International Journal of Pure and Applied Mathematics
Volume 35 No. 2 2007, 277-290

A MODIFIED LAGRANGIAN FUNCTION ALGRITHM FOR MINIMAX PROBLEMS WITH INEQUALITY CONSTRAINTS

Su-Xiang He[1 §], Yan-Li Zhao[2]

[1]School of Science
Wuhan University of Technology
Wuhan, 430070, P.R. CHINA
e-mail: suxianghe@163.com

[2]Wuhan National Laboratory for Optoelectronics
Huazhong University of Science and Technology
Wuhan, 430074, P.R. CHINA
e-mail: wtdzhaoyl@163.com

Abstract: A modified Lagrangian function (MLF) algorithm for solving minimax problems with inequality constraints, is presented. The algorithm alternatively minimizes the MLF in the primal space, and updates the Lagrange multipliers and controlling parameter. Under the mild conditions, the MLF algorithm is shown to converge Q-superlinearly and the error bounds of solution is also established. Finally the numerical results for several problems are reported.

AMS Subject Classification: 90C47, 49K35
Key Words: modified Lagrangian function algorithm, minimax problems, inequality constraints, numerical results

1. Introduction

In this paper we consider the following constrained minimax problems:

$$\begin{array}{l} \min F(x) = \max_{1\leq i\leq m} f_i(x)\,, \\ \text{s.t.} \quad g_j(x) \leq 0, \quad j = 1,\ldots,l, \end{array} \tag{1}$$

where $f_i : I\!R^n \longmapsto I\!R$ $(i = 1,\ldots,m)$, $g_j : I\!R^n \longmapsto I\!R$ $(j = 1,\ldots,l)$ are smooth real-valued functions. Evidently, the objective function $\max_{1\leq i\leq m} f_i(x)$ is non-

Received: December 27, 2006
§Correspondence author

smooth, i.e., the problem (1) is a nonsmooth optimization. Many applications in engineering, system analysis, management science, etc., exist for problem (1), for instance, see Ben-Tal and Nemirovsky [1], Erdmann [6] and Polak [9]. There are different smooth approaches that have been taken to solve problem (1). Many authors have derived algorithms for the solution of problem (1) by solving a sequence of either only unconstrained subproblems or constrained subproblems of the trust-region type, for instance, see Wang and Tang [11] and Erdmann [6].

In this paper, we reformulate problem (1) as an unconstrained smooth optimization problem by constructing a modified Lagrangian function (MLF) of the form

$$G(x,u,v,z,t) = |t|(\sum_{i=1}^{m} u_i e^{(f_i(x)-z)/|t|} + \sum_{j=1}^{l} v_j e^{g_j(x)/|t|}), \qquad (2)$$

where $(u_1,\ldots,u_m)^T \in R_+^m$, $(v_1,\ldots,v_l)^T \in R_+^l$, $z \in R$ is an estimation of the objective function $F(x)$, and $t \neq 0$ is a controlling parameter. The function has the similarity to that by Polyak [10] for unconstrained minimax problems. One of its features is that it inherits the same order of smoothness as the functions in problem (1). Under the standard second order optimality conditions, the MLF is shown to have a unique local minimizer when the corresponding Lagrange multipliers are very close to the optimal Lagrange multiplier of problem (1), z is very close to the optimal objective function and the controlling parameter $|t|$ is small enough.

The MLF algorithm consists of finding the unconstrained minimizer of the MLF (2) and updating the Lagrange multipliers, using the minimizer and the controlling parameter t. In contrast with the feasible direction algorithms, the attraction of the MLF algorithm is that it does not require feasible initial points. The advantage of MLF algorithm over the smoothing technique is that the controlling parameter related to the MLF algorithm does not need to be very small for it employs the Lagrange multipliers in MLF (2). It is demonstrated that under the mild conditions, the MLF algorithm is Q-superlinearly convergent. In the next section, the properties of MLF will be analyzed and the MLF algorithm for solving problem (1) will be described. Section 3 will present convergence theory and error bounds of solution for it. The numerical results for solving some typical minimax optimization problems, will be reported to show the validity of this algorithm in Section 4.

At last, we introduce the following notations.

$$\begin{aligned} R^m_+ &= \{x \in R^m | x_i \geq 0, i = 1, \ldots, m\}, \\ R^m_{++} &= \{x \in R^m | x_i > 0,\ i = 1, \ldots, m\}, \\ u_{(m-r_1)} &= (u_{r_1+1}, u_{r_1+2}, \ldots, u_m)^T \in R^{m-r_1}, \\ v_{(l-r_2)} &= (v_{r_2+1}, v_{r_2+2}, \ldots, v_l)^T \in R^{l-r_2}, \\ \bar{u} &= (u_1, u_2, \ldots, u_{r_1})^T \in R^{r_1}, \\ \bar{v} &= (v_1, v_2, \ldots, v_{r_2})^T \in R^{r_2}, \\ e_n &= (1, 1, \ldots, 1)^T \in R^n. \end{aligned}$$

2. The Properties of MLF and MLF Algorithm

Let (x^*, u^*, v^*) be the Kuhn-Tucker point of problem (1), $z^* = F(x^*)$, $I_1(x^*) = \{i | f_i(x^*) = F(x^*), i = 1, \ldots, m\}$ and $I_2(x^*) = \{j | g_j(x^*) = 0,\ j = 1, \ldots, l\}$. Let $L(x, u, v) = \sum_{i=1}^m u_i f_i(x) + \sum_{j=1}^l v_j g_j(x)$ denote the Lagrange function for problem (1). Set $\phi(x, u, t) = |t| \log \sum_{i=1}^m u_i e^{f_i(x)/|t|}$. The basic conditions are stated as follows, which will be used in the sequel proof.

(a) $f_i(x)$ $(i = 1, \ldots, m)$ and $g_j(x)$ $(j = 1, \ldots, l)$ are twice continuously differentiable.

(b) For convenience of statement, assume $I_1(x^*) = \{1, \ldots, r_1\}$ $(r_1 \leq m)$ and $I_2(x^*) = \{1, \ldots, r_2\}$ $(r_2 \leq l)$.

(c) (x^*, u^*, v^*) satisfies the Kuhn-Tucker conditions:

$$\begin{aligned} &\nabla_x L(x^*, u^*, v^*) = 0_n, g(x^*) \leq 0, \\ &u^* \in R^m_+,\ \sum_{i=1}^m u^*_i = 1,\ u^*_i(F(x^*) - f_i(x^*)) = 0,\ i = 1, \ldots, m, \\ &v^* \in R^l_+,\ v^*_j g_j(x^*) = 0,\ j = 1, \ldots, l. \end{aligned}$$

(d) Strict complementarity condition holds, i.e., $u^*_i > 0$ for $i \in I_1(x^*)$ and $v^*_j > 0$ for $j \in I_2(x^*)$.

(e) Linear independence constraint qualification holds in the sense that $\{\nabla f_i(x^*) | i \in I_1(x^*)\} \cup \{\nabla g_j(x^*) | j \in I_2(x^*)\}$ forms a set of linear independent vectors of R^n.

(f) There exists a constant $\lambda_0 > 0$ such that for all z satisfying $\nabla f_i(x^*)^T z = 0$, $i \in I_1(x^*)$ and $\nabla g_j(x^*)^T z = 0$, $j \in I_2(x^*)$, it holds that

$$z^T \nabla^2_x L(x^*, u^*, v^*) z \geq \lambda_0 \|z\|^2.$$

Now we study the properties of the modified Lagrangian function $G(x,u,v,z,t)$ and then establish a corresponding algorithm for problem (1).

Lemma 1. (see Bertsekas [2]) *Let $A \in I\!R^{n,n}$ be a symmetric matrix, $B \in I\!R^{r,n}$, $U = \mathrm{diag}_{1\le i\le r}(u_i) \in I\!R^{r,r}$, $u_i > 0 \quad (i=1,\dots,r)$ such that $By=0$ implies that $y^TAy \ge \lambda\|y\|^2$, where $\lambda > 0$ is a constant, then there exists $p_0 > 0$ such that for any $\mu \in (0,\lambda)$, we have*

$$x^T(A + pB^TUB)x \ge \mu\|x\|^2, \quad \forall x \in I\!R^n,$$

whenever $p > p_0$.

Proposition 1. *Assume that conditions (a) and (c) hold, then for $t \neq 0$,*

$$z^* = \phi(x^*, u^*, t).$$

Proposition 2. *Assume that conditions (a) and (d) hold, then it holds for $t \neq 0$ that*

$$\nabla_x G(x^*, u^*, v^*, z^*, t) = 0_n.$$

Proof. By calculation, we have

$$\nabla_x G(x,u,v,z,t) = \sum_{i=1}^m u_i e^{(f_i(x)-z)/|t|}\nabla f_i(x) + \sum_{j=1}^l v_j e^{g_j(x)/|t|}\nabla g_j(x).$$

It follows that

$$\begin{aligned}\nabla_x G(x^*, u^*, v^*, z^*, t) &= \sum_{i\in I_1(x^*)} u_i^*\nabla f_i(x^*) + \sum_{j\in I_2(x^*)} v_j^*\nabla g_j(x^*)\\ &= \nabla_x L(x^*, u^*, v^*) = 0_n. \quad \square\end{aligned}$$

Proposition 3. *Assume that conditions (a)-(f) are satisfied, then there exists $\hat{t} > 0$ such that $\nabla_x^2 G(x^*, u^*, v^*, z^*, t)$ is positive definite for any $|t| \in (0, \hat{t})$.*

Proof. Since

$$\begin{aligned}\nabla_x^2 G(x,u,v,z,t) &= \sum_{i=1}^m u_i e^{(f_i(x)-z)/|t|}\nabla^2 f_i(x) + \sum_{j=1}^l v_j e^{g_j(x)/|t|}\nabla^2 g_j(x)\\ &+ \frac{1}{|t|}\left(\sum_{i=1}^m u_i e^{(f_i(x)-z)/|t|}\nabla f_i(x)\nabla f_i(x)^T + \sum_{j=1}^l v_j e^{g_j(x)/|t|}\nabla g_j(x)\nabla g_j(x)^T\right),\end{aligned}$$

we have

$$\nabla_x^2 G(x^*, u^*, v^*, z^*, t) = \nabla_x^2 L(x^*, u^*, v^*) \\ + \frac{1}{|t|} \left(\sum_{i \in I_1(x^*)} u_i^* \nabla f_i(x^*) \nabla f_i(x^*)^T + \sum_{j \in I_2(x^*)} v_j^* \nabla g_j(x^*) \nabla g_j(x^*)^T \right).$$

It follows from condition (f) and Lemma 1 that the conclusion is true. □

Remark 1. From Proposition 3, we have that x^* is the unique local minimizer of $G(x, u, v, z, t)$ when (u, v, z) is very close to (u^*, v^*, z^*) and $|t|$ is small enough.

Hence it is sound to establish the following MLF algorithm.

Algorithm 1 *Step 1.* Choose $\eta \in (0, 1)$, $t^0 > 0$ to be small enough, $u^0 \in I\!R^m_{++}$, $\sum_{i=1}^m u_i^0 = 1$, $v^0 \in I\!R^l_{++}$, $z^0 \in I\!R^1_{++}$ and set $k = 0$.

Step 2. Solve $x^k = \text{argmin}_{x \in I\!R^n} G(x, u^k, v^k, z^k, t^k)$.

Step 3. If $u_i^k(f_i(x^k) - F(x^k)) = 0$ for $i = 1, ..., m$ and $v_j^k g_j(x^k) = 0$ for $j = 1, ..., l$, then stop and x^k is a solution to problem (1).

Step 4. Update u^k, v^k, z^k and t^k by

$$\begin{aligned} u_i^{k+1} &= u_i^k e^{f_i(x^k)/t^k} / \sum_{j=1}^m u_j^k e^{f_j(x^k)/t^k}, \; i = 1, ..., m, \\ v_j^{k+1} &= v_j^k e^{g_j(x^k)/t^k}, \; j = 1, ..., l, \\ z^{k+1} &= \phi(x^k, u^k, t^k), \\ t^{k+1} &= \eta t^k. \end{aligned}$$

Step 5. Let $k = k + 1$ and go to Step 2.

3. Convergence Results

This section will present the convergence theory for MLF algorithm.

Lemma 2. *Assume that conditions (a)-(d) are satisfied. Then there exists $\epsilon > 0$, such that for $i = r_1 + 1, \ldots, m$, the function*

$$\hat{u}_i(x, u, v, z, t) = \begin{cases} u_i e^{\frac{1}{|t|}(f_i(x) - z)}, & t \neq 0, \\ 0, & t = 0, \end{cases} \tag{3a}$$

and for $j = r_2 + 1, \ldots, l$, the function

$$\hat{v}_j(x, u, v, z, t) = \begin{cases} v_j e^{\frac{1}{|t|} g_j(x)}, & t \neq 0, \\ 0, & t = 0, \end{cases} \tag{3b}$$

are continuously differentiable on $S(x^*,\epsilon)\times S_+(u^*,\epsilon)\times S_+(v^*,\epsilon)\times(z^*-\epsilon,z^*+\epsilon)\times(-\epsilon,\epsilon)$.

Proof. At first, we establish the continuity of $\hat{u}_i(x,u,t)$ on $S(x^*,\epsilon)\times S_+(u^*,\epsilon)\times S_+(v^*,\epsilon)\times\ (z^*-\epsilon,z^*+\epsilon)\times\ (-\epsilon,\epsilon)$. Obviously, $\hat{u}_i(x,u,t)$ is continuous on $I\!R^n\times I\!R^m\times I\!R^l\times I\!R^1\times(I\!R^1\backslash\{0\})$, so we only need to prove the continuity of $\hat{u}_i(x,u,t)$ at $(x,u,v,z,0)$ when it is close to $(x^*,u^*,v^*,z^*,0)$.

Choose $\sigma>0$ satisfying

$$f_i(x^*)-F(x^*)\le-\sigma,\quad i=r_1+1,\dots,m.$$

By the continuity of $f_i(x)$ $(i=1,\dots,m)$ for $\delta\in(0,(2(m+l+1))^{-1}\sigma)$, there exists $\epsilon>0$ such that when $x\in S(x^*,\epsilon)$, it holds

$$|f_i(x)-f_i(x^*)|\le\delta,\quad i=1,\dots,m.$$

Choose ϵ being sufficiently small such that $u\in S_+(u^*,\epsilon)$ implies $|u_i-u_i^*|\le\delta$ $i=1,\dots,m$. Moreover, $z\in(z^*-\epsilon,z^*+\epsilon)$ implies $|F(x^*)-z|<\delta/2$. Then for $(x,u,v,z)\in S(x^*,\epsilon)\times S_+(u^*,\epsilon)\times S_+(v^*,\epsilon)\times(z^*-\epsilon,z^*+\epsilon)$, it holds

$$u_i\le\delta,\quad f_i(x)-z\le f_i(x^*)+\delta-z\le f_i(x^*)-z^*+3\delta/2<3\delta/2-\sigma.$$

for $i=r_1+1,\dots,m$. Hence for $t\ne0$ and for $i=r_1+1,\dots,m$, we have

$$0\ \le\ \hat{u}_i(x,u,\alpha)\ =\ u_ie^{|t|^{-1}(f_i(x)-z)}\ \le\ \delta e^{|t|^{-1}(3\delta/2-\sigma)}\ =\ \delta e^{-|t|^{-1}(\sigma-3\delta/2)}.$$

Since $\sigma>2(m+l+1)\delta$ and $\sigma-3\delta/2>0$, there exist two positive constants β_1 and β_2 dependent on σ and δ such that

$$0\le\hat{u}_i(x,u,t)\le\beta_2e^{-\beta_1|t|^{-1}},\quad i=r_1+1,\dots,m. \tag{4}$$

Thus for $i=r_1+1,\dots,m$, the following holds

$$\lim_{t\to0}\hat{u}_i(x,u,v,z,t)=\hat{u}_i(x,u,v,z,0)=0\ \text{ for }\ (x,u)\in S(x^*,\epsilon)\times S_+(u^*,\epsilon),$$

which means that $\hat{u}_i(x,u,v,z,t)$ is continuous at $(x,u,v,z,0)$.

Now we prove the continuous differentiability of $\hat{u}_i(x,u,v,z,t)$ for $i=r_1+1,\dots,m$. For $t\ne0$ and $i=r_1+1,\dots,m$, we obtain

$$\begin{aligned}\nabla_x\hat{u}_i(x,u,v,z,t)&=|t|^{-1}\hat{u}_i(x,u,v,z,t)\nabla f_i(x),\\ \frac{\partial\hat{u}_i(x,u,v,z,t)}{\partial u_i}&=e^{|t|^{-1}(f_i(x)-z)},\end{aligned}$$

$$\begin{aligned}
\frac{\partial \hat{u}_i(x,u,v,z,t)}{\partial u_j} &= 0, \quad i \neq j, \\
\frac{\partial \hat{u}_i(x,u,v,z,t)}{\partial t} &= -(t|t|)^{-1}\hat{u}_i(x,u,v,z,t) f_i(x), \\
\frac{\partial \hat{u}_i(x,u,v,z,t)}{\partial z} &= -|t|^{-1}\hat{u}_i(x,u,v,z,t).
\end{aligned}$$

Noting that $\hat{u}_i(x,u,0) = 0$, we get $\nabla_x \hat{u}_i(x,u,v,z,0) = 0_n$, $\nabla_u \hat{u}_i(x,u,v,z,0) = 0_m$ and $\frac{\partial \hat{u}_i(x,u,v,z,0)}{\partial z} = 0$. Obviously, $\hat{u}_i(x,u,t)$ is continuously differentiable on $I\!R^n \times I\!R^m \times I\!R^l \times I\!R^1 \times (I\!R^1 \backslash \{0\})$. Therefore, we only need to prove its differentiability at (x,u,v,z,t) when it is close to (x^*,u^*,v^*,z^*). From the above analysis, there exists $\epsilon > 0$, such that (4) are valid. Hence it holds

$$0 \leq \lim_{t \downarrow 0} \frac{\hat{u}_i(x,u,v,z,t) - \hat{u}_i(x,u,v,z,0)}{t} \leq \lim_{t \downarrow 0} \frac{\beta_2 \exp(-\beta_1 |t|^{-1})}{t} = 0.$$

Similarly,

$$\lim_{t \uparrow 0} \frac{\hat{u}_i(x,u,v,z,t) - \hat{u}_i(x,u,v,z,0)}{t} = 0.$$

Thus $\hat{u}_i(x,u,v,z,t)$ is differentiable at $t = 0$ with respect to t and $\partial \hat{u}_i(x,u,v,z,0)/\partial t = 0$.

Since $\|f_i(x)\|$ is bounded over $S(x^*,\epsilon)$ for $i = 1,\ldots,m$, there exist two positive constants β_3 and β_4 satisfying

$$\|\nabla_{x,u,v,z,t}\hat{u}_i(x,u,v,z,t)\| \leq \beta_4 |t|^{-2} \exp(-\beta_3 |t|^{-1}),$$

for $(x,u,v,z,t) \in S(x^*,\epsilon) \times S_+(u^*,\epsilon) \times S_+(v^*,\epsilon) \times (z^* - \epsilon, z^* + \epsilon) \times ((-\epsilon,\epsilon)\backslash\{0\})$, where $\nabla_{x,u,v,z,t}\hat{u}_i(.) = (\nabla_x \hat{u}_i(.), \nabla_u \hat{u}_i(.), \nabla_v \hat{u}_i(.), \partial \hat{u}_i(.)/\partial z, \partial \hat{u}_i(.)/\partial t)$.

Thus, we have

$$\begin{aligned}
\lim_{(x',u',v',z',t) \to (x,u,v,z,0)} \nabla_{x,u,v,z,t}\hat{u}(x',u',v',z',t) &= 0_{n+m+l+1+1} \\
&= \nabla_{x,u,v,z,t}\hat{u}(x,u,v,z,0),
\end{aligned}$$

for $(x,u,v,z) \in S(x^*,\epsilon) \times S_+(u^*,\epsilon) \times S_+(v^*,\epsilon) \times (z^* - \epsilon, z^* + \epsilon)$. That is, $\hat{u}_i(x,u,v,z,t)$ is continuously differentiable on $(x,u,v,z,t) \in S(x^*,\epsilon) \times S_+(u^*,\epsilon) \times S_+(v^*,\epsilon) \times (z^* - \epsilon, z^* + \epsilon) \times (-\epsilon,\epsilon)$.

The continuous differentiability of $\hat{v}_j(x,u,v,z,t)$ for $j = r_2+1,\ldots,l$ may be proved in the same way as that of $\hat{u}_i(x,u,v,z,t)$ for $i = r_1+1,\ldots,m$. □

Theorem 1. *Assume that conditions (a)-(f) are satisfied. Then there exist* $\epsilon^u > 0$, $\epsilon^v > 0$, $\epsilon^z > 0$, $\hat{t} \in (0,1)$ *and* $\epsilon^x > 0$ *such that for any* $(u,v,z,|t|) \in$

$S_+(u^*, \epsilon^u) \times S_+(v^*, \epsilon^y) \times (z^* - \epsilon, z^* + \epsilon) \times (0, \hat{t})$, *the following statements are true.*

(i) There exists $\tilde{x} = x(u, v, z, t) \in S(x^*, \epsilon^x)$ *such that*

$$\tilde{x} = \text{argmin}\{G(x, u, v, z, t) | x \in S(x^*, \epsilon^x)\};$$

(ii) For $\tilde{x}$ *in (i), let*

$$\tilde{\eta} = (\eta_1(\tilde{x}, u, v, z, t), \ldots, \eta_m(\tilde{x}, u, v, z, t))^T,$$

$$\tilde{\xi} = (\xi_1(\tilde{x}, u, v, z, t), \ldots, \xi_l(\tilde{x}, u, v, z, t))^T,$$

$$\tilde{\lambda} = (\lambda_1(\tilde{x}, u, v, z, t), \ldots, \lambda_m(\tilde{x}, u, v, z, t))^T$$

and $\tilde{z} = \phi(\tilde{x}, u, t)$, *where* $\eta_i(\tilde{x}, u, v, z, t) = u_i e^{\frac{1}{|t|} f_i(\tilde{x})}$ *for* $i = 1, \ldots, m$, $\xi_j(\tilde{x}, u, v, z, t) = v_j e^{\frac{1}{|t|} g_j(\tilde{x})}$ *for* $j = 1, \ldots, l$ *and*

$$\tilde{\lambda}_i(\tilde{x}, u, v, z, t) = \frac{\eta_i(\tilde{x}, u, v, z, t)}{\sum_{i=1}^m \eta_i(\tilde{x}, u, v, z, t)}$$

for $i = 1, \ldots, m$. *Then it holds*

$$\max\{\|\tilde{x} - x^*\|, \|\tilde{\eta} - u^*\|, \|\tilde{\xi} - v^*\|\} \leq \rho(u, v, z, t, u^*, v^*, z^*, c), \tag{5a}$$

and, for $i = 1, \ldots, r_1$,

$$\frac{-\rho(u, v, z, t, u^*, v^*, z^*, c) + u_i^*}{(1 + m\rho(u, v, z, t, u^*, v^*, z^*, c))} \leq \tilde{\lambda}_i \leq \frac{\rho(u, v, z, t, u^*, v^*, z^*, c) + u_i^*}{(1 - m\rho(u, v, z, t, u^*, v^*, z^*, c))}, \tag{5b}$$

for $j = r_1 + 1, \ldots, m$,

$$0 \leq \tilde{\lambda}_j \leq \frac{\rho(u, v, z, t, u^*, v^*, z^*, c)}{(1 - m\rho(u, v, z, t, u^*, v^*, z^*, c))}, \tag{5c}$$

where $\rho(u, v, z, t, u^*, v^*, z^*, c) = c(|t|(\|u - u^*\| + \|v - v^*\|) + |z - z^*\|)$, *and* $c > 0$ *is independent of parameter* t.

Proof. At first, we prove (i). Let

$$d(x, u, v, z, t) = \sum_{i=r_1+1}^m \hat{u}_i(x, u, v, z, t)\nabla f_i(x) + \sum_{j=r_2+1}^l \hat{v}_j(x, u, v, z, t)\nabla g_j(x).$$

Then by Lemma 2, one has $d(x, u, v, z, 0) = 0_n$, $\nabla_x d(x, u, v, z, 0) = 0_{n,n}$.

Define mapping

$$\Phi : I\!R^n \times I\!R^{r_1} \times I\!R^{r_2} \times I\!R^m \times I\!R^l \times I\!R^1 \times I\!R^1 \longmapsto I\!R^{n+r_1+r_2},$$

$$\Phi(x, \bar{u}, \bar{v}, u, v, z, t) = \begin{pmatrix} \sum_{i=1}^{r_1} \bar{u}_i \nabla f_i(x) + \sum_{j=1}^{r_2} \bar{v}_j \nabla g_j(x) + d(x, u, v, z, t) \\ f_1(x) - z - |t| \ln(\bar{u}_1 u_1^{-1}) \\ \vdots \\ f_{r_1}(x) - z - |t| \ln(\bar{u}_{r_1} u_{r_1}^{-1}) \\ g_1(x) - |t| \ln(\bar{v}_1 v_1^{-1}) \\ \vdots \\ g_{r_2} - |t| \ln(\bar{v}_{r_2} v_{r_2}^{-1}) \end{pmatrix}.$$

Thus, in view of the conditions (a)-(f), we have that $\Phi(x^*, u^*_{r_1}, v^*_{r_2}, u^*, v^*, z^*, 0) = 0$ and the Jacobi matrix

$$\nabla_{x,\bar{u},\bar{v}} \Phi(x^*, \bar{u}^*, \bar{v}^*, u^*, v^*, z^*, 0) = \begin{pmatrix} \nabla_x^2 L(x^*, u^*, v^*) & \nabla f_{(r_1)}(x^*) & \nabla g_{(r_2)}(x^*) \\ \nabla f_{(r_1)}(x^*)^T & 0_{r_1,r_1} & 0_{r_1,r_2} \\ \nabla g_{(r_2)}(x^*)^T & 0_{r_2,r_1} & 0_{r_2,r_2} \end{pmatrix}$$

is nonsingular, where $\nabla f_{(r_1)}(x^*) = (\nabla f_1(x^*), \ldots, \nabla f_{r_1}(x^*))$ and $\nabla g_{(r_2)}(x^*) = (\nabla g_1(x^*), \ldots, \nabla g_{r_2}(x^*))$. By the implicit theorem and condition (a), there exist $\delta^x > 0$, $\delta^{\bar{u}} > 0$,$\delta^{\bar{v}} > 0$, $\delta^u > 0$, $\delta^v > 0$, $\delta^z > 0$, and $t_1 > 0$ and unique smooth vector functions $x(u, v, z, t)$, $\bar{u}(u, v, z, t)$ and $\bar{v}(u, v, z, t)$ satisfying:

$$\begin{pmatrix} x(u, v, z, t) \\ \bar{u}(u, v, z, t) \\ \bar{v}(u, v, z, t) \end{pmatrix} : \quad D(u, v, z, t) \longmapsto S(x^*, \delta^x) \times S(\bar{u}^*, \delta^{\bar{u}}) \times S(\bar{v}^*, \delta^{\bar{v}}),$$

$$x(u^*, v^*, z^*, t) = x^*, \bar{u}(u^*, v^*, z^*, t) = \bar{u}^*, \bar{v}(u^*, v^*, z^*, t) = \bar{v}^*,$$

$$|t| \in (0, t_1),$$

for $(u, v, z, t) \in D(u, v, z, t)$, we have

$$\Phi(x(u, v, z, t), \bar{u}(u, v, z, t), \bar{v}(u, v, z, t), u, v, z, t) \equiv 0_{n+r_1+r_2}, \tag{6}$$

and there exists a constant ρ_1 such that

$$\|(\nabla_{x,\bar{u},\bar{v}} \Phi(x(u, v, z, t), \bar{u}(u, v, z, t), \bar{v}(u, v, z, t), u, v, z, t))^{-1}\| \leq \rho_1, \tag{7}$$

where $D(u,v,z,t) = \{(u,v,z,t)|(u,v,z,t) \in S(u^*,\delta^u) \times S(v^*,\delta^v) \times S(z^*,\delta^z) \times (0,t_1)\}$.

By (6), one gets

$$\textstyle\sum_{i=1}^{r_1} \bar{u}_i(u,v,z,t)\nabla f_i(x) + \sum_{j=1}^{r_2} \bar{v}_j(u,v,z,t)\nabla g_j(x) + d(x,u,v,z,t) = 0, \tag{8a}$$

and for $t \neq 0$

$$\bar{u}_i(u,v,z,t) = u_i e^{\frac{1}{|t|} f_i(x(u,v,z,t)-z)}, i = 1,\ldots,r_1, \tag{8b}$$

$$\bar{v}_j(u,v,z,t) = v_j e^{\frac{1}{|t|} g_j(x(u,v,z,t))}, \quad j = 1,\ldots,r_2. \tag{8c}$$

From (8a)-(8c), for $(u,v,z,t) \in S(u^*,\delta^u) \times S(v^*,\delta^v) \times S(z^*,\delta^z) \times (0,t_1)$, we have that

$$\nabla_x G(x(.),.) = \sum_{i=1}^{r_1} \bar{u}_i(.)\nabla f_i(x) + \sum_{j=1}^{r_2} \bar{v}_j(.)\nabla g_j(x) + d(x,u,v,z,t) = 0_n,$$

where (.) denotes (u,v,z,t).

Choose $\epsilon^x \in (0,\delta^x)$, $\epsilon^{\bar{u}} \in (0,\delta^{\bar{u}})$, $\epsilon^{\bar{v}} \in (0,\delta^{\bar{v}})$, $\epsilon^u \in (0,\delta^u)$, $\epsilon^v \in (0,\delta^v)$, $\epsilon^z \in (0,\delta^z)$, $\hat{t} \in (0,t_1)$ such that $\nabla_x^2 G(x,u,v,z,t)$ is positive definite for $(x,u,v,z,|t|) \in S(x^*,\epsilon^x) \times S_+(u^*,\epsilon^u) \times S_+(v^*,\epsilon^v) \times S(z^*,\epsilon^z) \times (0,\hat{t})$, and such that $x(S_+(u^*,\epsilon^u), S_+(v^*,\epsilon^v), S(z$ $\mathrm{int}S(x^*,\epsilon^x)$, $\bar{u}(S_+(u^*,\epsilon^u)$, $S_+(v^*,\epsilon^v)$, $S(z^*,\epsilon^v z), T_{sign}(\hat{t})) \subseteq \mathrm{int}S(\bar{u}^*,\epsilon^{\bar{u}})$, and $\bar{v}(S_+(u^*,\epsilon^u), S_+(v^*,\epsilon^v)$, $S(z^*,\epsilon^v)$, $T_{sign}(\hat{t})) \subseteq \mathrm{int}S(\bar{v}^*,\epsilon^{\bar{v}})$, where $T_{\mathrm{sign}}(\hat{t})$ is defined by $T_+(\hat{t}) = (0,\hat{t})$ and $T_-(\hat{t}) = (-\hat{t},0)$. Then we obtain $\nabla_x G(\hat{x},u,v,z,t) = 0$ by setting $\hat{x} = x(u,v,z,t)$ for $(u,v,z,t) \in S_+(u^*,\epsilon^u) \times S_+(v^*,\epsilon^v) \times S(z^*,\epsilon^z) \times T_{sign}(\hat{t})$. Hence by the choices of ϵ^x,$\epsilon^{\bar{u}}$, $\epsilon^{\bar{v}}$, ϵ^u, ϵ^v, ϵ^z, $\hat{t}$, we know that $\hat{x}$ is the unique minimizer point of $G(x,u,v,z,t)$ on $S(x^*,\epsilon^x)$, i.e., for $(u,v,z,|t|) \in S(u^*,\epsilon^u) \times S(v^*,\epsilon^v) \times S(z^*,\epsilon^z) \times (0,\hat{t})$,

$$\hat{x} = \mathrm{argmin}_{x \in S(x^*,\epsilon^x)} G(x,u,v,z,t),$$

which means that the statement (i) is true.

Now we turn to the proof of (ii). For $(u,v,z,|t|) \in S(u^*,\epsilon^u) \times S(v^*,\epsilon^v) \times S(z^*,\epsilon^z) \times (0,\hat{t})$, it follows from (6) that

$$\Phi(x(u,v,z,t), \bar{u}(u,v,z,t), \bar{v}(u,v,z,t), u,v,z,t) \equiv 0_{n+r_1+r_2}. \tag{9}$$

Differentiating the formula (9) with respect to (u,v,z), one gets

$$\nabla_{x,\bar{u},\bar{v}}\Phi(x(.),\bar{u}(.),\bar{v}(.),.)w(.) + \nabla_{u,v,z}\Phi(x(.),\bar{u}(.),\bar{v}(.),.)$$

$$\equiv 0_{(n+r_1+r_2)\times(m+l+1)},$$

where $w(.) = (\nabla_{u,v,z}x(.)^T, \nabla_{u,v,z}\bar{u}(.)^T, \nabla_{u,v,z}\bar{v}(.))^T$.

By calculation, we have

$$\nabla_{u,v,z}\Phi(x(u,v,z,t),\bar{u}(u,v,z,t),\bar{v}(u,v,z,t),u,v,z,t)$$
$$= \begin{pmatrix} \sum_{i=r_1+1}^m \nabla f_i(\tilde{x})\nabla_u \hat{u}_i(.)^T & \sum_{j=r_2+1}^l \nabla g_j(\tilde{x})\nabla_v \hat{v}_j(.)^T & \sum_{i=r_1+1}^m \frac{\partial \hat{u}_i(.)}{\partial z}\nabla f_i(\tilde{x}) \\ \frac{|t|}{u_1}\bar{e}_1^{(m)T} & 0_l^T & -1 \\ \vdots & \vdots & \vdots \\ \frac{|t|}{u_{r_1}}\bar{e}_{r_1}^{(m)T} & 0_l^T & -1 \\ 0_m^T & \frac{|t|}{v_1}\bar{e}_1^{(l)T} & 0 \\ \vdots & \vdots & \vdots \\ 0_m^T & \frac{|t|}{v_{r_1}}\bar{e}_{r_2}^{(l)T} & 0 \end{pmatrix}, \tag{10}$$

where $\bar{e}_i^{(m)}$ is the i-th unit vector of $I\!R^m$, $\bar{e}_j^{(l)}$ is the j-th unit vector of $I\!R^l$ and $(.) = (x(u,v,z,t),u,v,z,t)$. Noting for $i = r_1+1,\dots,m$ that $\nabla_u \hat{u}_i(x,u,v,z,0) = 0$ and for $j = r_2+1,\dots,l$ that $\nabla_v \hat{v}_j(x,u,v,z,0) = 0$, for $(u,v,z,|t|) \in S_+(u^*,\epsilon^u)\times S_+(v^*,\epsilon^v)\times S(z^*,\epsilon^z)\times(0,\hat{t})$, it follows that there exists $\rho_2 > 0$ satisfying

$$\|\nabla_u \hat{u}_i(x(u,v,z,,t),u,v,z,t)\| \le \rho_2|t|, \tag{11a}$$

$$\|\nabla_v \hat{v}_j(x(u,v,z,,t),u,v,z,t)\| \le \rho_2|t|. \tag{11b}$$

Further we note that when $u \in S(u^*,\epsilon^u)$ and $\epsilon^u \in (0, \min_{1\le i\le r_1}\{u_i^*\})$, it holds

$$u_i^{-1} \le \frac{1}{\min_{1\le j\le r_1}\{u_i^*\} - \epsilon^u},\ i = 1,\dots,r_1. \tag{12a}$$

Similarly, when $v \in S(v^*,\epsilon^v)$ and $\epsilon^v \in (0, \min_{1\le j\le r_2}\{v_j^*\})$, it holds

$$v_j^{-1} \le \frac{1}{\min_{1\le j\le r_2}\{v_j^*\} - \epsilon^v},\ j = 1,\dots,r_2. \tag{12b}$$

Taking into account that (10)-(12) and the boundedness of $\|\nabla f_i(x)\|$ and $\|\nabla g_j(x)\|$ in $S(x^*,\epsilon^x)$, we have that there exists $\rho_3 > 0$ in dependent of t such that for $(u,v,z,|t|) \in S_+(u^*,\epsilon^u)\times S_+(v^*,\epsilon^v)\times S(z^*,\epsilon^z)\times(0,\hat{t})$, it holds

$$\|\nabla_{u,v,z}\Phi(x(u,v,z,t),\bar{u}(u,v,z,t),\bar{v}(u,v,z,t),u,v,z,t)\|$$
$$\le \rho_3\|(|t|E_{n\times(m+l)}, e_n)\|,$$

where $E_{n\times(m+l)} = (e_n, \ldots, e_n) \in R^{n\times(m+l)}$. Let $\rho_4 = \rho_1\rho_3$, then we get

$$\begin{aligned}
&\|w(u,v,z,t)\| \\
&\qquad = \| - [\nabla_{x,\bar{u},\bar{v}}\Phi(x(u,v,z,t), \bar{u}(u,v,z,t), \bar{v}(u,v,z,t), u, v, z, t)]^{-1} \\
&\qquad\quad \times \nabla_{u,v,z}\Phi(x(u,v,z,t), \bar{u}(u,v,z,t), \bar{v}(u,v,z,t), u, v, z, t)\| \\
&\qquad\qquad\qquad\qquad \le \rho_4\|(|t|E_{n\times(m+l)}, e_n)\|.
\end{aligned}$$

For any $(u, v, z, |t|) \in S_+(u^*, \epsilon^u) \times S_+(v^*, \epsilon^v) \times S(z^*, \epsilon^z) \times (0, \hat{t})$, we have

$$\begin{aligned}
&\left\| \begin{bmatrix} x(u,v,z,t) - x^* \\ \bar{u}(u,v,z,t) - \bar{u}^* \\ \bar{v}(u,v,z,t) - \bar{v}^* \end{bmatrix} \right\| = \left\| \begin{bmatrix} x(u,v,z,t) - x(u^*,v^*,z^*,t) \\ \bar{u}(u,v,z,t) - \bar{u}(u^*,v^*,z^*,t) \\ \bar{v}(u,v,z,t) - \bar{v}(u^*,v^*,z^*,t) \end{bmatrix} \right\| \\
&= \left\| \int_0^1 w(u^* + s(u-u^*), v^* + s(v-v^*), z^* + s(z-z^*)t) \begin{bmatrix} u-u^* \\ v-v^* \\ z-z^* \end{bmatrix} \mathrm{d}s \right\| \\
&\qquad\qquad \le \rho_4(|t|(\|u-u^*\| + \|v-v^*\|) + |z-z^*|).
\end{aligned}$$

By (11) for $(u, v, z, |t|) \in S(u^*, \epsilon^u) \times S(v^*, \epsilon^v) \times S(z^*, \epsilon^z) \times (0, \hat{t})$, we obtain

$$\|\hat{u}_{(m-r_1)}(x(.),.) - u^*_{(m-r_1)}\| \quad \le \quad \rho_2(|t|(\|u-u^*\| + \|v-v^*\|) + |z-z^*|),$$

and

$$\|\hat{v}_{(l-r_2)}(x(.),.) - v^*_{(l-r_2)}\| \quad \le \quad \rho_2(|t|(\|u-u^*\| + \|v-v^*\|) + |z-z^*|).$$

Choose $c = 2\max\{\rho_2, \rho_4\}$, then for $(u, v, z, |t|) \in S_+(u^*, \epsilon^u) \times S_+(v^*, \epsilon^v) \times S(z^*, \epsilon^z) \times (0, \hat{t})$, we obtain

$$\max\{\|\tilde{x} - x^*\|, \|\tilde{\eta} - u^*\|, \|\tilde{\xi} - v^*\|\} \le \rho(u, v, z, t, u^*, v^*, z^*, c).$$

Assume that $\rho(.) \le \min_{1\le i\le r_1}\{u_i^*\}$, where $\rho(.)$ denotes $\rho(u, v, z, t, u^*, v^*, z^*, c)$. From (5a) we have that

$$0 \le -\rho(.) + u_i^* \le \tilde{\eta}_i \le \rho(.) + u_i^*, \quad i = 1, \ldots, r_1,$$

$$0 \le \tilde{\eta}_i \le \rho(.), \quad i = r_1 + 1, \ldots, m,$$

and

$$(1 + m\rho(.))^{-1} \le (\sum_{i=1}^m \tilde{\eta}_i)^{-1} \le (1 - m\rho(.))^{-1}, \quad i = 1, \ldots, r_1.$$

Thus, (5b) and (5c) hold. That is, (ii) is true. □

problem	iter	$\frac{1}{\lvert t\rvert}$	$\lVert\hat{x}-x^*\rVert$
Charalambous-Conn 1	4	16.105100	0.000002
Beal	16	9.189946	0.000000
Rosen-Suzuki	22	1.628055	0.000000
Wong 1	16	0.275698	0.000007
Wong 2	35	0.843073	0.000001
Wong 3	51	10.330395	0.000001

Table 1: Numerical results

4. Numerical Results

We performed numerical experimentation using the Algorithm 1 for many minimax problems with inequality constraints. Among them, the numerical results of seven problems (see Charalambous [3], Charalambous [4] and Dipillo [5]) are listed in the following tables. In the table, the first column lists the name of each problem, the second columns represents the number of outer iterations in which the multipliers u and v are updated, the third column gives the terminal value of $1/|t|$, the last column shows the gap between the approximate optimal solution via the Algorithm 1 and the optimal solution of each problem.

Remark 2. The numerical results show that, without the controlling parameter $1/|t|$ being very large, the approximate optimal solution to each problem is obtained by using the Algorithm 1.

Remark 3. The precision for each problem is 10^{-6}.

References

[1] A. Ben-Tal, A. Nemirovsky, *Convex Optimization in Engineering: Modeling Analysis, Algorithms, Technion*, Israel (1998).

[2] D.P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York (1982).

[3] C. Charalambous, Nonlinear least p-th optimization and nonlinear programming, *Mathematical Programming*, **12** (1977), 195-225.

[4] C. Charalambous, Acceleration of the least pth algorithm for minimax optimization with engineering applications, *Mathematical Programming*, **19** (1979), 270-297.

[5] G. Dipillo, L. Grippo, S. Lucidi, A smooth method for the finite minimax problem, *Mathematical Programming*, **60** (1993), 187-213.

[6] G. D. Erdmann, *A New Minimax Algorithm and its Application to Optics Problems*, Ph.D. Thesis, University of Minessota, USA (2003).

[7] S.X. He, A dual algorithm for minimax problems, *Journal of Applied Mathematics and Computing*, **17** (2005), 401-418.

[8] X.S. Li, An entropy-based aggregate method for minimax optimization, *Engineering Optimization*, **18** (1992), 277-285.

[9] E. Polak, On the mathematical foundations of nondifferentiable optimization in engineering design, *SIAM Review*, **29** (1987), 21-89.

[10] R. A. Polyak, Smooth optimization methods for minimax problems, *SIAM Journal on Control and Optimization*, **26** (1988), 1274-1286.

[11] Y.C. Wang, H.W. Tang, An entropy function method for min-max problems with constraints, *Numerical Mathematics A Journal of Chinese University*, **21** (1999), 132-139.

[12] L. W. Zhang, H. W. Tang, A maximum entropy algorithm with parameters for solving minimax problem, *Archives of Control Sciences*, **6 (XLII)** (1997), 47-59.