A note on the connectedness of chance constraints

R. Henrion
Weierstrass Institute for Applied Analysis and Stochastics
10117 Berlin
Germany

Abstract

We prove a result on connectedness of (functional) chance constraints $P(h(x) \geq g(\xi)) \geq p$, where the decision variable x belongs to a Banach space and h is assumed to be strictly quasiconcave. The derived characterization completely relies on a constraint qualification for the mapping h. There are no assumptions on the distribution of ξ involved. For the purpose of illustration, a generic application is briefly discussed. Limiting counter-examples are provided and a simple criterion for the constraint qualification to hold is given in case of linear h.

1 Introduction

The consideration of probabilistic or chance constraints $P(h(x,\xi) \leq 0) \geq p$ is a basic issue in stochastic optimization (e.g. [2], [4],[7]). Here, x is a decision variable, ξ is a random vector, P is a probability measure and $p \in (0,1)$ denotes some probability level at which the given system of inequalities is required to hold according to the distribution of ξ. Both, for numerical treatment and theoretical investigations (such as stability), convexity properties of the set of feasible x defined by chance constraints figure as a basic issue in this context. It is well-known, for instance, that if h is quasiconcave and ξ has a quasiconcave probability distribution, then the chance constraint defines a convex set. The importance of this fact results from the possibility of identifying quasiconcave (or r-concave or logconcave) probability distributions. It turns out that many of the prominent multivariate distributions like normal, Pareto, Dirichlet, Gamma, uniform (on convex, compact sets) etc. share this property (see [7]). In contrast, discrete distributions typically fall outside this class.

On the other hand, it might be of some interest to characterize more general structural properties of chance constraints, such as connectedness. Apart from the theoretical relevance of this question, it may also have some practical justification: for instance, restricting to stationary solutions of a problem, the convexity of constraint sets is not an indispensable prerequisite of applying methods from nonlinear optimization such as sequential quadratic programming. Disconnected feasible sets, however, will result in much more severe difficulties as compared to just a violation of convexity.

The aim of this note is to provide a characterization of connectedness of chance constraints. In contrast to the convexity result mentioned above, we have to assume a separable structure $h(x) \geq g(\xi)$ of the stochastic inequalities involved. But then it turns out, that the weaker property of connectedness can be characterized exclusively on the basis of an appropriate constraint qualification formulated for the mapping h. In particular, no assumptions on the distribution of the random vector are required. This is convenient as one usually has limited knowledge only about the true underlying distribution.

We proceed by recalling a few concepts and facts which are necessary for the derivation of the result: a function $f : X \to \mathbb{R}$, with X being a linear space, is strictly quasiconvex if for all $x, y \in X$ and all $\lambda \in (0, 1]$ the relation $f(x) < f(y)$ entails the inequality $f(\lambda x + (1 - \lambda)y) < f(y)$. f is strictly quasiconcave if $-f$ is strictly quasiconvex. It follows immediately that convex functions are strictly quasiconvex and that strictly quasiconvex functions are quasiconvex. The analogous concavity statements hold true as well. A multifunction $\Phi : V \rightrightarrows W$ between topological spaces is called lower semicontinuous at $v^0 \in V$ if for each open $B \subseteq W$ with $B \cap \Phi(v^0) \neq \emptyset$ there exists some open neighborhood A of v^0 such that $B \cap \Phi(v) \neq \emptyset$ for all $v \in A$. The following selection theorem by Michael, which is stated here in a simplified version (see [1]. Th. 2.3.1), figures as the key tool for the proof of our result:

Theorem 1 (Michael) *Let $\Phi : \Lambda \rightrightarrows X$ be a lower semicontinuous multifunction, where Λ is a compact metric space and X is a Banach space. If for all $\lambda \in \Lambda$ the sets $\Phi(\lambda)$ are nonempty and convex, then there exists a continuous selection of Φ, i.e. a continuous function $\varphi : \Lambda \to X$ such that $\varphi(\lambda) \in \Phi(\lambda)$ for all $\lambda \in \Lambda$.*

The following theorem provides a useful characterization of lower semicontinuity for multifunctions (see [1], Th. 3.1.6):

Theorem 2 *Let $M : Y \rightrightarrows X$ be a multifunction between Banach spaces which is defined by*

$$M(y) = \{x \in X | \alpha_i(x, y) \leq 0,\ i \in J\},$$

where J is a finite index set and the $\alpha_i : X \times Y \to \mathbb{R}$ are upper semicontinuous. For some fixed $\bar{y}$, assume that the functions $\alpha_i(\cdot, \bar{y})$ are strictly quasiconvex and that there exists some $x \in X$ with $\alpha_i(x, \bar{y}) < 0$ $(i \in J)$. Then, M is lower semicontinuous at $\bar{y}$.

2 Result

In the following, let $(\Omega, \mathcal{A}, P)$ be a probability space, $\xi : \Omega \to \mathbb{R}^s$ a random vector, $p \in (0,1)$ a probability level, X a Banach space and $h : X \to \mathbb{R}^m$, $g : \mathbb{R}^s \to \mathbb{R}^m$ mappings defining a (functional) chance constraint

$$M = \{x \in X | P(\omega \in \Omega : h(x) \geq g(\xi(\omega))) \geq p\}. \tag{1}$$

Here, the first inequality sign has to be understood componentwise. We note that the subsequent result is in terms of path connectedness and hence implies topological connectedness.

Theorem 3 *Assume that the mapping h in (1) is continuous and has strictly quasiconcave components h_i. Furthermore, let the following constraint qualification be satisfied*

$$\operatorname{Im} h \cap (t\mathbf{1} + \mathbb{R}^m_+) \neq \emptyset \quad \forall t \in \mathbb{R}, \tag{2}$$

where 'Im' denotes the image and $\mathbf{1} = (1, \ldots, 1)$. Then, the constraint set M in (1) is (path-) connected.

Proof. Let $x^1, x^2 \in M$ be arbitrarily given. In order to show the result, we have to construct a path in M joining x^1 and x^2, i.e. a continuous function $\pi : [0,1] \to M$ with $\pi(0) = x^1$ and $\pi(1) = x^2$. We put $y^1 := h(x^1)$ and $y^2 := h(x^2)$. By assumption, it holds that

$$P(\omega \in \Omega : y^i \geq g(\xi(\omega))) \geq p \quad i = 1, 2. \tag{3}$$

Next, define $y^* \in \mathbb{R}^m$ by $y^*_i := \max\{y^1_i, y^2_i\}$ $(i = 1, \ldots, m)$. Now,

$$\varphi(t) := \begin{cases} (1-2t)y^1 + 2ty^* & \text{for } t \in [0, \frac{1}{2}] \\ 2(1-t)y^* + (2t-1)y^2 & \text{for } t \in (\frac{1}{2}, 1] \end{cases}$$

is a continuous function from $[0,1]$ to $\mathbb{R}^m$ with $\varphi(0) = y^1$ and $\varphi(1) = y^2$. Furthermore, (3) implies that

$$P(\omega \in \Omega : \varphi(t) \geq g(\xi(\omega))) \geq p \quad \forall t \in [0,1], \tag{4}$$

which is verified for $t \in [0, \frac{1}{2}]$ by using that $\varphi(t) \geq y^1$ and for $t \in (\frac{1}{2}, 1]$ by using that $\varphi(t) \geq y^2$.

We define the multifunction $K : \mathbb{R}^m \rightrightarrows X$ by $K(y) := \{x \in X | h(x) \geq y\}$ and check the following properties to hold for all $y \in \mathbb{R}^m$:

1. $K(y)$ is closed and convex
2. $K(y) \neq \emptyset$ and K is lower semicontinuous at y.

1. follows immediately from h being continuous and having strictly quasiconcave components. For 2., choose an arbitrary $\bar{y} \in \mathbb{R}^m$ and put

$$t := 1 + \max\{\bar{y}_i | \, i = 1, \ldots, m\}.$$

Then (2) yields the existence of some $x \in X$ such that $h(x) \geq t\mathbf{1}$ which, in components, entails that $h_i(x) \geq t > \bar{y}_i$ for $i = \{1, \ldots m\}$. In other words, $K(\bar{y}) \neq \emptyset$. Setting $Y = \mathbb{R}^m$, $J = \{1, \ldots m\}$ and $\alpha(x, y) = y - h(x)$, the same relation allows to apply Theorem 2 in order to derive the lower semicontinuity of K at $\bar{y}$.

With the continuous function φ constructed above, we have that $\Lambda := \varphi([0,1])$ is a compact subset of $\mathbb{R}^m$. Exploiting the properties 1. and 2. of the multifunction K, Michael's selection theorem (Th. 1) provides us with a continuous selection of K, i.e., with some continuous function $\sigma : \Lambda \to X$ such that $\sigma(y) \in K(y)$ for all $y \in \Lambda$. By definition of K, this means that

$$h(\sigma(y)) \geq y \text{ for all } y \in \Lambda. \tag{5}$$

As a composition of continuous functions, $\Psi : [0,1] \to X$ defined by $\Psi := \sigma \circ \varphi$ is continuous itself. Now, the implication

$$g(z) \leq \varphi(t) \Longrightarrow g(z) \leq h(\sigma(\varphi(t))) = h(\Psi(t)) \quad (z \in \mathbb{R}^s, t \in [0,1]),$$

which relies on (5), yields

$$P(\omega \in \Omega : h(\Psi(t)) \geq g(\xi(\omega))) \geq P(\omega \in \Omega : \varphi(t) \geq g(\xi(\omega))) \geq p \quad \forall t \in [0,1]$$

according to (4). Hence, $\Psi(t) \in M$ for all $t \in [0,1]$. Finally, we construct the desired path as

$$\pi(t) := \begin{cases} 3t\Psi(0) + (1-3t)x^1 & \text{for } t \in [0, \frac{1}{3}] \\ \Psi(3t-1) & \text{for } t \in (\frac{1}{3}, \frac{2}{3}] \\ 3(1-t)\Psi(1) + (3t-2)x^2 & \text{for } t \in (\frac{2}{3}, 1] \end{cases}.$$

Obviously, $\pi : [0,1] \to X$ is a continuous function with $\pi(0) = x^1$ and $\pi(1) = x^2$. It remains to show that $\pi(t) \in M$ for all $t \in [0,1]$. This is clear for $t \in (\frac{1}{3}, \frac{2}{3}]$ from the corresponding property of Ψ verified above. Hence, we check the case $t \in [0, \frac{1}{3}]$ (the last case $t \in (\frac{2}{3}, 1]$ following analogously). Now, exploiting the (strict) quasiconcavity of the components h_i and (5), one gets for $t \in [0, \frac{1}{3}]$ and $i = 1, \ldots, m$:

$$\begin{aligned} h_i(\pi(t)) &= h_i(3t\Psi(0) + (1-3t)x^1) \geq \min\{h_i(\Psi(0)), h_i(x^1)\} \\ &= \min\{h_i(\sigma(y^1)), y_i^1\} = y_i^1. \end{aligned}$$

As a consequence of (4) and of $y^1 = \varphi(0)$, this implies the relation

$$P(\omega \in \Omega : h(\pi(t)) \geq g(\xi(\omega))) \geq P(\omega \in \Omega : y^1 \geq g(\xi(\omega))) \geq p.$$

In other words, $\pi(t) \in M$ as was to be shown. ■

We note that in case of a single stochastic inequality $h(x) \geq g(\xi(\omega))$ (i.e., $m = 1$) the structure of M in (1) is completely determined by the underlying deterministic inequality $h(x) \geq \alpha$ for some appropriate $\alpha \in \mathbb{R}$. Indeed, introducing the distribution function of the transformed random variable $g \circ \xi$ as

$$F(t) = P(\omega \in \Omega \mid g(\xi(\omega)) \leq t),$$

one gets

$$M = \{x \in X \mid F(h(x)) \geq p\} = \{x \in X \mid h(x) \geq F^{-1}(p)\},$$

where, $F^{-1}(p) = \inf\{t \mid F(t) \geq p\}$. In particular, no constraint qualification like (2) is necessary then in order to derive the result of Theorem 3: clearly, the quasiconvexity of h allone guarantees M to be even convex (hence connected). Furthermore, X need not be restricted to a Banach space but may be a general topological vector space in the case of a single inequality. However, starting from two inequalities ($m \geq 2$), it is necessary to impose the constraint qualification (2) in order to derive connectedness of M (see Example 1 below).

3 Connectedness of probabilistic storage level constraints

For the purpose of illustration, we briefly discuss an application of Theorem 3 in the context of probabilistic storage level constraints. Here, one deals with a reservoir which is stochastically fed with some substance on the one hand and emptied in a controlled manner on the other hand. For instance, the reservoir might be a water basin or lake being randomly filled with rain water (e.g. [6]). It could also be a feed tank in continuous distillation where stochastic inflows from previous industrial processes have to be operated (e.g. [3]). Finally, one could also think of a battery storing electrical energy randomly provided by a solar panel.

Typically, some level constraint has to be satisfied in the reservoir. It is assumed that some optimization problem, including the rate of extracting substance from the reservoir as a decision variable, has to be solved for a future time horizon $[0, T]$. However, since the rate of inflow to the reservoir during $[0, T]$ figures as a stochastic parameter which is unknown at the time of decision, the level restrictions have to be formulated as probabilistic or chance constraints. To be more precise, we make the following assumptions:

- The time horizon is subdivided into s pieces: $0 = t_0 < t_1 < \cdots < t_s = T$
- $\xi = (\xi_1, \cdots, \xi_s)$ is a random vector defined on some probability space $(\Omega, \mathcal{A}, P)$ and the component ξ_i of which refers to the amount of inflow to the reservoir during the interval $[t_{i-1}, t_i]$.
- $x \in L^\infty[0, T] =: X$ is the rate of extracting substance from the reservoir.
- $\bar{l}$ is an upper limit for the amount of substance in the reservoir which is required to be met with probability level $p \in (0, 1)$ at times t_i.

We do not insist in positive extraction rates x here, hence we allow for pumping back substance into the reservoir. Denoting by $l(t, \xi, x)$ the amount of substance in the reservoir at time t (given ξ and x), one has

$$l(t_i, \xi, x) = l_0 + \sum\nolimits_{j=1}^{i} \xi_j - \int_0^{t_i} x(\tau) d\tau \quad (i = 1, \ldots, s),$$

where l_0 refers to the initial amount of substance (at $t = 0$). Then, the upper level chance constraint writes as

$$P(\omega \in \Omega : l(t_i, \xi(\omega), x) \leq \bar{l}) \geq p.$$

As in (1), we denote by M the subset of those $x \in X$ satisfying the chance constraint. Clearly, M can be given the description of (1) with

$$g_i(z) := \sum\nolimits_{j=1}^{i} z_j + l_0 - \bar{l}; \quad h_i(x) := \int_0^{t_i} x(\tau)d\tau \quad (i = 1, \ldots, s).$$

Since h is a continuous linear mapping, it suffices to check the constraint qualification (2) in order to apply Theorem 3. To this aim, let $t \in \mathbb{R}$ be arbitrarily given. Putting

$$x(\tau) := \begin{cases} t/t_1 & \text{for } \tau \in [0, t_1] \\ 0 & \text{for } \tau \in (t_1, T] \end{cases},$$

one has $x \in X$ and $h_i(x) = t$ for $i = 1, \ldots, s$, hence (2) is satisfied. One may deduce now the connectedness of the set M of feasible extraction controls.

4 The linear case and limiting examples

Some few examples shall illustrate the use and the limitations of Theorem 3 in the finite dimensional case. The most important application arises, of course, when h is linear. In that special case, the constraint qualification (2) becomes a particularly simple condition:

Lemma 4 *If $h : \mathbb{R}^n \to \mathbb{R}^m$ is linear, i.e. $h(x) = Ax$, where A is a (m, n)-matrix, then (2) is satisfied provided that the rows of A are positively linear independent.*

Proof. Due to $t_1\mathbf{1} + \mathbb{R}^m_+ \subseteq t_2\mathbf{1} + \mathbb{R}^m_+$ for $t_2 \leq t_1$ and because of the linearity of β, it suffices to verify (2) for one single $t > 0$. The positive linear independence of the rows of A is equivalent to the existence of some $\xi \in \mathbb{R}^m$ with $[A^T\xi]_i > 0$, $(i = 1, \ldots, n)$. This makes the expression $\tau := t/\min\{[A^T\xi]_i \mid i = 1, \ldots, n\} > 0$ well defined. It follows that $A(\tau\xi) = \tau A\xi \geq t\mathbf{1}$, which implies (2). ∎

Of course, the linear independence of the rows of A implies their positive linear independence, but it is much too strong in general. For instance, Lemma 4 allows also to consider the situation $m > n$ as well which would be excluded when relying on linear independence of rows. The following example illustrates the link between the constraint qualification (2) and the connectedness of the constraint set:

Example 1 *In (1), let $X = \mathbb{R}$, $g =$ identity mapping on $\mathbb{R}^2$, $h(x) = (x, -x)^T$, $\xi \sim$ discrete distribution with mass 0.5 on each of the two points $(-1, 1)$, $(1, -1)$ and $p = 0.5$. According to the defined distribution of ξ, one has*

$$P((\xi_1, \xi_2) \leq (x, -x)) = \begin{cases} 0.5 & \textit{if } x = -1 \textit{ or } x = 1 \\ 0 & \textit{else} \end{cases}.$$

Consequently, the constraint set reduces to $M = \{-1, 1\}$ *which is obviously disconnected. This is reflected by the fact that the two 'rows' of the 'matrix'* $\binom{1}{-1}$ *defining the linear mapping* h *are positively linear dependent. Changing* h *to* $(x, x)^T$*, the rows of* $\binom{1}{1}$ *become positively linear independent now (although not linearly independent). Hence, Lemma 4 guarantees* M *to become a connected set now. Indeed,* $M = [1, \infty)$ *then.*

The next example demonstrates that additional deterministic constraints cannot be added to the definition of M in (1) without the risk of loosing connectedness:

Example 2 *Take* g*,* p *and the distribution of* ξ *as in the foregoing example, but define* $X = \mathbb{R}^2$ *and* $h = g$*. If we add to the definition of* M *in (1) the deterministic linear constraint* $x_1 + x_2 \leq 0$*, then* $M = \{(-1, 1), (1, -1)\}$ *becomes disconnected although the constraint qualification (2) is satisfied (since* h *is the identity mapping, see condition of Lemma 4) and, hence, the pure chance constraint is connected. The reason for violation of connectedness here is the fact that the intersection of connected sets needs no longer be so.*

Despite the last example, additional deterministic constraints of the form $\tilde{h}(x) \geq 0$ can well be incorporated into the result of Theorem 3 as long as $\tilde{h}$ is continuous and has strictly quasiconcave components (similar to h). In that case the joint (probabilistic and deterministic) constraint set M can be written as

$$M = \{x \in X | P(\omega \in \Omega : H(x) \geq G(\xi(\omega))) \geq p\},$$

where $H = (h, \tilde{h})$ and $G = (g, 0)$. Now, the constraint qualification (2) has just to be reformulated in terms of $(h, \tilde{h})$ rather than h alone.. Doing so in the last example, one arrives at H being a linear mapping defined by the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix},$$

the rows of which are positively linearly dependent. This explains the constraint set M becoming disconnected in the example. Changing the sign of constraint, i.e., $x_1 + x_2 \geq 0$, connectedness can be guaranteed by our result. Indeed, one has $M = [\{(-1, 1)\} + \mathbb{R}^2] \cup [\{(1, -1)\} + \mathbb{R}^2]$, which is connected although not convex.

References

[1] B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer, *Nonlinear Parametric Optimization*, Akademie-Verlag, Berlin, 1982.

[2] J.R. Birge and F. Louveaux, *Introduction to Stochastic Programming,* Springer, Berlin, 1999.

[3] R. Henrion and A. Möller, Optimization of a continuous distillation process under random inflow rate, Preprint No. 00-4 of the DFG-research program 'Echtzeit-Optimierung großer Systeme', submitted to: *Computers & Mathematics with Applications.*

[4] P. Kall and S.W. Wallace, *Stochastic Programming,* Wiley, Chichester, 1994.

[5] E. Michael, Continuous selections I, *Annals of Mathematics* **63** (1956), 361-182

[6] A. Prékopa and T. Szántai, On optimal regulation of a storage level with application to the water level regulation of a lake, European Journal of Operations Research **3** (1978), 175-189.

[7] A. Prékopa, *Stochastic Programming,* Kluwer, Dordrecht, 1995.