ON THE CALCULATION OF NASH EQUILIBRIUM POINTS WITH THE AID OF THE SMOOTHING APPROACH

G. Bouza Allende
University of Havana
gema@matcom.uh.cu

ABSTRACT
Let us consider a game with n players where the set of possible strategies depend on the decision of the other player. In this case, if the players behave rationally the solution is a point of generalized Nash equilibrium (GNE). These points can be obtained as solutions of a special class of bilevel programs. In this work, the bilevel problem is substituted by a simpler model which can be solved by the so called smoothing approach for mathematical programs with complementarity constraints. We discuss if the hypothesis for the convergence of this method are generically fulfilled or not.
Keywords: bilevel problems, generalized Nash equilibrium, generic set, mathematical programs with complementarity constraints.
MSC: 90C30.

RESUMEN
Sea $\mathcal{G}$ un juego de n jugadores, donde el conjunto de estrategias factibles de cada jugador depende de la decisión de los otros. Asumiendo racionalidad, la solución sería un punto de Nashm el cual puede modelarse como solución de un problema de dos niveles. Una forma de solución de este tipo de modelos es mediante el método de suavización. En este trabajo discutimos cuál será el comportamiento de este método en el caso generico.

1. Introduction

Let us consider the game $\mathcal{G} = \{I, X_1, \ldots, X_n, u_1, \ldots, u_n\}$ where $I = \{1, 2, \ldots, n\}$ is the set of players, $X_i = \{x_i \in \mathbb{R}^m \mid g_{i,k}(x_i) \geq 0,\ k = 1, \ldots, K_i\}$ is the set of feasible strategies of player $i \in I$ and u_i is the utility function of player i. Roughly speaking if $x_i \in X_i$ is the strategy chosen by player i and we denote as $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^{nm}$ the vector of strategies then $u_i : \mathbb{R}^{nm} \to \mathbb{R}$ represents the profit of player i if player $j = 1, \ldots, n$, uses strategy x_j. If we assume that that all players are rational, then we can assume that each player gives the best response to the strategies of the other players. Then the chosen strategy will be a Nash equilibrium. The formal definition will be the following.

Definition 1.1 $x^* \in X_1 \times X_2 \times, \ldots, \times X_n$ *is a Nash equilibrium point if*
For each $i = 1, \ldots, n$

$$u_i(x^*_{-i}, y) \leq u(x^*),\ \forall y \in X_i$$

where $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m)$.

We will suppose that X_i depends of x_{-i} and that the set has the following (classical) structure

$$X(x_{-i}) = \{y \in \mathbb{R}^m \mid g_{i,k}(y, x_{-i}) \geq 0,\ k = 1, \ldots, K_j\}.$$

Then, at a Nash equilibrium it holds that, for each $i = 1, \ldots, n$:

$$x_i^* \hookrightarrow arg \min_{y \in X(x_{-i})} -u_i(x_{-i}^*, y),\ \forall\, i = 1, \ldots, n.$$

For its practical solution, we will write the model in a bilevel form

$$\begin{array}{c} \min \|x-y\|^2 \\ \text{s.t. } x_i \hookrightarrow \begin{cases} arg \min -u_i(x_{-i}, z), \\ \text{s.t. } z \in X(y_{-i}), \end{cases} \quad i = 1 \ldots, n \end{array} \tag{1..1}$$

A natural way of dealing with this type of model is to substitute the lower level problem by a simpler condition, for instance the KKT-system. This lead us to a mathematical program with complementarity constraints which can be solved by SQP, regularization techniques and the smoothing approach. In this work we will study which is the expected behavior of the smoothing approach for this particular model. We will prove that under relative general assumptions, the set of feasible solutions constructed by the algorithm will be non empty.
The paper is organized as follows, first we will present some properties of bilevel problems and its relations with other class of problems. Then we will discuss the convergence of the algorithm. The last section is devoted to obtain the characteristics of the set of feasible solutions of the resulting problem for almost all quadratic perturbation of the involved functions and their consequences for the solution algorithm.

2. Preliminary aspects

In this section we are going to present some properties of bilevel problems and mathematical programs with complementarity problems. We will discuss on the relations between both types of problems and the convergence of a solution method based on a parametric approach.
Bilevel programs solve the model

$$P_{BL}: \quad \min_{x,y} f(x,y) \quad \text{s.t. } (x,y) \in \mathcal{M}_{BL} \tag{2..1}$$

$$\mathcal{M}_{BL} = \left\{ (x,y) \in \mathbb{R}^{n+m} \;\middle|\; \begin{array}{c} g_j(x,y) \geq 0,\ j \in J = \{1, \ldots, q\}, \\ \text{and} \quad y \text{ solves } Q(x) \end{array} \right\}.$$

Here

$$Q(x): \min_y \phi(x,y) \text{ s.t. } y \in Y(x) = \{y \in \mathbb{R}^m \mid v_i(x,y) \geq 0,\ i \in I = \{1, \ldots, l\}\}$$

and $(f, g_1, \ldots, g_q) \in [C^2]^{1+q}_{n+m}$ and $(\phi, v_1, \ldots, v_l) \in [C^3]^{1+l}_{n+m}$.
Bilevel problems form an important class of mathematical programs. They appear for example in Cournot equilibrium models, in Stakelberg Games (*cf.*, [1]), and in semi-infinite programming (see [12], [11]).
BL problems are difficult to solve. Note that to check if $(x,y) \in \mathcal{M}_{BL}$, we have to guarantee that y is a solution of a non-linear problem, which is not an easy task in the general case. Moreover the feasible set of a BL may be non closed (see *e.g.*, [12]). During the last 20 years, books and many papers are dedicated to this topic, see *e.g.*, [1], [8], [4] and the references therein.

If for all x the functions $\phi, -v_1, \dots, -v_l$ are convex in y and the LICQ holds for all x at $Y(x)$, then a BL problem can be reduced into an MPCC, namely the lower level constraint that y has to solve the program $Q(x)$, is replaced by the KKT-conditions

$$\begin{array}{rcll} \nabla_y \phi(x,y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x,y) & = & 0, & \\ v_i(x,y) & \geq & 0, & i = 1,\dots,l, \\ \lambda_i & \geq & 0, & i = 1,\dots,l, \\ \lambda_i v_i(x,y) & = & 0, & i = 1,\dots,l. \end{array} \tag{2..2}$$

So, instead of P_{BL}, we consider the program

$$P_{\text{FJBL}}: \qquad \min_{x,y,\lambda} f(x,y) \quad \text{s.t. } (x,y,\lambda) \in \mathcal{M}_{\text{FJBL}}, \tag{2..3}$$

where

$$\mathcal{M}_{\text{FJBL}} = \left\{ (x,y,\lambda) \in \mathbb{R}^{n+m+l+1} \mid \text{(2..2) holds and } g_j(x,y) \geq 0, j \in J \,. \right\}$$

Here, $\mathcal{M}_{\text{FJBL}}|_{\mathbb{R}^n \times \mathbb{R}^m}$ denotes the projection of $\mathcal{M}_{\text{FJBL}}$ into the space $\mathbb{R}^n \times \mathbb{R}^m$.
Problem P_{FJBL} represent specially structured *Mathematical programs with Complementarity Constraints* (MPCC).

$$\begin{array}{c} \min\limits_z \; f(z) \\ \text{s.t.} \quad z \in \mathcal{M}_{CC} \end{array}$$

$$\mathcal{M}_{CC} = \left\{ z \in \mathbb{R}^n \;\middle|\; \begin{array}{rcll} h_k(z) & = & 0, & k = 1,\dots,q_0, \\ g_j(z) & \geq & 0, & j = 1,\dots,q, \\ r_i(z)s_i(z) & = & 0, & i = 1,\dots,l, \\ r_i(z), s_i(z) & \geq & 0, & i = 1,\dots,l \end{array} \right\} \tag{2..4}$$

These problems have a less complicated structure than the original BL. In particular the feasible set $\mathcal{M}_{\text{FJBL}}$ is always closed. For literature on MPCC we refer the reader *e.g.*, to [10], [9], [3], [5] and [2].
Note that although it is not difficult to check if a point is feasible or not, classical constraint qualifications do not hold and the set of feasible solutions is has a disjunctive structure. So, new methods of solutions have appear. One of them is the so called smoothing approach. The idea is that for $\tau \to 0^+$, find $z(\tau)$ solutions of:

$$\begin{array}{c} \min\limits_z \; f(z) \\ \text{s.t.} \quad z \in \mathcal{M}_{CC}(\tau) \end{array}$$

$$\mathcal{M}_{CC}(\tau) = \left\{ z \in \mathbb{R}^n \;\middle|\; \begin{array}{rcll} h_k(z) & = & 0, & k = 1,\dots,q_0, \\ g_j(z) & \geq & 0, & j = 1,\dots,q, \\ r_i(z)s_i(z) & = & \tau, & i = 1,\dots,l, \\ r_i(z), s_i(z) & \geq & 0, & i = 1,\dots,l \end{array} \right\} \tag{2..5}$$

It is desired that such a solutions exists when $\tau \to 0$ and that $z(\tau) \to z(0)$ when $\tau \to 0^+$.
For the convergence conditions we need the following definition

Definition 2.2 $\overline{z}$ *is a* non-degenerate solution of MPCC *if it is a non-degenerate critical point of the relaxed problem:*

$$\begin{array}{rl} P_R(\overline{z}): & \min f(z) \\ \textit{s.t.} & z \in \mathcal{M}_R \end{array} \tag{2..6}$$

$$\mathcal{M}_R = \left\{ z \in \mathbb{R}^n \;\middle|\; \begin{array}{c} g_j(z) \geq 0, \quad j = 1,\dots,q, \\ \begin{array}{rclrcll} r_i(z) & = & 0, & s_i(z) & \geq & 0, & i \in I_r(\overline{z}), \\ s_i(x) & = & 0, & r_i(x) & \geq & 0, & i \in I_s(\overline{z}), \\ r_i(z) & \geq & 0, & s_i(z) & \geq & 0, & i \in I_{rs}(\overline{z}). \end{array} \end{array} \right\}$$

where
$I_{rs}(\overline{z}) = \{i \mid r_i(\overline{z}) = s_i(\overline{z}) = 0\}$, $I_r(\overline{z}) = \{i \mid r_i(\overline{z}) = 0,\ s_i(\overline{z}) > 0\}$, $I_s(\overline{z}) = \{i \mid s_i(\overline{z}) = 0,\ r_i(\overline{z}) > 0\}$ *and* $J_0(\overline{z}) = \{j \mid g_j(\overline{z}) = 0\}$,.

We have the following result

Theorem 1 (cf. Bouza-Still (06)) *If $z(0)$ is a non-degenerate solution (in the MPCC sense) $\|z(\tau) - z(0)\| = O(\sqrt{\tau})$.*

This result includes the fact that is the LICQ hold at $z \in \mathcal{M}_R$ for all possible combination of the feasible solutions set, then M_τ is non empty. Moreover it holds that for all neighborhood V such that $\mathcal{M} \cap V \neq \varnothing$, then for τ small enough $\mathcal{M}_\tau \cap V \neq \varnothing$.
We want to point out that for generic f, g, r, s, the smoothing algorithm converges and $\|z(\tau) - z(0)\| = O(\sqrt{\tau})$. It is a clear consequence of Theorem 1 and the following result:

Theorem 2 (cf. Scholtes-Stöhr (01)) *Generically with respect to f, g, r, s all solutions of MPCC are non-degenerate solutions (in the MPCC sense).*

3. Genericity result

In this section the main results of this paper are obtained. We will begin by presenting some characteristics of the particular bilevel model we are dealing with. Then we will see if the hypothesis of Theorem 1 are fulfilled generically.
Let us recall the MPCC model resulting after applying the KKT-approach to model (1..1):

$$\begin{array}{rccl} & \min \|x-y\|^2 & & \\ \text{s.t.} & -\nabla_{x_i} u_i(x) - \sum_{k=1}^{K_i} \lambda_{i,k} \nabla_{x_i} g_{i,k}(x_i, y_{-i}) & = & 0, \qquad i = 1, \ldots, n \\ & g_{i,k}(x_i, y_{-i}) & \geq & 0, \quad i = 1, \ldots, n, k = 1, \ldots, K_i, \\ & \lambda_{i,k}(x) & \geq & 0, \quad i = 1, \ldots, n, k = 1, \ldots, K_i, \\ & \lambda_{i,k} g_{i,k}(x_i, y_{-i}) & \geq & 0, \quad i = 1, \ldots, n, k = 1, \ldots, K_i, \end{array} \tag{3..1}$$

In order to have simpler notations we take as $(x_i, y_{-i}) = (y_1, \ldots, y_{i-1}, x_i, x_{i+1}, \ldots, x_n)$.
Note that since $X(y_{-i})$ depends on y_{-i}, we can not assure that LICQ is stable under perturbations of the functions $u_i, g_{i,k},\ i = 1, \ldots, n, k = 1, \ldots, K_i$, see [6] for the particular case $y_{-i} \in \mathbb{R}$.
If LICQ is not satisfied, then we can not guarantee that the feasible set of problem (3..1) contains M_{BL}. However it is a common assumption that the utility function u_i is concave and that the set if feasible solutions $X(y_{-i})$ is convex. Then we can suppose that $-g_{i,k}$ are convex and the lower level problem will be convex. Using a classical result of convex programs see [7], it holds that a stationary point is an optimal solution of the lower problem, and hence feasible of the original model. We have shown the following proposition implies

Proposition 1 *If (x, y, λ) feasible point of problem (3..1) then it is a feasible solution of (1..1) .*
The solutions of (3..1) are at least a feasible solution of (1..1).

Now let us prove the genericity result

Theorem 3 *If $(u, g) \in [C^n_{mn} \times C^{\sum_{i=1}^n K_i}_{mn}]^\infty$, then for almost all $(b, d) \in \mathbb{R}^{nm} \times \mathbb{R}^{\sum_{i=1}^n K_i}$, MPCC-LICQ holds at all the feasible solutions of the MPCC defined by the functions $(u_i(x) + b_i^T x, g_{i,k}(x) + d_{i,k}), i = 1, \ldots, n, k = 1, \ldots, K_i$,.*

Proof:
Note that a feasible point of the problem defined by the (perturbed) functions $(u_i(x) + b_i^T x, g_{i,k}(x) + d_{i,k}), i = 1, \ldots, n, k = 1, \ldots, K_i$, fulfills

$$\begin{array}{rcl} -\nabla_{x_i} u_i(x) + b - \sum_{k=1}^{K_i} \lambda_{i,k} \nabla_{x_i} g_{i,k}(x_i, y_{-i}) & = & 0, \qquad i = 1, \ldots, n \\ g_{i,k}(x_i, y_{-i}) + d & = & 0, \quad i = 1, \ldots, n, k \in I^i, \\ \lambda_{i,k}(x) & = & 0, \quad i = 1, \ldots, n, k \in J^i, , \end{array} \tag{3..2}$$

for some sets $I^i = \{k : g_{i,k}(x) + c^T x + d = 0\}$ and $J^i = \{k : \lambda_{i,k} = 0\}$. For simplicity we assume that I^i contains the first l_i indexes and J^i the last $K_i - l_{i,2},\ l_{i,2} \leq l_i$. Note that the Jacobian with respect to x, y, d, c, b

$$\begin{array}{cccc} \partial_{x,y} & \partial_\lambda & \partial_b & \partial_d \\ \hline \otimes & \otimes & I_{nm} & 0 \\ \otimes & 0 & 0 & \begin{array}{ccc} I_{|I^1|}|0 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_{|I^n|}|0 \end{array} \\ 0 & \begin{array}{ccc} 0|I_{|J^1|} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 0|I_{|J^n|}|0 \end{array} & 0 & 0 \end{array}$$

has full row rank. Then by the parameterized Sard Lemma [cf. Lemma [6] for almost every (b, d) matrix of the derivatives with respect to (x, y, λ) has full row rank. This implies the desired MPCC-LICQ. □
This result implies we can guarantee the non-emptiness of $\mathcal{M}_\tau$.
For the convergence of the method we need to prove that generically solutions of (1..1) are non-degenerate. This has some practical difficulties since the objective function has a very particular structure.

References

[1] Bard, J. (1998): **Practical bilevel optimization. Algorithms and applications** Nonconvex Optimization and its Applications, 30. Kluwer Academic Publishers, Dordrecht.

[2] Bouza, G. and Still, G. (2007): Mathematical programs with complementarity constraints: Convergence properties of a smoothing method **Mathematics of Operations Research**, 32(2):467–483.

[3] Chen, X. and Fukushima, M. (2004): A smoothing method for a mathematical program with P-matrix linear complementarity constraints **Computational Optimization and Applications**, 27(3):223–246.

[4] Dempe, S. (2003): Annotated bibliography on bilevel programming and mathematical programs with equilibrium constrains **Optimization**, 52(3):333–359.

[5] Facchinei, F., Houyuan, J., and Liqun, Q. (1999): A smoothing method for mathematical programs with equilibrium constraints **Mathematical Programming**, 85(1):107–134.

[6] Gómez, W., Guddat, J., Jongen, H. T., Rückmann, J., and C.Solano (2000): **Curvas criticas y saltos en la optimizacion no lineal** http://www.emis.de/monographs/curvas/index.html.

[7] Luenberger, D. (1984): **Linear and nonlinear programming** Addison-Wesley Inc., Reading, Massachusetts, Second edition.

[8] Luo, Z. Q., Pang, J. S., and Ralph, D. (1996): **Mathematical Programs with Equilibrium Constraints** Cambridge University Press, Cambridge, U.K.

[9] Ralph, D. and Wright, S. J. (2004): Some properties of regularization and penalization schemes for MPEC's **Optimization Methods and Software**, 19(5):527–556.

[10] Scholtes, S. (2001): Convergence properties of a regularization scheme for mathematical programs with complementarity constraints **SIAM Journal on Optimization**, 11(4):918–936.

[11] Stein, O. (2003): **Bi-level strategies in semi-infinite programming** Nonconvex Optimization and its Applications, 71. Kluwer Academic Publishers, Boston, MA.

[12] Stein, O. and Still, G. (2002): On generalized semi-infinite optimization and bilevel optimization **European Journal of Operational Research**, 142(3):444–462.

Received July, 2007.
Revised December 2007.