Optimisation under Uncertainty
Session 1/4

I Workshop de Otimização sob Incerteza - UFSCar

Fabricio Oliveira

Systems Analysis Laboratory
Department of Mathematics and Systems Analysis

Aalto University
School of Science

August 16, 2023

Outline of this lecture

Outline of this lecture

Introduction

Two-stage stochastic programming

Measures of quality: EVPI and VSS

Recourse types

Multi-stage problems

The farmer's problem [Birge and Louveaux, 2011]

A farmer has available 500 acres of land available for raising wheat, corn, and sugar beets.

- She needs at least 200 tons of wheat and 240 tons of corn for cattle feed;
- Cattle feed amounts can be raised on the farm or bought from a wholesale market;
- Sugar beet is raised for profit only. However, production above a 6000 ton quota has a lower sales price.

The farmer's problem [Birge and Louveaux, 2011]

A farmer has available 500 acres of land available for raising wheat, corn, and sugar beets.

- She needs at least 200 tons of wheat and 240 tons of corn for cattle feed;
- Cattle feed amounts can be raised on the farm or bought from a wholesale market;
- Sugar beet is raised for profit only. However, production above a 6000 ton quota has a lower sales price.

	Wheat	**Corn**	**Sugar beets**
Yield (ton/acre)	2.5	3	20
Planting cost ($/acre)	150	230	260
Selling price ($/ton)	170	150	36 (under 6000 ton)
			10 (above 6000 ton)
Purchase price ($/ton)	238	210	-

Table: Framer's problem data

The farmer's problem [Birge and Louveaux, 2011]

Let $I = \{1 : \text{wheat}, 2 : \text{corn}, 3 : \text{sugar beets}\}$. Then

- x_i - acres devoted to i;
- y_i - tons of i purchased, $i \in I \setminus \{3\}$;
- w_i - tons of i sold, $i \in I \cup \{4\}$, $\{4 : \text{sugar beets (over quota)}\}$.

The farmer's problem [Birge and Louveaux, 2011]

Let $I = \{1 : \text{wheat}, 2 : \text{corn}, 3 : \text{sugar beets}\}$. Then

- x_i - acres devoted to i;
- y_i - tons of i purchased, $i \in I \setminus \{3\}$;
- w_i - tons of i sold, $i \in I \cup \{4\}$, $\{4 : \text{sugar beets (over quota)}\}$.

The farmer's problem is:

$$\begin{aligned}
\text{min. } & 150x_1 + 230x_2 + 260x_3 + 238y_1 + 210y_2 \\
& - 170w_1 - 150w_2 - 36w_3 - 10w_4 \\
\text{s.t.: } & x_1 + x_2 + x_3 \leq 500 \\
& 2.5x_1 + y_1 - w_1 \geq 200 \\
& 3x_2 + y_2 - w_2 \geq 240 \\
& w_3 + w_4 \leq 20x_3 \\
& w_3 \leq 6000 \\
& x_i \geq 0, i \in I; y_i \geq 0, i \in I \setminus \{3\}; w_i \geq 0, i \in I \cup \{4\}.
\end{aligned}$$

The farmer's problem [Birge and Louveaux, 2011]

The optimal solution is given by:

Optimal strategy:

1. plant sugar beets to reach quota,
2. satisfy minimum requirements
3. if in excess, sell the excess as wheat; if in shortage, purchase corn.

	Wheat	**Corn**	**Sugar beets**
Surface	120	80	300
Yield	300	240	6000
Sales	100	-	6000
Purchase	-	-	-
Overall profit:			$118,600

Table: Optimal solution: average yields

The farmer's problem [Birge and Louveaux, 2011]

The optimal solution is given by:

Optimal strategy:

1. plant sugar beets to reach quota,
2. satisfy minimum requirements
3. if in excess, sell the excess as wheat; if in shortage, purchase corn.

	Wheat	**Corn**	**Sugar beets**
Surface	120	80	300
Yield	300	240	6000
Sales	100	-	6000
Purchase	-	-	-
Overall profit:			\$118,600

Table: Optimal solution: average yields

- The farmer is well aware that climate factors influence crop yields, which can fluctuate $\pm$ 20%.
- How can we take these scenarios into account for making land allocation decisions?

The farmer's problem [Birge and Louveaux, 2011]

Considering the scenarios individually, we obtain:

	Wheat	**Corn**	**Sugar beets**
Surface	183.33	66.67	250
Yield	550	240	6000
Sales	350	-	6000
Purchase	-	-	-
Overall profit:			$167,667

Table: Optimal solution: 20% higher yields

The farmer's problem [Birge and Louveaux, 2011]

Considering the scenarios individually, we obtain:

	Wheat	**Corn**	**Sugar beets**
Surface	183.33	66.67	250
Yield	550	240	6000
Sales	350	-	6000
Purchase	-	-	-
Overall profit:			\$167,667

Table: Optimal solution: 20% higher yields

	Wheat	**Corn**	**Sugar beets**
Surface	100	25	375
Yield	200	60	6000
Sales	-	-	6000
Purchase	-	180	-
Overall profit:			\$59,950

Table: Optimal solution: 20% lower yields

The farmer's problem [Birge and Louveaux, 2011]

As one may notice, the land allocation for sugar beets is the critical factor:

- Planting too much $\Rightarrow$ losses for selling above quota
- Planting too little $\Rightarrow$ opportunity losses

The farmer's problem [Birge and Louveaux, 2011]

As one may notice, the land allocation for sugar beets is the critical factor:

- Planting too much $\Rightarrow$ losses for selling above quota
- Planting too little $\Rightarrow$ opportunity losses

We can hedge against this uncertainty by taking a long-term perspective:

- We assume that each year one of these scenarios happens.
- We know they are equally likely to happen, but exactly which will happen cannot be known.
- Thus, maximise long-run profit $\Rightarrow$ maximise expected profit.

The farmer's problem [Birge and Louveaux, 2011]

Let $S = \{1 : \text{-20\%}, 2 : \text{avg.}, 3 : \text{+20\%}\}$ represent the yield scenarios.

The reformulated farmer's problem is:

$$
\begin{aligned}
\text{min. } & 150x_1 + 230x_2 + 260x_3 + \\
& \frac{1}{3}(238y_{11} + 210y_{21} - 170w_{11} - 150w_{21} - 36w_{31} - 10w_{41}) \\
& \frac{1}{3}(238y_{12} + 210y_{22} - 170w_{12} - 150w_{22} - 36w_{32} - 10w_{42}) \\
& \frac{1}{3}(238y_{13} + 210y_{23} - 170w_{13} - 150w_{23} - 36w_{33} - 10w_{43}) \\
\text{s.t.: } & x_1 + x_2 + x_3 \leq 500 \\
& 2x_1 + y_{11} - w_{11} \geq 200,\ 2.5x_1 + y_{12} - w_{12} \geq 200,\ 3x_1 + y_{13} - w_{13} \geq 200 \\
& 2.4x_2 + y_{21} - w_{21} \geq 240,\ 3x_2 + y_{22} - w_{22} \geq 240,\ 3.6x_2 + y_{23} - w_{23} \geq 240 \\
& w_{31} + w_{41} \leq 16x_3,\ w_{32} + w_{42} \leq 20x_3,\ w_{33} + w_{43} \leq 24x_3 \\
& w_{31} \leq 6000, w_{32} \leq 6000, w_{33} \leq 6000 \\
& x_i \geq 0, i \in I; y_{is} \geq 0, i \in I \setminus \{3\}, s \in S; w_{is} \geq 0, i \in I \cup \{4\}, s \in S.
\end{aligned}
$$

The farmer's problem [Birge and Louveaux, 2011]

The optimal solution becomes:

		Wheat	Corn	Sugar beets
	Surface	170	80	250
$s = 1$	**Yield**	340	192	4000
	Sales	140	-	4000
	Purchase	-	48	-
$s = 2$	**Yield**	422	240	5000
	Sales	225	-	5000
	Purchase	-	-	-
$s = 3$	**Yield**	510	288	6000
	Sales	310	48	6000
	Purchase	-	-	-
Overall profit:				\$108,390

Table: Optimal solution: all scenario yields

The farmer's problem [Birge and Louveaux, 2011]

By doing so, the farmer takes into account all scenarios simultaneously.

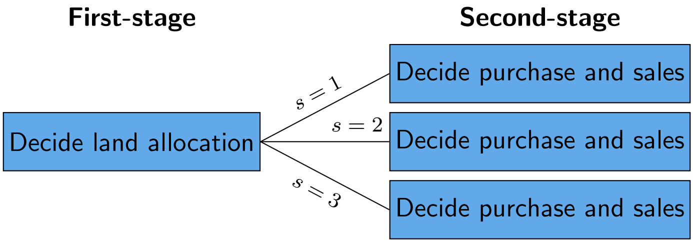
1. The farmer exploits the timing between making decisions and observing uncertainties;
2. This "hedging" comes with a "price" that can be estimated against perfect information performance.

The farmer's problem [Birge and Louveaux, 2011]

By doing so, the farmer takes into account all scenarios simultaneously.

1. The farmer exploits the timing between making decisions and observing uncertainties;
2. This "hedging" comes with a "price" that can be estimated against perfect information performance.

Effectively, this is achieved by incorporating decision stages into the model:



Two-stage stochastic programming

More formally, let:

- x - definite (first-stage) decisions;
- y - corrective or recourse (second-stage) decisions;
- ξ - random variable;
- $[q(\xi), T(\xi), W(\xi), h(\xi)]$ - random vector (data);

Two-stage stochastic programming

More formally, let:

- x - definite (first-stage) decisions;
- y - corrective or recourse (second-stage) decisions;
- ξ - random variable;
- $[q(\xi), T(\xi), W(\xi), h(\xi)]$ - random vector (data);

The formulation of a two-stage stochastic programming (2SSP) model is

$$\text{min. } c^\top x + \mathcal{Q}(x) \tag{1a}$$

$$\text{s.t.: } Ax = b \tag{1b}$$

$$x \geq 0, \tag{1c}$$

where $\mathcal{Q}(x) = \mathbb{E}_\xi \left[Q(x, \xi) \right]$ and

$$Q(x, \xi) = \left\{ \text{min. } q(\xi)^\top y : W(\xi)y = h(\xi) - T(\xi)x, y \geq 0 \right\}. \tag{2}$$

Outline of this lecture

Introduction

Two-stage stochastic programming

Measures of quality: EVPI and VSS

Recourse types

Multi-stage problems

Two-stage stochastic programming

In essence, we are assuming the following decision process

$$x \to \xi \to y(\xi, x)$$

in which y is chosen to minimise $\mathbb{E}_{\xi}\left[Q(x,\xi)\right]$, assuming a known probability distribution for ξ with support Ξ.

Two-stage stochastic programming

In essence, we are assuming the following decision process

$$x \to \xi \to y(\xi, x)$$

in which y is chosen to minimise $\mathbb{E}_\xi [Q(x, \xi)]$, assuming a known probability distribution for ξ with support Ξ.

Thus, we can pose problem (1) as the semi-infinite problem

$$\begin{aligned}
\text{min. } & c^\top x + \mathbb{E}_\xi [Q(x, \xi)] && (3a)\\
\text{s.t.: } & Ax = b, x \geq 0 && (3b)\\
& T(\xi)x + W(\xi)y(\xi) = h(\xi), \ \forall \xi \in \Xi && (3c)\\
& y(\xi) \geq 0, \ \forall \xi \in \Xi. && (3d)
\end{aligned}$$

Two-stage stochastic programming

There are two complicating factors in (3):

1. $\mathbb{E}_\xi \left[Q(x, \xi) \right]$;
2. $\forall \xi \in \Xi$.

Two-stage stochastic programming

There are two complicating factors in (3):

1. $\mathbb{E}_\xi \left[Q(x, \xi)\right]$;
2. $\forall \xi \in \Xi$.

Those are treated by means of discretisation, that is:

- In general, we assume Ξ to be a discrete and finite set;
- Each realisations $\xi_s \in \Xi$, for $s \in S \equiv \{1, \dots, |\Xi|\}$ is a scenario;
- Thus, $[q(\xi), T(\xi), W(\xi), h(\xi)]$ becomes a discrete and finite set of parameters:

$$
\begin{aligned}
& [q(\xi_1), T(\xi_1), W(\xi_1), h(\xi_1); \\
& \; q(\xi_2), T(\xi_2), W(\xi_2), h(\xi_2); \\
& \; \dots; \\
& \; q(\xi_{|\Xi|}), T(\xi_{|\Xi|}), W(\xi_{|\Xi|}), h(\xi_{|\Xi|})] \\
\Rightarrow \; & [q_s, T_s, W_s, h_s] = \xi_s, s \in S.
\end{aligned}
$$

Two-stage stochastic programming

Considering a finite and discrete set of scenarios, we can restate (3) as its deterministic equivalent

$$\begin{aligned} \text{min. } & c^\top x + \sum_{s \in S} P_s q_s^\top y_s && \text{(4a)} \\ \text{s.t.: } & Ax = b, x \geq 0 && \text{(4b)} \\ & T_s x + W_s y_s = h_s, \ \forall s \in S && \text{(4c)} \\ & y_s \geq 0, \ \forall s \in S. && \text{(4d)} \end{aligned}$$

Two-stage stochastic programming

Considering a finite and discrete set of scenarios, we can restate (3) as its deterministic equivalent

$$
\begin{aligned}
\text{min. } & c^\top x + \sum_{s \in S} P_s q_s^\top y_s && \text{(4a)} \\
\text{s.t.: } & Ax = b, x \geq 0 && \text{(4b)} \\
& T_s x + W_s y_s = h_s, \ \forall s \in S && \text{(4c)} \\
& y_s \geq 0, \ \forall s \in S. && \text{(4d)}
\end{aligned}
$$

Remark: notice how discretisation solves the tractability issues:

1. P_s is the probability associated with scenario s ($P_s = P(\xi = \xi_s)$). Thus $\mathbb{E}_\xi \left[Q(x, \xi) \right] = \sum_{s \in S} P_s q_s^\top y_s$.
2. (4) has a finite number of variables and constraints.

Two-stage stochastic programming

In the farmer's problem, there we assumed that $s \in \{1, 2, 3\}$:

- $q_s = q_{s'}$, $W_s = W_{s'}$, and $h_s = h_{s'}$, $\forall s, s' \in S \mid s \neq s'$;
- $T_s = [t_1(s), t_2(s)]$.

Two-stage stochastic programming

In the farmer's problem, there we assumed that $s \in \{1, 2, 3\}$:

- $q_s = q_{s'}$, $W_s = W_{s'}$, and $h_s = h_{s'}$, $\forall s, s' \in S \mid s \neq s'$;
- $T_s = [t_1(s), t_2(s)]$.

And thus, we had that:

$$
\begin{aligned}
Q_s(x) = \text{min.} \quad & 238y_1(s) - 170w_1(s) + 210y_2(s) - 150w_2(s) \\
& \quad - 36w_3(s) - 10w_4(s) \\
\text{s.t.:} \quad & t_1(s)x_1 + y_1(s) - w_1(s) \geq 200 \\
& t_2(s)x_2 + y_2(s) - w_2(s) \geq 240 \\
& w_3(s) + w_4(s) \leq t_3(s)x_3 \\
& w_3(s) \leq 6000 \\
& y_1(s), w_1(s), y_2(s), w_2(s), w_3(s), w_4(s) \geq 0.
\end{aligned}
$$

Outline of this lecture

Introduction

Two-stage stochastic programming

Measures of quality: EVPI and VSS

Recourse types

Multi-stage problems

The expected value of perfect information (EVPI)

The performance of the solution of a 2SSP can be compared against a so-called wait-and-see (WS) solution.

Let our 2SSP be compactly represented as

$$z = \min_x . \left\{ \mathbb{E}\left[F(x, \xi) \right] \right\},$$

where $F(x, \xi) = \left\{ c^\top x + \mathbb{E}\left[Q(x, \xi) \right] : x \in X \right\}$, $Q(x, \xi)$ is defined as (2), and $X = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$.

The expected value of perfect information (EVPI)

The performance of the solution of a 2SSP can be compared against a so-called wait-and-see (WS) solution.

Let our 2SSP be compactly represented as

$$z = \underset{x}{\mathsf{min.}} \ \{\mathbb{E}\left[F(x,\xi)\right]\},$$

where $F(x,\xi) = \left\{c^\top x + \mathbb{E}\left[Q(x,\xi)\right] : x \in X\right\}$, $Q(x,\xi)$ is defined as (2), and $X = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$.

A WS solution can be obtained from a perfect-foresight version

$$z^{\mathsf{WS}} = \mathbb{E}_\xi\left[\underset{x}{\mathsf{min.}} \ \{F(x,\xi)\}\right] = \mathbb{E}_\xi\left[F(x(\xi),\xi)\right],$$

where $x(\xi) = \mathrm{argmin}_x \{F(x,\xi)\}$.

The expected value of perfect information (EVPI)

The performance of the solution of a 2SSP can be compared against a so-called wait-and-see (WS) solution.

Let our 2SSP be compactly represented as

$$z = \underset{x}{\text{min.}} \ \{\mathbb{E}\left[F(x,\xi)\right]\},$$

where $F(x,\xi) = \left\{c^\top x + \mathbb{E}\left[Q(x,\xi)\right] : x \in X\right\}$, $Q(x,\xi)$ is defined as (2), and $X = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$.

A WS solution can be obtained from a perfect-foresight version

$$z^{\text{WS}} = \mathbb{E}_\xi\left[\underset{x}{\text{min.}} \ \{F(x,\xi)\}\right] = \mathbb{E}_\xi\left[F(x(\xi),\xi)\right],$$

where $x(\xi) = \text{argmin}_x \{F(x,\xi)\}$.

Then, the expected value of perfect information (EVPI) is:

$$EVPI = z - z^{\text{WS}}.$$

Value of stochastic solution (VSS)

We can also compare the solution of a 2SSP against a reference (first-stage) solution.

For that, let $\overline{\xi}$ be a reference scenario (realisation). Then

$$x(\overline{\xi}) = \operatorname*{argmin}_x F(x, \overline{\xi})$$

represents the optimal solution associated with that scenario.

Value of stochastic solution (VSS)

We can also compare the solution of a 2SSP against a reference (first-stage) solution.

For that, let $\overline{\xi}$ be a reference scenario (realisation). Then

$$x(\overline{\xi}) = \operatorname*{argmin}_x F(x, \overline{\xi})$$

represents the optimal solution associated with that scenario.

We can then calculate the performance of $x(\overline{\xi})$ against $\xi \in \Xi$:

$$z^{\mathsf{EV}} = \mathbb{E}_\xi \left[F(x(\overline{\xi}), \xi) \right].$$

If $\overline{\xi} = \mathbb{E}[\xi]$, we have the value of the stochastic solution (VSS):

$$VSS = z^{\mathsf{EV}} - z.$$

EVPI and VSS - general remarks

Some relevant observations:

- With minimisation as a reference, we have that

$$z^{\mathsf{WS}} \leq z \leq z^{\mathsf{EV}} \Rightarrow EVPI \geq 0, VSS \geq 0.$$

- Stronger statements are possible if assumptions on the 2SSP problem structure or the uncertainty terms are made.
- VSS: the higher the better;
- EVPI: the lower the better.

[1]Calculated by fixing the solution $x = (120, 80, 300)$ for each scenario and taking the average of the objective function values

EVPI and VSS - general remarks

Some relevant observations:

- With minimisation as a reference, we have that

$$z^{\mathsf{WS}} \leq z \leq z^{\mathsf{EV}} \Rightarrow EVPI \geq 0, VSS \geq 0.$$

- Stronger statements are possible if assumptions on the 2SSP problem structure or the uncertainty terms are made.
- VSS: the higher the better;
- EVPI: the lower the better.

For the farmer's example:

- $EVPI = -108,390 - \left(\frac{1}{3} \times -167,777 + \frac{1}{3} \times -118,000 + \frac{1}{3} \times -59,950\right) = \$7016$
- $z^{\mathsf{EV}} = -107,240$[1]; $VSS = -107,240 - (-108,390) = \$1150$

[1] Calculated by fixing the solution $x = (120, 80, 300)$ for each scenario and taking the average of the objective function values

Outline of this lecture

Types of recourse problems

It is common to classify 2SSP according to their recourse problem

$$Q(x, \xi) = \left\{ \text{min. } q(\xi)^\top y : W(\xi)y = h(\xi) - T(\xi)x, \ y \geq 0 \right\}.$$

Most common structures:

1. **Fixed recourse:** means that $W(\xi) = W$, $\forall \xi \in \Xi$.

Types of recourse problems

It is common to classify 2SSP according to their recourse problem

$$Q(x,\xi) = \left\{ \text{min. } q(\xi)^\top y : W(\xi)y = h(\xi) - T(\xi)x, \ y \geq 0 \right\}.$$

Most common structures:

1. **Fixed recourse:** means that $W(\xi) = W$, $\forall \xi \in \Xi$.
2. **Simple recourse:** In that case, $W = I$, reducing the recourse feasibility condition to $y = h(\xi) - T(\xi)x$.

   This implies that the recourse becomes

   $$Q(x,\xi) = h(\xi) - T(\xi)x.$$

Types of recourse problems

3. **Complete recourse:** relates to the feasibility of the recourse problem. If the 2SSP has complete recourse, then

$$
\begin{aligned}
&Q(x,\xi) < \infty,\ \forall \xi \in \Xi \iff \\
&\{y : W(\xi)y = h(\xi) - T(\xi)x, y \geq 0\} \neq \emptyset,\ \forall \xi \in \Xi
\end{aligned}
$$

Types of recourse problems

3. **Complete recourse:** relates to the feasibility of the recourse problem. If the 2SSP has complete recourse, then

$$Q(x, \xi) < \infty, \ \forall \xi \in \Xi \iff$$
$$\{y : W(\xi)y = h(\xi) - T(\xi)x, y \geq 0\} \neq \emptyset, \ \forall \xi \in \Xi$$

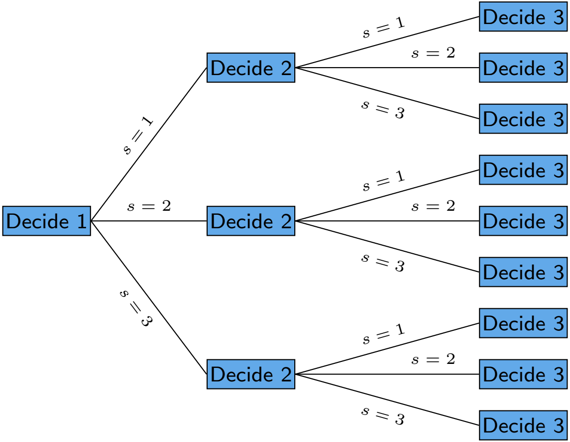
1. **Relatively complete recourse:** in this case, the feasibility of the recourse problem is conditioned on $x \in X$

$$Q(x, \xi) < \infty, \ \forall \xi \in \Xi, x \in X \iff$$
$$\{y : W(\xi)y = h(\xi) - T(\xi)x, y \geq 0\} \neq \emptyset, \ \forall \xi \in \Xi, x \in X.$$

Outline of this lecture

Beyond two decision stages

Many problems have in fact multiple decision points, in which decisions are made sequentially.



Beyond two decision stages

Multi-stage decision problems:

1. Consist of nested two-stage problems. This can be exploited in a dynamic programming fashion;
2. Likewise, presents exponential growth with scenarios per stage ($|\Xi|^{|H|-1}$, with $|\Xi|$-scenario stages $t \in [H]^2$);
3. Trade-off: future flexibility versus computational cost.

[2] $n \in [N] = n \in \{1, \dots, N\}$.

Beyond two decision stages

Multi-stage decision problems:

1. Consist of nested two-stage problems. This can be exploited in a dynamic programming fashion;
2. Likewise, presents exponential growth with scenarios per stage ($|\Xi|^{|H|-1}$, with $|\Xi|$-scenario stages $t \in [H]$[2]);
3. Trade-off: future flexibility versus computational cost.

Consider that we have H decision stages. Our decision process becomes:

$$x^1 \to \xi^2 \to x^2(\xi^2, x^1) \to \xi^3 \to x^3((\xi^2, \xi^3), (x^1, x^2)) \to \dots$$
$$\to \ \xi^H \to x^H((\xi^2, \dots, \xi^H), (x^1, \dots, x^{H-1}))$$

- ξ up to stage $t = 2, \dots, T$ represents a sequence of events
- Hereinafter, $x^t(\xi)$ is a shorthand for $x^t((\xi^2, \dots, \xi^t))$

[2] $n \in [N] = n \in \{1, \dots, N\}$.

Multi-stage decision problems

For $t = H$ we have:

$$\begin{aligned} Q^H(x^{H-1}, \xi^H) = \text{min. } & c^H(\xi)^\top x^H(\xi) \\ \text{s.t.: } & W^H(\xi)x^H(\xi) = h^H(\xi) - T^{H-1}(\xi)x^{H-1} \\ & x^H(\xi) \geq 0. \end{aligned}$$

Multi-stage decision problems

For $t = H$ we have:

$$\begin{aligned} Q^H(x^{H-1}, \xi^H) = \text{min. } & c^H(\xi)^\top x^H(\xi) \\ \text{s.t.: } & W^H(\xi) x^H(\xi) = h^H(\xi) - T^{H-1}(\xi) x^{H-1} \\ & x^H(\xi) \geq 0. \end{aligned}$$

For $t = 2, \dots, H-1$ we have:

$$\begin{aligned} Q^t(x^{t-1}, \xi^t) = \text{min. } & c^t(\xi)^\top x^t(\xi) + \mathcal{Q}^{t+1}(x^t) \\ \text{s.t.: } & W^t(\xi) x^t(\xi) = h^t(\xi) - T^{t-1}(\xi) x^{t-1} \\ & x^t(\xi) \geq 0. \end{aligned}$$

Multi-stage decision problems

For $t = H$ we have:

$$Q^H(x^{H-1}, \xi^H) = \text{min. } c^H(\xi)^\top x^H(\xi)$$
$$\text{s.t.: } W^H(\xi)x^H(\xi) = h^H(\xi) - T^{H-1}(\xi)x^{H-1}$$
$$x^H(\xi) \geq 0.$$

For $t = 2, \dots, H-1$ we have:

$$Q^t(x^{t-1}, \xi^t) = \text{min. } c^t(\xi)^\top x^t(\xi) + \mathcal{Q}^{t+1}(x^t)$$
$$\text{s.t.: } W^t(\xi)x^t(\xi) = h^t(\xi) - T^{t-1}(\xi)x^{t-1}$$
$$x^t(\xi) \geq 0.$$

We want to solve

$$\text{min. } c^{1\top}x^1 + \mathcal{Q}(x^1)$$
$$\text{s.t.: } W^1x^1 = h^1$$
$$x^1 \geq 0.$$

Example: 3SSP

A 3-stage formulation is given as:

$$
\begin{aligned}
\text{min. } & c^{1\top} x^1 + \sum_{\xi_s^2 \in S^2} P(\xi_s^2) \Big[c^2(\xi_s^2)^\top x^2(\xi_s^2) + \\
& \sum_{\xi_s^3 \in S^3(\xi_s^2)} P(\xi_s^3|\xi_s^2) \left(c^3(\xi_s^3|\xi_s^2)^\top x^3(\xi_s^3|\xi_s^2) \right) \Big] \\
\text{s.t.: } & T^1 x^1 = h^1 \\
& T(\xi_s^2) x^1 + W(\xi_s^2) x^2(\xi_s^2) = h(\xi_s^2), \forall \xi_s^2 \\
& T(\xi_s^3|\xi_s^2) x^2(\xi_s^2) + W(\xi_s^3|\xi_s^2) x^3(\xi_s^3|\xi_s^2) = h(\xi_s^3|\xi_s^2), \forall \xi_s^2, \xi_s^3|\xi_s^2 \\
& x^1 \geq 0 \\
& x^2(\xi_s^2) \geq 0, \ \forall \xi_s^2 \\
& x^3(\xi_s^3|\xi_s^2) \geq 0, \ \forall \xi_s^2, \xi_s^3|\xi_s^2.
\end{aligned}
$$

References

Birge, J. R. and Louveaux, F. (2011).
Introduction to stochastic programming.
Springer Science & Business Media.