MATHEMATICAL CONTROL THEORY
BANACH CENTER PUBLICATIONS, VOLUME 14
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

A METHOD OF REDUCTION OF CONSTRAINTS

S. ROLEWICZ

Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland

Let X, Y_1, Y_2, Z be real Banach spaces. We assume that Y_1, Y_2 are ordered by convex cones K_1, K_2. Let D be an open set in X. Let f be a real-valued function defined on D. Let G_1, G_2, H map D into Y_1, Y_2, Z; $G_1\colon D \to Y_1$, $G_2\colon D \to Y_2$, $H\colon D \to Z$. We assume that f, G_1, G_2, H are continuously Fréchet differentiable.

We consider the following problem:

$$f(x) \to \inf, \quad G_1(x) \leqslant 0, \quad G_2(x) \leqslant 0, \quad H(x) = 0. \tag{1}$$

We form the Lagrangian

$$L(x, \varphi_1, \varphi_2, \psi) = f(x) + \varphi_1\big(G_1(x)\big) + \varphi_2\big(G_2(x)\big) + \psi\big(H(x)\big). \tag{2}$$

Let $x_0 \in D$. Let $G_1(x_0) = G_2(x_0) = 0$. We assume that at x_0 the necessary optimality condition of the Kuhn–Tucker type holds. It means that there are linear continuous functionals $\varphi_1 \in Y_1^*$, $\varphi_2 \in Y_2^*$, $\psi \in Z^*$, $\varphi_1 \geqslant 0$, $\varphi_2 \geqslant 0$ such that the Fréchet differential of the Lagrangian is equal to 0:

$$d\big(L(x, \varphi_1, \varphi_2, \psi), \cdot\big) = 0 \tag{3}$$

(see [1]).

The main result of the paper is a proof of the fact that under certain conditions problem (1) can be reduced locally at x_0 to another problem with a smaller admissible set, namely to the problem

$$f(x) \to \inf, \quad G_1(x) = 0, \quad G_2(x) \leqslant 0, \quad H(x) = 0. \tag{4}$$

This result permits us to reduce problems concerning sufficient conditions and necessary conditions of higher order for problem (1) to similar considerations for problem (2).

THEOREM 1. *Let $X, Y_1, Y_2, Z, f, G_1, G_2, H$ be as above. Assume that at x_0 the necessary optimality condition of the Kuhn–Tucker type holds, i.e., there are $\varphi_1 \in Y_1^*$, $\varphi_2 \in Y_2^*$, $\varphi_1 \geqslant 0$, $\varphi_2 \geqslant 0$, $\psi \in Z^*$ such that we have* (3).

Suppose that φ_1 is uniformly positive, i.e., there is an $L > 0$ such that for $y \geqslant 0$

$$\|y\| \leqslant L\varphi_1(y) \tag{5}$$

(see [2]). Suppose that the differential

$$\big(dG_1(x_0, \cdot), dG_2(x_0, \cdot), dH(x_0, \cdot)\big)$$

is a surjection of X onto $Y_1 \times Y_2 \times Z$.

Then there is a neighbourhood Q of the point x_0 such that for each $x \in Q$ there is an $\hat{x} \in Q$ such that

$$G_1(\hat{x}) = 0, \quad G_2(\hat{x}) = G_2(x), \quad H(\hat{x}) = H(x) = 0 \tag{6}$$

and

$$f(x) \geqslant f(\hat{x}). \tag{7}$$

If, moreover, $G_1(x) \neq 0$, we have

$$f(x) > f(\hat{x}). \tag{7'}$$

Proof. Since $\big(dG_1(x_0, \cdot), dG_2(x_0, \cdot), dH(x_0, \cdot)\big)$ is a surjection, by the Ljusternik theorem [3], there are a constant $C > 0$ and a neighbourhood Q_1 of x_0 such that for all $x \in Q$ there is an x_1 such that (6) holds and

$$\|x - x_1\| \leqslant C\|G_1(x)\|. \tag{8}$$

By the continuity of Fréchet differentials there is a convex neighbourhood $Q_2 \subset Q_1$ of x_0 such that for $x \in Q_2$

$$\Big\|d\big(L(x, \varphi_1, \varphi_2, \psi), \cdot\big)\Big\| \leqslant \frac{1}{2CL}. \tag{9}$$

Observe that (9) implies that for any $x, y \in Q_2$

$$\|L(x, \varphi_1, \varphi_2, \psi) - L(y, \varphi_1, \varphi_2, \psi)\| \leqslant \frac{1}{2CL}\|x - y\|. \tag{10}$$

By the intersection theorem [5] there is a neighbourhood $Q \subset Q_2$ of x_0 such that for any $x \in Q$ there is an $\hat{x} \in Q$ such that (6) and (8) hold. By (4) and (6) and (8) we get

$$\begin{aligned} f(x) - f(\hat{x}) &= -\varphi_1\big(G_1(x)\big) + \big[f(x) + \varphi_1\big(G_1(x)\big) + \varphi_2\big(G_2(x)\big) + \\ &\qquad + \psi\big(H(x)\big)\big] - \big[f(\hat{x}) + \varphi_1\big(G_1(\hat{x})\big) + \varphi_2\big(G_2(\hat{x})\big) + \psi\big(H(\hat{x})\big)\big] \\ &\geqslant -\varphi_1\big(G_1(x)\big) + \tfrac{1}{2}\varphi_1\big(G_1(x)\big) \geqslant -\tfrac{1}{2}\varphi_1\big(G_1(x)\big) \\ &= \tfrac{1}{2}\big|\varphi_1\big(G_1(x)\big)\big|. \end{aligned} \tag{11}$$

Formula (11) gives trivially (7). By (6) we obtain (7'). Taking X, Y_1, Y_2, Z finite-dimensional, we obtain the result of paper [6]. As a consequence of Theorem 1 we obtain

COROLLARY 1. *Suppose that all the hypotheses of Theorem* 1 *are satisfied. Then a point x_0 is a solution of problem* (1) *if and only if it is a solution of problem* (2).

Now we shall give examples showing when the hypotheses of Theorem 1 are satisfied.

EXAMPLE 1. Let X be a real Banach space. Let Y_1, Y_2, Z be real finite-dimensional spaces. It means that $G_1(x)$, $G_2(x)$, $H(x)$ are of the following form:

$$\begin{aligned} G_1(x) &= \big(g_1^1(x), \ldots, g_n^1(x)\big), \\ G_2(x) &= \big(g_1^2(x), \ldots, g_m^2(x)\big), \\ H(x) &= \big(h_1(x), \ldots, h_l(x)\big). \end{aligned} \tag{12}$$

The differential $\big(dG_1(x_0, \cdot), dG_2(x_0, \cdot), dH(x_0, \cdot)\big)$ is a surjection if the gradients

$$\nabla g_1^1(x_0), \ldots, \nabla g_n^1(x_0); \quad \nabla g_1^2(x_0), \ldots, \nabla g_m^2(x_0); \quad \nabla h_1(x_0), \ldots, \nabla h_l(x_0)$$

are linearly independent.

EXAMPLE 2. Let Y_1 be an n-dimensional real space with the standard order. A functional $\varphi(y)$ is uniformly positive if and only if it is of the form

$$\varphi(y) = \sum_{i=1}^{n} \lambda_i y_i \tag{13}$$

where

$$\lambda_i > 0, \quad i = 1, 2, \ldots, n. \tag{14}$$

EXAMPLE 3. Let Y be $L^1(\Omega)$. Let $\varphi(y)$ be of the form

$$\varphi(y) = \int_{\Omega} \varphi(t) y(t)\, dt. \tag{15}$$

Then φ is uniformly positive if and only if

$$\operatorname{essinf} \varphi(t) > 0. \tag{16}$$

EXAMPLE 4. Let $Y = L^p(\Omega)$, $1 < p < 2$, with the standard order. If Y is infinite-dimensional, then there is no uniformly positive functional in Y. This follows trivially from Theorem 1.5 of [2].

The following examples show that the hypotheses of Theorem 1 are essential.

EXAMPLE 5. Let $X = R^3$. Let $Y_1 = R^3$. Let

$$g_1 = x_1 - x_2, \quad g_2 = x_1 + x_2,$$
$$g_3 = x_1 + x_3^2, \quad f = -x_1 - 2x_3^2.$$

Observe that at the point 0 we have the necessary optimality condition. Namely,

$$f'|_{x=x_0} + \tfrac{1}{3}(g_1'|_{x=x_0} + g_2'|_{x=x_0} + g_3'|_{x=x_0}) = 0.$$

Since

$$\{x\colon g_i(x) = 0,\ i = 1, 2, 3\} = \{(0, 0, 0)\},$$
$$\inf\{f(x)\colon g_i(x) = 0,\ i = 1, 2, 3\} = 0.$$

On the other hand, for any neighbourhood Q,

$$\inf\{f(x)\colon x \in Q,\ g_i(x) \leqslant 0,\ i = 1, 2, 3\} < 0.$$

The reason is that

$$(g_1'|_{x=0},\ g_2'|_{x=0},\ g_3'|_{x=0})$$

is not a surjection on R^3.

EXAMPLE 6. Let X, Y_1, g_1, g_2, f be as in Example 5. Let $h(x) := g_3$. In the same way as in Example 5 we obtain a counterexample showing that Theorem 1 does not hold if $(G'|_{x=x_0}, H'|_{x=x_0})$ is not a surjection.

The essentiality of uniform positivity results from the following

EXAMPLE 7. Let $X = R^3$, $g_1 = x_1$,

$$g_2 = -x_2 - x_1^2, \quad f = x_2.$$

Of course $(g_1'|_{x=0}, g_2'|_{x=0})$ is a surjection on R^2 and

$$f'|_{x=0} + 0 \cdot g_1'|_{x=x_0} + 1 \cdot g_2'|_{x=x_0} = 0.$$

Observe that for any neighbourhood Q of zero

$$\inf\{f(x)\colon x \in Q,\ g_i(x) \leqslant 0,\ i = 1, 2\} < 0.$$

On the other hand, $g_1(x) = g_2(x) = 0$ implies that $x = 0$ and

$$\inf\{f(x)\colon g_1(x) = 0,\ g_2(x) = 0\} = 0. \tag{17}$$

Let X, Y_1, Y_2, Z, f, G_1, G_2, H be as above. Suppose that at x_0 the necessary condition of optimality of the Kuhn–Tucker type holds, i.e., there are $\varphi_1 \in Y_1^*$, $\varphi_2 \in Y_2^*$, $\varphi_1 \geqslant 0$, $\varphi_2 \geqslant 0$, $\psi \in Z^*$ such that (3) holds. Let

$$T = \operatorname{Ker} d\big(G_1(x_0), \cdot\big) \cap \{h\colon dG_2(x_0, h) \leqslant 0\} \cap \operatorname{Ker} d\big(H(x_0), \cdot\big). \tag{18}$$

Let $K(x)$, $x \in X$, be a function uniformly continuous on bounded sets and homegeneous of order a, $a > 1$.

THEOREM 2. *Suppose that for all $x \in D \subset X$*

$$\|L(x, \varphi_1, \varphi_2, \psi) - L(x_0, \varphi_1, \varphi_2, \psi)\| \geqslant |K(x - x_0)|, \tag{19}$$

$$\operatorname{sign}[L(x, \varphi_1, \varphi_2, \psi) - L(x_0, \varphi_1, \varphi_2, \psi)] = \operatorname{sign} K(x - x_0).$$

If $K(h) > c > 0$ for $h \in T$, then x_0 is a local solution of problem (1). *If $K(h_0) < 0$ for some $h_0 \in T$, then x_0 is not a local solution of problem* (1).

Proof. Let

$$D_\varepsilon = \{h \in X\colon \operatorname{dist}(h, T) < \varepsilon \|h\|\}. \tag{20}$$

By the uniform continuity of $K(h)$ on bounded sets there is an $\varepsilon > 0$ such that

$$|K(h)| \geqslant \tfrac{1}{2} c \|h\|^a \quad \text{for} \quad h \in D. \tag{21}$$

Observe that there is a $\delta > 0$ such that

$$\{x\colon G_1(x) = 0, G_2(x) \leqslant 0, H(x) = 0\} \cap B(x_0, \delta) \subset D_\varepsilon \cap B(x_0, \delta), \tag{22}$$

where, as usual, $B(x_0, \delta)$ denotes the ball with radius δ and centre at x_0. Inclusion (22) and inequality (19) imply that x_0 is a solution of problem (4). Thus it is also a solution of problem (1).

Suppose now that there is an $h_0 \in T$ such that $K(h_0) < 0$. By the uniform continuity of K we can find an h such that

$$G_1(x_0 + h) = 0, \quad G_2(x_0 + h) \leqslant 0, \quad H(x_0 + h) = 0$$

and

$$K(h) < \tfrac{1}{2} K(h_0). \tag{23}$$

This trivially implies that x_0 is not a solution of problem (1).

Now we shall apply Theorem 2 to mathematical programming.

COROLLARY 2 [4] (see for example (4)). *Let $f, g_1, \ldots, g_n$ be twice continuously differentiable. Suppose that the gradients $\nabla g_1, \ldots, \nabla g_n$ are linearly independent.*

Suppose that there are Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_n \geqslant 0$ such that the Lagrangian

$$L(x, \lambda) = f(x) + \sum \lambda_i g_i(x) \tag{24}$$

has differential equal to 0 at x_0 and the Hessian (the second differential) at x_0 can be estimated as follows:

$$d^2\big(L(x_0, \lambda), h\big) \geqslant c\|h\|^2, \tag{25}$$

where $c > 0$ and $h \in T$. Then x_0 is a local solution of problem (1).

Proof. Polynomial operators are uniformly continuous on bounded sets. Let

$$Y_1 = \operatorname{lin}\{\nabla g_i:\ \lambda_i > 0\},$$
$$Y_2 = \operatorname{lin}\{\nabla g_i:\ \lambda_i = 0\}.$$

Let P_1, P_2 be projections onto Y_1 and Y_2, $P_1+P_2 = I$. We put $G_1 = P_1G$, $G_2 = P_2G$ and we use Theorem 2.

If X is finite-dimensional, we can replace (25) by

$$d^2\big(L(x,\lambda), h\big) > 0 \quad \text{for} \quad h \in T,\ h \neq 0, \tag{25'}$$

using compactness arguments.

In Corollary 2 linear independence of n gradients $\nabla g_1, \ldots, \nabla g_n$ is required. In fact it is not essential. The proof of this fact is based on the following

LEMMA 1. *Let $a_1, \ldots, a_n$ be linearly independent elements in a linear space Y. Let $b \in Y$ and*

$$a_1 + a_2 + \ldots + a_n + b \neq 0. \tag{26}$$

Then there is an index i such that the vectors

$$a_1, \ldots, a_{i-1}, a_i + b, a_{i+1}, \ldots, a_n$$

are linearly independent.

Proof. If $b \notin \operatorname{lin}(a_1, \ldots, a_n)$, we take $i = 1$ and trivially $a_1 + b, a_2, \ldots$ $\ldots, a_n$ are linearly independent.

Suppose that $b \in \operatorname{lin}(a_1, \ldots, a_n)$. Then $b = \beta_1 a_1 + \ldots + \beta_n a_n$. Since (26), there is a $\beta_i \neq -1$. In an obvious way

$$a_1, \ldots, a_{i-1}, a_i + b, a_{i+1}, \ldots, a_n$$

are linearly independent.

Now we can prove

COROLLARY 2′ [4]. *Let X be a Banach space. Let $f, g_1, \ldots, g_n$ be twice continuously differentiable real-valued functions. Let $g_1(x_0) = \ldots = g_n(x_0) = 0$. Suppose that there are Lagrange multipliers $\lambda_1, \ldots, \lambda_n \geqslant 0$ such that the Lagrangian*

$$L(x, \lambda) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) \tag{24}$$

has differential equal to 0 at x_0 and the Hessian at x_0 can be estimated as in (25). *Suppose that $\nabla f \neq 0$.*

Then x_0 is a local solution of problem (1).

Proof. Let $k = \dim\operatorname{lin}\big(\nabla g_1(x_0), \ldots, \nabla g_n(x_0)\big)$. We assume that $\nabla g_1(x_0), \ldots$ $\ldots, \nabla g_k(x_0)$ are ordered in such a way that $\nabla g_1(x_0), \ldots, \nabla g_k(x_0)$ are linearly independent and

$$
\begin{aligned}
&\lambda_i > 0, \quad i = 1, 2, \ldots, l, \quad l \leqslant k,\\
&\lambda_i = 0, \quad i = l+1, \ldots, k,\\
&\lambda_i > 0, \quad i = k+1, \ldots, m,\\
&\lambda_i = 0, \quad i = m+1, \ldots, n.
\end{aligned}
\tag{27}
$$

Using Lemma 1, we may assume without loss of generality that the gradients

$$\nabla g_1(x_0), \ldots, \nabla g_{l-1}(x_0), \sum_{i=l}^{n} \lambda_i \nabla g_i(x_0), \nabla g_{l+1}(x_0), \ldots, \nabla g_k(x_0)$$

are linearly independent.

Now we shall consider the following optimization problem:

$$f(x) \to \inf, \quad \bar{g}_i(x) \leqslant 0, \quad i = 1, 2, \ldots, k, \tag{28}$$

where

$$
\begin{aligned}
\bar{g}_i(x) &:= g_i(x), \quad i \neq l,\\
g_l(x) &= \sum_{i=l}^{n} \lambda_i g_i(x).
\end{aligned}
\tag{29}
$$

Of course the admissible set in problem (28) is bigger than the admissible set in problem (1). Therefore, if x_0 is a solution of problem (28), it is a solution of problem (1). Applying Lemma 2, it is easy to verify that if

$$d^2\big(L(x_0, \lambda), h\big) \geqslant c\|h\|^2, \quad c > 0, \quad \text{for all } h \in T_1, \tag{30}$$

where

$$
\begin{aligned}
T_1 = \bigcap_{i=1}^{l-1} \operatorname{Ker} d\big(g_1(x_0), \cdot\big) \cap \operatorname{Ker}\Big(\sum_{i=l}^{n} \lambda_i d\big(g_i(x_0), \cdot\big)\Big) \cap\\
\cap \{h\colon d\big(g_i(x_0), h\big) \leqslant 0, \ i = l+1, \ldots, k\},
\end{aligned}
\tag{31}
$$

then $T_1 = T$, where T is given by formula (18).

References

[1] K. J. Arrow, L. Hurwicz, H. Uzawa, *Studies in linear and nonlinear programming*, Stanford, Calif. 1958.

[2] M. A. Krasnosielskiĭ, *Positive solutions of operator equations*, Moscow 1962 [in Russian].

[3] L. A. Ljusternik, *On conditional extrema of functionals*, Mat. Sb. 41, 3 (1934), 390–401 [in Russian].

[4] F. Lempio, J. Zowe, *Higher order optimality conditions*, (to appear).

[5] S. Rolewicz, *On intersection of multifunctions*, Math. Operationsforsch. Statist. Ser. Optimization 11 (1980), 3–11.

[6] —, *On sufficient conditions of optimality in mathematical programming*, in: Proc. V Symp. Oper. Research, Köln 1980, Meth. of Oper. Res. 40, 1981, 149–152.