Bounds and Approximations in Stochastic Programming

Francesca Maggioni
Dept. of Management, Economics and Quantitative Methods,
University of Bergamo (ITALY)

Department of Electrical, Electronic, and Information Engineering Guglielmo Marconi,
University of Bologna,
May 12 2016

Motivation: why Stochastic Programming?

Stochastic Programming deals with optimization problems involving **uncertain data**.

Applications Fields:

- **Transportation** (uncertain demand, uncertain travel time);
- **Energy markets** (uncertain energy price, uncertain demand);
- **Finance** (uncertainty of investment returns and future liabilities);
- **Engineering** (accident probabilities);
- **Environment** (probability to satisfy regulamentations);
- etc...

Introduction to Stochastic Programming

- Suppose that we face an optimization problem in which some **parameters** are **uncertain**.
- Uncertain parameters depends on a **stochastic variable** ξ, defined on a probability space $(\Xi, \mathscr{A}, P)$. Let $\tilde{\xi}$ a realization of ξ.

Suppose now that some of the constraints are uncertain. The problem becomes

$$
\begin{aligned}
\min_x \quad & c^\top x \\
s.t. \quad & Ax = b \\
& T(\xi)x = h(\xi) \\
& x \geq 0
\end{aligned}
$$

Two-stage Stochastic Programming

We define a method to insert a penalty in the objective function with respect to the realization of the stochastic variable taking into account the error in the constraint.

Let's define $W \in \mathbb{R}^{r \times m}$ called **(fixed) recourse matrix** and $y(\tilde{\xi}) \in \mathbb{R}^m$ called **recourse variable**.

Now let's define the compensation vector:

$$W y(\tilde{\xi}) = h(\tilde{\xi}) - T(\tilde{\xi})x$$

and the **recourse function**:

$$Q(x, \tilde{\xi}) = \min_y \left\{ q(\tilde{\xi})^\top y : W y = h(\tilde{\xi}) - T(\tilde{\xi})x; \quad y \geq 0 \right\}$$

where $q(\tilde{\xi})$ is the cost of the introduction of the recourse vector y.

For every realization $\tilde{\xi}$:

$$Q(x,\tilde{\xi}) = \min_y \left\{ q(\tilde{\xi})^\top y : Wy = h(\tilde{\xi}) - T(\tilde{\xi})x; \quad y \geq 0 \right\}$$

and defining the recourse function $\mathcal{Q}(x) = E_\xi Q(x,\tilde{\xi})$, we get the following formulation

$$\begin{aligned} \min_x \quad & c^\top x + \mathcal{Q}(x) \\ s.t. \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

called **implicit representation** of the stochastic program (deterministic equivalent program).

The set of **decisions** is divided in two groups:

1. A number of decisions have to be taken before the realization $\tilde{\xi}$ of the random variable ξ (**first-stage decision** x)
2. A number of decision can be taken after the realization of the uncertainty (**second-stage decision** y)
   The sequence is: $x \longrightarrow \tilde{\xi} \longrightarrow y(\tilde{\xi}, x)$

Scenario-based Representation of a Two-Stage SP

In order to deal with computations it is useful to consider a finite number S of possible realizations of the future outcomes (**scenarios**) with probability π_k, $k = 1, \ldots, S$. The problem can be reformulated with a **Scenario-Based representation** as follows:

$$\begin{aligned}
\text{min}_x \quad & c^\top x + \sum_{k=1}^{S} \pi_k q_k^\top y_k \\
s.t. \quad & Ax = b \\
& T_1 x + W y_1 = h_1 \\
& T_2 x + W y_2 = h_2 \\
& \vdots \\
& T_S x + W y_S = h_K \\
& x \geq 0,\ y_k \geq 0, k = 1, \ldots, S
\end{aligned}$$

where $q_k, T_k, h_k,\ k = 1, \ldots, S$ are coefficients depending by scenario $\tilde{\xi}_k$ and $\sum_k \pi_k = 1$.

Mulststage Stochastic Programming

- Let us consider a **random process**

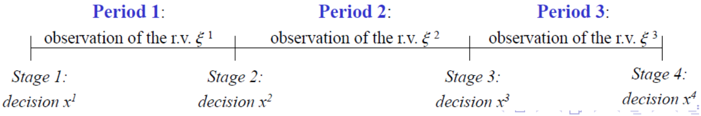
$$\boldsymbol{\xi}^{H-1} = (\xi^1, \xi^2, \ldots, \xi^{H-1}) \ ,$$

where ξ^t is defined on a **probability space** $(\Xi^t, \mathscr{A}^t, P)$, such that:
  - $\Xi = \Pi_t \Xi^t$ is the support;
  - P is a given probability distribution;
  - $\mathscr{A}^t$ (with $\mathscr{A}^{t-1} \subseteq \mathscr{A}^t$) is a sigma-algebra describing the evolving information set generated by the process over time;
- $E_{\boldsymbol{\xi}^{H-1}}$: **mathematical expectation** with respect to $\boldsymbol{\xi}^{H-1}$;
- $x = (x^1, x^2, \ldots, x^H)$: **decision vector** with decision x^t at time t is $\mathscr{A}^{t-1}-$ measurable and depending by the history up to t:

$$\begin{aligned} \text{decision}(x^1) \quad & \rightarrow \text{observation}(\xi^1) \rightarrow \text{decision}(x^2) \rightarrow \text{observation}(\xi^2) \rightarrow \ldots \\ \ldots \quad & \rightarrow \text{decision}(x^{t-1}) \rightarrow \text{observation}(\xi^{t-1}) \rightarrow \text{decision}(x^t). \end{aligned}$$

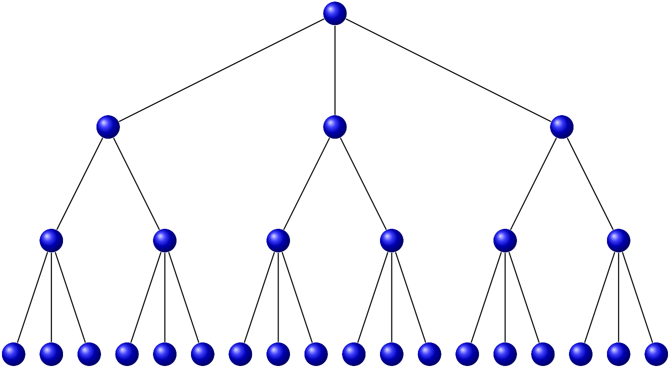
Multistage Linear Stochastic Program ($MLSP$)

$$
\begin{aligned}
RP \quad &:= \quad \min_{\mathbf{x}} E_{\boldsymbol{\xi}^{H-1}} z(\mathbf{x}, \boldsymbol{\xi}^{H-1}) \\
&:= \quad \min_{x^1} c^1 x^1 + E_{\xi^1}\Bigg[\min_{x^2} c^2 x^2(\boldsymbol{\xi}^1) + E_{\xi^2}\Bigg[\cdots + E_{\xi^{H-1}}\Bigg[\min_{x^H} c^H x^H\left(\boldsymbol{\xi}^{H-1}\right)\Bigg]\Bigg]\Bigg] \\
& \qquad A x^1 = h^1 \ , \\
& \qquad T^1(\boldsymbol{\xi}^1) x^1 + W^2(\boldsymbol{\xi}^1) x^2(\boldsymbol{\xi}^1) = h^2(\boldsymbol{\xi}^1) \ , \\
& \qquad \vdots \\
& \qquad T^{H-1}(\boldsymbol{\xi}^{H-1}) x^{H-1}(\boldsymbol{\xi}^{H-1}) + W^H(\boldsymbol{\xi}^{H-1}) x^H(\boldsymbol{\xi}^{H-1}) = h^H(\boldsymbol{\xi}^{H-1}) \ , \\
& \qquad x^1 \geq 0 \ ; \quad x^t(\boldsymbol{\xi}^{t-1}) \geq 0 \ , \ t = 2, \ldots, H;
\end{aligned}
$$



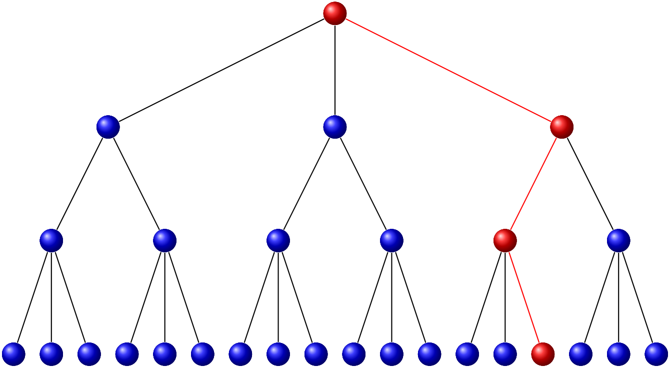
Scenario Tree Dictionary

- $\xi_1, \ldots, \xi_S$: possible realizations (or **scenarios**) of future outcomes $\boldsymbol{\xi}^{H-1}$;
- $\pi(\xi_k) = \pi_k :=$ **probability** of scenario $k = 1, \ldots, S$;



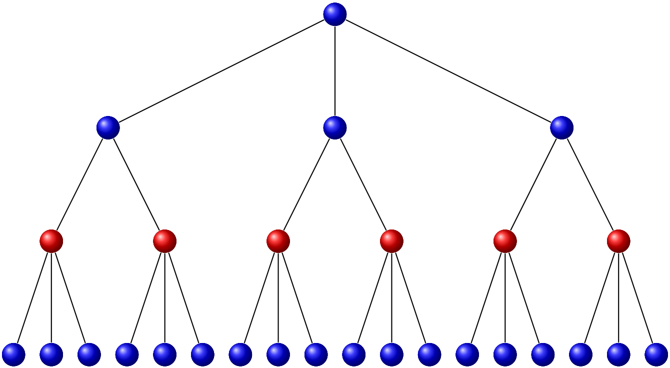
Scenario Tree Dictionary

- $\xi_1, \ldots, \xi_S$: possible realizations (or **scenarios**) of future outcomes $\boldsymbol{\xi}^{H-1}$;
- $\pi(\xi_k) = \pi_k :=$ **probability** of scenario $k = 1, \ldots, S$;



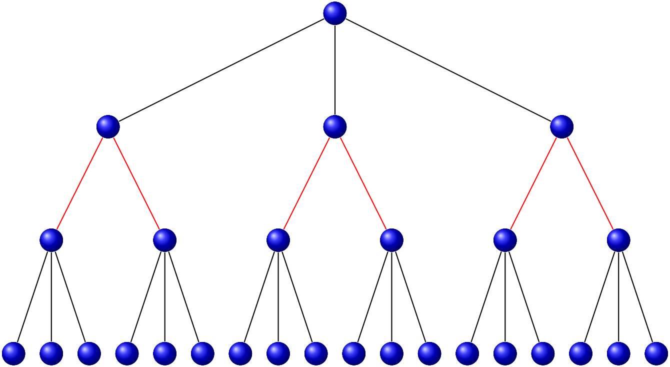
Scenario Tree Dictionary

Stage: decision time



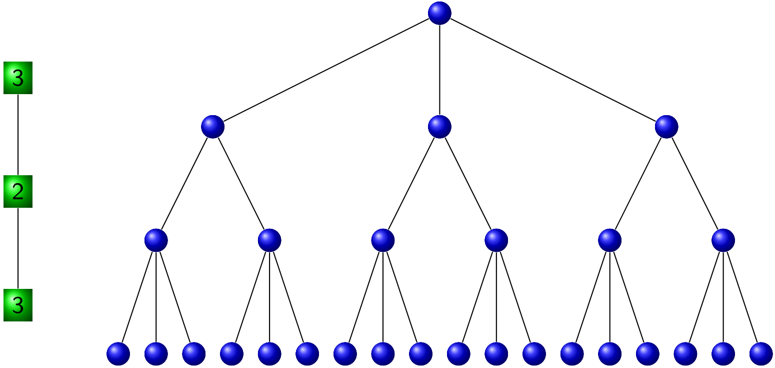
Scenario Tree Dictionary

Period: realization of the uncertainty



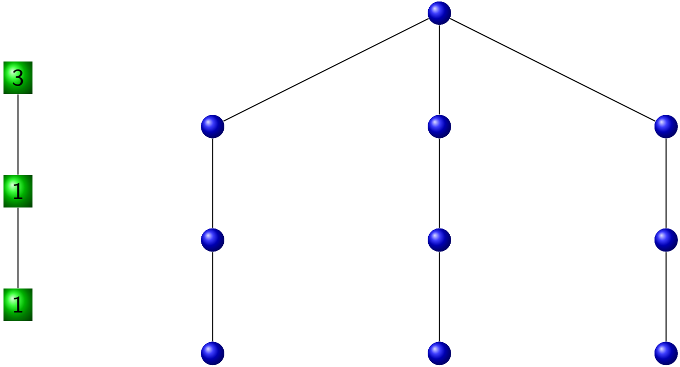
Scenario Tree Dictionary

Branching: number of child for each node in each period, 3-2-3



Scenario Tree Dictionary

If we have **no information** after the second stage we have a branching 3-1-1



Scenario Formulation of a MLSP

$$
\begin{aligned}
RP \quad := \quad & \min_{x} E_{\boldsymbol{\xi}^{H-1}} z(x, \boldsymbol{\xi}^{H-1}) \\
= \quad & \min_{x^1,\dots,x^H} c^1 x^1 + \sum_{k=1}^{S} \pi_k \left(c^2 x^2(\xi_k) + \cdots + c^H x^H(\xi_k) \right) \\
& \text{s.t. } A x^1 = h^1 \ , \\
& \qquad T^1(\xi_k^1) x^1(\xi_k) + W^2(\xi_k^1) x^2(\xi_k) = h^2(\xi_k^1) \ , \\
& \qquad \vdots \\
& \qquad T^{H-1}(\xi_k^{H-1}) x^{H-1}(\xi_k) + W^H(\xi_k^{H-1}) x^H(\xi_k) = h^H(\xi_k^{H-1}) \ , \\
& \quad x^1(\xi_k) \geq 0 \ ; \quad x^t(\xi_k) \geq 0 \ , \ t = 2, \dots, H \ , \quad k = 1, \dots, S \ , \\
& \qquad x^t(\xi_{j'}) = x^t(\xi_{j''}), \forall j', j'' \text{ for which } \xi_{j'}^t = \xi_{j''}^t \quad t = 2, \dots, H \ ;
\end{aligned}
$$

where have been explicity added the **non-anticipativity constraints**.

Data Sources and Scenario Generation Features

Three sources of data or combinations:

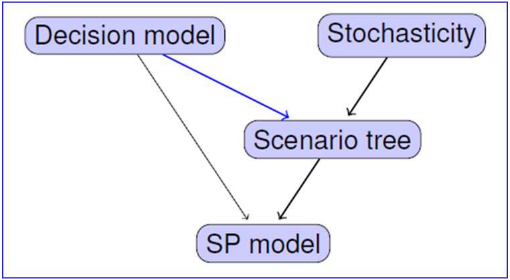
- Historical Data
- Simulation
- Expert Opinion

A *good* scenario generation method should:

- Be the most **automatic** as possible
- **Influence** the solution the least possible
- Let the solution **converge** to the true optima as the number of scenario increases
- Be as possible good for a given number of scenarios

Scenario Generation

- Choice of a good scenario generation method is **problem-dependent**
- Scenario generation is a part of **modelling process**



Why Bounds and Approximations in SP?

- Evaluation if it is worth the **additional computations** for the stochastic program versus simplified approaches.
- Deeper understanding of **expected value solution** and relation to the stochastic one.
- Useful because it could help to predict how the stochastic model will perform when is **not solvable**.
  - If the gap between lower and upper bounds for the objective value is acceptably small, one may even stop and never fully optimize.

The general idea behind construction of bounds we adopt, is that for every optimization problem of minimization type:

- **lower bound** to the optimal solution can be found by **relaxation of constraints**;
- **upper bound** to the optimal solution can be found by **inserting feasible solutions**.

Why Bounds and Approximations in SP?

- Evaluation if it is worth the **additional computations** for the stochastic program versus simplified approaches.
- Deeper understanding of **expected value solution** and relation to the stochastic one.
- Useful because it could help to predict how the stochastic model will perform when is **not solvable**.
  - If the gap between lower and upper bounds for the objective value is acceptably small, one may even stop and never fully optimize.

The general idea behind construction of bounds we adopt, is that for every optimization problem of minimization type:

- **lower bound** to the optimal solution can be found by **relaxation of constraints**;
- **upper bound** to the optimal solution can be found by **inserting feasible solutions**.

Information - Quality

We introduce three classes of measures:

1. Measures by **changing the probability measure** in MSP;
2. Measures by **inserting feasible solutions** in MSP
3. Measures of the **Quality** of EV Solution in MSP;

QUESTIONS

1. What is the value of **different level of information** in multistage SP?
2. Does the stochastic optimal solution **inherit** properties from the deterministic solution or are they totally different?

Information - Quality

We introduce three classes of measures:

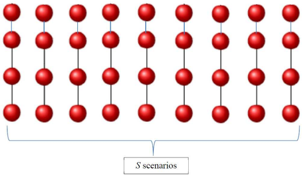
1. Measures by **changing the probability measure** in MSP;
2. Measures by **inserting feasible solutions** in MSP
3. Measures of the **Quality** of EV Solution in MSP;

QUESTIONS

1. What is the value of **different level of information** in multistage SP?
2. Does the stochastic optimal solution **inherit** properties from the deterministic solution or are they totally different?

Multistage Wait-and-see Problem (WS)

$$
\begin{aligned}
WS \quad := \quad & E_{\boldsymbol{\xi}^{H-1}} \min_{x^1(\boldsymbol{\xi}^{H-1}),\ldots,x^H(\boldsymbol{\xi}^{H-1})} c^1 x^1(\boldsymbol{\xi}^{H-1}) + \ldots + c^H x^H(\boldsymbol{\xi}^{H-1}) \\
& \text{s.t. } Ax^1 = h^1 \;, \\
& T^1(\xi^1)x^1(\boldsymbol{\xi}^{H-1}) + W^2(\xi^1)x^2(\boldsymbol{\xi}^{H-1}) = h^2(\xi^1) \;, \\
& \vdots \\
& T^{H-1}(\boldsymbol{\xi}^{H-1})x^{H-1}(\boldsymbol{\xi}^{H-1}) + W^H(\boldsymbol{\xi}^{H-1})x^H(\boldsymbol{\xi}^{H-1}) = h^H(\boldsymbol{\xi}^{H-1}) \;, \\
& x^1 \geq 0 \;; \quad x^t(\boldsymbol{\xi}^{H-1}) \geq 0 \;, \; t = 2, \ldots, H \;.
\end{aligned}
$$



The Expected Value Problem EV

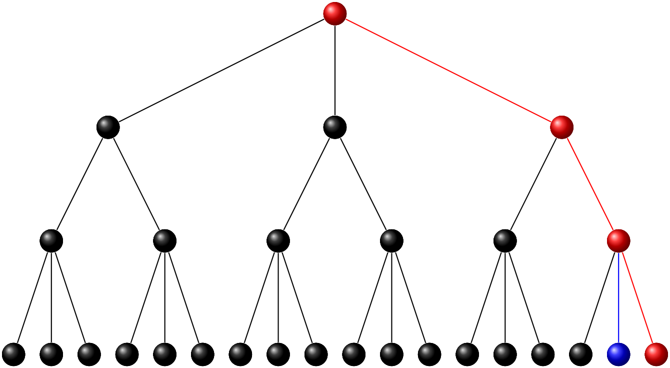
The **Expected Value problem** EV (with $\bar{\boldsymbol{\xi}} = (E\xi^1, E\xi^2, \ldots, E\xi^{H-1})$):

$$
\begin{aligned}
EV \quad := \quad & \min_{\mathbf{x}} z(x, \bar{\boldsymbol{\xi}}) \\
:= \quad & \min_{x^1,\ldots,x^H} c^1 x^1 + \cdots + c^H x^H \\
& \quad \text{s.t.} \quad A x^1 = h^1 \;, \\
& \qquad\quad T^1(\bar{\xi}^1) x^1 + W^2(\bar{\xi}^1) x^2 = h^2(\bar{\xi}^1) \;, \\
& \qquad\quad \vdots \\
& \qquad\quad T^{H-1}(\bar{\xi}^{H-1}) x^{H-1} + W^H(\bar{\xi}^{H-1}) x^H = h^H(\bar{\xi}^{H-1}) \;, \\
& \qquad\quad x^t \geq 0 \;, t = 1, \ldots, H.
\end{aligned}
$$

Theorem $\quad WS \leq RP \leq EEV \;,\quad$ **Madansky (1960)**

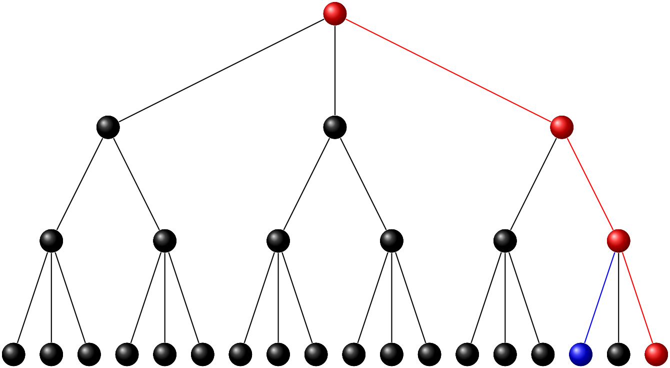
EEV: the solution of RP model, having $1^{st}-$stage variables fixed from EV.

Mulststage Pair Subproblem



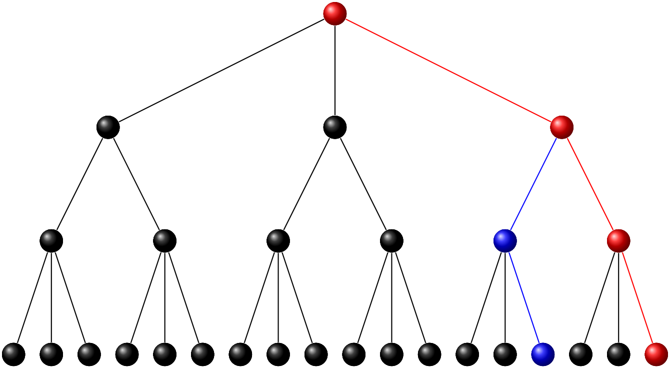
- Red: reference scenario (fixed) in the pair subproblem
- Blu: free scenario in the pair subproblem

Mulststage Pair Subproblem



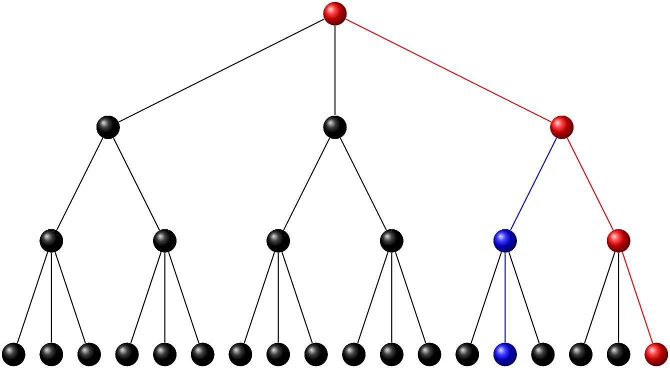
- Red: reference scenario (fixed) in the pair subproblem
- Blu: free scenario in the pair subproblem

Mulststage Pair Subproblem



- Red: reference scenario (fixed) in the pair subproblem
- Blu: free scenario in the pair subproblem

Mulststage Pair Subproblem



- Red: reference scenario (fixed) in the pair subproblem
- Blu: free scenario in the pair subproblem

We solve the following **Pair Subproblems** defined as:

$$\min z^P(x, \xi_r, \xi_k) := \min \left(c^1 x^1 + \sum_{t=2}^{\hat{H}-1} c^t x_r^t(\xi_r) + \sum_{t=\hat{H}}^{H} \left[\pi_r c^t x_r^t(\xi_r) + (1-\pi_r) c^t x_k^t(\xi_k) \right] \right)$$

$$\begin{aligned} \text{s.t.} \qquad & Ax^1 = h^1 \ , \\ & T_r^{t-1} x_r^{t-1}(\xi_r) + W_r^t x_r^t(\xi_r) = h_r^t(\xi_r) \ , \\ & T_k^{t-1} x_k^{t-1}(\xi_k) + W_k^t x_k^t(\xi_k) = h_k^t(\xi_k) \ , \\ & x^1 \geq 0 \ ; \quad x_r^t(\xi_r) \geq 0 \ , \quad x_k^t(\xi_k) \geq 0 \ , \quad t = 2, \ldots, H. \end{aligned}$$

The **Multistage Sum of Pairs Expected Values ($MSPEV$)** is:

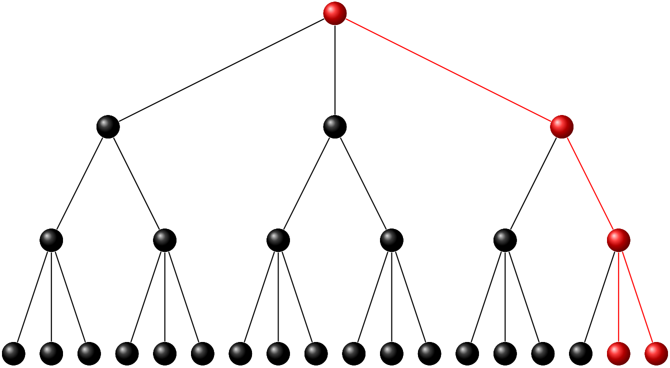
$$MSPEV := \frac{1}{1-\pi_r} \sum_{k=1, k \neq r}^{S} \pi_k \min z^P(x, \xi_r, \xi_k)$$

Proposition (Maggioni F., Allevi E. & Bertocchi M. (JOTA, 2014))

- *If $\xi_r \notin \Xi$, then $MSPEV = WS$.*
- *If $\xi_r \in \Xi$ then $WS \leq MSPEV \leq RP$.*

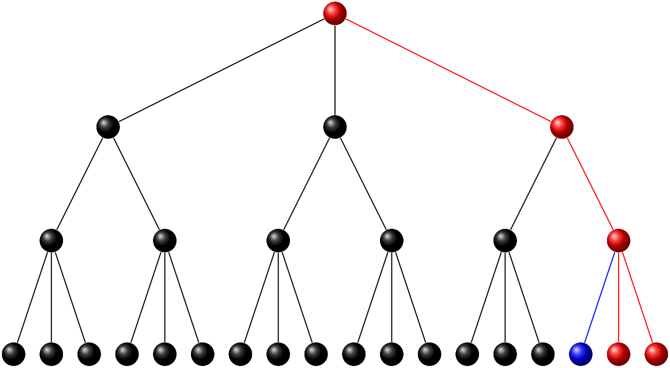
Multistage Group Subproblem

- Let $\mathscr{R} = \{1, \ldots, R\}$ be the index set of reference scenarios $(1 \leq R < S)$.
- Let $\mathscr{K} = \{R+1, \ldots, S\}$ be the index set of free scenarios.



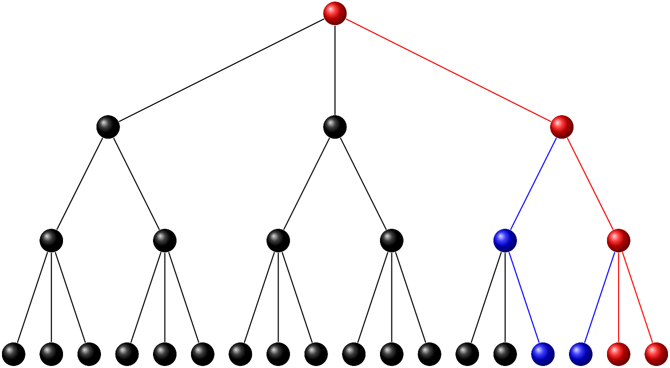
Multistage Group Subproblem

- Let $\mathscr{R} = \{1, \dots, R\}$ be the index set of reference scenarios $(1 \leq R < S)$.
- Let $\mathscr{K} = \{R+1, \dots, S\}$ be the index set of free scenarios.



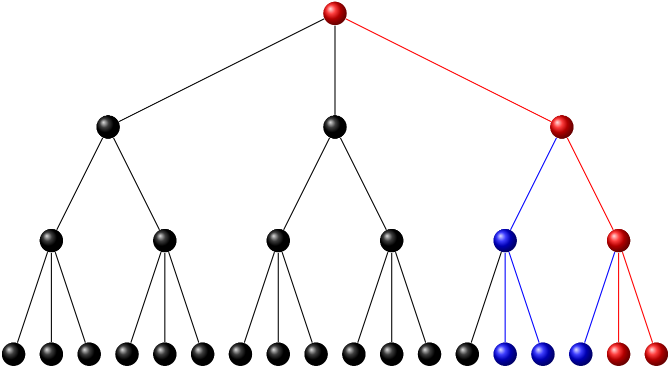
Multistage Group Subproblem

- Let $\mathscr{R} = \{1, \ldots, R\}$ be the index set of reference scenarios ($1 \leq R < S$).
- Let $\mathscr{K} = \{R+1, \ldots, S\}$ be the index set of free scenarios.



Multistage Group Subproblem

- Let $\mathscr{R} = \{1, \ldots, R\}$ be the index set of reference scenarios $(1 \leq R < S)$.
- Let $\mathscr{K} = \{R+1, \ldots, S\}$ be the index set of free scenarios.



Multistage Group Subproblem

The **Multistage Group Subproblem** $MGR(\Psi_k, R)$ is defined as follows:

$$
\begin{aligned}
\min_{x^1, x^2, \ldots, x^H} \quad & \left(c^1 x^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t x^t(\xi_r) \right) + \left(1 - \sum_{r=1}^{R} \pi_r\right) \sum_{i \in \Psi_k} \frac{\pi_i}{\pi(\Psi_k)} \sum_{t=2}^{H} c^t x^t(\xi_i) \right) \\
\text{s.t.} \quad & A x^1 = h^1 \ , \\
& T^{t-1}(\xi_r^{t-1}) x^{t-1}(\xi_r) + W^t(\xi_r^{t-1}) x^t(\xi_r) = h^t(\xi_r^{t-1}) \ , \ r \in \mathscr{R} \\
& T^{t-1}(\xi_i^{t-1}) x^{t-1}(\xi_i) + W^t(\xi_i^{t-1}) x^t(\xi_i) = h^t(\xi_i^{t-1}) \ , \ i \in \Psi_k \\
& x^1 \geq 0 \ ; \quad x^t(\xi_r) \geq 0, r \in \mathscr{R} \quad x^t(\xi_i) \geq 0, i \in \Psi_k \\
& x^t(\xi_{j'}) = x^t(\xi_{j''}), \forall j', j'' \in \mathscr{S} \text{ for which } \xi_{j'}^t = \xi_{j''}^t \quad t = 2, \ldots, H.
\end{aligned}
$$

where $\pi(\Psi_k) = \sum_{i \in \Psi_k} \pi_i$ is the probability assigned to every scenarios group Ψ_k.

Multistage Expected value of the Group Sub-problem

The *Multistage Expected value of the Group Sub-problem Objective* functions with k scenarios in each group and R fixed scenarios, is:

$$MEGSO(k,R) := \frac{1}{\binom{K-1}{k-1}\left(1-\sum_{r=1}^{R}\pi_r\right)}\left[\sum_{\Psi_k\in\mathcal{P}_k(\mathcal{K})}\pi(\Psi_k)\min z^R(\Psi_k)\right].$$

$MEGSO(k,R)$ is monotonically nondecreasing in k with R fixed and in R with k fixed.

Theorem (Maggioni F., Allevi E. & Bertocchi M. CMS (2016))

- *For any chosen fixed R, $1 \leq R < S$, the following chain of inequalities hold true:*

$$WS \leq MEGSO(1,R) \leq MEGSO(2,R) \leq \ldots \leq MEGSO(K,R) = RP\ .$$

- *Given an integer k, $1 \leq k \leq K$, the following chain of inequalities holds true:*

$$MEGSO(k,1) \leq MEGSO(k,2) \leq \cdots \leq MEGSO(k,S-k) = RP\ .$$

Complexity considerations

Theorem (Dyer, Stougie (2006))

- **Two-stage stochastic programming** *with discrete distributions on the parameters is* $\#P - hard$.
- **Multi-stage stochastic programming** *with discretely decision-dependent distributed parameters is* $PSPACE - hard$.

We consider the **worst case complexity** $c(n)$ of a tree with n nodes.

RP: Tree with height H and branching factor b_t:

- the number of **scenarios** is $b_1 \cdot b_2 \cdot \ldots \cdot b_{H-1}$;
- the number of **nodes** is $\sum_{t=1}^{H-1} \prod_{\tau=1}^{t} b_\tau + 1$.

$MGR(\Psi_k, R)$:

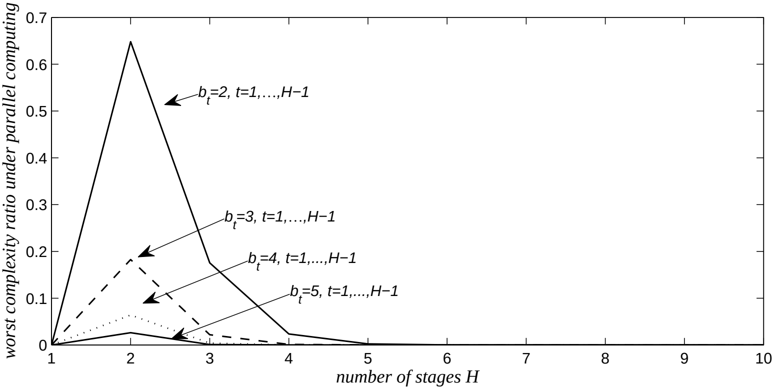
- the maximum number of **nodes** is $(R+k)(H-1)+1$;
- the number of **scenarios** is $k+R$.

Using:

$$c(n) = O(L \cdot n^3 / log(n))$$ **Anstreicher (1999)**

The **ratio between complexities** is

$$\frac{\kappa(MEGSO(k,R))}{\kappa(RP)} = \frac{\kappa((R+k)(H-1)+1)}{\kappa(\sum_{t=1}^{H-1}\prod_{\tau=1}^{t} b_\tau + 1)} \,.$$



Suppose that the stochastic program $Opt(\mathbb{P})$ is of the form:

$$Opt(\mathbb{P}): \quad v^*(\mathbb{P}) = \min\{\mathcal{R}_{\mathbb{P}}[z(x,\xi)] \,:\, x \in \mathbb{X}, x \lhd \mathscr{A}; \mathbb{P} \sim (\Xi, \mathscr{A}, P, \xi)\}$$

where $\mathcal{R}$ is a **risk functional** (like Expectation, or the Average Value-at-Risk).

Given the optimal value mapping

$$(\Xi, \mathscr{A}, P, \xi) \sim \mathbb{P} \longmapsto z^*(\mathbb{P}) = z^*(\Xi, \mathscr{A}, P, \xi)$$

we keep the **filtration and the process ξ fixed** and consider only the mapping

$$P \mapsto z^*(P).$$

One main structural property of some stochastic programs (including expectation type) is that this mapping is **concave**.
A **chain of bounds** can be derived as consequence of this property.

[Maggioni F. & Pflug G. SIOPT (2016)]

Upper bounds by inserting feasible solutions

The following upper bounds are introduced:

- The **Multistage Expected Value of R-reference Scenarios** is defined as

$$MEVRS^{1,R} := E_{\boldsymbol{\xi}^{H-1}} \min_{\mathbf{x}^{(2,H)}} z(\check{\mathbf{x}}^1_R, \mathbf{x}^{(2,H)}, \boldsymbol{\xi}^{H-1}) \ ,$$

  where $\check{\mathbf{x}}^1_R$ is the optimal first stage solution of the stochastic problem $\min_{\mathbf{x}} z(\mathbf{x}, \xi_1, \dots, \xi_R)$.

- The **Multistage Expectation of Group Subproblems** is defined as

$$MEGS(k,R) := \min_{\Psi_k \in \mathcal{P}_k(\mathscr{K}) \cup \mathscr{R}} (E_{\boldsymbol{\xi}^{H-1}} \min_{\mathbf{x}^{(2,H)}} z(\hat{x}^1_{\Psi_k,R}, \mathbf{x}^{(2,H)}, \boldsymbol{\xi}^{H-1})) \ ,$$

  where $\hat{x}^1_{\Psi_k,R}$ is the optimal first stage solution of $MGR(\Psi_k, R)$.

Proposition

For a fixed number R of reference scenarios and any $1 \leq k \leq K$ we have

$$RP \leq MEGS(k,R) \leq MEVRS^{1,R} \ .$$

Using MEGSO and MEGS in multistage mixed integer stochastic programming

Algorithm 1 Using $MEGSO(k,R)$ and $MEGS(k,R)$

Require: S, $R < S$, $K = S - R$, $\bar{\epsilon}$, $\bar{\gamma}$, $out\_of\_memory = False$

1: $k = 1$
2: $\epsilon = MEGS(k,R) - MEGSO(k,R)$
3: **while** $k < K \wedge CPU(MEGSO(k,R)) < \bar{\gamma} \wedge CPU(MEGS(k,R)) < \bar{\gamma} \wedge \epsilon \geq \bar{\epsilon} \wedge out\_of\_memory = False$ **do**
4: $\quad k = k + 1$
5: $\quad \epsilon = MEGS(k,R) - MEGSO(k,R)$
6: **end while**
7: **return** ϵ, $MEGS(k,R)$, $MEGSO(k,R)$

Measures of Quality of Deterministic Solution in SP

- The **classical evaluation of the expected value solution** $\bar{x}(\bar{\xi})$. We calculate $EEV =$ and compare it with RP using

$$VSS = EEV - RP$$

- Fix at zero all first stage variables which are at zero in the expected value solution and then solve the stochastic program (**skeleton**).
- Fix at zero only the first stage variables with **high reduced costs** in the expected value solution.

Measures of Quality of Deterministic Solution in SP

- The **classical evaluation of the expected value solution** $\bar{x}(\bar{\xi})$. We calculate $EEV =$ and compare it with RP using

$$VSS = EEV - RP$$

- Fix at zero all first stage variables which are at zero in the expected value solution and then solve the stochastic program (**skeleton**).
- Fix at zero only the first stage variables with **high reduced costs** in the expected value solution.

Measures of Quality of Deterministic Solution in SP

- The **classical evaluation of the expected value solution** $\bar{x}(\bar{\xi})$. We calculate $EEV =$ and compare it with RP using

$$VSS = EEV - RP$$

- Fix at zero all first stage variables which are at zero in the expected value solution and then solve the stochastic program (**skeleton**).
- Fix at zero only the first stage variables with **high reduced costs** in the expected value solution.

Measures of badness/goodness of deterministic solutions

- Let $\hat{x}$ be the solution of:

$$\begin{array}{ll} \min_{x \in \mathbb{X}} & E_{\boldsymbol{\xi}} z(x, \boldsymbol{\xi}) \\ \text{s.t.} & x_j = \bar{x}_j(\bar{\xi}), \quad j \in \mathcal{J}. \end{array}$$

where $\mathcal{J} = \{j \mid \bar{x}_j(\bar{\xi}) = 0\}$.

We compute:

$$ESSV = E_{\boldsymbol{\xi}}\left(z\left(\hat{x}, \boldsymbol{\xi}\right)\right)$$ **expected skeleton solution value**

and we compare it with RP by means of:

$$VSS \geq LUSS = ESSV - RP$$ **loss using skeleton solution**

[*Maggioni & Wallace (2012) AOR*]

- Let $\mathscr{R} = \{r_1, \ldots, r_j, \ldots, r_J\}$ be the set of **reduced costs** of the components $\bar{x}_j(\bar{\xi}),\ j \in \mathcal{J}$ of the expected value solution $\bar{x}(\bar{\xi})$ at zero or at their lower bound.

$$\mathscr{R}_p = \left\{ r_j : r^{min} + (p-1)\frac{(r^{max} - r^{min})}{N} \leq r_j \leq r^{min} + (p)\frac{(r^{max} - r^{min})}{N} \right\} .$$

Then let $\tilde{x}_p$ be the solution of:

$$\begin{array}{lll} \min_{x \in \mathbb{X}} & E_{\boldsymbol{\xi}} z\left(x, \boldsymbol{\xi}\right) & \\ \text{s.t.} & x_j = \bar{x}_j(\bar{\xi}), & j \in \mathcal{J}_p, \ldots, \mathcal{J}_N \ , \mathcal{J}_p = \left\{ j \mid r_j(\bar{x}_j(\bar{\xi})) \in \mathscr{R}_p \right\} \end{array}$$

The **generalized expected skeleton solution value**

$$GESSV(p, N) = E_{\boldsymbol{\xi}}\left(z\left(\tilde{x}_p, \boldsymbol{\xi}\right)\right) \ , \quad p = 1, \ldots, N$$

and **loss using p-generalized skeleton solution** is:

$$GLUSS(p, N) = GESSV(p, N) - RP \ , \quad p = 1, \ldots, N \ .$$

[Maggioni, Crainic, Perboli & Rei, CIRRELT-2015-21]

GLUSS properties

Proposition

For a fixed $N \in \mathbb{N} \setminus \{0, 1\}$

$$GLUSS(p, N) \geq GLUSS(p+1, N)\ , \quad p = 1, \ldots, N-1.$$

Proposition

For a given $N \in \mathbb{N} \setminus \{0\}$ *and a fixed* $p \in \mathbb{N} \setminus \{0\}$ *such that* $p = 1, \ldots, N$,

$$GLUSS(p, N+1) \geq GLUSS(p, N).$$

Proposition

For given $N_1, N_2 \in \mathbb{N} \setminus \{0\}$ *and* $p_1, p_2 \in \mathbb{N} \setminus \{0\}$, *with* $p_1 = 1, \ldots, N_1$, $p_2 = 1, \ldots, N_2$ *and such that* $\frac{p_1}{N_1} \leq \frac{p_2}{N_2}$

$$GLUSS(p_1, N_1) \geq GLUSS(p_2, N_2).$$

The problem

Multi-period supply planning problem: optimize vehicle-renting and transportation activities to satisfy demand in several destinations out of several origins in the presence of uncertainty on **demands** and the **cost of extra vehicles**.
Discounts for vehicles rented but not actually used are applied.

Sets:

$$\mathscr{K} = \{k : k = 1, \ldots, K\}$$, set of **suppliers**;

$$\mathscr{O}_k = \{i : i = 1, \ldots, O_k\}$$, set of **plant locations** of supplier $k \in \mathscr{K}$;

$$\mathscr{D} = \{j : j = 1, \ldots, D\}$$, set of **destination plants** (belonging to same company).

Deterministic Parameters:

- t_{ijk} , unit **transportation cost** from supplier $i \in \mathscr{O}_k, k \in \mathscr{K}$ to plant $j \in \mathscr{J}$;
- $\bar{b}_j$, average **buying cost** from an external source for plant $j \in \mathscr{J}$;
- q , vehicle **capacity**;
- g_j , **unloading capacity** at the customer $j \in \mathscr{D}$;
- v_k , **maximum requirement capacity** of supplier $k \in \mathscr{K}$;
- r_k , **minimum requirement capacity** of supplier $k \in \mathscr{K}$;
- l_j^0 , **initial inventory** of product at customer $j \in \mathscr{D}$;
- α , **discount**;

Set:

$$\mathscr{S} = \{s : s = 1, \ldots, S\} \quad , \quad \text{set of } \textbf{scenarios}$$

Stochastic Parameters:

d_j^s , **demand** of destination plant j at scenario $s \in \mathscr{S}$;
b_j^s , **buying cost** from external sources for destination plant j at scenario $s \in \mathscr{S}$;
p^s , **probability** of scenario $s \in \mathscr{S}$.

Variables:

- **First stage decision variables**

$$x_{ijk} \in \mathbb{N} \quad , \quad \text{number of } \textbf{vehicles booked}$$

- **Second stage decision variables**

$z_{ijk}^s \in \mathbb{N}$, number of **vehicles actually used**
$y_j^s \in \mathbb{R}^+$, **volume of product to purchase** from an external source

In the two-stage case, we get the following **mixed-integer stochastic programming model with recourse**:

$$\min_{(x_{ijk}),(y_j^s),(z_{ijk}^s)} q\sum_{k=1}^{K}\sum_{i=1}^{O_k}\sum_{j=1}^{D} t_{ijk}x_{ijk}+\sum_{s=1}^{S} p^s\left[\sum_{j=1}^{D} qb_j^s\, y_j^s-\alpha q\sum_{k=1}^{K}\sum_{i=1}^{O_k}\sum_{j=1}^{D} t_{ijk}\left(x_{ijk}-z_{ijk}^s\right)\right]$$

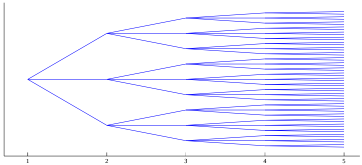
$$\begin{array}{rcl}
s.t. \quad q\sum_{k=1}^{K}\sum_{i=1}^{O_k} x_{ijk} & \leq & g_j\,, \quad j\in\mathscr{D}, \\
l_j^0+q\left(\sum_{k=1}^{K}\sum_{i=1}^{O_k} z_{ijk}^s+y_j^s\right)-d_j^s & \geq & 0\,, \quad j\in\mathscr{D}, \quad s\in\mathscr{S}, \\
z_{ijk}^s & \leq & x_{ijk}\,,\; i\in\mathscr{O}_k,\; k\in\mathscr{K},\; j\in\mathscr{D},\; s\in\mathscr{S}, \\
r_k\leq q\sum_{i\in O_k}\sum_{j=1}^{D} z_{ijk}^s & \leq & v_k\,, \quad k\in\mathscr{K}, \quad s\in\mathscr{S}, \\
x_{ijk} & \in & \mathbb{N}\,,\; i\in\mathscr{O}_k,\; k\in\mathscr{K},\; j\in\mathscr{D}, \\
y_j^s & \in & \mathbb{R}^+\,,\; j\in\mathscr{D},\; s\in\mathscr{S}, \\
z_{ijk}^s & \in & \mathbb{N}\,,\; i\in\mathscr{O}_k,\; k\in\mathscr{K},\; j\in\mathscr{D},\; s\in\mathscr{S},
\end{array}$$

In the multi-stage case the model becomes:

$$\min \sum_{t=1}^{H-1} \sum_{n=1}^{n_t} p^n q \sum_{k=1}^{K} \sum_{i=1}^{O_k} \sum_{j=1}^{D} t_{ij} x^n_{i(k)j} + \sum_{t=2}^{H} \sum_{n=1}^{n_t} p^n \left[\sum_{j=1}^{D} q b_j \, y^n_j - (1-\alpha) q \sum_{k=1}^{K} \sum_{i=1}^{O_k} \sum_{j=1}^{D} t_{ij} \left(x^{a(n)}_{i(k)j} - z^n_{i(k)j} \right) \right]$$

$$\begin{aligned}
s.t. \quad q \sum_{k=1}^{K} \sum_{i=1}^{O_k} x^n_{i(k)j} &\leq g_j \ , \ j \in \mathscr{D}, \ n \in \mathscr{N}^t, \ t \neq H \\
0 \leq l^{a(n)}_j + q \sum_{k=1}^{K} \sum_{i=1}^{O_k} z^n_{i(k)j} + y^n_j - d^n_j &\leq l_{\max} \ , \ j \in \mathscr{D}, \ n \in \mathscr{N}^t, \ t \neq 1 \\
z^n_{i(k)j} &\leq x^{a(n)}_{i(k)j} \ , \ i \in \mathscr{O}_k, \ k \in \mathscr{K}, \ j \in \mathscr{D}, \ n \in \mathscr{N}^t, \ t \neq 1 \\
r_k \leq q \sum_{j=1}^{D} z^n_{i(k)j} &\leq v_k \ , \ i \in \mathscr{O}_k, \ k \in \mathscr{K}, \ n \in \mathscr{N}^t, \ t \neq 1 \\
x^n_{i(k)j} &\in \mathbb{N} \ , \ i \in \mathscr{O}_k, \ k \in \mathscr{K}, \ j \in \mathscr{D}, \ n \in \mathscr{N}^t, \ t \neq H \\
y^n_j &\in \mathbb{N} \ , \ j \in \mathscr{D}, \ n \in \mathscr{N}^t, \ t \neq 1 \\
z^n_{i(k)j} &\in \mathbb{N} \ , \ \ i \in \mathscr{O}_k, \ k \in \mathscr{K}, \ j \in \mathscr{D}, \ n \in \mathscr{J}^t, \ t \neq 1
\end{aligned}$$

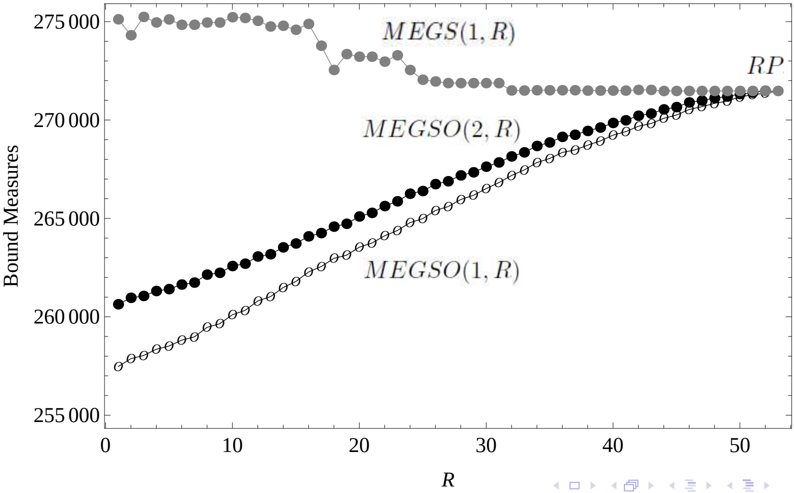
Benchmarch Scenario trees



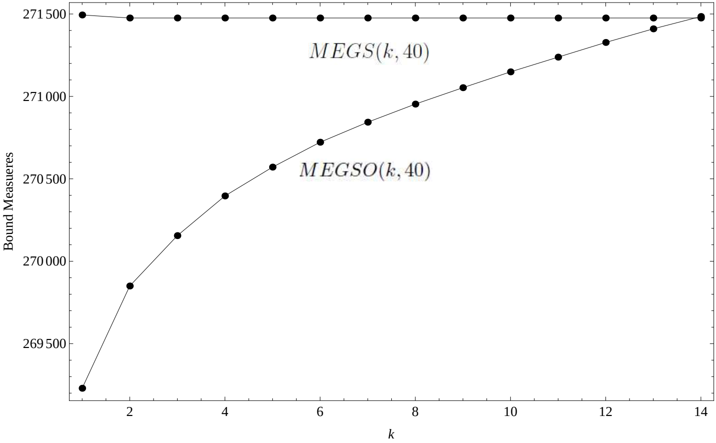
	$S = 54$	$S = 210$	$S = 840$
number of stages	5	4	5
number of nodes	94	260	1100
number of variables	66431	156105	685305
number of integer variables	63840	148320	652320
number of linear constraints	2591	140625	597375
CPU time (s)	12.6	52.3409	557.013

Table : Summary statistics of the three benchmark scenario trees respectively with 54, 210 and 840 scenarios.

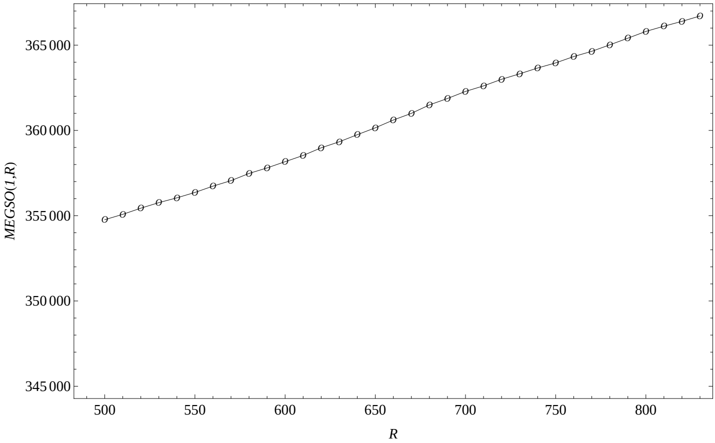
Lower and upper bounds for the supply planning problem



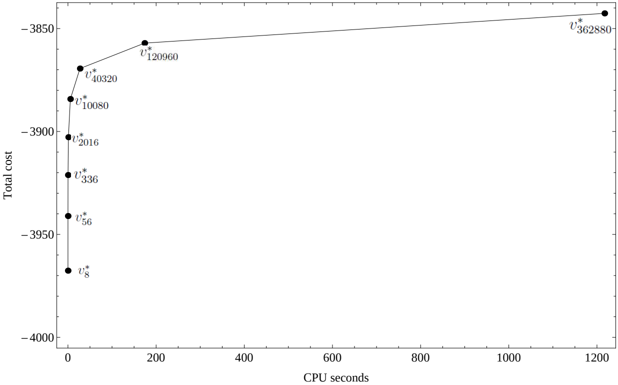
Lower and upper bounds for the supply planning problem



Lower and upper bounds for the supply planning problem



Lower bounds: 10-stage scenario tree with $S = 8 \times 7 \times 6 \times 6 \times 5 \times 4 \times 3 \times 3 \times 2 = 725760$ scenarios.



- The difference between the best upper and lower bounds obtained:

$$EEV^8 - v^*_{362880} = 299.45, \quad (7.79\% \; v^*_{362880}).$$

Quality measures: GLUSS

Table : Results of $GLUSS(p,3)$ and $GLUSS(p,10)$ for the STP as % from RP

Instance	VSS	$GLUSS(p,3)$			$GLUSS(p,10)$									
		1	2	3	1	2	3	4	5	6	7	8	9	10
1	∞	∞	0	0	∞	∞	0.006	0	0	0	0	0	0	0
2	∞	∞	0	0	∞	∞	0.006	0	0	0	0	0	0	0
3	∞	∞	0	0	∞	∞	0.008	0	0	0	0	0	0	0
4	∞	∞	0	0	∞	∞	0.084	0	0	0	0	0	0	0
5	∞	∞	0	0	∞	∞	0.001	0	0	0	0	0	0	0
6	∞	∞	0	0	∞	∞	0	0	0	0	0	0	0	0
7	∞	∞	0	0	∞	∞	0.01	0	0	0	0	0	0	0
8	∞	∞	0	0	∞	∞	0.006	0	0	0	0	0	0	0
9	∞	∞	0	0	∞	∞	0.002	0	0	0	0	0	0	0
10	∞	∞	0	0	∞	∞	0.002	0	0	0	0	0	0	0
Mean	∞	∞	0	0	∞	∞	0.0129	0	0	0	0	0	0	0

Highlights and general trends

We summarize the lessons learned from our experiments:

Which variables we have to fix?

1. to reduce the problem size, it would be preferable to **fix the largest possible number of variables**;
2. fixing too large a number may result in errors in terms of **feasibility and optimality**.

- **General trend**: fixing to 0 about **33%** of the non-basic variables with the highest reduced costs:
    1. we **reached the optimal stochastic solutions** without feasibility issues
    2. **reducing the computational time** up to one order of magnitude for the largest instance

- **Optimality**: the reduced costs obtained from a continuous relaxation of an integral problem **hint to the variables to make inactive** in order to guide the search for optimal solutions to stochastic programs

How to apply GLUSS to a new problem?

An **empirical method**

- **Divide** the reduced costs in $N = 3$ intervals and fix in the SP first stage solution the variables belonging to the third class;
- If there are **feasibility issues**, consider the removed interval and split it again into three sub-intervals;
- If one desires a **greater precision** split the out-of basis variables into a **larger number of bids**.

Conclusions

We have analyzed **lower** and **upper bounds** for multistage programs both in the **linear** and **non linear cases**:

- By changing the **probability measure** to get lower bounds;
- By **inserting a feasible solution** to get upper bounds;
  - By analizing the **Quality** of Deterministic Solution

The results apply also to the **integer case**.

Differences among the values allow to evaluate:

- if it is worth the **additional computations** for the stochastic program versus the simplified approaches proposed.
- give insight of what is **potentially wrong** with a solution coming from the **deterministic/approximated method** considered.

References

1. *Birge, Louveaux*: Introduction to stochastic programming, Springer-Verlag, New York, (2011).
2. *Kall, Wallace*: Stochastic Programming, Wiley-Interscience Series in Systems and Optimization, Wiley, Chichester, (1994).
3. *Madansky*: Inequalities for stochastic linear programming problems. *Manag. Sci.*, **6**, 197–204 (1960).
4. *Maggioni, Wallace*: Analyzing the quality of the expected value solution in stochastic programming. *AOR* **200**, 37–54 (2012).
5. *Maggioni, Allevi, Bertocchi*: Bounds in multistage linear stochastic programming. *JOTA* DOI: $10.1007/s10957-013-0450-1$ **163**(1) (2014).
6. *Maggioni, Pflug*: Bounds and Approximations for multistage stochastic programs *SIOPT* **26**(1), 831–855 (2016)
7. *Maggioni, Allevi, Bertocchi*: Monotonic bounds in multistage mixed-integer linear stochastic programming, *Computational Management Science* DOI 10.1007/s10287-016-0254-5 (2016)
8. *Maggioni, Crainic, Perboli, Rei*: The generalized skeleton solution: a new measure of the quality of the deterministic solution in stochastic programming, CIRRELT-2015-21 (under evaluation).