10

Reprinted from JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS Vol. 60, No. 2, September 1977

Measures as Lagrange Multipliers in Multistage Stochastic Programming

R. T. ROCKAFELLAR* AND R. J.-B. WETS†

Department of Mathematics, University of Washington, Seattle, Washington 98195

Submitted by H. J. Kushner

A duality theory is developed for multistage convex stochastic programming problems whose decision (or recourse) functions can be approximated by continuous functions satisfying the same constraints. Necessary and sufficient conditions for optimality are obtained in terms of the existence of multipliers in the class of regular Borel measures on the underlying probability space, these being decomposable, of course, into absolutely continuous and singular components with respect to the given probability measure. This provides an alternative to the approach where the multipliers are elements of the dual of $\mathscr{L}^\infty$ with an analogous decomposition. However, besides the existence of strictly feasible solutions, special regularity conditions are required, such as the "laminarity" of the probability measure, a property introduced in an earlier paper. These are crucial in ensuring that the minimum in the optimization problem can indeed be approached by continuous functions.

1. INTRODUCTION

In a stochastic programming problem, decisions must be taken in discrete time in response to the progressive observations of certain random variables, and in such a way as to minimize an overall expected cost subject to various constraints. In stage k, where $k = 1,..., N$, there is an R^{ν_k}-valued random variable ξ_k to be observed and a decision vector x_k in R^{n_k} to be determined. Let

$$\xi = (\xi_1, ..., \xi_N) \in R^{\nu_1} \times \cdots \times R^{\nu_N} = R^\nu,$$
$$x = (x_1, ..., x_N) \in R^{n_1} \times \cdots \times R^{n_N} = R^n.$$

The distribution of ξ is assumed to be given by a known (regular Borel) probability measure σ on a Borel subset Ξ of R^ν. The type of decision structure that is of interest is represented by a function $x: \Xi \to R^n$, called a *recourse*

* Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under AFOSR Grant No. 72-2269.

† Supported in part by the National Science Foundation under Grant MPS 75-07028.

ISSN 0022-247X

function (or in other contexts a decision rule, policy, or program), which is *nonanticipative* in the sense that $x_k(\xi)$ depends only on $\xi_1, ..., \xi_k$ and not on $\xi_{k+1}, ..., \xi_N$:

$$x(\xi) = (x_1(\xi_1), x_2(\xi_1, \xi_2), ..., x_N(\xi_1, ..., \xi_N)).$$

In the present paper we treat constriants of the form

$$x(\xi) \in X \quad \text{and} \quad f_i(\xi, x(\xi)) \leqslant 0 \qquad \text{for } i = 1, ..., m \quad \text{and} \quad \xi \in \Xi, \tag{1.1}$$

and the expected cost to be minimized is

$$\Phi(x) = E_\xi f_0(\xi, x(\xi)) = \int_\Xi f_0(\xi, x(\xi))\, \sigma(d\xi), \tag{1.2}$$

where

(a) *X is a nonempty convex subset of R^n, and each f_i for $i = 0, 1, ..., m$ is a real-valued function on $\Xi \times X$ such that $f_i(\xi, x)$ is convex in x for every ξ.*

For reasons explained below, we also make the following topological restrictions:

(b) *Ξ is compact and is the support of σ* (i.e., the smallest closed set of full σ-measure);

(c) *X is compact with nonempty interior,*

(d) *f_i is continuous on $\Xi \times X$ for $i = 0, 1, ..., m$.*

The continuity in (d) ensures in particular that $f_i(\xi, x(\xi))$ is a bounded, Borel-measurable function of ξ whenever $x(\xi)$ is a bounded, Borel measurable function of ξ. Furthermore, since X is bounded, only recourse functions x which are bounded will be needed.

Let $\mathcal{N}$ denote the linear space consisting of all bounded, Borel measurable functions $x: \Xi \to R^n$ which are nonanticipative. The basic problem we study here is

$$\begin{array}{c}\text{minimize the functional (1.2) over all} \\ x \in \mathcal{N} \text{ satisfying the constraints (1.1).}\end{array} \tag{P}$$

Note from the preceding remarks that the expected cost (1.2) is well defined. Furthermore, the functional Φ is convex, and it is to be minimized over a convex set. *Our aim is to characterize the optimal solutions to* (P) *in terms of Lagrange multipliers of some sort for the inequality constraints.*

A problem closely related to (P) is:

$$\begin{array}{c}\text{minimize the functional (1.2) over all } x \in \mathcal{N} \\ \text{satisfying the constraints (1.1) \textit{almost surely},}\end{array} \tag{P$_\infty$}$$

i.e., except perhaps for ξ in a subset of Ξ of measure zero with respect to σ. This can be regarded as a problem on the Banach space $\mathscr{L}_n^\infty = \mathscr{L}^\infty(\Xi, \mathscr{F}, \sigma)$ ($\mathscr{F}$ = Borel field on Ξ), with $\mathscr{N}$ replaced by the subspace $\mathscr{N}_\infty$ of $\mathscr{L}_n^\infty$ comprised of the functions equivalent to functions in $\mathscr{N}$.

In a series of papers [1–6], we have investigated Lagrange multipliers for (P_∞), paying particular attention to the two-stage case. We have shown, roughly speaking, that if at each stage one can make decisions without worrying about the possibility that certain future outcomes of the random variables might leave one with no feasible recourse, then multiplier functions y_i for the constraints $f_i(\xi, x(\xi)) \leqslant 0$ could be obtained under a "strict feasibility" assumption as elements of $\mathscr{L}^1 = \mathscr{L}^1(\Xi, \mathscr{F}, \sigma)$. In general, however, the best one could hope for would be multipliers as elements of the dual space $(\mathscr{L}^\infty)^*$, each of which could be identified with a pair consisting of a function $y_i \in \mathscr{L}^1$ and a certain "singular" component y_i°.

Here we present an alternative approach in which singular multipliers in $(\mathscr{L}^\infty)^*$ are avoided, yet no nonanticipativity condition on the dynamic behavior of the constraints is introduced. The multipliers appear instead as regular Borel *measures* p_i on Ξ, and these can be decomposed in the classical way into an absolutely continuous part with respect to σ and a singular part.

This approach, which was sketched in [7], depends on topological assumptions beyond those needed in our earlier treatment of (P_∞), not only the conditions on Ξ, X, and f_i introduced above, but also a certain property of σ. The reason is that we need to work simultaneously with (P), (P_∞) and still another problem:

$$\begin{aligned}&\text{minimize the functional (1.2) over all}\\ &x \in \mathscr{N}_c \text{ satisfying the constraints (1.1),}\end{aligned} \qquad (\mathrm{P_c})$$

where $\mathscr{N}_c$ consists of all the *continuous* nonanticipative functions $x: \Xi \to R^n$. Obviously, one always has

$$\inf(P_\infty) \leqslant \inf(\mathrm{P}) \leqslant \inf(\mathrm{P_c}), \qquad (1.3)$$

but without further assumptions on σ both inequalities in (1.3) may be strict, as demonstrated by counterexamples in [8]. It is essential to our approach here that equality hold throughout (1.3), because we gain our result for (P) by a marriage of an $\mathscr{L}^1$-multiplier duality theory for (P_∞) and a measure-multiplier duality theory for (P_c).

The further condition to be imposed on σ was developed in [8] with the present application in mind. We assume that

(e) *the probability measure σ is laminary* in the following sense. For any $S \subset R^{\nu}$ and index k, $1 \leqslant k < N$, let

$$\Lambda_k^S(\xi_1, ..., \xi_k) = \{(\xi_{k+1}, ..., \xi_N) \mid (\xi_1, ..., \xi_k, \xi_{k+1}, ..., \xi_N) \in S\},$$

$$S^k = \text{projection of } S \text{ on } R^{\nu_1} \times \cdots \times R^{\nu_k}$$
$$= \{(\xi_1, ..., \xi_k) \mid \Lambda_k^S(\xi_1, ..., \xi_k) \neq \varnothing\}.$$

The conditions for σ to be "laminary" are that

(1) the multifunction Λ_k^{Ξ} is lower semicontinuous relative to Ξ^k (hence actually continuous, since Ξ is compact), and

(2) whenever S is a Borel subset of Ξ with $\sigma(S) = \phi(\Xi) = 1$ such that S^k is also a Borel set in Ξ^k, then $\Lambda_k^S(\xi_1, ..., \xi_k)$ is dense in $\Lambda_k^{\Xi}(\xi_1, ..., \xi_k)$ for almost every $(\xi_1, ..., \xi_k)$ in S^k (with respect to the projection of σ on Ξ^k).

This assumption on σ is satisfied, for example, if σ can be expressed as

$$\sigma(d\xi) = \rho(\xi_1, ..., \xi_N)\, \sigma_1(d\xi_1) \cdots \sigma_N(d\xi_N),$$

where σ_k is a regular probability measure on R^{ν_k} with compact support Ξ_k, and the function ρ is positive on $\Xi_1 \times \cdots \times \Xi_N$. Trivially, it is also satisfied if σ is discrete (Ξ then being a finite set).

Under the foregoing assumptions and a constraint qualification ("strict feasibility" of (P_∞)), we obtain not only a characterization of optimal solutions to (P) in terms of "measure" multipliers, but also the existence of such optimal solutions and the assurance that they can be "approximated" by *continuous* recourse functions.

Restricting recourse functions to any specific class of functions always precludes a certain amount of generality. This is certainly the case if we demand that the recourse functions belong to $\mathscr{C}_n$, the space of R^n-valued continuous functions, but also if they must belong to $\mathscr{L}_n^p$, the space of R^n-valued functions that are p-summable, for $1 \leqslant p < \infty$. The spaces $\mathscr{C}_n$ and $\mathscr{L}_n^\infty$ have the distinct advantage of providing a more natural avenue to the derivation of necessary conditions for optimality. This might possibly be carried out for problems formulated in $\mathscr{L}_n^p$-spaces, $2 \leqslant p < \infty$, but it would certainly entail numerous refinements of the theory now available for multistage programs formulated in $\mathscr{L}^p$-spaces [9–11]. The fact that we are able to show the essential equivalence of (P_∞) and (P_c) with (P) under the above assumptions should therefore be regarded as one of the chief strengths of the present approach.

Although (P_c) is introduced here largely as an adjunct to the study of (P), it is noteworthy that, for large classes of stochastic programming problems, requiring the recourse functions to be in $\mathscr{C}_n$ does not actually engender any additional restriction. Typically, these are stochastic programs whose optimal solution, if it exists, is automatically in $\mathscr{C}_n$. A number of such problems have been identified in connection with the study of optimal decision rules; cf. [12–14]. Moreover, in a number of pratical problems, there might be an a priori specification of the class of acceptable recourse functions such as linear or continuous and piecewise linear, etc. Such a restricted class nearly always turns out to be a certain subspace of $\mathscr{C}_n$; see, for example, [15–17].

Further motivation for studying ($\mathrm{P_c}$) and its relation to (P) and (P_∞) lies in the area of computation. The possibility of approaching optimal recourse functions by continuous ones is likely to be important numerically, especially in algorithms based on discretization.

The equivalence of (P) and (P_∞), also reassuring for computational purposes, has previously been demonstrated for multistage "linear" stochastic programming problems satisfying rather weak regularity conditions [18, Proposition 4.8] and in abstract settings in [8, 19]. This turns out to be a minimal property for the actual derivation of the induced constraints when relatively complete recourse is not available [20, Sects. 4 and 5].

2. Lagrangian Function and Dual Problem

Let $\mathscr{M}$ denote the linear space consisting of all R^m-valued Borel measures $p = (p_1, \ldots, p_m)$ on Ξ, and let

$$\mathscr{P} = \{p \in \mathscr{M} \mid p \geqslant 0\}.$$

Let

$$\mathscr{X} = \{x \in \mathscr{N} \mid x(\xi) \in X \text{ for all } \xi \in \Xi\}.$$

Obviously $\mathscr{X}$ and $\mathscr{P}$ are convex. We define the Lagrangian L on $\mathscr{N} \times \mathscr{M}$ by

$$\begin{aligned} L(x, p) &= \int_\Xi [f_0(\xi, x(\xi))\, \sigma(d\xi) + f_1(\xi, x(\xi))\, p_1(d\xi) + \cdots + f_m(\xi, x(\xi))\, p_m(d\xi)] \\ &\qquad\qquad \text{if} \quad x \in \mathscr{X} \text{ and } p \in \mathscr{P}, \\ &= -\infty \qquad \text{if} \quad x \in \mathscr{X} \text{ but } p \notin \mathscr{P}, \\ &= +\infty \qquad \text{if} \quad x \notin \mathscr{X}. \end{aligned} \tag{2.1}$$

Since for $i = 0, \ldots, m$, $f_i(\xi, x(\xi))$ is a bounded, Borel measurable function of $\xi \in \Xi$ when $x \in \mathscr{X}$, it is clear that L is well defined and finite on $\mathscr{X} \times \mathscr{P}$ with $L(x, p)$ convex in x and concave in p.

It is elementary that

$$\begin{aligned} \sup_{p \in \mathscr{M}} L(x, p) &= \Phi(x) \qquad \text{in (1.2) if } x \text{ satisfies (1.1)}, \\ &= +\infty \qquad \text{otherwise}, \end{aligned} \tag{2.2}$$

and hence

$$\inf(P) = \inf_{x \in \mathcal{N}} \sup_{p \in \mathcal{M}} L(x, y). \tag{2.3}$$

Accordingly, we take the dual of (P) to be the problem

$$\text{maximize } g(p) \text{ over all } p \in \mathcal{M}, \tag{D}$$

where

$$g(p) = \inf_{x \in \mathcal{N}} L(x, p). \tag{2.4}$$

Let $\mathcal{M}_1$ denote the subspace of $\mathcal{M}$ consisting of the measures which are absolutely continuous with respect to the underlying probability σ. For $p = (p_1, ..., p_m) \geqslant 0$ in $\mathcal{M}_1$ and $x \in \mathcal{X}$, we can express L in terms of the Radon–Nikodym derivatives $y_i = dp_i/d\sigma$ by

$$L(x, p) = E_\xi \left\{ f_0(\xi, x(\xi)) + \sum_{i=1}^{m} y_i(\xi) f_i(\xi, x(\xi)) \right\}. \tag{2.5}$$

Note that the multiplier functions y_i then belong to $\mathcal{L}^1$. This provides a connection to the theory in [1–6]. For general $p \in \mathcal{P}$, we have a decomposition

$$p(d\xi) = y(\xi)\, \sigma(d\xi) + q(d\xi),$$

where q is singular with respect to σ, and the expression (2.5) is thus augmented by a singular term

$$\sum_{i=1}^{m} \int_{\Xi} f_i(\xi, x(\xi))\, q_i(d\xi).$$

We need to study the relationship between (D) and the problem

$$\text{maximize } g(p) \text{ in (2.4) over all } p \in \mathcal{M}_1. \tag{D$_1$}$$

Clearly,

$$\sup(\mathrm{D}_1) \leqslant \sup(D) \leqslant \inf(\mathrm{P}). \tag{2.6}$$

We say that (P) is *strictly feasible* if there exist $\tilde{x} \in \mathcal{N}$ and $\epsilon > 0$ such that

$$\tilde{x}(\xi) \in X \quad \text{and} \quad f_i(\xi, \tilde{x}(\xi)) \leqslant -\epsilon \qquad \text{for } i = 1,..., m \quad \text{and all } \xi \in \Xi, \tag{2.7}$$

where B is the closed unit ball of R^n. Similarly, (P_c) is strictly feasible if this condition holds for some $\tilde{x} \in \mathcal{N}_c$. If it merely holds almost surely for some $\tilde{x} \in \mathcal{N}$, we say that (P_∞) is strictly feasible and (P) is *essentially strictly feasible*.

Our main result may now be stated.

THEOREM 1. *Assume, in addition to conditions* (a), (b), (c), (d), *and* (e) *of Section* 1, *that* (P) *is essentially strictly feasible. Then*

$$\min(\mathrm{P}) = \max(\mathrm{D}) = \min(\mathrm{P}_\infty) = \inf(\mathrm{P}_c) = \sup(\mathrm{D}_1). \tag{2.8}$$

In particular, (P) *has at least one optimal solution* $\bar{x}$ *and* (D) *has at least one optimal solution* $\bar{p}$, *and the pairs* $(\bar{x}, \bar{p})$ *consisting of such solutions are precisely the saddle points of the Lagrangian L with respect to* $\mathcal{N} \times \mathcal{M}$ *(or equivalently, with respect to* $\mathcal{X} \times \mathcal{P}$).

3. Proof of the Main Result

The proof of Theorem 1 uses general duality (and minimax) theory in conjunction with a special result about continuous and measurable selections in the context of "feasible" nonanticipative recourse $x: \Xi \to R^n$. This result, which we proved in [8], makes its appearance here in the following form.

PROPOSITION 1. *Under the assumptions of Theorem* 1, *the problem* (P_c) *is strictly feasible and*

$$\min(\mathrm{P}) = \min(\mathrm{P}_\infty) = \inf(\mathrm{P}_c). \tag{3.1}$$

In fact, any feasible solution to (P_∞) *can be modified on a set of measure zero with respect to* σ *to obtain a feasible solution* x *to* (P), *and this can in turn be approximated by feasible solutions to* (P_c), *in the sense that for any* $\delta > 0$ *there exists a feasible solution* x' *to* (P_c) *such that*

$$\sigma(\{\xi \in \Xi \mid \delta < | x(\xi) - x'(\xi) |\}) < \delta. \tag{3.2}$$

Proof. Define the multifunction $D: \Xi \to R^n$ and the function $f: \Xi \times X \to R \cup \{+\infty\}$ by

$$D(\xi) = \{x \in R^n \mid x \in X \text{ and } f_i(\xi, x) \leqslant 0 \text{ for } i = 1,..., m\}, \tag{3.3}$$

$$\begin{aligned} f(\xi, x) &= f_0(\xi, x) \quad &\text{if} \quad & x \in D(\xi), \\ &= +\infty \quad &\text{if} \quad & x \notin D(\xi). \end{aligned} \tag{3.4}$$

Assumptions (a), (b), (c), and (d) imply f is a lower semicontinuous function on $\Xi \times R^n$, such that $f(\xi, x)$ is convex in x and uniformly bounded on the compact set

$$\operatorname{dom} f = \{(\xi, x) \mid x \in D(\xi)\} \subset \Xi \times X. \tag{3.5}$$

We demonstrate, in a moment, that

$$\begin{aligned}
&\text{(i)}\quad D \text{ is lower semicontinuous, and for all } \xi \in \Xi \text{ one has} \\
&\qquad \varnothing \neq \operatorname{int} D(\xi) = \{x \in \operatorname{int} X \mid f_i(\xi, x) < 0 \text{ for } i = 1,..., m\}; \\
&\text{(ii)}\quad \text{essential strict feasibility of (P) implies the existence of} \\
&\qquad \tilde{x}' \in \mathcal{N} \text{ and } \epsilon' > 0 \text{ such that } \tilde{x}'(\xi) + \epsilon' B \subset D(\xi) \text{ almost surely.}
\end{aligned} \tag{3.6}$$

Then, in view of the remarks already made, all the assumptions of [8, Theorem 2, Corollary to Theorem 1] will be fulfilled; these results assert in terms of f and D the desired relation (3.1) and the fact that there exists some $\tilde{x}'' \in \mathcal{N}_c$ and $\epsilon'' > 0$ such that $\tilde{x}''(\xi) + \epsilon'' B \subset D(\xi)$ for all $\xi \in \Xi$. The latter implies by (i) and the continuity of the functions f_i that (P_c) is strictly feasible. The remaining assertions of Proposition 1 are just details furnished by the actual proof of [8, Theorem 2] and by [8, Proposition 7].

We shall verify (ii) first. Let $\tilde{x} \in \mathcal{N}$ satisfy (2.7) almost surely, as in our assumption of essential strict feasibility. Choose a point $a \in \operatorname{int} X$ (as is possible by assumption (c)); let $\delta > 0$ be such that $a + \delta B \subset X$. Assumptions (b) and (d) guarantee the existence of $\alpha \in R$ such that

$$f_i(\xi, a) \leqslant \alpha \qquad \text{for all } \xi \in \Xi \quad \text{and} \quad i = 1,..., m.$$

For a yet undetermined $\lambda \in (0, 1)$, let $\tilde{x}' \in \mathcal{N}$ be the function defined by $\tilde{x}'(\xi) = (1 - \lambda)\,\tilde{x}(\xi) + \lambda a$. From convexity we have almost surely $\tilde{x}'(\xi) + \lambda\delta B \subset X$ and

$$f_i(\xi, \tilde{x}'(\xi)) \leqslant -(1 - \lambda)\,\epsilon + \lambda\alpha \qquad \text{for } i = 1,..., m. \tag{3.7}$$

Take λ sufficiently small that $-(1 - \lambda)\,\epsilon + \lambda\alpha < 0$. The functions f_i are uniformly continuous on the compact set $\Xi \times X$, so there exists $\epsilon' \in (0, \lambda\delta)$ such that for each ξ for which (3.7) holds and each x in $\tilde{x}'(\xi) + \epsilon' B$ we have $f_i(\xi, x) \leqslant 0$, as well as $x \in X$. Then $\tilde{x}'$ has the property described in (ii).

To establish (i), we again make use of $\tilde{x}$ and ϵ satisfying the essential strict feasibility assumption, the property (2.7) holding for all ξ in a certain subset S of Ξ with $\sigma(S) = 1$. Let

$$\Delta = \{(\xi, x) \in \Xi \times X \mid f_i(\xi, x) \leqslant -\epsilon \text{ for } i = 1,..., m\}.$$

The image of Δ under the projection $(\xi, x) \to \xi$ includes S and hence is dense in Ξ, since Ξ is, by assumption (b), the support of σ. But Δ is compact by the compactness of Ξ and X and the continuity of f_i. Therefore, the projection of Δ is all of Ξ; in other words, the set

$$\{x \in X \mid f_i(\xi, x) \leqslant -\epsilon \text{ for } i = 1,..., m\}$$

is nonempty for every $\xi \in \Xi$. Convexity then yields (3.7) (apply [21, Theorem 7.6] to the function $f^\xi(x) = \max_{i=1}^m f_i(\xi, x)$ if $x \in X$, $f^\xi(x) = +\infty$ if $x \notin X$). It follows next from (3.7) and the continuity of f_i that the set

$$\{(\xi, x) \in \Xi \times R^n \mid x \in \operatorname{int} D(\xi)\}$$

is open relative to $\Xi \times R^n$. Since $D(\xi)$ is convex, this implies D is lower semicontinuous [22, Lemma 2, p. 458]. The proof of Proposition 1 is now complete.

PROPOSITION 2. *Assumptions* (a)–(e) *imply* $\min(\mathrm{P}_\infty) = \sup(\mathrm{D}_1)$.

Proof. It is evident that for $x \in \mathcal{N}$ one has

$$\sup_{p \in \mathcal{M}_1} L(x, p) = \Phi(x) \qquad \text{in (1.2) if } x \text{ satisfies (1.1) almost surely,}$$
$$= +\infty \qquad \text{otherwise.}$$

Therefore

$$\inf(\mathrm{P}_\infty) = \inf_{x \in \mathcal{N}} \sup_{p \in \mathcal{M}_1} L(x, p),$$

whereas by definition

$$\sup(\mathrm{D}_1) = \sup_{p \in \mathcal{M}_1} \inf_{x \in \mathcal{N}} L(x, p).$$

We need to show that "min sup = sup inf" holds in this context. As seen from the discussion surrounding (2.5), this is equivalent to proving that

$$\min_{x \in \mathcal{X}_\infty} \sup_{y \in \mathcal{Y}_1} L'(x, y) = \sup_{y \in \mathcal{Y}_1} \inf_{x \in \mathcal{X}_\infty} L'(x, y), \tag{3.8}$$

where $\mathcal{X}_\infty$ is the subset of $\mathcal{L}_n^\infty = \mathcal{L}^\infty(\Xi, \mathcal{F}, \sigma; R^n)$ comprised of the functions equivalent to those in $\mathcal{X}$, $\mathcal{Y}_1$ is the nonnegative orthant of $\mathcal{L}_m^1 = \mathcal{L}^1(\Xi, \mathcal{F}, \sigma; R^m)$, and

$$L'(x, y) = E_\xi \left\{ f_0(\xi, x(\xi)) + \sum_{i=1}^m y_i(\xi) f_i(\xi, x(\xi)) \right\}.$$

(Here we use the fact that if x is equivalent to a function in $\mathcal{N}$ and satisfies $x(\xi) \in X$ almost surely, then x is equivalent to a function in $\mathcal{X}$. This is true by [8, Proposition 7], *since X is compact and σ is laminary*.) Since $\mathcal{X}_\infty$ and $\mathcal{Y}_1$ are convex and $L'(x, y)$ is convex in x and concave in y, we can obtain (3.8) from a minimax theorem of Fan [23, Theorem 2] by demonstrating, relative to the weak topology $w(\mathcal{L}_n^\infty, \mathcal{L}_n^1)$, that $\mathcal{X}_\infty$ is compact and $L'(x, y)$ is lower semicontinuous in x.

Fix any summable function $y: \Xi \to R_+^m$, $y = (y_1, ..., y_m)$, and define ϕ on $\Xi \times R^n$ by

$$\begin{aligned} \phi(\xi, x) &= f_0(\xi, x) + \sum_{i=1}^m y_i(\xi) f_i(\xi, x) \qquad &\text{if} \quad x \in X, \\ &= +\infty &\text{if} \quad x \notin X. \end{aligned}$$

Note from our assumptions (a), (c), and (d) that ϕ is lower semicontinuous in (ξ, x) and convex in x, and of course $\phi(\xi, x)$ is finite if and only if $x \in X$ (where $\operatorname{int} X \neq \varnothing$). It follows that ϕ is a normal convex integrand [24, Lemma 2]. Moreover, there exists by (b), (c), (d), and the summability of y a summable function $\mu: \Xi \to R_+$ such that

$$|\phi(\xi, x)| \leqslant \mu(\xi) \qquad \text{when} \qquad \phi(\xi, x) < +\infty.$$

The integral functional

$$I_\phi(x) = E_\xi\{\phi(\xi, x(\xi))\} \qquad \text{for } x \in \mathscr{L}_n^\infty$$

(which gives $L'(x, y)$ if $x \in \mathscr{N}_\infty$) is therefore well defined with values in $R \cup \{+\infty\}$, and it is bounded above by $E_\xi\{u(\xi)\}$ on

$$\begin{aligned} \operatorname{dom} I_\phi &= \{x \in \mathscr{L}_n^\infty \mid I_\phi(x) < +\infty\} \\ &= \{x \in \mathscr{L}_n^\infty \mid x(\xi) \in X \text{ almost surely}\}. \end{aligned} \tag{3.9}$$

Since $\operatorname{int} X \neq \varnothing$, we know from this and [24, Theorem 2], [22, Theorem 2], that I_ϕ is the conjugate of a certain integral functional on $\mathscr{L}_n^1$ (namely, I_{ϕ^*}), and hence, in particular, I_ϕ is lower semicontinuous relative to $w(\mathscr{L}_n^\infty, \mathscr{L}_n^1)$. (Thus $L'(x, y)$ is $w(\mathscr{L}_n^\infty, \mathscr{L}_n^1)$-lower semicontinuous as a function of $x \in \mathscr{X}_\infty$.) The set (3.9), which can also be expressed as the level set

$$\{x \in \mathscr{L}_n^\infty \mid I_\phi(x) \leqslant E_\xi\{\mu(\xi)\}\},$$

is not only $w(\mathscr{L}_n^\infty, \mathscr{L}_n^1)$-closed but bounded (since X is bounded), and therefore it is $w(\mathscr{L}_n^\infty, \mathscr{L}_m^1)$-compact. We have

$$\mathscr{X}_\infty = \mathscr{N}_\infty \cap \operatorname{dom} I_\phi,$$

and inasmuch as $\mathscr{N}_\infty$ is $w(\mathscr{L}_n^\infty, \mathscr{L}_n^1)$-closed as a subspace of $\mathscr{L}_n^\infty$, we can conclude from this that $\mathscr{X}_\infty$ is $w(\mathscr{L}_n^\infty, \mathscr{L}_n^1)$-compact, as required.

PROPOSITION 3. *Assumptions* (a), (b), (c), (d), *and the strict feasibility of* ($\mathrm{P_e}$) *imply that*

$$\inf(\mathrm{P_e}) = \inf_{x \in \mathscr{N}_e} \sup_{p \in \mathscr{U}} L(x, p) = \max_{p \in \mathscr{U}} \inf_{x \in \mathscr{N}_e} L(x, p). \tag{3.10}$$

Proof. This is a straightforward result of duality theory in the traditional vein. It corresponds to representing the constraints $f_i(\xi, x(\xi)) \leqslant 0$ by $-F(x) \in K$, where F is the mapping from

$$\mathscr{X}_c = \{x \in \mathscr{N}_c \mid x(\xi) \in X \text{ for all } \xi \in \Xi\} \tag{3.11}$$

to the Banach space $\mathscr{C}_n = \mathscr{C}(\Xi, R^m)$ (continuous functions) defined by

$$F(x) = (f_1(\cdot, x(\cdot)),..., f_m(\cdot, x(\cdot))),$$

and K is the nonnegative orthant of $\mathscr{C}_m$. The dual space of $\mathscr{C}_m$ can be identified with $\mathscr{M}$, and L is then the ordinary Lagrangian associated with minimizing the functional (1.2) over $\mathscr{X}_c$ subject to $-F(x) \in K$. Strict feasibility of ($\mathrm{P_c}$) means the existence of $\tilde{x} \in \mathscr{X}_c$ such that $-F(\tilde{x}) \in \operatorname{int} K$, and hence, it ensures the existence of a multiplier vector for the problem, i.e., the validity of (3.10).

PROPOSITION 4. *Under assumptions* (a), (b), (c), (d), *and* (e), *one has*

$$\inf_{x \in \mathscr{N}_c} L(x, p) = \inf_{x \in \mathscr{N}} L(x, p) \qquad \text{for all } p \in \mathscr{M}. \tag{3.12}$$

Proof. Assume first that $p \in \mathscr{P}$, and let θ be a regular Borel probability measure on Ξ with respect to which p and σ are absolutely continuous. (The support of θ is then Ξ.) Setting $\pi_i = dp_i/d\theta \geqslant 0$, where p_i is the ith component of p, we get the representation

$$L(x, p) = \int_{\Xi} f(\xi, x(\xi))\, \theta(d\xi) \qquad \text{for } x \in \mathscr{N}_c\,,$$

where

$$\begin{aligned} f(\xi, x) &= f_0(\xi, x) + \sum_{i=1}^{m} \pi_i(\xi) f_i(\xi, x) &\quad \text{if} \quad & x \in X, \\ &= +\infty &\quad \text{if} \quad & x \notin X. \end{aligned}$$

The same result invoked in the proof of Proposition 1, namely, [8, Theorem 2], when applied to this f asserts (3.12). Its hypothesis is satisfied almost trivially, since the effective domain $\{x \mid f(\xi, x) < +\infty\} = X$ is independent of ξ (with nonempty interior), and $f(\xi, x)$ is summable with respect to $\xi \in \Xi$ for each $x \in X$.

For $p \notin \mathscr{P}$, both sides in (3.12) are trivially infinite and equal, provided it is true that whenever $\mathscr{X} \neq \varnothing$ then $\mathscr{X}_c \neq \varnothing$ (where $\mathscr{X}_c$ is defined by (3.11)). The latter fact can again be obtained from [8, Theorem 2]: Take θ as above, but merely let $f(\xi, x) = 0$ for $x \in X$ and $f(\xi, x) = +\infty$ for $x \notin X$.

Proof of Theorem 1. This is simply a matter of putting together the conclusions of the four propositions. Let $\bar{p}$ denote an element of $\mathscr{M}$ for which the maximum in Proposition 3 is attained. Then $\bar{p}$ is actually an optimal solution to (D) by Proposition 4, and hence,

$$\inf(P_c) = \max(D).$$

At the same time we have

$$\inf(P_c) = \min(P) = \min(P_\infty) = \sup(D_1)$$

by Propositions 1 and 2, so (2.8) is valid and the theorem is proved.

REFERENCES

1. R. T. ROCKAFELLAR AND R. WETS, Stochastic convex programming: Basic duality, *Pacific J. Math.* **62** (1976), 173–195.
2. R. T. ROCKAFELLAR AND R. WETS, Stochastic convex programming: Singular multipliers and extended duality, *Pacific J. Math.* **62** (1976), 507–522.
3. R. T. ROCKAFELLAR AND R. WETS, Stochastic convex programming: Kuhn–Tucker condition, *J. Math. Econ.* **2** (1975), 349–370.
4. R. T. ROCKAFELLAR AND R. WETS, Stochastic convex programming: Relatively complete recourse and induced feasibility, *SIAM J. Control Optimization* **14** (1976), 574–589.
5. R. T. ROCKAFELLAR AND R. WETS, Nonanticipativity and $\mathscr{L}^1$ martingales in stochastic optimization problems, *in* Mathematical Programming Study No. 6 (1976), 170–187; also *in* "Stochastic Systems: Modeling, Identification, and Optimization," (R. Wets, Ed.), North-Holland, Amsterdam (1976).
6. R. T. ROCKAFELLAR AND R. WETS, Nonanticipative feasibility and $\mathscr{L}^1$-multipliers in multistage stochastic programming, *SIAM J. Control Optimization*, in press.
7. R. T. ROCKAFELLAR, Lagrange multipliers for an N-stage model in stochastic convex programming, *in* "Analyse Convexe et ses Applications 1974" (J.-P. Aubin, Ed.) Lecture Notes in Economics and Mathematical Systems, Vol. 102, pp. 180–187, Springer–Verlag, Berlin, 1974.
8. R. T. ROCKAFELLAR AND R. WETS, Continuous versus measurable recourse in N-stage stochastic programming, *J. Math. Anal. Appl.* **48** (1974), 836–859.
9. R. WETS, Problèmes duaux en programmation stochastique, *C. R. Acad. Sci. Paris Ser. A–B* **270** (1970), 47–50.
10. M. EISNER AND P. OLSEN, Duality for stochastic programming, interpreted as L. P. in L_p-space, *SIAM J. Appl. Math.* **28** (1974), 779–792.
11. M. EISNER AND P. OLSEN, Duality in probabilistic programming, *in* "Stochastic Programming: Proceedings of the 1974 Oxford International Conference" (M. Dempster, Ed.), Academic Press, New York, 1976.
12. D. WALKUP AND R. WETS, A note on decision rules for stochastic programming, *J. Comput. Sci.* **2** (1968), 305–311.
13. S. GARTSKA AND R. WETS, On decision rules in stochastic programming, *Mathematical Programming* **7** (1974), 117–143.
14. E. TSOY, "Decision Rules in Problems of Stochastic Programming," Doctoral thesis, Moscow, 1974 [in Russian].
15. M. EISNER, R. KAPLAN, AND T. SOOLEN, Admissible decision rules for the E-model of chance-constrained programming, *Management Sci.* **17** (1971), 337–353.

16. D. YUDIN AND E. TSOY, A priori decision rules in problems of multistage stochastic programming, *Ekon. Mat. Metody. SSSR* **9** (1973), 947–961 [in Russian].
17. D. YUDIN, "Mathematical Methods of Control under Uncertainty," Sovetskoe Radio, Moscow, 1974 [in Russian].
18. P. OLSEN, Multistage stochastic programming with recourse: The equivalent deterministic problem, *SIAM J. Control Optimization*, in press.
19. R. WETS, Induced constraints for stochastic optimization problems, *in* "Techniques of Optimization" (A. Balakrishnan, Ed.), pp. 433–443, Academic Press, New York, 1972.
20. R. WETS, Stochastic programs with fixed recourse: The equivalent deterministic Program, *SIAM Rev.* **16** (1974), 309–339.
21. R. T. ROCKAFELLAR, "Convex Analysis," Princeton Univ. Press, Princeton, N. J., 1970.
22. R. T. ROCKAFELLAR, Integrals which are convex functionals, II, *Pacific J. Math.* **39** (1971), 439–469.
23. K. FAN, Minimax theorems, *Proc. Nat. Acad. Sci. USA* **39** (1953), 42–47.
24. R. T. ROCKAFELLAR, Integrals which are convex functionals, *Pacific J. Math.* **24** (1968), 525–539.

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium