MATHEMATICAL CONTROL THEORY
BANACH CENTER PUBLICATIONS, VOLUME 14
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

AN APPROXIMATION METHOD IN STOCHASTIC OPTIMIZATION AND CONTROL

WERNER RÖMISCH

Sektion Mathematik, Humboldt-Universität zu Berlin, Berlin, DDR

1. Introduction

This paper investigates a constructive possibility for the approximation of stochastic optimization problems, including optimal control problems. More precisely, our basic problem is of the following type:

$$(1) \qquad J(u) := E[g(\omega, u)] \to \text{Min!} \quad \text{subject to } u \in C.$$

where C is a non-empty constraint set, $g\colon \Omega \times C \to \bar{R}$ is such that for all $u \in C$, $g(\cdot, u)$ represents a real random variable defined on a probability space $(\Omega, \mathfrak{A}, P)$ and E denotes the mean value. Chapter 2 discusses this type of decision problems in abstract spaces, where we often assume that C is a set with convergence. As examples we refer to two-stage problems of stochastic programming in a general setting (see [14], [16]) and to the optimal control of random operator equations in Banach spaces. In the latter case this paper extends works [22] and [26].

Chapter 3 shortly surveys approximation results for optimization problems with fixed constraint set, especially results from nonlinear parametric programming (see [1], [17], [18], [20], [21], [6]). These results are applied to the above class of stochastic optimization problems. The main problem in approximate solving (1) consists in the fact that usual optimization algorithms require the evaluation of $J(u)$ and (or) its derivative, which is impossible due to the difficulty of computing the mean value (see e.g. [13], [14]). These difficulties do not arise in so-called "discrete approximations" of (1), which contain only random variables with a discrete probability distribution (see [13], [14], [15], [23]). This is the reason why "discretization schemes" are proposed in Chapter 4. An important special case of such discretization schemes seems to be the approximation of the random variables entering problem (1) by suitable conditional expectations (see [13], [14], [26]). Some advantages of using conditional expectations are reported in Remark 6.

In Chapter 5 we apply the results to optimal control problems for random operator equations. Especially the case of a discretization scheme is discussed. We finish with a remark on the application to the optimal control of random differential equations (see also [26]).

At the beginning we explain some notations. N denotes the set of natural numbers, R the set of real numbers and $\bar{R} := R \cup \{-\infty\} \cup \{+\infty\}$. Usually $(\Omega, \mathfrak{A}, P)$ is a probability space. If X is a Banach space, we define

$$\mathscr{P}(X) := \{D \mid D \subseteq X, D \neq \emptyset\}$$

and by $\mathfrak{B}(X)$ the σ-algebra of Borel sets of X. $\mathfrak{A} \otimes \mathfrak{B}(X)$ denotes the smallest σ-algebra on $\Omega \times X$ containing $\{A \times B \mid A \in \mathfrak{A}, B \in \mathfrak{B}(X)\}$. $L_p(\Omega, \mathfrak{A}, P; X)$ $(1 \leqslant p < \infty)$ with the norm $\|\cdot\|_{L_p}$ is the usual Banach space of random variables, $L_p(\Omega, \mathfrak{A}, P) := L_p(\Omega, \mathfrak{A}, P; R)$. By a.s. we mean the notion "almost surely". C: $\Omega \to \mathscr{P}(X)$ will be called a *multifunction* (see [4], [33]) and the graph of C is denoted by

$$\operatorname{Gr} C := \{(\omega, x) \in \Omega \times X \mid x \in C(\omega)\}.$$

Finally we remark that we are often concerned with the situation of a non-empty set with convergence ϱ and that we use the notions "ϱ-continuous, ϱ-compact, ϱ-lim" (see [19], p. 90).

2. A class of decision problems in stochastic optimization

Throughout this paper we consider decision problems of stochastic optimization in abstract spaces (comp. [22], [32]). It is known that several problems of stochastic programming and stochastic control fit in this framework (see e.g. the examples below). We are given a probability space $(\Omega, \mathfrak{A}, P)$, a non-empty set C and a mapping g: $\Omega \times C \to \bar{R}$ such that for all $u \in C$, $g(\cdot, u)$ is a real random variable defined on $(\Omega, \mathfrak{A}, P)$. Let E denote the mean value on $(\Omega, \mathfrak{A}, P)$.

Then we consider the following stochastic optimization problem:

$$J(u) := E[g(\omega, u)] \to \text{Min!} \qquad \text{s.t. } u \in C$$

(s.t. for "subject to"). We note that under the usual convention ([25], p. 184, [11], p. IV.3) J has a well-defined value in $\bar{R}$.

Remark 1. (a) We will often assume that for all $u \in C$, $g(\cdot, u) \in L_1(\Omega, \mathfrak{A}, P)$. Then we have J: $C \to R$.

(b) If C is a set with convergence ϱ, then the following conditions are obviously sufficient for the ϱ-continuity of J:

(i) $g(\omega, \cdot)$: $C \to \bar{R}$ is ϱ-continuous a.s.;

(ii) there exists a random variable $c \in L_1(\Omega, \mathfrak{A}, P)$ such that for all $u \in C$, $|g(\omega, u)| \leqslant c(\omega)$ a.s.

(c) Similar conditions can be formulated for the differentiability of J, too. We refer to [2], [14], Theorem 3, [23], pp. 15, 16. If C is a convex subset of a linear space and $g(\omega, \cdot)$ is convex a.s., then J is a convex functional. For results concerning the lower semicontinuity of J we refer to [34], [11], p. IV.15. These results carry over to the case of a ϱ-lower semicontinuous functional J and thus an existence result for (1) on a ϱ-compact set C can be formulated.

(d) [31] contains a general approach to stochastic programming, which discusses (1) in the framework of a general notion of solution.

Examples. (a) Two-stage problems of stochastic programming ([14], [16]):

$$J(u) := E[g(\omega, u)] \to \text{Min!} \quad \text{s.t. } u \in C,$$

where

$$\begin{gathered} g(\omega, u) := \inf\{f(\omega, x, u) \mid x \in C_x(\omega, u)\}, \quad \omega \in \Omega, u \in C, \\ f\colon \Omega \times X \times C \to \bar{R}, \quad C_x\colon \Omega \times C \to \mathscr{P}(X), \\ X \text{ is a real separable Banach space.} \end{gathered}$$

If we assume that for all $u \in C$, $f(\cdot, \cdot, u)$: $\Omega \times X \to \bar{R}$ is a $\mathfrak{A} \otimes \mathfrak{B}(X)$-measurable function and $\operatorname{Gr} C_x(\cdot, u) \in \mathfrak{A} \otimes \mathfrak{B}(X)$, and that $(\Omega, \mathfrak{A}, P)$ is complete, it results from [4], Lemma III.39, that for all $u \in C$, $g(\cdot, u)$: $\Omega \to \bar{R}$ is measurable and the two-stage problem is well-defined.

We refer to [12], [13], [14], [15], [23], Chapter 10, for the structure and discussion of "discretization schemes" (see Def. 1) for such problems.

(b) Optimal control problems with random operator equations:

$$J(u) := E\big[g\big(\omega, x(\omega), u\big)\big] \to \text{Min!} \quad \text{s.t. } u \in C,$$

where

$$\begin{gathered} T\big(\omega, x(\omega), u\big) = 0, \quad x(\omega) \in C_x(\omega, u) \text{ a.s.}, \\ C_x(\cdot, u)\colon \Omega \to \mathscr{P}(X), \\ T(\cdot, \cdot, u)\colon \operatorname{Gr} C_x(\cdot, u) \to Y, \quad u \in C, \\ g(\cdot, \cdot, u)\colon \operatorname{Gr} C_x(\cdot, u) \to \bar{R}, \\ X, Y \text{ are real separable Banach spaces.} \end{gathered}$$

If we assume that for all $u \in C$ $g(\cdot, \cdot, u)$ and $T(\cdot, \cdot, u)$ are $\mathfrak{A} \otimes \mathfrak{B}(X)$-measurable, and additionally

$$S(\omega, u) := \{x \in C_x(\omega, u) |\ T(\omega, x, u) = 0\} \neq \emptyset \quad \text{a.s.},$$

then by [24], Theorem 1 the random operator equation

$$T(\omega, x, u) = 0$$

has a random solution $x\colon \Omega \to X$ for all $u \in C$ ([3], [8], [27], Chapt. 2), the mapping $g(\cdot, x(\cdot), u)\colon \Omega \to \bar{R}$ is measurable and the optimal control problem is well-defined.

3. Remarks on the approximation of optimization problems with fixed constraint set

In the following we collect some well-known results on the approximation (or "stability", "perturbation") of optimization problems with fixed constraint set.

We are given a non-empty set C, functionals J, $J_m\colon C \to \bar{R}$, $m \in N$, and we consider the problems:

$$(2) \qquad J(u) \to \text{Min!} \quad \text{s.t. } u \in C,$$

$$(2m) \qquad J_m(u) \to \text{Min!} \quad \text{s.t. } u \in C.$$

We define the optimal values

$$\varphi := \inf\{J(u) |\ u \in C\}, \qquad \varphi_m := \inf\{J_m(u) |\ u \in C\}, \qquad m \in N,$$

and the optimal set mappings

$$\psi := \{u \in C |\ J(u) = \varphi\}, \qquad \psi_m := \{u \in C |\ J_m(u) = \varphi_m\}, \qquad m \in N,$$

and ask for the convergence of the sequences $\{\varphi_m\}_{n \in N}$ to φ resp. $\{\psi_m\}_{m \in N}$ to ψ (in some sense).

First we give a simple lemma that results from [17], Lemmas 1.1 and 1.3.

LEMMA 1. *Let for every sequence* $\{u_m\}_{m \in N} \subset C$,

$$\lim_{m \to \infty} \big(J_m(u_m) - J(u_m)\big) = 0.$$

Then $\lim\limits_{m \to \infty} \varphi_m = \varphi$.

Remark 2. The assumption of Lemma 1 is equivalent to the uniform convergence of $\{J_m\}_{m \in N}$ to J (comp. [22], Lemma 3.1). Only by further assumptions on C and J, resp., it becomes possible to weaken the condition of Lemma 1 concerning the convergence of $\{J_m\}_{m \in N}$ to J (see Theorem 1).

Theorem 1. *Let C be a set with convergence ϱ and assume that*

(a) *C is ϱ-compact and J is ϱ-continuous;*

(b) *for all u, $u_m \in C$, $m \in N$, such that $\varrho\text{-}\lim_{m\to\infty} u_m = u$,*

$$\lim_{m\to\infty} J_m(u_m) = J(u).$$

Then we have

(i) $\lim_{m\to\infty} \varphi_m = \varphi$;

(ii) *for every sequence $u_m^* \in \psi_m$, $m \in N$, there exists an accumulation point (with respect to ϱ) $u^* \in \psi$.*

Proof. (i): We use Lemma 1 and choose a sequence $\{u_m\}_{m\in N}$ in C such that the condition of Lemma 1 is not fulfilled. By the assumptions of the Theorem this leads to a contradiction.

(ii): This assertion follows from the ϱ-compactness of C, from (b) and (i). ■

Remark 3. (a) Theorem 1 results e.g. from [1], Satz 4.2.2, [20], Satz 2.1 and 2.3. But in [1], [20], [21] the statements were proved for metric spaces. These proofs can easily be generalized to the present case of sets with convergence (see also [19], p. 154 ff.). For optimal control problems it seems to be advantageous to apply Theorem 1 in the case of weak convergence in Banach spaces (comp. [6], Chapter 2, [18], Satz 1.2, Chapter 5).

(b) The ϱ-compactness of C cannot be dropped in Theorem 1 (comp. various examples in [1]).

(c) Condition (b) in Theorem 1 represents something like a "discrete ϱ-convergence of $\{J_m\}_{m\in N}$ to J" ([30]). Therefore this condition can be replaced by:

(b′) For all $u \in C$, $\lim_{m\to\infty} J_m(u) = J(u)$; for all $u_m, v_m \in C$, $m \in N$, such that $\varrho\text{-}\lim_{m\to\infty} u_m = \varrho\text{-}\lim_{m\to\infty} v_m$, we have

$$\lim_{m\to\infty} |J_m(u_m) - J_m(v_m)| = 0.$$

("ϱ-consistency" and "ϱ-stability", [30], p. 290.)

(d) If we choose $u_m^* \in C$ such that $J_m(u_m^*) \leqslant \varphi_m + \varepsilon_m$, $m \in N$, where $\{\varepsilon_m\}_{m\in N}$ is a positive null sequence, the assertion of Theorem 1 remains valid.

4. On the approximation of stochastic optimization problems, discretization schemes and conditional expectations

As in Chapter 2 let us consider problem (1) and let additionally mappings g_m: $\Omega \times C \to \bar{R}$, $m \in N$, be given such that for all $u \in C$, $g_m(\cdot, u)$, $m \in N$, are real random variables. Then we define J_m: $C \to \bar{R}$, $J_m(u) := E[g_m(\omega, u)]$, $u \in C$, $m \in N$, and consider the problems

$$(1m) \qquad J_m(u) \to \text{Min!} \quad \text{s.t. } u \in C,\ m \in N.$$

Now our aim is to apply the results of Chapter 3 to obtain results for the approximation of (1) by problems $(1m)$, $m \in N$. Let φ, ψ and φ_m, ψ_m, $m \in N$, resp., be defined as above and let C be a set with convergence ϱ. Then the following Theorem is an obvious consequence of Theorem 1 and Remark 3(c).

THEOREM 2. *Let condition* (a) *of Theorem* 1 *and the following assumptions be fulfilled*:

For all $u \in C$, $\lim_{m\to\infty} \|g_m(\cdot, u) - g(\cdot, u)\|_{L_1} = 0$; *for all* $u_m, v_m \in C$, $m \in N$, *such that* $\varrho\text{-}\lim_{m\to\infty} u_m = \varrho\text{-}\lim_{m\to\infty} v_m$, *we have*

$$\lim_{m\to\infty} \|g_m(\cdot, u_m) - g_m(\cdot, v_m)\|_{L_1} = 0.$$

Then the assertion of Theorem 1 *is valid.*

The main concern when approximating (1) is that the problems $(1m)$, $m \in N$, are in some sense simpler to solve. It seems an essential possibility to suggest the following notion of a "discretization scheme" for (1).

DEFINITION 1. $\{\mathfrak{A}_m, g_m\}_{m\in N}$ is called a *discretization scheme* (for (1)) if for all $m \in N$ there exist

(i) a finite partition $A_{ml} \in \mathfrak{A}$, $l = 1, \ldots, m$, of Ω, i.e.,

$$\bigcup_{l=1}^{m} A_{ml} = \Omega \quad \text{and} \quad A_{ml} \cap A_{mk} = \emptyset, \quad l \neq k,$$

such that

$$\mathfrak{A}_m := \sigma(\{A_{ml}\}_{l=1,\ldots,m})$$

(here $\sigma(\mathfrak{E})$ denotes the smallest σ-algebra containing $\mathfrak{E} \subset \mathfrak{A}$);

(ii) g_{ml}: $C \to \bar{R}$, $l = 1, \ldots, m$, such that

$$g_m(\omega, \cdot) = g_{ml}, \quad \omega \in A_{ml}, \quad l = 1, \ldots, m.$$

Now let a discretization scheme $\{\mathfrak{A}_m, g_m\}_{m\in N}$ for (1) be given and we consider problems (1), (1m), $m \in N$.

Remark 4. (a) An advantage of discretization schemes seems to be that the functionals J_m, $m \in N$, have a special "deterministic" form:

$$J_m(u) = \sum_{l=1}^{m} g_{ml}(u)P(A_{ml}), \qquad u \in C.$$

Of course the probabilities $P(A_{ml})$, $l = 1, \ldots, m$, must be known. Nice properties of g_{ml}, $l = 1, \ldots, m$, like continuity, differentiability and convexity, yield the corresponding properties of J_m. Therefore the question of a suitable construction of discretization schemes arises (see Remark 5).

(b) Discretization schemes represent a well-known method in stochastic programming problems with recourse (see [12], [13], [14], [15], [23], Chapt. 10) and in more general decision problems of stochastic optimization ([22], Chapt. 6). The presented concept generalizes the known approaches in some sense.

In the following we refer to a simple, but essential possibility to construct discretization schemes.

Suppose we are given a real separable Banach space Z, a random variable

$$z\colon \Omega \to Z$$

with range

$$R(z) := \{z(\omega) |\ \omega \in \Omega\} \subseteq \tilde{Z} \subseteq Z,$$

and a function

$$g\colon \tilde{Z} \times C \to \bar{R}$$

such that for all $u \in C$

$$g(z(\cdot), u)\colon \Omega \to \bar{R}$$

is a real random variable. Then the problem

$$(1)' \qquad J(u) := E[g(z(\omega), u)] \to \text{Min!} \quad \text{s.t. } u \in C$$

is well-defined.

Remark 5. (a) If z is replaced in (1)$'$ by simple random variables $z_m\colon \Omega \to Z$ with m values and $R(z_m) \subseteq \tilde{Z}$, i.e., $z_m(\omega) = z_{ml} \in \tilde{Z}$, $\omega \in A_{ml}$, $l = 1, \ldots, m$, where $\{A_{ml}\}_{l=1,\ldots,m}$ is a partition of Ω, then discretization schemes result in a natural way. Especially one has

$$J_m(u) := E[g(z_m(\omega), u)] = \sum_{l=1}^{m} g(z_{ml}, u)P(A_{ml}), \qquad u \in C, m \in N.$$

Therefore it is suggested to replace the probability distribution of the data entering a stochastic optimization problem (1) by finite discrete distributions. By using Lemma 1 or Theorem 2 approximation results for (1)' are available.

(b) It seems to be a suitable possibility to replace z by conditional expectations with respect to certain finitely generated σ-algebras (see Remark 6). This way was suggested in the context of stochastic programming with recourse by Kall [13], [14] and in [28], [26], [27] for various problems.

Remark 6. As mentioned above, conditional expectations of z seem to be advantageous for the construction of discretization schemes (for (1)'). More precisely, if $z \in L_1(\Omega, \mathfrak{A}, P; Z)$ and if a sequence $\{\{A_{ml}\}_{l=1,\dots,m}\}_{m\in N} \subset \mathfrak{A}$ of partitions of Ω is given, we suggest to choose

$$z_m := E(z|\mathfrak{A}_m), \quad \mathfrak{A}_m := \sigma(\{A_{ml}\}_{l=1,\dots,m}), \quad m \in N,$$

$$z_m(\omega) := E(z|A_{ml}) := \frac{1}{P(A_{ml})} \int_{A_{ml}} z(\omega)\,dP, \quad \omega \in A_{ml},$$

$$l = 1, \dots, m.$$

In the following we state some reasons for this choice:

(i) In case that $\mathfrak{A}_m \subseteq \mathfrak{A}_{m+1}$, $m \in N$, and z is measurable with respect to $\sigma\big(\bigcup_{m\in N} \mathfrak{A}_m\big)$, we know from the well-known martingale convergence theorems (see [5], Theorems 1 and 4) that

$$\lim_{m\to\infty} \|z_m(\omega) - z(\omega)\|_Z = 0 \text{ a.s.},$$

$$\lim_{m\to\infty} \|z_m - z\|_{L_p} = 0 \quad \text{if} \quad z \in L_p(\Omega, \mathfrak{A}, P; Z), \quad 1 \leqslant p < \infty.$$

These convergence results can be used to verify the assumptions of Lemma 1 or Theorem 2 (for problem (1)').

(ii) If Z is a Hilbert space and $z \in L_2(\Omega, \mathfrak{A}, P; Z)$, it is known that for all $m \in N$ z_m minimizes the distance $\|z - y\|_{L_2}$ subject to $y \in L_2(\Omega, \mathfrak{A}_m, P; Z)$.

This fact is important for a wide class of applications in stochastic optimization, because one often has the following type of inequalities (see e.g. [14], Theorem 4 and Remark 5, [22], p. 325):

$$\big|E\big[g\big(z(\omega), u\big)\big] - E\big[g\big(y(\omega), u\big)\big]\big| \leqslant k(z, y, u)\|z - y\|_{L_2}, \quad u \in C,$$

where $k\colon L_2 \times L_2 \times C \to [0, \infty)$.

(iii) If $\tilde{Z}$ is convex and closed, $g(\cdot, u)$: $\tilde{Z} \to R$ is convex for $u \in C$ and $\mathfrak{A}_m \subseteq \mathfrak{A}_{m+1}$, then Jensen's inequality yields:

$$E\left[g\left(z_m(\omega), u\right)\right] \leqslant E\left[g\left(z_{m+1}(\omega), u\right)\right] \leqslant J(u)$$

(comp. [14], Theorems 5 and 7). This estimate provides a possibility to get error bounds. [14] contains results concerning computable a posteriori error bounds.

(iv) For wide classes of random variables z it is possible to produce such partitions $\{A_{ml}\}_{l=1,\ldots,m} \subset \mathfrak{A}$ of Ω, $m \in N$, that (i) holds and the computation of $P(A_{ml})$, $E(z|A_{ml})$, $l = 1, \ldots, m$, $m \in N$, is possible by using finite-dimensional distribution functions of z (see [27], Chapter 6). But, in general, the computation of multiple integrals is necessary. For important special classes of random variables there exist effective methods for the computation of the probabilities and conditional expectations. For instance, for the cases of multivariate normal and exponential random vectors we refer to [29], for Gaussian random variables in Banach spaces to [27], Chapter 6 and in the Banach space $Z := C([0,1])$ to [28], Chapter 6.

5. Approximation of optimal control problems with random operator equations

We consider optimal control problems of the following type:

$$J(u) := E\left[g\left(x(\omega)\right)\right] \to \text{Min!} \quad \text{s.t. } u \in C, \tag{3}$$

$$T\left(\omega, x(\omega), u\right) = 0, \quad x(\omega) \in C_x(\omega) \text{ a.s.},$$

where

g: $X \to R$ is continuous, C_x: $\Omega \to \mathcal{P}(X)$,

T: $\mathrm{Gr}\, C_x \times C \to Y$ has the property that for all $u \in C$ $T(\cdot, \cdot, u)$ is $\mathfrak{A} \otimes \mathfrak{B}(X)$-measurable and

$$S(\omega, u) := \{x |\ x \in C_x(\omega),\ T(\omega, x, u) = 0\} \neq \emptyset \text{ a.s.},$$

X, Y are real separable Banach spaces,
$(\Omega, \mathfrak{A}, P)$ is complete.

Problem (3) is somewhat simpler than the general problem in Example (b) in Chapter 2. We note that under the above assumptions for all $u \in C$ the random operator equation has a random solution x: $\Omega \to X$ ([24], Theorem 1, see also [3], [8], [27], Chapter 2). Thus $g\left(x(\cdot)\right)$: $\Omega \to R$ is measurable and J: $C \to \bar{R}$ is well-defined. $T(\cdot, \cdot, u)$, $u \in C$, represents a so-called random operator on the random domain C_x ([8]).

For the approximation of (3) let

$$C_{x,m}\colon \Omega\to\mathscr{P}(X),\qquad T_m\colon \operatorname{Gr} C_{x,m}\times C\to Y,\qquad m\in N,$$

be given such that for all $u\in C$ $T_m(\cdot,\cdot,u)$ is $\mathfrak{A}\otimes\mathfrak{B}(X)$-measurable and

$$S_m(\omega,u):=\{x\in C_{x,m}(\omega)|\ T_m(\omega,x,u)=0\}\neq\emptyset \text{ a.s.}$$

Then we consider additionally for all $m\in N$:

$$(3m)\qquad \begin{aligned} &J_m(u):=E\left[g\left(x_m(\omega)\right)\right]\to\text{Min!}\quad \text{s.t. } u\in C,\\ &T_m\left(\omega,x_m(\omega),u\right)=0,\quad x_m(\omega)\in C_{x,m}(\omega),\ \text{a.s.},\end{aligned}$$

where $x_m\colon \Omega\to X$ is measurable and $x_m(\omega)\in S_m(\omega,u)$ a.s. Now we apply the results of the preceding chapters to the approximation of (3) by problems $(3m)$, $m\in N$.

THEOREM 3. *Let the above and the following assumptions be fulfilled:*

(a) *C is a set with convergence ϱ and ϱ-compact;*

(b) *for all $u\in C$, $m\in N$ and almost all $\omega\in\Omega$, $S(\omega,u)$ and $S_m(\omega,u)$ are singletons and $S(\omega,\cdot)\colon C\to X$ is ϱ-continuous;*

(c) *there exists a random variable $c\in L_1(\Omega,\mathfrak{A},P)$ such that for almost all $\omega\in\Omega$ and all $x\in\tilde{X}(\omega)$, where*

$$\tilde{X}(\omega):=\{S(\omega,u)|\ u\in C\}\cup\{S_m(\omega,u)|\ u\in C,\ m\in N\}$$

we have

$$|g(x)|\leqslant c(\omega)$$

and g is uniformly continuous on $\bigcup_{\omega\in\Omega}\tilde{X}(\omega)$;

(d) *for all $u\in C$ we have $\lim\limits_{m\to\infty}\|S_m(\omega,u)-S(\omega,u)\|_X=0$ a.s.;*

(e) *for all sequences u_m, $v_m\in C$, $m\in N$, such that*

$$\varrho\text{-}\lim_{m\to\infty}u_m=\varrho\text{-}\lim_{m\to\infty}v_m,$$

we have

$$\lim_{m\to\infty}\|S_m(\omega,u_m)-S_m(\omega,v_m)\|_X=0\ \text{a.s.}$$

Then the assertion of Theorem 1 is valid for (3) and $(3m)$, $m\in N$, respectively.

Proof. We have to verify the conditions of Theorem 2. First we note that C is ϱ-compact and by (b), (c), $J\colon C\to R$, $J(u):=E\left[g\left(S(\omega,u)\right)\right]$ is ϱ-continuous. Further we have by (d), for all $u\in C$,

$$\lim_{m\to\infty}\left|g\left(S_m(\omega,u)\right)-g\left(S(\omega,u)\right)\right|=0\qquad\text{a.s.}$$

Therefore condition (c) yields

$$\lim_{m\to\infty} \|g(S_m(\cdot, u)) - g(S(\cdot, u))\|_{L_1} = 0.$$

Analogously it results from (e) and (c) that for all $u_m, v_m \in C$, $m \in N$, such that $\varrho\text{-}\lim_{m\to\infty} u_m = \varrho\text{-}\lim_{m\to\infty} v_m$, we have

$$\lim_{m\to\infty} \|g(S_m(\cdot, u_m)) - g(S_m(\cdot, v_m))\|_{L_1} = 0. \ \blacksquare$$

Remark 7. Let us discuss some of the assumptions of Theorem 3:

(a) For applications in optimal control it is a typical situation that the set of admissible controls is weakly compact. Therefore one has to choose ϱ as the weak convergence in a certain Banach space. In this case the ϱ-continuity of $S(\omega, \cdot)$ a.s. is often fulfilled in applications (see Remark 10, [10]). But, it turns out that condition (e) of Theorem 3 is the most essential assumption, which can be described as "uniform ϱ-continuity of $\{S_m(\omega, \cdot)\}_{m\in N}$" a.s.

(b) It is assumed that the occurring random operator equations have a locally unique random solution. In this case [27], Theorem 2, contains general sufficient conditions for the a.s.-convergence of random solutions ((d) in Theorem 3).

(c) Condition (c) of Theorem 3 represents an integrability property for g, which yields L_1-convergence if a.s.-convergence is already available. We remark that various sufficient conditions for (c) exist, e.g., growth conditions for g and integrable bounds for $\tilde{X}(\omega)$, $\omega \in \Omega$.

Remark 8. We mention that a direct application of Lemma 1 to obtain approximation results for (3) is possible, too. Obviously, for the convergence of the optimal values, the following condition is sufficient:

For all sequences $u_m \in C$, $m \in N$, we have

$$\lim_{m\to\infty} \|g(S_m(\cdot, u_m)) - g(S(\cdot, u_m))\|_{L_1} = 0.$$

If a condition like (c) in Theorem 3 is assumed, it is sufficient that

$$\lim_{m\to\infty} \|S_m(\omega, u_m) - S(\omega, u_m)\|_X = 0 \text{ a.s.}$$

Such a case was treated in [26], Theorem 2 and we refer to [22], Chapter 6 for similar conditions (in connection with a more special problem).

Now let us turn to the application of "discretization schemes". We assume that a discretization scheme $\{\mathfrak{A}_m, C_{x,m}, T_m\}_{m\in N}$ (see [27], Def. 2) is given, i.e., for all $m \in N$ there exist

(i) a finite partition $\{A_{ml}\}_{l=1,\dots,m} \subset \mathfrak{A}$ of Ω such that

$$\mathfrak{A}_m := \sigma(\{A_{ml}\}_{l=1,\dots,m});$$

(ii) $C_{ml} \in \mathscr{P}(X)$, $T_{ml}\colon C_{ml} \times C \to Y$, $l = 1, \dots, m$, such that

$$C_{x,m}(\omega) = C_{ml},$$

$$T_m(\omega, \cdot, \cdot) = T_{ml}, \qquad \omega \in A_{ml},\ l = 1, \dots, m.$$

Then we have $S_m(\omega, u) = \{x \in C_{ml} \mid T_{ml}(x, u) = 0\}$, $\omega \in A_{ml}$, $l = 1, \dots, m$, and random solutions x_m may be chosen as discrete random variables.

Remark 9. (a) In the case of the above-mentioned discretization schemes problem $(3m)$, $m \in N$, is of the following form:

$$J_m(u) = \sum_{l=1}^{m} g(x_{ml}) P(A_{ml}), \qquad u \in C,$$

$$T_{ml}(x_{ml}, u) = 0, \qquad x_{ml} \in C_{ml}, \qquad l = 1, \dots, m.$$

It turns out that this problem is a "deterministic" one and that by $\{\mathfrak{A}_m, C_{x,m}, T_m\}_{m \in N}$ a discretization scheme for (3) is induced (see Def. 1). But we observe that the computation of $J_m(u)$ requires to solve m deterministic operator equations (see [26], Remark 3 (a)).

An analogous situation appears in the case of discretization schemes for two-stage problems of stochastic programming (see [14], [15]).

(b) For a suitable construction of discretization schemes we refer to [27], Remark 10, Chapter 6, and we again suggest to approximate the stochastic input data by conditional expectations (Remark 6).

Remark 10. Finally, let us consider optimal control problems with random ordinary differential equations (see also [26] or [23], p. 174):

$$J(u) := E\big[g\big(x(\omega, t_1)\big)\big] \to \text{Min!} \quad \text{s.t. } u \in C \tag{4}$$

$$\dot{x}(\omega, t) = f\big(t, z(\omega, t), u(t), x(\omega, t)\big), \qquad t \in [t_0, t_1], \tag{4.1}$$

$$x(\omega, t_0) = x_0(\omega),$$

$$h\big(z(\omega, t), x(\omega, t)\big) \leqslant 0, \qquad t \in [t_0, t_1], \tag{4.2}$$

where

$C \subset L_2^r(t_0, t_1)$ is a set of admissible controls,

$g\colon R^n \to R$ is continuous,

$f\colon [t_0, t_1] \times R^s \times R^r \times R^n \to R^n$ is continuous and has the property that for all $u \in C$ there exists a unique random solution of (4.1), which satisfies (4.2) a.s.,

$h\colon R^s \times R^n \to R^p$ is continuous,

z: $\Omega \times [t_0, t_1] \to R^s$ is a stochastic process with continuous sample paths, and

x_0: $\Omega \to R^n$ is a random variable defined on a probability space $(\Omega, \mathfrak{A}, P)$.

We note that this type of problem fits into our class (3) and (without state constraints) was treated in [26]. As already mentioned, in [26] Lemma 1 is used for the approximation of (4). But, under suitable assumptions it seems to be possible to apply Theorem 3 to this situation and thus to extend [26] (see Remarks 7 and 8).

For the application of discretization schemes to (4) we refer to Remark 9, [27], Chapter 7, and [26].

Acknowledgement

The author wishes to thank the organizers of the Banach Center-Semester on "Mathematical Control Theory" for the excellent working conditions, thereby it was possible to prepare a previous version of the present paper during the author's stay in Warsaw in December 1980.

References

[1] B. Bank, J. Guddat, D. Klatte, B. Kummer, K. Tammer, *Nichtlineare parametrische Optimierung*, Humb.-Univ. Berlin, Sektion Mathematik, Seminarbericht Nr. 31, 1981.

[2] D. P. Bertsekas, *Stochastic optimization problems with nondifferentiable cost functionals*, J. Optimization Theory Appl. 12, 2 (1973), 218–231.

[3] A. T. Bharucha-Reid, *Random Integral Equations*, Academic Press, New York 1972.

[4] C. Castaing, M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics 580, Springer, 1977.

[5] S. D. Chatterji, *A note on the convergence of Banach-space valued martingales*, Math. Ann. 153 (1964), 142–149.

[6] J. W. Daniel, *The Approximate Minimization of Functionals*, Prentice-Hall, Inc., Englewood Cliffs, N. J. 1971.

[7] M. A. H. Dempster, Ed., *Stochastic Programming*, Academic Press, London 1980.

[8] H. W. Engl, *Random fixed point theorems*, in: *Nonlinear Equations in Abstract Spaces* (V. Lakshmikantham, Ed.), Academic Press, New York 1978, 67–80.

[9] —, *Existence of measurable optima in stochastic nonlinear programming and control*, Appl. Math. Optim. 5 (1979), 271–281.

[10] K. Glashoff, *Schwache Stetigkeit bei nichtlinearen Kontrollproblemen*, in: *Numerische Methoden bei Optimierungsaufgaben* (ed. L. Collatz, W. Wetterling), ISNM vol. 17, Birkhäuser, Basel 1973, 51–58.

[11] J.-B. Hiriart-Urruty, *Contributions à la programmation mathématique: cas déterministe et stochastique*, Thèse, Université de Clermont-Ferrand II (1977).

[12] P. Kall, *Approximations to stochastic programs with complete fixed recourse*, Numer. Math. 22 (1974), 333–339.

[13] P. Kall, *Computational methods for solving two-stage stochastic linear programming problems*, Z. Angew. Math. Phys. 30 (1979), 261–271.

[14] —, D. Stoyan, *Solving stochastic programming problems with recourse including error bounds*, Math. Operationsforsch. Statist. Ser. Optim. 13 (1982), 431–447.

[15] V. Kaňková, *Stability in the stochastic programming*, Kybernetika 14, 5 (1978), 339–349.

[16] —, *Optimization problem with parameter and its application to the problems of two-stage stochastic nonlinear programming*, ibid. 16, 5 (1980), 411–425.

[17] [O. Karma] О. Карма, *Об аппроксимации в задачах оптимизации*, Уч. зап. Тартуск. ун-та 448 (1978), 99-106.

[18] A. Kirsch, *Zur Störung von Optimierungsaufgaben unter besonderer Berücksichtigung von Optimalen Steuerungsproblemen*, Dissertation, Universität Göttingen (1978).

[19] R. Kluge, *Nichtlineare Variationsungleichungen und Extremalaufgaben*, VEB Deutscher Verlag der Wissenschaften, Berlin 1979.

[20] W. Krabs, *Stetige Abänderung der Daten bei nichtlinearer Optimierung und ihre Konsequenzen*, Methods of Operations Research 25 (1977), 93–113.

[21] B. Kummer, *Global stability of optimization problems*, Math. Operationsforsch. Statist., Ser. Optimization 8, 3 (1977), 367–383.

[22] K. Marti, *Approximationen der Entscheidungsprobleme mit linearer Ergebnisfunktion und positiv homogener, subadditiver Verlustfunktion*, Z. Wahrscheinlichkeitstheorie verw. Geb. 31 (1975), 203–233.

[23] —, *Approximationen stochastischer Optimierungsprobleme*, Mathematical systems in economics, vol. 43, Hain, Königsstein 1979.

[24] A. Nowak, *Random solutions of equations*, Trans. Eighth Prague Conf. on Information Theory, Statist. Decision Funct. and Random Processes, 1978, vol. B, Prague 1978, 77–82.

[25] R. T. Rockafellar, *Integral functionals, normal integrands and measurable selections*, in: *Nonlinear Operators and the Calculus of Variations*, Lecture Notes in Mathematics 543, Springer, 1976, 157–207.

[26] W. Römisch, *An approximation method in stochastic optimal control*, in: *Optimization Techniques*, Lecture Notes in Control and Information Sciences 22, Springer, 1980, 169–178.

[27] —, *On the approximate solution of random operator equations*, Wiss. Z. Humb.-Univ. Berlin, Math.-Nat. Reihe 30, 5 (1981), 455–462.

[28] W. Römisch, R. Schulze, *Kennwertmethoden für stochastische Volterrasche Integralgleichungen*, ibid. 28, 4 (1979), 523–533.

[29] R. Schulze, *Determinierte Simulation von Zufallsvektoren*, ibid. 30, 5 (1981), 449–454.

[30] F. Stummel, *Discrete convergence of mappings*; in: *Topics in Numerical Analysis* (J. J. H. Miller, Ed.), Academic Press, New York 1973, 285–310.

[31] K. Tammer, *Allgemeine Lösungsbegriffe für stochastische Optimierungsprobleme mit festem Restriktionsbereich*, Wiss. Z. Humb.-Univ. Berlin, Math.-Nat. Reihe 26, 5 (1977), 595–602.

[32] R. M. Van Slyke, R. Wets, *Stochastic programs in abstract spaces*, Stochastic Optimization and Control (H. F. Karreman, Ed.), Wiley, New York 1968, 25–45.

[33] D. H. Wagner, *Survey of measurable selection theorems*, SIAM J. Control Optimization 15, 5 (1977), 859–903.

[34] D. W. Walkup, R. Wets, *Stochastic programs with recourse. II: On the continuity of the objective*, SIAM J. Appl. Math. 17, 1 (1969), 98–103.