Rapid heuristic projection on simplicial cones *

A. Ekárt
Computer Science, Aston University
Aston Triangle, Birmingham B4 7ET
United Kingdom
email: a.ekart@aston.ac.uk

A. B. Németh
Faculty of Mathematics and Computer Science
Babeş Bolyai University, Str. Kogălniceanu nr. 1-3
RO-400084 Cluj-Napoca, Romania
email: nemab@math.ubbcluj.ro

S. Z. Németh
School of Mathematics, The University of Birmingham
The Watson Building, Edgbaston
Birmingham B15 2TT, United Kingdom
email: nemeths@for.mat.bham.ac.uk

January 19, 2010

Abstract

A very fast heuristic iterative method of projection on simplicial cones is presented. It consists in solving two linear systems at each step of the iteration. The extensive experiments indicate that the method furnishes the exact solution in more then 99.7 percent of the cases. The average number of steps is 5.67 (we have not found any examples which required more than 13 steps) and the relative number of steps with respect to the dimension decreases dramatically. Roughly speaking, for high enough dimensions the absolute number of steps is independent of the dimension.

1 Introduction

Projection on polyhedral cones is one of the important problems applied optimization is confronted with. Many applied numerical optimization methods uses projection on polyhedral cones as the main tool.

In most of them, projection is part of an iterative process which involve its repeated application (see e. g. problems of image reconstruction [1], nonlinear complementarity [4, 9], etc.). Hence, it is important to get fast projection algorithms.

* *1991 A M S Subject Classification.* Primary 90C33; Secondary 15A48, *Key words and phrases.* Metric projection on simplicial cones

The main streems of the current methods in use rely on the classical von Neumann algorithm (see e.g. the Dykstra algorithm [3, 2, 10]), but they are rather expensive for a numerical handling (see the numerical results in [7] and the remark preceding section 6.3 in [5]).

Finite methods of projections are of combinatorial nature which reduces their applicability to low dimensional ambient spaces.

Recently we have given a simple projection method exposed in note [8] for projecting on so called isotone projection cones. Isotone projection cones are special simplicial cones, and due to their good properties we can project on them in n steps, where n is the dimension of the ambient space. In the first part of that note we have explained our approach by considering the problem of projection on simplicial cones by giving an exact method based on duality. This method has combinatorial character and therefore it is inefficient. More recently we observed that a heuristic method based on the same ideas gives surprisingly good results. This note describes the theoretical foundation of the heuristic method and draws conclusions based on millions of numerical experiments.

Projection on polyhedral cones is a problem of high impact on scientific community.[1]

2 The simplicial cone and its polar

Let $\mathbb{R}^n$ be an n-dimensional Euclidean space endowed with a Cartesian reference system. We assume that each point of $\mathbb{R}^n$ is a column vector.

We shall use the term cone in the sense of closed convex cone. That is, the nonempty closed subset $K \subset \mathbb{R}^n$ in our terminology is a *cone*, if $K + K \subset K$, and $tK \subset K$ whenever $t \in \mathbb{R}$, $t \geq 0$.

Let $m \leq n$ and $\mathbf{e}_1, \ldots, \mathbf{e}_m$ be m elements in $\mathbb{R}^n$. Denote

$$\text{cone}\{\mathbf{e}_1, \ldots, \mathbf{e}_m\} = \{\lambda^1 \mathbf{e}_1 + \cdots + \lambda^m \mathbf{e}_m : \lambda^i \geq 0,\, i = 1, \ldots, m\},$$

the *cone engendered by* $\mathbf{e}_1, \ldots, \mathbf{e}_m$. Then,

$$\text{cone}\{\mathbf{e}_1, \ldots, \mathbf{e}_m\} = \{\mathbf{E}\mathbf{v} : \mathbf{v} \in \mathbb{R}^m_+\}, \tag{1}$$

where $\mathbf{E} = (\mathbf{e}_1, \ldots, \mathbf{e}_m)$ is the matrix with columns $\mathbf{e}_1, \ldots, \mathbf{e}_m$ and $\mathbb{R}^m_+$ is the non-negative orthant in $\mathbb{R}^m$.

Suppose that $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ are linearly independent elements. Then, the cone

$$\begin{array}{rcl} K & = & \text{cone}\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \\ & = & \{\lambda^1 \mathbf{e}_1 + \cdots + \lambda^n \mathbf{e}_n : \lambda^i \geq 0,\, i = 1, \ldots, n\} = \{\mathbf{E}\mathbf{v} : \mathbf{v} \in \mathbb{R}^m_+\}, \end{array} \tag{2}$$

with $\mathbf{E}$ the matrix from (1) for $m = n$, is called *simplicial cone*. Denote $N = \{1, 2, \ldots, n\}$.

The *polar* of K is the set

$$K^\circ = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^\top \mathbf{y} \leq 0,\, \forall \mathbf{y} \in K\}. \tag{3}$$

$K^* = -K^\circ$ is called the *dual* of K. K is called *subdual*, if $K \subset K^*$. This is equivalent to the condition $\mathbf{e}_\ell^\top \mathbf{e}_k \geq 0$, $\ell, k \in N$.

[1] see the popularity of the Wikimization page *Projection on Polyhedral Cone* at http://www.convexoptimization.com/wikimization/index.php/Special:Popularpages.

Lemma 1 *The polar of the simplicial cone (2) can be represented in the form*

$$K^\circ = \{\mu^1 \mathbf{u}_1 + \dots \mu^n \mathbf{u}_n : \mu^i \geq 0,\, i = 1, \dots, n\}, \tag{4}$$

where $\mathbf{u}_i (i = 1, \dots, n)$ *is a solution of the system*

$$\mathbf{e}_j^\top \mathbf{u}_i = 0,\ j = 1, \dots, n,\ j \neq i,$$

$$\mathbf{e}_i^\top \mathbf{u}_i = -1$$

($\mathbf{u}_i$ *is normal to the hyperplane* $\operatorname{span}\{\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_n\}$ *in the opposite direction to the halfspace that contains* $\mathbf{e}_i$*). Thus,*

$$K^\circ = \{\mathbf{U}\mathbf{x} : \mathbf{x} \in \mathbb{R}^n_+\}$$

with $\mathbb{R}^n_+ = \{\mathbf{x} = (x_1, \dots, x_n)^\top : x_i \geq 0,\ i = 1 \dots n\}$ *and*

$$\mathbf{U} = -(\mathbf{E}^{-1})^\top. \tag{5}$$

For simplicity we shall call $\mathbf{U}$ *the* polar matrix *of* $\mathbf{E}$*. The columns of* $\mathbf{U}$ *are* $\{\mathbf{u}_i \colon i = 1, \dots, n\}$.

Proof. Let $\mathbf{y} = \sum_{j=1}^n \mu^j \mathbf{u}_j$ and $\mathbf{z} = \sum_{i=1}^n \alpha^i \mathbf{e}_i$ for any non-negative real numbers α^i and μ^j. The inner produce of $\mathbf{y}$ and $\mathbf{z}$ is non-positive because

$$\mathbf{y}^\top \mathbf{z} = \sum_{i=1}^n \sum_{j=1}^n \alpha^i \mu^j \mathbf{e}_i^\top \mathbf{u}_j = -\sum_{i=1}^n \alpha^i \mu^i \leq 0$$

But $\mathbf{y}$ is an arbitrary element of the right hand side of (4) and $\mathbf{z}$ is an arbitrary element of K, thus we can conclude that the righ hand side of (4) is a subset of K°.

The vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ are linearly independent. (This can be verified by assuming the contrary, and by multiplying the subsequent relation by $\mathbf{e}_j$ to get a contradiction.) Hence for $\mathbf{y} \in K^\circ$ we have the representation

$$y = \beta^1 \mathbf{u}_1 + \dots + \beta^n \mathbf{u}_n.$$

By (3),

$$\mathbf{e}_k^\top \mathbf{y} = -\beta^k \leq 0$$

so $\beta^k \geq 0$ which prove that $\mathbf{y}$ is an element of the right hand side of (4). Thus we can conclude that K° is a subset of the right hand side of (4). □

The formula (5) of Lemma 1 is equivalent to the formula (380) of [1].

Corollary 1 *For each subset* I *of indices in* N*, the vectors* $\mathbf{e}_i, i \in I, \mathbf{u}_j, j \in I^c$ *(where* I^c *the complement of* I *with respect to* N*) are linearly independent.*

Proof. Assume that

$$\sum_{i \in I} \alpha^i \mathbf{e}_i + \sum_{j \in I^c} \beta^j \mathbf{u}_j = 0 \tag{6}$$

for some reals α^i and β^j. By the mutual orthogonality of the vectors $\mathbf{e}_i, i \in I$ and $\mathbf{u}_j, j \in I^c$ it follows, by multiplication of the relation (6) with $\sum_{i\in I} \alpha^i \mathbf{e}_i^\top$ and respectively with $\sum_{j\in I^c} \beta^j \mathbf{u}_j^\top$, that

$$\sum_{i\in I} \alpha^i \mathbf{e}_i = 0$$

and

$$\sum_{j\in I^c} \beta^j \mathbf{u}_j = 0.$$

Hence, $\alpha^i = \beta^j = 0$ must hold. □

The cone $K_0 \subset K$ is called a *face* of K if from $\mathbf{x} \in K_0, \mathbf{y} \in K$ and $\mathbf{x}-\mathbf{y} \in K$ it follows that $\mathbf{y} \in K_0$. The face K_0 is called a *proper face* of K, if $K_0 \neq K$.

Lemma 2 *If K is the cone (2) and K° is the cone (4), then for every subset of indices $I = \{i_1, ..., i_k\} \subset N$ the set*

$$F_I = \text{cone}\{\mathbf{e}_i : \ i \in I\} = \{\mathbf{x} \in K : \ \mathbf{x}^\top \mathbf{u}_j = 0 : \ i \in I^c\} \tag{7}$$

(with $F_I = \{\mathbf{0}\}$ if $I = \emptyset$) is a face of K. If $i_h \neq i_l$ whenever $h \neq l$, then F_I is for $k > 0$ a nonempty set in $\mathbb{R}^n$ of dimension k. (In the sense that F_I spans a subspace of $\mathbb{R}^n$ of dimension k.)

Every face of K is equal to F_I for some $I \subset N$. If $I \neq N$ then F_I is a proper face.

Proof.

The relation in (7) follows from the definition of the vectors $\mathbf{u}_j$ in Lemma 1, while the assertion on the dimension of F_I is obvious.

Suppose that $\mathbf{x} \in F_I$ and $\mathbf{y} \in K$ with $\mathbf{y} \leq \mathbf{x}$.

Then $(\mathbf{x}-\mathbf{y})^\top \mathbf{u}_j = -\mathbf{y}^\top \mathbf{u}_j \leq 0$, $\forall j \in I^c$ and $\mathbf{y}^\top \mathbf{u}_j \leq 0$, $\forall j \in N$, because $\mathbf{y} \in K$. Thus $\mathbf{y}^\top \mathbf{u}_j = 0$, $\forall j \in I^c$, hence $y \in F_I$, showing that F_I is a face.

Suppose that $\mathbf{x} \in F$ for F an arbitrary proper face of K. Since $\mathbf{x} \in K$, by the definition of the vectors $\mathbf{u}_j$, $\mathbf{x}^\top \mathbf{u}_j \leq 0$ for $j \in N$.

If $\mathbf{x}^\top \mathbf{u}_j < 0$, $\forall j \in N$, then there exists a positive scalar t with $(\mathbf{x}-t\mathbf{y})^\top \mathbf{u}_j \leq 0$, $\forall j \in N$. Hence, $\mathbf{x} - t\mathbf{y} \in K$ and thus $t\mathbf{y} \leq \mathbf{x}$. But then $t\mathbf{y} \in F$ and since F is a cone, $\mathbf{y} \in F$. This means that $K \subset F$, that is, F cannot be a proper face.

We have to show that F has a representation like (7). By the above reasoning, for each $\mathbf{x} \in F$ there exist some index $i \in N$ with with $\mathbf{x}^\top \mathbf{u}_i = 0$.

If $F = \{\mathbf{0}\}$ we have the representation (7) with $I = \emptyset$.

If $F \neq \{\mathbf{0}\}$, take $\mathbf{x}$ in the relative interior of F and let I be the complement in N of the maximal set of indices j with $\mathbf{x}^\top \mathbf{u}_j = 0$. ($I$ must be a nonempty, proper subset of N since $\mathbf{x} \neq \mathbf{0}$.)

Take $\mathbf{y} \in F$ arbitrarily. By the definition of $\mathbf{x}$, $\mathbf{x}-t\mathbf{y} \in F$ for some sufficiently small $t > 0$. Hence,

$$(\mathbf{x} - t\mathbf{y})^\top \mathbf{u}_i \leq 0, \ \ \forall i \in N. \tag{8}$$

By $\mathbf{y} \in F \subset K$ we also have $\mathbf{y}^\top \mathbf{u}_i \leq 0$, $\forall i \in N$. If $\mathbf{y}^\top \mathbf{u}_j < 0$ for some $j \in I^c$, then (8) would imply

$$\mathbf{x}^\top \mathbf{u}_j \leq t\mathbf{y}^\top \mathbf{u}_j < 0,$$

which is a contradiction. Hence, we must have $\mathbf{y}^\top \mathbf{u}_j = 0,\ \forall j \in I^c$; and accordingly

$$F \subset \{\mathbf{z} \in K :\ \mathbf{z}^\top \mathbf{u}_j = 0,\ \forall j \in I^c\}. \tag{9}$$

Suppose that $\mathbf{y} \in K$ and $\mathbf{y}^\top \mathbf{u}_j = 0,\ \forall j \in I^c$. From definition we have $\mathbf{x}^\top \mathbf{u}_i < 0$ for each $i \in I$, whereby for a sufficiently large $t > 0$,

$$(t\mathbf{x} - \mathbf{y})^\top \mathbf{u}_i \leq 0,\ \forall i \in N.$$

Hence, $t\mathbf{x}-\mathbf{y}$ is in the polar of K°, which by Farkas' lemma is K (This follows in fact, in our case, also by the symmetry of the vectors $\mathbf{e}_i$ and $\mathbf{u}_j$ in the formulae of Lemma 1.) Thus $\mathbf{0} \leq \mathbf{y} \leq t\mathbf{x}$, whereby $\mathbf{0} \leq (1/t)\mathbf{y} \leq \mathbf{x}$. Since F is a face of K, we have $(1/t)\mathbf{y} \in F$ and since it is also a cone, $\mathbf{y} \in F$. This proves the converse of the inclusion in (9) and completes the proof.

□

Thus a maximal proper face of K is of the form

$$K_{i_0} = \text{cone}\{\mathbf{e}_i : i \in N \setminus \{i_0\}\} = \text{cone}\{\mathbf{e}_i : i \in N, \mathbf{e}_i^\top \mathbf{u}_{i_0} = 0\},$$

hence it is also called *the face of K orthogonal to* $\mathbf{u}_{i_0}$. Similarly, we have a maximal proper face of K° orthogonal to some $\mathbf{e}_{j_0}$.

An equivalent result to the one presented in Lemma 2 is given by the Cone Table 1 on page 179 of [1].

Thus a maximal proper face of K is of the form

$$K_{i_0} = \text{cone}\{\mathbf{e}_i : i \in N \setminus \{i_0\}\} = \text{cone}\{\mathbf{e}_i : i \in N, \mathbf{e}_i^\top \mathbf{u}_{i_0} = 0\},$$

hence it is also called *the face of K orthogonal to* $\mathbf{u}_{i_0}$. Similarly, we have a maximal proper face of K° orthogonal to some $\mathbf{e}_{j_0}$.

Let $F = \text{cone}\{\mathbf{e}_i :\ i \in I\}$ and $F^\perp = \text{cone}\{\mathbf{u}_j :\ j \in I^c\}$. Then, from the above results it follows that

$$F = \{\mathbf{x} \in K :\ \mathbf{x}^\top \mathbf{u}_j = 0,\ j \in I^c\}$$

and

$$F^\perp = \{\mathbf{y} \in K^\circ :\ \mathbf{y}^\top \mathbf{e}_i = 0,\ i \in I\}.$$

The faces $F \subset K$ and $F^\perp \subset K^\circ$ of the above form are called a *pair of orthogonal faces* where $F^\perp$ is called the *orthogonal face of* F and F is called the *orthogonal face of* $F^\perp$.

3 Finite method of projection on a simplicial cone

For an arbitrary $\mathbf{u} \in \mathbb{R}^n$ denote $\|\mathbf{u}\| = \sqrt{\mathbf{u}^\top \mathbf{u}}$. Let $K \in \mathbb{R}^n$ be an arbitrary cone and K° its polar, and $C \subset \mathbb{R}^n$ an arbitrary closed convex set. Recall that the *projection mapping* $\mathbf{P}_C : H \to H$ on C is well defined by $\mathbf{P}_C\mathbf{x} \in C$ and

$$\|\mathbf{x} - \mathbf{P}_C\mathbf{x}\| = \min\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in C\}.$$

Then, Moreau's decomposition theorem asserts:

Theorem 1 *(Moreau, [6]) For* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ *the following statements are equivalent:*

(i) $\mathbf{z} = \mathbf{x} + \mathbf{y}, \mathbf{x} \in K, \mathbf{y} \in K^\circ$ *and* $\mathbf{x}^\top \mathbf{y} = 0$.

(ii) $\mathbf{x} = \mathbf{P}_K \mathbf{z}$ *and* $\mathbf{y} = \mathbf{P}_{K^\circ} \mathbf{z}$.

Suppose now, that K is a simplicial cone in $\mathbb{R}^n$. We shall use the representation (2) for K and the representation (4) for K°. Hence,

$$\mathbf{e}_i^\top \mathbf{u}_j = -\delta_j^i, i, j = 1, \ldots, n$$

where δ_j^i the Kronecker symbol. As a direct implication of Moreau's decomposition theorem and the constructions in the preceding section we have:

Theorem 2 *Let* $\mathbf{x} \in \mathbb{R}^n$. *For each subset of indices* $I \subset N$, $\mathbf{x}$ *can be represented in the form*

$$\mathbf{x} = \sum_{i \in I} \alpha^i \mathbf{e}_i + \sum_{j \in I^c} \beta^j \mathbf{u}_j \tag{10}$$

with I^c *the complement of* I *with respect to* N, *and with* α^i *and* β^j *real numbers. Among the subsets* I *of indices, there exists exactly one (the cases* $I = \emptyset$ *and* $I = N$ *are not excluded) with the property that for the coefficients in (10) one has* $\beta^j > 0, j \in I^c$ *and* $\alpha^i \geq 0, i \in I$. *For this representation it holds that*

$$\mathbf{P}_K \mathbf{x} = \sum_{i \in I} \alpha^i \mathbf{e}_i, \quad \alpha^i \geq 0, \tag{11}$$

and

$$\mathbf{P}_{K^\circ} \mathbf{x} = \sum_{j \in I^c} \beta^j \mathbf{u}_j, \quad \beta^j > 0. \tag{12}$$

Proof. The first assertion is the consequence of Corollary 1.

The projections $\mathbf{P}_K \mathbf{x}$ and $\mathbf{P}_{K^\circ} \mathbf{x}$ as elements of K and K°, respectively can be represented as

$$\mathbf{P}_K \mathbf{x} = \sum_{i=1}^{n} \alpha^i \mathbf{e}_i, \quad \alpha^i \geq 0 \tag{13}$$

and

$$\mathbf{P}_{K^\circ} \mathbf{x} = \sum_{j=1}^{n} \beta^j \mathbf{u}_j, \quad \beta^j \geq 0. \tag{14}$$

To prove existence, let $I^c = \{j \in N : \beta^j > 0\}$ and let I be the complement of I^c in the set N of indices. For an arbitrary element $\mathbf{z} \in \mathbb{R}^n$, denote $\mathbf{P}_K^\top \mathbf{z} = (\mathbf{P}_K \mathbf{z})^\top$. If $\alpha^j > 0$ would hold in (14), for some $j \in I^c$, then by Lemma 1 it would follow that $\mathbf{P}_K^\top \mathbf{x} \cdot \mathbf{P}_{K^\circ} \mathbf{x} < 0$, which contradicts the theorem of Moreau. Hence, (13) can be written in the form (11) and (14) can be written in the form (12). Therefore, Theorem 1 implies

$$\mathbf{x} = \mathbf{P}_K \mathbf{x} + \mathbf{P}_{K^\circ} \mathbf{x} = \sum_{i \in I} \alpha^i \mathbf{e}_i + \sum_{j \in I^c} \beta^j \mathbf{u}_j,$$

where $\alpha^i \geq 0$, $\forall i \in I$ and $\beta^j > 0$, $\forall j \in I^c$.

To prove uniqueness, suppose that in the representation (10) of $\mathbf{x}$ we have $\alpha^i \geq 0,\ \beta^i = 0$ for $i \in I$ and $\beta^j > 0,\ \alpha^j = 0$ for $j \in I^c$, where I is a subset of N, and I^c is the complement of I in N (the cases $I = \emptyset$ and $I = N$ are not excluded). Then representations (11) and (12) follow from Theorem 1 by using the mutual orthogonality of the vectors $\mathbf{e}_i,\ i \in I$ and $\mathbf{u}_j,\ j \in I^c$. From (11) and the uniqueness of the projection $\mathbf{P}_K\mathbf{x}$ it follows that I is unique.

□

From this theorem it follows that a given simplicial cone $K \subset \mathbb{R}^n$ determines a partition of the space $\mathbb{R}^n$ in 2^n cones in the sense that

$$\mathbb{R}^n = \bigcup_{I \subset N} \text{cone}\{\mathbf{e}_i, \mathbf{u}_j :\ i \in I,\ j \in I^c\}$$

and for two different sets I of indices the respective cones do not contain common interior points. The cones in the above union are exactly the sums of orthogonal faces.

This theorem suggests the following algorithm for finding the projection $\mathbf{P}_K\mathbf{x}$:

Step 1. For the subset $I \subset N$ we solve the following linear system in α^i

$$\mathbf{x}^\top \mathbf{e}_\ell = \sum_{i \in I} \alpha^i \mathbf{e}_i^\top \mathbf{e}_\ell,\ l \in I. \tag{15}$$

Step 2. Then, we select from the family of all subsets in N the subfamily Δ of subsets I for which the system possesses non-negative solutions.

Step 3. For each $I \in \Delta$ we solve the linear system in β^j

$$\mathbf{x}^\top \mathbf{u}_k = \sum_{j \in I^c} \beta^j \mathbf{u}_j^\top \mathbf{u}_k,\ k \in I^c. \tag{16}$$

By Theorem 2 among these systems there exists exactly one with non-negative solutions. By this theorem, for corresponding I and for the solution of the system (15), we must have

$$\mathbf{P}_K\mathbf{x} = \sum_{i \in I} \alpha^i \mathbf{e}_i.$$

This algorithm requires that we solve 2^n linear systems of at most n equations in Step 1 (15) and another $2^{|\Delta|}$ systems in Step 2 (16). (Observe that all these systems are given by Gram matrices, hence they have unique solutions.) Perhaps this great number of systems can be substantially reduced, but it still remains considerable.

Remark 1 *If K is subdual; that is, if $\mathbf{e}_k^\top \mathbf{e}_\ell \geq 0,\ k,\ l \in N$, the above algorithm can be reduced as follows: By supposing that we have got the representation (10) of x with non-negative coefficients, we multiply both sides of (10) by an arbitrary $\mathbf{e}_l^\top$. If $\mathbf{x}^\top \mathbf{e}_l < 0$ then l cannot be in I, otherwise the relations $\mathbf{u}_j \mathbf{e}_l = 0,\ j \in I^c$ and $\mathbf{e}_i^\top \mathbf{e}_l \geq 0,\ i \in I$ would furnish a contradiction. Thus, we have to look for the set I of indices (for which we have to solve the system (15)) among the subfamilies of $\{i \in N : \mathbf{x}^\top \mathbf{e}_i \geq 0\}$. (Arguments like this can be used, as it was done e. g. in [7] for the Dykstra algorithm, to eliminate some hyperplanes while comkputing successive approximations of the solution.)*

Obviously, the proposed method is inefficient. It was presented by A. B. Németh and S. Z. Németh in [8] as a preparatory matherial for an efficient algorithm for so called isotone projection cones *only. For isotone projection cones we can obtain the projection of a point in at most n steps, where n is the dimension of the space. Isotone projection cones are special simplicial cones. Even if there are important isotone projection cones in applications, they are rather particular in the family of simplicial cones.*

4 Heuristic method for projection onto a simplicial cone

Regardless the inconveniences of the above presented exact method, which follow from its combinatorial character, it suggests an interesting heuristic algorithm. To explain its intuitive background we consider again the simplicial cone

$$K = \text{cone}\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$$

and its polar

$$K^\circ = \text{cone}\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$$

given by Lemma 1.

Take an arbitrary $\mathbf{x} \in \mathbb{R}^n$. We are seeking the projection $\mathbf{P}_K x$.

If $\mathbf{e}_i^\top \mathbf{x} \leq 0,\ \forall\, i \in N$, then $\mathbf{x} \in K^\circ = \ker \mathbf{P}_K$, hence $\mathbf{P}_K \mathbf{x} = 0$.

If $\mathbf{u}_j^\top \mathbf{x} \leq 0,\ \forall\, j \in N$, then $\mathbf{x} \in K$, and hence $\mathbf{P}_K \mathbf{x} = \mathbf{x}$.

We can assume that $\mathbf{x} \notin K \cup K^\circ$. Hence, $\mathbf{x}$ projects in a proper face of K and in a proper face of K°.

Take an arbitrary family $I \subset N$ of indices. Then, the vectors

$$\mathbf{e}_i,\ \mathbf{u}_j :\ i \in I,\ j \in I^c$$

entgender by Corollary 1 a reference system in $\mathbb{R}^n$. Then,

$$\mathbf{x} = \sum_{i \in I} \alpha^i \mathbf{e}_i + \sum_{j \in I^c} \beta^j \mathbf{u}_j \tag{17}$$

with some $\alpha^i,\ \beta^j \in \mathbb{R}$. (As far as the family $I \subset N$ of indices is given, we can determine the coefficients α^i and β^j, according to Theorem 2, by solving the systems (15) and (16).)

If we have $\alpha^i \geq 0,\ \beta^j \geq 0 :\ i \in I,\ j \in I^c$, then from Theorem 2 we obtain

$$\mathbf{P}_K \mathbf{x} = \sum_{i \in I} \alpha^i \mathbf{e}_i$$

and

$$\mathbf{P}_{K^\circ} \mathbf{x} = \sum_{j \in I^c} \beta^j \mathbf{u}_j.$$

In this case $\mathbf{x}$ is projected onto face $F = \text{cone}\{\mathbf{e}_i :\ i \in I\}$ ortogonally along the subspace engendered by the elements $\{\mathbf{u}_j :\ j \in I^c\}$, roughly speaking, along the orthogonal face $F^\perp$ of F.

Suppose that $\beta^j < 0$ for some $j \in I^c$. Then, considering the reference system entgendered by $\mathbf{e}_i,\ \mathbf{u}_j,\ i \in I,\ j \in I^c$, $\mathbf{x}$ lies in its orthant with negative j^{th}

coordinate, that is in the direction of the vector $-\mathbf{u}_j$. By construction, $\mathbf{e}_j$ and $\mathbf{u}_j$ form an obtuse angle. Hence the angle of $\mathbf{e}_j$ and $-\mathbf{u}_j$ is an acute one. Thus there is a real chance that in a new reference system in which $\mathbf{e}_j$ replaces $\mathbf{u}_j$, the coordinate of $\mathbf{x}$ with respect to $\mathbf{e}_j$ has the same sign as its coordinate with respect to $-\mathbf{u}_j$, that is positive (or at least non-negative).

If we have $\alpha^i < 0$ for some $i \in I$, then by similar reasoning it seems to be advantageous to replace $\mathbf{e}_i$ with $\mathbf{u}_i$, and so on.

Thus, we arrive to the following step in our algorithm:

Substitute $\mathbf{u}_j$ with $\mathbf{e}_j$ if $\beta^j < 0$ and substitute $\mathbf{e}_i$ with $\mathbf{u}_i$ if $\alpha^i < 0$ and solve the systems (15) and (16) for the new configuration of indices I. We shall call this step an **iteration** of the heuristic algorithm.

Then, repeat the procedure for the new configuration of I and so on, until we obtain a representation (17) of $\mathbf{x}$ with all the coefficients non-negative.

5 Experimental results

The heuristic algorithm was programmed in Scilab, an open source platform for numerical computation.[2] Experiments were performed on numerical examples for 2, 3, 5, 10, 15, 20, 25, 30, 50, 75, 100, 200, 300, 500 dimensional cones. The algorithm was performed on 100000 random examples for each of the problem sizes $2, \ldots, 100$. Statistical analysis on a subset of 10000 examples from the set of 100000 examples for size 100 indicates no significant difference in overall results and performance, therefore we subsequenlty reduced the number of experiments on larger problem sizes. 10000 random examples were used for sizes 200 and 300 and 1000 examples for size 500, as the time needed by the algorithm increases with size. Table 1 shows the experimental results. For each problem size, the averages of all runs are shown, together with a confidence interval at confidence level 95% where appropriate.[3] The **Changes** column indicates the total number of swaps $\mathbf{u}_j$ for $\mathbf{e}_j$ and $\mathbf{e}_j$ for $\mathbf{u}_j$, respectively, before reaching the solution. The **Iterations** column indicates the number of iterations (as defined in Section 4) the algorithm performed before reaching a solution. The **Iterations with increases** column shows the percentage of iterations where the number of changes increased from the previous iteration. We noticed that in the majority of iterations the number of changes decreased, which led to the quick convergence of the algorithm in the vast majority of cases. In all examples the starting point for the search was the $\mathbf{e}_1, \ldots, \mathbf{e}_n$ base. The final column shows the percentage of problems where the algorithm was aborted due to going in a loop by allocating in some iteration a set of $\mathbf{e}_j$s and $\mathbf{u}_j$s that were encountered in a previous iteration. The percentage of loops was exponentially decreasing as the size increased and we did not observe any loops in any experiments on problem sizes of 30 or above. Overall, loops were observed in 0.1% of the experiments, so the heuristic algorithm was successful 99.9% of the time. A solution that we see for solving the problems that lead to a loop is to restart the algorithm from a different initial set of $\mathbf{e}_j$s and $\mathbf{u}_j$s. The problems ending in a loop were excluded from the detailed analysis that follows.

[2] http://www.scilab.org/

[3] Any difference less than ± 0.5 for integers and ± 0.1 for percentages, respectively, is not shown, as deemed irrelevant for the analysis.

Table 1: Number of changes, iterations, iterations where the number of changes increased and loops for the various cone dimensions

Size	**Changes**	**Iterations**	**Iterations with increases [%]**	**Loops [%]**
2	1	1	0	4.382
3	2	1	3.9 ± 0.1	4.278
5	4	2	13 ± 0.2	1.396
10	11	3	26 ± 0.3	0.273
15	17	4	30.3 ± 0.3	0.029
20	24	4	31.6 ± 0.3	0.007
25	30	4	31.2 ± 0.3	0.003
30	37	4	29.9 ± 0.3	–
50	64	5	26.8 ± 0.3	–
75	97	5	24.2 ± 0.3	–
100	131	5	23.8 ± 0.3	–
200	267 ± 1	6	19.9 ± 0.8	–
300	409 ± 1	6	26.2 ± 0.9	–
500	700 ± 5	7	25.7 ± 2.7	–

More detailed analysis of the three main performance indicators of changes, iterations and iterations with increases was performed using boxplots as shown in Figures 1, 2 and 3. Although the total number of changes performed increases linearly with problem size (at a rate of less than $2 \times n$, even if considering maximum numer of changes), this does not affect the performance substantially (see Figure 1, where the results were split into two parts for a clearer view).

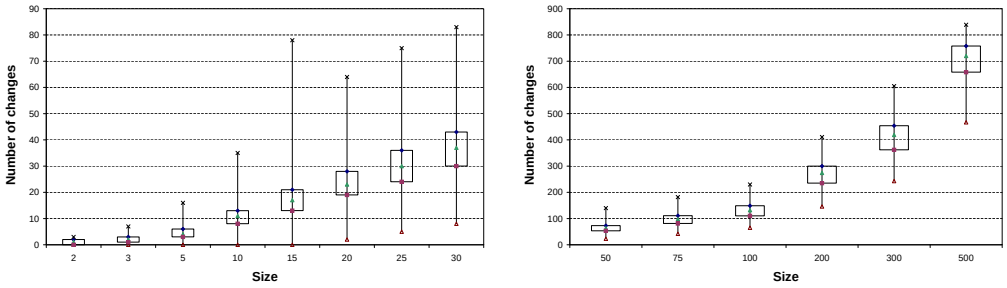

Figure 1: Boxplots of number of changes performed for the various cone dimensions

The number of iterations is the crucial indicator of performance. As shown in Figure 2 the number of iterations reaches at most 11 (for sizes 15 and 20), but in 75% of cases has a value of 7 or below. We ran a few experiments on larger sizes, up to 1750, and the largest number of iterations we obeserved was 13. Running experiments on very large problem sizes is problematic due to computer memory limitations and the Scilab built-in solving of linear systems.[4]

[4] Note that the time needed by one iteration substantially increases with problem size n as one iteration involves solving a linear system with n equations and n variables.

The major benefit of this heuristic algorithm is the small number of iterations even for very large number of cone dimensions.

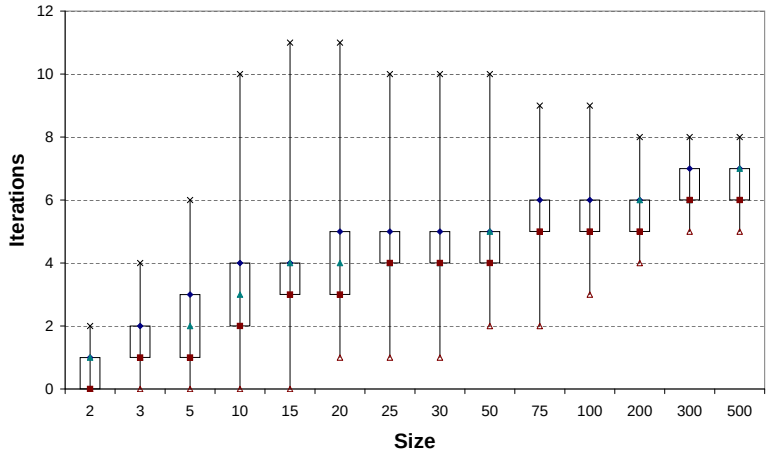

Figure 2: Boxplots of number of iterations needed for the various cone dimensions

As our heuristic algorithm seems to converge quickly, we wanted to know how frequently it deviates from the optimal path. An optimal path would consist of iterations with decreasing number of changes. Figure 3 shows the boxplots for the number of iterations where an increase in the number of changes took place. The maximum number of such iterations over all experiments is 4, but 75% of examples involved only one or no such iteration. This provides an explanation for the fast convergence: the algorithm very rarely deviates from the optimal path.

6 Conclusion

We presented a heuristic method of projection on simplicial cones based on Moreau's decomposition theorem. The heuristic algorithm presented in this note iteratively finds the projection onto a simplicial cone in a surprisingly small number of steps even for large cone dimensions in 99.9.% of the cases. We attribute the success to the fact that the algorithm rarely deviates from the optimal path, in every iteration it usually has to change less base values than in the previous iteration. We are planning to further extend the algorithm with random restart hoping to achieve 100% success rate.

Acknowledgements

S. Z. Németh was supported by the Hungarian Research Grant OTKA 60480. The authors are grateful to J. Dattorro for many helpful conversations.

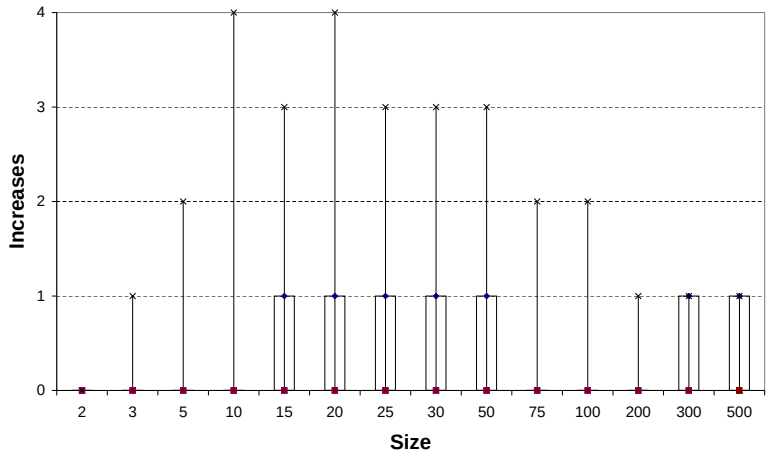

Figure 3: Boxplots of number of iterations with increases in number of changes needed for the various cone dimensions

References

[1] J. Dattorro. *Convex Optimization & Euclidean Distance Geometry.* $\mathcal{M}\varepsilon\beta oo$, 2005, v2009.04.11.

[2] F. Deutsch and H. Hundal. The rate of convergence of Dykstra's cyclic projections algorithm: the polyhedral case. *Numer. Funct. Anal. Optim.*, 15(5-6):537–565, 1994.

[3] R. L. Dykstra. An algorithm for restricted least squares regression. *J. Amer. Stat. Assoc.*, 78(384):273–242, 1983.

[4] G. Isac and A. B. Németh. Projection methods, isotone projection cones, and the complementarity problem. *J. Math. Anal. Appl.*, 153(1):258–275, 1990.

[5] T. Ming, T. Guo-Liang, F. Hong-Bin, and Ng. Kai Wang. A fast EM algorithm for quadratic optimization subject to convex constraints. *Statist. Sinica*, 17(3):945–964, 2007.

[6] J. J. Moreau. Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires. *C. R. Acad. Sci.*, 255:238–240, 1962.

[7] P. M. Morillas. Dykstra's algorithm with strategies for projecting onto certain polyhedral cones. *Applied Mathematics and Computation*, 167(1):635–649, 2005.

[8] A. B. Németh and S. Z. Németh. How to project onto an isotone projection cone. *Linear Algebra Appl.*, submitted.

[9] S. Z. Németh. Iterative methods for nonlinear complementarity problems on isotone projection cones. *J. Math. Anal. Appl.*, 350(1):340–347, 2009.

[10] X. Shusheng. Estimation of the convergence rate of Dykstra's cyclic projections algorithm in polyhedral case. *Acta Math. Appl. Sinica (English Ser.)*, 16(2):217–220, 2000.