The geometry of relative plausibilities

Fabio Cuzzolin
Computer Science Department, University of California at Los Angeles
cuzzolin@cs.ucla.edu

Abstract

The study of the interplay between belief and probability has recently been posed in a geometric setup, in which belief functions are represented as points of a simplex called belief space. Each belief function $b$ is associated with three different geometric entities, which in turn determine three different Bayesian functions related to $b$. In this paper, in particular, we study the geometry of relative plausibilities, and show how equidistribution can be possibly a trait d’union among all the Bayesian relatives of a belief function.

Keywords: Belief space, Bayesian relatives, relative plausibility, equidistribution, orthogonality.

1 Introduction

In the last decades a number of different uncertainty measures have been proposed, as alternatives or extensions of the classical probability theory. The theory of evidence is one the most popular formalism, extending quite naturally probabilities on finite spaces through the notion of belief function. Naturally enough, the connection between belief and probability plays a major role in the theory of evidence, and is the basement of a popular approach to evidential reasoning, Smets’ pignistic model [12]. In fact, the problem of finding a correct probabilistic approximation of a belief function has been widely studied [10, 1], mainly in order to find efficient implementations of the rule of combination, as in Tessem [13], or Lowrance et al.. In cooperative game theory, Shapley’s value can be considered as an approximation of a set function by an additive set function [8]. Voorbraak [14], on his side, proposed to adopt the so-called relative plausibility function $\tilde{pl}_b$, i.e. the unique probability that, given a belief function $b$ with plausibility $pl_b$, assigns to each singleton its normalized plausibility. He proved that $\tilde{pl}_b$ is a perfect representative of $b$ when combined with other probabilities, $\tilde{pl}_b \oplus p = b \oplus p \forall p \in \mathcal{P}$. Cobb and Shenoy [4] described the properties of the relative plausibility of singletons and discussed its nature of probability function which is equivalent to the original belief function.

The study of the interplay between belief and probability has recently been posed in a geometric setup [9, 2], in which belief functions are represented as points of a simplex called belief space [7]. Each belief function is associated with three different geometric entities, namely the line $(b, pl_b)$, the orthogonal complement $\mathcal{P}^\perp$ of the probabilistic subspace $\mathcal{P}$, and the simplex of consistent probabilities $\mathcal{P}[b] = \{ p \in \mathcal{P} : p(A) \geq b(A) \forall A \subset \Theta \}$. These in turn determine three different probabilities associated with $b$, i.e. the orthogonal projection $\pi[b]$ of $b$ onto $\mathcal{P}$, the barycenter of $\mathcal{P}[b]$, and the pignistic function $BetP[b]$. We recently showed that if the belief function $b$ is 2-additive then pignistic function, orthogonal projection and intersection of $(b, pl_b)$ with $\mathcal{P}$ coincide [5].
However, the analysis of the simplest binary case suggests that the relative plausibility function does not fit in this picture. In this paper we will study the geometry of relative plausibilities in the general case, pointing out its richness and complexity. We will see that their geometry can be described by three planes in the belief space, in turn related to an additional set of three Bayesian functions. As 2-additivity is not able to encompass both the above Bayesian functions and Voorbraak’s relative plausibility function, we show that equidistribution can be a trait d’union for all the “Bayesian relatives” of a belief function.

2 Geometry of belief functions

In the theory of evidence [11] degrees of (subjective) belief are represented as belief functions. A basic probability assignment (b.p.a.) over a finite set $\Theta$ (frame of discernment) is a function $m : 2^\Theta \to [0,1]$ such that $m(\emptyset) = 0$, $\sum_{A \subset \Theta} m(A) = 1$, $m(A) \geq 0 \forall A \subset \Theta$. Subsets of $\Theta$ associated with non-zero values of $m$ are called focal elements (i.e.). The belief function (b.f.) $b : 2^\Theta \to [0,1]$ associated with a basic probability assignment $m$ is simply $b(A) = \sum_{B \subset A} m(B)$. The basic probability assignment $m$ of a belief function $b$ can be uniquely recovered by means of the Möbius inversion formula $m(A) = \sum_{B \subset A} (-1)^{|A-B|} b(B)$. In particular, a probability function is a peculiar belief function called Bayesian b.f. assigning mass to singletons only: $m(A) = 0, |A| > 1$.

A dual representation of the evidence encoded by a belief function $b$ is the plausibility function, whose value $pl_b(A)$ expresses the amount of evidence not against a proposition $A$: $pl_b(A) = 1 - b(A^c) = \sum_{B \subset A} m(B) \geq b(A)$. From $pl_b$ we can derive a Bayesian belief function called relative plausibility of singletons [14] as follows:

$$pl_b : \Theta \to [0,1], \quad x \mapsto \frac{pl_b(x)}{\sum_{y \in \Theta} pl_b(y)}.$$

(1)

Consider now a frame of discernment $\Theta$ and introduce in the Euclidean space $\mathbb{R}^N$, $N = |2^\Theta| - 1$, an orthonormal reference frame $\{x_A\}_{A \subset \Theta, A \neq \emptyset}$ in which each coordinate $x_A$ measures the belief value $b(A)$. The belief space associated with $\Theta$ is then the set of points $\mathcal{B}_\Theta$ of $\mathbb{R}^N$ which correspond to a belief function. We assume the domain $\Theta$ fixed, and denote the belief space with $\mathcal{B}$. It is not difficult to prove [7] that $\mathcal{B}$ is convex. More precisely, if we call $b_A \doteq b \in \mathcal{B}$ s.t. $m(A) = 1$, $m(B) = 0 \forall B \neq A$ the unique belief function assigning all the mass to a single subset $A$ of $\Theta$, the belief space $\mathcal{B}$ is the convex closure $Cl$ of all the basis belief functions $b_A$:

$$\mathcal{B} = Cl(b_A, A \subset \Theta, A \neq \emptyset).$$

(2)

Furthermore, each belief function $b \in \mathcal{B}$ can be written as a convex sum as follows:

$$b = \sum_{A \subset \Theta, A \neq \emptyset} m(A)b_A.$$  

(3)

Since a probability is a belief function assigning non zero masses to singletons only, the set $\mathcal{P}$ of all the Bayesian belief functions is the simplex determined by the three plane in the belief space, in turn related to a set of three Bayesian functions $\mathcal{F} = Cl(b_{\emptyset}, x \in \Theta)$. Analogously, we call plausibility space the region $\mathcal{P}_L$ of $\mathbb{R}^N$ whose points correspond to admissible plausibility functions. It can be proved [6] that this is also a simplex $\mathcal{P}_L = Cl(pl_A, A \subset \Theta)$, whose vertices are

$$pl_A = -\sum_{B \subset A} (-1)^{|B|} b_B.$$  

(4)

The $A$-th vertex $pl_A$ of the plausibility space is the plausibility vector associated with the basis belief function $b_A$, $pl_A = pl_{b_A}$. Each plausibility function $pl_b$ can be uniquely expressed as a convex combination of basis plausibility functions (4) as follows:

$$pl_b = \sum_{A \subset \Theta} m(A)pl_A.$$  

(5)

Belief functions are then nothing but special points in an Euclidean space. On the other side, given an ordering $A_1, ..., A_N$ for the subsets of $\Theta$, each point $x = [x_1, ..., x_N] \in \mathbb{R}^N$
(where \(x'\) denotes the transpose of a vector \(x\)), \(N = 2^n - 1\), can be thought of as a function \(\zeta : 2^\Theta \setminus \emptyset \rightarrow \mathbb{R}\) s.t. \(\zeta(A) = x_{ord(A)}\), where \(ord(A)\) is the position of \(A\) in the ordering. As the Moebius transformation is invertible, for each of these functions \(\zeta\) there always exists another function \(m_\zeta : 2^\Theta \setminus \emptyset \rightarrow \mathbb{R}\) such that \(\zeta(A) = \sum_{B \subseteq A} m_\zeta(B)\). Each vector \(\zeta\) of \(\mathbb{R}^N\) can be thought of as a sum function, even though \(m_\zeta\) does not in general satisfy the positivity constraint \(m_\zeta(A) \geq 0 \forall A \in \Theta\). The section \(x_{ord(\emptyset)} = 1\) of \(\mathbb{R}^N\) corresponds to the constraint \(\zeta(\emptyset) = 1\), so that points of this section are sum functions meeting the normality axiom, \(\sum_{A \in \emptyset} m_\zeta(A) = 1\). We call them normalized sum functions (n.s.f.). Analogously to the case of belief functions, we call Bayesian normalized sum function a n.s.f. \(\zeta\) such that \(\sum_{x \in \Theta} m_\zeta(\{x\}) = 1\). Note that this implies \(\sum_{|A| > 1} m_\zeta(A) = 0\), but not necessarily \(m_\zeta(A) = 0 \forall |A| > 1\).

3 A geometric interplay of belief and probability

As \(b(\Theta) = pl_b(\Theta) = 1\) for all b.f. \(b\), we can neglect the coordinate \(x_0\) and think of \(B\) as a region of \(\mathbb{R}^{N-1}\), in which \(b_\emptyset = \emptyset = [0, \ldots, 0]'\). Figure 1 shows the geometry of belief and plausibility for a binary frame \(\Theta_2 = \{x, y\}\), where belief and plausibility vectors are points of a plane with coordinates

\[
b = [b(x) = m(x), b(y) = m(y)]
\]

\[
pl_b = [pl_b(x) = 1 - m(y), pl_b(y) = 1 - m(x)]
\]

respectively. These two simplices are symmetric with respect to \(P\). Furthermore, each pair of functions \((b, pl_b)\) determines a line which is orthogonal to \(P\), on which they lay on symmetric positions on the two sides of the Bayesian region. The orthogonal projection of \(b\) onto \(P\), the pignistic function \(BetP[b]\) and the center of mass \(P_b\) of the probabilities \(P[b] = \{p \in P: p(A) \geq b(A) \forall A \subset \Theta\}\) consistent with \(b\) coincide \(\pi[b] = BetP[b] = P[b]\). \(b\) is then associated with three geometric entities: the line \((b, pl_b)\), the set of consistent probabilities \(P[b]\), and the orthogonal complement \(P^\perp\) of \(P\). In [5] we proved that the line \((b, pl_b)\) is always orthogonal to \(P\), even in the general case. However, \((b, pl_b)\) does not intersect the probabilistic subspace in general, but there exists a unique Bayesian normalized sum function \(\zeta[b]\) such that

\[
\zeta[b] = \beta[b] \cdot pl_b + (1 - \beta[b]) \cdot b
\]

where

\[
\beta[b] = \frac{\sum_{|A| > 1} m(A)}{\sum_{|A| > 1} m(A) |A|}
\]

is a scalar value between 0 and 1 which depends on \(b\). \(\zeta[b]\)'s values on singletons are

\[
m_\zeta(x) = m(x) + \beta[b] \cdot \sum_{A \supset x, A \neq x} m(A).
\]

Hence, \(\zeta[b]\) is naturally associated with a Bayesian belief function, assigning the same mass (8) to each singleton, namely

\[
\sigma[b] \doteq \sum_{x \in \emptyset} m_\zeta[b](x) \cdot b_{\{x\}}.
\]

It is well known that \(P[b]\) is convex, and that its center of mass \(P[b]\) is in fact the pignistic function [3]. The three geometric loci which arise from the analysis of the binary case are in conclusion each associated with a different Bayesian b.f. \((b, pl_b) \leftrightarrow \sigma[b], P^\perp \leftrightarrow \pi[b], P[b] \leftrightarrow BetP[b] = P[b]\). The trait d’union between those Bayesian functions is again suggested by the binary case [5].

**Proposition 1.** If the belief function \(b\) is 2-additive \((m(A) = 0, |A| > 2)\) then pignistic
function, orthogonal projection onto \( \mathcal{P} \) and intersection of the line \((b, pl_b)\) with \( \mathcal{P} \) coincide, \( \text{Bet}P|b| = \bar{P}|b| = \pi|b| = \zeta|b| = \sigma|b| \).

However, it is clear from Figure 1 that Voorbraak’s relative plausibility function does not fit in this picture, as 2-additivity is then not general enough to encompass both \( \tilde{pl}_b \) and the above Bayesian functions. In the following sections we will study its geometry: for sake of simplicity, we will drop the dependence on \( b \) from the notations for \( \sigma, \zeta, \beta \).

4 Geometry of relative plausibility

From the definition of relative plausibilities (1) it follows a straightforward geometric interpretation of \( \tilde{pl}_b \). Let us call plausibility of singletons the quantity

\[
\tilde{pl}_b = \sum_{x \in \Theta} pl_b(x) \cdot b_{\{x\}}. \tag{10}
\]

\( \tilde{pl}_b \) can be written as \( \tilde{pl}_b = \sum_x pl_b(x) \cdot b_{\{x\}} + (1 - k_1) \cdot b_{\Theta} \) with \( k_1 = \sum_x pl_b(x) \) since \( b_{\Theta} = 0 \), in order for it to meet the normalization constraint. As \( 1 - k_1 \leq 0 \) \( \tilde{pl}_b \) is a normalized sum function. Now, by definition (1) \( \tilde{pl}_b = \bar{pl}_b/k_1 \). Since \( b_{\Theta} \) is the origin of the reference frame, \( \tilde{pl}_b \) lies on the segment \((\tilde{pl}_b, b_{\Theta})\). This in turn implies that \( \tilde{pl}_b = (\tilde{pl}_b, b_{\Theta}) \cap \mathcal{P} \) i.e. \( \tilde{pl}_b \) is the intersection of the line joining the vacuous belief function \( b_{\Theta} \) and the plausibility of singletons \( pl_b \) with the probabilistic subspace.

From the definition of \( \sigma \) (9) it follows that \( \sigma = \sum_x m(x) b_{\{x\}} + \beta \sum_x (pl_b(x) - m(x)) b_{\{x\}} = (1 - \beta) \sum_x m(x) b_{\{x\}} + \beta \sum_x pl_b(x) b_{\{x\}} \) so that, after defining

\[
\bar{b} = \sum_{x \in \Theta} m(x) b_{\{x\}} = \sum_{x \in \Theta} m(x) b_{\{x\}} + (1 - k_2) b_{\Theta}, \tag{11}
\]

with \( k_2 = \sum_{|A| > 1} m(A) \), Eq. (10) implies

\[
\sigma = (1 - \beta) \cdot \bar{b} + \beta \cdot \tilde{pl}_b \tag{12}
\]

where \( \bar{b} \in C l(b_{\Theta}, \mathcal{P}) = \mathcal{D} \) and \( \tilde{pl}_b \in \text{span}(\mathcal{D}) \), the affine (vector) space generated by the simplex \( \mathcal{D} \). \( \mathcal{D} \) is the simplex of discounted probabilities [11], i.e. belief functions assigning mass to singletons or \( \Theta \) only. \( \bar{b} \) is in fact the unique belief function assigning to \( \Theta \) all the mass \((1 - k_2)\) which \( b \) gives to non-singletons. In the general case, then, the functions on the line \((b, pl_b)\) are associated with other quantities laying on the line \((\tilde{b}, \tilde{pl}_b)\) in the affine subspace \( \text{span}(\mathcal{D}) \) of discounted probabilities \( b \mapsto \tilde{b}, pl_b \mapsto \tilde{pl}_b, \zeta \mapsto \sigma \) which are in the same relative positions on the segment (just compare (6) and (12)). Again, 2-additivity plays a role, since [5]

**Theorem 1.** \( \zeta|b| = \sigma|b| \) iff \( b \) is 2-additive, in which case \( \sigma|b| \) is the orthogonal projection \( \pi|b| \) of \( b \) onto \( \mathcal{P} \).

4.1 A three plane geometry

In the binary case \( \mathcal{D} = \mathcal{B} \), \( \tilde{b} = b \) and \( \tilde{pl}_b = pl_b \), so that the plausibility of singletons is a plausibility function (Figure 1). In the general case, though, \( \mathcal{D} = C l(b_{\Theta}, \mathcal{P}) \) is in correspondence with a “dual” simplex \( \hat{\mathcal{D}} \subset \mathcal{P}_\mathcal{L} \), namely the set of plausibilities associated with a discounted probability. As \( pl_{\Theta} \) is the dual of \( b_{\Theta} \), and each probability \( b_{\{x\}} = pl_{\{x\}} \) is the dual of itself, \( \hat{\mathcal{D}} = C l(pl_{\Theta}, \mathcal{P}) \). The functions \( \tilde{b}, \tilde{pl}_b \in \mathcal{D} \) then correspond to two dual quantities in \( \hat{\mathcal{D}} \) (recalling Equation (5))

\[
\hat{b} = \sum_{x \in \Theta} m(x) pl_{\{x\}} + (1 - k_2) pl_{\Theta} = \tilde{b} + (1 - k_2) pl_{\Theta}
\]

\[
\hat{pl}_b = \sum_{x \in \Theta} pl_b(x) pl_{\{x\}} + (1 - k_1) pl_{\Theta} = \tilde{pl}_b + (1 - k_1) pl_{\Theta}. \tag{13}
\]

**Theorem 2.** The line passing through the duals of plausibility of singletons (10) and belief of singletons (11) crosses \( \sigma|b| \) too, and

\[
\beta(\tilde{pl}_b - \tilde{b}) + \tilde{b} = \sigma|b| = \beta(\tilde{pl}_b - \tilde{b}) + b. \tag{14}
\]

The geometry of the relative plausibility function can then be described in terms of the three planes \((\tilde{pl}_b, \sigma, pl_b), (b_{\Theta}, \tilde{pl}_b, pl_{\Theta})\) and \((\tilde{b}, \tilde{pl}_b, pl_{\Theta})\) (see Figure 2), where

\[
\tilde{b} = \bar{b}/k_2 \tag{15}
\]

is another probability which is natural to call relative belief of singletons. More precisely, we have just seen how \( \sigma \) is the intersection of both \((\tilde{b}, \tilde{pl}_b)\) and \((\tilde{b}, \tilde{pl}_b)\) and lays in the
same relative position on the two lines. They then determine a plane \( (\hat{b}, \sigma, \hat{b}) = (\overrightarrow{pl}_b, \sigma, \overrightarrow{pl}_b) \). Furthermore, by definition \( pl_b - b_\Theta = (\overrightarrow{pl}_b - b_\Theta)/k_1 \) while (13) implies \( pl_b = \overrightarrow{pl}_b/k_1 = (\overrightarrow{pl}_b - (1 - k_1)pl_\Theta)/k_1 \) so that \( \overrightarrow{pl}_b - pl_\Theta = (\overrightarrow{pl}_b - pl_\Theta)/k_1 \). The relative plausibility function then lies in the same relative position on the two lines \((b_\Theta, pl_b)\) and \((pl_\Theta, \hat{b})\), which intersect exactly in \( pl_b, b_\Theta, pl_\Theta, \hat{b} \) and \( b_\Theta, pl_\Theta \) then determine another unique plane which we can denote with \((b_\Theta, \hat{b}, pl_\Theta)\).

Analogously, by definition \( b - b_\Theta = (\hat{b} - b_\Theta)/k_2 \) while (13) yields \( b - pl_\Theta = (\hat{b} - pl_\Theta)/k_2 \), so that the relative belief of singletons lies in the same relative position on the two lines \((b_\Theta, \hat{b})\) and \((pl_\Theta, \hat{b})\), which intersect exactly in \( \hat{b}, b_\Theta, pl_\Theta, \hat{b} \) and \( \hat{b} \) then determine a single plane denoted by \((b_\Theta, \hat{b}, pl_\Theta)\).

![Figure 2: Planes and angles describing the geometry of the relative plausibility function.](image)

### 4.2 A geometry of three angles

In the binary case, \( b = \hat{b} = \overrightarrow{pl}_b, pl_b = \overrightarrow{pl}_b = \hat{b} \) and all these quantities are coplanar. This suggests a description of the geometry of \( pl_b \) in terms of the three angles

\[
\phi_1[b] = \overrightarrow{pl}_b, \phi_2[b] = \overrightarrow{bpl}_b, \phi_3[b] = \overrightarrow{bbpl}_b.
\]

In fact, even though the line \((b, pl_b)\) is always orthogonal to \( \mathcal{P} \), \((\hat{b}, \overrightarrow{pl}_b)\) is not in general orthogonal to the probabilistic subspace:

\[
\langle pl_b - b, b_{(y)} - b_{(x)} \rangle \propto \cos(\phi_1) \neq 0, \quad y \neq x
\]

where \( \langle \cdot \rangle \) denotes the scalar product, and \{\( b_{(y)} - b_{(x)} \forall y \neq x \} \) are the basis vectors of \( \mathcal{P} \). Nevertheless, \( \phi_1 \) has an interesting interpretation in terms of belief.

**Theorem 3.** \((\hat{b}, \overrightarrow{pl}_b) \perp \mathcal{P} \) \( (\phi_1[b] = \pi/2) \) iff \( \forall x \sum_{A \supset x, A \neq x} m(A) = pl_b(x) - m(x) = \text{const.} \)

In this case, recalling the definition of \( \sigma (9) \),

\[
\sigma(x) = m(x) + \beta(pl_b(x) - m(x)) = m(x) + \frac{1-k_2}{k_1-k_2} pl_b(x) - m(x) = m(x) + \frac{1-k_2}{k_1-k_2} pl_b(x) - m(x)
\]

and \( \sigma \) assigns the mass originally given by \( b \) to non-singletons to all singletons on equal basis. This can be expressed in terms of another probability function \( R[b] \)

\[
R[b] = \frac{pl_b - \hat{b}}{k_1 - k_2} = \sum_x \frac{pl_b(x) - m(x)}{k_1 - k_2} b_{(x)}.
\]

**Corollary 1.** The dual line \((\hat{b}, \overrightarrow{pl}_b)\) is orthogonal to \( \mathcal{P} \) iff the Bayesian b.f. \( R[b] \) is the uninformative probability \( \mathcal{P} = \sum_x b_{(x)}/n \).

In fact \( \sum_{x \in \Theta} \sum_{A \supset x, A \neq x} m(A) = \sum_{x \in \Theta} (pl_b(x) - m(x)) = k_1 - k_2 \) so that the condition of Theorem 3 can be written as \( pl_b(x) - m(x) = \sum_{A \supset x, A \neq x} m(A) = \frac{k_1-k_2}{n} \) for all \( x \). Replacing this in (18) yields \( R[b] = \sum_{x \in \Theta} \frac{1}{n} b_{(x)} \). The value of \( \phi_2 \) also depends on this mysterious probability (18).

**Lemma 1.** \( \phi_2 \) is nil if and only if

\[
(1, R[b])^2 = \| R[b] \| ^2 (1, 1).
\]

where \( pl_b = 1 \) is the \( N \)-dim vector \([1, \ldots, 1] \).

**Theorem 4.** The angle \( \phi_2[b] \) is zero if and only if \( R[b] \) is parallel to \( pl_\Theta = 1 \).

Condition (19) has the form \( \langle A, B \rangle^2 = \| A \|^2 \| B \|^2 \cos^2(\overrightarrow{AB}) = \| A \|^2 \| B \|^2 \) i.e. \( \cos^2(\overrightarrow{AB}) = 1 \), with \( A = pl_\Theta, B = R[b] \). This yields \( \cos(R[b]pl_\Theta) = 1 \) or \( \cos(R[b]pl_\Theta) = -1 \), but as both \( R[b] \) and \( pl_\Theta \) have all positive components it must be \( \cos(R[b]pl_\Theta) = 1 \) so that \( R[b]pl_\Theta = 0 \). In this case the two lines \((\hat{b}, \overrightarrow{pl}_b)\) and \((b_\Theta, pl_b)\) coincide. However, \( R[b]pl_\Theta \) means \( R[b] = \alpha \cdot pl_\Theta \) for some scalar \( \alpha \) i.e. \( R[b] = -\alpha \cdot \sum_{A \supset \Theta} (-1)^{|A|} b_A \) (since \( pl_\Theta = -\sum_{A \supset \Theta} (-1)^{|A|} b_A \) by Equation (4)). On the other side, \( R[b] = \sum_{x \in \Theta} R[b](x) b_{(x)} \) is a probability and the two conditions are
never met together, unless $|\Theta| = 2$. In conclusion

**Corollary 2.** $\phi_2[b] \neq 0$ and the lines $(\bar{b}, \tilde{p}l_k), (\bar{b}, \tilde{p}l_k)$ never coincide $\forall b \in B_\Theta$ when $|\Theta| > 2$; instead $\phi_2[b] = 0 \forall b \in B_\Theta$ when $|\Theta| \leq 2$.

![Figure 3: Location of $R[b]$ in $\mathcal{P}$.](image)

The binary case is then really anomalous, a situation signaled by $R[b]$ being parallel to the vacuous plausibility $pl_\emptyset$. $R[b]$ has in fact a simple geometric interpretation, as Equation (18) implies $R[b](k_1 - k_2) = pl_k - \bar{b} = \tilde{p}l_k \cdot k_1 - \bar{b} \cdot k_2 = pl_k \cdot k_1 - \bar{b} \cdot k_2 + k_1 \cdot \bar{b} - k_1 \cdot \bar{b} = k_1 \cdot (\tilde{p}l_k - \bar{b}) + b(k_1 - k_2)$ from which

$$R[b] = \bar{b} + \frac{k_1}{k_1 - k_2} (\tilde{p}l_k - \bar{b}).$$  \hspace{1cm} (20)

$R[b]$ lies on the line joining $\bar{b}$ and $\tilde{p}l_k$, outside the segment. Figure 3 illustrates how its location depends on the ratio between the total belief of singletons $k_2$ and the total plausibility of singletons $k_1$.

Finally, the angle $\phi_3[b]$ between $\bar{b}$ and $\tilde{p}l_k$ nullifies iff $\bar{b} = \tilde{p}l_k$, which is equivalent to $m(x)/k_2 = pl_k(x)/k_1 \forall x \in \Theta$. In other words, relative plausibility and belief coincide iff mass and plausibility are distributed in the same way among the singletons.

Again, a necessary and sufficient condition for $\phi_3[b] = 0$ can be expressed in terms of $R[b]$, as $R[b](x) = (pl_k(x) - m(x))/(k_1 - k_2) = \frac{1}{k_1 - k_2} (k_2 m(x) - m(x)) = m(x)/k_2 \forall x$, i.e. $R[b] = \bar{b}$, with $R[b]$ “squashing” $\tilde{p}l_k$ onto $\bar{b}$ from outside. In this case the quantities $\tilde{p}l_k, \tilde{p}l_k, \tilde{p}l_k, \sigma[b], \bar{b}, \bar{b}$ all lie in the same plane.

## 5 Equidistribution, orthogonality and Bayesianity

This geometric analysis brings to full light the existence of several Bayesian functions linked to the original belief function by means of the balance between the overall plausibility $k_1$ of the elements of the frame, and the total mass $k_2$ assigned to them. This flag of the “non-Bayesianity” of $b$ is symbolized by the Bayesian function $R[b]$. Those functions must be added to the three other probabilities discussed in Section 3. However, 2-additivity falls short of unifying this extended set of Bayesian relatives of $b$. The binary case again provides us with an intuition of how equidistribution can provide such a common ground. From Figure 1, we can appreciate how belief functions with $m(x) = m(y)$ lay on the bisection of the first quadrant which is orthogonal to $\mathcal{P}$. Their relative plausibility is also equal to their orthogonal projection $\pi[b]$.

In fact, Theorem 3 can be interpreted in terms of equal distribution of mass among focal elements. Consider the quantity

$$M_k(x) = \sum_{A \supseteq x, |A| = k} m(A)$$  \hspace{1cm} (21)

representing the contribution of size $k$ subsets to the plausibility of $x \in \Theta$. Suppose that $M_k(x) = const = M_k \forall x \in \Theta$ for all $k = 2, ..., n - 1$ (trivially true for $k = n$). Then $\sum_{A \supseteq x, A \neq x} m(A) = \sum_{k-2}^n \sum_{|A| = k, A \supseteq x} m(A) = m(\Theta) + \sum_{k=2}^{n-1} M_k$, which is constant $\forall x \in \Theta$. In other words,

**Corollary 3.** If $M_k(x) = const$ for all $x \in \Theta$, $\forall k = 2, ..., n - 1$ then the dual line $(\bar{b}, \tilde{p}l_k)$ is orthogonal to $\mathcal{P}$.

It is worth to notice that [5]

**Proposition 2.** If the belief function $b$ is such that for all $x \in \Theta, k = 3, ..., n - 1$ $M_k(x) = const = M_k$ then pignistic function $P[b]$ and orthogonal projection coincide, $P[b] = \pi[b]$.

The quantity $M_k(x)$ is then somehow connected to the geometric condition of orthogonality in the belief space framework. In [5] we have shown that a belief function $b$ is or-
orthogonal to the probabilistic subspace \( P \) iff
\[
\sum_{B \ni y, B \ni x} m(B)21^{1-|B|} = \sum_{B \ni x, B \ni y} m(B)21^{1-|B|}.
\]

(22)

**Theorem 5.** If a belief function \( b \) is such that \( M_k(x) = \text{const} = M_k \) for all \( k = 1, \ldots, n-1 \) then \( b \) is orthogonal to the probabilistic subspace, \( b \perp P \).

In fact, condition (22) is equivalent to
\[
\sum_{B \ni y} m(B)21^{1-|B|} = \sum_{B \ni x} m(B)21^{1-|B|} \equiv
\sum_{k=1}^{n-1} \frac{1}{2^k} \sum_{|B|=k, B \ni y} m(B) = \sum_{k=1}^{n-1} \frac{1}{2^k} \sum_{|B|=k, B \ni x} m(B) \equiv
\sum_{k=1}^{n-1} \frac{1}{2^k} M_k(y) = \sum_{k=1}^{n-1} \frac{1}{2^k} M_k(x) \text{ for all } y \neq x, \text{ and if } M_k(x) = M_k(y) \forall y \neq x \text{ the condition is met.}
\]

In this case, confirming the intuition given by the binary case (Figure 1), all the Bayesian relatives of \( b \) converge to the same probability:

**Theorem 6.** If \( M_k(x) = \text{const} = M_k \) for all \( k = 1, \ldots, n-1 \) then \( pl_b = \pi[b] = \tilde{P}[b] = \bar{P} \).

To see this let us write the values of the pignistic function \( \tilde{P}[b](x) \) in terms of \( M_k(x) \) as
\[
\sum_{A \ni x} \frac{m(A)}{|A|} = \sum_{k=1}^{n} \sum_{A \ni x, |A|=k} \frac{m(A)}{k} = \sum_{k=1}^{n} \frac{M_k(x)}{k} \text{ which is constant under the hypothesis, yielding } \tilde{P}[b] = \bar{P}.
\]

But now, as \( pl_b(x) = \sum_{A \ni x} \frac{m(A)}{k} = \sum_{k=1}^{n-1} \frac{M_k(x)}{k} \) we get \( \tilde{pl}_b(x) = \frac{pl_b(x)}{k_1} = \frac{1}{\sum_{k=1}^{n-1} M_k(x) \text{ which is equal to } 1/n \text{ if } M_k(x) = M_k \forall k, x.}
\]

6 Discussion

From the geometrical point of view, each belief function is associated with two different families of Bayesian functions (the Bayesian relatives of \( b \)). A first group of probabilities \( \sigma[b], \pi[b], \tilde{P}[b] = BetP[b] \) coincide under the assumption of 2-additivity. From the geometric study of the relative plausibility of singletons a second family of probabilities \( \bar{b}, pl_b, R[b] \) arises which does not meet this condition. The geometry of \( pl_b \) can be described in terms of three planes and angles, which depend on the balance between total belief and total plausibility of singletons.

**Equidistribution** can provide a common ground for the two families. If subsets of the same size contribute equally to the plausibility of each singleton \( (M_k(x) = \text{const}) \),
\[
k = 3, \ldots, n \quad \Rightarrow \quad \tilde{P}[b] = \pi[b] = \tilde{pl}_b = \bar{P}[b] = \bar{P} \]

Equidistribution is then inherently linked to the geometric notion of orthogonality, and provides a unifying viewpoint for all the geometric entities we encountered in the study of the Bayesian relatives of \( b \): the orthogonal complement of \( P \), the probability \( P[b] \) consistent with \( b \), and the dual lines \( (b, pl_b) \) and \( (\bar{b}, pl_b) \). The next natural step will be the study of the behavior of all these functions under the rule of combination.

**Appendix**

**Proof.** (Theorem 2) From Equation (13) we have that \( \bar{b} - b = (1-k_2) pl_{\Theta} \) and \( pl_b - pl_{\Theta} = (1-k_1) pl_{\Theta} \). Hence, \( \beta(\bar{pl}_b - b) = \beta[pl_b + (1-k_1) pl_{\Theta} - \bar{b} - (1-k_2) pl_{\Theta}] = + (1-k_2) pl_{\Theta} = \beta(\bar{pl}_b - b) + (k_2 - k_1) pl_{\Theta} + \bar{b} + (1-k_2) pl_{\Theta} = \bar{b} + \beta(\bar{pl}_b - b) + pl_{\Theta}[\beta(k_2 - k_1) + 1-k_2] \text{ but } \beta(k_2 - k_1) + 1-k_2 = \frac{1-k_2}{k_2-k_1} (k_2-k_1) + 1-k_2 = 0 \text{ by definition of } \beta[b](7), \text{ and (14) is met.} \]

**Proof.** (Theorem 3) (17) can be written as \( \langle \bar{pl}_b - \bar{b}, b(y) - b(x) \rangle = \sum_{z \in \Theta} (pl_{b(z)} - m(z)) \cdot \langle b(z), b(x) \rangle = \sum_{z \in \Theta} (pl_{b(z)} - m(z)) \cdot (b(z), b(x)) \cdot \sum_{z \in \Theta} (pl_{b(z)} - m(z)) = (b(z), b(x)) \cdot \sum_{z \in \Theta} (pl_{b(z)} - m(z)) \cdot \langle b(z), b(y) \rangle - \langle b(z), b(y) \rangle = \sum_{z \in \Theta} (pl_{b(z)} - m(z)) \cdot \langle b(z), b(y) \rangle \) since it is easy to see by the definition of \( b_{\Theta} \) that \( \langle b_{\Theta}, b_{\Theta} \rangle = \langle b_{\Theta}, b_{\Theta} \rangle \).

We distinguish three cases: if \( z \neq x, y \text{ then } \sum_{z \in \Theta} = |z \cup y| = 2 \text{ and the difference } |b_{\Theta} - b_{\Theta}||^2 \text{ goes to zero. If } z = x \text{ then } |b_{\Theta} - b_{\Theta}||^2 = |b_{\Theta} - b_{\Theta}||^2 = 2^{n-2} - 1 - (2^{n-1} - 1) = -2^{n-2}, \text{ while if } z = y \text{ then } |b_{\Theta} - b_{\Theta}||^2 = |b_{\Theta} - b_{\Theta}||^2 = 2^{n-2} - 1 - (2^{n-1} - 1) = -2^{n-2}. \text{ Hence (17) is equal to } 2^{n-2}. \langle pl_b (y) - m(y) \rangle - 2^{n-2} \langle pl_b (x) - m(x) \rangle \text{ and the thesis follows as } \sum_{A \ni x, A \neq x} m(A) = pl_b (x) - m(x). \]

**Proof.** (Lemma 1) After noticing that \( \beta[b] = \frac{1-k_2}{k_2-k_1} \) we can write \( pl_b - \bar{b} = \bar{pl}_b + \bar{pl}_b - (pl_b - \bar{b}) \)
\[(1 - k_1)p\Theta + (1 - \beta)(\hat{p}b - \hat{b}) = \frac{k_1 - 1}{k_1 - k_2} \cdot (pl_b - \bar{b}) - (k_1 - 1)p\Theta, \text{so that} \langle \hat{p}b - \sigma, pl_b - \sigma \rangle = \langle pl_b - \sigma, pl_b - \sigma \rangle + (1 - k_1)\langle p\Theta, pl_b - \sigma \rangle. \] But now \langle \hat{p}b - \sigma, pl_b - \sigma \rangle = (1 - \beta)^2\langle pl_b - \bar{b}, pl_b - \bar{b} \rangle$, being $pl_b - \sigma = (1 - \beta)(pl_b - \bar{b})$, while $\langle p\Theta, pl_b - \sigma \rangle = (1 - \beta)\langle p\Theta, pl_b - \bar{b} \rangle$ where $1 - \beta = \frac{k_1 - 1}{k_1 - k_2}$, so that by definition of $R[b]$ (18) we have $\langle pl_b - \sigma, pl_b - \sigma \rangle = (k_1 - 1)^2\langle R[b], R[b] \rangle$ and $(1 - k_1)\langle p\Theta, pl_b - \sigma \rangle = -(k_1 - 1)^2\langle 1, R[b] \rangle$. The desired scalar product is then $\langle pl_b - \sigma, pl_b - \sigma \rangle = (k_1 - 1)^2(\langle R[b], R[b] \rangle - \langle 1, R[b] \rangle)$. Now,
\[
\cos(\pi - \phi_2) = \frac{\langle pl_b - \sigma, pl_b - \sigma \rangle}{\| pl_b - \sigma \| \| pl_b - \sigma \|}
\]
where $\| pl_b - \sigma \| = \langle pl_b - \sigma, pl_b - \sigma \rangle^{1/2} = (k_1 - 1)(\langle R[b], pl_b - p\Theta \rangle)^{1/2}$ which is equal to $(k_1 - 1)(\langle R[b], R[b] \rangle + \langle p\Theta, pl_b \rangle - 2\langle R[b], pl_b \rangle)^{1/2}$ and $\| pl_b - \sigma \| = \| (1 - \beta)(pl_b - \hat{b}) \| = (k_1 - 1)\| R[b] \|$. So, as $\phi_2[b] = 0$ iff $\cos(\pi - \phi_2[b]) = -1$ we can write the condition as
\[-1 = \frac{(k_1 - 1)^2\| R[b] \|^2 - \langle 1, R \rangle - (k_1 - 1)^2\| R[b] \|\sqrt{\langle R, R \rangle + \langle 1, 1 \rangle - 2\langle R, 1 \rangle}}{\langle 1, 1 \rangle - 2\langle R, 1 \rangle}
\]
that is equivalent to (after elevating to the square both numerator and denominator) $\| R[b] \|^2(\| R[b] \|^2 + \langle 1, 1 \rangle - 2\langle R[b], 1 \rangle) = \| R[b] \|^4 + \langle 1, R[b] \rangle^2 - 2\langle 1, R[b] \rangle\| R[b] \|^2$ and by erasing the common terms we have as desired.

References


