

DM872 – Spring 2020
Math Optimization at Work

Advanced Methods for MILP

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

[Partly based on slides by David Pisinger, DIKU (now DTU)]

Outline

1. Dantzig-Wolfe Decomposition

2. Delayed Column Generation

Outline

1. Dantzig-Wolfe Decomposition

2. Delayed Column Generation

Dantzig-Wolfe Decomposition

Motivation: Large difficult IP models

⇒ split them up into smaller pieces

Applications

- Cutting Stock problems
- Multicommodity Flow problems
- Facility Location problems
- Capacitated Multi-item Lot-sizing problem
- Air-crew and Manpower Scheduling
- Vehicle Routing Problems
- Scheduling

Leads to methods also known as:

- Branch-and-price (column generation + branch and bound)
- Branch-and-cut-and-price (column generation + branch and bound + cutting planes)

Dantzig-Wolfe Decomposition

From an original or **compact** formulation to an **extensive** formulation made of a **master problem** and a **subproblem**

- + Tighter bounds
- + Better control of subproblem
- Model may become (very) large

Delayed column generation


Write up the decomposed model gradually as needed


- Generate a few solutions to the subproblems
- Solve the master problem to LP-optimality
- Use the dual information to find most promising solutions to the subproblem
- Extend the master problem with the new subproblem solutions.

Motivation: Cutting stock problem

- Infinite number of raw stocks, having length L .
- Cut m piece types i , each having width w_i and demand b_i .
- Satisfy demands using least possible raw stocks.

Example:

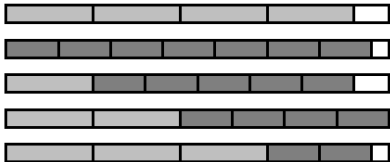
- $w_1 = 5, b_1 = 7$ 

- $w_2 = 3, b_2 = 3$ 

- Raw length $L = 22$



Some possible cuts



Formulation 1

minimize $u_1 + u_2 + u_3 + u_4 + u_5$

subject to $5x_{11} + 3x_{12} \leq 22u_1$

$5x_{21} + 3x_{22} \leq 22u_2$

$5x_{31} + 3x_{32} \leq 22u_3$

$5x_{41} + 3x_{42} \leq 22u_4$

$5x_{51} + 3x_{52} \leq 22u_5$

$x_{11} + x_{21} + x_{31} + x_{41} + x_{51} \geq 7$

$x_{12} + x_{22} + x_{32} + x_{42} + x_{52} \geq 3$

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

$x_{ij} \in \mathbb{Z}_+$

LP-relaxation gives solution value $z = 2$ with

$$u_1 = u_2 = 1, x_{11} = 2.6, x_{12} = 3, x_{21} = 4.4$$

Block structure

min	u_1		$+u_2$	$+u_3$	$+u_4$	$+u_5$	
8.1	x_{11}	x_{12}	$+x_{21}$	$+x_{31}$	$+x_{41}$	$+x_{51}$	≥ 7
			$+x_{22}$	$+x_{32}$	$+x_{42}$	$+x_{52}$	≥ 3
	$5x_{11} + 3x_{12} - 22u_1$		$5x_{21} + 3x_{22} - 22u_2$		$5x_{31} + 3x_{32} - 22u_3$		≤ 0
					$5x_{41} + 3x_{42} - 22u_4$		≤ 0
					$5x_{51} + 3x_{52} - 22u_5$		≤ 0

Formulation 2

The matrix A contains all different cutting patterns

All (undominated) patterns:

$$A = \begin{pmatrix} 4 & 0 & 1 & 2 & 3 \\ 0 & 7 & 5 & 4 & 2 \end{pmatrix}$$

Problem

minimize $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$

subject to $4\lambda_1 + 0\lambda_2 + 1\lambda_3 + 2\lambda_4 + 3\lambda_5 \geq 7$

$0\lambda_1 + 7\lambda_2 + 5\lambda_3 + 4\lambda_4 + 2\lambda_5 \geq 3$

$\lambda_j \in \mathbb{Z}_+$

LP-relaxation gives solution value $z = 2.125$ with

$$\lambda_1 = 1.375, \lambda_4 = 0.75$$

Due to integer property a lower bound is $\lceil 2.125 \rceil = 3$.

Optimal solution value is $z^* = 3$.

Round up LP-solution getting heuristic solution $z_H = 3$.

Decomposition

If model has “block” structure

$$\begin{aligned}
 \max \quad & c^1 x^1 + c^2 x^2 + \dots + c^K x^K \\
 \text{s.t.} \quad & A^1 x^1 + A^2 x^2 + \dots + A^K x^K = b \\
 & D^1 x^1 + \leq d_1 \\
 & + D^2 x^2 \leq d_2 \\
 & \dots \leq \vdots \\
 & D^K x^K \leq d_K \\
 & x^1 \in \mathbb{Z}_+^{n_1} \quad x^2 \in \mathbb{Z}_+^{n_2} \quad \dots \quad x^K \in \mathbb{Z}_+^{n_K}
 \end{aligned}$$

Lagrangian relaxation

Objective becomes

$$\begin{aligned}
 & c^1 x^1 + c^2 x^2 + \dots + c^K x^K \\
 & - \lambda (A^1 x^1 + A^2 x^2 + \dots + A^K x^K - b)
 \end{aligned}$$

Decomposed into

$$\begin{aligned}
 \max \quad & c^1 x^1 - \lambda A^1 x^1 + c^2 x^2 - \lambda A^2 x^2 + \dots + c^K x^K - \lambda A^K x^K + b \\
 \text{s.t.} \quad & D^1 x^1 + \leq d_1 \\
 & + D^2 x^2 \leq d_2 \\
 & \dots \leq \vdots \\
 & D^K x^K \leq d_K \\
 & x^1 \in \mathbb{Z}_+^{n_1} \quad x^2 \in \mathbb{Z}_+^{n_2} \quad \dots \quad x^K \in \mathbb{Z}_+^{n_K}
 \end{aligned}$$

Model is separable

Dantzig-Wolfe decomposition

If model has “block” structure

$$\begin{array}{llll}
 \max & c^1 x^1 & + & c^2 x^2 & + \dots + & c^K x^K \\
 \text{s.t.} & A^1 x^1 & + & A^2 x^2 & + \dots + & A^K x^K & = & b \\
 & D^1 x^1 & & & & & \leq & d_1 \\
 & & & + & D^2 x^2 & & \leq & d_2 \\
 & & & & & \dots & \leq & \vdots \\
 & & & & & & D^K x^K & \leq d_K \\
 & x^1 \in \mathbb{Z}_+^{n_1} & x^2 \in \mathbb{Z}_+^{n_2} & \dots & x^K \in \mathbb{Z}_+^{n_K}
 \end{array}$$

Describe each set $X^k, k = 1, \dots, K$

$$\begin{array}{llll}
 \max & c^1 x^1 & + & c^2 x^2 & + \dots + & c^K x^K \\
 \text{s.t.} & A^1 x^1 & + & A^2 x^2 & + \dots + & A^K x^K & = & b \\
 & x^1 \in X^1 & x^2 \in X^2 & \dots & x^K \in X^K
 \end{array}$$

where $X^k = \{x^k \in \mathbb{Z}_+^{n_k} : D^k x^k \leq d_k\}$

Assuming that X^k has finite number of points $\{x^{k,t}\} t \in T_k$

$$X^k = \left\{ \begin{array}{l} x^k \in \mathbb{R}^{n_k} : x^k = \sum_{t \in T_k} \lambda_{k,t} x^{k,t}, \\ \sum_{t \in T_k} \lambda_{k,t} = 1, \\ \lambda_{k,t} \in \{0, 1\}, t \in T_k \end{array} \right\}$$

Dantzig-Wolfe decomposition

Substituting X^k in original model getting *Master Problem*

$$\begin{aligned} \max & c^1 \left(\sum_{t \in T_1} \lambda_{1,t} x^{1,t} \right) + c^2 \left(\sum_{t \in T_2} \lambda_{2,t} x^{2,t} \right) + \dots + c^K \left(\sum_{t \in T_K} \lambda_{K,t} x^{K,t} \right) \\ \text{s.t.} & A^1 \left(\sum_{t \in T_1} \lambda_{1,t} x^{1,t} \right) + A^2 \left(\sum_{t \in T_2} \lambda_{2,t} x^{2,t} \right) + \dots + A^K \left(\sum_{t \in T_K} \lambda_{K,t} x^{K,t} \right) = b \end{aligned}$$

$$\sum_{t \in T_k} \lambda_{k,t} = 1 \quad k = 1, \dots, K$$

$$\lambda_{k,t} \in \{0, 1\}, \quad t \in T_k \quad k = 1, \dots, K$$

Strength of linear master model

Solving LP-relaxation of master problem, is equivalent to
(Wolsey Prop 11.1)

$$\begin{array}{ll} \max & c^1 x^1 + c^2 x^2 + \dots + c^k x^k \\ \text{s.t.} & A^1 x^1 + A^2 x^2 + \dots + A^k x^k = b \\ & x^1 \in \text{conv}(X^1) \quad x^2 \in \text{conv}(X^2) \quad \dots \quad x^k \in \text{conv}(X^k) \end{array}$$

Proof: Consider LP-relaxation

$$\begin{array}{ll} \max & c^1 \left(\sum_{t \in T_1} \lambda_{1,t} x^{1,t} \right) + c^2 \left(\sum_{t \in T_2} \lambda_{2,t} x^{2,t} \right) + \dots + c^K \left(\sum_{t \in T_K} \lambda_{K,t} x^{K,t} \right) \\ \text{s.t.} & A^1 \left(\sum_{t \in T_1} \lambda_{1,t} x^{1,t} \right) + A^2 \left(\sum_{t \in T_2} \lambda_{2,t} x^{2,t} \right) + \dots + A^K \left(\sum_{t \in T_K} \lambda_{K,t} x^{K,t} \right) = b \\ & \sum_{t \in T_k} \lambda_{k,t} = 1 \quad k = 1, \dots, K \\ & \lambda_{k,t} \geq 0, \quad t \in T_k \quad k = 1, \dots, K \end{array}$$

Informally speaking we have

- joint constraint is solved to LP-optimality
- block constraints are solved to IP-optimality

Strength of Lagrangian relaxation

- z^{LPM} be LP-solution value of master problem
- z^{LD} be solution value of lagrangian dual problem

(Theorem 11.2)

$$z^{LPM} = z^{LD}$$

Proof: Lagrangian relaxing joint constraint in

$$\begin{array}{llllll} \max & c^1 x^1 & + & c^2 x^2 & + \dots + & c^K x^K \\ \text{s.t.} & A^1 x^1 & + & A^2 x^2 & + \dots + & A^K x^K = b \\ & D^1 x^1 & + & & & \leq d_1 \\ & & + & D^2 x^2 & & \leq d_2 \\ & & & & \dots & \leq \vdots \\ & & & & & D^K x^K \leq d_K \\ x^1 \in \mathbb{Z}_+^{n_1} & x^2 \in \mathbb{Z}_+^{n_2} & \dots & x^K \in \mathbb{Z}_+^{n_K} \end{array}$$

Using result next page

$$\begin{array}{llllll} \max & c^1 x^1 & + & c^2 x^2 & + \dots + & c^k x^k \\ \text{s.t.} & A^1 x^1 & + & A^2 x^2 & + \dots + & A^k x^k = b \\ & x^1 \in \text{conv}(X^1) & x^2 \in \text{conv}(X^2) & \dots & x^k \in \text{conv}(X^k) \end{array}$$

Strength of Lagrangian Relaxation (section 10.2)

Integer Programming Problem

$$\begin{array}{ll}\text{maximize} & cx \\ \text{subject to} & Ax \leq b \\ & Dx \leq d \\ & x_j \in \mathbb{Z}_+, \quad j = 1, \dots, n\end{array}$$

Lagrange Relaxation, multipliers $\lambda \geq 0$

$$\begin{array}{ll}\text{maximize} & z_{LR}(\lambda) = cx - \lambda(Dx - d) \\ \text{subject to} & Ax \leq b \\ & x_j \in \mathbb{Z}_+, \quad j = 1, \dots, n\end{array}$$

for best multiplier $\lambda \geq 0$

$$\boxed{\max \left\{ cx : Dx \leq d, x \in \text{conv}(Ax \leq b, x \in \mathbb{Z}_+) \right\}}$$

Outline

1. Dantzig-Wolfe Decomposition

2. Delayed Column Generation

Delayed Column Generation

Delayed column generation, linear master

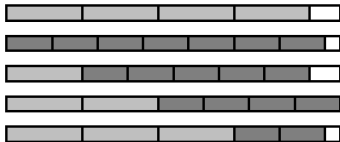
- Master problem can (and will) contain many columns
- To find bound, solve LP-relaxation of master
- Delayed column generation gradually writes up master

Delayed column generation, linear master

- $w_1 = 5, b_1 = 7$
- $w_2 = 3, b_2 = 3$
- Raw length $L = 22$



Some possible cuts



In matrix form

$$A = \begin{pmatrix} 4 & 0 & 1 & 2 & 3 & \cdots \\ 0 & 7 & 5 & 4 & 2 & \cdots \end{pmatrix}$$

LP-problem

$$\begin{aligned} \min \quad & cx \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

where

- $b = (7, 3)$,
- $x = (x_1, x_2, x_3, x_4, x_5, \dots)$
- $c = (1, 1, 1, 1, 1, \dots)$.

Revised Simplex Method

- $\max \{cx \mid Ax \leq b, x \geq 0\}$
- $B = \{1 \dots m\}$ basic variables
- $N = \{m+1 \dots m+n\}$ non-basic variables (will be set to lower bound 0)
- $A_B = [A_1 \dots A_m]$
- $A_N = [A_{m+1} \dots A_{m+n}]$

Standard form

$$\left[\begin{array}{cc|c|c} A_B & A_N & 0 & b \\ \hline c_B & c_N & 1 & 0 \end{array} \right]$$

$$Ax = A_N x_N + A_B x_B = b$$

$$A_B x_B = b - A_N x_N$$

$$x_B = A_B^{-1} b - A_B^{-1} A_N x_N$$

basic feasible solution:

- $x_N = 0$
- A_B lin. indep.
- $x_B \geq 0$

$$\begin{aligned} z = c^T x &= c_B^T (A_B^{-1} b - A_B^{-1} A_N x_N) + c_N^T x_N = \\ &= c_B^T A_B^{-1} b + (c_N^T - c_B^T A_B^{-1} A_N) x_N \end{aligned}$$

Canonical form

$$\left[\begin{array}{c|cc|c|c} I & A_B^{-1} A_N & 0 & A_B^{-1} b \\ \hline 0 & c_N^T - c_B^T A_B^{-1} A_N & 1 & -c_B^T A_B^{-1} b \end{array} \right]$$

In scalar form: the objective function is obtained by multiplying and subtracting constraints by means of multipliers π : $\pi = c_B^T A_B^{-1}$ (the dual variables)

$$z = \sum_{j=1}^m \left[c_j - \sum_{i=1}^m \pi_i a_{ij} \right] x_j + \sum_{j=m+1}^{m+n} \left[c_j - \sum_{i=1}^m \pi_i a_{ij} \right] x_j + \sum_{i=1}^m \pi_i b_i$$

Each basic variable has cost null in the objective function

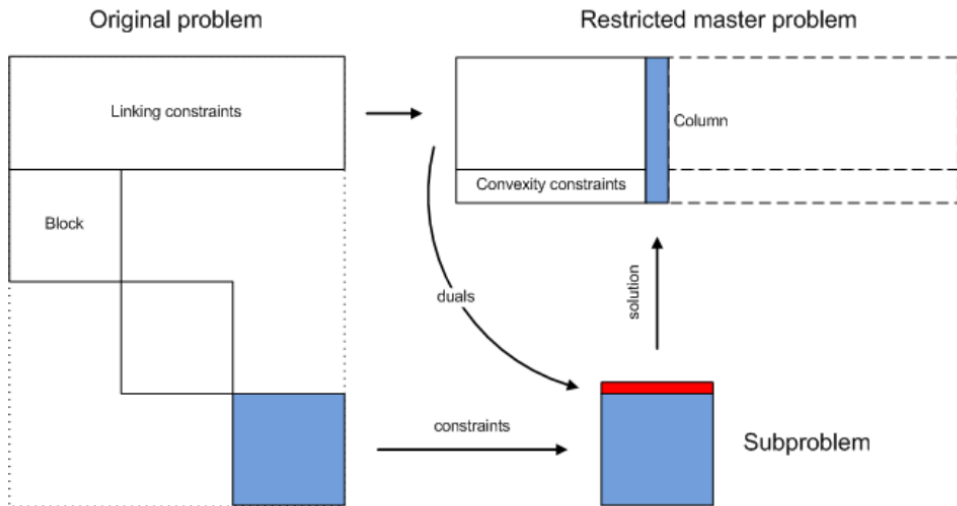
$$c_j - \sum_{i=1}^m \pi_i a_{ij} = 0 \quad j = 1, \dots, m$$

Reduced costs of non-basic variables:



$$\bar{c}_j = c_j - \sum_{i=1}^m \pi_i a_{ij} \quad j = m+1, \dots, m+n$$

If basis is optimal then $\bar{c}_j \leq 0$ for all $j = m+1, \dots, m+n$.

Dantzig Wolfe Decomposition with Column Generation



Delayed column generation (example)

- $w_1 = 5, b_1 = 7$ 
- $w_2 = 3, b_2 = 3$ 
- Row length $L = 22$

Initially we choose only the trivial cutting patterns

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 7 \end{pmatrix}$$

Solve LP-problem

$$\begin{array}{ll} \min & cx \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}$$

i.e.

$$\begin{pmatrix} 4 & 0 \\ 0 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$$

with solution $x_1 = \frac{7}{4}$ and $x_2 = \frac{3}{7}$.

The dual variables are $y = c_B A_B^{-1}$ i.e.

$$(1 \ 1) \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{7} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{7} \end{pmatrix}$$

Small example (continued)

Find entering variable

$$A = \begin{pmatrix} 1 & 2 & 3 & \cdots \\ 5 & 4 & 2 & \cdots \end{pmatrix} \quad \begin{array}{l} \frac{1}{4} \leftarrow y_1 \\ \frac{1}{7} \leftarrow y_2 \end{array}$$
$$c_N - yA_N = \left(1 - \frac{27}{28} \quad 1 - \frac{30}{28} \quad 1 - \frac{29}{28} \quad \cdots \right)$$

We could also solve optimization problem

$$\begin{array}{ll} \min & 1 - \frac{1}{4}x_1 - \frac{1}{7}x_2 \\ \text{s.t.} & 5x_1 + 3x_2 \leq 22 \\ & x \geq 0, \text{ integer} \end{array}$$

which is equivalent to knapsack problem

$$\begin{array}{ll} \max & \frac{1}{4}x_1 + \frac{1}{7}x_2 \\ \text{s.t.} & 5x_1 + 3x_2 \leq 22 \\ & x \geq 0, \text{ integer} \end{array}$$

This problem has optimal solution $x_1 = 2, x_2 = 4$.

Reduced cost of entering variable

$$1 - 2\frac{1}{4} - 4\frac{1}{7} = 1 - \frac{30}{28} = -\frac{1}{14} < 0$$

Small example (continued)

Add new cutting pattern to A getting

$$A = \begin{pmatrix} 4 & 0 & 3 \\ 0 & 7 & 2 \end{pmatrix}$$

Solve problem to LP-optimality, getting primal solution

$$x_1 = \frac{5}{8}, x_3 = \frac{3}{2}$$

and dual variables

$$y_1 = \frac{1}{4}, y_2 = \frac{1}{8}$$

Note, we do not need to care about “leaving variable”

To find entering variable, solve

$$\begin{aligned} \max \quad & \frac{1}{4}x_1 + \frac{1}{8}x_2 \\ \text{s.t.} \quad & 5x_1 + 3x_2 \leq 22 \\ & x \geq 0, \text{integer} \end{aligned}$$

This problem has optimal solution $x_1 = 4, x_2 = 0$.

Reduced cost of entering variable

$$1 - 4\frac{1}{4} - 0\frac{1}{8} = 0$$

Terminate with $x_1 = \frac{5}{8}, x_3 = \frac{3}{2}$, and $z_{LP} = \frac{17}{8} = 2.125$.

Questions

- Will the process terminate?

Always improving objective value. Only a finite number of basis solutions.

- Can we repeat the same pattern?

No, since the objective function is improved. We know the best solution among existing columns. If we generate an already existing column, then we will not improve the objective.

Tailing off effect

Column generation may converge slowly in the end

- We do not need exact solution, just lower bound
- Solving master problem for subset of columns does not give valid lower bound (why?)
- Instead we may use Lagrangian relaxation of joint constraint
- “guess” lagrangian multipliers equal to dual variables from master problem

Dual Bounds

Linear relaxation of the reduced master problem:

$$z_{LRMP} = \max \{ c\lambda \mid \bar{A}\lambda \leq b, \lambda \geq 0 \}$$

Note: $z_{LRMP} \not\geq z_{LMP}$ (LMP Lin. relax. master problem)

However, during colum generation we have access to a dual bound so that we can terminate the process when a desired solution quality is reached.

When we know that

$$\sum_{j \in J} \lambda_j \leq \kappa \quad J \text{ is the unrestricted set of columns}$$

for an optimal solution of the master, we cannot improve z_{RMP} by more than κ times the largest reduced cost obtained by the Pricing Problem (PP):

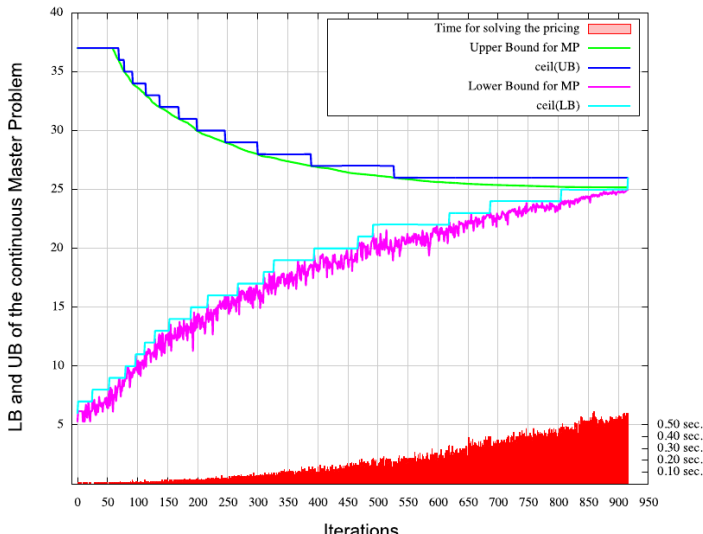
$$z_{LRMP} + \kappa z_{PP} \geq z_{LMP}$$

(It can be shown that this bound coincide with the Lagrangian dual bound.)

- with convexity constraints $\sum_{j \in J} \lambda_j \leq 1$ then $\kappa = 1$
- when $\mathbf{c} = \mathbf{1}$ we can set $\kappa = z_{LMP}$ and derive the better dual bound $\frac{z_{LRMP}}{1 - z_{PP}} \geq z_{LMP}$

Convergence in CG

In general the dual bound is not monotone during the iterations, for a problem of minimum:



Time in seconds for solving the pricing

Row and Column Generation

In problems with many rows we can generate them like done in column generation.

Cutting plane methods where the pricing problem is the separation problem.

Combining the two: column generation cannot ignore the missing rows. Existing approaches are problem specific.

Mixed Integer Linear Programs

- The primary use of column generation is in this context
(in LP simplex is better)
- column generation re-formulations often give much stronger bounds than the original LP relaxation
- Often column generation referred to as branch-and-price

- Master Problem
- Restricted Master Problem
- Subproblem or Pricing Problem
- Branch and cut:
Branch-and-bound algorithm using cuts to strengthen bounds.
- Branch and price:
Branch-and-bound algorithm using column generation to derive bounds.

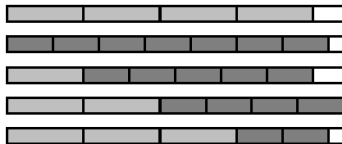
Branch-and-price

- LP-solution of master problem may have fractional solutions
- Branch-and-bound for getting IP-solution
- In each node solve LP-relaxation of master
- Subproblem may change when we add constraints to master problem
- Branching strategy should make subproblem easy to solve

Branch-and-price, example

The matrix A contains all different cutting patterns

$$A = \begin{pmatrix} 4 & 0 & 1 & 2 & 3 \\ 0 & 7 & 5 & 4 & 2 \end{pmatrix}$$



Problem

$$\text{minimize } \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$$

$$\text{subject to } 4\lambda_1 + 0\lambda_2 + 1\lambda_3 + 2\lambda_4 + 3\lambda_5 \geq 7$$

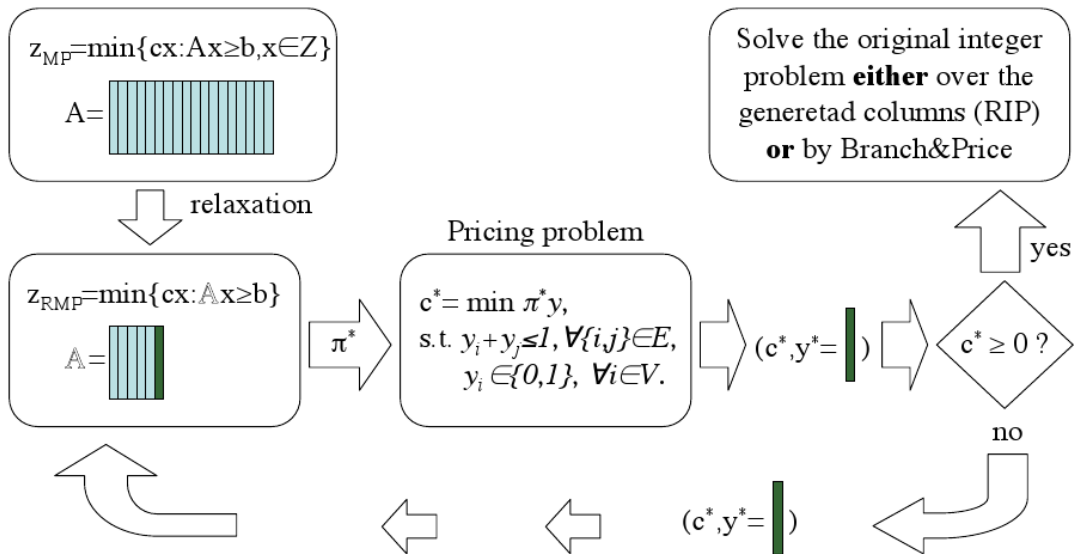
$$0\lambda_1 + 7\lambda_2 + 5\lambda_3 + 4\lambda_4 + 2\lambda_5 \geq 3$$

$$\lambda_j \in \mathbb{Z}_+$$

LP-solution $\lambda_1 = 1.375, \lambda_4 = 0.75$

Branch on $\lambda_1 = 0, \lambda_1 = 1, \lambda_1 = 2$

- Column generation may not generate pattern (4,0)
- Pricing problem is knapsack problem with pattern forbidden



Heuristic solution (eg, in sec. 12.6)

- Restricted master problem will only contain a subset of the columns
- We may solve restricted master problem to IP-optimality
- Restricted master is a “set-covering-like” problem which is not too difficult to solve

- Fisher M.L. (2004). **The Lagrangian relaxation method for solving integer programming problems.** *Management Science*, 50(12), pp. 1861–1871. This article originally appeared in *Management Science*, January 1981, Volume 27, Number 1, pp. 1-18, published by The Institute of Management Sciences.
- Lübbecke (2010). **Column Generation.** John Wiley & Sons, Inc.