

GENERALIZED CONVEXITY IN MATHEMATICAL PROGRAMMING

B. MOND

The role of convexity in the development of mathematical programming is reviewed. Many recent generalizations of convexity and their applications to optimization theory are discussed.

1. Introduction

As is well known, convexity plays a key role in mathematical programming. However, functions that are, in some sense, close to, but not quite, convex often have many of the important properties of convex functions. This, plus the fact that many practical problems involve functions that are 'almost' convex has led to many generalizations of the notion of convex functions. Here we survey some of these extensions of convexity and indicate briefly the role they play in optimization theory.

Most of the results in Sections 2 to 5 can be found in the books by Avriel [1], Mangasarian [16] and Martos [17] while the results in Section 6 on cone-convexity can be found in the book by Craven [8]. In view of this, no additional references or original sources will be given for results in these sections except for the new dual for non-linear programs in Section 5 that does not appear in any of the above-mentioned books. More detailed references will be given in the later sections.

Proofs will only be given where they are extremely brief and not generally available in the literature. It should be noted that most proofs

Received 7 October 1982. An invited lecture delivered to the Annual Congress of the South African Mathematical Society, October 1982.

utilize the differentiable form (2) of the definition of convexity.

2. Convexity

Let f denote a function from a convex set $C \subset \mathbb{R}^n$ into \mathbb{R} . f is said to be a *convex* function if for all $x, y \in C$, $0 \leq \lambda \leq 1$,

$$(1) \quad f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda)f(x).$$

If f is differentiable, an alternate and equivalent definition is

$$(2) \quad f(y) - f(x) \geq (y-x)^t \nabla f(x)$$

where $\nabla f(x)$ denotes the gradient (column) vector of partial derivatives of f with respect to x . Whereas (1) says that linear approximation over-estimates the value of the function, (2) says that a tangent to f at x lies on or below f .

By the *epigraph* of f , denoted by $E(f)$, we mean the points in \mathbb{R}^{n+1} that are on or above f ; that is,

$$(3) \quad E(f) = \{(x, \alpha) : x \in C, \alpha \in \mathbb{R}, f(x) \leq \alpha\}.$$

An alternate definition of a convex function is the following:

f is a convex function if and only if its epigraph is a convex set.

By the *lower level sets* of f we mean

$$(4) \quad L(f, \alpha) = \{x : x \in C, f(x) \leq \alpha\} \text{ for given } \alpha.$$

If f is a convex function, its lower level sets are convex for every $\alpha \in \mathbb{R}$. This is necessary but not sufficient for f to be convex. In order to obtain necessary and sufficient conditions for the convexity of f involving lower level sets, we define the generalized lower level sets of f as follows:

$$(5) \quad GL(f, \xi, \alpha) = \{x : x \in C, f(x) \leq \xi^t x + \alpha\}.$$

f is convex if and only if its generalized lower level sets $GL(f, \xi, \alpha)$ are convex sets for all $\xi \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

Another approach to convex functions, related to the definition (2), is the fact that convex functions are generated by the family of affine functions in the sense that a convex function f is the pointwise supremum

of affine functions, or equivalently, that the epigraph of f is supported at each point by the graph of an affine function.

Some properties of convex functions that we note are the following.

A local minimum of a convex function is always a global minimum. (This is a necessary but not a sufficient condition.) The class of convex functions is closed under addition and nonnegative scalar multiplication. Finally, the set of points for which a convex function attains a minimum is a convex set.

3. Applications of convexity

Consider the nonlinear programming problem

$$(6) \quad \text{minimize } f(x) \text{ subject to } g(x) \leq 0$$

where f is a differentiable function from R^n into R and g a differentiable function from R^n into R^m .

Kuhn and Tucker gave the following necessary conditions for feasible x_0 to be an optimal solution. If a constraint qualification is satisfied at optimal x_0 , then there exists $y \in R^m$ such that

$$(7) \quad \nabla y^t g(x_0) + \nabla f(x_0) = 0, \quad y^t g(x_0) = 0, \quad y \geq 0.$$

We now indicate the role convexity plays in

- (a) sufficiency of the Kuhn-Tucker conditions (7),
- (b) constraint qualifications for the necessary Kuhn-Tucker conditions,
- (c) dual programs of (6),
- (d) the relation between the existence of an optimal solution of (6) and the existence of a saddlepoint of the Lagrangean.

The relevant results are as follows.

- (a) If f and g are convex, then the existence of feasible x_0 and $y \in R^m$ such that the Kuhn-Tucker conditions are satisfied is sufficient for x_0 to be optimal for (6).

(b) (Generalized Slater Constraint Qualification.) If g is convex and, for all non-linear g_i , there exists feasible \bar{x} such that

$g_i(\bar{x}) < 0$, then for any optimal x_0 , there exists $y \in R^m$ satisfying (7).

(c) Wolfe gives the following dual of (6):

(8) maximize $f(u) + y^t g(u)$ subject to $\nabla y^t g(u) + \nabla f(u) = 0$, $y \geq 0$.

(6) and (8) are related as follows.

(Weak duality.) If f and g are convex, then for any feasible x of (6) and (y, u) of (8),

$$f(x) \geq f(u) + y^t g(u).$$

(Strong duality.) If, also, x_0 is optimal for (6), and a constraint qualification is satisfied at x_0 , then there exists optimal (y, u) , with $u = x_0$ and

$$(9) \quad f(x_0) = f(u) + y^t g(u).$$

(Converse duality.) If f and g are convex, (y_0, u_0) is optimal for (8) and the matrix

$$(10) \quad \nabla^2 [y_0^t g(x_0) + f(x_0)]$$

is non-singular, then $x = u_0$ is optimal for (6) and (9) is satisfied.

Another kind of duality involving conjugate functions is described in Avriel [1]. Here, too, convexity plays a fundamental role.

(d) The *Lagrangian* of (6) is defined by

$$(11) \quad L(x, y) \equiv f(x) + y^t g(x) \quad \text{for } y \geq 0.$$

(x^*, y^*) , $y^* \geq 0$, is said to be a *saddlepoint* of (11) if

$$(12) \quad L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \quad \text{for all } x, y \quad (y \geq 0).$$

The saddlepoint problem is related to the programming problem (6) as follows.

If (x^*, y^*) is a saddlepoint of $L(x, y)$, then x^* is an optimal solution of (6). If x^* is optimal for (6), a constraint qualification is satisfied at x^* , f and g are convex, then there exists y^* such that (x^*, y^*) is a saddlepoint of $L(x, y)$.

It is noteworthy that for the Kuhn-Tucker conditions convexity is only required for the sufficiency, not the necessary part of the theorem. By contrast, for the saddlepoint problem, sufficiency holds without any restrictions whatsoever; while convexity is needed for the necessary part of the result.

4. Pseudo-convex and quasi-convex functions

A differentiable function f is said to be *pseudo-convex* if, for all x, y in its domain,

$$(13) \quad (y-x)^t \nabla f(x) \geq 0 \Rightarrow f(y) - f(x) \geq 0.$$

An example of a pseudo-convex, but not convex, function is $f(x) \equiv x^3 + x$. A function f is said to be *quasi-convex*, if for all x, y in its domain and all $\lambda \in [0, 1]$ such that $\lambda x + (1-\lambda)y$ is in its domain,

$$(14) \quad f(\lambda x + (1-\lambda)y) \leq \max[f(y), f(x)].$$

If f is defined over a convex set, (14) is equivalent to

$$(15) \quad \text{all lower level sets of } f \text{ are convex.}$$

If f is differentiable it is quasi-convex if and only if

$$f(y) - f(x) \leq 0 \Rightarrow (y-x)^t \nabla f(x) \leq 0.$$

An example of a function that is quasi-convex, but not convex or pseudo-convex, is $f(x) = x^3$. If

$$(16) \quad f(x) \neq f(y) \Rightarrow f(\lambda x + (1-\lambda)y) < \max[f(x), f(y)], \quad 0 < \lambda < 1,$$

the function f is said to be *strictly quasi-convex* or *explicitly quasi-convex* (other names in the literature include semistrictly quasi-convex and functionally convex). If f is convex, it is pseudo-convex, strictly quasi-convex and quasi-convex.

Unlike convex functions, the sum of two pseudo-convex or two quasi-convex functions need not be pseudo-convex or quasi-convex, respectively.

An example of two pseudo-convex functions whose sum is not pseudo-convex are $x^3 + x$ and $-x$. If f is pseudo-convex on a convex set, it is strictly quasi-convex and quasi-convex. An example of a function (not defined on a convex set) that is pseudo-convex but not quasi-convex is $f(x) = x - x^2$ where $x = \{0, 1\}$.

If f is strictly quasi-convex and lower semi-continuous on a convex set, it is quasi-convex. An example of a function on a convex set that is strictly quasi-convex but not quasi-convex is

$$f(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Some of the importance of pseudo-convex and quasi-convex functions can be seen from the following result.

THEOREM. *If f is pseudo-convex or strictly quasi-convex, a local minimum is a global minimum.*

5. Application of pseudo-convex and quasi-convex functions

The use of pseudo-convex and quasi-convex, instead of convex, functions allows the weakening of the convexity requirements listed in Section 3 as follows.

(a) The sufficiency of the Kuhn-Tucker conditions (7) holds if f is pseudo-convex and each g_i such that $g_i(x_0) = 0$ is quasi-convex.

(b) The generalized Slater constraint qualification holds if all constraints are pseudo-convex and there exists a feasible \bar{x} such that, for all non-linear g_i , $g_i(\bar{x}) < 0$.

(c) (i) Converse duality holds if f is pseudo-convex and all g_i are quasi-convex. Weak and strong duality do not hold under these conditions as can be seen by the following counterexample:

$$(17) \quad \min x^3 + x, \text{ subject to } -x \leq -1.$$

The optimal is at $x = 1$, whereas the value of the Wolfe dual

$$(18) \quad \max u^3 + u + y(-u+1) \text{ subject to } 3u^2 + 1 - y = 0, y \geq 0,$$

is unbounded.

(ii) Weak and strong duality do hold if the Lagrangean, $f + y^t g$, is pseudo-convex. We give the simple proof for weak duality

$$(x-u)^t \nabla [y^t g(u) + f(u)] = 0 \Rightarrow y^t g(x) + f(x) - y^t g(u) - f(u) \geq 0 .$$

Therefore $f(x) \geq f(u) + y^t g(u)$.

(d) The requirements of the convexity of f and g in order for an optimal x^* of (6) to be part of a saddlepoint solution (12) can be weakened to the pseudo-convexity of the Lagrangean.

As pointed out, when dealing with the Wolfe dual (8), weak and strong duality require more stringent convexity requirements than converse duality. In order to lessen these convexity requirements, Mond and Weir [19] recently proposed the following dual to (6):

$$(19) \quad \max f(u) \quad \text{subject to} \quad \nabla y^t g(u) + \nabla f(u) = 0 , \\ y^t g(u) \geq 0 , \quad y \geq 0 .$$

THEOREM (Weak duality). *If for all feasible x of (6) and (y, u) of (19), f is pseudo-convex and $y^t g$ is quasi-convex, then*

$$f(x) \geq f(u) .$$

Proof. $y^t g(x) - y^t g(u) \leq 0 \Rightarrow (x-u)^t \nabla y^t g(u) \leq 0$. Therefore $(x-u)^t \nabla f(u) \geq 0 \Rightarrow f(x) \geq f(u)$.

THEOREM [19] (Strong duality). *If, also, x_0 is optimal for (6) and a constraint qualification is satisfied at x_0 , then there exists optimal (y, u) , with $u = x_0$, and the optimal values of primal and dual are equal.*

Thus, this new dual of (17) would be

$$(20) \quad \max u^3 + u \quad \text{subject to} \quad 3u^2 + 1 - y = 0 , \quad y(-u+1) \geq 0 , \quad y \geq 0 .$$

The optimal is attained at $u = 1$, $y = 4$.

6. Cone convexity

One extension of the different types of convexity we have encountered so far is to convex cones.

DEFINITION. $S \subset \mathbb{R}^m$ is a *convex cone* if

$$(21) \quad x, y \in S \Rightarrow \lambda x + \alpha y \in S \text{ for all } \lambda, \alpha \geq 0 .$$

Let S be a convex cone in \mathbb{C}^m . A function $g : \mathbb{C}^n \rightarrow \mathbb{C}^m$ is said to be *S-convex* if

$$(22) \quad -g(\lambda y + (1-\lambda)x) + \lambda g(y) + (1-\lambda)g(x) \in S$$

for all $x, y \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$.

If g is differentiable, this is equivalent to

$$(23) \quad g(y) - g(x) - (y-x)^t \nabla g(x) \in S .$$

The function g is said to be *S-pseudo-convex* if

$$(24) \quad (y-x)^t \nabla g(x) \in S \Rightarrow g(y) - g(x) \in S$$

and *S-quasi-convex* if

$$(25) \quad -g(y) + g(x) \in S \Rightarrow -(y-x)^t \nabla g(x) \in S .$$

Observe that, since the sum of two elements of S is also in S , an S -convex function is also S -pseudo-convex and S -quasi-convex. Note that if $S = \mathbb{R}_+^m$ (if g is a scalar function, $m = 1$), then S -convexity reduces to the usual definition of convexity.

Consider the problem

$$(26) \quad \text{minimize } f(x) \text{ subject to } -g(x) \in S$$

where S is a convex cone in \mathbb{R}^m . Its dual is

$$(27) \quad \text{maximize } f(u) + y^t g(u) \text{ subject to } \nabla y^t g(u) + \nabla f(u) = 0, \quad y \in S^*,$$

where S^* , the polar cone of S , is defined by

$$(28) \quad S^* = \{z \in \mathbb{R}^m : z^t s \geq 0 \text{ for all } s \in S\} .$$

In extending mathematical programming theory to problems with convex cones, it becomes necessary to distinguish between convex cones with an

infinite number of generators and polyhedral cones, that is, convex cones with a finite number of generators. In general, all results extend readily to problems with polyhedral cones. Thus, if x_0 is optimal to (26), the corresponding Kuhn-Tucker conditions are

$$(29) \quad \nabla y^t g(x_0) + \nabla f(x_0) = 0, \quad y \in S^*, \quad y^t g(x_0) = 0.$$

Whereas for polyhedral cones, the necessary conditions (29) always hold subject to an appropriate constraint qualification being satisfied, this is not true without further restrictions for cones with an infinite number of generators. Given that f is convex or pseudo-convex and g S -convex or S -quasi-convex, (29) and feasibility are sufficient for optimality. Similarly, if f is convex and g S -convex, all duality relationships between (26) and (27) hold if S is polyhedral. If S has an infinite number of generators, only weak duality holds without further restrictions.

7. Generalized convexity via generalized means

There is another natural generalization of convex functions that is rediscovered from time to time. Noting that the right hand side of the inequality

$$(30) \quad f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda)f(x), \quad 0 \leq \lambda \leq 1,$$

is just the weighted arithmetic mean of $f(y)$ and $f(x)$, the generalization is obtained by substituting other weighted means for the right side of (30).

Before formally stating the generalized convexity condition, we review some classical definitions and results involving generalized means. Let us assume, for the moment, that both $f(y)$ and $f(x)$ are positive. Then with $0 \leq \lambda \leq 1$, the r th mean of $f(y)$ and $f(x)$ is defined as follows:

$$(31) \quad M_r(f(y), f(x); \lambda) = \{\lambda[f(y)]^r + (1-\lambda)[f(x)]^r\}^{1/r} \quad \text{if } r \neq 0.$$

(Henceforth, we shall write $M_r(f; \lambda)$ for $M_r(f(y), f(x); \lambda)$.)

$$M_0(f(y), f(x); \lambda) = \lim_{r \rightarrow 0} M_r(f; \lambda) = [f(y)]^\lambda [f(x)]^{1-\lambda} ,$$

$$M_\infty(f(y), f(x); \lambda) = \lim_{r \rightarrow \infty} M_r(f; \lambda) = \max[f(y), f(x)] ,$$

$$M_{-\infty}(f(y), f(x); \lambda) = \lim_{r \rightarrow -\infty} M_r(f; \lambda) = \min[f(y), f(x)] .$$

THEOREM [14]. (a) If $f(y) = f(x)$, then $M_r(f; \lambda) = f(y) = f(x)$, $-\infty \leq r \leq \infty$.

(b) If $f(y) > f(x)$, then $f(y) > M_r(f; \lambda) > f(x)$, for all r, λ such that $-\infty < r < \infty$, $0 < \lambda < 1$.

(c) If $s > r$, then $M_s(f; \lambda) \geq M_r(f; \lambda)$ and the inequality is strict if $f(y) \neq f(x)$, $0 < \lambda < 1$.

Following the terminology of [2], we now extend the definition of convexity in (1) as follows. A function f is said to be r^+ convex if it satisfies, for all x, y in the domain of f ,

$$(32) \quad f(\lambda y + (1-\lambda)x) \leq M_r(f; \lambda) , \quad 0 \leq \lambda \leq 1 .$$

Note that (32) gives the usual definition of convexity for $r = 1$ and of quasi-convexity for $r = \infty$. Since $M_s(f; \lambda) \geq M_r(f; \lambda)$ for $s > r$, it follows that a function that is r^+ convex will also be s^+ convex for all $s > r$. Thus we have a continuous transition from the class of convex functions ($r = 1$) to the class of quasi-convex functions ($r = \infty$) via the intermediate class of r^+ convex functions with $1 < r < \infty$.

Recall now that we restricted our definition of $M_r(f; \lambda)$ to $f(y)$ and $f(x)$ positive. This was done so that $M_r(f; \lambda)$ will be defined for all r . In order to allow zero and negative values of f for all r , Avriel [2] and Martos [17] independently define r -convex functions as follows: f is said to be r -convex if for all r, λ , $-\infty \leq r \leq \infty$, $0 \leq \lambda \leq 1$, it satisfies

$$(33) \quad f(\lambda y + (1-\lambda)x) \leq \log M_r(e^{f(y)}, e^{f(x)}; \lambda) .$$

Thus

$$(34) \quad f(\lambda y + (1-\lambda)x) \leq \left. \begin{array}{ll} \log[\lambda e^{rf(y)} + (1-\lambda)e^{rf(x)}]^{1/r} & \text{if } r \neq 0 \\ \lambda f(y) + (1-\lambda)f(x) & \text{if } r = 0 \\ \max[f(y), f(x)] & \text{if } r = \infty \\ \min[f(y), f(x)] & \text{if } r = -\infty \end{array} \right\} .$$

Note that r -convexity, which is now no longer restricted to positive values of f , reduces to ordinary convexity when $r = 0$. As before r -convexity implies s -convexity for all $s > r$ and hence, for $0 < r < \infty$, (33) gives intermediary inequalities between convexity ($r = 0$) and quasi-convexity ($r = \infty$). In general, the sum of two r -convex functions or two r^+ -convex functions need not be r -convex or r^+ -convex.

Avriel [2] calls functions that satisfy (33) for $r < 0$, *superconvex* and, for $r > 0$, *subconvex*. Thus superconvexity implies convexity which implies subconvexity which in turn, implies quasi-convexity.

It is frequently difficult, algebraically, to deal with r -convex functions. However an r -convex function for finite r can be characterized in terms of ordinary convexity by the following result.

Let \hat{f} be defined by $e^{rf(x)}$. Then f is r convex with $r \neq 0$ if and only if \hat{f} is convex for $r > 0$ and concave for $r < 0$.

A technique for solving certain nonlinear programs involving r -convex constraints is given in Avriel [3]. In essence it involves solving a sequence of programs where the r -convex constraints are approximated by convex constraints.

A further extension of convexity by the use of generalized means is possible. Let θ be a continuous strictly increasing scalar function that includes $f(y)$ and $f(x)$ in its domain. Then f is said to be θ -convex if, for all x and y in the domain of f and all λ , $0 \leq \lambda \leq 1$,

$$(36) \quad f(\lambda y + (1-\lambda)x) \leq \theta^{-1}\{\lambda\theta[f(y)] + (1-\lambda)\theta[f(x)]\} .$$

Here θ^{-1} is the inverse function of θ . If $\theta(x) = x$, then (36)

reduces to (1), the usual definition of convexity. If $\theta(x) = x^r$ and $r \neq 0$ and $\theta(x) = \log x$ for $r = 0$, then (36) reduces to r^+ -convexity (32). If $\theta(x) = e^{rx}$ for $r \neq 0$ and $\theta(x) = x$ for $r = 0$, then

θ -convexity, (36), becomes r -convexity, (33).

So far we have extended the definition of convexity by generalizing the right side of (1). A further extension is possible by generalizing the left side of (1) as well. In essence, since $f(\lambda x+(1-\lambda)x)$, $0 \leq \lambda \leq 1$, consists of the values of f at all points on the straight line between x and y that are in the domain of f , we can consider other paths from x to y . Specifically, let $p_{x,y}(\lambda)$, where $p_{x,y}(0) = x$ and $p_{x,y}(1) = y$ represent a continuous path from x to y in R^n such that $f(p_{x,y}(\lambda))$, $0 \leq \lambda \leq 1$, is defined. Then f is said to be $(p-\theta)$ -convex if

$$(37) \quad f(p_{x,y}(\lambda)) \leq \theta^{-1}\{\lambda\theta[f(y)]+(1-\lambda)\theta[f(x)]\}$$

for all x and y in the domain of f , $0 \leq \lambda \leq 1$. A particular subclass of such functions where the path from x to y involves generalized mean value functions was explored in a most elegant way in Ben-Tal [5]. Specifically let h be a continuous one-to-one and onto function defined on a subset of R^n , including the domain of f and with values in R^n . Then f will be said to be $(h-\theta)$ -convex if, for all x, y , $0 \leq \lambda \leq 1$,

$$(38) \quad f(h^{-1}[\lambda h(y)+(1-\lambda)h(x)]) \leq \theta^{-1}\{\lambda\theta[f(y)]+(1-\lambda)\theta[f(x)]\}.$$

Some of the results obtained for $(h-\theta)$ -convex functions include the following [5], [1].

THEOREM. (a) *If f and g are $(h-\theta)$ -convex, then $\theta^{-1}\{\theta(f)+\theta(g)\}$ is also $(h-\theta)$ -convex.*

(b) *If f, h and θ are differentiable, then f is $(h-\theta)$ -convex if and only if, for all x, y ,*

$$\theta(f(y)) - \theta(f(x)) + \theta'(f(x)) \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f(y)}{\partial x_j} \frac{\partial h^{-1}(h(x))}{\partial x_i} (h_i(y)-h_i(x)).$$

This last result reduces to (2) if $\theta(f(x)) = f(x)$, $h(x) = x$. $(h-\theta)$ -convexity is related to ordinary convexity by the following result: f is $(h-\theta)$ -convex, if and only if \hat{f} given by

$$\hat{f}(y) = \theta(f(h^{-1}(y)))$$

is convex.

The application of $(h-\theta)$ convexity to nonlinear programming including some duality results can be found in [1] and [5].

8. Invex and sublinear functions

In the last section, the notion of convexity was extended by generalizing the definition (1). Another approach is to extend the definition of convexity (2) for differentiable functions. Hanson [12] introduced the class of functions that satisfies

$$(39) \quad f(y) - f(x) \geq h^t(x, y) \nabla f(x)$$

for all x, y and for some vector function h . Craven [9] calls such functions *invex* (for invariant convex) since $f = g \circ \theta$ will be invex if g is convex, θ is differentiable and θ' has full rank. Similarly, f is said to be *pseudo-invex* if, for some vector function $h(x, y)$ and all x, y ,

$$h^t(x, y) \nabla f(x) \geq 0 \Rightarrow f(y) - f(x) \geq 0$$

and *quasi-invex* if

$$f(y) - f(x) \leq 0 \Rightarrow h^t(x, y) \nabla f(x) \leq 0.$$

It follows that, by taking $h^t(x, y) = y - x$, convex functions are also invex and that invex functions are both pseudo-invex and quasi-invex.

Also, the sum of two functions that are invex for the same function $h(x, y)$ is also invex.

Hanson [12] proves that if, in (6), f and g are all invex for the same function h , the Kuhn-Tucker conditions (7) are sufficient. He also establishes duality between (6) and (8) when f and g are invex (for the same h) instead of convex. Indeed duality holds [13] if only the Lagrangean

$$(40) \quad L(x, y) = f(x) + y^t g(x), \quad y \geq 0,$$

is pseudo-invex. Weir [29] establishes duality between (6) and (19) if f is pseudo-invex and $y^t g$ is quasi-invex.

Mond and Hanson [18] and Craven [10] extend the concept of invex functions to convex cones. Thus, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be *S-invex*, where S is a convex cone in \mathbb{R}^m , if for all x and y in the domain of g , there exists some vector $h(x, y)$ such that

$$g(y) - g(x) - h^t(x, y)\nabla g(x) \in S.$$

S-pseudo-invex and *S-quasi-invex* functions can be defined in an analogous manner. This allows duality between (26) and (27) to be extended to *S-invex* functions.

Hanson and Mond [13] extend the generalization (39) still further by considering sublinear functionals. F is said to be sublinear if

$$(41) \quad (A) \quad F(a+b) \leq F(a) + F(b) \quad \text{for all } a, b,$$

$$(42) \quad (B) \quad F(\alpha x) = \alpha F(x) \quad \text{for all } x \text{ and every } \alpha \geq 0.$$

It follows from (B) that $F(0) = 0$.

Let f be a differentiable function satisfying

$$(43) \quad f(y) - f(x) \geq F_{y,x}[\nabla f(x)]$$

for all x, y and for some arbitrary given sublinear functional F . Sublinear functionals include linear functionals and, in particular, the special cases occurring on the right hand sides of (2) and (39). Thus (43) can be regarded as a generalization of convex and invex functions. Similarly corresponding to the definition of pseudo-convex (or pseudo-invex) and quasi-convex (or quasi-invex) functions one has the classes of functions satisfying

$$(44) \quad F_{y,x}[\nabla f(x)] \geq 0 \Rightarrow f(y) - f(x) \geq 0$$

and

$$(45) \quad f(y) - f(x) \leq 0 \Rightarrow F_{y,x}[f(x)] \leq 0.$$

Hanson and Mond [13] establish sufficiency of the Kuhn-Tucker conditions where there exists a sublinear functional F_{x,x_0} for all x such that

$$(44) \text{ is satisfied for } f \text{ and (45) for each } g_i \text{ such that } g_i(x_0) = 0.$$

Duality is established between (6) and (8) if the Lagrangean satisfies (44)

for some sublinear functional.

9. Additional generalizations

We mention briefly a number of other generalizations of convexity. Apparently based on an unpublished idea of C.R. Bector, Chandra [7], Lata [15] and Nehse [20] give duality results for functions satisfying

$$(46) \quad K(x, y)\{f(y)-f(x)\} \geq (y-x)^t \nabla f(x)$$

where $K(x, y)$ is an arbitrary positive scalar function. Such functions are called *strongly pseudo-convex*. If $K(x, y) = 1$, then (46) reduces to convexity (2). It follows, by taking

$$h(x, y) = [1/K(x, y)](y-x)$$

that strongly pseudo-convex functions are invex. Neshe [20] points out that the assumption in [7] and [15] that the sum of strongly pseudo-convex functions is strongly pseudo-convex is incorrect and, hence, that the duality results in [7] and [15] require further assumptions in order to be valid. The difficulty is that, unlike Hanson [12], Chandra [7] and Lata [15] do not require the same function $K(x, y)$ for all f and g_i .

Ben-Tal and Ben-Israel [6] generalize convexity by replacing the family of affine functions that support a convex function by a family of other functions. Let F be a family of functions $F : R^n \rightarrow R$. Then f is called *F-convex* if, for every x in the domain of f , there exists an $F \in F$ such that

$$f(x) = F(x) \quad \text{and} \quad f(z) \geq F(z) \quad \text{for all } x \neq z,$$

in which case F is a support of f at x . Although in general, for *F-convex* functions, a local minimum need not be a global minimum, Ben-Tal and Ben-Israel [6] do obtain some duality results for programs with *F-convex* functions.

In a recent paper, Doeringer [11] discusses *K-convex* functions. A function f on an interval I of the real line is called *K-convex*, where K is a nonnegative real number, if for any $x, y \in I$, $x < y$, and $0 \leq \lambda \leq 1$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)[f(y) + K].$$

It is clear that for $K = 0$, this becomes the usual definition of convexity (1). Although K -convex functions occur frequently in production and inventory problems, there has, so far, not been a great deal of research into such functions.

Finally, we point out that the survey of generalizations of convexity given here is not meant to be exhaustive. Indeed, Ponstein [21] lists seven different kinds of convexity (including variants of pseudo-convexity and quasi-convexity) and in [4] no less than nine different kinds of convexity, pseudo-convexity and quasi-convexity are given.

References

- [1] Mordecai Avriel, *Nonlinear programming, analysis and methods* (Prentice-Hall, Englewood Cliffs, New Jersey, 1976).
- [2] Mordecai Avriel, " r -convex functions", *Math. Programming* 2 (1972), 309-323.
- [3] M. Avriel, "Solution of certain nonlinear programs involving r -convex functions", *J. Optim. Theor. Appl.* 11 (1973), 159-174.
- [4] M. Avriel, W.E. Diewert, S. Schaible and W.T. Ziemba, "Introduction to concave and generalized concave functions", *Generalized concavity in optimization and economics*, 21-50 (Proc. NATO Advanced Study Institute, University of British Columbia, Vancouver, 1980. Academic Press [Harcourt Brace Jovanovich], New York and London, 1981).
- [5] A. Ben-Tal, "On generalized means and generalized convex functions", *J. Optim. Theor. Appl.* 21 (1977), 1-13.
- [6] A. Ben Tal and A. Ben Israel, " F -convex functions: properties and applications", *Generalized concavity in optimization and economics*, 301-334 (Proc. NATO Advanced Study Institute, University of British Columbia, Vancouver, 1980. Academic Press [Harcourt Brace Jovanovich], New York and London, 1981).
- [7] Suresh Chandra, "Strong pseudo-convex programming", *Indian J. Pure Appl. Math.* 3 (1972), 278-282.

- [8] B.D. Craven, *Mathematical programming and control theory* (Chapman and Hall, London; John Wiley & Sons, London; 1978).
- [9] B.D. Craven, "Duality for generalized convex fractional programs", *Generalized concavity in optimization and economics*, 473-489 (Proc. NATO Advanced Study Institute, University of British Columbia, Vancouver, 1980. Academic Press [Harcourt Brace Jovanovich], New York and London, 1981).
- [10] B.D. Craven, "Invex functions and constrained local minima", *Bull. Austral. Math. Soc.* 24 (1981), 357-366.
- [11] W. Doeringer, "A note on K -convex functions", *Z. Oper. Res. Ser. A-B* 26 (1982), 49-55.
- [12] Morgan A. Hanson, "On sufficiency of the Kuhn-Tucker conditions", *J. Math. Anal. Appl.* 80 (1981), 545-550.
- [13] M.A. Hanson, B. Mond, "Further generalizations of convexity in mathematical programming", *J. Inform. Optim. Sci.* 3 (1982), 25-32.
- [14] G.H. Hardy, J.E. Littlewood, G. Pólya, *Inequalities* (Cambridge University Press, Cambridge, 1959).
- [15] Manju Lata, "Strong pseudo-convex programming in Banach space", *Indian J. Pure Appl. Math.* 6 (1976), 45-48.
- [16] Olvi L. Mangasarian, *Nonlinear programming* (McGraw-Hill, New York, London, Sydney, 1969).
- [17] Béla Martos, *Nonlinear programming: theory and methods* (North-Holland, Amsterdam, Oxford; American Elsevier, New York; 1975).
- [18] B. Mond and M.A. Hanson, "On duality with generalized convexity" (Pure Mathematics Research Paper, 80-3. La Trobe University, Bundoora, Australia, 1980).
- [19] B. Mond and T. Weir, "Generalized concavity and duality", *Generalized concavity in optimization and economics*, 263-279 (Proc. NATO Advanced Study Institute, University of British Columbia, Vancouver, 1980. Academic Press [Harcourt Brace Jovanovich], New York and London, 1981).

- [20] Reinhard Nehse, "Strong pseudoconvex mappings in dual problems",
Math. Operationsforsch. Statist. Ser. Optim. 12 (1981), 483-491.
- [21] J. Ponstein, "Seven kinds of convexity", *SIAM Rev.* 9 (1967), 115-119.
- [22] Terence Weir, "Generalized convexity and duality in mathematical programming" (PhD thesis, La Trobe University, Bundoora, Australia, 1982).

Department of Pure Mathematics,
La Trobe University,
Bundoora,
Victoria,
Australia 3168;

Department of Applied Mathematics,
University of the Witwatersrand,
Johannesburg,
South Africa.