Journal of Mathematical Research & Exposition
Vol. 19, No. 2, 359- 366, May 1999

# A Note on Constrained Qualification for Bilevel Programming*

*Yang Qingzhi*
(Graduate School, University of Science and Technology of China, Beijing 100039)

**Abstract**: In this paper we research the constrained qualification for Bilevel programming. We show that the usual constrained qualifications in nonlinear programming fail to hold for more general Bilevel Program, and then we give a sufficient condition of "partial calmness" which is weak constrained qualification and can be satisfied for some Bilevel Programs.

**Keywords**: Bilevel programming, Kuhn-Tucker condition, constrained qualification, partial calmness

**Classification**: AMS(1991) 90C30/CLC O221.2

**Document code**: A **Article ID**: 1000-341X(1999)02-0359-08

## 1. Introduction

Bilevel Programming is a new branch of mathematical programming and has been developing since 1980's early. Due to its role is becoming distinquish in many applied mathematical areas, so it has been receiving great attention. Bilevel Programming Problem may be formulated as follows:

(BLPP) $$\min F(x, y),$$

where y solves the following problem

(LLP(x)) $$\min_y f(x, y)$$

$$\text{s.t.} \quad g_i(x, y) \leq 0, \quad i = 1, \cdots, s,$$
$$h_i(x, y) = 0, \quad j = 1, \cdots, p,$$

where $x \in R^n, y \in R^m$. $F(x, y)$, $f(x, y)$, $g_i(x, y)$ and $h_i(x, y)$ are all continuously differentiable functions defined on $R^{n+m}$, $s$, $p$ are nonnegtive integers.

## 2 On the constrained qualification

* **Received date**: 1996-01-03
**Biography**: Yang Qingzhi (1963- ), male, born in Anhui province. Ph.D., currenty an associate professor at Graduate School, University of Science and Technology of China.

It is an important research way to convert (BLPP) into common nonlinear programming problem and then to make use of ripe results in theory as well as algorithms in nonlinear programs.

Denote

$$\Omega = \{(x, y) \in R^{n+m}: g_i(x, y) \leq 0, i = 1, \cdots, s; h_j(x, y) = 0, j = 1, \cdots, p\},$$

$$\Omega(x) = \{y: (x, y) \in \Omega\},$$

$$Q(x) = \{y \in \Omega(x): y \text{ is optimal to (LLP(x))}\},$$

$$IR = \{(x, y) \in \Omega: y \in Q(x)\},$$

where induced region IR is feasible area of (BLPP) obviously.

Denote the value function of (LLP(x)) by V(x):

$$V(x) = \inf_y \{f(x, y): y \in \Omega(x)\}.$$

Then (BLPP) may be written into following equivalent nonlinear programming problem:

$$\text{(SLPP)} \qquad \min F(x, y)$$

$$\text{s. t.} \quad f(x, y) - V(x) = 0,$$

$$g_i(x, y) \leq 0, \quad i = 1, \cdots, s,$$

$$h_i(x, y) = 0, \quad j = 1, \cdots, p.$$

One of most important results in nonlinear programming theory is that the local optimum satisfies Kuhn-Tucker condition under certain constrained qualifications. Mangasarian-Fromovitz condition, linear independence condition and Slater condition are most popular constrained qualifications. Naturally we expect that these constrained qualifications still hold for (SLPP) under some assumptions. If this is true, then the corresponding Kuhn-Tucker condition may be obtained. However in [1] an unexpected result is showed: in a relatively mild case, abovementioned constrained qualifications fail to hold. To deal with this problem, a weaker constrained qualification originallly presented by Clarke for nonlinear program (see [4]), "partial calmness" was introduced to Bilevel Program by J. J. Ye etc. (see [1]). And it is showed that Kuhn-Tucker condition holds under this weaker constrained qualification. Moreover J. J. Ye et. showed that for some special situations, this weak constrained qualification is met. So it is suitable for (BLPP).

In the next we first illustrate that usual constrained qualifications in nonlinear programs fail to hold for (SLPP) in more general situation as well. Therefore this further demonstrates the essential difference between usual one-level nonlinear program and Bilevel Program. Then we give a sufficient condition which ensures the weaker constrained qualification "partial calmness" to hold.

Consider the lower level problem (LLP(x)) of Bilevel Program.

Denote

$$M_x^1(y) = \{(v,\pi) \in R^{p+s}: \nabla_y f(x,y) + \nabla_y h(x,y)^T v + \nabla_y g(x,y)^T \pi = 0,$$
$$\pi \geq 0,\ \pi^T g(x,y) = 0\},$$
$$M_x^0(y) = \{(v,\pi) \in R^{p+s}: \nabla_y h(x,y)^T v + \nabla_y g(x,y)^T \pi = 0,$$
$$\pi \geq 0, \pi^T g(x,y) = 0\}.$$

According to [1], $M_x^1(y)$ and $M_x^0(y)$ are termed as normal multiplier set and abnormal multiplier set at feasible point y, respectively.

Define $M_x^0 Q(x) = \bigcup_{y \in Q(x)} M_x^0(y)$. By [1] one has following results

**Lemma 2 1** ([1] Proposition 2 1) *Assume that $Q(x)$ is nonempty. If $M_x^0 Q(x) = \{0\}$, then value function $V(x)$ is Lipschitzian continuous near $x$, and one has following estimate for the Clarke generalized gradient*

$$\partial V(x) \subset \mathrm{co}\{\nabla_x f(x,y) + \nabla_x h(x,y)^T v + \nabla_x g(x,y)^T \pi: y \in Q(x), (v,\pi) \in M_x^1(y)\}. \tag{1}$$

**Theorem 2 2** ([1] Proposition 3 2) *Let $(x,y)$ be a solution of (BLPP). Assume that $M_x^0 Q(x) = \{0\}$ and equality holds in (1). Then there exists a nontrivial abnormal multiplier for (SLPP), i e, the set of abnormal multiplier corresponding to $(x,y)$ for (SLPP) contains a nonzero element.*

Above theorem implies that the usual constrained qualifications fail to hold under assumptions of theorem. The condition $M_x^0 Q(x) = \{0\}$ in Theorem 2 is to guarantee the continuity of V(x). The major assumption condition is that equality holds in (1). Naturally one may present questions:

1 When does equality hold in (1)?

2 Whether or not Theorem 2 2 still holds if the containing relation in (1) is strictly met, i e, equality doesn't hold in (1)?

For problem 1, there are following result.

**Lemma 2 3** ([2] Corollary 5 4) *Let $\Omega(x) = \{y \in R^m: (x,y) \in \Omega\}$ be uniformly compact near $x$, namely, there exists a bounded open set $B$ such that $\Omega(\bar{x}) \subset B$, for all $\bar{x} \in \cup(x)$, where $\cup(x)$ is a small neighborhood of $x$. Asssume that for any $y \in Q(x)$, $\nabla_y h_j(x,y)$ $(j = 1, \cdots, p)$, $\nabla_y g_i(x,y)$ $(i \in I(x,y))$ are linearly independent, then equality holds in (1), where $I(x,y) = \{(x,y) | g_i(x,y), i = 1, \cdots, s\}$*

It is not difficult to illustrate that equality in (1) generally does not hold provided the linear independence condition is removed.

For problem 2, we obtain following results which don't require that the equality holds in (1).

**Theorem 2 4** *Let $(x,y)$ be the solution of (BLPP), $Q(x)$ is a singleton set $\{y\}$ and $M_x^0 Q(x) = M_x^0(y) = \{0\}$. Then there exists nontrivial abnormal multiplier at $(x,y)$ for (SLPP).*

**Proof** Since $M_x^0(y)=\{0\}$, then $M_x^1(y)$ is a bounded closed convex set by [2]. So

$$\mathrm{co}\{\nabla_x f(x,y)+\nabla_x h(x,y)^T v+\nabla_x g(x,y)^T\pi:(v,\pi)\in M_x^1(y)\}$$
$$=\{\nabla_x f(x,y)+\nabla_x h(x,y)^T v+\nabla_x g(x,y)^T\pi:(v,\pi)\in M_x^1(y)\}.$$

By Lemma 2.1 one knows

$$\partial V(x)\subset\{\nabla_x f(x,y)+\nabla_x h(x,y)^T+\nabla_x(x,y)^T\pi:(v,\pi)\in M_x^1(x,y)\}.$$

Therefore for any $\xi\in\partial V(x)$, there exists $(\overline{v},\overline{\pi})\in M_x^1(y)$ such that

$$\xi=\nabla_x f(x,y)+\nabla_x h(x,y)^T\overline{v}+\nabla_x g(x,y)^T\overline{\pi}.$$

Due to $(\overline{v},\overline{\pi})\in M_x^1(y)$, so

$$0=\nabla_y f(x,y)+\nabla_y h(x,y)^T\overline{v}+\nabla_y g(x,y)^T\overline{\pi},\ \overline{\pi}\geq 0,\ \overline{\pi}^T g(x,y)=0.$$

Hence we deduce

$$\nabla f(x,y)-\xi\times\{0\}+\nabla h(x,y)^T\overline{v}+\nabla g(x,y)^T\overline{\pi}=0,$$

i.e.,

$$0\in\nabla f(x,y)-\partial V(x)\times\{0\}+\nabla h(x,y)^T\overline{v}+\nabla g(x,y)^T\overline{\pi}.$$

Obviously $(1,\overline{v},\overline{\pi})$ is a nontrivial abnormal multiplier at (x, y) for (SLPP). The proof is complete.

**Theorem 2.5** *Let $(x,y)$ be a solution of (BLPP). Suppose that $f(x,y)$, $g_i(x,y)$ $(i=1,\cdots,s)$ are all convex functions, $h_j(x,y)$ $(j=1,\cdots,p)$ are affine functions, $M_x^1Q(x)=\{0\}$. Then there exists nontrivial abnormal multiplier at $(x,y)$ for (SLPP).*

**Proof** The value function of (LLP(x)) is

$$V(x)=\min_y\{f(x,y):g_i(x,y)\leq 0,\ i=1,\cdots,s;\ h_j(x,y)=0,\ j=1,\cdots,p\}$$
$$=\min_{\alpha,y}\{f(\alpha,y):g_i(\alpha,y)\leq 0,\ i=1,\cdots,s;\ h_j(\alpha,y)=0,\ j=1,\cdots,p;\ x-\alpha=0\}.$$

Set $l_x(\alpha,y)\overset{\Delta}{=}x-\alpha$, then

$$V(x)=\min_{\alpha,y}\{f(\alpha,y):g_i(\alpha,y)\leq 0,\ i=1,\cdots,s;\ h_j(\alpha,y)=0,\ j=1,\cdots,p;\ l_x(\alpha,y)=0\}.$$

Define perturbed value function

$$\Phi(p,q_1,q_2)=\min_{\alpha,y}\{f(\alpha,y):g(\alpha,y)+p\leq 0,\ h(\alpha,y)+q_1=0,;\ l_x(\alpha,y)+q_2=0\}.$$

where $p,q_1,q_2$ are vector parameters.

Evidently $\Phi(0,0,0)=V(x)$.

Denote

$$\begin{aligned}\Lambda_1(x,y) &= \{(\lambda,\eta_1,\eta_2): \nabla f(x,y) + \nabla g(x,y)^T\lambda + \nabla h(x,y)^T\eta_1 + \nabla l_x(x,y)^T\eta_2 = 0,\\ &\quad \lambda \geq 0, \lambda^T g(x,y) = 0\}\\ &= \{(\lambda,\eta_1,\eta_2): (\lambda,\eta_1) \in M_x^1(y), \eta_2\\ &= \nabla_x f(x,y) + \nabla_x g(x,y)^T\lambda + \nabla_x h(x,y)^T\eta_1\}.\end{aligned}$$

Because $M_x^0(y) = \{0\}$, then $M_x^1(y)$ is a bounded compact convex set from [2]. So from preceding formula we conclude that $\Lambda_1(x,y)$ is a compact set. By Theorem 2.1 of [3] we obtain

$$\partial\Phi(0,0,0) \cap \Lambda_1(x,y) \neq \emptyset, \quad \forall y \in Q(x).$$

Moreover we know from Proposition 2.3.5 of [4] that if $\Phi(p,q_1,q_2)$ is convex in variable $q_2$, then for any element $(\lambda,\eta_1,\eta_2)$ in $\partial\Phi(0,0,0)$, one has

$$\eta_2 \in \partial_{q_2}\Phi(0,0,0) = \partial V(x).$$

Therefore it follows

$$\partial V(x) \cap \{\nabla_x f(x,y) + \nabla_x g(x,y)^T\lambda + \nabla_x h(x,y)^T\eta_1: (\lambda,\eta_1) \in M_x^1(y)\} \neq \emptyset$$

That is that there are $\xi \in \partial V(x)$ and $(\lambda,\eta_1) \in M_x^1(y)$ such that

$$\xi = \nabla_x f(x,y) + \nabla_x g(x,y)^T\lambda + \nabla_x h(x,y)^T\eta_1,$$

$$0 = \nabla_y f(x,y) + \nabla_y g(x,y)^T\lambda + \nabla_y h(x,y)^T\eta_1$$

$$\lambda \geq 0, \quad \lambda^T g(x,y) = 0$$

which is equivalent to

$$\nabla f(x,y) - \xi \times \{0\} + \nabla g(x,y)^T\lambda + \nabla h(x,y)^T\eta_1 = 0.$$

Therefore

$$0 \in \nabla f(x,y) - \partial V(x) \times \{0\} + \nabla g(x,y)^T\lambda + \nabla h(x,y)^T\eta_1.$$

One easily sees that $(1,\lambda,\eta_1)$ is a nontrivial abnormal multiplier.

Below it suffices to show that $\Phi(p,q_1,q_2)$ is convex in $q_2$.

Let $0<\lambda<1$. For $q_2'$ and $q_2''$, there exist corresponding $y_1$ and $y_2$ such that

$$\lambda\Phi(p,q_1,q_2') + (1-\lambda)\Phi(p,q_1,q_2'') = \lambda f(x+q_2',y_1) + (1-\lambda)f(x+q_2'',y_2).$$

where $y_1, y_2$ meet further

$$g(\alpha,y_1) + p \leq 0, \quad g(\alpha,y_2) + p \leq 0,$$

$$h(\alpha,y_1) + q_1 = 0, \quad h(\alpha,y_2) + q_1 = 0,$$

$$l_x(\alpha,y_1) + q_2' = 0, \quad l_x(\alpha,y_2) + q_2'' = 0.$$

By the convexity of $g(x, y)$ and linearity of $h(x, y)$ and $l_x(x, y)$ in y, we may deduce

$$g(\alpha, \lambda y_1 + (1-\lambda)y_2) + p \leq 0, h(\alpha, \lambda y_1 + (1-\lambda)y_2) + q_1 = 0,$$
$$l_x(\alpha, \lambda y_1 + (1-\lambda)y_2) + \lambda q_2' + (1-\lambda) q_2'' = 0,$$

Notice the convexity of f(x, y) one has

$$\begin{aligned}\lambda\Phi(p, q_1, q_2') + (1-\lambda)\Phi(p, q_1, q_2'') &\geq f(x + \lambda q_2' + (1-\lambda) q_2'', \lambda y_1 + (1-\lambda)y_2)\\ &\geq \min_{\alpha, y}\{f(\alpha, y): g(\alpha, y) + p \leq 0, h(\alpha, y) + q_1 = 0, l_x(\alpha, y) + \lambda q_2' + (1-\lambda) q_2'' = 0\}\\ &= \Phi(p, q_1, \lambda q_2' + (1-\lambda) q_2'').\end{aligned}$$

That is that $\Phi(p, q_1, q_2)$ is convex in $q_2$.

This completes the proof.

## 3. A sufficient condition of "partial calmness"

Due to the usual constrained qualifications fail to hold for Bilevel Program. Therefore in [1] a weak constrained qualification which first proposed by Clarke ([4]) for nonlinear program, "calmness", is introduced to Bilevel Program, and a new weaker constrained qualification "partial calmness" is defined. It proved to be suitable to Bilevel Program. In [1] it was shown that for Bilevel Program with linear lower level and a kind of special Bilevel Program --minimax prolem, "partial calmness" holds. Then Kuhn-Tucker condition may be obtained.

In this section, for general Bilevel Program a sufficient condition of "partial calmness" is presented by using a result of Rockafellar ([5]).

In the sequal assume that $F(x, y)$, $f(x, y)$, $g_i(x, y)$ and $h_i(x, y)$ are all twice continuously differentiable. Furthermore we assume that value function $V(x)$ is also twice continuously differentiable. The related conditions referres to [6].

Now consider following partially perturbed programming problem

$(SLPP_u)$ $\qquad \min F(x, y)$

$$\begin{aligned}\text{s.t.}\quad & f(x, y) - V(x) + u = 0,\\ & g_i(x, y) \leq 0, \quad i = 1, \cdots, s,\\ & h_i(x, y) = 0, \quad j = 1, \cdots, p,\end{aligned}$$

where $u \in R^1$ is perturbed parameter.

**Definition 3.1** *Let $(x, y)$ be a solution of $(SLPP)$. Then $(SLPP)$ is called "partial calmness" at $(x, y)$ provided there exist $\delta > 0$ and $\mu > 0$ such that for any $u \in (-\delta, \delta)$, one has that $F(x', y') - F(x, y) + \mu |u| \geq 0$, for any $(x', y') \in (x, y) + \delta B$ holds and $(x', y')$ is feasible to $(SLPP_u)$.*

For problem $(SLPP_u)$, define

$$l(u, x, y, \gamma) = F(x, y) + \delta^T g(x, y) + \eta^T h(x, y) + \eta_0 (f(x, y) - V(x) + u),$$

where $\delta, \eta$ are $s$-dimension and $p$-dimension vectors respectively, $\eta_0$ is real scalar, $\gamma = (\eta_0, \eta, \delta)$

We define following notations further

$$Y^1(x,y) = \{\gamma: \nabla_{x,y} l(u,x,y,\gamma) = 0, \delta \geq 0, \delta^T g(x,y) = 0\},$$
$$l_0(u,x,y,\gamma) = \delta^T g(x,y) + \eta^T h(x,y) + \eta_0 (f(x,y) - V(x) + u),$$
$$Y_1^0(x,y) = \{\gamma: \nabla_{x,y} l_0(u,x,y,\gamma) = 0, \delta \geq 0, \delta^T g(x,y) = 0\}.$$

**Definition 3.2** *Let $\{M_j\}_1^\infty$ be a family of sets in $R^d$, then $M$ is called the limited vallue of $\{M_j\}_1^\infty$ provided $\lim_{j\to} \operatorname{dist}(M_j, z) = \operatorname{dist}(M, z)$, $\forall z \in R^d$, where $dist(A, z) = \inf_{a\in A}\{\|a-z\|_2\}$.*

If $(x_i, y_i) \to (x,y)$, define

$$M_j = \{w : \nabla g_j(x_i,y_i)^T w = 0, j \in I(x,y), \nabla h_j(x_i,y_i)^T w = 0, j = 1,\cdots,p,$$
$$[\nabla (f(x_i,y_i) - V(x_i))]^T w = 0\},$$
$$\mu(x,y) = \{M : M \text{ is the limit of } \{M_j\}\},$$
$$\widetilde{Y}_0^2(x,y) = \{\gamma \in Y_0^1(x,y): \text{there exists } M \in \mu(x.\ y) \text{ such that}$$
$$w^T \nabla^2_{(x,y)} l_0(u,x,y,\gamma) w \geq 0, \forall w \in M\}.$$

Denote the value function of $(SLPP_u)$ by $p(u)$. The lower Hadamard directional derivative of $p(u)$ at $u$ is defined as

$$p_+(u; d) = \liminf_{\substack{d'\to d \\ t\downarrow 0}} \frac{p(u + td') - p(u)}{t}.$$

After making preceding preparations, let $X(u)$ be the solution set of (SLPP$_u$), then one has

**Theorem** 3.3([5] Theorem 6) *Let $u$ be a real scalar such that $(SLPP_u)$ possesses feasible solution, any each solution $(x,y)$ of $(SLPP_u)$ satisfies $\widetilde{Y}_0^2(x,y) = \{0\}$. Then for any given nonzero real number $d$, $p_+(u; d)$ is finite.*

By utilizing this theorem, the major result of this section may be obtained.

**Theorem 3.4** *Let $(x^*, y^*)$ be optimal to $(BLPP)$, and for each optimum $(\overline{x}, \overline{y})$ of $(SLPP)$, $\widetilde{Y}_0^2(\overline{x},\overline{y}) = \{0\}$ holds. Then $(SLPP)$ is "partial calmness" at $(x^*, y^*)$.*

**Proof** By assumptions and Theorem 3.3 it follows that $p_+(0; d)$ is finite. If this proposition is false, i.e., (SLPP) is not "partial calmness" at $(x,y)$, by the definition one has that there exists a sequence $u_i \to 0$ such that

$$\lim_{i\to\infty} \frac{F(x_i', y_i') - F(x^*, y^*)}{|u_i|} = -\infty,$$

where $(x_i', y_i')$ is feasible solution of (SLPP$_{u_i}$).

Because $p(0) = F(x,y)$, $p(u_i) \leq F(x_i', y_i')$, so

$$\lim_{u_i\to 0} \frac{p(u_i) - p(0)}{|u_i|} = -\infty.$$

Without loss of generality, assume that $d_i = u_i / |u_i| \to s$, $|s| = 1$, then

$$-\infty = \lim_{u_i \to 0} \frac{p(u_i) - p(0)}{|u_i|} = \lim_{u_i \to 0} \frac{p\left(|u_i| \frac{u_i}{|u_i|}\right) - p(0)}{|u_i|}$$

$$= \lim_{\substack{d_i \to s \\ |u_i| \to 0}} \frac{p(|u_i| d_i) - p(0)}{|u_i|}$$

$$\geq \liminf_{\substack{d_i \to s \\ t \downarrow 0}} \frac{p(th_i) - p(0)}{t} = p_+(0; s) > -\infty.$$

This is a contradiction. It implies that (SLPP) is "partial calmness" at $(x^*, y^*)$.

The concept "partial calmness" is a suited constrained qualification to obtain Kuhn-Tucker condition for (BLPP). However its definition depends on some unknown informations. Therefore whether or not there exist better and simpler constrained qualifications for (BLPP), this is worth further research.

## References


[1] Ye J J and Zhu D L. *Optimality conditions for bilevel programming problems* [J]. Optimization, 1995, **33**: 9-27.

[2] Gauvin J and Dubeau F. *Differential properties of the marginal function in mathematical programing* [J]. Mathematical Programming Study, 1982, **19**: 101-119.

[3] Gollan B. *Inner estimates for the generalized gradient of the optimal value function in nonlinear programming* [J]. Mathematical Programming Study, 1984, **22**: 132-146.

[4] Clarke H. *Optimization and Nonsmooth Analysis* [M]. Wiley and Sons, New York, 1983.

[5] Rockafellar R. *Directional differentiability of the optimal value function in a nonlinear programming* [J]. Mathematical Programming Study, 1984, **21**: 213-226.

[6] Shapiro A. *Second-order derivatives of extremal-value functions and optimality condition for semi-infinite program* [J]. Mathematical Operation Researchs, 1985, **10**: 207-219.

[7] Dempe S. *A necessary and a sufficient optimality condition for bilevel programming programs* [J]. Optimization, 1992, **25**: 341-354.

[8] Outrata J. *Necessary optimality conditions for Stackelberg problems* [J]. Journal of Optimizaiton Theory and Applications, 1993, **2**(76): 305-320.


# 关于双层规划约束规格的一个注记


杨 庆 之

(中国科学技术大学研究生院, 北京100080)


## 摘 要


本文表明了非线性规划中常见的约束规格对一般双层规划不成立, 并对双层规划可以满足的较弱的约束规格"部分平静", 给出了使其成立的充分条件.