Characterizing the set of coherent lower previsions with a finite number of constraints or vertices

Erik Quaeghebeur

Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA, USA

SYSTeMS Research Group, Ghent University, Gent, Belgium

1 Lower previsions

Lower previsions are uncertainty models that generalize classical expectation operators. Each is equivalent to a credal set, a set of probabilities. They are useful for situations in which it is unrealistic to assess a single probability distribution.

2 Setup & notation

The setup is finitary: Both the possibility space Ω and the set $\mathscr{K}$ of gambles, i.e., real-valued functions f on Ω whose lower prevision interests us, are finite.

Notation: id is the identity function, supp is the support operator, and $\langle \mu, \phi \rangle_{\mathscr{X}}$ is a shorthand for $\sum_{x \in \mathscr{X}} \mu_x \cdot \phi x$.

The vacuous lower prevision $\underline{P}_A$ relative to $A \subseteq \Omega$ is defined by $\underline{P}_A f = \min_{\omega \in A} f\omega$ and for ω in Ω by $\underline{P}_\omega := \underline{P}_{\{\omega\}}$.

3 Coherence

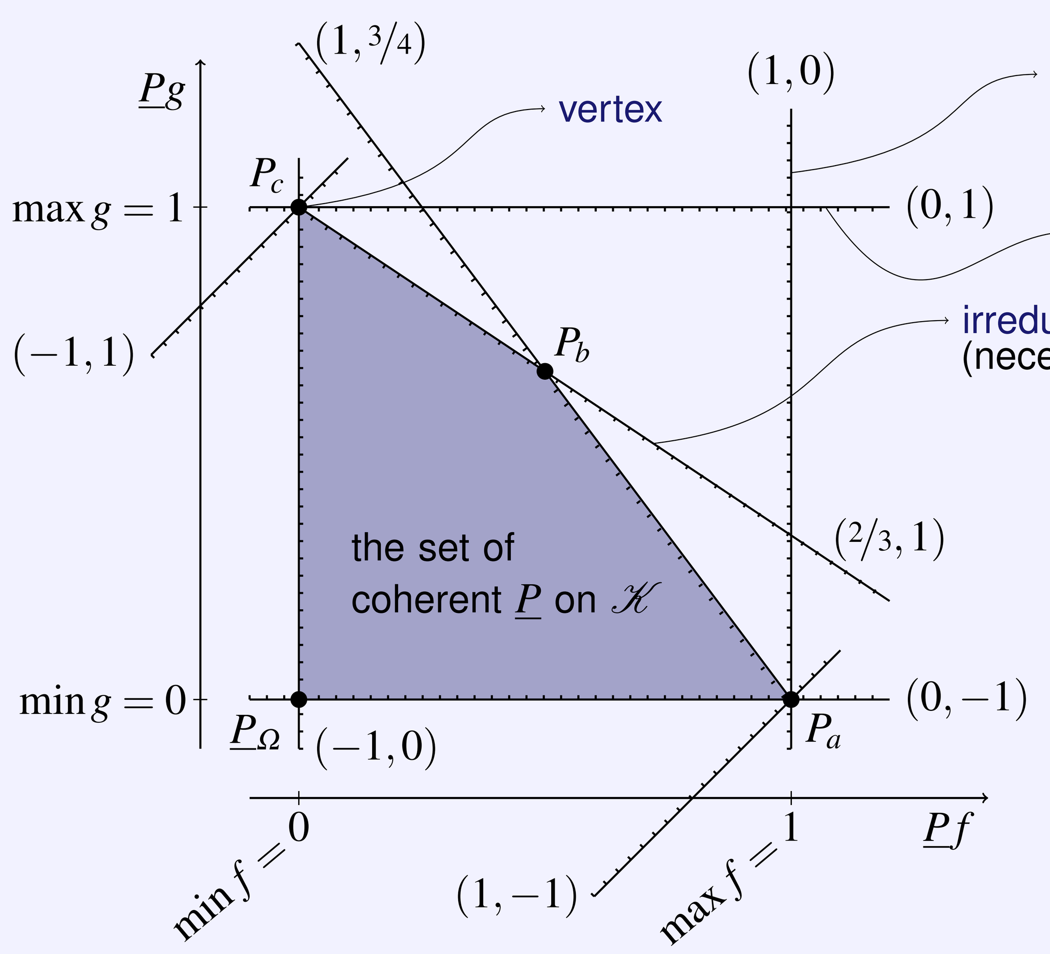
Coherence for lower previsions is the analogue of Kolmogorov's laws for probability measures or de Finetti's avoiding sure loss for linear previsions. (Def. 1: see paper.)

Definition 2. A lower prevision $\underline{P}$ on $\mathscr{K}$ is coherent iff $\langle \lambda, \underline{P} \rangle_{\mathscr{K}} \leq \max \langle \lambda, \mathrm{id} \rangle_{\mathscr{K}}$ for all coefficient vectors λ in $\mathbb{R}^{\mathscr{K}}$ with at most one strictly negative component.

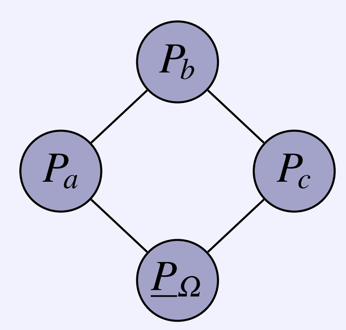
Each λ generates a linear constraint on $\underline{P}$; there are an infinity of them.

4 A toy example

Let $\Omega := \{a, b, c\}$ and $\mathscr{K} := \{f, g\}$ with $f := (1, 1/2, 0)$ and $g := (0, 2/3, 1)$. Each linear constraint—closed half-space—in $(\underline{P}f, \underline{P}g)$-space is completely determined by $\lambda = (\lambda_f, \lambda_g)$. We draw a sufficient finite subset of them.

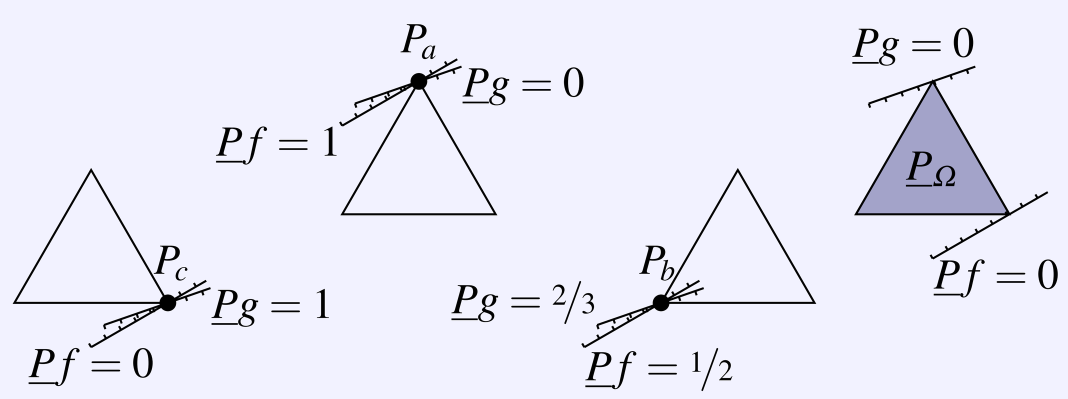


Adjacency graph



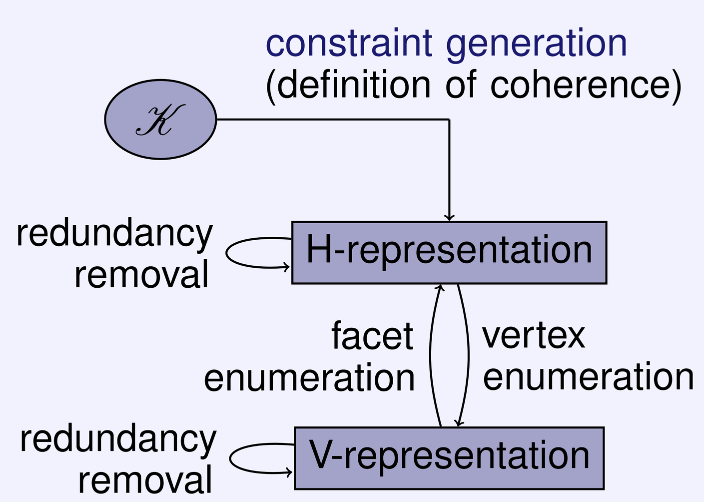
Credal sets

A probability mass function p belongs to the credal set of $\underline{P}$ if $\langle p, f \rangle_\Omega \geq \underline{P}f$ and $\langle p, g \rangle_\Omega \geq \underline{P}g$ hold.



5 Polytope theory

The Minkowski–Weyl theorem tells us that a polytope can be represented both by an H-representation, a set of linear constraints, and a V-representation, a set of vertices (points). Good implementations are available of algorithms for going from one to the other, and for removing redundant constraints (or points).



CONCEPTS & CONTEXT

6 Goal

Characterize the set of coherent lower previsions on $\mathscr{K}$, a polytope, in a finitary way: with a finite number of linear constraints or with a finite number of vertices.

For classical probability theory both characterizations are known: the set of all probability mass functions is the unit simplex. It is defined by the positivity and normalization constraints and it is the convex hull of the degenerate mass functions.

7 Approach

Reformulate Definition 2 into a criterion that generates a characterizing finite number of constraints and generate the characterizing finite number of vertices from it. This reformulation is done stepwise.

8 Normalization

In Definition 2, the coefficient vector λ appears on both sides of the constraint's inequality. It can therefore be normalized such that $\max \langle \lambda, \mathrm{id} \rangle_{\mathscr{K}} \in \{-1, 0, 1\}$.

It is a known consequence of coherence that $\underline{P}(\lambda \cdot g + \alpha) = \lambda \cdot \underline{P}g + \alpha$ for all gambles g in $\mathscr{K}$, all nonnegative real λ and all real α. Eliminating one gamble of pairs related by such an affine transformation from $\mathscr{K}$ does not essentially change the polytope of coherent lower previsions. So we can restrict attention to gambles in

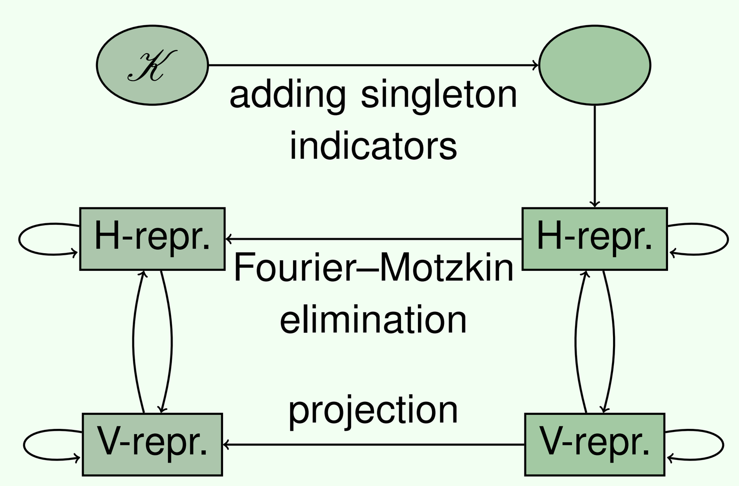
$$\mathscr{L} := \{g \in \mathbb{R}^\Omega : \min g = 0 \wedge \max g = 1\}.$$

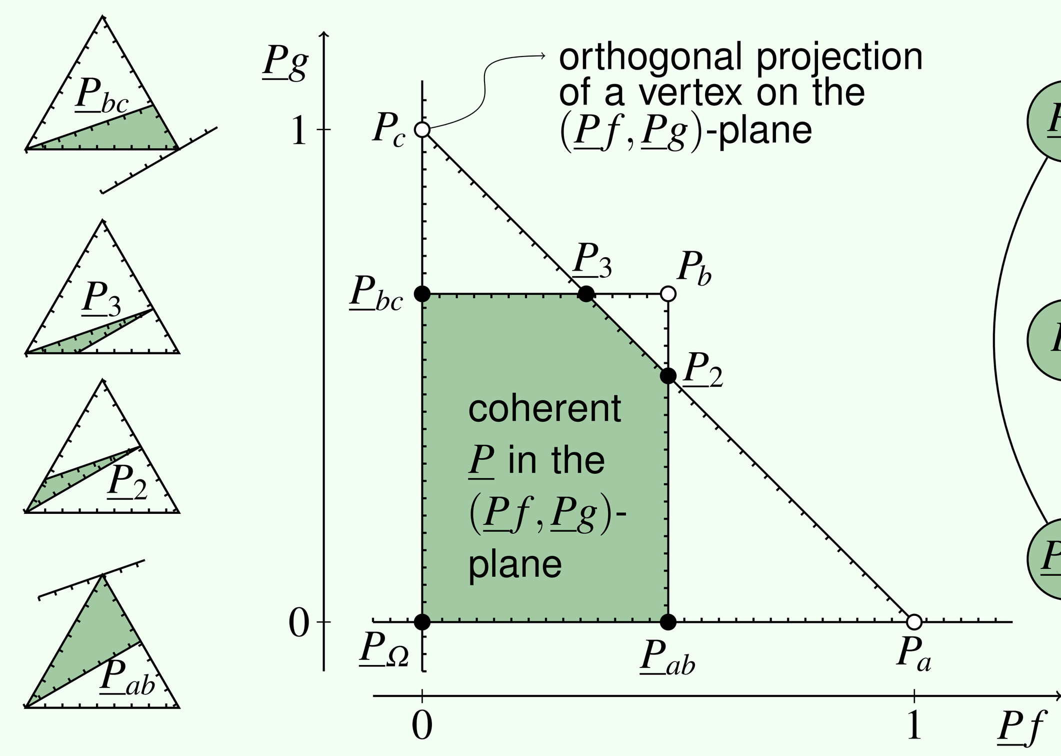
Dealing with such nonnegative gambles only, it follows from coherence that $\underline{P} \geq 0$. This allows us to restrict attention to λ whose components cannot be increased without increasing $\max \langle \lambda, \mathrm{id} \rangle_{\mathscr{K}}$, which can be either 0 or 1.

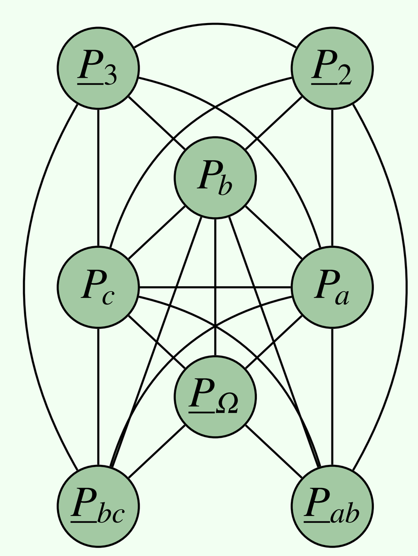
9 Singleton indicators

Singleton indicators are gambles that are one in some ω of Ω and zero elsewhere. Including them in $\mathscr{K}$ allows us to further restrict attention to the λ for which $\langle \lambda, \mathrm{id} \rangle_{\mathscr{K}}$ is constant, i.e., either 0 or 1.

Polytope theory algorithms allow us to return to the H- and V-representations of the original $\mathscr{K}$.







10 Linear independence

Two technical lemmas allow us to express each remaining constraint using a linearly independent subset of $\mathscr{K}$, which are but finite in number.

The final, finitary criterion we obtain is (Defs. 3–8: see paper.):

Definition 9. A lower prevision $\underline{P}$ on $\mathscr{K}$, a finite subset of $\mathscr{L}$ that contains all singleton indicators, is coherent iff

(i) $\underline{P} \geq 0$,

(ii) $\langle \lambda, \underline{P} \rangle_{\mathscr{N}} \leq \underline{P}f$ for all
- linearly independent subsets $\mathscr{N}$ of $\mathscr{K}$ such that $1 < |\mathscr{N}| \leq |\Omega|$,
- f in $\mathscr{K} \setminus \mathscr{N}$ s.t. $\operatorname{supp} f = \operatorname{supp} \mathscr{N}$,
- λ in $(\mathbb{R}_{>0})^{\mathscr{N}}$ s.t. $\langle \lambda, \mathrm{id} \rangle_{\mathscr{N}} = f$.

(iii) $\langle \lambda, \underline{P} \rangle_{\mathscr{N}} \leq 1$ for all
- linearly independent subsets $\mathscr{N}$ of $\mathscr{K}$ such that $1 < |\mathscr{N}| \leq |\Omega|$ and $\operatorname{supp} \mathscr{N} = \Omega$.
- λ in $(\mathbb{R}_{\neq 0})^{\mathscr{N}}$ s.t. $\langle \lambda, \mathrm{id} \rangle_{\mathscr{N}} = 1$ with at most one strictly negative component.

APPROACH & CHARACTERIZATION

11 Illustrative results

We present the combinatorics of a number of cases—sets of gambles $\mathscr{K}$—for which we used Definition 9 to generate a sufficient set of constraints, a redundancy removal algorithm to isolate the irredundant (necessary) ones, and a vertex enumeration algorithm to obtain the corresponding vertices, which are in this context called extreme coherent lower previsions.

Notation: $\#\lambda$ indicates the number of irredundant constraints (when of interest, the total number generated is in parentheses) and $\#\underline{P}$ the number of vertices.

All events

When $\mathscr{K}$ consists of the $2^{|\Omega|} - 2$ indicators for all nontrivial events:

$\lvert\Omega\rvert$	2	3	4	5
$\#\lambda$	3 (3)	9 (17)	48 (179)	285 (7351)
$\#\underline{P}$	3	8	402	> 1743093

Both $\#\lambda$ and $\#\underline{P}$ exhibit a combinatorial explosion. The resulting class of lower previsions is defined by an imprecise probability.

Values-based gambles

When we vary the number of gambles in $\mathscr{K}$ for the case $|\Omega| = 3$ by considering all $2 \cdot k \cdot |\Omega|$ gambles in $\mathscr{L}$ that take values in $\{\ell/k : 0 \leq \ell \leq k\}$:

k	2	3	4	5	6
$\lvert\mathscr{K}\rvert$	12	18	24	30	36
$\#\lambda$	15 (178)	21 (699)	27 (1796)	33 (3685)	39 (6582)
$\#\underline{P}$	49	180	455	928	1653

Observed patterns: $\#\lambda = 3 \cdot (2 \cdot k + 1)$ and $\#\underline{P} = (3 \cdot k + 1) \cdot (3 \cdot k^2 - 4 \cdot k + 3)$.

Singletons

When $\mathscr{K}$ consists of all $|\Omega|$ singleton indicators, $\#\lambda = \#\underline{P} = |\Omega| + 1$. The resulting class of lower previsions is defined by a lower probability mass function.

Singleton complements

When $\mathscr{K}$ consists of all $|\Omega|$ singleton complement indicators, $\#\lambda = 2 \cdot |\Omega| + 1$ and $\#\underline{P} = 2^{|\Omega|} - 1$. The resulting class of lower previsions is defined by an upper probability mass function.

Singletons and their complements

When $\mathscr{K}$ consists of all $2 \cdot |\Omega|$ singleton and singleton complement indicators:

$\lvert\Omega\rvert$	2	3	4	5	6	7	8	9	10
$\#\lambda$	3	9	16	20	24	28	32	36	40
$\#\underline{P}$	3	8	20	47	105	226	474	977	1991

Observed pattern: $\#\lambda = 4 \cdot |\Omega|$ (after a transient); $\#\underline{P}$ exhibits a combinatorial explosion. The resulting class of lower previsions is defined by so-called probability intervals.

12 Checking coherence

Definition 9 can be used for checking the coherence of lower previsions. As compared to other coherence checking approaches, it is especially useful when checking a large number of lower previsions on the same set of gambles, because the constraints only need to be generated once, ever.

Currently, constraint generation is practically feasible if both of $(|\Omega|, |\mathscr{K}|)$'s components are not too large. Concretely, $(6, 62)$ and $(11, 22)$ pose problems on a now-standard PC, whereas $(5, 30)$ and $(10, 20)$ can be handled. The generated sets contain a relatively large portion of redundant constraints, so there is room for improvement of the criterion in this regard.

13 Approximation

We can approximate a coherent lower previsions using its decomposition in terms of extreme coherent lower previsions. However, because their number is usually quite large and because both obtaining them and calculating a decomposition are computationally intensive, this idea is currently not practical for all but academic cases.

14 Implementation

We have implemented Definition 9, as a constraint generator, in Python, using the NumPy package for its linear algebra codes. As implementations of the necessary polytope theory algorithms, we use the freely available 'cdd' and 'lrs' bundles.

RESULTS & APPLICATIONS