Aggregation and Mixed Integer Rounding to Solve MILPs (Marchand and Wolsey)

Matthew Galati

SAS Institute - Analytical Solutions
Lehigh University - Department of Industrial and Systems Engineering

July 7, 2005

Classical Cutting Planes for MILP

- Special Structures (valid for certain relaxations of MILPs)
  - Knapsack / Gub Cover, Pack (many applications)
  - Flow Cover / Path (fixed charge network flow, lot-sizing, ...)
  - Cliques / Odd-Hole (set partitioning, covering, packing)
  - Implied Bound (logical implications between binary variables)
- Generic Cuts (valid for **any** MILP)
  - Gomory Mixed Integer
  - **Mixed Integer Rounding**

Mixed Integer Rounding—MIR

- Almost everything comes from considering the following very simple set, and very simple observation.

$$X = \{(x, y) \in \mathbb{R}_+ \times \mathbb{Z} \mid y \leq b + x\}$$

- Let $f = b - \lfloor b \rfloor$. Then a valid inequality for X is:

$$y \leq \lfloor b \rfloor + \frac{1}{1 - f} x$$

(Simple) Extension of MIR

$$X = \{(x, y) \in \mathbb{R}^2_+ \times \mathbb{Z}^{|N|}_+ \mid \sum_{j \in N} a_j y_j + x^+ \leq b + x^-\}$$

- $f = b - \lfloor b \rfloor$
- $f_j = a_j - \lfloor a_j \rfloor$
- The inequality

$$\sum_{j \in N} \left(\lfloor a_j \rfloor + \frac{(f_j - f)^+}{1 - f} \right) y_j \leq \lfloor b \rfloor + \frac{x^-}{1 - f}$$

  is valid for X.
- X is a one-row relaxation of a general *mixed* integer program, where all of the continuous variables have been aggregated into two variables (one with positive coefficients, one with negative coefficients).

Gomory Cuts are special cases of MIRs

- Consider the set

$$X^{=} = \{(x, y_0, y) \in \mathbb{R}^2_+ \times \mathbb{Z} \times \mathbb{Z}^{|N|}_+ \mid y_0 + \sum_{j \in N} a_j y_j + x^+ - x^- = b\}$$

  which is essentially the row of an LP tableau with y_0 the basic variable and x^+, x^- the sum of the continuous variables with positive and negative coefficients.

- Relax the equality to an inequality and apply MIR.

Proof.

$$y_0 + \sum_{j \in N} \left(\lfloor a_j \rfloor + \frac{(f_j - f)^+}{1 - f} \right) y_j \leq \lfloor b \rfloor + \frac{x^-}{1 - f} \quad (1)$$

$$b - \sum_{j \in N} a_j y_j - x^+ + x^- + \sum_{j \in N} \left(\lfloor a_j \rfloor + \frac{(f_j - f)^+}{1 - f} \right) y_j \leq \lfloor b \rfloor + \frac{x^-}{1 - f} \quad (2)$$

$$-b + \sum_{j \in N} a_j y_j + x^+ - x^- - \sum_{j \in N} \left(\lfloor a_j \rfloor + \frac{(f_j - f)^+}{1 - f} \right) y_j \geq -\lfloor b \rfloor - \frac{x^-}{1 - f} \quad (3)$$

$$\sum_{j \in N} f_j y_j + x^+ - x^- - \sum_{j \in N} \frac{(f_j - f)^+}{1 - f} y_j \geq f - \frac{x^-}{1 - f} \quad (4)$$

$$\sum_{j \in N} f_j y_j + x^+ + \frac{f}{1 - f} x^- - \sum_{j \in N_2} \frac{f_j - f}{1 - f} y_j \geq f \quad (5)$$

$$\sum_{j \in N_1} f_j y_j + x^+ + \frac{f}{1 - f} x^- + \sum_{j \in N_2} (f_j - \frac{f_j - f}{1 - f}) y_j \geq f \quad (6)$$

- (6) is the Gomory Mixed Integer Cut.

Residual Capacity Cuts are special cases of MIRs

$$X^F = \{(x, y) \in \mathbb{R}_+ \times \mathbb{Z}_+ \mid x \leq Cy, x \leq d\}$$

- X^F is a structure common to many network design models.
- The *Residual Capacity Inequality* for X^F is a MIR Inequality.

$$x \leq d - (d - C(\eta - 1))(\eta - y)$$

with $\eta = \lceil \frac{d}{C} \rceil$

Mixed Cover Cuts are special cases of MIRs

$$X^B = \{(s, y) \in \mathbb{R}_+ \times \mathbb{B}^{|N|} \mid \sum_{j \in N} a_j y_j \leq b + s\}$$

- X^B is the mixed 0-1 knapsack set, and is common in many applications.
- The *Mixed Cover Inequality* for X^B is a MIR Inequality.

$$\sum_{j \in E(C)} \min(a_j, \lambda) y_j \leq -\lambda + \sum_{j \in C} \min(a_j, \lambda) + s$$

where $C \subseteq N$ with $\sum_{j \in C} a_j = b + \lambda, \max_{j \in C} a_j > \lambda > 0$, and $E(C) = C \cup \{j \in N \setminus C \mid a_j \geq \max_{j \in C} a_j\}$.

Weight Inequalities are special cases of MIRs

$$X^W = \{(x,y) \in \mathbb{R}_+^{|P|} \times \mathbb{Z}_+^{|N|} \mid \sum_{j \in P} a_j x_j + \sum_{j \in N} a_j y_j \leq b, \quad x_j \leq u_j \text{ for } j \in P, \quad y_j \leq h_j \text{ for } j \in N\}$$

- X^W is a general mixed knapsack set.
- The *Weight Inequality* for X^W is a MIR Inequality.

$$\sum_{j \in Q} a_j x_j + \sum_{j \in T} a_j y_j + \sum_{j \in N \setminus T} (a_j - \mu)^+ y + j \leq b - \mu$$

where $T \subseteq N, Q \subseteq P$ with $\mu = b - \sum_{j \in Q} a_j u_j - \sum_{j \in T} a_j h_j > 0$ and $\max_{j \in N \setminus T} a_j > \mu$.

Mixed Knapsack Set and c-MIR Inequalities

$$X^{MK} = \{(s, y) \in \mathbb{R}_+ \times \mathbb{Z}_+^{|N|} \mid \sum_{j \in N} a_j y_j \leq b + s, y_j \leq u_j \text{ for } j \in N\}$$

- X^{MK} is called a *Mixed Knapsack Set.*
- X^{MK} is a one-row relaxation of a general *mixed* integer program, where all of the continuous variables have been aggregated into two variables (one with positive coefficients s^+, one with negative coefficients s^-. Then, the row is relaxed more by removing s^+ and setting $s = s^-$.
- Let (T, C) be some partition of N and $\delta > 0$. To form a *complemented MIR* (c-MIR) inequality for X^{MK}.
    1. Complement the variables in C with their upper bounds.
    2. Divide through by some scale factor $\delta > 0$.
    3. Apply mixed integer rounding to the resulting set.

Mixed Knapsack Set and c-MIR Inequalities

- Let $\beta = (b - \sum_{j \in C} a_j u_j)/\delta$.
- For a given partition (T, C) and scale factor δ, this results in the following set

$$X^{MK}_{(T,C),\delta} = \{(s, y) \in \mathbb{R}_+ \times \mathbb{Z}^{|N|}_+ \mid \sum_{j \in T} \frac{a_j}{\delta} y_j + \sum_{j \in C} \frac{-a_j}{\delta} y_j \leq \beta + \frac{s}{\delta}\}$$

with the following c-MIR inequality

$$\sum_{j \in T} G(\frac{a_j}{\delta}) y_j + \sum_{j \in C} G(\frac{-a_j}{\delta})(u_j - y_j) \leq \beta + \frac{s}{\delta(1-f)}$$

where $f = \beta - \lfloor \beta \rfloor$, and $G(d) = \lfloor d \rfloor + \frac{(f_d - f)^+}{1-f}$, with $f_d = d - \lfloor d \rfloor$.

Separation Procedure for MILPs

Classify the original constraints into mixed integer rows M, and add slacks to get equalities.

$$\sum_{j \in P} a^i_j x_j + \sum_{j \in N} g^i_j y_j = b_i \text{ for } i \in M$$

1. *Aggregation:* Combine rows $S \subseteq M$ to obtain a single mixed integer constraint.
2. *Bound Substitution:* Introduce slack variables for the simple or variable bounds to form a mixed knapsack set X^{MK}.
3. *c-MIR Separation:* Look for violated c-MIRs for the set X^{MK}.
4. If no violated inequality is found, and $|S| < MAXAGGR$, return to Step 1, else Stop.

Question: What is a good choice for $MAXAGGR$?

Step 1 Aggregation

Suppose a set of rows $S \subseteq M$ has been combined

$$\sum_{j \in P} \alpha_j x_j + \sum_{j \in N} \gamma_j y_j = \beta$$

1. Choose a list of variables that creates a link to a new constraint.

$$P^* = \{k \in P \mid \alpha_k \neq 0, l_k y_k^* < x_k^* < u_k y_k^* \text{ and } \exists r \in M \setminus S \text{ with } a_k^r \neq 0\}$$

2. Choose the variable furthest from its bounds.

$$\kappa = \text{arg}_{k \in P^*} \max\{\min\{x_k^* - l_k y_k^*, u_k y_k^* - x_k^*\}\}$$

3. Choose a row $r \in M \setminus S$ with nonzero linkage $a_\kappa^r \neq 0$.
4. Aggregate the row r and the original row so that the coefficient of x_κ becomes 0, that is, add $\frac{-\alpha_\kappa}{a_\kappa^r}$ times row r.

Step 2 Bound Substitution

1. Substitute either $x_j = l_j y_j + t_j$ or $y_j = u_j y_j - t_j$, leading to the new constraint

$$\sum_{j \in P} \delta'_j t_j + \sum_{j \in N} \gamma'_j y_j = \beta'$$

2. Let $s = \sum_{j \in P \,:\, \delta'_j < 0} (-\delta'_j) t_j \geq 0$, to obtain

$$X^{MK} = \{(s,y) \in \mathbb{R}_+ \times \mathbb{Z}_+^{|N|} \mid \sum_{j \in N} \gamma'_j y_j \leq \beta' + s, y_j \leq u_j \text{ for } j \in N\}$$

Question: How do we decide which bound to substitute?

Step 2 Bound Substitution

The authors suggest 3 heuristic suggestions:

- Minimize the difference between each continuous variable and its bound. That is, *take the one that is currently the tightest*.
- Minimize the value of $s^* = \sum_{j \in P \,:\, \delta'_j < 0} (-\delta'_j) t^*_j$
- Minimize the value of $\sum_{j \in P \,:\, \delta'_j > 0} (\delta'_j) t^*_j$

Step 3 c-MIR Separation

Given X^{MK} and a fractional point $(s*, y*)$, find a violated c-MIR inequality, if one exists.

$$X^{MK} = \{(s,y) \in \mathbb{R}_+ \times \mathbb{Z}_+^{|N|} \mid \sum_{j \in N} \gamma'_j y_j \leq \beta' + s, y_j \leq u_j \text{ for } j \in N\}$$

1. Choose $C = \{j \in N \mid y_j^* \geq \frac{u_j}{2}\}$ and $T = N \setminus C$.
2. Choose $\delta \in D = \{a_j \mid j \in N \text{ and } 0 \leq y_j^* < u_j\}$. Let $\delta^* \in D$ be the value which generates the *most violated* c-MIR inequality.
3. Try to increase the violation by dividing δ^* by $2, 4, 8$.
4. Try to increase the violation by successively complementing (switch from C to T) each variable between its bounds, ordered nondecreasing $|y_j^* - \frac{u_j}{2}|$.

Branch and Cut - Computational Tradeoffs

Cutting Planes

- Recognizing special structure. Generic cuts are usually weaker than structured cuts.
- The more (good) cuts, the better the bound, the smaller the search tree, the slower the LP solve.
  - So, restrict the number of cuts allowed. How?
  - If we restrict, then in what order should we generate them?
- How much time should we spend in the cut generation phase? (versus LP, branching, etc)
- How many cut passes should we attempt in root/leaf nodes?
- When do we stop cutting and start branching? Tail off.
- How often should we look for cuts? Every node? Every n nodes? Should this be different for different classes of cuts?

Branch and Cut - Computational Tradeoffs

Cutting Planes (Cont.)

- Density of cuts? The more dense, the harder LP factorize gets.
- Stability (max/min coefficients) of cuts? Can cause ill-conditioning, numerical issues, roundoff in LP.
- Can we share cuts across nodes (when is a cut globally valid?)
- What denotes a *good* cut?
- Should we keep all of our generated cuts in the LP?
  - Maybe we should remove slack ones each iteration?
  - Can the removed cuts ever come back into play? Should we keep them around? Cut pools.

Branching - many more considerations (see Jeff Linderoth's thesis)
Heuristics - many more considerations (see ?? - wide open area)
LP - many more considerations (see Ilog/Cplex)

Other Simple Ideas - MIRs

- *Joao Goncalves - Informs04* - Produce additional candidate sets X^{MK} by multiplying the aggregated row by -1.
- *Some Weird Guy at SAS* - In the step where we aggregate constraints, after constructing a variable link, we might have a choice of several rows. The authors choose arbitrarily (the first one). What if we choose the *tightest* given the current fractional point?
- What else?? There are many possible variations to consider.

Computational Experiments MIRs

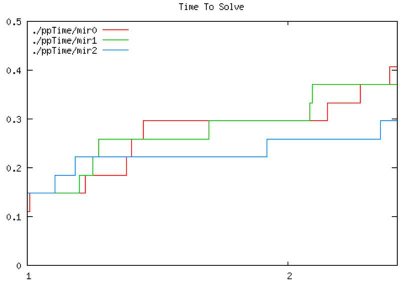
Some variants we consider:

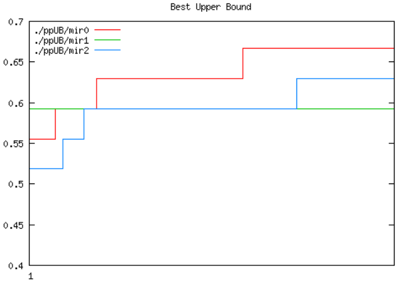
- Looking at additional candidates (multiply by -1).
- Choose the tightest row when aggregating.
- How many rows should we consider to aggregate?
- Which bound substitution heuristic gives the best results?
- Experiment:
  - 27 MIPLIB instances (those with some violated MIR)
  - 600s time cutoff, all other defaults

Computational Experiments MIRs

Looking at additional candidates (multiply by -1).

Run	Settings	NumSolve	AveCPU-M
mir0	NegOne, Tight, Agg=3, Sub=A	11	1.0%
mir1	NoNegOne, Tight, Agg=3, Sub=A	10	0.7%
mir2	no MIRs	8	

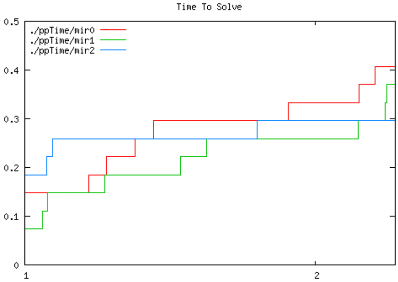


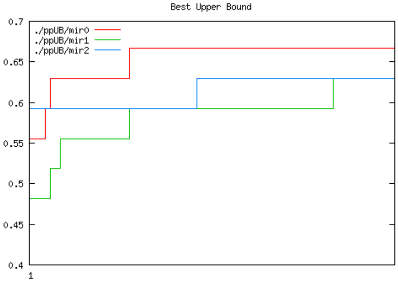


Computational Experiments MIRs

Choose the tightest row when aggregating.

Run	Settings	NumSolve	AveCPU-M
mir0	NegOne, Tight, Agg=3, Sub=A	11	1.0%
mir1	NegOne, NoTight, Agg=3, Sub=A	10	1.4%
mir2	no MIRs	8	

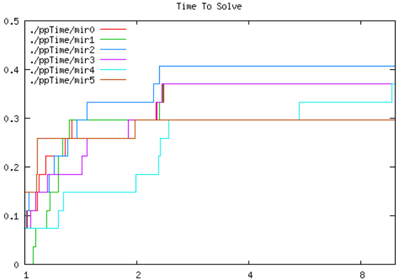


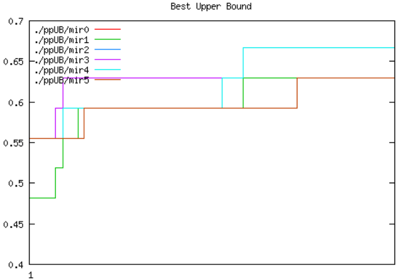


Computational Experiments MIRs

How many rows should we consider to aggregate?

Run	Settings	NumSolve	AveCPU-M
mir0	NegOne, Tight, Agg=1, Sub=A	10	0.8%
mir1	NegOne, Tight, Agg=2, Sub=A	10	0.9%
mir2	NegOne, Tight, Agg=3, Sub=A	11	1.0%
mir3	NegOne, Tight, Agg=4, Sub=A	10	1.3%
mir4	NegOne, Tight, Agg=5, Sub=A	10	1.3%
mir5	no MIRs	8	





Computational Experiments MIRs

Which bound substitution heuristic gives the best results?

Run	Settings	NumSolve	AveCPU-M
mir0	NegOne, Tight, Agg=3, Sub=A	11	1.0%
mir1	NegOne, Tight, Agg=3, Sub=B	10	1.2%
mir2	NegOne, Tight, Agg=3, Sub=C	11	2.3%
mir3	no MIRs	8