Journal of Convex Analysis
Volume 9 (2002), No. 1, 237–244

Convex Stochastic Duality and the "Biting Lemma"

Igor V. Evstigneev
*School of Economic Studies, University of Manchester,
Oxford Road, Manchester, M13 9PL, UK*
igor.evstigneev@man.ac.uk

Sjur D. Flåm
*Department of Economics, University of Bergen,
Fosswinckels gate 6, 5007 Bergen, Norway*
sjur.flaam@econ.uib.no

Received April 10, 2000
Revised manuscript received April 26, 2001

A standard approach to duality in stochastic optimization problems with constraints in L_∞ relies upon the Yosida - Hewitt theorem. We develop an alternative technique which employs only "elementary" means. The technique is based on an ϵ-regularization of the original problem and on passing to the limit as $\epsilon \to 0$ with the help of a simple measure-theoretic fact – the biting lemma.

Keywords: stochastic optimization, convex duality, constraints in L_∞, stochastic Lagrange multipliers, bounded sets in L^1, biting lemma, Gale's economic model

1991 Mathematics Subject Classification: 90C15, 52A41, 90C19, 90A16

1. Introduction

The paper analyzes dynamic problems of stochastic optimization in discrete time. The problems under consideration are concerned with maximizing concave functionals on convex sets of feasible strategies (programs). Feasibility is defined in terms of linear inequality constraints in L_∞ holding almost surely. The focus of the work is the existence of dual variables – stochastic Lagrange multipliers in L_1 – relaxing the constraints. Such Lagrange multipliers are important in various applications. In particular, they play key roles in the analysis of stability and sensitivity of solutions to stochastic optimization problems, as well as in the design of algorithms for computing these solutions [14]. Also, such multipliers often have clear interpretations, especially in models related to economics, which sheds additional light on the issues under study [1].

The most common approach to Lagrangian relaxation for the class of optimization problems in question is as follows. First, by using an infinite-dimensional version of the Kuhn - Tucker theorem, one constructs multipliers belonging to the dual, $(L_\infty)^*$, of the space L_∞. Linear functionals in $(L_\infty)^*$ are finitely additive measures. According to the Yosida - Hewitt theorem [23], any functional $\pi \in (L_\infty)^*$ can be decomposed into the sum $\pi = \pi^a + \pi^s$, where π^a is absolutely continuous and π^s is singular. Absolute continuity means that π^a is representable by a function in L_1. The functions in L_1 representing the absolutely continuous components of the functionals constructed turn out to be the desired Lagrange multipliers.

ISSN 0944-6532 / $ 2.50

The idea of the method outlined goes back to Dubovitskii and Milyutin [6] (who dealt originally with deterministic continuous-time problems). Variants of the method applicable to stochastic models were developed by Evstigneev [9], [10], Radner [18], Taksar [22], Rockafellar and Wets [19], [20], Flåm [12], Dempster [5], and others. Analogous techniques for equilibrium (rather than optimization) problems arising in economics were proposed by Bewley [2].

In the present paper, we suggest a different procedure for constructing the Lagrange multipliers. We first consider an appropriate "regularization" of the given problem, depending on a small parameter, $\epsilon > 0$. We construct $(L_\infty)^*$-multipliers for this version of the problem. Its regularity permits to show that the multipliers obtained are absolutely continuous, i.e., representable by functions in L_1. Having established this, we let $\epsilon \to 0$ and prove that the corresponding vector functions in L_1 are norm-bounded. Then we use an elementary measure-theoretic fact – the "biting lemma". By virtue of this lemma, any L_1-bounded sequence of functions contains a subsequence which can be modified on a family of sets with measures tending to zero so that the resulting sequence will be uniformly integrable. Any uniformly integrable set of functions in L_1 has a weak limit point. Such a limit point provides the sought-for vector of Lagrange multipliers.

The above procedure can be applied to a whole range of optimization problems with constraints in L_∞. In this paper, we have chosen – as a convenient vehicle for presenting the method – a stochastic analogue of Gale's [13] economic model. Stochastic versions of Gale's model represent a class of dynamic stochastic optimization problems which is investigated in much detail – see, e.g., Dynkin [7], Radner [18], Arkin and Evstigneev [1]. Our assumptions are milder than those in the literature cited (in particular, they do not imply that the optimal values are attained).

The paper is organized as follows. In Section 2, we describe the optimization problem and state the results. In Section 3, we present the proof of the main theorem.

2. The problem and the result

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq ... \subseteq \mathcal{F}_T$ a sequence of sub-σ-algebras of $\mathcal{F}$. Let d be a natural number. We denote by L_1 (resp. L_∞) the space of integrable (resp. essentially bounded) $\mathcal{F}$-measurable vector functions with values in the Euclidean space R^d. The analogous spaces of $\mathcal{F}_t$-measurable vector functions are denoted by $L_1(t)$ and $L_\infty(t)$. We write $\mathcal{X}_t$ for the standard nonnegative cone in $L_\infty(t)$.

In the model under consideration we are given: non-empty convex sets $Z_t \subseteq \mathcal{X}(t-1) \times \mathcal{X}(t)$, $t \in \{1, ..., T\}$, real-valued concave functionals $F_t(v)$, $v \in Z_t$, and a vector $y_0 \in \mathcal{X}_0$.

The optimization problem is as follows.

(**P**) Maximize

$$F(z) = \sum_{t=1}^{T} F_t(x_{t-1}, y_t)$$

over the set of sequences

$$z = \{(x_{t-1}, y_t)\}_{t=1}^{T} \tag{1}$$

satisfying

$$(x_{t-1}, y_t) \in Z_t,\ t \in \{1, 2, ..., T\}, \tag{2}$$

and

$$y_t \geq x_t,\ t \in \{0, ..., T-1\}. \tag{3}$$

All inequalities between random vectors are supposed to hold coordinatewise and almost surely (a.s.). If a sequence $z = \{(x_{t-1}, y_t)\}_{t=1}^T$ satisfies (2), we call z a *program* and write $z \in Z$. If z satisfies (3) as well, then z is termed a *feasible program*. The optimal value of problem (**P**) (which is not necessarily attained) is denoted by sup(**P**). We assume that sup(**P**)$< \infty$. Solutions to problem (**P**) are called *optimal programs*.

In the economic applications of the above model, elements (x, y) in Z_t are interpreted as *technological processes* (feasible input-output pairs). The sets Z_t are called *technology sets*. Inequalities (3) express *resource constraints*. The objective functionals $F_t(x, y)$ may represent, for example, *expected utilities* $Eu_t(\omega, x(\omega), y(\omega))$.

For a finite-dimensional vector $a = (a^i)$, we write $|a| = \sum |a^i|$. The L_1-norm of a random vector $v(\omega)$ is defined as $||v||_1 = E|v(\omega)|$, where E stands for the expectation with respect to the given probability P. We put $e = (1, ..., 1) \in R^d$. If Γ is a set in Ω, then χ_Γ stands for the indicator function of Γ. We impose the following assumptions.

(A.1) We have $(0, 0) \in Z_t$.

(A.2) If $(x, y) \in Z_t$, $(x', y') \in Z_t$, and $\Gamma \in \mathcal{F}_{t-1}$, then $\chi_\Gamma(x, y)+\ (1 - \chi_\Gamma)(x', y') \in Z_t$.

(A.3) There exists a program $\overset{o}{z}= \{(\overset{o}{x}_{t-1}, \overset{o}{y}_t)\}_{t=1}^T$ such that, for each $t = 0, ...,\ T-1$, we have $\overset{o}{y}_t \geq \overset{o}{x}_t + \delta e$, where $\delta > 0$ is some non-random constant (for $t = 0$, we define $\overset{o}{y}_0 = y_0$).

(A.4) For all $v, v' \in Z_t$,

$$\lim[F_t(v) - F_t(\chi_\Gamma v' + (1 - \chi_\Gamma)v)] = 0$$

as $P(\Gamma) \to 0$ $(\Gamma \in \mathcal{F}_{t-1})$.

Requirements (A.1), (A.2) and (A.4) are supposed to be fulfilled for all $t = 1, ..., T$. Condition (A.1) says that the sequence of zero inputs and outputs forms a feasible program (inactivity is feasible). Hypothesis (A.2) expresses a possibility of choice between two input-output pairs $(x, y) \in Z_t$ and $(x', y') \in Z_t$ depending on an event $\Gamma \in \mathcal{F}_{t-1}$. Assumption (A.3) guarantees the existence of a program which satisfies the resource constraints *with excess*. Condition (A.4) – a continuity property of $F_t(\cdot)$ – holds, for example, if $F_t(z) = Eu_t(\omega, z(\omega))$, where $u_t(\omega, b)$ is a measurable function of $\omega \in \Omega$ and $b \in R^{2d}$ meeting the following requirement: there exists a real-valued random variable $h_t(\omega) \geq 0$ such that $Eh_t(\omega) < \infty$ and $|u_t(\omega, v(\omega))| \leq h_t(\omega)$ for any $v \in Z_t$ and $\omega \in \Omega$.

Denote by $\mathcal{P}_t$ the standard nonnegative cone in $L_1(t)$. The main result, holding under assumptions (A.1) - (A.4), is as follows.

Theorem 2.1. *There exist* $p_0 \in \mathcal{P}_0$, ..., $p_{T-1} \in \mathcal{P}_{T-1}$ *such that*

$$F(z) + \sum_{t=0}^{T-1} Ep_t(y_t - x_t) \leq \sup(\mathbf{P}) \tag{4}$$

for all programs $z = \{(x_{t-1}, y_t)\}_{t=1}^T$.

The random vectors $p_0, ..., p_T$ described in Theorem 1 represent stochastic Lagrange multipliers relaxing constraints (3). As is well-known and easy to prove (see, e.g., [15], Section

8.6), the existence of such vectors is equivalent to the truth of the following *duality theorem*:

$$\sup(\mathbf{P}) = \min(\mathbf{P}^*),$$

where the *dual problem* $(\mathbf{P}^*)$ is defined as follows:

$(\mathbf{P}^*)$ Minimize

$$F^*(p) = \sup_{z \in Z}\{F(z) + \sum_{t=0}^{T-1} Ep_t(y_t - x_t)\}$$

over all $p = (p_0, ..., p_T) \in \mathcal{P}_0 \times ... \times \mathcal{P}_T$.

From (4), it follows immediately that a feasible program $\bar{z} = \{(\bar{x}_{t-1}, \bar{y}_t)\}_{t=1}^T$ is optimal if and only if

$$F(z) + \sum_{t=0}^{T-1} Ep_t(y_t - x_t) \leq F(\bar{z}) \tag{5}$$

for all $z \in Z$. Simple arguments show that (5) holds if and only if

$$F_t(x, y) + Ep_t y - Ep_{t-1} x \leq F_t(\bar{x}_{t-1}, \bar{y}_t) + Ep_t \bar{y}_t - Ep_{t-1}\bar{x}_{t-1} \ [(x, y) \in Z_t,\ t = 1, ..., T],$$

$$p_t(\bar{y}_t - \bar{x}_t) = 0 \ [t = 0, ..., T-1],$$

where $\bar{y}_0 = y_0$ and $p_T = 0$. In economic terms, this means that $p_0, ..., p_T$ are *supporting prices* associated with the resource constraints (see [1], Chapter 4).

In the proof of the above theorem, we will use the following fact.

Lemma 2.2 ("biting lemma"). *Let $\{q_m(\omega)\}$ ($\omega \in \Omega$, $m = 1, 2, ...$) be a sequence of random d-dimensional vectors such that $\{E|q_m|\}$ is bounded. Then there exist measurable sets $\Delta_1 \subseteq \Delta_2 \subseteq ...$ and natural numbers $m_1 < m_2 < ...$ such that $P(\Delta_k) \to 1$ and the sequence $q_{m_k}\chi_{\Delta_k}$ is uniformly integrable.*

Various versions of this lemma have been established by many authors. A number of important results related to it are contained in the papers by Castaing [3] and Saadoune and Valadier [21], where one can find further references. A proof is given, e.g., in [21], p. 349.

Remark 2.3. According to the Dunford – Pettis theorem (see, e.g., [16], Theorem II.23), every uniformly integrable sequence contains a subsequence converging in the weak topology $\sigma(L_1, L_\infty)$. Thus, there exist natural numbers $n_1 < n_2 < ...$, measurable sets $D_1 \subseteq D_2 \subseteq ...$, and a random vector $q \in L_1$ such that $P(D_k) \to 1$ and

$$q_{n_k}\chi_{D_k} \to q \ [\sigma(L_1, L_\infty)]. \tag{6}$$

3. Proof of the main theorem

1st step. Fix a sequence of numbers $1 > \epsilon_1 > \epsilon_2 > ... > 0$ converging to zero. For each $m \in \{1, 2, ...\}$, consider an auxiliary problem $(\mathbf{P}^m)$ which is formulated like $(\mathbf{P})$ with the only difference that Z_t is replaced by

$$Z_t^m = \{(x, y) \in Z_t : y \geq \epsilon_m \overset{o}{y}_t\}.$$

The definitions of programs and feasible programs for problem $(\mathbf{P}^m)$ are analogous to those for $(\mathbf{P})$ (with Z_t^m in place of Z_t). We write $z \in Z^m$ if z is a program for $(\mathbf{P}^m)$. Clearly $\sup(\mathbf{P}^m) \le \sup(\mathbf{P}^{m+1}) \le \sup(\mathbf{P})$. By virtue of (A.1) and (A.3), the sequence $\{\epsilon_m(\overset{o}{x}_{t-1}, \overset{o}{y}_t)\}$ belongs to Z^m, and, for any $\{(x_{t-1}, y_t)\} \in Z^m$, we have

$$y_t \ge \epsilon_m \overset{o}{y}_t,\ t = 0, ..., T. \tag{7}$$

(For $t = 0$, the last inequality holds because $\overset{o}{y}_0 = y_0$ and $\epsilon_m < 1$.) Furthermore,

$$\epsilon_m \overset{o}{y}_t - \epsilon_m \overset{o}{x}_t \ge \epsilon_m \delta e \ (t = 1, 2, ..., T), \text{ and } y_0 - \epsilon_m \overset{o}{x}_0 \ge \epsilon_m \delta e. \tag{8}$$

Viewing inequalities (3) as linear inequality constraints in the spaces $L_\infty(0)$, $L_\infty(1)$,..., $L_\infty(T-1)$, we can use an infinite-dimensional version of the Kuhn-Tucker theorem (see, e.g., Theorem 8.3.1 in [15]) and construct nonnegative linear functionals $\pi_t^m \in (L_\infty(t))^*$ such that

$$\sum_{t=1}^{T} F_t(x_{t-1}, y_t) + \sum_{t=0}^{T-1} < \pi_t^m, y_t - x_t > \le \sup(\mathbf{P}^m) \tag{9}$$

for any program $z = \{(x_{t-1}, y_t)\}_1^T \in Z^m$. The Slater constraint qualification, which allows to apply the above result, holds by virtue of (8).

2nd step. Let us show that the functionals π_t^m involved in (9) are in fact absolutely continuous, i.e., representable in the form $< \pi_t^m, x >= E p_t^m x$, $p_t^m \in \mathcal{P}_t$. Fix any $t \in \{0, 1, ..., T-1\}$ and any natural number m. Furthermore, fix any real number $\kappa > 0$ and consider some feasible program $z = \{(x_{t-1}, y_t)\}$ of problem $(\mathbf{P}^m)$ for which

$$F(z) \ge \sup(\mathbf{P}^m) - \kappa. \tag{10}$$

For any $\Gamma \in \mathcal{F}_t$, define a sequence $z^\Gamma = \{(x_{j-1}^\Gamma, y_j^\Gamma)\}$ by

$$(x_{j-1}^\Gamma, y_j^\Gamma) = \begin{cases} (x_{j-1}, y_j), & j \le t; \\ (1 - \chi_\Gamma)(x_{j-1}, y_j) + \chi_\Gamma \epsilon_m(\overset{o}{x}_{j-1}, \overset{o}{y}_j), & j > t. \end{cases}$$

Observe that z^Γ is a feasible program for $(\mathbf{P}^m)$. Indeed, if $j > t$, then $\Gamma \in \mathcal{F}_{j-1}$, and so

$$(x_{j-1}^\Gamma, y_j^\Gamma) = (1 - \chi_\Gamma)(x_{j-1}, y_j) + \chi_\Gamma \epsilon_m(\overset{o}{x}_{j-1}, \overset{o}{y}_j) \in Z_j$$

by virtue of (A.2). Further, we have $y_0 \ge x_0^\Gamma$ because $x_0^\Gamma = x_0$ if $t > 0$ and

$$x_0^\Gamma = (1 - \chi_\Gamma)x_0 + \chi_\Gamma \epsilon_m \overset{o}{x}_0 \le (1 - \chi_\Gamma)y_0 + \chi_\Gamma \epsilon_m \overset{o}{y}_0 \le y_0$$

if $t = 0$. If $T > j > 0$ and $j \ne t$, then the inequality $y_j^\Gamma \ge x_j^\Gamma$ holds because both $\{(x_{j-1}, y_j)\}$ and $\{(\overset{o}{x}_{j-1}, \overset{o}{y}_j)\}$ are feasible programs of problem $(\mathbf{P}^m)$. If $j = t$, then, by virtue of (7) and (8), we have

$$y_j^\Gamma - x_j^\Gamma \ge y_t - [(1 - \chi_\Gamma)x_t + \chi_\Gamma \epsilon_m \overset{o}{x}_t] \ge$$

$$(1 - \chi_\Gamma)(y_t - x_t) + \chi_\Gamma(y_t - \epsilon_m \overset{o}{x}_t) \ge \chi_\Gamma \epsilon_m(\overset{o}{y}_t - \overset{o}{x}_t) \ge \chi_\Gamma \epsilon_m \delta e \ge 0. \tag{11}$$

By applying (9) and (11) to z^Γ, we find

$$\epsilon_m \delta < \pi_t^m, \chi_\Gamma e > \leq \sum_{t=0}^{T-1} < \pi_t^m, y_t^\Gamma - x_t^\Gamma > \leq \sup(\mathbf{P}^m) - \sum_{t=1}^{T} F_t(x_{t-1}^\Gamma, y_t^\Gamma) \leq$$

$$\sum_{t=1}^{T} F_t(x_{t-1}, y_t) - \sum_{t=1}^{T} F_t(x_{t-1}^\Gamma, y_t^\Gamma) + \kappa,$$

where $y_0^\Gamma = \overset{o}{y}_0 = y_0$. Since $\kappa > 0$ is arbitrary, we obtain

$$\epsilon_m \delta < \pi_t^m, \chi_\Gamma e > \leq \sum_{t=1}^{T} F_t(x_{t-1}, y_t) - \sum_{t=1}^{T} F_t(x_{t-1}^\Gamma, y_t^\Gamma)\,.$$

By virtue of (A.4), the last expression tends to zero as $P(\Gamma) \to 0$. Thus $< \pi_t^m, \chi_\Gamma e > \to 0$ as $P(\Gamma) \to 0$. This implies that the functional π_t^m (≥ 0) is absolutely continuous, i.e., representable in the form $< \pi_t^m, x >= E p_t^m x$, $x \in L_\infty(t)$, where $p_t^m \in \mathcal{P}_t$ (see, e.g., [17], Corollary to Proposition IV.2.1).

By substituting $\{(\overset{o}{x}_{t-1}, \overset{o}{y}_t)\}$ into (9), we get

$$\delta E|p_t^m| = E p_t^m \delta e \leq E p_t^m (\overset{o}{y}_t - \overset{o}{x}_t) \leq \sup(\mathbf{P}) - \sum_{j=1}^{T} F_j(\overset{o}{x}_{j-1}, \overset{o}{y}_j).$$

This shows that the random vectors p_t^m are uniformly bounded in the L_1 norm.

3rd step. By using Lemma 2.2, we find sets $\Delta_t^m \in \mathcal{F}_t$ such that $P(\Delta_t^m) \to 1$, $\Delta_t^1 \subseteq \Delta_t^2 \subseteq ...$, and the sequence $\{\chi_{\Delta_t^m} p_t^m\}$ is uniformly integrable. Without loss of generality, we may assume in addition that $\Delta_0^m \supseteq \Delta_1^m \supseteq ... \supseteq \Delta_{T-1}^m$ (the set Δ_t^m can be replaced by $\Delta_0^m \cap \Delta_1^m \cap ... \cap \Delta_t^m$). Since the sequence $\{\chi_{\Delta_t^m} p_t^m\}$ is uniformly integrable, it contains a subsequence which converges in the topology $\sigma(L_1(t), L_\infty(t))$. This is true for each $t = 0, ..., \ T-1$. Thus, by passing to a subsequence, we may suppose, again without loss of generality, that, for each t, $\chi_{\Delta_t^m} p_t^m \to p_t$ $[\sigma(L_1(t), L_\infty(t))]$, where p_t is a nonnegative random vector in $L_1(t)$.

Let $z = \{(x_{t-1}, y_t)\}$ be any program of problem $(\mathbf{P})$ satisfying $y_t \geq \epsilon\, \overset{o}{y}_t$, where ϵ is some number in $(0,1)$. Then z is a program of problem $(\mathbf{P}^m)$ for all m for which $\epsilon_m < \epsilon$ (and so for all m large enough). Fix any such m and put $\Delta_t = \Delta_t^m$. Define

$$(\tilde{x}_{t-1}, \tilde{y}_t) = (1 - \chi_{\Delta_{t-1}})\epsilon(\overset{o}{x}_{t-1}, \overset{o}{y}_t) + \chi_{\Delta_{t-1}}(x_{t-1}, y_t),\ t = 1, ..., T.$$

By virtue of (A.2), $(\tilde{x}_{t-1}, \tilde{y}_t) \in Z_t$. Furthermore, $\tilde{y}_t \geq \epsilon\, \overset{o}{y}_t$, where $\epsilon > \epsilon_m$, and so $\tilde{z} \equiv \{(\tilde{x}_{t-1}, \tilde{y}_t)\}$ is a program of $(\mathbf{P}^m)$. Consequently,

$$\sum_{t=1}^{T} F_t(\tilde{x}_{t-1}, \tilde{y}_t) + \sum_{t=0}^{T-1} E p_t^m (\tilde{y}_t - \tilde{x}_t) \leq \sup(\mathbf{P}^m) \leq \sup(\mathbf{P}), \tag{12}$$

where $\tilde{y}_0 = y_0$.

Observe next that $(\tilde{y}_t - \tilde{x}_t) \geq \chi_{\Delta_t}(y_t - x_t)$. Indeed,

$$(\tilde{y}_t - \tilde{x}_t) - \chi_{\Delta_t}(y_t - x_t) \geq (1 - \chi_{\Delta_{t-1}})\epsilon\, \overset{o}{y}_t + \chi_{\Delta_{t-1} \setminus \Delta_t} y_t - (1 - \chi_{\Delta_t})\epsilon\, \overset{o}{x}_t \geq$$

$$(1-\chi_{\Delta_{t-1}})\epsilon \overset{o}{y}_t + \chi_{\Delta_{t-1}\setminus\Delta_t}\epsilon \overset{o}{y}_t - (1-\chi_{\Delta_t})\epsilon \overset{o}{x}_t =$$

$$(1-\chi_{\Delta_t})\,\epsilon(\overset{o}{y}_t - \overset{o}{x}_t) \geq 0,$$

because $\Delta_t \subseteq \Delta_{t-1}$ and $y_t \geq \epsilon \overset{o}{y}_t$. Consequently, from this and (12), we find

$$\sum_{t=1}^{T} F_t(\tilde{x}_{t-1}, \tilde{y}_t) + \sum_{t=0}^{T-1} Ep_t^m \chi_{\Delta_t^m}(y_t - x_t) \leq \sup(\mathbf{P}).$$

By passing to the limit (using (A.4)), we conclude

$$\sum_{t=1}^{T} F_t(x_{t-1}, y_t) + \sum_{t=0}^{T-1} Ep_t(y_t - x_t) \leq \sup(\mathbf{P}). \tag{13}$$

The above inequality was obtained under the assumption that $\{(x_{t-1}, y_t)\}$ is a program of $(\mathbf{P})$ satisfying $y_t \geq \epsilon \overset{o}{y}_t$ for some $\epsilon > 0$. Now consider any program $\{(x_{t-1}, y_t)\}$ of $(\mathbf{P})$. For each $\epsilon \in [0,1]$, set $(x_{t-1}^{(\epsilon)}, y_t^{(\epsilon)}) = \epsilon(\overset{o}{x}_{t-1}, \overset{o}{y}_t) + (1-\epsilon)(x_{t-1}, y_t)$. By using the result just obtained, we find

$$\sum_{t=1}^{T} F_t(x_{t-1}^{(\epsilon)}, y_t^{(\epsilon)}) + \sum_{t=0}^{T-1} Ep_t(y_t^{(\epsilon)} - x_t^{(\epsilon)}) \leq \sup(\mathbf{P}) \tag{14}$$

for all $\epsilon \in (0,1]$. The left-hand side of (14) is a concave function of $\epsilon \in [0,1]$. Hence, if inequality (14) holds for each $\epsilon \in (0,1]$, it holds for $\epsilon = 0$ as well. This proves (13) for any program of $(\mathbf{P})$.

The proof is complete.

Remark 3.1. We conjecture that the method described can be extended to stochastic models of economic equilibrium [11] generalizing the optimization model considered in this paper. In that, more general, setting, the counterpart of the problem of constructing stochastic Lagrange multipliers is the problem of constructing equilibrium states (solutions to certain variational inequalities). The conventional approach employing the Yosida–Hewitt theorem seems to be inapplicable in that context. It would be of interest to find out whether the technique based on the biting lemma could replace the conventional one in the framework of equilibrium models.

Acknowledgements. The authors thank an anonymous referee for helpful comments, which were taken into account in the final version of the paper. Financial support from the University of Bergen, NFR project Quantec, and from the grants INTAS 97-1050 and RFBR 00-01-00297 is gratefully acknowledged.

References

[1] V. I. Arkin, I. V. Evstigneev: Stochastic Models of Control and Economic Dynamics, Academic Press, London (1987).

[2] T. Bewley: Existence of equilibria in economies with infinitely many commodities, Journal of Economic Theory 4 (1972) 514–540.

[3] C. Castaing: Méthode de compacité et de décomposition applications: minimisation, convergence des martingales, lemme de Fatou multivoque, Ann. Mat. Pura Appl. 164 (1993) 51–75.

[4] N. Dunford, J. T. Schwartz: Linear operators, Part 1, General theory, New York, Interscience Publ. (1957).

[5] M. A. H. Dempster: On stochastic programming: II. Dynamic problems under risk, Stochastics 25 (1988) 15–42.

[6] A. Ya. Dubovitskii, A. A. Milyutin: Necessary conditions for a weak extremum in optimal control problems with mixed inequality constraints, USSR Computational Mathematics and Mathematical Physics 8 (1968) 24–98.

[7] E. B. Dynkin: Some probability models for a developing economy, Soviet Mathematics Doklady 12 (1971) 1422–1425.

[8] E. B. Dynkin: Stochastic concave dynamic programming, Math. USSR Sbornik 16 (1972) 501–515.

[9] I. V. Evstigneev: Optimal economic planning taking account of stationary random factors, Soviet Math. Doklady 13 (1972) 1357–1359.

[10] I. V. Evstigneev: Optimal stochastic programs and their stimulating prices, in: M.W. Loś and J. Loś (eds.), Mathematical Models in Economics, North Holland, Amsterdam (1974) 219–252.

[11] I. V. Evstigneev, M. I. Taksar: Stochastic equilibria on graphs II, Journal of Math. Economics 24 (1995) 383–406.

[12] S. D. Flåm: Nonanticipativity in stochastic programming, Journal of Optimization Theory and Applications 46 (1985) 23–30.

[13] D. Gale: On optimal development of a multisector economy, Review of Economic Studies 34 (1967) 1–18.

[14] P. Kall, S. Wallace: Stochastic Programming, Wiley, Chichester (1994).

[15] D. G. Luenberger: Optimization by Vector Space Methods, Wiley, New York (1969).

[16] Meyer P. A.: Probability and Potentials, Blaisdell, Waltham et al. (1966).

[17] J. Neveu: Mathematical Foundations of the Calculus of Probability Theory, Holden Day, San Francisco (1965).

[18] R. Radner: Optimal stationary consumption with stochastic production and resources, Journal of Economic Theory 6 (1973) 68–90.

[19] R. T. Rockafellar, R. J.-B. Wets: Stochastic convex programming: Kuhn–Tucker conditions, Journal of Mathematical Economics 2 (1975) 349–370.

[20] R. T. Rockafellar, R. J.-B. Wets: Nonanticipativity and L^1-martingales in stochastic optimization problems, in: R. J.-B. Wets (ed.), Stochastic Systems: Modeling, Identification and Optimization II, Math. Programming Study 6, North-Holland, Amsterdam (1976) 170–187.

[21] M. Saadoune, M. Valadier: Extraction of a "good" subsequence from a bounded sequence of integrable functions, J. Convex Analysis 2 (1995) 345–357.

[22] M. I. Taksar: Optimal planning over infinite time interval under random factors, in: M. W. Loś and J. Loś (eds.), Mathematical Models in Economics, North Holland, Amsterdam (1974) 289–298.

[23] K. Yosida, E. Hewitt: Finitely additive measures, Trans. Amer. Math. Soc. 72 (1952) 46–66.