Convexity of chance constraints with independent random variables *

René Henrion
Weierstrass Institute Berlin
Germany
email: henrion@wias-berlin.de

Cyrille Strugarek
EDF R&D
and
École Nationale des Ponts et Chaussées
France
email: strugare@cermics.enpc.fr

May 2, 2007

Abstract

We investigate the convexity of chance constraints with independent random variables. It will be shown, how concavity properties of the mapping related to the decision vector have to be combined with a suitable property of decrease for the marginal densities in order to arrive at convexity of the feasible set for large enough probability levels. It turns out that the required decrease can be verified for most prominent density functions. The results are applied then, to derive convexity of linear chance constraints with normally distributed stochastic coefficients when assuming independence of the rows of the coefficient matrix.

Key words: chance constraints, probabilistic constraints, stochastic programming, convexity, random matrix

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

Mathematics Subject Classification (2000):
90C15

1 Introduction

Many optimization problems in engineering or finance contain so-called chance constraints or probabilistic constraints of the form

$$\mathbb{P}(h(x,\xi) \geq 0) \geq p, \tag{1}$$

where $x \in \mathbb{R}^n$ is a decision vector, $\xi : \Omega \to \mathbb{R}^m$ is an m-dimensional random vector defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^s$ is a vector-valued mapping and $p \in [0,1]$ is some probability level. A compilation of practical applications in which constraints of the type (1) play a crucial role, may be found in the standard references [11], [12]. Not surprisingly, one of the most important theoretical questions related to such constraints is that of convexity of the set of decisions x satisfying (1). It is well-known ([11], Th. 10.2.1) that this set is convex provided that the law $\mathbb{P} \circ \xi^{-1}$ of ξ is a log-concave probability measure on $\mathbb{R}^m$ and that the components h_i of h are quasi-concave. The power of this result becomes evident in combination with a celebrated theorem by Prékopa stating that the law of ξ is log-concave whenever ξ has a log-concave density. As this is easily verified to hold true for many prominent multivariate distributions, this classical result guarantees convexity of the set of feasible decisions for a broad class of applications. The required quasi-concavity of the h_i is satisfied, for instance in the linear model $h(x,\xi) = Ax - B\xi$, where actually concavity of the h_i holds true.

In this paper, we shall be interested in chance constraints where random vectors appear separated from decision vectors, and which come as a special case of (1) by putting $h(x,\xi) = g(x) - \xi$. More precisely, we want to study convexity of a set of feasible decisions defined by

$$M(p) = \{x \in \mathbb{R}^n | \mathbb{P}(\xi \leq g(x)) \geq p\}, \tag{2}$$

where $g : \mathbb{R}^n \to \mathbb{R}^m$ is some vector-valued mapping. With $F : \mathbb{R}^m \to \mathbb{R}$ denoting the

distribution function of ξ, the same set can be rewritten as

$$M(p) = \{x \in \mathbb{R}^n | F(g(x)) \geq p\}. \tag{3}$$

We are interested in conditions on F and g such that $M(p)$ becomes a convex set for all $p \geq p^*$, where $p^* < 1$. Note that convexity for large enough p is a relevant feature because p is typically chosen to be close to one.

When trying to link the previously mentioned classical result to the special case of (2), in addition to the log-concavity of the law of ξ, we would have to impose quasi-concavity of the functions $g_i(x) - \xi_i$. Unfortunately, unlike concavity, quasi-concavity is not preserved under addition, so quasi-concavity of the components g_i is not sufficient here to ensure convexity of $M(p)$. To illustrate this fact consider the following example:

Example 1.1 *In (2), let ξ have a bivariate standard normal distribution with independent components, and let $g(x, y) := (e^x, e^y)$. Then, the components g_i are quasi-concave (as functions of x and y simultaneously). However, the set $M(0.5)$ fails to be convex (e.g., for $u := (1, -3)$ and $v = (-3, 1)$ one has that $u, v \in M(0.5)$ but $(u + v)/2 \notin M(0.5)$).*

On the other hand, concavity of the components g_i would do because then, $g_i(x) - \xi_i$ is a concave, hence quasi-concave function of the two variables x and ξ, simultaneously. In particular, convexity of $M(p)$ would hold true for all $p \in [0, 1]$ in Example 1.1 upon passing from g to $-g$. Therefore, the question arises, whether one can still derive convexity results for $M(p)$ in (2) when relaxing the strong requirement of concave components g_i. It turns out that this will be possible under the additional assumption of ξ having independent components. Then, roughly speaking, convexity can be derived for so-called r-concave g_i, a concept providing a parametrization of concavity properties between true concavity and quasi-concavity (see Section 2). As an application, we show that joint chance constraints defined by a normally distributed random matrix yield a convex set of feasible decisions provided the probability level is large enough and the rows of the random matrix are independently distributed. To the best of our knowledge, this result is new and may have an impact on solution procedures for problems of such kind by making available tools from convex optimization. We emphasize that the independence assumption is essential

for our approach. For other work on convexity properties of chance constraints where independence has been successfully exploited, we refer to [1], [4] and [7]. A Theorem by Bawa [1], for instance, provides a condition to ensure concavity of the product function

$$H(t) = F(t_1) \cdots F(t_m),$$

where F is a one-dimensional distribution function. This would be of interest in the context of (3) if all components ξ_i of the random vector had identical independent distributions. However, the interplay with relaxations of concavity of the g_i in (3) is not clear. The conditions we are going to impose on the distribution function F (or better: on the marginal distribution functions F_i) are related to the degree at which the corresponding densities f_i decrease asymptotically. This will ensure that the mappings $t \mapsto F_i(1/t^\alpha)$ become concave for an appropriate $\alpha > 0$.

2 Notation

We recall the definition of an r-concave function:

Definition 2.1 *A function $f : \mathbb{R}^s \to (0, \infty)$ is called r-concave for some $r \in [-\infty, \infty]$, if*

$$f(\lambda x + (1-\lambda)y) \geq [\lambda f^r(x) + (1-\lambda) f^r(y)]^{1/r} \quad \forall x, y \in \mathbb{R}^s, \forall \lambda \in [0,1]. \tag{4}$$

In this definition, the cases $r \in \{-\infty, 0, \infty\}$ are to be interpreted by continuity. In particular, 1-concavity amounts to classical concavity, 0-concavity equals log-concavity (i.e., concavity of $\log f$), and $-\infty$-concavity identifies quasi-concavity (this means that the right-hand side of the inequaltiy in the definition becomes $\min\{f(x), f(y)\}$). We recall, that an equivalent way to express log-concavity is the inequality

$$f(\lambda x + (1-\lambda)y) \geq f^\lambda(x) f^{1-\lambda}(y) \quad \forall x, y \in \mathbb{R}^s, \forall \lambda \in [0,1]. \tag{5}$$

For $r < 0$, one may raise (4) to the negative power r and recognize, upon reversing the inequality sign, that this reduces to convexity of f^r. If f is r^*-concave, then f is r-concave for all $r \leq r^*$. We shall be mainly interested in the case $r \leq 1$.

The following property is crucial in the context of this paper:

Definition 2.2 *We call a function $f : \mathbb{R} \to \mathbb{R}$* ***r-decreasing*** *for some $r \in \mathbb{R}$, if it is continuous on $(0, \infty)$ and if there exists some $t^* > 0$ such that the function $t^r f(t)$ is strictly decreasing for all $t > t^*$.*

Evidently, 0-decreasing means strictly decreasing in the classical sense. If f is a non-negative function like the density of some random variable, then r-decreasing implies r'-decreasing whenever $r' \leq r$. Therefore, one gets narrower families of r-decreasing density functions with $r \to \infty$. If f is not just continuous on $(0, \infty)$ but happens even to be differentiable there, then the property of being r-decreasing amounts to the condition

$$tf'(t) + rf(t) < 0 \quad \text{for all } t > t^*. \tag{6}$$

3 A Convexity Result

Lemma 3.1 *Let $F : \mathbb{R} \to [0, 1]$ be a distribution function with $(r + 1)$-decreasing density f for some $r > 0$. Then, the function $z \mapsto F(z^{-1/r})$ is concave on $(0, (t^*)^{-r})$, where t^* refers to Definition 2.2. Moreover, $F(t) < 1$ for all $t \in \mathbb{R}$.*

Proof. Let $h : \mathbb{R} \to \mathbb{R}$ be defined by $h(z) = F(z^{-1/r})$, for all $z > 0$. By definition, it holds that

$$h(z) = F(0) + \int_0^{z^{-1/r}} f(t)dt \quad \forall z > 0.$$

With the change of variables $t = u^{-1/r}$, the last equation rereads

$$h(z) = F(0) + r^{-1} \int_z^{+\infty} u^{-(1+1/r)} f(u^{-1/r})du.$$

Since f is continuous on $(0, \infty)$ by the very definition of r-decreasing functions, F and h are differentiable on the same interval. Consequently,

$$h'(z) = -r^{-1} z^{-(1+1/r)} f(z^{-1/r}).$$

Since, by assumption, $t \mapsto t^{r+1} f(t)$ is strictly decreasing on $(t^*, +\infty)$, one gets that $z \mapsto z^{-(1+1/r)} f(z^{-1/r})$ is strictly increasing on $(0, (t^*)^{-r})$. Summarizing, h' is strictly decreasing on $(0, (t^*)^{-r})$, whence h is concave on this interval.

Concerning the second statement, assume that $F(t) = 1$ for all $t \geq \tau$. Therefore, with F being a distribution function, it follows the contradiction $F'(t) = f(t) = 0$ for all $t > \tau$ to f being $(r+1)$-decreasing. ∎

Theorem 3.1 *For (2), we make the following assumptions for $i = 1, \ldots, m$:*

1. *There exist $r_i > 0$ such that the components g_i are $(-r_i)$-concave.*
2. *The components ξ_i of ξ are independently distributed with (r_i+1)-decreasing densities f_i.*

Then, $M(p)$ is convex for all $p > p^* := \max\{F_i(t_i^*) | 1 \leq i \leq m\}$, where F_i denotes the distribution function of ξ_i and the t_i^* refer to Definition 2.2 in the context of f_i being (r_i+1)-decreasing.

Proof. Let $p > p^*$, $\lambda \in [0,1]$ and $x, y \in M(p)$ be arbitrary. We have to show that $\lambda x + (1-\lambda)y \in M(p)$. Referring to the distribution functions F_i of ξ_i, we put

$$q_i^x := F_i(g_i(x)) < 1, \quad q_i^y := F_i(g_i(y)) < 1 \quad (i = 1, \ldots, m)\,, \tag{7}$$

where the strict inequalities rely on the second statement of Lemma 3.1. By assumption 2., the components of ξ are independent, hence the feasible set in (2) or (3), respectively, may be rewritten as

$$M(p) = \left\{ w \in \mathbb{R}^n \,\middle|\, \prod_{i=1}^m F_i(g_i(w)) \geq p \right\}. \tag{8}$$

In particular, by (7), the inclusions $x, y \in M(p)$ mean that

$$\prod_{i=1}^m q_i^x \geq p, \quad \prod_{i=1}^m q_i^y \geq p. \tag{9}$$

Now, (7), (9) and the definition of p^* entail that

$$1 > q_i^x \geq p > F_i(t_i^*) \geq 0, \quad 1 > q_i^y \geq p > F_i(t_i^*) \geq 0 \quad (i = 1, \ldots, m)\,. \tag{10}$$

For $\tau \in [0,1]$, we denote the τ-quantile of F_i by

$$\tilde{F}_i(\tau) := \inf\{z \in \mathbb{R} | F_i(z) \geq \tau\}.$$

Note that, for $\tau \in (0,1)$, $\tilde{F}_i(\tau)$ is a real number. Having a density, by assumption 2., the F_i are continuous distribution functions. As a consequence, the quantile functions $\tilde{F}_i(\tau)$ satisfy the implication

$$q > F_i(z) \Longrightarrow \tilde{F}_i(q) > z \quad \forall q \in (0,1)\, \forall z \in \mathbb{R}.$$

Now, (7) and (10) provide the relations

$$g_i(x) \geq \tilde{F}_i(q_i^x) > t_i^* > 0, \quad g_i(y) \geq \tilde{F}_i(q_i^y) > t_i^* > 0 \quad (i = 1, \ldots, m)\,. \tag{11}$$

In particular, for all $i = 1, \ldots, m$, it holds that

$$\left[\min\{\tilde{F}_i^{-r_i}(q_i^x), \tilde{F}_i^{-r_i}(q_i^y)\}, \max\{\tilde{F}_i^{-r_i}(q_i^x), \tilde{F}_i^{-r_i}(q_i^y)\}\right] \subseteq \left(0, (t_i^*)^{-r_i}\right). \tag{12}$$

Along with assumption 1., (11) yields for $i = 1, \ldots, m$:

$$\begin{aligned} g_i\left(\lambda x + (1-\lambda)y\right) &\geq \left(\lambda g_i^{-r_i}(x) + (1-\lambda) g_i^{-r_i}(y)\right)^{-1/r_i} \\ &\geq \left(\lambda \tilde{F}_i^{-r_i}(q_i^x) + (1-\lambda)\tilde{F}_i^{-r_i}(q_i^y)\right)^{-1/r_i}. \end{aligned} \tag{13}$$

The monotonicity of distribution functions allows to continue by

$$F_i\left(g_i\left(\lambda x + (1-\lambda)y\right)\right) \geq F_i\left(\left(\lambda \tilde{F}_i^{-r_i}(q_i^x) + (1-\lambda)\tilde{F}_i^{-r_i}(q_i^y)\right)^{-1/r_i}\right) \quad (i = 1, \ldots, m)\,. \tag{14}$$

Owing to assumption 2., Lemma 3.1 guarantees that the functions $z \mapsto F_i(z^{-1/r_i})$ are concave on $(0, (t_i^*)^{-r_i})$. In particular, these functions are log-concave on the indicated interval, as this is a weaker property than concavity (see Section 2). By virtue of (12) and (5), this allows to continue (14) as

$$F_i\left(g_i\left(\lambda x + (1-\lambda)y\right)\right) \geq \left[F_i\left(\tilde{F}_i(q_i^x)\right)\right]^{\lambda} \left[F_i\left(\tilde{F}_i(q_i^y)\right)\right]^{1-\lambda} \quad (i = 1, \ldots, m)\,.$$

Exploiting the fact that the F_i as continuous distribution functions satisfy the relation $F_i(\tilde{F}_i(q)) = q$ for all $q \in (0,1)$, and recalling that $q_i^x, q_i^y \in (0,1)$ by (10), we may deduce that

$$F_i\left(g_i\left(\lambda x + (1-\lambda)y\right)\right) \geq [q_i^x]^{\lambda}\, [q_i^y]^{1-\lambda} \quad (i = 1, \ldots, m)\,.$$

Passing to the product, it follows together with (9) that

$$\prod_{i=1}^m F_i\left(g_i\left(\lambda x + (1-\lambda)y\right)\right) \geq \prod_{i=1}^m [q_i^x]^{\lambda}\, [q_i^y]^{1-\lambda} = \left[\prod_{i=1}^m q_i^x\right]^{\lambda} \left[\prod_{i=1}^m q_i^y\right]^{1-\lambda} \geq p^{\lambda} p^{1-\lambda} = p.$$

Referring to (8), this shows that $\lambda x + (1-\lambda)y \in M(p)$. ∎

Remark 3.1 *The critical probability level p^* beyond which convexity can be guaranteed in Theorem 3.1, is completely independent of the mapping g, it just depends on the distribution functions F_i. In other words, for given distribution functions F_i, the convexity of $M(p)$ in (2) for $p > p^*$ can be guaranteed for a whole class of mappings g satisfying the first assumption of Theorem 3.1. Therefore, it should come at no surprise that, for specific mappings g even smaller critical values p^* may apply (see Example 4.1 below).*

In the following proposition, we establish the relation between log-concave distributions and distributions having an r-decreasing density. We recall that the class of log-concave distributions having a density coincides with the class of distributions having a log-concave density ([2], Th. 3.1).We also mention that most of the prominent distributions fall into this class.

Proposition 3.1 *Let $f : \mathbb{R} \to [0,1]$ be a log-concave and continuous density having an unbounded support in positive direction. Then, f is r-decreasing for all $r > 0$.*

Proof. By assumption, $\phi := \log f$ is a concave, possibly extended-valued function. As a consequence of concavity, there exists some $\tau > 0$ such that either $\phi(t) = -\infty$ for all $t > \tau$ or $\phi(t) > -\infty$ for all $t > \tau$. The first case amounts to $f(t) = 0$ for all $t > \tau$, which is a contradiction with our assumption of f having an unbounded support in positive direction. Consequently, ϕ is concave and real-valued on $[\tau, \infty)$. Moreover, as a continuous and log-concave density function, f must tend to zero at infinity, hence $\lim_{t\to\infty} \phi(t) = -\infty$. Along with the concavity of ϕ, this implies the existence of $\alpha < 0$ and $\beta \in \mathbb{R}$ such that

$$\phi(t) \leq \alpha t + \beta \quad \forall t \geq \tau. \tag{15}$$

Now, let $r > 0$ be arbitrary and put $h(t) := t^r f(t)$ for $t > 0$. Then, $\log h = r \log(\cdot) + \phi$ is also concave and real-valued on $[\tau, \infty)$. Assume there exists some $\tau^* > \tau$ such that $\log h(\tau^*) < \log h(\tau)$. By concavity of $\log h$, this function and, thus, h itself must then be strictly decreasing on $[\tau^*, \infty)$. In other words, f is r-decreasing as was to be

shown. Therefore, we are done if we can lead to a contradiction the opposite case, namely $\log h(t) \geq \log h(\tau)$ for all $t \geq \tau$. This is equivalent to

$$\phi(t) \geq \log h(\tau) - r \log t \quad \forall t \geq \tau. \tag{16}$$

We apply the general relation

$$-r \log t \geq -r \log s - rt/s - r \quad \forall t \geq s > 0$$

to $s := -2r/\alpha > 0$, where α refers to (15):

$$-r \log t \geq -r \log(-2r/\alpha) + \alpha t/2 - r \quad \forall t \geq s.$$

Combining this with (15) and (16), we arrive at the contradiction

$$K := \log h(\tau) - r \log(-2r/\alpha) - r - \beta \leq \alpha t/2 \quad \forall t \geq \max\{\tau, s\}$$

to the fact that K is a constant and $\alpha/2 < 0$. ■

Recalling that normal densities are log-concave, continuous and have unbounded support, we may combine Theorem 3.1 and Proposition 3.1, in order to obtain a useful characterization of convexity under normally distributed data:

Corollary 3.2 *In (2), let ξ have a regular multivariate normal distribution with independent components. Moreover, let each component g_i of g be $(-r_i)$-concave for some $r_i > 0$. Then, there exists some $p^* < 1$ such that $M(p)$ is convex for all $p > p^*$.*

4 Examples

The Cauchy distribution has a density

$$f(t) = \frac{a}{\pi(a^2 + t^2)} \quad (a > 0)$$

which is r-decreasing for any $r < 2$ but fails to be so for any $r \geq 2$. Most of the prominent one-dimensional distributions, however, have a density which is r-decreasing for any $r > 0$. Next, we want to calculate for some well-known one-dimensional distributions the t^*- and

$F(t^*)$- values needed in Theorem 3.1 for the computation of the critical probability level p^*. We start with the corresponding derivation of the normal distribution and collect the others in Table 1. To emphasize the dependence on the order r, we shall write t_r^* rather than just t^*.

Proposition 4.1 *Let ξ have a normal distribution with scalar parameters μ and $\sigma > 0$. Moreover, let $r > 0$ be arbitrarily given. Then, the corresponding density is r-decreasing with*

$$t_r^* = \frac{\sqrt{\mu^2 + 4r\sigma^2} + \mu}{2} \quad \textit{and} \quad F(t_r^*) = \Phi\left(\sqrt{r + \frac{1}{4}\left(\frac{\mu}{\sigma}\right)^2} - \frac{1}{2}\frac{\mu}{\sigma}\right),$$

where Φ denotes the distribution function of the standard normal distribution.

Proof. The calculation of the (optimal) t_r^*- value is straightforward from the representation of the normal density and (6). By definition,

$$F(t_r^*) = \mathbb{P}\left(\xi \leq t_r^*\right) = \mathbb{P}\left(\frac{\xi - \mu}{\sigma} \leq \frac{t_r^* - \mu}{\sigma}\right).$$

Since $\sigma^{-1}\left(\xi - \mu\right)$ has a standard normal distribution, one may continue as

$$F(t_r^*) = \Phi\left(\frac{t_r^* - \mu}{\sigma}\right) = \Phi\left(\sqrt{r + \frac{1}{4}\left(\frac{\mu}{\sigma}\right)^2} - \frac{1}{2}\frac{\mu}{\sigma}\right).$$

■

For the special case of a standard normal distribution ($\mu = 0, \sigma = 1$), one gets $t_r^* = \sqrt{r}$ and $F(t_r^*) = \Phi(\sqrt{r})$. As an illustration, we consider the following example:

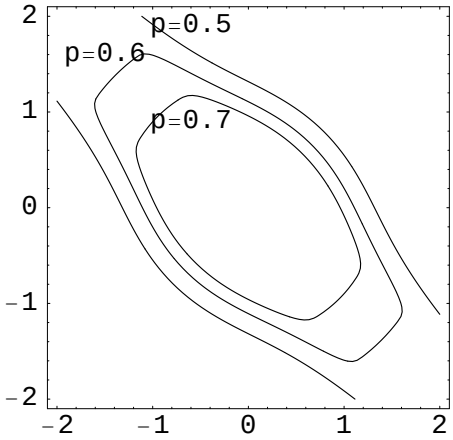
Example 4.1 *In (2) let ξ have a bivariate standard normal distribution: $\xi \sim \mathcal{N}(0, I_2)$. Moreover, put*

$$g_1(x,y) = \frac{1}{x^2 + y^2 + 0.1}, \quad g_2(x,y) = \frac{1}{(x+y)^2 + 0.1}.$$

Then, clearly, the components g_i are (-1)-concave (i.e. $1/g_i$ is convex). By assumption, the components of ξ have a one-dimensional standard normal distribution which, by Proposition 4.1, has a 2-decreasing density with $t^* = \sqrt{2}$. Now, Theorem 3.1 may be applied and we may derive convexity of the feasible set $M(p)$ in (2) beyond a critical

probability level $p^* = \Phi\left(\sqrt{2}\right) \approx 0.921$. *According to Remark 3.1, possibly some much smaller level could do with respect to convexity. This is confirmed for the example by Figure 1: obviously, the feasible set is convex for probabilities higher than* 0.7 *and nonconvex for probabilities lower than* 0.6*, so the true critical level in this example is somewhere in between* 0.6 *and* 0.7*. Note that the classical convexity theory could not be applied to this example because the components* g_i *are not concave (see Introduction). This is also supported by the observation that convexity fails for small probabilities.*

Figure 1: Illustration of the feasible set $M(p)$ for different levels p in an example.



In the example, convexity of the feasible set $M(p)$ could be guaranteed for all probability levels larger than 0.921. This may sound a strong requirement, but note that, in chance constraint programming, these levels are typically high, say 0.95 or 0.99. Moreover, the result of Proposition 4.1 strongly depends on the parameters μ and σ, and, more precisely, on their ratio. If this ratio becomes large, then $F(t_r^*)$ converges towards $\Phi(0)$. Hence, for the case of normal distributions with small relative standard deviations, the critical level p^* tends to 0.5.

Table 1 shows, how the t_r^*- value depends on r and on the parameters of the different distributions. For two distributions, a closed formula is available for the corresponding value $F(t_r^*)$ of the distribution function: First, for the exponential distribution, one gets $F(t_r^*) = 1 - e^{-r}$. Hence, reconsidering Example 4.1 with independent exponential rather than normal distributions, one could derive convexity of the set $M(p)$ for probabilities

Table 1: t_r^*- values in the definition of r-decreasing densities for a set of common distributions.

Law	Density	t_r^*
normal	$\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$	$\frac{\mu+\sqrt{\mu^2+4r\sigma^2}}{2}$
exponential	$\lambda\exp(-\lambda t) \quad (t>0)$	$\frac{r}{\lambda}$
Weibull	$abt^{b-1}\exp\left(-at^b\right) \quad (t>0)$	$\left(\frac{b+r-1}{ab}\right)^{1/b}$
Gamma	$\frac{b^a}{\Gamma(a)}\exp(-bt)\,t^{a-1} \quad (t>0)$	$\frac{a+r-1}{b}$
χ	$\frac{1}{2^{n/2-1}\Gamma(n/2)}t^{n-1}\exp\left(-\frac{t^2}{2}\right) \quad (t>0)$	$\sqrt{n+r-1}$
χ^2	$\frac{1}{2^{n/2}\Gamma(n/2)}t^{n/2-1}\exp\left(-\frac{t}{2}\right) \quad (t>0)$	$n+2r-2$
log-normal	$\frac{1}{\sqrt{2\pi}\sigma t}\exp\left(-\frac{(\log t-\mu)^2}{2\sigma^2}\right) \quad (t>0)$	$e^{\mu+(r-1)\sigma^2}$
Maxwell	$\frac{2t^2}{\sqrt{2\pi}\sigma^3}\exp\left(-\frac{t^2}{2\sigma^2}\right) \quad (t>0)$	$\sigma\sqrt{r+2}$
Rayleigh	$\frac{2t}{\lambda}\exp\left(-\frac{t^2}{\lambda}\right) \quad (t>0)$	$\sqrt{\frac{r+1}{2}\lambda}$

larger than $1-e^{-2}\approx 0.864$ which is a slightly better value than in the normal case. It is interesting to observe that the critical probability level for the exponential distribution does not depend on the parameter of this distribution. The second case with a closed formula is the Weibull distribution, where one calculates $F(t^*) = 1-e^{-(b+r-1)/b}$ (see Table 1 for the meaning of parameters). In general, no closed formula is available, but in concrete applications, the critical probability levels are easily read off from usual data tables or numerical routines.

5 Chance constraints with normally distributed stochastic matrices

In this section, we want to apply Theorem 3.1 in order to derive a convexity result for a more complicated chance constraint than (2). More precisely, consider the feasible set

$$M(p) = \{x \in \mathbb{R}^n | \mathbb{P}(\Xi x \leq a) \geq p\}, \tag{17}$$

where the rows ξ_i of the stochastic matrix Ξ have multivariate normal distributions according to $\xi_i \sim \mathcal{N}(\mu_i, \Sigma_i)$. Linear chance constraints of this type, having random coefficients, are of importance in many engineering applications (e.g., mixture problems). Note that, in contrast to (2), the random parameter and the decision vector are no longer separated but coupled in a multiplicative way. This makes the convexity analysis more involved. A classical result due to Kataoka [6] and Van de Panne and Popp [8] states that $M(p)$ is convex for $p \geq 0.5$ in the simple case where Ξ reduces to single row ($m = 1$). A much more precise characterization not only of convexity but also of compactness and nontriviality of $M(p)$ in this elementary situation was provided in [5]. Moreover, compactness of $M(p)$ could even be characterized there in the general case (m arbitrary). However, convexity in the general case remains an open question. Below, we shall provide a positive result under the assumption of Ξ having independent rows. This yields a complementary characterization to results by Prékopa and Burkauskas, who derived convexity under the assumption that all covariance and cross-covariance matrices of the columns or rows of Ξ, respectively, are proportional to each other (see [10] and [3]).

A direct application of Theorem 3.1 to (17) is not possible, since this type of chance constraint is different from (2). However, there exists a useful transformation of the one into the other. First, we need an auxiliary result:

Lemma 5.1 *For $\mu \in \mathbb{R}^n$ and positive definite matrix Σ of order (n, n), we put*

$$f(x) := \frac{\langle x, \Sigma x \rangle}{\left(a - \langle \mu, x \rangle\right)^2} \text{ defined on the domain } \Omega_1 := \{x | a - \langle \mu, x \rangle > 0\}.$$

Then, f is convex on the following open subset of Ω_1:

$$\Omega_2 := \left\{ x \,\middle|\, a - \langle \mu, x \rangle > 4\lambda_{\max}\lambda_{\min}^{-3/2} \|\mu\| \sqrt{\langle x, \Sigma x \rangle} \right\}.$$

Here, $\lambda_{\max}$ and $\lambda_{\min}$ denote the largest and smallest eigenvalues of Σ.

Proof. On Ω_1, the Hessian of f calculates as

$$D^2 f(x) = 2\left(a - \langle \mu, x \rangle\right)^{-4} \left[\left(a - \langle \mu, x \rangle\right)^2 \Sigma + 4\left(a - \langle \mu, x \rangle\right) \Sigma x \mu^T + 3\langle x, \Sigma x \rangle \, \mu\mu^T\right].$$

In order to verify the positive definiteness of $D^2 f$ on Ω_2, it is evidently sufficient to show this property for the matrix

$$\left(a - \langle \mu, x \rangle\right) \Sigma + 4\Sigma x \mu^T.$$

If $z \neq 0$ and $x \in \Omega_2$ are arbitrarily given, then, by definition of Ω_2,

$$\begin{aligned}
\left\langle z, \left[(a - \langle \mu, x \rangle) \Sigma + 4 \Sigma x \mu^T \right] z \right\rangle &= (a - \langle \mu, x \rangle) \langle z, \Sigma z \rangle + 4 \langle z, \Sigma x \rangle \langle \mu, z \rangle \\
&\geq \lambda_{\min} \|z\|^2 (a - \langle \mu, x \rangle) - 4 \|\Sigma x\| \|\mu\| \|z\|^2 \\
&> 4 \|z\|^2 \|\mu\| \left(\lambda_{\max} \lambda_{\min}^{-1/2} \sqrt{\langle x, \Sigma x \rangle} - \|\Sigma x\| \right) \\
&\geq 0.
\end{aligned}$$

Here, we exploited the relations

$$\left\langle x, \Sigma^2 x \right\rangle \leq \lambda_{\max}^2 \|x\|^2, \quad \lambda_{\min} \|x\|^2 \leq \langle x, \Sigma x \rangle .$$

■

The next simple proposition will be needed later on but is of independent interest as well because it makes no restrictions on the probability level p:

Proposition 5.1 *If $a \geq 0$ (componentwise) in (17), then $M(p)$ is starshaped with respect to the origin. In particular, $M(p)$ is a connected set.*

Proof. Since $a \geq 0$ by assumption, one immediately derives that $0 \in M(p)$. We have to show that, for arbitrary $x \in M(p)$ and arbitrary $\lambda \in [0, 1]$, it follows that $\lambda x \in M(p)$. This is evident for $\lambda = 0$. If $\lambda \in (0, 1]$, then

$$\mathbb{P}(\Xi(\lambda x) \leq a) = \mathbb{P}(\Xi x \leq \lambda^{-1} a) \geq \mathbb{P}(\Xi x \leq a) \geq p.$$

Here we used that $\lambda^{-1} a \geq a$ (componentwise) due to $a \geq 0$ and $\lambda \leq 1$. In other words, $\lambda x \in M(p)$. ■

Theorem 5.1 *In (17) we assume that the rows ξ_i of Ξ are pairwise independently distributed. Then, $M(p)$ is convex for*

$$p > \Phi \left(\max \left\{ \sqrt{3}, u^* \right\} \right), \tag{18}$$

where Φ is the one-dimensional standard normal distribution function,

$$u^* = \max_{i=1,\ldots,m} 4 \lambda_{\max}^{(i)} \left[\lambda_{\min}^{(i)} \right]^{-3/2} \|\mu_i\| .$$

and $\lambda_{\max}^{(i)}$ and $\lambda_{\min}^{(i)}$ refer to the largest and smallest eigenvalue of Σ_i.

Proof. The assumption of independent rows allows to rewrite the feasible set as

$$M(p) = \left\{ x \in \mathbb{R}^n \,\middle|\, \prod_{i=1}^m \mathbb{P}(\langle \xi_i, x\rangle \leq a_i) \geq p \right\}.$$

For $x \neq 0$ and $i = 1, \ldots, m$, we put

$$\eta_i(x) := \frac{\langle \xi_i - \mu_i, x\rangle}{\sqrt{\langle x, \Sigma_i x\rangle}} \sim \mathcal{N}(0,1); \quad g_i(x) := \frac{a_i - \langle \mu_i, x\rangle}{\sqrt{\langle x, \Sigma_i x\rangle}}.$$

Evidently, for $x \neq 0$, one has that $\langle \xi_i, x\rangle \leq a_i$ holds true if and only if $\eta_i(x) \leq g_i(x)$. Since the $\eta_i(x)$ have a standard normal distribution, one obtains

$$\mathbb{P}(\langle \xi_i, x\rangle \leq a_i) = \Phi(g_i(x)) \quad (\text{for } x \neq 0 \text{ and } i = 1, \ldots, m). \tag{19}$$

We introduce the following sets for $i = 1, \ldots, m$:

$$\begin{aligned} \Omega_1^{(i)} &: = \{x \in \mathbb{R}^n | a_i - \langle \mu_i, x\rangle > 0\} \\ \Omega_2^{(i)} &: = \left\{ x \in \mathbb{R}^n | a_i - \langle \mu_i, x\rangle > 4\lambda_{\max}^{(i)} \left[\lambda_{\min}^{(i)}\right]^{-3/2} \|\mu_i\| \sqrt{\langle x, \Sigma_i x\rangle} \right\}. \end{aligned}$$

The following inclusions hold true whenever p satisfies (18):

$$M(p)\backslash\{0\} \subseteq \Omega_2^{(i)} \subseteq \Omega_1^{(i)} \quad (i = 1, \ldots, m).$$

The second inclusion is trivial. To verify the first one, let $x \in M(p)\backslash\{0\}$ be arbitrary. Since $\Phi \leq 1$, one derives from (19) that

$$\Phi\left(g_i(x)\right) \geq \prod_{j=1}^m \Phi\left(g_j(x)\right) = \prod_{j=1}^m \mathbb{P}(\langle \xi_j, x\rangle \leq a_j) \geq p > \Phi\left(u^*\right) \quad (i = 1, \ldots, m).$$

With Φ being strictly increasing, this amounts to $g_i(x) > u^*$ and thus $x \in \Omega_2^{(i)}$ for $i = 1, \ldots, m$ by definition of u^*.

Next, on $\Omega_1^{(i)}$ define

$$f_i(w) := \frac{\langle w, \Sigma_i w\rangle}{\left(a_i - \langle \mu_i, w\rangle\right)^2} \quad (i = 1, \ldots, m).$$

Note that the f_i are finite-valued on $\Omega_1^{(i)}$. By Lemma 5.1, the f_i are convex on $\Omega_2^{(i)}$. On the other hand, the g_i are finite-valued and positive on $\Omega_1^{(i)}\backslash\{0\}$ and so in particular on $\Omega_2^{(i)}\backslash\{0\}$. From the respective definitions, it follows then that $f_i = g_i^{-2}$ on $\Omega_2^{(i)}\backslash\{0\}$.

Recalling that $p > 0$, by assumption, one gets that $0 \in M(p)$ if and only if $a_i \geq 0$ for all $i = 1, \ldots, m$. We proceed by case distinction:

First case: $\min_{i=1,\ldots,m} a_i < 0$

Then, $0 \notin M(p)$ and, by (19), $M(p) = \{x \in \mathbb{R}^n \,|\prod_{i=1}^m \Phi\left(g_i(x)\right) \geq p\}$. Hence, we are in the setting of (8) in Theorem 3.1 with $F_i := \Phi$ for $i = 1, \ldots, m$. From the remark below Proposition 4.1, we know that Φ has a 3 -decreasing density with critical value $t^* = \sqrt{3}$. Therefore, condition 2. of Theorem 3.1 is satisfied with $r_i := 2$ for $i = 1, \ldots, m$, and the statement of the Theorem will allow to derive convexity of $M(p)$ for all $p > \Phi(\sqrt{3})$ under the condition that the first assumption of Theorem 3.1 be fulfilled, i.e., the g_i are (-2)-concave. This point, however, deserves some attention because in contrast to the setting required in Theorem 3.1 and in Definition 2.1, our g_i are not defined on the whole space and may be not (-2)-concave on all of their domain. We shall proceed as follows: as in Theorem 3.1 we consider arbitrary $x, y \in M(p)$ and $\lambda \in [0, 1]$, and we show that

$$x_\lambda := \lambda x + (1 - \lambda)y \in M(p).$$

We have two options to do so. The first one is to check the relation of (-2)-concavity of the g_i for the concrete triple (x, y, x_λ):

$$g_i(x_\lambda) \geq \left(\lambda g_i^{-2}(x) + (1 - \lambda) g_i^{-2}(y)\right)^{-1/2}. \tag{20}$$

Indeed, this last relation corresponds to the first inequality in (13). A brief reinspection of the proof of Theorem 3.1 shows that, given all the necessary assumptions on the distribution functions, this inequality is all what is needed to derive that $x_\lambda \in M(p)$. However, it may happen, that (20) cannot be verified, for instance due to $x_\lambda = 0$, so that x_λ does not belong to the domain of the g_i. Then, we might be able to show $x_\lambda \in M(p)$ by a direct argument.

In a first step, we show that $x_\lambda \neq 0$. Assuming to the contrary, that $x_\lambda = 0$ and recalling that $0 \notin M(p)$ (so $x, y \neq 0$), it follows the existence of some $\alpha < 0$ such that $x = \alpha y$. Since, $x, y \in M(p) = M(p)\backslash\{0\} \subseteq \Omega_1^{(i)}$ for $i = 1, \ldots, m$, one derives from here the relation

$$|\langle \mu_i, y \rangle| < \min\left\{a_i, -\alpha^{-1} a_i\right\} \quad (i = 1, \ldots, m).$$

On the other hand, in the present first situation of case distinction, there exists at least one $a_i < 0$. Then, however, the right hand side of the last inequality becomes negative which yields a contradiction.

With $x, y \in M(p) = M(p)\backslash\{0\} \subseteq \Omega_2^{(i)}$ and the $\Omega_2^{(i)}$ being convex sets for $i = 1, \ldots, m$, it results that $x_\lambda \in \Omega_2^{(i)}$. The convexity of the f_i on $\Omega_2^{(i)}$ allows to continue as

$$f_i(x_\lambda) \leq \lambda f_i(x) + (1-\lambda) f_i(y) \quad (i = 1, \ldots, m).$$

On the other hand, we know that $x, y, x_\lambda \neq 0$, whence the f_i-values may be replaced by those of the g_i^{-2} (see above):

$$g_i^{-2}(x_\lambda) \leq \lambda g_i^{-2}(x) + (1-\lambda) g_i^{-2}(y) \quad (i = 1, \ldots, m).$$

Moreover, as the g_i are finite-valued and positive on $\Omega_2^{(i)}\backslash\{0\}$ (see above), so are the g_i^{-2}. This allows to raise the last inequality to the power $-1/2$ in order to derive at (20) as desired.

Second case: $\min\limits_{i=1,\ldots,m} a_i \geq 0$

Then, $0 \in M(p)$. Consequently, we may assume that $x_\lambda \neq 0$. This already excludes the case $x = y = 0$. Next suppose that, say, $x \neq 0$ and $y = 0$. Then, we may apply Proposition 5.1, to derive that $x_\lambda = \lambda x \in M(p)$. The case $y \neq 0$ and $x = 0$ follows by symmetry. Summarizing, we may assume that $x, y, x_\lambda \neq 0$ which allows to repeat the argumentation from the first case and then to invoke again (20) in order to verify that $x_\lambda \in M(p)$. ■

We note that the assumption of independent rows ξ_i in Theorem 5.1 does not mean independence of all entries of Ξ. Rather, the cross-covariance matrices $\mathrm{cov}\,(\xi_i, \xi_j)$ are required to be zero for $i \neq j$ whereas there are no restrictions for $i = j$.

Remark 5.1 *If the value u^* in Theorem 5.1 happens to be smaller than $\sqrt{3}$, (e.g., for mean vectors $\|\mu_i\|$ close to zero), then convexity of $M(p)$ can be derived for $p > \Phi(\sqrt{3}) \approx 0.958$.*

Acknowledgement:

The authors thank Prof. W. Römisch (Humboldt Universiy Berlin) for stimulating discussion on the topic of the paper.

References

[1] V.S. Bawa, "On chance-constrained programming problems with joint constraints," Manag. Sci. vol. 19, pp. 1326-1331, 1973.

[2] C. Borell, "Convex set functions in d-space," Period. Math. Hungar. vol. 6, pp. 111-136, 1975.

[3] A. Burkauskas, "On the convexity problem of probabilistic constrained stochastic programming problems," (in Hungarian) Alkalmz. Mat. Lapok vol. 12, pp. 77-90, 1986.

[4] N. Gröwe, "Estimated stochastic programs with chance constraints," European J. Oper. Res. vol. 101, pp. 285-305, 1997.

[5] R. Henrion, "Structural Properties of Linear Probabilistic Constraints," to appear in: Optimization.

[6] Kataoka, S., "A Stochastic Programming Model," Econometrica vol. 31, pp. 181-196, 1963 .

[7] B.L. Miller and H.M. Wagner, "Chance constrained programming with joint constraints," Oper. Res. vol. 13, pp. 930-945, 1965 .

[8] C. van de Panne and W. Popp, "Minimum-Cost Cattle Feed under Probabilistic Protein Constraints," Manag. Sci. vol. 9, pp. 405-430, 1963.

[9] A. Prékopa, "On logarithmic concave measures and functions," Acta Sci. Math. (Szeged) vol. 34, pp. 335-343, 1973.

[10] A. Prékopa, "Programming under probabilistic constraints with a random technology matrix," Optimization vol. 5, pp. 109-116, 1974.

[11] A. Prékopa, Stochastic Programming, Kluwer: Dordrecht, 1995.

[12] A. Prékopa, "Probabilistic Programming," Chapter V in: A. Ruszczynski and A. Shapiro (eds.): Stochastic Programming, Hamdbooks in Operations Research and Management Science, Vol. 10, Elsevier, Amsterdam, 2003.