Introduction to Stochastic Combinatorial Optimization

Stefanie Kosuch
PostDok at TCSLab
www.kosuch.eu/stefanie/

Guest Lecture at the CUGS PhD course
"**Heuristic Algorithms for Combinatorial Optimization Problems**"

What is the interest of Stochastic Combinatorial Optimization?

Combinatorial "real world problems" often subject to uncertainties

- Not all parameters known when decision has to be made: market fluctuations, available capacity...
- Own decision depends on future decision of other parties: competition, clients, government...
- Setting of problem might change: weather, location...

Definition

Stochastic Combinatorial Optimization concerns the study and resolution of Combinatorial Optimization problems that involve uncertainties.

Objectives of this lecture

- Give you examples of SCO-problems.
- Give you an idea of how uncertainties can be modeled (most common models).
- Give you an idea of why Stochastic Optimization is hard.
- Give you an idea of how SCO-problems can be solved.
- Give you an idea of why metaheuristics are important tools in SCO.

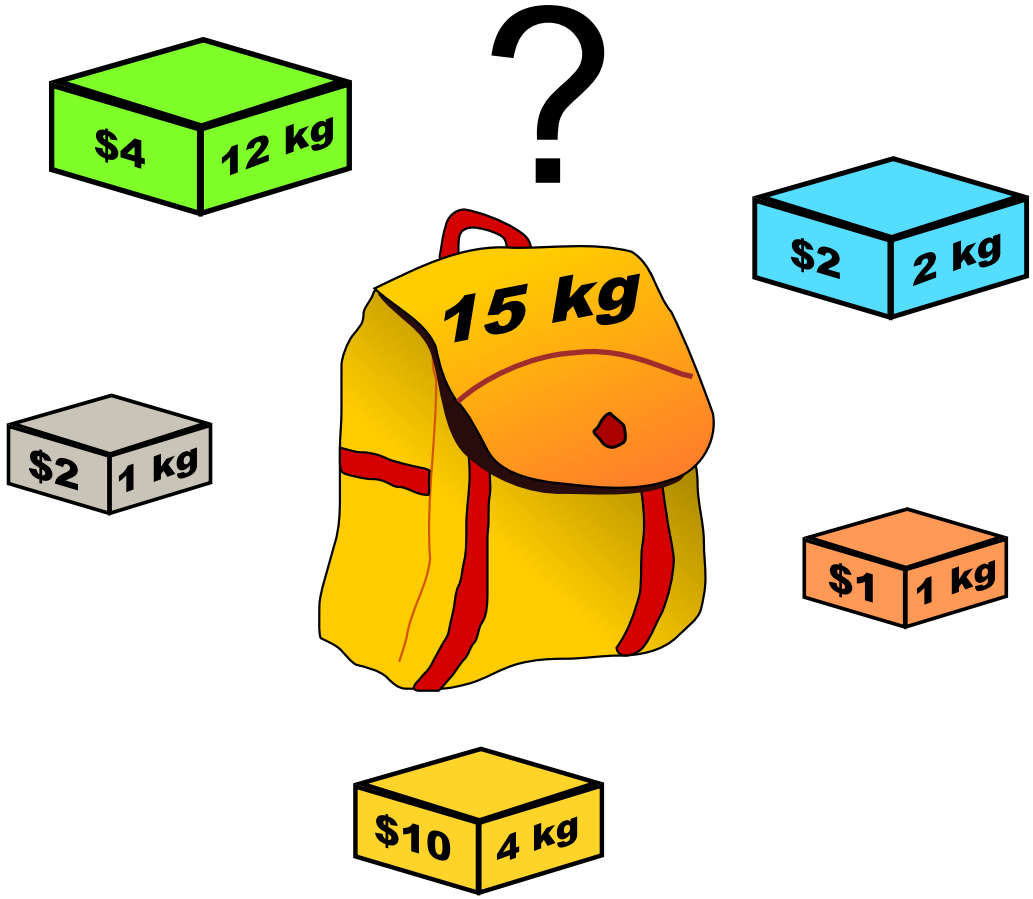
Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

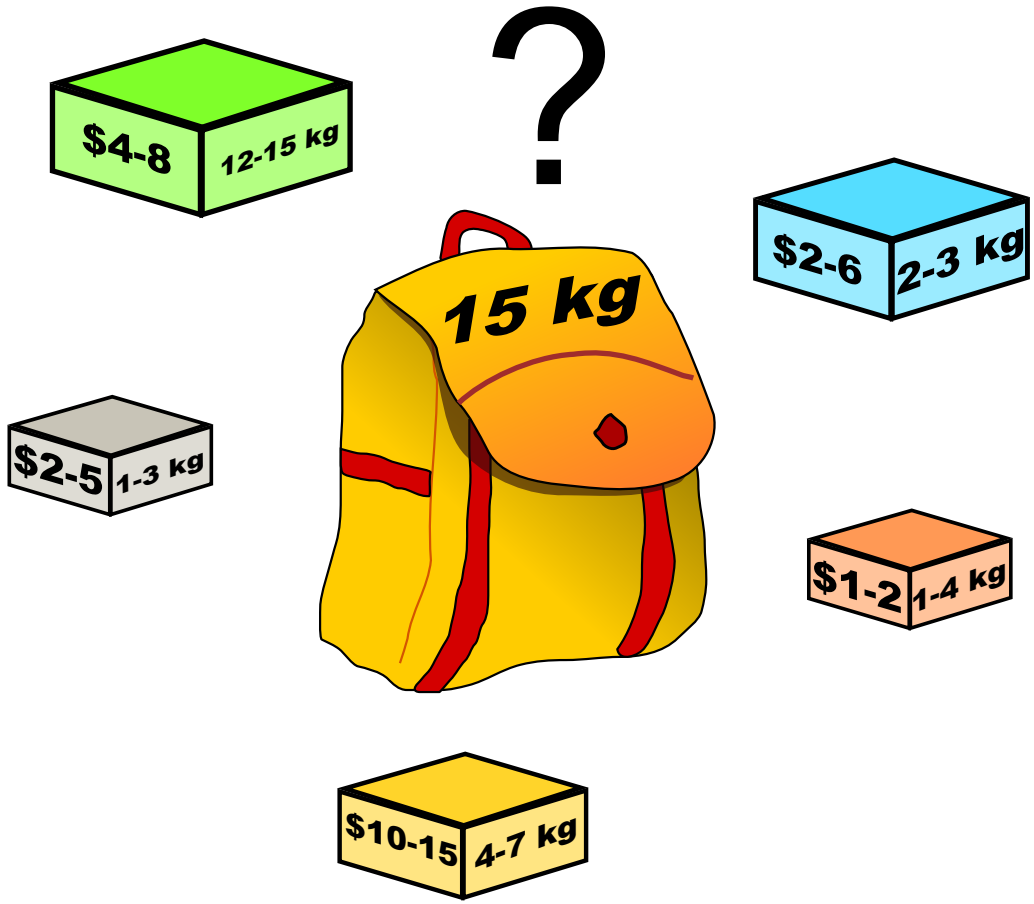
3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Deterministic Knapsack problem



Stochastic Knapsack problem



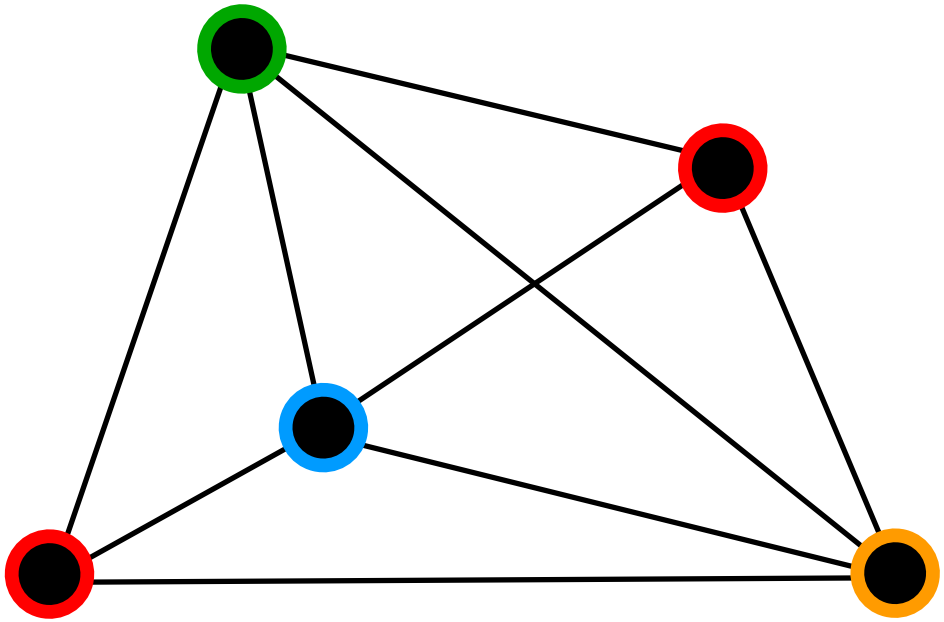
Possible ways to handle capacity constraint

- knapsack constraint violated $\Rightarrow$ penalty
- probability of capacity violation restricted
- decision can be corrected later (add. costs/reduced rewards)

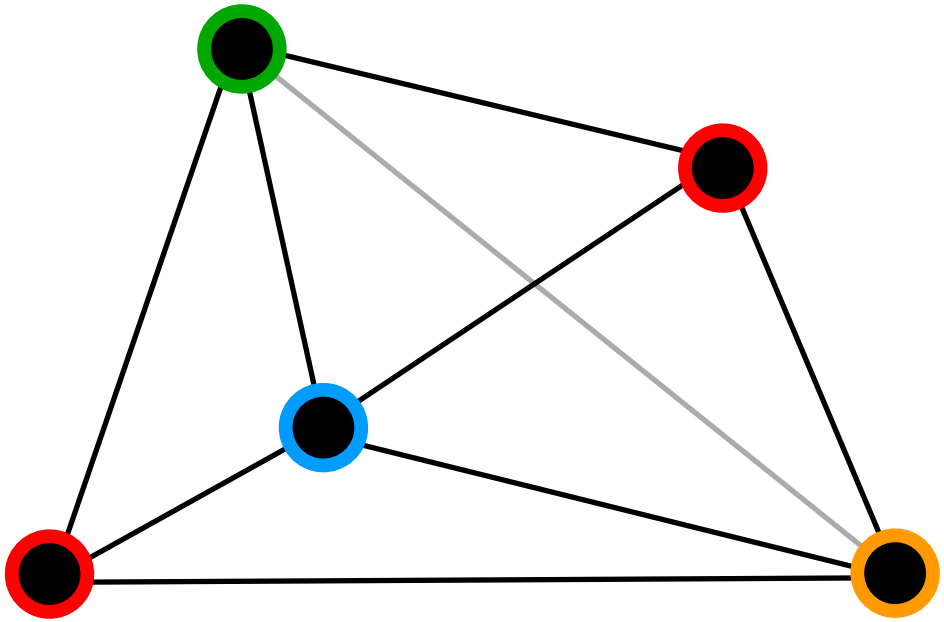
Possible ways to handle capacity constraint

- knapsack constraint violated $\Rightarrow$ penalty
- probability of capacity violation restricted
- decision can be corrected later (add. costs/reduced rewards)

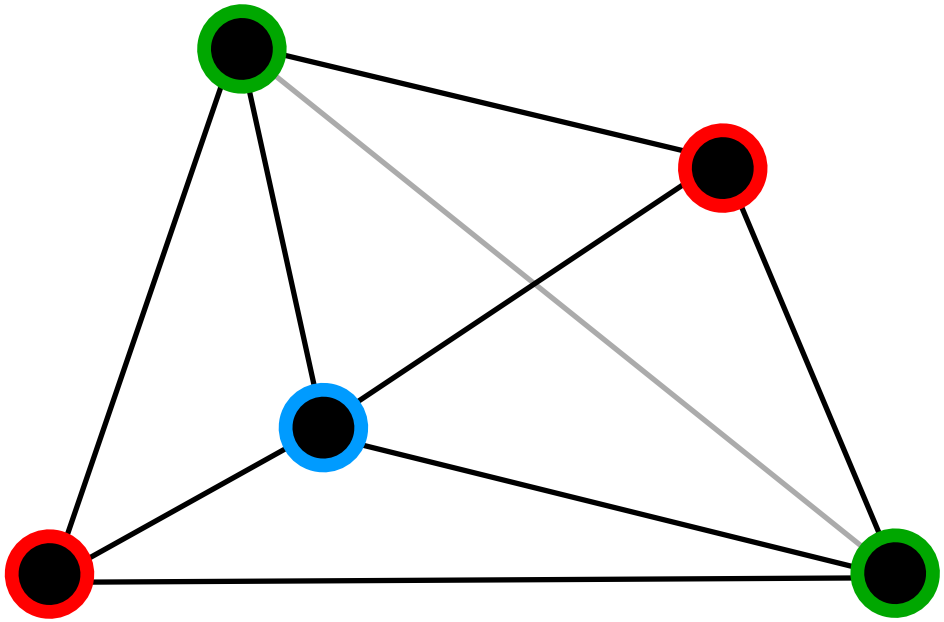
Deterministic Graph Coloring



Stochastic Graph Coloring



Stochastic Graph Coloring



Changing settings

- set of edges random
- set of vertices random

Changing parameters

- allowed number of colors random
- "cost" of colors random

Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Deterministic CO Model $\rightarrow$ Stochastic CO Model

$$\begin{array}{lll} \max\limits_{x\in\{0,1\}^n} & f(x) & \\ \text{s.t.} & g(x) \leq 0 & \end{array} \quad \rightarrow \quad \begin{array}{lll} \min\limits_{x\in\{0,1\}^n} & f(x,\chi) & \\ \text{s.t.} & g(x,\chi) \leq 0 & \end{array}$$

$\chi \in \Omega \subseteq \mathbb{R}^s$: vector with random entries

Outline

1 2 examples of SCO problems

2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Minimize an expectation

$$\min_{x \in \{0,1\}^n} \quad \mathbb{E}\left[f(x, \chi)\right]$$
$$\text{s.t.} \quad g(x) \leq 0$$

Advantages:

- Good result "on average"
- Objective function can often be reformulated deterministically

Disadvantages:

- We might encounter very "bad cases"

Minimize variance

$$\min_{x \in \{0,1\}^n} \quad Var\left[f(x, \chi)\right]$$
$$\text{s.t.} \quad g(x) \leq 0$$

Advantages:

- Outcome more concentrated around mean
- Possibility to reduce risk

Disadvantages:

- Makes not much sense without benchmark for expected costs

Minimize variance

$$\min_{x \in \{0,1\}^n} \quad \lambda Var\left[f(x,\chi)\right] + \mathbb{E}\left[f(x,\chi)\right]$$
$$\text{s.t.} \quad g(x) \leq 0$$

Advantages:

- Outcome more concentrated around mean
- Possibility to reduce risk

Minimize variance

$$\min_{x \in \{0,1\}^n} \quad Var\left[f(x,\chi)\right]^{\lambda} \cdot \mathbb{E}\left[f(x,\chi)\right]$$
$$\text{s.t.} \quad g(x) \leq 0$$

Advantages:

- Outcome more concentrated around mean
- Possibility to reduce risk

Robust optimization

$$\min_{x \in \{0,1\}^n} \max_{\chi \in \Omega} \quad f(x, \chi)$$
$$\text{s.t.} \quad g(x) \leq 0$$

Advantages:

- Worst case not too bad: Solution is *robust*

Disadvantages:

- $f(x, \cdot)$ needs to be bounded from above
- Worst case might be very improbable
- Average might be high

Outline

1 2 examples of SCO problems

2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Worst case model

$$\min_{x \in \{0,1\}^n} \quad f(x)$$
$$\text{s.t.} \quad g(x, \chi) \leq 0 \quad \forall \chi \in \Omega$$

Advantages:

- Absolutely robust solution

Disadvantages:

- Problem often infeasible or has only trivial solutions
- Solution at high costs
- Constraint forced to be satisfied in even very improbable cases

Chance-Constrained model

$$\min_{x \in \{0,1\}^n} \quad f(x)$$
$$\text{s.t.} \quad \mathbb{P}\{\exists i : g_i(x, \chi) > 0\} \leq \alpha$$

Advantages:

- Very improbable cases can be ignored
- Cost can be reduced

Disadvantages:

- No restriction of "magnitude" of allowed violation
- What happens if constraint is violated?

Simple-Recourse model

$$\min_{x \in \{0,1\}^{n_1}} \quad f(x) + \sum_{i=1}^{m} d_i \cdot \mathbb{E}\left[[g_i(x, \chi)]^+\right]$$

Advantages:

- Costs in case of violation taken into account
- "Magnitude" of violation can be controlled

Disadvantages:

- Probability of violation not restricted

Two-Stage model

$$\min_{x \in \{0,1\}^{n_1}} \quad f(x) + \mathbb{E}\left[Q(x,\chi)\right]$$

$$\text{s.t.} \quad Q(x,\chi) = \min_{y \in \{0,1\}^{n_2}} \quad \overline{f}(y)$$

$$\text{s.t.} \quad g(x,y,\chi) \leq 0$$

Advantages:

- Violation of constraint not permitted
- Corrections in case of violation taken into account

Disadvantages:

- Problem extremely hard to solve:
  - → Non-convex, non-continuous objective function
  - → No closed-form expression of objective function
  - → Second-stage problem $\mathcal{NP}$-hard

Deterministic Knapsack Problem

$$\max_{x \in \{0,1\}^n} \quad \sum_{i=1}^{n} r_i x_i$$

$$\text{s.t.} \quad \sum_{i=1}^{n} w_i x_i \leq c$$

Simple Recourse Knapsack Problem

$$\max_{x \in \{0,1\}^n} \sum_{i=1}^{n} r_i x_i - d \cdot \mathbb{E}\left[[\sum_{i=1}^{n} \chi_i x_i - c]^+ \right]$$

Two-Stage Knapsack Problem

$$(TSKP) \quad \max_{x\in\{0,1\}^n} \quad \sum_{i=1}^{n} r_i x_i + \mathbb{E}[\mathcal{Q}(x,\chi)]$$

$$\text{s.t.} \quad \mathcal{Q}(x,\chi) = \max_{y^+,y^-\in\{0,1\}^n} \sum_{i=1}^{n} \bar{r}_i y_i^+ - \sum_{i=1}^{n} d_i y_i^-,$$

$$\text{s.t.} \quad y_j^+ \le 1 - x_j, \quad j = 1,\dots,n,$$

$$y_j^- \le x_j, \quad j = 1,\dots,n,$$

$$\sum_{i=1}^{n} (x_i + y_i^+ - y_i^-)\chi_i \le c.$$

Outline

1 2 examples of SCO problems

2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Static Stochastic Optimization problems

- Random parameters revealed after decision has been made.
- For decision maker parameters are revealed "once for all".
- No corrective decision can be made.

- Minimize expectation and/or variance
- Robust/Worst Case Optimization
- Chance-Constrained Optimization
- Simple Recourse Model

Two-Stage Optimization problems

- For decision maker parameters are revealed "once for all".
- Random parameters revealed after first-stage decision has been made.
- Corrective decision can be made once the parameters are known.

- Two-Stage Model
- Simple Recourse Model

Multi-Stage Optimization problems

- Parameters are revealed in several stages.
- Corrective decision can be made in each stage.
- Which parameters are revealed in which stage generally defined.
- Decisions do only depend on already revealed parameters.

- Two-Stage Model
- Multi-Stage Model

Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Structural Difficulties

- Non-convexity
- Non-continuous objective functions
- No closed-form expression for objective function

Two-Stage model

$$\begin{aligned} \min_{x \in \{0,1\}^{n_1}} \quad & f(x) + \mathbb{E}\left[Q(x,\chi)\right] \\ \text{s.t.} \quad & Q(x,\chi) = \min_{y \in \{0,1\}^{n_2}} \quad \overline{f}(y) \\ & \qquad\qquad \text{s.t.} \quad g(x,y,\chi) \leq 0 \end{aligned}$$

Computational Difficulties

- Expectations and probabilities = (multi-dimensional) integrals!
- Evaluating objective function might be $\mathcal{NP}$-hard
- High number of binary decision variables and constraints

Two-Stage Knapsack Problem

$$(TSKP) \quad \max_{x\in\{0,1\}^n} \quad \sum_{i=1}^{n} r_i x_i + \mathbb{E}[\mathcal{Q}(x,\chi)]$$

$$\text{s.t.} \quad \mathcal{Q}(x,\chi) = \max_{y^+,y^-\in\{0,1\}^n} \sum_{i=1}^{n} \bar{r}_i y_i^+ - \sum_{i=1}^{n} d_i y_i^-,$$

$$\text{s.t.} \quad y_j^+ \le 1 - x_j, \quad j = 1,\dots,n,$$

$$y_j^- \le x_j, \quad j = 1,\dots,n,$$

$$\sum_{i=1}^{n} (x_i + y_i^+ - y_i^-)\chi_i \le c.$$

Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Idea

→ Reformulate problem as a deterministic optimization problem

→ Use already existing solvers to solve obtained problem

→ Adapt existing algorithms to the special structure of the obtained problem

Problem

Generally only possible under assumption of special distributions!

Ex. 1: Chance-constrained Knapsack prob. (normal distribution)

$$\max_{x \in \{0,1\}^n} \quad \sum_{i=1}^{n} r_i x_i$$

$$\text{s.t.} \quad \mathbb{P}\{\sum_{i=1}^{n} \chi_i x_i > c\} \leq \alpha$$

Assume:

$\chi \sim \mathcal{N}(\mu, \Sigma)$ and $\alpha < 0.5 \Rightarrow$

$$\mathbb{P}\{\sum_{i=1}^{n} \chi_i x_i > c\} \leq \alpha \Leftrightarrow \underbrace{\sum_{i=1}^{n} x_i \mu^i + \Phi^{-1}(1-\alpha)\|\Sigma^{1/2} x\| \leq c}_{\text{Second Order Cone Constraint!}}$$

Ex. 2: Chance-constrained Knapsack prob. (discrete distribution)

$$\max_{x\in\{0,1\}^n} \sum_{i=1}^{n} r_i x_i \qquad \text{s.t.} \qquad \mathbb{P}\{\sum_{i=1}^{n} \chi_i x_i > c\} \leq \alpha$$

Assume:

K outcomes $\chi^1, \ldots, \chi^K$ with prob.'s $p^1, \ldots, p^K$

Introduce:

K binary decision variables $z_1, \ldots, z_K$

Replace Chance-Constraint by

$$\sum_{i=1}^{n} \chi_i^k x_i \leq c + M z_k \quad \forall\, k = 1, \ldots, K,$$

$$\sum_{k=1}^{K} p_k z_k \leq \alpha$$

Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Idea of the SAA

$\rightarrow$ Sample K outcomes for random parameters

$\rightarrow$ Assign probability $1/K$ to each sample

$\rightarrow$ Replace distribution by finite sample

Advantages

- *Approximate* information about underlying distribution ✓
- *Approximate* closed-form expression for objective function ✓
- *Approximate* deterministic reformulation of problem ✓
- *Smaller* number of scenarios ✓

Idea of the SAA

→ Sample K outcomes for random parameters

→ Assign probability $1/K$ to each sample

→ Replace distribution by finite sample

Disadvantages

- Solution of SAA might be infeasible for original problem
- Solution of SAA might be non-optimal for original problem
- To approximate original problem high number of samples might be needed

Simple Recourse Knapsack Problem

$$\max_{x\in\{0,1\}^n} \sum_{i=1}^{n} r_i x_i - d \cdot \mathbb{E}\left[\left[\sum_{i=1}^{n} \chi_i x_i - c\right]^+\right]$$

Sample Average Approximation:

$$\max_{x\in\{0,1\}^n} \sum_{i=1}^{n} r_i x_i - d \cdot \sum_{k=1}^{K} \frac{1}{K}\left[\sum_{i=1}^{n} \chi_i^k x_i - c\right]^+$$

Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

"Positive" Features

- Basically same as for deterministic comb. opt.
- Based on sample average approximations
- Increase sample size to obtain convergence.
- Additional diversification due to randomness of samples

Difficulties

- Evaluation of objective function expensive
- Comparison of quality of solutions difficult
- SCO problems generally have a lot of local optima
- Values of local optima can be close

Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Summary

- Modeling SCO problems:
  - → Randomness in objective function?
  - → Randomness in constraint?
  - → Violation of constraint possible? Penalty?
  - → Correction of decision possible? How often?
- Solving SCO problems:
  - → Deterministic equivalent formulation
  - → Approximation of problem using sampling
  - → Meta-Heuristics

Greatest Challenges in Stochastic Combinatorial Optimization

- Modeling Combinatorial Real World problems with uncertainties as SCO problems
- Solve "realistic" sized problems in reasonable time
- Find more adapted solution techniques (for general distributions)
- Find efficient (Meta)Heuristics

Outline

1 2 examples of SCO problems
2 Modeling Stochastic Combinatorial Optimization Problems
- "Where" does the randomness occur?
  - Randomness occurs in the objective function f.
  - Randomness occurs in the constraint function g.
- When are the actual parameters revealed?
  - Parameters are revealed after decision has been made.
  - Parameters are revealed before corrective decision is made.
  - Parameters are revealed in several stages.

3 Solving Stochastic Combinatorial Optimization problems
- Problems/Difficulties
- Deterministic Reformulation
- Sample Average Approximation
- Metaheuristics for SCO problems

4 Conclusion
- Further Reading

Alexander Shapiro, Darinka Dentcheva, Andrzej Ruszczyński
Lectures on Stochastic Programming (2009)
pdf available online

András Prékopa
Stochastic Programming (1995).
Springer.

Anton J. Kleywegt, Alexander Shapiro, Tito Homem-de-Mello
The Sample Average Approximation Method for Stochastic Discrete Optimization (2002)
SIAM Journal on Optimization

Peter Kall, Stein W. Wallace
Stochastic Programming (1994)
pdf available online

Stochastic Programming Community
Stochastic Programming Community Home Page
`http://stoprog.org/`

Leonora Bianchi, Marco Dorigo, Luca Maria Gambardella, Walter J. Gutjahr
Metaheuristics in Stochastic Combinatorial Optimization: a Survey (2006)
pdf available online