A column generation algorithm for the bi-objective max-min knapsack problem

Cláudio Alves, Raid Mansi, Telmo Pinto, J. M. Valério de Carvalho

Algoritmi Research Center, University of Minho

{claudio,raid.mansi,telmo,vc}@dps.uminho.pt

International Workshop on Column Generation 2012
June 10 to 13, 2012
Bromont, Québec, Canada

Outline

Outline

Standard Knapsack

- set of I items
- each item with an associated weight w_i and a profit c_i
- knapsack with an integer capacity W

Objective: determine the subset of items that maximize the total profit without exceeding W

$$\max \sum_{i=1}^{I} c_i x_i$$

$$\text{s. t.} \sum_{i=1}^{I} w_i x_i \leq W,$$

$$x_i \in \{0,1\}, \quad i = 1, \ldots, I.$$

Max-min Knapsack Problem (MNK)

- The profit of items depends on the scenario
- The weight of items remains independent
- c_i^s - profit of an item i under the scenario s
- S - total number of scenarios

Objective: maximize the worst possible scenario

$$\max \quad \min_{s=1,\ldots,S} \left\{ \sum_{i=1}^{I} c_i^s x_i \right\} \tag{1}$$

s.t.

$$\sum_{i=1}^{I} w_i x_i \leq W, \tag{2}$$

$$x_i \in \{0,1\}, \quad i = 1, \ldots, I. \tag{3}$$

Literature

- Yu (1996): MNK is strongly NP-Hard for unbounded S, but weakly NP-hard for fixed S.
- Sbihi (2010): cooperative local search-based algorithm for multi-scenario MNK
- Taniguchi et al. (2008): heuristic and exact algorithms for multi-scenario MNK
- Taniguchi et al. (2009): instances with strongly correlated profits for 2-scenario MNK are difficult to solve, unless they are very small. NUOPT (solver competitive with CPLEX) times increase significantly with size of instance.

Bi-objective max-min knapsack problem

- Bi-objective: number of scenarios is equal to two

A standard LP formulation:

$$\max \quad z \tag{4}$$

s. t.

$$\sum_{i=1}^{I} c_i^1 x_i \geq z, \tag{5}$$

$$\sum_{i=1}^{I} c_i^2 x_i \geq z, \tag{6}$$

$$\sum_{i=1}^{I} w_i x_i \leq W, \tag{7}$$

$$x_i \in \{0,1\}, \quad i = 1, \ldots, I. \tag{8}$$

Outline

Decomposition of MNK

Original model

$$\max \quad z \tag{9}$$

s. t.

$$\sum_{i=1}^{I} c_i^1 x_i \geq z, \tag{10}$$

$$\sum_{i=1}^{I} c_i^2 x_i \geq z, \tag{11}$$

$$\sum_{i=1}^{I} w_i x_i \leq W, \tag{12}$$

$$x_i \in \{0,1\}, \quad i = 1, \ldots, I. \tag{13}$$

- Reformulation of (9)-(13) through the Dantzig-Wolfe decomposition principle.
- Master problem is defined from the constraints (10)-(11)
- The subproblem is defined from its knapsack constraint (12).

Decomposition of MNK (2)

- X - polyhedron defined by the knapsack constraint of the MNK
- $x = (x_1, x_2, \ldots, x_I) \in X$ can be defined as a convex combination of the extreme points of X (Minkowski theorem for bounded polyhedra)
- P - number of extreme points of X
- $E_p = (E_p^1, E_p^2, E_p^3, \ldots, E_p^I)$ - p^{th} extreme point, $p = 1, \ldots, P$

$$x = \sum_{p=1}^{P} \alpha_p E_p = (\sum_{p=1}^{P} \alpha_p E_p^1, \sum_{p=1}^{P} \alpha_p E_p^2, \ldots, \sum_{p=1}^{P} \alpha_p E_p^I), \tag{14}$$

$$\sum_{p=1}^{P} \alpha_p = 1, \tag{15}$$

$$\alpha_p \geq 0, p = 1, \ldots, P. \tag{16}$$

Master problem

- $\alpha_p,\ p = 1, \ldots, P$ - variables of the master

Master problem

$$\max \quad z \tag{17}$$

$$s.t.$$

$$\sum_{p=1}^{P} \alpha_p \left(\sum_{i=1}^{I} c_i^1 E_p^i\right) \geq z, \tag{18}$$

$$\sum_{p=1}^{P} \alpha_p \left(\sum_{i=1}^{I} c_i^2 E_p^i\right) \geq z, \tag{19}$$

$$\sum_{p=1}^{P} \alpha_p = 1, \tag{20}$$

$$\alpha_p \geq 0, \ \ p = 1, \ldots, P. \tag{21}$$

Subproblem

- Standard knapsack does not have the integrality property
- Bound provided by linear relaxation of the reformulated model will be at least as strong as the linear relaxation of original model

Dual variables of Master Problem

$$\max \quad z \tag{22}$$

s.t. — *Dual variables*

$$\sum_{p=1}^{P} \alpha_p \left(\sum_{i=1}^{I} c_i^1 E_p^i \right) \geq z, \qquad \pi_1 \tag{23}$$

$$\sum_{p=1}^{P} \alpha_p \left(\sum_{i=1}^{I} c_i^2 E_p^i \right) \geq z, \qquad \pi_2 \tag{24}$$

$$\sum_{p=1}^{P} \alpha_p = 1, \qquad \pi_0 \tag{25}$$

$$\alpha_p \geq 0, \ \ p = 1, \ldots, P. \tag{26}$$

Subproblem

Reduced cost of α_p

$$-\pi_1\left(\sum_{i=1}^{I} c_i^1 E_p^i\right) - \pi_2\left(\sum_{i=1}^{I} c_i^2 E_p^i\right) - \pi_0.$$

- attractive variable: reduced cost is positive

Subproblem

$$\begin{aligned} \min \quad & \sum_{i=1}^{I}(\pi_1 c_i^1 + \pi_2 c_i^2) y_i \\ s.t. \quad & \\ & \sum_{i=1}^{I} w_i y_i \leq W, \\ & y_i \in \{0,1\},\ i = 1, \ldots, I. \end{aligned}$$

Subproblem

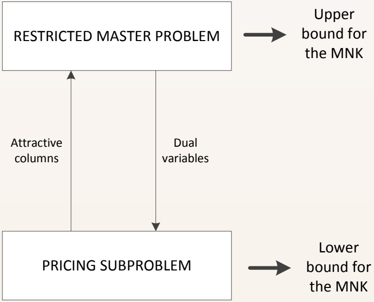
Each solution of subproblem is feasible for MNK

- its value can be checked against the upper bound to prove the optimality or to evaluate the tightness
- solution comes a single column (one convexity constraint)

Column generation

Bounds

- If the lower bound is equal to the upper bound (or the largest integer value smaller than this upper bound), then an optimal solution for the MNK has been found.



Outline

Instances

- weight w_i of item i is uniformly distributed over the integer interval $[1, 1000]$.
- values of items are generated according to:
  - UNCOR (uncorrelated): $c_i^s (s = 1, 2)$ are distributed independently and uniformly over $[1, 1000]$
  - WEAK (correlated): $c_i^s (s = 1, 2)$ are distributed independently and uniformly over $[w_i, w_i + 200]$
  - STRONG (strongly correlated): $c_i^1 := w_i + 100$ and c_i^2 is distributed uniformly over $[w_i, w_i + 200]$
- knapsack capacity $W := 500 I \rho, (\rho = \{0.25, 0.50, 0.75\}; \rho = 0.5$ means that approximately half of the items fit in the knapsack)

Computational Experiments

- Comparison between the proposed approach, CPLEX and the heuristic described in Taniguchi *et al.* (2009)
- Experiments conducted on a PC with 2.4GHz and 4 GB of RAM
- Two types of instances: correlated instances (20 sets) and strongly correlated instance (20 sets)
- Generation of 5 instances for each value of $I = \{250, 500, 1000, 2500\}$ with different $\rho = \{0.25, 0.50, 0.75\}$.

Terminology

- t_{orig}: average computing time in seconds needed by CPLEX to find an optimal solution for the original model (1)-(3) (and to prove its optimality);
- *it*: number of iterations of column generation algorithm;
- δ_{orig_cg}: average percentage difference between the upper bound given by the linear relaxation of (1)-(3) and the upper bound given by the linear relaxation of (22)-(26);
- δ_{opt_cg}: average percentage difference between the optimal solution of the problem and the lower bound given by our column generation algorithm;
- t_{cg}: average computing time needed by our column generation algorithm to solve the linear relaxation of (22)-(26) (in seconds);
- *opt*: number of times the column generation algorithm found an integer optimal solution to RMP.

Computational results: correlated instances (WEAK)

	Model (1)-(3)	Column Generation				
I	t_{orig}	it	δ_{orig_cg}	δ_{opt_cg}	t_{CG}	opt
250	0,0	2,4	0,3	0,4	0,0	4
500	0,0	7,6	0,2	0,0	0,2	2
1000	0,0	6,8	0,0	0,0	0,2	3
2500	0,0	6,6	0,0	0,0	0,6	4

Correlated instances

- Column generation upper bound improves slightly the LP upper bound for the first two sets of instances
- Column generation algorithm found and proved the optimality of the solutions for 13 instances at the root node.
- The lower bound obtained with the column generation algorithm is very close to the optimum
- both bounds from heuristic proposed by Taniguchi et al. (2009) are much worse than column generation bounds.
- CPLEX solves all the correlated instances very efficiently.

Computational results: strongly correlated instances (STRONG)

	Model (1)-(3)	Column Generation				
I	t_{orig}	it	δ_{orig_cg}	δ_{opt_cg}	t_{CG}	opt
250	61,6	4,4	2,1	0,0	0,2	2
500	261,8	5,4	2,0	0,0	0,2	2
1000	138,4	6,4	0,4	2,4	1,8	0
2500	182,2	7,8	0,1	4,8	5,2	1

Strongly Correlated Instances

- Computing time required by CPLEX to solve the original model increases significantly
- The average number of column generation iterations remains small
- both bounds from heuristic proposed by Taniguchi et al. (2009) are much worse than column generation bounds.

Outline

Conclusions

- Column generation approach for the bi-objective max-min knapsack problem
- Column generation algorithm clearly outperforms the state-of-the-art heuristic proposed by Taniguchi *et al.*, (2009), which takes approximately the same time as column generation.
- Both upper and lower bounds given by this heuristic are worse than those achieved with the original model

Future work

- Outline of a branch-and-price algorithm
  - fractional solution of RMP corresponds to fractional knapsack solution (convex combination of knapsack solutions)
  - branch on the original knapsack variables.
  - subproblem remains a knapsack problem.
- with narrow gaps, reduction of size of knapsack problems to core problem may allow solving exactly large instances of bi-objective MNK, in times competitive with CPLEX.
- heading to multi-scenario MNK.

Outline

Acknowledgements

- This work was supported by the Portuguese **Science and Technology Foundation** and by the **Algoritmi Research Center** of the University of Minho (Industrial and Systems Engineering Group).

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

ALGORITMI
UNIVERSIDADE DO MINHO