MATHEMATICAL PROGRAMMING WITH INCREASING CONSTRAINT FUNCTIONS*,†

WILLIAM P. PIERSKALLA

Southern Methodist University

The mathematical programming problem—find a non-negative n-vector x which maximizes $f(x)$ subject to the constraints $g^i(x) \geqq 0$, $i = 1, \cdots, m$—is investigated where $f(x)$ is assumed to be concave or pseudo-concave and the $g^i(x)$ are increasing functions. It is shown that under certain conditions on $g^i(x)$, the Kuhn-Tucker-Lagrange conditions are necessary and sufficient for the optimality of x^*. It is also shown that the $g^i(x)$ are a useful class of functions since, among other properties, they are closed under non-negative addition, under the addition of any scalar, and under multiplication of non-negative members of the class.

Examples of the above programming problem with increasing constraint functions are found in many chance-constrained programming problems.

The mathematical programming problem can be briefly described as follows: we seek a non-negative vector $x = (x_1, \cdots, x_n)$, $x \, \varepsilon \, E^n$, which

(1.1) $$\text{maximizes} \quad f(x)$$

subject to the constraints

(1.2) $$g^i(x) \geq 0 \qquad\qquad i = 1, \cdots, m$$

(1.3) $$x_j \geq 0 \qquad\qquad j = 1, \cdots, n$$

where f and g^i are real-valued differentiable functions. It was well established by Kuhn and Tucker [6] that if the constraint functions satisfy a constraint qualification,[1] then the Kuhn-Tucker Lagrange conditions (KTL)

(1.4) $$\nabla f(x^*) + \lambda \nabla g(x^*) + \mu = 0$$

(1.5) $$\mu x^* = 0$$

(1.6) $$\lambda g(x^*) = 0$$

(1.7) $$\lambda, \mu \geqq 0$$

are necessary conditions for x^* to maximize $f(x)$ subject to $g^i(x) \geq 0$, $x_j \geq 0$, where $\nabla g(x^*) = (\nabla g^1(x^*), \nabla g^2(x^*), \cdots, \nabla g^m(x^*))$ and $g(x^*) = (g^1(x^*), \cdots, g^m(x^*))$. They also showed that if f and g^i are concave functions, then the KTL are also sufficient conditions for the optimality of x^*. At a later date Arrow and Enthoven [2] extended these results under certain conditions to quasi-concave functions f and g^i.

In this paper we will show that certain increasing functions $g^i(x)$ have the property that the KTL are necessary and sufficient conditions to maximize a concave or pseudo-concave function $f(x)$. It should be emphasized at this point that the results of this paper broaden the class of constraint functions for which the KTL are necessary and

[1] The Kuhn-Tucker constraint qualification states: for any feasible point $\bar{x}$ on the boundary of the constraint set and for every direction h such that $\nabla g^i(\bar{x}) \cdot h \geqq 0$ for all i with $g^i(\bar{x}) = 0$ and $h_j \geqq 0$ for all j with $x_j = 0$ there exists a differentiable vector-valued function $\psi(\theta)$ such that $\psi(0) = \bar{x}, \psi(\theta) = x$ belongs to the constraint set for all $\theta > 0$ and θ sufficiently small and $\psi'(0) = h$.

sufficient. No attempt, at the present time, has been made to broaden the class of objective functions.

The following notational conventions [5] will be used throughout this paper: If $x = (x_1, \cdots, x_n)$ in E^n, then

x is *non-negative*, written $x \geq 0$, if $x_i \geq 0$ for all i,

x is *positive*, written $x > 0$, if $x_i > 0$ for all i,

x is *semipositive*, written $x \geq 0$, if $x_i \geq 0$ for all i,

and $x_i > 0$ for some i.

Furthermore, if x^1 and $x^2 \varepsilon E^n$, then we write $x^2 \geqq x^1$, $x^2 > x^1$, or $x^2 \geq x^1$, according as $x^2 - x^1$ is non-negative, positive, or semi-positive, respectively.

A scalar function $k(x)$ defined on a set $C \subset E^n$ is *increasing* if

$$(1.8) \qquad k(x^1) \leq k(x^2), \qquad \text{whenever} \quad x^1 \leqq x^2.$$

$k(x)$ is *strictly increasing* if $k(x^1) < k(x^2)$ whenever $x^1 \leq x^2$.

If it is assumed that $k(x)$ is differentiable on an open set C then an equivalent definition of an increasing function can be given by the lemma:

Lemma 1.1: If $k(x)$ is a differentiable scalar function on an open set $C \subset E^n$ the following are equivalent:

(1) $k(x)$ is an increasing function on C

(2) the directional derivative of k at x in the direction h, $D_h k(x)$, is non-negative for all $h \geq 0$ for all $x \varepsilon C$

(3) $\nabla k(x) \geqq 0$, for all $x \varepsilon C$.

Proof:

(1) $\rightarrow$ (2) Consider $x^2 \geq x^1$ or $x^2 - x^1 = h \geq 0$ and $\tau > 0$. Then

$$(k(x^1 + \tau h) - k(x^1))/\tau \geq 0 \qquad \text{for all} \quad \tau > 0.$$

Thus $\qquad \lim_{\tau \to 0} (k(x^1 + \tau h) - k(x^1))/\tau = D_h k(x^1) \geq 0$

(2) $\rightarrow$ (3) Let $\qquad \bar{h} = h/|h|;$ hence $\bar{h} \geq 0$

Then $D_h k(x) = \nabla k(x)\bar{h} \geq 0$ for all $\bar{h} \geq 0$. In particular pick $\bar{h} = e^i$, the ith unit vector for each $i = 1, \cdots, n$, then

$$\partial k(x)/\partial x_i \geq 0$$

for all $\qquad i = 1, \cdots, n$ and $\nabla k(x) \geqq 0$.

(3) $\rightarrow$ (1) Given $\nabla k(x) \geqq 0$ for all $x \varepsilon C$. By Taylor's Theoerm if $x^2 \geq x^1$ then for some θ, $0 \leq \theta \leq 1$

$$k(x^2) = k(x^1) + \nabla k(\theta x^2 + (1 - \theta)x^1)(x^2 - x^1)$$

But $x^2 - x^1 \geq 0$ and $\nabla k(\theta x^2 + (1 - \theta)x^1) \geqq 0$ thus $k(x^2) \geq k(x^1)$.   *Q.E.D.*

In general we will be most interested in $k(x)$ an increasing differentiable function defined for all $x \varepsilon E^n$; in this case the results of Lemma 1.1 will be used interchangeably with the definition of an increasing function.

For completeness we include the following definitions. A scalar function $f(x)$ defined on a convex subset $C \subset E^n$ is *concave* if

$$f(\alpha x^1 + (1 - \alpha)x^2) \geq \alpha f(x^1) + (1 - \alpha)f(x^2)$$

for all α such that $0 \leq \alpha \leq 1$ and x^1 and $x^2 \varepsilon C$.

A scalar function $f(x)$ defined on a convex subset $C \subset E^n$ is *pseudo-concave* if for every x^1 and $x^2 \, \varepsilon \, C$, $(x^2 - x^1)\nabla f(x^1) \leq 0$ implies $f(x^2) \leq f(x^1)$. (See [9])

A function $f(x)$ defined on a convex subset $C \, \varepsilon \, E^n$ is *quasi-concave* if for $x^1 \, \varepsilon \, C$ and $x^2 \, \varepsilon \, C$ and all α with $0 \leq \alpha \leq 1$, $f(\alpha x^1 + (1 - \alpha)x^2) \geq \min [f(x^1), f(x^2)]$. If $f(x)$ is an increasing, concave, pseudo-concave or quasi-concave function, $-f(x)$ is a decreasing, convex, pseudo-convex or quasi-convex function, respectively.

It should be noted that not all increasing (or decreasing) functions are monotonic quasi-concave or quasi-convex functions as the following example shows:
Let $X = \{x \mid x \geqq 0, x \, \varepsilon \, E^2\}$

$$
\begin{aligned}
k(x) &= x_1 x_2 - \tfrac{1}{4} \quad \text{for} \quad 0 \leq x_1 \leq 1, x_1 \leq x_2 \\
&= x_1 x_2 - \tfrac{1}{4} \quad \text{for} \quad 0 \leq x_2 \leq 1, x_2 \leq x_1 \\
&= (x_1 - 1)(x_2 - 1)^2 + x_2 - \tfrac{1}{4} \quad \text{for} \quad 1 \leq x_1, 1 \leq x_2, x_1 \leq x_2 \\
&= (x_1 - 1)^2(x_2 - 1) + x_1 - \tfrac{1}{4} \quad \text{for} \quad 1 \leq x_1, 1 \leq x_2, x_2 \leq x_1
\end{aligned}
$$

Now $k(x)$ is increasing and continuous on X but $k(x)$ is neither quasi-concave nor quasi-convex. Consider $x = (\tfrac{1}{2}, \tfrac{1}{2})$ then $k(x) = 0$; now let $x^1 = (\tfrac{3}{4}, \tfrac{1}{4})$ and $x^2 = (\tfrac{1}{4}, \tfrac{3}{4})$ then $k(x^1) = -\tfrac{1}{16}$, $k(x^2) = -\tfrac{1}{16}$ and $k(x) \nleq \max [k(x^1), k(x^2)]$ where $x = \tfrac{1}{2}x^1 + \tfrac{1}{2}x^2$. Thus $k(x)$ is not quasi-convex. Now consider $x = (\tfrac{3}{2}, \tfrac{3}{2})$, $k(x) = 88/64$ and let $x^1 = (\tfrac{7}{4}, \tfrac{5}{4})$, $x^2 = (\tfrac{5}{4}, \tfrac{7}{4})$ then $k(x^1) = 105/64 = k(x^2)$ and $k(x) \ngeq \min [k(x^1), k(x^2)]$ where $x = \tfrac{1}{2}x^1 + \tfrac{1}{2}x^2$. Thus $k(x)$ is not quasi-concave.

It is true, however, that if the domain of k is a convex subset $C \subset E^1$, then the definition of an increasing (or decreasing) function coincides with the definition of a monotonic quasi-concave function.

2. Some Further Properties of Increasing Functions

Increasing functions have some properties which make them useful. Although most of these properties are self-evident we state them in the form of a lemma in order to refer to them more easily.

Lemma 2.1: Let $\theta(\cdot)$ and $\psi(\cdot)$ be two continuous increasing functions defined for all $x \, \varepsilon \, C = E^n$ where C is a given set.

(1) $w(x) = \lambda_1 \theta(x) + \lambda_2 \psi(x)$ is an increasing function for $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$.

(2) $w(x) = \lambda_1 \theta(x) + \lambda_2$ is an increasing function for $\lambda_1 \geq 0$ and a decreasing function for $\lambda_1 \leq 0$.

(3) if in addition $\theta(\cdot)$ and $\psi(\cdot)$ are non-negative for all $x \, \varepsilon \, C$, then $w(x) = \theta(x)\psi(x)$ is an increasing function.

(4) $w(x) = \min \{\theta(x), \psi(x)\}$ is an increasing function.

(5) $w(x) = \max \{\theta(x), \psi(x)\}$ is an increasing function.

(6) if in addition $\theta(x) > 0$ for all $x \, \varepsilon \, C$, then (i) $w(x) = 1/\theta(x)$ is a decreasing function, and (ii) $w(x) = \ln \theta(x)$ is an increasing function.

(7) if $D \subset E^1$ is the range of ψ, if θ maps D into $F \subset E^1$ and if θ is an increasing function, then $w(x) = \theta(\psi(x))$ is an increasing function. (This part includes, *e.g.*, 6(ii) above.)

Since the results of this lemma follow directly from the definition of an increasing function, the proof will not be given.

There are many other properties of increasing functions in addition to those stated above. Furthermore, most of the foregoing properties generalize in the appropriate way to a sequence (finite or infinite) of increasing functions, $\{\theta^i(x)\}$. For example, part (1) can be stated: if $w(x) = \sum_{i=1}^{\infty} \lambda_i \theta^i(x)$ exists, and if all $\lambda_i \geq 0$, then $w(x)$ is an increasing function.

The objective of this paper is to give necessary and sufficient conditions for $x^* \geq 0$ to maximize $f(x)$ when the $g^i(x)$ are increasing functions. Before doing this, it should be pointed out that the $g^i(x)$ arise naturally from some constraints to the joint chance-constrained programming problem as defined by Symonds [13, 14] and by Wagner and Miller [15]. The constraints have the property that they are products of cumulative distribution functions and in many cases are merely increasing functions of the decision variables. We present this example directly.

3. A Chance-Constrained and a Joint Chance-Constrained Example

It was mentioned that an example of the type of mathematical programming problem given by (1.1), (1.2) and (1.3) could be obtained from chance-constrained programming. We quote the formulation and notation given in [15].

We seek an n-dimensional vector $x = (x_1, \cdots, x_n)$ to

$$(3.1) \qquad \text{maximize } cx$$

subject to

$$(3.2) \qquad Ax \leq b$$

$$(3.3) \qquad x \geq 0$$

$$(3.4) \qquad \text{Probability } \{\alpha_i x - \beta_i \geq \gamma_i\} \geq \eta_i, \qquad i \, \varepsilon \, S$$

where c is an n vector, b is a k vector, A is a $k \times n$ matrix, α_i is an n vector (which may contain random elements), β_i is a random variable, γ_i is a known constant, η_i is a specified probability level $0 < \eta_i \leq 1$, and S is the set of chance constraints.

The above model represents a chance-constrained programming problem with linear objective function. For a joint chance-constrained programming model we substitute

$$(3.5) \qquad \prod_{i \varepsilon S} \text{Probability } \{\alpha_i x - \beta_i \geq \gamma_i\} \geq \eta > 0$$

for the set of chance constraints (3.4). In this latter case it will be assumed (as shown in (3.5)) that the random variables ξ_i given by

$$(3.6) \qquad \xi_i = \alpha_i x - \beta_i, \qquad i \, \varepsilon \, S$$

are independent for all $i \, \varepsilon \, S$.

We now prove a lemma concerning the chance-constraints (3.4) or (3.5). First denote the sample space of a random vector $\alpha = (\alpha_1, \cdots, \alpha_m)$ by $\P = (Z, \Omega, p)$ where $Z \, \varepsilon \, E^m$ is the space of outcomes of an experiment, Ω is a parameter space and p is a function on $Z \times \Omega$ such that for fixed $\omega \varepsilon \Omega$, P_ω is a probability distribution on Z.

Lemma 3.1: Let $x^1, x^2 \, \varepsilon \, E^n$ such that $x^1 \leq x^2$. Let $q(\alpha; x)$ be a real-valued function of α and x. If for any $a \, \varepsilon \, Z$, $q(a; x)$ is an increasing function of x, then

$$P\{q(\alpha; x^2) \geq \gamma\} \geq P\{q(\alpha; x^1) \geq \gamma\}.$$

Proof: Let F_α be the cumulative distribution function of the vector random variable $\alpha = (\alpha_1, \cdots, \alpha_m)$. Thus

$$P\{q(\alpha; x^2) < \gamma\} = \int \cdots \int_{\{a \, : \, q(a;x^2) < \gamma\}} dF_{\alpha_1, \cdots, \alpha_m}(a_1, \cdots, a_m)$$

$$\leq \int \cdots \int_{\{a \, : \, q(a;x^1) < \gamma\}} dF_{\alpha_1, \cdots, \alpha_m}(a_1, \cdots, a_m) = P\{q(\alpha; x^1) < \gamma\}$$

since $q(a; x^2) \geq q(a; x^1)$ for all a implies

$$\{a{:}q(a; x^2) < \gamma\} \subset \{a{:}q(a; x^1) < \gamma\} \qquad\qquad Q.E.D.$$

Corollary 3.1 Let $\alpha = (\alpha_1, \cdots, \alpha_n, \beta)$, $\alpha_i(i = 1, \cdots, n)$ are nonnegative random variables, β is a random variable and $q(\alpha; x) = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n - \beta$ in Lemma 3.1, then $P\{\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n - \beta \geq \gamma\}$ is an increasing function of x.

Thus we see that if the α_i are nonnegative random variables, the chance-constraints given by (3.4) are merely increasing functions of x. Note that in lemma 2.1 we showed that adding or subtracting a constant from an increasing function yields an increasing function. Furthermore, the product of nonnegative increasing functions is a nonnegative increasing function. Thus using Corollary 3.1 we are able to state:

Corollary 3.2 If A is a matrix of nonnegative random variables, if β is a vector of random variables, if $\xi_i = \alpha_i x - \beta_i$ are independent for $i \, \varepsilon \, S$, then

$$k(x) = P\begin{Bmatrix} \alpha_1 \cdot x - \beta_1 \geq \gamma_1 \\ \vdots \\ \alpha_r \cdot x - \beta_r \geq \gamma_r \end{Bmatrix} - \eta$$

$$= \prod_{i=1}^{r} [1 - F_{\xi_i}(\gamma_i{:}x)] - \eta \geq 0 \qquad\qquad \text{for } 1 \geq \eta > 0$$

is an increasing function of x.

In this case we see that the constraint given by (3.5) is an increasing function of the decision variable x.

Thus the chance-constrained problem given by (3.1), (3.2), (3.3) and (3.4) and the joint chance-constrained problem (3.1), (3.2), (3.3) and (3.5) are merely examples of the more general problem given by (1.1), (1.2) and (1.3), provided of course the α_{ij}'s are nonnegative random variables.

4. Necessary Conditions

Before proceeding it will be convenient to introduce the following additional notation:

(1) Let C be any subset of E^n.
(2) Let $H_i = \{x \mid g^i(x) \geq 0, x \, \varepsilon \, C\}$ for $i = 1, \cdots, m$,
(3) Let $Y = \bigcap_{i=1}^{m} H_i$.
(4) Let $I(x) = \{i \mid g^i(x) = 0, \quad i = 1, \cdots, m, x \, \varepsilon \, C\}$.
(5) Let $J(x) = \{j \mid x_j = 0, \quad j = 1, \cdots, n, x \, \varepsilon \, C\}$.
(6) Let $D(x) = \{h \mid \nabla g^i(x)h \geq 0, \quad h_j \geq 0, \quad i \, \varepsilon \, I(x), \quad j \, \varepsilon \, J(x)\}$.

Thus if $C = X \equiv \{x \mid x \geq 0, x \, \varepsilon \, E^n\}$ then Y is the set of feasible points to the mathematical programming problem defined at the beginning of this paper; $I(\bar{x})$ and $J(\bar{x})$ are the set of indices i and j of binding constraints when the point $\bar{x}$ is given; and $D(\bar{x})$ is the set of locally constrained directions from $\bar{x}$.

We say that the function $g^i(x)$ has *property* Γ if for all $x \, \varepsilon \, C$, $g^i(x)$ is an increasing continuous function and if for any $x \, \varepsilon \, C$, $y \, \varepsilon \, C$ such that $y \geq x$ and $y_n > x_n$ then $g^i(y) > g^i(x)$. It should be mentioned that if $g^i(x)$ is a strictly increasing continuous function then $g^i(x)$ possesses property Γ automatically

Let $S \equiv \{s \mid g^i(s, y) = 0, \text{ for some } y \text{ with } (s, y) \, \varepsilon \, C, s \, \varepsilon \, E^{n-1}, y \, \varepsilon \, E^1\}$

Lemma 4.1: If $C \subset E^n$ is an open set and $n > 1$ and if for some $x \, \varepsilon \, C \, g(x) = 0$ where $g(\cdot)$ has property Γ, then every open neighborhood of x, $N(x)$, contains a point $z \, \varepsilon \, C$, $z \neq x$ such that $g(z) = 0$.

Proof: Without loss of generality, consider $N(x) \subset C$.

Let $W = \{w \mid w \geqq x, w \, \varepsilon \, N(x)\}$ and

$$V = \{v \mid v \leqq x, v \, \varepsilon \, N(x)\}.$$

It is clear that W and V have the same dimensionality as $N(x)$. Now $N(x)$ is a convex set. Consider any $v \, \varepsilon \, V$ and $w \, \varepsilon \, W$ then for all λ, $0 \leq \lambda \leq 1$ the point $z \, \varepsilon \, N(x)$ where $z = \lambda v + (1 - \lambda)w$. Since V and W have the full dimensionality of $N(x)$ we can choose $v_0 \, \varepsilon \, V$ and $w_0 \, \varepsilon \, W$ such that for all λ $(0 \leq \lambda \leq 1)$ we have $z_0 \neq x$, where $z_0 = \lambda v_0 + (1 - \lambda)w_0$. Now from property Γ, $g(w_0) \geq 0 \geq g(v_0)$. If $g(w_0) = 0$ or $g(v_0) = 0$ the lemma is proved; if not, then $g(w_0) > 0$ and $g(v_0) < 0$ and by a theorem of Bolzano [1], there exists a point $\bar{z}_0 = \bar{\lambda}v_0 + (1 - \bar{\lambda})w_0$ such that $g(\bar{z}_0) = 0$.   Q.E.D.

Lemma 4.2: If $g^i(x)$ has property Γ, then the level curve given by $g^i(x) = 0$, if it exists, defines a decreasing continuous function $x_n = \phi^i(x_1, \cdots, x_{n-1})$ where $g^i(x_1, \cdots, x_{n-1}, \phi^i(x_1, \cdots, x_{n-1})) = 0$.

Proof: The fact that $\phi^i(y)$ is uniquely defined as a scalar function of y follows immediately from property Γ; since if $x = (x_1, \cdots, x_{n-1}, x_n)$ and $\hat{x} = (x_1, \cdots, x_{n-1}, \hat{x}_n)$ are two points of C differing only in their last coordinate then $g^i(x) \neq g^i(\hat{x})$. Thus if the first $n - 1$ coordinates $x_1, \cdots, x_{n-1}$ of a point $x \, \varepsilon \, C$ are fixed there is only one choice of the last coordinate x_n such that $g^i(x) = 0$.

We will now show that $\phi^i(y)$ is a decreasing function by showing if $y = (y_1, \cdots, y_{n-1}) \leq (z_1, \cdots, z_{n-1})$ then $\phi^i(y) \geq \phi^i(z)$. If $y = z$ then $\phi^i(y) = \phi^i(z)$ trivially; so assume $y \leq z$. By the definition of $\phi^i(\cdot)$ $g^i(y, \phi^i(y)) = 0$ and $g^i(z, \phi^i(z)) = 0$, where $(y, \phi^i(y)) \, \varepsilon \, C$ and $(z, \phi^i(z)) \, \varepsilon \, C$. Consider the point $(y, \phi^i(z))$. If $(y, \phi^i(z)) \, \varepsilon \, C$ then consider any $\varepsilon > 0$ and the extended increasing function $\bar{g}^i(x)$ defined on E^n given by

$\bar{g}^i(x) = g^i(x)$ if $x \, \varepsilon \, C$

      $= \text{infimum } g(w) - \varepsilon$ if $x \, \varepsilon \, C$ and $\{w \mid w \geqq x\} \cap C \neq \text{empty set}$
      $\{w \mid w \geqq x\} \cap C$

      $= \text{supremum } g(u)$ if $x \, \varepsilon \, C$ and $\{w \mid w \geqq x\} \cap C = \text{empty set}$
      $u \, \varepsilon \, C$

Then

$\bar{g}^i(y, \phi^i(z)) = \text{infimum}$ $g^i(w) < \bar{g}^i(z, \phi^i(z))$
      $\{w \mid w \geqq (y, \phi^i(z))\} \cap C$ $= g^i(z, \phi^i(z))$

                                              $= 0$

                                              $= g^i(y, \phi^i(y))$

                                              $= \bar{g}^i(y, \phi^i(y))$

Now $\bar{g}^i(y, \phi^i(z))$ and $\bar{g}^i(y, \phi^i(y))$ have arguments differing *only* in their last coordinate and since $\bar{g}^i(\cdot)$ is an increasing function we must have $\phi^i(z) \leq \phi^i(y)$.
If $(y, \phi^i(z)) \, \varepsilon \, C$ then

(4.1) $g^i(y, \phi^i(z)) \leq g^i(z, \phi^i(z)) = 0 = g^i(y, \phi^i(y))$.

If $\phi^i(z) > \phi^i(y)$ then from property Γ, $g^i(y, \phi^i(z)) > g^i(y, \phi^i(y))$ which contradicts (4.1). Thus $\phi^i(z) \leq \phi^i(y)$ in both cases. Thus $\phi^i(\cdot)$ is a decreasing function. We now show $\phi^i(\cdot)$ is continuous on S.

From the preceding lemma 4.1, for any point $(z, \phi^i(z))$ ε S there exists a sequence of points $\{(y^j, \phi^i(y^j))\}_{j=1}^{\infty}$ ε S such that this sequence converges to the point $(z, \phi^i(z))$. Thus $\lim_{y^j \to z} \phi^i(y^j) = \phi^i(z)$ and $\phi^i(\cdot)$ is continuous on S. $Q.E.D.$

It is possible that there does not exist an x ε C such that $g^i(x) = 0$. In this case we have either $g^i(x) > 0$ for all x ε C and $g^i(x)$ is not a binding constraint over C or $g^i(x) < 0$ for all x ε C and the set H_i is empty. We are not interested in either of these cases. Therefore it will always be assumed that for all $i = 1, \cdots, m$ the H_i are not empty and for some i, the g^i are binding.

Corollary 4.1 If $g^i(x)$ defined on an open set C is an increasing function, $g^i(x)$ has continuous first partial derivatives (that is $g^i(x)$ ε C^1) and if $\partial g^i(x)/\partial x_n > 0$ for all x ε C, then $\phi^i(\cdot)$, if it exists, is a decreasing function and ϕ^i ε C^1 on S.

Proof: The fact that ϕ^i is decreasing follows from lemma 4.2 since g^i satisfies property Γ. Thus it is only necessary to show ϕ^i ε C^1 on S. We know that ϕ^i is uniquely defined by the set of points $g^i(x) = 0$. Consider any point $\bar{x}$ ε C such that $g^i(\bar{x}) = 0$. By the implicit function theorem, there exists a unique function $\psi(\cdot)$ ε C^1 such that $g^i(x_1, \cdots, x_{n-1}, \psi(x_1, \cdots, x_{n-1})) = 0$ for every $(x_1, \cdots, x_{n-1})$ ε $N(\bar{x}_1, \cdots, \bar{x}_{n-1})$ where $N(\bar{x}_1, \cdots, \bar{x}_{n-1})$ is a neighborhood of $(\bar{x}_1, \cdots, \bar{x}_{n-1})$ contained in C.

But the function ϕ^i also satisfies $g^i(x_1, \cdots, x_{n-1}, \phi^i(x_1, \cdots, x_{n-1})) = 0$ for $(x_1, \cdots, x_{n-1})$ ε $N(\bar{x}_1, \cdots, \bar{x}_{n-1})$; therefore $\phi^i = \psi$ in $N(\bar{x})$ since ψ is unique. Thus ϕ^i ε C^1 on $N(\bar{x}_1, \cdots, \bar{x}_{n-1})$. But x was an arbitrary point such that $g^i(\bar{x}) = 0$; hence ϕ^i ε C^1 on S. $Q.E.D.$

Theorem 4.1 Let $g^i(x)$ have property Γ on a convex set C. Then the set H_i is a convex set if and only if $\phi^i(\cdot)$ is a convex function on S.

Proof: (Necessity) We are given that H_i is convex and nonempty. Assume to the contrary that ϕ^i is not convex. Then there exists two points $x = (x_1, \cdots, x_{n-1}, \phi^i(x_1, \cdots, x_{n-1}))$, $y = (y_1, \cdots, y_{n-1}, \phi^i(y_1, \cdots, y_{n-1}))$ ε C [let us denote the first $n-1$ components of these vectors by $\hat{x}$ and $\hat{y}$, respectively] such that for some λ with $0 < \lambda < 1$

$$(4.2) \qquad \phi^i(\lambda\hat{x} + (1 - \lambda)\hat{y}) > \lambda\phi^i(\hat{x}) + (1 - \lambda)\phi^i(\hat{y}).$$

Let $\hat{z} = \lambda\hat{x} + (1 - \lambda)\hat{y}$. Then the points $(\hat{z}, \phi^i(\hat{z}))$ and

$$(\hat{z}, \lambda\phi^i(\hat{x}) + (1 - \lambda)\phi^i(\hat{y})) \ \varepsilon \ C$$

by the existence of ϕ^i and the convexity of C. Furthermore, from (4.2) $(\hat{z}, \phi^i(\hat{z})) \geq (\hat{z}, \lambda\phi^i(\hat{x}) + (1 - \lambda)\phi^i(\hat{y}))$. Therefore using property Γ and (4.2) we obtain

$$(4.3) \qquad g^i(\hat{z}, \phi^i(\hat{z})) > g^i(\hat{z}, \lambda\phi^i(\hat{x}) + (1 - \lambda)\phi^i(\hat{y})).$$

Now from the definition of ϕ^i we have $g^i(\hat{z}, \phi^i(\hat{z})) = 0$. Thus (4.3) can be written as:

$$(4.4) \qquad 0 = g^i(\hat{z}, \phi^i(\hat{z})) > g^i(\hat{z}, \lambda\phi^i(\hat{x}) + (1 - \lambda)\phi^i(\hat{y})).$$

Hence the point

$$(\hat{z}, \lambda\phi^i(\hat{x}) + (1 - \lambda)\phi^i(\hat{y})) \ \varepsilon \ H_i .$$

But the points $(\hat{x}, \phi^i(\hat{x}))$ and $(\hat{y}, \phi^i(\hat{y}))$ ε H_i and since H_i is convex, (4.4) yields the contradiction.

(Sufficiency) We are given that ϕ^i is convex. Let x, y ε H_i, we will show for any λ such that $0 \leq \lambda \leq 1$, $z = \lambda x + (1 - \lambda)y$ ε H_i. Let $\hat{x}$, $\hat{y}$ and $\hat{z}$ denote the first $n-1$ elements of x, y and z respectively.

Since g^i is an increasing function, then $g^i(x) \geq 0 = g^i(\hat{x}, \phi^i(\hat{x}))$ and $g^i(y) \geq 0 = g^i(\hat{y}, \phi^i(\hat{y}))$ imply

$$x \geq (\hat{x}, \phi^i(\hat{x})) \text{ and } y \geq (\hat{y}, \phi^i(\hat{y})).$$

This last result follows from that fact that x differs from $(\hat{x}, \phi^i(\hat{x}))$ only in its last component and $g^i(\cdot)$ is strictly increasing in its last component by property Γ (similarly for y). Therefore

(4.5) $$z \geq \lambda(\hat{x}, \phi^i(\hat{x})) + (1 - \lambda)(\hat{y}, \phi^i(\hat{y})).$$

From (4.5) and since g^i is an increasing function

$$g^i(z) \geq g^i(\lambda\hat{x} + (1 - \lambda)\hat{y}, \quad \lambda\phi^i(\hat{x}) + (1 - \lambda)\phi^i(\hat{y}))$$

$$\geq g^i(\lambda\hat{x} + (1 - \lambda)\hat{y}, \quad \phi^i(\lambda\hat{x} + (1 - \lambda)\hat{y})) = 0.$$

The last inequality follows from the convexity of ϕ^i. Therefore $z \, \varepsilon \, H_i$. Q.E.D.

Arrow and Enthoven show that a sufficient condition for ϕ^i to be convex is that $g^i(x)$ be quasi-concave. Thus for the case of increasing functions we only need $g^i(x)$ quasi-concave in a neighborhood of all x satisfying $g^i(x) = 0$. In our previous example of an increasing function which is neither quasi-concave nor quasi-convex, the function $\phi(x_1)$ is convex for $x_1 > 0$. That is, $g(x) = 0$ is given by $g(x) = x_1x_2 - \frac{1}{4} = 0$, and $x_2 = \phi(x_1) = 1/4x_1$.

Now from Theorem 4.1 it follows if the ϕ^i are convex, then Y is a convex set. Furthermore we can state:

Lemma 4.3: If Y is convex and has an interior and if for each $i = 1, \cdots, m$, $g^i(x)$ are increasing differentiable functions for $x \, \varepsilon \, Y$ and for any $\bar{x} \, \varepsilon$ boundary of Y, $\partial g^i(\bar{x})/\partial x_j > 0$ for some j and all i, then the KT Constraint Qualification holds.

Proof: Consider the points x, $y \, \varepsilon \, Y$ such that $x \neq y$ and the point $\bar{x} \, \varepsilon \, Y$ such that $\bar{x}$ is a boundary point of Y. Then the set $I(\bar{x}) \cup J(\bar{x})$ is not empty. Let $h^1 = x - \bar{x}$ and $h^2 = y - \bar{x}$ be two directions. Then h^1, $h^2 \, \varepsilon \, D(\bar{x})$. Furthermore for

$z = \lambda x + (1 - \lambda)y$ for all λ such that $0 \leq \lambda \leq 1$,

$$h = z - \bar{x} = \lambda h^1 + (1 - \lambda)h^2$$

belongs to $D(\bar{x})$. Then $D(\bar{x})$ has the same dimension as Y. Furthermore $D(\bar{x})$ is a closed convex cone.

Case 1: Let $h \, \varepsilon \, D(\bar{x})$ be an interior point of $D(\bar{x})$. Then since Y is a convex set there exists a $\theta(h) > 0$ such that the set of points $\{y \mid y = \bar{x} + \theta h, \text{ for all } \theta \text{ such that } 0 \leq \theta < \theta(h)\} \subset Y$.
Let $\psi(\theta) = \bar{x} + \theta h$ for $0 \leq \theta < \theta(h)$.

Then (i) $\psi(0) = \bar{x}$
     (ii) $\psi(\theta) = \bar{x} + \theta h \, \varepsilon \, Y$ for $0 \leq \theta < \theta(h)$
     (iii) $\lim_{\theta \to 0+} (\psi(\theta) - \psi(0))/\theta = h$.

Thus if h is an interior point of $D(\bar{x})$ then there exists a contained path $\psi(\theta)$.

Case 2: Let $h \, \varepsilon \, D(\bar{x})$ be a boundary point of $D(\bar{x})$. Since $D(\bar{x})$ has an interior there exists a sequence of interior points $\{h^i\}$ $h^i \, \varepsilon \, D(\bar{x})$ such that $\lim_{i \to \infty} h^i = h$. Furthermore for each h^i, there exists a $\theta_i > 0$ such that

$$y = \bar{x} + \theta h^i \, \varepsilon \, Y \text{ for all } \theta \text{ with } 0 \leq \theta < \theta_i$$

Without loss of generality assume $\theta_i < \theta_{i-1}$ and define

$$\psi(\theta) = \bar{x} + \theta h^i \qquad\qquad \text{for } \theta_{i+1} \leq \theta < \theta_i,$$
$$= \bar{x} \qquad\qquad\qquad \text{for } \theta = 0.$$

Then (i) $\psi(0) = \bar{x}$

(ii) $\psi(\theta) = \bar{x} + \theta h^i \ \varepsilon \ Y$ for $0 \leq \theta < \theta_i$

(iii) $\lim_{\theta \to 0+} (\psi(\theta) - \psi(0))/\theta = \lim_{\theta \to 0+, i \to \infty} (\bar{x} + \theta h^i - \bar{x})/\theta$

$$= h.$$

Thus, if h is a boundary point of $D(\bar{x})$, then $\psi(\theta)$ is a contained path. Therefore the KT Constraint Qualification holds.   Q.E.D.

In addition we can say that the Arrow, Hurwicz and Uzawa (AHU) Constraint Qualification W holds.[2] This latter condition follows from the fact that the KT Constraint Qualification implies the AHU Constraint Qualification W. In any event, however, the important result which follows from the preceding lemma is that the KTL (given by (1.4) − (1.7)) are a necessary condition for the optimality of x^*, i.e., if $x^* \geq 0$ maximizes $f(x)$ subject to $g^i(x) \geq 0$ for all $i = 1, \cdots, m$, then there exists a λ such that λ and x^* satisfy (1.4)–(1.7).

5. Sufficient Conditions

In this section we will show that if $g^i(x)$ is differentiable and has property Γ for $i = 1, \cdots, m$, if the H_i form convex sets, if f is a differentiable pseudo-concave function, and if λ^* and x^* are a solution to the KTL conditions, then x^* is optimal. Thus under these assumptions on f and g^i, the KTL are sufficient conditions for the optimality of x^*. We state this result as a theorem.

Theorem 5.1 If f is a differentiable pseudo-concave function for all $x \ \varepsilon \ C$, if $g^i(x)$ is differentiable and has property Γ for all $i = 1, \cdots, m$ and all $x \ \varepsilon \ C$, if the $\phi^i(\cdot)$ are convex functions for all $i = 1, \cdots, m$, and if x^* is a feasible point which satisfies the KTL (or equivalently $D_h f(x^*) \leq 0$ for all $h \ \varepsilon \ D(x^*)$), then x^* is optimal.

Proof: The fact that the KTL and the condition that $D_h f(x^*) \leq 0$ for all $h \ \varepsilon \ D(x^*)$ are equivalent follows directly from the application of Farkas' lemma to $D_h f(x^*) \leq 0$ for all $h \ \varepsilon \ D(x^*)$.

Since the ϕ^i are convex functions then the feasible set of points $Y = \bigcap_{i=1}^{n} H_i$ is convex. It is only necessary to show that if $D_h f(x^*) \leq 0$ for all $h \ \varepsilon \ D(x^*)$ then x^* maximizes $f(x)$ on Y.

For all $x \ \varepsilon \ Y$, $h = x - x^*$ belongs to $D(x^*)$ and $D_h f(x^*) \leq 0$ implies

$f(x^* + h) = f(x) \leq f(x^*)$.   Q.E.D.

Unfortunately the restriction that the ϕ^i's be convex for all i or equivalently that the H_i be convex sets is a strong requirement on the constraint functions. Actually in Theorem 5.1 and Lemma 4.3 it was only necessary that Y be a convex set, however, as was noted previously a sufficient condition for the latter to hold is that H_i be convex for all i.

As noted before the class of increasing functions has many useful properties, particularly for chance-constrained programming. The class is closed under non-negative addition of the functions, under the addition of any scalar, and under the product of non-negative functions. If in addition, the ϕ's are convex, then the KTL are necessary and sufficient conditions for the existence of the solution to the maximization problem when the objective function is pseudo-concave.

References

1. Apostol, T. M., *Mathematical Analysis*, Reading, Massachusetts, Addison-Wesley, 1957.
2. Arrow, Kenneth J. and Alain C. Enthoven, "Quasi-Concave Programming," *Econometrica*, Vol. 29, No. 4, (October 1961), pp. 779–800.
3. ——, Leonid Hurwicz and Hirofumi Uzawa, "Constraint Qualifications in Maximizations Problems," *Naval Research Logistics Quarterly*, Vol. 8, No. 2 (June 1961), pp. 175–191.
4. ——, —— and ——, *Studies in Linear and Nonlinear Programming*, Stanford, California, Stanford University Press, 1958.
5. Gale, David, *The Theory of Linear Economic Models*, McGraw-Hill Book Co., New York, 1960, p. 43.
6. Kuhn, H. W. and A. W. Tucker, "Nonlinear Programming," in J. Neyman (ed.) *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, California, University of California Press, 1951, pp. 481–492.
7. Karamardian, S., "Duality in Mathematical Programming," Technical Report ORC 66-2 (January 1966), Operations Research Center, University of California—Berkeley, California.
8. Kortanek, K. O. and J. P. Evans, "Pseudo-Concave Programming and Lagrange Regularity," *Operations Research*. Vol. 15, No. 5, pp. 882–891.
9. Mangasarian, O. S., "Pseudo-Convex Functions," *J. SIAM Control*, Series A, Vol. 3, No. 2, pp. 281–290.
10. Martos, B., "The Direct Power of Adjacent Vertex Programming Methods," *Management Science*, Vol. 12, No. 3 (November 1965), pp. 241–252.
11. Ponstein, J., "Seven Kinds of Convexity," *SIAM Review*, Vol. 9, No. 1, (January 1967), pp. 115–119.
12. Stoer, Josef, "Duality in Nonlinear Programming and the Minimax Theorem," *Numerische Mathmatik*, Vol. 5, (1963), pp. 371–379.
13. Symonds, G., "Stochastic Scheduling—Part I" Tech. Memo. No. 22, Operations Research Group, Case Institute of Technology, September 1964, Cleveland, Ohio.
14. ——, P. Doulliez, C. Steinmetz, and T. S. Chidambaram, "Stochastic Scheduling—Part IV," Tech Memo No. 39, Operations Research Group, Case Institute of Technology, August 1965, Cleveland, Ohio.
15. Wagner, H. and B. Miller, "Chance Constrained Programming with Joint Constraints," *Operations Research*, Vol. 13, No. 6, (November–December, 1966), pp. 930–945.