IE 331 OR: Optimization **KAIST, Spring 2023**
Lecture #19: Value at Risk (VaR) and Conditional Value at Risk (CVaR) May 9, 2023
Lecturer: Dabeen Lee

1 Outline

In this lecture, we cover

- Value at Risk (VaR),
- Conditional Value at Risk (CVaR),
- modeling with CVaR.

2 VaR and CVaR

Suppose that we have eight scenarios with equal probability. Consider a decision x with the following outcomes.

i	1	2	3	4	5	6	7	8
p_i	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8
$g(x, \xi_i)$	100	7	5	4	3	1	0	-2

How can we judge how "risky" the decision is? Note that

- Expectation: 29.5.
- Worst-case value: 100.

However, 7 out of the 8 scenarios have values at most 7.

- What about looking at the **2nd highest value** instead?

The second highest value is 7, and this perhaps better represents the risk of decision x. Consider an alternative decision x'.

i	1	2	3	4	5	6	7	8
p_i	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8
$g(x', \xi_i)$	20	7	5	4	3	1	0	-2

Here, the second-highest value for decision x' is also 7. However, we know that x has a worse value than x' in the worst case. Therefore, we should capture the difference between x and x' somehow.

- What about the **average of the two highest values**?

For x, we have $(100 + 7)/2 = 53.5$, while for x', we have $(20 + 7)/2 = 13/5$.

The following two risk measures essentially capture these ideas.

- **Value-at-Risk (VaR)**: Look at the k**th largest value** among $g(x, \xi_1), \ldots, g(x, \xi_N)$.
- **Conditional Value-at-Risk (CVaR)**: Look at the **average of the top k values** among $g(x, \xi_1), \ldots, g(x, \xi_N)$.

2.1 Risk measure 3: Value-at-Risk (VaR)

Assume that we have likelikhood weights p_i for each scenario ξ_i and the distribution $\hat{P}_N$ with

$$\mathbb{P}_{\xi\sim\hat{P}_N}\left[\xi = \xi_i\right] = p_i, \quad i \in [N].$$

Fix some $\alpha \in (0,1)$. Then the **Value-at-Risk** at level α or α**-VaR** is the risk measure defined as

$$\mathrm{VaR}_\alpha\left(g(x,\xi);\hat{P}_N\right) = \min\left\{t:\ \mathbb{P}_{\xi\sim\hat{P}_N}\left[g(x,\xi)\le t\right] > \alpha\right\}.$$

This is also referred to as the α**-quantile** of $\hat{P}_N$.

When $p_i = 1/N$ for $i \in [N]$ and $\alpha = 1 - k/N$, then $\mathrm{VaR}_\alpha\left(g(x,\xi);\hat{P}_N\right)$ is exactly the kth largest value among $g(x,\xi_1),\ldots,g(x,\xi_N)$.

Example 19.1. Suppose that we have

i	1	2	3	4	5	6
p_i	0.05	0.15	0.1	0.4	0.2	0.1
$g(x,\xi_i)$	10	8	6	3	2	-2
$\mathbb{P}_{\xi\sim\hat{P}_N}\left[g(x,\xi)\le g(x,\xi_i)\right]$	1	0.95	0.8	0.7	0.3	0.1

Then

- $\mathrm{VaR}_{0.98}\left(g(x,\xi);\hat{P}_N\right) = 10.$
- $\mathrm{VaR}_{0.95}\left(g(x,\xi);\hat{P}_N\right) = 10.$
- $\mathrm{VaR}_{0.85}\left(g(x,\xi);\hat{P}_N\right) = 8.$
- $\mathrm{VaR}_{0.8}\left(g(x,\xi);\hat{P}_N\right) = 8.$
- $\mathrm{VaR}_{0.7}\left(g(x,\xi);\hat{P}_N\right) = 6.$
- $\mathrm{VaR}_{0.6}\left(g(x,\xi);\hat{P}_N\right) = 3.$

2.2 Risk measure 4: Conditional Value-at-Risk (CVaR)

Assume that we have likelikhood weights p_i for each scenario ξ_i and the distribution $\hat{P}_N$ with

$$\mathbb{P}_{\xi\sim\hat{P}_N}\left[\xi = \xi_i\right] = p_i, \quad i \in [N].$$

Fix some $\alpha \in (0,1)$. Then the **Conditional Value-at-Risk** at level α or α**-CVaR** is the risk measure defined as

$$\begin{aligned}\mathrm{CVaR}_\alpha\left(g(x,\xi);\hat{P}_N\right) \quad := \quad \max \quad & \frac{1}{1-\alpha}\sum_{i\in[N]} z_i\cdot g(x,\xi_i)\\ \text{s.t.} \quad & 0\le z_i\le p_i, \quad i\in[N]\\ & \sum_{i\in[N]} z_i = 1-\alpha.\end{aligned}$$

What is this? To compute the value of α-CVaR, we need to solve a linear program. In fact, although it is a linear program, we can find an optimal solution by a **greedy algorithm** as follows.

1. Suppose that $g(x, \xi_1) \geq g(x, \xi_2) \geq \cdots \geq g(x, \xi_N)$.
2. Initialize budget $= 1 - \alpha$ and $i = 1$.
3. While budget > 0 do
   - Set $z_i = \min\{p_i,\ \text{budget}\}$.
   - budget $\leftarrow$ budget $- z_i$.
   - $i \leftarrow i + 1$.

Basically, to maximize the objective, we assign high weights to risky scenarios under the weight limit of p_i on scenario ξ_i for $i \in [N]$.

Example 19.2. Suppose that we have

i	1	2	3	4	5	6
p_i	0.05	0.15	0.1	0.4	0.2	0.1
$g(x, \xi_i)$	10	8	6	3	2	-2

Then

- $\text{CVaR}_{0.98}\left(g(x, \xi); \hat{P}_N\right) = (10 \times 0.02)/0.02 = 10.$
- $\text{CVaR}_{0.95}\left(g(x, \xi); \hat{P}_N\right) = (10 \times 0.05)/0.05 = 10.$
- $\text{CVaR}_{0.85}\left(g(x, \xi); \hat{P}_N\right) = (10 \times 0.05 + 8 \times 0.1)/0.15 = 8.666\cdots.$
- $\text{CVaR}_{0.8}\left(g(x, \xi); \hat{P}_N\right) = (10 \times 0.05 + 8 \times 0.15)/0.2 = 8.5.$
- $\text{CVaR}_{0.7}\left(g(x, \xi); \hat{P}_N\right) = (10 \times 0.05 + 8 \times 0.15 + 6 \times 0.1)/0.3 = 7.666\cdots.$
- $\text{CVaR}_{0.6}\left(g(x, \xi); \hat{P}_N\right) = (10 \times 0.05 + 8 \times 0.15 + 6 \times 0.1 + 3 \times 0.1)/0.4 = 6.5.$

More formally, the greedy algorithm is described as follows.

1. Order the scenarios so that
$$g(x, \xi_{\sigma(1)}) \geq \cdots \geq g(x, \xi_{\sigma(N)}).$$
   Note that the ordering σ depends on the decision x.
2. Find the largest k for which
$$\sum_{i=1}^{k} p_{\sigma(i)} \leq 1 - \alpha.$$
   Note that
$$\sum_{i=1}^{k+1} p_{\sigma(i)} > 1 - \alpha$$
   by definition.

3. Compute
$$V = \sum_{i=1}^{k} p_{\sigma(i)} p_{\sigma(i)} \cdot g(x, \xi_{\sigma(i)}) + \left(1 - \alpha - \sum_{i=1}^{k} p_{\sigma(i)}\right) g(x, \xi_{\sigma(k+1)}).$$
4. Then
$$\mathrm{CVaR}_\alpha \left(g(x,\xi); \hat{P}_N\right) = \frac{V}{1-\alpha}.$$

Why is it called "conditional" value-at-risk? An intuition for this is as follows. If
$$\mathbb{P}_{\xi \sim \hat{P}_N} \left[g(x,\xi) > \mathrm{VaR}_\alpha \left(g(x,\xi); \hat{P}_N\right)\right] = 1 - \alpha$$
then
$$\mathrm{CVaR}_\alpha \left(g(x,\xi); \hat{P}_N\right) = \mathbb{E}_{\xi \sim \hat{P}_N} \left[g(x,\xi) \mid g(x,\xi) > \mathrm{VaR}_\alpha \left(g(x,\xi); \hat{P}_N\right)\right].$$

3 Modeling with CVaR

To model CVaR, we take the dual of the linear program. By strong duality, we have
$$\begin{aligned} \mathrm{CVaR}_\alpha \left(g(x,\xi); \hat{P}_N\right) \quad := \quad \min \quad & t + \frac{1}{1-\alpha} \sum_{i \in [N]} p_i r_i \\ \text{s.t.} \quad & r \geq 0 \\ & t + r_i \geq g(x, \xi_i), \quad i \in [N]. \end{aligned}$$

Then it follows that
$$\mathrm{CVaR}_\alpha \left(g(x,\xi); \hat{P}_N\right) \leq 0$$
is equivalent to the constraints
$$\begin{aligned} t + \frac{1}{1-\alpha} \sum_{i \in [N]} p_i r_i \leq 0 \\ r \geq 0 \\ t + r_i \geq g(x, \xi_i), \quad i \in [N]. \end{aligned}$$

Therefore, if each $g(x, \xi_i)$ is linearly representable, then $\mathrm{CVaR}_\alpha \left(g(x,\xi); \hat{P}_N\right) \leq 0$ is also linearly representable. In summary,
$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & \mathrm{CVaR}_\alpha \left(g(x,\xi); \hat{P}_N\right) \leq 0 \\ & x \in \mathcal{X} \end{aligned}$$
is equivalent to
$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & t + \frac{1}{1-\alpha} \sum_{i \in [N]} p_i r_i \leq 0 \\ & r \geq 0 \\ & t + r_i \geq g(x, \xi_i), \quad i \in [N]. \end{aligned}$$