A generalized projection pursuit procedure and its significance level

Kanta NAITO

ABSTRACT. A generalized projection pursuit procedure which maximizes *nonellipticity* is proposed. Behaviors of the maxima of a generalized moment index are investigated under elliptically symmetric distributions. An approximation formula of the significant level is derived. Examples which illustrate the present concepts and results are discussed. Performances of the maxima of the proposed index under certain interesting structure are asymptotically evaluated.

0 Introduction

Exploratory projection pursuit aims to explore nonlinear structures of high-dimensional data through its projection to a low-dimensional subspace. A basic component of projection pursuit is its projection index which is a function of direction to which data is projected and is used to measure the departure from normality (see Huber (1985)). Let X be a p-dimensional random vector and α a p-dimensional unit vector. When we consider one-dimensional exploratory projection pursuit, a function $I(\alpha)$, which measures the departure of $\alpha'X$ from normality is employed as a projection index, where the prime stands for the transpose of a matrix. In the two-dimensional case, let α and β be two unit vectors satisfying $\alpha'\beta = 0$. Then the nonnormality of $(\alpha'X, \beta'X)$ is measured by $I(\alpha, \beta)$, a function of α and β. Various types of projection indices were discussed in Huber (1985), as functionals satisfying affine invariance property. Friedman (1987), Jones and Sibson (1987) and Hall (1989) proposed projection indices based on the orthogonal polynomials. Unified views of projection indices based on the orthogonal polynomials were given by Cook, Buja and Cabrera (1993) and Iwasaki (1989). Sun (1993) gave practical comparisons of Friedman index and Hall index. Friedman index and Jones and Sibson index (the monent index) have been extended to the case of two-dimensional projection pursuit, and recent work of Nason (1995) discussed the moment index for the three-dimensional case. Further, Posse (1995) proposed a new projection index for the two-dimensional case. Several other

---

projection indices have been proposed for various purposes such as smoothing; projection pursuit regression (PPR, Friedman and Stuetzle (1981)), projection pursuit density estimation (PPDE, Friedman, Stuetzle and Schroeder (1984)). We can see from their works that projection pursuit is one of the most powerful exploratory data analysis procedures and includes some potential interest as a computer intensive statistical method.

In this paper, we aim to generalize exploratory projection pursuit procedures. As mentioned in Friedman (1987), when we want to discover interesting structures, elliptically symmetric structure is the least interesting, and the framework that normal distribution is the least interesting is only due to the computational tractability of the projection index. From this point of view, we construct a projection index which maximizing *nonellipticity*. There are several elliptically symmetric structures such as multivariate *t*-distribution, contaminated normal distribution, symmetric Kotz type distribution and multivariate Pearson type distribution. The proposed pursuit procedure based on a specified elliptically symmetric structure regards its structure as the least interesting.

The index prososed in this paper, which is called *a generalized moment index*, includes the moment index proposed by Jones and Sibson (1987) as a special case. The moment index due to Jones and Sibson (1987) is based on the weighted sum of squares of skewness and kurtosis, which is an approximation to simultaneously proposed entropy index and is called a summary statistic for nonlinear structure. From a computational point of view, the usual or general moment index seems to be attractive since they have simple functional form. Another projection indices such as Friedman index, entropy index and Hall index have somewhat more complicated form than the moment index. Further, not only in computational and descriptive point of view but also in an inferential point of view, the moment index seems to be tractable.

The inferential theory of exploratory projection pursuit has not been well established yet. When "nonlinear structures" are discovered by some projection index $I(\alpha)$, it is important to know whether these are really structured or not. Projection pursuit is only a descriptive method unless we establish a criterion to evaluate these structures. This argument may be found in Miller (1985), Jones and Sibson (1987) and Friedman (1987). An approach to this question is to evaluate performances of $I(\alpha)$ under the least interesting structure. Let $a = \max_{\alpha \in \mathscr{S}^{p-1}} I(\alpha)$ be the maxima of a projection index $I(\alpha)$ from the observed data, where $\mathscr{S}^{p-1}$ is the unit sphere in R^p. If we can calculate a P value

$$p_{obs} = P\left\{ \max_{\alpha \in \mathscr{S}^{p-1}} I(\alpha) \geq a \right\} \tag{0.1}$$

under the least interesting structure, it is possible to judge whether apparent
nonlinearity is due to just the effect of noise or not. (0.1) is called *significance
level* in projection pursuit (see Sun (1991)). Under the case that normality is
the least interesting, Sun (1991) derived an approximation formula of (0.1)
when $I(\alpha)$ is Friedman index. Naito (1996) also derived an approximation
formula of (0.1) under when $I(\alpha)$ is a *higher order* version of the moment
index. (0.1) can be evaluated by Monte Carlo methods, however, in this case
they are computationally expensive. Thus deriving an approximation formula
of (0.1) seems to be very useful. In conjunction with this problem, Sun
(1993)'s result for the tail probability of the maxima of Gaussian random field
may be noted. The result was established by using *the tube method* briefly
discussed in Johansen and Johnstone (1990). Practically, Sun's approximation
formula for Friedman index was derived by using the result. Since this
corresponds to obtaining the distribution of a goodness of fit test statistic under
null hypothesis, in this paper, behaviors of the index under the least interesting
structure will be called evaluating *Null Behaviors*.

This paper is divided into two parts. Part I of this paper is mainly Null
Behaviors of a generalized moment index. The definition of a generalized
moment index and its asymptotic behaviors under the least interesting structure
(ellipticity) are discussed in Section 1. Using the result of Sun (1993), we will
derive an approximation formula of (0.1) for a generalized moment index in
Section 2. Section 3 gives some examples for elliptically symmetric distri-
butions and the quantities included in the theoretical results established in
Sections 1 and 2. A simple comparison of a generalized moment index with
the usual moment index through real data is given in conjunction with the
efficiency of pursuing nonellipticity. Monte Carlo experiments to check the
accuracy of the approximation formula under some elliptical structures are also
made in Section 3. Section 4 devotes to proofs and related calculations for
Sections 1 and 2.

Part II of this paper is concerned with the behaviors of a generalized
moment index under interesting structure. That is, the performances of the
index when a specified elliptical structure X is the least interesting are
investigated under another elliptical structure X^* satisfying certain condi-
tions. We call the behaviours *Nonnull Behaviors*. In Section 5, asymptotic
results of Nonnull Behaviors will be presented. Under certain conditions, it
will be shown that the proposed index converges weakly to a Gaussian random
field on $\mathscr{S}^{p-1}$. This result implies that the methods applied to Null Behaviors
can also be applied to Nonnull Behaviors. Proofs and the related calculations
for the results in Section 5 are presented in Section 6.

Part I Null Behaviors

1 A generalized moment index

1.1 Construction of the index

In this subsection, we construct a generalized moment index as a projection index. From an inferential point of view, it is important to know the magnitude of the departure from normality of high-dimensional data in hand. These may be near from normality, on the other hand, these may be far away from normality even if we transpose the data. When a data set seems to be distributed nearly as normal, giving some information about the position of the data in the family of elliptically symmetric distributions is useful, since we may deal the data set as a random sample from an elliptically symmetric distribution. Thus, the projection index should have the property that if a data set is nearly normal, then it detects some elliptically symmetric distribution, and if it is not, then it detects some nonlinear structure. These consideration motivates us to construct the index which maximizes non-ellipticity. Further, as noted in Section 0, some simple functional form is desirable to the index, since we want to establish the distributional results simultaneously. From these considerations, we proceed to generalize the moment index proposed by Jones and Sibson (1987) to maximize nonellipticity.

Let $X_1, \ldots, X_N$ be N independent observations on X with mean μ and nonsingular covariance matrix Σ. As usual, before imprementing the projection pursuit, we make to sphere the data by $S^{-1/2}(X_j - \overline{X})$ for $j = 1, \ldots, N$, where $\overline{X}$ and S are the sample mean vector and the sample covariance matrix, respectively, and $S^{-1/2}$ is a square root of S^{-1}. Let

$$H_k(x) = \exp\left(\frac{x^2}{2}\right)\left(\frac{-d}{dx}\right)^k \exp\left(-\frac{x^2}{2}\right)$$

be the k-th order Hermite polynomial. The moment index proposed by Jones and Sibson (1987) is defined as

$$\mathcal{MI}(\alpha) = \sum_{k=3}^{4}\left[\frac{1}{N}\sum_{j=1}^{N}\left\{\frac{H_k(\alpha'S^{-1/2}(X_j - \overline{X}))}{\sqrt{k!}}\right\}\right]^2,$$

which evaluates nonnormality in terms of up to forth moment (cumulant).

Define that $\theta_k = E[H_k(\alpha'\Sigma^{-1/2}(X - \mu))]$ for $k = 3, 4$, where the expectation is taken under the hypothesis that X has an elliptically symmetric distribution with mean vector μ and covariance matrix Σ. It is easy to see that $\theta_3 = 0$ and

$$\theta_4 = 3\left\{\frac{m_4}{p(p+2)} - 1\right\},$$

where $m_4 = E[\{(X - \mu)'\Sigma^{-1}(X - \mu)\}^2]$. Let

$$\sigma_3^2 = A - Var[H_3(\alpha'S^{-1/2}(X_1 - \bar{X}))],$$

$$\sigma_4^2 = A - Var[H_4(\alpha'S^{-1/2}(X_1 - \bar{X}))],$$

where $A - Var$ designates the asymptotic variance which is taken under the ellipticity hypothesis. A generalized moment index is defined as

$$\mathscr{GMI}(\alpha) = \sum_{k=3}^{4}\left[\frac{1}{N}\sum_{j=1}^{N}\left\{\frac{H_k(\alpha'S^{-1/2}(X_j - \bar{X})) - \theta_k}{\sigma_k}\right\}\right]^2. \qquad (1.1)$$

If $X \sim \mathscr{N}(\mu, \Sigma)$, p-variate normal distribution with mean μ and covariance matrix Σ, then $\theta_4 = 0$ and $\sigma_k^2 = k!$ for $k = 3, 4$ which imply $\mathscr{GMI} \equiv \mathscr{MI}$. So that we say $\mathscr{GMI}$, a generalized moment index, throughout this paper. From the definition of $\mathscr{GMI}$, we may see that this index also simultaneously evaluates nonellipticity via the weighted sum of squares of skewness and kurtosis.

From an inferential point of view, as noted in Section 0, it is important to judge whether the apparent structure based on $\mathscr{GMI}$ is real or due to some sample fluctuation. For this, we evaluate

$$p_{obs} = P\left\{\max_{\alpha \in \mathscr{S}^{p-1}} \mathscr{GMI}(\alpha) \geq z\right\}, \qquad (1.2)$$

where z is the maxima of $\mathscr{GMI}$ for the observed data set. In the following, we get an approximation formula of (1.2). However, it is difficult to get the exact formula of (1.2) in finite sample size N. Therefore, we attempt to derive some asymptotic results for the maxima of $\mathscr{GMI}$ as N tends to infinity. We proceed asymptotic theory for the maxima of $\mathscr{GMI}$ in the next subsection.

1.2 Asymptotic theory

Throughout this subsection we assume that the distribution of X is elliptically symmetric with center μ and ellipticity matrix Δ, that is, X is distributed as $\mu + A'Y$, where A is a nonsingular $p \times p$-matrix satisfying $A'A = \Delta$ and Y is a random vector having spherically symmetric distribution. Further we assume $P(X = \mu) = 0$. Since the maxima of $\mathscr{GMI}$ is invariant under affine linear transformations of the data, we may assume without loss of generality that X is distributed as a spherically symmetric with $E[XX'] = I_p$. This implies that $E[|X|^2] = p$, where $|\cdot|$ stands for the Euclidean norm. Generally, we define $m_k = E[|X|^k]$, for $k \geq 1$.

In this subsection, we investigate asymptotic behaviors of the maxima of $\mathscr{GMI}$ under the least interesting structure (ellipticity). For this, asymptotic results for some kind of random fields defined on $\mathscr{S}^{p-1}$ are needed. Let $C_2(\mathscr{S}^{p-1})$ be the separable Banach space of R^2 valued continuous functions defined on $\mathscr{S}^{p-1}$, endowed with the supremum norm. In what follows, $V_N(\cdot) \Rightarrow V(\cdot)$ means that the distribution of random element $V_N(\cdot)$ of $C_2(\mathscr{S}^{p-1})$ converges weakly to the distribution of a random element $V(\cdot)$ of $C_2(\mathscr{S}^{p-1})$. Let us introduce random fields

$$W_N(\alpha) = \begin{bmatrix} \tilde{W}_{3,N}(\alpha)/\sigma_3 \\ \tilde{W}_{4,N}(\alpha)/\sigma_4 \end{bmatrix}, \qquad Z_N(\alpha) = \begin{bmatrix} \tilde{Z}_{3,N}(\alpha)/\sigma_3 \\ \tilde{Z}_{4,N}(\alpha)/\sigma_4 \end{bmatrix}, \qquad (1.3)$$

$$\tilde{W}_N(\alpha) = \begin{bmatrix} \tilde{W}_{3,N}(\alpha) \\ \tilde{W}_{4,N}(\alpha) \end{bmatrix}, \qquad \tilde{Z}_N(\alpha) = \begin{bmatrix} \tilde{Z}_{3,N}(\alpha) \\ \tilde{Z}_{4,N}(\alpha) \end{bmatrix}, \qquad (1.4)$$

where

$$\tilde{W}_{3,N}(\alpha) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} H_3(\alpha' S^{-1/2}(X_j - \bar{X})),$$

$$\tilde{W}_{4,N}(\alpha) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} [H_4(\alpha' S^{-1/2}(X_j - \bar{X})) - \theta_4],$$

$$\tilde{Z}_{3,N}(\alpha) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} [(\alpha' X_j)^3 - 3\alpha' X_j],$$

$$\tilde{Z}_{4,N}(\alpha) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} [(\alpha' X_j)^4 - (\theta_4 + 3)\{2(\alpha' X_j)^2 - 1\}].$$

From the definition of $\mathscr{GMI}$, we see that

$$N\mathscr{GMI}(\alpha) = \left\{ \frac{\tilde{W}_{3,N}(\alpha)}{\sigma_3} \right\}^2 + \left\{ \frac{\tilde{W}_{4,N}(\alpha)}{\sigma_4} \right\}^2 = |W_N(\alpha)|^2. \qquad (1.5)$$

We present some asymptotic results in the following. Proofs are all given in Section 4.

LEMMA 1.1 *Let X have a spherically symmetric distribution with unit covariance matrix such that $m_8 < \infty$. Then*

$$\sup_{\alpha \in \mathscr{S}^{p-1}} |\tilde{Z}_N(\alpha) - \tilde{W}_N(\alpha)| \xrightarrow{P} 0$$

as $N \to \infty$.

From Lemma 1.1, we see that

$$\tilde{W}_N(\alpha) = \tilde{Z}_N(\alpha) + \mathscr{R}_N(\alpha),$$

where $\mathscr{R}_N(\alpha)$ converges to zero, as a member of $C_2(\mathscr{S}^{p-1})$, in probability. Therefore, it suffices to investigate the asymptotic performances of $\tilde{Z}_N(\cdot)$ for our purpose. Note that

$$E[\tilde{Z}_{3,N}(\alpha)] = E[\tilde{Z}_{4,N}(\alpha)] = 0,$$

for $\alpha \in \mathscr{S}^{p-1}$. Further, by direct computations, we can see that the covariance functions of $\tilde{Z}_{3,N}$ and $\tilde{Z}_{4,N}$ are respectively given as

$$v_3(\alpha, \bar{\alpha}) \equiv E[\tilde{Z}_{3,N}(\alpha)\tilde{Z}_{3,N}(\bar{\alpha})]$$

$$= \frac{6m_6}{p(p+2)(p+4)}(\alpha'\bar{\alpha})^3 + 9\left[1 - \frac{2m_4}{p(p+2)} + \frac{m_6}{p(p+2)(p+4)}\right](\alpha'\bar{\alpha})$$

$$= \tau_{3,1}(\alpha'\bar{\alpha})^3 + \tau_{3,2}(\alpha'\bar{\alpha}), \tag{1.6}$$

$$v_4(\alpha, \bar{\alpha}) \equiv E[\tilde{Z}_{4,N}(\alpha)\tilde{Z}_{4,N}(\bar{\alpha})]$$

$$= \frac{24m_8}{p(p+2)(p+4)(p+6)}(\alpha'\bar{\alpha})^4$$

$$+ 72\left[\frac{m_8}{p(p+2)(p+4)(p+6)} - \frac{2m_4m_6}{p^2(p+2)^2(p+4)} + \frac{m_4^3}{p^3(p+2)^3}\right](\alpha'\bar{\alpha})^2$$

$$+ 9\left[\frac{m_8}{p(p+2)(p+4)(p+6)} - \frac{4m_4m_6}{p^2(p+2)^2(p+4)}\right.$$

$$\left. + \frac{4m_4^3}{p^3(p+2)^3} - \frac{m_4^2}{p^2(p+2)^2}\right]$$

$$= \tau_{4,1}(\alpha'\bar{\alpha})^4 + \tau_{4,2}(\alpha'\bar{\alpha})^2 + \tau_{4,3}, \tag{1.7}$$

for $\alpha, \bar{\alpha} \in \mathscr{S}^{p-1}$, where

$$\tau_{3,1} = \frac{6m_6}{p(p+2)(p+4)},$$

$$\tau_{3,2} = 9\left[1 - \frac{2m_4}{p(p+2)} + \frac{m_6}{p(p+2)(p+4)}\right],$$

$$\tau_{4,1} = \frac{24m_8}{p(p+2)(p+4)(p+6)},$$

$$\tau_{4,2} = 72 \left[\frac{m_8}{p(p+2)(p+4)(p+6)} - \frac{2m_4 m_6}{p^2(p+2)^2(p+4)} + \frac{m_4^3}{p^3(p+2)^3} \right],$$

$$\tau_{4,3} = 9 \left[\frac{m_8}{p(p+2)(p+4)(p+6)} - \frac{4m_4 m_6}{p^2(p+2)^2(p+4)} + \frac{4m_4^3}{p^3(p+2)^3} - \frac{m_4^2}{p^2(p+2)^2} \right].$$

Further

$$E[\tilde{Z}_{3,N}(\alpha)\tilde{Z}_{4,N}(\bar{\alpha})] = 0, \tag{1.8}$$

for $\alpha, \bar{\alpha} \in \mathscr{S}^{p-1}$. Lemma 1.1 also gives the following formulas for σ_3^2 and σ_4^2:

$$\sigma_3^2 = A - Var[H_3(\alpha' S^{-1/2}(X_1 - \bar{X}))] = v_3(\alpha, \alpha) = \tau_{3,1} + \tau_{3,2}, \tag{1.9}$$

$$\sigma_4^2 = A - Var[H_4(\alpha' S^{-1/2}(X_1 - \bar{X}))] = v_4(\alpha, \alpha) = \tau_{4,1} + \tau_{4,2} + \tau_{4,3}. \tag{1.10}$$

Then the covariance functions of $Z_{3,N}$ and $Z_{4,N}$ are given as

$$c_3(\alpha, \bar{\alpha}) \equiv E[Z_{3,N}(\alpha)Z_{3,N}(\bar{\alpha})] = v_3(\alpha, \bar{\alpha})/\sigma_3^2, \tag{1.11}$$

$$c_4(\alpha, \bar{\alpha}) \equiv E[Z_{4,N}(\alpha)Z_{4,N}(\bar{\alpha})] = v_4(\alpha, \bar{\alpha})/\sigma_4^2. \tag{1.12}$$

THEOREM 1.2 *Under the conditions of Lemma 1.1, there exists a zero-mean Gaussian random field* $\tilde{Z}(\alpha), \alpha \in \mathscr{S}^{p-1}$, *with continuous sample paths and covariance kernel*

$$v(\alpha, \bar{\alpha}) = \begin{bmatrix} v_3(\alpha, \bar{\alpha}) & 0 \\ 0 & v_4(\alpha, \bar{\alpha}) \end{bmatrix}$$

for $\alpha, \bar{\alpha} \in \mathscr{S}^{p-1}$ *such that*

$$\tilde{Z}_N(\cdot) \Rightarrow \tilde{Z}(\cdot).$$

Therefore

$$\tilde{W}_N(\cdot) \Rightarrow \tilde{Z}(\cdot).$$

From the equations (1.3) and (1.4), we see that

$$W_N(\cdot) = \mathscr{D}\tilde{W}_N(\cdot), \qquad Z_N(\cdot) = \mathscr{D}\tilde{Z}_N(\cdot),$$

where $\mathscr{D} = diag\{1/\sigma_3, 1/\sigma_4\}$. So we get the following result.

COROLLARY 1.3 *Under the conditions of Lemma 1.1, there exists a zero-mean Gaussian random field* $Z(\alpha) = \mathscr{D}\tilde{Z}(\alpha)$, $\alpha \in \mathscr{S}^{p-1}$, *with continuous sample paths and covariance kernel*

$$c(\alpha, \bar{\alpha}) = \begin{bmatrix} c_3(\alpha, \bar{\alpha}) & 0 \\ 0 & c_4(\alpha, \bar{\alpha}) \end{bmatrix}$$

such that

$$Z_N(\cdot) \Rightarrow Z(\cdot).$$

Therefore

$$W_N(\cdot) \Rightarrow Z(\cdot).$$

From the affine invariance property of the maxima of $\mathscr{GMI}$ and the continuous mapping theorem, we obtain the next theorem.

THEOREM 1.4   *Let X have an elliptically symmetric distribution with mean μ and covariance matrix Σ such that*

$$E[\{(X - \mu)'\Sigma^{-1}(X - \mu)\}^4] < \infty.$$

Then

$$\max_{\alpha \in \mathscr{S}^{p-1}} N\mathscr{GMI}(\alpha) \Rightarrow \max_{\alpha \in \mathscr{S}^{p-1}} |Z(\alpha)|^2$$

as $N \to \infty$, where $Z(\cdot)$ is the Gaussian random field given in Corollary 1.3, and m_{2k} included in (1.11) and (1.12) is given by

$$m_{2k} = E[\{(X - \mu)'\Sigma^{-1}(X - \mu)\}^k],$$

for $k = 2, 3, 4$.

We have established an asymptotic result for the maxima of $\mathscr{GMI}$ in the above theorem. Here, we note σ_3^2 and σ_4^2 included in $\mathscr{GMI}$. These values may be determined, as noted in the equations (1.9) and (1.10), from the formulas

$$\sigma_3^2 = \frac{15m_6}{p(p+2)(p+4)} - \frac{18m_4}{p(p+2)} + 9, \tag{1.13}$$

$$\sigma_4^2 = \frac{105m_8}{p(p+2)(p+4)(p+6)} - \frac{180m_4 m_6}{p^2(p+2)^2(p+4)}$$

$$+ \frac{108m_4^3}{p^3(p+2)^3} - \frac{9m_4^2}{p^2(p+2)^2}. \tag{1.14}$$

Practically, it may be considered to use the natural estimators $\hat{\sigma}_3^2$ and $\hat{\sigma}_4^2$ given by

$$\hat{\sigma}_3^2 = \frac{15\hat{m}_6}{p(p+2)(p+4)} - \frac{18\hat{m}_4}{p(p+2)} + 9, \tag{1.15}$$

$$\hat{\sigma}_4^2 = \frac{105\hat{m}_8}{p(p+2)(p+4)(p+6)} - \frac{180\hat{m}_4 \hat{m}_6}{p^2(p+2)^2(p+4)}$$

$$+ \frac{108\hat{m}_4^3}{p^3(p+2)^3} - \frac{9\hat{m}_4^2}{p^2(p+2)^2}, \tag{1.16}$$

instead of σ_3^2 and σ_4^2, respectively, where

$$\hat{m}_{2k} = \frac{1}{N} \sum_{j=1}^{N} \{(X_j - \bar{X})' S^{-1} (X_j - \bar{X})\}^k,$$

for $k = 2, 3, 4$. By using these, we may propose another index

$$\mathscr{GMI}^*(\alpha) = \sum_{k=3}^{4} \left[\frac{1}{N} \sum_{j=1}^{N} \left\{ \frac{H_k(\alpha' S^{-1/2}(X_j - \bar{X})) - \theta_k}{\hat{\sigma}_k} \right\} \right]^2. \qquad (1.17)$$

Since $\hat{m}_{2k}$ converges almost surely to m_{2k} for $k = 2, 3, 4$, $\hat{\sigma}_3^2$ and $\hat{\sigma}_4^2$ converge almost surely to σ_3^2 and σ_4^2, respectively. This implies that Theorem 1.4 holds also for $\mathscr{GMI}^*$, that is

$$\max_{\alpha \in \mathscr{S}^{p-1}} N \mathscr{GMI}^*(\alpha) \Rightarrow \max_{\alpha \in \mathscr{S}^{p-1}} |Z(\alpha)|^2 \qquad (1.18)$$

as $N \to \infty$.

Further, relating to the approximation formula of (1.2), we consider the following index

$$\mathscr{GMI}^\dagger(\alpha) = \sum_{k=3}^{4} \left[\frac{1}{N} \sum_{j=1}^{N} \left\{ \frac{H_k(\alpha' S^{-1/2}(X_j - \bar{X})) - \theta_k}{\tilde{\sigma}_k} \right\} \right]^2, \qquad (1.19)$$

where $\tilde{\sigma}_k^2 = \max\{\sigma_k^2, \hat{\sigma}_k^2\}$ for $k = 3, 4$. Also for $\mathscr{GMI}^\dagger$, it holds that

$$\max_{\alpha \in \mathscr{S}^{p-1}} N \mathscr{GMI}^\dagger(\alpha) \Rightarrow \max_{\alpha \in \mathscr{S}^{p-1}} |Z(\alpha)|^2 \qquad (1.20)$$

as $N \to \infty$. From (1.20) and the fact that $\tilde{\sigma}_k^2$ is alway greater than or equal to $\hat{\sigma}_k^2$, we see that the maxima of $N \mathscr{GMI}^\dagger$ has the same limiting distribution as that of $N \mathscr{GMI}^*$, however, its convergence satisfies $\mathscr{GMI}^\dagger \leq \mathscr{GMI}^*$ in any N. Monte Carlo experiments in Section 3 show that the accuracy of the approximation formula derived in Section 2 for $\mathscr{GMI}^\dagger$ is better than the ones for $\mathscr{GMI}$ and $\mathscr{GMI}^*$.

In the rest of this Section, we give a representation of the limiting Gaussian random field. Note that for a fixed $\alpha \in \mathscr{S}^{p-1}$, $Z(\alpha) \sim \mathscr{N}(0, I_2)$, the bivariate standard normal distribution. Let us define a new random field derived from $|Z(\cdot)|$ as

$$\xi(\alpha, \beta) = \beta' Z(\alpha), \qquad (1.21)$$

for $\alpha \in \mathscr{S}^{p-1}$ and $\beta = (\beta_1, \beta_2) \in \mathscr{S}^1$. It has been shown that $Z(\cdot)$ is a Gaussian random field and hence $|Z(\cdot)|^2$ is a chi-square random field. From Theorem 1.4, our problem becomes to investigate performances of the maxima of a chi-

square random field. However, this problem can be reduced to investigate that of the Gaussian random field by the device

$$|Z(\alpha)| = \max_{\beta \in \mathscr{S}^1} \beta' Z(\alpha) = \max_{\beta \in \mathscr{S}^1} \xi(\alpha, \beta).$$

An important property of the Gaussian random field $\xi(\alpha, \beta)$ derived from $|Z(\cdot)|$ is that it has a finite Karhunen-Loève expansion.

PROPOSITION 1.5 *The limiting Gaussian random field* $\xi(\alpha, \beta)$, $\alpha \in \mathscr{S}^{p-1}$, $\beta \in \mathscr{S}^1$, *can be represented in the form*

$$\xi(\alpha, \beta) = \sqrt{\eta_{3,1}} \beta_1 \sum_{k=1}^{v(3)} \zeta_{3,k}(\alpha) N_{3,k} + \sqrt{\eta_{3,2}} \beta_1 \sum_{k=1}^{v(1)} \zeta_{1,k}(\alpha) N_{1,k}$$

$$+ \sqrt{\eta_{4,1}} \beta_2 \sum_{k=1}^{v(4)} \zeta_{4,k}(\alpha) N_{4,k} + \sqrt{\eta_{4,2}} \beta_2 \sum_{k=1}^{v(2)} \zeta_{2,k}(\alpha) N_{2,k} + \sqrt{\eta_{4,3}} \beta_2 N_0,$$

(1.22)

where

$$\eta_{3,1} = \frac{36 m_6}{\sigma_3^2 \{p(p+2)(p+4)\}^2},$$

$$\eta_{3,2} = \frac{9}{\sigma_3^2 p} \left[1 - \frac{2 m_4}{p(p+2)} + \frac{m_6}{p(p+2)^2} \right],$$

$$\eta_{4,1} = \frac{m_8}{\sigma_4^2} \left[\frac{24}{p(p+2)(p+4)(p+6)} \right]^2,$$

$$\eta_{4,2} = \frac{144}{\sigma_4^2} \left[\frac{m_8}{\{p(p+2)(p+4)\}^2} - \frac{2 m_6 m_4}{p^3 (p+2)^3 (p+4)} + \frac{m_4^3}{\{p(p+2)\}^4} \right],$$

$$\eta_{4,3} = \frac{9}{\sigma_4^2} \left[\frac{m_8}{\{p(p+2)\}^2} - \frac{4 m_4 m_6}{p^3 (p+2)^2} + \frac{4 m_4^3}{p^4 (p+2)^2} - \left\{ \frac{m_4^2}{p(p+2)} \right\}^2 \right],$$

N_0; $\{N_{l,k} : k = 1, \ldots, v(l), l = 1, 2, 3, 4\}$ *are independent standard normal random variables and* $\{\zeta_{l,k} : k = 1, \ldots, v(l), l = 1, 2, 3, 4\}$ *are linearly independent surface harmonics of degree* $l(l = 1, 2, 3, 4)$ *respectively, being orthonormal with respect to the uniform distribution on* $\mathscr{S}^{p-1}$, *where*

$$v(1) = p, \qquad v(2) = \frac{(p-1)(p+2)}{2}, \qquad v(3) = \frac{p(p-1)(p+4)}{6},$$

and

$$v(4) = \frac{p(p-1)(p+1)(p+6)}{24}.$$

2   Approximation formula

The purpose of this section is to derive an approximation formula of (1.2) under elliptically symmetric distributions. The fact that the limiting random field derived from $\mathscr{GMI}$ is Gaussian with mean zero and unit variance as shown in Theorem 1.4 and Proposition 1.5 motivates us to utilize the following theorem established by Sun (1993).

THEOREM 2.1 (SUN (1993))   *Suppose $Z(t)$ is a d-dimensional nonsingular differentiable Gaussian random field on a bounded d-dimensional set I, with mean zero, unit variance and covariance function $c(s, t)$. Under some regularity conditions for $c(s, t)$, as $z \to \infty$,*

$$P\left\{\max_{t \in I} Z(t) \geq z\right\} = \kappa_0 \psi\left(\frac{z^2}{2}, \frac{d+1}{2}\right) + \kappa_2 \psi\left(\frac{z^2}{2}, \frac{d-1}{2}\right)\{1 + o(1)\},$$

where

$$\psi(x, d) = \int_x^\infty y^{d-1} \exp(-y) dy$$

is an incomplete Gamma function and κ_0, κ_2 are two geometric constants which can be represented as

$$\kappa_0 = \frac{1}{2\pi^{(d+1)/2}} \int_I \|V(t)\|^{1/2} dt_1 \ldots dt_d, \tag{2.1}$$

$$\kappa_2 = \frac{1}{4\pi^{(d+1)/2}} \int_I \frac{1}{2} \{-S(t) - d(d-1)\} \|V(t)\|^{1/2} dt_1 \ldots dt_d. \tag{2.2}$$

$\|V(t)\|$ is the determinant of the $d \times d$ matrix

$$V(t) = \left\{\left.\frac{\partial^2 c(s, t)}{\partial s_i \partial t_j}\right|_{s=t}\right\}$$

and $S(t)$ is the scalar curvature of the manifold which has $V(t)$ as its metric tensor.

Let

$$\mathscr{S}^{p-1}/2 = \{\alpha = (\alpha_1, \ldots, \alpha_p)' \in \mathscr{S}^{p-1} : \alpha_p \geq 0\}.$$

Since $\mathscr{GMI}(\alpha) = \mathscr{GMI}(-\alpha)$, its maxima is unchanged if α is restricted to a semisphere $\mathscr{S}^{p-1}/2$. Then, by the results obtained in Section 1, we get the weak convergence property

$$\max_{\alpha \in \mathscr{S}^{p-1}/2} \sqrt{N\mathscr{GMI}(\alpha)} \Rightarrow \max_{\alpha \in \mathscr{S}^{p-1}/2} |Z(\alpha)|.$$

At first we evaluate the probability (1.2) as follows:

$$P\left\{\max_{\alpha \in \mathscr{S}^{p-1}} \mathscr{G}\mathscr{M}\mathscr{I}(\alpha) \geq z\right\} = P\left\{\max_{\alpha \in \mathscr{S}^{p-1}/2} \mathscr{G}\mathscr{M}\mathscr{I}(\alpha) \geq z\right\}$$

$$= P\left\{\max_{\alpha \in \mathscr{S}^{p-1}/2} \sqrt{N\mathscr{G}\mathscr{M}\mathscr{I}(\alpha)} \geq x\right\}$$

$$\approx P\left\{\max_{\alpha \in \mathscr{S}^{p-1}/2} |Z(\alpha)| \geq x\right\}$$

$$= P\left\{\max_{\alpha \in \mathscr{S}^{p-1}/2, \beta \in \mathscr{S}^1} \xi(\alpha, \beta) \geq x\right\}$$

$$= P\left\{\max_{t \in I} \xi(t) \geq x\right\},$$

where $\xi(\alpha, \beta)$ is defined in (1.21) and $x = \sqrt{Nz}$. The parameter space of the Gaussian random field $\xi(\alpha, \beta)$ is p-dimensional. It is possible to reparametrize $\{(\alpha, \beta)\}$ in terms of t by the following spherical polar coordinate transformation

$$\alpha = \alpha(\phi) = (\alpha_1(\phi), \ldots, \alpha_p(\phi))',$$

$$\phi = (\phi_1, \ldots, \phi_{p-1}) \in [0, \pi] \times \cdots \times [0, \pi] \equiv I_\phi \subset R^{p-1},$$

$$\beta = (\beta_1(\varphi), \beta_2(\varphi)),$$

$$\varphi \in [0, 2\pi] \equiv I_\varphi,$$

$$t = (\phi, \varphi) = (\phi_1, \ldots, \phi_{p-1}, \varphi) \in I_\phi \times I_\varphi \equiv I,$$

where $\varphi = 0$ and $\varphi = 2\pi$ represent the same point and

$$\xi(t) = \beta'(\varphi)Z(\alpha(\phi)),$$

for $t = (\phi_1, \ldots, \phi_{p-1}, \varphi) \in I$. Note that $\xi(t)$ is a nonsingular differentiable Gaussian random field with $E[\xi(t)] = 0$ and

$$r(s, t) \equiv E[\xi(s)\xi(t)]$$

$$= \beta'(\varphi)E[Z(\alpha(\phi))Z'(\alpha(\bar{\phi}))]\beta(\bar{\varphi})$$

$$= c_3(\alpha(\phi), \alpha(\bar{\phi}))\beta_1(\varphi)\beta_1(\bar{\varphi}) + c_4(\alpha(\phi), \alpha(\bar{\phi}))\beta_2(\varphi)\beta_2(\bar{\varphi}),$$

$$r(t, t) = 1,$$

for $t = (\phi, \varphi), s = (\bar{\phi}, \bar{\varphi}) \in I$. We can easily check the regularity conditions about the covariance function $r(s, t)$ for using Theorem 2.1. Especially, the critical radius of the tube of the manifold constructed from the Karhunen-

Loève expansion of $\xi(t)$ is positive since, as shown in Proposition 1.5, its expansion is finite (see Sun (1993)). Thus we can apply Theorem 2.1 to the maxima of $\xi(t), t \in I$.

For using Theorem 2.1, it is necessary to get the metric tensor matrix and two geometric constants concretely. In the following, we only summarize these quantities (detailed calculations are given in Section 4). Now, by a direct calculation, *the metric tensor matrix* is obtained as follows.

$$R(t) \equiv \left\{ \frac{\partial^2 r(s,t)}{\partial t_k \partial s_l} \bigg|_{t=s} \right\} = \left\{ \begin{array}{cc} R_{11}(t) & 0 \\ 0 & R_{22}(t) \end{array} \right\}, \qquad (2.3)$$

where

$$R_{11}(t) = C(\beta) diag \left\{ \sum_{l=1}^{p} \left(\frac{\partial \alpha_l}{\partial \phi_1} \right)^2, \ldots, \sum_{l=1}^{p} \left(\frac{\partial \alpha_l}{\partial \phi_{p-1}} \right)^2 \right\},$$

$$R_{22}(t) = \sum_{l=1}^{2} \left(\frac{\partial \beta_l}{\partial \varphi} \right)^2,$$

and

$$C(\beta) = \left\{ \frac{3\tau_{3,1} + \tau_{3,2}}{\sigma_3^2} \right\} \beta_1^2 + \left\{ \frac{4\tau_{4,1} + 2\tau_{4,2}}{\sigma_4^2} \right\} \beta_2^2.$$

Straightforward but somewhat lengthy computations yield that *the scalar curvature* is

$$\begin{aligned} S(t) = {} & - (p-1)(p-2)C(\beta)^{-1} \\ & + (p-1)C(\beta)^{-1} \left\{ \sum_{l=1}^{2} \left(\frac{\partial \beta_l}{\partial \varphi} \right)^2 \right\}^{-1} \frac{\partial^2 C(\beta)}{\partial \varphi^2} \qquad (2.4) \\ & + (p-1)(p/4-1)C(\beta)^{-2} \left\{ \sum_{l=1}^{2} \left(\frac{\partial \beta_l}{\partial \varphi} \right)^2 \right\}^{-1} \left\{ \frac{\partial C(\beta)}{\partial \varphi} \right\}^2. \end{aligned}$$

Therefore two geometric constants are given as

$$\begin{aligned} \kappa_0 &= \frac{1}{2\pi^{(p+1)/2}} \int_I \|R(t)\|^{1/2} d\phi_1 \cdots d\phi_{p-1} d\varphi \\ &= \frac{1}{2\pi^{(p+1)/2}} \int_I \{C(\beta)\}^{(p-1)/2} \left[\prod_{l=1}^{p-1} \left\{ \sum_{j=1}^{p} \left(\frac{\partial \alpha_j}{\partial \phi_l} \right)^2 \right\} \right]^{1/2} \\ &\quad \times \left\{ \sum_{l=1}^{2} \left(\frac{\partial \beta_l}{\partial \varphi} \right)^2 \right\}^{1/2} d\phi_1 \cdots d\phi_{p-1} d\varphi \qquad (2.5) \end{aligned}$$

$$= \frac{\omega_{p-1}}{4\pi^{(p+1)/2}} \int_{I_\varphi} \{C(\beta)\}^{(p-1)/2} \left\{ \sum_{l=1}^{2} \left(\frac{\partial \beta_l}{\partial \varphi} \right)^2 \right\}^{1/2} d\varphi$$

$$= \frac{\omega_{p-1}\omega_1}{4\pi^{(p+1)/2}} E_\beta[\{C(\beta)\}^{(p-1)/2}],$$

$$\kappa_2 = \frac{1}{4\pi^{(d+1)/2}} \int_I \frac{1}{2} \{-S(t) - p(p-1)\} \|R(t)\|^{1/2} d\phi_1 \cdots d\phi_{p-1} d\varphi$$

$$= \frac{\omega_{p-1}\omega_1}{8\pi^{(p+1)/2}} E_\beta[D\{p, C(\beta)\}],$$

(2.6)

where

$$D\{p, C(\beta)\} = -\frac{1}{2}p(p-1)\{C(\beta)\}^{(p-1)/2} + \frac{1}{2}(p-1)(p-2)\{C(\beta)\}^{(p-3)/2}$$

$$- \frac{1}{2}(p-1)\{C(\beta)\}^{(p-3)/2} \left\{ \sum_{l=1}^{2} \left(\frac{\partial \beta_l}{\partial \varphi} \right)^2 \right\}^{-1} \frac{\partial^2 C(\beta)}{\partial \varphi^2}$$

$$- \frac{1}{2}(p-1)(p/4-1)\{C(\beta)\}^{(p-5)/2} \left\{ \sum_{l=1}^{2} \left(\frac{\partial \beta_l}{\partial \varphi} \right)^2 \right\}^{-1} \left\{ \frac{\partial C(\beta)}{\partial \varphi} \right\}^2,$$

and ω_{p-1} is the surface area of $\mathscr{S}^{p-1}$ given by $\omega_{p-1} = 2\pi^{p/2}/\Gamma(p/2)$ and the expectation $E_\beta[\cdot]$ is taken for $\beta \sim Uniform(\mathscr{S}^1)$, the uniform distribution on $\mathscr{S}^1$.

Since the approximation mentioned above is based on the limiting random field, we see from (1.18) and (1.20) that similar argument also holds for $\mathscr{GMI}^*$ and $\mathscr{GMI}^\dagger$ defined by (1.17) and (1.19), respectively. We summarize the results as follows.

PROPOSITION 2.2 *Under the conditions of Theorem 1.4,*

$$P\left\{ \max_{\alpha \in \mathscr{S}^{p-1}} \mathscr{GMI}(\alpha) \geq z \right\} \approx \kappa_0 \psi\left(\frac{Nz}{2}, \frac{p+1}{2} \right) + \kappa_2 \psi\left(\frac{Nz}{2}, \frac{p-1}{2} \right), \quad (2.7)$$

$$P\left\{ \max_{\alpha \in \mathscr{S}^{p-1}} \mathscr{GMI}^*(\alpha) \geq z \right\} \approx \kappa_0 \psi\left(\frac{Nz}{2}, \frac{p+1}{2} \right) + \kappa_2 \psi\left(\frac{Nz}{2}, \frac{p-1}{2} \right), \quad (2.8)$$

$$P\left\{ \max_{\alpha \in \mathscr{S}^{p-1}} \mathscr{GMI}^\dagger(\alpha) \geq z \right\} \approx \kappa_0 \psi\left(\frac{Nz}{2}, \frac{p+1}{2} \right) + \kappa_2 \psi\left(\frac{Nz}{2}, \frac{p-1}{2} \right) \quad (2.9)$$

as $N \to \infty$.

3 Examples and Monte Carlo experiments

In this section we confirm the results in Section 2 through some elliptically symmetric distributions. Exploratory projection pursuit procedure based on

$\mathscr{GMI}$ is applied to real data in conjunction with showing efficiency of pursuing *nonellipticity*. Further, some Monte Calro experiments related to the approximation formulas in Proposition 2.2 are discussed.

3.1 Some examples

EXAMPLE 3.1 (Normal distribution). Projection pursuit discussed up to this point regards normal distribution as the least interesting structure. When $X \sim \mathscr{N}(\mu, \Sigma)$, a p-variate normal distribution with mean μ and covariance matrix Σ, since $(X - \mu)' \Sigma^{-1} (X - \mu)$ distributed as a chi-square distribution with p degree of freedom, we have

$$m_4 = E[\{(X - \mu)' \Sigma^{-1} (X - \mu)\}^2] = p(p + 2),$$

$$m_6 = E[\{(X - \mu)' \Sigma^{-1} (X - \mu)\}^3] = p(p + 2)(p + 4),$$

$$m_8 = E[\{(X - \mu)' \Sigma^{-1} (X - \mu)\}^4] = p(p + 2)(p + 4)(p + 6).$$

Therefore

$$\theta_4 = 0, \qquad \sigma_3^2 = 6, \qquad \sigma_4^2 = 24,$$

which have noted in Section 0. Further, the facts that

$$\tau_{3,1} = 6, \qquad \tau_{4,1} = 24, \qquad \tau_{3,2} = \tau_{4,2} = \tau_{4,3} = 0$$

give

$$C(\beta) = 3\beta_1^2 + 4\beta_2^2 = 4 - \beta_1^2, \tag{3.1}$$

since $\beta_1^2 + \beta_2^2 = 1$. It is well known fact that, for $\beta \sim Uniform(\mathscr{S}^1)$, β_1^2 in (3.1) is distributed as a Beta distribution $Beta(1/2, 1/2)$. Therefore we can obtain the values of $E_\beta[\{C(\beta)\}^{(p-1)/2}]$ by using the central moments of $Beta(1/2, 1/2)$ if p is odd. For even p, we can get these by numerical integral. The expressions of $E_\beta[\cdot]$ in (2.5) and (2.6) are useful because of, especially for large p, these can also be obtained by Monte Carlo. Table 3.1 tabulates the geometric constants κ_0 in (2.5) and κ_2 in (2.6) for several dimensions p.

EXAMPLE 3.2 (Contaminated normal distribution). If the random vector X has the density

$$f(x) = \frac{(1 - \varepsilon)}{(2\pi)^{p/2} |\Delta|^{1/2}} \exp\left[-\frac{1}{2}(x - \mu)' \Delta^{-1}(x - \mu)\right]$$

$$+ \frac{\varepsilon}{(2\pi)^{p/2} |c\Delta|^{1/2}} \exp\left[-\frac{1}{2c}(x - \mu)' \Delta^{-1}(x - \mu)\right]$$

for some vector $\mu \in R^p$, some symmetric positive definite matrix $\varDelta$ and some $c > 0$ we say that X has a contaminated normal distribution and denote it $X \sim CN_p(c, \mu, \varDelta)(0 < \varepsilon < 1$, see Anderson (1993, p.9)). We have

$$E[X] = \mu, \qquad E[(X - \mu)(X - \mu)'] = \{1 + \varepsilon(c - 1)\}\varDelta = \Sigma.$$

For $X \sim CN_p(c, \mu, \varDelta)$, since $(X - \mu)'\varDelta^{-1}(X - \mu)$ has the density

$$\frac{(1 - \varepsilon)}{2^{p/2}\Gamma(p/2)} \, t^{p/2-1}\exp(-t/2) + \frac{\varepsilon}{(2c)^{p/2}\Gamma(p/2)} \, t^{p/2-1}\exp(-t/(2c)),$$

we obtain

$$m_4 = p(p + 2)\frac{\{1 + \varepsilon(c^2 - 1)\}}{\{1 + \varepsilon(c - 1)\}^2},$$

$$m_6 = p(p + 2)(p + 4)\frac{\{1 + \varepsilon(c^3 - 1)\}}{\{1 + \varepsilon(c - 1)\}^3},$$

$$m_8 = p(p + 2)(p + 4)(p + 6)\frac{\{1 + \varepsilon(c^4 - 1)\}}{\{1 + \varepsilon(c - 1)\}^4}.$$

Therefore

$$\theta_4 = 3\left[\frac{\{1 + \varepsilon(c^2 - 1)\}}{\{1 + \varepsilon(c - 1)\}^2} - 1\right],$$

$$3\tau_{3,1} + \tau_{3,2} = 27\frac{\{1 + \varepsilon(c^3 - 1)\}}{\{1 + \varepsilon(c - 1)\}^3} - 18\frac{\{1 + \varepsilon(c^2 - 1)\}}{\{1 + \varepsilon(c - 1)\}^2} + 9,$$

$$\sigma_3^2 = 15\frac{\{1 + \varepsilon(c^3 - 1)\}}{\{1 + \varepsilon(c - 1)\}^3} - 18\frac{\{1 + \varepsilon(c^2 - 1)\}}{\{1 + \varepsilon(c - 1)\}^2} + 9,$$

$$4\tau_{4,1} + 2\tau_{4,2} = 240\frac{\{1 + \varepsilon(c^4 - 1)\}}{\{1 + \varepsilon(c - 1)\}^4} - 288\frac{\{1 + \varepsilon(c^2 - 1)\}\{1 + \varepsilon(c^3 - 1)\}}{\{1 + \varepsilon(c - 1)\}^5}$$

$$+ 144\frac{\{1 + \varepsilon(c^2 - 1)\}^3}{\{1 + \varepsilon(c - 1)\}^6},$$

$$\sigma_4^2 = 105\frac{\{1 + \varepsilon(c^4 - 1)\}}{\{1 + \varepsilon(c - 1)\}^4} - 180\frac{\{1 + \varepsilon(c^2 - 1)\}\{1 + \varepsilon(c^3 - 1)\}}{\{1 + \varepsilon(c - 1)\}^5}$$

$$+ 108\frac{\{1 + \varepsilon(c^2 - 1)\}^3}{\{1 + \varepsilon(c - 1)\}^6} - 9\frac{\{1 + \varepsilon(c^2 - 1)\}^2}{\{1 + \varepsilon(c - 1)\}^4}.$$

TABLE 3.1: Geometric Constants κ_0 and κ_2 (Normal Case)		
p	κ_0	κ_2
3	7.0	−9.4641016
4	11.650304	−29.805364
5	16.5	−68.0
6	20.698704	−125.07895
7	23.566667	−196.2
8	24.754117	−271.56853
9	24.270238	−339.31429
10	22.407414	−388.92865

TABLE 3.2: Geometric Constants κ_0 and κ_2 (Contaminated Normal Case)		
p	κ_0	κ_2
3	6.1537202	−8.177
4	9.6281624	−23.982534
5	12.840656	−51.242143
6	15.192911	−88.607571
7	16.339914	−131.0358
8	16.235653	−171.37926
9	15.077835	−202.71231
10	13.201537	−220.30817

Similar to Table 3.1, Table 3.2 tabulates the geometric constants κ_0 and κ_2 for several dimensions p under the case where $\varepsilon = 0.5$ and $c = 9.0$.

EXAMPLE 3.3 (Symmetric multivariate Pearson Type II distribution).   The random vector X is said to have a symmetric multivariate Pearson Type II distribution if X has the density

$$f(x) = \frac{\Gamma(p/2+m+1)}{\Gamma(m+1)\pi^{p/2}|\varDelta|^{1/2}} \{1-(x-\mu)'\varDelta^{-1}(x-\mu)\}^m \times I\{(x-\mu)'\varDelta^{-1}(x-\mu) \le 1\}$$

for some vector $\mu \in R^p$ and some symmetric positive definite matrix $\varDelta(m \in R, m > -1$ see Fang et $al.$ (1989, p. 89)). We shall denote this $X \sim MPII_p(m, \mu, \varDelta)$.

As shown in Baringhaus and Henze (1991, Examples 2.8, 3.7),

$$E[X] = \mu, \qquad E[(X - \mu)(X - \mu)'] = \frac{1}{p + 2m + 2} \varDelta = \varSigma,$$

and $(X - \mu)'\varDelta(X - \mu)$ has a Beta distribution $B(p/2, m + 1)$.   Further, we see for $X \sim MPII_p(m, \mu, \varDelta)$ that

$$m_4 = p(p + 2)\frac{p + 2m + 2}{p + 2m + 4},$$

$$m_6 = p(p + 2)(p + 4)\frac{(p + 2m + 2)^2}{(p + 2m + 4)(p + 2m + 6)},$$

$$m_8 = \frac{p(p + 2)(p + 4)(p + 6)(p + 2m + 2)^3}{(p + 2m + 4)(p + 2m + 6)(p + 2m + 8)}.$$

Thus direct calculations give

$$\theta_4 = 3\left[\frac{p+2m+2}{p+2m+4} - 1\right],$$

$$3\tau_{3,1} + \tau_{3,2} = \frac{27(p+2m+2)^2}{(p+2m+4)(p+2m+6)} - 18\left\{\frac{p+2m+2}{p+2m+4}\right\} + 9,$$

$$\sigma_3^2 = \frac{15(p+2m+2)^2}{(p+2m+4)(p+2m+6)} - 18\left\{\frac{p+2m+2}{p+2m+4}\right\} + 9,$$

$$4\tau_{4,1} + 2\tau_{4,2} = \frac{240(p+2m+2)^3}{(p+2m+4)(p+2m+6)(p+2m+8)}$$
$$- \frac{288(p+2m+2)^3}{(p+2m+4)^2(p+2m+6)} + 144\left\{\frac{p+2m+2}{p+2m+4}\right\}^3,$$

$$\sigma_4^2 = \frac{105(p+2m+2)^3}{(p+2m+4)(p+2m+6)(p+2m+8)}$$
$$- \frac{180(p+2m+2)^3}{(p+2m+4)^2(p+2m+6)}$$
$$+ 108\left\{\frac{p+2m+2}{p+2m+4}\right\}^3 - 9\left\{\frac{p+2m+2}{p+2m+4}\right\}^2.$$

EXAMPLE 3.4 (Symmetric multivariate Pearson Type VII distribution). If the random vector X has the density

$$f(x) = \frac{\Gamma(a)}{\Gamma(a-p/2)(\pi m)^{p/2}|\Delta|^{1/2}}\left\{1 + \frac{1}{m}(x-\mu)'\Delta^{-1}(x-\mu)\right\}^{-a}$$

for some vector $\mu \in R^p$ and some symmetric positive definite matrix Δ, we say that X has a symmetric multivariate Pearson Type VII distribution and we shall denote it $X \sim MPVII_p(m, \mu, \Delta)$ $(a > p/2, m > 0$, see Fang *et al.* (1989, p. 81)). In the case m is a positive integer and $a = (p+m)/2$, it is multivariate t-distribution. We have

$$E[X] = \mu, \qquad E[(X-\mu)(X-\mu)'] = \frac{m}{2a-p-2}\Delta = \Sigma(a > p/2+1).$$

Since $(X-\mu)'\Delta^{-1}(X-\mu)$ has the density

$$\frac{1}{B(p/2, a-p/2)}m^{-p/2}t^{p/2-1}\left(1 + \frac{t}{m}\right)^{-a}, \qquad t > 0,$$

straightforward calculations give

$$m_4 = p(p+2)\frac{2a-p-2}{2a-p-4},$$

$$m_6 = p(p+2)(p+4)\frac{(2a-p-2)^2}{(2a-p-4)(2a-p-6)},$$

$$m_8 = \frac{p(p+2)(p+4)(p+6)(2a-p-2)^3}{(2a-p-4)(2a-p-6)(2a-p-8)},$$

for $a > p/2 + 5$. So we get

$$\theta_4 = 3\left[\frac{2a-p-2}{2a-p-4} - 1\right],$$

$$3\tau_{3,1} + \tau_{3,2} = \frac{27(2a-p-2)^2}{(2a-p-4)(2a-p-6)} - 18\left\{\frac{2a-p-2}{2a-p-4}\right\} + 9,$$

$$\sigma_3^2 = \frac{15(2a-p-2)^2}{(2a-p-4)(2a-p-6)} - 18\left\{\frac{2a-p-2}{2a-p-4}\right\} + 9,$$

$$4\tau_{4,1} + 2\tau_{4,2} = \frac{240(2a-p-2)^3}{(2a-p-4)(2a-p-6)(2a-p-8)} - \frac{288(2a-p-2)^3}{(2a-p-4)^2(2a-p-6)}$$
$$+ 144\left\{\frac{2a-p-2}{2a-p-4}\right\}^3,$$

$$\sigma_4^2 = \frac{105(2a-p-2)^3}{(2a-p-4)(2a-p-6)(2a-p-8)} - \frac{180(2a-p-2)^3}{(2a-p-4)^2(2a-p-6)}$$
$$+ 108\left\{\frac{2a-p-2}{2a-p-4}\right\}^3 - 9\left\{\frac{2a-p-2}{2a-p-4}\right\}^2.$$

3.2 Lubischew's fleabeetle data

Next we apply exploratory projection pursuit procedure based on $\mathcal{GMI}$ to real data. In order to compare $\mathcal{GMI}$ with projection indices discussed in Jones and Sibson (1987), we adopt Lubischew's fleabeetle data (Lubischew (1962)). In Jones and Sibson (1987), comparisons of principal component analysis with one-dimensional projection pursuit based on entropy index were given by using this data set. For this data, $N = 74$ and $p = 6$. Here, we compare $\mathcal{GMI}$ under the case that normal is the least interesting, which is equivalent to $\mathcal{MI}$, with $\mathcal{GMI}$ under the case that contaminated normal (CNormal) with $\varepsilon = 0.5$, $c = 9.0$ is the least interesting. Empirical density plot of the least normal view and the least contaminated normal view are
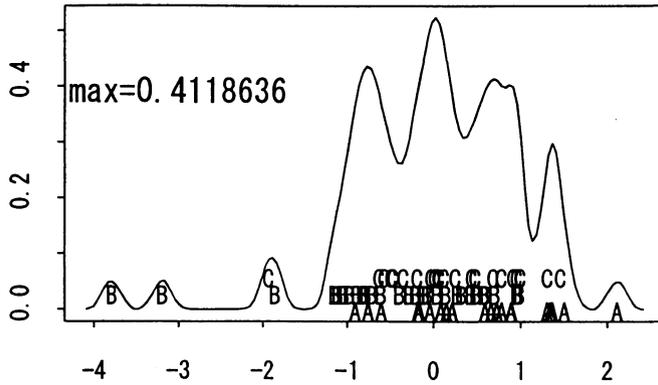
FIGURE 3.1 (a):   The least normal view obtained by $\mathcal{MI}$
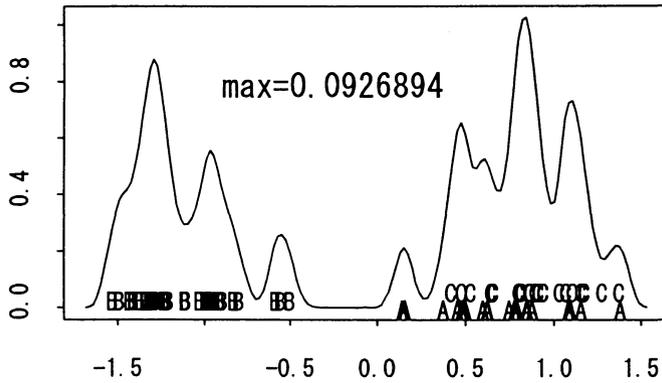
FIGURE 3.1 (b):   The least contaminated normal view obtained by $\mathcal{GMI}$

respectively shown in Figures 3.1 (a) and 3.1 (b).   In Figures 3.1 (a) and (b), utilized indices are respectively (1.1) with $\theta_4 = 0$, $\sigma_3^2 = 3!$ and $\sigma_4^2 = 4!$ and (1.1) with $\theta_4 = 1.92$, $\sigma_3^2 = 23.28$ and $\sigma_4^2 = 141.3996$.   These latter values can be obtained from Example 3.2.   Let α_{normal} and $\alpha_{cnormal}$ be the maximizers of $\mathcal{MI}$ and $\mathcal{GMI}$, respectively.   In fact, we get

$$\alpha_{normal} = (0.0932751, 0.1672886, -0.3639479, 0.0253959, -0.2715826, 0.8697437)',$$

$$\alpha_{cnormal} = (-0.6229537, 0.2468078, 0.0421696, 0.5716667, 0.0110745, 0.4714985)'.$$

The horizontal axes of Figures 3.1 (a) and (b) are α_{normal} and $\alpha_{cnormal}$, respectively.   In both Figures, "max" designates the values of the index for α_{normal} and $\alpha_{cnormal}$.   The labels A, B and C are corresponding to three different species (see, Tables 4, 5 and 6 in Lubischew (1962), respectively).   Figure 3.1
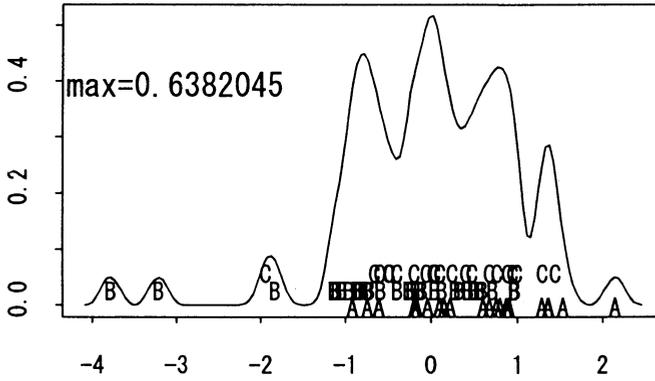
FIGURE 3.2 (a):   The least normal view obtained by $\mathcal{MI}^*$
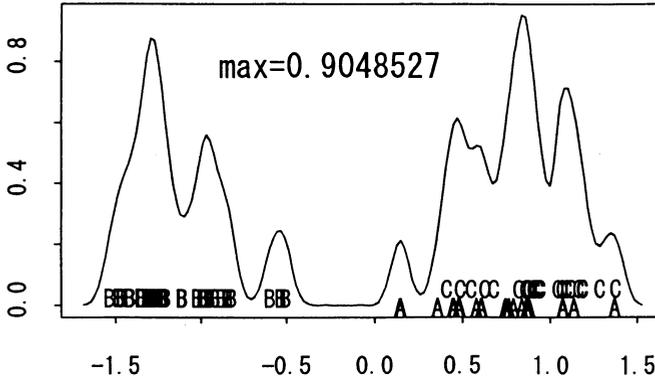
FIGURE 3.2 (b):   The least contaminated normal view obtained by $\mathcal{GMI}^*$

(b) is splitting species B from other two better than Figure 3.1 (a). This reveals not only exploratory projection pursuit based on the formalization that normal is the least interesting is not suffices to provide information of high-dimensional structure but also the fomalization that elliptical structure is the least interesting is efficient.

Similarly, Figures 3.2 (a) and (b) show empirical density plot based on $*$ version index $\mathcal{GMI}^*$, see (1.17). For this data, we get $\hat{\sigma}_3^2 = 4.409578$, $\hat{\sigma}_4^2 = 14.26718$ from (1.15) and (1.16), so that $\dagger$ version $\mathcal{GMI}^\dagger$ is equal to $\mathcal{GMI}$. In this case,

$$\alpha^*_{normal} = (0.1001596, 0.1681235, -0.3655418, 0.0345533, -0.2804055, 0.8650205)',$$

$$\alpha^*_{cnormal} = (-0.6279178, 0.2476047, 0.0545414, 0.5676038, 0.0110142, 0.4681250)',$$

where α^*_{normal} and $\alpha^*_{cnormal}$ are the maximizers of $\mathcal{MI}^*$ and $\mathcal{GMI}^*$, respectively. Also from Figures 3.2 (a) and (b), we observe that pursuing non-ellipticity is useful.

3.3 Monte Carlo experiments

In this subsection, we investigate the accuracy of the approximation formulas (2.7), (2.8) and (2.9) given in Proposition 2.2 through Monte Carlo experiments. Let

$$F(z) = P\left\{ \max_{\alpha \in \mathcal{S}^{p-1}} \mathcal{GMI}(\alpha) \geq z \right\},$$

$$F^*(z) = P\left\{ \max_{\alpha \in \mathcal{S}^{p-1}} \mathcal{GMI}^*(\alpha) \geq z \right\},$$

$$F^\dagger(z) = P\left\{ \max_{\alpha \in \mathcal{S}^{p-1}} \mathcal{GMI}^\dagger(\alpha) \geq z \right\}$$

and let

$$A(z) = \kappa_0 \psi\left(\frac{Nz}{2}, \frac{p+1}{2}\right) + \kappa_2 \psi\left(\frac{Nz}{2}, \frac{p-1}{2}\right)$$

be the value of the approximation formula in Proposition 2.2 for z. To check the accuracy of $A(z)$, we can only compare $A(z)$ with some simulated values $\hat{F}_M(z)$ of $F(z)$ for sample size M. Similarly, we compare $A(z)$ with some simulated values $\hat{F}^*_M(z)$ of $F^*(z)$ and $\hat{F}^\dagger_M(z)$ of $F^\dagger(z)$ for the same size M.

Suppose that $X_1^{(i)}, \ldots, X_N^{(i)}$ is the i-th observed data set with sample size N drawn from some specified elliptically symmetric distribution with mean μ and nonsingular covariance matrix Σ. Let

$$z_i = \max_{\alpha \in \mathcal{S}^{p-1}} \mathcal{GMI}(\alpha | X_1^{(i)}, \ldots, X_N^{(i)})$$

be the observed maxima of $\mathcal{GMI}$ based on the i-th data set, for $i = 1, \ldots, M$. The estimate $\hat{F}_M(z)$ for $F(z)$ is defined as

$$\hat{F}_M(z) = \frac{1}{M} \sum_{i=1}^{M} I(z_i \geq z).$$

Since our main interest is the significance level, we forcus on the accuracy under the case where z is in the tail of the distribution. We shall compare $\hat{F}_M(z)$ with $A(z)$ through the closeness of a and $A(\hat{a}_M)$, where $\hat{a}_M = \hat{F}_M^{-1}(a)$ is the quantile based on the estimate $\hat{F}_M(\cdot)$. The closeness of a and $A(\hat{a}_M)$

Kanta NAITO

TABLE 3.3: The accuracy of the approximation formulas (Normal Case)

		Normal		
		$p = 3$	$N = 200$	
a	SE	$A(\hat{a}_M^\dagger)$	$A(\hat{a}_M^\star)$	$A(\hat{a}_M)$
0.15	0.01129	0.1558130	0.0850384	0.0573070
0.1	0.00948	0.0753916	0.0459614	0.0174995
0.09	0.00904	0.0675008	0.0406268	0.0091794
0.08	0.00857	0.0553127	0.0350564	0.0066404
0.07	0.00806	0.0431159	0.0275108	0.0041428
0.06	0.00750	0.0365530	0.0191760	0.0019088
0.05	0.00689	0.0226535	0.0139100	0.0008007
0.04	0.00619	0.0146721	0.0104349	0.0002641
0.03	0.00539	0.0072111	0.0061774	0.0000899
0.02	0.00442	0.0042940	0.0037232	0.0000266
0.01	0.00314	0.0020497	0.0011117	0.0000013
		$p = 3$	$N = 500$	
a	SE	$A(\hat{a}_M^\dagger)$	$A(\hat{a}_M^\star)$	$A(\hat{a}_M)$
0.15	0.01129	0.1745950	0.1127960	0.0934352
0.1	0.00948	0.1088540	0.0690244	0.0357551
0.09	0.00904	0.0886893	0.0591318	0.0260286
0.08	0.00857	0.0809810	0.0511417	0.0220768
0.07	0.00806	0.0597809	0.0438919	0.0161603
0.06	0.00750	0.0511417	0.0331433	0.0115011
0.05	0.00689	0.0340495	0.0262270	0.0060777
0.04	0.00619	0.0262558	0.0207651	0.0032293
0.03	0.00539	0.0174575	0.0137071	0.0018384
0.02	0.00442	0.0096322	0.0075218	0.0005486
0.01	0.00314	0.0048497	0.0035916	0.0000329
		$p = 3$	$N = 1000$	
a	SE	$A(\hat{a}_M^\dagger)$	$A(\hat{a}_M^\star)$	$A(\hat{a}_M)$
0.15	0.01129	0.2229230	0.1516680	0.1444670
0.1	0.00948	0.1476660	0.0955310	0.0715722
0.09	0.00904	0.1299380	0.0844527	0.0603056
0.08	0.00857	0.1188060	0.0740134	0.0486144
0.07	0.00806	0.0922673	0.0599869	0.0302443
0.06	0.00750	0.0684738	0.0474915	0.0259835
0.05	0.00689	0.0500739	0.0414031	0.0219367
0.04	0.00619	0.0414031	0.0307235	0.0157401
0.03	0.00539	0.0257674	0.0205957	0.0086747
0.02	0.00442	0.0181965	0.0155692	0.0020360
0.01	0.00314	0.0106108	0.0074991	0.0002614

TABLE 3.4: The accuracy of the approximation formulas (Contaminated Normal Case)

	Contami.	Normal	$\varepsilon = 0.5$	$c = 9.0$
		$p = 3$	$N = 200$	
a	SE	$A(\hat{a}_M^\dagger)$	$A(\hat{a}_M^\star)$	$A(\hat{a}_M)$
0.15	0.01129	0.1718770	0.0777164	0.0769616
0.1	0.00948	0.1018090	0.0507668	0.0306819
0.09	0.00904	0.0867845	0.0456378	0.0236476
0.08	0.00857	0.0726342	0.0372194	0.0181558
0.07	0.00806	0.0630075	0.0298417	0.0123612
0.06	0.00750	0.0483675	0.0259224	0.0092919
0.05	0.00689	0.0373056	0.0164407	0.0060513
0.04	0.00619	0.0285515	0.0120531	0.0021956
0.03	0.00539	0.0181802	0.0089631	0.0008374
0.02	0.00442	0.0095219	0.0044727	0.0000534
0.01	0.00314	0.0031194	0.0027096	0.0000011
		$p = 3$	$N = 500$	
a	SE	$A(\hat{a}_M^\dagger)$	$A(\hat{a}_M^\star)$	$A(\hat{a}_M)$
0.15	0.01129	0.2060530	0.1053830	0.0974672
0.1	0.00948	0.1101210	0.0599274	0.0371190
0.09	0.00904	0.0981768	0.0552909	0.0315313
0.08	0.00857	0.0787344	0.0494322	0.0264778
0.07	0.00806	0.0646111	0.0408115	0.0180659
0.06	0.00750	0.0552909	0.0372266	0.0107216
0.05	0.00689	0.0457591	0.0285225	0.0070294
0.04	0.00619	0.0341372	0.0209803	0.0038401
0.03	0.00539	0.0260609	0.0136816	0.0026151
0.02	0.00442	0.0143398	0.0087518	0.0003624
0.01	0.00314	0.0065625	0.0052037	0.0000047
		$p = 3$	$N = 1000$	
a	SE	$A(\hat{a}_M^\dagger)$	$A(\hat{a}_M^\star)$	$A(\hat{a}_M)$
0.15	0.01129	0.1972380	0.1263580	0.1256910
0.1	0.00948	0.1072330	0.0711925	0.0557335
0.09	0.00904	0.0890502	0.0658864	0.0476321
0.08	0.00857	0.0749391	0.0581898	0.0351544
0.07	0.00806	0.0704033	0.0482156	0.0272595
0.06	0.00750	0.0587613	0.0425729	0.0216652
0.05	0.00689	0.0522590	0.0362011	0.0157389
0.04	0.00619	0.0402168	0.0325967	0.0095505
0.03	0.00539	0.0314704	0.0210523	0.0053822
0.02	0.00442	0.0199987	0.0145513	0.0023102
0.01	0.00314	0.0114957	0.0053734	0.0001032

indicates that $A(\cdot)$ is a good approximation to F. The same procedures are applied to * and † versions with the notations $\hat{a}_M^*$ and $\hat{a}_M^\dagger$, respectively. A rigorous standard error formula associated with $A(\hat{a}_M)$ is hard to obtain. However, as given in Sun (1991), the standard error of $A(\hat{a}_M)$ is asymptotically as $M \to \infty$,

$$SE = \left\{ \frac{a(1-a)}{M} \right\}^{1/2}.$$

We have done a great amount of simulation using quasi-Newton method by FORTRAN programming. Tables 3.3 and 3.4 are representative tables. Table 3.3 is the case where normal distribution is the least interesting. The values of the geometric constants included in $A(\cdot)$ are figured in Table 3.1. Further, Table 3.4 is the results of simulations in the case where contaminated normal distribution ($\varepsilon = 0.5, c = 9.0$) is the least interesting structure. In our simulation, we use $M = 1000$. We observe from Table 3.3 that the convergence of the maxima of $\mathcal{GMS}$ is so slowly that the limiting approximation is poor for $N = 200, 500$. However, the maxima of $\mathcal{GMS}^\dagger$ converges faster than that of $\mathcal{GMS}$ and $\mathcal{GMS}^*$, thus limiting approximations are better in even if $N = 500$. The trend that the approximation of $\mathcal{GMS}^\dagger$ is better than that of $\mathcal{GMS}^*$ and $\mathcal{GMS}$ can also be recognized from Table 3.4. Further, we note that its convergence under contaminated normal is faster than that under normal.

4   Proofs and related calculations (I)

This section includes proofs and detailed calculations of the results presented in the previous sections. We need the following result.

LEMMA 4.1 *Let X have a spherically symmetric distribution with unit covariance matrix such that $m_8 < \infty$. Then*

$$\sup_{\alpha \in \mathscr{S}^{p-1}} \left| \frac{1}{N} \sum_{j=1}^{N} (\alpha' X_j)^3 \right| \xrightarrow{P} 0$$

and

$$\sup_{\alpha \in \mathscr{S}^{p-1}} \left| \frac{1}{N} \sum_{j=1}^{N} (\alpha' X_j)^3 X_j - (\theta_4 + 3)\alpha \right| \xrightarrow{P} 0$$

as $N \to \infty$.

PROOF. Direct computation gives

$$\frac{1}{N}\sum_{j=1}^{N}(\alpha'X_j)^3 = \sum_{k,l,u=1}^{p}\alpha_k\alpha_l\alpha_u\left\{\frac{1}{N}\sum_{j=1}^{N}X_{jk}X_{jl}X_{ju}\right\},$$

where $X_j = (X_{j1},\ldots,X_{jp})'$. We see from Anderson (1993, p. 7) that

$$\frac{1}{N}\sum_{j=1}^{N}X_{jk}X_{jl}X_{ju} = o_P(1).$$

Thus

$$\sup_{\alpha\in\mathscr{S}^{p-1}}\left|\frac{1}{N}\sum_{j=1}^{N}(\alpha'X_j)^3\right| \le \left\{\sum_{k,l,u=1}^{p}\sup_{\alpha\in\mathscr{S}^{p-1}}|\alpha_k\alpha_l\alpha_u|\right\}o_P(1) = o_P(1).$$

For the second assertion, it suffices to show that, for arbirary $q(1 \le q \le p)$,

$$\sup_{\alpha\in\mathscr{S}^{p-1}}\left|\frac{1}{N}\sum_{j=1}^{N}(\alpha'X_j)^3 X_{jq} - (\theta_4+3)\alpha_q\right| \xrightarrow{P} 0.$$

For this, let f_q be a p-dimensional unit vector such that $f_q = (f_{q1},\ldots,f_{qp}) = (\delta_{q1},\ldots,\delta_{qp})'$, where δ_{kl} is Kronecker's symbol. We must show that

$$\sup_{\alpha\in\mathscr{S}^{p-1}}\left|\frac{1}{N}\sum_{j=1}^{N}(\alpha'X_j)^3(f_q'X_j) - (\theta_4+3)(\alpha'f_q)\right| \xrightarrow{P} 0. \qquad (4.1)$$

The left hand side of (4.1) can be expressed as

$$\sup_{\alpha\in\mathscr{S}^{p-1}}\left|\sum_{k,l,u,v=1}^{p}\alpha_k\alpha_l\alpha_u f_{qv}t_N(k,l,u,v)\right|,$$

where

$$t_N(k,l,u,v) = \frac{1}{N}\sum_{j=1}^{N}X_{jk}X_{jl}X_{ju}X_{jv} - \frac{m_4}{p(p+2)}(\delta_{kl}\delta_{uv}+\delta_{ku}\delta_{lv}+\delta_{kv}\delta_{lu}).$$

Since, by using formulas about the forth moment in Anderson (1993, p. 7),

$$t_N(k,l,u,v) = o_P(1),$$

for all k,l,u,v,

$$\sup_{\alpha\in\mathscr{S}^{p-1}}\left|\sum_{k,l,u,v=1}^{p}\alpha_k\alpha_l\alpha_u f_{qv}t_N(k,l,u,v)\right| \le \left\{\sum_{k,l,u,v=1}^{p}\sup_{\alpha\in\mathscr{S}^{p-1}}|\alpha_k\alpha_l\alpha_u f_{qv}|\right\}|o_P(1)| = o_P(1),$$

which completes the proof.

PROOF OF LEMMA 1.1.   Note that the assertion of Lemma 1.1 follows provided that

$$\sup_{\alpha \in \mathscr{S}^{p-1}} |\tilde{Z}_{k,N}(\alpha) - \tilde{W}_{k,N}(\alpha)|$$

converges to zero in probability for $k = 3, 4$.   We only prove the case $k = 4$.   For this, we identify any pair (a, B) of a vector $a = (a_1, \ldots, a_p)' \in R^p$ and a real $p \times p$ symmetric matrix $B = (b_{ij})$ with $p + p(p+1)/2$-dimensional vector

$$(a_1, \ldots, a_p, b_{11}, \ldots, b_{1p}, b_{22}, \ldots, b_{2p}, \ldots, b_{pp})'.$$

Let $\lambda = (0, I_p)$, the true values of parameters, and let $\hat{\lambda} = (\overline{X}, S^{-1/2})$, a consistent estimator of λ.   Now it follows by one term Taylor expansion of estimator that

$$\tilde{W}_{4,N}(\alpha) - \tilde{Z}_{4,N}(\alpha) = I_1(\alpha) + I_2(\alpha) + I_3(\alpha),$$

where

$$I_1(\alpha) = -\alpha' B^* (\sqrt{N}\overline{X}) \left\{ \frac{4}{N} \sum_{j=1}^{N} (\alpha' B^* (X_j - a^*))^3 \right\},$$

$$I_2(\alpha) = 2(\theta_4 + 3)\alpha' \left\{ \frac{1}{\sqrt{N}} \sum_{j=1}^{N} (X_j X_j' - I_p) \right\} \alpha,$$

$$I_3(\alpha) = \alpha' \{ \sqrt{N}(S^{-1/2} - I_p) \} \left\{ \frac{4}{N} \sum_{j=1}^{N} (\alpha' B^* (X_j - a^*))^3 (X_j - a^*) \right\}$$

and $\lambda^* = (a^*, B^*)$ satisfying $|\lambda^* - \lambda| < |\hat{\lambda} - \lambda|$.   Now it is easy to verify by Lemma 4.1 and Slutsky's theorem that

$$\sup_{\alpha \in \mathscr{S}^{p-1}} |I_1(\alpha)| \xrightarrow{P} 0$$

and that

$$\sup_{\alpha \in \mathscr{S}^{p-1}} \left| \frac{1}{N} \sum_{j=1}^{N} (\alpha' B^* (X_j - a^*))^3 (X_j - a^*) - (\theta_4 + 3)\alpha \right| \xrightarrow{P} 0.$$

The equality

$$S^{-1/2} - I_p = -S^{-1/2}(I_p + S^{1/2})^{-1}(S - I_p)$$

gives that

$$\sup_{\alpha \in \mathscr{S}^{p-1}} \sqrt{N} \left| \alpha' \left\{ S^{-1/2} - I_p + \frac{1}{2}(S - I_p) \right\} \alpha \right| = \sup_{\alpha \in \mathscr{S}^{p-1}} |\alpha' M_N U_N \alpha|$$

$$\leq \max\{|\kappa(1)|, |\kappa(p)|\},$$

where

$$M_N = \frac{1}{2} I_p - S^{-1/2}(I_p + S^{1/2})^{-1}, \qquad U_N = \sqrt{N}(S - I_p),$$

and $\kappa(1)$ and $\kappa(p)$ are the largest and the smallest eigenvalue of $M_N U_N$, respectively. Since $S^{-1/2}$ converges to I_p in probability, we get

$$\sup_{\alpha \in \mathscr{S}^{p-1}} \sqrt{N} \left| \alpha' \left\{ S^{-1/2} - I_p + \frac{1}{2}(S - I_p) \right\} \alpha \right| \xrightarrow{P} 0.$$

Thus

$$|I_2(\alpha) + I_3(\alpha)| = 2(\theta_4 + 3) \left| \alpha' \left\{ U_N - \frac{1}{\sqrt{N}} \sum_{j=1}^{N} (X_j X_j' - I_p) \right\} \alpha \right| + R_N,$$

where $R_N = R_N(\alpha)$ satisfies

$$\sup_{\alpha \in \mathscr{S}^{p-1}} |R_N(\alpha)| \xrightarrow{P} 0,$$

From the fact that

$$\sup_{\alpha \in \mathscr{S}^{p-1}} |\alpha'(\sqrt{N}\overline{X})\overline{X}'\alpha| \xrightarrow{P} 0,$$

it can be concluded that

$$\sup_{\alpha \in \mathscr{S}^{p-1}} |I_2(\alpha) + I_3(\alpha)| \xrightarrow{P} 0.$$

This completes the proof.

PROOF OF THEOREM 1.2. To compute $v(\alpha, \bar{\alpha})$, we can use the first part of the proof of Theorem 2.1 in Baringhaus and Henze (1991), which is based on the formula given in Fang *et al.* (1989, p.72). By using this, straightforward algebra gives (1.6), (1.7) and (1.8). Note that the metric space $(\mathscr{S}^{p-1}, |\cdot|)$ satisfies the metric entropy condition

$$\int_0^1 \{\log n(\mathscr{S}^{p-1}, \varepsilon)\}^{1/2} d\varepsilon < \infty,$$

where, for $\varepsilon > 0$, $n(\mathscr{S}^{p-1}, \varepsilon)$ stands for the smallest positive integer m such that $\mathscr{S}^{p-1}$ can be covered by m subsets each having diameter at most ε with respect to $|\cdot|$. Putting

$$\tilde{Z}_3(\alpha) = (\alpha'X)^3 - 3(\alpha'X),$$

$$\tilde{Z}_4(\alpha) = (\alpha'X)^4 - (\theta_4 + 3)\{2(\alpha'X)^2 - 1\},$$

for $\alpha \in \mathscr{S}^{p-1}$, we have

$$|\tilde{Z}_3(\alpha) - \tilde{Z}_3(\bar{\alpha})| \le 3\{|X|^3 + |X|\}|\alpha - \bar{\alpha}|,$$

and

$$|\tilde{Z}_4(\alpha) - \tilde{Z}_4(\bar{\alpha})| \le 4\{|X|^4 + (\theta_4 + 3)|X|^2\}|\alpha - \bar{\alpha}|,$$

for $\alpha, \bar{\alpha} \in \mathscr{S}^{p-1}$. Therefore, again by putting

$$\tilde{Z}(\alpha) = \begin{bmatrix} \tilde{Z}_3(\alpha) \\ \tilde{Z}_4(\alpha) \end{bmatrix},$$

for $\alpha \in \mathscr{S}^{p-1}$, we see that

$$|\tilde{Z}(\alpha) - \tilde{Z}(\bar{\alpha})| \le [3\{|X|^3 + |X|\} + 4\{|X|^4 + (\theta_4 + 3)|X|^2\}]|\alpha - \bar{\alpha}|.$$

Since $m_8 < \infty$, the proof is complete (see Araujo and Giné (1980)).

PROOF OF THEOREM 1.4. Let $Y = \Sigma^{-1/2}(X - \mu)$, $Y_j = \Sigma^{-1/2}(X_j - \mu)$ $(j = 1, \ldots, N)$, where $\Sigma^{-1/2}$ is a positive definite square root of Σ^{-1}. Let $\boldsymbol{GMI}(\alpha)$ and $w_N(\alpha)$ be the ones as in (1.1) and (1.3), respectively, with Y_j instead of X_j. Since Y_j is spherical and satisfying

$$m_8 = E[|Y|^8] = E[\{(X - \mu)'\Sigma^{-1}(X - \mu)\}^4] < \infty,$$

Lemme 1.1, Theorem 1.2 and Corollary 1.3 yield that

$$w_N(\cdot) \Rightarrow Z(\cdot),$$

where $Z(\cdot)$ is the Gaussian random field given in Corollary 1.3. Since the maxima of $\mathscr{GMI}$ is affine invariant and the mapping assigning to each R^2 valued continuous function on $\mathscr{S}^{p-1}$ the square of its maximum is continuous, we have

$$\max_{\alpha \in \mathscr{S}^{p-1}} N\mathscr{GMI}(\alpha) = \max_{\alpha \in \mathscr{S}^{p-1}} N\boldsymbol{GMI}(\alpha) = \max_{\alpha \in \mathscr{S}^{p-1}} |w_N(\alpha)|^2 \Rightarrow \max_{\alpha \in \mathscr{S}^{p-1}} |Z(\alpha)|^2,$$

which completes the proof.

PROOF OF PROPOSITION 1.5. The covariance function of the limiting Gaussian random field $\xi(\alpha, \beta)$ is

$$r((\alpha, \beta), (\bar{\alpha}, \bar{\beta})) = \frac{\beta_1 \bar{\beta}_1}{\sigma_3^2} \{\tau_{3,1}(\alpha' \bar{\alpha})^3 + \tau_{3,2}(\alpha' \bar{\alpha})\}$$

$$+ \frac{\beta_2 \bar{\beta}_2}{\sigma_4^2} \{\tau_{4,1}(\alpha' \bar{\alpha})^4 + \tau_{4,2}(\alpha' \bar{\alpha})^2 + \tau_{4,3}\},$$

for $\alpha, \bar{\alpha} \in \mathscr{S}^{p-1}$, $\beta, \bar{\beta} \in \mathscr{S}^1$, where $\tau_{3,1}, \tau_{3,2}, \tau_{4,1}, \tau_{4,2}$ and $\tau_{4,3}$ are given in Section 1. First we prove the equality (1.22) in the case of $p \geq 3$. Now the powers of $(\alpha' \bar{\alpha})$ can be expressed as

$$(\alpha' \bar{\alpha}) = \frac{C_1^\gamma(\alpha' \bar{\alpha})}{2\gamma},$$

$$(\alpha' \bar{\alpha})^2 = \frac{C_2^\gamma(\alpha' \bar{\alpha})}{2\gamma(\gamma + 1)} + \frac{1}{2(\gamma + 1)},$$

$$(\alpha' \bar{\alpha})^3 = \frac{3 C_3^\gamma(\alpha' \bar{\alpha})}{4\gamma(\gamma + 1)(\gamma + 2)} + \frac{3 C_1^\gamma(\alpha' \bar{\alpha})}{4\gamma(\gamma + 2)},$$

$$(\alpha' \bar{\alpha})^4 = \frac{3 C_4^\gamma(\alpha' \bar{\alpha})}{2\gamma(\gamma + 1)(\gamma + 2)(\gamma + 3)} + \frac{3 C_2^\gamma(\alpha' \bar{\alpha})}{2\gamma(\gamma + 1)(\gamma + 3)} - \frac{3}{4(\gamma + 1)(\gamma + 2)},$$

where $C_d^\gamma(t)$ is the Gegenbauer polynomial of degree d and order $\gamma = (p - 2)/2$ (see Stein and Weiss (1971, pp. 143, 149)). These can be obtained from equations

$$C_0^\gamma(t) = 1, \qquad C_1^\gamma(t) = 2\gamma t, \qquad C_2^\gamma(t) = 2\gamma(\gamma + 1)t^2 - \gamma,$$

$$C_3^\gamma(t) = \frac{4}{3}\gamma(\gamma + 1)(\gamma + 2)t^3 - 2\gamma(\gamma + 1)t,$$

$$C_4^\gamma(t) = \frac{2}{3}\gamma(\gamma + 1)(\gamma + 2)(\gamma + 3)t^4 - 2\gamma(\gamma + 1)(\gamma + 2)t^2 + \frac{1}{2}\gamma(\gamma + 1).$$

The addition theorem of the Gegenbauer polynomial gives that

$$C_d^\gamma(\alpha' \bar{\alpha}) = \left(1 + \frac{d}{\gamma}\right)^{-1} \sum_{l=1}^{v(d)} \zeta_{d,l}(\alpha) \zeta_{d,l}(\bar{\alpha}),$$

for $\alpha, \bar{\alpha} \in \mathscr{S}^{p-1}$, where $\{\zeta_{d,l}(\alpha) : l = 1, 2, \ldots, v(d)\}$ are linearly independent surface harmonics of degree d being orthonormal with respect to the uniform distribution on $\mathscr{S}^{p-1}$, and

$$v(d) = \frac{(p - 3 + d)!}{(p - 2)!(d - 1)!} + \frac{(p - 2 + d)!}{(p - 2)!d!}.$$

(see Erdélyi *et al.* (1953, p.243), and Stein and Weiss (1971)). By using the above equalities, we obtain the expression (1.22).

For $p = 2$, let $C_d^0(t)$ be the Chebyshev polynomial of degree d. We see from the relation

$$C_d^0(t) = \frac{d}{2} \lim_{\gamma \to 0} \frac{C_d^\gamma(t)}{\gamma}$$

for $d \geq 1$ (see Erdélyi *et al.* (1953, p.184)) that

$$\sum_{l=1}^{\nu(d)} \varphi_{d,l}(\alpha)\varphi_{d,l}(\bar{\alpha}) = 2C_d^0(\alpha'\bar{\alpha}).$$

And $C_0^0(t) = 1$, $C_1^0(t) = t$, $C_2^0(t) = 2t^2 - 1$, $C_3^0(t) = 4t^3 - 3t$, $C_4^0(t) = 8t^4 - 8t^2 + 1$ give the expression. This completes the proof.

CALCULATION OF THE METRIC TENSOR MATRIX.   The covariance function of the Gaussian random field $\xi(t)$ is

$$r(s,t) = c_3(\alpha(\phi), \alpha(\bar{\phi}))\beta_1(\varphi)\beta_1(\bar{\varphi}) + c_4(\alpha(\phi), \alpha(\bar{\phi}))\beta_2(\varphi)\beta_2(\bar{\varphi}),$$

for $t = (\phi, \varphi)$, $s = (\bar{\phi}, \bar{\varphi}) \in I$. By the spherical polar coordinate representation of α and β, we see that

$$\sum_{l=1}^{p} \frac{\partial \alpha_l}{\partial \phi_i} \alpha_l = \sum_{l=1}^{2} \frac{\partial \beta_l}{\partial \varphi} \beta_l = 0, \qquad \sum_{l=1}^{p} \frac{\partial \alpha_l}{\partial \phi_i} \frac{\partial \alpha_l}{\partial \phi_j} = \delta_{ij} \sum_{l=1}^{p} \left(\frac{\partial \alpha_l}{\partial \phi_i}\right)^2,$$

where δ_{ij} is Kronecker's symbol. We have

$$\left.\frac{\partial^2 r(s,t)}{\partial \phi_k \partial \bar{\phi}_l}\right|_{\phi=\bar{\phi}} = \left\{\frac{3\tau_{3,1} + \tau_{3,2}}{\sigma_3^2}\right\}\delta_{kl}\beta_1^2(\varphi) \sum_{l=1}^{p} \left(\frac{\partial \alpha_l}{\partial \phi_k}\right)^2$$

$$+ \left\{\frac{4\tau_{4,1} + 2\tau_{4,2}}{\sigma_4^2}\right\}\delta_{kl}\beta_2^2(\varphi) \sum_{l=1}^{p} \left(\frac{\partial \alpha_l}{\partial \phi_k}\right)^2, \qquad (4.2)$$

$$\left.\frac{\partial^2 r(s,t)}{\partial \varphi \partial \bar{\varphi}}\right|_{\varphi=\bar{\varphi}} = \sum_{l=1}^{2} \left(\frac{\partial \beta_l}{\partial \varphi}\right)^2, \qquad (4.3)$$

$$\left.\frac{\partial^2 r(s,t)}{\partial \phi_k \partial \bar{\varphi}}\right|_{s=t} = \left.\frac{\partial^2 r(s,t)}{\partial \bar{\phi}_k \partial \varphi}\right|_{s=t} = 0. \qquad (4.4)$$

Equations (4.2), (4.3) and (4.4) gives (2.3).

CALCULATION OF THE SCALAR CURVATURE.   We give a detailed calculation of scalar curvature with $R(t)$ as its metric tensor.   The inverse matrix $R^{-1}(t)$ of the metric tensor matrix $R(t) = (g_{ij}(t))$ is written as $R^{-1}(t) = (g^{ij}(t))$.   Let

$G^l_{ijk}(t)$ be Riemannian curvature tensor and let $G_{ij}(t) = \sum_{k=1}^d G^k_{ijk}(t)$ be Ricci curvature tensor (see Sun (1993)). The scalar curvature is defined as

$$S(t) = \sum_{i,j=1}^p g^{ij}(t) G_{ij}(t).$$

From diagonality of $R(t)$, it follows that

$$S(t) = \sum_{j,k=1}^p g^{jj}(t) G^k_{jjk}(t).$$

Let $\Phi = \{\phi_1, \ldots, \phi_{p-1}\}$ and let $\Psi = \{\varphi\}$. In what follow, $j \in \Phi$ and $k \in \Psi$ designate $t_j \in \{\phi_1, \ldots, \phi_{p-1}\}$ and $t_k = \varphi$, respectively. Using this notation, $S(t)$ is factorized as

$$S(t) = S_1(t) + S_2(t) + S_3(t) + S_4(t), \tag{4.5}$$

where

$$S_1(t) = \sum_{k \in \Phi, j \in \Phi} g^{jj} R^k_{jjk}, \qquad S_2(t) = \sum_{k \in \Psi, j \in \Phi} g^{jj} R^k_{jjk},$$

$$S_3(t) = \sum_{k \in \Phi, j \in \Psi} g^{jj} R^k_{jjk}, \qquad S_4(t) = \sum_{k \in \Psi, j \in \Psi} g^{jj} R^k_{jjk}.$$

Let

$$\eta_k = \sum_{l=1}^p \left(\frac{\partial \alpha_l}{\partial \theta_k}\right)^2, \qquad \xi = \sum_{l=1}^2 \left(\frac{\partial \beta_l}{\partial \varphi}\right)^2.$$

By calculating the Christoffel symbols included in Riemannian curvature, we obtain that

$$S_4(t) = 0. \tag{4.6}$$

Next we calculate S_1, S_2 and S_3. Similar computations yield that $S_i(t)$ $(i = 1, 2, 3)$ are factorized as

$$S_i(t) = \{C(\beta)\}^{-1} S_{i1}(t) + \{C(\beta)\}^{-2} S_{i2}(t), \tag{4.7}$$

where

$$S_{11}(t) = \sum_{k,j \in \Phi} \left[-\frac{1}{2}\eta_j^{-1}\eta_k^{-2}\left\{\frac{\partial \eta_k}{\partial \phi_j}\right\}^2 + \frac{1}{2}\eta_j^{-1}\eta_k^{-1}\left\{\frac{\partial^2 \eta_k}{\partial \phi_j^2}\right\} \right.$$
$$\left. + \frac{1}{2}\eta_j^{-1}\eta_k^{-1}\left\{\frac{\partial^2 \eta_j}{\partial \phi_k^2}\right\} - \frac{1}{2}\eta_j^{-1}\eta_k^{-2}\frac{\partial \eta_k}{\partial \phi_k}\frac{\partial \eta_j}{\partial \phi_k} + \frac{1}{4}\sum_{q=1}^{p-1}\eta_j^{-1}\eta_k^{-1}\eta_q^{-1}\frac{\partial \eta_j}{\partial \phi_q}\frac{\partial \eta_k}{\partial \phi_q} \right],$$

$$S_{12}(t) = \frac{1}{4}(p-1)(p-2)\xi^{-1}\left\{\frac{\partial C(\beta)}{\partial \varphi}\right\}^2,$$

$$S_{21}(t) = S_{31}(t) = \frac{p-1}{2}\xi^{-1}\frac{\partial^2 C(\beta)}{\partial \varphi^2},$$

$$S_{22}(t) = S_{32}(t) = -\frac{p-1}{4}\xi^{-1}\left\{\frac{\partial C(\beta)}{\partial \varphi}\right\}^2.$$

We see from easily but somewhat lengthy computations that $S_{11}(t) = -(p-1)(p-2)$. Thus we obtain that

$$\sum_{i=1}^{3} S_{i1}(t) = -(p-1)(p-2) + (p-1)\xi^{-1}\frac{\partial^2 C(\beta)}{\partial \varphi^2} \qquad (4.8)$$

and

$$\sum_{i=1}^{3} S_{i2}(t) = \frac{1}{4}(p-1)(p-4)\xi^{-1}\left\{\frac{\partial C(\beta)}{\partial \varphi}\right\}^2. \qquad (4.9)$$

So we get $S(t)$ presented in (2.4) by using the equalities from (4.5) to (4.9).

Part II   Nonnull Behaviors

5   Asymptotic behaviors under interesting structure

5.1   The problem

In the previous sections, we discussed the significance level of a generalized projection pursuit and derived its approximation formula. These considerations may be interpreted as the problem of investigating null performances of goodness of fit statistics. In the area of testing statistical hypothesis, it is important to investigate nonnull performances of test statistics. Relating to this, we are interested to get the distribution of test statistic under certain selected alternative and to evaluate its power.

In exploratory projection pursuit, above considerations are correspoding to evaluating behaviors of projection index under the interesting structure. However, it is not possible to define the interesting structure because of its variety. Thus we only deal with the interesting structure such that it is elliptically symmetric and is not the least interesting. That is, the performances of the index for the case that some elliptically symmetric structure X is the least interesting are investigated under another elliptically symmetric structure X^* satisfying certain conditions. More concretely, main purpose of

this section is to evaluate

$$P\left\{\max_{\alpha \in \mathscr{S}^{p-1}} \mathscr{GMI}(\alpha) \geq z\right\},\tag{5.1}$$

under X^*, where z is the maxima of $\mathscr{GMI}$, which is utilized for the case when X is the least interesting, based on a random sample drawn from X^*.

Similar to (1.2), it is difficult to evaluate (5.1) theoretically in finite N. Therefore some asymptotic results are needed. We proceed this in the next subsection.

5.2   Asymptotic consideration

Suppose that the least interesting elliptically symmetric structure is X with θ_4 so that the corresponding index is (1.1), that is,

$$\mathscr{GMI}(\alpha) = \sum_{k=3}^{4}\left[\frac{1}{N}\sum_{j=1}^{N}\left\{\frac{H_k(\alpha'S^{-1/2}(X_j - \bar{X})) - \theta_k}{\sigma_k}\right\}\right]^2.$$

Let X^* be another elliptically symmetric distributed random p-vector with mean μ^* and covariance matrix Σ^* such that

$$\theta_4^* = E[H_4(\alpha'(\Sigma^*)^{-1/2}(X^* - \mu^*))] \neq \theta_4$$

and $m_8 = E[\{(X^* - \mu^*)'(\Sigma^*)^{-1}(X^* - \mu^*)\}^4] < \infty$. In this subsection we discuss asymptotic behaviors of $\mathscr{GMI}$ under X^*. Note again that

$$N\mathscr{GMI}(\alpha) = \left\{\frac{\tilde{W}_{3,N}(\alpha)}{\sigma_3}\right\}^2 + \left\{\frac{\tilde{W}_{4,N}(\alpha)}{\sigma_4}\right\}^2,\tag{5.2}$$

for $\alpha \in \mathscr{S}^{p-1}$, where $\tilde{W}_{3,N}(\cdot)$ and $\tilde{W}_{4,N}(\cdot)$ are those in (1.4). We use the following trivial fact without proof.

LEMMA 5.1 *Let X have a spherically symmetric distribution with unit covariance matrix such that $m_8 < \infty$. Then*

$$\sup_{\alpha \in \mathscr{S}^{p-1}}\left|\frac{1}{\sqrt{N}}\left(\frac{\tilde{W}_{3,N}(\alpha)}{\sigma_3}\right)^2\right| \xrightarrow{P} 0$$

as $N \rightarrow \infty$.

Let $C_1(\mathscr{S}^{p-1})$ be the separable Banach space of real valued continuous functions defined on $\mathscr{S}^{p-1}$, endowed with the supremum norm. Let $X_1, \ldots, X_N$ be a random sample of size N drawn from X^* *not* from X. Then from Lemma 5.1 and (5.2), we see that

$$\sqrt{N}\left\{\mathscr{GMS}(\alpha) - \left(\frac{\theta_4^* - \theta_4}{\sigma_4}\right)^2\right\} = \sqrt{N}\left\{\left(\frac{\tilde{W}_{4,N}(\alpha)}{\sqrt{N}\sigma_4}\right)^2 - \left(\frac{\theta_4^* - \theta_4}{\sigma_4}\right)^2\right\} + R_N,$$

where $R_N = R_N(\alpha)$ converges, as a member of $C_1(\mathscr{S}^{p-1})$, to zero in probability. Let

$$(\sigma_4^*)^2 = A - Var[H_4(\alpha' S^{-1/2}(X_1 - \bar{X}))],$$

which is asymptotic variance taken with respect to the distribution of X^*.

It is easy to show that the limiting random field derived from $\mathscr{GMS}$ is also Gaussian. The result is stated in the following theorem.

THEOREM 5.2 *Suppose that X^* has an elliptically symmetric distribution with mean μ^* and covariance matrix Σ^* such that*

$$\theta_4^* = E[H_4(\alpha'(\Sigma^*)^{-1/2}(X^* - \mu^*))] \neq \theta_4$$

and

$$m_{2k} = E[\{(X^* - \mu^*)'(\Sigma^*)^{-1}(X^* - \mu^*)\}^k] < \infty,$$

for $k = 2, 3, 4$. Then we have

$$\max_{\alpha \in \mathscr{S}^{p-1}} \sqrt{N}\left\{\mathscr{GMS}(\alpha) - \left(\frac{\theta_4^* - \theta_4}{\sigma_4}\right)^2\right\}\left\{2\left|\frac{\theta_4^* - \theta_4}{\sigma_4}\right|\left(\frac{\sigma_4^*}{\sigma_4}\right)\right\}^{-1} \Rightarrow \max_{\alpha \in \mathscr{S}^{p-1}} Z_4(\alpha),$$

where $Z_4(\alpha), \alpha \in \mathscr{S}^{p-1}$, is a Gaussian random field with mean zero, unit variance and covariance function $c_4(\alpha, \bar{\alpha})$ given in (1.12), for $\alpha, \bar{\alpha} \in \mathscr{S}^{p-1}$.

The next representation of Z_4 is straightforwardly obtained from Proposition 1.5.

PROPOSITION 5.3 *The limiting Gaussian random field $Z_4(\alpha), \alpha \in \mathscr{S}^{p-1}$, can be represented in the form*

$$Z_4(\alpha) = \sqrt{\eta_{4,1}^*} \sum_{k=1}^{\nu(4)} \zeta_{4,k}(\alpha) N_{4,k} + \sqrt{\eta_{4,2}^*} \sum_{k=1}^{\nu(2)} \zeta_{2,k}(\alpha) N_{2,k} + \sqrt{\eta_{4,3}^*} N_0, \qquad (5.3)$$

where

$$\eta_{4,1}^* = \frac{m_8}{(\sigma_4^*)^2} \left[\frac{24}{p(p+2)(p+4)(p+6)} \right]^2,$$

$$\eta_{4,2}^* = \frac{144}{(\sigma_4^*)^2} \left[\frac{m_8}{\{p(p+2)(p+4)\}^2} - \frac{2m_6 m_4}{p^3(p+2)^3(p+4)} + \frac{m_4^3}{\{p(p+2)\}^4} \right],$$

$$\eta_{4,3}^* = \frac{9}{(\sigma_4^*)^2} \left[\frac{m_8}{\{p(p+2)\}^2} - \frac{4m_4 m_6}{p^3(p+2)^2} + \frac{4m_4^3}{p^4(p+2)^2} - \left\{ \frac{m_4}{p(p+2)} \right\}^2 \right],$$

$N_0; N_{2,k}, k = 1, \ldots, v(2); N_{4,k}, k = 1, \ldots, v(4)$ *are independent standard normal random variables and* $\{\zeta_{2,k} : k = 1, \ldots, v(2)\}$, $\{\zeta_{4,k} : k = 1, \ldots, v(4)\}$ *are linearly independent surface harmonics of degree 2 and 4, respectively, being orthonormal with respect to the uniform distribution on* $\mathscr{S}^{p-1}$, *where*

$$v(4) = \frac{p(p-1)(p+1)(p+6)}{24}, \qquad v(2) = \frac{(p-1)(p+2)}{2}.$$

Note that in above results, both of σ_4^* and θ_4^* do not depend on α. Let z be a fixed positive real number. Assume that X^* has an elliptically symmetric distribution with

$$\left(\frac{\theta_4^* - \theta_4}{\sigma_4} \right)^2 < z. \tag{5.4}$$

By arguments similar to the ones in Section 2, we evaluate the probability (5.1) as follows:

$$P\left\{ \max_{\alpha \in \mathscr{S}^{p-1}} \mathscr{GMI}(\alpha) \geq z \right\} \approx P\left\{ \max_{\alpha \in \mathscr{S}^{p-1}/2} Z_4(\alpha) \geq x \right\}$$

$$= P\left\{ \max_{\phi \in I_\phi} Z_4(\phi) \geq x \right\},$$

where Z_4 is the same one as in Theorem 5.2 and

$$x = \sqrt{N} \left\{ z - \left(\frac{\theta_4^* - \theta_4}{\sigma_4} \right)^2 \right\} \left\{ 2 \left| \frac{\theta_4^* - \theta_4}{\sigma_4} \right| \left(\frac{\sigma_4^*}{\sigma_4} \right) \right\}^{-1}.$$

From (5.4), x can be regarded as a point in tail of the distribution of the maxima of Z_4 for large N. The parameter space of the Gaussian random field Z_4 is $(p-1)$-dimensional. It is possible to reparametrize α in terms of ϕ by

the spherical polar coordinate transformation

$$\alpha = \alpha(\phi) = (\alpha_1(\phi), \ldots, \alpha_p(\phi))',$$

$$\phi = (\phi_1, \ldots, \phi_{p-1}) \in [0, \pi] \times \cdots \times [0, \pi] \equiv I_\phi \subset R^{p-1},$$

where $\phi_{p-1} = 0$ and $\phi_{p-1} = \pi$ represent the same point. We put $Z_4(\phi) \equiv Z_4(\alpha(\phi))$, for $\phi \in I_\phi$. We can see that $Z_4(\phi)$ is a $(p-1)$-dimensional non-singular differentiable Gaussian random field with mean zero, variance 1 and covariance function $c_4(\phi, \bar{\phi}) = c_4(\alpha(\phi), \alpha(\bar{\phi}))$, for $\phi, \bar{\phi} \in I_\phi$. The regularity conditions for $c_4(\cdot, \cdot)$ given in Sun (1993) can be easily checked. Especially the fact that $Z_4(\cdot)$ has a finite Karhunen-Loève expansion as in (5.3) guarantees the critical radius of the tube of the manifold derived from $c_4(\cdot, \cdot)$ is positive. So that Theorem 2.1 is also applicable to $Z_4(\phi), \phi \in I_\phi$.

By applying Theorem 2.1, we will get an approximation formula of (5.1). Similar to Section 3, it is necessary to check the accuracy of the approximation formula, however we only give theoretical results in the following.

Now the covariance function of $Z_4(\phi)$ is, by (1.12),

$$c_4(\phi, \bar{\phi}) = \{\tau_{4,1}(\alpha(\phi)'\alpha(\bar{\phi}))^4 + \tau_{4,2}(\alpha(\phi)'\alpha(\bar{\phi}))^2 + \tau_{4,3}\}/(\sigma_4^*)^2,$$

where $\tau_{4,1}, \tau_{4,2}$ and $\tau_{4,3}$ are given in Section 1. The metric tensor matrix is obtained as the following diagonal matrix:

$$\begin{aligned}
R(\phi) &\equiv \left\{ \left. \frac{\partial^2 c_4(\phi, \bar{\phi})}{\partial \phi \partial \bar{\phi}} \right|_{\phi = \bar{\phi}} \right\} \\
&= \left\{ \frac{4\tau_{4,1} + 2\tau_{4,2}}{(\sigma_4^*)^2} \right\} diag \left\{ \sum_{u=1}^{p} \left(\frac{\partial \alpha_u}{\partial \phi_1} \right)^2, \ldots, \sum_{u=1}^{p} \left(\frac{\partial \alpha_u}{\partial \phi_{p-1}} \right)^2 \right\}.
\end{aligned} \tag{5.5}$$

Thus it is easy to see that the scalar curvature is

$$S(\phi) = -(p-1)(p-2) \left\{ \frac{4\tau_{4,1} + 2\tau_{4,2}}{(\sigma_4^*)^2} \right\}^{-1} \tag{5.6}$$

(see Section 6). By using these, two geometric constants corresponding to (2.1) and (2.2) can be obtained as

$$\kappa_0^* = \frac{\omega_{p-1}}{4\pi^{p/2}} \left\{ \frac{4\tau_{4,1} + 2\tau_{4,2}}{(\sigma_4^*)^2} \right\}^{(p-1)/2}, \tag{5.7}$$

$$\kappa_2^* = \frac{(p-1)(p-2)\omega_{p-1}}{16\pi^{p/2}} \left\{ \frac{4\tau_{4,1} + 2\tau_{4,2}}{(\sigma_4^*)^2} \right\}^{(p-1)/2} \left[\left\{ \frac{4\tau_{4,1} + 2\tau_{4,2}}{(\sigma_4^*)^2} \right\}^{-1} - 1 \right]. \tag{5.8}$$

We summarize the above results as follow.

PROPOSITION 5.4 *Let z be a fixed positive real number. Suppose that X^* has an elliptically symmetric distribution with mean μ^*, covariance matrix Σ^* and*

$$\theta_4^* = E[H_4(\alpha'(\Sigma^*)^{-1/2}(X^* - \mu^*))] \neq \theta_4$$

such that

$$m_{2k} = E[\{(X^* - \mu^*)'(\Sigma^*)^{-1}(X^* - \mu^*)\}^k] < \infty,$$

for $k = 2, 3, 4$ and

$$\left(\frac{\theta_4^* - \theta_4}{\sigma_4}\right)^2 < z.$$

Then an approximation to (5.1) is that, for large N,

$$P\left\{ \max_{\alpha \in \mathscr{S}^{p-1}} \mathscr{GMI}(\alpha) \geq z \right\} \approx \kappa_0^* \psi\left(\frac{x^2}{2}, \frac{p}{2}\right) + \kappa_2^* \psi\left(\frac{x^2}{2}, \frac{p-2}{2}\right), \qquad (5.9)$$

where

$$x = \sqrt{N}\left\{ z - \left(\frac{\theta_4^* - \theta_4}{\sigma_4}\right)^2 \right\}\left\{ 2\left|\frac{\theta_4^* - \theta_4}{\sigma_4}\right|\left(\frac{\sigma_4^*}{\sigma_4}\right) \right\}^{-1},$$

κ_0^* *and* κ_2^* *are given in (5.7) and (5.8), respectively.*

We have obtained an approximation formula of (5.1). Checking the accuracy of the approximation formula (5.9) is left as the future work.

6 Proofs and related calculations (II)

This section gives the proofs and related calculations for Section 5.

PROOF OF THEOREM 5.2. Note that

$$\sqrt{N}\left\{ \left(\frac{\tilde{W}_{4,N}(\alpha)}{\sqrt{N}\,\sigma_4}\right)^2 - \left(\frac{\theta_4^* - \theta_4}{\sigma_4}\right)^2 \right\} = \left[\left(\frac{\tilde{W}_{4,N}(\alpha)}{\sqrt{N}\sigma_4}\right) + \left(\frac{\theta_4^* - \theta_4}{\sigma_4}\right) \right]\left(\frac{\sigma_4^*}{\sigma_4}\right)$$

$$\times \left[\frac{1}{\sqrt{N}} \sum_{j=1}^{N}\left\{ \frac{H_4(\alpha'S^{-1/2}(X_j - \bar{X})) - \theta_4^*}{\sigma_4^*} \right\} \right].$$

We see from Theorem 1.4 that

$$\frac{1}{\sqrt{N}} \sum_{j=1}^{N}\left\{ \frac{H_4(\alpha'S^{-1/2}(X_j - \bar{X})) - \theta_4^*}{\sigma_4^*} \right\}$$

converges weakly to $Z_4(\alpha)$, where $Z_4(\alpha)$, $\alpha \in \mathscr{S}^{p-1}$ is a Gaussian random field with mean zero, unit variance and covariance function $c_4(\alpha, \bar{\alpha})$, $\alpha, \bar{\alpha} \in \mathscr{S}^{p-1}$, given in Section 1. Since

$$\sup_{\alpha \in \mathscr{S}^{p-1}} \left| \left(\frac{\tilde{W}_{4,N}(\alpha)}{\sqrt{N}\sigma_4} \right) \right| \xrightarrow{P} \frac{\theta_4^* - \theta_4}{\sigma_4},$$

we obtain weak convergence

$$\sqrt{N} \left\{ \left(\frac{\tilde{W}_{4,N}(\cdot)}{\sqrt{N}\sigma_4} \right)^2 - \left(\frac{\theta_4^* - \theta_4}{\sigma_4} \right)^2 \right\} \Rightarrow 2 \left(\frac{\theta_4^* - \theta_4}{\sigma_4} \right) \left(\frac{\sigma_4^*}{\sigma_4} \right) Z_4(\cdot),$$

which derive the assertion. This completes the proof.

PROOF OF PROPOSITION 5.3. The proof is similar to that of Proposition 1.5, thus it is omitted.

CALCULATION OF THE SCALAR CURVATURE. By calculations similar to the ones in Section 4, we get the metric tensor (5.5)

$$R(\phi) \equiv \left\{ \left. \frac{\partial^2 c_4(\phi, \bar{\phi})}{\partial \phi \partial \bar{\phi}} \right|_{\phi = \bar{\phi}} \right\}$$

$$= \left\{ \frac{4\tau_{4,1} + 2\tau_{4,2}}{(\sigma_4^*)^2} \right\} diag \left\{ \sum_{u=1}^p \left(\frac{\partial \alpha_u}{\partial \phi_1} \right)^2, \ldots, \sum_{u=1}^p \left(\frac{\partial \alpha_u}{\partial \phi_{p-1}} \right)^2 \right\}.$$

Let

$$\varphi_k = \sum_{u=1}^p \left(\frac{\partial \alpha_u}{\partial \phi_k} \right)^2,$$

for $k = 1, \ldots, p-1$. In this case, by diagonality of $R(\phi)$, $S(\phi)$ is easily obtained. In fact easily but somewhat lengthy algebras yield

$$S(\phi) = \left\{ \frac{4\tau_1 + 2\tau_2}{(\sigma_4^*)^2} \right\}^{-1} \sum_{i,j=1}^{p-1} \left[-\frac{1}{2} \varphi_i^{-1} \varphi_j^{-2} \left(\frac{\partial \varphi_j}{\partial \phi_i} \right)^2 + \frac{1}{2} \varphi_i^{-1} \varphi_j^{-1} \frac{\partial^2 \varphi_j}{\partial \phi_i^2} \right.$$

$$- \frac{1}{2} \varphi_i^{-1} \varphi_j^{-2} \left(\frac{\partial \varphi_j}{\partial \varphi_j} \right) \left(\frac{\partial \varphi_i}{\partial \phi_j} \right) + \frac{1}{2} \varphi_i^{-1} \varphi_j^{-1} \frac{\partial^2 \varphi_i}{\partial \phi_j^2} \qquad (6.1)$$

$$\left. + \frac{1}{4} \sum_{l=1}^{p-1} \varphi_i^{-1} \varphi_j^{-1} \varphi_l^{-1} \left(\frac{\partial \varphi_l}{\partial \phi_l} \right) \left(\frac{\partial \varphi_j}{\partial \phi_l} \right) \right].$$

From similar calculations given in Section 4 of Naito (1996), the summation in

the right hand side of (6.1) is equal to $-(p-1)(p-2)$. So we get

$$S(\phi) = -(p-1)(p-2)\left\{\frac{4\tau_1 + \tau_2}{(\sigma_4^*)^2}\right\}^{-1}$$

which is equal to the right hand side of (5.6).

Acknowledgments

References

[1] T. W. Anderson, Nonnormal Multivariate distributions: Inference based on elliptically contoured distributions, In Multivariate Analysis: Future Direction (C. R. Rao, Editor), North-Holland, 1993.

[2] A. Araujo and E. Giné, The Central Limit Theorem for Real and Banach Valued Random Variables, Wiley, New York, 1980.

[3] L. Baringhaus and N. Henze, Limit distributions for measures of multivariate skewness and kurtosis based on projections. J. Multivariate Anal., **38** (1991), 51–69.

[4] D. Cook, A. Buja and J. Cabrera, Projection pursuit indexes based on orthonormal function expansions. J. Comput. Graph. Statist., **2** (1993), 225–250.

[5] A. Erdélyi, W. Magnus, F. Oberthettinger and F. G. Tricomi, Higher Transcendental Functions **2**, McGraw-Hill, New York, 1953.

[6] K. T. Fang, S. Kots and K. W. Ng, Symmetric Multivariate and Related Distributions, Chapman & Hall, London/New York, 1989.

[7] J. H. Friedman, Exploratory projection pursuit, J. Amer. Statist. Assoc., **82** (1987), 249–266.

[8] J. H. Friedman and W. Stuetzle, Projection pursuit regression, J. Amer. Statist. Assoc., **76** (1981), 817–823.

[9] J. H. Friedman, W. Stuetzle and A. Schroeder, Projection pursuit density estimation, J. Amer. Statist. Assoc., **79** (1984), 599–608.

[10] P. Hall, Polynomial-based projection pursuit, Ann. Statist., **17** (1989), 589–605.

[11] P. J. Huber, Projection Pursuit (with discussion), Ann. Statist., **13** (1985), 435–475.

[12] M. Iwasaki, Projection pursuit by polynomial-based indices, Japanese J. Appl. Statist., **18** (1989), 103–128. (in Japanese)

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

[14] M. C. Jones and R. Sibson, What is projection pursuit? (with discussion), J. R. Statist. Soc., A **150** (1987), 1–36.

[15] A. A. Lubischew, On the use of discrininant functions in taxonomy, Biometrics, **18** (1962), 455–477.

[16] W. Magnus, F. Oberthettinger and R. P. Soni, Fomulas and Theorems for the Special Functions of Mathematical Physics, Springer, New York, 1966.

[17] R. G. Miller, Discussion of paper by P. Huber, Ann. Statist., **13** (1985), 510–513.

[18] K. Naito, On the maxima of moment index for exploratory projection pursuit, TR No. 96–7, Statistical Research Group, Hiroshima University, 1996.

[19] G. P. Nason, Three-dimensional projection pursuit, Appl. Statist., **44** (1995), 411–430.

[20] C. Posse, Projection pursuit exploratory data analysis, Comput. Statist. & Data Analysis, **20** (1995), 669–687.

[21] E. Stein and G. Weiss, Introduction to Fourier Analysis on Euclidean Spaces, Princeton Univ. Press, Princeton, NJ, 1971.

[22] J. Sun, Significance level in exploratory projection pursuit, Biometrika, **78** (1991), 759–769.

[23] J. Sun, Tail probabilities of the maxima of gaussian random fields, Ann. Prob., **21** (1993), 34–71.

[24] J. Sun, Some practical aspects of exploratory projection pursuit, *SIAM* J. Sci. Comput., **14** (1993), 68–80.

Department of Mathematics
Hiroshima University
Higashi-Hiroshima, 739, Japan