Likelihood ratio test for the nonnegative polynomial cone hypotheses and associated confidence bands

Kato, Naohiro
The Graduate University for Advanced Studies
10-3 Midoricho, Tachikawa
Tokyo 190-8562, Japan
E-mail: nkato@ism.ac.jp

Kuriki, Satoshi
The Institute of Statistical Mathematics
10-3 Midoricho, Tachikawa
Tokyo 190-8562, Japan
E-mail: kuriki@ism.ac.jp

1. Introduction

Consider the polynomial regression model of degree n:

$$(1) \quad y_h = f(t;c) + \varepsilon_h, \quad f(t;c) = c^\top \psi(t) = \sum_{i=0}^{n} c_i t^i, \quad t \in T,$$

$h = 1, \ldots, N$, where $\psi(t) = \psi_n(t) = (1, t, \ldots, t^n)^\top$ and $c = (c_0, c_1, \ldots, c_n)^\top$ are column vectors in $\mathbb{R}^{n+1}$. The errors ε_h are independently distributed as the normal distribution $N(0, \sigma^2)$. $T \subseteq \mathbb{R}$ is the region of the explanatory variable t where the model (1) is defined. Typically, T is a bounded interval in $\mathbb{R}$. We assume that sufficient statistics of the model (1), that is, the ordinal least square estimator $\widehat{c}$ of c, and when σ^2 is unknown, the unbiased variance estimator $\widehat{\sigma}^2$ of σ^2 distributed independently of $\widehat{c}$ are available.

In this paper, we deal with the hypothesis of positivity, or the hypothesis of superiority:

$$(2) \quad f(t;c) \geq 0 \quad \text{for all } t \in T.$$

To state its statistical meaning, it is natural to consider a two sample problem. Let $f(t; c_{(j)}) = c_{(j)}^\top \psi(t)$ ($j = 0, 1$) be polynomial regression curves of two groups. The hypothesis that the polynomial curve of group 1 is always bounded below by, or superior to, group 0 is expressed as $f(t; c_{(1)}) \geq f(t; c_{(0)})$ for all $t \in T$. Taking the difference, we see that (2) stands for the hypothesis of superiority. This notion is particularly important in statistical tests for assessing new drugs (Liu, et al. (2009)).

The set of coefficients c satisfying (2) forms a closed convex cone

$$(3) \quad K = K_n = \left\{ c \in \mathbb{R}^{n+1} \mid c^\top \psi(t) \geq 0, \ \forall t \in T \right\}.$$

This is referred to as the cone of positive polynomials (nonnegative polynomial cone). The hypothesis (2) is rewritten as $c \in K$. Including this hypothesis, we consider hierarchical hypotheses

$$(4) \quad H_0 : c = 0, \quad H_1 : c \in K_n, \quad \text{and} \quad H_2 : c \in \mathbb{R}^{n+1} \ (c \text{ is unrestricted}).$$

We then formalize the test for positivity as the likelihood ratio test (LRT) for testing H_1 against H_2. In addition, we define a LRT for testing H_0 against H_1. In the context of the two sample problem, this is the test for the equality of two regression curves against the hypothesis of superiority. It is mathematically convenient to treat the two LRTs in a paratactic way.

The theory of LRTs for convex cone hypotheses has been developed under the name of order restricted inference (Robertson, et al. (1988)). The general theorem states that the null distribution of LRT statistics is a finite mixture of chi-square distributions (Shapiro (1988)). When the cone has piecewise smooth boundaries, Takemura and Kuriki (1997) proved that the weights (mixing probabilities) are expressed in terms of curvature measures on boundaries. This methodology is called the volume-of-tube method. However, there are few cones whose weights are obtained explicitly.

As the main result of the paper, we derive the weights associated with the cone of positive polynomials K, that is, the null distribution of the LRT for positivity. Thanks to the representation (parameterization) theorem for the positive polynomial cone and its dual cone developed in the framework of the Tchebycheff systems (Karlin and Studden (1966)), we evaluate the all weights when the degree n of the polynomial regression is less than or equal to 4.

Throughout the paper, descriptions on confidence bands, technical details including proofs, and numerical examples are omitted. See Kato and Kuriki (2011) for the details.

2. Likelihood ratio test statistics

We write the inner product and the norm as $\langle x, y\rangle_Q = x^\top Q y$ and $\|x\|_Q = \sqrt{\langle x, x\rangle_Q}$, where Q is a positive definite matrix. The orthogonal projection of x onto the set A with respect to the distance $\|\,\|_Q$ is denoted by $\Pi_Q(x|A) = \mathrm{argmin}_{y\in A}\|x - A\|_Q$. The subscript Q in $\langle\,,\,\rangle_Q$, $\|\,\|_Q$, and Π_Q will be omitted when it does not cause any confusion.

In the regression model (1) with σ^2 known, the least square statistics $\widehat{c}$ is the sufficient statistic, and we can restrict our attention to the inference based on $\widehat{c}$. The distribution of $\widehat{c}$ is the normal distribution $N_{n+1}(c, \Sigma)$, where $\Sigma = \sigma^2\Sigma_0$ with $\Sigma_0 = \left(\sum_{i=1}^N \psi(t_i)\psi(t_i)^\top\right)^{-1}$, the inverse of the design matrix. When σ^2 is unknown, the sufficient statistics is the pair $(\widehat{c}, \widehat{\sigma}^2)$, where $\widehat{\sigma}^2$ is the unbiased estimator of σ^2 calculated from the residuals, and is distributed proportionally to the chi-square distribution with $\nu = N - n - 1$ degrees of freedom.

Given the data $\widehat{c}$ distributed as $N_{n+1}(c, \Sigma)$ with $\Sigma = \sigma^2\Sigma_0$ known, the MLE of c under the hypothesis of positivity $H_1 : c \in K$ is the orthogonal projection $\widehat{c}_K$ of $\widehat{c}$ onto the cone K under the metric $\langle\,,\,\rangle_{\Sigma^{-1}}$. When σ^2 is unknown, the MLE is the orthogonal projection onto K under the metric $\langle\,,\,\rangle_{\widehat{\Sigma}^{-1}}$, $\widehat{\Sigma} = \widehat{\sigma}^2\Sigma_0$. This MLE is the same as that with Σ known. The MLEs of c under H_0 and H_2 are given as 0 and $\widehat{c}$, respectively. Acknowledging these facts, we get the LRT statistics.

Proposition 1. *When the variance σ^2 is known, the LRT statistics for H_0 against H_1, and H_1 against H_2 are given by*

$$\lambda_{01} = \|\widehat{c}_K\|^2_{\Sigma^{-1}} \quad \text{and} \quad \lambda_{12} = \|\widehat{c}\|^2_{\Sigma^{-1}} - \|\widehat{c}_K\|^2_{\Sigma^{-1}}, \tag{5}$$

respectively, where $\widehat{c}_K = \Pi_{\Sigma^{-1}}(\widehat{c}|K)$. When the variance σ^2 is unknown, and an independent and unbiased estimator $\widehat{\sigma}^2$ of σ^2 with ν degrees of freedom is available, the LRT statistics for H_0 against H_1, and H_1 against H_2 are given by

$$\beta_{01} = \frac{\|\widehat{c}_K\|^2_{\widehat{\Sigma}^{-1}}}{\|\widehat{c}\|^2_{\widehat{\Sigma}^{-1}} + \nu} \quad \text{and} \quad \beta_{12} = \frac{\|\widehat{c}\|^2_{\widehat{\Sigma}^{-1}} - \|\widehat{c}_K\|^2_{\widehat{\Sigma}^{-1}}}{\|\widehat{c}\|^2_{\widehat{\Sigma}^{-1}} - \|\widehat{c}_K\|^2_{\widehat{\Sigma}^{-1}} + \nu}, \tag{6}$$

respectively, where $\widehat{\Sigma} = \widehat{\sigma}^2\Sigma_0$, $\widehat{c}_K = \Pi_{\widehat{\Sigma}^{-1}}(\widehat{c}|K)$. The null hypotheses are rejected when the statistics are greater than critical points.

The hypothesis of positivity H_1 is a composite hypothesis. The proof of the following proposition is essentially given in Section 2.3 of Robertson, et al. (1988).

Proposition 2. *In both cases where σ^2 is known or unknown, the least favorable configurations of the LRTs for testing H_1 against H_2 is given by H_0, i.e., $c = 0$.*

3. The volume-of-tube method

Here, we give a brief summary of the volume-of-tube method. Historically, the distributions of the orthogonal projection of a zero-mean Gaussian random vector have been well studied in order restricted inference. From the general theory, the statistics λ_{01} and λ_{12} in (5) under $H_0 : c = 0$ have the following distribution:

$$(7)\quad P_{H_0}(\lambda_{01} \geq a,\, \lambda_{12} \geq b) = \sum_{i=0}^{n+1} w_i \bar{G}_i(a)\bar{G}_{n+1-i}(b),$$

where $\bar{G}_i$ is the upper probability of the chi-square distribution with i degrees of freedom. Note that $\bar{G}_0(a) = 1$ $(a < 0)$, 0 $(a \geq 0)$. In addition, the distribution of the LRTs β_{01} and β_{12} in (6) under H_0 is expressed as follows:

$$(8)\quad P_{H_0}(\beta_{01} \geq a,\, \beta_{12} \geq b) = \sum_{i=0}^{n+1} w_i \bar{B}_{\frac{i}{2},\frac{n+1-i+\nu}{2}}(a)\bar{B}_{\frac{n+1-i}{2},\frac{\nu}{2}}(b),$$

where $\bar{B}_{a,b}$ is the upper probability of the beta distribution with parameter (a, b).

The coefficients w_i appearing in (7) and (8) are nonnegative and satisfy $\sum_i w_i = 1$. The distribution (7) is a finite mixture distribution of the chi-square distributions refereed to as the chi-bar-square ($\bar{\chi}^2$) distribution.

When the cone K in (4) is polyhedral, that is, a finite intersection of half spaces, the weights $\{w_i\}$ can be understood in terms of the internal and external angles of each face of the cone. Moreover, in the general case where K is not polyhedral, Takemura and Kuriki (1997) proved that the weights $\{w_i\}$ are expressed as integrals of elementary symmetric polynomials of principle curvatures of the boundaries of the cone K. These integrals are not easy to handle in general. However, the weights of the two highest degrees and two lowest degrees, w_{n+1}, w_n, w_0, w_1 have relatively simple forms:

$$(9)\quad w_{n+1} = \frac{\mathrm{Vol}_n(K \cap \mathbb{S}^n)}{\mathrm{Vol}_n(\mathbb{S}^n)}, \quad w_n = \frac{\mathrm{Vol}_{n-1}(\partial K \cap \mathbb{S}^n)}{2\,\mathrm{Vol}_{n-1}(\mathbb{S}^{n-1})},$$
$$w_1 = \frac{\mathrm{Vol}^*_{n-1}(\partial K^* \cap (\mathbb{S}^n)^*)}{2\,\mathrm{Vol}_{n-1}((\mathbb{S}^{n-1})^*)}, \quad w_0 = \frac{\mathrm{Vol}^*_n(K^* \cap (\mathbb{S}^n)^*)}{\mathrm{Vol}^*_n(\mathbb{S}^n)},$$

where K^* is the dual cone of K referred to as the moment cone, and ∂K and ∂K^* are the boundaries, $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid \|x\|_{\Sigma^{-1}} = 1\}$, $(\mathbb{S}^n)^* = \{x \in \mathbb{R}^{n+1} \mid \|x\|_{\Sigma} = 1\}$ are the unit spheres, and Vol_d and Vol^*_d are d-dimensional volumes induced by the metrics $\langle\, ,\rangle_{\Sigma^{-1}}$ and $\langle\, ,\rangle_{\Sigma}$, respectively. The volume of the unit sphere is $\mathrm{Vol}_{d-1}(\mathbb{S}^{d-1}) = \mathrm{Vol}^*_{d-1}((\mathbb{S}^{d-1})^*) = 2\pi^{d/2}/\Gamma(d/2)$.

Moreover, a useful relation is known as a consequence of the Gauss-Bonnet theorem:

$$(10)\quad \sum_{i:\mathrm{odd}} w_i = \sum_{i:\mathrm{even}} w_i = \frac{1}{2}.$$

The distribution of β_{01} with $\nu = 0$ is interpreted as the volume formula of a spherical tubular neighborhood as below. Let $M = K \cap \mathbb{S}^n$ be the intersection of the cone K in (3) and the unit sphere. Define the spherical tube about M with the radius θ:

$$\mathrm{Tube}(M, \theta) = \Big\{ x \in \mathbb{S}^n \mid \min_{y \in M} \mathrm{dist}(x, y) \leq \theta \Big\}, \quad \mathrm{dist}(x, y) = \cos^{-1}\langle x, y\rangle.$$

Then, because $\beta_{01} = \|\Pi(\widehat{c}|K)\|^2/\|\widehat{c}\|^2 \geq \cos\theta \Leftrightarrow \widehat{c}/\|\widehat{c}\| \in \mathrm{Tube}(M, \theta)$, and $\widehat{c}/\|\widehat{c}\|$ is distributed uniformly on $\mathbb{S}^n$ under H_0, we see that

$$\frac{\mathrm{Vol}_n(\mathrm{Tube}(M, \theta))}{\mathrm{Vol}_n(\mathbb{S}^n)} = P_{H_0}(\beta_{01} \geq \cos\theta) = \sum_{i=0}^{n+1} w_i \bar{B}_{\frac{i}{2},\frac{n+1-i}{2}}(\cos\theta),$$

where Vol_n is the spherical volume defined on $\mathbb{S}^n$ induced by the metric $\langle\, ,\, \rangle = \langle\, ,\, \rangle_{\Sigma^{-1}}$ of the ambient space $\mathbb{R}^{n+1}$. This is the reason why our methodology is called the tube method (Kuriki and Takemura (2009)).

4. Representations for the cones K and K^*

In order to evaluate the volumes in (9), we need to introduce "local coordinates" of the cones K and K^*, and their boundaries. This is actually possible by means of the representations in the theory of Tchebycheff systems. Three cases can be considered: (i) $T = [a, b]$ (bounded), (ii) $T = [a, \infty)$, (iii) $T = (-\infty, \infty)$. We deal with the case (i) only here for simplicity. Let $\psi_n(t) = (1, t, \ldots, t^n)^\top$ ($|t| < \infty$), $(0, \ldots, 0, (\pm 1)^n)^\top$, ($t = \pm\infty$). Let $\mathbb{R}_+ = (0, \infty)$, and

(11) $\Delta_m = \Delta_m(T) = \{\tau = (\tau_1, \ldots, \tau_m) \in (\mathrm{int}\, T)^m \mid \tau_1 < \cdots < \tau_m\}$.

Proposition 3. *The moment cone K_n^* on $T = [a, b]$ has the following almost sure representations:*

(12) $$K_n^* = \phi_{n,n}^{(U)}\Big(\mathbb{R}_+^{[\frac{n+1}{2}]+1} \times \Delta_{[\frac{n}{2}]}\Big) = \phi_{n,n}^{(L)}\Big(\mathbb{R}_+^{[\frac{n}{2}]+1} \times \Delta_{[\frac{n+1}{2}]}\Big)$$

almost surely with respect to the $(n+1)$-dimensional Lebesgue measure, where

(13) $$\phi_{n,l}^{(U)}(\rho, \tau) = \begin{cases} \sum_{i=1}^m \rho_i \psi_n(\tau_i) + \rho_{m+1}\psi_n(b) & (l = 2m), \\ \rho_1\psi_n(a) + \sum_{i=1}^m \rho_{i+1}\psi_n(\tau_i) + \rho_{m+2}\psi_n(b) & (l = 2m+1), \end{cases}$$

(14) $$\phi_{n,l}^{(L)}(\rho, \tau) = \begin{cases} \rho_1\psi_n(a) + \sum_{i=1}^m \rho_{i+1}\psi_n(\tau_i) & (l = 2m), \\ \sum_{i=1}^{m+1} \rho_i\psi_n(\tau_i) & (l = 2m+1). \end{cases}$$

The maps $\phi_{n,n}^{(U)}$ and $\phi_{n,n}^{(L)}$ in (12) are diffeomorphic.

Remark 1. *The representation with (13) is called the upper representation. The representation with (14) is called the lower representation.*

Proposition 4. *Let Δ_m be defined in (11). Let $\phi_{n,l}^{(U)}$ and $\phi_{n,l}^{(L)}$ be defined in (13) and (14). The boundary of the moment cone ∂K_n^* has the following almost sure representation:*

(15) $$\partial K_n^* = \phi_{n,n-1}^{(L)}\Big(\mathbb{R}_+^{[\frac{n-1}{2}]+1} \times \Delta_{[\frac{n}{2}]}\Big) \sqcup \phi_{n,n-1}^{(U)}\Big(\mathbb{R}_+^{[\frac{n}{2}]+1} \times \Delta_{[\frac{n-1}{2}]}\Big)$$

almost surely with respect to the n-dimensional Hausdorff measure, where $\sqcup$ means disjoint union. The maps $\phi_{n,n-1}^{(U)}$ and $\phi_{n,n-2}^{(U)}$ in (15) are diffeomorphic.

Proposition 5. *The positive polynomial cone K_n on $T = [a, b]$ is represented as $K_n = \varphi_n\big(\mathbb{R}_+^2 \times \Delta_{n-1}\big)$ almost surely with respect to the $(n+1)$-dimensional Lebesgue measure. Here, the function $\varphi_n(\alpha, \gamma) \in \mathbb{R}^{n+1}$ with $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}_+^2$ and $\gamma = (\gamma_1, \ldots, \gamma_{n-1}) \in \Delta_{n-1}$ is the coefficients of the polynomial $p_n(t; \alpha, \gamma) = \varphi_n(\alpha, \gamma)^\top \psi_n(t)$ in t defined below:*

$$p_n(t; \alpha, \gamma) = \begin{cases} \alpha_1 \prod_{j=1}^m (t - \gamma_{2j-1})^2 + \alpha_2(t-a)(b-t)\prod_{j=1}^{m-1}(t - \gamma_{2j})^2 & (n = 2m), \\ \alpha_1(t-a)\prod_{j=1}^m (t - \gamma_{2j})^2 + \alpha_2(b-t)\prod_{j=1}^m (t - \gamma_{2j-1})^2 & (n = 2m+1). \end{cases}$$

The map φ_n is a diffeomorphism.

Proposition 6. *The boundary of the positive polynomial cone ∂K_n has the almost sure representation below. Define the functions $\varphi_n^{(i)}(\alpha, \gamma, \widetilde{\gamma}) \in \mathbb{R}^{n+1}$ with $\alpha \in \mathbb{R}_+^2$, $\gamma \in \Delta_{n-1-i}$, $\widetilde{\gamma} \in \mathbb{R}$, by the coefficient vectors of polynomials as $(t - \widetilde{\gamma})^i p_{n-i}(t; \alpha, \gamma) = \varphi_n^{(i)}(\alpha, \gamma, \widetilde{\gamma})^\top \psi_n(t)$ $(i = 1, 2)$. Then,*

$$\partial K_n = \varphi_n^{(2)}\Big(\mathbb{R}_+^2 \times \Delta_{n-3} \times T\Big) \sqcup \varphi_n^{(1)}\Big(\mathbb{R}_+^2 \times \Delta_{n-2}, a\Big) \sqcup \Big\{-\varphi_n^{(1)}\Big(\mathbb{R}_+^2 \times \Delta_{n-2}, b\Big)\Big\}$$

almost surely with respect to the n-dimensional Hausdorff measure, where $\sqcup$ means disjoint union. The maps φ_n, $\varphi_n^{(1)}(\cdot, a)$, $\varphi_n^{(1)}(\cdot, b)$, and $\varphi_n^{(2)}$ are diffeomorphisms.

5. Volume formulas and the weights

The diffeomorphic maps appearing in Propositions 3, 4, 5, and 6 are homogeneous functions with respect to their first arguments ρ and α. Therefore, by restricting the length of the first argument, we can construct almost sure representations for the intersections with the unit sphere. For example, $\phi^{(U)}_{n,n}(r\rho,\tau) = r\phi^{(U)}_{n,n}(\rho,\tau)$ for a constant $r > 0$, and we have

$$K^*_n \cap (\mathbb{S}^n)^* = \bar{\phi}^{(U)}_{n,n}\Big(\mathbb{S}_+^{[\frac{n+1}{2}]+1} \times \Delta_{[\frac{n}{2}]}\Big) \quad \text{a.s.},$$

where $\bar{\phi}^{(U)}_{n,l}(\rho,\tau) = \phi^{(U)}_{n,l}(\rho,\tau)/\|\phi^{(U)}_{n,l}(\rho,\tau)\|_\Sigma$, and $\mathbb{S}^m_+ = \{\rho = (\rho_i) \in \mathbb{R}^{m+1} \mid \sum \rho_i^2 = 1,\ \rho_i > 0\}$. Define $\bar{\phi}^{(L)}_{n,l}(\rho,\tau) = \phi^{(L)}_{n,l}(\rho,\tau)/\|\phi^{(L)}_{n,l}(\rho,\tau)\|_\Sigma$, $\bar{\varphi}_n(\alpha,\gamma) = \varphi_n(\alpha,\gamma)/\|\varphi_n(\alpha,\gamma)\|_{\Sigma^{-1}}$, $\bar{\varphi}^{(i)}_n(\alpha,\gamma,\widetilde{\gamma}) = \varphi^{(i)}_n(\alpha,\gamma,\widetilde{\gamma})/\|\varphi^{(i)}_n(\alpha,\gamma,\widetilde{\gamma})\|_{\Sigma^{-1}}$ $(i = 1,2)$, similarly.

In the proposition below, let $\theta = (\theta_i) \in \Theta_m$ be the local coordinates of $\mathbb{S}^m_+$. For example, $\rho = \rho(\theta) = \big(\theta_1,\ldots,\theta_m,\sqrt{1-\sum\theta_i^2}\big)$, $\theta \in \Theta_m = \mathbb{R}^m_+$. Another example is the polar coordinates.

Proposition 7. *Let $d\theta = \prod d\theta_i$ and $d\tau = \prod d\tau_i$ be the Lebesgue measures.*

$$\mathrm{Vol}^*(K^*_n \cap (\mathbb{S}^n)^*) = \int_{\Theta_{[\frac{n+1}{2}]}\times\Delta_{[\frac{n}{2}]}} \det\left\{\left(\frac{\partial\bar{\phi}^{(U)}_{n,n}(\rho,\tau)}{\partial(\theta,\tau)}\right)^\top \Sigma \left(\frac{\partial\bar{\phi}^{(U)}_{n,n}(\rho,\tau)}{\partial(\theta,\tau)}\right)\right\}^{\frac{1}{2}} d\theta\, d\tau,$$

$$\begin{aligned}\mathrm{Vol}(\partial K^*_n \cap (\mathbb{S}^n)^*) = &\int_{\Theta_{[\frac{n-1}{2}]}\times\Delta_{[\frac{n}{2}]}} \det\left\{\left(\frac{\partial\bar{\phi}^{(L)}_{n,n-1}(\rho,\tau)}{\partial(\theta,\tau)}\right)^\top \Sigma \left(\frac{\partial\bar{\phi}^{(L)}_{n,n-1}(\rho,\tau)}{\partial(\theta,\tau)}\right)\right\}^{\frac{1}{2}} d\theta\, d\tau \\ &+ \int_{\Theta_{[\frac{n}{2}]}\times\Delta_{[\frac{n-1}{2}]}} \det\left\{\left(\frac{\partial\bar{\phi}^{(U)}_{n,n-1}(\rho,\tau)}{\partial(\theta,\tau)}\right)^\top \Sigma \left(\frac{\partial\bar{\phi}^{(U)}_{n,n-1}(\rho,\tau)}{\partial(\theta,\tau)}\right)\right\}^{\frac{1}{2}} d\theta\, d\tau.\end{aligned}$$

Proposition 8. *Let $d\gamma = \prod d\gamma_i$ be the Lebesgue measure. Let $\alpha = (\cos\theta, \sin\theta)$.*

$$\mathrm{Vol}(K_n \cap \mathbb{S}^n) = \int_{(0,\frac{\pi}{2})\times\Delta_{n-1}} \det\left\{\left(\frac{\partial\bar{\varphi}_n(\alpha,\gamma)}{\partial(\theta,\gamma)}\right)^\top \Sigma^{-1} \left(\frac{\partial\bar{\varphi}_n(\alpha,\gamma)}{\partial(\theta,\gamma)}\right)\right\}^{\frac{1}{2}} d\theta\, d\gamma,$$

$$\begin{aligned}\mathrm{Vol}(\partial K_n \cap \mathbb{S}^n) = &\int_{(0,\frac{\pi}{2})\times\Delta_{n-3}\times T} \det\left\{\left(\frac{\partial\bar{\varphi}^{(2)}_n(\theta,\gamma,\widetilde{\gamma})}{\partial(\theta,\gamma,\widetilde{\gamma})}\right)^\top \Sigma^{-1} \left(\frac{\partial\bar{\varphi}^{(2)}_n(\theta,\gamma,\widetilde{\gamma})}{\partial(\theta,\gamma,\widetilde{\gamma})}\right)\right\}^{\frac{1}{2}} d\theta\, d\gamma\, d\widetilde{\gamma} \\ &+ \sum_{c\in\{a,b\}} \int_{(0,\frac{\pi}{2})\times\Delta_{n-2}} \det\left\{\left(\frac{\partial\bar{\varphi}^{(1)}_n(\alpha,\gamma,c)}{\partial(\theta,\gamma)}\right)^\top \Sigma^{-1} \left(\frac{\partial\bar{\varphi}^{(1)}_n(\alpha,\gamma,c)}{\partial(\theta,\gamma)}\right)\right\}^{\frac{1}{2}} d\theta\, d\gamma.\end{aligned}$$

Substituting the volumes obtained in Propositions 7 and 8 into (9), we get w_{n+1}, w_n, w_0, w_1. Combined with the Gauss-Bonnet theorem (10), all weights $\{w_i\}$ for $n \le 4$ are obtained.

$$(w_0,\ldots,w_{n+1}) = \begin{cases} \big(\frac{1}{2}-w_n, w_1, w_n, \frac{1}{2}-w_1\big) & (n=2), \\ \big(w_0, w_1, \frac{1}{2}-w_0-w_{n+1}, w_n, w_{n+1}\big) & (n=3), \\ \big(w_0, w_1, \frac{1}{2}-w_0-w_n, \frac{1}{2}-w_1-w_{n+1}, w_n, w_{n+1}\big) & (n=4). \end{cases}$$

6. A numerical procedure for MLE

To obtain the LRT statistics λ_{01} and λ_{12} in (5), we need to conduct the orthogonal projection onto the positive polynomial cone K. For this purpose, the following symmetric cone programming technique is useful.

The positive polynomial $p_n(t)$ of degree n on T is characterized in Proposition 5. This is a unique representation. Admitting the redundancy of the parameters, this polynomial is rewritten as

$$(16)\quad p_n(t) = \begin{cases} \psi_m(t)^\top Q_1 \psi_m(t) + (t-a)(b-t)\psi_{m-1}(t)^\top Q_2 \psi_{m-1}(t) & (n = 2m), \\ (t-a)\psi_m(t)^\top Q_1 \psi_m(t) + (b-t)\psi_m(t)^\top Q_2 \psi_m(t) & (n = 2m+1), \end{cases}$$

where Q_1 and Q_2 are symmetric positive semidefinite matrices. This polynomial (16) is obviously nonnegative on $T = [a, b]$. Conversely, the polynomial $p_n(t)$ in Proposition 5 can be written as (16). This representation is sometimes referred as the Markov-Lukacs theorem (Nesterov (2000)).

By arranging the terms, the polynomial $p_n(t)$ in (16) can be written as $p_n(t) = e(Q_1, Q_2)^\top \psi_n(t)$, where $e(Q_1, Q_2)$ is a column vector depending on Q_1 and Q_2. Using this representation, the orthogonal projection of a given vector $\widehat{c}$ onto the positive polynomial cone K is formalized as the optimization problem below:

$$\begin{array}{lll} \texttt{maximize} & -d & \\ \texttt{subject to} & d \geq \|\widehat{c} - c\|^2_{\Sigma^{-1}} & \text{(quadratic cone restriction)} \\ & c = e(Q_1, Q_2) & \text{(linear restriction)} \\ & Q_1, Q_2 \succeq 0 & \text{(PSD cone restriction)} \end{array}$$

This is an optimization problem with quadratic cone, linear, and positive semi-definite (PSD) cone restrictions. This can be solved in the framework of symmetric cone programming. Several public softwares are available (e.g., SeDuMi by Sturm (1999)).

REFERENCES

Karlin, S. and Studden, W. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Interscience Publishers, Wiley, New York.

Kato, N. and Kuriki, S. (2011). Tests for positivity in polynomial regressions based on the volume-of-tube formula. Preprint.

Kuriki, S. and Takemura, A. (2009). Volume of tubes and the distribution of the maximum of a Gaussian random field. *Selected Papers on Probability and Statistics*, AMS Translations Series 2, **227**, No. 2, 25–48.

Liu, W., Bretz, F., Hayter, A. J., and Wynn, H. P. (2009). Assessing non-superiority, non-inferiority or equivalence when comparing two regression models over a restricted covariate region. *Biometrics*, **65** (4), 1279–1287.

Nesterov, Y. (2000). Squared functoinal systems and optimization problems. In *High Performance Optimization* (eds. H. Frenk, K. Roos, T. Terlaky and S. Zhang), 405–440, Kluwer, Dordrecht, The Netherlands.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*, Wiley, Chichester.

Shapiro, A. (1988). Towards a unified theory of inequality constrained testing in multivariate analysis. *Internat. Statist. Rev.*, **56** (1), 49–62.

Sturm, J. F. (1999). Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, **11** (1), 625–653.

Takemura, A. and Kuriki, S. (1997). Weights of chi-bar-square distribution for smooth or piecewise smooth cone alternatives. *Ann. Statist.*, **25** (6), 2368–2387.