

Chebyshev Approximation of a Point Set by a Straight Line

M. Streng and W. Wetterling

Abstract. The problem of calculating the best approximating straight line—in the sense of Chebyshev—to a finite set of points in \mathbf{R}^n is considered. First- and second-order optimality conditions are derived and analysed. Lipschitz optimization techniques can be used to find a global minimizer.

1. Introduction and Motivation

In certain technical applications it is essential to calculate a best approximating manifold in the sense of Chebyshev to a given set of points in \mathbf{R}^n , with the requirement that this manifold fulfills certain constraints, e.g., with respect to curvature. An example of such a problem occurring in railway engineering is given in [10]. Other applications where the problem arises are in the fields of CAD/CAM and robotics.

As a first step toward the solution of these general problems we consider the best approximating straight line (one-dimensional manifold with curvature zero). Our approach via optimality conditions is that of classical optimization theory.

The paper is organized as follows: In Section 2 we translate the approximation problem into an optimization problem. In Sections 3 and 4 we derive first- and second-order optimality conditions, following the general framework as developed by Hettich and Jongen [4] and Ben-Tal [1]. The first- and second-order optimality conditions are both given solely in terms of the given point coordinates, the parameters of the approximating line, and the Lagrange multipliers. In Section 5 we analyse these second-order conditions and give a characterization of point configurations for which they are automatically fulfilled by a best approximant satisfying the first-order conditions. Finally in Section 6 we indicate how a global minimizer can be obtained by Lipschitz optimization techniques.

Approximation by a straight line is a special case of the general problem of approximation by linear manifolds. Work on the extension of our results to this

Date received: June 30, 1992. Date revised: May 17, 1993. Communicated by Dietrich Braess.

AMS classification: Primary 65D10; Secondary 41A50, 90C30.

Key words and phrases: Chebyshev approximation, Optimality conditions, Lipschitz optimization.

case is in progress. Gritzmann and Klee consider the problem in a more general framework from the viewpoint of complexity theory in [3] and several other papers announced there. The least-squares approximation of finite sets by linear manifolds has been investigated by Späth in [8].

2. Statement of the Problem

Let P be a finite nonvoid set of points x_1, \dots, x_m in \mathbf{R}^n . We want to find a straight line $x(t) = a + bt$ ($a, b \in \mathbf{R}^n, t \in \mathbf{R}$) which approximates P in the Chebyshev sense, that means the maximum distance of the points x_i to the line is minimized. In other words, the problem is to determine

$$(1) \quad \min_{a,b} \max_i \min_t \|x_i - a - bt\|$$

with $\|\cdot\|$ the Euclidean norm. A line $x(t)$ yielding a local minimum is called a *Chebyshev line* (for the set P). The existence of a Chebyshev line follows by a standard compactness argument. However, the minimum need not be unique. An example is the set of corner points of a regular polygon in \mathbf{R}^2 .

In order to reduce the amount of nonuniqueness in the parameters we assume a and b to be normalized by $b^T a = 0, b^T b = 1$. A simple calculation shows that $\|x_i - a - bt\|$ has a unique minimum for $t = t_i$, where $t_i = b^T(x_i - a)$. The square of the distance between the point x_i and the line $x(t) = a + bt$ is $(x_i - a)^T(I - bb^T)(x_i - a)$. Note that $I - bb^T$ is the orthogonal projection on the hyperplane orthogonal to b .

Problem (1) can now be written as an optimization problem with m inequality constraints and two equality constraints:

$$(2) \quad \begin{cases} \text{minimize } f^0(r, a, b) = r & \text{subject to} \\ f^i(r, a, b) = (x_i - a)^T(I - bb^T)(x_i - a) - r \leq 0 & (1 \leq i \leq m), \\ h^1(r, a, b) = a^T b = 0, \\ h^2(r, a, b) = b^T b - 1 = 0. \end{cases}$$

This problem is referred to as the *Chebyshev line optimization problem (CLOP)*. The notations correspond to those of Ben-Tal [1], whose results are used to derive the optimality conditions for the CLOP.

3. First-Order Optimality Conditions

In this and the following section first- and second-order optimality conditions for the CLOP are established. Constraint qualifications are not required for the necessary optimality conditions. The reason is that, in the optimality conditions of Fritz John type, the Lagrange multiplier associated with the objective function can be shown to be nonzero and hence may be taken equal to 1. (Note that the gradients of the functions h^1, h^2 are linearly independent vectors in \mathbf{R}^{2n+1} .)

The Lagrangian function L for the CLOP is given by

$$\begin{aligned} L(r, a, b, \lambda, \mu) &= \sum_{i=0}^m \lambda_i f^i(r, a, b) + \sum_{j=1}^2 \mu_j h^j(r, a, b) \\ &= \lambda_0 r + \sum_{i=1}^m \lambda_i [(x_i - a)^T (I - bb^T)(x_i - a) - r] \\ &\quad + \mu_1 a^T b + \mu_2 (b^T b - 1), \end{aligned}$$

with Lagrange multipliers $\lambda_i \geq 0$ and μ_j . For a point $(r, a, b) \in \mathbf{R}^{2n+1}$ the set of active indices is defined as

$$\mathcal{J} = \mathcal{J}(r, a, b) = \{i \mid f^i(r, a, b) = 0 \ (1 \leq i \leq m)\}.$$

Theorem 1 (First-Order Optimality Conditions). *Let (r, a, b) be a local minimizer for the CLOP. Then Lagrange multipliers λ_i ($i \in \mathcal{J}$) exist such that, with $t_i = b^T(x_i - a)$,*

$$(3) \quad \begin{cases} \sum_{i \in \mathcal{J}} \lambda_i = 1, \quad \lambda_i \geq 0 & (i \in \mathcal{J}), \\ \sum_{i \in \mathcal{J}} \lambda_i (x_i - a - bt_i) = 0, \\ \sum_{i \in \mathcal{J}} \lambda_i t_i (x_i - a - bt_i) = 0. \end{cases}$$

Proof. The Fritz John type optimality conditions for the CLOP are

$$\frac{\partial L}{\partial r} = \lambda_0 - \sum_{i=1}^m \lambda_i = 0,$$

$$(4) \quad \frac{\partial L}{\partial a} = -2 \sum_{i=1}^m \lambda_i (I - bb^T)(x_i - a) + \mu_1 b = 0,$$

$$(5) \quad \frac{\partial L}{\partial b} = -2 \sum_{i=1}^m \lambda_i (x_i - a) b^T (x_i - a) + \mu_1 a + 2\mu_2 b = 0.$$

These conditions hold with multipliers $\lambda_0 \geq 0$, $\lambda_i \geq 0$ ($i \in \mathcal{J}$), $\lambda_i = 0$ ($i \notin \mathcal{J} \cup \{0\}$), $\mu_1, \mu_2 \in \mathbf{R}$, not all of them being equal to zero. Left multiplication of (4) and (5) by b^T gives

$$\mu_1 = 0, \quad \mu_2 = \sum_{i \in \mathcal{J}} \lambda_i t_i^2.$$

Now $\lambda_0 = 0$ would imply that all multipliers are zero. Hence $\lambda_0 > 0$. We may assume $\lambda_0 = 1$ and obtain (3). ■

Consider the homogeneous equations in (3) as a linear system of $2n$ equations with $|\mathcal{J}|$ unknowns λ_i . The columns of the coefficient matrix are perpendicular

to $\begin{pmatrix} b \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ b \end{pmatrix}$, hence the column rank is $\leq 2n - 2$. The row rank is $\leq |\mathcal{J}| - 1$.

The situation where the row rank is maximal and hence the multipliers λ_i are uniquely determined can be viewed as the generic case. Thus generically $|\mathcal{J}| \leq 2n - 1$, generalizing the familiar and well-understood case of generically three active points determining a line of best approximation in \mathbf{R}^2 .

4. Second-Order Optimality Conditions

The constraints in (2) are clearly nonconvex. Several local minimizers and saddle points are possible. Therefore we consider second-order optimality conditions too.

Let (r, a, b) be a local minimizer for the CLOP. First we determine $D(r, a, b)$, the cone of critical directions at (r, a, b) as defined in [1]. A point in this cone is denoted by $d = (\tilde{r}, \tilde{a}, \tilde{b})$. It satisfies $\nabla f^i d \leq 0$ ($i \in \mathcal{J} \cup \{0\}$), $\nabla h^j d = 0$ ($j = 1, 2$). Explicitly,

$$\tilde{r} \leq 0,$$

$$(6) \quad -2\tilde{a}^T(I - bb^T)(x_i - a) - 2(x_i - a)^T \tilde{b} t_i - \tilde{r} \leq 0 \quad (i \in \mathcal{J}),$$

$$(7) \quad \tilde{a}^T b + a^T \tilde{b} = 0, \quad 2\tilde{b}^T b = 0.$$

The first-order conditions (3) hold with some, possibly not unique, $\lambda_i \geq 0$ ($i \in \mathcal{J}$). Multiplying (6) by λ_i , summing over $i \in \mathcal{J}$, and using (7) and (3) gives $\tilde{r} = 0$. Furthermore, (6) can be simplified to $(\tilde{a} + \tilde{b} t_i)^T (x_i - a - b t_i) \geq 0$ ($i \in \mathcal{J}$).

We define the set \mathcal{J}_+ of the indices $i \in \mathcal{J}$ with $\lambda_i > 0$ for some choice of the λ_i in Theorem 1. Using once more the first-order conditions we get $d = (\tilde{r}, \tilde{a}, \tilde{b}) \in D(r, a, b)$ iff

$$(8) \quad \begin{cases} \tilde{r} = 0, & \tilde{a}^T b + a^T \tilde{b} = 0, & \tilde{b}^T b = 0, \\ (\tilde{a} + \tilde{b} t_i)^T (x_i - a - b t_i) \begin{cases} = 0 & (i \in \mathcal{J}_+), \\ \geq 0 & (i \in \mathcal{J} \setminus \mathcal{J}_+). \end{cases} \end{cases}$$

Next the second partial derivatives of the Lagrangian function L are calculated. We give (in matrix notation) those which are not identically zero:

$$\begin{aligned} \frac{\partial^2 L}{\partial a^2} &= 2 \sum_{i \in \mathcal{J}} \lambda_i (I - bb^T), \\ \frac{\partial^2 L}{\partial a \partial b} &= 2 \sum_{i \in \mathcal{J}} \lambda_i [(x_i - a)b^T + t_i I] - \mu_1 I, \\ \frac{\partial^2 L}{\partial b^2} &= -2 \sum_{i \in \mathcal{J}} \lambda_i (x_i - a)(x_i - a)^T - 2\mu_2 I. \end{aligned}$$

The last step is the calculation of $d^T \nabla^2 L d$. After some algebra we obtain

$$(9) \quad d^T \nabla^2 L d = 2 \sum_{i \in \mathcal{J}} \lambda_i (\|\tilde{a} + \tilde{b} t_i\|^2 - (\tilde{b}^T x_i)^2).$$

By Theorem 3.2 of [1] with $\lambda_0 = 1$ we now have the following *necessary second-order optimality condition for the CLOP*:

Theorem 2. *If (r, a, b) is a local minimizer for the CLOP, then, corresponding to every critical direction $d = (\tilde{r}, \tilde{a}, \tilde{b}) = (0, \tilde{a}, \tilde{b}) \in D(r, a, b)$, multipliers $\lambda_i \geq 0$ ($i \in \mathcal{J}$) exist such that the first-order conditions (3) hold and $d^T \nabla^2 L d$ as given by (9) is nonnegative.*

Note that the complementarity statement in Ben-Tal's theorem is already implied by (8).

The following *sufficient second-order condition for the CLOP* follows from Theorem 4.1 of [1]:

Theorem 3. *Suppose that $D(r, a, b) \neq \{0\}$ and that, corresponding to every critical direction $d \in D(r, a, b)$, $d \neq 0$, multipliers $\lambda_i \geq 0$ ($i \in \mathcal{J}$) exist such that the first-order conditions (3) hold and $d^T \nabla^2 L d$ as given by (9) is positive. Then (r, a, b) is a strict local minimizer for the CLOP. In the case $D(r, a, b) = \{0\}$ the first-order conditions alone are sufficient.*

The normal situation for a given problem will be that the multipliers λ_i are uniquely determined and are all positive. In this case $\mathcal{J} = \mathcal{J}_+$ and $D(r, a, b)$ is a linear subspace of \mathbf{R}^{2n+1} (the tangential space of the manifold given by the equality constraints and the active inequality constraints). Now the second-order condition requires that the quadratic form $d^T \nabla^2 L d$ is positive semidefinite (necessary condition) or positive definite (sufficient condition) on the linear subspace $D(r, a, b)$.

Example 1 (In \mathbf{R}^3). $x_1 = (0, s, 0)$, $x_2 = (0, -s, 0)$, $x_3 = (1, 0, s)$, $x_4 = (1, 0, -s)$ with a parameter $s > 0$. For $a = 0$, $b = e_1 = (1, 0, 0)^T$ the first-order conditions (3) are satisfied with $\mathcal{J} = \{1, 2, 3, 4\}$, $t_1 = t_2 = 0$, $t_3 = t_4 = 1$, $\lambda_1 = \lambda_2 = \lambda$, $\lambda_3 = \lambda_4 = \frac{1}{2} - \lambda$ ($0 \leq \lambda \leq \frac{1}{2}$). $d = (\tilde{r}, \tilde{a}, \tilde{b})$ is in $D(r, a, b)$ iff $\tilde{r} = 0$, $\tilde{b} = (0, \tilde{b}_2, \tilde{b}_3)^T$, $\tilde{a} = (0, 0, -\tilde{b}_3)^T$. For these d we find

$$(10) \quad d^T \nabla^2 L d = \tilde{b}_2^2(2 - 4\lambda - 4\lambda s^2) + \tilde{b}_3^2(4\lambda + 4\lambda s^2 - 2s^2).$$

For $s < 1$ and the choice $\lambda = \frac{1}{4}$ the quadratic form (10) is positive definite, hence the sufficient second-order condition is satisfied and $x(t) = te_1$ is a strict local minimizer. For $s > 1$ either \tilde{b}_2^2 or \tilde{b}_3^2 has a negative factor in (10), hence the necessary second-order condition is not satisfied and $x(t) = te_1$ is not a local minimizer.

In Theorem 3 the case $D(r, a, b) = \{0\}$ has been mentioned. In that case the CLOP has a strongly unique minimum in a corner point of the feasible set. Therefore the number of active constraints is not less than the dimension: $|\mathcal{J}| + 2 \geq 2n + 1$, hence $|\mathcal{J}| \geq 2n - 1$. If the cornerpoint is nondegenerate we have $|\mathcal{J}| = 2n - 1$. See also the discussion of the generic case at the end of Section 3. The following example shows that $|\mathcal{J}| = 5$ is possible for $n = 3$.

Example 2. $x_1 = (0, 1, 0)$, $x_2 = (-3, 0, 1)$, $x_3 = (3, 0, -1)$, $x_4 = (11, -0.8, 0.6)$, $x_5 = (-11, -0.8, -0.6)$. The line $x(t) = a + bt$ with $a = (0, 0, 0)$, $b = e_1 = (1, 0, 0)$ is a Chebyshev line. It satisfies the first-order conditions with $\mathcal{J} = \{1, 2, 3, 4, 5\}$ and multipliers $\lambda_1 = 0.2$, $\lambda_2 = \lambda_3 = 0.275$, $\lambda_4 = \lambda_5 = 0.125$.

5. An Analysis of the Second-Order Conditions

Suppose a set P of points x_1, \dots, x_m is given and $x(t)$ is a straight line for which the first-order conditions (3) hold. The active points x_i ($i \in \mathcal{J}$) lie on the surface of a cylinder around the axis $x(t)$. The inactive points x_i ($i \notin \mathcal{J}$) lie inside this cylinder. The numbers $t_i = b^T(x_i - a)$ ($i = 1, \dots, m$) can be viewed as the coordinates of the projected points $a + bt_i$ on the line $x(t)$.

Suppose the radius of the cylinder (\sqrt{r} in our notation) is small compared with $\max(t_i) - \min(t_i)$. We expect intuitively that $x(t)$ will be a Chebyshev line and the second-order conditions will be satisfied. Theorem 4 shows that this is true. We define

$$\bar{t} = \sum_{i \in \mathcal{J}} \lambda_i t_i, \quad r_0 = \left(-\|a\| + \sqrt{\|a\|^2 + \sum_{i \in \mathcal{J}} \lambda_i (t_i - \bar{t})^2} \right)^2.$$

Theorem 4. Assume the first-order conditions (3) hold and $i, j \in \mathcal{J}$ with $t_i \neq t_j$, $\lambda_i > 0$, $\lambda_j > 0$, exist. If $0 < r < r_0$, then $d^T \nabla^2 L d > 0$ for all $d \in D(r, a, b) \setminus \{0\}$.

Proof. Let $d = (0, \tilde{a}, \tilde{b}) \in D(r, a, b)$ be given. Using (7) we obtain from (9) by orthogonal decomposition

$$\frac{1}{2} d^T \nabla^2 L d = \sum_{i \in \mathcal{J}} \lambda_i \|\tilde{a} + \tilde{b} t_i - b^T \tilde{a} b\|^2 - \sum_{i \in \mathcal{J}} \lambda_i [(\tilde{b}^T x_i)^2 - (b^T \tilde{a})^2].$$

We consider the two sums on the right-hand side separately. The first sum is equal to

$$\|\tilde{a} - b^T \tilde{a} b + \tilde{b}\|^2 + \sum_{i \in \mathcal{J}} \lambda_i (t_i - \bar{t})^2 \|\tilde{b}\|^2.$$

The terms in the second sum can be estimated by Schwarz's inequality:

$$\begin{aligned} |(\tilde{b}^T x_i)^2 - (b^T \tilde{a})^2| &= |(\tilde{b}^T x_i)^2 - (\tilde{b}^T a)^2| \\ &= |\tilde{b}^T(x_i + a) \tilde{b}^T(x_i - a)| \\ &= |\tilde{b}^T(x_i - a - bt_i + 2a) \tilde{b}^T(x_i - a - bt_i)| \\ &\leq \|\tilde{b}\|^2 (\|x_i - a - bt_i\| + 2\|a\|) \|x_i - a - bt_i\| \\ &= \|\tilde{b}\|^2 (\sqrt{r} + 2\|a\|) \sqrt{r}. \end{aligned}$$

Hence

$$\frac{1}{2} d^T \nabla^2 L d \geq \|\tilde{a} - b^T \tilde{a} b + \tilde{b}\|^2 + \|\tilde{b}\|^2 \left(\sum_{i \in \mathcal{J}} \lambda_i (t_i - \bar{t})^2 - (\sqrt{r} + 2\|a\|) \sqrt{r} \right).$$

The factor of $\|\tilde{b}\|^2$ is positive iff $0 < r < r_0$. If $d \neq 0$, then by (7) $\|\tilde{b}\|$ or $\|\tilde{a} - b^T \tilde{a} b + \tilde{b}\|$ or both are nonzero. ■

6. Numerical Solution by Lipschitz Optimization

In this section we show how a global solution of the CLOP can be obtained numerically by Lipschitz optimization. As an overall reference we mention the book by Horst and Tuy [6]. We apply this technique to our original problem:

$$(11) \quad \begin{cases} \min_{a,b} \max_i \|x_i - a - bt_i\| & \text{subject to} \\ a^T b = 0, \\ b^T b - 1 = 0. \end{cases}$$

The number of variables can be conveniently halved by considering a as a function of b in the following way:

$$(12) \quad a = \arg \min_a \max_i \|y_i - a\|,$$

where $y_i = (I - bb^T)x_i$. We call such a the Chebyshev point of the points $\{y_i\}$. Note that it automatically satisfies the constraint $a^T b = 0$. In fact, the problem to determine a is to find the smallest enclosing sphere to the points $\{y_i\}$. This (convex) problem can be solved by (modified) quadratic programming (see [2]) or by an adapted simplex method (see [9]). The choice for a is motivated by the fact that if (r, a, b) is a local minimizer to (2) (hence (a, b) is one for (11)), a is the unique minimizer for (12). From now on we write $a = a(b)$. Furthermore, we define $F_i(b) = \|x_i - a(b) - bt_i\|$ and $F(b) = \max_i F_i(b)$. Our aim is thus to find a global minimizer for $F(b)$.

We proceed by the specification of the set $M \subset \mathbf{R}^n$ in which we search for a global minimizer, a Lipschitz constant for F , and a strategy for subdivision of M .

The search domain M can be restricted to (e.g.) the northern hemisphere of S^{n-1} , the $(n - 1)$ -dimensional sphere in \mathbf{R}^n :

$$M = \{b \in \mathbf{R}^n \mid b^T b = 1, b^T e_n \geq 0\},$$

where e_n is the n th coordinate vector in \mathbf{R}^n .

Of the several global optimization methods known from recent literature, Lipschitz optimization seems to be especially suited for our problem. The reason is that F has $L = \max_i \|x_i\|$ as a Lipschitz constant, which comes up naturally if we think of a change in b as a rotation and which is sharp as can be seen from easy examples.

Theorem 5. $L = \max_i \|x_i\|$ is a Lipschitz constant for F .

Proof. Let $b_1, b_2 \in M$, and suppose $F(b_1) \geq F(b_2)$. Then

$$\begin{aligned} F(b_1) - F(b_2) &= \max_i \|x_i - a(b_1) - b_1 b_1^T(x_i - a(b_1))\| \\ &\quad - \max_i \|x_i - a(b_2) - b_2 b_2^T(x_i - a(b_2))\| \\ &\leq \max_i \|x_i - a(b_2) - b_1 b_1^T x_i\| \\ &\quad - \max_i \|x_i - a(b_2) - (b_1 + b_2 - b_1)(b_1 + b_2 - b_1)^T x_i\| \end{aligned}$$

by (12). Now suppose the first maximum is attained at $i = i_0$. Then we obtain

$$\begin{aligned} F(b_1) - F(b_2) &\leq \|x_{i_0} - a(b_2) - b_1 b_1^T x_{i_0}\| \\ &\quad - \|x_{i_0} - a(b_2) - b_1 b_1^T x_{i_0}\| \\ &\quad - [b_1(b_2 - b_1)^T + (b_2 - b_1)b_1^T + (b_2 - b_1)(b_2 - b_1)^T] x_{i_0} \| \\ &\leq \|[b_1(b_2 - b_1)^T + (b_2 - b_1)b_1^T + (b_2 - b_1)(b_2 - b_1)^T] x_{i_0}\| \\ &= \|b_2 b_2^T x_{i_0} - b_1 b_1^T x_{i_0}\|. \end{aligned}$$

We introduce two orthonormal vectors

$$u = (b_1 + b_2)/\|b_1 + b_2\|, \quad v = (b_1 - b_2)/\|b_1 - b_2\|.$$

Then we can show that

$$b_2 b_2^T - b_1 b_1^T = -\frac{1}{2}\|b_1 + b_2\| \|b_2 - b_1\| (uv^T + vu^T),$$

and hence

$$\begin{aligned} \|(b_2 b_2^T - b_1 b_1^T) x_{i_0}\| &= \frac{1}{2}\|b_1 + b_2\| \|b_2 - b_1\| \sqrt{(u^T x_{i_0})^2 + (v^T x_{i_0})^2} \\ &\leq \|b_2 - b_1\| \|x_{i_0}\| \leq \|b_2 - b_1\| \max_i \|x_i\|. \end{aligned} \quad \blacksquare$$

It should be noted that, for a given point-set P , the smallest possible Lipschitz constant is obtained by translating the points such that the origin becomes the Chebyshev point of P .

The iteration is initialized by partitioning the hemisphere M into 2^{n-1} orthants. In this way we obtain a list of subsets, called $\mathcal{M}_0 = \{M_{0,1}, \dots, M_{0,2^{n-1}}\}$. Each of these subsets has n vertices, and is diffeomorphic (by projection from the origin) to the regular simplex T_{n-1} with the same vertices. In the course of the iteration the partition is refined, and at each step k we get a new list $\mathcal{M}_k = \{M_{k,1}, \dots\}$. This refinement is carried out via subdivision of the corresponding simplices. If $n = 2$ the line segment T_1 has the obvious subdivision by equidistant points. If $n = 3$ the simplex T_2 can be subdivided into four congruent regular simplices. If $n > 3$ this is no longer possible but instead we use the standard barycentric subdivision of T_{n-1} into $(n-1)!$ subsimplices. In all cases the subdivision can be repeated recursively. The triangulations of the simplices can then be transformed to obtain subdivisions of the original subsets of \mathcal{M}_k . The

diameter of such a subset is equal to its longest edge (measured on the sphere S^{n-1}).

We use the following branch and bound algorithm (see also [5] and [6]):

1. At step k we have a list \mathcal{M}_k of length p_k of subsets still of interest:
 $\mathcal{M}_k = \{M_{k,1}, \dots, M_{k,p_k}\}, M_{k,j} \subset M, j = 1, \dots, p_k$.
2. For each subset $M_{k,j}, j = 1, \dots, p_k$, evaluate F at all its vertices. (Of course, if F has already been calculated at a vertex this old information can be used.)
3. For each subset $M_{k,j}$, upper and lower bounds $\alpha_{k,j}$ and $\beta_{k,j}$ for the minimum of F restricted to $M_{k,j}$ are given by

$$\alpha_{k,j} = \min_{v_{k,j}} F(v_{k,j}), \quad j = 1, \dots, p_k,$$

and

$$\beta_{k,j} = \max_{v_{k,j}} F(v_{k,j}) - L \text{ diameter}(M_{k,j}), \quad j = 1, \dots, p_k,$$

where $v_{k,j}$ runs over all vertices of $M_{k,j}$.

4. Upper and lower bounds α and β for the global minimum are $\alpha_k = \min_j \alpha_{k,j}$, and $\beta_k = \min_j \beta_{k,j}$. If $\alpha_k - \beta_k < \varepsilon$ or $\max_j \text{diameter}(M_{k,j}) < \delta$, stop. Otherwise go to step 5. (ε and δ are positive constants.)
5. Delete all $M_{k,j}$ from \mathcal{M}_k for which $\alpha_k < \beta_{k,j}$. Subdivide all remaining sets into smaller subsets. The resulting list is \mathcal{M}_{k+1} . Go to step 1.

For low dimensions ($n = 2, 3$) this algorithm works quite satisfactorily. However, when the dimension increases, it becomes rather time consuming. In practice this algorithm has been used in combination with a local search method: first a number of steps of the above algorithm are performed. The resulting best approximation to the global minimizer is used as a starting point for the SQP method, as implemented by Powell in a subroutine available from the

Table 1. Comparison (for dimensions $n = 2$ and $n = 3$) between Lipschitz optimization combined with the SQP method and the SQP method alone.*

	$n = 2$			$n = 3$		
	$l = 1$	$l = 2$	$l = 3$	$l = 1$	$l = 2$	$l = 3$
$m = 10$	7	5	1	15	4	1
$m = 20$	14	4	1	10	6	0
$m = 50$	16	3	2	17	2	0

* Both methods are used to find the Chebyshev-line for 25 samples of the indicated number of points and value for l . The number of samples in which the combined method gives a better approximation to the global minimizer is given. In other words: the SQP method alone converged in these cases to a critical point different from the global minimizer.

Harwell library (see [7]). We compared this two-stage method with the SQP method alone. In the test the points x_i are uniformly distributed in the n -dimensional block C obtained as the Cartesian product of an $(n - 1)$ -dimensional hypercube with an interval of length l . If l is large, a good approximation to a minimizer of the CLOP can easily be obtained, and the SQP method alone performs well. However, if $l \approx 1$, the two-stage method is much better, as is indicated in Table 1.

References

1. A. BEN-TAL (1980): *Second-order and related extremality conditions in nonlinear programming*. J. Optim. Theory Appl., **31**:143–165.
2. D. J. ELZINGA, D. W. HEARN (1972): *The minimum covering sphere problem*. Management Sci., **19**(1):96–104.
3. P. GRITZMANN, V. KLEE (1992): *Inner and outer j -radii of convex bodies in finite-dimensional normed spaces*. Discrete Comput. Geom., **7**:255–280.
4. R. HETTICH, H. TH. JONGEN (1976): *On first- and second-order conditions for local optima for optimization problems in finite dimensions*. Methods Oper. Res., **13**:82–94.
5. R. HORST, H. TUY (1987): *On the convergence of global methods in multiextremal optimization*. J. Optim. Theory Appl., **54**(2):253–271.
6. R. HORST, H. TUY (1990): *Global Optimization*. Berlin: Springer-Verlag.
7. M. J. D. POWELL (1982): *Extensions to subroutine VF02AD*. In: System Modeling and Optimization (R. F. Drenick, F. Kozin, eds.). Lecture Notes in Control and Information Sciences, vol. 38. Berlin: Springer-Verlag, pp. 529–538.
8. H. SPÄTH (1986): *Orthogonal least squares fitting with linear manifolds*. Numer. Math., **48**:441–445.
9. W. WETTERLING (1988): *Chebyshev Approximation by Parametrized Curves*. Memorandum No. 756, University of Twente, Enschede.
10. W. WETTERLING (1991): *Application of Chebyshev approximation to curve fitting*. Z. Angew. Math. Mech., **71**(7–8):291–292.

M. Streng
 Department of Applied Mathematics
 University of Twente
 P.O. Box 217
 7500 AE Enschede
 The Netherlands

W. Wetterling
 Department of Applied Mathematics
 University of Twente
 P.O. Box 217
 7500 AE Enschede
 The Netherlands