Étude de deux règles d'agrégation de préférences fondées sur les choix sur les sous-ensembles d'alternatives du point du vue de la rationalité et de l'algorithmique

Martin Durand, Fanny Pascual, Olivier Spanjaard

Sorbonne Université, LIP6, Paris, France

{martin.durand,fanny.pascual,olivier.spanjaard}@lip6.fr

Mots-clés : *choix social computationnel, décision collective, agrégation de préférences.*

1 Introduction

L'agrégation de préférences consiste en la création d'une préférence collective sur un ensemble d'alternatives à partir de préférences individuelles, exprimées sous la forme de rangements. L'ensemble de ces préférences est appelé profil de préférences. L'application la plus évidente se trouve en théorie du choix social, où les alternatives sont des candidats et les préférences individuelles correspondent aux rangements des candidats selon chaque votant : les procédures d'agrégation sont alors appelées règles de vote. Cependant l'agrégation de préférences a bien d'autres applications (agrégation de résultats de moteurs de recherche, apprentissage de préférences, ...).
Par exemple, on peut considérer un ensemble de 7 votants exprimant leurs préférences sur un ensemble de 3 candidats $\{1, 2, 3\}$:

— 3 votants : $1 \succ 2 \succ 3$ (3 votants préfèrent le candidat 1, puis 2, puis 3, dans cet ordre)
— 2 votants : $2 \succ 1 \succ 3$
— 2 votants : $2 \succ 3 \succ 1$

Notre but est alors de retourner un rangement de consensus, c'est-à-dire un rangement satisfaisant au mieux l'ensemble des votants.

Une des règles de vote les plus étudiées est la règle de Kemeny [4]. Cette règle retourne parmi tous les rangements possibles un rangement minimisant le nombre de désaccords sur les paires de candidats avec les préférences exprimées. On compte un désaccord entre un rangement r et une préférence exprimée p lorsque, pour un couple de candidats $\{a, b\}$, le candidat a est préféré au candidat b dans r alors qu'au contraire b est préféré à a dans p. En sommant sur tous les couples de candidats et sur toutes les préférences exprimées, on obtient le score de Kemeny, que l'on cherche à minimiser.

Une autre approche pour les procédures d'agrégation est de considérer qu'il existe un rangement « véritable » ou « correct » et que les préférences exprimées correspondent à des observations bruitées de ce rangement [2]. On cherche alors à retrouver le « vrai » rangement, qui nous est inconnu, à partir de ces observations bruitées. Étant donné un vrai rangement, il est plus ou moins probable d'observer un ensemble de préférences exprimées par les votants. On choisira donc le rangement pour lequel le profil de préférences observé est le plus probable : un tel rangement est appelé rangement de *maximum de vraisemblance*.

2 Problématique

La règle de Kemeny est très étudiée en raison de ses nombreuses bonnes propriétés axiomatiques. Cette règle a cependant un défaut qui a été soulevé à plusieurs reprises : plusieurs rangements assez différents entre eux peuvent être simultanément optimaux. Plusieurs règles inspirées de la règle de Kemeny ont donc été proposées. On s'intéresse à une règle récemment proposée, appelée règle de k-*wise* Kemeny [3]. On cherche cette fois, parmi tous les rangements possibles, un rangement minimisant le nombre de désaccords avec le profil de préférences sur les sous-ensembles de candidats de taille k ou moins. On comptabilise un désaccord sur un sous-ensemble de candidats entre un rangement r et une préférence p lorsque le candidat placé en tête de ce sous-ensemble dans r et dans p est différent. Lorsque $k = 2$, on retrouve la règle de Kemeny, car on minimise les désaccords sur les paires de candidats. Lorsque $k = 3$, on minimise les désaccords sur les paires et sur les triplets de candidats.
Cette règle pallie bien le défaut de la règle de Kemeny mentionné plus haut car on a moins souvent plusieurs rangements optimaux et ils sont plus semblables entre eux. Cependant, on peut se demander s'il est légitime de comptabiliser les désaccords sur tous les sous-ensembles de candidats de la même façon, indépendamment de leur taille. On a donc proposé et motivé deux règles d'agrégation proches de k-*wise* Kemeny, mais comptabilisant les désaccords sur les sous-ensembles de candidats différemment selon la taille du sous-ensemble considéré.

3 Contribution

On introduit d'abord la règle de k-*exact Kemeny*, qui retourne un rangement minimisant le nombre de désaccords avec le profil de préférences sur les sous-ensembles de candidats de taille exactement k. Dans ce cas, comme tous les sous-ensembles de candidats considérés sont de même cardinalité, on compte tous les désaccords avec le même poids. Nous en avons étudié les propriétés axiomatiques, notamment au regard du théorème d'Arrow [1], et nous avons montré qu'il est NP-difficile de trouver un rangement minimisant les désaccords sur les sous-ensembles de taille k. Nous avons également proposé un algorithme de résolution exacte et deux heuristiques, dont une heuristique polynomiale fondée sur une chaîne de Markov, construite à partir du profil de préférences, qui approche très bien la règle exacte.
Ensuite, on s'intéresse à une nouvelle version de la règle de k-*wise* Kemeny, en pondérant les désaccords en fonction de la taille du sous-ensemble de candidats considéré. Dans un premier temps, on détaille les hypothèses sur le comportement des votants permettant d'interpréter le rangement de k-wise Kemeny comme un rangement de maximum de vraisemblance. Ces hypothèses étant fortes, nous en avons étudiée une version relâchée plus satisfaisante. Nous avons ensuite défini une pondération de la règle de k-wise Kemeny cohérente avec ces nouvelles hypothèses. Nous avons finalement proposé une heuristique polynomiale pour approcher cette règle, elle aussi fondée sur une chaînes de Markov construite à l'aide du profil de préférences.
Toutes les heuristiques ont été évaluées expérimentalement sur des instances réelles et sur des instances synthétiques.

Références

[1] Kenneth Arrow. *Social choice and individual values. Yale University Press*, 2012.

[2] Vincent Conitzer, Matthew Rognlie, and Lirong Xia. *Preference functions that score rankings and maximum likelihood estimation.* In *IJCAI*, 2009.

[3] Hugo Gilbert, Tom Portoleau, and Olivier Spanjaard. *Beyond Pairwise comparisons in social choice : A setwise kemeny aggregation problem.* In *AAAI*, 2019.

[4] John G. Kemeny. *Mathematics without numbers. Daedalus*, 1959.