

Projet MADMC 2023–2024

Élicitation incrémentale et recherche locale pour le problème du sac à dos multi-objectifs

1. Introduction

On s'intéresse dans ce projet à la version suivante du problème du sac à dos multi-objectifs (objectifs à **maximiser**) :

Données :

- n objets,
- une valeur W (capacité du sac à dos),
- chaque objet i a un poids w_i
- chaque objet i est valué par un vecteur v^i à p dimensions (v_1^i, \dots, v_p^i)

Solutions réalisables : tout sous-ensemble d'objets, caractérisé par un vecteur $x = (x_1, \dots, x_n)$ ($x_i = 1$ si l'objet i est sélectionné, 0 sinon), tel que $\sum_{i=1}^n w_i x_i \leq W$.

But : déterminer une solution réalisable x maximisant une fonction d'agrégation paramétrée représentant les préférences du décideur sur les solutions. La fonction d'agrégation est de la forme $\Psi_\lambda(y(x)) = \Psi_\lambda(y_1(x), \dots, y_p(x))$, avec :

- λ : jeu de paramètres initialement inconnu parmi l'ensemble Λ des jeux de paramètres initialement admissibles ($\lambda \in \Lambda$)
- $y(x)$: évaluation d'une solution x donnée par $(\sum_{i=1}^n v_1^i x_i, \dots, \sum_{i=1}^n v_p^i x_i)$.

La fonction d'agrégation sera une fonction linéaire en ses paramètres, par exemple la somme pondérée, OWA ou l'intégrale de Choquet.

Le projet consiste à développer deux procédures de résolution du problème, fondées sur l'élicitation incrémentale, et à comparer leur efficacité.

2. Une première procédure de résolution

Dans cette partie, vous allez mettre en œuvre une procédure en deux phases pour déterminer la solution préférée du décideur. La procédure fonctionne de la manière suivante : (1) détermination d'une approximation des points non-dominés (au sens de Pareto), puis (2) génération de la solution préférée du décideur parmi les points non-dominés à l'aide d'une procédure d'élicitation incrémentale.

Lors de la phase (1), pour la détermination de l'approximation des points non-dominés, vous utiliserez la recherche locale de Pareto (PLS) vue en cours. Vous pouvez donc réutiliser les résultats

obtenus lors du TME sur les métaheuristiques du 10 octobre. Attention : on demande ici de traiter des instances pouvant comprendre plus de 2 objectifs. Il sera donc nécessaire d'adapter votre code si celui-ci ne permettait que de traiter des instances à 2 objectifs.

Notez aussi que la méthode PLS nécessite de nombreuses mises à jour d'ensembles de solutions mutuellement non-dominées. Pour des problèmes à plus de 2 objectifs, il est donc opportun d'utiliser des structures de données adéquates pour la mise à jour, telles que ND-Tree ou Quad-Tree (voir pour cela la dernière partie du cours du 3 octobre, qui est également disponible à partir de la vidéo de présentation suivante : <https://youtu.be/81c1XoR5UQk>).

Lors de la phase (2), à partir de l'approximation de l'ensemble de solutions non-dominées obtenu par PLS, vous appliquerez une méthode d'aide à la décision par élicitation incrémentale des poids des critères, pour une fonction d'agrégation $\Psi_\lambda(y(x))$ linéaire en ses paramètres (voir l'article articleProjet1.pdf, sections 3.1 et 3.2). On choisira un vecteur de poids qui reste inconnu de la procédure d'élicitation pour simuler les réponses du décideur à des questions portant sur ses préférences.

Pour cette partie, vous étudierez les deux points suivants :

- a) Évaluation du nombre de questions nécessaires pour identifier la solution optimale pour le décideur (parmi les solutions générées en phase 1).
- b) Étude de la variation du regret minimax en fonction du nombre de questions posées (pour calculer les regrets maximum, on utilisera le solveur Gurobi (<https://www.gurobi.com/>) qui est installé dans les salles en libre-service mais peut aussi être installé sur des machines personnelles en le téléchargeant depuis une adresse de l'université).

3. Une deuxième procédure de résolution

La deuxième procédure de résolution est une procédure où élicitation incrémentale et recherche locale sont combinées afin de réduire le nombre de solutions générées et de pouvoir résoudre des problèmes de plus grande taille.

Le principe est le suivant : comme pour PLS, en partant d'une solution admissible générée aléatoirement, vous appliquerez une fonction de voisinage permettant de générer de nouvelles solutions admissibles. En revanche, contrairement à PLS, vous déterminerez la meilleure solution parmi l'ensemble des voisins et la solution courante (solution à partir de laquelle le voisinage a été appliqué) à l'aide d'une procédure d'élicitation incrémentale. Si la meilleure solution n'est pas la solution courante mais une solution dans son voisinage, vous relancerez la procédure à partir de cette solution voisine. Sinon la procédure s'arrête car un optimum local a été obtenu (voir l'article articleProjet2.pdf, section 4.1).

4. Données et comparaison des deux méthodes

Vous trouverez dans la section Projet du site Moodle un fichier de données nommé "2KP200-TA-0.dat". Ce fichier contient les données d'une instance à 200 objets, où chaque objet est évalué selon 6 objectifs. Pour tester vos méthodes, vous pourrez utiliser des sous-ensembles de ces données, par exemple prendre uniquement les 20 premiers objets et les 3 premiers critères de façon à avoir un problème à $n=20$ objets et $p=3$ critères. La capacité W du sac à dos sera toujours prise égale à $\lfloor \frac{\sum_{i=1}^n w_i}{2} \rfloor$.

Pour l'élicitation incrémentale, vous testerez en priorité comme fonction d'agrégation la somme pondérée, ensuite éventuellement OWA avec poids décroissants et l'intégrale de Choquet avec une

capacité super-modulaire (les conditions imposées visent à favoriser les vecteurs de performances équilibrés).

Vous comparerez expérimentalement la première et la seconde procédure de résolution en termes de :

- temps de calcul,
- erreur par rapport à la solution optimale du décideur (obtenue en optimisant un programme linéaire ayant pour fonction objectif la fonction d'agrégation),
- nombre de questions posées.

Vous ferez la moyenne des résultats obtenus pour au moins 20 jeux de paramètres différents de la fonction d'agrégation.

5. Organisation et dates

Le travail est à effectuer en binôme. Le choix du langage de programmation est libre. Toutes les possibilités de visualisation sont les bienvenues.

Les projets doivent être soumis le **dimanche 21 janvier 2024** au plus tard à partir du lien de soumission du site Moodle. Votre livraison sera constituée d'une archive **zip** qui comportera les sources du programme, un fichier README détaillant comment compiler et exécuter le programme, et un rapport (un fichier au format **pdf**) comportant une description synthétique des procédures implantées et les résultats des différents tests numériques menés. Le plan du rapport suivra le plan du sujet.

Une **soutenance** avec transparents est prévue lors de la séance du **23 janvier**.