Technical University of Denmark

02435 Decision-Making Under Uncertainty

Assignment 1

Jorge Montalvo Arvizu, s192184

Spring 2020

Introduction

Part I - Modelling and duality

Task 1 - Modelling a mixed-integer linear program

1. Explicit mathematical program

Variables

x_R	Space of planted roses $[m^2]$
x_D	Space of planted dahlias $[m^2]$
x_G	Space of planted garden pinks $[m^2]$

$$\text{Max } Z = 25x_R + 30x_D + 20G \tag{1}$$

$$\text{s.t. } x_R + x_D + x_G \leq 5000 \tag{2}$$

$$2.5x_R + 5x_D + 2x_G \leq 20000 \tag{3}$$

$$90x_R + 100x_D + 120x_G \leq 700000 \tag{4}$$

$$0.3x_R + 0.5x_D + 0.2x_G \leq 2500 \tag{5}$$

$$x_R, x_D, x_G \geq 0 \tag{6}$$

The objective function (1) maximizes the income from selling the flowers at the end of the season, where each variable is multiplied by the selling price. Constraint (2) refers to the available space limitation, defined as the sum of all three flowers distributed over a 5000 m^2 space. Constraints (3) to (5) refer to the available resources limitation and the respective resource usage per flower, i.e. working hours, water, and fertilizer. Finally, constraint (6) enforces non-negativity of the variables.

2. If the nursery plants roses, a greenhouse is built at a cost of 20000 EUR. To address this extension, a new **binary** variable is created with its respective constraint:

New variable

y_R	1 if roses are planted (and greenhouse built), 0 otherwise.

$$y_R \leq x_R \leq 700y_R \tag{7}$$

This new constraint (7) limits the space of planted roses between 0 and 700, since it can only be planted inside the greenhouse. Also, the following objective function replaces the previous objective function (1) given the cost of the (possibly) built greenhouse:

$$\text{Max } Z = 25x_R + 30x_D + 20G - 20000y_R \tag{8}$$

3. The third extension requires a space of at least 200 m^2 of roses, if the greenhouse is built. The following equation replaces constraint (7) by setting a boundary on the space of planted roses between 200 and 700:

$$200y_R \leq x_R \leq 700y_R \tag{9}$$

4. Finally, if the greenhouse is built, the remaining space of the greenhouse can only be used for roses and garden pinks but not dahlia. Then a new variable is introduced to represent garden pinks inside the greenhouse.

New variable

x_{GG}	Space of planted garden pinks (inside greenhouse) $[m^2]$

This new variable interacts with the other variables through the following added and replaced constraints:

$$x_G + x_D \leq 5000 - 700y_R \tag{10}$$

$$x_R + x_{GG} \leq 700y_R \tag{11}$$

$$x_R, x_D, x_G, x_{GG} \geq 0 \tag{12}$$

Constraint (2) is divided and replaced by constraints (10) and (11). Constraint (10) limits the space of dahlias and 'normal' garden pinks if the greenhouse is built, i.e. by removing 700 m^2 from the available space. While constraint (11) limits the space of roses and 'greenhouse' garden pinks to only the greenhouse space. Finally, constraint (6) is replaced by (12) to add the enforcement of non-negativity to 'greenhouse' garden pinks variable. At the end, the model is formed by equations (8), (10), (11), (3), (4), (5), (9), and (12).

Task 2 - Primal-dual transformation

For easiness on the dual transformation, the problem is first converted into its standard form:[1]

$$\text{Max } 5x_1 - 4x_2 + 3(x_3^+ - x_3^-)$$

$$\text{s.t. } 2x_1 - 3x_2 - (x_3^+ - x_3^-) \leq 5$$

$$-4x_1 + x_2 - 2(x_3^+ - x_3^-) \leq -11$$

$$-3x_1 + 4x_2 + 2(x_3^+ - x_3^-) \leq 8$$

$$6x_1 - 5x_2 + (x_3^+ - x_3^-) \leq 1$$

$$-6x_1 + 5x_2 - (x_3^+ - x_3^-) \leq -1$$

$$x_1, x_2, x_3^+, x_3^- \geq 0$$

Then, the primal problem in standard form is transformed to the dual problem formulation:

$$\text{Min } 5y_1 - 11y_2 + 8y_3 + y_4 - y_5$$

$$\text{s.t. } 2y_1 - 4y_2 - 3y_3 + 6y_4 - 6y_5 \geq 5$$

$$-3y_1 + y_2 + 4y_3 - 5y_4 + 5y_5 \geq -4$$

$$-y_1 - 2y_2 + 2y_3 + y_4 - y_5 \geq 3$$

$$y_1 + 2y_2 - 2y_3 - y_4 + y_5 \geq 3$$

$$y_1, y_2, y_3, y_4, y_5 \geq 0$$

Part II - Stochastic Programming

Task 3 - General Questions

1. The structure is divided in two: "Here-and-now" decisions/variables are used in the first-stage of the two-stage stochastic program, while "wait-and-see" or "recourse" decisions/variables are used in the second stage. This is done because, in a problem formulation, some of the decisions have to be taken **before** we know how the uncertainty will realize. Then, **after** the uncertainty is revealed, we can make second-stage decisions.

2. The main advantage between stochastic programming and deterministic optimization is that all uncertainty information is used by including the probabilities via scenarios in the model to get the expected value of the solution. By taking into account second-stage decisions, it is possible to obtain an "educated" expectation of the solution by also taking into account the uncertainty, since we have two different types of decisions: before uncertainty and after uncertainty.

[1]However, it is also possible to use the SOB principles to directly transform the problem to the dual formulation when the problem is NOT in standard form.

3. It really depends on the definition of "importance" when formulating the problem, e.g. let's assume first-stage variables are build or no build of certain asset, while second-stage variables affect the revenue of the asset - one investor could argue that the important decisions are the build or no build status of the asset, while an analyst might argue that the second-stage variables will affect the revenue of given asset, thus is more important to her/him. Given that there's more interest in the effect of uncertainty (that's why there's a need for stochastic programming, otherwise deterministic optimization would've been used) second-stage or wait-and-see decisions are more important in uncertainty-related problems. These second-stage variables will affect the outcome of the whole problem and give a good estimate of the expected solution, including first-stage decisions.

Task 4 - Modelling of a two-stage stochastic program

1. General two-stage stochastic program:

Sets	
$\mathcal{I}$	Set of locations
$\mathcal{L}$	Set of capacity levels
$\mathcal{M}$	Set of markets
$\mathcal{F}$	Set of furniture products
$\mathcal{T}$	Set of time periods [years]
$\mathcal{S}$	Set of scenarios
Parameters	
$b_{m,f,t,s}$	Forecasted demand per market m, product f, time period t, and scenario s [units]
c^T	Transport costs [$ per km and production units]
$d^M_{i,m}$	Distance between location and market [km]
$d^L_{i,j}$	Distance between locations [km]
c_i	Yearly operational costs [$]
c^B_l	Building costs per capacity level [$]
k^P_l	Production capacity per capacity level [units]
k^S_l	Storage capacity per capacity level [units]
π_S	Probability per scenario [-]
D	Minimum distance between locations [km]
$bigM$	Big M [-]
Binary variables	
$build_i$	Building of facility at location i [1/0]
$cap_{i,l}$	Assignment of capacity level l to location i [1/0]
$assignment_{i,m}$	Assignment of market m to location i [1/0]
Positive variables	
$q^P_{f,i,t,s}$	Production units of product f, per facility i, time period t, and scenario s
$q^T_{f,i,j,t,s}$	Transported units of product f, between facility i and j, per time period t, and scenario s
$q^S_{f,i,m,t,s}$	Market supply units of product f, per facility i, market m, time period t, and scenario s
$s^{in}_{f,i,t,s}$	Inflow storage units of product f, per facility i, time period t, and scenario s
$s^{out}_{f,i,t,s}$	Outflow storage units of product f, per facility i, time period t, and scenario s
$s^{level}_{f,i,t,s}$	Storage level units per product f, facility i, time period t, and scenario s

Mathematical model:

$$\begin{aligned}\text{Min } & \sum_{i\in\mathcal{I}} 10 \cdot c_i build_i + \sum_{i\in\mathcal{I}}\sum_{l\in\mathcal{L}} cap_{i,l} c_l^B + \sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{I}}\sum_{i\in\mathcal{I}}\sum_{f\in\mathcal{F}} \pi_S c^T q^T_{f,i,j,t,s} d^L_{i,j} + \\ & \sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}}\sum_{m\in\mathcal{M}}\sum_{i\in\mathcal{I}}\sum_{f\in\mathcal{F}} \pi_S c^T q^S_{f,i,m,t,s} d^M_{m,i}\end{aligned} \tag{13}$$

$$\text{s.t. } build_i = \sum_{l\in\mathcal{L}} cap_{i,l} \qquad \forall i \in \mathcal{I} \tag{14}$$

$$\sum_{m\in\mathcal{M}} assignment_{i,m} \leq 15 \cdot build_i \qquad \forall i \in \mathcal{I} \tag{15}$$

$$\sum_{i\in\mathcal{I}} assignment_{i,m} = 1 \qquad \forall m \in \mathcal{M} \tag{16}$$

$$d^L_{i,j} \geq D(-1 + build_i + build_j) \qquad \forall i \in \mathcal{I}, \forall j \in \mathcal{I}, i \neq j \tag{17}$$

$$\sum_{f\in\mathcal{F}} q^P_{f,i,t,s} \leq \sum_{l\in\mathcal{L}} cap_{i,l} k^P_l \qquad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \forall s \in \mathcal{S} \tag{18}$$

$$\sum_{f\in\mathcal{F}} s^{level}_{f,i,t,s} \leq \sum_{l\in\mathcal{L}} cap_{i,l} k^S_l \qquad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \forall s \in \mathcal{S} \tag{19}$$

$$\sum_{m\in\mathcal{M}} q^S_{f,i,m,t,s} = q^P_{f,i,t,s} + s^{out}_{f,i,t,s} - s^{in}_{f,i,t,s} + \sum_{\substack{j\in\mathcal{I}\\ j\neq i}} q^T_{f,j,i,t,s} - \sum_{\substack{j\in\mathcal{I}\\ j\neq i}} q^T_{f,i,j,t,s} \qquad \forall f \in \mathcal{F}, \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \forall s \in \mathcal{S} \tag{20}$$

$$q^S_{f,i,m,t,s} = b_{m,f,t,s} assignment_{i,m} \qquad \forall f \in \mathcal{F}, \forall i \in \mathcal{I}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T}, \forall s \in \mathcal{S} \tag{21}$$

$$s^{level}_{f,i,t,s} = s^{level}_{f,i,t-1,s} + s^{in}_{f,i,t,s} - s^{out}_{f,i,t,s} \qquad \forall f \in \mathcal{F}, \forall i \in \mathcal{I}, \forall t \in \mathcal{T}\{\backslash 1\}, \forall s \in \mathcal{S} \tag{22}$$

$$s^{level}_{f,i,1,s} = s^{in}_{f,i,1,s} - s^{out}_{f,i,1,s} \qquad \forall f \in \mathcal{F}, \forall i \in \mathcal{I}, \forall s \in \mathcal{S} \tag{23}$$

$$q^T_{f,i,j,t,s} <= bigM \cdot build_i \qquad \forall f \in \mathcal{F}, \forall i \in \mathcal{I}, \forall j \in \mathcal{I}, \forall t \in \mathcal{T}, \forall s \in \mathcal{S} \tag{24}$$

The objective function (13) minimizes the cost of:

- operational costs over 10 years,
- one-time building costs,
- uncertainty-influenced transportation costs between locations,
- and uncertainty-influenced supply costs from location to markets.

Constraint (14) assigns the capacity level to each location, constraint (15) assigns up to 15 markets to each location, constraint (16) assigns one market to one location only, constraint (17) enforces a minimum distance between built locations, constraint (18) sets the production of each facility up to the production capacity level, constraint (19) restrains the storage level of each facility up to the storage capacity level, constraint (20) is the balance equation of each location (where the storage, transportation, production, and supply flows are balanced), constraint (21) is the market balance equation (where the supply flow should be equal to the demand of all markets connected to each location), constraint (22) establishes the storage level given outflows and inflows, constraint (23) establishes the storage level at t=1, and constraint (24) enforces transportation flows only between built locations. The value of $bigM$ was chosen as $100 \cdot max(b)$ as the maximum demand that could be transported.

The results of the model are as follows:

```
Location:1,open,4
Location:2,not_open,0
Location:3,not_open,0
Location:4,open,3
Location:5,open,4
Market:1, 5
Market:2, 5
Market:3, 5
Market:4, 1
Market:5, 5
Market:6, 1
Market:7, 4
Market:8, 5
Market:9, 1
Market:10, 4
Market:11, 5
Market:12, 1
Market:13, 5
Market:14, 1
Market:15, 1
Objective:173211.125
```

Task 5 - Evaluation of stochastic programs

To evaluate the model, it is useful to compare expected value of three different solution approaches to obtain a quality metric:

- Wait-and-see approach (WS)
- Recourse program (RP)
- Expected value solution (EEV)

On the wait-and-see approach (WS), all scenarios are solved individually and the resulting expected value is calculated, i.e. the optimal solution is applied for each scenario after uncertainty is resolved.

On the expected value solution (EEV), the first-stage solution is applied to all scenarios and then the recourse decision after uncertainty is resolved, i.e. there's no use of information about uncertainty.

Finally, the stochastic solution (RP) applies first-stage solution and second-stage decisions after uncertainty is resolved, i.e. it makes use of information about uncertainty.

The results of the model are as follows ($WS \leq RP \leq EEV$):

Solution	Expected Value [EUR]
WS	52718.362
RP	66794.400
EEV	157450.784

Table 1: Expected value of each solution

Given that the model is a minimization problem, the following formulas are used to calculate the Expected Value of Perfect Information (EVPI) and the Value of Stochastic Solution (VSS):

$$EVPI_{min} = RP - WS \tag{25}$$

$$VSS_{min} = EEV - RP \tag{26}$$

Equation (25) refers to the amount the modeler, stakeholder or decision-maker is willing to pay to obtain perfect information about the future (a priori), i.e. what WS is using. On the other hand, equation (26) refers to the advantage of using a stochastic approach versus a deterministic approach, i.e. using uncertainty information in the model via the random variables instead of probability-weight and fix the first-stage variables.[2] The results are as follows:

Solution	Expected Value [EUR]
EVPI	14076.037
VSS	90656.384

Table 2: Quality metrics

Given that the objective function was in terms of costs [in EUR], it is convenient to say that a decision-maker is willing to pay 14076.037 EUR to obtain perfect information about the future, while the advantage of using an stochastic approach vs. a deterministic approach is 90656.385 EUR.

Conclusion

Over the course of this assignment, the modelling of optimization problems was approached from an stochastic point of view, where uncertainty played an important part of the problem. The fundamentals of MILP and duality theory were overview in the first part of the assignment, while the second part was focused on stochastic programming (two-stage specifically) and it's evaluation. It's important to notice that a even though we may have forecasts and estimated probabilities of certain problem, it's of utmost priority to be able to model properly the problem and make use of fundamental theory of optimization and stochastic programming to obtain a good representation of reality. Also is the evaluation of these models of equal importance, since we might be able to just use a 'cheap' deterministic model instead a complex stochastic program if it is sufficient. Personally, this assignment was very useful to really grasp the theory of the classroom and practice with an applied problem (as well as increasing my knowledge of the programming language Julia).

[2] Conejo, A. J., Carrion, M., Morales, J. M. (2010). Decision Making Under Uncertainty in Electricity Markets. Springer Science+Business Media, Chapter 2.6 and 4