Smoothing Methods for Mathematical Programs with Equilibrium Constraints

Masao Fukushima
Department of Applied Mathematics and Physics
Graduate School of Informatics, Kyoto University
Kyoto 606-8501, Japan
fuku@amp.i.kyoto-u.ac.jp

Gui-Hua Lin*
Department of Applied Mathematics
Dalian University of Technology
Dalian 116024, China
lin_g_h@yahoo.com.cn

Abstract

In the recent optimization world, mathematical programs with equilibrium constraints (MPECs) have been receiving much attention and there have been proposed a number of methods for solving MPECs. In this paper, we provide a brief review of the recent achievements in the MPEC field and, as further applications of MPECs, we also mention the developments of the stochastic mathematical programs with equilibrium constraints (SMPECs).

1. Introduction

MPEC is a constrained optimization problem whose constraints include some parametric variational inequalities:

$$\begin{array}{ll} \text{minimize} & f(x,y) \\ \text{subject to} & (x,y) \in Z, \\ & y \text{ solves VI}(F(x,\cdot),C(x)). \end{array} \qquad (1)$$

Here, Z is a subset of $\Re^{n+m}$, $f : \Re^{n+m} \to \Re, F : \Re^{n+m} \to \Re^m$, $C : \Re^n \to 2^{\Re^m}$ are mappings, and $\mathrm{VI}(F(x,\cdot),C(x))$ denotes the *variational inequality* problem defined by the pair $(F(x,\cdot),C(x))$; that is, y solves $\mathrm{VI}(F(x,\cdot),C(x))$ if and only if $y \in C(x)$ and

$$(v-y)^T F(x,y) \geq 0, \quad \forall v \in C(x).$$

It is well-known [12] that, for a given variational inequality problem $\mathrm{VI}(G,Y)$, if the function G is the gradient mapping of a differentiable function $g : \Re^m \to \Re$ and the set Y is convex in $\Re^m$, then $\mathrm{VI}(G,Y)$ is a restatement of the first-order necessary conditions of optimality for the optimization problem

$$\begin{array}{ll} \text{minimize} & g(y) \\ \text{subject to} & y \in Y. \end{array}$$

Therefore, MPEC (1) can be regarded as a generalization of the so-called bilevel programming problem. Moreover, MPEC is also closely related to the well-known Stackelberg game, see [29, 31].

When $C(x) \equiv \Re^m_+$ for all x in problem (1), the parametric variational inequality constraints reduce to a parametric complementarity system and then problem (1) is equivalent to the following mathematical program with complementarity constraints (MPCC):

$$\begin{array}{ll} \text{minimize} & f(x,y) \\ \text{subject to} & (x,y) \in Z, \\ & y \geq 0,\ F(x,y) \geq 0, \\ & y^T F(x,y) = 0. \end{array} \qquad (2)$$

On the other hand, if the set-valued function C in problem (1) is given by

$$C(x) := \{y \in \Re^m |\ c(x,y) \leq 0\},$$

where $c : \Re^{n+m} \to \Re^s$ is continuously differentiable, then, under some suitable conditions, the variational inequality problem $\mathrm{VI}(F(x,\cdot),C(x))$ has an equivalent Karush-Kuhn-Tucker representation [33]:

$$\begin{array}{l} F(x,y) + \nabla_y c(x,y)\lambda = 0, \\ \lambda \geq 0, \quad c(x,y) \leq 0, \quad \lambda^T c(x,y) = 0, \end{array}$$

where λ is the Lagrange multiplier vector, and hence, problem (1) can be reformulated as a program like (2) under some conditions, see the monograph [29] for details. Hence, MPCCs constitute an important subclass of MPECs. For this reason, we particularly concentrate on this kind of MPECs.

MPEC plays a very important role in many fields such as engineering design, economic equilibrium, multilevel game, and mathematical programming theory itself. However, this problem is very difficult to deal with because, from the geometric point of view, its feasible region is not convex and not connected even in general, and in theory,

*Current address: Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. ghlin@amp.i.kyoto-u.ac.jp

its constraints fail to satisfy a standard constraint qualification such as the linear independence constraint qualification (LICQ) or the Mangasarian-Fromovitz constraint qualification (MFCQ) at any feasible point [4]. As a result, the developed nonlinear programming theory may not be applied to MPEC class directly. At present, a natural and popular approach is try to find some suitable approximations of an MPEC so that it can be solved by solving a sequence of ordinary nonlinear programs. Along this way, many methods have been developed in the literature. We will summarize these methods in Section 3.

Recently, as further applications of MPECs, stochastic mathematical programs with equilibrium constraints (SMPECs) have attracted people's attention. An SMPEC can be formulated as follows:

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x,y,\omega)] \\ \text{subject to} & (x,y) \in Z, \ \omega \in \Omega, \\ & y \text{ solves } \mathrm{VI}(F(x,\cdot,\omega), C(x,\omega)), \end{array} \tag{3}$$

where Z is a subset of $\Re^{n+m}$, Ω denotes the underlying sample space, E_ω means expectation with respect to the random variable $\omega \in \Omega$, and $f : \Re^{n+m} \times \Omega \to \Re, F : \Re^{n+m} \times \Omega \to \Re^m$, $C : \Re^n \times \Omega \to 2^{\Re^m}$ are mappings. Obviously, if Ω is a singleton, then problem (3) reduces to an ordinary MPEC, and so the SMPEC (3) can be thought as a generalization of the MPEC (1). The SMPEC (3) is also closely related to the so-called two-stage stochastic program with recourse [36]:

$$\begin{array}{ll} \text{minimize} & p(x) + E_\omega[Q(x,\omega)] \\ \text{subject to} & x \in X, \end{array} \tag{4}$$

where $p : \Re^n \to \Re$, $X \subseteq \Re^n$, and $Q : \Re^n \times \Omega \to \Re$ is defined by

$$Q(x,\omega) := \inf_{y \in Y(x,\omega)} g(y,\omega)$$

with $Y : \Re^n \times \Omega \to 2^{\Re^m}$ and $g : \Re^m \times \Omega \to \Re$. Many applications of problem (4) can be found in practice, especially in financial planning. See [2] for further details about problem (4).

Since an MPEC is already very hard to handle, SMPECs may be more difficult to deal with because the number of random events is usually very large in practice. The main developments of SMPECs will be reported in Section 4.

2. Preliminaries

For two vectors u and v in $\Re^s$, both $\min(u,v)$ and $\max(u,v)$ are understood to be taken componentwise. For a given function $G : \Re^s \to \Re^{s'}$ and a vector $u \in \Re^s$, $\nabla G(u)$ is the transposed Jacobian of G at u, whereas for a real valued function $g : \Re^s \to \Re$, $\nabla g(u)$ denotes the gradient vector of g at u. Moreover, we use

$$\mathcal{I}_G(u) := \{i \mid \ G_i(u) = 0\}$$

to stand for the active index set of G at u.

Consider the nonlinear programming problem

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & c_i(z) \le 0, \ \ i = 1, \cdots, l, \\ & c_i(z) = 0, \ \ i = l+1, \cdots, s, \end{array} \tag{5}$$

where $f : \Re^n \to \Re$ and $c : \Re^n \to \Re^s$ are twice continuously differentiable.

Definition 2.1 We say z to be a *stationary* point of problem (5) if it is feasible to (5) and there exists a Lagrange multiplier vector $\lambda \in \Re^s$ such that

$$\begin{array}{l} \nabla f(z) + \nabla c(z)\lambda = 0, \\ \lambda_i \ge 0, \ \lambda_i c_i(z) = 0, \quad i = 1, \cdots, l. \end{array}$$

Definition 2.2 Let z be a stationary point of problem (5) and λ be a Lagrange multiplier vector corresponding to z. We say the *weak second-order necessary condition* (WSONC) holds at z if we have

$$d^T \Big(\nabla^2 f(z) + \sum_{i=1}^{s} \lambda_i \nabla^2 c_i(z) \Big) d \ge 0$$

for any $d \in \mathcal{T}(z) := \{d \in \Re^n \mid \ d^T \nabla c_i(z) = 0, \ \ \forall \ i \in \mathcal{I}_c(z)\}$.

We next consider the MPCC

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0, \ h(z) = 0, \\ & G(z) \ge 0, \ H(z) \ge 0, \\ & G(z)^T H(z) = 0, \end{array} \tag{6}$$

where $f : \Re^n \to \Re, g : \Re^n \to \Re^p, h : \Re^n \to \Re^q$, and $G, H : \Re^n \to \Re^m$ are all twice continuously differentiable functions. Let $\mathcal{F}$ denote the feasible region of the above problem.

Definition 2.3 The MPEC-*linear independence constraint qualification* (MPEC-LICQ) is said to hold at $\bar{z} \in \mathcal{F}$ if the set of vectors

$$\Big\{ \nabla g_l(\bar{z}), \nabla h_r(\bar{z}), \nabla G_i(\bar{z}), \nabla H_j(\bar{z}) \ \Big| \ l \in \mathcal{I}_g(\bar{z}), \ r = 1, \cdots, q, \ i \in \mathcal{I}_G(\bar{z}), \ j \in \mathcal{I}_H(\bar{z}) \Big\}$$

is linearly independent.

This condition is not particularly stringent [41] and has been assumed often in the literature on MPCCs [9, 13, 15, 22, 39]. Note that this definition is different from the standard definition of LICQ in nonlinear programming theory that would require the gradient of the function $G(z)^T H(z)$ be linearly independent of the above vectors, which cannot happen in any case actually.

In the study of MPCCs, there are several kinds of stationarity defined for problem (6) [38].

Definition 2.4 We say $\bar{z} \in \mathcal{F}$ to be a *Bouligand or B-stationary* point of problem (6) if it satisfies

$$d^T \nabla f(\bar{z}) \geq 0, \ \forall d \in \mathcal{T}(\bar{z}, \mathcal{F}),$$

where $\mathcal{T}(\bar{z}, \mathcal{F})$ stands for the tangent cone of $\mathcal{F}$ at $\bar{z}$.

Definition 2.5 (1) $\bar{z} \in \mathcal{F}$ is called *weakly stationary* to problem (6) if there exist multiplier vectors $\bar{\lambda} \in \Re^p, \bar{\mu} \in \Re^q$, and $\bar{u}, \bar{v} \in \Re^m$ such that

$$\nabla f(\bar{z}) + \nabla g(\bar{z})\bar{\lambda} + \nabla h(\bar{z})\bar{\mu} - \nabla G(\bar{z})\bar{u} - \nabla H(\bar{z})\bar{v} = 0, \tag{7}$$

$$\bar{\lambda} \geq 0, \quad \bar{\lambda}^T g(\bar{z}) = 0, \tag{8}$$

$$\bar{u}_i = 0, \quad i \notin \mathcal{I}_G(\bar{z}), \tag{9}$$

$$\bar{v}_i = 0, \quad i \notin \mathcal{I}_H(\bar{z}). \tag{10}$$

(2) $\bar{z} \in \mathcal{F}$ is called a *Clarke or C-stationary* point of problem (6) if there exist multiplier vectors $\bar{\lambda} \in \Re^p, \bar{\mu} \in \Re^q$, and $\bar{u}, \bar{v} \in \Re^m$ such that (7)–(10) hold with

$$\bar{u}_i \bar{v}_i \geq 0, \quad i \in \mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z})$$

and we say $\bar{z}$ is *Mordukhovich or M-stationary* to problem (6) if, furthermore, either $\bar{u}_i > 0, \ \bar{v}_i > 0$ or $\bar{u}_i \bar{v}_i = 0$ for all $i \in \mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z})$.

(3) $\bar{z} \in \mathcal{F}$ is called a *strongly or S-stationary* point of problem (6) if there exist multiplier vectors $\bar{\lambda}, \bar{\mu}, \bar{u}$, and $\bar{v}$ such that (7)–(10) hold with

$$\bar{u}_i \geq 0, \ \ \bar{v}_i \geq 0, \quad i \in \mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}).$$

It is well-known [38] that, if the MPEC-LICQ holds at $\bar{z}$, B-stationarity is equivalent to S-stationarity. In general, in order to obtain a B-stationary point, some additional conditions are always assumed. The following is one of these conditions.

Definition 2.6 A weakly stationary point $\bar{z} \in \mathcal{F}$ of problem (6) is said to satisfy the *upper level strict complementarity* (ULSC) condition if there exist multiplier vectors $\bar{\lambda}, \bar{\mu}, \bar{u}$, and $\bar{v}$ satisfying (7)–(10) and

$$\bar{u}_i \bar{v}_i \neq 0, \quad i \in \mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}).$$

The ULSC condition is clearly weaker than the so-called *lower level strict complementarity* (LLSC) condition (which means $\mathcal{I}_G(\bar{z}) \cap \mathcal{I}_H(\bar{z}) = \emptyset$ and in this case, $\bar{z}$ is also said to be *nondegenerate*). Moreover, it is obvious that any M-stationary point of problem (6) satisfying the upper level strict complementarity condition must be a B-stationary point.

3. Methods for MPECs

There have been proposed several approaches such as relaxation approach, penalty function approach, active set identification approach, sequential quadratic programming (SQP) approach, interior point approach, and so on. Most methods are presented for the MPCC (2) or (6).

3.1. Relaxation Approach

It is the complementarity constraints that cause the main difficulties of an MPCC. In order to overcome this knotty problem, we may introduce some parameters to smooth or relax these constraints.

Consider the MPCC (6). Facchinei et al. [6] and Fukushima and Pang [9] make use of the smoothed Fischer-Burmeister function

$$\phi_\epsilon(a, b) := a + b - \sqrt{a^2 + b^2 + 2\epsilon^2} \tag{11}$$

to generate the following approximation of (6):

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \leq 0, \ h(z) = 0, \\ & \phi_{\epsilon_k}(G_i(z), H_i(z)) = 0, \\ & i = 1, \cdots, m, \end{array} \tag{12}$$

where $\epsilon_k > 0$ is a relaxation parameter. It is obvious that the function ϕ_{ϵ_k} is differentiable everywhere and

$$\phi_{\epsilon_k}(a, b) = 0 \quad \Leftrightarrow \quad a > 0, b > 0, ab = \epsilon_k.$$

Thus, we obtain a smooth approximation of problem (6). By letting $\epsilon_k \to 0$, we may expect to get a point with some kind of stationarity to the original MPCC.

Theorem 3.1 [9] *Let $\epsilon_k \downarrow 0$, z^k be a stationary point of problem (12), and the sequence $\{z^k\}$ converge to z^*. Suppose that the WSONC holds at each z^k, the MPEC-LICQ holds at z^*, and $\{z^k\}$ is asymptotically weakly nondegenerate. Then z^* is B-stationary to the MPCC (6).*

Roughly speaking, the asymptotically weak nondegeneracy of $\{z^k\}$ means that, for each $i \in \mathcal{I}_G(z^*) \cap \mathcal{I}_H(z^*)$, $G_i(z^k)$ and $H_i(z^k)$ approach zero in the same order of magnitude, see [9]. This property is obviously weaker than the

LLSC condition and even weaker than the ULSC condition [23].

Subsequently, some other relaxation methods are presented for solving (6). One is the regularization approximation suggested by Scholtes [39]:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0,\ h(z) = 0, \\ & G(z) \le 0,\ H(z) = 0, \\ & G_i(z)H_i(z) \le \epsilon_k, \\ & i = 1, \cdots, m, \end{array} \tag{13}$$

and another one was proposed by Lin and Fukushima [22] who employ a bi-hyperbola approximation:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0,\ h(z) = 0 \\ & G_i(z)H_i(z) \le \epsilon_k^2 \\ & (G_i(z) + \epsilon_k)(H_i(z) + \epsilon_k) \ge \epsilon_k^2 \\ & i = 1, \cdots, m. \end{array} \tag{14}$$

The main convergence result given in [39] can be stated as follows.

Theorem 3.2 [39] *Let $\epsilon_k \downarrow 0$, z^k be a stationary point of problem (13), and the sequence $\{z^k\}$ converge to z^*. Suppose the MPEC-LICQ holds at z^*. Then z^* is a C-stationary point of (6). If furthermore, the WSONC holds at each z^k and the ULSC holds at z^*, then z^* is B-stationary.*

This theorem remains valid for the method proposed in [22]. In addition, [22] gives the following result, where the conditions WSONC and ULSC are replaced by some new conditions which are new and relatively easy to verify in practice.

Theorem 3.3 [22] *Let $\epsilon_k \downarrow 0$ and z^k be a stationary point of problem (14) with Lagrangian multiplier vector $(\lambda_g^k, \lambda_h^k, \lambda_G^k, \lambda_H^k)$. Let β_k be the smallest eigenvalue of the matrix $\nabla_z^2 L_k(z^k, \lambda_g^k, \lambda_h^k, \lambda_G^k, \lambda_H^k)$, where L_k is the Lagrange function of problem (14). Suppose $\{z^k\}$ converge to z^* and the MPEC-LICQ holds at z^*. Then z^* is a B-stationary point of problem (6).*

On the other hand, Lin and Fukushima [25] study the MPCC (2) from another point of view. They use an expansive simplex instead of the nonnegative orthant involved in the complementarity constraints. In other words, the complementarity constraints are replaced by a variational inequality defined on an expansive simplex. It is well known that such a variational inequality problem can be represented by a finite number of inequalities. Based on this new idea, a relaxation method has been presented. It has been shown that the new method possesses similar properties to the ones introduced above, see [25] for more details.

3.2. Penalty Function Approach

Another way to deal with the complementarity constraints in (6) is to apply a penalty technique.

Noticing that problem (6) is equivalent to the problem

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0,\ h(z) = 0, \\ & \phi(G_i(z), H_i(z)) = 0, \\ & i = 1, \cdots, m, \end{array} \tag{15}$$

where ϕ is the Fischer-Burmeister function, i.e.,

$$\phi(a,b) := a + b - \sqrt{a^2 + b^2}, \tag{16}$$

Huang et al [15] suggested a method that penalizes all the constraints in problem (15); that is, the subproblem is an unconstrained optimization problem

$$\min_{z \in \Re^m} \ f_k(z),$$

where

$$f_k(z) \ := \ f(z) + \rho_k \Big(\sum_{i=1}^{p} (\max\{g_i(z), 0\})^2 + \sum_{i=1}^{q} (h_i(z))^2 + \sum_{i=1}^{m} (\phi(G_i(z), H_i(z)))^2 \Big)$$

and $\rho_k > 0$ is a penalty parameter. In addition, Hu and Ralph [13] proposed a method that penalizes the complementarity terms only; that is, the subproblem is a constrained optimization problem

$$\begin{array}{ll} \text{minimize} & f(z) + \rho_k G(z)^T H(z) \\ \text{subject to} & g(z) \le 0,\ h(z) = 0, \\ & G(z) \ge 0,\ H(z) \ge 0. \end{array} \tag{17}$$

These two methods possess similar properties. As in the standard nonlinear programming theory, a problem of the penalty approach is that the feasibility of a limit point to the original problem cannot be ensured in general. A comprehensive investigation of the feasibility of a point obtained by solving a sequence of the problems (17) has been made in [13]. Our computational experience shows that the penalty approach is effective for solving MPCCs.

Other penalty methods can be found in [14, 27, 29, 40] and the references therein.

3.3. Active Set Identification Approach

Most existing methods for MPCCs require to solve an infinite sequence of nonlinear programs. Recently, we proposed some hybrid algorithms that enable us to compute a solution or a point with some kind of stationarity to problem (6) by solving a finite number of nonlinear programs in [23, 24].

Consider the MPCC (6). We employ the model (12) to describe the method given in [23]. Suppose that $\epsilon_k \downarrow 0$, the sequence $\{z^k\}$ generated by solving (12) converges to z^*, and the MPEC-LICQ holds at z^*. Our idea is based on the fact that z^* is B-stationary to (6) if and only if z^* is stationary to the relaxed problem

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0,\ h(z) = 0, \\ & G_i(z) \ge 0,\ H_i(z) = 0,\ \ i \in \alpha(z^*) \\ & G_i(z) = 0,\ H_i(z) = 0,\ \ i \in \beta(z^*) \\ & G_i(z) = 0,\ H_i(z) \ge 0,\ \ i \in \gamma(z^*) \end{array} \tag{18}$$

with nonnegative multipliers associated with the constraints

$$G_i(z) = 0,\ H_i(z) = 0, \quad i \in \beta(z^*)$$

(see [11]), where

$$\begin{array}{rcl} \alpha(z^*) & := & \{i \mid G_i(z^*) > 0,\ H_i(z^*) = 0\}, \\ \beta(z^*) & := & \{i \mid G_i(z^*) = 0,\ H_i(z^*) = 0\}, \\ \gamma(z^*) & := & \{i \mid G_i(z^*) = 0,\ H_i(z^*) > 0\}. \end{array}$$

Problem (18) is no longer an MPCC and it is clear that, if z^* is an optimal solution of (6), then it must be an optimal solution of problem (18). Therefore, if we can obtain the index sets $\alpha(z^*)$, $\beta(z^*)$ and $\gamma(z^*)$, we may expect to get z^* by solving (18). So, our purpose is to identify $\alpha(z^*), \beta(z^*), \gamma(z^*)$ in finite steps.

We try to construct some index sets $\alpha^k, \beta^k, \gamma^k$ from the current point z^k such that $(\alpha^k, \beta^k, \gamma^k)$ is a partition of $\{1, \cdots, n\}$ and

$$\alpha^k = \alpha(z^*),\ \beta^k = \beta(z^*),\ \gamma^k = \gamma(z^*) \tag{19}$$

hold when k is large enough. At each iteration, we solve the problem

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g(z) \le 0,\ h(z) = 0, \\ & G_i(z) \ge 0,\ H_i(z) = 0,\ \ i \in \alpha^k, \\ & G_i(z) = 0,\ H_i(z) = 0,\ \ i \in \beta^k, \\ & G_i(z) = 0,\ H_i(z) \ge 0,\ \ i \in \gamma^k. \end{array} \tag{20}$$

Denote by $\hat{z}^k$ a stationary point of problem (20). If the Lagrange multipliers corresponding to the constraints

$$\begin{array}{ll} G_i(z) \ge 0,\ H_i(z) = 0, & i \in \alpha^k \cap \beta(\hat{z}^k), \\ G_i(z) = 0,\ H_i(z) = 0, & i \in \beta^k, \\ G_i(z) = 0,\ H_i(z) \ge 0, & i \in \gamma^k \cap \beta(\hat{z}^k) \end{array} \tag{21}$$

are all nonnegative, then $\hat{z}^k$ is a B-stationary point of problem (6) under the MPCC-LICQ assumption at the point.

The key to success is to define the index sets α^k, β^k and γ^k such that condition (19) holds when k sufficiently large. To this end, we employ an *identification function* $\rho : \Re^n \to [0, +\infty)$ satisfying

$$\lim_{k\to\infty} \rho(z^k) \ = \ 0 \tag{22}$$

and, for all k large enough,

$$\max_{i\in\beta(z^*)} \left\{ G_i(z^k), H_i(z^k) \right\} \le \rho(z^k), \tag{23}$$

$$\max_{i\in\alpha(z^*)\cup\gamma(z^*)} \left\{ \min\{G_i(z^k), H_i(z^k)\} \right\} \le \rho(z^k). \tag{24}$$

We present the following hybrid algorithm.

Algorithm H:

Step 0: Choose $\epsilon_0 > 0$ and set $k := 0$.

Step 1: Solve problem (12) and denote by z^k one of its stationary points. Set

$$\begin{array}{rcl} \alpha^k & := & \{\, i \mid\ G_i(z^k) > \rho(z^k),\ \ H_i(z^k) \le \rho(z^k)\ \}, \\ \beta^k & := & \{\, i \mid\ G_i(z^k) \le \rho(z^k),\ \ H_i(z^k) \le \rho(z^k)\ \}, \\ \gamma^k & := & \{1, \cdots, n\} \setminus (\alpha^k \cup \beta^k), \end{array}$$

and go to Step 2.

Step 2: Solve problem (20) to get a stationary point $\hat{z}^k$. If the Lagrange multipliers corresponding to the constraints (21) are all nonnegative, then terminate. Else, go to Step 3.

Step 3: If a stopping rule is satisfied at z^k, terminate. Otherwise, choose an $\epsilon_{k+1} \in (0, \epsilon_k)$ and let $k := k + 1$. Return to Step 1.

Theorem 3.4 [23] *Suppose the sequence $\{z^k\}$ generated by Algorithm H converges to z^* and is asymptotically weakly nondegenerate. For given $\tau > 0$ and $\sigma \in (0, 1)$, let*

$$\rho(z) := \tau \| \min(G(z), H(z)) \|^\sigma \ \text{ or } \ \tau \|\Phi_0(z)\|^\sigma,$$

where

$$\Phi_0(z) := (\phi(G_1(z), H_1(z)), \cdots, \phi(G_m(z), H_m(z)))^T.$$

Then conditions (22)–(24) are satisfied; i.e., ρ can serve as an identification function. Furthermore, condition (19) holds for all k large sufficiently.

Note that the subproblem (12) can be replaced by any models mentioned in the previous subsections.

Another two kinds of hybrid methods, one of which makes use of an index addition strategy and the other adopts an index subtraction strategy, have been presented in [24]. It has been shown that both can identify the correct index sets without the asymptotically weakly nondegeneracy assumption.

3.4. SQP Approach, etc.

SQP approach is another important way to modify the complementarity structure in an MPCC. In particular, Fukushima et al [8] consider the mathematical programs with linear complementarity constraints. Based on a reformulation of the complementarity constraints as a system of semismooth equations by means of the Fischer-Burmeister function (16), the authors proposed an SQP algorithm by applying a penalty technique. Global convergence of the algorithm has been established. Jiang and Ralph [18] consider the ordinary MPCC (2) and, with the help of the smoothed Fischer-Burmeister function (11), they also suggested some globally convergent SQP algorithms. More recently, Fletcher et al [7] presented a locally superlinearly convergent SQP method for an MPCC. In addition, a piecewise SQP approach can be found in [29].

Other methods proposed for MPCCs so far include the implicit programming methods [3, 29], interior point methods [28, 29], implementable active-set method [11], and nonsmooth methods [31].

4. Methods for SMPECs

Since in many practical problems, some elements may involve uncertain data, it is important to study the SMPECs. We next focus on the SMPCC subclass.

There are two kinds of SMPCCs studied in the literature: One is the *lower-level wait-and-see* model

$$\begin{array}{ll} \text{minimize} & E_\omega[f(x, y(\omega), \omega)] \\ \text{subject to} & (x, y(\omega)) \in Z, \;\; \omega \in \Omega, \\ & y(\omega) \geq 0, \; F(x, y(\omega), \omega) \geq 0, \\ & y(\omega)^T F(x, y(\omega), \omega) = 0, \end{array} \tag{25}$$

in which the upper-level decision x is made at once and the lower-level decision y may be made after a random event is observed. The other is the following *here-and-now* model that requires us to make all decisions before a random event is observed:

$$\begin{array}{ll} \text{minimize} & E_\omega[c(x, y, \omega) + d^T z(\omega)] \\ \text{subject to} & (x, y) \in Z, \\ & y \geq 0, \; F(x, y, \omega) + z(\omega) \geq 0, \\ & y^T (F(x, y, \omega) + z(\omega)) = 0, \\ & z(\omega) \geq 0, \;\; \omega \in \Omega. \end{array} \tag{26}$$

Here, $c : \Re^{n+m} \times \Omega \to \Re$, $z(\omega)$ is a recourse variable, and $d \in \Re^m$ is a constant vector with positive elements. We are particularly interested in the here-and-now decision model because it seems more realistic.

The lower-level wait-and-see model was first discussed in [36], including the study on the existence of solutions, the convexity and directional differentiability of the implicit objective function, and the links between the model and two-stage stochastic programs with recourse.

Lin et al [20] discussed both the lower-level wait-and-see and here-and-now decision problems. For the lower-level wait-and-see problem (25), they proposed a smoothing implicit programming method and established a comprehensive convergence theory. With the help of a penalty technique, they also suggested a similar method for the here-and-now decision problem (26).

Subsequently, [21, 26] consider the following special here-and-now problem:

$$\begin{array}{ll} \text{minimize} & f(x, y) + \sum_{\ell=1}^{L} p_\ell d^T z_\ell \\ \text{subject to} & g(x, y) \leq 0, \; h(x, y) = 0, \\ & y \geq 0, \; F_\ell(x, y) + z_\ell \geq 0, \\ & y^T (F_\ell(x, y) + z_\ell) = 0, \\ & z_\ell \geq 0, \;\; \ell = 1, \cdots, L, \end{array} \tag{27}$$

where p_ℓ means the probability of a random event ω_ℓ. It has been shown [21] that the stochastic complementarity problem may be formulated as this kind of SMPECs. By the duality theorem in nonlinear programming theory, we can show that problem (27) is equivalent to

$$\begin{array}{ll} \text{minimize} & f(x, y) + \sum_{\ell=1}^{L} p_\ell Q_\ell(x, y) \\ \text{subject to} & g(x, y) \leq 0, \; h(x, y) = 0, \; y \geq 0, \end{array} \tag{28}$$

where, for each ℓ, $Q_\ell : \Re^{n+m} \to [0, +\infty]$ is defined by

$$Q_\ell(x, y) := \sup_{u+ty \leq d, \; u \geq 0, \; t \leq 0} -(u + ty)^T F_\ell(x, y),$$

and (28) is further equivalent to

$$\begin{array}{ll} \text{minimize} & f(x, y) + \sum_{\ell=1}^{L} p_\ell d^T \max(-F_\ell(x, y), 0) \\ \text{subject to} & g(x, y) \leq 0, \; h(x, y) = 0, \;\; y \geq 0, \\ & y_i F_{\ell,i}(x, y) \leq 0, \;\; \forall i, \; \forall \ell, \end{array} \tag{29}$$

see [21] for details. However, on the one hand, the function Q_ℓ may be neither finite-valued nor differentiable everywhere in general, and on the other hand, the objective function in problem (29) is not differentiable everywhere and the problem has a great many constraints because L is usually very large in practice. Therefore, both (28) and (29) may not be easy to solve directly. In view of these flaws, [21] presented a smoothing penalty approach based on the reformulation (29) and [26] suggested a regularization method based on the reformulation (28). Convergence analysis has also been given.

5. Concluding Remarks

The methods reviewed in this paper primarily aim at computing a local optimal solution. There have been proposed a number of algorithms, typically based on a branch and bound technique, for finding a global optimal solution. Since MPEC is NP hard, such methods have limitations in the size of problems they can handle. Nevertheless it is definitely important to develop practically useful methods for finding a global optimal solution of MPEC.

The methods for SMPECs mentioned in Section 4 assume that random variables have discrete distribution. Even in this case, the problem may not be very tractable when the sample space contains a large number of events. Moreover, when a random variable has continuous distribution, the approaches described in this paper cannot be applied directly. One may possibly use some Monte Carlo type simulation techniques to generate approximations to the SMPEC. Development of practically effective methods for SMPECs will certainly enhance the importance of SMPECs as a practical modelling tool for real-world problems.

Recent development of methods for MPECs have mainly been concerned with static models. In the context of game theory, discrete or continuous time dynamic versions of Stackelberg games or bilevel optimization problems have been studied by a number of authors, see [1]. From the computational point of view, however, methods that can handle more general dynamic Stackelberg games are rather scarce. Moreover, the existing methods do not seem applicable to the case where the lower level constraints are represented as variational inequalities, rather than an optimization problem. It does not seem easy at all to deal with such general constraints, which brings us a very challenging subject.

Pang and Fukushima [35] have studied a multi-leader-follower game where leaders play a non-cooperative game while playing a Stackelberg-type game with followers. The resulting problem may thus be regarded as an *Equilibrium Program with Equilibrium Constraints (EPEC)*. This problem may in general fail to have a solution because of its inherent non-convexity. Therefore we need to introduce a reasonable solution concept that enable us to characterize a possible outcome of the game. Study on EPEC is still in its infancy and so much remains to be done in this extremely difficult but exciting problem.

References

[1] Başar, T. and Olsder, G.J., *Dynamic Noncooperative Game Theory*, Second Edition, Academic Press, New York, NY, 1995.

[2] Birge, J.R. and Louveaux, F., *Introduction to Stochastic Programming*, Springer, New York, 1997.

[3] Chen, X. and Fukushima, M., A smoothing method for a mathematical program with P-matrix linear complementarity constraints, *Computational Optimization and Applications*, to appear.

[4] Chen, Y. and Florian, M., The Nonlinear Bilevel Programming Problem: Formulations, Regularity and Optimality Conditions, *Optimization*, 32 (1995), pp. 193-209.

[5] Cottle, R.W., Pang, J.S., and Stone, R.E., *The Linear Complementarity Problem*, Academic Press, New York, NY, 1992.

[6] Facchinei, F., Jiang, H., and Qi, L., A smoothing method for mathematical programs with equilibrium constraints, *Mathematical Programming*, 85 (1999), pp. 107-134.

[7] Fletcher, R., Leyffer, S., Ralph, D., and Scholtes, S., Local Convergence of SQP Methods for Mathematical Programs with Equilibrium Constraints, Numerical Analysis Report, Department of Mathematics, University of Dundee, Dundee, Scotland, 2001.

[8] Fukushima, M., Luo, Z.Q., and Pang, J.S., A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints, *Computational Optimization and Applications*, 10 (1998), pp. 5-34.

[9] Fukushima, M. and Pang, J.S., Convergence of a smoothing continuation method for mathematical problems with complementarity constraints, Ill-posed Variational Problems and Regularization Techniques, *Lecture Notes in Economics and Mathematical Systems*, Vol. 477, M. Théra and R. Tichatschke (eds.), Springer-Verlag, Berlin/Heidelberg, 1999, pp. 105-116.

[10] Fukushima, M. and Pang, J.S., Some feasibility issues in mathematical programs with equilibrium constraints, *SIAM Journal on Optimization*, 8 (1998), pp. 673-681.

[11] Fukushima, M. and Tseng, P., An implementable active-set algorithm for computing a B-stationary point of the mathematical program with linear complementarity constraints, *SIAM Journal on Optimization*, 12 (2002), pp. 724-739.

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

[13] Hu, X. and Ralph, D., Convergence of a penalty method for mathematical programming with equilibrium constraints, *Journal of Optimization Theory and Applications*, to appear.

[14] Huang, X.X., Yang, X.Q., and Teo, K.L., Partial augmented Lagrangian method and mathematical programs with complementarity constraints, manuscript, Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, 2002.

[15] Huang, X.X., Yang, X.Q., and Zhu, D.L., A sequential smooth penalization approach to mathematical programs with complementarity constraints, manuscript, Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, 2001.

[16] Jiang, H. and Ralph, D., Extension of quasi-Newton methods to mathematical programs with complementarity constraints, *Computational Optimization and Applications*, 25 (2003), pp. 123-150.

[17] Jiang, H. and Ralph, D., QPECgen, a MATLAB generator for mathematical programs with quadratic objectives and affine variational inequality constraints, *Computational Optimization and Applications*, 13 (1999), pp. 25-59.

[18] Jiang, H. and Ralph, D., Smooth SQP methods for mathematical programs with nonlinear complementarity constraints, *SIAM Journal on Optimization*, 10 (2000), pp. 779-808.

[19] Kall, P. and Wallace, S.W., *Stochastic Programming*, John Wiley & Sons, Chichester, 1994.

[20] Lin, G.H., Chen, X., and Fukushima, M., Smoothing implicit programming approaches for stochastic mathematical programs with linear complementarity constraints, Technical Report 2003-006, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan (2003).

[21] Lin, G.H. and Fukushima, M., A class of stochastic mathematical programs with complementarity constraints: Reformulations and algorithms, Technical Report 2003-010, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan, 2003.

[22] Lin, G.H. and Fukushima, M., A modified relaxation scheme for mathematical programs with complementarity constraints, *Annals of Operations Research*, to appear.

[23] Lin, G.H. and Fukushima, M., Hybrid algorithms with active set identification for mathematical programs with complementarity constraints, Technical Report 2002-008, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan, 2002.

[24] Lin, G.H. and Fukushima, M., Hybrid algorithms with index addition and subtraction strategies for solving mathematical programs with complementarity constraints, Technical Report 2003-003, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan, 2003.

[25] Lin, G.H. and Fukushima, M., New relaxation method for mathematical programs with complementarity constraints, *Journal of Optimization Theory and Applications*, 118 (2003), pp. 81-116.

[26] Lin, G.H. and Fukushima, M., Regularization method for stochastic mathematical programs with complementarity constraints, Technical Report 2003-012, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto, Japan, 2003.

[27] Lin, G.H. and Fukushima, M., Some exact penalty results for nonlinear programs and their applications to mathematical programs with equilibrium constraints, *Journal of Optimization Theory and Applications*, 118 (2003), pp. 67-80.

[28] Liu, X. and Sun, J., Generalized stationary pointd and an interior point method for mathematical programs with equilibrium constraints, Preprint, National University of Singapore, Singapore, 2002.

[29] Luo, Z.Q., Pang, J.S., and Ralph, D., *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, United Kingdom, 1996.

[30] Luo, Z.Q., Pang, J.S., Ralph, D. and Wu, S.Q., Exact penalization and stationary conditions of mathematical programs with equilibrium constraints, *Mathematical Programming*, 75 (1996), pp. 19-76.

[31] Outrata, J.V., Kocvara, M., and Zowe, J., *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*, Kluwer Academic Publishers, Boston, MA, 1998.

[32] Outrata, J.V. and Zowe, J., A numerical approach to optimization problems with variational inequality constraints, *Mathematical Programming*, 68 (1995), pp. 105-130.

[33] Pang, J.S., Complementarity problems, *Handbook on Global Optimization*, R. Horst and P. Pardalos (eds.), Kluwer Academic Publishers, B.V. Dordrecht, 1994, pp. 271-338.

[34] Pang, J.S. and Fukushima, M., Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints, *Computational Optimization and Applications*, 13 (1999), pp. 111-136.

[35] Pang, J.-S. and Fukushima, M., Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games, *Computational Management Science*, to appear.

[36] Patriksson, M. and Wynter, L., Stochastic mathematical programs with equilibrium constraints, *Operations Research Letters*, 25 (1999), pp. 159-167.

[37] Rockafellar, R.T., *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[38] Scheel, H.S. and Scholtes, S., Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity, *Mathematics of Operations Research*, 25 (2000), 1-22.

[39] Scholtes, S., Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, *SIAM Journal on Optimization*, 11 (2001), pp. 918-936.

[40] Scholtes, S. and Stöhr, M., Exact penalization of mathematical programs with complementarity constraints, *SIAM Journal on Control and Optimization*, 37 (1999), pp. 617-652.

[41] Scholtes, S. and Stöhr, M., How stringent is the linear independence assumption for mathematical programs with complementarity constraints, *Mathematics of Operations Research*, 26 (2001), pp. 851-863.

[42] Zhang, J. and Liu, G., A new extreme point algorithm and its application in PSQP algorithms for solving mathematical programs with linear complementarity constraints, *Journal of Global Optimization*, 19 (2001), pp. 345-361.