Satisficing Models to Mitigate Uncertainty

Patrick Jaillet *

Department of Electrical Engineering and Computer Science
Laboratory for Information and Decision Systems
Operations Research Center
Massachusetts Institute of Technology

March 8, 2018

(*) with Sanjay Jena, Andy Ng, Melvyn Sim

outline

- optimization under uncertainty and satisficing - a quick review
- satisficing decision criteria - general representation theorem
- the t-model: a tractable probabilistic satisficing model
- numerical illustration - maximum coverage facility location problem

optimization under uncertainty

- deterministic optimization:
  $$\min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{c}'\boldsymbol{x} \;\; \text{s.t. } \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$$
- optimization under uncertainty $\tilde{\boldsymbol{z}} \in \mathcal{W}$:
  $$\min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{c}'\boldsymbol{x} \;\; \text{s.t. } \boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})$$
  - probability distributions available: stochastic optimization (Prékopa 1995; Birge and Louveaux 1997; etc. ....)
  - distributions unavailable:
    - robust optimization (Ben-Tal and Nemirovski 1999; Bertsimas and Sim 2004, Bertsimas et al. 2011)
    - distributionally robust optimization (Delage and Ye 2010; Wiesemann et al. 2014; etc ... this workshop)

satisficing

satisficing = satisfy + suffice

Simon (1959):

«... the entrepreneur may not care to maximize, but may simply want to earn a return that he regards as satisfactory [...] "satisfactory profits" is a concept more meaningfully related to the psychological notion of aspiration levels than to maximization...»

Simon, H. A. (1959). Theories of Decision-Making in Economics and Behavioral Science. The American Economic Review 49(3):253–83.

a first satisficing model under uncertainty: the p-model

Charnes and Cooper (1963)

- first to incorporate the idea of satisficing in mathematical programming under uncertainty:

$$\begin{array}{ll} \max & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X}. \end{array}$$

- randomly perturbed linear constraints $\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})$
  - tractable only for restricted special cases
  - general case intractable (Nemirovski and Shapiro 2006)

Charnes, A., and W. Cooper (1963) Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints. Operations Research 11(1):18–39.

chance-constrained optimization/programming

Charnes and Cooper (1959), close relation to the p-model

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \geq \Delta \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- satisficing criterion subject to a lower bound parameter $\Delta \in \mathbb{R}$
- objective is a deterministic cost function; $\boldsymbol{c} \in \mathbb{R}^N$ defines the objective function coefficients
- approximation by sample average approximation (SAA) methods
  - disadvantage: require large number of samples

Charnes, A., and W. Cooper (1959) Chance-Constrained Programming. Management Science 6(1):73–79.

robust optimization

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \;\; \forall \boldsymbol{z} \in \mathcal{U}(\Gamma) \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- $\boldsymbol{z}$ denotes realization of $\tilde{\boldsymbol{z}}$ from an *uncertainty set*, $\mathcal{U}(\Gamma)$ (typically $\mathcal{U}(\cdot)$ designed such that $\mathcal{U}(\alpha_1) \subseteq \mathcal{U}(\alpha_2) \subseteq \mathcal{W} \subseteq \mathbb{R}^K$ for all $0 \leq \alpha_1 \leq \alpha_2$).
- does not require the specification of a probability distribution, but instead a "budget of uncertainty" $\Gamma \in \mathbb{R}_+$
  - the level of uncertainty that must be tolerated
  - may not be easy to specify
- yields tractable formulations under reasonable conditions: e.g., if $\mathcal{U}(\Gamma)$ is described as norm-based sets $\mathcal{U}(\Gamma) = \{\boldsymbol{z} \in \mathcal{W} \mid ||\boldsymbol{z}|| \leq \Gamma\}$:
  - linear program for $||\cdot||_1$, $||\cdot||_\infty$ and D-Norm (Bertsimas et al. 2004)
  - second-order cone program for $||\cdot||_2$ norm

a satisficing model for robust optimization

- the p-model is a satisficing model for chance-constrained optimization:

$$\begin{array}{ll} \max & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \end{array}$$

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \geq \Delta \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

a satisficing model for robust optimization

- the p-model is a satisficing model for chance-constrained optimization:

$$\begin{array}{ll} \max & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \end{array}$$

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \geq \Delta \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

- can we define a satisficing model for robust optimization?

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \quad \forall \boldsymbol{z} \in \mathcal{U}(\Gamma) \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

a satisficing model for robust optimization

- the p-model is a satisficing model for chance-constrained optimization:

$$\begin{array}{ll} \max & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \end{array} \qquad \begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \geq \Delta \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

- can we define a satisficing model for robust optimization?

$\rightarrow$ the r-model:

$$\begin{array}{ll} \max & \{\alpha \mid \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \ \ \forall \boldsymbol{z} \in \mathcal{U}(\alpha)\} \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \end{array} \qquad \begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \ \ \forall \boldsymbol{z} \in \mathcal{U}(\Gamma) \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

some benefits of satisficing models

- the p-model is a satisficing model for chance-constrained optimization:

$$\begin{array}{ll} \max & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \end{array} \qquad \begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \geq \Delta \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

- can we define a satisficing model for robust optimization?

$\rightarrow$ the r-model:

$$\begin{array}{ll} \max & \{\alpha \mid \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \ \ \forall \boldsymbol{z} \in \mathcal{U}(\alpha)\} \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \end{array} \qquad \begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \ \ \forall \boldsymbol{z} \in \mathcal{U}(\Gamma) \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

some benefits of satisficing models

- the p-model is a satisficing model for chance-constrained optimization:

$$\begin{array}{ll} \max & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \cup \{\boldsymbol{c}'\boldsymbol{x} \leq B\} \end{array}$$

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right) \geq \Delta \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

- can we define a satisficing model for robust optimization?

  $\rightarrow$ the r-model:

$$\begin{array}{ll} \max & \{\alpha \mid \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \ \ \forall \boldsymbol{z} \in \mathcal{U}(\alpha)\} \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \cup \{\boldsymbol{c}'\boldsymbol{x} \leq B\} \end{array}$$

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \ \ \forall \boldsymbol{z} \in \mathcal{U}(\Gamma) \\ & \boldsymbol{x} \in \mathcal{X} \end{array}$$

satisficing decision criterion - definition

setting:

- $\tilde{\boldsymbol{z}}$ a K dimensional random vector that influences the entries of the function maps $\boldsymbol{A} : \mathbb{R}^K \mapsto \mathbb{R}^{M \times N}$ and $\boldsymbol{b} : \mathbb{R}^K \mapsto \mathbb{R}^M$.
- randomly perturbed linear constraints, $\boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z})$, where $\boldsymbol{z}$ is a random outcome of $\tilde{\boldsymbol{z}}$.
- $\mathcal{W} \subseteq \mathbb{R}^K$ the support of the random vector $\tilde{\boldsymbol{z}}$.

definition: satisficing decision criterion

a function $\nu : \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\}$ is a *satisficing decision criterion* if it has the following two properties. For all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$,

1. (satisficing dominance) if $\boldsymbol{A}(\boldsymbol{z})\boldsymbol{y} \geq \boldsymbol{b}(\boldsymbol{z})$ implies $\boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z})$ for all $\boldsymbol{z} \in \mathcal{W}$, then $\nu(\boldsymbol{x}) \geq \nu(\boldsymbol{y})$.
2. (infeasibility) if there does not exist $\boldsymbol{z} \in \mathcal{W}$ such that $\boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z})$, then $\nu(\boldsymbol{x}) = -\infty$.

satisficing decision criteria - previous examples

- the p-model is an optimization problem that maximizes a satisficing decision criterion $\nu_P : \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\}$ given by

$$\nu_P(\boldsymbol{x}) = \ln \mathbb{P}\left(\boldsymbol{A}(\tilde{\boldsymbol{z}})\boldsymbol{x} \geq \boldsymbol{b}(\tilde{\boldsymbol{z}})\right)$$

- the r-model is an optimization problem that maximizes a satisficing decision criterion $\nu_R : \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\}$ given by

$$\nu_R(\boldsymbol{x}) = \max\left\{\alpha \mid \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \;\; \forall \boldsymbol{z} \in \mathcal{U}(\alpha)\right\}$$

satisficing decision criteria - a general representation

a general representation of any satisficing decision criterion ν can be given by the following result:

theorem: general representation

consider a function $\nu : \mathbb{R}^N \mapsto \mathbb{R} \cup \{-\infty\}$ defined as

$$\nu(\boldsymbol{x}) = \max_{\boldsymbol{\alpha} \in \mathcal{S}} \{\rho(\boldsymbol{\alpha}) \mid \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \;\; \forall \boldsymbol{z} \in \mathcal{U}(\boldsymbol{\alpha})\} \tag{1}$$

for some function $\rho : \mathcal{S} \to \mathbb{R} \cup \{-\infty\}$ on domain $\mathcal{S} \subseteq \mathbb{R}^P$, and for some family of nonempty uncertainty sets $\mathcal{U}(\boldsymbol{\alpha}) \subseteq \mathcal{W}$ defined for all $\boldsymbol{\alpha} \in \mathcal{S}$; then the function ν is a satisficing decision criterion; moreover, any satisficing decision criterion can be represented in a form given by (1) with $\mathcal{S} \subseteq \mathbb{R}^N$.

the s-model: a general family of satisficing models

general s-model

$$\begin{aligned} \max \quad & \rho(\boldsymbol{\alpha}) \\ \text{s.t.} \quad & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq b(\boldsymbol{z}) \quad \forall \boldsymbol{z} \in \mathcal{U}(\boldsymbol{\alpha}) \\ & \boldsymbol{x} \in \mathcal{X} \\ & \boldsymbol{\alpha} \in \mathcal{S} \end{aligned}$$

- adjusts uncertainty sets $\mathcal{U}(\boldsymbol{\alpha})$ for which the constraints remain feasible
- maximizes $\rho(\boldsymbol{\alpha}) : S \to \mathbb{R}$
- careful design of $\rho(\boldsymbol{\alpha})$ and $\mathcal{U}(\boldsymbol{\alpha})$ can lead to meaningful and tractable models

the probabilistic s-model

- recap - the most general satisficing model:

general s-model

$$\begin{array}{rl} \max & \rho(\boldsymbol{\alpha}) \\ \text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq b(\boldsymbol{z}) \;\; \forall \boldsymbol{z} \in \mathcal{U}(\boldsymbol{\alpha}) \\ & \boldsymbol{x} \in \mathcal{X} \\ & \boldsymbol{\alpha} \in \mathcal{S} \end{array}$$

- how to combine useful aspects of both the p-model and the r-model?
- set $\rho(\boldsymbol{\alpha}) = \ln \mathbb{P}\left(\tilde{\boldsymbol{z}} \in \mathcal{U}(\boldsymbol{\alpha})\right)$

probabilistic s-model

$$\begin{array}{rl} \max & \ln \mathbb{P}\left(\tilde{\boldsymbol{z}} \in \mathcal{U}(\boldsymbol{\alpha})\right) \\ \text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \;\; \forall \boldsymbol{z} \in \mathcal{U}(\boldsymbol{\alpha}) \\ & \boldsymbol{x} \in \mathcal{X} \\ & \boldsymbol{\alpha} \in \mathcal{S} \end{array}$$

the t-model: a tractable probabilistic s-model

- uncertain parameters $\tilde{z}_k$, $k \in [K]$ are independently distributed real random variables with support $\mathcal{W}_k$; $\mathcal{W} = \times_{k=1}^{K} \mathcal{W}_k$
- uncertainty defined by affine functions:

$$\boldsymbol{a}_i(\boldsymbol{z}) = \boldsymbol{a}_i^0 + \sum_{k=1}^{K} \boldsymbol{a}_i^k z_k \text{ and } b_i(\boldsymbol{z}) = b_i^0 + \sum_{k=1}^{K} b_i^k z_k$$

- family of adjustable uncertainty sets ("box" type):

$$\mathcal{U}(\boldsymbol{\alpha}) = \mathcal{U}(\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}) = \left\{ \boldsymbol{z} \in \mathbb{R}^K : \boldsymbol{z} \in [\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}] \right\}.$$

t-model

$$\begin{aligned} \max \quad & \sum_{k \in [K]} \ln \mathbb{P} \left(\underline{\alpha}_k \leq \tilde{z}_k \leq \overline{\alpha}_k \right) \\ \text{s.t.} \quad & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \quad \forall \boldsymbol{z} \in [\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}] \\ & \boldsymbol{x} \in \mathcal{X}, \ \underline{\boldsymbol{\alpha}} \leq \overline{\boldsymbol{\alpha}}, \ \underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}} \in \mathcal{W} \end{aligned}$$

reformulation (*robust counterpart*) of the t-model

reformulation: t-model

$$
\begin{aligned}
\max \quad & \sum_{k\in[K]} \ln \mathbb{P}\left(\underline{\alpha}_k \le \tilde{z} \le \overline{\alpha}_k\right) \\
\text{s.t.} \quad & \sum_{j\in[N]} a_{ij}^0 x_j + \sum_{k\in[K]} v_{ik} \ge b_i^0 \quad \forall i \in [M] \\
& v_{ik} \le \sum_{j\in[N]} a_{ij}^k x_j \overline{\alpha}_k - b_i^k \overline{\alpha}_k \quad \forall i \in [M],\ k \in [K] \\
& v_{ik} \le \sum_{j\in[N]} a_{ij}^k x_j \underline{\alpha}_k - b_i^k \underline{\alpha}_k \quad \forall i \in [M],\ k \in [K] \\
& \boldsymbol{x} \in \mathcal{X}, \boldsymbol{v} \in \mathbb{R}^{M\times K}, \\
& \underline{\boldsymbol{\alpha}} \le \overline{\boldsymbol{\alpha}}, \quad \underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}} \in \mathcal{W}.
\end{aligned}
$$

- polynomial number of constraints (good)
- remaining difficulties:
  - non-linear objective function.
  - the terms $x_j \underline{\alpha}_k$ and $x_j \overline{\alpha}_k$ $\ j \in [N], k \in [K]$ are bilinear.

t-model for log-concave densities

the t-model is tractable if:

- the distributions of the random variables are described by log-concave densities
- constraints are linear: e.g., uncertainty in right-hand-side only or, decision variables $\boldsymbol{x}$ are binary

note: consequence for the non-linear objective function:

- if $\tilde{z}_k$ is log-concave, then $\ln \mathbb{P}\left(\underline{\delta} \leq \tilde{z}_k \leq \overline{\delta}\right)$ is a concave function of $(\underline{\delta}, \overline{\delta})$
- the objective function can be approximated by piecewise linear approximation of arbitrary accuracy (density cuts) $\rightarrow$ branch-and-cut

t-model for discrete distributions

- $\mathcal{W}_k = \left\{ \zeta_k^1, \zeta_k^2, \ldots, \zeta_k^{L(k)} \right\}$, $\mathbb{P}\left(\tilde{z}_k = \zeta_k^\ell \right) = p_k^\ell$
- outcomes ζ_k^ℓ sorted in non-decreasing order
- $\mathcal{U}(\boldsymbol{\alpha}) = \left\{ \boldsymbol{z} \in \mathcal{W} \mid \sum_{\ell \in [L(k)]} \zeta_k^\ell \underline{\alpha}_k^\ell \leq z_k \leq \sum_{\ell \in [L(k)]} \zeta_k^\ell \overline{\alpha}_k^\ell, \ \forall k \in [K] \right\}$

t-model for discrete distributions

$$\begin{array}{rl} \max & \ln \mathbb{P}(\tilde{\boldsymbol{z}} \in \mathcal{U}(\boldsymbol{\alpha})) \\ \text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} \geq \boldsymbol{b}(\boldsymbol{z}) \ \ \forall \boldsymbol{z} \in \mathcal{U}(\boldsymbol{\alpha}) \\ & \sum_{\ell \in [L(k)]} \underline{\alpha}_k^\ell = 1, \ \ \sum_{\ell \in [L(k)]} \overline{\alpha}_k^\ell = 1, \ \forall k \in [K] \\ & \sum_{\ell \in [L(k)]} \ell(\overline{\alpha}_k^\ell - \underline{\alpha}_k^\ell) \geq 0, \ \forall k \in [K] \\ & \underline{\boldsymbol{\alpha}}_k, \ \overline{\boldsymbol{\alpha}}_k \in \{0,1\}^{L(k)} \ \forall k \in [K], \ \ \boldsymbol{x} \in \mathcal{X}. \end{array}$$

monotone t-models

a t-model significantly simplifies if it is *monotone*:

definition

a t-model is *monotone* with respect to the uncertain parameters $\tilde{\boldsymbol{z}}$ if there exists a partition $\overline{\mathcal{K}}, \underline{\mathcal{K}} \subseteq [K]$, i.e., $\overline{\mathcal{K}} \cap \underline{\mathcal{K}} = \emptyset$, $\overline{\mathcal{K}} \cup \underline{\mathcal{K}} = [K]$ such that for all $k \in \overline{\mathcal{K}}$

$$\sum_{j \in [N]} a_{ij}^k x_j \leq b_i^k \quad \forall i \in [M], \boldsymbol{x} \in \mathcal{X}$$

and for all $k \in \underline{\mathcal{K}}$

$$\sum_{j \in [N]} a_{ij}^k x_j \geq b_i^k \quad \forall i \in [M], \boldsymbol{x} \in \mathcal{X}$$

a monotone t-model can also be turned into an adjustable t-model for multi-stage decision making !!

adjustable t-model for multi-stage decision making

- $(T+1)$-stage problem.
- first stage, decision $\boldsymbol{x}^0 \in \mathbb{R}^{N_0}$ is made before any uncertainty is realized.
- In subsequent stages, decisions made are $\boldsymbol{x}^1(\tilde{\boldsymbol{z}}_{\mathcal{T}_1}), \ldots, \boldsymbol{x}^T(\tilde{\boldsymbol{z}}_{\mathcal{T}_T})$, where the *recourse decision* $\boldsymbol{x}^t$ at stage $t+1$ is a measurable function $\boldsymbol{x}^t : \mathbb{R}^{|\mathcal{T}_t|} \mapsto \mathbb{R}^{N_t}$ that maps from the realization of the uncertain parameters $\tilde{\boldsymbol{z}}_{\mathcal{T}_t}$ to the appropriate action in $\mathbb{R}^{N_t}$.
- let

$$\boldsymbol{A}(\boldsymbol{z}) = \left[\boldsymbol{A}^0(\boldsymbol{z})\ \boldsymbol{A}^1(\boldsymbol{z})\ \ldots\ \boldsymbol{A}^T(\boldsymbol{z})\right],\ \ \boldsymbol{x}(\boldsymbol{z}) = \left(\boldsymbol{x}^0, \boldsymbol{x}^1(\boldsymbol{z}_{\mathcal{T}_1}), \ldots, \boldsymbol{x}^T(\boldsymbol{z}_{\mathcal{T}_T})\right)$$

of appropriate dimensions so that

$$\boldsymbol{A}(\boldsymbol{z})\boldsymbol{x}(\boldsymbol{z}) = \boldsymbol{A}^0(\boldsymbol{z})\boldsymbol{x}^0 + \sum_{t \in [T]} \boldsymbol{A}^t(\boldsymbol{z})\boldsymbol{x}^t(\boldsymbol{z}_{\mathcal{T}_t}).$$

adjustable t-model for multi-stage decision making, cont.

- formulate the adjustable T-model as follows:

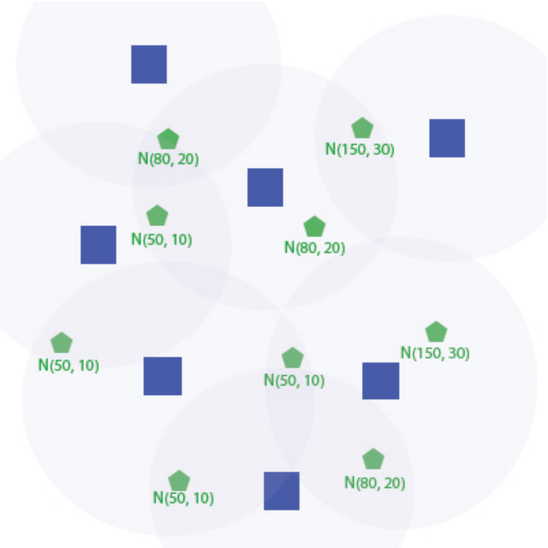
$$
\begin{array}{lll}
\max & \sum_{k\in[K]} \ln \mathbb{P}\left(\underline{\alpha}_k \le \tilde{z}_k \le \overline{\alpha}_k\right) & \\
\text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x}(\boldsymbol{z}) \ge \boldsymbol{b}(\boldsymbol{z}) & \forall \boldsymbol{z} \in [\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}] \\
& \boldsymbol{x}(\boldsymbol{z}) \in \mathcal{X} & \forall \boldsymbol{z} \in \mathcal{W} \\
& \boldsymbol{x}^t \in \mathcal{R}(|\mathcal{T}_t|, N_t) & \forall t \in [T] \\
& \underline{\boldsymbol{\alpha}} \le \overline{\boldsymbol{\alpha}},\ \underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}} \in \mathcal{W}, &
\end{array}
$$

  where $\mathcal{R}(m, n)$ denotes the family of all measurable functions that map from $\mathbb{R}^m$ to $\mathbb{R}^n$.
- under fixed recourse assumptions and our monotonicity condition, equivalent to solving a one-stage problem.

stochastic maximum coverage facility location problem

given:

- $\mathcal{I}$ candidate facility locations; $\mathcal{J}$ customer demands $\tilde{d}_j$
- network is "sparse": each customer can be covered by approx. 15% - 40% of all facilities
- available budget B; facility construction costs c_i; capacities a_i



stochastic maximum coverage facility location problem

select facilities that maximize the probability that all demands can be satisfied:

initial p-model for maximum coverage problem, hard to solve

$$
\begin{array}{lll}
\max & \ln \mathbb{P}\left(\sum_{i \in \mathcal{I}_j} y_{ij}(\tilde{\boldsymbol{z}}) \geq \tilde{z}_j \quad \forall j \in \mathcal{J} \right) & \\
\text{s.t.} & \sum_{j \in \mathcal{J}_i} y_{ij}(\boldsymbol{z}) \leq a_i x_i & \forall \boldsymbol{z} \in \mathcal{W}, i \in \mathcal{I} \\
& \sum_{i \in \mathcal{I}} c_i x_i \leq B & \\
& y_{ij}(\boldsymbol{z}) \geq 0 & \forall \boldsymbol{z} \in \mathcal{W}, i \in \mathcal{I}, j \in \mathcal{J}_i \\
& y_{ij}(.) \in \mathcal{R}(|\mathcal{J}|, 1) & \forall i \in \mathcal{I}, j \in \mathcal{J}_i \\
& x_i \in \{0, 1\} & \forall i \in \mathcal{I},
\end{array}
$$

t-models and monte carlo benchmarks

t-models:

- T-1: branch-and-cut for log-concave densities
    - maximizes the probability that each demand is met
    - assumes knowledge of the probability distribution
- T-2: sample based model (discrete distribution)
    - L data samples (scenarios)
    - maximizes # of outcomes that are feasible in constraints
    - no assumptions about probability distributions

SAA models with L data samples:

- P-1: maximizing feasibility probability
    - maximizes the number of feasible scenarios (obj.: P-model)
- E: minimizing expected demand shortfall
    - minimizes the expected demand shortfall (obj.: expected value)

computational settings

problem instances (total of 60):

- # customers $|\mathcal{J}| \in \{100, 250, 500, 1000, 2000\}$
- # facilities $|\mathcal{I}| \in \{0.5|\mathcal{J}|, |\mathcal{J}|, 2|\mathcal{J}|\}$
- network density $A_p \in \{15, 20, 30, 40\}$
- $\tilde{d}_j \sim N\left(\mu_j, (0.5\mu_j)^2\right); \quad \mu_j \sim U(1, 100)$
- budget B set 1.05 times the costs required to satisfy the average demand

computational settings:

- CPLEX 12.6.1 with standard parameters
- 12hrs computing time limit, 24gb memory limit
- evaluation via Monte Carlo simulation (100,000 samples)

customers - scalability of the model T-1

$\lvert\mathcal{J}\rvert$	succ. rate %	demand shortfall	time (minutes)
100	84.41	2.0	22.7
250	82.76	4.2	2.7
500	99.93	0.0	3.4
1000	96.60	1.2	30.7
2000	95.97	2.2	426.9
all	**92.06**	**1.9**	**98.5**

Table: Out-of-sample performance study for different problem sizes, reporting average success rate (%), average demand shortfall (in 10 units), average computing time (in minutes)

- solves all instances
- high success rates and low shortfalls for all problem sizes
- reasonable computing times

L - scalability: data samples based models

average over all 60 instances

L	T-2 succ. rate %	T-2 short fall	T-2 time	T-2 # ns	P-1 succ. rate %	P-1 short fall	P-1 time	P-1 # ns	E succ. rate %	E short fall	E time	E # ns
5	88.73	350.9	44.9	2	64.65	2,440.8	207.0	16	70.10	1,743.6	181.2	12
10	85.39	699.8	62.5	4	43.99	3,335.8	379.1	29	52.69	2,826.7	354.8	22
15	88.41	351.0	75.3	2	35.06	3,523.0	448.7	35	41.85	3,037.6	463.2	30
50	86.64	525.4	54.9	3	20.37	3,627.7	606.7	45	26.54	3,370.2	569.8	39

- # ns: number of instances without feasible solution
- models P-1 and E hard to solve as *L* increases
- model T-2 remains relatively stable

L - scalability & robustness

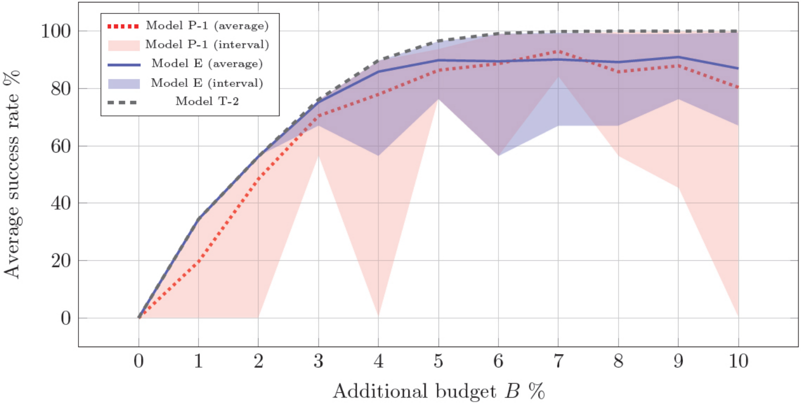
average [min, max] among averages of 10 replications increasing L: all instances

J	P-1	E	T-2				
	L=5	L=5	L=5	L=15	L=50	L=100	L=500
100	80.70	76.29	80.41	80.56	84.69	84.57	84.47
	[71.7, 86.8]	[66.4, 85.8]	[64.0, 87.3]	[70.7, 86.3]	[82.2, 86.9]	[81.3, 87.1]	[80.6, 86.8]
250	81.11	82.51	80.44	80.80	81.57	82.38	82.43
	[69.1, 82.8]	[82.3, 82.8]	[75.4, 82.7]	[75.6, 82.5]	[75.8, 82.7]	[81.9, 82.5]	[81.8, 82.5]
500	96.66	95.69	98.29	99.92	99.92	99.92	99.92
	[91.7, 99.9]	[86.9, 99.9]	[91.8, 99.9]	[99.9, 99.9]	[99.9, 99.9]	[99.9, 99.9]	[99.9, 99.9]
1000	47.90	70.96	96.58	96.62	96.63	96.63	96.63
	[23.7, 63.4]	[60.2, 88.5]	[96.5, 96.7]	[96.6, 96.6]	[96.6, 96.6]	[96.6, 96.6]	[96.6, 96.6]
2000	0.77	26.84	73.97	72.41	73.19	82.63	77.93
	[0.0, 7.7]	[20.3, 31.3]	[63.0, 86.5]	[63.0, 78.7]	[63.0, 86.6]	[63.0, 94.5]	[63.0, 94.5]
all	61.25	70.40	86.03	86.15	87.24	89.30	88.34
	[56.1, 65.2]	[67.6, 77.4]	[82.5, 88.8]	[84.0, 88.5]	[85.4, 88.9]	[85.8, 91.7]	[85.5, 91.9]

Table: Comparison of average [minimum, maximum] success rates (%) over all problem instances among 10 replications for Models P-1 and E with same sample size $L = 5$, and for Model T-2 with different sample sizes.

investment study: success rates

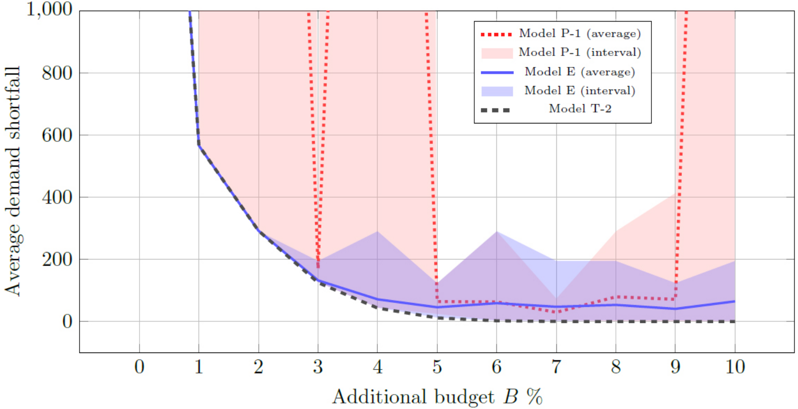
average intervals between min/max among 10 runs
($L = 5$, $|\mathcal{I}| = 500$, $|\mathcal{J}| = 1000$)



- P-1 and E models unstable; T-2 model stable at highest success rates

investment study: demand shortfall

average intervals between min/max among 10 runs
($L = 5$, $|\mathcal{I}| = 500$, $|\mathcal{J}| = 1000$)



- P-1 and E models unstable; T-2 model stable at lowest shortfalls

conclusions

contributions:

- introduction of the s-model
  - flexible adjustment of uncertainty sets
  - generalizes the p-model
  - provides link to general chance-constrains and robust optimization problems
- general framework: allows for many tractable implementations, e.g., the t-model
- exemplified for continuous and discrete/empiric distributions
  - log-concave density functions: cut-based solution methods
  - data sampling/discrete distributions: efficient reformulation to mixed-integer program

conclusions

contributions - computational experiments:

- maximum coverage facility location problem
- large problem instances
- benchmark approaches (SAA) cannot handle large sample sizes
- t-models scale well for all instances
  - knowledge about probability distribution helps
  - without available distributions, large sample size results in stable results for all instances

future research directions

- simple idea and easy to implement
  - relevant for decision makers in practice
  - applicable to many (difficult) problems
  - high performance $\rightarrow$ competitive alternative to traditional sampling methods
- cut-based method has been explored for NP-hard MIP
  - likely to be very quick for linear programs (cutting plane)
- further implementations of the S-model
  - other cases may yield tractable models for important problems
- scalability of data-driven approach
  - may handle even larger data sets when solved by advanced optimization methods - big data/machine learning?