620-362
Applied Operations Research

Column Generation

Department of Mathematics and Statistics
The University of Melbourne

Some contents of this presentation are adapted from year 2005 course notes for 620-362 Applied Operations Research, Department of Mathematics and Statistics, The University of Melbourne (compiled by Prof Natashia Boland and Dr Renata Sotirov)

Column Generation

Why is column generation used?

LPs with very large numbers of variables need to be solved.

Why formulate problems with large numbers of (integer) variables?

- The **bounds** provided by their LP-relaxations are much **better** than those of alternative compact formulations, .e.g the cutting stock gap is well-known to be ≤ 1 generally, so very little branching is required to get integer optimum.

- **Symmetry is eliminated or reduced** as compared to compact formulations, leading to much smaller B&B trees.

- **Complex constraints**, e.g. regulations specified in Enterprise Bargaining Agreement (EBA) can be relegated to a subproblem, and solved by specialised techniques.

What is column generation?

A method for solving LPs in which there are very large number of variables, without having to enumerate all the variables *a priori*.

Where has column generation been applied?

- Multi-commodity flow problems
- Cutting stock problems
- Binary cutting stock problems
- Air crew scheduling
- Crew rostering
- Aircraft fleeting and routing
- Vehicle routing
- Optical telecommunications design
- Ship scheduling and routing

A Quick revision on LP

1. LP: $z_{lp} = \min\{cx: Ax \geq b, x \geq 0, x \in R^n\}$

2. Very large LPS can be solved very efficiently with today's software, e.g. Cplex, Xpress^{MP}.
   – Usual method is still simplex method.

3. Dual LP: $w = \max\{ub: uA \leq c, u \geq 0, u \in R^m\}$

4. Simplex method maintains basis B for column space of A via subset of columns of A, indexed by $B \subseteq \{1,\ldots,n\}$.
   – It maintains feasible x and dual variable $u \geq 0$ with $ub = cx$ and strives to achieve $uA \leq c$.
   – Simplex iteration exchanges two columns in basis, i.e. $B := B \setminus \{i\} \cup \{j\}$, $i \in B$, $j \notin B$
   – To choose j: find most infeasible dual constraint, i.e. select
   $$j \in \arg\min (c_k - ua_k) < 0$$
   where $(c_k - ua_k)$ is the "reduced cost" of x_k.

Column Generation

What if A has a huge number of columns? How can we solve the LP?

Definition: We call the LP with a huge number of columns the Master LP.

Definition: If A^s is a matrix formed from a subset of the columns of A, and c^s is the corresponding subvector of c, the

$$LP^s: z^s_{LP} = \min\{c^s y: A^s y \geq b, y \geq 0\}$$

is called the Restricted LP.

Note: A^s must be chosen so that LP is feasible and bounded below.

Algorithm to solve the Master LP

Step 0:

Choose an initial set of columns A^s.

Step 1:

Solve the Restricted LP to obtain optimal basic solution *y* and corresponding duals *u*.

Step 2:

Note that *x* given by

$$x_i = y_i, \text{ if column } i \text{ is in } A^s; 0, \text{ otherwise}$$

is a basic feasible point for the Master LP and is optimal if

$$uA \leq c.$$

Solve column generation subproblem:

$$z^{sub} = \min\{c_i - ua_i : a_i \text{ is a column of } A\}$$

If $z^{sub} \geq 0$ then STOP: *x* is optimal for Master LP.
Otherwise add a_i such that $c_i - ua_i = z^{sub}$ to A^s and go to Step 1.

A Cutting Stock Problem

A cutting stock problem with 17m stock lengths and with orders for 25 pieces of length 3m, 20 pieces of length 5m and 15 pieces of 9m.

Let P be the set of all cutting patterns.

Let a_{ip} denote the number of pieces of length l_i cut in pattern p, where $l_1 = 3$, $l_2 = 5$ and $l_3 = 9$.

Let the demand for each length l_i be b_i, where $b_1 = 25$, $b_2 = 20$, $b_3 = 15$.

A Cutting Stock Problem

Pattern *p*	# of 3m (a_{1p})	# of 5m (a_{2p})	# of 9m (a_{3p})
Pattern 1	5	0	0
Pattern 2	4	1	0
Pattern 3	2	2	0
Pattern 4	2	0	1
Pattern 5	1	1	1
Pattern 6	0	3	0
Pattern 7	0	0	1

A Cutting Stock Problem

Variable:

x_p = number of times pattern p is cut

$$\min \quad \sum_{p \in P} x_p$$

$$s.t.$$

$$\sum_{p \in P} a_{ip} x_p \geq b_i, \quad i = 1,2,3$$

$$x \geq 0$$

Let A be the constraint matrix, i.e. $A = (a_{ip})$

A Cutting Stock Problem

Step 0:

Choose subset of columns corresponding to cutting patterns $(5\ 0\ 0)^T$, $(0\ 3\ 0)^T$, $(0\ 0\ 1)^T$, i.e.

$$A^s = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A Cutting Stock Problem

Step 1:

Solve

$$\text{Restricted LP}: \begin{cases} \min \quad y_1 + y_2 + y_3 \\ s.t. \\ 5y_1 \geq 25 \\ 3y_2 \geq 20 \\ y_3 \geq 15 \\ y \geq 0 \end{cases}$$

Optimal basic solution is

$y_1 = 5$, $y_2 = 6\ 2/3$, $y_3 = 15$

and optimal dual variables

$u = (1/5, 1/3, 1)$

A Cutting Stock Problem

Step 2:

Solve column generation subproblem:

$$\begin{cases} \min \quad 1-(1/5,1/3,1)a_p \\ s.t. \\ a_p \text{ is a column of the Master LP} \end{cases}$$

But what does it mean for a_p to be a column of the Master LP?

$a_p = (a_{1p}, a_{2p}, a_{3p})^T$ is integer, ≥ 0 and satisfies

$3a_{1p} + 5a_{2p} + 9a_{3p} \leq 17$

A Cutting Stock Problem

Therefore column generation subproblem is

$$z^{sub} = \begin{cases} \min \quad 1 - \left(\frac{1}{5} a_{1p} + \frac{1}{3} a_{2p} + a_{3p} \right) \\ s.t. \\ 3a_{1p} + 5a_{2p} + 9a_{3p} \leq 17 \\ a_{1p}, a_{2p}, a_{3p} \geq 0, \text{integer} \end{cases}$$

$$= 1 - \begin{cases} \max \quad \frac{1}{5} a_{1p} + \frac{1}{3} a_{2p} + a_{3p} \\ s.t. \\ 3a_{1p} + 5a_{2p} + 9a_{3p} \leq 17 \\ a_{1p}, a_{2p}, a_{3p} \geq 0, \text{integer} \end{cases}$$

Optimal solution:

$a_{1p} = a_{2p} = a_{3p} = 1$

z^{sub}

$= 1 – (1/5 + 1/3 + 1)$

$= -8/15 < 0$

A Cutting Stock Problem

Minimum reduced cost is negative, so column

$(1\ 1\ 1)^T$

should be added to A^s.

We return to Step 1 and solve the new restricted LP:

$$\text{Restricted LP}: \begin{cases} \min \quad y_1 + y_2 + y_3 + y_4 \\ s.t. \\ 5y_1 + y_4 \geq 25 \\ 3y_2 + y_4 \geq 20 \\ y_3 + y_4 \geq 15 \\ y \geq 0 \end{cases}$$

Optimal basic solution is

$y_1 = 2,\ y_2 = 1\ 2/3,\ y_3 = 0,\ y_4 = 15$

and optimal dual variables

$u = (1/5,\ 1/3,\ 7/15)$

A Cutting Stock Problem

Column generation subproblem is

$$z^{sub} = 1 - \begin{cases} \max \quad \frac{1}{5}a_{1p} + \frac{1}{3}a_{2p} + \frac{7}{15}a_{3p} \\ s.t. \\ 3a_{1p} + 5a_{2p} + 9a_{3p} \leq 17 \\ a_{1p}, a_{2p}, a_{3p} \geq 0, \text{integer} \end{cases}$$

Optimal solution to subproblem is

$a_1 = a_2 = a_3 = 1$

Therefore $z^{sub} = 1 - (1/5+1/3+7/15) = 0$

A Cutting Stock Problem

Thus there can be no negative reduced cost columns and optimal solution to Master LP is to use

$2 \times (5\ 0\ 0)^T$, $1\ 2/3 \times (0\ 3\ 0)^T$, $15 \times (1\ 1\ 1)^T$

Note:

This is not integer!!

But LP value = 2 + 1 2/3 + 15 = 18 2/3 is lower bound on problem.

Observation: Now

$2 \times (5\ 0\ 0)^T$, $2 \times (0\ 3\ 0)^T$, $15 \times (1\ 1\ 1)^T$

is an integer feasible solution with value 19.

Is this the optimal solution?

A Cutting Stock Problem

Sample Xpress^{MP} implementation

Human Resource Planning (HRP)

Tasks to be performed: 1, 2, …, n

h_i = time required for task i (hours)

H = max number of hours worker can work in a day

What is the minimum workforce required?

Master IP :

$$z = \min \sum_{s \in J'} x_s$$

$$s.t.$$

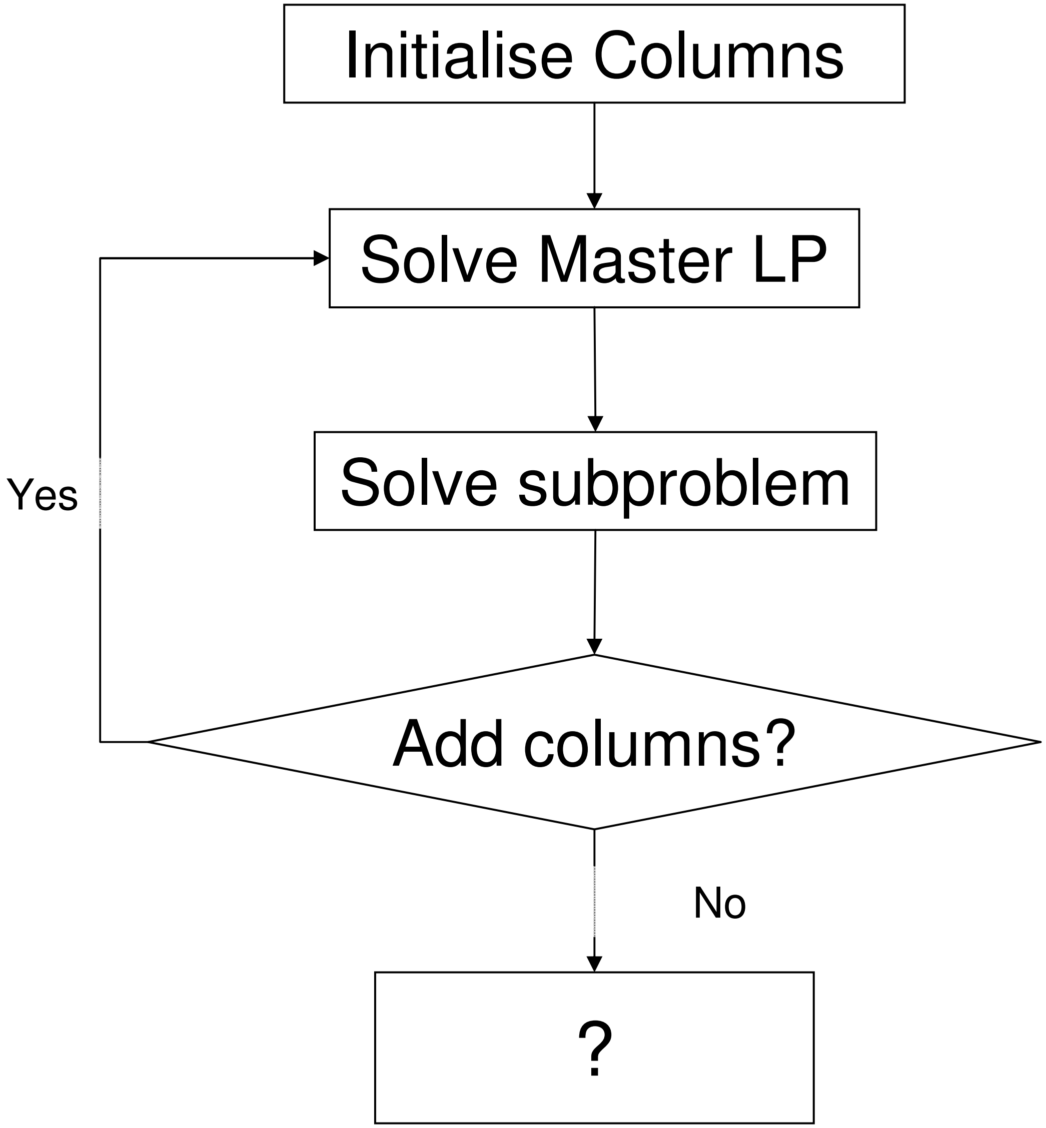
$$\sum_{s \in J' : i \in s} x_s = 1, \quad \forall i = 1, \ldots, n$$

$$x_s \in \{0,1\}, \quad \forall s \in J'$$

Subproblem :

$$z^{sub} = 1 - \begin{cases} \max \sum_{i=1}^{n} u_i y_i \\ s.t. \\ \sum_{i=1}^{n} h_i y_i \le H \\ y_i \in \{0,1\}, \quad \forall i = 1, \ldots, n \end{cases}$$

Human Resource Planning (HRP)



Human Resource Planning (HRP)

Sample Xpress^{MP} implementation

Human Resource Planning (HRP)

Scenario		HRP (Assignment)	HRP (Assignment - SymBreak)	HRP (Partition-Heuristic)	HRP (Partition-ColGen)
40 tasks	# constraints	64	87	40	40
	# variables	984	984	390	71
	LP value	22.25	22.25	23	23
	MIP value	23	23	23	23
	# nodes (# simplex iterations)	471 (204)	1,079 (443)	1 (76)	1 (46) 48 colgen iters
50 tasks	# constraints	84	117	50	50
	# variables	1,734	1,734	384	83
	LP value	30.425	30.425	33	33
	MIP value	33 (~7.8%, 60s)	33	33	33
	# nodes (# simplex iterations)	63,023 (242)	141 (530)	1 (70)	1 (60) 50 colgen iters

A Vehicle Routing Problem (VRP)

Let

N = fleet size

C = set of customers

R = set of routes

R_i = set of routes that serve customer i

c_r = cost of route $r = \sum_{l \in \{\text{links in route } r\}} c_l$

$$x_r = \begin{cases} 1, & \text{use route } r \\ 0, & \text{otherwise} \end{cases}$$

The Master IP:

$$\begin{cases} \min & \sum_{r \in R} c_r x_r \\ \text{s.t.} & \sum_{r \in R_i} x_r \geq 1, \forall i \in C \quad (\text{dual} : \pi_i) \\ & -\sum_{r \in R} x_r \geq -N \quad (\text{dual} : \mu) \\ & x_r \in \{0, 1\}, \forall r \in R \end{cases}$$

A Vehicle Routing Problem (VRP)

The subproblem is to find a feasible vehicle route that minimizes reduced cost. For the subproblem, let

$$y_{ij} = \begin{cases} 1, & \text{if route goes } i \rightarrow j \\ 0, & \text{otherwise} \end{cases}$$

The reduced cost for some route r^* is given by

$$\sum_{i,j} (c_{ij} - \pi_i) y_{ij} + \mu$$

A Vehicle Routing Problem (VRP)

The resulting sub-problem is a shortest path problem with negative cost cycle (see Figure 10). There may be further complicating constraints such as vehicle capacities and time windows.

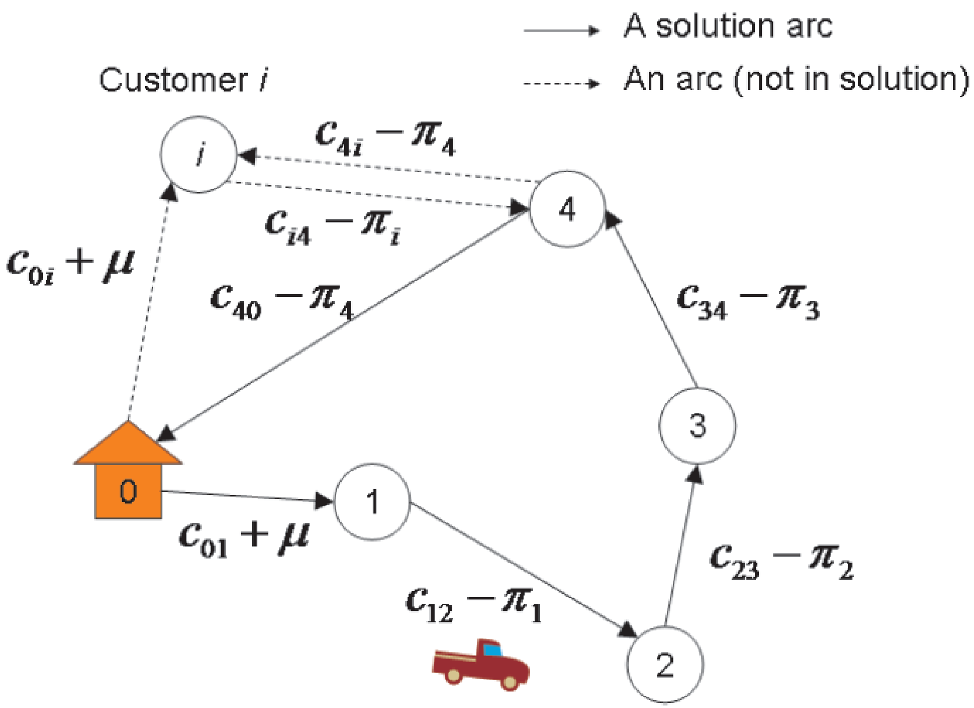


Figure 10: The subproblem for VRP.

An Air Crew Scheduling Problem

The objective here is to allocate air crews to flights such that total cost is minimised. A pairing is a sequence of flights flown by a crew. Consider the problem shown in Figure 11. Flights between Sydney, Melbourne and Adelaide are considered. Sydney and Melbourne are crew bases. There is a minimum half-an-hour turnaround time and there should be no more than 5 hours between consecutive flights. Pairings must start and end within one day, at the same crew base.

Let

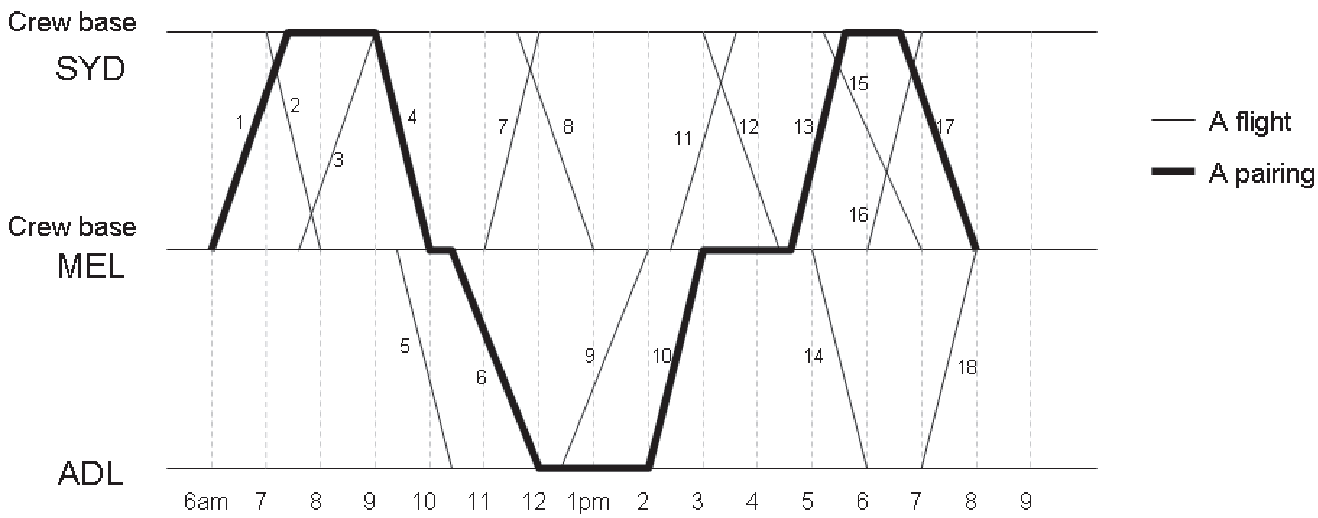
F = set of flights

P = set of crew pairings

P_f = set of pairings that use flights f

$$x_j = \begin{cases} 1, & \text{use pairing } j \text{ in solution} \\ 0, & \text{otherwise} \end{cases}$$

c_j = cost of pairing j

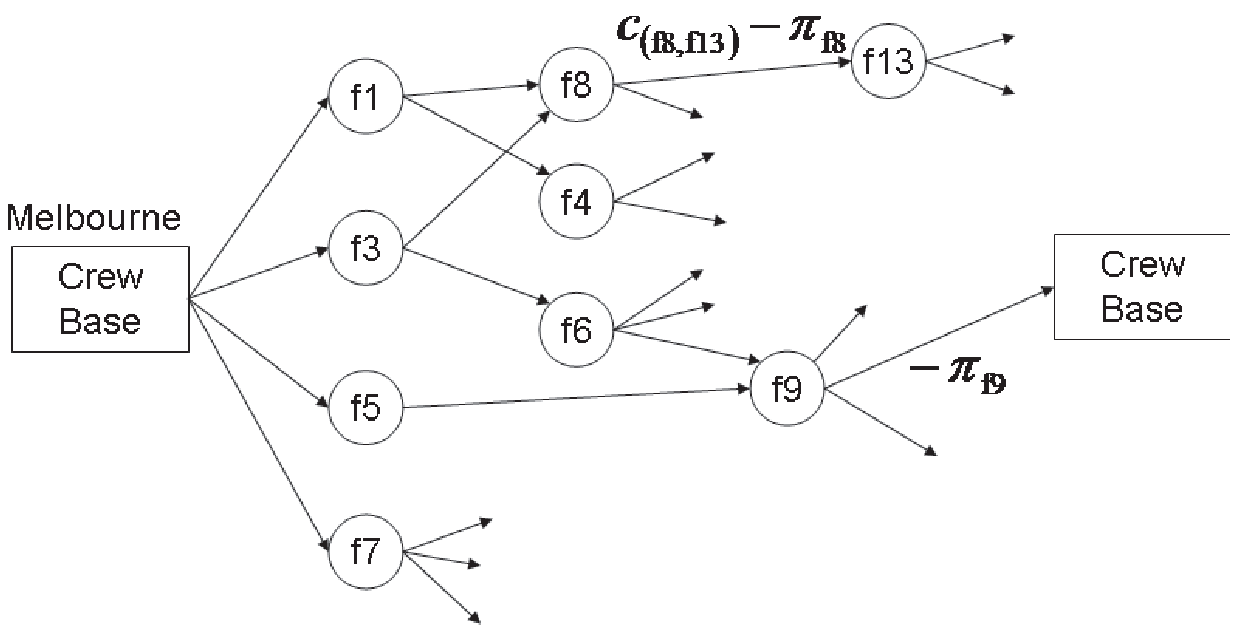


An Air Crew Scheduling Problem

The Master IP:

$$\begin{cases} \min & \sum_{j\in P} c_j x_j \\ \text{s.t.} & \sum_{j\in P_f} x_j = 1, \forall f \in F \quad (\pi_f) \\ & x_j \text{ binary } \forall j \in P \end{cases}$$

Subproblem: flight network (F^b, C^b) for all crew bases b (see Figure 12).
Minimum reduced cost pairing $\equiv$ shortest "valid" path.



Further reading

Winston Chapter 10.3

Next lecture…

Column Generation

&

Branch-and-Bound

Dantzig-Wolfe Decomposition