Decision Theory

Lecture 8

Michael Hartisch

Friedrich-Alexander Universität Erlangen-Nürnberg, Department Data Science June 10, 2024

Decision Theory

Organization

- decision theory only for Data Science specialization "optimization"
- evaluation

Decision Theory

Recap: What did we do?

- Decision under uncertainty
- Methods:
  - Minimax/Maximin
  - Minimax Regret
  - Hurwicz
  - Average
- Properties

Decision Theory

Decision Under Uncertainty

	s_1	s_2	$\dots$	s_n
a_1	e_{11}	e_{12}	$\dots$	e_{1n}
a_2	e_{21}	e_{22}	$\dots$	e_{2n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_m	e_{m1}	e_{m2}	$\dots$	e_{mn}

- Any of the $s_1, \dots, s_n$ might happen
- We do not know any likelihoods
- Which a_i to choose?

Decision Theory

Approach 1: Maximin/Minimax

- Maximisation problem: maximin
  - = maximise the smallest (worst) value
- Minimisation problem: minimax
  - = minimise the largest (worst) value
- Evaluate a row (decision) by its worst value
- Finds the best worst-case guarantee

Decision Theory

Approach 2: Minimax Regret

- For each column (scenario) find the best value
- Consider the difference to the best in each column
- Choose alternative that minimizes the largest difference

Example

Maximise:

	s_1	s_2
a_1	3	7
a_2	5	2

Regret:

	s_1	s_2
a_1	2	0
a_2	0	5

- Max regret of a_1 is 2
- Max regret of a_2 is 5
- Choose a_1

Decision Theory

Approach 3: Hurwicz's Pessimism-Optimism Index

- Given trade-off value $\alpha \in [0, 1]$
- For each alternative i:
  - Determine the worst outcome m_i
  - Determine the best outcome M_i
  - Calculate $\alpha m_i + (1 - \alpha)M_i$
- Compromise between best and worst
- $\alpha = 1$: optimise worst-case
- $\alpha = 0$: optimise best-case

Decision Theory

Approach 4: Laplace's Principle of Insufficient Reason (Average)

- Assumption:
  - If we don't know anything about probabilities, just assume all options are equally likely
- Optimise with respect to average value

Decision Theory

Example

Maximise:

	s_1	s_2	s_3	min	max	Hurwicz	regret	average
a_1	3	3	4					
a_2	2	3	1					
a_3	6	2	2					
a_4	5	2	3					

- Maximin:
- Hurwicz $\alpha = 1/2$:
- Regret:
- Average:

Decision Theory

Example

Maximise:

	s_1	s_2	s_3	min	max	Hurwicz	regret	average
a_1	3	3	4	3	4	7/2	3	10/3
a_2	2	3	1	1	3	4/2	4	6/3
a_3	6	2	2	2	6	8/2	2	10/3
a_4	5	2	3	2	5	7/2	1	10/3
max	6	3	4					

- Maximin: a_1
- Hurwicz $\alpha = 1/2$: a_3
- Regret: a_4
- Average: a_1, a_3 or a_4

Decision Theory

Axioms

- Axiom 1: Complete ranking
- Axiom 2: Independence of labelling
- Axiom 3: Independence of value scale
- Axiom 4: Strong domination
- Axiom 5: Independence of irrelevant alternatives
- Axiom 6: Independence of addition of a constant to a column
- Axiom 7: Independence of row permutation
- Axiom 8: Independence of column duplication

Decision Theory

Table of Results

	Minimax	Hurwicz	Regret	Average
Ax 1	✓	✓	✓	✓
Ax 2	✓	✓	✓	✓
Ax 3	✓	✓	✓	✓
Ax 4	✓	✓	✓	✓
Ax 5	✓	✓	✗	✓
Ax 6	✗	✗	✓	✓
Ax 7	✓	✓	✗	✓
Ax 8	✓	✓	✓	✗

Decision Theory

No Free Lunch

- Any criterion that fulfils
  - Axiom 1: Complete ranking
  - Axiom 4: Strong domination
  - Axiom 5: Independence of irrelevant alternatives
  - Axiom 6: Independence of addition of a constant to a column
  - Axiom 7: Independence of row permutation

  is equivalent to the average criterion
- The average criterion does not fulfil Axiom 8
- There is no criterion that fulfils all axioms

Decision Theory

Warm-Up

Find a counterexample to axiom 6: ”adding a constant to a column does not change the solution”

- Minimax does not satisfy axiom 6
- Hurwicz does not satisfy axiom 6

Decision Theory

Warm-Up

Minimize:

	s_1	s_2
a_1	1	1
a_2	2	0

- Minimax: choose a_1
- The same for Hurwicz $\alpha = 1$

	s_1	s_2
a_1	1	3
a_2	2	2

- Minimax: choose a_2
- The same for Hurwicz $\alpha = 1$

Decision Theory

Today

- Last session: Decision under Uncertainty
- Ordered Weighted Averaging (OWA)

Decision Theory

OWA

- Let a weight vector w be given
$$w = (w_1, \ldots, w_n)$$
with $w_j \in [0, 1]$, $\sum_{j \in [n]} w_j = 1$
- For $A = (a_1, \ldots, a_n)$, let the vector $B = (b_1, \ldots, b_n)$ be all entries from a ordered by size (from large to small)
- A function $F : \mathbb{R}^n \to \mathbb{R}$ is called an Ordered Weighted Averaging (OWA) operator if
$$F(a_1, \ldots, a_n) = w_1 b_1 + w_2 b_2 + \ldots + w_n b_n$$
- Weights w_i are bound to the position, not the scenario

Decision Theory

Example

- Let $w = (0.2, 0.3, 0.1, 0.4)$
- Find $F(6,\ 10,\ 3,\ 5)$
  - Ordered: $B = (10,\ 6,\ 5,\ 3)$

$$\begin{aligned}&F(6, 10, 3, 5)\\ =&0.2 \cdot 10 + 0.3 \cdot 6 + 0.1 \cdot 5 + 0.4 \cdot 3\\ =&5.5\end{aligned}$$

- Find $F(0,\ 7,\ 10,\ 2)$
  - Ordered: $B = (10,\ 7,\ 2,\ 0)$

$$\begin{aligned}&F(0, 7, 10, 2)\\ =&0.2 \cdot 10 + 0.3 \cdot 7 + 0.1 \cdot 2 + 0.4 \cdot 0\\ =&4.3\end{aligned}$$

Decision Theory

Example

Maximize. Calculate OWA for each alternative with $w = (1/2,\ 1/4,\ 0,\ 1/4)$

	s_1	s_2	s_3	s_4
a_1	6	3	5	2
a_2	4	8	2	0
a_3	3	2	5	6

- $D(a_1)$
- $D(a_2)$
- $D(a_3)$

Decision Theory

Example

Maximize. Calculate OWA for each alternative with $w = (1/2,\ 1/4,\ 0,\ 1/4)$

	s_1	s_2	s_3	s_4
a_1	6	3	5	2
a_2	4	8	2	0
a_3	3	2	5	6

- $D(a_1) = F(6, 3, 5, 2) = 0.5 \cdot 6 + 0.25 \cdot 5 + 0.25 \cdot 2 = 4.75$
- $D(a_2) = F(4, 8, 2, 0) = 0.5 \cdot 8 + 0.25 \cdot 4 + 0.25 \cdot 0 = 5$
- $D(a_3) = F(3, 2, 5, 6) = 0.5 \cdot 6 + 0.25 \cdot 5 + 0.25 \cdot 2 = 4.75$

Decision Theory

Properties

- $0 \leq F(A) \leq 1$ if all values are in $[0, 1]$
- Why?

$$F(A) = \sum_{i \in [n]} w_i b_i \leq \sum_{i \in [n]} w_i b_1 = b_1 \leq 1$$

$$F(A) = \sum_{i \in [n]} w_i b_i \geq \sum_{i \in [n]} w_i b_n = b_n \geq 0$$

Decision Theory

Properties

- Let $A = (a_1, \ldots, a_n)$
- Set $A' = (a'_1, \ldots, a'_n)$ with
$$a_j \geq a'_j \text{ for all } j \in [n]$$
- Then $F(A) \geq F(A')$
- If each individual value improves, the result improves

Why?

- Let B and B' be the corresponding sorted vectors
- Then $b_j \geq b'_j$
- Thus, $\sum_{i\in[n]} w_i b_i \geq \sum_{i\in[n]} w_i b'_i$

Decision Theory

Properties

- $F(a_1, \dots, a_n) = F(a_{\pi(1)}, \dots, a_{\pi(n)})$ for all permutations π
- Thus: order of scenarios doesn't matter

Why?

- According to the definition of OWA, values are sorted anyway

Decision Theory

Remark

- If we only assign weights w_j to each scenario j (not to positions), the property does not hold
- So:
$$G(a_1, \ldots, a_n) = \sum_{j \in [n]} w_j a_j$$
does not satisfy independence of column ordering
- Example: $G(a_1, a_2) = 0.7a_1 + 0.3a_2$
- $G(1, 0) \neq G(0, 1)$

Decision Theory

Special Cases

- Reminder: w_1 is the weight for the highest outcome, w_n for the lowest outcome
- What if ... (how to maximize)
  - $w_j = 1/n$ for all j?
  - $w_1 = 1$, $w_j = 0$ for all $j \neq 1$?
  - $w_n = 1$, $w_j = 0$ for all $j \neq n$?
  - $w_n = \alpha$, $w_1 = 1 - \alpha$, $w_j = 0$ for all $j \neq 0, n$?

Decision Theory

Special Cases

- Reminder: w_1 is the weight for the highest outcome, w_n for the lowest outcome
- What if ... (how to maximize)
  - $w_j = 1/n$ for all j? Average
  - $w_1 = 1$, $w_j = 0$ for all $j \neq 1$? Best case
  - $w_n = 1$, $w_j = 0$ for all $j \neq n$? Worst case
  - $w_n = \alpha$, $w_1 = 1 - \alpha$, $w_j = 0$ for all $j \neq 0, n$? Hurwicz
- OWA generalizes Average, Minimax, Hurwicz

Decision Theory

Properties

- Let $\underline{w} = (0, 0, \ldots, 0, 1)$ and $\overline{w} = (1, 0, 0, \ldots, 0)$
- Let $F_w(a_1, \ldots, a_n)$ be the OWA operator with respect to weight w
- Then:
  - $F_{\underline{w}}(a_1, \ldots, a_n) = \min_{j \in [n]} a_j$
  - $F_{\overline{w}}(a_1, \ldots, a_n) = \max_{j \in [n]} a_j$
  - $F_{\underline{w}}(a_1, \ldots, a_n) \leq F_w(a_1, \ldots, a_n) \leq F_{\overline{w}}(a_1, \ldots, a_n)$ for any w

Decision Theory

How do I determine w? First Approach

- Let $S_k = \sum_{j=1}^{k} w_j$
- $S_n = 1$
- $S_0 := 0$
- Assuming we have a sorted vector B such that $b_j = 1$ for $j \leq k$, and $b_j = 0$ otherwise
- Inquiry: How good do you find it if the best outcome is achieved in k out of n scenarios, and the remaining $n - k$ are worst possible?
- It is $F(B) = \sum_{j \in [n]} w_j b_j = \sum_{j=1}^{k} w_j = S_k$

Decision Theory

How do I determine w? First Approach

- S_k corresponds to the utility when k out of n scenarios are good
- $w_j = 1/n$: linear utility gain
- $\underline{w}$: no satisfaction unless all scenarios are good
- $\overline{w}$: satisfaction even if one scenario is already good
- If S_k can be determined, then set

$$w_j = S_j - S_{j-1}$$

Decision Theory

How do I determine w? Second Approach

- Given observations on how someone has decided
- Based on that, estimate weights
- Let $\mathcal{B} = (B^1, \dots, B^K)$ with $B^k = (b^k_1, \dots, b^k_n)$ be some alternatives we have observed so far
- Let $R \subseteq \mathcal{B} \times \mathcal{B}$ be pairs of alternatives where the decision-maker prefers the first alternative to the second alternative

How do I determine w? Second Approach

- OWA of B^k is $\sum_{j=1}^{n} w_j b_j^k$
- Problem: find w such that $F_w(B^i) \geq F_w(B^k)$ for all $(B^i, B^k) \in R$
- Might not be possible $\rightarrow$ Optimization problem

$$
\begin{aligned}
\max\ & z \\
\text{s.t. } & \sum_{j=1}^{n} w_j b_j^i - \sum_{j=1}^{n} w_j b_j^k \geq z && \forall (B^i, B^k) \in R \\
& \sum_{j=1}^{n} w_j = 1 \\
& w_j \geq 0 && \forall j \in [n]
\end{aligned}
$$

”Andness” and ”Orness”

- OWA can also be considered as a mixture of ”And” (andness) and ”Or” (orness)
- $w_n = 1$: each scenario must be good – ”and”
- $w_1 = 1$: one scenario must be good – ”or”
- Define ”orness” as

$$orness(w) = \frac{1}{n-1} \sum_{j \in [n]} w_j (n - j)$$

- $andness(w) = 1 - orness(w)$

Decision Theory

”Andness” and ”Orness”

$$orness(w) = \frac{1}{n-1} \sum_{j \in [n]} w_j (n-j)$$

- $\underline{w} = (0, 0, \ldots, 0, 1)$,

$$orness(\underline{w}) = \frac{1}{n-1} \cdot 0 = 0$$

- $\overline{w} = (1, 0, \ldots, 0)$

$$orness(\overline{w}) = \frac{1}{n-1} \cdot (n-1) = 1$$

Decision Theory

"Andness" and "Orness"

$$orness(w) = \frac{1}{n-1} \sum_{j \in [n]} w_j (n-j)$$

- $w = (1/n, 1/n, \ldots, 1/n)$,

$$\begin{aligned} orness(w) &= \frac{1}{n-1} \cdot \frac{1}{n} \cdot \sum_{j \in [n]} (n-j) \\ &= \frac{1}{n-1} \cdot \frac{1}{n} \cdot \frac{n(n-1)}{2} = \frac{1}{2} \end{aligned}$$

Decision Theory

"Andness" and "Orness"

$$orness(w) = \frac{1}{n-1} \sum_{j \in [n]} w_j (n - j)$$

- What does this weight vector do?

$$w = (0, 0, 1, 0, 0)$$

Decision Theory

”Andness” and ”Orness”

$$orness(w) = \frac{1}{n-1} \sum_{j \in [n]} w_j(n-j)$$

- What does this weight vector do?

$$w = (0, 0, 1, 0, 0)$$

- Chooses the median (the moderately good scenario)
- What is its *orness*?

”Andness” and ”Orness”

$$orness(w) = \frac{1}{n-1} \sum_{j \in [n]} w_j(n-j)$$

- What does this weight vector do?

$$w = (0, 0, 1, 0, 0)$$

- Chooses the median (the moderately good scenario)
- What is its *orness*?

$$orness(w) = \frac{1}{4} \cdot (5-3) = \frac{1}{2}$$

OWA for Combinatorial Problems

- So far: OWA for a single alternative
- Approach possible for decision matrix and combinatorial problems
- So far only maximization problems, also possible for minimization problems
- If maximizing: $w = (1, 0, \ldots, 0)$ chooses the largest – best case
- If minimizing: $w = (1, 0, \ldots, 0)$ chooses the smallest – worst case

Decision Theory

OWA for Combinatorial Problems

- Calculate minimal OWA solution with $w = (1/2, 1/2, 0, 0)$, choose 2 out of 3 items

item	c^1	c^2	c^3	c^4
1	3	4	3	5
2	4	2	5	1
3	4	2	4	5

- $\{1, 2\}$:
- $\{2, 3\}$:
- $\{1, 3\}$:

Decision Theory

OWA for Combinatorial Problems

- Calculate minimal OWA solution with $w = (1/2, 1/2, 0, 0)$, choose 2 out of 3 items

item	c^1	c^2	c^3	c^4
1	3	4	3	5
2	4	2	5	1
3	4	2	4	5

- $\{1,2\}$: $A = (7,6,8,6)$, $OWA(\{1,2\}) = 7.5$
- $\{2,3\}$: $A = (8,4,9,6)$, $OWA(\{2,3\}) = 8.5$
- $\{1,3\}$: $A = (7,6,7,10)$, $OWA(\{1,3\}) = 8.5$

Decision Theory

OWA for Combinatorial Problems

- Calculate minimal OWA solution with $w = (2/3, 1/3, 0, 0)$, choose 2 out of 3 items

item	c^1	c^2	c^3	c^4
1	3	7	2	4
2	2	1	8	3
3	3	4	7	1

- $\{1,2\}$:
- $\{2,3\}$:
- $\{1,3\}$:

Decision Theory

OWA for Combinatorial Problems

- Calculate minimal OWA solution with $w = (2/3, 1/3, 0, 0)$, choose 2 out of 3 items

item	c^1	c^2	c^3	c^4
1	3	7	2	4
2	2	1	8	3
3	3	4	7	1

- $\{1,2\}$: $A = (5, 8, 10, 7)$, $OWA(\{1,2\}) = 9.3$
- $\{2,3\}$: $A = (5, 5, 15, 4)$, $OWA(\{2,3\}) = 11.7$
- $\{1,3\}$: $A = (6, 11, 9, 5)$, $OWA(\{1,3\}) = 10.3$

Decision Theory

Approximation

- If minimizing, let the weights w_i be non-decreasing:

$$w_1 \geq w_2 \geq \ldots \geq w_n$$

- i.e., we are conservative / risk-averse: worse outcomes are always more important than better ones
- Worst-case with $w_1 = 1$ is a special case
- For item i, let $\hat{c}_{i1} \geq \hat{c}_{i2} \geq \ldots \hat{c}_{in}$ be the sorted cost vector over all scenarios
- Set $\hat{c}_i = \sum_{j \in [n]} w_j \hat{c}_{ij}$

Decision Theory

Approximation

- Solve with respect to scenario $\hat{c}$
- The solution is a $w_1 \cdot n$-approximation for OWA
- Meaning, the objective function value is at most a factor of $w_1 n$ larger than the optimum
- Worst-case: n-approximation

Decision Theory

Example

- Calculate minimal OWA solution with $w = (2/3, 1/3, 0, 0)$, choose 2 out of 3 items

item	c^1	c^2	c^3	c^4	$\hat{c}$
1	3	7	2	4	6.0
2	2	1	8	3	6.3
3	3	4	7	1	6.0

- $\hat{c}_1 = w_1 c_1^2 + w_2 c_1^4 + w_3 c_1^1 + w_4 c_1^3 = (2/3) \cdot 7 + (1/3) \cdot 4 = 6$
- Heuristic: include $\{1, 3\}$ (unfortunately not optimal)
- Approximation guarantee: $w_1 n = 2.7$

Decision Theory

Extension: WOWA

- We have seen so far: OWA generalizes Maximin, Hurwicz, Average
- A weighting of scenarios of the form

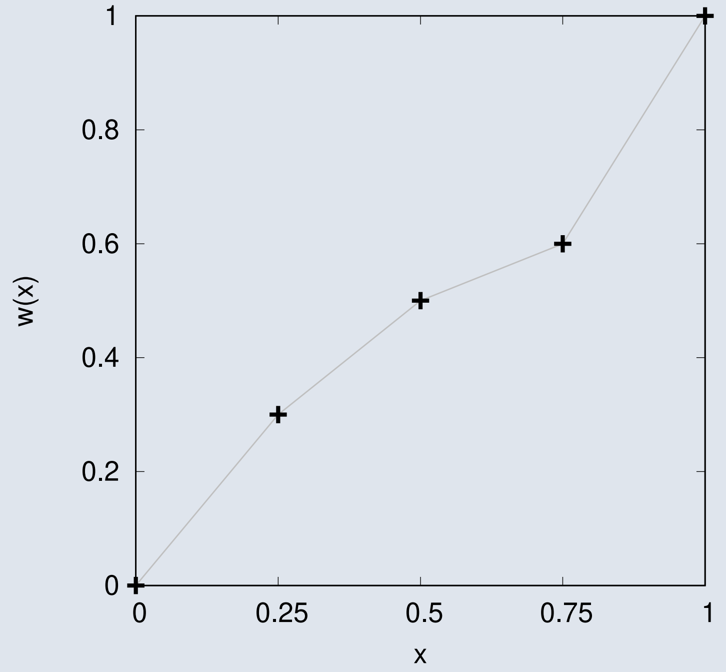
$$F(a_1, \ldots, a_n) = p_1 a_1 + p_2 a_2 + \ldots + p_n a_n$$

  is generally not possible
- Reminder: weights are bound to position and not to scenarios
- There is another generalization of OWA, called the Weighted Ordered Weighted Averaging (WOWA) Operator

Decision Theory

WOWA: Definition

- Weight vector $w = (w_1, w_2, \ldots, w_n)$
- $w(x)$ is an interpolating function of the points $(i/n, \sum_{j \leq i} w_j)$
- Example: $w = (0.3, 0.2, 0.1, 0.4)$



Decision Theory

WOWA: Definition

- Given:
  - $p_i \in [0, 1]$, $\sum_{i \in [n]} p_i = 1$
  - $w_i \in [0, 1]$, $\sum_{i \in [n]} w_i = 1$
- Define

$$WOWA(a_1, \ldots, a_n) = \sum_{i \in [n]} \omega_i a_{\pi(i)}$$

  with
  - π a permutation such that A is sorted in ascending order
  - $\omega_i = w(\sum_{j \leq i} p_{\pi(j)}) - w(\sum_{j < i} p_{\pi(j)})$
  - $w(x)$ is the interpolating function

Decision Theory

WOWA: Properties

- Let $p_i = 1/n$ for all $i \in [n]$
- Intuitively: all scenarios are equally important
- Then:

$$\begin{aligned}\omega_i &= w\left(\sum_{j\leq i} p_{\pi(j)}\right) - w\left(\sum_{j<i} p_{\pi(j)}\right)\\ &= w(i/n) - w((i-1)/n) = w_i\end{aligned}$$

- WOWA corresponds to the normal OWA

Decision Theory

WOWA: Properties

- Let $w_i = 1/n$ for all $i \in [n]$
- Intuitively: all positions are equally important
- Then:

$$\omega_i = w\left(\sum_{j\leq i} p_{\pi(j)}\right) - w\left(\sum_{j<i} p_{\pi(j)}\right)$$
$$= \sum_{j\leq i} p_{\pi(j)} - \sum_{j<i} p_{\pi(j)} = p_{\pi(i)}$$

- And WOWA corresponds to the weighting $\sum_{i\in[n]} p_i a_i$

Quiz

Question 1

True or false?

OWA is a generalization of Maximin, Regret, Hurwicz, and Average.

Question 2

For $A = (4, 8, 2, 6, 3, 0)$, what is OWA with $w = (0, 0, 1/2, 1/2, 0, 0)$?

Decision Theory

Quiz

Question 1

True or false?

OWA is a generalization of Maximin, Regret, Hurwicz, and Average.

Solution

False: it is not a generalization of Regret.

Question 2

For $A = (4, 8, 2, 6, 3, 0)$, what is OWA with $w = (0, 0, 1/2, 1/2, 0, 0)$?

Solution

It is the average of the two middle outcomes.

$$\begin{aligned} B &= (8, 6, 4, 3, 2, 0) \\ F(A) &= (1/2) \cdot 4 + (1/2) \cdot 3 \\ &= 3.5 \end{aligned}$$