Fondements Mathématiques pour l'Aide à la Décision

Rapport du Projet

Eco - score

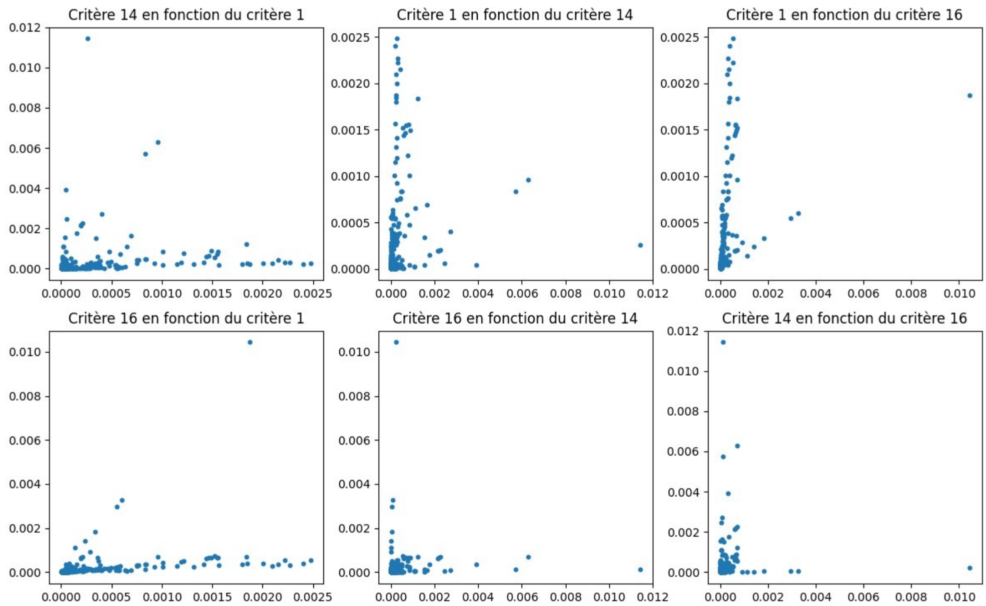
Emi - IM2D

El Yazid Kadiri - MIAGE

Carla Nachman - IM2D

Question 3

En utilisant les trois critères : impact sur le dérèglement climatique, impact sur l'épuisement des ressources en eau, et impact sur l'épuisement des ressources matières (critères 1, 14, et 16). Réaliser des graphiques pour illustrer les liens entre ces différents critères (e.g., avec le critère en abscisse et le critère en ordonnée).



Question 4

Quelles sont les propriétés vérifiées par la Pareto Dominance ? À quelle structure particulière, vue en cours, correspond-elle ? Vous devez apporter des justifications pour les propriétés vérifiées et celles non vérifiées.
La Pareto Dominance nous dit " X doit être au moins aussi bien que Y sur tous les critères, mais doit également être strictement meilleur que Y sur au moins un critère. Nous avons 16 critères.

Non-Réflexive :
X ne peut être strictement meilleure qu'elle-même. En effet, par l'absurde, si la Pareto Dominance était réflexive, on aurait pour tout i entre 1 et 16, xi>=xi ce qui est correcte, mais également, un io, tel que xio>xio ce qui est impossible.

a) Antisymétrie :
Montrons que la Pareto dominance n'est pas symétrique. En effet, si X domine Y, cela n'est pas équivalent au fait que Y domine X. Par l'absurde, X Pareto domine Y veut dire que pour tout i entre 1 et 16 : xi>=yi, et il existe un io tel que xio>yio. De la même manière, Y Pareto domine X : pour tout j entre 1 et 16 : yj>xj et il existe jo tel que yjo>xjo. Or pour tous les i entre 1 et 16, xi>=yi, y compris xjo>=yjo (puisque jo appartient aux indices entre 1 et 16). Avoir yjo>xjo est donc absurde ! La Pareto dominance n'est pas symétrique.

b) Asymétrie :
Partons de la définition : X Pareto domine Y implique non (Y Pareto domine X). X Pareto domine Y veut dire que pour tout i entre 1 et 16 : xi>=yi, et il existe un io tel que xio>yio. Ceci est équivalent à non (il existe un i entre 1 et 16 tel que xi<yi, et pour tout io appartenant à {1, ..., 16}, xio<=yio). Soit exactement la définition de non (Y Pareto domine X). Donc la relation est bien asymétrique.

a) Transitivité :
Si X domine Y et Y domine Z, alors X domine Z.
Montrons-le : X Pareto domine Y : pour tout i entre 1 et 16 : xi>=yi, et il existe un i1 parmi les i, tel que xi1>yi1. Y Pareto domine Z : pour tout i entre 1 et 16 : yi>=zi, et il existe un i2 parmi les i, tel que yi2>zi2.
Pour tous nos indices i, on a évidemment : xi>=yi>=zi, donc xi>=zi sur tous les critères.
Si i1=i2, on a directement : xi1>yi1>zi1
Sinon,
- Sur i1 on a xi1>yi1>=zi1, donc xi1>zi1
- Sur i2, on a xi2>=yi2>zi2, donc xi2>zi2

Ainsi, en prenant io=i1 ou io=i2, on a bien pour tous i entre 1 et 16 : xi>=zi et un indice io tel que xio>zio Ainsi X Pareto domine Z, et la transitivité est vérifiée.

b) Négativement-Transitive :
Elle ne peut être négativement transitive, car elle ne compare pas les éléments sur un critère ordinal.

Non complète : Il est possible que deux solutions X et Y ne soient pas comparables, c'est-à-dire qu'il n'y ait pas de relation de domination entre elles. En effet, si X et Y sont différents sur certains critères, mais qu'aucun des deux ne sont strictement meilleurs que l'autre sur tous les critères, alors il n'y a pas de relation de Pareto domination entre X et Y.

Ainsi, la relation de Pareto Dominance est non réflexive, asymétrique et transitive. C'est donc une relation d'ordre strict.

Question 7

Calculer le pourcentage d’ensembles {X, Y} de deux alternatives avec $X \in A /= Y \in A$ tels que X Pareto-domine Y. Comment ce pourcentage évolue-t-il si les alternatives n’étaient évaluées que sur les critères 1, 14, et 16 ?

Nous calculons le pourcentage d’ensemble {X, Y} appartenant à A^2, tels que X Pareto-domine Y, à l’aide d’une fonction “pourcentage” qui appelle les fonctions pareto_domine (de la Q°5) et pareto2 qui détermine si une alternative pareto-domine une autre sur les critères 1, 14 et 16 uniquement .
En exécutant “pourcentage(donnees)” nous obtenons :
Pourcentage pour tous les critères : 14.59 %
Pourcentage pour les critères 1, 14 et 16: 27.15 %
Ainsi, nous observons que le pourcentage est plus grand lorsque nous nous concentrons sur les critères 1, 14 et 16. Cela n’est pas forcément étonnant puisqu’on se concentre sur moins de critères et donc il y a “plus de chances” de respecter la pareto-dominance

Pour une lecture plus lisible, nous renommons w^ par wo.
wo = [0.2106, 0.0631, 0.0501, 0.0478, 0.0896, 0.0184, 0.0213, 0.062, 0.028, 0.0296, 0.0371, 0.0192, 0.0794, 0.0851, 0.0832, 0.0755]

Question 8

Calculer les taux de substitution entre les critères 1, 14, et 16 et donner les interprétations de ces taux de substitution.

Poids des critères (w) :

- Poids critère n°1 = 21.06 × 10^−2
- Poids critère n°14 = 8.51 × 10^−2
- Poids critère n°16 = 7.55 × 10^3

La formule pour calculer le taux de substitution entre deux critères i et j :
TS(x,y) = (Poids du critère x / Poids du critère y)
Donc on trouve les taux de substitutions suivants :

- Taux de substitution entre les critères 1 et 14 = 2.475
- Taux de substitution entre les critères 1 et 16 = 2.789 × 10^-5
- Taux de substitution entre les critères 14 et 16 = 1.127 × 10^-5

Interprétations :

· Le taux de substitution entre les critères 1 et 14 est de 2.475, ce qui indique que pour chaque unité de poids attribuée au critère 1 (impact sur le réchauffement climatique), on peut substituer 2.475 unités de poids du critère 14 (impact sur l'épuisement des ressources en eau) tout en maintenant le même niveau de préférence ou d'évaluation globale dans l'évaluation de l'impact environnemental des produits alimentaires. Cela signifie que l'impact sur l'épuisement des ressources en eau est considéré comme étant environ 2.475 fois moins important que l'impact sur le réchauffement climatique dans cette évaluation.

· Le taux de substitution entre les critères 1 et 16 est très faible, il est de 2.789 × 10^-5. Ce qui indique que les poids attribués aux critères 1 (impact sur le réchauffement climatique) et 16 (impact de l'empreinte matière) sont très différents et que les deux critères ne sont pas substituables facilement dans cette évaluation. Cela signifie que l'impact de l'empreinte matière est considéré comme étant très insignifiant par rapport à l'impact sur le réchauffement climatique dans cette évaluation.

· Ce taux de substitution est également très faible (1.127 × 10^-5), ce qui indique que les poids attribués aux critères 14 (impact sur l'épuisement des ressources en eau) et 16 (impact de l'empreinte matière) sont aussi très différents et que les deux critères ne sont pas substituables facilement dans cette évaluation. Cela signifie que l'impact de l'empreinte matière est considéré comme étant très insignifiant par rapport à l'impact sur l'épuisement des ressources en eau dans cette évaluation.

Question 9

Proposer un programme linéaire permettant de trouver le jeu de poids w (poids positifs, sommant à 1), tel que SP wo(X) ≥ SP wo(Y), et que // w − wo // 1 soit minimale.

Min ||w - wo||
Sc:
Σwi = 1
Σi=1a16 (wixi) > Σi=1a16 (woyi)
w ≥ 0

Avec wo = (0,2106 ; 0,0631 ; 0,0501; 0,0478; 0,0896; 0,0184; 0,0213, 0,062; 0,028; 0,0296; 0,0371; 0,0192; 0,0794; 0,0851; 0,0832; 0,0755)

Question 11

Réaliser un graphique associant à chaque valeur $x \in [0, 2]$, le nombre de paires (X, Y) telles que $L1\_inv(X, Y) \leq x$. Quelle critique peut-on faire sur l'interprétation de ce graphique ?

Pour cette question, nous avons eu du mal à exécuter notre programme. En effet, celui-ci tourne pendant plus de 5 minutes et nous n'avons donc pas de résultat pour le graphique. Les deux parties de code (L1_inv et graphique_L1) marchent séparément mais pas ensemble.

En revanche, nous pouvons tout de même faire une interprétation de cette méthode.
Plus le nombre de paires satisfaisant la condition est élevé pour une valeur de x donnée, plus la comparaison entre X et Y est considérée comme étant "robuste vis-à-vis des variations des poids".
Cependant, une critique importante de ce graphique est qu'il ne tient pas compte de la signification des poids eux-mêmes. Les poids sont souvent déterminés par des considérations subjectives ou des préférences personnelles, et le choix d'un jeu de poids particulier peut avoir un impact significatif sur les résultats de l'analyse multicritère.

Ce graphe nous permet d'observer comment la distance L1_inv se répartit entre les différentes paires de données. Mais il ne nous donne pas d'informations sur la répartition spatiale des (X, Y) dans l'espace des données, il peut donc être difficile de tirer des conclusions générales sur la qualité de l'estimation de L1_inv() à partir de ce graphique.

Question 12

Implémenter une méthode permettant d’évaluer les différentes alternatives par une moyenne pondérée ordonnée. Implémenter une méthode permettant d’évaluer les différentes alternatives par une moyenne géométrique pondérée. Proposer quelques jeux de poids qui vous semblent appropriés.
cf code,
Nous avons implémenté les deux méthodes et proposé les jeux de poids suivants:

- w1 = [1/16]*16 , un jeu de poid équitable
- w2 et w3, deux jeux de poids aléatoires

Ces 3 jeux somment tous a 1.

Question 13

Quelle propriété peut être satisfaite par ces deux agrégateurs qui n’est pas satisfaite par la somme pondérée ?

Les deux agrégateurs, la moyenne pondérée ordonnée et la moyenne géométrique pondérée, sont connus pour produire des solutions ou alternatives équilibrées lorsqu'ils sont utilisés avec des poids triés en ordre décroissant. En effet, ces deux agrégateurs donnent plus de poids aux éléments les moins élevés lorsque les poids sont triés dans cet ordre. Cette propriété est utile pour les situations où on souhaite une solution équilibrée.

Cependant, la somme pondérée ne satisfait pas cette propriété, car elle n'est pas influencée par l'ordre des poids. Ainsi, la somme pondérée peut donner plus de poids aux éléments les plus élevés si les poids sont triés du plus grand au plus petit, ce qui ne favorise pas nécessairement des solutions équilibrées.

De plus, il y a une seconde propriété qui peut être satisfaite par ces deux agrégateurs mais qui n’est pas satisfaite par la somme pondérée. En effet, la moyenne pondérée ordonnée et la moyenne géométrique pondérée garantissent la négative transitivité contrairement à la somme pondérée.

En effet, sachant que la propriété de négative transitivité est définie pour une relation binaire entre des éléments. C'est-à-dire que si deux éléments A et B sont en relation négative, et que B et C sont également en relation négative, alors A et C doivent également être en relation négative.

La moyenne pondérée ordonnée est donc négativement transitive car les éléments sont ordonnés selon leur importance et les coefficients de pondération sont choisis pour refléter cet ordre. Cette relation d'ordre implique donc une relation de positivité entre les coefficients de pondération et les éléments correspondants.

De plus, la moyenne géométrique pondérée est aussi négativement transitive car les éléments sont multipliés en utilisant des coefficients de pondération et comme tous les éléments sont positifs, la multiplication des éléments est également positive, ce qui implique une relation de positivité entre les éléments et les coefficients de pondération correspondants.

Cependant, la somme pondérée n'est pas forcément négativement transitive car l'addition de valeurs positives peut conduire à une valeur qui est positive, nulle ou alors négative, en fonction des coefficients de pondération choisis pour chaque élément. Ainsi, il n'y a pas de relation claire entre la relation de positivité ou de négativité entre les éléments et la somme pondérée correspondante.

Question 14

Montrer que dKT définit bien une distance sur les ordres forts définis sur A.
Pour montrer que dKT définit bien une distance sur les ordres forts définis sur A, il faut prouver que :
$\forall R,R'$
1) $dKT(R,R') \geqslant 0$
2) $dKT(R,R') = 0 \Leftrightarrow R = R'$
3) $dKT(R, R') = dKT(R', R)$

1)
On sait que toute relation d'ordre entre deux éléments est positive. On en déduit donc que $\forall R,R'$ $dKT(R,R') \geqslant 0$

2)
<=
$dKT(R,R) = \sum (i,j)\in A2\ 0.5(iRj \wedge jRi) + 0.5((iRj \wedge \neg(iRj)) \vee (\neg(jRi) \wedge jRi)) = 0$
Car $\forall(i,j)\in A2\ (iRj \wedge jRi) = 0 \wedge (iRj \wedge \neg(iRj)) = 0 \wedge (\neg(jRi) \wedge jRi) = 0$

=>
Soient R et R' deux ordres forts,
$dKT(R,R') = 0 \Leftrightarrow \sum (i,j)\in A2\ 0.5(iRj \wedge jR'i) + 0.5((iRj \wedge \neg(iR'j)) \vee (\neg(jRi) \wedge jR'i)) = 0$
$dKT(R,R') = 0 \Leftrightarrow \forall(i,j)\in A2\ iRj \wedge jR'i = 0 \wedge ((iRj \wedge \neg(iR'j)) \vee (\neg(jRi) \wedge jR'i)) = 0$
alors $\forall(i,j)\in A2$, $(iRj \wedge jR'i) = 0$ et $(iRj \wedge \neg(iR'j)) = 0$ et $(\neg(jRi) \wedge jR'i)=0$

donc $\forall(i,j)\in A2$,
$iRj=0$ ou $jR'i = 0$
et $iRj =0$ ou $\neg(iR'j) = 0$
et $\neg(jRi)=0$ ou $jR'i=0$
Ce qui implique alors que R et R' sont les mêmes relations $\forall(i,j)\in A2$

3)
on prend :$dKT(R,R') = \sum (i,j)\in A2\ 0.5(iRj \wedge jR'i) + 0.5((iRj \wedge \neg(iR'j)) \vee (\neg(jRi) \wedge jR'i))$
et on prend dKT(R',R) en inversant dans la somme les i et les j puisqu'ils sont quelconques :
$dKT(R',R) = \sum (i,j)\in A2\ 0.5(jR'i \wedge iRj) + 0.5((jR'i \wedge \neg(jRi)) \vee (\neg(iR'j) \wedge iRj))$
$\Leftrightarrow \sum (i,j)\in A2\ 0.5(iRj \wedge jR'i) + 0.5((\neg(jRi) \wedge jR'i) \vee (iRj\wedge \neg(iR'j)))$ car $a\wedge b = b\wedge a$
$\Leftrightarrow \sum (i,j)\in A2\ 0.5(iRj \wedge jR'i) + 0.5((iRj \wedge \neg(iR'j)) \vee (\neg(jRi) \wedge jR'i))$ car $a\vee b = b\vee a$
$\Leftrightarrow dKT(R,R')$

Question 15

Nous avons utilisé deux ordres forts pour faire nos calculs de distance de Kendall-Tau. La première est R1 qui vérifie si tous les critères d'une alternative sont strictement supérieurs aux critères d'une autre alternative, et R2 qui vérifie si tous les critères d'une alternative sont inférieurs ou égaux aux critères d'une autre alternative. Voici les résultats que nous obtenons en fonction des poids différents :

dKT(R1,R2) = 132.5 avec P1 = P2 = 0.5
dKT(R1,R2) = 79.5 avec P1 = 0.7 et P2 = 0.3
dKT(R1,R2) = 53 avec P1 = 0.8 et P2 = 0.2
dKT(R1,R2) = 212 avec P1 = 0.2 et P2 = 0.8

Question 16

Que pensez-vous de la formule logarithmique utilisée ? Quels avantages et désavantages y voyez-vous ? Essayez de proposer une alternative (sans la tester).

Un des avantages notable de la normalisation logarithmique de l'Éco-score est la facilité de lecture de celui- ci: On compresse l'échelle de mesure tout en préservant les écarts entre les valeurs.
L'échelle logarithmique est particulièrement adaptée pour rendre compte des ordres de grandeur dans les applications, elle montre sur un petit espace une large gamme de valeurs.

Cependant, il peut être difficile à interpréter intuitivement et peut conduire à des scores élevés pour des valeurs modestes de l'éco-score.
De plus, à partir de 3.11, l'éco score devient une valeur négative donc plus comprise entre 0 et 100. Il faut donc l'utiliser dans un cas précis de valeur pour ne pas fausser les résultats

De plus, la formule suit une courbe logarithmique, non linéaire.
Une alternative possible pourrait être d'utiliser une formule linéaire pour la transformation de la somme pondérée en score, telle que :

```
def score(x):
s = 100 * (x - min_SP) / (max_score - min_SP)
return(s)
```

où min_score et max_score correspondent aux valeurs minimales et maximales possibles de la somme pondérée.

Cette alternative a l'avantage d'être plus simple à interpréter pour les utilisateurs et de ne pas être sensible aux variations des valeurs des critères. Cependant, elle ne prend pas en compte de manière non linéaire les écarts de performance entre les alternatives.

17. Réaliser des tests similaires à ceux réalisés aux questions 9-11, pour voir combien d'alternatives peuvent changer de catégorie si le jeu de poids wo est modifié.

a) Proposer un programme linéaire permettant de trouver le jeu de poids w (poids positifs, sommant à 1), tel que scorew(X) $\geq$ scorew(Y), et que $/\!/ w - \hat{w} /\!/ 1$ soit minimale.

b) Implémenter ce programme linéaire dans une méthode, nommée L1_inv_eco, qui étant donné X,Y une paire d'alternatives telles que scorewo(X) < scorewo(Y), retourne la valeur optimale du programme linéaire précédant. Vous utiliserez le package Python mip.

c)Réaliser un graphique associant à chaque valeur $x \in [0, 2]$, le nombre de paires (X, Y) telles que L1_inv_eco(X, Y) $\leq$ x.
Quelle critique peut-on faire sur l'interprétation de ce graphique?

Comme pour la question 11, nous avons eu du mal à exécuter notre programme. En effet, celui-ci tourne pendant plus de 5 minutes et nous n'avons donc pas de résultat pour le graphique. Les deux parties de code (L1_inv_eco et graphique_L1_eco) marchent séparément mais pas ensemble.

Nous avons donc cherché un autre moyen (assez simple) pour voir combien d'alternatives peuvent changer de catégorie si le jeu de poids wo est modifié, en utilisant la formule logarithmique de l'éco score.