PACIFIC JOURNAL OF MATHEMATICS
Vol. 25, No. 3, 1968

A GENERAL CORRESPONDENCE BETWEEN DUAL MINIMAX PROBLEMS AND CONVEX PROGRAMS

R. TYRRELL ROCKAFELLAR

The Kuhn-Tucker theory of Lagrange multipliers centers on a one-to-one correspondence between nonlinear programs and minimax problems. This correspondence has been extended by Dantzig, Eisenberg and Cottle to one in which every minimax problem of a certain type gives rise to a pair of nonlinear programs dual to each other. The aim here is to show how, by forming conjugates of convex functions and saddle-functions (i.e. functions of two vector arguments which are convex in one argument and concave in the other), one can set up a more symmetric correspondence with even stronger duality properties. The correspondence concerns problems in quartets, each quartet being comprised of a dual pair of convex and concave programs and a dual pair of minimax problems. The whole quartet can be generated directly from any one of its members.

Our results grow out of a rather surprising fact, which we establish as Theorem 1: saddle-functions are really just convex functions to which Fenchel's conjugate operation has only been partly applied, i.e. only in some of the variables. In fact, there is a canonical one-to-one correspondence between the (minimax) equivalence classes of closed convex-concave functions K on $R^m \times R^n$ and the closed convex functions F on R^{m+n}. (The values of these functions are allowed to be, not only real numbers, but $+\infty$ and $-\infty$, so that, as explained in [8] and [9], there is no loss of generality in considering only functions which are everywhere defined on the given space.) The closed convex functions F on R^{m+n} can be paired with the closed concave functions G on R^{m+n} by reversing signs in the conjugacy correspondence for convex functions. At the same time, the equivalence classes of closed convex-concave functions K on $R^m \times R^n$ can be paired with the equivalence classes of closed concave-convex functions J or $R^m \times R^n$ by changing certain signs in the conjugacy correspondence for saddle-functions in [8].

Given four functions F, G, K, J corresponding to each other in this fashion we consider the problems

(I) minimize $F(x, v)$ subject to $x \geqq 0, v \geqq 0$;
(II) maximize $G(u, y)$ subject to $u \geqq 0, y \geqq 0$;
(III) minimaximize $K(x, y)$ subject to $x \geqq 0, y \geqq 0$;
(IV) maximinimize $J(u, v)$ subject to $u \geqq 0, v \geqq 0$.

(The formal definition of these problems is given in § 5. In (III) one minimizes in x and maximizes in y, whence "minimaximize." In (IV) one maximizes in u and minimizes in v.)

A concept of "stable solution" is introduced for these problems. We prove that, under very mild restrictions, all solutions are stable. The most interesting result is a *double duality theorem*: if any of the four problems has a stable solution, then all four have stable solutions and the four extrema are equal. Other theorems give criteria for the existence of solutions and characterize these solutions in terms of subdifferentials and complementary slackness.

It should be emphasized that constraints other than nonnegativity may be incorporated into the problems above by the device of infinity values. For example, let $f_0, f_1, \cdots, f_m$ be finite convex functions on R^m, and for $v = (\mu_1, \cdots, \mu_m) \in R^m$ let

$$F(x, v) = \begin{cases} f_0(x) \text{ if } f_i(x) + \mu_i \leqq 0 \text{ for } i = 1, \cdots, m, \\ +\infty \text{ if not.} \end{cases}$$

Then F is a closed convex function on R^{m+n}, and minimizing $F(x, v)$ subject to $x \geqq 0$ and $v \geqq 0$ as in (I) amounts to minimizing $f_0(x)$ subject to the constraints $x \geqq 0, f_i(x) \leqq 0, i = 1, \cdots, m$.

The nonnegativity conditions in our problems involve no real loss of generality, of course. As is well known, a free variable can always be expressed as a difference of two nonnegative variables, just as a linear equation can be expressed as a pair of weak linear inequalities, and these situations are dual to each other. If, for example, one wants to remove the nonnegativity constraint from the first component of x in (I), one has to remove the nonnegativity constraint likewise from the first component of x in (III) and at the same time strengthen the constraint on the first component of u (the variable dual to the first component of x in the sense of the complementary slackness conditions in Theorem 4) from $\geqq 0$ to $= 0$ in (II) and (IV). The theorems below are then applicable to the modified problems. This follows exactly as in the theory of linear programs.

Although in this paper we discuss only a four-way correspondence, a more extensive correspondence is actually implied. In problem (I), we have a function F on a space R^N, where each vector of R^N is decomposed into a component $x \in R^m$ and a component $v \in R^n$. Now there is nothing unique about this decomposition: we could just as easily partition the N canonical coordinates in R^N in some other way, so that each vector is decomposed into a component $x' \in R^{m'}$ and $v' \in R^{n'}$, $m' + n' = N$. This would have no effect on the dual problem (II), but it would lead to entirely different minimax problems (III) and (IV). Thus what we really have is a theory which represents (I) and (II) in a finite number of different ways as a dual pair of minimax problems.

We shall show elsewhere that such representations are closely related to the "simplex tableaux" encountered in the pivotal theory of linear and nonlinear programs.

2. Skew-conjugate functions.

We begin with a quick review of terminology. The main object is to set down various formulas involving the "skew-conjugate" operation. This is the same as the conjugate operation except for certain changes of sign, but it offers so many notational advantages in this particular context that we feel it warrants some explicit attention.

A *convex function* on f on R^m is, in our terminology, an everywhere-defined function f with values in the extended real interval $[-\infty, +\infty]$ such that

$$f(\lambda x_1 + (1-\lambda)x_2) \leqq \lambda\mu_1 + (1-\lambda)\mu_2$$

whenever

$$f(x_1) \leqq \mu_1 \in R, f(x_2) \leqq \mu_2 \in R\,, \qquad 0 < \lambda < 1\,.$$

The *closure* of such a function f is the function cl f defined by

$$(2.1)\qquad \begin{aligned} &(\mathrm{cl}\, f)(x) = \liminf_{z \to x} f(z) \quad (\text{if } f \text{ nowhere has the value } -\infty)\,, \\ &(\mathrm{cl}\, f)(x) = -\infty \text{ for all } x \text{ (if } f \text{ has the value } -\infty \text{ somewhere)}\,. \end{aligned}$$

When $\mathrm{cl}\, f = f$, we say f is *closed*. In particular, $\mathrm{cl}\, f$ is itself a closed convex function (cf. [3]).

A function g is *concave* if $-g$ is convex. The closure operation for concave functions is of course defined by

$$(2.2)\qquad \begin{aligned} &(\mathrm{cl}\, g)(x) = \limsup_{z \to x} g(z) \quad (\text{if } g \text{ nowhere has the value } +\infty)\,, \\ &(\mathrm{cl}\, g)(x) = +\infty \text{ for all } x \text{ (if } g \text{ has the value } +\infty \text{ somewhere)}\,. \end{aligned}$$

Again, g is *closed* if $\mathrm{cl}\, g = g$.

For any convex function f on R^m, the function g defined by

$$(2.3)\qquad g(u) = \inf_x \{f(x) - \langle x, u\rangle\}\,,$$

where $\langle x, u\rangle$ denotes the ordinary inner product of vectors x and u, will be called the *skew-conjugate* of f. It is really just the negative of the conjugate of f in [3] and [4]. Hence g is a closed concave function and

$$(2.4)\qquad (\mathrm{cl}\, f)(x) = \sup_u \{g(u) + \langle u, x\rangle\}\,.$$

Dually, starting with any concave function g on R^m, the skew-conjugate of g is the closed convex function f on R^m defined by

(2.5) $$f(x) = \sup_u \{g(u) + \langle u, x\rangle\} .$$

The skew-conjugate of f is then in turn $\mathrm{cl}\, g$. In particular, formulas (2.3) and (2.5) set up a one-to-one correspondence between the closed convex functions f on R^m and the closed concave functions g on R^m.

An everywhere-defined extended-real-valued function K on $R^m \times R^n$ is called a *convex-concave saddle-function* if $K(\cdot, y)$ is a convex function on R^m for each y, and $K(x, \cdot)$ is a concave function on R^n for each x. We denote by $\mathrm{cl}_1 K$ the function constructed by closing K in its first argument for each fixed value of the second argument, and similarly $\mathrm{cl}_2 K$. Both $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 K$ are again convex-concave saddle-functions, as we proved in [8]. The function F on R^{m+n} obtained by taking the skew-conjugate of K in its concave argument, i.e.

(2.6) $$F(x, v) = \sup_y \{K(x, y) + \langle v, y\rangle\} ,$$

will be called the *convex parent* of K. Dually, the *concave parent* G of K is defined by applying the skew-conjugate operation to the convex argument:

(2.7) $$G(u, y) = \inf_x \{K(x, y) - \langle x, u\rangle\} .$$

Two saddle-functions are said to be (*minimax*) *equivalent* if they have the same parent functions. A saddle-function K is *closed* if it is equivalent to $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 K$. (This differs from the definition in [8], but is equivalent to it by Corollary 2 to Theorem 1 in that paper.) The situation for concave-convex saddle-functions is virtually the same —the roles of the arguments are reversed.

As an important example [8] of an equivalence class of closed saddle-functions, let K be any finite continuous convex-concave function defined on $A \times B$, where A is a nonempty closed convex set in R^m and B is a nonempty closed convex set in R^n. Define $\underline{K}$ and $\bar{K}$ by

$$\underline{K}(x, y) = \begin{cases} K(x, y) & \text{if } x \in A, y \in B , \\ +\infty & \text{if } x \notin A, y \in B , \\ -\infty & \text{if } y \notin B \end{cases}$$

$$\bar{K}(x, y) = \begin{cases} K(x, y) & \text{if } x \in A, y \in B , \\ -\infty & \text{if } x \in A, y \notin B , \\ +\infty & \text{if } x \notin A . \end{cases}$$

The equivalence class consists of $\underline{K}$ and $\bar{K}$ and all the other convex-concave extensions of K to $R^m \times R^n$ lying between $\underline{K}$ and $\bar{K}$. In particular, if $A = R^m$ and $B = R^n$ then K is a closed saddle-function and is the sole member of its equivalence class.

The parents F and G of a convex-concave saddle-function K can be expressed in terms of the gradients of K in some cases. Suppose that K is finite and differentiable everywhere on $R^m R^n$. Let $\nabla_1 K(x, y)$ denote the gradient of $K(\cdot, y)$ at x, and let $\nabla_2 K(x, y)$ denote the gradient of $K(x, \cdot)$ at y. It can be seen that

$$F(x, v) = K(x, y) - \langle \nabla_2 K(x, y), y \rangle \quad \text{for } v = -\nabla_2 K(x, y)\,,$$
$$G(u, y) = K(x, y) - \langle \nabla_1 K(x, y), x \rangle \quad \text{for } u = \nabla_1 K(x, y)\,.$$

These gradient expressions are the basis of the duality theory in [2].

Given any closed convex-concave saddle-function K on $R^m \times R^n$, the functions

$$(2.8) \qquad \begin{aligned} \bar{J}(u, v) &= \inf_x \sup_y \{K(x, y) - \langle x, u \rangle + \langle v, y \rangle\}\,, \\ \underline{J}(u, v) &= \sup_y \inf_x \{K(x, y) - \langle x, u \rangle + \langle v, y \rangle\}\,, \end{aligned}$$

are closed concave-convex saddle-functions on $R^m \times R^n$ equivalent to each other, and they depend only on the equivalence class of K, as shown in [8]. Any function J in the class containing $\bar{J}$ and $\underline{J}$ will be called a *skew-conjugate* of K. For such a J, the functions

$$(2.9) \qquad \begin{aligned} \bar{K}(x, y) &= \inf_v \sup_u \{J(u, v) + \langle u, x \rangle - \langle y, v \rangle\} \\ \underline{K}(x, y) &= \sup_u \inf_v \{J(u, v) + \langle u, x \rangle - \langle y, v \rangle\} \end{aligned}$$

belong in turn to the equivalence class containing K. In this way, we get a skew-conjugate correspondence between the closed convex-concave saddle-functions on $R^m \times R^n$ and the closed concave-convex saddle-functions on $R^m \times R^n$ which is one-to-one up to equivalence.

3. Four-way correspondence.

It will be shown in this section that each equivalence class of closed saddle-functions on $R^m \times R^n$ is generated by a closed convex function on R^{m+n}. The first assertion of Lemma 1 has already been noted by Moreau [6].

LEMMA 1. *Let K be any saddle-function on $R^m \times R^n$, let F be its convex parent, and let G be its concave parent. Then F is a convex function on R^{m+n} and G is a concave function on R^{m+n}. Moreover, K is closed if and only if F and G are skew-conjugate to each other.*

Proof. We may assume K is convex-concave. By definition (2.6), F is a supremum of convex functions on R^{m+n}, one for each given value of y. The convexity of F is an easy consequence of this. One proves in the same way that G is concave. Now, since $F(x, \cdot)$ is the skew-conjugate of $K(x, \cdot)$ for each fixed x, the skew-conjugate of $F(x, \cdot)$ is in turn the closure of $K(x, \cdot)$. Thus

$$(\mathrm{cl}_2 K)(x, y) = \inf_v \{F(x, v) - \langle y, v\rangle\} \tag{3.1}$$

$$F(x, v) = \sup_y \{\mathrm{cl}_2 K(x, y) + \langle v, y\rangle\} . \tag{3.2}$$

For parallel reasons,

$$(\mathrm{cl}_1 K)(x, y) = \sup_u \{G(u, y) + \langle u, x\rangle\} \tag{3.3}$$

$$G(u, y) = \inf_x \{\mathrm{cl}_1 K(x, y) - \langle x, u\rangle\} . \tag{3.4}$$

According to (3.2) and (3.4), F is the convex parent of $\mathrm{cl}_2 K$ and G is the concave parent of $\mathrm{cl}_1 K$, always. Thus K is closed if and only if F is the convex parent of $\mathrm{cl}_1 K$ and G is the concave parent of $\mathrm{cl.} K$. That would mean by (3.1) and (3.3) that

$$F(x, v) = \sup_y \{\sup_u \{G(u, y) + \langle u, x\rangle\} + \langle v, y\rangle\} ,$$
$$G(u, y) = \inf_x \{\inf_v \{F(x, v) - \langle y, v\rangle\} - \langle x, u\rangle\} .$$

But this says that F and G are skew-conjugate as functions on R^{m+n}.

LEMMA 2. *Let F and G be functions on R^{m+n} skew-conjugate to each other, where F is convex and G is concave. Let*

$$\bar{K}(x, y) = \inf_v \{F(x, v) - \langle y, v\rangle\} ,$$
$$\underline{K}\{x, y) = \sup_u \{G(u, y) + \langle u, x\rangle\} .$$

Then $\bar{K}$ and $\underline{K}$ are convex-concave saddle-functions on $R^m \times R^n$ having F and G as parents (and hence $\bar{K}$ and $\underline{K}$ are in the same equivalence class).

Proof. Since $\bar{K}(x, \cdot)$ is the skew conjugate of the convex function $F(x, \cdot)$, it is concave. We must also prove for each y that $\bar{K}(\cdot, y)$ is a convex function. In other words, given

$$\bar{K}(x_1, y) \leqq \mu_1 \in R, \bar{K}(x_2, y) \leqq \mu_2 \in R , \qquad 0 < \lambda < 1 , \tag{3.5}$$

we must show that

$$K(\lambda x_1 + (1 - \lambda)x_2, y) \leqq \lambda\mu_1 + (1 - \lambda)\mu_2 . \tag{3.6}$$

For an arbitrary $\varepsilon > 0$, (3.5) and the definition of $\bar{K}$ imply the existence of some v_1 and v_2 in R^n such that

$$\mu_1 + \varepsilon \geqq F(x_1, v_1) - \langle y, v_1\rangle ,$$
$$\mu_2 + \varepsilon \geqq F(x_2, v_2) - \langle y, v_2\rangle .$$

Now F is convex, so it follows that

$$\lambda[\mu_1 + \varepsilon + \langle y, v_1\rangle] + (1 - \lambda)[\mu_2 + \varepsilon + \langle y, v_2\rangle]$$
$$\geqq F(\lambda x_1 + (1 - \lambda)x_2, \lambda v_1 + (1 - \lambda)v_2) .$$

Hence, for $v = \lambda v_1 + (1-\lambda)v_2$,

$$\lambda\mu_1 + (1-\lambda)\mu_2 + \varepsilon \geqq F(\lambda x_1 + (1-\lambda)x_2, v) - \langle y, v\rangle .$$

Since ε was arbitrary, this implies (3.6). Thus $\bar{K}$ is a saddle-function. As the skew-conjugate of G, F is closed. In particular, $F(x, \cdot)$ is closed for each x, and hence it is in turn the skew-conjugate of its skew-conjugate $\bar{K}(x, \cdot)$. Therefore F is the convex parent of $\bar{K}$. The concave parent of $\bar{K}$, on the other hand, is given by

$$\inf_x \{\bar{K}(x, y) - \langle x, \mu\rangle\} = \inf_{x,v}\{F(x, v) - \langle x, u\rangle - \langle y, v\rangle\} ,$$

which is just $G(u, y)$. The proof for $\underline{K}$ is analogous.

Lemmas 1 and 2 imply that every F and G skew-conjugate to each other on R^{m+n} are the parents of a unique equivalence class of closed convex-concave saddle-functions K on $R^m \times R^n$. But the same arguments must work for concave-convex saddle-functions, too. Specifically, suppose F and G are the convex and concave parents of the closed convex-concave saddle-function K. Then, as in Lemma 2, the functions

$$\begin{aligned} \bar{J}(u, v) &= \inf_x \{F(x, v) - \langle x, u\rangle\} , \\ \underline{J}(u, v) &= \sup_y \{G(u, y) + \langle v, y\rangle\} . \end{aligned} \tag{3.7}$$

will be closed concave-convex saddle-functions on $R^m \times R^n$ having F and G as parents. Substituting (2.6) and (2.7), we see that $\bar{J}$ and $\underline{J}$ are the skew-conjugates of K defined in (2.8). This proves the following theorem.

THEOREM 1. *There is a canonical four-way one-to-one correspondence between the closed convex functions F on R^{m+n}, the closed concave functions G on R^{m+n}, the (equivalence classes of) closed convex-concave saddle-functions K on $R^m \times R^n$ and the (equivalence classes of) closed concave-convex saddle functions J on $R^m \times R^n$. The relationships in this correspondence are that F and G are skew-conjugate to each other, K and J are skew-conjugate to each other, and F and G are the parents of K and J.*

4. Effective domains and subdifferentials. Henceforth we assume that F, G, K and J are four functions corresponding to each other in the manner described in Theorem 1. We shall also assume these functions are proper, i.e. we exclude the case in Theorem 1 where all four functions are identically $+\infty$ and the case where all four are $-\infty$. Then the sets

$$\begin{aligned} \operatorname{dom} F &= \{(x, v) \mid F(x, v) < +\infty\} , \\ \operatorname{dom} G &= \{(u, y) \mid G(u, y) > -\infty\} . \end{aligned}$$

are nonempty and convex in R^{m+n}. They are called the *effective domains* of F and G. The restriction of F to $\operatorname{dom} F$ is a finite convex function in the ordinary sense. The effective domain of K is defined by

$$\operatorname{dom} K = \operatorname{dom}_1 K \times \operatorname{dom}_2 K$$

where

$$\operatorname{dom}_1 K = \{x \mid K(x, y) < +\infty \text{ for all } y\} ,$$
$$\operatorname{dom}_2 K = \{y \mid K(x, y) > -\infty \text{ for all } x\} .$$

The latter are nonempty convex sets. It is shown in [8] that, for each $y \in \operatorname{dom}_2 K$, the effective domain of the convex function $K(\cdot, y)$ lies between $\operatorname{dom}_1 K$ and its closure. Likewise for the concave functions $K(x, \cdot)$. The restriction of K to the relative interior of $\operatorname{dom} K$ is a so-called relatively open saddle-element (the *kernel* of K) which completely determines the minimax equivalence class containing K. The effective domain of J is similarly defined by

$$\operatorname{dom} J = \operatorname{dom}_1 J \times \operatorname{dom}_2 J$$

where

$$\operatorname{dom}_1 J = \{u \mid J(u, v) > -\infty \text{ for all } v\} ,$$
$$\operatorname{dom}_2 J = \{v \mid J(u, v) > +\infty \text{ for all } u\} .$$

Since these effective domains enter into the hypotheses of two of our later theorems, it is helpful to know the following result about their relationship.

LEMMA 3.

$$\operatorname{dom}_1 K = \{x \mid (x, v) \in \operatorname{dom} F \textit{ for some } v\} ,$$
$$\operatorname{dom}_2 J = \{v \mid (x, v) \in \operatorname{dom} F \textit{ for some } x\} ,$$
$$\operatorname{dom}_1 J = \{u \mid (u, y) \in \operatorname{dom} G \textit{ for some } y\} ,$$
$$\operatorname{dom}_2 K = \{y \mid (u, y) \in \operatorname{dom} G \textit{ for some } u\} .$$

Proof. Since $F(x, \cdot)$ is the skew-conjugate of the convex function $K(x, \cdot)$, it is identically $+\infty$ if and only if $K(x, y) = +\infty$ for some y. This fact is equivalent to the formula for $\operatorname{dom}_1 K$. The other formulas can be established in the same way.

There is also a striking relationship between the subdifferentials of F, G, K and J which will be at the heart of our programming theory. These subdifferentials are most easily defined by first considering the case of a convex function f on R^m. A vector u is called a *subgradient* of f at x if

$$f(z) \geqq f(x) + \langle z - x, u \rangle \qquad \text{for all } z\,. \tag{4.1}$$

The set of these is a closed convex subset of R^m denoted by $\partial f(x)$. The map $x \to \partial f(x)$ is the *subdifferential* of f. If $f(x)$ is finite, the directional derivative

$$f'(x; z) = \lim_{\lambda \downarrow 0} [f(x + \lambda z) - f(x)]/\lambda \tag{4.2}$$

exists for all $z \in R^m$ and is a convex function of z. One has $u \in \partial f(x)$ if and only if

$$f'(x; z) \geqq \langle u, z \rangle \qquad \text{for all } z\,, \tag{4.3}$$

In particular, if f is differentiable in the ordinary sense at x, then $\partial f(x)$ consists of just one vector, the ordinary gradient $\nabla f(x)$. The situation for concave functions is the same, except that the inequalities need to be reversed in (4.1) and (4.3). (For more information about subdifferentials, we refer the reader to the bibliography in [9].) The meaning of $\partial F(x, v)$ and $\partial G(u, y)$ is now clear, since F and G are convex and concave functions on R^{m+n}. The subdifferential of K is defined by

$$\partial K(x, y) = \partial_1 K(x, y) \times \partial_2 K(x, y)$$

where, for each $y \in R^n$, $\partial_1 K(x, y)$ consists of the subgradients in R^m of the convex function $K(\cdot, y)$ at x, and so forth. One defines $\partial J(u, v)$ in the same way.

LEMMA 4. *The following conditions on a set of four vectors x, y, u, v, are equivalent:*

(a) $(u, y) \in \partial F(x, v)$,
(b) $(-x, -v) \in \partial G(u, y)$,
(c) $(u, -v) \in \partial K(x, y)$,
(d) $(-x, y) \in \partial J(u, v)$.

These conditions imply that the four values $F(x, v)$, $G(u, y)$, $K(x, y)$, $J(u, v)$, are finite.

Proof. If f and g are convex and concave on R^m, skew-conjugate to each other, and not identically $+\infty$ or $-\infty$, then the three conditions

$$u \in \partial f(x),\ -x \in \partial g(u),\ f(x) - g(u) \leqq \langle x, u \rangle\,, \tag{4.4}$$

are known to be equivalent and to imply the finiteness of $f(x)$ and

$g(u)$. One can quickly verify this using the definition of the skew-conjugate correspondence. Hence (a) and (b) are equivalent and imply the finiteness of $F(x, v)$ and $G(u, y)$. We next use the same basic fact to show (c) implies (a) and (b). By definition, (c) means that $u \in \partial_1 K(x, y)$ and $-v \in \partial_2 K(x, y)$, which itself means

$$K(z, y) \geqq K(x, y) + \langle z - x, u\rangle \qquad \text{for all } z ,$$
$$K(x, w) \leqq K(x, y) + \langle w - y, -v\rangle \qquad \text{for all } w .$$

We can write this equivalently as

$$\begin{aligned} K(x, y) - \langle x, u\rangle &\leqq G(u, y) , \\ K(x, y) + \langle y, v\rangle &\geqq F(x, v) , \end{aligned} \tag{4.5}$$

because F and G, being the parents of K, are given by (2.6) and (2.7). Of course, then

$$F(x, v) - G(u, y) \leqq \langle x, u\rangle + \langle y, v\rangle \tag{4.6}$$

so that (a) and (b) hold by the equivalences in (4.4). The finiteness of $K(x, y)$ also follows from this and (4.5). Now suppose conversely that (a) and (b) hold. In particular, y is a subgradient of $F(x, \cdot)$ at v, and the closed convex function $F(x, \cdot)$ is not identically $+\infty$ or $-\infty$. The skew-conjugate of $F(x, \cdot)$ is $\mathrm{cl}_2\, K(x, \cdot)$, so

$$F(x, v) - \mathrm{cl}_2\, K(x, y) \leqq \langle y, v\rangle ,$$

again by the equivalence of the conditions in (4.4). Likewise

$$\mathrm{cl}_1\, K(x, y) - G(u, y) \leqq \langle x, u\rangle .$$

These two inequalities imply (4.5), which we have shown to be equivalent to (c), because

$$\mathrm{cl}_1\, K(x, y) \leqq K(x, y) \leqq \mathrm{cl}_2\, K(x, y) .$$

The equivalence of (d) with (a) and (b) can be proved in the same way.

5. The four programming problems.

The four problems introduced in § 1 can be expressed more specifically as follows.

(I) Determine $\bar{x} \geqq 0$ and $\bar{v} \geqq 0$ such that $F(\bar{x}, \bar{v})$ is finite and $F(x, v) \geqq F(\bar{x}, \bar{v})$ for all $x \geqq 0$ and $v \geqq 0$.

(II) Determine $\bar{u} \geqq 0$ and $\bar{y} \geqq 0$ such that $G(\bar{u}, \bar{y})$ is finite and $G(u, y) \leqq G(\bar{u}, \bar{y})$ for all $u \geqq 0$ and $y \geqq 0$.

(III) Determine $\bar{x} \geqq 0$ and $\bar{y} \geqq 0$ such that $K(\bar{x}, \bar{y})$ is finite and $K(x, \bar{y}) \geqq K(\bar{x}, \bar{y}) \geqq K(\bar{x}, y)$ for all $x \geqq 0$ and $y \geqq 0$.

(IV) Determine $\bar{u} \geqq 0$ and $\bar{v} \geqq 0$ such that $J(\bar{u}, \bar{v})$ is finite and $J(u, \bar{v}) \leqq J(\bar{u}, \bar{v}) \leqq J(\bar{u}, v)$ for all $u \geqq 0$ and $v \geqq 0$.

Naturally, a pair $(\bar{x}, \bar{v})$ of the type described in (I) is called a *solution* to (I). The real number $F(\bar{x}, \bar{v})$ is then denoted by "min (I)". Similarly for (II), (III) and (IV). Note that we do not speak of solutions unless the extrema are finite.

We shall be interested mainly in what we call "stable" solutions. These are defined as follows. First consider the notationally simpler case of a convex function f on R^m. Let $\bar{x} \geqq 0$ be a point where f is finite, and let

$$M = \{\lambda(x - \bar{x}) \mid \lambda \geqq 0, x \geqq 0\} . \tag{5.1}$$

The infimum of f subject to $x \geqq 0$ is achieved at $\bar{x}$ if and only if the directional derivative function satisfies

$$f'(\bar{x}; z) \geqq 0 \qquad \text{for all } z \in M . \tag{5.2}$$

Now it can happen in certain peculiar situations that (5.2) holds, and yet

$$\inf_{z \in M} f'(\bar{x}; w + z) = -\infty \qquad \text{for some } w \notin M . \tag{5.3}$$

Then we say that the infimum is achieved *unstably* at $\bar{x}$ (otherwise: *stably*). This terminology is suggested by the fact that, if (5.2) and (5.3) hold, the infimum of f subject to $x \geqq \varepsilon w$ is a function of ε whose righthand derivative at $\varepsilon = 0$ is $-\infty$. In other words, the infimum would drop off at an initially infinite rate if the nonnegativity constraint on x were relaxed slightly. We shall not elaborate this here; a similar stability notion has been developed in detail in [9].

We speak of a solution $(\bar{x}, \bar{v})$ to (I) as *stable*, if the infimum of $F(x, v)$ subject to $x \geqq 0, v \geqq 0$ is achieved stably at $(\bar{x}, \bar{v})$ in the sense just defined. Stable solutions to (II), where G is concave instead of convex, are defined in the obviously analogous manner. Next consider (III). A solution $(\bar{x}, \bar{y})$ involves $K(x, \bar{y})$ having a minimum at $\bar{x}$ subject to $x \geqq 0$ and $K(\bar{x}, y)$ having a maximum at $\bar{y}$ subject to $y \geqq 0$. We therefore say $(\bar{x}, \bar{y})$ is a stable solution to (III) if these two separate extrema are stably achieved. The definition for (IV) is practically the same.

The following theorem gives elementary criteria for stability of solutions. (The relative interior of a convex set C is, of course, the interior of C with respect to its affine hull, the intersection of all subspace translates containing C.)

THEOREM 2. *If the relative interior of* dom F *contains at least one* $(x, v) \geqq 0$, *then all solutions to* (I) (*if they exist*) *are stable. If the relative interior of* dom G *contains at least one* $(u, y) \geqq 0$, *then* *all solutions to* (II) *are stable. If the relative interior of* dom K *contains at least one* $(x, y) \geqq 0$, *then all solutions to* (III) *are stable. If the relative interior of* dom J *contains at least one* $(u, v) \geqq 0$, *then all solutions to* (IV) *are stable.*

Proof. We start out again with a convex function f on R^m and its effective domain $\operatorname{dom} f = \{x \mid f(x) < +\infty\}$. Suppose that $\operatorname{dom} f$ contains some $x \geqq 0$ in its relative interior. Suppose also that the infimum of f subject to $x \geqq 0$ is finite and achieved at $\bar{x}$. Then f is certainly not identically $+\infty$, nor can it have the value $-\infty$ anywhere, since it is known that a convex function which takes on $-\infty$ must have this value throughout the relative interior of its effective domain. Thus f is a proper convex function in the sense of [7], so that Theorem 2 of that paper can be applied. According to this theorem, there exists some $\bar{u} \geqq 0$ such that

$$\inf_{x \geqq 0} f(x) = g(\bar{u})$$

where g is the skew-conjugate of f. Since $\langle \bar{x}, \bar{u} \rangle \geqq 0$, then

$$(5.4) \qquad f(\bar{x}) = \inf_x \{f(x) - \langle x, \bar{u} \rangle\} \leqq \inf_x \{f(x) - \langle x - \bar{x}, \bar{u} \rangle\} .$$

This says that $\bar{u}$ is a subgradient of f at $\bar{x}$. Taking $x = \bar{x}$ in (5.4), we also see that $\langle \bar{x}, \bar{u} \rangle = 0$, since both $\bar{x}$ and $\bar{u}$ are nonnegative. Thus $\langle z, \bar{u} \rangle \geqq 0$ for all $z \in M$, where M is the convex cone defined in (5.1). We therefore have for any w

$$(5.5) \qquad \inf_{z \in M} f'(x; w + z) \geqq \inf_{z \in M} \langle w + z, \bar{u} \rangle = \langle w, \bar{u} \rangle > -\infty .$$

Thus the infimum is achieved stably at $\bar{x}$, as we wanted to prove The assertion of the theorem about (I) differs only in notation from the fact just proved. The assertion about (II) follows analogously. A double application is needed to take care of (III). The relative interior of dom K contains a nonnegative element if and only if some $x \geqq 0$ belongs to the relative interior of $\operatorname{dom}_1 K$ and some $y \geqq 0$ belongs to the relative interior of $\operatorname{dom}_2 K$. Suppose $(\bar{x}, \bar{y})$ is a solution to (III). The relative interior of the effective domain of the convex function $K(\cdot, \bar{y})$ is the same as the relative interior of $\operatorname{dom}_1 K$, as shown in [8]. Taking $f(x) = K(x, \bar{y})$, we may conclude therefore that the infimum of $K(x, \bar{y})$ subject to $x \geqq 0$ is stably achieved at $\bar{x}$. Likewise, the supremum of $K(\bar{x}, y)$ subject to $y \geqq 0$ is stably achieved at $\bar{y}$. But this means that $(\bar{x}, \bar{y})$ is a stable solution to (III). The argument for (IV) is virtually the same.

6. Programming theorems. We can now establish our main results, a double duality theorem, a characterization theorem, and an existence theorem. The first two will have joint proof.

THEOREM 3. *If any one of the problems* (I), (II), (III), (IV), *has a stable solution, then all four problems have stable solutions and*

$$\min(\mathrm{I}) = \max(\mathrm{II}) = \operatorname{minimax}(\mathrm{III}) = \operatorname{maximin}(\mathrm{IV}) . \tag{6.1}$$

THEOREM 4. *The following conditions on a set of four vectors* $\bar{x}, \bar{y}, \bar{u}, \bar{v}$, *are equivalent.*

(a) $(\bar{x}, \bar{v})$ *is a stable solution to* (I) *and* $(\bar{u}, \bar{y})$ *is a stable solution to* (II).

(b) $(\bar{x}, \bar{y})$ *is a stable solution to* (III) *and* $(\bar{u}, \bar{v})$ *is a stable solution to* (IV).

(c) $\bar{x}, \bar{y}, \bar{u}, \bar{v}$, *satisfy one of the equivalent subdifferential conditions in Lemma* 4, *as well as the complementary slackness conditions*:

$$\bar{x} \geqq 0, \bar{u} \geqq 0, \langle \bar{x}, \bar{u} \rangle = 0, \bar{y} \geqq 0, \bar{v} \geqq 0, \langle \bar{y}, \bar{v} \rangle = 0 . \tag{6.2}$$

Proof. Once more consider a convex function f on R^m having a finite infimum subject to $x \geqq 0$. We shall show that this is achieved stably at $\bar{x} \geqq 0$ if and only if there exists some $\bar{u} \geqq 0$ such that $\bar{u} \in \partial f(\bar{x})$ and $\langle \bar{x}, \bar{u} \rangle = 0$. To start with, let us suppose that $\bar{x}$ and $\bar{u}$ have the latter properties. Then $\langle z, \bar{u} \rangle \geqq 0$ for all $z \in M$(the set in (5.1)), and (5.5) holds for every w. Thus the minimum is stably achieved at x by the argument already used in the proof of Theorem 2. Now suppose conversely that the infimum is achieved stably at $\bar{x}$. The function h defined by

$$h(w) = \inf_{z \in M} f'(\bar{x}; w + z) \tag{6.3}$$

then nowhere has the value $-\infty$. Furthermore, h is a convex function on R^m. This follows from the fact that $f'(\bar{x}; \cdot)$ is a convex function and the set M is convex. Namely, given

$$h(w_1) \leqq \mu_1 \in R, h(w_2) \leqq \mu_2 \in R, 0 < \lambda < 1 ,$$

and any $\varepsilon > 0$ there exists $z_1 \in M$ and $z_2 \in M$ such that

$$\mu_1 + \varepsilon \geqq f'(\bar{x}; w_1 + z_1) \quad \text{and} \quad \mu_2 + \varepsilon \geqq f'(\bar{x}; w_2 + z_2) .$$

Then we have

$$\lambda\mu_1 + (1 - \lambda)\mu_2 + \varepsilon \geqq f'(\bar{x}; \lambda(w_1 + z_1) + (1 - \lambda)(w_2 + z_2)) .$$

Since $\lambda z_1 + (1 - \lambda)z_2 \in M$ and ε was arbitrary, we conclude

$$\lambda\mu_1 + (1 - \lambda)\mu_2 \geqq h(\lambda w_1 + (1 - \lambda)w_2) ,$$

thereby completing the proof that h is convex. Fenchel showed in [3] (in the other notation) that a convex function which nowhere has the value $-\infty$ majorizes at least one (finite) affine function. Applied to the case at hand, this means there exists some $\bar{u} \in R^m$ and $\alpha \in R$ such that

$$(6.4) \quad \alpha + \langle \bar{u}, w \rangle \leqq h(w) \leqq f'(\bar{x}; w + z) \quad \text{for every} \quad z \in M \text{ and } w .$$

Taking $w = -z$ in (6.4), we see that $\langle \bar{u}, z \rangle \geqq 0$ for every $z \in M$. Hence $\bar{u} \geqq 0$ and $\langle \bar{x}, \bar{u} \rangle = 0$. Taking $z = 0$ in (6.4), we may also deduce at once that $\bar{u} \in \partial f(\bar{x})$. This finishes the demonstration of the stability condition stated at the outset of the proof. In the context of (I), the condition says that $(\bar{x}, \bar{v})$ is a stable solution if and only if (6.2) is satisfied for some $(\bar{u}, \bar{y}) \in \partial F(\bar{x}, \bar{v})$. (Here we are also making use of the fact that, by Lemma 4, the existence of a subgradient $(\bar{u}, \bar{y})$ implies $F(\bar{x}, \bar{v})$ is finite.) When the stability condition is applied in like manner to the other three problems, we get the following characterizations. In (II), $(\bar{u}, \bar{y})$ is a stable solution if and only if (6.2) is satisfied by some $(-x, -v) \in \partial G(\bar{u}, \bar{y})$. In (III), $(\bar{x}, \bar{y})$ is a stable solution if and only if (6.2) is satisfied by some $(\bar{u}, -\bar{v}) \in \partial K(\bar{x}, \bar{y})$. And in (IV), $(\bar{u}, \bar{v})$ is a stable solution if and only if (6.2) is satisfied by some $(-\bar{x}, \bar{y}) \in \partial J(\bar{u}, \bar{v})$. By virtue of the equivalences in Lemma 4, these characterizations prove Theorem 4, and all of Theorem 3 except (6.1). In the proof of Lemma 4, however, it was shown that the four equivalent subdifferential conditions on $\bar{x}, \bar{y}, \bar{u}, \bar{v}$, imply

$$(6.5) \quad K(\bar{x}, \bar{y}) - \langle \bar{x}, \bar{u} \rangle \leqq G(\bar{u}, \bar{y}) \quad \text{and} \quad K(\bar{x}, \bar{y}) + \langle \bar{y}, \bar{v} \rangle \geqq F(\bar{x}, \bar{v}) .$$

In the same way they imply

$$(6.6) \quad J(\bar{u}, \bar{v}) + \langle \bar{u}, \bar{x} \rangle \geqq F(\bar{x}, \bar{v}) \quad \text{and} \quad J(\bar{u}, \bar{v}) - \langle \bar{y}, \bar{v} \rangle \leqq G(\bar{u}, \bar{y}) .$$

On the other hand, we have

$$(6.7) \qquad F(\bar{x}, \bar{v}) - G(\bar{u}, \bar{y}) \geqq \langle \bar{x}, \bar{u} \rangle + \langle \bar{v}, \bar{y} \rangle$$

by definition of the skew-conjugate correspondence. Now $\langle \bar{x}, \bar{u} \rangle = 0$ and $\langle \bar{y}, \bar{v} \rangle = 0$ for stable solutions according to (6.2), so when (6.5), (6.6) and (6.7) are put together they yield

$$F(\bar{x}, \bar{v}) = G(\bar{u}, \bar{y}) = K(\bar{x}, \bar{y}) = J(\bar{u}, \bar{v}) .$$

This is (6.1), the desired conclusion.

THEOREM 5. *If there exist vectors* $x \geqq 0, y \geqq 0, u \geqq 0, v \geqq 0$ *such that* (x, v) *belongs to the relative interior of* dom F *and* (u, y) *belongs to the relative interior of* dom G, *then stable solutions exist for* (I), (II), (III), (IV).

Proof. Since F and G are skew-conjugate to each other, the hypothesis implies by Theorem 2 of [7] that (I) and (II) have solutions. These solutions are stable by Theorem 2 of the present paper. Problems (III) and (IV) then also have stable solutions by Theorem 3.

REFERENCES

1. A. Brøndsted, *Conjugate convex functions in topological vector spaces*, Mat. Fys. Medd. Dan. Vid. Selsk. **34** (Copenhagen, 1964).
2. G. Dantzig, E. Eisenberg and R. W. Cottle, *Symmetric dual nonlinear programs*, Pacific J. Math. **15** (1965), 809-812.
3. W. Fenchel, *On conjugate convex functions*, Canad. J. Math. **1** (1949), 73-77.
4. ———, *Convex Cones, Sets and Functions*, mimeographed lecture notes, Princeton Univ., 1951.
5. O. L. Mangasarian and J. Ponstein, *Minimax and duality in nonlinear programming*, J. Math. Anal. **11** (1965), 504-518.
6. J. J. Moreau, *Théorèmes 'inf-sup'*, C. R. Acad. Sc. Paris **258** (March, 1964), 2720-2722.
7. R. T. Rockafellar, *Duality theorems for convex functions*, Bull. Amer. Math. Soc. **70** (1964), 189-192.
8. ———, *Minimax theorems and conjugate saddle-functions*, Math. Scand. **14** (1964), 151-173.
9. ———, *Duality and stability in extremum problems involving convex functions*, Pacific J. Math. **21** (1967), 167-187.
10. J. Stoer, *Über einen Dualitätssatz der nichtlinearen Programmierung*, Num. Math. **6** (1964), 55-58.
11. A. Whinston, *Some applications of the conjugate function theory to duality, in Nonlinear Programming*, J. Abadie, ed., North-Holland, Amsterdam, 1967.

Received January 3, 1966. This work was supported in part by the Air Force Office of Scientific Research under grant AF-AFOSR-1202-67.

UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON