Mathematical Programming 1 (1971) 95–101. North-Holland Publishing Company

ON PSEUDO-CONVEX FUNCTIONS OF NONNEGATIVE VARIABLES *,**

Richard W. COTTLE and Jacques A. FERLAND
Stanford University, Stanford, California, U.S.A.

Received 20 October 1970
Revised manuscript received 18 February 1971

1. Introduction

In this paper, we prove the following conjecture stated by Bela Martos [6]:

> If the nonconvex quadratic function $\phi(x) = \frac{1}{2} x^{\mathrm{T}} D x + c^{\mathrm{T}} x$ is quasi-convex on the nonnegative orthant, it is pseudo-convex on the nonnegative orthant provided $c \neq 0$.

This is an extension of the results given earlier in [2], [4] and [5]. While our demonstration uses only ideas given in the latter papers, another proof can be extracted from the work of Arrow and Enthoven [1] as we shall show.

In section 2, we review briefly the general definitions of quasi-convexity and pseudo-convexity. We also recall there the definitions of positive subdefiniteness and strict positive subdefiniteness introduced by Martos in [4]. The results on pseudo-convex quadratic functions are summarized in section 3. Martos' conjecture is restated and proved in section 4. A second proof is given in section 5.

2. Basic definitions

A real-valued function ϕ defined on a convex set S is *quasi-convex* if and only if

* This paper was presented at the 7th Mathematical Programming Symposium 1970, The Hague, The Netherlands.

** Research was partially supported by Office of Naval Research, Contract N-00014-67-A0112-0011; National Science Foundation, Grant GP 9329. The second author was also supported by Hydro-Quebec and the University of Montreal.

$$\{x \in S | \phi(x) \leqslant \xi\} \text{ is convex for all } \xi \in E^1 .$$

For differentiable functions, the following "gradient-type inequality" is equivalent. For all $x, y \in S$

$$\phi(y) \leqslant \phi(x) \quad \text{implies} \quad \nabla\phi(x)(y-x) \leqslant 0. \tag{1}$$

A real-valued differentiable function ϕ is *pseudo-convex* on the set S if and only if for all $x, y \in S$.

$$\nabla\phi(x)(y-x) \leqslant 0 \quad \text{implies} \quad \phi(y) \geqslant \phi(x) . \tag{2}$$

Differentiable convex functions are pseudo-convex and pseudo-convex functions are quasi-convex [3]. In order to avoid treating convex functions as trivial special cases in what we are considering here, we speak of "merely quasi-convex (pseudo-convex) functions."

Another convenient term is "proper quadratic function" meaning an expression of the form $\frac{1}{2} x^{\mathrm{T}} Dx + c^{\mathrm{T}} x$ in which $c \neq 0$.

Recall that a nonzero, nonnegative vector x is called semipositive; we denote this by writing $x \geq 0$ (rather than $x \geqslant 0$ which means x is nonnegative and possibly 0). Naturally, x is seminegative ($x \leq 0$) if and only if $-x \geq 0$. The same kind of terminology applies to real matrices. For example, $D \leq 0$ means that D is nonpositive (entry-by-entry) but not the zero matrix.

In [4], Martos has identified a class of real symmetric matrices D and corresponding quadratic forms $\psi(x) = x^{\mathrm{T}} Dx$ called *positive subdefinite*. Their defining property is

$$x^{\mathrm{T}}Dx < 0 \quad \text{implies} \quad Dx \geq 0 \quad \text{or} \quad Dx \leq 0. \tag{3}$$

Moreover, the quadratic form ψ is *strictly positive subdefinite* if and only if

$$x^{\mathrm{T}}Dx < 0 \quad \text{implies} \quad Dx > 0 \quad \text{or} \quad Dx < 0. \tag{4}$$

It is evident that positive semi-definite quadratic forms are strictly positive subdefinite, and strictly positive subdefinite quadratic forms are positive subdefinite. Thus, in order to exclude the positive semi-definite quadratic forms, Martos inserts the word "merely" before "positive subdefinite".

3. Pseudo-convex quadratic functions

This section is a summary of results on pseudo-convexity for quadratic functions. We shall omit all the proofs since they can be found in [2], [4], or [5]. We also refer the reader to [2] and [4] for a matrix-theoretic characterization of merely positive subdefinite and strictly merely positive subdefinite matrices.

For the special case of a quadratic form, Martos gives the following criterion:

Theorem [4, Theorem 5]: The quadratic form $\psi(x) = x^{\mathrm{T}}Dx$ is pseudoconvex on the semipositive orthant, $E^n_+\backslash 0$, if and only if it is strictly positive subdefinite. (5)

In a subsequent paper, Martos studies quadratic functions and proves the following

Theorem [5, Theorem 3]: If $\phi(x) = \frac{1}{2}x^{\mathrm{T}}Dx + c^{\mathrm{T}}x$ is merely quasi-convex on the nonnegative orthant, and if the matrix $\begin{bmatrix} D & c \\ c^{\mathrm{T}} & 0 \end{bmatrix}$ has no row of zeros, it is merely pseudo-convex on the semi-positive orthant. (6)

Finally, the following result was shown by the authors.

Theorem [2, Theorem 6]: Let $\phi(x) = \frac{1}{2}x^{\mathrm{T}}Dx + c^{\mathrm{T}}x$ be a quadratic function on E^n such that $\begin{bmatrix} D & c \\ c^{\mathrm{T}} & 0 \end{bmatrix}$ contains no row of zeros. (7)
Then ϕ is pseudo-convex on $E^n_+\backslash 0$ if and only if the quadratic form

$$\psi(x, \xi) = \frac{1}{2} \begin{bmatrix} x \\ \xi \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} D & c \\ c^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} x \\ \xi \end{bmatrix}$$

is pseudo-convex on $E^{n+1}_+\backslash 0$ (or equivalently, strictly merely positive subdefinite).

4. Proof of Martos' conjecture

With this background, we can proceed to establish the truth of Martos' conjecture. First, we have to point out the similarity between [5, Theorem 3] and the statement of the conjecture. The condition prohibiting a row of zeros in the bordered matrix $\begin{bmatrix} D & c \\ c^{\mathrm{T}} & 0 \end{bmatrix}$ is replaced by $c \neq 0$, and the pseudo-convexity is extended to the nonnegative (as opposed to the semipositive) orthant.

The following property will be useful in proving the main result.

Lemma. Let the real symmetric matrix $\begin{bmatrix} D & c \\ c^{\mathrm{T}} & 0 \end{bmatrix}$ be merely positive subdefinite where D is a nonzero matrix of order n and c is a nonzero column vector. For any vector $v \in E^n$,

$$v^{\mathrm{T}} D v < 0 \quad \text{implies} \quad c^{\mathrm{T}} v \neq 0.$$

Proof. Since $D \neq 0$, it can be expressed without loss of generality in the form

$$D = \begin{bmatrix} \bar{D} & 0 \\ 0 & 0 \end{bmatrix}$$

where $\bar{D}$ is a symmetric matrix of order $\bar{n} \leqslant n$ having no row of zeros. As a consequence of [2, Theorem 7], we may write

$$\begin{bmatrix} D & c \\ c^{\mathrm{T}} & 0 \end{bmatrix} = \begin{bmatrix} \bar{D} & 0 & \bar{c} \\ 0 & 0 & 0 \\ \bar{c}^{\mathrm{T}} & 0 & 0 \end{bmatrix}, \qquad \bar{c} \neq 0.$$

By the inheritance property of positive subdefiniteness, the submatrix $\begin{bmatrix} \bar{D} & \bar{c} \\ \bar{c}^{\mathrm{T}} & 0 \end{bmatrix}$ is positive subdefinite. Since $\bar{D}$ has no row of zeros and $\bar{c} \neq 0$, it follows that $\begin{bmatrix} \bar{D} & \bar{c} \\ \bar{c}^{\mathrm{T}} & 0 \end{bmatrix}$ is strictly merely positive subdefinite.

Now, for contradiction, suppose there exists a vector $v \in E^n$ with the property $v^{\mathrm{T}} D v < 0$ and $c^{\mathrm{T}} v = 0$. By definition of $\bar{D}$ and $\bar{c}$,

$$v^{\mathrm{T}} D v = \bar{v}^{\mathrm{T}} \bar{D} \bar{v} < 0 \quad \text{and} \quad c^{\mathrm{T}} v = \bar{c}^{\mathrm{T}} \bar{v} = 0,$$

where $\bar{v}$ is the appropriate subvector of v. Hence

$$0 > \tfrac{1}{2} \bar{v}^{\mathrm{T}} \bar{D} \bar{v} + \bar{c}^{\mathrm{T}} \bar{v} = \frac{1}{2} \begin{bmatrix} \bar{v} \\ 1 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \bar{D} & \bar{c} \\ \bar{c}^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \bar{v} \\ 1 \end{bmatrix}.$$

Since $\begin{bmatrix} \bar{D} & \bar{c} \\ \bar{c}^{\mathrm{T}} & 0 \end{bmatrix}$ is strictly merely positive subdefinite, it follows from (4) that $\bar{c}^{\mathrm{T}}\bar{v} > 0$ or $\bar{c}^{\mathrm{T}}\bar{v} < 0$. Either of these contradicts the hypothesis that $\bar{c}^{\mathrm{T}}\bar{v} = 0$, and the proof is complete.

The main result will now follow by a slight modification of Martos' proof of (6).

Theorem 1. If the proper quadratic function $\phi(x) = \frac{1}{2}x^{\mathrm{T}}Dx + c^{\mathrm{T}}x$ is merely quasi-convex on the nonnegative orthant then it is pseudo-convex on the nonnegative orthant.

Proof. We must verify

$$\begin{aligned} &(x^{\mathrm{T}}D + c^{\mathrm{T}})(y - x) \geqslant 0 \\ &\text{implies} \quad \frac{1}{2}y^{\mathrm{T}}Dy + c^{\mathrm{T}}y \geqslant \frac{1}{2}x^{\mathrm{T}}Dx + c^{\mathrm{T}}x \end{aligned} \tag{8}$$

for all $x, y \in E^n_+$.

If $(x^{\mathrm{T}}D + c^{\mathrm{T}})(y - x) \geqslant 0$ and $(y - x)^{\mathrm{T}}D(y - x) \geqslant 0$, it follows directly that $\phi(y) \geqslant \phi(x)$, as required. Hence we assume $(y - x)^{\mathrm{T}}D(y - x) < 0$. Since ϕ is quasi-convex on E^n_+, it follows from [5, Theorem 1] that

$$\begin{bmatrix} D \\ c^{\mathrm{T}} \end{bmatrix} (y - x) \geqslant 0 \quad \text{or} \quad \begin{bmatrix} D \\ c^{\mathrm{T}} \end{bmatrix} (y - x) \leqslant 0.$$

First, suppose $\begin{bmatrix} D \\ c^{\mathrm{T}} \end{bmatrix} (y - x) \geqslant 0$. Since x and y are both nonnegative, we have $(x + y)^{\mathrm{T}}D(y - x) \geqslant 0$. Consequently,

$$\phi(y) - \phi(x) = \frac{1}{2}(x + y)^{\mathrm{T}}D(y - x) + c^{\mathrm{T}}(y - x) \geqslant 0.$$

We are left with the case $(y - x)^{\mathrm{T}}D(y - x) < 0$ and $\begin{bmatrix} D \\ c^{\mathrm{T}} \end{bmatrix} (y - x) \leqslant 0$. Since the hypothesis of the lemma are satisfied, it follows that $c^{\mathrm{T}}(y - x) < 0$. Because $x \in E^n_+$, we now have $x^{\mathrm{T}}D(y - x) + c^{\mathrm{T}}(y - x) < 0$ which means that (8) is true by default. This completes the proof.

Using Mangasarian's general result [3, Property 2], for the converse of the Theorem, we can state the

Corollary. On E^n_+, a proper quadratic function is pseudo-convex if and only if it is quasi-convex.

This result extends what we already know from [2] and [5] about characterizing pseudo-convex quadratic functions. From the standpoint of quadratic programming, the significance of this Corollary is that there is a (finitely testable) class of nonconvex objective functions for which Kuhn-Tucker stationary points yield global minima.

5. Another proof of Martos' conjecture

At a time before the property of pseudo-convexity had been explicitly identified, Arrow and Enthoven [1] gave a set of various conditions under which a quasi-convex function is pseudo-convex. Although the preceding observation may have a folkloric status, we are unaware of its presence in the literature.

Theorem 2. If f is a differentiable quasi-convex function on E^n_+, then f is pseudo-convex at any $x \in E^n_+$ where (i) $\nabla f(x) \neq 0$, and (ii) $\nabla^2 f(x)$ exists.

Proof. Let x be a point in E^n_+ satisfying (i) and (ii). Define $g(x) = \nabla f(x)(x - x)$ and consider the mathematical programming problem

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

$$\text{subject to } g(x) \geqslant 0$$

$$x \geqslant 0.$$

The Kuhn-Tucker conditions for this linearly-constrained problem are

$$\nabla f(x) - \lambda \nabla g(x) \geqslant 0$$

$$[\nabla f(x) - \lambda \nabla g(x)]x = 0$$

$$g(x) \geqslant 0$$

$$\lambda g(x) = 0$$

$$x \geqslant 0$$

$$\lambda \geqslant 0$$

and in virtue of the definition of g, these are obviously satisfied by

$(x, \lambda) = (\bar{x}, 1)$. Hence condition (c) of [1, Theorem 1] is fulfilled, and it follows that x is an optimal solution of (9). This means that for all $x \in E_+^n$

$$\nabla f(\bar{x})(x - \bar{x}) \geqslant 0 \quad \text{implies} \quad f(x) \geqslant f(\bar{x})$$

as required for the pseudo-convexity of f at $\bar{x}$.

Remark. The idea for writing down the program (9) can be found in the Arrow-Enthoven proof of [1, Theorem 5].

Now, for a quick alternate proof of Martos' conjecture, note that if $f(x) = \frac{1}{2}x^\mathrm{T}Dx + c^\mathrm{T}x$ is a merely quasi-convex proper quadratic function on E_+^n, we must have $D \leqslant 0$ and $c \leqslant 0$. Indeed, $D \lesseqgtr 0$ and $c \lesseqgtr 0$ follow directly from [5, Theorem 2] or [2, Theorem 5]; then $D \leqslant 0$ and $c \leqslant 0$ follow from the nonconvexity and the properness of f, respectively. Consequently, $\nabla f(x) = x^\mathrm{T}D + c^\mathrm{T} \leqslant 0$ for all $x \in E_+^n$. Clearly, Theorem 2 applies and f is pseudo-convex on E_+^n.

References

[1] K.J.Arrow and A.C.Enthoven, "Quasi-concave programming," *Econometrica* 29 (1961) 779–800.

[2] R.W.Cottle and J.A.Ferland, "Matrix-theoretic criteria for the quasi-convexity and pseudo-convexity of quadratic functions," Technical Report No. 70-6, Operations Research House, Stanford University (April 1970); to appear in: *Linear Algebra and Its Applications.*

[3] O.L.Mangasarian, "Pseudo-convex functions," *SIAM Journal on Control* 3 (1965) 281–290.

[4] B.Martos, "Subdefinite matrices and quadratic forms," *SIAM Journal on Applied Mathematics* 17 (1969) 1215–1223.

[5] B.Martos, "Quadratic programming with a quasi-convex objective function," *Operations Research* 19 (1971) 87–97.

[6] B.Martos, Private communication (June 24, 1970).