No. **621**

An Algorithm for Solving the Minimum Norm Point Problem over the Intersection of a Polytope and an Affine Set

by

Satoru Fujishige, Xiaojun Liu and Xiaodong Zhang

March 1995

An Algorithm for Solving the Minimum Norm Point Problem over the Intersection of a Polytope and an Affine Set

Satoru Fujishige, Xiaojun Liu and Xiaodong Zhang*

Institute of Socio-Economic Planning
University of Tsukuba
Tsukuba, Ibaraki 305, Japan

March 1995

Abstract

In this paper we propose an efficient algorithm for finding the minimum-norm point in the intersection of a polytope and an affine set in n-dimensional Euclidean space, where the polytope is given as the convex hull of finitely many points and the affine set as the intersection of k hyperplanes ($k \geq 1$).

Our algorithm given in this paper solves the problem by directly using the points and the hyperplanes rather than using the points of the intersection of the polytope and the affine set. When the number k of the hyperplanes expressing the affine set is equal to one, we can easily avoid the degeneracy, but this is not the case for $k \geq 2$. We give a subprocedure for treating the degenerate case. The subprocedure is interesting in its own right. We also show the practical efficiency of our algorithm by computational experiments.

1. Introduction

The problem of finding the minimum-norm point in the intersection of a convex polyhedron and a hyperplane has extensively been studied in some literatures. S. Fu-

*Current address: Computer Science Program, University of Texas at Dallas, Richardson, Texas 75083-0688 USA

jishige, H. Sato and P. Zhan [5] have developed a finite algorithm for finding the minimum-norm point in the intersection of a polytope and a hyperplane and K. Sekitani, J. M. Shi, Y. Yamamoto and K. Yamasaki [10] have also devised an algorithm in the intersection of a polytope and seveal hyperplanes.

The minimum-norm point problem has applications to a lot of practical problems such as the problems of mixing [7], positive linear approximation [12], the optimal control [1], the nondifferentiable function minimization [8], the submodular function minimization [2] and the portfolio selection [11].

In this paper we consider the problem of solving the minimum-norm point over the intersection of a polytope and an affine set, where the polytope is given as a convex hull of a finite point set $P = \{p_i \mid i \in I\}$ and the affine set is the intersection of k hyperplanes $H_i = \{x \mid (a_i, x) = l_i\}$ $(i = 1, \ldots, k)$ in the n-dimensional Euclidean space $\mathbf{R}^n$. The problem is described as follows:

$$
\begin{aligned}
\textbf{(P)} \quad \text{Minimize} \quad & ||x|| \\
\text{subject to} \quad & x = \sum_{i \in I} w_i p_i, \\
& \sum_{i \in I} w_i = 1, \quad w_i \geq 0 \ (i \in I), \\
& x \in V_k \, (= H_1 \cap H_2 \cap \cdots \cap H_k), \qquad (1.1)
\end{aligned}
$$

where $|| \cdot ||$ denotes the Euclidean norm.

An application of this model is the portfolio optimization treated by H. Takehara [11] and H. Konno and K. Suzuki [6]. Applications of this problem are found in fields similar to those of the minimum-norm point problem of a polytope. A parametric binary search algorithm is given in K. Sekitani, J. M. Shi, Y. Yamamoto and K. Yamasaki [10].

Our efficient algorithm is based on Wolfe's algorithm [13] and S. Fujishige, H. Sato and P. Zhan's algorithm[5]. We can solve the problem by using only original points $p_i \in P$ and the hyperplanes $H_i \, (i = 1, \cdots k)$. An interior point algorithm for solving the present problem is also given by H. Takehara [11] for k=1.

In Section 2 we give some preliminary results. In Section 3 we propose an algorithm for solving Problem (**P**) and prove the validity of the algorithm under the nondegeneracy assumption. In Section 4 we give an algorithm for avoiding the degeneracy. In Section 5 we also show how to solve relevant systems of equations. Finally, we carry out computational experiments to show the behavior of the proposed algorithm in Section 6.

2. Problem Description and Preliminaries

2.1. Definitions and Problem (P)

We consider an n-dimensional Euclidean space $\mathbf{R}^n$, where the Euclidean norm is denoted by $||\cdot||$. The inner product of two vectors a and x in $\mathbf{R}^n$ is expressed as (a,x) in the following. We denote the *convex hull* of a finite point set P in $\mathbf{R}^n$ by $\mathrm{C}(P)$ and the *affine hull* of P by $\mathrm{A}(P)$.

Let $H_1, H_2, \cdots, H_k$ be k hyperplanes expressed as $H_1 : (a_1, x) = l_1, H_2 : (a_2, x) = l_2, \cdots, H_k : (a_k, x) = l_k$. Denote by V_k the affine set given by the intersection $H_1 \cap H_2 \cap \cdots \cap H_k$. For a finite point set $P = \{p_i \mid i \in I\}$ in $\mathbf{R}^n$ consider the problem of finding the minimum-norm point in $\mathrm{C}(P) \cap V_k$, i.e.,

$$
\begin{array}{lll}
\textbf{(P)} & \text{Minimize} & ||x|| \\
 & \text{subject to} & x = \sum_{i\in I} w_i p_i, \\
 & & \sum_{i\in I} w_i = 1, \quad w_i \geq 0 \ \ (i \in I), \\
 & & x \in V_k (= H_1 \cap H_2 \cap \cdots \cap H_k).
\end{array} \tag{2.1}
$$

Let q be the minimum-norm point in V_k. Then we can easily see that Problem **(P)** described above is equivalent to the problem of finding a point $x \in \mathrm{C}(P) \cap V_k$ that minimizes $||x - q||$. Therefore, translating P and H_j $(j = 1, 2, \cdots, k)$ so that q becomes the origin $\mathbf{0}$, we assume without loss of generality throughout this paper that $l_j = 0$ $(j = 1, 2, \cdots, k)$ for the hyperplanes H_j $(j = 1, 2, \cdots, k)$. We also assume that the normal vectors a_j of H_j $(j = 1, 2, \cdots, k)$ are linearly independent.

For $Q \subseteq P$ and $x_0 \in \mathbf{R}^n$ we call (Q, x_0) a *corral* if Q is affinely independent and x_0 is the minimum-norm point of $\mathrm{A}(Q) \cap V_k$ and lies in the relative interior of $\mathrm{C}(Q)$. For any hyperplane $H : (a, x) = l$ with $l \neq 0$ we denote by H_+ the open halfspace determined by H and containing the origin $\mathbf{0}$.

We have an optimality condition for Problem **(P)** described as follows.

Theorem 2.1: *A feasible point $x_0 \in \mathrm{C}(P) \cap V_k$ is an optimal solution of Problem* **(P)** *if and only if*

$$(x_0, x) \geq ||x_0||^2 \quad (x \in \mathrm{C}(P) \cap V_k). \tag{2.2}$$

□

We also give the following optimality condition for Problem **(P)**, which follows from Theorem 2.1. We shall use the optimality condition given by the following theorem to construct an algorithm for solving Problem **(P)** in the sequel.

Theorem 2.2: *A feasible point* $x_0 \in \mathrm{C}(P) \cap V_k$ *is an optimal solution of Problem* **(P)** *if there exist* k *real numbers* $\alpha_1, \cdots, \alpha_k$ *such that*

$$(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_i) \geq ||x_0||^2 \quad (i \in I). \tag{2.3}$$

(Proof) Choose an arbitrary $\bar{x} \in \mathrm{C}(P) \cap V_k$. We can express $\bar{x}$ as $\bar{x} = \sum_{i \in I} w_i p_i$ with $w_i \geq 0$ $(i \in I)$ and $\sum_{i \in I} w_i = 1$. Since $\bar{x} \in \mathrm{C}(P) \cap V_k$, we have $(a_j, \bar{x}) = 0$ $(j = 1, \cdots, k)$. Therefore, using (2.3), we have

$$\begin{aligned}(x_0, \bar{x}) &= (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, \bar{x}) \\ &= (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, \sum_{i \in I} w_i p_i) \\ &\geq \sum_{i \in I} w_i ||x_0||^2 \\ &= ||x_0||^2. \end{aligned} \tag{2.4}$$

We see from Theorem 2.1 that x_0 is an optimal solution for Problem **(P)**. □

2.2. Problem ($\mathbf{P_0}$) and Its Simplification

In order to find an optimal solution for Problem **(P)** we need an initial corral (Q, x_0) to start our algorithm with. For this purpose we consider the following problem:

$$(\mathbf{P_0}) \quad \begin{array}{l} \text{Minimize } d(x, V_k) \\ \text{subject to } \ x = \sum_{i \in I} w_i p_i, \quad \sum_{i \in I} w_i = 1, \ w_i \geq 0 \quad (i \in I), \end{array} \tag{2.5}$$

where $d(x, V_k)$ is the Euclidean distance between x and V_k $(= H_1 \cap \cdots \cap H_k)$.
We also have an optimality condition for Problem ($\mathbf{P_0}$) described as follows.

Theorem 2.3: $x_0 \in \mathrm{C}(P)$ *is an optimal solution of Problem* ($\mathbf{P_0}$) *if and only if*

$$(x_0 - v_0, x - v_0) \geq ||x_0 - v_0||^2 \quad (x \in \mathrm{C}(P)) \tag{2.6}$$

(*or equivalently,* $(x_0 - v_0, p_i - v_0) \geq ||x_0 - v_0||^2$ *for any* $p_i \in P$), *where* v_0 *is the orthogonal projection of* x_0 *into* V_k.
(Proof) If $\mathrm{C}(P) \cap V_k \neq \emptyset$, then the present theorem holds. So, we assume that $\mathrm{C}(P) \cap V_k = \emptyset$ and $v_0 \neq x_0$. Suppose that $(x_0 - v_0, x - v_0) \geq ||x_0 - v_0||^2$ for any $x \in \mathrm{C}(P)$. Consider a hyperplane $H : (x_0 - v_0, x - v_0) = 0$. Since $v_0 \in V_k$ is the orthogonal projection of x_0 into V_k, we have $(x_0 - v_0, v - v_0) = 0$ for any $v \in V_k$. Hence, we have $V_k \subseteq H$ and for any $\bar{x} \in \mathrm{C}(P)$

$$d(\bar{x}, V_k) \geq d(\bar{x}, H) = \frac{(x_0 - v_0, \bar{x} - v_0)}{||x_0 - v_0||} \geq \frac{||x_0 - v_0||^2}{||x_0 - v_0||} = ||x_0 - v_0|| = d(x_0, V_k). \tag{2.7}$$

Therefore, x_0 is an optimal solution of Problem ($\mathbf{P_0}$). Conversely, suppose that x_0 is an optimal solution of Problem ($\mathbf{P_0}$), i.e., $d(\mathrm{C}(P), V_k) = ||x_0 - v_0||$. Define a hyperplane $H' : (x_0 - v_0, x - v_0) = ||x_0 - v_0||^2$. We see from the convexity that H' is a separating hyperplane between $\mathrm{C}(P)$ and V_k. It follows that for any $x \in \mathrm{C}(P)$ we have $(x_0 - v_0, x - v_0) \geq ||x_0 - v_0||^2$. □

According to Theorem 2.3 for Problem ($\mathbf{P_0}$), any $p_i \in P$ is decomposed into p'_i and p''_i in such a way that p'_i is parallel to and p''_i is orthogonal to V_k. In the same way, for $x \in \mathrm{C}(P)$, x is decomposed into x' and x'', where x' is parallel to and x'' is orthogonal to V_k. We have that $x \in \mathrm{C}(P)$ implies that $x'' \in \mathrm{C}(P'')$ where $P'' = \{p''_i \mid i \in I\}$. Therefore, Problem ($\mathbf{P_0}$) is equivalent to the problem described as follows:

$$(\mathbf{P''_0}) \quad \text{Minimize} \quad ||y|| \quad \text{subject to} \quad y \in \mathrm{C}(P''). \tag{2.8}$$

Problem ($\mathbf{P''_0}$) is an ordinary minimum-norm point problem, so that we can use Wolfe's algorithm [13] for solving it. We can then find an $x \in \mathrm{C}(P) \cap V_k$ and an initial corral (Q, x_0).

If the minimum-norm point y in $\mathrm{C}(P'')$ is not equal to $\mathbf{0}$, then Problem ($\mathbf{P}$) is not feasible. Otherewise, let the minimum-norm point $\mathbf{0}$ be expressed as

$$\mathbf{0} = \sum_{i \in I} w_i p''_i, \tag{2.9}$$

where $\sum_{i \in I} w_i = 1$, $w_i \geq 0$ $(i \in I)$. We have a feasible solution

$$x_0 = \sum_{i \in I} w_i p_i \in \mathrm{C}(P) \cap V_k \tag{2.10}$$

for Problem ($\mathbf{P}$). If necessary, we remove some p_i from $Q = \{p_i \mid w_i > 0, i \in I\}$ to make the obtained (Q, x_0) a corral, by a procedure similar to Steps 2 and 3 of the main algorithm to be shown in the next section.

3. An Algorithm and Its Validity

In this section we describe an algorithm and show its validity under the nondegeneracy assumption.

3.1. An Algorithm

We must first solve Problem ($\mathbf{P}_0''$) to find a feasible point $x_0 \in \mathrm{C}(P) \cap V_k$ and an initial corral (Q, x_0). Then, we update the obtained corral (Q, x_0) repeatedly by determining the values of the parameters $\alpha_1, \cdots, \alpha_k$ so that the hyperplane

$$\hat{H} : (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, x) = ||x_0||^2 \tag{3.1}$$

contains all the points $p_i \in Q$. If $(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_i) \geq ||x_0||^2$ for all $i \in I$, then x_0 is the optimal solution due to Theorem 2.2.

Now, an algorithm for solving Problem **(P)** is given as follows. When a current corral becomes degenerate, we carry out Procedure Degenerate. Procedure Degenerate and relevant concepts will be given in the next section.

An Algorithm
Input: A finite point set $P = \{p_i \mid i \in I\}$ in $\mathbf{R}^n$ and k hyperplanes $H_j : (a_j, x) = 0$ $(j = 1, \cdots, k)$.
Output: The minimum-norm point x_0 in $\mathrm{C}(P) \cap V_k$.
Step 0: Solve Problem ($\mathbf{P}_0''$) to find an initial corral (Q, x_0).
Step 1: If $x_0 = \mathbf{0}$, then stop. Otherwise determine the values of parameters $\alpha_1, \cdots, \alpha_k$ so that the hyperplane $\hat{H} : (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, x) = ||x_0||^2$ contains Q. If $\hat{H}$ is not uniquely determined, then perform Procedure Degenerate and go to the beginning of Step 1. If we have

$$(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_i) \geq ||x_0||^2 \quad (i \in I), \tag{3.2}$$

then stop (the current x_0 is the optimal solution of Problem **(P)**). Otherwise choose $p_{i*} \in P$ such that

$$(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_{i*}) = \min\{(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_i) \mid i \in I\}. \tag{3.3}$$

Put $Q := Q \cup \{p_{i*}\}$.
Step 2: Let y be the minimum-norm point in $\mathrm{A}(Q) \cap V_k$. If y is in the relative interior of $\mathrm{C}(Q)$, put $x_0 := y$ and go to Step 1.
Step 3: Let z be the point in the line segment $\mathrm{C}(Q) \cap \overline{x_0 y}$ nearest to y. Delete from Q the points p_i not in the minimal face of $\mathrm{C}(Q)$ on which z lies. Put $x_0 := z$ and go to Step 2.
(End)

We call the cycle of Steps 2 and 3 in the above algorithm a *minor cycle* and the cycle of Step 1 together with a possible repetition of the minor cycle a *major cycle*.

3.2. Validity of the Algorithm

Now, we show the validity of the main algorithm under the nondegeneracy assumption, i.e., when Procedure Degenerate is not invoked during the execution of the algorithm.

First, we have the following.

Lemma 3.1: *In Step 1, if $x_0 \neq \mathbf{0}$, then there exists a hyperplane $\hat{H}$: $(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, x) = ||x_0||^2$ that contains all points in Q.*

(Proof) Suppose that $Q = \{p_1, \cdots, p_l\}$. Q lies on $\hat{H}$ if and only if for each p_i in Q we have

$$(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_i) = ||x_0||^2. \tag{3.4}$$

That is,

$$\begin{aligned}
(a_1, p_1)\alpha_1 + (a_2, p_1)\alpha_2 + \cdots + (a_k, p_1)\alpha_k &= ||x_0||^2 - (x_0, p_1),\\
(a_1, p_2)\alpha_1 + (a_2, p_2)\alpha_2 + \cdots + (a_k, p_2)\alpha_k &= ||x_0||^2 - (x_0, p_2),\\
\vdots \quad \vdots \quad \vdots \\
(a_1, p_l)\alpha_1 + (a_2, p_l)\alpha_2 + \cdots + (a_k, p_l)\alpha_k &= ||x_0||^2 - (x_0, p_l).
\end{aligned} \tag{3.5}$$

Let A be the coefficient matrix of the system (3.5) of equations, i.e.,

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_l \end{pmatrix}_{l\times k} = \begin{pmatrix} (a_1, p_1) & (a_2, p_1) & \cdots & (a_k, p_1) \\ (a_1, p_2) & (a_2, p_2) & \cdots & (a_k, p_2) \\ \vdots & \vdots & & \vdots \\ (a_1, p_l) & (a_2, p_l) & \cdots & (a_k, p_l) \end{pmatrix}_{l\times k}. \tag{3.6}$$

Suppose that $\text{rank}(A) = r\ (\leq k)$. Without loss of generality we assume that the leading $r \times r$ submatrix of A is of rank r. Therefore, there exists k real numbers $\alpha_1, \cdots, \alpha_k$ that satisfy the first r equations of (3.5). Since an arbitrary jth $(r < j \leq l)$ row A_j of A can be expressed by the first r rows, there exist r real numbers $t_{1j}, \cdots, t_{rj}$ such that

$$A_j = \sum_{i=1}^{r} t_{ij} A_i \quad (r < j \leq l), \tag{3.7}$$

i.e.,

$$(a_m, p_j) = \sum_{i=1}^{r} t_{ij} (a_m, p_i) \quad (m = 1, \cdots, k) \tag{3.8}$$

or

$$(a_m, p_j - \sum_{i=1}^{r} t_{ij} p_i) = 0 \quad (m = 1, \cdots, k). \tag{3.9}$$

This implies that $p_j - \sum_{i=1}^{r} t_{ij} p_i \in V_k$ $(r < j \leq l)$. Since $x_0 \in \mathrm{A}(Q) \cap V_k$,

$$\sum_{i=1}^{r} t_{ij} x_0 + p_j - \sum_{i=1}^{r} t_{ij} p_i \in \mathrm{A}(Q) \cap V_k \quad (r < j \leq l). \tag{3.10}$$

Since x_0 is the minimum-norm point in $A(Q) \cap V_k$, we have

$$\begin{aligned} ||x_0||^2 &= (x_0, \sum_{i=1}^{r} t_{ij} x_0 + p_j - \sum_{i=1}^{r} t_{ij} p_i) \\ &= (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, \sum_{i=1}^{r} t_{ij} x_0 + p_j - \sum_{i=1}^{r} t_{ij} p_i) \\ &= \sum_{i=1}^{r} t_{ij} ||x_0||^2 + (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_j) \\ &\quad - \sum_{i=1}^{r} t_{ij} (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_i) \\ &= (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, p_j) \end{aligned} \tag{3.11}$$

for any j with $r < j \leq l$. Therefore, (3.5) holds for these $\alpha_1, \cdots, \alpha_k$. □

Note that by the proof of Lemma 3.1 we have actually proved the following.

Lemma 3.2: *For any $Q \subseteq P$, if x_0 is the minimum-norm point of $\mathrm{A}(Q) \cap V_k$ such that $x_0 \neq \mathbf{0}$, then there exists a hyperplane $\hat{H} : (x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, x) = ||x_0||^2$ that contains Q.* □

We also have the following.

Lemma 3.3: *When we choose $p_{i^*} \in P$ satisfying* (3.3) *in Step 1, we have*

$$\mathrm{A}(Q \cup \{p_{i^*}\}) \cap V_k \supset \mathrm{A}(Q) \cap V_k \quad \text{(strict inclusion)}. \tag{3.12}$$

(Proof) If $\mathrm{A}(Q \cup \{p_{i^*}\}) \cap V_k = \mathrm{A}(Q) \cap V_k$, then from Lemma 3.2 there exists a hyperplane H' of the form (3.1) that contains $Q \cup \{p_{i^*}\}$. Since H' contains Q, it follows from the uniqueness of the hyperplane $\hat{H}$ in Step 1 that we have $H' = \hat{H}$, which implies $p_{i^*} \in \hat{H}$, a contradiction. □

Lemma 3.4: *In Step 1 $Q \cup \{p_{i^*}\}$ is affinely independent.*
(Proof) In Step 1 we choose $p_{i^*} \notin \hat{H}$ that satisfies (3.3) and then obtain a new $Q := Q \cup \{p_{i^*}\}$, from which the present lemma follows. □

Now, we will show that $||x_0||$ strictly decreases every time x_0 is changed. First, we show

Lemma 3.5: *In Step 1, if we choose $p_{i^*} \in P$ satisfying $(x_0+\alpha_1 a_1+\cdots+\alpha_k a_k, p_{i^*}) < ||x_0||^2$, then for the minimum-norm point y in $A(Q\cup\{p_{i^*}\})\cap V_k$ we have $||y|| < ||x_0||$ for the current x_0.*
(Proof) From Lemma 3.3 we have $A(Q\cup\{p_{i^*}\})\cap V_k \supset A(Q)\cap V_k$. Since $p_{i^*} \in \hat{H}_+$, there exists a point $q \in A(Q\cup\{p_{i^*}\})\cap V_k$ such that $q \in \hat{H}_+$. Therefore, there also exists a point $q' \in A(Q\cup\{p_{i^*}\})\cap V_k$ such that $||q'|| < ||x_0||$, and hence, $||y|| < ||x_0||$. □

Unfortunately, if $\hat{H}$ is not uniquely determined by Q in Step 1, then the above way of selecting p_{i^*} may not ensure $||y|| < ||x_0||$. See the argument for the degenerate case given in the next section.

In Step 2 we find the minimum-norm point y in $A(Q\cup\{p_{i^*}\})\cap V_k$ by solving a system of linear equations.

Lemma 3.6: *Under the same assumption as in Lemma 3.5 y is uniquely expressed as $\sum_{p_i\in Q} w_i p_i + w_{i^*}p_{i^*}$ such that $w_{i^*} > 0$ and $\sum_{p_i\in Q} w_i + w_{i^*} = 1$.*
(Proof) The uniqueness follows from the affine independence of $Q\cup\{p_{i^*}\}$ due to Lemma 3.4. Also, since $y \in \hat{H}_+$, we have

$$\begin{aligned} ||x_0||^2 &> (x_0+\alpha_1 a_1+\cdots+\alpha_k a_k, y) \\ &= \sum_{p_i\in Q} w_i||x_0||^2 + w_{i^*}(x_0+\alpha_1 a_1+\cdots+\alpha_k a_k, p_{i^*}) \\ &= ||x_0||^2 + w_{i^*}\{(x_0+\alpha_1 a_1+\cdots+\alpha_k a_k, p_{i^*}) - ||x_0||^2\}. \end{aligned} \tag{3.13}$$

Since p_{i^*} also belongs to $\hat{H}_+$, we have from (3.13) $w_{i^*} > 0$. □

Lemma 3.7: *When we come to Step 2 from Step 3, we have $||y|| \le ||x_0||$.*
(Proof) Since $x_0 \in C(Q)\cap V_k \subseteq A(Q)\cap V_k$ and y is the minimum-norm point in $A(Q)\cap V_k$, we have the present lemma. □

Lemma 3.8: *In Step 3 we have $||z|| < ||x_0||$.*
(Proof) Let $z = \lambda y + (1-\lambda)x_0$. From Lemma 3.6 we have $\lambda \in (0,1]$. Hence,

$$\begin{aligned} ||z|| &= ||\lambda y + (1-\lambda)x_0|| \\ &\le \lambda||y|| + (1-\lambda)||x_0|| \\ &< \lambda||x_0|| + (1-\lambda)||x_0|| \\ &= ||x_0|| \end{aligned} \tag{3.14}$$

□

Lemma 3.9: *$||x_0||$ strictly decreases after each major cycle.*
(Proof) The present lemma follows from Lemmas 3.6, 3.7 and 3.8. □

Now, we show that the algorithm terminates in finitely many steps under the nondegeneracy assumption that we do not invoke Procedure Degenerate during the execution of the algorithm.

First, we have

Lemma 3.10: *During the execution of the algorithm the same Q does not appear more than once.*

(Proof) Since Q uniquely determines x_0, the present lemma follows from Lemma 3.9. □

Lemma 3.11: *Step 1 is executed finitely many times.*

(Proof) We have a new Q every time Step 1 is executed and the number of possible simplices Q is finite since each Q is a subset of the finite set P. □

Lemma 3.12: *The minor cycle in a major cycle is consecutively executed finitely many times.*

(Proof) In the minor cycle some points, at least one, in Q are removed in each minor cycle. So, the number of the consecutive minor cycles is at most $|Q| - 1$. □

Theorem 3.13: *Under the nondegeneracy assumption the algorithm solves Problem* **(P)** *in finitely many steps.*

(Proof) The theorem follows from Lemma 3.11, Lemma 3.12 and Theorem 2.2. □

4. The Degenerate Case

In this section we consider the degenerate case where in Step 1 of the main algorithm the hyperplane $\hat{H}$ is not uniquely determined. In this case $||x_0||$ may not strictly decrease without any additional procedure and the algorithm may not terminate in finitely many steps. We will propose Procedure Degenerate to overcome this difficulty.

4.1. Definitions

First, we introduce some definitions to be employed in this section.

For a corral (Q_0, x_0) and a point $p \in P$ denote by p' the vector in $\mathbf{R}^n$ such that the vector $p - x_0$ is decomposed into $p - x_0 = (p - p') + (p' - x_0)$ in such a way that $p - p'$ is parallel to and $p' - x_0$ is orthogonal to $H(x_0) \cap V_k$. Here, for any $\bar{x} \in \mathbf{R}^n$ $H(\bar{x})$ is the hyperplane containing $\bar{x}$ and having a normal vector $\bar{x}$. Note that p' depends only on x_0 of the current corral (Q_0, x_0). For $Q_0 \subseteq Q \subseteq P$ we define:

(I) $Q'_0 = \{p' \mid p \in Q_0\}, \quad Q' = \{p' \mid p \in Q\}.$

(II) Cone$(Q' - Q'_0)$: the cone generated by points in $Q' - Q'_0$.

(III) Any $y \in \mathrm{A}(Q' \cup \{x_0\})$ can be uniquely decomposed into y_1 and y_2 as $y = y_1 + y_2$ with $y_1 = w_0 x_0 + \sum_{p'_i \in Q'_0} w_i p'_i$ and $y_2 = \sum_{p'_i \in Q' - Q'_0} w_i p'_i$, where $w_0 + \sum_{p'_i \in Q'} w_i = 1$.

We call (Q, Q_0, x_0) a *degenerate corral* if it satisfies the following conditions (d.1)∼(d.4). It should be noted that this is a kind of conical version of a "corral". We will show that (d.1)∼(d.4) is equivalent to (d.1)∼(d.3) and (d.5) in Lemmas 4.7 and 4.9.

(d.1) (Q_0, x_0) is a corral with $x_0 \neq \mathbf{0}$ and a hyperplane of the following form containing Q_0 is not uniquely determined:

$$(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, x) = ||x_0||^2. \tag{4.1}$$

(d.2) $Q \supseteq Q_0$ is affinely independent and $|Q| \leq n$.

(d.3) When $Q \neq Q_0$,

(d.3a) $\dim(\mathrm{A}(Q' \cup \{x_0\})) = \dim(\mathrm{A}(Q'_0 \cup \{x_0\})) + |Q' - Q'_0|$,

(d.3b) for the minimum-norm point $\hat{x}$ in $\mathrm{A}(Q' \cup \{x_0\})$, decomposing $\hat{x}$ into $\hat{x} = \hat{x}_1 + \hat{x}_2$ as in (III) described above, $\hat{x}_2$ lies in the relative interior of Cone$(Q' - Q'_0)$.

(d.4) x_0 is the minimum-norm point of $\mathrm{A}(Q) \cap V_k$.

(d.5) For the minimum-norm point $\hat{x}$ in $\mathrm{A}(Q' \cup \{x_0\})$ we have $\hat{x} \neq \mathbf{0}$.

For (Q, Q_0, x_0) such that (d.1) and (d.2) are satisfied and that there exists a hyperplane of form (4.1) containing Q, define

$\hat{H}(Q)$: the hyperplane of the form (4.1) containing Q, where $\alpha_1, \cdots, \alpha_k$ should be determined so as to maximize the distance between the hyperplane and the origin.

4.2. Basic Properties in the Degenerate Case

We prove some lemmas on basic properties of the main algorithm in the degenerate case.

Lemma 4.1: *Let x_0 be a vector in V_k. Then any hyperplane that contains $H(x_0) \cap V_k$ but not V_k is of form* (4.1).
(Proof) Suppose that a hyperplane H_0 with a normal vector a_0 contains $H(x_0) \cap V_k$ but not V_k. H_0 is of the form $(a_0, x) = l$. Suppose that $a_0 = a_0' + a_0''$, where $a_0' \in V_k$ and a_0'' is orthogonal to V_k. Then, a_0'' is expressed as a linear combination of $a_1, \cdots, a_k$, i.e.,

$$a_0'' = \alpha_1' a_1 + \cdots + \alpha_k' a_k, \tag{4.2}$$

where $a_1, \cdots, a_k$ are the normal vectors of $H_1, \cdots, H_k$, respectively. Since H_0 does not contain V_k, we have $a_0' \neq \mathbf{0}$. By the assumption, $H_0 \supsetneq H(x_0) \cap V_k$ and $V_k \supsetneq H(x_0) \cap V_k$. Hence,

$$H_0 \cap \ V_k \supsetneq H(x_0) \cap V_k \neq \emptyset. \tag{4.3}$$

Since V_k is a linear subspace and $H_0 \cap V_k$ and $H(x_0) \cap V_k$ are both hyperplanes in V_k, we have from (4.3)

$$H_0 \cap V_k = H(x_0) \cap V_k. \tag{4.4}$$

For any $x, y \in H_0 \cap V_k$, we have $(a_0, x) = (a_0, y)$ since a_0 is normal to H_0 and $x, y \in H_0$. From the definition of a_0'', we have $(a_0'', x) = (a_0'', y) = 0$. Therefore, $(a_0', x) = (a_0', y)$. It follows that a_0' is normal to $H_0 \cap V_k = H(x_0) \cap V_k$ since $a_0' \in V_k$. Since a_0' and x_0 belong to V_k and are both normal to $H(x_0) \cap V_k$ which is a hyperplane in the subspace V_k, we have for some scalar $\alpha_0 \neq 0$

$$a_0' = \alpha_0 x_0. \tag{4.5}$$

Hence, from a general form $(a_0, x) = l$ of H_0, we get $(a_0' + a_0'', x) = l$, i.e., $(\alpha_0 x_0 + \alpha_1' a_1 + \cdots + \alpha_k' a_k, x) = l$. Since $\alpha_0 \neq 0$, H_0 is also expressed as

$$(x_0 + \alpha_1 x_1 + \cdots + \alpha_k a_k, x) = l'. \tag{4.6}$$

Moreover, since H_0 contains x_0, we have

$$l' = ||x_0||^2 \tag{4.7}$$

□

Lemma 4.2: *Suppose that (Q_0, x_0) is a corral. Then for any Q with $Q_0 \subseteq Q \subseteq P$ a hyperplane of the form*

$$(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, x) = ||x_0||^2 \tag{4.8}$$

contains Q if and only if it contains Q'.
(Proof) Let H_0 be a hyperplane of the form

$$(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, x) = ||x_0||^2. \tag{4.9}$$

For any $\bar{x} \in H(x_0) \cap V_k$, we have $(x_0, \bar{x}) = ||x_0||^2$ and $(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, \bar{x}) = ||x_0||^2$, since $\bar{x} \in V_k$. Hence, $\bar{x} \in H_0$. Consequently, $H_0 \supseteq H(x_0) \cap V_k$. Since $p - p'$ $(p \in Q)$ is parallel to $H(x_0) \cap V_k$, we have $x_0 + (p - p') \in H(x_0) \cap V_k$, so that $x_0 + (p - p') \in H_0$.

Now, if $Q' \subseteq H_0$, then for any $p \in Q$ we have $p = (x_0 + (p - p')) + p' - x_0 \in H_0$, i.e., $Q \subseteq H_0$. Conversely, if $Q \subseteq H_0$, then for any $p \in Q$ we have $p' = p + x_0 - (x_0 + (p - p')) \in H_0$, i.e., $Q' \subseteq H_0$. □

Lemma 4.3: *Suppose that* (Q, Q_0, x_0) *satisfies* (d.1), (d.2) *and* (d.5), *and that* $\hat{x}$ *is the minimum-norm point in* $\mathrm{A}(Q' \cup \{x_0\})$. *Then* $H(\hat{x})$ *contains* $H(x_0) \cap V_k$.
(Proof) Let A_0 be the affine set generated by x_0 and p' $(p \in Q)$. Since $\hat{x}$ is the minimum-norm point of A_0 and $\hat{x} \neq \mathbf{0}$, $H(\hat{x})$ contains A_0, i.e., $A_0 \subseteq H(\hat{x})$. Since x_0 is the minimum-norm point of $H(x_0)$ and $x_0 \in V_k$, x_0 is also the minimum-norm point of $H(x_0) \cap V_k$. Let A_1 be the affine set generated by x_0, p' $(p \in Q)$ and the vectors in $H(x_0) \cap V_k$. We show that $\hat{x}$ is also the minimum-norm point of the affine set A_1. For any $y \in A_1$ y is expressed as

$$y = x_0 + (z - x_0) + (w - x_0) \tag{4.10}$$

for some $z \in \mathrm{A}(Q' \cup \{x_0\})$ and $w \in H(x_0) \cap V_k$. Since both x_0 and $z - x_0$ are orthogonal to $w - x_0$, we have

$$\begin{aligned} ||y||^2 &= ||x_0 + (z - x_0) + (w - x_0)||^2 \\ &= ||x_0 + (z - x_0)||^2 + ||w - x_0||^2 \\ &\geq ||z||^2 \geq ||\hat{x}||^2. \end{aligned} \tag{4.11}$$

Hence, $\hat{x}$ is also the minimum-norm point of A_1. Therefore, $H(\hat{x})$ also contains $H(x_0) \cap V_k \subseteq A_1$. □

Lemma 4.4: *Suppose that* (Q, Q_0, x_0) *satisfies* (d.1), (d.2) *and* (d.5), *and that* $\hat{x}$ *is the minimum-norm point in* $\mathrm{A}(Q' \cup \{x_0\})$. *Then* $H(\hat{x}) = \hat{H}(Q)$.
(Proof) The present lemma follows from Lemmas 4.1~4.3. □

Lemma 4.5: *Suppose that* (Q, Q_0, x_0) *satisfies* (d.1)~(d.3) *and that* $\hat{x}$ *is the minimum-norn point in* $A(Q' \cup \{x_0\})$. *Then,* $\hat{x} \neq \mathbf{0}$ *if and only if the hyperplane* H_0 *of the form* (4.1) *containing* Q *exists.*
(Proof) If such a hyperplane H_0 exists, then we have $\mathbf{0} \notin H_0$ since $x_0 \neq \mathbf{0}$. From Lemma 4.2 $\hat{x}$ lies on H_0 and hence, $\hat{x} \neq \mathbf{0}$. Conversely, if $\hat{x} \neq \mathbf{0}$, from Lemma 4.3 $H(\hat{x})$ contains $H(x_0) \cap V_k$ and Q'. If $H(\hat{x})$ contains V_k, then we have $\mathbf{0} \in H(\hat{x})$, i.e., $\hat{x} = \mathbf{0}$, a contradiction. Hence, $H(\hat{x})$ does not contain V_k, and from Lemma 4.1, $H(\hat{x})$ is a hyperplane of the form (4.1) containing Q. □

Lemma 4.6: *For a degenerate corral (Q, Q_0, x_0) there exists a hyperplane of the form*

$$(x_0 + \alpha_1 a_1 + \cdots + \alpha_k a_k, x) = ||x_0||^2 \tag{4.12}$$

that contains $Q \cup Q'$.
(Proof) This lemma follows from Lemmas 3.2 and 4.2. □

Lemma 4.7: *Suppose that (Q, Q_0, x_0) is a degenerate corral and that $\hat{x}$ is the minimum-norn point in $A(Q' \cup \{x_0\})$. Then, $\hat{x} \neq \mathbf{0}$.*
(Proof) This follows from Lemma 4.5 and 4.6. □

Lemma 4.8: *For (Q, Q_0, x_0) satisfying* (d.3a) *and for any y in $A(Q' \cup \{x_0\})$ expressed as*

$$y = w_0 x_0 + \sum_{p'_i \in Q'_0} w_i p'_i + \sum_{p'_i \in Q' - Q'_0} w_i p'_i, \tag{4.13}$$

$$w_0 + \sum_{p'_i \in Q'_0} w_i + \sum_{p'_i \in Q' - Q'_0} w_i = 1, \tag{4.14}$$

the coefficients w_i for $p'_i \in Q' - Q'_0$ are uniquely determined.
(Proof) Suppose, on the contrary, that there is another different expression of y given by

$$y = \tilde{w}_0 x_0 + \sum_{p'_i \in Q'_0} \tilde{w}_i p'_i + \sum_{p'_i \in Q' - Q'_0} \tilde{w}_i p'_i, \tag{4.15}$$

$$\tilde{w}_0 + \sum_{p'_i \in Q'_0} \tilde{w}_i + \sum_{p'_i \in Q' - Q'_0} \tilde{w}_i = 1 \tag{4.16}$$

such that $\tilde{w}_i \neq w_i$ for some $p_i' \in Q' - Q_0'$. Then from (4.13) and (4.15) we have

$$\mathbf{0} = (w_0 - \tilde{w}_0) x_0 + \sum_{p'_i \in Q'_0} (w_i - \tilde{w}_i) p'_i + \sum_{p'_i \in Q' - Q'_0} (w_i - \tilde{w}_i) p'_i. \tag{4.17}$$

Choose any $p'_j \, (\in Q' - Q'_0)$ such that $w_j \neq \tilde{w}_j$. From (4.17) we have

$$p'_j = \frac{(\tilde{w}_0 - w_0) x_0 + \sum_{p'_i \in Q'_0} (\tilde{w}_i - w_i) p'_i + \sum_{p'_i \in Q' - (Q'_0 \cup \{p'_j\})} (\tilde{w}_i - w_i) p'_i}{w_j - \tilde{w}_j}. \tag{4.18}$$

We see from (4.14) and (4.16) that

$$\begin{aligned} &\{(\tilde{w}_0 - w_0) + \sum_{p'_i \in Q'_0} (\tilde{w}_i - w_i) + \sum_{p'_i \in Q' - (Q'_0 \cup \{p'_j\})} (\tilde{w}_i - w_i)\}/(w_j - \tilde{w}_j) \\ &= \{(1 - \tilde{w}_j) - (1 - w_j)\}/(w_j - \tilde{w}_j) \\ &= (w_j - \tilde{w}_j)/(w_j - \tilde{w}_j) = 1. \end{aligned} \tag{4.19}$$

Therefore, from (4.18) p'_j belongs to $\mathrm{A}((Q' - \{p'_j\}) \cup \{x_0\})$, which contradicts (d.3a). □

Lemma 4.9: *Suppose that* (Q, Q_0, x_0) *satisfies* (d.1), (d.2) *and* (d.3a). *Let* $\hat{x}$ *be the minimum-norm point of* $\mathrm{A}(Q' \cup \{x_0\})$ *expressed as*

$$\hat{x} = x_0 + \sum_{p_i \in Q} \beta_i (p_i' - x_0). \tag{4.20}$$

Also let y *be the minimum-norm point of* $\mathrm{A}(Q) \cap V_k$ *expressed as*

$$y = \sum_{p_i \in Q} \lambda_i p_i, \qquad \sum_{p_i \in Q} \lambda_i = 1. \tag{4.21}$$

Then $\hat{x} = \mathbf{0}$ *if and only if* $y \neq x_0$. *Furthermore,* $\hat{x} = \mathbf{0}$ *and* $\beta_i \geq 0$ *for each* $p_i \in Q - Q_0$ *if and only if* $y \neq x_0$ *and* $\lambda_i \geq 0$ *for each* $p_i \in Q - Q_0$.

(Proof) Suppose that for the minimum-norm point y in $\mathrm{A}(Q) \cap V_k$ we have $y \neq x_0$. From (4.21) we have

$$y' = \sum_{p_i \in Q} \lambda_i p_i', \qquad \sum_{p_i \in Q} \lambda_i = 1. \tag{4.22}$$

Since y' belongs to V_k and is orthogonal to $H(x_0) \cap V_k$, we have

$$y' = \gamma x_0, \qquad \gamma < 1. \tag{4.23}$$

(For, since $y \in H(x_0)_+$, we have $y' \in H(x_0)_+$.) From (4.22) and (4.23),

$$(1 - \gamma) x_0 + \sum_{p_i \in Q} \lambda_i (p_i' - x_0) = \mathbf{0}. \tag{4.24}$$

Hence,

$$x_0 + \sum_{p_i \in Q} (\lambda_i / (1 - \gamma))(p_i' - x_0) = \mathbf{0}. \tag{4.25}$$

Then we have $\mathbf{0} \in \mathrm{A}(Q' \cup \{x_0\})$ and hence $\hat{x} = \mathbf{0}$.

Now suppose that $\hat{x} = \mathbf{0}$. Define a vector z by

$$z = x_0 + \sum_{p_i \in Q} \beta_i (p_i - x_0), \tag{4.26}$$

where β_i's are those appearing in (4.20). It follows from (4.26) that

$$z \in \mathrm{A}(Q) \cap V_k \tag{4.27}$$

since $z' = \hat{x} = \mathbf{0}$ from (4.20). Also, from (4.20) and (4.26) we have

$$z = \sum_{p_i \in Q} \beta_i (p_i - p_i'). \tag{4.28}$$

Since $p_i - p_i'$ are parallel to $H(x_0)$, we have from (4.28)

$$z \in H(x_0) - x_0, \tag{4.29}$$

where $H(x_0) - x_0$ is the hyperplane obtained by translating $H(x_0)$ by $-x_0$, i.e., the hyperplane through the origin that is parallel to $H(x_0)$. Therefore, $z \notin H(x_0)$, so that $z \neq x_0$. Moreover, from (4.28) $z \in H(x_0)_+$, which implies that there exists a point q on the line segment between z and x_0 such that $q \neq x_0$ and $||q|| < ||x_0||$. Since z and x_0 (and hence q) belong to $\mathrm{A}(Q) \cap V_k$, it follows that $||y|| < ||x_0||$.

For the second part of this lemma, suppose that $x_0 + \sum_{p_i \in Q} \beta_i(p_i' - x_0) = \mathbf{0}$ and $\beta_i \geq 0$ $(p_i \in Q - Q_0)$. Since $\hat{x} = \mathbf{0}$, we have $y \neq x_0$, so that (4.25) holds. Because of (d.3a) β_i $(p_i \in Q - Q_0)$ are uniquely determined. It follows from (4.20), (4.23), (4.25) and the uniqueness of β_i for $p_i \in Q - Q_0$ that

$$\lambda_i = \beta_i(1 - \gamma) \geq 0 \quad (p_i \in Q - Q_0). \tag{4.30}$$

On the other hand, if $y \neq x_0$ and $\lambda_i \geq 0$ $(p_i \in Q - Q_0)$, then from (4.30), we have $\beta_i \geq 0$ $(p_i \in Q - Q_0)$. □

4.3. Procedure Degenerate and Its Validity

We now give Procedure Degenerate for avoiding the degeneracy.

Procedure Degenerate

Input: A degenerate corral $(Q, Q_0 = Q, x_0)$ and $\hat{x} = x_0$.

Step D.1: If there exists no point of P that lies in the open halfspace $\hat{H}(Q)_+$ $(= H(\hat{x})_+)$, then stop (the current x_0 is the desired minimum-norm point). Otherwise choose any point $p_{i^*} \in P \cap \hat{H}(Q)_+$ and put $Q := Q \cup \{p_{i^*}\}$.

Step D.2: Let $\hat{y}$ be the minimum-norm point in $\mathrm{A}(Q' \cup \{x_0\})$. Express $\hat{y}$ as $\hat{y} = w_0 x_0 + \sum_{p_i' \in Q'} w_i p_i'$ with $w_0 + \sum_{p_i' \in Q'} w_i = 1$ and decompose $\hat{y}$ into $\hat{y} = \hat{y}_1 + \hat{y}_2$ in such a way that $\hat{y}_1 = w_0 x_0 + \sum_{p_i' \in Q_0'} w_i p_i'$ and $\hat{y}_2 = \sum_{p_i' \in Q' - Q_0'} w_i p_i'$.

(2-1) If $\hat{y} = \mathbf{0}$ and $\hat{y}_2 \in \mathrm{Cone}(Q' - Q_0')$, then go to Step D.4.

(2-2) If $\hat{y} \neq \mathbf{0}$ and $\hat{y}_2 \in \mathrm{Cone}(Q' - Q_0')$, then delete from Q the points p corresponding to the points p' not in the minimal face of the cone $\mathrm{Cone}(Q' - Q_0')$ on which $\hat{y}_2$ lies, put $\hat{x} \leftarrow \hat{y}$ and go to Step D.1. (In this case (Q, Q_0, x_0) is a degenerate corral.)

(2-3) If $\hat{y}_2 \notin \mathrm{Cone}(Q' - Q_0')$, then go to Step D.3.

Step D.3: Let z be the point on the line segment $\overline{\hat{x}_2 \hat{y}_2}$ that is nearest to $\hat{y}_2$ under the constraint that z belongs to $\mathrm{Cone}(Q' - Q_0')$, where $\hat{x}_2$ is defined from $\hat{x}$ as in (d.3b) of the definition of degenerate corral. Delete from Q the points p corresponding to the points $p' \in Q' - Q_0'$ not in the minimal face of the cone $\mathrm{Cone}(Q' - Q_0')$ on which z lies. Let λ be the real number satisfying $z = \lambda \hat{x}_2 + (1 - \lambda)\hat{y}_2$. Put $\hat{x} := \lambda \hat{x} + (1 - \lambda)\hat{y}$ and go to Step D.2.

Step D.4: Let y be the minimum-norm point of $\mathrm{A}(Q) \cap V_k$. Repeat the minor cycle of the main algorithm until we get a new corral (Q, x_0). If the obtained corral (Q, x_0) is degenerate again, then go to Step D.1 with the current degenerate corral $(Q, Q_0 = Q, x_0)$. Otherwise return to the main algorithm with the nondegenerate corral (Q, x_0).

(End)

We show that $||\hat{x}||$ decreases every time $\hat{x}$ is updated and that Procedure Degenerate terminates in finitely many steps.

Lemma 4.10: *After Step D.1 we have*

$$\dim(\mathrm{A}(Q' \cup \{x_0\})) = \dim(\mathrm{A}(Q'_0 \cup \{x_0\})) + |Q' - Q'_0|. \tag{4.31}$$

(Proof) This lemma follows from the way of selecting p_i in Step D.1. □

Lemma 4.11: *When we enter Step D.1, the current (Q, Q_0, x_0) is a degenerate corral.*

(Proof) For the first execution of Step D.1 (Q, Q_0, x_0) is a degenerate corral. Suppose that in an execution of Step D.1, (Q, Q_0, x_0) is a degenerate corral and then we go to Step D.2 from Step D.1. The current (Q, Q_0, x_0) satisfies (d.1)∼(d.4) of the definition of degenerate corral except that $|Q| \leq n$. If we enter Step D.1 from Step D.3 or Step D.4, while repeating the minor cycle (formed by Steps D.2 and D.3 or Steps 2 and 3 of the main algorithm), some points, at least one, in $Q - Q_0$ have been removed in minor cycles. Hence, when entering Step D.1 from Step D.3 or Step D.4 the current (Q, Q_0, x_0) is a degenerate corral. If we enter Step D.1 from Step D.2, then the minimum-norm point y of $\mathrm{A}(Q) \cap V_k$ satisfies that $y = x_0$. From Lemma 3.2, there exists a hyperplane of the form (4.1) such that it contains Q. Since Q is always affinely independent, we have $|Q| \leq n$. The other conditions for a degenerate corral is easy to verify. □

Lemma 4.12: *Suppose that we have come to Step D.2 from Step D.1. Then the vector $\hat{y}$ in Step D.2 satisfies $||\hat{y}|| < ||\hat{x}||$.*

(Proof) Since $\hat{x}$ is the minimum-norm point of $H(\hat{x})$ and the point p_i chosen in Step D.1 belongs to $H(\hat{x})_+$, we have $||\hat{y}|| < ||\hat{x}||$. □

Lemma 4.13: *Suppose that we have come to Step D.2 from Step D.3. Then we have $||\hat{y}|| \leq ||\hat{x}||$.*

(Proof) Note that $\hat{x}$ belongs to the affine space generated by x_0 and p' $(p \in Q)$, in which $\hat{y}$ is the minimum-norm point. □

Lemma 4.14: *Let $\hat{z}$ be the point $\lambda\hat{x} + (1 - \lambda)\hat{y}$ obtained at the end of Step D.3. Then, in Step D.3 before updating $\hat{x}$ we have $||\hat{y}|| < ||\hat{x}||$ and $||\hat{z}|| < ||\hat{x}||$.*

(Proof) When we have come to Step D.3 from Step D.2, we have $\hat{y}, \hat{z} \neq \hat{x}$ since $\hat{x}_2$ lies in the cone generated by p' $(p \in Q - Q_0)$. We thus have $||\hat{y}|| < ||\hat{x}||$. Since $\hat{z}$ is expressed as $\hat{z} = \lambda\hat{x} + (1 - \lambda)\hat{y}$ with $0 < \lambda < 1$,

$$\begin{aligned} ||\hat{z}|| &= ||\lambda\hat{x} + (1 - \lambda)\hat{y}|| \\ &\leq \lambda||\hat{x}|| + (1 - \lambda)||\hat{y}|| \\ &< \lambda||\hat{x}|| + (1 - \lambda)||\hat{x}|| \\ &= ||\hat{x}||. \end{aligned} \tag{4.32}$$

□

Lemma 4.15: *When we enter Step D.4 let $\hat{z}$ be the point on the line segment $\overline{x_0 y}$ that is nearest to y under the constraint that $\hat{z}$ belongs to the convex hull* C(Q). *Then* $||\hat{z}|| < ||x_0||$.

(Proof) Since $\hat{z}$ can be expressed as $\hat{z} = \lambda y + (1-\lambda)x_0$ for some λ with $0 \le \lambda \le 1$, where $x_0 = \sum_{p_i \in Q_0} w_i p_i$, $w_i > 0$ ($p_i \in Q_0$) and $y = \sum_{p_i \in Q} \lambda_i p_i$. From the condition for entering Step D.4 and the proof of Lemma 4.9, we have $\lambda_i > 0$ ($p_i \in Q - Q_0$), which implies $\lambda > 0$, i.e., $\hat{z} \neq x_0$. □

Theorem 4.16: *Procedure Degenerate terminates in finitely many steps.*

(Proof) The current triple (Q, Q_0, x_0) at the beginning of an execution of Step D.1 is a degenerate corral. When we go to Step D.2 from Step D.1, the current (Q, Q_0, x_0) satisfies (d.1)~(d.4) of the definition of degenerate corral except that $|Q| \le n$. While repeating the minor cycle (formed by Steps D.2 and D.3), some points, at least one, in $Q - Q_0$ are removed in each minor cycle. Since (Q_0, Q_0, x_0) is a degenerate corral, after at most $|Q - Q_0|$ repetitions of the minor cycle the current (Q, Q_0, x_0) becomes a degenerate corral and we go to Step D.1.

Also, from Lemmas 4.12~4.14, $||\hat{x}||$ or $||x_0||$ strictly decreases every time we go to Step D.1. Therefore, we do not encounter any degenerate corral more than once. Hence, the procedure is finite. □

5. Solving the Equations

In Step 2 of the main algorithm we need to find the minimum-norm point in $\mathrm{A}(Q) \cap V_k$, where $V_k = H_1 \cap \cdots \cap H_k$. Without loss of generality and for simplicity, we assume that the hyperplanes are given as $H_i = \{x \mid x \in \mathbf{R}^n, x(n-k+i) = 0\}$ $(i = 1, \cdots, k)$. Now, the problem is equivalent to the following:

$$
\begin{aligned}
(\tilde{\mathbf{P}}) \quad \text{Minimize} \quad & \frac{1}{2}||x||^2 \\
\text{subject to} \quad & x = \sum_{p_i \in Q} v_i p_i, \\
& \sum_{p_i \in Q} v_i = 1, \\
& \sum_{p_i \in Q} v_i p_i(n-k+j) = 0 \quad (j = 1, \cdots, k).
\end{aligned} \tag{5.1}
$$

The Lagrangian function $L(v, \lambda_1, \cdots, \lambda_{k+1})$ is given by

$$L(v, \lambda_1, \cdots, \lambda_{k+1}) = \frac{1}{2} v^\top Q^\top Q v + \lambda_1 (\sum_{p_i \in Q} v_i - 1)$$

$$+ \sum_{j=1}^{k} \lambda_{j+1} \left(\sum_{p_i \in Q} v_i p_i(n-k+j) \right), \tag{5.2}$$

where λ_i $(i = 1, \cdots, k+1)$ are scalar variables. The optimal solution of Problem $(\tilde{\mathbf{P}})$ is given by solving

$$\nabla L(v, \lambda_1, \cdots, \lambda_{k+1}) = \mathbf{0}, \tag{5.3}$$

i.e.,

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & e^{\top} \\ 0 & 0 & \cdots & 0 & Q(n-k+1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & Q(n) \\ e & Q^{\top}(n-k+1) & \cdots & Q^{\top}(n) & Q^{\top}Q \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{k+1} \\ v \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \mathbf{0} \end{pmatrix}, \tag{5.4}$$

where $Q(i)$ is the ith row of Q and e is the vector with each component being equal to 1.

Concerning the coefficient matrix of (5.4), we have the following

Lemma 5.1: *Under the nondegeneracy assumption, the symmetric matrix*

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & e^{\top} \\ 0 & 0 & \cdots & 0 & Q(n-k+1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & Q(n) \\ e & Q^{\top}(n-k+1) & \cdots & Q^{\top}(n) & Q^{\top}Q \end{pmatrix}. \tag{5.5}$$

is nonsingular.
(Proof) Define

$$A = (Q^{\top}(n-k+1) \quad Q^{\top}(n-k+2) \quad \cdots \quad Q^{\top}(n)). \tag{5.6}$$

Then, because of the nondegeneracy assumption and the expression (3.6), we have

$$\text{rank } A = k. \tag{5.7}$$

Also define $Q^{\top} = (Q_1^{\top} \; Q_2^{\top})$, where $Q_2^{\top} = (Q^{\top}(n-k+1) \quad Q^{\top}(n-k+2) \quad \cdots \quad Q^{\top}(n))$. Then, from (5.6) and (5.7) we have $A = Q_2^{\top}$ and rank $Q_2 = k$ from (5.6). The symmetric matrix (5.5) is rewritten as:

$$\begin{pmatrix} 0 & \mathbf{0}^{\top} & e^{\top} \\ \mathbf{0} & O & Q_2 \\ e & Q_2^{\top} & Q^{\top}Q \end{pmatrix}. \tag{5.8}$$

We consider a system of homogeneous linear equations as follows:

$$\begin{pmatrix} 0 & \mathbf{0}^\top & e^\top \\ \mathbf{0} & O & Q_2 \\ e & Q_2^\top & Q^\top Q \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0}, \tag{5.9}$$

i.e.,

$$e^\top x_3 = 0, \tag{5.10}$$

$$Q_2 x_3 = \mathbf{0}, \tag{5.11}$$

$$x_1 e + Q_2^\top x_2 + Q^\top Q x_3 = \mathbf{0}. \tag{5.12}$$

Multiplying (5.12) by $x_3^\top$ from the left, we have

$$x_1 x_3^\top e + x_3^\top Q_2^\top x_2 + x_3^\top Q^\top Q x_3 = \mathbf{0}, \tag{5.13}$$

i.e.,

$$x_3^\top Q^\top Q x_3 = \mathbf{0}. \tag{5.14}$$

Therefore, we have

$$Q x_3 = \mathbf{0}. \tag{5.15}$$

Then, (5.12) becomes

$$x_1 e + Q_2^\top x_2 = \mathbf{0}. \tag{5.16}$$

Since the columns of Q are affinely independent, from (5.10) and (5.15) we have

$$x_3 = \mathbf{0}. \tag{5.17}$$

For Problem $(\tilde{\mathbf{P}})$, there exists a feasible solution v which satisfies:

$$e^\top v = 1 \tag{5.18}$$

$$Q_2 v = \mathbf{0}. \tag{5.19}$$

Multiplying (5.16) by $v^\top$ from the left, we have

$$x_1 v^\top e + v^\top Q_2^\top x_2 = \mathbf{0}. \tag{5.20}$$

From (5.18) $\sim$ (5.20) we have

$$x_1 = 0. \tag{5.21}$$

From (5.16) and (5.21) we also have

$$Q_2^\top x_2 = \mathbf{0}. \tag{5.22}$$

Since the columns of $Q_2^\top$ are linearly independent due to (5.8), we have

$$x_2 = \mathbf{0}. \tag{5.23}$$

It follows from (5.9), (5.17), (5.21) and (5.23) that the matrix (5.5) is nonsingular.

□

The system (5.4) of equations is rewritten as

$$e\lambda_1 + \sum_{j=1}^{k} Q^\top(n-k+j)\lambda_{j+1} + Q^\top Q v = \mathbf{0}, \tag{5.24}$$

$$e^\top v = 1, \tag{5.25}$$

$$Q(n-k+j)v = 0 \quad (j = 1, \cdots, k). \tag{5.26}$$

It follows from Lemma 5.1 that the system of equations (5.24) $\sim$ (5.26) has a unique solution $(\lambda_1, \lambda_2, \cdots, \lambda_{k+1}, v)$. In order to solve it, we first find the solution $(u_1, u_2, \cdots, u_{k+1})$ of the following system of linear equations:

$$(ee^\top + Q^\top Q)u_1 = e, \tag{5.27}$$

$$(ee^\top + Q^\top Q)u_{j+1} = Q^\top(n-k+j) \qquad (j = 1, \cdots, k), \tag{5.28}$$

where note that the common coefficient matrix of (5.27) and (5.28) is nonsingular since Q is affinely independent.

Lemma 5.2: *For the solutions $u_1, u_2, \cdots, u_{k+1}$ of (5.27) and (5.28), under the non-degeneracy assumption, there is a unique solution $(\alpha_1, \cdots, \alpha_k)$ of the following system of linear equations:*

$$Q(n-k+1)u_2\alpha_1 + Q(n-k+1)u_3\alpha_2 + \cdots + Q(n-k+1)u_{k+1}\alpha_k = -Q(n-k+1)u_1, \tag{5.29}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$Q(n)u_2\alpha_1 + Q(n)u_3\alpha_2 + \cdots + Q(n)u_{k+1}\alpha_k = -Q(n)u_1. \tag{5.30}$$

(Proof) Put $S = ee^\top + Q^\top Q$. From (5.27) and (5.28)

$$u_1 = S^{-1}e, \tag{5.31}$$

$$u_{j+1} = S^{-1}Q^\top(n-k+j) \qquad (j = 1, \cdots, k). \tag{5.32}$$

The coefficient matrix of (5.29)$\sim$(5.30) is given by

$$\begin{pmatrix} Q(n-k+1)S^{-1}Q^\top(n) & \cdots & Q(n-k+1)S^{-1}Q^\top(n-k+1) \\ \vdots & \vdots & \vdots \\ Q(n)S^{-1}Q^\top(n) & \cdots & Q(n)S^{-1}Q^\top(n-k+1) \end{pmatrix}$$

$$= \begin{pmatrix} Q(n-k+1) \\ \vdots \\ Q(n) \end{pmatrix} S^{-1} \begin{pmatrix} Q(n-k+1) \\ \vdots \\ Q(n) \end{pmatrix}^\top = Q_2 S^{-1} Q_2^\top. \tag{5.33}$$

Under nondegenrancy assumptiom, rank $Q_2 = k$ (i.e., Q_2 is of row full rank). Hence, the matrix (5.33) is nonsingular. □

Lemma 5.3: *Under the nondegeneracy assumption, for the solutions* $(u_1, \cdots, u_{k+1})$ *of* (5.27) *and* (5.28) *and* $(\alpha_1, \cdots, \alpha_k)$ *of* (5.29) ~ (5.30) *we have*

$$Q(i)(u_1 + \alpha_1 u_2 + \cdots + \alpha_k u_{k+1}) = 0 \quad (i = n-k+1, \cdots, n). \tag{5.34}$$

Define

$$v_\beta = u_1 + \alpha_1 u_2 + \cdots + \alpha_k u_{k+1}. \tag{5.35}$$

Then, $v^* = v_\beta / e^\top v_\beta$ *is the solution* v *of* (5.4).

(Proof) First, we show $e^\top v_\beta \neq 0$. We have from (5.27) and (5.28)

$$(ee^\top + Q^\top Q)v_\beta = e + \sum_{j=1}^{k} \alpha_j Q^\top(n-k+j), \tag{5.36}$$

i.e.,

$$Sv_\beta = e + \sum_{j=1}^{k} \alpha_j Q^\top(n-k+j) \tag{5.37}$$

For a feasible solution v of Problem ($\tilde{\mathbf{P}}$), multiplying (5.37) by $v^\top$ from the left, we have

$$v^\top S v_\beta = v^\top e + \sum_{j=1}^{k} \alpha_j v^\top Q^\top(n-k+j) = 1. \tag{5.38}$$

Hence,

$$v_\beta \neq \mathbf{0}. \tag{5.39}$$

Multiplying (5.36) by $v_\beta^\top$ from the left, we obtain

$$v_\beta^\top (e, Q^\top) \begin{pmatrix} e^\top \\ Q \end{pmatrix} v_\beta = v_\beta^\top e, \tag{5.40}$$

where note that $v_\beta^\top Q^\top(n-k+j) = 0 \ \ (j = 1, \ldots, k)$ because of (5.34). Since Q is affinely independent and $v_\beta \neq \mathbf{0}$, we have from (5.40) $e^\top v_\beta \neq 0$.

Now, we show that v^* is the required solution v. Clearly, v^* satisfies equations of (5.25) and (5.26). Rewrite (5.36) as

$$e(e^\top v_\beta - 1) - \sum_{j=1}^{k} \alpha_j Q^\top(n-k+j) + Q^\top Q v_\beta = \mathbf{0}, \tag{5.41}$$

divide the above equation by $e^{\mathsf{T}}v_\beta(\neq 0)$, and put $\lambda_1 = (e^{\mathsf{T}}v_\beta - 1)/e^{\mathsf{T}}v_\beta$ and $\lambda_{j+1} = \alpha_j/e^{\mathsf{T}}v_\beta$ $(j = 1, \ldots, k)$. We then see that v^* together with these λ_i $(i = 1, \cdots, k+1)$ defined here also satisfies (5.24). Under the nondegeneracy assumption, it follows from Lemma 5.1 that v^* is the unique solution v of (5.4). □

To determine v in (5.4) we first solve (5.27) and (5.28) and then (5.29)~(5.30), as shown in Lemma 5.3. Here, note that the coefficient matrices of (5.27) and (5.28) are the same and that its Cholesky factorization can effectively be updated by the technique given in [13]. The present two-stage method is effective for problems with a large n and a relatively small k, in particular.

6. Computational Experiments

Consider two types of test problems where m points p_j $(j = 1, \ldots, m)$ are given as follows.

Type 1: m points $p_j = (p_j(1), \ldots, p_j(n))^{\mathsf{T}}$ $(j = 1, \cdots, m)$ are chosen at random from the sample space where each component $p_j(i)$ $(i = 1, \ldots, n-1; j = 1, \ldots, m)$ is uniformly distributed on $\{1, \ldots, 50\}$ and $p_j(n-1), p_j(n)$ $(j = 1, \ldots, m)$ on $\{-50, \ldots, -1, 1, \ldots, 50\}$.

Type 2: m points $p_j = (p_j(1), \ldots, p_j(n))^{\mathsf{T}}$ $(j = 1, \cdots, m)$ are chosen at random from the sample space where each first component $p_j(1)$ $(j = 1, \ldots, m)$ is uniformly distributed on $[0.01 - 0.001, 0.01 + 0.001]$ and other components $p_j(k)$ $(k = 2, \ldots, n; j = 1, \ldots, m)$ uniformly on $[-1, 1]$.

Here, we consider only two hyperplanes, i.e., k=2, and assume that the hyperplanes are given as $H_1 = \{x \mid x \in \mathbf{R}^n, x(n-1) = 0\}$ and $H_2 = \{x \mid x \in \mathbf{R}^n, x(n) = 0\}$.

Figures 6.1 and 6.2 show sample behaviors of the algorithm for problems of Type 1 and Type 2 with m=1000 and n=20. We see that $||x_0||$ decreases rapidly within the first several major cycles.

Figure 6.3 shows ten-sample average behaviors of the number of major (or minor) cycles versus the dimension for the problem of Type 1 with m=500 and Figure 6.4 for the problem of Type 2.

Also, Figure 6.5 shows ten-sample average behaviors of the number of major (or minor) cycles versus the number of points for the problem of Type 1 with n=20 and Figure 6.6 for the problem of Type 2. These figures show that the growth rate of the number of cycles becomes almost zero.

Figure 6.7 shows ten-sample average behaviors of the running time (milliseconds) by S-4/10 model 41 versus the number of points for the problem of Type 1 with n=20 and Figure 6.8 for the problem of Type 2. These figures show that the running time required for solving the problem versus the number of points is almost linearly increasing.

The present computational experiments show the practicality of our algorithm.

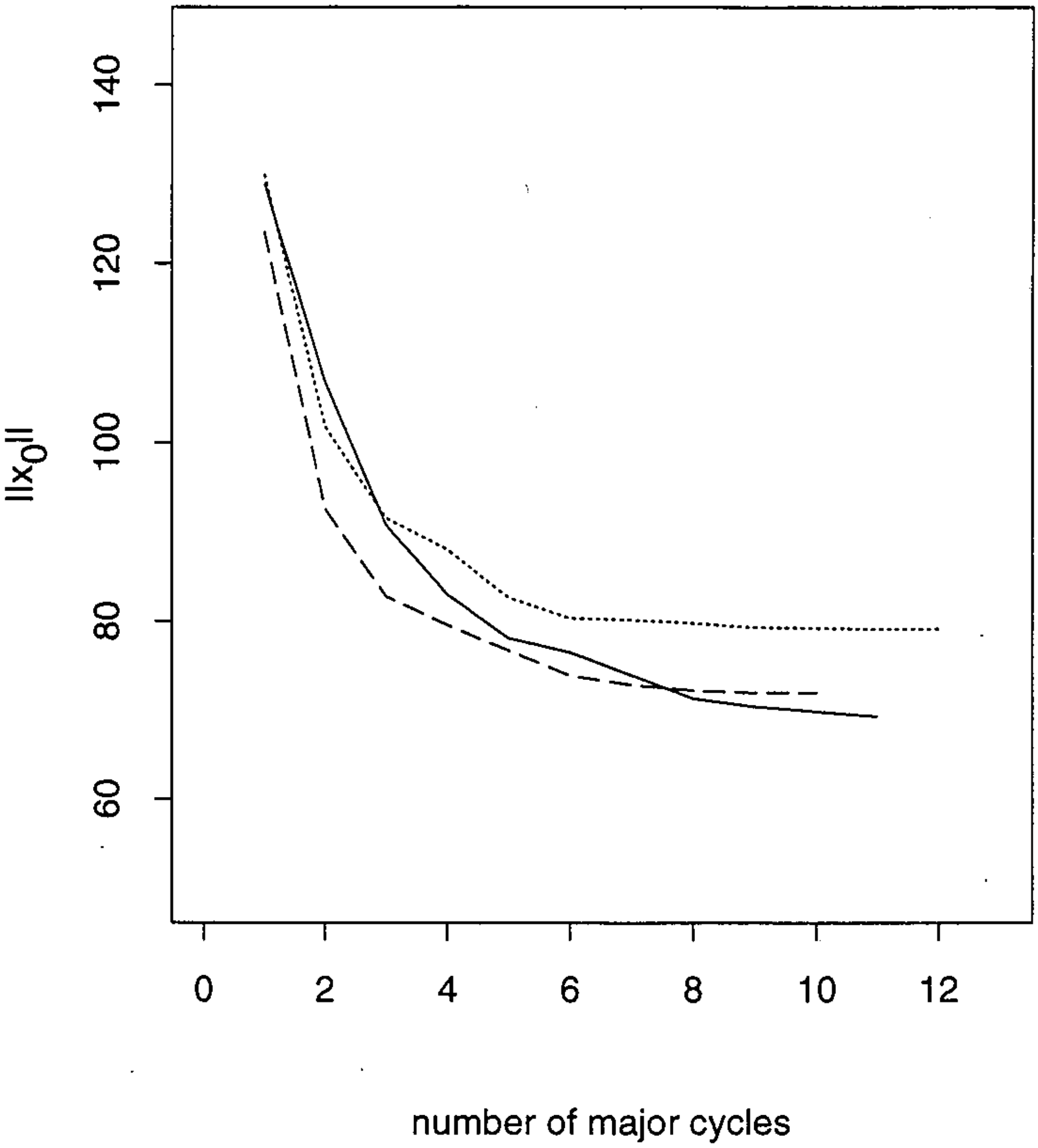

Figure 6.1: Three sample behaviors of Type 1 (m=1000, n=20).

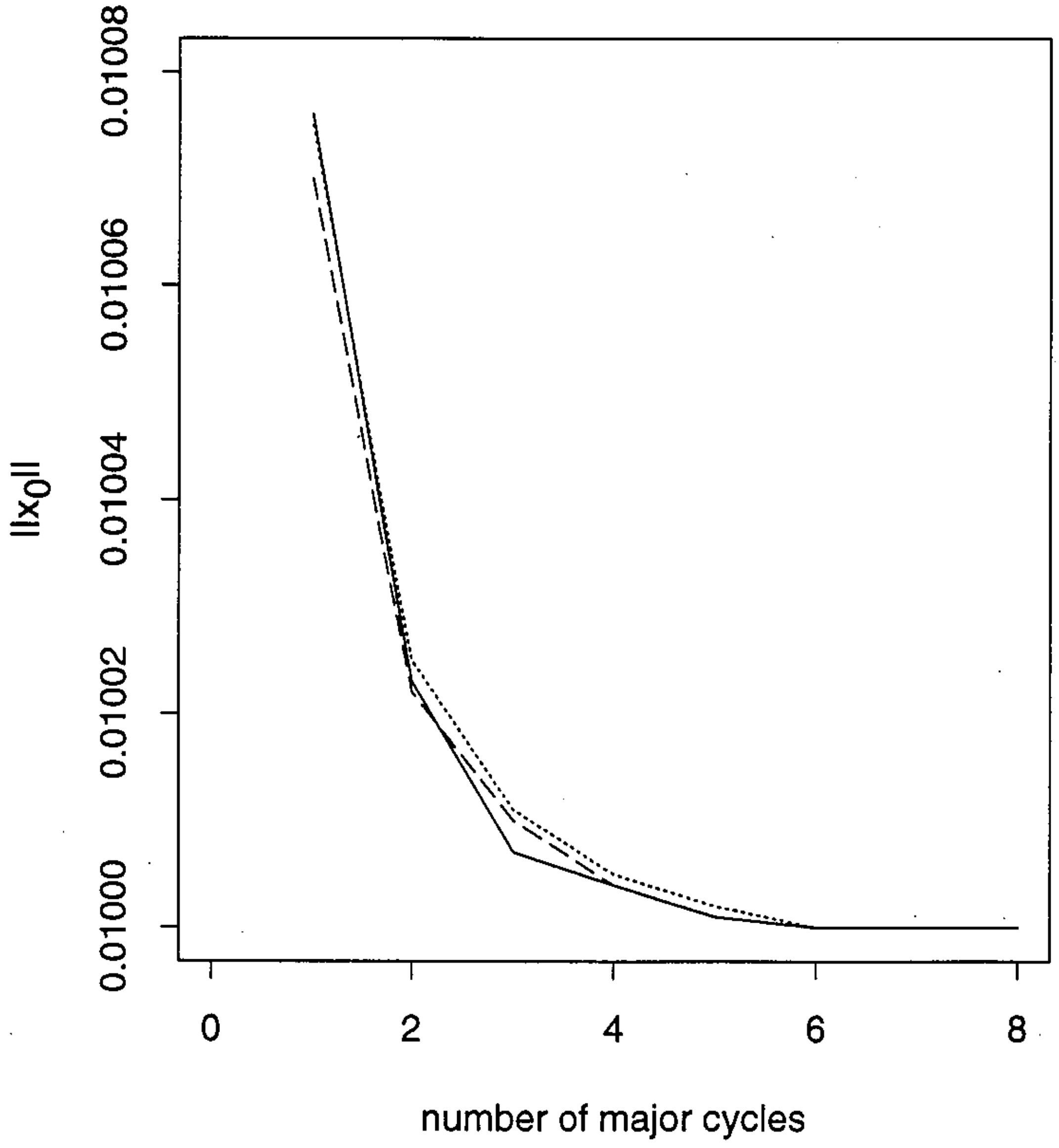

Figure 6.2: Three sample behaviors of Type 2 (m=1000, n=20).

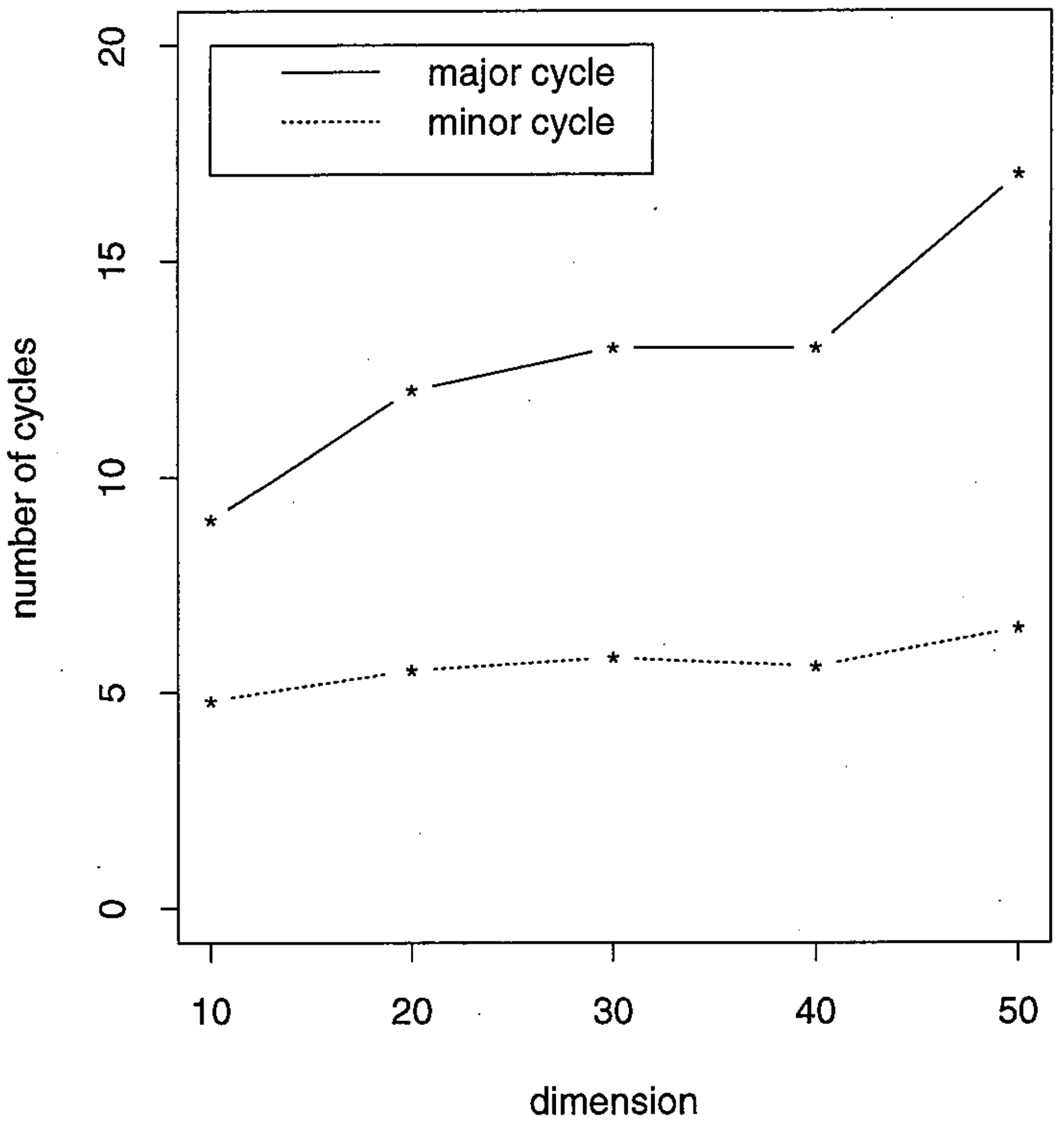

Figure 6.3: Ten-sample average behaviors for Type 1 (m=500).

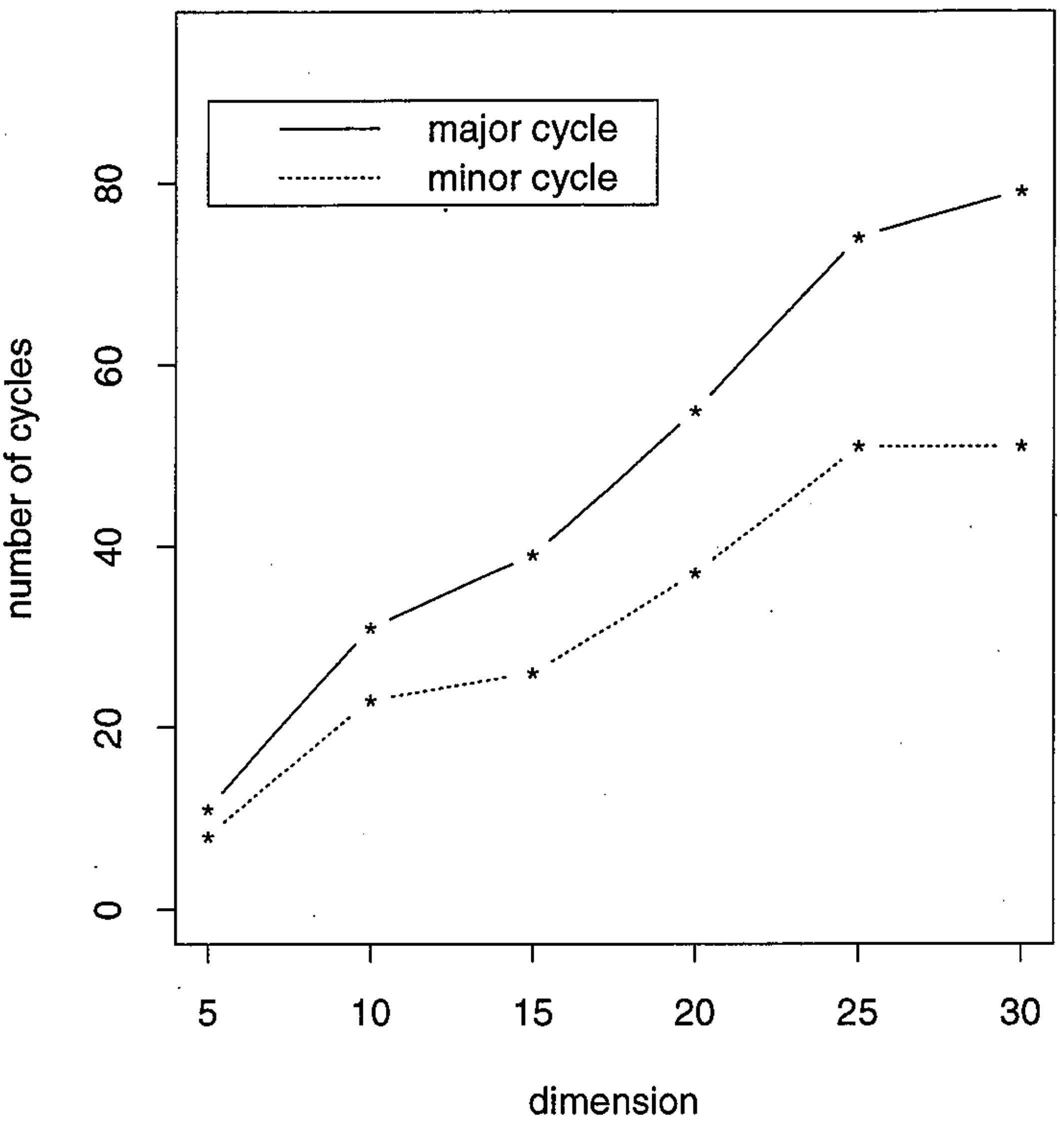

Figure 6.4: Ten-sample average behaviors for Type 2 (m=500).

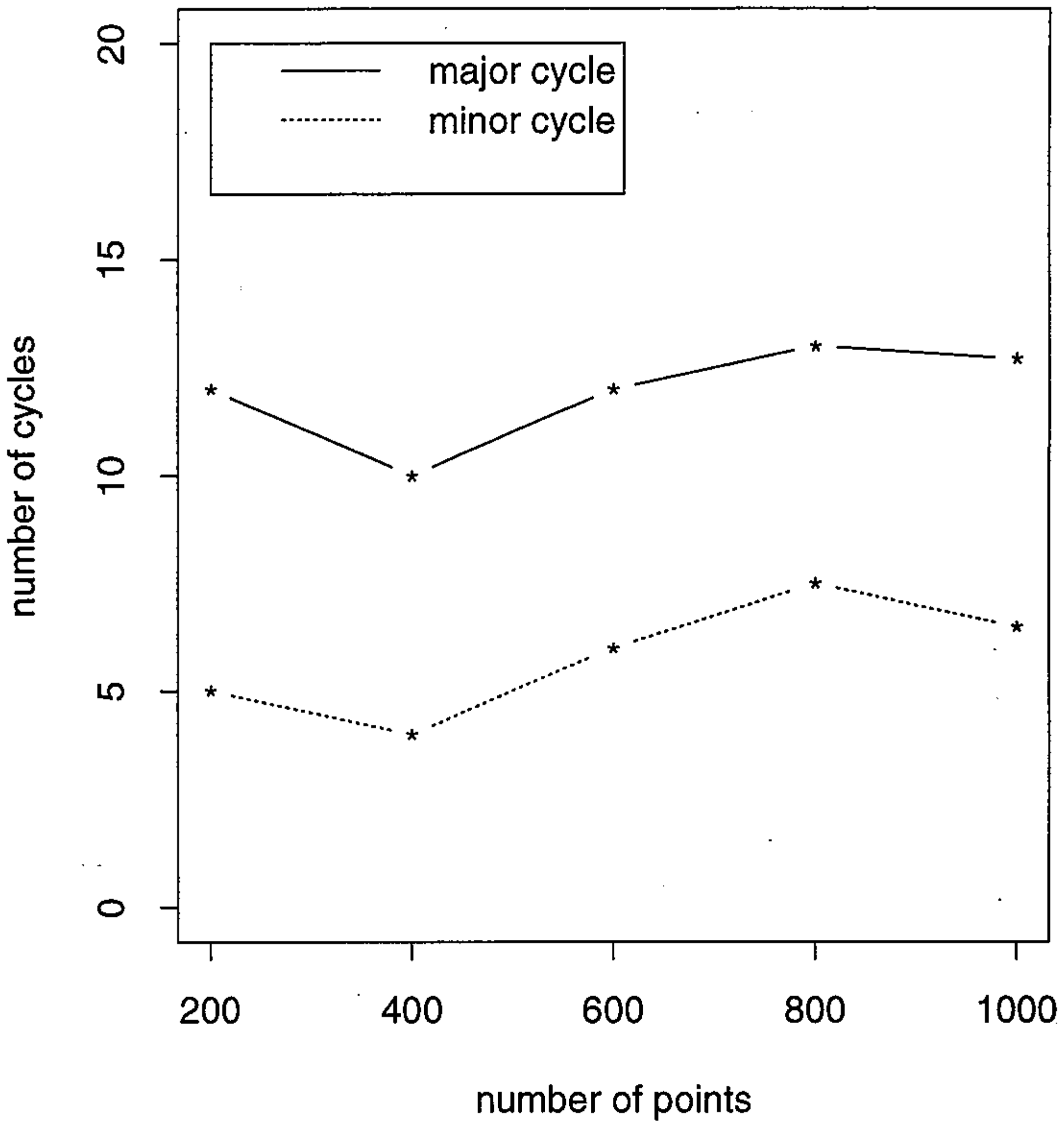

Figure 6.5: Ten-sample average behaviors for Type 1 (n=20).

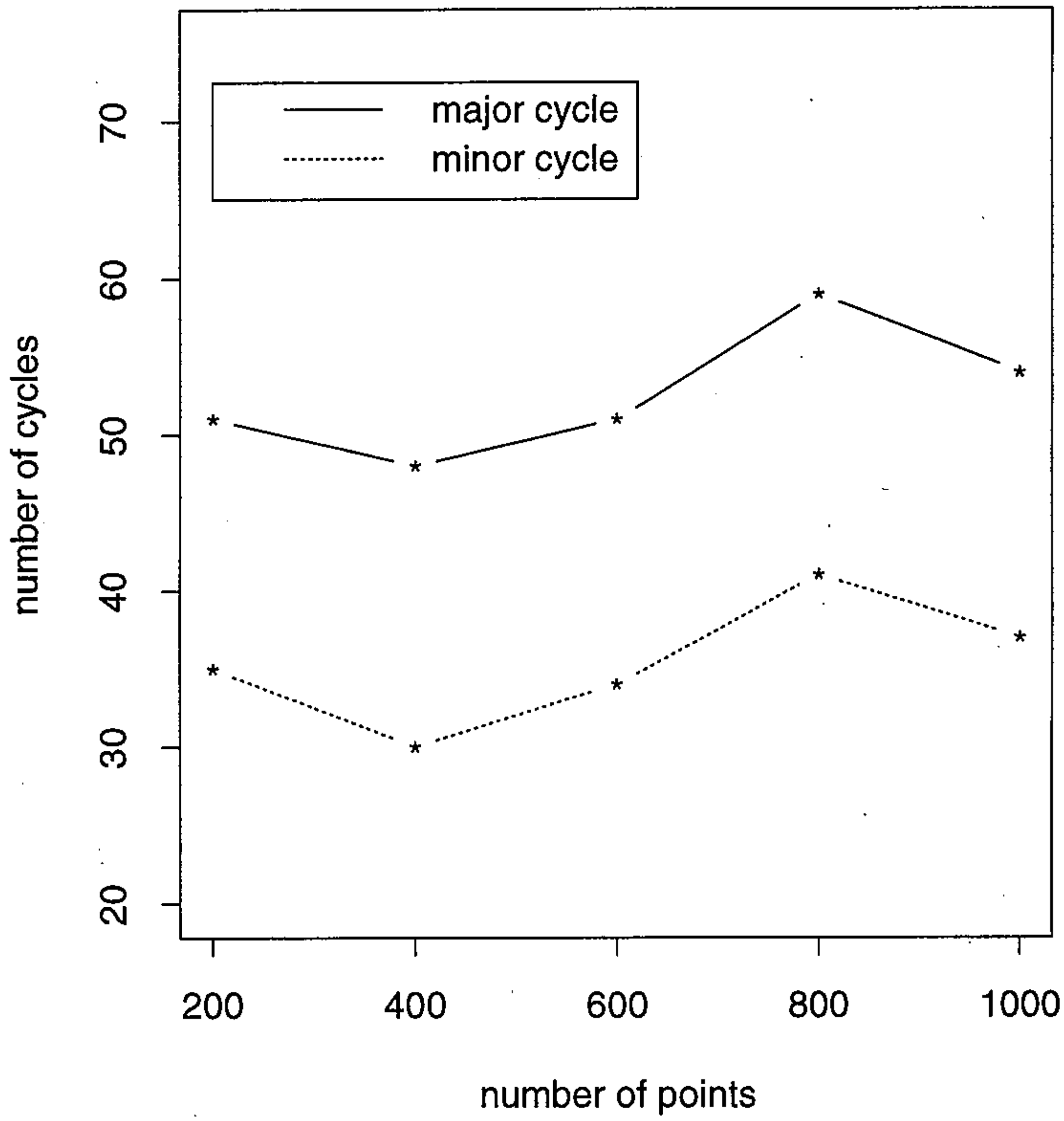

Figure 6.6: Ten-sample average behaviors for Type 2 (n=20).

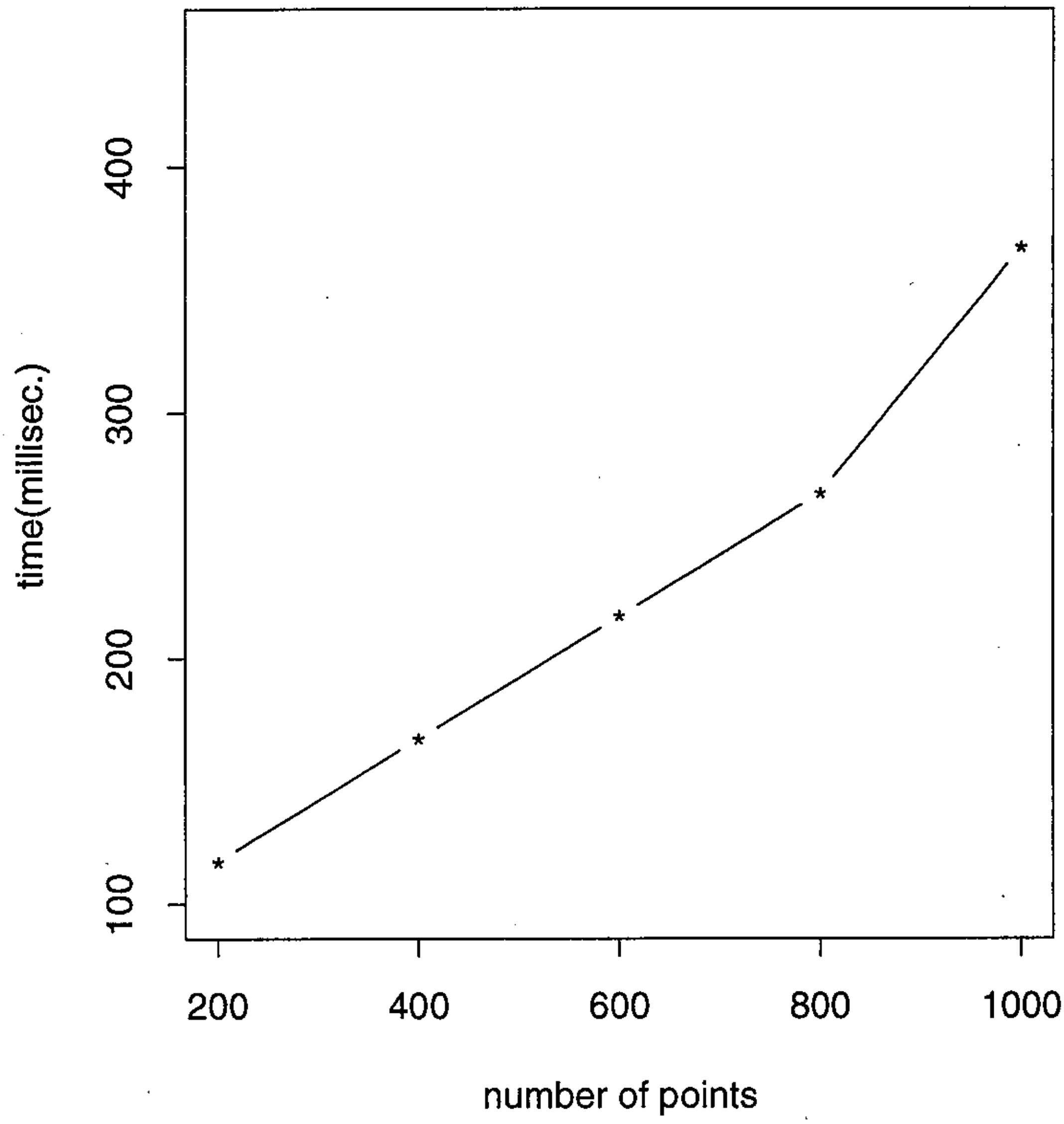

Figure 6.7: A ten-sample average behavior of the running time for Type 1 (n=20).

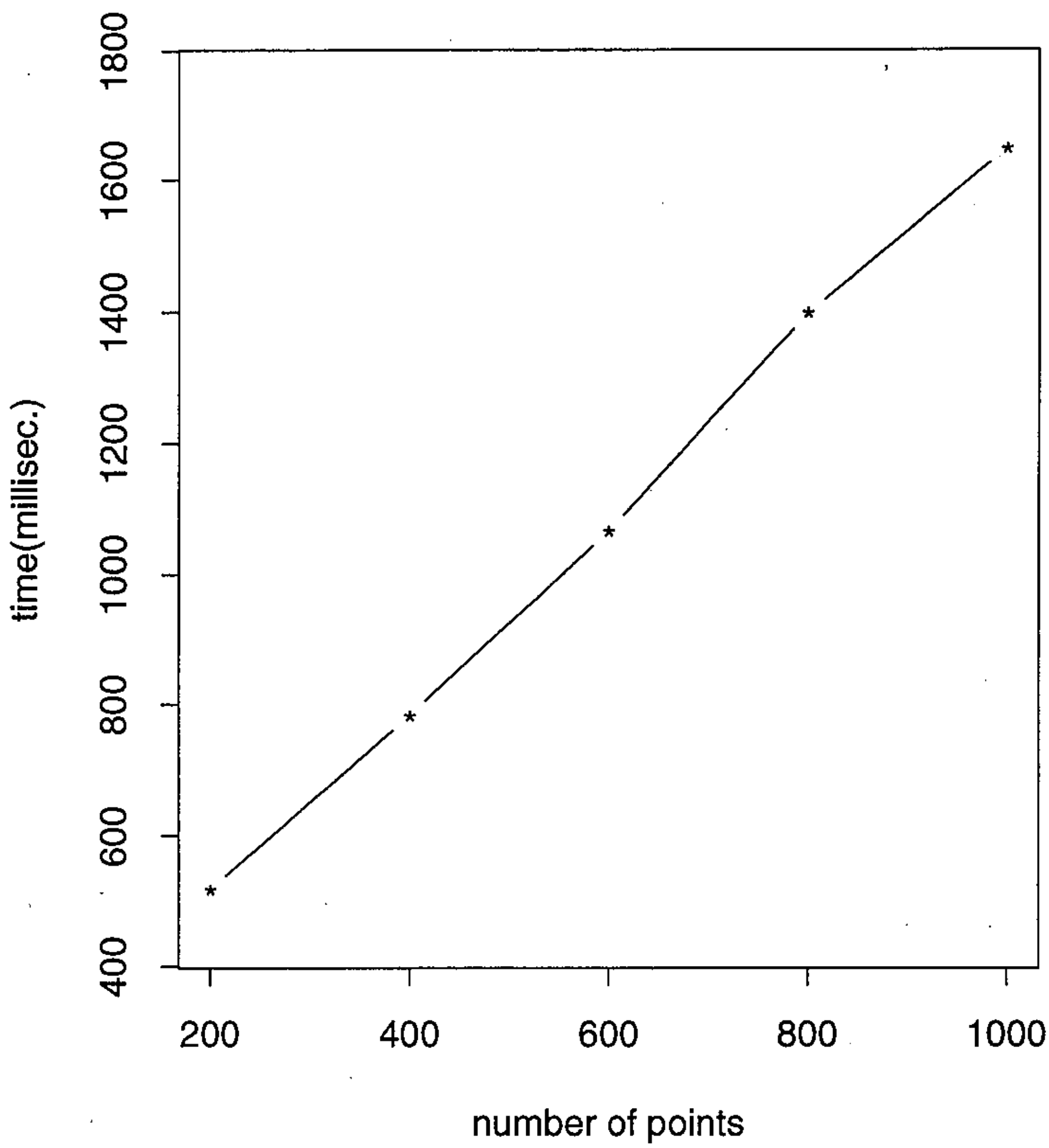

Figure 6.8: A ten-sample average behavior of the running time for Type 2 (n=20).

Acknowledgment

We are grateful to Professor Yoshitsugu Yamamoto of Institute of Socio-Economic Planning, University of Tsukuba for his useful comments.

References

[1] R. O. Barr: An efficient computational procedure for a generalized quadratic programming problem. *SIAM Journal on Control* **7** (1969) 415-429.

[2] S. Fujishige: *Submodular Functions and Optimization* (Annals of Discrete Mathematics, Vol.47, North-Holland, Amsterdam, 1991).

[3] S. Fujishige and P. Zhan: A dual algorithm for finding the minimum-norm point in a polytope. *Journal of the Operations Research Society of Japan* **33** (1990) 188-195.

[4] S. Fujishige and P. Zhan: A dual algorithm for finding a nearest pair of points in two polytopes. *Journal of the Operations Research Society of Japan* **35** (1992) 353-365.

[5] S. Fujishige, H. Sato and P. Zhan: An algorithm for finding the minimum-norm point in the intersection of a convex polyhedron and a hyperplane. *Japan Journal of Industrial and Applied Mathematics* **11** (1994) 245-264.

[6] H. Konno and K. Suzuki: A fast algorithm for solving large scale mean-variance models by compact factorization of covariance matrices. *Journal of the Operations Research Society of Japan* **35** (1992) 93-104.

[7] E. Klafszky, J. Mayer and T. Terlaky: On mathematical programming model of mixing. *European Journal of Operational Research* **42** (1989) 254-267.

[8] B. F. Mitchell, V. F. Dem'yanov and V. N. Malozemov: Finding the point of a polyhedron closest to the origin. *SIAM Journal on Control* **12** (1974) 19-26.

[9] K. Sekitani and Y. Yamamoto: A recursive algorithm for finding the minimum norm point in a polytope and a pair of closest points in two polytopes. *Mathematical Programming* **61** (1993) 233-249

[10] K. Sekitani, J.M. Shi, Y. Yamamoto and K. Yamasaki: Algorithm for the minimum-norm point in the intersection of a polytope and several hyperplanes. Discussion Paper No. 499, Institute of Socio-Economic Planning, University of Tsukuba (September 1992).

[11] H. Takehara: An interior point algorithm for large scale portfolio optimization. *Annals of Operations Research* **45** (1993) 373-386.

[12] D. R. Wilhelmsen: A nearest point algorithm for convex polyhedral cones and applications to positive linear approximation. *Mathematics of Computation* **30** (1976) 48-57.

[13] P. Wolfe: Finding the nearest point in a polytope. *Mathematical Programming* **11** (1976) 128-149.