THE FUNDAMENTAL QUADRANGLE

relating quantifications of various aspects of a random variable

---

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

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 1

QUANTIFICATIONS OF RISK IN STOCHASTIC OPTIMIZATION

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

Newcastle, Australia **February, 2010**

Uncertainty in Optimization

Decisions (optimal?) must be taken before the facts are all in:

- A bridge must be built to withstand floods, wind storms or earthquakes
- A portfolio must be purchased with incomplete knowledge of how it will perform
- A product's design constraints must be viewed in terms of "safety margins"

What are the consequences for optimization?
How may this affect the way problems are **formulated**?

The Fundamental Difficulty Caused by Uncertainty

A standard form of optimization problem **without uncertainty**:

minimize $c_0(x)$ over all $x \in S$ satisfying $c_i(x) \leq 0, \; i = 1, \ldots, m$
for a set $S \subset R^n$ and functions $c_i : S \mapsto R$

Incorporation of **future states** $\omega \in \Omega$ in the model:

the decision x must be taken before ω is known

Choosing $x \in S$ no longer fixes numerical values $c_i(x)$, but only fixes **functions on** Ω: $\underline{c}_i(x) : \omega \mapsto c_i(x, \omega), \quad i = 0, 1, \ldots, m$

Example: Linear Programming Context

minimize $c_0(x)$ over all $x \in S$ satisfying $c_i(x) \leq 0, \; i = 1, \dots, m$

Linear programming problem:

$$c_i(x) = a_{i1}x_1 + \cdots a_{in}x_n - b_i$$

minimize $a_{01}x_1 + \cdots a_{0n}x_n - b_0$ over $x = (x_1, \dots, x_n) \in S$
subject to $a_{i1}x_1 + \cdots a_{in}x_n - b_i \leq 0$ for $i = 1, \dots, m$,
where $S = \{x \,|\, x_1 \geq 0, \dots, x_n \geq 0 \text{ \& other conditions?}\}$

Effect of uncertainty:

$$c_i(x, \omega) = a_{i1}(\omega)x_1 + \cdots a_{in}(\omega)x_n - b_i(\omega)$$

There is **no single clear answer** to the question of how then to reconstitute the optimization **objective** and the **constraints**!

Stochastic Framework — Random Variables

Future state space Ω modeled with a probability structure:

$$(\Omega, \mathcal{F}, P), \quad P = \text{some probability measure}$$

Functions $X : \Omega \to R$ are interpreted as **random variables:**

cumulative distribution function $F_X : (-\infty, \infty) \to [0, 1]$

$$F_X(z) = \mathrm{prob}\,\{\omega \mid X(\omega) \leq z\}$$

expected value EX = mean value = $\mu(X)$

variance $\sigma^2(X) = E[(X - \mu(X))^2]$, standard deviation $\sigma(X)$

technical restriction imposed here: $X \in \mathcal{L}^2$ meaning $E[X^2] < \infty$

The functions $\underline{c}_i(x) : \omega \to c_i(x, \omega)$ are placed now in this picture:

choosing $x \in S$ yields **random variables** $\underline{c}_0(x), \underline{c}_1(x), \ldots, \underline{c}_m(x)$

Some Traditional Approaches

Recapturing optimization in the face of $\underline{c}_i(x) : \omega \to c_i(x, \omega)$

Approach 1: guessing the future

- identify $\bar{\omega} \in \Omega$ as the "best estimate" of the future
- minimize over $x \in S$:

$$c_0(x, \bar{\omega}) \text{ subject to } c_i(x, \bar{\omega}) \leq 0, \; i = 1, \ldots, m$$

- pro/con: simple and attractive, but dangerous—no hedging

Approach 2: worst-case analysis, "robust" optimization

- focus on the worst that might come out of each $\underline{c}_i(x)$:
- minimize over $x \in S$:

$$\sup_{\omega \in \Omega} c_0(x, \omega) \text{ subject to } \sup_{\omega \in \Omega} c_i(x, \omega) \leq 0, \; i = 1, \ldots, m$$

- pro/con: avoids probabilities, but expensive—maybe infeasible

Approach 3: relying on means/expected values

- focus on average behavior of the random variables $\underline{c}_i(x)$
- minimize over $x \in S$:

$$\mu(\underline{c}_0(x)) = E_\omega c_0(x,\omega) \text{ subject to}$$
$$\mu(\underline{c}_i(x)) = E_\omega c_i(x,\omega) \leq 0, \; i = 1,\dots,m$$

- pro/con: common for objective, but foolish for constraints?

Approach 4: safety margins in units of standard deviation

- improve on expectations by bringing standard deviations into consideration
- minimize over $x \in S$: for some choice of coefficients $\lambda_i > 0$

$$\mu(\underline{c}_0(x)) + \lambda_0\,\sigma(\underline{c}_0(x)) \text{ subject to}$$
$$\mu(\underline{c}_i(x)) + \lambda_i\,\sigma(\underline{c}_i(x)) \leq 0, \; i = 1,\dots,m$$

- pro/con: looks attractive, but a serious flaw will emerge

Approach 5: specifying probabilities of compliance

- choose probability levels $\alpha_i \in (0,1)$ for $i = 0, 1, \ldots, m$
- find lowest z such that, for some $x \in S$, one has

$$\text{prob}\left\{\underline{c}_0(x) \le z\right\} \ge \alpha_0,$$
$$\text{prob}\left\{\underline{c}_i(x) \le 0\right\} \ge \alpha_i \text{ for } i = 1, \ldots, m$$

- pro/con: popular and appealing, but flawed and controversial
  - no account is taken of the seriousness of violations
  - technical issues about the behavior of these expressions

Example: with $\alpha_0 = 0.5$, the **median** of $\underline{c}_0(x)$ would be minimized

Traditional usage: problems of **reliable design** in engineering

Quantification of Risk

How can the "risk" be measured in a random variable X?

orientation: $X(\omega)$ stands for a "cost" or loss

negative costs correspond to gains/rewards

The idea to be pursued here:

capture the "risk" in X by a **numerical surrogate** $\mathcal{R}(X)$

This leads to considering

functionals $\mathcal{R} : X \to \mathcal{R}(X)$ on the space of random variables

$\mathcal{R}$ = "risk quantifier" = "risk measure"

A Systematic Approach to Uncertainty in Optimization

When numerical values $c_i(x)$ become random variables $\underline{c}_i(x)$:

- choose risk quantifiers $\mathcal{R}_i$ for $i = 0, 1, \ldots, m$
- define the functions $\bar{c}_i$ on $\boldsymbol{R}^n$ by $\bar{c}_i(x) = \mathcal{R}_i(\underline{c}_i(x))$, and then
- minimize $\bar{c}_0(x)$ over $x \in S$ subject to $\bar{c}_i(x) \leq 0$, $i = 1, \ldots, m$.

Basic Guidelines

What axioms for numerical surrogates $\mathcal{R}(X) \in (-\infty, \infty]$?

Definition of coherency

$\mathcal{R}$ is a **coherent measure of risk** in the **basic** sense if

(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)

$\mathcal{R}$ coherent in the **extended** sense: axiom (R5) dropped

(from ideas of Artzner, Delbaen, Eber, Heath 1997/1999)

(R1)+(R2) $\Rightarrow$ $\mathcal{R}(X + C) = \mathcal{R}(X) + C$ for all X and constants C

(R2)+(R5) $\Rightarrow$ $\mathcal{R}(X + X') \leq \mathcal{R}(X) + \mathcal{R}(X')$ (subadditivity)

Associated Criteria for Risk Acceptability

For a "cost" random variable X, to what extent should outcomes $X(\omega) > 0$, in constrast to outcomes $X(\omega) \leq 0$, be tolerated?

preferences must be articulated!

Preference-based definition of acceptance

Given a choice of a risk measure $\mathcal{R}$:

the risk in X is deemed **acceptable** when $\mathcal{R}(X) \leq 0$

from (R1): $\mathcal{R}(X) \leq c \iff \mathcal{R}(X - c) \leq 0$

from (R3): $\mathcal{R}(X) \leq \sup X$ for all X,

so X is always acceptable when $\sup X \leq 0$

The Role of Coherency in Optimization

Reconstituted optimization problem:

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 $\underline{c}_i(x) : \omega \to c_i(x, \omega)$

Assumption for now: each $\mathcal{R}_i$ is **coherent** in the **basic** sense

Key properties associated with coherency

(a) (preservation of convexity)

$c_i(x, \omega)$ convex in x $\Longrightarrow$ $\bar{c}_i(x)$ convex in x

(b) (preservation of certainty)

$c_i(x, \omega)$ independent of ω $\Longrightarrow$ $\bar{c}_i(x)$ has that same value

(c) (insensitivity to scaling)

optimization is unaffected by rescaling of the units of the c_i's

(a) and (b) still hold for coherent measures in the extended sense

Coherency or Its Lack in Traditional Approaches

The case of Approach 1: guessing the future

$\mathcal{R}_i(X) = X(\bar{\omega})$ for a choice of $\bar{\omega} \in \Omega$ with prob > 0

$\mathcal{R}_i$ is **coherent**—but open to criticism

$\underline{c}_i(x)$ is deemed to be risk-acceptable if merely $c_i(x, \bar{\omega}) \leq 0$

The case of Approach 2: worst case analysis

$\mathcal{R}_i(X) = \sup X$

$\mathcal{R}_i$ is **coherent**—but very conservative

$\underline{c}_i(x)$ is risk-acceptable only if $c_i(x, \omega) \leq 0$ with prob $= 1$

The case of Approach 3: relying on expectations

$\mathcal{R}_i(X) = \mu(X) = EX$

$\mathcal{R}_i$ is **coherent**—but perhaps too "feeble"

$\underline{c}_i(x)$ is risk-acceptable as long as $c_i(x, \omega) \leq 0$ on average

The case of Approach 4: standard deviation units as safety margins

$\mathcal{R}_i(X) = \mu(X) + \lambda_i \sigma(X)$ for some $\lambda_i > 0$

$\mathcal{R}_i$ is **not coherent**: the monotonicity axiom (R3) fails!

$\Longrightarrow$ $\underline{c}_i(x)$ could be deemed more costly than $\underline{c}_i(x')$ even though $c_i(x,\omega) < c_i(x',\omega)$ with probability 1

$\underline{c}_i(x)$ is risk-acceptable as long as the mean $\mu(\underline{c}_i(x))$ lies below 0 by at least λ_i times the amount $\sigma(\underline{c}_i(x))$

The case of Approach 5: specifying probabilities of compliance

$\mathcal{R}_i(X) = q_{\alpha_i}(X)$ for some $\alpha_i \in (0,1)$, where

$q_{\alpha_i}(X) = \alpha_i$-quantile in the distribution of X

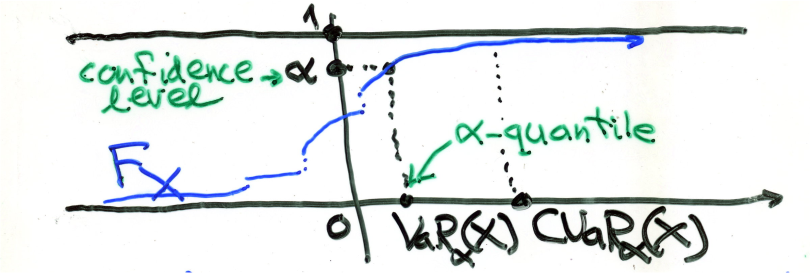
(to be explained)

$\mathcal{R}_i$ is **not coherent**: the convexity axiom (R2) fails!

$\Longrightarrow$ for portfolios, this could run counter to "diversification"

$\underline{c}_i(x)$ is risk-acceptable as long as $c_i(x,\omega) \leq 0$ with prob $\geq \alpha_i$

Quantiles and Conditional Value-at-Risk



α-quantile for X: $q_\alpha(X) = \min\{z \mid F_X(z) \geq \alpha\}$

value-at-risk: $\mathrm{VaR}_\alpha(X)$ same as $q_\alpha(X)$

conditional value-at-risk: $\mathrm{CVaR}_\alpha(X) =$ α-tail expectation of X

$$= \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\beta(X)d\beta \;\geq\; \mathrm{VaR}_\alpha(X)$$

THEOREM $\mathcal{R}(X) = \mathrm{CVaR}_\alpha(X)$ is a **coherent** measure of risk!

$\mathrm{CVaR}_\alpha(X) \nearrow \sup X$ as $\alpha \nearrow 1$, $\mathrm{CVaR}_\alpha(X) \searrow EX$ as $\alpha \searrow 0$

CVaR Versus VaR in Modeling

$$\text{prob}\{X \le 0\} \le \alpha \iff q_\alpha(X) \le 0 \iff \text{VaR}_\alpha(X) \le 0$$

Approach 5 recast: specifying probabilities of compliance

- focus on value-at-risk for the random variables $\underline{c}_i(x)$
- minimize $\text{VaR}_{\alpha_0}(\underline{c}_0(x))$ over $x \in S$ subject to
  $\text{VaR}_{\alpha_i}(\underline{c}_i(x)) \le 0, \; i = 1, \ldots, m$
- pro/con: seemingly natural, but "incoherent" in general

Approach 6: safeguarding with conditional value-at-risk

- conditional value-at-risk instead of value-at-risk for each $\underline{c}_i(x)$
- minimize $\text{CVaR}_{\alpha_0}(\underline{c}_0(x))$ over $x \in S$ subject to
  $\text{CVaR}_{\alpha_i}(\underline{c}_i(x)) \le 0, \; i = 1, \ldots, m$
- pro/con: coherent! also more cautious than value-at-risk

extreme cases: "$\alpha_i = 0$" $\sim$ expectation, "$\alpha_i = 1$" $\sim$ supremum

Minimization Formula for VaR and CVaR

$$\mathrm{CVaR}_\alpha(X) = \min_{C \in R} \left\{ C + \frac{1}{1-\alpha} E\left[\max\{0, X - C\} \right] \right\}$$

$$\mathrm{VaR}_\alpha(X) = \text{lowest } C \text{ in the interval giving the min}$$

Application to CVaR optimization: convert a problem like

$$\text{minimize } \mathrm{CVaR}_{\alpha_0}(\underline{c}_0(x)) \text{ over } x \in S \text{ subject to}$$
$$\mathrm{CVaR}_{\alpha_i}(\underline{c}_i(x)) \le 0, \quad i = 1, \ldots, m$$

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

$$\text{minimize } C_0 + \frac{1}{1-\alpha_0} E\left[\max\{0, \underline{c}_0(x) - C_0\} \right] \text{ while requiring}$$
$$C_i + \frac{1}{1-\alpha_i} E\left[\max\{0, \underline{c}_i(x) - C_i\} \right] \le 0, \quad i = 1, \ldots, m$$

Further Modeling Possibilities

Coherency-preserving combinations of risk measures

(a) If $\mathcal{R}_1, \ldots, \mathcal{R}_r$ are coherent and $\lambda_1 > 0, \ldots, \lambda_r > 0$ with $\lambda_1 + \cdots + \lambda_r = 1$, then

$$\mathcal{R}(X) = \lambda_1 \mathcal{R}_1(X) + \cdots + \lambda_r \mathcal{R}_r(X) \text{ is coherent}$$

(b) If $\mathcal{R}_1, \ldots, \mathcal{R}_r$ are coherent, then

$$\mathcal{R}(X) = \max\left\{\mathcal{R}_1(X), \ldots, \mathcal{R}_r(X)\right\} \text{ is coherent}$$

Example: $\mathcal{R}(X) = \lambda_1 \mathrm{CVaR}_{\alpha_1}(X) + \cdots + \lambda_r \mathrm{CVaR}_{\alpha_r}(X)$

Approach 7: safeguarding with CVaR mixtures

The CVaR approach already considered can be extended by replacing single CVaR expressions with weighted combinations

Continuous CVaR Mixtures and Risk Profiles

For any nonnegative **weighting** measure λ on $(0,1)$, a coherent measure of risk (in the basic sense) is given by

$$\mathcal{R}(X) = \int_0^1 \mathrm{CVaR}_\alpha(X)\, d\lambda(\alpha)$$

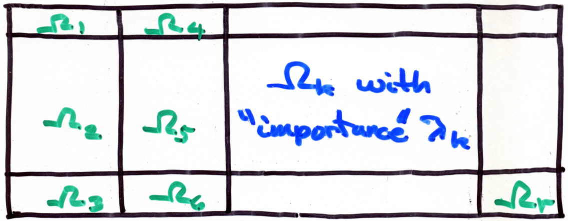
Spectral representation

Associate with λ the **profile** function $\varphi(\alpha) = \int_0^\alpha [1-\beta]^{-1}\, d\lambda(\beta)$
Then, as long as $\varphi(1) < \infty$, the above $\mathcal{R}$ has the expression

$$\mathcal{R}(X) = \int_0^1 \mathrm{VaR}_\beta(X)\varphi(\beta)\, d\beta$$

Risk Measures From Subdividing the Future

"robust" optimization modeling revisited with Ω subdivided



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

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

Approach 8: distributed worst-case analysis

Extend the ordinary worst-case model

$$\text{minimize } \sup_{\omega \in \Omega} c_0(x, \omega) \text{ subject to } \sup_{\omega \in \Omega} c_i(x, \omega) \leq 0 \,, i = 1, \ldots, m$$

by **distributing** each supremum **over subregions** of Ω, as above

Some References

[1] R. T. Rockafellar (2007), "Coherent approaches to risk in optimization under uncertainty," Tutorials in Operations Research INFORMS 2007, 39–61.

[2] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath (1999), "Coherent measures of risk," *Mathematical Finance* 9, 203–227.

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

[4] R.T. Rockafellar, S.P. Uryasev (2000), " Optimization of Conditional Value-at-Risk," *Journal of Risk* 2, 21–42.

[5] R.T. Rockafellar, S.P. Uryasev,, "Conditional value-at-risk for general loss distributions," *Journal of Banking and Finance* 26, 1443–1471.

[1], [4], [5], downloadable:

www.math.washington.edu/∼rtr/mypage.html