Statistica Sinica **19** (2009), 213-232

CONSTRUCTION OF EXACT SIMULTANEOUS CONFIDENCE BANDS IN MULTIPLE LINEAR REGRESSION WITH PREDICTOR VARIABLES CONSTRAINED IN AN ELLIPSOIDAL REGION

Wei Liu and Shan Lin

University of Southampton

Abstract: A simultaneous confidence band provides useful information on the plausible range of the unknown regression model. There are several recent papers using confidence bands for various inferential purposes; see, for example, Sun, Raz and Faraway (1999), Spurrier (1999), Al-Saidy et al. (2003), Liu, Jamshidian and Zhang (2004), and Piegorsch et al. (2005). Construction of simultaneous confidence bands has a history going back to Working and Hotelling (1929), and is often a hard problem when the region over which a confidence band is required is restricted and the number of predictor variables is more than one. This article considers the construction of exact $1-\alpha$ level one-sided and two-sided simultaneous confidence bands of hyperbolic shape for the normal-error multiple linear regression model when the predictor variables are constrained to a particular ellipsoidal region that is centered at the point of the means of the predictor variable values used in the experiment. MATLAB programs have been written for easy implementation of the constructions, and an illustrative example is provided.

Key words and phrases: Circular cone, linear regression, simultaneous confidence bands, statistical inference.

1. Introduction

Consider the multiple linear regression model

$$\mathbf{Y} = X\mathbf{b} + \mathbf{e},$$

where $\boldsymbol{Y}_{n\times 1}$ is the vector of observed responses, $X_{n\times p}$ is the design matrix with first column $(1,\ldots,1)^T$ and jth $(2 \le j \le p)$ column $(x_{1,j},\ldots,x_{n,j})^T$, $\mathbf{b} = (b_1,\ldots,b_p)^T$ is the vector of regression coefficients, and $\boldsymbol{e}_{n\times 1}$ is the error vector with $\boldsymbol{e} \sim N(0,\sigma^2 I)$ and σ^2 unknown. Assume X^TX is non-singular, so the least squares estimator of $\mathbf{b}$ is $\hat{\mathbf{b}} = (X^TX)^{-1}X^T\mathbf{Y}$. Let $\hat{\sigma}^2$ denote the mean square error with $\nu = n-p$ degrees of freedom. Then $\hat{\sigma}^2 \sim \sigma^2\chi^2_\nu/\nu$ and is independent of $\hat{\mathbf{b}}$.

Let $\mathbf{x} = (1,x_2,\ldots,x_p)^T$ and $\mathbf{x}_{(1)} = (x_2,\ldots,x_p)^T$. In this paper, the $p-1$ predictor variables x_i are assumed to satisfy no functional relationship, and so polynomial regression, for example, is excluded from the discussion. The results of the paper, however, can be used to construct conservative confidence bands when the predictor variables have functional relationships between them by simply ignoring these relationships. A simultaneous confidence band for the regression function

$$\mathbf{x}^T\mathbf{b} = b_1 + b_2x_2 + \cdots + b_px_p$$

on a given region $\mathcal{X}$ of the $p-1$ predictor variables $\mathbf{x}_{(1)} = (x_2, \ldots, x_p)^T$ provides useful information on the true but unknown regression model; a linear regression function is a plausible candidate for the unknown regression model if and only if it is contained completely inside the confidence band. There are several recent papers considering various applications of confidence bands; see for example Sun, Raz and Faraway (1999), Spurrier (1999), Al-Saidy et al. (2003), Liu, Jamshidian and Zhang (2004) and Piegorsch et al. (2005).

Construction of simultaneous confidence bands has a history going back to Working and Hotelling (1929). Scheffé (1953) provided a well known two-sided hyperbolic simultaneous confidence band when $\mathcal{X} = R^{p-1}$, that is, for the case where the $p-1$ predictor variables are not constrained at all.

For $p = 2$, Gafarian (1964) considered a two-sided confidence band with a constant width when the predictor variable is constrained to an interval; see also Miller (1981). His effort was followed by Bowden and Graybill (1966) and Bowden (1970) who considered two-sided confidence bands of other shapes; Piegorsch et al. (2000) considered the calculation of critical constants of a family of confidence bands from Bowden (1970). Wynn and Bloomfield (1971) and Uusipaikka (1983) provided exact two-sided hyperbolic confidence bands with width proportional to standard error when the predictor variable is restricted to an interval or union of intervals. Bohrer and Francis (1972) considered exact one-sided hyperbolic confidence bands when the predictor variable is constrained to an interval. Pan, Piegorsch and West (2003) also constructed exact one-sided hyperbolic confidence bands when the only predictor variable is constrained to an interval by using the idea of Uusipaikka (1983). Comparisons of different confidence bands for $p = 2$ have been considered by Naiman (1983), among others, under the average width criterion, and by Liu and Hayter (2007) under the minimum area confidence set criterion.

Construction of exact confidence bands is much harder for $p > 2$; there are at least two predictor variables and the region $\mathcal{X}$ may assume various forms. Bohrer (1967) considered a hyperbolic confidence band when the predictor variables are non-negative. One frequently used region $\mathcal{X}$ is the rectangular region

$$\mathcal{X}_R = \{{\mathbf{x}_{(1)}}^T : \ a_i \leq x_i \leq b_i, \ i = 2, \ldots, p\},$$

where $-\infty \leq a_i < b_i \leq \infty, i = 2, \ldots, p$ are given constants. Knafl, Sacks and Ylvisaker (1985) obtained an approximate two-sided hyperbolic confidence band when $p \leq 3$ by using an up-crossing inequality. This approach has been further developed by, among others, Naiman (1986, 1990) Johnstone and Siegmund (1989), Knowles and Siegmund (1989), Johansen and Johnstone (1990), Faraway and Sun (1995), and Sun and Loader (1994), to produce conservative or approximate two-sided hyperbolic simultaneous confidence bands for some linear regression models. All these methods are related to the tube method. In particular, Sun and Loader (1994) assumed that the $p-1$ predictor variables are functions of $q \geq 1$ independent variables (e.g. in polynomial regression models) and provided approximate two-sided hyperbolic band for the regression model when the q independent variables are constrained to intervals, $q = 1$ and 2. Software is available for implementing these approximations, see for example Loader (2004). Sun and Loader's (1994) approach has further been developed by Sun, Loader and McCormick (2000) for more general regression models, including generalized linear regression models. Naiman (1987) proposed a conservative two-sided hyperbolic confidence band by using simulation. Recently, Liu, Jamshidian, Zhang and Donnelly (2005a) proposed a simulation-based method for constructing a two-sided hyperbolic confidence band over $\mathcal{X}_R$ for a general $p \geq 2$; the critical constant can be calculated as accurately as required if the number of replications in the simulation is taken sufficiently large. The Matlab software for implementing this method is given by Jamshidian et al. (2005). This method can also be adapted to construct a one-sided hyperbolic confidence band over $\mathcal{X}_R$. Liu, Jamshidian, Zhang and Bretz (2005b) considered the construction of a two-sided constant width confidence band over $\mathcal{X}_R$ for a general $p \geq 2$ by using both numerical integration and simulation.

The focus of this paper is the construction of exact hyperbolic confidence bands over a specific ellipsoidal region $\mathcal{X}_E$ for a general $p > 2$. Let $X_{(1)}$ be the $n \times (p-1)$ matrix produced from the design matrix X by deleting the first column of X. Let $x_{\cdot j} = (1/n)\sum_{i=1}^{n} x_{ij}$ be the mean of the observed values of the jth predictor variable $(2 \leq j \leq p)$, and let $\bar{\mathbf{x}}_{(1)} = (x_{\cdot 2}, \ldots, x_{\cdot p})^T$. Define the $(p-1) \times (p-1)$ matrix

$$S = \frac{1}{n}\left(X_{(1)} - \underline{1}\bar{\mathbf{x}}_{(1)}^T\right)^T \left(X_{(1)} - \underline{1}\bar{\mathbf{x}}_{(1)}^T\right) = \frac{1}{n}\left(X_{(1)}^T X_{(1)} - n\bar{\mathbf{x}}_{(1)}\bar{\mathbf{x}}_{(1)}^T\right),$$

where $\underline{1}$ is an n-vector of 1's, the sample variance-covariance matrix of the $p-1$ predictor variables; it is non-singular since X is assumed to be of full column-rank. Now the ellipsoidal region is taken to be

$$\mathcal{X}_E = \left\{\mathbf{x}_{(1)} :\ \left(\mathbf{x}_{(1)} - \bar{\mathbf{x}}_{(1)}\right)^T S^{-1} \left(\mathbf{x}_{(1)} - \bar{\mathbf{x}}_{(1)}\right) \leq a^2\right\} \tag{1.1}$$

where $a > 0$ is a given constant.

One can show (see (2.4)below) that the variance of the fitted regression model at $\mathbf{x}$ is given by

$$\mathrm{Var}\,(\mathbf{x}^T\hat{\mathbf{b}}) = \frac{\sigma^2}{n}\left[1 + \left(\mathbf{x}_{(1)} - \bar{\mathbf{x}}_{(1)}\right)^T S^{-1}\left(\mathbf{x}_{(1)} - \bar{\mathbf{x}}_{(1)}\right)\right].$$

So, for all the $\mathbf{x}_{(1)}$ on the surface of the ellipsoid $\mathcal{X}_E$, $Var(\mathbf{x}^T\hat{\mathbf{b}})$ is given by $(\sigma^2/n)\left[1+a^2\right]$; the minimum value of $\mathrm{Var}\,(\mathbf{x}^T\hat{\mathbf{b}})$ is attained at $\mathbf{x} = (1, \bar{\mathbf{x}}_{(1)}^T)^T$. All the $\mathbf{x}_{(1)}$ on the surface of $\mathcal{X}_E$ can therefore be regarded as of equal 'distance', in terms of $\mathrm{Var}\,(\mathbf{x}^T\hat{\mathbf{b}})$, from $\bar{\mathbf{x}}_{(1)}$. Hence it is of interest to learn via a simultaneous confidence band about the regression model over $\mathcal{X}_E$ for a pre-specified a^2 value; note that the width of the two-sided hyperbolic band in (2.2)is a constant on the surface of $\mathcal{X}_E$. If the axes of $\mathcal{X}_E$ coincide with the axes of the coordinates, the design is called orthogonal and, in particular, if $\mathcal{X}_E$ is a sphere the design is called rotatable; see e.g., Atkinson and Donev (1992, p.48).

Halperin and Gurian (1968) constructed a conservative two-sided hyperbolic confidence band over $\mathcal{X}_E$ by using a result of Halperin et al. (1967). Casella and Strawderman (1980) were able to construct an exact two-sided hyperbolic confidence band over a region that is more general than $\mathcal{X}_E$, and this region was further studied by Seppanen and Uusipaikka (1992); see Section 4.2 below. Bohrer (1973) considered the construction of exact one-sided hyperbolic confidence bands over $\mathcal{X}_E$, while Hochberg and Quade (1975) considered the special case of $a = \infty$, that is, exact one-sided hyperbolic confidence bands over the whole space of the predictor variables. One interesting observation made by Bohrer (1973) is that the confidence level of the band can be expressed as a linear combination of several F probabilities. Wynn (1975) extended this observation to another region, though the calculation of the coefficients in the linear combination involves multiple integrals and is non-trivial.

Details given in Bohrer (1973) have some mistakes. This is re-studied, and two new methods are provided for the construction of exact one-sided hyperbolic confidence band over $\mathcal{X}_E$. These are presented in Section 3. Section 4 gives two new methods, together with the result from Casella and Strawderman (1980), for the construction of exact two-sided hyperbolic confidence band over $\mathcal{X}_E$. Section 5 contains a numerical example to illustrate the methodologies discussed in this paper.

2. Preliminaries

2.1. Transformation of the problem

In this subsection the original problems of the construction of exact one-sided and two-sided hyperbolic confidence bands over $\mathcal{X}_E$ are transformed to the formats that will be the starting points of Sections 3 and 4, respectively.

The problems are to construct a one-sided confidence band of the form

$$\mathbf{x}^T\mathbf{b} \geq \mathbf{x}^T\hat{\mathbf{b}} - r\hat{\sigma}\sqrt{\mathbf{x}^T(X^TX)^{-1}\mathbf{x}} \quad \text{for all } \mathbf{x}_{(1)} = (x_2, \ldots, x_p)^T \in \mathcal{X}_E, \tag{2.1}$$

and to construct a two-sided confidence band of the form

$$\mathbf{x}^T\mathbf{b} \in \mathbf{x}^T\hat{\mathbf{b}} \pm r\hat{\sigma}\sqrt{\mathbf{x}^T(X^TX)^{-1}\mathbf{x}} \quad \text{for all } \mathbf{x}_{(1)} = (x_2, \ldots, x_p)^T \in \mathcal{X}_E, \tag{2.2}$$

where $\mathcal{X}_E$ is defined in (1.1). In order to determine the critical constants r in (2.1)and (2.2)so that a confidence band has a confidence level equal to pre-specified $1-\alpha$, the key is to calculate the confidence level of the band for a given r: $\mathrm{P}\{U_1 \leq r\}$ and $\mathrm{P}\{U_2 \leq r\}$, where

$$U_1 = \sup_{\mathbf{x}_{(1)}\in\mathcal{X}_E} \frac{\mathbf{x}^T(\hat{\mathbf{b}}-\mathbf{b})}{\hat{\sigma}\sqrt{\mathbf{x}^T(X^TX)^{-1}\mathbf{x}}}\ , \qquad U_2 = \sup_{\mathbf{x}_{(1)}\in\mathcal{X}_E} \frac{|\mathbf{x}^T(\hat{\mathbf{b}}-\mathbf{b})|}{\hat{\sigma}\sqrt{\mathbf{x}^T(X^TX)^{-1}\mathbf{x}}}\ . \tag{2.3}$$

Note that an upper confidence band uses the same critical constant as the lower confidence band in (2.1).

Let $\boldsymbol{z} = \sqrt{n}(1, \bar{\mathbf{x}}_{(1)}^T)^T$. Note that $\boldsymbol{z}^T(X^TX)^{-1}\boldsymbol{z} = 1$ and hence there exists a $p \times (p-1)$ matrix Z such that $(\boldsymbol{z}, Z)^T(X^TX)^{-1}(\boldsymbol{z}, Z) = I_p$. It follows therefore that $\mathbf{N} = (\boldsymbol{z}, Z)^{-1}(X^TX)(\hat{\mathbf{b}}-\mathbf{b})/\sigma$ is a standard normal random vector of p dimensions. Also note that $\boldsymbol{z}^T(X^TX)^{-1}\mathbf{x} = 1/\sqrt{n}$. Let $\mathbf{w} = (\boldsymbol{z}, Z)^T(X^TX)^{-1}\mathbf{x} = (1/\sqrt{n}, \mathbf{w}_{(1)})$, where $\mathbf{w}_{(1)} = (w_2, \ldots, w_p) = Z^T(X^TX)^{-1}\mathbf{x}$. Then $\mathbf{x}^T(X^TX)^{-1}\mathbf{x} = \mathbf{w}^T\mathbf{w} = \| \mathbf{w} \|^2$. From this, and the fact that the region $\mathcal{X}_E$ in (1.1) can be expressed as

$$\mathcal{X}_E = \left\{ \mathbf{x}_{(1)} :\ \mathbf{x}^T(X^TX)^{-1}\mathbf{x} \leq \frac{1+a^2}{n} \right\}, \tag{2.4}$$

the possible values of $\mathbf{w}_{(1)}$, determined from the relationship $\mathbf{w} = (\boldsymbol{z}, Z)^T(X^TX)^{-1}\mathbf{x}$ when $\mathbf{x}_{(1)}$ varies over the region $\mathcal{X}_E$, form the set

$$\mathcal{W}_E = \left\{ \mathbf{w}_{(1)} : \| \mathbf{w} \|^2 \leq \frac{1+a^2}{n} \right\}. \tag{2.5}$$

The random variable U_1 in (2.3) can now be expressed as

$$\begin{aligned} U_1 &= \sup_{\mathbf{x}_{(1)}\in\mathcal{X}_E} \frac{\left\{(\mathbf{z}, Z)^T(X^TX)^{-1}\mathbf{x}\right\}^T \left\{(\mathbf{z}, Z)^{-1}(X^TX)(\hat{\mathbf{b}}-\mathbf{b})/\sigma\right\}}{(\hat{\sigma}/\sigma)\sqrt{\left\{(\mathbf{z}, Z)^T(X^TX)^{-1}\mathbf{x}\right\}^T \left\{(\mathbf{z}, Z)^T(X^TX)^{-1}\mathbf{x}\right\}}} \\ &= \sup_{\mathbf{w}_{(1)}\in\mathcal{W}_E} \frac{\mathbf{w}^T\mathbf{N}}{(\hat{\sigma}/\sigma) \| \mathbf{w} \|}\ . \end{aligned} \tag{2.6}$$

Furthermore, note from (2.6)that U_1 is invariant if $\mathbf{w}$ is replaced by $\mathbf{v} = u\mathbf{w}$ for any $u > 0$, and that

$$\begin{aligned}\mathcal{V}_E &= \{ \ \mathbf{v} = u\mathbf{w} = (v_1, \ldots, v_p)^T : \ u > 0, \ \mathbf{w}_{(1)} \in \mathcal{W}_E \ \} \\ &= \left\{ \mathbf{v} = (v_1, \ldots, v_p)^T : \| \mathbf{v} \| \leq v_1 \sqrt{1+a^2} \right\} = \{ \mathbf{v} : \ v_1 \geq c \| \mathbf{v} \| \} \subset R^p \quad (2.7)\end{aligned}$$

with $c = 1/\sqrt{1+a^2}$. We therefore have from (2.6)and (2.7)that

$$\mathrm{P}\{U_1 \leq r\} = \mathrm{P}\Big\{ \sup_{\mathbf{v} \in \mathcal{V}_E} \frac{\mathbf{v}^T \mathbf{N}}{(\hat{\sigma}/\sigma) \| \mathbf{v} \|} \leq r \Big\} = \mathrm{P}\Big\{ \mathbf{v}^T \Big\{ \frac{\mathbf{N}}{(\hat{\sigma}/\sigma)} \Big\} \leq r \| \mathbf{v} \| \ \ \forall \ \mathbf{v} \in \mathcal{V}_E \Big\}, (2.8)$$

where $\mathcal{V}_E$ is given in (2.7). This is the starting point of the three methods given in Section 3.

Similarly we have

$$\mathrm{P}\{U_2 \leq r\} = \mathrm{P}\Big\{ \sup_{\mathbf{v} \in \mathcal{V}_E} \frac{|\mathbf{v}^T \mathbf{N}|}{(\hat{\sigma}/\sigma) \| \mathbf{v} \|} \leq r \Big\} = \mathrm{P}\Big\{ \Big| \mathbf{v}^T \Big\{ \frac{\mathbf{N}}{(\hat{\sigma}/\sigma)} \Big\} \Big| \leq r \| \mathbf{v} \| \ \ \forall \ \mathbf{v} \in \mathcal{V}_E \Big\}, (2.9)$$

where $\mathcal{V}_E$ is given in (2.7). This is the starting point of the three methods given in Section 4.

2.2. Polar coordinates

Polar coordinates are used in several places below and so reviewed briefly here. For a p-dimensional vector $\mathbf{v} = (v_1, \ldots, v_p)^T$, define its polar coordinates $(R_{\mathbf{v}}, \theta_{\mathbf{v}1}, \ldots, \theta_{\mathbf{v},p-1})^T$ by

$$\begin{cases} v_1 = R_{\mathbf{v}} \cos \theta_{\mathbf{v}1} \\ v_2 = R_{\mathbf{v}} \sin \theta_{\mathbf{v}1} \cos \theta_{\mathbf{v}2} \\ v_3 = R_{\mathbf{v}} \sin \theta_{\mathbf{v}1} \sin \theta_{\mathbf{v}2} \cos \theta_{\mathbf{v}3} \\ \vdots \qquad \vdots \\ v_{p-1} = R_{\mathbf{v}} \sin \theta_{\mathbf{v}1} \sin \theta_{\mathbf{v}2} \cdots \sin \theta_{\mathbf{v},p-2} \cos \theta_{\mathbf{v},p-1} \\ v_p = R_{\mathbf{v}} \sin \theta_{\mathbf{v}1} \sin \theta_{\mathbf{v}2} \cdots \sin \theta_{\mathbf{v},p-2} \sin \theta_{\mathbf{v},p-1} \ , \end{cases}$$

where

$$\begin{cases} 0 \leq \theta_{\mathbf{v}1} \leq \pi \\ 0 \leq \theta_{\mathbf{v}2} \leq \pi \\ \vdots \qquad \vdots \\ 0 \leq \theta_{\mathbf{v},p-2} \leq \pi \\ 0 \leq \theta_{\mathbf{v},p-1} \leq 2\pi \\ R_{\mathbf{v}} \geq 0 \ . \end{cases}$$

The Jacobian of the transformation is

$$|J| = R_{\mathbf{v}}^{p-1} \sin^{p-2} \theta_{\mathbf{v}1} \sin^{p-3} \theta_{\mathbf{v}2} \cdots \sin \theta_{\mathbf{v},p-2}.$$

When $\mathbf{t} = \mathbf{N}/(\hat{\sigma}/\sigma)$, one can directly find the joint density function of $(R_{\mathbf{t}}, \theta_{\mathbf{t}1}, \ldots, \theta_{\mathbf{t},p-1})^T$. In particular, all the polar coordinates are independent random variables, the marginal density of $\theta_{\mathbf{t}1}$ is

$$f(\theta) = k \sin^{p-2}\theta, \quad 0 \le \theta \le \pi, \tag{2.10}$$

where $k = 1/(\int_0^\pi \sin^{p-2}\theta d\theta)$ is the normalizing constant, and the marginal distribution of $R_{\mathbf{t}}$ is

$$R_{\mathbf{t}} = \left\| \frac{\mathbf{N}}{\hat{\sigma}/\sigma} \right\| \sim \sqrt{pF_{p,\nu}}, \tag{2.11}$$

where $F_{p,\nu}$ denotes an F random variable that has p and ν degrees of freedom.

3. One-sided Hyperbolic Bands

3.1. The method of Bohrer

From (2.8), we have

$$\mathrm{P}\{U_1 \le r\} = \mathrm{P}\{\mathbf{N}/(\frac{\hat{\sigma}}{\sigma}) \in A_{r,1}\}, \tag{3.1}$$

where

$$A_{r,1} = A_{r,1}(c) = \{\mathbf{t} = (t_1, \ldots, t_p)^T : \ \mathbf{v}^T\mathbf{t} \le r\|\mathbf{v}\| \ \ \forall \ \mathbf{v} \text{ in } \mathcal{V}_E\}, \tag{3.2}$$

where $\mathcal{V}_E$ is given in (2.7). This is the form given in Bohrer (1973, p.647, expressions (1.1) and (1.2))

Note that $\mathcal{V}_E$ is a circular cone in R^P with its vertex at the origin and its central direction given by the v_1-axis. The half angle of this circular cone, i.e. the angle between any ray on the boundary of the cone and the v_1-axis, is $\theta^* = \arccos(c)$. Also note that $\mathbf{v}^T\mathbf{t} \le r\|\mathbf{v}\|$ in the definition of $A_{r,1}$ restricts $\mathbf{t}$ to the origin-containing side of the plane that is perpendicular to the vector $\mathbf{v}$, and r-distant from the origin in the direction of $\mathbf{v}$; see Figure 1(a). What is interesting, following the idea of Bohrer (1973), is that $A_{r,1}$ can be partitioned into three sets which can be expressed easily using polar coordinates.

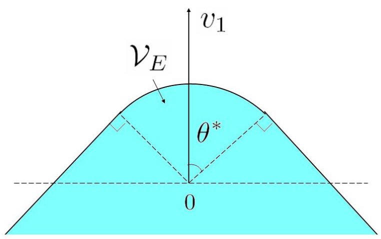

Figure 1(a). The set $A_{r,1}$ and the circular cone V_E.

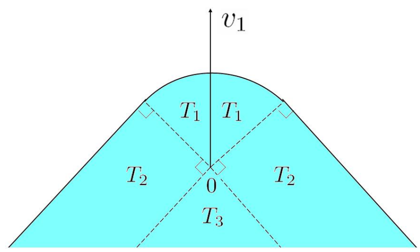

Figure 1(b). The partition of $A_{r,1} = T_1 + T_2 + T_3$.

Lemma 1. *We have* $A_{r,1} = T_1 + T_2 + T_3$ *where*

$$
\begin{aligned}
T_1 &= \{\mathbf{t} : 0 \le \theta_{\mathbf{t}1} \le \theta^*,\ R_{\mathbf{t}} \le r\},\\
T_2 &= \{\mathbf{t} : \theta_{\mathbf{t}1} - \theta^* \in (0, \pi/2]\,,\ R_{\mathbf{t}} \cos(\theta_{\mathbf{t}1} - \theta^*) \le r\}\,,\\
T_3 &= \{\mathbf{t} : \theta^* + \pi/2 < \theta_{\mathbf{t}1} \le \pi\}.
\end{aligned}
$$

This can be proved by introducing the notation f_j for $j = 1, \ldots, p-1$, as in Bohrer (1973); details are omitted here but available from the authors. Note that T_1 is the intersection of the circular cone $\mathcal{V}_E$ and the p-dimension ball centered at the origin with radius r, T_3 is the dual cone of $\mathcal{V}_E$, and $T_2 = T_1 \oplus T_3$; see Figure 1(b). In Bohrer (1973), $A_{r,1}$ is partitioned into four sets where the second and the third sets are in fact the same.

Now the three probabilities can be calculated as follows by using the distributional results in Section 2.2:

$$
\begin{aligned}
\mathrm{P}\{\mathbf{t} \in T_1\} &= \int_0^{\theta^*} k \sin^{p-2}\theta d\theta \cdot \mathrm{P}\{R_{\mathbf{t}} \le r\}\\
&= \int_0^{\theta^*} k \sin^{p-2}\theta d\theta \cdot \mathrm{P}\{pF_{p,\nu} \le r^2\}\\
&= \int_0^{\theta^*} k \sin^{p-2}\theta d\theta \cdot F_{p,\nu}(\frac{r^2}{p}),
\end{aligned}
\tag{3.3}
$$

where $F_{p,\nu}(\cdot)$ denotes the cdf of the random variable $F_{p,\nu}$;

$$
\mathrm{P}\{\mathbf{t} \in T_3\} = \int_{\theta^* + \frac{\pi}{2}}^{\pi} k \sin^{p-2}\theta d\theta = \int_0^{\frac{\pi}{2} - \theta^*} k \sin^{p-2}\theta d\theta; \tag{3.4}
$$

$$
\begin{aligned}
\mathrm{P}\{\mathbf{t} \in T_2\} &= \int_{\theta^*}^{\theta^* + \frac{\pi}{2}} k \sin^{p-2}\theta \cdot \mathrm{P}\{R_{\mathbf{t}} \cos(\theta - \theta^*) \le r\} d\theta\\
&= \int_0^{\frac{\pi}{2}} k \sin^{p-2}(\theta + \theta^*) \cdot \mathrm{P}\left\{R_{\mathbf{t}} \le \frac{r}{\cos\theta}\right\} d\theta\\
&= \int_0^{\frac{\pi}{2}} k \sin^{p-2}(\theta + \theta^*) \cdot F_{p,\nu}\left\{\frac{r^2}{p\cos^2\theta}\right\} d\theta.
\end{aligned}
\tag{3.5}
$$

The confidence level can therefore be calculated from

$$
\mathrm{P}\{U_1 \le r\} = \mathrm{P}\{\mathbf{t} \in T_1\} + \mathrm{P}\{\mathbf{t} \in T_2\} + \mathrm{P}\{\mathbf{t} \in T_3\}. \tag{3.6}
$$

One can express $\mathrm{P}\{\mathbf{t} \in T_2\}$ as a linear combination of several F probabilities by following Bohrer (1973). While this is interesting mathematically, expression (3.5) is easier for numerical calculation and is used in this paper.

When $a = \infty$ it is clear that $c = 0$ and $\theta^* = \pi/2$. In this case T_1 is a half ball in R^p and so

$$\mathrm{P}\{\mathbf{t} \in T_1\} = \frac{1}{2} F_{p,\nu}(\frac{r^2}{p});$$

T_2 is the half cylinder $T_2 = \{\mathbf{t} :\ t_1 < 0,\ \|\mathbf{t}_{(1)}\| \le r\}$ where $\mathbf{t}_{(1)} = (t_2, \ldots, t_p)^T$, and so

$$\mathrm{P}\{\mathbf{t} \in T_2\} = \frac{1}{2} F_{p-1,\nu}(\frac{r^2}{(p-1)});$$

T_3 is empty. Hence the confidence level is

$$\frac{1}{2} F_{p,\nu}(\frac{r^2}{p}) + \frac{1}{2} F_{p-1,\nu}(\frac{r^2}{(p-1)}),$$

which agrees with the result of Hochberg and Quade (1975, expression (2.4)).

3.2. An algebraical method

The key idea of this method is to find the supreme in (2.8) explicitly, as given in Lemma 2 below, from which the confidence level can be evaluated. This approach is similar to that of Casella and Strawderman (1980).

Lemma 2. *We have*

$$\sup_{\mathbf{v} \in \mathcal{V}_E} \frac{\mathbf{v}^T \mathbf{N}}{\|\mathbf{v}\|} = \begin{cases} \|\mathbf{N}\| & \text{if} \quad \mathbf{N} \in \mathcal{V}_E \\ \frac{qN_1 + \|\mathbf{N}_{(1)}\|}{\sqrt{q^2+1}} & \text{if} \quad \mathbf{N} \notin \mathcal{V}_E, \end{cases}$$

where N_1 *is the first element of* $\mathbf{N}$, $\mathbf{N}_{(1)} = (N_2, \ldots, N_p)^T$, *and* $q = \sqrt{c^2/(1-c^2)} = 1/a$.

This can be proved by using basic calculus and the details are omitted. From this lemma, letting $S = \hat{\sigma}/\sigma$, the confidence level of the band is given by

$$\begin{aligned} &\mathrm{P}\left\{ \sup_{\mathbf{v} \in \mathcal{V}_E} \frac{\mathbf{v}^T \mathbf{N}}{\|\mathbf{v}\|} \le rS \right\} \\ &= \mathrm{P}\left\{ \mathbf{N} \in \mathcal{V}_E,\ \|\mathbf{N}\| \le rS \right\} + \mathrm{P}\left\{ \mathbf{N} \notin \mathcal{V}_E,\ \frac{qN_1 + \|\mathbf{N}_{(1)}\|}{\sqrt{q^2+1}} \le rS \right\}. \\ &= \mathrm{P}\left\{ N_1 \ge q\|\mathbf{N}_{(1)}\|, \|N_1\|^2 + \|\mathbf{N}_{(1)}\|^2 \le r^2 S^2 \right\} \\ &\quad + \mathrm{P}\left\{ N_1 < q\|\mathbf{N}_{(1)}\|, \frac{qN_1 + \|\mathbf{N}_{(1)}\|}{\sqrt{q^2+1}} \le rS \right\}. \end{aligned}$$

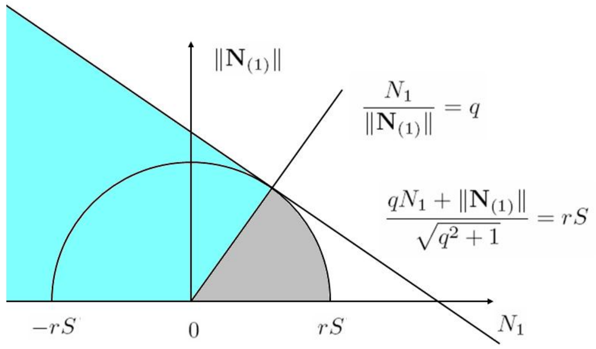

Figure 2. The regions used in Section 3.2.

The two regions

$$\left\{N_1 \geq q\|\mathbf{N}_{(1)}\|,\ \|N_1\|^2 + \|\mathbf{N}_{(1)}\|^2 \leq r^2S^2\right\}$$

and

$$\left\{N_1 < q\|\mathbf{N}_{(1)}\|,\ \frac{qN_1 + \|\mathbf{N}_{(1)}\|}{\sqrt{q^2+1}} \leq rS\right\}$$

are depicted in Figure 2 in the $(N_1, \|\mathbf{N}_{(1)}\|)$-coordinate system. Their union can be re-partitioned into two regions:

$$\{R^2 \leq r^2S^2\} \quad \text{where } R^2 = N_1^2 + \|\mathbf{N}_{(1)}\|^2,$$

$$\left\{r^2S^2 < R^2 < \infty,\ -\infty < \frac{N_1}{\|\mathbf{N}_{(1)}\|} \leq \frac{drS - b\sqrt{R^2 - r^2S^2}}{brS + d\sqrt{R^2 - r^2S^2}}\right\},$$

where $d = q/\sqrt{q^2+1} = c$ and $b = 1/\sqrt{q^2+1} = \sqrt{1-c^2}$. This follows from the fact that, when the point $(N_1, \|\mathbf{N}_{(1)}\|)$ varies on the segment of the circle $N_1^2 + \|\mathbf{N}_{(1)}\|^2 = R^2$ $(> r^2S^2)$, that is within the second region, $N_1/\|\mathbf{N}_{(1)}\|$ attains its minimum value $-\infty$ at the lower end of the circle-segment $(N_1, \|\mathbf{N}_{(1)}\|) = (-R, 0)$, and attains its maximum value $(drS - b\sqrt{R^2-r^2S^2})/(brS + d\sqrt{R^2-r^2S^2})$ at the upper end of the circle-segment whose coordinates can be solved from the simultaneous equations $N_1^2 + \|\mathbf{N}_{(1)}\|^2 = R^2$ and $(qN_1 + \|\mathbf{N}_{(1)}\|)/\sqrt{q^2+1} = rS$. Accordingly, the confidence level is

$$\mathrm{P}\{R^2 \leq r^2S^2\} + \mathrm{P}\left\{r^2S^2 < R^2 < \infty, -\infty < \frac{N_1}{\|\mathbf{N}_{(1)}\|} \leq \frac{drS - b\sqrt{R^2-r^2S^2}}{brS + d\sqrt{R^2-r^2S^2}}\right\}. \quad (3.7)$$

Now note that $R^2 = N_1^2 + \|\mathbf{N}_{(1)}\|^2$ and $N_1/\|\mathbf{N}_{(1)}\|$ are independent random variables, and both are independent of S. So the first probability in (3.7) is

$F_{p,\nu}(r^2/p)$, and the second probability in (3.7) is

$$\mathrm{P}\left\{r^2 < \frac{R^2}{S^2} < \infty, -\infty < \frac{N_1}{\|\mathbf{N}_{(1)}\|} \leq \frac{dr - b\sqrt{(R^2/S^2) - r^2}}{br + d\sqrt{(R^2/S^2) - r^2}}\right\} = \int_{\frac{r^2}{p}}^{\infty} g(w) dF_{p,\nu}(w),$$

where

$$g(w) = \mathrm{P}\left\{-\infty < \frac{N_1}{\|\mathbf{N}_{(1)}\|} \leq \frac{dr - b\sqrt{pw - r^2}}{br + d\sqrt{pw - r^2}}\right\}.$$

Next we express $g(w)$ in term of the cdf of an F distribution. Note that $dr - b\sqrt{pw - r^2} < 0 \iff w > r^2/(pb^2)$. Hence for $w > r^2/(pb^2)$ we have $(dr - b\sqrt{pw - r^2})/(br + d\sqrt{pw - r^2}) < 0$ and

$$\begin{aligned} g(w) &= \frac{1}{2}\mathrm{P}\left\{\frac{N_1^2}{\|\mathbf{N}_{(1)}\|^2} \geq \left(\frac{dr - b\sqrt{pw - r^2}}{br + d\sqrt{pw - r^2}}\right)^2\right\} \\ &= \frac{1}{2} - \frac{1}{2}F_{1,p-1}\left\{(p-1)\left(\frac{dr - b\sqrt{pw - r^2}}{br + d\sqrt{pw - r^2}}\right)^2\right\}; \end{aligned} \tag{3.8}$$

for $r^2/p < w \leq r^2/(pb^2)$ we have $(dr - b\sqrt{pw - r^2})/(br + d\sqrt{pw - r^2}) \geq 0$, and so

$$\begin{aligned} g(w) &= \mathrm{P}\{N_1 \leq 0\} + \mathrm{P}\left\{0 \leq \frac{N_1}{\|\mathbf{N}_{(1)}\|} \leq \frac{dr - b\sqrt{pw - r^2}}{br + d\sqrt{pw - r^2}}\right\} \\ &= \frac{1}{2} + \frac{1}{2}F_{1,p-1}\left\{(p-1)\left(\frac{dr - b\sqrt{pw - r^2}}{br + d\sqrt{pw - r^2}}\right)^2\right\}. \end{aligned} \tag{3.9}$$

Finally, the confidence level is given by

$$F_{p,\nu}\left(\frac{r^2}{p}\right) + \int_{\frac{r^2}{p}}^{\infty} g(w) dF_{p,\nu}(w) \tag{3.10}$$

with the function $g(w)$ being given by (3.8) and (3.9).

3.3. A method based on volume of tubular neighborhoods

This method is based on the volume of tubular neighborhoods of a spherical circular cone and similar to that used by Naiman (1986, 1990) and Sun and Loader (1994), among others. Due to the special form of the cone $\mathcal{V}_E$, the exact volume of its tubular neighborhoods can be calculated easily. From (2.8), the confidence level is given by

$$\mathrm{P}\left\{ \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{\mathbf{v}^T\mathbf{N}}{\|\mathbf{v}\|\|\mathbf{N}\|} \le r\frac{(\hat{\sigma}/\sigma)}{\|\mathbf{N}\|} \right\}, \tag{3.11}$$

where $\mathcal{V}_E$ is given in (2.7). Note that $\mathbf{N}/\|\mathbf{N}\|$ depends on only the $\theta_{\mathbf{N},i}$'s and $\|\mathbf{N}\| = R_{\mathbf{N}}$. Hence $\mathbf{N}/\|\mathbf{N}\|$ is independent of $\|\mathbf{N}\|$ and so $(\hat{\sigma}/\sigma)/\|\mathbf{N}\|$. Furthermore, the supreme in (3.11) is no larger than one, and $r/\sqrt{pw} < 1$ if and only if $w > r^2/p$. So the confidence level is

$$\begin{aligned} &1 - \int_0^\infty \mathrm{P}\left\{ \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{\mathbf{v}^T\mathbf{N}}{\|\mathbf{v}\|\|\mathbf{N}\|} > \frac{r}{\sqrt{pw}} \right\} dF_{p,\nu}(w) \\ &= 1 - \int_{\frac{r^2}{p}}^\infty \mathrm{P}\left\{ \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{\mathbf{v}^T\mathbf{N}}{\|\mathbf{v}\|\|\mathbf{N}\|} > \frac{r}{\sqrt{pw}} \right\} dF_{p,\nu}(w). \end{aligned} \tag{3.12}$$

The key to this method is to find the probability in (3.12). This is facilitated by the following observation.

Lemma 3. *Let* $0 < h < 1$, $\alpha = \arccos(h) \in (0, \pi/2)$, *and* $(R_{\mathbf{N}}, \theta_{\mathbf{N}1}, \ldots, \theta_{\mathbf{N},p-1})$ *be the polar coordinates of* $\mathbf{N}$ *as defined in Section* 2.2. *We have*

$$\left\{\mathbf{N}: \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{\mathbf{v}^T\mathbf{N}}{\|\mathbf{v}\|\|\mathbf{N}\|} > h\right\} = \{\mathbf{N}: \ \theta_{\mathbf{N}1} < \theta^* + \alpha\},$$

where θ^* *is defined in Section* 3.1.

Again the proof involves only calculus and details are omitted. From this lemma and the fact that the pdf of $\theta_{\mathbf{N}1}$ is given in (2.10) since $\theta_{\mathbf{N}1} = \theta_{\mathbf{t}1}$, we have for $0 < h < 1$,

$$\mathrm{P}\left\{ \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{\mathbf{v}^T\mathbf{N}}{\|\mathbf{v}\|\|\mathbf{N}\|} > h \right\} = \mathrm{P}\{\theta_{\mathbf{N}1} < \theta^* + \arccos(h)\} = \int_0^{\theta^*+\arccos(h)} k \sin^{p-2}\theta d\theta.$$

Substituting this expression for the probability in (3.12) with $h = r/\sqrt{pw}$, and changing the order of the double integration, gives the confidence level

$$\begin{aligned} &1 - \int_{\frac{r^2}{p}}^\infty \int_0^{\theta^*+\arccos(h)} k \sin^{p-2}\theta d\theta dF_{p,\nu}(w) \\ &= 1 - \int_0^{\theta^*} \int_{\frac{r^2}{p}}^\infty k \sin^{p-2}\theta dF_{p,\nu}(w) d\theta \\ &\quad - \int_{\theta^*}^{\theta^*+\frac{\pi}{2}} \int_{\frac{r^2}{p\cos^2(\theta-\theta^*)}}^\infty k \sin^{p-2}\theta dF_{p,\nu}(w) d\theta \end{aligned}$$

$$= 1 - \int_0^{\theta^*} k \sin^{p-2}\theta d\theta \cdot \mathrm{P}\left\{F_{p,\nu} > \frac{r^2}{p}\right\}$$
$$- \int_{\theta^*}^{\theta^*+\frac{\pi}{2}} k \sin^{p-2}\theta \cdot \mathrm{P}\left\{F_{p,\nu} > \frac{r^2}{p\cos^2(\theta-\theta^*)}\right\} d\theta.$$

Now by replacing the one in the last expression by

$$1 = \int_0^{\pi} k \sin^{p-2}\theta d\theta = \int_0^{\theta^*} k \sin^{p-2}\theta d\theta + \int_{\theta^*}^{\theta^*+\frac{\pi}{2}} k \sin^{p-2}\theta d\theta + \int_0^{\frac{\pi}{2}-\theta^*} k \sin^{p-2}\theta d\theta,$$

straightforward manipulation gives the confidence level

$$\int_0^{\theta^*} k \sin^{p-2}\theta d\theta \cdot F_{p,\nu}(\frac{r^2}{p})$$
$$+ \int_0^{\frac{\pi}{2}} k \sin^{p-2}(\theta+\theta^*) \cdot F_{p,\nu}\left\{\frac{r^2}{p\cos^2\theta}\right\} d\theta + \int_0^{\frac{\pi}{2}-\theta^*} k \sin^{p-2}\theta d\theta, \quad (3.13)$$

which amounts to (3.6).

Numerical computations for various values of a, r, p and ν confirm that (3.10) and (3.6) are equal. For example, both equal 0.77887 for $a = 2.0$, $r = 2.5$, $p = 6$ and $\nu = \infty$, and equal 0.95620 for $a = 1.5$, $r = 3.0$, $p = 4$ and $\nu = 20$. Among the methods for deriving the confidence level, the third given in Section 3.3 is the simplest. For numerical computation, (3.6) is easier to use than (3.10). A MATLAB program to compute the critical value r for simultaneous confidence level $1 - \alpha$ is available from the authors on request.

4. Two-Sided Hyperbolic Bands

4.1. A method based on Bohrer's approach

This method is similar to Bohrer's (1973) method for one-sided bands given in Section 3.1. From (2.9), the confidence level is

$$\mathrm{P}\Big\{\frac{\mathbf{N}}{(\frac{\hat{\sigma}}{\sigma})} \in A_{r,2}\Big\}, \quad (4.1)$$

where

$$A_{r,2} = A_{r,2}(c) = \{\mathbf{t} = (t_1, \ldots, t_p)^T : \ |\mathbf{v}^T\mathbf{t}| \le r\|\mathbf{v}\| \ \ \forall \ \mathbf{v} \text{ in } \mathcal{V}_E\}, \quad (4.2)$$

and $\mathcal{V}_E$ is given in (2.7).

Note that in the definition of $A_{r,2}$ in (4.2), each $|\mathbf{v}^T\mathbf{t}| \le r\|\mathbf{v}\|$ restricts $\mathbf{t}$ to the origin-containing stripe that is bounded by the two planes which are perpendicular to the vector $\mathbf{v}$ and r-distant from the origin; see Figure 3 where

$A_{r,2}$ is partitioned into four sets that can be expressed using polar coordinates, as given by the following lemma.

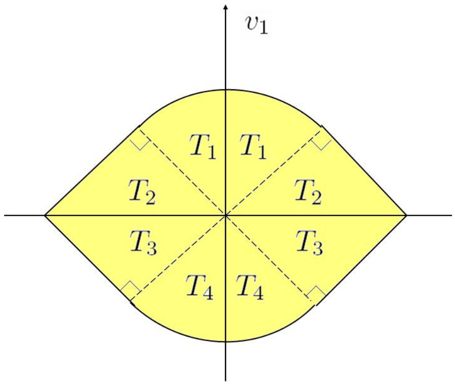

Figure 3. The set $A_{r,2}$ and its partition $A_{r,2} = T_1 + T_2 + T_3 + T_4$.

Lemma 4. *We have* $A_{r,2} = T_1 + T_2 + T_3 + T_4$ *where*

$$\begin{aligned} T_1 &= \{\mathbf{t} : 0 \le \theta_{\mathbf{t}1} \le \theta^*,\ R_{\mathbf{t}} \le r\}, \\ T_2 &= \{\mathbf{t} : \theta^* < \theta_{\mathbf{t}1} \le \frac{\pi}{2},\ R_{\mathbf{t}} \cos(\theta_{\mathbf{t}1} - \theta^*) \le r\}, \\ T_3 &= \{\mathbf{t} : \frac{\pi}{2} < \theta_{\mathbf{t}1} \le \pi - \theta^*,\ R_{\mathbf{t}} \cos(\pi - \theta^* - \theta_{\mathbf{t}1}) \le r\}, \\ T_4 &= \{\mathbf{t} : \pi - \theta^* < \theta_{\mathbf{t}1} \le \pi,\ R_{\mathbf{t}} \le r\}. \end{aligned}$$

The lemma can be proved using basic calculus, details are available from the authors. The four probabilities can be calculated as follows:

$$\begin{aligned} \mathrm{P}\{\mathbf{t} \in T_1\} = \mathrm{P}\{\mathbf{t} \in T_4\} &= \int_0^{\theta^*} k \sin^{p-2}\theta d\theta \cdot \mathrm{P}\{R_{\mathbf{t}} \le r\} \\ &= \int_0^{\theta^*} k \sin^{p-2}\theta d\theta \cdot \mathrm{P}\{pF_{p,\nu} \le r^2\} \\ &= \int_0^{\theta^*} k \sin^{p-2}\theta d\theta \cdot F_{p,\nu}\left(\frac{r^2}{p}\right) \ ; \end{aligned}$$

$$\begin{aligned} \mathrm{P}\{\mathbf{t} \in T_2\} = \mathrm{P}\{\mathbf{t} \in T_3\} &= \int_{\theta^*}^{\frac{\pi}{2}} k \sin^{p-2}\theta \cdot \mathrm{P}\{R_{\mathbf{t}} \cos(\theta - \theta^*) \le r\} d\theta \\ &= \int_0^{\frac{\pi}{2}-\theta^*} k \sin^{p-2}(\theta + \theta^*) \cdot \mathrm{P}\left\{R_{\mathbf{t}} \le \frac{r}{\cos\theta}\right\} d\theta \\ &= \int_0^{\frac{\pi}{2}-\theta^*} k \sin^{p-2}(\theta + \theta^*) \cdot F_{p,\nu}\left\{\frac{r^2}{p\cos^2\theta}\right\} d\theta. \end{aligned}$$

The confidence level is therefore given by

$$\int_0^{\theta^*} 2k\sin^{p-2}\theta d\theta \cdot F_{p,\nu}\left(\frac{r^2}{p}\right) + \int_0^{\frac{\pi}{2}-\theta^*} 2k\sin^{p-2}(\theta+\theta^*)\cdot F_{p,\nu}\left\{\frac{r^2}{p\cos^2\theta}\right\} d\theta. \quad (4.3)$$

4.2. The method of Casella and Strawderman

Using the results of Casella and Strawderman (1980), one can show that the simultaneous confidence level $\mathrm{P}\{U_2 \le r\}$ is given by

$$F_{p,\nu}\left(\frac{r^2}{p}\right) + \int_{\frac{r^2}{p}}^{\frac{r^2}{(b^2p)}} F_{1,p-1}\left\{(p-1)\left(\frac{cr - b\sqrt{pw-r^2}}{br + c\sqrt{pw-r^2}}\right)^2\right\} dF_{p,\nu}(w), \quad (4.4)$$

where $c = 1/\sqrt{1+a^2}$ and $b = a/\sqrt{1+a^2} = \sqrt{1-c^2}$ as before. We refer the reader to Casella and Strawderman (1980) for details, similar to those given in Section 3.2.

In fact, Casella and Strawderman (1980) considered the construction of an exact two-sided hyperbolic confidence band over a constrained region of the predictor variables whose simultaneous confidence level can be reduced to the form

$$\mathrm{P}\left\{\sup_{\mathbf{v}\in\mathcal{V}^*(m)} \frac{|\mathbf{v}^T\mathbf{N}|}{(\hat{\sigma}/\sigma)\parallel \mathbf{v}\parallel} \le r\right\}, \text{ where } \mathcal{V}^*(m) = \left\{\mathbf{v} : \sum_{i=1}^m v_i^2 \ge \frac{c^2}{1-c^2}\sum_{i=m+1}^p v_i^2\right\}$$

and $1 \le m \le p$ is a given integer. Note that $\mathcal{V}_E = \mathcal{V}^*(1)$. Seppanen and Uusipaikka (1992) provided an explicit form of the predictor variable region over which the two-sided hyperbolic confidence band has its simultaneous confidence level given by this form. It is noteworthy that, for $2 \le m \le p$, the predictor variable region corresponding to $\mathcal{V}^*(m)$ is not bounded, and so is of less interest in concrete applications. One can, however, find a conservative two-sided hyperbolic band over the rectangular predictor variable region $\mathcal{X}_R$: use the predictor variable regions corresponding to $\mathcal{V}^*(m)$ for $1 \le m \le p$ to bound the given $\mathcal{X}_R$, calculate the critical values for these $\mathcal{V}^*(m)$'s, and use the smallest calculated critical values as a conservative critical value for the confidence band over $\mathcal{X}_R$.

4.3. A method based on volume of tubular neighborhoods

From (2.9), the confidence level is given by

$$\mathrm{P}\left\{\sup_{\mathbf{v}\in\mathcal{V}_E} \frac{|\mathbf{v}^T\mathbf{N}|}{\parallel \mathbf{v}\parallel\parallel \mathbf{N}\parallel} \le r\frac{(\frac{\hat{\sigma}}{\sigma})}{\parallel \mathbf{N}\parallel}\right\}, \quad (4.5)$$

where $\mathcal{V}_E$ is given in (2.7). Similar to the one-sided case in Section 3.3, the confidence level is further equal to

$$1 - \int_{\frac{r^2}{p}}^{\infty} \mathrm{P}\left\{ \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{|\mathbf{v}^T\mathbf{N}|}{\| \mathbf{v} \|\| \mathbf{N} \|} > \frac{r}{\sqrt{pw}} \right\} dF_{p,\nu}(w). \tag{4.6}$$

The key to this method is to find the probability in (4.6), which hinges on the following result.

Lemma 5. *Let* $0 < h = r/\sqrt{pw} < 1$, $\alpha = \arccos(h) \in (0, \pi/2)$, *and* $(R_{\mathbf{N}}, \theta_{\mathbf{N}1}, \ldots, \theta_{\mathbf{N},p-1})$ *be the polar coordinates of* $\mathbf{N}$, *and* $\theta^* = \arccos(c)$. *We have*

$$\left\{\mathbf{N} : \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{\mathbf{v}^T\mathbf{N}}{\| \mathbf{v} \|\| \mathbf{N} \|} > h\right\} = \begin{cases} \{\mathbf{N} :\ \theta_{\mathbf{N}1} \in [0,\ \theta^* + \alpha] \cup [\pi - \theta^* - \alpha,\ \pi]\} & \text{if} \quad \theta^* + \alpha < \frac{\pi}{2} \\ \{\mathbf{N} :\ \theta_{\mathbf{N}1} \in [0,\ \pi]\} & \text{if} \quad \theta^* + \alpha \geq \frac{\pi}{2} \end{cases}.$$

The proof involves only calculus, details are omitted. From this lemma, that $\theta^* + \arccos(r/\sqrt{pw}) < \pi/2$ if and only if $w < r^2/(b^2p)$, and that the pdf of $\theta_{\mathbf{N}1}$ is given in (2.10), we have, for $r^2/p \leq w < r^2/(b^2p)$,

$$\begin{aligned} &\mathrm{P}\left\{ \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{\mathbf{v}^T\mathbf{N}}{\| \mathbf{v} \|\| \mathbf{N} \|} > \frac{r}{\sqrt{pw}} \right\} \\ &= \mathrm{P}\left\{ \ \theta_{\mathbf{N}1} \in [0,\ \theta^* + \arccos(\frac{r}{\sqrt{pw}})] \cup [\pi - \theta^* - \arccos(\frac{r}{\sqrt{pw}}),\ \pi] \right\} \\ &= 2\int_0^{\theta^* + \arccos(\frac{r}{\sqrt{pw}})} k \sin^{p-2}\theta d\theta, \end{aligned}$$

and for $w \geq r^2/(b^2p)$,

$$\mathrm{P}\left\{ \sup_{\mathbf{v}\in\mathcal{V}_E} \frac{\mathbf{v}^T\mathbf{N}}{\| \mathbf{v} \|\| \mathbf{N} \|} > \frac{r}{\sqrt{pw}} \right\} = \mathrm{P}\{ \ \theta_{\mathbf{N}1} \in [0,\ \pi]\} = 1.$$

Substituting these two expressions into (4.6), the confidence level is

$$1 - \int_{\frac{r^2}{p}}^{\frac{r^2}{(b^2p)}} \int_0^{\theta^* + \arccos(\frac{r}{\sqrt{pw}})} 2k \sin^{p-2}\theta d\theta dF_{p,\nu}(w) - \int_{\frac{r^2}{(b^2p)}}^{\infty} 1 dF_{p,\nu}(w). \tag{4.7}$$

By changing the order of integrations, it is straightforward to show that the double integral above is

$$\int_0^{\theta^*} 2k \sin^{p-2}\theta d\theta \cdot \left\{ F_{p,\nu}\left(\frac{r^2}{b^2p}\right) - F_{p,\nu}\left(\frac{r^2}{p}\right) \right\}$$

$$+\int_{\theta^*}^{\frac{\pi}{2}} 2k\sin^{p-2}\theta\cdot\left\{F_{p,\nu}\left(\frac{r^2}{b^2p}\right)-F_{p,\nu}\left(\frac{r^2}{p\cos^2(\theta-\theta^*)}\right)\right\}d\theta.$$

By substituting this into (4.7), it is clear that the confidence level is equal to the expression given in (4.3).

Numerical computations have been done to confirm that the results computed from (4.3) and (4.4) agree with the entries of Seppanen and Uusipaikka (1992, Table 1 for $r = 1$). A MATLAB program to compute the critical value r for given values of p, $\nu = n - p$, a and $1 - \alpha$ is available from the authors on request.

5. A Numerical Example

In this section, a portion of the acetylene data in Snee (1977) is used to illustrate the methodologies discussed in the paper; the same data set was also used by Casella and Strawderman (1980) and Naiman (1987). The two predictor variables are reactor temperature (x_2) and the ratio of H_2 to n-Heptane (x_3). The response variable (y) is conversion of n-Heptane to Acetylene. There are sixteen data points, so $p = 3$, $n = 16$ and $\nu = 13$. The fitted linear regression model is $y = -130.69 + 0.134x_2 + 0.351x_3$, $\hat{\sigma} = 3.624$, and $R^2 = 0.92$.

The observed values of x_2 range from 1,100 to 1,300 with average $x_{\cdot 2} = 1,212.5$, and the observed values of x_3 range from 5.3 to 23 with average $x_{\cdot 3} = 12.4$. The ellipsoidal region $\mathcal{X}_E$ is centered at $(x_{\cdot 2}, x_{\cdot 3})^T = (1,212.5, 12.4)^T$, and the size of $\mathcal{X}_E$ increases with the value of a. Figure 4 shows five $\mathcal{X}_E$'s corresponding to $a = 0.1(0.6)2.5$ respectively. The rectangular region indicates the observed range $[1,100, 1,300] \times [5.3, 23]$ of the predictor variables $(x_2, x_3)^T$.

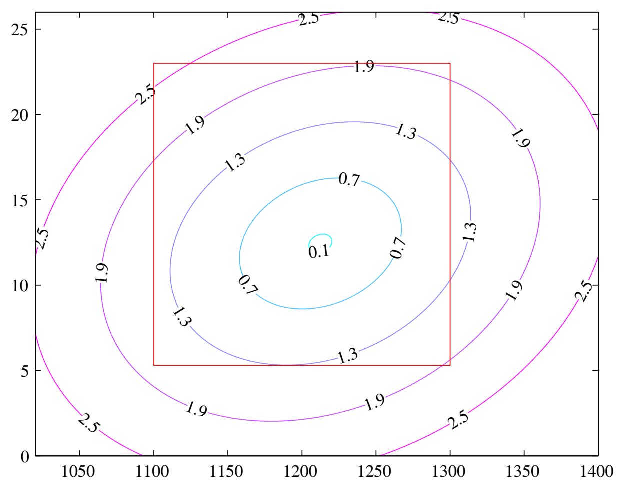

Figure 4. Several ellipsoidal regions $\mathcal{X}_E$.

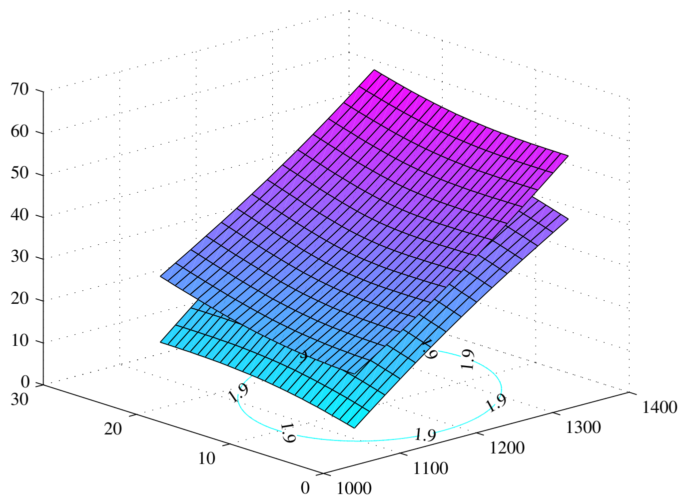

Figure 5. The 90% two-sided hyperbolic band over the $\mathcal{X}_E$.

For a given $\mathcal{X}_E$, one can use a two-sided confidence band to quantify the plausible range of the unknown regression model over $\mathcal{X}_E$. Suppose $a = 1.9$, so $\mathcal{X}_E$ is given by the second largest ellipse in Figure 4, and the simultaneous confidence level is $1 - \alpha = 90\%$. Then our MATLAB program calculates $r = 2.7229$ for the two-sided hyperbolic band. This confidence band is plotted in Figure 5; the band is given only by the part that is inside the cylinder in the y-direction, and has the $\mathcal{X}_E$ as the cross-section in the $(x_2, x_3)^T$-plane. Note from the discussion immediately below (1.1) that the width of this confidence band on the boundary of the $\mathcal{X}_E$ is

$$2r\hat{\sigma}\sqrt{\mathbf{x}^T(X^TX)^{-1}\mathbf{x}} = 2r\hat{\sigma}\sqrt{\frac{1+a^2}{n}} = 2 * 2.723 * 3.624 * \sqrt{\frac{1+1.9^2}{16}} = 10.594.$$

If one wants to learn how low the true model can plausibly be over $\mathcal{X}_E$, a lower hyperbolic band over $\mathcal{X}_E$ can be used. For this data set with $a = 1.9$ and $1 - \alpha = 90\%$, our MATLAB program calculates $r = 2.3697$ for the one-sided hyperbolic band.

Acknowledgements

We would like to thank the Editor, an associate editor and the referees for several critical and constructive comments.

References

Al-Saidy, O. M., Piegorsch, W. W., West, R. W. and Nitcheva, D. K. (2003). Confidence bands for low-dose risk estimation with quantal response data. *Biometrics* **59**, 1056-1062.

Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*. Oxford University Press.

Bohrer, R. (1967). On sharpening Scheffe bounds. *J. Roy. Statist. Soc. Ser. B* **29**, 110-114.

Bohrer, R. (1973). A multivariate t probability integral. *Biometrika* **60**, 647-654.

Bohrer, R. and Francis, G. K. (1972). Sharp one-sided confidence bands for linear regression over intervals. *Biometrika* **59**, 99-107.

Bowden, D. C. and Graybill, F. A. (1966). Confidence bands of uniform and proportional width for linear models. *J. Amer. Statist. Assoc.* **61**, 182-198.

Bowden, D. C. (1970). Simultaneous confidence bands for linear regression models. *J. Amer. Statist. Assoc.* **65**, 413-421.

Casella, G. and Strawderman, W. E. (1980). Confidence bands for linear-regression with restricted predictor variables. *J. Amer. Statist. Assoc.* **75**, 862-868.

Faraway, J. J. and Sun, J. (1995). Simultaneous confidence bands for linear regression with heteroscedastic errors. *J. Amer. Statist. Assoc.* **90**, 1094-1098.

Gafarian, A. V. (1964). Confidence bands in straight line regression. *J. Amer. Statist. Assoc.* **59**, 182-213.

Halperin, M. and Gurian, J. (1968). Confidence bands in linear regression with constraints on independent variables. *J. Amer. Statist. Assoc.* **63**, 1020-1027.

Halperin, M., Rastogi, S. C., Ho, I. and Yang, Y. Y. (1967). Shorter confidence bands in linear regression. *J. Amer. Statist. Assoc.* **62**, 1050-1067.

Hochberg, Y. and Quade, D. (1975). One-sided simultaneous confidence bounds on regression surfaces with intercepts. *J. Amer. Statist. Assoc.* **70**, 889-891.

Jamshidian, M., Liu, W., Zhang, Y. and Jamshidian, F. (2005). SimReg: a software including some new developments in multiple comparison and simultaneous confidence bands for linear regression models. *J. Statist. Soft.* **12**, 1-22.

Johansen, S. and Johnstone, I. (1990). Hotelling's theorem on the volume of tubes: some illustrations in simultaneous inference and data analysis. *Ann. Statist.* **18**, 652-684.

Johnstone, I. and Siegmund, D. (1989). On Hotelling's formula for the volume of tubes and Naiman's inequality. *Ann. Statist.* **17**, 184-194.

Knafl, G., Sacks, J. and Ylvisaker, D. (1985). Confidence bands for regression-functions. *J. Amer. Statist. Assoc.* **80**, 683-691.

Knowles, M. and Siegmund, D. (1989). On Hotelling's approach to testing for a nonlinear parameter in regression. *Internat. Statist. Rev.* **57**, 205-220.

Liu, W. and Hayter, A. J. (2007). Minimum area confidence set optimality for confidence bands in simple linear regression. *J. Amer. Statist. Assoc.* **102**, 181-190.

Liu, W., Jamshidian, M. and Zhang, Y. (2004). Multiple comparison of several regression models. *J. Amer. Statist. Assoc.* **99**, 395-403.

Liu, W., Jamshidian, M., Zhang, Y. and Donnelly, J. (2005a). Simulation-based simultaneous confidence bands for a multiple linear regression model when the covariates are constrained. *J. Comput. Graph. Statist.* **14**, 459-484.

Liu, W., Jamshidian, M., Zhang, Y. and Bretz, F. (2005b). Constant width simultaneous confidence bands for multiple linear regression when predictor variables are constrained in intervals. *J. of Statistical Computation and Simulation* **75**, 425-436.

Loader, C. (2004). The volume-of-tube formula: computational methods and statistical applications. Manuscript, Department of Statistics, Case Western Reserve University, USA.

Miller, R. G. (1981). *Simultaneous Statistical Inference*. 2nd edition. Springer-Verlag, New York.

Naiman, D. Q. (1983). Comparing Scheffé-type to constant-width confidence bands in regression. *J. Amer. Statist. Assoc.* **78**, 906-912.

Naiman, D. Q. (1986). Conservative confidence bands in curvilinear regression. *Ann. Statist.* **14**, 896-906.

Naiman, D. Q. (1987). Simultaneous confidence-bounds in multiple-regression using predictor variable constraints. *J. Amer. Statist. Assoc.* **82**, 214-219.

Naiman, D. Q. (1990). On volumes of tubular neighborhoods of spherical polyhedra and statistical inference. *Ann. Statist.* **18**, 685-716.

Pan, W., Piegorsch, W. W. and West, R. W. (2003). Exact one-sided simultaneous confidence bands via Uusipaikka's method. *Annals of the Institute of Statistical Mathematics* **55**, 243-250.

Piegorsch, W. W., West, R. W., Al-Saidy, O. M. and Bradley, K. D. (2000). Asymmetric confidence bands for simple linear regression over bounded intervals. *Comput. Statist. Data Anal.* **34**, 193-217.

Piegorsch, W. W., West, R. W., Pan, W. and Kodell, R. (2005). Low dose risk estimation via simultaneous statistical inferences. *J. Roy. Statist. Soc. Ser. C*, 245-258.

Scheffé, H. (1953). A method for judging all contrasts in analysis of variance. *Biometrika* **40**, 87-104

Seppanen, E. and Uusipaikka, E. (1992). Confidence bands for linear-regression over restricted regions. *Scand. J. Statist.* **19**, 73-81.

Snee, R. D. (1977). Validation of regression models: methods and example. *Technometrics* **19**, 415-428.

Spurrier, J. D. (1999). Exact confidence bounds for all contrasts of three or more regression lines. *J. Amer. Statist. Assoc.* **94**, 483-88.

Sun, J. and Loader, C. R. (1994). Simultaneous confidence bands for linear regression and smoothing. *Ann. Statist.* **22**, 1328-1346.

Sun, J., Loader, C. R. and McCormick, W. P. (2000). Confidence bands in generalized linear models. *Ann. Statist.* **28**, 429-460.

Sun, J., Raz, J. and Faraway, J. J. (1999). Simultaneous confidence bands for growth and response curves. *Statistica Sinica* **9**, 679-698.

Uusipaikka, E. (1983). Exact confidence bands for linear-regression over intervals. *J. Amer. Statist. Assoc.* **78**, 638-644.

Working, H. and Hotelling, H. (1929). Applications of the theory of error to the interpretation of trends. *J. Amer. Statist. Assoc.* **24**, 73-85.

Wynn, H. P. and Bloomfield, P. (1971). Simultaneous confidence bands in regression analysis. *J. Roy. Statist. Soc. Ser. B* **33**, 202-217

Wynn, H. P. (1975). Integrals for one-sided confidence bounds: A general result. *Biometrika* **62**, 393-396.

S^3RI and School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK.

E-mail: W.Liu@maths.soton.ac.uk

S^3RI and School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK.

E-mail: S.Lin@soton.ac.uk

(Received August 2006; accepted May 2007)