International Journal of Approximate Reasoning 101 (2018) 181–205

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

2-Monotone outer approximations of coherent lower probabilities

Ignacio Montes [a,*], Enrique Miranda [a], Paolo Vicig [b]

[a] University of Oviedo, Department of Statistics and Operations Research, Spain
[b] University of Trieste, DEAMS, Italy

ARTICLE INFO

Article history:
Received 20 April 2018
Received in revised form 10 July 2018
Accepted 12 July 2018
Available online 24 July 2018

Keywords:
Coherent lower probabilities
2-monotonicity
Coherent lower previsions
Probability intervals
Distortion models
Total variation distance

ABSTRACT

We investigate the problem of approximating a coherent lower probability on a finite space by a 2-monotone capacity that is at the same time as close as possible while not including additional information. We show that this can be tackled by means of a linear programming problem, and investigate the features of the set of undominated solutions. While our approach is based on a distance proposed by Baroni and Vicig, we also discuss a number of alternatives: quadratic programming, extensions of the total variation distance, and the Weber set from game theory. Finally, we show that our work applies to the more general problem of approximating coherent lower previsions.

1. Introduction

Since they were thoroughly studied by Peter Walley in [29] as an extension of Bruno de Finetti's work on subjective probability [9], coherent lower previsions have been considered one of the most general approaches to deal with imprecision and uncertainty.

They possess a number of advantages: first of all, they can be represented equivalently as closed and convex sets of probability measures, and as such they can be given an epistemic interpretation as a model for the imprecise knowledge of a probability measure. Moreover, the representation of these sets in terms of lower and upper envelopes allows for a number of mathematical advantages, for instance in the extension of the assessments to a greater domain.

Secondly, they can also be given a clear behavioural interpretation in terms of acceptable betting rates, thus extending de Finetti's approach to be able to deal with indecision, something that arises frequently in cases where the available knowledge is imprecise.

Thirdly, they include as particular cases most of the models of non-additive measures that have been proposed in the literature, such as probability intervals [7], belief functions [25] or possibility measures [14]. It is thus possible to work with these particular models using all the machinery that has already been developed for coherent lower previsions.

In spite of these advantages, coherent lower previsions (or their restrictions to events, called *coherent lower probabilities*) also have a number of drawbacks that hinder their use in practice: for instance, they have no general easy representation in terms of the extreme points of their associated sets of probabilities, and they sometimes lack some attractive mathematical properties possessed by some more specific models.

* Corresponding author.
E-mail addresses: imontes@uniovi.es (I. Montes), mirandaenrique@uniovi.es (E. Miranda), paolo.vicig@deams.units.it (P. Vicig).

https://doi.org/10.1016/j.ijar.2018.07.004

One alternative that somewhat solves these issues while being sufficiently general is to work with 2-monotone capacities, which can be easily determined on finite spaces by means of a finite number of extreme points [26] and that still include as particular cases many of the imprecise probability models from the literature [1]. It is therefore interesting to determine if we can approximate a coherent lower probability by a 2-monotone one with a minimal loss of information. This problem was already considered by Bronevich and Augustin in [3], and they gave two algorithms that provide an outer approximation that is optimal in the sense we shall discuss later on.

Besides the above mentioned motivations, we stress that the approximation problem we are concerned with is relevant for practical purposes too. In particular, it is related to the issue of how to exchange information among agents adopting different uncertainty formalisms discussed e.g. in [2]. In fact, it solves this problem in the case of a sending agent, or more generally an uncertainty interchange format common to an open community of agents, operating with coherent imprecise probabilities, and a receiving agent adopting 2-monotone probabilities or some special cases of theirs.

After recalling some preliminary concepts in Section 2, in Section 3 we study the problem of finding undominated outer approximations that minimize the distance to the original model, in the sense proposed by Baroni and Vicig in [2]. In Section 4, we focus on outer approximations by means of some particular subfamilies of 2-monotone capacities and prove that this problem has a unique solution in those cases. A comparison with other approaches is given in Section 5. In Section 6 we show that our results allow to solve also the problem of outer approximating coherent lower previsions. Finally, in Section 7 we take a brief look at the problem of finding *inner* approximations of a given coherent lower probability. Some additional comments are provided in Section 8. In order to ease the reading, proofs, as well as auxiliary lemmas, have been gathered in an appendix.

2. Preliminary concepts

Let $\mathcal{X}$ be a finite space with cardinality $|\mathcal{X}| = n$ and powerset $\mathcal{P}(\mathcal{X})$. A *lower probability* is a function $\underline{P} : \mathcal{K} \subseteq \mathcal{P}(\mathcal{X}) \to [0,1]$ defined on some subsets (events) of $\mathcal{X}$. Following Walley [29], a lower probability can be given a behavioural interpretation, so that $\underline{P}(A)$ is our supremum acceptable betting rate on event A (that is, the supremum amount of money we would pay for a gamble with reward 1 when A occurs and 0 when it does not). On the other hand, lower probabilities can also be given an epistemic interpretation, in situations of imprecise knowledge about the probability of some events. Then the lower probability is understood as a lower bound of an ideal, but unknown, probability measure, and our information about this measure can be equivalently represented by means of the set of probability distributions that are compatible with the information given by $\underline{P}$. This set is called *credal set*, and it is defined by:

$$\mathcal{M}(\underline{P}) = \{P \text{ probability measure} \mid P(A) \geq \underline{P}(A) \quad \forall A \in \mathcal{K}\}. \tag{1}$$

Some usual consistency requirements are imposed to lower probabilities. One of the simplest is *avoiding sure loss*, which means that there is at least one probability compatible with $\underline{P}$, or equivalently, that $\mathcal{M}(\underline{P}) \neq \emptyset$. A stronger requirement is *coherence*, which means that the bounds $\underline{P}$ gives for the probabilities of the different events are tight:

Definition 1. [29] A lower probability $\underline{P} : \mathcal{K} \subseteq \mathcal{P}(\mathcal{X}) \to [0,1]$ is called *coherent* when $\underline{P}(A) = \min\{P(A) : P \in \mathcal{M}(\underline{P})\}$ for every $A \in \mathcal{K}$.

The conjugate of a lower probability $\underline{P}$ on $\mathcal{K}$ is the function $\overline{P} : \mathcal{K}^c \subseteq \mathcal{P}(\mathcal{X}) \to [0,1]$ given by $\overline{P}(A) = 1 - \underline{P}(A^c)$ for every $A \in \mathcal{K}^c$, where $\mathcal{K}^c = \{A^c \mid A \in \mathcal{K}\}$. $\overline{P}$ is called an *upper probability*, and $\overline{P}(A)$ can be understood as an upper bound for the unknown probability of A, or, under the behavioural interpretation, as the infimum betting rate against A. The conjugacy relation implies that the credal set given in Eq. (1) can be equivalently represented by:

$$\mathcal{M}(\underline{P}) = \{P \text{ probability} \mid P(A) \leq \overline{P}(A) \quad \forall A \in \mathcal{K}^c\}.$$

In addition, $\underline{P}$ is coherent if and only if $\overline{P}(A) = \max_{P \in \mathcal{M}(\underline{P})} P(A)$ for every $A \in \mathcal{K}^c$.

Throughout this paper, we shall consider conjugate and coherent lower and upper probabilities $\underline{P}, \overline{P}$. Moreover, we shall assume that they are defined on the power set $\mathcal{P}(\mathcal{X})$. This assumption entails no loss of generality, since coherent models on a proper subset of $\mathcal{P}(\mathcal{X})$ can always be extended to the power set by means of the notion of *natural extension* (see [29] for more details):

$$\underline{P}(A) = \min\{P(A) : P \in \mathcal{M}(\underline{P})\} \quad \forall A \subseteq \mathcal{X}$$
$$\overline{P}(A) = \max\{P(A) : P \in \mathcal{M}(\underline{P})\} \quad \forall A \subseteq \mathcal{X}.$$

If $\underline{P}, \overline{P}$ are coherent and conjugate on $\mathcal{P}(\mathcal{X})$, they satisfy the following properties [29, Section 2.7.4]:

(C1) $0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1$ for every $A \subseteq \mathcal{X}$.
(C2) $\underline{P}(A) \leq \underline{P}(B)$ and $\overline{P}(A) \leq \overline{P}(B)$ for every $A, B \subseteq \mathcal{X}$ such that $A \subseteq B$.
(C3) $\overline{P}(A \cup B) \leq \overline{P}(A) + \overline{P}(B)$ for every $A, B \subseteq \mathcal{X}$.