Computational Optimization and Applications, 20, 267–279, 2001

Smoothing Method for Minimax Problems*

SONG XU song.xu@latticesemi.com
Lattice Semiconductor Corporations, 2680 Zanker Road, San Jose, CA 95134-2100, USA

Received July 21, 1999; Accepted February 7, 2000

Abstract. In this paper, we propose a smoothing method for minimax problem. The method is based on the exponential penalty function of Kort and Bertsekas for constrained optimization. Under suitable condition, the method is globally convergent. Preliminary numerical experiments indicate the promising of the algorithm.

Keywords: minimax, smoothing method, global convergence

1. Introduction

We consider the minimax optimization problem

$$\text{minimize } f(x), \tag{1}$$

where

$$f(x) = \max_{i=1,\dots,m} f_i(x). \tag{2}$$

We assume in this paper that the functions $f_1(x), \dots, f_m(x) : \mathbb{R}^n \to \mathbb{R}$ are twice continuous differentiable. Minimax problems arise in engineering design [23, 31], computer-aided-design [25] and circuit design [29] and optimal control [27, 28]. For a complete treatment of the minimax problems, see the books [11, 12, 24].

The objective function $f(x)$ has discontinuous first partial derivatives at points where two or more of the functions $f_i(x)$ are equal to $f(x)$ even if each $f_i(x)$ $(i = 1, \dots, m)$ has continuous first partial derivatives. To handle this difficulty, various methods have been proposed for minimax problems. See for example, Warren et al. [31], Osborne and Watson [21], Charalamous and Bandler [8], Charalamous and Conn [9], and Murray and Overton [20].

In this paper, we propose a smoothing method for solving the minimax problem. The method uses the smoothing function

$$f(x, \mu) = \mu \ln \sum_{i=1}^{m} \exp\left(\frac{f_i(x)}{\mu}\right), \tag{3}$$

to approximate the function $f(x)$. The function $f(x, \mu)$ provides a good approximation to

*Research supported in part by research grants from NSERC of Canada.

the function $f(x)$ in the sense that

$$f(x) \le f(x,\mu) \le f(x) + \mu \ln m;$$

for $\mu > 0$ (see proposition 2.2). The function is sometimes called the exponential penalty function [3, 4, 5, 17] or aggregate function [18], and has been used to in solving nonlinear programming problems [3, 4, 5, 17], the generalized linear complementarity problems [22] and the generalized nonlinear complementarity problems [19].

The proposed method solves a quadratic approximation of $f(x,\mu)$ for a decreasing sequence of μ. A line search procedure based on the merit function $f(x,\mu)$ is used to ensure the global convergence of the algorithm. The algorithm bears some similarity to the recently proposed smoothing methods for complementarity Problems [6, 7, 10, 16, 26, 30], and for mathematical programming with equilibrium constraints [13, 14],

The plan of the paper is as follows. In Section 2, we summarize some basic properties of $f(x,\mu)$. The algorithm and its global convergence is given in Section 3. We present some preliminary numerical results in Section 4. Finally, we end this paper by some remarks.

2. The smoothing function

Recall that x^* is a stationary point to the minimax problem (1) if there exists a vector $y^* = (y_1^*, \dots, y_m^*)$ such that

$$\sum_{j=1}^{m} y_j^* \nabla f_j(x^*) = 0, \quad y^* \ge 0, \quad \sum_{j=1}^{m} y_j^* = 1, \tag{4}$$

$$y_j^* = 0 \quad \text{if} \;\; f_j(x^*) < \max\{f_1(x^*), \dots, f_m(x^*)\}. \tag{5}$$

Proposition 2.1 [11]. *If x^* is a local minimum to the minimax problem* (1), *then x^* is a stationary point satisfying* (4) *and* (5). *Conversely, assume that $f(x)$ is convex, then if x^* is a stationary point, x^* is a global minimum to the minimax problem.*

We now list some properties of the smoothing function $f(x,\mu)$.

Proposition 2.2 [22]. *Suppose that $f_i(x)$ $(i = 1, \dots, m)$ are twice continuous differentiable, $f(x)$ and $f(x,\mu)$ are defined in* (2) *and* (3) *respectively, then*

(i) *$f(x,\mu)$ is increasing with respect to μ, and $f(x) \le f(x,\mu) \le f(x) + \mu \ln m$.*
(ii) *$f(x,\mu)$ is twice continuous differentiable for all $\mu > 0$, and*

$$\nabla_x f(x,\mu) = \sum_{i=1}^{m} \lambda_i(x,\mu) \nabla f_i(x), \tag{6}$$

$$\begin{aligned}\nabla_x^2 f(x,\mu) = &\sum_{i=1}^{m} \left(\lambda_i(x,\mu) \nabla^2 f_i(x) + \frac{1}{\mu} \lambda_i(x,\mu) \nabla f_i(x) \nabla f_i(x)^T \right) \\ &- \frac{1}{\mu} \left(\sum_{i=1}^{m} \lambda_i(x,\mu) \nabla f_i(x) \right) \left(\sum_{i=1}^{m} \lambda_i(x,\mu) \nabla f_i(x) \right)^T,\end{aligned} \tag{7}$$

where

$$\lambda_i(x,\mu) = \frac{\exp(f_i(x)/\mu)}{\Sigma_{j=1}^m \exp(f_j(x)/\mu)} \in (0,1), \quad \sum_{i=1}^m \lambda_i(x,\mu) = 1. \tag{8}$$

(iii) *For any* $x \in \mathbb{R}^n$ *and* $\mu > 0$, *it holds that* $0 \le f'_\mu(x,\mu) \le \ln m$.

We now give a condition under which the matrix $\nabla_x^2 f(x,\mu)$ is positive definite.

Lemma 2.1. *Suppose that* $f_i(x)$ $(i = 1,\ldots,m)$ *are twice continuous differentiable convex function with at least one of the functions having positive definite Hessian, then the matrix* $\nabla_x^2 f(x,\mu)$ *is positive definite for* $x \in \mathbb{R}^n$ *and* $\mu > 0$.

Proof: Let

$$\begin{aligned} A &= [\nabla f_1(x),\ldots,\nabla f_m(x)],\\ \Lambda &= \mathrm{diag}(\lambda_i(x,\mu)),\\ \Upsilon &= \begin{bmatrix} \lambda_1(x,\mu)\\ \vdots \\ \lambda_m(x,\mu)\end{bmatrix} [\lambda_1(x,\mu),\ldots,\lambda_m(x,\mu)]. \end{aligned}$$

Then

$$\begin{aligned} &\frac{1}{\mu}\left(\sum_{i=1}^m \lambda_i(x,\mu)\nabla f_i(x)\nabla f_i(x)^T - \left(\sum_{i=1}^m \lambda_i(x,\mu)\nabla f_i(x)\right)\left(\sum_{i=1}^m \lambda_i(x,\mu)\nabla f_i(x)\right)^T\right)\\ &\quad = \frac{1}{\mu}(A\Lambda A^T - A\Upsilon A^T) = \frac{1}{\mu}A(\Lambda - \Upsilon)A^T = \frac{1}{\mu}ADA^T, \end{aligned}$$

where $D = \Lambda - \Upsilon$. By Gersgorin theorem [15], the eigenvalues of D are located in the union of m discs

$$\bigcup_{i=1}^m \{z \in C : |z - d_{ii}| \le R_i(D)\},$$

where for $i = 1,\ldots,m$,

$$\begin{aligned} d_{ii} &= \lambda_i(x,\mu) - \lambda_i^2(x,\mu),\\ R_i(D) &= \sum_{j=1, j\ne i}^m \lambda_i(x,\mu)\lambda_j(x,\mu). \end{aligned}$$

By (8), we have

$$\begin{aligned} R_i(D) &= \sum_{j=1, j\neq i}^{m} \lambda_i(x,\mu)\lambda_j(x,\mu) \\ &= \sum_{j=1}^{m} \lambda_i(x,\mu)\lambda_j(x,\mu) - \lambda_i^2(x,\mu) \\ &= \lambda_i(x,\mu)\sum_{j=1}^{m} \lambda_j(x,\mu) - \lambda_i^2(x,\mu) \\ &= \lambda_i(x,\mu) - \lambda_i^2(x,\mu) \\ &= d_{ii}. \end{aligned}$$

So each eigenvalue of $\Lambda - \Upsilon$ is nonnegative, and therefore the matrix $\frac{1}{\mu}ADA^T$ is positive semi-definite. Since $\lambda_i(x,\mu) > 0$ and at least one of $\nabla^2 f_i(x)$ is positive definite, the matrix $\nabla_x^2 f(x,\mu)$ is positive definite. □

3. The algorithm and its convergence

The Algorithm

Step 0 (Initialization)
Choose x^0, $\mu_0 > 0$, $\beta \in (0,1)$, $\sigma \in (0,1)$ and $\rho \in (0,1)$. Let $k = 0$.
Step 1 (Direction Generation)
Solve Δx^k from the equation

$$\nabla_x^2 f(x^k,\mu_k)\Delta x^k = -\nabla_x f(x^k,\mu_k). \tag{9}$$

Step 2 (Step Size Determination)
If $\Delta x^k = 0$ set $x^{k+1} = x^k$, otherwise perform Armijo line search to find the smallest nonnegative integer $l > 0$ such that

$$f(x^k + \rho^l \Delta x^k, \mu_k) \leq f(x^k,\mu_k) + \sigma\rho^l(\nabla_x f(x^k,\mu_k))^T \Delta x^k. \tag{10}$$

Let $x^{k+1} = x^k + \rho^l \Delta x^k$.
Step 3 (Termination Check)
Terminate if a prescribed stopping rule is satisfied, otherwise let $\mu_{k+1} = \beta\mu_k$ and return to step 1 with k replaced by $k+1$.

We make the following assumptions

Assumption (A): The function $f(x)$ defined in (2) is coercive, that is, $f(x) \to +\infty$ as $\|x\| \to +\infty$.

Assumption (B): There exist constants $c_1, c_2 > 0$ such that

$$c_1\|z\|^2 \le z^T \nabla_x^2 f(x,\mu) z \ \le\ c_2\|z\|^2,$$

for all $z \in I\!R^n$ and $(x,\mu) \in \{(x,\mu) \mid f(x) \le f(x^0,\mu_0),\ 0 < \mu \le \mu_0\}$.

We now show that the algorithm is well-defined under the condition (B).

Proposition 3.1. *Assume that the condition* (B) *holds, then the algorithm is well-defined.*

Proof: From condition (B), the matrix $\nabla_x^2 f(x,\mu)$ is positive definite. So the equation (9) is well-defined and has a unique solution.

From (9), we have

$$(\nabla_x f(x,\mu))^T \Delta x^k = -(\Delta x^k)^T \nabla_x^2 f(x,\mu) \Delta x^k < 0,$$

and so Δx_k is a descent direction of $f(x,\mu)$ whenever $\Delta x^k \neq 0$ and therefore the Armijo line search procedure is finite terminating. □

Lemma 3.1. *Assume that conditions* (A) *and* (B) *hold. Let the sequences* $\{x^k\}$ *and* $\{\Delta x^k\}$ *be generated by the algorithm, then*

(a) *the sequence* $\{f(x^k,\mu^k)\}$ *is monotonically decreasing*,
(b) *the sequence*$\{f(x^{k+1},\mu_k)\}$ *and* $\{f(x^k,\mu_k)\}$ *are both convergent and have the same limit,*
(c) *the sequence* $\{x^k\}$ *is bounded, and*
(d) *the sequence* $\{\Delta x^k\}$ *is bounded.*

Proof: (a) From (10), we have

$$f(x^{k+1},\mu_k) \le f(x^k,\mu_k).$$

By part (i) of Proposition 2.2, $f(x,\mu)$ is increasing with respect to μ, So $f(x^{k+1},\mu_{k+1}) \le f(x^{k+1},\mu_k)$. Therefore

$$f(x^{k+1},\mu_{k+1}) \le f(x^{k+1},\mu_k) \le f(x^k,\mu_k)$$

for all k and the sequence $\{f(x^k,\mu_k)\}$ is monotonically decreasing.

(b) From Taylor expansion for μ, we have

$$f(x^{k+1},\mu_{k+1}) = f(x^{k+1},\mu_k) + f'_\mu(x^{k+1},\theta\mu_k)(\mu_{k+1}-\mu_k),$$

here $0 \le \theta \le 1$. So

$$f(x^{k+1},\mu_{k+1}) \le f(x^k,\mu_k) + |f'_\mu(x^{k+1},\theta\mu_k)|(1-\beta)\mu_k.$$

By (iii) of Proposition (2.2), $|f'_\mu(x^{k+1}, \theta\mu_k)| \le \ln m$. So

$$f(x^{k+1}, \mu_{k+1}) \le f(x^{k+1}, \mu_k) + (1-\beta)\mu_k \ \ln m,$$
$$f(x^{k+1}, \mu_{k+1}) - (1-\beta)\mu_k \ \ln m \le f(x^{k+1}, \mu_k) \le f(x^k, \mu_k).$$

Since $\mu_k \to 0$ as $k \to \infty$, we have that the sequence $\{f(x^{k+1}, \mu_k)\}$ and $\{f(x^k, \mu_k)\}$ are both convergent and have the same limit.

(c) By assumption (A), the set $\{x \mid f(x) \le f(x^0, \mu_0)\}$ is bounded. Since

$$x^k \in \{x \mid f(x, \mu) \le f(x^0, \mu_0)\} \subseteq \{x \mid f(x) \le f(x^0, \mu_0)\},$$

the sequence $\{x^k\}$ is bounded.

(d) By assumption (B) and (9),

$$\begin{aligned}(\Delta x_k)^T \nabla_x^2 f(x, \mu) \Delta x^k &= -\left(\Delta x^k\right)^T \nabla_x f(x^k, \mu_k),\\ c_1 \|\Delta x_k\|^2 &\le \|\Delta x^k\| \|\nabla_x f(x^k, \mu_k)\|,\\ \|\Delta x_k\| &\le \frac{1}{c_1} \|\nabla_x f(x^k, \mu_k)\|.\end{aligned}$$

By (c), $\{\|\nabla_x f(x^k, \mu_k)\|\}$ is bounded and therefore $\{\Delta x_k\}$ is bounded. □

We are now ready to prove our main convergence result.

Theorem 3.1. *Assume that conditions* (*A*) *and* (*B*) *hold. Let* $\{x^k\}$ *be the sequence generated by the algorithm. Then*

(1) *the sequence has a limit point*, *and*
(2) *every limit point is a stationary point of the minimax problem.*

Proof: (1) By (c) of Proposition 2.2, the sequence $\{x^k\}$ is bounded. So it has a limit point.

(2) Let $\bar{x}$ be a limit point of $\{x^k\}$. Without loss of generality, we assume that x^k converges to $\bar{x}$. By (10) and condition (B), we have

$$f(x^{k+1}, \mu_k) \le f(x^k, \mu_k) - c_1 \sigma \tau_k \|\Delta x^k\|^2.$$

By part (b) of Lemma 3.1, both $f(x^{k+1}, \mu_k)$ and $f(x^k, \mu_k)$ converge to the same limit, so

$$\lim_{k\to\infty} \tau_k \|\Delta x^k\| = 0.$$

We claim that

$$\lim_{k\to\infty} \|\Delta x^k\| = 0.$$

This clearly holds if $\lim_{k\to\infty}\inf\tau_k > 0$. Suppose $\lim_{k\to\infty}\inf\tau_k = 0$, without loss of generality, assume

$$\lim_{k\to\infty}\tau_k = 0.$$

By the boundedness of $\{\Delta x^k\}$, we may further assume that the sequence $\{\Delta x^k\}$ converges to $\widetilde{\Delta x}$. Since $\tau_k = \rho^{l_k}$ and $\rho \in (0,1)$,

$$\lim_{k\to\infty} l_k = \infty.$$

By the definition of l_k, we have

$$f(x^k + \rho^{l_k-1}\Delta x^k, \mu_k) > f(x^k, \mu_k) + \sigma\rho^{l_k-1}(\nabla_x f(x^k,\mu_k))^T\Delta x^k,$$
$$\frac{f(x^k + \rho^{l_k-1}\Delta x^k, \mu_k) - f(x^k,\mu_k)}{\rho^{l_k-1}} \geq \sigma(\nabla_x f(x^k,\mu_k))^T\Delta x^k.$$

By the mean value theorem, we have

$$(\nabla_x f(x^k + t_k\rho^{l_k-1}\Delta x^k, \mu_k))^T\Delta x^k \geq \sigma(\nabla_x f(x^k,\mu_k))^T\Delta x^k,$$

where $0 \leq t_k \leq 1$. Taking limit in the above equation, we obtain

$$(\nabla_x f(\bar{x}, 0^+))^T\widetilde{\Delta x} \geq \sigma(\nabla_x f(\bar{x}, 0^+))^T\widetilde{\Delta x}, \tag{11}$$

where $\nabla_x f(\bar{x}, 0^+) := \lim_{\mu\to 0^+}\nabla_x f(\bar{x},\mu)$. By (9) and condition (B), we have

$$\begin{aligned}(\nabla_x f(x^k,\mu_k))^T\Delta x^k &= -(\Delta x^k)^2\nabla_x^2 f(x^k,\mu_k)\Delta x^k\\ &\leq -c_1\|\Delta x^k\|^2.\end{aligned}$$

Taking the limit, we have

$$(\nabla_x f(\bar{x}, 0^+))^T\widetilde{\Delta x} \leq -c_1\|\widetilde{\Delta x}\|^2. \tag{12}$$

Combining (11) and (12), we have $\widetilde{\Delta x} = 0$. It follows from (9) that

$$\nabla_x f(\bar{x}, 0^+) = 0.$$

Let $B(\bar{x}) = \{i \in \{1,\ldots,m\} \mid f_i(\bar{x}) = f(\bar{x})\}$. Then

$$\lim_{\mu\to 0}\exp\left(\frac{f_i(\bar{x}) - f(\bar{x})}{\mu}\right) = 0, \quad i \notin B(\bar{x}),$$
$$\lambda_i(\bar{x}, 0^+) = 0 \text{ if } i \notin B(\bar{x}).$$

So

$$\nabla_x f(\bar{x}, 0^+) = \sum_{i=1}^{n} \lambda_i(\bar{x}, 0^+) \nabla f_i(\bar{x}) = 0,$$

where $\sum_{j=1}^{m} \lambda_i(\bar{x}, 0^+) = 1$ and $\lambda_i(\bar{x}, 0^+) = 0$ for $i \notin B(\bar{x})$. So $\bar{x}$ is a stationary point of the minimax problem. □

We now give a condition that ensures condition (B). We need the following definition.

Definition 3.1. Let $f : R^n \to R^n$ be continuously differentiable and let α be a positive scalar. If f satisfies the condition

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq \alpha \|x - y\|^2, \quad \text{for any } x, y \in R^n, \tag{13}$$

then f is called strongly convex.

It well known that if f is twice continuously differentiable, then the condition (13) is equivalent to the positive semidefiniteness of $\nabla^2 f(x) - \alpha I$ for every $x \in R^n$, where I is the identity matrix.

Theorem 3.2. *Assume that condition* (A) *holds. Let* $\{x^k\}$ *be the sequences generated by the algorithm and* $\bar{x}$ *be the limit of an arbitrary convergent subsequence of* $\{x^k\}$. *If* $f_i(x)$ $(i = 1, \ldots, m)$ *are strongly convex, then* $\bar{x}$ *is a solution to the minimax problem.*

Proof: By Theorem 3.1, we need only to prove that condition (B) holds. From condition (A), the set

$$\mathcal{C} := \{x \mid f(x) \leq f(x^0, \mu_0)\}$$

is bounded. Since $f_i(x)(i = 1, \ldots, m)$ are strongly convex, there exist $c_1 > 0$ and $c_2 > 0$ such that

$$c_1 \|z\|^2 \leq z^T \nabla^2 f_i(x) z \leq c_2 \|z\|^2$$

for $x \in \mathcal{C}$ and $i = 1, \ldots, m$. By Lemma 2.1 and (7), there exist $c_1' > 0$ and $c_2' > 0$ such that

$$c_1' \|z\|^2 \leq z^T \nabla_x^2 f(x, \mu) z \leq c_2' \|z\|^2$$

for all $x \in \mathcal{C}$ and $0 \leq \mu \leq \mu_0$. □

4. Numerical results

In this section, we present numerical results for the algorithm. The parameters are chosen to be $\beta = 0.5, \mu_0 = 10^2, \rho = 0.8, \sigma = 0.1, \epsilon = 10^{-6}$. We use the following stopping criteria in our implementation

$$|(\nabla_x f(x^k, \mu_k))^T \Delta x^k| \le \epsilon.$$

Since $\exp(\frac{f_i(x)}{\mu})$ tends to be large when μ approaches 0, to prevent overflow, special care has been taken in computing $f(x^k, \mu_k)$ and $\lambda_i(x^k, \mu_k)$.

$$\begin{aligned} f(x^k, \mu_k) &= \mu_k \ln \sum_{i=1}^{m} \exp\left(\frac{f_i(x^k)}{\mu_k}\right) \\ &= \mu_k \ln \exp \frac{f(x^k)}{\mu_k} \sum_{i=1}^{m} \exp\left(\frac{f_i(x^k) - f(x^k)}{\mu_k}\right) \\ &= f(x^k) + \mu_k \ln \sum_{i=1}^{m} \exp\left(\frac{f_i(x^k) - f(x^k)}{\mu_k}\right). \end{aligned}$$

$$\begin{aligned} \lambda_i(x^k, \mu_k) &= \frac{\exp\left(\frac{f_i(x^k)}{\mu_k}\right)}{\sum_{j=1}^{m} \exp\left(\frac{f_j(x^k)}{\mu_k}\right)} \\ &= \frac{\exp\left(\frac{f_i(x^k)-f(x^k)}{\mu_k}\right)}{\sum_{j=1}^{m} \exp\left(\frac{f_j(x^k)-f(x^k)}{\mu_k}\right)}. \end{aligned}$$

Example 1 [9].

$$\text{minimize}_{i=1,2,3}\ f_i(x),$$

where

$$\begin{aligned} &f_1(x) = x_1^2 + x_2^4, \\ &f_2(x) = (2 - x_1)^2 + (2 - x_2)^2, \\ &f_3(x) = 2\ \exp(-x_1 + x_2). \end{aligned}$$

The initial point is chosen to be $x_0 = (1, -0.1)$. After 22 iterations, we find solution

$$x_1 = 1.1390,\ \ x_2 = 0.8996.$$

Example 2 [9].

$$\text{minimize}_{i=1,2,3}\ f_i(x),$$

where

$$f_1(x) = x_1^4 + x_2^2,$$
$$f_2(x) = (2 - x_1)^2 + (2 - x_2)^2,$$
$$f_3(x) = 2\ \exp(-x_1 + x_2).$$

The initial point is chosen to be $x_0 = (1, -0.1)$. After 25 iterations, we find solution

$$x_1 = 1.0000,\ x_2 = 1.0000.$$

Consider the following nonlinear programming problem:

$$\begin{aligned}&\text{minimize } F(x)\\&\text{subject to } g_i(x) \geq 0 \quad i = 2, 3, \ldots, m.\end{aligned}$$

Consider the following minimax function:

$$f(x) = \max_{1 \leq i \leq m} f_i(x),$$

where

$$f_1(x) = F(x), \quad f_i(x) = F(x) - \alpha_i g_i(x), \quad 2 \leq i \leq m, \quad \alpha_i > 0, \quad 2 \leq i \leq m.$$

Bandler and Charalambous [2] proved that for sufficiently large α_i the optimum of the minimax function coincides with that of the nonlinear programming problem.

Example 3. (Rosen-Suzuki Problem).

$$F(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4,$$
$$g_2(x) = -x_1^2 - x_2^2 - x_3^3 - x_4^2 - x_1 + x_2 - x_3 + x_4 + 8,$$
$$g_3(x) = -x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 + 10,$$
$$g_4(x) = -x_1^2 - x_2^2 - x_3^2 - 2x_1 + x_2 + x_4 + 5.$$

We use

$$\alpha_2 = \alpha_3 = \alpha_4 = 10.$$

The solution is

$$x_1 = 0, x_2 = 1, x_3 = 2, x_4 = -1.$$

After 25 iterations, we find the solution

$$x_1 = 0.0000, x_2 = 1.0000, x_3 = 2.0000, x_4 = -1.0000.$$

Example 4 [1].

$$\begin{aligned}
F(x) &= (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 + 10x_5^6 + 7x_6^2 + x_7^4 \\
&\quad - 4x_6x_7 - 10x_6 - 8x_7, \\
g_2(x) &= -2x_1^2 - 3x_3^4 - x_3 - 4x_4^2 - 5x_5 + 127, \\
g_3(x) &= -7x_1 - 3x_2 - 10x_3^2 - x_4 + x_5 + 282, \\
g_4(x) &= -23x_1 - x_2^2 - 6x_6^2 + 8x_7 + 196, \\
g_5(x) &= -4x_1^2 - x_2^2 + 3x_1x_2 - 2x_3^2 - 5x_6 + 11x_7.
\end{aligned} \tag{14}$$

We use

$$\alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 10.$$

The solution is

$$\begin{aligned}
&x_1 = 2.3305, x_2 = 1.9514, x_3 = -0.47754, \\
&x_4 = 4.3657, x_5 = -0.62449, x_6 = 1.0381, x_7 = 1.5942.
\end{aligned}$$

The initial point

$$x^0 = (1, 2, 0, 4, 0, 1, 1).$$

After 27 iterations, we find the solution

$$\begin{aligned}
&x_1 = 2.3305, x_2 = 1.9514, x_3 = -0.4775, x_4 = 4.3657, \\
&x_5 = -0.6245, x_6 = 1.0381, x_7 = 1.5942.
\end{aligned}$$

5. Conclusion remark

In this paper, we propose a smoothing method for minimax problem and and prove that the method is globally convergent. Preliminary numerical experiments indicate the promising of the algorithm. However, it remains an open question what is the local convergence behavior of the method.

Acknowledgment

The author would like to thank the referees for their helpful suggestions.

References

1. J. Asaadi, "A computational comparison of some non-linear programs," Math. Programming, 1973.

2. J.W. Bandler and C. Charalambous, "Nonlinear programming using minimax techniques," J. Optimization Theory Appl., vol. 13, pp. 607–619, 1974.
3. D.P. Bertsekas, "Minimax methods based on approximation," in Proceeding of the 1976 Johns Hopkins Conference on Information Sciences and Systems, 1976.
4. D.P. Bertsekas, "A new algorithm for solution of nonlinear resistive networks involving diodes," IEEE Transactions on Circuit Theory, vol. CAS 23, pp. 599–608, 1976.
5. D.P. Bertsekas, "Approximation Procedures based on the method of multipliers," J. Optimization Theory and Applications, vol. 23, pp. 487–510, 1977.
6. J. Burke and S. Xu, "A non-interior predictor-corrector path following algorithm for the monotone linear complementarity problem," Mathematical Programming, vol. 87, pp. 113–130, 2000.
7. J. Burke and S. Xu, "The global linear convergence of a non-interior path-following algorithm for linear complementarity problem," Mathematics of Operations Research, vol. 23, pp. 719–734, 1998.
8. C. Charalamous and J.W. Bandler, "Nonlinear minimax optimization as a sequence of least pth optimization with finite values of p," Internat. J. Systems Sci., vol. 7, pp. 377–391, 1976.
9. C. Charalamous and A.R. Conn, "An efficient method to solve the minimax problem directly," SIAM J. Numer. Anal, vol. 15, pp. 162–187, 1978.
10. B. Chen and N. Xiu, "A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing functions," SIAM J. Optimization, Vol. 9, pp. 605–623, 1999.
11. V.F. Demyanov and V.N. Molozemov, Introduction to Minimax, Wiley: New York, 1974.
12. D.Z. Du and P.M. Pardalos (eds), Minimax and Applications, Kluwer Academic Publishers: Dordrecht, 1995.
13. F. Facchinei, H. Jiang and L. Qi, "A smoothing method for mathematical programs with equilibrium constraints," Mathematical Programming, vol. 85, pp. 107–134, 1999.
14. M. Fukushima, Z.-Q. Luo and J.-S. Pang, "A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints," Comput. Optim. Appl., vol. 10, pp. 5–34, 1998.
15. R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
16. K. Hotta and A. Yoshise, "Global convergence of a class of non-interior-point algorithms using Chen-Harker-Kanzow functions for nonlinear complementarity problems," Mathematical Programming, Vol. 86, pp. 105–133, 1999.
17. B.W. Kort and D.P. Bertsakas, "A new penalty function algorithm algorithm for constrained minimization," in Proceedings of the 1972 IEEE conferences on Decision and Control, New Orlean, Louisiana, 1972.
18. X.S. Li, "An aggregate function method for nonlinear programming," Science in China (A), vol. 34, pp. 1467–1473, 1991.
19. Z.-H. Lin and Hou-Duo Qi, "A non-interior homotopy system for generalized nonlinear complementarity problem," Preprint of the State key laboratory of scientific and engineering computing, Academic Sinica, Beijing, China, 1998.
20. W. Murray and M.L. Overton, "A projected Lagrangian algorithm for nonlinear minimax optimization," SIAM Journal on Scientific and Statistic Computing, vol. 1, pp. 345–370, 1980.
21. M.R. Osborne and G.A. Watson, "An algorithm for minimax approximation in the non-linear case," Comput. J., vol. 12, pp. 63–68, 1969.
22. J.-M. Peng and Z. Lin, "A non-interior continuation method for generalized linear complementarity problems," Mathematical Programming, Vol. 86, pp. 533–563, 1999.
23. E. Polak, "On the mathematical foundations of nondifferentiable optimization," SIAM Review, vol. 29, pp. 21–89, 1987.
24. E. Polak, Optimization: Algorithm and Consistent Approximations, Springer Verlag: New York, 1997.
25. E. Polak and J.E. Higgins and D.Q. Mayne, "A barrier function method for minimax problems," Mathematical Programming, vol. 64, pp. 277–294, 1994.
26. L. Qi and D. Sun and G. Zhou, "A new look at smoothing Newton methods for non-linear complementarity problems and box constrained variational inequalities," Mathematical Programming, vol. 87, pp. 1–35, 2000.
27. R.T. Rockafellar, "Linear-quadric programming and optimal control," SIAM Journal on Control and Optimization, vol. 25, pp. 781–814, 1987.

28. R.T. Rockafellar, "Computational schemes for large-scale problems in extended linear-quadratic programming," Mathematical Programming, vol. 48, pp. 447–474, 1990.
29. S.E. Sussman-Fort, "Approximate direct-search minimax circuit optimization," International Journal for Numerical Methods in Engineering, vol. 28, pp. 359–368, 1989.
30. P. Tseng, "Analysis of a non-interior continuation method based on chen-mangasarian smoothing functions for complementarity problems," in *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Masao Fukushima and Liqun Qi, (Eds.), Kluwer Academic. Publishers: Dordrecht, 1999, pp. 381–404.
31. A.D. Warren, L.S. Lasdon and D.F. Suchman, "Optimization in engineering design," Proc. IEEE, vol. 55, pp. 1885–1897, 1967.