Integral Simplex Using Decomposition

Presented by:
Issmail El Hallaoui

Joint Work With:
Abdelouahab Zaghrouti
François Soumis

International Workshop on Column Generation 2012
The set partitioning problem (SPP)

Happy example: wedding

http://www.coin-or.org/PuLP/CaseStudies/a_set_partitioning_problem.html
Let $G$ be the set of guests. Each guest must be assigned a table. The tables are “feasible” subsets of $G$. We wish to maximize the total \textit{happiness} of all of the tables.
The set partitioning problem (SPP)

Mathematical formulation

Minimize $c \mathbf{x}$

Constraint matrix with binary coefficients

$A \mathbf{x} = \mathbf{1}$

$x_j \in \{0, 1\}$

Usually solved by branch and bound and cut and price and ...
The set partitioning problem (SPP)

Applications

• Weddings all over the world
• Transportation

Pairings

Locomotives

Bus driver scheduling

Trucking

• Clustering: data clustering, sensor clustering, ...
The story

- The story began in 1969 with Trubin when he observed that the polytope $Q$ of SPP is *quasi-integral*.

  **Meaning that:** every edge of $\text{conv}(Q')$ is also an edge of $Q$ where $Q'$ is the set of its integer points.

\[
\begin{align*}
\text{conv}(Q') & : \quad x_1 \quad x_2 \quad x_3 \\
Q & : \quad x_1 \quad x_2 \quad x_3 \quad x_f \quad x_3
\end{align*}
\]
• **Interpretation**: there exists a path from $x_0$ to $x^*$ where vertices $x_j$ are all integer.

• **Trubin** (1969) proposed a variant of the simplex algorithm to solve this kind of problems.
The story

Balas and Padberg quoted in 1972 criticizing Trubin:

Trubin, using a completely different line of reasoning, shows that all edges of the convex hull of the feasible integer solutions to Q are also edges of the feasible set of Q’. This property (interesting in itself) then implies the existence of a path containing only integer vertices between any two integer vertices of the feasible set. However, neither the upper bound on the number of pivots required to get from one integer vertex to the other nor the existence of a minimum-length path whose associated objective function values form a monotonic sequence follows directly from Trubin's result.
The story

• Balas and Padberg (1972) proved the existence of a decreasing sequence of integer solutions leading to the optimal solution with at most $m$ pivots.

• But in practice...

\[
\begin{align*}
\text{Cost} \\
x_0 \\
x_1 \\
x^*
\end{align*}
\]

Huge degeneracy
Pivot on $\bar{a}_{ij} < 0$ !!!

\[
\begin{align*}
0 & \quad 1 \quad 2 \quad \text{Iterations}
\end{align*}
\]

Pivot on $\bar{a}_{ij} = 1$
The story

• Yemelichev, Kovalev and Kravtsov referred in 1984 (translation) to the algorithm as \textbf{Integral Simplex} for the first time.

• Thompson used in 2002 the integral simplex to solve instances with up to \textbf{163} constraints.

• Rönnberg and Larsson developed in 2009 an extension of the integral simplex method to the column generation context.
How?

Properties

The decomposition

The integral simplex using decomposition
The decomposition

$S_0 = \{A1, A2, A3\} \text{ of cost 107}$

Example: $G = \{g1, g2, g3, g4, g5, g6\}$

- $A4 = A2 \cup A3$
- $A4$ is compatible of lower cost 4 ($< 2 + 100$)
**Definition:** if $A_k$ is union of some columns of $S$, $A_k$ is said to be compatible with $S$.

Example: $G = \{g_1, g_2, g_3, g_4, g_5, g_6\}$

$S_0 = \{A_1, A_2, A_3\}$ of cost 107

$\overline{S}_0 = A - S_0$
The reduced problem:
contains columns *compatible* with the partition

The reduced problem improves the current integer solution in polynomial time.
How to escape the local optimum?

- Balas and Padberg *branch or pivot on negative coefficients* (degenerate variables) !!!
The combination $C_1 = \{A5, A6\}$ is compatible and minimal.

- $A5 \cup A6 = A1 \cup A4$
- The decomposition $g_1, g_2, g_4, g_4, g_5, g_6, g_3, g_5, g_6, g_4, g_5$,
**Definition:**

- A combination $C_k$ is said to be compatible if it is union of some columns of $S$.
- $C_k$ is minimal if it becomes incompatible by removing any column from it.
The decomposition

$S^* = \{A_5, A_6\}$ of cost 7

$S^* = A - S^*$

How could we find such combinations?
The complementary problem (CP): contains columns \textit{incompatible} with the partition.

\[ \min \bar{c} \mathbf{x} \]

Compatibility Matrix

\( M \mathbf{x} = 0 \)

\( \sum_j x_j = 1 \)

\( x_j \geq 0 \)

\( A_j \) such that \( x_j > 0 \) are \textit{disjoint.
Integral simplex using decomposition algorithm (ISUD)

- Initial integer Solution
- Reduced problem (RP)
  - Combinations of columns with negative reduced costs
  - Dual values
  - Integer pivots
- Complementary problem (CP)
**Proposition:** ISUD is a Dantzig-Wolfe decomposition of the set partitioning problem with $Z^* = Z_{DW}$.
**Theorem:** ISUD is exact and guarantees a decreasing sequence of integer solutions leading to the optimal solution.

Properties

**Convergence**

No degeneracy, no pivoting on negative coefficients
**Properties**

**Minimal combinations**

- $S_0$ : Set of columns of initial solution
- $S^*$ : Set of columns of optimal solution
- $C^*$ : Optimal combination ($C^* = S^* - S_0$)

**Proposition**

- $C^* = \bigcup_{i=1}^{k} C_i$ is union of minimal combinations
- The complementary problem (CP) finds minimal combinations

**ISUD is intrinsically parallelizable**
Properties

Integrality of combinations

- Minimal combination
- Min mean reduced cost
- Low degree of incomp.

ISUD favors integrality
If the current solution changes, the undesired combination (non-orthogonal) becomes incompatible. It cannot be generated by CP, the cut is no longer needed. Easy handling and less cutting.
The know how developed in metaheuristics could be recycled here
Numerical results

So what?

Conclusion

Open question
Numerical results

• Tests on large instances up to **1600** constraints (instead of **163** of Thompson 2002) and **500 000** variables.

• The complementary problem often finds combinations of
  – disjoint columns (**50%-90%**)  
  – small size (in average \(<= 10\) columns/combination)

• ISUD finds optimal solutions in **75%** of cases within **20** minutes.
  – CPLEX takes 10 hours on the easiest large instance (gap 0)
  – CPLEX finds no feasible solutions for the hardest ones
Open question

It’s all about...

Compatibility

\[ M \leq \n \Rightarrow \quad \exists c = 0 \]
The degree of incompatibility of a variable depends on M.

**Question:** could we find a compatibility matrix allowing to generate optimal disjoint combinations in polynomial time?
Conclusion

• Proof of concept showing high potential

• **Ongoing projects:**
  – Extensive experimentation and refinement
  – Local cuts for set partitioning problem
  – Parallel version of ISUD
  – ISUD with cost projection
Conclusion

The story continues...

Thank you