Decision analysis and risk-informed optimization for CI resilience

Yi-Ping Fang, Professor

Risk Resilience Reliability (R3) Research Group

Industrial Engineering Laboratory

CentraleSupélec, Université Paris-Saclay, France

yiping.fang@centralesupelec.fr

5/8/2025

Motivation

Prescriptive
What to do?

Predictive
What will happen?

Descriptive
What happened? Why?

World & data → **Information** → **Decision**

Motivation

- What to do is often not straightforward, instead very complicated, in real industrial systems.

- Maintenance planning of a wind farm
  - → Many turbines
  - → Limited repair resources
  - → Maintenance (preventive, corrective, opportunistic) cost v.s. profit
  - → Decide maintenance time, spare parts
  - → Even with RUL estimated perfectly...

- Decision theory + Optimization

A bit advertisement

- Master *Operation Research and Risk Analytics* at CentraleSupélec
  - → **Related courses**: *optimization of complex decisions, stochastic optimization, decision making and preference modelling, predictive maintenance*
- Using math and data to solve problems and make smart decisions, especially when dealing with uncertainty and risks

informs
Smarter Decisions for a Better World

EURO
THE ASSOCIATION OF EUROPEAN OPERATIONAL RESEARCH SOCIETIES

Highest paying college majors in the U.S. in 2023
10 years after graduation

Source: CNBC

		10+ years experience	≤ 5 years experience
1.	Petroleum engineering	$212,500	$97,500
2.	Operations research + Industrial engineering	$191,800	$98,300
3.	Interaction design	$173,600	$74,700
4.	Applied economics + management	$164,400	$76,500
5.	Building science	$163,100	$69,000
6.	Actuarial mathematics	$160,000	$70,700
6.	Operations research	$160,000	$92,200
8.	Systems engineering	$159,100	$87,000
9.	Optical science + engineering	$158,300	$79,600
10.	Information + computer science	$157,800	$76,000

Outline

- Risk measures
- Stochastic decision-making and optimization

  → Optimization model introduction

  → Stochastic optimization formalization
- Sampled average approximation
- Risk-averse stochastic optimization

St. Petersburg paradox

- From Nicolas Bernoulli's letter 1713
- **St. Petersburg game:** flip a fair coin until it comes up heads the 1st time. At that point the player wins $\$2^n$, where n is the number of times the coin was flipped. How much should one pay for playing this game?

Trial where first tails appears	Probability	Payout
1	$1/2$	2
2	$1/4$	4
3	$1/8$	8
4	$1/16$	16
⋮	⋮	⋮

- Decision theorists' advice: the expected value

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} P(X = n) \cdot \text{Payout}(n) = \sum_{n=1}^{\infty} \left(\frac{1}{2^n} \times 2^n \right) = \sum_{n=1}^{\infty} 1 = \infty$$

St. Petersburg paradox

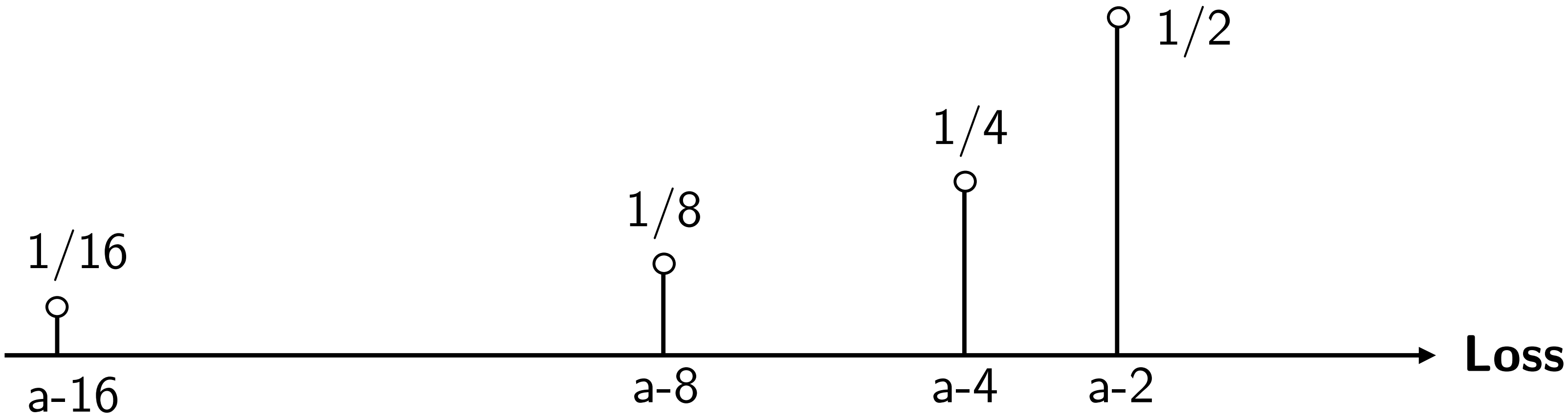
- The "paradox": it would be rational to pay any **finite** fee to play the game?
- Proposed resolutions, e.g., Daniel Bernoulli's expected utility

  $\rightarrow$ utility $= \log_{10}(2^n)$ (diminishing marginal utility)

  $\rightarrow$ expected utility $\approx \$4$

  $\rightarrow$ satisfactory?
- Why we have the paradox?

  $\rightarrow$ **Expectation neglects the risk of the (bad) outcomes**



Risk attitude

CentraleSupélec

- **Different attitudes** towards the risk of the outcomes

Game A:	Game B:
20% **+1000€**, 80% **−100 €**	100% **+120€**

- **Classifications**
  - → Risk averse fear loss and seek sureness
  - → Risk neutral are indifferent to the degree of risky outcome
  - → Risk seeking hope to "win big" and don't mind losing as much

Risk and risk measures

Risk:

The combination of the uncertainty of occurrence of “bad outcome”, and the severity of that outcome

- **Random outcomes**: r.v. $X : \Omega \to \mathbb{R}$

  $\to$ “Cost/loss” oriented X: high outcomes bad, low outcomes good

- **Risk measure**: a quantification be applied to X that elicits the level of “cost/loss” in X

$$\mathcal{R}(X) \text{ maps } X \text{ into } \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$$

- Risk $\neq$ Uncertainty (statistical dispersion)

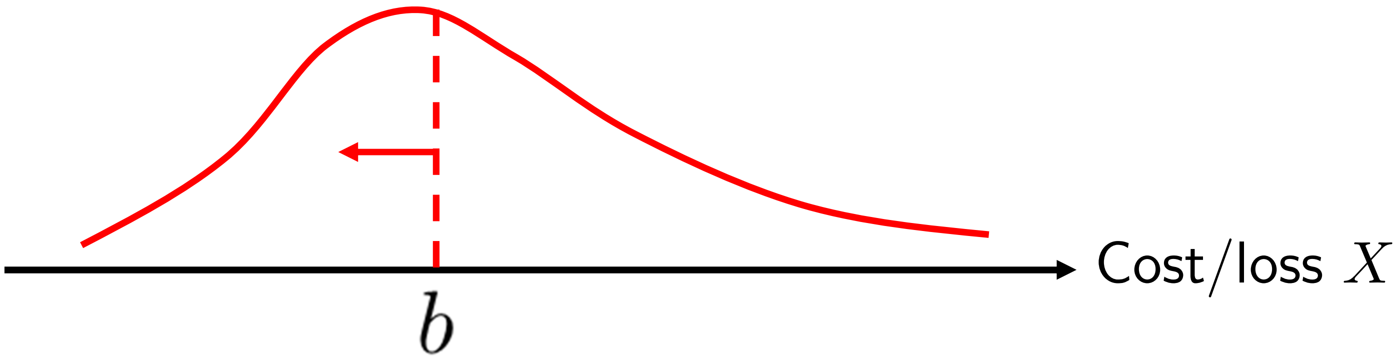
Some proposals

CentraleSupélec

- **Expectation**: $\mathcal{R}(X) = \mathbb{E}[X]$

  $\rightarrow$ $\mathcal{R}(X) \leq b \Leftrightarrow X \leq b$ on average

  $\rightarrow$ but not risk averse (risk measure?), perhaps too feeble



- **Focusing on worst cases**: $\mathcal{R}(X) = \sup X$

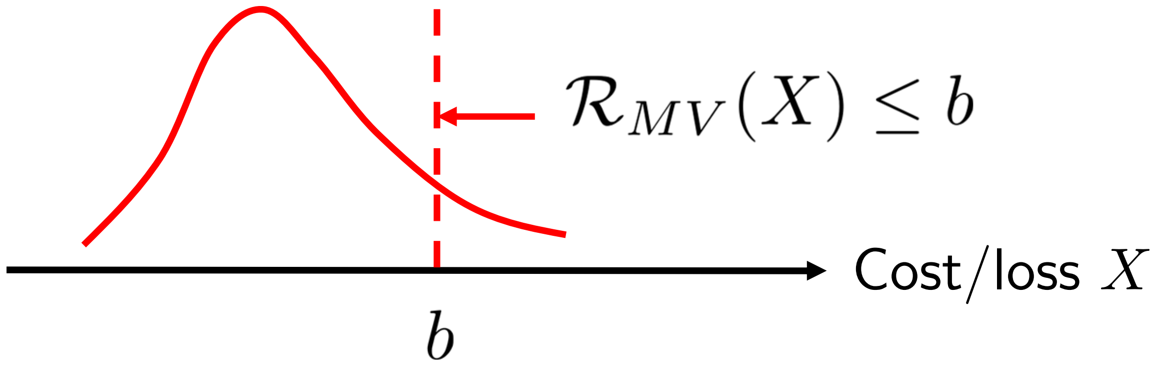
  $\rightarrow$ $\mathcal{R}(X) \leq b \Leftrightarrow X \leq b$ almost surely

  $\rightarrow$ Averse, perhaps overly conservative, infeasible

Some proposals

CentraleSupélec

- **Mean variance:** $\mathcal{R}_{MV}(X) = \mathbb{E}[X] + \lambda \mathbb{V}[X]$



$\rightarrow$ Widely used in finance (portfolio optimization, Harry Markowitz 1952)

$\rightarrow$ Possible drawback: variance is symmetric! penalizes high cost as well as low cost (profit?)

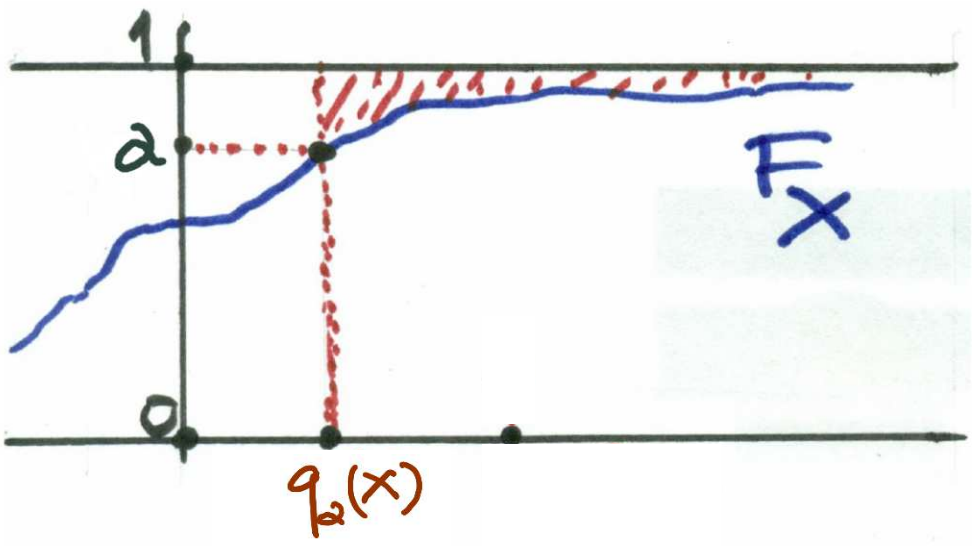
Value-at-Risk (VaR), α-quantile

- Loss/cost r.v. X associated with CDF $F(x)$
- For any $\alpha \in [0,1]$, the VaR on α is defined by

$$\mathbf{VaR}_\alpha(X) = q_\alpha(X) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$$

- For continuous CDF, $F(\mathbf{VaR}_\alpha) = \alpha$
- For strictly increasing and continuous CDF

$$\mathbf{VaR}_\alpha = F^{-1}(\alpha)$$



- Then

$$\mathcal{R}(X) = \mathbf{VaR}_\alpha(X) \leq b$$

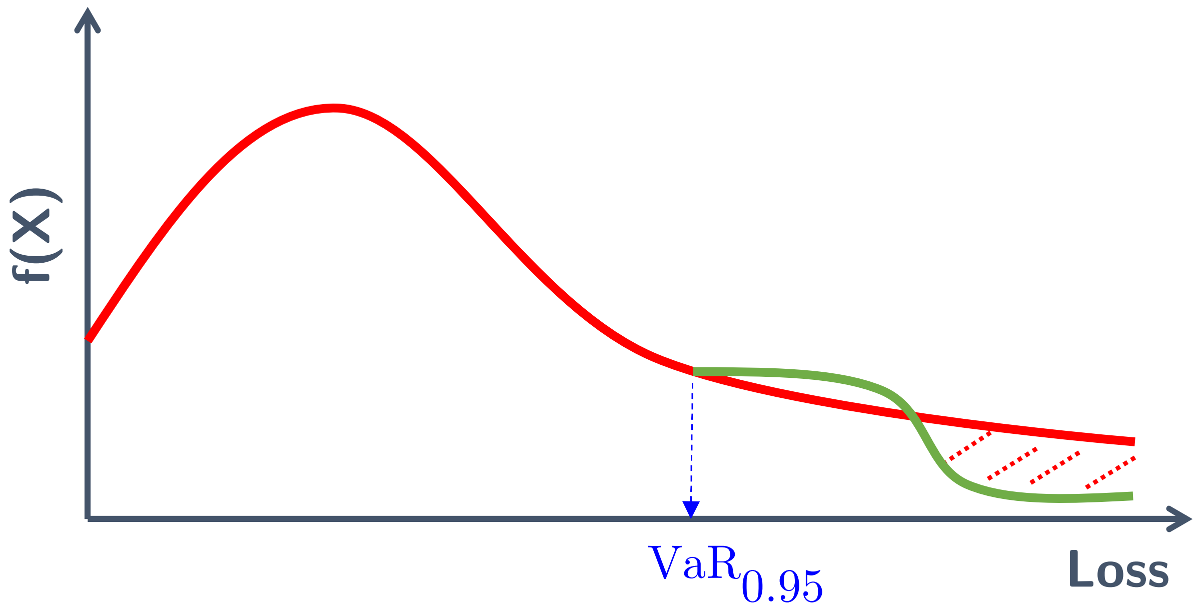
$$\Leftrightarrow \mathbb{P}(X \leq b) \geq \alpha$$

Value-at-Risk (VaR), α-quantile

CentraleSupélec

- Widely used (in Finance)
- VaR_α pays no attention to the magnitude of losses when the rare event of experiencing a loss above the level VaR_α occurs
- Bad mathematical & computational behavior: nonconvex, nonlinear

- J.P. Morgan 95%
- Bank of America 95%
- Citibank 95.4%
- Chase Manhattan 97.5%
- Basel Committee on Bank Supervision 99%



Conditional VaR, α-superquantile

- For $\alpha \in [0, 1]$, the CVaR on α is defined by

$$\mathbf{CVaR}_\alpha(X) = Q_\alpha(X) = \mathbb{E}[X | X \geq \mathbf{VaR}_\alpha(X)]$$

- Or, equivalently

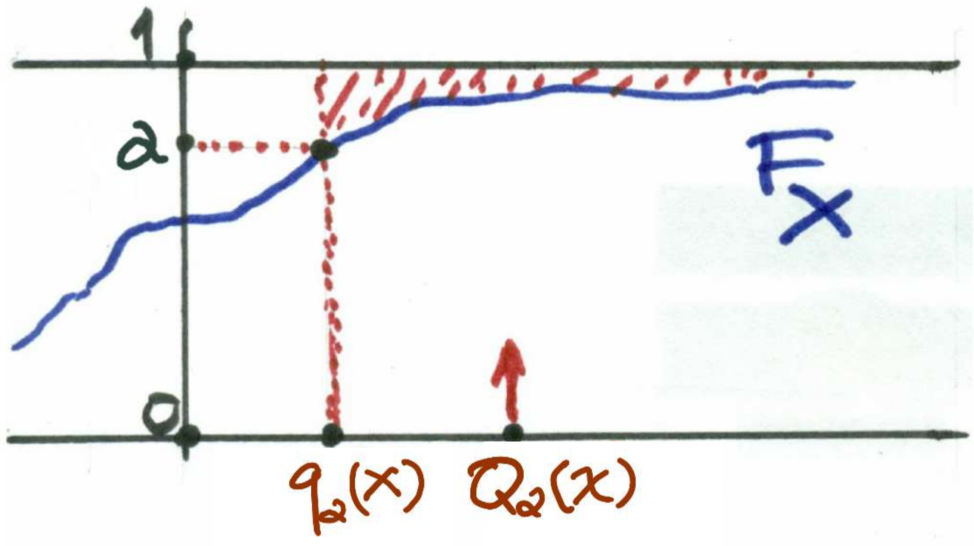
$$\mathbf{CVaR}_\alpha(X) = \frac{1}{1-\alpha} \int_{x \geq \mathbf{VaR}_\alpha} x \, dF(x)$$

- Good mathematical behaviors

- Measure the

$$\mathcal{R}(X) = Q_\alpha(X) \leq b$$

$\Leftrightarrow$ $X \leq b$ on average in upper α-tail



Coherent risk measures

Definition

A risk measure $\mathcal{R}(\cdot)$ is called coherent if it satisfies the following properties:

(A1) Convexity. $\beta \in (0,1)$, X_1 and X_2 random variables $\Rightarrow$

$$\mathcal{R}(\beta X_1 + (1-\beta)X_2) \leq \beta\mathcal{R}(X_1) + (1-\beta)\mathcal{R}(X_2)$$

(A2) Monotonicity. $X_1 \geq X_2$ a.s. $\Rightarrow \mathcal{R}(X_1) \geq \mathcal{R}(X_2)$

(A3) Translation invariance. If X is a r.v. and $a \in \mathbf{R}$, then

$$\mathcal{R}(X + a) = \mathcal{R}(X) + a$$

(A4) Positive homogeneity. $\mathcal{R}(tX) = t\mathcal{R}(X), \forall t > 0$

References: Artzner et al. (1999), Ruszczynski and Shapiro (2006)

Properties of risk measures

$$\mathbf{VaR}_\alpha(X) = \inf\{x \in \mathbf{R} : F(x) \geq \alpha\}$$

- Convex? No
- Monotone? Yes
- Translation invariant? Yes
- Positively homogeneous? Yes, if $\mathbf{P}(X \geq 0) = 1$

Mean variance $\mathcal{R}_{MV}(X) = \mathbb{E}[X] + \lambda \mathbb{V}[X]$

- Convex? Yes
- Monotone? No
- Translation invariant? Yes
- Positively homogeneous? No

Nonconvexity of VaR

- Suppose there are three equally likely outcomes (loss)

ω	Z^1	Z^2	$\frac{1}{2}Z^1 + \frac{1}{2}Z^2$
1	300	0	150
2	0	0	0
3	0	300	150
$\mathbf{VaR}_{0.6}$	0	0	150
$\mathbb{E}$	100	100	100

- $0.5\mathbf{VaR}_{0.6}(Z^1) + 0.5\mathbf{VaR}_{0.6}(Z^2) < \mathbf{VaR}_{0.6}(0.5Z^1 + 0.5Z^2)$

Nonmonotonicity of mean variance

- Suppose there are two outcomes: $\boldsymbol{P}(1) = 0.95, \boldsymbol{P}(2) = 0.05$

ω	Z^1	Z^2
1	10	10
2	10	0
$\mathbb{E}$	10	9.5
$\mathbb{V}$	0	4.75
$\mathbb{V}^{1/2}$	0	2.18

- $Z^1 \geq Z^2$ with probability 1, but mean variance (or mean standard-deviation) would prefer Z^1 for modest values of λ

Properties of risk measures

CVaR(·) is coherent (Rockafellar and Uryasev, 2000)

- Convex? Yes
- Monotone? Yes
- Translation invariant? Yes
- Positively homogeneous? Yes

Outline

CentraleSupélec

- Risk measures
- Stochastic decision-making and optimization

  → Optimization model introduction

  → Stochastic optimization formalization
- Sampled average approximation
- Risk-averse stochastic optimization

Stochastic DM problems for resilience

- Maintenance planning of wind farms
  → Maintenance time, spare parts
  → Uncertain lifetimes (RUL) of items

- Mitigate flooding risk of an NPP
  → Dike, what height?
  → Uncertain flooding level & outcomes

- Design resilient supply chain under unexpected perturbations (e.g., Covid)
  → Backup/secondary suppliers? inventory?
  → Uncertain perturbation type, frequency, outcomes

A simple example

- Use optimization to plan maintenance actions of a wind farm composed of 5 turbines based on their RUL forecasts

- Key question: "When should we schedule maintenance actions across time while minimizing cost?" Why:
  - → Some turbines are close to failure → act early
  - → Resources (technicians/equipment) are limited
  - → Other considerations in practice: opportunistic cost, overall system performance requirement, etc.

A simple example

- Problem setup: we manage 10 turbine over 8 time periods (e.g., weeks)
- Each turbine has
  - → a predicted RUL_i
  - → costs to maintain preventively c_i
  - → a higher cost if it fails $f_i > c_i$
- Constraints:
  - → max **2** maintenance slots per period
  - → Each is maintained **at most once** in the planning horizon
- Objectives: minimize total maintenance + failure costs

A simple example

- **Decision variables**: the quantities you can control or choose to achieve the best possible outcome

	$t = 1$	2	3	4	5	6	7	$T = 8$
M1	?	?	?	?	?	?	?	?
M2	?	?	?	?	?	?	?	?
…								
M10	?	?	?	?	?	?	?	?

$$x_{it} = \begin{cases} 1 & \text{if turbine } i \text{ is maintained in time } t \\ 0 & \text{otherwise} \end{cases}$$

- Binary decision variables

A simple example

- **Constraints**: on the decision variables
- Maintenance recourses limit per time

$$\sum_{i=1}^{10} x_{it} \leq 2,\ \forall t = 1, \ldots, 8$$

- Maintenance window: each turbine must be maintained at most once and only within its RUL

$$\sum_{t=1}^{\mathrm{RUL}_i} x_{it} \leq 1,\ \forall i = 1, \ldots, 10$$

A simple example

- **Objective function**: the quantities you can control or choose to achieve the best possible outcome

$$\min \sum_{i=1}^{10} \sum_{t=1}^{8} c_i x_{it} + \sum_{i=1}^{10} f_i \left(1 - \sum_{t=1}^{\mathsf{RUL}_i} x_{it} \right)$$

- First term = total maintenance cost
- Second term = total penalty for failure

A simple example

- The whole model:

$$
\begin{aligned}
\min_{x_{it}} \quad & \sum_{i=1}^{10} \sum_{t=1}^{8} c_i x_{it} + \sum_{i=1}^{10} f_i \left(1 - \sum_{t=1}^{\mathsf{RUL}_i} x_{it} \right) \\
\text{s.t.} \quad & \sum_{i=1}^{10} x_{it} \le 2, \; \forall t = 1, \ldots, 8 \\
& \sum_{t=1}^{\mathsf{RUL}_i} x_{it} \le 1, \; \forall i = 1, \ldots, 10 \\
& x_{it} \in \{0, 1\}, \; \forall i = 1, \ldots, 10, t = 1, \ldots, 8
\end{aligned}
$$

Optimization model concepts

CentraleSupélec

- **Key elements:**
  - → **Decision variables**: the quantities you can control or choose (e.g., maintenance time, height of a dike, supplier selection, etc.)
  - → **Constraints:** on decision variables (physical, operational, economical, etc.)
  - → **Objective function**: rewards (economic gain, system performance) or costs/loss
  - → Uncertainty: in evaluating the objective and the feasibility of constraints under different disruption scenarios

Optimization model concepts

- Deterministic mathematical optimization problem

$$\min_{x} \quad f_0(x)$$
$$\text{s.t.} \quad f_i(x) \leq 0,\ i = 1, \ldots, m$$

- Minimize a "cost" under constraints on the decision, which involve bounds on other "costs"
- $x := (x_1, x_2, \ldots, x_n) \rightarrow$ optimization variables
- $f_0(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ objective function
- $f_i(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \ldots, m$ constraint functions
- Feasible solution: x that satisfies the m constraints
- Optimal solution (set) $x^\star$ ($O^\star$) has the smallest objective value among all the feasible solutions

Model classification

- Often classified according to the types of the decision variables, constraints, and the objective function

$$\min_{x_{it}} \quad \sum_{i=1}^{10}\sum_{t=1}^{8} c_i x_{it} + \sum_{i=1}^{10} f_i\left(1 - \sum_{t=1}^{\mathrm{RUL}_i} x_{it}\right)$$

$$\text{s.t.} \quad \sum_{i=1}^{10} x_{it} \leq 2,\ \forall t = 1, \ldots, 8$$

$$\sum_{t=1}^{\mathrm{RUL}_i} x_{it} \leq 1,\ \forall i = 1, \ldots, 10$$

$$x_{it} \in \{0,1\},\ \forall i = 1, \ldots, 10, t = 1, \ldots, 8$$

- **Integer (binary) linear program**: decision variables are integers, constraint and objective functions are linear

Model classification

- Decision variable

  → Discrete (integer), continuous, mixed
- Unconstrained vs. constrained
- Constraint and objective function forms

  → Linear vs nonlinear

  → e.g., quadratic programming has quadratic objective function & linear constraints
- Convexity

  → Convex vs nonconvex optimization
- Deterministic vs. Stochastic
- **Example**: mixed integer quadratically constrained quadratic programs

Solving optimization problem

- **General optimization problem**

  → Very difficult to solve

  → Methods involve compromise, e.g., very long computation time, or not always finding the optimal solution

- **Convex optimization problems:** the clearest dividing line between efficiently solvable problems and numerical problems for which we have no hope to solve it easily, e.g.

  → Least square (LS) problems

  → Linear programming (LP) problems

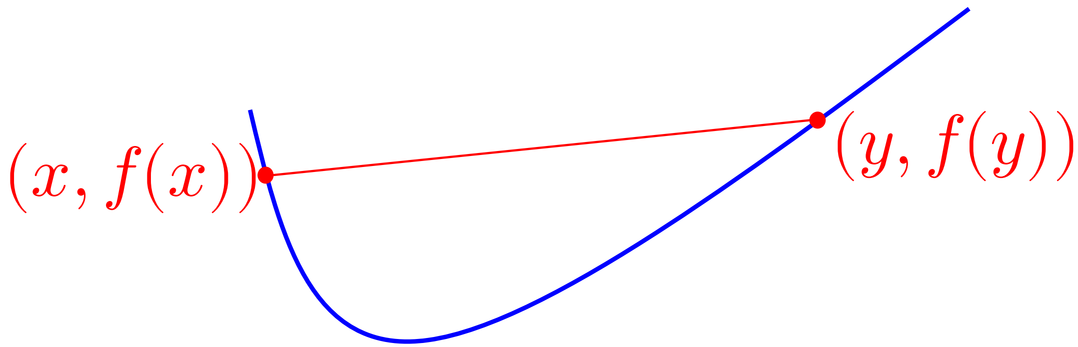
Convex optimization

$$\min_{x} \quad f_0(x)$$
$$\text{s.t.} \quad f_i(x) \leq 0, \; i = 1, \ldots, m$$

- **Objective and constraint functions** are convex:

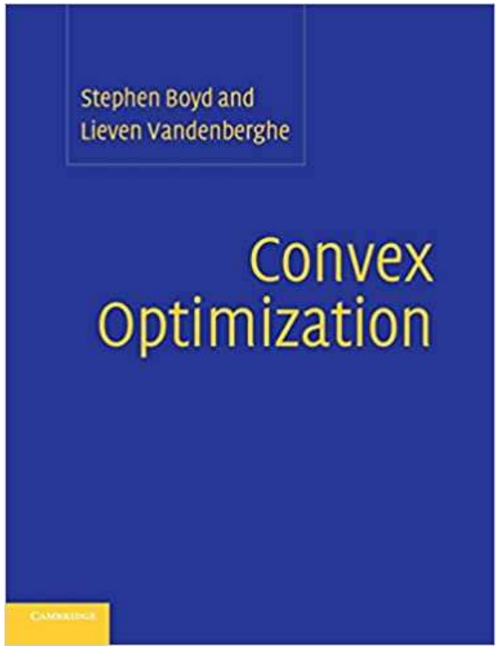
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

For all $x, y \in \mathbf{dom}\, f, \quad 0 \leq \theta \leq 1$



Convex optimization

- **Solving convex optimization**
  - → No analytical solution
  - → Reliable and efficient algorithms
  - → Almost a technology

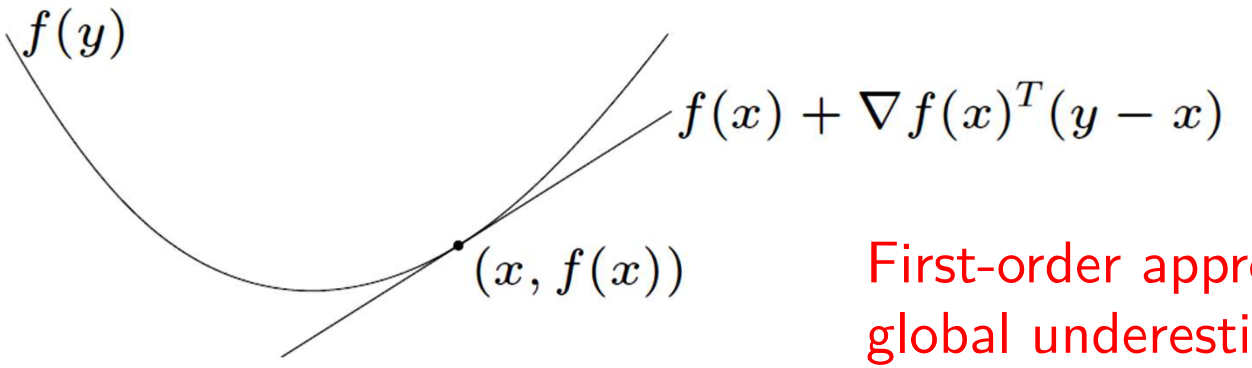


- **Using convex optimization**
  - → Often difficult to recognize
  - → Many tricks for transforming problems into convex form
  - → Surprisingly, many problems can be solved via convex optimization

The power of convexity

- **First-order condition**: differentiable f with convex domain is convex iff

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) \quad \text{for all } x, y \in \textbf{dom } f$$



First-order approximation of f is global underestimator

- **Theorem**: any locally optimal point of a convex problem is (globally) optimal

Least squares

$$\min_{x} \|Ax - b\|_2^2 = \sum_{i=1}^{k} (a_i^\top x - b_i)^2$$

- **Solving LS problems:**
  - $\rightarrow$ Analytical solution: $x^\star = (A^T A)^{-1} A^T b$
  - $\rightarrow$ Reliable and efficient algorithms and software
  - $\rightarrow$ A mature technology
- **Using LS:**
  - $\rightarrow$ LS problems are easy to recognize

Linear programming

$$\min_x \quad f_0(x) = c^\top x$$
$$\text{s.t.} \quad f_i(x) = a_i^\top x - b_i \leq 0, \ i = 1, \dots, m$$

- **Solving LP:**
  - → No analytical formula for solution
  - → Reliable and efficient algorithms and software
  - → A mature technology a mature technology, especially for problems of reasonable size
- **Using LP:**
  - → Not as easy to recognize as LS problems, e.g., problems involving piecewise-linear functions

Outline

CentraleSupélec

- Risk measures
- Stochastic decision-making and optimization
  - → Optimization model introduction
  - → Stochastic optimization formalization
- Sampled average approximation
- Risk-averse stochastic optimization

Optimization under uncertainty

- A stand optimization problem with uncertainty

$$\min_{x} \quad f_0(x, \xi)$$
$$\text{s.t.} \quad f_i(x, \xi) \leq 0, \; i = 1, \ldots, m$$

- Decision x must be made with uncertain parameter ξ
- Objective $f_0(x, \xi)$ cost/loss function (to be minimized), depending on our decision x, and the realization of uncertain nature ξ
- Same to the constraint functions $f_i(x, \xi)$

Optimization under uncertainty

- Example:
  - $\rightarrow$ x the height of a dike to be constructed
  - $\rightarrow$ $\xi =$ heights of the possible flooding
  - $\rightarrow$ $f_0(x, \xi)$ the economic loss associated with a given (x, ξ)

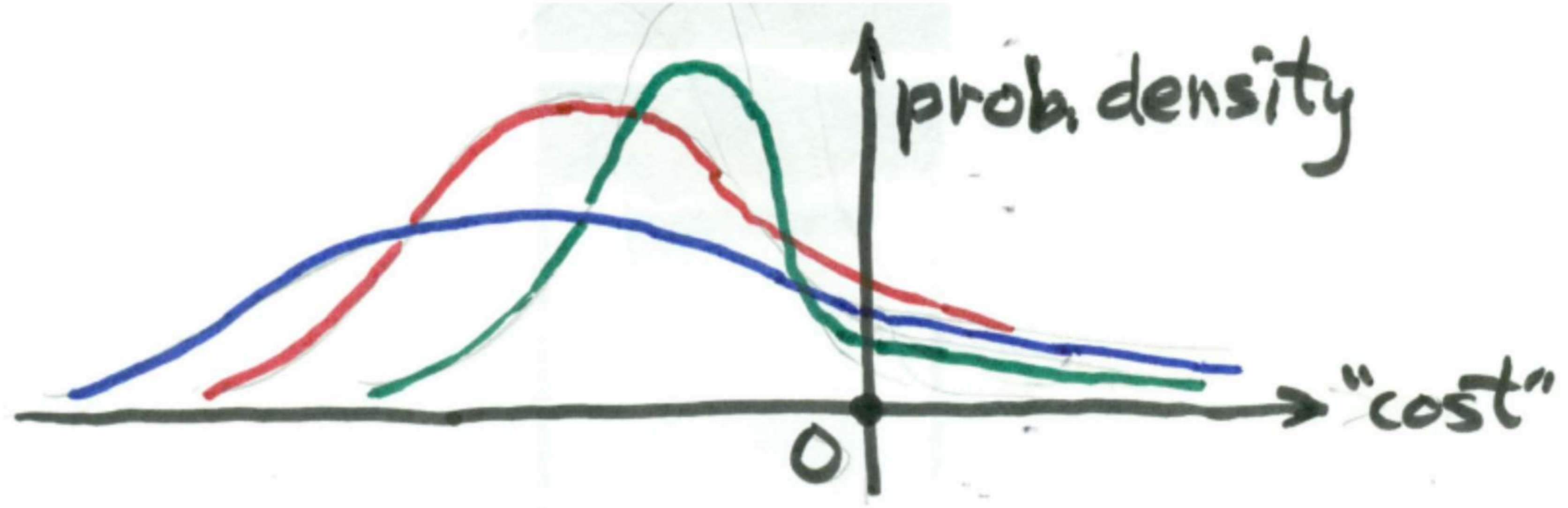
- How to model the uncertainty ξ?

- The Stochastic Programming (SP) approach
  - $\rightarrow$ $\xi \in \Xi =$ random variable with known probability law
  - $\rightarrow$ Then, $f_i(x, \xi)$ are random variables for any given x

Formulation challenges

Key issue in problem formulation

- The distribution of $f_i(x, \xi)$ can only be shaped by the choice of x
- How then can constraints and minimization be understood?



- Outcome < 0, if any, correspond to "reward/profit"

Remedy: risk measures

CentraleSupélec

Systematic prescription

- Articulate r.v. $f(x, \xi)$ numerically as $\mathcal{R}[f(x, \xi)]$ for a chosen risk measure $\mathcal{R}(\cdot)$

- **Constraints:** keeping $f_i(x, \xi)$ "acceptably" $\leq b_i$

  modeled as: constraint $\mathcal{R}[f_i(x, \xi)] \leq b_i$

- **Objective:** keeping $l_0(x, \xi)$ "as reasonably low as possible"

  modeled as: minimizing $\mathcal{R}[f_0(x, \xi)]$

- **Key assumption:** the probability distribution **P** of ξ is known (e.g., estimated from historical data)

SP with the expectation

- **Objective** is small on average

$$\min_x F_0(x) = \mathbb{E}[f_0(x, \xi)]$$

- **Constraints** are satisfied on average

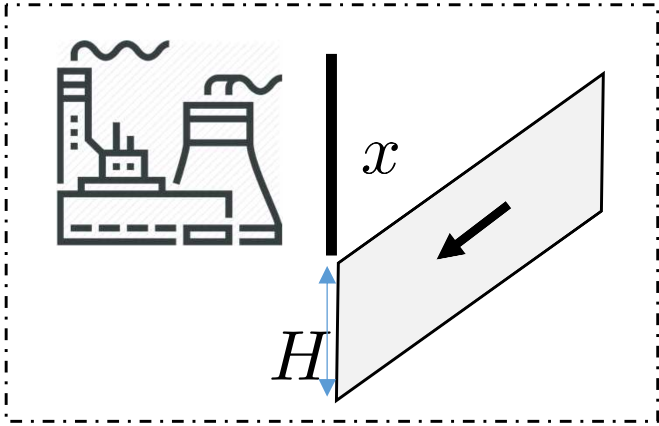
$$F_i(x) = \mathbb{E}[f_i(x, \xi)] \leq 0, \; i = 1, \ldots, m$$

- The expectation is suitable for "soft" constraint, e.g., quality of service, whose violation is tolerable
- F_i have analytical expressions only in very few cases; mostly, need to solve the problem approximately, e.g., sample average approximation (SAA) (Monte Carlo sampling)

SP example

Problem: protect a power plant from potential river flooding by building a dike with a height x meter

- Linear investment cost $c \cdot x$ (annualized)
- Uncertain annual maximum flooding water height H
- Annual loss is linear related to the height of overflow $H - x$ when it's positive with uncertain coefficient b



- **Objective**: minimize investment cost + expected economic loss

$$\min_{x} \quad cx + \mathbb{E}[b(H - x)_+]$$
$$\text{s.t.} \quad x \geq 0$$

Outline

CentraleSupélec

- Risk measures
- Stochastic decision-making and optimization
  - → Optimization model introduction
  - → Stochastic optimization formalization
- Sampled average approximation
- Risk-averse stochastic optimization

Sampled approximation

$$\min_x \quad L(x) = \mathbb{E}[l(x, \xi)]$$
$$\text{s.t.} \quad F_i(x) = \mathbb{E}[f_i(x, \xi)] \leq 0, \; i = 1, \ldots, m$$

- **Idea:** replace the expectation with a sampled model when the support Ξ is continuous
- Generate N i.i.d samples $\xi_1, \xi_2, \ldots, \xi_N$, each with $p_j = 1/N$
- Now solve the finite event deterministic SAA problem

$$\min_x \quad L_N(x) = \sum_{j=1}^{N} p_j l(x, \xi_j)$$
$$\text{s.t.} \quad F_{Ni}(x) = \sum_{j=1}^{N} p_j f_i(x, \xi_j) \leq 0, \; i = 1, \ldots, m$$

- Now, a deterministic optimization model

Sampled approximation

- Denote by $L_N^\star$ the optimal value of the SAA problem and $x_N^\star \in O_N^\star$ an optimal solution (set)
- How good is the quality of the sampled SAA solution
- Does it approach the true optimal value $L^\star$ and the true set of optimal solution $O^\star$ (as N increases)?
- How fast?
- How to validate the approximation in practice?

SAA and out-of-sample example

- Annual cost function of the dike design problem

$$\begin{aligned} \min_x \quad & 0.9x + \mathbb{E}[b(H-x)_+] \\ \text{s.t.} \quad & 0 \le x \le 3 \end{aligned}$$

$$H \sim \text{Gamma}(8, 4.5) \;\; b \sim \text{Lognorm}(0.7, 0.16^2)$$

$$\begin{aligned} \min_x \quad & 0.9x + \frac{1}{N}\sum_{i=1}^{N}[b_i(H_i - x)_+] \\ \text{s.t.} \quad & 0 \le x \le 3 \end{aligned}$$

- We solve SAA with $N=$ 10, 100, 1000, validation set uses M=100k
- Solved by **LP** (piecewise min can be transformed to LP)

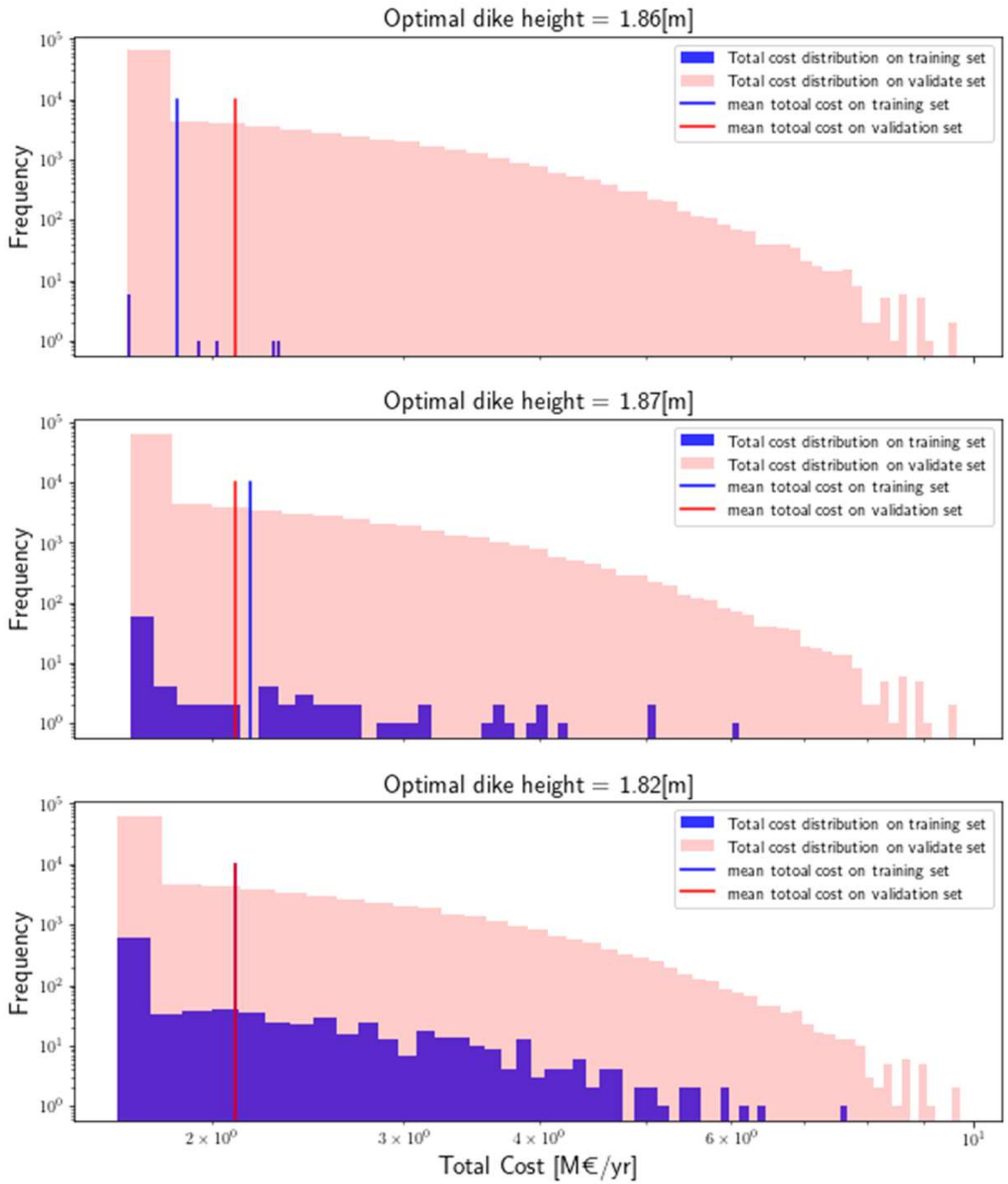
SAA and out-of-sample example

N	10	100	1000
$x_N^\star$	**1.86**	**1.87**	**1.82**
C_N	1.859	2.172	2.101
C^{val}	2.104	2.106	2.102

- $N \geq 1000$ is probably fine
- Computational cost increases very quickly along with N

  $\rightarrow$ Better sampling

  $\rightarrow$ Decomposition



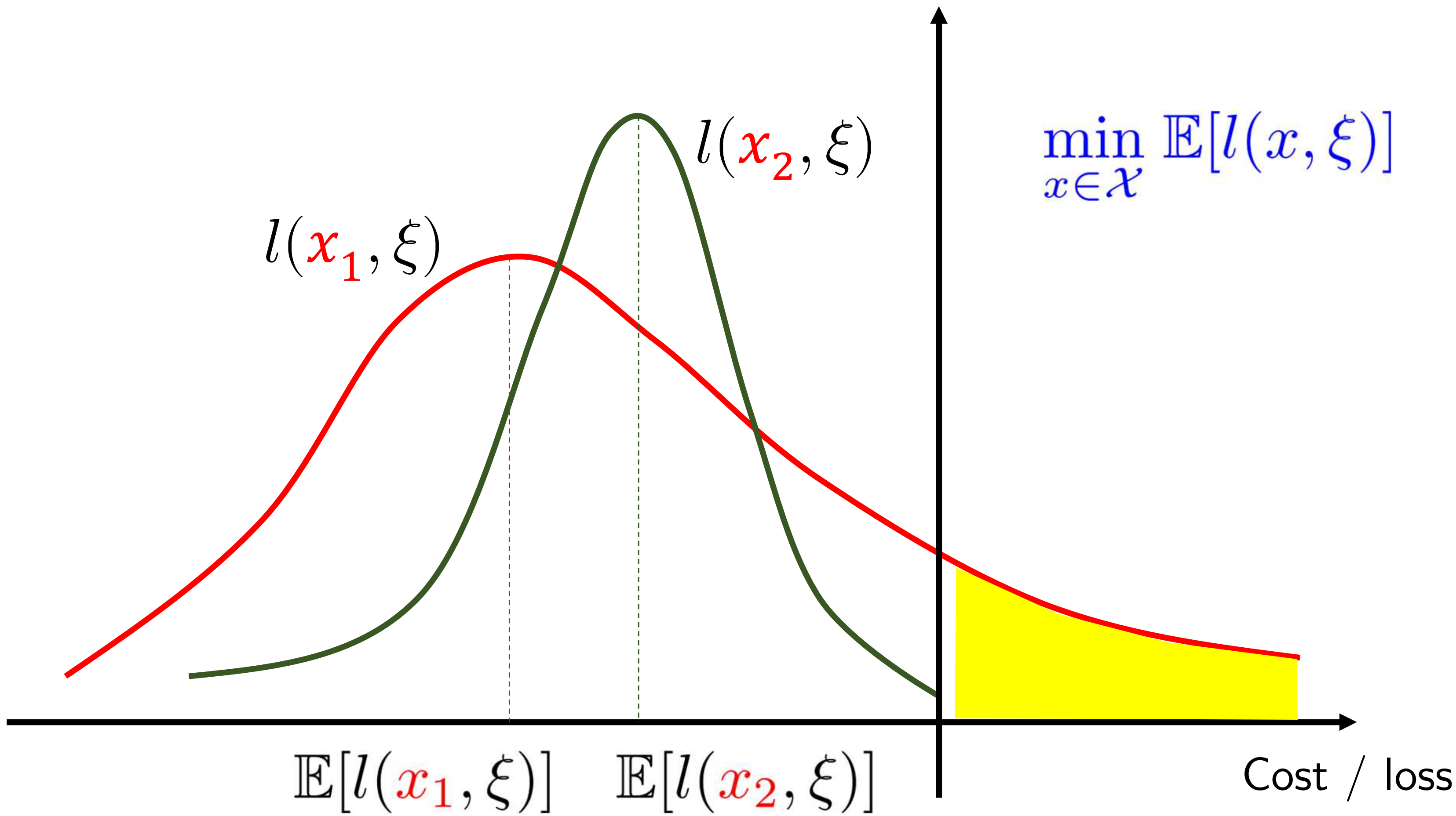
Outline

CentraleSupélec

- Risk measures
- Stochastic decision-making and optimization
  - → Optimization model introduction
  - → Stochastic optimization formalization
- Sampled average approximation
- Risk-averse stochastic optimization

Risk averse stochastic optimization

- To control of the **tail risk** in risk averse situations



VaR and chance constraints

- Let $l(x, \xi)$ represent a random loss function by definition:

$$\mathbf{VaR}_\alpha[l(x, \xi)] = \inf \{\gamma | \mathbb{P}[l(x, \xi) \leq \gamma] \geq \alpha\}$$

- **Chance constraint:**

$$\mathbf{VaR}_\alpha[l(x, \xi)] \leq b \Leftrightarrow \mathbb{P}[l(x, \xi) \leq b] \geq \alpha$$

- VaR in the **objective function**

$$\min_x \mathbf{VaR}_\alpha[l(x, \xi)] \quad \Leftrightarrow \quad \min_{x,\gamma} \gamma \quad \text{s.t. } \mathbb{P}[l(x, \xi) \leq \gamma] \geq \alpha$$

- **In general**, chance constraints are nonconvex and non-smooth, very difficult to optimize

CVaR optimization

- For $\alpha \in [0, 1]$, the CVaR on α is defined by

$$\mathbf{CVaR}_\alpha(X) = \mathbb{E}[X | X \geq \mathbf{VaR}_\alpha(X)] = \frac{1}{1-\alpha} \int_{x \geq \mathbf{VaR}_\alpha} x \, dF(x)$$

- Computing CVaR generally requires the computation of VaR
- CVaR-minimization might be even harder than VaR-minimization problem?

Under quite reasonable modeling assumptions, the opposite is true because that CVaR is coherent!

VaR vs CVaR in risk optimization

CentraleSupélec

CVaR is superior to VaR in risk optimization

- VaR is difficult to optimize numerically when losses are not normally distributed
- CVaR_α preserves convexity and can be expressed as a mimimization formulation
- $\mathrm{CVaR}_\alpha \geq \mathrm{VaR}_\alpha$ always holds, CVaR is a conservative approximation of VaR
- Able to control the expected tail risk

CVaR optimization

- Recall $l(x, \xi)$ represents a random loss function
- Consider the auxiliary function

$$F_\alpha(x, \gamma) = \gamma + \frac{1}{1-\alpha}\mathbb{E}\left\{[l(x,\xi) - \gamma]_+\right\}$$
$$= \gamma + \frac{1}{1-\alpha}\int_{\xi \in \Xi}[l(x,\xi) - \gamma]_+ \, p(\xi) d\xi$$

Theorem

1. For any fixed x, the function $\gamma \mapsto F_\alpha(x, \gamma)$ is convex.
2. $\mathbf{VaR}_\alpha(x)$ is a minimizer of the problem $\min_\gamma F_\alpha(x, \gamma)$.
3. $F_\alpha(x, \mathbf{VaR}_\alpha(x)) = \mathbf{CVaR}_\alpha(x)$.

CVaR optimization

CentraleSupélec

Proof: i) Since $[l(x,\xi)-\gamma]_+$ is a convex function in γ, it is true that this convexity is preserved by integrals and linear combination (Page 79, Boyd, S. and Vandenberghe, L., 2004. Convex optimization. Cambridge university press)

ii) Since $\min_\gamma F_\alpha(x,\gamma)$ is convex based on theorem 1(i), the KKT conditions are sufficient for optimality, i.e., we only need to check that $F_\alpha(x,\gamma)$ is stationary at $\gamma = \mathrm{VaR}_\alpha(x)$

$$\frac{\partial F_\alpha(x,\gamma)}{\partial \gamma}\Big|_{\gamma=\mathrm{VaR}_{\alpha(x)}} = 0$$

CVaR optimization

ii) Based on the following Proposition (Shapiro & Wardi 1994):

$E'[f(x)] = E[f'(x)]$ when the function $f(x)$ is convex w.p.1

we have

$$\frac{\partial F_\alpha(x,\gamma)}{\partial \gamma} = 1 + \frac{1}{1-\alpha}\int_{\xi\in\Xi}[-\mathbb{I}_{l(x,\xi)\geq\gamma}]p(\xi)d\xi$$

$$= 1 - \frac{P[l(x,\xi)\geq\gamma]}{1-\alpha}$$

$$P[l(x,\xi)\geq \mathrm{VaR}_\alpha(x)] = 1-\alpha$$

$$\frac{\partial F_\alpha(x,\gamma)}{\partial \gamma}|_{\gamma=\mathrm{VaR}_{\alpha(x)}} = 0$$

Shapiro, A., and Y. Wardi. "Nondifferentiability of the steady-state function in discrete event dynamic systems." *IEEE transactions on Automatic Control* 39, no. 8 (1994): 1707-1711.

CVaR optimization

iii) We have

$$F_\alpha(x, \mathrm{VaR}_\alpha(x)) = \mathrm{VaR}_\alpha(x) + \frac{1}{1-\alpha}\int_{\xi\in\Xi}[l(x,\xi) - \mathrm{VaR}_\alpha(x)]_+ p(\xi)d\xi$$

$$= \mathrm{VaR}_\alpha(x) + \boxed{\frac{1}{1-\alpha}\int_{\{\xi\in\Xi: l(x,\xi)\geq \mathrm{VaR}_\alpha(x)\}} l(x,\xi)p(\xi)d\xi}$$

$$-\frac{1}{1-\alpha}\int_{\{\xi\in\Xi: l(x,\xi)\geq \mathrm{VaR}_\alpha(x)\}} \mathrm{VaR}_\alpha(x)p(\xi)d\xi$$

$$= \mathrm{VaR}_\alpha(x) + \mathrm{CVaR}_\alpha(x)$$

$$-\frac{\mathrm{VaR}_\alpha(x)}{1-\alpha}\boxed{\int_{\{\xi\in\Xi: l(x,\xi)\geq \mathrm{VaR}_\alpha(x)\}} p(\xi)d\xi}$$

$$= \mathrm{VaR}_\alpha(x) + \mathrm{CVaR}_\alpha(x) - \mathrm{VaR}_\alpha(x)$$

CVaR optimization

The theorems imply

$$\min_{x\in\mathcal{X}} \mathbf{CVaR}_\alpha[l(x,\xi)] \quad \Leftrightarrow \quad \min_{x,\gamma\in\mathcal{X}\times\mathbb{R}} F_\alpha[l(x,\xi),\gamma]$$

$$\Leftrightarrow \quad \min_{x,\gamma\in\mathcal{X}\times\mathbb{R}} \gamma + \frac{1}{1-\alpha}\mathbb{E}\left\{[l(x,\xi)-\gamma]_+\right\}$$

- If the loss function $l(x,\xi)$ is convex in x, then $F_\alpha(x,\gamma)$ is convex in x and γ, the problem is convex optimization that can be well solved
- If $(x^\star,\gamma^\star)$ minimizes $F_\alpha(x,\gamma)$ over $\mathcal{X}\times\mathbb{R}$, then $F_\alpha(x^\star,\gamma^\star)$ is the optimal value of the CVaR minimization problem.

CVaR optimization

CentraleSupélec

- If CVaR is in the constraint

$$\mathbf{CVaR}_\alpha[l(x,\xi)] \leq L \;\Leftrightarrow\; \exists\gamma \in \mathbb{R}, F_\alpha[l(x,\xi),\gamma] \leq L$$

- Thus,

$$\begin{cases} \min\limits_{x\in\mathcal{X}} & f_0(x) \\ \text{s.t.} & \mathbf{CVaR}_\alpha[l(x,\xi)] \leq L \end{cases} \Leftrightarrow \begin{cases} \min\limits_{x,\gamma\in\mathcal{X}\times\mathbb{R}} & f_0(x) \\ \text{s.t.} & F_\alpha[l(x,\xi),\gamma] \leq L \end{cases}$$

- Convex optimization when $l(x,\xi)$ is convex in x

CVaR optimization

CentraleSupélec

- In the discrete case (e.g., sampling)

$$F_\alpha(x,\gamma) = \gamma + \frac{1}{1-\alpha}\sum_{k=1}^{N} p_k \left[l(x,\xi^k) - \gamma\right]_+$$

- Replace $[l(x,\xi^k) - \gamma]_+$ by additional variables z^k,

$$\min_{x\in\mathcal{X}} \mathbf{CVaR}_\alpha[l(x,\xi)] \;\Leftrightarrow\; \min_{x,\gamma\in\mathcal{X}\times\mathbb{R}} \gamma + \frac{1}{1-\alpha}\sum_{k=1}^{N} p_k z^k$$

$$\text{s.t. } z^k \geq 0,\ \forall k$$

$$z^k \geq l(x,\xi^k) - \gamma,\ \forall k$$

- When $l(x,\xi)$ affine in x, the CVaR minimization is an LP
- When $l(x,\xi)$ convex in x, convex optimization

CVaR optimization example

- Optimal dike construction

$$\min_{x} \quad 0.9x + \frac{1}{N}\sum_{k=1}^{N}[b_k(H_k - x)_+]$$

$$\text{s.t.} \quad 0 \le x \le 3$$

$c = 0.9,\ H \sim$ Gamma$(8, 4.5)$ and $b \sim$ Lognorm$(0.7, 0.16^2)$

$$\min_{x,\gamma \in [0,3]\times\mathbb{R}} \gamma + \frac{1}{(1-\alpha)N}\sum_{k=1}^{N} z^k$$

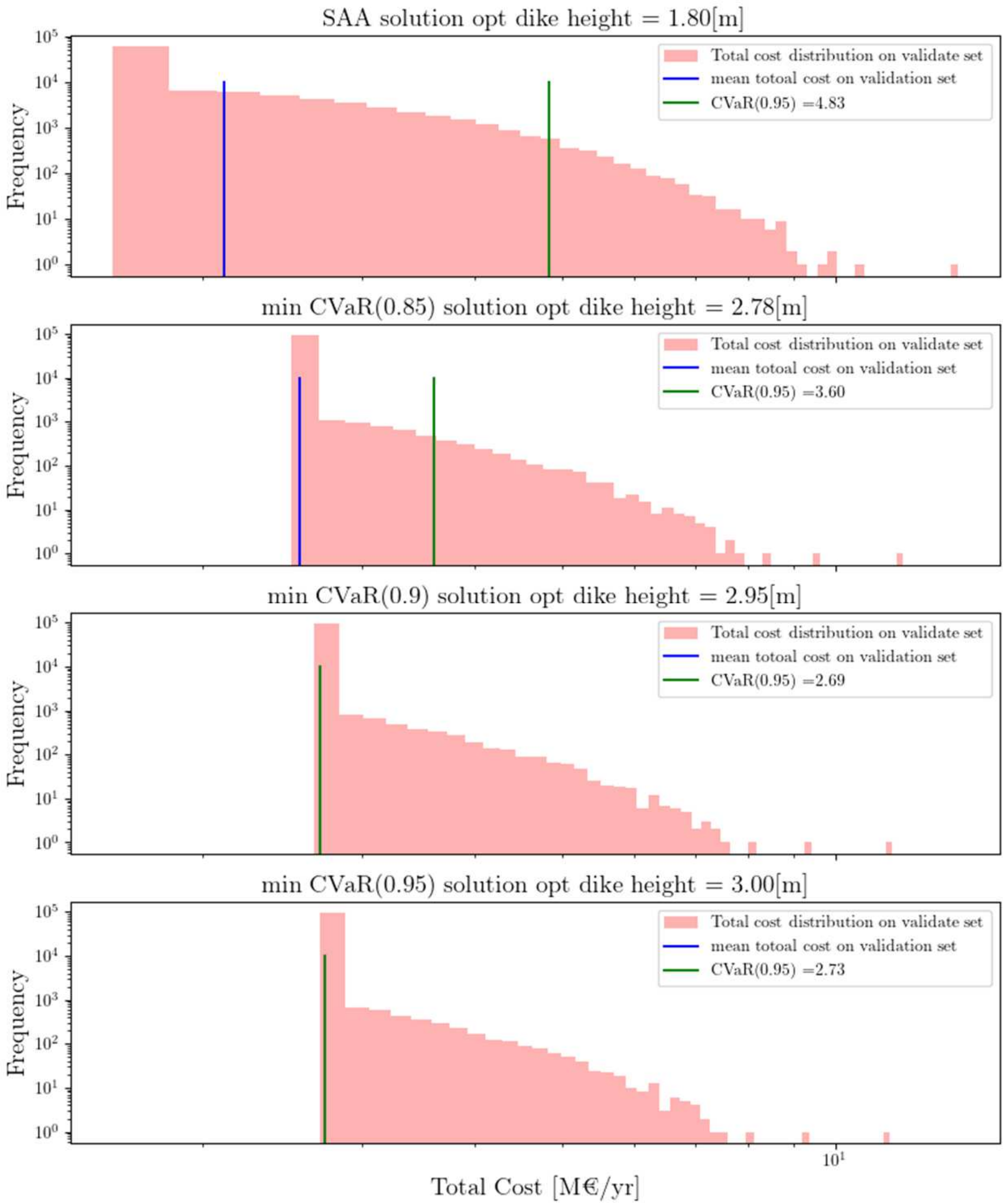
$$\text{s.t. } z^k \ge 0,\ \forall k$$

$$h^k \ge H_k - x,\ h_k \ge 0,\ \forall k$$

$$z^k \ge 0.9x + b_k h_k - \gamma,\ \forall k$$

CVaR optimization example

$(N = 2000)$	x [m]
SAA	1.80
$\text{CVaR}_{0.85}$	2.78
$\text{CVaR}_{0.9}$	2.95
$\text{CVaR}_{0.95}$	3.00



Exercises

- Decision analysis and risk-informed optimization for resilience

@yipingFANG

yiping.fang@centralesupelec.fr