A reverse Expected Shortfall/CVaR optimization formula

Zhanyi Jiao

Department of Statistics and Actuarial Science
University of Waterloo

25th International Congress on IME

July 13, 2022

What this talk is about

- Derive a reverse Expected Shortfall optimization formula.
- Compare the symmetries between ES optimization formula and the reverse one.
- Provide applications on worst-case risk under model uncertainty.
- Develop further theoretical results on reverse ES optimization formula
  - Reverse optimized certainty equivalents (OCE) formula.
  - Related Fenchel-Legendre transforms.

Contents

Preliminary

- $(\Omega, \mathcal{F}, \mathbb{P})$ atomless probability space.
- Let $\mathcal{X}$ be the set of integrable random variable, and X be the random loss.
- Left-quantile: $\mathrm{VaR}_\alpha^-(X) = \inf\{t \in \mathbb{R} : \mathbb{P}(X \le t) \ge \alpha\}$;
- Right-quantile: $\mathrm{VaR}_\alpha^+(X) = \inf\{t \in \mathbb{R} : \mathbb{P}(X \le t) > \alpha\}$. [1]
- Expected shortfall: $\mathrm{ES}_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_\beta^- \,\mathrm{d}\beta$. [2]
- Left-Expected shortfall: $\mathrm{ES}_\alpha^-(X) = \frac{1}{\alpha} \int_0^\alpha \mathrm{VaR}_\beta^-(X)\,\mathrm{d}\beta$

[1] $\mathrm{VaR}_0^-(X) = -\infty$ and $\mathrm{VaR}_1^+(X) = \infty$.
[2] $\mathrm{ES}_1(X) = \mathrm{VaR}_1^-(X)$.

ES/CVaR optimization formula

Theorem 1 (Rockafella and Uryasev, 2002)

For $X \in \mathcal{X}$ and $\alpha \in (0,1)$, it holds

$$\mathrm{ES}_\alpha(X) = \min_{t \in \mathbb{R}} \left\{ t + \frac{1}{1-\alpha}\mathbb{E}[(X-t)_+] \right\}, \quad (1)$$

and the set of minimizers for (1) is $[\mathrm{VaR}_\alpha^-(X), \mathrm{VaR}_\alpha^+(X)]$.

TITLE	CITED BY	YEAR
Optimization of conditional value-at-risk RT Rockafellar, S Uryasev Journal of risk 2, 21-42	7294	2000
Conditional value-at-risk for general loss distributions RT Rockafellar, S Uryasev Journal of banking & finance 26 (7), 1443-1471	4530	2002

[1] Source: https://scholar.google.ca/citations?user=Uwg1zpkAAAAJ&hl=en&oi=sra

Why is ES optimization formula such influential

- Efficient optimization techniques for portfolio allocation do not allow for direct controlling of percentiles of distribution.
    - ES optimization formula is convex w.r.t. t
    - It is possible to transform the problem into a linear program and find the global solution.
- It is difficult to handle ES_α because of VaR_α involved in its definition.
    - Minimizing the function w.r.t. t gives ES.
    - VaR is the minimum point of this function w.r.t. t.

Contents

Reverse ES optimization formula

Theorem 2 (Reverse ES optimization formula)

For $X \in \mathcal{X}$ and $t \in \mathbb{R}$, it holds

$$\mathbb{E}[(X-t)_+] = \max_{\alpha \in [0,1]} \left\{(1-\alpha)\left(\mathrm{ES}_\alpha(X) - t\right)\right\}, \tag{2}$$

and the set of maximizers for (2) is $[\mathbb{P}(X < t), \mathbb{P}(X \le t)]$.

Corollary 1

For $t \in \mathbb{R}$ and $X \in \mathcal{X}$, it holds

$$\mathbb{E}[X \wedge t] = \min_{\alpha \in [0,1]} \left\{\alpha \mathrm{ES}^-_\alpha(X) + (1-\alpha)t\right\}, \tag{3}$$

and the set of minimizers for (3) is $[\mathbb{P}(X < t), \mathbb{P}(X \le t)]$.

Symmetries between two formulas

(1) **Functional properties on $\mathcal{X}$**

- For a fixed $t \in \mathbb{R}$, the mapping $X \mapsto \mathbb{E}[(X-t)_+]$ is linear in the distribution of X and convex in the quantile of X.
- For a fixed $\alpha \in (0,1)$, the mapping $X \mapsto \mathrm{ES}_\alpha(X)$ is linear in the quantile of X and concave in the distribution of X.

(2) **Optimization problems**

- In the minimization (1) over $t \in \mathbb{R}$, the function $t \mapsto t + \frac{1}{1-\alpha}\mathbb{E}[(X-t)_+]$ is convex in t.
- In the maximization (2) over $\alpha \in [0,1]$, the function $\alpha \mapsto (1-\alpha)(\mathrm{ES}_\alpha(X) - t)$ is concave in α.

(3) **Solutions to the optimization problems**

(4) **Parametric forms**

Symmetries between two formulas (cont.)

Theorem 1 (ES/CVaR optimization formula)

For $X \in \mathcal{X}$ and $\alpha \in (0,1)$, it holds

$$\mathrm{ES}_\alpha(X) = \min_{t \in \mathbb{R}} \left\{ t + \frac{1}{1-\alpha} \mathbb{E}[(X-t)_+] \right\},$$

and the set of minimizers is $[\mathrm{VaR}_\alpha^-(X), \mathrm{VaR}_\alpha^+(X)]$.

Theorem 2 (Reverse ES optimization formula)

For $X \in \mathcal{X}$ and $t \in \mathbb{R}$, it holds

$$\mathbb{E}[(X-t)_+] = \max_{\alpha \in [0,1]} \left\{ (1-\alpha)\left(\mathrm{ES}_\alpha(X) - t\right) \right\},$$

and the set of maximizers is $[\mathbb{P}(X < t), \mathbb{P}(X \le t)]$.

Contents

Worst-case mean excess loss

Suppose that there is uncertainty about a random vector $\mathbf{X}$, assumed to be in a set $\mathcal{U}$, and $f : \mathbb{R}^d \to \mathbb{R}$ is a loss function. By the reverse ES optimization formula, the worst-case mean excess loss is computed by

$$\sup_{\mathbf{X}\in\mathcal{U}} \mathbb{E}[(f(\mathbf{X}) - t)_+] = \max_{\alpha\in[0,1]} \left\{(1-\alpha)\left(\sup_{\mathbf{X}\in\mathcal{U}} \mathrm{ES}_\alpha(f(\mathbf{X})) - t\right)\right\}.$$

Uncertainty set induced by moment information

- Uncertainty set induced by mean and a higher moment: for $p > 1$, $m \in \mathbb{R}$ and $v \geq 0$, denote by

$$\mathcal{L}^p(m, v) = \{X \in \mathcal{X} : \mathbb{E}[X] = m,\ \mathbb{E}[|X - m|^p] \leq v^p\}.$$

- The problem of $\sup_{X \in \mathcal{L}^p(m,v)} \rho(X)$ is better suited for $\rho = \mathrm{ES}_\alpha$ (see e.g., (Pesenti et al, 2020))
    - $\sup_{X \in \mathcal{L}^p(m,v)} \rho(X) = m + v \sup_{X \in \mathcal{L}^p(0,1)} \rho(X)$.
    - $\sup_{X \in \mathcal{L}^p(m,v)} \mathrm{ES}_\alpha(X) = m + v\alpha(\alpha^p(1-\alpha) + (1-\alpha)^p\alpha)^{-1/p}$

$\not\Rightarrow$ mean excess loss $\rho : X \mapsto \mathbb{E}[(X - t)_+]$.

Uncertainty set induced by moment information (cont.)

Proposition 3

For $p > 1$, $m, t \in \mathbb{R}$ and $v \geq 0$, we have

$$\sup_{X \in \mathcal{L}^p(m,v)} \mathbb{E}[(X-t)_+] = \max_{\alpha \in [0,1]} \Big\{(1-\alpha)(m-t) + v\left((1-\alpha)^{1-p} + \alpha^{1-p}\right)^{-1/p}\Big\}$$

In the most popular case $p = 2$, Proposition 3 gives

$$\sup_{X \in \mathcal{L}^2(m,v)} \mathbb{E}[(X-t)_+] = \frac{1}{2}\left(m - t + \sqrt{v^2 + (m-t)^2}\right),$$

which coincides with Jagannathan (1977).

Numerical example

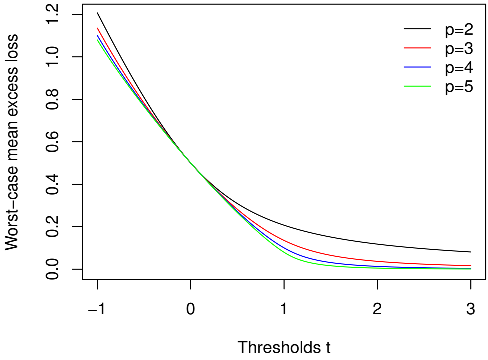


Figure: Worst-case mean excess loss with moment conditions in $\mathcal{L}^p(0,1)$: $\mathcal{L}^p(0,1) = \{X \in \mathcal{X} : \mathbb{E}[X] = 0,\ \mathbb{E}[|X|^p] \leq 1\}$

Uncertainty set induced by Wasserstein metrics

- Wasserstein metric of order $p \geq 1$:

$$W_p(F,G) = \inf_{X\sim F,\ Y\sim G} \left(\mathbb{E}[|X-Y|^p]\right)^{1/p}$$
$$= \left(\int_0^1 |F^{-1}(x) - G^{-1}(x)|^p \,\mathrm{d}x\right)^{1/p}.$$

- Wasserstein ball around X:

$$\{Y : W_p(F_X, F_Y) \leq \delta\}.$$

- Worst-case risk measure $\rho : \mathcal{X} \to \mathbb{R}$:

$$\sup\{\rho(Y) : W_p(F_X, F_Y) \leq \delta\}.$$

Uncertainty set induced by Wasserstein metrics (cont.)

Proposition 4

For $t \in \mathbb{R}$, $p \geq 1$, $\delta \geq 0$ and $X \in \mathcal{X}$, we have

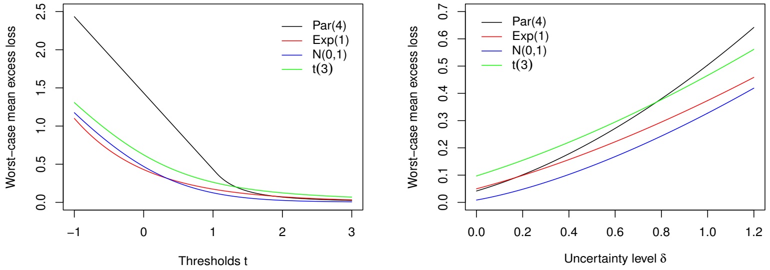
$$\sup\left\{\mathbb{E}[(Y-t)_+] : W_p(F_X, F_Y) \leq \delta\right\} = \max_{\alpha \in [0,1]} \left\{(1-\alpha)(\mathrm{ES}_\alpha(X) - t) + \delta(1-\alpha)^{1-1/p}\right\}.$$

Recall the reverse ES optimization formula:

$$\mathbb{E}[(X-t)_+] = \max_{\alpha \in [0,1]} \left\{(1-\alpha)\left(\mathrm{ES}_\alpha(X) - t\right)\right\}$$

The extra term $\delta(1-\alpha)^{1-1/p}$ compensates for model uncertainty.

Numerical example



(a) Changes with t (fixed $\delta = 0.1$) (b) Changes with δ (fixed $t = 2$)

Figure: Worst-case mean excess loss with Wasserstein uncertainty

Empirical analysis with insurance data

- `CASdatasets`: Normalized hurricane damages (`ushurricane`, 1900-2005); Normalized French commercial fire losses (`frecomfire`, 1982-1996) with same observations.
- Calculate the worst-case value of mean excess loss under uncertainty governed by the Wasserstein metric with $p = 2$.
- Fit the data with lognormal, Gamma and Weibull distributions as benchmark distributions.
- Let the uncertainty level δ vary in $[\delta_0, 2\delta_0]$, where δ_0 is the Wasserstein distance between the fitted distribution and the empirical distribution.
    - δ too large $\Rightarrow$ data become less relevant
    - δ too small $\Rightarrow$ lose the desired robustness.

Empirical analysis (fixed t)

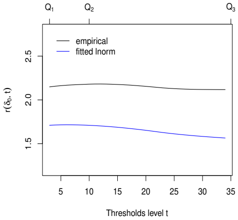
The ratio $r(\delta, t)$ of the worst-case mean excess loss to that of the benchmark distribution, defined by

$$r(\delta, t) = \frac{\sup\{\mathbb{E}[(Y-t)_+] : W_2(F_X, F_Y) \le \delta\}}{\mathbb{E}[(X-t)_+]}.$$

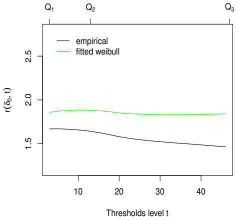
		δ_0	$1.2\delta_0$	$1.4\delta_0$	$1.6\delta_0$	$1.8\delta_0$	$2\delta_0$
Hurricane	Lognormal	1.708	1.839	1.985	2.132	2.279	2.425
	Weibull	1.853	2.012	2.193	2.352	2.534	2.715
	Gamma	1.964	2.149	2.334	2.539	2.724	2.950
Fire	Lognormal	1.358	1.431	1.505	1.582	1.657	1.735
	Weibull	1.400	1.481	1.564	1.649	1.733	1.819
	Gamma	1.456	1.548	1.644	1.740	1.837	1.937

Table: Values of $r(\delta, t_0)$ for the hurricane loss and the fire loss datasets.

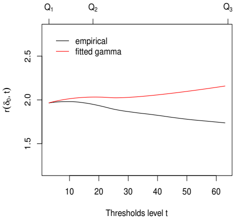
Empirical analysis (fixed δ_0)



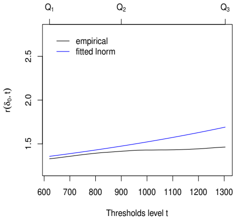
(a) Lognormal

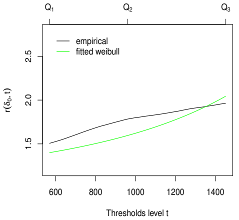


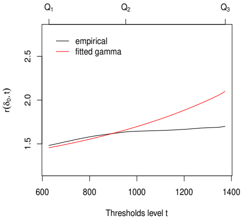
(b) Weibull



(c) Gamma







Contents

Optimized certainty equivalents (OCE)

Let V be the set of increasing and convex functions $v:\mathbb{R}\to\mathbb{R}$ satisfying $v(0)=0$, $\bar{v}\geq 1$ and $\lim_{t\to\infty} v'_+(-t)=0$ where $\bar{v}=\sup_{x\in\mathbb{R}} v'_+(x)$ and v'_+ is the right derivative of v. An OCE is a risk measure R defined by

$$R(X)=\inf_{t\in\mathbb{R}}\left\{t+\mathbb{E}[v(X-t)]\right\}, \qquad X\in\mathcal{X}_B.$$

($v=x_+/(1-\alpha)\Rightarrow$ ES optimization formula.)

Theorem 3 (Reverse OCE optimization formula)

For $X\in\mathcal{X}_B$, $t\in\mathbb{R}$ and $v\in V$, it holds

$$\mathbb{E}[v(X-t)]=\sup_{\beta\in(0,\bar{v}]}\left\{\beta(R^v_\beta(X)-t)\right\}.$$

where $R^v_\beta(X)=\inf_{t\in\mathbb{R}}\left\{t+\frac{1}{\beta}\mathbb{E}[v(X-t)]\right\}$.

($v=x_+\Rightarrow$ Reverse ES optimization formula.)

Related Fenchel-Legendre transforms

For a convex function $f : \mathbb{R} \to \mathbb{R}$, its Legendre-Fenchel transform is the function f^* on $\mathbb{R}$ defined by

$$f^*(\beta) = \sup_{t \in \mathbb{R}} \{t\beta - f(t)\}, \quad \beta \in \mathbb{R},$$

where β may be constrained to a subset of $\mathbb{R}$ such that f^* is real.

Related Fenchel-Legendre transforms

Proposition 5

(i) The Fenchel-Legendre transform of the convex quantile-based function $f_1(\alpha) = -(1-\alpha)\mathrm{ES}_\alpha(X)$, is given by

$$f_1^*(t) = \max_{\alpha \in [0,1]} \{\alpha t - f_1(\alpha)\} = \mathbb{E}[X \vee t].$$

(ii) The Fenchel-Legendre transform of the convex quantile-based function $f_2(\alpha) = \alpha \mathrm{ES}_\alpha^-(X)$, is given by

$$f_2^*(t) = \max_{\alpha \in [0,1]} \{\alpha t - f_2(\alpha)\} = \mathbb{E}[(t - X)_+].$$

Moreover, the set of maximizers for both maximization problems is $[\mathbb{P}(X < t), \mathbb{P}(X \leq t)]$.

Conclusion

- ES optimization formula v.s Reverse ES optimization formula
- Worst-case risk under model uncertainty
  - Uncertainty set induced by moments information.
  - Uncertainty set induced by Wasserstein metrics.
- Other related applications
  - Reverse OCE optimization formula.
  - Related Fenchel-Legendre transforms.

Reference

Guan, Y., Jiao, Z., & Wang, R. (2022)
A reverse Expected Shortfall optimization formula.
arXiv preprint arXiv:2203.02599

Jagannathan, R. (1977)
Minimax procedure for a class of linear programs under uncertainty.
Operations Research 25(1), 173-177.

Liu, F., Cai, J., Lemieux, C., & Wang, R. (2020)
Convex risk functionals: Representation and applications.
Insurance: Mathematics and Economics 90, 66-79.

Pesenti, S. M., Wang, Q., & Wang, R. (2020)
Optimizing distortion riskmetrics with distributional uncertainty
Available at SSRN 3728638.

Rockafellar, R. T., & Uryasev, S. (2002)
Conditional value-at-risk for general loss distributions.
Journal of banking & finance 26(7), 1443-1471.

Thank you!

Two important applications (cont.)

Proposition 1 (ES optimization)

Minimizing the expected shortfall $\phi_\alpha(\omega)$ with respect to ω is equivalent to minimizing $F_\alpha(\omega, t)$ over all $(\omega, t) \in W \times \mathbb{R}$

$$\min_{\omega \in W} \phi_\alpha(\omega) = \min_{(\omega,t) \in W \times \mathbb{R}} F_\alpha(\omega, t).$$

Proposition 2 (ES constraints)

Minimizing $g(\omega)$ over $\omega \in W$ satisfying $\phi_{\alpha_i}(\omega) \leq c_i$ for $i = 1, \dots, n$,

$$\Updownarrow$$

Minimizing $g(\omega)$ over $(\omega, \alpha_1, \dots, \alpha_n) \in W \times \mathbb{R} \times \cdots \times \mathbb{R}$ satisfying $F_\alpha(\omega, t) \leq c_i$