MS-E2134 - Decision making and problem solving
Assignment 3 **Due: 12:15 pm 9.12.2021**

General instructions:

- *This assignment paper is accompanied by two MATLAB files. Please note that* Assignment3Code1.m, *uses function* intlinprog *which is included in Optimization Toolbox of MATLAB.*
- *This assignment follows a continuous narrative. It is recommended that you read the whole assignment through once before starting to solve the problems.*
- *The assignment report should answer to all the questions/tasks that are denoted with bracketed numbers from (1) to (19). Use full sentences and justify your answers with, e.g., mathematical formulations, whenever needed. Include your completed MATLAB files* Assignment3Code1.m *and* Assignment3Code2.m *as electric attachments in your submission.*
- *Hand-written report is OK as long as it is readable.*
- *Attach a cover page (available on the "Assignments" subpage in MyCourses) to your report.*
- *Include your name and student number on the cover page of your report.*
- *Submit the report as a pdf-file on "Assignments" subpage in MyCourses.*
- *You may discuss the problems with your fellow students, but everyone must submit their individual solutions.* ***Copying is strictly forbidden.***
- *The total score of the assignment is scaled linearly to exam points, 11 points being the maximum. For reports submitted after the deadline, one exam point is reduced from the scaled score for each starting 12 h period after the deadline.*

Purpose

The purpose of this assignment is to deepen the students' understanding of multi-attribute value theory, implementation of value models and how these models can be used to support decision making. Furthermore, the purpose is also to demonstrate the relation of these models to multi-objective optimization. Of course, in a pedagogical assignment, we cannot provide each student with a real decision maker, whose preferences should be elicited. Instead, the elicitation process is simulated by giving only clues about the preferences in three levels: preference relations, value differences and verbal statements. While there exists specialized software to carry out the sensitivity/robustness analyses, this assignment demonstrates that such analyses can in many cases be carried out using only general math software that can solve optimization problems.

Background

Dr. Jones works in a small forestry centre and is faced with a task of developing an additive multiattribute value function for the evaluation of forest sites' biodiversity values. Dr. Jones has been given nine evaluation attributes that contribute to the sites' biodiversity values. The attributes measure a forest site's value per hectare. The attributes together with their measurement units and scales are in Table 1. In addition to differing in these attributes, the sites differ in area and cost of conservation. The sites' properties are in Table 2. The ultimate task of Dr. Jones is to select a combination of forest sites for conservation, subject to a budget constraint of 25000 euros. Your task is to help Dr. Jones in constructing the additive evaluation model and making a cost-effective decision on the selected sites.

Data and notation

Attribute # i	a_i
Site # j	x^j
Measurement of site x^j w.r.t. attribute a_i	x_i^j
Least preferred level w.r.t. attribute a_i	x_i^0
Most preferred level w.r.t. attribute a_i	x_i^*
[0,1]-normalized value of site x^j w.r.t. attribute a_i	$v_i(x_i^j)$
Attribute weight of a_i	w_i
Cost of site x^j (€)	c_j
Area of site x^j (ha)	A_j

Table 1. Attributes and measurement units.

Attribute	Attribute name	Measurement unit	x_i^0	x_i^*
a_1	Old aspens	m^3/ha	0	40
a_2	Other old broad-leaved trees	m^3/ha	0	30
a_3	Old scots pines	m^3/ha	2	30
a_4	Logs	m^3/ha	2	20
a_5	Conifer snags	m^3/ha	0	20
a_6	Broad-leaved snags	m^3/ha	0	20
a_7	Burned wood	m^3/ha	0	100
a_8	Endangered species*	number	0	100
a_9	Distance to the closest conservation site*	km	100	0

* also measured as value per hectare. Thus, for example, a distance of 1 km to the nearest conservation site is more valuable on a 1.1 ha site than on a 1.0 ha site.

Table 2. Sites' properties.

Site	c_j	A_j	x_1^j	x_2^j	x_3^j	x_4^j	x_5^j	x_6^j	x_7^j	x_8^j	x_9^j
x^1	2400	1.2	0	25	13	12	4	11	0	1	4
x^2	5700	3.0	12	11	8	3	15	5	0	2	0
x^3	4725	2.1	16	5	2	11	20	20	0	0	0.8
x^4	7800	3.0	25	27	4	2	17	2	0	0	1.2
x^5	1560	0.8	38	12	12	20	16	1	12	0	3.5
x^6	4500	2.0	40	1	24	2	19	16	0	1	1
x^7	6150	3.0	12	2	30	9	3	10	0	9	2
x^8	1728	0.9	16	24	3	9	8	10	0	5	4
x^9	1859	1.1	19	29	12	14	8	16	16	0	10
x^{10}	4800	2.4	5	14	22	6	4	4	2	1	0.6

Attribute-specific value functions

First, Dr. Jones specifies attribute-specific value functions. She has answered to the following elicitation questions, but your help is needed in converting her answers into value functions.

1. The value functions with regard to *old aspens* (a_1), *logs* (a_4), *conifer snags* (a_5), *broad-leaved snags* (a_6) and *burned wood* (a_7) are linear between the least and most preferred levels.
2. Dr. Jones makes the following comparisons:
   - In *other old broad-leaved trees* (a_2), all values between 0 m^3/ha and 2 m^3/ha are equally preferred. The improvement from 2 m^3/ha to 6 m^3/ha is equally preferred to the improvement from 6 m^3/ha to 15 m^3/ha, which is equally preferred to the improvement from 15 m^3/ha to 30 m^3/ha.
   - In *old scots pines* (a_3), the improvements from 2 m^3/ha to 9 m^3/ha and from 9 m^3/ha to 30 m^3/ha are equally preferred.

   Between these points of reference, the value function is approximated to be linear.
3. In *endangered species* (a_8), $v_8(k+1) - v_8(k) = 0.5\big(v_8(k) - v_8(k-1)\big)$ for $k = 1, \dots, 9$, $v_8(k) = v_8(k+1)$ for $k > 9$. (That is, $v_8(9) < v_8(10) = v_8(11)$.)
4. In *distance to the closest conservation site* (a_9), it applies $v_9(x_9) = B + A(1 - \exp[-(3 - x_9)/r])$ for $x_9 \in [0, 3]$ (A, B, and r are real-valued coefficients), and $v_9(x_9) = v_9(3)$ for any $x_9 > 3$. Furthermore, the improvement from 1 to 0 km is equally preferred to the improvement from 3 to 1 km.

(1) (9 pts) Report value functions for each attribute in closed form. Report mathematics applied.
- *HINT*: *The derivation of v_8 requires using the formula for the sum of a geometric series.*
- *HINT: In the derivation of v_9, the substitution $u := \exp(-1/r)$ at some point is helpful as it provides you with a simple cubic equation to be solved.*
- *HINT*: *Here's a list of values of the functions at given points to validate your results:*
  $v_1(5) = 0.125, v_2(14) = 0.6296, v_3(22) = 0.8095, v_4(6) = 0.2222,$
  $v_5(4) = 0.2, v_6(16) = 0.8, v_7(2) = 0.02, v_8(5) = 0.9697, v_9(0.6) = 0.6717$

(2) (2 pts) Compute and report attribute values (per hectare) for all alternatives, i.e., the values $v_i(x_i^j)$ for $i = 1, \dots, 9, j = 1, \dots, 10$.

(3) (2 pts) Plot value functions v_2 and v_9.

Attribute weights

Dr. Jones gives the following preference statements:

1. If the improvement from 0 to 20 m3/ha in *conifer snags* (a_5) is worth 40 points, then the corresponding improvement in *broad-leaved snags* (a_6) is worth 45 points.
2. Compared to the previous, the improvement from 2 to 20 m3/ha in *logs* (a_4) is worth 30 points.
3. Improvement from no endangered species to 1 endangered species in a_8 is equally preferred to an improvement of 8 m^3/ha in *logs* (a_4), whenever the amount of *logs* is between 2 and 12 m^3/ha.
4. Additional 1 m^3/ha of *burned wood* (a_7) is equally preferred to additional 1 m^3/ha of *logs* (a_4), whenever the amount of *burned wood* is between 0 and 99 m^3/ha and the amount of *logs* is between 2 and 19 m^3/ha.

5. Improvement from 0 to 40 m³/ha in *old aspens* (a_1) is equally preferred to improvements 0 to 30 m³/ha in *other old broad-leaved trees* (a_2) and 2 to 20 m³/ha in *old scots pines* (a_3) together.
6. Alternatives $(40, x_2, x_3, x_4, x_5, x_6, x_7, x_8, 1.2)$ and $(10, x_2, x_3, x_4, x_5, x_6, x_7, x_8, 0)$ are equally preferred.
7. $((x_1, x_2, x_3, 10, x_5, x_6, x_7, x_8, 1.2) \leftarrow x^0) \sim_d ((x_1, x_2, x_3, 18, x_5, x_6, x_7, x_8, 3) \leftarrow x^0)$.
8. Improvement from 0 to 15 m³/ha in *other old broad-leaved trees* (a_2) is equally preferred to the improvement from 2 to 30 m³/ha in *old scots pines* (a_3).

(4) (10 pts) Derive attribute weights that sum up to one. Report equations you needed in obtaining these weights.

Overall values

(5) (1 pts) Use the normalized scores from (2) and attribute weights from (4) and compute overall values for the sites. Note that the constructed model evaluates sites' properties per hectare – you need to multiply these values by area to get the overall values $V(x^j)$:

$$V(x^j) = A_j \sum_{i=1}^{9} w_i v_i(x_i^j) = \sum_{i=1}^{9} w_i A_j v_i(x_i^j)$$

Dr. Jones realizes that the most preferred level with regard to *burned wood* (a_7) is set so high that none of the sites is even close to achieving it. She decides to change this level to 40 m³/ha.

(6) (3 pts) Update normalized scores, compute new attribute weights w_i and overall values $V'(x^j)$. Confirm numerically that V' is an affine transformation of V.

From now on, *V* denotes the value function, where the most preferred level with regard to *burned wood* (a_7) is 40 m³/ha.

Site combination selection

Dr. Jones makes an assumption that the value $V(p)$ of a site combination $p \subseteq \{x^1, x^2, \dots, x^{10}\}$ is the sum of overall values of the sites that are included in this combination. The cost $c(p)$ of p is the sum of the costs of the sites included in p. Thus, it applies

$$V(p) = \sum_{x^j \in p} V(x^j), \quad c(p) = \sum_{x^j \in p} c_j$$

For example, the value of the site combination $p = \{x^1, x^3\}$ is $V(x^1) + V(x^3)$ and the cost is $c_1 + c_3$.

Dr. Jones plans to select the combination of sites, whose overall value is the highest, subject to the requirement that the site combination's cost $c(p)$ cannot exceed the budget of 25000 €.

(7) (2 pts) Formulate and report an integer linear programming (ILP) problem that solves this site combination and solve it. Which sites are recommended?
- *HINT: You can use for example MS Excel's Solver here. Select Simplex LP as the solving method in Solver. Furthermore, make sure that the box "Ignore Integer*

Constraints" in "Options"->"All methods" is unchecked. You will have a possibility to validate your results later in this assignment with MATLAB.

Multi-objective optimization applied to site combination selection

Dr. Jones has a meeting with her boss, Dr. Current, who had met with biodiversity experts. These experts had constructed a model that was very similar to that of Dr. Jones – but the experts had used slightly different weights! While thinking of ways that would relieve them from pondering too much the appropriate weights, Dr. Current realizes that

$$V(p) = \sum_{x^j \in p} V(x^j) = \sum_{x^j \in p} \sum_{i=1}^{9} w_i A_j \, v_i(x_i^j) = \sum_{i=1}^{9} w_i \sum_{x^j \in p} A_j v_i(x_i^j)$$

$$= \sum_{i=1}^{9} w_i \sum_{j=1}^{10} y_j A_j v_i(x_i^j) =: \sum_{i=1}^{9} w_i \, f_i(y), \qquad y_j \in \{0,1\} \; \forall \, j = 1, \ldots, 10$$

That is, the value of a given site combination $V(p)$ can be represented as a weighted sum of 9 functions f_i of the binary vector variable $y = (y_1, \ldots, y_{10})$ where $y_j = 1$ if the site x^j is included in the combination p and $y_j = 0$ if the site is not included in the combination. Dr. Current understands then immediately that this formulation of $V(p)$ is associated to the use of the weighted sum algorithm in the search of Pareto optimal solutions in multi-objective optimization (MOO). Based on her realization, Dr. Current suggests: "Let's model the site selection problem as a MOO problem and solve Pareto optimal site combinations. If we are lucky, there aren't that many of them and we can easily select a suitable site combination from among them. We can formulate the MOO problem as the following 9-objective binary integer linear programming model:"

$$\operatorname*{v-max}_{y \in \{0,1\}^{10}} \begin{bmatrix} \sum_{j=1}^{m} y_j A_j v_1(x_1^j) \\ \vdots \\ \sum_{j=1}^{m} y_j A_j v_9(x_9^j) \end{bmatrix} \qquad (*)$$

$$\text{s.t.}$$

$$\sum_{j=1}^{m} c_j y_j \leq 25000$$

To complete the remaining parts of the assignment, save the two MATLAB files to your computer. Make sure that all files are copied to the same folder. Set the folder as your Current Folder in MATLAB, write "help intlinprog" to MATLAB's Command Window, and familiarize yourself with the syntax of the function. If the command causes errors, you need to install the Optimization Toolbox.

(8) (4 pts) Complete the missing parts of the code Assignment3Code1.m (3 instances marked with XXXXX) and use it to find Pareto optimal solutions to the MOO problem. Include the resulting figures in your report.

(9) (1 pt) How many Pareto optimal site combinations can you find? Is the solution to (7) among them? Do the results change if you rerun the computations?

(10) (1 pt) Which site has the highest core index? Which site has the lowest core index?

Robustness analysis of site combination selection

After being blown away by the number of Pareto optimal site combinations discovered, Dr. Current now wants Dr. Jones to perform a sensitivity analysis on the original attribute weights. Dr. Jones browses through some literature and realizes that a multi-way sensitivity analysis of nine attributes is very difficult to conduct. Towards this end, she decides to rather loosen her preference statements. Let the attribute weights from (6) be denoted by w_i^*. She selects attribute a_7 to work as a reference and computes weight ratios $r_i = w_i^*/w_7^*$ for the other attributes (use the weights and the corresponding scores from (6)). She then allows the weights vary such that $w_i/w_7 \in [0.8r_i, 1.25r_i]$. The weights must sum up to one.

(11) (2 pts) Formulate and report a mathematical expression for the feasible region S of the attribute weights with linear constraints.

Dr. Jones realizes that her site selection model can no longer be applied. Your help is needed to derive decision recommendations in this setting, where attribute weights are only known to be in the set S.

The MATLAB code Assignment3Code2.m enumerates all possible site combinations. These are your decision alternatives now. Complete the missing parts of the code (6 instances marked with XXXXX) as follows:

- Multiply the normalized scores $v_i(x_i^j)$ by area A_j and add them to the code.
- Add the constraints on the attribute weights determined in (11) as matrix A_w and vector B_w so that $A_w w \leq B_w$.
- Complete the code so that it identifies *feasible* combinations p that do not exceed the budget constraint.
- Complete the code so that it identifies *dominated* combinations given the preference information represented by the set S.
    - *HINT: A feasible combination is dominated if another feasible combination has a greater overall value for all weights in S and strictly greater overall value for some weights in S. If a feasible combination is not dominated by any other feasible combination, it is non-dominated.*
    - *HINT: The code uses the MATLAB built-in function "linprog" to minimize and maximize the value difference of (all) pairs of feasible site combinations over S. You need to formulate the objective functions to these LPs (type "help linprog" into Command Window for help on the syntax).*
- Complete the code so that it identifies *dominated* combinations given the preference information represented by the set $S^0 = \{w \in \mathbb{R}^9 | w_i \in [0,1], \sum_{i=1}^{9} w_i = 1\}$ (That is, no preference information has been given).

(12) (2 pts) Present your A_w and B_w (e.g., as simple tables).
(13) (1 pt) What is the number of feasible site combinations?
(14) (1 pt) What is the number of non-dominated site combinations when S represents the set of feasible weights?
(15) (1 pt) Which sites belong to all of the non-dominated combinations determined in (14)?
(16) (1 pt) Why is the solution to (7) among the non-dominated combinations determined in (14)?

(17) (1 pt) What is the number of non-dominated site combinations when S^0 represents the set of feasible weights?

Mathematically, the set of Pareto optimal solutions related to Problem (*) is the same as the set of non-dominated site combinations when S^0 is the set of feasible weights.

(18) (1 pt) Did you find solutions in (8) which are not in the set of non-dominated combinations determined in (17)? Why can this happen?

(19) (1 pt) How many of the non-dominated combinations determined in (17) are not found from the set of solutions determined in (8)? What is the reason for this difference?