STOCHASTIC CONVEX PROGRAMMING: SINGULAR MULTIPLIERS AND EXTENDED DUALITY SINGULAR MULTIPLIERS AND DUALITY

R. T. ROCKAFELLAR AND R. J.-B. WETS

A two-stage stochastic programming problem with recourse is studied here in terms of an extended Lagrangian function which allows certain multipliers to be elements of a dual space $(\mathscr{L}^\infty)^*$, rather than an $\mathscr{L}^1$ space. Such multipliers can be decomposed into an $\mathscr{L}^1$-component and a "singular" component. The generalization makes it possible to characterize solutions to the problem in terms of a saddle-point, if the problem is strictly feasible. The Kuhn-Tucker conditions for the basic duality framework are modified to admit singular multipliers. It is shown that the optimal multiplier vectors in the extended dual problem are, in at least one broad case, ideal limits of maximizing sequences of multiplier vectors in the basic dual problem.

1. Introduction. This paper is a sequel to [3], where the following two-stage stochastic programming problem was investigated: minimize

$$(1.1) \qquad f_{10}(x_1) + \int_S f_{20}(s, x_1, x_2(s))\sigma(ds)$$

over all $x_1 \in R^{n_1}$ and $x_2 \in \mathscr{L}^\infty_{n_2} = \mathscr{L}^\infty(S, \Sigma, \sigma; R^{n_2})$ satisfying

$$(1.2) \qquad x_1 \in C_1 \quad \text{and} \quad f_{1i}(x_1) \leqq 0 \quad \text{for} \quad i = 1, \cdots, m_1$$

and almost surely

$$(1.3) \quad x_2(s) \in C_2 \quad \text{and} \quad f_{2i}(s, x_1, x_2(s)) \leqq 0 \quad \text{for} \quad i = 1, \cdots, m_2.$$

Here (S, Σ, σ) denotes a probability space, and the sets $C_1 \subset R^{n_1}$ and $C_2 \subset R^{n_2}$ are closed, convex, and nonempty. The functions f_{1i} for $i = 0, 1, \cdots, m_1$ and $f_{2i}(s, \cdot, \cdot)$ for $i = 0, 1, \cdots, m_2$ are finite and convex on all of R^{n_1} and $R^{n_1} \times R^{n_2}$, respectively. It is assumed for each $(x_1, x_2) \in R^{n_1} \times R^{n_2}$ the function $f_{2i}(\cdot, x_1, x_2)$ is measurable on S, in fact summable if $i = 0$ and bounded if $i = 1, \cdots, m_2$. (As pointed out in [1], these assumptions imply that for every $x_1 \in R^{n_1}$ and $x_2 \in \mathscr{L}^\infty_{n_2}$, $f_{2i}(s, x_1, x_2(s))$ is measurable in s, summable if $i = 0$ and essentially bounded if $i = 1, \cdots, m_2$).

The results in [3] concern the justification of this model and its analysis in terms of the Lagrangian function:

$$(1.4) \qquad L(x, y) = L_1(x_1, y_1) + \int_S L_2(s, x_1, x_2(s), y_2(s))\sigma(ds) \text{ if } x \in X_0$$

$$\text{and } y \in Y_0,$$

$$= -\infty \text{ if } x \in X_0 \text{ and } y \notin Y_0,$$

$$= +\infty \text{ if } x \notin X_0,$$

where

$$(1.5) \quad x = (x_1, x_2) \in X = R^{n_1} \times \mathscr{L}^\infty_{n_2} \text{ and } y = (y_1, y_2) \in Y = R^{m_1} \times \mathscr{L}^1_{m_2},$$

$$(1.6) \qquad X_0 = \{x \in X \mid x_1 \in C_1 \text{ and almost surely } x_2(s) \in C_2\},$$

$$(1.7) \qquad Y_0 = \{y \in Y \mid y_1 \geqq 0 \text{ and almost surely } y_2(s) \geqq 0\},$$

$$(1.8) \qquad L_1(x_1, y_1) = f_{10}(x_1) + \sum_{i=1}^{m_1} y_{1i} f_{1i}(x_1),$$

$$(1.9) \qquad L_2(s, x_1, x_2, y_2) = f_{20}(s, x_1, x_2) + \sum_{i=1}^{m_2} y_{2i} f_{2i}(s, x_1, x_2).$$

Of course, $L(x, y)$ is convex in $x \in X$, concave in $y \in Y$, and finite on $X_0 \times Y_0$.

It is easily seen that the stochastic programming problem above can be expressed equivalently as

P $\qquad\qquad$ minimize $f(x)$ over all $x \in X$,

where

$$(1.10) \qquad\qquad\qquad f(x) = \sup_{y \in Y} L(x, y).$$

A corresponding dual problem is therefore:

D $\qquad\qquad$ maximize $g(y)$ over all $y \in Y$,

where

$$(1.11) \qquad\qquad\qquad g(y) = \inf_{x \in X} L(x, y).$$

The general duality theory of [4] was applied to these problems in [3], and it was shown in particular that min **P** = sup **D** if C_1 and C_2 are bounded. The condition that $(\bar{x}, \bar{y})$ be a saddle-point of L on $X \times Y$ (or

equivalently on $X_0 \times Y_0$), which is always sufficient for $\bar{x}$ to solve **P** and $\bar{y}$ to solve **D**, is therefore also necessary in the latter case.

A version of the Kuhn-Tucker theorem is thereby established: assuming that C_1 and C_2 are bounded, and *assuming that the supremum in* **D** *is attained*, an element $\bar{x} \in X$ solves **P** if and only if there exists $\bar{y} \in Y$ such that $(\bar{x}, \bar{y})$ is a saddle-point of L. We have shown in another paper [5] how this saddle-point property can be expressed by certain "point-wise" Kuhn-Tucker conditions involving a function $p \in \mathscr{L}^1_{n_1}$ which assigns prices for making alterations in the first-stage decision x_1 after the observation of s.

The central question in this is whether the supremum in **D** is indeed attained. One might hope this is always true if **P** is *strictly feasible*, in the sense that for some $x \in X$ and $\epsilon > 0$ the constraints (1.2) and (almost surely) (1.3) are satisfied with $-\epsilon$ in place of 0. There are simple counterexamples, however [3, §5]. We have at least shown qualitatively in [3] that the supremum in **D** is "typically" attained, if C_1 and C_2 are bounded. But the issue remains of whether a verifiable condition can be furnished, whereby attainment is assured for an individual case.

The aim of this paper is to provide an answer in somewhat modified terms. The Lagrangian function is extended so that the multiplier functions $y_2 \in \mathscr{L}^1_{m_2}$ are replaced by more general elements of the dual space $(\mathscr{L}^\infty_{m_2})^*$, each of which can be decomposed into an $\mathscr{L}^1_{m_2}$ component and a "singular" component. This leads to a new dual problem $\tilde{\mathbf{D}}$, a natural extension of **D**. Strict feasibility of **P** is shown to imply $\inf \mathbf{P} = \max \tilde{\mathbf{D}}$ and hence a Kuhn-Tucker theorem in terms of the extended Lagrangian. It is demonstrated that the optimal multipliers in this setting, solutions to the extended dual $\tilde{\mathbf{D}}$, can be characterized fairly generally as the limits in a certain sense of maximizing sequences for the basic dual **D**. Finally, the Kuhn-Tucker conditions in [5] are generalized to the extended Lagrangian.

We shall show elsewhere [6] that these results, together with an analysis of so-called induced constraints, yield a reasonable condition for the supremum in **D** itself to be attained.

2. Singular functionals and the extended Lagrangian.
The second-stage inequality constraints in **P** can be expressed by saying that the functions

$$(2.1) \qquad \hat{f}_{2i}(x) \colon s \to f_{2i}(s, x_1, x_2(s)), \qquad i = 1, \cdots, m_2$$

(each of which is measurable and essentially bounded for $x = (x_1, x_2) \in X = R^{n_1} \times \mathscr{L}^\infty_{n_2}$) are all to be nonpositive as elements of $\mathscr{L}^\infty_1$. Symbolically:

$$(2.2) \qquad \hat{f}_{2i}(x) \leqq 0 \quad \text{for} \quad i = 1, \cdots, m_2 \quad (\text{in } \mathscr{L}^\infty_1).$$

The Lagrangian L associates with each of these constraints, in effect, a multiplier $y_{2i} \geqq 0$ in $\mathscr{L}_1^1$ via the natural pairing

$$(2.3) \quad \langle a, b \rangle = \int_S a(s)b(s)\sigma(ds) \quad \text{for} \quad a \in \mathscr{L}_1^\infty \quad \text{and} \quad b \in \mathscr{L}_1^1.$$

These multipliers y_{2i} thus represent continuous linear functionals on $\mathscr{L}_1^\infty$.

But, unless the probability space (S, Σ, σ) consists of a finite number of atoms, there exist continuous linear functionals on $\mathscr{L}_1^\infty$ not representable solely by elements of $\mathscr{L}_1^1$. The antithesis, so to speak, of a functional arising from $\mathscr{L}_1^1$ is a *singular* functional on $\mathscr{L}_1^\infty$. By this, one means a continuous linear functional b° with the property that the underlying probability space can be expressed as the union of an increasing sequence of measurable sets S_k, such that for each k one has $b^\circ(a) = 0$ for all the functions $a \in \mathscr{L}_1^\infty$ vanishing almost everywhere outside of S_k. The set of all singular functionals forms a linear space $\mathscr{S}_1 = \mathscr{S}_1(S, \Sigma, \sigma)$.

It is an important fact that the dual space $(\mathscr{L}_1^\infty)^*$ can be identified with $\mathscr{L}_1^1 \times \mathscr{S}_1$ under the pairing

$$(2.4) \qquad \langle a, (b, b^\circ) \rangle = \int_S a(s)b(s)\sigma(ds) + b^\circ(a).$$

In other words, the continuous linear functionals on $\mathscr{L}_1^\infty$ are the ones which can be represented as (2.4) for some (unique) $b \in \mathscr{L}_1^1$ and $b^\circ \in \mathscr{S}_1$. This decomposition, first applied in the study of optimization problems by Dubovitski and Miliutin [1], has been exploited in several papers on integral functionals and optimal control by Rockafellar [7], [8], [9]. Further applications to convex analysis have been pursued by Ioffe and Levin [2], who furnish in their Appendix I a complete proof in the more general case of $\mathscr{L}^\infty$-functions with values in a separable Banach space, rather than just R. Remarkably, singular functionals have not previously been employed in the theory of stochastic programming although they are essential, as we shall see, in the derivation of the strong forms of the duality results.

Singular functionals on $\mathscr{L}_m^\infty$ are defined in the same way as on $\mathscr{L}_1^\infty$. They correspond obviously to elements of $\mathscr{S}_m = \mathscr{S}_1 \times \cdots \times \mathscr{S}_1$ (m times). Let

$$(2.5) \qquad Y^\circ = \mathscr{S}_{m_2} \quad \text{and} \quad Y_0^\circ = \{y^\circ \in Y^\circ | y^\circ \geqq 0\}$$

where the relation $y^\circ \geqq 0$ means that each component $y_i^\circ \in \mathscr{S}_1$ is nonnegative as a functional on $\mathscr{L}_1^\infty$, i.e. satisfies

(2.6) $\qquad y_i^\circ(a) \geqq 0$ for all nonnegative $a \in \mathcal{L}_1^\infty$.

The *extended Lagrangian* $\tilde{L}$ on $X \times (Y \times Y_0^0)$ is defined by

$$(2.7) \quad \tilde{L}(x, y, y^\circ) = L_1(x_1, y_1) + \int_S L_2(s, x_1, x_2(s), y_2(s)) + L_2^\circ(x, y^\circ)$$

$$\text{if} \quad x \in X_0 \quad \text{and} \quad (y, y^\circ) \in Y_0 \times Y_0^\circ,$$

$$= -\infty \quad \text{if} \quad x \in X_0 \quad \text{and} \quad (y, y^\circ) \notin Y_0 \times Y_0^\circ,$$

$$= +\infty \quad \text{if} \quad x \in X_0,$$

where

$$(2.8) \qquad L_2^\circ(x, y^\circ) = \sum_{i=1}^{m_2} y_i^\circ(\hat{f}_{2i}(x)) = \sum_{i=1}^{m_2} y_i^\circ(f_{2i}(\,\cdot\,, x_1, x_2(\,\cdot\,))).$$

Thus each of the constraints (2.2) is now associated with a multiplier pair

$$(y_{2i}, y_i^\circ) \in \mathcal{L}_1^1 \times \mathcal{S}_1 \cong (\mathcal{L}_1^\infty)^*$$

It is clear that $\tilde{L}(x, y, y^\circ)$ is convex in x, concave in (y, y°), and finite on $X_0 \times (Y_0 \times Y_0^\circ)$. Moreover

$$(2.9) \qquad \tilde{L}(x, y, 0) \equiv L(x; y).$$

If x satisfies the second stage constraints (2.2) and $y^\circ \in Y_0^\circ$, then trivially $\tilde{L}(x, y, y^\circ) \leqq L(x, y)$. Therefore, the essential objective function f in **P**, defined by (1.10), can also be expressed by

$$(2.10) \qquad f(x) = \sup_{(y, y^\circ) \in Y \times Y^\circ} \tilde{L}(x, y, y^\circ).$$

In other words, $\tilde{L}$ is truly another Lagrangian associated with **P**. The problem

$$\tilde{\mathbf{D}} \qquad \text{maximize} \quad \tilde{g}(y, y^\circ) \quad \text{over all} \quad (y, y^\circ) \in Y \times Y^\circ,$$

where

$$(2.11) \qquad \tilde{g}(y, y^\circ) = \inf_{x \in X} \tilde{L}(x, y, y^\circ),$$

will be called the *extended dual* of **P**. Note that

$$(2.12) \qquad \tilde{g}(y, 0) \equiv g(y).$$

It is always true, therefore, that

(2.13) $\inf \mathbf{P} \geqq \sup \tilde{\mathbf{D}} \geqq \sup \mathbf{D},$

and any solution $(\bar{y}, \bar{y}^\circ)$ to $\tilde{\mathbf{D}}$ having $\bar{y}^\circ = 0$ is also a solution to $\mathbf{D}$. If C_1 and C_2 are bounded, then $\min \mathbf{P} = \sup \mathbf{D}$ according to the duality theorem in our previous paper [3], and hence in fact

(2.14) $\min \mathbf{P} = \sup \tilde{\mathbf{D}} = \sup \mathbf{D}.$

Formulas (2.10 and 2.11) yield also the general fact that $(\bar{x}, \bar{y}, \bar{y}^\circ) \in X \times (Y \times Y^\circ)$ *is a saddle-point of* $\tilde{L}$ *if and only if* $\bar{x}$ *gives the minimum in* $\mathbf{P}$, $(\bar{y}, \bar{y}^\circ)$ *gives the maximum in* $\tilde{\mathbf{D}}$, *and* $\min \mathbf{P} = \max \tilde{\mathbf{D}}$ (*finite*).

3. Perturbations. For purposes of applying general duality theory such as in [4] to the study of the relationship between $\mathbf{P}$ and $\tilde{\mathbf{D}}$, it is essential to demonstrate that the extended Lagrangian $\tilde{L}$ arises from a specific choice of perturbations of $\mathbf{P}$ and their topology.

In the case of the basic Lagrangian L, the perturbation space is

(3.1) $U = R^{m_1} \times \mathscr{L}^\infty_{m_2},$

and the function F is defined as follows: $F(x, u)$ for $x = (x_1, x_2) \in X$ and $u = (u_1, u_2) \in U$ is the value (1.1) if the perturbed constraints

(3.2) $x_1 \in C_1$ and $f_{1i}(x_1) \leqq u_{1i}$ for $i = 1, \cdots, m_1,$

and almost surely

(3.3) $x_2(s) \in C_2$ and $f_{2i}(s, x_1, x_2(s)) \leqq u_{2i}(s)$ for
                                                                $i = 1, \cdots, m_2,$

are satisfied; otherwise $F(x, u) = +\infty$. Then

(3.4) $F(x, 0) \equiv f(x).$

Assigning to U the product of the topology of R^{m_1} and the weak topology induced on $\mathscr{L}^\infty_{m_2}$ by $\mathscr{L}^1_{m_2}$, we can identify the dual of U with Y under the pairing

(3.5) $\langle u, y \rangle = u_1 \cdot y_1 + \int_S u_2(s) \cdot y_2(s) \sigma(ds).$

Then, as observed in [3], the formula

$$(3.6) \qquad L(x, y) = \inf_{u \in U} \{F(x, u) + \langle u, y \rangle\}$$

holds. In other words, the function L we have defined in (1.4) is the Lagrangian for **P** associated by general theory with this perturbational framework.

The function F is convex, proper, and lower semicontinuous on $X \times U$ with respect to the product of the above "weak* topology" on U and the similarly defined weak* topology on X [**3**, Proposition 3].

The framework changes if we assign to U the *strong topology*, i.e. the product of the topology of R^{m_1} and the norm topology of $\mathcal{L}^\infty_{m_2}$. Then the dual of U is identified instead with the space $Y \times Y^\circ$ under the pairing

$$(3.7) \qquad \langle u, (y, y^\circ) \rangle = \langle u, y \rangle + y^\circ(u).$$

PROPOSITION 1. *The Lagrangian for P associated with F and the latter pairing is* $\tilde{L}$:

$$(3.8) \qquad \tilde{L}(x, y, y^\circ) = \inf_{u \in U} \{F(x, u) + \langle u, (y, y^\circ) \rangle\}$$

Proof. Denote the right side of (3.8) by $M(x, y, y^\circ)$. Trivially $M(x, y, y^\circ) = +\infty$ if $x \notin X_0$, since then $F(x, u) = +\infty$ for all $u \in U$. Thus $\tilde{L}(x, y, y^\circ) = M(x, y, y^\circ)$ if $x \notin X_0$. Assume that $x \in X_0$. Then $M(x, y, y^\circ)$ is the infimum of the expression

$$(3.9) \quad f_{10}(x_1) + u_1 \cdot y_1 + \int_S [f_{20}(s, x_1, x_2(s)) + u_2(s) \cdot y_2(s)] \sigma(ds) + y^\circ(u_2)$$

over all $u_1 \in R^{m_1}$ and $u_2 \in \mathcal{L}^\infty_{m_2}$ satisfying (3.2) and (almost surely) (3.3). Setting

$$u_{1i} = w_{1i} + f_{1i}(x_1) \quad \text{for} \qquad\qquad i = 1, \cdots, m_1,$$
$$u_{2i}(s) = w_{2i}(s) + f_{2i}(s, x_1, x_2(s)) \quad \text{for} \qquad i = 1, \cdots, m_2,$$

we see that

$$(3.10) \quad M(x, y, y^\circ) = L_1(x_1, y_1) + \int_S L_2(s, x_1, x_2(s), y_2(s)) \sigma(ds) + N(y, y^\circ),$$

where L_1 and L_2 are given as before by (1.8) and (1.9), and

$$(3.11) \quad N(y, y^\circ) = \inf_{0 \leq w \in U} \left\{ w_1 \cdot y_1 + \int_S w_2(s) \cdot y_2(s) \sigma(ds) + y^\circ(w_2) \right\}.$$

To show that $M(x, y, y°) = \tilde{L}(x, y, y°)$, it suffices to show therefore that

$$N(y, y°) = 0 \quad \text{if} \quad (y, y°) \in Y_0 \times Y_0°,$$
$$= +\infty \quad \text{if} \quad (y, y°) \notin Y_0 \times Y_0°.$$

This is obvious from the definitions (1.7) and (2.5) of Y_0 and $Y_0°$, except for the following. Suppose that $b \in \mathcal{L}_1^1$ and $b° \in \mathcal{S}_1$ have the property that

$$(3.12) \quad \int_S a(s)b(s)\sigma(ds) + b°(a) \geqq 0 \text{ for all nonnegative } a \in \mathcal{L}_1^\infty.$$

We must show that then $b \geqq 0$ and $b° \geqq 0$. The argument is by contradiction. If it is not true that both $b \geqq 0$ in $\mathcal{L}_1^1$ and $b° \geqq 0$ in $\mathcal{S}_1$, then there exist nonnegative functions $\bar{a}$ and $\bar{\bar{a}}$ in $\mathcal{L}_1^1$ such that

$$(3.13) \qquad \int_S \bar{a}(s)b(s)\sigma(ds) + b°(\bar{\bar{a}}) < 0.$$

Let S be expressed as the union of an increasing family of measurable sets S_k associated with $b°$ as in the definition of singular functional. Define

$$a^k(s) = \bar{a}(s) \quad \text{if} \quad s \in S_k,$$
$$= \bar{\bar{a}}(s) \quad \text{if} \quad s \in S \backslash S_k.$$

Then $b°(a^k) = b°(\bar{\bar{a}})$ and

$$\lim_{k \to \infty} \int_S a^k(s)b(s)\sigma(ds) = \int_S \bar{a}(s)b(s)\sigma(ds),$$

so that by (3.13)

$$(3.14) \qquad \lim_{k \to \infty} \int_S a^k(s)b(s)\sigma(ds) + b°(a^k) < 0.$$

But also $a^k \geqq 0$, since $\bar{a} \geqq 0$ and $\bar{\bar{a}} \geqq 0$, so (3.14) is in conflict with (3.12). The proof is now complete.

4. Main results on duality. Proposition 1 enables us to analyse the relationship $\mathbf{P}$, $\hat{\mathbf{D}}$ and $\mathbf{D}$ in terms of the optimal value function

$$(4.1) \qquad\qquad \phi(u) = \inf_{x \in X} F(x, u),$$

giving for each $u \in U$ the infimum in the perturbed problem:

$$\mathbf{P}(u) \qquad\qquad \text{minimize} \quad F(x, u) \quad \text{over all} \quad x \in X.$$

Among the facts we obtain from general theory [4] (mostly right from the formulas above) are the following.

The function ϕ is convex on U and,

$$(4.2) \qquad \tilde{g}(y, y^\circ) = \inf_{u \in U} \{\phi(u) + \langle u, (y, y^\circ)\rangle\} = -\phi^*(-y, -y^\circ),$$

where ϕ^* denotes the conjugate of ϕ on $Y \times Y^\circ$ with respect to the pairing (3.7). Hence also

$$(4.3) \qquad \text{strong-cl } \phi(u) = \sup_{(y, y^\circ) \in Y \times Y^\circ} \{\tilde{g}(y, y^\circ) - \langle u, (y, y^\circ)\rangle\},$$

where strong-cl ϕ is the closure of ϕ with respect to the strong topology on U (coinciding with the conjugate of ϕ^* for the given pairing). Thus, whereas

$$(4.4) \qquad\qquad \inf \mathbf{P} = \phi(0)$$

by definition, we have

$$(4.5) \qquad\qquad \sup \tilde{\mathbf{D}} = \text{strong-cl } \phi(0).$$

This may be compared with the relation

$$(4.6) \qquad\qquad \sup \mathbf{D} = \text{weak}^*\text{-cl } \phi(0)$$

obtained in our previous paper [3], where weak*-cl ϕ is the closure of ϕ with respect to the weak topology on U induced by the pairing (3.5) with Y alone. In particular, then, if either $\inf \mathbf{P} < +\infty$ or $\sup \mathbf{D} > -\infty$, we have

$$\sup \tilde{\mathbf{D}} = \liminf_{u \to 0} \phi(u) \qquad \text{(strong topology)},$$

$$\sup \mathbf{D} = \liminf_{u \to 0} \phi(u) \qquad \text{(weak}^* \text{ topology)}.$$

As for solutions to $\tilde{\mathbf{D}}$, we have the fact that the equation

(4.7) $$\tilde{g}(\bar{y}, \bar{y}^\circ) = \max \tilde{\mathbf{D}} = \inf \mathbf{P}$$

holds if and only if

$$-(\bar{y}, \bar{y}^\circ) \in \partial \phi(0),$$

or in other words

(4.8) $$\phi(u) + \langle u, (\bar{y}, \bar{y}^\circ) \rangle \geq \phi(0) \quad \text{for all} \quad u \in U.$$

This means that $(\bar{y}, \bar{y}^\circ)$ represents a "vector of equilibrium prices" for the constraints in P with respect to the given system of perturbations: the expression

(4.9) $$F(x, u) + \langle u, \bar{y} \rangle + \bar{y}^\circ(u)$$

attains its minimum in (x, u) at $(\bar{x}, 0)$ if $\bar{x}$ is any solution to $\mathbf{P}$.

We can regard Y as a Banach space whose dual is U and whose bidual is $Y \times Y^\circ$. Thus Y is identified with a subspace of $Y \times Y^\circ$ (namely $Y \times \{0\}$) which is dense in the weak** topology, i.e the weak topology induced on $Y \times Y^\circ$ by the pairing with U. At the same time, the essential objective function g in D can be regarded as given on this dense subspace, and the essential objective function $\tilde{g}$ in $\tilde{\mathbf{D}}$ is then, according to (2.12), a certain extension of g to the whole space. The topological nature of this extension is clarified in an important case by the next result.

THEOREM 1. *Suppose the sets C_1 and C_2 are bounded. Then*

(4.10) $\tilde{g}(y, y^\circ) = \text{weak}^{**} - \limsup_{(z, 0) \to (y, y^\circ)} g(z)$ *for all* $(y, y^\circ) \in Y \times Y^\circ$.

Moreover, the following conditions on an element $(\bar{y}, \bar{y}^\circ) \in Y \times Y^\circ$ are equivalent:
   (a)   *$(\bar{y}, \bar{y}^\circ)$ solves $\tilde{\mathbf{D}}$;*
   (b)   *$(\bar{y}, \bar{y}^\circ)$ is the weak** limit of a maximizing (generalized) sequence for $\mathbf{D}$;*
   (c)   *$-(\bar{y}, \bar{y}^\circ) \in \partial \phi(0)$;*
   (d)   *For some (equivalently: for every) $\bar{x}$ solving $\mathbf{P}$, $(\bar{x}, \bar{y}, \bar{y}^\circ)$ is a saddle-point of the extended Lagrangian $\tilde{L}$.*

Proof. We rely on Theorem 3 of [3], from which we have not only

(2.14) but the properness and weak lower semicontinuity of the function ϕ in formula (4.2). The latter ensures, according to the basic theory of conjugate convex functions, that ϕ coincides with its biconjugate under the pairing between U and Y. Since

$$(4.11) \qquad \sup_{u \in U} \{\phi(u) + \langle u, y \rangle\} = \tilde{g}(y, 0) = g(y),$$

i.e. g is the concave conjugate of $-\phi$, this means that $-\phi$ is in turn the concave conjugate of g:

$$(4.12) \qquad -\phi(u) = \inf_{y \in Y} \{\langle u, y \rangle - g(y)\}.$$

Define $\bar{g}$ on $Y \times Y^\circ$ by

$$\bar{g}(y, y^\circ) = g(y) \quad \text{if} \quad y^\circ = 0,$$
$$= -\infty \quad \text{if} \quad y^\circ \neq 0.$$

Then $\bar{g}$ is concave, and we can re-express (4.12) as

$$(4.13) \qquad -\phi(u) = \inf_{(y, y^\circ) \in Y \times Y^\circ} \{\langle u, (y, y^\circ) \rangle - \bar{g}(y, y^\circ)\}.$$

But $\tilde{g}$ is the concave conjugate of $-\phi$ by (4.2). Therefore $\tilde{g} = \bar{g}^{**}$. Since $\tilde{g}$ is not identically $+\infty$ (obvious from (2.11), since $L(x, y, y^\circ) < +\infty$ if $x \in X_0$), it follows $\tilde{g}$ is the upper semicontinuous hull of $\bar{g}$ in any topology on $Y \times Y^\circ$ compatible with the pairing with U, and hence in particular in the weak** topology. This is exactly the meaning of (4.10). The equivalence of (a) and (b) is an immediate consequence of (4.10). The equivalence of (a), (c) and (d) follows from the equality (2.14) and the remarks above.

THEOREM 2. *If* **P** *is strictly feasible, then*

$$(4.14) \qquad \inf \mathbf{P} = \max \bar{\mathbf{D}} < +\infty,$$

and (assuming $\inf \mathbf{P} > -\infty$*) in terms of one-sided directional derivatives:*

$$(4.15) \quad \phi'(0; u) = \max\{\langle u, (\bar{y}, \bar{y}^\circ) \rangle \mid (\bar{y}, \bar{y}^\circ) \text{ solves } \bar{\mathbf{D}}\} \text{ for all } u \in U.$$

Furthermore, the following conditions are equivalent:
   (a) **P** *is strictly feasible;*

    (b) $\phi(0) < +\infty$, *and ϕ is strongly continuous on a neighborhood of* 0;

    (c)   *the level sets* $\{(y, y°) \mid \bar{g}(y, y°) \geqq \beta\}$, $\beta \in R$, *are all bounded in* $Y \times Y°$, *hence weak**-compact.*

Proof. Strict feasibility means by definition that $\phi(-\epsilon\mathbf{1}) < +\infty$ for some $\epsilon > 0$, where $\mathbf{1}$ is the element of $R^{m_1} \times \mathcal{L}^\infty_{m_2} = U$ whose components are all 1's. This is clearly equivalent to the seemingly weaker property that $\phi(u) < +\infty$ for some u in the interior of the nonpositive orthant of U. Since $\phi(u) \geqq \phi(u')$ if $u \leqq u'$, it follows that **P** is strictly feasible if and only if ϕ is bounded above on a strong neighborhood of 0. In this case, as observed in [**4**, Theorem 17], one has (4.14), (4.15) and (c). On the other hand, if (c) holds then from the conjugacy relation (4.2) and [**4**, Theorem 10] we know that the biconjugate of ϕ with respect to the pairing between U and $Y \times Y°$, namely strong-cl ϕ, is bounded above on a neighborhood of 0. But this implies $\phi(u) < +\infty$ for certain elements u of the interior of the nonpositive orthant of U, and hence, as we have seen, the strict feasibility of **P**.

    COROLLARY. *Suppose **P** is strictly feasible. Then, in order that $\bar{x}$ be a solution to **P**, it is necessary and sufficient that there exist $(\bar{y}, \bar{y}°) \in Y \times Y°$ such that $(\bar{x}, \bar{y}, \bar{y}°)$ is a saddle-point of the extended Lagrangian $\tilde{L}$. The pairs $(\bar{y}, \bar{y}°)$ are solutions to the extended dual $\tilde{D}$, and if C_1 and C_2 are bounded, they can also be regarded as idealized solutions to the basic dual **D**.*

    5.   Extended Kuhn-Tucker conditions. We have established in [**5**] that $(\bar{x}, \bar{y})$ is a saddle-point of L on $X \times Y$ if and only if the following properties, called the *basic Kuhn-Tucker conditions*, hold for some $p \in \mathcal{L}^1_{n_1}$:

    (a) $\bar{x}_1 \in C_1$ *and for $i = 1, \cdots, m_1$ one has* $\bar{y}_{1i} \geqq 0$, $f_{1i}(\bar{x}_1) \leqq 0$ *and* $\bar{y}_{1i} f_{1i}(\bar{x}_1) = 0$;

    (b)   *for almost all s, $\bar{x}_2(s) \in C_2$ and for $i = 1, \cdots, m_2$ one has* $\bar{y}_{2i}(s) \geqq 0$, $f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \leqq 0$ *and* $\bar{y}_{2i}(s) f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) = 0$;

    (c)   *the expression* $f_{10}(x_1) + \sum_{i=1}^{m_1} \bar{y}_{1i}(x_1) + x_1 \cdot \int_S p(s)\sigma(ds)$ *attains its minimum over the set C_1 at $\bar{x}_1$;*

    (d)   *for almost all s, the expression* $f_{20}(s, x_1, x_2) + \sum_{i=1}^{m_2} \bar{y}_{2i}(s) f_{2i}(s, x_1, x_2) - x_1 \cdot p(s)$ *attains its minimum in (x_1, x_2) over the set $R^{n_1} \times C_2$ at $(\bar{x}_1, \bar{x}_2(s))$.*

Supplementary conditions for a saddle point of the extended Lagrangian $\tilde{L}$ can be stated in terms of the function l on $R^{n_1} \times Y_0°$ defined by

$$(5.1) \quad l(x_1, y^\circ) = \inf\left\{\sum_{i=1}^{m_2} y_i^\circ(f_{2i}(\,\cdot\,, x_1, x_2(\,\cdot\,))) \;\Big|\; x_2 \in \mathcal{L}_{n_2}^\infty, x_2(s) \in C_2 \text{ a.s.}\right\}$$

(Recall that $Y_0^\circ = \{y^\circ \in \mathcal{S}_{m_2} \,|\, y_i^\circ \geq 0 \text{ for } i = 1, \cdots, m_2\}$.) It is easily seen that $l(x_1, y^\circ)$ is convex in $x_1 \in R^{n_1}$, concave in $y^\circ \in Y_0^\circ$, and nowhere $+\infty$.

We shall say that $(\bar{x}, \bar{y}, \bar{y}^\circ)$ satisfies the *extended Kuhn-Tucker conditions* if it satisfies the basic Kuhn-Tucker conditions with the term $l(x_1, \bar{y}^\circ)$ added to the expression in (c), i.e.:

(c′) *the expression* $f_{10}(x_1) + \sum_{i=1}^{m_1} \bar{y}_{1i} f_{1i}(x_1) + x_1 \displaystyle\int_S p(s)\sigma(ds) + $

$l(x_1, y^\circ)$ *attains its minimum over the set* C_1 *at* $\bar{x}_1$;

and if in addition the following conditions involving the singular parts of the multipliers hold:

$$(5.2) \quad \bar{y}_i^\circ \geq 0 \text{ and } \bar{y}_i^\circ(f_{2i}(\,\cdot\,, \bar{x}_1, \bar{x}_2(\,\cdot\,))) = 0 \text{ for } i = 1, \cdots, m_2,$$

$$(5.3) \quad l(\bar{x}_1, \bar{y}^\circ) = 0.$$

Conditions (a), (b), (c′) and (d) mean, of course, in view of the result cited above, that $(\bar{x}, \bar{y})$ is a saddle-point of the function $L + l(\,\cdot\,, \bar{y}^\circ)$, which is the basic Lagrangian for the modification of **P** in which f_{10} is replaced by $f_{10} + l(\,\cdot\,, \bar{y}^\circ)$.

THEOREM 3. *In order that $(\bar{x}, \bar{y}, \bar{y}^\circ)$ be a saddle-point of the extended Lagrangian $\tilde{L}$ on $X \times (Y \times Y^\circ)$, it is necessary and sufficient that the extended Kuhn-Tucker conditions hold for some $p \in \mathcal{L}_{n_1}^1$.*

Proof. In view of the remark prior to the theorem, it suffices to show that $(\bar{x}, \bar{y}, \bar{y}^\circ)$ is a saddle-point of $\tilde{L}$ if and only if $(\bar{x}, \bar{y})$ is a saddle-point of $L + l(\,\cdot\,, \bar{y}^\circ)$ and $(\bar{x}, \bar{y}^\circ)$ satisfies (5.2) and (5.3). By definition, $(\bar{x}, \bar{y}, \bar{y}^\circ)$ is a saddle-point of $\tilde{L}$ if and only if

$$(5.4) \quad \tilde{L}(\bar{x}, \bar{y}, \bar{y}^\circ) = \sup_{(y, y^\circ) \in Y \times Y^\circ} \tilde{L}(\bar{x}, y, y^\circ),$$

$$(5.5) \quad \tilde{L}(\bar{x}, \bar{y}, \bar{y}^\circ) = \inf_{x \in X} \tilde{L}(x, \bar{y}, \bar{y}^\circ).$$

These conditions imply in particular that $\bar{x} \in X_0$ and $(\bar{y}, \bar{y}^\circ) \in Y_0 \times Y_0^\circ$ (consider $(y, y^\circ) \in Y_0 \times Y_0^\circ$ in (5.4) and $x \in X_0$ in (5.5)). If $\bar{x} \in X_0$, the right side of (5.4) can be expressed as

$$\sup_{y \in Y} L(\bar{x}, y) + \sum_{i=1}^{m_2} \sup_{y_i^\circ \geq 0} y_i^\circ(f_{2i}(\,\cdot\,, \bar{x}_1, \bar{x}_2(\,\cdot\,))).$$

Assuming $\bar{x} \in X_0$, it is obvious from this that (5.4) is equivalent to (5.2) and the relation

$$(5.6) \qquad\qquad L(\bar{x}, \bar{y}) = \sup_{y \in Y} L(\bar{x}, y).$$

Trivially, (5.6) can also be written as

$$(5.7) \qquad\qquad [L(\bar{x}, \bar{y}) + l(\bar{x}_1, \bar{y}^\circ)] = \sup_{y \in Y} [L(\bar{x}_1, \bar{y}^\circ)],$$

if $l(\bar{x}_1, \bar{y}^\circ)$ is finite. To complete the proof of the theorem it will suffice therefore to show that (5.5) is equivalent to (5.3) and

$$(5.8) \qquad [L(\bar{x}, \bar{y}^\circ) + l(\bar{x}_1, \bar{y}^\circ)] = \inf_{x \in X} [L(x, \bar{y}) + l(x_1, \bar{y}^\circ)],$$

under the assumption that (5.2) holds and $\bar{y} \in Y_0$.

From (5.2) we have of course

$$(5.9) \qquad\qquad l(\bar{x}_1, \bar{y}^\circ) \leqq \sum_{i=1}^{m_2} y_i^\circ (f_{2i}(\,\cdot\,, \bar{x}_1, \bar{x}_2(\,\cdot\,))) = 0$$

and hence (5.3) holds if and only if the minimum in the definition (5.1) of $(\bar{x}_1, \bar{y}^\circ)$ is attained by $\bar{x}_2$. Thus (5.3) and (5.8) together amount to the assertion that the infimum of

$$(5.10)\, L_1(x_1, \bar{y}_1) + \int_S L_2(s, x_1, x_2'(s), \bar{y}_2(s))\sigma(ds) + \sum_{i=1}^{m_2} \bar{y}_i^\circ (f_{2i}(\,\cdot\,, x_1, x_2(\,\cdot\,)))$$

subject to

$$(5.11) \quad x_1 \in C_1, \quad x_2 \in \mathscr{L}_{n_2}^\infty, \quad x_2' \in \mathscr{L}_{n_2}^\infty, \quad x_2(s) \in C_2 \quad \text{and} \quad x_2'(s) \in C_2$$

almost surely is attained for $x_1 = \bar{x}_1$ and $x_2 = x_2' = \bar{x}_2$. On the other hand, (5.5) asserts that we have attainment of this infimum for $x_1 = \bar{x}_1$ and $x_2 = x_2' = \bar{x}_2$ if we make the *a priori* restriction that $x_2 = x_2'$. Thus the argument is reduced to showing for any $x_1 \in C_1$ that, with the notation

$$(5.12) \qquad A = \{x_2 \in \mathscr{L}_{n_2}^\infty | x_2(s) \in C_2 \text{ almost surely}\},$$

one has

$$(5.13) \quad \inf_{x_2 \in A} \left\{ \int_S L_2(s, x_1, x_2(s), \bar{y}_2(s)) \sigma(ds) + \sum_{i=1}^{m_2} \bar{y}_i^{\circ}(f_{2i}(\cdot, x_1, x_2(\cdot))) \right\}$$

$$= \inf_{x_2 \in A} \int_S L_2(s, x_1, x_2(s), \bar{y}_2(s)) \sigma(ds) + \inf_{x_2 \in A} \sum_{i=1}^{m_2} \bar{y}_i^{\circ}(f_{2i}(\cdot, x_1, x_2(\cdot))).$$

Since the inequality $\geqq$ certainly holds in (5.13), equality will follow if we demonstrate that for arbitrary $x_2' \in A$, $x_2'' \in A$ and $\epsilon > 0$, there exists $x_2 \in A$ such that

$$(5.14) \quad \int_S L_2(s, x_1, x_2(s), \bar{y}_2(s) \sigma(ds) + \sum_{i=1}^{m_2} \bar{y}_i^{\circ}(f_{2i}(\cdot, x_1, x_2(\cdot)))$$

$$\leqq \int_S L_2(s, x_1, x_2''(s), \bar{y}_2(s)) \sigma(ds) + \sum_{i=1}^{m_2} \bar{y}_i^{\circ}(f_{2i}(\cdot, x_1, x_2'(\cdot))) + \epsilon.$$

For each of the singular functionals $\bar{y}_i^{\circ}$, there exists by definition an increasing sequence of measurable sets S_{ik} with $\bigcup_{k=1}^{\infty} S_{ik} = S$, such that $\bar{y}_i^{\circ}(a) = 0$ if for some k the function $a \in \mathcal{L}_1^{\infty}$ vanishes almost everywhere outside of S_{ik}. The latter property implies that $\bar{y}_i^{\circ}(c) = \bar{y}_i^{\circ}(c')$ if c and c' agree almost everywhere outside of S_{ik}. For each index k, define

$$(5.15) \quad x_2^k(s) = x_2''(s) \quad \text{if} \quad s \in S_{ik} \quad \text{for} \quad i = 1, \cdots, m_2,$$

$$= x_2'(s) \quad \text{for all other } s.$$

Since $x_2' \in A$ and $x_2'' \in A$, we have $x_2^k \in A$. Moreover

$$f_{2i}(s, x_1, x_2^k(s)) = f_{2i}(s, x_1, x_2'(s)) \quad \text{if} \quad s \notin S_{ik},$$

so that

$$(5.16) \quad \sum_{i=1}^{m_2} \bar{y}_i^{\circ}(f_{2i}(\cdot, x_1, x_2^k(\cdot))) = \sum_{i=1}^{m_2} \bar{y}_i^{\circ}(f_{2i}(\cdot, x_1, x_2'(\cdot))).$$

At the same time, since

$$\lim_{k \to \infty} \sigma(S \setminus S_{ik}) = 0,$$

we have from (5.15) that

$$(5.17) \quad \lim_{k \to \infty} \int_S L_2(s, x_1, x_2^k(s), \bar{y}_2(s)) \sigma(ds) = \int_S L_2(s, x_1, x_2''(s), \bar{y}_2(s)) \sigma(ds).$$

From (5.16) and (5.17), we conclude that (5.14) holds for $x_2 = x_2^k$, if k is sufficiently large. This finishes the proof of Theorem 3.

REMARK. The last part of the proof of Theorem 3 has shown more generally that

$$(5.18)\quad \bar{g}(y, y^\circ) = \inf_{x \in X} \{L(x, y) + l(x_1, y^\circ)\} \quad \text{for} \quad (y, y^\circ) \in Y_0 \times Y_0^\circ.$$

Of course, $\bar{g}(y, y^\circ) = -\infty$ for $(y, y^\circ) \notin Y_0 \times Y_0^\circ$ by definition.

REFERENCES

1. A. Ya. Dubovitskii and A. A. Miliutin, *Necessary conditions for a weak extremum in problems of optimal control with mixed constraints of inequality type*, Zh. Vychisl. Mat. i Mat. Fys., **8** (1968), 725–779. (See also USSR Comp. Math. and Math. Phys., **8** (1968).)

2. A. D. Ioffe and V. I. Levin, *Subdifferentials of convex functions*, Trudy Mosk. Mat. Ob., (1972). (See also Transactions Moscow Math. Soc. (1972).)

3. R. T. Rockafellar and R. J.-B. Wets, *Stochastic convex programming: basic duality*, Pacific J. Math., **62** (1976).

4. R. T. Rockafellar, *Conjugate Duality and Optimization*, SIAM/CBMS lecture note series no. 16, 1974.

5. R. T. Rockafellar and R. J.-B. Wets, *Stochastic convex programming: Kuhn-Tucker conditions*, J. Math. Econ., **2** (1975), 349–370.

6. ———, *Stochastic convex programming: induced constraints*, SIAM J. Control Opt., **14** (1976).

7. R. T. Rockafellar, *Integrals which are convex functionals II*, Pacific J. Math., **39** (1971), 143–148.

8. ———, *Weak compactness of level sets of integral functionals*, Troisième Colloque d'Analyse Fonctionelle, H. G. Garnir (editor), Liège, 1971, 85–98.

9. ———, *State constraints in convex problems of Bolza*, SIAM J. Control, **10** (1972), 691–715.

UNIVERSITY OF WASHINGTON
AND
UNIVERSITY OF KENTUCKY