Mathematical Programming Study 6 (1976) 170–187.
North-Holland Publishing Company

NONANTICIPATIVITY AND $\mathscr{L}^1$-MARTINGALES IN STOCHASTIC OPTIMIZATION PROBLEMS

R. T. ROCKAFELLAR*
University of Washington, Seattle, Wash., U.S.A.

and

R. J.-B. WETS**
University of Kentucky, Lexington, Ky, U.S.A.

Received 1 July 1975
Revised manuscript received 11 August 1975

Necessary and sufficient conditions for optimality are derived for multistage stochastic programs. In particular it is shown that under some standard regularity conditions and a condition of "nonanticipative feasibility", a system of Lagrange multipliers, characterized by a martingale property, can be associated with the constraints of the problem. Nonanticipative feasibility is expressed in terms of the nonanticipativity of a certain multifunction and is shown to be related to the more familiar concept — in stochastic programming — of relatively complete recourse. It is also shown that this restriction renders possible the justification of the dynamic programming technique.

1. Introduction

We study a class of multistage stochastic optimization problems, motivated by the following heuristic model. First a random event is observed; this singles out an element ξ_1 of $\mathbf{R}^{\nu_1}$. Based on this observation, a decision x_1 is chosen which is an element of R^{n_1}. A second observation is then made yielding ξ_2 in $\mathbf{R}^{\nu_2}$, and a (*recourse*) *decision* x_2 in $\mathbf{R}^{n_2}$ is selected based on the information acquired so far (ξ_1, ξ_2). This continues until the N^{th} stage; at each stage k a new observation $\xi_k \in \mathbf{R}^{\nu_k}$ is made and the (recourse) decision x_k is selected as a function of $(\xi_1, \ldots, \xi_k)$. At the N^{th} stage, we determine $\xi_N \in \mathbf{R}^{\nu_N}$ and choose $x_N \in \mathbf{R}^{n_N}$. The result of this sequential decision process is a "cost" $f(x, \xi)$, where

* Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under AFOSR Grant Number 72-2269
** Supported in part by the National Science Foundation under grant MPS 75-07028

$$\xi = (\xi_1, \xi_2, \ldots, \xi_N) \in \mathbf{R}^\nu = \mathbf{R}^{\nu_1} \times \mathbf{R}^{\nu_2} \times \ldots \times \mathbf{R}^{\nu_N},$$

$$x = (x_1, x_2, \ldots, x_N) \in \mathbf{R}^n = \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \ldots \times \mathbf{R}^{n_N}.$$

The objective is to find a *decision rule* — also called a *recourse function* —

$$\xi \mapsto x(\xi)$$

which is *nonanticipative* (the decision x_k depends only on the past observations $\xi_1, \ldots, \xi_k$, but not on the future $\xi_{k+1}, \ldots, \xi_N$) and minimizes the expected cost. This is a *multistage* (or *N-stage*) *stochastic program* (*with recourse*).[1]

Constraints are introduced by allowing the function f to take on the value $+\infty$; $f(\xi, x) = +\infty$ means that if the sequence of events $\xi = (\xi_1, \xi_2, \ldots, \xi_N)$ does occur then the sequence of decisions $x = (x_1, x_2, \ldots, x_N)$ is unacceptable, or in other words x fails to satisfy the constraints of the problem. For a fixed $\xi \in \mathbf{R}^\nu$, we denote by $D(\xi)$ the *effective domain* of f with respect to x, i.e.,

$$D(\xi) = \operatorname{dom} f(\xi, \cdot) = \{x \mid f(\xi, x) < +\infty\}. \tag{1.1}$$

The map $D : \xi \mapsto D(\xi)$ is a multifunction on $\mathbf{R}^\nu$ yielding a description of the feasibility region as a function of the random events.

In a sequel to this paper [1] we deal with multistage stochastic programs in which the constraints appear explicitly in the formulation of the problem; here we concentrate on the restriction introduced by the requirement that recourse functions must be nonanticipative. The goal is to show that it is possible (under certain regularity conditions) to associate with nonanticipativity a price system (i.e., a system of Lagrange multipliers) which turns out to be a stochastic process with summable paths satisfying a martingale property. Our earlier results in this direction, sketched out in [2], involved a number of technical assumptions that limited the scope of the applications. The existence of a price system associated with nonanticipativity was first pointed out in [3, Section 4] in connection with our work on stochastic convex programs dealing with a special case with $N = 2$.

Price systems associated with constraints appearing in multistage stochastic programs have been investigated by Yudin [4], Eisner and Olsen [5, 6]

[1] If the sequential decision process actually starts with a decision rather than an observation of a random event or if it terminates with an observation of a random quantity with no subsequent decision allowed, the problem can nevertheless be made to fit into the mold layed out here, by introducing either a fictitious first stage random event or a fictitious decision in the last stage which would not affect the value of the objective function.

and also by Dynkin [7, 8, 9], Evstigneev [10, 11] and Radner [12] in the framework of models for optimal economic development. The model studied by Yudin [4] involves constraints on certain means, or in other words chance constraints, and does not quite fit in the class of problems described above. But it can be "reduced" to this case by a well-known equivalence [13] exploited by Eisner and Olsen, who exhibit in [6] a price system associated with chance constraints, using the results of [5] for a model without chance constraints. The model considered here is more general than the one studied by Eisner and Olsen [5] in the sense that it allows for convex objective and convex constraints rather than linear objective and linear constraints. Also, our main objective is to establish the *existence* of a "nice" price system, a question left open in [5]. In this sense, the results presented here are more closely connected to those of Dynkin, Evstigneev and Radner, who have similar preoccupations, at least in terms of the properties of the equilibrium price system to be associated with optimal growth programs. These problems are in fact multistage programs of a very specific nature. Our results are actually applicable to that class of problems and yield some refinements of the results of Dynkin et al.

2. Formulation

A rigorous formulation of the problem demands appropriate assumptions on the underlying probability space, on the function f and the associated multifunction D, as well as some restriction on the class of admissible recourse functions in particular in ensuring that the expected cost is in some sense well-defined.

A. *The probability space.* Let $(\Xi, \mathcal{F}, \sigma)$ denote the underlying probability space, with Ξ a Borel subset of $\mathbf{R}^\nu$, $\mathcal{F}$ the Borel field on Ξ and σ a regular probability measure on $(\Xi, \mathcal{F})$. We assume that the images of Ξ under projections $\mathbf{R}^\nu \mapsto \mathbf{R}^{\nu_1} \times \ldots \times \mathbf{R}^{\nu_k}$ are Borel, so that the "marginal" probability spaces induced by these projections are describable in the same terms. This will automatically be so, if for example Ξ is a product of nonnegative orthants, integer lattices, whole Euclidean spaces, etc.

B. *Admissible recourse functions.* For present purposes, we deem admissible as recourse functions only those (Borel) *measurable, essentially bounded* functions $x : \Xi \to \mathbf{R}^n$ which are *essentially nonanticipative* in the sense that for some (Borel) measurable set $\Xi' \subset \Xi$ with $\sigma(\Xi') = 1$ the restriction to Ξ'

of each of the component functions x_k for $k = 1, \ldots, N-1$ depends only on $(\xi_1, \ldots, \xi_k)$. The set of all these is denoted by $\mathcal{N}_\infty$. An equivalent definition of $\mathcal{N}_\infty$ is that it consists of the functions which differ by only a (Borel) null function from a (Borel) measurable function which is bounded and nonanticipative relative to Ξ itself. Note that $\mathcal{N}_\infty$ can be regarded as a linear subspace of $\mathscr{L}_n^\infty = \mathscr{L}^\infty(\Xi, \mathscr{F}, \sigma; \mathbf{R}^n)$, which is closed, not only in the norm topology, but in the weak topology w-$(\mathscr{L}_n^\infty, \mathscr{L}_n^1)$.

Our restriction to essentially bounded recourse functions is chiefly for technical reasons, but it plays a very substantial role in our approach. Certainly it involves no real loss of generality if the sets $D(\xi)$ are uniformly bounded for $\xi \in \Xi$, which is not unreasonable in real applications. Further justification is given in our paper [14, §3], along with a proof in the case $N = 2$ that the measurability restriction does not have some hidden, undesirable effect on the optimal value in the problem. Results in this context on the approximation of the minimum by means of *continuous* recourse functions when Ξ is bounded may be found in [15].

Let P_k denote the generic operator projecting a vector with L $(\geq k)$ components (each one being possibly a vector) into its first k components, e.g. for $1 \leq k \leq N$,

$$P_k x = P_k(x_1, x_2, \ldots, x_N) = (x_1, x_2, \ldots, x_k)$$

and

$$P_k \xi = P_k(\xi_1, \xi_2, \ldots, \xi_N) = (\xi_1, \xi_2, \ldots, \xi_k).$$

The *tail projection* of P_k is denoted by $(I - P_k)$, specifically, $(I - P_k)\xi = (\xi_k, \ldots, \xi_N)$. Nonanticipativity of a recourse function can be expressed in terms of projection, namely $\xi \mapsto x(\xi)$ is nonanticipative if for all $k = 1, 2, \ldots, N$ and all $\xi \in \Xi$ we have that $P_k x(\xi)$ depends only on $P_k \xi$. The set of admissible recourse functions is thus given equivalently by

$$\mathcal{N}_\infty = \{x \in \mathscr{L}_n^\infty \mid \text{for all } k = 1, \ldots, N,\ P_k x(\xi) = P_k x(\bar{\xi}) \text{ a.s. whenever } P_k \xi = P_k \bar{\xi}\}.$$

C. *The objective function.* The function $f : \Xi \times \mathbf{R}^n \to]-\infty, +\infty]$ is a *normal convex integrand*, by which we mean that

(i) for each ξ in Ξ, the function $x \mapsto f(\xi, x)$ is convex, proper (not identically $+\infty$) and lower semicontinuous;

(ii) there exists a countable collection U of measurable functions from Ξ

into $\mathbf{R}^n$ such that $\xi \mapsto f(\xi, u(\xi))$ is measurable for all $u \in U$, while for each fixed ξ in Ξ the set $U(\xi) \cap D(\xi)$ is dense in $D(\xi)$ where

$$U(\xi) = \{u(\xi) \in \mathbf{R}^n \mid u \in U\}. \tag{2.1}$$

Various criteria for normality of convex integrands are given in [16, 17, 18 and 19]. These properties of f imply in particular that the function $\xi \mapsto f(\xi, u(\xi))$ is measurable whenever u is a measurable function from Ξ to $\mathbf{R}^n$ [16, Corollary of Lemma 5]. A normal convex integrand is said to be *inf-compact* if for all $\xi \in \Xi$, the function $x \mapsto f(\xi, x)$ is also inf-compact, i.e., if for all $\alpha \in \mathbf{R}$, the (level) set $\{x \mid f(\xi, x) \leq \alpha\}$ is compact.

The multifunction $\xi \mapsto \operatorname{dom} f(\xi, \cdot) = D(\xi)$ associates with each $\xi \in \Xi$ a certain set $D(\xi) \subset \mathbf{R}^n$ of acceptable decisions. *We assume that the sets $D(\xi)$, $\xi \in \Xi$ are closed and uniformly bounded.* The interpretation of $D(\xi)$ as the set of acceptable decisions, a set typically determined by a system of inequalities involving continuous convex functions, makes $D(\xi)$ naturally closed. The uniform boundedness is a definite restriction, but an assumption of that type can not be avoided altogether when admissible recourse functions must be elements of $\mathscr{L}_n^\infty$. These assumptions, when combined with the fact that f is a normal convex integrand, imply that D is a *compact-convex valued, uniformly bounded, measurable multifunction* [18] and that f is an inf-compact normal convex integrand.

The objective function of the stochastic program is taken to be a functional I_f on $\mathscr{L}_n^\infty$ defined by

$$I_f(x) = \mathbf{E}\{f(\xi, x(\xi))\}. \tag{2.2}$$

For every measurable function, the value of the integral (expectation) is finite or $-\infty$ (in the standard sense) if $f(\cdot, x(\cdot))$ is majorized by a summable function of ξ; otherwise its value is set at $+\infty$. The effective domain of I_f is simply

$$\mathscr{D} = \{x \in \mathscr{L}_n^\infty \mid I_f(x) < +\infty\}. \tag{2.3}$$

The definition of the integral implies that

$$\mathscr{D} \subset \mathscr{D}_\sigma = \{x \in \mathscr{L}_n^\infty \mid x(\xi) \in D(\xi) \text{ a.s.}\}. \tag{2.4}$$

If we take as premise that a recourse function, which associates with almost all ξ, sequences of acceptable decisions, should have finite expected cost; then $\mathscr{D} = \mathscr{D}_\sigma$. This will certainly be the case if there exists a summable function $\mu : \Xi \to \mathbf{R}$ such that $x \in D(\xi)$ implies that $|f(\xi, x)| \leq \mu(\xi)$. *We shall assume that such a function μ exists*; we therefore say that x is *feasible* if $x \in \mathscr{L}_n^\infty$, x is nonanticipative and $x(\xi) \in D(\xi)$ a.s. or equivalently if

$$x \in \mathscr{D} \cap \mathcal{N}_\infty. \tag{2.5}$$

Observe that the postulated function μ satisfies

$$\mu(\xi) \geqq \inf_x f(\xi, x). \tag{2.6}$$

so that

$$I_{f^*}(0) = \mathbf{E}\{f^*(\xi, 0)\} \leqslant -\mathbf{E}\{\mu(\xi)\} < +\infty, \tag{2.7}$$

where $f^*(\xi, \cdot)$ is the conjugate of $f(\xi, \cdot)$, that is,

$$f^*(\xi, x^*) = \operatorname{Sup}\{\langle x, x^*\rangle - f(\xi, x) \mid x \in \mathbf{R}^n\}. \tag{2.8}$$

In view of the uniform boundedness of D, (2.7) also implies that $I_{f^*}(y) < +\infty$ for all $y \in \mathscr{L}^1_n$. (To see this, note that the condition $D(\xi) \subset \beta B$ for $\xi \in \Xi$, where $\beta > 0$ and B is the closed unit ball, implies that $f^*(\xi, x^*)$ is finite and Lipschitz in x^*, since

$$|f^*(\xi, x^{*\prime}) - f^*(\xi, x^*)| \leqslant \beta \|x^{*\prime} - x^*\| \tag{2.9}$$

for all $\xi \in \Xi$ and all $x^*, x^{*\prime}$ in $\mathbf{R}^n$. Thus, if $I_{f^*}(y) < +\infty$ for one $y \in \mathscr{L}^1_n$, then the same must hold for *every* $y \in \mathscr{L}^1_n$.)

The functional I_f is clearly convex. Moreover, from (2.7) and the existence of at least one feasible solution, it follows from Theorem 2 of [16] that I_f is also lower semicontinuous relative to the weak topology w-$(\mathscr{L}^\infty_n, \mathscr{L}^1_n)$. Actually it is *inf-compact* relative to this topology: the uniform boundedness of D implies that the level sets, which are w-$(\mathscr{L}^\infty_n, \mathscr{L}^\infty_1)$ closed, are contained in $\{x \in \mathscr{L}^\infty_n \mid \|x\| \leqslant \beta\}$ since $D(\xi) \subset \beta B$ for all ξ.

D. *Nonanticipative feasibility.* We define the multifunction $\xi \mapsto D(\xi)$ to be *nonanticipative* if for all $k = 1, \ldots, N$ and for all $\xi \in \Xi$ the projection of $D(\xi)$ on $\mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \ldots \times \mathbf{R}^{n_k}$, i.e., $P_k D(\xi)$, depends only on $P_k \xi$. Note that if D is single valued, this new definition reduces to the customary one for functions. *We shall assume that D is nonanticipative.*

To gain insight into this concept, suppose that up to stage k, the events $(\hat{\xi}_1, \ldots, \hat{\xi}_K)$ have been observed and $(x_1, \ldots, x_k) \in P_k D(\xi)$ an associated sequence of decisions, where ξ is an element of Ξ such that $P_k \xi = (\hat{\xi}_1, \ldots, \hat{\xi}_k)$. Nonanticipativity of D implies that there exists $x_{k+1}, \ldots, x_N$, such that $(x_1, \ldots, x_l) \in P_l D(\xi)$ for $l = k+1, \ldots, N$. This follows directly from the fact that $P_k D(\xi)$ only depends on $P_k \xi$ and that P_k is a projection. In other words, nonanticipativity of D means that the constraints on the choice of x_k in stage k depend only on the past decisions $x_1, \ldots, x_{k-1}$ and the realizations $(\hat{\xi}_1, \ldots, \hat{\xi}_k)$; there is no further restriction induced by the fact that feasible recourse must be possible in the future. Multistage stochastic programs that possess this property are termed *stochastic programs with relatively complete recourse.* Every stochastic program can be reduced to a stochastic program with relatively complete recourse by introducing explicitly the *induced constraints*, i.e., those additional constraints induced on $(x_1, \ldots, x_k)$ by the requirement of feasible future recourse. The properties of induced constraints [20] and the role they play in the theory of necessary and sufficient conditions [21] have been treated in detail for stochastic programs with $N = 2$.

E. *The Multistage stochastic program.* The problem at hand can be expressed as:

$$\text{P} \qquad \text{Find} \quad \inf I_f \quad \text{on } \mathscr{N}_\infty.$$

We shall assume that P is *feasible*, i.e. possesses at least one feasible solution. (It is said to be *strictly feasible* if there exists a function $\bar{x} \in \mathscr{N}_\infty$ and $\varepsilon > 0$ such that $\bar{x}(\xi) + \varepsilon B \subset D(\xi)$ for almost all ξ in Ξ.) In fact, if it is feasible it is also *solvable*. The existence of an optimal solution follows directly from the facts noted above, that relative to the topology w-$(\mathscr{L}^\infty_n, \mathscr{L}^1_n)$, I_f is inf-compact and $\mathscr{N}_\infty$ is closed.

3. The value function

An N-stage stochastic programming problem can be "reduced" to a k-stage stochastic program by relying—as in dynamic programming—on the value function. Let $(\Xi^k, \mathscr{F}^k, \sigma^k)$ be the marginal probability space associated with the random variables appearing in the first k stages with the notation $\xi^k = (\xi_1, \ldots, \xi_k) = P_k\xi$, $\Xi^k = P_k\Xi$. By $x^k = (x_1, \ldots, x_k)$ we denote an element of $(\mathbf{R}^{n^k} = \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \ldots \times \mathbf{R}^{n_k}$ and by $\mathscr{N}^k_\infty$ the class of essentially nonanticipative functions in $\mathscr{L}^\infty_{n^k} = \mathscr{L}^\infty(\Xi^k, \mathscr{F}^k, \sigma^k; \mathbf{R}^{n^k})$, $k = 1, \ldots, N$. For $k = 1, 2, \ldots, N$, we consider the following stochastic program:

$$\text{Find} \quad \inf I_{q_k} \quad \text{on } \mathscr{N}^k_\infty \tag{3.1}$$

where the *value function* q_k with domain $\Xi^k \times \mathbf{R}^{n^k}$, is defined recursively as follows: for $l = k+1, \ldots, N$ we set

$$q_{l-1}(\xi^{l-1}, x^{l-1}) = \boldsymbol{E}_{\xi^l}\left\{\inf_{x_l} q_l(\xi^l, x^l) \mid \xi^{l-1}\right\}, \tag{3.2}$$

$$q_N(\xi^N, x^N) = f(\xi, x). \tag{3.3}$$

Here $\mathbf{E}_{\xi^t}\{\cdot \mid \xi^{t-1}\}$ is (a version of) the conditional expectation with respect to the random variable ξ^t given the random variable ξ^{t-1}. The stochastic program (3.1) is well defined if the expression appearing in the left-hand side of (3.2) makes sense and is a normal convex integrand. The objective of this section is to show that for multistage stochastic programs as considered here, the stochastic program (3.1) is a program of the same type as P. Results along this line have been obtained by Olsen [22, 23], for "linear" multistage stochastic programs with convex criterion function, and by Evstigneev [24] who in a somewhat different framework derives the "equivalence" of optimal solutions; see Lemma 3 below and its application in Theorem 1.

Conditional expectations are defined up to an equivalence relation. However since $\mathcal{F}$ is the Borel field and σ is a regular Borel measure on Ξ, a theorem of Doob [25, Section 27] guarantees the existence of regular conditional probabilities so that a member of the equivalence class can be represented as an indefinite integral with respect to this regular conditional probability. Henceforth, conditional expectations will always be "regular" conditional expectations. If one does not adhere to this restricted class of conditioning, a number of inconsistencies may arise such as exemplified by Olsen in [22, Section 2].

The properties of q_k can be derived recursively. Consequently it will suffice to consider the notationally simpler case of a two-stage stochastic program and show that the reduction to a one-stage program preserves all the desired properties of the integrand. Thus here we set $N = 2$, $q_2 = f$ and simply write q for q_1. As a first step we intend to show that the function q from $\Xi_1 \times \mathbf{R}^{n_1}$ into $]-\infty, +\infty]$ and given by

$$q(\xi_1, x_1) = \mathbf{E}\left\{\inf_{x_2} f(\xi, x) \mid \xi_1\right\} \tag{3.4}$$

is an inf-compact normal convex integrand with dom $q(\xi_1, \cdot) = P_1 D(\xi)$.

Lemma 1. *Suppose that f is an inf-compact normal convex integrand. Then so is g for*

$$g(\xi, x_1) = \inf_{x_2} f(\xi, x_1, x_2). \tag{3.5}$$

Proof. Let $G_f(\xi) = \operatorname{epi} f(\xi, \cdot)$ and $G_g(\xi) = \operatorname{epi} g(\xi, \cdot)$. The condition that f be a normal convex integrand is equivalent to the condition that $\xi \mapsto G_f(\xi)$ be a measurable multifunction with nonempty closed convex values [18, Theorem 4]. Because of the inf-compactness of f, $G_g(\xi)$ is the nonempty closed convex set $QG_f(\xi)$, where Q is the projection: $(x_1, x_2, \alpha) \to (x_1, \alpha)$ from $\mathbf{R}^n \times \mathbf{R}$ into $\mathbf{R}^{n_1} \times \mathbf{R}$. Since the projection of a measurable multifunction is measurable (a direct consequence of the definition adopted in [18]), we conclude that $\xi \mapsto G_g(\xi)$ is measurable and hence again from [18, Theorem 4] that g is a normal convex integrand. The inf-compactness follows also from a projection argument, cf. for example [26].

A measurable multifunction is said to be *summable* if every measurable selector is summable. When D satisfies the conditions layed out in Section 2.C, D is a summable multifunction. A somewhat weaker condition is that there exists a summable function $\beta : \Xi \to R^n$ such that $D(\xi) \subset \beta(\xi) B$.

Lemma 2. *Suppose that* $g : \Xi \times \mathbf{R}^{n_1} \to]-\infty, +\infty]$ *is an inf-compact normal convex integrand where* $\Xi \subset \mathbf{R}^{n_1} \times \mathbf{R}^{n_2}$, *that* $\xi \mapsto \operatorname{dom} g(\xi, \cdot) = D(\xi) \neq \emptyset$ *is a nonanticipative summable multifunction. Suppose moreover that u is measurable and that* $u(\xi) \in D(\xi)$ *a.s. imply that* $g(\cdot, u(\cdot))$ *is summable. Then the function q defined by*

$$q(\xi_1, x_1) = \mathbf{E}\{g(\xi, x_1) \mid \xi_1\} \tag{3.6}$$

is an inf-compact normal convex integrand with $\operatorname{dom} q(\xi_1, \cdot) = P_1 D(\xi) = D_1(\xi_1)$. *Moreover if* u_1 *is measurable and* $u_1(\xi_1) \in D_1(\xi_1)$ *a.s., then* $q(\cdot, u_1(\cdot))$ *is summable.*

Proof. The inf-compactness and convexity of $x_1 \mapsto q(\xi_1, x_1)$ are a consequence of the inf-compactness and convexity of $g(\xi, \cdot)$ [26, Proposition 6]. The properness of $q(\xi_1, \cdot)$ follows from the fact that by nonanticipativity $x \in D(\xi)$ implies that $x_1 \in P_1 D(\xi) = D_1(\xi_1)$ and that $g(\cdot, x_1)$ is a summable function. This also yields the assertion about the effective domain of $q(\xi_1, \cdot)$.

Also if $v : \Xi_1 = P_1 \Xi \to \mathbf{R}^{n_1}$ is any measurable function then $q(\xi_1, v(\xi_1)) = \mathbf{E}\{g(\xi, v(\xi_1)) \mid \xi_1\}$ is a measurable function since it is a conditional expectation provided that $\xi \mapsto g(\xi, v(P_1\xi))$ is measurable; but this follows immediately from the fact that g is a normal convex integrand [14, Corollary to Lemma 5]. If moreover $v(\xi_1) \in D(\xi_1)$ a.s., then by hypothesis $g(\xi, v(P_1\xi))$ is summable in ξ and the summability of $q(\xi_1, v(\xi_1))$ results directly from (the extended) Fubini's theorem. It remains to establish the density condition. Let $V_1 = \{v \mid v(\xi_1) = \mathbf{E}\{u(\xi) \mid \xi_1\}, u \in U\}$ for U the countable collection of measurable functions invoked for the normality of g. We may assume that $u \in U$ implies that $u(\xi) \in D(\xi)$ for all $\xi \in \Xi$. (If this was not the case we could replace the collection U by the collection $U' = \{u' =$

$\text{prox}(u \mid \psi_D) \mid u \in D\}$ [16].) The elements of V_1 are measurable functions (Radon–Nikodym derivatives) well-defined (regular conditional expectations) and finite valued for all $\xi_1 \in \Xi_1$ (Fubini's Theorem) since D is a summable multifunction. The fact that $V_1(\xi_1)$ is dense in $D_1(\xi_1)$ now simply follows from the facts that $U(\xi)$ is dense in $D(\xi)$ and that $\mathbf{E}\{\cdot \mid \xi_1\}$ is a continuous linear surjection.

Lemma 3. *Suppose that f is a normal convex integrand defined on $\Xi \times \mathbf{R}^n$ and such that the multifunction $\xi \mapsto \text{dom}\, f(\xi, \cdot) = D(\xi)$ is nonanticipative with $\{D(\xi), \xi \in \Xi\}$ uniformly bounded. Let q be the value function defined by*

$$q(\xi_1, x_1) = \mathbf{E}\left\{\operatorname*{Inf}_{x_2} f(\xi, x) \mid \xi_1\right\} \tag{3.7}$$

and suppose that the stochastic program P

$$\textit{find} \quad \inf I_f \quad \textit{on}\ \mathcal{N}_\infty \tag{3.8}$$

admits an optimal solution. Then $\bar{x}$ solves (3.8) if and only if $\bar{x}_1 = P_1\bar{x}$ solves the stochastic program

$$\textit{find} \quad \inf I_q \quad \textit{on}\ \mathcal{N}^1_\infty = \mathscr{L}^\infty_{n_1} \tag{3.9}$$

and satisfies

$$f(\xi, \bar{x}(\xi)) = \operatorname*{Inf}_{x_2} f(\xi, \bar{x}_1(\xi), x_2) \quad a.s. \tag{3.10}$$

Proof. Let $\bar{x}$ be a minimum point of I_f on $\mathcal{N}_\infty$. To show (3.10) we first observe that the function

$$(\xi, x_2) \mapsto f(\xi, \bar{x}_1(\xi), x_2)$$

is an inf-compact normal convex integrand. The inf-compactness is a direct consequence of the boundedness of the $\{D(\xi), \xi \in \Xi\}$ and the lower semicontinuity of $\xi \mapsto f(\xi, \cdot)$. The normality is proved in a similar fashion to Lemma 1. (The inf-compactness allows us to replace Inf by Min in (3.10).) Thus the function

$$\xi \mapsto \alpha(\xi) = \operatorname*{Min}_{x_2} f(\xi, \bar{x}_1(\xi), x_2)$$

is measurable [16, Lemma 5] and almost surely finite ($\alpha \leq f(\cdot, \bar{x})$). If there exists a set $S \subset \Xi$ of positive measure such that $\alpha < f(\cdot, \bar{x})$ on S, let β be a real valued measurable function such that

$$\alpha(\xi) < \beta(\xi) < f(\xi, \bar{x}(\xi)) \tag{3.11}$$

for ξ in S. By [16, Lemma 6] there exists $\tilde{x}_2 : S \to \mathbf{R}^{n_2}$ a measurable function such that

$$f(\xi, \bar{x}_1(\xi), \tilde{x}_2(\xi)) \leq \beta(\xi) < f(\xi, \bar{x}(\xi)) \tag{3.12}$$

for all $\xi \in S$. Now extend the domain of $\tilde{x}_2$ to Ξ by setting $\tilde{x}_2(\xi) = \bar{x}_2(\xi)$ for $\xi \in \Xi \backslash S$. It follows that $I_f(\bar{x}_1, \tilde{x}_2) < I_f(\bar{x})$, which contradicts the optimality of $\bar{x}$. This proves (3.10) and also shows that

$$q(\xi_1, \bar{x}_1(\xi_1)) = \mathbf{E}\{f(\xi, \bar{x}(\xi)) \mid \xi_1\}, \tag{3.13}$$

since the nonanticipativity of $\bar{x}$ allows us to redefine $\bar{x}_1$ on a set of measure zero so that it is a well-defined function of ξ_1.

Now suppose that $\tilde{x}_1$ minimizes I_q on $\mathscr{L}^\infty_{n_1}$. In view of Lemmas 1 and 2 and the uniform boundedness of D, I_q is w-$(\mathscr{L}^\infty_{n_1}, \mathscr{L}^1_{n_1})$ inf-compact, which guarantees the existence of an optimal solution to (3.9). Then there exists an essentially bounded measurable function from Ξ to $\mathbf{R}^{n_2}$ such that

$$f(\xi, \tilde{x}_1(\xi), \tilde{x}_2(\xi)) = \operatorname*{Min}_{x_2} f(\xi, \tilde{x}_1(\xi), x_2(\xi)) \quad \text{a.s.} \tag{3.14}$$

To see this, first note that nonanticipative feasibility implies the feasibility of the mathematical program

$$\text{find} \quad \inf f(\xi, \tilde{x}_1(\xi), \cdot) \quad \text{on } \mathbf{R}^n \tag{3.15}$$

for all $\xi \in \Xi$ (this might involve a possible redefinition of $\tilde{x}_1$ on a set of measure zero so that $\tilde{x}_1(\xi_1) \in P_1 D(\xi)$ for all $\xi \in \Xi$). Existence of an optimal solution to (3.15) follows from the inf-compactness. The function $\xi \mapsto \tilde{\alpha}(\xi) = \operatorname{Min} f(\xi, \tilde{x}_1(\xi), \cdot)$ is measurable and thus the multifunction

$$\xi \mapsto \{x_2 \mid f(\xi, \tilde{x}_1(\xi), x_2) \leq \tilde{\alpha}(\xi)\}$$

is measurable with nonempty compact convex values, $(\xi, x_2) \mapsto f(\xi, \tilde{x}_1(\xi), x_2)$ being an inf-compact normal convex integrand [18, Corollary 4.3]. Hence there exists a measurable selector satisfying (3.14). Now let $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$; we have that $I_q(\tilde{x}_1) = I_f(\tilde{x})$. The proof is complete if we observe that by (3.13) $I_q(\bar{x}_1) = I_f(\bar{x})$ and that $I_q(\tilde{x}_1) < I_f(\bar{x}_1)$ would contradict the optimality of $\bar{x}$.

We now turn to the applications of these three lemmas to the class of multistage stochastic programs as defined in Section 2. This will justify the dynamic programming technique of telescoping an N-stage stochastic program into a k-stage program.

Theorem 1. *Consider the stochastic program*

P *find* $\inf I_f$ *on* $\mathcal{N}_\infty$,

where $f : \Xi \times \mathbf{R}^n \to]-\infty, +\infty]$ *is a normal convex integrand with* $\xi \mapsto D(\xi) = \operatorname{dom} f(\xi, \cdot)$ *a closed valued, uniformly bounded nonanticipative multifunction. Suppose that* P *is feasible and that there exists a summable function* μ *such that* $|f(\xi, x)| \leq \mu(\xi)$ *for all* $x \in D(\xi)$. *Then for all* $k = 1, 2, \ldots, N$, *the stochastic programs*

Q_k *find* $\inf I_{q_k}$ *on* $\mathcal{N}_\infty^k$

are well-defined. The value function q_k *can be expressed as*

$$q_k(\xi^k, x^k) = \mathbf{E}\left\{\inf_{x^l} q_l(\xi^l, x^l) \,\middle|\, P_k\xi^l = \xi^k, P_k x^l = x^k\right\} \tag{3.16}$$

for $k \leq l \leq N$. *It is a normal convex integrand on* $\Xi^k \times \mathbf{R}^{n^k}$, *with* $\xi^k \mapsto D^k(\xi^k) = \operatorname{dom} q(\xi^k, \cdot)$ *a closed valued, uniformly bounded, measurable multifunction and* I_{q_k} *is a proper, convex, lower semicontinuous (relative to the weak topology* w-$(\mathscr{L}^\infty_{n^k}, \mathscr{L}^1_{n^k})$) *functional on* $\mathscr{L}^\infty_{n^k}$. *Moreover,* P *is solvable and for all* $k = 1, \ldots, N$, *the programs* Q_k *are solvable. Finally, if* $\bar{x}$ *is an optimal solution of* P, *then* $P_k\bar{x}$ *solves* Q_k *and if* $\bar{x}^k$ *solves* Q_k *it can be "extended" to an optimal solution* $\bar{x}$ *of* P *such that* $P_k\bar{x} = \bar{x}^k$.

Proof. Repeated applications of Lemma 1 and Lemma 2 yield the assertions about the value function q_k except for the nonanticipativity of $D_k = \operatorname{dom} q_k$. But this follows simply from the fact that for $l \leq k$, $P_l D_k(\xi) = P_l P_k D(\xi)$ depends only on ξ^l. Summability of the multifunction D_k results from Fubini's Theorem.

The definition of the integral and the summability assumption on f imply that the feasibility set $\mathscr{D}$ is $\{x \in \mathscr{L}^\infty_n \mid x(\xi) \in D(\xi) \text{ a.s.}\}$. But then for $x \in \mathcal{N}_\infty \cap \mathscr{D}$, x^k is nonanticipative and almost surely $P_k x = x^k \in D_k = P_k D$, since otherwise there would be a subset of Ξ^k of nonzero measure for which x^k fails to belong to $P_k D$. By nonanticipativity of D, this in turn, would imply that x fails to belong to $D(\xi)$ almost surely. On the other hand if $x^k \in D_k$ a.s. then again by nonanticipativity of D there must exist a sequence of feasible decisions $(x_{k+1}, \ldots, x_N)$ whatever be the value of ξ, with $P_k\xi = \xi^k$, Section 2.D. This establishes a natural correspondence between feasible solutions to (3.14) and (3.15). The equivalence of the values of the two programs is a direct consequence of this equivalence and Lemma 3.

Finally, Lemma 3 applied recursively yields the assertion about the optimal solutions. The existence of optimal solutions to P and Q_k follows, as in Section 2.E, from the feasibility of P and the fact that I_f and I_{q_k} are inf-compact functionals to be minimized on closed subspaces.

4. Basic duality results

Let $\mathcal{M}_1$ denote the closed linear subspace of $\mathscr{L}^1_n$ consisting of those functions p which satisfy the following *martingale property*: for all $k = 1, \ldots, N$,

$$\mathbf{E}\{(I - P_k)p(\xi) \mid P_k\xi = \xi^k\} = 0. \tag{4.1}$$

We call such a function an $\mathscr{L}^1$*-martingale*. A direct application of the iterated conditional expectation formula shows that $\mathcal{M}_1$ is orthogonal to $\mathcal{N}_\infty$ [2].

Theorem 2 below shows that we can associate with the nonanticipativity restriction Lagrange multipliers which belong to $\mathcal{M}_1$. We proceed by a duality argument. We relate P to a dual program

D find $\sup - I_{f^*}$ on $\mathcal{M}_1$

and show that under certain conditions, there exist $\bar{x}$ and $\bar{p}$ solving P and D respectively, such that $\operatorname{Min} \mathrm{P} = I_f(\bar{x}) = -I_{f^*}(\bar{p}) = \operatorname{Max} \mathrm{D}$. Since

$$-I_{f^*}(p) = \mathbf{E}\{\operatorname{Min}[f(\xi, x) - x \cdot p(\xi) \mid x \in \mathbf{R}^n]\}$$

the result below gives us a function $\bar{p} \in \mathcal{M}_1$ such that $\bar{x}$ is optimal if and only if $\bar{x} \in \mathcal{N}_\infty$ and

$$\bar{x}(\xi) \in \arg\min[f(\xi, x) - x \cdot \bar{p}(\xi) \mid x \in \mathbf{R}^n] \quad \text{a.s.}$$

Thus for example if $f(\xi, \cdot)$ is strictly convex a.s., the unique optimal solution to P is

$$\bar{x}(\xi) = \arg\min[f(\xi, x) - x \cdot \bar{p}(\xi) \mid x \in \mathbf{R}^n],$$

in which case P is reduced by the multiplier function $\bar{p}$ to a *pointwise* minimization without regard for nonanticipativity.

To obtain the existence of Lagrange multipliers we always need a "constraint qualification". Strict feasibility will play that role. It is however noteworthy that by itself this condition does not lead to the existence of "nice" multipliers, but when combined with the nonanticipativity of D it yields the desired result. Without nonanticipativity, the multipliers would be in $(\mathscr{L}^\infty_n)^*$ rather than in $\mathscr{L}^1_n$.

Theorem 2. *Consider the stochastic program*

P *find* $\inf I_f$ *on* $\mathcal{N}_\infty$,

where $f : \Xi \times \mathbf{R}^n \to]-\infty, +\infty]$ *is a normal convex integrand with* $\xi \mapsto D(\xi) = \operatorname{dom} f(\xi, \cdot)$ *a closed valued, uniformly bounded nonanticipative multifunction. Suppose that* P *is strictly feasible and that there exists a summable function* μ *such that* $|f(\xi, x)| \leqslant \mu(\xi)$ *for all* $x \in D(\xi)$. *Then* P *and the associated (dual) program*

D *find* $\sup - I_{f^*}$ *on* $\mathcal{M}_1$

are solvable and

$$\operatorname{Min} \mathrm{P} = \operatorname{Max} \mathrm{D}. \tag{4.2}$$

Proof. Note that for any $x \in \mathcal{N}_\infty$ and $p \in \mathcal{M}_1$ we have that

$$I_f(x) + I_{f^*}(p) \geqslant \langle x, p \rangle = 0, \tag{4.3}$$

and hence it is trivially true that

$$\inf \mathrm{P} \geqslant \sup \mathrm{D}.$$

Moreover, if equality holds in (4.3), a condition which is equivalent by [16, Theorem 2] to the subdifferential relation $p \in \partial I_f(x)$, then x and p are optimal for the two problems and $\min \mathrm{P} = \max \mathrm{D}$.

The hypotheses of the theorem imply that P has an optimal solution, say $\bar{x}$. By way of Theorem 1, the same hypotheses also imply that for $k = 1, \ldots, N$, the value function q_k is well defined and that the stochastic program

Q_k find $\inf I_{q_k}$ on $\mathcal{N}^k_\infty$

admits $\bar{x}^k = P_k \bar{x}$ as an optimal solution.

The proof now proceeds by induction on N. The theorem is certainly true if $N = 1$, since then $\mathcal{N}^1_\infty = \mathscr{L}^\infty_{n_1}$, $\mathcal{M}_1 = \{0\}$ and

$$-I_{f^*}(0) = \min I_f.$$

Now, suppose that the theorem holds for all stochastic programs with $N - 1$ stages. Thus in particular we have that

$$\operatorname{Min} \quad I_{q'} \quad \text{on } \mathcal{N}'_\infty \quad = \quad \operatorname{Max} \quad -I_{(q')^*} \quad \text{on } \mathcal{M}'_1$$

where $'$ replaces $N-1$, i.e., $q' = q_{N-1}$ and in particular $\mathcal{M}'_1$ is the corresponding space of multipliers defined on $(\Xi', \mathscr{F}', \sigma')$ with values in $\mathbf{R}^{n'}$. Hence by induction there exists (Lagrange multiplier) $p' \in \mathcal{M}'_1$ such that the function $P_{N-1}\bar{x} = \bar{x}'$ (which minimizes $I_{q'}$ on $\mathcal{N}'_\infty$) also minimizes

$$I_{q'} - \langle \cdot, p' \rangle \tag{4.4}$$

on $\mathscr{L}^\infty_{n'}$. Let us define

$$l(\xi, x) = f(\xi, x) - \langle P_{N-1}x, p'(\xi) \rangle \tag{4.5}$$

and let $\mathscr{I}$ be the natural injection of $\mathscr{L}^\infty_{n'} = \mathscr{L}^\infty(\Xi', \mathscr{F}', \sigma'; \mathbf{R}^{n'})$ into $\mathscr{L}^\infty(\Xi, \mathscr{F}, \sigma; \mathbf{R}^{n'})$.

Set $\mathscr{W}' = \mathscr{I}(\mathscr{L}^\infty_{n'})$ and $\mathscr{W} = \mathscr{W}' \times \mathscr{L}^\infty(\Xi, \mathscr{F}, \sigma; \mathbf{R}^{n_N})$. Since $\bar{x}$ minimizes I_f on $\mathscr{N}_\infty$ and $\bar{x}' = P_{N-1}\bar{x}$ minimizes $I_{q'} - \langle \cdot, p' \rangle$ on $\mathscr{L}^\infty_{n'}$, $\bar{x}$ must minimize I_l on $\mathscr{W}$ by reverse argument.

Strict feasibility combined with the summability assumption on f imply that I_f, and also I_l, are continuous at a point $\tilde{x}$ in $\mathscr{N}_\infty$ [17, Theorem 2]. Hence by a version of Fenchel's Duality Theorem [27] there exists $w \in (\mathscr{L}^\infty_n)^*$ such that

$$w \in \partial I_l(\bar{x}), \tag{4.6}$$

and

$$w \text{ is orthogonal to } \mathscr{W}. \tag{4.7}$$

Applying Theorem 1 of [17] we see that there is a unique decomposition $w = w_a + w_s$ with $w_a \in \mathscr{L}^1_n$, the "absolutely continuous" component and w_s a "singular" functional on $\mathscr{L}^\infty_n$, such that

$$w_a \in \partial I_l(\bar{x})$$

and

$$w_s \text{ is orthogonal to } \operatorname{dom} I_l = \mathscr{D} \text{ at } \bar{x}.$$

From (4.7) one has

$$w_a = (w'_a, 0) = (P_{N-1}w_a, 0),$$

$$w_s = (w'_s, 0) = (P_{N-1}w_s, 0).$$

With $w'_a \in \mathscr{L}^1_{n'}$, $w'_s \in (\mathscr{L}^\infty_{n'})^*$ and $w' = w'_a + w'_s$ orthogonal to $\mathscr{W}'$, i.e.,

$$-w'_a = w'_s \quad \text{on } \mathscr{W}'. \tag{4.8}$$

The above allows us to conclude that $\mathscr{I}^* w'_s$ is a continuous linear functional on $\mathscr{L}^\infty_{n'}$ for $\mathscr{I}^*$ the adjoint of the injection $\mathscr{I}$. Thus $\mathscr{I}^* w'_s$ corresponds to some $z' \in \mathscr{L}^1_{n'}$ such that for all $x' \in \mathscr{L}^\infty_{n'}$,

$$(\mathscr{I}^* w_s)x' = w'_s(\mathscr{I}x') = \mathbf{E}_\xi\{\langle x'(\xi'), z'(\xi') \rangle\}. \tag{4.9}$$

Since $w_s = (w'_s, 0)$ is orthogonal to $\mathscr{D}$ we must also have that w'_s is orthogonal to $\mathscr{D}' = \operatorname{dom} I_{q'}$. This follows from the nonanticipativity of D (and the summability condition on f) since $\mathscr{D}'$ is then simply $P_{N-1}\mathscr{D}$. Let Q be the projection from $\mathscr{L}_n^\infty$ onto $\mathscr{L}_{n'}^\infty$ defined by

$$(Qx)(\xi') = \mathbf{E}\{P_{N-1}x(\xi) \mid P_{N-1}\xi = \xi'\}. \tag{4.10}$$

Then again by nonanticipativity of D (and the summability condition on f) $Q\mathscr{D} = \mathscr{D}'$ and in particular $Q\bar{x} = \bar{x}'$. With Q^* the adjoint map of Q and using the formula for iterated conditional expectation, we get

$$\begin{aligned}\langle Q^*z', x\rangle = \langle Qx, z'\rangle &= \mathbf{E}_{\xi'}\{\langle \mathbf{E}_\xi\{P_{N-1}x(\xi) \mid \xi'\}, z'(\xi')\rangle\} \\ &= \mathbf{E}_{\xi'}\{\langle x'(\xi'), z'(\xi')\rangle\}\end{aligned} \tag{4.11}$$

which in view of (4.9) yields that

$$Q^*z' = (\mathscr{I}^*z', 0) = (\bar{z}, 0). \tag{4.12}$$

Hence, for $\bar{z}(\xi) = z'(\xi')$ when $P_{N-1}\xi = \xi'$, we have that $(Q^*z') = (\bar{z}, 0)$ is orthogonal to $\mathscr{D}$ at $\bar{x}$. We also have that

$$w_s'(x') = \langle x', \bar{z}\rangle \quad \text{on} \quad \mathscr{W}'. \tag{4.13}$$

This, with the fact that $(w_a', 0) \in \partial I_l(\bar{x})$, implies that

$$(w_a' + \bar{z}, 0) \in \partial I_l(\bar{x}). \tag{4.14}$$

Moreover for $x \in \mathscr{W}$,

$$\langle x, (w_a' + z, 0)\rangle = w_a'(x') + w_s'(x') = w_a'(x') + \langle x', z'\rangle = 0. \tag{4.15}$$

Now let $\bar{v}$ be an element of $\mathscr{L}^1(\Xi, \mathscr{F}, \sigma; \mathbf{R}^n)$ corresponding to w_a' and set

$$\bar{r}(\xi) = \bar{v}(\xi) + \bar{z}(\xi). \tag{4.16}$$

The orthogonality of w' to $\mathscr{W}'$ yields the same property for $\bar{r}$, which is equivalent to asserting that

$$\mathbf{E}\{\bar{r}(\xi) \mid \xi'\} = \mathbf{E}\{\bar{v}(\xi) + \bar{z}(\xi) \mid \xi'\} = 0. \tag{4.17}$$

Thus we must have

$$\mathbf{E}\{\bar{z}(\xi) \mid \xi'\} = z'(\xi') = -\mathbf{E}\{\bar{v}(\xi) \mid \xi'\} \tag{4.18}$$

or equivalently,

$$\bar{r}(\xi) = \bar{v}(\xi) - \mathbf{E}\{\bar{v}(\xi) \mid \xi'\}. \tag{4.19}$$

Now also $(\bar{r}, 0) \in \partial I_l(\bar{x}) = \partial I_f(\bar{x}) - (p', 0)$. Set

$$\bar{p} = (\bar{r} + p', 0). \tag{4.20}$$

Clearly, $\bar{p} \in \mathcal{M}_1$, since $\bar{p} \in \mathscr{L}_n^1$ and

$$E\{\bar{p}(\xi) \mid \xi'\} = p'(\xi') \tag{4.21}$$

for $p' \in \mathcal{M}_1'$. Moreover $\bar{p} \in \partial I_f(\bar{x})$. This completes the proof of the theorem, since the latter relation is equivalent to $I_f(\bar{x}) + I_{f^*}(\bar{p}) = \langle \bar{x}, \bar{p} \rangle = 0$.

Corollary. *Under the same hypotheses as in the theorem, a recourse function $\bar{x}$ is optimal for the multistage stochastic program* P *if and only if there exists a $\bar{p} \in \mathcal{M}_1$ such that the pair $(\bar{x}, \bar{p})$ is a saddle point of the Lagrangian function $\mathscr{L}$ defined on $\mathscr{L}_n^\infty \times (\mathscr{L}_n^\infty)^*$ by*

$$L(x, p) = \begin{cases} I_f(x) - \langle x, p \rangle & \text{if} \quad p \in \mathcal{M}_1, \\ -\infty & \text{if} \quad p \notin \mathcal{M}_1. \end{cases} \tag{4.22}$$

References

[1] R. T. Rochafellar and R. Wets, "The optimal recourse problem in discrete time: $\mathscr{L}_1$-multipliers for inequality constraints", *SIAM Journal on Control and Optimization*, to appear.

[2] R. Wets, "On the relation between stochastic and deterministic optimization" in: A. Bensoussan and J. L. Lions, eds., *Control theory, numerical methods and computer systems modelling*, Springer Verlag Lecture Notes in Economics and Mathematical Systems, 107 (Springer, Berlin 1975) pp. 350–361.

[3] R. T. Rockafellar and R. Wets, "Stochastic convex programming: Kuhn–Tucker conditions", *Journal of Mathematical Economics* 2 (1975) 349–370.

[4] A. D. Yudin, "Duality in multistage stochastic programming", *Izvestija Akademii Nauk SSSR, Tekhnicheskaya Kibernetica* 6 (1973) 24–31 [in Russian; Engl. transl.: also *Engineers Cybernetics* 11 (1973) 908–915].

[5] M. Eisner and P. Olsen, "Duality for stochastic programming interpreted as L. P. in L_p-space", *SIAM Journal of Applied Mathematics* 28 (1975) 779–792.

[6] M. Eisner and P. Olsen, "Duality in probabilistic programming", in: M. Dempster, ed., *Stochastic Programming, Proceedings of the 1974 Oxford international conference* Academic Press, New York, 1976.

[7] E. B. Dynkin, "Stochastic concave dynamic programming", *Matematičeskič Sbornik* 87 (129) (1972) 490–503 [in Russian; Engl. transl.: *Mathematics of the USSR Sbornik* 16 (1972) 501–516.

[8] E. B. Dynkin, "Optimal programs and stimulating prices in stochastic models of economic growth", in: J. Łoś and M. W. Łoś, eds., *Mathematical models in economics*, (PWN-Polish Scientific Publishers, Warszawa, 1974) pp. 207–218.

[9] E. B. Dynkin, "Stochastic models of economic development", in: *American Mathematical Society Translations Series* 2 (1975).

[10] I. Evstigneev, "Optimal economic planning taking into account of stationary random factors", *Doklady Akademii Nauk SSR* 206 (1972) 1040–1942; [in Russian; Engl. transl.: *Soviet Mathematics Doklady* 13 (1972) 1357–1359.

[11] I. Evstigneev, "Optimal Stochastic Programs and their Stimulating Prices", in: J. Łoś and M. W. Łoś, eds., *Mathematical models in economics* PWN-Polish Scientific Publishers, Warszawa 1974) pp. 219–252.

[12] R. Radner, "Existence of equilibrium of plans, prices, and price expectations in a sequence of markets", *Econometrica* 40 (1972) 289–303.

[13] D. Walkup and R. Wets, "Stochastic programs with recourse: special forms", in: H. W. Kuhn, ed., *Proceedings of the Princeton symposium on mathematical programming*, (Princeton University Press, Princeton, N.J., 1970) pp. 139–162.

[14] R. T. Rockafellar and R. Wets, "Stochastic convex programming: basic duality", *Pacific Journal of Mathematics*, 62 (1976).

[15] R. T. Rockafellar and R. Wets, "Continuous versus measurable recourse in N-stage stochastic programming", *Journal of Mathematical Analysis and its Applications* 48 (1974) 836–859.

[16] R. T. Rockafellar, "Integrals which are convex functionals", *Pacific Journal of Mathematics* 24 (1968) 525–540.

[17] R. T. Rockafellar, "Integrals which are convex functionals, II", *Pacific Journal of Mathematics* 39 (1971) 439–469.

[18] R. T. Rockafellar, "Measurable dependence of convex sets and functions on parameters", *Journal of Mathematical Analysis and its Applications* 28 (1969) 4–25.

[19] R. T. Rockafellar, "Convex integral functionals and duality", in: E. Zarantonello, ed., *Contributions to nonlinear functional analysis* (Academic Press, New York 1971) pp. 215–236.

[20] R. Wets, "Induced constraints for stochastic optimization problems", in: A Balakrishnan, ed., *Techniques of optimization* (Academic Press, New York, 1972) pp. 433–443.

[21] R. T. Rockafellar and R. Wets, "Stochastic convex programming: relatively complete recourse and induced feasibility", *SIAM Journal on Control and Optimization* 14 (1976).

[22] P. Olsen, "When is a multistage stochastic programming problem well-defined", *SIAM Journal on Control and Optimization* 14 (1976).

[23] P. Olsen, "Multistage stochastic programming with recourse as mathematical programming in an L_p space", *SIAM Journal on Control and Optimization* 14 (1976).

[24] I. V. Evstigneev, "Measurable selection and dynamic programming", *Mathematics of Operations Research* 1 (1976).

[25] M. Loève, *Probability theory*, 3rd Ed. (Van Nostrand Princeton, N.J., 1963).

[26] R. Wets, "On inf-compact mathematical programs", in: *Conference on optimization techniques*, Part I, Lecture Notes in Computer Science (Springer, Berlin, 1974) pp. 426–436.

[27] R. T. Rockafellar, "Extension of Fenchel's duality theorem for convex functions", *Duke Mathematics Journal* 33 (1966) 81–89.