Relaxations of linear programming problems with first order stochastic dominance constraints

Nilay Noyan[a,*], Gábor Rudolf[a], Andrzej Ruszczyński[b]

[a]*RUTCOR, Rutgers University, 640 Bartholomew Rd., Piscataway, NJ 08854, USA*
[b]*Department of Management Science and Information Systems, Rutgers University, 640 Bartholomew Rd., Piscataway, NJ 08854, USA*

Abstract

Linear stochastic programming problems with first order stochastic dominance (FSD) constraints are non-convex. For their mixed 0–1 linear programming formulation we present two convex relaxations based on second order stochastic dominance (SSD). We develop necessary and sufficient conditions for FSD, used to obtain a disjunctive programming formulation and to strengthen one of the SSD-based relaxations.
© 2005 Published by Elsevier B.V.

1. Introduction

The notion of first order stochastic dominance (FSD), also called stochastic ordering, has been introduced in statistics (see [10,11]) and further applied and developed in economics (see [8,9,7,13,14]). A random variable X *dominates* another random variable Y in the first order, which we denote by $X \succcurlyeq_{(1)} Y$, if

$$F_X(\eta) \leqslant F_Y(\eta) \quad \text{for all } \eta \in \mathbb{R}, \tag{1}$$

where $F_X(\eta) = P[X \leqslant \eta]$ denotes the distribution function of a random variable X. It is

well-known (see, e.g., [12]) that the stochastic ordering relation $X \succcurlyeq_{(1)} Y$ can be equivalently expressed as follows: $\mathbb{E}[u(X)] \geqslant \mathbb{E}[u(Y)]$ for all non-decreasing functions $u : \mathbb{R} \to \mathbb{R}$ for which the above expectations are finite.

For two integrable random variables X and Y, $X \succcurlyeq_{(2)} Y$ denotes the second order stochastic dominance (SSD) relation:

$$\int_{-\infty}^{\eta} F_X(\alpha) \, d\alpha \leqslant \int_{-\infty}^{\eta} F_Y(\alpha) \, d\alpha \quad \text{for all } \eta \in \mathbb{R}. \tag{2}$$

It is easy to see that the FSD relation (1) implies SSD via integration over the interval $(-\infty, \eta)$. We are also going to use the following equivalent formulation of the SSD condition, featuring the expected shortfalls of

* Corresponding author.
E-mail address: noyan@rutcor.rutgers.edu (N. Noyan).

the random variables (see [4]):

$$\mathbb{E}([\eta - X]_+) \leqslant \mathbb{E}([\eta - Y]_+) \quad \text{for all } \eta \in \mathbb{R}. \qquad (3)$$

Dominance relations can be involved in stochastic optimization problems as constraints, allowing us to obtain solutions dominating some random reference outcomes. Such models have been introduced and analyzed in [4,6]. In [5] it is also proved that the convexification of the set defined by the FSD relation is equal to the set defined by the SSD relation, if the probability space is non-atomic.

In this paper we analyze FSD constraints, which are equivalent to a continuum of probabilistic constraints. Here the main challenge lies in the potential non-convexity of the associated feasible region.

In Section 2 we formulate a linear FSD-constrained optimization model as a mixed 0–1 programming problem with multiple knapsack constraints. In Section 3 a linear programming (LP) relaxation, based on SSD constraints, is introduced, which is different from the usual LP relaxation obtained by dropping the integrality restriction. In Section 4 we introduce a new necessary condition for FSD leading to yet another LP relaxation, which we call *interval second order stochastic dominance* (ISSD) relaxation. Section 5 provides an analysis of the relations between these relaxations. We conclude the paper by presenting in Section 6 a new necessary and sufficient condition for FSD. Using this condition we obtain a disjunctive programming formulation of the discrete FSD-constrained optimization model.

The expected value operator is denoted by $\mathbb{E}$. An abstract probability space is denoted by $(\Omega, \mathscr{F}, P)$, where Ω is the sample space, $\mathscr{F}$ is a σ-algebra on Ω and P is a probability measure on Ω. For a real number $\eta \in \mathbb{R}$ let $[\eta]_+ = \max(0, \eta)$ and $[\eta]_- = \min(0, \eta)$.

2. Linear optimization with first order constraints

Let $(\Omega, 2^\Omega, P)$ be a finite probability space, where $\Omega = \{\omega_1, \ldots, \omega_N\}$, with corresponding probabilities $p_1, \ldots, p_N$, and $\psi : \mathbb{R}^n \times \Omega \to \mathbb{R}$ an outcome mapping satisfying the condition that $\psi(\cdot, \omega)$ is a linear function for all $\omega \in \Omega$. For a given vector $\mathbf{z} \in \mathbb{R}^n$ let us define the mapping $\psi_{\mathbf{z}} : \Omega \to \mathbb{R}$ by $\psi_{\mathbf{z}}(\omega) = \psi(\mathbf{z}, \omega)$. Let Y be some random variable on Ω (*reference outcome*). We consider the following linear stochastic optimization model with an FSD constraint:

$$\begin{aligned}
\max \quad & \mathbf{c}^T \mathbf{z} \\
\text{s.t.} \quad & \psi_{\mathbf{z}} \succcurlyeq_{(1)} Y, \\
& \mathbf{z} \in Z,
\end{aligned} \qquad (4)$$

where Z is a compact polyhedron. Without the dominance constraint, problem (4) can be easily formulated as a linear programming problem. The stochastic dominance constraint, however, can render the feasible set non-convex in the general case.

Let $\hat{y}_i = Y(\omega_i)$, $i = 1, \ldots, N$, denote the realizations of the reference outcome Y. Without loss of generality we may assume $\hat{y}_1 \leqslant \hat{y}_2 \leqslant \cdots \leqslant \hat{y}_N$. Let us denote the different realizations of the reference outcome Y by $y_1 < y_2 < \cdots < y_D$, and let y_0 be an arbitrary number such that $y_0 < y_1$, implying $P[Y \leqslant y_0] = 0$.

Since the distribution function of the reference outcome Y is a right-continuous step function, it is easy to verify that the FSD constraint in (4) is equivalent to

$$P[\psi_{\mathbf{z}} < y_k] \leqslant P[Y \leqslant y_{k-1}], \quad k = 1, \ldots, D. \qquad (5)$$

We now reformulate problem (4) by introducing the binary variables

$$\beta_{i,k} = \begin{cases} 1 & \text{if } \psi(\mathbf{z}, \omega_i) < y_k, \\ 0 & \text{otherwise,} \end{cases}$$
$$i = 1, \ldots, N, \quad k = 1, \ldots, D. \qquad (6)$$

Let $M_{i,k} \in \mathbb{R}$ satisfy the inequalities

$$M_{i,k} \geqslant y_k - \min_{\mathbf{z} \in Z} \psi(\mathbf{z}, \omega_i),$$
$$i = 1, \ldots, N, \quad k = 1, \ldots, D. \qquad (7)$$

Since $\psi(\cdot, \omega_i)$ is linear for all $i = 1, \ldots, N$, and Z is compact, such $M_{i,k}$ values exist.

Now using (5)–(7) we can rewrite (4) as a mixed 0–1 linear programming problem, which we refer to as MBLP:

$$\begin{aligned}
\max \quad & \mathbf{c}^T \mathbf{z} \\
\text{s.t.} \quad & \sum_{i=1}^{N} p_i \beta_{i,k} \leqslant P[Y \leqslant y_{k-1}], \quad k = 1, \ldots, D, \\
& \psi(\mathbf{z}, \omega_i) + M_{i,k} \beta_{i,k} \geqslant y_k, \qquad (8)
\end{aligned}$$

$$i = 1, \ldots, N, \quad k = 1, \ldots, D, \qquad (9)$$
$$\beta_{i,k} \in \{0, 1\}, \quad i = 1, \ldots, N,$$
$$k = 1, \ldots, D, \qquad (10)$$
$$\mathbf{z} \in Z.$$

Note that we can select a sufficiently large common value $M_{i,k} = M$, $i = 1, \ldots, N$, $k = 1, \ldots, D$.

Inequalities (8) together with the integrality restriction (10) define a set of knapsack constraints. With every item $i \in \{1, \ldots, N\}$ we associate a weight $p_i > 0$, and every knapsack $k \in \{1, \ldots, D\}$ has a capacity $u_k = P[Y \leqslant y_{k-1}]$, $k = 1, \ldots, D$.

Our intention is to employ the probabilistic structure of the problem to derive new LP relaxations.

3. Relaxation based on second order constraints

Since FSD implies SSD and, as we are going to see below, SSD constraints can be represented by a system of linear inequalities, the SSD constraints can be viewed as valid inequalities for the feasible set of MBLP. This fact leads to an LP relaxation which is different from the usual one obtained by relaxing the integrality restriction. In Section 5 we provide examples which show that neither of these relaxations is stronger than the other.

Since Y has a discrete distribution, inequalities (3) are equivalent (see [4]) to

$$\mathbb{E}([y_k - X]_+) \leqslant \mathbb{E}([y_k - Y]_+), \quad k = 1, \ldots, D. \quad (11)$$

Let us introduce the following functions, representing the shortfall values of the outcome $\psi_{\mathbf{z}}$:

$$s_{\mathbf{z}}(i, k) = \begin{cases} 0 & \text{for } i = 1, \ldots, N, \\ & \quad k = 0, \\ [y_k - \psi(\mathbf{z}, \omega_i)]_+ & \text{for } i = 1, \ldots, N, \\ & \quad k = 1, \ldots, D. \end{cases}$$
$$(12)$$

When the policy $\mathbf{z}$ is clear from the context, we also use the simplified notation $s_{i,k} = s_{\mathbf{z}}(i, k)$. Then the next system of linear inequalities is equivalent to the SSD relation (11):

$$\sum_{i=1}^{N} p_i s_{i,k} \leqslant \mathbb{E}([y_k - Y]_+), \quad k = 1, \ldots, D. \quad (13)$$

Let $v_k = \mathbb{E}([y_k - Y]_+)$. By introducing the decision vector $\bar{\mathbf{s}} \in \mathbb{R}^{N(D+1)}$ to represent the shortfall values

we obtain the corresponding LP relaxation

$$\max \quad \mathbf{c}^{\mathrm{T}} \mathbf{z} \quad (14)$$

$$\text{s.t.} \quad \psi(\mathbf{z}, \omega_i) + \bar{s}_{i,k} \geqslant y_k,$$
$$i = 1, \ldots, N, \quad k = 1, \ldots, D, \quad (15)$$

$$\sum_{i=1}^{N} p_i \bar{s}_{i,k} \leqslant v_k, \quad k = 1, \ldots, D, \quad (16)$$

$$\bar{s}_{i,k} \geqslant 0, \quad i = 1, \ldots, N, \quad k = 1, \ldots, D, \quad (17)$$

$$\mathbf{z} \in Z. \quad (18)$$

We refer to the above problem as the SSD-based LP relaxation of MBLP.

Observation 1. *For every feasible solution* $(\mathbf{z}, \bar{\mathbf{s}})$ *of* (14)–(18) *we have* $\bar{\mathbf{s}} \geqslant \mathbf{s}_{\mathbf{z}}$, *and the pair* $(\mathbf{z}, \mathbf{s}_{\mathbf{z}})$ *is also feasible for* (14)–(18).

4. Relaxation based on interval second order constraints

The results of this section are based on the following condition on differences of expected shortfalls.

Proposition 1. *If* $X \succcurlyeq_{(1)} Y$ *then for all* $\eta_1 \leqslant \eta_2$ *in* $\mathbb{R}$,

$$\mathbb{E}([\eta_2 - X]_+) - \mathbb{E}([\eta_1 - X]_+)$$
$$\leqslant \mathbb{E}([\eta_2 - Y]_+) - \mathbb{E}([\eta_1 - Y]_+). \quad (19)$$

Proof. Similarly to the case of (2) and (3), we can rewrite (19) in the form

$$\int_{\eta_1}^{\eta_2} F_X(t) \, \mathrm{d}t \leqslant \int_{\eta_1}^{\eta_2} F_Y(t) \, \mathrm{d}t, \quad (20)$$

which immediately follows from (1) via integration on the interval $[\eta_1, \eta_2]$. $\quad \square$

Using the above result we can obtain a strengthening of the SSD-based LP relaxation by substituting $X = \psi_{\mathbf{z}}$, $\eta_1 = y_{k-1}$, $\eta_2 = y_k$ into (19).

Corollary 1. *If* $\psi_{\mathbf{z}} \succcurlyeq_{(1)} Y$, *then*

$$\mathbb{E}([y_k - \psi_{\mathbf{z}}]_+) - \mathbb{E}([y_{k-1} - \psi_{\mathbf{z}}]_+)$$

$$\leqslant \mathbb{E}([y_k - Y]_+) - \mathbb{E}([y_{k-1} - Y]_+),$$

$$k = 1, \dots, D. \tag{21}$$

In the discrete case condition (21) takes the form

$$\sum_{i=1}^{N} p_i s_{i,k} - \sum_{i=1}^{N} p_i s_{i,k-1} \leqslant v_k - v_{k-1},$$

$$k = 1, \dots, D. \tag{22}$$

By Observation 1 we obtain an LP relaxation of MBLP, based on the necessary condition (22):

$$\max \quad \mathbf{c}^{\mathsf{T}} \mathbf{z}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} p_i \bar{s}_{i,k} - \sum_{i=1}^{N} p_i \bar{s}_{i,k-1} \leqslant v_k - v_{k-1},$$

$$k = 1, \dots, D,$$

$$\psi(\mathbf{z}, \omega_i) + \bar{s}_{i,k} \geqslant y_k,$$

$$i = 1, \dots, N, \quad k = 1, \dots, D,$$

$$\bar{s}_{i,k} \geqslant 0, \quad i = 1, \dots, N, \quad k = 1, \dots, D,$$

$$\bar{s}_{i,0} = 0, \quad i = 1, \dots, N,$$

$$\mathbf{z} \in Z. \tag{23}$$

The next lemma provides an upper bound for the shortfalls, allowing us to further strengthen this relaxation. Let $\mathbf{e}_j = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$, with the 1 in the jth position.

Lemma 1. *For a policy* $\mathbf{z} \in Z$ *the corresponding shortfall vector* $\mathbf{s} = \mathbf{s}_{\mathbf{z}}$ *satisfies*

$$\mu s_{i,k} \leqslant \sum_{j=1}^{n} (z_j + m_j)[y_k - \mu \psi(\mathbf{e}_j, \omega_i)]_+$$

$$+ \mu m_j [\psi(\mathbf{e}_j, \omega_i)]_+,$$

$$i = 1, \dots, N, \quad k = 1, \dots, D,$$

where $m_j = -[\min\{z_j \mid \mathbf{z} \in Z\}]_-$ *and* $\mu = \min\{\sum_{j=1}^{n}(z_j + m_j) \mid \mathbf{z} \in Z\}$.

Proof. The result follows from the chain of inequalities

$$\mu s_{i,k} = \mu \left[y_k - \sum_{j=1}^{n} z_j \psi(\mathbf{e}_j, \omega_i) \right]_+$$

$$= \left[\mu y_k - \mu \sum_{j=1}^{n} (z_j + m_j) \psi(\mathbf{e}_j, \omega_i) \right.$$

$$\left. + m_j \psi(\mathbf{e}_j, \omega_i) \right]_+$$

$$\leqslant \left[\sum_{j=1}^{n} (z_j + m_j) y_k - \sum_{j=1}^{n} (z_j + m_j) \mu \psi(\mathbf{e}_j, \omega_i) \right.$$

$$\left. + \mu m_j \psi(\mathbf{e}_j, \omega_i) \right]_+$$

$$\leqslant \left[\sum_{j=1}^{n} (z_j + m_j)(y_k - \mu \psi(\mathbf{e}_j, \omega_i)) \right.$$

$$\left. + \mu m_j \psi(\mathbf{e}_j, \omega_i) \right]_+$$

$$\leqslant \sum_{j=1}^{n} (z_j + m_j)[y_k - \mu \psi(\mathbf{e}_j, \omega_i)]_+$$

$$+ \mu m_j [\psi(\mathbf{e}_j, \omega_i)]_+.$$

The last inequality relies on the fact that $z_j + m_j \geqslant 0$ and $\mu m_j \geqslant 0$ for $j = 1, \dots, n$. $\quad \square$

We can now strengthen the LP relaxation (23) by adding to it the valid inequalities

$$\mu \bar{s}_{i,k} \leqslant \sum_{j=1}^{n} (z_j + m_j)[y_k - \mu \psi(\mathbf{e}_j, \omega_i)]_+$$

$$+ m_j \mu [\psi(\mathbf{e}_j, \omega_i)]_+,$$

$$i = 1, \dots, N, \quad k = 1, \dots, D. \tag{24}$$

We will refer to this new formulation as the *interval second order stochastic dominance* (ISSD) relaxation.

5. Comparison of the three LP relaxations

The following theorem compares the strength of the various LP relaxations of MBLP introduced so far. For a set $W \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ let us introduce the notation $\text{Proj}(W) = \{\mathbf{z} \in \mathbb{R}^{n_1} : \exists \mathbf{v} \in \mathbb{R}^{n_2} \, (\mathbf{z}, \mathbf{v}) \in W)\}$.

Theorem 1. *Let us denote the feasible regions of the SSD-based LP relaxation, the ISSD relaxation and the usual LP relaxation of MBLP by* Q^{SSD}, Q^{ISSD} *and* Q, *respectively. Then the following relations hold*:

1. $\text{Proj}(Q^{\text{ISSD}}) \quad \subseteq \quad \text{Proj}(Q^{\text{SSD}})$, *in general* $\text{Proj}(Q^{\text{ISSD}}) \subsetneq \text{Proj}(Q^{\text{SSD}})$.
2. *In general* $\text{Proj}(Q^{\text{ISSD}}) \not\subseteq \text{Proj}(Q)$.
3. *In general* $\text{Proj}(Q) \not\subseteq \text{Proj}(Q^{\text{SSD}})$.

Proof. Let $Z = \{\mathbf{z} \in \mathbb{R}_+^2 : z_1 + z_2 = 1\}$, $\bar{\mathbf{z}} = (0, 1)$, $\hat{\mathbf{z}} = (1, 0)$, $\widetilde{\mathbf{z}} = (\frac{1}{2}, \frac{1}{2})$, and consider the probability spaces on ground sets $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\Omega' = \{\omega_1', \omega_2'\}$, with $P[\omega_1] = P[\omega_2] = P[\omega_3] = \frac{1}{3}$ and $P[\omega_1'] = P[\omega_2'] = \frac{1}{2}$.

1. Summing the interval dominance constraints in (23) over the index ranges $\{1, \dots, k\}, k = 1, \dots, D$, we obtain the SSD constraints, therefore $\text{Proj}(Q^{\text{ISSD}}) \subseteq \text{Proj}(Q^{\text{SSD}})$. Now consider on Ω the outcome function

$$\psi^{(1)}(\mathbf{z}, \omega_1) = 3z_1,$$
$$\psi^{(1)}(\mathbf{z}, \omega_2) = 4z_1 + 4z_2,$$
$$\psi^{(1)}(\mathbf{z}, \omega_3) = 5z_1 + 6z_2,$$

the reference outcome $Y = \psi_{\bar{\mathbf{z}}}^{(1)}$ and the decision vector (policy) $\hat{\mathbf{z}}$. It is easy to see that $\psi_{\hat{\mathbf{z}}}^{(1)}$ dominates Y in the second order, implying $\hat{\mathbf{z}} \in \text{Proj}(Q^{\text{SSD}})$. On the other hand, for any $(\hat{\mathbf{z}}, \bar{\mathbf{s}})$ feasible for ISSD, inequalities (24) imply $\bar{s}_{1,2} \leqslant 1$, $\bar{s}_{2,2} \leqslant 0$, $\bar{s}_{3,2} \leqslant 0$, and so $\sum_{i=1}^N p_i \bar{s}_{i,2} \leqslant \frac{1}{3}$. Since $\sum_{i=1}^N p_i \bar{s}_{i,3} \geqslant \sum_{i=1}^N p_i s_{i,3} = 2$, we have $\sum_{i=1}^N p_i \bar{s}_{i,3} - \sum_{i=1}^N p_i \bar{s}_{i,2} \geqslant \frac{5}{3} > \frac{4}{3} = v_3 - v_2$. Therefore, the constraint $\sum_{i=1}^N p_i \bar{s}_{i,k} - \sum_{i=1}^N p_i \bar{s}_{i,k-1} \leqslant v_k - v_{k-1}$ in (23) is violated for $k = 3$, which shows that $\hat{\mathbf{z}} \notin \text{Proj}(Q^{\text{ISSD}})$.

2. Consider on Ω the outcome function

$$\psi^{(2)}(\mathbf{z}, \omega_1) = 2z_1,$$
$$\psi^{(2)}(\mathbf{z}, \omega_2) = 5z_1 + 4z_2,$$
$$\psi^{(2)}(\mathbf{z}, \omega_3) = 5z_1 + 6z_2,$$

the reference outcome $Y = \psi_{\bar{\mathbf{z}}}^{(2)}$ and the decision vector (policy) $\hat{\mathbf{z}}$. It is easy to see that $\hat{\mathbf{z}} \in \text{Proj}(Q^{\text{ISSD}})$. On the other hand, the strongest formulation of the usual LP relaxation of MBLP is obtained by setting $M_{i,k} = y_k - \min_{\mathbf{z} \in Z} \psi(\mathbf{z}, \omega_i)$, yielding in our case $M_{1,3} = 6$, $M_{2,3} = 2$, $M_{3,3} = 1$. It follows from (9) that $\beta_{1,3} \geqslant \frac{2}{3}$, $\beta_{2,3} \geqslant \frac{1}{2}$ and $\beta_{3,3} \geqslant 1$ hold for any $(\hat{\mathbf{z}}, \beta) \in Q$, implying that inequality (8) cannot be satisfied for the decision vector $\hat{\mathbf{z}}$, thus $\hat{\mathbf{z}} \notin \text{Proj}(Q)$.

3. Consider on Ω' the outcome function

$$\psi^{(3)}(\mathbf{z}, \omega_1) = 10z_1 + 6z_2,$$
$$\psi^{(3)}(\mathbf{z}, \omega_2) = 4z_1 + 10z_2,$$

the reference outcome $Y = \psi_{\bar{\mathbf{z}}}^{(3)}$ and the decision vector (policy) $\widetilde{\mathbf{z}}$. It is easy to see that for the values $\hat{\beta}_{1,1} = \hat{\beta}_{2,1} = 0$, $\hat{\beta}_{1,2} = \hat{\beta}_{2,2} = \frac{1}{2}$ we have $(\hat{\beta}, \widetilde{\mathbf{z}}) \in Q$ for any $M_{i,k}$ values satisfying (7), while $\widetilde{\mathbf{z}} \notin \text{Proj}(Q^{\text{SSD}})$, as $\psi_{\widetilde{\mathbf{z}}}$ does not dominate Y in the second order. $\square$

6. A disjunctive programming formulation

We are going to show that the necessary condition for FSD presented in Proposition 1 is also sufficient, even when restricted to a finite set of pairs η_1, η_2. For a random variable X we write $x_i = X(\omega_i)$.

Theorem 2. $X \succeq_{(1)} Y$ *if and only if*

$$\mathbb{E}([y_k - X]_+) - \mathbb{E}([x_i - X]_+)$$
$$\leqslant \mathbb{E}([y_k - Y]_+) - \mathbb{E}([x_i - Y]_+) \qquad (25)$$

for all $i \in \{1, \dots, N\}$ *and* $k \in \{0, \dots, D\}$ *such that*

$$x_i < y_k \quad \text{and}$$
$$\{y_1, \dots, y_D, x_1, \dots, x_N\} \cap (x_i, y_k) = \emptyset, \qquad (26)$$

where (x_i, y_k) *is the open interval with endpoints* x_i *and* y_k.

Proof. The necessity of the condition follows immediately from Proposition 1. Now suppose that X does not dominate Y in the first order. Then there exists some $t^* \in \mathbb{R}$ for which $F_Y(t^*) < F_X(t^*)$. Recall that both functions $F_X(\cdot)$ and $F_Y(\cdot)$ are piecewise constant. Let (a, b) be the largest interval containing t^* and such

that $F_Y(t) < F_X(t)$ for $t \in (a, b)$:

$$a = \inf\{\tau : F_Y(t) < F_X(t), \ t \in [\tau, t^*]\},$$
$$b = \sup\{\tau : F_Y(t) < F_X(t), \ t \in [t^*, \tau]\}.$$

Both a and b are finite, because both functions are equal to 0 for sufficiently small t and equal to 1 for sufficiently large t. Since both functions F_X and F_Y are right continuous, $b > t^*$.

It follows that F_X has a jump at a and F_Y has a jump at b. Now we find a jump point $x_{i*} \in [a, b)$ of F_X and a jump point $y_{k*} \in (a, b]$ of F_Y such that none of these functions has jumps in the open interval (x_{i*}, y_{k*}). Such a pair exists: we can take x_{i*} to be the last jump point of F_X in the interval $[a, b)$ and y_{k*} to be the first jump point of F_Y in the interval $(x_{i*}, b]$.

The pair (i^*, k^*) satisfies (26) by construction. To complete the proof we observe that the pair (i^*, k^*) violates (25), which, similarly to (20), can be written as

$$\int_{x_{i*}}^{y_{k*}} F_X(t) \, dt \leqslant \int_{x_{i*}}^{y_{k*}} F_Y(t) \, dt. \qquad \square$$

In the following we are going to use for $X = \psi_{\mathbf{z}}$ a weaker form of Theorem 2.

Corollary 2. $\psi_{\mathbf{z}} \succcurlyeq_{(1)} Y$ *if and only if*

$$\mathbb{E}([y_k - \psi_{\mathbf{z}}]_+) - \mathbb{E}([\psi(\mathbf{z}, \omega_i) - \psi_{\mathbf{z}}]_+)$$
$$\leqslant \mathbb{E}([y_k - Y]_+) - \mathbb{E}([\psi(\mathbf{z}, \omega_i) - Y]_+)$$

for all $i \in \{1, \ldots, N\}$ *and* $k \in \{0, \ldots, D\}$ *such that* $\psi(\mathbf{z}, \omega_i) < y_k$.

Let us introduce the following decision vectors, representing the above shortfalls for a policy $\mathbf{z} \in Z$:

$$\xi_{i,k} = [\psi(\mathbf{z}, \omega_k) - \psi(\mathbf{z}, \omega_i)]_+,$$
$$i = 1, \ldots, N, \ k = 1, \ldots, N,$$
$$\chi_{i,k} = [\psi(\mathbf{z}, \omega_k) - y_i]_+,$$
$$i = 0, \ldots, D, \ k = 1, \ldots, N.$$

We conclude by presenting a disjunctive programming formulation of our problem. For an introduction to disjunctive programming (optimization under logical constraints involving linear inequalities) see [1,2].

Since the condition $y = \max(a, b)$ can be written in the disjunctive form

$$\begin{pmatrix} y = a \\ a \geqslant b \end{pmatrix} \vee \begin{pmatrix} y = b \\ b \geqslant a \end{pmatrix},$$

by Corollary 2 the FSD-constrained problem (4) can be formulated as a disjunctive programming problem:

$$\max \quad \mathbf{c}^{\mathrm{T}} \mathbf{z}$$

$$\text{s.t.} \quad (y_k \leqslant \psi(\mathbf{z}, \omega_i)) \vee \left(\sum_{j=1}^{N} p_j s_{j,k} - \sum_{j=1}^{N} p_j \xi_{j,i} \right.$$

$$\left. \leqslant v_k - \sum_{j=1}^{D} \sum_{l:\hat{y}_l = y_j} p_l \chi_{j,i} \right),$$

$$i = 1, \ldots, N, \ k = 0, \ldots, D,$$

$$\begin{pmatrix} s_{i,k} = y_k - \psi(\mathbf{z}, \omega_i) \\ y_k - \psi(\mathbf{z}, \omega_i) \geqslant 0 \end{pmatrix}$$
$$\vee \begin{pmatrix} s_{i,k} = 0 \\ y_k - \psi(\mathbf{z}, \omega_i) \leqslant 0 \end{pmatrix},$$
$$i = 1, \ldots, N, \ k = 0, \ldots, D,$$

$$\begin{pmatrix} \xi_{i,k} = \psi(\mathbf{z}, \omega_k) - \psi(\mathbf{z}, \omega_i) \\ \psi(\mathbf{z}, \omega_k) - \psi(\mathbf{z}, \omega_i) \geqslant 0 \end{pmatrix}$$
$$\vee \begin{pmatrix} \xi_{i,k} = 0 \\ \psi(\mathbf{z}, \omega_k) - \psi(\mathbf{z}, \omega_i) \leqslant 0 \end{pmatrix},$$
$$i = 1, \ldots, N, \ k = 1, \ldots, N,$$

$$\begin{pmatrix} \chi_{i,k} = \psi(\mathbf{z}, \omega_k) - y_i \\ \psi(\mathbf{z}, \omega_k) - y_i \geqslant 0 \end{pmatrix}$$
$$\vee \begin{pmatrix} \chi_{i,k} = 0 \\ \psi(\mathbf{z}, \omega_k) - y_i \leqslant 0 \end{pmatrix},$$
$$i = 0, \ldots, D, \quad k = 1, \ldots, N.$$

In an upcoming paper we shall obtain other disjunctions by restating the inequalities arising from the underlying combinatorial structure of the problem as logical conditions. This will lead to the generation of valid inequalities for the corresponding disjunctive relaxations, by applying the lift-and-project procedure developed in [3].

Acknowledgements

References

[1] E. Balas, Disjunctive programming: properties of the convex hull of feasible points, Technical Report MSRR 348, Carnegie Mellon University, 1974.

[2] E. Balas, Disjunctive programming, Ann. Discrete Math. 5 (1979) 3–51.

[3] E. Balas, S. Ceria, G. Cornuèjols, A lift-and-project cutting plane algorithm for mixed 0–1 programs, Math. Programming 58 (1993) 295–324.

[4] D. Dentcheva, A. Ruszczyński, Optimization with stochastic dominance constraints, SIAM J. Optim. 14 (2003) 548–566.

[5] D. Dentcheva, A. Ruszczyński, Convexification of stochastic ordering, Comptes Rendus de l'Academie Bulgare des Sciences 57 (4) (2004) 7–14.

[6] D. Dentcheva, A. Ruszczyński, Semi-infinite probabilistic optimization: first order stochastic dominance constraints, Optimization 53 (2004) 583–601.

[7] P.C. Fishburn, Utility Theory for Decision Making, Wiley, New York, 1970.

[8] J. Hadar, W. Russell, Rules for ordering uncertain prospects, Amer. Econom. Rev. 59 (1969) 25–34.

[9] G. Hanoch, H. Levy, The efficiency analysis of choices involving risk, Rev. Econom. Stud. 36 (1969) 335–346.

[10] E. Lehmann, Ordered families of distributions, Ann. Math. Statist. 26 (1955) 399–419.

[11] H.B. Mann, D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Statist. 18 (1947) 50–60.

[12] A. Müller, D. Stoyan, Comparison Methods for Stochastic Models and Risks, Wiley, Chichester, 2002.

[13] J.P. Quirk, R. Saposnik, Admissibility and measurable utility functions, Rev. Econom. Stud. 29 (1962) 140–146.

[14] M. Rothschild, J.E. Stiglitz, Increasing risk: I. A definition, J. Econom. Theory 2 (1969) 225–243.