On convergence rates of approximations in stochastic programming

Werner Römisch

Abstract:
Stochastic programming problems are viewed as parametric programs with respect to the involved probability measure μ. A general result about continuity properties of expectation functionals w.r.t. the Prokhorov metric in the space of all Borel probability measures is proved and used to obtain a quantitative continuity result for the optimal value of stochastic linear programs with complete fixed recourse w.r.t. μ. This result provides "convergence rates" (of the optimal values) for approximations by discrete probability measures, e.g. via conditional expectations.

1. Introduction

Let us consider the following class of stochastic programming problems:

$$\min \{ E[g(z(\omega),x)] \mid x \in \mathbb{R}^m, \ P[\{\omega \mid x \in X(z(\omega))\}] \geq \alpha \} \quad (1.1)$$

where z is a random variable (defined on some probability space $(\Omega,\mathcal{A},P)$, E denoting the expectation with respect to P) with values in $Z \subseteq \mathbb{R}^r$, g is a mapping from $Z \times \mathbb{R}^m$ into $\mathbb{R}$, X is a multifunction from Z into $\mathbb{R}^m$ and $\alpha \in [0,1]$, or, equivalently,

$$\min \{ \int_Z g(z,x)\, d\mu \mid x \in \mathbb{R}^m, \ \mu[\{z \in Z \mid x \in X(z)\}] \geq \alpha \} \quad (1.2)$$

where μ is the probability distribution of $z(\cdot)$, i.e., μ is a Borel probability measure on Z.
Note that a number of stochastic programs with recourse and (or) with chance constraints fit into this class (see [7]).

A well-known approach for solving (1.2) approximately consists in the use of discrete approximations, i.e., in approximating the probability distribution μ by discrete probability measures μ_n , $n \in \mathbb{N}$ (see e.g. [2], [6], [8], [9], [11], [13], [15], [18], [23], [24]). Usually, it is assumed that the sequence (μ_n) converges weakly to μ in the space $\mathcal{P}(Z)$ of all Borel probability measures on Z (see [1]). A theoretical basis

of this approach is provided by results about the convergence of the optimal values and optimal solutions of the approximate problems (containing μ_n, $n \in \mathbb{N}$) to those of the original problem (1.2) (using various concepts).
In this context, it is useful to note that (even if Z is a separable metric space) $\mathcal{P}(Z)$ equipped with the topology of weak convergence is metrizable, e.g. by the Prokhorov metric

$$\rho(\mu,\nu) := \inf\{\varepsilon > 0 \mid \mu(B) \leq \nu(B^{\varepsilon}) + \varepsilon \text{ for all Borel } B \subseteq Z\},$$

where $B^{\varepsilon} := \{z \in Z \mid \inf\{d(z,\tilde{z}) \mid \tilde{z} \in B\} < \varepsilon\}$, d denoting the metric on Z (see [1], [3], [14]).
Hence, (1.2) can be viewed as parametric optimization problem with parameter μ varying in the metric space $\mathcal{P}(Z)$. In particular, quantitative stability results of parametric programming (see e.g. [10]) may be applied to (1.2).
This motivates our interest in quantitative continuity results (at least) for the optimal value function $\varphi(\mu)$ (of (1.2)) w.r.t. μ in (subsets of) $\mathcal{P}(Z)$ and suggests a possible approach to obtaining estimates for $|\varphi(\mu) - \varphi(\mu_n)|$ in terms of $\rho(\mu,\mu_n)$, i.e., to "convergence rates" of $(|\varphi(\mu) - \varphi(\mu_n)|)$.
As a first step in this direction, we consider the case that $X(z) \equiv X_o$, where X_o is a subset of $\mathbb{R}^m$, and, hence, are interested in the quantitative continuity of

$$\varphi(\mu) := \inf\left\{ \int_Z g(z,x)\, d\mu \;\middle|\; x \in X_o \right\} \tag{1.3}$$

w.r.t. $\mu \in \mathcal{P}(Z)$. (The general problem (1.2) will be dealt with in [17].) In the spirit of [16] and [5], we obtain in Section 2 an estimate for

$$\left| \int_Z f(z)\, d(\mu - \nu) \right|$$

where μ and ν are probability measures on Z having certain finite moments, in terms of $\rho(\mu,\nu)$, the growth of the local Lipschitz constants of f, a tail estimate of μ, and the moments of μ and ν. We use this estimate for the study of continuity of (1.3) if, for all $x \in \mathbb{R}^m$, $g(\cdot,x)$ is locally Lipschitz continuous with Lipschitz constants which are independent of x, and apply this result to stochastic linear programming problems with complete fixed recourse (Theorem 5).

The latter result is related to those contained in [4], [20] (where the stability of solutions to stochastic programming problems with recourse with respect to parameters of the given probability distribution is studied) and in [22](where the case of empirical measures is considered and a version of the central limit theorem for the optimal solutions is given). Note that in [4], [20] and [22] the authors impose strong differentiability conditions on g and regularity conditions on μ .

2. Continuity properties of expectation functionals with respect to probability measures

Let Z be a separable metric space (with metric d) and $\mathcal{P}(Z)$ the space of all Borel probability measures on Z endowed with the Prokhorov metric ρ .
Let f be a mapping from Z into $\mathbb{R}$ such that

$$|f(z) - f(\tilde{z})| \leq L(\max\{d(z,0),\ d(\tilde{z},0)\})\ d(z,\tilde{z}) \quad (z,\tilde{z} \in Z) \qquad (2.1)$$

where $L: \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is a right-continuous monotonically increasing function and $0 \in Z$ is some distinguished element. For the following, we introduce some notations:

$$L_1(t) := L(t)t \quad , \text{ for all } t \in \mathbb{R}_+,$$

$$M_p(\mu) := \left[\int_Z (L_1(d(z,0)))^p \, d\mu \right]^{\frac{1}{p}} \quad (\mu \in \mathcal{P}(Z),\ 1 \leq p < +\infty),$$

$$\mathcal{P}_{L,p}(Z) := \{ \mu \in \mathcal{P}(Z) \mid M_p(\mu) < +\infty \} \quad (1 \leq p < +\infty).$$

Note that $\int_Z f(z)\, d\mu$ is finite if $\mu \in \mathcal{P}_{L,1}(Z)$ and moreover, if $\mu \in \mathcal{P}_{L,p}(Z)$, it follows that

$$\left[\int_Z |f(z)|^p d\mu \right]^{\frac{1}{p}} \leq |f(0)| + M_p(\mu) \quad .$$

Now, we ask for estimates of the form

$$\left| \int_Z f(z)\ d(\mu - \nu) \right| \leq \psi_f(\mu,\nu) \quad (\mu\ ,\ \nu \in \mathcal{P}_{L,p}(Z)) \qquad (2.2)$$

where $\psi_f(\mu,\nu)$ tends to zero if $\rho(\mu,\nu) \longrightarrow 0$.

Theorem 1:

Let $\mu, \upsilon \in \mathcal{P}_{L,p}(Z)$ for some $p \in]1, +\infty[$. Then

$$\left|\int_Z f(z)\, d(\mu - \upsilon)\right| \leq (C + 1)[\max\{2, L(h(\rho(\mu,\upsilon)) + \rho(\mu,\upsilon))\} * \rho(\mu,\upsilon)]^{1 - \frac{1}{p}} \quad (2.3)$$

where $C := 2|f(0)| + M_p(\mu) + M_p(\upsilon)$,

$$h(\delta) := \inf\{K > 0 \mid \mu(\{z \in Z \mid d(z,0) > K\}) \leq \delta\} \quad (\delta \in [0,1]).$$

Proof:

We consider the mapping $F: \mathcal{P}(Z) \longrightarrow \mathcal{P}(\mathbb{R})$ defined by $F(\mu)(B) := \mu(f^{-1}(B))$ for all Borel sets $B \subseteq \mathbb{R}$ and $\mu \in \mathcal{P}(Z)$. Let $\tilde{\rho}$ denote the Prokhorov metric in $\mathcal{P}(\mathbb{R})$. Then [5, Theorem 1] provides the estimate

$$\tilde{\rho}(F(\mu), F(\upsilon)) \leq \max\{2, L(h(\rho(\mu,\upsilon)) + \rho(\mu,\upsilon))\}\rho(\mu,\upsilon) . \quad (2.4)$$

By Strassen's theorem ([19, p. 438]) there exist real random variables ξ_1 and ξ_2 (defined on some probability space $(\Omega, \mathcal{A}, P)$) such that $F(\mu)$ and $F(\upsilon)$ are the probability distributions of ξ_1 and ξ_2, respectively, and that

$$\tilde{\rho}(F(\mu), F(\upsilon)) = \inf\{\alpha > 0 \mid P[\{\omega \mid |\xi_1(\omega) - \xi_2(\omega)| > \alpha\}] \leq \alpha\}. \quad (2.5)$$

Now, let $\alpha \in]0,1]$ be such that $P(A_\alpha) \leq \alpha$, where $A_\alpha := \{\omega \mid |\xi_1(\omega) - \xi_2(\omega)| > \alpha\}$. Then

$$\left|\int_Z f(z)\, d(\mu - \upsilon)\right| = \left|\int_{\mathbb{R}} t\, d(F(\mu) - F(\upsilon))\right| =$$

$$= \left|\int_\Omega (\xi_1(\omega) - \xi_2(\omega))\, dP\right| \leq \alpha + \int_{A_\alpha} |\xi_1(\omega) - \xi_2(\omega)|\, dP$$

$$\leq \alpha + P(A_\alpha)^{\frac{1}{q}} \left(\int_\Omega |\xi_1(\omega) - \xi_2(\omega)|^p\, dP\right)^{\frac{1}{p}}, \text{ where } \frac{1}{p} + \frac{1}{q} = 1,$$

$$\leq \alpha + \alpha^{\frac{1}{q}} \left[\left(\int_Z |f(z)|^p\, d\mu\right)^{\frac{1}{p}} + \left(\int_Z |f(z)|^p\, d\upsilon\right)^{\frac{1}{p}}\right]$$

$$\leq \alpha + C\alpha^{\frac{1}{q}} \leq (C + 1)\alpha^{1 - \frac{1}{p}} \quad (2.6)$$

(where we used Hölder's and Minkovsky's inequality, respectively). Now, taking the infimum over all $\alpha \in]0,1]$ with $P(A_\alpha) \leq \alpha$, (2.6), (2.5) and (2.4) imply (2.3). $\square$

Corollary 2:

Let $\mu, \upsilon \in \mathcal{P}_{L,p}(Z)$ for some $p \in]1, +\infty[$. Then (2.3) is valid with $h(\delta) := \inf\{ K > 0 \mid L(K)K = M_p(\mu)\delta^{-\frac{1}{p}}\}$ ($\delta \in [0,1]$).

Proof:

Since $\mu \in \mathcal{P}_{L,p}(Z)$, a Chebyshev-type estimate implies

$$\mu(\{ z \in Z \mid d(z,0) > K \}) \leq \frac{\int_Z (L_1(d(z,0)))^p \, d\mu}{(L_1(K))^p} = \left[\frac{M_p(\mu)}{L_1(K)}\right]^p$$

Thus, the assertion follows from Theorem 1. □

Corollary 3:

Let $L(t) := L_o t$, $t \in \mathbb{R}_+$ ($L_o > 0$), and $\mu, \upsilon \in \mathcal{P}(Z)$ be such that $\int_Z (d(z,0))^{2p} \, d\mu$ and $\int_Z (d(z,0))^{2p} \, d\upsilon$ are finite for some $p \in]1, +\infty[$. Then

$$\left|\int_Z f(z) \, d(\mu - \upsilon)\right| \leq (C+1)\max\{2, L_o + (L_o M_p(\mu))^{\frac{1}{2}}\} \varrho(\mu,\upsilon)^{r(p)} \quad (2.7)$$

where C is defined as in Theorem 1 and $r(p) := (1 - \frac{1}{2p})(1 - \frac{1}{p})$.

Proof:

We apply Corollary 2 and obtain

$$\left|\int_Z f(z) \, d(\mu - \upsilon)\right| \leq (C+1)\left[\max\{2, L_o((\frac{M_p(\mu)}{L_o})^{\frac{1}{2}} \varrho(\mu,\upsilon)^{-\frac{1}{2p}} + \varrho(\mu,\upsilon))\} * \right.$$

$$\left. * \varrho(\mu,\upsilon)\right]^{1 - \frac{1}{p}}$$

$$\leq (C+1)\left[\max\{2, L_o + (L_o M_p(\mu))^{\frac{1}{2}}\} \varrho(\mu,\upsilon)^{1 - \frac{1}{2p}}\right]^{1 - \frac{1}{p}}$$

and, hence, the estimate (2.7). □

Remark 4:

Note that Theorem 1 and its Corollaries give estimates of the type (2.2) at least on subsets of $\mathcal{P}_{L,p}(Z)$ where $M_p(\mu)$ is uniformly bounded.

Let g be a continuous mapping from $Z \times \mathbb{R}^m$, $Z \subseteq \mathbb{R}^r$, into $\mathbb{R}$ such that

$$|g(z,x) - g(\tilde{z},x)| \leq L(\max\{\|z\|, \|\tilde{z}\|\}) \, \|z - \tilde{z}\| \quad (z, \tilde{z} \in Z, \ x \in \mathbb{R}^m),$$

where $L: \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is a right-continuous monotonically increasing function and $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^r$. Let X_o be a compact subset of $\mathbb{R}^m$ and $\mu, \nu \in \mathcal{P}_{L,1}(Z)$. Then $\varphi(\mu)$ and $\varphi(\nu)$ (defined by (1.3)) are finite and there exists an $\bar{x} \in X_o$ such that

$$|\varphi(\mu) - \varphi(\nu)| \leq \left| \int_Z g(z,\bar{x}) \, d(\mu - \nu) \right| .$$

Hence, the above results can be used to estimate $|\varphi(\mu)-\varphi(\nu)|$.

3. An application to stochastic programs with complete fixed recourse

In this Section we give an application of the results obtained in Section 2 to stochastic linear programming problems with complete fixed recourse (see [7], [21] for excellent surveys). We consider the following problem:

$$\min \left\{ \int_Z g(z,x) \, d\mu \;\middle|\; x \in X_o \right\} , \tag{3.1}$$

where X_o is a compact subset of $\mathbb{R}^m$, and

$$g(z,x) := c^T x + \eta(a, b-Ax) \;, \quad z=(a,b,c,A) \in Z, \; x \in \mathbb{R}^m \tag{3.2}$$

$$Z := \{(a,b,c,A) \mid a \in \mathbb{R}^{\bar{n}}, \; b \in \mathbb{R}^d, \; c \in \mathbb{R}^m, \; A \in L(\mathbb{R}^m, \mathbb{R}^d), \; \{u \in \mathbb{R}^d \mid W^T u \leq a\} \neq \emptyset \},$$

$$\eta(a,v) := \inf \{a^T y \mid y \in \mathbb{R}^{\bar{n}}, \; Wy = v, \; y \geq 0\}, \; a \in \mathbb{R}^{\bar{n}}, \; v \in \mathbb{R}^d,$$

$W \in L(\mathbb{R}^{\bar{n}}, \mathbb{R}^d)$ is the "fixed recourse" matrix and μ is a given Borel probability measure on Z.

We assume that the stochastic program has complete (fixed) recourse, i.e., that for all $v \in \mathbb{R}^d$ the set $\{y \in \mathbb{R}^{\bar{n}} \mid Wy = v, \; y \geq 0\}$ is nonempty. Then $\eta(a,v) < +\infty$ for all $a \in \mathbb{R}^{\bar{n}}$, $v \in \mathbb{R}^d$. Hence, due to the duality theorem of linear programming $\eta(a,v)$ is finite for all $(a,v) \in \mathfrak{A} := \{a \in \mathbb{R}^{\bar{n}} \mid \{u \in \mathbb{R}^d \mid W^T u \leq a\} \neq \emptyset\} \times \mathbb{R}^d$, and

$$g(z,x) \in \mathbb{R} \quad \text{for all} \quad (z,x) \in Z \times \mathbb{R}^m.$$

Thus, (3.1) fits into the setting of the preceding Sections (note that Z is a closed (convex) subset of a finite-dimensional space and, hence, is a complete separable metric space). Using a well-known result of linear parametric programming ([12, Satz 8.8, p.219]) about the continuity and structure of the function η (on $\mathfrak{A}$), it can be shown (see [17] for details) that g is continuous and that there exists a constant $L_o > 0$ such that

$$|g(z,x)-g(\tilde{z},x)| \leq L_o \max\{\|z\|, \|\tilde{z}\|\}\|z - \tilde{z}\| \quad (z,\tilde{z} \in Z,\ x \in \mathbb{R}^m)$$

(where $\|\cdot\|$ is a norm on $\mathbb{R}^{\bar{n}} \times \mathbb{R}^d \times \mathbb{R}^m \times L(\mathbb{R}^m, \mathbb{R}^d)$).
Hence we can apply Corollary 3 (see Remark 4) and obtain

Theorem 5:

Let $\mu, \mu_n \in \mathcal{P}(Z)$, $n \in \mathbb{N}$, be such that $\int_Z \|z\|^{2p} d\mu$ and $\int_Z \|z\|^{2p} d\mu_n$, $n \in \mathbb{N}$, are finite for some $p \in]1, +\infty[$.
Then there is a constant $C_o > 0$ such that

$$|\varphi(\mu) - \varphi(\mu_n)| \leq C_o(1 + (\int_Z \|z\|^{2p} d\mu_n)^{\frac{1}{p}})\ \rho(\mu,\mu_n)^{r(p)} \quad (n \in \mathbb{N}) \quad (3.3)$$

where $\varphi(\mu)$ (and $\varphi(\mu_n)$, respectively) are defined by (1.3), C_o depends only on μ, g and p, and $r(p) := (1 - \frac{1}{2p})(1 - \frac{1}{p})$.

Proof:

With $f(z) := g(z,\bar{x})$, $z \in Z$, for some $\bar{x} \in X_o$ (see Remark 4), and $L(t) := L_o t$, $t \in \mathbb{R}_+$, we apply Corollary 3. □

A case of particular interest is that of discrete approximations for μ via conditional expectations (see [2, Sect. 4], [8], [9], [23]).

Corollary 6:

Let $\{Z_{nl}, l=1,\dots,n\}$ be a partition of Z into Borel sets and let μ_n be defined by

$$\mu_n(B) := \sum_{\substack{l=1 \\ m_{nl} \in B}}^{n} \mu(Z_{nl}), \quad \text{where} \quad m_{nl} := \frac{1}{\mu(Z_{nl})} \int_{Z_{nl}} z\, d\mu, \quad (3.4)$$

for all Borel subsets B of Z and all $n \in \mathbb{N}$.
Then $|\varphi(\mu) - \varphi(\mu_n)| \leq C_1\, \rho(\mu,\mu_n)^{r(p)}$, $r(p) := (1 - \frac{1}{2p})(1 - \frac{1}{p})$, for some constant C_1 depending only on μ, g, p.

Proof:

From Jensen's inequality we get $\int_Z \|z\|^{2p} d\mu_n \leq \int_Z \|z\|^{2p} d\mu$ for all $n \in \mathbb{N}$. Hence we can apply Theorem 5. $\square$

Remark 7:

By martingale convergence (μ_n) (defined by (3.4)) converges weakly to μ if $S_n := \{Z_{nl}, l=1,\dots,n\}$ is a refinement of S_{n-1}, i.e., $S_{n-1} \subset S_n$ $(n \geq 2)$, and if the σ-algebra generated by the algebra $\bigcup_{n \in \mathbb{N}} S_n$ contains all Borel subsets of Z.

If, in Corollary 6, $\int_Z \|z\|^p d\mu$ is finite for all $p \in]1,+\infty[$, then $|\varphi(\mu) - \varphi(\mu_n)| = O(\rho(\mu,\mu_n)^r)$ for all $r<1$.

Estimates for $|\varphi(\mu) - \varphi(\mu_n)|$ (for stochastic programs (3.1)) are also given e.g. in [6], [8], [11], but there in terms of the L_p-distance of the underlying random variables.

Acknowledgement: The author thanks Diethard Klatte (Humboldt-Universität Berlin and PH Halle) for valuable discussions.

References:

[1] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.

[2] J. Birge and R. Wets, Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse, IIASA Working Paper WP-83-111 (1983), Laxenburg, Austria and Mathematical Programming Study (to appear).

[3] R.M. Dudley, Distances of probability measures and random variables, Ann. Math. Statist. 39 (1968), 1563-1572.

[4] J. Dupačová, Stability in stochastic programming with recourse-estimated parameters, Math. Programming 28 (1984), 72-83.

[5] H.W. Engl and A. Wakolbinger, Continuity properties of the extension of a locally Lipschitz continuous map to the space of probability measures, Monatsh. Math. (to appear).

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

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

[8] P. Kall and D. Stoyan, Solving stochastic programming problems with recourse including error bounds, Math. Operationsforsch. Statist., Ser. Optimization, 13 (1982), 431-447.

[9] P. Kall, K. Frauendorfer and A. Ruszczyński, Approximation techniques in stochastic programming, Manuskript, Institut f. Operations Research, Universität Zürich, 1984.

[10] D. Klatte, Beiträge zur Stabilitätsanalyse nichtlinearer Optimierungsprobleme, Dissertation (B), Humboldt-Universität Berlin, Sektion Mathematik, 1984.

[11] K. Marti, Approximationen stochastischer Optimierungsprobleme, Verlag Anton Hain, Königstein, 1979.

[12] F. Nožička, J. Guddat, H. Hollatz and B. Bank, Theorie der linearen parametrischen Optimierung, Akademie-Verlag, Berlin, 1974.

[13] P. Olsen, Discretizations of multistage stochastic programming problems, Mathematical Programming Study 6 (1976), 111-124.

[14] Yu.V. Prokhorov, Convergence of random processes and limit theorems in probability theory, Teor. Verojatnost. i Primenen. 1 (1956), 177-238.

[15] W. Römisch, On discrete approximations in stochastic programming, Proc. 13. Jahrestagung Mathematische Optimierung, Vitte 1981, Humboldt-Univ. Berlin, Sekt. Mathematik, Seminarbericht Nr. 39 (1981), 166-175.

[16] W. Römisch and A. Wakolbinger, On Lipschitz dependence in systems with differentiated inputs, Math. Ann. 272 (1985), 237-248.

[17] W. Römisch and A. Wakolbinger, Obtaining convergence rates for approximations in stochastic programming (in preparation).

[18] G. Salinetti, Approximations for chance-constrained programming problems, Stochastics 10 (1983), 157-179.

[19] V. Strassen, The existence of probability measures with given marginals, Ann. Math. Statist. 36 (1965), 423-439.

[20] J. Wang, Distribution sensitivity analysis for stochastic programs with complete recourse, Math. Programming 31 (1985), 286-297.

[21] R. Wets, Stochastic programs with fixed recourse: the equivalent deterministic program, SIAM Review 16 (1974), 309-339.

[22] R. Wets, A statistical approach to the solution of stochastic programs with (convex) simple recourse, Preprint, University of Kentucky, Lexington, 1980.

[23] R. Wets, Stochastic programming: solution techniques and approximation schemes, in: Mathematical Programming: The State-of-the-Art 1982, Springer-Verlag, Berlin, 1983, 566-603.

[24] R. Wets, Numerical Challenges in Stochastic Optimization, Lecture at the International Conference on Stochastic Optimization, Kiev, Sept. 1984.

Author's address: Werner Römisch

Humboldt-Universität Berlin
Sektion Mathematik
DDR-1086 Berlin, PSF 1297