Operations Research

Lecture 7: Two-Stage Stochastic Programming

Shuming Wang

School of Economics & Management
University of Chinese Academy of Sciences
Email: wangshuming@ucas.edu.cn

- In all the previous optimization models, the decision $\boldsymbol{x}$ is made to tackle all the realizations of the uncertainty $\boldsymbol{\xi}$.
- This is a **static type** of decision making.
- In many practical situations, the decisions $\boldsymbol{y}(\omega)$ adapt to the uncertainty $\boldsymbol{\xi}(\omega)$, which therefore involves **dynamics**.

Outline

Outline

Consider a European farmer Tom who specializes in raising (a)wheat, (b)corn, and (c)sugar beets on his 500 acres of land. During the winter, Tom wants to decide how much land to devote to each crop.

(a) wheat

(b) corn

(c) sugarbeet

Farmer's problem

- Tom knows that at least 200 tons (T) of wheat and 240 T of corn are needed for cattle feed. These amounts can be raised on the farm or bought from a wholesaler. Any production exceeding the feeding requirement would be sold.
- The mean selling prices of wheat and corn (over the last decade) have been \$170 and \$150 per T, respectively. The purchase prices are 40% more than these.
- Another profitable crop is sugar beet, which he expects to sell at \$36 per T. However, the EC imposes a quota of 6000 T on sugar beet production: any amount in excess of 6000 T can only be sold at \$10 per T.

Farmer's problem

- Based on the past experience, the farmer knows that the mean yield on his land is roughly 2.5 T, 3 T, and 20 T per acre for wheat, corn, and sugar beets, respectively.
- The planning costs for wheat, corn, and sugar beets, are \$150, \$230, and \$260 per T, respectively.

Farmer's problem

These data are summarized in Table 1.

Table 1: Data for farmer's problem

	Wheat	Corn	Sugar Beets
Yield (T/acre)	2.5	3	20
Planning cost ($/acre)	150	230	260
Selling price ($/T)	170	150	36 under 6000 T
	–	–	10 above 6000 T
Purchase price ($/T)	238	210	–
Minimum requirement (T)	200	240	–
Total available land:	500 acres		

Farmer's problem

- ξ_1, ξ_2, ξ_3 are the yield (ton / unit land) for wheat, corn, and sugar beets, respectively, which usually are **uncertain variables related to the uncertain weather conditions.**
- x_1, x_2, x_3 are the land allocations for wheat, corn, and sugar beets, respectively, which **have to be decided before the uncertainty (yield) realized.**
- w_1, w_2 are the quantities of wheat and corn to sell, respectively, and w_3 and w_4 are the quantities of sugar beets to be sold at different prices s_3 and s_4, respectively.
- y_1 and y_2 are the quantities of wheat and corn to purchase, in case of the shortage to the reserve requirement.

Farmer's problem

Deterministic Base Model

When everything is known, we have:

$$
\begin{aligned}
\text{Max} \quad & 170w_1 + 150w_2 + 36w_3 + 10w_4 - 238y_1 - 210y_2 - 150x_1 - 230x_2 - 260x_3 \\
\text{s.t.} \quad & \sum_{i=1}^{3} x_i \leq 500 \\
& \hat{\xi}_1 x_1 + y_1 - w_1 \geq 200 \\
& \hat{\xi}_2 x_2 + y_2 - w_2 \geq 240 \\
& w_3 + w_4 \leq \hat{\xi}_3 x_3 \\
& w_3 \leq 6000 \\
& x_i, w_j \geq 0, i = 1,2,3, j = 1,2,3,4 \\
& y_j \geq 0, j = 1,2
\end{aligned}
$$

where $\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3$ are yields of wheat, corn, and sugar beets, respectively, and s_j, q_j and c_i are parameters.

Farmer’s problem

Uncertainty in Yields

Table 2: Distribution of uncertain yields (T/unit of land)

Scenarios	Yield of wheat (ξ_1)	Y. corn (ξ_2)	Y. sugar beets (ξ_3)	Probability
good	3.0	3.6	24	1/3
normal	2.5	3.0	20	1/3
bad	2.0	2.4	16	1/3
Mean value	2.5	3.0	20	

(d) good weather

(e) normal weather

(f) bad weather

It should be noted that decisions that "how much to sell (at what prices)? (w_1, w_2, w_3, w_4)" and "how much to buy"(y_1, y_2) are made after the realizations of yields coming out.

Decision Process:

$$(x_1, x_2, x_3) \longrightarrow \begin{cases} (\xi_1^1, \xi_2^1, \xi_3^1) \longrightarrow (w_1^1, w_2^1, w_3^1, w_4^1, y_1^1, y_2^1) \\ \\ (\xi_1^2, \xi_2^2, \xi_3^2) \longrightarrow (w_1^2, w_2^2, w_3^2, w_4^2, y_1^2, y_2^2) \\ \\ (\xi_1^3, \xi_2^3, \xi_3^3) \longrightarrow (w_1^3, w_2^3, w_3^3, w_4^3, y_1^3, y_2^3) \end{cases}$$

Farmer's problem

Stochastic Model

$$
\begin{aligned}
\text{Max} \quad & \frac{1}{3}\Big(170w_1^1 + 150w_2^1 + 36w_3^1 + 10w_4^1 - 238y_1^1 - 210y_2^1\Big) \\
& + \frac{1}{3}\Big(170w_1^2 + 150w_2^2 + 36w_3^2 + 10w_4^2 - 238y_1^2 - 210y_2^2\Big) \\
& + \frac{1}{3}\Big(170w_1^3 + 150w_2^3 + 36w_3^3 + 10w_4^3 - 238y_1^3 - 210y_2^3\Big) - 150x_1 - 230x_2 - 260x_3 \\
\text{s.t.} \quad & \sum_{i=1}^{3} x_i \le 500, x_i \ge 0, i = 1,2,3 \\
& \begin{array}{lll}
3.0x_1 + y_1^1 - w_1^1 \ge 200; & 2.5x_1 + y_1^2 - w_1^2 \ge 200; & 2.0x_1 + y_1^3 - w_1^3 \ge 200 \\
3.6x_2 + y_2^1 - w_2^3 \ge 240; & 3.0x_2 + y_2^2 - w_2^3 \ge 240; & 2.4x_2 + y_2^3 - w_2^3 \ge 240 \\
w_3^1 + w_4^1 \le 24x_3; & w_3^2 + w_4^2 \le 20x_3; & w_3^3 + w_4^3 \le 16x_3 \\
w_3^1 \le 6000; & w_3^2 \le 6000; & w_3^3 \le 6000
\end{array} \\
& w_j^k \ge 0, j = 1,2,3,4, k = 1,2,3 \\
& y_j^k \ge 0, j = 1,2, k = 1,2,3
\end{aligned}
$$

Stochastic solution

Table 3: Solution obtained from the stochastic problem

First-stage decision	x_1	x_2	x_3			
	170	80	250			
Scenario/Second-stage decision	w_1^k	w_2^k	w_3^k	w_4^k	y_1^k	y_2^k
Good weather ($k = 1$)	310	48	6000	0	0	0
Normal weather ($k = 2$)	225	0	5000	0	0	0
Bad weather ($k = 3$)	140	0	4000	0	0	48

The stochastic solution obtained is $(x_1, x_2, x_3) = (170, 80, 250)$.

Farmer's problem

Mean value scenario problem

$$
\begin{aligned}
\text{Max} \quad & 170w_1 + 150w_2 + 36w_3 + 10w_4 - 238y_1 - 210y_2 - 150x_1 - 230x_2 - 260x_3 \\
\text{s.t.} \quad & \sum_{i=1}^{3} x_i \leq 500 \\
& \mathbf{2.5}x_1 + y_1 - w_1 \geq 200 \\
& \mathbf{3.0}x_2 + y_2 - w_2 \geq 240 \\
& w_3 + w_4 \leq \mathbf{20}x_3 \\
& w_3 \leq 6000 \\
& x_i, w_j \geq 0, i = 1,2,3, j = 1,2,3,4 \\
& y_j \geq 0, j = 1,2
\end{aligned}
$$

where $\bar{\xi_1} = 2.5, \bar{\xi_2} = 3.0$ and $\bar{\xi_3} = 20$ are the mean values of ξ_1, ξ_2 and ξ_3, respectively.

Farmer's Problem

Mean Value Solution (MVS)

Table 4: Solution obtained from the stochastic problem

First-stage decision	x_1	x_2	x_3			
	120	80	300			
Second-stage decision	w_1	w_2	w_3	w_4	y_1	y_2
	100	0	6000	0	0	0
Objective (profit)	118600					

The mean value solution is denoted by $(\bar{x}_1, \bar{x}_2, \bar{x}_3) = (120, 80, 300)$

Farmer's problem

Evaluating the mean value solution $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ by following the practical decision process

Evaluation of MV Solution

$$
\begin{aligned}
\text{Max} \quad & \sum_{k=1}^{3} p_k \left[\sum_{j=1}^{4} s_j w_j^k - \sum_{j=1}^{2} q_j y_j^k \right] - \sum_{i=1}^{3} c_i \bar{x}_i \\
\text{s.t.} \quad & \widehat{\xi}_1^k \bar{x}_1 + y_1^k - w_1^k \geq 200, k = 1, 2, 3 \\
& \widehat{\xi}_2^k \bar{x}_2 + y_2^k - w_2^k \geq 240, k = 1, 2, 3 \\
& w_3^k + w_4^k \leq \widehat{\xi}_3^k \bar{x}_3, k = 1, 2, 3 \\
& w_3^k \leq 6000, k = 1, 2, 3 \\
& w_j^k \geq 0, j = 1, 2, 3, 4, k = 1, 2, 3 \\
& y_j^k \geq 0, j = 1, 2, k = 1, 2, 3
\end{aligned}
$$

where $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ is the mean value solution, $\widehat{\xi}_i^k$ is the sample of ξ_i in

Table 5: Mean profit obtained in different scenarios

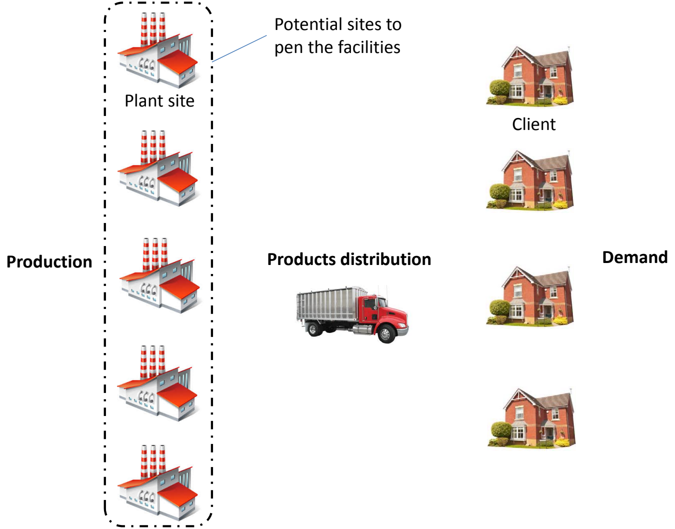
First-stage decision	Mean value solution	Stochastic solution
$\boldsymbol{x} = (x_1, x_2, x_3)$	(120,80,300)	(170,80, 250)
Mean profit	107240	108390

- If we use the mean value solution in decision making process, we lose value 108390-107240=1150!
- This amount of value is called the value of stochastic solution (VSS).

Outline

Facility Location-Allocation Problem

Background



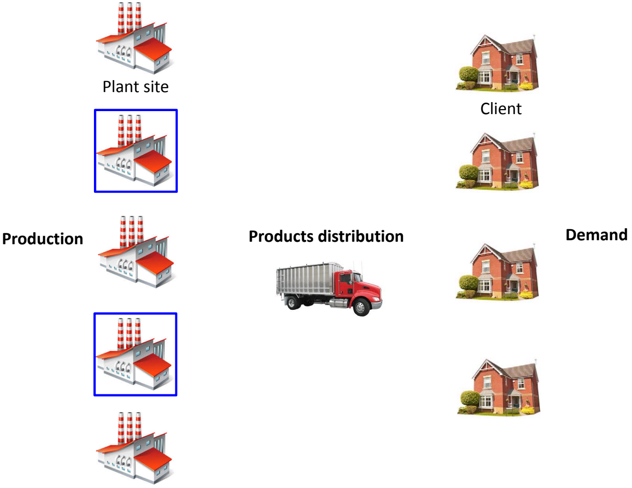


Figure 1: Location decision

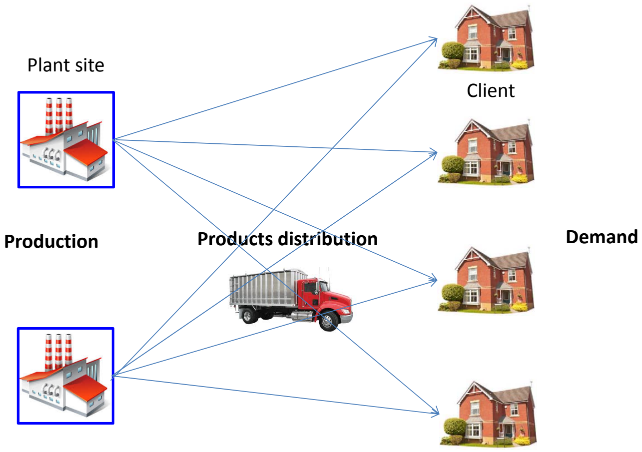


Figure 2: Productions allocation process

Indices:

- i is the index of facilities, $i \in [N]$
- j is the index of clients, $j \in [M]$

Parameters:

- r_j is the unit price charged to client j
- s_i is the capacity of facility i
- c_i is the fixed cost for opening the facility i
- t_{ij} is the unit transportation cost from i to j
- $\tilde{d}_j$ is **random demand** of client j
- $\tilde{v}_i$ is the **random unit variable cost** of operating facility i

Decisions:

- x_i binary decision variable indicating opening the facility i
- y_{ij} the quantity supplied from facility i to client j

Objective:

$$\operatorname*{Max}_{\boldsymbol{x}} \quad \underbrace{\mathbb{E}_{\omega}\Big[\ \overbrace{\mathcal{Q}\big(\boldsymbol{x}, \tilde{\boldsymbol{d}}(\omega), \tilde{\boldsymbol{v}}(\omega)\big)}^{\text{Operational return}}\ \Big]}_{\text{Expected operational return}} - \underbrace{\sum_{i=1}^{N} c_i x_i}_{\text{Fixed cost}}$$

where

$$\mathcal{Q}\Big(\boldsymbol{x}, \boldsymbol{d}(\omega), \boldsymbol{v}(\omega)\Big) := \operatorname*{Max}_{\boldsymbol{y} \in \mathbf{F}(\omega, \boldsymbol{x})} \sum_{i=1}^{N} \sum_{j=1}^{M} \big(r_j - \tilde{v}_i(\omega) - t_{ij}\big) y_{ij}.$$

Two types of decisions:

- Location decision $\boldsymbol{x}$: strategic decision (long-term decision).
- Products allocations $\boldsymbol{y}$: operations decision (short-term decision, scenario dependent).

Balance constraints:

- Each customer's demand cannot be over-served, but it is possible that the demand is not fully satisfied:

$$\sum_{i=1}^{N} y_{ij} \leq \tilde{d}_j(\omega), j \in [M].$$

- The total supply from one facility to all clients cannot exceed the capacity of the facility:

$$\sum_{j=1}^{M} y_{ij} \leq s_i x_i, i \in [N]$$

$$\left.\begin{array}{rl}\underset{\boldsymbol{x}}{\text{Max}} & \mathbb{E}\Big[\mathcal{Q}(\boldsymbol{x},\tilde{\boldsymbol{d}},\tilde{\boldsymbol{v}})\Big]-\sum_{i=1}^{N}c_ix_i \\ \text{subject to} & x_i\in\{0,1\}, i\in[N],\end{array}\right\} \quad (1)$$

where $\mathcal{Q}(\boldsymbol{x},\tilde{\boldsymbol{d}},\tilde{\boldsymbol{v}})$ is the second-stage value function, which given each realization $(\tilde{\boldsymbol{d}}(\omega),\tilde{\boldsymbol{v}}(\omega))$ solves an optimization sub-problem:

$$\left.\begin{array}{rl}\mathcal{Q}\Big(\boldsymbol{x},\boldsymbol{d}(\omega),\boldsymbol{v}(\omega)\Big):=\underset{\boldsymbol{y}}{\text{Max}} & \sum_{i=1}^{N}\sum_{j=1}^{M}\big(r_j-\tilde{v}_i(\omega)-t_{ij}\big)y_{ij} \\ \text{s.t.} & \\ & \sum_{i=1}^{N}y_{ij}\le\tilde{d}_j(\omega), j\in[M], \\ & \sum_{j=1}^{M}y_{ij}\le s_ix_i, i\in[N], \\ & y_{ij}\ge 0, i\in[N], j\in[M].\end{array}\right\} \quad (2)$$

$$
\begin{aligned}
\underset{\boldsymbol{x}}{\text{Max}} \quad & \mathbb{E}\left[\underset{\boldsymbol{y}\in\mathbf{F}(\omega,\boldsymbol{x})}{\text{Max}} \sum_{i=1}^{N}\sum_{j=1}^{M}\big(r_j-\tilde{v}_i(\omega)-t_{ij}\big)y_{ij}\right]-\sum_{i=1}^{N}c_i x_i \\
\text{subject to} \quad & x_i\in\{0,1\}, i\in[N],
\end{aligned}
$$

where $\mathbf{F}(\omega,\boldsymbol{x})$ is the scenario-dependent feasible set, defined by

$$
\mathbf{F}(\omega,\boldsymbol{x}) := \left\{ \boldsymbol{y} \;\middle|\; \begin{array}{l} \sum_{i=1}^{N} y_{ij} \leq \tilde{d}_j(\omega), j\in[M] \\ \sum_{j=1}^{M} y_{ij} \leq s_i x_i, i\in[N] \\ y_{ij}\geq 0, i\in[N], j\in[M] \end{array} \right\}
$$

Given samples $\{(\hat{\boldsymbol{d}}^k, \hat{\boldsymbol{v}}^k), k \in [K]\}$, we have the following Formulation:

Sample Formulation for Two-Stage Location Problem

$$
\begin{aligned}
\max_{\boldsymbol{x}, \boldsymbol{y}^k} \quad & \frac{1}{K} \sum_{k=1}^{K} \left[\sum_{i=1}^{N} \sum_{j=1}^{M} \left(r_j - \hat{v}_i^k - t_{ij} \right) y_{ij} \right] - \sum_{i=1}^{N} c_i x_i \\
\text{s.t.} \quad & \sum_{i=1}^{N} y_{ij}^k \leq \hat{d}_j^k, j \in [M], k \in [K] \\
& \sum_{j=1}^{M} y_{ij}^k \leq s_i x_i, i \in [N], k \in [K] \\
& y_{ij}^k \geq 0, i \in [N], j \in [M], k \in [K] \\
& x_i \in \{0, 1\}, i \in [N].
\end{aligned}
$$

Given samples $\{(\hat{\boldsymbol{d}}^k, \hat{\boldsymbol{v}}^k), k \in [K]\}$, we have the following Formulation:

Sample Formulation for Two-Stage Location Problem

$$
\begin{aligned}
\max_{\boldsymbol{x}, \boldsymbol{y}^k} \quad & \frac{1}{K} \sum_{k=1}^{K} \left[\sum_{i=1}^{N} \sum_{j=1}^{M} \left(r_j - \hat{v}_i^k - t_{ij} \right) y_{ij} \right] - \sum_{i=1}^{N} c_i x_i \\
\text{s.t.} \quad & \sum_{i=1}^{N} y_{ij}^k \leq \hat{d}_j^k, j \in [M], k \in [K] \\
& \sum_{j=1}^{M} y_{ij}^k \leq s_i x_i, i \in [N], k \in [K] \\
& y_{ij}^k \geq 0, i \in [N], j \in [M], k \in [K] \\
& x_i \in \{0, 1\}, i \in [N].
\end{aligned}
$$

- The SAA formulation is a mixed integer LP for our location problem.
- The number of integers does not grow up with the number of samples. (Different from our Chance-Constrained Model!)

Outline

Two-Stage Stochastic (Linear) Program with Recourse

$$\min_{\boldsymbol{x}} \quad \mathbb{E}_{\mathbb{P}}\Big[\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi})\Big] + \boldsymbol{c}'\boldsymbol{x} \tag{3a}$$

$$\text{s.t.} \quad \boldsymbol{x} \in \mathcal{X} := \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}, \tag{3b}$$

where $\boldsymbol{\xi}$ is the random parameter, and $\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi})$ is the second stage value function which given $\boldsymbol{x}$ and realization $\boldsymbol{\xi}(\omega)$ of the random parameter solves the following sub-problem:

$$\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\omega)) := \min_{\boldsymbol{y}} \quad \boldsymbol{q}_\omega'\boldsymbol{y} \tag{4a}$$

$$\text{s.t.} \quad \boldsymbol{T}_\omega\boldsymbol{x} + \boldsymbol{W}_\omega\boldsymbol{y} = \boldsymbol{h}_\omega, \tag{4b}$$

$$\boldsymbol{y} \geq 0. \tag{4c}$$

Two-Stage Stochastic (Linear) Program with Recourse

$$\min_{\boldsymbol{x}} \quad \mathbb{E}_{\mathbb{P}}\Big[\mathcal{Q}(\boldsymbol{x},\boldsymbol{\xi})\Big] + \boldsymbol{c}'\boldsymbol{x} \tag{3a}$$

$$\text{s.t.} \quad \boldsymbol{x} \in \mathcal{X} := \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}, \tag{3b}$$

where $\boldsymbol{\xi}$ is the random parameter, and $\mathcal{Q}(\boldsymbol{x},\boldsymbol{\xi})$ is the second stage value function which given $\boldsymbol{x}$ and realization $\boldsymbol{\xi}(\omega)$ of the random parameter solves the following sub-problem:

$$\mathcal{Q}(\boldsymbol{x},\boldsymbol{\xi}(\omega)) := \min_{\boldsymbol{y}} \quad \boldsymbol{q}_\omega'\boldsymbol{y} \tag{4a}$$

$$\text{s.t.} \quad \boldsymbol{T}_\omega\boldsymbol{x} + \boldsymbol{W}_\omega\boldsymbol{y} = \boldsymbol{h}_\omega, \tag{4b}$$

$$\boldsymbol{y} \geq 0. \tag{4c}$$

- $\boldsymbol{T}_\omega := \boldsymbol{T}(\boldsymbol{\xi}(\omega)), \boldsymbol{q}_\omega := \boldsymbol{q}(\boldsymbol{\xi}(\omega)), \boldsymbol{W}_\omega := \boldsymbol{W}(\boldsymbol{\xi}(\omega)), \boldsymbol{h}_\omega := \boldsymbol{h}(\boldsymbol{\xi}(\omega))$

Integrated Formulation

$$
\begin{aligned}
\underset{\boldsymbol{x}}{\text{Min}} \quad & \mathbb{E}_{\mathbb{P}}\left[\underset{\boldsymbol{y}}{\text{Min}}\left\{\boldsymbol{q}_\omega'\boldsymbol{y} \;\middle|\; \begin{array}{l} \boldsymbol{T}_\omega\boldsymbol{x}+\boldsymbol{W}_\omega\boldsymbol{y}=\boldsymbol{h}_\omega, \omega\in\Omega \\ \boldsymbol{y}\geq 0 \end{array}\right\}\right]+\boldsymbol{c}'\boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \boldsymbol{x}\geq \boldsymbol{0}.
\end{aligned}
$$

Integrated Formulation

$$\begin{array}{ll}\underset{\boldsymbol{x}}{\text{Min}} & \mathbb{E}_{\mathbb{P}}\left[\underset{\boldsymbol{y}}{\text{Min}}\left\{\boldsymbol{q}_\omega'\boldsymbol{y}\;\middle|\;\begin{array}{l}\boldsymbol{T}_\omega\boldsymbol{x}+\boldsymbol{W}_\omega\boldsymbol{y}=\boldsymbol{h}_\omega,\omega\in\Omega\\ \boldsymbol{y}\geq 0\end{array}\right\}\right]+\boldsymbol{c}'\boldsymbol{x}\\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geq\boldsymbol{0}.\end{array}$$

- $\boldsymbol{x}$ is the first-stage decision, or here-and-now decision.
- Matrix $\boldsymbol{T}_\omega$ and vectors $\boldsymbol{q}_\omega$ and $\boldsymbol{h}_\omega$ are uncertain that contains random entries.
- Matrix $\boldsymbol{W}_\omega$ is called Recourse Matrix
- $\boldsymbol{y}$ is called the second-stage (wait-and-see) decision which adapts the realizations of uncertainty $\boldsymbol{T}_\omega$ and vectors $\boldsymbol{q}_\omega$ and $\boldsymbol{h}_\omega$

An key condition for a two-stage SP to be well defined:

$$\text{Min}\left\{\mathbb{E}_{\mathbb{P}}\left(\left[\mathcal{Q}(\boldsymbol{x},\boldsymbol{\xi})\right]_{+}\right), \mathbb{E}_{\mathbb{P}}\left(\left[\mathcal{Q}(\boldsymbol{x},\boldsymbol{\xi})\right]_{-}\right)\right\} < \infty, \forall \boldsymbol{x} \in \mathcal{X}.$$

Some important particular situations:

Fixed and Complete Recourse

- Fixed Recourse: $\boldsymbol{W}_{\omega} \equiv \boldsymbol{W}, \forall \omega \in \Omega$
- Complete Recourse: The polyhedron

$$\mathbf{P}(\boldsymbol{\eta}) := \left\{\boldsymbol{y} : \mathbf{W}\boldsymbol{y} = \boldsymbol{\eta}, \boldsymbol{y} \geq \mathbf{0}\right\} \neq \emptyset, \forall \boldsymbol{\eta}$$

$\Leftrightarrow \Pi(\boldsymbol{q}) := \{\boldsymbol{p} : \mathbf{W}'\boldsymbol{p} \leq \boldsymbol{q}\}$ is bounded, $\forall \boldsymbol{q} \Leftrightarrow$ Recession Cone $= \{\mathbf{0}\}$

Example (Simple Recourse)

- $\boldsymbol{W}_\omega \equiv \boldsymbol{W}, \mathbf{T}_\omega \equiv \mathbf{T}, \boldsymbol{q}_\omega \equiv \boldsymbol{q}, \forall \omega \in \Omega$
- $\boldsymbol{W} = [\mathbf{I}; -\mathbf{I}]$

- H.W.: Simple recourse is fixed and complete recourse

An key condition for a two-stage SP to be well defined:

$$\mathrm{Min}\left\{\mathbb{E}_{\mathbb{P}}\left(\left[\mathcal{Q}(\boldsymbol{x},\boldsymbol{\xi})\right]_+\right),\mathbb{E}_{\mathbb{P}}\left(\left[\mathcal{Q}(\boldsymbol{x},\boldsymbol{\xi})\right]_-\right)\right\} < \infty, \forall \boldsymbol{x} \in \mathcal{X}.$$

Some important particular situations:

Relative Complete Recourse

$$\mathbb{P}\left\{\omega \in \Omega : \mathcal{Q}(\boldsymbol{x},\boldsymbol{\xi}(\omega)) < \infty\right\} = 1, \forall \boldsymbol{x} \in \mathcal{X}.$$

- The general problem is difficult to solve due to the task of multiple integration

$$\mathbb{E}_{\mathbb{P}}\Big[\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi})\Big].$$

- The general problem is difficult to solve due to the task of multiple integration

$$\mathbb{E}_{\mathbb{P}}\left[\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi})\right].$$

- A more practical way to handle the two-stage SP (3a)-(3b) is to use the available samples of the uncertain parameters $\boldsymbol{\xi}$, that is $\{\boldsymbol{\xi}(\hat{\omega}_i), i \in [N]\}$.
- The idea behind this approach is to use the sample distribution

$$\mathbb{P}\{\boldsymbol{\xi} = \boldsymbol{\xi}(\hat{\omega}_i)\} = \frac{1}{N},\ i \in [N]$$

to approximate the true distribution of $\boldsymbol{\xi}$, and we solve the problem with this sample discrete distribution.

SAA Formulation

$$\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{y}_i} \quad & \left[\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{q}_i'\boldsymbol{y}_i\right] + \boldsymbol{c}'\boldsymbol{x} && (5)\\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}, && (6)\\
& \boldsymbol{T}_i\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}_i = \boldsymbol{h}_i, i \in [N] && (7)\\
& \boldsymbol{y}_i \geq 0, i \in [N], && (8)
\end{aligned}$$

where we drop the "ω" and denote $\boldsymbol{q}_i := \boldsymbol{q}_{\omega_i}$ for the concise sake, the same rule applies to $\boldsymbol{T}_i$ and $\boldsymbol{h}_i$.

- This approach has a name: *Sample Average Approximation* or SAA for short.
- With SAA formulation, at least we can solve approximately the SP problem as an LP.

Outline

L-shaped Solution Approach

We have good news and bad news for the SAA approach:

- Good news: the SAA model (5)-(7) is a Linear Programming!
- Bad news: the size of problem (5)-(7) could become HUGE! For example, when the components of $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_M)$ is i.i.d. and each $\xi_m, m \in [M]$ is a 2-point distribution. Then $N = 2^M$ which is exponential in the dimension of $\boldsymbol{\xi}$.

Therefore, we introduce a decomposition approach, namely *L*-shaped algorithm, which can handle the large scale SAA problem in an iterative fashion, and in each iteration we only need to solve a small scale problem.

L-shaped Solution Approach

Lemma (Subgradient (revisited))

A function $f : \Re^n \mapsto \Re$ is convex if and only if for any $\hat{\boldsymbol{x}} \in \Re^n$, there exists a vector $\boldsymbol{s} \in \Re^n$ such that

$$f(\boldsymbol{x}) \geq f(\hat{\boldsymbol{x}}) + \boldsymbol{s}'(\boldsymbol{x} - \hat{\boldsymbol{x}}),$$

for all $\boldsymbol{x} \in \Re^n$.

- *The vector $\boldsymbol{s}$ is called a* **Subgradient** *of f at $\hat{\boldsymbol{x}}$.*
- *The set of all subgradients of f at $\hat{\boldsymbol{x}}$ is denoted by $\partial f(\hat{\boldsymbol{x}})$ and is called the* **Subdifferential** *of f at $\hat{\boldsymbol{x}}$.*
- *Especially, when f is differentiable, then the Subdifferential reduces to the one point set of gradient $\partial f(\hat{\boldsymbol{x}}) = \{\nabla f(\hat{\boldsymbol{x}})\}$.*

L-shaped Solution Approach

Under the samples condition $\{\omega_i, i \in [N]\}$, we consider the SAA formulation of two-stage SP:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{Min}} \quad & \mathcal{G}(\boldsymbol{x}) + \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned}$$

where $\mathcal{G}(\boldsymbol{x}) := \frac{1}{N}\sum_{i=1}^{N} \mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\hat{\omega}_i))$ and

$$\begin{aligned} \mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\omega_i)) := \underset{\boldsymbol{y}}{\text{Min}} \quad & \boldsymbol{q}_i'\boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{T}_i\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y} = \boldsymbol{h}_i, \\ & \boldsymbol{y} \geq 0. \end{aligned}$$

- Assumption: Fixed Recourse $\boldsymbol{W}$.
- Assumption: for each feasible $\boldsymbol{x}$ and sample ω_i, $\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\omega_i)) < \infty$.

- Applying the strong duality of the LP to $\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\omega_i))$, we can have

$$\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\omega_i)) = \operatorname*{Max}_{\boldsymbol{p}} \left\{ \boldsymbol{p}'(\boldsymbol{h}_i - \boldsymbol{T}_i \boldsymbol{x}) : \boldsymbol{W}'\boldsymbol{p} \le \boldsymbol{q}_i \right\},$$

which is a convex function in $\boldsymbol{x}$ for each $\omega_i \in \Omega$.

- We can have an optimal solution $\boldsymbol{p}_i^*$ to each dual sub-problem i, that is

$$\boldsymbol{p}_i^* := \operatorname*{argmax}_{\boldsymbol{p}} \left\{ \boldsymbol{p}'(\boldsymbol{h}_i - \boldsymbol{T}_i \boldsymbol{x}) : \boldsymbol{W}'\boldsymbol{p} \le \boldsymbol{q}_i \right\},$$

therefore $\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\omega_i)) = (\boldsymbol{h}_i - \boldsymbol{T}_i \boldsymbol{x})'\boldsymbol{p}_i^*$, and

$$-\boldsymbol{T}_i'\boldsymbol{p}_i^* \in \partial \mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\omega_i))$$

Furthermore, we can determine a subgradient of $\mathcal{G}(\boldsymbol{x})$ at $\hat{\boldsymbol{x}}$:

$$\boldsymbol{s} := -\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{T}_i' \hat{\boldsymbol{p}}_i^* \in \frac{1}{N} \sum_{i=1}^{N} \partial \mathcal{Q}(\hat{\boldsymbol{x}}, \boldsymbol{\xi}(\omega_i)) = \partial \mathcal{G}(\hat{\boldsymbol{x}}).$$

- By property of subgradient, in the two-stage SP:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{Min}} & \mathcal{G}(\boldsymbol{x}) + \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \mathcal{X} := \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}, \end{array}$$

we can replace the original $\mathcal{G}(\boldsymbol{x})$ with $\mathcal{G}(\hat{\boldsymbol{x}}) + \boldsymbol{s}'(\boldsymbol{x} - \hat{\boldsymbol{x}})$ which is a linear function of $\boldsymbol{x}$, and obtain a lower bound problem

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{Min}} \left[\mathcal{G}(\hat{\boldsymbol{x}}) + \boldsymbol{s}'(\boldsymbol{x} - \hat{\boldsymbol{x}})\right] + \boldsymbol{c}'\boldsymbol{x} \leq \underset{\boldsymbol{x}\in\mathcal{X}}{\text{Min}}\, \mathcal{G}(\boldsymbol{x}) + \boldsymbol{c}'\boldsymbol{x}.$$

L-shaped Solution Approach

- Imagine now we had L subgradients $\boldsymbol{s}_\ell$ of $\mathcal{G}(\boldsymbol{x})$ at $\hat{\boldsymbol{x}}_\ell, \ell \in [L]$, then we can solve the following lower bound problem which is an LP:

$$\begin{aligned} \underset{\boldsymbol{x},\theta}{\text{Min}} \quad & \theta + \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}, \\ & \theta \geq \mathcal{G}(\hat{\boldsymbol{x}}_\ell) + \boldsymbol{s}'_\ell(\boldsymbol{x} - \hat{\boldsymbol{x}}_\ell), \ell = 1, 2, \cdots, L. \end{aligned}$$

- Whenever new subgradient $\hat{\boldsymbol{s}}$ is obtained for new $\hat{\boldsymbol{x}}$, we can add the *Cut*:

$$\mathcal{G}(\hat{\boldsymbol{x}}) + \hat{\boldsymbol{s}}'(\boldsymbol{x} - \hat{\boldsymbol{x}})$$

  into the constrain to improve the tightness of the lower bound problem.

When to STOP:

- Lower bound:

$$\begin{aligned} \theta^* + \boldsymbol{c}'\boldsymbol{x}^* \in \operatorname*{argmin}_{\boldsymbol{x},\theta} \quad & \theta + \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}, \\ & \theta \geq \mathcal{G}(\hat{\boldsymbol{x}}_\ell) + \boldsymbol{s}_\ell'(\boldsymbol{x} - \hat{\boldsymbol{x}}_\ell), \ell = 1, 2, \cdots, L. \end{aligned}$$

- Upper bound:

$$\mathcal{G}(\boldsymbol{x}^*) + \boldsymbol{c}'\boldsymbol{x}^*.$$

- STOP if

$$\mathcal{G}(\boldsymbol{x}^*) - \theta^* \leq \epsilon.$$

STEP 0. Initial Setting.

1. Given the samples $\{w_i, i \in [N]\}$ and ϵ.
2. Find a lower bound θ_0 (e.g. -M) for

$$\mathcal{G}(\boldsymbol{x}) = \frac{1}{N} \sum_{i=1}^{N} \mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\hat{\omega}_i)).$$

3. Denote $\mathcal{C} = \{(\boldsymbol{x}, \theta) \mid \boldsymbol{x} \in \mathcal{X}, \theta \geq \theta_0\}$

STEP 1. Master Problem Solving.

1. Solving the master problem (lower bound problem):

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{Min}} \left\{\boldsymbol{c}'\boldsymbol{x} + \theta \mid (\boldsymbol{x},\theta)\in\mathcal{C}\right\}.$$

2. Return a solution $(\hat{\boldsymbol{x}},\hat{\theta})$.

- Remark:

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{Min}}\, \mathcal{G}(\boldsymbol{x}) + \boldsymbol{c}'\boldsymbol{x} \geq \underset{\boldsymbol{x}\in\mathcal{X}}{\text{Min}} \left\{\boldsymbol{c}'\boldsymbol{x} + \theta \mid (\boldsymbol{x},\theta)\in\mathcal{C}\right\} = \boldsymbol{c}'\hat{\boldsymbol{x}} + \hat{\theta}.$$

L-shaped Solution Approach

STEP 2. Identifying the Optimality Cuts.

1. Evaluating

$$\mathcal{G}(\hat{\boldsymbol{x}}) = \frac{1}{N}\sum_{i=1}^{N} \mathcal{Q}(\hat{\boldsymbol{x}}, \boldsymbol{\xi}(\omega_i)).$$

2. If $\mathcal{G}(\hat{\boldsymbol{x}}) \le \hat{\theta} + \epsilon$, STOP: $\hat{\boldsymbol{x}}$ is an ϵ-optimal solution.
3. Otherwise. for each $i \in [N]$ solve the dual sub-problem and obtain

$$\hat{\boldsymbol{p}}_i^* \in \underset{\boldsymbol{p}}{\operatorname{argmax}} \left\{ \boldsymbol{p}'(\boldsymbol{h}_i - \boldsymbol{T}_i \hat{\boldsymbol{x}}) : \boldsymbol{W}'\boldsymbol{p} \le \boldsymbol{q}_i \right\}.$$

   We can then form a subgradient: $\hat{\boldsymbol{s}} = -\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{T}_i' \hat{\boldsymbol{p}}_i^* \in \partial \mathcal{G}(\hat{\boldsymbol{x}})$. Then we find a lower bound for $\mathcal{G}(\boldsymbol{x})$ for any $\boldsymbol{x}$: $\mathcal{G}(\hat{\boldsymbol{x}}) + \hat{\boldsymbol{s}}'(\boldsymbol{x} - \hat{\boldsymbol{x}})$.
4. Update the optimality cut:

$$\mathcal{C} = \mathcal{C} \cap \left\{ (\boldsymbol{x}, \theta) : \theta \ge \mathcal{G}(\hat{\boldsymbol{x}}) + \hat{\boldsymbol{s}}'(\boldsymbol{x} - \hat{\boldsymbol{x}}) \right\}.$$

   Then go to **STEP 1.**

- The master (lower bound) problem

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\mathrm{Min}} \left\{\boldsymbol{c}'\boldsymbol{x}+\theta \mid (\boldsymbol{x},\theta)\in\mathcal{C}\right\} \nearrow \underset{\boldsymbol{x}\in\mathcal{X}}{\mathrm{Min}}\, \mathcal{G}(\boldsymbol{x})+\boldsymbol{c}'\boldsymbol{x}.$$

  improves when more cuts "$\mathcal{G}(\hat{\boldsymbol{x}})+\hat{\boldsymbol{s}}'(\boldsymbol{x}-\hat{\boldsymbol{x}})$" are added into feasible set $\mathcal{C}$ at each iteration.
- Computational Task 1: Solving the master problem as a small scaled LP.
- Computational Task 2: Evaluating $\mathcal{G}(\hat{\boldsymbol{x}})$ by solving N small scaled LPs $\mathcal{Q}(\hat{\boldsymbol{x}},\boldsymbol{\xi}(\omega_i)), i=1,2,\cdots,N$.
- Computational Task 3: Determining the subgradient $\hat{\boldsymbol{s}}$ by solving a small scaled dual sub-problem $\hat{\boldsymbol{p}}_i^*$.

- How to simplify the algorithm, if $\boldsymbol{q}_\omega \equiv \boldsymbol{q}$ and $\{\boldsymbol{p} : \boldsymbol{W}'\boldsymbol{p} \leq \boldsymbol{q}\}$ is bounded?
- How to modify the algorithm, if the assumption: for each feasible $\boldsymbol{x}$ and sample ω_i, $\mathcal{Q}(\boldsymbol{x}, \boldsymbol{\xi}(\omega_i)) < \infty$, is relaxed?
  Recall that when considering the following LP:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{Max}} & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{x} \in \boldsymbol{P} = \{\boldsymbol{x} \in \Re^n \mid \boldsymbol{A}\boldsymbol{x} \geq \mathbf{b}\}. \end{array}$$

Theorem (Optimality & Extreme Ray)

Suppose $\boldsymbol{P}$ has at least one extreme point. Then the optimal cost is $+\infty$ if and only if some **extreme ray** $\boldsymbol{d} \in \boldsymbol{P}$ *satisfies* $\mathbf{c}'\boldsymbol{d} > 0$.

- is Two-Stage SP the only choice for Adaptive Optimization?
- Only Two-Stages?

Outline

Conclusion

- Farmer's Problem and Facility Location-Allocation Problem as examples of Two-Stage Stochastic Programming
- Two types of decisions, dynamics induced by the uncertainty.
- SAA formulation as a practical approach, yet not scalable.
- Decomposition approach to handle the large scaled cases.

Reference and Further Reading

1. J.R. Birge, F. Louveaux, *Introduction to Stochastic Programming*, Springer, 1997.
2. A. Shapiro, A. Nemirovski, On complexity of stochastic programming problems, in: *On Continuous Optimization: Current Trends and Applications*, pp. 111-144, V. Jeyakumar and A.M. Rubinov (Eds.), Springer, 2005.
3. A. Shapiro, D. Dentcheva, A. Ruszczynski, Lectures on Stochastic Programming: Modeling and Theory, SIAM, Philadelphia, 2009.