On quantitative stability in stochastic programming problems with recourse

Michal Houda
Charles University of Prague
Department of Probability and Mathematical Statistics
Sokolovská 83, 186 75 Praha 8-Karlín
houda@karlin.mff.cuni.cz

Abstract

The paper deals with the quantitative stability of stochastic programs with recourse with respect to a space of probability measures. In detail, the paper deals with the stability of the problems where the operator of mathematical expectation appears in the objective function, the constraint set is deterministic, and the space is equipped with the Kolmogorov or Wasserstein metrics. A numerical illustration on the farmer's problem is presented: different bounds on optimal values can be obtained using the two considered metrics.

Keywords

Stochastic programs with recourse, quantitative stability, Wasserstein metrics, Kolmogorov metrics, farmer's problem.

1 Introduction

Random inputs have to be taken into account by a variety of optimization problems and the randomness is introduced into mathematical models through the underlying probability distribution rather than the random variable itself (usually unknown). Furthermore, the distribution is assumed to be completely known in a large class of problems (because of some favorable properties of the model). In real-life applications, two frequent cases happen: first, the probability distribution is not known and hence estimated (in a more or less precise way); second, the probability distribution is too complicated to be applied directly and it is subject of approximation.

It is of interest to answer a question if the estimated (approximated) distribution and thus solutions obtained are still relevant to the original problem. Achieved results on the *stability* of optimal value and optimal solution set can help to do it. First, we have to quantify a distance between the two distributions (original and approximated), i. e. to introduce a suitable metrics on the space of probability measures (also called *probability metrics*). This "suitability" is very closely related to the nature of the problem, as we can see in the second section of the paper. Our attention is paid to the Wasserstein and Kolmogorov metrics.

This paper deals with the stability of problems where the operator of mathematical expectation appears in the objective function and the constraint set is deterministic. This class of problems is often called as problems with penalty or *with recourse* and are widely applied in a number of economic models. The third section of the paper review some quantitative stability theorems involving this class of problems: we are especially interested in Lipschitz continuity properties of optimal value. A great attention has already been paid to this problem in the literature; recall e. g. papers [3], [7], [8], etc. The last section of the paper is devoted to a numerical illustration of these properties: a variation of the popular farmer's problem is taken in order to show stability bounds with the two metrics considered.

2 Probability metrics

Let $\mathcal{P}(\Xi)$ be the space of all Borel probability measures defined on some Borel set $\Xi \subset \mathbb{R}^s$. Let $\mu, \nu \in \mathcal{P}(\Xi)$ and F, G be their corresponding distribution functions. The *Kolmogorov metrics* is defined by

$$\mathcal{K}(\mu,\nu) := \sup_{z\in\Xi} |F(z) - G(z)|.$$

To define the Wasserstein metrics, let $p \in [1; +\infty)$ and consider the space $\mathcal{P}_p(\Xi) := \{\nu \in \mathcal{P}(\Xi) : \int_\Xi \|\xi\|^p \nu(\mathrm{d}\xi) < +\infty\}$ where $\|\cdot\|$ denotes the Euclidean norm. The *p-Wasserstein metrics* is defined for $\mu, \nu \in \mathcal{P}_p(\Xi)$ by

$$W_p(\mu,\nu) := \left[\inf\left\{\int_{\Xi\times\Xi} \|\xi - \tilde{\xi}\|^p \ \eta(\mathrm{d}\xi \times \mathrm{d}\tilde{\xi}) : \eta \in D(\mu,\nu)\right\}\right]^{1/p},$$

where $D(\mu,\nu)$ is the set of all probability measures of $\mathcal{P}(\Xi\times\Xi)$ for which μ and ν are marginal distributions. It can be shown (see [9]), that for $s = 1$

$$W_p(\mu,\nu) = \left(\int_0^1 |F^{-1}(t) - G^{-1}(t)|^p \mathrm{d}t\right)^{1/p}, \ W_1(\mu,\nu) = \int_{-\infty}^{+\infty} |F(t) - G(t)| \mathrm{d}t.$$

Figure 1 illustrates that an application of the metrics can fail even if we deal with rather common case. First figure illustrates approximation of a discrete distribution with unknown mass points, the second one a distribution with heavy tails (F and G take forms of $F(z) := \varepsilon/(1-z)$, $G(z) := 0$ on $[-K; 0)$ with $\varepsilon \in (0; 1)$ and $K \gg 0$, $F(z) = G(z)$ are arbitrary on $[0; +\infty)$). We refer to the book [6] for further details and references about these and other probability metrics.

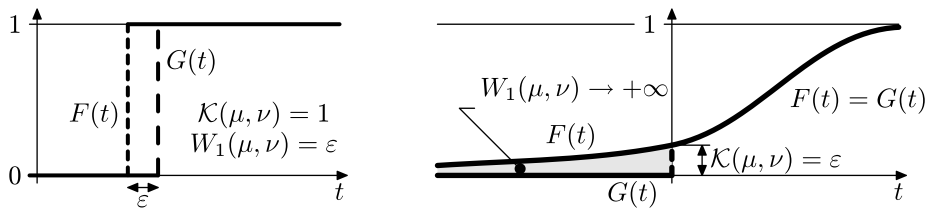

Figure 1: Kolmogorov and Wasserstein metrics as unsuitable metrics

3 Stability results

For $\mu \in \mathcal{P}_1(\Xi)$ consider a problem

$$\text{minimize} \int_\Xi g(x;\xi)\mu(\mathrm{d}\xi) \quad \text{subject to } x \in X \tag{1}$$

where $X \subset \mathbb{R}^n$ is a compact set, the function g is uniformly continuous on $\mathbb{R}^n \times \mathbb{R}^s$ and Lipschitz in ξ for all $x \in X$ with a Lipschitz constant L not depending on x. Denote $\varphi(\mu)$ the optimal value and $\psi(\mu)$ the optimal solution set of the problem (1).

Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is *strongly convex* on a convex set $V \subset \mathbb{R}^n$ with parameter $\sigma > 0$, if for any $x_1, x_2 \in V$ and any $\lambda \in [0; 1]$ we have $f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda)f(x_2) - \frac{1}{2}\sigma\lambda(1-\lambda)\|x_1 - x_2\|^2$.

Theorem 1 *For any $\nu \in \mathcal{P}_1(\Xi)$ we have*

$$|\varphi(\mu) - \varphi(\nu)| \le L\, W_1(\mu,\nu).$$

If moreover X is a convex set and $g(\cdot,\xi)$ is strongly convex function on X with parameter $\sigma > 0$, then

$$\|\psi(\mu) - \psi(\nu)\| \le \frac{8}{\sigma} L\, W_1(\mu,\nu).$$

This is a little generalized assertion of Römisch and Schultz proven originally for the complete linear recourse (see [7], [2]). Due to the strong convexity property, we can consider $\psi(\nu)$ as a (unique) point of $\mathbb{R}^n$ instead of a singleton in $\mathbb{R}^n$. A similar result employing the Kolmogorov metrics can be stated (see e. g. [5]). However, we now pay our attention to the case where (1) has the following separable structure

$$\text{minimize} \sum_{i=1}^{s} \int_{\Xi_i} g_i(x; \xi_i)\mu_i(\mathrm{d}\xi_i) \quad \text{subject to } x \in X \tag{2}$$

where for all $i = 1, \ldots, s$, μ_i are one-dimensional probability measures with support $\Xi_i \subset \mathbb{R}$, g_i are uniformly continuous on $\mathbb{R}^n \times \mathbb{R}$ and Lipschitz in ξ_i for all $x \in X$ with Lipschitz constants L_i not depending on x.

Theorem 2 *([2]) For any $\nu \in \mathcal{P}_1(\Xi)$ in (2) we have*

$$|\varphi(\mu) - \varphi(\nu)| \leq \sum_{i=1}^{s} L_i\, W_1(\mu_i, \nu_i).$$

Theorem 3 *([4]) Assume (i) μ_i are absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$ with $f_i(\cdot)$ as density functions; (ii) Ξ_i are closed (non-degenerated) intervals on $\mathbb{R}$; (iii) there exists real constants $\vartheta_i > 0$ such that $\vartheta_i \leq f_i(\xi)$, $\xi_i \in \Xi_i$. Then for any $\nu \in \mathcal{P}_1(\Xi)$ we have*

$$|\varphi(\mu) - \varphi(\nu)| \leq \sum_{i=1}^{s} \frac{1}{\vartheta_i} L_i \mathcal{K}(\mu_i, \nu_i).$$

If moreover X is a convex set and function $\sum g_i(\cdot, \xi)$ is strongly convex function on X with parameter $\sigma > 0$, then

$$\|\psi(\mu) - \psi(\nu)\|^2 \leq \sum_{i=1}^{s} \frac{8}{\sigma \vartheta_i} L_i \mathcal{K}(\mu_i, \nu_i).$$

4 Farmer's problem

Consider the following farmer's problem (see [1]): a farmer specializes in raising three crops, grain, corn, and sugar beets on a land of 500 acres. He requires an amount of 200 tons (T) of wheat and 240 T of corn for his needs (both can be raised or bought elsewhere). Assume that production costs, and selling and purchase prices are known (see Table 1), and that there is a quota of 6000 T on sugar beets above which the selling price is considerably reduced. Finally consider the yields for each crop to be independent random variables with uniform distribution with ranges stated in Table 1 again.

		Notation	*grain*	*corn*	*sugar beets*
Production costs	\$/acre	c	150	230	260
Selling price	\$/T	q^+	170	150	36
(quota exceeded)		r^+	–	–	10
Purchase price	\$/T	q^-	238	210	40
Min. requierements	T	b	200	240	0
Yield (range)	T/acre	ξ	$[2; 3]$	$[2.4; 3.6]$	$[16; 24]$
Total available land	T/acre	a		500	
Production quota	T	l	–	–	6000
Area for each crop	T/acre	x		1^{st} stage variables	
Sales	T	y^+		2^{nd} stage variables	
(quota exceeded)		w^+			
Purchase	T	y^-		2^{nd} stage variables	

Table 1: Input values for the farmer's problem

The farmer's goal is to devote an optimal amount of land for each crop in order to minimize his costs. We can formulate the above problem as a two-stage stochastic linear program (subscripts 1–3 refer to grain, corn, and sugar beets respectively): minimize $c^T x + E_\mu Q(x;\xi)$ subject to $x_1 + x_2 + x_3 \le a,\ x \ge 0$, where[1]

$$\begin{aligned} Q(x;\xi) = \min\{q^T y \mid \xi_1 x_1 - y_1^+ + y_1^- &\ge b_1 \\ \xi_2 x_2 - y_2^+ + y_2^- &\ge b_2 \\ \xi_3 x_3 - y_3^+ - w_3^+ + y_3^- &\ge b_3 \\ y_3^+ \le 6000;\ x, y^+, y^-, w_3^+ &\ge 0\ \} \end{aligned}$$

This favourable form of the second stage problem gives us possibility to rewrite Q as $Q(x;\xi) = Q_1(x_1;\xi_1) + Q_2(x_2;\xi_2) + Q_3(x_3;\xi_3)$. Functions Q_1, Q_2, Q_3 are Lipschitz in ξ. with Lipschitz constants $L_1 = 119000, L_2 = 105000, L_3 = 20000$.

In order to simulate a random yield, we generate a sample of length N (from the uniform distribution). We set up empirical distribution functions, calculate values of Kolmogorov and Wasserstein metrics with respect to the original uniform distribution and thus obtain values of the upper bound for the optimal value (right-hand sides in Theorems 2 and 3). We have made 100 repetitions of each of samples and summarize values of the mean and the 95% confidence interval for the upper bound in Table 2. In this case of uniform distribution, the Wasserstein metrics seems to delimit a stability interval in better way.

[1] We denote $x = (x_1, x_2, x_3)^T$ (and similar for y^+, y^-, etc.), $y = (y^+, w_3^-, y^-)^T$, $q = (-q^+, -r_3^+, q^-)^T$.

	Wasserstein metrics' bound	*Kolmogorov metrics' bound*
$N = 10$	(37824) 39716 (41608)	(99923) 103673 (107423)
$N = 50$	(17082) 17986 (18890)	(46378) 48179 (49979)
$N = 100$	(12025) 12664 (13304)	(33219) 34400 (35579)

Table 2: Mean bounds for optimal values and theirs 95% confidence intervals

References

[1] Birge, J. R., and Louveaux, F. (1997), *Introduction to Stochastic Programming,* Springer-Verlag, New York.

[2] Houda, M. (2001), Stabilita a odhady v úlohách stochastického programování (Speciální případy), diploma thesis, Charles University, Faculty of Mathematics and Physics, Prague, in Czech.

[3] Kaňková, V. (1994), On stability in two-stage stochastic nonlinear programming, in: *Asymptotic Statistics* (P. Mandl and M. Hušková eds.), Physica-Verlag, Heidelberg, 329–340.

[4] Kaňková, V. (1994), On the stability in stochastic programming — generalized simple recourse problems, *Informatica* 5, 55–78.

[5] Kaňková, V., and Houda, M. (2002), A note on quantitative stability and empirical estimates in stochastic programming, in: *Operations Research Proceedings 2002* (U. Leopold-Wildburger, F. Rendl, and G. Wäsher eds.), Springer-Verlag, Heidelberg, 413–418

[6] Rachev, S. T. (1991), *Probability Metrics and the Stability of Stochastic Models*, Wiley, Chichester.

[7] Römisch, W., and Schultz, R. (1993), Stability of solutions for stochastic programs with complete recourse, *Mathematics of Operations Research* 18, 590–609.

[8] Shapiro, A. (1994), Quantitative stability in stochastic programming, *Mathematical Programming* 67, 99–108.

[9] Vallander, S. S. (1973), Calculation of the Wasserstein distance between probability distributions on the line, *Theory of Probability and its Applications* 18, 784–786.