Stochastic Programming Problems with Recourse via Empirical Estimates

Vlasta Kaňková

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic
kankova@utia.cas.cz

1 Introduction

Let $\xi := \xi(\omega)\,(s \times 1)$ be a random vector defined on a probability space $(\Omega,\, \mathcal{S},\, P)$; $F,\, P_F$ the distribution function and the probability measure corresponding to the random vector ξ. Let, moreover, $g_0(x,\, z)$, $g_0^1(y,\, z)$ be functions defined on $R^n \times R^s$ and $R^{n_1} \times R^s$; $f_i(x,\, z)$, $g_i(y)$, $i = 1, \ldots, m$ functions defined on $R^n \times R^s$ and R^{n_1}; $h := h(z)\,(m \times 1)$ a vector function defined on R^s, $h^{'}(z) = (h_1(z), \ldots, h_m(z))$; $X \subset R^n$, $Y \subset R^{n_1}$ be nonempty sets. Symbols $x\,(n \times 1)$, $y := y\,(x,\, \xi)\,(n_1 \times 1)$ denote decision vectors. (R^n denotes the n–dimensional Euclidean space, $h^{'}$ a transposition of the vector function h.)

Stochastic programming problems with recourse (in a rather general setting) can be introduced as the following problem:

Find

$$\varphi(F) = \min_{x \in X} \mathsf{E}_F \{ g_0(x,\, \xi) + \min_{\{y \in Y:\, g_i(y) \le h_i(\xi) - f_i(x,\xi),\, i=1,\ldots,m\}} g_0^1(y,\, \xi)\}, \tag{1}$$

where E_F denotes the operator of mathematical expectation corresponding to F.

A special case of the problem (1) is a stochastic programming problem with linear recourse, where $Y = R^{n_1}$ and, furthermore,

$$\varphi(F) = \min_{x \in X} \mathsf{E}_F \{ g_0(x,\, \xi) + \min_{\{y \in R^{n_1}:\, Wy = h - Tx,\, y \ge 0\}} q^{'} y\} \tag{2}$$

with $q := q(\xi)\,(n_1 \times 1)$, $T := T(\xi)\,(m \times n)$, $W := W(\xi)\,(m \times n_1)$, $m \leq n_1$, $m \leq n$ (generally) random vectors and matrices.

If we denote

$$\begin{aligned} Q(x,\, \xi) \, &= \min_{\{y \in Y:\, g_i(y) \,\leq\, h_i(\xi) - f_i(x,\, \xi),\, i=1,\, \ldots,\, m\}} g_0^1(y,\, \xi)\}, \\ f_0(x,\, \xi) &= g_0(x,\, \xi) + Q(x,\, \xi), \end{aligned} \tag{3}$$

then evidently the problem (1) is covered by a more general problem:

Find

$$\varphi(F) = \inf\{\mathsf{E}_F f_0(x,\, \xi) | x \in X\}, \tag{4}$$

with $f_0(x,\, z)$ arbitrary real valued function defined on $R^n \times R^s$.

In applications very often the “underlying” distribution function F has to be replaced by an empirical distribution function F^N. Evidently, then the solution is sought with respect to the “empirical” problem:

Find

$$\varphi(F^N) = \inf\{\mathsf{E}_{F^N} f_0(x,\, \xi) | x \in X\}. \tag{5}$$

If $\mathcal{X}(F)$, $\mathcal{X}(F^N)$ denote the optimal solution sets of the problems (1) and (5), then under rather general assumptions $\varphi(F^N)$, $\mathcal{X}(F^N)$ are “good” stochastic estimates of $\varphi(F)$, $\mathcal{X}(F)$ (see e.g. [1], [4], [5], [12], [13]). There were introduced assumptions guaranteing the consistency, asymptotic normality and convergence rate. Especially, it means in the last case that

$$P\{\omega : N^{\beta} |\varphi(F) - \varphi(F^N)| > t\} \longrightarrow_{(N \longrightarrow \infty)} 0 \quad \text{for} \quad t > 0,\, \beta \in (0,\, \frac{1}{2}). \tag{6}$$

To obtain the relation (6), the Hoeffding inequality (see e.g. [2], [5]), large deviation (see e.g. [4]), Talagrand approach (see e.g. [10]) and the stability results (see e.g. [11]) have been employed. To obtain new assertions, we employ stability results [8] based on the Wasserstein metric determined by $\mathcal{L}_1$ norm in R^s. Consequently, our results are based on the assumption of thin tails of one–dimensional marginal distribution functions $F_i(z)$, $i = 1,\, \ldots,\, s$ corresponding to $F(z)$.

2 Some Auxiliary Assertions

Let $\mathcal{P}(R^s)$ denote the set of all Borel probability measures on R^s, $s \geq 1$; $\mathcal{M}_1(R^s) = \{P \in \mathcal{P}(R^s) : \int\limits_{R^s} \|z\|_s^1 P(dz) < \infty\}$, $\| \cdot \|_s^1$ the $\mathcal{L}_1$ norm in R^s. First, we recall a little generalized result of [7].

Proposition 1. Let G be an arbitrary s–dimensional distribution function such that $P_G \in \mathcal{M}_1(R^s)$. Let, moreover, $P_F \in \mathcal{M}_1(R^s)$, $f_0(x, z)$ be defined on $R^n \times R^s$. If for every $x \in X$, $f_0(x,z)$ is a Lipschitz function of $z \in R^s$ with the Lipschitz constant $L(x)$ (corresponding to $\mathcal{L}_1$ norm), then

$$|\mathsf{E}_F f_0(x,\xi) - \mathsf{E}_G f_0(x,\xi)| \le L(x) \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i \quad \text{for} \quad x \in X.$$

(Symbols F_i, G_i, $i = 1, \ldots, s$ denote one–dimensional distribution functions corresponding to F, G.)

Evidently, Proposition 1 reduces (from the mathematical point of view) stability results considered with respect to s–dimensional distribution functions to one–dimensional case. The next assertion has been proven in [8].

Proposition 2. Let $s = 1$, $t > 0$, $\bar{R} > 0$. If

1. P_F is absolutely continuous with respect to the Lebesgue measure on R^1,
2. there exists $\psi(N, t) := \psi(N, t, \bar{R})$ such that the empirical distribution function F^N fulfils for $N = 1, 2, \ldots$ the relation

$$P\{\omega : |F(z) - F^N(z)| > t\} \le \psi(N, t) \quad \text{for every} \quad z \in (-\bar{R}, \bar{R}),$$

then for $\frac{t}{4\bar{R}} < 1$, $N = 1, 2, \ldots$ it holds that

$$\begin{aligned} &P\{\omega : \int_{-\infty}^{\infty} |F(z) - F^N(z)| dz > t\} \le \\ &(\tfrac{12\bar{R}}{t} + 1)\psi(N, \tfrac{t}{12\bar{R}}, \bar{R}) + P\{\omega : \int_{-\infty}^{-\bar{R}} F(z) dz > \tfrac{t}{3}\} + \\ &P\{\omega : \int_{\bar{R}}^{\infty} (1 - F(z)) dz > \tfrac{t}{3}\} + 2NF(-\bar{R}) + 2N(1 - F(\bar{R})). \end{aligned} \tag{7}$$

To recall the next auxiliary assertion (proven in [9]), let $\bar{\xi}$, $\bar{\eta}$ be random values defined on $(\Omega, \mathcal{S}, P)$. We denote by $F_{(\bar{\xi}, \bar{\eta})}$, $F_{\bar{\xi}}$, $F_{\bar{\eta}}$ the distribution functions of the random vector $(\bar{\xi}, \bar{\eta})$ and marginal distribution functions of $\bar{\xi}$ and $\bar{\eta}$.

Lemma. Let $\bar{\zeta} = \bar{\xi}\,\bar{\eta} := \bar{\xi}(\omega)\bar{\eta}(\omega)$, $F_{\bar{\zeta}}$ denote the distribution function of $\bar{\zeta}$. If

1. $P_{F_{\bar{\xi}}}$, $P_{F_{\bar{\eta}}}$ are absolutely continuous with respect to the Lebesgue measure on R^1 (we denote by $f_{\bar{\xi}}$, $f_{\bar{\eta}}$ the probability densities corresponding to $F_{\bar{\xi}}$, $F_{\bar{\eta}}$),
2. there exist constants $C_1^{\bar{\xi}}$, $C_2^{\bar{\xi}}$, $C_1^{\bar{\eta}}$, $C_2^{\bar{\eta}} > 0$ and $T^{'} > 0$ such that

$$f_{\bar{\xi}}(z) \le C_1^{\bar{\xi}} \exp\{-C_2^{\bar{\xi}}|z|\} \quad \text{for} \quad z \in (-\infty, \ -T^{'}) \bigcup (T^{'}, \ \infty),$$
$$f_{\bar{\eta}}(z) \le C_1^{\bar{\eta}} \exp\{-C_2^{\bar{\eta}}|z|\} \quad \text{for} \quad z \in (-\infty, \ -T^{'}) \bigcup (T^{'}, \ \infty),$$

then, there exist constants $C_1^{\bar{\zeta}}$, $C_2^{\bar{\zeta}} > 0$, $\bar{T} > 1$ such that for $z > \bar{T}$

$$F_{\bar{\zeta}}(-z) \le \frac{C_1^{\bar{\zeta}}}{C_2^{\bar{\zeta}}} \exp\{-C_2^{\bar{\zeta}}\sqrt{z}\}, \quad (1 - F_{\bar{\zeta}}(z)) \le \frac{C_1^{\bar{\zeta}}}{C_2^{\bar{\zeta}}} \exp\{-C_2^{\bar{\zeta}}\sqrt{z}\}.$$

3 Convergence Rate

Let $\{\xi^i\}_{i=1}^{\infty}$ be a sequence of independent s–dimensional random vectors with a common distribution function F, F^N be determined by $\{\xi^i\}_{i=1}^{N}$.

3.1 General Case

Theorem 1. [8] Let $t > 0$, X be a compact set. If

1. P_{F_i}, $i = 1, \ldots, s$ are absolutely continuous with respect to the Lebesgue measure on R^1 (we denote by f_i, $i = 1, \ldots, s$ the probability densities corresponding to F_i),
2. there exist constants $C_1, C_2 > 0$ and $T > 0$ such that for $i = 1, \ldots, s$

$$f_i(z_i) \le C_1 \exp\{-C_2|z_i|\} \quad \text{for} \quad z_i \in (-\infty, \ -T) \bigcup (T, \ \infty),$$

3. $f_0(x, z)$ (defined by the relation (3)) is a uniformly continuous, Lipschitz (with respect to $\mathcal{L}_1$ norm) function of $z \in R^s$, the Lipschitz constant L is not depending on $x \in X$,

then

$$P\{\omega : \ N^{\beta}|\varphi(F^N) - \varphi(F)| > t\} \longrightarrow_{(N \longrightarrow \infty)} 0 \quad \text{for } \beta \in (0, \ \frac{1}{2}). \qquad (8)$$

Remarks.

1. Some cases, under which $f_0(x, z)$ (defined by (3)) fulfils the assumption 3 of Theorem 1, are introduced e.g. in [6].

2. If $Q(x, z)$ corresponds to the case (2) (with q and simultaneously with at least one of h or T random), then evidently, the assumption 3 of Theorem 1 has not to be fulfilled (for more details see e.g. [3]).

3.2 Stochastic Programming Problems with Linear Recourse

Considering the linear case (2), we assume:

A.1 a. W is a deterministic matrix,
b. W is a complete recourse matrix (for the definition of the complete recourse matrix see e.g. [3]),

A.2 there exists $u \in R^m$ such that $u^{'}W \leq q$ a.s.

Theorem 2. [8] Let $t > 0$, X be a compact set, the assumptions A.1, A.2 and the assumptions 1, 2 of Theorem 1 be fulfilled. If

1.
$$f_0(x,\, \xi) = g_0(x,\, \xi) + Q(x,\, \xi)$$
$$Q(x,\, z) = \min_{\{y \in R^{n_1}:\, Wy = h - Tx,\, y \geq 0\}} q^{'}y\},$$

2. $g_0(x, z)$ is a uniformly continuous, Lipschitz (with respect to $\mathcal{L}_1$ norm) function of $z \in R^s$, the Lipschitz constant L is not depending on $x \in X$,

then

$$P\{\omega :\, N^{\beta}|\varphi(F) - \varphi(F^N)| > t\} \longrightarrow_{(N \longrightarrow \infty)} 0 \quad \text{for} \quad t > 0,\, \beta \in (0,\, \frac{1}{2}).$$

Proof. Employing the assertion of Propositions 1, 2, Lemma and the technique employed in [8] we obtain the assertion of Theorem 2. □

4 Conclusion

The paper deals with the convergence rate of the optimal value of the empirical estimates in the case of the stochastic programming with recourse. It is known that if X is a convex, nonempty, compact set and either $f_0(x,\, z)$ a strongly convex (with a parameter $\rho > 0$) function on X or some growth conditions ([8], [12]) are fulfilled, then also

$$P\{\omega :\, N^{\beta}\|x(F^N) - x(F))\|^2 > t\} \longrightarrow_{(N \longrightarrow \infty)} 0 \quad \text{for } t > 0,\, \beta \in (0,\, \frac{1}{2}). \tag{9}$$

To see the conditions under which $Q(x, z)$ is a strongly convex function on X see e.g. [11].

Acknowledgment The research was supported by the Czech Science Foundation under Grants 402/07/1113, 402/08/0107 and 402/06/0990.

References

1. Dupačvá J, Wets RJB (1984) Asymptotic behaviour of statistical estimates and optimal solutions of stochastic optimization problems. Ann Statist 16: 1517–1549
2. Hoeffding W (1963) Probability inequalities for sums of bounded random variables. Journal of Americ Statist Assoc 58: 13–30
3. Kall P (1976) Stochatic linear programming. Springer, Berlin
4. Kaniovski YM, King AJ, Wets RJB (1995) Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems. Annals of Oper Res 56: 189–208
5. Kaňková V (1978) An approximative solution of stochastic optimization problem. In: Trans 8th Prague Conf Academia, Prague: 349–353
6. Kaňková V (1999) Convexity, Lipschitz property and differentiability in two–stage stochastic nonlinear programming problems. Aportaciones Matematicas Serie Comunicaciones 24
7. Kaňková V, Houda M (2006) Empirical estimates in stochastic programming. In: Proceedings of Prague Stochastics 2006. (M. Hušková and M. Janžura, eds.). MATFYZPRESS, Prague: 426–436.
8. Kaňková V (2007) Empirical estimates via stability in stochastic programming. Research Report UTIA 2007, No. 2192
9. Kaňková V (2008) A remark on nonlinear functionals and empirical estimates. In: Proceedings of Quantitative Methods in Economics (Multiple Criteria Decision Making XIV). The Slovak Society for Operations Research and University of Economics in Bratislava 2008 (to appear)
10. Pflug GCh (2003) Stochastic optimization and statistical inference. In: Stochastic Programming (A. Ruszczynski and A. A. Shapiro, eds.). Handbooks in Operations Research and Managemennt Science, Vol 10. Elsevier, Amsterodam
11. Römisch W, Schulz R (1993) Stability of solutions programs with complete recourse. Matematics of Operations Research 18:
12. Shapiro A (1994) Quantitative stability in stochastic programming. Math Program 67: 99–108
13. Wets RJB (1974) A statistical approach to the solution of stochastic programs with (convex) simple recourse. Research Report University Kentucky USA