Distributionally robust chance constrained geometric optimization

Jia Liu\textsuperscript{a}, Abdel Lisser\textsuperscript{b}, Zhiping Chen\textsuperscript{a}

\textsuperscript{a}: School of Mathematics and Statistics, Xi’an Jiaotong University, Xi’an, Shaanxi, 710049, P. R. China,
jialiu@xjtu.edu.cn; zchen@xjtu.edu.cn

\textsuperscript{b}: Laboratoire de Recherche en Informatique (LRI), Université Paris Sud - XI, Bât. 650, 91405 Orsay Cedex, France,

lisser@lri.fr

July 15, 2019

Abstract

This paper discusses distributionally robust geometric programs with individual and joint chance constraints. Seven groups of uncertainty sets are considered: uncertainty sets with first two order moments information, uncertainty sets constrained by the Kullback-Leibler divergence distance with a normal reference distribution or a discrete reference distribution, uncertainty sets with known first moments or known first two order moments information and nonnegative support, and the joint uncertainty sets for the product of random variables. For the seven groups of uncertainty sets, we find tractable reformulations of the distributionally robust geometric programs with both individual and joint chance constraints. Finally, numerical tests are carried out on a shape optimization problem.

\textbf{Keywords:} Distributionally robust optimization; Chance constraints; Uncertainty sets; Geometric programming

1 Introduction

Geometric programs are a type of optimization problems characterized by objective and constraint functions which have a special form. A geometric program can be formulated
as:

\[(GP) \quad \min_t g_0(t) \text{ s.t. } g_k(t) \leq 1, \; k = 1, \ldots, K, \; t \in \mathbb{R}_+^M \quad (1)\]

with

\[g_k(t) = \sum_{i=1}^{I_k} c^k_i \prod_{j=1}^{M} t_j^{a^k_{ij}}, \; k = 0, \ldots, K. \quad (2)\]

Usually, \(c^k_i \prod_{j=1}^{M} t_j^{a^k_{ij}}\) is called a monomial, and \(g_k(t)\) is known as a posynomial. The coefficients \(c^k_i, i = 1, \ldots, I_k, k = 0, \ldots, K\), are nonnegative real numbers. Slightly different from Dupačová’s formulation (2009), the posynomials in different constraints might have different parameters \(c^k_i\) and \(a^k_{ij}\).

**Example:** A classic example of geometric programs is the shape optimization problem widely applied in the design and construction of structural mechanics and in the optimal control of distributed parameter systems (Boyd et al., 2007). A three-dimensional example of the shape optimization problem can be formulated as

\[
\begin{align*}
\max_{h,w,\zeta} & \quad hw\zeta \\
\text{s.t.} & \quad 2hw + 2h\zeta \leq A_{\text{wall}}, \\
& \quad w\zeta \leq A_{\text{flr}}, \\
& \quad \alpha \leq hw^{-1} \leq \beta, \\
& \quad \gamma \leq \zeta w^{-1} \leq \delta.
\end{align*}
\]

Here, we maximize the volume of a box-shaped structure with height \(h\), width \(w\) and depth \(\zeta\). There are two constraints on the total wall area \(2(hw + h\zeta)\), and the floor area \(w\zeta\), respectively. Moreover, there are some lower and upper bounds on the aspect ratios \(h/w\) and \(\zeta/w\).

Geometric programs are not convex with respect to \(t\). However, they are convex with respect to \(\{r : r_j = \log t_j, \; j = 1, \ldots, M\}\) (Ecker, 1980; Boyd et al., 2007).

In real world applications, the parameters of monomials may not be known in advance, e.g., the wall area \(A_{\text{wall}}\) or the floor area \(A_{\text{flr}}\) in the shape optimization problem. As a result, stochastic geometric programs are proposed to model geometric problems with random parameters. Most of the literature on stochastic geometric programs consider random \(c^k_i, \forall k, i\) but deterministic \(a^k_{ij}, \forall k, i, j\). There are two reasons for this: firstly, the shape of monomials is generally fixed while their combination coefficients might be random, e.g., the shape optimization problems (see for instance, Rao, 1996, Liu et al., 2009).
2016). Secondly, it is more tractable to handle a linear combination of random variables than a product of parameters with random exponents.

Rao (1996) and Dupačová (2009) formulated the following chance constrained geometric program with random \( c^k_i, \forall k, i \), and study the properties of its deterministic equivalent model:

\[
\begin{align*}
(ISGP) \quad & \min_{t \in \mathbb{R}_{++}^M} \mathbb{E}_{F_0} \left[ \sum_{i=1}^{I_0} c^0_i \prod_{j=1}^{M} t^{a_{ij}}_j \right] \\
& \text{s.t.} \quad \mathbb{P}_{F_k} \left( \sum_{i=1}^{I_k} c^k_i \prod_{j=1}^{M} t^{a_{ij}}_j \leq 1 \right) \geq 1 - \epsilon_k, \ k = 1, \ldots, K,
\end{align*}
\]

where \( F_0 \) is the probability distribution of \( c^0 = [c^0_1, c^0_2, \ldots, c^0_{I_0}] \), \( F_k \) is the probability distribution of \( c^k = [c^k_1, c^k_2, \ldots, c^k_{I_k}] \), \( k = 1, \ldots, K \), \( \epsilon_k \in (0, 1) \) is the tolerance probability for \( k^{th} \) constraint, \( k = 1, \ldots, K \).

In the same vein, Liu et al. (2016) studied stochastic geometric problems with joint chance constraints,

\[
\begin{align*}
(JSGP) \quad & \min_{t \in \mathbb{R}_{++}^M} \mathbb{E}_{F_0} \left[ \sum_{i=1}^{I_0} c^0_i \prod_{j=1}^{M} t^{a_{ij}}_j \right] \\
& \text{s.t.} \quad \mathbb{P}_{F} \left( \sum_{i=1}^{I_k} c^k_i \prod_{j=1}^{M} t^{a_{ij}}_j \leq 1, \ k = 1, \ldots, K \right) \geq 1 - \epsilon,
\end{align*}
\]

where \( F \) is the joint probability distribution of \( c^1, c^2, \ldots, c^K \).

A basic assumption in stochastic geometric programming problems is that the probability distribution of the random parameters is known beforehand. However, the underlying distribution of the random parameters cannot be exactly specified in advance for many real life problems. Distributionally robust optimization technique is thus proposed to measure the ambiguity of the distribution and has been successfully applied in many areas of optimization, see for instance, Scarf (1958), Žáčková (1966), El Ghaoui et al. (2003) and Delage and Ye (2010).

Considering the ambiguity of \( F_0, F_k \) or \( F \), we study in this paper the following distributionally robust geometric programs with individual or joint chance constraints.
\[(IRGP) \quad \min_{t \in \mathbb{R}^{+M}} \sup_{F_0 \in \mathcal{F}_0} \mathbb{E}_{F_0} \left[ \sum_{i=1}^{I_0} c_i^0 \prod_{j=1}^{M} t_{ij}^0 \right] \quad (7)\]

\[\text{s.t.} \quad \inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( \sum_{i=1}^{I_k} c_i^k \prod_{j=1}^{M} t_{ij}^k \leq 1 \right) \geq 1 - \epsilon_k, \quad k = 1, \ldots, K. \quad (8)\]

\[(JRGB) \quad \min_{t \in \mathbb{R}^{+M}} \sup_{F_0 \in \mathcal{F}_0} \mathbb{E}_{F_0} \left[ \sum_{i=1}^{I_0} c_i^0 \prod_{j=1}^{M} t_{ij}^0 \right] \quad (9)\]

\[\text{s.t.} \quad \inf_{F \in \mathcal{F}} \mathbb{P}_F \left( \sum_{i=1}^{I_k} c_i^k \prod_{j=1}^{M} t_{ij}^k \leq 1, \quad k = 1, \ldots, K \right) \geq 1 - \epsilon, \quad (10)\]

where \(\mathcal{F}_0, \mathcal{F}_k, k = 1, \ldots, K\) and \(\mathcal{F}\) are the uncertainty sets, which contain all possible distributions of \(F_0, F_k, k = 1, \ldots, K\) and \(F\), respectively.

In the next section, we briefly review the existing work in geometric optimization and distributionally robust optimization.

## 2 Literature review

Geometric programs have been studied for several decades, they were first introduced by Duffin et al. in the late 1960s. Applications of geometric programs can be found in several survey papers, namely Peterson (1976), Ecker (1980) and Boyd et al. (2007).

Stochastic geometric programs with individual chance constraints (ISGP) were discussed in Rao (1996) and Dupačová (2009), where the authors showed that the chance constraints (4) are equivalent to several deterministic constraints involving posynomials and common additional slack variables. They assume that the parameters \(a_{ij}^k, \forall k, i, j\), are deterministic parameters whilst the parameters \(c_i^k, \forall k, i\), are uncorrelated normally distributed random variables. In the same vein, Liu et al. (2016) studied stochastic geometric programs with joint chance constraints (JSGP) under the similar setting, and proposed tractable approximations by using piecewise linear functions and a sequential convex optimization algorithm.

Considering the uncertainty of both \(a_{ij}^k\) and \(c_i^k\), Jagannathan (1990) formulated a stochastic geometric program with random \(c_i^k\) and \(a_{ij}^k\) parameters as a multiplicative recourse problem. He showed that the deterministic equivalent is in general a convex program. Hsiung et al. (2005) studied a parametrically robust geometric programming problem and found its relationship with stochastic geometric programs. Furthermore, Hsiung
et al. (2008) derived approximation methods for the parametrically robust geometric programming problem, based on polyhedral and ellipsoidal uncertainty sets. Chasseiny and Goerigk (2014) showed that this parametrically robust geometric programming problem is co-NP hard in its natural posynomial form.

One can naturally consider the distributionally robust geometric programs when facing the ambiguity of the distributions, where the choice of the uncertainty set is always a core issue. The most adopted one is the uncertainty set with known mean and variance of the random variable (Lobo and Boyd, 2000, El Ghaoui et al., 2003). Delage and Ye (2010) proposed an uncertainty set based on the unknown first two order moments. Cheng et al. (2014) and Rujeerapaiboon et al. (2015) further considered unknown moments based uncertainty sets in the distributionally robust chance constrained knapsack problem and in the worst-case Value-at-Risk problem, respectively.

Moreover, several uncertainty sets are studied in the literature where the size of sets is controlled by a probabilistic distance; e.g., $\phi$-divergence (Ben-Tal et al., 2013), Wasserstein metric (Pflug and Wozabal, 2007, Esfahani and Kuhn, 2018), Kullback-Leibler distance as a special case of $\phi$-divergence (Hu and Hong, 2013). Jiang and Guan (2016) considered a chance constrained robust optimization problem under $\phi$-divergence and proposed related tractable reformulations.

In some practical problems, the support information should also be considered in the uncertainty sets. Delage and Ye (2010) studied the distributionally robust expected utility function with an uncertainty set based on the first two order moments and bounded support. Hanasusanto et al. (2017) studied distributionally robust joint linear chance constraints with an uncertainty set characterized by the mean and support of the uncertainties and by an upper bound on their dispersion. Other studies concerning support information are mostly with moments information (Wiesemann et al. 2014, Nataranjan et al. 2010) and use the duality theory to reformulate the distributionally robust expectation or chance constraint as semi-infinite constraints. Natarajan et al. (2011) analyzed a distributionally robust mixed 0-1 linear program with the mean vector and the second-order moment matrix information of the nonnegative objective coefficients. They showed that computing a tight upper bound for this problem requires to solve an NP-hard completely positive program.

In this paper, we consider the distributionally robust geometric programs ($IRGP$) and ($JRG$) under six groups of uncertainty sets. The first group considers the uncertainty of the distribution with known or unknown first two order moments. The second and third groups of uncertainty sets consider the uncertainty of the density of the random
parameters constrained by the Kullback-Leibler divergence; we use the normal distribution and data-driven samples as reference distributions, respectively. The forth and fifth groups of uncertainty sets are based on first order moments or first two order moments, together with nonnegative support constraints. All the above five groups of sets are formulated with random parameters $c^k$ while $a^k$ are deterministic real numbers. The last group of uncertainty sets are formulated with random parameters $c^k$ and $a^k$. For all the uncertainty sets, we reformulate (IRGP) and (JRG) as convex programs or under tractable forms.

3 Uncertainty sets with first two order moments

In this section, we consider the uncertainties of random variables in (IRGP) and (JRG), in terms both of the distribution and of the first two order moments.

3.1 Individual chance constraints case

We consider following uncertainty sets $\mathcal{F}_k$, $k = 0, \ldots, K$, for (IRGP).

Assumption 1. The uncertainty set for $c^k$ is

$$\mathcal{F}_k = \left\{ F_k \left| \begin{array}{l} (\mathbb{E}_{F_k}[c^k] - \mu^k)^\top (\Gamma^k)^{-1} (\mathbb{E}_{F_k}[c^k] - \mu^k) \leq \pi_1^k; \\
\text{Cov}_{F_k}[c^k] \preceq D \pi_2^k \Gamma^k \end{array} \right. \right\},$$

where $\text{Cov}_{F_k}[c^k] = \mathbb{E}_{F_k} \left[ (c^k - \mathbb{E}_{F_k}[c^k]) (c^k - \mathbb{E}_{F_k}[c^k])^\top \right]$, $\mu^k = [\mu_1^k, \mu_2^k, \ldots, \mu_{I_k}^k]^\top$ and $\Gamma^k = \{\sigma_{i,j}^k, i, j = 1, \ldots, I_k, k = 0, 1, \ldots, K\}$. $\mu_i^k$ is the reference value of the expected value of $c_i^k$, $i = 1, \ldots, I_k$, $k = 0, \ldots, K$. $\sigma^k_{i,j}$ is the reference value of the covariance between $c_i^k$ and $c_j^k$, $i, j = 1, \ldots, I_k$, $k = 0, \ldots, K$. We assume that $\mu^k \geq 0$ and $\Gamma^k$ is a positive semidefinite matrix, $k = 0, 1, \ldots, K$. $\pi_1^k \in \mathbb{R}_+$, $\pi_2^k \geq 1$, are two nonnegative scale parameters controlling the size of the uncertainty sets. $A \preceq_D B$ means $B - A$ is positive semidefinite.

The constraints in the uncertainty sets were proposed by Delage and Ye (2010).

In the particular case where $\pi_1^k = 0$, $\pi_2^k = 1$, $k = 0, 1, \ldots, K$, the first constraint in $\mathcal{F}_k$ is reduced to a known mean vector constraint $\mathbb{E}_{F_k}[c^k] = \mu^k$; The second constraint is reduced to $\text{Cov}_{F_k}[c^k] \preceq_D \Gamma^k$, which is further equivalent to a known covariance matrix constraint in many distributionally robust problems, see for instance El Ghaoui et al. (2003), and Chen et al. (2011).

Theorem 1. Given Assumption 1, (IRGP) is equivalent to
Then, problem (IRGP) can be reformulated as

\[
\begin{align*}
\min_{t \in \mathbb{R}_{++}^M} & \sum_{i=1}^{l_0} \mu_i^0 \prod_{j=1}^{M} t_{ij}^{a_{ij}^0} + \sqrt{\frac{1}{\pi_1^0}} \sum_{i=1}^{l_0} \sum_{p=1}^{M} \sigma_{i,p}^0 \prod_{j=1}^{M} t_{ij}^{a_{ij}^0 + a_{pj}^0} \\
\text{s.t.} & \sum_{i=1}^{l_k} \mu_i^k \prod_{j=1}^{M} t_{ij}^{a_{ij}^k} + \left( \sqrt{\frac{1 - \epsilon_k}{\epsilon_k} \sqrt{\pi_2^k} + \sqrt{\pi_1^k}} \right) \sum_{i=1}^{l_k} \sum_{p=1}^{M} \sigma_{i,p}^k \prod_{j=1}^{M} t_{ij}^{a_{ij}^k + a_{pj}^k} \leq 1, \ k = 1, \ldots, K.
\end{align*}
\]

**Proof.** We first denote

\[ w^k = \left[ \prod_{j=1}^{M} t_{ij}^{a_{ij}^k}, \ldots, \prod_{j=1}^{M} t_{ij}^{a_{ij,k^k}} \right], \ k = 0, 1, \ldots, K. \]

Then, problem (IRGP) can be reformulated as

\[
\begin{align}
\min_{t \in \mathbb{R}_{++}^M} & \sup_{F_0 \in \mathcal{F}_0} \mathbb{E}_{F_0} \left[ (c^0)^\top w^0 \right] \\
\text{s.t.} & \inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right) \geq 1 - \epsilon_k, \ k = 1, 2, \ldots, K.
\end{align}
\]

Based on the structure of the uncertainty set \( \mathcal{F}_k \), the \( k^{th} \) constraint can be written as

\[
\inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right) = \inf_{(\mu, \Gamma) \in \mathcal{U}_k} \inf_{F_k \in \mathcal{F}(\mu, \Gamma)} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right),
\]

where

\[ \mathcal{F}(\mu, \Gamma) = \left\{ F_k \mid \mathbb{E}_{F_k} [c^k] = \mu, \text{Cov}_{F_k} [c^k] = \Gamma \right\} \]

and

\[ \mathcal{U}_k = \left\{ (\mu, \Gamma) \mid (\mu - \mu)^\top (\Gamma^k)^{-1} (\mu - \mu) \leq \pi_k^k, \ \Gamma \preceq D \pi^k \Gamma_k \right\}. \]

Consider the inner optimization problem \( \inf_{F_k \in \mathcal{F}(\mu, \Gamma)} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right) \), \( (c^k)^\top w^k \) has the mean value \( \mu^\top w^k \) and the variance \( (w^k)^\top \Gamma w^k \). Moreover, for any distribution \( F_\xi \) of a random variable \( \xi \) with mean value \( \mu^\top w^k \) and variance \( (w^k)^\top \Gamma w^k \), there exists a distribution \( F_k \) of a random vector \( c^k \) with mean value \( \mu \) and covariance matrix \( \Gamma \) such that \( \mathbb{P}_{F_\xi}(\xi \in B) = \mathbb{P}_{F_k}( (c^k)^\top w^k \in B) \) for every Borel set \( B \subseteq \mathbb{R} \) (Yu et al., 2009). By taking \( B = (-\infty, 1] \), we have

\[
\inf_{F_k \in \mathcal{F}(\mu, \Gamma)} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right) = \inf_{F_{\xi_k} \in \mathcal{F}_{\xi_k}} \mathbb{P}_{F_{\xi_k}} \left( \xi_k \leq 1 \right), \ k = 1, 2, \ldots, K.
\]
where \( \mathcal{F}_{\xi_k} = \{ F_{\xi_k} | \mathbb{E}_{F_{\xi_k}}(\xi_k) = \mu^\top w^k, \text{Var}_{F_{\xi_k}}(\xi_k) = (w^k)\Gamma w^k \} \). For the sake of simplicity, we denote \( \mu_{\xi_k} = \mu^\top w^k \) and \( \sigma^2_{\xi_k} = (w^k)^\top \Gamma w^k \).

According to one-sided Chebyshev inequality, we have

\[
\inf_{F_{\xi_k} \in \mathcal{F}_{\xi_k}} \mathbb{P}_{F_{\xi_k}}(\xi_k \leq 1) = \begin{cases} 
1 - \frac{1}{1 + (\mu_{\xi_k} - 1)^2/\sigma^2_{\xi_k}}, & \text{if } \mu_{\xi_k} \leq 1, \\
0, & \text{otherwise}.
\end{cases}
\]

The tightness of the inequality under \( \mu_{\xi_k} \leq 1 \) can be seen from a two-point distribution given in Rujeerapaiboon et al. (2018).

When \( \mu_{\xi_k} = \mu^\top w^k > 1 \),

\[
\inf_{F_k \in \mathcal{G}(\mu, \Gamma)} \mathbb{P}_{F_k} ((c^k)^\top w^k \leq 1) = 0,
\]

which leads constraint (12) to be infeasible.

When \( \mu_{\xi_k} = \mu^\top w^k \leq 1 \),

\[
\inf_{F_k \in \mathcal{G}(\mu, \Gamma)} \mathbb{P}_{F_k} ((c^k)^\top w^k \leq 1) = 1 - \frac{1}{1 + (\mu^\top w^k - 1)^2/((w^k)^\top \Gamma w^k)},
\]

and constraint (12) is equivalent to

\[
\inf_{(\mu, \Gamma) \in \mathcal{W}_k} 1 - \frac{1}{1 + (\mu^\top w^k - 1)^2/((w^k)^\top \Gamma w^k)} \geq 1 - \epsilon_k.
\]

Constraint (13) can be reformulated as

\[
f_k(w^k) \geq \sqrt{\frac{1 - \epsilon_k}{\epsilon_k}},
\]

where

\[
f_k(w^k) = \min_{\mu, \Gamma} \frac{1 - \mu^\top w^k}{\sqrt{(w^k)^\top \Gamma w^k}}
\]

s.t. \( \mu - \mu^k \leq \pi_1^k \), \( \Gamma \preceq \pi_2^k \Gamma^k \).

The problem (15)-(17) can be separated into two optimization problems:

\[
f_k(w^k) = \frac{1 + v_1(w^k)}{\sqrt{v_2(w^k)}},
\]
where
\[ v_1(w^k) = \min_{\mu} -\mu^T w^k \] (19)
\[ \text{s.t. } (\mu - \mu^k)^T (\Gamma^k)^{-1} (\mu - \mu^k) \leq \pi_1^k, \] (20)
and
\[ v_2(w^k) = \max_{\Gamma} (w^k)^T \Gamma w^k \] (21)
\[ \text{s.t. } \Gamma \preceq_D \pi_2^k \Gamma^k. \] (22)

As \( \Gamma \preceq_D \Gamma^k \) implies \( u^T \Gamma u \leq u^T \Gamma^k u \) for any \( u \in \mathbb{R}^n \), we know the optimal solution \( \Gamma^* \) of problem (21)-(22) is \( \pi_2^k \Gamma^k \), and the maximum value is \( v_2(w^k) = \pi_2^k (w^k)^T \Gamma^k w^k \).

Meanwhile, by setting the Lagrange function
\[ L(\mu, \varepsilon) = -(w^k)^T \mu - \varepsilon \left( (\mu - \mu^k)^T (\Gamma^k)^{-1} (\mu - \mu^k) - \pi_1^k \right), \] \( \varepsilon \leq 0 \)
we can find the first order necessary optimality condition for problem (19)-(20) is
\[ \begin{cases} 
\frac{\partial L(\mu, \varepsilon)}{\partial \mu} = -w^k - 2\varepsilon (\Gamma^k)^{-1} (\mu - \mu^k) = 0, \\
\varepsilon \left( (\mu - \mu^k)^T (\Gamma^k)^{-1} (\mu - \mu^k) - \pi_1^k \right) = 0. 
\end{cases} \] (23)

Solving (23) gives \( \mu^* = \mu^k - \frac{1}{2\varepsilon^*} (\Gamma^k) w^k \) and \( \varepsilon^* = -\sqrt{\frac{1}{4\pi_1^k}} (w^k)^T \Gamma^k w^k \). The corresponding optimum value is
\[ v_1(w^k) = -(w^k)^T \mu^k - \sqrt{\pi_1^k} \sqrt{(w^k)^T \Gamma^k w^k}. \]

Taking the optimum values of the two sub-problems back to (14) and (18), constraint (14) is equivalent to
\[ \frac{1 - (w^k)^T \mu^k - \sqrt{\pi_1^k} \sqrt{(w^k)^T \Gamma^k w^k}}{\sqrt{\pi_2^k} \sqrt{(w^k)^T \Gamma^k w^k}} \geq \sqrt{\frac{1 - \epsilon_k}{\epsilon_k}}, \]
which can be reformulated as
\[ (w^k)^T \mu^k + \left( \sqrt{\frac{1 - \epsilon_k}{\epsilon_k}} \sqrt{\pi_1^k} + \sqrt{\pi_2^k} \right) \sqrt{(w^k)^T \Gamma^k w^k} \leq 1. \] (24)

By the positive semidefiniteness of \( \Gamma^k \) and the positivity of \( \sigma_{\epsilon_k} \), (24) implies \( (w^k)^T \mu^k \leq 1 \). Hence, (24) is equivalent to the \( k \)th constraint in (12).
Moreover, by using similar transformation techniques and solution methods, we can formulate the objective function as

\[
\sup_{F_0} \mathbb{E}_{F_0} \left[(c^0)^\top w^0 \right] = \max_{\mu} \mu^\top w^0 \\
\text{s.t.} \quad (\mu - \mu^0)^\top (\Gamma^0)^{-1} (\mu - \mu^0) \leq \pi^1_0, \\
= (w^0)^\top \mu^0 + \sqrt{\pi^0_1} \sqrt{(w^0)^\top \Gamma^0 w^0}
\] (25)

Finally, taking \((\mu^k)^\top w^k = \sum_{i=1}^{I_k} \mu^k_i \pi^k_j t^a_{ij} \) and \((w^k)^\top \Gamma^k w^k = \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma^k_{i,p} \pi^k_j t^a_{ij} + t^a_{pj}, k = 0, 1, \ldots, K, \) into (25) and (24) finalizes the proof of the theorem.

\[\square\]

Although problem \((\text{IRGP}_1)\) is not convex with respect to \(t\), we can transform it into the following optimization problem by letting \(r_j = \log(t_j), \ j = 1, \ldots, M:\)

\[
(\text{IRGP}_{1s}) \quad \min_{r \in \mathbb{R}^M} \sum_{i=1}^{I_k} \mu^0_i \exp \left\{ \sum_{j=1}^{M} a^0_{ij} r_j \right\} + \sqrt{\pi^0_1} \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma^0_{i,p} \exp \left\{ \sum_{j=1}^{M} (a^0_{ij} + a^0_{pj}) r_j \right\}} \\
\text{s.t.} \quad \left( \frac{1 - \epsilon_k}{\epsilon_k} \sqrt{\pi^2_k + \pi^1_k} \right) \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma^k_{i,p} \exp \left\{ \sum_{j=1}^{M} (a^k_{ij} + a^k_{pj}) r_j \right\}} \\
+ \sum_{i=1}^{I_k} \mu^k_i \exp \left\{ \sum_{j=1}^{M} a^k_{ij} r_j \right\} \leq 1, \ k = 1, \ldots, K.
\]

\textbf{Proposition 1.} If \(\sigma^k_{i,p} \geq 0, \) for any \(i, p = 1, \ldots, I_k, k = 0, 1, \ldots, K, \) \((\text{IRGP}_{1s})\) is a convex programming problem.

\textit{Proof.} This can be observed from the convexity of the sum of several convex exponential terms.

\[\square\]

Proposition 1 is a sufficient condition for the convexity of \((\text{IRGP}_{1s})\). They are two possible situations: some random vector components \(c^k_i\) can be pairwise uncorrelated whilst others can be positive correlated, like the conditions in Bawa (1973). Moreover, the condition in Proposition 1 suggests that \((\text{IRGP}_{1s})\) can be rewritten in order to satisfy
the rules of disciplined convex programming and thus several solvers, e.g., CVX, Ipopt or ECOS, can be used to solve \((JRGP_{1s})\). If some \(\sigma_{i,p}^k < 0\), for instance there are two random variables \(c_i^k\) and \(c_p^k\) which are negatively correlated, \((JRGP_{1s})\) is a Difference of Convex (DC) programming problem. In this case, global optimization algorithms, e.g., successive convex approximation methods, can be used to deal with the nonconvex problems.

### 3.2 Joint chance constraints case

We further consider an uncertainty set for \((JRGP)\), which considers the uncertainties in terms of the distribution and of the first two order moments. In addition, we assume that the marginal distributions are pairwise independent.

**Assumption 2.** The uncertainty set \(\mathcal{F} = \mathcal{F}_1 \times \cdots \times \mathcal{F}_K\), and for any distribution \(F\) in \(\mathcal{F}\), its marginal distributions \(F_1, \ldots, F_K\) are pairwise independent. Moreover,

\[
\mathcal{F}_k = \left\{ F_k \left| \begin{array}{c}
(\mathbb{E}_{F_k}[c_k] - \mu_k)^\top (\Gamma_k)^{-1} (\mathbb{E}_{F_k}[c_k] - \mu_k) \leq \pi_{1k}^k, \\
\text{Cov}_{F_k}[c_k] \preceq D \pi_{2k}^k \Gamma_k \\
\end{array} \right\}, \quad k = 0, 1, \ldots, K.
\]

**Theorem 2.** Given Assumption 2, \((JRGP)\) is equivalent to

\[
(JRGP_1) \quad \min_{t \in \mathbb{R}^{M_+}, y \in \mathbb{R}^K_+} \left\{ \sum_{i=1}^{I_0} \mu_i^0 \prod_{j=1}^{M} t_{ij}^{0} + \sqrt{\pi_{10}^0} \sum_{i=1}^{I_0} \sum_{p=1}^{I_k} \sigma_{i,p}^{0} \prod_{j=1}^{M} t_{ij}^{0} + a_{ij}^{0} \right\}
\]

s.t.

\[
\left\{ \sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^{M} t_{ij}^{k} + \sqrt{\pi_{1k}^k} \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^{k} \prod_{j=1}^{M} t_{ij}^{k} + a_{ij}^{k} \right\} + \sqrt{\frac{y_k}{1 - y_k}} \sqrt{\pi_{2k}^k} \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^{k} \prod_{j=1}^{M} t_{ij}^{k} + a_{ij}^{k} \leq 1, \quad k = 1, \ldots, K.
\]

\[
\prod_{k} y_k \geq 1 - \epsilon, \quad 0 \leq y_k \leq 1, \quad k = 1, \ldots, K.
\]

**Proof.** Firstly, the reformulation of the objective function is directly obtained by Theorem 1.

Based on the independence assumption of the marginal distributions, we have

\[
\inf_{F \in \mathcal{F}} \mathbb{P}_F \left( \sum_{i=1}^{I_k} \prod_{j=1}^{M} t_{ij}^{k} \leq 1, \quad k = 1, \ldots, K \right) = \prod_{k=1}^{K} \inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( \sum_{i=1}^{I_k} \prod_{j=1}^{M} t_{ij}^{k} \leq 1 \right)
\]
By introducing auxiliary variables $y_k \in \mathbb{R}_+$, $k = 1, \ldots, K$, (10) can be equivalently transformed into

$$\inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( \sum_{i=1}^{I_k} c_i^k \prod_{j=1}^{M} t_{ij}^{a_{ij}^k} \leq 1 \right) \geq y_k, \ k = 1, \ldots, K,$$

and

$$\prod_{k=1}^{K} y_k \geq 1 - \epsilon, \ 1 \geq y_k \geq 0, \ k = 1, \ldots, K.$$

We have from Theorem 1 that under uncertainty sets $\mathcal{F}_k$, constraint (26) is equivalent to

$$\sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^{M} t_{ij}^{a_{ij}^k} + \left( \sqrt{\pi_1^k} + \sqrt{\pi_2^k} \sqrt{\frac{y_k}{1-y_k}} \right) \left[ \sum_{i=1}^{I_k} \sum_{p=1}^{l_k} \sigma_{i,p}^{k} \prod_{j=1}^{M} t_{ij}^{a_{ij}^k + a_{pj}^k} \right] \leq 1, \ k = 1, \ldots, K.$$

Together with the reformulation of the objective function in Theorem 1, this gives the conclusion of this theorem.

\[\square\]

However, $(JRG\!P_1)$ is still not a convex programming problem due to the product of a nonlinear function and the square root of a posynomial. By applying the standard variable transformation widely used in geometric programming problems to $(JRG\!P_1)$, we have the following equivalent reformulation of $(JRG\!P_1)$ with new variables $r_j = \log(t_j)$, $j = 1, \ldots, M$ and $x_k = \log(y_k)$, $k = 1, \ldots, K$:

$$(JRG\!P_1_{s}) \quad \min_{r \in \mathbb{R}^M, x \in \mathbb{R}^K} \sum_{i=1}^{I_0} \mu_i^0 \exp \left\{ \sum_{j=1}^{M} a_{ij}^0 r_j \right\} + \sqrt{\pi_1^0} \left[ \sum_{i=1}^{I_0} \sum_{p=1}^{l_0} \sigma_{i,p}^{0} \exp \left\{ \sum_{j=1}^{M} (a_{ij}^0 + a_{pj}^0) r_j \right\} \right]
\text{s.t.} \quad \sum_{i=1}^{I_k} \mu_i^k \exp \left\{ \sum_{j=1}^{M} a_{ij}^k r_j \right\} + \sqrt{\pi_1^k} \left[ \sum_{i=1}^{I_k} \sum_{p=1}^{l_k} \sigma_{i,p}^{k} \exp \left\{ \sum_{j=1}^{M} (a_{ij}^k + a_{pj}^k) r_j \right\} \right]
+ \sqrt{\pi_2^k} \left[ \sum_{i=1}^{I_k} \sum_{p=1}^{l_k} \sigma_{i,p}^{k} \exp \left\{ \sum_{j=1}^{M} (a_{ij}^k + a_{pj}^k) r_j \right\} \right] \leq 1, \ s = 1, \ldots, S, \ k = 1, \ldots, K,$$

$$\sum_{k=1}^{K} x_k \geq \log(1 - \epsilon), \ x_k \leq 0, \ k = 1, \ldots, K,$$
Notice that \( \log \left( \frac{x_k}{1-x_k} \right) = x_k - \log (1 - e^{x_k}) \) is a convex function. Similarly to Proposition 1, if \( \alpha_{i,p}^k \geq 0 \) for all \( i, p, k \), \((JRGP_{1s})\) is a convex programming problem (a sufficient condition).

**Remark 1.** In the case when \( \pi_1^k = 0, \pi_2^k = 1, k = 0, 1, ..., K \), \( \mathcal{F}_k \) in Assumptions 1 and 2 are reduced to the uncertainty sets with known first two order moments information. In our numerical tests, we also investigate the case with known moments. Hence, we denote \((IRGP_1)\) and \((JRGP_1)\) with \( \pi_1^k = 0, \pi_2^k = 1, k = 0, 1, ..., K \), by \((IRGP_2)\) and \((JRGP_2)\), respectively.

In the case where \( \pi_1^k \geq 0 \) and \( \pi_2^k \geq 1, k = 0, 1, ..., K \), \((IRGP_2)\) and \((JRGP_2)\) provide upper bounds of \((IRGP_1)\) and \((JRGP_1)\), respectively.

4 Uncertainty sets with a reference distribution

In this section, we consider uncertainty sets proposed by Ben-Tal et al. (2013), for \((IRGP)\) and \((JRGP)\), which control the distance between the true distribution and the reference distribution of \( c^k \).

4.1 Case \((IRGP)\) with reference distributions

First we consider \((IRGP)\) under the following K-L distance based uncertainty set.

**Assumption 3.** The uncertainty sets are

\[
\mathcal{F}_k = \{ F_k \mid D_{KL}(F_k||F_k^0) \leq \kappa_k \}, \quad k = 0, ..., K.
\]

where \( D_{KL} \) is the Kullback-Leibler divergence distance

\[
D_{KL}(F_k||F_k^0) = \int_{\Omega} \phi \left( \frac{f_{F_k}(c^k)}{f_{F_k^0}(c^k)} \right) f_{F_k^0}(c^k) dc^k,
\]

\( F_k^0 \) is the reference distribution of \( c^k \), \( f_{F_k}(c^k) \) and \( f_{F_k^0}(c^k) \) are the density functions of the true distribution and the reference distribution of \( c^k \) on \( \Omega \), \( \kappa_k \) is a parameter controlling the size of the uncertainty set, \( k = 0, ..., K \). \( \phi(t) = t \log t - t + 1, \) for \( t \geq 0 \), and \( \phi(t) = \infty \), otherwise.

Given the above K-L distance based uncertainty sets, we have the following reformulations of the objective function and the constraints.
Proposition 2. Given Assumption 3, the objective function in (7) is equivalent to

\[
\inf_{\alpha \in [0, \infty)} \alpha \log \mathbb{E}_{F_0} \left[ \exp \left\{ \left( \sum_{i=1}^{I_0} c_i^{0} \prod_{j=1}^{M} t_{ij}^{0} \right) / \alpha \right\} \right] + \alpha \kappa_0.
\]

Proof. The proposition follows from Theorem 1 in Hu and Hong (2013).

Proposition 3. Given Assumption 3, the constraint (10) is equivalent to

\[
P_{F_0} \left( \sum_{i=1}^{I_k} c_i^{k} \prod_{j=1}^{M} t_{ij}^{k} \leq 1, \ k = 1, \ldots, K \right) \geq 1 - \epsilon',
\]

where \( \epsilon' = 1 - \inf_{x \in (0,1)} \left\{ \frac{e^{-\kappa_k x^{1-r_k} - 1}}{x-1} \right\} \).

Proof. The proposition follows from Theorem 1 and Proposition 4 in Jiang and Guan (2016).

As a special case, given Assumption 3, the constraint (8) is equivalent to

\[
P_{F_k} \left( \sum_{i=1}^{I_k} c_i^{k} \prod_{j=1}^{M} t_{ij}^{k} \leq 1 \right) \geq 1 - \epsilon'_k, \ k = 1, \ldots, K,
\]

where \( \epsilon'_k = 1 - \inf_{x \in (0,1)} \left\{ \frac{e^{-\kappa_k x^{1-r_k} - 1}}{x-1} \right\}, \ k = 1, \ldots, K. \)

In the following, we consider two kinds of reference distributions.

4.1.1 Case (IRGP) with normal reference distribution

Suppose the reference distribution \( F_k^0 \) follows a normal distribution with mean vector \( \mu^k = [\mu_1^k, \mu_2^k, \ldots, \mu_{I_k}^k]^\top \) and covariance matrix \( \Gamma^k = \{ \sigma_{i,j}^k, \ i,j = 1, \ldots, I_k \}, k = 0, 1, \ldots, K. \) Moreover, we assume that \( \mu^k \geq 0 \) and \( \Gamma^k \) is a positive semidefinite matrix, \( k = 0, 1, \ldots, K. \) Then we have
Theorem 3. Given Assumption 3 and normal distribution assumption for \( F_0^k \), \( k = 0, 1, \ldots, K \), \((IRGP)\) is equivalent to

\[
(\text{IRGP}_{3N}) \quad \min_{\mathbf{t} \in \mathbb{R}^M_{++}} \sum_{i=1}^{I_0} \mu_i^0 \prod_{j=1}^{M} t_j^{0i} + \sqrt{2\kappa_0} \sqrt{\sum_{i=1}^{I_0} \sum_{p=1}^{M} \sigma_{i,p}^0 \prod_{j=1}^{M} (t_j^{0i} + a_{pj})}
\]

s.t. \[
\sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^{M} t_j^{a_{ij}} + \Phi^{-1}(1 - \epsilon'_k) \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{M} \sigma_{i,p}^k \prod_{j=1}^{M} (t_j^{a_{ij}} + a_{pjj})} \leq 1, \quad k = 1, \ldots, K.
\]

Here, \( \Phi^{-1}(. \) is the quantile of the standard normal distribution \( N(0,1) \).

Proof. Similarly to Theorem 1, by denoting

\[
w_k = \left[ \prod_{j=1}^{M} t_j^{a_{j1}}, \ldots, \prod_{j=1}^{M} t_j^{a_{jk}} \right], \quad k = 0, 1, \ldots, K,
\]

\((IRGP)\) can be reformulated as (11)-(12).

By Proposition 2 and Proposition 3, (11)-(12) are equivalent to

\[
\min_{\mathbf{t} \in \mathbb{R}^M_{++}} \inf_{\alpha \in [0, \infty)} \alpha \log \mathbb{E}_{F_0^0} \left[ \exp \left\{ \left( (c^0)^\top w^0 \right) / \alpha \right\} \right] + \alpha \kappa_0,
\]

s.t. \[
\mathbb{P}_{F_0^0} \left( ((c^0)^\top w^k) \leq 1 \right) \geq 1 - \epsilon'_k, \quad k = 1, 2, \ldots, K,
\]

where \( \epsilon'_k \) is defined in Proposition 3.

Moreover, by the normal distribution assumption of \( F_0^0 \) for \( c^0 \), \( (c^0)^\top w^0 / \alpha \) follows a normal distribution with mean value \( \frac{1}{\alpha} (\mu^0)^\top w^0 \) and variance \( \frac{1}{\alpha^2} (w^0)^\top \Gamma_0 w^0 \). Whilst \( \exp \left\{ (c^0)^\top w^0 / \alpha \right\} \) follows a log-normal distribution with mean value \( \exp \left\{ \frac{1}{\alpha} (\mu^0)^\top w^0 + \frac{1}{2\alpha^2} (w^0)^\top \Gamma_0 w^0 \right\} \).

Therefore, the inner level optimization problem in (27) is equivalent to

\[
\inf_{\alpha \in [0, \infty)} \alpha \left( \frac{1}{\alpha} (\mu^0)^\top w^0 + \frac{1}{2\alpha^2} (w^0)^\top \Gamma_0 w^0 \right) + \alpha \kappa_0.
\]

The optimal solution of (29) is \( \alpha^* = \sqrt{\frac{(w^0)^\top \Gamma_0 w^0}{2\kappa_0}} \), and the optimal value of (29) is

\[
(\mu^0)^\top w^0 + \sqrt{2\kappa_0 (w^0)^\top \Gamma_0 w^0}.
\]

Moreover, we know from Dupačová (2009) that, when \( F_0^0 \) follows a multivariate normal distribution, (28) is equivalent to

\[
(\mu^k)^\top w^k + \Phi^{-1}(1 - \epsilon'_k) \sqrt{(w^k)^\top \Gamma^0 w^k} \leq 1, \quad k = 1, \ldots, K.
\]
Finally, taking $(\mu^k)^\top w^k = \sum_{i=1}^{I_k} \mu_i^k \prod_{j=1}^{M} t_j^{a_{ij}}$ and $(w^k)^\top \Gamma^k w^k = \sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma_{i,p}^k \prod_{j=1}^{M} t_j^{a_{ij} + a_{pj}}$, $k = 0, 1, \ldots, K$, into (30) and (31) finalizes the proof of the theorem.

The main difference between (IRGP$_1$) and (IRGP$_3$) lays in the parameter before the square root in the constraint. In this case, we can reformulate (IRGP$_3$) using the log transformation of the decision variables (Liu et al., 2016). Moreover, the obtained reformulation is convex when $\epsilon'_k \leq 0.5$ and $\sigma_{i,p}^k \geq 0$ for all $i, p, k$.

4.1.2 Case (IRGP) with data-driven reference distribution

We consider discrete reference distribution which comes from the historical data.

**Theorem 4.** Given Assumption 3, we further assume $F_k^0$ follows a discrete distribution with $H$ possible scenarios $\tilde{c}^k(1), \tilde{c}^k(2), \ldots, \tilde{c}^k(H)$, along with their respective probabilities $1/H$, $k = 0, 1, \ldots, K$. Then, problem (IRGP) is equivalent to

\[
\text{(IRGP)$_{3D}$} \quad \min_{r \in \mathbb{R}^M, \alpha \in [0, \infty), \varsigma} \quad \alpha \log \left( \frac{1}{H} \sum_{h=1}^{H} \exp \left\{ \left( \sum_{i=1}^{I_k} \tilde{c}^0_i(h) \exp \left\{ \sum_{j=1}^{M} a_{ij}^0 r_j \right\} / \alpha \right) \right\} \right) + \alpha \kappa_0,
\]

s.t.

\[
\frac{1}{H} H \sum_{h=1}^{H} \left( 1 - \varsigma_h^k \right) \geq 1 - \epsilon'_k, \quad k = 1, 2, \ldots, K,
\]

\[
\sum_{i=1}^{I_k} \tilde{c}^k_i(h) \exp \left\{ \sum_{j=1}^{M} a_{ij}^k r_j \right\} \leq M \varsigma_h^k + 1, \quad h = 1, \ldots, H, \quad k = 1, 2, \ldots, K,
\]

\[
\varsigma_h^k \in \{0, 1\}, \quad h = 1, \ldots, H, \quad k = 1, 2, \ldots, K.
\]

**Proof.** According to Theorem 1, (IRGP) can be reformulated as (11)-(12).

By Proposition 2 and Proposition 3, (11)-(12) are equivalent to

\[
\min_{\varsigma \in [0, \infty), \alpha \in [0, \infty)} \quad \alpha \log \mathbb{E} F_0^k \left[ \exp \left\{ (c^0)^\top w^0 / \alpha \right\} \right] + \alpha \kappa_0,
\]

s.t.

\[
\mathbb{E} F_k^0 \left( 1_{(-\infty, 0]}(\tilde{c}^k(h)^\top w^k - 1) \right) \geq 1 - \epsilon'_k, \quad k = 1, 2, \ldots, K,
\]

where $\epsilon'_k$ is defined in Proposition 3. $1_A(x)$ is the indicator function: $1_A(x) = 1$, if $x \in A$ and $1_A(x) = 0$, if $x \notin A$. 

16
As $F_k$ follows a discrete distribution, (32)-(33) can be reformulated as
\[
\min_{t \in \mathbb{R}_+^M, \alpha \in [0, \infty)} \alpha \log \left( \frac{1}{H} \sum_{h=1}^H \exp \left\{ \left( \bar{c}_0(h)^\top t^0 \right) / \alpha \right\} \right) + \alpha \kappa_0, 
\]
\[
\text{s.t.} \quad \frac{1}{H} \sum_{h=1}^H \left( 1 - \varsigma^k_h \right) \geq \epsilon'_k, \quad k = 1, 2, \ldots, K. 
\]
(34)

Constraint (35) is equivalent to
\[
\frac{1}{H} \sum_{h=1}^H \left( 1 - \varsigma^k_h \right) \geq \epsilon'_k, \quad k = 1, 2, \ldots, K, 
\]
\[
\bar{c}_k(h)^\top t^k \leq M \varsigma^k_h + 1, \quad h = 1, 2, \ldots, H, \quad k = 1, 2, \ldots, K, 
\]
\[
\varsigma^k_h \in \{0, 1\}, \quad h = 1, 2, \ldots, H, \quad k = 1, 2, \ldots, K, 
\]
(36)

where $\varsigma$ is an auxiliary variable and $M$ is a large number commonly used in the Big-M method.

If we take $w^k = \left[ \prod_{j=1}^M t_{ij}^{a_{ij}^h}, \ldots, \prod_{j=1}^M t_{ij}^{a_{ij}^k} \right]$, $k = 0, 1, \ldots, K$, then (32)-(33) are equivalent to
\[
\min_{t \in \mathbb{R}_+^M, \alpha \in [0, \infty), \varsigma} \alpha \log \left( \frac{1}{H} \sum_{h=1}^H \exp \left\{ \left( \sum_{i=1}^I \bar{c}_i^0(h) \prod_{j=1}^M t^0_{ij} \right) / \alpha \right\} \right) + \alpha \kappa_0, 
\]
\[
\text{s.t.} \quad \frac{1}{H} \sum_{h=1}^H \left( 1 - \varsigma^k_h \right) \geq 1 - \epsilon'_k, \quad k = 1, 2, \ldots, K, 
\]
\[
\sum_{i=1}^I \bar{c}_i^k(h) \prod_{j=1}^M t_{ij}^{a_{ij}^k} \leq M \varsigma^k_h + 1, \quad h = 1, 2, \ldots, H, \quad k = 1, 2, \ldots, K, 
\]
\[
\varsigma^k_h \in \{0, 1\}, \quad h = 1, 2, \ldots, H, \quad k = 1, 2, \ldots, K. 
\]
(37)

Finally, the change of variable $r_j = \log(t_j)$, $j = 1, \ldots, M$, gives the equivalent form, and complete the proof.

Notice that (IRGP3D) is a mixed integer programming problem, it can be solved by available free or commercial softwares.
4.2 Case \((JRGP)\) with reference distributions

We consider uncertainty sets for \((JRGP)\) with a normal or data-driven reference distribution.

**Assumption 4.** The uncertainty sets are

\[
\mathcal{F}_0 = \{ F_0 \mid D_{KL}(F_0||F^0_0) \leq \kappa_0 \} \quad \text{and} \quad \mathcal{F} = \{ F \mid D_{KL}(F||F^0) \leq \kappa \},
\]

where \(D_{KL}\) is defined in Assumption 3, \(F^0_0\) is the reference distribution for \(c_0\), \(F^0\) is the reference joint distribution for \(c^1, c^2, \ldots, c^K\), such that \(F^0 = F^0_1 \times \cdots \times F^0_K\) and the marginal distributions \(F^0, \ldots, F^0_K\) are pairwise independent.

**Theorem 5.** Given Assumption 4, we assume that \(F^0_0, F^0_1, \ldots, F^0_K\) are normally distributed with mean vector \(\mu^k = [\mu^k_1, \mu^k_2, \ldots, \mu^k_I]^T\) and covariance matrix \(\Gamma^k = \{\sigma^k_{i,j}, i,j = 1, \ldots, I_k\}, k = 0, 1, \ldots, K\). Then \((JRGP)\) is equivalent to

\[
\begin{align*}
\min_{t \in \mathbb{R}_{++}^M, y \in \mathbb{R}_+^K} & \quad I_0 \sum_{i=1}^I \mu^0_i \prod_{j=1}^M l^{0}_{ij} + \sqrt{2}\kappa_0 \sqrt{\sum_{i=1}^I \sum_{p=1}^I \sigma^0_{i,p} \prod_{j=1}^M t^{0}_{ij}^{a_{i,p}} + \Phi^{-1}(y_k)} \\
\text{s.t.} & \quad I_k \sum_{i=1}^I \mu^k_i \prod_{j=1}^M l^{k}_{ij} + \Phi^{-1}(y_k) \sqrt{\sum_{i=1}^I \sum_{p=1}^I \sigma^k_{i,p} \prod_{j=1}^M t^{k}_{ij}^{a_{i,p}} + 2 \log(\Phi^{-1}(e^{x_k}))} \leq 1, \quad k = 1, \ldots, K, (40) \\
& \quad \prod_{k=1}^K y_k \geq 1 - \epsilon', \quad 0 \leq y_k \leq 1, \quad k = 1, \ldots, K. (41)
\end{align*}
\]

where \(\epsilon' = 1 - \inf_{x \in (0,1)} \left\{ \frac{e^{-x^2} - 1}{x^2} \right\} \).

**Proof.** The proof can be done using Theorem 3 and the approach of Liu et al. (2016).

The standard variable transformation \(r_j = \log(t_j), \quad j = 1, \ldots, M\) and \(x_k = \log(y_k), \quad k = 1, \ldots, K\) applied to \((JRGP_{3N})\) leads to a posynomial objective function and the following constraint:

\[
\begin{align*}
\sum_{i=1}^{I_k} \mu^k_i \exp \left\{ \sum_{j=1}^M a^k_{ij} r_j \right\} + \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^I \sigma^k_{i,p} \exp \left\{ \sum_{j=1}^M \left( (a^k_{ij} + a^k_{pj}) r_j + 2 \log(\Phi^{-1}(e^{x_k})) \right) \right\}} \leq 1, \quad k = 1, \ldots, K, \quad (42)
\end{align*}
\]
Notice that due to the non-elementary quantile function \( \log(\Phi^{-1}(e^{\sigma_k})) \), it is difficult to deal with this constraint. The following lemma gives a sufficient condition for the convexity of the reformulated problem.

**Lemma 1.** \( \Phi^{-1}(x) \geq 0 \) for \( x \geq \frac{1}{2} \). \( \log(\Phi^{-1}(x)) \) is monotone increasing and convex on \([\Phi(1), 1)\).

Hence, the reformulated programming problem with constraint (42) is a convex problem if \( \epsilon' \leq 1 - \Phi(1) \) and \( \sigma_{i,p}^k \geq 0 \), for all \( i, p = 1, \ldots, I_k \). Notice that the condition on \( \epsilon' \) for the convexity is more strict for the joint chance constraint than for the individual chance constraint.

Moreover, for the data-driven robust geometric programming with joint chance constraints, we have

**Theorem 6.** Given Assumption 4, we assume that \( F^0_k \) is a discrete distribution with \( H \) possible values \( \tilde{c}^k(h), h = 1, \ldots, H \), associated with their probabilities \( \frac{1}{H}, k = 0, 1, \ldots, K \). Then (JRGP) is equivalent to

\[
\begin{align*}
\text{(JRGP}_3\text{D)} \quad \min_{r \in \mathbb{R}^M, \alpha \in [0, \infty), \alpha, x} & \quad \alpha \log \left( \frac{1}{H} \sum_{h=1}^{H} \exp \left\{ \left( \sum_{i=1}^{I_0} \tilde{c}_i^0(h) \exp \left\{ \sum_{j=1}^{M} a_{ij}^0 r_j \right\} \right) / \alpha \right\} \right) + \alpha \kappa_0, \\
\text{s.t.} & \quad \sum_{k=1}^{K} x_k \geq \log(1 - \epsilon'), \\
& \quad \frac{1}{H} \sum_{h=1}^{H} (1 - \varsigma_h^k) \geq e^{x_k}, \quad k = 1, 2, \ldots, K, \\
& \quad \sum_{i=1}^{I_k} \tilde{c}_i^k(h) \exp \left\{ \sum_{j=1}^{M} a_{ij}^k r_j \right\} \leq M \varsigma_h^k + 1, \quad h = 1, \ldots, H, \quad k = 1, 2, \ldots, K, \\
& \quad \varsigma_h^k \in \{0, 1\}, \quad h = 1, 2, \ldots, H.
\end{align*}
\]

*Proof.* Let \( w^k = \left[ \prod_{j=1}^{M} t_j^{a_{ij}^k}, \ldots, \prod_{j=1}^{M} t_j^{a_{ij}^k} \right], \quad k = 0, 1, \ldots, K \). By Proposition 2 and Proposition 3, (JRGP) can be reformulated as

\[
\begin{align*}
\min_{t \in \mathbb{R}^M_+} & \quad \inf_{\alpha \in [0, \infty]} \alpha \log \mathbb{E}_{F^0_k} \left[ \exp \left\{ \left( (c^0)^\top w^0 \right) / \alpha \right\} \right] + \alpha \kappa_0, \\
\text{s.t.} & \quad \mathbb{E}_{F^0_k} \left[ \prod_{k=1}^{K} \left( 1_{(-\infty,0]}(\tilde{c}^k(h)^\top w^k - 1) \right) \right] \geq 1 - \epsilon',
\end{align*}
\]

where \( \epsilon' \) is defined in Theorem 5.
Since the distributions $F^0_k, k = 1, \ldots, K$ are independent, and discretely distributed, we introduce a new auxiliary variable $\varsigma$ and use the Big-M method. Therefore, constraints (43)-(44) can be reformulated as

$$
\min_{t \in \mathbb{R}^M_{++}, \alpha \in [0, \infty), y, \varsigma} \alpha \log \left( \frac{1}{H} \sum_{h=1}^{H} \exp \left\{ \left( \tilde{c}^0(h)^\top w^0 \right) / \alpha \right\} \right) + \alpha \kappa_0,
$$

s.t.

$$
\prod_{k=1}^{K} y_k \geq 1 - \epsilon', \\
\frac{1}{H} \sum_{h=1}^{H} (1 - \varsigma^k_h) \geq y_k, \ k = 1, 2, \ldots, K, \\
\tilde{c}^k(h)^\top w^k \leq M \varsigma^k_h + 1, \ h = 1, 2, \ldots, H, \ k = 1, 2, \ldots, K, \\
\varsigma^k_h \in \{0, 1\}, \ h = 1, 2, \ldots, H.
$$

Combining $w^k = \left[ \prod_{j=1}^{M} t_j^{a_{kj}^h}, \ldots, \prod_{j=1}^{M} t_j^{a_{kj}^j} \right]$, $k = 0, 1, \ldots, K$ with $r_j = \log(t_j), \ j = 1, \ldots, M, \ x_k = \log(y_k), \ k = 1, \ldots, K$, gives the final reformulation.

Problem ($JRG P_{3D}$) is a mixed integer nonlinear programming problem which can be solved for instance by BONMIN or Pajarito softwares.

## 5 Uncertainty sets with known first order moment and nonnegative support

In the previous sections, we consider the full support for the random vector $c^k$ in the uncertainty sets. However, the vector $c^k$ should be nonnegative in geometric optimization (Boyd et al., 2007). In this section, we take this a step further by considering uncertainty sets with a nonnegative support and either the first-order moment or the first two order moments.
5.1 Individual chance constraints case

We consider the uncertainty sets $\mathcal{F}_k$, $k = 0, ..., K$, with known first order moment information and nonnegative support.

**Assumption 5.** The uncertainty sets are

$$\mathcal{F}_k = \{ F_k \mid \mathbb{E}_{F_k}[c^k] = \mu^k, \mathbb{P}[c^k \geq 0] = 1 \}, \; k = 0, 1, ..., K,$$

where $\mu^k > 0$, $k = 0, 1, ..., K$.

In the particular case of nonnegative support, if $\mu^k = 0$ then $c^k = 0$. Hence, we consider only the case where $\mu^k > 0$, $k = 0, 1, ..., K$ in the uncertainty set.

**Theorem 7.** Given Assumption 5, $(IRGP)$ is equivalent to the following geometric program:

$$(IRGP_4) \min_{t \in \mathbb{R}^{M^+}, \lambda, \alpha, \beta} \sum_{i=1}^{I_0} \mu_i^0 \prod_{j=1}^M t^{0}_{ij}$$

s.t. \[(1 - \epsilon_k)\lambda_k^{-1} - \lambda_k^{-1} \beta_k \mu^k \leq 1, \; k = 1, ..., K, \]

$$\beta_k < 0, 0 < \lambda_k \leq 1, \; k = 1, ..., K,$$

$$\lambda_k^{-1} \alpha_k \geq 1, \; k = 1, ..., K,$$

$$(-\beta_k)^{-1} \alpha_k \prod_{j=1}^M t_{ij}^{k} \leq 1, \; i = 1, \ldots, I_k, \; k = 1, ..., K.$$ 

**Proof.** We first denote

$$w^k = \left[ \prod_{j=1}^M a_{ij}^k, \ldots, \prod_{j=1}^M a_{ik_j}^k \right], \; k = 0, 1, \ldots, K.$$ 

Then, problem $(IRGP)$ can be reformulated as

$$\min_{t \in \mathbb{R}^{M^+}, F_0 \in \mathcal{F}_0} \sup_{F_k \in \mathcal{F}_k} \mathbb{E}_{F_0} \left[ (c^0)^\top w^0 \right]$$

s.t. \[\inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right) \geq 1 - \epsilon_k, \; k = 1, 2, ..., K. \] (46)

The reformulation of the objective function is obvious. The key point is the reformulation of the inner probability optimization problems:

$$\inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right), \; k = 1, 2, ..., K.$$ (47)
For each $k = 1, 2, \ldots, K$, given Assumption 5, the $k^{th}$ optimization problem in (47) can be reformulated as:

$$\inf_{P} \int_{c^{k} \geq 0} 1_{(c^{k})^\top w^{k} \leq 1} dP(c^{k}) \quad \text{(48)}$$

$$\text{s.t.} \quad \int_{c^{k} \geq 0} c^{k} dP(c^{k}) = \mu^{k}, \quad \int_{c^{k} \geq 0} dP(c^{k}) = 1, \quad \text{(49)}$$

where $1_{A}$ is the indicator function defined as $1_{A} = 1$ when $A$ holds and 0 otherwise.

From the Lagrange function

$$L(P; \beta, \lambda) = \int_{c^{k} \geq 0} 1_{(c^{k})^\top w^{k} \leq 1} dP(c^{k}) + \beta^\top \left( \mu^{k} - \int_{c^{k} \geq 0} c^{k} dP(c^{k}) \right) + \lambda \left( 1 - \int_{c^{k} \geq 0} dP(c^{k}) \right),$$

we can find the dual problem of (48)-(50),

$$\sup_{\beta \in \mathbb{R}^{d_k}, \lambda \in \mathbb{R}} \beta^\top \mu^{k} + \lambda \quad \text{(51)}$$

$$\text{s.t.} \quad 1_{(c^{k})^\top w^{k} \leq 1} - \beta^\top c^{k} - \lambda \geq 0, \quad \forall c^{k} \geq 0. \quad \text{(52)}$$

Constraint (52) can be reformatted as

$$\beta^\top c^{k} + \lambda \leq 1, \quad \forall c^{k} \geq 0, \quad \text{(53)}$$

$$\beta^\top c^{k} + \lambda \leq 0, \quad \forall c^{k} \geq 0 : \quad (c^{k})^\top w^{k} > 1. \quad \text{(54)}$$

Constraint (53) is equivalent to $\beta \leq 0$ and $\lambda \leq 1$. As we are seeking for the supremum in (51)-(52), we can replace the strict inequality in (54) by an inequality. Hence, (54) can be replaced by $f^{*} \leq 0$, where

$$f^{*} = \sup_{c^{k} \in \mathbb{R}^{d_k}} \beta^\top c^{k} + \lambda \quad \text{(55)}$$

$$\text{s.t.} \quad (c^{k})^\top w^{k} \geq 1, \quad c^{k} \geq 0. \quad \text{(56)}$$

The dual problem of (55)-(57) is

$$\inf_{\alpha \geq 0} \lambda - \alpha \quad \text{(58)}$$

$$\text{s.t.} \quad \beta + \alpha w^{k} \leq 0. \quad \text{(59)}$$
If we take the dual form (58)-(59) back to (51)-(52) and (53)-(54), we obtain a reformulation of the $k^{th}$ constraint in (46)

$$\begin{align*}
\beta^T \mu^k + \lambda &\geq 1 - \epsilon_k, \\
\beta &\leq 0, \lambda \leq 1, \\
\lambda &\leq \alpha, \\
\beta + \alpha w^k &\leq 0.
\end{align*}$$

(60) \hspace{2cm} (61) \hspace{2cm} (62) \hspace{2cm} (63)

As $\epsilon_k < 1$, $\beta \leq 0$ and $\mu^k > 0$, constraint (60) implies $\lambda > 0$, which further implies $\alpha \geq \lambda > 0$. As $w^k > 0$ (notice that $w^k = \prod_{j=1}^{M} t_j \cdot j$ and $t_j$ are strictly positive), Constraint (63) implies $\beta \leq -\alpha w^k < 0$.

Hence, (60)-(63) can be rewritten in the following form

$$\begin{align*}
(1 - \epsilon_k) \lambda^{-1} - \lambda^{-1} \beta^T \mu^k &\leq 1, \\
\beta &< 0, 0 < \lambda \leq 1, \\
\lambda^{-1} \alpha &\geq 1, \\
\beta_i + \alpha w^k_i &\leq 0, \ i = 1, \ldots, I_k.
\end{align*}$$

For each $k$, we have such a reformulation for the $k^{th}$ individual chance constraint.

If we take back $w^k = \left[ \prod_{j=1}^{M} t_j^{a_{ij}}, \ldots, \prod_{j=1}^{M} t_j^{a_{ik}} \right]$, $k = 1, \ldots, K$, we obtain the reformulation for $(IRGP)$.

$\square$

As $\lambda > 0$ and $\beta < 0$, we can reformulate $(IRGP_k)$ as a convex programming problem, using new auxiliary variables, $\tilde{\lambda} = \log(\lambda)$, $\tilde{\beta}^k_i = \log(-\beta^k_i)$, $\tilde{\alpha} = \log(\alpha)$, and $r_j = \log(t_j), j = 1, \ldots, M$. 

23
(JRGP$_{4r}$) $\min_{r \in \mathbb{R}^M, \tilde{\lambda}, \tilde{\alpha}, \tilde{\beta}} \quad \sum_{i=1}^{I_0} \mu_i^0 \exp \left\{ \sum_{j=1}^{M} a_{ij}^0 r_j \right\}$

s.t. $\quad (1 - \epsilon_k) e^{-\tilde{\lambda}_k} + \sum_{i=1}^{I_k} \exp \left\{ -\tilde{\lambda}_k - \tilde{\beta}_i^k + \log(\mu_i^k) \right\} \leq 1, \quad k = 1, \ldots, K,$

$\quad \tilde{\lambda}_k \leq 0, \quad k = 1, \ldots, K,$

$\quad \tilde{\lambda}_k \leq \tilde{\alpha}_k, \quad k = 1, \ldots, K,$

$\quad \tilde{\alpha}_k + \sum_{j=1}^{M} a_{ij}^k r_j - \tilde{\beta}_i^k \leq 0, \quad i = 1, \ldots, I_k, \quad k = 1, \ldots, K.$

5.2 Joint chance constraints case

We consider the uncertainty set with first order moment and nonnegative support for (JRGP). We assume that the marginal distributions in the uncertainty sets are pairwise independent.

Assumption 6. The uncertainty set $\mathcal{F} = \mathcal{F}_1 \times \cdots \times \mathcal{F}_K$, and for any distribution $F$ in $\mathcal{F}$, its marginal distributions $F_1, \ldots, F_K$ are pairwise independent. Moreover,

$\mathcal{F}_k = \{ F_k \mid \mathbb{E}_{F_k}[c^k] = \mu_k^k, \mathbb{P}[c^k \geq 0] = 1 \}, \quad k = 0, 1, \ldots, K,$

where $\mu_k^k > 0, \quad k = 0, 1, \ldots, K$.

Theorem 8. Given Assumption 6, (JRGP) is equivalent to the following geometric program:

(JRGP$_4$) $\min_{t \in \mathbb{R}^+_M, y \in \mathbb{R}^+_K, \alpha, \beta, \lambda} \quad \sum_{i=1}^{I_0} \mu_i^0 \prod_{j=1}^{M} \xi_j^0$

s.t. $\prod_{k=1}^{K} y_k \geq 1 - \epsilon, \quad 0 \leq y_k \leq 1, \quad k = 1, \ldots, K,$

$\quad y_k \lambda_k^{-1} - \lambda_k^{-1} \beta_k^\top \mu_k^k \leq 1, \quad k = 1, \ldots, K,$

$\quad \beta_k < 0, 0 < \lambda_k \leq 1, \quad k = 1, \ldots, K,$

$\quad \lambda_k^{-1} \alpha_k \geq 1, \quad k = 1, \ldots, K,$

$\quad (-\beta_i^k)^{-1} \alpha_k \prod_{j=1}^{M} \xi_j^k \leq 1, \quad i = 1, \ldots, I_k, \quad k = 1, \ldots, K.$

24
Proof. The proof of this theorem is trivial as the random variables in different rows are assumed to be independent. Hence, the joint chance constraint can be reformulated by a series of individual chance constraints and a monomial constraint. The rest of the proof follows from Theorem 7.

Problem \((JRGP_4)\) can be reformulated as a convex programming using additional auxiliary variables:

\[
\begin{align*}
(JRGP_{4r}) \quad \min_{r, x, \tilde{\lambda}, \tilde{\alpha}, \tilde{\beta}} & \quad \sum_{i=1}^{I_0} \mu_i^0 \exp \left\{ \sum_{j=1}^{M} a_{ij}^0 r_j \right\} \\
\text{s.t.} & \quad \sum_{k=1}^{K} x_k \geq \log(1 - \epsilon), \ x_k \leq 0, \ k = 1, \ldots, K, \\
& \quad \exp \left\{ x_k - \tilde{\lambda}_k \right\} + \sum_{i=1}^{I_k} \exp \left\{ -\tilde{\lambda}_k + \tilde{\beta}_i^k + \log(\mu_i^k) \right\} \leq 1, \ k = 1, \ldots, K, \\
& \quad \tilde{\lambda}_k \leq 0, \ k = 1, \ldots, K, \\
& \quad \tilde{\lambda}_k \leq \tilde{\alpha}_k, \ k = 1, \ldots, K, \\
& \quad \tilde{\alpha}_k + \sum_{j=1}^{M} a_{ij}^k r_j - \tilde{\beta}_i^k \leq 0, \ i = 1, \ldots, I_k, \ k = 1, \ldots, K.
\end{align*}
\]

It has \(K\) posynomial constraints, one posynomial objective function and some linear constraints.

6 Uncertainty sets with known first two order moments and nonnegative support

In this section, we consider an uncertainty set with first two order moment constraints and a nonnegative support constraint.

6.1 Individual chance constraints case

We consider the uncertainty sets \(\mathcal{F}_k, \ k = 0, \ldots, K\), with known first two order moment information and nonnegative support.
Assumption 7. The uncertainty sets are

\[ \mathcal{F}_k = \{ F_k \mid \mathbb{E}_{F_k}[c^k] = \mu^k, \, \text{Cov}_{F_k}[c^k] \preceq_S \Gamma^k, \, \mathbb{P}[c^k \geq 0] = 1 \}, \, k = 0, 1, \ldots, K. \]

We assume that \( \mu^k > 0 \) and \( \Gamma^k \) is a positive semidefinite matrix, \( k = 0, 1, \ldots, K \). \( A \succeq_S B \) means \( A - B \in \mathbb{S}_n^+ \), where \( \mathbb{S}_n^+ \) is the \( n \)-dimensional positive semidefinite cone.

Theorem 9. Given Assumption 7, (IRGP) is equivalent to the following optimization problem:

\[
\min_{t,Y,\beta,\epsilon,\lambda,w} \sum_{i=1}^{l_0} \mu_i^0 \prod_{j=1}^M t_{ij}^{0j} \quad (64)
\]

\[
\text{s.t.} \quad \prod_{j=1}^M t_{ij}^{0j} \leq w_k^i, \, i = 1, \ldots, I_k, \, k = 1, \ldots, K, \quad (65)
\]

\[
\left( -Y_k - \frac{1}{2} \beta_k, -\frac{1}{2} \beta_k^\top 1 - \lambda \right) \in \mathcal{COP}^n, \, k = 1, \ldots, K, \quad (66)
\]

\[
\left( -Y_k - \frac{1}{2} (\beta_k + \epsilon_k w^k), -\frac{1}{2} (\beta_k + \epsilon_k w^k)^\top \epsilon_k - \lambda_k \right) \in \mathcal{COP}^n, \, k = 1, \ldots, K \quad (67)
\]

\[
\beta_k^\top \mu^k + \lambda_k + < Y_k, \Gamma^k + \mu^k (\mu^k)^\top > \geq 1 - \epsilon_k, \, k = 1, \ldots, K, \quad (68)
\]

\[
-Y_k \succeq_S 0, \, k = 1, \ldots, K, \quad (69)
\]

\[
\epsilon_k \geq 0, \, k = 1, \ldots, K. \quad (70)
\]

\( \mathcal{COP}^n \) is the copositive cone, i.e., \( \mathcal{COP}^n = \{ A \in \mathbb{S}_n \mid x^\top Ax \geq 0, \, \forall x \geq 0 \} \).

Remark: The semidefinite constraint \( \text{Cov}_{F_k}[c^k] \preceq_S \Gamma^k \) in Assumption 7 can also be replaced by an equality constraint \( \text{Cov}_{F_k}[c^k] = \Gamma^k \). In this case, (IRGP) is not a reformulation of (IRGP). However, if we replace constraint (69) in (IRGP) by a copositive constraint \( -Y_k \in \mathcal{COP}^n \), (IRGP) provides a lower bound of (IRGP). The inequivalence comes from switching the min and the max operators in (86) and (87) stated hereafter.

Proof. We first denote

\[
w^k = \left[ \prod_{j=1}^M t_{1j}^{a_{1j}} \cdots \prod_{j=1}^M t_{kj}^{a_{kj}}, \ldots \right], \quad k = 0, 1, \ldots, K.
\]

Then, problem (IRGP) can be reformulated as

\[
\min_{t \in \mathbb{R}_+^M} \sup_{F_0 \in \mathcal{F}_0} \mathbb{E}_{F_0} \left[ (c^0)^\top w^0 \right] \quad (71)
\]

\[
\text{s.t.} \quad \inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right) \geq 1 - \epsilon_k, \, k = 1, 2, \ldots, K. \quad (72)
\]
The reformulation of the objective function is obvious. The key point is the reformulation of the inner probability optimization problem:

$$\inf_{F_k \in \mathcal{F}_k} \mathbb{P}_{F_k} \left( (c^k)^\top w^k \leq 1 \right) \geq 1 - \epsilon_k, \ k = 1, 2, ..., K. \quad (73)$$

Given Assumption 7, the $k^{th}$ optimization problem in (73), for $k = 1, 2, ..., K$, can be reformulated as

$$\inf_{F_k} \int_{c^k \geq 0} 1_{(c^k)^\top w^k \leq 1} dF_k(c^k) \quad (74)$$
$$\text{s.t.} \int_{c^k \geq 0} c^k dF_k(c^k) = \mu^k, \quad (75)$$
$$\int_{c^k \geq 0} c^k (c^k)^\top dF_k(c^k) \preceq S \Gamma^k + \mu^k (\mu^k)^\top, \quad (76)$$
$$\int_{c^k \geq 0} dF_k(c^k) = 1. \quad (77)$$

The Lagrange function can be written as

$$L(F_k; \beta, \lambda, Y) = \int_{c^k \geq 0} 1_{(c^k)^\top w^k \leq 1} dF_k(c^k) + \beta^\top \left( \mu^k - \int_{c^k \geq 0} c^k dF_k(c^k) \right) + \lambda \left( 1 - \int_{c^k \geq 0} dF_k(c^k) \right)$$
$$+ <Y, \Gamma^k + \mu^k (\mu^k)^\top - \int_{c^k \geq 0} c^k (c^k)^\top dF_k(c^k)>, \quad (78)$$

and the dual problem of (74)-(77) is

$$\sup_{\beta \in \mathbb{R}^I, \lambda \in \mathbb{R}, Y \in S^n} \beta^\top \mu^k + \lambda + <Y, \Gamma^k + \mu^k (\mu^k)^\top > \quad (79)$$
$$\text{s.t.} \quad 1_{(c^k)^\top w^k \leq 1} - \beta^\top c^k - \lambda - <Y, c^k (c^k)^\top > \geq 0, \ \forall c^k \geq 0, \quad (80)$$
$$-Y \succeq 0. \quad (81)$$

Constraint (79) can be reformulated as

$$<Y, c^k (c^k)^\top > + \beta^\top c^k + \lambda \leq 1, \ \forall c^k \geq 0, \quad (81)$$
$$<Y, c^k (c^k)^\top > + \beta^\top c^k + \lambda \leq 0, \ \forall c^k \geq 0 : \ (c^k)^\top w^k > 1. \quad (82)$$

We write the quadratic form in (81) in a matrix format,

$$\sup_{c^k \geq 0} \begin{pmatrix} c^k \\ 1 \end{pmatrix}^\top \begin{pmatrix} Y & \frac{1}{2} \beta \\ \frac{1}{2} \beta^\top & \lambda - 1 \end{pmatrix} \begin{pmatrix} c^k \\ 1 \end{pmatrix} \leq 0. \quad (83)$$
Let \( x = \left( \begin{array}{c} c^k \\ 1 \end{array} \right) / \left\| \left( \begin{array}{c} c^k \\ 1 \end{array} \right) \right\| \), constraint (83) can be rewritten as

\[
\inf_{x \geq 0, \|x\| = 1} x^T (-A) x \geq 0,
\]

where, \( A = \left( \begin{array}{cc} Y & \frac{1}{2} \beta \\ \frac{1}{2} \beta^T & \lambda - 1 \end{array} \right) \).

From Proposition 5.1 in Hiriart-Urruty and Seeger (2010), we know that,

\[
\inf_{x \geq 0, \|x\| = 1} x^T (-A) x \geq 0,
\]

\[
\Rightarrow \sup \{ z \in R : -A - z 1_n 1_n^T \in \mathcal{OP}^n \} \geq 0,
\]

\[
\Rightarrow -A \in \mathcal{OP}^n.
\]

Therefore, constraint (83) is equivalent to

\[
\left( \begin{array}{cc} -Y & -\frac{1}{2} \beta \\ -\frac{1}{2} \beta^T & 1 - \lambda \end{array} \right) \in \mathcal{OP}^n. \tag{84}
\]

For the second group of constraint (82), we first rewrite (82) as

\[
\sup_{c^k \geq 0, (c^k)^T w^k \geq 1} < Y, c^k (c^k)^T > + \beta^T c^k + \lambda \leq 0,
\]

which can be reformulated in the following form,

\[
\max_{c^k \geq 0, (c^k)^T w^k \geq 1} < Y, c^k (c^k)^T > + \beta^T c^k + \lambda \leq 0. \tag{85}
\]

Introducing a Lagrange multiplier \( \varepsilon \) for the constraint \((c^k)^T w^k \geq 1\), constraint (85) is equivalent to

\[
\max \min_{c^k \geq 0, \varepsilon \geq 0} < Y, c^k (c^k)^T > + \beta^T c^k + \lambda + \varepsilon ((c^k)^T w^k - 1) \leq 0, \tag{86}
\]

From (80), we know that \(-Y\) is positive semidefinite. Hence, the maxmin problem (86) is equivalent to

\[
\min \max_{\varepsilon \geq 0} < Y, c^k (c^k)^T > + \beta^T c^k + \lambda + \varepsilon ((c^k)^T w^k - 1) \leq 0, \tag{87}
\]

This constraint is equivalent to

\[ \exists \varepsilon \in [0, +\infty] \text{ s.t. } \max_{c^k \geq 0} < Y, c^k (c^k)^T > + \beta^T c^k + \lambda + \varepsilon ((c^k)^T w^k - 1) \leq 0, \]
which can be rewritten in the form of a matrix,

$$\max_{c_k \geq 0} \left( \begin{array}{c} c_k \\ 1 \end{array} \right)^T \left( \begin{array}{cc} Y & \frac{1}{2}(\beta + \varepsilon w^k) \\ \frac{1}{2}(\beta + \varepsilon w^k)^T & \lambda - \varepsilon \end{array} \right) \left( \begin{array}{c} c_k \\ 1 \end{array} \right) \leq 0. \quad (88)$$

Similarly to the reformulation of (83), and given Proposition 5.1 in Hiriart-Urruty and Seeger (2010), problem (88) is equivalent to

$$\left( \begin{array}{cc} -Y & -\frac{1}{2}(\beta + \varepsilon w^k) \\ -\frac{1}{2}(\beta + \varepsilon w^k)^T & \varepsilon - \lambda \end{array} \right) \in \mathcal{COP}^n. \quad (89)$$

So far, (81) is equivalent to (84) and (82) is equivalent to (89). Then, we replace the optimal value of (78) in (73). Together with the dual feasible constraint (80), we have the following reformulation of the distributionally robust individual chance constraint (73),

$$\beta^T \mu^k + \lambda + <Y, \Gamma^k + \mu^k(\mu^k)^T > \geq 1 - \varepsilon_k,$$

$$\left( \begin{array}{cc} -Y & -\frac{1}{2}(\beta + \varepsilon w^k) \\ -\frac{1}{2}(\beta + \varepsilon w^k)^T & \varepsilon - \lambda \end{array} \right) \in \mathcal{COP}^n,$$

$$\left( \begin{array}{cc} -Y & -\frac{1}{2}\beta \\ -\frac{1}{2}\beta^T & 1 - \lambda \end{array} \right) \in \mathcal{COP}^n,$$

$$-Y \succeq S 0,$$

$$\varepsilon \geq 0.$$

Furthermore, $A \in \mathcal{COP}^n$ and $B \succeq A$ (entrywise) imply $B \in \mathcal{COP}^n$. Let $w^k$ be the slack variables such that $\prod_{j=1}^M t_i^{b_j} \leq w^k_i$, $i = 1, \ldots, I_k$, $k = 1, \ldots, K$. Together with the aforementioned constraints, we obtain a reformulation of the distributionally robust individual chance constraint (73).

This reformulation is obtained for each $k$, the proof follows from substituting these constraints in (71)-(72).

$$\square$$

Although the copositive cone is convex, the bilinear term $\varepsilon_k w^k$ in (67) makes the constraint (67) nonconvex. Here, $\varepsilon_k$ is the dual variable for the constraints in (86). In order to get a tractable reformulation, we decompose the optimization problem into $2^K$ sub-problems corresponding to all possible cases where $\varepsilon_k = 0$ or $\varepsilon_k > 0$. We introduce
the index sets \( J_d, d = 1, \ldots, 2^K \), where \( J_d \) is a non-repeating subset of \( \{1, 2, \ldots, K\} \). We denote the complement set of \( J_d \) as \( J_d^C \). We let \( \varepsilon_k = 0 \) for \( k \in J_d \) and \( \varepsilon_k > 0 \) for \( k \in J_d^C \).

When \( \varepsilon_k = 0 \), the bilinear term \( \varepsilon_k w^k \) in (67) is equal to zero and thus (67) implies (66).

When \( \varepsilon_k > 0 \), we use the auxiliary variables, \( \bar{\varepsilon}_k = 1/\varepsilon_k, \bar{Y}_k = Y_k/\varepsilon_k, \bar{\beta}_k = \beta_k/\varepsilon_k, \) and \( \bar{\lambda}_k = \lambda_k/\varepsilon_k \), to put the bilinear term into a geometric form. For the purpose of the reformulation, we use the positive homogeneity of the copositive cone and the positive semidefinite cone. Then, we can get a reformulation of (66)-(70) under the case of \( \varepsilon_k > 0 \),

\[
\begin{align*}
\begin{pmatrix}
-\bar{Y}_k & -\frac{1}{2}\bar{\beta}_k \\
-\frac{1}{2}\bar{\beta}_k^T & -\bar{\lambda}_k
\end{pmatrix} & \in \mathcal{COP}^n, \\
\begin{pmatrix}
-\bar{Y}_k & -\frac{1}{2}(\bar{\beta}_k + w^k) \\
-\frac{1}{2}(\bar{\beta}_k + w^k)^T & 1 - \bar{\lambda}_k
\end{pmatrix} & \in \mathcal{COP}^n, \\
\bar{\beta}_k \mu^k + \bar{\lambda}_k + <\bar{Y}_k, \Gamma^k + \mu^k(\mu^k)^T > & \geq (1 - \varepsilon_k)\bar{\varepsilon}_k, \\
-\bar{Y}_k & \succeq 0, \\
\bar{\varepsilon}_k & \geq 0.
\end{align*}
\]

We then put the two cases together and denote the \( d \)th sub-optimization problem with respect to sub-index set \( J_d \) and its complement set \( J_d^C \), as

\[
v(d) = \min_{t,Y,\beta,\varepsilon,\lambda,w} \sum_{i=1}^{I_0} \mu_i^0 \prod_{j=1}^{M} t_j^{a_{ij}} \\
\text{s.t.} \quad \prod_{j=1}^{M} t_j^{a_{ij}} \leq w^i, \ i = 1, \ldots, I_k, \ k = 1, \ldots, K, \\
\begin{pmatrix}
-\bar{Y}_k & -\frac{1}{2}\bar{\beta}_k \\
-\frac{1}{2}\bar{\beta}_k^T & -\bar{\lambda}_k
\end{pmatrix} & \in \mathcal{COP}^n, \ k \in J_d, \\
\begin{pmatrix}
-\bar{Y}_k & -\frac{1}{2}\bar{\beta}_k \\
-\frac{1}{2}\bar{\beta}_k^T & -\bar{\lambda}_k
\end{pmatrix} & \in \mathcal{COP}^n, \ k \in J_d^C, \\
\begin{pmatrix}
-\bar{Y}_k & -\frac{1}{2}(\bar{\beta}_k + w^k) \\
-\frac{1}{2}(\bar{\beta}_k + w^k)^T & 1 - \bar{\lambda}_k
\end{pmatrix} & \in \mathcal{COP}^n, \ k \in J_d^C, \\
\bar{\beta}_k \mu^k + \bar{\lambda}_k + <\bar{Y}_k, \Gamma^k + \mu^k(\mu^k)^T > & \geq (1 - \varepsilon_k)\bar{\varepsilon}_k, \ k \in J_d^C, \\
-\bar{Y}_k & \succeq 0, \ k = 1, \ldots, K, \\
\varepsilon_k & \geq 0, \ k \in J_d^C.
\]

30
and the optimum value \((IRGP_3)\) is equal to \(\min_{d=1,\ldots,2^K} \{v(d)\}\). Here, we unify the notations where \(Y_k, \tilde{\beta}_k, \tilde{\lambda}_k, \tilde{\epsilon}_k, k \in J_d^C\), are denoted by \(Y_k, \beta_k, \lambda_k, \epsilon_k\).

We can rewrite \(v(d)\) as a convex programming problem by using the auxiliary variables \(r_j = \log(t_j), j = 1, \ldots, M\), where the monomials \(\prod_{j=1}^M t_j^{a_{ij}}\) are reformulated as exponential terms \(\exp\left\{\sum_{j=1}^M a_{ij} r_j\right\}\).

Although we can reformulate \(v(d)\) as a convex optimization problem, it is a NP-hard problem (Murty and Kabadi, 1987).

6.2 Joint chance constraints case

We consider an uncertainty set for \((JRGP)\), which characterizes the uncertainties in terms of the nonnegative support and of the first two order moments. In addition, we assume that the marginal distributions are pairwise independent.

**Assumption 8.** The uncertainty set \(\mathcal{F} = \mathcal{F}_1 \times \cdots \times \mathcal{F}_K\), and for any distribution \(F\) in \(\mathcal{F}\), its marginal distributions \(F_1, \ldots, F_K\) are pairwise independent. Moreover,

\[
\mathcal{F}_k = \{F_k \mid \mathbb{E}F_k[c^k] = \mu^k, \text{Cov}_{F_k}[c^k] \preceq \Gamma^k, \mathbb{P}[c^k \geq 0] = 1\}, \quad k = 0, 1, \ldots, K,
\]

where \(\mu^k > 0, k = 0, 1, \ldots, K\).

**Theorem 10.** Given Assumption 8, \((JRGP)\) is equivalent to the following geometric program:

\[
(JRGP_5) \quad \min_{t, y, \bar{Y}, \tilde{\beta}, \tilde{\epsilon}, \tilde{\lambda}, w} \sum_{i=1}^{I_0} \mu_i^0 \prod_{j=1}^M a_{ij}^0
\]

\[
s.t. \quad (66)-(67), (69)-(70),
\]

\[
\beta_k^T \mu^k + \lambda_k + Y_k, \Gamma^k + \mu^k (\mu^k) \geq y_k, \quad k = 1, \ldots, K,
\]

\[
\prod_{k=1}^K y_k \geq 1 - \epsilon, \quad 0 \leq y_k \leq 1, \quad k = 1, \ldots, K.
\]

**Proof.** The proof of this theorem is trivial as the row vector random variables are assumed to be independent. Hence, the joint chance constraint can be reformulated by a series of individual chance constraints and a monomial constraint. The details of the proof follows from Theorem 9.
(JRGP_5) is not a convex optimization problem due to the introduction of the monomial \( \prod_k y_k \), which lead the variable transformation technique fails to work in this case. One can use the sequence convex approximation algorithm to solve the nonconvex optimization problem or use SOCP approximations (Cheng and Lisser, 2012).

7 Uncertainty sets for both \( c^k \) and \( a^k_{i,j} \)

In the previous sections, we consider the parameters \( c^k \) as random variables and \( a^k_{i,j} \) as deterministic parameters. Although only the parameters \( c^k \) are random in a great majority of practical geometric optimization problems (Rao, 1996; Dušačová, 2009; Liu et al., 2016), we focus in this section on the case where both \( c^k \) and \( a^k_{i,j} \) are random. From a practical point of view, it is reasonable to assume that the random exponents \( a^k \) should be integer or at least be selected in finite support, and thus follows a discrete distribution with \( M_k \) scenarios. Moreover, we assume that for each realization of \( a^k \), a distribution or an uncertainty set of the distribution of \( c^k \) is given.

7.1 Individual chance constraints case

We focus on uncertainty sets \( \mathcal{G}_k, k = 0, 1, \ldots, K \) where both \( a^k \) and \( c^k, k = 0, 1, \ldots, K \) are random. The marginal distributions of \( c^k, a^k \), and the joint distributions of \( c^k \) and \( a^k \) are denoted by \( F_k, F^a_k \) and \( G_k, k = 0, 1, \ldots, K \), respectively.

Assumption 9. The uncertainty sets for the joint distribution of \( c^k \) and \( a^k \) are

\[
\mathcal{G}_k = \left\{ G_k \left| \begin{array}{l} G_k(a^k = a^k(m)) = p^k_m, m = 1, \ldots, M_k, \\
G_k(c|a^k = a^k(m)) = F^m_k(c), \forall c \in \text{supp}(F_k), \; F^m_k \in \mathcal{F}^m_k, m = 1, \ldots, M_k. \end{array} \right\} \right., k = 0, 1, \ldots, K.
\]

\( \mathcal{F}^m_k \) could be an uncertainty set defined in Assumptions 1, 3, 5, or 7, \( p^k_m \) is the appearing probability of \( m \)th scenario of \( a^k \) such that \( \sum_{m=1}^{M_k} p^k_m = 1 \).

Theorem 11. Given Assumption 9, (IRGP) is equivalent to the following chance con-
strained program:

\[(IRGP_a)\]

\[
\min_{t \in \mathbb{R}^{M \times I}} \sum_{m=1}^{M_0} p^0_m \left( \sup_{F^m_0 \in \mathcal{F}_0} \mathbb{E}_{F^m_0} \left[ \sum_{i=1}^{I_0} c^0_i \prod_{j=1}^{M} t^0_{ij}(m) \right] \right) 
\]

\[
\text{s.t.} \quad \sum_{m=1}^{M_k} p^k_m z^k_m \geq 1 - \epsilon_k, \quad k = 1, \ldots, K, \quad (91)
\]

\[
\inf_{F^m_k \in \mathcal{F}^m_k} \mathbb{P}_{F^m_k} \left( \sum_{i=1}^{I_k} c^k_i \prod_{j=1}^{M} t^k_{ij}(m) \leq 1 \right) \geq z^k_m, \quad m = 1, \ldots, M_k, k = 1, \ldots, K, \quad (92)
\]

\[
z^k_m \in [0, 1], \quad m = 1, \ldots, M_k, k = 1, \ldots, K. \quad (93)
\]

**Proof.** We can formulate the \(k\)th individual constraint of \(IRGP\) in the following form,

\[
\mathbb{P}_{G_k} \left( \sum_{i=1}^{I_k} c^k_i \prod_{j=1}^{M} t^k_{ij}(m) \leq 1 \right) \geq 1 - \epsilon_k, \quad \forall G_k \in \mathcal{G}_k, \quad (94)
\]

For a given \(G_k \in \mathcal{G}_k\), we can use the law of total probability to rewrite the left-hand side of (94) as the sum of the conditional probabilities with respect to \(F^m_k\) subject to \(a^k = a^k(m)\). Then, if \(F^m_k\) is any marginal distribution with respect to Assumption 9, we have the following reformulation of (94),

\[
\sum_{m=1}^{M_k} p^k_m \mathbb{P}_{F^m_k} \left( \sum_{i=1}^{I_k} c^k_i \prod_{j=1}^{M} t^k_{ij}(m) \leq 1 \right) \geq 1 - \epsilon_k, \quad \forall F^m_k \in \mathcal{F}^m_k, m = 1, \ldots, M_k. \quad (95)
\]

The introduction of the slack variables \(z^k_m \in [0, 1], m = 1, \ldots, M_k, k = 1, \ldots, K\), and the infimum operator in the summation, lead to the result of the theorem. The reformulation of the objective function can be obtained by using the law of total probability.

\[\square\]

Given the uncertainty set \(\mathcal{F}^m_k\) defined in Assumptions 1, 3, 5, 7, we can use Theorems 1, 3, 4, 7, 9 to reformulate constraint (92) as a deterministic constraint or a group of deterministic constraints.

For instance, if we consider the uncertainty set \(\mathcal{F}^m_k\) with known first two order moments, \(\mu^k_i(m), i = 1, \ldots, I_k\) and \(\{\sigma_{i,p}^k(m), i, p = 1, \ldots, I_k\}\), defined in Assumption 1 with
\[ \pi^k_1 = 0, \pi^k_2 = 1, k = 0, 1, \ldots, K, \] we can reformulate \((IRGP_a)\) as

\[
(\text{IRGP}_{a1}) \quad \min_{t \in \mathbb{R}^M_{+}, z} \quad \sum_{i=1}^{I_0} \mu^0_i \prod_{j=1}^{M} t^0_{ij}(m) \tag{95}
\]

subject to

\[
\sum_{m=1}^{M_k} p^k_m z^k_m \geq 1 - \epsilon_k, \quad k = 1, \ldots, K, \tag{96}
\]

\[
\sum_{i=1}^{I_k} \mu^k_i(m) \prod_{j=1}^{M} t^k_{ij}(m) + \sqrt{\frac{z^k_m}{1 - z^k_m}} \sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma^{k}_{i,p}(m) \prod_{j=1}^{M} t^k_{ij}(m) + a^k_p(m)} \leq 1, \quad m = 1, \ldots, M_k, k = 1, \ldots, K, \tag{97}
\]

\[
z^k_m \in [0, 1], \quad m = 1, \ldots, M_k, k = 1, \ldots, K. \tag{98}
\]

where \(\mu^0_i = \sum_{m=1}^{M_0} p^0_m \mu^0_i(m)\). Substituting \(r_j = \log(t_j), j = 1, \ldots, M\) in \((\text{IRGP}_{a1})\) leads to

\[
(\text{IRGP}_{a1r}) \quad \min_{t \in \mathbb{R}^M_{+}, z} \quad \sum_{i=1}^{I_0} \mu^0_i \exp \left\{ \sum_{j=1}^{M} a^0_{ij}(m)r_j \right\} \tag{99}
\]

subject to

\[
\sqrt{\sum_{i=1}^{I_k} \sum_{p=1}^{I_k} \sigma^{k}_{i,p}(m) \exp \left\{ \sum_{j=1}^{M} (a^k_{ij}(m) + a^k_{ip}(m))r_j + \log \left( \frac{z^k_m}{1 - z^k_m} \right) \right\}} \]

\[
+ \sum_{i=1}^{I_k} \mu^k_i(m) \exp \left\{ \sum_{j=1}^{M} a^k_{ij}(m)r_j \right\} \leq 1, \quad m = 1, \ldots, M_k, \quad k = 1, \ldots, K. \tag{100}
\]

\[(96),(98). \tag{101}\]

**Proposition 4.** If \(\epsilon_k \leq \frac{1}{2} \min_m \{p^k_m\}, \) and \(\sigma^k_{i,p} \geq 0, \) for any \(i, p = 1, \ldots, I_k, k = 1, \ldots, K,\) \((\text{IRGP}_{a1r})\) is a convex programming problem.

**Proof.** The only term that affects the convexity is \(\log \left( \frac{z^k_m}{1 - z^k_m} \right).\) For \(z^k_m \geq 0.5, \) \(\log \left( \frac{z^k_m}{1 - z^k_m} \right)\) is convex. Together with (96), \(\epsilon_k \leq \frac{\min_m p^k_m}{2} \) implies \(z^k_m \geq 0.5, \) \(\forall m = 1, \ldots, M_k.\) Hence, this condition guarantees the convexity.

\[\square\]
7.2 Uncertainties of $p^k_m$

In the case where the distribution of $a^k$ is not known, we consider the following uncertainty sets $G_k$, $k = 0, \ldots, K$.

**Assumption 10.** The uncertainty sets for the joint distribution of $c^k$ and $a^k$ are

$$G_k = \left\{ G_k(a^k = a^k(m)) = p^k_m, \ m = 1, \ldots, M_k, \ p^k \in \mathcal{P}_k, \ G_k(c|a^k = a^k(m)) = F^m_k(c), \forall c \in \text{supp}(F_k), \ F^m_k \in \mathcal{F}_k^m, \ m = 1, \ldots, M_k \right\}, \ k = 0, 1, \ldots, K.$$ 

where $\mathcal{P}_k$ is the uncertainty set of the distribution of $a^k$.

For example, we can consider the following two kinds of $\mathcal{P}_k$ (Zhu and Fukushima, 2009):

- box based uncertainty,

  $$\mathcal{P}_k = \{ p^k | p^k = \bar{p}^k + \eta_k, \ e^\top \eta_k = 0, \ \underline{\eta}_k \leq \eta_k \leq \bar{\eta}_k \},$$

- ellipsoidal based uncertainty,

  $$\mathcal{P}_k = \{ p^k | p^k = \bar{p}^k + A_k \eta_k, \ e^\top A_k \eta_k = 0, \ \bar{p}^k + A_k \eta_k \geq 0, \ ||\eta_k|| \leq 1 \}.$$ 

Given Assumption 10, we can get a similar reformulation of $(IRGP)$ as $(IRGP_a)$. However, constraint (96) should hold for all $p^k \in \mathcal{P}_k$.

Taking $\mathcal{P}_k$ as the box uncertainty or the ellipsoidal uncertainty, constraint (96) can be replaced by $\min_{p^k \in \mathcal{P}_k} (z^k)^\top p^k \geq 1 - \epsilon_k$. The latter semi-infinite optimization problem can be reformulated either by as a set of linear constraints or by a set of second order constraints depending whether the box based uncertainty or the ellipsoidal one is considered.

7.3 Joint chance constraints case

We focus on the uncertainty sets $G$ for $(JRG)$ where both $a$ and $c$ are random, and pairwise independent. The joint distributions of $c, a$ and the joint distributions of $c$ and $a$ are denoted by $F, F_a$ and $G$, respectively.

**Assumption 11.** The uncertainty set $G = G_1 \times \cdots \times G_K$ where $G_k$ is a marginal uncertainty set defined in Assumption 9. For any distribution $G$ in $G$, its marginal distributions $G_1, \ldots, G_K$ are pairwise independent.
Theorem 12. Given Assumption 11, (JRGP) is equivalent to the following program:

\[
(I_{RGP}) \quad \min_{t \in \mathbb{R}_{++}, y, z} \quad \sum_{m=1}^{M_0} P_m^0 \left( \sup_{F_m^0 \in \mathcal{F}_{m,0}^0} \mathbb{E}_{F_m^0} \left[ \sum_{i=1}^{I_0} C_m^0 \prod_{j=1}^{M} t_{a_{ij}^0}^{0} \right] \right) \tag{102}
\]

s.t. \[
\prod_{k=1}^{K} y_k \geq (1 - \epsilon), \quad 0 \leq y_k \leq 1, \quad k = 1, \ldots, K, \tag{103}
\]

\[
\sum_{m=1}^{M_k} P_m^k z_m^k \geq y_k, \quad k = 1, \ldots, K, \tag{104}
\]

Proof. We can reformulate the joint chance constraint as a series of individual chance constraint, this can be achieved through the introduction of the auxiliary variables \(y_k\). The proof follows from Theorems 2 and 11.

Notice that if we consider \(\mathcal{F}_k^m\) according to Assumptions 1, 3, 5, 7, we can come-up with deterministic equivalent reformulations of (JRGP) by the mean of Theorems 1, 3, 4, 7, 9, respectively.

As an example, we choose \(\mathcal{F}_k^m\) with respect to Assumption 1 with \(\pi_{1k} = 0, \pi_{2k} = 1, k = 0, 1, \ldots, K\). We can reformulate \((I_{RGP_a})\) as

\[
(I_{RGP_{a1}}) \quad \min_{t \in \mathbb{R}_{++}, y, z} \quad \sum_{i=1}^{I_0} \mu_i^0 \prod_{j=1}^{M} t_{a_{ij}^0}^{0} \tag{106}
\]

s.t. \[
(97), (98), (103), (104) \tag{107}
\]

Substituting \(r_j = \log(t_j), j = 1, \ldots, M\) in \((I_{RGP_{a1}})\) leads to

\[
(I_{RGP_{a1r}}) \quad \min_{r, x, z} \quad \sum_{i=1}^{I_0} \mu_i^0 \exp \left\{ \sum_{j=1}^{M} a_{ij}^{0} r_j \right\} \tag{108}
\]

s.t. \[
\sum_{k=1}^{K} x_k \geq \log(1 - \epsilon), \quad x_k \leq 0, \quad k = 1, \ldots, K, \tag{109}
\]

\[
\sum_{m=1}^{M_k} P_m^k z_m^k \geq e^{x_k}, \quad k = 1, \ldots, K, \tag{110}
\]

(98), (100).
Similarly to \((IRGP_{a1r})\), a sufficient condition for the convexity of \((JRGP_{a1r})\) is \(\epsilon \leq \frac{1}{2} \min_{k,m} \{p^k_m\}\) and \(\sigma^k_{i,p} \geq 0\), for any \(i, p = 1, \ldots, I_k, k = 1, \ldots, K\).

8 Numerical experiments

In this section, we consider a multi-dimensional distributionally robust shape optimization problem with individual chance constraints

\[
(RSOP_I) \quad \min_{x_1, \ldots, x_m} \prod_{i=1}^{m} x_i^{-1} \tag{108}
\]

\[
s.t. \quad \inf_{F_{wall} \in \mathcal{F}_{wall}} \mathbb{P}_{F_{wall}} \left[ \sum_{j=1}^{m-1} \frac{m-1}{A_j} x_1 \prod_{i=2, i \neq j}^{m} x_i \right] \leq \beta_{wall} \geq 1 - \epsilon_{wall} \tag{109}
\]

\[
\inf_{F_{flr} \in \mathcal{F}_{flr}} \mathbb{P}_{F_{flr}} \left[ \frac{1}{A_{flr}} \prod_{j=2}^{m} x_j \leq \beta_{flr} \right] \geq 1 - \epsilon_{flr} \tag{110}
\]

\[
x_i x_j^{-1} \leq \gamma_{i,j}, \quad \forall i \neq j \tag{111}
\]

and with a joint chance constraint

\[
(RSOP_J)
\]

\[
\min_{x_1, \ldots, x_m} \prod_{i=1}^{m} x_i^{-1} \tag{112}
\]

\[
s.t. \quad \inf_{F \in \mathcal{F}} \mathbb{P}_F \left[ \sum_{j=1}^{m-1} \frac{m-1}{A_j} x_1 \prod_{i=2, i \neq j}^{m} x_i \right] \leq \beta_{wall}, \quad \frac{1}{A_{flr}} \prod_{j=2}^{m} x_j \leq \beta_{flr} \geq 1 - \epsilon \tag{113}
\]

\[
x_i x_j^{-1} \leq \gamma_{i,j}, \quad \forall i \neq j \tag{114}
\]

which are extensions of the three-dimensional shape optimization problems introduced in Section 1.

In all our experiments, we assume that \(1/A_{flr}\) and \(1/A_j, j = 1, \ldots, m - 1\) are pairwise independent. Unless otherwise noted, we set \(m = 100, \gamma_{i,j} = 2, \forall i \neq j, \beta_{wall} = 2, \beta_{flr} = 2, \epsilon_{wall} = 0.02, \epsilon_{flr} = 0.0306\) and \(\epsilon = 0.05\). We set the mean value of \(1/A_{flr}\) to 0.01, the variance of \(1/A_{flr}\) to 0.01, the mean value of \(1/A_j\) to 0.01, \(j = 1, \ldots, m - 1\), the variance of \(1/A_j\) to 0.01, \(j = 1, \ldots, m - 1\), all the covariances between \(1/A_{flr}\) and \(1/A_j, j = 1, \ldots, m - 1\), are zero. Moreover, we set \(\pi^k_1 = 0.0001, \pi^k_2 = 1.2, k = 1, 2, \kappa_0 = \kappa = \kappa_1 = \kappa_2 = 0.02\). We use MOSEK and BONMIN solvers with CVX package to solve our different reformulation problems within Matlab R2012b, on a PC with a
Table 1: Optimal values of (IRGP) and (JRGP)

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>$\epsilon_{wall}$</th>
<th>$\epsilon_{flr}$</th>
<th>(IRGP$_2$)</th>
<th>(JRGP$_2$)</th>
<th>(ISGP)</th>
<th>(JSGP)</th>
<th>(IRGP$_{3N}$)</th>
<th>(JRGP$_{3N}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.045</td>
<td>0.0052</td>
<td>289.98</td>
<td>277.05</td>
<td>313.50</td>
<td>299.35</td>
<td>138.24</td>
<td>135.61</td>
</tr>
<tr>
<td>0.05</td>
<td>0.040</td>
<td>0.0104</td>
<td>305.27</td>
<td>277.05</td>
<td>320.71</td>
<td>299.35</td>
<td>141.10</td>
<td>135.61</td>
</tr>
<tr>
<td>0.05</td>
<td>0.035</td>
<td>0.0155</td>
<td>323.66</td>
<td>277.05</td>
<td>330.44</td>
<td>299.35</td>
<td>144.29</td>
<td>135.61</td>
</tr>
<tr>
<td>0.05</td>
<td>0.030</td>
<td>0.0206</td>
<td>346.47</td>
<td>277.05</td>
<td>350.44</td>
<td>299.35</td>
<td>147.87</td>
<td>135.61</td>
</tr>
<tr>
<td>0.05</td>
<td>0.025</td>
<td>0.0250</td>
<td>375.74</td>
<td>277.05</td>
<td>375.74</td>
<td>299.35</td>
<td>151.97</td>
<td>135.61</td>
</tr>
<tr>
<td>0.05</td>
<td>0.020</td>
<td>0.0300</td>
<td>415.34</td>
<td>277.05</td>
<td>407.52</td>
<td>299.35</td>
<td>156.85</td>
<td>135.61</td>
</tr>
<tr>
<td>0.05</td>
<td>0.015</td>
<td>0.0355</td>
<td>473.29</td>
<td>277.05</td>
<td>450.97</td>
<td>299.35</td>
<td>162.89</td>
<td>135.61</td>
</tr>
<tr>
<td>0.05</td>
<td>0.010</td>
<td>0.0404</td>
<td>570.38</td>
<td>277.05</td>
<td>514.50</td>
<td>299.35</td>
<td>171.02</td>
<td>135.61</td>
</tr>
<tr>
<td>0.05</td>
<td>0.005</td>
<td>0.0452</td>
<td>789.58</td>
<td>277.05</td>
<td>620.92</td>
<td>299.35</td>
<td>184.01</td>
<td>135.61</td>
</tr>
</tbody>
</table>

2.6 Ghz Intel Core i7-5600U CPU and 12.0 GB RAM. In Subsections 8.1 and 8.2, we compare optimal values of individual and joint chance constraints with respect to the first three groups of uncertainty sets. Subsection 8.3 shows the effect of ignoring nonnegative support constraint in the uncertainty set. Subsection 8.4 presents the effect of ignoring the randomness of $a_{i,j}$ in the uncertainty set.

8.1 Comparison between individual and joint chance constraints

We first compare the optimal values between the individual and joint distributionally robust chance constraints, namely (IRGP$_2$) and (JRGP$_2$) based on the first two order moments uncertainty set, (IRGP$_1$) and (JRGP$_1$) under the uncertainty set with unknown first two order moments, (IRGP$_{3N}$) and (JRGP$_{3N}$) based on normal distribution as a reference distribution. Moreover, we consider normally distributed stochastic shape optimization problems with individual and joint chance constraints denoted by (ISGP) and (JSGP), respectively (Liu et al., 2016). We use the stochastic solution methods to find the optimal solutions for (ISGP) and (JSGP) (Liu et al. (2016)).

We set nine groups of $\epsilon_{wall}$ and $\epsilon_{flr}$ such that $(1 - \epsilon_{wall})(1 - \epsilon_{flr}) = 1 - \epsilon$. Columns 1-3 in Table 1 show the values of $\epsilon$, $\epsilon_{wall}$ and $\epsilon_{flr}$, respectively.

Then, we solve (IRGP$_2$), (IRGP$_1$), (ISGP) and (IRGP$_{3N}$) with different values of $\epsilon_{wall}$ and $\epsilon_{flr}$. The optimal values are given in Columns 4,6,8, and 10 of Table 1, respectively. Columns 5, 7, 9, and 11 of Table 1 show the optimal values of (JRGP$_2$) and (JRGP$_1$), (JSGP) and (JRGP$_{3N}$) with $\epsilon = 0.05$, respectively. The CPU times for solving (JRGP$_2$), (JRGP$_1$), (JSGP) and (JRGP$_{3N}$) are within 120 seconds.

From Table 1, we can see that the optimal values of (IRGP$_2$), (IRGP$_1$) or (IRGP$_{3N}$) are always larger than the optimal value of (JRGP$_2$), (JRGP$_1$) or (JRGP$_{3N}$), respectively. This phenomenon means that the individual chance constrained problems are generally
more conservative than the joint chance constrained problems.

The optimal values of \((IRGP_1)\) or \((JRGP_1)\) are larger than or equal to the optimal values of \((IRGP_2)\) or \((JRGP_2)\), respectively. This is due to the fact that the uncertainty sets with unknown moments contain the uncertainty sets with known moments under the same reference values of mean vector and covariance. Hence, \((IRGP_1)\) and \((JRGP_1)\) are more conservative than \((IRGP_2)\) or \((JRGP_2)\).

Moreover, the optimal values of \((IRGP_3N)\) or \((JRGP_3N)\) are smaller than or equal to the optimal values of \((IRGP_2)\) and \((IRGP_1)\) or \((JRGP_2)\) and \((JRGP_1)\), respectively. This means that the uncertainty set with known or unknown moments information is wider in this example than the uncertainty set bounded by a distance to a reference normal distribution.

8.2 Satisfaction of the probability constraints

In this subsection, we compare the performances of the optimal values of the above eight optimization problems regarding the satisfaction of the probability constraints when there is a bias between the real distribution and the reference distribution or between the real moments and the reference moments.

We randomly generate 100 groups of simulated distributions for \(1/A_{flr}\) and \(1/A_j\), \(j = 1, \ldots, m-1\). We suppose that all the simulated distributions are normal distributions with different mean values and variances. In each group, the distributions of \(1/A_{flr}\) and \(1/A_j\), \(j = 1, \ldots, m-1\) are uncorrelated. The values of the means and the variances are random numbers generated by ‘rand’ Matlab function with parameter 0.01.

Then, we use the optimal solutions of \((IRGP_2)\), \((IRGP_1)\), \((ISGP)\) and \((IRGP_{3N})\) to compute the values of left-hand side of (109) and (110) with the 100 groups of simulated distributions. We denote the two terms in the left-hand side by \(P_{F_{\text{wall}}}\) and \(P_{F_{flr}}\) for short, which represent the satisfaction probabilities for the wall constraint and the floor constraint, respectively. Figure 1 shows the values of \(P = P_{F_{\text{wall}}} \times P_{F_{flr}}\) with different simulated distributions, as well as the percentage of simulated distributions, under which \(P_{F_{\text{wall}}} \geq 0.98\) and \(P_{F_{flr}} \geq 0.9694\).

We choose 8 groups out of the 100 groups, and show the values of \(P_{F_{\text{wall}}}\) and \(P_{F_{flr}}\) in Table 2. The first and second columns in Table 2 present the simulated distributions of \(1/A_{flr}\) and \(1/A_1\). Column 3 designed by “…” refers to the columns \(1/A_2\ldots 1/A_{100}\) implicitly. The other columns present the corresponding values of \(P_{F_{\text{wall}}}\) and \(P_{F_{flr}}\) with the optimal solutions of \((IRGP_2)\), \((IRGP_1)\), \((ISGP)\) and \((IRGP_{3N})\), respectively.
Figure 1: Values of $P$
In Figure 1 and Table 2, the optimal solutions of \((ISGP)\) and \((IRGP_{3N})\) fail to guarantee \(P_{F_{wall}} \geq 0.98\) and \(P_{F_{flr}} \geq 0.9694\) in most of the cases. This means that the optimal solutions of \((ISGP)\) and \((IRGP_{3N})\) are generally not feasible when the underlying distribution is not estimated accurately.

Meanwhile, the values of \(P_{F_{wall}}\) and \(P_{F_{flr}}\) obtained with the moments based distributionally robust models, \((IRGP_2)\) and \((IRGP_1)\), are always larger than or equal to 0.98 and 0.9694. This means that these approaches can efficiently guarantee the satisfaction of the chance constraints when the distributions are not precisely known in advance.

We also test the optimal solutions of the joint chance constrained problems, \((JRG_{P_2})\), \((JRG_{P_1})\), \((JSGP)\) and \((JRG_{P_{3N}})\). We use the 100 groups of simulated normal distributions to compute the values of the left-hand side of constraint (113) with the optimal solutions of \((JRG_{P_2})\), \((JRG_{P_1})\), \((JSGP)\) and \((JRG_{P_{3N}})\). We denote the values by \(\mathbb{P}\) for short. Figure 2 shows the values of \(\mathbb{P}\) of \((JSGP)\), \((JRG_{P_2})\), \((JRG_{P_1})\) and \((JRG_{P_{3N}})\) under different simulated distributions, as well as the percentage of simulated distributions, under which \(\mathbb{P} \geq 0.95\). Table 3 shows the results of 8 groups out of the 100 groups of simulated distributions.

Figure 2: Values of \(\mathbb{P}\)

From Figure 2 and Table 3, we can see that the results of joint chance constrained models are comparable with the results of individual chance constrained models. The values of \(\mathbb{P} = P_{F_{wall}} \times P_{F_{flr}}\) obtained with individual chance constrained problems are generally larger than the values of \(\mathbb{P}\) obtained with the corresponding joint chance constrained problems.
### Table 3: Values of $\mathbb{P}$ with $\epsilon = 0.05$

<table>
<thead>
<tr>
<th>Simulated distribution</th>
<th>Values of $\mathbb{P}$</th>
<th>(1/A_{\text{wall}})</th>
<th>(1/A_1)</th>
<th>(\cdots)</th>
<th>(JRGP(_2))</th>
<th>(JRGP(_1))</th>
<th>(JSGP)</th>
<th>(JRGP(_3N))</th>
</tr>
</thead>
<tbody>
<tr>
<td>N(0.0199, 0.0385)</td>
<td></td>
<td>0.9047</td>
<td>0.9383</td>
<td>0.4332</td>
<td>0.5321</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(0.0220, 0.0457)</td>
<td></td>
<td>0.9177</td>
<td>0.9475</td>
<td>0.4670</td>
<td>0.5655</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(0.0279, 0.0521)</td>
<td></td>
<td>0.9208</td>
<td>0.9499</td>
<td>0.4688</td>
<td>0.5684</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(0.0112, 0.0502)</td>
<td></td>
<td>0.9281</td>
<td>0.9549</td>
<td>0.4911</td>
<td>0.5900</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(0.0279, 0.0344)</td>
<td></td>
<td>0.9318</td>
<td>0.9579</td>
<td>0.4914</td>
<td>0.5920</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(0.0396, 0.0173)</td>
<td></td>
<td>0.9373</td>
<td>0.9622</td>
<td>0.4910</td>
<td>0.5945</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(0.0158, 0.0335)</td>
<td></td>
<td>0.9483</td>
<td>0.9691</td>
<td>0.5432</td>
<td>0.6422</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(0.0162, 0.0229)</td>
<td></td>
<td>0.9508</td>
<td>0.9709</td>
<td>0.5469</td>
<td>0.6466</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 8.3 Effect of ignoring nonnegative support restriction

In this subsection, we test the distributionally robust chance constrained optimization problem with nonnegative support constraint: \((IRGP_4)\) and \((JRGP_4)\). We assume that \(1/A_j\) and \(1/A_{flr}\) are positively distributed.

We get the optimal solution of \((IRGP_4)\) and \((JRGP_4)\) by solving the equivalent reformulations \((IRGP_{4r})\) and \((JRGP_{4r})\).

Letting $\epsilon_{\text{wall}}$ to range from 0.005 to 0.045, and $\epsilon_{flr} = 1 - (1 - \epsilon)/(1 - \epsilon_{\text{wall}})$, we test \((IRGP_4)\) nine times. The optimal values of the tested models are given in Table 4.

For the sake of comparability, we consider \((IRGP_4)\) and \((JRGP_4)\) without the nonnegative support constraint in the uncertainty set. Therefore, it is easy to see that a distributionally robust chance constraint $\inf_{\mathbb{P}} \mathbb{P}(x \leq \beta) \geq 1 - \epsilon$ with only the mean constraint, i.e., $\mathbb{E}[x] = \mu > 0$, is infeasible when $\epsilon$ is not equal to 1. It is easy to check such infeasibility by choosing a two pointed distribution such that $x = \beta + (\mu - \beta)/\delta$ with probability $\delta$ and $x = \beta$ with probability $1 - \delta$ ($\mu < \beta$). For this distribution with mean value $\mu$, $\mathbb{P}(x \leq \beta) = \delta$, where $\delta$ could be set close enough to 0.

From Table 4, we can find that the effect of considering nonnegative support in a distributionally robust chance constrained optimization problem is very important as it may directly affect the feasibility of the chance constraints. Additionally, it shows that ignoring the nonnegative support constraint in an uncertainty set could lead an excess in terms of conservatism. As for joint and individual problems, the former always gives a lower bound to the latter when $(1 - \epsilon_{\text{wall}})(1 - \epsilon_{flr}) = 1 - \epsilon$. 

82
Table 4: Optimal values of \((IRGP_4)\) and \((JRGP_4)\) with or without nonnegative support constraints

<table>
<thead>
<tr>
<th>(\epsilon)</th>
<th>0.05</th>
<th>0.05</th>
<th>0.05</th>
<th>0.05</th>
<th>0.05</th>
<th>0.05</th>
<th>0.05</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\epsilon_{wall})</td>
<td>0.045</td>
<td>0.04</td>
<td>0.035</td>
<td>0.03</td>
<td>0.025</td>
<td>0.02</td>
<td>0.015</td>
<td>0.01</td>
</tr>
<tr>
<td>(\epsilon_{flr})</td>
<td>0.0052</td>
<td>0.0104</td>
<td>0.0155</td>
<td>0.0206</td>
<td>0.0256</td>
<td>0.0306</td>
<td>0.0355</td>
<td>0.0404</td>
</tr>
</tbody>
</table>

With nonnegative support \((IRGP_4)\) \((JRGP_4)\)

| | 2337.4 | 2632.7 | 3012.9 | 3520.5 | 4232.4 | 5302.4 | 7090.5 | 10679.4 | 21508.8 |

Without nonnegative support

| individual* | Infeasible |
| joint* | Infeasible |

*: with only first order moment constraint

8.4 Effect of ignoring randomness of \(a_{i,j}^k\)

In this subsection, we test \((IRGP_{a1})\) and \((JRGP_{a1})\), the distributionally robust chance constrained optimization problem with random \(c^k\) as well as random \(a_{i,j}^k\).

We consider two variants of the distributionally robust individual and joint chance constrained shape optimization problem,

\[(RSOP^I_a)\]

\[
\begin{align*}
\min_{x_1, \ldots, x_m} & \quad \prod_{i=1}^m x_i^{-1} \\
\text{s.t.} & \quad \inf_{F_{A,a_{wall}} \in \mathcal{F}_{A,a_{wall}}} \mathbb{P} \left[ \sum_{j=1}^{m-1} \left( \frac{m-1}{A_j} x_1^{a_{wall}} \prod_{i=2, i \neq j}^m x_i^{a_{wall}} \right) \leq \beta_{wall} \right] \geq 1 - \epsilon_{wall} \\
& \quad \inf_{F_{A_{flr}, a_{flr}} \in \mathcal{F}_{A_{flr}, a_{flr}}} \mathbb{P} \left[ \frac{1}{A_{flr}} \prod_{j=2}^m x_j^{a_{flr}} \leq \beta_{flr} \right] \geq 1 - \epsilon_{flr} \\
& \quad x_i x_j^{-1} \leq \gamma_{i,j}, \quad \forall i \neq j,
\end{align*}
\]

\[(RSOP^a_j)\]

\[
\begin{align*}
\min_{x_1, \ldots, x_m} & \quad \prod_{i=1}^m x_i^{-1} \\
\text{s.t.} & \quad \inf_{F_{A,a} \in \mathcal{F}_{A,a}} \mathbb{P} \left[ \sum_{j=1}^{m-1} \left( \frac{m-1}{A_j} x_1^{a_{wall}} \prod_{i=2, i \neq j}^m x_i^{a_{wall}} \right) \leq \beta_{wall}, \quad \frac{1}{A_{flr}} \prod_{j=2}^m x_j^{a_{flr}} \leq \beta_{flr} \right] \geq 1 - \epsilon, \\
& \quad x_i x_j^{-1} \leq \gamma_{i,j}, \quad \forall i \neq j.
\end{align*}
\]
Table 5: Optimal values of \((IRGP_{a1})\) and \((JRGP_{a1})\)

<table>
<thead>
<tr>
<th>(\epsilon_{\text{wall}})</th>
<th>(\epsilon_{\text{flr}})</th>
<th>(\epsilon_{\text{wall}})</th>
<th>(\epsilon_{\text{flr}})</th>
<th>(\epsilon_{\text{wall}})</th>
<th>(\epsilon_{\text{flr}})</th>
<th>(\epsilon_{\text{wall}})</th>
<th>(\epsilon_{\text{flr}})</th>
<th>(\epsilon_{\text{wall}})</th>
<th>(\epsilon_{\text{flr}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.005</td>
<td>0.0104</td>
<td>0.0155</td>
<td>0.0206</td>
<td>0.0256</td>
<td>0.0306</td>
<td>0.0355</td>
<td>0.0404</td>
<td>0.0452</td>
<td></td>
</tr>
<tr>
<td>((IRGP_{a1}))</td>
<td>20513.7</td>
<td>25315.0</td>
<td>31586.8</td>
<td>39822.6</td>
<td>50692.7</td>
<td>65106.5</td>
<td>84670.4</td>
<td>120878.0</td>
<td>225584.0</td>
</tr>
<tr>
<td>((JRGP_{a1}))</td>
<td>16822.6</td>
<td>16822.6</td>
<td>16822.6</td>
<td>16822.6</td>
<td>16822.6</td>
<td>16822.6</td>
<td>16822.6</td>
<td>16822.6</td>
<td>16822.6</td>
</tr>
</tbody>
</table>

We suppose \(a_{\text{flr}}\) and \(a_{\text{wall}}\) follow a joint two-pointed distribution, i.e., \(a_{\text{flr}} = a_{\text{wall}} = 1\) with probability \(2/3\) and \(a_{\text{flr}} = a_{\text{wall}} = 1/2\) with probability \(1/3\). In this case, \((RSOP_{I}^{a})\) and \((RSOP_{J}^{a})\) correspond to \((IRGP_{a1})\) and \((JRGP_{a1})\) we investigated in Section 7. This means that, each monomial in \((RSOP_{I}^{a})\) and \((RSOP_{J}^{a})\) could be linear with respect to \(x_{i}\) with probability 66.6% or be a square root of \(x_{i}\) with probability 33.3%.

Letting the \(\epsilon_{\text{wall}}\) to range from 0.005 to 0.045, and \(\epsilon_{\text{flr}} = 1 - (1 - 0.05)/(1 - \epsilon_{\text{wall}})\), we test \((IRGP_{a1})\) nine times. The optimal values of the tested models are given in Table 5.

If we set \(a_{\text{flr}} = 1\) and \(a_{\text{wall}} = 1\), i.e., without randomness, \((RSOP_{I}^{a})\) and \((RSOP_{J}^{a})\) are reduced to \((RSOP_{I})\) and \((RSOP_{J})\), respectively. In addition, the corresponding distributionally robust optimization problems \((IRGP_{a1})\) and \((JRGP_{a1})\) are reduced to \((IRGP_{2})\) and \((JRGP_{2})\), respectively. The corresponding optimal values of \((IRGP_{2})\) and \((JRGP_{2})\) with the moments information are given in Table 1.

In order to show the feasibility of the optimal solution with respect to the biased different simulated distributions against the reference distribution, we use the 100 simulated distributions for \(1/A_{\text{flr}}\) and \(1/A_{j}, \ j = 1, \ldots, m - 1\), which are generated in Section 8.1.

We select a subset of instances for illustration purposes where \(\epsilon_{\text{wall}} = 0.02\), \(\epsilon_{\text{flr}} = 0.0306\), and use the optimal solutions of \((IRGP_{a1})\) and \((JRGP_{a1})\) together with two models \((IRGP_{2})\) and \((JRGP_{2})\) where \(a_{ij}^{k}\) is deterministic, to compute the values of the left-hand side of (116), (117) and (120). This computation is performed on the generated 100 groups of simulated distributions of \(1/A_{\text{flr}}\) and \(1/A_{j}, \ j = 1, \ldots, m - 1\), and the two-point distribution of \(a_{\text{flr}}\) and \(a_{\text{wall}}\).

We denote the three terms in the left-hand side by \(P_{F_{\text{wall}}}, P_{F_{\text{flr}}}\) and \(P_{\text{joint}}\) for short, which represent the feasibility for the wall constraint and the floor constraint, individually or jointly, respectively.

We draw the product probability \(P_{F_{\text{wall}}} \times P_{F_{\text{flr}}}\) for \((IRGP_{a1})\) and \((IRGP_{2})\) in Figure 3, and the joint probability \(P_{\text{joint}}\) for \((JRGP_{2})\) and \((JRGP_{a1})\) in Figure 4.

The y-axis values in Figure 3 show the probability in the simulated distributions such that \(P_{F_{\text{wall}}} \geq 0.98\) and \(P_{F_{\text{flr}}} \geq 0.9694\) whilst the y-axis values in Figure 4 show the
probability in simulated distributions such that $\mathbb{P}_{joint} \geq 0.95$.

From Table 1 and Table 5, we can see that the optimal values of distributionally robust optimization models considering randomness of $a_{ij}$ are larger than the optimal values of distributionally robust optimization models without considering the randomness of $a_{ij}$.

However, Figure 3 and Figure 4 show that, when $a_{ij}$ is random in the real environment, distributionally robust optimization models without considering the randomness of $a_{ij}$ can not guarantee the feasibility of the chance constraints. In contrary, $IRGP_{a1}$ and $JRGP_{a1}$ which take the randomness of $a_{ij}$ into consideration can efficiently reduce the infeasibility for the chance constraints.

**Conclusion**

In this paper, we discuss distributionally robust geometric programs with individual and joint chance constraints. We consider seven kinds of the state-of-the-art uncertainty sets either with $c$ as a random parameter and $a$ as a deterministic one, or with $c$ and $a$ as random parameters. We also considered the full support as well as the nonnegative support for the random parameters. We derive deterministic reformulations of distributionally robust geometric programs with individual or joint chance constraints. We perform intensive numerical experiments for all our models, and show the efficiency of our approaches.
Acknowledgement

This research was supported by Programme Cai Yuanpei under number 34593YE.

References


