Decomposition Part, Week 1

ISE Department
COR@L Lab
Lehigh University
zelihakca@lehigh.edu

01/26/07
We will Cover:

- Dantzig-Wolfe decomposition and column generation in IP.
- 2 applications of DW decomposition: maximum independent set problem, multistage stochastic IP.
- Lagrangian relaxation (this week or next week).
- Formulation of Benders decomposition.
- Global optimization of nonconvex MINP with decomposable structures.
- Any suggestions? Additions? Not interesting?
- Volunteers?
- Suggestions for the structure of the seminar?
Introduction

Decomposition In General

1. DW Decomposition
2. Lagrangian relaxation
3. Benders decomposition

- All are large scale (models with large number of constraints/variables) optimization algorithms.
- We can find better bounds for branch and bound using these approaches.
- For the first 2, alternative approaches are cutting plane, variable redefinition.
- Basically, first 2 decomposition includes 3 steps:
  - Decompose system of inequalities into two parts.
  - Find the convex hull of the second system.
  - Optimize first system over this convex hull.
Why decomposition can find a better bound?


\[
\text{min}_{cx}
\]
\[
\text{s.t. } Ax \leq b
\]
\[
x \in \mathbb{Z}^n
\]

Let \( Q = \{ x \in \mathbb{R}^n | Ax \leq b \} = \{ x \in \mathbb{R}^n | A'' x \leq b'' , A' x \leq b' \} \).

Let \( F = Q \cap \mathbb{Z}^n \) (feasible integer points) and \( P \) be the convex hull of \( F \).

Let \( Q' = \{ x \in \mathbb{R}^n | A' x \leq b' \} \), \( F' = Q' \cap \mathbb{Z}^n \) (feasible integer points) and \( P' \) be the convex hull of \( F' \).

Let \( Q'' = \{ x \in \mathbb{R}^n | A'' x \leq b'' \} \).

\( z_{LP} = \text{min}_{x \in Q} \{ cx \} \).

\( z_{DECOMP} = \text{min}_{x \in P'} \{ cx | A'' x \leq b'' \} = \text{min}_{x \in P' \cap Q''} \{ cx \} \).
Why decomposition can find a better bound?

Dantzig Wolfe Decomposition

- Assume \( P' \) is bounded. Let \( E \subseteq F' \) be the set of extreme points of \( P' \).

\[
    P' = \{ x \in \mathbb{R}^n \mid x = \sum_{s \in E} s \lambda_s, \sum_{s \in E} \lambda_s = 1, \lambda_s \geq 0 \ \forall s \in E \}.
\]

- Dantzig Wolfe formulation and bound:

\[
    z_{DW} = \min_{x \in \mathbb{R}^n} \{ c^T x \mid A''x \geq b'', x = \sum_{s \in E} s \lambda_s, \sum_{s \in E} \lambda_s = 1, \lambda_s \geq 0 \ \forall s \in E \}.
\]

- Substitute new variables:

\[
    z_{DW} = \min_{\lambda \in \mathbb{R}^E_+} \{ c^T \left( \sum_{s \in E} s \lambda_s \right) \mid A'' \left( \sum_{s \in E} s \lambda_s \right) \geq b'', \sum_{s \in E} \lambda_s = 1 \}.
\]

- \(|E|\) may be very large, thus \( E \) should be generated dynamically.

Introduction

Cutting Strips Problem

\[
\begin{align*}
\min & \quad \sum_{k=1}^{K} L^k \left( W^k y^k - \sum_{i=1}^{p} w_i z_i^k \right) \\
\text{[P^cs]} & \quad \text{s.t.} \\
& \quad \sum_{k=1}^{K} L^k z_i^k \geq d_i, \quad i = 1, \ldots, p, \\
& \quad \sum_{i=1}^{p} w_i z_i^k \leq W^k y^k, \quad k = 1, \ldots, K, \\
& \quad y^k \in \{0, 1\}, \quad k = 1, \ldots, K, \\
& \quad z_i^k \in \mathbb{N}, \quad i = 1, \ldots, p, \quad k = 1, \ldots, K.
\end{align*}
\]

Let $Q^k$ be the set of feasible cutting patterns for sheet $k$, $\lambda_q^k$ be the number of times pattern $q^k$ is selected.

$$X^{cs} = \left\{ (y^k, z^k)_{k=1,\ldots,K} \in \mathbb{N}^{K(1+p)} : \right\}$$

$$\sum_{i=1}^{p} w_i z_i^k \leq W^k y^k \text{ for } k = 1, \ldots, K$$

$$= \left\{ (y^k, z^k)_{k=1,\ldots,K} \in \mathbb{R}^{K(1+p)} : \right\}$$

$$y^k = \sum_{q^k \in Q^k} \lambda_q^k, \quad z^k = \sum_{q^k \in Q^k} q^k \lambda_q^k,$$

$$\sum_{q^k \in Q^k} \lambda_q^k \leq 1 \quad \forall k, \quad \lambda_q^k \in \{0, 1\} \quad \forall q^k \in Q^k, k$$
Dantzig Wolfe Decomposition of Cutting Strips Problem

\[
\min \sum_{k=1}^{K} \sum_{q^k \in Q^k} c^k_q \lambda^k_q
\]

\[
[M^{cs}] \quad \text{s.t.} \quad \sum_{k=1}^{K} \sum_{q^k \in Q^k} L^k q^k_i \lambda^k_q \geq d_i, \quad i = 1, \ldots, p, \quad (6)
\]

\[
\sum_{q^k \in Q^k} \lambda^k_q \leq 1, \quad k = 1, \ldots, K,
\]

\[\lambda^k_q \in \{0, 1\}, \quad q^k \in Q^k, k = 1, \ldots, K.\]
Master problem may include infinitely many cutting patterns:

\[ Q^k = \{ q^k \in \mathbb{N}^p : \sum_{i} w_i q_i^k \leq W \} \]

Start with “a” set of feasible cutting patterns → RMP.

Generate the column with most negative reduced cost to improve the bound.

Let \( \pi_i \) be the dual variable for demand constraint \( i \), \( \mu_k \) be the dual variable for convex comb. constr. Subproblem:

\[
\zeta^k(\pi, \mu) = \min L^k \left( W^k - \sum_{i=1}^{p} (w_i + \pi_i) z_i \right) + \mu_k \\
\text{s.t.} \\
\sum_{i=1}^{p} w_i z_i \leq W^k, \\
z_i \in \mathbb{N}, \quad i = 1, \ldots, p,
\]
Lower bound for DW Decomposition LP

If column generation subproblem is solved to optimality:

\[ z_{DW}(LP)^{LB} = \text{(obj. value of dual master)} \]
\[ + \text{(obj. function value of optimal column generation subproblem)} \]

Let \( q \) be the optimal solution to the column generation problem.

\[ z_{DW}(LP)^{LB} = \left( \sum_{i=1..p} d_i \pi_i + \sum_{k=1..K} \mu_k \right) + \left( \sum_{k=1..K} c_q q_k - \sum_{k=1..K,i=1..p} L^k q_i - \sum_{k=1..K} \mu_k \right) \]
Some Points

- Lagrangian relaxation and Dantzig Wolfe decomposition is used to find better bounds for the MIP.
- Lagrangian relaxation, Dantzig Wolfe decomposition, cutting plane, variable redefinition can be used to get the same bound.
- Cutting plane: convex hull of second system $P'$ is defined by finding the facet defining inequalities of the system.
- Master LP (from Dantzig Wolfe Decomp.) = dual formulation of the Lagrangian dual that results from dualizing the $A''x \leq b''$.
- Dantzig Wolfe decomposition leads to models with large number of variables which requires column generation algorithm.
- Variable redefinition: develop an alternative formulation $Z$ for the polyhedron $P'$. 
The first DW example is **convexification**.

Cutting Strip problem uses **discretization**.

**Convexification:** $\lambda_s > 0$,

**Discretization:** $\lambda_s \in \{0, 1\}$ if it is an extreme point, $\lambda_s \in \mathbb{N}$ if it is an extreme ray.

Convexification: $x = \sum_s s\lambda_s$ does not imply $x$ is integer. To express integrality, must return to original variables. Branching should be in original variables.

Discretization: $x = \sum_s s\lambda_s$ results integer variables. Can do branching or write cuts in terms of $\lambda_s$.

Both give the same LP relaxation of master.

Both has the same IP master if variables are binary.

Discretization $\neq$ convexification if some variables are general integer.
General form of master problem.
\[ \sum_s \lambda_s = 1 \] can be changed with \( \sum_s \lambda_s \leq 1 \) if 0 vector is a feasible solution for the problem.

**Independent subsystems:** \( Dx \leq d \rightarrow D^k x \leq d^k, \; k = 1, \ldots, K. \)
\[
\sum_{q \in Q(k)} \lambda_q = 1, \; k = 1, \ldots, K.
\]

**Identical independent subsystems:** \( Dx \leq d \rightarrow D^k x \leq d^k, \; k = 1, \ldots, K, \) but \( D^k = \bar{D}, \; d^k = \bar{d} \) for \( k = 1 \ldots K. \)
\[
\sum_{q \in \bar{Q}} \lambda_q = K.
\]
Let $P$ be the original problem, $M$ be the master problem.
Both have the same set of feasible integer points.
Since the representation of solution is different in both formulations, a solution $x$ for $P$ may not result a unique solution $\lambda$ for $M$.
Cases when the solutions do not corresponds to a unique solution is given in the paper.
$Z_{LP}(P) \leq Z_{LP}(M) \leq Z_{IP}$.
$Z_{LP}(P) < Z_{LP}(M)$ if subsystem does not have integrality property.
At each branch and bound node, a **feasible** LP solution is required. Especially with branching it becomes difficult to maintain feasibility. One way is to use artificial variables that will always result feasible solutions.

Designing **branching rules**: Branching rule must be incorporated into subproblem (column generation problem).

**RMP**: Optimize RMP to get dual variables. Dual solution is not **unique** if primal is degenerate. Dual variables effect the generated column.

- Initialization is necessary with simplex and for other methods it can reduce **heading-in effect** of initially producing bad duals which may cause irrelevant columns.
- Solution with primal, primal-dual simplex, barrier?
- Simplex based column generation has **tailing off effect** (poor convergence).