Error Bounds for Finite Approximations of Coherent Lower Previsions

Damjan Škulj

University of Ljubljana

ISIPTA '17, Lugano
12 July 2017

Contents

University *of Ljubljana*

Contents

University of Ljubljana

Finite (imprecise) probability spaces

We study models with the following elements:

- sample space $\mathcal{X}$: a finite set with elements $x \in \mathcal{X}$;
- gamble: any map $f\colon \mathcal{X} \to \mathbb{R}$ or a vector in $\mathbb{R}^{\mathcal{X}}$;
- an arbitrary set of gambles $\mathcal{K}$;
- (precise) probability vector $p \in \mathbb{R}^{\mathcal{X}}$ satisfying $p(x) \geq 0\, \forall x \in \mathcal{X}$ and $\sum_{x\in\mathcal{X}} p(x) = 1$;
- linear prevision (expectation functional) $P\colon \mathcal{K} \to \mathbb{R}$ of the form $P(f) = \sum_{x\in\mathcal{X}} p(x)f(x) = p \cdot f$ where p is a precise probability vector;
- coherent lower prevision $\underline{P}\colon \mathcal{K} \to \mathbb{R}$ is a lower envelope of linear previsions.

University *of Ljubljana*

Coherent lower previsions and lower expectation functionals

A coherent lower prevision $\underline{P}\colon \mathcal{K} \to \mathbb{R}$ can be expressed as a lower envelope of linear previsions

$$\underline{P}(f) = \min_{P \in \mathcal{M}(\underline{P})} P(f),$$

where $\mathcal{M}(\underline{P})$ is the credal set of $\underline{P}$:

$$\mathcal{M}(\underline{P}) = \{P\colon P(f) \geq \underline{P}(f) \forall f \in \mathcal{K}\}.$$

A coherent lower prevision can be extended to a lower expectation functional $\underline{E}\colon \mathbb{R}^{\mathcal{X}} \to \mathbb{R}$, which is a coherent lower prevision defined everywhere in $\mathbb{R}^{\mathcal{X}}$.

The minimal coherent extension is called the natural extension.

Lower expectation functionals therefore form a family of coherent lower previsions.

Contents

University *of Ljubljana*

Partially specified coherent lower prevision

- Let $\underline{P}$ be a coherent lower prevision on a set of gambles $\mathcal{H}$ (often $\mathcal{H} = \mathbb{R}^{\mathcal{X}}$).
- But we only know the values of $\underline{P}(f)\ \forall f \in \mathcal{K} \subset \mathcal{H}$.
- What can we say about $\underline{P}(h)$ for $h \in \mathcal{H} - \mathcal{K}$?
- The natural extension of $\underline{P}|_{\mathcal{K}}$ is often our best guess.

Problem

What is the maximal possible error that we make by taking the natural extension (or any other extension) instead of the true value $\underline{P}(h)$?

University of Ljubljana

Reformulation of the problem

Reformulation 1

What is the maximal possible distance between two coherent extensions of $\underline{P}|_{\mathcal{K}}$ to $\mathcal{H} \supset \mathcal{K}$?

Reformulation 2

What is the maximal possible distance between two coherent lower previsions on $\mathcal{H}$ which coincide on $\mathcal{K} \subseteq \mathcal{H}$?

Special case

What is the maximal distance between the natural extension and any other coherent extension of a coherent lower prevision $\underline{P}$ on $\mathcal{K}$?

Common examples

Imprecise probability models are often approximated by:

coherent lower probabilities (interval probabilities) $L(A)$ is the lower probability of an event A; i.e. $\mathcal{K} = \{1_A : A \subseteq \mathcal{X}\}$;

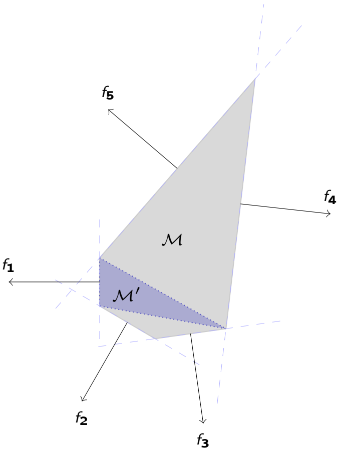
probability intervals intervals are given for the probabilities of atomic events $[l(x), u(x)]$; $\mathcal{K} = \{1_{\{x\}} : x \in \mathcal{X}\} \cup \{1_{\mathcal{X}-\{x\}} : x \in \mathcal{X}\}$;

The above models are often considered as good approximations of the completely specified coherent lower previsions.

Graphical illustration

Lower previsions $\underline{P}$ and $\underline{P}'$ with the credal sets $\mathcal{M}$ and $\mathcal{M}'$ respectively coincide on the set of gambles $\mathcal{K} = \{f_1, \ldots, f_5\}$.

(Note that $\underline{P}$ is the natural extension of $\underline{P}|_{\mathcal{K}}$.)



University of Ljubljana

Contents

University of Ljubljana

Mathematical formulation

Let $\underline{P}$ be a coherent lower prevision specified on a finite set of gambles $\mathcal{K}$.

- Let $\underline{P}_1$ and $\underline{P}_2$ be two extensions to $\mathbb{R}^{\mathcal{X}}$.
- The distance between $\underline{P}_1$ and $\underline{P}_2$ is defined as

$$d(\underline{P}_1, \underline{P}_2) = \max_{h \in \mathbb{R}^{\mathcal{X}}} \frac{|\underline{P}_1(h) - \underline{P}_2(h)|}{\|h\|},$$

where $\|\cdot\|$ denotes the Euclidean norm.

University *of Ljubljana*

The maximal distance to the natural extension

The following result simplifies the problem.

Theorem

Let

- $\underline{P}$ *be a coherent lower prevision specified on a finite set of gambles* $\mathcal{K}$*;*
- $\underline{E}$ *its natural extension;*
- $\underline{P}_1$ *and* $\underline{P}_2$ *another two extensions to* $\mathbb{R}^{\mathcal{X}}$*.*

Then

$$d(\underline{P}_1, \underline{P}_2) \leq max(d(\underline{P}_1, \underline{E}), d(\underline{P}_2, \underline{E}))$$

We thus try to find an upper bound for the rhs over all coherent extensions $\underline{P}_1$.

Contents

University of Ljubljana

Contents

University *of Ljubljana*

Bounding the distance

Recall that extrema w.r.t. credal sets are found in extreme points. Therefore:

$$d(\underline{P}, \underline{E}) = \max_{h \in \mathbb{R}^{\mathcal{X}}} \frac{\underline{P}(h) - \underline{E}(h)}{\|h\|}$$
$$= \max_{h \in \mathbb{R}^{\mathcal{X}}} \max_{E \in \mathrm{ext}\mathcal{M}(\underline{E})} \min_{P \in \mathrm{ext}\mathcal{M}(\underline{P})} \frac{P(h) - E(h)}{\|h\|},$$

where $\mathrm{ext}\cdot$ denotes the set of extreme points of a credal set.

Unfortunately, only the set of extreme points of $\mathcal{M}(\underline{E})$ is known, while $\mathcal{M}(\underline{P})$ is unspecified, as well as its extreme points.

We do assume that $\underline{P}$ is coherent, though. What does it tell us?

University *of Ljubljana*

A consequence of coherence

$\underline{P}$ is a coherent extension of $\underline{P}|_{\mathcal{K}}$, and therefore there must exist some $P \in \mathcal{M}(\underline{P})$ so that $P(f) = \underline{P}(f)$ for every $f \in \mathcal{K}$.

Thus, the face $\mathcal{M}_f = \{P \in \mathcal{M}(\underline{E}) \colon P(f) = \underline{P}(f)\}$ must intersect $\mathcal{M}(\underline{P})$.

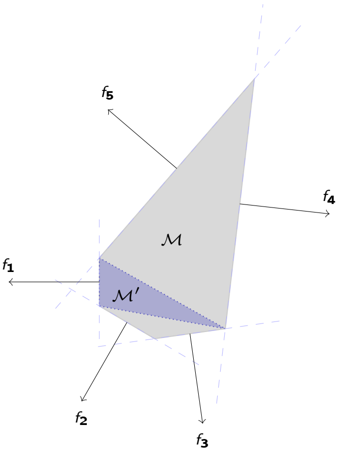
Consequently, the part of the expression used in the maximizing formula can be bounded as follows:

$$\min_{P \in \mathrm{ext}\,\mathcal{M}(\underline{P})} P(h) \leq \min_{f \in \mathcal{K}} \max_{P \in \mathrm{ext}\,\mathcal{M}_f} P(h)$$

Since $\mathcal{M}_f$ is a face of $\mathcal{M}(\underline{E})$, the rhs in the above inequality is obtainable in terms of extreme points of $\mathcal{M}(\underline{E})$.

Graphical illustration

Notice that $\mathcal{M}'$ intersects every face $\mathcal{M}_f$.
Otherwise, the corresponding lower prevision would not be coherent.



A final form of the optimization problem

Using the above estimates, we can now state:

$$d(\underline{P}, \underline{E}) \leq \max_{E \in \mathrm{ext}\mathcal{M}(\underline{E})} \min_{f \in \mathcal{K}} \max_{P \in \mathrm{ext}\mathcal{M}_f} \max_{h \in \mathbb{R}^{\mathcal{X}}} \frac{P(h) - E(h)}{\|h\|}$$

It is sufficient to restrict to those h that satisfy: $E(h) = \underline{E}(h)$, whence we obtain more restrictive error bound:

$$d(\underline{P}, \underline{E}) \leq \max_{E \in \mathrm{ext}\mathcal{M}(\underline{E})} \min_{f \in \mathcal{K}} \max_{P \in \mathrm{ext}\mathcal{M}_f} \max_{\substack{h \in \mathbb{R}^{\mathcal{X}} \\ E(h) = \underline{E}(h)}} \frac{P(h) - E(h)}{\|h\|}$$

We will therefore solve a maximization problem in a set called normal cone of E.

University of Ljubljana

Contents

University of Ljubljana

Normal cones

Let $\mathcal{M}$ be a credal set and $E \in \mathcal{M}$ an extreme point.

The set

$$N_{\mathcal{M}}(E) = \{f : E(f) = \underline{P}(f)\}$$

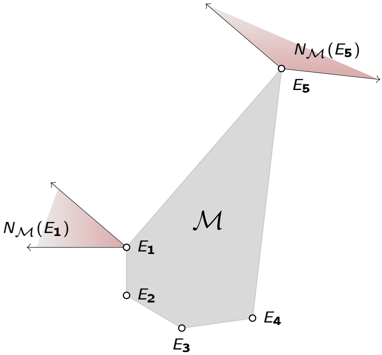
is called the normal cone of $\mathcal{M}$ at point E.

The normal cone is the set of all gambles that reach minimal expectation at E.

University of Ljubljana

Example: normal cones

Normal cones $N_{\mathcal{M}}(E_i)$ at the extreme points are the positive hulls of the normal vectors of adjacent faces.



University of Ljubljana

Contents

University *of Ljubljana*

Distance between extreme points

Let E be an extreme point of a credal set $\mathcal{M}$ and P another linear prevision in $\mathcal{M}$.

We will need to find the maximal possible distance

$$d_E(E, P) = \max_{h \in N_{\mathcal{M}}(E)} \frac{|P(h) - E(h)|}{\|h\|}.$$

The above distance is called the normed distance of P from E.

The reason for only considering elements of the normal cone is that in expression $\underline{P}(h)$ only those gambles will reach the minimal value in E.

Setting up the problem

Let $h \in N_{\mathcal{M}}(E)$. We can represent it as a positive combination:

$$h = \sum_{i \in I} \alpha_i f_i$$

where $I = \{i \colon f_i \in \mathcal{K}, E(f_i) = \underline{P}(f_i)\}$.
Recall that P and E are themselves vectors too, and therefore we can write:

$$P(h) - E(h) = (P - E) \cdot h = D \cdot h$$

We can also decompose

$$f_i = \lambda_i D + u_i.$$

We thus obtain vectors $\underline{\alpha} = (\alpha_i)_{i \in I}$ and $\underline{\lambda} = (\lambda_i)_{i \in I}$ and a matrix U whose rows are u_i.

We have:

$$h = (\underline{\alpha} \cdot \underline{\lambda})D + \underline{\alpha}U$$
$$\|h\|^2 = \|D\|^2\underline{\alpha\lambda\lambda}^t\underline{\alpha}^t + \underline{\alpha}UU^t\underline{\alpha}^t$$
$$P(h) - E(h) = D \cdot (\underline{\alpha} \cdot \underline{\lambda})D = (\underline{\alpha} \cdot \underline{\lambda})\|D\|^2.$$

Further denote $\Pi = \|D\|^2\underline{\lambda\lambda}^t + UU^t$, which is a symmetric positive semi-definite matrix.

Thus we would like to minimize the expression

$$\frac{(\underline{\alpha} \cdot \underline{\lambda})\|D\|^2}{\sqrt{\underline{\alpha}\Pi\underline{\alpha}^t}}$$

with respect to $\underline{\alpha}$.

Quadratic programming formulation

Since we may always multiply vector $\underline{\alpha}$ by a positive constant, we can always ensure the numerator in

$$\frac{(\underline{\alpha} \cdot \underline{\lambda})\|D\|^2}{\sqrt{\underline{\alpha}\Pi\underline{\alpha}^t}}$$

to be equal 1.

In this case, we can maximize the above expression by minimizing the norm:

$$\underline{\alpha}\Pi\underline{\alpha}^t$$

subject to

$$\begin{aligned}(\underline{\alpha} \cdot \underline{\lambda})\|D\|^2 &= 1\\ \underline{\alpha} &\geq 0\end{aligned}$$

University *of Ljubljana*

Final thoughts

- The method is practically applicable; unfortunately, highly computationally complex...
- Approximate methods might be computationally more efficient.

Thank you for your attention!!

Questions...

University of Ljubljana