Bilevel Integer Linear Programming

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

Université Bordeaux, 16 December 2008

Thanks: Work supported in part by the National Science Foundation

Outline

Motivation

- The modeling framework of standard mathematical programming assumes a decision problem with a single decision-maker and a single objective.
- Many real-world decision problems involve multiple, independent decision-makers (DMs) and multiple, possibly conflicting objectives.
- Related modeling frameworks
  - Multiobjective programming
  - Nash games
  - Stackelberg games

Nash and Stackelberg Games

- Many game theoretic models can be formulated as optimization problems involving multiple decision makers.
- In a *Nash game*, the players are treated as equals and take simultaneous action.
- One is concerned with finding a *Nash equilibrium*, in which the action of each player is optimal, given the actions of all other players.
- In a *Stackelberg game*, there is a dominant player, called the *leader*, who acts first and other players react.
- In this case, one is concerned with determining the leader's decision, given the assumption that the *followers* will react optimally.

Applications of Stackelberg Games

- Hierarchical decision sytems
  - Government agencies
  - Large corporations with multiple susidiaries
  - Markets with a single “market-maker.”
- Parties in direct conflict
  - Zero sum games
  - Interdiction problems
- Modeling “robustness”: leader represents external phenomena that canot be controlled.
  - Weather
  - External market conditions
- Controlling optimized systems: follower represents a system that is optimized by its nature.
  - Electrical networks
  - Biological systems

Example: Tunnel Closures (Maurizio Bruglieri)

- The EU wishes to close certain international tunnels to trucks in order to increase security.
- The response of the trucking companies to a given set of closures will be to take the shortest remaining path.
- Each travel route has a certain “risk” associated with it and the EU’s goal is to minimize the riskiest path used after tunnel closures are taken into account.
- This is a classical Stackelberg game.

Example: Robust Facility Location (Snyder (2006))

- We wish to locate a set of facilities, but we want our decision to be robust with respect to possible disruptions.
- The disruptions may come from natural disasters or other external factors that cannot be controlled.
- Given a set of facilities, we will operate them according to the solution of an associated optimization problem.
- Under the assumption that at most k of the facilities will be disrupted, we want to know what the worst case scenario is.
- This is a Stackelberg game in which the leader is not a cognizant DM.

Example: Atrial Fibrilation Ablation (Phillips)

- Atrial fibrilation is a common form of heart arrhythmia that may be the result of impulse cycling within macroreentrant circuits.
- AF ablation procedures are intended to block these unwanted impulses from reaching the AV node.
- This is done by surgically removing some pathways.
- Since electrical impulses travel via the path of lowest resistance, we can model their flow using a mathematical program.
- If we wish to determine the least disruptive strategy for ablation, this is a Stackelberg game.
- In this case, the follower is not a cognizant DM.

Example: Electricity Networks (Bienstock and Verma (2008))

- As we know, electricity networks operate according to principles of optimization.
- Given a network, determining the power flows is an optimization problem.
- Suppose we wish to know the minimum number of links that need to be removed from the network in order to cause a failure.
- This, too, can be viewed as a Stacklerberg game.
- Note that neither the leader nor the follower is a cognizant DM in this case.

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.
- The follower may have several equally favorable reactions to the action of 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.

Multilevel Programming

- Multilevel programming is a generalization of standard mathematical programming that applies to hierarchical decision systems.
- In a *multilevel program*
  - The variables are divided into groups controlled by separate DMs.
  - The objective and constraints of each DM can involve the variables of DMs at higher levels in the hierarchy.
- In principle, the assumption of DM rationality allows the multilevel program to be viewed as single optimization problem.
- We focus on the case of *bilevel programs*, which can be used to model Stackleberg games.

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

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

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

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

We utilize the following notation:

Notation

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

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 $X \supset \mathbb{Z}^{n_1}$, the problem may not have a solution in general because the feasible set may not be closed.

Further Generalizations

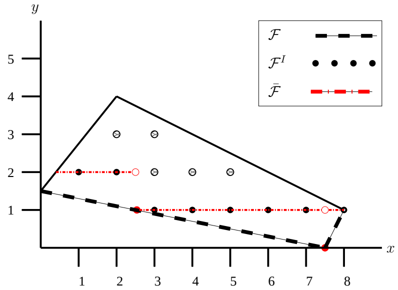
- The follower's variables may appear in the leader's constraints (see, e.g., Audet et al. (1997)).
- The follower's objective may also parameterized (see Dempe (2001)).

Example

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

$$\max_{x \in X} \quad x + 10y$$

$$\text{subject to} \quad y \in \operatorname{argmin}\{y : -25x + 20y \leq 30,\ x + 2y \leq 10,\ 2x - y \leq 15,\ 2x + 10y \geq 15,\ y \in Y\}$$



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.

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.
- This transforms the original bilevel optimization problem into a standard mathematical program.
- The optimality conditions for the lower-level optimization problem are

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

- 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.

Discrete BLPs

- When some of the variables are discrete, the situation is a bit more difficult.
- If the follower's optimization problem has a strong dual, then we can, in principle, reformulate as a single-level optimization problem.
- At best, it will now be an integer MPEC, however.

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.

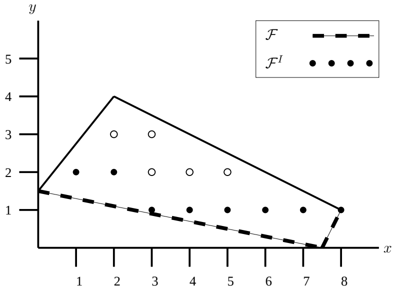
Properties 2 and 3 result from the fact that the set of feasible solutions for the original MILP is contained in the feasible set of the relaxation.

THIS IS NOT THE CASE FOR MIBLP

Example

Consider the following instance of (MIBLP) again:

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



From the figure, we can see that

1. $\mathcal{F} \subseteq \Omega$, $\mathcal{F}^I \subseteq \Omega^I$, and $\Omega^I \subseteq \Omega$
2. $\mathcal{F}^I \not\subseteq \mathcal{F}$

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.

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

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

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).

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

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

Towards a Branch and Bound Algorithm

The end goal is to develop a branch-and-bound algorithm that generalizes concepts from mixed integer linear programming.

Components

- Bounding methods (required)
- Branching methods (required)
- Search strategies (required)
- Preprocessing methods (optional)
- Primal heuristics (optional)

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

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. $y \in M^I(x)$.

Simple Observations

The following observations are used to derive two valid classes of inequalities.

Observation 1

Any inequality (π_1, π_2, π_0) valid for Ω^I is valid for $\mathcal{F}^I$.

Observation 2

Let $(x, y) \in \Omega^I$ such that $y \notin M^I(x)$. If the inequality (π_1, π_2, π_0) is valid for $\Omega^I \setminus \{(x, y)\}$ it is also valid for $\mathcal{F}^I$.

- The first result is derived from the relationship $\mathcal{F}^I \subseteq \Omega^I$ and allows us to separate fractional solutions to (LR).
- The second proposition states that we can separate points that are integer but not bilevel feasible.

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

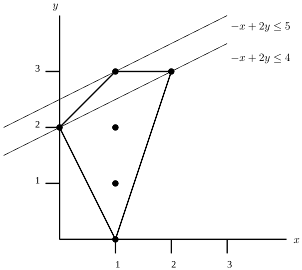
A Valid Inequality

$\sum_{i \in I} a_i' x + g_i' y \leq \sum_{i \in I} b_i - 1$ 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 $\mathcal{F}^I$.

A Complete Algorithm

- 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 a MILP is a function $z : \mathbb{R} \rightarrow \mathbb{R} \cup \{\pm\infty\}$ 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 $d \in \mathbb{R}^m$,

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

- Note that the value function is an optimal solution for the subadditive dual for any right-hand side.
- If we could express the value function in closed form, we could solve the subadditive reformulation.
- It is not known how to do this in general.

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 correction 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:

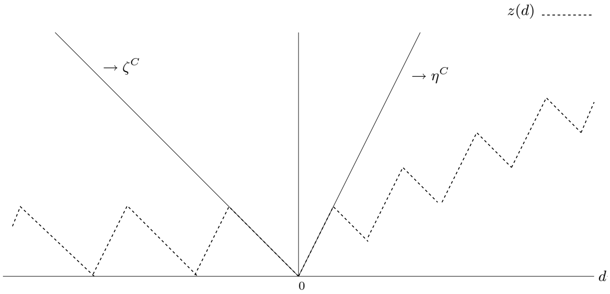
- z is composed of a finite number of linear segments on any closed interval
- the slope of each of these linear segments is given by one of two possible values.

For illustration purposes, we henceforth assume that the lower-level problem (1) contains a single equality constraint. For convenience, we also assume the upper-level contains only equality constraints.

Value Function Structure

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

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



Jeroslow Formula

- Let $M \in \mathbb{Z}_+$ be such that for any $t \in T$, $\frac{Ma_j}{a_t} \in \mathbb{Z}$ for all $j \in I$.
- Then there is a Gomory function g such that

$$z(d) = \min_{t \in T}\{g(\lfloor d \rfloor_t) + \frac{c_t}{a_t}(d - \lfloor d \rfloor_t)\}, \quad \lfloor d \rfloor_t = \frac{a_t}{M}\left\lfloor \frac{Md}{a_t} \right\rfloor, \quad \forall d \in \mathbb{R}$$

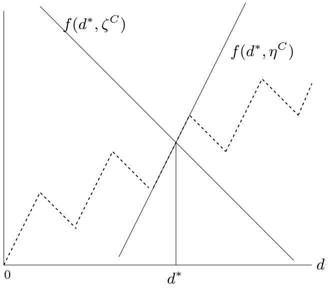
- Such a Gomory function can be obtained from the value function of a related PILP.
- For $t \in T$, setting

$$\omega_t(d) = g(\lfloor d \rfloor_t) + \frac{c_t}{a_t}(d - \lfloor d \rfloor_t) \; \forall d \in \mathbb{R},$$

  we can write

$$z(d) = \min_{t \in T} \omega_t(d) \; \forall d \in \mathbb{R}$$

Key Insight



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

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

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:

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

and

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

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*:

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

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

Numerical Example

MIBLP Example

$$
\begin{aligned}
\min \quad & 2x_1 + \frac{1}{4}x_2 + \frac{1}{2}x_3 + \frac{3}{4}x_5 + x_6 \\
& + \frac{1}{4}y_1 + \frac{3}{4}y_2 + \frac{1}{2}y_3 + y_4 + \frac{1}{2}y_6 \\
\text{subject to} \quad & x_1 + \frac{1}{2}x_2 - x_3 + \frac{3}{4}x_4 + \frac{3}{2}x_5 + \frac{1}{4}x_6 = 6 \\
& x_1, x_2, x_3 \in \mathbb{Z}_+, x_4, x_5, x_6 \in \mathbb{R}_+ \\
& y \in \operatorname{argmin} \Big\{ \frac{1}{2}y_1 + \frac{1}{2}y_3 + 2y_4 + y_5 + \frac{3}{4}y_6 : \\
& \qquad \frac{1}{2}x_1 - \frac{1}{4}x_2 + x_3 - x_4 + \frac{3}{4}x_5 + \frac{1}{4}x_6 \\
& \qquad\quad + y_1 - \frac{3}{2}y_2 + \frac{3}{2}y_3 + y_4 - y_5 + \frac{1}{3}y_6 = 4, \\
& \qquad y_1, y_2, y_3 \in \mathbb{Z}_+, y_3, y_5, y_6 \in \mathbb{R}_+ \Big\},
\end{aligned}
$$

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

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

Current Work: The General Case

- At present, we do not have an implementation for the general case.
- We would like to avoid branching on integer variables as suggested by Moore and Bard (1990).
- Instead, we would like to generalize the methods based on the structure of the value function of the lower level problem.
- In principle, there are many things that can be done using subadditive duality.
- The practicality of these remains to be seen.

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*

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

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.*

Current Work: 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.

Current Work: Interdiction Problems

A special case of interest is the *mixed integer interdiction problem* (MIPINT)

Mixed Integer Interdiction

$$\max_{x \in \mathcal{P}_U^I} \min_{y \in \mathcal{P}_L^I(x)} dy \qquad \text{(MIPINT)}$$

where

$$\mathcal{P}_U^I = \left\{x \in \mathbb{B}^n \mid A^1 x \leq b^1\right\}$$
$$\mathcal{P}_L^I(x) = \left\{y \in \mathbb{Z}^p \times \mathbb{R}^{n-p} \mid G^2 y \geq b^2, y \leq u(e-x)\right\}.$$

- When the follower's problem has network structure, we have a *network interdiction problem*.
- Existing literature focuses on variants of network interdiction problem.
- The model above allows for lower-level systems described by general MILPs.

Network Interdiction

To illustrate the case of network interdiction, we consider the problem of *maximizing the shortest path*.

Maximum Shortest Path Problem (Israeli and Wood, 2002)

- The follower tries to move through a network along a shortest path.
- The leader tries to maximize the length of that shortest path by interdicting network arcs.

Notation

- $G = (N, A)$ is a graph in which the follower moves a commodity from the source node $s \in N$ and the sink node $t \in N$.
- $\delta^-(i)$ and $\delta^+(i)$ is the set of arcs directed out of and into node i, respectively.
- $0 < c_k < \infty$ is the length of arc $k \in A$.
- $0 < r_k$ is the resource required to interdict arc $k \in A$.
- r_0 is the total *interdiction budget*.

Maximizing the Shortest Path

One formulation of this problem is:

Maximum Shortest Path Formulation

$$
\begin{aligned}
\max_{x \in X} \quad & \sum_{k \in A} (c_k + x_k d_k) y_k \\
\text{subject to} \quad & y \in \operatorname{argmin} \Big\{ \quad \sum_{k \in A} (c_k + x_k d_k) y_k : \\
& \qquad \sum_{k \in \delta^-(i)} y_k - \sum_{k \in \delta^+(i)} y_k = \begin{cases} 1 & \text{for } i = s \\ 0 & \forall i \in N \setminus \{s, t\} \\ -1 & \text{for } i = t \end{cases}, \\
& \qquad y_k \geq 0 \quad \forall k \in A \ \Big\},
\end{aligned}
$$

where $X = \{x \in \mathbb{B}^{|A|} \mid rx \leq r_0\}$ and d_k is the delay if arc k is interdicted.

Current Work: Bilevel Branching (w/ A. Lodi, S. Smriglio, and F. Rossi)

- Consider the problem of determining a branching disjunction that produces maximal bound improvement.
- If the bound is obtained by solving a mathematical prorgram, then this problem can be formulated as a bilevel program.

Current Work: Cuts for Interdiction Problems

- In the case of an interdiction problem, all upper level variables are binary.
- So-called “no good” cuts can be added after each bilevel feasibility check.
- These have had a big impact on the size of the search tree.
- We have been discussing how to generalize these cuts.
- The special structure of these problems also lends itself to customized versions of other components.

Current Work: Primal Heuristics

- For these problems, feasible solutions are relatively easy to find
- For example, replacing the upper-level objective with the lower-level objective *and optimizing over* Ω^I will produce a feasible solution.
- We can try and improve these solutions by adding cuts of the form $c^1x + d^2y \geq L$, where L is the objective value of the current incumbent solution.
- Note that once cuts are added to the original set of linear constraints, we are not guaranteed feasiblity and must test for bilevel feasibility as usual.
- However, due to the nature of the bilevel feasiblity check, we are always (eventually) guaranteed the generation of a feasible solution with this method.

Current Work: Primal Heuristics (cont.)

- Using sensitivity information on the lower level problem, local search methods can be implemented.
- We are currently working on a method utilizing objective cuts for the lower-level problem.
- From the bilevel feasiblity check, we have feasible solution $(\hat{x}, y^*)$.
- Using this information, we can optimize over Ω^I with the cut $d^2 y \leq d^2 y^*$ in an attempt to generate a good feasible solution.
- This is in an attempt to drive the solution towards feasiblity.
- As in the previous method, we must check for bilevel feasibility, but are guaranteed a feasible solution.

Current Work: Preprocessing

- We are currently working on a preprocessing method that allows us to fix variables before entering the branch-and-bound phase of our algorithm.
- This method is similar to methods used in preprocessing MILPs.
- We utilize information from the optimal basis of the original LP relaxation $z_{LP} = \max_{x,y \in \Omega} c^1 x + d^2 y$.
- Let $(\hat{x}, \hat{y})$ be a solution to this LP and $\underline{z}$ be a lower bound on the upper-level objective value,. Then, for all j such that
  - $|\bar{c}_j| \leq \underline{z} - z_{LP}$,
  - $x_j \in X$, and
  - x_j is nonbasic at its upper or lower bound,

  we can fix x_j to its current value.
- As in the primal heuristic methods, one way to determine $\underline{z}$ is to optimize over Ω^I with respect to the lower-level objective.

Conclusions and Future Work

- Preliminary testing to date has revealed that these problems can be extremely difficult to solve.
- What we have implemented so far has only scratched the surface.
- Nevertheless, we are encouraged by what we have seen.
- Much work remains to be done.
  - Preprocessing procedures
  - Reduced cost tightening
  - Generalization of known MILP search strategies (e.g., *best-estimate* of Benichou et al. (1971)).
  - More sophisticated branching rules (i.e., strong branching).

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.

Benichou, M., J. Gauthier, P. Girodet, G. Hentges, G. Ribiere, and O. Vincent 1971. Experiments in mixed-integer linear programming. *Mathematical Programming* **1**, 76–94.

Bienstock, D. and A. Verma 2008. The n-k problem in power grids: New models, formulations and computation. Available at `http://www.columbia.edu/~dano/papers/nmk.pdf`.

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.

References II

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

Israeli, E. and R. Wood 2002. Shortest path network interdiction. *Networks* **40**(2), 97–111.

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

Snyder, L. V. 2006. Facility location under uncertainty: A review. *IIE Transactions* **38**(7), 537–554.