

r -convex transformability in nonlinear programming problems

E. GALEWSKA, M. GALEWSKI

Abstract. We show that for r -convex transformable nonlinear programming problems the Karush-Kuhn-Tucker necessary optimality conditions are also sufficient and we provide a method of solving such problems with the aid of associated r -convex ones.

Keywords: nonlinear programming problem, r -convex transformability, sufficiency, the Karush-Kuhn-Tucker conditions

Classification: 49K99, 90C26

1. Introduction

The aim of the paper is to prove that the Karush-Kuhn-Tucker necessary optimality conditions (KKT conditions) are also sufficient for the r -convex transformable (r -convexifiable) nonsmooth nonlinear programming problem (PI) which reads:

Find a point $\bar{x} \in \mathbb{R}^n$, if it exists, such that

$$f_0(\bar{x}) = \min_{x \in S_I} f_0(x)$$

where

$$S_I = \{x \in \mathbb{R}^n \mid f(x) \leq 0\},$$

$f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$, $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are locally Lipschitz and r -convex transformable, see Definition 2.1.

In order to solve Problem (PI) one should find KKT points. However, our approach consists in seeking the KKT points of the above r -convex transformable problem using an associated r -convex problem (PC) which is much easier to be solved. As it appears the procedure provided in the paper has applications also when the problem (PI) is actually r -convex. In that case the r -convex problem could be chosen to be an easier one to be solved.

For definition of r -convexity, see [3]. For the locally Lipschitz case, we say that $f : X \rightarrow \mathbb{R}$, where $X \subset \mathbb{R}^n$ is a nonempty open convex set, is r -convex if for all $x, u \in X$,

$$\frac{1}{r}e^{rf(x)} \geq \frac{1}{r}e^{rf(u)} \left[1 + r(x - u)^T \xi \right] \quad \text{if } r \neq 0$$

and

$$f(x) - f(u) \geq (x - u)^T \xi \quad \text{if } r = 0$$

for any $\xi \in \partial f(u)$.

The r -convexity at a point has now an obvious meaning. Following the concept of r -convexity, T. Antczak defined a notion of r -invexity and investigated it in a vast number of papers, see for example [1], [2] and the references therein. We discuss the relation between r -convex transformability and r -invexity in the last part of Section 5.

2. r -convex transformability

We shall make clear now what we mean by r -convex transformability, compare [6], [9].

Definition 2.1. A locally Lipschitz function $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is said to be r -convex transformable or r -convexifiable (with respect to φ) provided there exists a C^1 diffeomorphism $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with C^1 inverse such that the composed function $f \circ \varphi^{-1}$ is r -convex.

If a function f is differentiable, then it suffices to assume that φ is only differentiable. In this paper we consider the case when $r \neq 0$. For the case $r = 0$, see [7], [8].

Some results concerning r -convex transformability and its applications in non-linear programming problems could be found in [4]. But our ones seem to be more applicable and provide the algorithms for solving the problems considered. Since r -convex transformable functions possess many r -convex-like properties it seems that some techniques used for r -convex functions will also apply. Our paper investigates a few of these relationships. We shall start with a sufficient condition for r -convex transformability and provide two simple criteria which allow one to check whether a twice-differentiable function f is r -convex transformable with respect to a certain φ . The criteria follow by applying the second order convexity condition to the function $x \mapsto \exp(r(f \circ \varphi^{-1})(x))$.

Proposition 2.2. Let a function $f : \mathbb{R} \rightarrow \mathbb{R}$ and a diffeomorphism $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ be twice differentiable and let $\bar{x} \in \mathbb{R}$. If $\varphi'(\bar{x}) \neq 0$ and

$$r f'(\varphi^{-1}(\bar{x})) \left[r f'(\varphi^{-1}(\bar{x})) - \varphi''(\bar{x}) \right] + r f''(\varphi^{-1}(\bar{x})) \geq 0$$

then f is r -convex transformable with respect to φ at \bar{x} .

Proposition 2.3. Let a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and a diffeomorphism $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be twice differentiable and let $\bar{x} \in \mathbb{R}^n$. Assume that the derivatives $\frac{\partial \varphi^{-1}(\bar{x})}{\partial x_i}$,

$\frac{\partial^2 \varphi^{-1}(\bar{x})}{\partial x_i \partial x_j}$ exist for any $i, j = 1, \dots, n$ and denote by A the $n \times n$ matrix with elements

$$a_{ij} = rf'(\varphi^{-1}(\bar{x})) \left[rf'(\varphi^{-1}(\bar{x})) \frac{\partial \varphi^{-1}(\bar{x})}{\partial x_i} \frac{\partial \varphi^{-1}(\bar{x})}{\partial x_j} - \frac{\partial^2 \varphi^{-1}(\bar{x})}{\partial x_i \partial x_j} \right] + rf''(\varphi^{-1}(\bar{x})) \frac{\partial \varphi^{-1}(\bar{x})}{\partial x_i} \frac{\partial \varphi^{-1}(\bar{x})}{\partial x_j}, \quad i, j = 1, \dots, n.$$

If A is positive semidefinite, then f is r -convex transformable with respect to φ at \bar{x} .

Since we are interested in r -convex transformability at KKT points only, the above proposition provides, together with an algorithm for checking the positive definiteness of the matrix in [5], quite a useful tool.

Example 2.4. A function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ given by

$$f(x_1, x_2) = \ln \left[(x_1 + x_1^3)^2 + (x_2 + x_2^3)^2 + 1 \right]$$

is r -convex transformable for all $r \in \mathbb{R}$. To prove this we need to find a diffeomorphism φ and check that the function $g = f \circ \varphi^{-1}$ is actually r -convex. Here $\varphi(x_1, x_2) = (x_1 + x_1^3, x_2 + x_2^3)$ and g is given by the formula $g(y_1, y_2) = \ln [y_1^2 + y_2^2 + 1]$. Hence by definition of r -convexity it follows that g is r -convex for any $r \neq 0$. Let us observe that the function f itself is not r -convex for all r . Indeed, the function

$$h(x_1, x_2) = \exp(rf(x_1, x_2)) = \left((x_1 + x_1^3)^2 + (x_2 + x_2^3)^2 + 1 \right)^r$$

is not convex, for instance, for $r = 1/10$ which follows by applying the second order convexity test.

It is worth to stress here that the class of r -convex transformable functions is wider than the class of r -convex functions. Hence our results apply not only for r -convex problems but also for such problems that have enough r -convexity-like properties.

3. The sufficiency of the KKT conditions

We shall prove that the assumption of r -convex transformability leads to the sufficient optimality conditions in the nonlinear programming problems.

Theorem 3.1. *Let functions f_0, f be r -convex transformable with respect to the same φ and let \bar{x} be a KKT point for problem (PI). Then \bar{x} is a solution to problem (PI).*

PROOF: Let us put $g_i = f_i \circ \varphi^{-1}$ for $i = 0, 1, 2, \dots, m$ and take any feasible x . Then $f_i = g_i \circ \varphi$ for $i = 0, 1, 2, \dots, m$ and

$$(3.1) \quad g_i(\varphi(x)) \leq 0 \quad \text{for } i = 1, 2, \dots, m.$$

Since \bar{x} is a KKT point for problem (PI) we obtain from [10] that there exists a Lagrange multiplier $\bar{\lambda} \in \mathbb{R}^m, \bar{\lambda} \geq 0$, such that

$$(3.2) \quad 0 \in \partial f_0(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \partial f_i(\bar{x}),$$

$$(3.3) \quad \bar{\lambda}_i f_i(\bar{x}) = 0 \quad \text{for } i = 1, 2, \dots, m.$$

Hence there exist subgradients $\tau \in \partial f_0(\bar{x})$ and $v_i \in \partial f_i(\bar{x})$ for $i = 1, 2, \dots, m$ such that

$$(3.4) \quad \tau + \sum_{i=1}^m \bar{\lambda}_i v_i = 0.$$

By the r -convex transformability of f_0 it follows that, for any $r \neq 0$,

$$(3.5) \quad \frac{1}{r} e^{r g_0(\varphi(x))} \geq \frac{1}{r} e^{r g_0(\varphi(\bar{x}))} (1 + r(\varphi(x) - \varphi(\bar{x}))^T \zeta)$$

where $\zeta \in \partial_y g_0(\varphi(\bar{x}))$. An application of the chain rule yields $\zeta = (\nabla_x \varphi(\bar{x}))^{-1} \tau$. From (3.4) we thus get

$$(3.6) \quad \begin{aligned} \frac{1}{r} e^{r g_0(\varphi(x))} &\geq \frac{1}{r} e^{r g_0(\varphi(\bar{x}))} \\ &- e^{r g_0(\varphi(\bar{x}))} (\varphi(x) - \varphi(\bar{x}))^T \left(\sum_{i=1}^m \bar{\lambda}_i (\nabla_x \varphi(\bar{x}))^{-1} v_i \right). \end{aligned}$$

By (3.6) and the chain rule we now obtain

$$(3.7) \quad \begin{aligned} \frac{1}{r} e^{r g_0(\varphi(x))} &\geq \frac{1}{r} e^{r g_0(\varphi(\bar{x}))} \\ &- e^{r g_0(\varphi(\bar{x}))} (\varphi(x) - \varphi(\bar{x}))^T \left(\sum_{i=1}^m \bar{\lambda}_i \zeta_i \right) \end{aligned}$$

where $\zeta_i \in \partial_y g_i(\varphi(\bar{x}))$ for $i = 1, 2, \dots, m$. Thus by the r -convex transformability of f_i for $i = 1, 2, \dots, m$ and by (3.7) we get

$$(3.8) \quad \begin{aligned} & \frac{1}{r}e^{rg_0(\varphi(x))} - \frac{1}{r}e^{rg_0(\varphi(\bar{x}))} \\ & \geq \frac{1}{r}e^{rg_0(\varphi(\bar{x}))} \sum_{i=1}^m \bar{\lambda}_i \left(1 - e^{rg_i(\varphi(x)) - rg_i(\varphi(\bar{x}))} \right). \end{aligned}$$

Now we show that

$$(3.9) \quad \frac{1}{r}e^{rg_0(\varphi(\bar{x}))} \sum_{i=1}^m \bar{\lambda}_i \left(1 - e^{rg_i(\varphi(x)) - rg_i(\varphi(\bar{x}))} \right) \geq 0.$$

Let us first consider the case when $r > 0$. Of course $\frac{1}{r}e^{rg_0(\varphi(\bar{x}))} > 0$. Since $\bar{\lambda}_i g_i(\varphi(\bar{x})) = 0$ for $i = 1, 2, \dots, m$ we obtain that either $\bar{\lambda}_i = 0$ or $g_i(\varphi(\bar{x})) = 0$ for $i = 1, 2, \dots, m$. By the above arguments and since $rg_i(\varphi(x)) \leq 0$ it follows that $\bar{\lambda}_i(1 - e^{rg_i(\varphi(x)) - rg_i(\varphi(\bar{x}))})$ is either nonnegative or zero. The case when $r < 0$ follows in a similar manner. Hence (3.8) follows. By (3.8) and (3.9) it follows that

$$(3.10) \quad \frac{1}{r}e^{rg_0(\varphi(x))} \geq \frac{1}{r}e^{rg_0(\varphi(\bar{x}))}$$

hence $f_0(x) \geq f_0(\bar{x})$ i.e. \bar{x} is a solution to problem (PI). □

A thorough analysis of the above proof shows that we may impose weaker conditions on the inequality constraints. It also reveals that the KKT points to both the problems are related by the diffeomorphism φ . We shall deal with that questions in the next section, where we formulate an associated problem (PC) which is “less nonlinear” and equivalent to problem (PI) in a certain sense.

4. An associated problem

In this section we shall give new sufficient conditions for optimality in problem (PI) under less restrictive hypotheses on the data than r -convex transformability. We shall use an associated r -convex problem (PC) to find the candidates for the solution to the problem considered. Similarly to well known convexity- or invexity-type assumptions we shall make our assumptions solely at KKT points.

Let \bar{x} be a KKT point for problem (PI) with $\bar{\lambda}$ being the vector of Lagrange multipliers. Denote by I the set of the indices of all active constraint functions at \bar{x} , i.e.

$$(4.1) \quad I := \{1 \leq i \leq m \mid f_i(\bar{x}) < 0\}.$$

Let f_0 be r -convex transformable at the point \bar{x} with respect to $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and let $\bar{y} := \varphi(\bar{x})$. Define the functions $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ as follows

$$(4.2) \quad g_i := f_i \circ \varphi^{-1} \quad \text{for } i = 0, 1, 2, \dots, m.$$

We impose the following assumptions on the constraint functions, for any feasible y and all $\zeta_i \in \partial g_i(\bar{y})$, $i \in I$:

$$(4.3) \quad \zeta_i^T (y - \bar{y}) \leq 0.$$

The conditions imposed on g_i , $i \in I$, are direct generalizations of the notion of r -convex transformability at a point. This may be viewed as a version of quasiconvexity, although it is more general, compare [11]. Indeed consider a function $f(x) = x^4 - x^2 + 1$ over a set $X = (-1, +\infty)$. Take a point $\bar{x} = -1$. Then $f'(\bar{x}) = -2$ and relation (4.3) reads $-2(x + 1) \leq 0$. While for the quasiconvexity of f at \bar{x} it is required that $f(x) \leq f(\bar{x}) \Rightarrow f'(\bar{x})(x - \bar{x}) \leq 0$ for all $x \in X$ and $f(x) \leq f(\bar{x})$ for $x \in [-1, 1]$.

Now we are in position to consider an associated problem (PC):

Find an $\bar{y} \in S_C$, if it exists, such that

$$g_0(\bar{y}) = \min_{y \in S_C} g_0(y)$$

where

$$S_C = \{y \in \mathbb{R}^n \mid g(y) \leq 0\},$$

$g_0 : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ satisfy (4.3).

The KKT points of problems (PI) and (PC) are connected each to other in the following way.

Proposition 4.1. *Let f_0 be r -convex transformable at \bar{x} with respect to φ and let (4.3) hold. The point $\bar{x} \in S_I$ is a KKT point for (PI) iff $\bar{y} = \varphi(\bar{x}) \in S_C$ is a KKT point for (PC). Moreover, the vector of Lagrange multipliers remains the same.*

PROOF: It follows easily from the r -convex transformability assumption and the properties of a coordinate transform φ , compare [7]. □

Now we may formulate and prove the sufficient optimality condition for problem (PI).

Proposition 4.2. *Let \bar{x} be a KKT point for problem (PI) and let f_0 be r -convex transformable at \bar{x} with respect to φ . Assume that (4.3) holds. Then \bar{x} is a solution to problem (PI).*

PROOF: Reasoning as in the proof of Theorem 3.1 we get

$$(4.4) \quad \frac{1}{r}e^{rg_0(\varphi(x))} \geq \frac{1}{r}e^{rg_0(\varphi(\bar{x}))} - e^{rg_0(\varphi(\bar{x}))} (\varphi(x) - \varphi(\bar{x}))^T \left(\sum_{i=1}^m \bar{\lambda}_i \zeta_i \right),$$

where $\zeta_i \in \partial_y g_i(\varphi(\bar{x}))$ for $i = 1, 2, \dots, m$. Thus by the r -convex transformability assumption, (4.3) and properties of φ we get

$$(4.5) \quad \frac{1}{r}e^{rg_0(\varphi(x))} \geq \frac{1}{r}e^{rg_0(\varphi(\bar{x}))}.$$

Therefore

$$(4.6) \quad g_0(\varphi(x)) \geq g_0(\varphi(\bar{x}))$$

which means that

$$(4.7) \quad f_0(x) \geq f_0(\bar{x}).$$

□

From the above results it follows that the solution to (PI) may be obtained from the solution of (PC). The procedure is as follows. Problem (PI) is transformed to an r -convex problem (PC) which is usually easier to be solved. The values of the objective functions at optimal points for both problems are equal. If we are interested in finding the KKT point for (PI) these can be obtained by the same transformation φ .

5. Applications and remarks

We shall start with a concrete example:

Example 5.1. Consider problem (PI)

$$\log \left((x^3 + x)^2 + 1 \right) \rightarrow \min$$

subject to

$$\log (x^3 + x + 1) - 10 \leq 0$$

over $C = \{x : x \geq 1\}$. We show that this problem is r -convex transformable. We put $\varphi(x) = x^3 + x$ and consider the following 1-convex problem

$$\log(y^2 + 1) \rightarrow \min$$

subject to

$$\log(y + 1) - 10 \leq 0$$

over $C_1 = \{y : y \geq 2\}$. Its solution is obviously $y = 2$ and in a consequence an x satisfying $x^3 + x = 2$, i.e. $x = 1$.

To end the paper we provide a few remarks upon some other types of problems that can be considered by our approach.

The differentiable case. In case the functions involved are differentiable we have similar results but we may include equality constraint, i.e. problem (PI) now reads:

Find a point $\bar{x} \in \mathbb{R}^n$, if it exists, such that

$$f_0(\bar{x}) = \min_{x \in S_I} f_0(x)$$

where

$$S_I = \{x \in \mathbb{R}^n \mid f(x) \leq 0, h(x) = 0\}.$$

In case the equality constraint appears it must be continuously differentiable anyway. The assumptions are as follows:

Let f_0 be r -convexifiable at a point \bar{x} with respect to φ and let for all active constraints f_i the functions $g_i = f_i \circ \varphi^{-1}$ satisfy for any feasible y

$$(5.1) \quad \nabla g_i(\bar{y})^T (y - \bar{y}) \leq 0$$

and for all $j = 1, 2, \dots, k$ let h_j be such that $p_j = h_j \circ \varphi^{-1}$ satisfies for any feasible y

$$(5.2) \quad \text{sgn}(\bar{v}_i) \nabla p_j(\bar{y})^T (y - \bar{y}) \leq 0,$$

where \bar{v} is the Lagrange multiplier associated with the equality constraint.

The associated problem (PC) now reads:

Find a point $\bar{y} \in \mathbb{R}^n$, if it exists, such that

$$g_0(\bar{y}) = \min_{y \in S_C} g_0(y)$$

where

$$S_C = \{y \in \mathbb{R}^n \mid g(y) \leq 0, p(y) = 0\}.$$

With the above assumptions, the results concerning solvability of the problem stated follow:

Proposition 5.2. *A point $\bar{x} \in S_I$ is a KKT point for (PI) iff $\bar{y} = \varphi(\bar{x}) \in S_C$ is a KKT point for (PC). Moreover, the vector of Lagrange multipliers remains the same. Both the points \bar{x} and \bar{y} are solutions to problems (PI) and (PC), respectively.*

Problems with nonnegative variable. The approach presented in the paper is very useful in case of problems with a nonnegative variable, i.e., problem (PIE)

$$\text{minimize } f_0(x)$$

over

$$S_I = \{x \in \mathbb{R}^n \mid f(x) \leq 0, x \geq 0, h(x) = 0\}.$$

Such problems are very important from the applicational point of view. But it is extremely difficult to provide sufficient optimality conditions different from convexity or its standard generalizations due to the presence of a linear inequality constraint. Thus it appears that r -convex transformability may prove to be quite a useful tool in that case contrary to invexity, compare [11] for a suitable example. The additional hypothesis we have to impose on the diffeomorphism φ is quite natural and satisfied in most cases, i.e., we require that $\varphi(x) \geq 0$ iff $x \geq 0$. This allows us to prove the following:

Theorem 5.3. *A point \bar{x} at which problem (PIE) is r -convex transformable, i.e. (5.1) and (5.2) are satisfied, is its global solution if and only if the point $\varphi(\bar{x})$ is a global solution of problem (PIC)*

$$\text{minimize } g_0(y)$$

over

$$S_C = \{y \in \mathbb{R}^n \mid g(y) \leq 0, y \geq 0, p(y) = 0\}.$$

The proof is similar to the proofs of Propositions 4 and 5 from [8] and it relies on defining the equivalent problem

$$\text{minimize } f_0(x)$$

over

$$S_{IE} = \{x \in \mathbb{R}^n \mid f(x) \leq 0, \varphi(x) \geq 0, h(x) = 0\}$$

and later reasoning as in the above.

Connections with r -invexity. It is well known that in some cases the invex functions are convex transformable, provided the scale functions η satisfies certain assumptions, see [6] and [7], [8] for a more applied approach. Hence it is obvious that a similar result would hold in our case. This explains the connections of our results with those of [1], [2].

Indeed, let us define a problem (PI) to be *generalized r -invex* at a point \bar{x} provided that there exists a function $\eta : S_I \rightarrow \mathbb{R}^n$ such that for all $x \in S_I$ and for all active constraint functions f_i the following relations hold ($r \neq 0$):

$$\begin{aligned} \frac{1}{r}e^{rf_0(x)} &\geq \frac{1}{r}e^{rf_0(\bar{x})}(1 + r(\eta(x))^T \nabla f_0(\bar{x})), \\ 0 &\geq \eta(x)^T \nabla f_i(\bar{x}) \quad \text{for } i \in I. \end{aligned}$$

By the properties of φ and standard calculus it follows that

Proposition 5.4. *Let f_0 be r -convex transformable at \bar{x} with respect to φ and let (4.3) hold. Then problem (PI) is generalized r -invex at \bar{x} with respect to function η given by the formula*

$$\eta(x) = (\nabla_x \varphi(\bar{x}))^{-1}(\varphi(x) - \varphi(\bar{x})).$$

Moreover, assuming generalized r -invexity with respect to the above η we obtain that (4.3) is satisfied at \bar{x} .

Remark 1. We may restrict our considerations to functions defined on a certain convex set $C \subset \mathbb{R}^n$ by adding an additional constraint $x \in C$. In that case the diffeomorphism φ should be such that the set $\varphi(C)$ is also convex. All results provided above are valid with that additional assumption. In order to make our approach readable we have not included that assumption. However the addition of an assumption $x \in C$ would result in some technical changes in the calculations above, for example in formula (3.2).

REFERENCES

- [1] Antczak T., *(p, r)-invex sets and functions*, J. Math. Anal. Appl. **263** (2001), 355–379.
- [2] Antczak T., *Lipschitz r -invex functions and nonsmooth programming*, Numer. Funct. Anal. Optim. **23** (2002), no. 3–4, 265–283.
- [3] Avriel M., *r -convex functions*, Math. Programming **2** (1972), 309–323.
- [4] Avriel M., Diewert W.E., Schaible S., Zang I., *Generalized Concavity*, Plenum Press, New York, 1988.
- [5] Bazaraa M.S., Sherali H.D., Shetty C.M., *Nonlinear Programming. Theory and Algorithms*, J. Wiley, New York, 1993.
- [6] Craven B.D., *Relationships between invex properties*, in: Recent Trends in Optimization Theory, Agrawal R.P., Ed., World Scientific, Singapore, 1995.
- [7] Galewski M., *On some connection between invex and convex problems in nonlinear programming*, Control and Cybernetics **30** (2001), no. 1, 11–22.
- [8] Galewski M., *A note on invex problems with nonnegative variable*, European J. Oper. Res. **163** (2005), no. 2, 565–568.
- [9] Hanson M.A., Mond B., *Convex transformable programming problems and invexity*, J. Inform. Optim. Sci. **8** (1987), no. 2, 201–207.

- [10] Hiriart-Urruty J.B., *Refinements of necessary optimality conditions in nondifferentiable programming*, Appl. Math. Optim. **5** (1979), 63–82.
- [11] Martin D.H., *The essence of invexity*, J. Optim. Theory Appl. **47** (1985), 65–76.

FACULTY OF MATHEMATICS, UNIVERSITY OF LODZ, BANACHA 22, 90-238 LODZ, POLAND

E-mail: emlynar@imul.uni.lodz.pl
galewski@imul.uni.lodz.pl

(Received November 24, 2004, revised January 13, 2005)