Effective Scenarios in Multistage Stochastic Optimization

Tito Homem-de-Mello

School of Business
Universidad Adolfo Ibañez
Santiago, CHILE

Joint work with Hamed Rahimian (Northwestern University) and Güzin Bayraksan (Ohio State Univerisity)

Multi-Stage Stochastic Optimization for Clean Energy Transition
Oaxaca 2019

Outline

Outline

Motivation

Suppose we have a cost function $h(x, \omega)$ that depends on the decisions x and the uncertainty $\omega \in \Omega := \{\omega_1, \dots, \omega_n\}$.

- Each ω_i is a scenario.

Motivation

Suppose we have a cost function $h(x, \omega)$ that depends on the decisions x and the uncertainty $\omega \in \Omega := \{\omega_1, \ldots, \omega_n\}$.

- Each ω_i is a scenario.

We want to make the best decision that protects us from the risk of high costs.

Motivation

Suppose we have a cost function $h(x, \omega)$ that depends on the decisions x and the uncertainty $\omega \in \Omega := \{\omega_1, \dots, \omega_n\}$.

- Each ω_i is a scenario.

We want to make the best decision that protects us from the risk of high costs.

That is, we want to choose x that minimizes (some function of) the highest costs.

Motivation

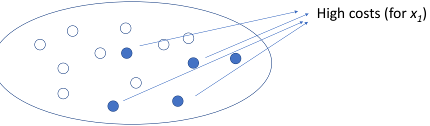
The question

What are the scenarios in Ω that really "matter" for the problem?

Motivation

The question

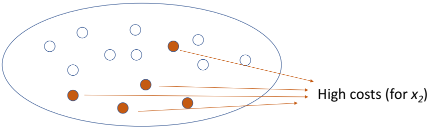
What are the scenarios in Ω that really "matter" for the problem?



Motivation

The question

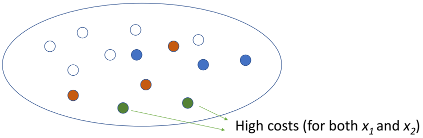
What are the scenarios in Ω that really “matter” for the problem?



Motivation

The question

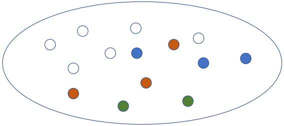
What are the scenarios in Ω that really "matter" for the problem?



Motivation

The question

What are the scenarios in Ω that really "matter" for the problem?



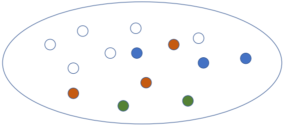
What are the important scenarios?

- Blue? (x_1)
- Red? (x_2)
- Green? (x_1 and x_2)
- All of them?
- How about other feasible points x?

Motivation

The question

What are the scenarios in Ω that really "matter" for the problem?



What are the important scenarios?

- Blue? (x_1)
- Red? (x_2)
- Green? (x_1 and x_2)
- All of them?
- How about other feasible points x?

An informal definition: a scenario is important if its removal changes the optimal objective value of the problem.

Why is this relevant?

The knowledge of which scenarios are "important" can be useful to the decision maker and/or for algorithmic purposes. For example,

Why is this relevant?

The knowledge of which scenarios are "important" can be useful to the decision maker and/or for algorithmic purposes. For example,

- To provide insights about the system.
  - E.g., which components are more stressed.

Why is this relevant?

The knowledge of which scenarios are “important” can be useful to the decision maker and/or for algorithmic purposes. For example,

- To provide insights about the system.
  - E.g., which components are more stressed.
- To help the decision maker to choose an appropriate level of risk (Rahimian, Bayraksan, HdM, 2019).

Why is this relevant?

The knowledge of which scenarios are "important" can be useful to the decision maker and/or for algorithmic purposes. For example,

- To provide insights about the system.
  - E.g., which components are more stressed.
- To help the decision maker to choose an appropriate level of risk (Rahimian, Bayraksan, HdM, 2019).
- To guide scenario generation/reduction methods.
  - Better representation of the uncertainty can lead to more efficient solution methods for risk-averse problems (Arpón, HdM, Pagnoncelli, 2018).

Why is this relevant?

The knowledge of which scenarios are "important" can be useful to the decision maker and/or for algorithmic purposes. For example,

- To provide insights about the system.
  - E.g., which components are more stressed.
- To help the decision maker to choose an appropriate level of risk (Rahimian, Bayraksan, HdM, 2019).
- To guide scenario generation/reduction methods.
  - Better representation of the uncertainty can lead to more efficient solution methods for risk-averse problems (Arpón, HdM, Pagnoncelli, 2018).
- To accelerate Decomposition Algorithms.

Motivation

Address the following research question in the context of multistage stochastic programs

Motivation

Address the following research question in the context of multistage stochastic programs

- What uncertain scenarios are *important* to a multistage stochastic programming model?
  - How to define *important* scenarios?
  - How to identify *important* scenarios?
  - Do these correspond to the highest cost scenarios?

Motivation

Address the following research question in the context of multistage stochastic programs

- What uncertain scenarios are *important* to a multistage stochastic programming model?
    - How to define *important* scenarios?
    - How to identify *important* scenarios?
    - Do these correspond to the highest cost scenarios?

Let's first look at the two-stage case.

Outline

Static/Two-Stage DRSP

We shall look at the problem from a distributionally robust perspective, i.e.,

$$\min_{x \in \mathcal{X}} \left\{ f(x) := \max_{\mathbf{p} \in \mathcal{P}} \mathbb{E}_{\mathbf{p}} \left[h(x, \omega) \right] \right\},$$

where

- $\mathcal{X} \subseteq \mathbb{R}^n$ is a deterministic and non-empty convex compact set,
- Ω is sample space, assumed finite
- $h : \mathcal{X} \times \Omega \mapsto \mathbb{R}$ is an integrable convex random function, i.e., for any $x \in \mathcal{X}$, $h(x, \cdot)$ is integrable, and $h(\cdot, \omega)$ is convex q-almost surely,

Static/Two-Stage DRSP

We shall look at the problem from a distributionally robust perspective, i.e.,

$$\min_{x \in \mathcal{X}} \left\{ f(x) := \max_{\mathbf{p} \in \mathcal{P}} \mathbb{E}_{\mathbf{p}} \left[h(x, \omega) \right] \right\},$$

where

- $\mathcal{P}$ is the ambiguity set of distributions, a subset of all probability distributions on Ω, which may contain, e.g., all distributions that are not far from some reference distribution — e.g., the empirical distribution corresponding to the data.
- Useful in particular when we don't have full confidence that $\mathbf{p}$ is the "correct" probability distribution.

Assessment Problem of "Removed" Scenarios

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios:

$$\mathcal{P}^{\mathsf{A}}(\mathcal{F}) := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0,\ \omega \in \mathcal{F}\}.$$

Assessment Problem of "Removed" Scenarios

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios:

$$\mathcal{P}^{\mathsf{A}}(\mathcal{F}) := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0,\ \omega \in \mathcal{F}\}.$$

The Assessment problem of scenarios in $\mathcal{F}$ is

$$\min_{x \in \mathcal{X}} \left\{ f^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\}.$$

Assessment Problem of "Removed" Scenarios

Consider "removing" a set $\mathcal{F} \subset \Omega$ of scenarios:

$$\mathcal{P}^{\mathsf{A}}(\mathcal{F}) := \{\mathbf{p} \in \mathcal{P} : p_\omega = 0,\ \omega \in \mathcal{F}\}.$$

The Assessment problem of scenarios in $\mathcal{F}$ is

$$\min_{x \in \mathcal{X}} \left\{ f^{\mathsf{A}}(x; \mathcal{F}) = \max_{\mathbf{p} \in \mathcal{P}^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h_\omega(x) \right\}.$$

If inner max of the Assessment Problem is infeasible: $f^{\mathsf{A}}(x; \mathcal{F}) = -\infty$.

Effective/Ineffective Scenarios in DRSP

(Rahimian, Bayraksan, HdM, 2018)

Definition (Effective Subset of Scenarios)

A subset $\mathcal{F} \subset \Omega$ is called effective if by its "removal" the optimal value of the Assessment problem becomes strictly smaller than the optimal value of DRSP; i.e., if

$$\min_{x \in \mathcal{X}} f^{\mathsf{A}}(x; \mathcal{F}) < \min_{x \in \mathcal{X}} f(x)$$

.

Effective/Ineffective Scenarios in DRSP

(Rahimian, Bayraksan, HdM, 2018)

Definition (Effective Subset of Scenarios)

A subset $\mathcal{F} \subset \Omega$ is called effective if by its "removal" the optimal value of the Assessment problem becomes strictly smaller than the optimal value of DRSP; i.e., if

$$\min_{x \in \mathcal{X}} f^{\mathsf{A}}(x; \mathcal{F}) < \min_{x \in \mathcal{X}} f(x)$$

.

Definition (Ineffective Subset of Scenarios)

A subset $\mathcal{F} \subset \Omega$ that is not effective is called ineffective.

A particular case: the Total Variation distance

IDEA: Define the ambiguity set $\mathcal{P}$ as the set of distributions that are not too far from a given reference distribution q.

A particular case: the Total Variation distance

IDEA: Define the ambiguity set $\mathcal{P}$ as the set of distributions that are not too far from a given reference distribution q.

- This idea has been popular in the literature

A particular case: the Total Variation distance

IDEA: Define the ambiguity set $\mathcal{P}$ as the set of distributions that are not too far from a given reference distribution q.

- This idea has been popular in the literature
- Many different ways to measure the distance between distributions (ϕ-divergences, Wasserstein distances, etc.); e.g. Ben-Tal et al. (2013); Blanchet et al. (2016); Esfahani and Kuhn (2015); Gao and Kleywegt (2016); Jiang and Guan (2015); Pflug and Wozabal (2007),...

A particular case: the Total Variation distance

IDEA: Define the ambiguity set $\mathcal{P}$ as the set of distributions that are not too far from a given reference distribution q.

- This idea has been popular in the literature
- Many different ways to measure the distance between distributions (ϕ-divergences, Wasserstein distances, etc.); e.g. Ben-Tal et al. (2013); Blanchet et al. (2016); Esfahani and Kuhn (2015); Gao and Kleywegt (2016); Jiang and Guan (2015); Pflug and Wozabal (2007),...
- We work with the Total Variation distance, defined as

$$d_{\mathrm{TV}}(p,q) \;:=\; \frac{1}{2}\sum_k |p_k - q_k|.$$

DRSO formulation with Total Variation distance

$$\min_{x \in \mathcal{X}} \left\{ f_\gamma(x) := \max_{p \in \mathcal{P}_\gamma} \sum_{k=1}^{n} p_k h(x, \omega_k) \right\}$$

where

$$\mathcal{P}_\gamma := \left\{ p : \frac{1}{2} \sum_{k=1}^{n} |q_k - p_k| \le \gamma \quad \sum_{k=1}^{n} p_k = 1 \quad p_k \ge 0,\ \forall k \right\}$$

- Hedge against the worst probability distribution

DRSO formulation with Total Variation distance

$$\min_{x\in\mathcal{X}} \left\{ f_\gamma(x) := \max_{p\in\mathcal{P}_\gamma} \sum_{k=1}^{n} p_k h(x,\omega_k) \right\}$$

where

$$\mathcal{P}_\gamma := \left\{ p : \frac{1}{2}\sum_{k=1}^{n} |q_k - p_k| \le \gamma \right.$$

$$\sum_{k=1}^{n} p_k = 1$$

$$\left. p_k \ge 0,\ \forall k \right\}$$

- Hedge against the worst probability distribution
- Restrict total variation distance from reference distribution

DRSO formulation with Total Variation distance

$$\min_{x\in\mathcal{X}} \left\{ f_\gamma(x) := \max_{p\in\mathcal{P}_\gamma} \sum_{k=1}^{n} p_k h(x,\omega_k) \right\}$$

where

$$\mathcal{P}_\gamma := \left\{ p : \frac{1}{2}\sum_{k=1}^{n} |q_k - p_k| \le \gamma \quad \sum_{k=1}^{n} p_k = 1 \quad p_k \ge 0,\ \forall k \right\}$$

- Hedge against the worst probability distribution
- Restrict total variation distance from reference distribution
- Ensure a probability distribution

A risk-averse interpretation

A key property of DRSO with variation distance is given below:

Proposition (Jiang and Guan (2015))

$$f_\gamma(x) \;=\; \gamma \sup_{\omega\in\Omega} h(x,\omega) + (1-\gamma)\mathrm{CVaR}_\gamma\left[h(x,\omega)\right]$$

So we see that DRSO with total variation is equivalent to a risk-averse problem with $\mathcal{R}[Z] = \gamma \sup Z + (1-\gamma)\mathrm{CVaR}_\gamma[Z]$. Moreover,

- For $\gamma = 0$, $f_\gamma(x) = \mathbb{E}_q\left[h(x,\omega)\right]$ (risk-neutral problem).
- For $\gamma = 1$, $f_\gamma(x) = \sup_{\omega\in\Omega} h(x,\omega)$ (robust optimization)

Analyzing the removal of scenarios

Recall that removing a subset of scenarios $\mathcal{F}$ means solving the problem

$$\min_{x \in \mathcal{X}} \left\{ f_\gamma^{\mathsf{A}}(x, \mathcal{F}) := \max_{\mathbf{p} \in \mathcal{P}_\gamma^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^{\mathsf{c}}} p_\omega h(x, \omega) \right\}.$$

where $\mathcal{P}_\gamma^{\mathsf{A}}(\mathcal{F}) := \{\mathbf{p} \in \mathcal{P}_\gamma : p_\omega = 0,\ \omega \in \mathcal{F}\}$.

Analyzing the removal of scenarios

Recall that removing a subset of scenarios $\mathcal{F}$ means solving the problem

$$\min_{x \in \mathcal{X}} \left\{ f_\gamma^{\mathsf{A}}(x, \mathcal{F}) := \max_{\mathbf{p} \in \mathcal{P}_\gamma^{\mathsf{A}}(\mathcal{F})} \sum_{\omega \in \mathcal{F}^c} p_\omega h(x, \omega) \right\}.$$

where $\mathcal{P}_\gamma^{\mathsf{A}}(\mathcal{F}) := \{\mathbf{p} \in \mathcal{P}_\gamma : p_\omega = 0,\ \omega \in \mathcal{F}\}$.

BUT:

- We don't want to re-solve the problem!
- The difficulty to compare the minimum of the function f_γ^{A} with that of the original function f_γ is that the optimization problem in the assessment problem has extra constraints ($p_\omega = 0$ for $\omega \in \mathcal{F}$).

Analyzing the removal of scenarios (cont.)

Theorem (Risk-averse interpretation of removal of scenarios, Rahimian, Bayraksan and HdM 2018)

$$f_\gamma^A(x, \mathcal{F}) \;=\; \gamma \sup_{\omega \in \mathcal{F}^c} h(x, \omega) + (1-\gamma)\mathrm{CVaR}_{\gamma_{\mathcal{F}}}\left[h(x,\omega) \,|\, \mathcal{F}^c\right],$$

where $\gamma_{\mathcal{F}} := \frac{\gamma - q(\mathcal{F})}{1 - q(\mathcal{F})}$.

Analyzing the removal of scenarios (cont.)

Theorem (Risk-averse interpretation of removal of scenarios, Rahimian, Bayraksan and HdM 2018)

$$f_\gamma^A(x, \mathcal{F}) \; = \; \gamma \sup_{\omega \in \mathcal{F}^c} h(x, \omega) + (1-\gamma)\mathrm{CVaR}_{\gamma_\mathcal{F}} \left[h(x,\omega) \,|\, \mathcal{F}^c \right],$$

where $\gamma_\mathcal{F} := \frac{\gamma - q(\mathcal{F})}{1-q(\mathcal{F})}$.

Now it's easier to compare the minima of f_γ^A and f_γ!

Results for DRSP with Total Variation Distance

(Rahimian, Bayraksan, HdM, 2018)

Consider an optimal solution $(x^*, \mathbf{p}^*) \in \mathcal{X} \times \mathcal{P}_\gamma$ to DRSP-V:

$$x^* \in \underset{x \in \mathcal{X}}{\operatorname{argmin}} \; \mathbb{E}_{\mathbf{p}*} [h(x, \omega)]$$

$$\mathbf{p}^* := \mathbf{p}^*(x^*) \in \underset{\mathbf{p} \in \mathcal{P}_\gamma}{\operatorname{argmax}} \; \mathbb{E}_{\mathbf{p}} [h(x^*, \omega)]$$

- Typically, the effective scenarios are those ω for which $p^*_\omega > 0$ (the tail of $h(x^*, \cdot)$), and the ineffective scenarios are those ω for which $p^*_\omega = 0$.

Results for DRSP with Total Variation Distance

(Rahimian, Bayraksan, HdM, 2018)

Consider an optimal solution $(x^*, \mathbf{p}^*) \in \mathcal{X} \times \mathcal{P}_\gamma$ to DRSP-V:

$$x^* \in \underset{x \in \mathcal{X}}{\text{argmin}}\ \mathbb{E}_{\mathbf{p}*}\left[h(x,\omega)\right]$$

$$\mathbf{p}^* := \mathbf{p}^*(x^*) \in \underset{\mathbf{p} \in \mathcal{P}_\gamma}{\text{argmax}}\ \mathbb{E}_{\mathbf{p}}\left[h(x^*,\omega)\right]$$

- Typically, the <u>effective</u> scenarios are those ω for which $p^*_\omega > 0$ (the tail of $h(x^*, \cdot)$), and the <u>ineffective</u> scenarios are those ω for which $p^*_\omega = 0$.
- There is however an important exception:
  - Multiple scenarios ω such that $h(x^*, \omega) = \mathrm{VaR}_{\gamma,q}\left[h(x^*, \cdot)\right]$.

Results for DRSP with Total Variation Distance

(Rahimian, Bayraksan, HdM, 2018)

Consider an optimal solution $(x^*, \mathbf{p}^*) \in \mathcal{X} \times \mathcal{P}_\gamma$ to DRSP-V:

$$x^* \in \underset{x \in \mathcal{X}}{\text{argmin}}\ \mathbb{E}_{\mathbf{p}*}\left[h(x, \omega)\right]$$

$$\mathbf{p}^* := \mathbf{p}^*(x^*) \in \underset{\mathbf{p} \in \mathcal{P}_\gamma}{\text{argmax}}\ \mathbb{E}_{\mathbf{p}}\left[h(x^*, \omega)\right]$$

- Typically, the effective scenarios are those ω for which $p^*_\omega > 0$ (the tail of $h(x^*, \cdot)$), and the ineffective scenarios are those ω for which $p^*_\omega = 0$.
- There is however an important exception:
  - Multiple scenarios ω such that $h(x^*, \omega) = \mathrm{VaR}_{\gamma, q}\left[h(x^*, \cdot)\right]$.
  - Note that in the above case $\mathbf{p}^*$ is not unique.

Results for DRSP with Total Variation Distance

(Rahimian, Bayraksan, HdM, 2018)

Consider an optimal solution $(x^*, \mathbf{p}^*) \in \mathcal{X} \times \mathcal{P}_\gamma$ to DRSP-V:

$$x^* \in \underset{x \in \mathcal{X}}{\text{argmin}}\ \mathbb{E}_{\mathbf{p}*}[h(x, \omega)]$$

$$\mathbf{p}^* := \mathbf{p}^*(x^*) \in \underset{\mathbf{p} \in \mathcal{P}_\gamma}{\text{argmax}}\ \mathbb{E}_{\mathbf{p}}[h(x^*, \omega)]$$

- Typically, the <u>effective</u> scenarios are those ω for which $p^*_\omega > 0$ (the tail of $h(x^*, \cdot)$), and the <u>ineffective</u> scenarios are those ω for which $p^*_\omega = 0$.
- There is however an important exception:
  - Multiple scenarios ω such that $h(x^*, \omega) = \mathrm{VaR}_{\gamma, q}[h(x^*, \cdot)]$.
  - Note that in the above case $\mathbf{p}^*$ is not unique.
- When we can determine the (in)effectiveness of a scenario, we say it satisfies the identifying conditions.

Related concepts

Supporting constraints (Campi and Garatti, 2018): Consider the problem

$$\min\ c^T x \ \text{ s.t. } \ x \in X_{\delta^i},\ i = 1, \ldots, n,$$

where $\delta^1, \ldots, \delta^n$ are scenarios. Then, X_{δ^j} is a supporting constraint if its removal changes the optimal solution.

Related concepts

Supporting constraints (Campi and Garatti, 2018): Consider the problem

$$\min\ c^T x \ \text{ s.t. } \ x \in X_{\delta^i},\ i = 1, \ldots, n,$$

where $\delta^1, \ldots, \delta^n$ are scenarios. Then, X_{δ^j} is a supporting constraint if its removal changes the optimal solution.

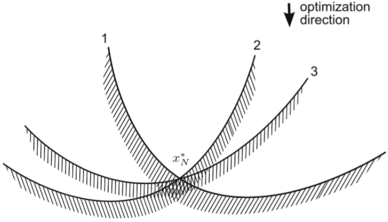


Fig. 2 Constraints 1, 2, and 3 are active, but 1 is the only support constraint, since removing 2 while maintaining 1 and 3 or removing 3 while maintaining 1 and 2 does not change the solution x_N^*. If the sole support constraint is maintained, then the solution moves to a lower value

Related concepts

Supporting constraints (Campi and Garatti, 2018): Consider the problem

$$\min\ c^T x \ \text{ s.t. } \ x \in X_{\delta^i},\ i = 1, \ldots, n,$$

where $\delta^1, \ldots, \delta^n$ are scenarios. Then, X_{δ^j} is a supporting constraint if its removal changes the optimal solution.

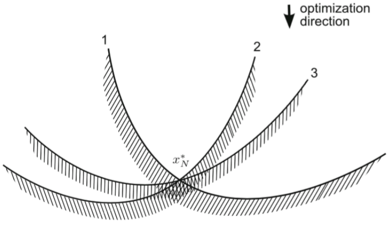


Fig. 2 Constraints 1, 2, and 3 are active, but 1 is the only support constraint, since removing 2 while maintaining 1 and 3 or removing 3 while maintaining 1 and 2 does not change the solution x_N^*. If the sole support constraint is maintained, then the solution moves to a lower value

Campi and Garatti (2018) use this notion to estimate the the probability that a new randomly selected constraint X_δ is violated by x_n^*.

Related concepts (cont.)

Strictly monotone risk measures (Shapiro, 2017): A coherent risk measure ρ is strictly monotone if, for any random variable Z, all the maximizers $\mathbf{p}^*$ of

$$\rho(Z) = \max_{\mathbf{p} \in \mathcal{Q}} \sum_{k=1}^{n} p_k Z(\omega_k)$$

satisfy $p^*_\omega > 0$ for all $\omega \in \Omega$.

UNIVERSIDAD ADOLFO IBÁÑEZ

Related concepts (cont.)

Strictly monotone risk measures (Shapiro, 2017): A coherent risk measure ρ is strictly monotone if, for any random variable Z, all the maximizers $\mathbf{p}^*$ of

$$\rho(Z) = \max_{\mathbf{p} \in \mathcal{Q}} \sum_{k=1}^{n} p_k Z(\omega_k)$$

satisfy $p^*_\omega > 0$ for all $\omega \in \Omega$.

- Shapiro (2017) shows (in a more general setting) that strict monotonicity is required to establish a one-to-one correspondence between

$$\underset{x}{\operatorname{argmin}}\ h(x, \cdot) \quad \text{and} \quad \underset{\chi}{\operatorname{argmin}}\ \rho(h(\chi(\cdot), \cdot)).$$

Related concepts (cont.)

Strictly monotone risk measures (Shapiro, 2017): A coherent risk measure ρ is strictly monotone if, for any random variable Z, all the maximizers $\mathbf{p}^*$ of

$$\rho(Z) = \max_{\mathbf{p} \in \mathcal{Q}} \sum_{k=1}^{n} p_k Z(\omega_k)$$

satisfy $p^*_\omega > 0$ for all $\omega \in \Omega$.

- Shapiro (2017) shows (in a more general setting) that strict monotonicity is required to establish a one-to-one correspondence between

$$\operatorname*{argmin}_x \ h(x, \cdot) \quad \text{and} \quad \operatorname*{argmin}_\chi \ \rho(h(\chi(\cdot), \cdot)).$$

- Strict monotonicity means that $\mathcal{Q}$ is contained in the interior of the simplex $\left\{\mathbf{p} \,:\, p_\omega \geq 0, \ \sum_\omega p_\omega = 1\right\}$.

Related concepts (cont.)

Strictly monotone risk measures (Shapiro, 2017): A coherent risk measure ρ is strictly monotone if, for any random variable Z, all the maximizers $\mathbf{p}^*$ of

$$\rho(Z) = \max_{\mathbf{p} \in \mathcal{Q}} \sum_{k=1}^{n} p_k Z(\omega_k)$$

satisfy $p^*_\omega > 0$ for all $\omega \in \Omega$.

- Shapiro (2017) shows (in a more general setting) that strict monotonicity is required to establish a one-to-one correspondence between

$$\operatorname*{argmin}_x \, h(x, \cdot) \quad \text{and} \quad \operatorname*{argmin}_\chi \, \rho(h(\chi(\cdot), \cdot)).$$

- Strict monotonicity means that $\mathcal{Q}$ is contained in the interior of the simplex $\left\{\mathbf{p} \,:\, p_\omega \geq 0, \, \sum_\omega p_\omega = 1\right\}$.
- Therefore, in such cases all scenarios are effective.

Outline

General Formulation of MSSP

$$\min_{x_1, x_2, \ldots, x_T} \mathbb{E}\left[g_1(x_1, \xi_1) + g_2(x_2, \xi_2) + \ldots + g_T(x_T, \xi_T)\right]$$
$$\text{s.t.} \quad x_t \in \mathcal{X}_t := \mathcal{X}_t(x_{[t-1]}, \xi_{[t]}), \; t = 1, 2, \ldots T,$$

where

- $\xi_{[t]}$ and $x_{[t]}$: history of stochastic process and decisions up to stage t
- $x_t := x_t(\xi_{[t]})$: decision made at each stage
- $\mathcal{X}_t := \mathcal{X}_t(x_{[t-1]}, \xi_{[t]})$: feasibility set in stage t
- $g_t(x_t, \xi_t)$: cost of decision x_t given the realized uncertainty ξ_t at stage t

General Formulation of MSSP

$$\min_{x_1, x_2, \ldots, x_T} \mathbb{E}\left[g_1(x_1, \xi_1) + g_2(x_2, \xi_2) + \ldots + g_T(x_T, \xi_T)\right]$$
$$\text{s.t.} \quad x_t \in \mathcal{X}_t := \mathcal{X}_t(x_{[t-1]}, \xi_{[t]}), \; t = 1, 2, \ldots T,$$

where

- $\xi_{[t]}$ and $x_{[t]}$: history of stochastic process and decisions up to stage t
- $x_t := x_t(\xi_{[t]})$: decision made at each stage
- $\mathcal{X}_t := \mathcal{X}_t(x_{[t-1]}, \xi_{[t]})$: feasibility set in stage t
- $g_t(x_t, \xi_t)$: cost of decision x_t given the realized uncertainty ξ_t at stage t

General Formulation of MSSP

$$\min_{x_1, x_2, \ldots, x_T} \mathbb{E}\left[g_1(x_1, \xi_1) + g_2(x_2, \xi_2) + \ldots + g_T(x_T, \xi_T)\right]$$
$$\text{s.t.} \quad x_t \in \mathcal{X}_t := \mathcal{X}_t(x_{[t-1]}, \xi_{[t]}), \; t = 1, 2, \ldots T,$$

where

- $\xi_{[t]}$ and $x_{[t]}$: history of stochastic process and decisions up to stage t
- $x_t := x_t(\xi_{[t]})$: decision made at each stage
- $\mathcal{X}_t := \mathcal{X}_t(x_{[t-1]}, \xi_{[t]})$: convex feasibility set in stage t
- $g_t(x_t, \xi_t)$: convex cost of decision x_t given the realized uncertainty ξ_t at stage t

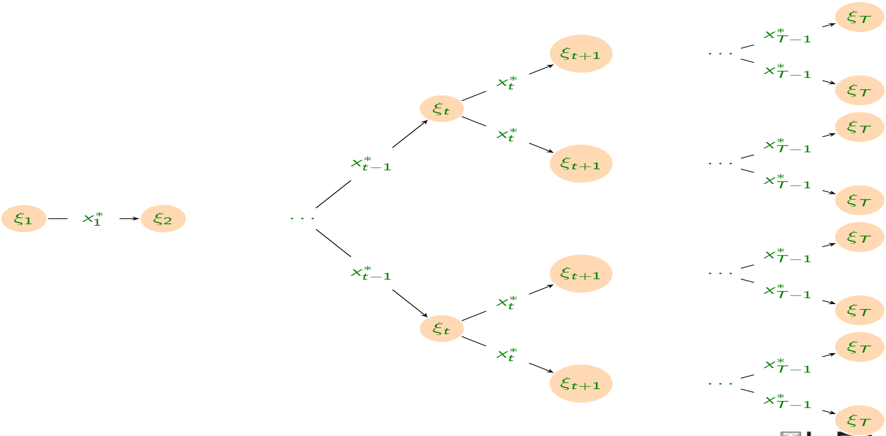
Nested Formulation of MSSP

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \mathbb{E}_{\mathbf{q}_2|\xi_{[1]}} \left[\min_{x_2 \in \mathcal{X}_2} g_2(x_2, \xi_2) + \mathbb{E}_{\mathbf{q}_3|\xi_{[2]}} \left[\ldots + \mathbb{E}_{\mathbf{q}_T|\xi_{[T-1]}} \left[\min_{x_T \in \mathcal{X}_T} g_T(x_T, \xi_T) \right] \ldots \right] \right]$$

- $\mathbf{q}_{t|\xi_{[t-1]}}$: conditional distribution of stage t, conditioned on $\xi_{[t-1]}$
- $\mathbb{E}_{\mathbf{q}_t|\xi_{[t-1]}}[\cdot]$: conditional expectation w.r.t. $\mathbf{q}_{t|\xi_{[t-1]}}$

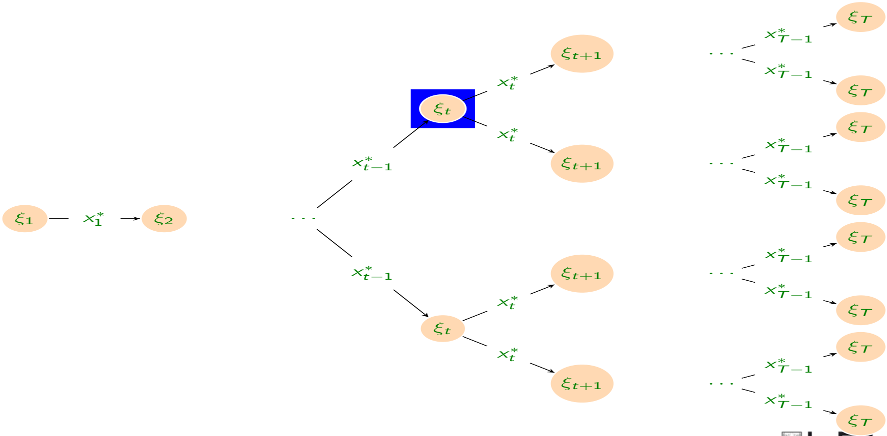
Toward a Nested Formulation of Multistage DRSP

Given a scenario tree and a nominal distribution on the tree



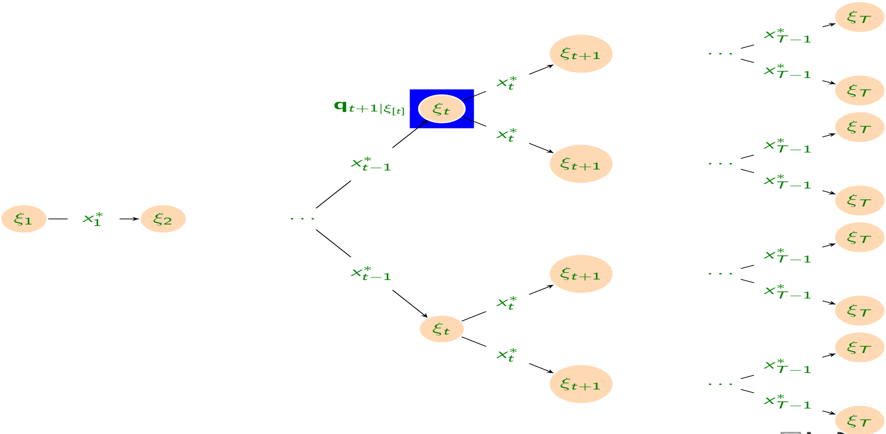
Toward a Nested Formulation of Multistage DRSP

Given a scenario tree and a nominal distribution on the tree



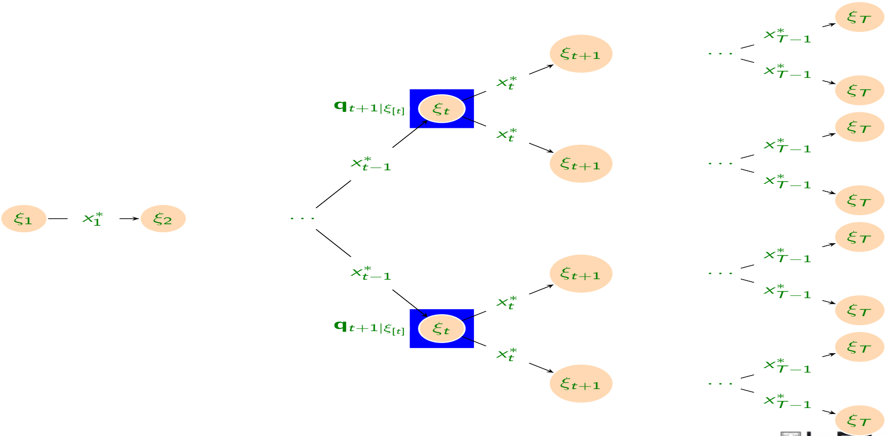
Toward a Nested Formulation of Multistage DRSP

Given a scenario tree and a nominal distribution on the tree



Toward a Nested Formulation of Multistage DRSP

Given a scenario tree and a nominal distribution on the tree



Nested Formulation of Multistage DRSP

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \mathbb{E}_{\mathbf{q}_2|\xi_{[1]}} \left[\min_{x_2 \in \mathcal{X}_2} g_2(x_2, \xi_2) + \mathbb{E}_{\mathbf{q}_3|\xi_{[2]}} \left[\cdots + \mathbb{E}_{\mathbf{q}_T|\xi_{[T-1]}} \left[\min_{x_T \in \mathcal{X}_T} g_T(x_T, \xi_T) \right] \cdots \right] \right],$$

Nested Formulation of Multistage DRSP

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \bigvee^{\max_{\mathbf{p}_2 \in \mathcal{P}_{2|\xi_{[1]}}}} \mathbb{E}_{\mathbf{p}_2} \Bigg[\min_{x_2 \in \mathcal{X}_2} g_2(x_2, \xi_2) + \bigvee^{\max_{\mathbf{p}_3 \in \mathcal{P}_{3|\xi_{[2]}}}} \mathbb{E}_{\mathbf{p}_3} \Big[\ldots + \bigvee^{\max_{\mathbf{p}_T \in \mathcal{P}_{T|\xi_{[T-1]}}}} \mathbb{E}_{\mathbf{p}_T} \Big[\min_{x_T \in \mathcal{X}_T} g_T(x_T, \xi_T) \Big] \ldots \Big] \Bigg],$$

Nested Formulation of Multistage DRSP

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \max_{\mathbf{p}_2 \in \mathcal{P}_{2|\xi_{[1]}}} \mathbb{E}_{\mathbf{p}_2} \Bigg[\min_{x_2 \in \mathcal{X}_2} g_2(x_2, \xi_2) + \max_{\mathbf{p}_3 \in \mathcal{P}_{3|\xi_{[2]}}} \mathbb{E}_{\mathbf{p}_3} \Big[\ldots + \max_{\mathbf{p}_T \in \mathcal{P}_{T|\xi_{[T-1]}}} \mathbb{E}_{\mathbf{p}_T} \Big[\min_{x_T \in \mathcal{X}_T} g_T(x_T, \xi_T) \Big] \ldots \Big] \Bigg],$$

where

$\mathcal{P}_{t|\xi_{[t-1]}}$ is the conditional ambiguity set for stage-t probability measure, conditioned on $\xi_{[t-1]}$.

How to Construct the Ambiguity Set (Multistage)?

- *Moment*-based sets: distributions with similar moments
  (Shapiro, 2012), (Xin et al., 2013), (Xin and Goldberg, 2015)
- *Distance*-based sets: sufficiently close distributions to a nominal distribution with respect to a distance
  - *Nested distance (Wasserstein metric):* (Pflug and Pichler, 2014), (Analui and Pflug, 2014), Duque and Morton (2019)
  - *Modified* χ^2 *distance:* (Philpott et al., 2017)
  - L_∞ *norm:* (Huang et al., 2017)
  - *General theory:* (Shapiro, 2016; 2017; 2018)

How to Construct the Ambiguity Set (Multistage)?

- *Moment*-based sets: distributions with similar moments

  (Shapiro, 2012), (Xin et al., 2013), (Xin and Goldberg, 2015)

- *Distance*-based sets: sufficiently close distributions to a nominal distribution with respect to a distance
  - *Nested distance (Wasserstein metric):* (Pflug and Pichler, 2014), (Analui and Pflug, 2014), Duque and Morton (2019)
  - *Modified* χ^2 *distance:* (Philpott et al., 2017)
  - L_∞ *norm:* (Huang et al., 2017)
  - *General theory:* (Shapiro, 2016; 2017; 2018)
  - **This talk: total variation distance**

Multistage DRSP with Total Variation Distance (DRSP-V)

At stage t, given $\xi_{[t-1]}$, instead of considering one ("nominal") distribution $\mathbf{q}_{t|\xi_{[t-1]}}$,

Consider all distributions $\mathbf{p}_t$ in

$$\mathcal{P}_{t|\xi_{[t-1]}} = \Bigg\{ \mathbf{p}_t : \mathrm{V}(\mathbf{p}_t, \mathbf{q}_{t|\xi_{[t-1]}}) := \frac{1}{2} \int_{\Xi_{t|\xi_{[t-1]}}} \left| \mathbf{p}_t - \mathbf{q}_{t|\xi_{[t-1]}} \right| \, d\nu \le \gamma_t, \quad \int_{\Xi_{t|\xi_{[t-1]}}} \mathbf{p}_t \, d\nu = 1, \quad \mathbf{p}_t \ge 0 \Bigg\},$$

where $\Xi_{t|\xi_{[t-1]}}$ is the sample space of stage t, given $\xi_{[t-1]}$.

Effective/Ineffective Scenarios in Multistage DRSP

How to extend those notions to the Multistage case?

Effective/Ineffective Scenarios in Multistage DRSP

How to extend those notions to the Multistage case?

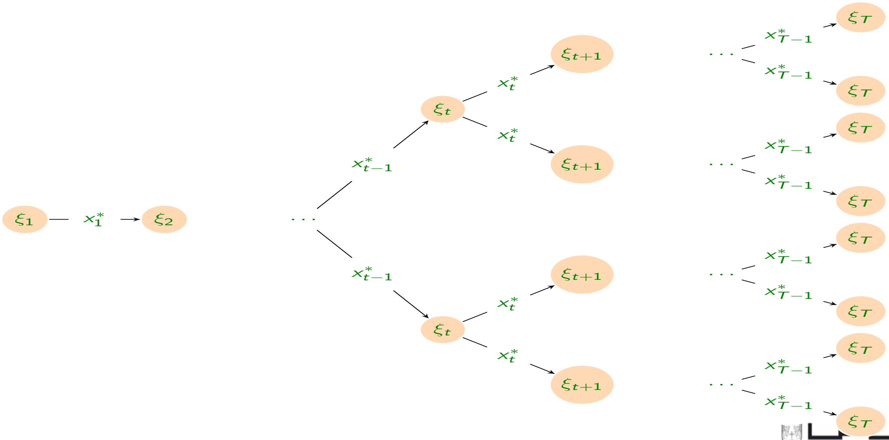
Recall the notation

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \max_{\mathbf{p}_2 \in \mathcal{P}_{2|\xi_{[1]}}} \mathbb{E}_{\mathbf{p}_2} \Bigg[\min_{x_2 \in \mathcal{X}_2} g_2(x_2, \xi_2) + \max_{\mathbf{p}_3 \in \mathcal{P}_{3|\xi_{[2]}}} \dots$$

$$\dots + \max_{\mathbf{p}_T \in \mathcal{P}_{T|\xi_{[T-1]}}} \mathbb{E}_{\mathbf{p}_T} \left[\min_{x_T \in \mathcal{X}_T} g_T(x_T, \xi_T) \right] \dots \Bigg] \Bigg]$$

Effective/Ineffective Scenarios in Multistage DRSP?

What is the effectiveness of a scenario (path)?



Effective Scenarios in Multistage DRSP:

Effectiveness of a Scenario Path

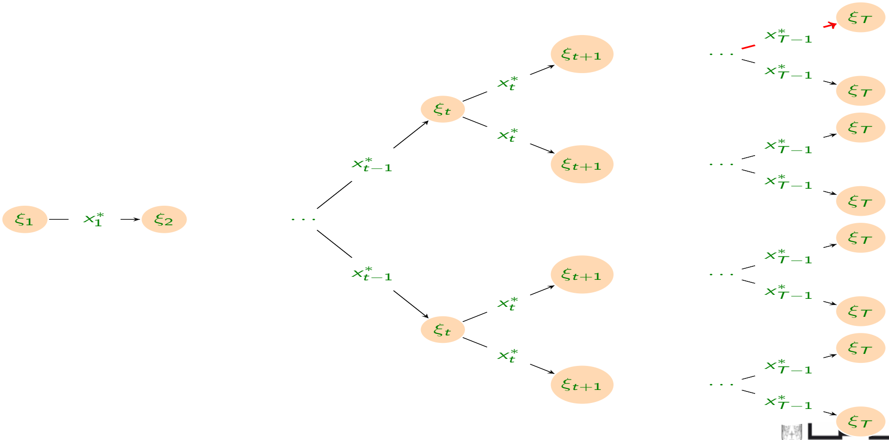
Definition (**Effective Scenario Path**)

A scenario path $\{\xi_t\}_{t=1}^{T}$ is called effective if by its "removal" the optimal value of the new problem is strictly smaller than the optimal value of multistage DRSP.

NOTE: Removing a scenario path is defined by forcing the probability of ξ_T to be zero.

Effective/Ineffective Scenarios in Multistage DRSP?

Removing (for instance) the uppermost scenario:



Effective/Ineffective Scenarios in Multistage DRSP?

Questions

- How to check the effectiveness of a scenario (path)?
- Can we have similar results to the 2-stage case?

Effective/Ineffective Scenarios in Multistage DRSP?

Questions

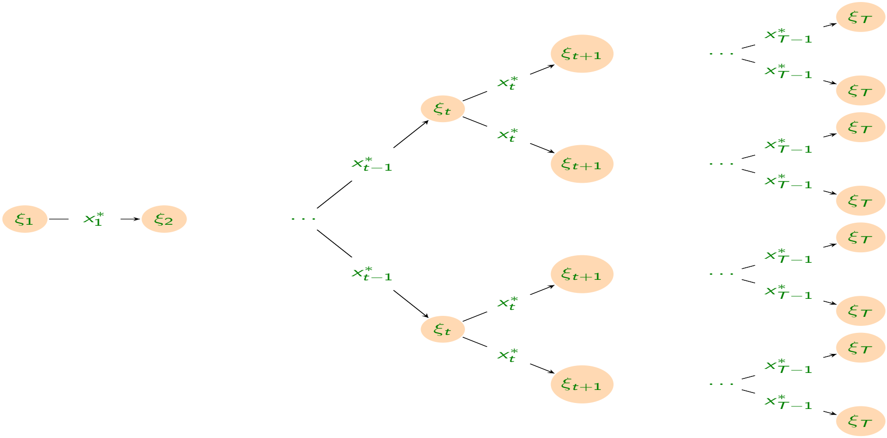
- How to check the effectiveness of a scenario (path)?
- Can we have similar results to the 2-stage case?

Main Idea

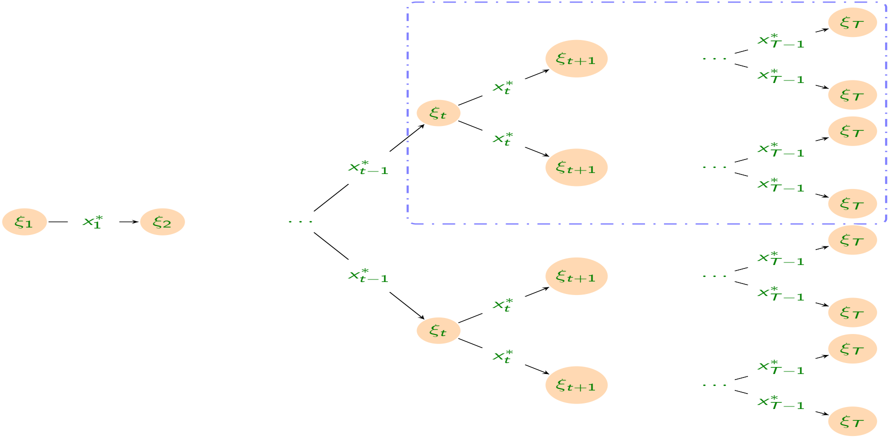
- Look at realizations conditioned on their history of decisions and stochastic process

$\rightarrow$ At an optimal policy x^*, if we look at stage t, **given** $x^*_{[t-1]}$ and $\xi_{[t]}$, previous definitions on effective/ineffective scenarios hold **conditionally**.

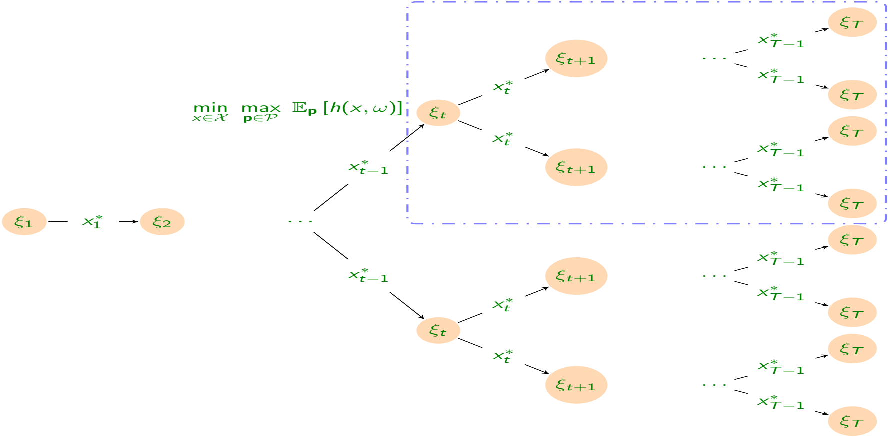
Effective/Ineffective Scenarios in Multistage DRSP?



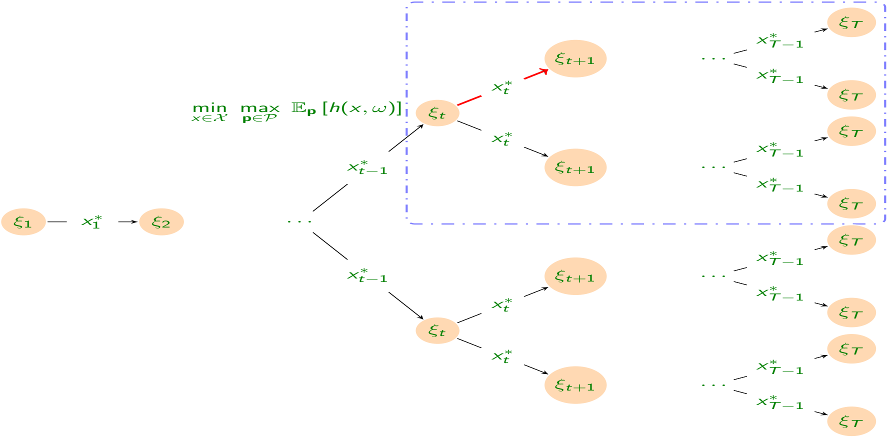
Effective/Ineffective Scenarios in Multistage DRSP?



Effective/Ineffective Scenarios in Multistage DRSP?



Effective/Ineffective Scenarios in Multistage DRSP?



Effective Scenarios in Multistage DRSP:

Conditional Effectiveness

Definition (**Conditionally Effective Realization**)

At an optimal policy $x^* := [x_1^*, \ldots, x_T^*]$, a realization of ξ_{t+1} in stage $t+1$ is called conditionally effective, given $x_{[t-1]}^*$ and $\xi_{[t]}$, if by its removal the optimal stage-t cost function (immediate cost + cost-to-go function) of the new problem is strictly smaller than the optimal value of the original stage-t problem in multistage DRSP.

Use Conditional Effectiveness of Realizations in Multistage DRSP-V

AIM: Propose easy-to-check conditions

Theorem [Conditionally Multistage ← Two-stage]

The identifying conditions for effective/ineffective scenarios in two-stage DRSP-V are valid conditions to identify conditionally effective/ineffective scenarios in multistage DRSP-V.

Effectiveness of Scenario Paths in Multistage DRSP-V

Consider a scenario path $\{\xi_t\}_{t=1}^T$.

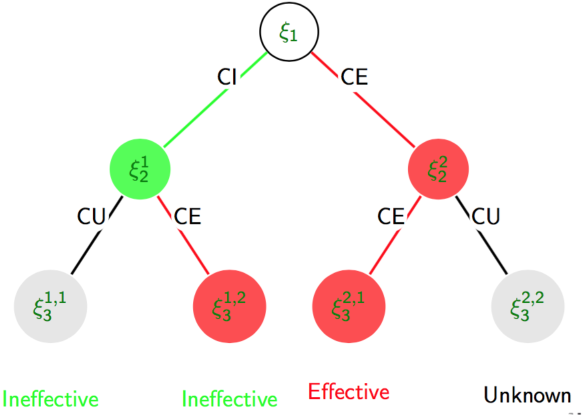
Theorem

If ξ_t is conditionally effective by the identifying conditions, for all $t = 1, \ldots, T$, then, the scenario path $\{\xi_t\}_{t=1}^T$ is effective.

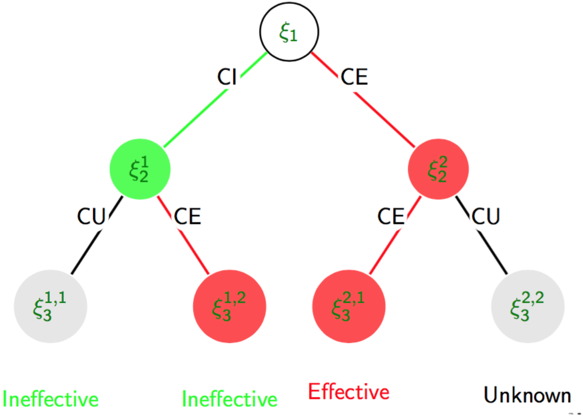
Theorem

If ξ_T is **not trivially** conditionally effective (i.e., too large nominal conditional probability) and there exists t, $t = 1, \ldots, T$, such that ξ_t is conditionally ineffective by the identifying conditions, then, the scenario path $\{\xi_t\}_{t=1}^T$ is ineffective.

Identifying Conditions for Effectiveness of Scenario Paths



Identifying Conditions for Effectiveness of Scenario Paths



NOTE: The effective scenarios are not necessarily the ones with highest cost! (reason: there is no rectangularity).

Solution approach

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \max_{\mathbf{p}_2 \in \mathcal{P}_{2|\xi_{[1]}}} \mathbb{E}_{\mathbf{p}_2} \left[\min_{x_2 \in \mathcal{X}_2} g_2(x_2, \xi_2) + \ldots + \max_{\mathbf{p}_T \in \mathcal{P}_{T|\xi_{[T-1]}}} \mathbb{E}_{\mathbf{p}_T} \left[\min_{x_T \in \mathcal{X}_T} g_T(x_T, \xi_T) \right] \right]$$

Solution approach

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \max_{\mathbf{p}_2 \in \mathcal{P}_{2|\xi_{[1]}}} \mathbb{E}_{\mathbf{p}_2} \underbrace{\left[\min_{x_2 \in \mathcal{X}_2} g_2(x_2, \xi_2) + \ldots + \max_{\mathbf{p}_T \in \mathcal{P}_{T|\xi_{[T-1]}}} \mathbb{E}_{\mathbf{p}_T} \left[\min_{x_T \in \mathcal{X}_T} g_T(x_T, \xi_T) \right] \right]}_{Q_2(x_1, \xi_{[2]})}$$

First-stage cost function

$$\min_{x_1 \in \mathcal{X}_1} \; g_1(x_1, \xi_1) + \max_{\mathbf{p}_2 \in \mathcal{P}_{2|\xi_{[1]}}} \mathbb{E}_{\mathbf{p}_2} \left[Q_2(x_1, \xi_{[2]}) \right]$$

Solution approach

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \max_{\mathbf{p}_2 \in \mathcal{P}_{2|\xi_{[1]}}} \mathbb{E}_{\mathbf{p}_2} \underbrace{\left[\min_{x_2 \in \mathcal{X}_2} g_2(x_2, \xi_2) + \underbrace{\ldots + \max_{\mathbf{p}_T \in \mathcal{P}_{T|\xi_{[T-1]}}} \mathbb{E}_{\mathbf{p}_T} \underbrace{\left[\min_{x_T \in \mathcal{X}_T} g_T(x_T, \xi_T) \right]}_{Q_T(x_{T-1}, \xi_{[T]})}}_{Q_3(x_2, \xi_{[3]})} \right]}_{Q_2(x_1, \xi_{[2]})}$$

First-stage cost function

$$\min_{x_1 \in \mathcal{X}_1} g_1(x_1, \xi_1) + \max_{\mathbf{p}_2 \in \mathcal{P}_{2|\xi_{[1]}}} \mathbb{E}_{\mathbf{p}_2} \left[Q_2(x_1, \xi_{[2]}) \right]$$

stage-t cost function

$$Q_t(x_{t-1}, \xi_{[t]}) := \min_{x_t \in \mathcal{X}_t} g_t(x_t, \xi_t) + \max_{\mathbf{p}_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \mathbb{E}_{\mathbf{p}_{t+1}} \left[Q_{t+1}(x_t, \xi_{[t+1]}) \right]$$

A Cutting Plane Approach

stage-t cost function

$$Q_t(x_{t-1}, \xi_{[t]}) = \min_{x_t \in \mathcal{X}_t} g_t(x_t, \xi_t) + \max_{\mathbf{p}_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \mathbb{E}_{\mathbf{p}_{t+1}} \left[Q_{t+1}(x_t, \xi_{[t+1]}) \right]$$

A Cutting Plane Approach

stage-t cost function

$$Q_t(x_{t-1}, \xi_{[t]}) = \min_{x_t \in \mathcal{X}_t} \; g_t(x_t, \xi_t) + \alpha_t$$

$$\text{s.t.} \quad \alpha_t \geq \max_{\mathbf{p}_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \mathbb{E}_{\mathbf{p}_{t+1}} \left[Q_{t+1}(x_t, \xi_{[t+1]}) \right]$$

A Cutting Plane Approach

stage-t cost function

$$Q_t(x_{t-1}, \xi_{[t]}) = \min_{x_t \in \mathcal{X}_t} \; g_t(x_t, \xi_t) + \alpha_t$$
$$\text{s.t.} \quad \alpha_t \geq \mathbb{E}_{\mathbf{p}_{t+1}} \left[Q_{t+1}(x_t, \xi_{[t+1]}) \right], \quad \mathbf{p}_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}$$

A Cutting Plane Approach

stage-t cost function

$$Q_t(x_{t-1}, \xi_{[t]}) = \min_{x_t \in \mathcal{X}_t} \; g_t(x_t, \xi_t) + \alpha_t$$
$$\text{s.t.} \quad \alpha_t \geq \mathbb{E}_{\mathbf{p}_{t+1}} \left[Q_{t+1}(x_t, \xi_{[t+1]}) \right], \quad \mathbf{p}_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}$$

For multistage DRSP-V,

- $\mathcal{P}_{t+1|\xi_{[t]}}$ is a polyhedron $\Longrightarrow$ Finite convergence

A Cutting Plane Approach

stage-t cost function

$$Q_t(x_{t-1}, \xi_{[t]}) = \min_{x_t \in \mathcal{X}_t} \; g_t(x_t, \xi_t) + \alpha_t$$
$$\text{s.t.} \quad \alpha_t \geq \mathbb{E}_{\mathbf{p}_{t+1}}\left[Q_{t+1}(x_t, \xi_{[t+1]})\right], \quad \mathbf{p}_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}$$

For multistage DRSP-V,

- $\mathcal{P}_{t+1|\xi_{[t]}}$ is a polyhedron $\Longrightarrow$ Finite convergence

This idea can be applied to any polyhedral ambiguity set, with finite convergence guaranteed

How to Generate Distributional Cuts?

Distribution Separation Problem

For a fixed $x_t \in \mathcal{X}_t$, solve

$$\max_{p_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \mathbb{E}_{\mathbf{p}_{t+1}} \left[Q_{t+1}(x_t, \xi_{[t+1]}) \right]$$

How to Generate Distributional Cuts?

Distribution Separation Problem

For a fixed $x_t \in \mathcal{X}_t$, solve

$$\max_{p_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \int_{\Xi_{t+1|\xi_{[t]}}} \mathbf{p}_{t+1} Q_{t+1}(x_t, \cdot)\, d\,\nu$$

How to Generate Distributional Cuts?

Distribution Separation Problem

For a fixed $x_t \in \mathcal{X}_t$, solve

$$\max_{p_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \int_{\Xi_{t+1|\xi_{[t]}}} \mathbf{p}_{t+1} Q_{t+1}(x_t, \cdot)\, d\,\nu$$

For multistage DRSP-V,

- $\mathcal{P}_{t+1|\xi_{[t]}}$ is a polytope $\Longrightarrow$ Optimum is obtained at an extreme point

How to Generate Distributional Cuts?

Distribution Separation Problem

For a fixed $x_t \in \mathcal{X}_t$, solve

$$\max_{p_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \int_{\Xi_{t+1|\xi_{[t]}}} \mathbf{p}_{t+1} Q_{t+1}(x_t, \cdot)\, d\,\nu$$

For multistage DRSP-V,

- $\mathcal{P}_{t+1|\xi_{[t]}}$ is a polytope $\Longrightarrow$ Optimum is obtained at an extreme point

Challenge

- We do not have $Q_{t+1}(x_t, \xi_{[t+1]})$

How to Generate Distributional Cuts?

Distribution Separation Problem

For a fixed $x_t \in \mathcal{X}_t$, solve

$$\max_{p_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \int_{\Xi_{t+1|\xi_{[t]}}} \mathbf{p}_{t+1} \bar{Q}_{t+1}(x_t, \cdot)\, d\nu$$

For multistage DRSP-V,

- $\mathcal{P}_{t+1|\xi_{[t]}}$ is a polytope $\Longrightarrow$ Optimum is obtained at an extreme point

Challenge

- We do not have $Q_{t+1}(x_t, \xi_{[t+1]})$

But...

- We can use an upper bound $\bar{Q}_{t+1}(x_t, \xi_{[t+1]})$

Primal Decomposition Algorithm

Main Idea

- Combine Nested L-shaped method and Distribution Separation problem

Forward Pass

- Obtain $x = [x_1, \ldots, x_T]$
- Use upper bound on $Q_{t+1}(x_t, \xi_{[t+1]})$, $t = T-1, \ldots, 1$ to obtain $\mathbf{p} = [p_T, \ldots, p_2]$

Backward Pass

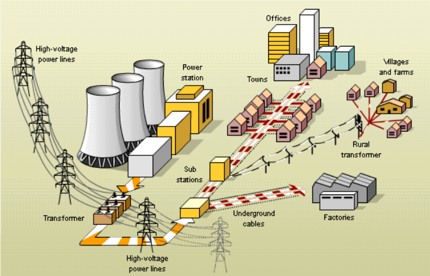
- Refine outer approximations on $Q_{t+1}(x_t, \xi_{[t+1]})$ and $\max_{\mathbf{p}_{t+1} \in \mathcal{P}_{t+1|\xi_{[t]}}} \mathbb{E}_{\mathbf{p}_{t+1}} \left[Q_{t+1}(x_t, \xi_{[t+1]}) \right]$

Outline

Two-stage: capacity expansion

We studied a DRSO version of PGP2, described in Higle and Sen (1994).

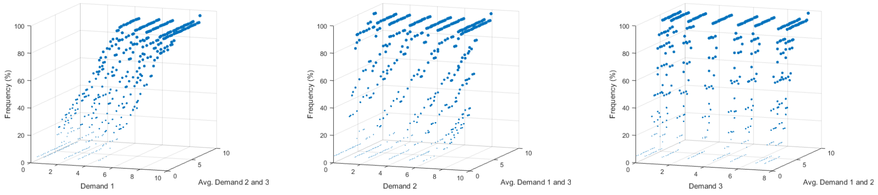
Distribution and Use of Electricity



Power network with uncertain demand:

- First-stage decisions: What capacities to install at the generators?
- Second-stage decisions: Purchase additional capacities to fulfill unmet demands

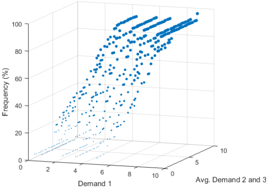
Results, PGP2



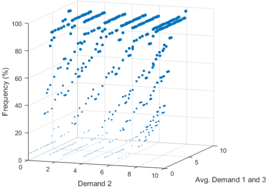
(a) (b) (c)

- The graphs display the effective scenarios of one source of demand vs. the other two sources, for each level of robustness.

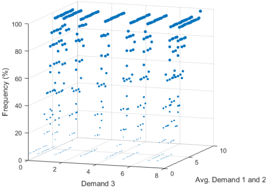
Results, PGP2



(a)



(b)



(c)

- The graphs display the effective scenarios of one source of demand vs. the other two sources, for each level of robustness.
- We see that source 1 is the critical one (even high values of sources 2 and 3 do not necessarily lead to critical scenarios).

Multistage: water resources allocation (Zhang, Rahimian, Bayraksan, 2016)

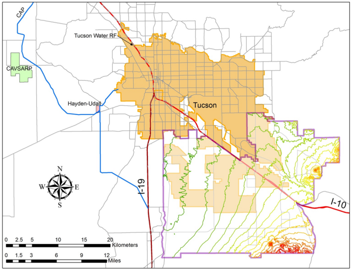


Figure: The southeastern region of Tucson, AZ.

How to best allocate Colorado River water among different users while meeting uncertain water demand and not exceeding uncertain water supply over the next 16 years?

Multistage: water resources allocation (Zhang, Rahimian, Bayraksan, 2016)

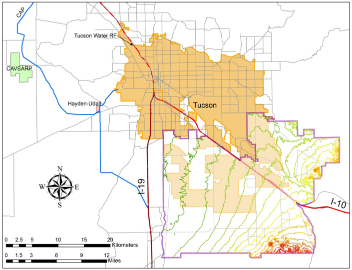


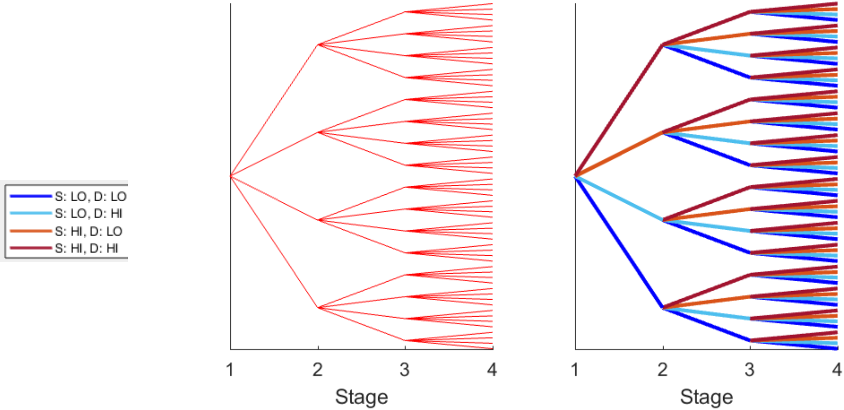
Figure: The southeastern region of Tucson, AZ.

How to best allocate Colorado River water among different users while meeting uncertain water demand and not exceeding uncertain water supply over the next 16 years?

The problem shown here has 4 Stages, $4^4 = 64$ scenarios of (demand, supply).

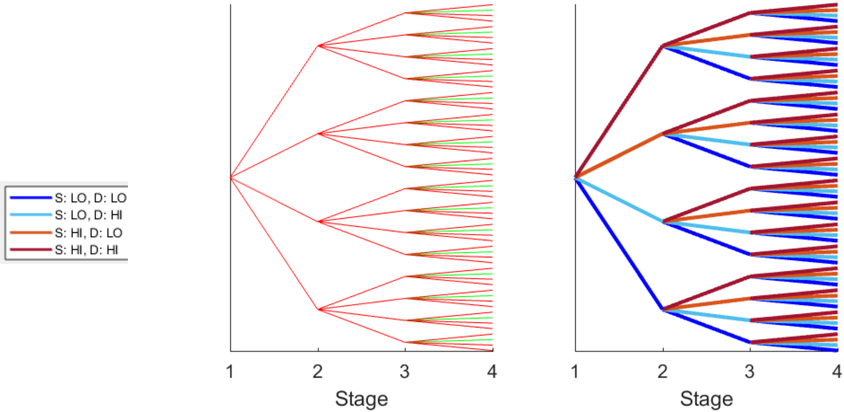
Effective/ineffective scenarios

$$\gamma = 0.0$$



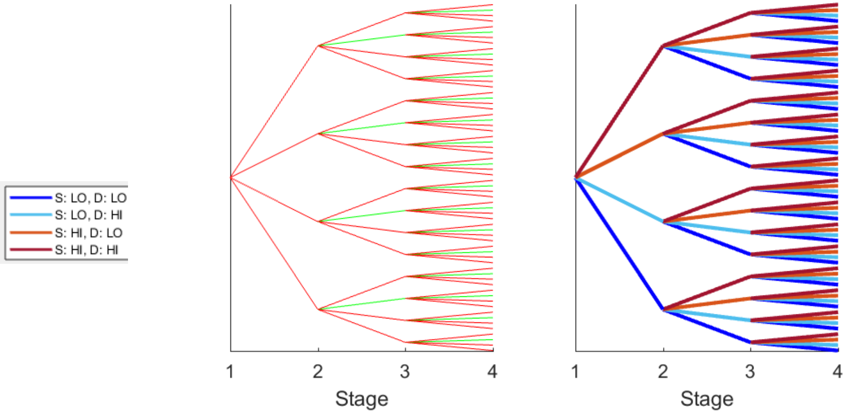
Effective/ineffective scenarios

$\gamma = 0.35$



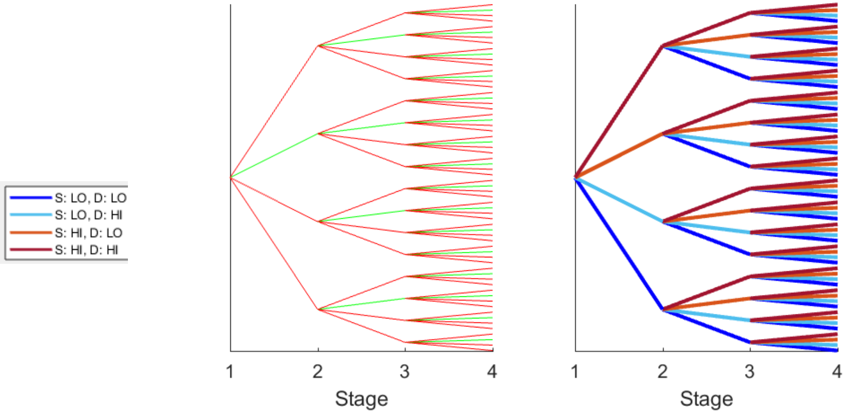
Effective/ineffective scenarios

$\gamma = 0.4$



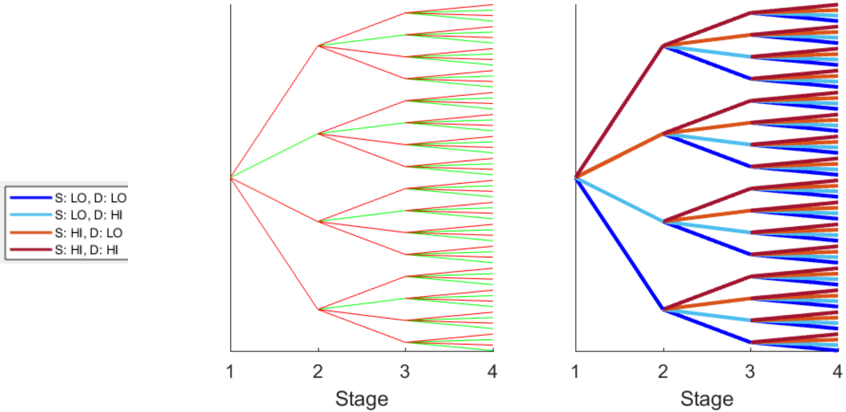
Effective/ineffective scenarios

$\gamma = 0.45$



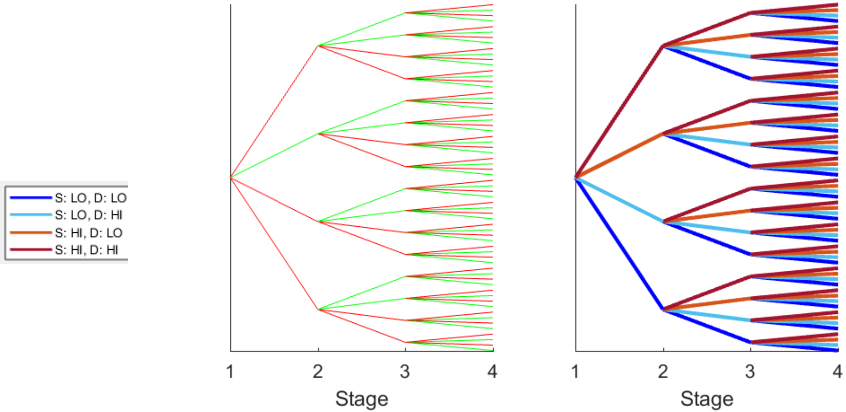
Effective/ineffective scenarios

$$\gamma = 0.5$$



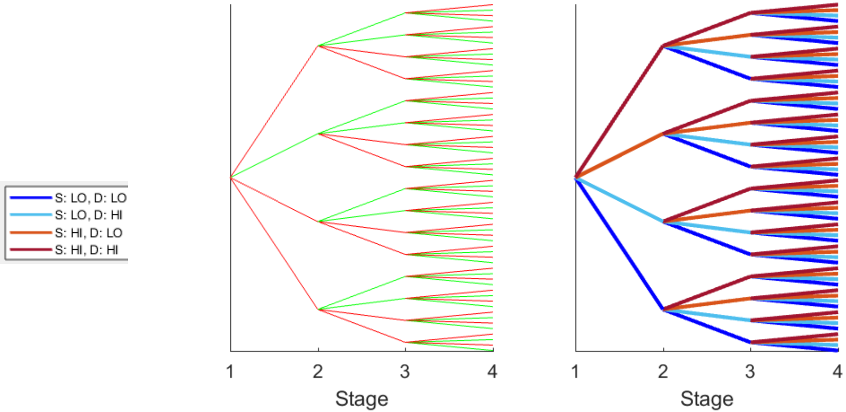
Effective/ineffective scenarios

$\gamma = 0.55$



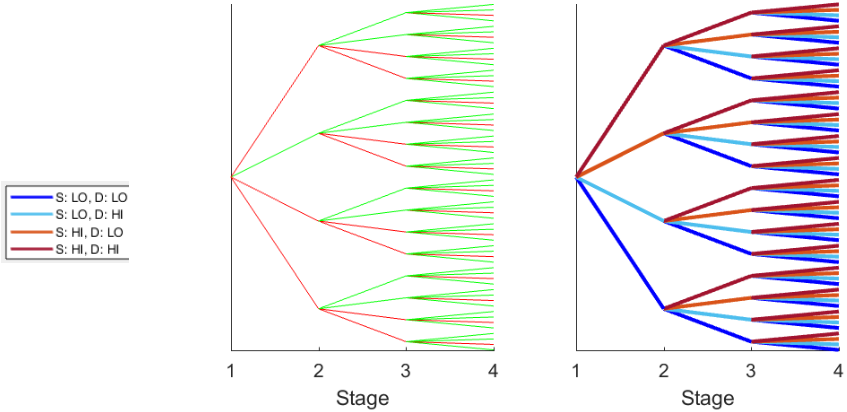
Effective/ineffective scenarios

$$\gamma = 0.6$$



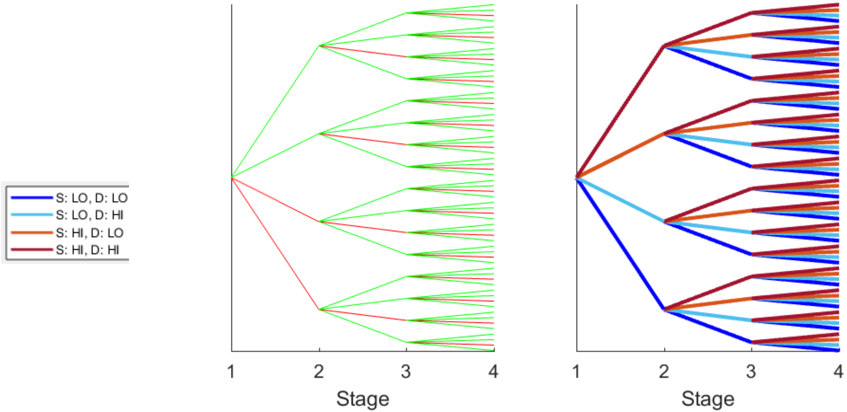
Effective/ineffective scenarios

$$\gamma = 0.65$$



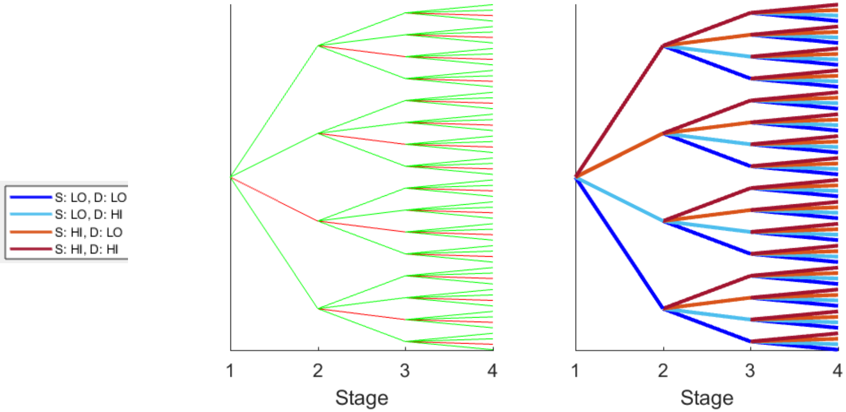
Effective/ineffective scenarios

$$\gamma = 0.7$$

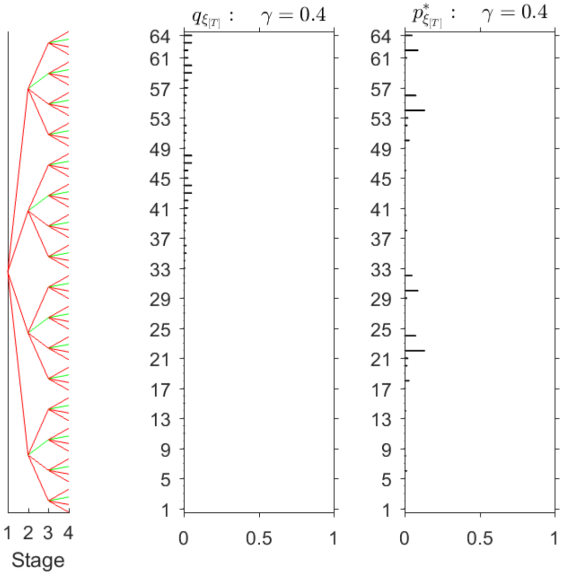


Effective/ineffective scenarios

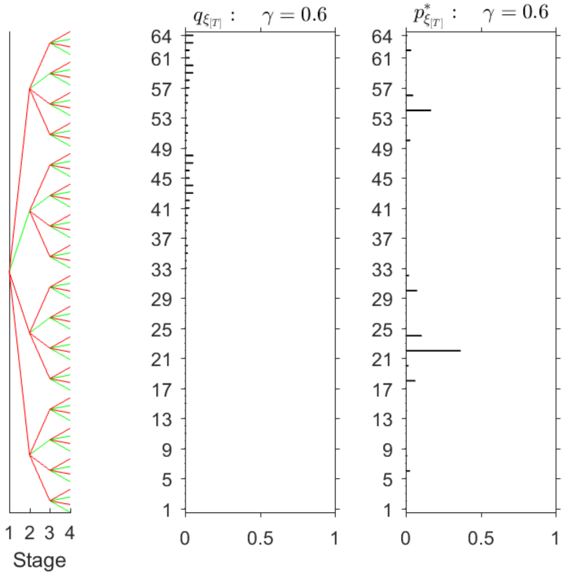
$$\gamma = 1.0$$



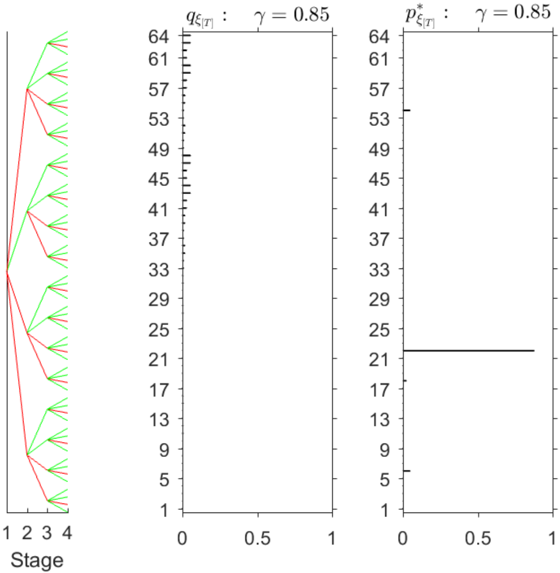
Probabilities



Probabilities



Probabilities



Outline

Conclusion and Future Research

- Effective scenarios can provide insight into the underlying uncertainties of the problems, both for managerial as well as for computational purposes.

Conclusion and Future Research

- Effective scenarios can provide insight into the underlying uncertainties of the problems, both for managerial as well as for computational purposes.
- Concepts from the two-stage case can be extended to the multistage case via (time-)decomposition.

Conclusion and Future Research

- Effective scenarios can provide insight into the underlying uncertainties of the problems, both for managerial as well as for computational purposes.
- Concepts from the two-stage case can be extended to the multistage case via (time-)decomposition.
- Future work: Extension to other distances, connection with strictly monotone risk measures.

Acknowledgements and References

Gratefully acknowledge support of FONDECYT 1171145, Chile.

References:

- Rahimian, H., G. Bayraksan, and T. Homem-de-Mello, "Identifying Effective Scenarios in Distributionally Robust Stochastic Programs with Total Variation Distance," *Mathematical Programming* 173(1-2): 393–430, 2019.
- Rahimian, H., G. Bayraksan, and T. Homem-de-Mello, "Distributionally Robust Newsvendor Problems with Variation Distance," *European Journal of Operational Research* 279, 854–868, 2019.
- Rahimian, H., G. Bayraksan, and T. Homem-de-Mello, "Effective Scenarios in Multistage Distributionally Robust Stochastic Programs with Total Variation Distance," *Working paper*.

Thank you!

(tito.hmello@uai.cl)

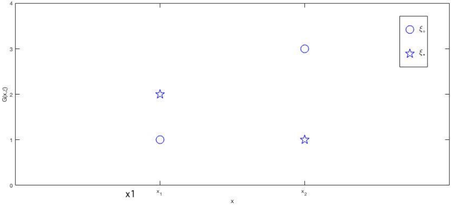
The role of convexity

While the above results appear intuitive, they rely strongly on convexity of the objective function.

The role of convexity

While the above results appear intuitive, they rely strongly on convexity of the objective function.

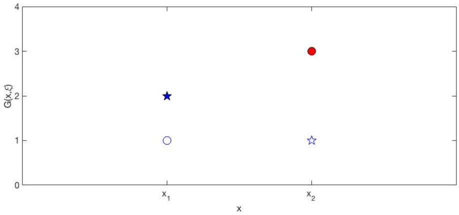
Suppose $\Omega = \{\omega_\circ, \omega_\star\}$, and the function h defined on $X = \{x_1, x_2\}$:



The role of convexity

While the above results appear intuitive, they rely strongly on convexity of the objective function.

Suppose $\Omega = \{\omega_\circ, \omega_\star\}$, and the function h defined on $X = \{x_1, x_2\}$:

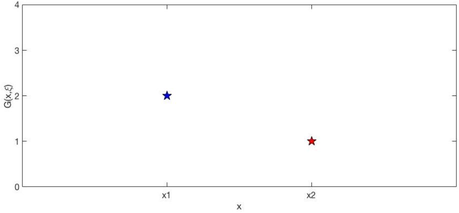


If we want to protect ourselves against the highest cost, then the optimal solution is $\bar{x} = x_1$, with opt. value=2 (scen. $\omega_\star$). Moreover, $p^*(\omega_\star) = 1$.

The role of convexity

While the above results appear intuitive, they rely strongly on convexity of the objective function.

Suppose $\Omega = \{\omega_\circ, \omega_\star\}$, and the function h defined on $X = \{x_1, x_2\}$:



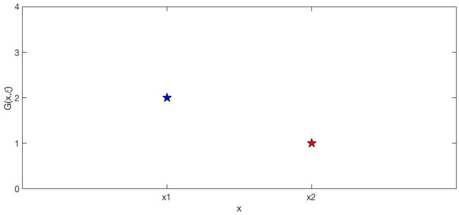
If we want to protect ourselves against the highest cost, then the optimal solution is $\bar{x} = x_1$, with opt. value=2 (scen. $\omega_\star$). Moreover, $p^*(\omega_\star) = 1$.

- But if we remove $\omega_\circ$, then the solution becomes x_2, and the optimal value is 1.

The role of convexity

While the above results appear intuitive, they rely strongly on convexity of the objective function.

Suppose $\Omega = \{\omega_\circ, \omega_\star\}$, and the function h defined on $X = \{x_1, x_2\}$:



If we want to protect ourselves against the highest cost, then the optimal solution is $\bar{x} = x_1$, with opt. value=2 (scen. $\omega_\star$). Moreover, $p^*(\omega_\star) = 1$.

- But if we remove $\omega_\circ$, then the solution becomes x_2, and the optimal value is 1.
- So, both scenarios are effective — the problem is lack of convexity.