COHERENT APPROACHES TO RISK IN OPTIMIZATION UNDER UNCERTAINTY

Terry Rockafellar
University of Florida, Gainesville
University of Washington, Seattle

INFORMS tutorial: Seattle, 5 November 2007

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 and solved?

How can "risk" properly be taken into account, with attention paid to the attitudes of the optimizer?

How should the future, where the essential uncertainty resides, be modeled with respect to decisions and information?

The Fundamental Difficulty Caused by Uncertainty

with simple modeling of the future

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: $\underline{c}_i(x) : \omega \mapsto c_i(x, \omega), \quad i = 0, 1, \ldots, m$

Optimization objectives and constraints must be reconstrued in terms of such function, but how? There is no universal answer...

Various approaches: old/new? good/bad? yet to be discovered?
Adaptations to attitudes about "risk"?

Example: Linear Programming Context

Problem without uncertainty: $c_i(x) = a_{i1}x_1 + \cdots a_{in}x_n - b_i$

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

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

Portfolio illustration with financial instruments $j = 1, \ldots, n$

$r_j(\omega)$ = rate of return, x_j = weight in the portfolio
portfolio rate of return $= x_1 r_1(\omega) + \cdots + x_n r_n(\omega)$

Constraints: $x \in S = \{(x_1, \ldots, x_n) \mid x_j \ge 0,\ x_1 + \cdots + x_n = 1\}$

Uncertain ingredients to incorporate in optimization model:

$c_0(x, \omega) = -[x_1 r_1(\omega) + \cdots + x_n r_n(\omega)]$
(conversion to "cost" orientation for minimization)

$c_1(x, \omega) = q(\omega) - [x_1 r_1(\omega) + \cdots + x_n r_n(\omega)]$, q = benchmark
(shortfall below benchmark, desired outcome ≤ 0)

Probabilistic Framework — Random Variables

Future state space Ω modeled with a probability structure:

$(\Omega, \mathcal{A}, P)$, P =probability measure

"true"? "subjective"? or merely for reference?

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

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

$$F_X(z) = P\{\omega \mid X(\omega) \leq z\}$$

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

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

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

Corresponding convergence criterion as $k = 1, 2, \ldots \infty$:

$$X_k \to X \iff \mu(X_k - X) \to 0 \text{ and } \sigma(X_k - X) \to 0$$

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 $\underline{c}_0(x), \underline{c}_1(x), \ldots, \underline{c}_m(x)$

No-Distinction Principle for Objectives and Constraints

Is there an intrinsic reason why uncertainty/risk in an objective should be treated differently than uncertainty/risk in a constraint?

NO, because of well known, elementary reformulations

Given an optimization problem in standard format:

$$\text{minimize } c_0(x) \text{ over } x \in S \text{ with } c_i(x) \le 0,\ i = 1, \ldots, m$$

augment $x = (x_1, \ldots, x_n)$ by another variable x_{n+1}, and in terms of

$\tilde{x} = (x, x_{n+1}) \in \tilde{S} = S \times \mathbb{R}$,
$\tilde{c}_i(\tilde{x}) = c_i(x)$ for $i = 1, \ldots, m$,
$\tilde{c}_0(\tilde{x}) = x_{n+1}$, $\quad \tilde{c}_{m+1}(\tilde{x}) = c_0(x) - x_{n+1}$

pass equivalently to the reformulated problem:

$$\text{minimize } \tilde{c}_0(\tilde{x}) \text{ over } \tilde{x} \in \tilde{S} \text{ with } \tilde{c}_i(\tilde{x}) \le 0,\ i = 1, \ldots, m, m+1$$

Uncertainty in $c_0, c_1, \ldots, c_m$ will not affect the objective with $\tilde{c}_0$.
It will only affect the constraints with $\tilde{c}_1, \ldots, \tilde{c}_m, \tilde{c}_{m+1}$.

Some Traditional Approaches

Aim: recapturing optimization in the face of $\underline{c}_i(x) : \omega \to c_i(x, \omega)$
each approach followed uniformly, for emphasis in illustration

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 $x \in S$:

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

- 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 $\lambda_i > 0$

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

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

The idea here: find the lowest z such that, for some $x \in S$,
$\underline{c}_0(x) - z, \underline{c}_1(x), \ldots, \underline{c}_m(x)$ will be ≤ 0 except in λ_i-upper tails

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

$$P\{\underline{c}_0(x) \leq z\} \geq \alpha_0, \quad P\{\underline{c}_i(x) \leq 0\} \geq \alpha_i \text{ for } i = 1, \ldots, m$$

- 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

Additional modeling ideas:

- Staircased variables: $\underline{c}_i(x)$ propagated to $\underline{c}_i^k(x) = \underline{c}_i(x) - d_i^k$ for a series of thresholds d_i^k, $k = 1, \ldots, r$ with different compliance conditions placed on having these "subvariables" $\underline{c}_i^k(x)$ be ≤ 0
- Expected penalty expressions like $E[\,\psi(\underline{c}_0(x))\,]$
- Stochastic programming, dynamic programming

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

- Idea 1: assess the "risk" in X in terms of how **uncertain** X is:

  $\longrightarrow$ **measures $\mathcal{D}$ of deviation from constancy**

- Idea 2: capture the "risk" in X by a **numerical surrogate** for overall cost/loss: $\longrightarrow$ **measures $\mathcal{R}$ of potential loss**

$\longrightarrow$ our concentration, for now, will be on Idea 2

A General Approach to Uncertainty in Optimization

In the context of the numerical values $c_i(x) \in \boldsymbol{R}$ being replaced by random variables $\underline{c}_i(x) \in \mathcal{L}^2$ for $i = 0, 1, \ldots, m$:

- choose measures $\mathcal{R}_i$ of the risk of potential loss,
- 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

For a functional $\mathcal{R}$ that assigns to each random "cost" $X \in \mathcal{L}^2$ a numerical surrogate $\mathcal{R}(X) \in (-\infty, \infty]$, what axioms?

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)

$\mathcal{R}$ is a coherent measure of risk in the *extended* sense if it satisfies (R1)–(R4), not necessarily (R5)

(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?

There is no single answer—this has to depend on **preferences!**

Preference-based definition of acceptance

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

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

(examples to come will illuminate this concept of Artzner et al.)

Notes:

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 $\sup X \leq 0$

(i.e., when there is **no chance** of an outcome $X(\omega) > 0$)

Consequences of Coherency for Optimization

For $i = 0, 1, \ldots, m$ let $\mathcal{R}_i$ be a coherent measure of risk in the basic sense, and consider the reconstituted 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 $\underline{c}_i(x) : \omega \to c_i(x, \omega)$

Key properties

(a) (preservation of convexity) If $c_i(x, \omega)$ is convex with respect to x, then the same is true for $\bar{c}_i(x)$

(so convex programming models persist)

(b) (preservation of certainty) If $c_i(x, \omega)$ is a value $c_i(x)$ independent of ω, then $\bar{c}_i(x)$ is that same value

(so features not subject to uncertainty are left undistorted)

(c) (insensitivity to scaling) The optimization problem 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

Assessing the risk in each $\underline{c}_i(x)$ as $\mathcal{R}_i(\underline{c}_i(x))$ for a choice of $\mathcal{R}_i$

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 $c_i(x, \bar{\omega}) \leq 0$ 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 $c_i(x, \bar{\omega}) \leq 0$ 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)

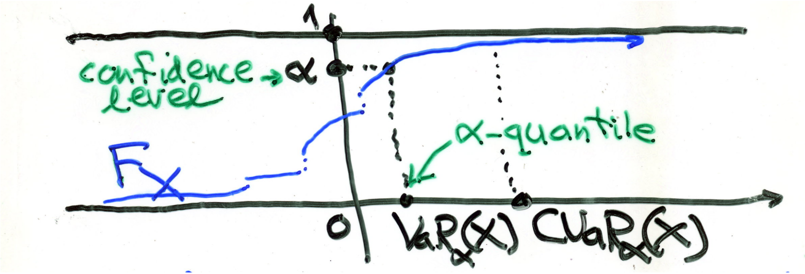
$\mathcal{R}_i$ is **not coherent**: the convexity axiom (R2) fails!

$\Longrightarrow$ 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 $\geq \alpha_i$

What further alternatives, remedies?

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: $\mathrm{CVaR}_\alpha(X) = \alpha$-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

$$P\{X \leq 0\} \leq \alpha \iff q_\alpha(X) \leq 0 \iff \mathrm{VaR}_\alpha(X) \leq 0$$

Approach 5 recast: specifying probabilities of compliance

- focus on value-at-risk for the random variables $\underline{c}_i(x)$
- minimize $\mathrm{VaR}_{\alpha_0}(\underline{c}_0(x))$ over $x \in S$ subject to
  $\mathrm{VaR}_{\alpha_i}(\underline{c}_i(x)) \leq 0, \ i = 1, \ldots, m$
- 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 $\mathrm{CVaR}_{\alpha_0}(\underline{c}_0(x))$ over $x \in S$ subject to
  $\mathrm{CVaR}_{\alpha_i}(\underline{c}_i(x)) \leq 0, \ i = 1, \ldots, m$
- pro/con: coherent! also more cautious than value-at-risk

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

Some Elementary Portfolio Examples

securities $j = 1, \ldots, n$ with rates of return $\underline{r}_j$ and weights x_j

$$S = \{x = (x_1, \ldots, x_n) \,|\, x_j \geq 0,\ x_1 + \cdots + x_n = 1\}$$

rate of return of x-portfolio: $\underline{r}(x) = -[x_1 \underline{r}_1 + \cdots + x_n \underline{r}_n]$

$\underline{c}_0(x) = -\underline{r}(x)$, $\underline{c}_1(x) = \underline{q} - \underline{r}(x)$ with $\underline{q} \equiv -0.04$ here

Problems 1(a)(b)(c): expectation objective, CVaR constraints

(a) minimize $E[\underline{c}_0(x)]$ over $x \in S$
(b) minimize $E[\underline{c}_0(x)]$ over $x \in S$ subject to $\mathrm{CVaR}_{0.8}(\underline{c}_1(x)) \leq 0$
(b) minimize $E[\underline{c}_0(x)]$ over $x \in S$ subject to $\mathrm{CVaR}_{0.9}(\underline{c}_1(x)) \leq 0$

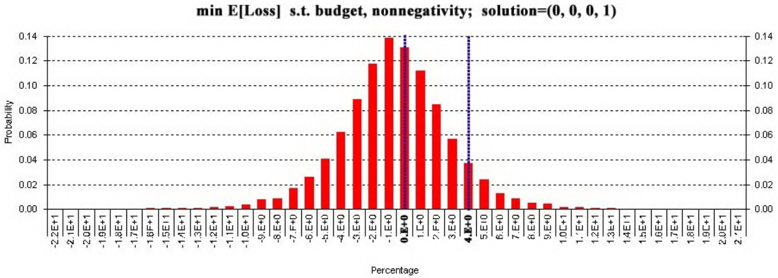
Problems 2(a)(b)(c): CVaR objectives, no benchmark constraints

(a) minimize $E[\underline{c}_0(x)]$ over $x \in S$ $\quad E[\underline{c}_0(x)] = \mathrm{CVaR}_{0.0}(\underline{c}_0(x))$
(b) minimize $\mathrm{CVaR}_{0.8}(\underline{c}_0(x))$ over $x \in S$
(c) minimize $\mathrm{CVaR}_{0.9}(\underline{c}_0(x))$ over $x \in S$

Portfolio Rate-of-Loss Contours, Problems 1(a)(b)(c)

Solutions computed with *Portfolio Safeguard* software, available for evaluation from American Optimal Decisions www.AOrDa.com

Results for Problem 1(a)

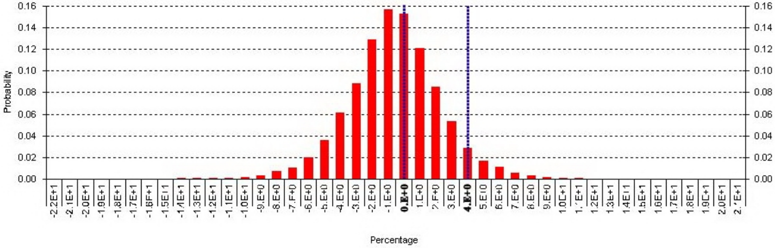


Solution vector: the portfolio weights for four different stocks

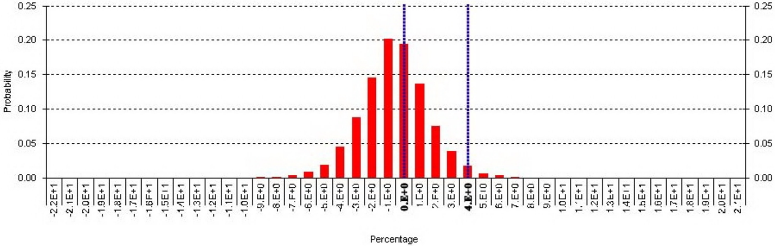
Note that in this case all the weight goes to the risky fourth stock

Results for Problems 1(b) and 1(c)

min E[Loss] s.t. CVaR{80%}(Loss) <=0.04, budget, nonnegativity; solution=(0.17, 0.04, 0.18, 0.61)



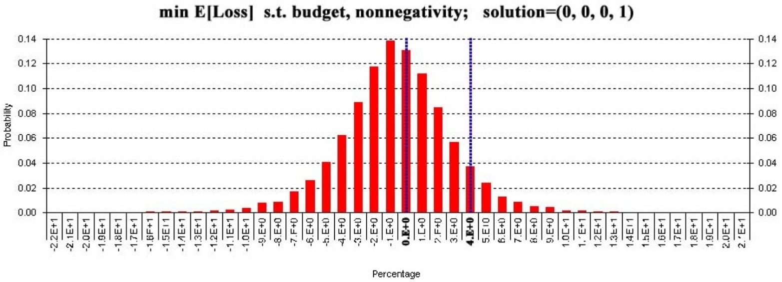
min E[Loss] s.t. CVaR{90%}(Loss) <=0.04, budget, nonnegativity; solution=(0.44, 0.40, 0.09, 0.07)



Portfolio Rate-of-Loss Contours, Problems 2(a)(b)(c)

Solutions computed with *Portfolio Safeguard* software, available for evaluation from American Optimal Decisions www.AOrDa.com

Results for Problem 2(a), same as Problem 1(a)

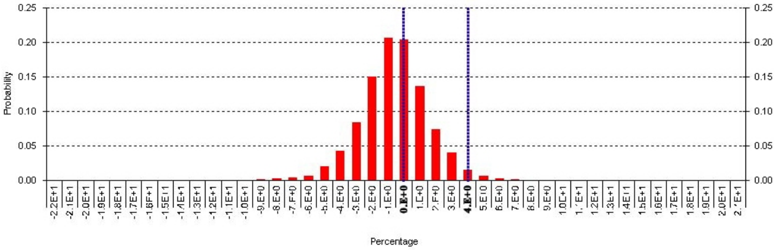


Solution vector: the portfolio weights for four different stocks

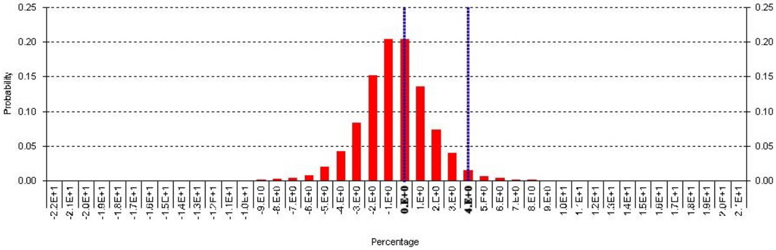
Again, in this case all the weight goes to the risky fourth stock

Results for Problems 2(b) and 2(c)

min CVaR{80%}(Loss) s.t. budget, nonnegativity; solution=(0.47, 0.53, 0, 0)



min CVaR{90%}(Loss) s.t. budget, nonnegativity; solution=(0.49, 0.51, 0, 0)



Minimization Formula for VaR and CVaR

$$\mathrm{CVaR}_\alpha(X) = \min_{C\in R}\Big\{ C + \tfrac{1}{1-\alpha} E\Big[\max\{0, X - C\}\Big]\Big\}$$
$$\mathrm{VaR}_\alpha(X) = \text{lowest } C \text{ in the interval giving the min}$$

min values behave better parametrically than minimizing points!

Application to CVaR optimization: convert a problem like

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

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

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

Important case: this converts to **linear programming** when

(1) each $c_i(x, \omega)$ depends linearly on x,

(2) the future state space Ω is finite

(as is common in financial modeling, for instance)

Further Modeling Possibilities

additional sources of coherent measures of risk

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

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

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

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

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

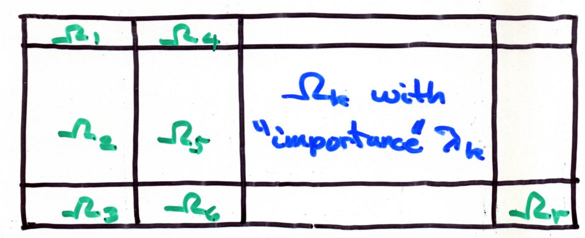
Approach 7: safeguarding with CVaR mixtures

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

"Continuous sums" are available too and relate to "risk profiles"

Risk Measures From Subdividing the Future

"robust" optimization modeling revisited with Ω subdivided



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

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

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

Safety Margins Reconsidered

Traditional approach to an expected cost EX being safely below 0:

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

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

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

General deviation measures for quantifying uncertainty

$\mathcal{D}$ is a **measure of deviation** (in the *basic* sense) if

(D1) $\mathcal{D}(X) = 0$ for $X \equiv C$ constant, $\mathcal{D}(X) > 0$ otherwise

(D2) $\mathcal{D}((1-\lambda)X + \lambda X') \leq (1-\lambda)\mathcal{D}(X) + \lambda\mathcal{D}(X')$ for $\lambda \in (0,1)$ (convexity)

(D3) $\mathcal{D}(X) \leq c$ when $X_k \to X$ with $\mathcal{D}(X_k) \leq c$ (closedness)

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

It is a **coherent** measure of deviation if it also satisfies

(D5) $\mathcal{D}(X) \leq \sup X - EX$ for all X

deviation measures in the *extended* sense: (D4) dropped

Risk Measures Paired With Deviation Measures

$\mathcal{R}$ is a **loss-averse** measure of risk if it satisfies the axioms for coherency, except perhaps (R3) (monotonicity), and

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

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

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

and moreover $\mathcal{R}$ is coherent $\iff$ $\mathcal{D}$ is coherent

Approach 9: safety margins with coherency

- replace standard deviation by coherent deviation measures $\mathcal{D}_i$
- minimize over $x \in S$: for some choice of coefficients $\lambda_i > 0$

$$E[\underline{c}_0(x)] + \lambda_0 \mathcal{D}_0(\underline{c}_0(x)) \text{ subject to}$$

$$E[\underline{c}_i(x)] + \lambda_i \mathcal{D}_i(\underline{c}_i(x)) \leq 0, \; i = 1, \ldots, m$$

- pro/con: coherency in the model has been restored

Risk Envelope Characterization of Coherency

A subset $\mathcal{Q}$ of $\mathcal{L}^2$ is a **coherent risk envelope** is 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 underlying probability measure P on Ω of an alternative probability measure P' on Ω : $\quad E_{P'}[X] = E[XQ]$

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

THEOREM: There is a one-to-one correspondence $\mathcal{R} \leftrightarrow \mathcal{Q}$ between coherent measures of risk $\mathcal{R}$ (in the basic sense) and coherent risk envelopes Q, which is furnished by the relations

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

Some examples: $\quad \mathcal{R}(X) = EX \;\leftrightarrow\; \mathcal{Q} = \{1\}$

$\mathcal{R}(X) = \sup X \;\leftrightarrow\; \mathcal{Q} = \left\{ \text{all } Q \geq 0 \text{ with } EQ = 1 \right\}$

$\mathcal{R}(X) = \mathrm{CVaR}_\alpha(X) \;\leftrightarrow\; \mathcal{Q} = \left\{ \text{all } Q \geq 0 \text{ with } Q \leq \alpha^{-1},\; EQ = 1 \right\}$