Statistica Sinica **8**(1998), 767-784

OPERATING TRANSFORMATION RETRANSFORMATION ON SPATIAL MEDIAN AND ANGLE TEST

Biman Chakraborty, Probal Chaudhuri and Hannu Oja*

*Indian Statistical Institute and * University of Oulu*

Abstract: An affine equivariant modification of the spatial median constructed using an adaptive transformation and retransformation procedure has been studied. It has been shown that this new estimate of multivariate location improves upon the performance of nonequivariant spatial median especially when there are correlations among the real valued components of multivariate data as well as when the scales of those components are different (e.g. when data points follow an elliptically symmetric distribution). For such correlated multivariate data the proposed estimate is more efficient than the non-equivariant vector of coordinatewise sample medians, and it outperforms the sample mean vector in the case of heavy tailed non-normal distributions. As an extension of the methodology, we have proposed an affine invariant modification of the well-known angle test based on the transformation approach, which is applicable to one sample multivariate location problems. We have observed that this affine invariant test performs better than the noninvariant angle test and the coordinatewise sign test for correlated multivariate data. Also, for heavy tailed non-normal multivariate distributions, the test outperforms Hotelling's T^2 test. Finite sample performance of the proposed estimate and the test is investigated using Monte Carlo simulations. Some data analytic examples are presented to demonstrate the implementation of the methodology in practice.

Key words and phrases: Affine transformation, efficiency, elliptically symmetric distribution, equivariant estimate, invariant test, multivariate median, multivariate sign test.

1. Introduction: Spatial Median, Angle Test and Their Lack of Equivariance and Invariance

For a set of n data points $\mathbf{X}_1, \ldots, \mathbf{X}_n$ in $\mathbb{R}^d$, there are many versions of multivariate median and multivariate sign test that have been proposed and investigated in the existing literature. Readers are referred to Small (1990), Hettmansperger, Nyblom and Oja (1992) and Chaudhuri and Sengupta (1993) for some recent detailed reviews. Among different versions of the multivariate median, the vector of coordinatewise median and the spatial median, which is defined as the vector $\hat{\Phi}_n$ that minimizes the sum $\sum_{i=1}^n \|\mathbf{X}_i - \Phi\|$ of Euclidean distances of $\Phi \in \mathbb{R}^d$ from the data points $\mathbf{X}_i$'s, are perhaps the simplest ones, and they appear to have the greatest antiquity in the literature. At the beginning of the 20th century, Hayford (1902) considered the possibility of using the vector of medians of orthogonal coordinates to locate the 'geographical center' of the US population. In view of the difficulty arising from the fact that such a concept of multivariate location depends on the choice of the orthogonal coordinate system to be laid on the map of the country and its lack of equivariance under orthogonal transformations, Scates (1933) proposed the use of the spatial median for the same purpose. Many other authors (see e.g. Gini and Galvani (1929) and Haldane (1948)) have independently considered the spatial median as a generalization of the univariate median in different situations. In one sample location problems, the tests that are naturally associated with the vector of coordinatewise median and the spatial median are the coordinatewise sign test and the angle test respectively. The former test has been studied extensively by Bennett (1962), Bickel (1965), Chatterjee (1966), Puri and Sen (1971), etc. while the latter one, which is a test based on the direction vectors $\mathbf{U}(\mathbf{X}_i) = \|\mathbf{X}_i\|^{-1}\mathbf{X}_i$ $(1 \leq i \leq n, \mathbf{X}_i \neq 0)$, has been considered by Brown (1983, 1985), Mottonen and Oja (1995), etc.

One serious drawback of the coordinatewise median as well as the spatial median is that neither of them is equivariant under arbitrary affine transformation of the data. In addition to being an undesirable geometric feature, this lack of equivariance is known to have some negative impact on the statistical performance of these two location estimates especially when the real valued components of the multivariate data are substantially correlated. This is also a handicap for both the coordinatewise sign test and the angle test none of which is invariant under general affine transformation of the data. This issue was first articulated by Bickel (1964, 1965), and subsequently investigated by Brown and Hettmansperger (1987, 1989). The spatial median is known to have rather impressive and somewhat counter-intuitive efficiency properties for multidimensional data generated from a spherically symmetric probability distribution, and this has been discussed in detail in Chaudhuri (1992) (see also Brown (1983)). However, the performance of the spatial median as well as the angle test tends to be quite poor compared to other affine equivariant or invariant procedures when there is significant deviation from spherical symmetry caused by the presence of correlation among observed variables (e.g. when the underlying distribution is elliptically symmetric). While the vector of the coordinatewise median is equivariant under coordinatewise scale transformation of the data, the lack of equivariance in the spatial median under such transformations makes it useless when in practice different variables are measured in different scales. Similarly, while the coordinatewise sign test can be used for data consisting of variables measured in different scales, it is not possible to use the angle test on such data due to its lack of invariance under coordinatewise scale transformation.

Recently Chakraborty and Chaudhuri (1996, 1998) have proposed a resolution for this problem of lack of affine equivariance and loss in statistical efficiency in the case of the coordinatewise median using a data based transformation and retransformation strategy. They converted the nonequivariant coordinatewise median into an affine equivariant location estimate and thereby repaired some of its undesirable features, which led to a gain in statistical efficiency. Their principal idea originated from the concept of a 'data driven coordinate system' introduced by Chaudhuri and Sengupta (1993) as an effective general tool for constructing affine invariant versions of the multivariate sign test. In this article, our objective is to explore an affine equivariant version of the multivariate median and an affine invariant version of the multivariate sign test, which are constructed by applying the transformation retransformation technique to the spatial median and the angle test respectively. It is appropriate to note here that there are other versions of the affine equivariant multivariate median (see e.g. Tukey (1975), Oja (1983), Liu (1990)) and affine invariant multivariate sign test (see e.g. Hodges (1955), Blumen (1958), Brown and Hettmansperger (1989), Brown, Hettmansperger, Nyblom and Oja (1992), Randles (1989), Oja and Nyblom (1989), Chaudhuri and Sengupta (1993), Hettmansperger, Nyblom and Oja (1994)) that have been proposed and extensively studied in the literature. However, we have been motivated to use the transformation and retransformation approach primarily by its appealing geometric interpretation and computational simplicity in the resulting statistical test and estimate as well as the elegant mathematical theory for their statistical properties. Our proposed procedures are quite easy to implement in analyzing data for making statistical inference in practice. We demonstrate in Section 3 how one can conveniently estimate the sampling variation in the proposed location estimate by resampling techniques such as the bootstrap. In the same Section, we will indicate how one can estimate the P-value when our test is applied to a data set by simulating the null distribution of the proposed test statistic. As we see in Section 2, a very encouraging common feature of both of our test and location estimates is that they inherit the impressive efficiency properties of the angle test and the spatial median in spherically symmetric multivariate models and can be extended to more general elliptically symmetric situations.

2. Transformation and Retransformation: Methodology and Motivation

Let us begin by observing a simple geometrical fact about any given affine transformation of a set of multivariate observations. For a nonsingular $d \times d$ matrix $\mathbf{A}$ and any $\mathbf{b} \in \mathbb{R}^d$, the transformation that maps $\mathbf{X}_i$ into $\mathbf{A}\mathbf{X}_i + \mathbf{b}$ for

$1 \leq i \leq n$ essentially expresses the original data in terms of a new coordinate system determined by $\mathbf{A}$ and $\mathbf{b}$. The new origin is located at $-\mathbf{A}^{-1}\mathbf{b}$, and depending on whether $\mathbf{A}$ is an orthogonal matrix or not, this new coordinate system may or may not be an orthonormal system. The fundamental idea that lies at the root of data based transformation and retransformation is to form an appropriate 'data driven coordinate system' (see also Chaudhuri and Sengupta (1993)) and to express all the data points in terms of that coordinate system first. This is tantamount to making an affine transformation of the data. Then one computes a location estimate or a test statistic based on those transformed data points. Finally the location estimate is retransformed to express everything back in terms of the original coordinate system (see also Chakraborty and Chaudhuri (1996, 1998)). Now, in order to form a 'data driven coordinate system', we need $d+1$ data points in $\mathbb{R}^d$, one of which determines the origin, and the lines joining that origin to the remaining d data points form various coordinate axes. In order to get a valid coordinate system, these $d+1$ points must satisfy some 'non-singularity' or 'affine independence' condition that we will gradually consider. However, it is not necessary for this 'data driven coordinate system' to be an orthonormal system. We now discuss in detail and in more precise terms how this transformation and retransformation technique converts spatial median into an affine equivariant estimate of multivariate location.

2.1. An affine equivariant version of multivariate median

Suppose that we have n data points $\mathbf{X}_1, \ldots, \mathbf{X}_n$ in $\mathbb{R}^d$, and assume that $n > d+1$. Let $\alpha = \{i_0, i_1, \ldots, i_d\}$ denote a subset of size $d+1$ of $\{1, \ldots, n\}$. Consider the points $\mathbf{X}_{i_0}, \mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_d}$, which form a 'data driven coordinate system' as described above, and the $d \times d$ matrix $\mathbf{X}(\alpha)$ containing the columns $\mathbf{X}_{i_1} - \mathbf{X}_{i_0}, \ldots, \mathbf{X}_{i_d} - \mathbf{X}_{i_0}$ can be taken as the transformation matrix for transforming the data points $\mathbf{X}_j$ such that $1 \leq j \leq n$, $j \notin \alpha$ in order to express them in terms of the new coordinate system as $\mathbf{Z}_j^{(\alpha)} = \{\mathbf{X}(\alpha)\}^{-1}\mathbf{X}_j$. From now on, all vectors in this paper are assumed to be column vectors unless specified otherwise, and the superscript T is used to denote transpose of vectors and matrices. Clearly, we need $\mathbf{X}(\alpha)$ to be an invertible matrix, and this is guaranteed if the $\mathbf{X}_i$'s are generated as i.i.d. observations from a distribution that is absolutely continuous w.r.t the Lebesgue measure on $\mathbb{R}^d$. One can now compute the spatial median $\hat{\Phi}_n^{(\alpha)}$ of the $\mathbf{Z}_j^{(\alpha)}$'s by minimizing the sum $\sum_{j \notin \alpha} \|\mathbf{Z}_j^{(\alpha)} - \Phi\|$. Finally, in order to express things back in terms of the original coordinate system, we need to retransform $\hat{\Phi}_n^{(\alpha)}$ into $\hat{\Theta}_n^{(\alpha)} = \mathbf{X}(\alpha)\hat{\Phi}_n^{(\alpha)}$, which is our desired location estimate. In view of its construction, it is clear that $\hat{\Theta}_n^{(\alpha)}$ is an affine equivariant estimate of location. A question that naturally arises at this point is how to choose the 'data driven coordinate system' or equivalently the data based transformation matrix $\mathbf{X}(\alpha)$. An answer to this question is provided in the following Theorem, which we prove in the Appendix.

Since in the univariate case there is a unique concept of the median and the sign test both of which are already extensively studied in the literature, for the rest of the paper we assume that $d \geq 2$ and write $g(\mathbf{x})$ to denote the elliptically symmetric density $\{\det(\Sigma)\}^{-1/2} f(\mathbf{x}^T \Sigma^{-1} \mathbf{x})$, where Σ is a $d \times d$ positive definite matrix and $f(\mathbf{x}^T \mathbf{x})$ is a continuous spherically symmetric density around the origin in $\mathbb{R}^d$. The $\mathbf{X}_i$'s are assumed to be i.i.d. observations with common elliptically symmetric density $g(\mathbf{x} - \Theta)$, where $\Theta \in \mathbb{R}^d$ is the location of elliptic symmetry for the data.

Theorem 2.1. *For any given subset α of $\{1, \ldots, n\}$ with size $d+1$ and given the $\mathbf{X}_i$'s with $i \in \alpha$, the conditional asymptotic distribution of $n^{1/2}(\hat{\Theta}_n^{(\alpha)} - \Theta)$ is d-variate normal with zero mean and a variance covariance matrix $\Delta\{f, \Sigma, \mathbf{X}(\alpha)\}$ that depends on f, Σ and the transformation matrix $\mathbf{X}(\alpha)$. Here the positive definite matrix Δ is such that the difference $\Delta\{f, \Sigma, \mathbf{A}\} - \Delta\{f, \Sigma, \mathbf{B}\}$ is non-negative definite (i.e. $\Delta\{f, \Sigma, \mathbf{A}\} \geq_{n.n.d} \Delta\{f, \Sigma, \mathbf{B}\}$) for any f, Σ and any two $d \times d$ invertible matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{B}^T \Sigma^{-1} \mathbf{B} = \lambda \mathbf{I}_d$, where $\lambda > 0$ is a constant and $\mathbf{I}_d$ is the $d \times d$ identity matrix. Further, for any such $\mathbf{B}$, we have $\Delta\{f, \Sigma, \mathbf{B}\} = c(d, f)\Sigma$, where $c(d, f) = \pi^{-1} d(d-1)^{-2}\{f_1(0)\}^{-2}[\Gamma\{(d-1)/2\}]^{-2}\{\Gamma(d/2)\}^2$, f_1 being the univariate marginal of the spherically symmetric density f on $\mathbb{R}^d$.*

The main message communicated by the above Theorem is that we need to choose $\mathbf{X}(\alpha)$ in such a way that $\{\mathbf{X}(\alpha)\}^T \Sigma^{-1} \mathbf{X}(\alpha)$ becomes as close as possible to a matrix of the form $\lambda \mathbf{I}_d$, which is a diagonal matrix with all diagonal entries equal. In other words, the coordinate system represented by the transformation matrix $\Sigma^{-1/2}\mathbf{X}(\alpha)$ should be as orthonormal in nature as possible. The expression for $c(d, f)$ in Theorem 2.1 implies that when $\{\mathbf{X}(\alpha)\}^T \Sigma^{-1} \mathbf{X}(\alpha)$ is chosen to be close to a diagonal matrix with all diagonal entries equal, the asymptotic efficiency of the estimate $\hat{\Theta}_n^{(\alpha)}$ becomes close to that of the spatial median under spherically symmetric models (i.e. when $\Sigma = \sigma^2 \mathbf{I}_d$), and it will be more efficient than the spatial median in elliptically symmetric models (see Chaudhuri (1992)). It is known that for spherically symmetric data the rotationally equivariant spatial median is more efficient than the vector of coordinatewise medians, which lacks rotational equivariance (see Brown (1983), Chaudhuri (1992)), and with a proper selection of $\mathbf{X}(\alpha)$, $\hat{\Theta}_n^{(\alpha)}$ too will have similar superior performance. Another implication of Theorem 2.1 is that with appropriate choice of $\mathbf{X}(\alpha)$, the estimate $\hat{\Theta}_n^{(\alpha)}$ will be more (or less) efficient than the sample mean vector depending on whether the tail of the density f is 'heavy' (or 'light').

We discuss a procedure for choosing the transformation matrix $\mathbf{X}(\alpha)$ from the data when we discuss numerical examples in Section 3, and there we compare the finite sample performance of our estimate with that of some other well-known estimates of multivariate location. But before that let us close this Section by noting that an alternative affine equivariant modification of the spatial median has been considered in the literature by other authors (see e.g. Isogai (1985), Rao (1988)), who computed the spatial median based on multivariate observations transformed by the square root of the variance covariance matrix. While transforming the data points by the square root of the sample variance covariance matrix is a popular approach, the resulting co-ordinate system does not have any simple and natural geometric interpretation. Further, such a transformation cannot lead to an affine invariant modification of other non-equivariant location estimates such as the coordinatewise median, as has been discussed in Chakraborty and Chaudhuri (1996), and the limitation of that approach is primarily due to the fact that there does not exist an affine equivariant square root of the variance covariance matrix. As we see in the next section, the strategy of transforming the data points using an appropriately chosen $\mathbf{X}(\alpha)$ leads to an affine invariant modification of the well-known angle test, which turns out to be 'distribution free' in nature in the sense that the null distribution of the test statistic under elliptically symmetric model does not depend on the unknown density f. This is not achievable by transforming the observations using the square root of the sample variance covariance matrix. Our 'data driven coordinate system' is a widely applicable tool for converting non-equivariant (or non-invariant) procedures into equivariant (or invariant) procedures, which is not limited to only the spatial median. Besides, it has a very nice and intuitively meaningful geometric interpretation, and an attractive feature of this data based transformation retransformation strategy is the clean and elegant mathematical theory associated with the approach, which provides an effective guideline for implementation of the methodology for analyzing data and making statistical inference.

2.2. An affine invariant multivariate sign test

As in the preceding Section, let us assume that the $\mathbf{X}_i$'s are i.i.d. observations generated from the elliptically symmetric density $g(\mathbf{x} - \Theta)$ on $\mathbb{R}^d$. Suppose that we have two competing hypotheses $H_0 : \Theta = 0$ and $H_A : \Theta \neq 0$ concerning the center of elliptic symmetry of the distribution. Consider once again the transformed observations $\mathbf{Z}_j^{(\alpha)}$ for $j \notin \alpha$ and $1 \leq j \leq n$, and define the test statistic $\mathbf{T}_n^{(\alpha)} = \sum_{j \notin \alpha} \|\mathbf{Z}_j^{(\alpha)}\|^{-1} \mathbf{Z}_j^{(\alpha)}$. We now state a Theorem that summarizes the main features of this test statistic.

Theorem 2.2. *Under the null hypothesis $H_0 : \Theta = 0$, the conditional distribution of $\mathbf{T}_n^{(\alpha)}$ given the $\mathbf{X}_i$'s with $i \in \alpha$ does not depend on f, and it depends on Σ*

through $\Sigma^{-1/2}\mathbf{X}(\alpha)$. Further, in large samples, the conditional null distribution of $n^{-1/2}\mathbf{T}_n^{(\alpha)}$ is approximately normal with zero mean and a variance covariance matrix $\Psi\{\Sigma^{-1/2}\mathbf{X}(\alpha)\}$ that depends on $\Sigma^{-1/2}\mathbf{X}(\alpha)$. When $\log f$ is twice differentiable almost everywhere (w.r.t. Lebesgue measure) on $\mathbb{R}^d$ and satisfies the Cramer type regularity conditions, the alternatives $H_A^{(n)} : \Theta = n^{-1/2}\Phi$ such that $\Phi \in \mathbb{R}^d$ and $\Phi \neq 0$ form a contiguous sequence, and the conditional limiting distribution of $n^{-1/2}\mathbf{T}_n^{(\alpha)}$ under that sequence of alternatives is normal with the same variance covariance matrix Ψ and a mean vector $\Lambda\{f, \Sigma^{-1/2}\Phi, \Sigma^{-1/2}\mathbf{X}(\alpha)\}$ that depends on f, $\Sigma^{-1/2}\Phi$ and $\Sigma^{-1/2}\mathbf{X}(\alpha)$. Also, the limiting conditional power of the test under such a sequence of contiguous alternatives depends monotonically on the noncentrality parameter

$$\begin{aligned}&\delta\{f, \Sigma^{-1/2}\Phi, \Sigma^{-1/2}\mathbf{X}(\alpha)\}\\ &= [\Lambda\{f, \Sigma^{-1/2}\Phi, \Sigma^{-1/2}\mathbf{X}(\alpha)\}]^T[\Psi\{\Sigma^{-1/2}\mathbf{X}(\alpha)\}]^{-1}\Lambda\{f, \Sigma^{-1/2}\Phi, \Sigma^{-1/2}\mathbf{X}(\alpha)\},\end{aligned}$$

where δ is such that for any f, Φ, Σ and any two invertible matrices $\mathbf{A}$ and $\mathbf{B}$, we have $\delta\{f, \Sigma^{-1/2}\Phi, \Sigma^{-1/2}\mathbf{B}\} \geq \delta\{f, \Sigma^{-1/2}\Phi, \Sigma^{-1/2}\mathbf{A}\}$ whenever $\mathbf{B}^T\Sigma^{-1}\mathbf{B} = \lambda\mathbf{I}_d$ for some $\lambda > 0$.

We prove this Theorem in the Appendix. But before that note that one of the main implications of this Theorem is that whatever f may be, it is possible to simulate the conditional finite sample null distribution of $\mathbf{T}_n^{(\alpha)}$ after obtaining an appropriate estimate of Σ in small sample situations, and one can use the normal approximation when the sample size is large. It is also noteworthy that in order to maximize the power of the test that rejects H_0 for large values of $\{\mathbf{T}_n^{(\alpha)}\}^T[\Psi\{\Sigma^{-1/2}\mathbf{X}(\alpha)\}]^{-1}\mathbf{T}_n^{(\alpha)}$, one needs to choose $\mathbf{X}(\alpha)$ in such a way that $\{\mathbf{X}(\alpha)\}^T\Sigma^{-1}\mathbf{X}(\alpha)$ becomes as close as possible to a diagonal matrix with all diagonal entries equal (especially for alternatives close to the null). It will become transparent when we present the proof of Theorem 2.2 that by choosing $[\mathbf{X}(\alpha)]^T\Sigma^{-1}\mathbf{X}(\alpha)$ very close to a matrix of the form $\lambda\mathbf{I}_d$, asymptotic Pitman efficiency of the test can be made close to that of the angle test in the spherically symmetric model (i.e. when $\Sigma = \sigma^2\mathbf{I}_d$), and it will be more efficient than the angle test in elliptically symmetric models. Also, the test will be more (or less) efficient than the standard test of location based on Hotelling's T^2 statistic (which is a test based on the sample mean vector) if the tail of the density f is 'heavy' (or 'light'). In the following Section, we demonstrate how the simulated conditional null distribution of $\mathbf{T}_n^{(\alpha)}$ can be used to determine the critical region of the test for a pre-specified level of significance and also to estimate the P-value corresponding to a given value of $\{\mathbf{T}_n^{(\alpha)}\}^T[\Psi\{\Sigma^{-1/2}\mathbf{X}(\alpha)\}]^{-1}\mathbf{T}_n^{(\alpha)}$ computed from the observed data. Also, there we compare the finite sample performance of our test with some other standard one sample tests for multivariate location.

3. Simulation Studies and Data Analysis

It is quite clear from our main results and discussion in the preceding Section that we need to choose $\mathbf{X}(\alpha)$ in such a way that $\{\mathbf{X}(\alpha)\}^T\Sigma^{-1}\mathbf{X}(\alpha)$ becomes as close as possible to a matrix of the form $\lambda\mathbf{I}_d$. Since Σ will be unknown in practice, we have to estimate that from the data and we will need a consistent and affine equivariant estimate (say $\hat{\Sigma}$). When the variables observed in the data have finite population variances, we can use the usual variance covariance matrix for this purpose. In any case, after obtaining $\hat{\Sigma}$, we try to choose $\mathbf{X}(\alpha)$ in such a way that the eigenvalues of the positive definite matrix $\{\mathbf{X}(\alpha)\}^T\Sigma^{-1}\mathbf{X}(\alpha)$ become as equal as possible. To achieve this, our strategy is to minimize either the ratio between the arithmetic mean and the geometric mean or that between the geometric mean and the harmonic mean of the eigenvalues. Note that a major advantage in using such a criterion is that it does not involve explicit computation of the eigenvalues of the matrix. Arithmetic and harmonic means of the eigenvalues can be obtained from the trace of the matrix and that of its inverse respectively, while the geometric mean can be computed from its determinant. In our numerical studies, we have observed that the criteria based on different ratios yield more or less similar results. Instead of minimizing the ratio over all possible subsets α with size $d+1$ of $\{1,\ldots,n\}$, one can substantially reduce the amount of computation by stopping the search for optimal $\mathbf{X}(\alpha)$ as soon as the ratio becomes smaller than $1+\epsilon$, where ϵ is a preassigned small positive number. In our simulations and data analysis, we did not observe such an approach to cause any significant change in the statistical performance of the procedures though there was considerable gain in the speed of computation. Of course, there are other different ways to achieve this goal of making $\{\mathbf{X}(\alpha)\}^T\Sigma^{-1}\mathbf{X}(\alpha)$ as close as possible to a diagonal matrix with all diagonal entries equal. We have adopted a specific strategy that is computationally convenient and has been observed to work fairly well in our numerical investigations. Note that once $\mathbf{X}(\alpha)$ is chosen, we can compute the spatial median $\hat{\Phi}_n^{(\alpha)}$ from the transformed observations $\mathbf{Z}_j^{(\alpha)}$'s using any of the standard algorithms discussed in the literature (see e.g. Gower (1974), Chaudhuri (1996))

We now discuss a simulation study that was undertaken with the objective of comparing the finite sample performance of $\hat{\Theta}_n^{(\alpha)}$ with that of the sample mean vector and the vector of coordinatewise sample medians. We have used sample size $n = 30$, considered the cases $d = 2$ and 3 and generated data from three different distributions, namely multivariate normal, multivariate Laplace (i.e. when $f(\mathbf{x}^T\mathbf{x}) = k\exp\{-(\mathbf{x}^T\mathbf{x})^{1/2}\}$) and multivariate t with 3 degrees of freedom. Keeping in mind location equivariance as well as equivariance under coordinatewise scale transformation of each of the three multivariate location estimates considered, we decided to generate data from the elliptically symmetric density $g(\mathbf{x} - \Theta)$, where Θ was taken to be the zero vector and Σ was taken to be the matrix with each diagonal entry equal to one and each off diagonal entry equal to ρ. The value of ρ was chosen from the interval $[0, 1)$. We denote by e_1 and e_2 the efficiencies of $\hat{\Theta}_n^{(\alpha)}$ compared with the sample mean vector and the vector of coordinatewise sample medians respectively. For two competing estimates $\hat{\Phi}_1$ and $\hat{\Phi}_2$ of a d-dimensional location parameter Φ, we define the efficiency of the former estimate compared with the latter one as the dth root of the ratio between $\det\{E(\hat{\Phi}_2 - \Phi)(\hat{\Phi}_2 - \Phi)^T\}$ and $\det\{E(\hat{\Phi}_1 - \Phi)(\hat{\Phi}_2 - \Phi)^T\}$ (see e.g. Bickel (1964), Chakraborty and Chaudhuri (1998)). The results are reported in Tables 3.1 and 3.2. In each case, we have estimated the efficiencies e_1 and e_2 based on 10,000 Monte Carlo replications for $d = 2$ and using 5,000 Monte Carlo replications for $d = 3$. Since both of $\hat{\Theta}_n^{(\alpha)}$ and sample mean vector are affine equivariant estimates, the value of e_1 remains constant for different values of ρ. The superior performance of $\hat{\Theta}_n^{(\alpha)}$ for non-normal elliptically symmetric distributions (especially when ρ is large) is quite apparent in the results given in Tables 3.1 and 3.2.

Table 3.1. Finite sample efficiency of affine equivariant modification of the spatial median for $n = 30$ and $d = 2$

Distribution		ρ					
		0.00	0.75	0.80	0.85	0.90	0.95
Normal	e_1	0.7153	0.7153	0.7153	0.7153	0.7153	0.7153
	e_2	1.1313	1.4418	1.5243	1.6447	1.8285	2.1747
Laplace	e_1	1.2849	1.2849	1.2849	1.2849	1.2849	1.2849
	e_2	1.0861	1.3779	1.4655	1.5877	1.7688	2.1172
t with 3 d.f.	e_1	1.7676	1.7676	1.7676	1.7676	1.7676	1.7676
	e_2	1.0628	1.3551	1.4379	1.5512	1.7291	2.0769

Table 3.2. Finite sample efficiency of affine equivariant modification of the spatial median for $n = 30$ and $d = 3$.

Distribution		ρ					
		0.00	0.75	0.80	0.85	0.90	0.95
Normal	e_1	0.7319	0.7319	0.7319	0.7319	0.7319	0.7319
	e_2	1.1649	1.5883	1.7140	1.8725	2.1219	2.6873
Laplace	e_1	1.1023	1.1023	1.1023	1.1023	1.1023	1.1023
	e_2	1.1701	1.6078	1.7271	1.9041	2.1757	2.7461
t with 3 d.f.	e_1	1.6725	1.6725	1.6725	1.6725	1.6725	1.6725
	e_2	1.1395	1.5725	1.6830	1.8538	2.1097	2.6413

Let us next consider two real data sets and try to investigate the resulting performance of $\hat{\Theta}_n^{(\alpha)}$. One of the primary reasons for using the transformation retransformation technique is that once the optimal data based transformation matrix $\mathbf{X}(\alpha)$ is chosen, it is quite easy to compute $\hat{\Theta}_n^{(\alpha)}$ as it requires only the computation of the spatial median based on the transformed observations $\mathbf{Z}_j^{(\alpha)}$'s. An important consequence of this is that one can conveniently use resampling techniques such as the bootstrap (see e.g. Efron (1982)) to estimate the conditional sampling variation of $\hat{\Theta}_n^{(\alpha)}$ given the $\mathbf{X}_i$'s with $i \in \alpha$ (i.e. after $\mathbf{X}(\alpha)$ is fixed). In each of the two examples discussed below, we have used 10,000 bootstrap replications to estimate the sampling variation and the efficiency of our transformation retransformation estimate, and it took only a few seconds on a workstation equipped with a standard FORTRAN compiler. It will be appropriate to note here that it is exceedingly difficult to estimate sampling variations of many of the other affine equivariant versions of the multivariate median proposed in the literature (e.g. Tukey (1975), Oja (1983) and Liu (1990)). Due to complex computational problems associated with those versions of the median in high or even moderately high dimensions, it becomes virtually impossible to apply the bootstrap or any other resampling techniques on them (see also Chakraborty and Chaudhuri (1998)).

Table 3.3. Transformation retransformation estimates and the results of bootstrap analysis of Fisher's Iris data

Species	Estimates and estimated (RMSE's)				Estimated efficiency
	Sepal length	Sepal width	Petal length	Petal width	
Setosa	5.0148	3.4180	1.4684	0.2376	$e_1^* = 1.0308$
	(0.0488)	(0.0648)	(0.0221)	(0.0137)	$e_2^* = 1.2482$
Versicolor	5.9111	2.8001	4.2733	1.3256	$e_1^* = 0.6607$
	(0.1178)	(0.0656)	(0.0961)	(0.0422)	$e_2^* = 2.4361$
Virginica	6.5421	2.9864	5.4953	2.0428	$e_1^* = 0.7494$
	(0.0926)	(0.0516)	(0.0802)	(0.0514)	$e_2^* = 1.9220$

Example 3.1. This example deals with the well-known Iris data analyzed by R. A. Fisher and many other famous statisticians. In the data set, there are three different species, namely Iris Setosa, Iris Versicolor and Iris Virginica, and each data point consists of four measurements, namely sepal length, sepal width, petal length and petal width. There are fifty observations for each species. Table 3.3 gives our location estimate and its root mean squared error (RMSE) as estimated by the bootstrap for each variable separately for different species. We have denoted by e_1^* and e_2^* the bootstrap estimates of the efficiencies of our affine equivariant modification of spatial median as compared with the sample mean

vector and the vector of coordinatewise sample medians respectively. It is interesting to note that while there is a gain in efficiency when compared with the non-equivariant vector of coordinatewise median in all three species, when compared with the affine equivariant sample mean, there is gain only in the case of Iris Setosa, and there is a definite loss in efficiency in each of the other two cases. The entire analysis seems to make a very good case for using affine equivariant procedures.

Table 3.4. Transformation retransformation estimates and the results of bootstrap analysis of urine data

Estimates and estimated (RMSE's)				Estimated
Specific Gravity	pH	Conductivity	Osmolarity	efficiency
1.025	7.90	721.0	23.6	$e_1^* = 1.0921$
(0.0016)	(0.1201)	(54.5453)	(1.6232)	$e_2^* = 2.8250$

Example 3.2. The data set used in this example was originally obtained from the laboratory of James S. Elliot, M.D. of the Urology Section, Veteran's Administration Medical Center, Palo Alto, California and the Division of Urology, Stanford University School of Medicine, Stanford, California, and it is reported in Andrews and Herzberg (1985). We have considered four physical characteristics of thirty three urine specimens with calcium oxalate crystals. These characteristics are : specific gravity (i.e. the density of urine relative to water), pH (i.e. the negative logarithm of hydrogen ion concentration), osmolarity (which is proportional to the concentration of molecules in the solution) and conductivity (which is proportional to the concentration of charged ions in the solution). As one would expect, the correlations among these variables are fairly high and the variables are measured in widely different scales. Table 3.4 summarizes the results of the bootstrap analysis of this data set. The values of e_1^* and e_2^* indicate considerable gain in efficiency over the non-equivariant vector of coordinatewise medians and a small gain over the sample mean vector.

At this point we turn our attention to the affine invariant test introduced and discussed in Section 2.2. It will be clear from the proof of Theorem 2.2 that once the transformation matrix $\mathbf{X}(\alpha)$ is fixed, the conditional null distribution of $\mathbf{T}_n^{(\alpha)}$ is the same as that of $\sum_{j=1}^{(n-d-1)} \|\mathbf{Y}_j\|^{-1}\mathbf{Y}_j$, where the $\mathbf{Y}_j$'s are i.i.d. observations generated from the elliptically symmetric density $\det\{\mathbf{Y}(\alpha)\}f[\mathbf{y}^T\{\mathbf{Y}(\alpha)\}^T\mathbf{Y}(\alpha)\mathbf{y}]$ and $\mathbf{Y}(\alpha) = \Sigma^{-1/2}\mathbf{X}(\alpha)$. Further, elliptic symmetry implies that the distribution of $\|\mathbf{Y}_j\|^{-1}\mathbf{Y}_j$ is uniform on the ellipse, which is completely determined by the matrix $\{\mathbf{Y}(\alpha)\}^T\mathbf{Y}(\alpha) = \{\mathbf{X}(\alpha)\}^T\Sigma^{-1}\mathbf{X}(\alpha)$ and does not depend on f. Hence one can simulate the conditional null distribution of $\mathbf{T}_n^{(\alpha)}$ by taking f to be any specific spherically symmetric density (e.g. the normal density) on $\mathbb{R}^d$. Of course, the actual Σ will be unknown in practice, and one can use a consistent affine invariant estimate $\hat{\Sigma}$ while simulating the null distribution. In the following example, we demonstrate simulation based estimation of the P-value when our test is applied to a real data set.

Example 3.3. Merchants, Halprin, Hudson, Kilburn, McKenzie, Hurst and Bermazohn (1975) investigated changes in pulmonary functions of twelve workers after they were exposed to cotton dust for six hours. Table 3.5 gives the changes in forced vital capacity (FVC), forced expiratory volume (FEV_3) and closing capacity (CC) for these twelve workers. When Hotelling's T^2 test is applied to this data, the P-value computed using the F distribution turns out to be 0.051.

Table 3.5. Changes in pulmonary functions of twelve workers exposed to cotton dust for six hours.

Subject	FVC	FEV_3	CC
1	−0.11	−0.12	−4.3
2	0.02	0.08	4.4
3	−0.02	0.03	7.5
4	0.07	0.19	−0.3
5	−0.16	−0.36	−5.8
6	−0.42	−0.49	14.5
7	−0.32	−0.48	−1.9
8	−0.35	−0.30	17.3
9	−0.10	−0.04	2.5
10	0.01	−0.02	−5.6
11	−0.01	−0.17	2.2
12	−0.26	−0.30	5.5

On the other hand the coordinatewise sign test yields a P-value of 0.300 based on a χ^2 approximation with 3 d.f. We estimated the P-value of our test based on a simulation of the conditional null distribution of $\mathbf{T}_n^{(\alpha)}$ using 10,000 Monte Carlo replications, and it turns out to be 0.0721. For simulating the null distribution, we have chosen f to be the multivariate spherically symmetric normal density and estimated Σ by the usual variance covariance matrix. Figures in Table 3.5 indicate presence of correlation among the variables, and the scale of the third variable is very different from that of each of the other two. The close agreement between the P-values obtained from two affine invariant tests is noteworthy, and the high P-value produced by the non-invariant coordinatewise sign test is an indication of its failure to detect the deviation from the null hypothesis.

Table 3.6. Finite sample power of affine invariant modification of angle test and its competitors for $n = 30$, $d = 2$ and nominal level of significance = 5%.

Distribution	Test statistic	ρ	$\delta = (\Theta^T \Sigma^{-1} \Theta)^{1/2}$ 0.0	0.3	0.6	0.9	1.2	1.5
Normal	$\mathbf{T}_n^{(\alpha)}$	–	0.049	0.207	0.679	0.954	0.998	1.000
	T^2	–	0.049	0.262	0.809	0.992	1.000	1.000
	S_n	0.00	0.047	0.185	0.620	0.928	0.997	1.000
		0.75	0.043	0.192	0.652	0.938	0.986	0.968
		0.85	0.043	0.197	0.655	0.928	0.954	0.893
		0.95	0.033	0.183	0.614	0.834	0.798	0.661
Laplace	$\mathbf{T}_n^{(\alpha)}$	–	0.050	0.149	0.421	0.740	0.923	0.982
	T^2	–	0.045	0.124	0.363	0.688	0.908	0.981
	S_n	0.00	0.046	0.129	0.387	0.696	0.887	0.969
		0.75	0.047	0.134	0.403	0.726	0.912	0.970
		0.85	0.047	0.138	0.403	0.722	0.894	0.947
		0.95	0.036	0.123	0.379	0.665	0.795	0.812
t with 3 d.f.	$\mathbf{T}_n^{(\alpha)}$	–	0.051	0.178	0.553	0.883	0.988	0.999
	T^2	–	0.041	0.151	0.478	0.798	0.936	0.979
	S_n	0.00	0.048	0.160	0.509	0.844	0.974	0.997
		0.75	0.042	0.169	0.547	0.873	0.977	0.981
		0.85	0.044	0.174	0.546	0.866	0.956	0.935
		0.95	0.033	0.151	0.513	0.778	0.823	0.752

We close this Section with the results of a simulation study that was carried out to compare the finite sample powers of our affine invariant test, which is based on the statistic $\mathbf{T}_n^{(\alpha)}$ with that of the well-known Hotelling's T^2 test and the non-invariant sign test, which is based on the coordinatewise sign test statistic denoted by S_n. We have used sample size $n = 30$, and for nominal level 5%, we estimated the power in each case from 5,000 Monte Carlo replications for $d = 2$ and from 3,000 Monte Carlo replications for $d = 3$. Note that the values of the standard deviation of the sample proportion based on 5,000 i.i.d. Bernoulli trials with $p = 0.05$ (i.e. the nominal level), 0.15, 0.30 and 0.50 are 0.003, 0.005, 0.006 and 0.007 respectively, and these values based on 3,000 i.i.d. Bernoulli trials with the same values of p are 0.004, 0.006, 0.008 and 0.009 respectively. Different distributions, which were used in the simulation study, were chosen exactly in the same way as in the simulation study reported at the beginning of this Section, where we compared the performance of various multivariate location estimates. The results are presented in Tables 3.6 and 3.7. For Hotelling's T^2 test the critical value at nominal 5% level was determined from the F distribution table, and for the coordinatewise sign test, we used a χ^2 approximation (with 2 d.f. and 3 d.f. for $d = 2$ and 3 respectively) for the distribution of the test statistic. In the case of $\mathbf{T}_n^{(\alpha)}$, we chose to simulate its distribution as we have done in Example 3.3, and for this purpose we used the spherically symmetric normal density and 8,000 Monte Carlo replications in each case. The matrix Σ was estimated using the usual variance covariance matrix. Since both of our proposed test and Hotelling's T^2 test are affine invariant in nature, their powers do not depend on ρ and depend only on the noncentrality parameter $\delta = (\Theta^T\Sigma^{-1}\Theta)^{1/2}$. Figures in Tables 3.6 and 3.7 convincingly demonstrate superior performance of our affine invariant modification of the angle test in non-normal elliptically symmetric distributions especially for large values of ρ.

Table 3.7. Finite sample power of affine invariant modification of angle test and its competitors for $n = 30$, $d = 3$ and nominal level of significance = 5%.

Distribution	Test statistic	ρ	$\delta = (\Theta^T\Sigma^{-1}\Theta)^{1/2}$ 0.0	0.3	0.6	0.9	1.2	1.5
Normal	$\mathbf{T}_n^{(\alpha)}$	–	0.051	0.174	0.604	0.928	0.994	1.000
	T^2	–	0.050	0.215	0.723	0.977	1.000	1.000
	S_n	0.00	0.047	0.144	0.512	0.879	0.986	1.000
		0.75	0.039	0.148	0.557	0.901	0.968	0.917
		0.85	0.033	0.149	0.549	0.869	0.883	0.759
		0.95	0.019	0.118	0.447	0.658	0.569	0.364
Laplace	$\mathbf{T}_n^{(\alpha)}$	–	0.051	0.104	0.258	0.497	0.757	0.916
	T^2	–	0.039	0.085	0.235	0.494	0.741	0.915
	S_n	0.00	0.043	0.080	0.213	0.417	0.655	0.844
		0.75	0.033	0.077	0.225	0.454	0.705	0.881
		0.85	0.029	0.071	0.216	0.448	0.693	0.852
		0.95	0.018	0.049	0.174	0.364	0.545	0.639
t with 3 d.f.	$\mathbf{T}_n^{(\alpha)}$	–	0.051	0.158	0.493	0.826	0.965	0.995
	T^2	–	0.037	0.130	0.420	0.748	0.916	0.972
	S_n	0.00	0.038	0.123	0.415	0.762	0.938	0.989
		0.75	0.037	0.129	0.459	0.806	0.943	0.947
		0.85	0.031	0.128	0.459	0.787	0.883	0.843
		0.95	0.018	0.102	0.376	0.617	0.615	0.493

Acknowledgement

Research of the first two authors has been partially supported by a grant from Indian Statistical Institute. Research of the third author is partially supported by a grant from the Academy of Finland. The authors are thankful to an anonymous referee for a careful reading of an earlier version of the paper and several helpful comments.

Appendix : The Proofs

Proof of Theorem 2.1. First observe that in view of affine equivariance of $\hat{\Theta}_n^{(\alpha)}$, it is enough to consider the case when $\Theta = 0$ and $\Sigma = \mathbf{I}_d$. Then $g(\mathbf{x} - \Theta)$

reduces to the spherically symmetric density $f(\mathbf{x}^T\mathbf{x})$. Now, for a given subset α with size $d+1$ of $\{1,\ldots,n\}$ and given the $\mathbf{X}_i$'s for which $i \in \alpha$, the transformed observations $\mathbf{Z}_j^{(\alpha)}$'s are conditionally independent, and they are identically distributed with common elliptically symmetric density $h\{\mathbf{z}|\mathbf{X}(\alpha)\} = \det\{\mathbf{X}(\alpha)\}f[\mathbf{z}^T\{\mathbf{X}(\alpha)\}^T\mathbf{X}(\alpha)\mathbf{z}]$. Now for a random vector $\mathbf{Z}$ with density h, elliptic symmetry around the origin implies that the distribution of $\|\mathbf{Z}\|^{-1}\mathbf{Z}$ does not depend on f but on $\mathbf{X}(\alpha)$. Consider now the matrices $\mathbf{C}\{\mathbf{X}(\alpha)\} = E_h(\|\mathbf{Z}\|^{-2}\mathbf{Z}\mathbf{Z}^T)$ and $\mathbf{D}\{f,\mathbf{X}(\alpha)\} = E_h\{\|\mathbf{Z}\|^{-1}(\mathbf{I}_d - \|\mathbf{Z}\|^{-2}\mathbf{Z}\mathbf{Z}^T)\}$. Then it follows from Chaudhuri (1992) that given $\mathbf{X}(\alpha)$, the conditional limiting distribution of $n^{1/2}\hat{\Phi}_n^{(\alpha)}$, where $\hat{\Phi}_n^{(\alpha)}$ is the spatial median based on the transformed observations $\mathbf{Z}_j^{(\alpha)}$'s, is normal with zero mean and $[\mathbf{D}\{f,\mathbf{X}(\alpha)\}]^{-1}\mathbf{C}\{\mathbf{X}(\alpha)\}[\mathbf{D}\{f,\mathbf{X}(\alpha)\}]^{-1}$ as the variance covariance matrix. Further, elliptic symmetry of h around the origin implies that $\mathbf{D}\{f,\mathbf{X}(\alpha)\} = \mu(d,f)\mathbf{G}\{\mathbf{X}(\alpha)\}$, where μ is a positive constant depending on dimension d and f, and $\mathbf{G}$ is a positive definite symmetric matrix depending on $\mathbf{X}(\alpha)$ only. Finally, since $\hat{\Theta}_n^{(\alpha)} = \mathbf{X}(\alpha)\hat{\Phi}_n^{(\alpha)}$, the conditional limiting distribution of $n^{1/2}\hat{\Theta}_n^{(\alpha)}$ must be normal with variance covariance matrix

$$\begin{aligned}&\mathbf{X}(\alpha)[\mathbf{D}\{f,\mathbf{X}(\alpha)\}]^{-1}\mathbf{C}\{\mathbf{X}(\alpha)\}[\mathbf{D}\{f,\mathbf{X}(\alpha)\}]^{-1}\{\mathbf{X}(\alpha)\}^T\\ &= \{\mu(d,f)\}^{-2}\mathbf{X}(\alpha)[\mathbf{G}\{\mathbf{X}(\alpha)\}]^{-1}\mathbf{C}\{\mathbf{X}(\alpha)\}[\mathbf{G}\{\mathbf{X}(\alpha)\}]^{-1}\{\mathbf{X}(\alpha)\}^T\end{aligned}$$

which we can write as $\Delta\{f,\mathbf{I}_d,\mathbf{X}(\alpha)\}$, where by affine equivariance we have $\Delta\{f,\Sigma,\mathbf{A}\} = \Sigma^{1/2}\Delta\{f,\mathbf{I}_d,\Sigma^{-1/2}\mathbf{A}\}\Sigma^{1/2}$.

Next, observe that it is enough to prove the non-negative definite ordering of Δ stated in the Theorem for $\Sigma = \mathbf{I}_d$ and $\mathbf{B}^T\mathbf{B} = \mathbf{I}_d$ because when $\mathbf{B}^T\mathbf{B} = \lambda\mathbf{I}_d$, $\Delta\{f,\mathbf{I}_d,\mathbf{B}\}$ is a diagonal matrix that does not depend on the value of λ or the specific choice of $\mathbf{B}$. Also, for any nonsingular $\mathbf{A}$,

$$\Delta\{f,\mathbf{I}_d,\mathbf{A}\} = \{\mu(d,f)\}^{-2}\mathbf{A}\{\mathbf{G}(\mathbf{A})\}^{-1}\mathbf{C}(\mathbf{A})\{\mathbf{G}(\mathbf{A})\}^{-1}\mathbf{A}^T,$$

and hence, in order to prove the non-negative definite ordering of Δ, we can choose f to be any specific density as its effect appears only through the scalar factor $\mu(d,f)$. In particular, we can choose $f(\mathbf{x}^T\mathbf{x})$ to be the multivariate Laplace density $k\exp\{-(\mathbf{x}^T\mathbf{x})^{1/2}\}$. Then it is straight forward to verify that for a random vector $\mathbf{Z}$ with density $h(\mathbf{z}|\mathbf{A}) = \det(\mathbf{A})f(\mathbf{z}^T\mathbf{A}^T\mathbf{A}\mathbf{z})$, we must have

$$\begin{aligned}(\mathbf{A}^T)^{-1}\mathbf{D}(f,\mathbf{A})\mathbf{A}^{-1} &= (\mathbf{A}^T)^{-1}E\{\|\mathbf{Z}\|^{-1}(\mathbf{I}_d - \|\mathbf{Z}\|^{-2}\mathbf{Z}\mathbf{Z}^T)\}\mathbf{A}^{-1}\\ &= \mathbf{COV}\Big\{\|\mathbf{Z}\|^{-1}(\mathbf{A}^T)^{-1}\mathbf{Z}\|\mathbf{A}\mathbf{Z}\|^{-1}\mathbf{A}\mathbf{Z}\Big\},\end{aligned}$$

where $\mathbf{COV}$ denotes the covariance matrix between two random vectors. Also, $(\mathbf{A}^T)^{-1}\mathbf{C}(\mathbf{A})\mathbf{A}^{-1}$ is nothing but the dispersion matrix of $\|\mathbf{Z}\|(\mathbf{A}^T)^{-1}\mathbf{Z}$. Now, the non-negative definiteness of the difference

$$\begin{aligned}&\Delta(f,\mathbf{I}_d,\mathbf{A}) \;-\; \Delta(f,\mathbf{I}_d,\mathbf{B})\\ &= \mathbf{A}\{\mathbf{D}(f,\mathbf{A})\}^{-1}\mathbf{C}(\mathbf{A})\{\mathbf{D}(f,\mathbf{A})\}^{-1}\mathbf{A}^T - \mathbf{B}\{\mathbf{D}(f,\mathbf{B})\}^{-1}\mathbf{C}(\mathbf{B})\{\mathbf{D}(f,\mathbf{B})\}^{-1}\mathbf{B}^T\end{aligned}$$

follows from the simple fact that for any two d-dimensional random vectors $\mathbf{U}$ and $\mathbf{V}$, the difference

$$\{\mathbf{COV}(\mathbf{V},\mathbf{U})\}^{-1}\mathbf{DISP}(\mathbf{V})\{\mathbf{COV}(\mathbf{U},\mathbf{V})\}^{-1} - \{\mathbf{DISP}(\mathbf{U})\}^{-1}$$

is non-negative definite, where **DISP** denotes dispersion matrix, and all the matrices involved are invertible. Finally, the expression of $c(d,f)$ stated in the Theorem follows from a direct algebraic computation using the asymptotic distribution of the spatial median in spherically symmetric models (see e.g. Brown (1983), Chaudhuri (1992)).

Proof of Theorem 2.2. First note that in view of the affine invariance of the test statistic $\mathbf{T}_n^{(\alpha)}$, it is enough to prove the entire Theorem only for $\Sigma = \mathbf{I}_d$. Now, since given $\mathbf{X}(\alpha)$ the transformed observations $\mathbf{Z}_j^{(\alpha)}$'s are conditionally independent and they are identically distributed with the elliptically symmetric density $\det\{\mathbf{X}(\alpha)\}f[\mathbf{z}^T\{\mathbf{X}(\alpha)\}^T\mathbf{X}(\alpha)\mathbf{z}]$, the conditional distribution of $\mathbf{T}_n^{(\alpha)} = \sum_{j\notin\alpha}\|\mathbf{Z}_j^{(\alpha)}\|^{-1}\mathbf{Z}_j^{(\alpha)}$ does not depend on f and depends only on $\mathbf{X}(\alpha)$. This actually follows from what we have already seen in the proof of Theorem 2.1. Next, the asymptotic normality of the conditional null distribution of $n^{-1/2}\mathbf{T}_n^{(\alpha)}$ follows by a straightforward application of the central limit theorem, and the variance covariance matrix Ψ is equal to the matrix $\mathbf{C}$ defined in the proof of Theorem 2.1.

When $\log f$ is twice differentiable almost everywhere in $\mathbb{R}^d$ and satisfies Cramer type conditions (what we really need here is the square integrability of the first derivatives and the rth power ($r > 1$) integrability of the second derivatives of $\log f$ under the density f), the alternatives $H_A^{(n)}$ form a contiguous sequence in view of some straight forward analysis using Le Cam's first lemma (see Hajek and Sidak (1967)). Further, some standard analysis using Le Cam's third lemma (see Hajek and Sidak (1967)) and the spherical symmetry of the density f imply that under the sequence $H_A^{(n)}$, the conditional limiting distribution of $n^{-1/2}\mathbf{T}_n^{(\alpha)}$ given $\mathbf{X}(\alpha)$ is normal with mean equal to $\beta(d,f)\mathbf{H}\{\mathbf{X}(\alpha)\}\Phi = \Lambda\{f,\Phi,\mathbf{X}(\alpha)\}$ and $\Psi\{\mathbf{X}(\alpha)\}$ as the variance covariance matrix. Here β is a scalar multiple that depends only on the dimension d and the density f, and the $d\times d$ matrix $\mathbf{H}\{\mathbf{X}(\alpha)\}$ is equal to $\mathbf{COV}[\|\{\mathbf{X}(\alpha)\}^{-1}\mathbf{Z}\|^{-1}\{\mathbf{X}(\alpha)\}^{-1}\mathbf{Z}\ ,\ \|\mathbf{Z}\|^{-1}\mathbf{Z}]$, where $\mathbf{Z}$ is a d-dimensional random vector with density $f(\mathbf{z}^T\mathbf{z})$. This immediately implies that the conditional limiting distribution of $n^{-1}\{\mathbf{T}_n^{(\alpha)}\}^T[\Psi\{\Sigma^{-1/2}\mathbf{X}(\alpha)\}]^{-1}\mathbf{T}_n^{(\alpha)}$ under $H_A^{(n)}$ is noncentral χ^2 with d d.f. and noncentrality parameter $\delta\{f,\Phi,\mathbf{X}(\alpha)\}$, which is defined in the statement of the Theorem. Consequently the limiting conditional power of the test under the sequence of contiguous alternatives will be a monotonically increasing function of this δ. Finally the ordering of δ stated in the Theorem will follow if we can show that

$$\{\mathbf{H}(\mathbf{A})\}^T\{\Psi(\mathbf{A})\}^{-1}\mathbf{H}(\mathbf{A}) \ \leq_{n.n.d}\ \{\mathbf{H}(\mathbf{B})\}^T\{\Psi(\mathbf{B})\}^{-1}\mathbf{H}(\mathbf{B})\ ,$$

for any two nonsingular matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{B}^T\mathbf{B} = \lambda\mathbf{I}_d$. The proof of this nonnegative definite ordering of matrices follows from straightforward arguments that are very similar to those used in the second half of the proof of Theorem 2.1.

References

Andrews, D. F. and Herzberg, A. M. (1985). *Data : A Collection of Problems from Many Fields for the Student and Research Worker.* Springer-Verlag, New York.

Bennet, B. M. (1962). On multivariate sign tests. *J. Roy. Statist. Soc. Ser. B* **24**, 159-161.

Bickel, P. J. (1964). On some alternative estimates for shift in the P-variate one sample problem. *Ann. Math. Statist.* **35**, 1079-1090.

Bickel, P. J. (1965). On some ssymptotically nonparametric competitors of Hotelling's T^2. *Ann. Math. Statist.* **36**, 160-173.

Blumen, I. (1958). A new bivariate sign test. *J. Amer. Statist. Assoc.* **53**, 448-456.

Brown, B. M. (1983). Statistical use of the spatial median. *J. Roy. Statist. Soc. Ser. B* **45**, 25-30.

Brown, B. M. (1985). Spatial median. In *Encyclopedia of Statistical Science* vol. 8, 574-575. Wiley, New York.

Brown, B. M. and Hettmansperger, T. P. (1987). Affine invariant rank methods in the bivariate location model. *J. Roy. Statist. Soc. Ser. B* **49**, 301-310.

Brown, B. M. and Hettmansperger, T. P. (1989). An affine invariant bivariate version of the sign test. *J. Roy. Statist. Soc. Ser. B* **51**, 117-125.

Brown, B. M., Hettmansperger, T. P., Nyblom, J. and Oja, H. (1992). On certain bivariate sign tests and medians. *J. Amer. Statist. Assoc.* **87**, 127-135.

Chakraborty, B. and Chaudhuri, P. (1996). On a transformation and re-transformation technique for constructing affine equivariant multivariate median. *Proc. Amer. Math. Soc.* **124**, 2539-2547.

Chakraborty, B. and Chaudhuri, P. (1998). On an adaptive transformation-retransformation estimate of multivariate location. *J. Roy. Statist. Soc. Ser. B* **60**, 145-157.

Chatterjee, S. K. (1966). A bivariate sign test for location. *Ann. Math. Statist.* **37**, 1771-1782.

Chaudhuri, P. (1992). Multivariate location estimation using extension of R-estimates through U-statistics type approach. *Ann. Statist.* **20**, 897-916.

Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *J. Amer. Statist. Assoc.* **91**, 862-872.

Chaudhuri, P. and Sengupta, D. (1993). Sign tests in multidimension : inference based on the geometry of the data cloud. *J. Amer. Statist. Assoc.* **88**, 1363-1370.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans.* SIAM, Philadelphia.

Gini, C. and Galvani, L. (1929). Di talune estensioni dei concetti di media ai caratteri qualitativi. *Metron*, **8**. Partial English translation in *J. Amer. Statist. Assoc.* **25**, 448-450.

Gower, J. C. (1974). The mediancenter. *J. Roy. Statist. Soc. Ser. C* **23**, 466-470.

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

Haldane, J. B. S. (1948). Note on the median of a multivariate distribution. *Biometrika* **35**, 414-415.

Hayford, J. F. (1902). What is the center of an area, or the center of a population ? *J. Amer. Statist. Assoc.* **8**, 47-58.

Hettmansperger, T. P., Nyblom, J. and Oja, H. (1992). On multivariate notions of sign and rank. In L_1 *Statistical Analysis and Related Methods* (Editor by Y. Dodge), 267-278. North Holland : Amsterdam.

Hettmansperger, T. P., Nyblom, J. and Oja, H. (1994). Affine invariant multivariate one sample sign tests. *J. Roy. Statist. Soc. Ser. B* **56**, 221-234.

Hodges, J. L. (1955). A bivariate sign test. *Ann. Math. Statist.* **26**, 523-527.

Isogai, T. (1985). Some extension of Haldane's multivariate median and its application. *Ann. Inst. Statist. Math.* **37**, 289-301.

Liu, R. Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18**, 405-414.

Merchants, J. A., Halprin, G. M., Hudson, A. R., Kilburn, K. H., McKenzie, W. N., Jr., Hurst, D. J. and Bermazohn, P. (1975). Responses to cotton dust. *Arch. Environmental Health* **30**, 222-229.

Mottonen, J. and Oja, H. (1995). Multivariate spatial sign and rank methods. *J. Nonparametr. Statist.* **5**, 201-213.

Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statist. Probab. Lett.* **1**, 327-332.

Oja, H. and Nyblom, J. (1989). Bivariate sign tests. *J. Amer. Statist. Assoc.* **84**, 249-259.

Puri, M. L. and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* Wiley, New York.

Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections. *J. Amer. Statist. Assoc.* **84**, 1045-1050.

Rao, C. R. (1988). Methodology based on the L_1-norm in statistical inference. *Sankhyā Ser. A* **50**, 289-313.

Scates, D. E. (1933). Locating the median of the population in the United States. *Metron* **11**, 49-66.

Small, C. G. (1990). A survey of multidimensional medians. *Internat. Statist. Rev.* **58**, 263-277.

Tukey, J. W. (1975). Mathematics and picturing data. In *Proceedings of the International Congress of Mathematicians, Vancouver 1974*, vol 2 (Edited by R. D. James), 523-531. Canadian mathematical Congress.

Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, 203 B. T. Road, Calcutta 700035, India.

E-mail: res9421@isical.ac.in

E-mail: probal@isical.ac.in

Department of Mathematical Sciences, University of Oulu, FIN-90570, Oulu, Finland.

E-mail: hannuoja@cc.oulu.fi

(Received September 1996; accepted March 1997)