Aalto University
School of Science

Project portfolio selection under incomplete information of future scenarios and utility

Johannes Mäkinen
Presentation *11*
22.10.2021

MS-E2191 Graduate Seminar on Operations Research
Fall 2021

Content

Introduction

- Multiple investment projects:
- Each project can have: multiple attributes, several resource constraints, project interdependencies,…
  - One-time investment can be large, we have to assess risk

- In addition, **exogenous uncertainties** (which are not influenced by the projects) e.g. macroeconomic developments
  - Rate of industry growth: ”[Finnish forest] … The growth target supports climate objectives and creates opportunities for the use of wood resources”. –Business Finland
  - Raw material shortages: ”And right now, there just aren't enough of [microchips] to meet industry demand. As a result, many popular products are in short supply”. - BBC
  - Pandemics…

Content

Basics

- Possible investment projects $X^0 = \{x^1, \dots x^m\}$
  - One-time investments that require only initial budgeting

- Each project has an outcome in *n* disjoint scenarios $\Omega = \{s_1, \dots, s_n\}$
  - Value of project x^j in scenario s_i is $x^j(s_i)$

- Portfolio value in scenario s_i
  - $X(s_i) = \sum_{x^j \in X} x^j(\mathrm{s_i})$, where $\mathrm{X} = \{X \mid X \subseteq X^0\}$

Basics

- The probability of scenario s_i is p_i
  - Scenario probabilities $p = (p_1, \dots, p_n)^T$ belong to the set
    $P^0 = \{p \in \mathbb{R}^n | p_i \geq 0, \sum_{i=1}^n p_i = 1\}$

- The expected value of portfolio X is $\mathbb{E}_p[X] = \sum_{i=1}^n p_i X(s_i)$

- The optimization problem is thus:

$$\max_{X \in X_F} \mathbb{E}_p[u(X)] = \max_{z \in \{0,1\}^\wedge m} \left\{ \sum_{i=1}^n p_i u \left(\sum_{j=1}^m z_j x^j(s_i) \right) | Az \leq B \right\}$$

$z(X) \in \{0,1\}^m$ *is such that* $z_j(X) = 1$ iff $x^j \in X$

The problem statement visualised

Expert estimates on actions' scenario values and relative costs.

j	Action title	Uncertain supply of plantation wood	Wood has no use in energy production	Uncertain/ volatile wood prices	Molecule level high-tech paper products	Nano-science has failed	All wood is certified	Wild wood markets dominate	Strong environmental regulation	Cost a_j
		$x^j(s_1)$	$x^j(s_2)$	$x^j(s_3)$	$x^j(s_4)$	$x^j(s_5)$	$x^j(s_6)$	$x^j(s_7)$	$x^j(s_8)$	
1	CO2 technology (T)	0	4	0	0	4	0	0	7	2.5
2	Flexible production (O)	6	2	5	6	0	1	1	6	5
3	Less price sensitive (O)	0	4	0	4	4	4	0	0	1
4	Focus on South America (O)	1	0	0	0	5	1	0	1	5
5	Vertical integration (O)	7	2	6	5	5	6	0	1	2.5
6	Deinked pulp (O)	6	4	2	0	0	7	6	2	5
7	Erasable paper (T)	3	0	1	6	0	0	0	0	10
8	Production in China (O)	0	5	0	4	6	0	0	1	5
9	Multi-fibre units (T)	7	3	5	6	5	5	6	1	5
10	Bio refineries (O)	4	0	4	4	4	2	0	1	5
11	Smart papers (T)	0	3	2	2	0	0	0	0	10
12	Hybrid media (T)	3	3	3	5	2	0	0	1	10
13	Broad portfolio (O)	0	3	4	6	4	1	4	1	2.5
14	Small mills (O)	0	0	1	2	0	2	0	5	5
15	Fusions (S)	0	3	0	0	5	0	4	1	2.5
16	Secure raw material (S)	6	3	7	3	3	7	7	1	5
17	Technology company (S)	3	4	0	7	0	0	0	0	10
18	Traditional technology (T)	0	0	3	0	6	0	0	6	2.5
19	Adaptable production (O)	5	0	4	0	0	1	5	1	5
20	Mass-customisation (O)	0	4	0	0	2	0	0	1	1
21	Mini-mills (O)	1	1	0	1	0	6	1	4	5
22	Brand paper (O)	0	2	0	0	0	0	0	0	1
23	Region portfolio (O)	0	0	0	3	6	1	6	3	2.5
24	Paper collection (S)	0	2	3	1	3	7	6	1	5

Source: Liesiö, J., & Salo, A. (2012). Scenario-based portfolio selection of investment projects with incomplete probability and utility information.

Content

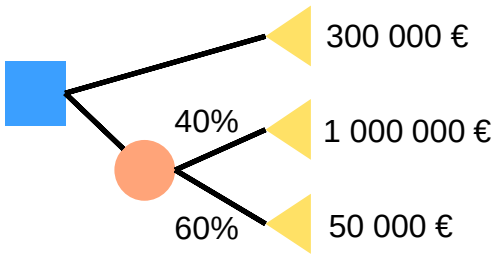
Modelling incomplete information

- Feasible probabilities $P := \{p \in P^0 | A_p p \leq B_p\}$
  - A_p and B_p are derived from statements about the scenario probabilities
    - *s_1 more probable than s_2* $\rightarrow p_1 \geq p_2$
    - *$\{s_1, s_2\}$ less probable than $\{s_3, s_4, s_5\}$* $\rightarrow p_3 + p_4 + p_5 \geq p_1 + p_2$
    - $\underline{p_i} \leq p_i \leq \overline{p_i}$

- If multiple experts give different probability estimates, the set P can be defined as the convex hull of the estimates

Modelling incomplete information

- Feasible utilities $U \subseteq U^0$, $U^0 = \{u: \mathbb{R} \rightarrow [0,1] | u(t) \geq u(t') \; \forall t \geq t'\}$
  - Elicitation of incomplete information is easier than standard elicitation:



If the uncertain option is preferred
→ u(300) ≤ 0.4u(1000)+0.6u(50)

<u>This is enough</u> as any preference between alternatives already implies constraints on the model.

Thus, alternatives need not to be adjusted until the DM is indifferent between the two

- Though, again the experts can give differing preferences
- The information set is **S**= $\boldsymbol{P} \times \boldsymbol{U}$

Content

Dominance structures

Definition. Portfolio X dominates $X'$with regard to information set $S = P \times U$,denoted $X \succ_S X'$ if

$\mathbb{E}_p[u(X)] \geq \mathbb{E}_p[u(X')]\ \forall\ (p, u) \in S$ and
$\mathbb{E}_p[u(X)] > \mathbb{E}_p[u(X')]$ for some $(p, u) \in S$

- Portfolio dominates another if
  1. its expected utility is at least as high for all feasible scenario probabilities and utility functions
  2. there exist some scenario probabilities and a utility function for which it has a strictly higher expected utility

Dominance structures

Theorem 1: Let $P \subseteq P^0, U \subseteq U^0$ and choose portfolios $X, X' \in \chi$. Then

i. $X \succ_{P \times U} X' \leftrightarrow X \succ_{ext(P) \times U} X'$,

ii. $X \succ_{P^0 \times U} X' \leftrightarrow X(s_i) \geq X'(s_i)\ \forall i \in \{1, \dots, n\}$,

Where at least one of the inequalities in (*ii*) is strict for some $i \in \{1, \dots, n\}$

- Portfolio dominance can be checked by comparing the expected utilities at the extreme points of the set of feasible scenario probabilities
- If no constraints on probabilities ($P = P^0$), every extreme point occurs with a probability of 1

Content

Computation of non-dominated portfolios

- First, we can discard all dominated portfolios and determine the set $X_N(P^0 \times U^0)$, because $X_N(P \times U) \subseteq X_N(P^0 \times U^0)$

- Pareto-optimal solutions to the MOZOLP problem

$$\max_z \{Cz | Az \leq B, z \in \{0,1\}^m\}$$

Where $C \in \mathbb{R}^{n \times m}$ with $[C]_{ij} = x^j(s_i)$ contains the projects' scenario-specific values.

Computation of non-dominated portfolios

- Second, with our information set $S = P \times U \subseteq P^0 \times U^0$, we can determine the min. and max. expected utility differences for each pair of portfolios X and X' at each <u>extreme point</u> $p \in ext(P)$

$$\mathbb{E}_p[u(X)] - \mathbb{E}_p[u(X')] = \sum_{i=1}^{n} p_i u\big(X(s_i)\big) - \sum_{i=1}^{n} p_i u\big(X'(s_i)\big)$$

- Now, $X \succ_S X'$ iff:
  1. $min.\{\mathbb{E}_p[u(X)] - \mathbb{E}_p[u(X')]\} \geq 0 \; \forall \; p \in ext(P)$
  2. $max.\{\mathbb{E}_p[u(X)] - \mathbb{E}_p[u(X')]\} > 0$ for some p $\in ext(P)$

Computation of non-dominated portfolios

	s_1	s_2
$X(s_i)$	5	3
$X'(s_i)$	2	5

- Consider a two-scenario problem:
  - Constrain $P = \{p = (p_1, p_2) \in P^0 \mid p_1 \in [0.4,0.5]\}$
    → $ext(P) = \{(0.4,0.6), (0.5,0.5)\}$
  - Set sorted values to $\hat{v} = (2,3,5)$ & $\hat{u} = u(\hat{v}), \hat{u} \in [0,1]^3, \hat{u}_j \leq \hat{u}_{j+1}$

Now, at extreme point (0.4,0.6), the expected utility difference is:

$$\mathbb{E}[u(X)] - \mathbb{E}[u(X')] = 0.4u(5) + 0.6u(3) - 0.4u(2) - 0.6u(5)$$
$$= -0.4u(2) + 0.6u(3) - 0.2u(5)$$

For increasing utility functions, the above attains its minimum -0.2 at $\hat{u} = (0,0,1)^T$ → X does not dominate X'

Computation of non-dominated portfolios

- We can introduce additional constraints to restrict utility functions to be concave:

$$\frac{\hat{u}_j - \hat{u}_{j-1}}{\hat{v}_j - \hat{v}_{j-1}} \geq \frac{\hat{u}_{j+1} - \hat{u}_j}{\hat{v}_{j+1} - \hat{v}_j} \leftrightarrow \frac{u(3) - u(2)}{3 - 2} \geq \frac{u(5) - u(3)}{5 - 3} \leftrightarrow -2u(2) + 3u(3) - u(5) \geq 0$$

$$u(2) = 0, u(5) = 1 \rightarrow u(3) \geq \frac{1}{3}$$

The expected utility difference is minimised when $u(3) = \frac{1}{3}$,(and maximised when $u(3) = 1$)

$$p = (0.4,0.6): \mathbb{E}[u(X)] - \mathbb{E}[u(X')] = -0.4u(2) + 0.6u(3) - 0.2u(5) = 0$$

$$p = (0.5,0.5): \mathbb{E}[u(X)] - \mathbb{E}[u(X')] = -0.5u(2) + 0.5u(3) = \frac{1}{6}$$

→ X dominates X'

Computation of non-dominated portfolios

- Quick recap on what we just did:
  1. Find non-dominated portfolios without any information on probabilities or utilities
  2. Use our information set to narrow down the set of non-dominated portfolios
     - *We only need to look at the extreme points of the probability space!*
  3. Add constraints to our utility functions to express the risk preferences

Content

Application

- Explore the possible future operation environments of the global forest industry

- Experts estimated:
  - 24 actions (possible investment projects)
  - 8 different scenarios
  - Cost of each action
  - Value of each completed action in each scenario

- Portfolio cost is constrained $\leq \frac{1}{3} \cdot$ (sum of all actions)

Application

Expert estimates on actions' scenario values and relative costs.

j	Action title	Uncertain supply of plantation wood	Wood has no use in energy production	Uncertain/ volatile wood prices	Molecule level high-tech paper products	Nano-science has failed	All wood is certified	Wild wood markets dominate	Strong environmental regulation	Cost a_j
		$x^j(s_1)$	$x^j(s_2)$	$x^j(s_3)$	$x^j(s_4)$	$x^j(s_5)$	$x^j(s_6)$	$x^j(s_7)$	$x^j(s_8)$	
1	CO2 technology (T)	0	4	0	0	4	0	0	7	2.5
2	Flexible production (O)	6	2	5	6	0	1	1	6	5
3	Less price sensitive (O)	0	4	0	4	4	4	0	0	1
4	Focus on South America (O)	1	0	0	0	5	1	0	1	5
5	Vertical integration (O)	7	2	6	5	5	6	0	1	2.5
6	Deinked pulp (O)	6	4	2	0	0	7	6	2	5
7	Erasable paper (T)	3	0	1	6	0	0	0	0	10
8	Production in China (O)	0	5	0	4	6	0	0	1	5
9	Multi-fibre units (T)	7	3	5	6	5	5	6	1	5
10	Bio refineries (O)	4	0	4	4	4	2	0	1	5
11	Smart papers (T)	0	3	2	2	0	0	0	0	10
12	Hybrid media (T)	3	3	3	5	2	0	0	1	10
13	Broad portfolio (O)	0	3	4	6	4	1	4	1	2.5
14	Small mills (O)	0	0	1	2	0	2	0	5	5
15	Fusions (S)	0	3	0	0	5	0	4	1	2.5
16	Secure raw material (S)	6	3	7	3	3	7	7	1	5
17	Technology company (S)	3	4	0	7	0	0	0	0	10
18	Traditional technology (T)	0	0	3	0	6	0	0	6	2.5
19	Adaptable production (O)	5	0	4	0	0	1	5	1	5
20	Mass-customisation (O)	0	4	0	0	2	0	0	1	1
21	Mini-mills (O)	1	1	0	1	0	6	1	4	5
22	Brand paper (O)	0	2	0	0	0	0	0	0	1
23	Region portfolio (O)	0	0	0	3	6	1	6	3	2.5
24	Paper collection (S)	0	2	3	1	3	7	6	1	5

Source: Liesiö, J., & Salo, A. (2012). Scenario-based portfolio selection of investment projects with incomplete probability and utility information.

Application

- Five different information sets were constructed using
  - Rank ordering:

    $$P^R = \{p \in P^0 | p_3 \geq p_2 \geq p_4 \geq p_8 \geq p_5 \geq p_7 \geq p_6 \geq p_1\}$$
  - Incomplete rank ordering:

    $$P^{IR} = \{p \in P^0 | p_{i_a} \geq p_{i_b} \geq p_{i_c} \geq p_{i_d} \ \forall\ i_a \in \{2,3\}, i_b \in \{4,8\}, i_c \in \{5,7\}, i_d \in \{6,1\}\}$$
  - Centroid set:

    $$\hat{p} = \sum_{p \in ext(P^R)} p/n$$
  - Linear utility $\{u_L\}$
  - Strictly increasing concave utility U^A

Application

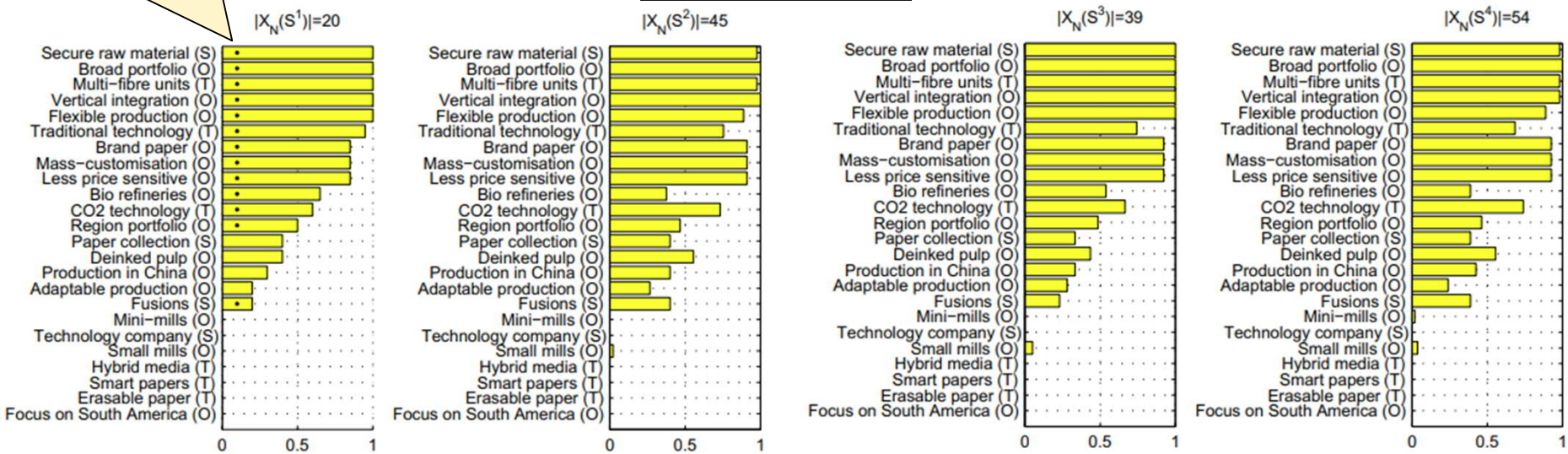
$$\{\hat{p}\} \times \{u^L\} \subset S^1 = P^R \times \{u^L\} \subset S^2 = P^{IR} \times \{u^L\}$$

$$\cap \qquad\qquad\qquad \cap$$

$$S^3 = P^R \times U^A \subset S^4 = P^{IR} \times U^A$$

Point estimate portfolio (dots) was *very* sensitive to probability estimates

Core indices of actions



Source: Liesiö, J., & Salo, A. (2012). Scenario-based portfolio selection of investment projects with incomplete probability and utility information.

Summary

- The framework produces non-dominated portfolios with incomplete information on scenario probabilities and/or risk preferences
- Interactive exploration of non-dominated portfolios by adding more elicited information
- Sensitive to point estimated probabilities

References

Liesiö, J., & Salo, A. (2012). Scenario-based portfolio selection of investment projects with incomplete probability and utility information. *European Journal of Operational Research*, *217*(1), 162-172.

Finland updates national forest strategy, (*2019,* February 18), *Business Finland,*
https://www.businessfinland.fi/en/whats-new/news/2019/finland-updates-national-forest-strategy

Chris Baraniuk (2021, August 27), '*Why is there a chip shortage?', BBC,*
https://www.bbc.com/news/business-58230388

COVID-19 frontpage, World Health Organisation,
https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Homework

- Task 1
  - Describe briefly a decision making problem where you are faced with exogenous uncertainties.
    - Slides 21-22 but way simpler, 3-5 investment projects (actions) and 3-5 scenarios
    - Remember to explain why the scenarios could affect the actions

- Task 2
  - List at least three (3) reasons why the framework should be used in project portfolio selection
    - There are multiple good reasons listed in the paper

Send the answers to *johannes.makinen@aalto.fi*
with title "MS-E2191_HW_11" by 29.10.2021 9.00

Why do we need to check only the extreme points?

Theorem 1. *Let $P \subseteq P^0$, $U \subseteq U^0$ and choose portfolios $X, X' \in \mathcal{X}$. Then*

$$(i) X \succ_{P \times U} X' \iff X \succ_{ext(P) \times U} X',$$
$$(ii) X \succ_{P^0 \times U} X' \iff X(s_i) \geqslant X'(s_i) \ \forall i \in \{1, \ldots, n\},$$

where at least one of the inequalities in (ii) is strict for some $i \in \{1, \ldots, n\}$.

Liesiö, J., & Salo, A. (2012). Scenario-based portfolio selection of investment projects with incomplete probability and utility information.

Proof of Theorem 1. Part (i) '$\Leftarrow$': Assume $X \succ_{ext(P) \times U} X'$ which implies

$$\mathbb{E}_p[u(X)] \geqslant \mathbb{E}_p[u(X')] \forall p \in \{p^1, \ldots, p^t\}, \quad u \in U,$$

where $\{p^1, \ldots, p^t\} = ext(P)$. Any $p \in P$ is a linear combination of these extreme points, i.e., $p = \sum_{k=1}^{t} \alpha_k p^k$, where $\alpha_k \geqslant 0$. For any $(p, u) \in P \times U$:

$$\begin{aligned}
\mathbb{E}_p[u(X)] - \mathbb{E}_p[u(X')] &= \sum_{i=1}^{n} p_i [u(X(s_i)) - u(X'(s_i))] \\
&= \sum_{i=1}^{n} \sum_{k=1}^{t} \alpha_k p_i^k [u(X(s_i)) - u(X'(s_i))] \\
&= \sum_{k=1}^{t} \alpha_k \sum_{i=1}^{n} p_i^k [u(X(s_i)) - u(X'(s_i))] \\
&= \sum_{k=1}^{t} \alpha_k (\mathbb{E}_{p^k}[u(X)] - \mathbb{E}_{p^k}[u(X')]) \geqslant 0,
\end{aligned}$$

since all terms of the sum are non-negative. Thus $\mathbb{E}_p[u(X)] \geqslant \mathbb{E}_p[u(X')]$ for all $(p, u) \in (P \times U)$ and the inequality is strict for some $p \in ext(P) \subset P$ and $u \in U$, which implies $X \succ_{P \times U} X'$.'$\Rightarrow$': Assume $X \succ_{P \times U} X'$, which implies $\mathbb{E}_p[u(X)] \geqslant \mathbb{E}_p[u(X')]$ for all $(p, u) \in ext(P) \times U$, since $ext(P) \subset P$. Furthermore, exists $p \in P$, $p = \sum_{k=1}^{t} \alpha_k p^k$, such that $0 < \mathbb{E}_p[u(X)] - \mathbb{E}_p[u(X')] = \sum_{k=1}^{t} \alpha_k (\mathbb{E}_{p^k}[u(X)] - \mathbb{E}_{p^k}[u(X')])$. Thus $\mathbb{E}_{p^k}[u(X)] > \mathbb{E}_{p^k}[u(X')]$ for some $p^k \in ext(P)$ which implies $X \succ_{ext(P) \times U} X'$.

Part (ii) Since the extreme points $\{p^1, \ldots, p^n\}$ of P^0 are of the form $p_i^i = 1$, $p_j^i = 0 \ \forall j \neq i$, i) implies that dominance $X \succ_{(P^0 \times U)} X'$ holds if and only if $u(X(s_i)) \geqslant u(X'(s_i)) \forall i \in \{1, \ldots, n\}$, $u \in U$ (with the inequality strict for some i and u). Since $U \subseteq U^0$ contains only increasing utility functions (and at least one strictly increasing) the condition is equal to $X(s_i) \geqslant X'(s_i) \forall i \in \{1, \ldots, n\}$ with a strict inequality for at least one i. $\square$