Tutorial on stochastic programming: from two-stage to multi-stage risk averse stochastic programming

PGMO - LASON

Bernardo Pagnoncelli & Tito Homem-de-Mello

Escuela de Negocios
Universidad Adolfo Ibáñez
Santiago, Chile

January 31st, 2013

Two stage stochastic programming

- Started in the 50's. Important early works: Dantzig ('55), Beale ('55), Walkup and Wets ('67).
- Increased interest in the last 20 years due to computational advances.
- Mature field: Kall and Wallace '94, Birge and Louveaux '97, Shapiro et al. '09.
- Applications: Finance, Energy, Transportation, Production Planning, Telecommunications, Forestry, ...

Two stage stochastic programming

- Started in the 50's. Important early works: Dantzig ('55), Beale ('55), Walkup and Wets ('67).
- Increased interest in the last 20 years due to computational advances.
- Mature field: Kall and Wallace '94, Birge and Louveaux '97, Shapiro et al. '09.
- Applications: Finance, Energy, Transportation, Production Planning, Telecommunications, Forestry, ...

- Started in the 50's. Important early works: Dantzig ('55), Beale ('55), Walkup and Wets ('67).
- Increased interest in the last 20 years due to computational advances.
- Mature field: Kall and Wallace '94, Birge and Louveaux '97, Shapiro et al. '09.
- Applications: Finance, Energy, Transportation, Production Planning, Telecommunications, Forestry, ...

Two stage stochastic programming

- Started in the 50's. Important early works: Dantzig ('55), Beale ('55), Walkup and Wets ('67).
- Increased interest in the last 20 years due to computational advances.
- Mature field: Kall and Wallace '94, Birge and Louveaux '97, Shapiro et al. '09.
- Applications: Finance, Energy, Transportation, Production Planning, Telecommunications, Forestry, ...

Two stage stochastic programming

decision x $\rightsquigarrow$ realization ξ $\rightsquigarrow$ Recourse action y.

$$\min_{x \in X} \{cx + \mathbb{E}[Q(x,\xi)]\},$$

where

$$Q(x,\xi) = \min_{y \in Y} \{qy | Tx + Wy \geq h\}.$$

An example: the newsvendor model

- A newsvendor must decide how many newspapers x he/she will buy at price c.
- The sold quantity is y and the selling price is r.
- Unsold newspapers (w) can be salvaged at value v.
- The demand ξ is a nonnegative random variable with cumulative distribution F.
- The goal is minimize costs (or maximize profits).

An example: the newsvendor model

- A newsvendor must decide how many newspapers x he/she will buy at price c.
- The sold quantity is y and the selling price is r.
- Unsold newspapers (w) can be salvaged at value v.
- The demand ξ is a nonnegative random variable with cumulative distribution F.
- The goal is minimize costs (or maximize profits).

An example: the newsvendor model

- A newsvendor must decide how many newspapers x he/she will buy at price c.
- The sold quantity is y and the selling price is r.
- Unsold newspapers (w) can be salvaged at value v.
- The demand ξ is a nonnegative random variable with cumulative distribution F.
- The goal is minimize costs (or maximize profits).

An example: the newsvendor model

- A newsvendor must decide how many newspapers x he/she will buy at price c.
- The sold quantity is y and the selling price is r.
- Unsold newspapers (w) can be salvaged at value v.
- The demand ξ is a nonnegative random variable with cumulative distribution F.
- The goal is minimize costs (or maximize profits).

An example: the newsvendor model

- A newsvendor must decide how many newspapers x he/she will buy at price c.
- The sold quantity is y and the selling price is r.
- Unsold newspapers (w) can be salvaged at value v.
- The demand ξ is a nonnegative random variable with cumulative distribution F.
- The goal is minimize costs (or maximize profits).

Formulation

$$\min_{x \geq 0} \left\{ cx + \mathbb{E}\left[Q(x, \xi) \right] \right\},$$

$$\begin{aligned} Q(x,\xi) = \text{minimize} \quad & -ry - vw \\ \text{subject to} \quad & y \leq \xi, \\ & y + w \leq x, \\ & y, w \geq 0. \end{aligned}$$

The exact solution

- The exact solution is given by

$$F^{-1}\left(\frac{r-c}{r-v}\right)$$

where $F^{-1}(\cdot)$ is the (generalized) inverse distribution function of ξ.

	$U^d[1,10]$	Exp(10)
x^*	[2,3]	2.23
v^*	-1.5	-1.07

- The exact solution is given by

$$F^{-1}\left(\frac{r-c}{r-v}\right)$$

where $F^{-1}(\cdot)$ is the (generalized) inverse distribution function of ξ.

	$U^d[1,10]$	Exp(10)
x^*	[2,3]	2.23
v^*	-1.5	-1.07

Risk aversion

- Minimize the expected cost is just one possible criterion.
- What if bad outcomes are extremely undesirable?
- In the newsvendor problem, what if staying with excess inventory is catastrophic?
- In the uniform case, buying 3 newspapers is optimal but there is a 20% chance of overstocking.

Risk aversion

- Minimize the expected cost is just one possible criterion.
- What if bad outcomes are extremely undesirable?
- In the newsvendor problem, what if staying with excess inventory is catastrophic?
- In the uniform case, buying 3 newspapers is optimal but there is a 20% chance of overstocking.

Risk aversion

- Minimize the expected cost is just one possible criterion.
- What if bad outcomes are extremely undesirable?
- In the newsvendor problem, what if staying with excess inventory is catastrophic?
- In the uniform case, buying 3 newspapers is optimal but there is a 20% chance of overstocking.

Risk aversion

- Minimize the expected cost is just one possible criterion.
- What if bad outcomes are extremely undesirable?
- In the newsvendor problem, what if staying with excess inventory is catastrophic?
- In the uniform case, buying 3 newspapers is optimal but there is a 20% chance of overstocking.

Risk measures

- A risk measure is a function from a space of random variables into the real numbers.
- A risk measure should capture dispersion and protect the decision maker against extreme losses.
- Classical ones: variance and the Value-at-Risk (VaR).
- Coherent risk measures (Artzner et al. '99) and the Conditional Value-at-Risk (CVaR) ('00) (Rockafellar and Uryasev).

Risk measures

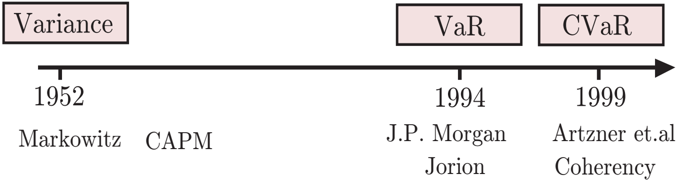
- A risk measure is a function from a space of random variables into the real numbers.
- A risk measure should capture dispersion and protect the decision maker against extreme losses.
- Classical ones: variance and the Value-at-Risk (VaR).
- Coherent risk measures (Artzner et al. '99) and the Conditional Value-at-Risk (CVaR) ('00) (Rockafellar and Uryasev).

Risk measures

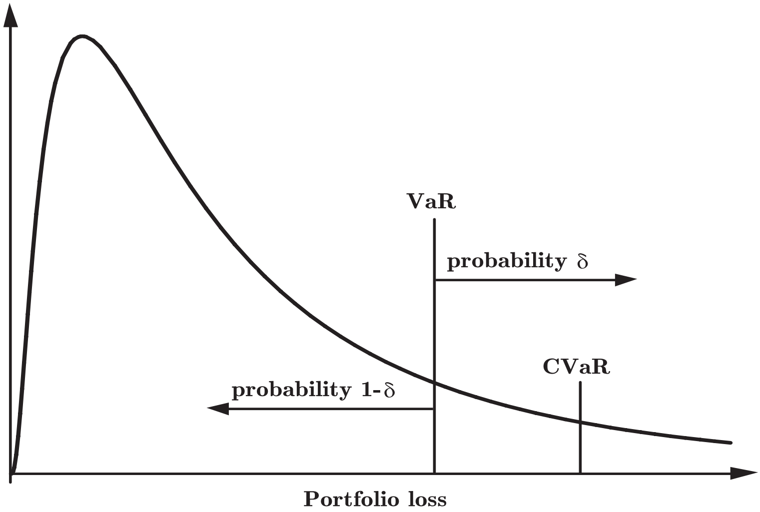
- A risk measure is a function from a space of random variables into the real numbers.
- A risk measure should capture dispersion and protect the decision maker against extreme losses.
- Classical ones: variance and the Value-at-Risk (VaR).
- Coherent risk measures (Artzner et al. '99) and the Conditional Value-at-Risk (CVaR) ('00) (Rockafellar and Uryasev).

- A risk measure is a function from a space of random variables into the real numbers.
- A risk measure should capture dispersion and protect the decision maker against extreme losses.
- Classical ones: variance and the Value-at-Risk (VaR).
- Coherent risk measures (Artzner et al. '99) and the Conditional Value-at-Risk (CVaR) ('00) (Rockafellar and Uryasev).

Risk measures



How bad is bad? Conditional Value-at-Risk



- The Value-at-Risk:

$$\text{VaR}_\alpha[X] = \inf\{x : \mathbb{P}(X \leq x) \geq 1 - \alpha\}, \quad \alpha \in (0,1).$$

- Conditional Value-at-Risk, Average Value-at-Risk, Expected Tail Loss and Expected Shortfall are all the same thing!

- Formally, we define $\text{CVaR}_\alpha[X] = \inf_{t \in \mathbb{R}} \left\{ t + \frac{1}{1-\alpha} \mathbb{E}\left[X - t\right]_+ \right\} =$ (Cont. case) $= \mathbb{E}\left[X \mid X > \text{VaR}_\alpha\right]$.

- Average Value-at-Risk (AVaR)[X] $= \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_\gamma[X] d\gamma$.

- The Value-at-Risk:

$$\text{VaR}_\alpha[X] = \inf\{x : \mathbb{P}(X \leq x) \geq 1 - \alpha\}, \quad \alpha \in (0, 1).$$

- Conditional Value-at-Risk, Average Value-at-Risk, Expected Tail Loss and Expected Shortfall are all the same thing!

- Formally, we define $\text{CVaR}_\alpha[X] = \inf_{t \in \mathbb{R}} \left\{ t + \frac{1}{1-\alpha} \mathbb{E}\left[X - t\right]_+ \right\} =$ (Cont. case) $= \mathbb{E}\left[X \mid X > \text{VaR}_\alpha\right]$.

- Average Value-at-Risk (AVaR)[X] $= \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_\gamma[X] d\gamma$.

CVaR definiton

- The Value-at-Risk:

$$\text{VaR}_\alpha[X] = \inf\{x : \mathbb{P}(X \leq x) \geq 1 - \alpha\}, \quad \alpha \in (0,1).$$

- Conditional Value-at-Risk, Average Value-at-Risk, Expected Tail Loss and Expected Shortfall are all the same thing!

- Formally, we define $\text{CVaR}_\alpha[X] = \inf_{t \in \mathbb{R}} \left\{ t + \frac{1}{1-\alpha} \mathbb{E}\left[X - t\right]_+ \right\} =$ (Cont. case) $= \mathbb{E}\left[X \mid X > \text{VaR}_\alpha\right]$.

- Average Value-at-Risk (AVaR)[X] $= \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_\gamma[X] d\gamma$.

CVaR definiton

- The Value-at-Risk:

$$\text{VaR}_\alpha[X] = \inf\{x : \mathbb{P}(X \leq x) \geq 1-\alpha\}, \quad \alpha \in (0,1).$$

- Conditional Value-at-Risk, Average Value-at-Risk, Expected Tail Loss and Expected Shortfall are all the same thing!

- Formally, we define $\text{CVaR}_\alpha[X] = \inf_{t\in\mathbb{R}} \left\{ t + \frac{1}{1-\alpha}\mathbb{E}\left[X-t\right]_+ \right\} =$ (Cont. case) $= \mathbb{E}\left[X \mid X > \text{VaR}_\alpha\right]$.

- Average Value-at-Risk (AVaR)[X] $= \frac{1}{1-\alpha}\int_\alpha^1 \text{VaR}_\gamma[X]d\gamma$.

1) $\rho(X + c) = \rho(X) + c$.

2) $X \le Y \Rightarrow \rho(X) \le \rho(Y)$.

3) $\rho(\lambda X) = \lambda \rho(X)$ for $\lambda \ge 0$.

4) $\rho(X + Y) \le \rho(X) + \rho(Y)$.

A risk measure that satisfies axioms 1) – 4) is called *coherent*. Other example (Mean Deviation Risk of order p):

$$\rho(X) := \mathbb{E}[X] + c(\mathbb{E}[|X - \mathbb{E}[X]|^p])^{1/p}, c > 0, p \in [0, \infty).$$

1) $\rho(X + c) = \rho(X) + c$.

2) $X \leq Y \Rightarrow \rho(X) \leq \rho(Y)$.

3) $\rho(\lambda X) = \lambda \rho(X)$ for $\lambda \geq 0$.

4) $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

A risk measure that satisfies axioms 1) – 4) is called *coherent*. Other example (Mean Deviation Risk of order p):

$$\rho(X) := \mathbb{E}[X] + c(\mathbb{E}[|X - \mathbb{E}[X]|^p])^{1/p}, c > 0, p \in [0, \infty).$$

Coherent risk measures

1) $\rho(X + c) = \rho(X) + c$.

2) $X \leq Y \Rightarrow \rho(X) \leq \rho(Y)$.

3) $\rho(\lambda X) = \lambda \rho(X)$ for $\lambda \geq 0$.

4) $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

A risk measure that satisfies axioms 1) – 4) is called *coherent*. Other example (Mean Deviation Risk of order p):

$$\rho(X) := \mathbb{E}[X] + c(\mathbb{E}[|X - \mathbb{E}[X]|^p])^{1/p}, c > 0, p \in [0, \infty).$$

1) $\rho(X + c) = \rho(X) + c$.

2) $X \le Y \Rightarrow \rho(X) \le \rho(Y)$.

3) $\rho(\lambda X) = \lambda\rho(X)$ for $\lambda \ge 0$.

4) $\rho(X + Y) \le \rho(X) + \rho(Y)$.

A risk measure that satisfies axioms 1) – 4) is called *coherent*. Other example (Mean Deviation Risk of order p):

$$\rho(X) := \mathbb{E}[X] + c(\mathbb{E}[|X - \mathbb{E}[X]|^p])^{1/p}, c > 0, p \in [0, \infty).$$

Coherent risk measures

1) $\rho(X + c) = \rho(X) + c$.

2) $X \leq Y \Rightarrow \rho(X) \leq \rho(Y)$.

3) $\rho(\lambda X) = \lambda \rho(X)$ for $\lambda \geq 0$.

4) $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

A risk measure that satisfies axioms 1) – 4) is called *coherent*. Other example (Mean Deviation Risk of order p):

$$\rho(X) := \mathbb{E}[X] + c(\mathbb{E}[|X - \mathbb{E}[X]|^p])^{1/p}, c > 0, p \in [0, \infty).$$

Risk averse newsvendor

$$\min_{x \geq 0} \{cx + \text{CVaR}_\alpha [Q(x,\xi)]\} = \min_{x \geq 0} \left\{ cx + \left(t + \frac{1}{(1-\alpha)} \mathbb{E}[Q(x,\xi)] \right) \right\},$$

$$\begin{aligned} Q(x,\xi) = \text{minimize} \quad & u \\ \text{subject to} \quad & y \leq \xi, \\ & y + w \leq x, \\ & u \geq -rw - vy - t, \\ & u, y, w \geq 0. \end{aligned}$$

Risk averse two stage stochastic programming with the CVaR

$$\min_{x \geq 0} \left\{ cx + \left(t + \frac{1}{1-\alpha} \mathbb{E}\left[Q(x,\xi) \right] \right) \right\},$$

$$\begin{aligned} Q(x,\xi) = \text{minimize} \quad & (qy - t)_+ \\ \text{subject to} \quad & Tx + Wy \geq h, \\ & y, w \geq 0. \end{aligned}$$

Numerical results

α	x^*
$0.5 \leq \alpha < 1$	1
$0 \leq \alpha < 0.5$	2

Table: $U^d[1, 10]$.

α	x^* / x^*_{RN}
.95	4%
.9	9%
.8	18%
.5	46%
.1	91%

Table: Exp(10).

Numerical results

α	x^*
$0.5 \leq \alpha < 1$	1
$0 \leq \alpha < 0.5$	2

Table: $U^d[1, 10]$.

α	x^* / x^*_{RN}
.95	4%
.9	9%
.8	18%
.5	46%
.1	91%

Table: Exp(10).

- Stochastic Dynamic Programming. Or ...
- Markov Decision Process, or
- Multistage Stochastic Programming, or
- Intertemporal Consumption, or
- Life-Cycle Consumption, or...

They all want to solve the same problem: optimal decision making over time, often under uncertainty.

- Stochastic Dynamic Programming. Or ...
- Markov Decision Process, or
- Multistage Stochastic Programming, or
- Intertemporal Consumption, or
- Life-Cycle Consumption, or...

They all want to solve the same problem: optimal decision making over time, often under uncertainty.

- Stochastic Dynamic Programming. Or ...
- Markov Decision Process, or
- Multistage Stochastic Programming, or
- Intertemporal Consumption, or
- Life-Cycle Consumption, or...

They all want to solve the same problem: optimal decision making over time, often under uncertainty.

Multi-period choice under uncertainty

- Stochastic Dynamic Programming. Or ...
- Markov Decision Process, or
- Multistage Stochastic Programming, or
- Intertemporal Consumption, or
- Life-Cycle Consumption, or...

They all want to solve the same problem: optimal decision making over time, often under uncertainty.

- Stochastic Dynamic Programming. Or ...
- Markov Decision Process, or
- Multistage Stochastic Programming, or
- Intertemporal Consumption, or
- Life-Cycle Consumption, or...

They all want to solve the same problem: optimal decision making over time, often under uncertainty.

Multi-period choice under uncertainty

- Stochastic Dynamic Programming. Or ...
- Markov Decision Process, or
- Multistage Stochastic Programming, or
- Intertemporal Consumption, or
- Life-Cycle Consumption, or...

They all want to solve the same problem: optimal decision making over time, often under uncertainty.

Multi-period choice under uncertainty

- Stochastic Dynamic Programming. Or ...
- Markov Decision Process, or
- Multistage Stochastic Programming, or
- Intertemporal Consumption, or
- Life-Cycle Consumption, or...

They all want to solve the same problem: optimal decision making over time, often under uncertainty.

- Extensive research in portfolio selection, hydrothermal scheduling, production planning and others.
- Popular algorithms include the Nested L-Shaped (Birge ’85), SDDP (Pereira and Pinto ’91), Progressive Hedging (Rockafellar and Wets ’91), SAA (Shapiro ’03, ’06), ADP (Powell ’07).
- The effectiveness of each algorithm is highly problem dependent.
- General purpose algorithms are not readily applicable.

- Extensive research in portfolio selection, hydrothermal scheduling, production planning and others.
- Popular algorithms include the Nested L-Shaped (Birge ’85), SDDP (Pereira and Pinto ’91), Progressive Hedging (Rockafellar and Wets ’91), SAA (Shapiro ’03, ’06), ADP (Powell ’07).
- The effectiveness of each algorithm is highly problem dependent.
- General purpose algorithms are not readily applicable.

Multistage stochastic programming literature

- Extensive research in portfolio selection, hydrothermal scheduling, production planning and others.
- Popular algorithms include the Nested L-Shaped (Birge '85), SDDP (Pereira and Pinto '91), Progressive Hedging (Rockafellar and Wets '91), SAA (Shapiro '03, '06), ADP (Powell '07).
- The effectiveness of each algorithm is highly problem dependent.
- General purpose algorithms are not readily applicable.

Multistage stochastic programming literature

- Extensive research in portfolio selection, hydrothermal scheduling, production planning and others.
- Popular algorithms include the Nested L-Shaped (Birge '85), SDDP (Pereira and Pinto '91), Progressive Hedging (Rockafellar and Wets '91), SAA (Shapiro '03, '06), ADP (Powell '07).
- The effectiveness of each algorithm is highly problem dependent.
- General purpose algorithms are not readily applicable.

General formulation of MSSP with random RHS

Assume $\{\xi_1, \ldots, \xi_T\}$ is a stochastic process, ξ_0 is a constant.

$$\max \ \mathbb{E}_{\xi_1,\ldots,\xi_T}\left[c_0'x_0 + c_1'x_1 + \ldots + c_T'x_T\right]$$

subject to [MSSP]

$$\begin{aligned}
A_0x_0 &\le \xi_0 \\
A_1x_1 &\le \xi_1 - B_0x_0 \\
&\vdots \\
A_Tx_T &\le \xi_T - \sum_{m=0}^{T-1} B_m x_m,
\end{aligned}$$

x_t depends only on $\xi_0, \ldots, \xi_t$.

Recursive formulation of MSSP

$$\begin{aligned} \max\ & c_0^T x_0 + \mathbb{E}_{\xi_1}\left[Q_1(x_0,\xi_1)\right] \\ \text{subject to}\ & \\ A_0 x_0\ & \le \xi_0. \end{aligned} \qquad \text{[MSSP-R]}$$

The function Q_1 is defined recursively as

$$\begin{aligned} Q_t(x_0,\ldots,x_{t-1},\xi_1,\ldots,\xi_t)\ =& \\ & \max_{x_t}\ c_t^T x_t + \mathbb{E}_{\xi_{t+1}}\left[Q_{t+1}(x_0,\ldots,x_t,\xi_1,\ldots,\xi_{t+1})\,|\,\xi_1,\ldots,\xi_t\right] \\ \text{subject to}& \\ & A_t x_t\ \le \xi_t - \sum_{m=0}^{t-1} B_m x_m, \end{aligned}$$

$t = 1,\ldots,T$. Also, $Q_{T+1} \equiv 0$.

- Assume you are a retailer who sells one product and needs to decide now how much inventory to buy, at price $c = \$2$.
- There will be two selling opportunities: in the second stage the product can be sold at price $s_1 = \$3$ and on the third stage the product can be sold for $s_2 = \$10$.
- At the end of the horizon unsold units are discarded.
- Demand is given by a binary tree.

- Assume you are a retailer who sells one product and needs to decide now how much inventory to buy, at price $c = \$2$.
- There will be two selling opportunities: in the second stage the product can be sold at price $s_1 = \$3$ and on the third stage the product can be sold for $s_2 = \$10$.
- At the end of the horizon unsold units are discarded.
- Demand is given by a binary tree.

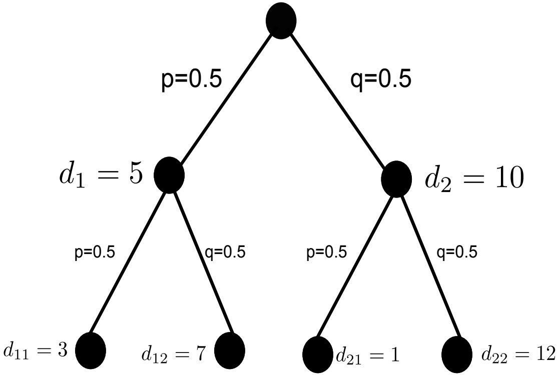
A 3-stage inventory problem

- Assume you are a retailer who sells one product and needs to decide now how much inventory to buy, at price $c = \$2$.
- There will be two selling opportunities: in the second stage the product can be sold at price $s_1 = \$3$ and on the third stage the product can be sold for $s_2 = \$10$.
- At the end of the horizon unsold units are discarded.
- Demand is given by a binary tree.

A 3-stage inventory problem

- Assume you are a retailer who sells one product and needs to decide now how much inventory to buy, at price $c = \$2$.
- There will be two selling opportunities: in the second stage the product can be sold at price $s_1 = \$3$ and on the third stage the product can be sold for $s_2 = \$10$.
- At the end of the horizon unsold units are discarded.
- Demand is given by a binary tree.

A scenario tree



Separated risk per stage

The risk neutral formulation:

$$\begin{aligned} \min cx + \mathbb{E}_1[-s_1 y] + \mathbb{E}_2[-s_2 z] \\ \text{s.t.} \quad & y \le D, \\ & y \le x, \\ & z + y \le x, \\ & y \le D. \end{aligned}$$

A possible risk averse formulation of this problem can be written as follows:

$$\begin{aligned} \min cx + \rho_1(-s_1 y) + \rho_2(-s_2 z) \\ \text{s.t.} \quad & y \le D, \\ & y \le x, \\ & z + y \le x, \\ & y \le D. \end{aligned}$$

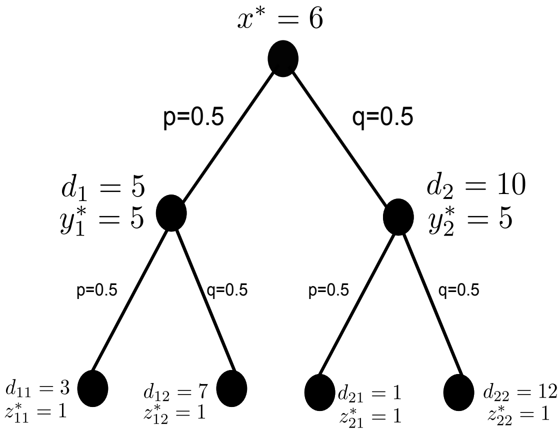
Separated risk per stage

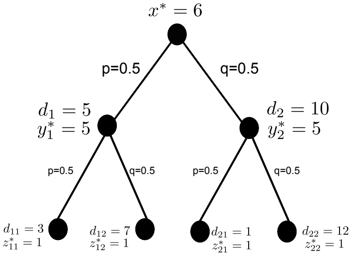
The risk neutral formulation:

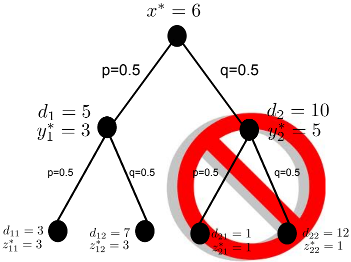
$$\begin{aligned} \min cx + \mathbb{E}_1\left[-s_1 y\right] + \mathbb{E}_2\left[-s_2 z\right] & \\ \text{s.t.} \quad & y \leq D, \\ & y \leq x, \\ & z + y \leq x, \\ & y \leq D. \end{aligned}$$

A possible risk averse formulation of this problem can be written as follows:

$$\begin{aligned} \min cx + \rho_1\left(-s_1 y\right) + \rho_2\left(-s_2 z\right) & \\ \text{s.t.} \quad & y \leq D, \\ & y \leq x, \\ & z + y \leq x, \\ & y \leq D. \end{aligned}$$


$x^* = 6$
p=0.5
q=0.5
$d_1 = 5$
$y_1^* = 5$
$d_2 = 10$
$y_2^* = 5$
p=0.5
q=0.5
p=0.5
q=0.5
$d_{11} = 3$
$z_{11}^* = 1$
$d_{12} = 7$
$z_{12}^* = 1$
$d_{21} = 1$
$z_{21}^* = 1$
$d_{22} = 12$
$z_{22}^* = 1$


$x^* = 6$
p=0.5
q=0.5
$d_1 = 5$
$y_1^* = 5$
$d_2 = 10$
$y_2^* = 5$
p=0.5
q=0.5
p=0.5
q=0.5
$d_{11} = 3$
$z_{11}^* = 1$
$d_{12} = 7$
$z_{12}^* = 1$
$d_{21} = 1$
$z_{21}^* = 1$
$d_{22} = 12$
$z_{22}^* = 1$


$x^* = 6$
p=0.5
q=0.5
$d_1 = 5$
$y_1^* = 3$
$d_2 = 10$
$y_2^* = 5$
p=0.5
q=0.5
p=0.5
q=0.5
$d_{11} = 3$
$z_{11}^* = 3$
$d_{12} = 7$
$z_{12}^* = 3$
$d_{21} = 1$
$z_{21}^* = 1$
$d_{22} = 12$
$z_{22}^* = 1$

Consistency

- There are several definitions in the literature.

- We adopt the one described in Shapiro '09, which is equivalent to the definition proposed in Carpentier et al. '12.

- If $1 \leq t_1 < t_2 \leq T$ and $\bar{x}_\tau(\xi_{[t_1,\tau]}), \tau = t_1, \ldots, T$ is *an* optimal solution of MSPP for $t = t_1$ conditional on a realization $\xi_1, \ldots, \xi_{t_1}$ of the process, then $\bar{x}_\tau(\xi_{[t_1,\tau]}), \tau = t_2, \ldots, T$ is *an* optimal solution of MSSP for $t = t_2$ conditional on a realization $\xi_1, \ldots, \xi_{t_1}, \xi_{t_1+1}, \ldots, \xi_{t_2}$ of the process.

- Informally, if you solve the problem today and find solutions for each node, you should find the same solutions if you re-solve tomorrow given what was observed and decided today.

Consistency

- There are several definitions in the literature.
- We adopt the one described in Shapiro '09, which is equivalent to the definition proposed in Carpentier et al. '12.
- If $1 \leq t_1 < t_2 \leq T$ and $\bar{x}_\tau(\xi_{[t_1,\tau]}), \tau = t_1, \ldots, T$ is *an* optimal solution of MSPP for $t = t_1$ conditional on a realization $\xi_1, \ldots, \xi_{t_1}$ of the process, then $\bar{x}_\tau(\xi_{[t_1,\tau]}), \tau = t_2, \ldots, T$ is *an* optimal solution of MSSP for $t = t_2$ conditional on a realization $\xi_1, \ldots, \xi_{t_1}, \xi_{t_1+1}, \ldots, \xi_{t_2}$ of the process.
- Informally, if you solve the problem today and find solutions for each node, you should find the same solutions if you re-solve tomorrow given what was observed and decided today.

- There are several definitions in the literature.
- We adopt the one described in Shapiro '09, which is equivalent to the definition proposed in Carpentier et al. '12.
- If $1 \leq t_1 < t_2 \leq T$ and $\bar{x}_\tau(\xi_{[t_1,\tau]}), \tau = t_1, \ldots, T$ is *an* optimal solution of MSPP for $t = t_1$ conditional on a realization $\xi_1, \ldots, \xi_{t_1}$ of the process, then $\bar{x}_\tau(\xi_{[t_1,\tau]}), \tau = t_2, \ldots, T$ is *an* optimal solution of MSSP for $t = t_2$ conditional on a realization $\xi_1, \ldots, \xi_{t_1}, \xi_{t_1+1}, \ldots, \xi_{t_2}$ of the process.
- Informally, if you solve the problem today and find solutions for each node, you should find the same solutions if you re-solve tomorrow given what was observed and decided today.

- There are several definitions in the literature.
- We adopt the one described in Shapiro '09, which is equivalent to the definition proposed in Carpentier et al. '12.
- If $1 \leq t_1 < t_2 \leq T$ and $\bar{x}_\tau(\xi_{[t_1,\tau]}), \tau = t_1, \ldots, T$ is *an* optimal solution of MSPP for $t = t_1$ conditional on a realization $\xi_1, \ldots, \xi_{t_1}$ of the process, then $\bar{x}_\tau(\xi_{[t_1,\tau]}), \tau = t_2, \ldots, T$ is *an* optimal solution of MSSP for $t = t_2$ conditional on a realization $\xi_1, \ldots, \xi_{t_1}, \xi_{t_1+1}, \ldots, \xi_{t_2}$ of the process.
- Informally, if you solve the problem today and find solutions for each node, you should find the same solutions if you re-solve tomorrow given what was observed and decided today.

$$\max \; c_0^T x_0 + \rho_{\xi_1}\left[Q_1(x_0, \xi_1)\right]$$

subject to [Risk-MSSP]

$$A_0 x_0 \le \xi_0.$$

The function Q_1 is defined recursively as

$$Q_t(x_0, \ldots, x_{t-1}, \xi_1, \ldots, \xi_t) =$$

$$\max_{x_t} \; c_t^T x_t + \rho_{\xi_{t+1}}\left[Q_{t+1}(x_0, \ldots, x_t, \xi_1, \ldots, \xi_{t+1}) \,|\, \xi_1, \ldots, \xi_t\right]$$

subject to

$$A_t x_t \le \xi_t - \sum_{m=0}^{t-1} B_m x_m,$$

$t = 1, \ldots, T$. Also, $Q_{T+1} \equiv 0$.

A promising alternative

- We have been studying the so-called $m - CVaR$ (Pflug and Ruszczynski '05)

$$\text{mCVaR}_{\alpha}(c^T x) = \sum_{t=1}^{T} \mathbb{E}_t \left[\text{CVaR}_{\alpha_t}[c_t x_t] | \xi_1, \ldots, \xi_{t-1} \right].$$

- It a promising candidate for several reasons:

1. It is midway between a separated and a nested formulation.
2. One can understand how risk is being measured
3. It can be converted into a modified risk neutral problem, which is suitable for SDDP.

A promising alternative

- We have been studying the so-called $m - CVaR$ (Pflug and Ruszczynski '05)

$$m\text{CVaR}_{\alpha}(c^T x) = \sum_{t=1}^{T} \mathbb{E}_t \left[\text{CVaR}_{\alpha_t}[c_t x_t] | \xi_1, \ldots, \xi_{t-1} \right].$$

- It a promising candidate for several reasons:
  1. It is midway between a separated and a nested formulation.
  2. One can understand how risk is being measured
  3. It can be converted into a modified risk neutral problem, which is suitable for SDDP.

A promising alternative

- We have been studying the so-called $m - CVaR$ (Pflug and Ruszczynski '05)

$$m\text{CVaR}_\alpha(c^T x) = \sum_{t=1}^{T} \mathbb{E}_t \left[\text{CVaR}_{\alpha_t}[c_t x_t] | \xi_1, \ldots, \xi_{t-1}\right].$$

- It a promising candidate for several reasons:

1. It is midway between a separated and a nested formulation.
2. One can understand how risk is being measured
3. It can be converted into a modified risk neutral problem, which is suitable for SDDP.

A promising alternative

- We have been studying the so-called $m - CVaR$ (Pflug and Ruszczynski '05)

$$mCVaR_\alpha(c^T x) = \sum_{t=1}^{T} \mathbb{E}_t \left[CVaR_{\alpha_t}[c_t x_t] | \xi_1, \ldots, \xi_{t-1} \right].$$

- It a promising candidate for several reasons:
  1. It is midway between a separated and a nested formulation.
  2. One can understand how risk is being measured
  3. It can be converted into a modified risk neutral problem, which is suitable for SDDP.

- We have been studying the so-called $m - CVaR$ (Pflug and Ruszczynski '05)

$$mCVaR_{\alpha}(c^T x) = \sum_{t=1}^{T} \mathbb{E}_t \left[CVaR_{\alpha_t}[c_t x_t] | \xi_1, \ldots, \xi_{t-1} \right].$$

- It a promising candidate for several reasons:

1. It is midway between a separated and a nested formulation.
2. One can understand how risk is being measured
3. It can be converted into a modified risk neutral problem, which is suitable for SDDP.

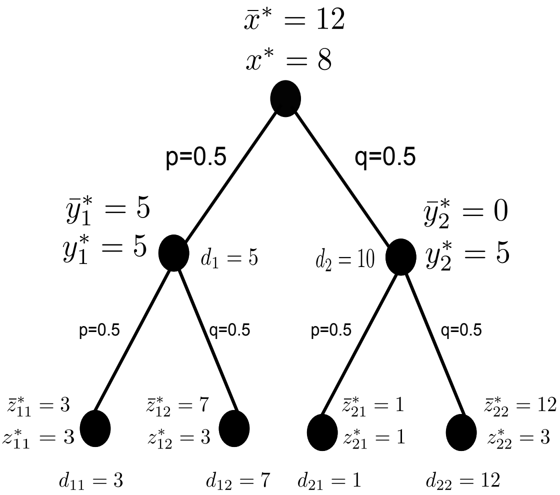
- We have been studying the so-called $m - CVaR$ (Pflug and Ruszczynski '05)

$$m\mathrm{CVaR}_{\alpha}(c^T x) = \sum_{t=1}^{T} \mathbb{E}_t \left[\mathrm{CVaR}_{\alpha_t}[c_t x_t] | \xi_1, \ldots, \xi_{t-1}\right].$$

- It a promising candidate for several reasons:

  1. It is midway between a separated and a nested formulation.
  2. One can understand how risk is being measured
  3. It can be converted into a modified risk neutral problem, which is suitable for SDDP.

Solution


$\bar{x}^* = 12$
$x^* = 8$
p=0.5
q=0.5
$\bar{y}_1^* = 5$
$y_1^* = 5$
$d_1 = 5$
$d_2 = 10$
$\bar{y}_2^* = 0$
$y_2^* = 5$
p=0.5
q=0.5
p=0.5
q=0.5
$\bar{z}_{11}^* = 3$
$z_{11}^* = 3$
$\bar{z}_{12}^* = 7$
$z_{12}^* = 3$
$\bar{z}_{21}^* = 1$
$z_{21}^* = 1$
$\bar{z}_{22}^* = 12$
$z_{22}^* = 3$
$d_{11} = 3$
$d_{12} = 7$
$d_{21} = 1$
$d_{22} = 12$

Merci!