TECHNICAL NOTE

A Finite Algorithm for Finding the Projection of a Point onto the Canonical Simplex of $\mathbb{R}^n$

C. MICHELOT

Communicated by A. V. Fiacco

Abstract. An algorithm of successive location of the solution is developed for the problem of finding the projection of a point onto the canonical simplex in the Euclidean space $\mathbb{R}^n$. This algorithm converges in a finite number of steps. Each iteration consists in finding the projection of a point onto an affine subspace and requires only explicit and very simple computations.

Key Words. Nonlinear programming, quadratic programming, projection onto a simplex, optimality conditions.

1. Introduction

It is very important to know how to compute the projection of a point onto a polyhedron. This problem arises in many contexts, in particular in constrained linear optimization methods, such as gradient projection methods (Refs. 1-3). A great deal of papers have been devoted to this problem, and the authors applied themselves to describing stable and finite algorithms (Refs. 4-8).

In this paper, we present a finite projection algorithm for the canonical simplex of $\mathbb{R}^n$. This procedure is recursive. At each iteration, using Lagrange multipliers, we locate the solution and we are going to solve the same problem in a strict lower-dimensional space. The maximum number of iterations is equal to the initial dimension of the space. Moreover, each iteration requires only explicit and very simple computations.
2. Preliminaries

We denote by \( X = \mathbb{R}^n \) [resp., \( Y = \mathbb{R}^m \)] the \( n \)-dimensional [resp., \( m \)-dimensional] Euclidean space, by \( \langle \cdot, \cdot \rangle_n \) and \( \| \cdot \|_n \) [resp., \( \langle \cdot, \cdot \rangle_m \) and \( \| \cdot \|_m \)] the usual inner product and the Euclidean norm of \( X \) [resp., \( Y \)]. Let \( x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \). Then, we write \( x \geq 0 \) if \( x_i \geq 0 \) for all \( i = 1, 2, \ldots, n \). \( A^T \) denotes the transpose of the \( m \times n \) matrix \( A \).

Let \( K \subset X \) be a nonempty, closed, convex set; and let \( c \) be a point such that \( c \notin K \). We denote by \( p_K(c) \) the projection of \( c \) onto \( K \). \( p_K(c) \) is the optimal solution to the following problem (P):

\[
(P) \quad \inf \frac{1}{2} \| x - c \|_n^2, \quad x \in K.
\]

Let us suppose that \( K \) is a polyhedron defined by

\[
K = \{ x \in X | Ax = b, x \geq 0 \},
\]

where \( A \) is an \( m \times n \) matrix with full row rank \( m \) and \( b \) is a vector of \( Y \).

For solving (P), we can introduce the Lagrangian \( L \), defined by

\[
L(x; (\lambda, \mu)) = \frac{1}{2} \| x - c \|_n^2 + \langle \lambda, Ax - b \rangle_m - \langle \mu, x \rangle_n, \quad x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m, \mu \geq 0, \quad \text{otherwise}.
\]

Then, we have the following well-known result (Ref. 9).

**Theorem 2.1.** \( x^* \) is an optimal solution to (P) if and only if there exist \( \lambda^* \in Y, \mu^* \in X \) such that \( (x^*; (\lambda^*, \mu^*)) \) is a saddle point of \( L \).

Let \( V \subset X \) be the affine subspace defined by

\[
V = \{ x \in X | Ax = b \}.
\]

Then, we have the following result.

**Corollary 2.1.** Let \( c \in V, c \notin K \). Then, \( x^* \in K \) is an optimal solution to (P) if and only if there exists \( \mu^* \in X \) such that

\[
x^* - c + A^T(AA^T)^{-1}A\mu^* = \mu^*, \tag{1}
\]

\[
\langle x^*, \mu^* \rangle_n = 0, \tag{2}
\]

\[
\mu^* \geq 0. \tag{3}
\]

**Proof.** \( (x^*; (\lambda^*, \mu^*)) \) is a saddle point of \( L \) if and only if the following statements are satisfied:

\[
x^* \in K, \tag{4}
\]

\[
x^* - c + A^T\lambda^* = \mu^*, \tag{5}
\]

\[
\langle x^*, \mu^* \rangle_n = 0, \tag{6}
\]

\[
\mu^* \geq 0. \tag{7}
\]
If we premultiply both sides of Eq. (5) by $A$, we get
\[ AA^T \lambda^* = A\mu^*. \]

Since $A$ is of rank $m$, the operator $AA^T$ has an inverse. Then, we can substitute for $\lambda^*$ into Eq. (5), and the result follows.

In fact, we can always assume that $c \in V$ in view of the following theorem.

**Theorem 2.2.** Let $V \subset X$ be an affine subspace; let $K \subset V$ be a non-empty, closed, convex set; and let $c \not\in V$. Then, we have
\[ p_K(c) = p_K(p_V(c)). \]

**Proof.** From the characterization of the projection of $c$ on $K$, we have
\[ \langle c - p_K(c), x - p_K(c) \rangle_n = 0, \quad \text{for all } x \in K. \]
This inequality can be rewritten as
\[ \langle c - p_V(c), x - p_K(c) \rangle_n + \langle p_V(c) - p_K(c), x - p_K(c) \rangle_n = 0, \quad \text{for all } x \in K. \]
Now, we have $x - p_K(c) \in V$ and $c - p_V(c) \in V^\perp$, where $V^\perp$ is the orthogonal subspace to $V$. This implies that
\[ \langle p_V(c) - p_K(c), x - p_K(c) \rangle_n = 0, \quad \text{for all } x \in K. \]
Then, the proof is complete.

3. **Projection of a Point onto the Canonical Simplex**

In this section, we shall give a method for solving (P) in the following case:
\[ (P) \quad \inf_{x \in K} \frac{1}{2} \|x - c\|_n^2, \quad x \in K, \]
where
\[ K = \left\{ x \in X \middle| \sum_{i=1}^n x_i = 1, x \geq 0 \right\}. \]

We begin with some additional notations. We denote
\[ I_n = \{1, 2, \ldots, n\}, \]
\[ V = \left\{ x \in X \middle| \sum_{i=1}^n x_i = 1 \right\}. \]
For an arbitrary subset $I$ of $I_n$,

$X_I = \{x \in X | x_i = 0, \text{ for all } i \in I\}$

$V_I = X_I \cap V$

$K_I = X_I \cap K$

$n_I = \dim(X_I)$

where $X_I$ is a linear subspace of $X$.

Let $c \in X_I$, $c \notin K_I$. Consider the following problem (P$_I$):

$$(P_I) \quad \inf \frac{1}{2} \|x - c\|^2_n, \quad x \in K_I.$$

From Corollary 2.1, $x^* \in K_I$ is an optimal solution to (P$_I$) if and only if there exists $\mu^* \in X_I$ such that

$$x^*_i - c_i + \left(\sum_{j} \mu^*_j\right) / n_I = \mu^*_i, \quad \text{for all } i \notin I,$$

$$\langle x^*, \mu^* \rangle_n = 0,$$

$$\mu^* \geq 0.$$  

By means of these optimality conditions, we obtain the main result.

**Theorem 3.1.** Let $I \subset I_n$, $c \in V_I$, and put $\tilde{I} = \{i \notin I | c_i < 0\}$. Then, if $\tilde{I} \neq \emptyset$, we have $p_{K_I}(c) \in K_{I \cup \tilde{I}}$.

**Proof.** Put $x^* = p_{K_I}(c)$, and let $i \in \tilde{I}$. By assumption, $i \notin I$. The left-hand side of equality (8) is strictly positive. Then, we have $x^*_i = 0$, since the corresponding multiplier $\mu^*_i$ is different from zero. This means that $x^* \in K_{I \cup \tilde{I}}$.  

Now, Theorem 2.2 and Theorem 3.1 give a very simple procedure, in three stages, for solving (P): Find the projection onto an affine subspace $V_I$; locate $p_{K}(c)$ in a convex set $K_J$ with $I \subset J$; find the projection onto the linear subspace $X_J$; replace $I$ by $J$, and repeat the process. We obtain in more detail the following algorithm.

4. Algorithm

**Description**

**Step 1.** Initialization. Put $I = \emptyset$ and $x = c$.

**Step 2.** Iteration. Let $I \subset I_n$, and let $x \in X_I$, such that $x \notin V_I$. Compute $\tilde{x} = P_{V_I}(x)$. If $\tilde{x} \geq 0$, then stop: $\tilde{x} = p_{K_I}(c)$. Otherwise, replace $I$ by $I \cup \{i | \tilde{x}_i < 0\}$, and replace $c$ by $P_{X_I}(\tilde{x})$. 
Convergence. The algorithm converges with at most \( n \) iterations, since the dimension \( n_t \) of the subspace \( X_t \) decreases by at least one unit.

Implementation. From a practical point of view, the algorithm requires the computation of \( \bar{x} = P_{\varphi_t}(x) \), which is explicitly made by the relation:

\[
\bar{x}_i = x_i - \left( \sum_{j \in I} x_j - 1 \right) / n_t, \quad \text{if } i \notin I, \\
\bar{x}_i = 0, \quad \text{if } i \in I.
\]

The algorithm requires also the computation of \( x = P_{\varphi_0}(\bar{x}) \), which consists only in putting

\[
x_i = \bar{x}_i, \quad \text{if } \bar{x}_i > 0, \\
x_i = 0, \quad \text{otherwise}.
\]

5. Concluding Remarks

In general, the algorithm is very efficient, given that at each iteration the dimension \( n_t \) of the subspace \( X_t \) often decreases by several units. It is also to be noted that the method can be generalized to the case where the convex set \( K \) is given by

\[
K = \left\{ x \in X \left| \sum_{i=1}^{n} \alpha_i x_i = 1, x \geq 0 \right. \right\},
\]

where \( \alpha_i \) is a strictly positive real, \( i = 1, 2, \ldots, n \).

References


