Stochastic Optimization
Two-stage problems

V. Leclère

October 5 2017

École des Ponts ParisTech

UNIVERSITÉ PARIS-EST

Presentation Outline

1. Two-stage Stochastic Programming
2. Some information frameworks
3. L-Shaped decomposition method

Presentation Outline

1. Two-stage Stochastic Programming
2. Some information frameworks
3. L-Shaped decomposition method

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 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 $\boldsymbol{u}_1$ 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 ξ, there exists a control $\boldsymbol{u}_1$ that is admissible.
- For a lot of algorithm relatively complete recourse is a condition of convergence. It means that there is no induced constraints.

Presentation Outline

1. Two-stage Stochastic Programming
2. Some information frameworks
3. L-Shaped decomposition method

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 :

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

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 :

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

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:
  - Lagrangian decomposition methods (like Progressive-Hedging algorithm).
  - Benders decomposition methods (like L-shaped or nested-decomposition methods).
- $V^{anticipative}$ 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 ξ^s and taking the means of the value of the deterministic problem.

Presentation Outline

1 Two-stage Stochastic Programming

2 Some information frameworks

3 L-Shaped decomposition method

Linear 2-stage stochastic program

Consider the following problem

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

With associated Extended Formulation

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

Linear 2-stage stochastic program

Consider the following problem

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

With associated Extended Formulation

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

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 :

$$\forall u_0 \geq 0, \quad \forall s \in [\![1, S]\!] \qquad Au_0 = b \quad \Longrightarrow \quad \exists u_1^s \geq 0, \quad W^s u_1^s = h^s - T^s u_0.$$

Decomposition of linear 2-stage stochastic program

We rewrite the extended formulation as

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

where $Q(u_0) = \sum_{s=1}^{S} p^s Q^s(u_0)$ with

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

Note that $Q(u_0)$ is a polyhedral function of u_0, hence $\theta \geq Q(u_0)$ can be rewritten $\theta \geq \alpha_k^\top u_0 + \beta_k, \forall k$.

Obtaining (optimality) cuts

Recall that

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

can also be written (through strong duality)

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

Obtaining (optimality) cuts II

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

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

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

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

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

Obtaining (optimality) cuts III

Thus we have that,

$$\forall u_0', \qquad Q^s(u_0') \geq \underbrace{h^s \cdot \lambda^s_{u_0}}_{\beta^s_k} \underbrace{-\lambda^s \cdot T^s}_{\alpha^s_k} u_0'.$$

Recall that,

$$\forall u_0', \qquad Q(u_0') = \sum_{s=1}^{S} p^s Q^s(u_0')$$

thus

$$\forall u_0', \qquad Q(u_0') \geq \sum_{s=1}^{S} p^s (\alpha^s_k u_0' + \beta^s_k)$$

or

$$\forall u_0', \qquad Q(u_0') \geq \underbrace{\Big(\sum_{s=1}^{S} p^s \alpha^s_k\Big)}_{\alpha_k} u_0' + \underbrace{\sum_{s=1}^{S} p^s \beta^s_k}_{\beta_k}$$

L-shaped method

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

$$\begin{aligned} \min_{Au_0=b,u_0\geq 0} \quad & c^\top u_0 + \theta \\ s.t. \quad & \theta \geq \alpha_k u_0 + \beta_k \qquad \forall k \in [\![1, K]\!] \end{aligned}$$

3. Solve N slave dual problems, with optimal dual solution λ^s

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

4. construct new cut with

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