Decomposition Techniques and Disjunctive Linear Programming for Fixed-Income Portfolio Selection

by

Katherine G. Wyatt

A dissertation submitted to the Graduate Faculty in Mathematics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1997

UMI Number: 9732986

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted for the Graduate Faculty in Mathematics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

4/29/97
Date

Chair of Examining Committee

4/30/97
Date

Executive Officer

Supervisory Committee:

Ken McAloon

Joseph Malkevitch

Carol Tretkoff

THE CITY UNIVERSITY OF NEW YORK

Abstract

Decomposition Techniques and Disjunctive Linear Programming for Fixed-Income Portfolio Selection

by

Katherine G. Wyatt

Advisor: Professor Ken McAloon

Drawing on Sharpe's work in the '60s and '70s, researchers have developed a family of portfolio selection models that use absolute deviation, instead of variance, as a measure of dispersion of returns. We define a class of linear programs, called *step-shaped programs*, and show that programs for these absolute deviation models are step-shaped. The addition of logical requirements to programs in this class leads to the definition of *disjunctive step-shaped programs.* Variable decomposition for linear step-shaped programs is discussed, and our results for variable decomposition of disjunctive step-shaped programs are presented. Algorithms for the solution of disjunctive step-shaped programs are outlined and verified. We describe a hybrid method of decomposition introduced by Van Roy, called *cross decomposition*, and show the effects of applying this method to step-shaped programs. A new model for fixed-income portfolio selection, the *absolute deviation*

trade-off model, is introduced and the linear and disjunctive step-shaped programs that describe this model are detailed and analyzed. We include an overview of the financial ideas behind our model.

Acknowledgements

I am indebted to my advisors Prof. Ken McAloon and Prof. Carol Tretkoff for their encouragement and support. My many discussions with them helped focus my research, and the Logic-Based Systems Lab at Brooklyn College provided a warm and stimulating research environment. My thanks also to Prof. Joseph Malkevitch for serving on my committee. I am grateful to Prof. Robert Cowen for his early encouragement, which gave me the impetus to pursue this degree. Finally, my thanks to my daughters Nora, Emily, and Rachel Bloch, for their support and love.

This work was partially supported by Office of Naval Research Grant N000-14-96-1-1057.

Contents

1 Step-shaped programs

Portfolio selection models that use the absolute deviation of returns on different scenarios or time periods as the measure of risk have been developed by several researchers [Ko] [KoYa] [WoVZZe] [ZeKa]. Development of these models followed work by Sharpe [Sh67] [Sh71] using absolute deviation of returns in finding the characteristic or market line for a particular security or portfolio. The absolute deviation models are of interest because they can be formulated as programs with linear objective functions, as opposed to models that use variance as the measure of risk. The reader is invited to turn to Section 7 and Section 8 for an overview of the relevant financial material that motivates these models.

The programs that correspond to these models share several family characteristics. First, auxiliary variables are introduced to represent the absolute value of the difference between sampled and expected returns; second, there are constraints on the primary variables in which the auxiliary variables don't appear; third, the constraints in which both the primary and auxiliary variables appear do not further constrain the primary variables; and, finally, the only constraints on the auxiliary variables alone are bounding constraints restricting their values to the nonnegative orthant. We will call the primary variables *portfolio variables* and the auxiliary variables *risk variables.*

These characteristics motivate the following definitions.

DEFINITION: Let $R = \{x \in \Re^n : Ax \leq b\}$ and let $Dx + Ey \leq f$ be a system of linear inequalities. Then $Dx + Ey \leq f$ is *free* for $x \in R$ if for every $x \in R\ \exists y \geq 0$ such that (x, y) satisfies $Dx + Ey \leq f$. □

Let $x \in \Re_+^N$ and $y \in \Re_+^S$. Consider the following linear program (P):

$$\begin{array}{llrcl} (P) & \max & \{cx + dy\} & = & z_P \\ & \text{s.t.} & A_1^1 x + A_1^2 y & \leq & r \\ & & A_2^1 x + A_2^2 y & \leq & b \\ & & A_3^1 x + A_3^2 y & \leq & 0 \end{array}$$

DEFINITION: Let $R = \{x \in \Re_+^N : A_1^1 x \leq r\}$. A program (P) is a *step-shaped program* if the coefficient matrix for (P) can be written as above and if, in this formulation for (P), (1) A_1^2 and A_3^1 are zero submatrices, and (2) the constraints $A_2^1 x + A_2^2 y \leq b$ are free for $x \in R$. □

From the characterization above, we see that a program for an absolute deviation portfolio selection model has constraints on the portfolio variables alone, constraints that link the portfolio and the risk variables that are free for the portfolio variables, and bounding constraints on the risk variables alone. Therefore, these are step-shaped programs.

Programs for absolute deviation models typically have two constraints that link the portfolio and risk variables for every scenario. In applications it is necessary to include as many scenarios as possible, so these problems quickly become very large; analysis of decomposition techniques applied to step-shaped programs can lead to explicit solutions of relatively large programs. We will first discuss decomposition of linear step-shaped programs; disjunctive step-shaped programs are defined in Section 3, and decomposition methods for disjunctive step-shaped programs are also analyzed in that section.

We will need two definitions for the discussion that follows. Let (LP) be a linear program. Then we define $F(LP)$ to be the region in $\Re^n$ containing points feasible for (LP) and $v(LP)$ is the optimal objective function value of (LP).

Notation: We adopt the convention that a variable without a subscript (e.g., x) is a vector in $\Re^n$, and a variable with a subscript (e.g., x_j) is a real number and the jth coordinate of a vector in $\Re^n$. The only exception is when the subscripted variable has a vector symbol above it, in which case (e.g., $\vec{u}_i$) it is also a vector in $\Re^n$.

2 Decomposition techniques

Dantzig and Wolfe introduced a technique for decomposition of large-scale linear programs in [DaWo]. Their method is based on a partition of the problem's constraint matrix and the use of LP-duality to solve a master problem with many fewer constraints and simple subproblems. Held and Karp [HeKa70][HeKa71] developed the Lagrangian relaxation method, which uses decomposition of the constraint set to solve large mixed integer problems. Benders [Be] used a partition of the variable set to solve mixed integer problems. Van Slyke and Wets [VSWe] reported on a method called L-shaped decomposition, which is akin to Benders decomposition, for large-scale and stochastic linear problems. Finally, Van Roy introduced a method, cross decomposition, that combines variable (Benders) and constraint (Lagrangian) decomposition [VR]. Schrijver [Sc] describes Benders decomposition and Lagrangian relaxation. We will focus on variable decomposition and cross decomposition.

The form of the coefficient matrix for step-shaped programs is close to the Van Slyke and Wets description of L-shaped programs. The shape of the nonzero submatrices we are considering differs from the L-shaped form, however, in that we need to explicitly consider the bounding constraints on y since we will be interested in dual multipliers for the constraints. In fact, our definition of *free* makes precise an informal property discussed by [VR]. We will first outline variable decomposition of linear step-shaped programs.

2.1 Variable decomposition for linear step-shaped programs

Let (P) be a linear step-shaped program:

$$\begin{array}{llcl} (P) & \max & \{cx + dy\} & = & z_P \\ & \text{s.t.} & A_1^1 x + A_1^2 y & \leq & r \\ & & A_2^1 x + A_2^2 y & \leq & b \\ & & A_3^1 x + A_3^2 y & \leq & 0 \end{array}$$

Let $x \in \Re_+^N$ and $y \in \Re_+^S$. Let m_1 denote the number of rows in $\left(A_1^1 \; A_1^2 \right)$, let m_2 denote the number of rows in $\left(A_2^1 \; A_2^2 \right)$, and let m_3 denote the number of rows in $\left(A_3^1 \; A_3^2 \right)$. Then A_1^2 is an $m_1 \times S$ zero submatrix, and A_3^1 is an $m_3 \times N$ zero submatrix.

Let $M = m_1 + m_2 + m_3$, let $u \in \Re_+^M$ and let π be a projection operator. Then $\pi_1(u) = \vec{u}_1 \in \Re_+^{m_1}$, $\pi_2(u) = \vec{u}_2 \in \Re_+^{m_2}$, and $\pi_3(u) = \vec{u}_3 \in \Re_+^{m_3}$ and $u = \langle \vec{u}_1, \vec{u}_2, \vec{u}_3 \rangle$ is the same dimension as points feasible for the dual of (P).

To carry out variable decomposition, we rearrange the constraints for (P), and write the program as

$$\max_{\{x \geq 0\}} \{cx + \max_{y \in G} dy\}$$

$$G = \{y \in \Re_+^S : A_1^2 y \leq r - A_1^1 x, A_2^2 y \leq b - A_2^1 x, A_3^2 y \leq 0\}$$

Then by LP-duality, we can write (P) as

$$\max_{\{x \geq 0\}} \{cx + \min_{u \in J} \{\vec{u_1}(r - A_1^1 x) + \vec{u_2}(b - A_2^1 x) + \vec{u_3} 0\}\}$$

$$\text{such that } \; 0 \leq \vec{v}_1^s(r - A_1^1 x) + \vec{v}_2^s(b - A_2^1 x) + \vec{v}_3^s 0$$

$$\forall(\vec{v}_1^s, \vec{v}_2^s, \vec{v}_3^s) \text{ extreme rays of } \{u \in \Re_+^M : \vec{u}_1 A_1^2 + \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 \geq 0\}$$

$$J = \{u \in \Re_+^M : \vec{u}_1 A_1^2 + \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$$

Notice that since A_1^2 is a zero matrix, $\vec{u}_1 = \pi_1(u)$ is unconstrained in the description of J and for each $s = 1 \dots m_1, (\vec{v}_1^s, \vec{v}_2^s, \vec{v}_3^s)$ is the extreme ray where for the sth coordinate we have $\vec{v}_1^s = 1$, all other coordinates of $\vec{v}_1^s = 0$, all coordinates of $\vec{v}_2^s = 0$, and all coordinates of $\vec{v}_3^s = 0$. Since the constraints $A_2^1 x + A_2^2 y \leq r$ are free for $x \in \Re^N$, then if x is feasible for the constraints $A_1^1 x \leq r$, we have that the constraints $\vec{v}_1^s(r - A_1^1 x) + \vec{v}_2^s(b - A_2^1 x) + \vec{v}_3^s * 0$ are automatically satisfied. So as long as we ensure that x is feasible by including the constraints $A_1^1 x \leq r$, we can work in J', which is the projection of J, where $J' = \{\vec{u}_2, \vec{u}_3 : \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$ and drop the constraints involving v^s for all s.

Therefore, since by weak duality $dy \leq \vec{u}_2(b - A_2^1 x)$ for all feasible x, y, and $\vec{u}_2$, a program equivalent to (P) is the following one, where one continuous variable λ is substituted for dy and is bound by the dual objective function value at extreme points of J'. Let H be an *a priori* bound for y.

$$\begin{array}{llrcl}
(MP) & \max & \{cx + \lambda\} & = & z_{MP} \\
 & s.t. & A_1^1 x & \leq & r \\
 & & |\lambda| & \leq & dH \\
 & & \lambda & \leq & \vec{u}_2^t(b - A_2^1 x) \\
 & & & & \forall u^t = \langle \vec{u}_2^t, \vec{u}_3^t \rangle
\end{array}$$

Here u^t is an extreme point of $\{u \in \Re_+^{m_2+m_3} : \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$.

(MP) is the master problem for the variable decomposition reformulation of (P). The variable decomposition method for solving (MP) is (briefly) as follows. Assume $R = \{x \in \Re_+^N : A_1^1 x \leq r\} \neq \emptyset$. At the first iteration, solve the relaxation (MP^0) of (MP) by dropping the constraints

$$\lambda \leq \vec{u}_2^t(b - A_2^1 x) \ \forall u^t = \langle \vec{u}_2^t, \vec{u}_3^t \rangle = \text{extreme points of } \{u \in \Re_+^{m_2+m_3} : \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$$

from (MP). Let (x^0, λ^0) be the optimal solution. If we have

$$(1) \quad \min\{\vec{u}_2(b - A_2^1 x) : \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d, \langle \vec{u}_2, \vec{u}_3 \rangle \geq 0\} \geq \lambda^0$$

then $\lambda^0 = dy^0$ is the optimal solution of the program

$$(2) \quad \max\{dy : A_2^2 y \geq b - A_2^1 x^0, A_3^2 y \geq 0\}$$

and (x^0, y^0) is optimal for (P). If

$$\lambda^0 > \min\{\vec{u}_2(b - A_2^1 x^0) : \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d, \langle \vec{u}_2, \vec{u}_3 \rangle \geq 0\}$$

then (x^0, λ^0) is not feasible for (MP) and we have a valid cut that can be added to (MP^0)

$$\lambda \leq \min\{\vec{u}_2(b - A_2^1 x) : \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$$

to produce the relaxed master problem at stage 1, (MP^1).

At each iteration, this method either proves a point optimal or generates a cut to be added to the current relaxed master problem. (1) is called the dual subproblem $(DSP(x))$ and (2) is called the primal subproblem $(SP(x))$.

Using this method, we replace (P) with a series of relaxations of (MP) and easier subproblems. Solving (P) is carried out by first solving a relaxation of the master problem (MP^k) at stage k and then solving a subproblem

$(DSP(x^k))$ to obtain the extreme point that will figure in the test for optimality, and if this fails, in an inequality that is added for stage $k+1$. In this way the relaxations of (MP) are progressively tightened. Candidate optimal points are tested at each stage, and once the optimum has been found the algorithm ends. Subproblems are defined for a fixed x^k, the solution to the current relaxation of (MP). The solution to the primal subproblem $(SP(x^k))$ provides the optimal value for y^k.

$$\begin{array}{llrcl}
(DSP(x^k)) & \min & \{\vec{u}_2(b - A_2^1 x^k)\} & = & z_{DSP} \\
& \text{s.t.} & \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 & = & d \\
& & u & \geq & 0 \\
(SP(x^k)) & \max & \{dy\} & = & z_{SP} \\
& \text{s.t.} & A_2^1 x^k + A_2^2 y & \leq & b \\
& & A_3^2 y & \leq & 0
\end{array}$$

Note that the feasible region for $(DSP(x^k))$ does not depend on x, *i.e.* is the same for all x^k chosen.

2.2 Decomposition algorithm for linear programs

The decomposition algorithm consists of the following steps:

Initialization: Set $k := 0$. Let H be a bound for y. The initial master problem is:

$$\begin{array}{lrrcl}
(MP^0) & \max & z_{MP^0} & = & cx + \lambda \\
& \text{subject to} & A_1^1 x & \leq & r \\
& & |\lambda| & \leq & dH
\end{array}$$

1. Solve (MP^k) for (x^k, λ^k). If $\{x \in \Re_+^N : A_1^1 x \leq r\} = \emptyset$, end. (P) is infeasible.

2. Solve (DSP(x^k)) and obtain $\langle \vec{u}_2^k, \vec{u}_3^k \rangle$.

3. Use value of $\vec{u}_2^k$ found to test for optimality of (x^k, λ^k). A point is optimal if it is dual feasible, *i.e*, (x^k, λ^k) is optimal if

$$\lambda^k \leq \vec{u}_2^k(b - A_2^1 x^k).$$

4. If (x^k, λ^k) is optimal, solve $(SP(x^k))$ to obtain y^k, and end.

5. If (x^k, λ^k) fails test for optimality, then add cut with respect to $\vec{u}_2^k$ to (MP^k) to obtain (MP^{k+1}):

$$\lambda \leq \vec{u}_2^k(b - A_2^1 x)$$

   This is a valid cut since $\vec{u}_2^k$ is a projection of an extreme point of the feasible region of (DSP) for all x.

6. Set $k := k + 1$. Return to (1).

The following two theorems draw on standard results for variable decomposition.

Theorem 1 *If $\lambda^k \leq \vec{u}_2^k(b - A_2^1 x^k)$ at some stage k for the extreme point projection $\vec{u}_2^k$, then (x^k, y^k), where y^k is the optimal solution to $(SP(x^k))$, is optimal for (P).*

Proof. If $\lambda^k \leq \vec{u}_2^k(b - A_2^1 x^k)$ then since

$$\vec{u}_2^k(b - A_2^1 x^k) \leq \vec{u}_2^t(b - A_2^1 x^k)$$

for any other extreme point projection $\vec{u}_2^t$, then $\lambda^k = \vec{u}_2^k(b - A_2^1 x^k)$ by optimality of (x^k, λ^k). Further, since dy^k is the optimal solution to $(SP(x^k))$

$$dy^k = \vec{u}_2^k(b - A_2^1 x^k) = \lambda^k.$$

If y^k is feasible for $(SP(x^k))$ and $A_1^1 x^k \leq r$, then (x^k, y^k) is feasible for (P); thus (x^k, y^k) optimal for (MP^k) , which is a relaxation of (MP), means that (x^k, y^k) is optimal for (P) by the equivalence of (P) and (MP). □

Theorem 2 *The decomposition algorithm terminates.*

Proof. Assume that $z_P < \infty$ and that $\{(x,y) : (x,y)$ is feasible for $(P)\} \neq \emptyset$. The feasible region for (DSP) is bounded below and there are a finite number of extreme points. If we can show that no cut with respect to a particular dual multiplier (extreme point) is introduced twice then it will be clear that the algorithm will terminate. Let m and n be different iterations of the algorithm, where $m < n$. Assume that (x^n, λ^n) is the optimal solution to (MP^n) where $n-1$ cuts have already been added. Assume also that solving (DSP(x^n)) produces $\langle \vec{u}_2^n, \vec{u}_3^n \rangle$ as the optimal dual multiplier for x^n. Consider the case where $\vec{u}_2^n = \vec{u}_2^m$, a projection which figures in the cut added at the mth iteration of (MP). If this case obtains, (x^n, λ^n) will be found optimal, and a cut involving $\vec{u}_2^n$ will not be introduced. Thus the decomposition algorithm will terminate since cuts are not repeated and each cut involves one of a finite number of extreme points.

If $\{(x,y) : (x,y)$ is feasible for $(P)\} = \emptyset$, then since the constraints $A_2^1 x + A_2^2 y \leq b$ are free for $x \in R = \{x \in \Re_+^N : A_1^1 x \leq r\}$, it must be that $R = \emptyset$. In this case, the algorithm will end at stage 0. □

3 Variable decomposition of programs with disjunctive requirements

3.1 Disjunctive linear programs

DEFINITION: A polyhedral set is the solution set for a finite system of linear inequalities, and a finite union of polyhedral sets is called a disjunctive set. A *disjunctive (linear) program* is the problem of maximizing (or minimizing) a linear functional over the intersection of a family of disjunctive sets.

For notation, let M_i, for each $i, i = 1, \ldots, C$, equal the cardinality of $\{H_{ij}\}$, where each H_{ij} is a polyhedral set and $j \leq M_i$. Let $Ax \leq b$ be a system of linear inequalities, where A is an $m \times n$ matrix, $x \in \Re^n$, b is an $m \times 1$-vector, and let $F = \{x \in \Re^n : Ax \leq b\}$. Then we will denote the disjunctive set

$$F \cap (\bigcap_i \bigcup_j H_{ij}) = \bigcap_i \bigcup_j (F \cap H_{ij})$$

as $\cap_i \cup_j F_{ij}$. This set is the intersection over $i = 1, \ldots, n$ of the union of the polyhedral sets F_{ij}. F is called the *linear relaxation* of this disjunctive program.

We can write a disjunctive linear program as

$$\begin{array}{lll} (DP) & \max & cx \\ & \text{s.t.} & Ax \leq b \\ & & x \in \cap_i \cup_j F_{ij} \end{array}$$

where $c \in \Re^n$ and A is an $m \times n$ matrix, b is an $m \times n$-vector, and $\cup_j F_{ij}$ is a disjunctive set for each $i, i = 1, \ldots, n$. The intersection of the disjunctive sets F_{ij} can be described by a logical formula in conjunctive normal form (CNF):

$$\bigwedge_{i \leq N} (\vee_{j \leq M_i} (x \in F_{ij}))$$

or, equivalently, in disjunctive normal form (DNF):

$$\bigvee_{h \in Q} (x \in K_h)$$

where $|Q| = \prod_i nM_i$. If $M_i = M_j = M$ for $i, j \leq n$, then $|Q| = M^n$. We shall call $K_h, \forall h \in Q$, a *disjunctive feasible region*, since $K_h \subseteq \cap_i \cup_j F_{ij} \;\; \forall h \in Q$. The set of vectors $\cup_{h \in Q} K_h$ is a disjunctive set. See [McTr] for a discussion of disjunctive linear programming.

3.2 Solving disjunctive linear programs

Balas [Ba74][Ba85] describes a linear program which is equivalent to a given disjunctive program if the linear relaxation of the disjunctive program has a finite optimum. Using our notation, this formulation is as follows.

Let Q be the index set for $\cup K_h$, where each K_h is a disjunctive feasible region. Then the disjunctive program (DP) is equivalent to the following linear program:

$$\begin{array}{lll} (\mathcal{LP}) & \max & \sum_{h \in Q} c\alpha^h \\ & \text{s.t.} & A^h \alpha^h - b^h \alpha_0^h \leq 0 \;\; \forall h \in Q \\ & & \sum_{h \in Q} \alpha_0^h = 1 \\ & & (\alpha^h, \alpha_0^h) \geq 0 \;\; \forall h \in Q \end{array}$$

in the sense that

1. If x is a vertex of the convex closure of $\cup_{h \in Q} K_h$, then there exists $k \in Q$ such that α defined by $(\alpha^k, \alpha_0^k) = (x, 1)$ and $(\alpha^h, \alpha_0^h) = (0, 0)\ \forall h \in Q - \{k\}$ is a vertex of $F(\mathcal{LP})$.

2. If α is a vertex of $F(\mathcal{LP})$, then there exists $k \in Q$ such that $\alpha_0^k = 1$, $(\alpha^h, \alpha_0^h) = (0, 0)$ for $h \in Q - \{k\}$ and $x = \alpha^k$ is a vertex of K_k.

3. x is an optimal solution to (DP) iff α as defined in (1) is an optimal solution to $(\mathcal{LP})$.

Let $q = |Q|$. Then $(\mathcal{LP})$ has $q \times (n+1)$ variables (without counting possible slack variables) and $(q \times m) + 1$ constraints, where A is $m \times n$.

The coefficient matrix for $(\mathcal{LP})$ is block angular with q blocks. Clearly, if q is large $(\mathcal{LP})$ is a daunting linear program, even if the linear programs corresponding to each disjunct, $\max\{cx : x \in K_h\}$, are solved separately.

3.3 Decomposition for disjunctive step-shaped programs

Benders decomposition [Be] was developed for mixed integer programs and involves decomposition with respect to a set of "complicating" variables, that is, ones required to take on integer values. Linear subproblems are solved for a fixed integer-valued variable, and the optimal solution over all feasible integer values is found. An account of Benders decomposition can be found in [Sc]. Hooker [Ho] used Benders decomposition after expressing a problem with logical requirements as a $0-1$ program. Here, we will generalize Benders decomposition to disjunctive programs; *i.e.*, we will solve linear problems with respect to variables required to take on values in certain disjunctive sets. Since each disjunctive set is defined by a "large" linear step-shaped program, we can then apply variable decomposition to solving each of these linear programs. Note that step-shaped programs are *hereditary* in the sense that adding constraints on x yields another step-shaped program.

Assume we have a problem with the constraint set of the step-shaped linear problem (P), but with additional logical requirements for x. Assume also that these requirements can be represented by the condition $x \in \cap_i \cup_j F_{ij}$ where F_{ij} is a polyhedral set for all i and j. Then the program below is a step-shaped disjunctive program.

$$\begin{array}{lll} (L) & \max & cx + dy \\ & \text{s.t.} & A_1^1 x \leq r \\ & & A_2^1 x + A_2^2 y \leq b \\ & & A_3^2 y \leq 0 \\ & & x \in \cap_i \cup_j F_{ij} \end{array}$$

Proposition 1 *Let* $\{(x,y) : (x,y)$ *is feasible for* $(L)\} \neq \emptyset$. *Then the program* (L) *is equivalent to the following formulation for Benders decomposition:*

$$(1) \quad \max_{x \in \cap_i \cup_j F_{ij}} \{cx + \min_{u \in J}\{\vec{u}_1(r - A_1^1 x) + \vec{u}_2(b - A_2^1 x) + \vec{u}_3 0\}\}$$

where $J = \{u \in \Re^{3S+N+2} : \vec{u}_1 A_1^2 + \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$.

Proof. We can write (1) as

$$\max_{x \in \cap_i \cup_j F_{ij}} G(x)$$

where

$$G(x) = cx + \min_{u \in J}\{\vec{u}_1(r - A_1^1 x) + \vec{u}_2(b - A_2^1 x) + \vec{u}_3 s\}.$$

For each fixed $\bar{x}$, the linear program with $G(\bar{x})$ as its objective function and constraint set $\{u \in J\}$ is dual to

$$\begin{array}{lll} \max & c\bar{x} + & dy \\ \text{s.t.} & A_1^2 y & \le r - A_1^1 \bar{x} \\ & A_2^2 y & \le b - A_2^1 \bar{x} \\ & A_3^2 y & \le 0 \end{array}$$

The result follows from LP duality. □

We shall express the linear program corresponding to each *disjunctive region* of (L) separately. For each K_h, we have a linear step-shaped program (L_h).

$$\begin{array}{lll} (L_h) & \max & cx + dy \\ & \text{s.t.} & A_1^1 x \le r \\ & & A_2^1 x + A_2^2 y \le b \\ & & A_3^2 y \le 0 \\ & & x \in K_h \end{array}$$

We can combine the constraints $A_1^1 x \le r$ with the requirement $x \in K_h$ and rewrite the constraint set as $E^h x \le e^h$. Then we have

$$\begin{array}{lll} (L_h) & \max & cx + dy \\ & \text{s.t.} & E^h x \le e^h \\ & & A_2^1 x + A_2^2 y \le b \\ & & A_3^2 y \le 0 \end{array}$$

Since this is a linear step-shaped program, the discussion in Section 2.1 applies, and the full master problem for Benders decomposition of (L_h) is:

$$\begin{array}{lll} (ML_h) & \max & cx + \lambda \\ & \text{s.t.} & E^h x \le e^h \\ & & \lambda \le \vec{u}_2^t (b - A_2^1 x) \end{array}$$

for $\vec{u}_2^t = \pi(u^t)$, where $u^t = \langle \vec{u}_2^t, \vec{u}_3^t \rangle$ is an extreme point of the set $\{u \in \Re^{m_2+m_3} : \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$. The dual and primal subproblems are

$$\begin{array}{lll} (DSP_h\bar{x}) & \min & \vec{u}_2(b - A_2^1\bar{x}) + c\bar{x} \\ & \text{s.t.} & \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d \\ & & u \geq 0 \end{array}$$

$$\begin{array}{lll} (SP_h\bar{x}) & \max & c\bar{x} + dy \\ & \text{s.t.} & A_2^1\bar{x} + A_2^2 y \leq b \\ & & A_3^2 y \leq 0 \end{array}$$

Notice that the feasible region for the dual subproblem $(DSP_h\bar{x})$ is the same for every (L_h) and is in fact identical to the feasible region for the dual subproblem for the linear problem (P) which is a relaxation of (L). Let U be the collection of inequalities that are in a full variable decomposition master problem for (ML_h); *i.e.*, U contains all inequalities of the form $\lambda \leq \vec{u}_2(b - A_2^1 x)$ where $(\vec{u}_2, \vec{u}_3)$ is an extreme point of (DSP). So U contains all the inequalities for a full variable decomposition master problem for (P). Since $\cup_{h \in Q} K_h$ consists only of points feasible for the original program (P), we have $\cup_h K_h \subseteq F(P)$. Therefore, we consider two types of master problems: an *outer* master problem that only contains inequalities from U, and an *inner* master problem that contains inequalities from U plus inequalities that define disjunctive requirements. (MP) is an outer master problem; (ML_n) is an inner master problem at node n.

Relaxations of (MP) contain a subset of the inequalities from U and relaxations of (ML_h) contain a subset of inequalities from U in addition to a subset of the inequalities that define the requirement $x \in K_h$. So the inner master problem at node n, where n is an ancestor of leaf h, is a relaxation of (ML_h) and of (MP). We can test for optimality for any $x \in \cup_{h \in Q} K_h$ both for some L_h and for (P). We also can use the fact that if a point x satisfies $x \in \cup_{h \in Q} K_h$ and tests optimal for an outer master problem and therefore for (P), then this point must be optimal for (L).

3.4 Branch-and-bound search and decomposition

The constraint set for each polyhedral set K_h has a linear step-shaped coefficient matrix and if $i \in \{1, \ldots, N\}$ and $j \in \{1, \ldots, M\}$ there are M^N linear step-shaped programs to solve in any exhaustive search of this problem. Since the number of programs is exponential in the number of variables with disjunctive requirements, it is necessary computationally to shorten this search. We will use a branch-and-bound method to direct the search among the disjunctive regions and interleave the steps of a best-bound search with the cut generation steps of Benders (variable) decomposition. The point of this interleaving is to use the decomposition algorithm to direct the search for the maximal objective function among the disjunctive sets, and to use the maximum objective function value of any disjunctive feasible set as a lower bound to eliminate branches of the search tree. Both breadth-first and depth-first forms of branch-and-bound search are computationally important; we will show that either form of search can be combined with the Benders cut generation. A breadth-first search is described first.

At the first level of the breadth-first search, we find an optimal solution to the linear relaxation of the disjunctive problem. If this optimal solution satisfies the logical requirements, the algorithm terminates. If not, the search branches on a requirement i that is not satisfied, and replaces the relaxation with M linear programs, each of which defines a disjunctive feasible set for the ith requirement. This process continues until an optimal solution is found.

Algorithm for combined decomposition and breadth-first search:

Initialization: $k := 0$. Assume we have u^0, an extreme point of the feasible region for (DSP). Set lower bound $:= 0$. Let (MP^0) be the initial outer master problem; $(MP^0) = (L_0^1)$.

$$P_0^0 = \{x \in \Re_+^N : A_1^1 x \leq r\}$$

$$(L_0^1) \quad \max_{x \in P_0^0} z$$

$$z \leq cx + u_2^0(b - A_2^1 x)$$

Level $k (k > 0)$:

1. Let (L_j^k), for all j, be the program consisting of the current outer master problem plus the inequalities defining $x \in P_j^k$.
2. Solve (L_j^k) for each P_j^k. If for any (L_j^k) the initial objective function value is less than the current lower bound, or if (L_j^k) is infeasible, discard this set and problem. Calculate the optimal solution for

each P_j^k using Benders decomposition procedure with each (L_j^k) a relaxed inner master problem. The optimal solution is $x^{k,j}$ and optimal objective function value is $z^{k,j}$. The cuts that are added during the solution of each (L_j^k) are called *subcuts*. For each j, check if $x^{k,j}$ is disjunctive feasible – if it is, update the lower bound.

3. Find $\max\{x^{k,j}\} = x^k$ with objective function value $= z^k$ in set P^k. If x^k is disjunctive feasible, we're done. If not, find i such that $x^k \notin \cup_{j<N} F_{ij}$, and form $\{P_m^{k+1}\}$ for $m = 0, \ldots, M$ by replacing P^k with

$$Q_r^{k+1} = \{x \in P^k : x \in F_{ij}\} \text{ for } r = 1, \ldots, M$$

Each Q_r^{k+1} is a linear step-shaped program.

4. Add a master cut, a cut wrt u^k, where u^k is the optimal solution to $(DSP(x^k))$, to the current outer master problem (MP^k) and set $k := k + 1$ and return to (1).

To verify our algorithm, we show that only valid cuts are added to the relaxed inner and outer master problems, and that an optimal solution to (L) is optimal for a subproblem at some level of the search.

Proposition 2 *Let $z^{k,j}$ be the objective function value for a level k subproblem (L_j^k) where $F(L_j^k) \subseteq P_j^k$. Then the following inequality is valid for all extreme points u^t of $F(DSP)$:*

$$z^{k,j} \leq cx + \vec{u}_2^t(b - A_2^1 x) \quad \forall x \in P_j^k.$$

Proof. Consider the level 1 problem (L_0^1) which is equivalent to (P), the relaxation obtained by dropping all the disjunctive requirements. Let z be a nonnegative variable representing the objective function value of (L_0^1). Then $z \leq cx + \vec{u}_2^t(b - A_2^1 x)$ for all $x \in F(P)$ and for all extreme points of $F(DSP)$. (L_0^1) is a relaxation of every (L_j^k) so $v(L_j^k) \leq v(L_0^1)$ for every k and j, or $z^{k,j} \leq z$ for all subproblems (L_j^k), since $\{x \in P_j^k\} \subseteq \{x \in F(L_0^1)\}$ for all sets P_j^k. □

Theorem 3 *If $(\bar{x}, \bar{y})$ is optimal for the program (L) with disjunctive requirements, then $(\bar{x}, \bar{y})$ is optimal for some level k problem for some k and for all feasible problems descendent from it.*

Proof. Assume $(\bar{x}, \bar{y})$ is optimal for (L). Then $\bar{x}$ is a point in the union of regions that satisfy each disjunctive requirement *i.e.*, $\bar{x}$ is a point in at least one K_h, where $x \in K_h$ satisfy one clause in a DNF formulation of the logical requirements. If $(\bar{x}, \bar{y})$ is optimal for (L), then $\bar{x}$ achieves the maximal objective function value over all $x \in K_h$ and there is a point (x^*, y^*) where $cx^* + dy^* = c\bar{x} + d\bar{y}$ and (x^*, y^*) is an optimal point of $K_h \cap F(P)$. There is some smallest k such that x^* is the optimal solution across all (L_j^k) and a smallest k' such that x^* is the optimal solution to some problem $(L_j^{k'})$. Then x^* is optimal for all problems descendent from $(L_j^{k'})$. □

So if (x^*, y^*) is an optimal solution to (P), then (x^*, y^*) is an optimal point of (L_0^1) and all programs descendent from (L_0^1).

Corresponding to (L_j^k) is the problem (LF_j^k) where k disjuncts $F_{ij(i)}$ have been selected:

$$\begin{array}{lll} (LF_j^k) & \max & cx + dy \\ & \text{s.t.} & A_1^1 x \leq r \\ & & A_2^1 x + A_2^2 y \leq b \\ & & A_3^2 y \leq 0 \\ & & x \in F_{i,j(i)} \qquad i \text{ selected} \end{array}$$

Let (x^k, λ^k) be an optimal solution for (L_j^k) and u^k the solution to $(DSPx^k)$. Then

$$\lambda^k \leq \vec{u}_2^k (b - A_2^1 x^k)$$

and x^k is the projection in $\Re^N$ of a point $(x^k, y^k) \in \Re^{N+S}$ such that $dy^k = \lambda^k$ and (x^k, y^k) is optimal for (LF_j^k).

At each level, a determination of the set with the best solution is made; if this solution is disjunctive feasible, an optimal solution has been found and the algorithm terminates. If this 'best' solution is not disjunctive feasible, the 'child' programs are formed. At any level, relaxed inner master problems have to be solved only for the new child programs; the values of the other sets at a level are carried over.

Only master cuts, those with u^k, the optimal solution to $(DSPx^k)$ where x^k is the best solution at a level, are carried over from level to level and accumulated in the outer master problem. The other subcuts generated during the solution of a particular inner master problem (L_j^k) are not carried to the next level. The master cuts are only added to the new programs formed at level k if x^k is not disjunctive feasible. It may be that subcuts will be repeated from level to level, but the master cuts will not be repeated.

The following proposition shows that it is not necessary to add the k-level master cut to all the sets at a level.

Proposition 3 *Assume we have completed stage k of the decomposition and branch-and-bound algorithm, i.e. we have found z^k, the best objective function value across all sets at stage k. So z^k is the optimal objective function value for the program (L^k), which describes some set P_j^k. We have (x^k, λ^k) the optimal solution to (L^k) and u^k the optimal solution to $(DSPx^k)$. Let $(L\bar{P})$ be a stage k inner master problem different from (L^k), and let $\bar{x}$ be its optimal solution, $\bar{z}$ its optimal value, and $\bar{u}$ the optimal solution to $(DSP\bar{x})$. Then we have*

$$\bar{z} \leq c\bar{x} + u^k(b - A_2^1\bar{x})$$

Proof. By definition, $\bar{u}$ is the solution to

$$\min_{\vec{u}_2 \geq 0}\{\vec{u}_2(b - A_2^1\bar{x}) + c\bar{x} : \vec{u}_2 A_2^2 \geq d\}$$

Therefore, for any extreme point $\tilde{u}$ of $F(DSP)$.

$$c\bar{x} + \vec{\bar{u}}_2(b - A_2^1\bar{x}) \leq c\bar{x} + \vec{\tilde{u}}_2(b - A_2^1\bar{x})$$

Thus, a cut involving u^k, *i.e.*, $cx + \lambda \leq \vec{u}_2^k(b - A_2^1 x)$, will be satisfied by $\bar{x}$ and $\bar{z}$ of $(L\bar{P})$. □

Theorem 4 *The combined Benders decomposition and breadth-first disjunctive search algorithm will terminate.*

Proof. Assume, first, that the linear relaxation (P) of the program (L) with disjunctive requirements is feasible and (L) itself is feasible. If $M =$ the number of sets feasible for each disjunctive requirement and $N =$ the number of such requirements, then there are M^N possible disjunctive feasible sets and the solution to (L) will be found in one of them. Let $R =$ the number of extreme points of the feasible region for the decomposition subproblem (DSP). $R < \infty$ since (DSP) describes a polyhedron. At most R master cuts can be added during the algorithm since the master cuts accumulate and cannot be repeated. And at most $R - k$ subcuts can be added to any stage k relaxed inner master problem. Every relaxed inner master problem at each stage k will end by the proof of termination for the linear Benders decomposition case. Assume that at every stage of the algorithm the optimal solution for that stage is not disjunctive feasible. Then at every stage new programs are formed and the optimal solution across the level is selected. The sets at any stage are disjoint and correspond to nodes of a search tree. No set is formed twice since the sets are disjoint and therefore at most M^N sets will be formed and the algorithm will terminate.

If the linear relaxation (P) is feasible, but (L) is infeasible, then the disjuncts will all lead to infeasibility. If (P) is infeasible, then the algorithm will terminate at stage 0, as noted in Section 2.2. □

The algorithm will terminate with the optimal solution if no sets containing optimal points are removed.

Theorem 5 *Pruning will not remove optimal points.*

Proof. A set P_j^k is eliminated from the stage k search if the initial solution to (L_j^k) is strictly less than the current lower bound. The initial program (L_j^k) is a relaxation of both the full master problem for the set P_j^k and a relaxation of $(L_j^k)^*$, the program with the constraints of (L_j^k) and the full set of disjunctive requirements. Therefore, if the initial value for (L_j^k) is less than the current lower bound, the optimal objective function value for $(L_j^k)^*$ is less than the current lower bound, and the optimal solution to $(L_j^k)^*$ cannot be the optimal solution to (L). □

3.5 Decomposition and depth-first search

Let n be the index of a node in a search tree that uses the parameters M and N described earlier. Then there are at most $\sum_{i=1}^{N-1} M^i$ nodes and associated with each node is a linear program (L_n) with constraints that describe the disjunctive feasible regions selected at node n. There are M^N leaves to the search tree, and each of these leaves corresponds to a linear program (L_h), where all the disjunctive requirements have been satisfied. We can think of the set $K_h = \{x : x \text{ is feasible for } (L_h)\}$ in terms of a DNF formulation of the logical requirements for (L). Then $K_h = \cup_{i \leq N} F_{i,j(i)}$ where $j(i) \leq M$ is the index for a particular domain for x_i that satisfies the logical requirement i. If Q is the index set for $\cup K_h$ then $|Q| = M^N$.

In combined depth-first-search and variable decomposition, a solution is found at a node and then tested for disjunctive feasibility. If the solution is in one of the disjunctive regions, then it is tested for optimality at a node.

If it does not satisfy the disjunctive requirements, then a new disjunctive set is branched on. If a point x satisfies the disjunctive requirements, there is some K_h such that $x \in K_h$; and whether this point is optimal for (ML_h) or not, its objective function value can be used as a lower bound. Further, an inner master problem at node n, (ML_n^k) at stage k is a relaxation of (ML_n^{k+1}) and all nodes descendent from it. So if a disjunctive feasible point is optimal for (ML_n^k) then all children of (ML_n^k) can be pruned; if the initial objective function value for (ML_n^k) is less than the current lower bound, (ML_n^k) and its children can be pruned. The linear problems and their corresponding full inner master problems for variable decomposition are as follows:

$$
\begin{array}{lll}
(L_h) & \max & cx + dy \\
 & \text{s.t.} & A_1^1 x \le r \\
 & & x \in K_h \\
 & & A_2^1 x + A_2^2 y \le b \\
 & & A_3^2 y \le 0
\end{array}
\qquad
\begin{array}{lll}
(ML_h) & \max & cx + \lambda \\
 & \text{s.t.} & A_1^1 x \le r \\
 & & x \in K_h \\
 & & \lambda \le \vec{u}_2^t (b - A_2^1 x) \\
 & \forall & u^t \text{ extreme points of } F(DSP)
\end{array}
$$

$$
\begin{array}{lll}
(L_n) & \max & cx + dy \\
 & \text{s.t.} & A_1^1 x \le r \\
 & & x \in \cup_j F_{i,j(i)}, i \text{ selected} \\
 & & A_2^1 x + A_2^2 y \le b \\
 & & A_3^2 y \le 0
\end{array}
\qquad
\begin{array}{lll}
(ML_n) & \max & cx + dy \\
 & \text{s.t.} & A_1^1 x \le r \\
 & & x \in \cup_j F_{i,j(i)}, i \text{ selected} \\
 & \forall & u^t \text{ extreme points of } F(DSP)
\end{array}
$$

The algorithm is as follows:

Initialization: Assume we have u^0 an extreme point for the feasible region for (DSP). Let $n = 1, \ldots, T$ index the nodes of the search tree. Set $k := 0$. Select the root node as the first node to examine. Set lower bound equal to $-\infty$. Let $(ML^0) = (L_0^1)$ be the current outer master problem.

$$\begin{array}{lll}(L_0^1) & \max & z \\ & \text{s.t.} & A_1^1 x \leq r \\ & & z \leq cx + \vec{u}_2(b - A_2^1 x)\end{array}$$

Stage $k(k > 0)$:

(1) Solve the relaxed inner master problem (ML_n) at node n to get $(\bar{x}_n, \bar{\lambda}_n)$ and the relaxed objective function value $\bar{z}_n$. If $\bar{z}_n \geq$ lower bound, go to (2). If $\bar{z}_n <$ lower bound, go to (3). If (ML_n) is infeasible, go to (3).

(2) Check if $\bar{x} \in \cup_{h \in Q} K_h$.

(a) If $\bar{x}$ is disjunctive feasible, then solve $(DSP\bar{x})$ to get $\bar{u}$. Test $(\bar{x}_n, \bar{\lambda}_n)$ for optimality for (ML_n).

(i) If $(\bar{x}_n, \bar{\lambda}_n)$ is optimal for (ML_n) then update lower bound if possible and go to (3). If all nodes have been explored, stop. If the lower bound is greater than or equal to 0, then lower bound is optimal.

(ii) If $(\bar{x}_n, \bar{\lambda}_n)$ is not optimal for (ML_n) then add cut with respect to $\bar{u}$ to (ML_n).

(b) If $\bar{x}_n \notin \cup_{h \in Q} K_h$, then branch on a new disjunctive set; *i.e.*, select a node descendent from n.

Go to (1).

(3) Set $k{:=}k+1$. Backtrack to last unexplored node m. Go to (1). If all nodes have been explored, stop. The lower bound is optimal.

Theorem 6 *Assume that the linear relaxation (P) of the disjunctive program (L) is feasible. If (L) is feasible, then the algorithm combining variable decomposition and depth-first search will terminate with the optimal solution. If (L) is infeasible, then the algorithm will end with no solution.*

Proof. If no nodes are pruned in the search for the optimal solution to (L), then all M^N leaves in the search tree will be explored and the optimal objective function value will be found for each disjunctive feasible set. In this case, the algorithm clearly terminates and the optimal solution is the best objective function value found at a leaf. Since any nodes pruned by this algorithm achieved relaxed objective function values strictly less than the current lower bound, and since any parent node is a relaxation of its descendents, leaves which contain optimal solutions are not removed. Therefore, the algorithm will terminate with the optimal value.

However, if there is no feasible solution to (L), then every inner master problem at a leaf is infeasible. This infeasibility may be detected at a node higher up in the search tree, in which case the descendents of this node will not be explored. But since every disjunctive region is considered, the algorithm will terminate with no solution. □

3.5.1 Verifying Optimality

In practice, we wish to verify optimality for the disjunctive step-shaped program (L) considering as few nodes as possible. Consider the following subroutine (SB) that uses the Benders decomposition method for a linear program

(P).

1. Solve a relaxed master problem for (P) for $(\bar{x}, \bar{\lambda})$.

2. Solve $(DSP\bar{x})$ for the optimal dual multiplier $\bar{u}$.

3. Test $(\bar{x}, \bar{\lambda})$ for optimality.

4. If $(\bar{x}, \bar{\lambda})$ is not optimal, add cut with respect to $\bar{u}$ to the relaxed master problem.

5. Solve the new master problem for (x^*, λ^*).

Lemma 1 *If in the subroutine (SB) outlined above, the point x^* found at step (5) is equal to $\bar{x}$ found at step (1), then (x^*, λ^*) is optimal for (P).*

Proof. By definition, $\bar{u}$ is an optimal solution to the problem

$$\min\{\vec{u}_2(b - A_2^1\bar{x}) + c\bar{x} : \vec{u}_2A_2^2 + \vec{u}_3A_3^2 = d\}$$

and (x^*, λ^*) is an optimal solution to the master problem from step (1) with the added inequality $\lambda \leq \bar{\vec{u}}_2(b - A_2^1x)$. So if $\bar{x} = x^*$, and u^* is an optimal solution of $(DSPx^*)$, then $\bar{\vec{u}}_2(b - A_2^1\bar{x}) = \vec{u}_2^*(b - A_2^1x^*)$ and we have $\lambda^* \leq \bar{\vec{u}}_2(b - A_2^1\bar{x})$, since $\bar{x} = x^*$. Therefore $\lambda^* \leq \vec{u}_2^*(b - A_2^1x^*)$ and (x^*, λ^*) is optimal for (P). □

Assume a node has been selected at the kth stage of the combined depth-first and decomposition algorithm. Let $(\hat{x}, \hat{\lambda})$ be the initial solution of the inner master problem (ML_n) which consists of the inequalities from the current outer master problem (MP^k) plus the constraints describing the region feasible at node n. Let $\hat{u}$ be the optimal solution to $(DSP\hat{x})$.

Proposition 4 *Let $\hat{x}$ as described above be disjunctive feasible. Consider the following operation: add a cut with respect to $\hat{u}$ to the outer problem MP^k and apply subroutine (SB) to this problem. If $x^* = \bar{x}$ in the subroutine, and if $\hat{x} = x^*$, then $(\hat{x}, \hat{y})$, where $d\hat{y} = \lambda^*$, is optimal for (L).*

Proof. We have (x^*, λ^*) optimal for (P) by the proof of the preceding lemma, and since $\hat{x} = x^* \in \cup_{h \in Q} K_h$ x^* satisfies the disjunctive requirements for (L). (P) is a relaxation of (L), so (x^*, y^*), where $dy^* = \lambda^*$ is optimal for (L). □

4 Constraint decomposition and cross decomposition

The cross decomposition method of Van Roy [VR] combines Benders (variable) decomposition and Lagrangian (constraint) decomposition. We will develop the Lagrangian relaxation method of constraint decomposition first, since this formulation is necessary for cross decomposition.

In the Lagrangian relaxation method [HeKa70] [HeKa71] the constraint set of a linear problem (often an integer linear problem) is partitioned. Consider a general linear programming problem:

$$\begin{array}{lll} (LP) & \max & cx \\ & \text{s.t.} & A_1x \le b_1 \\ & & A_2x \le b_2 \end{array}$$

where A_1 and A_2 have order $m_1 \times n$ and $m_2 \times n$, respectively. By linear programming duality,

$$v(LP) = \max\{cx : A_1x \le b_1, A_2x \le b_2\}$$

$$\le \min_{u \ge 0}\{ub_1 + \max\{cx - uA_1x : A_2x \le b_2\}\}$$

The minimization problem above is the *Lagrangian dual* or *Lagrangian relaxation* and the components of u are *Lagrangian multipliers*. Note that the constraints $A_1x \le b_1$ have been added as a penalty function to the objective function of (LP) and the Lagrangian relaxation provides an upper bound for (LP). The master problem for Lagrangian (constraint) decomposition of

(LP) is

$$\begin{array}{lll} (ML) & \min & \eta \\ & \text{s.t.} & \eta \geq cx^t - u(b_1 - A_1 x^t) \\ & & u \geq 0 \end{array}$$

for all extreme points x^t of $\{x : A_2 x \leq b_2\}$. The Lagrangian subproblem is:

$$\begin{array}{lll} (SD) & \max & cx + \bar{u}(b_1 - A_1 x) \\ & \text{s.t.} & A_2 x \leq b_2 \\ & & x \geq 0 \end{array}$$

Master problems are solved for fixed x that are extreme points of the feasible region for the subproblem; subproblems are solved for fixed Lagrangian multipliers that are optimal solutions to the master problem. Inequalities for the extreme points x^t accumulate in the master problem.

Unlike the situation with variable decomposition, the Lagrangian relaxation may not be equivalent to the original problem (P).

DEFINITION: Let $d = v(ML) - v(LP)$. Then d is called the *duality gap* of the Lagrangian relaxation (ML).

The interesting case, of course, is when $d > 0$.

The cross decomposition method of Van Roy combines Benders (variable) decomposition and Lagrangian (constraint) decomposition in one method, where the coefficient matrix A is partitioned horizontally and vertically. Instances of the portfolio problem typically involve a large number of scenarios, so it is appealing to try this technique, since the disjunctive requirements make the portfolio variables complicated, and the number of scenarios make the constraints that define risk also complicating. We will first outline the cross decomposition method and then discuss some of the problems with applying it to step-shaped programs.

In cross decomposition the multiplier for the Lagrangian subproblem is an extreme point of $F(DSP)$, the feasible region for the dual subproblem for Benders (variable) decomposition. Variable and constraint decomposition methods are dual to each other; this can be exploited in cross decomposition where the master problem of one type is equivalent to the subproblem of the other when the input variable is fixed. To describe cross decomposition and these equivalences, we will use Van Roy's notation. A representation of the coefficient matrix for a typical problem (LP) in variable and constraint decomposition is as follows:

$$A = \begin{bmatrix} A^1 & A^2 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{bmatrix}$$

$$\begin{array}{lll} (LP) & \max & cx + dy \\ & \text{s.t.} & A_1^1 x + A_1^2 y \leq b_1 \\ & & A_2^1 x + A_2^2 y \leq b_2 \end{array}$$

We will also need the following facts about the relationships among problem formulations.

(P) original problem (max)

(LD) Lagrangian Dual (min)

(MP_F) full Benders master problem (including cuts for all extreme points of dual region) (max)

(MD_F) full Lagrangian master problem (including cuts for all extreme points of primal feasible region) (min)

(MDx_2) relaxation of (MD_F) with x_2 fixed (min) $(\equiv (DSPx_2) \equiv (SPx_2))$

(MPu_2) relaxation of (MP_F) with u_2 fixed (max) $(\equiv (SDu_2))$

$$(MP_F) \equiv (P) \quad v(MP_F) = v(P)$$

$$(MD_F) \equiv (LD) \quad v(MD_F) = v(LD)$$

$v(LD) \geq v(P)$ since the Lagrangian problem is a relaxation of (P).

(MDx_2) is a relaxation of (MD_F) so $v(MDx_2) \leq v(MD_F) = v(LD)$.

(MPu_2) is a relaxation of (MP_F) so $v(MPu_2) \geq v(MP_F) = v(P)$.

$v(SDu_2) \geq v(MD_F)$ since $(MD_F) = \min_{u_2 \geq 0} v(SDu_2)$.

$v(SPx_2) = v(DSPx_2) \leq v(MP_F)$ since $(MP_F) = \max_{x_2 \geq 0} v(SPx_2)$.

Ordering of optimal objective functions is as follows:

$$\begin{array}{ccc} v(SDu_2) & = & v(MPu_2) \\ \geq & & \geq \\ v(LD) & = & v(MD_F) \\ \geq & & \geq \\ v(P) & = & v(MP_F) \\ \geq & & \geq \\ v(DSPx_2) & = & v(MDx_2) \end{array}$$

An equivalence between problems can be defined as follows: A problem (Q) is equivalent to (Q') with respect to a subset of variables, say z, if the optimal solutions of (Q) for z are optimal in (Q') for z and vice versa.

To use the constraint decomposition subproblem and variable decomposition subproblems instead of the constraint master problem and subproblem or variable master problem and subproblems, we have to first show that if a variable is fixed for the master constraint decomposition problem (MD) and the variable decomposition subproblem (SP) or (DSP), then the variables common to these two problems will have identical values at optimality and the optimal objective function values for the master constraint problem and the variable subproblem are equal. Then we have to show the corresponding equivalence for the master variable decomposition problem (MP) and the constraint decomposition subproblem (SD).

However, since a master problem accumulates cuts involving extreme points of the feasible region for its subproblem, we have to consider the changes in the feasible region for the subproblem when we fix a variable in the master problem. Define $F(\bar{SD}x_2)$ as the feasible region for the subproblem that has x_2 fixed and $(\bar{MD}x_2)$ as the master problem that has x_2 fixed. Then $(\bar{MD}x_2)$ has cuts involving extreme points of $F(\bar{SD}x_2)$, *i.e.* extreme points (x_1, x_2) that may or may not be extreme points of $F(SD)$, the feasible region for (SD) that does not have x_2 fixed. If (x_1, x_2) is not an extreme point of $F(SD)$ then (x_1, x_2) can be expressed as a convex combination of the extreme points of $F(SD)$, and therefore $v(\bar{MD}x_2) \geq v(MDx)$, the master problem without x_2 fixed.

Similarly, define $F(D\bar{S}Pu_2)$ as the feasible region for the variable sub-problem that has u_2 fixed and $(\bar{M}Pu_2)$ as the variable master problem that has u_2 fixed. Since $(\bar{M}Pu_2)$ has cuts involving the extreme points of $F(D\bar{S}Pu_2)$, $v(\bar{M}Pu_2) \leq v(MPu)$, which has cuts involving only the extreme points of $F(DSP)$, where u_2 is not fixed.

Let $x = (x_1, x_2)$ and $u = (u_1, u_2)$. Then we can demonstrate equivalence as defined above for $(\bar{M}Dx_2)$ and $(D\bar{S}Px_2)$, where x_2 is fixed, and correspondingly for $(\bar{M}Pu_2)$ and $(\bar{S}Du_2)$, where u_2 is fixed. The following is based on [VR].

4.1 Proof of equivalences

Fix u_2. $(\bar{M}Pu_2) \equiv (\bar{S}Du_2)$ with respect to x_2 and $v(\bar{M}Pu_2) = v(\bar{S}Du_2)$.

$$(\bar{M}Pu_2) \quad \max_{x_2 \in Z} x_0$$

$$\text{s.t. } u^t b + (c^2 - u^t A^2) x_2 \geq x_0$$

(where u^t is an extreme point of $F(D\bar{S}Pu_2)$, $u^t = (u^{t_1}, u_2)$). This is equivalent to

$$\max_{x_2 \in Z} \inf\{u^t b + (c^2 - u^t A^2) x_2\}$$

$$\equiv \quad \max_{x_2 \in Z} \left\{ \begin{array}{lll} \min_{u_1 \geq 0} & & ub + (c^2 - uA^2) x_2 \\ & \text{s.t.} & u_1 A_1^1 \geq c^1 - u_2 A_2^1 \end{array} \right\}$$

Since u_2 is fixed, inf is min with respect to u_1, where (u_1, u_2) is an extreme point of $F(DSPx_2)$. We can take the dual of the inner min since x_2 and u_2 are fixed. We then have

$$max_{x_2 \in Z} \left\{ \begin{array}{ll} \max_{x_1 \geq 0} & (c^1 - u_2 A_2^1)x_1 + u_2 b_2 + (c^2 - u_2 A_2^2)x_2 \\ \text{s.t.} & A_1^1 x_1 \leq b_1 - A_1^2 x^2 \end{array} \right\}$$

$$\equiv \max_{x_2 \in Z, x_1 \geq 0} \left\{ \begin{array}{c} (c - u_2 A_2)x + u_2 b_2 \\ A_1 x \leq b_1 \end{array} \right\} \equiv (\bar{S}Du_2)$$

Claim: Fix x_2. $(\bar{M}Dx_2) \equiv (D\bar{S}Px_2)$ with respect to u_2 and $v(\bar{M}Dx_2) = v(D\bar{S}Px_2)$.

$$(\bar{M}Dx_2) \quad \min_{u_2 \geq 0} u_0$$

$$\text{s.t. } cx^t + u_2(b_2 - A_2 x^t) \leq u_0$$

(where $x^t = (x_1^t, x_2)$ is an extreme point of $F(\bar{S}Dx_2)$).

$$\equiv \min_{u_2 \geq 0} \sup\{cx^t + u_2(b_2 - A_2 x^t)\}$$

$$\equiv \min_{u_2 \geq 0} \left\{ \begin{array}{ll} \max_{x_1 \geq 0} & cx + u_2(b_2 - A_2 x) \\ \text{s.t.} & A_1^1 x_1 \leq b_1 - A_1^2 x_2 \end{array} \right\}$$

Because $x^t = (x^{t_1}, x_2)$ is an extreme point of $F(\bar{S}Dx_2)$ with x_2 fixed and u_2 is fixed in inner max we can take this dual and express it in terms of $(u_1, u_2) = u$ and constants x_2, u_2:

$$\min_{u_2 \geq 0} \left\{ \begin{array}{ll} \min_{u_1 \geq 0} & u_1(b_1 - A_1^2 x_2) + c^2 x_2 + u_2(b_2 - A_2^2 x_2) \\ \text{s.t.} & u_1 A_1^1 \geq c^1 - u_2 A_2^1 \end{array} \right\}$$

$$\equiv \min_{u \geq 0} \left\{ \begin{array}{ll} & u(b - A^2 x_2) + c^2 x_2 \\ \text{s.t.} & uA^1 \geq c^1 \end{array} \right\} \equiv (D\bar{S}Px_2)$$

The fact that the Lagrangian subproblem is equivalent to the Benders master problem and the Benders subproblem is equivalent to the Lagrangian master problem means that the optimal objective function values are equal for these equivalent problems. The equality of the optimal objective function values translates into a description of a kind of complementary slackness for cross decomposition. Since x_1 appears in the Lagrangian problems and not in the Benders problem and u_1 appears in the Benders problem and not in the Lagrangian problem, we have

$$(c^1 - \bar{u}_2 A_2^1) x_1 = u_1 (b_1 - A_1^2 \bar{x}_2).$$

Van Roy described conditions for optimality of the subproblem solutions in the following two lemmas [VR].

Lemma 2 *Let u^0 be an optimal dual solution of (SPx_2^0) and let x^+ be an optimal solution of (SDu_2^0). Then $x_2^0 \neq x_2^+$ unless $v(SPx_2^0) = v(P)$.*

Proof. Assume $x_2^0 = x_2^+$ and $v(SPx_2^0) < v(P)$. We have $v(SPx_2^0) \leq v(P)$ always since the value of the full Benders master problem equivalent to (P) is the maximum over all $x_2 \geq 0$ of $v(SPx_2)$. Now u^0 is an optimal solution of $(DSPx_2^0)$ so u_2^0 is optimal in (MDx_2^0) and

$$v(MDx_2^0) = c^1 x_1 + c^2 x_2^0 + u_2^0 (b_2 - A_2^1 x_1 - A_2^2 x_2^0).$$

where (x_1, x_2^0) is an extreme point of $F(\bar{SD}x_2)$. But x^+ is optimal for (SDu_2^0) so

$$v(SDu_2^0) = cx^+ + u_2^0(b_2 - A_2x^+) = v(MDx) \text{ without } x_2 \text{ fixed}$$

$$\leq v(\bar{M}Dx_2^0) = v(SPx_2^0).$$

or,

$$v(SPx_2^0) = v(MDx_2^0) \geq v(MDx) \geq v(P).$$

However, x^+ is also optimal in (MPu_2^0) by the equivalence of (MPu_2^0) and (SDu_2^0) and $v(MPu_2^0) \geq v(P)$ since (MPu_2^0) is a relaxation of (MP_F) which is equivalent to (P). Therefore, we have

$$v(SDu_2^0) = v(MPu_2^0) \geq v(P) > v(SPx_2^0).$$

This contradicts what we found above, so it must be that $v(SPx_2^0) = v(P)$. □

Lemma 3 *Let x^0 be an optimal solution of (SDu_2^0) and let u^+ be an optimal dual solution of (SPx_2^0). Then $u_2^0 \neq u_2^+$ unless $v(SDu_2^0) = v(P)$.*

Proof. Assume $u_2^+ = u_2^0$ and $v(SDu_2^0) > v(P)$. We have $v(SDu_2^0) \geq v(P)$ always since $v(MD_F)$ is minimum over $u_2 \geq 0$ of $v(SDu_2^0)$ and $v(MD_F) = v(LD) \geq v(P)$. We have x_2^0 an optimal solution of (MPu_2^0) and

$$v(MPu_2^0) = u_1b_1 + u_2^0b_2 + (c^2 - u_1A_1^2 - u_2^0A_2^2)x_2^0$$

where (u_1, u_2^0) is an extreme point of $F(D\bar{S}Pu_2^0)$ and since $u_2^+ = u_2^0$

$$v(MPu_2^0) = u_1b_1 + u_2^+b_2 + (c^2 - u_1A_1^2 - u_2^+A_2^2)x_2^0.$$

But u^+ is optimal for $(DSPx_2^0)$ so

$$v(DSPx_2^0) = c^2x_2^0 + u^+b - u^+A^2x_2^0 = v(MPu) \text{ without } u_2^0 \text{ fixed}$$

$$\geq v(MPu_2^0) = v(SDu_2^0) > v(P).$$

However, $v(DSPx_2^0) \leq v(MP_F) = v(P)$. So the assumption $v(SDu_2^0) > v(P)$ contradicts the optimality of u^+ for $(DSPx_2^0)$. □

Conditions under which there will be a zero duality gap are outlined in the theorem below from Van Roy [VR].

Theorem 7 *Let (x^*, y^*) be an optimal solution of (P) and let u_2^* be an optimal solution of (LD). Then the following are equivalent:*

(a) *The Lagrangian relaxation relative to the constraints $(A_2^1x + A_2^2y \leq b)$ has no duality gap, i.e., $v(P) = v(LD)$.*

(b) *There is an optimal solution u^+ of $(DSPx^*)$ with $v(SDu_2^+) = v(DSPx^*)$.*

(c) *There is an optimal solution (x^+, y^+) of $(SD(u_2^*))$ with $v(DSPx^+) = v(SD(u_2^*))$.*

Proof (a) → (b): Assume $v(P) = v(LD)$. We have $(DSP(x^*)) \equiv (MD(x^*))$ with respect to u_2 so if u^+ is optimal for $(DSP(x^*))$ then it is also optimal for $(MD(x^*))$.

$$v(MD(x^*)) = v_0^* = cx^* + dy + u_2^+(b - (A_2^1x^* + A_2^2y))$$

Then $v_0^* = v(MD(x^*)) = v(DSP(x^*)) = v(P)$ since $v(MP_F) = \max_{y \geq 0}\{v(SP(x^*))\}$ and x^* is optimal for (P). Suppose $v(SD(u_2^+)) \neq v(DSP(x^*))$, i.e.,

$$v(SD(u_2^+)) = cx^k + dy^k + u_2^+(b - (A_2^1 x^k + A_2^2 y^k)) > v_0^*$$

where (x^k, y^k) is an extreme point of $F(SD)$. Consider the master problem $(MD(x^*, k))$ which is the result of adding a cut involving (x^k, y^k) to $(MD(x^*))$. This cut would be

$$cx^k + dy^k + u_2(b - (A_2^1 x^k + A_2^2 y^k)) \leq v_0^*$$

since $v(MD(x^*, k)) \leq v(MD(x^*)) = v_0^*$ and $v(MD(x^*, k)) \leq v(LD) = v_0^*$. We know that u_2^+ is not feasible for $(MD(x^*, k))$. Let u_2^{++} be the optimal solution to $(MD(x^*, k))$. Then we have u_2^{++} feasible for $(MD(x^*))$ and optimal since

$$cx^k + dy^k + u_2^{++}(b - (A_2^1 x^k + A_2^2 y^k)) = v_0^*.$$

Therefore, since $(DSP(x^*)) \equiv (MD(x^*))$ with respect to u_2, there must be (u_2^{++}, u_3^{++}) optimal for $(DSP(x^*))$. There are a finite number of extreme points for $F(DSP(x^*))$ and it must be the case that there is some u_2^+ that is an extreme point of $F(DSP(x^*))$ and is optimal for $(DSP(x^*))$ such that $v(SD(u_2^+) = v(DSP(x^*))$.

Proof (a) $\rightarrow$ (c): Assume $v(P) = v(LD)$. We have $(SD(u_2^*)) \equiv (MP(u_2^*))$ with respect to x, so

$$v(MP(u_2^*)) = u_2^*(b - A_2^1 x^+) + cx^+ = w_0^*.$$

Then $w_0^* = v(MP(u_2^*)) = v(SD(u_2^*)) = v(LD) = v(P)$ since $v(LD) \leq v(SD(u_2^*) = v(MP(u_2^*)) \geq v(P)$. Suppose $v(DSP(x^+)) \neq v(SD(u_2^*))$. Then

$$v(DSP(x^+)) = cx^+ + u_2^k(b - A_2^1 x^+) < w_0^*.$$

where (u_2^k, u_3^k) is an extreme point of $F(DSP)$. Then we can form a new master problem, $(MP(u_2^*, k))$, by adding the following cut involving u_2^k to $(MP(u_2^*))$:

$$u_2^k(b - A_2^1 x) + cx \geq w_0^*$$

Because (u_2^k, u_3^k) is an extreme point of $F(DSP)$, $v(MP(u_2^*, k)) \geq v(MP(u_2^*))$. Clearly x^+ is not feasible for this new master problem. Let x^{++} be the optimal solution to $(MP(u_2^*))$. Then x^{++} is feasible and optimal for $(MP(u_2^*))$, so there must be a point (x^{++}, y^{++}) optimal for $(SD(u_2^*))$ such that $v(DSP(x^{++})) = v(SD(u_2^*))$. □

4.2 General Cross Decomposition Algorithm

Consider an algorithm where we solve consecutively

$$(DSPx^k), (SDu_2^{k+1}), (DSPx^{k+2}), (SDu_2^{k+3}), (DSPx^{k+4}), \ldots.$$

If x^k is equal to $x^{k+1}(= x^{k+2})$, then $v(DSPx^k) = v(P)$. So if x^k is not optimal, cycling could occur only after three iterations of the algorithm. Therefore, cycling will be prevented if a master problem is solved after every four iterations of the algorithm. A master problem can also be solved if the current solution does not improve the objective function value.

1. Initialize: Set $k := 0$. Set δ(flag for (LD) optimality) to 0, and set ξ(cycling counter) to 1. Let $\bar{v}_P$ = the best primal solution = $\bar{x}_0 = -\infty$ and $\bar{v}_D$ = best dual solution = $\bar{u}_0 = +\infty$. Select $u_2^0 \geq 0$.

2. Constraint (Dual) Subproblem.

   (a) Set $k := k + 1$. Solve (SDu_2^k); let x^k be an optimal solution.

   (b) If $\bar{v}_D > v(SDu_2^k)$, update $\bar{v}_D = v(SDu_k^2)$ and set $\xi = \xi + 1$. If $\bar{v}_D \leq \bar{v}_P$, stop. $\bar{x} = x^k$ is an optimal solution of (P). If $\bar{v}_D \leq \bar{x}_0$, set $\delta = 1$; i.e., (LD) is solved.

   (c) Convergence test CT_P: if $\xi = 4$, go to step 4a. Otherwise set $x^{k+1} := x^k$ and go to (3).

3. Variable (Primal) Subproblem

   (a) Set $k := k + 1$. Solve $(DSPx_2^k)$ and (SPx_2^k); let u^k and x_1^k be the optimal solutions.

   (b) If $\bar{v}_P < v(DSPx^k)$, update $\bar{v}_P = v(DSPx_2^k)$ and set $\bar{x} = x^k$ and set $\xi = \xi + 1$. If $\bar{v}_P \geq \bar{v}_D$ or $\bar{v}_P \geq \bar{u}_0$, stop; $\bar{x}$ is an optimal solution of (P).

   (c) Convergence test CT_D: If $\delta \neq 0$, go to step 4b. If $\xi = 4$, go to step 4b. Otherwise set $u_2^{k+1} = u_2^k$ and go to (2).

4. Master system.

(a) Solve the dual master problem (MD); let $(\bar{u}^0, u_2^{k+1})$ be an optimal solution. Set $\xi := 1$. If $\bar{v}_D > \bar{u}_0$, go to step 2. Otherwise, set $\delta = 1$ and go to (4b).

(b) Solve the primal master problem (MP); let $(\bar{x}_0, x_2^{k+1})$ be an optimal solution. Set $\xi := 1$. If $\bar{v}_D \leq \bar{x}_0$, stop; $\bar{x} = x_2^{k+1}$ is an optimal solution of (P). Otherwise, go to step 3.

4.2.1 Test for optimality in cross decomposition algorithm

Consider the test for optimality in the variable decomposition algorithm. Let (x^k, τ^k) be the optimal solution to (MPu^{k-1}) where u^{k-1} has been chosen (if $k = 1$) or is the solution to $(DSPx^{k-1})$. The basic step in the variable decomposition algorithm is as follows:

1. Solve $(DSPx_2^k)$ to get $u^k = (u_1^k, u_2^k)$.

2. Test (x^k, τ^k) for optimality: if

$$\tau^k \leq cx^k + u_1^k(b_1 - A_1^1 x^k) + u_2^k(b_2 - A_1^2 x^k)$$

then (x^k, τ^k) is optimal for (P).

3. If (x^k, τ^k) is not optimal, add cut w.r.t. u^k to master problem and go to (1).

We can write this test as

$$\tau^k \leq u_1^k(b_1 - A_1^1 x^k) + u_2^k(b_2 - A_1^2)x^k) + cx^k.$$

But $\tau^k = v(MPu^{k-1}) = v(SDu^{k-1}) = \bar{v}_D$ and

$$\bar{v}_P = v(DSPx_2^k) = u_1^k(b_1 - A_1^1 x^k) + u_2^k(b_1 - A_1^2 x^k) + cx^k.$$

So the cross decomposition test for optimality $\bar{v}_D \leq \bar{v}_P$ is a restatement of the variable decomposition test for optimality.

5 Cross decomposition for step-shaped disjunctive programs

5.1 Lagrangian relaxation

Recall the step-shaped program (L).

$$\begin{array}{lll} (L) & \max & cx + dy \\ & \text{s.t.} & A_1^1 x \leq r \\ & & A_2^1 x + A_2^2 y \leq b \\ & & A_3^2 y \leq 0 \\ & & x \in \cap_i \cup_j F_{ij} \end{array}$$

Theorem 8 *Let (L) be the step-shaped disjunctive program above. Then we have the following inequality:*

$$v(L) = \max_{\substack{x \in \cap \cup F_{ij} \\ A_1^1 x \leq r \\ A_2^1 x + A_2^2 y \leq b \\ A_3^2 y \leq 0}} \{cx + dy\} \leq \min_{u \geq 0} \left\{ ub + \max_{\substack{x \in \cap \cup F_{ij} \\ A_1^1 x \leq r \\ A_3^2 y \leq 0}} cx + dy - u(A_2^1 x + A_2^2 y)) \right\}$$

Proof. Let (x^*, y^*) be an optimal solution to (L) and let u^* be an optimal solution to the minimization above. Then

$$\begin{aligned} cx^* + dy^* &= u^*(A_2^1 x^*) + u^* A_2^2 y^* + cx^* - u^* A_2^1 x^* + dy^* - u^* A_2^2 y^* \\ &= u^*(A_2^1 x^* + A_2^2 y^*) + (c - u^* A_2^1)x^* + (d - u^* A_2^2)y^* \\ &\leq u^* b + cx^* - u^* A_2^1 x^* + dy^* - u^* A_2^2 y^* \\ &\leq u^* b + \max_{\substack{x \in \cap \cup F_{ij} \\ A_1^1 x \leq r \\ A_3^2 y \leq 0}} cx + dy - u^*(A_2^1 x + A_2^2 y)) \end{aligned}$$

since F_{ij} is a polyhedral set for all i, j and

$$\{x : x \in \cap \cup F_{ij}, A_1^1 x \leq r, A_2^1 x + A_2^2 y \leq b, A_3^2 y \leq 0\}$$

$$\subseteq \{x : x \in \cap \cup F_{ij}, A_1^1 x \leq r, A_3^2 y \leq 0\}. \quad \square$$

Therefore, the Lagrangian relaxation for (L) is

$$(LLD) \quad \min_{u \geq 0} \left\{ \max_{\substack{x \in \cap \cup F_{ij} \\ A_1^1 x \leq r \\ A_3^2 y \leq 0}} cx + dy + u(b - (A_2^1 x + A_2^2 y)) \right\}$$

The Lagrangian master problem for (L) is

$$\begin{aligned} (MLL) \quad \min \quad & \eta \\ \text{s.t.} \quad & \eta \geq cx^t + dy^t + u(b - (A_2^1 x^t + A_2^2 y^t)) \\ & u \geq 0 \end{aligned}$$

for all extreme points of $\{(x, y) \in \Re^{N+S} : A_1^1 x \leq r, A_3^2 y \leq 0\}$.

The subproblem is

$$\begin{aligned} (SLD) \quad \max \quad & cx + dy + \bar{u}(b - (A_2^1 x + A_2^2 y)) \\ \text{s.t.} \quad & A_1^1 x \leq r \\ & A_3^2 y \leq 0 \\ & x \in \cap \cup F_{ij} \end{aligned}$$

However, in this formulation, every subproblem is a disjunctive linear program with as many variables as our original problem. We will need to apply Lagrangian relaxation to the linear program for each disjunctive region, *i.e.*, to each (L_h).

$$\begin{array}{lll}(L_h) & \max & cx + dy \\ & \text{s.t.} & A_1^1 x \le r \\ & & A_2^1 x + A_2^2 y \le b \\ & & A_3^2 y \le 0 \\ & & x \in K_h\end{array}$$

As before, we can combine the constraints $A_1^1 x \le r$ with the requirements $x \in K_h$ and rewrite this constraint set as $E^h x \le e^h$. The Lagrangian master problem and subproblem with respect to the constraints $A_2^1 x + A_2^2 y \le b$ are below.

$$\begin{array}{lll}(ML_h) & \min & \eta \\ & \text{s.t.} & \eta \ge cx^t + dy^t + u(b - (A_2^1 x^t + A_2^2 y^t)) \\ & & u \ge 0\end{array}$$

for all extreme points of $\{(x, y) \in \Re^{N+S} : E^h x \le e^h, A_3^2 y \le 0\}$.

$$\begin{array}{lll}(SD_h) & \max & cx + dy + \bar{u}(b - (A_2^1 x + A_2^2 y)) \\ & \text{s.t.} & E^h x \le e^h \\ & & A_3^2 y \le 0\end{array}$$

5.2 Cross decomposition for (L)

We will rearrange the rows of the coefficient matrix A for (L_h) to ease the formulation for cross decomposition. Let $A_1^1 x + A_1^2 y \le r$ represent the constraint set $E^h x \le e^h$. Then A_1^2 is an $N+2 \times S$ zero matrix. After rearranging, we have the coefficient matrix A^*:

$$A^* = \begin{bmatrix} A_1^1 & A_1^2 \\ A_3^1 & A_3^2 \\ A_2^1 & A_2^2 \end{bmatrix}$$

Let the submatrices be denoted as below:

$$A_1^* = \begin{bmatrix} A_1^1 & A_1^2 \\ A_3^1 & A_3^2 \end{bmatrix} \quad A_2^* = \begin{bmatrix} A_2^1 & A_2^2 \end{bmatrix} \quad A^{1*} = \begin{bmatrix} A_1^1 \\ A_3^1 \\ A_2^1 \end{bmatrix} \quad A^{2*} = \begin{bmatrix} A_1^2 \\ A_3^2 \\ A_2^2 \end{bmatrix}$$

$$A = \begin{bmatrix} A_1^* \\ A_2^* \end{bmatrix} = \begin{bmatrix} A^{1*} & A^{2*} \end{bmatrix}$$

Cross decomposition is carried out with respect to a set of complicating variables and a set of complicating constraints. If we select A_2^* as the coefficients of the complicating constraints then the formulation of the constraint decomposition master problem and subproblem remain as before.

However, to show the cross decomposition equivalences for (L_h) it is necessary to express the variable decomposition master problem and subproblems in terms of the full $3S + N + 2$-dimension dual variables $\langle \vec{u}_1, \vec{u}_2, \vec{u}_3 \rangle$ instead of just the projections $(\vec{u}_2, \vec{u}_3)$. If x is chosen as the complicating variable, (L_h) could be written as:

$$\max_{x \geq 0} \left\{ \begin{array}{llcl} \max & cx + dy & & \\ \text{s.t.} & E^h x + 0y & \leq & e^h \\ & A_3^1 x + A_3^2 y & \leq & 0 \\ & A_2^1 x + A_2^2 y & \leq & b \end{array} \right\}$$

where x is assumed fixed in the inner maximization.

Then (L_h) is equivalent to

$$\max_{x \geq 0} \left\{ \min_{u \geq 0} \vec{u}_1(e^h - E^h x) + \vec{u}_2(b - A_2^1 x) + cx \right\}$$

$$\text{where } \vec{u}_1 A_1^2 + \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d$$

The master problem and subproblems then for variable decomposition are:

$$\begin{array}{llrcl}(MP_h) & \max & \tau & & \\ & \text{s.t.} & \tau & \leq & \vec{u}_1^t e^h + \vec{u}_2^t b - \vec{u}_1^t E^h x - \vec{u}_2^t A_2^1 x + cx\end{array}$$

$\forall u^t = \langle \vec{u}_1^t, \vec{u}_2^t, \vec{u}_3^t \rangle$ extreme points of $\{u \in \Re^M : \vec{u}_1 A_1^2 + \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$

$$\begin{array}{llrcl}(SP_h(\bar{x})) & \max & c\bar{x} + dy & & \\ & \text{s.t.} & A_2^1 x + A_2^2 y & \leq & b \\ & & A_3^2 y & \leq & 0\end{array}$$

$$\begin{array}{llrcl}(DSP_h(\bar{x})) & \min & \vec{u}_1(e^h - E^h\bar{x}) + \vec{u}_2(b - A_2^1\bar{x}) & & \\ & \text{s.t.} & \vec{u}_1 A_1^2 + \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 & = & d\end{array}$$

Define $(MP_h(\bar{\vec{u}}_2))$ as the relaxation of the master problem that involves all cuts that have $\bar{\vec{u}}_2$ fixed.

Claim: For fixed $\bar{\vec{u}}_2$, $(MP_h(\bar{\vec{u}}_2))$ gives the same optimal value for x as $(SD_h(\bar{\vec{u}}_2))$.

$$\begin{array}{llrcl}(MP_h(\bar{\vec{u}}_2)) & \max & \tau & & \\ & \text{s.t.} & \tau & \leq & \vec{u}_1(e^h - E^h x) + \bar{\vec{u}}_2(b - A_2^1 x) + cx\end{array}$$

$\langle \vec{u}_1, \bar{\vec{u}}_2, \vec{u}_3 \rangle$ an extreme point of the feasible region for $(DSP_h(\bar{x}))$

We can write $(MP_h(\bar{\vec{u}}_2))$ as

$$\max_{x \geq 0} \left\{ \begin{array}{lrcl} \min_{\vec{u}_1, \vec{u}_3 \geq 0} & \vec{u}_1(e^h - E^h\bar{x}) + \bar{\vec{u}}_2(b - A_2^1\bar{x}) + \vec{u}_3\vec{0} + c\bar{x} & & \\ \text{s.t.} & \vec{u}_1 A_1^2 + \bar{\vec{u}}_2 A_2^2 + \vec{u}_3 A_3^2 & \geq & d \end{array} \right\}$$

Since $\bar{x}$ and $\bar{\vec{u}}_2$ are fixed in the inner minimization above, when we take the dual of this inner minimization, we have

$$\max_{x \geq 0} \left\{ \begin{array}{lrcl} \max_{y \geq 0} & (d - \bar{\vec{u}}_2 A_2^2)y + \bar{\vec{u}}_2(b - A_2^1\bar{x}) + c\bar{x} & & \\ \text{s.t.} & y(0) & \leq & (e^h - E^h\bar{x}) \\ & yA_3^2 & \leq & 0 \end{array} \right\}$$

The problem now is stated in terms of x and y so it is equivalent to

$$\begin{array}{lll} (SD_h(\vec{\bar{u}}_2)) & \max & cx + dy + \vec{\bar{u}}_2(b - (A_2^1 x + A_2^2 y)) \\ & \text{s.t.} & E^h x \le e^h \\ & & A_3^2 y \le 0 \end{array}$$

It remains to show that for fixed $(x, y), (MD_h(x^t, y^t))$ gives the same optimal value for $(\vec{u}_2)$ as $(DSP_h(x^t))$. If $\bar{x}$ is fixed in $(MD_h(\bar{x}, y^t))$ then $(\bar{x}, y)$ is a solution of $(SD_h(\vec{\bar{u}}_2))$ for some fixed $\vec{\bar{u}}_2$ and $(\bar{x}, y)$ is an extreme point of the feasible region of $(SD_h(\vec{\bar{u}}_2))$. Define $(MD_h(\bar{x}))$ as the relaxation of the constraint master problem (MD) that has cuts involving all the extreme points $(\bar{x}, y^t)$ of the feasible region of $(SD_h(\vec{\bar{u}}_2))$ that is restricted by having $\bar{x}$ fixed. Then we have

$$(MD_h(\bar{x})) \quad \min_{\vec{u}_2 \ge 0} v^0$$

$$\text{such that } v^0 \ge c\bar{x} + dy^t + \vec{u}_2(b - (A_2^1 \bar{x} + A_2^2 y))$$

which is equivalent to

$$\min_{\vec{u}_2 \ge 0} \left\{ \begin{array}{lll} \max & c\bar{x} + dy + \vec{\bar{u}}_2(b - (A_2^1 x + A_2^2 y)) & \\ \text{s.t.} & 0y \le e^h - E^h \bar{x} & \\ & A_3^2 y \le 0 & \end{array} \right\}$$

since every $(\bar{x}, y)$ is an extreme point of the feasible region of $(SD_h \vec{\bar{u}}_2)$. By duality, and since $\bar{x}$ and $\vec{\bar{u}}_2$ are fixed in the inner max, we have

$$\min_{\vec{u}_2 \ge 0} \left\{ \begin{array}{ll} & \vec{u}_1(e^h - E^h \bar{x}) + \vec{u}_2(b - A_2^1 \bar{x}) + c\bar{x} + \vec{u}_3 \vec{0} \\ \text{s.t.} & \vec{u}_1 A_1^2 + \vec{u}_3 A_3^2 = d - \vec{\bar{u}}_2 A_2^2 \end{array} \right\}$$

which is equivalent to $(DSP_h \bar{x})$ and therefore dual to $(SP_h \bar{x})$.

From the equivalences we have demonstrated between $(MP_h\vec{\bar{u}}_2)$ and $(SD_h\vec{\bar{u}}_2)$ we have

$$\begin{aligned} v(MP_h\vec{\bar{u}}_2) &= \vec{u}_1(e^h - E^h x) + \vec{\bar{u}}_2(b - A_2^1 x) + cx \\ &= v(SD_h\vec{\bar{u}}_2) \\ &= cx + dy + \vec{\bar{u}}_2(b - A_2^1 x - A_2^2 y) \end{aligned}$$

This implies that $\vec{u}_1(e^h - E^h x) = (d - \vec{\bar{u}}_2 A_2^2)y$ at optimality. In our earlier discussion (2.1) we determined that, because of the zero submatrix A_1^2, the optimal solution to (DSP) would have $\vec{u}_1(r - A_1^1 x) = 0$ for all x feasible for (P). Since each (L_h) has the same form as (P) , we have that at optimality for (L_h), $\vec{u}_1(E^h x - e^h) = 0$ and, therefore, if $y_i \neq 0$, then $d_i = \vec{\bar{u}}_2 A_{2i}^2$.

6 Different formulations for cross decomposition

Cross decomposition relies on an equivalence between the constraint (Lagrangian) subproblem and the variable (Benders) master problem, and an equivalence between the variable (Benders) subproblem and the constraint master problem. Thus it avoids the necessity of accumulating cuts at every iteration of the algorithm. There are problems with this approach, however. If there is a strict duality gap, *i.e.*, $v(LD) - v(P) \neq 0$, then the constraint subproblem cannot be used in the test for optimality. The cross decomposition algorithm alternates between solving a Benders subproblem and a constraint subproblem; unconstrained dual variables in the Benders subroutine can cause a strict duality gap in the Lagrangian subproblem. Also, if any of the relaxed constraints are free for a subset of the variables, then decomposition can cause unboundedness in the primal variables; this leads to duality gaps and/or nonoptimal solutions. Analyzing a program (LP) in the context of cross decomposition illustrates some of the problems with this technique.

The presence of zero submatrices in the coefficient matrix can lead to unconstrained dual variables. To ease discussion of zero submatrices we will a general representation of the coefficient matrices for variable and constraint decomposition.

$$A = \begin{bmatrix} A^1 & A^2 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} A_1^1 & A_1^2 \\ A_2^1 & A_2^2 \end{bmatrix}$$

$A : (M + P) \times (N + S)$

$A^1 : (M + P) \times N$

$A^2 : (M + P) \times S$

$A_1 : M \times (N + S)$

$A_2 : P \times (N + S)$

$$\begin{array}{lll} (LP) & \max & cx + dy \\ & \text{s.t.} & A_1^1 x + A_1^2 y \leq b_1 \\ & & A_2^1 x + A_2^2 y \leq b_2 \end{array}$$

Theorem 9 *Let A^1 be the coefficient matrix for the variables (x) fixed in variable (Benders) decomposition, and let A_1 be the coefficient matrix of con-*

straints to be relaxed and added as a penalty function in constraint decomposition. Then if there is a zero submatrix of dimension $(k \times S), k \geq 1$, in matrix A_1^2 there will be a strict duality gap between the original problem and the constraint master problem, unless the constraint set that remains defines a vertex that is an optimal solution to the original problem.

Proof. The master problems and subproblems for this formulation are:

$$\begin{array}{lll} (MP(u^t)) & \max \tau & \\ & \tau & \leq cx + dy + \vec{u}_1^t(b_1 - A_1^1 x) + \vec{u}_2^t(b_2 - A_2^1 x) \\ & & \forall(\vec{u}_1^t, \vec{u}_2^t) \text{ extreme points of } F(DSP) \end{array}$$

$$\begin{array}{lll} (DSP(\bar{x})) & \min & c\bar{x} + \vec{u}_1(b_1 - A_1^1\bar{x}) + \vec{u}_2(b_2 - A_2^1\bar{x}) \\ & \text{s.t.} & \vec{u}_1 A_1^2 + \vec{u}_2 A_2^2 \geq d \end{array}$$

$$\begin{array}{lll} (SP(\bar{x})) & \max & c\bar{x} + dy \\ & \text{s.t.} & A_1^2 y \leq b_1 - A_1^1\bar{x} \\ & & A_2^2 y \leq b_2 - A_2^1\bar{x} \end{array}$$

$$\begin{array}{lll} (MD(x^t, y^t)) & \min \delta & \\ & \delta & \geq cx^t + dy^t + \vec{u}_1(b_1 - A_1^1 x^t - A_1^2 y^t) \\ & & \forall(x^t, y^t) \text{ extreme points of } F(SD) \end{array}$$

$$\begin{array}{lll} (SD(\bar{u}_1)) & \max & cx + dy + \vec{\bar{u}}_1(b_1 - A_1^1 x - A_1^2 y) \\ & \text{s.t.} & A_2^1 x + A_2^2 y \leq b_2 \end{array}$$

Consider, in turn, that there is a zero submatrix in A_1^2, A_2^2, and A^1. First, assume that k rows of A_1^2 are all zero; label these rows $m_1, \ldots, m_k$. Then the dual variables $u_{m_1}, \ldots, u_{m_k}$ are unconstrained in $(DSP(\bar{x}))$ for all x. We can add constraints for the extreme rays $\langle \vec{v}_1^s, \vec{v}_2^s \rangle$ of $F(DSP)$; *i.e.*, add the constraint

$$\vec{v}_1^s(b_1 - A_1^1\bar{x}) + \vec{v}_2^s(b_2 - A_2^1\bar{x}) \geq 0$$

for each extreme ray of $\{u \in R^{M+P} : \vec{u}_1(b_1 - A_1^1\bar{x}) + \vec{u}_2(b_2 - A_2^1\bar{x}) \geq 0\}$. Adding the constraints for the extreme rays will ensure that $v(DSP) > -\infty$; however there are only nonnegativity constraints on $u_{m_1}, \ldots, u_{m_k}$. So if $\bar{\vec{u}}_{1,j} = 0$ for any coordinate j is optimal for $(DSP\bar{x})$ and is then used as input to $(SD(\bar{\vec{u}}_1))$ then points that do not satisfy the constraints $A_{1,m_i}^1 x + A_{1,m_i}^2 y \leq b_{1,i}$ for $i = 1, \ldots, k$ can be selected as optimal solutions. A duality gap will occur if any of these constraints are necessary to define an optimal vertex solution to the original problem.

Next, if there are k rows of zeros in A_2^2, then k coordinates of $\vec{u}_2$ will be constrained only by nonnegativity constraints. However, this will not lead to a duality gap because $\vec{u}_2$ does not appear in the constraint subproblem or master problem. In this case, it is possible to solve (DSP) for the projection of $\vec{u}_2$ in R_+^{M+P-k}, instead of the full-dimensional $\langle \vec{u}_1, \vec{u}_2 \rangle$.

It is possible that there are zero submatrices in A_1^1 or A_2^1. However, in this case, if the rows of zeros are labelled $m_1, \ldots, m_k$, the coefficient of u_{m_i} in (DSP) will be the same as the coefficient of the variable u_{m_i} in the problem dual to (LP). So if (DSP) is unbounded, then this dual is unbounded and therefore (LP) is infeasible.

By a similar analysis, if y is selected as the complicating variable and A_2 as the matrix of complicating constraints, then there will be a duality gap with (LP) if there is a $(k \times N), k \geq 1$, zero submatrix in A_2^1, and any of the k constraints corresponding to this zero submatrix are necessary to define an optimal vertex solution to the original problem. □

Step-shaped programs have zero submatrices of the dimensions noted above, as well as a subset of constraints that is free for a subset of the variables. We have the following results concerning cross decomposition for these programs.

Lemma 4 *Let (LP) be the following step-shaped program:*

$$\begin{array}{lll} (LP) & \max & \alpha x + \beta y \\ & s.t. & Ax \leq c \\ & & -y \leq 0 \\ & & Dx + Ey \leq f \end{array}$$

Assume that the set $R = \{x \in \Re^N : Ax \leq c\} \neq \emptyset$ and R is bounded. Let the set of constraints $Dx + Ey \leq f$ be free for $x \in R$.

I. *If the Lagrangian relaxation for (LP) is formed with respect to the system $Dx + Ey \leq f$, then the Lagrangian subproblem for (LP) is*

$$\begin{array}{lll} (SD\bar{u}) & \max & \alpha x + \beta y + \bar{u}(f - Dx - Ey) \\ & s.t. & Ax \leq c \\ & & -y \leq 0 \end{array}$$

and three cases apply:

1. *If $\beta - \bar{u}E > 0$, then $(SD\bar{u})$ is unbounded.*
2. *If $\beta - \bar{u}E < 0$, then the optimal solution to $(SD\bar{u})$ will be $(x^*, 0)$ where x^* is the optimal solution to $\max\{\alpha x + \bar{u}(f - Dx) : Ax \leq c\}$.*
3. *If $\beta - \bar{u}E = 0$, then (x^*, y) is optimal for $(SD\bar{u})$ where x^* is the optimal solution to $\max\{\alpha x + \bar{u}(f - Dx) : Ax \leq c\}$ and $y \in \Re^S_+$.*

II. *Let the Lagrangian relaxation be formed with respect to the constraints $Ax \leq c$, $-y \leq 0$. Then the Lagrangian subproblem for (LP) is*

$$(SD^{*}\bar{u}) \quad \max \quad \alpha x + \beta y + \bar{\vec{u}}_1(c - Ax) + \bar{\vec{u}}_2 y$$
$$s.t. \quad Dx + Ey \leq f$$

If $\bar{u}_1 = 0$, then x is unconstrained in $(SD^{*}\bar{u})$.

Proof.

I. The only constraints on y in $(SD\bar{u})$ are the nonnegativity constraints. If the optimal objective function value for $(SD\bar{u})$ is finite, then the optimal solution will either be the extreme point $(x^{*}, 0)$ or there will be a ray of optimal solutions (x^{*}, y).

II. The system $Dx + Ey \leq f$ is free for $x \in \Re^{N}$. □

Lemma 5 *Let (LP) be as above. Assume that the cross decomposition algorithm is being carried out and that the variable decomposition problems with respect to x have been formulated. Then if $\bar{x}$ is such that $A\bar{x} < c$, the optimal solution to $(DSP\bar{x})$ will be $u = (\vec{u}_1, \vec{u}_2 u_3) = (0, \vec{u}_2, \vec{u}_3)$.*

Proof. $(DSP\bar{x})$ is the program

$$\min \quad \alpha\bar{x} + \vec{u}_1(c - Ax) + \vec{u}_2(f - Dx) + \vec{u}_3\vec{0}$$

$$\text{s.t.} \quad \vec{u}_1[0] + \vec{u}_2 E - \vec{u}_3 = \beta \quad , u \geq 0$$

By assumption $(c - Ax) > 0$. Therefore the optimal solution of $(DSP\bar{x})$ will be $u = (0, \vec{u}_2, \vec{u}_3)$. □

Theorem 10 *Assume that the cross decomposition algorithm is being carried out on the step-shaped program* (LP) *as defined in Lemma 4; variable decomposition is being done with respect to* x *and constraint decomposition is being done with respect to* $Ax \leq c, -y \leq 0$. *Let* $\bar{u}$ *be the optimal solution to* $(DSPx)$ *at iteration* m *of the cross decomposition algorithm, and assume that the optimal solution to* $(SD\bar{u})$ *is* $(\bar{x}, \bar{y})$ *where* $A\bar{x} < c$. *If there exists* $x^* \in \Re^N_+, y^* \in \Re^S_+$ *such that* $Ax^* > c, Dx^* + Ey^* \leq f$, *and* $\alpha x^* + \beta y^* > \alpha x + \beta y$ *for all* (x, y) *such that* $x \in R$ *and* $Dx + Ey \leq f$, *then either the dual subproblem* (DSP) *or the constraint master problem* (MD) *will be unbounded at iteration* $m + 2$ *of the cross decomposition algorithm.*

Proof. Let $\bar{x}$ be the optimal solution to $(SD\bar{u})$ or $(MP\bar{u})$ at iteration m where $A\bar{x} < c$. Then the solution to $(DSP\bar{x})$ will be $u_i^* = 0$ for k coordinates i of u^*. Consider the subproblem (SDu^*) at iteration $m + 1$. Since all $x \in \Re^N$ are feasible for (SDu^*), let (x^*, y^*) where $Ax^* > c$ be optimal for (SDu^*). Then at iteration $m + 2$,

$$v(MD) = \alpha x^* + \beta y^* + \vec{u}_1(c - Ax^*) + \vec{u}_3 y$$

$$= v(DSPx^*) = \alpha x^* + \vec{u}_1(c - Ax^*) + \vec{u}_2(f - Dx^*) = -\infty$$

since $\vec{u}_1$ is unconstrained in $(DSPx)$ and since $(DSPx) \equiv (MDx)$. □

6.1 Example of infeasibility in cross decomposition

Consider the following linear step-shaped program and its variable decomposition (with respect to x_1) and constraint decomposition (with respect to

$x_1 \leq 3$) master problems and subproblems.

$$\begin{array}{lll} (P) & \max & 3x_1 + 5x_2 \\ & \text{s.t.} & x_1 \leq 3 \\ & & x_1 + 2x_2 \leq 5 \\ & & x_2 \leq 2 \end{array}$$

$$\begin{array}{lll} (MP\bar{u}) & \max & \tau \\ & & \tau \leq 3x_1 + \bar{u}_1 * (3 - x_1) + \bar{u}_2 * (5 - x_1) + \bar{u}_3 * (2) \\ & & \forall (u_1, u_2, u_3) \text{ extreme points of } F(DSP) \end{array}$$

$$\begin{array}{lll} (DSP\bar{x}) & \min & u_1(3 - \bar{x}_1) + u_2(5 - \bar{x}_1) + u_3(2) \\ & \text{s.t.} & u_1(0) + u_2(2) + u_3 >= 5 \end{array}$$

$$\begin{array}{lll} (SP\bar{x}) & \max & 3\bar{x}_1 + 5x_2 \\ & \text{s.t.} & 0 \leq 3 - \bar{x}_1 \\ & & 2x_2 \leq 5 - \bar{x}_1 \\ & & x_2 \leq 2 \end{array}$$

$$\begin{array}{lll} (MDx_1, x_2) & \min & \delta \\ & \text{s.t.} & \delta \geq 3\bar{x}_1 + 5\bar{x}_2 + u_1(3 - \bar{x}_1) \\ & & \forall (x_1, x_2) \text{ extreme points of } F(SD) \end{array}$$

$$\begin{array}{lll} (SD\bar{u}_1) & \max & 3x_1 + 5x_2 + \bar{u}_1(3 - x_1) \\ & \text{s.t.} & x_1 + 2x_2 \leq 5 \\ & & x_2 \leq 2 \end{array}$$

In cross decomposition one alternates between solving an instance of (DSP) and an instance of $(SD\bar{u}_1)$ within four iterations of the algorithm. A master problem is solved after every four iterations. Consider carrying out this algorithm with the problem above: the extreme points of $F(DSP)$ are $(0,5,0)$ and $(0,0,5)$. Let $(\bar{x}_1, \bar{x}_2) = (1,2)$. Then the solution to $(DSP(1))$ is

$(\bar{u}_1, \bar{u}_2, \bar{u}_3) = (0, 0, 5)$. However, at the next iteration of the algorithm, $\bar{u}_1 = 0$ is used as input to $(SD(0))$. The extreme points of $F(SD)$ are $(0,0), (1,2)$, and $(5,0)$ and $\max\{3x_1 + 5x_2 + 0 * (3 - x_1)\}$ is at $x_1 = 5$ and $x_2 = 0$.

In the following step, either the Lagrangian master problem $(MD(5,0))$ or the Benders subproblem $(DSP(5))$ is solved. In either case, the objective function value is $-\infty$ and u_1 is unbounded.

In conclusion, the form of the constraint set for step-shaped programs can lead to infeasibility and unboundedness when the cross decomposition method is used. The presence of unconstrained dual variables in the Benders (variable) decomposition subproblem sets the stage for the selection of an infeasible point as optimal in the Lagrangian subproblem. If this occurs, then both the Lagrangian (constraint) master problem and variable (Benders) subproblem will be unbounded at the next step in the cross decomposition algorithm. If the constraint set for a step-shaped program also contains a subset of free constraints, and these constraints are relaxed in the Lagrangian subproblem, then infeasibility and unboundedness can also occur. When these constraints are added to the objective function as a penalty term, then, since some variables are not constrained by these inequalities, infeasible solutions for these variables are not penalized. However, if these free constraints remain in the constraint set for the Lagrangian subproblem, then again an infeasible solution can be found if the inequalities that do constrain the variables have been relaxed. In this case, also, infeasibility leads to unboundedness at the next step of the algorithm.

7 Fixed-Income Portfolios

(**Note:** A glossary of financial terms used in this section begins on page 64.)

Fixed-income portfolios are made up of instruments that involve cash payments at set dates followed by payment of principal at maturity. One way to express the worth of a fixed-income instrument is as the profit that accrues over a fixed holding period from re-investing the cash payments at the current spot interest rate [WoVZZe] [FaFa95]. This necessitates the development of a model for the evolution of the term structure of interest rates. Risk is the expectation of less than acceptable profit from an investment. Following Markowitz [Ma], variance of returns has been considered the measure of risk for equity instruments. However, Sharpe suggested mean absolute deviation of expected returns as an appropriate measure of risk. One of the advantages of this measure is the fact that in calculating mean absolute deviation less weight is given to outliers [Sh71]. Finally, since both the price and the profit from a fixed-income instrument depends on future interest rate movements, expected returns are usually calculated from an analysis of projected interest rates. Scenario analysis for fixed-income instruments often takes the place of historical analysis of returns for stock portfolios.

7.1 Fixed-income instruments

Fixed-income portfolios can contain U.S. Treasury securities, corporate bonds, mortgage-backed and asset-backed securities, and interest rate derivatives. These all make cash payments at set dates. The simplest fixed-income security is a zero-coupon Treasury bill, which pays its face value at maturity. Coupon-bearing Treasury bonds pay a percentage of the face value (the coupon rate) at set dates and the principal plus coupon at maturity. Certain issues of government bonds are callable, and municipal and corporate bonds are coupon-bearing and sometimes have call provisions for the issuer. Municipal bonds are often tax-free for residents of the issuing municipality. Convertible bonds include an option that allows the purchaser to convert the security to equity holdings. Mortgage-backed and asset-backed securities make varying cash payments since prepayment can affect the underlying cash stream. Valuing this cash stream involves valuing the option homeowners have to call the mortgage. Interest rate options and derivatives have cash payments based on domestic and foreign interest rate movements.

Cash-flow models are used to calculate the expected present value of holding a fixed-income security for a certain period. If the security has no embedded options, then the present value is simply the sum of the discounted future cash flows. However, if a fixed-income security has a call provision or a put provision, then the contribution of these options has to be included in the value of the security. The value of convertible bonds and mortgage-backed securities, for instance, includes the value of their embedded options.

7.2 Models of the term structure of interest rates

Many researchers have worked on models of the term structure of interest rates. Some models, like that of Black, Derman, and Toy, are recombining binomial lattices [BlDeTo]; however, these assume that the evolution of interest rates is path independent. Hull and White [HuWh] have developed several models; other researchers in this field are Vasicek [Va], Cox, Ingersoll, and Ross [CoInRo], and Ho and Lee [HoLe]. We will use the Heath-Jarrow-Morton model [HeJaMo] [Ja]. This model takes as input a set of forward rates and estimates of volatility and generates spot rates for every node of a binomial tree. The model uses arbitrage pricing theory and notions of equilibrium to develop the term structure of interest rates. In our model, these spot rates are used to value accrued reinvestment from holding-period cash-flows and to calculate the expected present value of outstanding securities at the end of the holding period for the portfolio.

7.3 Measures of risk

Markowitz pioneered portfolio analysis using variance of returns as the definition of risk [Ma]. He defined efficient portfolios as ones such that there is no portfolio with greater return and the same or less risk, and there is no portfolio with less risk and the same or greater return. An efficient portfolio is thus a Pareto optimal portfolio. In mean variance portfolio analysis, the ratio of expected return to variance is maximized to find the set of efficient portfolios. These problems are usually solved as quadratic programming problems. Markowitz also used semi-variance, the average of the squared deviations below the mean, as a measure of risk in portfolio analysis.

Sharpe proposed a linear programming algorithm for stock portfolio optimization in [Sh67]. He used the weighted sum of the absolute values of the sensitivities of a portfolio's securities to some common market index as the measure of the portfolio's risk. He presented a linear program for which the objective function was the parameterized difference between a portfolio's expected return and the absolute value of its sensitivity to the chosen index. This program was used in an algorithm to produce a *border* of efficient portfolios. Later, in [Sh71], Sharpe discussed using mean absolute deviation as a measure of dispersion of returns in analyzing portfolios. Konno [Ko] [KoYa] advocated using mean absolute deviation of returns in portfolio selection, since mean absolute deviation can be represented by a piecewise linear risk function that can be minimized using linear programming techniques.

The traditional measure of risk for a fixed-income instrument is its *duration*, or the sensitivity of its price to a change in interest rates [ElGr]. Duration is usually expressed as the first order term in the Taylor series expansion of the present value function for a security's cash flows, and thus is a good approximation for small changes in yield when the changes in interest rates correspond to parallel shifts in the yield curve. Researchers have pointed out, however, that there is no justification for assuming that shifts in the yield curve will be parallel. Therefore, a better measure of interest-rate risk for fixed-income instruments is based on changes in yield corresponding to the evolution of the term structure of interest rates [Ja] [FaFa95]. Zenios and Kang used mean absolute deviation of returns on interest-rate scenarios in portfolio optimization for mortgage-backed securities [ZeKa]. Hiller and

Eckstein used an objective function that maximizes the trade-off between expected return and the average downside deviation on scenarios in a model to select the optimal asset/liability-dedicated fixed-income portfolio [HiEc].

One method of optimal portfolio selection involves measuring risk as the deviation from a benchmark portfolio or an index of assets. Equity index funds may, for example, interpret risk as deviation from the return of the S&P500. In the fixed-income market, a portfolio manager may pick a particular portfolio that he feels will have a desired performance as his benchmark; or the benchmark may be a fixed-income index, *e.g.* the Lehman Government Corporate Index. Sometimes the Treasury yield curve will be used as a benchmark. However, once a benchmark is chosen, selecting an optimal portfolio involves minimizing the risk of returns different from the benchmark's, *i.e.* finding the portfolio that tracks the benchmark best. Usually, this difference in returns is calculated on many scenarios using one of the previously mentioned measures of risk. Worzel *et al* minimized the mean absolute deviation of portfolio and benchmark returns in [WoVZZe]. Here we will use mean absolute deviation from a benchmark, on interest-rate scenarios, as the measure of risk; and the objective we will maximize is the difference between expected return and risk. Mean absolute deviation is linearized and thus selecting the optimal portfolio is now a large linear problem.

7.4 Logical requirements for portfolio selection

The choice of an optimal portfolio is often subject to constraints other than budget ones. There may be integral trading amount requirements for certain instruments, *e.g.* Treasury bills have to be purchased in increments of $5,000, after an initial investment of $10,000. Investors may demand a certain level of diversification, defined according to disjunctions, or a certain mix of credit grades for their portfolios. Some institutions and individuals have to satisfy regulatory agency requirements. There are limits, for instance, on the assets and liabilities pension funds can hold. All these are logical requirements; they can be modeled with disjunctions that describe allowable choices. However, introducing logical requirements into a model usually makes an otherwise linear and convex problem into a nonconvex programming problem, a disjunctive linear programming problem.

7.4.1 Glossary

arbitrage pricing theory A theory that takes the prices of a primary set of traded assets as given, as well as their stochastic evolution, and then prices a secondary set of traded assets. It prices the secondary assets by constructing a portfolio of the primary assets that replicates the returns of the secondary asset. To prevent riskless profit or arbitrage, the price of the replicating portfolio must equal the price of the secondary asset.

asset-backed securities Bonds or notes backed by loan paper or accounts receivable originated by banks, credit card companies, or other providers of credit.

call provision If a bond has a call provision or feature, the issuer retains the right to retire the debt, fully or partially, before the scheduled maturity date. It permits the borrower, should market rates fail, to replace the bond issue with a lower-interest-cost issue.

credit grade Bond ratings published by organizations that do bond analysis; ratings reflect analysis of probability of future payments, *i.e.* risk of default. Ratings run from AAA to D (S&P); the highest grade is called gilt-edge or maximum security.

dedication A portfolio strategy that matches monthly cash flows from a portfolio of bonds to a prespecified set of monthly cash requirements or liabilities.

dispersion of returns The spread of returns, or the extent to which a distribution is concentrated around a single value. Dispersion is usually measured by fractiles (quartiles, etc.).

downside deviation The amount that a portfolio's return is below a certain level.

equity Ownership interest possessed by shareholders in a corporation – stock as opposed to bonds.

forward rate The rate contracted at time t for the period $(T, T+1)$.

$$f(t,T) = \frac{\text{price of zero-coupon bond at } t \text{ that matures at T}}{\text{price of zero-coupon bond at time } t \text{ that matures at } T+1}$$

holding period The period of time a portfolio is held without selling.

index A statistical composite that measures changes in the economy or financial markets.

interest rate risk The risk for an investor who has to sell a fixed-income security before maturity that interest rates will rise; *i.e.* that he will realize a capital loss at sale.

mortgage-backed securitiy A security backed by mortgages; investors receive payments out of the interest and principal on the underlying mortgages.

option-adjusted spread The amount that yield of a security with embedded options varies from some well-known instrument, typically, the forward curve.

Pareto optimality The concept that resources are optimally distributed when an individual cannot move into a better position without putting someone else in a worse position.

present value The value today of a future payment, or stream of payments, discounted at some appropriate compound interest or discount rate.

put provision A provision that grants the investor the right to sell the issue back to the issuer at par value at designated dates.

reinvestment risk The variability in the returns from reinvestment from a given strategy due to changes in the market rates.

scenario A projected sequence of events; in fixed-income analysis, a particular sequence of interest-rate movements.

short sale The sale of a security or commodity futures contract not owned by the seller.

spot rate The rate contracted at time t on a one-period riskless loan starting immediately; equal to the forward rate with $T = t$.

term structure The relationship between yield and maturity.

volatility The characteristic of a security, commodity, or market to rise or fall sharply in price within a short period.

yield curve A graph that plots the yields of all bonds of the same quality with their maturities, ranging from the shortest to the longest available.

8 The absolute deviation trade-off model for fixed-income portfolio selection

Konno [KoYa], Zenios and Kang [ZeKa], and Worzel *et al* [WoVZZe] have developed models for portfolio optimization that use the absolute value of the deviation of a portfolio's returns from the mean as a measure of risk. Their models have as objective the minimization of the expected absolute deviation of returns. Hiller and Eckstein [HiEc] introduced a model that used downside deviation from a liability stream as the measure of risk. Auxiliary variables are introduced in each of these models to represent the amount of risk, and the programs for all these models are examples of step-shaped programs. Each model has two sets of variables, portfolio variables and risk variables; the constraints linking the two sets are free for the portfolio variables; and the coefficient matrices for these models can be partitioned so that there is a submatrix with zero entries for all the risk variables and a submatrix with zero entries for all the portfolio variables.

The absolute deviation trade-off model is an extension of these earlier mean absolute deviation models. We use the L_1 distance from a given benchmark as the measure of risk and maximize the trade-off between the expected present value of reinvested cash flows and the difference between benchmark and portfolio returns. This objective function is similar in spirit to one utilized by Hiller and Eckstein [HiEc]. We will first detail the linear version (P) of our model; we present the disjunctive version in Section 9.2. One obvious advantage in using a mean absolute deviation model instead of a quadratic model is that a MAD model can be formulated as a linear program. Another advantage is that disjunctive linear programs can be solved by techniques such as those of Section 3. Quadratic $0-1$ and general integer programming methods are still in development and encounter complexity theoretic problems [GJ]. Also, in many cases, optimal portfolios for these LP-based models have many fewer nonzero holdings than do portfolios found optimal for quadratic models [Ko] [Bi]. This is important in portfolio optimization because transaction costs occur for every separate purchase. In fact, we show in Section 9.1 that among the extreme points of the polyhedron representing feasible solutions of (P) are portfolios with a minimal number of nonzero holdings.

For our model we assume that we have a fixed universe of available instruments and that there is a cash-flow model available to calculate the present value of each security's cash payments and embedded options. Further, we have fixed a benchmark with known returns on the set of scenarios. We assume that the only constraints are that all asset holdings are nonnegative, *i.e.*, there is no short selling, and there is a set budget for the investment. We also assume that we have a model of the term structure of interest rates of the Heath-Jarrow-Morton type and that from this model we have computed short rates at every node of a binomial tree. Also, we have calculated the value of reinvesting the cash flows from each of the available assets at the current spot rate at every node. If there are N available assets and the holding period for the portfolio comprises M dates on which cash flows are collected, then there are 2^M scenarios. If the number of scenarios becomes too large, sampling techniques or Monte Carlo simulation methods are used in conjunction with the basic model. In this research, we have used explicit computation on a discrete model of the future evolution of spot rates [Ja].

For every asset, the present value of investing the cash flows at the current spot rate over the holding period of the portfolio is calculated. This figure is then "normalized" by multiplying by the price of the asset (per $100 face) divided by the budget amount. A feasible portfolio is a linear combination of the assets for which the price does not exceed the budget. A variable y_i is introduced for every scenario i to model the absolute value of the difference between a portfolio and the benchmark yielded by this scenario. Two constraints are added to the constraint set for every scenario:

$$\forall i \quad (\sum_{j=1}^{N} ret_{ij}x_j) - y_i \leq benchmark_i$$

$$\forall i \quad (-\sum_{j=1}^{N} ret_{ij}x_j) - y_i \leq -benchmark_i$$

The objective function then is the difference between expected return and expected risk:

$$\max \frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{N} ret_{ij}x_j - \frac{1}{S}\sum_{i=1}^{S} y_i$$

At optimality, since the y_i are nonnegative, subtracting the average of the y_i is equivalent to minimizing the y_i, or forcing y_i to the absolute value of the difference between benchmark and portfolio on each scenario. We now have a linear problem with linear constraints instead of a piecewise linear problem. The definitions and constraints for the model are as follows:

Definitions:

S number of scenarios: 2^T where T is investment horizon

N number of securities available

B budget limit

ret_{ij} present value of investing eash flow from security j on scenario i

$benchmark_i$ return from benchmark portfolio on scenario i

p_j current price of \$100 face amount of security j

x_j amount held of security j

y_i absolute value of the deviation of the portfolio return from the benchmark return on scenario i

v minimum trade amount

w minimum investment amount

The linear program that models the problem of selecting the best tracking portfolio is:

$$\text{(P)} \quad \max\{\frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{N} ret_{ij}x_j - \frac{1}{S}\sum_{i=1}^{S} y_i\} = z_P$$

$$
\begin{array}{lrcl}
\text{s.t.} & p_1x_1+\ldots+p_Nx_N & \le & B \\
 & -x_1-\ldots-x_N & \le & -d \\
\forall j & -x_j & \le & 0 \\
\forall i & ret_{i1}x_1+\ldots+ret_{iN}x_N-y_i & \le & benchmark_i \\
\forall i & -ret_{i1}x_1-\ldots-ret_{iN}x_N-y_i & \le & -benchmark_i \\
\forall i & -y_i & \le & 0
\end{array}
$$

(P) can be written formally as:

$$
\begin{array}{llll}
(P) & \max & \{cx+dy\} & = & z_P \\
 & \text{s.t.} & A_1^1x+A_1^2y & \le & r \\
 & & A_2^1x+A_2^2y & \le & b \\
 & & A_3^1x+A_3^2y & \le & 0
\end{array}
$$

or, in full detail,

$$
\begin{array}{lccclccclc}
A_1^1 = & p_1 & \cdots & p_N & A_1^2 = & 0 & \cdots & 0 & r = & B \\
 & -1 & -1 & -1 & & 0 & \cdots & 0 & & -w \\
 & -1 & \cdots & 0 & & 0 & \cdots & 0 & & 0 \\
 & \vdots & -1 & \vdots & & \vdots & 0 & \vdots & & \vdots \\
 & 0 & \cdots & -1 & & 0 & \cdots & 0 & & 0 \\
A_2^1 = & ret_{11} & \cdots & ret_{1N} & A_2^2 = & -1 & \cdots & 0 & b = & benchmark_1 \\
 & \vdots & \vdots & \vdots & & \vdots & -1 & \vdots & & \vdots \\
 & ret_{S1} & \cdots & ret_{SN} & & 0 & \cdots & -1 & & benchmark_S \\
 & -ret_{11} & \cdots & ret_{1N} & & -1 & \cdots & \cdots & & -benchmark_1 \\
 & \vdots & \vdots & \vdots & & 0 & -1 & 0 & & \vdots \\
 & -ret_{S1} & \cdots & -ret_{SN} & & 0 & \cdots & -1 & & -benchmark_S \\
A_3^1 = & 0 & \cdots & 0 & A_3^2 = & -1 & \cdots & 0 & & 0 \\
 & \vdots & 0 & \vdots & & \vdots & -1 & \vdots & & \vdots \\
 & 0 & \cdots & 0 & & 0 & \cdots & -1 & & 0
\end{array}
$$

We have the following results for our model:

Lemma 6 *The optimal objective function value for (P) is less than or equal to the average over all scenarios of the return of the benchmark portfolio.*

Proof. Assume there is some portfolio $\bar{x}$ for which

$$\frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{N} ret_{ij}x_j > \frac{1}{S}\sum_{i=1}^{S} benchmark_i$$

Let $\bar{z}$ be the objective function value for $\bar{x}$. Consider the case first where for all i such that $y_i > 0$, $\sum_{j=1}^{N} ret_{ij}\bar{x}_j > benchmark_i$. Then $y_i = \sum_{i=1}^{N} ret_{ij}\bar{x}_j - benchmark_i$ and

$$\begin{aligned}
\bar{z} &= \frac{1}{S}\left\{\sum_{j=1}^{N} ret_{ij}\bar{x}_j - (\sum_{j=1}^{N} ret_{ij}\bar{x}_j - benchmark_i)\right\} \\
&= \frac{1}{S}\sum_{i=1}^{S} benchmark_i
\end{aligned}$$

Next consider the case when for some y_k such that $y_k > 0$ we have $benchmark_k > \sum_{j=1}^{N} ret_{kj}\bar{x}_j$. Then $y_k = benchmark_k - \sum_{j=1}^{N} ret_{kj}\bar{x}_j$ and

$$\begin{aligned}
\bar{z} &= \frac{1}{S}\left\{\sum_{i\neq k} ret_{ij}\bar{x}_j - \left(\sum_{j=1}^{N} ret_{ij}\bar{x}_j - benchmark_i\right) + \left(\sum_{j=1}^{N} ret_{kj}\bar{x}_j\right) - y_k\right\} \\
&= \frac{1}{S}\left\{\sum_{i\neq k} benchmark_i + \left(\sum_{j=1}^{N} ret_{ij}\bar{x}_j\right) - y_k\right\} \\
&< \frac{1}{S}\sum_{i=1}^{S} benchmark_i
\end{aligned}$$

since $\sum_{j=1}^{N} ret_{kj} - y_k < benchmark_k$. □

Lemma 7 *For $x \in \Re^N$, let $f_i(x) : \Re^N \to \Re$ for $i = 1, \ldots, S$ be the sequence of functions $f_i(x) = |benchmark_i - \sum_{j=1}^{N} ret_{ij}x_j|$. Then for every $x \in \Re^N$, the point $(x_1, \ldots, x_N, f_1(x), \ldots, f_S(x))$ satisfies S inequalities from the constraint set $A_2^1 x + A_2^2 y \leq b$ of (P) with equality.*

Proof. Let $x \in \Re^N$. Consider the vector $(x_1, \ldots, x_N, y_1, \ldots, y_S) \in \Re^{N+S}$, where $y_i = f_i(x)$ for all i. If on the ith scenario, $\sum_{j=1}^{N} ret_{ij} x_j > benchmark_i$, then the ith constraint in $A_2^1 x + A_2^2 y \leq b$ is satisfied with equality, and the $(S+i)$th constraint

$$-\sum_{j=1}^{N} ret_{ij} x_j - y_i \leq -benchmark_i$$

is redundant. If $benchmark_i > \sum_{j=1}^{N} ret_{ij} x_j$, then the $(S+i)$th constraint is satisfied with equality and the ith constraint

$$\sum_{j=1}^{N} ret_{ij} x_j - y_i \leq benchmark_i$$

is redundant. If both the ith constraint and the $(S+i)$th constraint are satisfied with equality for any i, then $y_i = 0$ since otherwise $A_2^1 x + A_2^2 y \leq b$ is inconsistent. □

9 Characterizing optimal portfolios

Recall the coefficient matrix for the linear portfolio selection problem (P) of Section 8. Let M be the set of vectors feasible for (P). M is a polyhedron. Let C be the characteristic cone for M. By definition, C is the set

$$C = \{(z_1, z_2) \in \Re^{N+S} : (x, y) + (z_1, z_2) \in M \; \forall \, (x, y) \in M\}$$

In our case, we have

$$C = \left\{ (z_1, z_2) \in \Re^{N+S} : \begin{array}{lcl} A_1^1 z_1 & \leq & 0 \\ A_2^1 z_1 + A_2^2 z_2 & \leq & 0 \\ A_3^2 z_2 & \leq & 0 \end{array} \right\}$$

Therefore, $\{(z_1, z_2) \in C\} = \{(0, z_2) \in \Re^{N+S} : z_2 \geq 0\}$. The lineality space of M is

$$C \cap -C = \left\{ (z_1, z_2) \in \Re^{N+S} : \begin{array}{lcl} A_1^1 z_1 & = & 0 \\ A_2^1 z_1 + A_2^2 z_2 & = & 0 \\ A_3^2 z_2 & = & 0 \end{array} \right\} = 0$$

So M is pointed and C is pointed. M can be decomposed into a polytope Q and the cone C, where the polytope Q = the convex hull of the vertices of M. Let $G = \{x \in \Re_+^N : A_1^1 x \leq r\}$.

Theorem 11 *Assume $w > 0$. If (x^*, y^*) is a vertex of M, then (x^*, y^*) satisfies S constraints of $\{A_2^1 x + A_2^2 y \leq b\}$ with equality. Therefore, if (x, y) is a vertex of M, $y_i = f_i(x) \; \forall \, i$.*

Proof. The rank of the coefficient matrix for M is $N + S$ since the matrix has $N + S$ columns and an $N \times N$ diagonal submatrix disjoint from an $S \times S$ diagonal submatrix. Let (x, y) be a vertex of M. Then (x, y) satisfies $N + S$

constraints with equality, where if $Ax + By \leq c$ is the set of constraints satisfied with equality, then $Ax + By = c$ has $N + S$ independent rows.

Assume (x, y) satisfies fewer than S constraints from $A_2^1 x + A_2^2 y \leq b$ with equality. Denote the submatrix consisting of the first S rows of $\left(A_2^1 \; A_2^2 \right)$ as A^* and the submatrix consisting of the second S rows of this matrix as A^{**}. At most N constraints of G can be satisfied with equality by any point x: either $N - 1$ bounding constraints and the budget constraint, or $N - 1$ bounding constraints and the minimum investment constraint. Of the remaining constraints, only $2S$ can be in a system of independent equations since $A_3^2 = \left(0 \; I \right) * A^* = \left(0 \; I \right) * A^{**}$. If (x, y) is a vertex and (x, y) satisfies fewer than S constraints from $A_2^1 x + A_2^2 y \leq b$ with equality, then (x, y) must satisfy at least one bounding constraint for y with equality. But for every constraint of $A_3^2 y \leq 0$ satisfied with equality by (x, y) there is a constraint of $A_2^1 x + A_2^2 y \leq b$ satisfied with equality if (x, y) is feasible, and only one of these can be included in a system of independent equations. Therefore, it must be that if (x, y) is a vertex of M, then (x, y) satisfies S constraints of $A_2^1 x + A_2^2 y \leq b$ with equality. □

Therefore, if (x, y) is a vertex of M, $y_i = f_i(x)$, $\forall\, i$.

Theorem 12 *Let (x^*, y^*) be a vertex of M. Then either x^* is a vertex of G or for at least one coordinate i of $y^*, y_i^* = 0$.*

Proof. First consider the case when x^* is a vertex of G. Then x^* satisfies N constraints from G with equality. Next consider the case where (x', y') is a vertex of M and x' is not a vertex of G. Then (x', y') must satisfy N

constraints from $H = \{A_1^1 x \leq r,\ A_3^2 y \leq 0\}$ with equality, but fewer than N from $\{A_1^1 x \leq r\}$. Therefore, at least one constraint from $\{A_3^2 y \leq 0\}$ is held with equality by (x', y'). If (P) has degenerate solutions, then there will be vertices $(\bar{x}, \bar{y})$ of M in which $\bar{x}$ is a vertex of G and for some i, $\bar{y}_i = 0$. □

Theorem 13 *An optimal vertex solution to the linear portfolio selection problem will either have full investment in one asset or will achieve the same return as the benchmark on at least one scenario.*

Proof. The vertices of G are the intersections of the halfspaces represented by the constraints $A_1^1 x \leq r$. These constraints are:

(1) $\sum_{j=1}^{N} p_j x_j \leq B$

(2) $\sum_{j=1}^{N} x_j \geq w$

(3) $x_1 \geq 0$

⋮

(N+2) $x_N \geq 0$

The vertices of intersections of these halfspaces will be points $x = x_1, \ldots, x_N$ where $x_j = 0$ for $N-1$ of the N coordinates and $x_k = w$ or $x_k = B/p_k$ for the remaining coordinate. Let $x^* = (0, \ldots, B/p_i, 0, \ldots, 0)$, *i.e.* the portfolio with the total budget invested in the ith asset, and $x' = (0, \ldots, 0, \ldots, B/p_j, 0)$ the portfolio with the budget invested in the jth asset. Let $0 \leq \lambda \leq 1$. Then a portfolio $\bar{x}$ which satisfies $p_i\bar{x}_i + p_j\bar{x}_j = B$ can be expressed as

$\bar{x} = \lambda x^* + (1-\lambda)x'$. Similarly, let $\tilde{x} = (0,\dots,w,\dots,0)$, or the portfolio with only the minimum amount w invested in the ith asset, and $\dot{x} = (0,\dots,0,\dots,w,0)$, or the portfolio with only the minimum amount invested in the jth asset. Then a portfolio $\hat{x}$ satisfying $\hat{x}_i + \hat{x}_j = w$ can be expressed as $\hat{x} = \lambda\tilde{x} + (1-\lambda)\dot{x}$. Therefore, the vertices of G are single-asset portfolios. □

9.1 Multiple optimal portfolios

A condition for uniqueness of the solution to (P) is nondegeneracy in the dual to (P) [GiMuWr]. Since the constraints in (P) correspond to constraints in (MP^0) and $(SP(x^k))$, the constraints in (P) active at optimality will also be active in (MP^0) and $(SP(x^k))$ at optimality, where k is the final stage of the decomposition algorithm. Therefore, the presence of fewer than S coordinates of the optimal dual multiplier (u_2^k, u_3^k) at nonzero values implies degeneracy in $(DSP(x^k))$ and thus multiple optimal solutions to (P). Counting the number of nonzero coordinates of u^k can provide a test for uniqueness of the solution to (P).

Generally, if there are multiple optimal solutions, we know that more than $N+S$ constraints of (P) are satisfied with equality. So if an optimal portfolio is not a pure portfolio, then for some $i, i \in 1,\dots,S$, $y_i = 0$.

The occurrence of multiple optimal solutions in this problem provides information about the range of optimal investments: a kind of "efficient frontier" of investments with different levels of risk and return is delineated by these multiple solutions. If the solution to (P) is unique, then introducing a parameter τ to vary allowable levels of risk can provide a range of investments from more to less risky. In this approach, τ is increased until the portfolio with the least risk is found optimal *c.f* [HiEc]. Conversely, introduction of the parameter τ will lead to discovery of multiple optimal solutions with different levels of risk if they exist.

9.2 The portfolio problem with disjunctive requirements

As mentioned earlier, fixed-income investments usually have to be made in set amounts and at minimum levels. Requiring that all nonzero holdings in a portfolio are at least v is one example of a logical or disjunctive requirement. If we add this logical requirement to (P), we have a disjunctive program that models selecting the optimal tracking portfolio that respects the minimum trade amount requirement. The disjunctive requirement applies only to the asset holdings and not to portfolio risk so the constraints on x and y together and on y alone remain the same.

A program with disjunctive requirements on the quantities modeled by the x variables, linear constraints on the x and y variables together, and bounding constraints on the y variables, can be expressed in the general form:

$$\begin{array}{lrcl} \max & \{cx + dy\} & & \\ \text{s.t.} & Ex & \leq & e \\ & Gx + Hy & \leq & g \\ & x & \in & \cap_i \cup_j F_{ij} \end{array}$$

By way of example, if we express the program that results from adding threshold requirements to (P) in this form, we have the following program (L). $Ex \leq e$ describes the constraints on x alone.

$$\text{(L)} \quad \max\{\frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{N} ret_{ij}x_j - \frac{1}{S}\sum_{i=1}^{S} y_i\} = z_L$$

$$\begin{array}{lrcl}
\text{s.t.} & Ex & \leq & e \\
& \forall i \;\; ret_{i1}x_1 + \ldots ret_{iN}x_N - y_i & \leq & benchmark_i \\
& \forall i \;\; - ret_{i1}x_1 - \ldots - ret_{iN}x_N - y_i & \leq & -benchmark_i \\
& \forall i \;\; - y_i & \leq & 0 \\
& \forall j \;\; x_j = 0 \vee x_j \geq v & &
\end{array}$$

As discussed in Section 3, the optimal portfolio for (L) can be found by solving a series of linear programs (L_k) where each (L_k) corresponds to a particular leaf of (L); *i.e.*, each (L_k) defines a feasible region where every portfolio is in the kth disjunctive region (out of M^N possible regions) and satisfies the disjunctive requirements.

$$(\text{L}_k) \quad \max\,\{\frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{N} ret_{ij}x_j - \frac{1}{S}\sum_{i=1}^{S} y_i\} = z_{L_k}$$

$$\begin{array}{lrcl}
\text{s.t.} & p_1x_1 + \ldots + p_Nx_N & \leq & B \\
& -x_1 - \ldots - x_N & \leq & -w \\
& \forall j \;\; \alpha_{k(j)}x_j & \leq & \beta_{k(j)} \\
& \forall i \;\; ret_{i1}x_1 + \ldots + ret_{iN}x_N - y_i & \leq & benchmark_i \\
& \forall i \;\; - ret_{i1}x_1 - \ldots - ret_{iN}x_N - y_i & \leq & benchmark_i \\
& \forall i \;\; - y_i & \leq & 0
\end{array}$$

where $\alpha_{k(j)} = 1$ if $\beta_{k(j)} = 0$, *i.e.* if the holdings in asset j equal 0; $\alpha_{k(j)} = -1$ if $\beta_{k(j)} = -v$, *i.e.* if there is a nonzero holding in asset j. The only difference between the constraints for (P) and any (L_k) is in the rows $E^k x \leq e^k$, where

$$E^k = \begin{matrix} p_1 & \cdots & p_N \\ -1 & -1 & -1 \\ \alpha_{k(1)} & \cdots & 0 \\ \vdots & \alpha_{k(j)} & \vdots \\ 0 & \cdots & \alpha_{k(N)} \end{matrix} \qquad e^k = \begin{matrix} B \\ -w \\ \beta_{k(1)} \\ \beta_{k(j)} \\ \beta_{k(N)} \end{matrix}$$

(L) is a disjunctive step-shaped program and each (L_k) is a step-shaped program so we can use either the combined breadth-first decomposition algorithm or the combined depth-first decomposition algorithm to solve (L). In either case, the full master problem for a particular (L_k) is:

$$\begin{array}{llrcl} (\mathrm{ML}_k) & \max & cx + \lambda & = & z_{ML_k} \\ & \text{s.t.} & E^k x & \leq & e^k \\ & & \lambda & \leq & \vec{u}_2^t(b - A_2^1 x) \\ & & & & \text{for } \vec{u}_2^t = \pi_2(u^t) \end{array}$$

where $u^t = \langle \vec{u}_2^t, \vec{u}_3^t \rangle$ is an extreme point of the set $\{u \in R_+^{3S} : \vec{u}_2 A_2^2 + \vec{u}_3 A_3^2 = d\}$.

Recall that M denotes the set of vectors feasible for (P). Let M_k equal the set of vectors feasible for (L_k), for each k. Let $H_k = \{x \in \Re^N : E^k x \leq e^k\}$, where the rank of E is N. Since a vertex of M_k must satisfy $N+S$ constraints from $\{E^k x \leq e^k, A_2^1 x + A_2^2 y \leq b, A_3^2 y \leq 0\}$ with equality, we have, as in Theorem 12, that if (x, y) is a vertex of M_k, either x is is a vertex of H_k or $y_i = 0$ for at least one i. Each M_k is a polyhedron that differs from M only in that each coordinate of x is either equal to 0 or greater than v, the minimum trade amount. So the set of vertices of the convex hull of $\cup_k M_k$ contains vertices (x^*, y^*) of M such that $x_i^* = 0$ or $x_i^* \geq v$, $\forall\, i$.

Consider first that $w \geq v$. Then (x, y) such that x is a vertex of G and $y_i = f_i(x)$ for all i are among the vertices of the convex hull of $\cup_k M_k$ since for each of these vertices, the vector x satisfies N constraints with equality for some H_k. The only other possible vertices of $\cup_k M_k$ are points (x, y) where $x = x_1, \ldots, x_N$ such that $x_i = v$ for some i and $x_j = 0$ for all $j \neq i$ and $y_i = f_i(x)$ for all i. If these points are feasible for $E^k x \leq e^k$ they are also among the vertices of the convex hull of $\cup_k M_k$.

Next, let $w < v$. Then (x, y) such that x is a vertex of G and $x_i = B/p_i$ and $x_j = 0$ for all $j \neq i$ and $y_i = f_i(x)$ for all i are the only vertices of M that are vertices of the convex hull of $\cup_k M_k$.

So the vertices of the convex hull of $\cup_k M_k$ are (x, y) such that x is a portfolio invested in only one asset or $(x, y) \in M_k$ and $y_i = 0$ for at least one coordinate i.

Another logical requirement imposed on portfolios is that the optimal portfolio maintain a certain level of diversification. We used $0-1$ variables $D_1, \ldots, D_N$ in our formulation of this disjunctive requirement. The program (L^*) that models the portfolio selection problem with minimum trade requirements and diversification requirements is as follows:

$$
\begin{array}{llrcl}
(L^*) & \max & \frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{N} ret_{ij}x_j & - & \frac{1}{S}\sum_{i=1}^{S} y_i \\
& \text{s.t.} & \sum_{j=1}^{N} p_j x_j & \leq & B \\
& & -\sum_{j=1}^{N} x_j & \leq & -w \\
& & x & \in & \cap_i \cup_j F_{ij} \\
& \forall j & 0 \leq D_j & \leq & 1 \\
& \forall j & x_j & \leq & M D_j \\
& \forall j & -x_j & \leq & -v D_j \\
& & -\sum_{j=1}^{N} D_j & \leq & -m
\end{array}
$$

where M is an upper bound for x and we have $x \in \cap_i \cup_j F_{ij}$ if for all j, $x_j = 0 \vee x_j \geq v$ and the sum of nonzero coordinates of x is at least m.

When this program is used in branch-and-bound search, the requirement at a node that $\forall j \;\; 0 \leq D_j \leq 1$ forces the sum of nonzero holdings to be at least m. At each node, there is a linear program (L_n^*) with the coefficient matrix R^h in place of E^h of (L).

$$
\begin{array}{rcrcrcll}
R^h & = & p_1x_1 & \dots & p_Nx_N & \leq & B & = r^h \\
 & & -x_1 & \dots & -x_N & \leq & -w & \\
 & & x_1 & \dots & 0 & \leq & MD_{k(1)} & \\
 & & 0 & x_j & 0 & \leq & MD_{k(j)} & \\
 & & 0 & \dots & x_N & \leq & MD_{k(N)} & \\
 & & -x_1 & \dots & 0 & \leq & -vD_{k(1)} & \\
 & & 0 & -x_j & 0 & \leq & -vD_{k(j)} & \\
 & & 0 & \dots & -x_N & \leq & -vD_{k(N)} &
\end{array}
$$

where for $j \in H \subset \{1, \dots, N\}$, $D_j = 0 \vee D_j = 1$ and $\sum_{j=1}^{N} D_j \geq m$. Again, this problem can be solved using either the combined breadth-first and decomposition algorithm or the combined depth-first and decomposition algorithm (see Sections 3.4 and 3.5). If we consider the polyhedron of vectors feasible for (R^h), we see that the vertices of G, which represent single-asset portfolios, have been removed. The vertices now include only portfolios with more than m nonzero holdings.

Let T_h equal the set of vectors feasible for (L_n^*). If we set the sum of nonzero holdings equal to m, then the vertices of T_h are m-asset portfolios. Included in the vertices of T_h are the portfolios (x, y) where $x = (v, \dots, v, \dots, \frac{B-2v}{p_j}, \dots, 0)$, or the portfolios which have minimum holdings in two assets and the balance of the budget invested in the jth asset, and where $y_i = f_i(x)$ for all i. These are the vertices (x, y) such that x is a vertex of $\{x \in \Re^N : R^h x \leq r^h\}$. Other vertices of T_h are points (x, y) such that $(x, y) \in T_h$ and $y_i = 0$ for at least one i.

If for any (L_n^*) the subroutine (SB) of Section (3.4.1) succeeds, then we know that an optimal portfolio for (L_n^*) is optimal for (P), which is the linear relaxation of (L^*). Therefore, if we are concerned with finding multiple optimal solutions to (L), carrying out this subroutine at each unpruned leaf can give us a set of disjunctive feasible optimal portfolios, if such portfolios exist, where the portfolios exhibit a range of levels of risk.

9.3 Cross decomposition for the disjunctive portfolio problem

Consider the portfolio selection problem (L) with minimum trade requirements of the previous section. Then the Lagrangian relaxation of (L) with respect to the constraints $A_2^1 x + A_2^1 y \leq b$ is the following problem, where $\vec{u}_2$ is the Lagrangian multiplier:

$$(LLD) \quad \min_{\vec{u}_2 \geq 0} \left\{ \begin{array}{lrcl} \max_{x \in \cap \cup F_{ij}} & cx + dy + \vec{u}_2(b - (A_2^1 x + A_2^2 y)) & & \\ \text{s.t.} & A_1^1 x & \leq & r \\ & A_3^2 y & \leq & 0 \end{array} \right\}$$

Let (L_h) be the linear program that describes the portfolio selection problem on a particular feasible region where all the portfolios satisfy the logical requirements; *i.e.*, $x \in K_h$, where

$$K_h = \cup_{i \leq N} F_{i,j(i)}$$

and $j(i) < S$ is the index for a particular set containing x_i that satisfies the logical requirement i. We have for (L_h):

$$
\begin{array}{lll}
(L_h) & \max & cx + dy \\
& \text{s.t.} & E^h x \leq e^h \\
& & A_2^1 x + A_2^2 y \leq b \\
& & A_3^2 y \leq 0
\end{array}
$$

The Lagrangian dual (LL_h) for (L_h) with respect to $A_2^1 x + A_2^2 y \leq b$ is

$$
(LL_h) \quad \min_{\vec{u}_2 \geq 0} \left\{ \begin{array}{lrl} \max_{(x,y)\geq 0} & cx + dy + \vec{u}_2(b - (A_2^1 x + A_2^2 y)) & \\ & E^h x & \leq e^h \\ & A_3^2 y & \leq 0 \end{array} \right\}
$$

If we look at the portfolio problem (L_h) in the context of cross decomposition, then we see that there is an $(N + 2 \times S)$ zero submatrix in A_1^2 and an $S \times N$ zero submatrix in A_3^1. So by Theorem 9 of Section 6, if we hope to avoid a duality gap between the Lagrangian dual (LL_h) and (L_h) we must either choose (1) x as our complicating variable and the rows $A_2^1 x + A_2^2 y \leq b$ as the complicating constraints or (2) y as the complicating variable and, again, $A_2^1 x + A_2^2 y \leq b$ as the complicating constraints.

For (1) the constraint subproblem is

$$
\begin{array}{lll}
(SD(\vec{u}_1)) & \max & cx + dy + \vec{u}_1(b_1 - A_2^1 x - A_2^2 y) \\
& \text{s.t.} & E^h x \leq e^h \\
& A_3^2 y \leq 0 &
\end{array}
$$

However, y is not constrained in this subproblem. Depending on the value of the sum of the coefficients of y in the objective function, there are three cases for the optimal y: if this sum is greater than 0, then $v(SD\vec{u}_1)$ is unbounded; if this sum is less than 0, then $y_i = 0$ will be optimal always and there may be a duality gap; and if the sum of the coefficients of y_i is equal to 0, then again there may be a duality gap since y_i is not constrained in $(SD\vec{\bar{u}}_1)$. Further, if the dual multiplier $\vec{u}_2$ is feasible for (DSP), then we know that $dy - \vec{u}_2 A_2^2 y \leq 0$. This follows from the fact that if $\vec{u}_2$ is the solution to $(DSP\bar{x})$ for some $\bar{x}$ feasible for (L_h) then we have $\vec{u}_2 A_2^2 \geq d$.

Consider the case when $y = 0$ in (LL_h). Then

$$cx + dy + \vec{u}_2(b - (A_2^1 x + A_2^2 y)) = cx + \vec{u}_2(b - A_2^1 x),$$

so for a fixed $\vec{\bar{u}}_2$ feasible for (DSP), the optimal portfolio will be an extreme point of the polyhedron defined by

$$\begin{aligned} E^h x &\leq e^h \\ A_3^2 y &\leq 0 \end{aligned} .$$

Let $p^* = (x_1^*, \ldots, x_N^*)$ be the optimal solution to (LL_h) when y is fixed at 0, and let z^* be the objective function value for this portfolio. Then

$$z^* = cx^* + \vec{\bar{u}}_2(b - A_2^1 x^*)$$

$$\geq cx + \vec{\bar{u}}_2(b - A_2^1 x) \geq cx + \vec{\bar{u}}_2(b - (A_2^1 x + A_2^2 y)) + dy$$

for any feasible (x, y) where $y > 0$. This last inequality holds for all $\vec{u}_2$ feasible for (DSP). Therefore, in this problem, $y = 0$ will always be optimal since the only constraints for y in (LL_h) are nonnegativity constraints. So if $y > 0$ is optimal for (L_h) there will be a duality gap between (L_h) and (LL_h). If $\vec{u}_2$ is not feasible for (DSP), *i.e.*, if $dy - \vec{u}_2 A_2^2 y > 0$, then (1) has an unbounded objective function value.

The last formulation to consider is (2). Here the variable and constraint subproblems are:

$$
\begin{array}{lll}
(DSP(\bar{y})) & \min & \vec{u}_1(e^h) + \vec{u}_2(b - A_2^2\bar{y}) - \vec{u}_3 A_3^2 y + d\bar{y} \\
 & \text{s.t.} & \vec{u}_1 E^h + \vec{u}_2 A_2^1 \geq c
\end{array}
$$

$$
\begin{array}{lll}
(SD(\bar{\vec{u}}_2)) & \max & cx + dy + \bar{\vec{u}}_2(b - (A_2^1 x + A_2^2 y \leq b)) \\
 & \text{s.t.} & E^h x \leq e^h \\
 & & A_3^2 y \leq 0
\end{array}
$$

Now, because $\{A_2^1 x + A_2^2 y \leq b\}$ are free for $x \in \Re^N$, there is a feasible y for every $x \in R^N$, and y is not constrained in $(SD(\bar{\vec{u}}_2))$. So, as above, either $\vec{u}_2$ is feasible for (DSP) and for the optimal y we will have that $dy - \vec{u}_2 A_2^2 y = 0$ or $(SD\vec{u}_2)$ is unbounded.

References

[Ba74] Balas, Egon, Disjunctive programming: properties of the convex hull of feasible points, MSRR No. 348, Graduate School of Industrial Administration, Carnegie-Mellon University (1974).

[Ba85] Balas, Egon, Disjunctive programming and a hierarchy of relaxations for discrete optimization problems, *SIAM J. Alg. Disc. Meth.* **6**:3, 466-486 (1985).

[BaRe] Martin Baxter and Andrew Rennie, *Financial Calculus: An introduction to derivative pricing.* Cambridge University Press, Cambridge (1996).

[Be] Benders, J. F., Partitioning procedures for solving mixed variables programming problems, *Numerische Mathematik* **4**, 238-252 (1962).

[Bi] Bienstock, Daniel, Computational study of a family of mixed-integer quadratic programming problems, *Mathematical Programming* **74**, 121-140 (1996).

[BlDeTo] Black, Fischer, Eugene Derman, and William Toy, A one-factor model of interest rates and its application to Treasury bond options, *Financial Analysts Journal*, 33-39, January-February 1990.

[CoInRo] Cox, J.C, J.E. Ingersoll, and S.A. Ross, A theory of the term structure of interest rates, *Econometrica* **53**:2, 385-407 (1985).

[Da] George B. Dantzig, *Linear Programming and Extensions.* The Rand Corporation, Santa Monica, California, 1963.

[DaIn] Dantzig, George B., and Infanger, Gerd, Multi-stage stochastic linear programs for portfolio optimization, *Annals of Operations Research* **45**, 59-76 (1993).

[DaWo] Dantzig, George B. and Philip Wolfe, Decomposition priniciple for linear programs, *Operations Research* **8**:1, 101-111 (1960).

[ElGr] Edwin J. Elton and Martin J. Gruber, *Modern Portfolio Theory and Investment Analysis.* 4th Ed. John Wiley & Sons, Inc., New York, 1991.

[FaFa95] *The Handbook of Fixed Income Securities.* 4th Ed. Frank J. Fabozzi and T. Dessa Fabozzi, eds. Irwin Professional Publishing, New York, 1995.

[Fl] R. Fletcher, *Practical Methods of Optimization.* John Wiley & Sons, Ltd., Chichester, 1987.

[Fr] Joel Franklin, *Methods of Mathematical Economics.* Springer-Verlag New York Inc., 1980.

[FlRK] Flippo, Olaf E., and Rinnooy Kan, Alexander H.G., Variable decomposition, constraint decomposition and cross decomposition in general mathematical programming, *NATO ASI Series* **F 82**, 1-18 (1992).

[GaJo] Michael R. Garey and David S. Johnson, *Computers and Intractability, A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, New York, 1979.

[GiMuWr] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Numerical Linear Algebra and Optimization Volume 1.* Addison-Wesley Publishing Company, Redwood City, California, 1991.

[HeJaMo] Heath, David, Robert Jarrow, and Andrew Morton, Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation, *Econometrica* **60**:1, 77-105 (1992).

[HeKa70] Held, M. and R. M. Karp, The Traveling Salesman problem and minimum spanning trees, *Operations Research* **18**, 1138-1162 (1970).

[HeKa71] The Traveling Salesman problem and minimal spanning trees: part II, *Mathematical Programming* **1**, 6-25 (1971).

[HiEc] Hiller, Randall S., and Jonathan Eckstein, Stochastic dedication: designing fixed income portfolios using massively parallel Benders decomposition, *Management Science* **39**:11, 1422-1437 (1993)

[HoLe] Ho, Thomas S. Y., and S.B. Lee, Term structure movements and pricing of interest rate claims, *Journal of Finance* **41**, 1011-1029 (1986).

[Ho] Hooker, J. N., and Yan, H., Verifying logic circuits by Benders decomposition, in *Principles and Practice of Constraint Programming.* Vijay Saraswat and Pascal Van Hentenryck, eds. MIT Press, 1995.

[HuWh] Hull, John, and A. White, One-factor interest rate models and the valuation of interest rate derivative securities, *Journal of Financial and Quantitative Analysis* (1993).

[Ja] Robert A. Jarrow, *Modelling Fixed Income Securities and Interest Rate Options*, The McGraw-Hill Companies, Inc., New York, 1996.

[JoSu] Johnson, Ellis L., and Suhl, Uwe H., Experiments in integer programming, *Discrete Applied Mathematics* **2**, 39-55 (1980).

[Ko] Konno, Hiroshi, Piecewise linear risk function and portfolio optimization, *Journal of the Operations Research Society of Japan* **33**:2, 139-156 (1990).

[KoIn] Konno, Hiroshi, and Michimori Inori, Bond portfolio optimization by bilinear fractional programming, *Journal of the Operations Research Society of Japan* **32**:2, 143-158 (1989).

[KoYa] Konno, Hiroshi, and Hiroaki Yamazaki, Mean-absolute deviation portfolio optimization and its applications to Tokyo stock market, *Management Science* **37**:5, 519-531 (1991).

[Ma] Harry M. Markowitz, *Portfolio Selection: Efficient Diversification of Investments.* Cowles Foundation for Research in Economics at Yale University, New Haven, 1959.

[McTr] Ken McAloon and Carol Tretkoff, *Optimization and Computational Logic*, Wiley, New York, 1996.

[NeWo] George L. Nemhauser and Laurence A. Wolsey, *Integer and Combinatorial Optimization*, Wiley, New York, 1988.

[Sc] Alexander Schrijver, *Theory of Linear and Integer Programming*, Wiley, Chichester, 1986.

[Sh67] Sharpe, William F., A linear programming algorithm for mutual fund portfolio selection, *Management Science* **13**:7, 499-510 (1967).

[Sh71] Sharpe, William F., Mean-absolute-deviation characteristic lines for securities and portfolios, *Management Science* **18**: 2, B1-B13 (1971).

[StWi] Josef Stoer and Christoph Witzgall, *Convexity and Optimization in Finite Dimensions I.* Springer-Verlag, Berlin, 1970.

[TaTa] Tabata, Yoshio, and Takeda, Eiji, Bicriteria optimization problem of designing an index fund, *Journal of the Operational Research Society* **46**, 1023-1032 (1995).

[VR] Van Roy, Tony J., Cross decomposition for mixed integer programming, *Mathematical Programming* **25**, 46-63 (1983).

[Va] Vasicek, O., An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**, 177-188 (1977).

[VSWe] Van Slyke, R. M., and Roger Wets, L-Shaped linear programs with applications to optimal control and stochastic programming, *SIAM J. Appl. Math.* **17**:4, 638-663 (1969).

[WoVZZe] Worzel, Kenneth J., Vassidou-Zeniou, Christiana, and Zenios, Stavros A., Integrated simulation and optimization models for tracking indices of fixed-income securities, *Operations Research* **42**:2, 223-233 (1994).

[Ze93] *Financial Optimization*, Stavros A. Zenios, ed. Cambridge University Press, Cambridge, 1993.

[Ze95] Zenios, Stavros A., Asset/liability management under uncertainty for fixed-income securities, *Annals of Operations Research* **59**, 77-97 (1995).

[ZeKa] Zenios, Stavros A., and Kang, Pan, Mean-absolute deviation portfolio optimization for mortgage-backed securities, *Annals of Operations Research* **45**, 433-450 (1993).