Efficient Point Methods for Probabilistic Optimization Problems

Darinka Dentcheva[*], Bogumila Lai[†] and Andrzej Ruszczyński[‡]

August 8, 2003

Abstract

We consider nonlinear stochastic programming problems with probabilistic constraints. The concept of a p-efficient point of a probability distribution is used to derive equivalent problem formulations, and necessary and sufficient optimality conditions. We analyze the dual functional and its subdifferential. Two numerical methods are developed based on approximations of the p-efficient frontier. The algorithms yield an optimal solution for problems involving r-concave probability distributions. For arbitrary distributions, the algorithms provide upper and lower bounds for the optimal value and nearly optimal solutions. The operation of the methods is illustrated on a cash matching problem with a probabilistic liquidity constraint.

Keywords: Stochastic Programming, Convex Programming, Probabilistic Constraints, Dual Methods

1 Introduction

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, be concave functions, and let $\mathcal{D} \subseteq \mathbb{R}^n$ be a closed convex set. We consider the convex programming problem

$$\begin{aligned} &\max f(x) \\ \text{subject to } &g(x) \ge Y, \\ &x \in \mathcal{D}. \end{aligned}$$

For two vectors a and b the inequality $a \le b$ is understood componentwise.

[*]Stevens Institute of Technology, Department of Mathematics, Hoboken, NJ 07030, U.S.A.

[†]Stevens Institute of Technology, Department of Mathematics, Hoboken, NJ 07030, U.S.A.

[‡]Rutgers University, Department of Management Science and Information Systems, Piscataway, NJ 08854, U.S.A.

If the vector $Y \in \mathbb{R}^m$ is random, we require that $g(x) \geq Y$ shall hold at least with some prescribed probability $p \in (0,1)$, rather than *for all* possible realizations of the right hand side. This leads to the *nonlinear programming problem with probabilistic constraints*:

$$\begin{aligned} &\max f(x) \\ &\text{subject to } \mathbb{P}\big[g(x) \geq Y\big] \geq p, \\ &x \in \mathcal{D}, \end{aligned} \tag{1}$$

where the symbol $\mathbb{P}$ denotes probability. For a detailed presentation of the theory and numerical methods for linear models with one probabilistic constraint on finitely many inequalities, we refer the reader to [13].

The formulation of the problem with probabilistic constraints is in harmony with the basic statistical principles used in testing statistical hypotheses and other statistical decisions. It is also in agreement with the decision principles used in actuarial mathematics, engineering, finance, etc. Problems with reliability constraints are of great practical importance in telecommunication, transportation, network design and operation, engineering structure design, electronic manufacturing problems, etc. In finance, the concept of *Value at Risk* represents a probabilistic constraint on the asset–liability balance of the company. For recent publications in this area we refer the Reader to [3], [6], [5],and [14].

2 p-Efficient Points

Let us define the set

$$\mathcal{Z}_p = \big\{y \in \mathbb{R}^m : \mathbb{P}[Y \leq y] \geq p\big\}. \tag{2}$$

Clearly, problem (1) can be compactly rewritten as

$$\begin{aligned} &\max f(x) \\ &\text{subject to } g(x) \in \mathcal{Z}_p, \\ &x \in \mathcal{D}. \end{aligned} \tag{3}$$

Lemma 2.1 *For every $p \in (0,1)$ the set $\mathcal{Z}_p$ is nonempty and closed.*

Proof: The assertion follows from the monotonicity and the right continuity of the distribution function. □

Let F denote the probability distribution function of Y, and let F_i be the marginal probability distribution function of the ith component Y_i.
We recall the concept of a p-efficient point, which we studied in the context of discrete distributions and linear problems in the papers [1, 2].

Definition 2.2 *Let $p \in (0,1]$. A point $v \in \mathbb{R}^m$ is called a* p-efficient *point of the probability distribution function F, if $F(v) \geq p$ and there is no $y \leq v$, $y \neq v$ such that $F(y) \geq p$.*

Obviously, for a scalar random variable Y and for every $p \in (0,1]$ there is exactly one p-efficient point: the smallest v such that $F(v) \geq p$. Since $F(v) \leq F_i(v_i)$ for every $v \in \mathbb{R}^m$ and $i = 1, \ldots, m$, we obtain that the set of p-efficient points is bounded from below.

Lemma 2.3 *Let $p \in (0,1]$ and let l_i be the p-efficient point of the one-dimensional marginal distribution F_i, $i = 1, \ldots, m$. Then every $v \in \mathbb{R}^m$ such that $F(v) \geq p$ must satisfy the inequality $v \geq l = (l_1, \ldots, l_m)$.*

In [1] the following fact is shown: For each $p \in (0,1)$ the set of p-efficient points of an integer random vector is nonempty and finite. For a general random vector the set of p-efficient points may be unbounded and not closed.

Let $p \in (0,1)$ and let v^j, $j \in J$, be *all* p-efficient points of Y, where J is an arbitrary set. We define the cones

$$K_j = v^j + \mathbb{R}^m_+, \quad j \in J.$$

The following result can be derived from the Phelps theorem [7, Lemma 3.12] about the existence of conical support points, but we provide an easy direct proof.

Proposition 2.4 $\mathcal{Z}_p = \bigcup_{j \in J} K_j$.

Proof: If $y \in \mathcal{Z}_p$ then either y is p-efficient or there exists a vector w such that $w \leq y$, $w \neq y$, $w \in \mathcal{Z}_p$. By Lemma 2.3, one must have $l \leq w \leq y$. The set $Z_1 := \{z \in \mathcal{Z}_p : l \leq z \leq y\}$ is compact by the closedness of $\mathcal{Z}_p$. Thus, there exists $w^1 \in Z_1$ with the minimal first coordinate. If w^1 is a p-efficient point, then $y \in w^1 + \mathbb{R}^m_+$, what had to be shown. Otherwise, we define $Z_2 := \{z \in \mathcal{Z}_p : l \leq z \leq w^1\}$, and choose a point $w^2 \in Z_2$ with the minimal second coordinate. Proceeding in the same way, we shall find the minimal element w^m in the set $\mathcal{Z}_p$ with $w^m \leq w^{m-1} \leq \cdots \leq y$. Therefore, $y \in w^m + \mathbb{R}^m_+$, and this completes the proof. □

By virtue of Proposition 2.4 we obtain (for $0 < p < 1$) the following *disjunctive semi-infinite* formulation of problem (3):

$$\begin{aligned} &\max f(x) \\ &\text{subject to } g(x) \in \bigcup_{j \in J} K_j, \\ &x \in \mathcal{D}. \end{aligned} \tag{4}$$

Its main advantage is an insight into the nature of the non-convexity of the feasible set. The main difficulty is the implicit character of the disjunctive constraint.

Let S stand for the simplex in $\mathbb{R}^{m+1}$, $S = \{u \in \mathbb{R}^{m+1} : \sum_{j=1}^{m+1} u_j = 1, u_j \geq 0\}$. We define the convex hull of the p-efficient points:

$$E = \Big\{ \sum_{j=1}^{m+1} \lambda_j v^{k_j} : \lambda \in S,\ k_j \in J \Big\}.$$

The convex hull of $\mathcal{Z}_p$ has a semi-infinite disjunctive representation as well.

Lemma 2.5 $\operatorname{co} \mathcal{Z}_p = E + \mathbb{R}^m_+$.

Proof: By Proposition 2.4 every point $y \in \operatorname{co} \mathcal{Z}_p$ can be represented as a convex combination of points in the cones K_j. By the theorem of Caratheodory we can write $y = \sum_{j=1}^{m+1} \alpha_j (v^j + w^j)$, where $w^j \in \mathbb{R}^m_+$, $\alpha_j \geq 0$, $j = 1, \dots, m+1$, and $\sum_{j=1}^{m+1} \alpha_j = 1$. The vector $w = \sum_{j=1}^{m+1} \alpha_j w^j \in \mathbb{R}^m_+$. Therefore, $y \in \sum_{j=1}^{m+1} \alpha_j v^j + \mathbb{R}^m_+$. □

Proposition 2.6 *For every $p \in (0,1)$ the set $\operatorname{co} \mathcal{Z}_p$ is closed.*

Proof: Consider a sequence $\{z^k\}$ of points of $\operatorname{co} \mathcal{Z}_p$ which is convergent to a point $\bar{z}$. We have

$$z^k = \sum_{i=1}^{m+1} \alpha_i^k y_i^k,$$

with $y_i^k \in \mathcal{Z}_p$, $\alpha_i^k \geq 0$, and $\sum_{i=1}^{m+1} \alpha_i^k = 1$. By passing to a subsequence, if necessary, we can assume that the limits

$$\bar{\alpha}_i = \lim_{k \to \infty} \alpha_i^k$$

exist for all $i = 1, \dots, m+1$. By Lemma 2.3 all points y_i^k are bounded below by some vector l. For simplicity of notation we may assume that $l = 0$.

Let $I = \{i : \bar{\alpha}_i > 0\}$. Clearly, $\sum_{i \in I} \bar{\alpha}_i = 1$. We obtain

$$z^k \geq \sum_{i \in I} \alpha_i^k y_i^k.$$

We observe that $0 \leq \alpha_i^k y_i^k \leq z^k$ for all $i \in I$ and all k. Since $\{z^k\}$ is convergent and $\alpha_i^k \to \bar{\alpha}_i > 0$, each sequence $\{y_i^k\}$, $i \in I$, is bounded. Therefore we can assume that each of them is convergent to some limit $\bar{y}_i$, $i \in I$. By virtue of Lemma 2.1 $\bar{y}_i \in \mathcal{Z}_p$. Passing to the limit in the last displayed inequality we obtain

$$\bar{z} \geq \sum_{i \in I} \bar{\alpha}_i \bar{y}_i \in \operatorname{co} \mathcal{Z}_p.$$

Due to Lemma 2.5, $\bar{z} \in \operatorname{co} \mathcal{Z}_p$. □

Proposition 2.7 *For every $p \in (0,1)$ the set of extreme points of $\operatorname{co} \mathcal{Z}_p$ is nonempty and it is contained in the set of p-efficient points.*

Proof: The set $\operatorname{co} \mathcal{Z}_p$ is included in $l + \mathbb{R}^m_+$, by virtue of Lemma 2.3. Therefore it does not contain any line. Since it is closed by Proposition 2.6, it has at least one extreme point.

Let w be an extreme point of $\operatorname{co} \mathcal{Z}_p$. Thus $\operatorname{co} \mathcal{Z}_p \setminus \{w\}$ is convex and w can be separated from this set. Consequently, there exists $u \neq 0$ such that

$$\langle u, w \rangle \leq \langle u, z \rangle \quad \text{for all} \quad z \in \operatorname{co} \mathcal{Z}_p.$$

Since $\mathbb{R}^m_+$ is the recession cone of $\operatorname{co}\mathcal{Z}_p$, we have $u \geq 0$. Suppose that w is not a p-efficient point. Then Proposition 2.4 implies that there exists a p-efficient point $v \leq w$, $v \neq w$. Since $w + \mathbb{R}^m_+ \subset \operatorname{co}\mathcal{Z}_p$, the point w is a convex combination of v and $w + (w - v)$. Consequently, w cannot be extreme. □

3 Lagrangian Relaxation

Let us split variables in problem (3):

$$\begin{aligned} \max\ & f(x) \\ & g(x) \geq z, \\ & x \in \mathcal{D}, \\ & z \in \mathcal{Z}_p. \end{aligned} \tag{5}$$

Associating Lagrange multipliers $u \in \mathbb{R}^m_+$ with constraints (5) we obtain the Lagrangian function:

$$L(x, z, u) = f(x) + \langle u, g(x) - z \rangle.$$

The dual functional has the form

$$\Psi(u) = \sup_{(x,z) \in \mathcal{D} \times \mathcal{Z}_p} L(x, z, u) = h(u) - d(u),$$

where

$$h(u) = \sup\{f(x) + \langle u, g(x) \rangle \mid x \in \mathcal{D}\}, \tag{6}$$

$$d(u) = \inf\{\langle u, z \rangle \mid z \in \mathcal{Z}_p\}. \tag{7}$$

For any $u \in \mathbb{R}^m_+$ the value of $\Psi(u)$ is an upper bound on the optimal value F^* of the original problem. The best Lagrangian upper bound will be given by

$$D^* = \inf_{u \geq 0} \Psi(u). \tag{8}$$

For $u \not\geq 0$ one has $d(u) = -\infty$, because the set $\mathcal{Z}_p$ contains a translation of K_+. The function $d(\cdot)$ is concave and one can easily see that

$$d(u) = \inf\{\langle u, z \rangle \mid z \in \operatorname{co}\mathcal{Z}_p\}. \tag{9}$$

Let us consider the *convex hull problem*:

$$\begin{aligned} \max\ & f(x) \\ & g(x) \geq z, \\ & x \in \mathcal{D}, \\ & z \in \operatorname{co}\mathcal{Z}_p. \end{aligned} \tag{10}$$

We shall make the following assumption.
Constraint Qualification Condition. *There exist points* $x^0 \in \mathcal{D}$ *and* $z^0 \in \operatorname{co} \mathcal{Z}_p$ *such that* $g(x^0) > z^0$.

If the Constraint Qualification Condition is satisfied, from the duality theory in convex programming [11, Corollary 28.2.1] we know that there exists $\hat{u} \geq 0$ at which the minimum in (8) is attained, and $D^* = \Psi(\hat{u})$ is the optimal value of the convex hull problem (10).

We now study in detail the structure of the dual functional Ψ. We shall characterize the solution sets of the two subproblems (6) and (7), which provide values of the dual functional. Let us define the following sets:

$$V(u) = \{v \in \mathbb{R}^m : \langle u, v \rangle = d(u) \text{ and } v \text{ is a } p\text{-efficient point}\}, \tag{11}$$

$$C(u) = \{d \in \mathbb{R}^m_+ : d_i = 0 \text{ if } u_i > 0,\ i = 1, \dots, s\}. \tag{12}$$

Lemma 3.1 *For every* $u > 0$ *the solution set of (7) is nonempty. Foe every* $u \geq 0$ *it has the following form:* $\hat{Z}(u) = V(u) + C(u)$.

Proof: Let us at first consider the case $u > 0$. Then every recession direction d of $\mathcal{Z}_p$ satisfies $\langle u, d \rangle > 0$. Since $\mathcal{Z}_p$ is closed, a solution to (7) must exist. Suppose that a solution z to (7) is not a p-efficient point. By virtue of Proposition 2.4, there is a p-efficient $v \in \mathcal{Z}_p$ such that $v \leq z$, and $v \neq z$. Thus, $\langle u, v \rangle < \langle u, z \rangle$, which is a contradiction.

In the general case $u \geq 0$, the solution set of the problem to (7), if it is nonempty, always contains a p-efficient point. Indeed, if a solution z is not p-efficient, we must have a p-efficient point v dominated by z, and $\langle u, v \rangle \leq \langle u, z \rangle$ holds by the nonnegativity of u. Consequently, $\langle u, v \rangle \leq \langle u, z \rangle$ for all p-efficient $v \leq z$, which is equivalent to $z \in \{v\} + C(u)$, as required.

If the solution set of (7) is empty then $V(u) = \emptyset$ and the assertion is true as well. □

The last result allows us to calculate the subdifferential of d in a closed form.

Lemma 3.2 *For every* $u \geq 0$ *one has* $\partial d(u) = \operatorname{co} V(u) + C(u)$. *If* $u > 0$ *then* $\partial d(u) \neq \emptyset$.

Proof: From (7) we obtain $d(u) = -\delta^*_{\mathcal{Z}_p}(-u)$, where $\delta^*_{\mathcal{Z}_p}(\cdot)$ is the support function of $\mathcal{Z}_p$ and, consequently, of $\operatorname{co} \mathcal{Z}_p$. This fact follows from the structure of $\mathcal{Z}_p$ described Proposition 2.4, by virtue of Corollary 16.5.1 in [11]. Thus

$$\partial d(u) = \partial \delta^*_{\mathcal{Z}_p}(-u).$$

Recall that $\operatorname{co} \mathcal{Z}_p$ is closed, by Proposition 2.6. Using [11, Thm 23.5], we observe that $s \in \partial \delta^*_{\mathcal{Z}_p}(-u)$ if and only if $\delta^*_{\mathcal{Z}_p}(-u) + \delta_{\operatorname{co} \mathcal{Z}_p}(s) = -\langle s, u \rangle$, where $\delta_{\operatorname{co} \mathcal{Z}_p}(\cdot)$ is the indicator function of $\operatorname{co} \mathcal{Z}_p$. It follows that $s \in \operatorname{co} \mathcal{Z}_p$ and $\delta^*_{\mathcal{Z}_p}(-u) = -\langle s, u \rangle$. Consequently,

$$\langle s, u \rangle = d(u). \tag{13}$$

Since $s \in \operatorname{co} \mathcal{Z}_p$ we can represent it as follows:

$$s = \sum_{j=1}^{m+1} \alpha_j e^j + w,$$

where e^j, $j = 1, \ldots, m+1$, are extreme points of $\operatorname{co} \mathcal{Z}_p$ and $w \geq 0$. Using Proposition 2.7 we conclude that e^j are p-efficient points. Moreover

$$\langle s, u \rangle = \sum_{j=1}^{m+1} \alpha_j \langle u, e^j \rangle + \langle u, w \rangle \geq d(u), \tag{14}$$

because $\langle u, e^j \rangle \geq d(u)$ and $\langle u, w \rangle \geq 0$. Combining (13) and (14) we conclude that $\langle u, e^j \rangle = d(u)$ for all j, and $\langle u, w \rangle = 0$. Thus $s \in \operatorname{co} V(u) + C(u)$.

Conversely, if $s \in \operatorname{co} V(u) + C(u)$ then (13) holds true. This implies that $s \in \partial d(u)$, as required.

The set $\partial d(u)$ is nonempty for $u >$ by virtue of Lemma 3.1. □

Let us turn now to the function $h(\cdot)$. Define the set of maximizers in (6),

$$X(u) = \{x \in \mathcal{D} : f(x) + \langle u, g(x) \rangle = h(u)\}.$$

Lemma 3.3 *Assume that the set $\mathcal{D}$ is compact. The subdifferential of the function h is described as follows for every $u \in \mathbb{R}^m$:*

$$\partial h(u) = \operatorname{co} \{g(x) : x \in X(u)\}.$$

Proof: The function h is convex on $\mathbb{R}^m$. Since the set $\mathcal{D}$ is compact and f and g are concave, the set $X(u)$ is compact. Therefore, the subdifferential of $h(u)$ for every $u \in \mathbb{R}^m$ is the closure of $\operatorname{co} \{g(x) : x \in X(u)\}$ (see [4, Chapter VI,Lemma 4.4.2]). By the compactness of $X(u)$ and concavity of g, the set $\{g(x) : x \in X(u)\}$ is closed. Therefore, we can omit taking the closure in the description of the subdifferential of $h(u)$. □

This analysis provides the basis for the following necessary and sufficient optimality conditions for problem (8).

Theorem 3.4 *Assume that the Constraint Qualification Condition is satisfied and that the set $\mathcal{D}$ is compact. A vector $u \geq 0$ is an optimal solution of (8) if and only if there exists a point $x \in X(u)$, points $v^1, \ldots, v^{m+1} \in V(u)$ and scalars $\beta_1 \ldots, \beta_{m+1} \geq 0$ with $\sum_{j=1}^{m+1} \beta_j = 1$, such that*

$$g(x) - \sum_{j=1}^{m+1} \beta_j v^j \in C(u). \tag{15}$$

where $C(u)$ is given by (12).

Proof: Since $-C(u)$ is the normal cone to the positive orthant at $u \geq 0$, the necessary and sufficient optimality condition for (8) has the form

$$\partial \Psi(u) \cap C(u) \neq \emptyset \tag{16}$$

(cf. [11, Thm. 27.4]). Since $\operatorname{int}\operatorname{dom} d \neq \emptyset$ and $\operatorname{dom} h = \mathbb{R}^m$ we have $\partial\Psi(u) = \partial h(u) - \partial d(u)$. Using Lemma 3.2 and Lemma 3.3, we conclude that there exist

$$\begin{aligned}
&p\text{-efficient points } v^j \in V(u), \quad j = 1, \ldots, m+1, \\
&\beta^j \geq 0, \quad j = 1, \ldots, m+1, \quad \sum_{j=1}^{m+1} \beta_j = 1, \\
&x^j \in X(u), \quad j = 1, \ldots, m+1, \\
&\alpha^j \geq 0, \quad j = 1, \ldots, m+1, \quad \sum_{j=1}^{m+1} \alpha_j = 1,
\end{aligned} \tag{17}$$

such that

$$\sum_{j=1}^{m+1} \alpha_j g(x^j) - \sum_{j=1}^{m+1} \beta_j v^j \in C(u). \tag{18}$$

Let us define

$$x = \sum_{j=1}^{m+1} \alpha_j x^j.$$

By the convexity of $X(u)$ we have $x \in X(u)$.

From the concavity of f and g_j we obtain

$$f(x) + \sum_{i=1}^{m} u_i g_i(x) \geq f(x^j) + \sum_{i=1}^{m} u_i g_i(x^j), \quad j = 1, \ldots, m+1.$$

In view of (17),

$$f(x) + \sum_{i=1}^{m} u_i g_i(x) \leq f(x^j) + \sum_{i=1}^{m} u_i g_i(x^j), \quad j = 1, \ldots, m+1.$$

Consequently,

$$f(x) + \sum_{i=1}^{m} u_i g_i(x) = f(x^j) + \sum_{i=1}^{m} u_i g_i(x^j), \quad j = 1, \ldots, m+1. \tag{19}$$

Multiplying the last equation by α_j and adding we obtain

$$f(x) + \sum_{i=1}^{m} u_i g_i(x) = \sum_{j=1}^{m+1} \alpha_j \Big[f(x^j) + \sum_{i=1}^{m} u_i g_i(x^j) \Big].$$

Since $g_i(x) \geq \sum_{j=1}^{m+1} \alpha_j g_i(x^j)$, substituting into the above equation, we obtain

$$f(x) \leq \sum_{j=1}^{m+1} \alpha_j f(x^j).$$

If $g_i(x) \geq \sum_{j=1}^{m+1} \alpha_j g_i(x^j)$ and $u_i > 0$ for some i, the above inequality becomes strict, in contradiction to the concavity of f. Thus, for all $u_i > 0$ we have $g_i(x) = \sum_{j=1}^{m+1} \alpha_j g_i(x^j)$, and it follows that

$$g(x) - \sum_{j=1}^{m+1} \alpha_j g(x^j) \in C(u).$$

Therefore relation (18) can be simplified as (15), as required.

To prove the converse implication assume that we have $x \in X(u)$, points $v^1, \dots, v^{m+1} \in V(u)$ and scalars $\beta_1 \dots, \beta_{m+1} \geq 0$ with $\sum_{j=1}^{m+1} \beta_j = 1$, such that (15) holds true. By Lemma 3.2 and Lemma 3.3 we have

$$g(x) - \sum_{j=1}^{m+1} \beta_j v^j \in \partial\Psi(u).$$

Thus (15) implies (16), which is a necessary and sufficient optimality condition for (8). □

Since the set of p-efficient points is not known, we need a numerical method for solving the convex hull problem (10) or its dual (8).

4 The dual method

The idea of our first numerical method is to solve the dual problem (8) using the information about the subgradients of the dual functional Ψ to generate convex piecewise-linear approximations of Ψ. Suppose that the values of the functional Ψ at certain points u^j, $j = 1, \dots, k$, are available. Moreover, we assume that the corresponding solutions $v^1, \dots, v^k$ and $x^1, \dots, x^k$ of the two problems (9) and (6) are available as well. According to Lemma 3.1 we can assume that v^j, $j = 1, \dots, k$ are p-efficient points. By virtue of Lemma 3.3 and Lemma 3.2 the following function $\Psi_k(\cdot)$ is a lower bound of Ψ:

$$\Psi_k(u) := \max_{1 \leq j \leq k} \left[\Psi(u^j) + \langle g(x^j) - v^i, u - u^j \rangle \right].$$

Minimizing $\Psi_k(u)$ over $u \geq 0$, we obtain the next iterate u^{k+1}. For the purpose of numerical tractability, we shall impose an upper bound $b \in \mathbb{R}$ on the dual variables u_j. We define the feasible set of the dual problem as follows:

$$U := \{u \in \mathbb{R}^m : 0 \leq u_i \leq b,\ i = 1, \dots, m\}$$

where b is a sufficiently large number. We also use $\varepsilon > 0$ as a stopping test parameter.

The Dual Method

Step 0: Select a vector $u^1 \in U$. Set $\Psi_0(u^1) = -\infty$ and $k = 1$.

Step 1: Calculate

$$h(u^k) = \max\{f(x) + \langle u^k, g(x)\rangle \mid x \in \mathcal{D}\}, \tag{20}$$

$$d(u^k) = \min\{\langle u^k, z\rangle \mid z \in \operatorname{co} \mathcal{Z}_p\}. \tag{21}$$

Let x^k be the solution of problem (20) and v^k be the solution of problem (21).

Step 2: Calculate $\Psi(u^k) = h(u^k) - d(u^k)$. If $\Psi(u^k) \le \Psi_{k-1}(u^k) + \varepsilon$ then stop; otherwise continue.

Step 3: Define

$$\Psi_k(u) = \max_{1\le j\le k} \big[\Psi(u^j) + \langle g(x^j) - v^j, u - u^j\rangle\big],$$

and find a solution u^{k+1} of the problem:

$$\min_{u\in U} \Psi_k(u).$$

Step 4: Increase k by one and go to Step 1.

A few comments are in order. Problem (20) is a convex nonlinear problem, and it can be solved by a suitable numerical method for nonlinear optimization. Problem (21) requires a dedicated numerical method. In particular applications, specialized methods may provide its efficient numerical solution. Alternatively, one can approximate the random vector Y by finitely many realizations (scenarios), and use the general method suggested in [12] for solving the approximate problem.

Theorem 4.1 *Suppose that $\varepsilon = 0$. Then the sequences $\Psi(u^k)$ and $\Psi_k(u^k)$, $k = 1, 2, \ldots$, converge to the optimal value of problem (8). Moreover, every accumulation point of the sequence $\{u^k\}$ is an optimal solution of (8).*

Proof: The convergence of the method follows from a standard argument about cutting plane methods for convex optimization (see, e.g., [4, Thm. 4.2.3]). □

It follows from the above theorem that for every $\varepsilon > 0$ the dual method has to stop after finitely many iterations at some step k for which

$$\Psi(u^k) - \varepsilon \le \Psi_{k-1}(u^k) \le \min_{u\in U} \Psi(u). \tag{22}$$

Let us define the set of active cutting planes at u^k:

$$J = \big\{ j \in \{1, \ldots, k-1\} : \Psi(u^j) + \langle g(x^j) - v^j, u^k - u^j\rangle = \Psi_{k-1}(u^k)\big\}.$$

The subdifferential of $\Psi_{k-1}(\cdot)$ has the form

$$\partial\Psi_{k-1}(u) = \Big\{ s \in \mathbb{R}^m : s = \sum_{j\in J} \alpha_j \big(g(x^j) - v^j\big),\ \sum_{j\in J} \alpha_j = 1,\ \alpha_j \geq 0,\ j \in J \Big\}.$$

Since u^k is a minimizer of $\Psi_{k-1}(\cdot)$, there must exist a subgradient s such that

$$s \in C(u^k).$$

Thus there exist nonnegative α_j totaling 1 such that

$$\sum_{j\in J} \alpha_j \big(g(x^j) - v^j\big) \in C(u^k). \tag{23}$$

By the definition of Ψ,

$$\Psi(u^j) = f(x^j) + \langle u^j, g(x^j)\rangle - \langle u^j, v^j\rangle.$$

Substituting this into the definition of the set J we obtain that

$$\Psi_{k-1}(u^k) = f(x^j) + \langle g(x^j) - v^j, u^k\rangle, \quad j \in J.$$

Multiplying both sides by α_j and summing up we conclude that

$$\Psi_{k-1}(u^k) = \sum_{j\in J} \alpha_j f(x^j) + \Big\langle \sum_{j\in J} \alpha_j \big(g(x^j) - v^j\big), u^k \Big\rangle.$$

This combined with (23) yields

$$\Psi_{k-1}(u^k) = \sum_{j\in J} \alpha_j f(x^j). \tag{24}$$

Define

$$\bar{x} = \sum_{j\in J} \alpha_j x^j, \quad \bar{z} = \sum_{j\in J} \alpha_j v^j.$$

Clearly, $\bar{x} \in \mathcal{D} \cap \operatorname{co} \mathcal{Z}_p$. Using the concavity of g and (23) we see that

$$g(\bar{x}) \geq \sum_{j\in J} \alpha_j g(x^j) \geq \sum_{j\in J} \alpha_j v^j = \bar{z}.$$

Thus the point $(\bar{x}, \bar{z})$ is feasible for the convex hull problem (10).

It follows from the concavity of f and (24) that

$$f(\bar{x}) \geq \sum_{j\in J} \alpha_j f(x^j) = \Psi_{k-1}(u^k).$$

By the stopping test (22),

$$f(\bar{x}) \geq \Psi(u^k) - \varepsilon. \tag{25}$$

Since the value of $\Psi(u)$ is an upper bound for the objective value at any feasible point (x, z) of the convex hull problem, we conclude that $(\bar{x}, \bar{z})$ is an ϵ-optimal solution of this problem.

The above construction can be carried out at every iteration k. In this way we obtain a certain sequence $(\bar{x}^k, \bar{v}^k)$, $k = 1, 2, \ldots$. Since the sequence $\{\bar{x}^k\}$ is contained in a compact set and each $(\bar{x}^k, \bar{z}^k)$ is feasible for the convex hull problem (10), the sequence $\{\bar{z}^k\}$ is included in a compact set as well. Thus the sequence $\{(\bar{x}^k, \bar{v}^k)\}$ has accumulation points. It follows from Theorem 4.1 and from (25) that every accumulation point of the sequence $\{(\bar{x}^k, \bar{v}^k)\}$ is a solution of the convex hull problem (10).

The algorithm presented in this section ia based on a cutting plane approximation of the entire dual functional. The primal method of the next section involves approximations of the functional $d(\cdot)$ only.

5 The primal method

The primal algorithm extends to general distributions our earlier idea for discrete distributions, presented in [1]. The method consists of an iterative generation of p-efficient points and the solution of a restriction of problem (1). The restriction is based on the disjunctive representation of $\operatorname{co} \mathcal{Z}_p$ by the p-efficient points generated so far.

We assume that we know a compact set B containing all p-efficient points v such that there exists $x \in \mathcal{D}$ satisfying $v \leq g(x)$. It may be just a box with the lower bound at l, the vector of p-efficient points of all marginal distributions of Y, and with the upper bound above the maxima of $g_i(x)$ over $x \in \mathcal{D}$, $i = 1, \ldots, m$. Such a box exists by the compactness of $\mathcal{D}$. We also use a stopping test parameter $\varepsilon > 0$.

We denote the simplex in $\mathbb{R}^k$ by S_k, i.e.,

$$S_k := \{\lambda \in \mathbb{R}^k : \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1\}.$$

The Primal Method

Step 0: Select a p-efficient point $v^1 \in B$ such that there exists $\tilde{x} \in \mathcal{D}$ satisfying $g(\tilde{x}) > v^1$. Set $J_1 = \{1\}$, $k = 1$.

Step 1: Solve the *master problem*

$$\max f(x) \tag{26}$$

$$g(x) \geq \sum_{j \in J_k} \lambda_j v^j, \tag{27}$$

$$x \in \mathcal{D}, \ \lambda \in S_k. \tag{28}$$

Let u^k be the vector of Lagrange multipliers associated with the constraint (27).

Step 2: Calculate $d_k(u^k) = \min_{j \in J_k} \langle u^k, v^j \rangle$.

Step 3: Find a p-efficient solution v^{k+1} of the subproblem:

$$\min_{z \in \mathcal{Z}_p \cap B} \langle u^k, z \rangle$$

and calculate $d(u^k) = \langle v^{k+1}, u^k \rangle$.

Step 4: If $d(u^k) \geq d_k(u^k) - \varepsilon$ then stop; otherwise set $J_{k+1} = J_k \cup \{k+1\}$, increase k by one, and go to Step 1.

The first p-efficient point v^1 can be found by solving the subproblem at Step 3 for some $u \geq 0$. All master problems will be solvable, if the first one is solvable, which is assumed at Step 0. Moreover, all master problems satisfy Slater's constraint qualification condition with the point $\tilde{x}$ and $\tilde{\lambda} = (1, 0, \ldots, 0)$. Therefore it is legitimate to assume at Step 1 that we obtain a vector of Lagrange multipliers associated with (27).

Theorem 5.1 *Let $\varepsilon = 0$. The sequence $\{f(x^k)\}$, $k = 1, 2, \ldots$ converges to the optimal value of the convex hull problem (10). Every accumulation point $\hat{x}$ of the sequence $\{x^k\}$ is an optimal solution of problem (10), with $z = g(\hat{x})$.*

Proof: We formulate the dual problem to the master problem (26)–(28).

The dual functional is defined as follows:

$$\Phi_k(u) = \sup \Big\{ f(x) + \langle u, g(x) - \sum_{j \in J_k} \lambda_j v^j \rangle : \ x \in \mathcal{D}, \ \lambda \in S_k \Big\} = h(u) - d_k(u),$$

where $h(u)$ is the same as in (6) and

$$d_k(u) = \inf_{\lambda \in S_k} \sum_{j \in J_k} \lambda_j \langle u, v^j \rangle. \tag{29}$$

It is clear that $d_k(u) = \min_{j \in J_k} \langle u, v^j \rangle \geq d(u)$, where $d(u)$ is as in (7). Thus the function $\Phi_k(u)$ is a lower bound of the dual functional $\Psi(u)$, i.e.,

$$\Phi_k(u^k) \leq \Psi(u^k).$$

We observe that the sequence $\{f(x^k)\}$ is monotonically increasing because the feasible set of problem (26)–(28) increases. By duality, the sequence $\{\Phi_k(u^k)\}$ is monotonically increasing as well.

For $\delta > 0$ consider the set K_δ of iteration numbers k for which

$$\Phi_k(u^k) + \delta \leq \Psi(u^k).$$

Suppose that $k \in K_\delta$. We obtain the following chain of inequalities for all $j \leq k$:

$$\delta \leq \Psi(u^k) - \Phi_k(u^k) = -d(u^k) + d_k(u^k) = - \min_{z \in \mathcal{Z}_p \cap B} \langle u^k, z \rangle + \min_{j \in J_k} \langle u^k, v^j \rangle$$
$$\leq \langle u^k, v^j - v^{k+1} \rangle \leq \|u^k\| \cdot \|v^j - v^{k+1}\|.$$

We shall show later that there exists $M > 0$ such that $\|u^k\| \le M$ for all k. Therefore

$$\|v^{k+1} - v^j\| \ge \delta/M \quad \text{for all} \quad k \in K_\delta \quad \text{and all} \quad j = 1, \dots, k.$$

It follows from the compactness of the set B that the set K_δ is finite for every $\delta > 0$. Thus, we can find a subsequence $\mathcal{K}$ such that

$$\Psi(u^k) - \Phi_k(u^k) \to 0, \quad k \in \mathcal{K}.$$

Since for all k

$$\Psi(u^k) \ge \min_{u \ge 0} \Psi(u) \ge \min_{u \ge 0} \Phi_k(u) = \Phi_k(u^k), \tag{30}$$

and the sequence $\{\Phi_k(u^k)\}$ is nondecreasing, we conclude that

$$\lim_{k \to \infty} \Phi_k(u^k) = \min_{u \ge 0} \Psi(u).$$

We also have $\Phi_k(u^k) = f(x^k)$ and thus the sequence $\{f(x^k)\}$ is convergent to the optimal value of the convex hull problem (10). Since $\{x^k\}$ is included in $\mathcal{D}$, it has accumulation points and every accumulation point $\hat{x}$ is a solution of (10), with $z = g(\hat{x})$.

It remains to show that the multipliers u^k are uniformly bounded. To this end observe that the Lagrangian

$$L_k(x, \lambda, u^k) = f(x) + \big\langle u^k, g(x) - \sum_{j=1}^{k} \lambda_j v^j \big\rangle$$

achieves its maximum in $\mathcal{D} \times S_k$ at x^k and some λ^k. The optimal value is equal to $f(x^k)$ and it is bounded above by the optimal value μ of the convex hull problem (10).

The point $\tilde{x}$ and $\tilde{\lambda} = (1, 0, \dots, 0)$ is in $\mathcal{D} \times S_k$. Therefore

$$L_k(\tilde{x}, \tilde{\lambda}, u^k) \le \mu.$$

It follows that

$$\langle u^k, g(\tilde{x}) - v^1 \rangle \le \mu - f(\tilde{x}).$$

Recall that $g(\tilde{x}) - v^1 > 0$. Therefore u^k is an element of the compact set

$$U = \{u \in \mathbb{R}^m : \langle u, g(\tilde{x}) - v^1 \rangle \le \mu - f(\tilde{x}), \ u \ge 0\}.$$

□

If we use $\varepsilon > 0$ at Step 4, then relations (30) guarantee that the current solution x^k is ε-optimal for the convex hull problem (10).

Under the assumption that the distribution function of the random vector Y is r-concave for some $r \in \overline{\mathbb{R}}$ (see [8, 10] and the references therein), the suggested algorithms provide an optimal solution of problem (1). Otherwise, we obtain an upper bound of the optimal value. Moreover, the solution point $\hat{x}$ determined by both algorithms satisfies the constraint $g(x) \in \operatorname{co} \mathcal{Z}_p$, and may not satisfy the probabilistic constraint. We now suggest an approach to determine a primal feasible solution.

Both algorithms end with a collection of p-efficient points. In the primal algorithm, we consider the multipliers λ_j of the master problem (26)-(28). We define $A = \{j \in J : \lambda_j > 0\}$. In the dual algorithm, we consider the active cutting planes in the last approximation, and set $A = \{j \in J : \beta_j > 0\}$, where J and β_j are determined in the proof of Theorem 4.1.

In both cases, if A contains only one element, the point $\hat{x}$ is feasible and therefore optimal for the disjunctive formulation (4). If, however, there are more elements in A, we need to find a feasible point. A natural possibility is to consider the *restricted* disjunctive formulation:

$$\begin{aligned} &\max\ f(x) \\ &\text{subject to } g(x) \in \bigcup_{j \in A} K_j, \\ &x \in \mathcal{D}. \end{aligned} \tag{31}$$

It can be solved by simple enumeration of all cases for $j \in A$:

$$\begin{aligned} &\max\ f(x) \\ &\text{subject to } g(x) \geq v^j, \\ &x \in \mathcal{D}. \end{aligned} \tag{32}$$

An alternative strategy would be to solve the corresponding bounding problem (32) every time a new p-efficient point is generated. If L_j denotes the optimal value of (32), the lower bound at iteration k is

$$\bar{L}^k = \max_{0 \leq j \leq k} L_j. \tag{33}$$

6 Numerical Illustration

To illustrate the operation of the two methods presented in the paper we consider the following cash matching problem. We have random liabilities L_t in periods $t = 0, 1, \ldots, T$ and a basket of n bonds. The payment of bond j in period t is denoted by a_{jt}. It is zero for t before the purchase of the bond and for t greater than the maturity time of the bond. At the time of purchase a_{jt} is the negative of the price of the bond, at the following periods it is equal to the coupon payment, and at the time of maturity it is equal to the face value plus the coupon payment. Out initial capital equals C.

The objective is to design a bond portfolio such that the probability of covering the liabilities over the entire period $0, 1, \ldots, T$ is at least p. Subject to this condition, we want to maximize the final cash on hand, guaranteed with probability p.

Let us introduce the cumulative liabilities

$$Y_t = \sum_{\tau=0}^{t} L_\tau, \quad t = 0, \ldots, T.$$

Denoting by x_j the amount invested in bond j, $j = 1, \dots, n$, we observe that the cumulative cash flows up to time t can be expressed as follows:

$$z_0 = C + \sum_{j=1}^{n} a_{j0} x_j,$$

$$z_t = z_{t-1} + \sum_{j=1}^{n} a_{jt} x_j, \quad t = 1, \dots, T.$$

The problem takes on the form

$$\begin{aligned} &\max\ z_T \\ &\text{subject to } \mathbb{P}\big[z_t \geq Y_t,\ t = 0, \dots, T\big] \geq p, \\ &\qquad z_0 = C + \sum_{j=1}^{n} a_{j0} x_j, \\ &\qquad z_t = z_{t-1} + \sum_{j=1}^{n} a_{jt} x_j, \quad t = 1, \dots, T, \\ &\qquad x \geq 0. \end{aligned}$$

Let us observe that expressing the probabilistic liquidity constraint in terms of cumulative cash flows and cumulative liabilities allows us to obtain a constraint with a random right hand side only. Since the vector Y has a joint normal distribution, which is log-concave, the resulting problem is convex. Thus both methods described in this paper yield optimal solutions of the problem.

We have used data on 72 government bonds and AAA corporate bonds ranging from 6-month treasury bills (which do not pay coupons, but sell at discount) to 5-year bonds paying coupons each 6-months. The liabilities were assumed to be normally distributed with expectation \$2,000,000 and standard deviation \$100,000. The initial capital was $C = 20,000,000$ and the number of 6-month periods $T = 10$. The probability $p = 0.95$. To facilitate the numerical solution of the method, the distribution of the liabilities was approximated by $N = 100$ equally likely scenarios.

Both methods were implemented in the AMPL modeling language. The search for new p-efficient points in both methods (problem (21) in the dual, and Step 3 in the primal) was implemented as a simple binary optimization problem with a knapsack constraint. Other subproblems were solved by the CPLEX linear programming solver.

The dual method terminated after 34 iterations finding the optimal portfolio of 9 bonds of different maturities. The primal method found exactly the same solution after just 3 iterations. In both cases the computation time on a 1.7GHz PC was less than a minute.

References

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

[2] D. Dentcheva, A. Prékopa and A. Ruszczyński, Bounds for integer stochastic programs with probabilistic constraints, *Discrete Applied Mathematics* 124 (2002) 55-65.

[3] K. Dowd, *Beyond Value at Risk. The Science of Risk Management.* Wiley, New York, 1997.

[4] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I and II*, Springer-Verlag, Berlin 1993.

[5] W. Ogryczak and A. Ruszczyński, On consistency of stochastic dominance and mean–semideviation models, *Mathematical Programming* 89 (2001) 217–232.

[6] W. Ogryczak and A. Ruszczyński, Dual stochastic dominance and related mean–risk models, *SIAM Journal on Optimization* 13 (2002) 60–78.

[7] R.R. Phelps, Convex functions, monotone operators, and differentiability, Lecture Notes in Mathematics 1364, Springer-Verlag, 1989.

[8] A. Prékopa, *Stochastic Programming*, Kluwer, Dordrecht, Boston, 1995.

[9] A. Prékopa, B. Vizvári, and T. Badics, Programming Under Probabilistic Constraint with Discrete Random Variable, in: *New Trends in Mathematical Programming*, L. Grandinetti et al. (Ed.), Kluwer, Dordrecht, Boston, 1998, pp. 235–255.

[10] Y. Rinott, On Convexity of Measures, *Annals of Probability* 4 (1976) 1020–1026.

[11] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

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

[13] A. Ruszczyński and A. Shapiro, *Stochastic Programming*, Elsevier, to appear in 2003.

[14] S. Uryasev and R.T. Rockafellar, Conditional value-at-risk: optimization approach. in *Stochastic optimization: algorithms and applications* (Gainesville, FL, 2000), 411–435, Appl. Optim., 54, Kluwer Academic Publishers, Dordrecht, 2001.