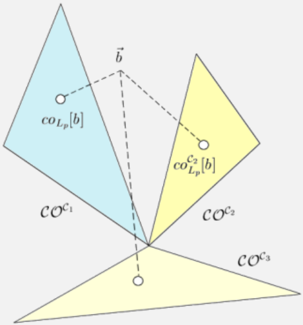
Consonant approximations in the belief space

Fabio Cuzzolin

Department of Computing and Communication Technologies
Oxford Brookes University

BELIEF 2012 - Compiegne, France
May 2012

Outline



1. Consonant approximation
2. A geometric approach to consonant approximation
3. Consonant approximation in the belief space
4. Interpretation and comparisons
5. Conclusions

Belief versus possibility measures

- a **belief function** $b : 2^\Theta \to [0,1]$ is such that,

$$b(A) = \sum_{B \subseteq A} m(B)$$

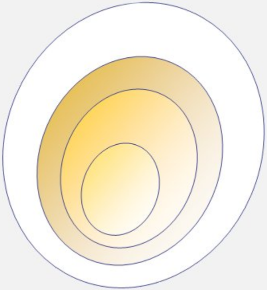
where $m : 2^\Theta \to [0,1]$ is a basic belief (probability) assignment s.t.

$$m(\emptyset) = 0, \sum_{A \subseteq \Theta} m(A) = 1, m(A) \geq 0 \ \forall A \subseteq \Theta$$

- a **possibility measure** $Pos : 2^\Theta \to [0,1]$ is such that $Pos(\bigcup_i A_i) = \sup_i Pos(A_i)$ for any family of subsets $\{A_i\}$
- the dual *necessity* measure is defined as $Nec(A) = 1 - Pos(A^c)$

Approximating a belief function with a possibility

- necessity measures have as counterparts in the theory of evidence **consonant** belief functions
- their focal elements are nested: $A_1 \subset \cdots \subset A_m$, $A_i \subseteq \Theta$



- they are **less computationally expensive** than b.f.s $\rightarrow$
- allow to **make decisions in the possibility framework**
- arguments for **inferring** consonant belief functions [Shafer76]

Previous work: outer approximations

- **outer** consonant approximations [Dubois&Prade 90]
- those consonant belief functions *co* such that $co(A) \leq b(A)$ $\forall A \subseteq \Theta$
- for each possible maximal chain $A_1 \subset \cdots \subset A_n$, $|A_i| = i$ of focal elements ...
- ... the **maximal outer consonant** approximation has mass $m_{\max}(A_i) = b(A_i) - b(A_{i-1})$
- mirrors the behavior of the vertices of the credal set of probabilities dominating a belief function [Chateauneuf, Miranda& Grabish]
- later extended by Baroni to capacities [Baroni 04]

Previous work: isopignistic approximation

- completely different approximation in Smets' Transferable Belief Model [Smets94,05]
- **pignistic transform** for decision making: $BetP(x) = \sum_{A \supseteq x} \frac{m_b(A)}{|A|}$
- **isopignistic" approximation**: the unique consonant belief function whose pignistic probability is identical to that of b [Dubois, Aregui]
- its contour function is:

$$pl_{iso}(x) = \sum_{x' \in \Theta} \min \Big\{ BetP[b](x), BetP[b](x') \Big\}.$$

- mass assignment: $m_{iso}(A_i) = i \cdot (BetP[b](x_i) - BetP[b](x_{i+1}))$ where $\{x_i\} = A_i \setminus A_{i-1}$

Goal: consonant transformation

- consonant transformations can be built by solving a **minimization problem**

$$co[b] = \arg \min_{co \in \mathcal{CO}} dist(b, co). \qquad (1)$$

- *dist* is some distance measure between belief functions, and $\mathcal{CO}$ denotes the collection of all consonant b.f.s
- *lugging in different distance functions in (1) we get different consistent transformations
- approach supported here: **distance between belief functions as vectors** of belief values:

$$\vec{b} = [b(A), \emptyset \subsetneq A \subseteq \Theta]$$

Geometry of belief functions

- each belief function can be seen as a vector:

$$\vec{b} = [b(A), \emptyset \subsetneq A \subsetneq \Theta]'$$

- the collection $\mathcal{B}$ of such vectors is a “simplex” (in rough words a higher-dimensional triangle), the **belief space**

$$\mathcal{B} = Cl(\vec{b}_A, \emptyset \subsetneq A \subseteq \Theta)$$

  which is the convex closure of all “categorical” belief functions such that $\vec{b}_A(A) = 1$, $\vec{b}_A(B) = 0 \; \forall B \neq A$

- alternatively we can use mass vectors $\vec{m}_b = [m_b(A), \emptyset \subsetneq A \subseteq \Theta]'$, living in a **mass space**: $\mathcal{M} = Cl(\vec{m}_A, \emptyset \subsetneq A \subseteq \Theta)$

Binary example

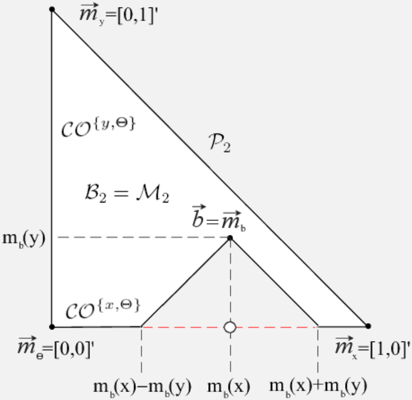


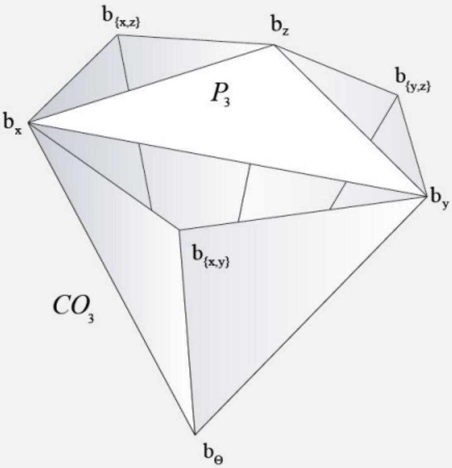
Figure: The belief/mass space $\mathcal{B}_2 = \mathcal{M}_2$ for a binary frame is the triangle above in $\mathbb{R}^2$.
Consonant b.f.s live in the **union of the two segments** $\mathcal{CO}^{\{x,\Theta\}}$ and $\mathcal{CO}^{\{y,\Theta\}}$.

The consonant simplex

- the geometry of consonant belief functions can be described by resorting to the notion of **simplicial complex**
- a collection Σ of simplices of arbitrary dimensions possessing the following properties:
  1. if a simplex belongs to Σ, then all its faces of any dimension belong to Σ;
  2. the intersection of any two simplices is a face of both the intersecting simplices.
- the region $\mathcal{CO}$ of consistent belief functions in the belief space **is a simplicial complex**:

$$\mathcal{CO} = \bigcup_{x \in \Theta} Cl(\vec{b}_A, A \ni x)$$

The consonant simplex: ternary case



Approximation on a simplicial complex

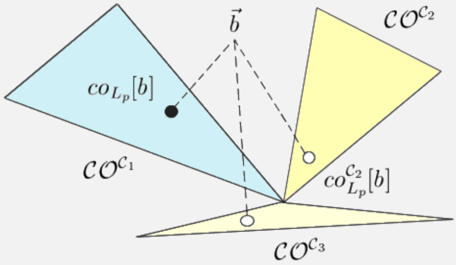


Figure: To minimize the distance of a point from a simplicial complex, we need to find **all the partial solutions on all the maximal simplices** in the complex (white circles), and compare these partial solutions to **select a global optimum** (black circle).

Choice of norm: L_p norms

- close relation between consonant belief functions and L_p norms
- consonant b.f.s relate to possibility distributions, which relate to the L_∞ (max) norm
- have been employed in different problems such as probability and consistent transformation [Cuzzolin SMC-B 07,09]
- in the consistent case generate focussed consistent transformations [Cuzzolin ECSQARU'11]
- classical L_p norms in $\mathcal{B}$

$$
\begin{aligned}
\|\vec{b} - \vec{b}'\|_{L_1} &= \sum_{A \subseteq \Theta} |b(A) - b'(A)|, \\
\|\vec{b} - \vec{b}'\|_{L_2} &= \sqrt{\sum_{\emptyset \subsetneq B \subsetneq \Theta} (b(B) - b'(B))^2} \\
\|\vec{b} - \vec{b}'\|_{L_\infty} &= \max_{A \subseteq \Theta} \{|b(A) - b'(A)|\}
\end{aligned}
$$

Other norms

- a number of other norms can be introduced
- generalizations to belief functions of the classical Kullback-Leibler divergence of two probability distributions P, Q:
  $D_{KL}(P|Q) = \int_{-\infty}^{\infty} p(x) \log(\frac{p(x)}{q(x)}) dx$
- measures based on information theory such as fidelity and entropy-based norms [Jousselme IJAR'11]
- many others have been proposed [diaz,jiang,khatibi,shi], exhaustive analysis huge task!

L_1 approximation

Theorem

*the **partial** L_1 **consonant approximations in** $\mathcal{B}$ with maximal chain of focal elements $\mathcal{C} = \{A_1 \subset \cdots \subset A_n, |A_i| = i\}$ are the co.b.f.s whose mass vectors live in the polytope:*

$$Cl\Big(\big[b^1, b^2 - b^1, \cdots, b^i - b^{i-1}, \cdots, 1 - b^{n-1}\big]' \ \Big| \ b^i \in \big\{\gamma^i_{int1}, \gamma^i_{int2}\big\} \ \forall i \Big),$$

where $\gamma^i_{int1}, \gamma^i_{int2}$ are the innermost (median) elements of the list

$$\mathcal{L}_i = \{b(A), A \supseteq A_i, A \not\supset A_{i+1}\}. \tag{2}$$

- the innermost values of the above lists (2) cannot be analytically identified in full generality
- the **global** L_1 approximation(s) maximize(s) the cumulative lower halves of the lists of belief values $\mathcal{L}_i$

L_2 approximation

Theorem

*the **partial** L_2 **consonant approximation of** b **in** $\mathcal{B}$ with maximal chain $\mathcal{C} = \{A_1 \subset \cdots \subset A_n\}$ is unique, and has mass*

$$m_{co^{\mathcal{C}}_{L_2}[b]}(A_i) = ave(\mathcal{L}_i) - ave(\mathcal{L}_{i-1}) \quad \forall i = 1, ..., n,$$

where

$$ave(\mathcal{L}_i) = \frac{1}{2^{|A^c_{i+1}|}} \sum_{A \supseteq A_i, A \not\supset A_{i+1}} b(A)$$

is the average of the list $\mathcal{L}_i$ (2), $\mathcal{L}_0 \doteq \{0\}$.

- global L_2 approximation not trivial, has not been addressed yet

L_∞ approximation

Theorem

*the **partial** L_∞ **consonant approximations of** b **in** $\mathcal{B}$ with maximal chain of focal elements $\mathcal{C} = \{A_1 \subset \cdots \subset A_n\}$ are the co.b.f.s whose mass vectors live in the polytope*

$$Cl\Bigg(\quad \big[b^1, \cdots, b^i - b^{i-1}, \cdots, 1 - b^{n-1}\big]' \,\bigg|\, b^i = \frac{b(A_i)+b(\{x_{i+1}\}^c)}{2} + +\{-b(A_1^c), b(A_1^c)\}\ \forall i \Bigg).$$

- its **barycenter** has mass $m_{co^{\mathcal{C}}_{L_\infty}[b]}(A_i) = \frac{b(A_i)-b(A_{i-1})}{2} + \frac{pl_b(x_i)-pl_b(x_{i+1})}{2}, \quad 2 \leq i \leq n-1$
- it is the **average of the maximal outer consonant approximation and the "contour-based" consonant approximation**

Interpretation as generalized maximal outer approximations

- all L_p approximations in $\mathcal{B}$ are functions of the list $\mathcal{L}_i$

$$
\begin{array}{rcl}
m_{co^{\mathcal{C}}_{max}[b]}(A_i) & = & \min(\mathcal{L}_i) - \min(\mathcal{L}_{i-1}); \;\; m_{co^{\mathcal{C}}_{con}[b]}(A_i) = \max(\mathcal{L}_i) - \max(\mathcal{L}_{i-1}); \\
m_{co^{\mathcal{C}}_{L_1}[b]}(A_i) & = & (int_1(\mathcal{L}_i) + int_2(\mathcal{L}_i))/2 - (int_1(\mathcal{L}_{i-1}) + int_2(\mathcal{L}_{i-1}))/2; \\
m_{co^{\mathcal{C}}_{L_2}[b]}(A_i) & = & ave(\mathcal{L}_i) - ave(\mathcal{L}_{i-1}); \\
m_{co^{\mathcal{C}}_{L_\infty}[b]}(A_i) & = & (\max(\mathcal{L}_i) + \min(\mathcal{L}_i))/2 - (\max(\mathcal{L}_{i-1}) + \min(\mathcal{L}_{i-1}))/2
\end{array}
$$

- for any two elements of the chain $A_i \subset A_{i+1}$ there is a number of "intermediate" focal elements A containing A_i but not A_{i+1}
- if 2^Θ were a totally ordered set, $\mathcal{L}_i$ would contain only $b(A_i)$ and all the L_p approximations would reduce to $co^{\mathcal{C}}_{max}[b]$
- they can all be seen as different *generalizations of the maximal outer consonant approximation*
- not, in general, outer approximations in the usual sense

Admissibility, negative masses

- geometric approximation in the belief space generates solutions which are in general only **partially** admissible
- some of them can have negative masses
- **sufficient conditions** on the desired maximal chain $\mathcal{C}$ under which they are indeed admissible can be given
- the contour-based approximation is admissible **if and only if the chain is generated by singletons sorted by their plausibility values**
- same for barycenter of L_∞ approximations
- similar conditions hold for L_1, L_2

Comparison with approximations in the mass space

- L_p consonant approximations in the *mass* space are associated with different but related *mass redistribution* processes
- i.e., the mass outside the desired chain of focal elements is re-assigned in some way to the elements of the chain
- their relationships with classical outer approximations and the isopignistic transform are weak
- have natural interpretation in terms of belief revision's general imaging
- some partial approximations (such as L_1 and L_2) are always entirely admissible and should be preferred

Approximations in a ternary example

- illustrate the different approximations in the toy case of a **ternary frame**, $\Theta = \{x, y, z\}$
- assuming we want the partial approximation with maximal chain $\mathcal{C} = \{\{x\} \subset \{x, y\} \subset \Theta\}$
- we illustrate the different partial consonant approximations in the simplex $Cl(\vec{m}_x, \vec{m}_{x,y}, \vec{m}_\Theta)$
- example belief function with masses

$$\begin{array}{lll} m_b(x) = 0.2, & m_b(y) = 0.3, & m_b(z) = 0, \\ m_b(x, y) = 0, & m_b(x, z) = 0.5, & m_b(y, z) = 0 \end{array}$$

Approximations in a ternary example

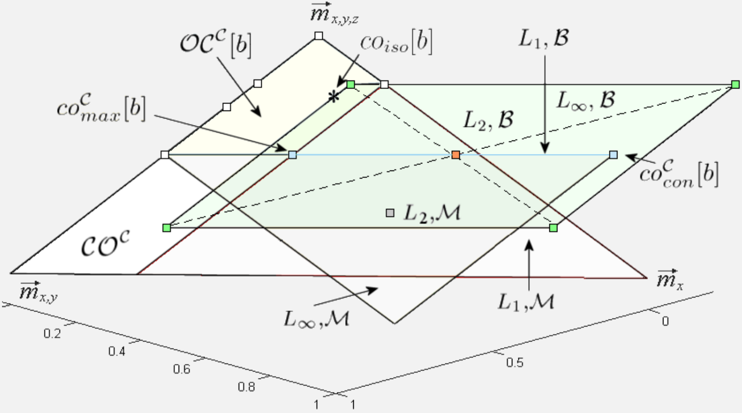


Figure: geometric L_p partial consonant approximations in the mass $\mathcal{M}$ and belief $\mathcal{B}$ spaces, isopignistic and outer consonant approximations for the example belief function

Conclusions

- consonant approximation is an alternative approach to limiting the computational cost of belief calculus
- consonant approximation can be induced by distance minimization
- we analyzed consonant transformations induced by L_p norms in the space of belief vectors
- distinguish between partial approximations on a specific chain, and global approximations
- two different semantics in the belief versus the mass space

Conclusions

- in the belief space, they can all be seen as different **generalizations of the maximal outer consonant approximation**
- partial solutions are completely admissible under sufficient conditions
- in the mass space, they amount to **some form of imaging**, and are completely admissible in most cases
- we adopted a purely ”geometric” approach to the problem, but **...**
- **...** huge field, dissimilarity measures coming from information-theoretic arguments