QÜESTIIÓ, Vol 20, 2 pp. 215-222, 1996

A STABILITY THEOREM IN NONLINEAR BILEVEL PROGRAMMING**

SHOU-YANG WANG*, QIAN WANG† and LUIS COLADAS URIA‡

In this short paper, we are concerned with the stability of nonlinear bilevel programs. A stability theorem is proven and an example is given to illustrate this theorem.

Keywords: Bilevel programming, stability.

Bilevel programming, a nested optimization problem, emerged as the appropriate model to formulate a hierarchical decision making situation where the higher level in the hierarchy can only influence rather than dictate the choices of the lower level (Bard, 1984; Bialas and Karwan, 1984; Wang and Lootsma, 1994). Most of the investigations in this field are focused on optimality conditions and algorithms (see comments made in Chen and Florian, 1995; Wang, Wang and Romano-Rodríguez, 1994). Since a parametric solution or error bounds on a solution with perturbed data are typically of great interest both in practical applications and in theoretical characterizations (Fiacco, 1983), to study stability of an optimal solution to a bilevel programming problem is certainly a very important topic in bilevel programming.

** Supported in part by NSFC and MADIS. This paper was completed during the first author's stay at Departamento de Estadística e Investigación Operativa, Universidad de Santiago de Compostela, Spain. He is grateful to the financial support for his visit by Consejo Superior de Investigaciones Científicas (CSIC) and Universidad de Santiago de Compostela. The authors are very grateful to the referees' valuable comments and suggestions.

* Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, China.

† Department of Mathematics, University of Nebraska at Omaha, Omaha, Nebraska 68182-0243, USA.

‡ Departamento de Estadística e Investigación Operativa, Universidad de Santiago de Compostela, 15706 Santiago de Compostela, Spain.

– Article rebut el juliol de 1995.

– Acceptat el maig de 1996.

The bilevel programming problem with parameter considered in this paper is stated as $(BLP(\varepsilon))$:

$$\underset{x\in X}{\text{minimize}}\, F(x,y,\varepsilon) \text{ where } y \text{ solves}$$

$$\underset{y\in Y}{\text{minimize}}\, f(x,y,\varepsilon)$$

$$\text{subject to } g_i(x,y,\varepsilon) \leq 0, \quad i = 1,\cdots,m$$

where $x \in R^{n_1}$ and $y \in R^{n_2}$ are controlled by the leader and the follower respectively, X and Y are closed convex sets in R^{n_1} and R^{n_2} respectively, ε is a parameter vector in R^k, $F(x,y,\varepsilon)$ and $f(x,y,\varepsilon)$ are the objective functions of the leader and the follower respectively, $g_i(x,y,\varepsilon), i = 1,\cdots,m$, are the constraint functions.

When the parameter vector ε is identical with the zero vector (*i.e.*, no data is perturbed in the model), the problem $(BLP(\varepsilon))$ can be written as (BLP):

$$\underset{x\in X}{\text{minimize}}\, F(x,y) \text{ where } y \text{ solves}$$

$$\underset{y\in Y}{\text{minimize}}\, f(x,y)$$

$$\text{subject to } g_i(x,y) \leq 0, \quad i = 1,\cdots,m$$

where $F(x,y) = F(x,y,0)$, $f(x,y) = f(x,y,0)$ and $g_i(x,y) = g_i(x,y,0), i = 1,\cdots,m$. This type of bilevel programming problems have been extensively studied by many authors. Refer to Ben-Ayed (1993) and Wang and Wang (1994) for a survey.

For a given $x \in X$, we denote the inner program as $P(x,\varepsilon)$ and define

$$Y(x,\varepsilon) = \{y \mid y \text{ is a minimum of } P(x,\varepsilon)\},$$

$$\bar{S}(\varepsilon) = \{(x,y) \in X \times Y \mid g_i(x,y,\varepsilon) \leq 0, i = 1,\cdots,m, \text{ and } y \in Y(x,\varepsilon)\}$$

and

$$L(x,y,\varepsilon,u) = f(x,y,\varepsilon) + \sum_{i=1}^{m} u_i g_i(x,y,\varepsilon)$$

We make the following assumptions:

(i) the bilevel problem is well–posed, *i.e.*, $Y(x,\varepsilon)$ is a singleton and the unique element is denoted as $y(x,\varepsilon)$;

(ii) F, f and $g_i (i = 1,\cdots,m)$ are twice continuously differentiable in y, their gradients with respect to y and $g_i (i = 1,\cdots,m)$ are continuously differentiable in both x and ε, f is convex in y;

(iii) for any x, the second–order sufficient conditions for a minimum of $P(x,\varepsilon)$ holds at $y(x,\varepsilon)$, with associated Lagrange multipliers $u(x,\varepsilon)$ *i.e.*, for any $s \neq 0$ that satisfies

$$s^T \nabla_y g_i(x, y(x,\varepsilon),\varepsilon) = 0, i \in I_1(x,\varepsilon)$$
$$s^T \nabla_y g_i(x, y(x,\varepsilon),\varepsilon) \leq 0, i \in I_2(x,\varepsilon)$$

$s^T \nabla^2_{yy} L(x, y(x,\varepsilon,\varepsilon, u(x,\varepsilon)))s > 0$ holds where $I_1(x,\varepsilon) \triangleq \{j \mid g_j(x,y(x,\varepsilon),\varepsilon) = 0, u_j(x,\varepsilon) > 0\}$ and $I_2(x,\varepsilon) \triangleq \{j \mid g_j(x,y(x,\varepsilon),\varepsilon) = 0, u_j(x,\varepsilon) = 0\}$;

(iv) the gradients $\nabla_y g_i(x,y(x,\varepsilon),\varepsilon), i \in I_0(x,\varepsilon) \triangleq \{j \mid g_j(x,y(x,\varepsilon),\varepsilon) = 0\}$ are linearly independent;

(v) strict complementary slackness holds, *i.e.*, $u_i(x,\varepsilon) > 0$ when $i \in I_0(x,\varepsilon)$;

(vi) $F(x,y,\varepsilon)$ is continuous on $X \times Y \times R^k$ and X and Y are compact.

A pair $(x^*(\varepsilon), y^*(\varepsilon))$ is said to be *an optimal solution to* $(BLP(\varepsilon))$ if it satisfies (i) $y^*(\varepsilon) \in Y(x^*(\varepsilon),\varepsilon)$ and (ii) $F(x^*(\varepsilon), y^*(\varepsilon),\varepsilon) \leq F(x,y,\varepsilon)$ for any $(x,y) \in \bar{S}(\varepsilon)$.

Define

$$M(x) = \begin{bmatrix} \nabla^2_{yy} L(x,y(x,0),0,u(x,0)) & \nabla_y g_1(x,y(x,0),0), & \cdots & \nabla_y g_m(x,y(x,0),0) \\ u_1 \nabla_y g_1^T(x,y(x,0),0) & g_1(x,y(x,0),0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ u_m \nabla_y g_m^T(x,y(x,0),0) & 0 & \cdots & g_m(x,y(x,0),0) \end{bmatrix}$$

and

$$N(x) = [-\nabla^2_{\varepsilon x} L(x,y(x,0),0,u(x,0)), -u_1 \nabla_\varepsilon g_1(x,y(x,0),0), \cdots, -u_m \nabla_\varepsilon g_m(x,y(x,0),0)]^T.$$

Lemma 1

For any given $x \in X$, $M(x)$ is nonsingular.

Proof

Without loss of generality, let $I_0(x,0) = \{1,\cdots,p\}, I \setminus I_0(x,0) = \{p+1,\cdots,m\}$. Denote

$$G = (\nabla_y g_1(x,y(x,0),0), \cdots, \nabla_y g_p(x,y(x,0),0)),$$
$$\bar{G} = (\nabla_y g_{p+1}(x,y(x,0),0), \cdots, \nabla_y g_m(x,y(x,0),0)),$$
$$U = diag(u_1(x,0), \cdots, u_p(x,0))$$

and

$$D = diag(g_{p+1}(x,y(x,0),0), \cdots, g_m(x,y(x,0),0)).$$

Then

$$M(x) = \begin{pmatrix} \nabla^2_{yy}L(x,y(x,0),0,u(x,0)) & G & \bar{G} \\ UG^T & 0 & 0 \\ 0 & 0 & D \end{pmatrix}.$$

From Assumption (v) and the assumption of $I_0(x,0)$, $u_i(x,0) > 0, i = 1,\cdots,p$ and $g_j(x,y(x,0),0) < 0, j = p+1,\cdots,m$. So the matrices U and D are nonsingular. Hence, it is only required to show that matrix

$$\begin{pmatrix} \nabla^2_{yy}L(x,y(x,0),0,u(x,0)) & G \\ UG^T & 0 \end{pmatrix}$$

is nonsingular. This is equivalent to prove that the following system

$$\nabla^2_{yy}L(x,y(x,0),0,u(x,0))s - Gz = 0 \tag{1a}$$

$$UG^T s = 0 \tag{1b}$$

has the unique solution $s = 0, z = 0$.

From (1b), we get $G^T s = 0$. Hence, s satisfies Assumption (iii). Multiplying (1a) by s, we have

$$s^T\nabla^2_{yy}L(x,y(x,0),0,u(x,0))s - s^T Gz = 0,$$

$$s^T\nabla^2_{yy}L(x,y(x,0),0,u(x,0))s = 0.$$

By Assumption (iii), we get $s = 0$. Thus, $Gz = 0$. Owing to Assumption (iv), the column rank of G is full. Hence, $z = 0$. Therefore, $M(x)$ is nonsingular.

■

Lemma 2

For any given $\bar{x} \in X$, the following first-order approximation

$$\begin{bmatrix} y(\bar{x},\varepsilon) \\ u(\bar{x},\varepsilon) \end{bmatrix} = \begin{bmatrix} y(\bar{x},0) \\ u(\bar{x},0) \end{bmatrix} + M(\bar{x})^{-1}N(\bar{x})\varepsilon + o(||\varepsilon||) \tag{2}$$

holds in a neighborhood of $\varepsilon = 0$.

Proof

From Assumption (iii), we know that

$$\nabla_y L(x,y,\varepsilon,u) = 0 \tag{3a}$$

$$u_i g_i(x,y,\varepsilon) = 0, i = 1,\cdots,m \tag{3b}$$

hold at $(\bar{x}, y(\bar{x}, 0), 0, u(\bar{x}, 0))$. By Lemma 1, the inverse of the Jacobian of the vector-valued function $(\nabla_y L(x, y, \varepsilon, u), u_1 g_1(x, y, \varepsilon), \cdots, u_m g_m(x, y, \varepsilon))$ with respect to (y, u) exists. Hence, the assumptions of the implicit function theorem with respect to (3) are satisfied and we can conclude that in a neighborhood of $\varepsilon = 0$, there exists a unique continuously differentiable function $(y(\bar{x}, \varepsilon), u(\bar{x}, \varepsilon))$ satisfying (3). This implies that for any ε near 0, $y(\bar{x}, \varepsilon)$ is a Kuhn–Tucker point of $P(x, \varepsilon)$ with associated Lagrange multipliers $u(\bar{x}, \varepsilon)$.

The gradient of $(y(\bar{x}, \varepsilon), u(\bar{x}, \varepsilon))$ with respect to ε at $\varepsilon = 0$ is $M(\bar{x})^{-1} N(\bar{x})$. So the conclusion of this lemma holds.

■

Lemma 3

$F(x, y, \varepsilon)$ is uniformly continuous on $X \times Y \times N_0(\varepsilon)$ and $M(x)^{-1} N(x)$ is uniformly bounded on X, where $N_0(\varepsilon)$ is a neighborhood of $\varepsilon = 0$.

Proof

It is not difficult to show $M(x)$ and $N(x)$ are continuous on X. Since $M(x)$ is nonsingular for all $x \in X$, $M(x)^{-1} N(x)$ is continuous on X. Hence, we can get this result from the properties that continuous functions are uniformly continuous and uniformly bounded on compact sets.

■

Let (x^*, y^*) be the unique optimal solution of problem $(BLP(0))$. Then, we can prove the following main result.

Theorem 1

Suppose Assumption (i)–(vi) are satisfied. Then for any given positive number ν, there exists a δ such that when $\|\varepsilon\| < \delta$,

$$|F(x^*(\varepsilon), y^*(\varepsilon), \varepsilon) - F(x^*, y^*, 0)| < \nu.$$

Proof

Let $(x^*(\varepsilon), y^*(\varepsilon))$ be the optimal solution of $(BLP(\varepsilon))$, then $y^*(\varepsilon) = y(x^*(\varepsilon), \varepsilon)$. Denote $\sigma_1 = |F(x^*(\varepsilon), y^*(\varepsilon), \varepsilon) - F(x^*(\varepsilon), y(x^*(\varepsilon), 0), 0)|$ and $\sigma_2 = |F(x^*, y(x^*, \varepsilon), \varepsilon) - F(x^*, y^*, 0)|$. Since

$$F(x^*(\varepsilon), y(x^*(\varepsilon), 0), 0) \geq F(x^*, y^*, 0)$$

and

$$F(x^*,y(x^*,\varepsilon),\varepsilon) \geq F(x^*(\varepsilon),y^*(\varepsilon),\varepsilon),$$

$$|F(x^*(\varepsilon),y^*(\varepsilon),\varepsilon) - F(x^*,y^*,0)| \leq \max\{\sigma_1,\sigma_2\}.$$

From Lemma 2, for $x^*(\varepsilon)$ and x^*, we have the following first-order approximations in a neighborhood of $\varepsilon = 0$,

$$\begin{bmatrix} y(x^*(\varepsilon),\varepsilon) \\ u(x^*(\varepsilon),\varepsilon) \end{bmatrix} = \begin{bmatrix} y(x^*(\varepsilon),0) \\ u(x^*(\varepsilon),0) \end{bmatrix} + M(x^*(\varepsilon))^{-1}N(x^*(\varepsilon))\varepsilon + o(||\varepsilon||) \tag{4}$$

$$\begin{bmatrix} y(x^*,\varepsilon) \\ u(x^*,\varepsilon) \end{bmatrix} = \begin{bmatrix} y(x^*,0) \\ u(x^*,0) \end{bmatrix} + M(x^*)^{-1}N(x^*)\varepsilon + o(||\varepsilon||) \tag{5}$$

Because of the uniformly continuity of $F(x,y,\varepsilon)$, for any given positive number v, there exists a δ_1 such that when $|(y^*(\varepsilon),\varepsilon) - (y(x^*(\varepsilon),0),0)| < \delta_1$, we have $\sigma_1 < v$. By the uniformly boundedness of $M(x)^{-1}N(x)$ and (4), there exists a δ_2 such that when $||\varepsilon|| < \delta_2$, $|(y^*(\varepsilon),\varepsilon) - (y(x^*(\varepsilon),0),0)| < \delta_1$ holds. With an almost same analysis, we can find a δ_3 such that when $||\varepsilon|| < \delta_3, \sigma_2 < v$. Let $\delta = \min\{\delta_2,\delta_3\}$. We can conclude that for any given v, there exist a δ such that when $||\varepsilon|| < \delta$, $|F(x^*(\varepsilon),y^*(\varepsilon),\varepsilon) - F(x^*,y^*,0)| < v$.

■

Now we give an example to illustrate the above theorem.

Example 1

Consider the following bilevel programming problem ($P(\varepsilon)$):

$$\underset{x\in X}{\text{minimize}}\, x^2 + 2\varepsilon_1 y \text{ where } y \text{ solves}$$

$$\underset{y\in Y}{\text{minimize}}\, y$$

$$\text{subject to } y - x \geq \varepsilon_2$$

where $x \in R^1$ and $y \in R^1$ are controlled by the leader and the follower respectively, $X = \{x \in R^1 \mid |x| \leq M\}$ and $Y = \{y \in R^1 \mid |y| \leq 2M\}$, M is a given positive number and ε is a parameter vector in R^2 satisfying $|\varepsilon_2| < M$.

For any given $x \in X$, the unique optimal solution of the inner problem

$$\underset{y\in Y}{\text{minimize}}\, y$$

$$\text{subject to } y - x \geq \varepsilon_2$$

is $y^*(x,\varepsilon) = x+\varepsilon_2$. So the problem $(P(\varepsilon))$ can be reformulated as

$$\underset{x\in X}{\text{minimize}}\, x^2 + 2\varepsilon_1(x+\varepsilon_2).$$

It can be easily shown that the optimal solution of this minimization problem is $x^*(\varepsilon) = -\varepsilon_1$ and the optimal objective value $F(x^*(\varepsilon),y^*(\varepsilon)) = 2\varepsilon_1\varepsilon_2 - \varepsilon_1^2$. When $\varepsilon_1 = \varepsilon_2 = 0$, $(P(\varepsilon))$ is reduced to the following bilevel programming problem $(P(0))$:

$$\underset{x\in X}{\text{minimize}}\, x^2$$

$$\underset{y\in Y}{\text{minimize}}\, y$$

$$\text{subject to } y - x \geq 0.$$

It is obvious that the optimal objective value of this problem is $F(x^*,y^*,0) = 0$.

It is not hard to verify that Assumptions (i) - (vi) are satisfied. By Theorem 1, for any given positive number v, there exists a δ such that when $||\varepsilon|| < \delta$,

$$|F(x^*(\varepsilon),y^*(\varepsilon),\varepsilon) - F(x^*,y^*,0)| < v.$$

In fact, if we choose $\delta = \sqrt{\frac{v}{3}}$, then

$$|F(x^*(\varepsilon),y^*(\varepsilon),\varepsilon) - F(x^*,y^*,0)| = |2\varepsilon_1\varepsilon_2 - \varepsilon_1^2|$$

$$\leq |\varepsilon_1|^2 + 2|\varepsilon_1\varepsilon_2| \leq ||\varepsilon||^2 + 2||\varepsilon||^2 = 3||\varepsilon||^2 < v.$$

REFERENCES

[1] **Bard, J.F.** (1984). «Optimality conditions for the bilevel programming problem», *Naval Research Logistics Quarterly*, **31**, 13–26.

[2] **Ben–Ayed, O.** (1993). «Bilevel linear programming», *Computers and Operations Research*, **20**, 485–501.

[3] **Bialas, W.F. and Karwan, M.H.** (1984). «Two–level linear programming», *Management Science*, **30**, 1004–1020.

[4] **Chen, Y. and Florian, M.** (1995). «The nonlinear bilevel programming problem: formulations, regularity and optimality conditions», *Optimization*, **32**, 193–209.

[5] **Fiacco, A.V.** (1983). *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, New York.

[6] **Wang, Q. and Wang, S.Y.** (1994). «Bilevel programs with multiple potential reactions», *Journal of Systems Science and Systems Engineering*, **3**, 269–278.

[7] **Wang, S.Y. and Lootsma, F.A.** (1994). «A hierarchical optimization model of resource allocation», *Optimization*, **28**, 351–361.

[8] **Wang, S.Y., Wang, Q. and Romano–Rodríguez, S.** (1994). «Optimality conditions and an algorithm for linear–quadratic bilevel programs», *Optimization*, **31**, 127–139.