KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0210

INFLUENCE FUNCTION AND ASYMPTOTIC EFFICIENCY OF THE AFFINE EQUIVARIANT RANK COVARIANCE MATRIX

by
E. OLLILA
C. CROUX
H. OJA

D/2002/2376/10

Influence Function and Asymptotic Efficiency of the Affine Equivariant Rank Covariance Matrix

Esa Ollila[1] Christophe Croux[2] Hannu Oja[1]

January 23, 2002

[1]*Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35, FIN-40351 Jyväskylä, Finland*
(esa.ollila@maths.jyu.fi, hannu.oja@maths.jyu.fi)

[2]*Department of Applied Economics, K.U. Leuven, Naamsestraat 69, B-3000 Leuven, Belgium*
(christophe.croux@econ.kuleuven.ac.be)

Abstract

Visuri *et al.* (2001) proposed and illustrated the use of the affine equivariant rank covariance matrix (RCM) in classical multivariate inference problems. The RCM was shown to be asymptotically multinormal but explicit formulas for the limiting variances and covariances were not given yet. In this paper the influence functions and the limiting variances and covariances of the RCM and the corresponding scatter estimate are derived in the multivariate elliptic case. Limiting efficiencies are given in the multivariate normal and t distribution cases. The estimates based on the RCM are highly efficient in the multinormal case, and for heavy tailed distribution, perform better than those based on the regular covariance matrix.

Key words: Elliptic distribution; influence function; multivariate analysis; multivariate rank; scatter matrix

1 Introduction

Ranks and signs are frequently used in statistical analysis to obtain procedures which are less sensible to the model assumptions. Computing statistical quantities based on ranks instead of on the original observations can result in more nonparametric and robust methods. A simple example hereof is the Spearmann rank correlation. When observations are multivariate, it is not so obvious anymore how the sign and the rank of an observation are defined. In the paper the affine equivariant multivariate extension of the concept of rank as proposed by Brown and Hettmansperger (1987) is considered. This concept of rank is based on the Oja (1983) median, and has been successfully applied to multivariate analogies of one-sample, two-sample and multisample sign and rank tests. For example, nonparametric and robust competitors of MANOVA have been developed in Hettmansperger *et al.* (1998). For a review of statistical methods based on ranks we refer to Hettmansperger and McKean (1998) and Oja (1999).

The approach based on multivariate signs and ranks has recently been extended to other classical multivariate inference problems, such as principal component analysis, canonical correlation analysis and multivariate regression analysis. These developments are based on the affine equivariant multivariate sign and rank covariance matrices, as defined in Visuri *et al.* (2000). For the affine equivariant sign covariance matrix (SCM), the asymptotic distribution and asymptotic variances were obtained by Ollila *et al.* (2001b). Knowledge of the limit distribution of the SCM allowed to obtain asymptotic results for multivariate regression based on the SCM (Ollila *et al.* 2001a). Multivariate inference based on the affine equivariant rank covariance matrix (RCM) was proposed, outlined and illustrated in Visuri *et al.* (2001). Their simulation studies and examples showed that the estimates based on the RCM enjoy very good efficiency properties, at the price of not being highly robust with respect to extreme outliers. The asymptotic variances of the RCM, however, were not computed yet. These asymptotic variances are the key-quantities for determining the asymptotic distribution of estimators for principal components analysis, canonical correlation analysis and multivariate regression based on the rank covariance matrix.

The main contribution of this paper is that the limiting variance of the RCM have been computed. Moreover, we also obtained an expression for the influence function of the RCM. This influence function is seen to be approximately linear, in contrast with the influence function of the regular covariance matrix, the latter being quadratic. The RCM is therefore more robust than the classical covariance matrix, but has still an unbounded influence function. Despite that, we will show that the RCM remains quite efficient at heavy tailed distributions.

In Section 2 the concept and properties of the affine equivariant rank based on the Oja objective function (1983) are briefly reviewed. The rank covariance matrix and corresponding scatter matrix estimator are defined in Section 3. Influence functions of the estimators, at elliptical model distributions, are given in Section 4 . The limiting variances and covariances of the estimates in the elliptic case are presented in Section 5. The paper is closed with some final comments in Section 6.

2 Affine Equivariant Ranks

Let $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ be a k-variate data set. Then the volume of the k-variate simplex determined by $k+1$ vertices $\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_k}$ (k data points) and $\boldsymbol{x}$ is a constant $(1/k!)$ times

$$V(\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_k}, \boldsymbol{x}) = \text{abs}\left\{\det\begin{pmatrix} 1 & \cdots & 1 & 1 \\ \boldsymbol{x}_{i_1} & \cdots & \boldsymbol{x}_{i_k} & \boldsymbol{x} \end{pmatrix}\right\} \tag{1}$$

and the affine equivariant Oja (1983) median minimizes the criterion function

$$V(\boldsymbol{x}; X) = \text{ave}\{V(\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_k}, \boldsymbol{x})\},$$

where the average is taken over all possible k-subsets $\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_k}$ with $1 \leq i_1 < \ldots < i_k \leq n$. The multivariate centered **rank function** is defined as the gradient

$$\boldsymbol{R}(\boldsymbol{x}; X) = \nabla_{\boldsymbol{x}} V(\boldsymbol{x}; X).$$

Note that in the univariate case $V(x; X) = \text{ave}|x - x_i|$, the mean deviation, gives the univariate median and the centered rank function $R(x; X) = \text{ave}\{\text{sign}(x - x_i)\}$. The **observed ranks**

$$\boldsymbol{R}_i = \boldsymbol{R}(\boldsymbol{x}_i; X), \quad i = 1, \ldots, n,$$

are centered, so $\sum_i \boldsymbol{R}_i = \boldsymbol{0}$, and affine equivariant in the sense that if the ranks $\boldsymbol{R}_i^*$ are calculated from the transformed observations $\boldsymbol{x}_i^* = A\boldsymbol{x}_i + \boldsymbol{b}$, with A a nonsingular $k \times k$-matrix and $\boldsymbol{b}$ a k-vector, then

$$\boldsymbol{R}_i^* = \text{abs}\{\det(A)\}(A^{-1})^T \boldsymbol{R}_i.$$

The population counterparts are as follows. If $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is a random sample from a k-variate distribution with cdf F with finite first order moments, then the expected volume of the simplex is a constant $(1/k!)$ times

$$V(\boldsymbol{x}; F) = E_F[V(\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_k}, \boldsymbol{x})].$$

The multivariate centered **population rank function** is then

$$\boldsymbol{R}(\boldsymbol{x}; F) = \nabla_{\boldsymbol{x}} V(\boldsymbol{x}; F). \tag{2}$$

Naturally also the population rank function is affine equivariant and $E_F[\boldsymbol{R}(\boldsymbol{x};F)] = \mathbf{0}$. The empirical rank function $\boldsymbol{R}(\boldsymbol{x};X)$ converges uniformly in probability to the population rank function $\boldsymbol{R}(\boldsymbol{x};F)$. In the univariate case $R(x;F) = 2F(x) - 1$. See Oja (1999), Hettmansperger *et al.* (1998) and Visuri *et al.* (2001).

Consider now the population rank function at a spherical model distribution G. If $\boldsymbol{x}$ follows a spherically symmetric distribution G, then its radius $r = \|\boldsymbol{x}\|$ and direction $\boldsymbol{u} = \boldsymbol{x}/\|\boldsymbol{x}\|$ are independent and $\boldsymbol{u}$ is uniformly distributed on the periphery of a unit sphere. Now $V(\boldsymbol{x};G)$ depends on $\boldsymbol{x} = r\boldsymbol{u}$ only through r. Hence, in this case, we may write $V(\boldsymbol{x};G) = V_0(r;G)$. Then the population rank function at G is simply

$$\boldsymbol{R}(\boldsymbol{x};G) = \nabla_{\boldsymbol{x}} V_0(r;G) = q(r;G)\,\boldsymbol{u} \tag{3}$$

with $q(r;G) = V_0'(r;G)$. Expressions for the functions $V_0(r;G)$ and $q(r;G)$ at spherical normal and t-distributions are given in Lemma 1 of the Appendix. Next consider the elliptical case. Let the distribution G of $\boldsymbol{z}$ be spherical with mean vector $\mathbf{0}$ and covariance matrix I_k and write $\boldsymbol{x} = \Sigma^{1/2}\boldsymbol{z} + \boldsymbol{\mu}$, with Σ a positive definite $k \times k$-matrix. Then the distribution F of $\boldsymbol{x}$ is elliptically symmetric with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. Due to affine equivariance, the population rank function of F at $\boldsymbol{x} = \Sigma^{1/2}\boldsymbol{z} + \boldsymbol{\mu}$ is

$$\boldsymbol{R}(\boldsymbol{x};F) = \text{abs}\{\det(\Sigma^{1/2})\}\Sigma^{-1/2}\boldsymbol{R}(\boldsymbol{z};G).$$

3 The Rank Covariance Matrix (RCM)

Let $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_n$ be the observed ranks for a k-variate data set $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$. The **rank covariance matrix (RCM)** is then

$$\widehat{D} = \text{ave}\{\boldsymbol{R}_i\boldsymbol{R}_i^T\}.$$

Since the ranks are centered, the RCM is nothing else but a usual convariance matrix computed from the ranks instead of from the original observations. It is affine equivariant in the sense that if the the RCM $\widehat{D}^*$ is calculated from the transformed observations $\boldsymbol{x}_i^* = A\boldsymbol{x}_i + \boldsymbol{b}$, with nonsingular A, then

$$\widehat{D}^* = \det(A)^2(A^{-1})^T\widehat{D}A^{-1}. \tag{4}$$

Visuri *et al.* (2001) showed that if $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is a random sample from a k-variate distribution with cdf F with finite first order moments, then the rank covariance matrix converges in probability to the **population rank covariance matrix**,

$$D(F) = E_F[\boldsymbol{R}(\boldsymbol{x};F)\boldsymbol{R}^T(\boldsymbol{x};F)]. \tag{5}$$

Consider now a spherical distribution G. Then (3) together with $E_G[\boldsymbol{u}\boldsymbol{u}^T] = I_k/k$ yields

$$D(G) = \frac{c_G^2}{k} I_k\,, \tag{6}$$

where $c_G^2 = E_G[q^2(r;G)]$. For values of c_G^2 in multivariate spherical normal and t distribution cases, see Lemma 2 in the Appendix. Take now $\boldsymbol{z} \sim G$, with G spherical with mean vector $\boldsymbol{0}$ and covariance matrix I_k. Then $\boldsymbol{x} = \Sigma^{1/2}\boldsymbol{z} + \boldsymbol{\mu}$ follows an elliptically symmetric distribution F with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. By (6) and the affine equivariance property (4) we get

$$D(F) = \det(\Sigma)\Sigma^{-1/2}\,D(G)\,\Sigma^{-1/2} = (c_G^2/k)\det(\Sigma)\Sigma^{-1}. \tag{7}$$

Thus, at elliptical models, the RCM is proportional to the inverse of the regular covariance matrix. (This is in fact true also in a much wider class of distributions, the so called location-scale model, Visuri *et al.* 2001.) The inverse of the RCM is therefore an estimator of a multiple of the scatter matrix Σ and we may say that it estimates the shape of Σ.

But RCM also carries information about the size of the date cloud, and one can construct an affine equivariant **scatter matrix functional** $C(F)$ based on the RCM:

$$C(F) = \left[\frac{\det\{D(F)\}}{c_G^2/k}\right]^{1/(k-1)} D(F)^{-1}. \tag{8}$$

It is immediate to check that $C(F) = \Sigma$ at elliptical distributions F, so C is Fisher consistent for Σ at elliptical models. For example, at the k-variate normal model ($G = \Phi_k$)

$$\widehat{C} = \left[\frac{\det(\widehat{D})}{c_{\Phi_k}^2/k}\right]^{1/(k-1)} \widehat{D}^{-1}$$

(with $c_{\Phi_k}^2$ given in Lemma 1 of the Appendix) is a consistent estimator of the population covariance matrix Σ. Moreover, the above estimator is affine equivariant in the sense that $\widehat{C}^*$ computed from the transformed observations $\boldsymbol{x}_i^* = A\boldsymbol{x}_i + \boldsymbol{b}$ verifies

$$\widehat{C}^* = A\widehat{C}A^T\,. \tag{9}$$

Therefore we may call $\widehat{C}$ a scatter matrix estimator, which can be compared with other estimators of multivariate scatter (see Maronna and Yohai (1998) for an overview of different scatter matrix estimators).

4 Influence Functions in the Elliptical Case

In this section we derive the influence functions of the RCM functional D and the associated scatter functional C at elliptical models. The influence function (IF) of a functional T at a distribution F measures the effect of an infinitesimal contamination at a single point $\boldsymbol{x}$ on T. For that, consider the contaminated distribution

$$F_\varepsilon = (1-\varepsilon)F + \varepsilon\Delta_{\boldsymbol{x}}$$

where $\Delta_{\boldsymbol{x}}$ is a distribution putting all its mass at $\boldsymbol{x}$. Then the influence function is defined by (see Hampel *et al.*, 1986)

$$\mathrm{IF}(\boldsymbol{x};T,F) = \lim_{\varepsilon\downarrow 0}\frac{T(F_\varepsilon)-T(F)}{\varepsilon} = \frac{\partial}{\partial\varepsilon}T(F_\varepsilon)\big|_{\varepsilon=0}.$$

The next theorem gives an expression for the influence function of the RCM at a spherical model distribution. The proof is quite technical and can be found in the Appendix.

Theorem 1 *Let G be a k-variate spherical distribution. Then the influence function of the RCM functional D at G is given by*

$$\begin{aligned}\mathrm{IF}(\boldsymbol{x};D,G) &= \left\{q^2(r;G)+\gamma(r;G)-\eta(r;G)\right\}\boldsymbol{u}\boldsymbol{u}^T - \left\{2k+1-\frac{\eta(r;G)}{c_G^2/k}\right\}D(G)\\ &= \alpha(r;G)\,\boldsymbol{u}\boldsymbol{u}^T - \beta(r;G)\,D(G)\end{aligned}$$

where $r=\|\boldsymbol{x}\|$, $\boldsymbol{u}=\|\boldsymbol{x}\|^{-1}\boldsymbol{x}$, $D(G)=(c_G^2/k)I_k$ and $q(r;G)$ is defined by (3). Furthermore

$$\begin{aligned}\gamma(r;G) &= 2kE_G\left[\frac{(z_1-r)z_1}{\|\boldsymbol{z}_r\|\,\|\boldsymbol{z}\|}\left\{V_0(\rho_{\boldsymbol{z}};G_{k-1})-\rho_{\boldsymbol{z}}\,q(\rho_{\boldsymbol{z}};G_{k-1})\right\}q(\|\boldsymbol{z}\|;G)\right],\\ \eta(r;G) &= 2kE_G\left[\frac{z_2^2}{\|\boldsymbol{z}_r\|\,\|\boldsymbol{z}\|}\left\{V_0(\rho_{\boldsymbol{z}};G_{k-1})+\frac{(r^2-\rho_{\boldsymbol{z}}^2)}{\rho_{\boldsymbol{z}}}\,q(\rho_{\boldsymbol{z}};G_{k-1})\right\}q(\|\boldsymbol{z}\|;G)\right]\end{aligned}$$

where $\boldsymbol{z}=(z_1,\dots,z_k)^T\sim G$, G_{k-1} is the $k-1$-variate spherical distribution of $(z_1,\dots,z_{k-1})$,

$$\boldsymbol{z}_r = (z_1-r,z_2,\dots,z_k)^T \quad \text{and} \quad \rho_{\boldsymbol{z}}^2 = r^2\left\{1-\frac{(z_1-r)^2}{\|\boldsymbol{z}_r\|^2}\right\}.$$

The influence function of scatter functional $C(F)$ is obtained next. Croux and Haesbroeck (2000) showed that the influence function of any affine equivariant scatter estimator at a spherical model G may be expressed as

$$\mathrm{IF}(\boldsymbol{x};C,G) = \tilde{\alpha}(\|\boldsymbol{x}\|;G)\boldsymbol{u}\boldsymbol{u}^T - \tilde{\beta}(\|\boldsymbol{x}\|;G)I_k, \tag{10}$$

with again $\boldsymbol{u} = \boldsymbol{x}/\|\boldsymbol{x}\|$, for some real-valued functions $\tilde{\alpha}$ and $\tilde{\beta}$ depending on the estimator and the model. Using Theorem 1 and (8), the functions $\tilde{\alpha}$ and $\tilde{\beta}$ of the RCM scatter functional C are easily obtained using matrix differentiation rules.

Corollary 1 *Using the notations of Theorem 1, the influence function of the scatter functional C at a k-variate spherical distribution G is determined by setting in* (10)

$$\tilde{\alpha}(r;G) = -\frac{\alpha(r;G)}{c_G^2/k} \quad \textit{and} \quad \tilde{\beta}(r;G) = \frac{1}{k-1}\left\{2k+1-\frac{q^2(r;G)}{c_G^2/k}-\frac{\gamma(r;G)}{c_G^2/k}\right\}.$$

In Figure 1a we pictured the functions $\tilde{\alpha}(r;G)$ of the regular covariance matrix estimator and the scatter estimator C based on the RCM at the bivariate normal model. (Note that both estimators are comparable, since they estimate the same population quantity Σ.) The $\tilde{\alpha}$ of the regular covariance matrix is quadratic in the radius r, while that of $C(F)$ is approximately linear for large r. This shows that the RCM will give more protection to outliers than the regular covariance matrix. Surely, its influence function remains unbounded, so the RCM is not a robust method in the strict sense. The RCM resembles an L_1-based method: more robust than an L_2 based approach, very efficient (as we will see in the next section), but not highly robust.

In Figure 1b we see the $\tilde{\beta}(r;G)$ functions of the regular covariance matrix estimator and the scatter estimator C based on the RCM at the bivariate normal model. This function is much less important, since it does not intervene in the influence function of the off-diagonal elements of C. It only measures the influence on the estimation of the size of the scatter matrix, not on the shape. For example, the influence function of the correlation matrix estimator associated with C will solely depend on $\tilde{\alpha}(r;G)$ (Croux and Haesbroeck 2000, Ollila *et al.* 2001b).

Remark: Due to the affine equivariance properties, the influence functions of the RCM functional D and the associated scatter matrix functional C at an elliptical distribution F with mean $\boldsymbol{\mu}$ and covariance Σ are simply given by

$$\mathrm{IF}(\boldsymbol{x};D,F) = \det(\Sigma)\Sigma^{-1/2}\,\mathrm{IF}\big(\Sigma^{-1/2}(\boldsymbol{x}-\boldsymbol{\mu});D,G\big)\,\Sigma^{-1/2}$$

and

$$\mathrm{IF}(\boldsymbol{x};C,F) = \Sigma^{1/2}\mathrm{IF}\big(\Sigma^{-1/2}(\boldsymbol{x}-\boldsymbol{\mu});C,G\big)\Sigma^{1/2}.$$

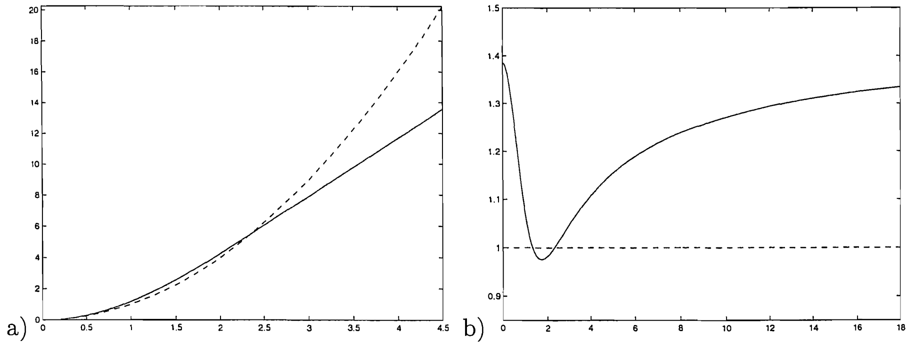

Figure 1: Functions a) $\tilde{\alpha}(r;G)$ and b) $\tilde{\beta}(r;G)$ of the RCM scatter estimator (solid line) and the regular covariance estimator (dashed line) at the bivariate normal model ($G=\Phi_2$).

5 Limiting Variances and Covariances in the Elliptical Case

We start this section with some notational conventions. We use "vec" as operator working on matrices: vec(A) vectorize matrix A by stacking the columns of the matrix on top of each other. A commutation matrix, $I_{k,k}$, is a $k^2 \times k^2$ block matrix with (i,j)-block being equal to a $k \times k$ matrix having 1 at entry (j,i) and zero elsewhere. Finally, the Kronecker product of two $k \times k$ matrices A and B, denoted by $A \otimes B$, is a $k^2 \times k^2$-block matrix with $k \times k$-blocks, the (i,j)-block equal to $a_{ij}B$. For more information on Kronecker products, commutation matrices and the vec-operator, the reader is referred to Magnus and Neudecker (1988).

Visuri *et al.* (2001) showed that, if the observations come from a k-variate distribution with finite second order moments, then the limiting distribution of $\sqrt{n}\,\mathrm{vec}(\widehat{D}-D)$ is multivariate normal with zero mean. To compute the asymptotic covariance matrix, we use

$$\mathrm{ASV}(\widehat{D};F) = E[\mathrm{vec}\{\mathrm{IF}(\boldsymbol{x};D,F)\}\mathrm{vec}\{\mathrm{IF}(\boldsymbol{x};D,F)\}^T].$$

The structure of the influence function of D at spherical distributions, and the symmetry properties of G imply that $\mathrm{ASV}(\widehat{D};G)$ will only depend on two numbers: $\mathrm{ASV}(\widehat{D}_{11};G)$ and $\mathrm{ASV}(\widehat{D}_{12};G)$. The asymptotic covariances between on-diagonal elements are all equal to

$$\mathrm{ASC}(\widehat{D}_{11},\widehat{D}_{22};G) = \mathrm{ASV}(\widehat{D}_{11};G) - 2\mathrm{ASV}(\widehat{D}_{12};G),$$

while all the other limiting covariances are zero. Similar developments also hold true for $\widehat{C}$ and we get

Corollary 2 *The covariance matrices of the limiting distribution of $\sqrt{n}\,\mathrm{vec}(\widehat{D}-D)$ and $\sqrt{n}\,\mathrm{vec}(\widehat{C}-C)$ at a spherical distribution G are given by*

$$\mathrm{ASV}(\widehat{D}_{12};G)(I_{k^2}+I_{k,k})+\mathrm{ASC}(\widehat{D}_{11},\widehat{D}_{22};G)\mathrm{vec}(I_k)\mathrm{vec}(I_k)^T$$

and

$$\mathrm{ASV}(\widehat{C}_{12};G)(I_{k^2}+I_{k,k})+\mathrm{ASC}(\widehat{C}_{11},\widehat{C}_{22};G)\mathrm{vec}(I_k)\mathrm{vec}(I_k)^T,$$

respectively.

Using the affine equivariance (and properties of vec-operator and Kronecker product), the limiting covariance matrix of $\sqrt{n}\mathrm{vec}(\widehat{D}-D)$ at an elliptical F with mean vector $\boldsymbol{\mu}$ and covariance matrix $\Sigma = AA^T$, being the distribution of $A\boldsymbol{z}+\boldsymbol{\mu}$ with $\boldsymbol{z}\sim G$, is given by

$$\begin{aligned}\mathrm{ASV}(\widehat{D};F) &= \det(A)^4(A^{-1}\otimes A^{-1})^T\mathrm{ASV}(\widehat{D};G)(A^{-1}\otimes A^{-1})\\ &= \frac{k^2}{c_G^4}\left[\mathrm{ASV}(\widehat{D}_{12};G)(I_{k^2}+I_{k,k})(D\otimes D)+\mathrm{ASC}(\widehat{D}_{11},\widehat{D}_{22};G)\mathrm{vec}(D)\mathrm{vec}(D)^T\right]\end{aligned}$$

and the limiting covariance matrix of $\sqrt{n}\mathrm{vec}(\widehat{C}-\Sigma)$ by

$$\mathrm{ASV}(\widehat{C}_{12};G)(I_{k^2}+I_{k,k})(\Sigma\otimes\Sigma)+\mathrm{ASC}(\widehat{C}_{11},\widehat{C}_{22};G)\mathrm{vec}(\Sigma)\mathrm{vec}(\Sigma)^T.$$

In fact, the above expressions for the asymptotic covariance matrices are valid for any asymptotically normal affine equivariant scatter matrix estimate $\widehat{C}$, not only for the RCM-based one. Therefore we will measure the relative efficiency of a scatter matrix estimate $\widehat{C}$ with respect to the regular covariance matrix estimate $\widehat{S}$ by the two ratios

$$\mathrm{ARE}(\widehat{C}_{11},\widehat{S}_{11};G)=\frac{\mathrm{ASV}(\widehat{S}_{11};G)}{\mathrm{ASV}(\widehat{C}_{11};G)} \quad\text{and}\quad \mathrm{ARE}(\widehat{C}_{12},\widehat{S}_{12};G)=\frac{\mathrm{ASV}(\widehat{S}_{12};G)}{\mathrm{ASV}(\widehat{C}_{12};G)},$$

being the asymptotic relative efficiencies of on-diagonal and off-diagonal elements of the scatter matrix estimate $\widehat{C}$ with respect to the sample covariance matrix $\widehat{S}$. In our case, the limiting variances are readily obtained from Theorem 1 and Corollary 1. For example, $\mathrm{ASV}(\widehat{C}_{12};G)=(k^2/c_G^4)\mathrm{ASV}(\widehat{D}_{12};G)$, where

$$\mathrm{ASV}(\widehat{D}_{12};G)=E_G\big[\mathrm{IF}(\boldsymbol{x};D_{12},G)^2\big]=E_G\big[\alpha^2(\|\boldsymbol{x}\|;G)u_1^2u_2^2\big]=\frac{E_G[\alpha^2(\|\boldsymbol{x}\|;G)]}{k(k+2)},$$

Dimension	A. 5	Degrees of freedom 6	8	15	∞	B. 5	Degrees of freedom 6	8	15	∞
2	2.05	1.47	1.20	1.06	0.99	2.33	1.61	1.27	1.08	0.98
3	2.00	1.44	1.18	1.05	0.99	2.26	1.57	1.24	1.07	0.98
5	1.92	1.38	1.14	1.02	0.97	2.14	1.50	1.20	1.04	0.96
8	1.88	1.37	1.14	1.04	0.99	2.11	1.48	1.20	1.06	0.99
10	1.86	1.36	1.13	1.03	0.97	2.08	1.48	1.19	1.05	0.97

Table 1: Asymptotic Relative Efficiency of the off-diagonal (panel A) and on-diagonal (panel B) elements of the RCM-based scatter matrix estimator with respect to the classical covariance matrix estimator at multivariate $t_{k,\nu}$ distributions for several values of the dimension k and the degrees of freedom ν. The column $\nu = \infty$ corresponds to the standard normal distribution.

which can be calculated using numerical integration or Monte-Carlo techniques.

In Table 1 the on-diagonal and off-diagonal asymptotic relative efficiencies of the RCM-based scatter matrix estimator are obtained for multivariate t-distributions $t_{k,\nu}$. Dimensions $k = 2, 3, 5, 8, 10$ and degrees of freedom $\nu = 5, 6, 8, 15$ are considered. Efficiencies for multivariate normal distributions ($\nu \to \infty$) are also given. First we note that the ARE for the RCM-based scatter matrix estimator is surprisingly high at the normal distribution. There is almost no loss in efficiency, all ARE being above 97%. These numbers are clearly superior to the efficiencies of high breakdown robust estimators like the Minimum Covariance Determinant estimator or S-estimators (tabulated in Croux and Haesbroeck 1999).

For multivariate t-distributions, the RCM-based scatter matrix estimator outperforms the classical covariance matrix. The gain gets large when the degrees of freedom increase, i.e. when the distribution gets heavier tails. Ollila *et al.* (2001b) also reported the on-diagonal and off-diagonal efficiencies of the scatter estimate based on the affine equivariant Sign Covariance Matrix (SCM). The efficiency of the rank based estimates are somewhat better in all reported cases, except for dimension 5 where SCM has a slightly higher efficiency.

6 Final Comments

Classical multivariate analysis is based on the sample mean vector and sample covariance matrix. To robustify the inference procedures, the mean vector and covariance matrix have often been replaced by robust affine equivariant location vector and scatter matrix estimates. Influence functions are then used for robustness considerations and derivations of the limiting variances and covariances of the estimates.

At elliptical models, two quantities, namely efficiencies of the on-diagonal and off-diagonal elements, fully characterize the efficiency properties of an affine equivariant scatter matrix. In the multivariate multiple regression problem, for example, the off-diagonal efficiency gives the efficiency of the regression coefficient estimate based on the scatter matrix estimate. In principal component analysis, the on-diagonal and off-diagonal efficiencies yield the efficiencies of the corresponding eigenvalue and eigenvector estimates. In the canonical correlation analysis, the efficiency of the canonical correlations is given by the off-diagonal efficiency, and the efficiency of the canonical vectors depend on both on-diagonal and off-diagonal efficiencies. See e.g. Croux and Haesbroeck (2000), Van Aelst *et al.* (2000), Croux *et al.* (2001) and Taskinen *et al* (2002). The asymptotic efficiencies of the on-diagonal and off-diagonal elements of the RCM have now been obtained in this paper.

Comparing between different robust estimators of multivariate scatter is a difficult job. The attractivity of the rank covariance matrix can be found in its high efficiency, even at heavy tailed distribution, and in its close relationship to existing rank concepts, which gives it a non-parametric flavor. Of course, the RCM is not meant to be a competitor with high breakdown scatter matrices in terms of robustness. It is also remarkable, that no location estimate is needed to construct the RCM. C-programs for calculating the ranks and the rank covariance matrix are available on web site http://www.cc.jyu.fi/~esaolli/.

A Appendix

A.1 Expressions for $q(r;G)$, $V_0(r;G)$ and c_G^2 at Spherical Normal and t Distributions

Before stating the expressions, we recall the definition of a power series important in our construction:

Definition 1 *A generalized hypergeometric series is defined as*

$$ {}_pF_q(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; z) = \sum_{i=0}^{\infty} \frac{(a_1)_i (a_2)_i \ldots (a_p)_i}{(b_1)_i (b_2)_i \ldots (b_q)_i} \frac{z^i}{i!}, $$

where $(c)_i = c(c+1)\cdots(c+i-1) = \Gamma(c+i)/\Gamma(c)$.

The next 2 Lemmas states the expressions for $V_0(r;G)$, $q(r;G)$, and c_G^2 for G a multivariate normal and a t-distribution.

Lemma 1 *In the k-variate standard normal case,* $G = \Phi_k$,

$$V_0(r;\Phi_k) = \frac{2^{k/2}\Gamma(\frac{k+1}{2})}{\sqrt{\pi}} \exp(-kr^2/2)\, {}_1F_1\left(\frac{k+1}{2}; \frac{k}{2}; \frac{kr^2}{2}\right),$$

$$q(r;\Phi_k) = r\,\frac{2^{k/2}\Gamma(\frac{k+1}{2})}{\sqrt{\pi}} \exp(-kr^2/2)\, {}_1F_1\left(\frac{k+1}{2}; \frac{k+2}{2}; \frac{kr^2}{2}\right).$$

In the k-variate spherical t-distribution with ν *degrees of freedom,* $G = t_{k,\nu}$,

$$V_0(r;t_{k,\nu}) = \frac{c_{k,\nu}}{(1+r^2/\nu)^{k(\nu-1)/2}}\, {}_2F_1\left(\frac{k+1}{2}, \frac{k(\nu-1)}{2}; \frac{k}{2}; \frac{r^2/\nu}{1+r^2/\nu}\right),$$

$$q(r;t_{k,\nu}) = \frac{c_{k,\nu}\, r(\nu-1)}{\nu(1+r^2/\nu)^{k(\nu-1)/2+1}}\, {}_2F_1\left(\frac{k+1}{2}, \frac{k(\nu-1)+2}{2}; \frac{k+2}{2}; \frac{r^2/\nu}{1+r^2/\nu}\right),$$

where

$$c_{k,\nu} = \frac{\nu^{k/2}\Gamma(\frac{k+1}{2})\Gamma^k(\frac{\nu-1}{2})}{\Gamma^k(\frac{\nu}{2})\sqrt{\pi}}.$$

Lemma 2 *The constant* c_G^2 *in (6) for the standard normal distribution* $G = \Phi_k$ *and for student distributions* $G = t_{k,\nu}$ *is given by*

$$c_{\Phi_k}^2 = \frac{k2^k\Gamma^2(\frac{k+1}{2})}{\pi(k+1)}\, {}_2F_1\left(\frac{1}{2}, \frac{1}{2}; \frac{k+2}{2}; \frac{k^2}{(k+1)^2}\right),$$

$$c_{t_{k,\nu}}^2 = d_{k,\nu} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\Gamma(\frac{k+1}{2}+i)\Gamma(\frac{k(\nu-1)+2}{2}+i)\Gamma(\frac{k+1}{2}+j)\Gamma(\frac{k(\nu-1)+2}{2}+j)\Gamma(\frac{k+2}{2}+i+j)}{\Gamma(\frac{k+2}{2}+i)\Gamma(\frac{k+2}{2}+j)\Gamma(\frac{2k\nu-k+\nu+4}{2}+i+j)\,i!\,j!},$$

where

$$d_{k,\nu} = \frac{\nu^{k-1}\Gamma(\frac{k}{2})\Gamma^{2k}(\frac{\nu-1}{2})\Gamma(\frac{k+\nu}{2})\Gamma(\frac{2k(\nu-1)+\nu+2}{2})}{\Gamma^2(\frac{k(\nu-1)}{2})\Gamma^{2k+1}(\frac{\nu}{2})\pi}.$$

For example, $c_{\Phi_k}^2 = 0.712, 7.681, 203.749$ for dimensions $k = 2, 4, 6$ respectively. These results are as in Möttönen *et al.* (1998), but slightly simplified.

A.2 Proofs and Additional Lemmas

The prove the Theorem 1 we need the following Lemma:

Lemma 3 *Let r be a constant scalar and write $\boldsymbol{v} = (1, 0, \dots, 0)^T$ for a unit k-vector. Let G denote a cdf of a k-variate spherical random vector $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \dots, x_{ik})^T$ and let G_{k-1} denote a cdf of a $(k-1)$-variate spherical random (sub)vector $\boldsymbol{x}_i' = (x_{i2}, \dots, x_{ik})^T$. Then*

$$E_G[V(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k, \boldsymbol{x})|\boldsymbol{x}_1 = r\boldsymbol{v}, \boldsymbol{x}] = \|\boldsymbol{x} - r\boldsymbol{v}\| V_0(\rho_{\boldsymbol{x}}; G_{k-1}),$$

where

$$\rho_{\boldsymbol{x}}^2 = r^2 \left\{ 1 - \frac{(x_1 - r)^2}{\|\boldsymbol{x} - r\boldsymbol{v}\|^2} \right\}$$

and $V(\cdot)$ and $V_0(\cdot)$ are defined by (1) and (3), respectively.

Proof of Lemma 3: Let P be a $k \times k$ *orthogonal* (rotation) matrix (hence $PP^T = P^T P = I_k$ and $\text{abs}\{\det(P^T)\} = 1$) such that $P(\boldsymbol{x} - r\boldsymbol{v}) = (\|\boldsymbol{x} - r\boldsymbol{v}\|, 0, \dots, 0)^T$. Then

$$P = \begin{pmatrix} \boldsymbol{p}_1^T \\ P_2 \end{pmatrix} = \begin{pmatrix} \|\boldsymbol{x} - r\boldsymbol{v}\|^{-1}(\boldsymbol{x} - r\boldsymbol{v})^T \\ P_2 \end{pmatrix}$$

where P_2 is a $(k-1) \times k$-matrix. By symmetry, it is equivalent to solve the expectation:

$$\begin{aligned}
&E_G[V(\boldsymbol{x}_1, P^T\boldsymbol{x}_2 \dots, P^T\boldsymbol{x}_k, \boldsymbol{x})|\boldsymbol{x}_1 = r\boldsymbol{v}, \boldsymbol{x}] \\
&= E_G\left[\text{abs}\left\{\det\left(P^T\boldsymbol{x}_2 - \boldsymbol{x}_1 \quad \dots \quad P^T\boldsymbol{x}_k - \boldsymbol{x}_1 \quad \boldsymbol{x} - \boldsymbol{x}_1\right)\right\} |\boldsymbol{x}_1 = r\boldsymbol{v}, \boldsymbol{x}\right] \\
&= E_G\left[\text{abs}\left\{\det\left(\boldsymbol{x}_2 - P\boldsymbol{x}_1 \quad \dots \quad \boldsymbol{x}_k - P\boldsymbol{x}_1 \quad P(\boldsymbol{x} - \boldsymbol{x}_1)\right)\right\} |\boldsymbol{x}_1 = r\boldsymbol{v}, \boldsymbol{x}\right] \text{abs}\{\det(P^T)\} \\
&= E_G\left[\text{abs}\left\{\det\begin{pmatrix} x_{21} - \boldsymbol{p}_1^T\boldsymbol{x}_1 & \dots & x_{k1} - \boldsymbol{p}_1^T\boldsymbol{x}_1 & \|\boldsymbol{x} - \boldsymbol{x}_1\| \\ \boldsymbol{x}_2' - P_2\boldsymbol{x}_1 & \dots & \boldsymbol{x}_k' - P_2\boldsymbol{x}_1 & \boldsymbol{0} \end{pmatrix}\right\} \middle| \boldsymbol{x}_1 = r\boldsymbol{v}, \boldsymbol{x}\right] \\
&= \|\boldsymbol{x} - r\boldsymbol{v}\| \, E_G\left[\text{abs}\left\{\det\left(\boldsymbol{x}_2' - P_2\boldsymbol{x}_1 \quad \dots \quad \boldsymbol{x}_k' - P_2\boldsymbol{x}_1\right)\right\} |\boldsymbol{x}_1 = r\boldsymbol{v}\right] \\
&= \|\boldsymbol{x} - r\boldsymbol{v}\| \, E_G\left[\text{abs}\left\{\det\begin{pmatrix} 1 & \dots & 1 & 1 \\ \boldsymbol{x}_2' & \dots & \boldsymbol{x}_k' & P_2\boldsymbol{x}_1 \end{pmatrix}\right\} \middle| \boldsymbol{x}_1 = r\boldsymbol{v}\right] \\
&= \|\boldsymbol{x} - r\boldsymbol{v}\| \, V_0(\|P_2 r\boldsymbol{v}\|; G_{k-1}).
\end{aligned}$$

From $\|Pr\boldsymbol{v}\|^2 = r^2$, we obtain the relation

$$(\boldsymbol{p}_1^T r\boldsymbol{v})^2 + \|P_2 r\boldsymbol{v}\|^2 = r^2$$

and as $\boldsymbol{p}_1^T r\boldsymbol{v} = \|\boldsymbol{x} - r\boldsymbol{v}\|^{-1}(x_1 - r)r$, it follows that

$$\|P_2 r\boldsymbol{v}\|^2 = r^2 \left\{ 1 - \frac{(x_1 - r)^2}{\|\boldsymbol{x} - r\boldsymbol{v}\|^2} \right\} = \rho_{\boldsymbol{x}}^2,$$

which completes the proof. □

Proof of Theorem 1: Let $X = \{\boldsymbol{x}_{i_1}, \dots, \boldsymbol{x}_{i_k}, \boldsymbol{x}_{j_1}, \dots, \boldsymbol{x}_{j_k}, \boldsymbol{x}_i\}$ denote the data set of $2k+1$ i.i.d. observations from the k-variate spherical distribution G. Further, write $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_k)$ for the k-sets of indices. Then, let the the scalar $d_0(I)$ and the k-vector $\boldsymbol{d}(I) = (d_1(I), \dots, d_k(I))^T$ denote the cofactors corresponding to the last column of the matrix

$$\det \begin{pmatrix} 1 & \dots & 1 & 1 \\ \boldsymbol{x}_{i_1} & \dots & \boldsymbol{x}_{i_k} & \boldsymbol{x}_i \end{pmatrix}$$

and write $S_I(\boldsymbol{x}_i) = \text{sign}\{d_0(I) + \boldsymbol{d}(I)^T \boldsymbol{x}_i\}$. Then, by reversing the order of the expectation and the differentiation, equation (2) can be rewritten as

$$\boldsymbol{R}(\boldsymbol{x}_i; G) = E_G[S_I(\boldsymbol{x}_i)\boldsymbol{d}(I)|\boldsymbol{x}_i].$$

Then note that

$$\begin{aligned} D(G) &= E_{\boldsymbol{x}_i}[\boldsymbol{R}(\boldsymbol{x}_i; G)\boldsymbol{R}(\boldsymbol{x}_i; G)] = E_{\boldsymbol{x}_i}\left[E_G[S_I(\boldsymbol{x}_i)S_J(\boldsymbol{x}_i)\boldsymbol{d}(I)\boldsymbol{d}(J)^T|\boldsymbol{x}_i]\right] \\ &= E_G[S_I(\boldsymbol{x}_i)S_J(\boldsymbol{x}_i)\boldsymbol{d}(I)\boldsymbol{d}(J)^T]. \end{aligned}$$

It is now straightforward to show that the influence function of D at G is

$$\begin{aligned} &\text{IF}(\boldsymbol{x}; D, G) \\ &= \frac{\partial}{\partial \varepsilon} \int \cdots \int \int S_I(\boldsymbol{x}_i)S_J(\boldsymbol{x}_i)\boldsymbol{d}(I)\boldsymbol{d}(J)^T \, dG_\varepsilon(\boldsymbol{x}_{i_1}) \cdots dG_\varepsilon(\boldsymbol{x}_{j_k})dG_\varepsilon(\boldsymbol{x}_i)\Big|_{\varepsilon=0} \\ &= \boldsymbol{R}(\boldsymbol{x}; G)\boldsymbol{R}^T(\boldsymbol{x}; G) + 2kE_G[S_I(\boldsymbol{x}_i)S_J(\boldsymbol{x}_i)\boldsymbol{d}(I)\boldsymbol{d}(J)^T|\boldsymbol{x}_{i_1} = \boldsymbol{x}] - (2k+1)D(G), \end{aligned} \quad (11)$$

where $G_\varepsilon = (1-\varepsilon)G + \varepsilon\Delta_{\boldsymbol{x}}$ is the contaminated distribution.

Next we derive the influence function (11) of D for a point in the direction of the first axis, that is, we set $\boldsymbol{x} = r\boldsymbol{v}$, where $\boldsymbol{v} = (1, 0, \dots, 0)^T$. Using the fact that a spherical random variable $\boldsymbol{x} = (x_i)_{1 \le i \le k}$ and $(s_i x_{\pi(i)})_{1 \le i \le k}$ has the same distribution for arbitrary $s_i \in \{-1, 1\}$ and permutation π of $\{1, 2, \dots, k\}$, one immediately finds that

$$2k\, E_G[S_I(\boldsymbol{x}_i)S_J(\boldsymbol{x}_i)\boldsymbol{d}(I)\boldsymbol{d}(J)^T|\boldsymbol{x}_{i_1} = r\boldsymbol{v}] = \begin{pmatrix} \gamma(r; G) & \boldsymbol{0}^T \\ \boldsymbol{0} & \eta(r; G)I_{k-1} \end{pmatrix},$$

where

$$\begin{aligned} \gamma(r; G) &= 2k\, E_G[S_I(\boldsymbol{x}_i)S_J(\boldsymbol{x}_i)d_1(I)d_1(J)|\boldsymbol{x}_{i_1} = r\boldsymbol{v}] \\ &= 2k\, E_{\boldsymbol{x}_i}\left[E_G[S_I(\boldsymbol{x}_i)d_1(I)|\boldsymbol{x}_{i_1} = r\boldsymbol{v}, \boldsymbol{x}_i]\, q(\|\boldsymbol{x}_i\|; G)\frac{x_{i1}}{\|\boldsymbol{x}_i\|}\right] \end{aligned}$$

as $E_G[S_J(\boldsymbol{x}_i)\boldsymbol{d}(J)|\boldsymbol{x}_i] = \boldsymbol{R}(\boldsymbol{x}_i;G) = q(\|\boldsymbol{x}_i\|;G)\boldsymbol{x}_i/\|\boldsymbol{x}_i\|$ by (3). Similarly, we obtain

$$\eta(r;G) = 2kE_{\boldsymbol{x}_i}\left[E_G[S_I(\boldsymbol{x}_i)d_2(I)|\boldsymbol{x}_{i_1} = r\boldsymbol{v}, \boldsymbol{x}_i]\, q(\|\boldsymbol{x}_i\|;G)\frac{x_{i2}}{\|\boldsymbol{x}_i\|}\right].$$

Next, note that

$$\begin{aligned}
E_G[S_I(\boldsymbol{x}_i)d_1(I)|\boldsymbol{x}_{i_1} = r\boldsymbol{v}, \boldsymbol{x}_i] &= \frac{\partial}{\partial x_{i1}} E_G[V(\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_k}, \boldsymbol{x}_i)|\boldsymbol{x}_{i_1} = r\boldsymbol{v}, \boldsymbol{x}_i] \\
&= \frac{\partial}{\partial x_{i1}} \|\boldsymbol{x}_i - r\boldsymbol{v}\| V_0(\rho_{\boldsymbol{x}_i}; G_{k-1}) \\
&= \frac{x_{i1} - r}{\|\boldsymbol{x}_i - r\boldsymbol{v}\|}\{V_0(\rho_{\boldsymbol{x}_i}; G_{k-1}) - \rho_{\boldsymbol{x}_i}\, q(\rho_{\boldsymbol{x}_i}; G_{k-1})\},
\end{aligned}$$

where the first equality follows from reversing the order of expectation and differentiation, the second equality follows from Lemma 3 and the third equality follows by simple diffrerentiation rules (use the chain rule to obtain $\partial V_0(\rho_{\boldsymbol{x}_i}; G_{k-1})/\partial x_{i1} = q(\rho_{\boldsymbol{x}_i}; G_{k-1})\partial \rho_{\boldsymbol{x}_i}/\partial x_{i1}$ (as $V_0' = q$) and $\partial \rho_{\boldsymbol{x}_i}/\partial x_{i1} = -(x_{i1} - r)\rho_{\boldsymbol{x}_i}\|\boldsymbol{x}_i - r\boldsymbol{v}\|^{-2}$). Similarly, one can show that

$$E_G[S_I(\boldsymbol{x}_i)d_2(I)|\boldsymbol{x}_{i_1} = r\boldsymbol{v}, \boldsymbol{x}_i] = \frac{x_{i2}}{\|\boldsymbol{x}_i - r\boldsymbol{v}\|}\left\{V_0(\rho_{\boldsymbol{x}_i}; G_{k-1}) + \frac{(r^2 - \rho^2_{\boldsymbol{x}_i})}{\rho_{\boldsymbol{x}_i}} q(\rho_{\boldsymbol{x}_i}; G_{k-1})\right\}.$$

This then yields the stated expressions for $\gamma(r;G)$ and $\eta(r;G)$.

Then, as $\boldsymbol{R}(r\boldsymbol{v};G) = q(r;G)\boldsymbol{v}$, we may now write the influence function (11) of D for a point $\boldsymbol{x} = r\boldsymbol{v}$ as

$$\begin{aligned}
\mathrm{IF}(r\boldsymbol{v}; D, G) &= q^2(r;G)\boldsymbol{v}\boldsymbol{v}^T + \begin{pmatrix}\gamma(r;G) & \boldsymbol{0}^T \\ \boldsymbol{0} & \eta(r;G)I_{k-1}\end{pmatrix} - (2k+1)D(G) \\
&= \{q^2(r;G) + \gamma(r;G) - \eta(r;G)\}\boldsymbol{v}\boldsymbol{v}^T + \eta(r;G)I_k - (2k+1)D(G) \qquad (12)
\end{aligned}$$

An influence point in an arbitrary direction is obtained by setting $\boldsymbol{x} = Pr\boldsymbol{v} = r\boldsymbol{u}$ for a well chosen orthogonal ($PP^T = I_k$) rotation matrix $P = [\boldsymbol{u}\ \boldsymbol{u}_2 \cdots \boldsymbol{u}_k]$. The influence function is then given by

$$\mathrm{IF}(\boldsymbol{x}; D, G) = P\, \mathrm{IF}(r\boldsymbol{v}; D, G)\, P^T,$$

which, by (12) and relations $P\boldsymbol{v} = \boldsymbol{u}$, $PP^T = I_k$ and $D(G) = (c_G^2/k)I_k$ reduces to

$$\mathrm{IF}(\boldsymbol{x}; D, G) = \left\{q^2(r;G) + \gamma(r;G) - \eta(r;G)\right\}\boldsymbol{u}\boldsymbol{u}^T - \left\{2k + 1 - \frac{\eta(r;G)}{c_G^2/k}\right\} D(G),$$

which completes the proof. □

Proof of Corollary 2: First we note that for a random vector $\boldsymbol{u} = (u_1, \ldots, u_k)^T$ uniformly distributed on the perihery of the unit sphere, one has that

$$E[\boldsymbol{u}\boldsymbol{u}^T] = k^{-1}I_k,\ E[u_i^2u_j^2] = (k(k+2))^{-1},\ E[u_i^4] = 3(k(k+2))^{-1},\ E[u_i^3u_j] = 0, \qquad (13)$$

where u_i and u_j are distinct elements of $\boldsymbol{u}$. Let $\boldsymbol{x} \sim G$ and write $r = \|\boldsymbol{x}\|$, $\boldsymbol{u} = \boldsymbol{x}/\|\boldsymbol{x}\|$ and recall that r and $\boldsymbol{u}$ are independent. Then,

$$\begin{aligned}
\text{ASV}(\widehat{C};G) &= E[\text{vec}\{\text{IF}(\boldsymbol{x};C,F)\}\text{vec}\{\text{IF}(\boldsymbol{x};C,F)\}^T] \\
&= E[\text{vec}\{\tilde{\alpha}(r;G)\boldsymbol{u}\boldsymbol{u}^T - \tilde{\beta}(r;G)I_k\}\text{vec}\{\tilde{\alpha}(r;G)\boldsymbol{u}\boldsymbol{u}^T - \tilde{\beta}(r;G)I_k\}^T] \\
&= E[\tilde{\alpha}^2(r;G)]\, E[\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)^T] \\
&\quad - E[\tilde{\alpha}(r;G)\tilde{\beta}(r;G)]\, E[\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)\text{vec}(I_k)^T] \\
&\quad - E[\tilde{\alpha}(r;G)\tilde{\beta}(r;G)]\, E[\text{vec}(I_k)\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)^T] \\
&\quad + E[\tilde{\beta}^2(r;G)]\, \text{vec}(I_k)\text{vec}(I_k)^T.
\end{aligned}$$

Using (13), it is easy to show that

$$E[\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)^T] = (k(k+2))^{-1}\left[(I_{k^2} + I_{k,k}) + \text{vec}(I_k)\text{vec}(I_k)^T\right]$$

and

$$E[\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)\text{vec}(I_k)^T] = E[\text{vec}(I_k)\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)^T] = k^{-1}\text{vec}(I_k)\text{vec}(I_k)^T,$$

so $\text{ASV}(\widehat{C};G)$ can be written simply as

$$\text{ASV}(\widehat{C};G) = \text{ASV}(\widehat{C}_{12};G)(I_{k^2} + I_{k,k}), + \text{ASC}(\widehat{C}_{11},\widehat{C}_{22};G)\text{vec}(I_k)\text{vec}(I_k)^T,$$

since

$$\text{ASV}(\widehat{C}_{12};G) = E[\text{IF}(\boldsymbol{x};C_{12},G)^2] = E[\tilde{\alpha}^2(r;G)]E[u_1^2u_2^2] = E[\tilde{\alpha}^2(r;G)](k(k+2))^{-1}$$

and

$$\begin{aligned}
\text{ASC}(\widehat{C}_{11},\widehat{C}_{22};G) &= E[\{\tilde{\alpha}(r;G)u_1^2 - \tilde{\beta}(r;G)\}\{\tilde{\alpha}(r;G)u_2^2 - \tilde{\beta}(r;G)\}] \\
&= (k(k+2))^{-1}E[\tilde{\alpha}^2(r;G)] - 2k^{-1}E[\tilde{\alpha}^2(r;G)\tilde{\beta}(r;G)] + E[\tilde{\beta}^2(r;G)].
\end{aligned}$$

Naturally, the developements for $\text{ASV}(\widehat{D};G)$ are similar. □

References

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

Croux, C., and Haesboeck, G. (1999). Influence function and efficiency of the minimum covariance determinant scatter matrix estimator. *J. Multivariate Anal.,* **71**, 161-190.

Croux, C and Haesbroeck, G. (2000). Principal component analysis based on robust estimators of the covariance or correlation matrix : influence functions and efficiencies. *Biometrika*, **87**, 3, 603-618.

Croux, C., Dehon, C., Rousseeuw, P.J., and Van Aelst, S. (2001). Robust estimation of the conditional median at elliptical models. *Statist. Probab. Lett.*, **51** 361-368.

Hampel, F., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. (1986). *Robust Statistics: The Approach Based on Influence Functions.* Wiley, New York.

Hettmansperger, T.P. and McKean, J. (1998). *Robust Nonparametric Statistical Methods.* Arnold, London.

Hettmansperger, T.P., Möttönen, J. and Oja, H. (1998). Affine invariant multivariate rank tests for several samples. *Statist. Sinica*, **8**, 765-800.

Magnus, J.R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Wiley, New York.

Maronna, R.A., and Yohai, V. (1998). Robust Estimation of Multivariate Location and Scatter, in *Encyclopedia of Statistical Sciences Update Volume 2*, edited by S. Kotz, C. Read, and D. Banks. John Wiley, New York: 589-596.

Möttönen, J., Hettmansperger, T.P., Oja, H. and Tienari, J. (1998). On the efficiency of affine invariant multivariate rank tests. *J. Multivariate Anal.*, **66**, 118-132.

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

Oja, H. (1999). Affine equivariant multivariate sign and rank tests and corresponding estimates: A review. *Scand. J. Statist.*, **26**, 319-343.

Ollila, E., Hettmansperger, T.P. and Oja, H. (2001a). Estimates of regression coefficients based on the sign covariance matrix. Submitted. http://www.cc.jyu.fi/~esaolli/.

Ollila, E., Oja, H. and Croux, C. (2001b). The Affine equivariant sign covariance matrix: Asymptotic behavior and efficiencies. Submitted. http://www.cc.jyu.fi/~esaolli/.

Taskinen, S., Croux, C., Kankainen, A., Ollila, E. and Oja, H. (2002). Canonical analysis based on scatter and shape matrices. In preparation.

Van Aelst, S., Van Driessen, K., and Rousseeuw, P. (2000). A robust method for multivariate regression. In: *Data Analysis, Classification, and Related Methods*, eds. by Kiers, H.A.L. et al. Springer, Heidelberg.

Visuri,S., Koivunen, V., and Oja, H. (2000). Sign and rank covariance matrices. *J. Statist. Plann. Inference*, **91**, 557-575.

Visuri,S., Ollila, E., Koivunen, V., Möttönen, J. and Oja, H. (2001). Affine equivariant multivariate rank methods. To appear in *J. Statist. Plann. Inference.* http://www.cc.jyu.fi/~esaolli/.