Risk Quadrangle and Applications in Statistics, Data Mining and Portfolio Optimization

Stan Uryasev

jointly with **Anton Malandii** and **Cheng Peng**

Stony Brook University

IAQF & Thalesians Seminar Series

March 5, 2024

SPIR Internship Projects for Financial Firms in NY State

NY State provides matching funds for internship SPIR projects. Stony Brook University's Quantitative Finance Program within the Department of Applied Mathematics and Statistics is seeking for industrial partners. This opportunity is available only to companies with a presence in NY State. Drop an email to Prof. Stan Uryasev <**Stanislav.Uryasev@stonybrook.edu**> or Prof. Pawel Polak <**Pawel.Polak@stonybrook.edu>** to learn more.

Key Advantages of the SPIR Collaboration:

1. **50% Time Employment**: Engage with our graduate students at low cost due to 50% position for an entire year!

2. **Matching State Funding:** The NY state may match your funding, amplifying the impact of your investment.

3. **Hassle-Free Paperwork:** SPIR coordinator handles the paperwork.

4. **Legal Status for International Students:** No worry about visas or special programs (like OPT or CPT). Work done at your company is considered a part of study and does not require work permit or special visa.

5. **Interaction with Faculty:** Benefit from collaboration with university faculty. You can onboard a professor directly into your project, if of interest.

6. **Basic Setting:** MS student in the Quantitative Finance program studies for 2 years. The first year covers fundamental courses at the university. The second year a student works 50% time for the company and 50% studies at the university.

Link describing the application process:
https://www.stonybrook.edu/commcms/spir/how-to-apply/index.php

Outline

The Fundamental Risk Quadrangle

Developed by [Rockafellar & Uryasev, 2013], [Rockafellar & Royset, 2015]

X = **"cost or loss":** quantity to be minimized or kept below given levels

	risk $\mathcal{R} \longleftrightarrow \mathcal{D}$ deviation	
optimization	$\uparrow\downarrow \quad \mathcal{S} \quad \downarrow\uparrow$	*estimation*
	regret $\mathcal{V} \longleftrightarrow \mathcal{E}$ error	

$\mathcal{R}(X)$ elicits the "level of cost" in X to use in making comparisons

$\mathcal{D}(X)$ measures the "nonconstancy" in X as its uncertainty

$\mathcal{E}(X)$ measures the "nonzeroness" in X for use in estimates

$\mathcal{V}(X)$ measures the "regret" in facing the mix of outcomes of X

$\mathcal{S}(X)$ is a "statistic" associated with X through $\mathcal{E}$ and $\mathcal{V}$

Relationship Formulae

Quartet $(\mathcal{R}, \mathcal{D}, \mathcal{V}, \mathcal{E})$ of risk, deviation, regret, and error is called a risk quadrangle if it satisfies

(1) **Error Projection:** $\mathcal{D}(X) = \min_C \{ \mathcal{E}(X - C) \}$

(2) **Certainty Equivalence:** $\mathcal{R}(X) = \min_C \{ C + \mathcal{V}(X - C) \}$

(3) **Centerness:** $\mathcal{R}(X) = \mathcal{D}(X) + \mathbb{E}[X] \quad \mathcal{V}(X) = \mathcal{E}(X) + \mathbb{E}[X]$

Quartet $(\mathcal{R}, \mathcal{D}, \mathcal{V}, \mathcal{E})$ is bound by the statistic $\mathcal{S}(X)$

$$\mathcal{S}(X) = \operatorname*{argmin}_C \{ \mathcal{E}(X - C) \} = \operatorname*{argmin}_C \{ C + \mathcal{V}(X - C) \}$$

Stochastic Framework

Space of elementary outcomes: Ω with elements $\omega \in \Omega$ ("scenarios")

$\mathcal{A}_0 = \sigma$-algebra of subsets of Ω $\quad$ $\mathbb{P}_0 =$ probability measure on $\mathcal{A}$

Random variables: functions $X : \Omega \to \mathbb{R}$

$$\mathbb{E}[X] = \int_\Omega X(\omega) d\mathbb{P}_0(\omega) \qquad \sigma^2(X) = \mathbb{E}[(X - \mathbb{E}X)^2]$$

Function space setting: $X \in \mathcal{L}^2 := \mathcal{L}^2(\Omega)$

Hilbert space of random variables with finite mean and variance

$$\langle X, Y \rangle = \mathbb{E}[XY] \qquad \|X\|^2 = \mathbb{E}[X^2]$$

α-**Quantile:** "value-at-risk" in finance

$q_\alpha(X) = \mathrm{VaR}_\alpha(X) =$ "$F_X^{-1}(x)$" $=$ generalized inverse of CDF

α-**Superquantile:** "conditional value-at-risk" in finance

$\bar{q}_\alpha(X) = \mathrm{CVaR}_\alpha(X) =$ "$\mathbb{E}[X | X \geq q_\alpha(X)]$" $= \frac{1}{1-\alpha} \int_\alpha^1 q_\beta(X) d\beta$

Measures of Risk, Regret, Error, and Deviation

$\mathcal{R}(X)$ quantifies the risk

(R1) $\mathcal{R}(C) = C$ (R2) convexity (R3) closedness
(R4) aversity: $\mathcal{R}(X) > \mathbb{E}[X]$ for nonconstant X

$\mathcal{V}(X)$ quantifies the anti-utility

(V1) $\mathcal{V}(0) = 0$ (V2) convexity (V3) closedness
(V4) aversity: $\mathcal{V}(X) > \mathbb{E}[X]$ for nonconstant X

$\mathcal{E}(X)$ generalizes norm $\|X\|$

(E1) $\mathcal{E}(0) = 0$ (E2) convexity (E3) closedness
(E4) positiveness: $\mathcal{E}(X) > 0$ for nonzero X

$\mathcal{D}(X)$ generalizes standard deviation $\sigma(X)$

(D1) $\mathcal{D}(C) = 0$ (D2) convexity (D3) closedness
(D4) positiveness: $\mathcal{D}(X) > 0$ for nonconstant X

Example: Quantile Quadrangle

Quantile $q_\alpha(X)$ is a characteristic of a random variable X calculated by minimizing the Koenker-Bassett (KB) error $\mathcal{E}_\alpha(X)$ [Koenker & Bassett, 1978]

$$\mathcal{E}_\alpha(X) = \mathbb{E}[\alpha X_+ + (1-\alpha)X_-] \qquad q_\alpha(X) = \operatorname*{argmin}_{C\in\mathbb{R}} \{\mathcal{E}_\alpha(X-C)\}$$

where $\alpha \in (0,1)$ is a parameter defining asymmetry and $X_\pm = \max\{0, \pm X\}$

Quantile Quadrangle [Rockafellar & Uryasev, 2013]

$$\begin{aligned}
\mathcal{S}_\alpha(X) &= q_\alpha(X) = \textit{quantile}\\
\mathcal{R}_\alpha(X) &= \bar{q}_\alpha(X) = \textit{CVaR (superquantile)}\\
\mathcal{D}_\alpha(X) &= \bar{q}_\alpha(X - \mathbb{E}[X]) = \textit{CVaR deviation}\\
\mathcal{V}_\alpha(X) &= \frac{1}{1-\alpha}\mathbb{E}[X_+] = \textit{regret penalty}\\
\mathcal{E}_\alpha(X) &= \mathbb{E}\left[\frac{\alpha}{1-\alpha}X_+ + X_-\right] = \textit{normalized KB error}
\end{aligned}$$

$$\bar{q}_\alpha(X) = \min_C \quad C + \mathcal{V}_\alpha(X-C) = \min_C \quad C + \frac{1}{1-\alpha}\mathbb{E}[(X-C)_+]$$

Example: Applications in Stochastic Optimization

- $\boldsymbol{x}$ = decision vector $\quad \boldsymbol{x} \in \mathcal{X}$ = feasible closed convex set $\quad \mathcal{X} \subseteq \mathbb{R}^d$
- $L(\boldsymbol{x})$ = parametric random variable $\quad L(\cdot)$ = convex or linear function

Minimizing Risk through Regret via Certainty Equivalence

$$\min_{\boldsymbol{x}\in\mathcal{X}} \mathcal{R}(L(\boldsymbol{x})) = \min_{\boldsymbol{x}\in\mathcal{X}} \min_{C} \; C + \mathcal{V}(L(\boldsymbol{x}) - C)$$

$$= \min_{\boldsymbol{x}\in\mathcal{X}, C} \; C + \mathcal{V}(L(\boldsymbol{x}) - C)$$

$\Longrightarrow$ **efficient optimization algorithms!**

CVaR Optimization [Rockafellar & Uryasev, 2000]

$$\min_{\boldsymbol{x}\in\mathcal{X}} \bar{q}_\alpha(L(\boldsymbol{x})) = \min_{\boldsymbol{x}\in\mathcal{X}, C} C + \frac{1}{1-\alpha}\mathbb{E}[(L(\boldsymbol{x}) - C)_+]$$

Expectile Optimization [Malandii, Kuzmenko, Uryasev, 2024]

$$\min_{\boldsymbol{x}\in\mathcal{X}} e_K(L(\boldsymbol{x})) = \min_{\boldsymbol{x}\in\mathcal{X}, C} C + \left(\mathbb{E}\left[L(\boldsymbol{x}) - C + \frac{1}{K}(L(\boldsymbol{x}) - C)_+\right]\right)_+$$

Example: Applications in Statistical Estimation

- Y = regressant (target variable)
- $\boldsymbol{X} = (X_1, \ldots, X_d)$ = regressors (factors)
- $Z_f = Y - f(\boldsymbol{X}) - C$ = residual
- $\bar{Z}_f = Y - f(\boldsymbol{X})$ = residual w/o intercept

Generalized Regression

$$\min_{f \in \mathcal{F}, C} \mathcal{E}(Z_f) = \min_{f \in \mathcal{F}} \underbrace{\min_{C} \ \mathcal{E}(\bar{Z}_f - C)}_{error\ \ projection}$$

$$= \min_{f \in \mathcal{F}} \ \mathcal{D}(\bar{Z}_f) \text{ and } C \in \mathcal{S}(\bar{Z}_f) = \operatorname*{argmin}_{C} \ \mathcal{E}(\bar{Z}_f - C)$$

Generalized Regression Theorem

Regression as Error Minimization

$$\min_{f\in\mathcal{F},C\in\mathbb{R}} \mathcal{E}(Z_f) \qquad (1)$$

Regression as Deviation Minimization

$$\min_{f\in\mathcal{F}} \mathcal{D}(\bar{Z}_f) \text{ and } C \in \mathcal{S}(\bar{Z}_f) \qquad (2)$$

Theorem [Rockafellar, Uryasev, Zabarankin, 2008], [Rockafellar & Uryasev, 2013]

$\mathcal{E} =$ regular error of expectation type or with representable risk identifier (in linear case)
$\mathcal{D} =$ regular deviation corresponding to $\mathcal{E}$
$\mathcal{S} =$ statistic corresponding to $\mathcal{E}$
$\mathcal{F} =$ class of functions $f : \mathbb{R}^d \to \mathbb{R}$ such that there exists

$$f(\boldsymbol{x}) + C \in \mathcal{S}(Y|\boldsymbol{x}) \text{ a.s. for } \boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^d$$
$$\text{where } Y|\boldsymbol{x} = Y_{\boldsymbol{X}=\boldsymbol{x}} \text{ (conditional distribution)}$$

Then

- (1) $\iff$ (2)
- optimal $f, C \implies f(\boldsymbol{X}) + C \in \mathcal{S}(Y|\boldsymbol{X})$ a.s.

Examples of Regressions

Error Minimization

Least Squares

$$\min_{f,C} \mathbb{E}[Z_f^2]$$

Expectile Regression $q \in (0,1)$

$$\min_{f,C} \mathbb{E}[q(Z_f)_+^2 + (1-q)(Z_f)_-^2]$$

Median Regression

$$\min_{f,C} \mathbb{E}[|Z_f|]$$

Quantile Regression $\alpha \in (0,1)$

$$\min_{f,C} \mathbb{E}[\tfrac{\alpha}{1-\alpha}(Z_f)_+ + (Z_f)_-]$$

Deviation Minimization

Least Squares

$$\min_{f} \sigma^2(\bar{Z}_f) \text{ and } C = \mathbb{E}[\bar{Z}_f]$$

Expectile Regression

$$\min_{f} \sigma_q^2(\bar{Z}_f) \text{ and } C = e_q[\bar{Z}_f]$$

Median Regression $\alpha = 0.5$

$$\min_{f} \bar{q}_\alpha(\bar{Z}_f - \mathbb{E}[\bar{Z}_f]) \text{ and } C = q_\alpha(\bar{Z}_f)$$

Quantile Regression

$$\min_{f} \bar{q}_\alpha(\bar{Z}_f - \mathbb{E}[\bar{Z}_f]) \text{ and } C = q_\alpha(\bar{Z}_f)$$

ε-SVR: Discrete Case

Given: training data $X^\ell = (\boldsymbol{x}_i, y_i)_{i=1}^{\ell}$, $\boldsymbol{x}_i \in \mathbb{R}^n$, $y_i \in \mathbb{R}$, $\boldsymbol{y} = (y_1, \dots, y_\ell)$

Find: hyperplane $f_{\boldsymbol{w},b}(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{x} + b$, $(\boldsymbol{w}, b) \in \mathbb{R}^{n+1}$ that optimally fits the data

Let $z = z(\boldsymbol{w}, b)$ be a random variable taking with equal probabilities $p = 1/\ell$ the components $(y_i - f_{\boldsymbol{w},b}(\boldsymbol{x}_i))_{i=1}^{\ell}$, $C > 0$, $\nu \in (0, 1]$

Denote $\mathcal{L}(\xi) = \max\{0, |\xi| - \varepsilon\} = [|\xi| - \varepsilon]_+ =$ "ε-insensitive loss function"

The ε-SVR [Vapnik, 1995]

$$\min_{\boldsymbol{w},b} \quad \mathbb{E}[\mathcal{L}(z(\boldsymbol{w}, b))] + \frac{1}{2C}\|\boldsymbol{w}\|_2^2$$

ν-SVR: Discrete Case

The ν-SVR [Schölkopf et al., 2000] (improvement of ε-SVR)

$$\min_{\boldsymbol{w},b,\varepsilon} \quad \nu\left(\varepsilon + \frac{1}{\nu}\mathbb{E}[\mathcal{L}(z(\boldsymbol{w},b))]\right) + \frac{1}{2C}\|\boldsymbol{w}\|_2^2$$

Let $\nu = 1-\alpha$. Consider

$$\min_{\varepsilon} \quad (1-\alpha)\left(\varepsilon + \frac{1}{1-\alpha}\mathbb{E}[|z(\boldsymbol{w},b)| - \varepsilon]_+\right) = (1-\alpha)\bar{q}_\alpha(|z(\boldsymbol{w},b)|)$$

$= \langle\langle z(\boldsymbol{w},b)\rangle\rangle_\alpha = \mathrm{CVaR\ norm}$ [Bertsimas et.al., 2014; Mafusalov & Uryasev, 2016]

The ν-SVR [Malandii & Uryasev, 2022]

$$\min_{\boldsymbol{w},b} \quad \langle\langle z(\boldsymbol{w},b)\rangle\rangle_\alpha + \frac{1}{2C}\|\boldsymbol{w}\|_2^2$$

The ν-SVR [Takeda et al., 2010]

$$\min_{\boldsymbol{w},b} \quad \mathrm{CVaR}_\alpha\left(\left|\boldsymbol{y} - C\frac{f(\boldsymbol{w},b)}{\|\boldsymbol{w}\|}\right|\right)$$

SVR: General Case

$\boldsymbol{X} = (X_1, \ldots, X_n)^T$ = vector of random variables (**regressors**)
Y = target random variable (**regressant**)
$Z(\boldsymbol{w}, b) = Y - (\boldsymbol{w}^T \boldsymbol{X} + b)$ = **residual**

The ν-SVR ($C > 0,\ \alpha \in [0,1)$)

$$\min_{\boldsymbol{w},b} \quad \langle\!\langle Z(\boldsymbol{w}, b) \rangle\!\rangle_\alpha + \frac{1}{2C} \|\boldsymbol{w}\|_2^2$$

CVaR Norm Quadrangle [Mafusalov & Uryasev, 2016]

$$\mathcal{S}_1(X) = \frac{1}{2}\left(q_{(1-\alpha)/2}(X) + q_{(1+\alpha)/2}(X)\right)$$
$$\mathcal{R}_1(X) = \frac{1}{2}\left((1+\alpha)\bar{q}_{(1-\alpha)/2}(X) + (1-\alpha)\bar{q}_{(1+\alpha)/2}(X)\right)$$
$$\mathcal{D}_1(X) = \frac{1}{2}\left((1+\alpha)\bar{q}_{(1-\alpha)/2}(X - \mathbb{E}X) + (1-\alpha)\bar{q}_{(1+\alpha)/2}(X - \mathbb{E}X)\right)$$
$$\mathcal{V}_1(X) = \langle\!\langle X \rangle\!\rangle_\alpha + \mathbb{E}X$$
$$\mathcal{E}_1(X) = \langle\!\langle X \rangle\!\rangle_\alpha$$

Quantile Symmetric Average Quadrangle

The ε-SVR ($C > 0,\ \varepsilon \geq 0$)

$$\min_{\boldsymbol{w},b} \quad \mathbb{E}[|Z(\boldsymbol{w}, b)| - \varepsilon]_+ + \frac{1}{2C} \|\boldsymbol{w}\|_2^2$$

$(\mathcal{R}_1, \mathcal{D}_1, \mathcal{V}_1, \mathcal{E}_1) =$ CVaR Norm Quadrangle with statistic $\mathcal{S}_1$

Quantile Symmetric Average Quadrangle [Malandii & Uryasev, 2022]

$$\mathcal{S}_2(X) = \bigcup_{\alpha \in \mathcal{A}_x} \mathcal{S}_1(X),$$

$$\mathcal{R}_2(X) = \mathcal{R}_1(X) - (1-\alpha)x, \quad \forall\, \alpha \in \mathcal{A}_x$$

$$\mathcal{D}_2(X) = \mathcal{D}_1(X) - (1-\alpha)x, \quad \forall\, \alpha \in \mathcal{A}_x$$

$$\mathcal{V}_2(X) = \mathbb{E}[|X| - x]_+ + \mathbb{E}X,$$

$$\mathcal{E}_2(X) = \mathbb{E}[|X| - x]_+ = \text{Vapnik error}$$

$$\mathcal{A}_x = \left\{ \alpha \in [0,1) \,\middle|\, \frac{1}{2}(q^-_{(1+\alpha)/2}(X) - q^-_{(1-\alpha)/2}(X)) \leq x \leq \frac{1}{2}(q^+_{(1+\alpha)/2}(X) - q^+_{(1-\alpha)/2}(X)) \right\}$$

Vapnik error is of expectation type $\implies$ estimation

Equivalence?

SVR: Equivalence

Equivalence: duality is the key

Proposition (The ε-SVR and ν-SVR equivalence) [Malandii & Uryasev, 2022]

1. Let $(\boldsymbol{w}^*, b^*)$ be an optimal solution of ν-SVR for some $\alpha \in [0, 1)$. Then $(\boldsymbol{w}^*, b^*)$ is also an optimal solution vector of ε-SVR for each $\varepsilon \in q_\alpha(|Z(\boldsymbol{w}^*, b^*)|)$
2. Let $(\boldsymbol{w}^*, b^*)$ be an optimal solution vector of ε-SVR for some $\varepsilon \geq 0$. Then $(\boldsymbol{w}^*, b^*)$ is also an optimal solution vector of ν-SVR for each $\alpha \in [\mathbb{P}(|Z(\boldsymbol{w}^*, b^*)| < \varepsilon), \mathbb{P}(|Z(\boldsymbol{w}^*, b^*)| \leq \varepsilon))$

Equivalence in discrete case was studied by [Chang & Lin, 2001]

Estimation: asymptotically unbiased estimator

$$f_{\boldsymbol{w}^*, b^*}(\boldsymbol{X}) \in \frac{1}{2}(q_{(1+\alpha)/2}(Y|\boldsymbol{X}) + q_{(1-\alpha)/2}(Y|\boldsymbol{X}))$$

boundary case: $x = 0 \iff \alpha = 0$

$$f_{\boldsymbol{w}^*, b^*}(\boldsymbol{X}) \in \text{med}(Y|\boldsymbol{X})$$

SVR as Deviation Minimization

$(\mathcal{S}_1, \mathcal{D}_1) \in$ **CVaR Norm Quadrangle**

$(\mathcal{S}_2, \mathcal{D}_2) \in$ **Quantile Symmetric Average Quadrangle**

The ε-SVR

$\min_{\boldsymbol{w}} \quad \mathcal{D}_2(\bar{Z}(\boldsymbol{w})) + \frac{1}{2C} \|\boldsymbol{w}\|_2^2$
and take $b^* \in \mathcal{S}_2(\bar{Z}(\boldsymbol{w}^*))$

The ν-SVR

$\min_{\boldsymbol{w}} \quad \mathcal{D}_1(\bar{Z}(\boldsymbol{w})) + \frac{1}{2C} \|\boldsymbol{w}\|_2^2$
and take $b^* \in \mathcal{S}_1(\bar{Z}(\boldsymbol{w}^*))$

The optimal intercept b^* is a known function (statistic) of an optimal w^*

SVR: Dual Formulation

Notation: $\mathbf{1}_\ell = (1, \ldots, 1)^T \in \mathbb{R}^\ell,\ \boldsymbol{y} = (y_1, \ldots, y_\ell)^T,\ \hat{\mathbf{X}} = (\boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_\ell^T) \in \mathbb{R}^{\ell \times n}$

Primal Problem

$$\mathbf{p}^\star := \min_{\boldsymbol{w}, b, \boldsymbol{z}} \quad \langle\langle \boldsymbol{z} \rangle\rangle_\alpha + \frac{1}{2C} \|\boldsymbol{w}\|_2^2$$

$$\text{s.t.} \quad \boldsymbol{z} = \boldsymbol{y} - \hat{\mathbf{X}}\boldsymbol{w} - \mathbf{1}_\ell b$$

Dual Problem [Malandii & Uryasev, 2022]

$$\mathbf{d}^\star := \max_{\boldsymbol{\mu}} \quad \boldsymbol{\mu}^T \boldsymbol{y} - \frac{1}{2} \boldsymbol{\mu}^T \hat{\mathbf{X}} \hat{\mathbf{X}}^T \boldsymbol{\mu}$$

$$\text{s.t.} \quad \|\boldsymbol{\mu}\|_1 \le C(1 - \alpha), \quad \|\boldsymbol{\mu}\|_\infty \le \frac{C}{\ell}$$

$$\boldsymbol{\mu}^T \mathbf{1}_\ell = 0$$

Other variants were developed by [Vapnik, 1995], [Schölkopf et al., 2000]

Kernelization

The objective function in the dual problem depends on feature vectors only through their inner product

Kernel Function: $k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$

Kernel Matrix: $\mathbf{K} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)_{i,j=1}^{\ell}$

Kernelized Dual Problem

$$\max_{\boldsymbol{\mu}} \quad \boldsymbol{\mu}^T \boldsymbol{y} - \frac{1}{2}\boldsymbol{\mu}^T \mathbf{K} \boldsymbol{\mu}$$

$$\text{s.t.} \quad \|\boldsymbol{\mu}\|_1 \leq C(1-\alpha), \quad \|\boldsymbol{\mu}\|_\infty \leq \frac{C}{\ell}$$

$$\boldsymbol{\mu}^T \mathbf{1}_\ell = 0$$

Nonlinear Extension of SVR!

Closing Part II: Key Takeaways

- SVR is the asymptotically unbiased estimator of

$$\frac{1}{2}(q_{(1+\alpha)/2}(Y|\boldsymbol{X}) + q_{(1-\alpha)/2}(Y|\boldsymbol{X})$$

- Symmetric average of two conditional quantiles allows for estimating a wide range of statistics including the mean

- SVR minimizes CVaR norm with a regularization penalty

Expectile Quadrangle (Asymmetric Variance Error)

Expectile $e_q(X)$ is a characteristic of a random variable X calculated by minimizing the asymmetric variance error $\mathcal{E}_q(X)$ [Newey & Powell, 1987]

$$\mathcal{E}_q(X) = \mathbb{E}[qX_+^2 + (1-q)X_-^2] \qquad e_q(X) = \operatorname*{argmin}_{C\in\mathbb{R}} \{\mathcal{E}_q(X-C)\}$$

where $q \in (0,1)$ is a parameter defining asymmetry

Expectile Quadrangle [Kuzmenko, 2020], [Malandii, Kuzmenko, Uryasev, 2024]

$$\mathcal{S}_q(X) = e_q(X) = \textit{expectile}$$
$$\mathcal{R}_q(X) = \mathcal{D}_q(X) + \mathbb{E}[X] = \textit{asymmetric risk}$$
$$\mathcal{D}_q(X) = \mathbb{E}\left[q(X - e_q(X))_+^2 + (1-q)(X - e_q(X))_-^2\right] = \textit{asymmetric deviation}$$
$$\mathcal{V}_q(X) = \mathcal{E}_q(X) + \mathbb{E}[X] = \textit{asymmetric regret}$$
$$\mathcal{E}_q(X) = \mathbb{E}[qX_+^2 + (1-q)X_-^2] = \textit{asymmetric variance error}$$

Special Case: $e_{1/2}(X) = \mathbb{E}[X] \implies$ Mean Quadrangle

Expectile Quadrangle (Piecewise Linear Error)

First-order optimality condition for $\min_C \ \mathcal{E}_q(X - C)$:

$$C - \mathbb{E}[X] = \frac{1}{K}\mathbb{E}[(X - C)_+]$$

where $K = \frac{1-q}{2q-1}$, $q \in (1/2, 1)$ For $q \in (0, 1/2)$ $e_K(X) = -e_{-1-K}(-X)$

Piecewise Linear Error

$$e_K(X) = \mathop{\mathrm{argmin}}_C \ \max\left\{C - \mathbb{E}[X], \frac{1}{K}\mathbb{E}[(X - C)_+]\right\}$$

Expectile Quadrangle [Kuzmenko, 2020], [Malandii, Kuzmenko, Uryasev, 2024]

$\mathcal{S}_K(X) = e_K(X)$ = *expectile*

$\mathcal{R}_K(X) = e_K(X)$ = *expectile*

$\mathcal{D}_K(X) = e_K(X - \mathbb{E}[X])$ = *expectile deviation*

$\mathcal{V}_K(X) = \left(\mathbb{E}\left[X + \frac{1}{K}X_+\right]\right)_+$ = *piecewise linear regret*

$\mathcal{E}_K(X) = \max\left\{-\mathbb{E}[X], \frac{1}{K}\mathbb{E}[X_+]\right\}$ = *piecewise linear error*

Expectile Deviation vs CVaR Deviation

Kusuoka-type representation of expectile [Kuzmenko, 2020; Bellini, 2014]

$$e_K(X - \mathbb{E}[X]) = \sup_{F_X(\mathbb{E}[X])<\alpha<1} \frac{(1-\alpha)}{K + (1-\alpha)} \bar{q}_\alpha(X - \mathbb{E}[X])$$

Proposition (Quasiconcavity of Scaled CVaR Deviation) [Malandii et al., 2024]

The function

$$\frac{(1-\alpha)\bar{q}_\alpha(X - \mathbb{E}[X])}{K + (1-\alpha)}$$

is a strictly quasiconcave function of $\alpha \in (0,1)$ for $K > 0$

Optimal α^*

$$e_K(X) \in q_{\alpha^*}(X) \qquad \alpha^* = F_X(e_K(X))$$

Expectile Deviation vs CVaR Deviation: Implications

Theorem (Equivalence of Expectile and CVaR Deviation Minimization) [Malandii et al., 2024]

$\mathcal{X} \subseteq L^2(\Omega) =$ feasible closed convex set for

$$\min_{X \in \mathcal{X}} \quad \mathcal{D}_K(X) \quad K > 0$$

$$\min_{X \in \mathcal{X}} \quad \mathcal{D}_\alpha(X) \quad \alpha \in (0,1)$$

Then $\forall\, K > 0 \;\exists\, \alpha \in (0,1)$ such that

$$\operatorname*{argmin}_{X \in \mathcal{X}} \; \mathcal{D}_K(X) = \operatorname*{argmin}_{X \in \mathcal{X}} \; \mathcal{D}_\alpha(X)$$

Corollary (Expectile and CVaR Minimization with Expectation Constraint) [Malandii et al., 2024]

$\mathcal{X}_\mu = \{X \in \mathcal{X} : \mathbb{E}[X] = \mu \in \mathbb{R}\}$. Then $\forall\, K > 0 \;\exists\, \alpha \in (0,1)$ such that

$$\operatorname*{argmin}_{X \in \mathcal{X}} \; \mathcal{R}_K(X) = \operatorname*{argmin}_{X \in \mathcal{X}} \; \mathcal{R}_\alpha(X)$$

Expectile Optimization

- $\boldsymbol{x}$ = decision vector $\quad \boldsymbol{x} \in \mathbb{R}^d$
- $L(\boldsymbol{x})$ = parametric random variable
- $L(\cdot)$ = convex or linear function

Optimization Problem

$$\min_{\boldsymbol{x}} \quad e_K(L(\boldsymbol{x}))$$

Certainty Equivalence [Malandii, Kuzmenko, Uryasev, 2024]

$$\begin{aligned} e_K(L(\boldsymbol{x})) &= \min_C \quad C + \mathcal{V}_K(L(\boldsymbol{x}) - C) \\ &= \min_C \quad C + \left(\mathbb{E}[L(\boldsymbol{x}) - C + \frac{1}{K}(L(\boldsymbol{x}) - C)_+]\right)_+ \end{aligned}$$

Reduction to Convex Programming

$$\min_{\boldsymbol{x}} e_K(L(\boldsymbol{x})) = \min_{\boldsymbol{x}, C} C + \left(\mathbb{E}[L(\boldsymbol{x}) - C + \frac{1}{K}(L(\boldsymbol{x}) - C)_+]\right)_+$$

Reduction of expectile optimization to linear programming was developed by [Colombo, 2018]

Expectile and CVaR Portfolio Optimization: Equivalence

- $L(\boldsymbol{x}) = -(x_1 X_1 + \ldots + x_d X_d)$ $X_i =$ random asset returns
- $\mathcal{X}_{\boldsymbol{\mu}} = \{\boldsymbol{x} \in \mathbb{R}^d : \sum_{i=1}^{d} x_i = 1,\ \mathbb{E}[-L(\boldsymbol{x})] = \mu\}$

Then $\forall\ K > 0\ \exists\ \alpha \in (0,1)$ such that

$$\underset{\boldsymbol{x} \in \mathcal{X}_\mu}{\operatorname{argmin}}\ e_K(L(\boldsymbol{x})) = \underset{\boldsymbol{x} \in \mathcal{X}_\mu}{\operatorname{argmin}}\ \bar{q}_\alpha(L(\boldsymbol{x}))$$

Expectile portfolio optimization $\iff$ **CVaR portfolio optimization**

Expectile and Omega Portfolio Optimization: Equivalence

Omega Function

$$\Omega_B(X) = \frac{\mathbb{E}[X-B]_+}{\mathbb{E}[X-B]_-} = \frac{\text{expected over-performance}}{\text{expected under-performance}} \text{ relative to benchmark } B$$

Portfolio Optimization Equivalence [Wagner & Uryasev, 2019]

$\mathcal{X} = \{X \in \mathcal{L}^2 : \mathbb{E}[X] \geq \mu\}$ = feasible set of random variables for

$$\min_{X \in \mathcal{X}} e_K(X) \quad K < -1$$

$$\min_{X \in \mathcal{X}} \Omega_B(X) \quad B \in \mathbb{R}$$

- If $X^* \in \underset{X \in \mathcal{X}}{\operatorname{argmin}}\, e_K(X) \implies X^* \in \underset{X \in \mathcal{X}}{\operatorname{argmin}}\, \Omega_B(X) \quad B = -e_K(X^*)$
- If $X^* \in \underset{X \in \mathcal{X}}{\operatorname{argmin}}\, \Omega_B(X) \implies X^* \in \underset{X \in \mathcal{X}}{\operatorname{argmin}}\, e_K(X) \quad K = (1 + \Omega_B(X^*))^{-1}$

Expectile Estimation with Linear Regression

Given: Y = regressant $\quad \boldsymbol{X}$ = regressor

Denote: $Z(C_0, \boldsymbol{C}) = Y - C_0 - \boldsymbol{C}^\top \boldsymbol{X}$ = regression residual

Generalized Linear Regression Problem (GLRP)

$$\min_{C_0, \boldsymbol{C}} \quad \mathcal{E}(Z(C_0, \boldsymbol{C}))$$

Expectile Estimation with Linear Regression (Cont'd)

Asymmetric Variance (AV) Error

$$\min_{C_0, \boldsymbol{C}} \; \mathbb{E}\Big[q(Z(C_0, \boldsymbol{C}))_+^2 + (1-q)(Z(C_0, \boldsymbol{C}))_-^2\Big]$$

Piecewise Linear (PL) Error

$$\min_{C_0, \boldsymbol{C}} \max \left\{ -\mathbb{E}[Z(C_0, \boldsymbol{C})], \frac{1}{K}\mathbb{E}[(Z(C_0, \boldsymbol{C}))_+] \right\}$$

Normalized KB Error

$$\min_{C_0, \boldsymbol{C}} \; \mathbb{E}\left[\frac{\alpha}{1-\alpha}(Z(C_0, \boldsymbol{C}))_+ + (Z(C_0, \boldsymbol{C}))_-\right]$$

Equivalence of regressions with Piecewice Linear and Asymmetric Variance Errors

Theorem I [Malandii, Kuzmenko, Uryasev, 2024]

Let $X_i,\ \ Y,\ \varepsilon =$ random variables defined on $\mathcal{L}^2(\Omega) \quad i = 1,\dots,d$

$Y = a_0 + \boldsymbol{a}^\top \boldsymbol{X} + \varepsilon \quad a.s. =$ true regression law $\quad (a_0, \boldsymbol{a}) \in \mathbb{R}^{d+1}$

$(a)\ X_i \perp\!\!\!\perp \varepsilon \quad (b)\ e_K(\varepsilon) = 0 \quad (\perp\!\!\!\perp =$ independent$)$

Then

- $(a_0, \boldsymbol{a}) \in \underset{C_0, \boldsymbol{C}}{\operatorname{argmin}}\ \mathcal{E}_q(Z(C_0, \boldsymbol{C})) \quad q = \frac{K+1}{2K+1}$
- $(a_0, \boldsymbol{a}) \in \underset{C_0, \boldsymbol{C}}{\operatorname{argmin}}\ \mathcal{E}_K(Z(C_0, \boldsymbol{C}))$

Equivalence of Regressions II

Equivalence of regressions with Piecewice Linear and Asymmetric Variance Errors

Theorem II [Malandii, Kuzmenko, Uryasev, 2024]

Consider the Generalized Linear Regression Problem

S_{AV} = set of optimal solutions to GLRP with AV error
S_{PL} = set of optimal solutions to GLRP with PL error

Let $X_i \perp\!\!\!\perp Z(C_0, \boldsymbol{C}) \quad \forall\, (C_0, \boldsymbol{C}) \in S_{AV} \qquad i = 1, \ldots, d$

Then

$$S_{AV} \subseteq S_{PL}$$

Equivalence of Regressions III

Equivalence of regressions with Piecewice Linear and Koenker-Bassett Errors

Theorem III [Malandii, Kuzmenko, Uryasev, 2024]

Consider the Generalized Regression Problem

Let $K > 0$

$\hat{S}_{PL}$ = set of optimal solutions with PL error

$(f^*, C) \in \hat{S}_{PL}$

$\hat{S}_{KB}$ = set of optimal solutions with KB error with $\alpha^* = F_{Z_{f^*}}(0)$

Then

$$\hat{S}_{PL} \subseteq \hat{S}_{KB}$$

Equivalence of Three Regressions

Corollary I

$(i)\ (a_0, \boldsymbol{a}) \in \underset{C_0, \boldsymbol{C}}{\operatorname{argmin}}\ \mathcal{E}_q(Z(C_0, \boldsymbol{C})) \quad q = \frac{K+1}{2K+1}$

$(ii)\ (a_0, \boldsymbol{a}) \in \underset{C_0, \boldsymbol{C}}{\operatorname{argmin}}\ \mathcal{E}_K(Z(C_0, \boldsymbol{C}))$

$(iii)\ (a_0, \boldsymbol{a}) \in \underset{C_0, \boldsymbol{C}}{\operatorname{argmin}}\ \mathcal{E}_\alpha(Z(C_0, \boldsymbol{C})) \quad \alpha = F_\varepsilon(0)$

Corollary II

Let $(C_0^*, \boldsymbol{C}^*) \in S_{AV}$

S_{KB} = set of optimal solutions to GRLP with KB error, $\alpha^* = F_{Z(C_0^*, \boldsymbol{C}^*)}(0)$

Then

$$S_{AV} \subseteq S_{PL} \subseteq S_{KB}$$

Closing Part III: Key Takeaways

$$\mathcal{E}_q(X) = \mathbb{E}[qX_+^2 + (1-q)X_-^2]$$

$$\mathcal{E}_K(X) = \max\left\{-\mathbb{E}[X], \frac{1}{K}\mathbb{E}[(X)_+]\right\}$$

- Expectile $e_q(X) = \underset{C}{\operatorname{argmin}}\{\mathcal{E}_q(X - C)\}$ is law invariant, coherent, elicitable risk measure for $q \in (1/2, 1)$
- There are 2 Expectile Quadragles $\implies$ 2 equivalent errors ($\mathcal{E}_q$ and $\mathcal{E}_K$) for expectile estimation
- CVaR and expectile portfolio optimization problems are equivalent
- Expectile and quantile regressions are equivalent

Biased Mean Quadrangle

Superexpectation [Rockafellar & Royset, 2014]

$$\mathbb{E}_X(x) = \mathbb{E}[(X-x)_+] + x = \max_{\alpha\in[0,1]} \Big\{ \alpha x + (1-\alpha)\bar{q}_\alpha(X) \Big\} \qquad x \in \mathbb{R}$$

Superexpectation Deviation [Malandii & Uryasev, 2024]

$$\mathbb{E}_{X-\mathbb{E}[X]}(x) - x_+ = \mathbb{E}[(X-\mathbb{E}[X]-x)_+] - x_- = \max_{\alpha\in[0,1]} \Big\{ (1-\alpha)\mathcal{D}_\alpha(X) - ((1-\alpha)x_+ + \alpha x_-) \Big\}$$

Superexpectation Error [Malandii & Uryasev, 2024]

$$\mathcal{E}_x(X) = \max_{\alpha\in[0,1]} \Big\{ (1-\alpha)\mathcal{E}_\alpha(X) - ((1-\alpha)x_+ + \alpha x_-) \Big\}$$

Biased Mean Quadrangle [Malandii & Uryasev, 2024]

$\mathcal{S}_x(X) = x + \mathbb{E}[X] =$ *biased mean*

$\mathcal{R}_x(X) = \mathbb{E}_{X-\mathbb{E}[X]}(x) - x_+ + \mathbb{E}[X] =$ *superexpectation risk*

$\mathcal{D}_x(X) = \mathbb{E}_{X-\mathbb{E}[X]}(x) - x_+ =$ *superexpectation deviation*

$\mathcal{V}_x(X) = \max\{\mathbb{E}[X_-] - x_+, \mathbb{E}[X_+] - x_-\} + \mathbb{E}[X] =$ *superexpectation regret*

$\mathcal{E}_x(X) = \max\{\mathbb{E}[X_-] - x_+, \mathbb{E}[X_+] - x_-\} =$ *superexpectation error*

Biased Mean Quadrangle (Cont'd)

Optimal α

$$\underset{\alpha\in[0,1]}{\text{argmax}}\left\{(1-\alpha)\mathcal{D}_\alpha(X)-((1-\alpha)x_+ + \alpha x_-)\right\} = [\mathbb{P}(X < x+\mathbb{E}[X]), \mathbb{P}(X \leq x+\mathbb{E}[X])]$$

Equivalently

$$x+\mathbb{E}[X] \in q_\alpha(X)$$

Special case: $x = 0$

Mean Quadrangle (piecewise linear error) [Malandii & Uryasev, 2024]

$$\mathcal{S}_0(X) = \mathbb{E}[X] = \textit{mean}$$
$$\mathcal{R}_0(X) = \mathbb{E}[(X-\mathbb{E}[X])_+] + \mathbb{E}[X] = \textit{mean} - \textit{semidiviation risk}$$
$$\mathcal{D}_0(X) = \mathbb{E}[(X-\mathbb{E}[X])_+] = \frac{1}{2}\mathbb{E}[|X-\mathbb{E}[X]|] = \textit{mean absolute deviation}$$
$$\mathcal{V}_0(X) = \max\{\mathbb{E}[X_-], \mathbb{E}[X_+]\} + \mathbb{E}[X]$$
$$\mathcal{E}_0(X) = \max\{\mathbb{E}[X_-], \mathbb{E}[X_+]\}$$

Equivalence of Regressions IV

Equivalence of regressions with Superexpectation and Koenker-Bassett Errors

Theorem IV [Malandii & Uryasev, 2024]

Consider the Generalized Regression Problem

Let $x \in \mathbb{R}$

$\hat{S}_{SE}$ = set of optimal solutions with SE error

$(f^*, C) \in \hat{S}_{SE}$

$\hat{S}_{KB}$ = set of optimal solutions with KB error with $\alpha^* = F_{Z_{f^*}}(0)$

Then

$$\hat{S}_{SE} \subseteq \hat{S}_{KB}$$

Equivalence of Regressions V & Robust Mean Estimation

Equivalence of regressions with Superexpectation and Mean Squared Errors

Theorem V [Malandii & Uryasev, 2024]

Let $X_i,\ Y,\ \varepsilon =$ random variables defined on $\mathcal{L}^2(\Omega)\quad i = 1,\dots,d$

$Y = a_0 + \boldsymbol{a}^\top \boldsymbol{X} + \varepsilon\quad a.s. =$ true regression law $\quad (a_0, \boldsymbol{a}) \in \mathbb{R}^{d+1}$

$(a)\ X_i \perp\!\!\!\perp \varepsilon\quad (b)\ \mathbb{E}[\varepsilon] = 0\quad (\perp\!\!\!\perp =$ independent$)$

Then

- $(a_0, \boldsymbol{a}) \in \underset{C_0, \boldsymbol{C}}{\operatorname{argmin}}\ \mathcal{E}_q(Z(C_0, \boldsymbol{C}))\quad q = \frac{1}{2}$
- $(a_0, \boldsymbol{a}) \in \underset{C_0, \boldsymbol{C}}{\operatorname{argmin}}\ \mathcal{E}_x(Z(C_0, \boldsymbol{C}))\quad x = 0$

Minimize Standard Deviation

$$\min_{\boldsymbol{C}}\ \sigma(\bar{Z}(\boldsymbol{C}))$$

Minimize Mean Absolute Deviation

$$\min_{\boldsymbol{C}}\ \mathbb{E}[|\bar{Z}(\boldsymbol{C}) - \mathbb{E}[\bar{Z}(\boldsymbol{C})]|]$$

$$C_0 = \mathbb{E}[\bar{Z}(\boldsymbol{C})]$$

Closing Part IV: Key Takeaways

$$\mathcal{E}_\alpha(X) = \mathbb{E}\left[\frac{\alpha}{1-\alpha}X_+ + X_-\right]$$

$$\mathcal{E}_x(X) = \max\left\{\mathbb{E}[X_+] - x_-, \mathbb{E}[X_-] - x_+\right\}$$

$$\mathcal{E}_q(X) = \frac{1}{2}\mathbb{E}[X^2] \quad q = 1/2$$

- Biased mean and quantile regressions are equivalent
- Biased mean regression allows estimating the mean via linear programming

General Observation: The Universal Utility of Quantiles

CVaR Regression [Rockafellar et al., 2014]

Goal: estimate $\bar{q}_\alpha(Y)$

- Rockafellar Error $\mathcal{E}(Y) = \min_{B_1,\dots,B_r} \left\{ \sum_{k=1}^{r} \lambda_k \mathcal{E}_{\alpha_k}(Y - B_k) \,\middle|\, \sum_{k=1}^{r} \lambda_k B_k = 0 \right\}$
- Statistic $S(Y) = \sum_{k=1}^{r} \lambda_k q_{\alpha_k}(Y) \qquad \lambda_k \geq 0,\ \sum_{k=1}^{r} \lambda_k = 1,\ \alpha \leq \alpha_k \leq 1$

Support Vector Regression [Vapnik, 1995], [Schölkopf et al., 2000]

Goal: estimate $\mathbb{E}[Y]$ in a sparse way (using the least amount of data possible)

- CVaR Norm $\langle\langle Y \rangle\rangle_\alpha = \bar{q}_\alpha(|Y|)$
- Vapnik Error $\mathcal{E}(Y) = \mathbb{E}[|Y| - x]_+ \quad x \in \mathbb{R}$
- Statistic $\mathcal{S}(Y) = \frac{1}{2}(q_{(1+\alpha)/2}(Y) + q_{(1-\alpha)/2}(Y))$

$\varphi-$Divergence Quadrangle: Examples

CVaR Portfolio Optimization

$$\min_{\boldsymbol{x}\in\mathcal{X}_\mu} \bar{q}_\alpha(L(\boldsymbol{x})) = \min_{\boldsymbol{x}\in\mathcal{X}_\mu} \max_{Q\in\mathcal{Q}_\varphi} \mathbb{E}[L(\boldsymbol{x})Q]$$

Quantile Regression

$$\min_{f,C} \mathcal{E}_\alpha(Z_f) = \min_f \bar{q}_\alpha(Z_f) - \mathbb{E}[Z_f] = \min_f \max_{Q\in\mathcal{Q}_\varphi} \mathbb{E}[Z_f Q] - \mathbb{E}[Z_f] \text{ and } C = q_\alpha(Z_f)$$

$$\mathcal{Q}_\varphi = \{Q \in \mathcal{L}^2(\Omega) \mid \mathbb{E}[Q] = 1,\ \mathbb{E}[\varphi(Q)] \leq \beta\} \quad \varphi(x) = \begin{cases} 0, & x \in [0, \frac{1}{1-\alpha}] \\ +\infty, & \text{otherwise} \end{cases}$$

Link to Distributionally Robust Optimization (Probability Alternatives)

Measures $\mathbb{P}$ can be represented by densities $\frac{d\mathbb{P}}{d\mathbb{P}_0}$ wrt $\mathbb{P}_0$ (underlying measure)

$$\mathbb{E}[QL(\boldsymbol{x})] = \mathbb{E}_{\mathbb{P}}[L(\boldsymbol{x})] \text{ for } Q = \frac{d\mathbb{P}}{d\mathbb{P}_0}$$

[sets $\mathcal{P}$ of alternatives $\mathbb{P}$] $\longleftrightarrow$ [sets $\mathcal{Q}$ of $Q \in \mathcal{L}^2_+ \mid \mathbb{E}[Q] = 1$]

$\varphi-$Divergence Quadrangle: Basic Definitions

φ-divergence Function

A convex lsc function $\varphi : \mathbb{R} \to \bar{\mathbb{R}}$ is a divergence function if $\varphi(1) = 0$

φ-divergence Uncertainty Sets

For a radius $\beta > 0$

$$\mathcal{Q}_{\varphi,\beta} = \{Q \in \mathcal{L}^2 |\ \mathbb{E}[Q] = 1,\ \mathbb{E}[\varphi(Q)] \leq \beta\}$$

Remark (domain of φ)

If $\varphi(x) = +\infty$ for $x < 0$ then $Q \in \mathcal{Q}_{\varphi,\beta}$ might be interpreted as a density

$\varphi-$Divergence Quadrangle: Dual Representation

$\varphi-$Divergence Quadrangle [Peng, Malandii, Uryasev, 2024]

$$\mathcal{R}_{\varphi,\beta}(X) = \sup_{Q \in \mathcal{Q}^{\mathcal{R}}_{\varphi,\beta}} \mathbb{E}[XQ]$$

$$\mathcal{V}_{\varphi,\beta}(X) = \sup_{Q \in \mathcal{Q}^{\mathcal{V}}_{\varphi,\beta}} \mathbb{E}[XQ]$$

$$\mathcal{D}_{\varphi,\beta}(X) = \sup_{Q \in \mathcal{Q}^{\mathcal{R}}_{\varphi,\beta}} \mathbb{E}[X(Q-1)]$$

$$\mathcal{E}_{\varphi,\beta}(X) = \sup_{Q \in \mathcal{Q}^{\mathcal{V}}_{\varphi,\beta}} \mathbb{E}[X(Q-1)]$$

$$\mathcal{S}_{\varphi,\beta}(X) = \operatorname*{argmin}_{C} \sup_{Q \in \mathcal{Q}^{\mathcal{V}}_{\varphi,\beta}} \mathbb{E}[(X-C)(Q-1)]$$

$$\mathcal{Q}^{\mathcal{R}}_{\varphi,\beta} = \{Q \in \mathcal{L}^2(\Omega) | \mathbb{E}[Q] = 1,\ \mathbb{E}[\varphi(Q)] \leq \beta\}$$

$$\mathcal{Q}^{\mathcal{V}}_{\varphi,\beta} = \{Q \in \mathcal{L}^2(\Omega) | \mathbb{E}[\varphi(Q)] \leq \beta\}$$

$\varphi-$Divergence Quadrangle: Primal Representation

Risk $\mathcal{R}_{\varphi,\beta}(X)$ studied by [Shapiro, 2017], [Dommel & Pichler, 2020]

$\varphi^* =$ convex conjugate of φ

$\varphi-$Divergence Quadrangle [Peng, Malandii, Uryasev, 2024]

$$\mathcal{R}_{\varphi,\beta}(X) = \inf_{\substack{C\in\mathbb{R},\\ t>0}} t\left\{ C + \beta + \mathbb{E}\left[\varphi^*\left(\frac{X}{t} - C\right)\right]\right\}$$

$$\mathcal{D}_{\varphi,\beta}(X) = \inf_{\substack{C\in\mathbb{R},\\ t>0}} t\left\{ C + \beta + \mathbb{E}\left[\varphi^*\left(\frac{X}{t} - C\right) - \frac{X}{t}\right]\right\}$$

$$\mathcal{V}_{\varphi,\beta}(X) = \inf_{t>0} t\left\{ \beta + \mathbb{E}\left[\varphi^*\left(\frac{X}{t}\right)\right]\right\}$$

$$\mathcal{E}_{\varphi,\beta}(X) = \inf_{t>0} t\left\{ \beta + \mathbb{E}\left[\varphi^*\left(\frac{X}{t}\right) - \frac{X}{t}\right]\right\}$$

$$\mathcal{S}_{\varphi,\beta}(X) = \operatorname*{argmin}_{C} \inf_{t>0} t\left\{ \frac{C}{t} + \beta + \mathbb{E}\left[\varphi^*\left(\frac{X - C}{t}\right)\right]\right\}$$

References

- Bellini, F. et al. (2014). "Generalized quantiles as risk measures". In: *Insurance: Mathematics and Economics* 54, pp. 41–48.
- Bertsimas, Dimitris, Dessislava Pachamanova, and Melvyn Sim (2011). "Robust linear optimization under general norms". In: *Operations Research Letters* 39.4, pp. 281–286. DOI: 10.1016/j.orl.2011.05.009.
- Chang, C. C. and C. J. Lin (2001). "Training ν-support Vector Regression: Theory and Algorithms". In: *Neural Computation* 14.8, pp. 1959–1977.
- Colombo, C. (2018). "Portfolio optimization with expectiles". In: *University of Milano*. URL: https://doi.org/10.13140/RG.2.2.35097.67685.
- Dommel, Paul and Alois Pichler (Mar. 2020). *Convex Risk Measures based on Divergence*. Papers 2003.07648. arXiv.org. URL: https://ideas.repec.org/p/arx/papers/2003.07648.html.
- Mafusalov, A. and S. Uryasev (2016). "CVaR (Superquantile) Norm: Stochastic Case". In: *European Journal of Operational Research* 249, pp. 200–208.

Malandii, Anton, Viktor Kuzmenko, and Stan Uryasev (2024). "Expectile Risk Quadrangles and Applications". In: *Working Paper.*

Malandii, Anton and Stan Uryasev (2022). "Support Vector Regression: Risk Quadrangle Framework". In: *Working Paper.*

— (2024). "Biased Mean Quadrangles and Applications". In: *Working Paper.*

Newey, W. K. and J. L. Powell (1987). "Asymmetric least squares estimation and testing". In: *Econometrica* 55.4, pp. 819–847. URL: https://doi.org/10.2307/1911031.

Peng, Cheng, Anton Malandii, and Stan Uryasev (2024). "Risk Quadrangle and Distributionally Robust Regression Based on φ-Divergence". In: *Working Paper.*

Rockafellar, R Tyrrell and Johannes O Royset (2015). "Measures of residual risk with connections to regression, risk tracking, surrogate models, and ambiguity". In: *SIAM Journal on Optimization* 25.2, pp. 1179–1208.

Rockafellar, R Tyrrell, Stan Uryasev, and Michael Zabarankin (2008). "Risk tuning with generalized linear regression". In: *Mathematics of Operations Research* 33.3, pp. 712–729.

Rockafellar, R. Tyrrell and Johannes O. Royset (2014). "Random variables, monotone relations, and convex analysis". In: *Mathematical Programming* 148, pp. 297–331.

Rockafellar, R. Tyrrell and Stan Uryasev (2013). "The fundamental risk quadrangle in risk management, optimization and statistical estimation". In: *Surveys in Operations Research and Management Science* 18.1-2, pp. 33–53. DOI: 10.1016/j.sorms.2013.03.001.

Schölkopf, B. et al. (2000). "New Support Vector Algorithms". In: *Neural Computation* 12.5, pp. 1207–1245.

Shapiro, Alexander (2017). "Distributionally robust stochastic programming". In: *SIAM Journal on Optimization* 27.4, pp. 2258–2275.

Takeda, Akiko, Jun-ya Gotoh, and Masashi Sugiama (2010). "Support vector regression as conditional value-at-risk minimization with application to financial time-series analysis". In: *2010 IEEE International Workshop on Machine Learning for Signal Processing*. IEEE, pp. 118–123.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Data Mining and Knowledge Discovery, Springer.

Wagner, A. and S. Uryasev (2019). "Portfolio Optimization with Expectile and Omega Functions". In: *Proceedings of the 2019 Winter Simulation Conference*. Ed. by N. Mustafee et al., pp. 926–937.

Case Studies

Problems:

- Hedging Portfolio of Options by Minimizing Expectile
- CVaR and Expectile Portfolio Optimization
- Expectile Estimation with Linear Regression Using Three Errors
  - Artificial Data
  - Style Classification Data
  - Portfolio Hedging Data + Cardinality Constraints
- Biased Mean Regression Using Three Errors
  - Artificial Data
- Support Vector Regression
- φ-divergence Quadrangle:
  - Portfolio Optimization
  - Regression
  - Classification

Solvers:

- Portfolio Safeguard (PSG)
- Gurobi

Case Study: Hedging Portfolio of Options

Problem: hedging a Portfolio of Options by a Portfolio of Stocks and Options by minimizing Expectile with a budget constraint

Data: 45,000 scenarios & 121 instruments

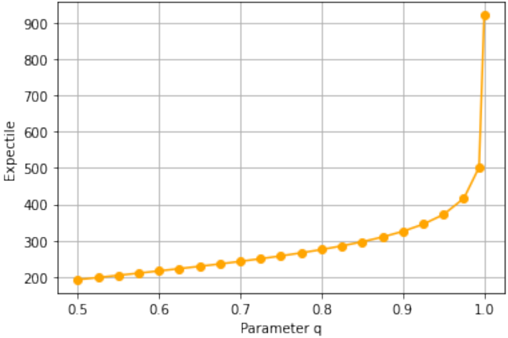


Figure 1: Hedging portfolio of options with expectile risk: expectile as a function $q \in (1/2, 1)$

Case Study: CVaR and Expectile Portfolio Optimization

Problem: optimizing a Portfolio of Assets by minimizing Expectile and CVaR with budget and expected return constraints

Data: 10,000 scenarios & 4 instruments

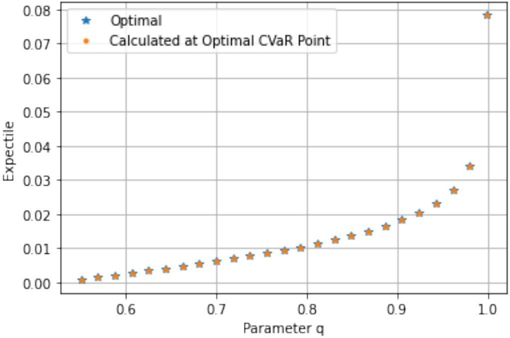


Figure 2: CVaR and Expectile portfolio optimization: expectile as a function $q \in (1/2, 1)$

Case Study: Expectile Regression on Artificial Data

Model: $Y = X + \varepsilon,\ X \sim \mathcal{N}(0,1),\ \varepsilon \sim \mathcal{N}(-0.7657751, 1),\ q = 0.875$

Problem: minimize asymmetric variance (AV), piecewise linear (PL), and normalized KB errors

Data: 100 samples with the size of $n =$ 100, 500, 1000, 5000, 10000, 50000, 100000, 500000

Sample Size	Average (AV)	Average (PL)	Average (KB)
100	0.146534	0.177659	0.201121
500	0.063652	0.071750	0.075280
1000	0.043098	0.049517	0.054172
5000	0.019807	0.023852	0.025737
10000	0.015490	0.017866	0.019351
50000	0.006986	0.007596	0.008282
100000	0.005190	0.005695	0.006174
500000	0.001947	0.002294	0.002536

Case Study: Style Classification with Expectile Regression

Problem: Fidelity Magellan Fund style classification with expectile regression using three errors

Data: 4 indices & 1,264 scenarios

#		AV Error	PL Error	KB Error
1.	Parameter value	$q = 0.875$	$K = 1/6$	$\alpha^* = 0.80063$
2.	Expectile C_0^* (*"intercept"*) (explanatory variables = 0)	0.033323	0.033323	0.033333
3.	Error (explanatory variables=0)	7.32e-4	0.036522	0.013980
4.	Optimal point:			
5.	RLV	0.557645	0.549268	0.549968
6.	RLG	0.494235	0.508173	0.508088
7.	RUJ	-0.063137	-0.071745	-0.072769
8.	RUO	-0.006424	-0.005496	-0.005520
9.	intercept	3.754e-3	3.9798e-3	4.035e-3
10.	Errors at optimal point for AV	1.4e-5	4.895e-3	1.792e-3
11.	Errors at optimal point for PL	1.5e-5	4.88e-3	1.786e-3
12.	Errors at optimal point for KB	1.5e-5	4.926e-3	1.786e-3

Case Study: Portfolio Tracking with Cardinality Constraints

Problem: expectile estimation with PL error ($K = 1/6$) and AV error ($q = 0.875$), unconstrained and constrained (card = 25)

Data: 45,000 scenarios & 121 instruments

Solver	Problem	Objective Value*	Objective Value$^{@}$	Solving Time (s)
PSG	$\mathcal{E}_K$ (unconstrained)	32.330525	32.531606	18.5
PSG	$\mathcal{E}_K$ (constrained)	32.374030	32.560966	109.5
Gurobi	$\mathcal{E}_K$ (constrained)	32.374030	32.560966	202.2
PSG	$\mathcal{E}_q$ (unconstrained)	567.537406	575.342945	12.6
Gurobi	$\mathcal{E}_q$ (constrained)	568.894529	576.076514	55683.0

Case Study: SE Regression on Artificial Data

Model: $Y = X + \varepsilon,\ X \sim \mathcal{N}(0,1),\ \varepsilon \sim$ skewed $\mathcal{N}(10,0,1),\ \alpha = 0.57$

Problem: minimize Mean Squared (MS), Superexpectation (SE), and normalized KB errors

Data: 100 samples with the size of $n =$ 100, 500, 1000, 5000, 10000, 50000, 100000, 500000

Sample Size	Average (MS)	Average (SE)	Average (KB)
100	0.053674	0.072555	0.089225
500	0.023166	0.033096	0.040961
1000	0.016722	0.022237	0.028069
5000	0.007264	0.009865	0.012039
10000	0.004465	0.007651	0.009193
50000	0.002167	0.003265	0.003907
100000	0.001476	0.001782	0.002466
500000	0.000751	0.001008	0.001330

Case Study: Robustness of SE Regression

Model: $Y = X + \varepsilon,\ X \sim \mathcal{U}([1,10]),\ \varepsilon \sim \mathcal{N}(0,1) +$ outliers

Problem: minimize Mean Squared (MS), Superexpectation (SE), and normalized (KB) ($\alpha = 0.5$) errors

Data: 45 observations, 3 outliers

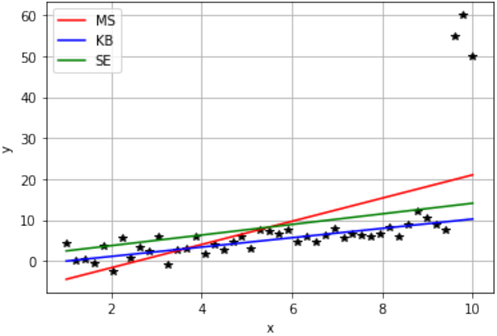


Figure 3: Performance of errors in the presence of outliers

Case Study: SVR: Equivalence & Dual Formulation

True Law: $f(x) = x, \quad x \in [0, 1]$

Uniform Partitioning: $0 = x_1 < x_2 < \ldots < x_\ell = 1$

Model: $y_i = x_i + \epsilon_i, \quad i = 1, \ldots, \ell, \ \epsilon_i \sim \mathrm{Laplace}(0, 1)$

Parameters: $\alpha = 0.6, \ \ell = 1000, \ C = 1$

Regularization Coefficient: $\lambda = \frac{1}{2C\ell}$

Solver: PSG

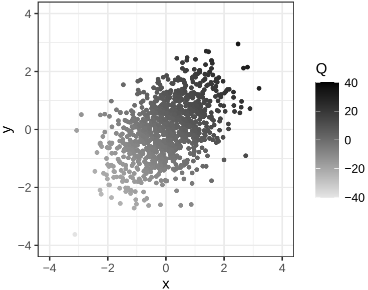
Method	Objective Function	b^* $\boldsymbol{w}^*$	α	ε	Time (sec)
ν-SVR (primal)	error	0.020 0.932	0.6	0.914	0.01
ε-SVR (primal)	error	0.020 0.932	0.600	0.914	0.01
ν-SVR (primal)	deviation	0.019 0.932	0.6	0.914	0.01
ν-SVR (dual)	error	0.019 0.932	0.6	0.914	0.09

Table 1: Optimization outputs: SVR, primal and dual formulations.

Case Study: φ-divergence Quadrangle: Portfolio Optimization

Problem: Illustration of Markowitz Portfolio Optimization as robust optimization ($\beta = 100$)
Data: $1,000$ samples from a bivariate Gaussian distribution with $\sigma_i = 1,\ \rho = 0.5,\ \mu_i = 0$

Risk envelope in Markowitz portfolio optimization. The darker the color of the point, the greater the value of $Q^*(w)$. The optimal portfolio weight is $(0.509, 0.491)$.



Case Study: φ-divergence Quadrangle: Regression

Problem: Illustration of Least Squares as robust optimization ($\beta = 100$)
Data: $1,000$ samples from a bivariate Gaussian distribution with $\sigma_i = 1,\ \rho = 0.5,\ \mu_i = 0$

Risk envelope in the least squares regression. The darker the color of the point, the greater the value $Q^*(w)$. The straight line is the least squares regression line $y = 0.47x + 0.00197$.