Efficient $L1$-Based Probability Assessments Correction: Algorithms and Applications to Belief Merging and Revision

Marco Baioletti, Andrea Capotorti
Dipartimento di Matematica e Informatica
Università degli Studi di Perugia
{baioletti,capot}@dipmat.unipg.it

1. INTRODUCTION

A probability assessment on a finite domain is a quadruple $\pi = (V, U, p, \mathfrak{C})$, where

- $V = \{X_1, \ldots, X_n\}$ is a finite set of propositional variables
- U is a subset of V that contains the effective events taken into consideration,
- $p : U \to [0, 1]$ is a function which assigns a probability value to each variable in U
- $\mathfrak{C}$ is a finite set of logical constraints which lie among all the variables in V.

We use the De Finetti's notion of coherent probability assessment.

When a probability assessment $\pi = (V, U, p, \mathfrak{C})$ is not coherent, then there exist many different ways to find a coherent probability assessment π' which is as close as possible to π.

A possible solution is to revise only the probability values, i.e. $\pi' = (V, U, p', \mathfrak{C})$, and to use a distance defined only in terms of p and p'. Chosen a distance d on $\mathbb{R}^n$, a d-correction of a probability assessment $\pi = (V, U, p, \mathfrak{C})$ is a vector p' such that

- $\pi' = (V, U, p', \mathfrak{C})$ is coherent
- $d(p, p')$ is minimized

We denote $\mathcal{C}_d(\pi)$ the sets of all the d-correction of π.

In general, given a probability assessment π, $\mathcal{C}_d(\pi)$ could have more than one element and in this case the operation of correcting a probability assessment leads to an imprecise probability model, the so called "credal set".

In this work we focus on the L_1 distance $d_1(p, p') = \sum_{i=1}^{n} |p(X_i) - p'(X_i)|$ and we denote $\mathcal{C}_{d_1}(\pi)$ as $\mathcal{C}(\pi)$.
L_1 distance is interesting because the optimization problem can be converted into a MIP problem, i.e. a linear problem with integer and real variables. This is a great computational advantage with respect other distances that imply implementation of non linear (quadratic, logarithmic, etc.) optimizations tools.

The mixed integer program $\mathcal{P}1$ is built similarly to the method described in [Cozman].
Let $U = \{X_1, \ldots, X_n\}$ and $m = |\mathfrak{C}|$.
The real variables of $\mathcal{P}1$ are

- b_{ij}, for $i = 1, \ldots, n$ and $j = 1, \ldots, n+1$.
- q_j, for $j = 1, \ldots, n+1$
- r_i, s_i, for $i = 1, \ldots, n$

all of them are non-negative (as usual in linear programming).
The program $\mathcal{P}1$ has also the integer variables a_{ij}, for $i = 1, \ldots, n$ and $j = 1, \ldots, n+1$, constrained to be 0 or 1.
The constraints of $\mathcal{P}1$ are

- $\sum_{j=1}^{n+1} b_{ij} = p(X_i) + (r_i - s_i)$ for each $i = 1, \ldots, n$
- $0 \leq b_{ij} \leq a_{ij}, \quad a_{ij} - 1 + q_j \leq b_{ij} \leq q_j$ for $i = 1, \ldots, n$ and $j = 1, \ldots, n+1$,
- $\sum_{i=1}^{n+1} q_j = 1$
- $r_i \leq 1, \quad s_i \leq 1$ for $i = 1, \ldots, n$

Moreover, for each clause c_i (for $i = 1, \ldots, m$), where $c_i = \bigvee_{h \in H_i} X_h \vee \bigvee_{l \in L_i} \neg X_l$, and for each $j = 1, \ldots, n+1$ the linear constraint $\sum_{h \in H_i} a_{h,j} + \sum_{l \in L_i} (1 - a_{l,j}) \geq 1$ is added.
Finally, the objective function to be minized is $\sum_{i=1}^{n} (r_i + s_i)$

The actual implementations of MIP solvers make possible to solve probability correction problems of reasonable size in a feasible amount of time.

The optimal value δ for the objective function corresponds to the minimum possible correction on p and any coherent probability assessment $\pi' = (V, U, p', \mathfrak{C})$ such that $d_1(p, p') = \delta$ is a possible solution, i.e. p' is an element of $\mathcal{C}(\pi)$.

We describe the procedure **Correct** to find all the extremal elements $q_1, \ldots, q_s$ of $\mathcal{C}(\pi)$.

Let $\mathcal{Q}$ be the set of all vectors $q \in \mathbb{R}^n$ such that the probability assessments $(V, U, q, \mathfrak{C})$ are coherent. $\mathcal{Q}$ forms a convex polytope whose extremal points are exactly the atoms, i.e. all truth-value assignments α which satisfy the logical constraints $\mathfrak{C}$.

Let $\mathcal{B}_\pi(\delta)$ be the ball of all vectors $q \in \mathbb{R}^n$ such that $d(p, q) \leq \delta$, with p the numerical probability assessment present in π. $\mathcal{B}_\pi(\delta)$ is a convex set whose extremal points are the points $p \pm \delta e_i$, where e_i is the i-th vector of the canonical basis, for $i = 1, \ldots, n$.

Note that $\mathcal{C}_\pi = \mathcal{Q} \cap \mathcal{B}_\pi(\delta)$.

As a starting point, let us find a particular element $\bar{p} \in \mathcal{C}(\pi)$, which has the property that

$$\max_{i=1,\ldots,n} |\bar{p}(X_i) - p(X_i)|$$

is minimum, among all the coherent assessments such that $d_1(\bar{p}, p) = \delta$.
The vector $\bar{p}$ is the solution of a MIP problem $\mathcal{P}2$, which is similar to $\mathcal{P}1$, but with the following differences:

- the real variable z;
- the constraint $\sum_{i=1}^{n} (r_i + s_i) = \delta$;

- the constraints $r_i + s_i \leq z$, for $i = 1, \dots, n$;
- the objective function z, which has to be minimized.

$\bar{\pi} = (V, U, \bar{p}, \mathfrak{C})$ differs from π by δ and tries to spread this difference as much as possible among the variables of U: $\bar{p}$ is, in some sense, the most "central" point of $\mathcal{C}(\pi)$.

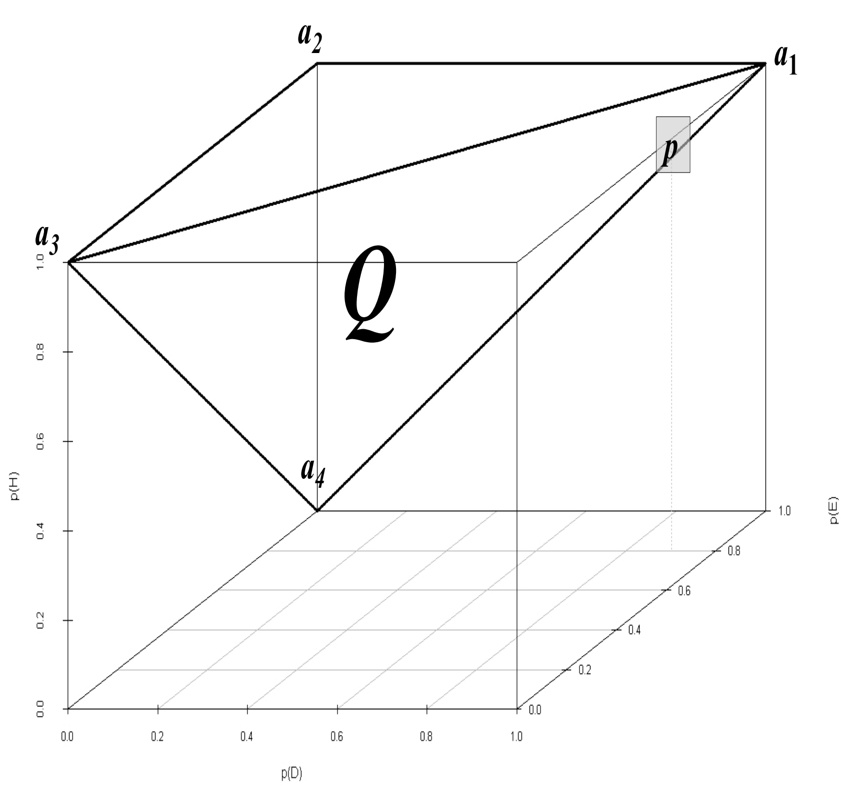
Using $\bar{p}$, it is possible to find the face F_1 of the polytope $\mathcal{Q}$ where $\mathcal{C}(\pi)$ lies: F_1 is a convex set with at most $n+1$ atoms as extremal points, which can be found as a part of the solutions of $\mathcal{P}2$ (i.e. the optimal values of a_{ij}).
By looking at the signs of $\bar{p}(X_i) - p(X_i)$, for $i = 1, \dots, n$, it is also possible to determine the face F_2 of $\mathcal{B}_\pi(\delta)$ which contains $\mathcal{C}(\pi)$: F_2 is a convex set with at most n extremal points of the form $p + sign(\bar{p}(X_j) - p(X_j)) \cdot \delta \cdot e_j$
Finally, the extremal points $X = \{q_1, \dots, q_s\}$ of $\mathcal{C}(\pi)$ can be easily found by means of the following procedure.

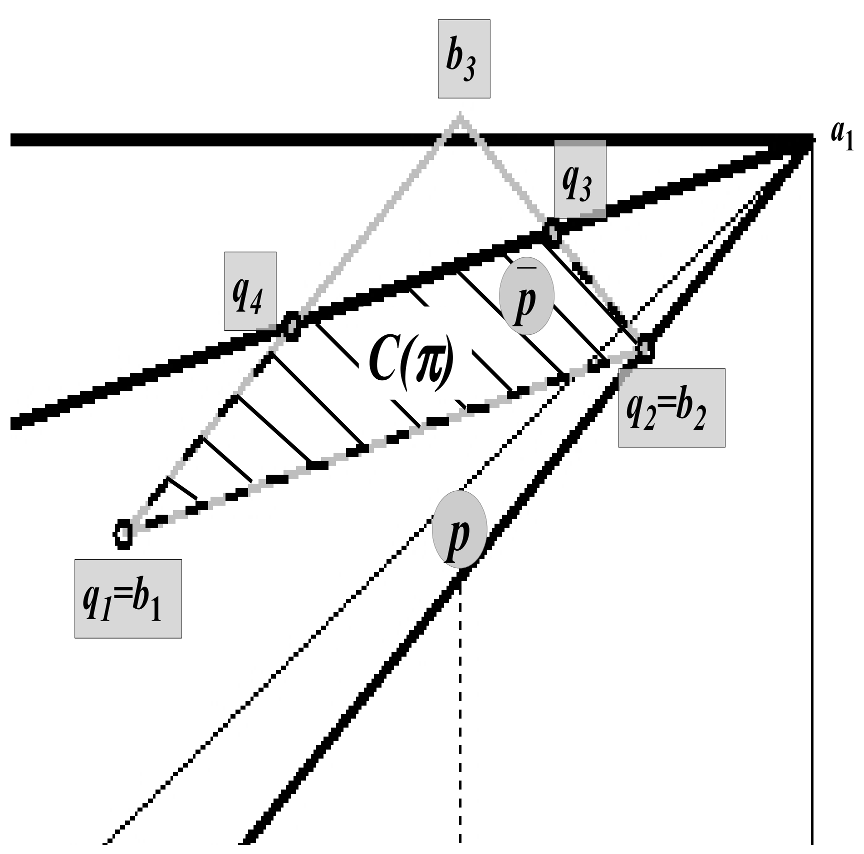
- let E_1 be the extremal points of F_1
- let E_2 be the extremal points of F_2
- compute H_1 as the H-representation of F_1
- compute H_2 as the H-representation of F_2
- let $H = H_1 \cup H_2$, the H-representation of $F_1 \cap F_2 = \mathcal{C}(\pi)$
- compute X as the V-representation of H

where the V-representation of a convex set C is the set of its extremal points, while the H-representation of C is a set H of half-spaces such that $C = \bigcap\limits_{h \in H} h$.
It is possible to convert from the V-representation of C to its H-representation by means of a face enumeration algorithm, while the inverse conversion is performed by a vertex enumeration algorithm. Both steps can be computed in polynomial time.

3. EXAMPLE

Suppose we have the following incoherent probability assessment $\pi = (V, U, \bar{p}, \mathfrak{C})$, where $V = U = \{E, H, D\}$, $p(D) = 0.9$, $p(E) = 0.8$ and $p(H) = 0.9$, $\mathfrak{C} = \{E \vee H, \neg D \vee E, \neg D \vee H\}$.
$\mathcal{C}(\pi)$ is the convex set whose extremal points are $q_1 = b1 = (0.7, 0.8, 0.9), q_2 = b2 = (0.9, 1, 0.9), q_3 = (0.9, 0.9, 1), q_4 = (0.8, 0.8, 1)$.





4. MERGING

The merging operation of two coherent probability assessments $\pi_1 = (V, U, p, \mathfrak{C})$ and $\pi_2 = (V, W, q, \mathfrak{D})$ produces a third probability assessment π_3 which is the coherent "fusion" of π_1 and π_2. We propose two approaches to merging: a "weighted combination" of the two assessments, or a "assignment to duplicates".

The first approach requires to create a non contradictory probability assessment from π_1 and π_2, by choosing a unique probability values for each variable in $U \cap W$ as a weighted average of p and q, i.e. chosen a weighting coefficient $\omega \in [0, 1]$, we define $\pi_1 +_\omega \pi_2$ as the probability assessment $(V, U \cup W, r, \mathfrak{C} \cup \mathfrak{D})$, where $r : U \cup W \to [0, 1]$ is now defined

$$r(x) = \begin{cases} p(x) & \text{if } x \in U \setminus W \\ q(x) & \text{if } x \in W \setminus U \\ \omega p(x) + (1-\omega)q(x) & \text{if } x \in U \cap W \end{cases}$$

Finally, the merging operation of π_1 and π_2 is

$$\pi_1 \oplus_\omega \pi_2 = \text{Correct}(\pi_1 +_\omega \pi_2)$$

A different approach is to create a probability assessment which maintains both numerical values and to solve the apparent contradiction by adding a new logical variable X_i', for each event $X_i \in U \cap W$ such that $p(X_i) \neq q(X_i)$, and assigning the values $r(X_i) = p(X_i)$ and $r(X_i') = q(X_i)$. Moreover, the logical constraint $X_i = X_i'$ is added to $\mathfrak{C} \cup \mathfrak{D}$.
Indeed, the assessment so obtained $\pi_1 + \pi_2$ is obviously incoherent and the merging operation of π_1 and π_2 is computed as

$$\pi_1 \oplus_I \pi_2 = \text{Correct}(\pi_1 + \pi_2).$$

The main difference between the two approaches is that the $\oplus_I$ tries to automatically solve the contradiction, while $\oplus_\omega$ needs an explicit way of solving it. The approach of $\oplus_\omega$ is in some sense a supervised one, because the user must explicitly provide a weight ω, while $\oplus_I$ adopts an unsupervised approach.

5. REVISION

Suppose that the coherent probability assessment $\pi_1 = (V, U, p, \mathfrak{C})$ represents our current belief state and a new reliable information arrives, represented by the probability assessment $\pi_2 = (V, W, q, \mathfrak{D})$.

We want to update π_1 with the new available information π_2, with the idea that

- we assume that the new information is correct
- we allow to revise, as less as possible, π_1 in order to adapt it to the new information

The revision operation can be performed as follows. First, π_1 and π_2 are merged together with the operator $+_0$, thus in the case of contradiction, the values from π_2 are used. Second, the resulting assessment is corrected by forbidding any change the probabilities of the variables in W.

The revision of π_1 with π_2 is then computed as

$$\pi_1 \star \pi_2 = Correct2(\pi_1 +_0 \pi_2, W)$$

where the procedure Correct2 is a small modification of the procedure Correct, in which it is possible to specify that the probability values of a given set of variables W are fixed