ON SUFFICIENT CONDITIONS FOR OPTIMALITY

S. Rolewicz, Warszawa *)

Let $f, g_1, \ldots, g_m$ be continuously differentiable real valued functions defined on a domain Ω of n-dimensional real space $\mathbb{R}^n$. We consider the following optimization problem.

$$f(x) \to \inf$$

(1)

$$g_i(x) \leq 0 , \quad x \in \Omega .$$

Let $x_0 \in \Omega$. We assume that at x_0 all constraints g_i are active, i.e. $g_i(x_0) = 0$.

THEOREM 1 ([9]): *Suppose that at the point* x_0 *all gradients of* g_i, ∇g_i, *are linearly independent. Suppose that at* x_0 *Kuhn-Tucker necessary conditions for optimality hold, i.e. there are* $\lambda_i \geq 0$ *such that*

(2) $$\nabla (f + \Sigma \lambda_i g_i) \Big|_{x_0} = 0 .$$

If all $\lambda_i > 0$, $i = 1,2,\ldots,m$, *then* x_0 *is a local minimum of problem* (1) *if and only if it is a local minimum of the following equality problem*

$$f(x) \to \inf$$

(3)

$$g_i(x) = 0 .$$

The proof of Theorem 1 is elementary and uses only the implicit functions theorem. Theorem 1 gives a very useful algorithm for reducing a problem of sufficient condition for problem (1) to well-known classical

---

*) This work was partially supported by Monash University, Clayton, Victoria.

problem (3). In this way we can obtain sufficient condition of optimality of order higher than 2. It is important that this algorithm needs only to invert one fixed matrix determined by the gradients.

Now we shall present a simple example

EXAMPLE 1 ([9]). Let

$$g_1(x,y,z) = -(x + y) + z^2$$

$$g_2(x,y,z) = - y + z^4$$

$$f(x,y,z) = x + 2y - x^2 + y^2 - z^2 .$$

It is easy to check that for (0,0,0) Kuhn-Tucker conditions hold for $\lambda_1 = \lambda_2 = 1$.

Using the theorem we can replace problem (1) by problem (3). Thus $y = z^4$, $x = z^2 - z^4$ and $f = x + 2y - x^2 + y^2 - z^2 = 2z^6$. It implies that (0,0,0) is a local minimum of problem (1).

Replacing f by

$$f_\alpha = x + 2y = \alpha x^2 + y^2 - z^2$$

we are able to prove that for $\alpha > 1$ the corresponding problem does not have local minimum, at (0,0,0) . In both cases the corresponding conditions are of the order higher than 2. (6 in the first case 4 in the second one.)

This basic theorem can be extended in the following way.

THEOREM 2 ([9]): *Suppose that at the point* x_0 *all gradients of* g_i , ∇g_i *are linearly independent. Suppose that at* x_o *the Kuhn-Tucker necessary conditions for optimality hold. Suppose that* $\lambda_i > 0$, $i = 1,2,\ldots,p$, $\lambda_i = 0$ *for* $i = p+1,\ldots,m$. *Then* x_0 *is a local minimum of the problem* (1) *if and only if it is a local minimum of the following problem*

$$f(x) \to \inf$$

(4) $$g_i(x) = 0 \qquad i = 1,2,\dots,p$$

$$g_i(x) \le 0 \qquad i = p+1,\dots,m .$$

There is a natural question. The existing second order sufficient conditions [3], [7] (historical discussion of the subject is well presented in [4]) did not request linear independence of the gradients ∇g_i but positiveness of the second differential on the set

(5) $$T = \bigcap_{i=1}^{m} T_i$$

where

$$T_i = \{x : (\nabla g_i, x) = 0\} \quad \text{if} \quad \lambda_i > 0$$

(6)

$$T_i = \{x : (\nabla g_i, x) \le 0\} \quad \text{if} \quad \lambda_i = 0 .$$

Observe that in fact the set T can be described by a linearly independent subset ∇g_{i_j} where $\operatorname{span}(\nabla g_{i_j}) = \operatorname{span}(\nabla g_i)$. Thus basing on Theorem 2 we can obtain the following

THEOREM 3 ([12]): *Let* $f, g_1, \dots, g_m$ *be k-time continuously differentiable functions defined on a domain* $\Omega \subset \mathbb{R}^n$. *Suppose* at $x_0 \in \Omega$ *all constraints are active, i.e.* $g_i(x_0) = 0$, $i = 1,2,\dots,m$. *Suppose that there are* $\lambda_1,\dots,\lambda_m \ge 0$ *such that the differentials of the Lagrangian*

$$L(x) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

are equal to zero till the order $k-1$ *for all* $h \in \mathbb{R}^n$, *i.e.*

$$d^i L(x_0, h) \equiv 0 \qquad i = 1,2,\dots,k-1$$

and that

$$d^k L(x_0, h) > 0 \qquad \text{for} \quad h \in T$$

where T *is defined by* (4). *Thus* x_0 *is a local minimum of problem* (1).

For $k = 2$ the result was known much earlier [3], [7].

The natural question which arises is about extensions of the theorems given above on Banach spaces. It can be done in the following way

THEOREM 4 ([10]): *Let* X, Y_1, Y_2, Z *be Banach spaces. Let* Y_1, Y_2 *be ordered spaces. Let* Ω *be a domain in* X. *Let* F *be a real valued function defined on* Ω.

Let $G_1 : \Omega \to Y_1$, $G_2 : \Omega \to Y_2$, $H : X \to Z$ *be continuously differentiable operators. We consider the following problem.*

(7)
$$\begin{aligned} F(x) &\to \inf \\ G_1(x) &\le 0 \\ G_2(x) &\le 0 \\ H(x) &= 0 . \end{aligned}$$

Suppose that at the point x_0 *all constraints are active*

$$G_1(x_0) = G_2(x_0) = H(x_0) = 0 .$$

Suppose that the differential $\nabla G_1 \times \nabla G_2 \times \nabla H$ *is a surjection of* X *onto* $Y_1 \times Y_2 \times Z$. *Suppose that there are linear functionals*

$$\phi_1 \in Y_1^* , \quad \phi_2 \in Y_2^* , \quad \psi \in Z^*$$

such that the differential of the Lagrangian

(8) $$d(L(x),h) = d(F(x) + \phi_1(F_1(x)) + \phi_2(G_1(x)) + \psi(H(x)):h) = 0 .$$

If ϕ_1 *is uniformly positive, i.e. there is* $C > 0$ *such that*

(9) $$\|y\| \le C\phi_1(y)$$

for all $y \in Y_1$, $y \ge 0$, *then* x_0 *is a local solution of problem* (6) *if and only if it is a local solution of the following equality problem*

$$F(x) \to \inf$$
$$G_1(x) = 0$$
$$G_2(x) \le 0$$
$$H(x) = 0 \; .$$

The hypothesis that ϕ_1 , ϕ_2 , ψ are linear is not essential; it is enough that they are odd.

Theorem 4 can be generalized to Lipschitz functions in the following way.

THEOREM 5 ([10]): *Let* X , Y_1 , Y_2 , Z , Ω *be as in Theorem 4. Let* $G_1 : X \to Y_1$, $G_2 : X \to Y_2$, $H : X \to Z$. *We shall not assume continuity of those operators, but we assume that the multi-function*

$$\Gamma(y_1,y_2,z) = \{x : G_1x = y_1,\ G_2x = y_2,\ Hx = z\}$$

is locally Hausdorf continuous, i.e. for each neighbourhood Q *of* x_0 *there is a neighbourhood* Q_1 *of* x_0 , $Q_1 \subset Q$ *and a neighbourhood* W *of* $(G_1(x_0)\ ,\ G_2(x_0)\ ,\ H(x_0))$ *in the space* $Y_1 \times Y_2 \times Z$ *and a constant* $K > 0$ *such that*

$$d(\Gamma(y_1,y_2,z) \cap Q_1,\ \Gamma(\bar{y}_1,\bar{y}_2,\bar{z}) \cap Q_1)$$
$$\le K\|(y_1,y_2,z) - (\bar{y}_1,\bar{y}_2,\bar{z})\|_1 \; ,$$

where d *denotes the Hausdorf distance of spaces and* $\| \; \|_1$ *denotes an arbitrary norm in* $Y_1 \times Y_2 \times Z$, *coinciding on* Y_1 , *i.e.* $\|(y,0,0)\|_1 = \|y\|_{Y_1}$.

If there are linear continuous functionals $\phi_1 \in Y_k^*$, $\phi_2 \in Y_2^*$, $\psi \in Z^*$ *such that* $\phi_1 \ge 0$, $\phi_2 \ge 0$ *and* ϕ_1 *is uniformly positive (i.e.* (9) *holds) and the function*

$$F(x) + \phi_1(G_1(x)) + \phi_2(G_2(x)) + \psi(H(x))$$

satisfies the Lipschitz condition with constant M , *then* x_0 *is a local* *solution of the problem* (6) *if and only if it a local solution of problem* (7) *provided*

$$(10) \qquad MKC < 1 .$$

Theorem 5 gives Theorem 4 via Ljusternik theorem [5]. Other conditions warranting Γ is locally Lipschitz, can be found in the papers [1], [2], [8].

Theorem 5 presented above can also be used to obtain results of sufficient conditions for Pareto minimization.

THEOREM 6 ([13]): *Let* X, Y_1, Y_2, Z, P *be Banach spaces. Suppose that the spaces* Y_1, Y_2, P *are ordered. Let* U *be a domain in* X *and let* F, G_1, G_2, H *be continuously differentiable operators* $F : U \to P$, $G_1 : U \to Y_1$, $G_2 : U \to Y_2$, $H : U \to Z$. *We are looking for a local Pareto minimum of the following problem*

$$(11) \qquad \begin{aligned} F(x) &\to \inf \\ G_1(x) &\leq 0 \\ G_2(x) &\leq 0 \\ H(x) &= 0 . \end{aligned}$$

Suppose that

(i) *there are continuous linear functionals*

$$\alpha \in P^*, \quad \lambda_1 \in Y_1^*, \quad \lambda_2 \in Y_2^*, \quad \gamma \in Z^*$$

such that

$$\alpha(\nabla F) + \lambda_1(\nabla G_1) + \lambda_2(\nabla G_2) + \gamma(\nabla H) = 0 ,$$

where ∇F, ∇G_1, ∇G_2, ∇H *are the differentials of* F, G_1, G_2, H *taken at the point* x_0 *(this is called a necessary condition of optimality of the Kuhn-Tucker type).*

(ii) *the functionals* α, λ_1, λ_2 *are positive, and* α_1, λ_1 *are* *uniformly positive i.e., there are positive constant* c_α, c_1, c_2 *such that for* $p \geq 0$, $y_1 \geq 0$,

$$\|p\| \leq c_\alpha \, \alpha(p)$$

$$\|y_1\| \leq c_1 \, \lambda_1(Y_1) \, .$$

(iii) *the constraints are active at* x_0 *, i.e.,*

$$G_1(x_0) = 0 \, , \quad G_2(x_0) = 0 \, , \quad H(x_0) = 0 \, ;$$

(iv) F *is a surjection on* P *and* $(\nabla G_1, \nabla G_2, \nabla H)$ *is a surjection on* $Y_1 \times Y_2 \times H$.

(v) *the space* $L_1 = \ker \nabla F$ *and the halfsubspace*

$$L_2 = \ker \nabla G_1 \cap \ker \nabla H \cap \{x : \nabla G_2(x) \leq 0\}$$

have a positive gap d , i.e.;

$$d = \max(\inf\{\|x - y\|, \ x \in L_1, \ y \in L_2, \ \|x\| = 1\}, \ \inf\{\|x - y\|, \ x \in L_1, \ y \in L_2, \ \|y\| = 1\}) > 0 \, .$$

Then x_0 *is a local Pareto minimum of problem* (11).

In the theorem presented above condition (v) is very restrictive. Unfortunately simple examples [14] show that this condition is essential.

REFERENCES

[1] S. Dolecki, 'Semicontinuity in constrained optimization', *Control and Cybernetics* 7 (1978) Ia, No. 2, 5-16, Ib, No. 3, 17-26.

[2] S. Dolecki and S. Rolewicz, 'Exact penalties for local minima', *SIAM Journal of Control and Optim.* 17 (1979), 596-606.

[3] M.R. Hesteness, 'An indirect sufficiency proof for the problem of Bolza in nonparametric form', *Trans. Amer. Math. Soc.* 62 (1947), 509-535.

[4] F. Lempio and J. Zowe, 'Higher order optimality conditions' (preprint of Math. Institute, Univ. Bayreuth).

[5] L.A. Ljusternik, 'On conditional estrema of functionals', (in Russian), *Matem. Sc.* 41 (1934), 390-401.

[6] H. Maurer and J. Zowe, 'First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems', *Math. Programming* 16 (1979), 98-110.

[7] E.J. McShane, 'Sufficient conditions for a weak relative minimum in the problem of Bolza', *Trans. Amer. Math. Soc.* 52 (1942), 344-379.

[8] S. Rolewicz, 'On intersections of multifunctions', *Math. Operforschung und Statistik, Ser. Optimization* 11 (1980), 3-11.

[9] S. Rolewicz, 'On sufficient conditions of optimality in mathematical programming', *Oper. Res. Verf.* 40 (1981), 149-152.

[10] S. Rolewicz, 'On sufficient condition of optimality for Lipschitzian functions', *Proc. Conf. Game Theory and Mathematical Economy*, ed. O. Moeschlin and D. Pallaschke, North-Holland Publishing Co. 1981, 351-355.

[11] S. Rolewicz, 'On sufficient condition of vector optimization', *Oper. Res. Verf. (Methods of Operation Research)* 43 (1982), 151-157.

[12] S. Rolewicz, 'On sufficient conditions of optimality of second order', *Ann. Pol. Math.* 42 (1983), 297-300.

[13] S. Rolewicz, 'On Pareto optimization in Banach spaces', *Stud. Math.* (to appear).

[14] S. Rolewicz, 'Remarks on sufficient conditions of optimality of vector optimization', *Math. Operationsforschung and Statistik, Series Optimization* (in print).