Multistage stochatic programming

Example and main definitions

Welington de Oliveira

BAS Lecture 13, April 19, 2016, IMPA

Set YouTube resolution to 480p

for best viewing

The flower-girl problem

- Flower-girl sells roses at price p and has to buy them at cost c before she starts selling.
- Flowers left over at the end of the day can be stored and sold the next day, when she starts selling the old roses.
- The roses cannot be carried over more than one additional day at the end of which they are thrown away.
- The demand is random, $h_t(\xi_t)$ denotes demand on the t-th day. The flower-girl wants to maximize her total expected profit.

Horizon is related to the number of days for which the flower-girl continues selling roses without break.

The flower-girl problem

Assume that the flower-girl:

- sells roses only during weekend
- orders an amount x_1 on Friday evening
- registers demand $h_2(\xi_2)$ on Saturday
- stores unsold roses (at no cost)
- possibly buys $x_2(\xi_2)$ new roses.

Denote $s(\xi_2)$ the stock left for the second day, and $w(\xi_2, \xi_3)$ the amount of unsold (wasted) roses at the end of the second day, which is also affected by the demand $h_3(\xi_3)$ on Sunday.

The flower-girl problem

First stage: Friday night

- Variable x_1: amount of flower to be ordered by the flower-girl
- Constraint: $x_1 \geq 0$

Second stage: Saturday

- Demand is revealed: $h_2(\xi_2)$
- Variables $s(\xi_2)$ and $x_2(\xi_2)$: stock and (possibly) new purchase order
- Constraints: $s(\xi_2) \geq 0,\ x_2(\xi_2) \geq 0$ and $x_1 - s(\xi_2) \leq h_2(\xi_2)$

Third stage: Sunday

- Demand is revealed: $h_3(\xi_3)$
- Variables $w(\xi)$: unsold roses will be throw away
- Constraints: $w(\xi_2, \xi_3) \geq 0$ and $x_2(\xi_2) + s(\xi_2) - w(\xi_2, \xi_3) \leq h_3(\xi_3)$

Objective function

The flower-girls wants to maximize her profit:

$$(p - c)\, x_1 + 0\, s(\xi_2) + (p - c)\, x_2(\xi_2) - c\, w(\xi_2, \xi_3)$$

which is

$$(p - c)\, (x_1 + \, x_2(\xi_2)) - c\, w(\xi_2, \xi_3)$$

If the flower demands $h_2(\xi_2)$ and $h_3(\xi_3)$ were known in advance, what would be the optimal solution?

The flower-girl problem

Objective function

The flower-girls wants to maximize her profit:

$$(p-c)\,x_1 + 0\,s(\xi_2) + (p-c)\,x_2(\xi_2) - c\,w(\xi_2,\xi_3)$$

which is

$$(p-c)\,(x_1 + x_2(\xi_2)) - c\,w(\xi_2,\xi_3)$$

If the flower demands $h_2(\xi_2)$ and $h_3(\xi_3)$ were known in advance, what would be the optimal solution?

$$x_1 = h_2(\xi_2) \quad \text{and} \quad x_2(\xi_2) = h_3(\xi_3)$$

yielding the profit: $(p-c)\,(h_2(\xi_2) + h_3(\xi_3))$

But the flower-girl does not know the future demand...

How should she proceed?

She could, for instance, make a decision that is "on average optimal".

The flower-girl problem

Let the random vector be $\xi = (\xi_2, \xi_3)$.

General formulation

$$\begin{cases} \max & \mathbb{E}_\xi \Big[(p-c)\,x_1 + (p-c)\,x_2(\xi_2) - c\,w(\xi) \Big] \\ \text{s.a.} & x_1 \geq 0 \\ & x_1 - s(\xi_2) \leq h_2(\xi_2) \\ & s(\xi_2),\ x_2(\xi_2) \geq 0 \\ & x_2(\xi_2) + s(\xi_2) - w(\xi) \leq h_3(\xi_3) \\ & w(\xi) \geq 0\,. \end{cases}$$

Notice that the variables x_2, s, x_3 and w are functions of ξ.
These variables define a policy.
Variable w is taking into account the temporal dependence!

The flower-girl problem should be more realistically formulated as an integer stochastic program... But let's keep it simple.

The flower-girl problem

Let the random vector be $\xi = (\xi_2, \xi_3)$.

General formulation

$$
\begin{cases}
\max & (p-c)\,x_1 + \mathbb{E}_\xi\Big[(p-c)\,x_2(\xi_2) - c\,w(\xi)\Big] \\
\text{s.a.} & x_1 \geq 0 \\
& x_1 - s(\xi_2) \leq h_2(\xi_2) \\
& s(\xi_2),\ x_2(\xi_2) \geq 0 \\
& x_2(\xi_2) + s(\xi_2) - w(\xi) \leq h_3(\xi_3) \\
& w(\xi) \geq 0\,.
\end{cases}
$$

Notice that the variables x_2, s, x_3 and w are functions of ξ.
These variables define a policy.
Variable w is taking into account the temporal dependence!

The flower-girl problem should be more realistically formulated as an integer stochastic program... But let's keep it simple.

Let the random vector be $\xi = (\xi_2, \xi_3)$.

Nested formulation

$$\max_{x_1 \geq 0} (p-c)\,x_1 + \mathbb{E}_{\xi_2}\left[\max_{\substack{s,\,x_2 \geq 0 \\ x_1 - s \leq h_2(\xi_2)}} (p-c)\,x_2 + \mathbb{E}_{\xi_3|\xi_2}\left[\max_{\substack{w \geq 0 \\ x_2 + s - w \leq h_3(\xi_3)}} -c\,w\right]\right]$$

Is the temporal dependence being taken into account?

Let the random vector be $\xi = (\xi_2, \xi_3)$.

NESTED FORMULATION

$$\max_{x_1 \geq 0} (p-c)\, x_1 + \mathbb{E}_{\xi_2} \left[\max_{\substack{s,\, x_2 \geq 0 \\ x_1 - s \leq h_2(\xi_2)}} (p-c)\, x_2 + \mathbb{E}_{\xi_3|\xi_2} \left[\max_{\substack{w \geq 0 \\ x_2 + s - w \leq h_3(\xi_3)}} -c\, w \right] \right]$$

Is the temporal dependence being taken into account?
Yes, by the Nested structure (conditional expectation).

- $\mathbb{E}_{\xi_3|\xi_2}$ is the expected value with respect to the conditional probability given ξ_2.

The flower-girl problem

Let the random vector be $\xi = (\xi_2, \xi_3)$.

Dynamic formulation

- State 3:

$$Q_3(\,(x_2, s),\ (\xi_2, \xi_3)\,) = \max_{\substack{w \geq 0 \\ x_2+s-w \leq h_3(\xi_3)}} -c\,w$$

- Stage 2:

$$Q_2(x_1,\ \xi_2) = \max_{\substack{s,\, x_2 \geq 0 \\ x_1 - s \leq h_2(\xi_2)}} (p-c)\,x_2 + \mathbb{E}_{\xi_3|\xi_2}[Q_3(\,(x_2, s),\ (\xi_2, \xi_3)\,)]$$

- Stage 1:

$$\max_{x_1 \geq 0} (p-c)\,x_1 + \mathbb{E}_{\xi_2}[Q_2(x_1,\ \xi_2)]$$

The dynamic equations dictates the temporal dependence.

The flower-girl problem

In practical terms, the random variables ξ_2 and ξ_3 need to be approximated by a scenario tree.

- $\xi_2^{\iota 2}$ is a realization of ξ_2, having probability p_{ι_2}, $\iota_2 \in \Omega_2$
- $\xi_3^{\iota 3}$ is a realization of ξ_3. The transfer probability is $p_{\iota_2 \iota_3}$, $\iota_2 \in \Omega_2$ and $\iota_3 \in \Omega_3$

Nested formulation

$$\max_{x_1 \geq 0} (p-c)\, x_1 + \sum_{\iota_2 \in \Omega_2} p_{\iota_2} \left[\max_{\substack{s,\, x_2 \geq 0 \\ x_1 - s \leq h_2(\xi_2^{\iota 2})}} (p-c)\, x_2 + \sum_{\iota_3 \in \Omega_3} p_{\iota_2 \iota_3} \left[\max_{\substack{w \geq 0 \\ x_2 + s - w \leq h_3(\xi_3^{\iota 3})}} -c\, w \right] \right]$$

Stage × Period: the flower-girl problem

The flower-girl problem has:

- Three stages
  - On Friday the flower-girl decides x_1
  - Depending on the demand $h_2(\xi_2)$ on Saturday, she possibly buys more flowers $x_2(\xi_2)$
  - As a result of the demand $h_3(\xi_3)$ on Sunday, she possibly throws flowers away

  The problem is thus a 3-stage stochastic linear programming problem
- Three periods
  - Friday, Saturday and Sunday

Therefore, stage is equal to period. Right or wrong?

Stage × Period: the flower-girl problem

The flower-girl problem has:

- Three stages
  - On Friday the flower-girl decides x_1
  - Depending on the demand $h_2(\xi_2)$ on Saturday, she possibly buys more flowers $x_2(\xi_2)$
  - As a result of the demand $h_3(\xi_3)$ on Sunday, she possibly throws flowers away

  The problem is thus a 3-stage stochastic linear programming problem
- Three periods
  - Friday, Saturday and Sunday

Therefore, stage is equal to period. Right or wrong?
Ans: Wrong!

Stage × Period in stochastic programs

Stages do not necessarily correspond to time periods!

- The main variables are the first-stage decisions (these are the decisions that have to be made before the future is revealed
- Second-stage decisions are allowed to adapt to additional information available at the end of first-stage
- Third-stage decisions are allowed to adapt to additional information available at the end of second-stage
- Mutatis mutandis.

In the flower-girl problem, the time horizon $T = 3$. In two-stage programs the time horizon is $T = 2$. From now on we will assume $T > 2$ the horizon time.

Stages do not necessarily correspond to time periods!

Two-stage multiperiod programs

- There exist two-stage programs with multiperiods.
- In multiperiod two-stage problems decisions at all time instances $t = 1, ..., T$ are made at once, no further information is expected.

Example. Suppose that the flower-girl cannot buy roses (from a supplier) on Saturdays. She needs the decide on Friday the amount of flowers to sell on Saturday and Sunday.

Two-stage multiperiod programs can be dealt with in a similar manner than two-stage programs.

Hence, let's focus on multistage stochastic programs. We'll need some definitions...

Sigma-algebra

Let Ξ be an arbitrary set, and let 2^{Ξ} represents its *power-set*. Then a subset $\mathcal{F} \subset 2^{\Xi}$ is called a σ-algebra if it satisfies the following properties:

1. $\Xi \in \mathcal{F}$
2. $\mathcal{F}$ is closed under complementation:

$$\text{if } A \in \mathcal{F} \text{ then } \Xi \backslash A \in \mathcal{F}$$

3. $\mathcal{F}$ is closed under countable unions:

$$\text{if } A_1, A_2, A_3, \ldots \text{ are in } \mathcal{F} \text{ then } A_1 \cup A_2 \cup A_3 \cup \ldots \text{ belongs to } \mathcal{F}$$

Elements of a σ-algebra are called measurable sets

$(\Xi, \mathcal{F})$ is a measurable space - (or sample space)

Sigma-algebra

Let $\Xi \subset \Re^d$ be a given set, and $\mathcal{F}_t$ be associate σ-algebras, for $t = 1, \ldots, T$.

Definition

The sequence $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_T)$ form a filtration if

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_T \, .$$

Note: elementary elements of $\mathcal{F}_t$ are subsets of Ξ.

Sigma-algebra generated by a function

Let $(\Xi, \mathcal{F})$ and $(\Omega, \mathcal{G})$ be two measurable spaces and $x : \Xi \to \Omega$ be a given function.
The σ-algebra generated by x, denoted by $\sigma(x)$, is the collection of all preimages[1]:

$$\sigma(x) := \{x^{-1}(S) : S \in \mathcal{G}\}.$$

A function $x : \Xi \to \Omega$ is measurable w.r.t. $\mathcal{F}$ iff $\sigma(x) \in \mathcal{F}$.

We'll employ the notation $x \lhd \mathcal{F}$ to denote that the function x is measurable w.r.t $\mathcal{F}$.

[1] Remark: $x^{-1}(S) := \{\xi \in \Xi : x(\xi) \in S\}$

Probability measure

A function $P : \mathcal{F} \to \Re_+$ is called a *measure* on $(\Xi, \mathcal{F})$ if every collection $A_i \in \mathcal{F}$ such that $A_i \cap A_j$ for all $i \neq j$, we have

$$P(\cup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N} A_i} P(A_i)$$

In this definition we're assuming that $P(A_i)$ is finite for all $A_i \in \mathcal{F}$ (and therefore $P(\Xi) < \infty$).

Probability measure

- If $P(\Xi) = 1$, then P is said to be a probability measure
- A measurable space $(\Xi, \mathcal{F})$ equipped with a probability measure P is called a *probability space* $(\Xi, \mathcal{F}, P)$.

Random variable, random vector and stochastic process

- A measurable function ξ from a probability space $(\Xi, \mathcal{F}, P)$ to $\Re$ is called a *random variable*
- A measurable function ξ from a probability space $(\Xi, \mathcal{F}, P)$ to $\Re^m$ is called a *random vector*

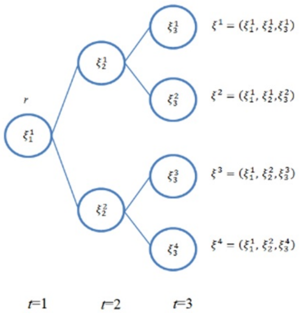
In multistage stochastic programs (T-SP) the uncertain data is revealed gradually over time $t = 1, \ldots, T$

- We view the sequence $\xi_1, \xi_2, \ldots, \xi_T$ as a sequence of random vectors $\xi_t \in \Re^{n_t}$.
- The sequence $\xi = (\xi_1, \ldots, \xi_T)$ is a stochastic process: a sequence of random vectors with a specific probability distribution
- $\xi_{[t]} := (\xi_1, \ldots, \xi_t)$ denotes the history of ξ up to time t
- the stochastic process is stage-wise independent if $P(\xi_t | \xi_{[t-1]}) = P(\xi_t)$
- In T-SP the decision vector x_t, chosen at time t, may depend on $\xi_{[t]}$
  - $x_t = x_t(\xi_{[t]})$ is a stochastic process as well
  - x_t is a function belonging to appropriate functional space

Scenario trees

- Assume that the stochastic process $\xi = (\xi_1, \ldots, \xi_T)$ has a finite number K of realizations
- Each realization (sequence) is called a scenario $\xi^i = (\xi_1^i, \ldots, \xi_T^i)$

It is convenient to depict all the possible K sequences in a form of a *scenario tree*

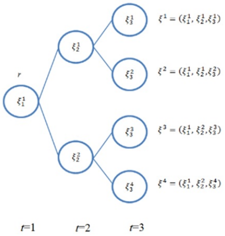


A tree is a directed graph without cycles.

Scenario trees

- Each scenario $\xi^i = (\xi^i_1, \ldots, \xi^i_T)$ has a probability $p_i > 0$ associated
- The value of a given scenario ξ^i at stage t (i.e., ξ^i_t) is denoted a node of tree
- The set of all nodes at stage t is denoted by Ω_t
- Ω_1 contains only the *root node*
- The total number of scenario is $K = |\Omega_T|$
- We sometimes use the short hand ι to denote a node: $\iota \in \Omega_t$
- Each scenario ξ^i has a unique node at stage $t = 1, 2, \ldots, T$
- The ancestor of a node $\iota \in \Omega_t$ is $a(\iota) \in \Omega_{t-1}$ (the root node does have a ancestor)
- The set of descendants (children) of a node $\iota \in \Omega_t$ is denoted by C_ι
- $\Omega_{t+1} = \cup_{\iota \in \Omega_t} C_\iota$, and $C_\iota \cap C_{\iota'} = \emptyset$ if $\iota \neq \iota'$
- Nodes in Ω_T does not have children. Such nodes are called leaf nodes.

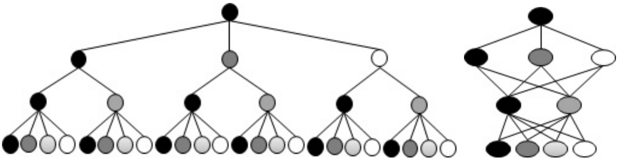
Scenario trees



- $\mathcal{S}^{(\iota)}$ is the set of all scenarios passing through node ι
- The probability of node ι is $p^{(\iota)} := \mathbb{P}[\mathcal{S}^{(\iota)}]$
- Conditional probability $\rho_{a\iota} = \frac{p^{(\iota)}}{p^{(a)}}$ if $a = a(\iota)$
- The probability of reaching a node $\iota \in \Omega_t$ is
  $p^{(\iota)} = \rho_{\iota_1 \iota_2} \rho_{\iota_2 \iota_1} \ldots \rho_{\iota_{t-1} \iota_t} = \mathbb{P}[\mathcal{S}^{(\iota)}]$

Stagewise independence

A stochastic process is stagewise independent if $P(\xi_t|\xi_{[t-1]}) = P(\xi_t)$



Figures by Felipe B. Rodríguez

A stochastic process ξ_i, $t = 1, 2, \ldots, T$, that can take a finite number of different values is a *Markov chain* if

$$\mathbb{P}[\xi_{t+1} = \xi^j_{t+1}|\xi_{[t]} = \xi^i_{[t]}] = \mathbb{P}[\xi_{t+1} = \xi^j_{t+1}|\xi_t = \xi^i_t] = \rho_{ij}$$