Decomposition Methods for Large-Scale Discrete Optimization

Matthew V Galati

Ted K Ralphs

Department of Industrial and Systems Engineering

Lehigh University, Bethlehem, PA, USA

November 20, 2002

Outline

- Preliminaries, Decomposition Methods
- Decomposition Algorithm
- Implementation and Extensions
- Summary

Preliminaries

- Consider the following integer linear program (ILP):

$$\min\{c^\top x : x \in \mathcal{F}\}$$

where

$$\begin{aligned} \mathcal{F} &= \{x \in \mathbb{Z}^n : A'x \geq b', A''x \geq b''\} & \mathcal{Q} &= \{x \in \mathbb{R}^n : A'x \geq b', A''x \geq b''\} \\ \mathcal{F}' &= \{x \in \mathbb{Z}^n : A'x \geq b'\} & \mathcal{Q}' &= \{x \in \mathbb{R}^n : A'x \geq b'\} \\ & & \mathcal{Q}'' &= \{x \in \mathbb{R}^n : A''x \geq b''\} \end{aligned}$$

- Denote $\mathcal{P} = conv(\mathcal{F})$ and $\mathcal{P}' = conv(\mathcal{F}')$.
- Assume that optimization (separation) over $\mathcal{P}$ is *difficult*.
- Assume that optimization (separation) over $\mathcal{P}'$ can be done *effectively*.

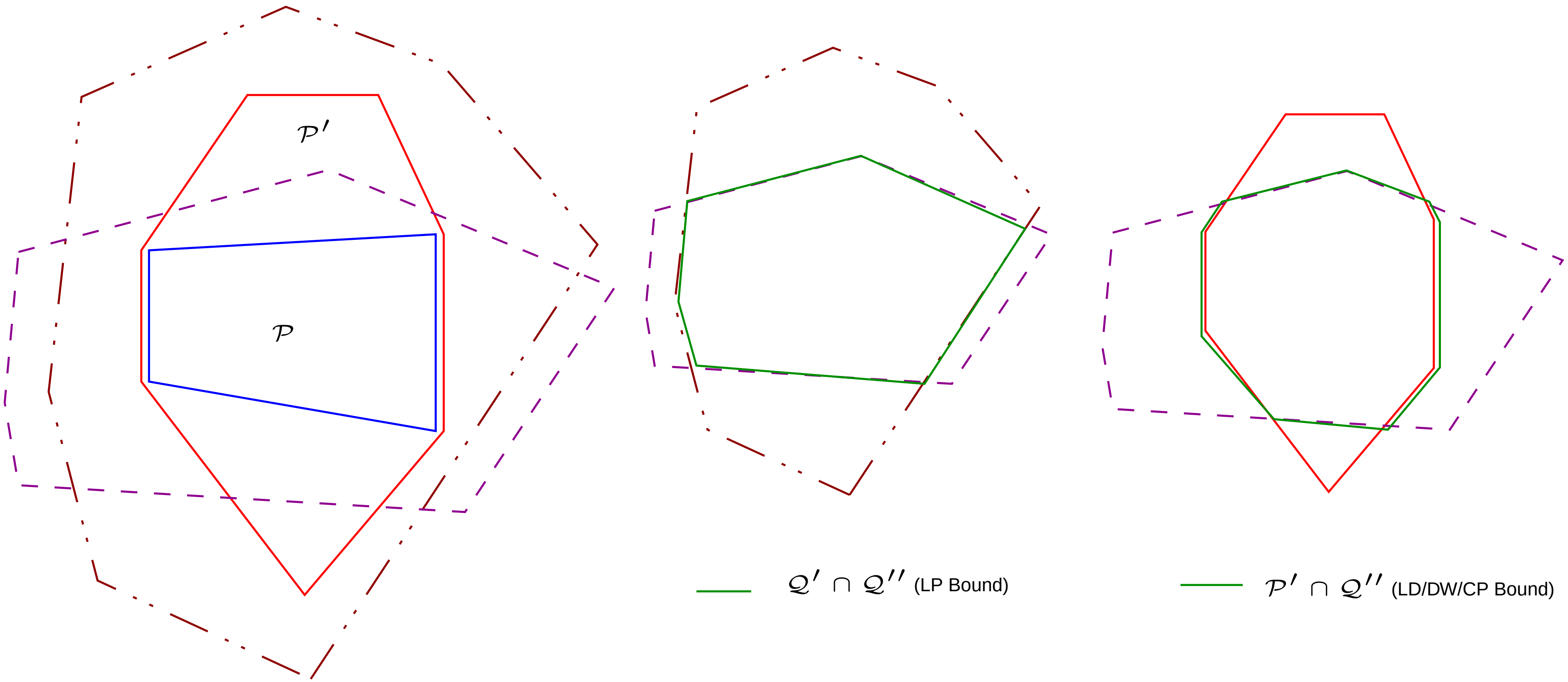
Decomposition Methods

- Develop lower bounds to embed in a branch and bound (BB) framework.
- Goal: Improve the bound given by the initial LP relaxation

$$\min\{c^\top x : x \in \mathcal{Q}\}$$

- Traditional Decomposition Methods
  - Dantzig-Wolfe Decomposition
  - Lagrangian Relaxation
  - Cutting Plane Methods

Polyhedra, LP Bound, LD/DW/CP Bound



$\mathcal{P} = conv(\{x \in \mathbb{Z}^n : Ax \geq b\})$

$\mathcal{P}' = conv(\{x \in \mathbb{Z}^n : A'x \geq b'\})$

$\mathcal{Q}' = \{x \in \mathbb{Q}^n : A'x \geq b'\}$

$\mathcal{Q}'' = \{x \in \mathbb{Q}^n : A''x \geq b''\}$

Dantzig-Wolfe Decomposition

- Explicitly enforce membership in $\mathcal{Q}''$
- Implicitly enforce membership in $\mathcal{P}'$

$$\min\{c(\sum_{f \in \mathcal{F}'} f\lambda_f) : A''(\sum_{f \in \mathcal{F}'} f\lambda_f) \geq b'', \sum_{f \in \mathcal{F}'} \lambda_f = 1, \lambda_f \geq 0 \; \forall f \in \mathcal{F}'\}$$

- The *optimal fractional solution* $\hat{x}$ and *optimal decomposition* $\hat{\lambda}$

$$\hat{x} = \sum_{f \in \mathcal{F}'} f\hat{\lambda}_f \in \mathcal{P}'$$

- Solution method: column generation
- Subproblem: optimize over $\mathcal{P}'$

Lagrangian Relaxation

- Explicitly enforce membership in $\mathcal{F}'$
- Implicitly enforce membership in $\mathcal{Q}''$

$$
\begin{aligned}
L_R(u) &= \min\{(c - uA'')f + ub'' : f \in \mathcal{F}'\} \\
L_D &= \max_{u \geq 0} L_R(u)
\end{aligned}
$$

- Solution method: subgradient optimization ($L_R(u)$ is concave in u)
- Subproblem: $L_R(u)$ - optimize over $\mathcal{P}'$

- Rewriting L_D as a linear program is dual to the Dantzig-Wolfe LP

$$
\max_{u_0, u \geq 0} \{u_0 : u_0 \leq (c - uA'')f + ub'', \forall f \in \mathcal{F}'\}
$$

Cutting Plane Methods

- Explicitly enforce membership in $\mathcal{Q}''$
- Implicitly enforce membership in $\mathcal{P}'$
- Weyl's Theorem - existence of $[A'_I, b'_I]$ that describe $\mathcal{P}'$

$$\min\{cx : A'_I x \geq b'_I, A''x \geq b''\}$$

- Solution method: cutting planes
- Subproblem: separation over $\mathcal{P}'$

Equivalence of Traditional Decomposition Methods

- Equivalence of bounds (yet implementation very different)

$$\max_{u \geq 0} \min\{(c - uA'')f + ub'' : f \in \mathcal{F}'\} \quad (LD)$$

$$= \max_{u_0, u \geq 0} \{u_0 : u_0 \leq (c - uA'')f + ub'', \forall f \in \mathcal{F}'\}$$

$$= \min\{c(\sum_{f \in \mathcal{F}'} f\lambda_f) : A''(\sum_{f \in \mathcal{F}'} f\lambda_f) \geq b'', \sum_{f \in \mathcal{F}'} \lambda_f = 1, \lambda_f \geq 0 \; \forall f \in \mathcal{F}'\} \quad (DW)$$

$$= \min\{cx : x \in \mathcal{P}', A''x \geq b''\}$$

$$= \min\{cx : A'_I x \geq b'_I, A''x \geq b''\} \quad (CP)$$

$$\geq \min\{cx : A'x \geq b', A''x \geq b''\} \quad (LP)$$

Improving the LD/DW/CP Bound

- Let $\mathcal{L}$ be a class of valid inequalities for $\mathcal{P}$.
- Consider the setting:
  - The optimization problem over $\mathcal{P}$ is difficult.
  - The optimization problem over $\mathcal{P}'$ can be solved effectively.
  - Separation of a fractional solution from $\mathcal{P}$ using members of $\mathcal{L}$ is difficult.
  - Separation of a member of $\mathcal{F}'$ from $\mathcal{P}$ using members of $\mathcal{L}$ can be solved effectively.
- Does this occur in practice? Yes

Vehicle Routing Problem

ILP Formulation:

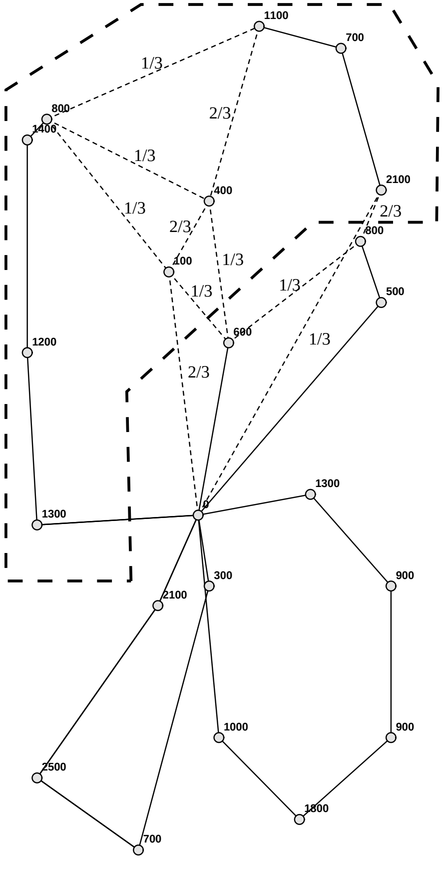
$$\begin{array}{rcll} \sum_{e\in\delta(0)} x_e & = & 2k & & (1)\\ \sum_{e\in\delta(i)} x_e & = & 2 & \forall i \in V \setminus \{0\} & (2)\\ \sum_{e\in\delta(S)} x_e & \geq & 2b(S) & \forall S \subset V \setminus \{0\},\ |S| > 1 & (3) \end{array}$$

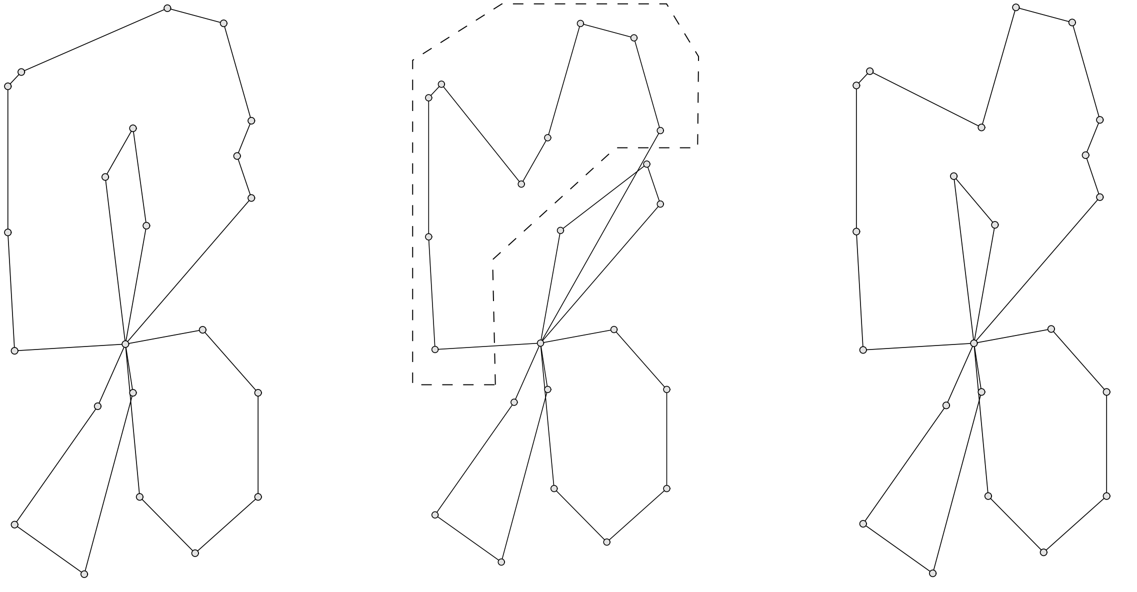
$$\begin{array}{rcl} b(S) & = & \text{lower bound on the number of trucks required to service } S \\ & = & \lceil (\sum_{i\in S} d_i)/C \rceil \text{ (normally)} \end{array}$$

- Relaxations:
  - Multiple Traveling Salesman Problem: Set $C = \sum_{i\in S} d_i$.
  - k-Tree: Set $C = \sum_{i\in S} d_i$. Relax (2) but leave $\sum_{e\in E} x_e = n + k$.
- Separation of the GSECs (3) in this formulation is $\mathcal{NP}$-Complete.
- Given the incidence vector of an MTSP or a k-Tree ($\in \mathcal{F}'$), we can easily determine whether it satisfies all of these inequalities.

Example of Decomposition

VRP/k-TSP





Steiner Problem in Graphs

ILP Formulation:

$$\begin{array}{lcll} \sum_{e \in \bar{E}} x_e & = & |\bar{V}| - 1 & & (4) \\ \sum_{e \in \bar{E}(S)} x_e & \leq & |S| - 1 & \forall S \subseteq \bar{V} & (5) \\ x_{0i} + x_{ij} & \leq & 1 & \forall i \in V \setminus T, \{i, j\} \in \delta(i) & (6) \end{array}$$

- Relaxation: Minimum Spanning Tree - Relax (6).
- Lifted Subtour Elimination Constraints:

$$\begin{array}{lcll} \sum_{e \in \bar{E}(S)} x_e + \sum_{i \in S \setminus T} x_{0i} & \leq & |S| - 1 & \forall S \subseteq \bar{V}, S \cap T \neq \emptyset & (\mathcal{L}_1) \\ \sum_{e \in \bar{E}(S)} x_e + \sum_{i \in S \setminus k} x_{0i} & \leq & |S| - 1 & \forall S \subseteq \bar{V}, S \cap T \neq \emptyset, k \in S & (\mathcal{L}_2), \\ \sum_{e \in \bar{E}(X,Y)} x_e & \leq & 1 & \forall X, Y \subseteq \bar{V}, \ \ X \cap T \neq \emptyset & \\ & & & Y \cap T \neq \emptyset, \ \ X \cap Y = \emptyset & (\mathcal{L}_3) \end{array}$$

- Separation of $\mathcal{L}_1, \mathcal{L}_2$ and $\mathcal{L}_3$ for a fractional point: $O(|\bar{V}|^4)$ - Gomory-Hu.
- Separation of $\mathcal{L}_1, \mathcal{L}_2$ and $\mathcal{L}_3$ for a MST: $O(|\bar{E}|)$ - Breadth-First Search.

Dynamic Cut Generation in Lagrangian Relaxation

- Consider a solution ($f \in \mathcal{F}'$) to the Lagrangian dual L_D.
- Attempt to separate f from $\mathcal{P}$ using the class of inequalities $\mathcal{L}$.
- If successful, dualize the violated inequalities on the fly.
- Origins of relax and cut:
  - Christofidos and Balas, “A Restricted Lagrangean Approach to the TSP”, *Mathematical Programming* 1981
  - M.L. Fisher, “Optimal Solution of VRPs Using Minimum K-Trees”, Wharton Tech Report 1990
- Applications: Couple Constrained Assignment Problem [Aboudi 91], Steiner Problem in Graphs [Lucena 92], Quadratic Knapsack [Palmeria et al 99], Edge-Weighted Clique Problem [Hunting et al 01] and Rectangular Partitions [Calheiros et al 02] (coming - survey paper by Lucena).
- Guignard/Ralphs: Adding these inequalities might not improve the bound.
- Why? Using subgradient method, cannot obtain the optimal fractional solution - so, without reoptimization cannot tell if bound will improve.

Dynamic Cut Generation using Dantzig-Wolfe Decomposition

Proposition 1 *Let* $\hat{x} \in \mathbb{R}^n$ *such that* $\hat{x} = \sum_{f \in \mathcal{F}'} f \lambda_f, \sum_{f \in \mathcal{F}'} \lambda_f = 1$, *and* $\lambda_f \geq 0 \; \forall f \in \mathcal{F}'$. *If* $(a, b) \in \mathcal{L}$ *and* $a^\top \hat{x} < b$, *then for some* $f \in \mathcal{F}'$ *with* $\lambda^f > 0$, $a^\top f < b$.

- Consider an optimal fractional solution $\hat{x} \in \mathbb{R}^n$ to the D-W LP, and let $D = \{f \in \mathcal{F}' : \hat{\lambda}_f > 0\}$.
- Attempt to separate each $f \in D$ from $\mathcal{P}$ using the class of inequalities $\mathcal{L}$.
- In this case, we can check for improvement in the bound without reoptimization.
- Origins:
  - T.K. Ralphs, “Parallel Branch and Cut for VRP”, Cornell Thesis 1995
  - L. Kopman, “A New Generic Separation Routine and Its Application In a Branch and Cut Algorithm for the CVRP”, Cornell Thesis 1999
  - T.K. Ralphs, L. Kopman, W.R. Pulleyblank and L.E.Trotter Jr., “On the CVRP”, *Mathematical Programming* 2001.

Outline

- Preliminaries, Decomposition Methods
- Decomposition Algorithm
- Implementation and Extensions
- Summary

Decomposition-Based Separation Algorithm

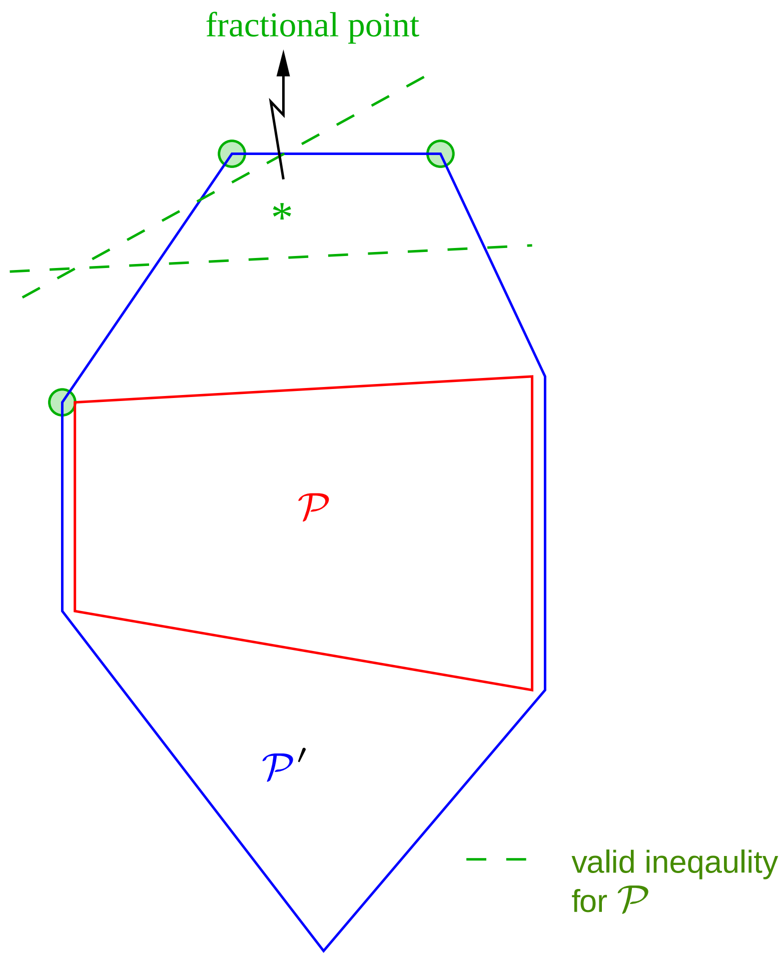
Input: $\hat{x} \in \mathbb{R}^n$
Output: A valid inequality for $\mathcal{P}$ which is violated by $\hat{x}$, if one is found.

- **Step 0.** Apply separation algorithms and heuristics for $\mathcal{P}$ and $\mathcal{P}'$. If one of these returns a violated inequality, then output the violated inequality and STOP.
- **Step 1.** Otherwise, attempt to decompose $\hat{x}$ into a convex combination of members of $\mathcal{F}'$ by solving the LP

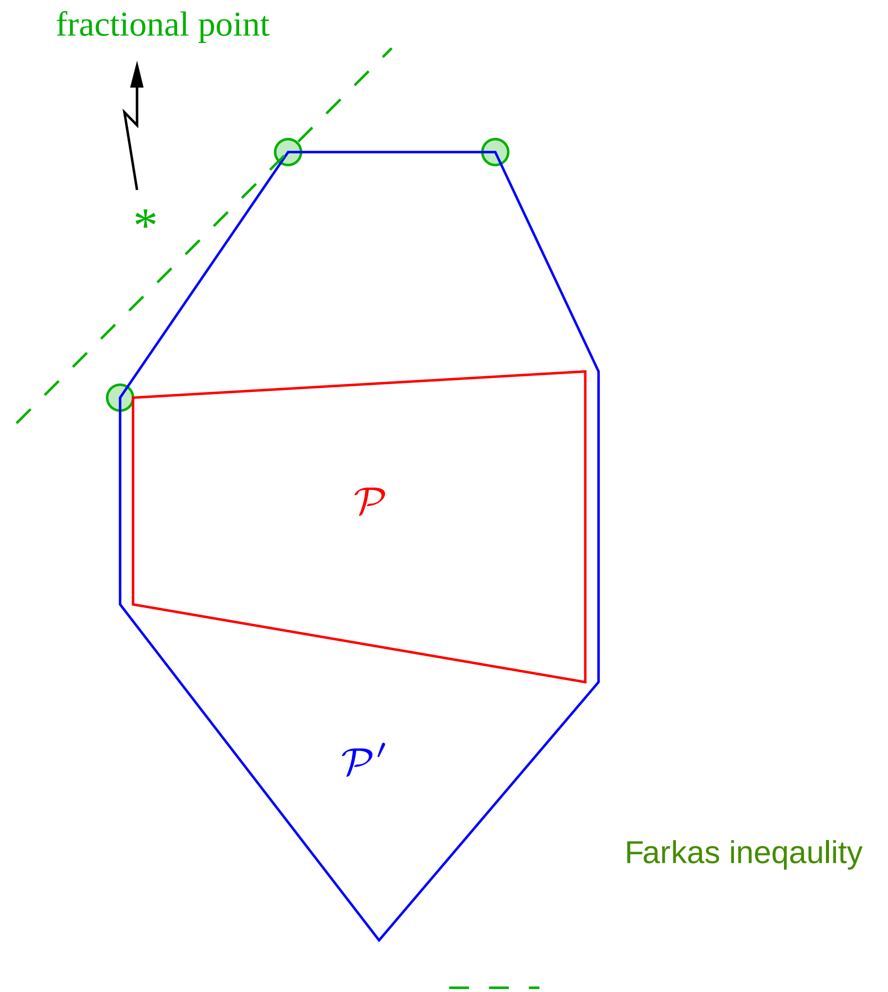
$$\max\{\mathbf{0}^\top \lambda \ : \ T\lambda = \hat{x}, \mathbf{1}^\top \lambda = 1, \lambda \geq 0\}, \tag{2}$$

  where T is a matrix whose columns are members of $\mathcal{F}'$.
- **Step 2a.** If a decomposition $\hat{\lambda}$ exists, let D represent the columns of T participating in the convex combination of $\hat{x}$. Scan the members of $\mathcal{F}'$ corresponding to the columns in D. For each inequality in $\mathcal{L}$ violated by a column of D, check whether it is also violated by $\hat{x}$. If a constraint violated by $\hat{x}$ is encountered, output it and STOP.
- **Step 2b.** If a decomposition does not exist, output a Farkas inequality $(\alpha, -\gamma)$ for $\mathcal{P}'$ that is violated by $\hat{x}$ and STOP.

Example of Separation Using Decomposition



(a) $\hat{x} \in \mathcal{P}'$



(b) $\hat{x} \notin \mathcal{P}'$

Column Generation Subroutine

Input: $\hat{x} \in \mathbb{R}^n$

Output: Either (1) a decomposition of $\hat{x}$ into members of $\mathcal{F}'$
or (2) a valid inequality for $\mathcal{P}'$ violated by $\hat{x}$

- **Step 1.0**. Generate a matrix T' containing a small subset of members of $\mathcal{F}'$ and set $T = T'$.
- **Step 1.1**. Solve (2) using the dual simplex. If this LP is feasible, output the members of $\mathcal{F}'$ participating in the decomposition, then STOP.
- **Step 1.2**. Otherwise, let r be the row in which the dual unboundedness condition was discovered, and let (α, γ) be the r^{th} row of the basis inverse. Optimize over $\mathcal{P}'$ with cost vector α. Let $t^* \in \mathcal{F}'$ be the result.
- **Step 1.3**. If $\alpha t^* + \gamma < 0$, then t^* is a column eligible to enter the basis. Add t^* to T and go to 1.1. Otherwise, the LP is infeasible and $\hat{x} \notin \mathcal{P}'$. Output the Farkas inequality $(\alpha, -\gamma)$ and STOP.

Outline

- Preliminaries, Decomposition Methods
- Decomposition Algorithm
- Implementation and Extensions
- Summary

Questions

- How can we better deal with the column generation subroutine?
  - Can we do a better job initializing our search?
  - Can we be smarter about adding members to T?
- How can we be smarter in our search for *effective* members of $\mathcal{L}$?
- What is the trade-off between the strength of the relaxation and the ease of finding violated members of $\mathcal{L}$?
- Can we take advantage of distributed systems?

Implementation Extensions

- Assume COP, i.e., $x \in \mathbb{B}^n$. Let $E = \{1, ..., n\}$.
- Consistency Condition - restrict column generation of T

Proposition 2 *Assume* $\hat{x} = \sum_{f \in \mathcal{F}'} f \lambda_f, \sum_{f \in \mathcal{F}'} \lambda_f = 1$ *and* $\lambda_f \geq 0 \; \forall f \in \mathcal{F}'$. *Let* $D = \{f \in F' : \lambda_f > 0\}$. *Then, for each* $f \in D, \forall e \in E$, *the following two conditions must be true*

$$\text{if } \hat{x}_e = 1, \text{ then } f_e = 1;$$

$$\text{if } \hat{x}_e = 0, \text{ then } f_e = 0$$

- Define $M \geq \max\{\gamma - \alpha x : x \in \mathcal{P}\}$ (a LB for opt over $\mathcal{P}$ with cost α).

$$\sigma_e = \begin{cases} M & \text{if } e \in E_0 = \{e : \hat{x}_e = 0\}; \\ -M & \text{if } e \in E_1 = \{e : \hat{x}_e = 1\}; \\ \alpha & \text{otherwise} \end{cases}$$

for $e \in E$ and $\beta = \gamma - \sum_{e \in E_1} (-M - \alpha_e)$.

Implementation Extensions

- Lifting the Farkas Inequalities
  - bigM Method
  - Sequential Lifting
- For general ILPs, decomposition into members of $\mathcal{F}$ [Kopman 99]
  - Column generation subproblem is an optimization problem over $\mathcal{P}$!!
  - Given a proof (Farkas cut) that $\hat{x} \notin \mathcal{P}$, this will improve the bound.
  - D. Applegate, R. Bixby, V. Chvátal, and W. Cook, "TSP Cuts Which Do Not Conform to the Template Paradigm", Computational Combinatorial Optimization 2001

Implementation Extensions

- Initialization of T:
  - Consider the solution to a LP relaxation $\hat{x}_0 = T_0\lambda$, where T_0 is the set of enumerated columns from which the optimal decomposition was found.
  - Let $\hat{x}_1$ be the solution to the next LP relaxation.
  - If $\|\hat{x}_1 - \hat{x}_0\| \leq \epsilon$, initialize $T_1 = T_0$,
    else, continue as before (setting T_1 randomly / brute force).
- Column Pool:
  - Analogous to *cut pools* in branch and cut.
  - Keep a set of columns that have participated in a decomposition *recently* and use this (or part of this) to initialize T.
- Is it necessary to be exact in solving the column generation problem? No
  - Try optimizing over $\mathcal{P}'$ heuristically first - all we need is a column that destroys the proof of infeasibility.
  - Also, solve a relaxed version of the feasibility problem
    $\max\{\mathbf{0}^\top\lambda \;:\; \hat{x} - \epsilon \leq T\lambda \leq \hat{x} + \epsilon, \mathbf{1}^\top\lambda = 1, \lambda \geq 0\}$,

Implementation Extensions

- Metrics for Effectiveness
  - Consider a solution $\hat{x} \in \mathbb{R}^n$ to the initial LP, $D = \{f \in \mathcal{F}' : \hat{\lambda}_f > 0\}$.
  - The class $\mathcal{L}$ might contain an exponential number of inequalities that violate f, only some of which violate $\hat{x}$.
  - Sort D from least to greatest by normalized distance : $\|\hat{x} - f\|$
  - Spend more computation time enumerating inequalities for extreme points *closer* to the fractional point.
- Parallel Implementation
  - Column generation subproblem is an optimization problem over $\mathcal{P}'$ - `for each` dual ray.
  - Separation problem `for each` extreme point in $D \subseteq \mathcal{F}'$.

DECOMP Library

- An abstract base class in C++ which can be derived for a specific application.
- *pure virtual methods*, the user must derive
  - ```
    virtual double optimize_over_relaxation(
                   const vector<double> & cost,
                   vector<variable> & solution_relax)
    ```
  - ```
    virtual void separate_member_of_relaxation(
                 const vector<variable> & solution_lp,
                 const vector<variable> & solution_relax,
                 vector<rc_vector> & new_rows)
    ```
- *virtual methods*, the user can derive
  - ```
    virtual void find_initial_columns_decomp(
                 vector<rc_vector> & initial_columns)
    ```
  - ```
    virtual void lift_farkas_inequality(
                 rc_vector & farkas)
    ```

Outline

- Preliminaries, Decomposition Methods
- Decomposition Algorithm
- Implementation and Extensions
- Summary

Summary

- DECOMP Library - base algorithm
- Computational Study - previous work / relax and cut literature
  - VRP (k-TSP or k-Tree : GSECs, Combs, Multistars)
  - Couple Constrained Assignment Problem (AP : SVIs)
  - Steiner Problem in Graphs (Minimum Spanning Tree : Lifted SECs)
- Computational Study - proposed new applications
  - Edge-Weighted Clique Problem (Tree Relaxation : Trees, Cliques)
  - 2-Connected Network with Bounded Rings (2-CN : ??)
  - VRP with Time Windows (kTSPTW : k-Path Cuts) - Galexis
  - Generalized Assignment Problem (m-Knap or AP : SVIs) - FabTime
  - Service Constrained Network Flow (Network Flow : ??) - IBM
  - Generalized Minimum Spanning Tree (MST : ??)

Summary

- Implementation - previous extensions (generalize from COPs to ILPs)
  - Consistency condition
  - Lifting of the Farkas inequality
- Alternative methods for lifting the Farkas inequalities.
- Explore the use of decomposition into members of $\mathcal{F}$.
- Implementation - proposed new extensions
  - Column pool
  - Heuristic solution to column generation subproblem
  - Metric for effectiveness
  - Parallel implementation
- Theoretical ties between relax and cut and DECOMP.