Weak Structural Dependence in Chance-Constrained Programming

Michal Houda[1]

Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, houda@karlin.mff.cuni.cz

Abstract. In chance-constrained optimization problems, a solution is assumed to be feasible only with certain, sufficiently high probability. For computational and theoretical purposes, the convexity property of the resulting constraint set is treated. It is known, for example, that a suitable combination of a concavity property of the probability distribution and concavity of constraint mappings are sufficient conditions to the convexity of the resulting constraint set. Recently, new concavity condition of the probability distribution - r-decreasing density - has been developed. Henrion and Strugarek (2006) show, under the assumption of independence of constraint rows, that this condition on marginal densities allows us, on the other side, weaken the concavity of constraint mappings. In this contribution we present a relaxation of the independence assumption in favour of a specific weak-dependence condition. If the independence assumption is not fulfiled, the resulting constraint set is not due to be convex. However, under a weak-dependence assumption, the non-convex problem can be approximated by a convex one. Applying stability results on optimal values and optimal solutions, we show that optimal values and optimal solutions remain stable under assumptions common in stochastic programming. This implies desirable consequences, because convex problems are easiest to compute and also many theoretical results are based on convexity assumptions. We accompany the shown results by simple example to illustrate the concept of the presented approximation.

1 Introduction

Theory of optimization is very important area to model real-life economic and engineer problems. These can be modelled as deterministic, with ad-hoc selected or estimated parameters but there are many known cases leading to false results due to presence of uncertainty that cannot be neglected. In such cases various approaches dealing with uncertainty are applied, e. g., traditional sensitivity analysis, parametric programming, robust programming techniques. In this paper we discuss stochastic programming approach, notament its important part: chance-constrained programming.

Consider an optimization problem of the form

$$\text{minimize } c(x) \text{ subject to } h(x;\xi) \geq 0 \tag{1}$$

where $x \in \mathbb{R}^m$ is a decision vector, $\xi \in \Xi \subset \mathbb{R}^s$ is a parameter of the problem, $c : \mathbb{R}^m \to \mathbb{R}$ is real and $h : \mathbb{R}^m \times \mathbb{R}^s \to \mathbb{R}^d$ vector-valued mappings. The parameter ξ is not known precisely – suppose that it is a random vector on some underlying probability space $(\Omega, \mathcal{A}, \Pr)$ with known probability distribution μ defined on the support Ξ. We can now require the constraints of the problem to be satisfied only with some prescribed but sufficiently high probability $p \in [0; 1]$. We obtain so-called chance-constrained optimization problem:

$$\text{minimize } c(x) \text{ subject to } \Pr\{h(x; \xi) \geq 0\} \geq p \tag{2}$$

There are many known results from both the theoretical and application area of chance-constrained programming, see e. g. [4]. One of two main known issues discussed by the stochastic programming community is the question of convexity of the feasible set in (2). Denote

$$M(p) \coloneqq \Pr\{h(x; \xi) \geq 0\} \geq p \tag{3}$$

Convexity of the set $M(p)$ is of high importance from the theoretical as from the computational point of view. There are known results starting with the trivial one: if $h(\cdot, \xi)$ is convex for all ξ then the sets $M(0)$, $M(1)$ are convex. Unfortunately, these sets are not of our usual interest (they represent all possible, and all almost sure realizations respectively).

An answer to the question posed above, nowadays considered as classical one, is first developed in [3]. The author uses a concept of parameterization of the concavity of a function.

Definition 1. *A function $f : \mathbb{R}^d \to (0; +\infty)$ is called r-concave for some $r \in [-\infty; +\infty]$ if*

$$f(\lambda x + (1 - \lambda)y) \geq [\lambda f^r(x) + (1 - \lambda)f^r(y)]^{1/r} \tag{4}$$

is valid for each $x, y \in \mathbb{R}^d$ and each $\lambda \in [0; 1]$. The cases $r = -\infty, 0, +\infty$ are treated by continuity.

Our interest focuses on values of $r \leq 1$. For $r = 1$, f is concave in its classical sense. 0-concave function is also called log-concave (as $log f$ is concave in such case), $-\infty$-concave function is known as quasi-concave function. Further, in [3] you can find the definition of r-concave probability measure which is used as assumption for the following result:

Proposition 1 ([5], Theorems 2.5 and 2.11). *If μ is absolutely continuous (with respect to Lebesgue measure), log-concave (or r-concave for $r \geq -1/s$) measure, and the one-dimensional components of h are quasi-concave functions of (x, ξ) then $M(p)$ is convex set.*

For our purposes it is sufficient to say that a log-concave (r-concave) measure is implied by a log-concave ($\frac{r}{1-rs}$-concave) density. Many multivariate distributions (normal, beta, Wishart, etc.) share this property hence many chance-constrained problems can be solved by means of convex optimization.

In this paper we focus on the chance-constrained problem with random right-hand side only:

$$\min c(x) \text{ subject to} \Pr\{g(x) \geq \xi\} \geq p. \tag{5}$$

Here, we have set $h(x;\xi) = g(x) - \xi$ with g having appropriate dimensions, and according to Proposition 1 we require h to be quasi-concave. Unfortunately, quasi-concavity is not preserved under addition and we have to require $g(x)$ to be concave (quasi-concave is not sufficient) to satisfy this requirement (see [4] again).

Recently, Henrion and Strugarek [1] proposed an idea to relax concavity condition of g and make more stringent concavity condition on the probability distribution μ. They defined the notion of so-called r-decreasing density as follows:

Definition 2 ([1], Definition 2.2). *A function $f: \mathbb{R} \to \mathbb{R}$ is called r-decreasing for some $r \in \mathbb{R}$ if*

1. *it is continuous on $(0;+\infty)$, and*
2. *there exists a threshold $t^* > 0$ such that $t^r f(t)$ is strictly decreasing for all $t > t^*$.*

In [1], the notion of r-decreasng densities is used to replace the original assumption on r-concavity of the distribution of μ as follows:

Theorem 1 ([1], Theorem 3.1). *If*

1. *there exist $r_i > 0$ such that the components g_i of g are $(-r_i)$-concave,*
2. *the components ξ_i of ξ have $r_i + 1$-decreasing densities, and*
3. *the components ξ_i of ξ are independently distributed,*

then $M(p)$ is convex for all $p > p^* := \max_i F_i(t_i^*)$ where F_i denotes the distribution function of ξ_i (one-dimensional marginals of ξ) and t_i^* refer to the definition of $r_i + 1$-decreasing probability density.

2 Weak dependence of the rows

An important assumption – the row independence (assumption 3 in the previous theorem) – could be relaxed in the following way. We define an α coefficient of dependence as

$$\alpha := \sup_z \Big| F(z) - \prod_i F_i(z_i) \Big| \tag{6}$$

where F is the distribution function of the vector ξ, F_i are the corresponding one-dimensional marginal distribution functions and $z = (z_1, \ldots, z_s) \in \mathbb{R}^s$.

It is a modified version of the strong-mixing coefficient, well known from the theory of random processes. To simplify the notation we use the notion of

α-dependence and α coefficient in the sense of the above definition and not in the classical sense which we will not need anymore in the remaining part of the paper.

In Problem (5) we allow for a small structural dependence in the following way. Recall that the set $M(p)$ of feasible solution can be written as

$$M(p) := \{x \in X | F(g(x)) \geq p\} \tag{7}$$

where F is the distribution function of the random right-hand side ξ. For our purposes we replace $M(p)$ in (5) by another set defined as

$$M'(p) = \{x \in X | \prod_{i=1}^{s} F_i(g_i(x)) \geq p\} \tag{8}$$

If the components ξ_i of ξ are independently distributed then the two sets $M(p)$ and $M'(p)$ are equal. This is not true for the dependent case but the following proposition is valid

Proposition 2. *If the components ξ_i of ξ in (5) are α-dependent (in the sense of (6) then*

$$M'(p+\alpha) \subset M(p) \subset M'(p-\alpha) \subset M(p-2\alpha). \tag{9}$$

The proof of this proposition is given in [2]. For sufficiently high p the (possibly) non-convex set is bounded from both sides by convex sets by the following theorem:

Theorem 2. *If*

1. *there exist $r_i > 0$ such that the components g_i of g are $(-r_i)$-concave,*
2. *the components ξ_i of ξ have $r_i + 1$-decreasing densities,*
3. *the components ξ_i of ξ are α-dependently distributed, and*
4. *$p > \max_i F_i(t_i^*) + \alpha$*

*then $M(p)$ is bounded (from both sides) by **convex** sets $M'(p+\alpha)$ and $M'(p-\alpha)$.*

For small values of α, classical stability results of stochastic programming apply. For example, if $c(x)$ is Lipschitz continuous, the constraints are metrically regular and α sufficiently small, then the optimal values and optimal solutions remains (locally) stable. For details see [2] again.

3 Comparison of the dependent and independent case

Consider the following optimization problem

$$\text{minimize } x + y \text{ subject to} \tag{10}$$

$$g_1(x,y) = \frac{1}{x^2+y^2+0.1} \geq \xi_1$$
$$g_2(x,y) = \frac{1}{(x+y)^2+0.1} \geq \xi_2$$

and assume that the random vector ξ is normally distributed with zero mean and the variance matrix Σ. We consider two cases:
(a) independent case with

$$\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and (b) weak dependent case with

$$\Sigma = \begin{pmatrix} 1 & 0.1 \\ 0.1 & 1 \end{pmatrix}$$

In this example the g_i's are (-1)-concave, ξ_i's have 2-decreasing densities with the treshold $t^* = \sqrt{2}$, the critical probability level is $p^* = \Psi(\sqrt{2}) = 0.921$ and the weak-dependence coefficient for the dependent case is $\alpha = 0.017$.

The overall shape of the collection of sets $M'(p)$ is given on Figure 3. Each individual set $M'(p)$ is given as horizontal cut on the specified level p (z-axis). The contour lines of these sets are depicted on the following figure; the symbol φ denotes the corresponding optimal values.

For the chosen normal distribution, convexity of the feasible set is assured theoretically at the probability level of 0.921 in the independent case. As Figure 3 shows, the actual probability level in the example is much more smaller, around the value of 0.7. In the weak dependent case, these tresholds (theoretical and actual) are shifted towards the center of feasibility sets (center of image), and the optimal values and optimal solutions (depicted as points on the last figure) remain stable as the value of α-coefficient is small. Interesting behaviour around the actual treshold value is still open question from the theoretical point of view.

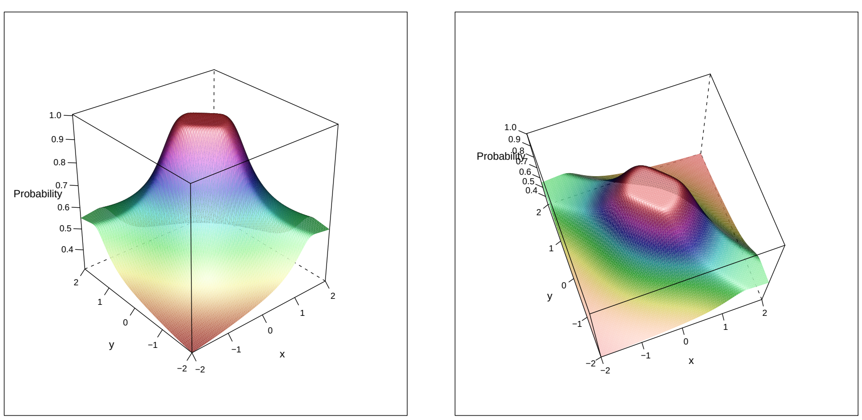

Fig. 1. Collection of sets $M'(p)$

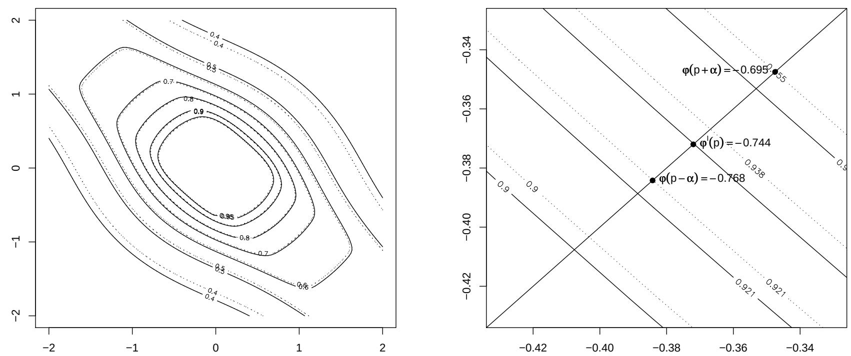

Fig. 2. Contour lines for $M'(p)$ (solid) and $M(p)$ (dotted) sets

References

1. Henrion, R., Strugarek, C.: Convexity of chance constraints with independent random variables. Stochastic Programming E-Print Series (SPEPS) **9** (2006) To appear in Computational Optimization and Applications.
2. Houda, M.: Convexity and dependence in chance-constrained programming. Research Report 2190, Institute of Information Theory, Academy of Science of the Czech Republic (August 2007)
3. Prékopa, A.: A class of stochastic programming decision problems. Mathematische Operations forschung und Statistik (1972) 349–354
4. Prékopa, A.: Stochastic Programming. Akadémiai Kiadó, Budapest (1995)
5. Prékopa, A.: Probabilistic programming. In Ruszczynski, A., Shapiro, A., eds.: Stochastic Programming. Handbooks in Operations Research and Management Science. Volume 10. Elsevier, Amsterdam (2003) 267–352