Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

ON PROJECTION OF A POSITIVE DEFINITE MATRIX ON A CONE OF NONNEGATIVE DEFINITE TOEPLITZ MATRICES*

KATARZYNA FILIPIAK†, AUGUSTYN MARKIEWICZ‡, ADAM MIELDZIOC‡, AND ANETA SAWIKOWSKA‡§

Abstract. We consider approximation of a given positive definite matrix by nonnegative definite banded Toeplitz matrices. We show that the projection on linear space of Toeplitz matrices does not always preserve nonnegative definiteness. Therefore we characterize a convex cone of nonnegative definite banded Toeplitz matrices which depends on the matrix dimensions, and we show that the condition of positive definiteness given by Parter [*Numer. Math. 4*, 293–295, 1962] characterizes the asymptotic cone. In this paper we give methodology and numerical algorithm of the projection basing on the properties of a cone of nonnegative definite Toeplitz matrices.

This problem can be applied in statistics, for example in the estimation of unknown covariance structures under the multi-level multivariate models, where positive definiteness is required. We conduct simulation studies to compare statistical properties of the estimators obtained by projection on the cone with a given matrix dimension and on the asymptotic cone.

Key words. Banded Toeplitz matrix, Projection, Convex cone, Estimation, Covariance structure.

AMS subject classifications. 15B05, 65F35, 62H12.

1. Introduction.

The estimation of covariance matrix of an m-variate data that has some special structure has been considered in the literature. Some of the covariance structures that have received attention are linear, e.g. the covariance matrix belongs to the linear subspaces of symmetric matrices; cf. Anderson (1973). The standard estimation procedure starts with the estimation of unstructured covariance matrix with the use of sufficient statistics, say $\mathbf{S}$, and then it is approximated by particular structure. One of the approximation criteria is the Frobenius norm, which leads to the orthogonal projection of preliminary estimate of unstructured covariance matrix $\mathbf{S}$ onto the given linear subspace. This approach is used for example by Ohlson and von Rosen (2010) in the context of estimation or by Cui et al. [2] in the context of regularization. Nevertheless such projection does not need to preserve nonnegative definiteness that is required for covariance matrices (see Example 1.1). Therefore, it is necessary to project not onto the linear subspace but onto the convex cone of non-negative definite matrices from this subspace; cf. e.g. Ingram and Marsh [5]. It is worth noting that the definiteness is preserved for example when it does exist the orthogonal basis of the subspace consisting of non-negative definite matrices, such as uniform structure or orthogonal block structure; cf. e.g. Vanleeuwen et al. [8].

In this paper we focus on the subspace of banded Toeplitz matrices for which we show that the projection on them do not preserve definiteness (see Example 1.2).

Let us denote the space of symmetric matrices by $\mathcal{S}$, and let the $p+1$-dimensional linear subspace of $\mathcal{S}$

*Received by the editors on February 28, 2018. Accepted for publication on October 7, 2018. Handling Editor: Simo Puntanen.

†Insitute of Mathematics, Poznań University of Technology, Poznań, Poland (katarzyna.filipiak@put.poznan.pl).

‡Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland (amark@up.poznan.pl, adam.mieldzioc@mail.up.poznan.pl, aneta.sawikowska@mail.up.poznan.pl).

§Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

of $m \times m$ banded Toeplitz matrices be denoted by $\mathcal{T}_m(p)$, that is

$$\mathcal{T}_m(p) = \{ \begin{bmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_p & 0 & \dots & 0 \\ \alpha_1 & \alpha_0 & \alpha_1 & \dots & \alpha_p & \ddots & \vdots \\ \vdots & \alpha_1 & \alpha_0 & \alpha_1 & & \ddots & 0 \\ \alpha_p & & \ddots & \ddots & \ddots & & \alpha_p \\ 0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \alpha_1 \\ 0 & \dots & 0 & \alpha_p & \dots & \alpha_1 & \alpha_0 \end{bmatrix}_{mxm} : \alpha_0, \alpha_1, \dots, \alpha_p \in \mathbb{R} \}.$$

In other words every matrix $\mathbf{A}^t$ from $\mathcal{T}_m(p)$ can be expressed as a linear combination of identity matrix $\mathbf{I}$ and symmetric orthogonal matrices $\mathbf{T}_i$, $i = 1, \dots, p$, such that the ith superdiagonal and ith subdiagonal entries are equal to 1 and all other elements are equal to 0. Note that $\{\mathbf{I}, \mathbf{T}_1, \dots, \mathbf{T}_p\}$ is an orthogonal basis of the linear space $\mathcal{T}_m(p)$.

In this paper the Frobenius norm $||\mathbf{X}||_F^2 = \operatorname{tr}(\mathbf{X}^\top \mathbf{X})$ is considered as an approximation criterion. The projection $\mathbf{A}^t$ of $\mathbf{A}$ onto the linear subspace $\mathcal{T}_m(p)$, can be given as

$$\mathbf{A}^t = \alpha_0 \mathbf{I} + \sum_{i=1}^{p} \alpha_i \mathbf{T}_i$$

with $\alpha_0 = \operatorname{tr}\mathbf{A}/m$ and $\alpha_i = \operatorname{tr}(\mathbf{A}\mathbf{T}_i)/(2(m-i))$, $i = 1, \dots, p$; cf. e.g. Cui et al. [2]. It is worth to note that every α_i is obtained as the average of ith superdiagonal and subdiagonal.

The following example shows that even if $\mathbf{A}$ is positive definite, the projection $\mathbf{A}^t$ does not need to be positive definite.

Example 1.1. *Let $m = 4$ and $\mathbf{A} = \mathbf{I} + 0.9\mathbf{T}_1 + 0.9\mathbf{T}_2 + 0.9\mathbf{T}_3$. The projection of $\mathbf{A}$ onto $\mathcal{T}_4(1)$ gives $\mathbf{A}^t = \mathbf{I} + 0.9\mathbf{T}_1$. It is easy to calculate that $\mathbf{A}$ is positive definite (with the eigenvalues belonging to the set $\{3.7, 0.1\}$), whilst the projection $\mathbf{A}^t$ is not definite (with the eigenvalues $\{2.46, 1.56, 0.44, -0.46\}$).*

The aim of this paper is to find the best approximation of $\mathbf{A}$ over the set of nonnegative definite banded Toeplitz matrices, say $\mathcal{T}_m^{\geq}(p)$, that is

$$\min_{\mathbf{A}^t \in \mathcal{T}_m^{\geq}(p)} ||\mathbf{A} - \mathbf{A}^t||_F = ||\mathbf{A} - \mathbf{A}^{\geq}||_F. \tag{1.1}$$

In other words, in the case when $\mathbf{A}^t$ is not nonnegative definite, it is necessary to project not onto the linear subspace $\mathcal{T}_m(p)$ but onto the convex cone of non-negative definite matrices from this subspace, $\mathcal{T}_m^{\geq}(p)$. Thus, to solve (1.1) first we project $\mathbf{A}$ onto $\mathcal{T}_m(p)$, then we check if $\mathbf{A}^t$ is nonnegative definite and finally, if the definiteness is not preserved, we project $\mathbf{A}^t$ onto the convex cone $\mathcal{T}_m^{\geq}(p)$.

Observe that every matrix from the convex cone $\mathcal{T}_m^{\geq}(1)$ must satisfy the condition

$$-\frac{1}{2\cos(\pi/(m+1))} \leq \frac{\alpha_1}{\alpha_0} \leq \frac{1}{2\cos(\pi/(m+1))}; \tag{1.2}$$

cf. [4, Sec. 28.5, p. 522]. Nevertheless, for $p > 1$ analytic characterization of a cone is much more complicated as a general characterization of nonnegative definiteness Toeplitz matrix expressed in terms of nonnegativity

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

its m leading minors. Parter [7, Remark II] proposed the following characterization of positive definiteness of banded Toeplitz matrix for arbitrary m: a given banded Toeplitz matrix is positive definite if and only if

$$f(t) = \alpha_0 + 2\sum_{k=1}^{p} \alpha_k \cos(kt) \geq 0 \tag{1.3}$$

for every $t \in \mathbb{R}$. However, as follows from the proof of Remark II of Parter [7], the above condition is satisfied if $m \to \infty$. The following example shows that Parter's condition is not true in general.

Example 1.2. *Let* $\mathbf{A}^t = 10\mathbf{I} + \mathbf{T}_1 - 5\mathbf{T}_2 \in \mathcal{T}_5(2)$. *Since* $\mathbf{A}^t$ *is positive definite, the function* $f(t) = 10 + 2\cos(t) - 10\cos(2t)$ *must be nonnegative for every* $t \in \mathbb{R}$. *Thus, let assume* $t = \pi$. *Then* $f(\pi) = -2$. *Contradiction.*

Observe however, that condition (1.3) is still wrongly used as characterization of positive definiteness of Toeplitz banded matrix for a finite m; cf. e.g. Lin et al. [6, p. 133]. From the other hand, from sufficiency in Parter's characterization it follows that the asymptotic cone given by Parter [7] is included in the convex cone $\mathcal{T}_m^{\geq}(p)$.

Since the general characterization of convex cone $\mathcal{T}_m^{\geq}(p)$ for $p > 1$ is not available, we determine it numerically. The algorithm for finding the solution of (1.1) is given in Section 2.2.

In statistical applications it is allowed to consider singular $\mathbf{A}$ (in so-called high-dimensional problems). Observe however, that the best approximation of $\mathbf{A}$ does not need to be singular. In such a case matrix $\mathbf{A}^t$ should belong to the cone interior. This problem is considered in Section 3.

2. Projection onto $\mathcal{T}_m^{\geq}(p)$. In this section we will use the vector norm of $\mathbf{x} = (x_0, x_1, \ldots, x_p)^\top \in \mathbb{R}^{p+1}$ defined for a given m as

$$||\mathbf{x}||^2 = mx_0^2 + 2(m-1)x_1^2 + 2(m-2)x_2^2 + \cdots + 2(m-p)x_p^2. \tag{2.4}$$

The above norm represents the Frobenius norm of Toeplitz matrix, with $(\mathbf{x}, \mathbf{0}_{m-p-1}^T)^T$ being its first column.

Observe that if $p = 1$ condition (1.2) for nonnegative definiteness of $\mathbf{A}^t$ is imposed on the components α_0 and α_1. Thus, instead of $\mathcal{T}_m(1)$ we can consider the space $\mathbb{R}^2$ of the components of $\mathbf{A}^t$ with the norm defined in (2.4). To characterize a cone $\mathcal{T}_m^{\geq}(1)$, we use a 2−dimensional cone $\mathcal{C}_m(1) \subset \mathbb{R}^2$, which is a space of components of $\mathbf{A}^t$ satisfying (1.2). Observe, that since (1.2) depends on the order m of a matrix, the cone $\mathcal{C}_{m_2}(1) \subset \mathcal{C}_{m_1}(1)$ for every $m_1 < m_2$. It can be also noted that the increase of m results in the increase of the number of conditions for positive definiteness of a given matrix expressed for example by principal minors.

Similarly, in Section 2.2 we give the characterization of a cone $\mathcal{T}_m^{\geq}(p)$ by imposing conditions on components of a matrix from $(p+1)$−dimensional space $\mathcal{T}_m(p)$. We denote by $\mathcal{C}_m(p) \subset \mathbb{R}^{p+1}$ a cone of components of nonnegative definite Toeplitz matrices with $p > 1$. Since the condition for matrix definiteness depends on the order m of this matrix, it is obvious that $\mathcal{C}_{m_2}(p) \subset \mathcal{C}_{m_1}(p)$ for every $m_1 < m_2$. Additionally, if $m \to \infty$, the cone $\mathcal{C}_m(p)$ (denoted by $\mathcal{C}_\infty(p)$) can be described by (1.3); cf. Parter [7].

All the figures presented in the following sections are prepared with the use of *Mathematica*.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

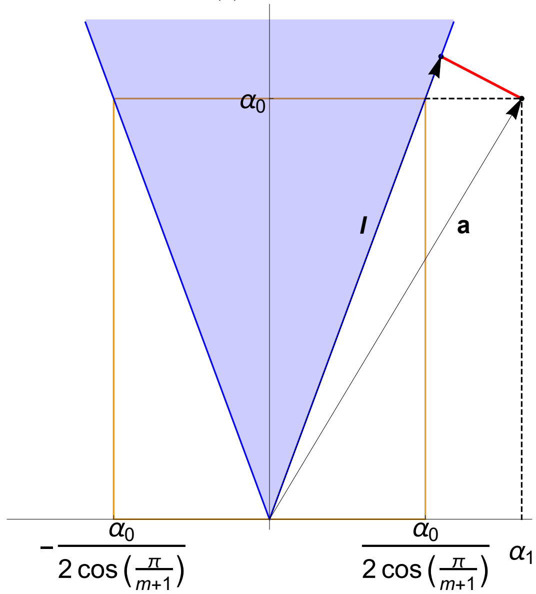
2.1. Projection onto $\mathcal{T}_m^{\geq}(1)$. In this section we determine the solution of (1.1) for tridiagonal matrices.

Algorithm 1 – outline:

1. Project $\mathbf{A}$ onto the space generated by $\mathbf{I}$ and $\mathbf{T}_1$ and get $\mathbf{A}^t = \alpha_0\mathbf{I} + \alpha_1\mathbf{T}_1$. If $\mathbf{A}^t \in \mathcal{T}_m^{\geq}(1)$, then $\mathbf{A}^t = \mathbf{A}^{\geq}$.
2. If $\mathbf{A}^t \notin \mathcal{T}_m^{\geq}(1)$, then project $\mathbf{A}^t$ onto the cone $\mathcal{T}_m^{\geq}(1)$ and get $\mathbf{A}^{\geq}$ in the following way:
   2.1. define vector $\mathbf{a} = (\alpha_0, \alpha_1)^\top \in \mathbb{R}^2$;
   2.2. project $\mathbf{a}$ onto the cone $\mathcal{C}_m(1)$ with respect to the vector norm defined in (2.4) and get the vector $\boldsymbol{\ell} = (\nu_0, \nu_1)^\top$;
   2.3. the solution is $\mathbf{A}^{\geq} = \nu_0\mathbf{I} + \nu_1\mathbf{T}_1$.

The above algorithm is illustrated on Figure 1. For better illustration (conformed with the cones $\mathcal{C}_4(2)$ on Figure 2) the vertical axis represents the values of α_0 components and the horizontal one the values of α_1. It is worth noting that the projection of vector $\mathbf{a}$ onto the cone $\mathcal{C}_m(1)$ does not need to be orthogonal, as it is the projection with respect to the norm defined in (2.4).

FIGURE 1. *The cone $\mathcal{C}_m(1)$ and illustration of the algorithm*



EXAMPLE 2.1. *Let $\mathbf{A}$ be the same as in Example 1.1. Since $\mathbf{A}^t \in \mathcal{T}_4(1)$ is not nonnegative definite , we have to project $\mathbf{A}^t$ onto $\mathcal{T}_4^{\geq}(1)$. We obtain $\mathbf{A}^{\geq} = 1.166\mathbf{I} + 0.721\mathbf{T}_1$ with the eigenvalues $\{2.33, 1.61, 0.721, 0\}$ and $||\mathbf{A} - \mathbf{A}^{\geq}||_F = 2.272283097$.*

2.2. Projection onto $\mathcal{T}_m^{\geq}(p)$, $p > 1$. The aim of this section is to present general algorithm for projection of a given nonnegative definite matrix onto the convex cone $\mathcal{T}_m^{\geq}(p)$.

One of the properties of $\mathcal{C}_m(p)$ is its symmetry with respect to the space generated by $\mathbf{e}_{2k-1}$, $k = 1, \ldots, \lfloor \frac{p+2}{2} \rfloor$, where $\mathbf{e}_i$ is the ith column of $\mathbf{I}_{p+1}$. It follows from the property of Toeplitz matrices given by

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

Abdikalykov et al. [1]. He showed that for every $\mathbf{A}^t = \alpha_0\mathbf{I} + \sum_{i=1}^{p} \alpha_i\mathbf{T}_i$

$$\mathbf{F}\mathbf{A}^t\mathbf{F} = \alpha_0\mathbf{I} + \sum_{i=1}^{p}(-1)^p\alpha_i\mathbf{T}_i, \qquad \mathbf{F} = \mathrm{diag}(1, -1, 1, -1, ..., \pm 1),$$

which implies that $\det(\mathbf{F}\mathbf{A}^t\mathbf{F}) = \det(\mathbf{A}^t)$ and thus, if $\mathbf{A}^t$ is nonnegative definite (nnd), then also $\mathbf{F}\mathbf{A}^t\mathbf{F}$ is nnd.

In the case of $\mathcal{C}_m(p)$ the symmetry means that if $\mathbf{x} \in \mathcal{C}_m(p)$ than also $\mathbf{R}\mathbf{x} \in \mathcal{C}_m(p)$, where $\mathbf{R}$ is the reflector that maps $\mathbf{x}$ to its reflection about the hyperplane generated by $\mathbf{e}_{2k-1}$, $k = 1, \ldots, \lfloor \frac{p+2}{2} \rfloor$, i.e., $\mathbf{R} = \mathrm{diag}(1, -1, 1, -1, ..., \pm 1) \in \mathbb{R}^{p+1}$. Symmetry for $p = 1, 2$ can be seen on Figures 1 and 2. For $p > 1$, the property of symmetry reduces the number of possible directions of decreasing norm in the Algorithm 2, point 2.6.

We now illustrate the problem geometrically for $\mathcal{C}_m(2) \subset \mathbb{R}^3$. Figure 2 represent the cone for $m = 4$ and asymptotic cone described by (1.3). Note, that the asymptotic cone for $p = 2$ is the sum of a polyhedral cone generated by points $(1, \pm 2/3, 1/6)$ and $(1, 0, -0.5)$, and an oblique elliptical cone $\frac{\alpha_1^2}{0.5} + \frac{(\alpha_2 - 0.25\alpha_0)^2}{0.0625} = \alpha_0^2$, $\alpha_0 > 0$.

FIGURE 2. *The cone $\mathcal{C}_4(2)$ and the asymptotic cone $\mathcal{C}_\infty(2)$ and both cones together*

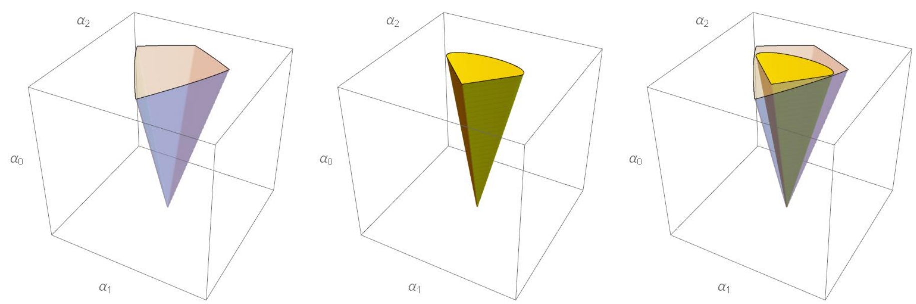


Figure 3 represents the intersections of $\mathcal{C}_m(2)$ and the plane $\alpha_0 = 1$ for $m \geq 4$ with respective intersection of asymptotic cone inside.

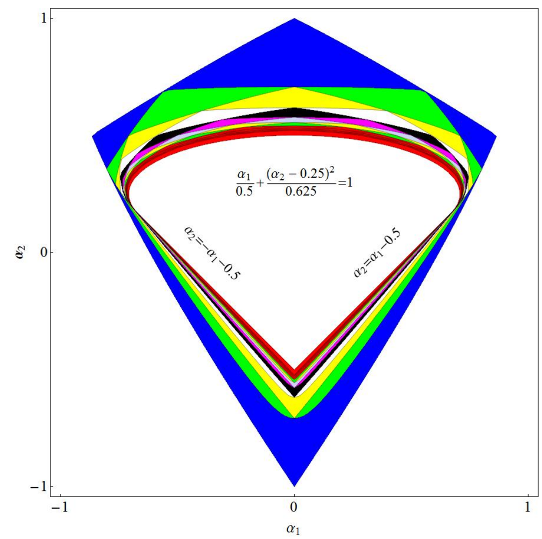
In general, the best approximation $\mathbf{A}^{\geq}$ of $\mathbf{A}$ over the cone $\mathcal{T}_m^{\geq}(p)$ can be found using the following algorithm.

Algorithm 2 – outline:

1. Project $\mathbf{A}$ onto the space $\mathcal{T}_m(p)$ and get $\mathbf{A}^t = \alpha_0\mathbf{I} + \alpha_1\mathbf{T}_1 + \alpha_2\mathbf{T}_2 + \cdots + \alpha_p\mathbf{T}_p$. If $\mathbf{A}^t \in \mathcal{T}_m^{\geq}(p)$ then $\mathbf{A}^t = \mathbf{A}^{\geq}$.
2. If $\mathbf{A}^t \notin \mathcal{T}_m^{\geq}(p)$ then:
   2.1. define vector $\mathbf{a}_0 = (\alpha_0, \alpha_1, \ldots, \alpha_p)^\top \in \mathbb{R}^{p+1}$;
   2.2. take $\mathbf{b}_0 = (\alpha_0, 0, \ldots, 0)^\top \in \mathcal{C}_m(p)$.
   2.3. determine vector $\mathbf{c}_0$ which is a convex combination of $\mathbf{a}_0$ and $\mathbf{b}_0$, such that $\mathbf{c}_0$ lies on the edge of $\mathcal{C}_m(p)$; that is $\mathbf{c}_0 = (\alpha_0, c_1, c_2, \ldots, c_p)^\top$;
   2.4. determine an extreme ray ℓ_0 defined by vector $\mathbf{c}_0$; it means that every vector $\boldsymbol{\ell}$ that belongs to ℓ_0 can be expressed as $(\alpha_0 t, c_1 t, c_2 t, \ldots, c_p t)^\top$ with $t \in \mathbb{R}$;

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

FIGURE 3. *The intersections of $\mathcal{C}_m(2)$, $m \geq 4$, and $\mathcal{C}_\infty(2)$ with the plane $\alpha_0 = 1$ (blue color for $m = 4$, green for $m = 5$, yellow for $m = 6$, etc.).*



2.5. project vector $\mathbf{a}_0$ onto ℓ_0 with respect to the norm defined in (2.4), that is find the minimum of $||\boldsymbol{\ell} - \mathbf{a}_0||$ with respect to parameter t and get $\boldsymbol{\ell}_0 = (\nu_0^0, \nu_1^0, \ldots, \nu_p^0)^\top$; calculate $||\boldsymbol{\ell}_0 - \mathbf{a}_0||$;
2.6. verify, if the norm $||\boldsymbol{\ell}_0 - \mathbf{a}_0||$ decreases in the neighborhood of the point $(\nu_0^0, \nu_1^0, \ldots, \nu_p^0)$, considering $\lfloor \frac{p}{2} \rfloor$ possible directions; thus, for every combination of $\kappa_j \in \{-1, 0, 1\}$, $j \in Z = \{3, 5, \ldots, 2\lfloor \frac{p}{2} \rfloor + 1\}$ construct vectors $\mathbf{b}_j^{(\kappa_j)} = \mathbf{b}_0 + \varepsilon \sum_{j \in Z} \kappa_j \mathbf{e}_j$, where $\varepsilon > 0$; determine vectors $\mathbf{c}_j^{(\kappa_j)}$ which is a convex combination of $\mathbf{a}_0$ and $\mathbf{b}_j^{(\kappa_j)}$ lying on the edge of $\mathcal{C}_m(p)$, determine extreme rays $\ell_j^{(\kappa_j)}$ defined by $\mathbf{c}_j^{(\kappa_j)}$, project $\mathbf{a}_0$ onto $\ell_j^{(\kappa_j)}$ and get $\boldsymbol{\ell}_j^{(\kappa_j)}$, and calculate $||\boldsymbol{\ell}_j^{(\kappa_j)} - \mathbf{a}_0||$; among the vectors $\mathbf{b}_j^{(\kappa_j)}$ indicate a vector $\mathbf{b}_j^{(\kappa_j)}$ such that $||\boldsymbol{\ell}_j^{(\kappa_j)} - \mathbf{a}_0||$ is minimal and denote this vector by $\mathbf{b}_1$ with respective $\boldsymbol{\ell}_1$;
2.7. repeat i times point 2.6 with $\mathbf{b}_j^{(\kappa_j)} = \mathbf{b}_i + \varepsilon \sum_{j \in Z} \kappa_j \mathbf{e}_j$, getting the vectors $\mathbf{b}_{i+1}$ and $\boldsymbol{\ell}_{i+1} = (\nu_0^{i+1}, \nu_1^{i+1}, \ldots, \nu_p^{i+1})^\top$ until $||\boldsymbol{\ell}_{i+1} - \mathbf{a}_0|| > ||\boldsymbol{\ell}_i - \mathbf{a}_0||$;
2.8. the solution is $\mathbf{A}^{\geq} = \nu_0^i \mathbf{I} + \nu_1^i \mathbf{T}_1 + \cdots + \nu_p^i \mathbf{T}_p$.

It is worth noting that ε in point 2.6 of the above algorithm can be understood as the precision of the numerical solution and thus its value should be rather small.

EXAMPLE 2.2. *Let $\mathbf{A}$ be as in Example 1.1. Projecting $\mathbf{A}$ onto $\mathcal{T}_4(2)$ we obtain $\mathbf{A}^t = \mathbf{I} + 0.9\mathbf{T}_1 + 0.9\mathbf{T}_2$ with the eigenvalues $\{3.305, 1.000, 0.100, -0.405\}$. Projecting $\mathbf{A}^t$ onto $\mathcal{T}_4^{\geq}(2)$ we get $\mathbf{A}^{\geq} = 1.187\mathbf{I} + 0.826\mathbf{T}_1 + 0.72\mathbf{T}_2$ with the eigenvalues $\{3.20, 1.20, 0.347, 0\}$ and $||\mathbf{A} - \mathbf{A}^{\geq}||_F = 1.38659684$.*

Observe, that to project $\mathbf{A}^t$ onto the asymptotic cone, denoted by $\mathcal{T}_\infty(p)$, it is enough to replace in Algorithm 2 vectors $\mathbf{b}_0, \mathbf{c}_0 \in \mathcal{C}_m(p)$ by respective vectors belonging to $\mathcal{C}_\infty(p)$.

Next example shows that for relatively small m, the matrix $\mathbf{A}^{\geq}$ which follows from the projection of $\mathbf{A}^t$ onto the cone $\mathcal{T}_m^{\geq}(2)$ and the matrix $\mathbf{A}^{\infty}$, which follows from the projection of $\mathbf{A}^t$ onto the asymptotic cone $\mathcal{T}_\infty(2)$, are relatively close to each other. It means that the Frobenius norms $||\mathbf{A} - \mathbf{A}^{\geq}|_F$ and $||\mathbf{A} - \mathbf{A}^{\infty}||_F$ normalized by $||\mathbf{A}||_F$, do not differ much.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

Example 2.3. *Let $m = 20$ and $\mathbf{A} = \mathbf{I} + 0.9\sum_{i=1}^{19}\mathbf{T}_i$ with the eigenvalues belonging to the set $\{3.7, 0.1\}$. Projecting $\mathbf{A}$ onto $\mathcal{T}_{20}(2)$ we get*

$$\mathbf{A}^t = \mathbf{I} + 0.9\mathbf{T}_1 + 0.9\mathbf{T}_2$$

which is not definite (one of the eigenvalues is equal to -1.02183856944 whilst the remaining are positive). Projecting $\mathbf{A}^t$ onto $\mathcal{T}_{20}^{\geq}(2)$ we get $\mathbf{A}^{\geq} = 1.382\mathbf{I} + 0.763\mathbf{T}_1 + 0.587\mathbf{T}_2$. It can be checked that all of the eigenvalues of $\mathbf{A}^{\geq}$ are nonnegative and $||\mathbf{A} - \mathbf{A}^{\geq}||_F/||\mathbf{A}||_F = 0.882$.
Projecting now $\mathbf{A}^t$ onto $\mathcal{T}_{\infty}^{\geq}(2)$ we get $\mathbf{A}^{\infty} = 1.406\mathbf{I} + 0.765\mathbf{T}_1 + 0.577\mathbf{T}_2$. It can be checked that all of the eigenvalues of $\mathbf{A}^{\geq}$ are positive and $||\mathbf{A} - \mathbf{A}^{\geq}||_F/||\mathbf{A}||_F = 0.883$.

Observe, that $\mathbf{A}^{\geq}$ obtained as the result of the algorithm should be singular or ill-conditioned, as it lies close to the border of the cone. Since very often, especially in statistics, more interesting is to get positive definite banded Toeplitz matrix, the problem of avoiding singularity or ill-conditioning of the projection can be solved by the projection onto the asymptotic cone $\mathcal{T}_{\infty}(p)$.

The problem described in the paper raised from the statistical problem of estimation of structured covariance matrix. Observe however, that the matrix obtained by the projection onto the space of Toeplitz matrices, does not need to be even definite (see e.g. Example 1). Therefore, Algorithms 1 and 2 can be used for arbitrary matrices, not necessarily positive definite.

3. Simulations. One of the possible applications of the proposed method is the estimation of the covariance matrix structured as a banded Toeplitz matrix. In the literature usually the maximum likelihood estimators are studied in the context of estimation of structured covariance matrices. However, for a Toeplitz structure the maximum likelihood estimation is challenging and no general closed-form solution is known; cf. e.g. Dembo et al. [3]. Therefore in this section we compare statistical properties of the projections onto $\mathcal{C}_m(p)$ and $\mathcal{C}_{\infty}(p)$ using simulations. All numerical calculations are performed in R, using our own script.

Let the matrix of observations $\mathbf{X}$ is matrix-variate normally distributed, i.e., $\mathbf{X} \sim N_{n,m}(\boldsymbol{\mu}^{\top} \otimes \mathbf{1}_n, \boldsymbol{\Sigma}, \mathbf{I}_n)$ with unknown $\boldsymbol{\Sigma}$ structured as banded Toeplitz matrix. For simulations we assume (n, m) as $(10, 5)$, $(10, 10)$ or $(20, 10)$, $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{I}_m + 0.5\mathbf{T}_1 + 0.4\mathbf{T}_2$ or $\boldsymbol{\Sigma} = \mathbf{I}_m + 0.25\mathbf{T}_1 - 0.25\mathbf{T}_2$. For every pair (n, m) we simulate the matrix of observations and determine the sample covariance matrix $\mathbf{S} = \frac{1}{n-1}\mathbf{X}\mathbf{Q}_{\mathbf{1}_n}\mathbf{X}^{\top}$, where $\mathbf{Q}_{\mathbf{1}_n} = \mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n^{\top}$, 1000 times. In Tables 1 and 2 we present the average, the standard deviation (s.d.) and the quadratic risk of proposed estimators obtained as projections onto $\mathcal{C}_m(2)$ (matrix $\mathbf{S}^{\geq} = \nu_0\mathbf{I}_m + \nu_1\mathbf{T}_1 + \nu_2\mathbf{T}_2$) and onto $\mathcal{C}_{\infty}(2)$ (matrix $\mathbf{S}^{\infty} = \theta_0\mathbf{I}_m + \theta_1\mathbf{T}_1 + \theta_2\mathbf{T}_2$). The risk R of the estimators $\mathbf{S}^{\geq}$ and $\mathbf{S}^{\infty}$, which can be expressed as the sum of the risks of respective components of $\mathbf{S}^{\geq}$ and $\mathbf{S}^{\infty}$, as well as the averaged normalized distance d between the estimators and the sample covariance matrix $\mathbf{S}$, that is $||\mathbf{S} - \mathbf{S}^{\geq}||_F/||\boldsymbol{\Sigma}||_F$ and $||\mathbf{S} - \mathbf{S}^{\infty}||_F/||\boldsymbol{\Sigma}||_F$, respectively, are also given. We shall mention that in the above simulations about 70% of projections $\mathbf{S}^t$ belong to the asymptotic cone $\mathcal{T}_m^{\geq}(2)$ and thus are positive definite.

Note, that for $n = m = 10$, the sample covariance matrix $\mathbf{S}$ is singular. The following simulations show that Algorithm 2 can be also used for high-dimensional case. Observe however, that the resulting matrix, $\mathbf{S}^{\geq}$, can be still singular or ill-conditioned, and to avoid this problem projection of $\mathbf{S}^t$ onto the asymptotic cone can be suggested.

From algebraic point of view, if we are looking for the estimator of banded Toeplitz matrix, we should project $\mathbf{S}$ onto the cone $\mathcal{T}_m^{\geq}(p)$. As expected, the distance d for $\mathbf{S}^{\geq}$ is always smaller than the distance d for $\mathbf{S}^{\infty}$. Moreover, the components of $\mathbf{S}^{\geq}$ are usually less biased than the components of $\mathbf{S}^{\infty}$, whilst the variability are comparable. Observe however, that the risk R of the estimators $\mathbf{S}^{\geq}$ and $\mathbf{S}^{\infty}$ are close to each

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

Table 1

Mean, standard deviation and risk of the components of $\mathbf{S}^{\geq}$ *and* $\mathbf{S}^{\infty}$ *for 1000 simulation trials, as well as the risk of* $\mathbf{S}^{\geq}$ *and* $\mathbf{S}^{\infty}$ *and the normalized distance* d *for* $\mathbf{\Sigma} = \mathbf{I}_m + 0.5\mathbf{T}_1 + 0.4\mathbf{T}_2$

n	m		ν_0	ν_1	ν_2	θ_0	θ_1	θ_2
10	5	mean	1.0112	0.4946	0.3948	1.0518	0.4836	0.3604
		s.d.	0.2776	0.2534	0.2346	0.3176	0.2420	0.1903
		risk	0.0772	0.0643	0.0550	0.1035	0.0589	0.0378
		R		0.1965			0.2002	
		d		0.5164			0.5245	
10	10	mean	1.0149	0.4941	0.3937	1.0299	0.4902	0.3814
		s.d.	0.2060	0.1749	0.1398	0.2183	0.1719	0.1262
		risk	0.0426	0.0306	0.0196	0.0486	0.0297	0.0163
		R		0.0928			0.0946	
		d		0.8025			0.8039	
20	10	mean	1.0116	0.4955	0.3910	1.0265	0.4917	0.3790
		s.d.	0.2110	0.1753	0.1396	0.2230	0.1721	0.1268
		risk	0.0447	0.0307	0.0196	0.0504	0.0297	0.0165
		R		0.0928			0.0966	
		d		0.8071			0.8085	

Table 2

Mean, standard deviation and risk of the components of $\mathbf{S}^{\geq}$ *and* $\mathbf{S}^{\infty}$ *for 1000 simulation trials, as well as the risk of* $\mathbf{S}^{\geq}$ *and* $\mathbf{S}^{\infty}$ *and the normalized distance* d *for* $\mathbf{\Sigma} = \mathbf{I}_m + 0.25\mathbf{T}_1 - 0.25\mathbf{T}_2$

n	m		ν_0	ν_1	ν_2	θ_0	θ_1	θ_2
10	5	mean	0.9897	0.2415	-0.2481	1.0132	0.2197	-0.2243
		s.d.	0.2243	0.1450	0.1989	0.2393	0.1401	0.1759
		risk	0.0504	0.0211	0.0396	0.0574	0.0206	0.0316
		R		0.1111			0.1096	
		d		0.6449			0.6493	
10	10	mean	1.0055	0.2463	-0.2416	1.0144	0.2378	-0.2330
		s.d.	0.1693	0.0999	0.1250	0.1742	0.0996	0.1181
		risk	0.0287	0.0100	0.0157	0.0306	0.0101	0.0142
		R		0.0544			0.0549	
		d		0.9662			0.9672	
20	10	mean	1.0099	0.2441	-0.2436	1.0189	0.2354	-0.2348
		s.d.	0.1696	0.0997	0.1252	0.1740	0.0990	0.1187
		risk	0.0289	0.0100	0.0157	0.0306	0.0100	0.0143
		R		0.0546			0.0549	
		d		0.9670			0.9679	

other in general, and in one case the risk of $\mathbf{S}^{\infty}$ is even smaller than the risk of $\mathbf{S}^{\geq}$. Moreover, since $\mathbf{S}^{\geq}$ belongs to the edge of $\mathcal{C}_m(p)$, it is always nonnegative definite, whilst $\mathbf{S}^{\infty}$ is always positive definite. Finally, another advantage of $\mathbf{S}^{\infty}$ is that the procedure of its determining is much simpler than for $\mathbf{S}^{\geq}$. For example for $\mathcal{C}_{\infty}(2)$ it is enough to check three conditions for definiteness, whilst for $\mathcal{C}_m(2)$ the number of conditions

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 33, pp. 74-82, Month 2018.

to check depends on the number of steps in the algorithm. For $p > 2$ the algorithm for determining $\mathbf{S}^{\geq}$ is much more time consuming than determination of $\mathbf{S}^{\infty}$.

Summing up, in the problem of estimation of structured covariance matrix, the projection of $\mathbf{S}$ onto the asymptotic cone $\mathcal{T}_{\infty}^{\geq}(p)$ can be recommended instead of the projection onto the cone $\mathcal{T}_{m}^{\geq}(p)$. Thus, the approach presented by Lin et al. [6, p. 133] is statistically permissible.

4. Acknowledgments. This research is partially supported by Statutory Activities No. 04/43/DSPB/0101 (Katarzyna Filipiak).

REFERENCES

[1] A.K. Abdikalykov, V.N. Chugunov and Kh.D. Ikramov. Unitary congruence automorphisms of the space of Toeplitz matrices. *Linear and Multilinear Algebra*, 63:1195–1203, 2015.

[2] X. Cui, C. Li, J. Zhao, L. Zeng, D. Zhang, and J. Pan. Covariance structure regularization via Frobenius norm discrepancy. *Linear Algebra Appl.*, 510:124–145, 2016.

[3] A. Dembo, C.L. Mallows, and L.A. Shepp. Embedding nonnegative definite Toeplitz Matrices in nonnegative definite circulant matrices, with application to covariance estimation. *IEEE Trans. on Information Theory*, 35:1206–1212, 1989.

[4] N.J. Higham. *Accuracy and Stability of Numerical Algorithms* (2nd ed). SIAM, Philadelphia, USA, 2002.

[5] J.M. Ingram and M.M. Marsh. Projection onto convex cones in Hilbert space. *J. Approx. Theory*, 64:343–350, 1991.

[6] L. Lin, N. Higham and J. Pan. Covariance Structure Regularization via Entropy Loss Function. *Computational Statistics and Data Analysis*, 72:315-327, 2014.

[7] S.V. Parter. An observation on the numerical solution of difference equations and a theorem of Szegö. *Numer. Math.*, 4:293–295, 1962.

[8] Vanleeuwen, D.M., D.S. Birkes, J.F. Seely, 1999. Balance and orthogonality in designs for mixed classification models. *Ann. Statist.*, 27:1927–1947.