On the use of copulae for the joint probabilistic constrained problems

Michal Houda and Abdel Lisser

Laboratoire de Recherche en Informatique
Université Paris Sud XI

PGMO Conference
Paris, October 3–4, 2013

General chance constrained optimization problem

$$\min c(x) \quad \text{s.t.} \quad G(x) := \mathbb{P}\{g_k(x;\xi) \geq 0,\ k \in \mathcal{K}\} \geq p,\ x \in X_0 \tag{1}$$

$x \in \mathbb{R}^n$ ... decision vector
$X_0 \subset \mathbb{R}^n$ ... deterministic constraints
$\xi \in \mathbb{R}^s$ ... random vector
p ... fixed (high) probability
$\mathcal{K}$ ... index set for constraints g_k

Issues to address

- convexity of the feasible set
- computational effort.

$$\min c(x) \quad \text{s.t.} \quad G(x) = \mathbb{P}\{g_k(x;\xi) \geq 0,\ k \in \mathcal{K}\} \geq p,\ x \in X_0 \tag{1}$$

For simplicity, we assume

$$X_0 = \mathbb{R}^n \qquad c(x) = c^T x \qquad \mathcal{K} = \{1, \ldots, K\}.$$

Very classical result

Proposition 1

The feasible set is convex $\Leftrightarrow$ $G(x)$ is quasi-concave (on X_0).

We will concentrate in two important special cases

1. nonlinear constraints with random right-hand side
2. linear constraints with random coefficient matrix

Constraints with random right hand side

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{g_k(x) \geq \xi_k,\ k \in \mathcal{K}\} \geq p \tag{2}$$

i. e.

$$g_k(x;\xi) = g_k(x) - \xi_k$$
$$G(x) = \mathbb{P}\{g_k(x) \geq \xi,\ k \in \mathcal{K}\} \geq p,$$

Denote

$$M(p) := \{x \mid \mathbb{P}\{g_k(x) \geq \xi_k,\ k \in \mathcal{K}\} \geq p\}$$

... set of feasible solutions of the problem (2).

Constraints with linear random matrix

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{\Xi x \leq h\} \geq p \tag{3}$$

i. e.

$$\begin{aligned} \xi &= ({\Xi_1}^T, \ldots, {\Xi_k}^T) \\ g_k(x;\xi) &= h_k - \Xi x \\ G(x) &= \mathbb{P}\{\Xi_k x \leq h_k,\ k \in \mathcal{K}\} \geq p, \end{aligned}$$

Denote

$$X(p) := \{x \mid \mathbb{P}\{\Xi x \leq h\} \geq p\}$$

... set of feasible solutions of the problem (3).

Outline of the talk

1. Introduction — problem formulation
   - finished
2. Dependence of constraint rows
   - overcoming row independence conditions
   - notion of copulae
3. Convexity of the feasible set
   - classical results (first)
   - extended result
4. Approximation
   - formulating the problem as a second-order cone programming (SOCP) problem

Dependence

- Many practical results concerning chance-constrained programming is based on the assumption of **independence** of random rows of the problem.
- The proof are usually based on the separability property of the form

$$F(\xi) = \prod_{k \in \mathcal{K}} F_k(\xi_k)$$

  where F is the distribution function of ξ and F_k its marginals.
- first approach (Houda (2009)): approximate the problem employing the difference

$$F(\xi) - \prod_{k \in \mathcal{K}} F_k(\xi_k)$$

- this presentation: use of copulae (Nelsen (2006))

Definition 2

The copula is the distribution function $C : [0;1]^K \to [0;1]$ of some K-dimensional random vector whose marginals are uniformly distributed on $[0;1]$.

Proposition 3 (Sklar's theorem)

For any K-dimensional distribution function $F : \mathbb{R}^K \to [0;1]$ with marginals $F_1, \ldots, F_K$, there exists a copula C such that

$$\forall z \in \mathbb{R}^K \quad F(z) = C\big(F_1(z_1), \ldots, F_K(z_K)\big).$$

If, moreover, F_k are continuous, then C is uniquely given by

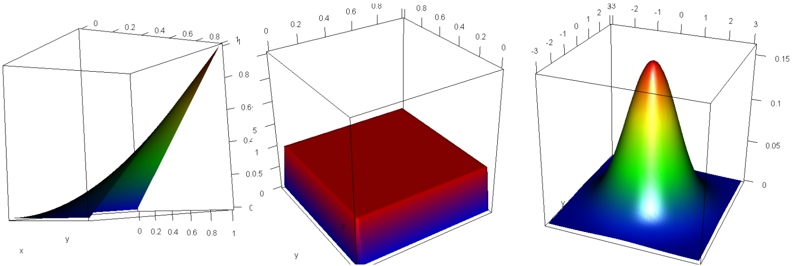
$$C(u) = F\left(F_1^{-1}(u_1), \ldots, F_K^{-1}(u_K)\right).$$

Otherwise, C is uniquely determined on $\operatorname{range} F_1 \times \cdots \times \operatorname{range} F_K$.

Assume $u = (u_1, \ldots, u_K)^T \in \mathbb{R}^K$

1. independence (product) copula

$$\Pi(u) := \prod_{k=1}^{K} u_k,$$



3. maximum (comonotone) copula

$$M(u) := \min_k \{u_k\}$$

We also define

$$W(u) := \max\left\{\sum_{k=1}^{K} u_k - K + 1, 0\right\}$$

(W fails to be a copula for $K > 2$)

Proposition 4 (Fréchet-Hoeffding bounds)

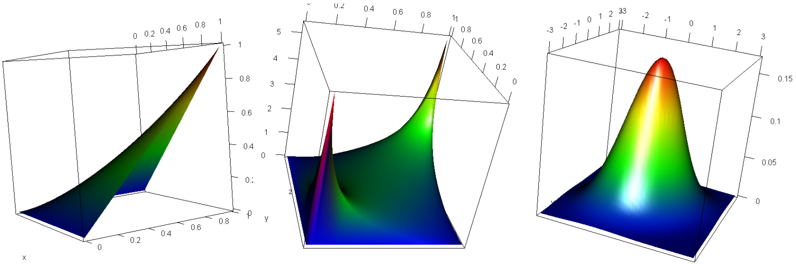
For every copula C and every $u \in \operatorname{dom} C$

$$W(u) \leq C(u) \leq M(u)$$

3. Gaussian (normal) copula

$$C^{\Sigma}(u) := \prod_{k} \Phi^{\Sigma}\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_K)\right)$$

where Φ^{Σ} is the distribution function of multivariate normal distribution with zero mean, unit variance and correlation matrix Σ



Definition 5

A copula C is called Archimedean if there exists a continuous strictly decreasing function $\psi : [0;1] \to \mathbb{R}_+$, called generator of C, such that $\psi(1) = 0$ and

$$C(u) = \psi^{-1}\left(\sum_{i=1}^{n} \psi(u_i)\right). \quad (4)$$

If $\lim_{u\to 0} \psi(u) = +\infty$ then C is called a strict Archimedean copula and ψ is called a strict generator.

Properties of copula generator

- ψ is convex
- ψ^{-1} is continuous strictly decreasing convex on $[0; \psi(0)]$

Proposition 6

Let $\psi : [0;1] \to \mathbb{R}^+$ *be convex, strictly decreasing function with* $\psi(1) = 0$, $\lim_{u\to 0} \psi(u) = +\infty$, *and*

$$(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}t^k} \psi^{-1}(t) \geq 0 \ \forall k = 0, 1, \ldots, K \text{ and } \forall t \in \mathbb{R}^+.$$

Then ψ *is a strict copula generator.*

Proposition 7

Let $\psi : [0;1] \to \mathbb{R}^+$ *be convex, strictly decreasing function with* $\psi(1) = 0$, $\lim_{u\to 0} \psi(u) = +\infty$. *Then* ψ *is a strict copula generator for every* $K \geq 2$ *iif*

$$(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}t^k} \psi^{-1}(t) \geq 0 \ \forall k = 0, 1, \ldots, \text{ and } \forall t \in \mathbb{R}^+.$$

Such function ψ^{-1} is called completely monotonic.

Practical advantages of Archimedean copulae:

- associative class (separability)
- explicit formulae of generators for most common copulae
- model dependence through one "nice" generator function and usually one or a small number of parameters
- many families adapted to a concrete problem setting (Nelsen (2006) provides a table of 22 one-parameter families of Archimedean copulae)

Some known shortcomings:

- Gaussian copula is not Archimedean
- all copula margins are the same $\Rightarrow$ only symmetric dependence structures described
- limited area of dependence structures (due to small number of parameters)
- less ability to capture negative dependence

1. independence (product) copula

$$\psi(t) := -\ln t,$$

2. Gumbel-Hougaard copulae, with $\theta \geq 1$

$$\psi_\theta(u) := (-\ln t)^\theta$$

Limiting cases

$$C_1 = \Pi \qquad C_{+\infty} = M$$

[see animation]

3. Clayton copulae, with with $\theta \geq 0$

$$\psi_\theta(u) := -\frac{1}{\theta}(t^{-\theta} - 1)$$

Limiting cases

$$C_0 = \Pi \qquad C_{+\infty} = M$$

[see animation]

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{g_k(x) \geq \xi_k,\ k \in \mathcal{K}\} \geq p \tag{2}$$

Theorem 8

If the joint distribution of ξ is driven by an Archimedean copula with the generator ψ then the feasible set of chance-constrained problem (2) can be equivalently written as

$$M(p) = \Big\{ x \in X \mid \exists y_k \geq 0,\ \sum_{k=1}^{K} y_k = 1 : \\ \psi[F_k(g_k(x))] \leq \psi(p) y_k \text{ for every } k \in \mathcal{K} \Big\}. \tag{5}$$

where F_k are one-dimensional distribution functions of components ξ_k of ξ.

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{\Xi x \leq h\} \geq p \tag{3}$$

- If the rows ${\Xi_k}^T$ are independent and follow $N_n(\mu_k; \Sigma_k \succ 0)$ then

$$\xi_k(x) := \frac{\Xi_k^T x - \mu_k^T x}{\sqrt{x^T \Sigma_k x}}, \qquad g_k(x) := \frac{h_k - \mu_k^T x}{\sqrt{x^T \Sigma_k x}},$$

  so

$$G(x) = \mathbb{P}\{g_k(x) \geq \xi_k(x),\ k \in \mathcal{K}\}$$

  where $\xi_k(x) \sim N(0; 1)$.

- If $K = 1$ (only one constraint) then

$$X(p) = \left\{ x \in X \mid {\mu_1}^T x + \Phi^{-1}(p)\sqrt{x^T \Sigma_1 x} \leq h_1 \right\}.$$

- To extend the result to $K > 1$ (more constraint rows): introduce auxiliary variables y_k. If the rows are independent, then (Cheng, Lisser(2012))

$$X(p) = \Bigg\{ x \in X \mid \exists y_k \geq 0,\ \sum_{k=1}^{K} y_k = 1 :$$

$$\qquad {\mu_k}^T x + \Phi^{-1}\left(p^{y_k}\right)\sqrt{x^T \Sigma_k x} \leq h_k \text{ for every } k \in \mathcal{K} \Bigg\}.$$

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{\Xi x \leq h\} \geq p \tag{3}$$

Theorem 9

Suppose $\Xi_k^T \sim N(\mu_k, \Sigma_k \succ 0)$. *Then the feasible set of* (3) *is*

$$X(p) = \Big\{ x \in X \mid \exists y_k \geq 0,\ \sum_{k=1}^{K} y_k = 1 : \\ {\mu_k}^T x + \Phi^{-1}\left(\psi^{-1}(y_k \psi(p))\right) \sqrt{x^T \Sigma_k x} \leq h_k,\ k \in \mathcal{K} \Big\}$$

where

1. Φ *is one-dimensional standard normal distribution function, and*
2. ψ *is a generator of an Archimedean copula describing the dependence properties of the rows of the matrix* Ξ.

Definition 10

$f(\cdot)$ is r-concave iif for any x, y such that $f(x) > 0, f(y) > 0$ and $\lambda \in [0; 1]$ we have

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

cases $r = -\infty, 0, +\infty$ treated by continuity.

$r = -\infty$	$RHS = \min\{f(x), f(y)\}$	... f quasi-concave
$r < 0$		... f^r convex
$r = 0$	$RHS = f^\lambda(x) f^{1-\lambda}(y)$	... f log-concave ($\log f$ concave)
$r > 0$		... f^r concave
$r = 1$		... f concave
$r = +\infty$	$RHS = \max\{f(x), f(y)\}$	... f quasi-convex

If f is r^*-concave, it is also r-concave for all $r \leq r^*$
$\Rightarrow$ every r-concave function is quasi-concave.

Prékopa (1971) (log-concave measures)

Definition 11

Probability measure $\mathbb{P}$ (on $\mathcal{B}(\mathbb{R}^s)$) is r-concave iif for any Borel convex subset A, B such that $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$ and $\lambda \in [0;1]$ we have

$$\mathbb{P}(\lambda A + (1-\lambda)B) \geq [\lambda \mathbb{P}^r(A) + (1-\lambda)\mathbb{P}^r(B)]^{1/r}$$

cases $r = -\infty, 0, +\infty$ treated by continuity.

Properties:

1. r-concave probability measure has r-concave distribution function
2. quasi-concave measure $\mathbb{P}$ on $\mathbb{R}^s$ with $\dim \operatorname{supp} \mathbb{P} = s$ has a density
3. r-concave density $\Leftrightarrow$ $\frac{r}{1+mr}$-concave measure on convex subset $\Omega \subset \mathbb{R}^s$ of dim $m > 0$ (so for $r > -\frac{1}{m}$)

Borell (1975), Brascamp, Lieb (1976)

Theorem 12

If

1. $g_k(\cdot,\cdot), k \in \mathcal{K}$ ... *quasi-concave jointly in both arguments*
2. $\xi \in \mathbb{R}^s$ ... *random variable with r-concave probability distribution*

then the feasible set

$$X_p := \{x \mid G(x) \geq p\}$$

is convex and closed.

- consequence of a general theorem by Prékopa (1971), see also Prékopa (1995) Shapiro, Ruszczyński, Dentcheva (2009))
- in fact, (ii) means quasi-concave probability distribution (or $-\frac{1}{s}$-concave density);
- many prominent multivariate distributions has log-concave (or at least quasi-concave and/or for some parameters) distribution so that (ii) is not tie (uniform, normal, Wishart, Beta, Dirichlet, Gamma, Cauchy, Pareto)

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{g_k(x) \geq \xi_k,\ k \in \mathcal{K}\} \geq p \qquad (2)$$

Recall $g_k(x;\xi) = g_k(x) - \xi_k$

- $g_k(x;\xi)$ required quasi-concave in both arguments
- Issue: quasi-concavity not preserved under additions
- Classical result: if $g_k(\cdot)$ concave $\Rightarrow$ $g_k(\cdot;\cdot)$ concave $\Rightarrow$ assumption (i) of Theorem 12 fulfilled.

Idea: Henrion, Strugarek (2008)

- relax concavity requirement on $g_k(\cdot)$
- strengthen quasi-concavity requirement on distribution of ξ

Definition 13 (Henrion, Strugarek (2008))

$f_\xi(\cdot)$ is an r-decreasing density of one-dimensional random variable ξ if

1. f_ξ is continuous on $(0; +\infty)$
2. $t^r f_\xi(t)$ is (strictly) decreasing for all $t > t^*$ and some treshold $t^* > 0$

- $r = 0$ ... density decreasing in classical sense
- f_ξ r^*-decreasing for some $r^* \Rightarrow r$-decreasing for all $r \geq r^*$

Lemma 14

The normal distribution has r-decreasing density for every $r > 0$ with threshold $t^* = \frac{1}{2}\left(\mu + \sqrt{\mu^2 + 4r\sigma^2}\right)$

- for standard normal distribution: $t^* = \sqrt{r}$

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{g_k(x) \geq \xi_k,\ k \in \mathcal{K}\} \geq p \tag{2}$$

Proposition 15 (Henrion, Strugarek (2008))

If

1. $g_k(\cdot)$ *are* $(-r_k)$*-concave with* $r_k > 0$
2. ξ_k *are independent with* $(r_k + 1)$*-decreasing densities with thresholds* $t_k^* > 0$
3. $p > \max\{F_k(t_k^*)\}$

Then the feasible set $M(p)$ *of* (2) *is convex.*

- Henrion and Strugarek(2011) generalized the result for dependency given by so-called *log-exp-concave* copulae
- our result similar, but based on properties of Archimedean copulae

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{g_k(x) \geq \xi_k,\ k \in \mathcal{K}\} \geq p \tag{2}$$

Theorem 16

If

1. $g_k(\cdot)$ *are* $(-r_k)$*-concave with* $r_k > 0$
2. *the marginal distributions functions* F_k *have* $(r_k + 1)$*-decreasing densities with thresholds* t_k^*
3. *the joint distribution of* ξ *is driven by an Archimedean copula with a strict generator* ψ *such that* ψ^{-1} *is completely monotonic;*
4. $p > \max\{F_k(t_k^*)\}$

Then the feasible set $M(p)$ *of* (2) *is convex.*

- independent, Gumbel-Hougaard, Clayton (and many other) copulae satisfy assumption (iii) of Theorem 16

Convexity of the feasible set: main result

Chance-constrained problem with dependent constraint rows

$$\min c^T x \quad \text{s.t.} \quad G(x) = \mathbb{P}\{\Xi x \leq h\} \geq p \tag{3}$$

Theorem 17

If

1. *the rows* $\Xi_k^T \sim N(\mu_k, \Sigma_k \succ 0)$*;*
2. *the joint distribution function of standardized rows is driven by an Archimedean copula with a generator* ψ

then the problem (3) *can be equivalently written as*

$$\min c^T x \quad s.t. \quad {\mu_k}^T x + \Phi^{-1}\left(\psi^{-1}(y_k \psi(p))\right)\sqrt{x^T \Sigma_k x} \leq h_k,$$
$$\sum\nolimits_k y_k = 1, \; y_k \geq 0, \; k \in \mathcal{K}.$$

Moreover, if

3. ψ^{-1} *is completely monotonic;*
4. $p > p^* := \Phi\left(\max\{\sqrt{3}, 4\lambda_{max}^{(k)}[\lambda_{min}^{(k)}]^{-3/2}\|\mu_k\|\}\right)$*, where* $\lambda_{max}^{(k)}, \lambda_{min}^{(k)}$ *are largest and lowest eigenvalues of the matrices* Σ_k*,*

then the problem is convex.

Lemma 18

The function $y \mapsto \Phi^{-1}\left(\psi^{-1}(y_k \psi(p))\right)$ *is convex.*

Consequence: linear approximation results in

1. piecewise tangent approximation ($\Rightarrow$ outer bound for feasible solutions)
2. piecewise linear approximation ($\Rightarrow$ inner bound for feasible solutions)

Hence we obtain two SOCP problems of the form

$$\min c^T x \quad \text{s.t.} \quad \begin{aligned} & {\mu_k}^T x + \sqrt{{z^k}^T \Sigma_k z^k} \le h_k, \\ & z^k \ge a_{kj} x + b_{kj} w^k, \ k \in \mathcal{K}, \ j \in \mathcal{J} \\ & \textstyle\sum_k w^k = x, \ w^k \ge 0, \ z^k \ge 0, \ k \in \mathcal{K}. \end{aligned}$$

$\mathcal{J}$ is the index set of partition points of $[0;1]$, a_{kj} and b_{kj} are given by the linear approximation above.

Details not presented here (technique described in Cheng, Lisser(2012)).

Conclusions

In this presentation

- Equivalent deterministic formulation of the linear chance-constrained problems with random matrix of dependent rows was given.
- Row dependence modeled through Archimedean copulae.
- Convexity of the feasible set proven for sufficiently high probabilities.
- Inner and outer approximation of the feasible set formulated as SOCP problems.

Under actual investigation

- possible extensions over normal distribution
- checking different class of Archimedean copulae for the convexity
- copula estimation
- numerical tests (comparison with other numerical methods)
- application in power industry

References I

Nelsen, R. J. (2006).
An Introduction to Copulas.
Springer, New York. 2nd edition.

Prékopa, A. (1995).
Stochastic Programming.
Akadémiai Kiadó, Budapest.

Ruszczyński, A., Shapiro, A. (2003)
Stochastic Programming. Handbooks in Operations Research and Management Science, vol. 10.
Elsevier, Amsterdam.

Shapiro, A., Dentcheva, D., Ruszczyński, A. (2009).
Lectures on Stochastic Programming: Modelling and Theory.
MPS-SIAM, Philadelphia.

References II

Borell, C. (1975).
Convex set functions in d-space.
Periodica Mathematica Hungarica **6**, 111–136.

Brascamp, H. J., Lieb, E. H. (1976).
On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equations.
Journal of Functional Analysis **22**, 366–389.

Prékopa, A. (1971).
Logarithmic concave measures with applications to stochastic programming.
Acta Scientiarium Mathematicarum (Szeged) **32**, 301–316.

References III

Cheng, J. and Lisser, A. (2012)
A second-order cone programming approach for linear programs with joint probabilistic constraints.
Operations Research Letters **40**(5), 325–328.

Henrion, R. and Strugarek, C. (2008).
Convexity of chance constraints with independent random variables.
Computational Optimization and Applications, **41**(2), 263–276.

Henrion, R. and Strugarek, C. (2011)
Convexity of chance constraints with dependent random variables: The use of copulae.
In *Stochastic Optimization Methods in Finance and Energy* (Bertocchi, M., Consigli, G., and Dempster, M. A. H. eds.), *International Series in Operations Research & Management Science* **163**, Springer, New York, 427–439.

Houda, M. (2009)
Stability and Estimates for Stochastic Programs.
Ph.D. thesis. Charles University in Prague, Faculty of Mathematics and Physics.