Structural Properties of Linear Probabilistic Constraints *

R. Henrion
Weierstrass Institute, Mohrenstr. 39, 10117 Berlin, Germany

Abstract

The paper provides a structural analysis of the feasible set defined by linear probabilistic constraints. Emphasis is laid on single (individual) probabilistic constraints. A classical convexity result by Van de Panne/Popp and Kataoka is extended to a broader class of distributions and to more general functions of the decision vector. The range of probability levels for which convexity can be expected is exactly identified. Apart from convexity, also nontriviality and compactness of the feasible set are precisely characterized at the same time. The relation between feasible sets with negative and with nonnegative right-hand side is revealed. Finally, an existence result is formulated for the more difficult case of joint probabilistic constraints.

Key words.
probabilistic constraints, stochastic programming, chance constraints,
stochastic optimization

Mathematics Subject Classification (2000):
90C15

1 Introduction

Many optimization problems in engineering sciences involve stochastic linear constraints of the form

$$\Xi x \leq \eta, \tag{1}$$

where x is an n-dimensional decision vector, Ξ is a stochastic matrix of order (m, n) and η is a fixed or stochastic random vector of dimension m (see [12], for instance). Typically, 'here-and-now' decisions have to be taken, which means that the random parts of (1) are observed only after deciding upon x. Thus, no matter how x is chosen, a sure feasibility with respect to (1) cannot be guaranteed. However, depending on the distribution of Ξ (and η whenever stochastic), it is possible to choose x in a way to keep the probability of violating (1) small. More precisely, one can turn (1) into a probabilistic constraint

$$P(\Xi x \leq \eta) \geq p, \tag{2}$$

*This work was supported by the DFG Research Center MATHEON *Mathematics for key technologies* in Berlin

where P is a probability measure and $p \in [0,1]$ is some probability level (typically close to 1) at which (1) is required to hold. Inequality (2) is also referred to as a *joint probabilistic constraint* as it takes into account the probability of the entire system (1) to be satisfied. In general, joint probabilistic constraints are difficult to handle and both their algorithmic treatment and their theoretical investigation keep posing a lot of challenging questions (see [9] for a comprehensive introduction and [10] for a review on recent work in this area). It is much easier, although not justified in all situations, to turn each single inequality of (1) into an *individual probabilistic constraint* as follows:

$$P\left(\langle \xi_i, x\rangle \leq \eta_i\right) \geq p_i, \quad (i = 1, \dots, m). \tag{3}$$

Here, the ξ_i refer to the rows of Ξ and now the probability levels may differ for each constraint.

For algorithmic purposes it is of much interest to know whether or not the set of feasible decisions x satisfying (3) is convex. As the intersection of convex sets remains convex, this issue boils down to the investigation of a single linear probabilistic constraint

$$M = \{x \in \mathbb{R}^n | P\left(\langle \xi, x\rangle \leq \eta\right) \geq p\},$$

where ξ is an n-dimensional random vector and η is a scalar (possibly random). The convexity of M has been investigated first in the classical papers by Van de Panne and Popp [8] and by Kataoka [5]. They have shown that M is a convex subset of $\mathbb{R}^n$ provided that ξ has a nondegenerate multivariate normal distribution and that $p \geq 0.5$. This frequently cited result leaves open a lot of questions. First, one could ask about distributions different from normal ones or about more general functions of x under which the same result can be maintained. Second, it is clear that the feasible set M becomes smaller when the level p is increased towards 1. Hence, the important observation that M is convex for p large enough has to be coupled with the question of nontriviality because the empty set is convex too. Third, also large sets like $\mathbb{R}^n$ may be convex. This raises the question if there exists a range of small values of p which guarantees convexity as well. Finally, apart from convexity and triviality, compactness of M is another issue of theoretical and algorithmic interest. Nonempty and compact feasible sets guarantee the existence of solutions and allow to derive stability results for solutions when the usually unknown distribution of ξ has to be approximated on the basis of estimations or historical observations (see [2]).

The purpose of this paper is to provide a detailed structural analysis to linear chance constraints and to give a fairly precise answer to the questions posed. The classical results of [8] and [5] can be extended to the class of elliptically symmetric distributions and to certain component-wise convex mappings of x. In the classical setting of normal distributions, it will be possible to exactly identify the range of p-values for which convexity, triviality and compactness (or nonconvexity, nontriviality and unboundedness) hold true. It is interesting to observe, that these results strongly depend on whether the right-hand side η is negative or nonnegative. Under this case distinction, all structural results become rather different and seemingly independent. However, they are not as independent as they might look like. Roughly speaking, the first main result of this paper states that, for negative right-hand side and large values of p the feasible set looks like the complement of the feasible set for nonnegative right-hand side and small values of p. In the more demanding situation of optimization problems involving joint probabilistic constraints as in (2), an existence theorem can be derived from the case of single constraints. More precisley, this theorem allows exactly to calculate a critical p- level above which compactness and nonemptiness of a joint probabilistic constraint can be guaranteed. Such result is not only interesting with respect to the existence of solutions but also concerning stability of solution sets under perturbation (approximation) of the given probability distribution.

2 Results

In the following, we shall consider constraint sets

$$M_p^\alpha := \{x \in \mathbb{R}^n | P(\langle q(x), \xi \rangle \leq \alpha) \geq p\} \quad (\alpha \in \mathbb{R}, p \in (0,1)). \tag{4}$$

Here, ξ is an s-dimensional random vector defined on some probability space $(\Omega, \mathcal{A}, P)$ and $q : \mathbb{R}^n \to \mathbb{R}^s$ is a mapping from the space of decision vectors to the space of realizations of the random vector. The indices α and p shall emphasize the fact that we are going to analyze the structure of the feasible set as a function of the right-hand side of the considered stochastic inequality and of the probability level p. Putting $q(x) = x$, one gets back to the classical linear probabilistic constraint set M_p^α with deterministic right-hand side. Choosing $q(x) = (x, -1)$ and considering the extended $(s+1)$-dimensional random vector (ξ, η), M_p^0 recovers the constraint set with stochastic right-hand side (see introduction). In this latter case, q is an affine linear mapping which will figure as an assumption in several subsequent results. As an immediate consequence of the definition (4), one has the following properties:

$$M_{p_1}^\alpha \subseteq M_{p_2}^\alpha \qquad \forall \alpha \in \mathbb{R}\ \forall p_1, p_2 \in (0,1) : p_1 \geq p_2 \tag{5}$$

$$q^{-1}(0) \subseteq M_p^\alpha \qquad \forall \alpha \geq 0\ \forall p \in (0,1) \tag{6}$$

$$q^{-1}(0) \subseteq \left(M_p^\alpha\right)^c \qquad \forall \alpha < 0\ \forall p \in (0,1) \tag{7}$$

Moreover, the M_p^α are closed subsets of $\mathbb{R}^n$ under mild assumptions. Indeed, we may refer to the following consequence of a general closedness characterization provided in [11] (Prop. 3.1), where we keep the meaning of ξ and P:

Lemma 2.1 *Let $g : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^m$ be a vector-valued mapping with lower semicontinuous (in both variables simultaneously) components. Then, the set $\{x \in \mathbb{R}^n | P(g(x,\xi) \leq 0) \geq p\}$ is closed.*

Corollary 2.1 *If in (4), q is a mapping with lower semicontinuous components, then M_p^α is closed for all $\alpha \in \mathbb{R}$ and all $p \in (0,1)$.*

2.1 On the relation between positive and negative right-hand side

Before investigating properties of M_p^α, like convexity, nontriviality and compactness, we want to identify the structural relation between constraint sets with positive and negative right-hand side. The following theorem tells us that, up to closure and translation, the sets M_p^α are identical to the complements of the 'dual' sets $M_{1-p}^{-\alpha}$. Convexity and compactness are examples for properties which are not affected by translation or closure.

Theorem 2.1 *Let the distribution of ξ be absolutely continuous with respect to the Lebesgue measure, and let the support of ξ be all of $\mathbb{R}^s$. Furthermore, assume that q is a surjective, affine linear mapping. Then, there exists some $d \in \mathbb{R}^n$ such that*

$$M_p^\alpha = \{d\} - \mathrm{cl}\left\{\left(M_{1-p}^{-\alpha}\right)^c\right\} \quad \forall \alpha \neq 0\ \forall p \in (0,1).$$

Proof. We fix arbitrary $\alpha \neq 0$, $p \in (0,1)$ and start by observing that the function

$$x \longmapsto P(\langle q(x), \xi \rangle \leq \alpha) \tag{8}$$

is continuous at each $x \notin q^{-1}(0)$. Indeed, this condition, together with the fact that q is continuous, ensures that the set-valued mapping

$$T_y := \{u \in \mathbb{R}^s | \langle q(y), u\rangle \leq \alpha\}$$

satisfies $\lim_{y\to x} T_y = T_x$. Here, the set convergence is taken in the Kuratowski-Painlevé sense. Along with the assumption, that ξ has an absolutely continuous distribution, this ensures that $\lim_{y\to x} P(\xi \in T_y) = P(\xi \in T_x)$, whenever all the T_y and T_x are closed and convex (see [7], Th. 3, Lemma 1 and Proof of Th. 4).

To proceed with the proof of our Theorem, we may assume that $q(x) = Ax + b$ for some matrix A having full rank. Put

$$d := -2A^T \left(AA^T\right)^{-1} b.$$

As a consequence, one has that $-q(x) = q(d-x)$ for all $x \in \mathbb{R}^n$ and, in particular that $x \in q^{-1}(0)$ if and only if $d - x \in q^{-1}(0)$. For arbitrary $x \notin q^{-1}(0)$, the following equivalences hold true:

$$\begin{aligned} P(\langle q(x), \xi\rangle \leq \alpha) \geq p \quad &\Leftrightarrow \quad P(\langle q(x), \xi\rangle > \alpha) \leq 1-p \\ &\Leftrightarrow \quad P(-\langle q(x), \xi\rangle < -\alpha) \leq 1-p \\ &\Leftrightarrow \quad P(\langle q(d-x), \xi\rangle \leq -\alpha) \leq 1-p. \end{aligned} \tag{9}$$

Here, the last equivalence relies on the fact that $q(d-x) \neq 0$, so that $\langle q(d-x), \cdot\rangle = -\alpha$ defines a hyperplane in $\mathbb{R}^s$, which has probability zero by our assumption on the distribution of ξ. Next, we verify the following identity:

$$\mathrm{cl}\left\{\left(M_{1-p}^{-\alpha}\right)^c\right\} = \{z \in \mathbb{R}^n | P(\langle q(z), \xi\rangle \leq -\alpha) \leq 1-p\} \quad \forall z \notin q^{-1}(0). \tag{10}$$

For $z \in \mathrm{cl}\left\{\left(M_{1-p}^{-\alpha}\right)^c\right\}$, there exists a sequence $z_n \to z$ such that

$$P(\langle q(z_n), \xi\rangle \leq -\alpha) < 1-p.$$

This entails the inclusion '$\subseteq$' in (10) via the continuity of the function (8). For the reverse inclusion, let z be given such that $z \notin q^{-1}(0)$ and

$$P(\langle q(z), \xi\rangle \leq -\alpha) \leq 1-p.$$

With

$$z_n := z - \frac{\operatorname{sgn}\alpha}{n} A^T \left(AA^T\right)^{-1} (Az+b),$$

one gets that $z_n \to z$ and

$$q(z_n) = Az_n + b = Az + b - \frac{\operatorname{sgn}\alpha}{n}(Az+b) = \begin{cases} (1-n^{-1})q(z) & \text{if } \alpha > 0 \\ (1+n^{-1})q(z) & \text{if } \alpha < 0. \end{cases}$$

Consequently, in case that $\alpha > 0$, one arrives at the inclusion

$$\{u \in \mathbb{R}^s | \langle q(z_n), u\rangle \leq -\alpha\} = \{u \in \mathbb{R}^s | \langle q(z), u\rangle \leq -\alpha(1-n^{-1})^{-1}\} \subseteq \{u \in \mathbb{R}^s | \langle q(z), u\rangle \leq -\alpha\}.$$

Thus,

$$\begin{aligned} 1-p &\geq P\left(\langle q(z), \xi\rangle \leq -\alpha\right) \\ &= P\left(\langle q(z_n), \xi\rangle \leq -\alpha\right) + P\left(-\alpha(1-n^{-1})^{-1} < \langle q(z_n), \xi\rangle \leq -\alpha\right). \end{aligned}$$

Now, since the strip

$$\{u \in \mathbb{R}^s | -\alpha(1-n^{-1})^{-1} < \langle q(z_n), u\rangle \leq -\alpha\}$$

has a nonempty interior, its probability must be strictly positive according to our assumption that the support of ξ is all of $\mathbb{R}^s$. Thus, we get

$$1-p > P\left(\langle q(z_n), \xi\rangle \leq -\alpha\right)$$

which amounts to saying that $z_n \in \left(M_{1-p}^{-\alpha}\right)^c$. An analogous argumentation applies to the case $\alpha < 0$ upon using the respective definition of z_n. This establishes (10).

Applying (10) to (9) with $z = d - x \notin q^{-1}(0)$, we may summarize the preceding considerations in the form

$$x \in M_p^\alpha \backslash q^{-1}(0) \Longleftrightarrow x \in \left[\{d\} - \mathrm{cl}\left\{\left(M_{1-p}^{-\alpha}\right)^c\right\}\right] \backslash q^{-1}(0). \tag{11}$$

In order to finish the proof, it remains to verify the equivalence

$$x \in M_p^\alpha \cap q^{-1}(0) \Longleftrightarrow x \in \left[\{d\} - \mathrm{cl}\left\{\left(M_{1-p}^{-\alpha}\right)^c\right\}\right] \cap q^{-1}(0). \tag{12}$$

If $x \in M_p^\alpha \cap q^{-1}(0)$, then also $d - x \in q^{-1}(0)$ and $\alpha \geq 0$ by (7). Since $\alpha \neq 0$, it follows that $-\alpha < 0$ and $d - x \in \left(M_{1-p}^{-\alpha}\right)^c$, again by (7). This proves the implication '$\Longrightarrow$' in (12). Conversely, let

$$x \in \left[\{d\} - \mathrm{cl}\left\{\left(M_{1-p}^{-\alpha}\right)^c\right\}\right] \cap q^{-1}(0).$$

Once more, $d - x \in q^{-1}(0)$. By definition, there is a sequence $x_n \to d - x$ with $x_n \in \left(M_{1-p}^{-\alpha}\right)^c$. Assume first, that there is a subsequence of x_n, which we do not relabel, such that $x_n \notin q^{-1}(0)$. Then, also $d - x_n \notin q^{-1}(0)$, so that we can apply (11) to $d - x_n$ rather than x. This yields that $d - x_n \in M_p^\alpha$. On the other hand, M_p^α is closed according to Corollary 2.1. It follows that

$$d - x_n \to x \in M_p^\alpha \cap q^{-1}(0),$$

which establishes the reverse implication in (12) for a special case. It remains to check the case when $d - x_n \in q^{-1}(0)$ for all n. Then, also $x_n \in q^{-1}(0)$ for all n. The assumption $\alpha \leq 0$ would lead to the contradiction $x_n \in M_{1-p}^{-\alpha}$ via (6). So, $d - x_n \in M_p^\alpha$, again by (6). The same closedness argument as in the first special case yields that $x \in M_p^\alpha \cap q^{-1}(0)$. This completes the proof.

■

The following example illustrates, why we have to insist on the condition $\alpha \neq 0$ in Theorem 2.1:

Example 2.1 *In dimension one, let $q(x) = x$ and ξ have a standard normal distribution. Then, $M_{0.5}^0 = \mathbb{R}$. All assumptions of Theorem 2.1 are met except that $\alpha = 0$. If the theorem would hold true, there should exist some $d \in \mathbb{R}$ such that*

$$M_{0.5}^0 = \{d\} - \mathrm{cl}\left\{\left(M_{0.5}^0\right)^c\right\} = \{d\} - \mathrm{cl}\,\emptyset = \emptyset,$$

which is a contradiction.

2.2 Convexity

We recall the class of elliptically symmetric distributions, whose density (if it exists) is given by

$$f(x) = (\det \Sigma)^{-1/2} g\left(\left\langle x-\theta, \Sigma^{-1}(x-\theta)\right\rangle\right),$$

where Σ is a positive definite matrix and g is some nonnegative function. In particular, the s-dimensional normal distribution belongs to this class with mean vector θ, covariance matrix Σ and

$$g(t) = (2\pi)^{-s/2} \exp(-t/2).$$

However, the class of elliptically symmetric distributions is much broader than just multivariate normal ones and incorporates, for instance, multivariate versions of student or exponential distributions ([1], [3]). The characteristic function of an elliptically symmetric distribution has the form

$$\phi(t) = \exp(i\langle t, \theta\rangle)\, h(\langle t, \Sigma t\rangle)$$

for some scalar function h, called the 'characteristic generator' of this distribution.

In the following, we use the symbol $\|\cdot\|_C$ for the norm induced by a positive definite matrix C, i.e.: $\|\cdot\|_C = \sqrt{\langle x, Cx\rangle}$. Moreover, for a 1-dimensional distribution function F we define its p-quantile as

$$F^{-1}(p) = \inf\{t | F(t) \geq p\}.$$

Lemma 2.2 *In (4), let q be arbitrary and let ξ have an elliptically symmetric distribution with parameters Σ, θ, where Σ is positive definite. Denote by h its characteristic generator. Then*

$$M_p^\alpha = \{x \in \mathbb{R}^n | F^{-1}(p)\, \|q(x)\|_\Sigma + \langle \theta, q(x)\rangle \leq \alpha\},$$

where F is the 1-dimensional distribution function induced by the characteristic function $\phi(\tau) := h\left(\tau^2\right)$.

Proof. The characteristic function of ξ is

$$\phi_\xi(t) = \exp(i\langle t, \theta\rangle)\, h\left(\|t\|_\Sigma^2\right).$$

Let $x \in \mathbb{R}^n \backslash q^{-1}(0)$ be arbitrary. Then, the scaled random variable

$$\eta_x := \frac{\langle q(x), \xi - \theta\rangle}{\sqrt{\langle q(x), \Sigma q(x)\rangle}} = \left\langle \frac{q(x)}{\|q(x)\|_\Sigma}, \xi \right\rangle - \left\langle \frac{q(x)}{\|q(x)\|_\Sigma}, \theta \right\rangle$$

is a well-defined affine linear transformation of ξ. Following the general calculus rule

$$\phi_{\langle c, \xi\rangle + d}(\tau) = \exp(i\tau d) \cdot \phi_\xi(\tau c)$$

for characteristic functions, that of η_x calculates as

$$\phi_{\eta_x}(\tau) = \exp\left(-i\tau \frac{\langle q(x), \theta\rangle}{\|q(x)\|_\Sigma}\right) \phi_\xi\left(\frac{\tau}{\|q(x)\|_\Sigma} q(x)\right) = h\left(\tau^2\right).$$

In particular, the distribution of η_x does not depend on x. Its distribution function is given by F as introduced in the statement of this lemma. It follows that

$$\begin{aligned} P(\langle q(x), \xi\rangle \leq \alpha) \geq p \quad &\Leftrightarrow \quad P\left(\eta_x \leq \frac{\alpha - \langle \theta, q(x)\rangle}{\|q(x)\|_\Sigma}\right) \geq p \Leftrightarrow F\left(\frac{\alpha - \langle \theta, q(x)\rangle}{\|q(x)\|_\Sigma}\right) \geq p \\ &\Leftrightarrow \quad F^{-1}(p)\,\|q(x)\|_\Sigma + \langle \theta, q(x)\rangle \leq \alpha. \end{aligned}$$

Now, the assertion results from (6) and (7) upon observing that the last inequality holds true for all $x \in q^{-1}(0)$ if $\alpha \geq 0$ and is violated for all $x \in q^{-1}(0)$ if $\alpha < 0$.

■

Proposition 2.1 *Let, in addition to the setting of Lemma 2.2, one of the following assumptions hold true:*

- *q is affine linear*

  or

- *q has nonnegative, convex components, $\theta_i \geq 0$ for $i = 1, \ldots, s$ and all elements of Σ are nonnegative..*

Then, M_p^α is convex for all $\alpha \in \mathbb{R}$ and all $p > 0.5$. If, moreover, the random vector ξ in Lemma 2.2 has a strictly positive density, then M_p^α is convex for all $\alpha \in \mathbb{R}$ and all $p \geq 0.5$.

Proof. By Lemma 2.2, we are done if we can show that both functions

$$\langle \theta, q(x)\rangle \quad \text{and} \quad F^{-1}(p)\,\|q(x)\|_\Sigma$$

are convex. This is obvious for $\langle \theta, q(x)\rangle$ without restrictions on θ in case that q is affine linear and for θ with nonnegative components in case that the components of q are convex. Let us turn to the second term now: Since F is a one-dimensional symmetric distribution function, it follows that $F(0) = 0.5$. Therefore, $F^{-1}(p) \geq 0$ for $p > 0.5$ and also $F^{-1}(p) \geq 0$ for $p = 0.5$, in the case that F has a strictly positive density. It remains to verify thus, that $\|q(\cdot)\|_\Sigma$ is a convex function. This is evident in case that q is affine linear. For the alternative case, recall that, for any fixed $x \notin q^{-1}(0)$, the optimization problem

$$\max\{\langle q(x), y\rangle \mid \|y\|_{\Sigma^{-1}} = 1\}$$

has the solution

$$y^* = \|q(x)\|_\Sigma^{-1}\, \Sigma q(x).$$

Since, by assumption, all components of q and all elements of Σ are nonnegative, the components of y^* are nonnegative too. This allows to write that

$$\|q(x)\|_\Sigma = \langle q(x), y^*\rangle = \max\{\langle q(x), y\rangle \mid \|y\|_{\Sigma^{-1}} = 1\} = \max\{\langle q(x), y\rangle \mid \|y\|_{\Sigma^{-1}} = 1, y \in \mathbb{R}^s_+\}.$$

for all $x \notin q^{-1}(0)$. The same identity

$$\|q(x)\|_\Sigma = \max\{\langle q(x), y\rangle \mid \|y\|_{\Sigma^{-1}} = 1, y \in \mathbb{R}^s_+\}$$

holds trivially true in case that $x \in q^{-1}(0)$, hence it is valid for all $x \in \mathbb{R}^n$. For $y \in \mathbb{R}^s_+$, $\langle q(\cdot), y\rangle$ is convex by the assumed convexity of the components of q. Summarizing, $\|q(\cdot)\|_\Sigma$ is convex as a maximum of convex functions $\langle q(\cdot), y\rangle$.

■

When reducing Proposition 2.1 to a nondegenerate multivariate normal distribution of ξ, then its first statement evidently recovers the classical convexity result of [5],[8] with random or deterministic right-hand side (see introduction and beginning of Section 2). The first statement of Proposition 2.1 was shown in [4] based on the concept of so-called α-nuclei. In contrast, our proof essentially relies on the representation Lemma 2.2. This representation allows, in the second statement of Proposition 2.1, to generalize the convexity result to nonlinear functions q of the decision vector. A different extension of the classical results to the class of log-concave symmetric distributions has been obtained in [6]. As the elliptically symmetric distributions considered here, the log-concave symmetric distributions also contain multivariate normal distributions (but apart from it also uniform distributions over symmetric, convex, compact sets).

From now on we shall assume, for simplicity, that the random vector ξ has a nondegenerate multivariate normal distribution with mean vector μ and (positive definite) covariance matrix Σ: $\xi \sim \mathcal{N}(\mu, \Sigma)$. Then, by Lemma 2.2,

$$M_p^\alpha = \{x \in \mathbb{R}^n | \Phi^{-1}(p) \, \|q(x)\|_\Sigma + \langle \mu, q(x)\rangle \leq \alpha\}, \tag{13}$$

where Φ denotes the distribution function of the one-dimensional standard normal distribution and $\Phi^{-1}(p)$ its p-quantile.

Proposition 2.1 tells us for which range of p-values convexity of the constraint set may be expected. It does not imply, however, nonconvexity of this set for the remaining p-values. The following proposition clarifies, under which circumstances nonconvexity may be derived.

Proposition 2.2 *Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ with positive definite Σ and let q be a surjective affine linear mapping. Then, M_p^α is nonconvex in any of the following two situations:*

$$\alpha < 0, \; p < 0.5$$

or

$$\alpha \geq 0, \; \Phi\left(-\|\mu\|_{\Sigma^{-1}}\right) < \; p < 0.5.$$

Proof. First, let $\alpha < 0$ and $p < 0.5$, whence $\Phi^{-1}(p) < 0$. We choose $\delta \neq 0$ such that $\langle \delta, \mu\rangle = 0$. By surjectivity of q, there is some h such that $q(h) = q(0) + \delta$. Again by surjectivity of q, we may choose some $x^* \in q^{-1}(0)$. By virtue of (7), one has that $x^* \notin M_p^\alpha$. For $t \in \mathbb{R}$, put $x_t := x^* + th$. The affine linearity of q implies that

$$q(x^t) = q(x^*) + t(q(h) - q(0)) = q(x^*) + t\delta.$$

Then,

$$\Phi^{-1}(p) \, \|q(x_t)\|_\Sigma + \langle \mu, q(x_t)\rangle = \Phi^{-1}(p) \, \|q(x^*) + t\delta\|_\Sigma + \langle \mu, q(x^*)\rangle \, .$$

Since $\delta \neq 0$ and $\Phi^{-1}(p) < 0$, it follows that

$$\lim_{t \to \infty} \Phi^{-1}(p) \, \|q(x_t)\|_\Sigma + \langle \mu, q(x_t)\rangle = \lim_{t \to -\infty} \Phi^{-1}(p) \, \|q(x_t)\|_\Sigma + \langle \mu, q(x_t)\rangle = \infty.$$

Consequently, for $|t|$ large enough, one has

$$\Phi^{-1}(p) \, \|q(x_t)\|_\Sigma + \langle \mu, q(x_t)\rangle \leq \alpha,$$

which means that $x_t \in M_p^\alpha$ according to (13). In particular, there is some $\tau > 0$ such that $x_\tau, x_{-\tau} \in M_p^\alpha$. On the other hand,

$$\frac{x_\tau + x_{-\tau}}{2} = x^* \notin M_p^\alpha.$$

Therefore, M_p^α is not convex.

Now, let $\alpha \geq 0$ and $\Phi\left(-\|\mu\|_{\Sigma^{-1}}\right) < p < 0.5$. In particular, $\mu \neq 0$, beacause otherwise $\Phi\left(-\|\mu\|_{\Sigma^{-1}}\right) = 0.5$. For each $t \in \mathbb{R}$, the surjectivity of q allows to choose some y_t such that $q(y_t) = t\Sigma^{-1}\mu$. Then,

$$\begin{aligned}\Phi^{-1}(p)\,\|q(y_t)\|_\Sigma + \langle \mu, q(y_t)\rangle &= \Phi^{-1}(p)t\,\|\mu\|_{\Sigma^{-1}} + t\,\|\mu\|_{\Sigma^{-1}}^2 \\ &= \left(\Phi^{-1}(p) + \|\mu\|_{\Sigma^{-1}}\right) t\,\|\mu\|_{\Sigma^{-1}} \to_{t\to\infty} \infty,\end{aligned}$$

where the convergence towards infinity relies on the fact that $\mu \neq 0$ and on the fact that the expression in parentheses is strictly positive by our assumption on the admissible range of p. Hence, for t large enough, the expression above will exceed α. In other words, by (13), for t large enough, $y_t \notin M_p^\alpha$. We fix such a point and call it $\bar{x}$. Now, we may repeat exactly the same argumentation as in the first part of this proof but with x^* replaced by $\bar{x}$. This allows again to find points $x_\tau, x_{-\tau} \in M_p^\alpha$ such that

$$\frac{x_\tau + x_{-\tau}}{2} = \bar{x} \notin M_p^\alpha,$$

and hence, convexity of M_p^α is violated once more.

■

2.3 Non-emptiness and compactness

So far, we have characterized the convexity of the constraint set. It has to be taken into account, however, that M_p^α might be trivially convex in being identical either to the empty set or to the whole space. Therefore, a characterization of triviality is of interest as well.

Proposition 2.3 *Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ with positive definite Σ. Then,*

$$\begin{aligned} M_p^\alpha &= \mathbb{R}^n \quad \forall \alpha \geq 0 \quad \forall p \leq \Phi\left(-\|\mu\|_{\Sigma^{-1}}\right) \\ M_p^\alpha &= \emptyset \quad \forall \alpha < 0 \quad \forall p \geq \Phi\left(\|\mu\|_{\Sigma^{-1}}\right). \end{aligned}$$

Moreover, if q is surjective, then

$$\begin{aligned} M_p^\alpha &\neq \emptyset \quad \forall \alpha \geq 0 \quad \forall p \in (0,1) \\ M_p^\alpha &\neq \emptyset \quad \forall \alpha < 0 \quad \forall p < \Phi\left(\|\mu\|_{\Sigma^{-1}}\right). \end{aligned}$$

Proof. A generalized version of the Cauchy-Schwarz inequality (for symmetric, positive definite matrices) yields the relation

$$|\langle \mu, q(x)\rangle| \leq \|q(x)\|_\Sigma\,\|\mu\|_{\Sigma^{-1}}. \tag{14}$$

From here, for arbitrary $x \in \mathbb{R}^n$, one obtains the following pair of inequalities by case distinction:

$$\Phi^{-1}(p)\,\|q(x)\|_\Sigma + \langle \mu, q(x)\rangle$$
$$\begin{cases} \leq \|q(x)\|_\Sigma\left(\Phi^{-1}(p) + \|\mu\|_{\Sigma^{-1}}\right) \leq 0 \leq \alpha & \forall \alpha \geq 0\ \forall p \leq \Phi\left(-\|\mu\|_{\Sigma^{-1}}\right) \\ \geq \|q(x)\|_\Sigma\left(\Phi^{-1}(p) - \|\mu\|_{\Sigma^{-1}}\right) \geq 0 > \alpha & \forall \alpha < 0\ \forall p \geq \Phi\left(\|\mu\|_{\Sigma^{-1}}\right). \end{cases}$$

By virtue of (13), this proves the first part of our Corollary. The first statement of the second part of the corollary is evident from (6) because $q^{-1}(0) \neq \emptyset$ by the assumed surjectivity of q. Concerning the last statement, define for each $t > 0$ some x_t such that $q(x_t) = -t\Sigma^{-1}\mu$ (which is possible again by surjectivity of q). For any $p < \Phi\left(\|\mu\|_{\Sigma^{-1}}\right)$ and any α, it follows that

$$\frac{\alpha - \langle \mu, q(x_t)\rangle}{\|q(x)\|_\Sigma} = \frac{\alpha + t\,\|\mu\|^2_{\Sigma^{-1}}}{t\,\|\mu\|_{\Sigma^{-1}}} \underset{t\to\infty}{\rightarrow} \|\mu\|_{\Sigma^{-1}} > \Phi^{-1}(p).$$

Consequently, there is some x_t such that

$$\Phi^{-1}(p)\,\|q(x_t)\|_\Sigma + \langle \mu, q(x_t)\rangle < \alpha.$$

By Lemma (13), this amounts to saying that $x_t \in M_p^\alpha$.

■

Remark 2.1 *Note that the very first statement of Proposition 2.3 confirms that, for $\alpha \geq 0$, M_p^α is convex not just for $p \geq 0.5$ according to Proposition 2.1 but also for $p \leq \Phi\left(-\|\mu\|_{\Sigma^{-1}}\right)$.*

For algorithmic purposes, not only convexity of the constraint set is of interest but also its compactness. This, together with the non-emptiness characterized in Proposition 2.3, will guarantee the existence of solutions.

Proposition 2.4 *Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ with positive definite Σ. Moreover, let $q : \mathbb{R}^n \to \mathbb{R}^n$ be a homeomorphism (i.e., a bijective mapping such that q and q^{-1} are continuous). Then, for any $\alpha \in \mathbb{R}$, M_p^α is unbounded whenever $p < \Phi\left(\|\mu\|_{\Sigma^{-1}}\right)$ and compact whenever $p > \Phi\left(\|\mu\|_{\Sigma^{-1}}\right)$. If $p = \Phi\left(\|\mu\|_{\Sigma^{-1}}\right)$, then M_p^α is unbounded in the case that $\alpha \geq 0$ and is compact (actually empty) in the case that $\alpha < 0$.*

Proof. Let $0.5 < p < \Phi\left(\|\mu\|_{\Sigma^{-1}}\right)$. In particular, $\mu \neq 0$, because otherwise $\Phi\left(\|\mu\|_{\Sigma^{-1}}\right) = 0.5$. Also, by assumption, $\Phi^{-1}(p) < \|\mu\|_{\Sigma^{-1}}$. For each $t \geq 0$, put $y_t := q^{-1}(-t\Sigma^{-1}\mu)$. Then,

$$\Phi^{-1}(p)\,\|q(y_t)\|_\Sigma + \langle \mu, q(y_t)\rangle = \left(\Phi^{-1}(p) - \|\mu\|_{\Sigma^{-1}}\right) t\,\|\mu\|_{\Sigma^{-1}} \to_{t\to\infty} -\infty.$$

Hence, there is some t_0 such that, by (13), $y_t \in M_p^\alpha$ for all $t \geq t_0$. In other words,

$$q^{-1}\left(-[t_0,\infty)\cdot\Sigma^{-1}\mu\right) \subseteq M_p^\alpha$$

Since $\mu \neq 0$, one also has that $\Sigma^{-1}\mu \neq 0$. Therefore, $-[t_0,\infty)\cdot\Sigma^{-1}\mu$ is an unbounded set and $q^{-1}\left(-[t_0,\infty)\cdot\Sigma^{-1}\mu\right)$ is unbounded too because q is a homeomorphism. Consequently, M_p^α is an unbounded set. If $p \leq 0.5$ then M_p^α becomes even larger due to (5). This proves the first part of our proposition.

If $\alpha < 0$ and $p \geq \Phi\left(\|\mu\|_{\Sigma^{-1}}\right)$, then $M_p^\alpha = \emptyset$ by Proposition 2.3, so compactness follows trivially in this situation. Next, let $\alpha \geq 0$ and $p > \Phi\left(\|\mu\|_{\Sigma^{-1}}\right)$, whence $\Phi^{-1}(p) > \|\mu\|_{\Sigma^{-1}}$. The closed ball (w.r.t. the norm induced by Σ)

$$B := \{y \mid \|y\|_\Sigma \leq \left(\Phi^{-1}(p) - \|\mu\|_{\Sigma^{-1}}\right)^{-1}\alpha\}$$

is compact, hence $q^{-1}(B)$ is compact too. On the other hand, for $x \in M_p^\alpha$, one derives from (14) and (13) that

$$\|q(x)\|_\Sigma\left(\Phi^{-1}(p) - \|\mu\|_{\Sigma^{-1}}\right) \leq \Phi^{-1}(p)\,\|q(x)\|_\Sigma + \langle \mu, q(x)\rangle \leq \alpha,$$

whence $q(x) \in B$. In other words, $M_p^\alpha \subseteq q^{-1}(B)$. As a closed subset of a compact set, M_p^α has to be compact too (for closedness see continuity of the constraint function in (8)). Finally, let $\alpha \geq 0$ and $p = \Phi\left(\|\mu\|_{\Sigma^{-1}}\right)$. If $\mu = 0$, then $\Phi^{-1}(p) = 0$ and $M_p^\alpha = \mathbb{R}^n$ according to (13). In the case $\mu \neq 0$, one could reapeat the construction of y_t in the beginning of this proof in order to derive that

$$\Phi^{-1}(p) \, \|q(y_t)\|_\Sigma + \langle \mu, q(y_t) \rangle = 0 \leq \alpha \quad \forall t \geq 0.$$

Then, $q^{-1}\left(-[0,\infty) \cdot \Sigma^{-1}\mu\right) \subseteq M_p^\alpha$ and unboundedness of M_p^α would result in the same way as above.

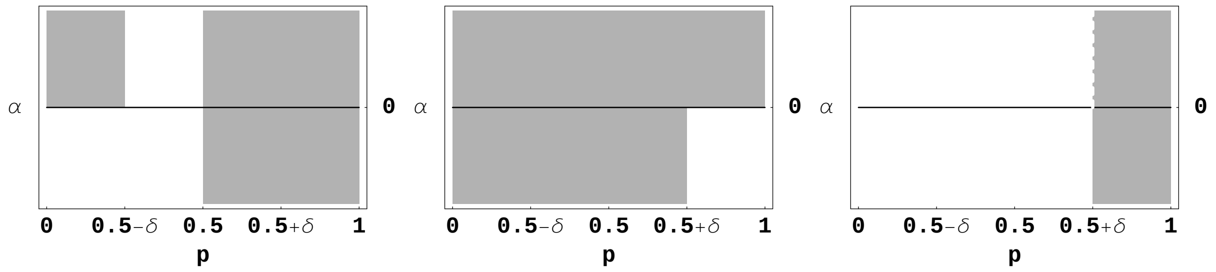
■

The following theorem provides a compilation of the results obtained so far. In order to collect a maximum of information, we restrict the functions q to the class of regular affine linear mappings, i.e., $q(x) = Ax + b$, with some regular matrix A. This class satisfies all assumptions made so far and covers in particular the case of linear chance constraints with stochastic coefficients and deterministic or stochastic right-hand side. The results on convexity, non-emptiness and compactness proven in the previous sections, are exhaustive in the sense that they completely determine, for which constellations of α and p the feasible sets M_p^α will be convex or nonconvex, empty or nonempty, compact or unbounded. In this sense, a full structural characterization is established. Let us define the following regions in the (p, α)- plane:

$$\mathcal{R}^{\text{conv(non}\emptyset\text{,comp)}} = \{(p,\alpha) | M_p^\alpha \text{ is convex (nonempty, compact)}\}.$$

For the purpose of abbreviation, denote $\delta := \Phi\left(\|\mu\|_{\Sigma^{-1}}\right) - 0.5$ and observe that $\delta \geq 0$ and that $\Phi\left(-\|\mu\|_{\Sigma^{-1}}\right) = 0.5 - \delta$.

Figure 1: Illustration of the regions of convexity (left), non-emptiness (middle) and compactness (right) in the (p, α)- plane.



Theorem 2.2 *In (4), let q be a regular affine linear mapping and let $\xi \sim \mathcal{N}(\mu, \Sigma)$ with positive definite Σ. Then,*

$$\begin{aligned}
\mathcal{R}^{conv} &= \{[0, 0.5-\delta] \times [0,\infty)\} \cup \{[0.5, 1] \times (-\infty, \infty)\} \\
\mathcal{R}^{non\emptyset} &= \{[0,1] \times [0,\infty)\} \cup \{[0, 0.5+\delta] \times (-\infty, 0)\} \\
\mathcal{R}^{comp} &= \{[0.5+\delta, 1] \times (-\infty, 0)\} \cup \{(0.5+\delta, 1] \times (0,\infty)\}.
\end{aligned}$$

Proof. Follows from Proposition 2.1 (first statement), Remark 2.1, Proposition 2.2, Proposition 2.3 and Proposition 2.4.

■

The regions $\mathcal{R}^{\text{conv}}$, $\mathcal{R}^{\text{non}\emptyset}$ and $\mathcal{R}^{\text{comp}}$ are illustrated in Figure 1.

Remark 2.2 *In the special case that $\mu = 0$, one derives that M_p^α is convex for all $\alpha \geq 0$ and* *__all__* *$p \in (0,1)$.*

2.4 Application to problems with joint probabilistic constraints

It is obvious to apply the previously obtained results for single probabilistic constraints like (4) to systems of individual probabilistic constraints like (3) because the feasible set of the latter system is just the intersection of the feasible sets induced by the single contsraints. Therefore, in this section, we shall address the more complicated case of joint probabilistic constraints as in (2). Consider the feasible set

$$M = \{x \in \mathbb{R}^n | P\left(\Xi q(x) \leq a\right) \geq p\} \quad (p \in (0,1)). \tag{15}$$

defined by a stochastic matrix Ξ of order (m,n) and a deterministic right-hand side $a \in \mathbb{R}^m$. Here, $q : \mathbb{R}^n \to \mathbb{R}^n$ refers to a (possibly nonlinear) mapping of the decision vector. By

$$M^i := \{x \in \mathbb{R}^n | P(\langle q(x), \xi_i\rangle \leq a_i) \geq p\} \quad (i = 1, \ldots, m),$$

we denote the feasible set induced by the i-th row of Ξ. Of course, M is not just the intersection of the M^i. However, for any i, one has the obvious inclusion $M \subseteq M^i$. This simple fact allows to derive the following useful compactness condition for joint probabilistic constraints:

Theorem 2.3 *In (15), assume that the rows ξ_i of Ξ are normally distributed according to $\xi_i \sim \mathcal{N}(\mu_i, \Sigma_i)$ with positive definite covariance matrices Σ_i for $i = 1, \ldots, m$. Moreover, let q be a homeomorphism (e.g., $q(x) = x$). Then, M is compact provided that*

$$p > \min_{i=1,\ldots,m} \Phi\left(\|\mu_i\|_{\Sigma_i^{-1}}\right).$$

Proof. According to the assumption, there exists some $i \in \{1, \ldots, m\}$ such that $p > \Phi\left(\|\mu_i\|_{\Sigma_i^{-1}}\right)$. Then, M^i is compact by Proposition 2.4. Consequently, M is bounded due to $M \subseteq M^k$. By Lemma 2.1, M is also closed. Summarizing, M is compact.

■

As an immediate corollary to Theorem 2.3, one derives the following existence result for the optimization problem

$$\min\{f(x) | x \in M\} \tag{16}$$

with joint probabilistic constraints:

Corollary 2.2 *In (16), let f be lower semicontinuous. Let M satisfy the hypotheses of Theorem 2.3 in the special case that $q(x) = x$. Moreover let $a \geq 0$ (componentwise). Then, there exists a solution to (16) provided that*

$$p > \min_{i=1,\ldots,m} \Phi\left(\|\mu_i\|_{\Sigma_i^{-1}}\right).$$

Proof. The assumptions $a \geq 0$ and $q(x) = x$ imply that $0 \in M$. Hence, M is nonempty. The result follows from Theorem 2.3 via the Weierstrass Theorem.

■

Theorem 2.3 and Corollary 2.2 hold true for large enough probability levels p which are typically encountered in applications of probabilitic constraints. Moreover, the required level is easily calculated just on the basis of the parameters μ_i and Σ_i. The additional condition of $a \geq 0$ in Corollary 2.2 is needed to ensure nonemptiness of the feasible set (which does not affect the compactness result of Theorem 2.3). From the reverse point of view, a general condition for emptiness can be derived as follows:

Theorem 2.4 *The feasible set M in (15) is empty if*

$$p \geq \min_{i \in I} \Phi\left(\|\mu_i\|_{\Sigma_i^{-1}}\right),$$

where $I := \{i \in \{1, \ldots, m\} | a_i < 0\}$.

Proof. With the same inclusion as used in the proof of Theorem 2.3, one may apply the first statement of Proposition 2.3.

■

We note that compactness and nonemptiness of feasible sets are crucial conditions not only for existence but also for stability of solutions and optimal values in problems like (16) when approximating the underlying, usually unknown probability distribution by another one which may be based on historical data (see [2]). Often, there is no chance directly to check the nonemptiness and compactness of a feasible set defined by a pure probabilistic constraint. Theorem , however, confirms that, for sufficiently high probability levels p, this assumption holds true in our case and, moreover, the notion 'sufficiently high' can be easily quantified excatly.

Acknowledgement
The author wishes to thank Prof. A. Seeger (University of Avignon) for proposin, and moreover, the term 'sufficiently high' can be easily quantified.g an improved version of Proposition 2.1

References

[1] Chmielewski, M.A., Elliptically Symmetric Distributions: A Review and Bibliography, Int. Stat. Rev. 49 (1981), 67-74.

[2] Henrion, R. and Römisch, W., Hölder and Lipschitz stability of solution sets in programs with probabilistic constraints, Math. Programming, **100** (2004), 589-611.

[3] Johnson, N.L. and Kotz, S., *Distributions in Statistics: Continuous Multivariate Distributions*, Wiley, New York, 1972.

[4] Kan, Y., Application of the Quantile Optimization, In: K. Marti (Ed.): Stochastic Optimization Techniques, *Lect. Notes. Econ. Math. Syst.*, **Vol. 513**, Springer, Berlin, 2002, pp. 285-308.

[5] Kataoka, S., A Stochastic Programming Model, *Econometrica* **31** (1963), 181-196.

[6] Lagoa, C.M., Li, X. and Sznaier, M., Probabilistically Constrained Linear Programs and Risk-Adjusted Controller Design, to appear in SIAM J. Control Optim.

[7] Lucchetti, R., Salinetti, G. and Wets, R.J.-B., Uniform Convergence of Probability Measures: Topological Criteria, *J. Multivariate Anal.* **51** (1994), 252-264.

[8] Panne, C. van de, Popp, W., Minimum-Cost Cattle Feed under Probabilistic Protein Constraints, *Manag. Sci.* **9** (1963itManag. Sci. **9** (1963), 405-430.

[9] Prékopa, A., Stochastic Programming, Kluwer, Dordrecht, 1995.

[10] Prékopa, A., Probabilistic Programming, Chapter 5 in [12].

[11] Römisch, W. and Schultz, R., Stability Analysis for Stochastic Programs, *Annals of Operations Research* **30** (1991), 241-266.

[12] Ruszczyński, A. and Shapiro, A. (Eds.), *Stochastic Programming*, Handbooks in OR&MS, **Vol. 10**, Elsevier, Amsterdam, 2003.