Stochastic Optimization

Summer School Grenoble
Optimization under uncertainty
July 4-7, 2016

Abdel Lisser
lisser@lri.fr

University of Paris Sud

5 juillet 2016

Outline of the course

1. **Introduction — problem formulation, random LP**
2. Recourse models
    - Recourse, two-stage stochastic programming, farmer and newsvendor problems
3. Multi-stage models
4. Distributionally robust shortest path problem
5. Distributionally robust knapsack problem

Stochastic programming

Définition

- The word "programming" refers to optimization and decision making.
- Stochastic concerns decision making in uncertain environment.
- Stochastic Optimization $\Longrightarrow$ mathematical programming with random variables.

Stochastic programming

Introduction

- Stochastic programming = deterministic mathematical programming + random parameters.
- Consider the following deterministic mathematical programming problem :

$$
\begin{array}{lrcll}
\min & g_0(x) & & & \\
\text{s.t} & g_i(x) & \leq & 0 & i = 1, \ldots, m \\
& x \in X \subset \mathbb{R}^n. & & &
\end{array}
$$

- **Assumption** : the parameters of functions $g_i, i = 0, \ldots, m$ are known in advance.
- Is this assumption realistic ?

Nonrealistic assumption

- measure errors
- uncertainty on the future
  - Climate conditions
  - Finance markets
  - Competition
  - Technology progress
  - Availability of the data...

Optimization problem with a random parameters

$$\begin{array}{rccl} \min_{x \in X} & g_0(x,\xi) & & \\ \text{s.t} & g_i(x,\xi) & \leq & 0 \quad i = 1,\ldots,m \end{array}$$

- ξ is a random vector.

Uncertainty

Fundamental assumption

- Probability distribution is known in advance.
- This might be considered as a limitation, but
  - Partial description of the distribution could be useful $\Longrightarrow$ distributionally robust optimization.
  - If there is no information on the distribution available, we can use alternative approaches, e.g., robust optimization.

Scenario approach

- This approach allows to handle easily uncertainty, it is the first step to learn optimization with stochasticity.
- The scenario approach assumes that there exists a finite number of potential decisions. Each decision is called a realization of the random parameter or scenario, and has a related probability.
  - Demand for a given product is "low, medium or high".
  - Season is "dry" or "wet".
  - Market is "high" or "low".
- The scenarios are realizations of random variables discretely distributed. Continuous distributions are more realistic but much more difficult to handle.

The farmer problem

Illustration from Introduction to stochastic programming, Birge and Louveaux

- A farmer can plant either corn, wheat or beans.
- For the sake of simplicity, we assume that the season will either be dry or wet.
- If the season is wet, corn is more profitable.
- If the season is dry, wheat is more profitable.
- The following table gives the different strategies according to the climate.

	corn	wheat	beans
wet	100	70	80
dry	-10	40	35

Farmer problem

Illustration

- Let p be the probability that the season will be wet. The expected profit for the different cultures is :
  - Corn : $-10 + 110p$
  - Wheat : $40 + 30p$
  - Beans : $35 + 45p$

Half corn—half wheat

- Let $p = 0.5$, what is the best strategy ?
- Plant a half corn, half wheat ?
  - Expected profit : $0.5(-10 + 110(0.5)) + 0.5(40 + 30(0.5)) = 50$
- Is this strategy optimal ?

- This is not optimal because planting all beans provides a higher profit !
- Expected profit : $35 + 45(0.5) = 57.5$
- The profit increases by 15% for the optimal strategy when compared to the average strategy.

Conclusions

- Average solutions are not the best ones.
- We can't replace the random parameters by their means in case of uncertainty.
- Needs Stochastic programming to deal with stochasticity.

A random LP - From Luedtke

- Given the LP defined by :

$$
\begin{array}{lll}
\min & x_1 + x_2 & \\
\text{s.c} & \omega_1 x_1 + x_2 & \geq 7 \\
 & \omega_2 x_1 + x_2 & \geq 4 \\
 & x_1, x_2 \geq 0 & ,
\end{array}
$$

  where $\omega_1 \sim \mathcal{U}[1,4], \omega_2 \sim \mathcal{U}[1/3,1]$
- How to solve the problem ?
- What does it mean solving this problem ?
  - When should we solve it ?
  - When x should be decided ?
  - Before or after the realization of ω ?

Random LP. . .

- Wait-and-see approach consists in deciding on x after the realization of ω, i.e., solving deterministic LP.
- In real life, wait-and-see is not appropriate !
- One should decide on x before the realization of the random values of ω.

Random LP. . .

- To handle the uncertainty, three possibilities :
  - Estimate the uncertainty
  - Probabilistic constraints
  - Penalize shortfall

Non stochastic approaches

Three strategies to estimate the risk :

Risk

- Unbiased : Replace the random variables by their mean values.
- Pessimistic : worse case values for ω.
- Optimistic : choose the best values for ω.

No risk vs optimistic vs pessimistic strategies

$\omega_1 \sim \mathcal{U}[1,4], \omega_2 \sim \mathcal{U}[\frac{1}{3},1]$

- No risk strategy.
- $\hat{\omega} \equiv \mathbb{E}(\omega) = (\frac{5}{2}, \frac{2}{3})$

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.c} \quad & \tfrac{5}{2}x_1 + x_2 \geq 7 \\ & \tfrac{2}{3}x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{aligned}$$

- $\hat{v} = \frac{50}{11}$
- $(\hat{x}_1, \hat{x}_2) = (\frac{18}{11}, \frac{32}{11})$

- Pessimistic strategy
- $\hat{\omega} \equiv (\omega) = (1, \frac{1}{3})$

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.c} \quad & 1x_1 + x_2 \geq 7 \\ & \tfrac{1}{3}x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{aligned}$$

- $\hat{v} = 7$
- $(\hat{x}_1, \hat{x}_2) = (0, 7)$

- Optimistic strategy
- $\hat{\omega} \equiv (\omega) = (4, 1)$

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.c} \quad & 4x_1 + x_2 \geq 7 \\ & 1x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{aligned}$$

- $\hat{v} = 4$
- $(\hat{x}_1, \hat{x}_2) = (4, 0)$

No risk vs optimistic vs pessimistic strategies

$\omega_1 \sim \mathcal{U}[1,4], \omega_2 \sim \mathcal{U}[\frac{1}{3}, 1]$

- No risk strategy.
- $\hat{\omega} \equiv \mathbb{E}(\omega) = (\frac{5}{2}, \frac{2}{3})$

$$\begin{aligned} \min \quad & x_1 + x_2 & \\ \text{s.c} \quad & \tfrac{5}{2}x_1 + x_2 & \geq 7 \\ & \tfrac{2}{3}x_1 + x_2 & \geq 4 \\ & x_1, x_2 & \geq 0 \end{aligned}$$

- $\hat{v} = \frac{50}{11}$
- $(\hat{x}_1, \hat{x}_2) = (\frac{18}{11}, \frac{32}{11})$

- Pessimistic strategy
- $\hat{\omega} \equiv (\omega) = (1, \frac{1}{3})$

$$\begin{aligned} \min \quad & x_1 + x_2 & \\ \text{s.c} \quad & 1x_1 + x_2 & \geq 7 \\ & \tfrac{1}{3}x_1 + x_2 & \geq 4 \\ & x_1, x_2 & \geq 0 \end{aligned}$$

- $\hat{v} = 7$
- $(\hat{x}_1, \hat{x}_2) = (0, 7)$

- Optimistic strategy
- $\hat{\omega} \equiv (\omega) = (4, 1)$

$$\begin{aligned} \min \quad & x_1 + x_2 & \\ \text{s.c} \quad & 4x_1 + x_2 & \geq 7 \\ & 1x_1 + x_2 & \geq 4 \\ & x_1, x_2 & \geq 0 \end{aligned}$$

- $\hat{v} = 4$
- $(\hat{x}_1, \hat{x}_2) = (4, 0)$

No risk vs optimistic vs pessimistic strategies

$\omega_1 \sim \mathcal{U}[1,4], \omega_2 \sim \mathcal{U}[\frac{1}{3},1]$

- No risk strategy.
- $\hat{\omega} \equiv \mathbb{E}(\omega) = (\frac{5}{2}, \frac{2}{3})$

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.c} \quad & \tfrac{5}{2}x_1 + x_2 \geq 7 \\ & \tfrac{2}{3}x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{aligned}$$

- $\hat{v} = \frac{50}{11}$
- $(\hat{x}_1, \hat{x}_2) = (\frac{18}{11}, \frac{32}{11})$

- Pessimistic strategy
- $\hat{\omega} \equiv (\omega) = (1, \frac{1}{3})$

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.c} \quad & 1x_1 + x_2 \geq 7 \\ & \tfrac{1}{3}x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{aligned}$$

- $\hat{v} = 7$
- $(\hat{x}_1, \hat{x}_2) = (0, 7)$

- Optimistic strategy
- $\hat{\omega} \equiv (\omega) = (4, 1)$

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.c} \quad & 4x_1 + x_2 \geq 7 \\ & 1x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{aligned}$$

- $\hat{v} = 4$
- $(\hat{x}_1, \hat{x}_2) = (4, 0)$

Advantages vs Drawbacks

+ Easy problems to solve : LP with the same size as the random LP.
+ Only one global information on the uncertainty is considered.
- Only one realization of the random variable is taken into account.
- There might be some values of ω where x is not feasible.

Chance/probabilistic constraints

- More realistic approach !
- Let α_1, α_2 and α be large prescribed probabilities.
- The LP constraints can be written as :

$$P\{\omega_1 x_1 + x2 \geq 7\} \geq \alpha_1$$

$$P\{\omega_2 x_1 + x2 \geq 4\} \geq \alpha_2$$

- or

$$P\{\omega_1 x_1 + x2 \geq 7, \omega_2 x_1 + x2 \geq 4\} \geq \alpha$$

- If $\alpha_1 = \alpha_2 = \alpha = 1$, the stochastic LP is equivalent to the deterministic problem.

Penalize infeasibility

- We accept infeasibility but we penalize expected shortage.
- Notation :
  - $x^+ \equiv \max(0, x)$: the positive part of x.
  - $x^- \equiv \max(0, -x)$: the negative part of x.
- For the constraint $\omega_1 x_1 + x_2 \geq 7$, the shortfall is $(\omega_1 x_1 + x_2 - 7)^-$
- We assign the penalties for each constraint, e.g., q_1, q_2 .
- We have the following problem :

$$\min_{x \in \mathbb{R}_+^2} \left\{ x_1 + x_2 + q_1 \mathbb{E}_{\omega_1} \left[(\omega_1 x_1 + x_2 - 7)^- \right] + q_2 \mathbb{E}_{\omega_2} \left[(\omega_2 x_1 + x_2 - 4)^- \right] \right\}$$

Penalize infeasibility

- The following problem

$$\min_{x\in\mathbb{R}^2_+}\{x_1+x_2+q_1\mathbb{E}_{\omega_1}[(\omega_1 x_1+x_2-7)^-]+q_2\mathbb{E}_{\omega_2}[(\omega_2 x_1+x_2-4)^-]\}$$

- is equivalent to

$$\min_{x\in\mathbb{R}^2_+}\left\{x_1+x_2+\mathbb{E}_\omega\left[\min_{y\in\mathbb{R}^2_+} q_1y_1+q_2y_2:\omega_1x_1+x_2+y_1\geq 7;\omega_2x_1+x_2+y_2\geq 4\right]\right\}$$

- which can be expressed in terms of x as

$$\min_{x\in\mathbb{R}^2_+}\{x_1+x_2+\mathcal{Q}(x_1,x_2)\}$$

- where

$$\mathcal{Q}(x_1,x_2)=\mathbb{E}_\omega\left[\min_{x\in\mathbb{R}^2_+} q_1y_1+q_2y_2:y_1\geq 7-\omega_1x_1-x_2;y_2\geq 4-\omega_2x_1-x_2\right]$$

Recourse problem

- $\mathcal{Q}(x_1, x_2)$ is called recourse function.
- For a given x_1, x_2, we check the recourse which is in this case the shortfall.
- y_1, y_2 are called recourse variables and correspond to the shortfall.
- Broadly speaking, in stochastic optimization we solve the recourse problem instead of penalizing shortfall.
- The strategy "here-and-now" implies the following two period decision process :
  - We make decision on x now (first stage) ;
  - Random event occurs ;
  - We make a recourse action to correct the first stage decision (second stage or recourse).

Outline of the course

1. Introduction — problem formulation, random LP
2. **Recourse** models
   - Recourse, two-stage stochastic programming, farmer and newsvendor problems
3. Multi-stage models
4. Distributionally robust shortest path problem
5. Distributionally robust knapsack problem

The farmer problem

- A farmer has 500 acres where he can grow wheat, corn and beans.
- He needs 200T of wheat, 240T of corn to feed his cattle.
- The farmer can either plant on his land or bought from the market.
- Excess production of wheat can be sold for \$170/T, and \$150/T for corn.
- Any shortfall must be bought from the market at a cost of \$238/T for the wheat and \$210/T for the corn.
- The farmer can also sell beans for \$36/T for the first 6000 first tons and \$10/T for 6000 tons excess due to economic quotas.
- Question : can you help the farmer to design his production planning ?

Problem data

- Planting surface : 500 acres

Culture	Wheat	Corn	Beans
Yield (T/acre)	2.5	3	20
Planting cost ($/T)	150	230	260
Selling price ($/T)	170	150	$36(\leq 6000\,T)$
			$10(> 6000\,T)$
Purchasing price ($/T)	238	210	–
Minimum requirements (T)	200	240	–

LP formulation

- x_1, x_2, x_3 : surfaces to plant for wheat, corn and beans respectively.
- y_1, y_2 : tons of wheat and corn purchased.
- w_1, w_2 : tons of wheat and corn sold.
- w_3, w_4 : tons of beans sold at favorable price and lower price respectively.

LP Formulation

$$\begin{array}{lll}
\min & 150x_1 + 230x_2 + 260x_3 + & \\
 & 238y_1 - 170w_1 + 210y_2 - 150w_2 - 36w_3 - 10w_4 & \\
\text{s.t} & x_1 + x_2 + x_3 & \leq 500 \\
 & 2.5x_1 + y_1 - w_1 & \geq 200 \\
 & 3x_2 + y_2 - w_2 & \geq 240 \\
 & 20x_3 - w_3 - w_4 & = 0 \\
 & w_3 & \leq 6000 \\
 & x_1, x_2, x_3, y_1, y_2, w_1, w_2, w_3, w_4 & \geq 0
\end{array}$$

What is the solution of this LP ?

LP solution with expected yield

Plant	Wheat	Corn	Beans
Surfaces (acres)	120	80	300
Production (T)	300	240	6000
Sales (T)	100	–	6000
Purchase (T)	0	0	0

Results

- Total profit : \$118600

Deterministic solution is satisfied ?

- Can the farmer be happy with the deterministic solution ?
  - Plant beans with respect to quotas.
  - Plant land necessary for feeding his cattle with wheat and corn.
  - The remaining surface is dedicated to wheat in order to sell excess production.

Quid of climate conditions ?

- The farmer knows that the yield depends on the climate conditions, and are not always $R = (2.5, 3, 20)$.
- It is reasonable to consider assumptions with a good weather and bad weather :
  - Good weather : $1.2R$
  - Bad weather : $0.8R$

Formulation - Good weather conditions

$$
\begin{array}{lll}
\min & 150x_1 + 230x_2 + 260x_3 + & \\
& 238y_1 - 170w_1 + 210y_2 - 150w_2 - 36w_3 - 10w_4 & \\
\text{s.t} & x_1 + x_2 + x_3 & \leq 500 \\
& 3x_1 + y_1 - w_1 & \geq 200 \\
& 3.6x_2 + y_2 - w_2 & \geq 240 \\
& 24x_3 - w_3 - w_4 & = 0 \\
& w_3 & \leq 6000 \\
& x_1, x_2, x_3, y_1, y_2, w_1, w_2, w_3, w_4 & \geq 0
\end{array}
$$

Good weather Solution

Plant	Wheat	corn	Beans
Surfaces (acres)	183.33	66.67	250
Production (T)	550	240	6000
Sales (T)	350	–	6000
Purchase (T)	0	0	0

Results

- Total Profit : $167667

Formulation - Bad weather

$$
\begin{array}{lll}
\min & 150x_1 + 230x_2 + 260x_3 + & \\
 & 238y_1 - 170w_1 + 210y_2 - 150w_2 - 36w_3 - 10w_4 & \\
\text{s.t} & x_1 + x_2 + x_3 & \leq 500 \\
 & 2x_1 + y_1 - w_1 & \geq 200 \\
 & 2.4x_2 + y_2 - w_2 & \geq 240 \\
 & 16x_3 - w_3 - w_4 & = 0 \\
 & w_3 & \leq 6000 \\
 & x_1, x_2, x_3, y_1, y_2, w_1, w_2, w_3, w_4 & \geq 0
\end{array}
$$

Solution Bad condition

Plant	Wheat	Corn	Beans
Surfaces (acres)	100	25	375
Production (T)	200	60	6000
Sales (T)	0	0	6000
Purchase (T)	0	180	0

Results

- Total Profit : $59950

Conclusions

- Decisions depend strongly on the climate change and yield.
- It is impossible to take reasonable decision in one step, i.e. decide the plant surface assignment without taking into account climate conditions which are not know in advance.
- Decide on (x_1, x_2, x_3) now, and decide on $(w_i, i = 1, \dots, 4, y_j, j = 1, 2)$ (sales and purchase) which comes in the future.
- Alternative : scenarios

Maximize the expected profit

- Assume that the three yield scenarios are equiprobable (good weather, average climate conditions, bad weather). The prices are the same.
- Let $s = 1, 2, 3$ be the scenario indices associated with sales and purchase variables w_{is}, y_{is}.
- $s = 1$ if the weather conditions are good, 2 for average climate conditions, and 3 for bad weather.
- Example : w_{22} is the quantity of corn sold in scenario 2.

Maximize the expected profit

$$
\begin{array}{lll}
\min & 150x_1 + 230x_2 + 260x_3 + \sum_{s=1}^{3} \frac{1}{3}(238y_{1s} - 170w_{1s} + & \\
 & 210y_{2s} - 150w_{2s} - 36w_{3s} - 10w_{4s}) & \\
\text{s.t} & x_1 + x_2 + x_3 & \leq 500 \\
 & 3x_1 + y_{11} - w_{11} & \geq 200 \\
 & 2.5x_1 + y_{12} - w_{11} & \geq 200 \\
 & 2x_1 + y_{13} - w_{13} & \geq 200 \\
 & 3.6x_2 + y_{21} - w_{21} & \geq 240 \\
 & 3x_2 + y_{22} - w_{22} & \geq 240 \\
 & 2.4x_2 + y_{23} - w_{23} & \geq 240 \\
 & 24x_3 - w_{31} - w_{41} & = 0 \\
 & 20x_3 - w_{32} - w_{42} & = 0 \\
 & 16x_3 - w_{33} - w_{43} & = 0 \\
 & w_{31}, w_{32}, w_{33} & \leq 6000 \\
 & x_1, x_2, x_3, y_{1s}, y_{2s}, w_{1s}, w_{2s}, w_{3s}, w_{4s} & \geq 0, s = 1, 3
\end{array}
$$

Optimal solution

	Plant	Wheat	corn	Beans
First step	surface (acres)	170	80	250
$s = 1$	Production (T)	510	288	6000
	Sales (T)	310	48	6000
	Purchase (T)	0	0	0
$s = 2$	Production (T)	425	240	5000
	Sales (T)	225	0	5000
	Purchase (T)	0	0	0
$s = 3$	Production (T)	340	192	4000
	Sales (T)	140	0	4000
	Purchase (T)	0	48	0

- Total expected profit : $108390

Wait and See

Wait and See

- The farmer has already solved the three scenarios with a respect to the yield assumptions.
- This strategy is called Wait and See

	$0.8R$	R	$1.2R$
Corn	25	80	66.67
Wheat	100	120	183.33
Beans	375	300	250
Profit	59950	118600	167667

Solution characteristics

Solution analysis

- If the farmer could have some knowledge about the best scenario, his problem would be

$$\frac{1}{3}(167667) + \frac{1}{3}(118600) + \frac{1}{3}(59950) = \$115406$$

.

- In this case, we have a decision under perfect information.
- The difference between this solution and the stochastic solution is called Expected Value of Perfect Information.

Solution characteristics

Solution analysis

- If the farmer uses the average information, the profit will be

$$\frac{1}{3}(118600 + 55120 + 148000) = \$107240$$

(results obtained by taking the values of first level variables as $x_1 = 120, x_2 = 80, x_3 = 300$ (expected value solution),i.e., a loss of \$1150 compared to to the stochastic solution 108390.

- This difference is known as Value of the Stochastic Solution.

Solution characteristics

Solution analysis

- Using stochastic programming gives more profit to the farmer $1150 compared to the average solution.
- land surface planting are called first stage variables whilst sales and purchase variables are called second order variables.

Solution . . .

Analysis

- The best solution assign surfaces to the bean culture in order to avoid selling at low prices.
- Corn is planted with the objective to cover the farmer requirements in the average case.
- The remaining surface is assigned to the wheat.
- The best solution is impossible to find ! Stochastic models give balanced solutions between the different scenarios.

Random LP

- Consider the following LP parameterized with the random vector ω :

$$
\begin{aligned}
\min \quad & c^T x \\
\text{s.t} \quad & \\
& Ax = b \\
& T(\omega)x = h(\omega) \\
& x \in X
\end{aligned}
$$

where $X = \{x \in \mathcal{R}^n : l \leq x \leq u\}$

Random LP

- We need to solve an optimization problem where the decision on x should be taken before the realization of the random event ω.
- We should know the distribution of ω defined in Ω.
- In the recourse models, random constraints are considered as soft constraints. They could not be satisfied but the violation cost has an influence on the choice of x.

The new LP

- However, we can reasonably solve the following problem :

$$
\begin{aligned}
\min \quad & c^T x + \mathbb{E}[q_+^T s(\omega) + q_-^T t(\omega)] \\
\text{s.t} \quad & \\
& Ax = b \\
& T(\omega)x + s(\omega) - t(\omega) = h(\omega), \qquad \forall \omega \in \Omega \\
& x \in X.
\end{aligned}
$$

Recourse

- Broadly speaking, we can react appropriately thanks to a recourse structure.
- The recourse allows to correct a first stage decision. Recourse is defined by three components :
  1. A set of feasible recourse actions $Y = \{y \in \mathbb{R}^p : y \geq 0\}$.
  2. Recourse cost vector q.
  3. Recourse $m \times p-$matrix denoted W.

Formulation with recourse

- The new LP can be written as :

$$
\begin{array}{ll}
\min & c^T x + \mathbb{E}[q^T y] \\
\text{s.t} & \\
& Ax = b \\
& T(\omega)x + Wy(\omega) = h(\omega), \qquad \forall \omega \in \Omega \\
& x \in X \\
& y \in Y
\end{array}
$$

- In this case, the matrix W is not dependent on the scenarios.
- when W is unique, the recourse is called fixed recourse.

Definitions

$$\min_{x \in X : Ax = b} \{c^T x + \mathbb{E}_\omega[\min_{y \in Y}\{q^T y : Wy = h(\omega) - T(\omega)x\}]\}$$

- The recourse function can be written :
- $v : \mathbb{R}^m \mapsto \mathbb{R}$
  - $v(z) \equiv \min_{y \in Y}\{q^T y : Wy = z\}$
  - This function describes the random constraints violations $T(\omega)x = h(\omega)$ as function of the vecteur z.
- Minimum recourse function (expected value function) $\mathcal{Q} : \mathbb{R}^n \mapsto \mathbb{R}$
  - $\mathcal{Q}(x) \equiv \mathbb{E}_\omega[v(h(\omega) - T(\omega)x)]$
  - For any decision x, this function gives the expected recourse cost.

- The stochastic problem with recourse can be written as :

$$\min_{x \in X} \{c^T x + \mathcal{Q}(x) : Ax = b\}$$

- This problem is not linear in $\mathbb{R}^n$.
- Solving this problem depends on the mathematical properties $\mathcal{Q}$ i.e.
  - Is it linear, convex, continuous, differentiable... ?

Analytical model

$$
\begin{array}{lll}
\min\limits_{x\in\mathbb{R}^n, y(\omega)\in\mathbb{R}^n} & \mathbb{E}[c^T x + q^T y(\omega)] & \\
\text{s.t} & & \\
 & Ax = b & \text{First level} \\
 & T(\omega)x + Wy(\omega) = h(\omega), \forall \omega \in \Omega & \text{Second level} \\
 & x \in X, y(\omega) \in Y &
\end{array}
$$

- If we consider $\Omega = \{\omega_1, \omega_2, \dots, \omega_S\} \subseteq \mathbb{R}^r$ is discrete set, then
- $P(\omega = \omega_s) = p_s, \forall s = 1, 2, \dots, S$
- $T_s \equiv T(\omega_s), h_s = h(\omega_s)$

DEP

$$
\begin{array}{cccccccc}
\min & c^T x & +p_1 q^T y_1 & +p_2 q^T y_2 & + \dots & +p_s q^T y_s & & \\
\text{s.t} & & & & & & & \\
Ax & & & & & & = & b \\
T_1 x & +Wy_1 & & & & & = & h_1 \\
T_2 x & & +Wy_2 & & & & = & h_2 \\
\vdots & & & \ddots & & & & \\
T_S x & & & & +Wy_s & & = & h_s \\
x \in X & y_1 \in Y & y_2 \in Y & & y_s \in Y & & &
\end{array}
$$

DEP

- y_s is the recourse action to be taken if the scenario ω_s occurs.
- The Deterministic Equivalent Problem is an LP which is easy to solve BUT its size is the main drawback :
  - Number of variables : $n + pS$
  - Number of constraints : $m1 + mS$
- Despite the large size of the matrix, it has however a particular structure easy to handle, i.e. block diagonal matrix.

Newsvendor problem

- the newsvendor problem consists in deciding the number of newspapers to buy in order to maximize the profit.
- The newsvendor does not know the daily demand in advance.
- Each newspaper costs c and could be sold q.
- Unsold newspapers can be returned at the end of the day for a price r.
- Given the probability distribution $F(t) = P(\omega \leq t)$, how many newspapers should the newsvendor buy in order to maximize his profit ?

Newsvendor problem

- According to the recourse definition, the newsvendor problem can be written as :

$$\max_{x \geq 0} \{-cx + \mathcal{Q}(x)\}$$

- $\mathcal{Q}(x)$ is the expected profit the newspaper can make if he buys x newspapers :

$$\mathcal{Q}(x) = \mathbb{E}_{\omega} Q(x, \omega)$$

- The recourse function can be also written as

$$\mathcal{Q}(x) = v(h(\omega) - T(\omega)x)$$

Recourse

- $\mathcal{Q}(x, \omega)$ is the expected profit if the newsvendor buys x newspapers for a demand ω.
- Let y be the variable representing the effective sales and v the unsold newspapers. The recourse problem can be written :

$$\begin{aligned} \mathcal{Q}(x, \omega) = \quad \min \quad & -qy(\omega) - rw(\omega) \\ \text{s.t.} \quad & y(\omega) \leq \omega \\ & y(\omega) + w(\omega) \leq x \\ & y(\omega), w(\omega) \geq 0. \end{aligned}$$

- another formulation :

$$\mathcal{Q}(x,\omega) = \begin{array}{ll} qx & x \leq \omega \\ q\omega + r(x-\omega) & x \geq \omega \end{array} \tag{1}$$

$$\begin{aligned}
&\mathcal{Q}(x) = \mathbb{E}_\omega \mathcal{Q}(x,\omega) = \int_{-\infty}^{\infty} \mathcal{Q}(x,\omega) dF(\omega) \\
&= \int_{-\infty}^{x} (q\omega + r(x-\omega)) dF(\omega) + \int_{x}^{\infty} qx dF(\omega) \\
&= (q-r) \int_{-\infty}^{x} \omega dF(\omega) + rx \int_{-\infty}^{x} dF(\omega) + qx \int_{x}^{\infty} dF(\omega) \\
&= (q-r) \int_{-\infty}^{x} \omega dF(\omega) + rxF(x) + qx(1 - F(x))
\end{aligned}$$

Karuch-Kuhn-Tucker conditions (KKT)

- Calculate the derivative of $Q(x)$: $\mathcal{Q}'(x) = q - (q-r)F(x)$
- The optimal solution satisfy

$$-c + q - (q-r)F(x) = 0$$

- x^* is the optimal solution when $F(x) = \left(\frac{q-c}{q-r}\right)$ i.e.

$$x^* = F^{-1}\left(\frac{q-c}{q-r}\right)$$

Numerical application

- $c = 0.15$
- $q = 0.25$
- $r = 0.02$
- $\omega \sim \mathcal{N}(650, 80)$.
- $x^* = F^{-1}(0.1/0.23) = 636.86$

Two-stage stochastic problem with a fixed recourse

$$
\begin{aligned}
\min \quad & c^T x + \mathbb{E}_\omega[q^T y] \\
\text{s.t.} \quad & Ax = b \\
& T(\omega)x + Wy(\omega) = h(\omega), \qquad \forall \omega \in \Omega \\
& x \in X \\
& y(\omega) \in Y.
\end{aligned}
$$

- $\mathcal{Q}(x, \omega) = \min_{y \in Y}\{q^T y : Wy = h(\omega) - T(\omega)x\}$

- Assume that $Y = \mathbb{R}^n_+$.

$$v(z) = \min_{y \in \mathbb{R}^p_+} \{q^T y : Wy = z\}, z \in \mathbb{R}^m$$

- for a fixed z, we solve the LP problem to evaluate $v(z)$.
- If duality theory holds $\forall z \in \mathbb{R}^m$, $-\infty < v(z) < \infty$
- According to duality theory,
  $v(z) = \min_{y \in \mathbb{R}^p_+} \{q^T y : Wy = z\} = \max_{t \in \mathbb{R}^m} \{z^T t : W^T t \leq q\}$
- $v(z)$ is convex for $z \in \mathbb{R}^m$.

Properties

- for a fixed ω, $\mathcal{Q}(x,\omega)$ is convex.
- Recall that $\mathcal{Q}(x,\omega) \equiv v(h(\omega) - T(\omega)x)$

$$\begin{aligned} & \lambda(\mathcal{Q}(x_1,\omega) + (1-\lambda)\mathcal{Q}(x_2,\omega) \\ = & \lambda v(h(\omega) - T(\omega)x_1) + (1-\lambda)v(h(\omega) - T(\omega)x_2) \\ \geq & v(\lambda(h(\omega) - T(\omega)x_1) + (1-\lambda)(h(\omega) - T(\omega)x_2) \\ = & v(h(\omega) - T(\omega)(x_1) + (\lambda x_1 + (1-\lambda))x_2) \\ = & \mathcal{Q}(\lambda x_1 + (1-\lambda))x_2,\omega) \end{aligned}$$

Récapitulatif : equivalent formulations

$$\min_{x\in\mathbb{R}^n_+ : Ax=b} \left\{ c^T x + \mathbb{E}_\omega \left[\min_{y\in\mathbb{R}^p_+} \{ q^T y : Wy = h(\omega) - T(\omega)x \} \right] \right\}$$

$$\min_{x\in\mathbb{R}^n_+ : Ax=b} \left\{ c^T x + \mathbb{E}_\omega v\left(h(\omega) - T(\omega)x\right) \right\}$$

$$\min_{x\in\mathbb{R}^n_+ : Ax=b} \left\{ c^T x + \mathbb{E}_\omega \mathcal{Q}(x,\omega) \right\}$$

$$\min_{x\in\mathbb{R}^n_+} \left\{ c^T x + \mathcal{Q}(x) : Ax = b \right\}$$

Relatively complete recourse

- First stage feasibility problem : $K_1 = \{x \in \mathbb{R}^n_+ : Ax = b\}$
- Second stge feasibility problem : $K_2 = \{x | \mathcal{Q}(x) < \infty\}$

  We can rewrite the problem as

$$\min\{c^T x + \mathcal{Q}(x) : x \in K_1 \cap K_2\}$$

Def. We call relatively complete recourse problem if $K_1 \subseteq K_2$

- Consequently, if x is a first stage feasible solution, then $\mathcal{Q}(x, \omega) < \infty$.

Complete recourse

- Denote $K_2(\omega) = \{x | \mathcal{Q}(x, \omega) < \infty\}$ the set of feasible points for a given realization of ω.
- $K_2 = \cap_{\omega \in \Omega} K_2(\omega)$

Def. A recourse problem is called complete if $\forall z \in \mathbb{R}^m, v(z) < \infty$ i.e. $\forall z \in \mathbb{R}^m, \exists y \in \mathbb{R}^p_+ : Wy = z$.

$\implies \forall x, T(\omega), h(\omega), \mathcal{Q}(x, \omega) < \infty$ as $z = h - Tx$

- The complete recourse is also a property of W

i.e. If the columns of W generate $\mathbb{R}^m$, then $\forall z \in \mathbb{R}^m, \exists y \in \mathbb{R}^p_+ : Wy = z$.

Simple recourse

- Assume that $\mathcal{Q}(x,\omega) = \min_{y \in \mathbb{R}^p_+} \{q^T y : Wy = h(\omega) - T(\omega)x\}$ is not feasible for a given x.
- It suffices to set $W = [I, -I]$, in this case we have a simple recourse.
- The second stage problem can be written as :

$$\begin{aligned} \mathcal{Q}(x,\omega) = \quad \min y \quad & (q^+)^T y^+ + (q^-)^T y^- \\ \text{s.c.} \quad & y^+ - y^- = h(\omega) - T(\omega)x \\ & y^+, y^- \geq 0. \end{aligned}$$

Recourse problems properties

Theorem 1

If $f_1(x), f_2(x), \ldots, f_q(x)$ is a set of convex functions, then $M = \max\{f_1(x), f_2(x), \ldots, f_q(x)\}$ is also a convex function.

$\Longrightarrow$ $\mathcal{Q}(x, \omega) \equiv v(h(\omega) - T(\omega)x)$ is convex.

- $\mathcal{Q}(x) \equiv \mathbb{E}_\omega \mathcal{Q}(x, \omega)$ is convex.
- These definitions hold also either for discrete functions or continuous.

Properties of recourse problem (KKT)

Theorem 2

Given convex functions $f : \mathbb{R}^n \longmapsto \mathbb{R}$ *and* $g_i : \mathbb{R}^n \longmapsto \mathbb{R}, i = 1, \ldots, m$. $\hat{x}$ *is an optimal solution (under certain regularization conditions) of* $\min_{x \in \mathbb{R}^n_+} \{f(x) : g_i(x) \leq 0, \forall i = 1, \ldots, m\}$ *iif*

- $g_i(x) = 0, \forall i = 1, \ldots, m$
- $\exists \lambda_1, \ldots, \lambda_m \in \mathbb{R}, \mu_1, \ldots, \mu_n \in \mathbb{R}_+$ *tel que*
  - $0 \in \partial f(\hat{x} + \sum_{i=1}^m \lambda_i \partial g_i(\hat{x}) - \sum_{j=1}^n \mu_j$
  - $\mu_j \geq 0, \forall j = 1, \ldots, n$
  - $\mu_j \hat{x}_j = 0, \forall j = 1, \ldots, n$

Application KKT to the recourse problem

$$\min_{x \in \mathbb{R}^n_+ : Ax = b} \left\{ c^T x + \mathcal{Q}(x) : Ax = b \right\}$$

Theorem 3

$\hat{x} \in K_1$ is optimal iif

- $\exists \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^n_+$ *such that*
  - $0 \in c + \partial \mathcal{Q}(\hat{x}) + A^T \lambda - \mu$
  - $\mu^T \hat{x} = 0$

où encore

$$-c - A^T \lambda + \mu \in \partial \mathcal{Q}(\hat{x})$$

Outline of the course

1. Introduction — problem formulation, random LP
2. Recourse models
   - Recourse, two-stage stochastic programming, farmer and newsvendor problems
3. **Multi**-stage models
4. Distributionally robust shortest path problem
5. Distributionally robust knapsack problem

Production problem -from Beasley-

Multi-stage problem

- A company has to make a decision about the amount of a product χ to produce.
- Each unit of χ that we make costs us $2e$. χ is made to meet demand from customers in the next time period.
- Customer demand must be met. We have the flexibility to buy in the product from an external supplier for $3e$ per unit.
- Demande D of χ is a random variable defined by :
  - $P(D = 500) = 0.6$ for scenario 1.
  - $P(D = 700) = 0.4$ for scenario 2.
- How much should we make now before we know what customer demand is ?

Solving method

- Let $\Omega = \{1, 2\}$, the set of scenarios.
- The problem variables are :
  - x_1 : the number of units of χ to produce now.
  - $y_{2\omega}$: the number of units of χ to buy from the market at the second stage for the scenarios $\omega = 1, 2$.
  - The stochastic program can be written as :

$$\min_{x \geq 0} \{2x_1 + \mathbb{E}_\omega [Q(x, \omega)]\}$$

where

$$Q(x, \omega) = [\min_{y_{2\omega}} \{q(\omega)^T y_{2\omega} : T(\omega)x_1 + W(\omega)y_{2\omega} = h(\omega)\}]$$

Recourse function

Recourse

$$Q(x,1) = \min_{y_{21}} \quad 3y_{21} \quad \text{s. t.} \quad y_{21} \geq 500 - x_1, \quad y_{21} \geq 0.$$

$$Q(x,2) = \min_{y_{22}} \quad 3y_{22} \quad \text{s. t.} \quad y_{22} \geq 700 - x_1, \quad y_{22} \geq 0.$$

Deterministic equivalent problem

$$\begin{aligned} \min_{x,y} \quad & 2x_1 + (0.6)3y_{21} + (0.4)3y_{22} \\ \text{s. t.} \quad & x_1 + y_{21} \geq 500 \\ & x_1 + y_{22} \geq 700 \\ & x_1, y_{21}, y_{22} \geq 0. \end{aligned}$$

3-stage problem

Three stages

- The previous example can be extended to a 3-stage problem.
- Assume that after the realization of the second level, owe can still produce additional quantities x_2 of χ to satisfy the demand at the third level.
- The units bought at level 2 represent inventories for the level 3.

A 3-stage problem

Demand scenarios

- Assume that demand D_1 for the second level is :
  - 500 with probability 0.6
  - 700 with probability 0.4
- and the demand D_2 for the third level is :

$$(D_1 = 500) : \begin{cases} 600 & \text{avec } p = 0.3 \\ 700 & \text{avec } p = 0.7. \end{cases} \qquad (D_1 = 700) : \begin{cases} 900 & \text{avec } p = 0.2 \\ 800 & \text{avec } p = 0.8. \end{cases}$$

3-stage problem

Demand scenarios

- Demand vector (D_1, d_2) implies $2^2 = 4$ scenarios.
- If D_1 et D_2 are independent random variables, we have the following joint probabilities :

1. $P(D_1 = 500, D_2 = 600) = 0.6 * 0.3 = 0.18$ Scenario 1
2. $P(D_1 = 500, D_2 = 700) = 0.6 * 0.7 = 0.42$ Scenario 2
3. $P(D_1 = 700, D_2 = 900) = 0.4 * 0.2 = 0.08$ Scenario 3
4. $P(D_1 = 700, D_2 = 800) = 0.4 * 0.8 = 0.32$ Scenario 4

3-stage problem

Variables

- let $\Omega = \{1, 2, 3, 4\}$ be the set of scenarios.
- x_1 : first stage production.
- $y_{2\omega}$: quantity to buy at the second level, $\omega = \{1, 2, 3, 4\}$
- $x_{2\omega}$: quantity to produce at the second level for the scenarios $\omega = \{1, 2, 3, 4\}$
- $y_{3\omega}$: quantity to buy at the third level, $\omega = \{1, 2, 3, 4\}$

3-stage problem

Problem data

- To satisfy the demand at the second level, we have the following constraints :

$$\begin{cases} x_1 + y_{2\omega} & \geq 500 \qquad \omega = 1, 2 \\ x_1 + y_{2\omega} & \geq 700 \qquad \omega = 3, 4. \end{cases}$$

- The inventory constraints are (units left over at the second stage) :

$$\left. \begin{array}{lll} x_1 + y_{2\omega} - & 500 & \omega = 1, 2 \\ x_1 + y_{2\omega} - & 700 & \omega = 3, 4. \end{array} \right\}$$

3-stage problem

Problem data

- To satisfy the demand of the third level, we have the following constraints :

$$\left.\begin{array}{llr} x_1 + y_{2\omega} - 500 + x_{2\omega} + y_{3\omega} & \geq 600 & \omega = 1 \\ x_1 + y_{2\omega} - 500 + x_{2\omega} + y_{3\omega} & \geq 700 & \omega = 2 \\ x_1 + y_{2\omega} - 700 + x_{2\omega} + y_{3\omega} & \geq 900 & \omega = 3 \\ x_1 + y_{2\omega} - 700 + x_{2\omega} + y_{3\omega} & \geq 800 & \omega = 4 \end{array}\right\}$$

Problem formulation

Objective function

- The objective function can be formulated by the date of the following table :

Scenario	Probability	Cost
1	0.18	$2x_{21} + 3y_{21} + 3y_{31}$
2	0.42	$2x_{22} + 3y_{22} + 3y_{32}$
3	0.08	$2x_{23} + 3y_{23} + 3y_{33}$
4	0.32	$2x_{24} + 3y_{24} + 3y_{34}$

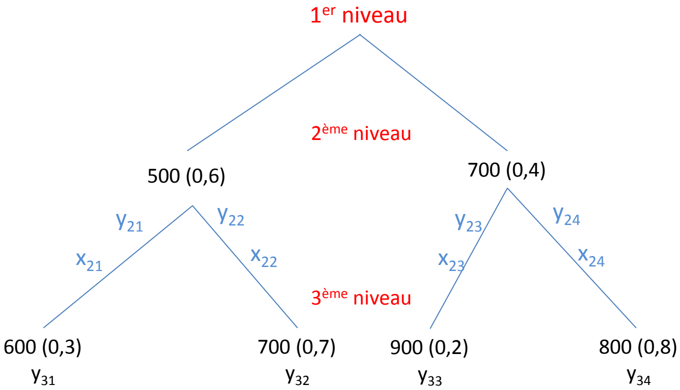
Problem formulation

SP

- The stochastic program is :

$$
\begin{array}{lll}
(SP_1)\ \min\limits_{x,y} & 2x_1 & \\
 & +(0.18)(2x_{21} + 3y_{21} + 3y_{31}) + (0.42)(2x_{22} + 3y_{22} + 3y_{32}) & \\
 & +(0.08)(2x_{23} + 3y_{23} + 3y_{33}) + (0.32)(2x_{24} + 3y_{24} + 3y_{34}) & \\
\text{s. t.} & x_1 + y_{21} & \geq 500 \\
 & x_1 + y_{22} & \geq 500 \\
 & x_1 + y_{23} & \geq 700 \\
 & x_1 + y_{24} & \geq 700 \\
 & x_1 + y_{21} + x_{21} + y_{31} & \geq 1100 \\
 & x_1 + y_{22} + x_{22} + y_{32} & \geq 1200 \\
 & x_1 + y_{23} + x_{23} + y_{33} & \geq 1600 \\
 & x_1 + y_{24} + x_{24} + y_{34} & \geq 1500 \\
 & x_1,\, y_{2\omega},\, y_{3\omega} \geq 0,\ \omega = 1, 4. &
\end{array}
$$

Scenario tree representation



Problem formulation

Remarks

- Are the constraints of (SP_1) sufficient to distinguish the scenarios 1 and 2 at the second level ?
- These two scenarios have the same history until the second level : decision x_1 at level 1 and observed realization of 500 at level 2.
- Scenarios 1 and 2 do not differ only at the third level. It would seem logical therefore that y_{21} and y_{22} must be equal, we add the following constraints :
  1. For scenarios 1 and 2 :

$$\begin{cases} y_{21} & = y_{22} \\ x_{21} & = x_{22}. \end{cases}$$

  2. For scenarios 3 and 4 :

$$\begin{cases} y_{23} & = y_{24} \\ x_{23} & = x_{24}. \end{cases}$$

- These constraints are called **nonanticipativity** constraints. They imply that we cannot anticipate the future.

Problem formulation

SP with nonanticipativity constriants

- The SP can be written :

$$
\begin{array}{lll}
(SP_2)\ \min\limits_{x,y} & 2x_1 & \\
& +(0.18)(2x_{21}+3y_{21}+3y_{31})+(0.42)(2x_{22}+3y_{22}+3y_{32}) & \\
& +(0.08)(2x_{23}+3y_{23}+3y_{33})+(0.32)(2x_{24}+3y_{24}+3y_{34}) & \\
\text{s. t.} & x_1+y_{21} & \geq 500 \\
& x_1+y_{22} & \geq 500 \\
& x_1+y_{23} & \geq 700 \\
& x_1+y_{24} & \geq 700 \\
& x_1+y_{21}+x_{21}+y_{31} & \geq 1100 \\
& x_1+y_{22}+x_{22}+y_{32} & \geq 1200 \\
& x_1+y_{23}+x_{23}+y_{33} & \geq 1600 \\
& x_1+y_{24}+x_{24}+y_{34} & \geq 1500 \\
& y_{21}-y_{22} & =0 \\
& x_{21}-x_{22} & =0 \\
& y_{23}-y_{24} & =0 \\
& x_{23}-x_{24} & =0 \\
& x_1,\ y_{2\omega},\ y_{3\omega}\ \geq 0,\ \omega=1,4. &
\end{array}
$$

Three-stage Stochastic optimization problem

Solutions with and without nonanticipativity constraints

Without nonanticipativity	With nonanticipativity
$z^* = 2620$	$z^* = 2664$
$x_1 = 960$	$x_1 = 945$
$x_{21} = 140$	$x_{21} = 254$
$x_{22} = 240$	$x_{22} = 254$
$y_{21} = 0$	$y_{21} = 0$
$y_{22} = 0$	$y_{22} = 0$
$y_{23} = 0$	$y_{23} = 0$
$y_{24} = 0$	$y_{24} = 0$
$x_{23} = 640$	$x_{23} = 554$
$x_{24} = 540$	$x_{24} = 554$
$y_{33} = 0$	$y_{33} = 100$
$y_{34} = 0$	$y_{34} = 0$

Multistage models, from Ruszczynski

Sequential decision process x_t in stages $t = 1, \dots, T$ with learning (observations) of the random data D_t :

$$
\begin{array}{ll}
 & \text{decision}(x_1) \\
\text{observation } D_2 := (c_2, A_{21}, A_{22}, b_2) & \text{decision}(x_2) \\
\vdots & \\
\text{observation } D_T := (c_T, A_{T,T-1}, A_{TT}, b_T) & \text{decision}(x_T)
\end{array}
$$

Objective : Design the decision process in order to minimize the expected value of the total cost.

$$
\begin{array}{llllllll}
\min & \mathbb{E}[c_1x_1 & +c_2x_2 & +c_3x_3 & \dots & +c_Tx_T] & \\
\text{s.t.} & A_{11}x_1 & & & & & = b_1 \\
 & A_{21}x_1 & +A_{22}x_2 & & & & = b_2 \\
 & & A_{32}x_2 & +A_{33}x_3 & & & = b_3 \\
 & \dots & \dots & \dots & \dots & \dots & \dots \\
 & & & & A_{T,T-1}x_{T-1} & +A_{TT}x_T & = b_T \\
 & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & \dots & x_T \geq 0. &
\end{array}
$$

Assume that x_1, A_{11} and b_1 are known.

Nonanticipativity

$$
\begin{array}{lllllll}
\min & \sum_{s=1}^{S} p_s[c_1 x_1^s & +c_2^s x_2^s & +c_3^s x_3^s & \dots & +c_T^s x_T^s] & \\
\text{s. t.} & A_{11} x_1^s & & & & & = b_1 \\
 & A_{21}^s x_1^s & +A_{22}^s x_2^s & & & & = b_2^s \\
 & & A_{32}^s x_2^s & +A_{33}^s x_3^s & & & = b_3^s \\
 & \dots & \dots & \dots & \dots & \dots & \dots \\
 & & & & A_{T,T-1}^s x_{T-1}^s & +A_{TT}^s x_T & {}^s = b_T^s \\
 & x_1^s \geq 0 & x_2^s \geq 0 & x_3^s \geq 0 & \dots & x_T^s \geq 0, & s = 1, \dots, S
\end{array}
$$

- Feasibility constraints : $(x_1^s, x_1^s, \dots, x_T^s) \in F^s$ (feasibility set for scenario s)
- Nonanticipativy constraints : $x_t^s = x_t^\sigma$ for all s, σ s.t. $D_{[1,t]}^s = D_{[1,t]}^\sigma, t = 1, \dots, T$

References

Bibliography

- John Birge and François Louveaux, *Introduction to stochastic programming*, Springer Verlag, 1997.
- Peter Kall and Stein Wallace, *Stochastic Programming*, John Wiley & Sons, 1994.
- Andrzej Ruszczynski et Alexander Shapiro, *Stochastic Programming*, Elsevier, 2003.
- Jianqiang Cheng, Janny Leung, Abdel Lisser, *New reformulations of distributionally robust shortest path problem*, Computers & Operations Research 74 : 196-204 (2016).
- Jianqiang Cheng, Erick Delage and Abdel Lisser, *Distributionally Robust Stochastic Knapsack Problem*, SIAM Journal on Optimization 24(3) : 1485-1506 (2014).