Introduction to One and Two-Stage Stochastic and Robust Optimization

Michel De Lara
Cermics, École des Ponts ParisTech
France

École des Ponts ParisTech

December 3, 2018

Outline of the presentation

In decision-making, risk and time are bedfellows,
but for the fact that an uncertain outcome is revealed after the decision.
The talk moves along the number of decision stages: 1,2, more

Working out static examples

Two-stage stochastic programming problems

Outline of the presentation

Working out static examples

Two-stage stochastic programming problems

Working out classical examples

We will work out classical examples in Stochastic Optimization

- the blood-testing problem

  static, only risk

- the newsvendor problem

  static, only risk

Outline of the presentation

Working out static examples

The blood-testing problem

The newsvendor problem

Two-stage stochastic programming problems

Moving from deterministic convex piecewise linear programs

Moving from linear programs with constraints

Scenario decomposition resolution methods

Recalls on duality and Lagrangian decomposition

Progressive Hedging

The blood-testing problem (R. Dorfman) is a static stochastic optimization problem

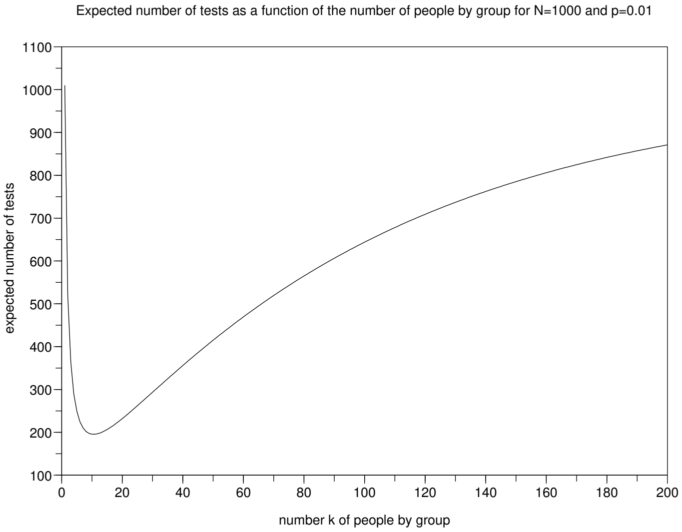
- Data:
  - A large number N of individuals are subjected to a blood test
  - The probability that the test is positive is p, the same for all individuals (a positive test means that the target individual has a specific disease; the prevalence of the disease in the population is p)
  - Individuals are stochastically independent
- Blood-testing method: the blood samples of k individuals are pooled and analyzed together
  - If the test is negative, this one test suffices for the k individuals
  - If the test is positive, each of the $k > 1$ individuals must be tested separately, and $k + 1$ tests are required, in all
- Optimization problem:
  - Find the value of k which minimizes the expected number of tests
  - Find the minimal expected number of tests

What is the optimal number of individuals in a group that minimizes the expected number of tests?

- For the first pool $\{1, \ldots, k\}$, the test is
  - negative with probability $(1-p)^k$ (by independence) $\rightarrow$ 1 test
  - positive with probability $1-(1-p)^k \rightarrow k+1$ tests
- When the pool size k is small,
  compared to the number N of individuals,
  the blood samples $\{1, \ldots, N\}$ are split in approximately N/k groups,
  so that the expected number of tests is

$$J(k) \approx \frac{N}{k}[1 \times (1-p)^k + (k+1) \times (1-(1-p)^k)]$$

The expected number of tests displays a marked hollow



In army practice, R. Dorfman achieved savings up to 80%

- The expected number of tests is

$$J(k) \approx \frac{N}{k}[1 \times (1-p)^k + (k+1) \times (1-(1-p)^k)]$$

- For small p,

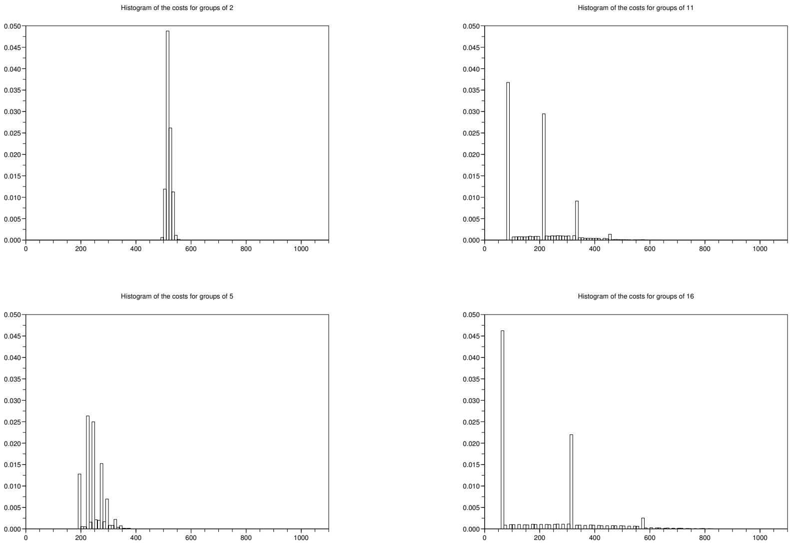
$$J(k)/N \approx 1/k + kp$$

  - so that the optimal number of individuals per group is $k^\star \approx 1/\sqrt{p}$
  - and the minimal expected number of tests is about

$$J^\star \approx J(k^\star) \approx 2N\sqrt{p} < N$$

- William Feller reports that, in army practice,
  R. Dorfman achieved savings up to 80%,
  compared to making N tests (the worst case solution)
  (take $p = 1/100$, giving $k^\star = 11 \approx 1/\sqrt{1/100} = 10$ and $J^\star \approx N/5$)

The optimal number of tests is a random variable



What about risk?

- The optimal number of individuals per group is 11
  if one minimizes the mathematical expectation $\mathbb{E}$
  of the number of tests
  (see also the top right histogram above)
- But if one minimizes the Tail Value at Risk at level $\lambda = 5\%$
  of the number of tests (more on $TVaR_\lambda$ later),
  numerical calculation show that, in the range from 2 to 33,
  the optimal number of individuals per group is 5
  (see also the bottom left histogram above)
- The bottom left histogram is more tight (less spread)
  than the top right histogram

Outline of the presentation

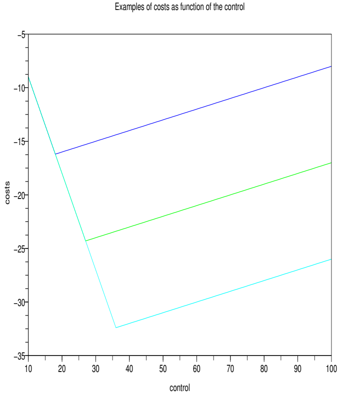
The "newsboy problem" is now coined the "newsvendor problem" ;-)

The (single-period) newsvendor problem stands as a classic in stochastic optimization

- Each morning, the newsvendor must decide how many copies $u \in \mathbb{U} = \{0, 1, 2 \ldots\}$ of the day's paper to order:
  u is the decision variable
- The newsvendor will meet a demand $w \in \mathbb{W} = \{0, 1, 2 \ldots\}$:
  the variable w is the uncertainty
- The newsvendor faces an economic tradeoff
  - she pays the unitary purchasing cost c per copy
  - she sells a copy at price p
  - if she remains with an unsold copy, it is worthless (perishable good)
- The newsvendor's costs $j(u, w)$ depend both
  on the decision u and on the uncertainty w:

$$j(u, w) = \underbrace{cu}_{\text{purchasing}} - \underbrace{p \min\{u, w\}}_{\text{selling}} = \max\{cu - pu, cu - pw\}$$

What is an "optimal" solution to the newsvendor problem?



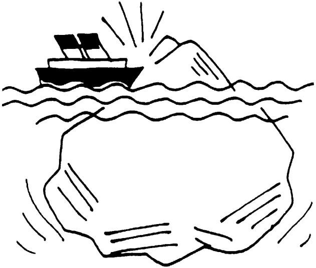
If you solve

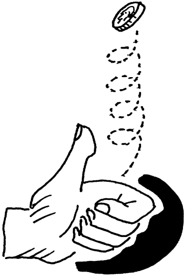
$$\min_{u \in \mathbb{U}} j(u, w)$$

the optimal solution is $u^\star = w$...
which depends on the unknown quantity w!

So, what do you suggest an "optimal" solution?

For you, Nature is rather random or hostile?





The newsvendor reveals her attitude towards risk in how she aggregates outcomes with respect to uncertainty

- In the robust or pessimistic approach,
  the (paranoid?) newsvendor minimizes the worst costs

$$\min_{u \in \mathbb{U}} \underbrace{\max_{w \in \overline{\mathbb{W}}} j(u, w)}_{\text{worst costs } J(u)}$$

  as if Nature were malevolent

- In the stochastic or expected approach, the newsvendor solves

$$\min_{u \in \mathbb{U}} \underbrace{\mathbb{E}_{\mathbf{W}}[j(u, \mathbf{W})]}_{\text{expected costs } J(u)}$$

  as if Nature played stochastically (casino)

If the newsvendor minimizes the worst costs

- We suppose that
  - the demand w belongs to a set $\overline{\mathbb{W}} = [\![w^\flat, w^\sharp]\!]$
  - the newsvendor knows the set $[\![w^\flat, w^\sharp]\!]$
- The worst costs are

$$J(u) = \max_{w \in \overline{\mathbb{W}}} j(u,w) = \max_{w \in [\![w^\flat, w^\sharp]\!]} [cu - p\min\{u,w\}] = cu - p\min\{u, w^\flat\}$$

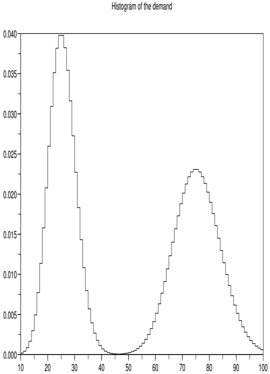
- Show that the order $u^\star = w^\flat$ minimizes the above expression $J(u)$
- Once the newsvendor makes the optimal order $u^\star = w^\flat$, the optimal costs are

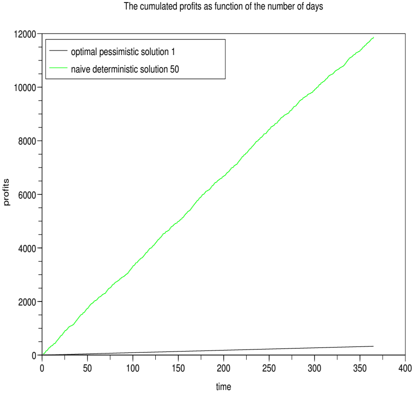
$$j(u^\star, \cdot) : w \in [\![w^\flat, w^\sharp]\!] \mapsto -(p-c)w^\flat$$

which, here, are no longer uncertain

Does it pay to be so pessimistic?

Not if demands are drawn independently from a probability distribution





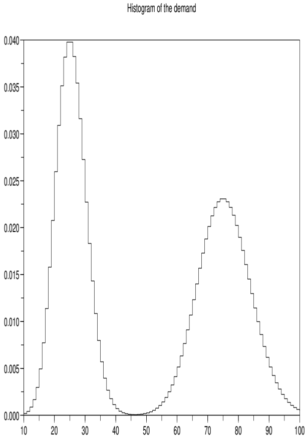
If the newsvendor minimizes the expected costs

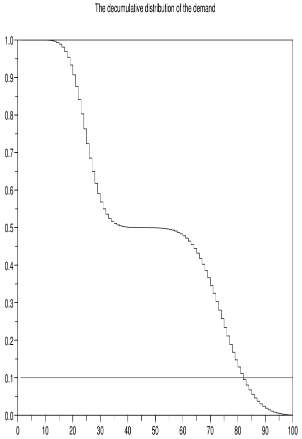
- We suppose that
  - the demand is a random variable, denoted $\mathbf{W}$
  - the newsvendor knows the probability distribution $\mathbb{P}_{\mathbf{W}}$ of the demand $\mathbf{W}$
- The expected costs are

$$J(u) = \mathbb{E}_{\mathbf{W}}[j(u, \mathbf{W})] = \mathbb{E}_{\mathbf{W}}[cu - p \min\{u, \mathbf{W}\}]$$

- Find an order $u^{\star}$ which minimizes the above expression $J(u)$
  - by calculating $J(u+1) - J(u)$
  - then using the *decumulative distribution function* $u \mapsto \mathbb{P}(\mathbf{W} > u)$

Here is an example of probability distribution and of decumulative distribution for the demand





Here stand some steps of the computation

$$\begin{aligned} J(u) &= cu - p\mathbb{E}[\min\{u, \mathbf{W}\}] \\ \min\{u, \mathbf{W}\} &= u\mathbf{1}_{\{u<\mathbf{W}\}} + \mathbf{W}\mathbf{1}_{\{u\geq\mathbf{W}\}} \\ \min\{u+1, \mathbf{W}\} &= (u+1)\mathbf{1}_{\{u+1\leq\mathbf{W}\}} + \mathbf{W}\mathbf{1}_{\{u+1>\mathbf{W}\}} \\ &= (u+1)\mathbf{1}_{\{u<\mathbf{W}\}} + \mathbf{W}\mathbf{1}_{\{u\geq\mathbf{W}\}} \\ \min\{u+1, \mathbf{W}\} - \min\{u, \mathbf{W}\} &= \mathbf{1}_{\{u<\mathbf{W}\}} \end{aligned}$$

$$J(u+1) - J(u) = c - p\mathbb{E}[\mathbf{1}_{\{u<\mathbf{W}\}}] = c - p\mathbb{P}(\mathbf{W} > u) \uparrow \text{ with } u$$

- If $\mathbb{P}(\mathbf{W} > 0) = 1$, then $J(1) - J(0) = c - p < 0$
- $J(u+1) - J(u) \rightarrow_{u\rightarrow+\infty} c > 0$

Characterization of the optimal decision $u^\star$

- Define the cut-off decisions $u^{\star\flat}$ and $u^{\star\sharp}$ by

$$u^{\star\flat} = \max\{u \ , \ \mathbb{P}(\mathbf{W} > u) > \frac{c}{p}\} \quad \Big(u \leq u^{\star\flat} \iff J(u+1) < J(u)\Big)$$

$$u^{\star\sharp} = \min\{u \ , \ \mathbb{P}(\mathbf{W} > u) < \frac{c}{p}\} \quad \Big(u \geq u^{\star\sharp} \iff J(u+1) > J(u)\Big)$$

- An optimal decision $u^\star$ satisfies

$$u^\star \in \{u^{\star\flat}+1, \ldots, u^{\star\sharp}\} \text{ and } J(u^\star) = \min\{J(u^{\star\flat}+1), J(u^{\star\sharp})\}$$

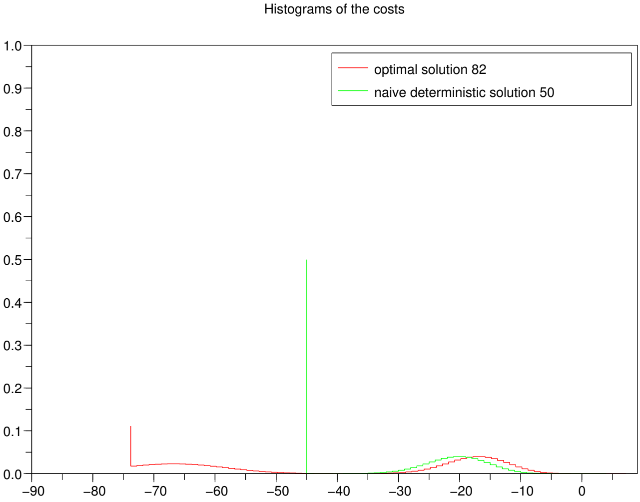
- The optimal decision $u^\star$ is unique if and only if $u^{\star\flat}+1 = u^{\star\sharp}$, that is, if and only if

$$\mathbb{P}(\mathbf{W} > u^\star - 1) > \frac{c}{p} > \mathbb{P}(\mathbf{W} > u^\star)$$

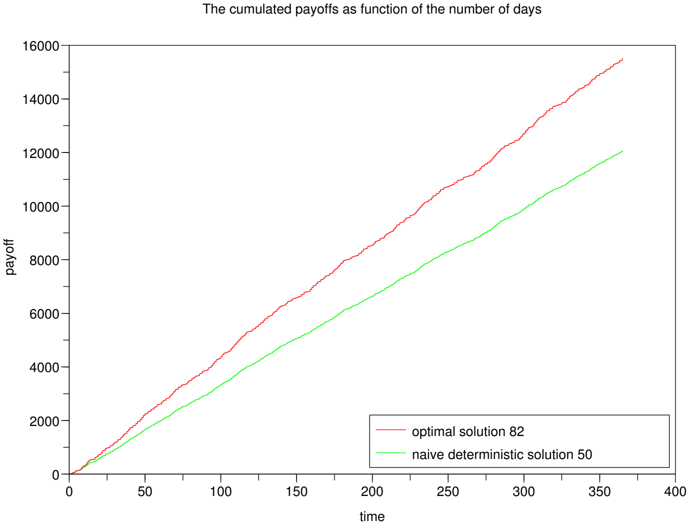
- Once the newsvendor makes the optimal order $u^\star$, the optimal costs are the random variable

$$j(u^\star, \mathbf{W}) = cu^\star - p\min\{u^\star, \mathbf{W}\}$$

The distribution of the optimal costs displays lower costs than with the naive deterministic solution $u = \mathbb{E}[\mathbf{W}]$



The cumulated profits over 365 days reveal that it pays to do stochastic optimization



When deterministic optimization is (wrongly) optimistic

- If you plug the mean value $\overline{\mathbf{W}} = \mathbb{E}_{\mathbf{W}}[\mathbf{W}]$ in the criterion

$$j(u, w) = cu - p \min\{u, w\}$$

instead of the random variable $\mathbf{W}$

- you obtain a deterministic optimization problem

$$\min_{u \in \mathbb{U}} j(u, \overline{\mathbf{W}}) = j(\overline{\mathbf{W}}, \overline{\mathbf{W}}) = (c - p)\overline{\mathbf{W}}$$

whose minimal value $\min_{u \in \mathbb{U}} j(u, \overline{\mathbf{W}})$
is overly and wrongly optimistic

- because, *on the proper stochastic benchmark* $\mathbb{E}_{\mathbf{W}}[j(\cdot, \mathbf{W})]$, the deterministic optimal solution $\overline{u} = \overline{\mathbf{W}} = 50$ produces poorer results than the stochastic optimal solution $u^{\star} = 82$

$$\underbrace{j(\overline{u}, \overline{\mathbf{W}}) = \min_{u \in \mathbb{U}} j(u, \overline{\mathbf{W}})}_{-44.968856} \leq \underbrace{\mathbb{E}_{\mathbf{W}}[j(u^{\star}, \mathbf{W})] = \min_{u \in \mathbb{U}} \mathbb{E}_{\mathbf{W}}[j(u, \mathbf{W})]}_{-41.259519} \leq \underbrace{\mathbb{E}_{\mathbf{W}}[j(\overline{u}, \mathbf{W})]}_{-32.498824}$$

When deterministic optimization is (wrongly) optimistic

Proposition

Let $\mathbf{W}$ be a random variable with mean $\overline{\mathbf{W}} = \mathbb{E}_{\mathbf{W}}[\mathbf{W}]$.
Suppose that $w \mapsto j(u, w)$ is convex, for all decision u. Then,

$$\inf_{u \in \mathbb{U}} j(u, \mathbb{E}_{\mathbf{W}}[\mathbf{W}]) \leq \inf_{u \in \mathbb{U}} \mathbb{E}_{\mathbf{W}}[j(u, \mathbf{W})]$$

Supposing that the infima are minima

$$j(\overline{u}, \overline{\mathbf{W}}) = \min_{u \in \mathbb{U}} j(u, \overline{\mathbf{W}}) \leq \mathbb{E}_{\mathbf{W}}[j(u^\star, \mathbf{W})] = \min_{u \in \mathbb{U}} \mathbb{E}_{\mathbf{W}}[j(u, \mathbf{W})]$$

we have

$$\overbrace{j(\overline{u}, \overline{\mathbf{W}}) \leq}^{\text{overly optimistic}} \quad \mathbb{E}_{\mathbf{W}}[j(u^\star, \mathbf{W})] \underbrace{\leq \mathbb{E}_{\mathbf{W}}[j(\overline{u}, \mathbf{W})]}_{\text{wrongly optimistic}}$$

Where do we stand after having worked out two examples?

- When you move from deterministic optimization to optimization under uncertainty, you come accross the issue of risk attitudes
- Risk is in the eyes of the beholder ;-)
  and materializes in the a priori knowledge on the uncertainties
  - either probabilistic/stochastic
    - independence and Bernoulli distributions in the blood test example
    - uncertain demand faced by the newsvendor
      modeled as a random variable
  - or set-membership
    - uncertain demand faced by the newsvendor modeled by a set

Where have we gone till now? And what comes next

- We have seen two examples of optimization problems
  with a single deterministic decision variable,
  and with a criterion including a random variable
- Now, we will turn to optimization problems
  with two decision variables,
  the first one deterministic and the second one random

Outline of the presentation

Working out static examples

Two-stage stochastic programming problems

What awaits us

- We lay out two ways to move
  from one-stage deterministic optimization problems
  to two-stage stochastic linear programs
  - in one, we start from a deterministic convex piecewise linear program (without constraints)
  - in the other, we start from a deterministic linear program with constraints
- We show how we can also obtain two-stage risk-averse programs, when we handle risk by means of the Tail Value at Risk
- We show a scenario decomposition resolution method
  adapted to two-stage stochastic programs that are strongly convex
- We outline the Progressive Hedging resolution method,
  adapted to two-stage stochastic linear programs

Outline of the presentation

We revisit the newsvendor problem

Writing the newsvendor problem as a linear program, in three steps

We consider the stochastic optimization problem

$$\min_{u \in \mathbb{U}} J(u) = \mathbb{E}_{\mathbb{P}}[j(u, \mathbf{W})]$$

where

$$j(u, w) = cu - p \min\{u, w\}$$

and we show in three steps how to rewrite this problem as a linear program

Step 1: exploiting convex piecewise linearity of the criterion

First, we write:

$$
\begin{aligned}
j(u,w) &= cu - p\min\{u,w\} \\
&= \max\{cu - pu, cu - pw\} \\
&= \min_{v\in\mathbb{R}}\{v \mid v \geq cu - pu \;,\;\; v \geq cu - pw\}
\end{aligned}
$$

Step 2: exploiting convexity of the mathematical expectation

- We suppose that the demand $\mathbf{W}$ can take a finite number S of possible values w_s, $s \in \mathbb{S}$
- where s denotes a *scenario* in the finite set $\mathbb{S}$ (S=card($\mathbb{S}$))
- and we denote π_s the probability of scenario s, with

$$\sum_{s \in \mathbb{S}} \pi_s = 1 \text{ and } \pi_s \geq 0 \ , \quad \forall s \in \mathbb{S}$$

Step 2: exploiting convexity of the mathematical expectation

Second, we deduce

$$
\begin{aligned}
J(u) &= \mathbb{E}_{\mathbb{P}}[j(u, \mathbf{W})] \\
&= \sum_{s \in \mathbb{S}} \pi_s j(u, w_s) \\
&= \sum_{s \in \mathbb{S}} \pi_s \min_{v_s \in \mathbb{R}} \{ v_s \mid v_s \geq cu - pu \ , \quad v_s \geq cu - pw_s \} \\
&= \min_{(v_s)_{s \in \mathbb{S}} \in \mathbb{R}^S} \sum_{s \in \mathbb{S}} \pi_s v_s \\
&\quad \text{under the constraints} \\
&\quad v_s \geq cu - pu \ , \quad v_s \geq cu - pw_s \ , \quad \forall s \in \mathbb{S}
\end{aligned}
$$

Step 3: exploiting $\min \min = \min$

Third, we minimize with respect to the original decision $u \in \mathbb{U}$

$$\min_{u \in \mathbb{U}} J(u) = \min_{u \in \mathbb{U}, (v_s)_{s \in \mathbb{S}} \in \mathbb{R}^S} \sum_{s \in \mathbb{S}} \pi_s v_s$$

$$v_s \geq cu - pu \; , \quad \forall s \in \mathbb{S}$$

$$v_s \geq cu - pw_s \; , \quad \forall s \in \mathbb{S}$$

This is a linear program,
especially when we consider that the decision can take continuous values:
$u \in \mathbb{U} = [0, u^\sharp]$

The newsvendor problem example
is a special case of a general mechanism

From convex piecewise linear to linear programming

- The convex piecewise linear program

$$\min_{x \in \mathbb{R}^n} \max_{i=1,\ldots,m} \left\langle c^i \, , x \right\rangle + b^i$$

- can be written as the linear program

$$\min_{x \in \mathbb{R}^n} \min_{v \in \mathbb{R}} v$$

$$v \geq \left\langle c^i \, , x \right\rangle + b^i \, , \qquad i = 1, \ldots, m$$

From stochastic convex piecewise linear programming to stochastic linear programming

- The stochastic convex piecewise linear program

$$\min_{x \in \mathbb{R}^n} \sum_{s \in \mathbb{S}} \pi_s \max_{i=1,\ldots,m} \left\langle c_s^i , x \right\rangle + b_s^i$$

- can be written as the stochastic linear program

$$\min_{x \in \mathbb{R}^n} \min_{(v_s)_{s \in \mathbb{S}} \in \mathbb{R}^S} \sum_{s \in \mathbb{S}} \pi_s v_s$$

$$v_s \geq \left\langle c_s^i , x \right\rangle + b_s^i \; , \quad i = 1, \ldots, m \; , \;\; s \in \mathbb{S}$$

What happens if we want to minimize risk, not mathematical expectation?

- Instead of minimizing the mathematical expectation

$$\mathbb{E}[\mathbf{C}] \quad (= \sum_{s \in \mathbb{S}} \pi_s \mathbf{C}_s)$$

- we want to minimize the Tail Value at Risk (at level $\lambda \in [0, 1[$), given by the Rockafellar-Uryasev formula

$$TVaR_\lambda[\mathbf{C}] = \inf_{r \in \mathbb{R}} \left\{ \frac{\mathbb{E}[(\mathbf{C} - r)_+]}{1 - \lambda} + r \right\}$$

- whose limit cases are mean and worst case

$$TVaR_0[\mathbf{C}] = \mathbb{E}[\mathbf{C}]$$
$$TVaR_1[\mathbf{C}] = \lim_{\lambda \to 1} TVaR_\lambda[\mathbf{C}] = \sup_{\omega \in \Omega} \mathbf{C}(\omega)$$

Minimizing the Tail Value at Risk of costs: convex piecewise linear programming formulation

- The risk-averse stochastic convex piecewise linear program

$$\min_{x \in \mathbb{R}^n} \min_{r \in \mathbb{R}} \left\{ r + \frac{1}{1-\lambda} \sum_{s \in \mathbb{S}} \pi_s \Big(\max_{i=1,\ldots,m} \left\langle c_s^i \,, x \right\rangle + b_s^i - r \Big)_+ \right\}$$

- can be written as the convex piecewise linear program

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \min_{r \in \mathbb{R}} \min_{(u_s)_{s \in \mathbb{S}} \in \mathbb{R}^S} \quad & r + \frac{1}{1-\lambda} \sum_{s \in \mathbb{S}} \pi_s (u_s - r)_+ \\
& u_s \geq \left\langle c_s^1 \,, x \right\rangle + b_s^1 \,, \quad \forall s \in \mathbb{S} \\
& \vdots \\
& u_s \geq \left\langle c_s^m \,, x \right\rangle + b_s^m \,, \quad \forall s \in \mathbb{S}
\end{aligned}$$

Minimizing the Tail Value at Risk of costs: linear programming formulation

- The risk-averse stochastic convex piecewise linear program

$$\min_{x \in \mathbb{R}^n} \min_{r \in \mathbb{R}} \left\{ r + \frac{1}{1-\lambda} \sum_{s \in \mathbb{S}} \pi_s \Big(\max_{i=1,\ldots,m} \left\langle c_s^i \,, x \right\rangle + b_s^i - r \Big)_+ \right\}$$

- can be written as the linear program

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \min_{r \in \mathbb{R}} \min_{(v_s)_{s \in \mathbb{S}} \in \mathbb{R}^S} \quad & r + \frac{1}{1-\lambda} \sum_{s \in \mathbb{S}} \pi_s v_s \\
& v_s \geq \left\langle c_s^1 \,, x \right\rangle + b_s^1 - r \,, \quad \forall s \in \mathbb{S} \\
& \vdots \\
& v_s \geq \left\langle c_s^m \,, x \right\rangle + b_s^m - r \,, \quad \forall s \in \mathbb{S} \\
& v_s \geq 0 \,, \quad \forall s \in \mathbb{S}
\end{aligned}$$

How to use risk-averse stochastic programming in practice?

- Denote by $x^\star_\lambda$ the (supposed unique) solution
- As $1-\lambda$ measures the upper probability of risky events, start with $\lambda = 0$ and display, to the decision-maker, the risk-neutral solution $x^\star_0$ and the probability distribution (histogram) of the random costs

$$s \mapsto \max_{i=1,\dots,m} \left\langle c^i_s \, , x^\star_0 \right\rangle + b^i_s$$

- Then move to the confidence level $\lambda = 0.99$ (only events with probability less than 1% are considered), and do the same
- For a range of possible values for λ, display, to the decision-maker, the solution $x^\star_\lambda$ and the histogram of the random costs

$$s \mapsto \max_{i=1,\dots,m} \left\langle c^i_s \, , x^\star_\lambda \right\rangle + b^i_s$$

- The decision-maker should choose his confidence level λ

We can also minimize the mean costs, while controlling for large costs

- Instead of only minimizing the mathematical expectation

$$\mathbb{E}[\mathbf{C}] \quad (= \sum_{s \in \mathbb{S}} \pi_s \mathbf{C}_s)$$

- we add the constraint that the Tail Value at Risk (at level $\lambda \in [0, 1[$) is not too large

$$TVaR_\lambda[\mathbf{C}] = \inf_{r \in \mathbb{R}} \left\{ \frac{\mathbb{E}[(\mathbf{C} - r)_+]}{1 - \lambda} + r \right\} \leq C^\sharp$$

- We can also choose to minimize a mixture

$$\theta\mathbb{E}[\mathbf{C}] + (1-\theta)\,TVaR_\lambda[\mathbf{C}] = \inf_{r \in \mathbb{R}} \left\{ \theta\mathbb{E}[\mathbf{C}] + (1-\theta)\frac{\mathbb{E}[(\mathbf{C} - r)_+]}{1 - \lambda} + (1-\theta)r \right\}$$

Minimizing a mixture: convex piecewise linear programming formulation

- The risk-averse stochastic convex piecewise linear program

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n}\min_{r\in\mathbb{R}} \Bigg\{ & \theta \sum_{s\in\mathbb{S}} \pi_s \max_{i=1,\dots,m} \left\langle c_s^i \,, x \right\rangle + b_s^i \\
& +(1-\theta)r + \frac{1-\theta}{1-\lambda} \sum_{s\in\mathbb{S}} \pi_s \left(\max_{i=1,\dots,m} \left\langle c_s^i \,, x \right\rangle + b_s^i - r \right)_+ \Bigg\}
\end{aligned}
$$

- can be written as the convex piecewise linear program

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n}\min_{r\in\mathbb{R}}\min_{(u_s)_{s\in\mathbb{S}}\in\mathbb{R}^S} \quad & \sum_{s\in\mathbb{S}} \pi_s \left\{ \theta u_s + (1-\theta) r + \frac{1-\theta}{1-\lambda}(u_s - r)_+ \right\} \\
& u_s \geq \left\langle c_s^1 \,, x \right\rangle + b_s^1 \,, \quad \forall s \in \mathbb{S} \\
& \vdots \\
& u_s \geq \left\langle c_s^m \,, x \right\rangle + b_s^m \,, \quad \forall s \in \mathbb{S}
\end{aligned}
$$

Minimizing a mixture: linear programming formulation

- The risk-averse stochastic convex piecewise linear program

$$\min_{x\in\mathbb{R}^n}\min_{r\in\mathbb{R}}\Big\{\theta\sum_{s\in\mathbb{S}}\pi_s\max_{i=1,\ldots,m}\left\langle c_s^i\,,x\right\rangle+b_s^i$$

$$+(1-\theta)r+\frac{1-\theta}{1-\lambda}\sum_{s\in\mathbb{S}}\pi_s\Big(\max_{i=1,\ldots,m}\left\langle c_s^i\,,x\right\rangle+b_s^i-r\Big)_+\Big\}$$

- can be written as the linear program

$$\min_{x\in\mathbb{R}^n}\min_{r\in\mathbb{R}}\min_{(u_s)_{s\in\mathbb{S}}\in\mathbb{R}^S}\min_{(v_s)_{s\in\mathbb{S}}\in\mathbb{R}^S}\sum_{s\in\mathbb{S}}\pi_s\left\{\theta u_s+(1-\theta)r+\frac{1-\theta}{1-\lambda}v_s\right\}$$

$$u_s\geq\left\langle c_s^1\,,x\right\rangle+b_s^1\;,\quad\forall s\in\mathbb{S}$$

$$\vdots$$

$$u_s\geq\left\langle c_s^m\,,x\right\rangle+b_s^m\;,\quad\forall s\in\mathbb{S}$$

$$v_s\geq u_s-r\;,\quad\forall s\in\mathbb{S}$$

$$v_s\geq 0\;,\quad\forall s\in\mathbb{S}$$

How to use risk-averse stochastic programming in practice?

- Denote by $x^{\star}_{\lambda,\theta}$ the (supposed unique) solution
- As $1-\lambda$ measures the upper probability of risky events, let the decision-maker choose a confidence level λ — $\lambda = 0.99$ (only events with probability less than 1% are considered), $\lambda = 0.95$, $\lambda = 0.90$, for instance
- Start with $\theta = 0$ and display, to the decision-maker, the risk-neutral solution $x^{\star}_{\lambda,0}$ (which does not depend on λ) and the probability distribution (histogram) of the random costs

$$s \mapsto \max_{i=1,\ldots,m} \left\langle c^i_s \, , x^{\star}_{\lambda,0} \right\rangle + b^i_s$$

- Increase θ from 0 to 1, and display, to the decision-maker, the solution $x^{\star}_{\lambda,\theta}$ and the histogram of the random costs

$$s \mapsto \max_{i=1,\ldots,m} \left\langle c^i_s \, , x^{\star}_{\lambda,\theta} \right\rangle + b^i_s$$

- The decision-maker reveals his confidence level λ and his mixture $(\theta, 1-\theta)$ as he selects his prefered histogram

Outline of the presentation

Suppose you had to manage a day-ahead energy market
You would have to fix reserves by night
and adjust in the morning with recourse energies

From linear to stochastic programming

- The linear program

$$\min_{x \in \mathbb{R}^n} \langle c \,, x \rangle$$
$$Ax + b \quad \geq 0 \quad (\in \mathbb{R}^m)$$

- becomes a stochastic program

$$\min_{x \in \mathbb{R}^n} \sum_{s \in \mathbb{S}} \pi_s \langle c_s \,, x \rangle$$
$$A_s x + b_s \quad \geq 0 \,, \quad \forall s \in \mathbb{S}$$

- We observe that there are as many (vector) inequalities as there are possible scenarios $s \in \mathbb{S}$

$$A_s x + b_s \geq 0 \,, \quad \forall s \in \mathbb{S}$$

and these inequality constraints can delineate an empty domain for optimization

Recourse variables need be introduced for feasability issues

- We denote by $s \in \mathbb{S}$ any possible value of the random variable ξ, with corresponding probability π_s
- and we introduce a recourse variable $y = (y_s)_{s \in \mathbb{S}}$ and the program

$$
\min_{x,(y_s)_{s\in\mathbb{S}}} \sum_{s\in\mathbb{S}} \pi_s \Big(\langle c_s \,, x \rangle + \langle p_s \,, y_s \rangle \Big)
$$

$$
\begin{aligned}
y_s &\geq 0 \,, \quad \forall s \in \mathbb{S} \\
A_s x + b_s + y_s &\geq 0 \,, \quad \forall s \in \mathbb{S}
\end{aligned}
$$

- so that the inequality $A_s x + b_s + y_s \geq 0$ is now possible, at (unitary recourse) price vector $p = (p_s, s \in \mathbb{S})$
- Observe that such stochastic programs are huge problems, with solution $\big(x, (y_s)_{s\in\mathbb{S}}\big)$, but remain linear

Minimizing the Tail Value at Risk of costs: linear programming formulation

- The risk-averse stochastic linear program with recourse

$$\min_{x,(y_s)_{s\in\mathbb{S}}} \min_{r\in\mathbb{R}} \left\{ r + \frac{1}{1-\lambda} \sum_{s\in\mathbb{S}} \pi_s \Big(\langle c_s \, , x \rangle + \langle p_s \, , y_s \rangle \Big)_+ \right\}$$

- can be written as the linear program

$$\begin{aligned} \min_{x,(y_s)_{s\in\mathbb{S}}} \min_{r} \min_{(v_s)_{s\in\mathbb{S}}} \quad & r + \frac{1}{1-\lambda} \sum_{s\in\mathbb{S}} \pi_s v_s \\ v_s - \langle c_s \, , x \rangle - \langle p_s \, , y_s \rangle & \geq 0 \, , \quad \forall s \in \mathbb{S} \\ v_s & \geq 0 \, , \quad \forall s \in \mathbb{S} \\ y_s & \geq 0 \, , \quad \forall s \in \mathbb{S} \\ A_s x + b_s + y_s & \geq 0 \, , \quad \forall s \in \mathbb{S} \end{aligned}$$

Minimizing a mixture: linear programming formulation

- The risk-averse stochastic linear program with recourse

$$\min_{x,(y_s)_{s\in\mathbb{S}}} \min_{r\in\mathbb{R}} \Big\{ \theta \sum_{s\in\mathbb{S}} \pi_s \Big(\langle c_s , x \rangle + \langle p_s , y_s \rangle \Big) + (1-\theta) r + \frac{1-\theta}{1-\lambda} \sum_{s\in\mathbb{S}} \pi_s \Big(\langle c_s , x \rangle + \langle p_s , y_s \rangle \Big)_+ \Big\}$$

- can be written as the linear program

$$\min_{x,(y_s)_{s\in\mathbb{S}}} \min_{r} \min_{(u_s,v_s)_{s\in\mathbb{S}}} \sum_{s\in\mathbb{S}} \pi_s \Big\{ \theta u_s + (1-\theta) r + \frac{1-\theta}{1-\lambda} v_s \Big\}$$

$$\begin{aligned}
u_s - \langle c_s , x \rangle - \langle p_s , y_s \rangle &\geq 0 , \quad \forall s \in \mathbb{S} \\
v_s - u_s + r &\geq 0 , \quad \forall s \in \mathbb{S} \\
v_s &\geq 0 , \quad \forall s \in \mathbb{S} \\
y_s &\geq 0 , \quad \forall s \in \mathbb{S} \\
A_s x + b_s + y_s &\geq 0 , \quad \forall s \in \mathbb{S}
\end{aligned}$$

Roger Wets example

http://cermics.enpc.fr/~delara/TEACHING/

CEA-EDF-INRIA_2012/Roger_Wets1.pdf

Two-step stochastic programs with recourse can become deterministic non-smooth convex problems

- Define

$$\underbrace{Q_s(x)}_{\text{value function}} = \min\{\langle p_s , y\rangle , A_s x + b_s + y \geq 0\}$$

which is a convex function of x, non-smooth

- so that the original two-step stochastic program with recourse

$$\min_{x,(y_s)_{s\in\mathbb{S}}} \sum_{s\in\mathbb{S}} \pi_s \big[\langle c_s , x\rangle + \langle p_s , y_s\rangle \big]$$

$$\begin{aligned} y_s &\geq 0 , \quad \forall s \in \mathbb{S} \\ A_s x + b_s + y_s &\geq 0 , \quad \forall s \in \mathbb{S} \end{aligned}$$

- now becomes the deterministic non-smooth convex problem

$$\min_x \sum_{s\in\mathbb{S}} \pi_s \big[\langle c_s , x\rangle + Q_s(x) \big]$$

- An optimal solution is now more likely to be an inner solution (more robust)

A quadratic toy problem

Let $c > 0\ ,\ \ d_1 \geq 0\ ,\ \ d_2 \geq 0$

- Show that the (worst case) optimization problem

$$\min_{x \in \mathbb{R}} \frac{1}{2} c x^2$$
$$x \geq d_1$$
$$x \geq d_2$$

  has (worst case) solution

$$\overline{x} = \max\{d_1, d_2\}$$

- What happens if we allow room for recourse?

A quadratic toy problem with recourse

Let $c > 0$, $d_1 \geq 0$, $d_2 \geq 0$, $p_1 > 0$, $p_2 > 0$

- Show that the (stochastic) optimization problem

$$\min_{(x,y_1,y_2)\in\mathbb{R}^3} \frac{1}{2}\Big(cx^2 + p_1 y_1^2 + p_2 y_2^2\Big)$$
$$x + y_1 = d_1$$
$$x + y_2 = d_2$$

  has a solution $x^\star$ given by

$$x^\star = \frac{p_1}{c + p_1 + p_2} d_1 + \frac{p_2}{c + p_1 + p_2} d_2 + \frac{c}{c + p_1 + p_2} 0$$

- Therefore, $x^\star$ belongs to the convex generated by $\{0, d_1, d_2\}$, that is,

$$x^\star \in [0, \max\{d_1, d_2\}]$$

- Compare with the (worst case) solution $\overline{x} = \max\{d_1, d_2\}$

Where have we gone till now? And what comes next

- We have arrived at optimization problems
  with two decision variables
  - a first one deterministic
  - a second one random (as it is indexed by the scenarios)
- We have handled the risk neutral case, but also the risk averse case,
  with risk measures displaying good mathematical properties
  (like the Tail Value at Risk)
- We will now present resolution methods that, somehow surprisingly,
  relax the assumption that the first decision variable is deterministic

Outline of the presentation

We start with a two-stage stochastic optimization problem formulated on a tree

$$\text{Criterion} \quad j : \underbrace{\mathbb{X}}_{\text{initial decision}} \times \overbrace{\mathbb{Y}}^{\text{recourse variable}} \times \underbrace{\mathbb{S}}_{\text{scenario}} \to \mathbb{R}$$

and set-valued mapping $\mathcal{Y} : \mathbb{X} \times \mathbb{S} \to 2^{\mathbb{Y}}$

- Stochastic optimization problem

$$\min_{x,(y_s)_{s\in\mathbb{S}}} \sum_{s\in\mathbb{S}} \pi_s j_s(x, y_s)$$
$$x \in \mathbb{X}$$
$$y_s \in \mathcal{Y}_s(x) \;, \quad \forall s \in \mathbb{S}$$

- Solutions $(x, (y_s)_{s\in\mathbb{S}})$ are naturally indexed by a tree
  - with one root
  - and $S = \text{card}\,\mathbb{S}$ leaves

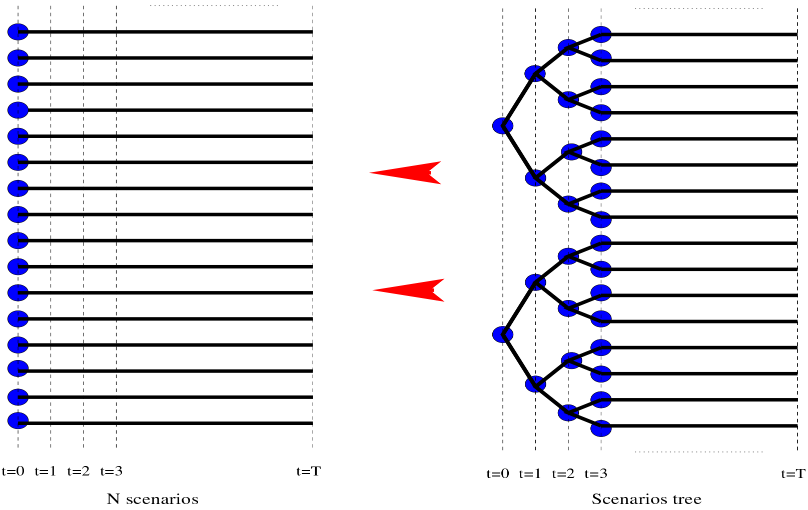
We transform the two-stage stochastic optimization problem by extending the solution space

- We consider initial decisions $(x_s)_{s \in \mathbb{S}}$ and the problem

$$
\begin{aligned}
&\min_{x,(x_s)_{s\in\mathbb{S}},(y_s)_{s\in\mathbb{S}}} \sum_{s\in\mathbb{S}} \pi_s j_s\big(x_s, y_s\big) \\
&x_s \in \mathbb{X} \ , \quad \forall s \in \mathbb{S} \\
&y_s \in \mathcal{Y}_s(x_s) \ , \quad \forall s \in \mathbb{S} \\
&x_s = x \ , \quad \forall s \in \mathbb{S} \\
&x \in \mathbb{X}
\end{aligned}
$$

- This problem has the same solutions $\big(x, (y_s)_{s\in\mathbb{S}}\big)$ as the original one

Scenarios can be organized like a fan or like a tree



We transform the two-stage stochastic optimization problem from a tree to a fan

- We consider initial decisions $(x_s)_{s \in \mathbb{S}}$ and the problem

$$\begin{aligned}
&\min_{(x_s)_{s \in \mathbb{S}}, (y_s)_{s \in \mathbb{S}}} \sum_{s \in \mathbb{S}} \pi_s j_s\big(x_s, y_s\big) \\
&x_s \in \mathbb{X} \;, \quad \forall s \in \mathbb{S} \\
&y_s \in \mathcal{Y}_s(x_s) \;, \quad \forall s \in \mathbb{S} \\
&x_s = \textstyle\sum_{s' \in \mathbb{S}} \pi_{s'} x_{s'} \;, \quad \forall s \in \mathbb{S}
\end{aligned}$$

- Solutions $(x_s, y_s)_{s \in \mathbb{S}}$ are naturally indexed by a fan

Primal and dual problems

- The primal problem is

$$
\begin{aligned}
&\min_{(x_s, y_s)_{s \in \mathbb{S}}} \max_{(\lambda_s)_{s \in \mathbb{S}}} \sum_{s \in \mathbb{S}} \pi_s \Big(j_s(x_s, y_s) + \lambda_s \big(x_s - \sum_{s' \in \mathbb{S}} \pi_{s'} x_{s'} \big) \Big) \\
&x_s \in \mathbb{X} \; , \quad \forall s \in \mathbb{S} \\
&y_s \in \mathcal{Y}_s(x_s) \; , \quad \forall s \in \mathbb{S}
\end{aligned}
$$

- The dual problem is

$$
\begin{aligned}
&\max_{(\lambda_s)_{s \in \mathbb{S}}} \min_{(x_s, y_s)_{s \in \mathbb{S}}} \sum_{s \in \mathbb{S}} \pi_s \Big(j_s(x_s, y_s) + \lambda_s \big(x_s - \sum_{s' \in \mathbb{S}} \pi_{s'} x_{s'} \big) \Big) \\
&x_s \in \mathbb{X} \; , \quad \forall s \in \mathbb{S} \\
&y_s \in \mathcal{Y}_s(x_s) \; , \quad \forall s \in \mathbb{S}
\end{aligned}
$$

We can translate the multipliers λ_s in the dual problem

- Denote by $\boldsymbol{X} : \mathbb{S} \to \mathbb{X}$ the random variable $\boldsymbol{X}(s) = x_s$, $s \in \mathbb{S}$
- Denote by $\boldsymbol{\Lambda} : \mathbb{S} \to \mathbb{R}$ the random variable $\boldsymbol{\Lambda}(s) = \lambda_s$, $s \in \mathbb{S}$

$$
\begin{aligned}
&\sum_{s \in \mathbb{S}} \pi_s \lambda_s \big(x_s - \sum_{s' \in \mathbb{S}} \pi_{s'} x_{s'} \big) \\
=&\mathbb{E}\big[\boldsymbol{\Lambda}\big(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}]\big)\big] \\
=&\mathbb{E}\big[\boldsymbol{\Lambda}\boldsymbol{X}\big] - \mathbb{E}[\boldsymbol{\Lambda}]\mathbb{E}[\boldsymbol{X}] \\
=&\mathbb{E}\big[\big(\boldsymbol{\Lambda} - \mathbb{E}[\boldsymbol{\Lambda}]\big)\boldsymbol{X}\big] \\
=&\sum_{s \in \mathbb{S}} \pi_s \underbrace{\big(\lambda_s - \sum_{s' \in \mathbb{S}} \pi_{s'} \lambda_{s'}\big)}_{\text{projected multiplier } \overline{\lambda}_s} x_s
\end{aligned}
$$

Restricting the multiplier

Then the dual problem is

$$\max_{(\lambda_s)_{s\in\mathbb{S}}} \min_{(x_s,y_s)_{s\in\mathbb{S}}} \sum_{s\in\mathbb{S}} \pi_s \Big(j_s\big(x_s, y_s\big) + \big(\lambda_s - \sum_{s'\in\mathbb{S}} \pi_{s'} \lambda_{s'}\big) x_s \Big)$$

$$x_s \in \mathbb{X}\ , \quad \forall s \in \mathbb{S}$$

$$y_s \in \mathcal{Y}_s(x_s)\ , \quad \forall s \in \mathbb{S}$$

The dual problem can be decomposed scenario by scenario

- The dual problem

$$\max_{(\lambda_s)_{s\in\mathbb{S}}} \min_{(x_s,y_s)_{s\in\mathbb{S}}} \sum_{s\in\mathbb{S}} \pi_s \Big(j_s(x_s, y_s) + \big(\lambda_s - \sum_{s'\in\mathbb{S}} \pi_{s'} \lambda_{s'}\big) x_s \Big)$$

$$x_s \in \mathbb{X} \ , \quad \forall s \in \mathbb{S}$$

$$y_s \in \mathcal{Y}_s(x_s) \ , \quad \forall s \in \mathbb{S}$$

- is equivalent to

$$\max_{(\lambda_s)_{s\in\mathbb{S}}} \sum_{s\in\mathbb{S}} \pi_s \min_{(x_s,y_s)} \Big(j_s(x_s, y_s) + \big(\lambda_s - \sum_{s'\in\mathbb{S}} \pi_{s'} \lambda_{s'}\big) x_s \Big)$$

$$x_s \in \mathbb{X} \ , \quad \forall s \in \mathbb{S}$$

$$y_s \in \mathcal{Y}_s(x_s) \ , \quad \forall s \in \mathbb{S}$$

Under proper assumptions
— to be seen later, as they require recalls in duality theory —
the dual problem can be solved by an algorithm "à la Uzawa"
yielding the following
scenario decomposition algorithm

Scheme of the scenario decomposition algorithm

Data: step $\rho > 0$, initial multipliers $\left\{\boldsymbol{\lambda}_s^{(0)}\right\}_{s\in\mathbb{S}}$ and mean first decision $\overline{\boldsymbol{x}}^{(0)}$;

Result: optimal first decision $\boldsymbol{x}$;

repeat

forall *scenarios* $s \in \mathbb{S}$ **do**

Solve the deterministic minimization problem for scenario s, with a penalization $+\boldsymbol{\lambda}_s^{(k)}\left(\boldsymbol{x}_s^{(k+1)} - \overline{\boldsymbol{x}}^{(k)}\right)$, and obtain optimal first decision $\boldsymbol{x}_s^{(k+1)}$;

Update the mean first decisions

$$\overline{\boldsymbol{x}}^{(k+1)} = \sum_{s\in\mathbb{S}} \pi_s \boldsymbol{x}_s^{(k+1)} \;;$$

Update the multiplier by

$$\boldsymbol{\lambda}_s^{(k+1)} = \boldsymbol{\lambda}_s^{(k)} + \rho\left(\boldsymbol{x}_s^{(k+1)} - \overline{\boldsymbol{x}}^{(k+1)}\right)\,, \quad \forall s \in \mathbb{S}\;;$$

until $\boldsymbol{x}_s^{(k+1)} - \sum_{s'\in\mathbb{S}} \pi_{s'} \boldsymbol{x}_{s'}^{(k+1)} = 0\;, \quad \forall s \in \mathbb{S}$;

Outline of the presentation

Recalls and exercises on continuous optimization

http://cermics.enpc.fr/~delara/TEACHING/slides_optimization.pdf

Thanks to Pierre Carpentier (ENSTA ParisTech),
we provide a geometric interpretation of Uzawa's algorithm

Uzawa's geometric interpretation

For the sake of simplicity, we consider here equality constraints:

$$u^{(k+1)} \in \underset{u \in \mathcal{U}^{ad}}{\arg\min} \; J(u) + \left\langle \lambda^{(k)} , Y(u) \right\rangle ,$$
$$\lambda^{(k+1)} = \lambda^{(k)} + \rho \Theta(u^{(k+1)}) .$$

The minimization step in u is equivalent to:

$$\min_{v \in \mathcal{V}} \; \min_{u \in \mathcal{U}^{ad}} \; J(u) + \left\langle \lambda^{(k)} , v \right\rangle \quad \text{s.t.} \quad Y(u) - v = 0 .$$

Introducing the perturbation function G given by

$$G(v) \; = \; \min_{u \in \mathcal{U}^{ad}} \; J(u) \quad \text{s.t.} \quad Y(u) - v = 0 ,$$

the minimization step writes equivalently:

$$\min_{v \in \mathcal{V}} G(v) + \left\langle \lambda^{(k)} , v \right\rangle .$$

Uzawa's geometric interpretation II

Recall that the initial problem consists in obtaining $G(0)$...

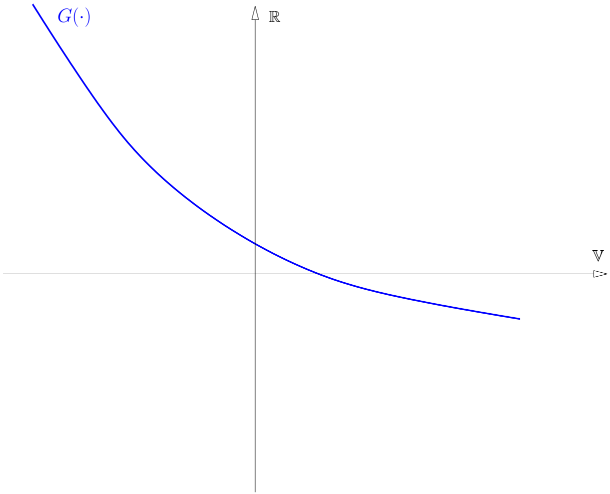
With the help of the function G, Uzawa's algorithm writes:

$$v^{(k+1)} \in \underset{v \in \mathcal{V}}{\arg\min}\, G(v) + \langle \lambda^{(k)}, v \rangle ,$$
$$\lambda^{(k+1)} = \lambda^{(k)} + \rho v^{(k+1)} .$$

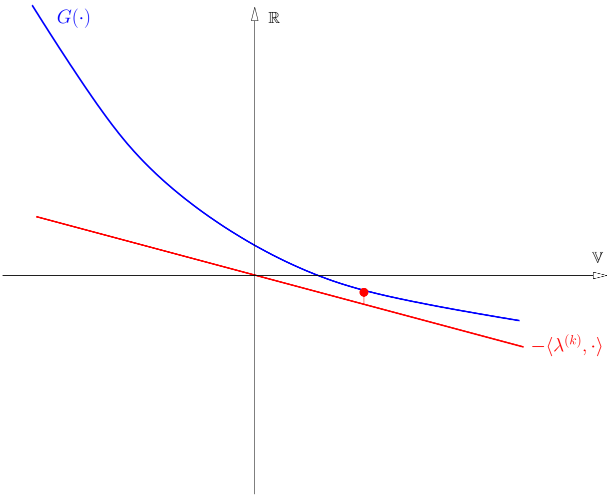
From a (conceptual) geometric point of view, it amounts to:

- **Step (a)**: minimize the gap between $G(\cdot)$ and $\langle -\lambda^{(k)}, \cdot \rangle$.
- **Step (b)**: adjust the slope $-\lambda^{(k)}$ if $v^{(k+1)} \neq 0$.

Uzawa's geometric interpretation III

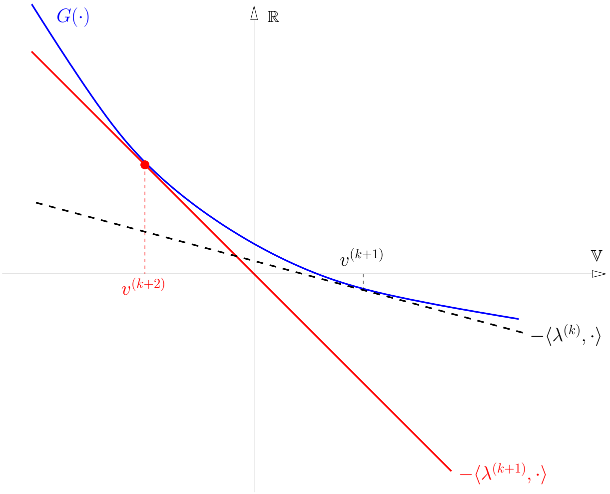


Uzawa's geometric interpretation III

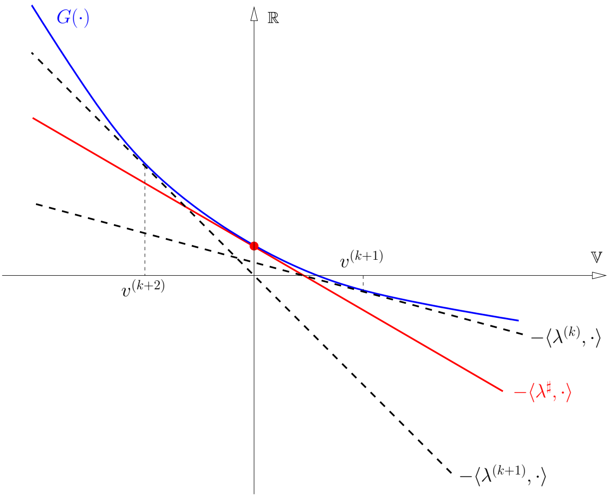


Uzawa's geometric interpretation III

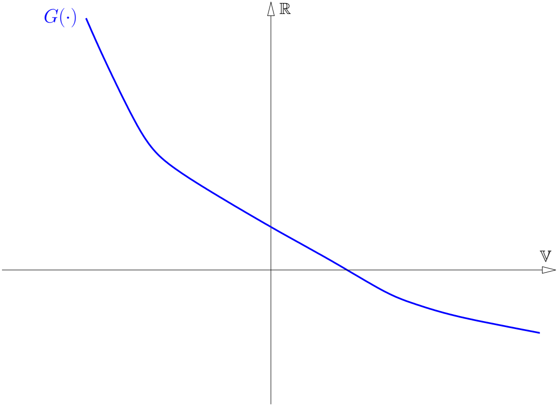
Uzawa's geometric interpretation III



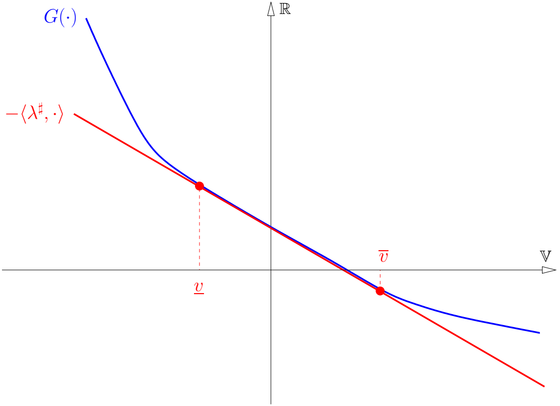
Uzawa's geometric interpretation III



Uzawa's geometric interpretation IV

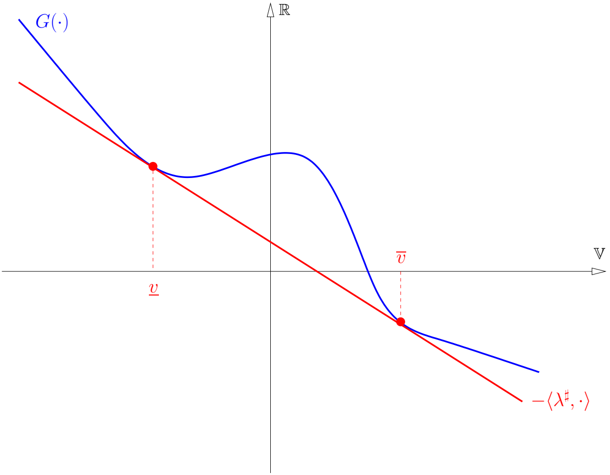


Uzawa's geometric interpretation IV

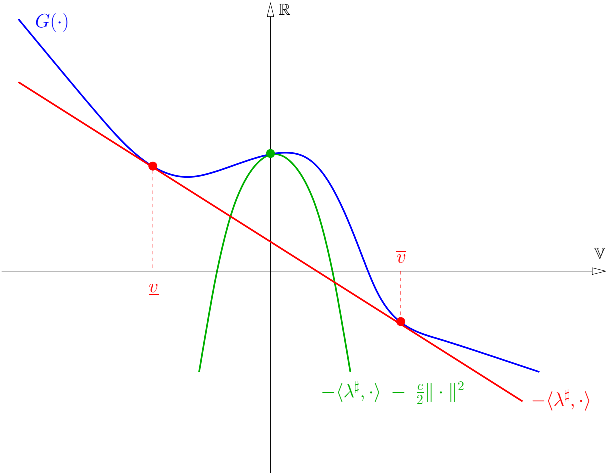


Even if $\{\lambda^{(k)}\}_{k\in\mathbb{N}}$ converges towards $\lambda^\star$, the constraint level $v^{(k)}$ oscillates between $\underline{v}$ and $\overline{v}$, but the value $v^\star = 0$ is never reached

Uzawa's geometric interpretation V



Uzawa's geometric interpretation V



In the non convex case, use an augmented Lagrangian. . .

Outline of the presentation

Progressive Hedging

Rockafellar, R.T., Wets R. J-B.
Scenario and policy aggregation in optimization under uncertainty,
Mathematics of Operations Research, 16, pp. 119-147, 1991

http://cermics.enpc.fr/~delara/TEACHING/

CEA-EDF-INRIA_2012/Roger_Wets4.pdf

The “plus” of Progressive Hedging

- In addition to the variables x_s, we introduce a new variable $\overline{x}$, so that the non-anticipativity constraint becomes $x_s = \overline{x}$
- We dualize this constraint with an augmented Lagrangian term, yielding to an optimization problem with variables $x_{\cdot}, \overline{x}, \lambda$
- When the multiplier λ is fixed,
  we minimize the primal problem which, unfortunately,
  is not separable with respect to scenarios s
- Luckily, we recover separability by solving sequentially “à la Gauss-Seidel”

$$\min_{x_{\cdot}} \mathcal{L}(x_{\cdot}, \overline{x}^{(k)}, \lambda^{(k)})$$
$$\min_{\overline{x}} \mathcal{L}(x_{\cdot}^{(k+1)}, \overline{x}, \lambda^{(k)})$$

because the first problem is separable with respect to scenarios s

Scheme of the Progressive Hedging algorithm

Data: penalty $r > 0$, initial multipliers $\left\{\boldsymbol{\lambda}_s^{(0)}\right\}_{s\in\mathbb{S}}$ and mean first decision $\overline{\boldsymbol{x}}^{(0)}$;

Result: optimal first decision $\boldsymbol{x}$;

repeat

forall *scenarios* $s \in \mathbb{S}$ **do**

Solve the deterministic minimization problem for scenario s, with penalization $+\boldsymbol{\lambda}_s^{(k)}\left(\boldsymbol{x}_s^{(k+1)} - \overline{\boldsymbol{x}}^{(k)}\right) + \frac{r}{2}\left\|\boldsymbol{x}_s^{(k+1)} - \overline{\boldsymbol{x}}^{(k)}\right\|^2$, and obtain optimal first decision $\boldsymbol{x}_s^{(k+1)}$;

Update the mean first decisions

$$\overline{\boldsymbol{x}}^{(k+1)} = \sum_{s\in\mathbb{S}} \pi_s \boldsymbol{x}_s^{(k+1)} \; ;$$

Update the multiplier by

$$\boldsymbol{\lambda}_s^{(k+1)} = \boldsymbol{\lambda}_s^{(k)} + r\left(\boldsymbol{x}_s^{(k+1)} - \overline{\boldsymbol{x}}^{(k+1)}\right) \; , \quad \forall s \in \mathbb{S} \; ;$$

until $\boldsymbol{x}_s^{(k+1)} - \sum_{s'\in\mathbb{S}} \pi_{s'} \boldsymbol{x}_{s'}^{(k+1)} = 0 \; , \quad \forall s \in \mathbb{S}$;

What land have we covered?

- We have introduced one and two-stage optimization problems under uncertainty
- Thanks to a general framework, using risk measures, stochastic and robust optimization appear as (important) special cases
- We have presented resolution methods by scenario decomposition for two-stage optimization problems
- Dealing with multi-stage optimization problems requires specific tools, as is the notion of state

"Self-promotion, nobody will do it for you" ;-)