Journal of Convex Analysis
Volume 2 (1995), No.1/2, 1–17

Consistency of Minimizers and the SLLN for Stochastic Programs[1]

Zvi Artstein[2]
Department of Theoretical Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel.
e-mail: mtarts@weizmann.weizmann.ac.il

Roger J-B Wets
Department of Mathematics, University of California, Davis, CA 95616, U.S.A.
e-mail: rjb@ucdavis.edu

Received 21 June 1994
Revised manuscript received 1 February 1995

Dedicated to R. T. Rockafellar on his 60th Birthday

A general strong law of large numbers for stochastic programs is established. It is shown that solutions and approximate solutions may not be consistent with the strong law in general, but consistency holds locally, or when the decision space is compact. An additional integrability condition implies the uniform consistency of approximate solutions. The results are applied in the context of linear recourse models.

1. Introduction

The paper examines relations between solutions of a stochastic optimization problem, and the solutions of large sampled versions of the problem. We consider an abstract stochastic program of the form

$$\underset{x\in X}{\text{minimize}}\ E^{P(d\xi)}(f(x,\xi)) \tag{$*$}$$

where $E^{P(d\xi)}$ is the expectation operator with respect to the probability measure P over the space Ξ of random elements. The decision space here is taken as a metric space. For a given sequence $\xi_1,\dots,\xi_n$ of realizations of the random variable we form the deterministic problem

$$\underset{x\in X}{\text{minimize}}\ \frac{1}{n}(f(x,\xi_1)+\cdots+f(x,\xi_n)). \tag{$**$}$$

When the ξ_i are drawn according to the distribution P, and independently, a reasoning in line of a strong law of large numbers indicates toward relations between solutions of $(*)$

[1] Research supported by a grant from the U.S.-Israel Binational Science Foundation.
[2] Incumbent of the Hettie H. Heineman Professorial Chair in Mathematics.

ISSN 0944-6532 / \$ 2.50 © Heldermann Verlag

and those of (∗∗) for large n. Such relations, especially the robustness and the statistical consistency, are explored in the sequel. We note that the abstract form (∗) incorporates many concrete stochastic optimization problems, see e.g. Wets [15]; a linear recourse model is analyzed in the closing section.

As a preliminary consideration we examine the functions on the space X which determine the minimization problems (∗) and (∗∗), namely

$$(Ef)(x) = E^{P(d\xi)}(f(x,\xi)) \tag{1.1}$$

(we suppress the superscript P on the operator E when no confusion may arise), and for $\xi_1, \dots, \xi_n$ fixed

$$F(x; \xi_1, \dots, \xi_n) = \frac{1}{n}\sum_{j=1}^{n} f(x, \xi_j). \tag{1.2}$$

Under quite relaxed conditions (e.g. boundedness from below by an integrable function), if the ξ_j are independent and P-distributed, the values $F(x; \xi_1, \dots, \xi_n)$ converge almost surely to $(Ef)(x)$, for x fixed. This follows from the standard strong law of large numbers, see e.g. Ash [1, Section 7].

But a more general property holds. Under some conditions, the functions $F(\cdot\,; \xi_1, \dots, \xi_n)$ converge almost surely to the function $(Ef)(\cdot)$, this when the convergence is with respect to the epi-convergence of functions, which can be interpreted as a one-sided version of uniform convergence. Such results are given in King and Wets [9], Attouch and Wets [3], Castaing and Ezzaki [5] and Hess [7]; we elaborate on that in Section 2 where such a strong law of large numbers for functions is verified under somewhat weaker conditions.

In the rest of the paper we are interested in the robustness and consistency of optimal solutions and near optimal solutions of (∗). We ask, in particular, whether employing an optimal solution of (∗) for the problem (∗∗) with large n, yields with high probability a good approximate solution of (∗∗). The same question is examined with respect to uniformly small deviations from solutions of (∗).

The consistency of optimal solutions is examined in Section 3. We give a counterexample for global consistency; but provide a local consistency result, and establish the global consistency when the decision space is compact.

The uniform consistency of approximate solutions is examined in Section 4. Here even local robustness fails, unless the underlying conditions are strengthened. We display such sufficient conditions, and in the closing section demonstrate their applicability in the context of linear recourse models.

2. A Law of Large Numbers

We start by setting the conditions under which we work. We then recall the notion of epi-convergence, and state and prove the corresponding strong law of large numbers.

We assume that X is a complete separable metric space. We assume that (Ξ, Σ, P) is a probability space, and that the σ-field Σ is complete with respect to P, namely, a subset of a null set in Σ also belongs to Σ (this assumption is for technical convenience only; standard techniques can be used to eliminate it).

Two assumptions are placed on the cost functions $f(x, \xi)$ as follows.

Assumption 2.1. The function

$$f(x, \xi): X \times \Xi \to (-\infty, \infty]$$

is measurable on $X \times \Xi$ (where on X the Borel field is taken), and $f(\cdot, \xi)$ for ξ fixed is lower semicontinuous in x, namely $x_k \to x_0$ implies $\liminf f(x_k, \xi) \geq f(x_0, \xi)$.

Assumption 2.2. For each $x_0 \in X$ there exists an open set N_0 in X and an integrable function $\alpha_0(\xi): \Xi \to (-\infty, \infty)$ such that $x_0 \in N_0$ and for almost all $\xi \in \Xi$ the inequality

$$f(x, \xi) \geq \alpha_0(\xi) \tag{2.1}$$

holds for all $x \in N_0$.

The assumptions are quite relaxed, and are satisfied by the typical stochastic programs. In fact, in quite a number of cases stochastic programs exhibit continuity in the x variable; the semicontinuity allows in turn to model constraints, as we can set $f(x, \xi) = \infty$ for the nonfeasible cases. We also note that the terminology random lower semicontinuous functions is often used in the literature to describe functions that satisfy Assumption 2.1.

Note that Fatou's lemma (see e.g. Ash [1, 1.6.8(a)]), implies that $(Ef)(x)$ is lower semicontinuous under Assumptions 2.1 and 2.2. Also, it is clear then that $F(x; \xi_1, \ldots, \xi_n)$ is lower semicontinuous in the x variable. We recall now the concept of epi-convergence for lower semicontinuous functions. Consult with Attouch [2] and Rockafellar and Wets [12] for a thorough analysis of the concept and its relation with minimization problems.
Consider the sequence of lower semicontinuous functions

$$F_k(x): X \to (-\infty, \infty].$$

The sequence $F_k(\cdot)$ epi-converges to $F_0(\cdot)$ if the following two properties hold for every $x_0 \in X$.

(I) $\liminf F_k(x_k) \geq F_0(x_0)$ whenever $x_k \to x_0$, and

(II) $\lim F_k(y_k) = F_0(x_0)$ for at least one sequence $y_k \to x_0$.

Epi-convergence of lower semicontinuous functions is equivalent (as it is easy to see) to the set convergence of their epigraphs. To this end we introduce the notation

$$\text{epi } h = \{(x, r): r \geq h(x)\} \tag{2.2}$$

for the epigraph of the function $h(\cdot)$. When $h(\cdot)$ is lower semicontinuous the set epi h is a closed set in $X \times (-\infty, \infty]$. Under some conditions on X, e.g. local compactness, epi-convergence is metrizable; we shall not need this property.

We state and prove now a strong law of large numbers with respect to epi-convergence. Such results were initiated in King and Wets [9], where the strong law was verified for convex integrands in a euclidean space. Attouch and Wets [3] have established the law for a separable Banach space, and general lower semicontinuous functions. Here we further relax the conditions, and work within a metric space; our approach though is in line with that of [3], the difference lies only in the way the equi-Lipschitz approximations are taken. Also worth mentioning in this respect are the contributions of Castaing and Ezzaki [5] and Hess [7] (we thank a referee for pointing them out to us). In [5] the authors use the Lipschitz approximations to provide limit laws for martingales and supperadditive sequences. In [7] the result is verified for a general, not necessarily complete, metric space. In turn the minorization condition is somewhat stronger. We note that a particular case of the epi-convergence would be the standard vector-valued slln into the space of continuous functions with the sup norm. But in our case the vector-valued methods seem not to apply. Also note the related observations by Plachky [10].

For convenience we state the results using a sequence $\xi_1, \xi_2, \ldots$ of independent and identically distributed samplings from Ξ; then almost sure properties refer to the denumerable power of (Ω, Σ, P). An easy reduction would cover the case of a sequence $f_j(x, \xi)$ of i.i.d. cost functions.

Theorem 2.3. *Let $\xi_1, \xi_2, \ldots$ be a sequence of independent and identically P-distributed drawings from Ξ. Under Assumptions* 2.1 *and* 2.2*, the sequence $F(\cdot\,; \xi_1, \ldots, \xi_n)$ almost surely epi-converges to $(Ef)(\cdot)$.*

Proof. We first prove that almost surely property (I) of epi-convergence holds.
Consider a fixed open set N_0 in X, on which Assumption 2.2 holds. We verify first that almost surely property (I) holds for the restriction of $F(\cdot\,; \xi_1, \ldots, \xi_n)$ and of $(Ef)(\cdot)$ to N_0.
One possibility is that on a set of ξ of positive measure the function $f(x, \xi) = \infty$ identically. Then the result is trivial, as $(Ef)(x) = \infty$ for $x \in N_0$, and clearly almost surely $F(x; \xi_1, \ldots, \xi_n) = \infty$ for n large. Hence we proceed under the assumption that almost surely $f(x, \xi) < \infty$ for some $x \in N_0$.
We claim that a sequence $g_k(x, \xi)$ exists, of functions

$$g_k(x, \xi) \colon N_0 \times \Xi \to (-\infty, \infty]$$

satisfying the following properties

(i) Each g_k is measurable on $N_0 \times \Xi$.

(ii) Each g_k is a Lipschitz function of the variable x, with Lipschitz constant independent of the variable ξ.

(iii) $g_k(x, \xi) \geq \alpha_0(\xi)$, and $g_k(x, \xi)$ converges monotonically to $f(x, \xi)$ as $k \to \infty$.

In order to provide the sequence g_k we employ the very useful construction of G. Beer [4]. Define first

$$g_0(x, \xi) = \alpha_0(\xi). \tag{2.3}$$

Suppose that $g_k(x, \xi)$ is given. Define

$$\varphi_k(x, \xi) = \inf\{d(x, y) + |g_k(x, \xi) - r| \colon y \in N_0, r \geq f(y, \xi)\} \tag{2.4}$$

where $d(\cdot, \cdot)$ is the metric on X; namely, $\varphi_k(x, y)$ is the distance (of an L_1 nature) of $(x, g_k(x, \xi))$ from the epigraph of $f(\cdot, \xi)$ restricted to N_0. Since $f(\cdot, \xi)$ is not identically $+\infty$, the function $\varphi_k(x, \xi)$ is finite. Define

$$g_{k+1}(x, \xi) = g_k(x, \xi) + \varphi_k(x, \xi). \tag{2.5}$$

We verify now the three desired properties. The measurability of $g_{k+1}(x, \xi)$ follows, by induction, from that of $\varphi_k(x, \xi)$. To establish the measurability of the latter, consider the set-valued map epi $f(\cdot, \xi)$, namely a map from Ξ to the closed subsets of $N_0 \times (-\infty, \infty]$. By Lemma VII-1 of Castaing and Valadier [6, page 196], this set-valued function has a measurable graph. Further, by the Castaing Representation, see Rockafellar [11, Thm. 1B] or Castaing and Valadier [6, Section 2], a sequence $s_k(\xi)\colon \Xi \to N_0 \times (-\infty, \infty]$ of measurable point-valued functions exists, such that

$$\text{epi } f(\cdot, \xi) = \text{closure of } \{s_1(\xi), s_2(\xi), \ldots\} \tag{2.6}$$

(here epi $f(\cdot,\xi)$ is the epigraph of the function $f(x,\xi)$ for ξ fixed). Each $s_j(\xi)$ is of the form $(x_j(\xi), r_j(\xi))$. Since for an arbitrary fixed c

$$\{(x,\xi)\colon \varphi_k(x,\xi) < c\} = \bigcup_{j=1}^{\infty} \Big\{(x,\xi)\colon d(x, x_j(\xi)) + |g_k(x,\xi) - r_j(\xi)| < c\Big\}$$

and since $x_j(\cdot)$, $g_k(\cdot,\cdot)$, $r_j(\cdot)$ are measurable, it indeed follows that $\varphi_k(\cdot,\cdot)$ is measurable. This verifies property (i).

Property (ii) follows directly from the construction. Indeed, $g_0(x,\xi)$ is constant in x, namely has a Lipschitz constant 0. Suppose that L is a Lipschitz constant for $g_k(\cdot,\xi)$. From the definition of $\varphi_k(x,\xi)$ in (2.4) it follows that it is Lipschitz with constant $L+1$; hence by (2.4) the value $2L+1$ is a Lipschitz constant for $g_{k+1}(\cdot,\xi)$ for ξ fixed. Simple recursion shows then that $2^k - 1$ is a Lipschitz constant for $g_k(\cdot,\xi)$, and property (ii) is proved.

Property (iii) is clear, and as was mentioned, it was observed by Beer [4]. This completes the proof of the claim.

With the aid of the sequence $g_k(x,\xi)$ we can verify Property (I) on N_0 as follows.
For x and k fixed, define

$$G_k(x;\xi_1,\ldots,\xi_n) = \frac{1}{n}\sum_{j=1}^{n} g_k(x,\xi_j) \tag{2.7}$$

since $g_k(x,\xi) \ge \alpha_0(\xi)$. The standard strong law of large numbers implies that for x fixed, almost surely $G_k(x;\xi_1,\ldots,\xi_n)$ converge to

$$(Eg_k)(x) = E^{P(d\xi)}(g_k(x,\xi)). \tag{2.8}$$

Let x_i, $i = 1,2,\ldots$ be a dense sequence in N_0. By the countability, almost surely $G_k(x_i;\xi_1,\ldots,\xi_n)$ converge as $n \to \infty$ to $(Eg_k)(x_i)$ for all $i = 1,2,\ldots$, say that this holds for the collection Θ_k of sequences. The Lipschitz property in (ii) implies that for $(\xi_1,\xi_2,\ldots)$ in Θ_k the convergence of $G_k(x;\xi_1,\ldots,\xi_n)$ to $(Eg_k)(x)$ holds for all $x \in N_0$. Finally, on the intersection of Θ_k for $k = 0,1,\ldots$, we have that $G_k(x;\xi_1,\ldots,\xi_n)$ converge as $n \to \infty$ to $(Eg_k)(x)$ for all $x \in N_0$ and all k. And note that the intersection, which we denote by Θ, has a full measure. We are ready to prove that Property (I) holds for every $F(\cdot\,;\xi_1,\ldots,\xi_n)$ with $(\xi_1,\xi_2,\ldots)$ in Θ, namely almost surely.

Let $x_0 \in N_0$ be arbitrary. Suppose first that $(Ef)(x_0) < \infty$. The monotonic convergence property (iii) implies that for k large the value $(Eg_k)(x_0)$ is close to $(Ef)(x_0)$, say

$$(Ef)(x_0) - (Eg_k)(x_0) < \epsilon.$$

Since for a fixed k, the Lipschitz constant of $g_k(\cdot, \xi)$ is shared by all the ξ (it is $2^k - 1$), it follows that $(Eg_k)(\cdot)$ is also Lipschitz, and the same is true for each $G_k(\cdot, \xi_1, \ldots, \xi_n)$. For sequences in Θ, the values $G_k(\cdot\, ; \xi_1, \ldots, \xi_n)$ converge to $(Eg_k)(\cdot)$, and since

$$F(x; \xi_1, \ldots, \xi_n) \geq G_k(x; \xi_1, \ldots, \xi_n)$$

always, it follows that whenever $x_n \to x_0$

$$\liminf F(x_n; \xi_1, \ldots, \xi_n) \geq (Ef)(x_0) - \epsilon \tag{2.9}$$

for all sequences in Θ. Since ϵ is arbitrarily small it follows that (2.9) holds also with $\epsilon = 0$. This concludes the proof for $(Ef)(x_0) < \infty$. The case $(Ef)(x_0) = \infty$ is similar. The only modification is to replace the phrase ϵ arbitrarily small by $(Eg_k)(x_0)$ arbitrarily large. We leave out the details.
So far we verified that almost surely Property (I) holds for $x \in N_0$. Since X is assumed separable, a countable number of such neighborhoods N_0 cover X, this by Assumption 2.2. Therefore on a set of full measure, namely the intersection of the sequence of the corresponding Θ, the inequality (2.9) holds with $\epsilon = 0$ for all $x_0 \in X$ and sequences $x_n \to x_0$. This concludes the first part of the proof.

We now verify that almost surely Property (II) of epi-convergence holds. We consider the epigraph

$$\mathrm{epi}(Ef)(\cdot) \tag{2.10}$$

which is a set in $X \times (-\infty, \infty]$ (by Assumption 2.2), and a dense sequence (x_i, r_i) in it. Note that $r_i = \infty$ is allowed. The lower semicontinuity of $(Ef)(\cdot)$ implies in particular that the sequence $(x_i, (Ef)(x_i))$ is then dense in the lower boundary of the epigraph (2.10). Namely, for each $x_0 \in X$, a subsequence of $(x_i, (Ef)(x_i))$ converges to $(x_0, (Ef)(x_0))$. We apply now the standard strong law of large numbers for each x_i separately. The countability implies that on a set Θ of sequences $(\xi_1, \xi_2, \ldots)$ of full measure, $F(x_i; \xi_1, \ldots, \xi_n)$ converge to $(Ef)(x_i)$ as $n \to \infty$, for each $i = 1, 2, \ldots$. Clearly, for each element in Θ, if x_0 is given, a subsequence, say y_n, of x_i can be deduced (related to the aforementioned subsequence of x_i), such that

$$\lim F(y_n; \xi_1, \ldots, \xi_n) = (Ef)(x_0). \tag{2.11}$$

This verifies Property (II) of epi-convergence on Θ, that is almost surely. This concludes the proof. □

3. Robustness and Consistency of Minimizers

The epi-convergence almost surely of $F(\cdot\, ; \xi_1, \ldots, \xi_n)$ to $(Ef)(\cdot)$ implies a number of properties for the convergence of the minimizers and of the infima. For instance, the set of approximate solutions of $(**)$ almost surely converges topologically to the corresponding set of approximate solutions of $(*)$. See Rockafellar and Wets [12] for an elaborate discussion. Here we go beyond such statements and examine robust properties of solutions and approximate solutions.

We define first the notion of asymptotic minimizers, and discuss its relation to the robustness and consistency. We show then that for a general decision space, a minimizer of

$(*)$ may not be robust with respect to independent samples. We establish consistency for a compact X, and provide a local consistency result for a general space.

A point x_0 is an ϵ-minimizer of $H(\cdot)$ if

$$H(x_0) - \min_{x \in X} H(x) \le \epsilon. \tag{3.1}$$

A point x_0 is an asymptotic ϵ-minimizer of the sequence $H_k(\cdot)$ if

$$\limsup_{k \to \infty} \Big(H_k(x_0) - \min_{x \in X} H_k(x) \Big) \le \epsilon. \tag{3.2}$$

A 0-minimizer is called a minimizer, and likewise for an asymptotic 0-minimizer.

If $H_k(\cdot)$ converges in some sense to $H_0(\cdot)$, it it natural to inquire whether a minimizer of $H_0(\cdot)$ is an asymptotic minimizer of the sequence $H_k(\cdot)$. This is the robustness of the minimizer with respect to the convergence. We note that minimizers are not, in general, robust with respect to the epi-convergence of lower semicontinuous functions. For instance, for x scalar let $H_0(x) = 1$ if $x < 0$ and $H_0(x) = x$ if $x \ge 0$. Let $H_k(x) = H_0(x - k^{-1})$. Then $H_k(\cdot)$ epi-converge to $H_0(\cdot)$, yet the minimizer $x^* = 0$ of $H_0(\cdot)$ is not an asymptotic minimizer of $H_k(\cdot)$.

The situation may be different in the framework of the almost sure convergence given in Theorem 2.3, as the strong law can be applied to $f(x^*, \cdot)$ with x^* being the solution of $(*)$. Asymptotic minimization is then called consistency, in line with the terminology in statistical estimates.

We find, however, that the consistency of the minimizer of $(*)$ may fail in general.

Example 3.1. Let $\Xi = [0, 1]$ with the Lebesgue measure. Let $X = \{0, 1, 2, \ldots\}$ with the line distance. We define

$$f(0, \xi) = 0. \tag{3.3}$$

We proceed successively. In the jth step ($j \ge 2$) we define $f(k, \xi)$ for $\binom{j^2}{j}$ numbers, say for the numbers

$$m = k(j) + 1, \ldots, k(j) + \binom{j^2}{j}. \tag{3.4}$$

This is done as follows. We divide Ξ into j^2 intervals of equal length. To each group of j such intervals we associate a different number from (3.4), and hence the $\binom{j^2}{j}$ numbers in the sequence. Denote the union of the j intervals associated with the number m by $I_{j,m}$. Define

$$f(m, \xi) = \begin{cases} -1 & \text{if } \xi \in I_{j,m} \\ 2 & \text{otherwise.} \end{cases} \tag{3.5}$$

This for $m = k(j) + 1, \ldots, k(j) + \binom{j^2}{j}$.

The cost function f satisfies, of course, all the assumptions. The expectation is easily computed as

$$(Ef)(m) = (2j^2 - 3j)j^{-2} \quad \text{if} \quad k(j) + 1 \le m \le k(j) + \binom{j^2}{j}. \tag{3.6}$$

Therefore $x^* = 0$ is the unique minimizer of $(Ef)(\cdot)$, with $(Ef)(0) = 0$, and $(Ef)(m) \geq 1$ for $m > 0$. But, for every sample $(\xi_1, \ldots, \xi_j)$ there is at least one natural number m with

$$F(m; \xi_1, \ldots, \xi_j) = -1. \tag{3.7}$$

Indeed, such an m can be found among the numbers in (3.4). This completes the example.

The situation is different when the decision space is compact. We show it following a standard, yet useful, lemma.

Lemma 3.2. *If X is compact, then*

$$\min_{x \in X} H(x)$$

is continuous with respect to epi-convergence.

Proof. A stronger result is provided in Attouch [2, Theorem 2.11].

Theorem 3.3. *Suppose that X is compact. Let x^* be a minimizer (an ϵ-minimizer) of $(Ef)(\cdot)$. Then almost surely x^* is an asymptotic minimizer (respectively an asymptotic ϵ-minimizer) of $F(\cdot\,; \xi_1, \ldots, \xi_n)$.*

Proof. The standard strong law of large numbers when applied to $f(x^*, \cdot)$ with x^* fixed, implies that on a set Θ_1 of sequences $(\xi_1, \xi_2, \ldots)$, of full measure the following limit holds

$$\lim_{n \to \infty} F(x^*; \xi_1, \ldots, \xi_n) = (Ef)(x^*). \tag{3.10}$$

The slln of Theorem 2.3 together with Lemma 3.2 imply that on a set Θ_2 of sequences, of full measure, the following limit occurs

$$\lim_{n \to \infty} \left(\min_{x \in X} F(x; \xi_1, \ldots, \xi_n) \right) = \min_{x \in X} (Ef)(x). \tag{3.11}$$

The two limits imply directly the desired asymptotic minimization properties on the intersection of Θ_1 and Θ_2 which is also of full measure. □

Here is an alternative formulation of the preceding result. The probability on the n-product $\Xi \times \Xi \times \cdots \times \Xi$ is the product probability.

Proposition 3.4. *Suppose that X is compact. Let x^* be an ϵ-minimizer of $(Ef)(\cdot)$. Then for every $\eta > 0$,*

$$\mathrm{Prob}\Big\{ F(x^*; \xi_1, \ldots, \xi_n) - \Big(\min_{x \in X} F(x; \xi_1, \ldots, \xi_n) + \epsilon \Big) \geq \eta \Big\} \tag{3.12}$$

tends to 0 as $n \to \infty$; for x^* being a minimizer the same estimate holds with $\epsilon = 0$.

Proof. Follows from Theorem 3.3 in a standard way. □

A formulation similar to that of (3.12) holds locally near a minimizer of $(Ef)(\cdot)$, even in the noncompact case, as follows.

Proposition 3.5. *Let x^* be an ϵ-minimizer of $(Ef)(\cdot)$. For every $\eta > 0$ there are $\delta > 0$ and n_0 such that*

$$\text{Prob}\Big\{\min_{x\in X}(Ef)(x) - \Big(\min_{x\in N_\delta(x^*)} F(x;\xi_1,\dots,\xi_n) + \epsilon\Big) \geq \eta\Big\} < \eta \tag{3.13}$$

for all $n \geq n_0$, where $N_\delta(x^*)$ denotes the δ-neighborhood of x^* in X. For x^* a minimizer the same estimate holds with $\epsilon = 0$.

Proof. On a set Θ_1 of full measure of sequences $(\xi_1, \xi_2, \dots)$ we have

$$F(x^*;\xi_1,\dots,\xi_n) \to (Ef)(x^*). \tag{3.14}$$

The slln in Theorem 2.3 implies that on a set Θ_2 of full measure

$$\liminf F(x_n;\xi_1,\dots,\xi_n) \geq (Ef)(x^*) \tag{3.15}$$

whenever $x_n \to x^*$. In particular, for each fixed sequence in Θ_2, a number $\delta > 0$ exists and n_0 such that

$$F(x;\xi_1,\dots,\xi_n) - F(x^*;\xi_1,\dots,\xi_n) \leq \eta \tag{3.16}$$

if $n \geq n_0$ and $x \in N_\delta(x^*)$. A simple exhaustion argument would show that (3.16) holds with arbitrary large probability for δ small and n_0 high enough. Together with (3.14) and noting that x^* is an ϵ-minimizer the proof is complete. □

4. Uniform Consistency of Approximations

The consistency in Theorem 3.3 applies for a fixed minimizer, or for a fixed approximate minimizer. In this section we inquire about the uniform consistency of approximations. For instance, we ask if with high probability an arbitrary approximate solution of (∗) yields an approximate solution of (∗∗) for large n. We find that this uniformity is false even locally for X compact, unless further conditions are imposed.

Example 4.1. Let $X = \{0, 1, \frac{1}{2}, \frac{1}{3}, \dots\}$ with the metric of the line. Let $\Xi = [0, 1]$ with the Lebesgue measure. We construct an integrand $g(x, \xi)$ by modifying Example 3.1 as follows. At $x = 0$ we use the same value, namely

$$g(0, \xi) = 0. \tag{4.1}$$

For $x = \frac{1}{m}$ we define

$$g(\tfrac{1}{m}, \xi) = \begin{cases} 0 & \text{if } f(m, \xi) = 2 \\ 1 & \text{if } f(m, \xi) = -1 \end{cases} \tag{4.2}$$

with f as defined in Example 3.1. The underlying assumptions hold, including the lower semicontinuity at $x = 0$. It is easy to see that

$$(Eg)(\tfrac{1}{m}) = \tfrac{1}{j} \quad \text{if} \quad k(j) + 1 \leq m \leq k(j) + \binom{j^2}{j} \tag{4.3}$$

with $k(j)$ given in Example 3.1. Since $(Eg)(0) = 0$ it follows that $(Eg)(\cdot)$ is continuous at its minimizer $x^* = 0$. For each sample $(\xi_1, \ldots, \xi_n)$, however, there are points $\frac{1}{m} \in X$ arbitrarily close to 0, with $F(\frac{1}{m}; \xi_1, \ldots, \xi_n) = 1$. This shows that with probability one an arbitrary approximate solution for $(*)$ may fail to provide an approximate solution of $(**)$.

Remark 4.2. Note that the approximation in the preceding example breaks down in one direction, namely for an ϵ-minimizer x of $(*)$ it may occur that $F(x; \xi_1, \ldots, \xi_n) - (Ef)(x^*)$ is large. The smallness of $(Ef)(x^*) - F(x; \xi_1, \ldots, \xi_n)$ is guaranteed in the compact case by Theorem 3.3, and for local perturbations by Proposition 3.5.

An appropriate continuity condition would be sufficient for the desired uniform consistency. Indeed, notice that the integrand in Example 4.1 oscillates rapidly near the minimizer $x^* = 0$. We set first some terminology.

We consider the problem $(*)$ and denote by X^* the ensemble of minimizers of $(Ef)(\cdot)$, namely the solutions of $(*)$. Note that under Assumptions 2.1 and 2.2 the function $(Ef)(\cdot)$ is lower semicontinuous, hence X^* is closed.

Let Y be a subset of X. We say that Y is a manifold of approximations with respect to X^* if $y_k \in Y$ and $y_k \to x^*$ with $x^* \in X^*$, imply that $(Ef)(y_k)$ converge to $(Ef)(x^*)$. Namely on $Y \cup X^*$ the function $(Ef)(\cdot)$ is continuous at X^*.

A manifold of approximations may not contain a full neighborhood of X^*; in fact, in realistic problems it may occur that only perturbations from the minimizers in prescribed directions yield good approximations. The following result refers to a manifold of approximations (which may yet be part of a larger manifold) which enjoys additional properties. What we in fact show is that under the additional properties the approximations in this manifold are statistically consistent.

Theorem 4.3. *Let Y be a manifold of approximations. Suppose that Y is compact and that the following conditions hold.*
(a) *$y_k \in Y$, $y_k \to x^*$ and $x^* \in X^*$ then $f(y_k, \xi) \to f(x^*, \xi)$ for almost every ξ;*
(b) *for each $x^* \in Y \cap X^*$ a neighborhood N_0 of x^* exists, and an integrable function $\beta(\xi)$ such that $|f(y, \xi)| \le \beta(\xi)$ for all $y \in N_0 \cap Y$.*
Then for every $\epsilon > 0$ there exists a $\delta > 0$ such that on a set Θ of sequences $(\xi_1, \xi_2, \ldots)$ of full measure,

$$\limsup_{n\to\infty} \Big(\max_{y \in Y_\delta} F(y; \xi_1, \ldots, \xi_n) - \min_{x \in X} F(x; \xi_1, \ldots, \xi_n) \Big) \le \epsilon \tag{4.4}$$

when Y_δ is the set of δ-minimizers of $(Ef)(\cdot)$ in Y.

Proof. The statement allows for $Y \cap X^*$ to be empty, but then the claim holds trivially. We therefore assume that there are elements in Y which are minimizers of $(Ef)(\cdot)$ on X. We denote $Y^* = Y \cap X^*$.

The lower semicontinuity of $(Ef)(\cdot)$ and the compactness imply that if y_k is a sequence with $(Ef)(y_k)$ converging to $\min(Ef)(\cdot)$, then y_k must converge to Y^*. Therefore we can replace the statement "Y_δ is the set of δ-minimizers of $(Ef)(\cdot)$" by the statement "Y_δ is in the closed δ-neighborhood of Y^* in Y".

Consider now the upper closure of $f(y,\xi)$ with respect to Y, namely the function

$$h(y,\xi) = \limsup_{z\in Y, z\to y} f(z,\xi).$$

Then $h(y,\xi)$ is upper semicontinuous in the y variable on Y. We claim that $h(y,\xi)$ is measurable. To prove the claim let $s_k(\xi) = (y_k(\xi), r_k(\xi))$ be a sequence of measurable functions from Ξ into $Y \times (-\infty, \infty]$, such that

$$\text{epi } f(\cdot,\xi) = \text{closure } \{s_1(\xi), s_2(\xi), \dots\}$$

for all ξ. Such a sequence exists by the Castaing Representation, see Rockafellar [11] or the arguments leading to (2.6). The lower semicontinuity of $f(y,\xi)$ in y implies that

$$\Big(y_k(\xi), f(y_k(\xi),\xi)\Big)$$

is almost surely dense in the graph of $f(\cdot,\xi)$. Therefore $h(y,\xi)$ can also be defined as

$$h(y,\xi) = \limsup_{y_i(\xi)\to y} f(y_k(\xi),\xi)$$

with $y_i(\xi)$ an arbitrary subsequence of $y_k(\xi)$. The measurability of $y_k(\cdot)$, and hence of $f(y_k(\cdot),\cdot)$, yields then the measurability of $h(y,\xi)$ in a routine way as follows. For every $c \in (-\infty,\infty)$ the inverse set $h^{-1}(-\infty,c)$ can be expressed as follows

$$\begin{aligned}&\{(y,\xi) : h(y,\xi) < c\} = \\ &\bigcup_{\delta>0}\bigcap_{\epsilon>0}\bigcup_{k=1}^{\infty}\Big\{(y,\xi) : d(y,y_k(\xi)) < \epsilon\Big\} \cap \Big\{(y,\xi) : f(y,\xi) < c-\delta\Big\}\end{aligned} \tag{4.5}$$

with $d(\cdot,\cdot)$ the metric on Y. The union over $\delta > 0$ and the intersection over $\epsilon > 0$ can be replaced by denumerable operations with $\delta_i \to 0$ and $\epsilon_i \to 0$. Then clearly the right hand side of (4.5) is a measurable set, and hence $h^{-1}(-\infty,c)$ is measurable, which implies the measurability of h.

We note, however, that the continuity assumption (a) implies that $h(y^*,\xi) = f(y^*,\xi)$ for $y^* \in Y^*$, and furthermore (a) holds when $h(y,\xi)$ replaces $f(y,\xi)$. The integrability assumptions (b) and Assumption 2.2 imply that $(Eh)(y)$ is continuous at each $y^* \in Y^*$. Therefore, for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|(Eh)(y) - (Ef)(y^*)| \le \epsilon \tag{4.6}$$

when $d(y,y^*) \le \delta$ for some $y^* \in Y^*$.

We apply now Theorem 3.3 with a decision space being Y_δ, and with the upper semicontinuous integrand $h(y,\xi)$; namely, we apply the result for the upper semicontinuous case, with maximizers replacing minimizers. The inequality (4.6) implies that y^* in Y^* is an ϵ-maximizer on Y_δ. By Theorem 3.3, on a set Θ of sequences $(\xi_1,\xi_2,\dots)$, with full measure, the decision y^* is an asymptotic ϵ-maximizer for

$$H(y;\xi_1,\dots,\xi_n) = \frac{1}{n}\sum_{j=1}^{n} h(y,\xi_j)$$

on Y_δ. Since $F(y;\xi_1,\dots,\xi_n) \le H(y;\xi_1,\dots,\xi_n)$ it follows that y^* is then also an asymptotic ϵ-maximizer on Y_δ. Namely,

$$\liminf\Big(F(y^*;\xi_1,\dots,\xi_n) - \max_{\delta\in Y_\delta} F(y;\xi_1,\dots,\xi_n)\Big) \ge -\epsilon$$

which in a form closer to (4.4) can be written as

$$\limsup\Big(\max_{y\in Y_\delta} F(y;\xi_1,\dots,\xi_n) - F(y^*;\xi_1,\dots,\xi_n)\Big) \le \epsilon. \tag{4.7}$$

The previous inequality together with the convergence almost surely of $F(y^*;\xi_1,\dots,\xi_n)$ to $(Ef)(y^*)$, verifies (4.4) almost surely. □

An alternative formulation of the uniform consistency of approximations is as follows.

Proposition 4.4. *Under the conditions of Theorem 4.3, for every $\epsilon > 0$ there exists a $\delta > 0$, such that for every $\eta > 0$*

$$\mathrm{Prob}\Big\{\max_{y\in Y_\delta} F(y;\xi_1,\dots,\xi_n) - \min_{x\in X} F(x;\xi_1,\dots,\xi_n) \ge \epsilon + \eta\Big\} < \eta \tag{4.8}$$

for n large enough, with Y_δ the set of δ-minimizers of $(Ef)(\cdot)$.

Proof. Follows from (4.4) using standard arguments. □

We note that indeed in Example 4.1, the probability in equation (4.8) is equal to 1 whenever $\epsilon + \eta < 1$. Note also that the integrability condition (b) in the statement of Theorem 4.3 cannot be dropped, even if the continuity in (a) prevails. Indeed, the values $f(y,\xi)$ may be made large for y close to y^* on subsets of Ξ with small measure, along the construction of Example 4.1, such that (4.8) is violated. Finally note that conditions (a) and (b) in Theorem 4.3 actually imply that Y is a manifold of approximations, so stating this fact explicitly as an assumption could be dropped.

5. Linear Recourse

The results of the preceding sections are illustrated here in the context of the two stage linear recourse model. For an overview of the theory and applications of such models consult with Wets [14], [15], and Kall [8].

Consider the problem

$$\begin{aligned}&\text{minimize } \langle c,x\rangle + E^{P(d\xi)}(Q(x,\xi))\\ &\text{subject to } Ax = b, \qquad x \ge 0\end{aligned} \tag{5.1}$$

where $x \in \mathbb{R}^k$, the k dimensional euclidean space, A, c and b are a fixed matrix and vectors (and $\langle\cdot,\cdot\rangle$ denotes scalar multiplication), and P is a probability distribution over a space Ξ of random elements, as in previous sections. The cost $Q(x,\xi)$ is determined via a recourse procedure, and given by

$$Q(x,\xi) = \inf_z\{\langle q(\xi), z\rangle\colon\ W(\xi)z = d(\xi) + T(\xi)x,\ z \ge 0\} \tag{5.2}$$

with $z \in \mathbb{R}^\ell$ and q, W, d and T are random matrices and vectors with the appropriate dimensions; in particular $q(\cdot)$, $W(\cdot)$, $d(\cdot)$ and $T(\cdot)$ are measurable.
In terms of the problem formulation in $(*)$ we have

$$f(x,\xi) = \begin{cases} \langle c, x\rangle + Q(x,\xi) & \text{if } Ax = b,\ x \geq 0 \\ \infty & \text{otherwise.} \end{cases} \tag{5.3}$$

Note that $f(x,\xi)$ may equal $+\infty$ even if the constraints on x are met; this happens when the recourse is not feasible, i.e., for no $z \geq 0$ the equality $W(\xi)z = d(\xi)+T(\xi)z$ is satisfied. Also note that the formulation (5.3) may not fit the general framework of this paper as $Q(x,\xi)$ may be equal to $-\infty$. This happens when the recourse decision is degenerate. We refer the reader to Wets [14] and references therein, and to Kall [8] for conditions that guarantee that $Q(x,\xi) > \infty$ for almost every ξ. Here, for concreteness, we impose a condition which implies this property, as follows. We use M^T to denote the transpose of the matrix M.

Assumption 5.1. (Integrable dual feasibility) There exists a vector u such that $c \geq A^T u$. There exists a measurable function $v(\xi)$ such that $q(\xi) \geq W^T(\xi)v(\xi)$ and such that $\langle d(\xi), v(\xi)\rangle$ and $T^T(\xi)v(\xi)$ are both integrable.

Remark 5.2. The role of the integrable dual feasilibility conditions will be apparent in the proof of the following proposition. We note here that most practical examples satisfy this condition (see [14], [15], [8] and references therein). For instance, if $c \geq 0$ and $q(\xi) \geq 0$ for almost every ξ then the integrable dual feasibility holds with $u = 0$ and $v(\xi) = 0$.

Proposition 5.3. *Under Assumption* 5.1 *the linear recourse problem* (5.1)–(5.2) *satisfies Assumptions* 2.1 *and* 2.2.

Proof. The lower semicontinuity of $f(x,\xi)$ in x follows from the more general result of Wets [14, Proposition 7.5]; note that for ξ fixed the lower semicontinuity of $Q(x,\xi)$ can be verified by checking that $\min\{\langle q, z\rangle\colon\ Wz = s,\ z \geq 0\}$ is lower semicontinuous in s, for q and W fixed. The lower semicontinuity of $f(x,\xi)$ in x follows then from the closedness of the constraints $Ax = b$, $x \geq 0$.

To prove the joint measurability of $f(x,\xi)$ consider first a set-valued map, say Γ_1, which assigns to the data (x, W, d, T) the set $\{z\colon\ Wz = d + Tx\}$ if $Ax = b$ and $x \geq 0$, and the empty set otherwise. The set-valued map Γ_1 may have empty values, but it certainly has a closed graph, hence Γ_1 is measurable but moreover, the inverse $\Gamma_1^-(C)$ of a compact set C is a closed set (see e.g. Castaing and Valadier [6] for the definitions). Define now

$$\Gamma(x,\xi) = \Gamma_1(x, W(\xi), d(\xi), T(\xi)).$$

By the measurability of the data in (5.2) the set-valued map Γ is measurable jointly in (x,ξ), this since the inverse $\Gamma^-(C)$ of a compact set is the inverse of the closed set $\Gamma_1^-(C)$ by the function $(x, W(\xi), d(\xi), T(\xi))$ which is assumed to be measurable. Finally note that

$$f(x,\xi) = \min\{\langle q(\xi), z\rangle\colon\ z \in \Gamma(x,\xi)\}$$

which implies (say by using the Castaing Representation) the measurability of $f(x,\xi)$ jointly in (x,ξ) in case the underlying σ-field is complete (see [6, Chapter III]), a property that we assume throughout.
The remaining properties needed to complete the proof are that $f(x,\xi) > -\infty$ and that the integral bound in Assumption 2.2 holds. For this we employ Assumption 5.1. Indeed, $\langle c,x\rangle \geq \langle A^T u, x\rangle$, hence by the equality $Ax = b$ we get $\langle c,x\rangle \geq \langle u,b\rangle$, which gives a uniform lower bound on the cost of the primary decision. The recourse decision can be estimated as follows

$$\begin{aligned}\langle q(\xi), z\rangle &\geq \langle W^T(\xi), z\rangle\\ &= \langle v(\xi), W(\xi)z\rangle\\ &= \langle v(\xi), d(\xi) - T^T(\xi)x\rangle.\end{aligned}$$

The integrability requirement in Assumption 5.1 implies that $Q(x,\xi) > -\infty$, and furthermore, there is an integral bound as required in Assumption 2.2, locally in x. □

With the preceding proposition, the strong law of large numbers verified in Theorem 2.3 applies. An immediate consequence is that if $x^n = x^n(\xi_1,\ldots,\xi_n)$ is a solution of

$$\begin{array}{ll}\text{minimize} & \langle c,x\rangle + \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\langle q(\xi_j), z_j\rangle\\ \text{subject to} & Ax = b\\ & T(\xi_j)x + W(\xi_j)z_j = d(\xi_j) \quad j = 1,\ldots,n\\ & x \geq 0\\ & z_j \geq 0 \quad j = 1,\ldots,n\end{array} \tag{5.4}$$

then almost surely any cluster point of x^n is a solution of the stochastic problem (5.1). Note that (5.4) is a large scale linear program, depending on a stochastic parameter.

We turn now to check the consistency of minimizers and ϵ-minimizers (that could as well be computed, with large probability, using (5.4) with a large enough sample). We note that the underlying conditions of Theorem 3.3 and Propositions 3.4 and 3.5 are satisfied. Hence consistency of minimizers and ϵ-minimizers is guaranteed with respect to decisions in any compact subset of $\mathbb{R}^k$.
Next we wish to analyze the uniform consistency of approximations for the linear recourse, in line with the approach in Section 4. To this end define

$$D = \left\{x\colon\; E^{P(d\xi)}(f(x,\xi)) < \infty\right\} \tag{5.5}$$

namely D is the effective domain of the cost function. Note that

$$D = \left\{x\colon\; E^{P(d\xi)}(Q(x,\xi)) < \infty\right\} \cap \{x\colon\; Ax = b,\; x \geq 0\}. \tag{5.6}$$

The domain D may not be compact, and even its intersection with compact subsets of $\mathbb{R}^k$ may not be compact; this was noted in Walkup and Wets [13] and will be also demonstrated in an example below. The following observation is of help.

Proposition 5.4. *Let K be a compact polyhedral subset of D. Then there exists an integrable function $\beta(\xi)$: $\Xi \to [0, \infty)$, such that for all $x \in K$, the bound $f(x, \xi) \le \beta(\xi)$ holds almost everywhere.*

Proof. The compact polyhedral set K is a convex combination of a finite number of points, say $x_1, \ldots, x_r$. Let $z_j(\xi)$ be an optimal recourse decision for the point x_j, namely $W(\xi)z_j(\xi) = d(\xi) + T(\xi)x_j$, and $\langle q(\xi), z_j(\xi)\rangle = Q(x_j, \xi)$. Such a recourse exists and is measurable by the completeness of the σ-field, see Kall [8]. The linearity of the constraints implies that the convex combination

$$\alpha_1 z_1(\xi) + \cdots + \alpha_r z_r(\xi)$$

is an admissible recourse decision for the primal decision

$$\alpha_1 x_1 + \cdots + \alpha_r x_r.$$

Since the latter combinations exhaust K, it follows that $\beta_1(\xi) = \max\{\langle q(\xi), z_j(\xi)\rangle : j = 1, \ldots, r\}$ is an integrable bound of $Q(x, \xi)$ for $x \in K$, and therefore $\beta(\xi) = \beta_1(\xi) + \max\{\langle c, x\rangle : x \in K\}$ is the desired integrable bound. $\square$

We note that the preceding argument also proves that any polyhedral subset K of D is a manifold of approximations for the recourse problem, as defined in Section 4. This property also follows from the fact, proved in Wets [14], that $E^{P(d\xi)}(Q(x, \xi))$ is continuous on polyhedral subsets of D.

With the preceding arguments we can conclude that the uniform consistency of approximations to solutions of the stochastic linear recourse, holds whenever the approximations are restricted to a set contained in a polyhedral subset of D. Indeed, Proposition 5.4 implies that condition (b) of Theorem 4.3 holds, and as noted before, any polyhedral subset of D is a manifold of approximations.

It is worthwile to note that under additional conditions the entire domain D is polyhedral. For instance, if P is supported on a finite set; or when W is fixed and the data $(q(\cdot), d(\cdot), T(\cdot))$ have a finite variance and form a polyhedral set in the linaer space of (q, d, T), see [14, Theorem 4.7]; or if W and T are fixed and $(q(\cdot), d(\cdot))$ are square integrable, see [14, Theorem 4.10]. In general, D is a convex set which may not be closed, and $E^{P(d\xi)}(f(x, \xi))$ may not be continuous on D, as noted in Walkup and Wets [13]. We provide here an example which demonstrates the phenomenon, and illustrates the preceding considerations.

Example 5.5. Let the primary decision $x = (x_1, x_2)$ be two dimensional, and let the recourse decision $z = (z_1, z_2, z_3)$ be three dimensional. The random element ξ takes values in $[1, \infty)$, with an underlying probability that will be determined later.

Let the recourse cost be given by

$$Q(x, \xi) = \min\{z_1 : \ z_1 + z_2 = \xi^2 x_1, \ z_2 + z_3 = (\xi^2 - \xi)x_2, \ z \ge 0\} \tag{5.7}$$

and suppose that the goal is to minimize the expected value of the recourse, with $x \ge 0$ (namely $c = b = A = 0$).

It is easy to see that the optimal decision, regardless of the probability, is $x = 0$, and then $Q(0, \xi) = 0$ identically. For a general x, it is clear that

$$Q(x, \xi) = \max(0, \xi^2 x_1 - (\xi^2 - \xi)x_2). \tag{5.8}$$

Consider now the particular probability P on $[1,\infty)$ that assigns the value 2^{-k} to the point $\xi = 2^k$, this for $k = 1, 2, \dots$, (note that then the data in (5.7) does not have finite variance). With this probability it is easy to see that

$$D = \{(x_1, x_2)\colon\ 0 \le x_1 < x_2\} \cup \{(0,0)\}. \tag{5.9}$$

Indeed, for $x_1 \ge x_2 > 0$ we have $Q(x,\xi) = \xi^2(x_1 - x_2) + \xi x_2$ and the expectation is $+\infty$. If $x_1 < x_2$ then for $\xi > x_2(x_2 - x_1)^{-1}$ we have $Q(x,\xi) = 0$, hence the expectation is finite. The lower semicontinuity of the expected cost $E^{P(d\xi)}(Q(x,\xi))$ implies then that as decisions $x^j \in \mathbb{R}^2$ converge to $x^0 = (x_1, x_1)$ and $x_1 > 0$, then $E^{P(d\xi)}(Q(x^j,\xi))$ converge to $+\infty$. This implies that not every compact subset of D that includes $(0,0)$ is a manifold of approximations. Indeed, on points x^j that converge to 0, but get closer and closer to the diagonal $x_1 = x_2$ in $\mathbb{R}^2$, the value may converge to $+\infty$ as $j \to \infty$. (In the concrete example (5.7) it is not difficult to compute the expectation, and come up with such a sequence. In fact, on the cone in $\mathbb{R}^2_+$ determined by the relation $(1 - 2^{-\ell})x_2 \le x_1 \le (1 - 2^{-(\ell+1)})x_2$ the value can be obtained by a finite summation of $2^{-k}Q(x, 2^k)$ over $k = 1, \dots, 2^\ell$, resulting in

$$E^{P(d\xi)}(Q(x,\xi)) = (x_1 - x_2)(2^{\ell+1} - 2) + \ell x_2.$$

Hence if the sequence $x^j = (x_{1,j}, x_{2,j})$ is determined by $x_{2,j} = j^{-1/2}$, $x_{1,j} = (1 - 2^{-j})x_{2,j}$, the expectations tend to $+\infty$ as $x^j \to 0$.)

A conclusion of the theory, however, is that on any polyhedral subset of (5.9), in particular on a set $\{(x_1, x_2)\colon\ 0 \le x_1 \le (1-\delta)x_2 \le \Delta\}$, with $\delta > 0$, the consistency of the minimizer $x = 0$, and the uniform consistency of approximations of $x = 0$, hold.

References

[1] R.B. Ash: Real Analysis and Probability. Academic Press, New York, 1972.

[2] H. Attouch: Variational Convergence of Functions and Operators. Pitman, London, 1984.

[3] H. Attouch, R.J-B Wets: Epigraphical processes: laws of large numbers for lsc functions. To appear.

[4] G. Beer: A geometric algorithm for approximating semicontinuous functions. J. Approximation Theory 49 (1987), 31–40.

[5] C. Castaing, F. Ezzaki: Variational inequalities for integrands martingales and additive random sequences. Seminaire d'Analyse Convexe Montpellier 22 (1992), n^o 1.

[6] C. Castaing, M. Valadier: Convex Analysis and Measurable Multifunctions. Lecture Notes in Mathematics 580, Springer-Verlag, Berlin, 1977.

[7] C. Hess: Epi-convergence of sequences of normal integrands and strong consistency of the maximum likelihood estimator. In: XXIV^{es} Journées de Statistique, Bruxelles (18–22 May, 1992), pp. 254–256; also Cahiers de Mathématiques de la Decision n^o 9121, Université Paris Dauphine (1991).

[8] P. Kall: Stochastic Linear Programming. Springer-Verlag, Berlin, 1976.

[9] A. King, R.J-B Wets: Epi-consistency of convex stochastic programs. Stochastics and Stochastics Reports 34 (1991), 83–92; erratum in 37 (1991) 259–260.

[10] D. Plachky: A version of the slln universal under mappings. Research Memorandum, Institute of Mathematical Statistics, University of Münster.

[11] R.T. Rockafellar: Integral functionals, normal integrands and measurable selections. In: Nonlinear Operators and the Calculus of Variations (J. Gossez and L. Waclbroek, eds.), Lecture Notes in Mathematics 543, Springer-Verlag, Berlin, 1976, pp. 157–207.

[12] R.T. Rockafellar, R.J-B Wets: Variational analysis: an introduction. In: Multifunctions and Integrands Stochastic Analysis and Optimization (G. Salinetti, ed.), Lecture Notes in Mathematics 1091, Springer-Verlag, Berlin, 1984, pp. 1–54.

[13] D.W. Walkup, R.J-B Wets: Stochastic programs with recourse. SIAM J. Appl. Math. 15 (1967), 1299–1314.

[14] R.J-B Wets: Stochastic programs with fixed recourse: The equivalent deterministic program. SIAM Review 16 (1974), 309–339.

[15] R.J-B Wets: Stochastic programming. In: Handbook for Operations Research and Management Science, Vol. 1 (G. Nemhanser and A. Rinooy Kan, eds.), Elsevier Science Publishers, Amsterdam, 1989, pp. 573–629.

HIER :

Leere Seite
18