JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 35, 58–66 (1971)

Convex Inversion*

A. WARBURTON AND W. T. ZIEMBA

Faculty of Commerce and Business Administration, University of British Columbia Vancouver, B. C., Canada

April 8, 1970

Submitted by Richard Bellman

1. INTRODUCTION

This paper considers the problem of determining when the n equations $f_i(x, y) = 0$, $i = 1,..., n$, $x \in R^n$, $y \in R^m$ have a globally unique solution $x_i = h_i(y)$ and h_i is convex. Conditions of this nature have application in a number of areas, and in particular, in dynamic programming (see Section 3 below).

Gale and Nikadô [6] have proved some global univalence theorems that can be applied to the problem. These theorems imply that the h_i exist and are unique C^1 functions. Under the assumption that the h_i exist and are C^1, we present necessary and sufficient conditions on the f_i for the convexity of the h_i. Additionally, sufficient conditions on the f_i are given that insure that the h_i exist, are C^1, and are convex functions.

The use of these results is illustrated via analyses of a reasonably general class of dynamic programming problems and a specialized Leontief production system.

Let Ω be a convex subset of R^k. A function $g : \Omega \to R$ is convex on Ω if for all points x^0, $x^1 \in \Omega$ and all $\lambda \in [0, 1]$,

$$g\{\lambda x^0 + (1 - \lambda) x^1\} \leqslant \lambda g(x^0) + (1 - \lambda) g(x^1).$$

If g is differentiable on Ω, an equivalent definition is that g is convex on Ω iff $g(x^1) - g(x^0) \geqslant \nabla g(x^0)' (x^1 - x^0) \; \forall x^0, x^1 \in \Omega$, where ∇ denotes the gradient operator. The function g is said to be quasiconvex on Ω if $\{x \mid g(x) \leqslant \alpha\}$ is a convex set for all $\alpha \in R$, or equivalently if

$$g\{\lambda x^0 + (1 - \lambda) x^1\} \leqslant \max\{g(x^0), g(x^1)\} \qquad \forall \lambda \in [0, 1], x^0, x^1 \in \Omega.$$

* This research was supported in part by the National Research Council under constract NRC-A7147.

If g is differentiable on Ω, then g is quasiconvex on Ω iff

$$g(x^1) \leqslant g(x^0) \Rightarrow \nabla g(x^0)' (x^1 - x^0) \leqslant 0 \qquad \forall x^0, x^1 \in \Omega.$$

The function g is said to be concave or quasiconcave if $-g$ is convex or quasiconvex, respectively. These definitions are standard; for more detail see Mangasarian [8].

Let $f = (f_1, ..., f_n) : \Omega \to R^n$. The vector valued function f is defined to be convex on Ω iff each f_i is convex on Ω. Similarly, f is defined to be quasi-convex on Ω if each f_i is quasiconvex on Ω. This definition of quasiconvexity implies that if f is quasiconvex, then the set $\{x \mid x \in \Omega, f_i(x) \leqslant \alpha_i, i = 1, ..., n\}$ is convex for all $\alpha = (\alpha_1, ..., \alpha_n) \in R^n$. The converse does not hold, however.

Hence if f is differentiable it is convex iff $f_i(x^1) - f_i(x^0) \geqslant \nabla f_i(x^0)' (x^1 - x^0)$, $i = 1, ..., n$ and $\forall x^0, x^1 \in \Omega$ and it is quasiconvex iff

$$f_i(x^1) \leqslant f_i(x^0) \Rightarrow \nabla f_i(x^0)' (x^1 - x^0) \leqslant 0, \qquad i = 1, ..., n \quad \text{and} \quad \forall x^0, x^1 \in \Omega.$$

Let $f : R^n \times R^m \to R^n : (x, y) \to f(x, y)$ be a differentiable function. Denote by $Jf(x^0, y^0)$ the Jacobian matrix

$$\begin{bmatrix} Dx_1 f_1(x^0, y^0), ..., Dx_n f_1(x^0, y^0), & Dy_1 f_1(x^0, y^0), ..., Dy_m f_1(x^0, y^0) \\ \vdots \qquad\qquad \vdots & \vdots \qquad\qquad \vdots \\ Dx_1 f_n(x^0, y^0), ..., Dx_n f_n(x^0, y^0), & Dy_1 f_n(x^0, y^0), ..., Dy_m f_n(x^0, y^0) \end{bmatrix},$$

where $Dz_i f_j(x^0, y^0)$ denotes the partial derivative of f_j with respect to z_i evaluated at (x^0, y^0).

Let $J_x f(x^0, y^0)$ be the $n \times n$ matrix formed by the first n columns, and $J_y f(x^0, y^0)$ the $n \times m$ matrix formed by the last m columns, of $Jf(x^0, y^0)$, hence

$$Jf(x^0, y^0) = [J_x f(x^0, y^0), J_y f(x^0, y^0)].$$

Using the Jacobian notation, the criterion for quasiconvexity of a vector valued function becomes: $f = (f_1, ..., f_n) : \Omega \to R^n$ is quasiconvex on Ω iff $f(y^1) \leqslant f(y^0) \Rightarrow Jf(y^0)(y^1 - y^0) \leqslant 0 \ \forall y^0, y^1 \in \Omega$. Also f is convex on Ω iff $f(y^1) - f(y^0) \geqslant Jf(y^0)(y^1 - y^0) \ \forall y^0, y^1 \in \Omega$.

2. Global Inversion

The basic problem is to determine conditions on $f : R^n \times R^m \to R^n$, such that the equation $f(x, y) = 0$ has a globally unique convex solution $x = h(y)$; that is whenever $f(x, y) = 0$, then $x = h(y)$ with h convex and $f\{h(y), y\} = 0$.

The following lemma will be useful below:

LEMMA 1 [4]. *Suppose* $f : R^n \times R^m \to R^n$ *and that* $f(x^0, y^0) = 0$. *If* f *satisfies the conditions of the implicit function theorem, then the solution* $x = h(y)$ *of the equation* $f(x, y) = 0$, *unique in a neighborhood of* y^0, *has its Jacobian given by*

$$Jh(y^0) = -J_x^{-1}f(x^0, y^0)\, J_y f(x^0, y^0),$$

where the ij-th *element of* $Jh(y^0)$ *is* $Dy_j h_i(y^0)$, $i = 1,\ldots, n$, $j = 1,\ldots, m$, *and* $x^0 = h(y^0)$.

The following lemma provides sufficient conditions for h to be of class C^1. A function $f : \Omega \to R^n$ is said to be univalent if $x \neq y \Rightarrow f(x) \neq f(y)$ $\forall x, y \in \Omega$ (and hence f^{-1} exists, where $f^{-1}[f(x)] = x$).

LEMMA 2. *Let* $f : R^n \times R^m \to R^n : (x, y) \to f(x, y)$ *be of class* C^1 *on some open subset* $A \times B$ *of* $R^n \times R^m$. *Let* $S \equiv \{y \in B : f(x, y) = 0$ *for some* $x \in A\}$ *and suppose* $S \neq \phi$. *If the function* $g_y : A \to R^n : x \to f(x, y)$ *is univalent for each* $y \in S$, *then for each* $y \in S$, *there exists a unique* $h(y) \in A$ *such that* $f\{h(y), y\} = 0$, *and* h *is of class* C^1 *on* S *if* $|J_x f(x^0, y^0)| \neq 0$ *for any* $(x^0, y^0) \in h(S) \times S$.

Proof. Let $y \in S$, since g_y is univalent, there is exactly one $x \in A$ such that $f(x, y) = g_y(x) = 0$. Hence a unique function $h : S \to A$ is defined such that $f(h(y), y) = 0$. To show differentiability of h, let $y^0 \in S$ and let $x^0 = h(y^0)$. Then $f(x^0, y^0) = 0$ and $|J_x f(x^0, y^0)| \neq 0$. By the implicit function theorem, there exists a unique function ψ defined on a neighborhood $N \subset S$ of y^0 such that $x = \psi(y)$ and $f(\psi(y), y) = 0$ for all $y \in N$. But h also has this property. Hence the restriction of h to N is ψ. Since ψ is continuously differentiable at y^0, so is h.

A necessary and sufficient condition for the convexity of h is contained in the following theorem:

THEOREM 1. *Let* $f : R^n \times R^m \to R^n$ *be of class* C^1 *on a convex subset* $A \times B$ *of* $R^n \times R^m$. *Suppose that* $f(x, y) = 0$ *has a unique solution* $x = h(y)$ *of class* C^1 *on* S. *Then* h *is convex on* S *iff*

$$J_x^{-1}f(x^0, y^0)\, Jf(x^0, y^0)\,(x^1 - x^0, y^1 - y^0) \geqslant 0, \quad \forall (x^1, y^1), (x^0, y^0) \in h(S) \times S. \tag{1}$$

Proof. Let (x^0, y^0), $(x^1, y^1) \in h(S) \times S$. Now

$$J_x^{-1}f(x^0, y^0)\, Jf(x^0, y^0) = J_x^{-1}f(x^0, y^0)\,[J_x f(x^0, y^0), J_y f(x^0, y^0)].$$

By Lemma 1

$$J_x^{-1}f(x^0, y^0)\, J_y f(x^0, y^0) = -\, Jh(y^0).$$

Hence

$$J_x^{-1}f(x^0, y^0)\, Jf(x^0, y^0) = [I_n\,, -\, Jh(y^0)],$$

where I_n is the identity matrix of order n. Using this we obtain

$$J_x^{-1}f(x^0, y^0)\, Jf(x^0, y^0)\, (x^1 - x^0, y^1 - y^0)$$

$$= \begin{bmatrix} (x_1{}^1 - x_1{}^0) - \sum_{k=1}^{n} Dy_k h_1(y^0)\,(y_k{}^1 - y_k{}^0) \\ \vdots \qquad\qquad\qquad \vdots \\ (x_n{}^1 - x_n{}^0) - \sum_{k=1}^{n} Dy_k h_n(y^0)\,(y_k{}^1 - y_k{}^0) \end{bmatrix}.$$

But $x^1 = h(y^1)$, $x^0 = h(y^0)$ if (x^0, y^0), $(x^1, y^1) \in h(S) \times S$.

Also

$$\sum_{k=1}^{n} Dy_k h_j(y^0)\,(y_k{}^1 - y_k{}^0) = \nabla h_j(y^0)'\,(y^1 - y^0) \qquad \text{for } j = 1,\ldots, n.$$

Hence (1) may be written as

$$h(y^1) - h(y^0) - Jh(y^0)\,(y^1 - y^0).$$

By definition, this expression is $\geqslant 0$ iff h is convex on S.

Theorem 3 below presents conditions on f that imply that h exists and is convex. The result is utilized in the examples in Section 3 below. The proof utilizes a global univalence result due to Gale and Nikadô [6] and a property of Leontief matrices. If A is an $n \times n$ nonnegative matrix whose row sums do not exceed unity, then $I - A$ is said to be Leontief. Additionally, if $I - A$ is a P matrix, i.e., a matrix all of whose principal minor determinants are positive, then it is well known [5] that $(I - A)^{-1}$ exists and is nonnegative.

THEOREM 2 (Gale-Nikaidô) [6]. *Let D be a rectangular subset of R^k and $f: D \to R^k$ be a differentiable mapping. If $Jf(x)$ is a P matrix of Leontief type $\forall x \in D$ then f is univalent.*

THEOREM 3. *Suppose $f: D \times B \to R^n$ is quasiconcave and of class C^1 on $D \times B$, where D is a rectangular subset of R^n and B is an open convex subset of R^m. Suppose further that $J_x f(x, y)$ is a P matrix of Leontief type on $D \times S$, then h exists, is convex, and of class C^1 on S, if $S \neq \varnothing$.*

Proof. As in Lemma 2, consider the function $g_y : D \to R^n$ defined by $g_y(x) = f(x, y)$. Clearly $Jg_{y0}(x^0) = J_x f(x^0, y^0)$. Now $J_x f(x^0, y^0)$ is a P matrix of Leontief type $\forall x^0 \in D$, $y^0 \in S$. Hence by Theorem 2, g_y is univalent $\forall y \in S$. By Lemma 2, h exists and is of class C^1 on S since $| J_x f(x^0, y^0)|$ is nonzero on $h(S) \times S$. Now f is identically zero on $h(S) \times S$, hence by the quasiconcavity property $Jf(x^0, y^0)\,(x^1 - x^0, y^1 - y^0) \geqslant 0$, $\forall (x^0, y^0)$, $(x^1, y^1) \in h(S) \times S$. Since $J_x f(x^0, y^0)$ is a P matrix of Leontief type on $h(S) \times S$, it has a positive inverse there. Hence,

$$J_x^{-1} f(x^0, y^0)\, Jf(x^0, y^0)\,(x^1 - x^0, y^1 - y^0) \geqslant 0$$

and h is convex by Theorem 1.

3. Examples

The results of Section 2 may be used to demonstrate the concavity of the return functions in certain dynamic programming problems. Consider the T-period economic planning problem

$$\max_{x_1 \in K_1} \{u_1(x_1, y_1) + \max_{x_2 \in K_2} [u_2(x_2, y_2) + \cdots + \max_{x_T \in K_T} u_T(x_T, y_T)]\}, \tag{2}$$

subject to

$$\begin{aligned} H_1(x_1, y_1) &= 0, \\ H_2(x_1, x_2, y_1, y_2) &= 0, \\ &\vdots \\ H_T(x_1, \ldots, x_T, y_1, \ldots, y_T) &= 0, \end{aligned}$$

where u_t is the (additive) utility that accrues from policy decision vector x_t and state vector occurence y_t. The H_t represent the structural equations of the model used to describe the economic system and the K_t represent the sets of allowable policy decisions.

A function $v : R^n \to R^1$ is said to be nondecreasing if

$$x_i^1 \leqslant x_i^2 \forall i \Rightarrow v(x_1^1, \ldots, x_n^1) \leqslant v(x_1^2, \ldots, x_n^2).$$

Theorem 4. *Suppose that the K_t are convex sets, the y_t are defined on the rectangular sets $\Omega_t \subset R^{m_t}$, and that the u_t are concave and nondecreasing. The H_t are assumed to be continuously differentiable, quasiconvex and nonincreasing in $(x, \ldots, x_t, y_1, \ldots, y_t)$ and $J_{y_t} H_t(\cdot)$ is Leontief and is a P matrix on $\Omega_t \times M_t$, where*

$$M_t \equiv \{(x_1, \ldots, x_t, y_1, \ldots, y_{t-1}) \mid H_t(x_1, \ldots, x_t, y_1, \ldots, y_t) = 0, \textit{ for some } y_t \in \Omega_t\}.$$

Then (2) *is equivalent to the problem*

$$\max_{x_1 \in K_1} v_1(x_1), \tag{3}$$

where v_1 is concave and nondecreasing.

The proof of Theorem 4 will utilize the following three lemmas:

LEMMA 3. *Let $v(x) \equiv u[z_1(x),..., z_k(x)]$, where $v : C \to R$, $u : R^k \to R$, $z_i : C \to R$, $i = 1,..., k$, $C \subset R^n$. If C is a convex set, u and the z_i are concave and nondecreasing, then v is concave and nondecreasing in C.*

Proof. Concavity was proved in Ref. [1]. Let x^1, $x^2 \in C$, $x^2 \geqslant x^1$. Since z_i is nondecreasing $z_i(x^2) \geqslant z_i(x^1)\ \forall_i$, and since u is nondecreasing

$$v(x^2) \equiv u[z_1(x^2),..., z_k(x^2)] \geqslant u[z_1(x^1),..., z_k(x^1)] \equiv v(x^1).$$

Hence v is nondecreasing.

LEMMA 4. *Let $v(x) \equiv \max_{y \in K}\{w(x, y)\}$, where $w : C \times K \to R$, $v : C \to R$ and $C \subset R^n$, $K \subset R^m$ are convex sets. Suppose that for each $x \in C$ the maximum is obtained for some $y_0(x) \in K$. If w is a concave function and is nondecreasing in x for each fixed y, then v is concave and nondecreasing.*

Proof. Concavity was proved in [7]. Let x^1, $x^2 \in C$, $x^2 \geqslant x^1$. Since w is nondecreasing in x for each fixed $y \in K$, $w\{x^2, y\} \geqslant w\{x^1, y\}$. Hence

$$v(x^2) \equiv \max_{y \in K} w(x^2, y) \geqslant \max_{y \in K} w(x^1, y) \equiv v(x^1),$$

thus v is nondecreasing.

LEMMA 5. *The assumptions of Theorem 3 coupled with the additional assumption that f is nondecreasing in y, imply that h is a nonincreasing function.*

Proof. By Lemma 1, the Jacobian of h is given by

$$Jh(y^0) = -J_x^{-1}f(x^0, y^0)\, J_y f(x^0, y^0).$$

Since f is nondecreasing in y, $J_y f(x^0, y^0) \geqslant 0$. Also $J_x^{-1}f(x^0, y^0) \geqslant 0$. Hence $Jh(y^0) \leqslant 0$, i.e., h is nonincreasing.

Proof of Theorem 4. In period T under the assumptions the problem is

$$\begin{aligned}\max_{x_T \in K_T} & u_T[x_T, I_T(x_1, ..., x_T, y_1, ..., y_{T-1})] \\ &\equiv \max_{x_T \in K_T} w_T(x_1, ..., x_T, y_1, ..., y_{T-1}) \\ &\equiv v_{T-1}(x_1, ..., x_{T-1}, y_1, ..., y_{T-1}),\end{aligned}$$

where I_t exists and is concave and nondecreasing (by Theorem 3 and Lemma 5).

Now w_T is concave and nondecreasing by Lemma 3 and v_{T-1} is concave and nondecreasing by Lemma 4.

Suppose $v_t(x_1, ..., x_t, y_1, ..., y_t)$ is concave and nondecreasing. Then the problem in period t is

$$\begin{aligned}
&\max_{x_t \in K_t} [u_t(x_t, y_t) + v_t(x_1, ..., x_t, y_1, ..., y_t)] \\
&\quad = \max_{x_t \in K_t} [u_t\{x_t, I_t(x_1, ..., x_t, y_1, ..., y_{t-1})\} \\
&\qquad + v_t\{x_1, ..., x_t, y_1, ..., y_{t-1}, I_t(x_1, ..., x_t, y_1, ..., y_{t-1})\}] \\
&\quad \equiv \max_{x_t \in K_t} w_t(x_1, ..., x_t, y_1, ..., y_{t-1}) \equiv v_{t-1}(x_1, ..., x_{t-1}, y_1, ..., y_{t-1}),
\end{aligned}$$

where I_t exists and is concave and nondecreasing (by Theorem 3 and Lemma 5), and w_t is concave and nondecreasing by Lemma 3, and v_{t-1} is concave and nondecreasing by Lemma 4, and the fact that nonnegative sums of concave nondecreasing functions are concave and nondecreasing.

A more specialized example in which the hypotheses of Theorem 3 are satisfied is the following simple linear input-output production system. Consider an economy in which each of n goods $G_1, ..., G_n$ is produced by a single activity. Let $x \equiv (x_1, ..., x_n)$ and $p \equiv (p_1, ..., p_n)$ be output and price vectors, respectively, and a_{ij} be the amount of G_j which it is necessary to consume in order to produce one unit of G_i. The demand for G_j is $g_j(p_j) = b_j - d_j p_j$ ($d_j \geqslant 0$). The optimization problem is to determine nonnegative (normalized) prices p and output levels x to maximize the utility u of revenue $p'g(p)$ minus production costs $c(x)$. Since net production is $x - Ax$, output levels of $\{g_1(p_1), ..., g_n(p_n)\}$ require that (4) $x - Ax = g(p)$. Let $a(x, p) = x - Ax - g(p)$, which is quasiconcave in x and p by linearity, and assume that $J_x a(x, p) = I - A$ is a P matrix. Thus (4) has the solution $x = b(p) = (I - A)^{-1} g(p)$, where b is convex (since it is linear) in p. Suppose that u is concave and nondecreasing, that is the economy is risk averse and that c is convex meaning that marginal production costs are nondecreasing. The problem is to maximize

$$u\{p'g(p) - c[b(p)]\}, \tag{5}$$

subject to

$$p \geqslant 0, \qquad \sum_{i=1}^{n} p_i = 1, \qquad g(p) \geqslant 0.$$

Since $b(p)$ is convex, $-c[b(p)]$ is concave and since $p'g(p)$ is concave, (5) is a concave program in p.

4. Remarks

Theorem 3 generalizes for functions of class C^1 an earlier 1-dimensional result of Pierskalla [9] to the n-dimensional case.

Several results, e.g., Refs. [2] [6], other than Theorem 2, imply that f satisfies the conditions of Lemma 2; hence that h exists and is C^1. Moreover, in some of the above cases restrictions of the domain of f may be relaxed. For instance, if it is assumed that the symmetric part of $J_x f(x, y)$ is everywhere positive in $D \times S$, then h exists, is C^1, and D need only be convex, rather than rectangular.

In these cases, however, h may fail to be convex.

When the variables are separable in $f(x, y) = 0$ and a solution $x = h(y)$ exists then this solution has a simple form. Thus, if

$$f(x, y) = f_1(x) + f_2(y) = 0,$$

then

$$x = f^{-1}\{f_2(y)\} \qquad \text{and} \qquad h(y) = f^{-1}[f_2(y)].$$

In the linear case we have $Ax + By = 0$, hence $x = - A^{-1}By$.

A similar result to that in Theorem 4 may be obtained if the H_t are functions of random vectors, say, $\xi_1, \ldots, \xi_t$ and u_t depends upon ξ_t. Assume that the objective in each period t is to maximize the expected utility of state vector realizations, random vectors, and decisions made in $\tau = t, \ldots, T$. Replace $u_t(x_t, y_t)$ and $H_t(x_1, \ldots, x_t, y_1, \ldots, y_t)$ by $\tilde{u}_t(x_t, y_t, \xi_t)$ and $\tilde{H}_t(x_1, \ldots, x_t, y_1, \ldots, y_t, \xi_1, \ldots, \xi_t)$, where $\forall(\xi_1, \ldots, \xi_t) \in \Xi_1 \times \cdots \times \Xi_t$, $\tilde{u}_t$ and $\tilde{H}_t$ satisfy the assumptions imposed upon u_t and H_t, respectively. Using a similar proof one may reduce this problem to (3). When the H_t are linear this is similar to a result of Dantzig [3]. Note, however, that if any maxima or conditional expectations are unbounded from above, then so will v_1.

References

1. C. Berge, "Topological Spaces," MacMillan Book Co., New York, 1963.
2. M. Berger and M. Berger, "Perspectives in Nonlinearity," W. A. Benjamin Inc., New York, 1968.
3. G. B. Dantzig, "Linear Programming Under Uncertainty," *Management Sci.* 1 (1955), 197–206.
4. J. Dieudonné, "Foundations of Modern Analysis," Academic Press, New York, 1960.
5. D. Gale, "The Theory of Linear Economic Models," McGraw-Hill, New York, 1960.
6. D. Gale and H. Nikaidô, The Jacobian Matrix and Global Univalence of Mappings, *Math. Ann.* **159** (1965), 81–93.

7. D. L. IGLEHART, Capital Accumulation and Production for the Firm: Optimal dynamic policies, *Management Sci.* **12** (1965), 193–205.
8. O. L. MANGASARIAN, "Nonlinear Programming," McGraw-Hill Book Co., New York, 1969.
9. W. P. PIERSKALLA, Mathematical Programming with Increasing Constraint Functions, *Management Sci.* **15** (1969), 416–425.