Author(s): Taskinen, Sara; Kankainen, Annaliisa; Oja, Hannu

Title: Sign test of independence between two random vectors

Year: 2003

Version:

Please cite the original version:

Taskinen, S., Kankainen, A., & Oja, H. (2003). Sign test of independence between two random vectors. Statist. Probab. Lett., 62(1), 9-21. https://doi.org/10.1016/S0167-7152(02)00399-1

SIGN TEST OF INDEPENDENCE BETWEEN TWO RANDOM VECTORS

Sara Taskinen, Annaliisa Kankainen and Hannu Oja

Department of Mathematics and Statistics, University of Jyväskylä
P.O. Box 35, Fin-40351 Jyväskylä, Finland.

January 3, 2003

Abstract

A new affine invariant extension of the quadrant test statistic (Blomqvist, 1950) based on spatial signs is proposed for testing the hypothesis of independence. In the elliptic case, the new test statistic is asymptotically equivalent to the interdirection test by Gieser and Randles (1997), but is easier to compute in practice. Limiting Pitman efficiencies and simulations are used to compare the test to the classical Wilks' test.

Keywords: Affine invariance, Pitman efficiency, Quadrant test, Robustness, Wilks' test

1 INTRODUCTION

In the paper, we consider the interrelations between two sets of variables, the p-vector $\boldsymbol{x}^{(1)}$ and q-vector $\boldsymbol{x}^{(2)}$. Let

$$\begin{pmatrix} \boldsymbol{x}_1^{(1)} \\ \boldsymbol{x}_1^{(2)} \end{pmatrix}, \begin{pmatrix} \boldsymbol{x}_2^{(1)} \\ \boldsymbol{x}_2^{(2)} \end{pmatrix}, \ldots, \begin{pmatrix} \boldsymbol{x}_n^{(1)} \\ \boldsymbol{x}_n^{(2)} \end{pmatrix},$$

be a random sample from an unknown k variate distribution ($k = p + q$). The null hypothesis to be tested is H_0: $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ are independent.

If the observations come from a k-variate normal distribution with partitioned covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

the hypothesis of independence may be formulated as $\Sigma_{12} = 0$. If

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}.$$

is the partitioned sample covariance matrix, then the likelihood ratio test statistic (Wilks, 1935) is given by

$$W_n = |I_q - S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}|.$$

An asymptotically equivalent test statistic is the Pillai's (1955) test statistic

$$T_n = Tr(S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}).$$

Note that T_n is affine invariant, that is, invariant under the group of transformations

$$\begin{pmatrix} A\boldsymbol{x}_1^{(1)} \\ B\boldsymbol{x}_1^{(2)} \end{pmatrix}, \begin{pmatrix} A\boldsymbol{x}_2^{(1)} \\ B\boldsymbol{x}_2^{(2)} \end{pmatrix}, \ldots, \begin{pmatrix} A\boldsymbol{x}_n^{(1)} \\ B\boldsymbol{x}_n^{(2)} \end{pmatrix},$$

where A is nonsingular $p \times p$ matrix and B nonsingular $q \times q$ matrix. If $p = q = 1$, the regular squared correlation coefficient is obtained, and T_n may be seen as an extension of the multiple correlation coefficient. See Mardia et al. (1997, Section 6.5.4), for a discussion on that. Note that statistic T_n can be used to test the hypothesis of independence for nonnormal cases also as independence implies that $\Sigma_{12} = 0$ (if it exists). Under H_0 (and some general assumptions), $nT_n \rightarrow_d \chi^2_{pq}$.

For $p = q = 1$, nonparametric approaches to the independence problems include classical quadrant test, Spearman's rho and Kendall's tau, see for example Lehmann (1998, Section 7) and references therein. Puri and Sen (1971) considered extensions where the test statistics, analogous to T_n, are obtained replacing the regular covariance matrix by the covariance matrix of marginal (centered) signs or ranks (or scored ranks). Unfortunately, these extensions of quadrant test and Spearman's rho miss the desirable affine invariance property. Gieser and Randles (1997) introduced an invariant extension of the quadrant test. In their construction, they first centered separately the $\boldsymbol{x}^{(1)}$-vectors and the $\boldsymbol{x}^{(2)}$-vectors (using affine equivariant Oja (1983) median). Next they calculated the interdirection proportions a_{ij} and b_{ij} between pairs of centered $\boldsymbol{x}^{(1)}$-vectors and centered $\boldsymbol{x}^{(2)}$-vectors, respectively. Finally, the test statistic is

$$Q_n \;=\; pq \;\mathrm{ave}_{i,j}\left\{\cos(\pi a_{ij})\cos(\pi b_{ij})\right\},$$

where the average is taken over all possible $i, j = 1, \ldots, n$. Again, under H_0 and under some general assumptions (both $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ elliptically symmetric), $nQ_n \to_d \chi^2_{pq}$. Gieser and Randles also found asymptotic relative efficiencies of the Q_n test as compared to the classical Wilks' test (or the asymptotically equivalent T_n test). However, the computation of the interdirection proportions is time consuming, and may cause problems for large sample sizes and in high dimensions.

In the paper, a new invariant test based on multivariate standardized spatial signs is proposed. The standardized spatial signs described in Section 2 are based on the approach launched in Randles (2000) and Hettmansperger and Randles (2002). The test statistic, based on the covariance matrix between the standardized spatial signs of the $\boldsymbol{x}^{(1)}$- and $\boldsymbol{x}^{(2)}$-vectors and again analogous to T_n, is introduced in Section 3. Also asymptotic theory is developed to provide approximations for the limiting distributions. In Section 4 the limiting Pitman efficiencies are given in the multivariate normal distribution, t distribution and contaminated normal distribution cases and simulations are used to compare finite sample powers. In Section 5 the theory is illustrated by an example and the paper is closed with some comments in Section 6. The proofs are postponed to the Appendix.

2 STANDARDIZED SPATIAL SIGNS

To construct invariant test procedures for the independence problem, one has to estimate (explicitly or implicitly) the location centres and covariance structures of the marginal multivariate distributions of the $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$. Assume that the distribution of the $\boldsymbol{x}$ is p-variate and elliptic with symmetry centre $\boldsymbol{\mu}$ and covariance matrix Σ (if they exist). Matrix $V = (k/Tr(\Sigma))\Sigma$ is then the related *shape matrix*.

Hettmansperger and Randles (2002) introduced intuitively appealing location vector and shape matrix estimates based on the spatial sign concept.

The spatial sign of vector $\boldsymbol{x}$ is defined as

$$\boldsymbol{S}(\boldsymbol{x}) = \begin{cases} \|\boldsymbol{x}\|^{-1}\boldsymbol{x}, & \boldsymbol{x} \neq \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{x} = \boldsymbol{0}, \end{cases}$$

where $\|\boldsymbol{x}\| = (\boldsymbol{x}^T\boldsymbol{x})^{1/2}$ is the (Euclidean) length of the vector $\boldsymbol{x}$. Note that $\boldsymbol{S}(\boldsymbol{x})$ is a unit vector in the direction of $\boldsymbol{x}$ and therefore a natural multivariate extension of the sign concept. Spatial signs $\boldsymbol{S}(\boldsymbol{x}_i)$, $i = 1, \ldots, n$, are not affine equivariant, however.

In Hettmansperger and Randles (2002) estimates $\widehat{\boldsymbol{\mu}}$ and $\widehat{V}$ are obtained as follows. ($\widehat{V}^{-1/2}$ is here taken to be symmetric.)

Definition 1. *The location vector estimate $\widehat{\boldsymbol{\mu}}$ and the shape matrix estimate $\widehat{V}$ are choices such that the spatial signs of standardized observations*

$$\boldsymbol{z}_i = \widehat{V}^{-1/2}(\boldsymbol{x}_i - \widehat{\boldsymbol{\mu}}), \quad i = 1, \ldots, n$$

satisfy

$$\text{ave}_i\{\boldsymbol{S}(\boldsymbol{z}_i)\} \;=\; \boldsymbol{0} \quad \textit{and} \quad \text{ave}_i\left\{\boldsymbol{S}(\boldsymbol{z}_i)\boldsymbol{S}^T(\boldsymbol{z}_i)\right\} \;=\; \frac{1}{p}I_p.$$

The vectors

$$\boldsymbol{S}_i = \boldsymbol{S}(\boldsymbol{z}_i), \quad i = 1, \ldots, n$$

are called the standardized signs.

The resulting shape matrix estimate $\widehat{V}$ is the so called Tyler's M-estimate (1987) and the location estimate vector $\widehat{\boldsymbol{\mu}}$ is a transformation-retransformation spatial median utilizing Tyler's transformation matrix. The estimates are naturally affine equivariant. For the transformation retransformation technique, see also Chakraborty et al. (1998). For the properties of these and related location vector and shape matrix estimates, see also Ollila, Hettmansperger and Oja (2002).

We will use the standardized signs in the test constructions and then need the following equivariance property.

Lemma 1. *The standardized sign vectors $\boldsymbol{S}_i$ are affine equivariant in the sense that if the $\boldsymbol{S}_i^*$ are calculated from $\boldsymbol{x}_i^* = A\boldsymbol{x}_i + \boldsymbol{b}$, $i = 1, \ldots, n$ with a nonsingular $k \times k$ matrix A and k-vector $\boldsymbol{b}$, then $\boldsymbol{S}_i^* = P\boldsymbol{S}_i$, $i = 1, \ldots, n$, where transformation matrix $P = (A\widehat{V}A^T)^{-1/2}A\widehat{V}^{1/2}$ is orthogonal.*

3 NEW TEST STATISTIC AND ITS LIMITING DISTRIBUTION

We are now ready to introduce our test statistic. As earlier, let

$$\begin{pmatrix} \boldsymbol{x}_1^{(1)} \\ \boldsymbol{x}_1^{(2)} \end{pmatrix}, \begin{pmatrix} \boldsymbol{x}_2^{(1)} \\ \boldsymbol{x}_2^{(2)} \end{pmatrix}, \ldots, \begin{pmatrix} \boldsymbol{x}_n^{(1)} \\ \boldsymbol{x}_n^{(2)} \end{pmatrix}$$

be a random sample from an unknown k-variate distribution ($k = p+q$) and the null hypothesis to be tested is H_0: $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ are independent.

First, we construct standardized spatial sign vectors separately for the $\boldsymbol{x}^{(1)}$- and the $\boldsymbol{x}^{(2)}$-vectors and obtain

$$\begin{pmatrix} \boldsymbol{S}_1^{(1)} \\ \boldsymbol{S}_1^{(2)} \end{pmatrix}, \begin{pmatrix} \boldsymbol{S}_2^{(1)} \\ \boldsymbol{S}_2^{(2)} \end{pmatrix}, \ldots, \begin{pmatrix} \boldsymbol{S}_n^{(1)} \\ \boldsymbol{S}_n^{(2)} \end{pmatrix}.$$

The sign vectors $\boldsymbol{S}_i^{(1)}$ and $\boldsymbol{S}_i^{(2)}$ thus satisfy

$$\text{ave}_i\{\boldsymbol{S}_i^{(1)}\} = 0 \quad \text{and} \quad \text{ave}_i\{\boldsymbol{S}_i^{(1)}\boldsymbol{S}_i^{(1)^T}\} = p^{-1}I_p$$

and

$$\text{ave}_i\{\boldsymbol{S}_i^{(2)}\} = 0 \quad \text{and} \quad \text{ave}_i\{\boldsymbol{S}_i^{(2)}\boldsymbol{S}_i^{(2)^T}\} = q^{-1}I_q.$$

Next write

$$H = \text{ave}_i \left\{ \boldsymbol{S}_i^{(1)}\boldsymbol{S}_i^{(2)^T} \right\}$$

for the covariance matrix between the marginal standardized signs. The new test statistic is then given by the next

Definition 2. *The standardized spatial sign test statistic for testing the hypothesis of independence H_0 is given by*

$$U_n = pq \; ||H||^2$$

where $||H||^2 = Tr(H^T H)$ is the squared matrix norm (Frobenius norm).

As the standardized signs are affine equivariant, the invariance of the test statistic easilly follows.

Lemma 2. *U_n is affine invariant.*

The null distribution of the test statistic is given by

Theorem 1. *Under H_0 and for elliptic $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$, the limiting distribution of nU_n is a chi-squared distribution with pq degrees of freedom.*

Next we derive the limiting distribution of U_n under alternative sequences. As U_n is affine invariant, we restrict to the spherical case only. Let thus $\boldsymbol{x}_i^{(1)}$ and $\boldsymbol{x}_i^{(2)}$ be independent with spherical marginal distributions and write

$$\begin{pmatrix} \boldsymbol{y}_i^{(1)} \\ \boldsymbol{y}_i^{(2)} \end{pmatrix} = \begin{pmatrix} (1-\Delta)I_p & \Delta M_1 \\ \Delta M_2 & (1-\Delta)I_q \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_i^{(1)} \\ \boldsymbol{x}_i^{(2)} \end{pmatrix}, \tag{1}$$

where $\Delta = \delta/\sqrt{n}$. The sequence of alternatives is similar to that used in Gieser and Randles (1997). If U_n^* is calculated from these transformed observations in (1), we get

Theorem 2. *For* $\max(p,q) > 1$*, the limiting distribution of* nU_n^* *is a noncentral chi-squared distribution with pq degrees of freedom and noncentrality parameter*

$$\frac{\delta^2}{pq}||c_1 M_1 + c_2 M_2^T||^2,$$

where

$$c_1 = (p-1)E(||\boldsymbol{x}_i^{(2)}||)E(||\boldsymbol{x}_i^{(1)}||^{-1}) \quad \text{and} \quad c_2 = (q-1)E(||\boldsymbol{x}_i^{(1)}||)E(||\boldsymbol{x}_i^{(2)}||^{-1}).$$

For the case $p = q = 1$, see Taskinen et al. (2003).

4 LIMITING AND FINITE-SAMPLE EFFICIENCIES

We next compare U_n to Wilks' likelihood ratio test W_n and the test based on interdirections Q_n through limiting efficiencies and simple simulation studies.

As all the limiting distributions are of the same type, χ^2_{pq}, the efficiency comparisons (Pitman efficiencies) may be based on noncentrality parameters only. First note that, for elliptic $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$, U_n is asymptotically equivalent to the test based on interdirections. See (A.1) in Gieser and Randles (1997). For the asymptotic efficiencies in the exponential power family, see then Gieser and Randles (1997). Gieser (1993) considered also the t distribution case in his unpublished dissertation thesis. The efficiency comparison is now made in the multivariate normal distribution, t distribution and contaminated normal distribution cases. For simplicity, we assume that $M_1 = M_2^T$.

Since W_n has a limiting noncentral chi-squared distribution with pq degrees of freedom and noncentrality parameter $\delta^2||M_1 + M_2^T||^2$, the asymptotic efficiencies are simply

$$\text{ARE}(U_n, W_n) = \frac{(c_1 + c_2)^2}{4pq} \tag{2}$$

with c_1 and c_2 given in Theorem 2. For a k-dimensional standardized t distribution with ν degrees of freedom

$$E_{t_\nu}(||\boldsymbol{x}_i||) = \frac{\sqrt{\nu-2}\,\Gamma\left(\frac{\nu-1}{2}\right)\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\Gamma\left(\frac{k}{2}\right)} \quad \text{and} \quad E_{t_\nu}(||\boldsymbol{x}_i||^{-1}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)\Gamma\left(\frac{k-1}{2}\right)}{\sqrt{\nu-2}\,\Gamma\left(\frac{\nu}{2}\right)\Gamma\left(\frac{k}{2}\right)}$$

and the resulting limiting efficiencies for selected degrees of freedom and selected dimensions are listed in Table 1. Note that since limiting multinormality of the regular covariance matrix holds if the forth moments of underlying distribution are finite, W_n has a limiting distribution only when $\nu \geq 5$. In multinormal cases the asymptotic efficiency of U_n (and Q_n) is

Table 1: ARE(U_n, W_n) at different p- and q-variate t distributions for selected $\nu = \nu_1 = \nu_2$.

		p					
	q	2	3	5	8	10	∞
	2	0.790	0.854	0.911	0.945	0.957	1.006
	3		0.923	0.984	1.022	1.034	1.087
$\nu = 5$	5			1.050	1.090	1.103	1.160
	8				1.131	1.145	1.203
	10					1.159	1.219
	∞						1.281
	2	0.689	0.745	0.795	0.824	0.835	0.877
	3		0.805	0.859	0.891	0.902	0.948
$\nu = 10$	5			0.916	0.950	0.962	1.012
	8				0.986	0.999	1.050
	10					1.011	1.063
	∞						1.117
	2	0.617	0.667	0.711	0.738	0.747	0.785
	3		0.721	0.769	0.798	0.807	0.849
$\nu = \infty$	5			0.820	0.851	0.861	0.905
	8				0.883	0.894	0.940
	10					0.905	0.951
	∞						1

low but gets higher as the dimensions p and q increase. For heavy-tailed distributions the efficiencies are low in the case of small dimensions, but as the dimensions increase, new test outperforms the W_n test.

We also consider the contaminated normal distribution and assume that $p = q$ and $M_1 = M_2^T$. Then with probability $(1 - \epsilon)$ the $\boldsymbol{x}^{(1)}$-vectors (and $\boldsymbol{x}^{(2)}$- vectors) are drawn from a standard normal distribution $N_p(\mathbf{0}, I_p)$, and with probability ϵ from $N_p(\mathbf{0}, c^2 I_p)$. The cdf is thus $F(\boldsymbol{x}) = (1 - \epsilon)\,\Phi_p(\boldsymbol{x}) + \epsilon\,\Phi_p(\boldsymbol{x}/c)$, where $c > 0$ and Φ_p is the cdf of $N_p(\mathbf{0}, I_p)$. Now in (2)

$$c_1 = c_2 = (p-1)(1 - \epsilon + \epsilon\, c)(1 - \epsilon + \epsilon/c) E_{\Phi_p}(||\boldsymbol{x}_i^{(1)}||) E_{\Phi_p}(||\boldsymbol{x}_i^{(1)}||^{-1}).$$

Asymptotic relative efficiencies ARE(U_n, W_n) are illustrated in Figure 1 as a function of c, for $p = 5$ and for $\epsilon = 0.05, 0.10, 0.20$. For heavy-tailed cases (large ϵ and/or large c), the new test is much more efficient than the Wilks' likelihood ratio test.

Finally, a simulation study was used to compare the finite sample powers of three test statistics. The critical values used in test constructions were based on the chi-square approximations to the null distributions. The empirical powers were calculated for $p = q = 3$ (Figure 2) and for $p = q = 5$

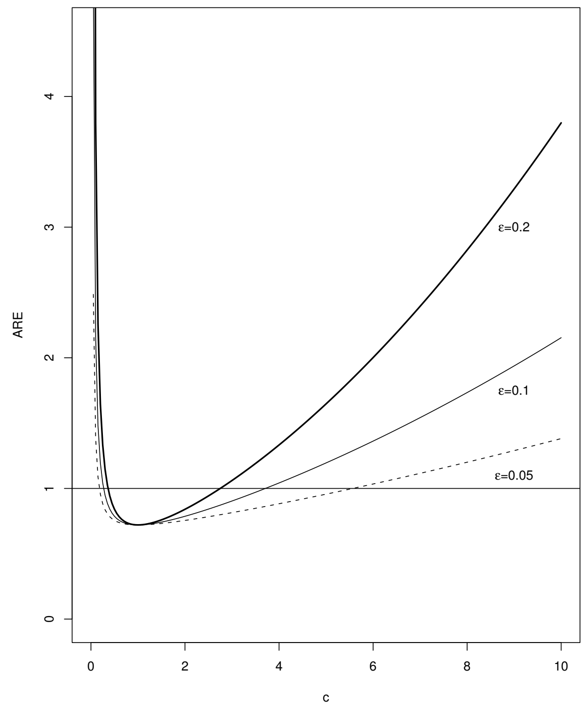

Figure 1: ARE(U_n, W_n) as a function of c at the contaminated normal model for $p = q = 5$ and for $\epsilon = 0.05, 0.10, 0.20$.

(Figure 3). 1500 independent $\boldsymbol{x}^{(1)}$- and $\boldsymbol{x}^{(2)}$-samples of sizes $n = 50$ and 150 were generated from a multivariate standard normal distribution, from a standardized t distribution with $\nu_1 = \nu_2 = 5$ and from a contaminated normal distribution with $\epsilon = 0.2$ and $c = 6$. Finally, the transformation in (1) with $M_1 = M_2^T = I_r$, where $r = p = q$, was applied for chosen values of δ to introduce dependence into the model.

In computations, test statistic W_n was multiplied by Bartlett's correction factor $1 - (2r + 3)/2n$, and for Q_n the data vectors were centered using marginal transformation-retransformation spatial medians (Hettmansperger and Randles, 2002). In most cases, with the exception of $n = 50$ and $r = 3$, the interdirection proportions, needed for Q_n, were estimated using a sample of 1000 hyperplanes (instead of using all possible $\binom{n-2}{p-1}$ hyperplanes).

Consider first the simulation results in the case $r = 3$ (Figure 2). In the multinormal case W_n is clearly best, as one can expect. In the considered t distribution case W_n is slightly better than the other tests, but U_n and Q_n perform very well in the contaminated normal case. For small samples, Q_n seems to be a bit more powerful than U_n. As $r = 5$ (Figure 3), the empirical powers of U_n and Q_n increase as compared to the W_n. As $n = 50$, Q_n does well also in the normal case, but again, as n increases, no significant differences can be seen between U_n and Q_n. The size of the U_n test is reasonably close to the aimed size 0.05 in all cases whereas the Q_n and especially the W_n test levels often seem to exceed 0.05.

5 A REAL DATA EXAMPLE

The theory is now illustrated by a simple example. Consider the diabetes data (see Rencher, 1998, p. 17 and references therein), that contains 76 5-variate observations, namely glucose intolerance, insulin response to oral glucose, insulin resistance, relative weight and fasting plasma glucose. We wish to test the hypothesis that the first three variables are independent of the remaining two variables.

First, p-values given by W_n, U_n and Q_n were calculated for the original data set using the chi-square approximation of the null distributions. Then, to illustrate that the test statistics U_n and Q_n are robust against outliers, the p-values were recalculated for a data set, where the first relative weight measurement 0.81 was replaced by 8.1 (simulating a printing error). The obtained p-values are listed in Table 2. Note that in the original case all statistics produce very small p-values, but the contamination changes the p-value given by W_n drastically; only small changes can be seen in the other two cases.

Table 2: p-values for the original and contaminated data sets.

	Original	Contaminated
Statistic	p-value	p-value
W_n	0.0003	0.265
U_n	0.021	0.037
Q_n	0.020	0.037

6 FINAL REMARKS

In the paper we proposed a new robust and affine invariant extension of the quadrant test statistic to test the hypothesis that two random vectors are independent. If the marginal distributions of $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ are elliptic, the new test is asymptotically equivalent with the interdirection quadrant statistic (Gieser and Randles, 1997) but is much easier to compute. The new test is not very efficient in the multinormal case but performs well for heavy-tailed distributions. This is illustrated with limiting and finite sample efficiency studies.

A related testing problem is that of testing the null hypothesis $H_0: \Sigma_{12} = 0$, which says that all the canonical correlations between the $\boldsymbol{x}^{(1)}$- and $\boldsymbol{x}^{(2)}$-variables are zero. For testing H_0 in the elliptic case, any shape matrix V for $(\boldsymbol{x}^{(1)^T}, \boldsymbol{x}^{(2)^T})^T$ may be used in the test construction. If sample statistic V is partitioned as

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix},$$

then a test statistic, analogous to Pillai's test statistic, may be defined as

$$V_n = Tr(V_{11}^{-1} V_{12} V_{22}^{-1} V_{21}).$$

Again, V_n is affine invariant, and nV_n has a central (noncentral) chi-squared distribution with pq degrees of freedom under the null hypothesis (alternative sequences of type (1)). In the multinormal case, $H_0: \Sigma_{12} = 0$ is of course a hypothesis of independence also, and the tests based on shape matrices may be used for testing independency.

If V is the Tyler's M-estimate (Tyler, 1987; Hettmansperger and Randles, 2002) and $M_1 = M_2^T$, the asymptotic efficiency of V_n relative to U_n is simply

$$\text{ARE}(V_n, U_n) = \frac{4pq(p+q)}{(c_1 + c_2)^2 (p+q+2)},$$

where c_1 and c_2 are given in Theorem 2. The V_n test then outperforms the U_n test in the multinormal case; see asymptotic relative efficiencies listed in Table 3.

Table 3: $\text{ARE}(V_n, U_n)$ at different p- and q-variate normal distributions.

	p					
q	2	3	5	8	10	∞
2	1.081	1.071	1.094	1.129	1.147	1.273
3		1.041	1.041	1.061	1.073	1.178
5			1.017	1.019	1.024	1.104
8				1.007	1.007	1.064
10					1.005	1.051
∞						1

Acknowledgements

The authors wish to thank a referee for several helpful comments and suggestions. The research was partially supported by the Academy of Finland.

References

Blomqvist, N. (1950), On a measure of dependence between two random variables, *Ann. Math. Statist.*, **21**, 593–600.

Chakraborty, B., Chaudhuri, P. and Oja, H. (1998), Operating transformation retransformation on spatial median and angle test, *Statistica Sinica*, **8**, 767–784.

Gieser, P.W. (1993), A new nonparametric test for independence between two sets of variates, unpublished Ph.D. thesis, University of Florida, Gainesville.

Gieser, P.W. and Randles, R.H. (1997), A nonparametric test of independence between two vectors, *J. Amer. Statist. Assoc.*, **92**, 561–567.

Hettmansperger, T.P. and Randles, R.H. (2002), A practical affine equivariant multivariate median, *Biometrika*, **89**, 4, 851–860.

Lehmann, E.L. (1998), *Nonparametrics: Statistical Methods Based on Ranks* (Prentice Hall, Upper Saddle River, NJ).

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1997), *Multivariate Analysis* (Academic Press, London).

Ollila, E., Hettmansperger, T.P. and Oja, H. (2002), Affine equivariant multivariate sign methods. Under revision.

Oja, H. (1983), Descriptive statistics for multivariate distributions, *Statistics & Probability Letters*, **1**, 327–332.

Pillai, K.C.S. (1955), Some new test criteria in multivariate analysis, *Ann. Math. Statist.*, **26**, 117–121.

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

Randles, R.H. (2000). A simpler, affine-invariant, multivariate, distribution-free sign test, *J. Amer. Statist. Assoc.*, **95**, 1263–1268.

Rencher, A.C. (1998), *Multivariate Statistical Inference and Applications* (Wiley, New York).

Taskinen, S., Kankainen, A. and Oja, H. (2003), Tests of independence based on sign and rank covariances, in: R. Dutter, P. Filzmoser, U. Gather and P.J. Rousseeuw, eds. *Developments in Robust Statistics. International Conference on Robust Statistics 2001*, (Springer-Verlag, Heidelberg) pp. 387–403.

Tyler, D.E. (1987), A distribution-free M-estimator of multivariate scatter, *Ann. Statist.*, **15**, 234–251.

Wilks, S.S. (1935), On the independence of k sets of normally distributed statistical variables, *Econometrica*, **3**, 309–326.

Appendix: Proofs of the results

Proof of Lemma 1 Using $\widehat{V}^* = k\,[Tr(A\widehat{V}A^T)]^{-1}A\widehat{V}A^T$ and $\widehat{\boldsymbol{\mu}}^* = A\widehat{\boldsymbol{\mu}}+\boldsymbol{b}$, it is straightforward to show as in Randles (2000) that

$$\boldsymbol{S}_i^* = \boldsymbol{S}(\widehat{V}^{*^{-1/2}}(\boldsymbol{x}_i^* - \widehat{\boldsymbol{\mu}}^*)) = (A\widehat{V}A^T)^{-1/2}A\widehat{V}^{1/2}\boldsymbol{S}_i =: P\boldsymbol{S}_i.$$

Transformation matrix P is clearly orthogonal.

Proof of Lemma 2 The proof follows from Lemma 1 and the fact that $Tr(P_2H^TP_1^TP_1HP_2^T) = Tr(H^TH)$ for any orthogonal matrices P_1 and P_2.

Proof of Theorem 1 In the proof, we use the general delta method. As the test is affine invariant, it is enough to consider the spherical case with $\boldsymbol{\mu}_1 = \mathbf{0}$, $V_1 = I_p$, $\boldsymbol{\mu}_2 = \mathbf{0}$, $V_2 = I_q$. Let us first show that under H_0,

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{S}_i^{(1)}\boldsymbol{S}_i^{(2)^T} - \frac{1}{n}\sum_{i=1}^{n}\frac{\boldsymbol{x}_i^{(1)}\boldsymbol{x}_i^{(2)^T}}{||\boldsymbol{x}_i^{(1)}||||\boldsymbol{x}_i^{(2)}||}\right) \xrightarrow{P} 0.$$

Write the standardized location and shape estimates for $\boldsymbol{x}^{(1)}$ as

$$\boldsymbol{\mu}_1^* = \sqrt{n}\widehat{\boldsymbol{\mu}}_1 \quad \text{and} \quad V_1^* = \sqrt{n}(\widehat{V}_1 - I_p).$$

Then

$$\widehat{V}_1^{-1/2}(\boldsymbol{x}_i^{(1)} - \widehat{\boldsymbol{\mu}}_1) = \boldsymbol{x}_i^{(1)} - \frac{1}{2\sqrt{n}}V_1^*\boldsymbol{x}_i^{(1)} - \frac{1}{\sqrt{n}}\boldsymbol{\mu}_1^* + o_P(n^{-1/2})$$

and

$$\boldsymbol{S}_i^{(1)} = \frac{\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(1)}||} + \frac{1}{\sqrt{n}}\frac{\boldsymbol{x}_i^{(1)^T}\boldsymbol{\mu}_1^*}{||\boldsymbol{x}_i^{(1)}||^2}\frac{\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(1)}||} + \frac{1}{2\sqrt{n}}\frac{\boldsymbol{x}_i^{(1)^T}V_1^*\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(1)}||^2}\frac{\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(1)}||} - \frac{1}{\sqrt{n}}\frac{\boldsymbol{\mu}_1^*}{||\boldsymbol{x}_i^{(1)}||} - \frac{1}{2\sqrt{n}}\frac{V_1^*\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(1)}||} + o_P(n^{-1/2}).$$

Similarly

$$\boldsymbol{S}_i^{(2)} = \frac{\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(2)}||} + \frac{1}{\sqrt{n}}\frac{\boldsymbol{x}_i^{(2)^T}\boldsymbol{\mu}_2^*}{||\boldsymbol{x}_i^{(2)}||^2}\frac{\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(2)}||} + \frac{1}{2\sqrt{n}}\frac{\boldsymbol{x}_i^{(2)^T}V_2^*\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(2)}||^2}\frac{\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(2)}||} - \frac{1}{\sqrt{n}}\frac{\boldsymbol{\mu}_2^*}{||\boldsymbol{x}_i^{(2)}||} - \frac{1}{2\sqrt{n}}\frac{V_2^*\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(2)}||} + o_P(n^{-1/2}),$$

and it is easy to see, considering termwise sums of $\boldsymbol{S}_i^{(1)}\boldsymbol{S}_i^{(2)^T}$, that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{S}_i^{(1)}\boldsymbol{S}_i^{(2)^T} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\boldsymbol{x}_i^{(1)}\boldsymbol{x}_i^{(2)^T}}{||\boldsymbol{x}_i^{(1)}||||\boldsymbol{x}_i^{(2)}||} + o_P(1). \qquad (3)$$

The next step is to note that as $H = (h_{ij})$,

$$n\,pq\,Tr(H^T H) = \sum_{i=1}^{p}\sum_{j=1}^{q}\left(\frac{\sqrt{n}h_{ij}}{1/\sqrt{pq}}\right)^2.$$

Finally (3) implies that the limiting distribution of $\sqrt{n}\text{vec}(H)$ is $N_{pq}(\mathbf{0}, \frac{1}{pq}I_{pq})$, where $\text{vec}(H)$ is obtained by stacking the columns of H on top of each other, and consequently $nU_n \xrightarrow{d} \chi^2_{pq}$.

Proof of Theorem 2 The sequence of alternatives for $\Delta = \delta/\sqrt{n}$ is contiguous to the null hypothesis $\Delta = 0$. Therefore (see the proof of Theorem 1) also

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{\boldsymbol{S}}_i^{(1)}\tilde{\boldsymbol{S}}_i^{(2)^T} - \frac{1}{n}\sum_{i=1}^{n}\frac{\boldsymbol{y}_i^{(1)}\boldsymbol{y}_i^{(2)^T}}{||\boldsymbol{y}_i^{(1)}||||\boldsymbol{y}_i^{(2)}||}\right) \xrightarrow{P} 0,$$

where $\tilde{\boldsymbol{S}}_i^{(1)}$ and $\tilde{\boldsymbol{S}}_i^{(2)}$ are the standardized spatial sign vectors of $\boldsymbol{y}_i^{(1)}$ and $\boldsymbol{y}_i^{(2)}$. As

$$\boldsymbol{y}_i^{(1)} = \boldsymbol{x}_i^{(1)} - \frac{\delta}{\sqrt{n}}\boldsymbol{x}_i^{(1)} + \frac{\delta}{\sqrt{n}}M_1\boldsymbol{x}_i^{(2)} \quad \text{and} \quad \boldsymbol{y}_i^{(2)} = \boldsymbol{x}_i^{(2)} - \frac{\delta}{\sqrt{n}}\boldsymbol{x}_i^{(2)} + \frac{\delta}{\sqrt{n}}M_2\boldsymbol{x}_i^{(1)}$$

we get

$$\frac{\boldsymbol{y}_i^{(1)}}{||\boldsymbol{y}_i^{(1)}||} = \frac{\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(1)}||} - \frac{\delta}{\sqrt{n}}\frac{\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(1)}||}\frac{\boldsymbol{x}_i^{(1)^T}M_1\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(1)}||^2} + \frac{\delta}{\sqrt{n}}\frac{M_1\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(1)}||} + o_P(n^{-1/2})$$

and

$$\frac{\boldsymbol{y}_i^{(2)}}{||\boldsymbol{y}_i^{(2)}||} = \frac{\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(2)}||} - \frac{\delta}{\sqrt{n}}\frac{\boldsymbol{x}_i^{(2)}}{||\boldsymbol{x}_i^{(2)}||}\frac{\boldsymbol{x}_i^{(2)^T}M_2\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(2)}||^2} + \frac{\delta}{\sqrt{n}}\frac{M_2\boldsymbol{x}_i^{(1)}}{||\boldsymbol{x}_i^{(2)}||} + o_P(n^{-1/2}).$$

But then in the spherical case,

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\boldsymbol{S}}_i^{(1)}\tilde{\boldsymbol{S}}_i^{(2)^T} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\boldsymbol{x}_i^{(1)}\boldsymbol{x}_i^{(2)^T}}{||\boldsymbol{x}_i^{(1)}||||\boldsymbol{x}_i^{(2)}||} - \frac{\delta}{n}\sum_{i=1}^{n}\frac{\boldsymbol{x}_i^{(1)}\boldsymbol{x}_i^{(1)^T}M_2^T\boldsymbol{x}_i^{(2)}\boldsymbol{x}_i^{(2)^T}}{||\boldsymbol{x}_i^{(1)}||||\boldsymbol{x}_i^{(2)}||^3} \\
&+ \frac{\delta}{n}\sum_{i=1}^{n}\frac{\boldsymbol{x}_i^{(1)}\boldsymbol{x}_i^{(1)^T}M_2^T}{||\boldsymbol{x}_i^{(1)}||||\boldsymbol{x}_i^{(2)}||} - \frac{\delta}{n}\sum_{i=1}^{n}\frac{\boldsymbol{x}_i^{(1)}\boldsymbol{x}_i^{(1)^T}M_1\boldsymbol{x}_i^{(2)}\boldsymbol{x}_i^{(2)^T}}{||\boldsymbol{x}_i^{(2)}||||\boldsymbol{x}_i^{(1)}||^3} + \frac{\delta}{n}\sum_{i=1}^{n}\frac{M_1\boldsymbol{x}_i^{(2)}\boldsymbol{x}_i^{(2)^T}}{||\boldsymbol{x}_i^{(2)}||||\boldsymbol{x}_i^{(1)}||} \\
&+ o_P(1) \xrightarrow{P} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\boldsymbol{x}_i^{(1)}\boldsymbol{x}_i^{(2)^T}}{||\boldsymbol{x}_i^{(1)}||||\boldsymbol{x}_i^{(2)}||} \\
&+ \frac{\delta}{pq}\left((p-1)E(||\boldsymbol{x}_i^{(2)}||)E(||\boldsymbol{x}_i^{(1)}||^{-1})M_1 + (q-1)E(||\boldsymbol{x}_i^{(1)}||)E(||\boldsymbol{x}_i^{(2)}||^{-1})M_2^T\right),
\end{aligned}$$

which completes the proof.

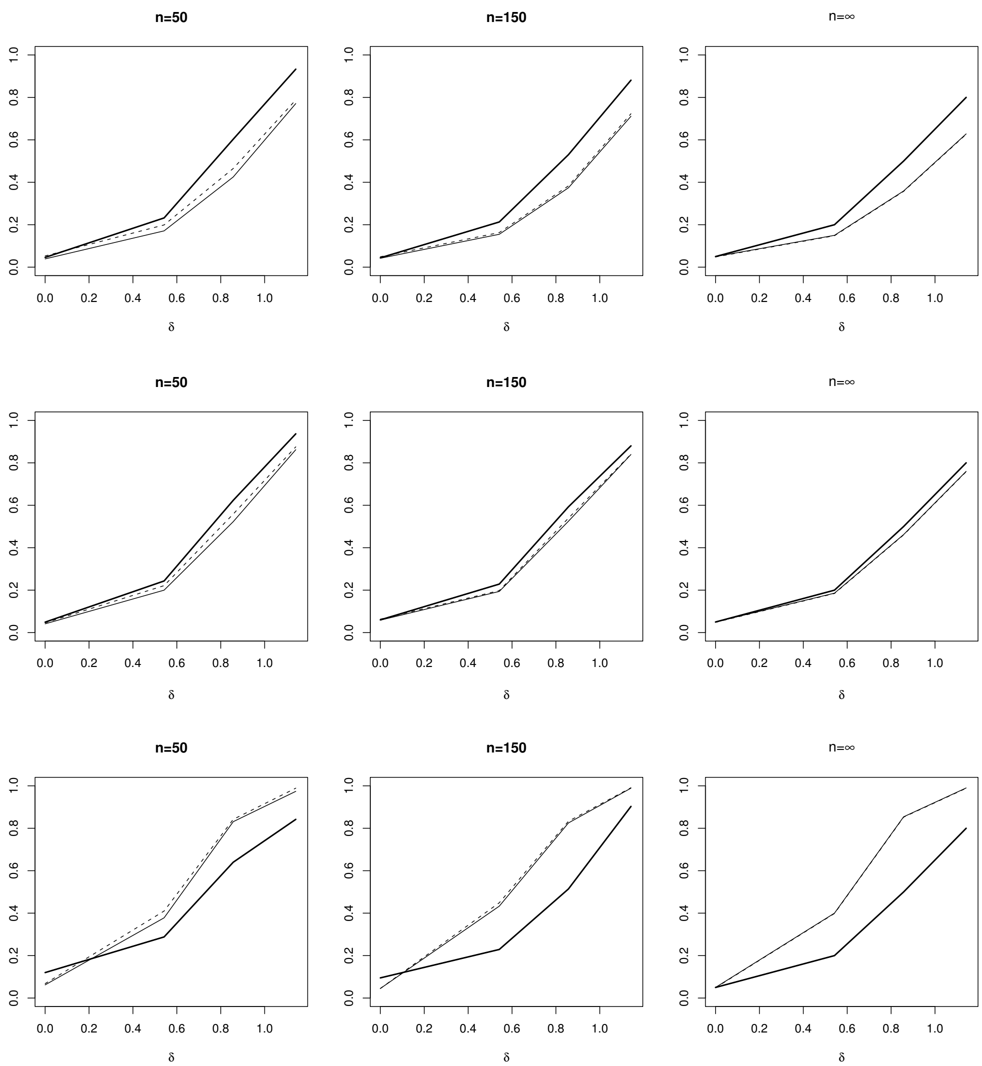

Figure 2: Empirical powers for $p = q = 3$ using the multivariate normal distribution (first row), multivariate t distribution with five degrees of freedom (second row) and contaminated normal distribution with $\epsilon = 0.2$ and $c = 6$ (third row). The thick solid line denotes W_n, the thin solid line U_n and the dotted line Q_n.

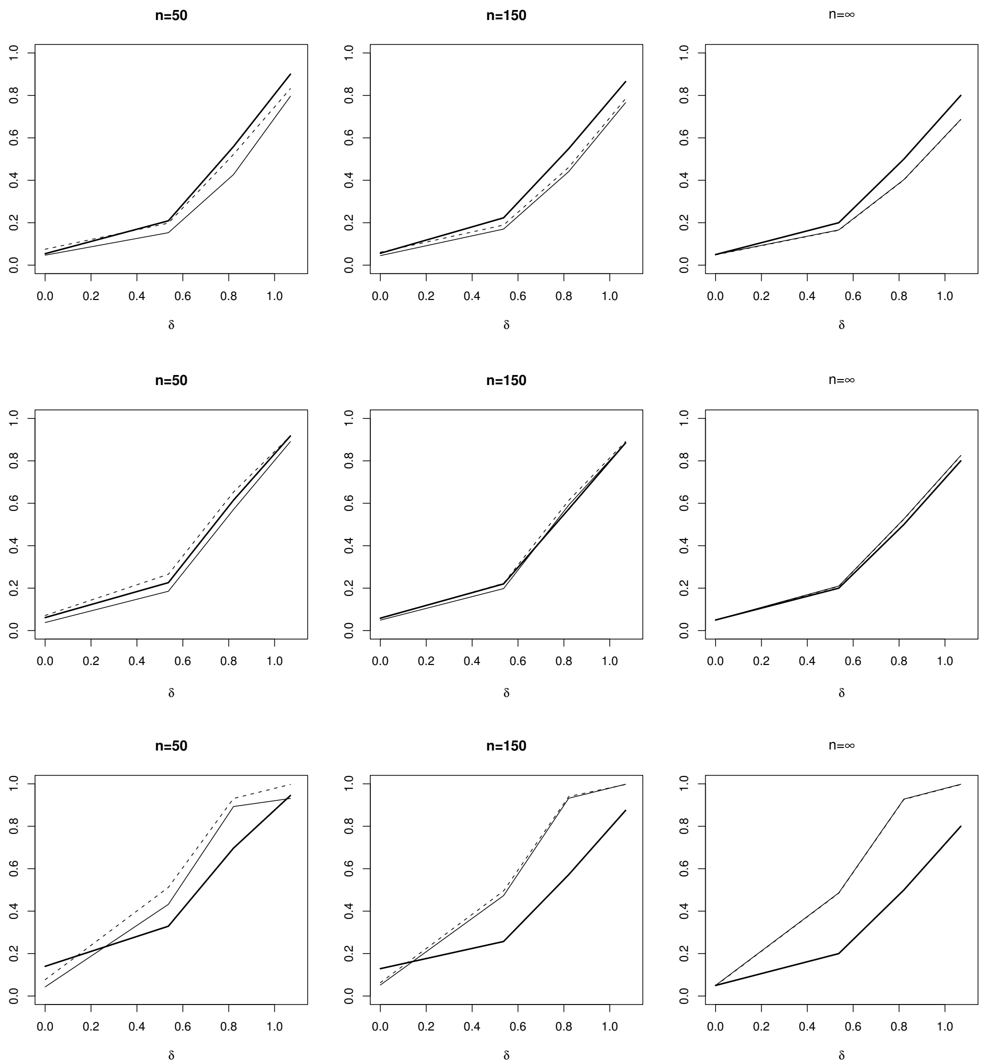

Figure 3: Empirical powers for $p = q = 5$ using the multivariate normal distribution (first row), multivariate t distribution with five degrees of freedom (second row) and contaminated normal distribution with $\epsilon = 0.2$ and $c = 6$ (third row). The thick solid line denotes W_n, the thin solid line U_n and the dotted line Q_n.