J.SIAM Control
Ser. A, Vol. 3, No. 2
Printed in U.S.A., 1965

PSEUDO-CONVEX FUNCTIONS*

O. L. MANGASARIAN†

Abstract. The purpose of this work is to introduce pseudo-convex functions and to describe some of their properties and applications. The class of all pseudo-convex functions over a convex set C includes the class of all differentiable convex functions on C and is included in the class of all differentiable quasi-convex functions on C. An interesting property of pseudo-convex functions is that a local condition, such as the vanishing of the gradient, is a global optimality condition. One of the main results of this work consists of showing that the Kuhn-Tucker differential conditions are sufficient for optimality when the objective function is pseudo-convex and the constraints are quasi-convex. Other results of this work are a strict converse duality theorem for mathematical programming and a stability criterion for ordinary differential equations.

1. Introduction. Throughout this work, we shall be concerned with the real, scalar, single-valued, differentiable function $\theta(x)$ defined on the nonempty open set D in the m-dimensional Euclidean space E^m. We let C be a subset of D and let ∇_x denote the $m \times 1$ partial differential operator

$$\nabla_x = \left[\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_m}\right]',$$

where the prime denotes the transpose. We say that $\theta(x)$ is *pseudo-convex* on C if for every x^1 and x^2 in C,

$$(1.1) \qquad (x^2 - x^1)'\nabla_x\theta(x^1) \geqq 0 \quad \text{implies} \quad \theta(x^2) \geqq \theta(x^1).$$

We say that $\theta(x)$ is *pseudo-concave* on C if for every x^1 and x^2 in C,

$$(1.2) \qquad (x^2 - x^1)'\nabla_x\theta(x^1) \leqq 0 \quad \text{implies} \quad \theta(x^2) \leqq \theta(x^1).$$

Thus $\theta(x)$ is pseudo-concave if and only if $-\theta(x)$ is pseudo-convex. In the subsequent paragraphs we shall confine our remarks to pseudo-convex functions. Analogous results hold for pseudo-concave functions by the appropriate multiplication by -1.

We shall relate the pseudo-convexity concept to the previously established notions of convexity, quasi-convexity [1], [2] and strict quasi-convexity [3], [5].

The function $\theta(x)$ is said to be *convex* on C, [2], if C is convex and if for every x^1 and x^2 in C,

$$(1.3) \qquad \theta(\lambda x^1 + (1 - \lambda)x^2) \leqq \lambda\theta(x^1) + (1 - \lambda)\theta(x^2)$$

* Received by the editors March 4, 1965.
† Shell Development Company, Emeryville, California.

for every λ such that $0 \leqq \lambda \leqq 1$. Equivalently, $\theta(x)$ is convex on C if

$$\theta(x^2) - \theta(x^1) \geqq (x^2 - x^1)'\nabla_x\theta(x^1) \tag{1.4}$$

for every x^1 and x^2 in C.

The function $\theta(x)$ is said to be *quasi-convex* on C, [1], [2], if C is convex and if for every x^1 and x^2 in C,

$$\theta(x^2) \leqq \theta(x^1) \quad \text{implies} \quad \theta(\lambda x^1 + (1 - \lambda)x^2) \leqq \theta(x^1) \tag{1.5}$$

for every λ such that $0 \leqq \lambda \leqq 1$. Equivalently, $\theta(x)$ is quasi-convex on C if

$$\theta(x^2) \leqq \theta(x^1) \quad \text{implies} \quad (x^2 - x^1)'\nabla_x\theta(x^1) \leqq 0. \tag{1.6}$$

The function $\theta(x)$ is said to be *strictly quasi-convex* on C, [3], [5], if C is convex and if for every x^1 and x^2 in C, $x^1 \neq x^2$,

$$\theta(x^2) < \theta(x^1) \quad \text{implies} \quad \theta(\lambda x^1 + (1 - \lambda)x^2) < \theta(x^1) \tag{1.7}$$

for every λ such that $0 < \lambda < 1$. It has been shown [5] that every lower semicontinuous strictly quasi-convex function is quasi-convex but not conversely.

In the next section we shall give some properties of pseudo-convex functions and show how these properties can be used to generalize some previous results of mathematical programming, duality theory and stability theory of ordinary differential equations. Theorem 1 generalizes the Arrow-Enthoven version [1, Theorem 1] of the Kuhn-Tucker differential sufficient optimality conditions for a mathematical programming problem. Theorem 2 gives a generalization of Huard's converse duality theorem of mathematical programming [4, Theorem 2] and Theorem 3 generalizes a stability criterion for equilibrium points of nonlinear ordinary differential equations [8, Theorem 1].

2. Properties of pseudo-convex functions and applications. In this section we shall give some properties of pseudo-convex functions and some extensions of the results of mathematical programming and ordinary differential equations.

PROPERTY 0. *Let $\theta(x)$ be pseudo-convex on C. If $\nabla_x\theta(x^0) = 0$, then x^0 is a global minimum over C.*

Proof. For any x in C,

$$(x - x^0)'\nabla_x\theta(x^0) = 0,$$

and hence by (1.1),

$$\theta(x) \geqq \theta(x^0),$$

which establishes the property.

PROPERTY 1. *Let C be convex. If $\theta(x)$ is convex on C, then $\theta(x)$ is pseudo-convex in C, but not conversely.*

Proof. If $\theta(x)$ is convex on C, then by (1.4),

$$(x^2 - x^1)'\nabla_x\theta(x^1) \geqq 0 \qquad \text{implies} \qquad \theta(x^2) \geqq \theta(x^1),$$

which is precisely (1.1). That the converse is not necessarily true can be seen from the example

$$\theta(x) \equiv x + x^3, \qquad x \in E^1,$$

which is pseudo-convex on E^1 but not convex.[1]

PROPERTY 2. *Let C be convex. If $\theta(x)$ is pseudo-convex on C, then $\theta(x)$ is strictly quasi-convex (and hence quasi-convex) on C, but not conversely.*

Proof. Let $\theta(x)$ be pseudo-convex on C. We shall assume that $\theta(x)$ is not strictly quasi-convex on C and show that this leads to a contradiction. If $\theta(x)$ is not strictly quasi-convex on C, then it follows from (1.7) that there exist $x^1 \neq x^2$ in C such that

$$\theta(x^2) < \theta(x^1), \tag{2.1}$$

and

$$\theta(x) \geqq \theta(x^1), \tag{2.2}$$

for some $x \in L$, where

$$L = \{x \mid x = \lambda x^1 + (1 - \lambda)x^2, 0 < \lambda < 1\}. \tag{2.3}$$

Hence there exists an $\bar{x} \in L$ such that

$$\theta(\bar{x}) = \max_{x \in \bar{L}} \theta(x), \tag{2.4}$$

where

$$\bar{L} = L \cup \{x^1, x^2\}. \tag{2.5}$$

Now define

$$f(\lambda) = \theta((1 - \lambda)x^1 + \lambda x^2), \qquad 0 \leqq \lambda \leqq 1. \tag{2.6}$$

Hence

$$\theta(\bar{x}) = f(\bar{\lambda}), \tag{2.7}$$

where

$$\bar{x} = (1 - \bar{\lambda})x^1 + \bar{\lambda}x^2, \qquad 0 < \bar{\lambda} < 1. \tag{2.8}$$

[1] To see that $x + x^3$ is pseudo-convex, note that $\nabla_x\theta(x) = 1 + 3x^2 > 0$. Hence $(x - x^0)'\nabla_x\theta(x^0) \geqq 0$ implies that $x \geqq x^0$ and $x^3 \geqq (x^0)^3$, and thus

$$\theta(x) - \theta(x^0) = (x + x^3) - (x^0 + (x^0)^3) \geqq 0.$$

We have from (2.4) through (2.7) that $f(\lambda)$ achieves its maximum at $\bar{\lambda}$. Hence it follows by the differentiability of $\theta(x)$ and the chain rule that

$$(x^2 - x^1)'\nabla_x \theta(\bar{x}) = \frac{df(\bar{\lambda})}{d\lambda} = 0. \tag{2.9}$$

Since

$$x^2 - \bar{x} = x^2 - (1 - \bar{\lambda})x^1 - \bar{\lambda}x^2 = (1 - \bar{\lambda})(x^2 - x^1), \tag{2.10}$$

it follows from (2.9) and (2.10) and the fact that $\bar{\lambda} < 1$, that

$$(x^2 - \bar{x})'\nabla_x\theta(\bar{x}) = 0. \tag{2.11}$$

But by the pseudo-convexity of $\theta(x)$, (2.11) implies that

$$\theta(x^2) \geqq \theta(\bar{x}). \tag{2.12}$$

Hence from (2.1) and (2.12),

$$\theta(x^1) > \theta(\bar{x}),$$

which contradicts (2.4). Hence $\theta(x)$ must be strictly quasi-convex on C.

That the converse is not necessarily true can be seen from the example

$$\theta(x) \equiv x^3, \qquad x \in E^1,$$

which is strictly quasi-convex on E^1, but not pseudo-convex.

Property 3. *Let C be convex. If $\theta(x)$ is pseudo-convex on C, then every local minimum*[2] *is a global minimum.*

Proof. By Property 2, $\theta(x)$ is strictly quasi-convex on C. Now if $\bar{x}$ is a local minimum, then

$$\theta(\bar{x}) \leqq \theta(x) \qquad \text{for every} \quad x \in N(\bar{x}) \cap C, \tag{2.13}$$

where $N(\bar{x})$ is some neighborhood of $\bar{x}$. Let x be any point in C, but not in $N(\bar{x}) \cap C$. Then there exists a $\bar{\lambda}$, $0 < \bar{\lambda} < 1$, such that

$$\tilde{x} = ((1 - \bar{\lambda})\bar{x} + \bar{\lambda}x) \in N(\bar{x}) \cap C.$$

Now if $\theta(x) < \theta(\bar{x})$, then by the strict quasi-convexity of $\theta(x)$,

$$\theta(\bar{x}) > \theta(\tilde{x}),$$

which contradicts (2.13). Hence $\theta(x) \geqq \theta(\bar{x})$, which proves Property 3.

Theorem 1. *Let $\theta(x)$, $g_1(x)$, $\cdots$, $g_n(x)$ be differentiable functions on E^m. Let C be a convex set in E^m and $\theta(x)$ be pseudo-convex on C and $g_1(x)$, $\cdots$, $g_n(x)$ be quasi-convex on C. If there exist an $x^0 \in C$ and $y^0 \in E^n$ satisfy-*

[2] A local minimum is an $\bar{x} \in C$ such that $\theta(\bar{x}) \leqq \theta(x)$ for all $x \in N(\bar{x}) \cap C$, where $N(\bar{x})$ is some neighborhood of $\bar{x}$.

ing the Kuhn-Tucker differential conditions [7], *namely,*

$$\nabla_x\theta(x^0) + \nabla_x\sum_{i=1}^{n} y_i^0 g_i(x^0) = 0, \tag{2.14}$$

$$\sum_{i=1}^{n} y_i^0 g_i(x^0) = 0, \tag{2.15}$$

$$g_i(x^0) \leqq 0, \qquad i = 1, \cdots, n, \tag{2.16}$$

$$y_i^0 \geqq 0, \qquad i = 1, \cdots, n, \tag{2.17}$$

then

$$\theta(x^0) = \min_{x\in C}\{\theta(x) \mid g_i(x) \leqq 0, i = 1, \cdots, n\}. \tag{2.18}$$

Proof. The proof is similar to part of the proof of [1, Theorem 1]. Let

$$I = \{i \mid g_i(x^0) < 0\}.$$

Hence $g_i(x^0) = 0$ for $i \notin I$. From (2.15), (2.16) and (2.17) it follows that

$$y_i^0 = 0 \quad \text{for} \quad i \in I. \tag{2.19}$$

Let

$$R = \{x \mid g_i(x) \leqq 0, i = 1, 2, \cdots, n, x \in C\}.$$

Then $g_i(x) \leqq g_i(x^0)$ for $i \notin I$, $x \in R$. Hence by the quasi-convexity of the g_i's on R it follows from (1.6) that

$$(x - x^0)'\nabla_x g_i(x^0) \leqq 0 \quad \text{for} \quad i \notin I, x \in R. \tag{2.20}$$

Hence by (2.20) and (2.17) we have that

$$(x - x^0)'\nabla_x \sum_{i\notin I} y_i^0 g_i(x^0) \leqq 0 \quad \text{for} \quad x \in R, \tag{2.21}$$

and from (2.19) we have

$$(x - x^0)'\nabla_x \sum_{i\in I} y_i^0 g_i(x^0) = 0 \quad \text{for} \quad x \in R. \tag{2.22}$$

Hence (2.21) and (2.22) imply

$$(x - x^0)'\nabla_x \sum_{i=1}^{n} y_i^0 g_i(x^0) \leqq 0 \quad \text{for} \quad x \in R,$$

which in turn implies, by (2.14), that

$$(x - x^0)'\nabla_x \theta(x^0) \geqq 0 \quad \text{for} \quad x \in R. \tag{2.23}$$

But by the pseudo-convexity of $\theta(x)$ on R, (2.23) implies that

$$\theta(x) \geqq \theta(x^0) \quad \text{for} \quad x \in R.$$

For the case when the set I is empty, the above proof is modified by deleting (2.19), (2.22) and references thereto. For the case when $I = \{1, 2, \cdots, n\}$ the above proof is modified by deleting that part of the proof *between* (2.19) and (2.22) and references thereto.

It should be noted here that the above theorem is indeed a generalization of Arrow and Enthoven's result [1, Theorem 1]. Every case covered there is covered by the above theorem, but not conversely. An example of a case not covered by Arrow and Enthoven is the following one:

$$\min_{x \in E^1} \{-e^{-x^2} \mid -x \leqq 0\}.$$

Another application of pseudo-convex functions may be found in duality theory. Consider the primal problem

$$\text{(PP)} \qquad \min_{x \in E^m} \{\theta(x) \mid g(x) \leqq 0\},$$

where $\theta(x)$ is a scalar function on E^m and $g(x)$ is an $n \times 1$ vector function on E^m. For the above problem Wolfe [10] has defined the dual problem as

$$\text{(DP)} \qquad \max_{x \in E^m, y \in E^n} \{\psi(x, y) \mid \nabla_x \psi(x, y) = 0, y \geqq 0\},$$

where

$$\psi(x, y) \equiv \theta(x) + y'g(x).$$

Under appropriate conditions Wolfe has shown [10, Theorem 2] that if x^0 solves (PP), then x^0 and some y^0 solve (DP). Conversely, under somewhat stronger conditions, Huard [4, Theorem 2] showed that if (x^0, y^0) solves (DP), then x^0 solves (PP). Both Wolfe and Huard required, among other things, that $\theta(x)$ and the components of $g(x)$ be convex. We will now show that Huard's theorem can be extended to the case where $\theta(x)$ is pseudo-convex and the components of $g(x)$ are quasi-convex, and that Wolfe's theorem is not amenable to such an extension.

THEOREM 2. (*Strict converse duality theorem*)[3]. *Let $\theta(x)$ be a pseudo-convex function on E^m and let the components of $g(x)$ be differentiable quasi-convex functions on E^m.*

(a) *If (x^0, y^0) solves* (DP) *and $\psi(x, y^0)$ is twice continuously differentiable with respect to x in a neighborhood of x^0, and if the Hessian of $\psi(x, y^0)$ with respect to x is nonzero at x^0, then x^0 solves* PP.

(b) *Let x^0 solve* (PP) *and let $g(x) \leqq 0$ satisfy the Kuhn-Tucker constraint qualification* [7]. *It does not necessarily follow that x^0 and some y^0 solve* (DP).

Proof. (a) The assumption that the Hessian of $\psi(x, y^0)$ with respect to x is nonzero at x^0 insures the validity of the following Kuhn-Tucker neces-

[3] For the difference between "duality" and "strict duality," the reader is referred to [9].

sary conditions for some $v^0 \in E^m$:

$$\nabla_x\psi(x^0, y^0) + \nabla_x v^{0\prime}\nabla_x\psi(x^0, y^0) = 0,$$
$$\nabla_y\psi(x^0, y^0) + \nabla_y v^{0\prime}\nabla_x\psi(x^0, y^0) \leqq 0,$$
$$y^{0\prime}\nabla_y\psi(x^0, y^0) + y^{0\prime}\nabla_y v^{0\prime}\nabla_x\psi(x^0, y^0) = 0,$$
$$y^0 \geqq 0,$$
$$\nabla_x\psi(x^0, y^0) = 0.$$

The first and last equations above, together with the assumption that the Hessian of $\psi(x, y^0)$ is nonzero at x^0, imply that $v^0 = 0$. Hence the above necessary conditions become:

$$\nabla_x\psi(x^0, y^0) = 0,$$
$$\nabla_y\psi(x^0, y^0) = g(x^0) \leqq 0,$$
$$y^{0\prime}\nabla_y\psi(x^0, y^0) = y^{0\prime}g(x^0) = 0,$$
$$y^0 \geqq 0.$$

But from Theorem 1, with $C = E^m$, these conditions are sufficient for x^0 to be a solution of (PP).

(b) This part of the theorem will be established by means of the following counter-example:

(PP1) $$\min_{x\in E^1}\{-e^{-x^2} \mid -x + 1 \leqq 0\},$$

(DP1) $$\max_{x\in E^1, y\in E^1}\{-e^{-x^2} - yx + y \mid 2xe^{-x^2} - y = 0, y \geqq 0\}.$$

The solution of (PP1) is obviously $x^0 = 1$, whereas (DP1) has no maximum solution but has a zero supremum.

Finally, we give an application of pseudo-concavity outside the realm of of mathematical programming. In particular, we extend a stability criterion for equilibrium points of ordinary differential equations [8, Theorem 1].

THEOREM 3. (*Stability criterion*). *Let*

$$\dot{x} = f(t, x)$$

be a system of ordinary differential equations, where x and f are m-dimensional vectors and $0 \leqq t < \infty$. Let $f(t, x)$ be continuous in the (x, t) space and let $f(t, 0) = 0$ for $0 \leqq t < \infty$, so that $x = 0$ is an equilibrium point. If $x'f(t, x)$ is a pseudo-concave function of x on E^m for $0 \leqq t < \infty$, then $x = 0$ is a stable equilibrium point.

Proof. Consider the Lyapunov function

$$V(x, t) = \tfrac{1}{2}x'x,$$

which is obviously positive definite. It follows that for $0 \leqq t < \infty$,

$$\dot{V} = x'\dot{x} = x'f(t, x) \leqq 0,$$

where the last inequality follows from the pseudo-concavity of $x'f(t, x)$ in x and the fact that $f(t, 0) = 0$. Hence by Lyapunov's stability theorem [6], $x = 0$ is a stable equilibrium point.

It should be noted that the above proof would not go through had we merely required that $x'f(t, x)$ be quasi-concave instead of pseudo-concave.

3. Remarks on pseudo-convex functions. Properties 1 and 2 and the fact that every differentiable strictly quasi-convex function is also quasi-convex [5] establish a hierarchy among differentiable functions that is depicted in Fig. 1. In other words, if we let S_1, S_2, S_3, and S_4 represent the sets of all differentiable functions defined on a *convex* set C in E^m that are, respectively, convex, pseudo-convex, strictly quasi-convex, and quasi-convex, then

$$S_1 \subset S_2 \subset S_3 \subset S_4 .$$

Functions belonging to S_1, S_2, or S_3 share the property that a local minimum is a global minimum. Functions belonging to S_4 do not necessarily have this property. The Kuhn-Tucker differential conditions are sufficient for optimality, (see (2.18)), provided that $g_i(x)$, $i = 1, \cdots, n$ belong to S_4 and $\theta(x)$ belongs to S_1 or S_2, but not if $\theta(x)$ belongs to S_3 or S_4. It seems that the pseudo-convexity of $\theta(x)$ and the quasi-convexity of $g_i(x)$ are the weakest conditions that can be imposed so that relations (2.14) to (2.17) are sufficient for optimality.

There does not seem to be a simple extension of the concept of pseudo-convexity to nondifferentiable functions. This may be due to the fact that pseudo-convexity eliminates inflection points, and such points are easily described by derivatives, but not otherwise.

Finally, it should be remarked that the convexity of the set C is inherent in the definition of quasi-convexity. In contrast, the convexity of C is not needed in the definition of pseudo-convexity. Thus, without the convexity of C, we may have a pseudo-convex function that is not quasi-convex. For example, over the nonconvex set

$$C = \{x \mid x \in E^1, x \neq 0\},$$

the function

$$\theta(x) = \begin{cases} x & \text{for} \quad x < 0, \\ x + 1 & \text{for} \quad x > 0, \end{cases}$$

is pseudo-convex but obviously not quasi-convex, since C is nonconvex.

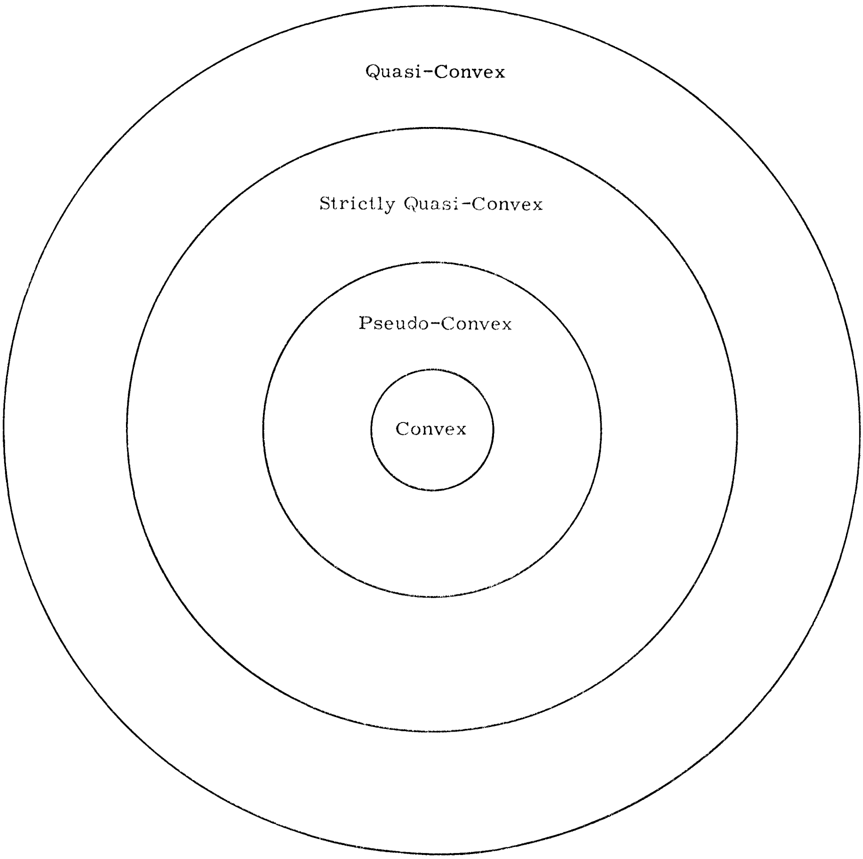

FIG. 1

4. Acknowledgement. I am indebted to my colleagues, S. Karamardian and J. Ponstein, for stimulating discussions on this paper.

REFERENCES

[1] K. J. ARROW AND A. C. ENTHOVEN, *Quasi-concave programming*, Econometrica, 29 (1961), pp. 779–800.

[2] C. BERGE, *Topological Spaces*, Macmillan, New York, 1963.

[3] M. A. HANSON, *Bounds for functionally convex optimal control problems*, J. Math. Anal. Appl., 8 (1964), pp. 84–89.

[4] P. HUARD, *Dual programs*, IBM J. Res. Develop., 6 (1962), pp. 137–139.

[5] S. KARAMARDIAN, *Duality in mathematical programming*, forthcoming.

[6] N. N. KRASOVSKII, *Stability of Motion*, Stanford University Press, Stanford, California, 1963.

[7] H. W. KUHN AND A. W. TUCKER, *Nonlinear programming*, Proceedings of the

Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, 1951, pp. 481–492.

[8] O. L. MANGASARIAN, *Stability criteria for nonlinear ordinary differential equations*, this Journal, 1 (1963), pp. 311–318.

[9] O. L. MANGASARIAN AND J. PONSTEIN, *Minmax and duality in nonlinear programming*, International Symposium on Mathematical Programming, London, 1964; J. Math. Anal. Appl., to appear.

[10] P. WOLFE, *A duality theorem for nonlinear programming*, Quart. Appl. Math., 19 (1961), pp. 239–244.