Entropic Regularization Approach for Mathematical Programs with Equilibrium Constraints

Ş. İlker Birbil * Shu-Cherng Fang † Jiye Han ‡

February 18, 2003

Erasmus Universiteit Rotterdam

Abstract

A new smoothing approach based on entropic perturbation is proposed for solving mathematical programs with equilibrium constraints. Some of the desirable properties of the smoothing function are shown. The viability of the proposed approach is supported by a computational study on a set of well-known test problems.

Keywords Mathematical programs with equilibrium constraints, smoothing approach, entropic regularization.

*Erasmus Research Institute of Management (ERIM), Erasmus University, Postbus 1738, 3000 DR Rotterdam, The Netherlands. E-mail: sibirbil@few.eur.nl

†Industrial Engineering and Operations Research, North Carolina State University, Raleigh, NC 26695-7906, USA. E-mail: fang@eos.ncsu.edu

‡Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, P.R. China. E-mail: hjyyyf@ihw.com.cn

1 Introduction

A mathematical program with equilibrium constraints (MPEC) is an optimization problem including essentially a set of parametric variational inequality or complementarity type constraints [Harker and Pang, 1988]. The "inner problem" corresponding to the variational inequality or complementarity type constraints, suggested the word "equilibrium", since in many engineering and economic applications, these constraints refer to a certain equilibrium phenomena. MPEC problems, which are known to be very difficult, arise in numerous areas such as network design, regional science, transportation planning, and game theory. We refer the interested reader to the extensive monograph by Luo et al. [1997].

There have been successful approaches to deal with this problem. One approach is based on general penalization techniques. This approach has been widely used and many good results have been reported [Ishizuka and Aiyoshi, 1992a,b]. Another approach is the application of heuristics. These heuristics are fairly general and do not require strict assumptions in handling MPEC problems [Suwansirikul et al., 1987, Friesz et al., 1992].

Outrata and Zowe [1995] reformulated the optimization problems with monotone variational inequalities. Their reformulation led to a nonsmooth Lipschitz optimization problem and hence benefited from the well-developed theory of nonsmooth optimization [Clarke, 1983]. In their paper, Outrata and Zowe successfully solved several academic problems and applied their approach to compute the well-known Stackelberg-Cournot-Nash equilibria [Murphy et al., 1982].

Recently, Kanzow and Jiang [1998] considered a continuation method for solving the monotone variational inequality problems. Their idea was based on solving a sequence of perturbed problems. The main tool for the perturbed problems was a specific smoothing function, which was also used by the authors in one of their earlier work [Kanzow, 1996]. They reported promising results on variational inequality and convex optimization problems.

Along this line, Facchinei et al. [1999] made use of a smoothing function from Kanzow and Jiang [1998] to solve MPEC problems. In addition to proposing an efficient algorithm, they showed its global convergence. Furthermore, they composed a set of test problems from the literature. Their numerical results on this set of problems demonstrated remarkable performance.

The primary aim of this paper is to exploit the application of a regularization approach to solve the MPEC problems. In particular, we are interested in using the entropy functions. The paper starts with the reformulation of the MPEC problem with nonsmooth constraints. To deal with the nonsmooth constraints, a regularization approach using the entropy function is applied. We validate the proposed approach by conducting a numerical study on a set of test problems collected from the literature. Our test results have shown that the proposed approach is able to converge to the best known results. In some cases, even finds better solutions than those reported in the literature.

The paper is organized as follows. Section 2 introduces an MPEC problem and a nonsmooth reformulation. Section 3 describes the proposed regularization approach with a discussion of its properties. The computational results over a set of test problems are reported in Section 4. The paper is concluded in Section 5.

2 Preliminaries

We adopt the notation $\Re^n$ to denote the n-dimensional real vector space and $\|\cdot\|$ to denote the Euclidean norm in this space. In the sequel, the vectors in a real vector space are viewed as column vectors and the vector $(x^T, y^T, z^T, \lambda^T)^T \in \Re^n \times \Re^m \times \Re^l \times \Re^l$ is usually abbreviated by (x, y, z, λ).

Let us first define a variational inequality problem with variables $x \in \Re^n$ and $y \in \Re^m$. Given a continuously differentiable function $F : \Re^{n+m} \to \Re^m$ and a set valued mapping

$$C(x) \triangleq \{y \in \Re^m : g_i(x, y) \geq 0, i = 1, 2, \cdots, l\} \tag{1}$$

with $g : \Re^{n+m} \to \Re^l$ being twice continuously differentiable and concave in the second variable, a variational inequality problem, VI(F, C) seeks for the solution set $S(x)$ such that $y \in S(x)$ if and only if $y \in C(x)$ and

$$(v - y)^T F(x, y) \geq 0, \text{ for all } v \in C(x). \tag{2}$$

Then, a typical MPEC becomes

$$\begin{array}{ll} \min & f(x, y) \\ \text{s.t.} & x \in X, \\ & y \in S(x), \end{array} \tag{MPEC}$$

where $f : \Re^{n+m} \to \Re$ is a continuously differentiable function and X is a nonempty set in $\Re^n$ [Luo et al., 1997].

Before proceeding to the subsequent analysis, let us first recall several definitions and results given by Harker and Pang [1990].

Definition 1 Given $x \in X$, let F_x denote the vector-valued function $F(x, \cdot) : C(x) \to \Re^n$, then F_x is

(a) monotone if

$$(y_1 - y_2)^T (F_x(y_1) - F_x(y_2)) \geq 0 \text{ for all } y_1, y_2 \in C(x), \tag{3}$$

(b) strongly monotone with modulus $\alpha > 0$ if

$$(y_1 - y_2)^T (F_x(y_1) - F_x(y_2)) \geq \alpha \|y_1 - y_2\|^2 \text{ for all } y_1, y_2 \in C(x). \tag{4}$$

Theorem 1 [Harker and Pang, 1990, Corollary 3.2] Let $C(x)$ be a nonempty, closed convex set and $F(x, \cdot)$ be a strongly monotone and continuous mapping. Then the solution set of the variational inequalities VI(F,C) consists of a unique point.
Furthermore, define the set of *active constraints* as

$$I(x, y) \triangleq \{i : g_i(x, y) = 0\}. \tag{5}$$

Then, the following definition facilitates an important assumption particularly for (KKT) optimality conditions.

Definition 2 A vector y^* satisfies the linear independence constraint qualification (LICQ) if the gradients of the active inequality constraints are linearly independent.

The following *blanket* assumptions similar to the ones in Outrata and Zowe [1995] are made throughout the paper.

(A1) $C(x) \neq \emptyset$ for all $x \in A$, where A is an open bounded set in $\Re^n$ such that X is contained in A.

(A2) $C(x)$ is uniformly compact on A, i.e., there exists an open bounded set $B \subset \Re^m$ such that $C(x) \subset B$ for all $x \in A$.

(A3) $F(x, \cdot)$ is strongly monotone.

(A4) X is nonempty and compact in $\Re^n$.

(A5) At each $x \in X$ and $y \in S(x)$, the partial gradients $\nabla_y g_i(x, y)$, $i \in I(x, y)$, satisfy (LICQ).

As a consequence of Theorem 1, the above assumptions imply that for every x in X, there exists a unique solution to VI(F, C). More specifically, assumption (A5) implies that every solution $y \in S(x)$ satisfies the KKT conditions

$$\begin{array}{c} F(x,y) - \nabla_y g(x,y)^T \lambda = 0, \\ g(x,y) \geq 0,\ \lambda \geq 0,\ \lambda^T g(x,y) = 0, \end{array} \tag{6}$$

where the multiplier $\lambda \in \Re^l$ is uniquely determined. Moreover, (2) and (6) are equivalent [Harker and Pang, 1990, Proposition 2.2]. Therefore, we can reformulate the problem (MPEC) as a standard nonlinear complementarity constrained optimization problem

$$\begin{array}{lc} \min & f(x,y) \\ \text{s.t.} & x \in X, \\ & F(x,y) - \nabla_y g(x,y)^T \lambda = 0, \\ & g(x,y) \geq 0,\ \lambda \geq 0,\ \lambda^T g(x,y) = 0. \end{array} \tag{NLC}$$

Though the problem (NLC) gives a smooth formulation, in general it does not satisfy any standard constraint qualification. In addition, the complementarity constraints are very difficult to handle [Falk and Liu, 1995]. Thus, following Facchinei et al. [1999], we consider a nonsmooth reformulation.

$$\begin{array}{lc} \min & f(x,y) \\ \text{s.t.} & x \in X, \\ & F(x,y) - \nabla_y g(x,y)^T \lambda = 0, \\ & g(x,y) - z = 0, \\ & \min(z, \lambda) = 0, \end{array} \tag{7}$$

where $z \in \Re^l$ is added to simplify the notation and the "min" operator is applied componentwise to the vectors z and λ. This problem can be further simplified to

$$\begin{array}{lc} \min & f(x,y) \\ \text{s.t.} & x \in X, \\ & H(x,y,z,\lambda) = 0, \end{array} \tag{NSM}$$

where $H : \Re^{n+m+2l} \to \Re^{m+2l}$ is defined as

$$H(x,y,z,\lambda) \triangleq \begin{pmatrix} F(x,y) - \nabla_y g(x,y)^T \lambda \\ g(x,y) - z \\ \min(z,\lambda) \end{pmatrix}. \tag{8}$$

The following proposition shows the equivalence of the problems (MPEC) and (NSM).

Theorem 2 [Facchinei et al., 1999] (x^*, y^*) is a global (local) solution of the problem (MPEC) if and only if there exists a vector (z^*, λ^*), which is a global (local) solution of the problem (NSM).

One major difficulty encountered in solving the problem (NSM) is the nondifferentiability of the "$\min$" function. One approach uses the *smoothing methods* to approximate the "$\min$" function [Chen and Mangasarian, 1995, Qi and Chen, 1995]. In particular, Kanzow [1996] used the following Chen-Harker-Kanzow-Smale (CHKS) smoothing function

$$\phi_\mu(a, b) = a + b - \sqrt{(a-b)^2 + 4\mu} \tag{9}$$

where μ is the real parameter. Kanzow [1996] applied this aprroach successfully for solving linear complementarity problems, and then for solving monotone variational inequalities [Kanzow and Jiang, 1998]. Recently, Facchinei et al. [1999] benefited from this function in handling MPEC problems.

Here, we exploit the use of *entropic regularization*, since this particular approach has been shown to have useful properties [Fang and Wu, 1996, Li and Fang, 1997, Fang et al., 1997].

3 Smoothing by Entropic Regularization

The smoothing function $\phi_p : \Re^2 \rightarrow \Re$ is defined as

$$\phi_p(a, b) \triangleq -\frac{1}{p} \ln\{e^{-pa} + e^{-pb}\} \tag{10}$$

where $p > 0$ is the real parameter. Notice that, for each $p \neq 0$, $\phi_p(a, b)$ is a C^∞ function. Furthermore, the following fundamental result shows that (10) defines a smooth approximation for the "$\min$" function.

Lemma 1 Let $a, b \in \Re$ then for any $\epsilon > 0$, there exists $P(\epsilon) > 0$ such that

$$\mid \phi_p(a, b) - min(a, b) \mid \leq \epsilon \text{ for all } p. \geq P(\epsilon) \tag{11}$$

Proof. Assume that $\min(a, b) = a$, then

$$\phi_p(a, b) = -\frac{1}{p} \ln\{e^{-pa} + e^{-pb}\} \tag{12}$$

$$= -\frac{1}{p} \ln\{e^{-pa}[1 + e^{-p(b-a)}]\} \tag{13}$$

$$= -\frac{1}{p} \ln\{e^{-pa}\} - \frac{1}{p} \ln\{1 + e^{-p(b-a)}\} \tag{14}$$

$$= a - \frac{1}{p} \ln\{1 + e^{-p(b-a)}\} \tag{15}$$

By plugging this into (11) we have

$$\mid a - \frac{1}{p} \ln\{1 + e^{-p(b-a)}\} - a \mid = \frac{1}{p} \ln\{1 + e^{-p(b-a)}\} \leq \frac{1}{p} \ln 2 \tag{16}$$

Therefore, the result follows with $P(\epsilon) = \frac{\ln 2}{\epsilon}$. ∎

In nonlinear optimization, one of the desirable properties of a function is its convexity. Our next result shows that proposed function, $-\phi_p$ is convex.

Lemma 2 Given any $p > 0$, ϕ_p is a concave function on $\Re^2$.

Proof. For any $u = (a_1, b_1), v = (a_2, b_2) \in \Re^2$ and $\alpha \in (0, 1)$, define

$$T_p(u, v) \triangleq \phi_p(\alpha a_1 + (1-\alpha)a_2, \alpha b_1 + (1-\alpha)b_2) \tag{17}$$

Then we have

$$T_p(u, v) = -\frac{1}{p} \ln\{e^{-\alpha p a_1} e^{-(1-\alpha)p a_2} + e^{-\alpha p b_1} e^{-(1-\alpha)p b_2}\} \tag{18}$$

$$= -\frac{1}{p} \ln\{(e^{-pa_1})^\alpha (e^{-pa_2})^{1-\alpha} + (e^{-pb_1})^\alpha (e^{-pb_2})^{1-\alpha}\}. \tag{19}$$

The Hölder Inequality [Kazarinoff, 1961, page 67] implies that

$$(e^{-pa_1})^\alpha (e^{-pa_2})^{1-\alpha} + (e^{-pb_1})^\alpha (e^{-pb_2})^{1-\alpha} \le (e^{-pa_1} + e^{-pb_1})^\alpha (e^{-pa_2} + e^{-pb_2})^{1-\alpha}. \tag{20}$$

Consequently,

$$T_p(u, v) \ge -\frac{\alpha}{p} ln\{e^{-pa_1} + e^{-pb_1}\} - \frac{1-\alpha}{p} ln\{e^{-pa_2} + e^{-pb_2}\} \tag{21}$$

$$= \alpha\phi_p(a_1, b_1) + (1-\alpha)\phi_p(a_2, b_2). \tag{22}$$

Therefore, ϕ_p is a concave function on $\Re^2$. ∎

Remark 1. Notice that a locally Lipschitz function is said to be regular if the directional derivative exists at all points and it is equal to its Clarke's directional derivative [Clarke, 1983]. As a consequence of Lemma 2, ϕ_p is a locally Lipschitz and regular function. This remark will be recalled in the sequel.

We now give a smooth reformulation of the problem (NSM). Let us define a nonlinear mapping $H_p : \Re^{n+m+2l} \to \Re^{m+2l}$,

$$H_p(x, y, z, \lambda) \triangleq \begin{pmatrix} F(x, y) - \nabla_y g(x, y)^T \lambda \\ g(x, y) - z \\ \Phi_p(z, \lambda) \end{pmatrix} \tag{23}$$

where

$$\Phi_p(z, \lambda) \triangleq (\phi_p(z_1, \lambda_1), \cdots, \phi_p(z_l, \lambda_l))^T \in \Re^l. \tag{24}$$

Thus, for each $p \in \Re$, we have

$$\begin{array}{ll} \min & f(x, y) \\ \text{s.t.} & x \in X, \\ & H_p(x, y, z, \lambda) = 0. \end{array} \tag{SM$_p$}$$

Notice that as a result of Lemma 1, problem (SM$_p$) is a smooth perturbation of the problem (NSM). Let the feasible set of problem (SM$_p$) be $\mathcal{F}_p \in \Re^{n+m+2l}$. To illustrate the formulation with entropic regularization, the following example is given.

Example 1 Monotone quasi-variational inequality problem [Harker, 1991].

$$
\begin{aligned}
& f(x,y) = \tfrac{1}{2}((x_1 - y_1)^2 + (x_2 - y_2)^2) \\
& X = [0,10] \times [0,10] \\
& F(x,y) = \begin{pmatrix} -34 + 2y_1 + \frac{8}{3}y_2 \\ -24.25 + 1.25y_1 + 2y_2 \end{pmatrix} \\
& g_1(x,y) = -x_2 - y_1 + 15 \\
& g_2(x,y) = -x_1 - y_2 + 15
\end{aligned}
\tag{25}
$$

The corresponding problems of (NLC) and (SM$_p$) can be derived as

$$
\begin{array}{ll}
\min & \frac{1}{2}((x_1 - y_1)^2 + (x_2 - y_2)^2) \\
\text{s.t.} & 0 \le x_1, x_2 \le 10 \\
& -34 + 2y_1 + \frac{8}{3}y_2 + \lambda_1 = 0 \\
& -24.25 + 1.25y_1 + 2y_2 + \lambda_2 = 0 \\
& -x_2 - y_1 + 15 \ge 0 \\
& -x_1 - y_2 + 15 \ge 0 \\
& \lambda_1, \lambda_2 \ge 0 \\
& \lambda_1(-x_2 - y_1 + 15) = 0 \\
& \lambda_2(-x_1 - y_2 + 15) = 0
\end{array}
\tag{26}
$$

and

$$
\begin{array}{ll}
\min & \frac{1}{2}((x_1 - y_1)^2 + (x_2 - y_2)^2) \\
\text{s.t.} & 0 \le x_1, x_2 \le 10 \\
& -34 + 2y_1 + \frac{8}{3}y_2 + \lambda_1 = 0 \\
& -24.25 + 1.25y_1 + 2y_2 + \lambda_2 = 0 \\
& -x_2 - y_1 + 15 - z_1 = 0 \\
& -x_1 - y_2 + 15 - z_2 = 0 \\
& -\frac{1}{p}\ln\{e^{-pz_1} + e^{-p\lambda_1}\} = 0 \\
& -\frac{1}{p}\ln\{e^{-pz_2} + e^{-p\lambda_2}\} = 0,
\end{array}
\tag{27}
$$

respectively.

Theorem 3 Given any $p \neq 0$ and $(x, y, z, \lambda) \in \mathcal{F}_p$, the Jacobian of H_p with respect to the variables (y, z, λ) is nonsingular.

Proof. Since ϕ_p, $p \neq 0$, is continuously differentiable on $\Re^2$, the operator H_p is also continuously differentiable. Hence, its Jacobian is given by

$$
\nabla H_p(x,y,z,\lambda) = \begin{pmatrix} Q & 0 & -A^T \\ A & -I_l & 0 \\ 0 & D_1 & D_2 \end{pmatrix},
\tag{28}
$$

where

$$Q \triangleq \nabla_y F(x,y) - \sum_{i \in I} \lambda_i \nabla_y^2 g_i(x,y) \tag{29}$$

$$A \triangleq \nabla_y g(x,y) \tag{30}$$

$$D_1 \triangleq \text{diag}\left(\frac{\partial \phi_p}{\partial z_1}(z_1,\lambda_1),\cdots,\frac{\partial \phi_p}{\partial z_l}(z_l,\lambda_l)\right) \tag{31}$$

$$D_2 \triangleq \text{diag}\left(\frac{\partial \phi_p}{\partial \lambda_1}(z_1,\lambda_1),\cdots,\frac{\partial \phi_p}{\partial \lambda_l}(z_l,\lambda_l)\right) \tag{32}$$

and I_l is the l-dimensional identity matrix. Notice that since $F(x,\cdot)$ is strongly monotone, $\nabla_y F(x,\cdot)$ is positive definite. Moreover since all functions $g_i(x,\cdot)$ are concave, their Hessian matrices $\nabla_y^2 g_i(x,\cdot)$ are negative semidefinite for $i \in I$. Therefore, the matrix Q is positive definite. For all $i = 1,2,\cdots,l$, since

$$\frac{\partial \phi_p(z_i,\lambda_i)}{\partial z_i} = \frac{e^{-pz_i}}{e^{-pz_i}+e^{-p\lambda_i}} \in (0,1) \text{ and} \tag{33}$$

$$\frac{\partial \phi_p(z_i,\lambda_i)}{\partial \lambda_i} = \frac{e^{-p\lambda_i}}{e^{-pz_i}+e^{-p\lambda_i}} \in (0,1), \tag{34}$$

the diagonal matrices D_1 and D_2 are positive definite. To show the nonsingularity of ∇H_p, assume that $\nabla H_p u = 0$ for some vector $u = (u^{(1)},u^{(2)},u^{(3)}) \in \Re^m \times \Re^l \times \Re^l$. Then

$$Qu^{(1)} - A^T u^{(3)} = 0 \tag{35}$$

$$Au^{(1)} - u^{(2)} = 0 \tag{36}$$

$$D_1 u^{(2)} + D_2 u^{(3)} = 0 \tag{37}$$

Substituting $u^{(2)}$ from (36) into (37) leads to

$$u^{(3)} = -D_2^{-1} D_1 A u^{(1)}. \tag{38}$$

Plugging $u^{(3)}$ into (35) yields

$$Qu^{(1)} + A^T D_2^{-1} D_1 A u^{(1)} = (Q + A^T D_2^{-1} D_1 A)u^{(1)} = 0. \tag{39}$$

Notice that $D_2^{-1} D_1$ and Q are positive definite, hence $u^{(1)} = 0$. From (37) and (38), we have $u^{(3)} = 0$ and $u^{(2)} = 0$. Therefore, ∇H_p is nonsingular. ∎

Remark 2. A vector function is said to be locally Lipschitz and regular if each of its components is locally Lipschitz and regular. We have already mentioned in Remark 1 that ϕ_p satisfies this property. Since all the remaining components of H_p are continuously differentiable, we can conclude that H_p is a locally Lipschitz and regular function.

Note that locally Lipschitz and regular functions are useful for the application of implicit function theorem, which leads to the convergence properties of the nonlinear programming algorithms [Outrata and Zowe, 1995, Kanzow, 1996].

4 Numerical Results

Note that specific algorithms can be developed for the entropic regularization approach. However, we are interested in investigating the performance of the proposed approach by using a publicly available software. In general, people tend to benefit from commercial or freeware software instead of creating their own tool. In this respect, we decided to use NEOS server, which provides access to many optimization solvers through the internet. Especially, the Kestrel interface enables the remote solution of optimization problems within the AMPL and GAMS modeling languages [Dolan and Munson, 2001]. We used AMPL as our modeling language, which is also available through the internet [AMPL, 2001].

SNOPT is one of the recently updated solvers on the NEOS server. It employs a sparse sequential quadratic programming algorithm with (quasi) Newton approximations. In particular, SNOPT allows the nonlinear constraints to be violated (if necessary) and minimizes the sum of such violations [Gill et al., 2000]. Therefore, we selected it to be the solver for our numerical study. We remark that we did not seek for optimal parameter settings for SNOPT but used the default parameters.

Recently, Facchinei et al. [1999] created a set of frequently cited problems, and reported their results. They developed their own algorithms and conducted computations by using a commercial package. We took advantage of their work and applied our approach to this set of problems. For fair comparision, we solved all the test problems by using both their smoothing approach and the proposed approach.

In Facchinei et al. [1999], the perturbation function (9) is used. We set the parameter, μ to be the initial choice of 1.0e-4, since they reported that their approach did not seem to depend critically on this choice. However, our limited experience showed that reducing the value of this parameter increases the precision at the cost of excessive iterations. Thus, we did not adjust this parameter unless we observed radical differences between our results and the results that they reported. We abbreviate their approach by FSA and entropic regularization approach by ERA in the subsequent discussion.

In Table 1, the first column numbers the test problems in the same order as Facchinei et al. reported. But, we initiated each problem with more different starting points than they did. Thus, the second column includes their starting points as well as additional starting points (which produced interesting results). Columns three and six give the best objective function values (f^*) reported by SNOPT. All the test problems are minimization problems, hence the smaller objective function values mean the better solutions are obtained. Corresponding to these objective function values, columns four and seven give the optimum solutions.

Notice that for both FSA and ERA, the required computational work at each iteration is about the same. Therefore, the number of iterations given in columns five and eight shows how quickly the solver converges to the reported objective function value. The last column gives the parameter p for entropic regularization approach. In each problem we initially set this parameter to 1.0e4.

Observe that, except for problem 11, FSA gave the same results as reported by Facchinei et al. [1999]. Note that, ERA demonstrated a remarkable performance and was able to solve each test problem in a reasonable number of iterations. For problems 1 to 5, FSA and ERA produced the same results. However, for problem 6, ERA perfomed exceptionally well. In fact, this result has never been reported in any work that we know of [Outrata and Zowe, 1995, Riesz et al., 1990]. To make sure, we reduced the parameter of FSA down to 1.0e-12, but the result did not change. For problem 7, although ERA was able to converge to the vicinity of the reported optimum solution, the precision of the objective function value is not as good as FSA.

For problems 8.1 to 10 both approaches were able to converge to the same optimum solutions. For problem 11, we reduced the parameter of FSA down to 1.0e-12. The optimum value was reached with the first starting point, but FSA failed with the remaining two starting points. ERA did not observe any difficulty in converging to the optimal solution with the first two starting points, but stuck in the local optimum with the last starting point.

One more observation worth mentioning. ERA slows down when an initial starting point is selected to be the 0 vector and the parameter p is not large enough (for problems 2, 9 and 11). Because in this case, $\phi_p(0,0) = -\frac{\ln 2}{p}$, and hence excessive iterations are spent to satisfy the violated equality constraint.

In our computational study, we observed that overflow problems may arise due to the exponential function in (10). We experienced this phenomenon for problems 6,7, 9 and 10. However, except problem 7, the results were not affected. Consequently, for this particular set of problems, we did not tackle this issue explicitly.

To handle the overflow problem, the smoothing function (10) can be slightly modified without changing its properties. For example, when we compute the exponential function e^{-pa} with $pa < 0$ and very large $|pa|$, the overflow problem occurs in the smoothing function. This problem can be avoided by using the following equality:

$$\phi_p(a,b) \triangleq -\frac{1}{p} \ln\{e^{-pa+pc} + e^{-pb+pc}\} + c \tag{40}$$

where c is any value satisfying $c \leq \min(a,b)$. Notice that both $-pa+pc \leq 0$ and $-pb+pc \leq 0$, hence the overflow problem is handled effectively.

		FSA			ERA			
No.	Start	f^*	x^*	It.	f^*	x^*	It.	p
1	0.0	3.2078	4.0605	26	3.2077	4.0604	32	1.0e4
	10.0	3.2078	4.0605	29	3.2077	4.0604	29	
2	0.0	3.4494	5.1536	28	3.4494	5.1536	33	1.0e4
	10.0	3.4494	5.1536	27	3.4494	5.1536	19	
3	0.0	4.6042	2.3894	21	4.6042	2.3894	29	1.0e4
	10.0	4.6042	2.3894	21	4.6041	2.3894	28	
4	0.0	6.5926	1.3731	30	6.5926	1.3731	23	1.0e4
	10.0	6.5926	1.3731	22	6.5926	1.3731	28	
5	(0.0, 0.0)	-0.9999	(0.5010, 0.5010)	15	-0.9999	(0.5010, 0.5010)	13	1.0e4
	(2.0, 2.0)	-0.9999	(0.5010, 0.5010)	19	-0.9999	(0.5010, 0.5010)	21	
6	0.0	-3266.6666	93.3333	7	-4512.5000	95.0000	10	1.0e2
	100.0	-3266.6666	93.3333	6	-4512.5000	95.0000	6	
	200.0	-3266.6666	93.3333	6	-4512.5000	95.0000	6	
7	(25.0, 25.0)	4.9996	(25.0012, 30.0001)	44	5.0151	(24.9978, 30.01812)	35	1.0e2
	(50.0, 50.0)	4.9996	(25.0012, 30.0001)	43	5.0151	(24.9978, 30.01812)	41	
8.1	0.0	-343.3452	55.5513	19	-343.3452	55.5513	19	1.0e4
	150.0	-343.3452	55.5513	23	-343.3452	55.5513	23	
8.2	0.0	-203.1550	42.5383	25	-203.1550	42.5383	24	1.0e4
	150.0	-203.1550	42.5383	22	-203.1550	42.5383	22	
8.3	0.0	-68.1356	24.1451	22	-68.1356	24.1450	20	1.0e4
	150.0	-68.1356	24.1451	22	-68.1356	24.1450	21	
8.4	0.0	-19.1540	12.3726	22	-19.1540	12.3726	20	1.0e4
	150.0	-19.1540	12.3726	21	-19.1540	12.3726	21	
8.5	0.0	-3.1611	4.7536	22	-3.1611	4.7536	20	1.0e4
	150.0	-3.1611	4.7536	21	-3.1611	4.7536	21	
8.6	0.0	-346.8931	50.0000	16	-346.8932	50.0000	21	1.0e4
	50.0	-346.8931	50.0000	17	-346.8931	50.0000	17	
8.7	0.0	-224.0371	39.7915	21	-224.0371	39.7913	25	1.0e4
	40.0	-224.0371	39.7915	17	-224.0372	39.7913	15	
8.8	0.0	-80.7859	24.2571	21	-80.7859	24.2571	29	1.0e4
	30.0	-80.7859	24.2571	19	-80.7859	24.2571	19	
8.9	0.0	-22.8371	13.0196	21	-22.8371	13.0196	20	1.0e4
	30.0	-22.8371	13.0196	19	-22.8371	13.0196	19	
8.10	0.0	-5.3491	6.0023	20	-5.3491	6.0023	20	1.0e4
	20.0	-5.3491	6.0023	19	-5.3491	6.0023	19	
9	(0.0, 0.0)	4.19e-12	(5.0000, 9.0000)	16	6.34e-11	(5.0000, 9.0000)	15	1.0e2
	(5.0, 5.0)	4.20e-13	(5.0000, 8.9999)	11	7.20e-14	(5.0000, 8.9999)	11	
	(10.0, 10.0)	2.23e-10	(9.0004, 5.9995)	27	4.50e-11	(9.0976, 5.9023)	31	
	(10.0, 0.0)	2.38e-10	(9.0004, 5.9995)	24	8.32e-11	(9.0937, 5.9062)	32	
	(0.0, 10.0)	6.19e-12	(5.0000, 9.0000)	18	6.19e-12	(5.0000, 9.0000)	19	
10	(0.0, 0.0, 0.0, 0.0)	-6600.0000	(7.0, 3.0, 12.0, 18.0)	83	-6600.0000	(7.0, 3.0, 12.0, 18.0)	59	1.0e2
	(0.0, 5.0, 0.0, 20.0)	-6600.0000	(7.0, 3.0, 12.0, 18.0)	81	-6600.0000	(7.0, 3.0, 12.0, 18.0)	68	
	(5.0, 0.0, 15.0, 10.0)	-6600.0000	(7.0, 3.0, 12.0, 18.0)	76	-6600.0000	(7.0, 3.0, 12.0, 18.0)	62	
	(5.0, 5.0, 15.0, 15.0)	-6600.0000	(7.0, 3.0, 12.0, 18.0)	79	-6600.0000	(7.0, 3.0, 12.0, 18.0)	66	
	(10.0, 5.0, 15.0, 10.0)	-6600.0000	(7.0, 3.0, 12.0, 18.0)	53	-6600.0000	(7.0, 3.0, 12.0, 18.0)	44	
11	(0.0, 0.0)	-12.6787	(0.0000, 2.0000)	40	-12.6787	(0.0000, 2.0000)	53	1.0e1
	(0.0, 2.0)	-8.6364	(1.8003, 0.0371)	20	-12.6787	(0.0000, 2.0000)	21	
	(2.0, 0.0)	-8.6364	(1.8003, 0.0371)	23	-10.3567	(2.0000, 0.0000)	38	

Table 1: Comparison of the entropic regularization approach with Facchinei et al. [1999]

5 Conclusion

We have proposed a regularization approach using an entropy function for solving mathematical programs with equilibrium constraints. The proposed approach has been shown that it can be easily implemented. Instead of developing any specific software, we have used a publicly available solver with default parameters to compare the proposed approach with recently reported results of Facchinei et al. [1999] over a set of test problems. Our numerical results have shown that proposed approach has a very promising performance and is able to converge to the best known results. Moreover, the proposed approach has even found better solutions than those reported in the literature.

References

AMPL. AMPL home page, 2001. http://www.ampl.com.

C. Chen and O.L. Mangasarian. Smoothing methods for convex inequalities and linear complementarity problems. *Mathematical Programming*, 71:51–69, 1995.

F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983.

E.D. Dolan and T.S. Munson. The Kestrel interface to the NEOS server, 2001. http://www-neos.mcs.anl.gov/neos/ftp/kestrel.ps.

F. Facchinei, H. Jiang, and L. Qi. A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming*, 85:107–134, 1999.

J.E. Falk and J. Liu. On bilevel programming, part I: general nonlinear cases. *Mathematical Programming*, 48:47–72, 1995.

S.-C. Fang, J.R. Rajasekera, and H.-S. Tsao. *Entropy Optimization and Mathematical Programming*. Kluwer Academic Publishers, Norwell, 1997.

S.-C. Fang and S.-Y. Wu. Solving min-max problems and linear semi-infinite programs. *Comput. Math. Applicat.*, 32:87–93, 1996.

T.L. Friesz, H.-J. Cho, N.J. Mehta, R.L. Tobin, and G. Anandalingam. A simulated annealing approach to the network design problem with variational inequality constraints. *Transportation Science*, 26:18–26, 1992.

P.E. Gill, W. Murray, M.A. Saunders, A. Drud, and E. Kalvelagen. GAMS/SNOPT: An SQP algorithm for large-scale constrained optimization, 2000. http://www.gams.com/docs/solver/snopt.pdf.

P.T. Harker. Generalized Nash games and quasi-variational inequalities. *Eur. J. Oper. Res.*, 54: 81–94, 1991.

P.T. Harker and J.-S. Pang. On the existence of optimal solutions to mathematical program with equilibrium constraints. *Operations Research Letters*, 7:61–64, 1988.

P.T. Harker and J.-S. Pang. Finite-dimensional and variational inequality and nonlinear complementarity problems: A survey of theory, algorithms, and applications. *Mathematical Programming*, 48:161–220, 1990.

Y. Ishizuka and E. Aiyoshi. Double penalty method for bilevel optimization problems. In G. Anandalingam and T.L. Friesz, editors, *Hierarchical Optimization. Annals of Operations Research*, volume 34, pages 73–88. Baltzer Science Publishers, 1992a.

Y. Ishizuka and E. Aiyoshi. Regularizations for two-level optimization problems. In *Advances in Optimization*, pages 239–255. Springer Verlag, 1992b.

C. Kanzow. Some noninterior continuation methods for linear complementarity problems. *SIAM J. Matrix Anal. Appl.*, 17:851–868, 1996.

C. Kanzow and H. Jiang. A continuation method for (strongly) monotone variational inequalities. *Mathematical Programming*, 89:103–125, 1998.

N.D. Kazarinoff. *Analytic Inequalities*. Holt, Rinehart and Winston, New York, 1961.

X.-S. Li and S.-C. Fang. On the entropic regularization method for solving min-max problems with applications. *Mathematical Methods of Operations Research*, 46:119–130, 1997.

Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1997.

F.H. Murphy, H.D. Sherali, and A.L. Soyster. A mathematical programming approach for determining oligopolistic market equilibrium. *Mathematical Programming*, 24:92–106, 1982.

J. Outrata and J. Zowe. A numerical approach to optimization problems with variational inequality constraints. *Mathematical Programming*, 68:105–130, 1995.

L. Qi and X. Chen. A globally convergent successive approximation methods for nonsmooth equations. *SIAM J. Control Optim.*, 33:402–418, 1995.

T.L. Riesz, R.L. Tobin, H.-J. Chao, and N.J. Mehta. Sensitivity analysis based heuristic algorithms for mathematical programs with variational inequalities. *Mathematical Programming*, 48:265–284, 1990.

C. Suwansirikul, T.L. Friesz, and R.L. Robin. Equilibrium decomposed optimization: A heuristic for the continuous equilibrium network design problem. *Transportation Science*, 21:254–263, 1987.