Bilevel Integer Linear Programming

TED RALPHS
SCOTT DENEGRE
ISE Department
COR@L Lab
Lehigh University
ted@lehigh.edu

COR@L
COMPUTATIONAL OPTIMIZATION RESEARCH AT LEHIGH

MOPTA 2009, Lehigh University, 19 August 2009

Thanks: Work supported in part by the National Science Foundation

Surgeon General's Warning
This talk contains half-baked ideas. Consumption may result in periods of confusion and/or insomnia.

Outline

Motivation

- A standard mathematical program models a decision to be made by a *single* decision-maker with a *single* objective at a *single* point in time.
- Many real-world decision problems involve
  - multiple, independent decision-makers (DMs),
  - multiple, possibly conflicting objectives, and/or
  - hierarchical/multi-stage decisions.
- Modeling frameworks
  - Multiobjective Programming $\Leftarrow$ multiple objectives, single DM
  - Mathematical Programming with Recourse $\Leftarrow$ multiple stages, single DM
  - Multilevel Programming $\Leftarrow$ multiple stages, multiple DMs
- *Multilevel programming* generalizes standard mathematical programming by modeling hierarchical decision problems, such as Stackelberg games.
- In this talk, we focus on bilevel programming models, particularly with discrete variables.
- Such models arises in a remarkably wide array of applications.

Applications

- Hierarchical decision systems
  - Government agencies
  - Large corporations with multiple subsidiaries
  - Markets with a single "market-maker."
  - Decision problems with recourse
- Parties in direct conflict
  - Zero sum games
  - Interdiction problems
- Modeling "robustness": leader represents external phenomena that cannot be controlled.
  - Weather
  - External market conditions
- Controlling optimized systems: follower represents a system that is optimized by its nature.
  - Electrical networks
  - Biological systems

Basic Framework

- For the remainder of the talk, we consider systems in which there are two DMs, a *leader* or *upper-level* DM and a *follower* or *lower-level* DM.
- We assume *individual rationality* of the two DMs.
- This means roughly that the leader has the ability to predict the reaction of the follower to a given course of action.
- For simplicity, we also assume that for every action by the leader, the follower has a feasible reaction.
- The follower may in fact have more than one equally favorable reaction to a given action by the leader.
- These alternatives may not be equally favorable to the leader.
- We assume that the leader may choose among the follower's alternatives.
- This assumption is reasonable if the players have a "semi-cooperative" relationship.

Bilevel Linear Programming

Formally, a *bilevel linear program* is described as follows.

- $x \in X \subseteq \mathbb{R}^{n_1}$ are the *upper-level variables*
- $y \in Y \subseteq \mathbb{R}^{n_2}$ are the *lower-level variables*

Bilevel Linear Program

$$\max \left\{ c^1 x + d^1 y \mid x \in \mathcal{P}_U \cap X, y \in \operatorname{argmin}\{d^2 y \mid y \in \mathcal{P}_L(x) \cap Y\} \right\}$$

The *upper-* and *lower-level feasible regions* are:

$$\mathcal{P}_U = \left\{ x \in \mathbb{R}_+ \mid A^1 x \leq b^1 \right\} \text{ and}$$
$$\mathcal{P}_L(x) = \left\{ y \in \mathbb{R}_+ \mid G^2 y \geq b^2 - A^2 x \right\}.$$

Notation

We utilize the following notation:

Notation

$$
\begin{aligned}
\Omega &= \left\{(x,y) \in \mathbb{R}_+^{n_1} \times \mathbb{R}_+^{n_2} \mid x \in \mathcal{P}_U, y \in \mathcal{P}_L(x)\right\} \\
\Omega^I &= \Omega \cap X \times Y \\
M(x) &= \operatorname{argmin}\{d^2 y \mid y \in \mathcal{P}_L(x)\} \\
M^I(x) &= M(x) \cap Y \\
\mathcal{F} &= \{(x,y) \mid x \in \mathcal{P}_U, y \in M(x)\} \\
\mathcal{F}^I &= \left\{(x,y) \mid x \in \mathcal{P}_U^I, y \in M^I(x)\right\}
\end{aligned}
$$

Special Cases

- When $X = \mathbb{R}^{n_1}$ and $Y = \mathbb{R}^{n_2}$, we have a *continuous* BLP (usually just a BLP).
- When $X = \mathbb{Z}^{p_1} \cap \mathbb{R}^{n_1 - p_1}$ and/or $Y = \mathbb{Z}^{p_2} \cap \mathbb{R}^{n_2 - p_2}$, then we have a *mixed integer* BLP.
- When does a solution exist?

Existence of Solutions (Dempe, 2001)

- In the continuous case, if Ω is nonempty and bounded, then there is a solution.
- This suffices also in the case that $X = \mathbb{Z}^{n_1}$.
- If $X \supset \mathbb{Z}^{n_1}$ and $Y \supset \mathbb{R}^{n_1}$, the problem may not have a solution in general because the feasible set may not be closed.

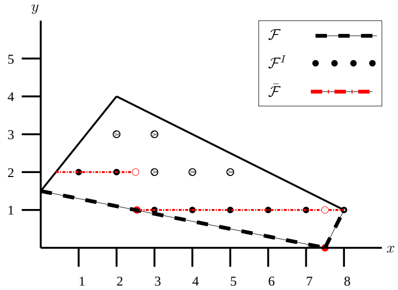
Further Generalizations

- The follower's variables may appear in the leader's constraints (see, e.g., Audet et al. (1997)).
- The follower's objective may also parameterized (see Dempe (2001)).

Example

The following instance of (MIBLP) is from Moore and Bard (1990).

$$
\begin{aligned}
\max_{x \in X} \quad & x + 10y \\
\text{subject to} \quad & y \in \operatorname{argmin} \{ y : -25x + 20y \leq 30 \\
& \qquad x + 2y \leq 10 \\
& \qquad 2x - y \leq 15 \\
& \qquad 2x + 10y \geq 15 \\
& \qquad y \in Y \; \}
\end{aligned}
$$



1. For $X = \mathbb{R}_+$ and $Y = \mathbb{R}_+$, the feasible set is $\mathcal{F}$ and the solution is $(8, 1)$ with objective value 18.
2. For $X = \mathbb{Z}_+$ and $Y = \mathbb{Z}_+$, the feasible set is $\mathcal{F}^I$ and the solution is $(2, 2)$ with objective value 22.
3. For $X = \mathbb{R}_+$ and $Y = \mathbb{Z}_+$, the feasible set is $\bar{\mathcal{F}}$ and there is no solution. The infimum of the objective values is 22.5.

An Important Special Case

- If $d^1 = -d^2$, we can view this as a *mathematical program with recourse*.
- Two-stage stochastic programs with recourse (discrete distribution) are a special case.
- The resulting problem can be solved as a standard mathematical program.
- For the case when $Y = \mathbb{R}^n$, we can solve the problem by Benders Decomposition.
- Note that the value function of the lower-level problem is convex in the upper-level variables, so we can also reformulate as a convex program
- This is a useful way of visualizing the situation.
- This is really the only tractable case.

Technical Assumptions

We make the following assumptions in order to ensure the problem is well-posed and has a solution.

Assumptions

1. For every action by the leader, the follower has a rational reaction ($\mathcal{P}_L(x) \cap Y \neq \emptyset$ for all $x \in \mathcal{P}_U \cap X$).
2. The follower is semi-cooperative (the leader may choose among alternative members of $M^I(x)$).
3. The feasible set $\mathcal{F}^I$ is nonempty and compact.

The BLP can now be simply stated as:

Bilevel Linear Program

$$\max_{(x,y)\in\mathcal{F}^I} c^1 x + d^1 y. \tag{MIBLP}$$

Solving Continuous BLPs

- In the continuous case, the lower-level problem can be replaced with its optimality conditions.
- The optimality conditions for the lower-level optimization problem are

$$
\begin{aligned}
G^2 y &\geq b^2 - A^2 x \\
uG^2 &\leq d^2 \\
u(b^2 - G^2 - A^2 x) &= 0 \\
(d^2 - uG^2)y &= 0 \\
u, y &\in \mathbb{R}_+
\end{aligned}
$$

- Note that this is a special case of a class of non-linear mathematical programs known as *mathematical programs with equilibrium constraints* (MPECs).
- This can be solved in a number of ways, including converting it to standard integer program.
- Note that in this case, the value function of the lower-level problem is piecewise linear, but not necessarily convex.

Discrete BLPs

- When some of the variables are discrete, the situation is a bit more difficult.
- Because the duals that exist for general integer programs are not tractable in general, we cannot use the same approach as we did for the continuous case.
- In fact, going from the continuous case to the discrete case in the bilevel setting poses significantly different challenges than for standard MILPs.

Duality for Mixed Integer Linear Programs

Let $\Gamma^m = \{F : \mathbb{R}^m \Rightarrow \mathbb{R} \mid F \text{ is subadditive and nonincreasing}, F(0) = 0\}$. Then the subadditive dual is

Subadditive Dual Problem

$$
\begin{aligned}
\max \quad & F(d) \\
& F(a_j) \leq c_j \quad j \in [1, p_2] \\
& \bar{F}(a_j) \leq c_j \quad j \in [p_2 + 1, n_2] \\
& F \in \Gamma^{m_2}
\end{aligned}
$$

where a_j is the j^{th} column of A and

$$\bar{F}(d) = \limsup_{\delta \to 0^+} \frac{F(\delta d)}{\delta}.$$

Reformulation with Optimality Conditions

In principle, we can use subadditive duality to obtain optimality conditions for the lower-level problem (reformulation shown here is for the pure integer case).

$$
\begin{aligned}
\max_{x,y,F} \quad & c^1x + d^1y \\
\text{subject to} \quad & A^1x \le b^1 \\
& A^2x + G^2y \ge b^2 \\
& F(g_j^2) \le d_j^2, \quad \forall j = 1, \ldots, n_2 \\
& (F(g_j^2) - d_j^2)y_j = 0, \quad \forall j = 1, \ldots, n_2 \\
& \sum_{j=1}^{n_2} F(g_j^2)y_j = F(b^2 - A^2x) \\
& x \in \mathbb{Z}_+^{n_1}, y \in \mathbb{Z}_+^{n_2}, F \in \Gamma^{m_2}.
\end{aligned}
$$

This is analogous to the reformulation in the continuous case, but is intractable in general.

Towards a Branch and Bound Algorithm

- The seemingly obvious thing to do then is to develop a branch-and-bound algorithm.
- Our approach is to try as much as possible to directly generalize concepts from mixed integer linear programming.

Components

- Bounding methods ($\Leftarrow$ this talk)
- Branching methods ($\Leftarrow$ this talk)
- Search strategies
- Preprocessing methods
- Primal heuristics

- In the remainder of the talk, we address development of these components.

Contrast with Mixed Integer Linear Programming

In algorithms for solving standard mixed integer linear programs (MILPs), we frequently use the following properties.

Properties

1. If the continuous relaxation has no feasible solution, then neither does the original problem.
2. If the continuous relaxation has a solution, then its objective value is a valid upper bound on that of the original problem.
3. If the solution to the continuous relaxation is integral, then it is optimal for the original problem.

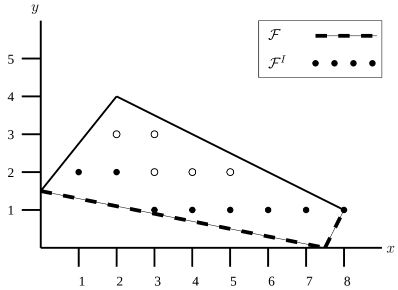
Properties 2 and 3 result from the fact that the set of feasible solutions for the original MILP is contained in the feasible set of the relaxation.

THIS IS NOT THE CASE FOR MIBLP

Example

Consider the following instance of (MIBLP) again:

$$
\begin{aligned}
\max_{x \in \mathbb{Z}_+} \quad & x + 10y \\
\text{subject to} \quad & y \in \operatorname{argmin} \{ y : \\
& \qquad -25x + 20y \leq 30 \\
& \qquad x + 2y \leq 10 \\
& \qquad 2x - y \leq 15 \\
& \qquad 2x + 10y \geq 15 \\
& \qquad y \in \mathbb{Z}_+ \ \}
\end{aligned}
$$



From the figure, we can see that

1. $\mathcal{F} \subseteq \Omega$, $\mathcal{F}^I \subseteq \Omega^I$, and $\Omega^I \subseteq \Omega$
2. $\mathcal{F}^I \nsubseteq \mathcal{F}$

Properties of MIBLPs

In this example:

- Optimizing over $\mathcal{F}$ yields the *integer* solution $(8, 1)$, with the upper-level objective value 18.
- Imposing integrality yields the solution $(2, 2)$, with upper-level objective value 22

From this we can make two important observations:

1. The objective value obtained by relaxing integrality is not a valid bound on the solution value of the original problem since we may have

$$\max_{(x,y)\in\mathcal{F}} c^1x + d^1y < \max_{(x,y)\in\mathcal{F}^I} c^1x + d^1y.$$

2. Even when solutions to $\max_{(x,y)\in\mathcal{F}} c^1x + d^1y$ are in $\mathcal{F}^I$, they are not necessarily optimal.

Thus, only *Property 1* remains valid.

Bounding Methods

Relaxing integrality conditions *and* the requirement $y \in M^I(x)$ yields the relaxation

$$\max_{(x,y)\in\Omega} c^1x + d^1y. \tag{LR}$$

- The resulting bound can be used in combination with a standard variable branching scheme to yield an algorithm that solves (MIBLP).
- Unfortunately, the bound is too weak to be effective on interesting problems.

Idea!

Strengthen the linear relaxation with inequalities valid for $\mathcal{F}^I$ to improve the bound.

Valid Inequalities for MIBLP

Definition

An inequality defined by (π_1, π_2, π_0) is *valid* for $\mathcal{F}^I$ if $\pi_1 x + \pi_2 y \leq \pi_0$ for all $(x, y) \in \mathcal{F}^I$.

- Unless $\text{conv}(\mathcal{F}^I) = \Omega$, $\exists$ inequalities that are valid for $\mathcal{F}^I$, but are violated by some members of Ω.
- To generate these inequalities, we must exploit information *not* contained in the linear description of Ω.
- For a point (x, y) to be feasible for an MIBLP, it must satisfy three conditions:

Bilevel Feasibility Conditions

1. $(x, y) \in \Omega$,
2. $(x, y) \in X \times Y$, and
3. $y \in M^I(x)$.

Cutting Plane Approach

Let $(\hat{x}, \hat{y})$ be a solution to

$$\max_{(x,y)\in\Omega} c^1 x + d^1 y. \tag{LR}$$

- If $(\hat{x}, \hat{y}) \notin X \times Y$, then Condition 2 is violated $\Rightarrow$ apply MILP cutting plane techniques to separate $(\hat{x}, \hat{y})$ from Ω^I.
- If $(\hat{x}, \hat{y}) \in X \times Y \;\Rightarrow$ check whether Condition 3 is satisfied.
- Fix $x = \hat{x}$ and solve the lower-level problem

$$\min_{y \in \mathcal{P}_L^I(\hat{x})} d^2 y \tag{1}$$

with the fixed upper-level solution $\hat{x}$.

Bilevel Feasibility Check

Let y^* be the solution to (1).

- $(\hat{x}, y^*)$ is bilevel feasible $\Rightarrow c^1\hat{x} + d^1 y^*$ is a valid upper bound on the optimal value of the original MIBLP
- Either
  1. $d^2\hat{y} = d^2 y^* \Rightarrow (\hat{x}, \hat{y})$ is bilevel feasible.
  2. $d^2\hat{y} > d^2 y^* \Rightarrow$ generate a valid inequality violated by $(\hat{x}, \hat{y})$.

Bilevel Feasibility Cut

Let

$$A := \begin{bmatrix} A^1 \\ A^2 \end{bmatrix}, \quad G := \begin{bmatrix} 0 \\ G^2 \end{bmatrix}, \quad \text{and} \quad b := \begin{bmatrix} b^1 \\ b^2 \end{bmatrix}.$$

A basic feasible solution $(\hat{x}, \hat{y}) \in \Omega^I$ to (LR) is the *unique* solution to

$$a_i' x + g_i' y = b_i, \quad i \in I$$

where I is the set of active constraints at $(\hat{x}, \hat{y})$.

This implies that

$$\left\{ (x, y) \in \Omega^I \mid \sum_{i \in I} a_i' x + g_i' y = \sum_{i \in I} b_i \right\} = \left\{ (\hat{x}, \hat{y}) \right\}$$

and $\sum_{i \in I} a_i' x + g_i' y \leq \sum_{i \in I} b_i$ is valid for Ω.

Bilevel Feasibility Cut (cont.)

The face of Ω^I induced by $\sum_{i \in I} a_i' x + g_i' y \leq \sum_{i \in I} b_i$ does not contain any other members of Ω^I

$\Rightarrow$ If $X = \mathbb{Z}^{n_1}$ and $Y = \mathbb{Z}^{n_2}$ (i.e., the *pure integer* case), we can "push" the hyperplane until it meets the next integer point without separating any additional members of Ω^I.

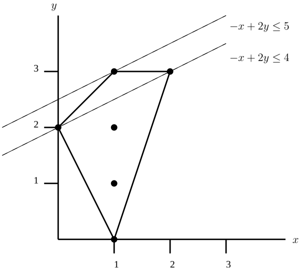
A Valid Inequality

$\sum_{i \in I} a_i' x + g_i' y \leq \sum_{i \in I} b_i - 1$ is valid for $\Omega^I \setminus \{(x, y)\}$.

- Observation 2 $\Rightarrow$ inequality is valid for $\mathcal{F}^I$.
- Similar in spirit to Gomory's procedure for standard ILPs.

A Simple Example

$$\max_x \min_y \{y \mid -x + y \leq 2, -2x - y \leq -2, 3x - y \leq 3, y \leq 3, x, y \in \mathbb{Z}_+\}.$$



- The *bilevel infeasible* point $(1, 3)$ is an optimal solution to the LP

$$\max_x \{y \mid -x + y \leq 2, -2x - y \leq -2, 3x - y \leq 3, y \leq 3, x, y \in \mathbb{R}_+\}.$$

- The inequality $-x + 2y \leq 4$ separates $(1, 3)$ from $\mathcal{F}^I$.

Cutting Plane Algorithm (Pure Integer Case)

- This approach can be implemented simply within a standard MILP solver framework simply by adding the aforementioned cut generation method.
- These inequalities are only valid for the pure integer case, however, and we would like to consider the general case.
- The cuts generated by this method are also not very deep. In order to solve problems of interesting size, we would like to generate deeper cuts.
- In order to derive stronger disjunctions that can be used for branching and/or cutting , we must look more closely at violations of Condition 3.

Idea!

When the current relaxed solution is bilevel infeasible, derive disjunctions from the local structure of the value function.

The Value Function

The value function of an MILP is a function $z : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ that returns the optimal value of the program as a function of the right-hand side vector.

MILP Value Function

$$z(d) = \min_{x \in S(d)} cx, \tag{2}$$

where, for a given right-hand side vector $d \in \mathbb{R}^m$,

$$S(d) = \{x \in \mathbb{Z}^p_+ \times \mathbb{R}^{n-p}_+ \mid Ax \leq d\}.$$

- If we knew the value function, we could replace the lower-level problem with a single non-linear constraint.
- Note that the value function is an optimal solution for the subadditive dual for any right-hand side.
- Constructing the value function is difficult, but we can take advantage of local structure.

Value Function Characterization

Blair and Jeroslow (1977) and Blair (1995) show that z is

- piecewise polyhedral, and
- can be expressed as the sum of the value function of a related pure integer program and a linear interpolation term obtained from the coefficients of the continuous variables.

In Guzelsoy and Ralphs (2008), the case of a MILP with a single constraint is considered. Under this special case:

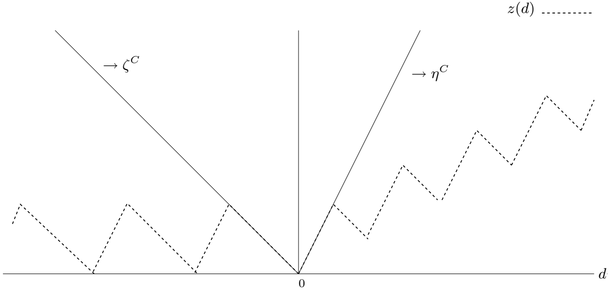
- z is composed of a finite number of linear segments on any closed interval
- the slope of each of these linear segments is given by one of two possible values.

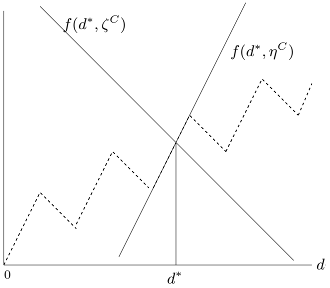
For illustration purposes, we now examine the case of a single equality constraint at the lower level.

Value Function Structure (Single Constraint)

Let $C^+ = \{i \in C \mid a_i > 0\}$ and $C^- = \{i \in C \mid a_i < 0\}$, and

$$\eta^C = \min\left\{\frac{c_i}{a_i} \mid i \in C^+\right\} \text{ and } \zeta^C = \max\left\{\frac{c_i}{a_i} \mid i \in C^-\right\}.$$





For any $d \leq d^*$,

$$z(d) \leq \max\{f(d^*, \zeta^C), f(d^*, \eta^C)\} = f(d^*, \zeta^C).$$

Similarly, for any $d \geq d^*$,

$$z(d) \leq \max\{f(d^*, \zeta^C), f(d^*, \eta^C)\} = f(d^*, \eta^C).$$

In Our Context. . .

Let $(\hat{x}, \hat{y}) \in X \times Y$ be a solution to (LR), and

$$z(b^2 - A^2\hat{x}) = \max\{d^2 y \mid G^2 y = b^2 - A^2\hat{x}, y \in Y\}.$$

Suppose, $(\hat{x}, \hat{y})$ is *not* bilevel feasible (i.e., $d^2\hat{y} > z(b^2 - A^2\hat{x})$). Then:

1. For any x such that $b^2 - A^2 x \leq b^2 - A^2\hat{x}$,

$$d^2 y \leq f(b^2 - A^2\hat{x}, \zeta^C).$$

2. For any x such that $b^2 - A^2 x \geq b^2 - A^2\hat{x}$,

$$d^2 y \leq f(b^2 - A^2\hat{x}, \eta^C).$$

Bilevel Feasibility Branching

Thus, we have the following disjunction.

Bilevel Feasibility Disjunction

$$b^2 - A^2 x \leq b^2 - A^2 \hat{x} \quad \text{AND} \quad d^2 y \leq f(b^2 - A^2 \hat{x}, \zeta^C)$$

$$\text{OR}$$

$$b^2 - A^2 x \geq b^2 - A^2 \hat{x} \quad \text{AND} \quad d^2 y \leq f(b^2 - A^2 \hat{x}, \eta^C).$$

This can immediately be used to develop a stronger branching scheme when solutions $(\hat{x}, \hat{y}) \in \Omega^I$ such that $\hat{y} \notin M^I(\hat{x})$ are found.

A Disjunctive Cut Approach

Consider the two polyhedra that result if we impose this disjunction on the original set of constraints in Ω. This yields the polyhedra:

$$P^1 = \left\{ \begin{array}{lcl} A^1x & = & b^1 \\ A^2x + G^2y & = & b^2 \\ A^2x & \geq & A^2\hat{x} \\ -\zeta^C A^2 x - d^2 y & \geq & -\zeta^C A^2 \hat{x} - d^2 y^* \\ x, y & \geq & 0 \end{array} \right\}$$

and

$$P^2 = \left\{ \begin{array}{lcl} A^1x & = & b^1 \\ A^2x + G^2y & = & b^2 \\ -A^2x & \geq & -A^2\hat{x} \\ -\eta^C A^2 x - d^2 y & \geq & -\eta^C A^2 \hat{x} - d^2 y^* \\ x, y & \geq & 0. \end{array} \right\}$$

Constructing the Disjunctive Cut

Let (u^i, v^i, w^i, z^i) be multipliers for the constraints in polyhedron P^i. The following inequalities are valid for P^1:

$$u^1A^1x + v^1A^2x + w^1A^2x - z^1\zeta^C A^2x + v^1G^2y - z^1d^2y \geq$$
$$u^1b^1 + v^1b^2 + w^1A^2\hat{x} - z^1(\zeta^C A^2\hat{x} + d^2y^*)$$

and P_2:

$$u^2A^1x + v^2A^2x - w^2A^2x - z^2\eta^C A^2x + v^2G^2y - z^2d^2y \geq$$
$$u^2b^1 + v^2b^2 - w^2A^2\hat{x} - z^2(\eta^C A^2\hat{x} + d^2y^*).$$

It is well-known that, given these inequalities, we can construct an inequality $\alpha x + \beta y \geq \gamma$ that is valid for $\text{conv}(P^1 \cup P^2)$ by selecting α, β, and γ such that

$$\alpha \geq \max\{\pi_1^1, \pi_1^2\}, \quad \beta \geq \max\{\pi_2^1, \pi_2^2\}, \quad \text{and} \quad \gamma \leq \min\{\pi_0^1, \pi_0^2\}.$$

Linear Description of the Set of Valid Inequalities

Thus, the inequality $\alpha x + \beta y \geq \gamma$ is valid for $\text{conv}(P^1 \cup P^2)$ if

$$\begin{aligned}
\alpha - (u^{1^+} - u^{1^-})A^1 - (v^{1^+} - v^{1^-})A^2 - w^1 A^2 + z^1 \zeta^C A^2 &\geq 0 \\
\alpha - (u^{2^+} - u^{2^-})A^1 - (v^{2^+} - v^{2^-})A^2 + w^2 A^2 + z^2 \eta^C A^2 &\geq 0 \\
\beta - (v^{1^+} - v^{1^-})G^2 + z^1 d^2 &\geq 0 \\
\beta - (v^{2^+} - v^{2^-})G^2 + z^2 d^2 &\geq 0 \\
\gamma - (u^{1^+} - u^{1^-})b^1 - (v^{1^+} - v^{1^-})b^2 - w^1 A^2 \hat{x} + z^1(\zeta^C A^2 \hat{x} - d^2 y^*) &\leq 0 \\
\gamma - (u^{2^+} - u^{2^-})b^1 - (v^{2^+} - v^{2^-})b^2 + w^2 A^2 \hat{x} + z^2(\eta^C A^2 \hat{x} - d^2 y^*) &\leq 0 \\
u^{1^+}, u^{1^-}, u^{2^+}, u^{2^-}, v^{1^+}, v^{2^-}, w^1, w^2, z^1, z^2 &\geq 0.
\end{aligned}$$

Cut Generating LP

To find the deepest cut, we can solve the *cut generation LP*:

$$
\begin{aligned}
\min \quad & \alpha\hat{x} + \beta\hat{y} - \gamma \\
\text{s.t.} \quad & \alpha - (u^{1^+} - u^{1^-})A^1 - (v^{1^+} - v^{1^-})A^2 - w^1A^2 + z^1\zeta^C A^2 \geq 0 \\
& \alpha - (u^{2^+} - u^{2^-})A^1 - (v^{2^+} - v^{2^-})A^2 + w^2A^2 + z^2\eta^C A^2 \geq 0 \\
& \beta - (v^{1^+} - v^{1^-})G^2 + z^1d^2 \geq 0 \\
& \beta - (v^{2^+} - v^{2^-})G^2 + z^2d^2 \geq 0 \\
& \gamma - (u^{1^+} - u^{1^-})b^1 - (v^{1^+} - v^{1^-})b^2 - w^1A^2\hat{x} + z^1(\zeta^C A^2\hat{x} + d^2y^*) \leq 0 \\
& \gamma - (u^{2^+} - u^{2^-})b^1 - (v^{2^+} - v^{2^-})b^2 + w^2A^2\hat{x} + z^2(\eta^C A^2\hat{x} + d^2y^*) \leq 0 \\
& u^{1^+} + u^{1^-} + v^{1^+} + v^{1^-} + w^1 + z^1 + u^{2^+} + u^{2^-} + v^{2^+} + v^{2^-} + w^2 + z^2 = 1 \\
& u^{1^+}, u^{1^-}, u^{2^+}, u^{2^-}, v^{1^+}, v^{1^-}, v^{2^+}, v^{2^-}, w^1, w^2, z^1, z^2 \geq 0,
\end{aligned}
$$

similar to that used in constructing lift-and-project cuts.

Numerical Example

MIBLP Example

$$
\begin{aligned}
\min \quad & 8x_1 + x_2 + 2x_3 + 3x_5 + 4x_6 \\
& + y_1 + 3y_2 + 2y_3 + 4y_4 + 2y_6 \\
\text{subject to} \quad & 4x_1 + 2x_2 - 4x_3 + 3x_4 + 6x_5 + x_6 = 24 \\
& x_1, x_2, x_3 \in \mathbb{Z}_+, x_4, x_5, x_6 \in \mathbb{R}_+ \\
& y \in \operatorname{argmin} \{ \, 2y_1 + 2y_3 + 8y_4 + 4y_5 + 3y_6 : \\
& \qquad\qquad 2x_1 - x_2 + 4x_3 - 4x_4 + 3x_5 + x_6 \\
& \qquad\qquad\quad + 4y_1 - 6y_2 + 6y_3 + 4y_4 - 4y_5 + \frac{4}{3}y_6 = 16, \\
& \qquad\qquad y_1, y_2, y_3 \in \mathbb{Z}_+, y_3, y_5, y_6 \in \mathbb{R}_+ \, \},
\end{aligned}
$$

Numerical Example (cont.)

- The initial LP lower bound is 3, with a solution of

$$x_4 = 8, \quad y_1 = 12.$$

- The cutting plane procedure yields the cut

$$\frac{3}{4}x_1 + \frac{7}{24}x_2 + \frac{1}{3}x_4 + \frac{9}{8}x_5 + \frac{5}{24}x_6 - \frac{1}{3}y_1 - \frac{1}{6}y_2 - \frac{1}{8}y_3 - \frac{1}{12}y_4 - \frac{1}{12}y_g \geq \frac{10}{3}.$$

- This yields a new lower bound of 3.224 and a solution of

$$x_4 = 0.8163, \quad x_5 = 3.5918, \quad y_1 = 2.1225.$$

- The optimal value is 3.25 and an optimal solution is

$$x_5 = 4, \quad y_1 = 1.$$

The General Case

- This method can be generalized in principle.
- The Jeroslow formula tells us what the local structure looks like.
- In fact, we can derive local structure from the branch-and-bound tree constrcuted when we do the bilevel feasibility check.
- There is an obvious combinatorial explosion.

Jeroslow Formula for General MILP

Let the set $\mathbb{E}$ consist of the index sets of dual feasible bases of the linear program

$$\min\{\frac{1}{M}c_C x_C \ : \ \frac{1}{M}A_C x_C = b, x \geq 0\}$$

where $M \in \mathbb{Z}_+$ such that for any $E \in \mathbb{E}$, $MA_E^{-1}a^j \in \mathbb{Z}^m$ for all $j \in I$.

Theorem 1 (Jeroslow Formula) *There is a* $g \in \mathscr{G}^m$ *such that*

$$z(d) = \min_{E \in \mathbb{E}} g(\lfloor d \rfloor_E) + v_E(d - \lfloor d \rfloor_E) \ \ \forall d \in \mathbb{R}^m \ \ \textit{with} \ \ \mathcal{S}(d) \neq \emptyset,$$

where for $E \in \mathbb{E}$, $\lfloor d \rfloor_E = A_E \lfloor A_E^{-1} d \rfloor$ *and* v_E *is the corresponding basic feasible solution.*

Implementation

The Mixed Integer Bilevel Solver (MibS) implements the branch and bound framework described here using software available from the Computational Infrastructure for Operations Research (COIN-OR) repository.

COIN-OR Components Used

- The COIN High Performance Parallel Search (CHiPPS) framework to perform the branch and bound.
- The COIN Branch and Cut (CBC) framework for solving the MILPs.
- The COIN LP Solver (CLP) framework for solving the LPs arising in the branch and cut.
- The Cut Generation Library (CGL) for generating cutting planes within CBC.
- The Open Solver Interface (OSI) for interfacing with CBC and CLP.

Conclusions and Future Work

- Preliminary testing to date has revealed that these problems can be extremely difficult to solve in practice.
- What we have implemented so far has only scratched the surface.
- Currently, we are focusing on special cases where we can get traction.
  - Interdiction problems
  - Stochastic integer programs
- Much work remains to be done.
  - Preprocessing procedures
  - Heuristics
  - More sophisticated branching rules

References I

Audet, C., P. Hansen, B. Jaumard, and G. Savard 1997. Links between linear bilevel and mixed 0-1 programming problems. *Journal of Optimization Theory and Applications* **93**(2), 273–300.

Blair, C. 1995. A closed-form representation of mixed-integer program value functions. *Mathematical Programming* **71**, 127–136.

Blair, C. and R. Jeroslow 1977. The value function of a mixed integer program: I. *Discrete Mathematics* **19**, 121–138.

Dempe, S. 2001. Discrete bilevel optimization problems. Technical Report D-04109, Institut fur Wirtschaftsinformatik, Universitat Leipzig, Leipzig, Germany.

Guzelsoy, M. and T. Ralphs 2008. The value function of a mixed-integer linear program with a single constraint. Technical report, Lehigh University.

Moore, J. and J. Bard 1990. The mixed integer linear bilevel programming problem. *Operations Research* **38**(5), 911–921.