Research Article

Exact Penalization in Stochastic Programming—Calmness and Constraint Qualification

Martin Branda

Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University in Prague, Sokolovská 83, 186 75 Prague, Czech Republic

Correspondence should be addressed to Martin Branda; branda@karlin.mff.cuni.cz

Received 31 October 2013; Accepted 17 January 2014; Published 23 February 2014

Academic Editor: Shelton Peiris

Copyright © 2014 Martin Branda. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We deal with the conditions which ensure exact penalization in stochastic programming problems under finite discrete distributions. We give several sufficient conditions for problem calmness including graph calmness, existence of an error bound, and generalized Mangasarian-Fromowitz constraint qualification. We propose a new version of the theorem on asymptotic equivalence of local minimizers of chance constrained problems and problems with exact penalty objective. We apply the theory to a problem with a stochastic vanishing constraint.

1. Introduction

Constraints of real life optimization problems often depend on a random vector. If we know its probability distribution or if we have a good estimate, we can use various stochastic programming formulations to solve the problem and to get a highly reliable solution with respect to the realizations of the random vector. In chance constrained programming problems, a probability for fulfilling the random constraints is prescribed; see Prékopa [1, 2] for an extensive review of the main results. In general, chance constrained problems are difficult to solve. Below, we give an overview of the main recent results in this area. P-level efficient points (pLEPs) introduced by Prékopa [3] can be seen as generalizations of the quantiles for multivariate probabilistic distributions and have been shown to be useful for solving chance constrained problems with random right-hand sides. The new definition of pLEPs proposed by Lejeune [4] uses elements from the combinatorial pattern recognition field. He demonstrated that the collection of pLEPs can be represented as a disjunctive normal form (DNF) and proposed a mixed-integer programming formulation allowing for the construction of the DNF. Bruni et al. [5] presented a heuristic approach EXPRESS for solving mixed-integer programming formulations of chance constrained problems. The heuristic is based on alternating phases of exploration using simplified integer problems and feasibility repairing using penalization. Computational results on a probabilistically constrained version of the multiknapsack problem confirmed good solution quality and time performance of the heuristic. Dentcheva and Martinez [6] analyzed nonlinear stochastic optimization problems with probabilistic constraints described by continuously differentiable nonconvex functions. They described the tangent and the normal cone to the level sets of the probability function and formulated first- and second-order optimality conditions. Moreover, they developed an augmented Lagrangian method for discrete distributions. Henrion and Möller [7] provided an explicit gradient formula for linear chance constraints under a multivariate Gaussian distribution and used the formula to solve a transportation network problem with stochastic demands. An important property of the feasibility set is convexity which can be ensured, for example, by log-concavity of the underlying distribution; see Shapiro et al. [8] for a review and generalizations. A recent result in this area has been proposed by Ninh and Prékopa [9] who investigated log-concavity of the compound distributions, which have many applications in insurance, finance, hydrology, and so forth.
Sample approximation techniques were studied by Luedtke and Ahmed [10], Branda [11] who derived exponential rates of convergence of the approximated solutions to the true one. Convergence of approximated probabilistic functions expressed as quantiles is investigated also by Miao et al. [12]. The underlying probability distribution is usually not known precisely and possible influence of its improper choice has to be investigated. Dupačová and Kopa [13] accessed robustness of the probabilistic risk functions using the contamination technique and derived approximate contamination bound for the optimal value.

In this paper, we investigate an approach which uses penalized random constraints incorporated into the objective function. It can be verified that under some conditions we obtain an optimal solution with desirable reliability by increasing the penalty parameter and solving the resulting problems. It can be even shown that under mild conditions the penalty approach and the chance constrained problem are asymptotically equivalent; see Branda [14, 15], Branda and Dupačová [16], and Ermoliev et al. [17]. Recently, the equivalence was stated under exact penalization where a finite penalty parameter is sufficient to obtain a local optimal solution which is permanently feasible with respect to the random constraints; see Branda [18]. The exact penalization property is ensured by a very general calmness property of the underlying optimization problem; compare Burke [19, 20], Clarke [21]. However, it is difficult or even impossible to verify the problem calmness directly. Thus, in this paper, some necessary conditions for stochastic programming problems are introduced and discussed. These conditions are based on the relations between constraint set calmness, local error bound property, and constraint qualification conditions. Moreover, we perform the whole analysis on the problems with general abstract constraints. This setting is important, for example, for the problems with vanishing constraints which are extended in this paper to take into account randomness. The deterministic case was considered, for example, by Hoheisel et al. [22].

There is an extensive development of the exact penalty methods for deterministic optimization in recent years. Antczak [23] established the exactness of the $l_1$ penalty function method used for solving nonsmooth constrained optimization problems with both inequality and equality constraints under locally Lipschitz invexity assumptions. Antczak [24] used the exact minimax penalty function method for solving a general nondifferentiable extremum problem with both inequality and equality constraints. Lian and Han [25] proposed a method to smooth the square-order exact penalty function for inequality constrained optimization and showed convergence of their algorithm under mild conditions. Liu et al. [26] solved nonconvex optimization problems using a penalty algorithm based on smooth functions that approximate the usual exact penalty function. Under a generalized Mangasarian-Fromovitz constraint qualification condition they proved the convergence of the generated points to Karush-Kuhn-Tucker points of the original problem. Zaslavski [27] developed an approach in order to prove the existence of an exact penalty involving Palais-Smith condition and nonexistence of a solution which is a critical point for the constraint map. In his recent paper, Zaslavski [28] studied a class of constrained minimization problems on complete metric spaces; he established the generalized exact penalty property and obtained an estimate of the exact penalty. Applications of the exact penalty method in stochastic programming are limited to problems with second-order stochastic dominance constraints; see Meskarian et al. [29], Sun et al. [30].

The paper is organized as follows. The basic notation and problem formulations are proposed in Section 2. In Section 3, the necessary conditions for the problem calmness are summarized. In Section 4, we propose a theorem which states an asymptotic equivalence of the chance constrained problems and problems with exact penalty objective. An example of the problem with a stochastic vanishing constraint is given in Section 5. Section 6 concludes the paper.

2. Problem Statement

Let $\xi$ be a real random vector on the probability space $(\Omega, \mathcal{F}, P)$ with values in $\mathbb{R}^k$ and known probability measure $P$. A general chance constrained problem with an abstract random constraint can be formulated as

$$
\psi_\varepsilon = \min_x f(x)
\quad \text{s.t.} \quad P(F(x, \xi) \in \Lambda) \geq 1 - \varepsilon,
$$

(1)

where $f: \mathbb{R}^n \to \mathbb{R}$, $F: \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$, and $F(x, \xi)$ are continuous in $x$ for arbitrary $\xi$ and measurable in $\xi$ for arbitrary $x$, $\Lambda \subseteq \mathbb{R}^m$ is closed, and $\varepsilon \in [0, 1]$ is a small prescribed level. We will focus on the case when the distribution is finite discrete with realizations $\xi^1, \ldots, \xi^k$ and positive probabilities $p_1, \ldots, p_k$ such that $\sum_{s=1}^k p_s = 1$. Then the problem can be rewritten as

$$
\min_x f(x)
\quad \text{s.t.} \quad \sum_{s=1}^k p_s I(F(x, \xi^s) \in \Lambda) \geq 1 - \varepsilon,
$$

(2)

where $I$ denotes the indicator function which is equal to one if the condition in the brackets is satisfied and to zero otherwise. We will use the following modified $l_1$-norm for a vector $u \in \mathbb{R}^{ms}$ defined as

$$
\|u\| = \sum_{s=1}^k p_s \sum_{s=1}^m |u_{is}|
$$

(3)

which is necessary for further steps and to show the asymptotic equivalence of the considered stochastic programming problems; compare Branda [18].

An important role is put to the problem where all random constraints are fulfilled. To simplify the notation, we set

$$
\mathcal{F}_S(x) = \begin{pmatrix} F(x, \xi^1) & \cdots & F(x, \xi^k) \end{pmatrix}^T,
$$

(4)
and $\Lambda^S = x^S_{s=1} \Lambda$. For $u \in \mathbb{R}^{m_S}$, we formulate a perturbed mathematical program:

$$\min_x f(x)$$

s.t. \[ \mathcal{F}^S(x) + u \in \Lambda^S. \]

The unperturbed problem for which $u = 0$ corresponds to the case when $\varepsilon = 0$ in the problem (2); that is, the feasible solutions are permanently feasible with respect to the considered realizations of the random vector $\xi$. We denote by

$$d_{\Lambda^S}(u) = \min_{u' \in \Lambda^S} \|u - u'\|$$

the distance of a point $u$ and the set $\Lambda^S$ which will serve as the penalty function later on.

We will employ two concepts of calmness. The first is related directly to the perturbed problem (5); compare Burke [19, Definition 1.1].

**Definition 1.** Let $x^*$ be feasible for the unperturbed problem, that is, for $u = 0$. Then the problem (5) is said to be calm at $x^*$ if there exist constant $\tilde{N} \geq 0$ (modulus) and $\varepsilon > 0$ (radius) such that for all $(x, u) \in \mathbb{R}^n \times \mathbb{R}^{m_S}$ satisfying $x \in B_{\varepsilon}(x^*)$ and $\mathcal{F}^S(x) + u \in \Lambda^S$, one has

$$f(x) + \tilde{N} \|u\| \geq f(x^*).$$

Note that if the problem (5) is calm at $x^*$, then $x^*$ is necessarily a local solution to the unperturbed problem.

The following proposition was formulated by Burke [19, Theorem 1.1], Theorem 1.1, and gives the equivalence of the exact penalty property and problem calmness.

**Proposition 2.** Let $x^*$ be feasible for the unperturbed problem, that is, for $u = 0$. Then the unperturbed problem is calm at $x^*$ with modulus $\tilde{N}$ and radius $\varepsilon > 0$ if and only if $x^*$ is a local minimum of the function

$$f(x) + N \cdot d_{\Lambda^S}(\mathcal{F}^S(x))$$

over the neighbourhood $B_{\varepsilon}(x^*)$ for all $N \geq \tilde{N}$.

The form of the penalty problem follows from the theorem:

$$\varphi_N = \min_x f(x) + N \cdot d_{\Lambda^S}(\mathcal{F}^S(x)).$$

### 3. Necessary Conditions for Problem Calmness

In this section, we give necessary conditions to ensure calmness of the problem (5) which can be hardly verified directly from its definition. First, we review the Lipschitz-like properties of the set-valued mappings $M : \mathbb{R}^m \mapsto \mathbb{R}^n$ with graphs defined by

$$\text{Gph } M = \{(y, x) \in \mathbb{R}^m \times \mathbb{R}^n : x \in M(y)\};$$

see, for example, Henrion and Outrata [31], Rockafellar and Wets [32] for further details.

**Definition 3.** Let $M : \mathbb{R}^m \mapsto \mathbb{R}^n$ be a set-valued mapping. One says that $M$ fulfills

1. local Lipschitz property at $\bar{y}$ if and only if $\exists L, \varepsilon > 0$,

$$d_{M(y)}(x) \leq L \cdot d(y, 1_y), \quad \forall x \in M(y) \cap B_{\varepsilon}(\bar{y}),$$

2. Aubin property at $\bar{y}$ if and only if $\exists L, \varepsilon > 0$,

$$d_{M[\mathcal{F}]}(x) \leq L \cdot d(\bar{y}, y), \quad \forall x \in M(y) \cap B_{\varepsilon}(\bar{y}),$$

3. local upper Lipschitz property at $(\bar{y}, \bar{x})$ if and only if $\exists L, \varepsilon > 0$,

$$d_{M[\mathcal{F}]}(x) \leq L \cdot d(\bar{y}, y), \quad \forall x \in M(y) \cap B_{\varepsilon}(\bar{x}),$$

4. (graph) calmness at $(\bar{y}, \bar{x})$ in $\text{Gph } M$ if and only if $\forall L, \varepsilon > 0$,

$$d_{M[\mathcal{F}]}(x) \leq L \cdot d(\bar{y}, y), \quad \forall x \in M(y) \cap B_{\varepsilon}(\bar{x}).$$

It can be seen directly from the definitions that the calmness is the weakest of the conditions and is implied by the local upper Lipschitz property or by the Aubin property, whereas these two conditions are ensured by the local Lipschitz property.

For $u \in \mathbb{R}^{m_S}$, define by

$$M(u) = \{x \in \mathbb{R}^n : \mathcal{F}^S(x) + u \in \Lambda^S\}$$

the perturbation of the (original) constraint set $M(0) = (\mathcal{F}^S)^{-1}(\Lambda)$. The following proposition gives us an important result which shows that the graph calmness together with the local Lipschitz continuity of the objective function is sufficient conditions for the problem calmness.

**Proposition 4.** Let $x^*$ be a local minimizer such that $M$ is calm at $(0, x^*)$. Let the objective function $f$ be locally Lipschitz continuous around $x^*$. Then the original problem is calm at $x^*$ with modulus which is equal to the product of the modulus of the constraint set and the objective function.

**Proof.** See Hoheisel et al. [22, Proposition 3.5 and its proof].
In general, it is not possible to obtain such strong result. Thus, we have to focus on conditions which ensure the calmness at least locally. These conditions are based on constraint qualification; see, for example, Bazaraa et al. [33], Kruger et al. [34] for a review. We start with the definitions of important cones. The Fréchet normal cone to $\Pi \subseteq \mathbb{R}^p$ at $a \in \text{cl} \Pi$ is defined as

$$\tilde{N}_\Pi (a) = \left\{ y \in \mathbb{R}^p : \lim_{a' \to a} \sup_{\Pi} \frac{\langle y, a' - a \rangle}{\|a' - a\|} \leq 0 \right\}. \quad (16)$$

Note that in finite dimensional setting the Fréchet normal cone corresponds to the polar of the tangent cone. Limiting normal cone to $\Pi$ at $a$ is defined by $N_\Pi (a) = \lim \sup_{a' \to a} \tilde{N}_\Pi (a)$, where $a \in \text{cl} \Pi$ and the limes-superior is considered in the Painlevé-Kuratowski sense; see [32].

Definition 6. Let $\mathcal{F}^S$ be continuously differentiable. Generalized Mangeras-Fr´echet constraint qualification (GMFCQ) is fulfilled at $x^*$ if

$$\lambda^T (\mathcal{F}^S)' (x^*) = 0, \quad \lambda \in N_{\mathcal{F}^S} (\mathcal{F}^S (x^*)) \implies \lambda = 0. \quad (17)$$

The relevance of the GMFCQ to our analysis is confirmed by the following proposition.

Proposition 7. Let $\mathcal{F}^S$ be continuously differentiable. If the generalized Mangeras-Fr´echet constraint qualification is fulfilled at $x^*$, then the perturbation map $M$ is calm at $(0, x^*)$. Moreover, if $f$ is continuously differentiable, then the problem is calm at $x^*$.

Proof. For the first part see Flegel et al. [35, Corollary 5]. Note that Mordukhovich [36] even showed that GMFCQ implies the stronger Aubin property. The second part follows directly from Proposition 4. $\square$

Another important property, which is equivalent to graph calmness, is the existence of a local error bound of the constraint set.

Proposition 8. Calmness of the constraints system (15) at $(0, x^*)$ is equivalent to the existence of a local error bound:

$$d_{(\mathcal{F}^S)'(A_N)} (x) \leq L \cdot d_A (\mathcal{F}^S (x)), \quad \forall x \in B_\varepsilon (x^*), \quad (18)$$

for some constants $L > 0$ and $\varepsilon > 0$.

Proof. See Hoheisel et al. [22, Proposition 3.4]. $\square$

4. An Asymptotic Equivalence of Stochastic Programming Problems

In this section, we propose a new version of the theorem which concerns asymptotic equivalence of chance constrained problems and problems with penalty objective. We use the notion of optimal values for local optimal values; that is, the vector of decision variables $x$ is restricted to $B_\varepsilon (x^*)$ for some local optimal solution $x^*$ of (5) in the following theorem.

**Theorem 9.** One considers the chance constrained problem (1) and the problem with the penalty objective (9). One assumes that

(i) the problem (5) has a local solution $x^*$,

(ii) $f, F(\cdot, \xi)$ are continuously differentiable around the local solution $x^*$,

(iii) the GMFCQ is satisfied at the local solution $x^*$.

For arbitrary $y \in (0, 1)$, $N > 0$, and $\varepsilon \in (0, 1)$ put

$$\alpha_N (x) = N \cdot d_{\mathcal{F}^S} (\mathcal{F}^S (x)), \quad (19)$$

$$\beta_\varepsilon (x) = \frac{1}{\varepsilon^T} \cdot d_{\mathcal{F}^S} (\mathcal{F}^S (x)), \quad (20)$$

Then for any prescribed $\varepsilon \in (0, 1)$ there always exists $N \leq \tilde{N}$ large enough so that minimization (9) generates local optimal solutions $x_N$ which also satisfy the probabilistic constraint with the given $\varepsilon$, where $\tilde{N}$ denotes the calmness modulus of the problem (9).

Moreover, bounds on the local optimal value $\psi_\varepsilon$ of (1) based on the local optimal value $\psi_N$ of (9) and vice versa can be constructed:

$$\psi_\varepsilon (x_N) - \beta_\varepsilon (x_N) \leq \psi_\varepsilon (x_N) \leq \psi_N - \alpha_N (x_N),$$

$$\psi_\varepsilon (x_N) + \alpha_N (x_N) \leq \psi_N \leq \psi_\varepsilon (x_N) + \beta_\varepsilon (x_N) \leq \psi_\varepsilon (x_N) + \beta_\varepsilon (x_N) \leq \psi_N,$$  

with

$$\lim_{N \to \tilde{N}} \alpha_N (x_N) = \lim_{N \to \tilde{N}} \psi_N = \lim_{\varepsilon \to \varepsilon_0} \psi_\varepsilon (x_N) = 0.$$  

for any sequences of the local optimal solutions of the chance constrained problem $x_N^{\text{CP}}$ and the problem with penalty objective $x_N$, where $\varepsilon_0 \in \text{min}_p p_s$.

Proof. The proof is analogous to Theorem 3 proposed by Branda [18] if we realize that the GMFCQ together with differentiability implies calmness of the problem (5) and that the local solution of (5) is permanently feasible. $\square$

5. Problem with a Stochastic Vanishing Constraint

Let $g_1 (x, \xi), g_2 (x, \xi) : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ be continuously differentiable in $x$ for arbitrary $\xi$ and measurable in $\xi$ for arbitrary $x$. Then the joint chance constrained problem with a stochastic vanishing constraint can be formulated as

$$\min_{x} f (x)$$

s.t. $P (g_1 (x, \xi) \geq 0, g_1 (x, \xi) g_2 (x, \xi) \leq 0) \geq 1 - \varepsilon.$  

(23)
Note that if the function \( g_1 \) is equal to zero for some \( x \) and \( \xi \), then the constraint \( g_2(x, \xi) \leq 0 \) vanishes. If we denote
\[
\Lambda = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_1 x_2 \leq 0\}
\]
\[
F(x, \xi) = g_1(x, \xi), g_2(x, \xi)^T : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^2,
\]
we can rewrite the chance constraint in the abstract form as follows:
\[
P(F(x, \xi) \in \Lambda) \geq 1 - \varepsilon.
\]
Let \( \xi \) have a finite discrete distribution with realizations \( \xi^1, \ldots, \xi^\gamma \). Denote \( \Lambda^\xi = x^\xi \in \Lambda, \mathcal{F}^\xi(x) = (F(x, \xi^1), \ldots, F(x, \xi^\gamma))^T \). For a given permanently feasible point \( x^* \) we define the following partitioning of the set of scenarios:
\[
S_{00} = \{ s : g_1(x^*, \xi^s) > 0, g_2(x^*, \xi^s) = 0 \},
S_{0-} = \{ s : g_1(x^*, \xi^s) > 0, g_2(x^*, \xi^s) < 0 \},
S_{0+} = \{ s : g_1(x^*, \xi^s) = 0, g_2(x^*, \xi^s) > 0 \},
S_{00} = \{ s : g_1(x^*, \xi^s) = 0, g_2(x^*, \xi^s) = 0 \},
S_{0+} = \{ s : g_1(x^*, \xi^s) = 0, g_2(x^*, \xi^s) < 0 \}.
\]
According to Hoheisel et al. [22, Theorem 4.2], the GMFCQ is satisfied at \( x^* \) if the gradients
\[
\nabla_x g_1(x^*, \xi^s), \quad s \in S_{00}, \cup S_{00}
\]
are linearly independent and there is a vector \( d \in \mathbb{R}^n \) such that
\[
d^T \nabla_x g_1(x^*, \xi^s) > 0, \quad s \in S_{0-}
\]
\[
d^T \nabla_x g_1(x^*, \xi^s) = 0, \quad s \in S_{00}, \cup S_{00}
\]
\[
d^T \nabla_x g_2(x^*, \xi^s) < 0, \quad s \in S_{00}, \cup S_{00}
\]
Moreover, it follows from Hoheisel et al. [22, Corollary 4.7], that the penalty can be expressed under our settings in the form
\[
d_{\mathcal{F}^\xi(x)}(s) = \sum_{s} p_s \max \left\{ 0, -g_1(x, \xi^s), \min \left\{ g_1(x, \xi^s), g_2(x, \xi^s) \right\} \right\}
\]
\[
= \mathbb{E} \left[ \max \left\{ 0, -g_1(x, \xi), \min \left\{ g_1(x, \xi), g_2(x, \xi) \right\} \right\} \right].
\]
If the GMFCQ is fulfilled and the objective function is continuously differentiable, then the problem (5) is calm and the corresponding problem with the penalty objective derived above is exact. Thus, in this case Theorem 9 applies to the problem with a stochastic vanishing constraint.

6. Conclusions

In this paper, we have discussed conditions which ensure exact penalization in stochastic programming. We have started with a general calmness condition of the underlying problem, which is equivalent to the exact penalty property. However, this condition cannot be verified directly; therefore necessary conditions have been proposed. We have shown that the local Lipschitz continuity of the objective function together with the calmness of the constraint set, which is implied, for example, by generalized Mangasarian-Fromovitz constraint qualification, is sufficient. These findings have enabled us to propose a new version of the theorem concerning asymptotic equivalence of the chance constrained problems and the problems with exact penalty objective, which is the main contribution of this paper. The theoretical results have been applied to the problem with a stochastic vanishing constraint where a sufficient condition for fulfilling the GMFCQ has been formulated.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by the Czech Science Foundation under the Grant GP13-03749P.

References


Advances in Decision Sciences


Submit your manuscripts at
http://www.hindawi.com