Stochastic Optimization
IDA PhD course 2011ht

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

4. Lecture: Simple Recourse Problems
03. November 2011

Outline

Deterministic Opt. Model → Stochastic Programming Model

$$\begin{array}{ll} \max_{x \in X} & f(x) \\ \text{s.t.} & G(x) \leq 0 \end{array} \quad \rightarrow \quad \begin{array}{ll} \min_{x \in X} & f(x) \\ \text{s.t.} & G(x, \chi) \leq 0 \end{array}$$

$\chi \in \Omega \subseteq \mathbb{R}^s$: random vector
$G : \mathbb{R}^n \times \mathbb{R}^s \rightarrow \mathbb{R}^m$
$G(x, \chi) = (g_1(x, \chi), \ldots, g_m(x, \chi))$

Question

What means "feasible solution" in case of uncertain parameters in the constraint function(s)?

- Feasible in all possible cases → conservative/infeasible/trivial
- Feasible with high probability → chance-constraint ✓
- Average violation not too bad → simple recourse
- Feasible after correction has been made → multi-stage

Linköping University

Outline

Linköping University

Chance-Constrained Stochastic Optimization Problem

$$\begin{aligned} \min_{x \in X} \quad & f(x) \\ \text{s.t.} \quad & \mathbb{P}\{G(x, \chi) \leq 0\} \geq p \end{aligned}$$

Outline

Definition (Concave Function)

A function $f : \mathbb{R}^n \to \mathbb{R}$ is called **concave** if for all $x, y \in dom(f)$ and all $t \in [0, 1]$ it is

$$f(tx + (1 - t)y) \geq tf(x) + (1 - t)f(y)$$

Definition (Log-Concave Function)

A function $f : \mathbb{R}^n \to \mathbb{R}$ is called **logarithmic concave (log-concave)** if $\log(f)$ is a concave function.

Property

A function $f : \mathbb{R}^n \to \mathbb{R}$ is log-concave iff for all $x, y \in dom(f)$ and all $\theta \in [0, 1]$ it is

$$f(\theta x + (1 - \theta)y) \geq f(x)^{\theta} f(y)^{(1-\theta)}$$

Linköping University

Definition (Log-Concave Function)

A function $f : \mathbb{R}^n \to \mathbb{R}$ is called log-concave if $\log(f)$ is a concave function.

Definition (Log-Concave Probability Distribution)

A continuous probability distribution is called a **log-concave probability distribution** if the corresponding density function is log-concave.

Outline

Theorem (Prékopa '72)

Let $g_i(x, y)$ $(i = 1, \dots, m)$ be **concave functions** *on $\mathbb{R}^n \times \mathbb{R}^s$ (where x is an n-dimensional and y an s-dimensional vector). Let further χ be an s-dimensional random vector with* **logarithmic concave probability distribution***. Then, the left hand side x-function of the joint chance-constraint*

$$\mathbb{P}\{g_i(x, \chi) \geq 0, i = 1, \dots, r\} \geq p \tag{1}$$

is **logarithmic concave** *in the entire space $\mathbb{R}^n$.*

Corollary (Tamm '77)

Let $g_i(x, y)$ *(*$i = 1, \ldots, m$*) be* **quasi-concave functions** *on* $\mathbb{R}^n \times \mathbb{R}^s$ *(where* x *is an n-dimensional and* y *an s-dimensional vector). Let further* χ *be an s-dimensional random vector with* **logarithmic concave probability distribution**. *Then the joint chance-constraint*

$$\mathbb{P}\{g_i(x, \chi) \geq 0, i = 1, \ldots, r\} \geq p \quad (2)$$

defines a **convex set**.

Outline

Linköping University

Definition (Quasi-Concave Function)

A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called **quasi-concave** if for all $x, y \in dom(f)$ and all $t \in [0, 1]$ it is

$$f(tx + (1-t)y) \geq min(f(x), f(y))$$

Generalization by Tamm ('76/'77)

Prékopa's results stay valid if the g_i's are only quasi-concave!

Corollary (Tamm '77)

Let $g_i(x, y)$ $(i = 1, \ldots, m)$ be **quasi-concave functions** *on $\mathbb{R}^n \times \mathbb{R}^s$ (where x is an n-dimensional and y an s-dimensional vector). Let further χ be an s-dimensional random vector with* **logarithmic concave probability distribution**. *Then the joint chance-constraint*

$$\mathbb{P}\{g_i(x, \chi) \geq 0, i = 1, \ldots, r\} \geq p \tag{2}$$

defines a **convex set**.

Examples of log-concave probability distributions

- Uniform distribution
- Normal distributions
- Exponential distribution
- Laplace distribution
- ...

Linköping University

Outline

Linköping University

Basic Idea(s)

- Allow violation
- Restrict average "amount of violation"
- Introduce penalty per "violation unit"
- Find good trade-off: cost $\leftrightarrow$ penalty

Simple-Recourse problem

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

- $d_i > 0 \; \forall i \in \{1, \ldots, m\}$
- $[x]^+ = max(0, x)$

Question

Why not choosing an expectation-constrained model?

Expectation-Constrained Optimization Problem

$$\min_{x \in X} \quad f(x)$$
$$\text{s.t.} \quad \mathbb{E}\left[[g_i(x,\chi)]^+\right] \leq \delta_i \quad \forall i = 1,\ldots,m$$

Answer

- No difference between "small" and "big" average violation.
- Single objective.
- No "flexibility".

Linköping University

Advantages:

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

Disadvantages:

- Probability of violation not restricted
- Computation of $\mathbb{E}\left[[g_i(x,\chi)]^+\right]$ in gen. difficult

Outline

Linköping University

Discrete Finite Distribution

Simple-Recourse Problem

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

$\chi^1, \ldots, \chi^K \in \mathbb{R}^s$: scenarios
$\mathbb{P}\{\chi = \chi^k\} := p^k$: probabilities

Reformulate Problem Deterministically

$$\min_{x \in X} \quad f(x) + \sum_{i=1}^{m} d_i \cdot \sum_{k=1}^{K} [g_i(x,\chi^k)]^+$$

Linköping University

Discrete Finite Distribution II

Deterministically reformulated problem

$$\min_{x \in X} \quad f(x) + \sum_{i=1}^{m} d_i \cdot \sum_{k=1}^{K} [g_i(x, \chi^k)]^+$$

Good news

- $f(\cdot)$, $g_i(\cdot, \hat{\chi})$ convex $(\forall \hat{\chi}, \forall i \in \{1, \ldots, n\})$
  $\Longrightarrow$ Objective function convex
- $f(\hat{\chi})$, $g_i(\cdot, \hat{\chi})$ linear $(\forall \hat{\chi}, \forall i \in \{1, \ldots, n\})$
  $\Longrightarrow$ Objective function piecewise linear

Problems

- ?

g_i linear / Normal Distribution / Independence

Simple-Recourse Problem

$$\min_{x \in X} \quad f(x) + \sum_{i=1}^{m} d_i \cdot \mathbb{E}\left[[a^i(\chi)x - b_i]^+\right]$$

$a^i(\chi) \in \mathbb{R}^n$: random vector with ind. normally distr. entries
$a^i_j(\chi) \sim \mathcal{N}(\mu^i_j, \sigma^i_j)$

g_i linear / Normal Distribution / Independence II

φ: Density function of standard normal distribution
Φ: Cumulative distribution function of standard normal distribution
$\mu_x^i := \sum_{j=1}^n \mu_j^i x_j$
$\sigma_x^i := \sqrt{\sum_{j=1}^n (\sigma_j^i)^2 x_j^2}$

Deterministically reformulated problem

$$\min_{x \in X} \quad f(x) + \sum_{i=1}^m d_i \cdot \left[\sigma_x^i \cdot \varphi\left(\frac{b_i - \mu_x^i}{\sigma_x^i}\right) - (b_i - \mu_x^i) \cdot \left[1 - \Phi\left(\frac{b_i - \mu_x^i}{\sigma_x^i}\right)\right]\right]$$

Linköping University

g_i linear / Normal Distribution / Independence III

Deterministically reformulated problem

$$\min_{x \in X} \quad f(x) + \sum_{i=1}^{m} d_i \cdot \left[\sigma_x^i \cdot \varphi\left(\frac{b_i - \mu_x^i}{\sigma_x^i}\right) - (b_i - \mu_x^i) \cdot \left[1 - \Phi\left(\frac{b_i - \mu_x^i}{\sigma_x^i}\right)\right]\right]$$

Good news

- Objective function evaluation easy.
- f convex $\Longrightarrow$ Objective function convex.
- Objective function differentiable.

Problems

- No analytic description.

G linear / Poisson Distribution / Independence

Simple-Recourse Problem

$$\min_{x \in X} \quad f(x) + d \cdot \mathbb{E}\left[[\chi^T x - b]^+\right]$$

$\chi \in \mathbb{R}^n$: random vector with independent Poisson distr. entries
μ: Vector of means of χ

$\chi^T x$: Poisson with mean $\hat{\mu} = \mu^T x$

Deterministically reformulated problem

$$\min_{x \in X} \quad f(x) + d \cdot \sum_{k=b+1}^{\infty} \frac{e^{-\hat{\mu}} \hat{\mu}^k}{k!}(k - b)$$
$$\text{s.t.} \quad \hat{\mu} = \mu^T x$$

Outline

Linköping University

Theorem

Let

- $\gamma(x) := \mathbb{E}\left[[g(x,\chi)]^+\right]$
- $\mathbb{E}\left[g(x,\chi)\right] < \infty \ \forall x \in \mathbb{R}^n$

Then the following holds:

$$g(\cdot,\hat{\chi}) \text{ is convex } \forall \hat{\chi} \quad \Longrightarrow \quad \gamma \text{ is convex}$$

Theorem

Let

- $\gamma(x) := \mathbb{E}\left[[a(\chi)x - b]^+\right]$
- Φ*: cumulative distribution function of of* $a(\chi)x - b$
- $\mathbb{E}\left[a(\chi)x - b\right] < \infty \ \forall x \in \mathbb{R}^n$

Then the following holds:

1. γ *is convex*
2. *For all* $x \in \mathbb{R}^n$ $\gamma(x) < \infty$.
3. γ *is Lipschitz continuous.*
4. γ *is differentiable wherever* Φ *is continuous.*
5. $a(\chi)x - b$ *discretely distributed* $\Rightarrow \gamma$ *is piecewise linear.*

Linköping University

QUESTIONS?

What about next week?