THE FUNDAMENTAL QUADRANGLE

relating quantifications of various aspects of a random variable

---

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

optimization | estimation

Lecture 1: optimization, the role of $\mathcal{R}$

Lecture 2: estimation, the roles of $\mathcal{E}$, $\mathcal{D}$, $\mathcal{S}$

Lecture 3: tying both together along with $\mathcal{V}$ and duality

Lecture 3

RISK VERSUS DEVIATION, REGRET AND ENTROPIC DUALITY

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

Newcastle, Australia **February, 2010**

Aversity in Risk

toward a fundamental connection with deviation measures

Recall axioms for coherent measures of risk

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

(R2) $\mathcal{R}((1-\lambda)X + \lambda X') \leq (1-\lambda)\mathcal{R}(X) + \lambda\mathcal{R}(X')$ for $\lambda \in (0,1)$ (convexity)

(R3) $\mathcal{R}(X) \leq \mathcal{R}(X')$ when $X \leq X'$ (monotonicity)

(R4) $\mathcal{R}(X) \leq c$ when $X_k \to X$ with $\mathcal{R}(X_k) \leq c$ (closedness)

(R5) $\mathcal{R}(\lambda X) = \lambda\mathcal{R}(X)$ for $\lambda > 0$ (positive homogeneity)

basic sense:(R5) yes, **extended** sense:(R5) no

Another important category of risk measures

$\mathcal{R}$ is an **averse** measure of risk if it satisfies (R1), (R2), (R4) and

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

basic sense: (R5) yes, **extended** sense: (R5) no

Risk Measures Paired With Deviation Measures

- Many risk measures are both coherent and averse
  $\mathcal{R}(X) = \mathrm{CVaR}_\alpha(X), \quad \mathcal{R}(X) = \sup X$
- Some risk measures are coherent but not averse
  $\mathcal{R}(X) = EX, \quad \mathcal{R}(X) = X(\bar{\omega})$
- Some risk measures are averse but not coherent
  $\mathcal{R}(X) = EX + \lambda\sigma(X)$ (to be seen shortly)

Coherency in deviation: require $\mathcal{D}(X) \leq \sup X - EX$ for all X

THEOREM: deviation versus risk

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 $\mathcal{R}$ is coherent $\iff$ $\mathcal{D}(X)$ is coherent

Note: coherency fails for deviation measures $\mathcal{D}(X) = \lambda\sigma(X)$!
$\Longrightarrow$ risk measures $\mathcal{R}(X) = \mu(X) + \lambda\sigma(X)$ aren't coherent

Safety Margins Revised

Recall the traditional approach to $\mu(X)$ being "safely" below 0:

$\mu(X) + \lambda\sigma(X) \leq 0$ for some $\lambda > 0$ scaling the "safety"

but $\mathcal{R}(X) = \mu(X) + \lambda\sigma(X)$ is not **coherent**

Can the coherency be restored if $\sigma(X)$ is replaced by some $\mathcal{D}(X)$?

Yes! $\mathcal{R}(X) = \mu(X) + \lambda\mathcal{D}(X)$ is coherent when $\mathcal{D}$ is coherent

Safety margin modeling with coherency

In the safeguarding problem model

minimize $\bar{c}_0(x)$ over $x \in S$ with $\bar{c}_i(x) \leq 0$ for $i = 1, \ldots, m$

where $\bar{c}_i(x) = \mathcal{R}_i(\underline{c}_i(x))$ for $\underline{c}_i(x) : \omega \to c_i(x, \omega)$

coherency is obtained with

$\mathcal{R}_i(X) = \mu(X) + \lambda_i\mathcal{D}_i(X)$ for $\lambda_i > 0$ and $\mathcal{D}_i$ coherent

Risk Envelope Characterization of Coherency

for coherent risk measures in the **basic** sense

A subset $\mathcal{Q}$ of $\mathcal{L}^2$ is a **coherent risk envelope** if it is nonempty, closed and convex, and $\quad Q \in \mathcal{Q} \quad \Longrightarrow \quad Q \geq 0,\ EQ = 1$

Interpretation: Any such Q is the "density" relative to the probability measure P on Ω of an alternative probability measure P' on Ω : $\quad E_{P'}[X] = E[XQ],\ \ Q = dP'/dP$

[specifying $\mathcal{Q}$] $\longleftrightarrow$ [specifying a comparison set of measures P']

Theorem: basic dualization

$\exists$ **one-to-one** correspondence $\mathcal{R} \longleftrightarrow \mathcal{Q}$ between coherent risk measures $\mathcal{R}$ in the **basic** sense and coherent risk envelopes Q:

$$\mathcal{R}(X) = \sup_{Q \in \mathcal{Q}} E[XQ], \qquad \mathcal{Q} = \left\{ Q \,\middle|\, E[XQ] \leq \mathcal{R}(X) \text{ for all } X \right\}$$

Conclusion: basic coherency $=$ "customized" worst-case analysis

Some Risk Envelope Examples

recall that "1" $=$ density Q of underlying P with respect to itself

$$\mathcal{R}(X) = EX \longleftrightarrow \mathcal{Q} = \{1\}$$

$$\mathcal{R}(X) = \sup X \longleftrightarrow \mathcal{Q} = \{\, \text{all } Q \geq 0,\ EQ = 1 \,\}$$

$$\mathcal{R}(X) = \mathrm{CVaR}_\alpha(X) \longleftrightarrow \mathcal{Q} = \{Q \geq 0,\, EQ = 1,\, Q \leq (1-\alpha)^{-1}\}$$

$$\mathcal{R}(X) = \sum_{k=1}^{r} \lambda_k \mathcal{R}(X) \longleftrightarrow \mathcal{Q} = \{\, \sum_{k=1}^{r} \lambda_k Q_k \,\big|\, Q_k \in \mathcal{Q}_k \}$$

Dual characterization of aversity:

- $\mathcal{R} \longleftrightarrow \mathcal{Q}$ as before, but $Q \in \mathcal{Q} \;\not\Rightarrow\; Q \geq 0$
- must have $1 \in \mathcal{Q}$ "strictly"

Entropic Characterization of Extended Coherency

what happens for coherent $\mathcal{R}$ without positive homogeneity?

Generalized entropy

Call a functional $\mathcal{I}$ on $\mathcal{L}^2$ an entropic distance when

(I1) $\mathcal{I}$ is convex and lower semicontinuous

(I2) $\mathcal{I}(Q) < \infty \quad \Longrightarrow \quad Q \geq 0,\ EQ = 1$

(I3) $\inf \mathcal{I} = 0 \quad \Longrightarrow \quad \mathrm{cl}(\mathrm{dom}\,\mathcal{I})$ is a risk envelope $\mathcal{Q}$

Theorem: extended dualization with conjugacy

$\exists$ **one-to-one** correspondence $\mathcal{R} \longleftrightarrow \mathcal{I}$ between coherent risk measures $\mathcal{R}$ in the **extended** sense and entropic distances $\mathcal{I}$:

$$\mathcal{R}(X) = \sup_Q \{E[XQ] - \mathcal{I}(Q)\}, \quad \mathcal{I}(Q) = \sup_X \{E[XQ] - \mathcal{R}(X)\}$$

Previous correspondence: $\mathcal{I} =$ "indicator" of $\mathcal{Q}$

Aversity: (I3) demands $\mathcal{I}(1) = 0$ with $1 \in \mathcal{Q}$ "strictly"

A Particularly Interesting Example

A pairing with Bolzmann-Shannon entropy

$\mathcal{R}(X) = \log E[e^X]$ coherent and averse corresponds to
$\mathcal{I}(Q) = E[Q \log Q]$ when $Q \geq 0,\ EQ = 1$ but $= \infty$ otherwise

How does this fit into the fundamental quadrangle?

- $\mathcal{D}(X) = \log E[\,e^{(X-EX)}\,]$ deviation measure paired with $\mathcal{R}$
- $\mathcal{E}(X) = E[e^X - X - 1]$ error measure projecting to $\mathcal{D}$
- $\mathcal{S}(X) = \log[e^X] = \mathcal{R}(X)\,!$ the "statistic" associated with $\mathcal{E}$

$\longrightarrow$ some development to be pursued in regression?

Expected Utility

Utility in finance: having a big role in traditional theory

X = incoming money in future, random variable

$u(x)$ = "utility" (in present terms) of getting future amount x

u generally concave, nondecreasing

$u(X(\omega))$ = utility of amount received in state $\omega \in \Omega$

$E[u(X)]$ = expected utility, something to consider maximizing

Importance of a threshold: X = gain/loss against **benchmark**

incrementally, people hate losses more than they love gains!

Normalization of utility: $x > 0$ rel. gain, $x < 0$ rel. loss

$u(0) = 0,\ u'(0) = 1$ for differentiable u, but the latter is equivalent without differentiability to $u(x) \leq x$ for all x

Resulting interpretation:

$u(x)$ = the amount of present money deemed to be acceptable in lieu of getting the future amount x

Translation to Minimization Framework

Utility replaced by regret: $v(x) = -u(-x)$

$v(x)$ = the regret in contemplating a future loss x

= the amount of present money deemed necessary as compensation for a relative loss x in the future

v is convex, nondecreasing, with $v(0) = 0$, $v(x) \geq x$

Converted context:

X = relative loss in future, random variable

$E[v(X)]$ = expected regret   something to consider minimizing

Insurance interpretation:

$E[v(X)]$ = the amount to charge (with respect to v) for covering the uncertain future loss X

Observations: about $\mathcal{V}(X) = E[v(X)]$ as a functional on $\mathcal{L}^2$

$\mathcal{V}$ is convex, nondecreasing, with $\mathcal{V}(0) = 0$, $\mathcal{V}(X) \geq EX$

Quantifications of Regret in General

expressions $\mathcal{V}(X)$ for potential losses X, not just of form $E[v(X)]$

Coherency in regret

Call $\mathcal{V}$ a **coherent** measure of regret if

(V1) $\mathcal{V}(0) = 0$

(V2) $\mathcal{V}((1-\lambda)X + \lambda X') \leq (1-\lambda)\mathcal{V}(X) + \lambda\mathcal{V}(X')$ (convexity)

(V3) $\mathcal{V}(X) \leq \mathcal{V}(X')$ when $X \leq X'$ (monotonicity)

(V4) $\mathcal{V}(X) \leq c$ when $X_k \to X$ with $\mathcal{V}(X_k) \leq c$ (closedness)

(V5) $\mathcal{V}(\lambda X) = \lambda\mathcal{V}(X)$ for $\lambda > 0$ (positive homogeneity)

Aversity in regret

Call $\mathcal{V}$ an **averse** measure of regret if (V3) is relinquished, but

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

basic sense: (V5) yes, **extended** sense: (V5) no

A Trade-off That Identifies Risk

For $\mathcal{V}$ = some measure of regret consider the expression:

$C + \mathcal{V}(X - C)$ for a future loss X and constants C

Interpretation: accept a certain loss C, thereby shifting the threshold and only regetting a residual future loss $X - C$

Theorem: derivation of risk from regret

Given an **averse** regret measure $\mathcal{V}$, define $\mathcal{R}$ and $\mathcal{S}$ by

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

Then
- $\mathcal{R}$ is an **averse** risk measure (coherent for $\mathcal{V}$ coherent)
- $\mathcal{S}(X)$ is a nonempty closed interval (singleton?)

CVaR example: $\mathcal{V}(X) = E[\frac{1}{1-\alpha} X_+]$

$$\mathcal{R}(X) = \min_C \{C + \tfrac{1}{1-\alpha} E[X - C]_+\} = \mathrm{CVaR}_\alpha(X)$$

$\longrightarrow$ the key minimization rule with argmin $= \mathrm{VaR}_\alpha(X) = q_\alpha(X)$

Completing the Fundamental Quadrangle of Risk

Error versus regret

The simple relations

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

provide a **one-to-one** correspondence between error measures $\mathcal{E}$ and **averse** regret measures $\mathcal{V}$ (with $V(C) < \infty$?), where

$$\mathcal{V} \text{ is coherent} \quad \Longleftrightarrow \quad \mathcal{E}(-X) \leq EX \text{ when } X \geq 0$$

Moreover, the $\mathcal{R}$ from $\mathcal{V}$ is **paired** with the $\mathcal{D}$ from $\mathcal{E}$, and in the minimization formulas giving statistics $\mathcal{S}$,

$$\text{the } \mathcal{S}(X) \text{ from } \mathcal{V} \to \mathcal{R} \quad = \quad \text{the } \mathcal{S}(X) \text{ from } \mathcal{E} \to \mathcal{D}$$

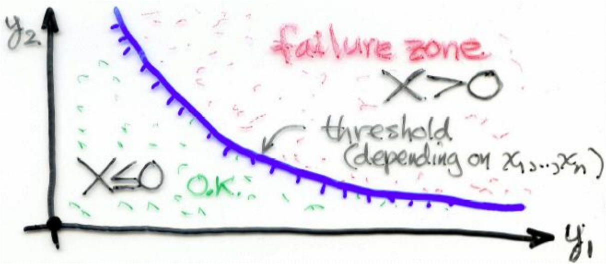
Expectation version:

$$\mathcal{V}(X) = E[v(X)] \longleftrightarrow \mathcal{E}(X) = E[\varepsilon(X)]$$
$$\varepsilon(x) = v(x) - x, \quad v(x) = x + \varepsilon(x)$$

Further Development From an Engineering Perspective

Uncertain "cost": $X = c(x_1, \ldots, x_n; Y_1, \cdots, Y_r)$

$x_1, \ldots, x_n$ =design variables, $Y_1, \ldots, Y_r$ =stochastic parameters



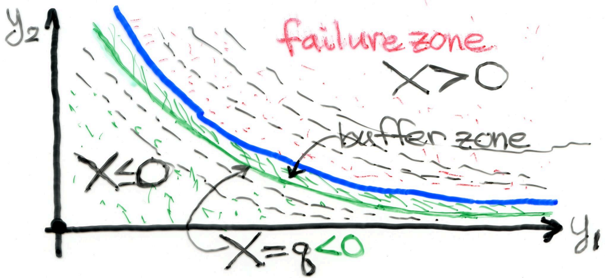
Probability of failure: $p_f = \text{prob}\{X > 0\}$

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

  Both p_f and the threshold **shift** with changes in $x_1, \ldots, x_n$

Buffered Failure — Enhanced Safety

Uncertain "cost": $X = c(x_1, \ldots, x_n; Y_1, \cdots, Y_r)$



Buffered probability of failure: $P_f = \text{prob}\{X > q\}$

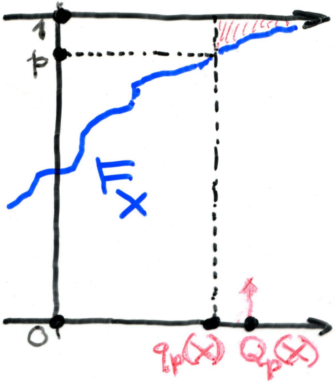
q determined so as to make $E[X \mid X > q] = 0$

Suggestion: adjust failure modeling to P_f in place of p_f

safer by integrating tail information, and

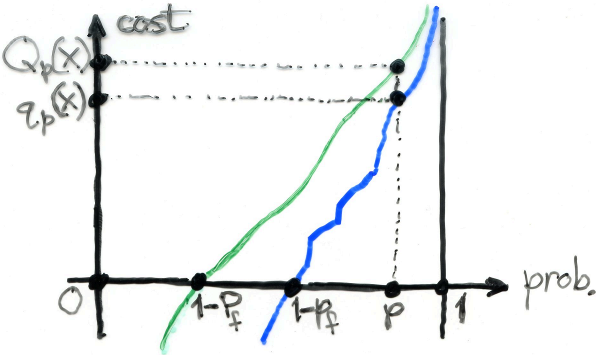
easier also to work with in optimization!

Quantiles and "Superquantiles"



quantile: $q_p(X) = F_X^{-1}(p) = \mathrm{VaR}_p(X)$
superquantile: $Q_p(X) = E[\,X\,|X > q_p(X)\,] \;= \mathrm{CVaR}_p(X)$
terms in finance: value-at-risk and conditional value-at-risk

Diagram of Relationships



$$q_p(X) = F_X^{-1}(p), \qquad Q_p(X) = \frac{1}{1-p}\int_p^1 q_s(X)\,ds$$

$q_p(X)$ can depend poorly on p, but $Q_p(X)$ depends smoothly on p

failure modeling: p_f determined by $q_p(X) = 0,\ p = 1 - p_f$

P_f determined by $Q_p(X) = 0,\ p = 1 - P_f$

Comparison of Roles in Optimization

Key fact: $\mathcal{R}(X) = Q_p(X)$ is **coherent** but $\mathcal{R}(X) = q_p(X)$ is not!

Constraint $p_f(\, c(x_1, \ldots, x_n, Y_1, \ldots, Y_m)\,) \leq 1 - p$ corresponds to
$q_p(\, c(x_1, \ldots, x_n, Y_1, \ldots, Y_m)\,) \leq 0$

Constraint $P_f(\, c(x_1, \ldots, x_n, Y_1, \ldots, Y_m)\,) \leq 1 - p$ corresponds to
$Q_p(\, c(x_1, \ldots, x_n, Y_1, \ldots, Y_m)\,) \leq 0$

Minimizing $q_p(\, c(x_1, \ldots, x_n, Y_1, \ldots, Y_m)\,)$ corresponds to
finding $x_1, \ldots, x_n$ with lowest C such that
$c(x_1, \ldots, x_n, Y_1, \ldots, Y_m) \leq C$ with probability $< 1 - p$

Minimizing $Q_p(\, c(x_1, \ldots, x_n, Y_1, \ldots, Y_m)\,)$ corresponds to
finding $x_1, \ldots, x_n$ with lowest C such that, even in the $1 - p$ worst fraction of cases, $c(x_1, \ldots, x_n, Y_1, \ldots, Y_m) \leq C$ on average

Some References

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

[2] A. Ben-Tal, M. Teboulle (2007), "An old-new concept of convex risk measures: the optimized certainty equivalent," *Mathematical Finance* 17, 449–476.

[3] R. T. Rockafellar (1974), *Conjugate Duality and Optimization*, No. 16 in the Conference Board of Math. Sciences Series, SIAM, Philadelphia.

[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*. downloadable:
faculty.nps.edu/joroset/docs/RockafellarRoyset_RESS.pdf

NEW BOOK NOW AVAILABLE