SAFEGUARDING IN OPTIMIZATION UNDER UNCERTAINTY

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

a common situation:

- decision must be taken now
- consequences not known until later

uncertainty tied to the future
→ risk!

the challenge for optimization:

- how to understand constraints and objectives?
- how to characterize solutions and compute them?

safeguarding:
approaches for handling risks, related to "safety margins"

- engineering design
- systems management
- financial planning ....

UNCERTAINTY IN OPTIMIZATION

decisions $\longrightarrow$ uncertain consequences

standard problem with certainty:

choose $x \in S \subset \mathbb{R}^n$ to

minimize $f_0(x)$ subject to $f_i(x) \leq 0$, $i = 1, \ldots$

basic uncertainty:

Ω = space of future states ω

$\leftarrow$ probability structure $(\Omega, \mathcal{F}, P)$

$\begin{cases} \text{choose } x \in S \text{ now} \\ \text{observe } \omega \in \Omega \text{ later} \end{cases}$

$\longrightarrow f_i(x, \omega)$ for $i = 0, 1, \ldots, m$

decision x only determines

<u>functions</u>: $\underbrace{\omega \longmapsto f_i(x, \omega)}_{\underset{\sim}{f}_i(x) \text{ "random variable"}}$

How should a problem of optimization be formulated in this context?

How to think about objectives, constraints

TRADITIONAL APPROACHES

1. guessing the future:
   $\bar{\omega}$ = "estimate" of future state in Ω
   minimize $f_0(x,\bar{\omega})$ subject to
   $f_i(x,\bar{\omega}) \le 0,\ i=1,\dots,m$
   common, but dangerous

2. relying on expectations:
   minimize $E_\omega\{f_0(x,\omega)\}$ subject to
   $E_\omega\{f_i(x,\omega)\} \le 0,\ i=1,\dots,m$
   focuses on long-range behavior

3. worst-case analysis:
   minimize $\sup_\omega f_0(x,\omega)$ subject to
   $\sup_\omega f_i(x,\omega) \le 0,\ i=1,\dots,m$
   pessimistic, maybe infeasible

4. probabilistic constraints:
   find lowest α such that
   $\text{prob}\{f_0(x,\omega) \le \alpha\} \ge \pi_0$
   $\text{prob}\{f_i(x,\omega) \le 0\} \ge \pi_i,\ i=1,\dots,m$
   $\pi_0, \pi_1, \dots, \pi_m$ = selected probabilities
   $\longrightarrow$ technical troubles,
   conceptual controversies

A GENERAL FRAMEWORK

random variables $X : \Omega \to \mathbb{R}$

$$X \in \mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{A}, P)$$

orientation:

$X(\omega)$ = a kind of "cost" incurred in state ω

lower = better, higher = worse

risk functionals: $\mathcal{R} : \mathcal{L}^1 \to (-\infty, \infty]$

$\mathcal{R}(X)$ = single number assigned to the random variable X

a "numerical surrogate" for X

axioms for "coherency":

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

(R2) $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ for $\lambda > 0$

(R3) $\mathcal{R}(X_1 + X_2) \leq \mathcal{R}(X_1) + \mathcal{R}(X_2)$

(R4) $\mathcal{R}(X_1) \leq \mathcal{R}(X_2)$ when $X_1 \leq X_2$ a.s.

(R5) $\{X \mid \mathcal{R}(X) \leq 0\}$ is closed in $\mathcal{L}^1$

$\Longrightarrow$ $\mathcal{R}$ is convex, positively homogeneous, lower semicontinuous

$$\mathcal{R}(0) = 0, \quad \mathcal{R}(X + C) = \mathcal{R}(X) + C$$

coherency: from Artzner et al., 1999 (with small modifications)

OPTIMIZATION MODEL

context: decision $x \in S \rightarrow$ random variable $\underset{\sim}{f}_i(x) \in \mathcal{L}^1$

$$\underset{\sim}{f}_i(x) : \omega \mapsto f_i(x,\omega)$$

problem formulation:

minimize $F_0(x) := \mathcal{R}_0(\underset{\sim}{f}_0(x))$ subject to

$$F_i(x) := \mathcal{R}_i(\underset{\sim}{f}_i(x)) \leq 0, \; i = 1, \dots, m$$

$\mathcal{R}_0, \mathcal{R}_1, \dots, \mathcal{R}_m$ = selected risk functionals

PROPOSITION 1:

If $f_i(x,\omega)$ is convex in x, then

$F_i(x) = \mathcal{R}_i(\underset{\sim}{f}_i(x))$ is convex function of

$\longrightarrow$ convexity is preserved, when present

PROPOSITION 2:

If $f_i(x,\omega) \equiv f_i(x)$ (no uncertainty), the

$F_i(x) = \mathcal{R}_i(\underset{\sim}{f}_i(x)) = f_i(x)$

$\longrightarrow$ certainty is preserved, when present

INITIAL EXAMPLES

seen through risk functionals

1. guessing the future: $\bar{\omega} \in \Omega$ = discrete, finite set

$$f_i(x,\bar{\omega}) = \mathcal{R}(\underset{\sim}{f}_i(x)) \text{ for}$$
$$\mathcal{R}(\underset{\sim}{X}): \underset{\sim}{X} \longrightarrow \underset{\sim}{X}(\bar{\omega})$$
satisfies axioms R1–R5

2. relying on expectations:

$$E_\omega\{f_i(x,\omega)\} = \mathcal{R}(\underset{\sim}{f}_i(x)) \text{ for}$$
$$\mathcal{R}(\underset{\sim}{X}): \underset{\sim}{X} \to E\underset{\sim}{X}$$
satisfies axioms R1–R5

3. worst-case analysis:

$$\sup_\omega f_i(x,\omega) = \mathcal{R}(\underset{\sim}{f}_i(x)) \text{ for}$$
$$\mathcal{R}(\underset{\sim}{X}): \underset{\sim}{X} \to \sup \underset{\sim}{X}$$ "essential" su (maybe ∞)

satisfies axioms R1–R5

4. probabilistic constraints:

$$\text{prob}\{f_i(x,\omega) \le 0\} \ge \pi_i \iff \mathcal{R}_i(\underset{\sim}{f}_i(x)) \le 0$$
for $\mathcal{R}_i(\underset{\sim}{X}) = VaR_{\pi_i}(\underset{\sim}{X})$ "value-at-risk"

satisfies axioms R1, R2, R4, R5, but not R3

convexity lacking, not preserved

VARIANCE EXAMPLE

related to Markovitz model in finance

$\sigma(X)$ = standard deviation

$\lambda_0, \lambda_1, \dots, \lambda_m$ = parameters > 0

minimize $E[\underset{\sim}{f}_0(x)] + \lambda_0 \sigma(\underset{\sim}{f}_0(x))$

subject to $E[\underset{\sim}{f}_i(x)] + \lambda_i \sigma(\underset{\sim}{f}_i(x)) \leq 0$, $i = 1 \dots$

"safeguarding with deviation margins"

(again, constraints and objective could really be treated differently)

This corresponds to choosing:

$$\mathcal{R}_i(X) = EX + \lambda_i \sigma(X)$$

Here, $\mathcal{R}_i$ satisfies axioms R1, R2, R3, R5 but not R4!

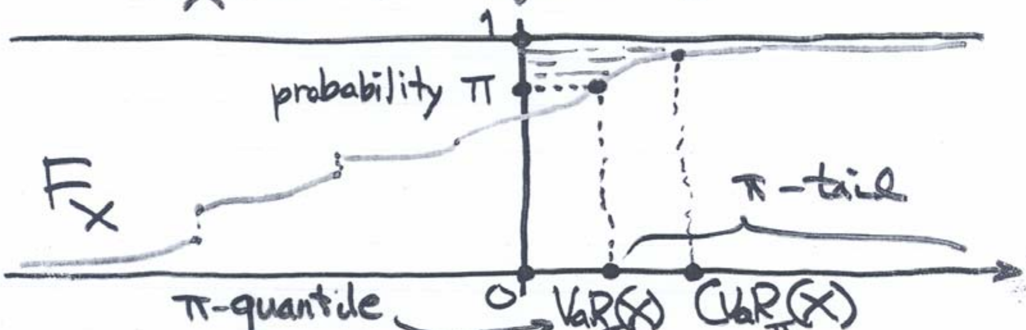
$\hookrightarrow$ such modeling is "incoherent" (more later)

CONDITIONAL VALUE-AT-RISK

Rockafellar & Uryasev, 2000 & 2002

$F_X : (-\infty, \infty) \to [0,1]$ distribution function for random variable $X \in \mathcal{L}^1$

$F_X(z) = \text{prob}\{X \leq z\}$



Definition 1: $CVaR_\pi(X)$ = conditional expectation of X subject to $X \geq VaR_\pi(X)$

needs refinement if F_X has jump at $VaR_\pi(X)$

Definition 2: $CVaR_\pi(X) = \frac{1}{1-\pi} \int_\pi^1 VaR_\tau(X)\, d\tau$

Observation: $CVaR_\pi(X) \leq 0 \Rightarrow \underbrace{\text{prob}\{X \leq 0\} \geq \pi}_{VaR_\pi(X) \leq 0}$

THEOREM $\mathcal{R}(X) = CVaR_\pi(X)$ is a risk functional satisfying R1-R4

coherent, convexity preserving!

CVaR VERSUS VaR

improvement of prob. constraints

constraint comparisons:

- $VaR_{\pi_i}(\underset{\sim}{f}_i(x)) \le 0$ means
  $f_i(x,\omega) \le 0$ with probability π_i
- $CVaR_{\pi_i}(\underset{\sim}{f}_i(x)) \le 0$ means
  $f_i(x,\omega) \le 0$ with probability π_i
  <u>and</u> even in the remaining cases one has $f_i(x,\omega) \le 0$ on average

objective comparisons:

- minimizing $VaR_{\pi_0}(\underset{\sim}{f}_0(x))$ means finding lowest α such that $f_0(x,\omega) \le \alpha$ with probability π
- minimizing $CVaR_{\pi_0}(\underset{\sim}{f}_0(x))$ means finding lowest α such that $f_0(x,\omega) \le 0$ with probability π
  <u>and</u> even in the remaining cases, one has $f_i(x,\omega) \le \alpha$ on average

MINIMIZATION SHORTCUT

$$CVaR_\pi(X) = \min_{z\in\mathbb{R}} \left\{ z + \frac{1}{1-\pi} E_\omega\{\max[0, X(\omega)-z]\} \right\}$$

$$VaR_\pi(X) = \underset{z\in\mathbb{R}}{\operatorname{argmin}} \{ \text{--- - same - ---} \}$$

take lowest point if not unique

this explains the different behaviors!

application to optimization:

$$\begin{cases} \underset{x\in S}{\text{minimize}}: CVaR_{\pi_0}(\underset{\sim}{f}_0(x)) \text{ subject to} \\ \quad CVaR_{\pi_i}(\underset{\sim}{f}_i(x)) \le 0, \; i=1\ldots m \end{cases}$$

$$\begin{cases} \underset{\substack{x\in S \\ z_0, z_1, \ldots z_m \in\mathbb{R}}}{\text{minimize}}: z_0 + \frac{1}{1-\pi_0} E_\omega\{\max[0, f_0(x,\omega)-z_0]\} \\ \quad \text{subject to} \\ \quad z_i + \frac{1}{1-\pi_i} E_\omega\{\max[0, f_i(x,\omega)-z_i]\} \le 0 \\ \quad i=1\ldots m \end{cases}$$

→ converts to linear programming when $\begin{cases} \Omega \text{ is discrete, finite} \\ f_i(x,\omega) \text{ is linear in } x \end{cases}$

→ the values of $CVaR_{\pi_i}(\underset{\sim}{f}_i(x))$ are computed automatically as part of the optimization.

STOCHASTIC DOMINANCE
connections with VaR, CVaR

first-order dominance: $\precsim_1$

$$X \precsim_1 Y \iff E[u(X)] \geq E[u(Y)]$$
for all nonincreasing utilities u

second-order dominance: $\precsim_2$

$$X \precsim_2 Y \iff E[u(X)] \geq E[u(Y)]$$
for all [nonincreasing concave] utilities

$$X \leq Y \Rightarrow X \precsim_1 Y \Rightarrow X \precsim_2 Y$$

THEOREM

$$X \precsim_1 Y \iff VaR_\pi(X) \leq VaR_\pi(Y) \text{ for all}$$
$$X \precsim_2 Y \iff CVaR_\pi(X) \leq CVaR_\pi(Y) \text{ for all}$$

consistency of $\mathcal{R}$ with respect to $\precsim_2$:

$\mathcal{R}(X) \leq \mathcal{R}(Y)$ in particular if $X \precsim_2 Y$

otherwise one could have $\mathcal{R}(X) > \mathcal{R}(Y)$ even though $E[u(X)] \geq E[u(Y)]$ for all nonincreasing concave utilities u./

conclusion:

this holds $\iff$ $\mathcal{R}(X)$ depends only on the function $\pi \mapsto CVaR_\pi(X$

satisfied by $\mathcal{R}(X) = CVaR_\pi(X)$, but not by $\mathcal{R}(X) = VaR_\pi(X$

MORE RISK FUNCTIONALS

mixed CVaR:

- $\mathcal{R}(X) = \sum_{k=1}^{r} \lambda_k \, CVaR_{\pi_k}(X)$ (satisfies all axioms)
  $\lambda_k \geq 0, \ \sum_{k=1}^{r} \lambda_k = 1$
- $\mathcal{R}(X) = \int_0^1 CVaR_\pi(X) \, d\lambda(\pi)$ ← weighting measure

  related to "risk profiles" in dual utility theory

general rules:

- $\mathcal{R}_k$ coherent for $k = 1 \dots r$
  $\Longrightarrow \mathcal{R}(X) = \sum_{k=1}^{r} \lambda_k \mathcal{R}_k(X)$ coherent
  $\lambda_k \geq 0, \ \sum_{k=1}^{r} \lambda_k = 1$
- $\mathcal{R}_i$ coherent for $i \in I$ (index set)
  $\Longrightarrow \mathcal{R}(X) = \sup_{i \in I} \mathcal{R}_i(X)$ coherent

limit cases:

- $CVaR_\pi(X) \longrightarrow \sup X$ as $\pi \nearrow 1$
- $CVaR_\pi(X) \longrightarrow EX$ as $\pi \searrow 0$

consistency with stochastic dominance $\succsim_2$:

DUALITY

representation of coherent risk functionals through convex analysis

Definition: risk envelope $\mathcal{Q}$
= a nonempty subset of $\mathcal{L}^\infty = \mathcal{L}^\infty(\Omega, \mathcal{F}, P)$
satisfying
(Q1) $\mathcal{Q}$ is closed and convex
(Q2) every $Q \in \mathcal{Q}$ has $Q \geq 0$, $E[$

interpretation:
each $Q \in \mathcal{Q}$ is a probability density relative to the underlying prob. measure P

$$E[XQ] = \int_\Omega X(\omega) \underbrace{Q(\omega)\, dP(\omega)}_{dP'(\omega)}$$

alternat
prob. mea

THEOREM

$$\left\{\begin{array}{l}\text{risk functionals } \mathcal{R} \\ \text{satisfying R1–R5}\end{array}\right\} \xleftrightarrow{\text{one-to-one}} \left\{\begin{array}{l}\text{risk envelopes } \\ \text{satisfying Q1–}\end{array}\right.$$

$$\mathcal{R}(X) = \sup_{Q \in \mathcal{Q}} E[XQ]$$

$$\mathcal{Q} = \{Q \in \mathcal{L}^\infty \mid E[XQ] \leq \mathcal{R}(X), \forall X \in \mathcal{L}^1\}$$

interpretation:
$\mathcal{R}(X)$ = worst expectation $E_{P'}[X]$
over the alternative probabilit
measures P' corresponding to
the densities $Q \in \mathcal{Q}$

$Q = \frac{dP'}{dP}$

RISK ENVELOPE EXAMPLES

1. relying on expectations:

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

$$\mathcal{R}(X) = E_{P'}[X] \iff \mathcal{Q} = \{\frac{dP'}{dP}\}$$

2. worst-case analysis:

$$\mathcal{R}(X) = \sup X \iff \mathcal{Q} = \{\text{all } Q \geq 0 \text{ with } EQ$$

all prob. densities

3. conditional-value-at-risk:

$$\mathcal{R}(X) = CVaR_\pi(X) \iff \mathcal{Q} = \{Q \mid 0 \leq Q \leq \frac{1}{\pi}, EQ =$$

some general rules:

$$\mathcal{R} = \lambda_1 \mathcal{R}_1 + \cdots + \lambda_m \mathcal{R}_m \iff \mathcal{Q} = \lambda_1 \mathcal{Q}_1 + \cdots + \lambda_m \mathcal{Q}_m$$

$$\lambda_k \geq 0, \ \lambda_1 + \cdots + \lambda_m = 1$$

$$\mathcal{R} = \max\{\mathcal{R}_1, \ldots, \mathcal{R}_m\} \iff \mathcal{Q} = \mathrm{conv}\{\mathcal{Q}_1 \cup \cdots \cup \mathcal{Q}_m$$

$\longrightarrow$ subgradients, dual problems, characterizations of optimality

GENERALIZED DEVIATIONS

a different aspect of risk

deviation functionals $\mathcal{D}: \mathcal{L}^1 \to [0,\infty]$

$\mathcal{D}(X)$ = an assessment of the uncertainty in the random variable X

axioms for generalized deviation:

(D1) $\mathcal{D}(C)=0$ for constants C

(D2) $\mathcal{D}(\lambda X)=\lambda\mathcal{D}(X)$ for $\lambda>0$

(D3) $\mathcal{D}(X_1+X_2)\le\mathcal{D}(X_1)+\mathcal{D}(X_2)$

(D4) $\mathcal{D}(X)>0$ for nonconstant X

(D5) $\{X \mid \mathcal{D}(X)\le 1\}$ is closed in $\mathcal{L}^1$

classical example: $\mathcal{D}(X)=\sigma(X)$ or $\lambda\sigma(X)$

THEOREM

the relations $\begin{cases}\mathcal{D}(X)=\mathcal{R}(X-EX)\\ \mathcal{R}(X)=EX+\mathcal{D}(X)\end{cases}$

give a one-to-one correspondence between

- $\mathcal{D}$ satisfying (D1)–(D5) and also
  $\mathcal{D}(X)\le \sup X - EX$
- $\mathcal{R}$ satisfying (R1)–(R5) and also
  $\mathcal{R}(X)\le \sup X$

fails for $\mathcal{D}(X)=\lambda\sigma(X)$
(incoherency)

GENERALIZED ERRORS

error-functionals $\mathcal{E} : \mathcal{L}^2 \to [0, \infty)$

$\mathcal{E}(X) =$ error relative to $X \neq 0$

axioms for generalized error:

(E1) $\mathcal{E}(0) = 0$, $\mathcal{E}(X) > 0$ for $X \neq C$

(E2) $\mathcal{E}(\lambda X) = \lambda \mathcal{E}(X)$ for $\lambda > 0$

(E3) $\mathcal{E}(X_1 + X_2) \leq \mathcal{E}(X_1) + \mathcal{E}(X_2)$

associated deviation:

$$\mathcal{D}(X) = \min_C \mathcal{E}(X - C)$$

associated "statistic":

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

examples:

$\mathcal{E}(X) = [E X^2]^{1/2}$, $\mathcal{D}(X) = \sigma(X)$, $\mathcal{S}(X) =$

$\mathcal{E}(X) =$ K-B error, $\mathcal{D}(X) = CVaR_\alpha(X -$

$\mathcal{S}(X) = VaR_\alpha(X)$

etc.

$\longrightarrow$ "generalized regression"

SOME REFERENCES

1. Rockafellar, Uryasev
"Conditional value-at-risk for general loss distributions"
Journal of Banking and Finance 26 (2002)

2. Rockafellar, Uryasev, Zabarankin
"Deviation measures in risk analysis"
Finance and Stochastics 9 (2005)

3. Rockafellar, Uryasev, Zabarankin
"Master funds in portfolio analysis with general deviation measures"
Journal of Banking and Finance 29 (2005)

4. Rockafellar, Uryasev, Zabarankin
"Optimality conditions in portfolio analysis with general deviation measures"
Mathematical Programming 68 (2005)

These papers can be downloaded from:
www.ise.ufl.edu/uryasev