) \text{ for } v, w \in \mathbb{R}^d \text{ and } t \in [0,1]\}, \\
\mathcal{F}_m &= \{g : \mathbb{R}^d \to \mathbb{R} \mid g(v) \le g(w) \text{ for } v = (v(1), \ldots, v(d)) \text{ and } w = (w(1), \ldots, w(d)) \in \mathbb{R}^d \text{ with } v(k) \le w(k), 1 \le j \le d\} \\
\mathcal{F}_p &= \{g : \mathbb{R}^d \to \mathbb{R} \mid g(v) \ge 0 \text{ for } v \in \mathbb{R}^d\}.
\end{aligned}$$

When the context is clear, H_c^0, H_m^0, and H_p^0 will be referred to as the null hypotheses and H_c^a, H_m^a, and H_p^a the alternative hypotheses.

The proposed test procedure is described as follows.

Test of Convexity: When one suspects that the unknown function f_* is convex, a natural step to take in order to estimate f_* is to fit a convex function $\hat{f}_c : \mathbb{R}^d \to \mathbb{R}$ to the data set $((X_i, \overline{Y}_i) : 1 \le i \le n)$, where $\overline{Y}_i = \sum_{j=1}^r Y_{ij}/r$ for $1 \le i \le n$, which minimizes the sum of squared distances between the fit and the data set. The fitted function $\hat{f}_c$ can be defined by the solution to the following infinite-dimensional minimization problem:

$$\text{Minimize} \qquad \frac{1}{n}\sum_{i=1}^n \left(\overline{Y}_i - g(X_i)\right)^2 \tag{1}$$

over $g \in \mathcal{F}_c$. Problem (1) turns out to be equivalent to the following finite-dimensional quadratic programming problem:

$$\begin{aligned}
&\text{Minimize} \qquad \frac{1}{n}\sum_{i=1}^n \left(\overline{Y}_i - g(X_i)\right)^2 \qquad (2)\\
&\text{Subject to} \qquad g(X_j) \ge g(X_i) + \xi_i^T (X_j - X_i), \quad 1 \le i, j \le n
\end{aligned}$$

in the decision variables $g(X_i) \in \mathbb{R}$ and $\xi_i \in \mathbb{R}^d$ for $1 \le i \le n$, where ξ_i^T denotes the transpose of ξ_i (Lemma 2.5 of Seijo & Sen, 2011).

When f_* is truly convex, the mean square error (MSE), defined by

$$\frac{1}{n}\sum_{i=1}^n \left(\overline{Y}_i - \hat{f}_c(X_i)\right)^2,$$

converges to 0 as $r \to \infty$ because $\sum_{i=1}^n (\overline{Y}_i - \hat{f}_c(X_i))^2/n \le \sum_{i=1}^n (\overline{Y}_i - f_*(X_i))^2/n = \sum_{i=1}^n \bar{\epsilon}_i/n \to \infty$ as $n \to \infty$ by the minimizing property of $\hat{f}_c$ and the strong law of large numbers, where $\bar{\epsilon}_i = \sum_{j=1}^r \epsilon_{ij}/r$ for $1 \le i \le n$. Proposition 1 of this paper also shows that the rate of convergence is of order $1/r$. However, when f_* is not convex, the MSE will possibly converge to a certain positive number as $r \to \infty$ because $\hat{f}_c$ converges to a function that is different from f_*. Figure 2 shows a graph of f_*, which is not convex, and the function to which $\hat{f}_c$ converges as $r \to \infty$.

The test statistic of our proposed procedure is therefore the MSE multiplied by r as follows:

$$\text{TS}_c \quad = \quad \frac{r}{n}\sum_{i=1}^n \left(\overline{Y}_i - \hat{f}_c(X_i)\right)^2 \tag{3}$$

The asymptotic behavior of the MSE suggests that we do not reject H_c^0 if the test statistic TS_c diverges to infinity as $r \to \infty$. The critical value will be derived from Propositions 1 and 2 in Section 2 at a prescribed value of the Type II error.

Test of Monotonicity: For a test of monotonicity, we will use a similar procedure. When one suspects that f_* is monotone, one can fit a monotone function $\hat{f}_m : \mathbb{R}^d \to \mathbb{R}$ to the data set, which minimizes the sum of squared errors. The fitted function $\hat{f}_m$ is the solution to the following quadratic program:

$$\begin{aligned} &\text{Minimize} && \frac{1}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - g(X_i)\right)^2 \\ &\text{Subject to} && g(X_i) \le g(X_j) \text{ for } X_i = (X_i(1), \ldots, X_i(d)) \text{ and } X_j = (X_j(1), \ldots, X_j(d)) \\ &&& \qquad \text{satisfying } X_i(k) \le X_j(k) \text{ for } k = 1, \ldots, d \end{aligned} \tag{4}$$

in the decision variables $g(X_1), \ldots, g(X_n) \in \mathbb{R}$. The proposed test statistic is then defined by

$$\mathrm{TS}_m = \frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_m(X_i)\right)^2 \tag{5}$$

and H_m^0 is not rejected if the test statistic TS_m diverges to infinity as $r \to \infty$.

Test of Positivity: For the test of positivity, the proposed test statistic is given by

$$\mathrm{TS}_p = \frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_p(X_i)\right)^2, \tag{6}$$

where $\hat{f}_p$ is the solution to the following quadratic program:

$$\begin{aligned} &\text{Minimize} && \frac{1}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - g(X_i)\right)^2 \\ &\text{Subject to} && g(X_i) \ge 0, 1 \le i \le n \end{aligned} \tag{7}$$

in the decision variables $g(X_i)$ for $1 \le i \le n$. H_p^0 is not rejected if the test statistic TS_p diverges to infinity as $r \to \infty$.

Tests of convexity/monotonicity/positivity have been widely studied in the statistics literature. Various types of hypothesis tests are proposed with different test statistics. However, most work in the literature has focused on observing the behavior of the test statistic as $n \to \infty$ with a fixed value of r (Yatchew, 1992; Hall & Jeckman, 2000; Baraud et al., 2005) or imposed a condition that requires the normality of the ϵ_{ij}'s (Bartholomew, 1959; Shapiro, 1988; Baraud et al., 2005). For example, the MSE has been studied as a test statistics in Shapiro (1988), but the behavior of the test statistic is studied only for the case where $n \to \infty$ with r fixed. Empirical studies suggest that the weights used in Shapiro (1988) are difficult to compute exactly, so the test procedure proposed by Shapiro (1988) is computationally burdensome (Sen & Silvapulle, 2002). Bartholomew (1959) used the MSE as a test statistic, but he assumed that the ϵ_{ij}'s are normally distributed and did not consider the case where $r \to \infty$. Yetchew (1992) also considered the MSE as a test statistic, but did not consider the case where $r \to \infty$, and focused only on the case where f_* is defined on the one-dimensional set $\mathbb{R}$. Others have used various types of test statistics to test a function's convexity/monotonicity/positivity. For example, Ghosal et al. (2000) use a locally weighted version of Kendall's tau statistic as a test statistic, whereas Wang & Meyer (2011) use regression splines and their derivatives to define a test statistic.

In this paper, we take a different point of view from the existing literature. Even though increasing n to infinity may result in a good estimator of the true function $f_*(x)$ over *all* $x \in \mathbb{R}^d$, increasing r to infinity can provide simpler and more practical tests for convexity/monotonicity/positivity detection that are easier to implement. We thus observe the asymptotic behavior of our test statistic as $r \to \infty$ and derive the critical value accordingly. The critical value turns out to be a percentile of $(\sigma^2/n)\chi_n^2$, where χ_n^2 follows the chi-square distribution with n degrees of freedom. Considering the fact that σ^2 can be easily estimated from the sample variance of $Y_{11}, \ldots, Y_{1r}$, the critical value can be readily computed from the data set.

The situation where r is large arises frequently in practice. In particular, this situation arises when f_* is an unknown function that we want to estimate using "computer simulation" and when we are able to select any point x in the domain of f_*, get an estimate of $f_*(x)$ through computer simulation, and repeat this procedure as many times as we wish. For example, when f_* is the price of a stock option that is contingent on the price $x \in \mathbb{R}$ of the underlying stock, we can use computer simulation to get an estimate of $f_*(x)$ at any point x as many times as we wish, and hence, r can be made as large as we wish.

Our main contribution is therefore proposing a simple and practical test procedure that is based on the idea that the MSE converges to 0 as $r \to \infty$ if f_* is convex, and diverges to infinity as $r \to \infty$ if f_* is not convex. The proposed procedure does not require estimation of any additional parameters. Furthermore, it does not rely on any assumptions regarding the distribution of the X_i's and the ϵ_{ij}'s. The test statistic and the critical value can be easily computed from the data set.

This paper is organized as follows. In Section 2, we prove that the proposed test achieves a prescribed power as $r \to \infty$. We also describe the proposed test procedure in more detail. In Section 3, we apply the proposed test to different types of f_*, and observe the conclusion of our test as $r \to \infty$. The numerical results in Section 3 illustrate that the proposed test successfully rejects the null hypothesis when the alternative hypothesis is true for r sufficiently large. Concluding remarks are included in Section 4.

2. The Asymptotic Behavior of the Test Statistics and the Proposed Test Procedure

In order to analyze the behavior of the test statistics, we will impose some probabilistic assumptions on the ϵ_{ij}'s. In particular, we require the following assumptions:

A1. Given $X_1, X_2, \ldots$, the ϵ_{ij}'s are iid random values.

A2. $\mathbb{E}\left(\epsilon_{ij} \mid X_1, \ldots, X_n\right) = 0$ for $1 \le i \le n$ and $1 \le j \le r$.

A3. $\mathbb{E}\left(\epsilon_{ij}^2 \mid X_1, \ldots, X_n\right) = \sigma^2$ for $1 \le i \le n$ and $1 \le j \le r$ and for some $\sigma < \infty$

We first establish in Proposition 1 the fact that the asymptotic distribution of the test statistics defined in (3), (5), and (6) as $r \to \infty$ is similar to the distribution of $(\sigma^2/n)\chi_n^2$.

Proposition 1 *Assume A1–A3. Then, for a fixed n,*

$$\limsup_{r\to\infty} \mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_c(X_i)\right)^2 > \tau \;\middle|\; f_* \in \mathcal{F}_c\right) \le \mathbb{P}\left((\sigma^2/n)\chi_n^2 > \tau\right),$$
$$\limsup_{r\to\infty} \mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_m(X_i)\right)^2 > \tau \;\middle|\; f_* \in \mathcal{F}_m\right) \le \mathbb{P}\left((\sigma^2/n)\chi_n^2 > \tau\right),$$
$$\limsup_{r\to\infty} \mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_p(X_i)\right)^2 > \tau \;\middle|\; f_* \in \mathcal{F}_p\right) \le \mathbb{P}\left((\sigma^2/n)\chi_n^2 > \tau\right)$$

for any $\tau \ge 0$, where χ_n^2 follows the chi-squared distribution with n degrees of freedom.

Proof. We begin by proving the existence and the uniqueness of $\hat{f}_c$, $\hat{f}_m$, and $\hat{f}_p$. The existence and the uniqueness of $\hat{f}_c$ is proven in Lemma 2.3 of Seijo & Sen (2011). To prove the existence of $\hat{f}_m$, we note that Problem (4) is a minimization problem of a coercive function over a non-empty closed subset of $\mathbb{R}^n$. By Theorem 2.32 on page 25 of Beck (2014), the solution to Problem (4) exists. To prove the uniqueness of the solution, suppose on the contrary that Problem (4) has two distinct minimizers, say $v = (v(1), \ldots, v(n))$ and $w = (w(1), \ldots, w(n))$. Since $\varphi : \mathbb{R}^n \to \mathbb{R}$, defined by $\varphi(z) = \sum_{i=1}^{n}(\overline{Y}_i - z(i))^2/n$ for $z = (z(1), \ldots, z(n)) \in \mathbb{R}^n$, is strictly convex and $(1/2)v + (1/2)w$ is a feasible solution to Problem (4), we must have $\varphi((1/2)v + (1/2)w) < (1/2)\varphi(v) + (1/2)\varphi(w)$. Since $\varphi(v) = \varphi(w)$, we have $\varphi((1/2)v + (1/2)w) < \varphi(v)$, which contradicts the fact that v is a minimizer of Problem (4). Therefore, Problem (4) has a unique minimizer. The existence and the uniqueness of $\hat{f}_p$ follows using similar arguments.

Now, we are ready to prove the main statement of Proposition 1. Suppose $f_* \in \mathcal{F}_c$. Then,

$$\frac{1}{n}\sum_{i=1}^{n}\left(\hat{f}_c(X_i) - \overline{Y}_i\right)^2 \le \frac{1}{n}\sum_{i=1}^{n}\left(f_*(X_i) - \overline{Y}_i\right)^2 \tag{8}$$

since f_* is convex and $\hat{f}_c$ minimizes $\sum_{i=1}^{n}\left(\overline{Y}_i - g(X_i)\right)/n$ over all convex functions g. Next, we will prove that

$$\frac{r}{n}\sum_{i=1}^{n}\left(f_*(X_i) - \overline{Y}_i\right)^2 = \frac{r}{n}\sum_{i=1}^{n}\bar{\epsilon}_i^2 \Rightarrow \frac{\sigma^2}{n}\chi_n^2 \tag{9}$$

as $r \to \infty$. To prove (9), we first note that, for any $\eta \in \mathbb{R}$,

$$\mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\bar{\epsilon}_i^2 > \eta\right) = \mathbb{E}\left(\mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\bar{\epsilon}_i^2 > \eta \;\middle|\; X_1, \ldots, X_n\right)\right)$$

and that

$$\mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\bar{\epsilon}_i^2 > \eta \,\middle|\, X_1,\ldots,X_n\right) \to \mathbb{P}\left((\sigma^2/n)\chi_n^2 > \eta\right)$$

almost surely as $r \to \infty$ by the weak law of large numbers (A1, A2, and A3) and the continuous mapping theorem. Applying the bounded convergence theorem to $\mathbb{P}((r/n)\sum_{i=1}^{n}\bar{\epsilon}_i^2 > \eta \mid X_1,\ldots,X_n)$ yields

$$\mathbb{E}\left(\mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\bar{\epsilon}_i^2 > \eta \,\middle|\, X_1,\ldots,X_n\right)\right) \to \mathbb{P}\left((\sigma^2/n)\chi_n^2 > \eta\right),$$

and hence, (9) follows.

Combining (8) and (9) yields

$$\mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_c(X_i)\right)^2 > \tau\right) \le \mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - f_*(X_i)\right)^2 > \tau\right) \to \mathbb{P}\left((\sigma^2/n)\chi_n^2 \ge \tau\right)$$

as $r \to \infty$, and hence, the first inequality of Proposition 1 follows. The rest of Proposition 1 uses similar arguments. □

Proposition 1 enables us to suggest the following test procedure.

Proposed Hypothesis Test

1. Using the data set $((X_i, \overline{Y}_i) : 1 \le i \le r)$, compute the test statistic TS_c from (3) for a test of convexity, TS_m from (5) for a test of monotonicity, and TS_p from (5) for a test of positivity.

2. Let β be the prescribed value of the Type II error. In other words, β is the desired probability of not rejecting the null hypothesis when the alternative hypothesis is true. Let $z_{1-\beta}$ be the $100(1-\beta)$th percentile of $(\sigma^2/n)\chi_n^2$. When σ^2 is not known, σ^2 can be estimated from the sample variance of $Y_{11},\ldots,Y_{1r}$, i.e., $\sum_{j=1}^{r}(Y_{1j} - \overline{Y}_1)^2/(r-1)$.

3. If the test statistic is less than or equal to $z_{1-\beta}$, then reject the null hypothesis in favor of the alternative hypothesis. Otherwise, do not reject the null hypothesis.

The following proposition shows that the proposed test achieves the prescribed power as $r \to \infty$.

Proposition 2 *Assume A1–A2. Then*

$$\liminf_{r\to\infty} \mathbb{P}\left(\textit{Reject } H_c^0 \mid f_* \in \mathcal{F}_c\right) \ge 1-\beta,$$
$$\liminf_{r\to\infty} \mathbb{P}\left(\textit{Reject } H_m^0 \mid f_* \in \mathcal{F}_m\right) \ge 1-\beta,$$
$$\liminf_{r\to\infty} \mathbb{P}\left(\textit{Reject } H_p^0 \mid f_* \in \mathcal{F}_p\right) \ge 1-\beta.$$

Proof. Since

$$\mathbb{P}\left(\text{Reject } H_c^0 \,\middle|\, f_* \in \mathcal{F}_c\right) = \mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_c(X_i)\right)^2 \le z_{1-\beta} \,\middle|\, f_* \in \mathcal{F}_c\right) = 1 - \mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_c(X_i)\right)^2 > z_{1-\beta} \,\middle|\, f_* \in \mathcal{F}_c\right),$$

it follows, by Proposition 1, that

$$\liminf_{r\to\infty} \mathbb{P}\left(\text{Reject } H_c^0 \,\middle|\, f_* \in \mathcal{F}_c\right) = 1 - \limsup_{r\to\infty} \mathbb{P}\left(\frac{r}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \hat{f}_c(X_i)\right)^2 > z_{1-\beta} \,\middle|\, f_* \in \mathcal{F}_c\right) > 1 - \mathbb{P}\left((\sigma/r)\chi_r^2 > z_{1-\beta}\right) = 1-\beta.$$

The rest of Proposition 2 uses similar arguments. □

3. Numerical Results

In this section, we apply the proposed test procedure to various types of f_*. We conduct the proposed hypothesis test for each case of f_*, and observe whether the null hypothesis is rejected or not as we increase r. By repeating the procedure multiple times for each case of f_*, we estimate the proportion of time that the null hypothesis is rejected. Numerical results in Sections 3.1, 3.2, and 3.3 display that the proportion of time that H_c^0, H_m^0, or H_p^0 is rejected converges to 1 as r increases

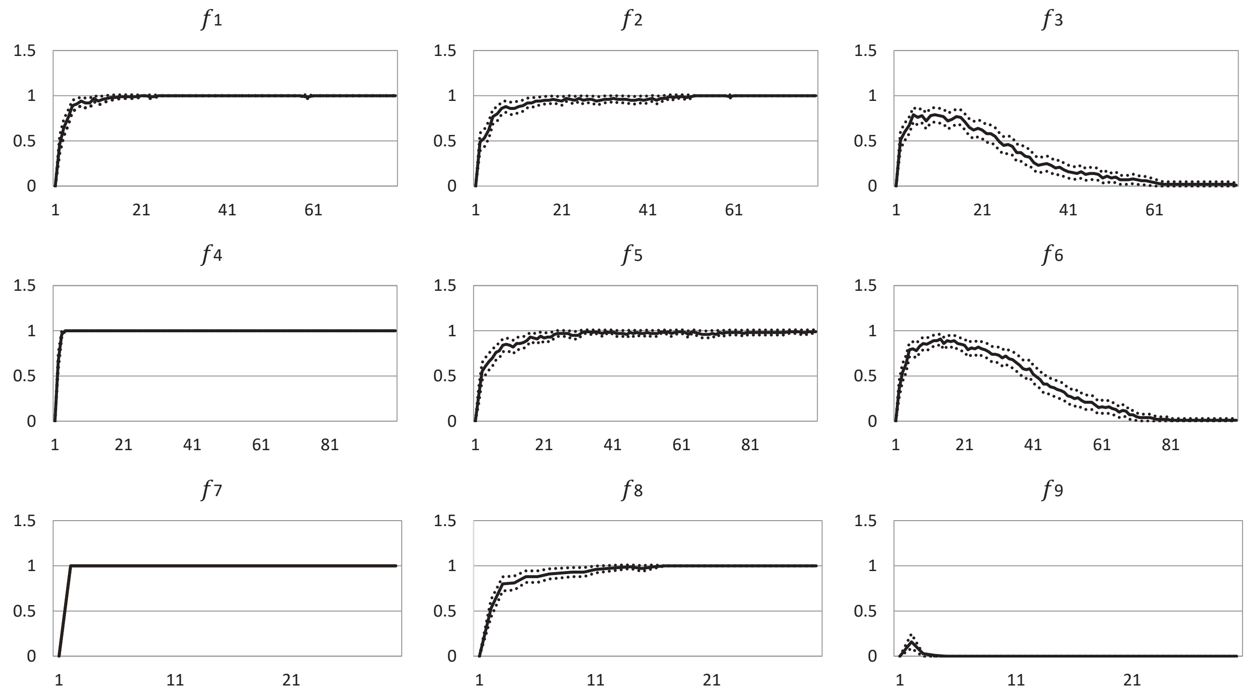

Figure 2. The dotted lines are the lower and upper limits for the 95% confidence interval of the proportion of time rejecting H_c^0 (upper three graphs), H_m^0 (middle three graphs), or H_p^0 (bottom three graphs) in the case where f_* is f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, or f_9 from top left to bottom right. The solid lines are the centers of the 95% confidence intervals. The horizontal axis is r, the number of observations made at each point in the domain, in all graphs. $n = 64$.

to infinity when f_* is convex, monotone, or positive, respectively, whereas the proportion of time that H_c^0, H_m^0, or H_p^0 is rejected converges to 0 as r increases to infinity when f_* is not convex, not monotone, or not positive, respectively. These results support Proposition 2 in Section 3, which claims that the power of the proposed test converges to 1 as $r \to \infty$. They also suggest that the type I error converges to 0 as $r \to \infty$ for n sufficiently large.

We conducted all simulations using a 64-bit computer with an Intel(R) Core(TM) i7-6600U CPU at 6 GHz and a memory of 237 GB. We programmed all simulations in MATLAB R2010a.

3.1 Test of Convexity

We consider the case where f_* is one of the following test functions:

$f_1 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_1(x) = x(1)^2 + x(2)^2 + x(3)^2$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$,

$f_2 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_2(x) = 3x(1) - x(2) + 2x(3)$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$, and

$f_3 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_3(x) = x(1)^2 - x(2)^2 + x(1)x(3)$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$.

For f_1, f_2, and f_3, we let $\{X_1, \ldots, X_n\}$ be given by

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

We then generate the Y_{ij}'s from $Y_{ij} = f_k(X_i) + U_{ij}(-1, 1)$ for $1 \le i \le n$, $1 \le j \le r$, and $1 \le k \le 3$, where the $U_{ij}(-1, 1)$'s are iid random variables uniformly distributed between -1 and 1. We next compute $\hat{f}_c$ by solving the quadratic programming problem in (2) using CVX, a package for specifying and solving convex programs (Grant & Boyd, 2014), and the test statistic TS_c by using Equation (3). When conducting the proposed test procedure, β is set as 0.05. The proposed test procedure is repeated 2,000 times. The 95% confidence interval of the proportion of time that H_c^0 is rejected is computed using these 2,000 trials and is reported in Table 1 when $n = 8$, in Table 2 when $n = 27$, and in Table 3 when $n = 64$ for a variety of r values. Figure 2 reports the 95% confidence interval of the proportion of time that H_c^0 is rejected, based on 100 iid replications, when $n = 64$ for a variety of r values.

Tables 1, 2 and 3 and Figure 2 show that the proportion of time that H_c^0 is rejected becomes close to 1 for the convex functions, f_1 and f_2, and to 0 for the non-convex function f_3 (when n is sufficiently large) as $r \to \infty$.

Table 1. The 95% confidence interval of the proportion of time rejecting H_c^0 in the case where f_* is f_1, f_2, and f_3 when n = 8.

r	$f_* = f_1$	$f_* = f_2$	$f_* = f_3$
5	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
10	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
50	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
100	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00

Table 2. The 95% confidence interval of the proportion of time rejecting H_c^0 in the case where f_* is f_1, f_2, and f_3 when n = 27.

r	$f_* = f_1$	$f_* = f_2$	$f_* = f_3$
5	1.00 ± 0.00	0.92 ± 0.01	0.63 ± 0.02
10	1.00 ± 0.00	0.98 ± 0.01	0.32 ± 0.02
50	1.00 ± 0.00	1.00 ± 0.00	0.00 ± 0.00
100	1.00 ± 0.00	1.00 ± 0.00	0.00 ± 0.00

Table 3. The 95% confidence interval of the proportion of time rejecting H_c^0 in the case where f_* is f_1, f_2, and f_3 when n = 64.

r	$f_* = f_1$	$f_* = f_2$	$f_* = f_3$
5	0.83 ± 0.02	0.75 ± 0.02	0.70 ± 0.02
10	0.95 ± 0.01	0.88 ± 0.01	0.74 ± 0.02
50	1.00 ± 0.00	0.99 ± 0.00	0.06 ± 0.01
100	1.00 ± 0.00	1.00 ± 0.00	0.00 ± 0.00

3.2 Test of Monotonicity

We next consider the case where f_* is one of the following test functions:

$f_4 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_4(x) = x(1) + x(2) + x(3)$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$,

$f_5 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_5(x) = 0$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$, and

$f_6 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_6(x) = \min(x(1) + x(2) + x(3), -0.2x(1) - 0.2x(2) - 0.2x(3))$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$, where $\min(a, b)$ denotes the minimum of a and b for $a, b \in \mathbb{R}$.

For f_4, f_5, and f_6, we let $\{X_1, \ldots, X_n\}$ be given by

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

We then generate the Y_{ij}'s from $Y_{ij} = f_k(X_i) + U_{ij}(-1, 1)$ for $1 \le i \le n$, $1 \le j \le r$, and $4 \le k \le 6$, where the $U_{ij}(-1, 1)$'s are iid random variables uniformly distributed between -1 and 1. We next compute $\hat{f}_m$ by solving the quadratic programming problem in (4) using CVX, and the test statistic TS_m by using Equation (5). When conducting the proposed test procedure, β is set as 0.05. The proposed test procedure is repeated 2,000 times. The 95% confidence interval of the proportion of time rejecting H_m^0 is computed using these 2,000 trials and is reported in Table 4 when $n = 8$, in Table 5 when $n = 27$, and in Table 6 when $n = 64$ for a variety of r values. Figure 2 reports the 95% confidence interval of the proportion of time that H_m^0 is rejected, based on 100 iid replications, when $n = 64$ for a variety of r values.

Tables 4, 5 and 6, and Figure 2 show that the proportion of time rejecting H_m^0 becomes close to 1 for the monotone functions, f_4 and f_5, and to 0 for the non-monotone function f_6 (when n is sufficiently large) as $r \to \infty$.

3.3 Test of Positivity

We consider the case where f_* is one of the following test functions:

$f_7 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_7(x) = x(1) + x(2) + x(3) + 3$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$,

$f_8 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_8(x) = 0$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$, and

$f_9 : \mathbb{R}^3 \to \mathbb{R}$ given by $f_9(x) = x(1) + x(2) + x(3)$ for $x = (x(1), x(2), x(3)) \in \mathbb{R}^3$.

Table 4. The 95% confidence interval of the proportion of time rejecting H_m^0 in the case where f_* is f_4, f_5, and f_6 when $n = 8$.

r	$f_* = f_4$	$f_* = f_5$	$f_* = f_6$
2	0.95 ± 0.01	0.54 ± 0.02	0.64 ± 0.02
10	1.00 ± 0.00	0.95 ± 0.01	0.98 ± 0.01
100	1.00 ± 0.00	0.99 ± 0.00	0.74 ± 0.02

Table 5. The 95% confidence interval of the proportion of time rejecting H_m^0 in the case where f_* is f_4, f_5, and f_6 when $n = 27$.

r	$f_* = f_4$	$f_* = f_5$	$f_* = f_6$
2	0.76 ± 0.02	0.36 ± 0.02	0.44 ± 0.02
10	1.00 ± 0.00	0.88 ± 0.01	0.85 ± 0.02
100	1.00 ± 0.00	0.98 ± 0.01	0.00 ± 0.00

Table 6. The 95% confidence interval of the proportion of time rejecting H_m^0 in the case where f_* is f_4, f_5, and f_6 when $n = 64$.

r	$f_* = f_4$	$f_* = f_5$	$f_* = f_6$
2	0.65 ± 0.02	0.31 ± 0.02	0.38 ± 0.02
10	1.00 ± 0.00	0.77 ± 0.02	0.89 ± 0.01
100	1.00 ± 0.00	0.98 ± 0.01	0.00 ± 0.00

For f_7, f_8, and f_9, we let $\{X_1, \ldots, X_n\}$ be given by

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

We then generate the Y_{ij}'s from $Y_{ij} = f_k(X_i) + U_{ij}(-1, 1)$ for $1 \leq i \leq n$, $1 \leq j \leq r$, and $7 \leq k \leq 9$, where the $U_{ij}(-1, 1)$'s are iid random variables uniformly distributed between -1 and 1. We next compute $\hat{f}_p$ by solving the quadratic programming problem in (7) using CVX, and the test statistic TS_m by using Equation (6). When conducting the proposed test procedure, β is set as 0.05. The proposed test procedure is repeated 2,000 times. The 95% confidence interval of the proportion of time rejecting H_p^0 is computed using these 2,000 trials and is reported in Table 7 when $n = 8$, in Table 8 when $n = 27$, and in Table 9 when $n = 64$ for a variety of r values. Figure 2 reports the 95% confidence interval of the proportion of time that H_p^0 is rejected, based on 100 iid replications, when $n = 64$ for a variety of r values.

Tables 7, 8 and 9, and Figure 2 show that the proportion of time rejecting H_p^0 becomes close to 1 for the positive functions, f_7 and f_8, and to 0 for the non-positive function f_9 (when n is sufficiently large) as $r \to \infty$.

3.4 Comparisons with an Existing Method

In this section, we apply our proposed method and the method proposed by Yatchew (1992) to test for a function's convexity, and compare their performance.

We assume that f_* is one of the following test functions:

$f_{10} : \mathbb{R} \to \mathbb{R}$ given by $f_{10}(x) = x^2$ for $x \in \mathbb{R}$, and

$f_{11} : \mathbb{R} \to \mathbb{R}$ given by $f_{11}(x) = 2.3(x - 0.5)x(x + 0.5)$ for $x \in \mathbb{R}$.

We let $X_i = -1 + (2i - 1)/n$ for $1 \leq i \leq n$.

To apply our proposed method, we generate the Y_{ij}'s from $Y_{ij} = f_*(X_i) + N_{ij}(0, 1^2)$ when $f_* = f_{10}$ or $Y_{ij} = f_*(X_i) + N_{ij}(0, 0.5^2)$ when $f_* = f_{11}$ for $1 \leq i \leq n$ and $1 \leq j \leq r$, where the $N_{ij}(0, 1^2)$'s and the $N_{ij}(0, 0.5^2)$'s are iid random variables normally distributed with a mean of 0 and variances of 1 and 0.5^2, respectively. We next compute $\hat{f}_c$ by solving the quadratic programming problem in (2) using CVX, and the test statistic TS_c by using Equation (3). When conducting the proposed test procedure, β is set as 0.05. The proposed test procedure is repeated 100 times. The 95% confidence interval of the proportion of time that H_c^0 is rejected is computed using these 100 trials and is reported in Tables 10 and 11 for a variety of r values when $n = 5$.

We next apply the method proposed by Yatchew (1992). In this method, only one observation of $f_*(x)$ is made at each

Table 7. The 95% confidence interval of the proportion of time rejecting H_p^0 in the case where f_* is f_7, f_8, and f_9 when $n = 8$.

r	$f_* = f_7$	$f_* = f_8$	$f_* = f_9$
2	1.00 ± 0.00	0.54 ± 0.02	0.12 ± 0.01
10	1.00 ± 0.00	0.96 ± 0.01	0.00 ± 0.00
20	1.00 ± 0.00	0.98 ± 0.01	0.00 ± 0.00

Table 8. The 95% confidence interval of the proportion of time rejecting H_p^0 in the case where f_* is f_7, f_8, and f_9 when $n = 27$.

r	$f_* = f_7$	$f_* = f_8$	$f_* = f_9$
2	1.00 ± 0.00	0.45 ± 0.02	0.03 ± 0.01
10	1.00 ± 0.00	0.97 ± 0.01	0.00 ± 0.00
20	1.00 ± 0.00	0.99 ± 0.00	0.00 ± 0.00

Table 9. The 95% confidence interval of the proportion of time rejecting H_p^0 in the case where f_* is f_7, f_8, and f_9 when $n = 64$.

r	$f_* = f_7$	$f_* = f_8$	$f_* = f_9$
2	1.00 ± 0.00	0.41 ± 0.02	0.01 ± 0.00
10	1.00 ± 0.00	0.96 ± 0.01	0.00 ± 0.00
20	1.00 ± 0.00	1.00 ± 0.00	0.00 ± 0.00

point x in the domain of f_*. So, we generate the Y_{i1}'s from $Y_{i1} = f_*(X_i)+N_i(0, 1^2)$ when $f_* = f_{10}$ or $Y_{i1} = f_*(X_i)+N_i(0, 0.5^2)$ when $f_* = f_{11}$ for $1 \le i \le n$, where the $N_i(0, 1^2)$'s and the $N_i(0, 0.5^2)$'s are iid random variables normally distributed with a mean of 0 and variances of 1 and 0.5^2, respectively. We then compute the following test statistic proposed by Yatchew (1992):

$$\frac{n^{1/2}}{(2\eta)^{1/2}}(\bar{\sigma}_n^2 - \hat{\sigma}_n^2), \tag{10}$$

where $\bar{\sigma}_n^2 = \sum_{i=1}^n (Y_{i1} - \bar{f}(X_i))^2/n$, $\hat{\sigma}_n^2 = \sum_{i=1}^n (Y_{i1} - \hat{f}(X_i))^2/n$, $\bar{f}(X_1), \ldots, \bar{f}(X_n)$ is the solution to the following quadratic program:

$$\begin{array}{lll} \text{Minimize} & \displaystyle\sum_{i=1}^{n}(Y_{i1} - g(X_i))^2/n & \\ \text{Subject to} & |g(X_i)| \le L_0, & 1 \le i \le n \\ & \dfrac{|g(X_i) - g(X_j)|}{|X_i - X_j|} \le L_1, & 1 \le i, j \le n, i \ne j \\ & \left|\dfrac{g(X_i) - g(X_j)}{X_i - X_j} - \dfrac{g(X_j) - g(X_k)}{X_j - X_k}\right| /|X_i - X_k| \le L_2/2, & 1 \le i, j, k \le n, i \ne j, j \ne k, k \ne i \\ & g(X_i) \le \dfrac{X_j - X_i}{X_k - X_i} g(X_k) + \dfrac{X_k - X_j}{X_k - X_i} g(X_i), & 1 \le i, j, k \le, X_i \le X_j \le X_k \end{array}$$

over $g(X_1), \ldots, g(X_n) \in \mathbb{R}$, $\hat{f}(X_1), \ldots, \hat{f}(X_n)$ is the solution to the following quadratic program:

$$\begin{array}{lll} \text{Minimize} & \displaystyle\sum_{i=1}^{n}(Y_{i1} - g(X_i))^2/n & \\ \text{Subject to} & |g(X_i)| \le L_0, & 1 \le i \le n \\ & \dfrac{|g(X_i) - g(X_j)|}{|X_i - X_j|} \le L_1, & 1 \le i, j \le n, i \ne j \\ & \left|\dfrac{g(X_i) - g(X_j)}{X_i - X_j} - \dfrac{g(X_j) - g(X_k)}{X_j - X_k}\right| /|X_i - X_k| \le L_2/2, & 1 \le i, j, k \le n, i \ne j, j \ne k, k \ne i \end{array}$$

Table 10. The 95% confidence interval of the proportion of time rejecting H_c^0 in the case where $f_* = f_{10}$.

$n \times r$	Proposed Method ($n = 5$)	Yatchew ($r = 1$)
10	0.76 ± 0.08	0.00 ± 0.00
20	0.95 ± 0.04	0.37 ± 0.09
30	1.00 ± 0.00	0.73 ± 0.04
40	1.00 ± 0.00	0.95 ± 0.04
50	1.00 ± 0.00	0.93 ± 0.05

Table 11. The 95% confidence interval of the proportion of time rejecting H_c^0 in the case where $f_* = f_{11}$.

$n \times r$	Proposed Method ($n = 5$)	Yatchew ($r = 1$)
20	0.56 ± 0.10	0.27 ± 0.09
40	0.55 ± 0.10	0.35 ± 0.09
60	0.39 ± 0.10	0.37 ± 0.09
80	0.28 ± 0.09	0.24 ± 0.08
100	0.13 ± 0.07	0.22 ± 0.08
120	0.07 ± 0.05	0.26 ± 0.09

over $g(X_1), \ldots, g(X_n) \in \mathbb{R}$, and $\eta = (\sum_{i=1}^{n}(Y_{i1} - \hat{f}(X_i))^4/n) - \hat{\sigma}_n^4$. L_0, L_1, and L_2 are the upper bounds on $|f_*|$, the absolute value of the first derivative of f_*, and the absolute value of the second derivative of f_*, respectively. We set $L_0 = 40, L_1 = 40$, and $L_2 = 80$ for f_{10}, and $L_0 = 3, L_1 = 9$, and $L_2 = 18$ for f_{11}. Once the test statistic is evaluated from Equation (10), H_c^0 is rejected if the test statistic is between the $100(\beta/2)$th percentile and the $100(1 - \beta/2)$th percentile of the standard normal distribution with $\beta = 0.05$. This procedure is repeated 100 times. The 95% confidence interval of the proportion of time that H_c^0 is rejected is computed using these 100 trials and is reported in Tables 10 and 11 for a variety of r and n values.

The results in Tables 10 and 11 show that the proposed method exhibits good performance in identifying the convexity of non-convexity of a function.

4. Conclusions

In this paper, we proposed a new method of testing for a function's convexity/monotonicity/positivity. The proposed method differs from the existing methods in the literature in that it observes the behavior of the test statistic as $r \to \infty$ rather than as $n \to \infty$. Propositions 1 and 2 establish that the proposed method successfully detects a function's convexity/monotonicity/positivity and achieves a prescribed value of the type II error as $r \to \infty$. An interesting point that can be raised next is whether the proposed test procedure can successfully detect a function's non-convexity/non-monotonicity/non-positivity. Our numerical results in Section 3 indicate successful detection of non-convexity/non-monotonicity/non-positivity when n is large enough to capture the overall shape of the underlying function f_*. Therefore, a promising research topic for the future is the study of the probability of the Type I error of the proposed test procedure when both r and n increase to infinity.

References

Barlow, R. E., Bartholomew, D. J., Bremner, J. M., & Brunk, H. D. (1972). *Statistical inference under order restrictions.* New York: John Wiley.

Bartholomew, D. J. (1959). A test of homogeneity for ordered alternatives. *Biometrika, 46*, 36–48. https://doi.org/10.1093/biomet/46.1-2.36

Baraud, Y., Huet, S., & Laurent, B. (2005). Testing convex hypotheses on the mean of a Gaussian vector. Application to testing qualitative hypotheses on a regression function. *The Annals of Statistics, 33*(1), 214–257. https://doi.org/10.1214/009053604000000896

Beck, A. (2014). *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB, volume 19.* SIAM. https://doi.org/10.1137/1.9781611973655

Ghosal, S., Sen, A., & van der Vaart, A. W. (2000). Testing monotonicity of regression. *The Annals of Statistics, 28*(4), 1054–1082. https://doi.org/10.1214/aos/1015956707

Grant, M., & Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.1. Retrieved June 2017 from http://cvxr.com/cvx

Hall, P., & Heckman, N. E. (2000). Testing for monotonicity of a regression mean by calibrating for linear functions. *The Annals of Statistics, 28*(1), 20–39.

Keynes, J. M. (1935). *The general theory of employment, interest, and money.* New York: Harvest/HBJ.

Seijo, E., & Sen, B. (2011). Nonparametric least squares estimation of a multivariate convex regression function. *The Annals of Statistic, 39*(3), 1633C-1657. https://doi.org/10.1214/10-AOS852

Sen, P. K., & Silvapulle, M. J. (2002). An appraisal of some aspects of statistical inference under inequality constraints. *Journal of Statistical Planning and Inference, 107*, 3–43. https://doi.org/10.1016/S0378-3758(02)00242-2

Shapiro, A. (1988). Towards a unified theory of inequality constrained testing in multivariate analysis. *International Statistical Review, 56*(1), 49–62. https://doi.org/10.2307/1403361

Wang, J. C., & Meyer, M. C. (2011). Testing the monotonicity or convexity of a function using regression splines. *Canadian Journal of Statistics, 39*, 89–107. https://doi.org/10.1002/cjs.10094

Yatchew, A. J. (1992). Nonparametric regression tests based on least squares. *Econometric Theory, 8*(4), 435–451. https://doi.org/10.1017/S0266466600013153

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/).