Computational Mathematics and Mathematical Physics, Vol. 45, No. 11, 2005, pp. 1915–1922.
Translated from Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, Vol. 45, No. 11, 2005, pp. 1991–1999.

Convergence of the Suitable Affine Subspace Method for Finding the Least Distance to a Simplex

E. A. Nurminskii

Institute for Automation and Control Processes, Far East Division, Russian Academy of Sciences, ul. Radio 5, Vladivostok, 690041 Russia

e-mail: nurmi@dvo.ru

Received March 28, 2005

Abstract—A minimum-length vector is found for a simplex in a finite-dimensional Euclidean space. The algorithm of successive projections onto affine subspaces containing suitable subsimplices of the initial simplex is shown to have a globally higher-than-linear convergence rate. Results of numerical experiments are presented.

Keywords: projection, minimum-norm element, simplex.

1. INTRODUCTION

The problem of determining the minimum-norm element in a given set is frequently encountered in theoretical foundations of many applied mathematics areas and in applications that are often characterized by large dimensions and severe limitations on the CPU time required for problem solving. Motivated by these requirements, several methods based on efficient projection algorithms have been created. An overview of them can be found, for example, in [1].

Since the problem has a large dimension, we have to solve it by iterative methods. As a result, the important problem of increasing the convergence rate of such algorithms arises. One of the basic approaches can be described as follows. The set C on which a projection is constructed is represented as the intersection of "simple" sets C_i ($i = 1, 2, \ldots, N$), which are half-spaces, cubes, simple cones, etc., and iterative algorithms involving easily computable projections onto C_i are constructed. At the same time, the internal description of C as the convex hull of a family of simple sets naturally arises in some problems:

$$C = \operatorname{co}\{C_i, i = 1, 2, \ldots, N\} = \left\{\bigcup_{i=1}^{N} \lambda_i C_i : \sum_{i=1}^{N} \lambda_i = 1, \lambda_i \geq 0, i = 1, 2, \ldots, N\right\},$$

and the minimum-norm element in C can also be sought using algorithms based on individual projections onto C_i and/or simple modifications of these sets (see, for example, [2]). This poses the problem of constructing an effective projection onto C_i. As a step in the direction of its efficient solution, in [3], an algorithm of projection onto C_i was proposed when the latter are simplices, i.e., the convex hulls of affinely independent vectors of a finite-dimensional Euclidean space $\mathbb{E}$. The algorithm uses projections onto the affine hulls of subsimplices of the initial simplex and, to avoid confusion with successive projection methods [4, 5], we will refer to it as the affine subspace method (ASM).

In [3], the finite convergence of ASM was proved and promising numerical results were presented that demonstrated the high performance of the method. However, no theoretical estimate of the convergence was given.

The goal of this paper is to establish a global higher-than-linear convergence rate of this algorithm and to demonstrate some promising numerical results.

2. BASIC CONCEPTS AND NOTATIONS

The problem of searching for a minimum-norm element is considered in a finite-dimensional Euclidean space $\mathbb{E}$ with the conventional norm $\|x\| = \sqrt{xx}$ associated with the scalar product xy. Hereafter, it is assumed that the dimension of $\mathbb{E}$ is n.

Along with the convex hull mentioned above, we will also use the affine hull.

Definition 1. For $C \subset \mathbb{E}$, the set aff$\{C\}$ defined as

$$\text{aff}\{C\} = \left\{ c = \sum_{i=1}^{n+1} \lambda_i c^i, \sum_{i=1}^{n+1} \lambda_i = 1, c^i \in C, i = 1, 2, \ldots, n+1 \right\}$$

is called the *affine hull* of C.

The affine hull aff$\{C\}$ is the smallest affine subspace containing C. A set of points$\{x^1, x^2, \ldots, x^m\}$ of $\mathbb{E}$ is called affinely independent if

$$\sum_{i=1}^{m} \lambda_i x^i = 0, \quad \sum_{i=1}^{m} \lambda_i = 0$$

imply $\lambda_i = 0$ for $i = 1, 2, \ldots, m$.

Let ri$\{A\}$ denote the relative interior of A, i.e., the interior with respect to aff$\{A\}$.

Every affine subspace A can be assigned a certain linear space L_A, which is called the associated space. In this case, it can be defined as

$$L_A = \left\{ x = \sum_{i=1}^{n} \lambda_i a^i, \text{ where } \sum_{i=1}^{n} \lambda_i = 0, a^i \in A, i = 1, 2, \ldots, n \right\}.$$

In $\mathbb{E}$, we consider the problem of finding the least distance from a given convex set X to the origin:

$$\max_{z \in X} \|z\|^2 = \|z^*\|^2, \quad z^* \in X. \tag{1}$$

It is assumed that X is a simplex, i.e., a convex polyhedron specified by affinely independent vertices:

$$X = \text{co}\{x^i, i \in \mathcal{N} \triangleq \{1, 2, \ldots N\}\}. \tag{2}$$

Affine independence implies that $N \le n + 1$. The conditions for the optimality of z^* in problem (1) have the form of the simplest variational inequality

$$z^*(x - z^*) \ge 0 \quad \forall x \in X. \tag{3}$$

In the case of simplex (2), they are reduced to the verification of N inequalities: $z^* x^i \ge \|z^*\|^2$, $i \in \mathcal{N}$.

The algorithm proposed below makes use of projections onto the affine hulls of subsimplices of X; i.e., it uses solutions to problems of the form

$$\min_{z \in \text{aff}(S_I)} \|z\|^2, \tag{4}$$

where $S_I = \text{co}\{x^i, i \in I \subset \mathcal{N}\}$. An important role is played by the following easy-to-check property: if z^* is a solution to problem (4), then $z^* y = 0$ for any $y \in L_{\text{aff}(S_I)}$.

To describe the algorithm and analyze its convergence, we introduce the following useful concept.

Definition 2. The set $S_I = \text{co}\{x^i, i \in I\}$ defined by an index set $I \subset \mathcal{N}$ is called a suitable subsimplex if

$$\min_{z \in \text{aff}\{S_I\}} \|z\|^2 = \min_{z \in S_I} \|z\|^2.$$

Obviously, for example, any singleton set I defines a suitable subsimplex.

The set of points $\{x^i, i \in I\}$, together with I describing $S = \text{co}\{x^i, i \in I\}$, is referred to as a basis. The basis of a suitable subsimplex is called a suitable basis. The number of vectors in I is denoted by $|I|$.

The affine hull of a suitable subsimplex is called a suitable affine subspace.

3. THE AFFINE SUBSPACE METHOD

The algorithm consists of iterations, each of which begins when there is a suitable subsimplex of the initial simplex (a basis, used as input data) and is completed when a new suitable subsimplex (new basis) is constructed whose distance to the origin is strictly smaller than that in the preceding subsimplex.

By using a structural programming pseudolanguage, the method can be described as follows:

Initialization. Let I_0 be an initial basis generating a suitable subsimplex $S_{I_0} \triangleq S_0$ and an affine subspace H_0:

$$S_0 = \mathrm{co}\{x^i, i \in I_0\}, \quad H_0 = \mathrm{aff}\{x^i, i \in I_0\} = \mathrm{aff}\{S_0\}.$$

The iteration index is set equal to zero: $k = 0$.

Basic Iteration

Step 1. *Projection onto H_k.* Solve the problem

$$\min_{z \in H_k} \|z\|^2 = \|z^k\|^2. \tag{5}$$

Step 2. *Verification of optimality.* By virtue of (3), if $x^i z^k \ge \|z^k\|^2$ for all $i \in \mathcal{N} \setminus I_k$, then z^k is a solution to problem (1), since $x^i z^k \ge \|z^k\|^2$ for $i \in I_k$ by construction. The algorithm halts.

Step 3. Initialization of inner iteration for constructing a new basis (executed if the optimality conditions for z^k are not satisfied). Choose an arbitrary $i_k \in \mathcal{N} \setminus I_k$ such that

$$x^{i_k} z^k < \|z^k\|^2. \tag{6}$$

Set the inner iteration index equal to zero: $s = 0$. Set $J_s = I_k \cup \{i_k\}$ and $w^s = z^k$.

Step 4. *Inner iteration.*

(a) *Projection onto a modified basis.* Form a new subsimplex $T_s = \mathrm{co}\{x^i, i \in J_s\}$ and a new affine subspace $G_s = \mathrm{aff}\{T_s\}$ and solve the auxiliary projection problem

$$\min_{y \in G_s} \|y\|^2 = \|y^s\|^2. \tag{7}$$

(b) *Feasibility verification.* If $y^s \in T_s$ (i.e., T_s is a suitable subsimplex), then set $I_{k+1} = J_s$ and $H_{k+1} = G_s$ and exit the inner iteration.

(c) *Basis correction* (executed if $y^s \notin T_s$). Set

$$u(\mu) = \mu y^s + (1-\mu) w^s \tag{8}$$

and find a maximum number μ such that $u(\mu) \in T_s$; i.e.,

$$\mu_s = \max_{u(\mu) \in T_s} \mu. \tag{9}$$

By construction, the point $u(\mu)$ with $\mu = \mu_s$ belongs to the relative interior of a minimal edge of T_s, which in turn defines the set of its extreme points x^i, $i \in J_{s+1}$; i.e.,

$$u(\mu_s) = \sum_{i \in J_{s+1}} \theta_i x^i \triangleq w^{s+1}, \tag{10}$$

where $J_{s+1} \subset \mathcal{N}$ and $\sum_{i \in J_{s+1}} \theta_i = 1$, $\theta_i > 0$ for $i \in J_{s+1}$. Increase the inner iteration index by one ($s \Rightarrow s + 1$) and return to the beginning of the inner iteration.

Remark. Since $| |J_{s+1}| < |J_s|$, the inner iteration is executed a finite number of times and necessarily produces a suitable subsimplex, in the extreme case, a singleton.

Step 5. *End of the inner iteration.* Increase the outer iteration index by one ($k \Rightarrow k + 1$) and return to the beginning of the outer iteration.

End of the basic iteration.

End of the algorithm.

It can be seen from the description that an iteration step of the algorithm begins with a suitable basis I_k generating the corresponding suitable affine subspace $H_k = \mathrm{aff}\{x^i, i \in I_k\}$ and a suitable subsimplex $S_k = \mathrm{co}\{x^i, i \in I_k\}$ and is completed after a new suitable basis I_{k+1} with the corresponding affine subspace $H_{k+1} = \mathrm{aff}\{x^i, i \in I_{k+1}\}$ and with a suitable subsimplex $S_{k+1} = \mathrm{co}\{x^i, i \in I_{k+1}\}$ is constructed.

It will be shown later that ASM (i) converges in a finite number of iterations and (ii) has a globally linear convergence rate. Rigorous statements of these results are given below.

4. CONVERGENCE OF THE AFFINE SUBSPACE METHOD

The following theorem states the convergence of ASM and provides an estimate for its convergence rate.

Theorem 1. *There is* $\bar{k}$ *such that*

$$\min_{z \in H_{\bar{k}+1}} \|z\| = \min_{z \in X}\|z\| = \|z^*\|, \quad z^* \in X,$$

and there is $\rho \in [0, 1)$ *satisfying*

$$\min_{z \in H_{k+1}} \|z\| - \|z^*\| \le \rho(\min_{z \in H_k}\|z\| - \|z^*\|)$$

for $k = 1, 2, \ldots, \bar{k}$.

Proof. First, we show that the inner iterations, at least, do not increase $\|w^s\|$. Indeed, in (8), we have y^s, $w^s \in H_s$, and $\|y^s\| \le \|w^s\|$ and, since the norm is convex,

$$\|u(\mu)\|^2 \le \mu\|y^s\|^2 + (1-\mu)\|w^s\|^2 \le \mu\|w^s\|^2 + (1-\mu)\|w^s\|^2 = \|w^s\|^2.$$

Setting $\mu = \mu_s$ yields $u(\mu_s) = w^{s+1}$; therefore, $\|w^{s+1}\| \le \|w^s\|$.

Consequently, if the inner iteration is exited, for example, at the $\bar{s}$th step, then

$$\|z^{k+1}\| = \min_{z \in H_{k+1}} \|z\| = \|y^{\bar{s}}\| = \|w^{\bar{s}}\| \le \|w^{\bar{s}-1}\| \le \ldots \le \|w^0\| = \|z^k\|, \tag{11}$$

which guarantees that the sequence $\|z^k\|$ ($k = 0, 1, \ldots, \bar{k}$) at least does not increase.

To prove that it strictly decreases, it suffices, in view of (11), to show that the strict inequality $\|w^{s+1}\| < \|w^s\|$ holds for at least one $0 \le s < \bar{s}$ at the inner iteration. In particular, we show that it holds for $s = 0$. Then $w^0 = z^k \in \operatorname{ri}\{T_0\}$. Let

$$\delta = \min_{i \in I_k = J_0} \lambda_i > 0$$

in the representation

$$w^0 = \sum_{i \in J_0} \lambda_i x^i.$$

Suppose that y^0 solves problem (7) at $s = 0$. Then, the following is true:

(A) y^0 is either representable as

$$y^0 = \theta x^{i_k} + (1-\theta)h^0 \text{ for some } \theta \ne 1, \quad h^0 \in G_0,$$

(B) or as

$$y^0 = x^{i_k} + g \text{ for some } g \in L_{G_0} = L_0^{\Delta},$$

where L_{G_0} (otherwise L_0) is the linear subspace associated with G_0.

We show that $\theta > 0$ in case (A). Indeed, let

$$y^0 = \theta x^{i_k} + (1-\theta)h^0,$$

or

$$0 = \theta(x^{i_k} - y^0) + (1-\theta)(h^0 - y^0).$$

Multiplying this equality by $w^0 - y^0$ gives

$$0 = \theta(x^{i_k} - y^0)(w^0 - y^0) + (1-\theta)(h^0 - y^0)(w^0 - y^0). \tag{12}$$

Note that

$$(x^{i_k} - y^0)(w^0 - y^0) = x^{i_k}(w^0 - y^0),$$

since y^0 is orthogonal to L_0, which follows from the optimality of y^0 as a solution to problem (7). Furthermore,

$$x^{i_k}(w^0 - y^0) = x^{i_k}w^0 - x^{i_k}y^0 \le \|w^0\|^2 - x^{i_k}y^0$$

according to the choice of x^{i_k} in (6). In turn,

$$\begin{aligned}\|w^0\|^2 - x^{i_k}y^0 &= \|w^0 - y^0\|^2 - x^{i_k}y^0 + 2w^0y^0 - \|y^0\|^2 = \|w^0 - y^0\|^2 - x^{i_k}y^0 \\ + w^0y^0 + w^0y^0 - \|y^0\|^2 &= \|w^0 - y^0\|^2 - y^0(x^{i_k} - w^0) + y^0(w^0 - y^0) = \|w^0 - y^0\|^2,\end{aligned} \tag{13}$$

since both $x^{i_k} - w^0$ and $w^0 - y^0$ belong to L_0 and, hence, are orthogonal to y^0.

Furthermore,

$$\begin{aligned}(h^0 - y^0)(w^0 - y^0) &= (h^0 - y^0 + w^0 - w^0)(w^0 - y^0) = \|w^0 - y^0\|^2 + (h^0 - w^0)(w^0 - y^0) \\ &= \|w^0 - y^0\|^2 + w^0(h^0 - w^0) - y^0(h^0 - y^0) = \|w^0 - y^0\|^2,\end{aligned} \tag{14}$$

since $w^0(h^0 - w^0) = 0$ by virtue of the optimality of w^0 on H_k and $y^0(h^0 - y^0) = 0$ by virtue of the optimality of y^0 on G_0, as mentioned above.

Then, if $\theta < 0$, Eq. (12), combined with (13) and (14), yields

$$0 \ge \theta\|w^0 - y^0\|^2 + (1 - \theta)\|w^0 - y^0\|^2 = \|w^0 - y^0\|^2 > 0,$$

which is a contradiction.

The equality $\theta = 0$ is ruled out by choosing x^{i_k}. Indeed, if this is not the case, then $y^0 = h^0$ for some $h^0 \in H_k$ and

$$\|w^0\|^2 \le \|y^0\|^2 = \min_{z \in \operatorname{aff}\{x^{i_k}, H_k\}} \|z\|^2 \le \min_{\lambda \in [0, 1]} \|(1 - \lambda)w^0 + \lambda x^{i_k}\|^2 = \|(1 - \lambda_*)w^0 + \lambda_* x^{i_k}\|^2.$$

Direct calculation shows that

$$\lambda_* = \min\left\{1, \frac{\|w^0\|^2 - x^{i_k}w^0}{\|w^0 - x^{i_k}\|^2}\right\} > 0$$

in view of (6). Moreover, $(1 - \lambda_*)w^0 + \lambda_* x^{i_k} \ne w^0$ and the strict convexity of the norm implies that

$$\|w^0\|^2 \le \|y^0\|^2 \le \|(1 - \lambda_*)w^0 + \lambda_* x^{i_k}\|^2 < \|w^0\|^2,$$

which again leads to a contradiction.

Note that $\theta > 0$ implies $\|y^0\| < \|w^0\|$. According to (8)–(10), let

$$w^1 = (1 - \mu_0)w^0 + \mu_0 y^0,$$

where

$$\mu_0 = \max_{\mu \in [0, 1]} \mu : \mu w^0 + (1 - \mu)y^0 \in T_0.$$

Since

$$w^0 = \sum_{i \in I_k} \lambda_i x^i + 0 \cdot x^{i_k}, \quad y^0 = \sum_{i \in I_k} \bar{\lambda}_i x^i + \theta x^{i_k}$$

with $\lambda_i \geq \delta > 0$, $|\bar{\lambda}_i| < \Delta$ ($i \in I_k$), and $\theta > 0$, we conclude that μ_0 satisfies the estimate

$$\mu_0 = \min_{\bar{\lambda}_i < 0,\, i \in I_k} \frac{\lambda_i}{\lambda_i - \bar{\lambda}_i} \geq \frac{\delta}{\delta + \Delta} > 0,$$

where Δ is an easy-to-estimate constant. Then

$$\|w^1\| \leq \mu_0 \|w^0\| + (1 - \mu_0)\|w^0\| < \|w^0\|, \tag{15}$$

as required.

Case (B) (i.e., $y^0 = x^{i_k} + g$, where $g \in L_0$) is much easier. By construction, we have $[x^{i_k}, w^0] \subset T_0$ and, since $w^0 \in \mathrm{ri}\{S_k\} \subset H_k$, there exists $\gamma > 0$ such that $w^0 + \gamma g \in S_k \subset T_0$. Therefore, $[x^{i_k}, w^0 + \gamma g] \subset T_0$.

Then, picking $\bar{\mu} = \gamma/(1 + \gamma)$, we obtain

$$\bar{\mu} y^0 + (1 - \bar{\mu}) w^0 = \bar{\mu}(x^{i_k} + g) + (1 - \bar{\mu}) w^0$$

$$= \frac{\gamma}{1+\gamma}(x^{i_k} + g) + \left(1 - \frac{\gamma}{1+\gamma}\right) w^0 = \frac{\gamma}{1+\gamma} x^{i_k} + \frac{1}{1+\gamma}(w^0 + \gamma g) \in T_0,$$

and, hence,

$$\mu_0 = \max_{\mu \in [0,1]} \mu : (1 - \mu) w^0 + \mu y^0 \in T_0 \geq \bar{\mu} > 0,$$

which, as earlier, implies (15). If $\|z^{k+1}\| < \|z^k\|$ (which follows from (11) and (15)), we conclude that the algorithm is finite because the number of suitable subsimplices of X is finite and none of them can be used twice. Moreover, the algorithm cannot stop at a nonoptimal point since the generated sequence is strictly monotone if optimality conditions (3) are violated.

To prove that the method converges exponentially, we show that $\|y^0\|^2 \leq (1 - \gamma^2)\|z^k\|^2$ for some $\gamma^2 \in (0, 1]$. This follows from the estimates

$$\|y^0\|^2 = \min_{y \in G_0} \|y\|^2 \leq \min_{\lambda} \|z^k + \lambda(x^{i_k} - z^k)\|^2 = \|z^k\|^2 + \min_{\lambda}\{2\lambda z^k(x^{i_k} - z^k) + \lambda^2 \|x^{i_k} - z^k\|^2\}$$

$$= \|z^k\|^2 + \|z^k\|\|x^{i_k} - z^k\| \min_{\lambda}\{\delta\lambda^2 - 2\gamma\lambda\},$$

where

$$\delta = \frac{\|x^{i_k} - z^k\|}{\|z^k\|}, \quad \gamma = \frac{z^k(z^k - x^{i_k})}{\|x^{i_k} - z^k\|\|z^k\|} \in (0, 1].$$

Therefore,

$$\|y^0\| \leq \|z^k\|^2 - \|z^k\|\|x^{i_k} - z^k\|\gamma^2/\delta = (1 - \gamma^2)\|z^k\|^2.$$

Now, we can estimate

$$\|z^{k+1}\|^2 = \|z^{k+1} - y^0\|^2 + \|y^0\|^2 \leq (1-q)\|z^k - y^0\|^2 + \|y^0\|^2 \leq (1-q)\|z^k - y^0\|^2 + (1-q)\|y^0\|^2 + q\|y^0\|^2$$

$$= (1-q)(\|z^k - y^0\|^2 + \|y^0\|^2) + q\|y^0\|^2 \tag{16}$$

$$= (1-q)\|z^k\|^2 + q\|y^0\|^2 \leq (1-q)\|z^k\|^2 + q(1-\gamma^2)\|z^k\|^2 = (1 - q\gamma^2)\|z^k\|^2 = \rho^2\|z^k\|^2.$$

It follows that

$$\|z^{k+1}\| - \min_{z \in X}\|z\| = \|z^{k+1}\| - \|z^*\| \leq \rho\|z^k\| - \|z^*\| < \rho(\|z^k\| - \|z^*\|), \tag{17}$$

as required.

Note that, by virtue of strict inequality (17), the resulting convergence rate estimate is a better-than-linear convergence rate for $\|z^k\| - \|z^*\|$, which is also confirmed by the numerical experiments described below.

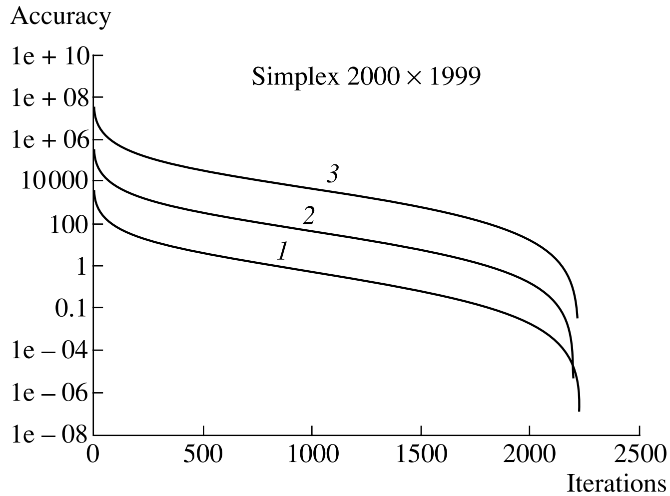

Fig. 1. Curves for (*1*) $\sigma^2 = 10$, (*2*) $\sigma^2 = 1000$, and (*3*) $\sigma^2 = 10\,000$; the shift parameter is $\delta = 0.001$.

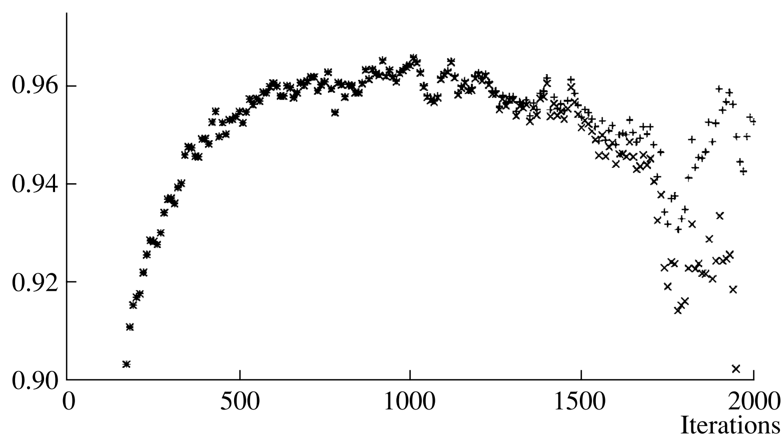

Fig. 2. Notes: (+) for r1, (×) for r2.

5. NUMERICAL EXPERIMENTS

In the numerical experiments with the algorithm, the initial simplex was generated at random to be a set of vectors with components that are independent uniformly distributed random variables. To generate stress tests, we used scaling so the last coordinate of each vector had a much smaller range of variation than the remaining coordinates. More precisely, in each test, the initial data set was a family of $n - 1$ n-dimensional vectors of the form $x = (\xi_1, \ldots, \xi_n)$ with their components ξ_i calculated as

$$\xi_i = \begin{cases} \sigma(\zeta_i - 0.5), & i = 1, 2, \ldots, n-1, \\ \sigma^{-1}\zeta_n + \delta, & i = n, \end{cases}$$

where σ is the scaling parameter, δ is the shift parameter along the nth coordinate axis, and ζ_i ($i = 1, 2, \ldots, n$) are independent random variables uniformly distributed over [0, 1].

Despite the simplicity of the test, the minimum-norm element seems rather difficult to find. At least, the simple algorithm of [6], when used on the problem with a relatively moderate dimension of 100 variables and 100 simplex vertices, could satisfy the optimality conditions only to 10^{-2} accuracy after more than one million iterations.

The algorithm described in this paper has been implemented in the language Octave [7], which is a freely distributed matrix-vector software that is fairly convenient for such experiments: it took about 150 lines to code the algorithm. More details on the source code can be found in [8], which provides a detailed description of the algorithm's current version by using the literate programming technique.

Nevertheless, the program only was used to check the reliability and numerical efficiency of the algorithm in terms of the number of iterations required for deriving a solution. No special attempts were made to enhance its efficiency at this stage. For example, the most computationally intensive task—the projection onto the affine subspaces G_s in (7)—was performed by directly solving the system of necessary and (in this case also sufficient) optimality conditions

$$P_s u - \theta e = 0,$$

$$eu = 1.$$

The system was solved for u (which are the expansion coefficients in the projection of zero onto G_s in terms of the basis of this subspace) and for θ (which is dual to the normalization conditions for these coefficients). Here, $e = (1, \ldots, 1)$ and P_s is the Gram matrix of the basis vectors of G_s. The solution of the optimality conditions is reduced primarily to the inversion of P_s, which, in the simplest case, was run from a cold start. Of course, this is a major reserve for improving the efficiency of the algorithm. Specifically, preliminary experiments have shown that the computational (time) complexity of the algorithm can be reduced from $O(n^{3.5})$, as obtained in [3], to $O(n^2)$. However, this point requires further investigations because the effect of error accumulation arises in this case.

A typical behavior of the algorithm can be demonstrated by minimizing the distance to a simplex with 1999 vertices in a 2000-dimensional space, which is a fairly representative example. Figure 1 shows (on a logarithmic scale in the Oy axis) the convergence of $\|z^k\| - \|z^*\|$ to zero for various values of the scaling factor σ. It can well be seen that the convergence rate is exponential at most of the iterations and acceleration occurs at the beginning and end of the run. Moreover, the parameters of the exponential convergence rate exhibit remarkable stability with respect to scaling: the convergence rate remains nearly the same as σ^2 varies by three orders of magnitude.

Figure 2 presents a more detailed exposition of the results concerning the linear convergence parameters. The data set r1 describes the variation in the norm of z^k, which, according to (16), must decrease no more slowly than a geometric progression with the ratio ρ. Figure 2 (the data set r1) shows the values of $\|z^{k+1}\|/\|z^k\|$ for the basic set of iterations k. The data set r2 represents the values of $(\|z^{k+1}\| - \|z^*\|)/(\|z^k\| - \|z^*\|)$ for the same data. The acceleration of convergence in terms of the deviation from the optimal value can well be seen at the final stage of the computations. For such a dimension, a linear convergence parameter of about 0.95 is fairly reasonable or, at least, comparable with the analogous theoretical estimates obtained for ellipsoid methods and projection algorithms. A detailed analysis of the dependence of this factor on the dimension and parameters of the problem is still to be performed in the future.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 04-07-90287v) and program No. 17 of the Presidium of the Russian Academy of Sciences.

REFERENCES

1. H. H. Bauschke and J. M. Borwein, "On Projection Algorithms for Solving Convex Feasibility Problems," SIAM Rev. **38**, 367–426 (1996).
2. E. A. Nurminskii, "Parallel Method for Projection onto the Convex Hull of a Family of Sets," Izv. Vyssh. Uchebn. Zaved., Mat. **12** (499), 78–82 (2003).
3. E. A. Nurminskii, "Method of Sequential Projections for Solving the Least-Distance Problem for Simplexes," Elektron. Zh. Issledovano v Rossii **160**, 1732–1739 (2004); http://zhurnal.ape.relarn.ru/articles/2004/160.pdf.
4. L. M. Bregman, "Determination of a Common Point of Convex Sets by Sequential Projection," Dokl. Akad. Nauk SSSR **162**, 487–490 (1965).
5. L. G. Gurin, B. T. Polyak, and E. V. Raik, "Projection Methods for Finding a Common Point of Convex Sets," Zh. Vychisl. Mat. Mat. Fiz. **7**, 1211–1228 (1967).
6. V. F. Dem'yanov and V. N. Malozemov, *Introduction to Minimax* (Nauka, Moscow, 1972) [in Russian].
7. GNU Octave, http://www.octave.org.
8. E. A. Nurminskii, "The Method of Affine Subspaces for Searching for a Minimal-Length Vector in a Simplex," http://www.iacp.dvo.ru/lab_11/e-prints.