Arcwise-convex functions on surfaces

T. RAPCSÁK[1]) (Budapest)

Abstract. The paper considers the characteristics of the arcwise-convex functions, then gives on this basis second-order necessary and sufficient conditions to enable a function to be arcwise-convex on a surface.

Key Words: arcwise-convex functions, nonlinear programming, differential geometry.

1. Introduction

The convex programming problems are an important class of nonlinear programming problems. The importance of the class is due to the circumstance that here any local optimum is at the same time global, therefore the algorithms are substantially more efficient than in other cases. In the field of mathematical programming many authors have dealt with the generalization of the convexity concept. A survey of some classes of generalized convex functions can be found in Avriel (Ref. 1). These classes include quasi-convex, strictly quasi-convex, strongly quasi-convex, pseudo-convex and strictly pseudo-convex functions. A unifying property of these functions is that their level sets are convex.

Ortega and Rheinboldt (Ref. 2) extended the families of quasi-convex and strongly quasi-convex functions by considering continuous arcs instead of line segments in the definitions. Arcwise-convex and arcwise-connected functions are defined and some of their properties are shown in Avriel (Ref. 1). In the paper of Avriel and Zang (Ref. 3) the generalized arcwise-connected functions are defined and characterized with respect to the local-global minimum properties. The main results are that under some mild regularity assumptions a local-global property of a function implies its class inclusion in one of the generalized arcwise-connected functions. A more general version of the local-global properties is derived in Martin's works (Ref. 4, 5, 6) where the concept of connectedness is utilized, instead of arcwise-connectedness. The relationship between the local-global properties and the lower-semicontinuous point-to-set mappings has been discussed by Zang and Avriel (Ref. 7.).

Pointwise convex functions in the nondifferentiable case are introduced by Brosowski (Ref. 8, 9).

[1]) Senior research worker, Department of Applied Mathematics, Computer and Automation Institute of the Hungarian Academy of Sciences, *Budapest*, Hungary.

The purpose of this paper is to set second-order necessary and sufficient conditions which characterize the arcwise-convex functions on surfaces. This situation comes about in the non-convex programming problems of a certain type (e.g. a non-linear programming problem constrained by equality conditions where the conditions define a surface).

2. The arcwise-convex functions

Following ORTEGA and RHEINBOLDT (Ref. 2), AVRIEL (Ref. 1), AVRIEL and ZANG (Ref. 3), we first extend the concept of convexity.

Definition 2.1. A set $C \subset R^n$ is said to be arcwise-connected if for every pair of points $\underline{x}_1 \in C$, $\underline{x}_2 \in C$ there is at least one continuous vector-valued function $\underline{x}(t)$, called an arc, defined on the unit interval $[0, 1] \subset R$ for which

(1) $$\underline{x}(t) \in C, \quad \forall t \in [0, 1] \quad \text{and} \quad \underline{x}(0) = \underline{x}_1, \quad \underline{x}(1) = \underline{x}_2.$$

Note that $\underline{x}(t)$ generally depends on the points $\underline{x}_1$, $\underline{x}_2$ and the set C, and for a pair of points $\underline{x}_1$, $\underline{x}_2$ in an arcwise-connected set there may exist more than one single arc.

It is obvious that a convex set is also arcwise-connected and

(2) $$\underline{x}(t) = (1-t)\underline{x}_1 + t\underline{x}_2, \quad 0 \leqq t \leqq 1.$$

Definition 2.2. Let $C \subset R^n$ be an arcwise-connected set. A subset of the connecting arcs is called a family of the feasible arcs or shortly, feasible arcs if for every pair of points in C there exists a connecting arc belonging to the selected subset.

Definition 2.3. Let $C \subset R^n$ be an arcwise-connected set and F a family of the feasible arcs. The real-valued function $f(\underline{x})$, defined on the set C, is then said to be arcwise-convex (arcwise-connected) for the set F if for every pair of points $\underline{x}_1 \in C$, $\underline{x}_2 \in C$ and a feasible arc $\underline{x}(t)$ we have

(3) $$f(\underline{x}(t)) \leqq (1-t) f(\underline{x}_1) + t f(\underline{x}_2), \quad 0 \leqq t \leqq 1.$$

This idea can be further extended (Ref. 1.3).

The first statement, which results e.g. from the more general theorems of AVRIEL and ZANG (Ref. 3, 7), shows the importance of the arcwise-convexity.

Lemma 2.1. *If $f(\underline{x})$ is arcwise-convex (for a set of functions F) on C and $\underline{x}_0$ is a local minimum, then $\underline{x}_0$ is also a global minimum point.*

Assume further that $C \subset R^n$ is an open, connected set. Hence we get that C is also arcwise-connected (Ref. 1).

Lemma 2.2. *Let $f(\underline{x})$ be a differentiable, scalar function on C and the feasible arcs be differentiable. Then $f(\underline{x})$ is arcwise-convex if and only if for every pair of points*

$\underline{x}_1 \in C$, $\underline{x}_2 \in C$ and $\underline{x}(t)$ $(\underline{x}(0)=\underline{x}_1,\ \underline{x}(1)=\underline{x}_2)$ feasible arcs

$$f(\underline{x}_2)-f(\underline{x}_1) \geqq \nabla f(\underline{x}_1)\dot{\underline{x}}(0). \tag{4}$$

(In the lemma $\nabla f(\underline{x}_1)$ and $\dot{\underline{x}}(0)$ mean respectively the gradient of $f(\underline{x})$ at the point $\underline{x}_1$ and the derivative of $\underline{x}(t)$ by t at the point 0.)

Lemma 2.3. *Let $f(\underline{x})$ be a scalar, twice continuously differentiable function on C and the feasible arcs be twice continuously differentiable. Then $f(\underline{x})$ is arcwise-convex if and only if for every feasible arc $\underline{x}(t)$*

$$\frac{d^2}{dt^2} f(\underline{x}(t)) \geqq 0, \quad 0 \leqq t \leqq 1. \tag{5}$$

The proofs of Lemma 2.2 and Lemma 2.3 being similar to the convex case will be omitted.

If the feasible arcs are given in arc length parameter, then between two arbitrarily chosen points $\underline{x}_1, \underline{x}_2 \in C$ there are in form $\underline{x}(s)$, $\underline{x}(0)=\underline{x}_1$, $\underline{x}(s_0)=\underline{x}_2$. Introducing the notation $\underline{x}(s)=\underline{x}(\lambda s_0)=\tilde{\underline{x}}(\lambda)$, $0 \leqq \lambda \leqq 1$ we obtain the case discussed above.

Further the differentiation by the arc length will be noted by a prime, i.e.

$$\frac{d\underline{x}(s)}{ds} = \underline{x}'(s). \tag{6}$$

3. Generalization of convexity properties of functions to surface

In this section it will be examined when a $\underline{x}(\underline{u})$, $\underline{u} \in U_k$ surface of dimension k can be called convex in any sense. In differential geometry the generalization of convexity led to geodesically convex surfaces and sets. A surface (set) is called geodesically convex when the geodesic of minimal length connecting any of its two points also belongs to the surface (set) (Ref. 10, 11). Ref. 12 deals with the problem when a surface of dimension $(n-1)$ is convex. The approach of this paper is different because in addition to the surface it is given also a vector field which is determined in a nonlinear programming problem by the objective function.

First the differential geometric interpretation of the convexity concept of the functions is needed. Let $Z \subset R^n$ be an open, convex set and $f(\underline{x})$ a twice continuously differentiable convex function defined on Z. It is well-known that $f(\underline{x})$ is convex on Z if and only if its Hessian matrix is positive semidefinite in every point. Consider now in the R^{n+1} space the $y-f(\underline{x})=0$ level surface.

This surface can be written in the following form:

$$\underline{x}(\underline{u}) = \begin{cases} x_1 = u_1 \\ x_2 = u_2 \\ \vdots \\ x_n = u_n \\ x_{n+1} = f(\underline{u}) \end{cases}, \quad \underline{u} \in U_n = Z. \tag{7}$$

The parameter-line tangents of the surface (7) are

$$\frac{\partial \underline{x}}{\partial u_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \dfrac{\partial f}{\partial u_1} \end{pmatrix}, \quad \frac{\partial \underline{x}}{\partial u_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ \dfrac{\partial f}{\partial u_2} \end{pmatrix}, \ldots, \quad \frac{\partial \underline{x}}{\partial u_n} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ \dfrac{\partial f}{\partial u_n} \end{pmatrix}. \tag{8}$$

These vectors are linearly independent, therefore (7) determines a surface in differential geometrical sense. The normal vector of the hypersurface (7) in an arbitrary point is given by

$$\underline{n} = (-\nabla f(\underline{x}), 1). \tag{9}$$

Determine the matrix B of the second fundamental form of the surface (7). According to the definition the elements of B are

$$b_{ij} = \frac{\partial^2 \underline{x}}{\partial u_i \partial u_j} \cdot \frac{1}{\|\underline{n}\|} \underline{n}, \quad i, j = 1, \ldots, n \tag{10}$$

where $\|\underline{n}\|$ means the norm of $\underline{n}$.

Hence it follows that

$$B = \frac{1}{\|\underline{n}\|} H \tag{11}$$

where H is the Hessian matrix of $f(\underline{x})$.

As a result, the convexity is composed of the global characteristics of the basic set (Z is a convex, open set) and of the local properties of the function $f(\underline{x})$ (the Hessian matrix or the matrix of the second fundamental form of the hypersurface (7) is positive semidefinite in every point).

Consider now a surface of dimension k. The normal vector space of this surface is of dimension $(n-k)$ and as the second fundamental form can be interpreted not only in one direction, a direction field should be given too.

Definition 3.1. Let $\underline{x}(\underline{u})$, $\underline{u} \in U_k$ be a surface of dimension k and V be a vector field on the surface. (The vectors of the vector field should not necessarily be contained in the normal space.) The surface $\underline{x}(\underline{u})$ is said to be convex (concave) in the direction of the vector field V, if the matrix of the second fundamental form using the vector of the vector field V instead of $\underline{n}$ is positive (negative) semidefinite in every point.

By the surface (7) the vector field V consisted of the vectors $(-\nabla f(\underline{x}), 1)$.

The above definition is more general than the classic convexity notion as the basic set is not convex, the surface is no special level surface and also the vector field V can be arbitrary.

4. Arcwise-convex functions on surfaces

The best feature of convex programming problems in nonlinear programming is that any local minimum is at the same time global such that the algorithms are substantially more efficient than in other cases. But the property of this kind is not known if there are also equality conditions in the nonlinear programming problem. In this section we intend, if the constraints of the nonlinear programming problem determine a surface, to set necessary and sufficient conditions to ensure for a function to be arcwise-convex on surfaces. For this purpose we utilize the results of the preceding sections.

Consider the $\underline{x}(\underline{u})\in R^n$, $\underline{u}\in U_k$ surface of dimension k and the twice continuously differentiable function $f(\underline{x})$ defined on the open $C\subset R^n$ $(\underline{x}(\underline{u})\in C,\ \underline{u}\in U_k)$ set. By the differential geometric investigations the local coordinate system in the points of the surface consists of vectors spanning the tangent vector space and the so-called normal space orthogonal to the tangent space. Assume without loss of generality the choice of a local, orthonormal coordinate system in which the component in normal directions of the vector $\nabla f(\underline{x})$ (denoted by $\nabla f(\underline{x})_N$) shows in the direction of the first normally-directed coordinate axis. Thus

$$\nabla f(\underline{x}) = \nabla f(\underline{x})_T + \nabla f(\underline{x})_N \tag{12}$$

where $\nabla f(\underline{x})_T$ means the component in tangential directions.

Theorem 4.1. *Let the feasible arcs on the surface $\underline{x}(\underline{u})$ be the geodesics given in arc length parameter. Then the necessary and sufficient condition of the arcwise-convexity (geodesically convexity) of the function $f(\underline{x})$ on the surface $\underline{x}(\underline{u})$ is the positive semidefiniteness of the following matrix in every point:*

$$H|_{TM} + |\nabla f(\underline{x})_N| B_{\nabla f_N} \tag{13}$$

where $H|_{TM}$ is the Hessian matrix of the function $f(\underline{x})$ restricted to the tangent space of $\underline{x}(\underline{u})$,

$B_{\nabla f_N}$ is the second fundamental form of $\underline{x}(\underline{u})$ in the normal direction of the vector $\nabla f(\underline{x})$.

Proof. Using Lemma 2.3 and the subsequent remark it is sufficient to prove that considering any feasible arc (geodesic) between two arbitrarily chosen points $\underline{x}_1, \underline{x}_2\in C$

$$\frac{d^2}{dt^2} f(x(ts_0)) \geq 0, \quad 0 \leq t \leq 1 \quad (\underline{x}(0) = \underline{x}_1,\ \underline{x}(s_0) = \underline{x}_2). \tag{14}$$

As the feasible arcs are given in arc length parameter, it follows that

$$\frac{d}{dt} f(\underline{x}(ts_0)) = \nabla f(\underline{x}(s))\underline{x}'(s)s_0, \tag{15}$$

$$\frac{d^2}{dt^2} f(\underline{x}(ts_0)) = \underline{x}'(s)^T \nabla^2 f(\underline{x}(s))\underline{x}'(s)s_0^2 + \nabla f(\underline{x}(s))\underline{x}''(s)s_0^2.$$

As $\underline{x}(\underline{u})$ is a surface we obtain for the derivatives of the curve $\underline{x}(\underline{u}(s))$ the following expressions:

$$\frac{d\underline{x}(\underline{u}(s))}{ds} = \frac{\partial \underline{x}}{\partial u_i} u_i', \tag{16}$$

$$\frac{d^2\underline{x}(\underline{u}(s))}{ds} = \frac{\partial^2 \underline{x}}{\partial u_i \partial u_j} u_i' u_j' + \frac{\partial \underline{x}}{\partial u_i} u_i'',$$

where according to the Einstein convention the identical index occurring twice in one term is understood a summation also without writing out the summation sig n.

According to the Gauss equations

$$\frac{\partial^2 \underline{x}}{\partial u_i \partial u_j} = \Gamma_{ij}^{\sigma} \frac{\partial \underline{x}}{\partial u_\sigma} + b_{ij}^{\gamma} \underline{n}_\gamma, \quad i, j = 1, \dots, k \tag{17}$$

where the Γ_{ij}^{σ} quantities are the Christoffel symbols of the second kinds and the b_{ij}^{γ} quantities are the elements of the matrix of the second fundamental form in the corresponding normal directions.

Using the Gauss equations in the second equality of (16) we obtain that

$$\frac{d^2\underline{x}(\underline{u}(s))}{ds} = (\Gamma_{ij}^{\sigma} u_i' u_j' + u_\sigma'') \frac{\partial \underline{x}}{\partial u_\sigma} + b_{ij}^{\gamma} u_i' u_j' \underline{n}_\gamma. \tag{18}$$

The equation of the geodesics in arc length parameter is

$$u_\sigma'' = -\Gamma_{ij}^{\sigma} u_i' u_j' \tag{19}$$

and since $\underline{x}(\underline{u}(s))$ is geodesics, the expression in parenthesis disappears in equation (18).

Hence we have the following:

$$\nabla f(\underline{x}(s))\, \underline{x}''(s) = (\nabla f(\underline{x})_T + \nabla f(\underline{x})_N)\, \underline{x}''(s) = |\nabla f(\underline{x})_N|\, b_{ij}^1 u_i' u_j'. \tag{20}$$

(In equality (20) it was used that $\nabla f(\underline{x})_T$ is orthogonal to every normal vector and $\nabla f(\underline{x})_N$ is in the direction of the first normal coordinate axis.)

It follows from the equality (20) that the expression $\nabla f(\underline{x}(s))\underline{x}''(s)$ is nonnegative if and only if the surface $\underline{x}(\underline{u})$ is convex in the direction of the vector field $\nabla f(\underline{x})_N$.

The function $f(\underline{x})$ is convex such that the Hessian matrix is in every point positive semidefinite, therefore on the basis of (15) the inequalities (14) are in fact fulfilled i.e. the function $f(\underline{x})$ is arcwise-convex on the surface $\underline{x}(\underline{u})$.

Corollary 4.1. Under the conditions of Theorem 4.1 every local minimum of the function $f(\underline{x})$ on the surface $\underline{x}(\underline{u})$ is global minimum, too.

Corollary 4.2. *Let the feasible arcs on the surface $\underline{x}(\underline{u})$ be the geodesics given in arc length parameter. Then a sufficient conditions for the function $f(\underline{x})$ to be arcwise-convex (geodesically convex) on the surface $\underline{x}(\underline{u})$ is that the function should be con-*

vex on the set C and the surface $\underline{x}(\underline{u})$ should be convex in the direction of the vector field $\nabla f(\underline{x})_N$.

The PROOF is similar to the proof of Theorem 4.1.

If instead of the surface $\underline{x}(u)$, we consider a Riemannian manifold immersed in a Euclidean n-space, the statements remain valid.

References

[1] M. AVRIEL, Nonlinear Programming — Analysis and Methods, *Prentice Hall, Inc. N. J.*, 1976.

[2] J. M. ORTEGA and W. C. RHEINBOLDT, Iterative Solution of Nonlinear Equations in Several variables, *Academic Press, N. Y.*, 1970.

[3] M. AVRIEL and I. ZANG, Generalized Arcwise Connected Functions and Characterizations of Local-Global Properties, *Journal of Optimization Thery and Applications*, **32** (1980), 407—425.

[4] D. H. MARTIN, Connected Level Sets, and Uniqueness in Optimization — Part I, *National Research Institute for Mathematical Sciences, Pretoria*, CSIR Special Report WISK 276, September 1977.

[5] D. H. MARTIN, Connected Level Sets, and Uniqueness in Optimization — Part II, *National Research Institute for Mathematical Sciences, Pretoria*, CSIR Special Report WISK 277, October 1977.

[6] D. H. MARTIN, Connected Level Sets, Minimizing Sets, and Uniqueness in Optimization. *Journal of Optimization Theory and Applications*, 36 (1982), 71—93.

[7] I. ZANG and M. AVRIEL, On Functions Whose Local Minima Are Global, *Journal of Optimization Theory and Applications*, **16** (1975), 183—191.

[8] B. BROSOWSKI, Parametric Semi-Infinite Optimization, *Verlag Peter Lang, Frankfurt Am Maine, Bern*, 1982.

[9] B. BROSOWSKI, Parametric Optimization and Approximation *(to appear)*.

[10] N. J. HICKS, Notes on Differential Geometry, *D. Van Nostrand Company, Princeton, New Jersey, Toronto, New York, London*, 1965.

[11] M. SPIVAK, A Comprehensive Introduction to Differential Geometry, *Publish or Perish, Berkeley*, 1979.

[12] F. NOZICKA, Spherical Mappings of Convex Sets and Convex Parametrical Optimization, Survey of Mathematical Programming Proceedings of the 9-th International Mathematical Programming Symposium, Vol. 1., *Edited by A. Prékopa. Akadémiai Kiadó, Budapest*, (1979), 299—307.

[13] T. RAPCSÁK, On the arcwise-convexity, *Alkalmazott Matematikai Lapok* No. **10** (1984), 115—123 (in Hungarian).

[14] A. RAPCSÁK and L. TAMÁSSY, Differential Geometry, *Tankönyvkiadó*, Budapest, 1967 (in Hungarian).

[15] B. MARTOS, Nonlinear Programming Theory and Methods, *Akadémiai Kiadó*, Budapest, 1975.

[16] O. L. MANGASARIAN, Nonlinear Programming, *McGraw-Hill-Book Company*, New York, 1969.

(Received November 3, 1984)