ELSEVIER Linear Algebra and its Applications 335 (2001) 167–181

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Obtuse cones and Gram matrices with non-negative inverse

Andrzej Cegielski

Institute of Mathematics, Technical University, ul. Podgórna 50, 65-246 Zielona Góra, Poland

Received 24 November 1999; accepted 30 January 2001

Submitted by R.A. Brualdi

Abstract

We study some properties of Gram matrices with non-negative inverse which lead to the constructions of obtuse cones. These constructions can find applications in optimization problems, e.g., in methods of convex feasibility problems or of convex minimization problems. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Gram matrix; Positive definite matrix with non-negative elements; Obtuse cone

1. Introduction

In optimization methods we often have to solve the following problem:

Given a system of linear inequalities $A^{\mathrm{T}}x \leqslant b$, where A is a matrix of type $n \times m$, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$, and an approximation $\overline{x} \in \mathbb{R}^n$ of a solution of this system. Find x^+ which essentially better approximates a solution $x^* \in M = \{x : A^{\mathrm{T}}x \leqslant b\}$ or detect that $M = \emptyset$.

The best possibility would be to take $x^+ = P_M\overline{x}$ but the evaluation of such a projection is often too expensive. On the other hand, x^+ evaluated as a projection $P_{\{x:a_i^{\mathrm{T}}x\leqslant b_i\}}\overline{x}$, where a_i is the ith column of A and $i \in I = \{1, \ldots, m\}$ such that $a_i^{\mathrm{T}}\overline{x} > b_i$ is often not essentially a better approximation of a solution x^* than $\overline{x}$. There is also a compromise: choose appropriate columns $J \subset I$ of the matrix A (denote by A_J the submatrix of A which consists of the chosen columns $J \subset I$ and by b_J the subvector of b which consists of the coordinates $J \subset I$) and evaluate

E-mail address: a.cegielski@im.pz.zgora.pl (A. Cegielski).

0024-3795/01/$ - see front matter © 2001 Elsevier Science Inc. All rights reserved.
PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 2 8 4 - 1

$$x^+ = P_{\{x:A_J^{\mathrm{T}}x \leqslant b_J\}}\overline{x}. \quad (1)$$

Now, a new problem arises in this context: how to choose $J \subset I$ such that x^+ evaluated by (1) essentially better approximates a solution x^* than $\overline{x}$ and such that x^+ can easily be evaluated. It seems natural to choose J such that $A_J^{\mathrm{T}}x > b_J$ or at least $A_J^{\mathrm{T}}x \geqslant b_J$. Suppose that A_J has full column rank and evaluate $x' = P_{\{x:A_J^{\mathrm{T}}x=b_J\}}\overline{x}$. Of course, x' is not necessarily equal to x^+ given by (1). Nevertheless, it can be easily shown that equality $x' = x^+$ holds if and only if $\lambda := (A_J^{\mathrm{T}}A_J)^{-1}(A_J^{\mathrm{T}}\overline{x} - b_J) \geqslant 0$ (in fact, λ is the Lagrange multipliers vector for the problem $\min\{\frac{1}{2}\|x-\overline{x}\|^2 : A_J^{\mathrm{T}}x \leqslant b_J\}$). Since we have assumed $A_J^{\mathrm{T}}x \geqslant b_J$ we see that $\lambda \geqslant 0$ when $(A_J^{\mathrm{T}}A_J)^{-1} \geqslant 0$. This observation leads to a study of the properties of a full column rank matrix whose Gram matrix has non-negative inverse. Such matrices are strictly connected with the so-called acute cones (see, e.g. [3,4,10,12]). We are interested in our study on the properties of such matrices which can help us to construct obtuse cones. The projections evaluated by (1) have in these cases interesting properties which lead to an essentially better approximation x^+ of a solution than the actual one.

We start with the notations and definitions which will be occurring in the paper. We use the following notations:

- $x = (x_1, \ldots, x_n)^{\mathrm{T}}$—an element of $\mathbb{R}^n$,
- $\langle\cdot,\cdot\rangle$—the usual scalar product in $\mathbb{R}^n$, i.e., $\langle x, y\rangle = x^{\mathrm{T}}y = \sum_{j=1}^n x_j y_j$, for $x, y \in \mathbb{R}^n$,
- $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$—the Euclidean norm in $\mathbb{R}^n$,
- e_i—ith element of the standard basis in $\mathbb{R}^n$,
- $P_D x$—the metric projection of x onto a given closed convex subset D.

Furthermore, for a given matrix A of type $n \times m$, we denote:

- $\mathrm{Lin}\, A$—the linear subspace generated by the columns of A,
- $(\mathrm{Lin}\, A)^{\perp}$—the linear subspace orthogonal to $\mathrm{Lin}\, A$,
- $\mathrm{cone}\, A$—the cone generated by the columns of A,
- $\mathrm{r}\,(A)$—the rank of A,

and we write:

- $A \geqslant 0$ (>0) if all elements of A are non-negative (positive),
- A has full column rank if $\mathrm{r}\,(A) = m$ (or in other words, the columns of A are linearly independent).

Finally, for a given cone C we define the dual cone of C by

$$C^* = \{s \in \mathbb{R}^n : \langle s, x\rangle \leqslant 0 \text{ for all } x \in C\}.$$

The matrices we consider in the paper are real matrices.

Definition 1.1. A cone C is said to be *acute* if $\langle x, y\rangle \geqslant 0$ for all $x, y \in C$. A cone C is said to be *obtuse* (in $\mathrm{Lin}\, C$) if $C^* \cap \mathrm{Lin}\, C$ is an acute cone.

Definition 1.2. Let A be a matrix of type $n \times m$. The matrix $A^{\mathrm{T}}A$ is said to be the Gram matrix of A.

Remark 1.3. It is well known that the Gram matrix of A is non-negative definite and that it is positive definite if and only if A has full column rank.

Definition 1.4. Let A be a matrix with full column rank. The matrix

$$A^+ = (A^\mathrm{T} A)^{-1} A^\mathrm{T} \tag{2}$$

is said to be the (Moore–Penrose) *pseudoinverse* of A.

Definition 1.5. A matrix G is said to be monotone if

$$Gx \geqslant 0 \Rightarrow x \geqslant 0.$$

The following result which connects the introduced objects is well known (see, e.g. [1–4,6,10,12]). Nevertheless, we present in Appendix A a proof of the following lemma.

Lemma 1.6. *Let A have full column rank. The following conditions are equivalent*:
(i) cone A *is obtuse*,
(ii) cone $A^{+\mathrm{T}}$ *is acute*,
(iii) $(\text{cone } A)^* \cap \text{Lin } A \subset -\text{cone } A$,
(iv) $(A^\mathrm{T} A)^{-1} \geqslant 0$,
(v) $A^\mathrm{T} A$ *is monotone.*

The paper is organized as follows. In Section 2, we present some properties of Gram matrices with non-negative inverse, which lead to the constructions of an obtuse cone. These constructions consist in an addition of one vector to a given system of vectors generating an obtuse cone. Such properties and constructions are known (see, e.g. [3,4,10,12]). Therefore, Section 2 has an auxiliary character. The main result of this section, Theorem 2.3, will be generalized in the following section in Lemma 3.6 and in Theorem 3.11. In Section 3, we present the sufficient and necessary conditions for a block matrix $A = [A_1, A_2]$ having a Gram matrix with non-negative inverse. These conditions lead to the construction of an obtuse cone which consists in addition of several vectors to a given system of vectors generating an obtuse cone. In Section 4, we present the results of Section 3 in terms of the QR-factorization. Finally, in Appendix A, we present a proof of Lemma 1.6.

2. One-dimensional update of obtuse cones

Definition 2.1. A matrix A of type $n \times m$ is said to be *strongly admissible* if there exists $x \in \mathbb{R}^n$ such that $A^\mathrm{T} x > 0$. A matrix A of type $n \times m$ is said to be *weakly admissible* if there exists $x \in \mathbb{R}^n$ such that $A^\mathrm{T} x \geqslant 0$ with at least one positive coordinate.

Denote by A_1 a matrix of type $n \times m_1$ with columns $a_i,\ i = 1, \ldots, m_1$, by a an element of $\mathbb{R}^n$ and by A the $n \times m$ matrix $[A_1, a]$, where $m = m_1 + 1$.

Lemma 2.2. *Let* $A = [A_1, a]$ *be weakly admissible and let* $x \in \mathbb{R}^n$ *and* $j \in \{1, \ldots, m\}$ *be such that* $A^{\mathrm{T}}x \geqslant 0$ *with positive* jth *coordinate. If*
(i) A_1 *has full column rank*,
(ii) $A_1^+ a \leqslant 0$ *with negative* jth *coordinate if* $j < m$,
then A *has full column rank.*

Proof. Suppose $\mathrm{r}\,(A) < m$. Then, by the Kronecker–Capelli theorem, $a = A_1\alpha$ for some $\alpha \in \mathbb{R}^{m_1}$ since A_1 has full column rank. Furthermore, this representation is unique. Let $x \in \mathbb{R}^n$ and $j \leqslant m$ be such that $A^{\mathrm{T}}x \geqslant 0$ with positive jth coordinate. Such x and j exist since A is weakly admissible. By (ii) we have

$$A_1^+ a = A_1^+ A_1 \alpha = \alpha \leqslant 0$$

with $\alpha_j < 0$ if $j < m$. The above claims leads to the following inequalities:

$$0 \leqslant a^{\mathrm{T}}x = \alpha^{\mathrm{T}} A_1^{\mathrm{T}} x \leqslant 0,$$

where the first inequality is strict if $j = m$ and the second one is strict if $j < m$. The contradiction shows that A has full column rank. □

Theorem 2.3. *Let* $A = [A_1, a]$ *be a matrix with full column rank. If*
(i) $(A_1^{\mathrm{T}} A_1)^{-1} \geqslant 0$,
(ii) $A_1^+ a \leqslant 0$,
then
(iii) $(A^{\mathrm{T}} A)^{-1} \geqslant 0$.

Remark 2.4. A proof of a lemma which is formulated equivalently to Theorem 2.3 can be found in [2, Lemma 5.6.B]. In the quoted lemma, the condition $A_1^+ a \leqslant 0$ has an equivalent form $a \in -\mathrm{cone}\, A_1 + (\mathrm{Lin}\, A_1)^{\perp}$. Theorem 2.3 also follows from [10, Lemma 3.2]. In Section 3 we will prove a generalization of this theorem.

Corollary 2.5. *Let* $A = [A_1, a]$ *be weakly admissible and let* $j \in \{1, \ldots, m\}$ *be such that the* jth *coordinate of* $A^{\mathrm{T}}x$ *is positive for some* j, *and for some* $x \in \mathbb{R}^n$. *If*
(i) A_1 *has full column rank*,
(ii) $(A_1^{\mathrm{T}} A_1)^{-1} \geqslant 0$,
(iii) $A_1^+ a \leqslant 0$ *with negative* jth *coordinate if* $j < m$,
then
(iv) A *has full column rank and*
(v) $(A^{\mathrm{T}} A)^{-1} \geqslant 0$.

Proof. The corollary follows directly from Theorem 2.3 and Lemma 2.2. □

Remark 2.6. Compare also [10, Lemma 3.2], where the above corollary is proved, or [4, Remark 10], where a proof of the above corollary is suggested.

Corollary 2.7. *Let* $A = [A_1, a]$ *be strongly admissible. If*

(i) A_1 *has full column rank*,

(ii) $(A_1^{\mathrm{T}} A_1)^{-1} \geqslant 0$,

(iii) $A_1^{+} a \leqslant 0$,

then

(iv) A *has full column rank and*

(v) $(A^{\mathrm{T}} A)^{-1} \geqslant 0$.

Proof. Take $j = m$ in Corollary 2.5 which can be done since A is strongly admissible. □

Remark 2.8. A proof of Corollary 2.7 can also be found in [10, Lemma 3.2] or in [4, Theorem 5].

Corollary 2.9. *Let* $A = [A_1, a]$ *be a matrix with full column rank. If*

(i) $(A_1^{\mathrm{T}} A_1)^{-1} \geqslant 0$,

(ii) $A_1^{\mathrm{T}} a \leqslant 0$,

then

(iii) $(A^{\mathrm{T}} A)^{-1} \geqslant 0$.

Proof. From assumptions (i) and (ii) it follows that

$$A_1^{+} a = (A_1^{\mathrm{T}} A_1)^{-1} A_1^{\mathrm{T}} a \leqslant 0.$$

The corollary now follows from Theorem 2.3 □

Remark 2.10. Other proofs of Corollary 2.9 can be found in [12, Corollary 4.1] or in [2, Theorem 5.4.A].

Corollary 2.11. *Let* $A = [A_1, a]$ *be weakly admissible and let the* jth *coordinate of* $A_1^{\mathrm{T}} x$ *be positive for some* $x \in \mathbb{R}^n$. *If*

(i) A_1 *has full column rank*,

(ii) $(A_1^{\mathrm{T}} A_1)^{-1} \geqslant 0$,

(iii) $A_1^{\mathrm{T}} a \leqslant 0$ *with negative* jth *coordinate*,

then

(iv) *A has full column rank and*

(v) $(A^{\mathrm{T}}A)^{-1} \geqslant 0$.

Proof. From assumptions (ii) and (iii) it follows that

$$A_1^+ a = (A_1^{\mathrm{T}} A_1)^{-1} A_1^{\mathrm{T}} a \leqslant 0.$$

Denote by α_i the ith coordinate of $A_1^+ a$, by b_i the ith coordinate of $A_1^{\mathrm{T}} a$ and by b_{ik} the (i, k)th element of $(A_1^{\mathrm{T}} A_1)^{-1}$. Of course $b_{jj} > 0$ since all diagonal elements of a positive definite matrix are positive. By (ii) and (iii) we have

$$\alpha_j = \sum_{k=1}^{m_1} b_{jk} b_k \leqslant b_{jj} b_j < 0.$$

Now, we obtain by Lemma 2.2 that A has full column rank, and by Theorem 2.3 that $(A^{\mathrm{T}}A)^{-1} \geqslant 0$. □

Remark 2.12. Similarly as in Corollary 2.7 one can also formulate the following version of Corollary 2.11.

Corollary 2.13. *If the Gram matrix G of A has full column rank and non-positive off-diagonal elements, then $G^{-1} \geqslant 0$.*

Proof. The corollary can be verified by the repeated application of Corollary 2.9. □

Remark 2.14. Similarly as in Corollary 2.11 one can formulate versions of Corollary 2.13 by an appropriate use of strong or weak admissibility instead of linear independency.

Remark 2.15. Corollary 2.13 follows also from [5, Theorem 4.3] and from [11, Theorem 5′] which deal with the so-called *Minkowski matrices* (non-singular matrices with non-positive off-diagonal elements and with non-negative inverse). Some other formulations of Corollary 2.13 in terms of cones can be found in [2,7–9,12].

Remark 2.16. Theorem 2.3 allows more general constructions of obtuse cones than those obtained by application of Corollaries 2.9 or 2.13 which can be seen in the example below.

Example 2.17. Let $A = [a_1, a_2, a_3]$, where $a_1 = (1, 0, 0)^{\mathrm{T}}$, $a_2 = (-1, -\frac{1}{4}, \frac{1}{4})^{\mathrm{T}}$, $a_3 = (-1, \frac{1}{4}, \frac{1}{4})^{\mathrm{T}}$. One can see that cone A is obtuse since

$$(A^{\mathrm{T}}A)^{-1} = \begin{bmatrix} 17 & 8 & 8 \\ 8 & 8 & 0 \\ 8 & 0 & 8 \end{bmatrix} \geqslant 0.$$

On the other hand, $a_2^{\mathrm{T}} a_3 > 0$ and cone A cannot be constructed by an application of Corollaries 2.9 or 2.13. Nevertheless, cone A can be constructed by an iterative application of Theorem 2.3: take in the first iteration $A_1 = a_1,\ a = a_2$, and in the second one $A_1 = [a_1, a_2],\ a = a_3$.

3. A general construction of obtuse cones

We start with the following example which shows that there exist obtuse cones which cannot be obtained by an application of the constructions presented in Section 2.

Example 3.1. Consider the following matrix G:

$$G = \begin{bmatrix} \frac{175}{88} & \frac{65}{88} & -\frac{25}{22} & -\frac{25}{22} \\ \frac{65}{88} & \frac{175}{88} & -\frac{25}{22} & -\frac{25}{22} \\ -\frac{25}{22} & -\frac{25}{22} & \frac{175}{88} & \frac{65}{88} \\ -\frac{25}{22} & -\frac{25}{22} & \frac{65}{88} & \frac{175}{88} \end{bmatrix}.$$

G is positive definite and its inverse is equal to

$$G^{-1} = \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{5} & 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 & \frac{1}{5} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{5} & 1 \end{bmatrix},$$

i.e., by Lemma 1.6, G is the Gram matrix of a system of vectors $A = [a_1, a_2, a_3, a_4]$ which generates an obtuse cone. Nevertheless, none of the 3-element subsystems of A generate an obtuse cone, because if one cancels the ith row and the ith column from the matrix G ($i = 1, 2, 3, 4$), then the obtained matrix does not have non-negative inverse:

$$\begin{bmatrix} \frac{175}{88} & -\frac{25}{22} & -\frac{25}{22} \\ -\frac{25}{22} & \frac{175}{88} & \frac{65}{88} \\ -\frac{25}{22} & \frac{65}{88} & \frac{175}{88} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{24}{25} & \frac{2}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{3}{4} & -\frac{1}{20} \\ \frac{2}{5} & -\frac{1}{20} & \frac{3}{4} \end{bmatrix},$$

$$\begin{bmatrix} \frac{175}{88} & \frac{65}{88} & -\frac{25}{22} \\ \frac{65}{88} & \frac{175}{88} & -\frac{25}{22} \\ -\frac{25}{22} & -\frac{25}{22} & \frac{175}{88} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{3}{4} & -\frac{1}{20} & \frac{2}{5} \\ -\frac{1}{20} & \frac{3}{4} & \frac{2}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{24}{25} \end{bmatrix}.$$

Therefore, the system A which generates an obtuse cone cannot be constructed by an application of Theorem 2.3 or of any of its corollaries presented in Section 2.

Observe that there exist 2-element subsystems $[a_1, a_3]$ and $[a_2, a_4]$ of A which generate obtuse cones since

$$\begin{bmatrix} \frac{175}{88} & -\frac{25}{22} \\ -\frac{25}{22} & \frac{175}{88} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{56}{75} & \frac{32}{75} \\ \frac{32}{75} & \frac{56}{75} \end{bmatrix} \geqslant 0.$$

In this section, we present a construction of obtuse cones which is more general than that obtained by an application of Theorem 2.3.

Lemma 3.2. *A matrix B is the pseudoinverse of a full column rank matrix A if and only if*
(i) $BA = I$ (*i.e.*, *B is a left inverse of* A);
(ii) $B = \Lambda A^{\mathrm{T}}$ *for some square matrix* Λ (*i.e.*, *rows of* $B \in \operatorname{Lin} A$).

Proof. The necessity of (i) and (ii) is obvious. Now, we prove the sufficiency. By (i) and (ii) we have $I = BA = \Lambda A^{\mathrm{T}} A$, i.e., $\Lambda = (A^{\mathrm{T}} A)^{-1}$. Hence, $B = \Lambda A^{\mathrm{T}} = (A^{\mathrm{T}} A)^{-1} A^{\mathrm{T}}$. □

Let A_1, A_2 be matrices of types $n \times m_1$ and $n \times m_2$, respectively, and let the matrix $A = [A_1, A_2]$ have full column rank. One can decompose A_1 into a sum of a part linearly dependent on the columns of A_2 and a part orthogonal to $\operatorname{Lin} A_2$. We have

$$A_1 = A_2 A_2^{+} A_1 + A_2^{\perp}, \tag{3}$$

where

$$A_2^{\perp} = (I - A_2 A_2^{+}) A_1. \tag{4}$$

There hold the obvious equalities

$$A_2^{\perp\mathrm{T}} A_2 = 0 \tag{5}$$

and

$$A_2^{\perp\mathrm{T}} A_1 = A_2^{\perp\mathrm{T}} A_2^{\perp}. \tag{6}$$

Analogously, the matrix A_2 can be decomposed into a sum of a part linearly dependent on the columns of A_1 and a part orthogonal to $\operatorname{Lin} A_1$. So, we have

$$A_2 = A_1 A_1^{+} A_2 + A_1^{\perp}, \tag{7}$$

where

$$A_1^{\perp} = (I - A_1 A_1^{+}) A_2 \tag{8}$$

and

$$A_1^{\perp\mathrm{T}} A_1 = 0, \tag{9}$$

$$A_1^{\perp\mathrm{T}} A_2 = A_1^{\perp\mathrm{T}} A_1^{\perp}. \tag{10}$$

Furthermore, by equalities

$$A_2^{\perp}\alpha = A_1\alpha - A_2A_2^{+}A_1\alpha = A\begin{bmatrix}\alpha \\ -A_2^{+}A_1\alpha\end{bmatrix},$$

where $\alpha \in \mathbb{R}^{m_1}$, one easily obtains from the definition that the columns of $A_2^{\perp}$ are linearly independent since the columns of A are linearly independent. Therefore $\mathrm{r}\,(A_2^{\perp}) = m_1$, consequently, there exist matrices $A_2^{\perp+}$, $(A_2^{\perp\mathrm{T}}A_2^{\perp})^{+}$ and $(A_2^{\perp\mathrm{T}}A_1)^{+}$. Correspondingly, $\mathrm{r}\,(A_1^{\perp}) = m_2$ and there exist matrices $A_1^{\perp+}$, $(A_1^{\perp\mathrm{T}}A_1^{\perp})^{+}$ and $(A_1^{\perp\mathrm{T}}A_2)^{+}$.

Lemma 3.3. *The matrix*

$$A' = \begin{bmatrix} A_1^{+} - A_1^{+}A_2A_1^{\perp+} \\ A_2^{+} - A_2^{+}A_1A_2^{\perp+} \end{bmatrix}$$

is the pseudoinverse of the matrix A.

Proof. First we prove that $A' = \Lambda A^{\mathrm{T}}$ for

$$\Lambda = \begin{bmatrix} (A_1^{\mathrm{T}}A_1)^{-1} + A_1^{+}A_2(A_1^{\perp\mathrm{T}}A_1^{\perp})^{-1}(A_1^{+}A_2)^{\mathrm{T}} & -A_1^{+}A_2(A_1^{\perp\mathrm{T}}A_1^{\perp})^{-1} \\ -A_2^{+}A_1(A_2^{\perp\mathrm{T}}A_2^{\perp})^{-1} & (A_2^{\mathrm{T}}A_2)^{-1} + A_2^{+}A_1(A_2^{\perp\mathrm{T}}A_2^{\perp})^{-1}(A_2^{+}A_1)^{\mathrm{T}} \end{bmatrix}.$$

By (7), (3) and (2) we have

$$\begin{aligned} \Lambda A^{\mathrm{T}} &= \begin{bmatrix} A_1^{+} - A_1^{+}A_2(A_1^{\perp\mathrm{T}}A_1^{\perp})^{-1}(A_2 - A_1A_1^{+}A_2)^{\mathrm{T}} \\ A_2^{+} - A_2^{+}A_1(A_2^{\perp\mathrm{T}}A_2^{\perp})^{-1}(A_1 - A_2A_2^{+}A_1)^{\mathrm{T}} \end{bmatrix} \\ &= \begin{bmatrix} A_1^{+} - A_1^{+}A_2(A_1^{\perp\mathrm{T}}A_1^{\perp})^{-1}A_1^{\perp\mathrm{T}} \\ A_2^{+} - A_2^{+}A_1(A_2^{\perp\mathrm{T}}A_2^{\perp})^{-1}A_2^{\perp\mathrm{T}} \end{bmatrix} = A'. \end{aligned}$$

Furthermore, by equalities (9), (10), (5) and (6), and by the definition of the pseudo-inverse we easily obtain $A'A = I$. The lemma follows now from Lemma 3.2. □

Lemma 3.4. *The matrix*

$$A'' = \begin{bmatrix} A_2^{\perp+} \\ A_1^{\perp+} \end{bmatrix}$$

is the pseudoinverse of the matrix A.

Proof. The matrix A'' can be presented in the form $A'' = \Lambda A^{\mathrm{T}}$ for

$$\Lambda = \begin{bmatrix} (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} & -(A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} (A_2^{+} A_1)^{\mathrm{T}} \\ -(A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} (A_1^{+} A_2)^{\mathrm{T}} & (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} \end{bmatrix}.$$

Indeed, by equalities (3), (7) and (2) we have

$$\begin{aligned} \Lambda A^{\mathrm{T}} &= \begin{bmatrix} (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} A_1^{\mathrm{T}} - (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} (A_2^{+} A_1)^{\mathrm{T}} A_2^{\mathrm{T}} \\ -(A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} (A_1^{+} A_2)^{\mathrm{T}} A_1^{\mathrm{T}} + (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} A_2^{\mathrm{T}} \end{bmatrix} \\ &= \begin{bmatrix} (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} (A_1 - A_2 A_2^{+} A_1)^{\mathrm{T}} \\ (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} (A_2 - A_1 A_1^{+} A_2)^{\mathrm{T}} \end{bmatrix} \\ &= \begin{bmatrix} (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} A_2^{\perp\mathrm{T}} \\ (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} A_1^{\perp\mathrm{T}} \end{bmatrix} = \begin{bmatrix} A_2^{\perp+} \\ A_1^{\perp+} \end{bmatrix}. \end{aligned}$$

Furthermore, by equalities (3), (7), (5) and (9), we obtain

$$\begin{aligned} A'' A &= \begin{bmatrix} A_2^{\perp+} \\ A_1^{\perp+} \end{bmatrix} [A_1, A_2] = \begin{bmatrix} A_2^{\perp+} A_1 & A_2^{\perp+} A_2 \\ A_1^{\perp+} A_1 & A_1^{\perp+} A_2 \end{bmatrix} \\ &= \begin{bmatrix} (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} A_2^{\perp\mathrm{T}} (A_2 A_2^{+} A_1 + A_2^{\perp}) & (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} A_2^{\perp\mathrm{T}} A_2 \\ (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} A_1^{\perp\mathrm{T}} A_1 & (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} A_1^{\perp\mathrm{T}} (A_1 A_1^{+} A_2 + A_1^{\perp}) \end{bmatrix} \\ &= \begin{bmatrix} (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} A_2^{\perp\mathrm{T}} A_2^{\perp} & 0 \\ 0 & (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} A_1^{\perp\mathrm{T}} A_1^{\perp} \end{bmatrix} = \begin{bmatrix} I_{m_1} & 0 \\ 0 & I_{m_2} \end{bmatrix} = I. \end{aligned}$$

The lemma follows now from Lemma 3.2. □

Corollary 3.5. *There holds the equality* $A' = A''$.

Proof. The assertion follows from Lemmas 3.3 and 3.4. Namely, the pseudoinverse of a matrix is uniquely determined. □

Observe that, for a matrix B with full column rank, there holds the equality

$$B^{+} B^{+\mathrm{T}} = (B^{\mathrm{T}} B)^{-1}. \tag{11}$$

Present the matrix $(A^{\mathrm{T}} A)^{-1}$ in the form of a block matrix

$$(A^{\mathrm{T}} A)^{-1} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix},$$

where $D_{11},\ D_{12},\ D_{21},\ D_{22}$ are matrices of types $m_1 \times m_1, m_1 \times m_2, m_2 \times m_1$ and $m_2 \times m_2$, respectively. Similarly, present the matrices A' and A'' in Lemmas 3.3 and 3.4 in the form of block matrices:

$$A' = \begin{bmatrix} A'_1 \\ A'_2 \end{bmatrix},$$

$$A'' = \begin{bmatrix} A''_1 \\ A''_2 \end{bmatrix},$$

where $A'_1,\ A''_1$ are matrices of type $m_1 \times n$ and $A'_2,\ A''_2$ are matrices of type $m_2 \times n$.

Lemma 3.6. *There hold equalities*

$$D_{11} = (A_1^{\mathrm{T}} A_1)^{-1} + A_1^{+} A_2 (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} (A_1^{+} A_2)^{\mathrm{T}} = (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1}, \tag{12}$$

$$D_{12} = -A_1^{+} A_2 (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1}, \tag{13}$$

$$D_{21} = -A_2^{+} A_1 (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1}, \tag{14}$$

$$D_{22} = (A_2^{\mathrm{T}} A_2)^{-1} + A_2^{+} A_1 (A_2^{\perp\mathrm{T}} A_2^{\perp})^{-1} (A_2^{+} A_1)^{\mathrm{T}} = (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1}, \tag{15}$$

consequently, $(A^{\mathrm{T}}A)^{-1} \geqslant 0$ *if and only if all matrices given by* (12)–(15) *are non-negative.*

Proof. By equality (11), we have

$$(A^{\mathrm{T}}A)^{-1} = \begin{bmatrix} A'_1 \\ A'_2 \end{bmatrix} \begin{bmatrix} A'^{\mathrm{T}}_1 & A'^{\mathrm{T}}_2 \end{bmatrix}$$

and

$$(A^{\mathrm{T}}A)^{-1} = \begin{bmatrix} A''_1 \\ A''_2 \end{bmatrix} \begin{bmatrix} A''^{\mathrm{T}}_1 & A''^{\mathrm{T}}_2 \end{bmatrix}.$$

Now, we evaluate the successive blocks of the matrix $(A^{\mathrm{T}}A)^{-1}$. We have

$$\begin{aligned} D_{11} &= A'_1 A'^{\mathrm{T}}_1 \\ &= A_1^{+} A_1^{+\mathrm{T}} - A_1^{+} A_1^{\perp+\mathrm{T}} A_2^{\mathrm{T}} A_1^{+\mathrm{T}} - A_1^{+} A_2 A_1^{\perp+} A_1^{+\mathrm{T}} \\ &\quad + A_1^{+} A_2 A_1^{\perp+} A_1^{\perp+\mathrm{T}} A_2^{\mathrm{T}} A_1^{+\mathrm{T}}. \end{aligned}$$

By (2) and by (9) $A_1^{+} A_1^{\perp+\mathrm{T}} = 0$. Furthermore, by (11), we get

$$A_1^{\perp+} A_1^{\perp+\mathrm{T}} = (A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1}$$

and

$$A_1^+ A_1^{+\mathrm{T}} = (A_1^\mathrm{T} A_1)^{-1}.$$

Therefore

$$D_{11} = (A_1^\mathrm{T} A_1)^{-1} + A_1^+ A_2 (A_1^{\perp\mathrm{T}} A_1^\perp)^{-1} (A_1^+ A_2)^\mathrm{T}.$$

On the other hand,

$$D_{11} = A_1'' A_1''^\mathrm{T} = (A_2^{\perp\mathrm{T}} A_2^\perp)^{-1}.$$

Using similar arguments we obtain equalities (15). Furthermore, by (7), (5) and (6) we see

$$\begin{aligned}
D_{12} &= A_1'' A_2''^\mathrm{T} = A_2^{\perp+} A_1^{\perp+\mathrm{T}} \\
&= (A_2^{\perp\mathrm{T}} A_2^\perp)^{-1} A_2^{\perp\mathrm{T}} A_1^\perp (A_1^{\perp\mathrm{T}} A_1^\perp)^{-1} \\
&= (A_2^{\perp\mathrm{T}} A_2^\perp)^{-1} A_2^{\perp\mathrm{T}} (A_2 - A_1 A_1^+ A_2)(A_1^{\perp\mathrm{T}} A_1^\perp)^{-1} \\
&= -(A_2^{\perp\mathrm{T}} A_2^\perp)^{-1} A_2^{\perp\mathrm{T}} A_1 A_1^+ A_2 (A_1^{\perp\mathrm{T}} A_1^\perp)^{-1} \\
&= -(A_2^{\perp\mathrm{T}} A_2^\perp)^{-1} A_2^{\perp\mathrm{T}} A_2^\perp A_1^+ A_2 (A_1^{\perp\mathrm{T}} A_1^\perp)^{-1} \\
&= -A_1^+ A_2 (A_1^{\perp\mathrm{T}} A_1^\perp)^{-1}.
\end{aligned}$$

Equality (14) can be proved analogously. □

Corollary 3.7. *There holds the equality*

$$(A_2^{\perp\mathrm{T}} A_2^\perp)^{-1} (A_2^+ A_1)^\mathrm{T} = A_1^+ A_2 (A_1^{\perp\mathrm{T}} A_1^\perp)^{-1}.$$

Proof. The matrix $(A^\mathrm{T} A)^{-1}$ is symmetric. Therefore, for the blocks D_{12} and D_{21} given by (13) and (14) the equality $D_{21}^\mathrm{T} = D_{12}$ is fulfilled. □

Corollary 3.8. *There holds the equality*

$$D_{11} = (A_1^\mathrm{T} A_1)^{-1} + D_{12} D_{22}^{-1} D_{21}.$$

Proof. The corollary follows easily from Lemma 3.6 and from the symmetry of the matrix $(A^\mathrm{T} A)^{-1}$. □

Corollary 3.9. *Let $A = [A_1, a]$ be a matrix with full column rank, where A_1 is an $n \times m_1$ matrix and $a \in \mathbb{R}^n$. Furthermore, let $a^\perp = (I - A_1 A_1^+)a$. Then the matrix $(A^\mathrm{T} A)^{-1}$ has the representation*

$$(A^\mathrm{T} A)^{-1} = \begin{bmatrix} (A_1^\mathrm{T} A_1)^{-1} + (A_1^+ a)(A_1^+ a)^\mathrm{T}/\|a^\perp\|^2 & -A_1^+ a/\|a^\perp\|^2 \\ -(A_1^+ a)^\mathrm{T}/\|a^\perp\|^2 & 1/\|a^\perp\|^2 \end{bmatrix}$$

and, consequently, $(A^\mathrm{T} A)^{-1}$ is non-negative if and only if

(i) $A_1^+ a \leqslant 0$ *and*

(ii) $(A_1^\mathrm{T} A_1)^{-1} + (A_1^+ a)(A_1^+ a)^\mathrm{T}/\|a^\perp\|^2 \geqslant 0.$

Proof. The corollary follows easily if one takes $A_2 = a$ in equalities (12)–(15). □

Remark 3.10. Theorem 2.3 follows now from Corollary 3.9. Furthermore, the results of [10, Lemma 3.2] and [4, Theorem 5] follow from Corollary 3.9 and Lemma 2.2.

Theorem 3.11. *Let* $A = [A_1, A_2]$ *be a matrix with full column rank. If*
(i) $(A_1^{\mathrm{T}} A_1)^{-1} \geqslant 0$,
(ii) $A_1^{+} A_2 \leqslant 0$,
(iii) $(A_1^{\perp\mathrm{T}} A_1^{\perp})^{-1} \geqslant 0$,
then $(A^{\mathrm{T}} A)^{-1} \geqslant 0$.

Proof. The theorem follows easily from Lemma 3.6. □

4. Obtuse cones and QR-factorization

In this section we present a form of the matrix $(A^{\mathrm{T}} A)^{-1}$ based on the QR-factorization of the block matrix $A = [A_1, A_2]$. Let $A = [A_1, A_2]$ be a matrix of type $n \times (m_1 + m_2)$ with full column rank and let $A = QR$ be a QR-factorization of the matrix $A = [A_1, A_2]$. Let $Q = [Q_1, Q_2]$, where Q_1 and Q_2 are orthogonal matrices of types $n \times m_1$ and $n \times m_2$, respectively. Furthermore, let

$$R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix},$$

where $R_{11},\ R_{12},\ R_{22}$ are matrices of types $m_1 \times m_1,\ m_1 \times m_2$ and $m_2 \times m_2$ and 0 is a matrix of type $m_2 \times m_1$ with all elements equal to 0. The following equalities are obvious:

$$A_1 = Q_1 R_{11}, \tag{16}$$

$$A_1^{+} = R_{11}^{-1} Q_1^{\mathrm{T}} \tag{17}$$

and

$$A_2 = A_1 R_{11}^{-1} R_{12} + Q_2 R_{22}. \tag{18}$$

Theorem 4.1. *There holds the equality*

$$(A^{\mathrm{T}} A)^{-1} = \begin{bmatrix} (R_{11}^{\mathrm{T}} R_{11})^{-1} + R_{11}^{-1} R_{12} (R_{22}^{\mathrm{T}} R_{22})^{-1} R_{12}^{\mathrm{T}} R_{11}^{-\mathrm{T}} & -R_{11}^{-1} R_{12} (R_{22}^{\mathrm{T}} R_{22})^{-1} \\ -(R_{22}^{\mathrm{T}} R_{22})^{-1} R_{12}^{\mathrm{T}} R_{11}^{-\mathrm{T}} & (R_{22}^{\mathrm{T}} R_{22})^{-1} \end{bmatrix}$$

and, consequently, $(A^{\mathrm{T}} A)^{-1}$ *is non-negative if and only if*
(i) $(R_{22}^{\mathrm{T}} R_{22})^{-1} \geqslant 0$,
(ii) $(R_{22}^{\mathrm{T}} R_{22})^{-1} R_{12}^{\mathrm{T}} R_{11}^{-\mathrm{T}} \leqslant 0$,
(iii) $(R_{11}^{\mathrm{T}} R_{11})^{-1} + R_{11}^{-1} R_{12} (R_{22}^{\mathrm{T}} R_{22})^{-1} R_{12}^{\mathrm{T}} R_{11}^{-\mathrm{T}} \geqslant 0$.

Proof. By equalities (8), (16)–(18) and by the properties of orthogonal matrices we have

$$
\begin{aligned}
A_1^{\perp\mathrm{T}} A_1^{\perp} &= A_2^{\mathrm{T}}(I - A_1 A_1^{+})^{\mathrm{T}}(I - A_1 A_1^{+})A_2 \\
&= A_2^{\mathrm{T}}(I - A_1 A_1^{+})^{\mathrm{T}} A_2 \\
&= R_{12}^{\mathrm{T}} R_{11}^{-\mathrm{T}} A_1^{\mathrm{T}} A_1 R_{11}^{-1} R_{12} + R_{12}^{\mathrm{T}} R_{11}^{-\mathrm{T}} A_1^{\mathrm{T}} Q_2 R_{22} \\
&\quad - R_{12}^{\mathrm{T}} R_{11}^{-\mathrm{T}} A_1^{\mathrm{T}} Q_1 Q_1^{\mathrm{T}} A_1 R_{11}^{-1} R_{12} - R_{12}^{\mathrm{T}} R_{11}^{-\mathrm{T}} A_1^{\mathrm{T}} Q_1 Q_1^{\mathrm{T}} Q_2 R_{22} \\
&\quad + R_{22}^{\mathrm{T}} Q_2^{\mathrm{T}} A_1 R_{11}^{-1} R_{12} + R_{22}^{\mathrm{T}} Q_2^{\mathrm{T}} Q_2 R_{22} \\
&= R_{22}^{\mathrm{T}} R_{22}.
\end{aligned}
$$

Furthermore, by equalities (17) and (18) and by the properties of orthogonal matrices, we obtain

$$A_1^{+} A_2 = R_{11}^{-1} R_{12}.$$

The theorem follows now from Lemma 3.6 and Corollary 3.7. □

Appendix A

Proof of Lemma 1.6 (i) $\Leftrightarrow$ (ii) Using the well-known equality $\mathbb{R}^n = \operatorname{Lin} A + (\operatorname{Lin} A)^{\perp}$ one can prove that $(\operatorname{cone} A)^* = -\operatorname{cone} A^{+\mathrm{T}} + (\operatorname{Lin} A)^{\mathrm{T}}$. As a consequence one obtains $\operatorname{cone} A^{+\mathrm{T}} = -(\operatorname{cone} A)^* \cap \operatorname{Lin} A$. Now the equivalence (i) $\Leftrightarrow$ (ii) follows directly from Definition 1.1.

(i) $\Rightarrow$ (iii) Suppose (i) holds and (iii) does not hold. Let $y \in (\operatorname{cone} A)^* \cap \operatorname{Lin} A$ and $y \notin -\operatorname{cone} A$. Then, by the separation theorem there exists z such that $\langle z, x\rangle \geqslant 0$ for all $x \in -\operatorname{cone} A$, and $\langle z, y\rangle < 0$. It is well known that $z = \overline{z} + z^{\perp}$, where $\overline{z} \in \operatorname{Lin} A$ and $z^{\perp} \in (\operatorname{Lin} A)^{\perp}$. We have $\overline{z} \in (\operatorname{cone} A)^*$ since $\langle \overline{z}, x\rangle = \langle z - z^{\perp}, x\rangle = \langle z, x\rangle \geqslant 0$ for all $x \in -\operatorname{cone} A$. Now, it follows from (i) that $\langle z, y\rangle = \langle \overline{z} + z^{\perp}, y\rangle = \langle \overline{z}, y\rangle \geqslant 0$ since $z^{\perp} \in (\operatorname{Lin} A)^{\perp}$ and $\overline{z}, y \in (\operatorname{cone} A)^* \cap \operatorname{Lin} A$. We have obtained a contradiction which proves that (i) $\Rightarrow$ (iii).

(iii) $\Rightarrow$ (iv) Suppose that $(\operatorname{cone} A)^* \cap \operatorname{Lin} A \subset -\operatorname{cone} A$ and that $b_{ij} < 0$ for some element b_{ij} of the matrix $(A^{\mathrm{T}}A)^{-1}$. Let $y = A(A^{\mathrm{T}}A)^{-1}(-e_j)$. Then $A^{\mathrm{T}} y = -e_j \leqslant 0$, i.e., $y \in (\operatorname{cone} A)^* \cap \operatorname{Lin} A$. Therefore, by the assumption, $y \in -\operatorname{cone} A$, i.e. $y = A\alpha$ for some $\alpha \leqslant 0$. Furthermore, α is uniquely determined since A has full column rank. Now we have $A(A^{\mathrm{T}}A)^{-1}(-e_j) = y = A\alpha$, and consequently, $\alpha = (A^{\mathrm{T}}A)^{-1}(-e_j)$, i.e., $\alpha_i = -b_{ij} > 0$, which is a contradiction.

(iv) $\Rightarrow$ (v) Let $y = A^{\mathrm{T}}Ax \geqslant 0$. Then, by the assumption, $x = (A^{\mathrm{T}}A)^{-1}y \geqslant 0$, i.e., $A^{\mathrm{T}}A$ is monotone.

(v) $\Rightarrow$ (i) Let $x, y \in (\operatorname{cone} A)^* \cap \operatorname{Lin} A$. Then $A^{\mathrm{T}}x \leqslant 0$, $A^{\mathrm{T}}y \leqslant 0$ and $x = A\alpha$, $y = A\beta$ for some $\alpha, \beta \in \mathbb{R}^m$. Hence $A^{\mathrm{T}}A\alpha \leqslant 0$, $A^{\mathrm{T}}A\beta \leqslant 0$ and, consequently, $\alpha \leqslant 0$ since $A^{\mathrm{T}}A$ is monotone. Now we have $x^{\mathrm{T}}y = \alpha^{\mathrm{T}}A^{\mathrm{T}}A\beta \geqslant 0$, i.e., by Definition 1.1, $\operatorname{cone} A$ is obtuse. □

References

[1] A. Berman, R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York, 1979.
[2] A. Cegielski, Relaxation Methods in Convex Optimization Problems, Higher College of Engineering, Series Monographs, No. 67, Zielona Góra, 1993 (Polish).
[3] A. Cegielski, Projection onto an acute cone and convex feasibility problems, in: J. Henry, J.-P. Yvon (Eds.), Lecture Notes in Control and Information Science, vol. 197, Springer, London, 1994, pp. 187–194.
[4] A. Cegielski, A method of projection onto an acute cone with level control in convex minimization, Math. Programming 85 (1999) 469–490.
[5] M. Fiedler, V. Pták, On matrices with non-positive off-diagonal elements and positive principal minors, Czech. Math. J. 12 (1962) 382–400.
[6] J.L. Goffin, The relaxation method for solving systems of linear inequalities, Math. Oper. Res. 5 (1980) 388–414.
[7] K.C. Kiwiel, Block-iterative surrogate projection methods for convex feasibility problems, Linear Algebra Appl. 215 (1995) 225–259.
[8] K.C. Kiwiel, The efficiency of subgradient projection methods for convex optimization, part I: General level methods, SIAM J. Control Optim. 34 (1996) 660–676.
[9] K.C. Kiwiel, The efficiency of subgradient projection methods for convex optimization, part II: Implementations and extensions, SIAM J. Control Optim. 34 (1996) 677–697.
[10] K.C. Kiwiel, Monotone Gram matrices and deepest surrogate inequalities in accelerated relaxation methods for convex feasibility problems, Linear Algebra Appl. 252 (1997) 27–33.
[11] Ky Fan, Topological proofs for certain theorems on matrices with non-negative elements, Monatsh. Math. 62 (1958) 219–236.
[12] M.J. Todd, Some remarks on the relaxation method for linear inequalities, Technical Report 419, Cornell University, Cornell, Ithaca, 1979.