

Convex Optimization II

Lecture 14: Risk Averse Optimization

Hamed Shah-Mansouri

Department of Electrical Engineering
Sharif University of Technology

1400-2

OUTLINE

- The newsvendor problem
- Chance constrained optimization
- Value at Risk
- Convex Approximation
- Conditional Value at Risk

MOTIVATIONS

- In stochastic programming problems, we aim to optimize the cost on average.
- However, the instances of real decisions may incur a cost that can be quite different from the optimal-on-average cost.
- A natural question is whether we can control the risk of the cost to be not too high.

[1] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modelling and Theory*. Society for Industrial and Applied Mathematics (SIAM), 2nd Ed., 2014.

[2] A. Shapiro and A. Philpott, "A Tutorial on Stochastic Programming", [Online]. Available.

http://www.isye.gatech.edu/people/faculty/Alex_Shapiro/TutorialSP.pdf

THE NEWSVENDOR PROBLEM

- Recall the newsvendor problem from the last lecture, where we denoted the cost of ordering x with demand d as $G(x, d)$.
- For a particular realization of the demand D , the cost $G(\bar{x}, D)$ can be quite different from the optimal-on-average cost $\mathbb{E}G(\bar{x}, D)$.
- By introducing a new constraint, we can manage the risk of having a too high cost.
- We may add the constraint

$$G(x, D) \leq \gamma$$

to be satisfied for all possible realizations of the demand D .

- That is, we want to make sure that the total cost will be at most γ in all possible circumstances.

CHANCE CONSTRAINT

Question: What is the drawback of including the above constraint?

- This could be quite restrictive. In particular, if there is at least one realization d resulting in cost greater than γ , then the corresponding problem has no feasible solution.
- In such situations, it makes more sense to introduce the constraint that the probability of $G(x, D) > \gamma$ is less than a specified value (significance level) $\alpha \in (0, 1)$.
- This leads to the so-called **chance constraint**

$$\mathbf{Prob} (G(x, D) > \gamma) \leq \alpha$$

or equivalently

$$\mathbf{Prob} (G(x, D) \leq \gamma) \geq 1 - \alpha$$

BASIC STOCHASTIC PROGRAMMING PROBLEM

$$\begin{aligned} & \underset{x}{\text{minimize}} \quad F_0(x) = \mathbb{E}f_0(x, \omega) \\ & \text{subject to} \quad F_i(x) = \mathbb{E}f_i(x, \omega) \leq 0, \quad i = 1, \dots, m \end{aligned}$$

- Problem data are f_i and distribution of ω
- If $f_i(x, \omega)$ are convex in x for each ω , $F_i(x)$ are convex, so is the stochastic programming problem.

Question: How to solve the above optimization problem? [Approximation with chance constraint](#)

CHANCE CONSTRAINT AND PERCENTILE OPTIMIZATION

- Chance Constraint

$$\mathbf{Prob} (f_i(x, \omega) \leq 0) \geq \eta_i$$

- ▶ η_i is called *confidence level*
- ▶ generally interested in $\eta_i = 0.9, 0.95, 0.99$.

- Percentile Optimization

minimize γ

subject to $\mathbf{Prob} (f_0(x, \omega) \leq \gamma) \geq \eta$

VALUE AT RISK

- Value-at-risk of random variable z , at level η :

$$\mathbf{VaR}_\eta(z) = \inf \{ \gamma \mid \mathbf{Prob}(z \leq \gamma) \geq \eta \}$$

- Chance Constraint

$$\mathbf{Prob}(f_i(x, \omega) \leq 0) \geq \eta$$

can be written as

$$\mathbf{VaR}_\eta(f_i(x, \omega)) \leq 0$$

EXAMPLE: PORTFOLIO OPTIMIZATION

- n investment opportunities, with random return rates r_1, \dots, r_n .
- $\mathbf{x} \in \mathbb{R}^n$ gives portfolio allocation (fraction)
- Portfolio return is $\mathbf{r}^T \mathbf{x}$, where r_i follows normal distribution $\mathcal{N}(\bar{r}, \sigma)$ (a more realistic model is log-normal)
- The objective is to maximize the expected return subject to limit on the probability of loss.

EXAMPLE: PORTFOLIO OPTIMIZATION

The portfolio maximization problem is

$$\begin{aligned} & \text{maximize } \mathbb{E} \mathbf{r}^T \mathbf{x} \\ & \text{subject to } \mathbf{Prob}(\mathbf{r}^T \mathbf{x} \leq 0) \leq 1 - \eta \\ & \quad \mathbf{1}^T \mathbf{x} = 1 \end{aligned}$$

can be expressed as a convex problem.

Linear combination with normally distributed parameter also gives a convex VaR constraint.

EXAMPLE: PORTFOLIO OPTIMIZATION

- $n = 10$ assets, $1 - \eta = 0.05$, a normal distribution

portfolio	$\mathbb{E} \mathbf{r}^T \mathbf{x}$	$\mathbf{Prob}(\mathbf{r}^T \mathbf{x} \leq 0)$
optimal	7.51	5.0 %
no VaR	10.66	20.3 %
uniform portfolio	3.41	18.9 %

Source: J. Duchi

CHANCE CONSTRAINTS FOR LOG-CONCAVE DISTRIBUTIONS

Suppose

- ω has log-concave distribution $p(\omega)$
- Set $\mathcal{C} = \{(x, \omega) \mid f(x, \omega) \leq 0\}$ is convex in (x, ω)

Then, the following function is log-concave.

$$\mathbf{Prob}(f(x, \omega) \leq 0) = \int_{\mathcal{C}} p(\omega) d\omega$$

So, the chance constraint

$$\mathbf{Prob}(f(x, \omega) \leq 0) \geq \eta$$

can be expressed as the convex set

$$\log \mathbf{Prob}(f(x, \omega) \leq 0) \geq \log \eta$$

CHANCE CONSTRAINED OPTIMIZATION

Question: How to deal with non-convex VaR constraints?

CONVEX APPROXIMATION OF VAR

- assume $f_i(x, \omega)$ is convex in x .
- suppose $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is nonnegative convex nondecreasing, with $\phi(0) = 1$
- step function

$$\mathbf{1}(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z \leq 0 \end{cases}$$

- for any $\alpha_i > 0$, $\phi(z/\alpha_i) \geq \mathbf{1}(z > 0)$ for all z , so

$$\mathbb{E}\phi(f_i(x, \omega)/\alpha_i) \geq \mathbf{Prob}(f_i(x, \omega) > 0)$$

- Hence,

$$\mathbb{E}\phi(f_i(x, \omega)/\alpha_i) \leq 1 - \eta$$

ensures chance constraint $\mathbf{Prob}(f_i(x, \omega) \leq 0) \geq \eta$ holds.

CONVEX APPROXIMATION OF VAR

$$\alpha_i \mathbb{E} \phi(f_i(x, \omega)/\alpha_i) \leq \alpha_i(1 - \eta)$$

- Perspective function $v\phi(u/v)$ is convex in (u, v) for $v > 0$, nondecreasing in u .
- So composition $\alpha_i \mathbb{E} \phi(f_i(x, \omega)/\alpha_i)$ is convex in (x, α_i) for $\alpha_i > 0$
- Hence, the above conservative approximation ensures chance constraint is **convex** in x and α_i .

We can optimize over x and $\alpha_i > 0$ via **convex optimization**.

The smaller the function $\phi(\cdot)$ is, the better this approximation will be.

Theory says that $\phi(u) = (u + 1)_+$ is a best choice of such a function.

CONVEX APPROXIMATION OF VAR

Markov chance constraint bound: Set $\phi(u) = (u + 1)_+$
Convex approximation constraint:

$$\text{Prob}(f_i(x, \omega) > 0) \leq \mathbb{E}(f_i(x, \omega)/\alpha_i + 1)_+$$

which can be written as

$$\mathbb{E}(f_i(x, \omega) + \alpha_i)_+ \leq \alpha_i(1 - \eta)$$

Chebyshev chance constraint bound: Set $\phi(u) = (u + 1)_+^2$
Convex approximation constraint:

$$\mathbb{E}(f_i(x, \omega) + \alpha_i)_+^2 / \alpha_i \leq \alpha_i(1 - \eta)$$

CONVEX APPROXIMATION OF VAR

Traditional Chebyshev chance constraint bound: Dropping + projection
Convex approximation constraint:

$$\alpha_i \mathbb{E}(f_i(x, \omega) / \alpha_i + 1)^2 \leq \alpha_i(1 - \eta)$$

which can be written as

$$\mathbb{E}(f_i(x, \omega) + \alpha_i)^2 \leq \alpha_i^2(1 - \eta)$$

and

$$2\mathbb{E}(f_i(x, \omega)) + 1/\alpha_i \mathbb{E}(f_i(x, \omega))^2 + \alpha_i \eta \leq 0$$

Minimizing over α_i gives

$$\mathbb{E}(f_i(x, \omega)) + (\eta \mathbb{E}(f_i(x, \omega))^2)^{1/2} \leq 0$$

Question: Why do we minimize over α_i ?

EXAMPLE

- maximize a linear revenue function (say) subject to random linear constraints holding with probability η :

$$\begin{array}{ll}\text{maximize} & c^T x \\ \text{subject to} & \mathbf{Prob}(\max(Ax - b) \leq 0) \geq \eta\end{array}$$

with variable $x \in \mathbf{R}^n$; $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$ random (Gaussian)

- Markov/CVaR approximation:

$$\begin{array}{ll}\text{maximize} & c^T x \\ \text{subject to} & \mathbf{E}(\max(Ax - b) + \alpha)_+ \leq \alpha(1 - \eta)\end{array}$$

with variables $x \in \mathbf{R}^n$, $\alpha \in \mathbf{R}$

Source: J. Duchi

EXAMPLE

- Chebyshev approximation:

$$\begin{array}{ll}\text{maximize} & c^T x \\ \text{subject to} & \mathbf{E}(\max(Ax - b) + \alpha)_+^2 / \alpha \leq \alpha(1 - \eta)\end{array}$$

with variables $x \in \mathbf{R}^n$, $\alpha \in \mathbf{R}$

- optimal values of these approximate problems are lower bounds for original problem

Source: J. Duchi

EXAMPLE

- instance with $n = 5$, $m = 10$, $\eta = 0.9$
- solve approximations with sampling method with $N = 1000$ training samples, validate with $M = 10000$ samples
- compare to solution of deterministic problem

$$\begin{array}{ll}\text{maximize} & c^T x \\ \text{subject to} & \mathbf{E} Ax \leq \mathbf{E} b\end{array}$$

- estimates of $\mathbf{Prob}(\max(Ax - b) \leq 0)$ on training/validation data

	$c^T x$	train	validate
Markov	3.60	0.97	0.96
Chebyshev	3.43	0.97	0.96
deterministic	7.98	0.04	0.03

Source: J. Duchi

CONDITIONAL VALUE AT RISK

CVaR or the *expected shortfall* at η % level is the expected return on the portfolio in the worst $\eta\%$ of cases.

$$\mathbf{CVaR}_\eta(z) = \inf_{\beta} (\beta + 1/(1 - \eta) \mathbb{E}(z - \beta)_+)$$

Take the 1st derivative w.r.t. β

$$1 - 1/(1 - \eta) \mathbf{Prob}(z \geq \beta) = 0$$

$$\mathbf{Prob}(z \geq \beta^*) = 1 - \eta \Rightarrow \beta^* = \mathbf{VaR}_\eta(z)$$

CONDITIONAL VALUE AT RISK

Conditional tail expectation (or expected shortfall)

$$\begin{aligned}\mathbb{E}(z \mid z \geq \beta^*) &= \mathbb{E}(\beta^* + (z - \beta^*) \mid z \geq \beta^*) \\ &= \beta^* + \mathbb{E}((z - \beta^*)_+) / \mathbf{Prob}(z \geq \beta^*) \\ &= \mathbf{CVaR}_\eta(z)\end{aligned}$$

CONDITIONAL VALUE AT RISK

Recall the Markov chance constraint bound: Set $\phi(u) = (u + 1)_+$

Convex approximation constraint:

$$\mathbb{E}(f_i(x, \omega) + \alpha_i)_+ \leq \alpha_i(1 - \eta)$$

Minimizing over α_i

$$\inf_{\alpha_i > 0} \left(\frac{\mathbb{E}(f_i(x, \omega) + \alpha_i)_+}{1 - \eta} - \alpha_i \right)$$

This is $\mathbf{CVaR}_\eta(f_i(x, \omega))$.

So, convex approximation replaces $\mathbf{VaR}_\eta(f_i(x, \omega)) \leq 0$ with $\mathbf{CVaR}_\eta(f_i(x, \omega)) \leq 0$, which is convex.

SUMMARY

- Control the risk of the cost to be not too high.
- Chance constrained optimization ensures that constraints are met with high probability.
- VaR is measure of the risk of loosing a constraint.
- Convex approximation of VaR yields convex optimization problems.
- CVaR considers some of the worst-case scenarios.