



Computer and Automation Institute,
Hungarian Academy of Sciences

Research Division

H-1518 Budapest, P.O.Box 63.

ON THE PROJECTION ONTO A FINITELY GENERATED CONE

Ujvári, M.

WP 2007-5
August, 2007

Laboratory of
Operations Research and
Decision Systems

LORDS

On the projection onto a finitely generated cone

Miklós Ujvári *

Abstract. In this paper we study the properties of the projection onto a finitely generated cone. We show for example that this map is made up of finitely many linear parts with a structure resembling the facial structure of the finitely generated cone. An economical algorithm is also presented for calculating the projection of a fixed vector, based on Lemke's algorithm to solve a linear complementarity problem. Some remarks on the conical inverse (a generalization of the Moore-Penrose generalized inverse) conclude the paper.

Mathematics Subject Classifications (2000). 90C20, 90C49

1 Introduction

A standard way to generalize concepts in convex analysis is to replace subspaces with polyhedral cones, polyhedral cones with closed convex cones, closed convex cones with closed convex sets, and closed convex sets with closed convex functions. In this paper we make the first step on this way in the case of the concept of the projection onto a subspace, and examine the properties of the projection onto a finitely generated cone. (For higher levels of generality, see [2].)

Let A be an m by n real matrix. Let $\text{Im } A$ resp. $\text{Ker } A$ denote the range space (that is the image space) and the null space (that is the kernel) of the matrix A . It is well-known that $(\text{Im } A)^\perp = \text{Ker } (A^T)$ and $(\text{Ker } A)^\perp = \text{Im } (A^T)$ where T denotes transpose and $^\perp$ denotes orthogonal complement ([7]).

The projection map $p_{\text{Im } A}$ onto the subspace $\text{Im } A$ can be defined as follows: for every vector $y \in \mathcal{R}^m$, $p_{\text{Im } A}(y)$ is the unique vector Ax_* such that

$$\|y - Ax_*\| = \min_{x \in \mathcal{R}^n} \|y - Ax\|, x_* \in \mathcal{R}^n.$$

It is well-known that $p_{\text{Im } A} : \mathcal{R}^m \rightarrow \mathcal{R}^m$ is a linear map: there exists a unique matrix $P_{\text{Im } A} \in \mathcal{R}^{m \times m}$ such that

$$p_{\text{Im } A}(y) = P_{\text{Im } A} y \quad (y \in \mathcal{R}^m).$$

*H-2600 Vác, Szent János utca 1. HUNGARY

The matrix $P_{\text{Im } A}$ is symmetric (that is $P_{\text{Im } A}^T = P_{\text{Im } A}$), idempotent (that is $P_{\text{Im } A}^2 = P_{\text{Im } A}$), consequently positive semidefinite (that is $y^T P_{\text{Im } A} y \geq 0$ for every $y \in \mathcal{R}^m$). Also the equalities

$$P_{\text{Im } A} \cdot P_{\text{Im } B} = 0, P_{\text{Im } A} + P_{\text{Im } B} = I$$

hold for any matrix B such that $\text{Ker}(A^T) = \text{Im } B$ ([7]). (Here I denotes the identity matrix.)

Analogous results hold also in the case of the projection onto a finitely generated cone. Before stating the corresponding theorem we fix some further notation.

Let $\text{Im}_+ A$ resp. $\text{Ker}_+ A$ denote the so-called finitely generated cone

$$\text{Im}_+ A := \{Ax : 0 \leq x \in \mathcal{R}^n\}$$

and the polyhedral cone

$$\text{Ker}_+ A := \{x \in \mathcal{R}^n : Ax \geq 0\}.$$

A reformulation of Farkas' Lemma ([5]) says that $(\text{Im}_+ A)^* = \text{Ker}_+(A^T)$ and $(\text{Ker}_+ A)^* = \text{Im}_+(A^T)$ where $*$ denotes dual cone. Thus polyhedral cones are the duals of the finitely generated cones, and vice versa. Furthermore, by the Farkas-Weyl-Minkowski Theorem ([5]), for every matrix A there exists a matrix B such that $\text{Im}_+ A = \text{Ker}_+(B^T)$, or, dually, $\text{Ker}_+(A^T) = \text{Im}_+ B$. In other words, the finitely generated cones are polyhedral cones, and vice versa.

The projection map onto $\text{Im}_+ A$ can be defined similarly as in the case of $\text{Im } A$: let $p_{\text{Im}_+ A}(y)$ be the unique vector Ax_* such that

$$\|y - Ax_*\| = \min_{0 \leq x \in \mathcal{R}^n} \|y - Ax\|, 0 \leq x_* \in \mathcal{R}^n.$$

For finitely generated cones which are not subspaces this map is not linear anymore, but is made up of linear parts, and these parts have the properties described already in the case of the projection onto a subspace. To state these facts precisely, the definitions of the faces, complementary faces and polyhedral partition are needed.

Let C be a convex set in \mathcal{R}^d . A convex set $F \subseteq C$ is called an *extremal subset* (or shortly a *face*) of the set C , if F does not contain an inner point of a line segment from C without containing the endpoints of the line segment, that is

$$c_1, c_2 \in C, 0 < \varepsilon < 1, \varepsilon c_1 + (1 - \varepsilon)c_2 \in F \text{ implies } c_1, c_2 \in F.$$

We will denote by $F \triangleleft C$ the fact that F is a face of C .

If K is the finitely generated cone $\text{Im}_+ A$ then its faces are finitely generated cones also, and there are only finitely many of them. The faces of (the finitely generated cone) K^* are exactly the *complementary faces* F^Δ of the faces F of K , defined as

$$F^\Delta := F^\perp \cap (K^*) \quad (F \triangleleft K).$$

The complementary face F^Δ can be alternatively defined as

$$F^\Delta = \{f_0\}^\perp \cap (K^*)$$

where f_0 is an arbitrary element of $\text{ri } F$, the relative interior of the face F . Also $(F^\Delta)^\Delta = F$ holds for every face F of K ([6]).

It is not difficult to verify using a standard separation argument that if \mathcal{R}^d is the finite union of closed convex sets C_i with nonempty and pairwise disjoint interiors then the sets C_i are necessarily polyhedrons. In this case we call the set $\{C_i\}$ a *polyhedral partition* of \mathcal{R}^d .

Now we can state our main result,

THEOREM 1.1. *Let A be an m by n real matrix. Let K denote the finitely generated cone $\text{Im}_+ A$. Then the following statements hold:*

a) *The cones $\{F - F^\Delta : F \triangleleft K\}$ form a polyhedral partition of \mathcal{R}^m . The map p_K is linear restricted to the members of this partition, that is for every $F \triangleleft K$ there exists a unique matrix P_F such that*

$$p_K(f - g) = P_F(f - g) \quad (f \in F, g \in F^\Delta).$$

b) *For every $F \triangleleft K$, and every basis B of F , $P_F = P_{\text{Im } B}$. Specially, the matrices P_F ($F \triangleleft K$) are symmetric, idempotent and positive semidefinite.*

c) *The map P is a bijection between the sets $\{F : F \triangleleft K\}$ and $\{P_F : F \triangleleft K\}$; and it preserves the usual partial ordering on these sets, that is $F_1 \subseteq F_2$ if and only if $P_{F_2} - P_{F_1}$ is positive semidefinite.*

d) *For every face F of K ,*

$$P_F \cdot P_{F^\Delta}^* = 0, P_F + P_{F^\Delta}^* = I.$$

(Here $P_{F^\Delta}^*$ denotes the matrices defined by p_{K^*} .)

In Section 2 we will prove Theorem 1.1. In Section 3 we describe an algorithm for calculating the projection $p_{\text{Im}_+ A}(y)$ for a fixed vector $y \in \mathcal{R}^m$. The method is based on Lemke's algorithm to solve a linear complementarity problem LCP ([1]). After writing the problem as an LCP, using the structure of the problem our algorithm calculates with $r(A)$ by $2r(A)$ matrices instead of n by $2n$ matrices ($r(A)$ denotes rank). Finally, in Section 4 we describe a concept closely related to the projection $p_{\text{Im}_+ A}$: the conical inverse $A^<$. Theoretical and algorithmical properties of the conical inverse are largely unexplored and need further research.

2 Proof of the main theorem

In this section we will prove Theorem 1.1. First we state three lemmas and propositions that will be needed in the proof of statement a) in Theorem 1.1.

The first lemma describes a well-known characterization of the projection of a vector onto a closed convex cone, now specialized to the case of a finitely generated cone (see [2], Proposition 3.2.3).

LEMMA 2.1. *For every vector $y \in \mathcal{R}^m$ there exists a unique vector $k \in \mathcal{R}^m$ such that*

$$k \in K, k - y \in K^*, k^T(k - y) = 0. \quad (1)$$

This vector k equals $p_K(y)$ then.

As an immediate consequence, we obtain

PROPOSITION 2.1. *Let F be a face of K . Let C_F denote the set of vectors y such that $p_K(y) \in \text{ri } F$. Then*

$$C_F = (\text{ri } F) - F^\Delta. \quad (2)$$

Proof. Let C denote the set on the right hand side of (2). First we will show that $C_F \subseteq C$. Let y be an element of C_F , and let k denote the vector $p_K(y)$. Then $k \in \text{ri } F$; and, by (1), $k - y \in \{k\}^\perp \cap K^*$, that is $k - y \in F^\Delta$. We can see that $y = k - (k - y)$ is an element of C , and the inclusion $C_F \subseteq C$ is proved.

Conversely, if $k \in \text{ri } F$ and $k - y \in F^\Delta$, then (1) holds, so $k = p_K(y)$, and we can see that $y \in C_F$. This way we have proved the inclusion $C \subseteq C_F$ as well. \square

The closure of the set C_F defined in Proposition 2.1 is $\text{cl } C_F = F - F^\Delta$. The next lemma states that this finitely generated cone is full-dimensional (or equivalently has nonempty interior).

LEMMA 2.2. *The linear hull of the set $F - F^\Delta$ is \mathcal{R}^m , for every face F of K .*

Proof. Let B be a matrix such that $K = \text{Ker}_+(B^T)$. It is well-known (see [5]) that then there exists a partition (B_1, B_2) of the columns of B such that

$$\begin{aligned} F &= \{y : B_1^T y \geq 0, B_2^T y = 0\}, \\ \text{lin } F &= \{y : B_2^T y = 0\}, \\ \text{ri } F &= \{y : B_1^T y > 0, B_2^T y = 0\}. \end{aligned}$$

(Here lin denotes linear hull.) Let $f_0 \in \text{ri } F$. Then $F^\Delta = \{f_0\}^\perp \cap K^*$, and the latter set can be easily seen to be equal to $\text{Im}_+ B_2$. Thus the linear hull of F^Δ is $\text{Im } B_2$, the orthogonal complement of the linear hull of F . The linear hull of the set $F - F^\Delta$, being the sum of these two subspaces, is \mathcal{R}^m . \square

It is well-known that relative interiors of the faces of a convex set form a partition of the convex set (see [3]): they are pairwise disjoint, and their union

is the whole convex set. From this observation easily follows that the sets C_F are pairwise disjoint, and their union is the whole space. Consequently their closures, the sets $\text{cl } C_F$ ($F \triangleleft K$), have pairwise disjoint interiors (as the interior of $\text{cl } C_F$ equals the interior of the convex C_F), cover the whole space, and (by Lemma 2.2) are full-dimensional. We obtained a proof of

PROPOSITION 2.2. *The sets $F - F^\Delta$ ($F \triangleleft K$) form a polyhedral partition of \mathcal{R}^m . \square*

We call a set $C \subseteq \mathcal{R}^m$

- *positively homogeneous* if $0 < \lambda \in \mathcal{R}$, $y \in C$ implies $\lambda y \in C$;
- *additive* if $y_1, y_2 \in C$ implies $y_1 + y_2 \in C$.

Similarly, we call a map $p : C \rightarrow \mathcal{R}^m$, defined on a positively homogeneous, additive set C ,

- *positively homogeneous* if $0 < \lambda \in \mathcal{R}$, $y \in C$ implies $p(\lambda y) = \lambda p(y)$;
- *additive* if $y_1, y_2 \in C$ implies $p(y_1 + y_2) = p(y_1) + p(y_2)$.

For example,

PROPOSITION 2.3. *The sets C_F defined in Proposition 2.1 are positively homogeneous, additive sets; and the map p_K restricted to C_F is a positively homogeneous, additive map.*

Proof. The first half of the statement follows from Proposition 2.1: the sets $(\text{ri } F) - F^\Delta$ are obviously positively homogeneous, additive sets.

To prove the remaining part of the statement, let $y \in C_F$. Then by Proposition 2.1 there exist $f_0 \in \text{ri } F$, $g \in F^\Delta$ such that $y = f_0 - g$. Also, for $0 < \lambda \in \mathcal{R}$, $\lambda y \in C_F$. Again, by Proposition 2.1 there exist $f_0(\lambda) \in \text{ri } F$, $g(\lambda) \in F^\Delta$ such that $\lambda y = f_0(\lambda) - g(\lambda)$. Necessarily,

$$\lambda p_K(y) = \lambda f_0 = f_0(\lambda) = p_K(\lambda y),$$

and we have proved the positive homogeneity of the map p_K restricted to the set C_F . Additivity can be similarly verified, so the proof is finished. \square

We can see that the sets C_F are full-dimensional, positively homogeneous, additive sets, and the map p_K restricted to the set C_F is a positively homogeneous, additive map. Such maps have a unique linear extension as the following lemma states.

LEMMA 2.3. *Let C be a positively homogeneous, additive set in \mathcal{R}^m such that the linear hull of the set C is the whole space \mathcal{R}^m . Let $p : C \rightarrow \mathcal{R}^m$ be a positively homogeneous, additive map. Then there exists a unique linear map $\ell : \mathcal{R}^m \rightarrow \mathcal{R}^m$ such that $\ell(y) = p(y)$ for every $y \in C$.*

Proof. Let us choose a basis $\{y_1, \dots, y_m\}$ from the set C , and let us define the map ℓ as follows:

$$\ell \left(\sum_{i=1}^m \lambda_i y_i \right) := \sum_{i=1}^m \lambda_i p(y_i) \quad (\lambda_1, \dots, \lambda_m \in \mathcal{R}).$$

We will show that the restriction of this linear map ℓ to the set C is the map p . Let C_0 denote the set of the linear combinations of the vectors y_1, \dots, y_m with positive coefficients. Then C_0 is an open set, and $p(y) = \ell(y)$ for every $y \in C_0$. Fix $y_0 \in C_0$, and let y be an arbitrary element from the set C . Then there exists a constant $0 < \varepsilon < 1$ such that the vector $y_\varepsilon := \varepsilon y + (1 - \varepsilon)y_0$ is in the set C_0 . By positive homogeneity and additivity of the map p ,

$$p(y_\varepsilon) = \varepsilon p(y) + (1 - \varepsilon)p(y_0).$$

On the other hand, by linearity of the map ℓ ,

$$\ell(y_\varepsilon) = \varepsilon \ell(y) + (1 - \varepsilon)\ell(y_0).$$

Here $\ell(y_\varepsilon) = p(y_\varepsilon)$ and $\ell(y_0) = p(y_0)$, so we have $\ell(y) = p(y)$; the map ℓ meets the requirements.

Finally, uniqueness of the map ℓ is trivial, as ℓ must have fixed values for the elements of the full-dimensional set C . \square

Now we can describe the proof of Theorem 1.1.

Proof of part a) in Theorem 1.1: By Proposition 2.3 and Lemma 2.3 existence and uniqueness of matrices P_F follow such that

$$p_K(y) = P_F y \quad (y \in C_F).$$

It is well-known (see Proposition 3.1.3 in [2]), that the map p_K is continuous, so we have actually

$$p_K(y) = P_F y \quad (y \in \text{cl } C_F).$$

We have seen already (see Proposition 2.2) that the sets $\text{cl } C_F = F - F^\Delta$ ($F \triangleleft K$) form a polyhedral partition of \mathcal{R}^m , thus the proof of statement a) in Theorem 1.1 is complete.

Proof of part b) in Theorem 1.1: Let F be a face of the cone K . Let B be a basis of the face F , and let B^Δ be a basis of the complementary face F^Δ . Then every vector $y \in \mathcal{R}^m$ can be written in the form

$$y = Bv + B^\Delta w, \quad v \in \mathcal{R}^{\dim F}, \quad w \in \mathcal{R}^{\dim F^\Delta}.$$

Multiplying this equality from the left with the matrices P_F and B^T , respectively, we obtain the equalities

$$P_F y = Bv, \quad B^T y = B^T Bv.$$

These equalities imply

$$P_F y = Bv = B(B^T B)^{-1} B^T y = P_{\text{Im } B} y.$$

We have $P_F = P_{\text{Im } B}$, and the proof of statement b) in Theorem 1.1 is finished also.

Proof of part c) in Theorem 1.1: First, notice that the map P is trivially a bijection between the sets $\{F : F \triangleleft K\}$ and $\{P_F : F \triangleleft K\}$. (Injectivity follows from the obvious fact that if $F_1 \neq F_2$, for example there exists $y \in F_1 \setminus F_2$, then $P_{F_1} = P_{F_2}$ would imply $P_{F_2} y = P_{F_1} y = y$, and thus $y \in F_2$, which is a contradiction.)

Hence we have to verify only that $F_1, F_2 \triangleleft K$, $F_1 \subseteq F_2$ implies that $P_{F_2} - P_{F_1}$ is positive semidefinite. Let B_1 be a basis of the face F_1 , and let B_2 be a basis of the face F_2 such that $B_1 \subseteq B_2$. Then for every $y \in \mathcal{R}^m$, by the definition of the projection map,

$$\|y - P_{\text{Im } B_1} y\|^2 \geq \|y - P_{\text{Im } B_2} y\|^2.$$

This inequality, by part b) in Theorem 1.1 imply that

$$y^T P_{F_2} y \geq y^T P_{F_1} y \quad (y \in \mathcal{R}^m),$$

that is the positive semidefiniteness of the matrix $P_{F_2} - P_{F_1}$, which was to be shown.

Proof of part d) in Theorem 1.1: Let y_1, \dots, y_m be a basis in the set C_F , and let $y_1^\Delta, \dots, y_m^\Delta$ be a basis in the set $C_{F^\Delta}^*$.

Then to prove that $P_F \cdot P_{F^\Delta}^* = 0$, it is enough to show that

$$y_i^T P_F \cdot P_{F^\Delta}^* y_j^\Delta = 0 \quad (1 \leq i, j \leq m).$$

In other words we have to show that the vectors $p_K(y_i)$ and $p_{K^*}(y_j^\Delta)$ are orthogonal. This follows from the fact that the former vectors are in F , while the latter vectors are in F^Δ .

To prove the equality $P_F + P_{F^\Delta}^* = I$, it is enough to verify that

$$y_i^T (P_F + P_{F^\Delta}^*) y_j^\Delta = y_i^T y_j^\Delta \quad (1 \leq i, j \leq m).$$

In other words that

$$p_K(y_i)^T y_j^\Delta + y_i^T p_{K^*}(y_j^\Delta) = y_i^T y_j^\Delta \quad (1 \leq i, j \leq m),$$

or equivalently that

$$(y_i - p_K(y_i))^T (y_j^\Delta - p_{K^*}(y_j^\Delta)) = 0 \quad (1 \leq i, j \leq m).$$

This latter equality is the consequence of the fact that the vectors $y_i - p_K(y_i)$ are in F^Δ while the vectors $y_j^\Delta - p_{K^*}(y_j^\Delta)$ are in $(F^\Delta)^\Delta = F$ and so are orthogonal.

This way we have finished the proof of part d), and the proof of Theorem 1.1 as well. \square

We conclude this section with an example. Let K be the finitely generated cone

$$K := \{(y_1, y_2) \in \mathcal{R}^2 : y_1 \geq y_2 \geq 0\}.$$

Then K has four faces,

$$F_1 = \{0\}, F_2 = \{(\lambda, \lambda) : \lambda \geq 0\}, F_3 = \{(\lambda, 0) : \lambda \geq 0\}, F_4 = K.$$

The dual of K is

$$K^* = \{(y_1, y_2) \in \mathcal{R}^2 : y_2 \geq -y_1 \leq 0\}$$

with faces

$$F_1^\Delta = K^*, F_2^\Delta = \{(\lambda, -\lambda) : \lambda \geq 0\}, F_3^\Delta = \{(0, \lambda) : \lambda \geq 0\}, F_4^\Delta = \{0\}.$$

The reader can easily verify that

$$P_{F_1} = 0, P_{F_2} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, P_{F_3} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, P_{F_4} = I;$$

$$P_{F_1^\Delta}^* = I, P_{F_2^\Delta}^* = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, P_{F_3^\Delta}^* = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, P_{F_4^\Delta}^* = 0$$

and that the statements of Theorem 1.1 hold in this special case.

3 Algorithm for computing the projection

In this section we will describe an algorithm for calculating the projection of a fixed vector onto a finitely generated cone. The algorithm economically solves a certain type of linear complementarity problems as well.

Let A be an m by n real matrix, and let K denote the finitely generated cone $\text{Im}_+ A$. Let us fix a vector $y \in \mathcal{R}^m$. To compute the projection $p_K(y)$, by Lemma 2.1, we have to find a vector $x \in \mathcal{R}^n$ such that

$$x \geq 0, Ax - y \in \text{Ker}_+(A^T), (Ax)^T(Ax - y) = 0.$$

This problem can be rewritten as a linear complementarity problem $LCP(A)$: find vectors $z, x \in \mathcal{R}^n$ such that

$$z - A^T Ax = -A^T y; x, z \geq 0; z^T x = 0.$$

A finite version of Lemke's algorithm ([1]) can be applied to solve $LCP(A)$; if (x, z) is a solution, then $Ax = p_K(y)$ is the projection we searched for.

However, a more economical algorithm can be constructed to find the projection of a vector; economical in the sense that instead of solving the n -dimensional $LCP(A)$, it solves a sequence of $r(A)$ -dimensional LCPs, $LCP(B_1)$, $LCP(B_2), \dots$ where B_1, B_2, \dots are bases of the matrix A . (A matrix B is called a *basis* of the matrix A , if B is an m by $r(A)$ submatrix of A , and $r(B) = r(A)$.)

Before describing this algorithm, we prove three propositions and lemmas that will be needed in the proof of the correctness and finiteness of the algorithm.

Let B be a basis of the matrix A , with corresponding basis tableau $T(B)$. (The *basis tableau* $T(B)$ corresponding to a basis B contains the unique coefficients $t_{ij} \in \mathcal{R}$ such that

$$a_j = \sum_i \{t_{ij}a_i : a_i \in B\} \quad (a_j \in A),$$

where a_j denotes the j -th column vector of the matrix A .)

With these notations,

PROPOSITION 3.1. *If (x, z) is a solution of $LCP(B)$, then $Bx = p_{\text{Im}_+ B}(y)$. Furthermore, if*

$$a_j^T(Bx - y) \geq 0 \quad (a_j \in A \setminus B) \quad (3)$$

holds, then $Bx = p_{\text{Im}_+ A}(y)$ as well.

Proof. The statement follows from Lemma 2.1, see also the beginning of this section. \square

If (3) does not hold, then there exist column vectors $a_j \in A \setminus B$ such that $a_j^T(Bx - y) < 0$. Choose one of them which minimizes the inner product with the vector $Bx - y$: let j_* be an index such that

$$a_{j_*}^T(Bx - y) = \min \{a_j^T(Bx - y) : a_j \in A \setminus B\}, \quad a_{j_*} \in A \setminus B. \quad (4)$$

This vector a_{j_*} will enter the basis B .

From the definition of the index j_* immediately follows

LEMMA 3.1. *The minimum in (4) is less than 0. \square*

Now, we will choose the vector a_{i^*} that leaves the basis B . Let i^* denote an index such that

$$a_{i^*}^T(Bx - y) = \max \{a_i^T(Bx - y) : a_i \in B, t_{ij_*} \neq 0\}, \quad a_{i^*} \in B, \quad t_{i^*j_*} \neq 0. \quad (5)$$

Remember that $Bx = p_{\text{Im}_+ B}(y)$; so $a_i^T(Bx - y) \geq 0$ holds for every vector $a_i \in B$. Hence the maximum in (5) is at least 0. But more can be said:

LEMMA 3.2. *The maximum in (5) is greater than 0.*

Proof. The vector a_{j_*} can be written in the form

$$a_{j_*} = \sum \{t_{ij_*} a_i : a_i \in B\}.$$

If $a_i^T(Bx - y) = 0$ would hold for every index i such that $a_i \in B$, $t_{ij_*} \neq 0$, then $a_{j_*}^T(Bx - y) = 0$ would follow, contradicting Lemma 3.1. \square

It is well-known (see [7]) that $t_{i^*j_*} \neq 0$ implies

PROPOSITION 3.2. *The submatrix*

$$\hat{B} := (B \setminus \{a_{i^*}\}) \cup \{a_{j_*}\}$$

is a basis of the matrix A . Furthermore, the corresponding basis tableau $T(\hat{B})$ can be obtained from the basis tableau $T(B)$ by pivoting on the (i^, j_*) -th position of the latter tableau.*

We will show that the new basis \hat{B} is an improvement over B : the distance between $p_{\text{Im}_+ \hat{B}}(y)$ and y is less than the distance between $p_{\text{Im}_+ B}(y)$ and y . To verify this statement we need the following lemma:

LEMMA 3.3. *The vectors $p_{\text{Im}_+ B}(y)$ and a_{j_*} are elements of the cone $\text{Im}_+ \hat{B}$.*

Proof. As for the solution (x, z) of $LCP(B)$, $z^T x = 0$, $z, x \geq 0$, necessarily $x_i z_i = 0$ for all index i . As by Lemma 3.2 $z_{i^*} = a_{i^*}^T(Bx - y) > 0$, we have $x_{i^*} = 0$. We can see that $Bx \in \text{Im}_+(B \setminus \{a_{i^*}\})$; consequently $Bx \in \text{Im}_+ \hat{B}$, that is $p_{\text{Im}_+ B}(y) \in \text{Im}_+ \hat{B}$. The statement that $a_{j_*} \in \text{Im}_+ \hat{B}$ is trivial. \square

PROPOSITION 3.3. *It holds that*

$$\|p_{\text{Im}_+ \hat{B}}(y) - y\| < \|p_{\text{Im}_+ B}(y) - y\|.$$

Proof. Let S denote the following set of vectors:

$$S := \{\varepsilon Bx + (1 - \varepsilon)a_{j_*}; 0 < \varepsilon < 1, \varepsilon \in \mathcal{R}\}.$$

By Lemma 3.3, $S \subseteq \text{Im}_+ \hat{B}$. Furthermore, from Lemma 3.1 it can be easily seen that there exists a vector $s \in S$ such that the distance between y and s is less than the distance between y and $Bx = p_{\text{Im}_+ B}(y)$. As the distance between y and $p_{\text{Im}_+ \hat{B}}(y)$ is not greater than the distance between y and s , the statement follows. \square

Now we can describe the algorithm and prove its correctness.

Algorithm: Let B be a given basis of the matrix A , with corresponding basis tableau $T(B)$. (Initially, this basis can be any basis of the matrix A .)

Solve $LCP(B)$ using a finite version of Lemke's algorithm. This way a solution (x, z) is obtained, $p_{\text{Im}_+ B}(y) = Bx$.

If the inequality $a_j^T(Bx - y) \geq 0$ holds for every j such that $a_j \in A \setminus B$, then stop: $p_{\text{Im}_+ A}(y) = Bx$.

If there exist at least one index j such that $a_j \in A \setminus B$ and $a_j^T(Bx - y) < 0$, then choose j_* according to (4) and choose i^* according to (5).

Let $\hat{B} := (B \setminus \{a_{i^*}\}) \cup \{a_{j_*}\}$ be the new basis, update $T(B)$ by pivoting on its (i^*, j_*) -th position, and repeat the above step until $p_K(y)$ is reached.

THEOREM 3.1. *The algorithm described above finds $p_K(y)$ after finite number of steps.*

Proof. The correctness of the algorithm follows from Propositions 3.1 and 3.2. The finiteness of the algorithm is a consequence of Proposition 3.3: as the distance between the vectors $p_{\text{Im}_+ B}(y)$ and y decreases with each step, so there can be no repetition in the sequence of the bases, and there are only finitely many of the bases of A . \square

Finally, we remark that the above algorithm can be applied to solve economically any LCP of the form

$$(LCP) : \text{ find } x, z \text{ such that } z - Mx = q; z, x \geq 0; z^T x = 0$$

where M is a symmetric positive semidefinite matrix, $q \in \text{Im } M$. To see this it is enough to note that a matrix A can be found such that $M = A^T A$ using Cholesky decomposition ([7]). Then $\text{Im } M = \text{Im}(A^T)$, so $q = -A^T y$ for some vector y . This way we have rewritten (LCP) as $LCP(A)$, and the algorithm discussed in this section can be applied to solve the problem.

4 Remarks on the conical inverse

In this section we describe a concept closely related to the projection onto a finitely generated cone and also to the Moore-Penrose generalized inverse of a matrix.

Let A be an m by n real matrix. For every vector $y \in \mathcal{R}^m$ there exists a unique vector $x_* \in \mathcal{R}^n$ such that

$$\|x_*\| = \min\{\|x\| : Ax = p_{\text{Im } A}(y), x \in \mathcal{R}^n\}.$$

Surprisingly, the dependence of the vector x_* on the vector y turns out to be linear ([7]): there exists a unique matrix, called the *Moore-Penrose generalized inverse* of the matrix A and denoted by A^- such that $x_* = A^- y$ for every $y \in \mathcal{R}^m$.

It is also well-known (see [4]) that the Moore-Penrose generalized inverse can be alternatively defined as the unique matrix $A^- \in \mathcal{R}^{n \times m}$ satisfying the

four conditions: a) $AA^-A = A$; b) $A^-AA^- = A^-$; c) $(AA^-)^T = AA^-$; d) $(A^-A)^T = A^-A$.

Similarly as in the case of the projection map, this concept can also be generalized via replacing the subspace $\text{Im } A$ with the finitely generated cone $\text{Im}_+ A$. The map defined this way is called the *conical inverse* of the matrix A , and is denoted by $A^<$. Thus for every vector $y \in \mathcal{R}^m$, $A^<(y)$ is the unique vector in \mathcal{R}^n satisfying

$$\|A^<(y)\| = \min\{\|x\| : Ax = p_{\text{Im}_+ A}(y), x \geq 0\}; A^<(y) \geq 0.$$

The next proposition describes a certain relation between the two inverses defined above.

PROPOSITION 4.1. *For every vector $y \in \mathcal{R}^m$, it holds that*

$$(A, -A)^<(y) = (\max\{A^-y, 0\}, -\min\{A^-y, 0\})$$

where the *max* and *min* are meant elementwise.

Proof. Let us consider the following two programs:

$$(\hat{P}) : \quad \min_{x_1, x_2 \geq 0} \|y - Ax_1 + Ax_2\|$$

and

$$(P) : \quad \min_x \|y - Ax\|.$$

The variable transformations

$$x := x_1 - x_2 \text{ resp. } x_1 := \max\{x, 0\}, x_2 := -\min\{x, 0\}$$

show the equivalence of programs (\hat{P}) and (P) .

Furthermore, it can be easily seen that

- if x is an optimal solution of program (P) then the vector

$$(x_1, x_2) := (\max\{x, 0\} + p, -\min\{x, 0\} + p)$$

is an optimal solution of program (\hat{P}) for any vector $p \geq 0$;

- if (x_1, x_2) is an optimal solution of program (\hat{P}) then the vector $x := x_1 - x_2$ is an optimal solution of program (P) such that

$$(x_1, x_2) = (\max\{x, 0\} + p, -\min\{x, 0\} + p)$$

for some vector $p \geq 0$.

Consequently, the shortest optimal solution of (\hat{P}) will be

$$(x_1, x_2) = (\max\{x, 0\}, -\min\{x, 0\})$$

where x is the shortest optimal solution of (P) ; which was to be shown. \square

Thus any algorithm for calculating the conical inverse can be used for calculating the Moore-Penrose generalized inverse. Conversely also, as the following proposition shows.

PROPOSITION 4.2. *The vector x equals $A^{\langle}(y)$ if and only if for some vector z , (x, z) is a solution of the following linear complementarity problem:*

$$(LCP^{\langle}) : \quad \begin{cases} \text{find } x, z \text{ such that } x, z \geq 0; z^T x = 0; \\ x - (I - A^- A)z = A^- p_{\text{Im}^+ A}(y). \end{cases}$$

Proof. Let $x_0 \geq 0$ be a vector such that $Ax_0 = p_{\text{Im}^+ A}(y)$. To find $A^{\langle}(y)$ we have to find the unique optimal solution of the following convex quadratic program:

$$(QP) : \quad \min \frac{1}{2} \|x\|^2, Ax = Ax_0, x \geq 0.$$

By the Kuhn-Tucker conditions (see [1]), the vector $x \geq 0$ is an optimal solution of program (QP) if and only if there exists a vector $z \geq 0$ such that $z^T x = 0$,

$$x - z \in \text{Im}(A^T), x - x_0 \in \text{Ker } A. \quad (6)$$

It is well-known that

$$P_{\text{Im}(A^T)} = A^- A, P_{\text{Ker } A} = I - A^- A,$$

so (6) can be rewritten as

$$(I - A^- A)(x - z) = 0, A^- Ax = A^- Ax_0. \quad (7)$$

It is easy to see that (7) holds if and only if (x, z) satisfies the following equality

$$x - (I - A^- A)z = A^- Ax_0.$$

We can see that $x = A^{\langle}(y)$ if and only if there exists a vector z such that (x, z) is a solution of (LCP^{\langle}) ; the proof is complete. \square

Finally, we mention two open problems concerning the conical inverse:

- Is it true, that similarly to the case of the projection, the conical inverse is also made up from linear parts? (This statement is trivial if the m by n matrix A has rank $r(A) = n$.)

- We can see from Proposition 4.2 that the conical inverse for a fixed vector y can be calculated via solving an n -dimensional LCP. Is it possible to construct an algorithm to compute $A^{\leftarrow}(y)$ more economically, similarly as in the case of the projection map?

Conclusion. In this paper we examined the properties of the projection onto a finitely generated cone. Our main result shows that this map is made up of linear parts (the map is linear if and only if we project onto a subspace) with a structure resembling the facial structure of the finitely generated cone we project onto. Also we presented an algorithm for computing the projection of a fixed vector. The algorithm is economical in the sense that it calculates with matrices whose size depends on the dimension of the finitely generated cone and not on the number of the generating vectors of the cone. Some remarks and open problems concerning the conical inverse conclude the paper.

Acknowledgements. I am indebted to Tamás Rapcsák and András Frank for the several consultations.

References

1. BAZARAA, M. S. AND SHETTY, C. M., *Nonlinear programming, theory and algorithms*, John Wiley & Sons, New York, 1979.
2. HIRIART-URRUTY, J. AND LEMARÉCHAL, C., *Convex analysis and minimization algorithms I.*, Springer-Verlag, Berlin, 1993.
3. ROCKAFELLAR, R. T., *Convex analysis*, Princeton University Press, Princeton, 1970.
4. RÓZSA, P., *Linear algebra and its applications*, Tankönyvkiadó, Budapest, 1991 (in Hungarian).
5. SCHRIJVER, A., *Theory of linear and integer programming*, John Wiley & Sons, New York, 1986.
6. STOER, J. AND WITZGALL, C., *Convexity and optimization in finite dimensions I.*, Springer-Verlag, Berlin, 1970.
7. STRANG, G., *Linear algebra and its applications*, Academic Press, New York, 1980.

8. UJVÁRI, M., *Applications of semidefinite programming in combinatorial optimization*, ELTE Eötvös Kiadó, 2001 (in Hungarian).
9. UJVÁRI, M., *Convex analysis*, Unpublished manuscript, 2007 (in Hungarian). URL: <http://www.oplab.sztaki.hu/tanszek/letoltes.htm>