Semi-Infinite Relaxations for a Dynamic Knapsack Problem

Alejandro Toriello

joint with Daniel Blado, Weihong Hu

Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

CMO-BIRS Workshop
November 3, 2015
Problem Statement

- Knapsack with capacity $b > 0$ and item set $N = \{1, \ldots, n\}$. Each item $i$ has
  1. deterministic value $c_i$,
  2. independent random size $A_i \geq 0$ with known distribution.

- When attempting to insert $i$:
  - If $i$ fits collect $c_i$, update capacity.
  - Else process ends.

- Policy may depend on remaining items and remaining capacity.
  - Goal is to maximize expected value.

- Problem is at least NP-hard, some versions PSPACE-hard (Vondrák, 05).
Outline

Pertinent Past Work

Approximation and Bound

Computational Experiments

Extensions and Conclusions
Brief Literature Review

- **Derman/Lieberman/Ross (78):** Sizes are exponential r.v.’s.
  - Greedy policy w.r.t. $c_i/E[A_i]$ is optimal.

- **Dean/Goemans/Vondrák (04,08):** Two LP bounds with polynomially many variables.
  - Linear knapsack, polymatroid, both within constant gap.
  - Greedy approximate policies.

- **Gupta/Krishnaswamy/Molinaro/Ravi (11), Ma (14):** Integer sizes, LP bounds of pseudo-polynomial size.
  - Randomized policies based on LP optimal solutions.
  - Extensions to models with correlated random item values, preemption, multi-armed bandits.

- **Other work, e.g. Bhalgat/Goel/Kanna (11), Li/Yuan (13), Bansal/Nagarajan (14).**
Linear Knapsack Bound
Dean/Goemans/Vondrák (08)

- Use $x_i$, probability policy attempts to insert $i$:

$$\max_x \sum_{i \in N} c_i x_i$$

s.t. $$\sum_{i \in N} x_i \mathbb{E}[A_i] \leq b; \quad 0 \leq x_i \leq 1, \quad i \in N.$$
Use $x_i$, probability policy attempts to insert $i$:

$$\max_x \sum_{i \in N} c_i x_i \mathbb{P}(A_i \leq b)$$

s.t. $$\sum_{i \in N} x_i \mathbb{E}[A_i] \leq b; \quad 0 \leq x_i \leq 1, \quad i \in N.$$
Linear Knapsack Bound
Dean/Goemans/Vondrák (08)

- Use $x_i$, probability policy attempts to insert $i$:

$$
\max_x \sum_{i \in N} c_i x_i \mathbb{P}(A_i \leq b)
$$

s.t. $\sum_{i \in N} x_i \mathbb{E}[\min\{b, A_i\}] \leq b; \quad 0 \leq x_i \leq 1, \quad i \in N.$

- “Mean truncated size” $\mathbb{E}[\min\{b, A_i\}]$: $A_i$ above $b$ is irrelevant (insertion will fail).
Linear Knapsack Bound
Dean/Goemans/Vondrák (08)

- Use $x_i$, probability policy attempts to insert $i$:

$$\max_x \sum_{i \in N} c_i x_i \mathbb{P}(A_i \leq b)$$

subject to

$$\sum_{i \in N} x_i \mathbb{E}[\min\{b, A_i\}] \leq 2b; \quad 0 \leq x_i \leq 1, \quad i \in N.$$ 

- “Mean truncated size” $\mathbb{E}[\min\{b, A_i\}]$: $A_i$ above $b$ is irrelevant (insertion will fail).

- Bound intuition: In worst case, policy exactly fills knapsack, then attempts to insert very large item.
  - Worst-case gap is $32/7$.
  - Polymatroid bound is extension of same idea.
Dynamic Programming Formulation

State: remaining items, remaining capacity
\((M, s)\) for \(M \subseteq N, \ s \in [0, b]\).

Actions: attempt to insert \(i \in M\).

- Bellman recursion is
  \[
  v^*_M(s) = \max_{i \in M} \mathbb{P}(A_i \leq s)(c_i + \mathbb{E}[v^*_{M \setminus i}(s - A_i)|A_i \leq s]),
  \]
  \[
  v^*_\emptyset(s) = 0.
  \]

- In doubly infinite LP form:
  \[
  \min_v \ v_N(b)
  \]
  s.t. \(v_{M \cup i}(s) \geq \mathbb{P}(A_i \leq s)(c_i + \mathbb{E}[v_M(s - A_i)|A_i \leq s]), \ i \in N, \ M \subseteq N \setminus i, \ s \in [0, b]\)
  \[
  v_M : [0, b] \to \mathbb{R}_+, \ M \subseteq N.
  \]
Value Function Approximation

- Any feasible solution to LP yields upper bound.
- Use affine approximation

\[ v_M(s) \approx qs + r_0 + \sum_{i \in M} r_i, \]

where

- \( q \) is marginal value of capacity,
- \( r_i \) is item \( i \)'s "inherent" value,
- \( r_0 \) is value of process continuing ("staying alive").
Lemma

The best bound given by $v_M(s) \approx qs + \sum_{i \in M \cup 0} r_i$ is the semi-infinite LP

$$\min_{q,r \geq 0} q b + r_0 + \sum_{i \in N} r_i$$

s.t. $q \mathbb{E}[\min\{s, A_i\}] + r_0 \mathbb{P}(A_i > s) + r_i \geq c_i \mathbb{P}(A_i \leq s)$,

$i \in N$, $s \in [0, b]$. 

Value Function Approximation

Lemma

The best bound given by \( v_M(s) \approx qs + \sum_{i \in M \cup 0} r_i \) is the semi-infinite LP

\[
\min_{q, r \geq 0} \quad qb + r_0 + \sum_{i \in N} r_i
\]

s.t. \( q \mathbb{E}[\min\{s, A_i\}] + r_0 \mathbb{P}(A_i > s) + r_i \geq c_i \mathbb{P}(A_i \leq s), \quad i \in N, \ s \in [0, b]. \)

Proof sketch.

\[
v_{M \cup i}(s) - \mathbb{P}(A_i \leq s) \mathbb{E}[v_M(s - A_i)|A_i \leq s]
\approx qs - \mathbb{P}(A_i \leq s) \mathbb{E}[q(s - A_i)|A_i \leq s] \quad \text{(focusing on } q) \n= qs \mathbb{P}(A_i > s) + q \mathbb{P}(A_i \leq s) \mathbb{E}[A_i|A_i \leq s] = q \mathbb{E}[\min\{s, A_i\}]\]
Theorem

*The LP’s finite-support dual is solvable and has zero duality gap:*

\[
\max_{x \geq 0} \sum_{i \in N} \sum_{s \in [0,b]} c_i x_{i,s} \mathbb{P}(A_i \leq s)
\]

\[\text{s.t.} \sum_{i \in N} \sum_{s \in [0,b]} x_{i,s} \mathbb{E}[\min\{s, A_i\}] \leq b, \quad \text{(exp. frac. size under } b)\]

\[\sum_{i \in N} \sum_{s \in [0,b]} x_{i,s} \mathbb{P}(A_i > s) \leq 1 \quad \text{(one exp. failure; cf. Ma 14)}\]

\[\sum_{s \in [0,b]} x_{i,s} \leq 1 \quad \text{(insert } i \text{ once)}\]

\[x \text{ has finite support.}\]

- \(x_{i,s}\): probability policy attempts to insert \(i\) when \(s\) capacity remains.
Multiple-Choice Linear Knapsack Bound

Pricing problem

\[
\min_{q,r \geq 0} \left\{ qb + \sum_{i \in N \cup \{0\}} r_i : q \mathbb{E}[\min\{s, A_i\}] + r_0 \mathbb{P}(A_i > s) + r_i \geq c_i \mathbb{P}(A_i \leq s), \forall i \in N, s \in [0, b] \right\}
\]

- Pricing/separation: Given \( q, r \), for each \( i \) solve

\[
\min_{s \in [0,b]} \left\{ q \mathbb{E}[\min\{s, A_i\}] - (c_i + r_0) \mathbb{P}(A_i \leq s) \right\}.
\]

Mean truncated size is concave in \( s \). If CDF is piecewise convex, check only endpoints of convex intervals.

- Applies to discrete, uniform distributions
- Polynomially many variables.

- Other distributions (e.g. exponential, conditional normal) have closed-form solution.
- Check at most countably many points in general.
Multiple-Choice Linear Knapsack Bound
Pricing problem: Exponential distribution example

\[
\min_{s \in [0, b]} \left\{ q \mathbb{E}[\min\{s, A_i\}] - (c_i + r_0) \mathbb{P}(A_i \leq s) \right\}
\]

- Suppose \( A_i \sim \text{exp}(\lambda) \):

\[
\mathbb{P}(A_i \leq s) = 1 - e^{-\lambda s}
\]

\[
\mathbb{E}[\min\{s, A_i\}] = \mathbb{P}(A_i \leq s)/\lambda.
\]

Thus

\[
q \mathbb{E}[\min\{s, A_i\}] - (c_i + r_0) \mathbb{P}(A_i \leq s)
\]

\[
= (q/\lambda - c_i - r_0) \mathbb{P}(A_i \leq s)
\]

minimized at \( s \in \{0, b\} \).
Multiple-Choice Linear Knapsack Bound

- So if sizes are exponentially distributed, the bound is

\[
\max_x \sum_{i \in N} c_i x_{i,b} \mathbb{P}(A_i \leq b)
\]

s.t. \[
\sum_{i \in N} x_{i,b} \mathbb{E}[\min\{b, A_i\}] \leq b
\]

\[
0 \leq x_{i,b} \leq 1, \quad i \in N.
\]

This is DGV linear knapsack with capacity cut in half.

- Applies to other size distributions, e.g. conditional normal, uniform, geometric.

Theorem

*The MCLK bound dominates the DGV knapsack bound on any instance.*

- Conjecture: MCLK also dominates DGV polymatroid.
Computational Experiments

- Generated instances from deterministic knapsack instances.
  - 8 small, \( n \in [5, 24] \): people.sc.fsu.edu/~jburkardt
  - 10 large, \( n = 100 \): www.diku.dk/~pisinger/codes.html (uncorrelated)

- For a deterministic size \( a_i \), generated:
  - Exponential \( 1/a_i \)
  - Uniform \( [0, 2a_i] \) and \( [a_i/2, 3a_i/2] \)
  - Conditional normal \( (a_i, a_i/3) \)

- Bound comparison: average of deterministic knapsack over 400 simulations (“perfect information relaxation”).
  - Not reporting: DGV polymatroid bound not competitive.

- Benchmark: Adaptive greedy policy w.r.t. \( c_i \mathbb{P}(A_i \leq s) \) \( \mathbb{E}[\min\{s, A_i\}] \) (basic version studied in DGV).
Computational Experiments

Geometric gap mean

<table>
<thead>
<tr>
<th></th>
<th>Small</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PIR</td>
<td>MCLK</td>
</tr>
<tr>
<td>Exponential*</td>
<td>48%</td>
<td>5%</td>
</tr>
<tr>
<td>Uniform 1</td>
<td>41%</td>
<td>12%</td>
</tr>
<tr>
<td>Uniform 2</td>
<td>26%</td>
<td>12%</td>
</tr>
<tr>
<td>Normal</td>
<td>30%</td>
<td>12%</td>
</tr>
</tbody>
</table>

* Greedy benchmark is optimal (Derman/Lieberman/Ross 78).

- MCLK gives consistently better bound across instance types. Tighter for most small, all large instances.
- All gaps improve as number of items increases.
  - See an averaging effect as $n$ grows.
- Especially stark advantage for exponential instances.
Extensions

- Correlated value: Much of analysis applies, but must use conditional value $E[C_i | A_i \leq s]$ (GKMR 11, Ma 14).

- If items have integer support: Use non-parametric pseudo-polynomial approximation

  $$v_M(s) \approx \sum_{i \in M} r_i + \sum_{\sigma=0}^{s} w_\sigma.$$

  - Yields Ma bound (14).
  - Can use to show Ma bound dominates GKMR bound (strengthen Ma’s result).

- Policies: MCLK and pseudo-polynomial bounds can be used for policy design.
  - E.g. from value function approximation, “rounding”, ad hoc methods.
Conclusions

- MCLK bound has theoretical guarantees and good empirical performance on various item size distributions.
  - Gets better as number of items increases. Asymptotically optimal? (We have a rough proof.)

- Value function approximation is systematic way to generate bounds for dynamic problems.

- Big picture questions:
  1. Exact algorithms: cutting planes, branching?
  2. Extend to general “stochastic and dynamic” IP (Vondrák 05).

atoriello@isye.gatech.edu
www.isye.gatech.edu/~atoriello3