RISK MODELING IN MANAGEMENT AND ENGINEERING

R. T. Rockafellar
University of Washington, Seattle
University of Florida, Gainesville

Workshop on Optimization and Risk Management
The Hong Kong Polytechnic University
18 December 2015

Decisions in the Face of Uncertain Outcomes

Reliable design in engineering:

a structure must be engineered to withstand various impacts
possible impacts can only be estimated
potential for failure must be kept low

Logistics in operations research:

supplies must be stockpiled to meet demands
distribution and replacement expense should be kept at bay
demands and costs are random variables, but partly guesswork

Asset management in finance:

a portfolio must be chosen with intelligent safeguarding
performance known at best from history and economic factors
the threat of losses cannot be eliminated

The inescapable modeling challenge in stochastic optimization:

appropriate formulation of **constraints** and **objectives**

Uncertain “Costs” (Viewed Abstractly)

“costs” = quantities to be minimized or kept below given levels

General “cost” expression in decision-making:

$c(x, v)$ with $x =$ **decision** vector, $v =$ **data** vector

$$x = (x_1, \ldots, x_n), \qquad v = (v_1, \ldots, v_m)$$

Stochastic uncertainty:

v is replaced by a **random variable** vector $V = (V_1, \ldots, V_m)$

then the “cost” becomes a random variable: $\underline{c}(x) = c(x, V)$

Key consequence:

the distribution of $\underline{c}(x)$ can only be **shaped** by the choice of x

but how then can constraints or minimization be understood?

A Broad Pattern for Handling Risk in Optimization

Risk measures: functionals $\mathcal{R}$ that "quantify the risk" in a random variable X by a value $\mathcal{R}(X)$ ("risk" $\neq$ "uncertainty")

Systematic prescription

Faced with an uncertain "cost" $\underline{c}(x) = c(x, V)$ articulate it numerically as $\bar{c}(x) = \mathcal{R}(\underline{c}(x))$ for a choice of risk measure $\mathcal{R}$

Constraints: keeping $\underline{c}(x)$ "adequately" $\leq b$
modeled as: constraint $\bar{c}(x) = \mathcal{R}(\underline{c}(x)) \leq b$

Objectives: keeping $\underline{c}(x)$ "as low as reasonable"
modeled as: minimizing $\bar{c}(x) = \mathcal{R}(\underline{c}(x))$
$=$ choosing (x, b) such that $\bar{c}(x) \leq b$ for lowest b

Supporting theory: exploration of examples and guidelines
but note: choosing $\mathcal{R}$ expresses a decision-maker's preferences

Some Familiar Approaches Subject to Pros and Cons

[**random** cost $\underline{c}(x) = c(x, V)$ reduced to a **numerical** cost $\bar{c}(x)$]

Focusing on worst cases:

$\bar{c}(x) = \sup[\underline{c}(x)]$ [taking $\mathcal{R}(X) = \sup X$ (ess. sup)]

then $\bar{c}(x) \le b \iff \underline{c}(x) \le b$ almost surely

Passing to expectations:

$\bar{c}(x) = \mu[\underline{c}(x)] = E[\underline{c}(x)]$ [taking $\mathcal{R}(X) = \mu(X) = EX$]

then $\bar{c}(x) \le b \iff \underline{c}(x) \le b$ "on average"

Adopting a safety margin:

$\bar{c}(x) = \mu[\underline{c}(x)] + \lambda\sigma[\underline{c}(x)]$ [taking $\mathcal{R}(X) = \mu X + \lambda\sigma(X)$]

then $\bar{c}(x) \le b$ unless in tail beyond λ standard deviations

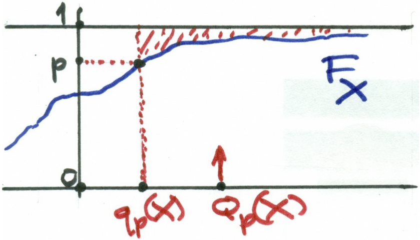
Looking at quantiles:

$\bar{c}(x) = q_p[\underline{c}(x)]$ [taking $\mathcal{R}(X) = p$-quantile of X]

then $\bar{c}(x) \le b \iff \mathrm{prob}\{\underline{c}(x) \le b\} \ge p$

Quantiles and "Superquantiles": VaR and CVaR

F_X = cumulative distribution function for random variable X



Quantile: "value-at-risk" in finance

$$q_p(X) = \mathrm{VaR}_p(X) = F_X^{-1}(p)$$

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

$$Q_p(X) = \mathrm{CVaR}_p(X) = E[\,X\,|X \geq q_p(X)\,] = \frac{1}{1-p}\int_p^1 q_t(X)dt$$

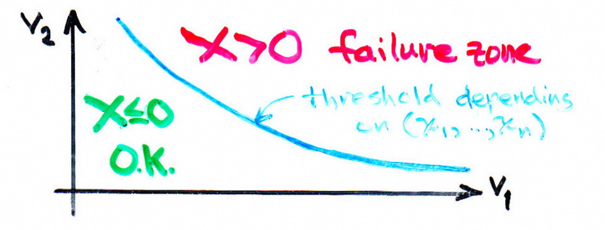
Replacing quantiles by superquantiles:

$\bar{c}(x) = Q_p[\underline{c}\,(x)]$ [taking $\mathcal{R}(X) = p$-superquantile of X]

then $\bar{c}(x) \leq b \iff \underline{c}\,(x) \leq b$ on average even in p-quantile-tail

An Engineering Perspective in Reliable Design

a "cost" triggering "failure": $X = c(x_1, \dots, x_n; V_1, \cdots, V_r)$



Probability of failure: $p_f(X) = \mathrm{prob}\{X > 0\}$

$$p_f(X) \leq 1 - p \quad \Longleftrightarrow \quad q_p(X) \leq 0$$

Troubles with this quantile concept:

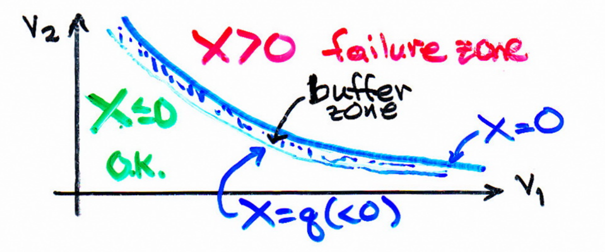
- How to compute or at least estimate?
- How to cope with dependence on $x_1, \dots, x_n$ in optimization?

both p_f and the threshold **shift** with changes in $x_1, \dots, x_n$

Alternative to consider: switch to a **superquantile** concept

Buffered Failure — Enhanced Safety

uncertain cost: $X = c(x_1, \ldots, x_n; V_1, \cdots, V_r)$



Buffered probability of failure: $P_f(X) = \mathrm{prob}\{X > q\}$

with q determined so as to make $E[X \mid X > q] = 0$!

$$P_f(X) \leq 1 - p \iff Q_p(X) \leq 0$$

Suggestion: adjust failure modeling to $P_f(X)$ in place of $p_f(X)$

safer by integrating tail information, and

easier also to work with in optimization! why?

Minimization Formula for VaR and CVaR

$$\mathrm{CVaR}_p(X) = \min_{C \in R} \left\{ C + \frac{1}{1-p} E\Big[\max\{0, X - C\} \Big] \right\} \quad p \in (0,1)$$
$$\mathrm{VaR}_p(X) = \mathrm{argmin} \;\; \text{(if unique, otherwise the lowest)}$$

Application to CVaR models: convert a problem in x like

minimize $\mathrm{CVaR}_{p_0}(\underline{c}_0(x))$ subject to [hard constraints and]
$$\mathrm{CVaR}_{p_i}(\underline{c}_i(x)) \leq 0, \;\; i = 1, \ldots, m$$

into a problem in x and auxiliary variables $C_0, C_1, \ldots, C_m$:

$$\text{minimize } C_0 + \frac{1}{1-p_0} E\Big[\max\{0, \underline{c}_0(x) - C_0\} \Big] \text{ while requiring}$$
$$C_i + \frac{1}{1-p_i} E\Big[\max\{0, \underline{c}_i(x) - C_i\} \Big] \leq 0, \;\; i = 1, \ldots, m$$

Important case: this converts to **linear programming** when
(1) each $\underline{c}_i(x) = c_i(x, V)$ depends **linearly** on x,
(2) the future state space Ω is modeled as **finite**

Axiomatization of Risk

Key axioms for risk measures: $\mathcal{R} : \mathcal{L}^p(\Omega, \mathcal{A}, P) \to (-\infty, \infty]$

(R1) $\mathcal{R}(C) = C$ for all constants C

(R2) convexity, (R3) lower semicontinuity

Supplementary properties of interest:

(R4) positive homogeneity: $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ for $\lambda > 0$

(R5) monotonicity: $\mathcal{R}(X) \leq \mathcal{R}(X')$ when $X \leq X'$

(R6) aversity: $\mathcal{R}(X) > EX$ for nonconstant X

Coherent measure of risk: $\mathcal{R}$ satisfying (R1), (R2), (R3), (R5)

Averse measure of risk: $\mathcal{R}$ satisfying (R1), (R2), (R3), (R6)

Artzner et al. (2000) introduced coherency without aversity

Preservation of convexity under coherency of $\mathcal{R}$

$\underline{c}(x) = c(x, V)$ convex in x $\implies$ $\bar{c}(x) = \mathcal{R}(\underline{c}(x))$ convex in x

a further advantage of coherency: it promotes **dualizations**

Coherency or its Absence

- $\mathcal{R}(X) = q_p(X) = \mathrm{VaR}_p(X)$ fails (R2), (R3), and (R6)!
- $\mathcal{R}(X) = Q_p(X) = \mathrm{CVaR}_p(X)$ satisfies **all** axioms
- $\mathcal{R}(X) = \sup X$ satisfies **all** axioms
- $\mathcal{R}(X) = \mu(X) = EX$ fails (R6) (coherency without aversity)
- $\mathcal{R}(X) = \mu(X) + \lambda\sigma(X)$, $\lambda > 0$, fails (R5) (no monotonicity)

Mixtures of risk measures

$$\mathcal{R}(X) = \sum_{k=1}^{r} \lambda_k \mathcal{R}_k(X) \;\; \text{with} \;\; \lambda_k > 0, \;\; \sum_{k=1}^{r} \lambda_k = 1$$

- $\mathcal{R}_k$ coherent for $k = 1, \ldots, r \quad \Longrightarrow \quad \mathcal{R}$ coherent
- $\mathcal{R}_k$ averse for $k = 1, \ldots, r \quad \Longrightarrow \quad \mathcal{R}$ averse

a versatile example: enjoying both coherency and aversity

$$\mathcal{R}(X) = \sum_{k=1}^{r} \lambda_k Q_{p_k}(X) = \sum_{k=1}^{r} \lambda_k \, \mathrm{CVaR}_{p_k}(X)$$

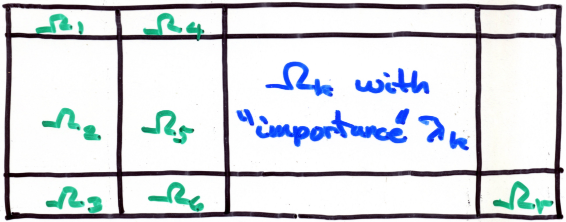
"continuous" mixtures work also $\longleftrightarrow$ dual utility risk profiles

Relaxations From Worst-Case Analysis

Interpolations with CVaR:

$\mathrm{CVaR}_p(X) \nearrow \sup X$ as $p \nearrow 1$, $\quad \mathrm{CVaR}_p(X) \searrow EX$ as $p \searrow 0$

Partitioning the future: Ω as a union of "likelihood regions" Ω_k



$\lambda_k > 0$ for $k = 1, \ldots, r, \quad \lambda_1 + \cdots + \lambda_r = 1$

$$\mathcal{R}(X) = \lambda_1 \Big[\sup_{\omega \in \Omega_1} X(\omega) \Big] \cdots + \lambda_r \Big[\sup_{\omega \in \Omega_r} X(\omega) \Big]$$ is **coherent**

Axiomatization of Deviation, a Partner to Risk

$\mathcal{D}(X)$ quantifies the **nonconstancy/uncertainty** in X

Key axioms for deviation measures: $\mathcal{D} : \mathcal{L}^1(\Omega, \mathcal{A}, P) \to [0, \infty]$

(D1) $\mathcal{D}(C) = 0$, but $\mathcal{D}(X) > 0$ for nonconstant X

(D2) convexity, (D3) lower semicontinuity

Supplementary properties of interest:

(D4) positive homogeneity: $\mathcal{D}(\lambda X) = \lambda \mathcal{D}(X)$ for $\lambda > 0$

(D5) upper range boundedness: $\mathcal{D}(X) \leq \sup X - EX$

Risk measures in relation to deviation measures

A **one-to-one** correspondence $\mathcal{D} \longleftrightarrow \mathcal{R}$ between deviation measures $\mathcal{D}$ and **averse** risk measures $\mathcal{R}$ is furnished by

$$\mathcal{R}(X) = EX + \mathcal{D}(X), \qquad \mathcal{D}(X) = \mathcal{R}(X - EX),$$

where moreover **coherency** is characterized by

$$\mathcal{R}(X) \text{ satisfies (R5)} \iff \mathcal{D}(X) \text{ satisfies (D5)}$$

$\mathcal{D}(X) = \lambda\sigma(X)$ fails (D5), so $\mathcal{R}(X) = EX + \lambda\sigma(X)$ isn't coherent!

Optimization and Estimation in Coping With Uncertainty

support for decision-making in a stochastic environment

Optimization:

minimize a “cost” expression under constraints on the decision
the constraints could involve bounds on other “costs”
the “costs” may have a background in statistical analysis

Estimation:

approximate some quantity from empirical/historical data
minimize an error expression to get regression coefficients
different interpretations of “error” yield different results

Interplay:

- optimization problems involving uncertainty depend on estimation methodology even in coming to a formulation
- estimation problems are optimization of a special sort
- new and deeper connections are now coming to light

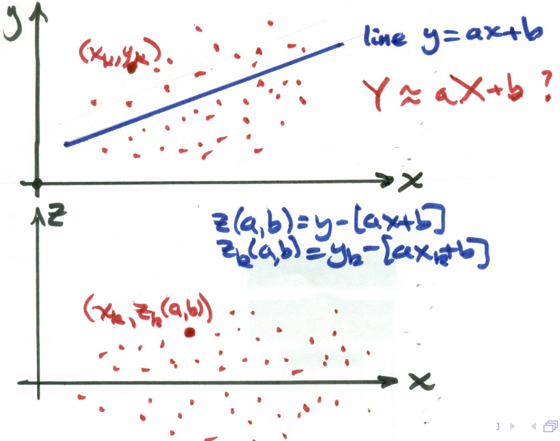
Databases: the Role of Estimation

a one-dimensional linear version for initial simplicity

Available information: a large(?) collection of pairs (x_k, y_k)

Perspective: empirical distribution in x, y space of r.v.'s X, Y

Approximation: $Y \approx aX + b$ error gap: $Z(a,b) = Y - [aX + b]$



Regression From a General Point of View

Y = random variable (scalar) to be understood in terms of
$X_1, \ldots, X_n$ = some "more basic" variables (e.g., "factors")

Approximation scheme: $Y \approx f(X_1, \ldots, X_n)$ for $f \in \mathcal{F}$

$\mathcal{F}$ = some specified class of functions $f : \boldsymbol{R}^n \to \boldsymbol{R}$

for instance $f(x_1, \ldots, x_n) = c_0 + c_1 x_1 + \cdots + c_n x_n$

corresponding error variable: $Z_f = Y - f(X_1, \ldots, X_n)$

Regression problem:

minimize $\mathcal{E}(Z_f)$ over all $f \in \mathcal{F}$ for some "error measure" $\mathcal{E}$

$\mathcal{E}$ quantifies the "**nonzeroness**" of the random variable Z_f

Important issue for connecting with optimization:

- $\underline{c}(x) = c(x, V)$ may only be supported by a database
- regression is essential then to get a "formula" for $\underline{c}(x)$
- for using $\bar{c}(x) = \mathcal{R}(\underline{c}(x))$, shouldn't this be "tuned" to $\mathcal{R}$?

Example: an Application to Composition of Alloys

Alloy model: a mixture of various metals

amounts of chief ingredients: $x = (x_1, \ldots, x_n)$ "design"

amounts of other ingredients: $v = (v_1, \ldots, v_m)$ "contaminants"

a "characteristic" to be controlled: y ideally kept ≤ 0, say

due to uncertainty, a quantile constraint may be envisioned

Background information: $y = c(x, v)$? no available formula!

there is only a database in (x, v, y)-space, $\{(x^k, v^k, y^k)\}_{k=1}^N$

- view the database as an empirical distribution for random variables $X = (X_1, \ldots, X_n)$, $V = (V_1, \ldots, V_m)$, Y
- use regression of Y on $X_1, \ldots, X_N$ to get a function $y = \bar{c}(x)$
- then impose the constraint $\bar{c}(x) \leq 0$ on the design x

shouldn't the regression adapt then to the intended constraint?

The Effect of Regression in General

$\mathcal{E}$ = some measure of **error**: (E1) $\mathcal{E}(X) > 0$ when $X \not\equiv 0$
(E2) convexity, (E3) lower semicontinuity

for instance $\mathcal{E}(X) = E\{\varepsilon(X)\}$ or $\mathcal{E}(X) = \min_{w \in W} E\{\varepsilon(w, X)\}$

Error projection (with respect to constants C)

Let $\mathcal{D}(X) = \min_C \mathcal{E}(X - C)$, $\mathcal{S}(X) = \operatorname{argmin}_C \mathcal{E}(X - C)$.

Then $\mathcal{D}$ is a measure of deviation and $\mathcal{S}$ the associated "statistic"

$\mathcal{S}(X)$ is the constant C "nearest" to X with respect to $\mathcal{E}$

Regression problem "decomposition" (when $f \in \mathcal{F} \Rightarrow f + C \in \mathcal{F}$)

minimizing $\mathcal{E}(Z_f)$ over $f \in \mathcal{F}$ corresponds to minimizing $\mathcal{D}(Z_f)$
under the constraint $\mathcal{S}(Z_f) = 0$ where $Z_f = Y - f(X_1, \ldots, X_n)$

basic examples: least-squares regression, quantile regression

The Fundamental Quadrangle of Risk: A New Paradigm

an array of functionals to be applied to random "costs" X
serving to connect methodologies usually viewed separately

$$\begin{array}{ccccc} & \text{risk } \mathcal{R} & \longleftrightarrow & \mathcal{D} \text{ deviation} & \\ \text{optimization} & \uparrow\downarrow & \mathcal{S} & \downarrow\uparrow & \text{estimation} \\ & \text{regret } \mathcal{V} & \longleftrightarrow & \mathcal{E} \text{ error} & \end{array}$$

$\mathcal{R}(X)$ quantifies the "overall" cost in X
$\mathcal{D}(X)$ quantifies the nonconstancy in X
$\mathcal{E}(X)$ quantifies the nonzeroness in X
$\mathcal{V}(X)$ quantifies the net regret in outcomes $X > 0$ versus $X \leq 0$
$\mathcal{S}(X)$ is the "statistic" associated with X through $\mathcal{E}$

dualizations yield many insights and connect with relative entropy

Regret and Utility

$\mathcal{V}(X)$ quantifies (net) regret in outcomes $X > 0$ versus $X \leq 0$

Key axioms for regret measures: $\mathcal{V} : \mathcal{X} \to (-\infty, \infty]$

(V1) $\mathcal{V}(0) = 0$, (V2) convexity, (V3) closedness

Supplementary properties of interest:

(V4) positive homogeneity: $\mathcal{V}(\lambda X) = \lambda \mathcal{V}(X)$ for $\lambda > 0$

(V5) monotonicity: $\mathcal{V}(X) \leq \mathcal{V}(X')$ when $X \leq X'$

(V6) aversity: $\mathcal{V}(X) > EX$ for nonconstant X

Coherent measure of regret: $\mathcal{V}$ with (V1), (V2), (V3), (V5)

Averse measure of regret: $\mathcal{V}$ with (V1), (V2), (V3), (V6)

Connection with utility: $\mathcal{V} \longleftrightarrow \mathcal{U}$ expressed by

$$\mathcal{V}(X) = -\mathcal{U}(-X), \quad \mathcal{U}(Y) - \mathcal{V}(-Y)$$

- this is **relative** utility because of the $\mathcal{V}$ threshold: $\mathcal{U}(0) = 0$

Expectation case: $\mathcal{V}(X) = E[v(X)]$, $\mathcal{U}(Y) = E[u(Y)]$, where

$v(x) = -u(-x)$ for a utility $u : (-\infty, \infty) \to [-\infty, \infty)$

Risk Assessed From Regret

Trade-off formula

From a regret measure $\mathcal{V}$, a risk measure $\mathcal{R}$ can be derived by

$$\mathcal{R}(X) = \min_C \{ C + \mathcal{V}(X - C) \}$$

and in general this operation preserves aversity and coherency

Interpretation: write off a **certain** loss amount C and then only worry about the **uncertain residual** loss $X - C$

Ideal amount to write off: $\mathcal{S}(X) = \operatorname{argmin}_C \{ C + \mathcal{V}(X - C) \}$

Example: $\mathcal{R}(X) = Q_p(X)$ and $\mathcal{S}(X) = q_p(X)$ emerge from $\mathcal{V}(X) = \frac{1}{1-p} E[\max\{0, X\}]$ (classical penalty for loss)

Important challenge: inversion $\mathcal{R} \to \mathcal{V}$ more generally nonunique, but "natural" antecedents are now broadly known

Finalizing the Quadrangle

$$\begin{array}{ccccc} & \text{risk } \mathcal{R} & \longleftrightarrow & \mathcal{D} \text{ deviation} & \\ \text{optimization} & \uparrow\downarrow & \mathcal{S} & \downarrow\uparrow & \text{estimation} \\ & \text{regret } \mathcal{V} & \longleftrightarrow & \mathcal{E} \text{ error} & \end{array}$$

To consider now: $\mathcal{V} \longleftrightarrow \mathcal{E}$ as set forth through

$$\mathcal{E}(X) = \mathcal{V}(X) - EX, \qquad \mathcal{V}(X) = EX + \mathcal{E}(X)$$

expectation case: $\mathcal{E}(X) = E[\varepsilon(X)]$ for $\varepsilon : (-\infty, \infty] \to [0, \infty]$

Key consequence for $\mathcal{S}$ under $\mathcal{V} \longleftrightarrow \mathcal{E}$ and $\mathcal{R} \longleftrightarrow \mathcal{D}$

The derivations $\mathcal{E} \to \mathcal{D}$, $\mathcal{E} \to \mathcal{S}$, transform now exactly to

$$\mathcal{R}(X) = \min_C \{ C + \mathcal{V}(X - C) \}, \quad \mathcal{S}(X) = \operatorname*{argmin}_C \{ C + \mathcal{V}(X - C) \}$$

Final links: nonunique but "natural" inversions $\mathcal{D} \to \mathcal{E}$, $\mathcal{R} \to \mathcal{V}$

The Mean-Based Quadrangle

articulated with a scaling parameter $\lambda > 0$

$$\mathcal{S}(X) = EX = \mu(X)$$
$$= \text{mean}$$

$$\mathcal{E}(X) = \lambda ||X||_2$$
$$= L^2\text{-error, scaled}$$

$$\mathcal{D}(X) = \lambda\,\sigma(X)$$
$$= \text{standard deviation, scaled}$$

$$\mathcal{R}(X) = EX + \lambda\,\sigma(X)$$
$$= \text{standard-deviational tail risk}$$

$$\mathcal{V}(X) = EX + \lambda ||X||_2$$
$$= L^2\text{-regret}$$

properties: **aversity** with **convexity** but not coherency

The Quantile-Based Quadrangle

at any probability level $p \in (0,1)$

$$\mathcal{S}(X) = q_p(X) = \mathrm{VaR}_p(X)$$
$$= \text{quantile}$$

$$\mathcal{R}(X) = Q_p(X) = \mathrm{CVaR}_p(X)$$
$$= \text{superquantile}$$

$$\mathcal{D}(X) = Q_p(X - EX) = \mathrm{CVaR}_p(X - EX)$$
$$= \text{superquantile deviation}$$

$$\mathcal{E}(X) = E[\tfrac{p}{1-p}X_+ + X_-]$$
$$X_+ = \max\{0, X\},\ \ X_- = \max\{0, -X\}$$
$$= \text{"normalized" Koenker-Bassett error}$$

$$\mathcal{V}(X) = \tfrac{1}{1-p}E[X_+]$$
$$= \text{quantile-scaled absolute loss}$$

properties: **aversity** with **coherency**

Some References

[1] H. Föllmer, A. Schied (2002, 2004), *Stochastic Finance*.

[2] R.T. Rockafellar, S.P. Uryasev (2002), "Conditional Value-at-Risk for General Loss Distributions," *J. Banking and Finance* 26, 1443–1471.

[3] R.T. Rockafellar, S. Uryasev, M. Zabarankin (2008), "Risk tuning in generalized linear regression," *Mathematics of Operations Research* 33, 712–729.

[4] R. T. Rockafellar, J. O. Royset (2010), "On buffered failure probability in design and optimization of structures," *Journal of Reliability Engineering and System Safety* 95, 499–510.

[5] R.T. Rockafellar, S.P. Uryasev (2013), "The fundamental risk quadrangle in risk management, optimization and statistical estimation," *Surveys in OR and Management Science* 18, 33–55.

my website: www.math.washington.edu/∼rtr/mypage.html