# Stochastic Optimization
# Decomposition Methods for Two-stage problems

V. Leclère

October 8 2018

École des Ponts ParisTech

UNIVERSITÉ PARIS-EST

# Presentation Outline

# Presentation Outline

# One-Stage Problem

Assume that $\boldsymbol{\xi}$ has a discrete distribution [1], with $\mathbb{P}(\boldsymbol{\xi} = \xi^s) = p^s > 0$ for $s \in [\![1, S]\!]$. Then, the one-stage problem

$$\begin{aligned} \min_{u_0} \quad & \mathbb{E}\Big[L(u_0, \boldsymbol{\xi})\Big] \\ s.t. \quad & g(u_0, \boldsymbol{\xi}) \leq 0, \qquad \mathbb{P} - a.s \end{aligned}$$

can be written

$$\begin{aligned} \min_{u_0} \quad & \sum_{s=1}^{S} p^s L(u_0, \xi^s) \\ s.t \quad & g(u_0, \xi^s) \leq 0, \qquad \forall s \in [\![1, S]\!]. \end{aligned}$$

[1] If the distribution is continuous we can sample and work on the sampled distribution, this is called the Sample Average Approximation approach with lots of guarantee and results

# Newsvendor problem (continued)

We assume that the demand can take value $\{d^s\}_{i\in[\![1,n]\!]}$ with probabilities $\{p^s\}_{i\in[\![1,n]\!]}$.

# Newsvendor problem (continued)

We assume that the demand can take value $\{d^s\}_{i\in[\![1,n]\!]}$ with probabilities $\{p^s\}_{i\in[\![1,n]\!]}$.
In this case the stochastic newsvendor problem reads

$$\begin{aligned} \min_{u} \quad & \sum_{s=1}^{S} p^s \Big( cu - p \min(u, d^s) \Big) \\ s.t. \quad & u \geq 0 \end{aligned}$$

# Recourse Variable

In most problem we can make a correction $u_1$ once the uncertainty is known:

$$u_0 \rightsquigarrow \boldsymbol{\xi}_1 \rightsquigarrow \boldsymbol{u}_1.$$

As the recourse control $u_1$ is a function of $\boldsymbol{\xi}$ it is a random variable. The two-stage optimization problem then reads

$$\begin{aligned}
\min_{u_0, \boldsymbol{u}_1} \quad & \mathbb{E}\Big[L(u_0, \boldsymbol{\xi}, \boldsymbol{u}_1)\Big] \\
s.t. \quad & g(u_0, \boldsymbol{\xi}, \boldsymbol{u}_1) \leq 0, \qquad \mathbb{P} - a.s \\
& \sigma(\boldsymbol{u}_1) \subset \sigma(\boldsymbol{\xi})
\end{aligned}$$

- $u_0$ is called a first stage control
- $\boldsymbol{u}_1$ is called a second stage control. It is a random variable.

# Recourse Variable

In most problem we can make a correction $u_1$ once the uncertainty is known:

$$u_0 \rightsquigarrow \boldsymbol{\xi}_1 \rightsquigarrow \boldsymbol{u}_1.$$

As the recourse control $u_1$ is a function of $\boldsymbol{\xi}$ it is a random variable. The two-stage optimization problem then reads

$$\begin{aligned} \min_{u_0, \boldsymbol{u}_1} \quad & \mathbb{E}\Big[L(u_0, \boldsymbol{\xi}, \boldsymbol{u}_1)\Big] \\ s.t. \quad & g(u_0, \boldsymbol{\xi}, \boldsymbol{u}_1) \leq 0, \qquad \mathbb{P} - a.s \\ & \sigma(\boldsymbol{u}_1) \subset \sigma(\boldsymbol{\xi}) \end{aligned}$$

- $u_0$ is called a first stage control
- $\boldsymbol{u}_1$ is called a second stage control. It is a random variable.

# Two-stage Problem

The extensive formulation of

$$\begin{aligned} \min_{u_0, \boldsymbol{u}_1} \quad & \mathbb{E}\Big[L(u_0, \boldsymbol{\xi}, \boldsymbol{u}_1)\Big] \\ s.t. \quad & g(u_0, \boldsymbol{\xi}, \boldsymbol{u}_1) \leq 0, \qquad \mathbb{P} - a.s \end{aligned}$$

is

$$\begin{aligned} \min_{u_0, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{n} p^s L(u_0, \xi^s, u_1^s) \\ s.t \quad & g(u_0, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1, S]\!]. \end{aligned}$$

It is a deterministic problem that can be solved with standard tools or specific methods.

# Two-stage newsvendor problem I

We can represent the newsvendor problem in a 2-stage framework.

- Let $u_0$ be the number of newspaper bought in the morning.
- let $u_1$ be the number of newspaper sold during the day.

# Two-stage newsvendor problem I

We can represent the newsvendor problem in a 2-stage framework.

- Let $u_0$ be the number of newspaper bought in the morning.
  $\rightsquigarrow$ first stage control
- let $u_1$ be the number of newspaper sold during the day.
  $\rightsquigarrow$ second stage control

# Two-stage newsvendor problem

We can represent the newsvendor problem in a 2-stage framework.

- Let $u_0$ be the number of newspaper bought in the morning.
  $\rightsquigarrow$ first stage control
- let $u_1$ be the number of newspaper sold during the day.
  $\rightsquigarrow$ second stage control

The problem reads

$$\begin{aligned}
\min_{u_0, \boldsymbol{u}_1} \quad & \mathbb{E}\Big[cu_0 - p\boldsymbol{u}_1\Big] & \\
s.t. \quad & u_0 \geq 0 & \\
& \boldsymbol{u}_1 \leq u_0 & \mathbb{P}-as \\
& \boldsymbol{u}_1 \leq \boldsymbol{d} & \mathbb{P}-as \\
& \sigma(\boldsymbol{u}_1) \subset \sigma(\boldsymbol{d}) &
\end{aligned}$$

# Two-stage newsvendor problem II

In extensive formulation the problem reads

$$
\begin{aligned}
\min_{u_0, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s \big( c u_0 - p u_1^s \big) \\
s.t. \quad & u_0 \geq 0 \\
& u_1^s \leq u_0 && \forall s \in [\![1, S]\!] \\
& u_1^s \leq d^s && \forall s \in [\![1, S]\!]
\end{aligned}
$$

Note that there are as many second-stage control $u_1^s$ as there are possible realization of the demand $\boldsymbol{d}$, but only one first-stage control $u_0$.

# Two-stage newsvendor problem II

In extensive formulation the problem reads

$$
\begin{aligned}
\min_{u_0, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s \big( c u_0 - p u_1^s \big) \\
s.t. \quad & u_0 \geq 0 \\
& u_1^s \leq u_0 && \forall s \in [\![1,S]\!] \\
& u_1^s \leq d^s && \forall s \in [\![1,S]\!]
\end{aligned}
$$

Note that there are as many second-stage control $u_1^s$ as there are possible realization of the demand $\boldsymbol{d}$, but only one first-stage control $u_0$.

# Recourse assumptions

- We say that we are in a complete recourse framework, if for all $u_0$, and all possible outcome $\boldsymbol{\xi}$, every control $\boldsymbol{u}_1$ is admissible.
- We say that we are in a relatively complete recourse framework, if for all $u_0$, and all possible outcome $\xi$, there exists a control $\boldsymbol{u}_1$ that is admissible.
- For a lot of algorithms relatively complete recourse is a condition of convergence. It means that there is no induced constraints.

# Presentation Outline

# Two-stage framework : three information models

Consider the problem

$$\min_{\boldsymbol{u}_0, \boldsymbol{u}_1} \mathbb{E}\big[L(\boldsymbol{u}_0, \boldsymbol{\xi}, \boldsymbol{u}_1)\big]$$

- Open-Loop approach : $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$ are deterministic. In this case both controls are choosen without any knowledge of the alea $\boldsymbol{\xi}$. The set of control is small, and an optimal control can be found through specific method (e.g. Stochastic Gradient).
- Two-Stage approach : $\boldsymbol{u}_0$ is deterministic and $\boldsymbol{u}_1$ is measurable with respect to $\boldsymbol{\xi}$. This is the problem tackled by the Stochastic Programming approach.
- Anticipative approach : $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$ are measurable with respect to $\boldsymbol{\xi}$. This approach consists in solving one deterministic problem per possible outcome of the alea, and taking the expectation of the value of this problems.

# Two-stage framework : three information models

Consider the problem

$$\min_{\boldsymbol{u}_0, \boldsymbol{u}_1} \mathbb{E}\big[L(\boldsymbol{u}_0, \boldsymbol{\xi}, \boldsymbol{u}_1\big]$$

- Open-Loop approach : $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$ are deterministic. In this case both controls are choosen without any knowledge of the alea $\boldsymbol{\xi}$. The set of control is small, and an optimal control can be found through specific method (e.g. Stochastic Gradient).
- Two-Stage approach : $\boldsymbol{u}_0$ is deterministic and $\boldsymbol{u}_1$ is measurable with respect to $\boldsymbol{\xi}$. This is the problem tackled by the Stochastic Programming approach.
- Anticipative approach : $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$ are measurable with respect to $\boldsymbol{\xi}$. This approach consists in solving one deterministic problem per possible outcome of the alea, and taking the expectation of the value of this problems.

# Two-stage framework : three information models

Consider the problem

$$\min_{\boldsymbol{u}_0, \boldsymbol{u}_1} \mathbb{E}\big[L(\boldsymbol{u}_0, \boldsymbol{\xi}, \boldsymbol{u}_1\big]$$

- Open-Loop approach : $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$ are deterministic. In this case both controls are choosen without any knowledge of the alea $\boldsymbol{\xi}$. The set of control is small, and an optimal control can be found through specific method (e.g. Stochastic Gradient).
- Two-Stage approach : $\boldsymbol{u}_0$ is deterministic and $\boldsymbol{u}_1$ is measurable with respect to $\boldsymbol{\xi}$. This is the problem tackled by the Stochastic Programming approach.
- Anticipative approach : $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$ are measurable with respect to $\boldsymbol{\xi}$. This approach consists in solving one deterministic problem per possible outcome of the alea, and taking the expectation of the value of this problems.

# Information models for the Newsvendor I

Open-loop :

$$
\begin{aligned}
\min_{u_0,u_1} \quad & \sum_{s=1}^{S} p^s\big(cu_0 - pu_1\big) \\
s.t. \quad & u_0 \geq 0 \\
& u_1 \leq u_0 \\
& u_1 \leq d^s & \forall s \in [\![1, S]\!]
\end{aligned}
$$

# Information models for the Newsvendor II

Two-stage :

$$
\begin{aligned}
\min_{u_0, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s \left( c u_0 - p u_1^s \right) \\
s.t. \quad & u_0 \geq 0 \\
& u_1^s \leq u_0 && \forall s \in [\![1, S]\!] \\
& u_1^s \leq d^s && \forall s \in [\![1, S]\!]
\end{aligned}
$$

# Information models for the Newsvendor III

Anticipative :

$$\begin{aligned} \min_{\{u_0^s, u_1^s\}_{s\in[\![1,n]\!]}} \quad & \sum_{s=1}^{S} p^s \big(cu_0 - pu_1^s\big) & \\ s.t. \quad & u_0^s \geq 0 & \forall s \in [\![1, S]\!] \\ & u_1^s \leq u_0 & \forall s \in [\![1, S]\!] \\ & u_1^s \leq d^s & \forall s \in [\![1, S]\!] \end{aligned}$$

# Comparing the information models

The three information models can be written this way :

$$
\begin{aligned}
\min_{\{u_0^s, u_1^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) & \\
s.t. \quad & u_0^s \geq 0 & \forall s \in [\![1,S]\!] \\
& u_1^s \leq u_0 & \forall s \in [\![1,S]\!] \\
& u_1^s \leq d^s & \forall s \in [\![1,S]\!] \\
& u_0^s = u_0^{s'} & \text{for 2-stage and OL} \\
& u_1^s = u_1^{s'} & \text{for OL}
\end{aligned}
$$

Hence, by simple comparison of constraints we have

$$V^{anticipative} \leq V^{2-stage} \leq V^{OL}.$$

# Comparing the information models

The three information models can be written this way :

$$
\begin{aligned}
\min_{\{u_0^s,u_1^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s,\xi^s,u_1^s) & \\
s.t. \quad & u_0^s \geq 0 & \forall s \in [\![1,S]\!] \\
& u_1^s \leq u_0 & \forall s \in [\![1,S]\!] \\
& u_1^s \leq d^s & \forall s \in [\![1,S]\!] \\
& u_0^s = u_0^{s'} & \text{for 2-stage and OL} \\
& u_1^s = u_1^{s'} & \text{for OL}
\end{aligned}
$$

Hence, by simple comparison of constraints we have

$$V^{anticipative} \leq V^{2-stage} \leq V^{OL}.$$

# Solving the problems

- $V^{OL}$ can be approximated through specific methods (e.g. Stochastic Gradient).
- $V^{2-stage}$ is obtained through Stochastic Programming specific methods. There are two main approaches:
  - Benders decomposition methods (like L-shaped or nested-decomposition methods).
  - Lagrangian decomposition methods (like Progressive-Hedging algorithm).
- $V^{anticipative}$ is difficult to compute exactly but can be estimated through Monte-Carlo approach by drawing a reasonable number of realizations of $\boldsymbol{\xi}$, solving the deterministic problem for each realization $\xi^s$ and taking the means of the value of the deterministic problem.

# Presentation Outline

# Two-stage Problem

The extensive formulation of

$$\begin{aligned} \min_{u_0, \boldsymbol{u}_1} \quad & \mathbb{E}\Big[ L(u_0, \boldsymbol{\xi}, \boldsymbol{u}_1) \Big] \\ s.t. \quad & g(u_0, \boldsymbol{\xi}, \boldsymbol{u}_1) \leq 0, \qquad \mathbb{P} - a.s \end{aligned}$$

is

$$\begin{aligned} \min_{u_0, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0, \xi^s, u_1^s) \\ s.t \quad & g(u_0, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1, S]\!]. \end{aligned}$$

It is a deterministic problem that can be solved with standard tools or specific methods.

# Splitting variables

Extended Formulation

$$\begin{aligned} \min_{u_0, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) \\ s.t \quad & g(u_0, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1, S]\!]. \end{aligned}$$

Can be written

$$\begin{aligned} \min_{u_0^s, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) \\ s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1, S]\!] \\ & u_0^s = u_0^{s'} \qquad \forall s, s' \end{aligned}$$

# Splitting variables

Extended Formulation

$$\begin{aligned} \min_{u_0, \{u_1^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) \\ s.t \quad & g(u_0, \xi^s, u_1^s) \leq 0, \qquad && \forall s \in [\![1, S]\!]. \end{aligned}$$

Can be written

$$\begin{aligned} \min_{u_0^s, \{u_1^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) \\ s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, \qquad && \forall s \in [\![1, S]\!] \\ & u_0^s = \sum_{s'} p^{s'} u_0^{s'} && \forall s \end{aligned}$$

# Dualizing non-anticipativity constraint I

$$
\begin{aligned}
\min_{u_0^s, \{u_1^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) & \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, & \forall s \in [\![1,S]\!] \\
& u_0^s = \sum_{s'} p^{s'} u_0^{s'} & \forall s
\end{aligned}
$$

is equivalent to

$$
\begin{aligned}
\min_{u_0^s, \{u_1^s\}_{s\in[\![1,S]\!]}} \max_{\{\lambda^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) + p^s \lambda^s \Big( u_0^s - \sum_{s'} p^{s'} u_0^{s'} \Big) \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1,S]\!]
\end{aligned}
$$

# Dualizing non-anticipativity constraint I

$$
\begin{aligned}
\min_{u_0^s, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, && \forall s \in [\![1, S]\!] \\
& u_0^s = \sum_{s'} p^{s'} u_0^{s'} && \forall s
\end{aligned}
$$

is equivalent to

$$
\begin{aligned}
\min_{u_0^s, \{u_1^s\}_{s \in [\![1,S]\!]}} \ \max_{\{\lambda^s\}_{s \in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) \\
& + \sum_{s=1}^{S} p^s \lambda^s u_0^s - \sum_{s,s'} p^s \lambda^s p^{s'} u_0^{s'} \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1, S]\!]
\end{aligned}
$$

# Dualizing non-anticipativity constraint I

$$
\begin{aligned}
\min_{u_0^s, \{u_1^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, && \forall s \in [\![1, S]\!] \\
& u_0^s = \sum_{s'} p^{s'} u_0^{s'} && \forall s
\end{aligned}
$$

is equivalent to

$$
\begin{aligned}
\min_{u_0^s, \{u_1^s\}_{s\in[\![1,S]\!]}} \max_{\{\lambda^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) \\
& + \sum_{s=1}^{S} p^s \lambda^s u_0^s - \sum_{s'} \mathbb{E}\left[\lambda\right] p^{s'} u_0^{s'} \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1, S]\!]
\end{aligned}
$$

# Dualizing non-anticipativity constraint

$$
\begin{aligned}
\min_{u_0^s, \{u_1^s\}_{s\in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) & \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, & \forall s \in [\![1,S]\!] \\
& u_0^s = \sum_{s'} p^{s'} u_0^{s'} & \forall s
\end{aligned}
$$

is equivalent to

$$
\begin{aligned}
\min_{u_0^s, \{u_1^s\}_{s\in [\![1,S]\!]}} \max_{\{\lambda^s\}_{s\in [\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s L(u_0^s, \xi^s, u_1^s) & \\
& + \sum_{s=1}^{S} p^s \Big(\lambda^s - \mathbb{E}\big[\lambda\big]\Big) u_0^s & \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, & \forall s \in [\![1,S]\!]
\end{aligned}
$$

# Dualizing non-anticipativity constraint II

Thus, the dual problem reads

$$\max_{\lambda} \quad \min_{u_0^s, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad \sum_{s=1}^{S} p^s \Big( L(u_0^s, \xi^s, u_1^s) + \Big(\lambda^s - \mathbb{E}[\lambda]\Big) u_0^s \Big)$$

$$s.t \quad g(u_0^s, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1, S]\!]$$

The inner minimization problem, for $\lambda$ given, can decompose scenario by scenario, by solving $S$ deterministic problem

$$\min_{u_0^s, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad \Big( L(u_0^s, \xi^s, u_1^s) + \lambda^s u_0^s \Big)$$

$$s.t \quad g(u_0^s, \xi^s, u_1^s) \leq 0$$

# Dualizing non-anticipativity constraint II

Thus, the dual problem reads

$$
\begin{aligned}
\max_{\lambda:\mathbb{E}[\lambda]=0} \min_{u_0^s,\{u_1^s\}_{s\in[\![1,S]\!]}} \quad & \sum_{s=1}^{S} p^s \Big( L(u_0^s, \xi^s, u_1^s) + \Big(\lambda^s \qquad\qquad \Big) u_0^s \Big) \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0, \qquad \forall s \in [\![1, S]\!]
\end{aligned}
$$

The inner minimization problem, for $\lambda$ given, can decompose scenario by scenario, by solving $S$ deterministic problem

$$
\begin{aligned}
\min_{u_0^s,\{u_1^s\}_{s\in[\![1,S]\!]}} \quad & \Big( L(u_0^s, \xi^s, u_1^s) + \lambda^s u_0^s \Big) \\
s.t \quad & g(u_0^s, \xi^s, u_1^s) \leq 0
\end{aligned}
$$

# Price of information

- By weak duality, any $\lambda$ such that $\mathbb{E}\big[\lambda\big] = 0$ will give a lower bound on the 2-stage problem, computed as

$$\sum_{s=1}^{S} p^s \quad \min_{u_0^s, \{u_1^s\}_{s \in [\![1,S]\!]}} \quad \Big( L(u_0^s, \xi^s, u_1^s) + \lambda^s u_0^s \Big)$$
$$s.t \quad g(u_0^s, \xi^s, u_1^s) \leq 0$$

- $\lambda = 0$ lead to the anticipative lower-bound
- If problem is convex, and under some qualification assumptions, there exists an optimal $\lambda^*$, called the price of information, such that the lower bound is tight.

# Progressive Hedging Algorithm

The progressive hedging algorithm build on this decomposition in the following way.

1. Set a price of information $\{\lambda^s\}_{s\in[\![1,S]\!]}$ such that $\mathbb{E}[\lambda] = 0$
2. For each scenario solve

$$\begin{aligned} \min_{u_0^s,\{u_1^s\}_{s\in[\![1,S]\!]}} \quad & L(u_0^s,\xi^s,u_1^s) + \lambda^s u_0^s \\ s.t \quad & g(u_0^s,\xi^s,u_1^s) \leq 0 \end{aligned}$$

3. Compute the mean first control $\bar{u}_0 := \sum_{s=1}^S p^s u_0^s$
4. Update the price of information with

$$\lambda^s := \lambda^s + \rho(u_0^s - \bar{u}_0)$$

5. Go back to 2.

# Progressive Hedging Algorithm

The progressive hedging algorithm build on this decomposition in the following way.

1. Set a price of information $\{\lambda^s\}_{s\in[\![1,S]\!]}$ such that $\mathbb{E}\big[\lambda\big] = 0$
2. For each scenario solve

$$\begin{aligned} \min_{u_0^s,\{u_1^s\}_{s\in[\![1,S]\!]}} \quad & L(u_0^s,\xi^s,u_1^s) + \lambda^s u_0^s + \rho\| u_0^s - \bar{u}_0\|^2 \\ s.t \quad & g(u_0^s,\xi^s,u_1^s) \leq 0 \end{aligned}$$

3. Compute the mean first control $\bar{u}_0 := \sum_{s=1}^S p^s u_0^s$
4. Update the price of information with

$$\lambda^s := \lambda^s + \rho(u_0^s - \bar{u}_0)$$

5. Go back to 2.

# Presentation Outline

# Linear 2-stage stochastic program

Consider the following problem

$$\begin{aligned} \min \quad & \mathbb{E}\Big[c^\top u_0 + \boldsymbol{q}^\top \boldsymbol{u}_1\Big] \\ s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \\ & \boldsymbol{T} u_0 + \boldsymbol{W} \boldsymbol{u}_1 = \boldsymbol{h}, \quad \boldsymbol{u}_1 \geq 0, \qquad \mathbb{P} - a.s. \\ & u_0 \in \mathbb{R}^n, \quad \sigma(\boldsymbol{u}_1) \subset \sigma\underbrace{(\boldsymbol{q}, \boldsymbol{T}, \boldsymbol{W}, \boldsymbol{h})}_{\boldsymbol{\xi}} \end{aligned}$$

Which we rewrite

$$\begin{aligned} \min_{u_0 \geq 0} \quad & c^\top u_0 + \mathbb{E}\Big[Q(u_0, \xi)\Big] \\ s.t. \quad & Au_0 = b \end{aligned}$$

with

$$\begin{aligned} Q(u_0, \xi) := \min_{u_1 \geq 0} \quad & q_\xi^\top u_1 \\ s.t. \quad & W_\xi u_1 = h_\xi - T_\xi u_0 \end{aligned}$$

# Linear 2-stage stochastic program

Consider the following problem

$$
\begin{aligned}
\min \quad & \mathbb{E}\Big[c^\top u_0 + \boldsymbol{q}^\top \boldsymbol{u}_1\Big] \\
s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \\
& \boldsymbol{T} u_0 + \boldsymbol{W} \boldsymbol{u}_1 = \boldsymbol{h}, \quad \boldsymbol{u}_1 \geq 0, \qquad \mathbb{P} - a.s. \\
& u_0 \in \mathbb{R}^n, \quad \sigma(\boldsymbol{u}_1) \subset \sigma(\underbrace{\boldsymbol{q}, \boldsymbol{T}, \boldsymbol{W}, \boldsymbol{h}}_{\boldsymbol{\xi}})
\end{aligned}
$$

Which we rewrite

$$
\begin{aligned}
\min_{u_0 \geq 0} \quad & c^\top u_0 + \mathbb{E}\Big[Q(u_0, \boldsymbol{\xi})\Big] \\
s.t. \quad & Au_0 = b
\end{aligned}
$$

with

$$
\begin{aligned}
Q(u_0, \xi) := \min_{u_1 \geq 0} \quad & q_\xi^\top u_1 \\
s.t. \quad & W_\xi u_1 = h_\xi - T_\xi u_0
\end{aligned}
$$

# Linear 2-stage stochastic program

Consider the following problem

$$\begin{aligned} \min \quad & \mathbb{E}\Big[c^\top u_0 + \boldsymbol{q}^\top \boldsymbol{u}_1\Big] \\ s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \\ & \boldsymbol{T} u_0 + \boldsymbol{W} \boldsymbol{u}_1 = \boldsymbol{h}, \quad \boldsymbol{u}_1 \geq 0, \qquad \mathbb{P} - a.s. \\ & u_0 \in \mathbb{R}^n, \quad \sigma(\boldsymbol{u}_1) \subset \sigma(\underbrace{\boldsymbol{q}, \boldsymbol{T}, \boldsymbol{W}, \boldsymbol{h}}_{\boldsymbol{\xi}}) \end{aligned}$$

Which we rewrite

$$\begin{aligned} \min_{u_0 \geq 0} \quad & c^\top u_0 + \mathbb{E}\Big[Q(u_0, \boldsymbol{\xi})\Big] \\ s.t. \quad & Au_0 = b \end{aligned}$$

with

$$\begin{aligned} Q(u_0, \xi) := \min_{u_1 \geq 0} \quad & q_\xi^\top u_1 \\ s.t. \quad & W_\xi u_1 = h_\xi - T_\xi u_0 \end{aligned}$$

# Linear 2-stage stochastic program : Extensive Formulation

The associated extensive formulation read

$$\begin{aligned} \min \quad & c^\top u_0 + \sum_{s=1}^{S} p^s \, q^s \cdot u_1^s \\ s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \\ & T^s u_0 + W^s u_1^s = h^s, \quad u_1^s \geq 0, \forall s \end{aligned}$$

Which we rewrite

$$\begin{aligned} \min_{u_0} \quad & c^\top u_0 + \sum_{s} p^s Q^s(u_0) \\ s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \end{aligned}$$

with

$$\begin{aligned} Q^s(u_0) := \min_{u_1 \geq 0} \quad & q^s \cdot u_1 \\ s.t. \quad & W^s u_1 = h^s - T^s u_0 \end{aligned}$$

# Linear 2-stage stochastic program : Extensive Formulation

The associated extensive formulation read

$$\begin{aligned} \min \quad & c^\top u_0 + \sum_{s=1}^{S} p^s \, q^s \cdot u_1^s \\ s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \\ & T^s u_0 + W^s u_1^s = h^s, \quad u_1^s \geq 0, \forall s \end{aligned}$$

Which we rewrite

$$\begin{aligned} \min_{u_0} \quad & c^\top u_0 + \sum_{s} p^s Q^s(u_0) \\ s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \end{aligned}$$

with

$$\begin{aligned} Q^s(u_0) := \min_{u_1 \geq 0} \quad & q^s \cdot u_1 \\ s.t. \quad & W^s u_1 = h^s - T^s u_0 \end{aligned}$$

# Linear 2-stage stochastic program : Extensive Formulation

The associated extensive formulation read

$$
\begin{aligned}
\min \quad & c^\top u_0 + \sum_{s=1}^{S} p^s \, q^s \cdot u_1^s \\
s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \\
& T^s u_0 + W^s u_1^s = h^s, \quad u_1^s \geq 0, \forall s
\end{aligned}
$$

Which we rewrite

$$
\begin{aligned}
\min_{u_0} \quad & c^\top u_0 + \sum_s p^s Q^s(u_0) \\
s.t. \quad & Au_0 = b, \quad u_0 \geq 0
\end{aligned}
$$

with

$$
\begin{aligned}
Q^s(u_0) := \min_{u_1 \geq 0} \quad & q^s \cdot u_1 \\
s.t. \quad & W^s u_1 = h^s - T^s u_0
\end{aligned}
$$

# Relatively complete recourse

We assume here relatively complete recourse. Without this assumption we would need feasability cuts (see Bender's decomposition method).

Here, relatively complete recourse means that, for $u_0 \geq 0$ :

$$Au_0 = b \quad \Longrightarrow \quad Q_s(u_0) < +\infty, \quad \forall s \in [\![1, S]\!]$$

# Decomposition of linear 2-stage stochastic program

We rewrite the extended formulation as

$$\begin{aligned} \min \quad & c^\top u_0 + \sum_s p^s \theta^s \\ s.t. \quad & Au_0 = b, \quad u_0 \geq 0 \\ & \theta^s \geq Q^s(u_0) \qquad\qquad u_0 \in \mathbb{R}^n, \quad \forall s \end{aligned}$$

Note that $Q^s(u_0)$ is a polyhedral function of $u_0$, hence $\theta^s \geq Q^s(u_0)$ can be rewritten $\theta \geq \alpha_k^s \cdot u_0 + \beta_k^s, \forall k$.
The decomposition approach consists in constructing iteratively cut coefficients $\alpha_k^s$ and $\beta_k^s$.

# Obtaining (optimality) cuts I

Recall that

$$Q^s(u_0) := \min_{u_1^s \in \mathbb{R}^n} \quad q^s \cdot u_1^s$$
$$s.t. \quad W^s u_1^s = h^s - T^s u_0, \quad u_1^s \geq 0$$

can also be written (through strong duality by relatively complete recourse assumption)

$$(D_{u_0}) \quad Q^s(u_0) = \max_{\lambda^s \in \mathbb{R}^m} \quad \lambda^s \cdot \left(h^s - T^s u_0\right)$$
$$s.t. \quad (W^s)^\top \lambda^s \leq q^s$$

# Obtaining (optimality) cuts II

$$
(D_{u_0}) \quad Q^s(u_0) = \max_{\lambda^s \in \mathbb{R}^m} \quad \lambda^s \cdot (h^s - T^s u_0) \qquad s.t. \quad (W^s)^\top \lambda^s \leq q^s
$$

admits for optimal solution $\lambda^s_{u_0}$.
Consider another control $u_0'$, we have

$$
(D_{u_0'}) \quad Q^s(u_0') = \max_{\lambda^s \in \mathbb{R}^m} \quad \lambda^s \cdot (h^s - T^s u_0') \qquad s.t. \quad (W^s)^\top \lambda^s \leq q^s
$$

As $\lambda^s_{u_0}$ is admissible for $(D_{u_0})$ it is also admissible for $(D_{u_0'})$, hence

$$
Q^s(u_0') \geq \lambda^s_{u_0} \cdot (h^s - T^s u_0').
$$

# Obtaining (optimality) cuts III

To sum up we have seen that, for any admissible first stage solution, we can construct an exact cut for $Q^s$ by solving the dual of the second stage problem.

More precisely, let $u_0^k \geq 0$ be such that $Au_0^k = b$. Let $\lambda_k^s$ be an optimal dual solution. Then, setting

$$\alpha_k^s := -(T^s)^\top \lambda_k^s \quad \text{and} \quad \beta_k^s := (\lambda_k^s)^\top h^s$$

we have

$$\begin{cases} Q^s(u_0') \geq \alpha_k^s \cdot u_0' + \beta_k^s & \forall u_0' \geq 0, Au_0 = b \\ Q^s(u_0^k) = \alpha_k^s \cdot u_0^k + \beta_k^s \end{cases}$$

# L-shaped method (multi-cut version)

1. We have a collection of $K \times S$ cuts, such that $Q^s(u_0) \geq \alpha_k^s \cdot u_0 + \beta_k^s$.
2. Solve the master problem, with optimal primal solution $u_0^{K+1}$.

$$\begin{aligned}
\min_{u_0 \geq 0} \quad & c^\top u_0 + \sum_{s=1}^{S} p^s \theta^s \\
s.t. \quad & A u_0 = b \\
& \theta^s \geq \alpha_k^s u_0 + \beta_k^s \qquad \forall k \in [\![1, K]\!], \quad \forall s \in [\![1, S]\!]
\end{aligned}$$

3. Solve $S$ slave dual problems, with optimal dual solution $\lambda_{K+1}^s$

$$\begin{aligned}
\max_{\lambda^s \in \mathbb{R}^m} \quad & \lambda^s \cdot \left(h^s - T^s u_0^{K+1}\right) \\
s.t. \quad & W^s \cdot \lambda^s \leq q^s
\end{aligned}$$

4. construct $S$ new cuts with

$$\alpha_{K+1}^s := -(T^s)^\top \lambda^s, \qquad \beta_{K+1}^s := h^s \cdot \lambda_{K+1}^s$$

# L-shaped method (multi-cut version) : bounds

- At any iteration of the L-shaped method we can easily determine upper and lower bound over our problem.
- Indeed, $u_0^K$ is an admissible firt stage solution, and $Q^s(u_0^K)$ is the value of a slave problem. Thus the value of admissible solution $u_0^k$ is simply given by

$$UB = c^\top u_0^K + \sum_{s=1}^{S} p^s Q^s(u_0^K).$$

- Furthermore, $Q_K^s(u_0) \geq \max_{k \leq K} \alpha_k^s \cdot u_0 + \beta_k^s$, thus the value of the master problem is always a lower bound over the value of the SP problem :

$$LB = c^\top u_0^K + \sum_{s=1}^{S} p^s \theta_K^s.$$

# L-shaped method (single-cut version)

1. We have a collection of $K$ cuts, such that $Q(u_0) \geq \alpha_k \cdot u_0 + \beta_k$.
2. Solve the master problem, with optimal primal solution $u_0^{K+1}$.

$$
\begin{aligned}
\min_{u_0 \geq 0} \quad & c^\top u_0 + \theta \\
s.t. \quad & A u_0 = b \\
& \theta \geq \alpha_k u_0 + \beta_k \qquad \forall k \in [\![1, K]\!]
\end{aligned}
$$

3. Solve $S$ slave dual problems, with optimal dual solution $\lambda^s_{K+1}$

$$
\begin{aligned}
\max_{\lambda^s \in \mathbb{R}^m} \quad & \lambda^s \cdot \left( h^s - T^s u_0^{K+1} \right) \\
s.t. \quad & W^s \cdot \lambda^s \leq q^s
\end{aligned}
$$

4. construct new cut with

$$
\alpha_{K+1} := -\sum_{i=1}^{S} p^s \, (T^s)^\top \lambda^s, \qquad \beta_{K+1} := \sum_{i=1}^{S} p^s \, h^s \cdot \lambda^s.
$$