Rank Covariance Matrix For A Partially Known Covariance Matrix

Kristi Kuljus and Dietrich von Rosen

Research Report
Centre of Biostochastics

Swedish University of Agricultural Sciences

Report 2006:1
ISSN 1651-8543

Rank Covariance Matrix
For A Partially Known Covariance Matrix

Kristi Kuljus[1]

Department of Mathematics
Uppsala University
PO Box 480, SE-751 06 Uppsala, Sweden

Dietrich von Rosen

Centre of Biostochastics
Swedish University of Agricultural Sciences
Box 7013, SE-750 07 Uppsala, Sweden

Abstract

Classical multivariate methods are often based on the sample covariance matrix, which is very sensitive to outlying observations. One alternative to the covariance matrix is the affine equivariant rank covariance matrix (RCM) that has been studied for example in Visuri *et al.* (2003). In this article we assume that the covariance matrix is partially known and study how to estimate the corresponding RCM. We use the properties that the RCM is affine equivariant and that the RCM is proportional to the inverse of the regular covariance matrix, and reduce the problem of estimating the RCM to estimating marginal rank covariance matrices. This is a great advantage when the dimension of the original data vectors is large.

Key words: multivariate ranks, rank covariance matrix, elliptical distributions, affine equivariance.

[1]E-mail address to the correspondence author: kristi@math.uu.se

1 Introduction

In classical multivariate analysis it is assumed that the underlying distribution is multivariate normal. In this case the sample mean and covariance matrix are sufficient statistics and UMVU estimators and standard multivariate methods are based on the sample covariance matrix. However, the sample covariance matrix is very sensitive to outlying observations, which leads to searching for more robust methods. Besides, the assumption about multivariate normal distribution is restrictive, because it does not allow the data to come from a distribution with heavier tails. In this work we assume that our data are elliptically distributed. Simple simulation studies show how sensitive the sample covariance matrix is to observations from the tails, when we have data from a multivariate distribution with heavier tails, e.g. a t-distribution. This affirms again that alternatives to the standard sample covariance matrix are needed.

In the univariate case robust statistics are often based on ordering the data and defined via rank statistics. In the multivariate case the concept of ordering becomes much more complicated. Since there is no obvious method for complete ordering, different restricted ordering or subordering principles are used. For the issue of multivariate ordering principles, see the classical paper by Barnett (1976). Since there is no unique way of ordering multivariate data, the definition of rank can also be extended to the multivariate case in several ways. Puri & Sen (1985) gave the simplest definition: they defined multivariate ranks through the usual univariate ranks for each marginal sample, i.e. via coordinatewise ranks. In this article we will concentrate on the affine equivariant ranks based on the Oja median (see Oja, 1983). The definition of these affine equivariant ranks can be found for example in Visuri *et al.* (2003). The rank covariance matrix based on these ranks has a nice property: it is proportional to the inverse of the regular covariance matrix. It follows that the RCM can be used in many applications instead of the regular covariance matrix.

Visuri *et al.* (2003) studied the use of the RCM in classical multivariate analysis methods, e.g. multivariate analysis of variance, principal component analysis, multivariate regression analysis and canonical correlation analysis. However, Visuri *et al.* (2003) did not make any assumptions about the structure of the underlying covariance matrix. In this paper we assume that we have zero blocks in the covariance matrix. We are going to study the problem where we partition a random vector into three subvectors and where the covariance matrix of any two of the three subvectors is a null-matrix. We are interested in how to estimate the corresponding rank covariance matrix in this case. Zero blocks in the covariance matrix mean a restriction to the parameter space. Knowing that some of the parameters are equal to zero gives us additional information about the parameter space and we should use this information in estimating the rest of the parameters.

In Section 2 we give some definitions and results concerning the elliptical distributions and the affine equivariant RCM. In Section 3 we study how zero blocks in a covariance matrix affect the corresponding RCM. To estimate this “structured" RCM, we will use the affine equivariance property of the RCM and find such a linear transformation that the RCM of the transformed data will be block-diagonal. Section 4 deals with how to estimate a block-diagonal RCM using the corresponding marginal vectors and marginal rank covariance matrices. In the end of Section 4 we present two examples of how one can use the RCM instead of the usual covariance matrix.

2 On elliptical distributions and affine equivariant ranks

We give some definitions and results that we will need throughout this work. In this paper, we consider the class of elliptical distributions. Let $X \stackrel{d}{=} Y$ denote that X and Y have the same distribution. Elliptical distributions or elliptically symmetric distributions belong to the class of location-scale families which are defined as follows.

Definition 2.1. *Let* $\mathbf{z} : p \times 1$ *be reflection and permutation invariant, i.e.* $\mathbf{Hz} \stackrel{d}{=} \mathbf{z}$ *for every reflection and permutation matrix* $\mathbf{H}$. *Then the corresponding location-scale family is given by the distributions of*

$$\mathbf{x} = \mathbf{Az} + \mathbf{b},$$

where $\mathbf{A} : p \times p$ *is nonsingular and* $\mathbf{b}$ *is a* p*-vector.*

The location-scale family corresponding to a reflection and permutation invariant vector $\mathbf{z}$ is obtained from $\mathbf{z}$ through affine transformations, which are linear transformations followed by a translation. A reflection matrix is a diagonal matrix with diagonal elements equal to 1 or -1, a permutation matrix is obtained from an identity matrix by permuting its rows or columns.

The class of elliptical distributions can be defined in several ways. This class includes for example the class of multivariate normal distributions and multivariate t-distributions. One may say that elliptically symmetric distributions are extensions of the multivariate normal distribution. In Fang & Zhang (1990) we can find the following definition for the class of elliptically contoured distributions.

Definition 2.2. *If the characteristic function of a p-dimensional random vector* $\mathbf{x}$ *has the form*

$$\exp(i\mathbf{t}'\boldsymbol{\mu})\phi(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}),$$

for some function ϕ, *where* $\boldsymbol{\mu} : p \times 1$, $\boldsymbol{\Sigma} : p \times p$ *and* $\boldsymbol{\Sigma} \geq 0$, *we say that* $\mathbf{x}$ *is elliptically distributed with parameters* $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ *and* ϕ *and we write* $\mathbf{x} \sim \mathcal{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$. *In particular, when* $\boldsymbol{\mu} = \mathbf{0}$ *and* $\boldsymbol{\Sigma} = \mathbf{I}_p$, $\mathcal{EC}_p(\mathbf{0}, \mathbf{I}_p, \phi)$ *is called a spherical distribution and denoted by* $\mathcal{S}_p(\phi)$.

A spherically distributed vector $\mathbf{y} : k \times 1$ is invariant with respect to orthogonal transformations, i.e. if $\mathcal{O}(k)$ denotes the set of $k \times k$ orthogonal matrices, then for every $\boldsymbol{\Gamma} \in \mathcal{O}(k)$

$$\boldsymbol{\Gamma}\mathbf{y} \overset{d}{=} \mathbf{y}.$$

In Fang *et al.* (1990) the function ϕ is called the characteristic generator.

Next we present some properties of elliptical distributions which can all be found in Fang *et al.* (1990), for example. One fundamental property of elliptical distributions is that they are invariant under the group of affine transformations which are linear transformations followed by a translation.

Theorem 2.1. *Assume that* $\mathbf{x} \sim \mathcal{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ *with* $\mathrm{rank}(\boldsymbol{\Sigma}) = k$, *let* $\mathbf{B}$ *be a* $p \times m$*-matrix and* $\boldsymbol{\nu}$ *an* m*-vector. Then*

$$\boldsymbol{\nu} + \mathbf{B}'\mathbf{x} \sim \mathcal{EC}_m(\boldsymbol{\nu} + \mathbf{B}'\boldsymbol{\mu}, \mathbf{B}'\boldsymbol{\Sigma}\mathbf{B}, \phi).$$

From this theorem it follows that the marginal vectors of an elliptically distributed vector are also elliptically distributed. Partition $\mathbf{x}$, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ as

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \tag{1}$$

where $\mathbf{x}_1 : m \times 1$, $\boldsymbol{\mu}_1 : m \times 1$, $\boldsymbol{\Sigma}_{11} : m \times m$, $0 < m < p$. Then the following corollary holds.

Corollary 2.2. *Assume that* $\mathbf{x} \sim \mathcal{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$, *then* $\mathbf{x}_1 \sim \mathcal{EC}_m(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \phi)$ *and* $\mathbf{x}_2 \sim \mathcal{EC}_{p-m}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22}, \phi)$.

The last result that we state concerns the parameters of an elliptical distribution.

Theorem 2.3. *Assume that* $\mathbf{x} \sim \mathcal{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ *and that the covariance matrix of* $\mathbf{x}$ *exists. Then*

$$\mathrm{E}\,\mathbf{x} = \boldsymbol{\mu}, \ \mathrm{Cov}\,\mathbf{x} = -2\phi'(0)\boldsymbol{\Sigma}.$$

In this work we will concentrate on affine equivariant ranks based on the Oja median (Oja, 1983). Affine equivariance means that something is equivariant with respect to affine transformations. An equivariant estimator of a parameter or parametric function for example is an estimator, that is transformed in the same way as the parameter or parametric function, when the observations are transformed. For a description of the principle of equivariance, see Lehmann & Casella (1998) and Bondesson (1982).

The Oja median is a spatial median. Since it is the corner-stone behind the definition of the affine equivariant ranks, we present its definition here.

Definition 2.3. *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be a sample from* $\mathbf{x}$ *in* $\mathcal{R}^p$. *Let* $V(\mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_p}, \boldsymbol{\mu})$ *denote the volume of the simplex with vertices* $\mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_p}, \boldsymbol{\mu}$ *in* $\mathcal{R}^p$, *where* $1 \leq i_1 < i_2 < \ldots < i_p \leq n$. *Then the Oja simplex median of the sample* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *is a point* $\hat{\boldsymbol{\mu}}$ *which minimizes*

$$\sum_{i_1 < \ldots < i_p} V(\mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_p}, \boldsymbol{\mu}) \quad ,$$

where the sum is taken over all subsets of integers $1 \leq i_1 < i_2 < \ldots < i_p \leq n$.

We can now look at a more generalized situation. Let $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be a data set of p-variate observation vectors and let $I = \{i_1 \ldots, i_p\}$ refer to a subset of p observations in $\mathbf{X}$. Consider the hyperplane determined by the points in I, which is given by the following equation:

$$\left\{ \mathbf{x} \in \mathcal{R}^p : \ \det \begin{pmatrix} 1 & \cdots & 1 & 1 \\ \mathbf{x}_{i_1} & & \mathbf{x}_{i_p} & \mathbf{x} \end{pmatrix} = 0 \right\}. \qquad (2)$$

The determinant of the $(p+1) \times (p+1)$ matrix in (2) can also be written as $d_0(I) + \mathbf{d}'(I)\mathbf{x}$, where $\mathbf{d}(I)$ is a p-dimensional vector. If $\mathbf{x}$ does not belong to the hyperplane, the set I along with $\mathbf{x}$ determine a p-variate simplex with volume

$$V_I(\mathbf{x}) = \frac{1}{p\,!} \,\mathrm{abs}\{\, d_0(I) + \mathbf{d}'(I)\,\mathbf{x}\,\},$$

where abs$\{a\}$ denotes the absolute value of a. To define multivariate ranks, the centered rank function is defined at first. This definition is based directly on the definition of the Oja median (see Definition 2.3). The multivariate Oja median minimizes the average of the volumes of all possible simplices over I, so it minimizes also

$$D(\mathbf{x}) = p\,!\,\text{ave}_I\{V_I(\mathbf{x})\}.$$

The Oja median is the solution of the gradient function $\nabla D(\mathbf{x}) = \mathbf{0}$. The following definitions of multivariate affine equivariant ranks and the corresponding rank covariance matrix (RCM) can be found for example in Visuri *et al.* (2003).

Definition 2.4. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a random sample from a distribution with c.d.f. F. Then the centered rank function at $\mathbf{x}$ is defined as*

$$\mathbf{R}(\mathbf{x}) = \nabla D(\mathbf{x}) = \text{ ave}_I\{S_I(\mathbf{x})\,\mathbf{d}(I)\},$$

where

$$S_I(\mathbf{x}) = \text{ sgn }(d_0(I) + \mathbf{d}'(I)\,\mathbf{x}).$$

The corresponding population rank function at $\mathbf{x}$ is defined as

$$\mathbf{R}_F(\mathbf{x}) = \text{E}_F\{S_I(\mathbf{x})\,\mathbf{d}(I)\},$$

assuming that the order of the expectation and the differentiation can be reversed. Here the expectation is taken over the vectors in I.

The sample rank vectors are given by

$$\mathbf{R}_i = \mathbf{R}(\mathbf{x}_i), \quad i = 1, \ldots, n.$$

The $\mathbf{R}_i$-vectors satisfy $\sum_i \mathbf{R}_i = \mathbf{0}$, i.e. they are centered. Since the affine equivariant ranks are vectors, they have both magnitude and direction. The most important property of the rank vectors is their affine equivariance. This property is presented in the following lemma (see Visuri *et al.*, 2003).

Lemma 2.4. *Let $\mathbf{R}_1^*, \ldots, \mathbf{R}_n^*$ be the centered ranks for the transformed observations $\mathbf{x}_i^* = \mathbf{A}\mathbf{x}_i + \mathbf{b}$, $i = 1, \ldots, n$, where $\mathbf{A} : p \times p$ is nonsingular. Let $\mathbf{R}_1, \ldots, \mathbf{R}_n$ be the centered rank vectors for the untransformed observations $\mathbf{x}_i$, $i = 1, \ldots, n$. Then*

$$\mathbf{R}_i^* = \text{abs }\{|\mathbf{A}|\}\,(\mathbf{A}^{-1})'\,\mathbf{R}_i, \;\; i = 1, \ldots, n. \tag{3}$$

Here $|\mathbf{A}|$ denotes the determinant of $\mathbf{A}$. Next we give the definitions of the sample and theoretical rank covariance matrices.

Definition 2.5. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a sample from a random vector $\mathbf{x}$ with c.d.f. F. The sample rank covariance matrix is defined as*

$$\hat{\mathbf{D}} = \text{ave}_i\{\mathbf{R}_i \mathbf{R}_i'\}. \tag{4}$$

The corresponding theoretical rank covariance matrix is defined as

$$\mathbf{D}_F = \text{E}_F[\mathbf{R}_F(\mathbf{x})\mathbf{R}_F'(\mathbf{x})].$$

The rank covariance matrices are affine equivariant (see Visuri *et al.*, 2003), this follows directly from equations (3) and (4).

Theorem 2.5. *Let $\hat{\mathbf{D}}^*$ be the sample RCM calculated from the transformed observations $\mathbf{x}_i{}^* = \mathbf{A}\mathbf{x}_i + \mathbf{b}$, $i = 1, \ldots, n$, for some nonsingular $\mathbf{A} : p \times p$. Let $\hat{\mathbf{D}}$ be the sample RCM for the observations $\mathbf{x}_i$, $i = 1, \ldots, n$. Then*

$$\hat{\mathbf{D}}^* = |\mathbf{A}|^2 (\mathbf{A}^{-1})' \hat{\mathbf{D}} (\mathbf{A}^{-1}).$$

The affine equivariance property holds also for the theoretical RCM:

$$\mathbf{D}_F^* = |\mathbf{A}|^2 (\mathbf{A}^{-1})' \mathbf{D}_F (\mathbf{A}^{-1}).$$

The following theorem from Visuri *et al.* (2003) presents the important relationship between a regular covariance matrix and the corresponding rank covariance matrix in location-scale families. We will use this result several times in our work.

Theorem 2.6. *Let $\mathbf{z}$ be a p-dimensional permutation and reflection invariant vector with $\text{Cov}\,\mathbf{z} = \mathbf{I}_p$. Let*

$$\mathbf{x} = \mathbf{P}\boldsymbol{\Lambda}^{1/2}\mathbf{z} + \boldsymbol{\mu}$$

be a random vector in the corresponding location-scale family. Here $\mathbf{P} : p \times p$ is an orthogonal matrix, $\boldsymbol{\Lambda} : p \times p$ is a diagonal matrix and $\boldsymbol{\mu} : p \times 1$ a symmetry center. Then $\text{Cov}\,\mathbf{x} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}' = \boldsymbol{\Psi}$ and the population rank covariance matrix of $\mathbf{x}$ equals to

$$\mathbf{D}_F = \omega\, |\boldsymbol{\Lambda}|\, \mathbf{P}\boldsymbol{\Lambda}^{-1}\mathbf{P}' = \omega\, |\boldsymbol{\Psi}|\, \boldsymbol{\Psi}^{-1}, \tag{5}$$

where ω is a constant that depends on the distribution of $\mathbf{z}$ only.

Equation (5) shows that the rank covariance matrix is proportional to the inverse of the corresponding covariance matrix. It is interesting to notice that the constant ω does not depend only on the type of distribution, it also depends on the dimension of the distribution. One can say that ω contains the information about the class of distributions that we are working with. The expressions for calculating this constant in the multivariate normal distribution and t-distribution are given in Ollila *et al.* (2004). For example, for normal distributions the constant ω equals 0.356, 1.920 and 33.958 for dimensions $k = 2, 4$ and 6, respectively.

3 Rank covariance matrix corresponding to a partially known covariance matrix

Suppose we have a random vector $\mathbf{y} = (\mathbf{y}_1', \mathbf{y}_2', \mathbf{y}_3')'$ from a family of elliptical distributions consisting of three subvectors $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$. Let the mean and covariance matrix of this random vector be

$$\boldsymbol{\mu} = (\boldsymbol{\mu}_1', \boldsymbol{\mu}_2', \boldsymbol{\mu}_3')' \qquad \text{and} \qquad \boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Psi}_{11} & \boldsymbol{\Psi}_{12} & \boldsymbol{\Psi}_{13} \\ \boldsymbol{\Psi}_{21} & \boldsymbol{\Psi}_{22} & \boldsymbol{\Psi}_{23} \\ \boldsymbol{\Psi}_{31} & \boldsymbol{\Psi}_{32} & \boldsymbol{\Psi}_{33} \end{pmatrix}, \tag{6}$$

respectively, where $\boldsymbol{\Psi}$ is positive definite and symmetric. Suppose that $\boldsymbol{\Psi}_{23} = \mathbf{0}$ (and $\boldsymbol{\Psi}_{32} = \mathbf{0}$), i.e. the subvectors $\mathbf{y}_2$ and $\mathbf{y}_3$ are uncorrelated. We denote the corresponding rank covariance matrix by

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} & \mathbf{D}_{13} \\ \mathbf{D}_{21} & \mathbf{D}_{22} & \mathbf{D}_{23} \\ \mathbf{D}_{31} & \mathbf{D}_{32} & \mathbf{D}_{33} \end{pmatrix}.$$

In Theorem 2.6 we stated that a covariance matrix and the corresponding rank covariance matrix for a distribution from a location-scale family are related as follows:

$$\mathbf{D} = \omega\, |\boldsymbol{\Psi}|\, \boldsymbol{\Psi}^{-1}, \tag{7}$$

where the constant ω depends on the distribution of $(\mathbf{y}_1', \mathbf{y}_2', \mathbf{y}_3')'$. Using equation (7) and the assumption that $\boldsymbol{\Psi}_{23} = \mathbf{0}$, we obtain the next theorem.

Theorem 3.1. *Suppose that we have a random vector $(\mathbf{y}_1', \mathbf{y}_2', \mathbf{y}_3')'$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Psi}$ from a location-scale family. Suppose that $\boldsymbol{\Psi}_{23} = \mathbf{0}$. Then the following condition holds for the rank covariance matrix of $\mathbf{y}$:*

$$\mathbf{D}_{23} = \mathbf{D}_{21}\, \mathbf{D}_{11}^{-1}\, \mathbf{D}_{13}.$$

Proof. From matrix theory we know that if $\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$ and $\mathbf{A}^{-1}$ exists, then

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{E}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{E}^{-1} \\ -\mathbf{E}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{E}^{-1} \end{pmatrix}, \tag{8}$$

where $\mathbf{E} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$ is the Schur complement of $\mathbf{A}_{11}$. From (7) it follows that $\mathbf{\Psi} = c\,\mathbf{D}^{-1}$ for some constant c.

Using this relationship we obtain:

$$\begin{aligned} \begin{pmatrix} \mathbf{\Psi}_{22} & \mathbf{\Psi}_{23} \\ \mathbf{\Psi}_{32} & \mathbf{\Psi}_{33} \end{pmatrix} &= c\left[\begin{pmatrix} \mathbf{D}_{22} & \mathbf{D}_{23} \\ \mathbf{D}_{32} & \mathbf{D}_{33} \end{pmatrix} - \begin{pmatrix} \mathbf{D}_{21} \\ \mathbf{D}_{31} \end{pmatrix} \mathbf{D}_{11}^{-1} \begin{pmatrix} \mathbf{D}_{12} & \mathbf{D}_{13} \end{pmatrix}\right]^{-1} \\ &= c\begin{pmatrix} \mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12} & \mathbf{D}_{23} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{13} \\ \mathbf{D}_{32} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{12} & \mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13} \end{pmatrix}^{-1}. \end{aligned}$$

Applying (8), we get

$$\mathbf{\Psi}_{23} = -c\,(\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12})^{-1}(\mathbf{D}_{23} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{13})(\mathbf{E}^*)^{-1},$$

where

$$\begin{aligned} \mathbf{E}^* &= (\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}) - \\ &\quad -(\mathbf{D}_{32} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{12})(\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12})^{-1}(\mathbf{D}_{23} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}). \end{aligned}$$

It follows that if $\mathbf{\Psi}_{23} = \mathbf{0}$, then $\mathbf{D}_{23} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{13} = \mathbf{0}$, and the conclusion follows. □

So our problem becomes how to estimate $\mathbf{D}$ if we know that $\mathbf{D}_{23} = \mathbf{D}_{21}\,\mathbf{D}_{11}^{-1}\,\mathbf{D}_{13}$ (and $\mathbf{D}_{32} = \mathbf{D}_{31}\,\mathbf{D}_{11}^{-1}\,\mathbf{D}_{12}$). This relationship is additional information about the RCM and should be used in estimating $\mathbf{D}$. To solve this problem, we will use the affine equivariance property (Theorem 2.5) of the rank covariance matrix.

Let $\mathbf{y}_1, \ldots, \mathbf{y}_n$ be a random sample from an elliptical vector $\mathbf{y}$. The affine equivariance property states that if $\mathbf{D}^*$ is calculated from the transformed observations

$$\mathbf{y}_i^* = \mathbf{A}\mathbf{y}_i + \mathbf{b}, \qquad i = 1, \ldots, n,$$

for some nonsingular matrix $\mathbf{A}$, then

$$\mathbf{D}^* = |\mathbf{A}|^2\,(\mathbf{A}^{-1})'\,\mathbf{D}\,\mathbf{A}^{-1}.$$

We begin by transforming our random vector $\mathbf{y}$. Let the transformation matrix be

$$\mathbf{A} = \begin{pmatrix} \mathbf{I} & \mathbf{D}_{11}^{-1}\mathbf{D}_{12} & \mathbf{D}_{11}^{-1}\mathbf{D}_{13} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}. \tag{9}$$

Then $|\mathbf{A}| = 1$ and

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{D}_{11}^{-1}\mathbf{D}_{12} & -\mathbf{D}_{11}^{-1}\mathbf{D}_{13} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

We choose a transformation matrix of this form because then the rank covariance matrix of the transformed data will become block-diagonal. When a rank covariance matrix is block-diagonal, the subranks of the corresponding components are uncorrelated. The rank covariance matrix for the transformed data is

$$\mathbf{D}^* = (\mathbf{A}^{-1})'\mathbf{D}\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13} \end{pmatrix},$$

since $|\mathbf{A}|^2 = 1$. Here $\mathbf{D}_{23}^* = \mathbf{0}$, because $\mathbf{D}_{23} = \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$. Thus, we have proved the following result.

Theorem 3.2. *Let* $\mathbf{y}$ *be elliptically distributed with a mean and covariance matrix given in* (6). *Let* $\mathbf{A}$ *be defined as in* (9). *Then the rank covariance matrix of the transformed vector*

$$\mathbf{A}\mathbf{y} = [(\mathbf{y}_1 + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{y}_2 + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{y}_3)', \mathbf{y}_2', \mathbf{y}_3']'$$

equals

$$\mathbf{D}^* = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13} \end{pmatrix}. \tag{10}$$

The covariance matrix of the transformed vector is equal to

$$\mathbf{\Psi}^* = \mathbf{A}\mathbf{\Psi}\mathbf{A}' = \begin{pmatrix} \mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Psi}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Psi}_{33} \end{pmatrix}. \tag{11}$$

Because of the relationship (7) we can write

$$\mathbf{D}^* = \omega\,|\boldsymbol{\Psi}^*|\,(\boldsymbol{\Psi}^*)^{-1} \tag{12}$$

for some constant ω that depends on the distribution of $\mathbf{y}$. As we mentioned earlier, this constant does not depend only on the family of distributions to which our random vector belongs, it depends also on the size of the vector.

4 Estimating a block-diagonal rank covariance matrix

In this section we study the problem of estimating a block-diagonal RCM. To emphasize different components of the problem, we divide the section into three subsections.

In the first subsection we show how our problem can be reduced in dimension: we can use marginal sample vectors to estimate the diagonal blocks $\mathbf{D}_{11}$, $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$. The circumstance that we can use the marginal sample vectors and the corresponding marginal rank covariance matrices in the estimation process, is very important. The computation of the sample RCM's is computer-intensive due to the number of hyperplanes considered. The number of possible hyperplanes one has to regard in order to estimate an RCM increases fast with increasing size of data vectors. Therefore, it is always an advantage to reduce the dimension of data vectors when working with rank covariance matrices.

In the second subsection we deal with estimating the constants ω_1, ω_2 and ω_3. We show that our new estimates of the blocks in the original rank covariance matrix $\mathbf{D}$ contain the constant ω. This entails that in the applications that require the estimate of the RCM up to a proportionality constant, we do not need to bother estimating the constant ω.

In the third subsection we present a couple of examples of how one can use the RCM.

4.1 Estimating a block-diagonal rank covariance matrix via marginal rank covariance matrices

We have reduced our problem to estimating a rank covariance matrix $\mathbf{D}^*$ that is block-diagonal, see (10). When we have estimated the diagonal blocks $\mathbf{D}_{11}$, $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ of $\mathbf{D}^*$, and $\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$, we can estimate all the blocks in the original rank covariance matrix $\mathbf{D}$. From (12) it is easy to see how the diagonal blocks of $\mathbf{D}^*$ and $\mathbf{\Psi}^*$ are related:

$$
\begin{aligned}
\mathbf{D}_{11} &= \omega\,|\mathbf{\Psi}_{22}|\,|\mathbf{\Psi}_{33}|\,|\mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31}| \\
&\quad \times (\mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31})^{-1}, \qquad (13)\\
\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12} &= \omega\,|\mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31}|\,|\mathbf{\Psi}_{33}|\,|\mathbf{\Psi}_{22}|\,\mathbf{\Psi}_{22}^{-1}, \\
\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13} &= \omega\,|\mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31}|\,|\mathbf{\Psi}_{22}|\,|\mathbf{\Psi}_{33}|\,\mathbf{\Psi}_{33}^{-1}.
\end{aligned}
$$

If a random vector $(\mathbf{y}_1', \mathbf{y}_2', \mathbf{y}_3')'$ has a multivariate normal distribution and its covariance matrix $\mathbf{\Psi}$ is block-diagonal, then we know that the components $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$ are independent. In this case marginal sample vectors are used to estimate $\mathbf{\Psi}_{11}$, $\mathbf{\Psi}_{22}$ and $\mathbf{\Psi}_{33}$, respectively. For other elliptical distributions a block-diagonal covariance matrix does not imply independence of the components. Suppose, anyway, that we want to use marginal sample vectors to estimate the blocks $\mathbf{D}_{11}$, $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ in our transformed rank covariance matrix.

Theorem 4.1. *Let the marginal vectors* $\mathbf{y}_1 + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{y}_2 + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{y}_3$, $\mathbf{y}_2$ *and* $\mathbf{y}_3$ *be vectors of dimensions* p_1, p_2 *and* p_3, *respectively. Denote the population rank covariance matrices of these random vectors by* $\mathbf{M}_1$, $\mathbf{M}_2$ *and* $\mathbf{M}_3$, *correspondingly. Then the following equations hold:*

$$
\begin{aligned}
\mathbf{M}_1 &= \omega_1\,|\mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31}| \\
&\quad \times (\mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31})^{-1}, \\
\mathbf{M}_2 &= \omega_2\,|\mathbf{\Psi}_{22}|\,\mathbf{\Psi}_{22}^{-1}, \qquad (14)\\
\mathbf{M}_3 &= \omega_3\,|\mathbf{\Psi}_{33}|\,\mathbf{\Psi}_{33}^{-1}.
\end{aligned}
$$

Proof. From Theorem 2.1 it follows that $\mathbf{Ay}$ is elliptical with the covariance matrix given in (11). Due to Corollary 2.2, the marginal vectors $\mathbf{y}_1 + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{y}_2 + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{y}_3$, $\mathbf{y}_2$ and $\mathbf{y}_3$ are also elliptically distributed with the covariance matrices $\mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31}$, $\mathbf{\Psi}_{22}$ and $\mathbf{\Psi}_{33}$, respectively. So the equations in (14) follow from (7). □

Notice that if the dimensions p_1, p_2 and p_3 are not the same, then the constants ω_1, ω_2 and ω_3 differ. We will discuss estimation of these constants in the following subsection. Meanwhile, let us assume that ω, ω_1, ω_2 and ω_3 are known.

By expressing the blocks $\mathbf{D}_{11}$, $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ through the marginal rank covariance matrices $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$, we obtain one of the main results of this work.

Theorem 4.2. *The block matrices* $\mathbf{D}_{11}$*,* $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ *and* $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ *of the block-diagonal matrix* $\mathbf{D}^*$ *in* (10) *can be expressed in terms of the marginal rank covariance matrices* $\mathbf{M}_1$*,* $\mathbf{M}_2$ *and* $\mathbf{M}_3$ *as follows:*

$$\begin{aligned} \mathbf{D}_{11} &= \omega\frac{1}{\omega_1}\mathbf{M}_1 \sqrt[p_2-1]{\frac{|\mathbf{M}_2|}{\omega_2^{p_2}}} \sqrt[p_3-1]{\frac{|\mathbf{M}_3|}{\omega_3^{p_3}}}, \\ \mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12} &= \omega\frac{1}{\omega_2}\mathbf{M}_2 \sqrt[p_1-1]{\frac{|\mathbf{M}_1|}{\omega_1^{p_1}}} \sqrt[p_3-1]{\frac{|\mathbf{M}_3|}{\omega_3^{p_3}}}, \\ \mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13} &= \omega\frac{1}{\omega_3}\mathbf{M}_3 \sqrt[p_1-1]{\frac{|\mathbf{M}_1|}{\omega_1^{p_1}}} \sqrt[p_2-1]{\frac{|\mathbf{M}_2|}{\omega_2^{p_2}}}. \end{aligned} \tag{15}$$

Proof. From the equations in (14) it follows that

$$\begin{aligned} |\boldsymbol{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\boldsymbol{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\boldsymbol{\Psi}_{31}| &= \sqrt[p_1-1]{\frac{|\mathbf{M}_1|}{\omega_1^{p_1}}}, \\ |\boldsymbol{\Psi}_{22}| &= \sqrt[p_2-1]{\frac{|\mathbf{M}_2|}{\omega_2^{p_2}}}, \\ |\boldsymbol{\Psi}_{33}| &= \sqrt[p_3-1]{\frac{|\mathbf{M}_3|}{\omega_3^{p_3}}}, \end{aligned} \tag{16}$$

and that

$$\begin{aligned} (\boldsymbol{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\boldsymbol{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\boldsymbol{\Psi}_{31})^{-1} &= |\boldsymbol{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\boldsymbol{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\boldsymbol{\Psi}_{31}|^{-1} \\ &\quad \times \omega_1^{-1}\mathbf{M}_1, \\ \boldsymbol{\Psi}_{22}^{-1} &= \omega_2^{-1}|\boldsymbol{\Psi}_{22}|^{-1}\mathbf{M}_2, \\ \boldsymbol{\Psi}_{33}^{-1} &= \omega_3^{-1}|\boldsymbol{\Psi}_{33}|^{-1}\mathbf{M}_3. \end{aligned} \tag{17}$$

The equations in (15) now follow by substituting the expressions in (16) and (17) into (13). □

It is important to stress that Theorem 4.2 gives us a tool for presenting block-diagonal RCM's via respective marginal RCM's, which are of smaller size.

Before we go further with how to estimate all the blocks in the original rank covariance matrix $\mathbf{D}$, we describe briefly what we have done. We started with an elliptically distributed vector $\mathbf{y}$ with mean $\boldsymbol{\mu}$ and covariance matrix

$\mathbf{\Psi}$ partitioned as in (6). We assumed that $\mathbf{\Psi}_{23} = \mathbf{0}$, i.e. that the covariance matrix $\mathbf{\Psi}$ is partially known. In Theorem 3.1 we showed that this results in $\mathbf{D}_{23} = \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ for the corresponding RCM. The purpose was to use this additional information in the estimation of $\mathbf{D}$. We began with transforming our vector $\mathbf{y}$ so that the RCM of the transformed vector became block-diagonal, see Theorem 3.2. Next we decided to use an idea from the theory for normal distributions, namely to use the rank covariance matrices $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ of the marginal vectors $\mathbf{y}_1 + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{y}_2 + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{y}_3$, $\mathbf{y}_2$ and $\mathbf{y}_3$, respectively, for estimating the blocks $\mathbf{D}_{11}$, $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ in $\mathbf{D}^*$. In Theorem 4.2 we showed that these blocks can be expressed in terms of $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$.

To estimate $\mathbf{D}_{11}$, $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$, we need to estimate the marginal rank covariance matrices $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$. In (4) we defined the sample RCM. Esa Ollila (see Visuri *et al.* 2003) has written C-programs for calculating the rank vectors and the RCM, so estimating $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ is not a problem, when we have observation vectors from $\mathbf{y}_1 + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{y}_2 + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{y}_3$, $\mathbf{y}_2$ and $\mathbf{y}_3$. Since $\mathbf{M}_1$ is the RCM of $\mathbf{y}_1 + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{y}_2 + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{y}_3$, we need to estimate at first the matrices $\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$, to estimate $\mathbf{M}_1$. Since $\mathbf{D}^*$ is block-diagonal, the subranks of the components are uncorrelated. This means that we can obtain estimates of $\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ by solving a regression problem. Set $\mathbf{B}_2 = -\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{B}_3 = -\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$. Then $\mathbf{y}_{1i} - \mathbf{B}_2\mathbf{y}_{2i} - \mathbf{B}_3\mathbf{y}_{3i}$, $i = 1, \ldots, n$, are nothing else but error vectors. So we have to solve a multivariate regression problem to estimate $\mathbf{M}_1$. When we have estimates for $\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$, $\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$, $\mathbf{D}_{11}$, $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$, we can get the rest of the estimates.

The process of estimation of a structured RCM can be summarized in the following steps:

1) estimate $\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ in $\mathbf{y}_1 + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{y}_2 + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{y}_3$ by solving a multivariate regression problem;

2) then the marginal rank covariance matrices $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ can be estimated;

3) thereafter the estimates of $\mathbf{D}_{11}$, $\mathbf{D}_{22}^* = \mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33}^* = \mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ can be obtained using the formulas in (15);

4) the rest of the estimates can be obtained as follows:

$$\begin{aligned}
\widehat{\mathbf{D}}_{12} &= \widehat{\mathbf{D}}_{11}\widehat{\mathbf{D}_{11}^{-1}\mathbf{D}_{12}}, \\
\widehat{\mathbf{D}}_{13} &= \widehat{\mathbf{D}}_{11}\widehat{\mathbf{D}_{11}^{-1}\mathbf{D}_{13}}, \\
\widehat{\mathbf{D}}_{22} &= \widehat{\mathbf{D}_{22}^{*}} + \widehat{\mathbf{D}}_{21}\widehat{\mathbf{D}_{11}^{-1}\mathbf{D}_{12}}, \\
\widehat{\mathbf{D}}_{33} &= \widehat{\mathbf{D}_{33}^{*}} + \widehat{\mathbf{D}}_{31}\widehat{\mathbf{D}_{11}^{-1}\mathbf{D}_{13}}, \\
\widehat{\mathbf{D}}_{23} &= \widehat{\mathbf{D}}_{21}\widehat{\mathbf{D}_{11}^{-1}\mathbf{D}_{13}}.
\end{aligned}$$

Finally, we have estimated all the blocks in the original rank covariance matrix $\mathbf{D}$.

4.2 Estimating the constants ω_1, ω_2 and ω_3

We can see from formulas in (15) that $\mathbf{D}_{11}$, $\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$ have one common proportionality constant ω, so we can write

$$\begin{aligned}
\mathbf{D}_{11} &= \omega\,\mathbf{K}_1, \\
\mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12} &= \omega\,\mathbf{K}_2, \\
\mathbf{D}_{33} - \mathbf{D}_{31}\mathbf{D}_{11}^{-1}\mathbf{D}_{13} &= \omega\,\mathbf{K}_3,
\end{aligned}$$

where the matrices $\mathbf{K}_1$, $\mathbf{K}_2$ and $\mathbf{K}_3$ are functions of $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ and ω_1, ω_2, ω_3, see (15). We can express the blocks in the original RCM through ω, $\mathbf{K}_1$, $\mathbf{K}_2$, $\mathbf{K}_3$, $\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$:

$$\begin{aligned}
\mathbf{D}_{12} &= \omega\,\mathbf{K}_1\,\mathbf{D}_{11}^{-1}\mathbf{D}_{12}, \\
\mathbf{D}_{13} &= \omega\,\mathbf{K}_1\,\mathbf{D}_{11}^{-1}\mathbf{D}_{13}, \\
\mathbf{D}_{22} &= \omega\,[\mathbf{K}_2 + (\mathbf{D}_{11}^{-1}\mathbf{D}_{12})'\,\mathbf{K}_1'\,(\mathbf{D}_{11}^{-1}\mathbf{D}_{12})], \\
\mathbf{D}_{33} &= \omega\,[\mathbf{K}_3 + (\mathbf{D}_{11}^{-1}\mathbf{D}_{13})'\,\mathbf{K}_1'\,(\mathbf{D}_{11}^{-1}\mathbf{D}_{13})], \\
\mathbf{D}_{23} &= \omega\,(\mathbf{D}_{11}^{-1}\mathbf{D}_{12})'\,\mathbf{K}_1'\,(\mathbf{D}_{11}^{-1}\mathbf{D}_{13}).
\end{aligned} \tag{18}$$

It follows that in the applications that require the estimate of the rank covariance matrix up to a proportionality constant, it is enough to estimate $\mathbf{K}_1$, $\mathbf{K}_2$, $\mathbf{K}_3$, $\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ and $\mathbf{D}_{11}^{-1}\mathbf{D}_{13}$, and one does not need to be concerned about estimating the constant ω. However, since $\mathbf{K}_1$, $\mathbf{K}_2$ and $\mathbf{K}_3$ are functions of $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$ and ω_1, ω_2, ω_3, we still need to estimate the constants ω_1, ω_2 and ω_3, which connect the marginal rank covariance matrices $\mathbf{M}_1$, $\mathbf{M}_2$ and

$\mathbf{M}_3$ with the corresponding covariance matrices. Estimating ω_1, ω_2 and ω_3 is a difficult problem. One advantage that we have, compared to the situation of estimating ω, is that for the marginal covariance matrices we do not have any restrictions concerning the parameter space like we had for the covariance matrix $\mathbf{\Psi}$ of $\mathbf{y}$, i.e. the zero blocks. We can estimate ω_1, ω_2 and ω_3 using marginal sample vectors, which is also an advantage, because estimating the affine equivariant RCM is computer-intensive.

The constant ω in $\mathbf{D} = \omega\,|\mathbf{\Psi}|\mathbf{\Psi}^{-1}$ can be expressed through the radius r in the respective spherical distribution. If a random vector $\mathbf{x}$ follows a spherically symmetric distribution F_0, it can be represented as $\mathbf{x} = r\,\mathbf{u}$, where $r = \|\,\mathbf{x}\,\|$ and $\mathbf{u} = \mathbf{x}/\|\,\mathbf{x}\,\|$ and r and $\mathbf{u}$ are independent. The population rank function is then given by

$$\mathbf{R}_{F_0}(\mathbf{x}) = q_{F_0}(r)\,\mathbf{u} \tag{19}$$

for some function $q_{F_0}(r)$, see Möttönen *et al.* (1998) or Visuri *et al.* (2003), for example. The expressions for $q_{F_0}(r)$ for the multivariate normal and t-distribution can be found in Ollila *et al.* (2004).

Lemma 4.3. *Let* $\mathbf{x} : p\times 1$ *follow a spherical distribution* F_0*. Then the constant* ω *in* (7) *is given by*

$$\omega = \frac{\mathrm{E}\{q^2_{F_0}(r)\}}{(\mathrm{E}r^2)^{p-1}}p^{p-2}\,.$$

Proof. Since $\mathbf{x} = r\,\mathbf{u}$, $\mathrm{Cov}\,\mathbf{x} = (\mathrm{E}\,r^2/p)\,\mathbf{I}_p$. On one hand, from (7) it follows that

$$\mathbf{D}_{F_0} = \omega\frac{(Er^2)^{p-1}}{p^{p-1}}\,\mathbf{I}_p\,.$$

On the other hand, (19) gives that

$$\mathbf{D}_{F_0} = \frac{E\{q^2_{F_0}(r)\}}{p}\,\mathbf{I}_p\,,$$

and the result follows. □

We can see that ω is a function of the radius in the spherical distribution corresponding to our elliptical distribution. The constants ω_1, ω_2 and ω_3 are accordingly functions of the radii in the respective marginal spherical distributions.

One possible way to define the estimator of ω is to use relationship (7), which gives

$$\hat{\omega} = \frac{|\,\hat{\mathbf{D}}\,|^{1/p}}{|\,\hat{\mathbf{\Psi}}\,|^{(p-1)/p}}\,,$$

where $\mathbf{D}$ and $\mathbf{\Psi}$ are $p \times p$-matrices. Hence, the estimates of ω_1, ω_2 and ω_3 can be obtained via the marginal covariance matrices and the corresponding RCM's as follows:

$$\begin{aligned}
\hat{\omega}_1 &= \frac{|\, \hat{\mathbf{M}}_1 \,|^{1/p_1}}{|\, \hat{\mathbf{\Psi}}^*_{11} \,|^{(p_1-1)/p_1}}, \\
\hat{\omega}_2 &= \frac{|\, \hat{\mathbf{M}}_2 \,|^{1/p_2}}{|\, \hat{\mathbf{\Psi}}_{22} \,|^{(p_2-1)/p_2}}, \\
\hat{\omega}_3 &= \frac{|\, \hat{\mathbf{M}}_3 \,|^{1/p_3}}{|\, \hat{\mathbf{\Psi}}_{33} \,|^{(p_3-1)/p_3}},
\end{aligned}$$

where $\mathbf{\Psi}^* = \mathbf{\Psi}_{11} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Psi}_{21} + \mathbf{D}_{11}^{-1}\mathbf{D}_{13}\mathbf{\Psi}_{31}$, see (11).

4.3 Examples

In this section we present two examples of applications where it is enough to estimate the RCM corresponding to a regular covariance matrix up to a proportionality constant. When we estimate the rank covariance matrix corresponding to a partially known covariance matrix, we do not have to bother estimating the constant ω, see (18).

I. Principal component analysis

One application where the analysis can be based on the RCM instead of the covariance matrix, is principal component analysis. Suppose that we have a p-variate random vector $\mathbf{z}$ from a location-scale family. In location-scale families the population RCM is proportional to the inverse of the regular covariance matrix,

$$\mathbf{D} = \omega \, |\mathbf{\Psi}| \, \mathbf{\Psi}^{-1},$$

which can also be written as

$$\mathbf{D} = \omega \, |\mathbf{\Lambda}| \, \mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}',$$

where $\mathbf{\Lambda} : p \times p$ is the diagonal matrix of eigenvalues of $\mathbf{\Psi}$ and $\mathbf{P} : p \times p$ is the matrix of eigenvectors of $\mathbf{\Psi}$, see Theorem 2.6. The eigenvectors of the theoretical RCM equal the eigenvectors of the ordinary covariance matrix, and the eigenvalues of the RCM are proportional to the inverses of the eigenvalues of the covariance matrix. Suppose that we estimate the RCM up to the proportionality constant ω, i.e. if $\mathbf{D} = \omega\mathbf{K}$, we estimate $\mathbf{K}$. Let $\mathbf{K} = \mathbf{PLP}'$ be the orthogonal decomposition of $\mathbf{K}$, where $\mathbf{L} = \text{diag}(l_1, \dots, l_p)$ is the diagonal matrix of eigenvalues of $\mathbf{K}$. Assume that $l_1 \leq l_2 \leq \dots \leq l_p$. The eigenvalues of $\mathbf{D}$ are given by $\omega l_1 \leq \omega l_2 \leq \dots \leq \omega l_p$. The eigenvalues $\lambda_1, \dots, \lambda_p$ of the regular covariance matrix $\mathbf{\Psi}$ are proportional to the $(\omega l_1)^{-1}, \dots, (\omega l_p)^{-1}$, respectively, so we can write

$$\lambda_1 = \frac{c}{\omega l_1}, \; \dots \; \lambda_p = \frac{c}{\omega l_p},$$

for some constant c, $\lambda_1 \geq \dots \geq \lambda_p$. In principal component analysis one is interested in k first principal components with the largest variances, where k is determined so that these k components account for a certain percentage of the total variance. Often this percentage is chosen to be 85%. In this case k is determined from the inequality

$$\frac{\frac{c}{\omega l_1} + \dots + \frac{c}{\omega l_k}}{\frac{c}{\omega l_1} + \dots + \frac{c}{\omega l_p}} = \frac{1/l_1 + \dots + 1/l_k}{1/l_1 + \dots + 1/l_p} \geq 0.85.$$

We can see that it is enough to estimate $\mathbf{D} = \omega\mathbf{K}$ up to the proportionality constant. The k first principal components are given by $y_1 = \mathbf{p}_1'\mathbf{z}$, $y_2 = \mathbf{p}_2'\mathbf{z}, \dots$, $y_k = \mathbf{p}_k'\mathbf{z}$, where $\mathbf{p}_1, \dots, \mathbf{p}_k$ are the eigenvectors corresponding to $l_1, \dots, l_k$.

II. Canonical correlation analysis

Another application where the RCM can be used instead of the covariance matrix, is canonical correlation analysis, see also Visuri *et al.* (2003). Let $\mathbf{z}_1$ and $\mathbf{z}_2$ be vectors of dimensions p and q, respectively, $p \leq q$. Partition the covariance matrix $\mathbf{\Psi}$ of $\mathbf{z} = (\mathbf{z}_1', \mathbf{z}_2')'$ and the corresponding rank covariance matrix $\mathbf{D}$ as follows:

$$\mathbf{\Psi} = \begin{pmatrix} \mathbf{\Psi}_{11} & \mathbf{\Psi}_{12} \\ \mathbf{\Psi}_{21} & \mathbf{\Psi}_{22} \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{pmatrix},$$

where $\mathbf{\Psi}_{11}$ and $\mathbf{D}_{11}$ are $p \times p$-matrices. Let the inverses of $\mathbf{\Psi}$ and $\mathbf{D}$ be

$$\mathbf{\Psi}^{-1} = \begin{pmatrix} \mathbf{\Psi}^{11} & \mathbf{\Psi}^{12} \\ \mathbf{\Psi}^{21} & \mathbf{\Psi}^{22} \end{pmatrix}, \quad \mathbf{D}^{-1} = \begin{pmatrix} \mathbf{D}^{11} & \mathbf{D}^{12} \\ \mathbf{D}^{21} & \mathbf{D}^{22} \end{pmatrix},$$

respectively. It can be shown that

$$\mathbf{\Psi}_{11}^{-1}\mathbf{\Psi}_{12}\mathbf{\Psi}_{22}^{-1}\mathbf{\Psi}_{21} = [(\mathbf{\Psi}^{11})^{-1}\mathbf{\Psi}^{12}(\mathbf{\Psi}^{22})^{-1}\mathbf{\Psi}^{21}]' = [\mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{D}_{22}^{-1}\mathbf{D}_{21}]', \quad (20)$$

$$\mathbf{\Psi}_{22}^{-1}\mathbf{\Psi}_{21}\mathbf{\Psi}_{11}^{-1}\mathbf{\Psi}_{12} = [(\mathbf{\Psi}^{22})^{-1}\mathbf{\Psi}^{21}(\mathbf{\Psi}^{11})^{-1}\mathbf{\Psi}^{12}]' = [\mathbf{D}_{22}^{-1}\mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}]'.$$

Since $\mathbf{C}$ and $\mathbf{C}'$ have the same eigenvalues, it follows that $\mathbf{\Psi}_{11}^{-1}\mathbf{\Psi}_{12}\mathbf{\Psi}_{22}^{-1}\mathbf{\Psi}_{21}$ and $\mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{D}_{22}^{-1}\mathbf{D}_{21}$ have the same eigenvalues. We can conclude that the canonical correlations found using $\mathbf{\Psi}$ and $\mathbf{D}$ are exactly the same. From (20) it follows again that it is enough to estimate $\mathbf{D}$ up to a proportionality constant for calculating canonical correlations.

5 Summary

In this article we have studied the problem of estimating the affine equivariant rank covariance matrix based on the Oja median, when the corresponding covariance matrix is partially known. Specifically, we have assumed that the covariance matrix contains zero blocks, i.e. any two subvectors of the observation vector are uncorrelated. We have presented one way of estimating the rank covariance matrix, when the additional knowledge about the covariance matrix has been taken into account. An advantage of the presented estimators is, that the marginal rank covariance matrices can be used in the estimation process. When the size of the data vector is large, reducing the dimension is very useful, because calculating the rank covariance matrices is computer-intensive. One drawback of our estimators is that they include the constants ω, ω_1, ω_2 and ω_3, which characterize the relationship between the rank covariance matrices and the respective covariance matrices. The problem of estimating these constants needs to be studied further. The properties of our estimators and their numerical behavior also need to be further studied.

References

Barnett, V. (1976). The ordering of multivariate data. *Journal of the Royal Statistical Society, A*, **139**, 318–355.

Bondesson, L. (1982). Equivariant estimators. In *Encyclopedia of Statistical Sciences*, (Kotz, S. & Johnson, N.L., eds.) Vol.**2**, 538–540. Wiley, New York.

Fang, K.-T., Kotz, S. & Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.

Fang, K.-T. & Zhang, Y.-T. (1990). *Generalized Multivariate Analysis*. Science Press and Springer-Verlag.

Lehmann, E.L., Casella, G. (1998). *Theory of Point Estimation*. Springer, New York.

Möttönen, J., Hettmansperger, T.P., Oja, H. & Tienari, J. (1998). On the efficiency of affine invariant multivariate rank tests. *Journal of Multivariate Analysis*, **66**, 118–132.

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

Ollila, E., Croux, C. & Oja, H. (2004). Influence function and asymptotic efficiency of the affine equivariant rank covariance matrix. *Statistica Sinica*, **14**, 297–316.

Puri, M.L. & Sen, P.K. (1985). *Nonparametric Methods in General Linear Models*. Wiley, New York.

Visuri, S., Ollila, E., Koivunen, V., Möttönen, J. & Oja, H. (2003). Affine equivariant multivariate rank methods. *Journal of Statistical Planning and Inference*, **114**, 161–185.