A SURVEY ON PROBABILITY-CONSTRAINED OPTIMIZATION PROBLEMS

XIAODI BAI
School of Management, Fudan University
Shanghai 200433, China

XIAOJIN ZHENG
School of Economics and Management
Tongji University, Shanghai 200092, China

XIAOLING SUN
School of Management, Fudan University
Shanghai 200433, China

ABSTRACT. Probabilistically constrained optimization problems are an important class of stochastic programming problems with wide applications in finance, management and engineering planning. In this paper, we summarize some important solution methods including convex approximation, DC approach, scenario approach and integer programming approach. We also discuss some future research perspectives on the probabilistically constrained optimization problems.

1. Introduction. A general optimization problem can be expressed as

\[
(P) \quad \min f(x) \\
\text{s.t. } c_i(x, \xi) \leq 0, \quad i = 1, \ldots, m, \\
x \in X,
\]

where \( f : \mathbb{R}^n \to \mathbb{R} \) and \( c_i(\cdot, \xi) : \mathbb{R}^n \to \mathbb{R} \) (\( i = 1, \ldots, m \)) are real-valued functions, \( \xi \in \mathbb{R}^s \) is a parameter vector and \( X \) is a convex closed subset of \( \mathbb{R}^n \). When \( \xi \) is fixed coefficient vector and \( f \) and \( c_i \) are convex functions, problem \( (P) \) is a conventional deterministic convex optimization problem which can be solved by many efficient algorithms (see, e.g., [8, 27]). However, in many applications of optimization problems in finance, engineering and management, the parameter vector \( \xi \) is often uncertain. For example, \( \xi \) may represent the returns of stocks, the market demands or agricultural production which can be regarded as random variables with full or partial information about the distributions.

A simple way to handle the parameter uncertainty in \( (P) \) is to ignore the uncertainty of \( \xi \) by replacing it with an estimation or a guess of \( \xi \). For instance, we may use the mean value of \( \xi \), which is an unbiased estimation of \( \xi \), to replace \( \xi \) in \( (P) \).

2000 Mathematics Subject Classification. Primary: 91C15; Secondary: 90C26.

Key words and phrases. Stochastic programming, probabilistic constraints, convex approximation method, DC approach, integer programming approach.

This research was supported by National Natural Science Foundation of China under grants 10971034 and 11101092, and by the Joint NSFC/RGC grants under grant 71061160506.
and solve the resulting problem as a deterministic optimization problem. This approach, however, may lead to an infeasible solution to problem \((P)\) and we are not able to control the probability of this infeasibility. Another approach is to require the constraints \(c_i(x, \xi) \leq 0\) \((i = 1, \ldots, m)\) to be satisfied for all possible values of \(\xi\), so that the optimal solution (if any) of the resulting problem is always feasible to the original problem \((P)\). However, requiring the constraints to be satisfied for all possible values of \(\xi\) is often infeasible or very expensive and the optimal solution, if exists, is extremely conservative and thus often useless in practical situations. When the parameters are known within some sets such as bounds or ellipsoids, the problem \((P)\) can be tackled by robust optimization (see [4]).

Probabilistically or chance constrained programming was first proposed by Charnes and Cooper [10] to deal with the parameter uncertainty in optimization problems. The idea is to guarantee the satisfaction of the constraints with a given probability, i.e., \(c_i(x, \xi) \leq 0\) is replaced by probability constraint \(\mathbb{P}\{c(x, \xi) \leq 0\} \geq 1 - \alpha\), where \(c(x, \xi) = (c_1(x, \xi), \ldots, c_m(x, \xi))^T\) and \(0 < \alpha < 1\) is the risk level. The probability constrained problem has the following form:

\[
\begin{align*}
(PCP) \quad & \min f(x) \\
& \text{s.t. } \mathbb{P}\{c(x, \xi) \leq 0\} \geq 1 - \alpha, \\
& x \in X.
\end{align*}
\]

1.1. Two Examples. We now give two examples to show the applications of problem \((PCP)\).

**Example 1.** (Portfolio selection) Let \(\xi = (\xi_1, \ldots, \xi_n)^T\) be the random returns of \(n\) risky assets with mean value vector \(\mu\) and covariance matrix \(\Sigma\). One of the classical mean-variance portfolio selection models of Markowitz is to minimize the variance while requiring that the mean value of the portfolio is equal to or higher than a prescribed return level \(R\). A more reasonable constraint for controlling the random return is to use the following probabilistic constraint:

\[
\mathbb{P}(\xi^T x \geq R) \geq 1 - \alpha,
\]

where \(\alpha \in (0, 1)\) is the risk level. This constraint is also called Value-at-Risk (VaR) constraint in financial engineering. The VaR-constrained portfolio selection model can be expressed as:

\[
\begin{align*}
& \min x^T \Sigma x \\
& \text{s.t. } \mathbb{P}(\xi^T x \geq R) \geq 1 - \alpha, \\
& x \in X,
\end{align*}
\]

where \(X\) is a group of deterministic linear constraints. VaR-based portfolio selection models have drawn much attention in the last decade in financial engineering (see [1, 2, 5, 7, 15]).

**Example 2.** (Transportation problem) Suppose that a product has \(n\) suppliers and \(m\) major clients. The capacity of supplier \(i\) is \(M_i\), \(i = 1, \ldots, n\), the unit shipping costs from supplier \(i\) to client \(j\) is \(c_{ij}\). Suppose that the demand vector
\( \xi = (\xi_1, \ldots, \xi_m)^T \) is random. The transportation problem can be modeled as

\[
\min \sum_{i=1}^n \sum_{j=1}^m c_{ij} x_{ij}
\]

s. t. \( \mathbb{P} \left\{ \sum_{i=1}^n x_{ij} \geq \xi_j, \ j = 1, \ldots, m \right\} \geq 1 - \alpha, \)

\[
\sum_{j=1}^m x_{ij} \leq M_i, \ i = 1, \ldots, n,
\]

\[
x_{ij} \geq 0, \ i = 1, \ldots, n, \ j = 1, \ldots, m,
\]

where \( x_{ij} \geq 0 \) is the amount of product shipped from supplier \( i \) to client \( j \). This is a stochastic version of the classical transportation problem (see [12, 14, 24, 30]). Similar probabilistically constrained linear programming has been used in inventory/production management in multi-period supply chain management (see [22]).

1.2. Two convex cases. We now discuss two special cases where problem \((PCP)\) can be converted into a convex program. For simplicity, we suppose that \( m = 1 \) and \( c(x, \xi) = c_1(x, \xi) \).

We first consider the log-concave distribution case where

- \( \xi \) is a continuous random vector with log-concave density function \( p(\xi) \);
- The set \( W = \{(x, \xi) \mid c(x, \xi) \leq 0\} \) is convex in \((x, \xi)\) space.

Let \( 1_W \) denote the indicator function of set \( W \). Then \( \log(1_W) = 0 \) if \((x, \xi) \in W\) and \(-\infty\) otherwise. Thus \( 1_W \) is a log-concave function of \((x, \xi)\). It follows that

\[
\mathbb{P}\{c(x, \xi) \leq 0\} = \int 1_W \cdot p(\xi) d\xi
\]

is also log-concave because \( \log(1_W \cdot p(\xi)) = \log(1_W) + \log(p(\xi)) \) is a concave function of \((x, \xi)\). Now, the probabilistic constraint \( \mathbb{P}\{c(x, \xi) \leq 0\} \geq 1-\alpha \) can be equivalently expressed as the following convex constraint:

\[
\log[\mathbb{P}\{c(x, \xi) \leq 0\}] \geq \log(1-\alpha).
\]

Thus, the feasible set of \((PCP)\) is convex.

Another convex case of \((PCP)\) is linear programming with normally distributed parameters. Suppose \( c(x, \xi) = \xi^T x - b \) where \( \xi \sim N(\bar{a}, \Sigma) \). Then the probabilistic constraint \( \mathbb{P}(\xi^T x \leq b) \geq 1 - \alpha \) can be written as

\[
\mathbb{P}\left( \frac{\xi^T x - \bar{a}^T x}{\sqrt{x^T \Sigma x}} \leq \frac{b - \bar{a}^T x}{\sqrt{x^T \Sigma x}} \right) \geq 1 - \alpha.
\]

Since \( (\xi^T x - \bar{a}^T x)/\sqrt{x^T \Sigma x} \sim N(0, 1) \), the probability in the left-hand side of the above inequality is simply \( \Phi((b - \bar{a}^T x)/\sqrt{x^T \Sigma x}) \), where

\[
\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt
\]

is the cumulative distribution function of a standard Gaussian random variable. Thus, the probability constraint \( \mathbb{P}(\xi^T x - b \leq 0) \geq 1 - \alpha \) is equivalent to

\[
\frac{b - \bar{a}^T x}{\sqrt{x^T \Sigma x}} \geq \Phi^{-1}(1 - \alpha)
\]

or

\[
\bar{a}^T x + \Phi^{-1}(1 - \alpha) \|\Sigma^{1/2} x\|_2 \leq b,
\]
which is a second-order cone constraint when $\alpha \in (0, 0.5)$. Note that $\Phi^{-1}(1-\alpha) > 0$ when $\alpha \in (0, 0.5)$.

1.3. **Major difficulties.** The major difficulties of the probabilistically constrained problem are due to the nonconvexity of the feasible set. Except for some special cases such as the two convex cases discussed in the previous subsection, the constraints $\mathbb{P}\{c(x, \xi) \leq 0\} \geq 1 - \alpha$ are in general nonconvex (see [16, 17, 20, 28]).

**Example 3.** Consider the following 2-dimensional example:

$$
\begin{align*}
\min & \quad f(x) := 2x_1^2 + 3x_2^2 + 2x_1x_2 \\
\text{s. t.} & \quad \mathbb{P}\{\xi^T x \geq 0.4\} \geq 0.6, \\
& \quad x \in X = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \leq 1, \ x_1^2 + 2x_2^2 \leq 1, \ 0 \leq x_1, \ x_2 \leq 1\},
\end{align*}
$$

where $\xi \in \mathbb{R}^2$ is a discrete random variable whose distribution is given in Table 1. Figure 1 shows the feasible set $S$ from which we can see that $S$ is nonconvex.

**Table 1.** Distribution of $\xi$

| scenario | 1 | 2 | 3 | 4 | 5 | 6 |
|----------|---|---|---|---|---|---|
| $\xi_1^1$ | 0 | 0.5 | 0.6 | 0.8 | 1 | 2 |
| $\xi_2^1$ | 1 | 2 | 0.7 | 0.5 | 0 | 0.4 |
| $p_i$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |

![Figure 1. Feasible set of Example 3.](image)

In this paper, we will review some important approaches and recent progress on probabilistically constrained optimization problems. In particular, we will investigate convex approximation methods, DC approach, integer programming methods and scenario approximation. We will also discuss research perspectives and challenging problems in this field.
2. **Convex Approximation Approaches.** The idea of convex approximation is to construct convex inner (conservative) approximation to the nonconvex feasible set of \((PCP)\). Nemirovski and Shapiro [25] gave a unified method to derive convex approximations to \((PCP)\). For simplicity, we consider the case of \(m = 1\), i.e., \(c(x, \xi) = c_1(x, \xi)\) is a real-valued function of \(x\). Also, we assume that \(c(x, \xi)\) is a convex function of \(x\) for any \(\xi\).

Note that \(\mathbb{P}\{c(x, \xi) \leq 0\} \geq 1 - \alpha\) is equivalent to \(\mathbb{P}\{c(x, \xi) > 0\} \leq \alpha\). We consider in the following how to construct a convex function \(\phi(x)\) such that \(\mathbb{P}\{c(x, \xi) > 0\} \leq \phi(x)\) for all \((x, \xi)\) and hence \(\phi(x) \leq \alpha\) is a convex (inner) approximation of the probabilistic constraint.

Let \(\psi : \mathbb{R} \to \mathbb{R}\) be a nonnegative convex and nondecreasing function satisfying the following property:

\[
\psi(z) \geq \psi(0) = 1, \quad \forall z > 0.
\]

Then, for any \(t > 0\) and random variable \(Z\), we have

\[
\mathbb{E}[\psi(tZ)] \geq \mathbb{E}[1_{(0, +\infty)}(tZ)] = \mathbb{P}(tZ > 0) = \mathbb{P}(Z > 0).
\]

Letting \(Z = c(x, \xi)\) and replacing \(t\) by \(t^{-1}\) in the above equation, we obtain

\[
\mathbb{P}\{c(x, \xi) > 0\} \leq \mathbb{E}[\psi(t^{-1}c(x, \xi))], \quad \forall t > 0. \tag{1}
\]

Define \(\Psi(x, t) := t\mathbb{E}[\psi(t^{-1}c(x, \xi))]\). Note that \(t\psi(t^{-1}u)\) is the perspective function of \(\psi\) which is convex in \((u, t)\) for \(t > 0\) and nondecreasing in \(u\). So the composite function \(\Psi(x, t)\) is convex in \((x, t)\) for \(t > 0\). It follows from (1) that \(\Psi(x, t) \leq t\alpha\) implies that \(\mathbb{P}\{c(x, \xi) > 0\} \leq \alpha\). Therefore,

\[
\Psi(x, t) \leq t\alpha, \tag{2}
\]

is a convex conservative approximation for any \(t > 0\).

Furthermore, we can take minimum for \(t > 0\) in (2) to obtain the tightest convex approximation:

\[
\inf_{t > 0} \{\Psi(x, t) - t\alpha\} \leq 0 \tag{3}
\]

Note that the left-hand side is still a convex function of \(x\) since \(\Psi(x, t)\) is convex in \((t, x)\) for \(t > 0\). Accordingly, the convex approximation to problem \((PCP)\) is

\[
\min f(x) \\
\text{s.t. } \inf_{t > 0} \{\Psi(x, t) - t\alpha\} \leq 0, \\
x \in X.
\]

We now discuss some special cases of (2) and (3) when different \(\psi(z)\) is used.

2.1. **CVaR approximation.** Let \(\psi(z) = [1 + z]_+\), where \([a]_+ = \max\{0, a\}\). The convex approximation (3) has the following form

\[
\inf_{t > 0} \{\mathbb{E}[t + c(x, \xi)]_+ - t\alpha\} \leq 0, \tag{4}
\]

where \(t > 0\) can be replaced by \(t \in \mathbb{R}\).

Define Conditional Value-at-Risk (CVaR) of a random variable \(z\) (see [29]):

\[
CVaR_{1-\alpha}(z) := \inf_{t \in \mathbb{R}} \left[ t + \frac{1}{\alpha} \mathbb{E}[z - t]_+ \right].
\]

It is easy to see that (4) is equivalent to \(CVaR_{1-\alpha}(c(x, \xi)) \leq 0\). Thus (4) is also called CVaR approximation. It can be proved that CVaR is the best convex approximation to the probabilistic constraint (see [11]).
Assume that $\xi$ has finite discrete distribution, i.e., $\xi$ takes values $\xi^i, i = 1, \ldots, N$, with equal probability. The CVaR approximation to the problem (PCP) can be written as the following convex program:

\[
(CVaR) \quad \min f(x) \\
\text{s.t. } t + \frac{1}{\alpha N} \sum_{i=1}^{N} y_i \leq 0, \\
y_i \geq 0, \quad y_i \geq c_j(x, \xi^i) - t, \quad i = 1, \ldots, N, \quad j = 1, \ldots, m, \\
x \in X, \quad t \in \mathbb{R}.
\]

2.2. Chebyshev approximation. Letting $\psi(z) = [(1 + z)_+]^2$ in (1), we obtain the following Chebyshev bound:

\[
P\{c(x, \xi) > 0\} \leq \mathbb{E}[(c(x, \xi)/t + 1)_+]^2, \quad \forall t > 0.
\]

The corresponding convex approximation to the probabilistic constraint is

\[
\mathbb{E}[(c(x, \xi)/t + 1)_+]^2 \leq t\alpha, \quad \forall t > 0,
\]

which can be written as

\[
\mathbb{E}[(c(x, \xi) + t)_+]^2/t \leq t\alpha, \quad \forall t > 0. \tag{5}
\]

Note that $\mathbb{E}(u^2_i) \leq \mathbb{E}(u^2)$. Replacing $(c(x, \xi) + t)_+$ by $c(x, \xi) + t$ in (5), we obtain a more conservative approximation:

\[
\mathbb{E}[(c(x, \xi)/t + 1)^2] \leq t\alpha, \quad \forall t > 0,
\]

which can be written as

\[
2\mathbb{E}[c(x, \xi)] + (1/t)\mathbb{E}[c(x, \xi)^2] + t(1 - \alpha) \leq 0, \quad \forall t > 0.
\]

Minimizing the left-hand side over all $t \in \mathbb{R}$ gives $t^* = \{\mathbb{E}[c(x, \xi)^2]/(1 - \alpha)\}^{\frac{1}{2}}$. This $t^*$ gives the tightest approximation:

\[
\mathbb{E}[c(x, \xi)] + \{(1 - \alpha)\mathbb{E}[c(x, \xi)^2]\}^{\frac{1}{2}} \leq 0.
\]

The above approximation only depends on the first and the second moments of $c(x, \xi)$.

3. DC Approach. Hong et al. [18] proposed a novel approach to obtain tighter approximation to the probabilistic function $\mathbb{P}(c(x, \xi) > 0)$ using DC (difference of two convex) functions. Successive convex approximation methods can be then employed to solve the resulting DC optimization problem.

Recall that $\mathbb{P}(z > 0) = \mathbb{E}[1_{(0, +\infty)}(z)]$. We thus focus on constructing DC approximation to the indicator function $1_{(0, +\infty)}(z)$. Let

\[
\psi(z, t) = t^{-1}[z + t]_+, \quad \phi(z, t) = t^{-1}[z]_+, \quad t > 0.
\]

Then, the DC function $\pi(z, t) := \psi(z, t) - \phi(z, t)$ is a DC approximation to the indicator function $1_{(0, +\infty)}(z)$, which is tighter than CVaR approximation function $\psi(z, t)$. See Figs. 2-5.

Now, let $g_1(x, t) = \mathbb{E}[t + c(x, \xi)]_+$, $g_2(x) = g_1(x, 0) = \mathbb{E}[c(x, \xi)]_+$. Let

\[
\tilde{p}(x) = \inf_{t > 0} \{t^{-1}[g_1(x, t) - g_2(x)]\}.
\]
Under some mild conditions, we can prove that (PCP) is actually equivalent to the following problem (see [18]):

$$\min f(x)$$

s.t. $\tilde{p}(x) \leq \alpha$,

$$x \in X.$$ 

Taking a small $t = \varepsilon > 0$, the above problem can be approximated by the following DC optimization problem:

$$(DC) \min f(x)$$

s.t. $g_1(x, \varepsilon) - g_2(x) \leq \varepsilon \alpha$,

$$x \in X.$$ 

Combining with Monte Carlo method to estimate the expectation and gradient, a sequential convex method was proposed in [18] to solve the DC optimization problem.

4. Integer Programming Methods. In this section, we assume that the random vector $\xi$ has finite discrete distribution, i.e., $P(\xi = \xi^j) = p_j$, $j = 1, \ldots, N$, $\sum_{j=1}^{N} p_j = 1$. Let $u^b_j$ be an upper bound vector of $c(x, \xi^j)$ over $X$, $j = 1, \ldots, N$. 
Then, the problem (PCP) can be reformulated as the following mixed-integer 0-1 programming (see [30]):

\[
\begin{align*}
\text{(MIP)} \quad & \min f(x) \\
\text{s.t.} \quad & c(x, \xi^j) \leq z_j u_j^b, \quad j = 1, \ldots, N, \\
& \sum_{j=1}^N p_j z_j \leq \alpha, \\
& x \in X, \quad z \in \{0, 1\}^N.
\end{align*}
\]

When \( f \) and \( c_i(\cdot, \xi^i) \) are convex functions, and \( X \) is a convex set, \( \text{(MIP)} \) is a convex mixed-integer 0-1 programming problem which can be solved by continuous relaxation based branch-and-bound or branch-and-cut methods. For instance, when \( f(x) \) is a linear function or quadratic function of \( x \), \( c_i(x, \xi) \) is linear in \( x \) for each \( i \), and \( X \) is defined by linear constraints, we can use commercial software such as CPLEX to solve \( \text{(MIP)} \). However, due to the poor quality of the continuous relaxation of \( \text{(MIP)} \) and its subproblems during the branch-and-bound process, a direct application of the branch-and-bound method to \( \text{(MIP)} \) is often inefficient and can only deal with problems with small size.

Zheng et al. [32] considered a quadratic case of (PCP) where \( f(x) = x^T Q x + c^T x, \quad Q \succeq 0, \quad c(x, \xi) = R - \xi^T B x, \) where \( R \in \mathbb{R}, \xi \in \mathbb{R}^m, B \in \mathbb{R}^{m \times n} \). Suppose that \( \xi \) takes values \( \xi_i, i = 1, \ldots, N, \) with equal probability. Let \( l_i^b \) and \( u_i^b \) be the lower bound and upper bound of \( \xi_i^T B x \) over \( X, \) \( i = 1, \ldots, N. \) Then, (PCP) can be written as a mixed-integer 0-1 convex quadratic programming problem:

\[
\begin{align*}
\text{(MIQP)} \quad & \min x^T Q x + c^T x \\
\text{s.t.} \quad & (\xi_i^T B x) \geq R + y_i (l_i^b - R), \quad i = 1, \ldots, N, \\
& e^T y \leq K, \quad y \in \{0, 1\}^N, \\
& x \in X,
\end{align*}
\]

where \( K = \lfloor N \alpha \rfloor \) and \( e \) is the column vector of all ones.

In order to obtain tighter MIQP reformulation whose continuous relaxation is tighter than that of (MIQP), we are going to construct a semidefinite program (SDP) relaxation of (MIQP) via Lagrangian decomposition and then derive a new second-order cone program (SOCP) reformulation of (MIQP).

Let \( v_i = (\xi_i^T B x) \) for \( i = 1, \ldots, N. \) Let

\[
\Theta = \{ \theta \in \mathbb{R}^N \mid Q - \sum_{i=1}^N \theta_i B^T \xi_i (\xi_i^T B) \succeq 0 \}.
\]

For any \( \theta \in \Theta \), consider the following equivalent formulation of problem (MIQP):

\[
\begin{align*}
\text{(P_0)} \quad & \min x^T (Q - \sum_{i=1}^N \theta_i B^T \xi_i (\xi_i^T B)) x + c^T x + \sum_{i=1}^N \theta_i v_i^2 \\
\text{s.t.} \quad & v_i = (\xi_i^T B x), \quad i = 1, \ldots, N, \\
& v_i \geq R + y_i (l_i^b - R), \quad i = 1, \ldots, N, \\
& e^T y \leq K, \quad y \in \{0, 1\}^N, \\
& x \in X, \quad l_i^b \leq v \leq u_i^b.
\end{align*}
\]
where \( v = (v_1, \ldots, v_N)^T \). Note that the objective function in \((P_\theta)\) is decomposed as the sum of a convex nonseparable quadratic function of \( x \) and a separable quadratic function of \( v \). The constraint \( v_i = (\xi_i)^T B x \) in \((P_\theta)\) can be viewed as a link constraint for the \( i \)th scenario. Associating a multiplier \( \lambda_i \) to each constraint \( v_i = (\xi_i)^T B x \), we obtain the following Lagrangian relaxation problem of \((P_\theta)\):

\[
d(\lambda) = \min \ x^T (Q - \sum_{i=1}^{N} \theta_i B^T \xi_i (\xi_i)^T B) x + c^T x + \sum_{i=1}^{N} \theta_i v_i^2
\]

\[
+ \sum_{i=1}^{N} \lambda_i (v_i - (\xi_i)^T B x)
\]

s.t. \( v_i \geq R + y_i (l_i^b - R), \ y_i \in \{0, 1\}, \ i = 1, \ldots, N, \)
\( x \in X, \ e^T y \leq K, \ l_i^b \leq v \leq u_i^b, \)
\[
= \ d_1(\lambda) + d_2(\lambda),
\]

where

\[
d_1(\lambda) = \min_{x \in X} \{ x^T (Q - \sum_{i=1}^{N} \theta_i B^T \xi_i (\xi_i)^T B)x + (c - \sum_{i=1}^{N} \lambda_i B^T \xi_i) x \},
\]

\[
d_2(\lambda) = \min_{i=1}^{N} \theta_i v_i^2 + \lambda_i v_i
\]

s.t. \( v_i \geq R + y_i (l_i^b - R), \ y_i \in \{0, 1\}, \ i = 1, \ldots, N, \)
\( c^T y \leq K, \ l_i^b \leq v \leq u_i^b. \)

For any \( \theta \in \Theta \) and \( \lambda \in \mathbb{R}^N \), we have \( d(\lambda) \leq v(P_\theta) = v(P) \). The dual problem of \((P_\theta)\) is

\[
(D_\theta) \ \max_{\lambda \in \mathbb{R}^N} \ d(\lambda).
\]

Let \((QP)\) denote the continuous relaxation of \((MIQP_\theta)\). It is easy to show that for any \( \theta \in \Theta \cap \mathbb{R}^N_+ \), it holds that \( v(D_\theta) \geq v(QP) \). It was shown in [32] that \((D_\theta)\) can be reduced to an SDP problem. Moreover, for any fixed \( \theta \geq 0 \) and \( \theta \in \Theta \), problem \((D_\theta)\) can be actually reduced to an SOCP problem. The following tighter MIQP reformulation is then derived in [32]:

\[
(MIQP) \ \ \min_{\theta \in \mathbb{R}^N} \ x^T (Q - \sum_{i=1}^{N} \theta_i B^T \xi_i (\xi_i)^T B)x + c^T x + \sum_{i=1}^{N} \theta_i (w_i^2 + \phi_i - R^2 y_i)
\]

s.t. \( (\xi_i)^T B x = w_i + z_i - y_i R, \ i = 1, \ldots, N, \)
\( c^T y \leq K, \ y \in \{0, 1\}^N, \)
\( l_i^b y_i \leq z_i \leq R y_i, \ \phi_i y_i \geq z_i^2, \ \phi_i \geq 0, \ i = 1, \ldots, N, \)
\( x \in X, \ R \leq w_i \leq u_i^b, \ i = 1, \ldots, N, \)

where \( \theta \in \mathbb{R}^N \) satisfies \( Q - \sum_{i=1}^{N} \theta_i B^T \xi_i (\xi_i)^T B \succeq 0 \). Different choices of \( \theta \) were discussed in [32].

When the functions involved in \((PCP)\) are linear and \( X \) is defined by linear constraints, and the probabilistic constraint has the special form: \( \mathbb{P}\{Tx \geq \xi\} \geq 1 - \alpha \), valid inequalities and strengthened mixed-integer linear programming reformulations were presented in [24] and [19].
5. Other Methods. Several other methods have been proposed in the literature for general or special probabilistically constrained optimization problems:

- Scenario method ([9, 26]);
- Sample average method ([23]);
- p-efficient point method ([12, 13, 14, 21]);
- MIP reformulations for stochastic set covering problem ([6, 31]).

We now briefly introduce the first two methods.

5.1. Scenario method. Scenario method directly uses the finite number of samples or scenarios $\xi_j$ ($j = 1, \ldots, N$) in the probabilistic constraints, resulting in a deterministic optimization problem:

$$
(SP) \quad \min f(x) \\
\text{s.t. } c(x, \xi_j) \leq 0, \ j = 1, \ldots, N, \\
x \in X.
$$

Two advantages of scenario method: (1) it has no requirement for the distribution of $\xi$; (2) if $f$, $c$ and $X$ are convex, $(SP)$ is also a convex optimization problem. Therefore, when $N$, the number of samples, is not large, problem $(SP)$ can be easily solved. However, the disadvantage of scenario method is also obvious: because different samples lead to different constraints, the problem $(SP)$ itself is random and hence the optimal solution is also random. An important problem is how to guarantee that the optimal solution of $(SP)$ is also feasible to $(PCP)$ with high probability. It was shown in [9] that if the number of samples satisfies

$$
N \geq \frac{2}{\alpha} \log(\frac{1}{\delta}) + 2n + \frac{2n}{\alpha} \log(\frac{2}{\alpha}),
$$

then the optimal solution of $(SP)$ is feasible to $(PCP)$ with probability $1 - \delta \in (0, 1)$ (see also [26]).

5.2. Sample average method. Sample average method is a simple way to estimate the probability function $P\{c(x, \xi) \leq 0\}$. Suppose $\xi^1, \ldots, \xi^N$ are $N$ independent samples of the random vector $\xi$. For any given $\epsilon \in (0, 1)$, consider the following sample average problem:

$$
(SAP) \quad \min f(x) \\
\text{s.t. } \frac{1}{N} \sum_{j=1}^{N} 1_{(\infty, 0)}(c(x, \xi^j)) \geq 1 - \epsilon, \\
x \in X.
$$

Obviously, when $\epsilon = 0$, problem $(SAP)$ is exactly the scenario approximation problem $(SP)$. It was shown in [23] that when $\epsilon > \alpha$, $(SAP)$ generates a lower bound of the optimal solution of the original problem with probability converging to 1 at exponential speed. When $\epsilon < \alpha$, under certain condition, the optimal solution of the sample average problem $(SAP)$ is feasible to the original problem with high probability.

6. Conclusions and Research Perspectives. Parameter uncertainty or ambiguity arises in many practical applications of optimization models. Optimization under uncertainty has drawn much attention in recent years and is one of the most important and challenging topics in modern optimization and operations research.
Probabilistic or chance constraint is one of the popular ways to deal with the uncertainty. Probabilistically constrained programming has been successfully used in modeling optimization problems in many application areas such as finance, engineering and management. However, in contrast to the well-developed theory, methods and software for (deterministic) convex optimization problems, great research efforts are still needed for the study of probabilistically constrained optimization.

When the parameter ambiguity is described by some simple uncertainty set, the optimization problem under set uncertainty can be dealt with by robust optimization methods. When the uncertainty sets are of special forms such as bounds or ellipsoids, the resulting robust problem can be reduced to second-order cone programming or semidefinite programming problems which can be efficiently solved by interior-point methods. The major disadvantage of robust optimization method is that the optimal solution of the robust model is often too conservative to be useful in practical situations.

There are many challenging problems in probabilistically constrained optimization. One of the interesting research topics is the study of approximation methods for the mixed-integer programming models with large-size samples or scenarios. When there is no prior knowledge about the distribution of the random variables, it is a natural choice to use historical data or samples of simulation to approximate the distribution. The resulting model is mixed-integer 0-1 programming with special structures. Due to the large size of the samples, it is very difficult or impossible to solve the model to global optimality even using the most powerful commercial mixed-integer programming solvers. A heuristic method based on penalty decomposition and alternating direction method was proposed in [3] to solve the mixed-integer programming model of the probabilistically constrained problem. Further efforts are needed to design more efficient approximation methods for probabilistically constrained problems arising from real-world applications such as finance, management and engineering.

**REFERENCES**

[1] G. J. Alexander and A. M. Baptista, *A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model*, Management Science, 50 (2004), 1261–1273.

[2] G. J. Alexander, A. M. Baptista and S. Yan, *Mean-variance portfolio selection with ‘at-risk’ constraints and discrete distributions*, Journal of Banking & Finance, 31 (2007), 3761–3781.

[3] X. D. Bai, J. Sun, X. J. Zheng and X. L. Sun, *A penalty decomposition method for probabilistically constrained convex programs*, Technical report, School of Management, Fudan University, P. R. China, 2012. Available from: http://www.optimization-online.org/DB-FILE/2012/04/3448.pdf.

[4] A. Ben-Tal, L. El Ghaoui and A. S. Nemirovski, “Robust Optimization,” Princeton University Press, 2009.

[5] S. Benati and R. Rizzi, *A mixed integer linear programming formulation of the optimal mean/Value-at-Risk portfolio problem*, European Journal of Operational Research, 176 (2007), 423–434.

[6] P. Beraldi and A. Ruszczyński, *The probabilistic set-covering problem*, Operations Research, 50 (2002), 956–967.

[7] P. Bonami and M. A. Lejeune, *An exact solution approach for portfolio optimization problems under stochastic and integer constraints*, Operations Research, 57 (2009), 650–670.

[8] S. P. Boyd and L. Vandenberghe, “Convex Optimization,” Cambridge University Press, 2004.

[9] G. C. Calafiore and M. C. Campi, *The scenario approach to robust control design*, IEEE Transactions on Automatic Control, 51 (2006), 742–753.

[10] A. Charnes and W. W. Cooper, *Chance-constrained programming*, Management Science, 6 (1959), 73–79.
[11] W. Chen, M. Sim, J. Sun and C. P. Teo, *From CVaR to uncertainty set: Implications in joint chance-constrained optimization*, Operations Research, 58 (2010), 470–485.

[12] M. S. Cheon, S. Ahmed and F. Al-Khayyal, *A branch-reduce-cut algorithm for the global optimization of probabilistically constrained linear programs*, Mathematical Programming, 108 (2006), 617–634.

[13] D. Dentcheva and G. Martínez, *Augmented Lagrangian method for probabilistic optimization*, Annals of Operations Research, 200 (2012), 109-130.

[14] D. Dentcheva, A. Prékopa and A. Ruszczyński, *Concavity and efficient points of discrete distributions in probabilistic programming*, Mathematical Programming, 89 (2000), 55–77.

[15] A. A. Gaivoronski and G. Pflug, *Value at risk in portfolio optimization: Properties and computational approach*, Journal of Risk, 7 (2005), 1–31.

[16] R. Henrion, *Structural properties of linear probabilistic constraints*, Optimization, 56 (2007), 425–440.

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

[18] L. J. Hong, Y. Yang, and L. Zhang, *Sequential convex approximations to joint chance constrained programs: A Monte Carlo approach*, Operations Research, 59 (2011), 617–630.

[19] S. Küçükyavuz, *On mixing sets arising in chance-constrained programming*, Mathematical Programming, 132 (2012), 31–56.

[20] C. M. Lagoa, X. Li, and M. Sznajer, *Probabilistically constrained linear programs and risk-adjusted controller design*, SIAM Journal on Optimization, 15 (2005), 938–951.

[21] M. Lejeune and N. Noyan, *Mathematical programming approaches for generating p-efficient points*, European Journal of Operational Research, 207 (2010), 590–600.

[22] M. A. Lejeune and A. Ruszczyński, *An efficient trajectory method for probabilistic inventory production-distribution problems*, Operations Research, 55 (2007), 378–394.

[23] J. Luedtke and S. Ahmed, *A sample approximation approach for optimization with probabilistic constraints*, SIAM Journal on Optimization, 19 (2008), 674–699.

[24] J. Luedtke, S. Ahmed and G. L. Nemhauser, *An integer programming approach for linear programs with probabilistic constraints*, Mathematical Programming, 122 (2010), 247–272.

[25] A. Nemirovski and A. Shapiro, *Convex approximations of chance constrained programs*, SIAM Journal on Optimization, 17 (2006), 969–996.

[26] A. Nemirovski and A. Shapiro, *Scenario approximations of chance constraints*, in “Probabilistic and Randomized Methods for Design Under Uncertainty” (eds. G. Calafiore and F. Dabbene), Springer, (2006), 3–47.

[27] J. Nocedal and S. J. Wright, “Numerical Optimization,” Springer, 1999.

[28] A. Prékopa, *Probabilistic programming*, in “Stochastic Programming,” Handbooks in Operations Research and Management Science (eds. A. Ruszczyński and A. Shapiro), Elsevier, (2003), 267–351.

[29] R. T. Rockafellar and S. Uryasev, *Optimization of conditional value-at-risk*, Journal of Risk, 2 (2000), 21–42.

[30] A. Ruszczyński, *Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra*, Mathematical Programming, 93 (2002), 195–215.

[31] A. Saxena, V. Goyal and M. A. Lejeune, *MIP reformulations of the probabilistic set covering problem*, Mathematical Programming, 121 (2010), 1–31.

[32] X. J. Zheng, X. L. Sun, D. Li and X. T. Cui, *Lagrangian decomposition and mixed-integer quadratic programming reformulations for probabilistically constrained quadratic programs*, European Journal of Operational Research, 221 (2012), 38–48.

Received November 2011; 1st revision June 2012; final revision October 2012.

E-mail address: xdbaifudan.edu.cn
E-mail address: xjzheng@tongji.edu.cn
E-mail address: xls@fudan.edu.cn