Compatibility, coherence and the running intersection property

Enrique Miranda Marco Zaffalon

University of Oviedo IDSIA

WPMSIIP '2018

Goals of the work

- To determine necessary and sufficient conditions for the compatibility of a number of marginal models with some joint.
- To extend the result based on the running intersection property to the conditional case.
- To get a more efficient manner to compute the natural extension.

Disjoint assessments

A simple scenario is when we have disjoint sets of variables: if we are given marginal probability measures P_1, P_2, P_3 on $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3$, then we can find a joint P on $\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3$ simply by applying independence: we make

$$P := P_1 \times P_2 \times P_3.$$

In the finite case, we just make the product of the mass functions.

What if they are not disjoint?

If we consider assessments P_{12} on $\mathcal{X}_1 \times \mathcal{X}_2$ and P_{23} on $\mathcal{X}_2 \times \mathcal{X}_3$, then a necessary condition for the existence of a compatible joint is that P_{12}, P_{23} induce the same marginal on $\mathcal{X}_2$:

$$P_{12}(A) = P_{23}(A) \ \forall A \subseteq \mathcal{X}_2.$$

In fact, we could always define in the finite case

$$P(x_1, x_2, x_3) = P_{12}(x_1, x_2) \cdot P_{23}(x_3|x_2),$$

where $P_{23}(x_3|x_2)$ is derived from P_{23} using Bayes' rule.

What if they are not disjoint?

If we consider assessments P_{12} on $\mathcal{X}_1 \times \mathcal{X}_2$ and P_{23} on $\mathcal{X}_2 \times \mathcal{X}_3$, then a necessary condition for the existence of a compatible joint is that P_{12}, P_{23} induce the same marginal on $\mathcal{X}_2$:

$$P_{12}(A) = P_{23}(A) \ \forall A \subseteq \mathcal{X}_2.$$

In fact, we could always define in the finite case

$$P(x_1, x_2, x_3) = P_{12}(x_1, x_2) \cdot P_{23}(x_3|x_2),$$

where $P_{23}(x_3|x_2)$ is derived from P_{23} using Bayes' rule.

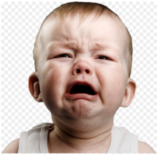
$\hookrightarrow$ So is this enough?

Pairwise compatibility $\not\Rightarrow$ global compatibility

Actually it is not: consider $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = \{0,1\}$ and the marginals P_{12}, P_{13}, P_{23} given by:

$$P_{12}(0,0) = P_{12}(1,1) = 0.5, \;\; P_{12}(0,1) = P_{12}(1,0) = 0$$
$$P_{13}(0,0) = P_{12}(1,1) = 0.5, \;\; P_{12}(0,1) = P_{12}(1,0) = 0$$
$$P_{23}(0,0) = P_{12}(1,1) = 0, \;\;\; P_{12}(0,1) = P_{12}(1,0) = 0.5$$

They are pairwise compatible (all of them have uniform marginals), but not globally compatible: P_{12} implies $X_1 = X_2$, P_{13} implies $X_1 = X_3$ and P_{23} implies $X_2 \neq X_3$, and these three things are incompatible!



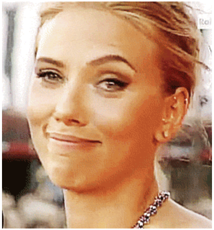
Running intersection property (Beeri et al.)

The key is in the running intersection property (RIP): we say that indices $S_1, \dots, S_r$ satisfy RIP when

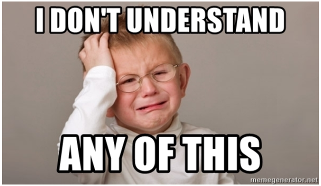
$$S_i \cap (\cup_{j<i} S_j) \subseteq S_{j^*} \text{ for some } j^* < i.$$

Then if we have marginals $P_{S_1}, \dots, P_{S_r}$ on $\mathcal{X}_{S_1}, \dots, \mathcal{X}_{S_r}$,

$$P_{S_1}, \dots, P_{S_r} \text{ globally compatible} \Leftrightarrow \begin{cases} P_{S_1}, \dots, P_{S_r} \text{ pairwise compatible} \\ S_1, \dots, S_r \text{ satisfy } RIP. \end{cases}$$



Why the hell??



The key here is that the RIP condition allows us to establish an order in the marginals we are given, and then we can apply the law of total probability by adding some assumptions of independence between sets of variables.

First generalisation

Now we are going to try to generalise the result in a number of ways:

- When the possibility spaces are infinite.
- When the marginals are imprecise.
- When we have conditional information.

NO FEAR

Sets of desirable gambles

A gamble on $\mathcal{X}$ is a bounded real-valued function $f : \mathcal{X} \to \mathbb{R}$. We denote the set of all gambles on $\mathcal{X}$ by $\mathcal{L}(\mathcal{X})$, and let $\mathcal{L}^+ := \{f \gneq 0\}$, the set of positive gambles.

Given $\mathcal{X}_1, \ldots, \mathcal{X}_n$ and $S \subseteq \{1, \ldots, n\}$, we let $\mathcal{X}_S := \times_{j \in S} \mathcal{X}_j$.

A gamble f on $\mathcal{X}^n$ is S-measurable if $f(x) = f(y)$ for every $x, y \in \mathcal{X}^n$ such that $\pi_S(x) = \pi_S(y)$, and we denote by $\mathcal{K}_S$ the set of $\mathcal{X}_S$-measurable gambles.

There exists a one-to-one correspondence between $\mathcal{L}(\mathcal{X}_S)$ and $\mathcal{K}_S$.

Coherent sets of gambles

$\mathcal{D} \subseteq \mathcal{L}(\mathcal{X})$ is coherent when $0 \notin \mathcal{D}$ and $\mathcal{D} = \text{posi}(\mathcal{D} \cup \mathcal{L}^+)$, where posi denotes the set of positive linear combinations.

In particular, we say that a set $\mathcal{D} \subseteq \mathcal{K}_S$ is coherent relative to $\mathcal{K}_S$ when the set $\mathcal{D}' \subseteq \mathcal{L}(\mathcal{X}_S)$ that we can make a one-to-one correspondence with, is coherent.

$\mathcal{D}$ avoids partial loss when it is included in some coherent set of gambles. The smallest such set is called its natural extension, and it is $\mathcal{E} = \text{posi}(\mathcal{L}^+ \cup \mathcal{D})$.

Compatibility for sets of desirable gambles

Consider subsets $S_1, \ldots, S_r$ of $\{1, \ldots, n\}$, and let $\mathcal{D}_j \subseteq \mathcal{L}(\mathcal{X}^n)$ be coherent with respect to $\mathcal{K}_{S_j} := \mathcal{K}_j$.

Given $i \neq j$ in $\{1, \ldots, r\}$, we say that $\mathcal{D}_i, \mathcal{D}_j$ are pairwise compatible if and only if

$$\mathcal{D}_i \cap \mathcal{K}_j = \mathcal{D}_j \cap \mathcal{K}_i.$$

- If $S_1, \ldots, S_r$ satisfy RIP and $\mathcal{D}_1, \ldots, \mathcal{D}_r$ are pairwise compatible, then there exists a coherent set of desirable gambles $\mathcal{D} \subseteq \mathcal{L}(\mathcal{X}^n)$ that is globally compatible with $\mathcal{D}_1, \ldots, \mathcal{D}_r$, in the sense that $\mathcal{D} \cap \mathcal{K}_j = \mathcal{D}_j \; \forall j = 1, \ldots, r$.

Coherent lower previsions

A lower prevision on is $\mathcal{L}(\mathcal{X})$ a functional $\underline{P} : \mathcal{L}(\mathcal{X}) \to \mathbb{R}$. $\underline{P}$ is called coherent when for any $f, g \in \mathcal{L}$ and any $\lambda > 0$:

(C1) $\underline{P}(f) \geq \inf_{x \in \mathcal{X}} f(x)$;

(C2) $\underline{P}(\lambda f) = \lambda \underline{P}(f)$;

(C3) $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$.

When $\mathcal{K} = \mathcal{L}(\mathcal{X})$ and (C3) holds with equality for every $f, g \in \mathcal{L}(\mathcal{X})$, $\underline{P}$ is called a linear prevision and is denoted by P.

A coherent set of desirable gambles $\mathcal{D}$ induces a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\mathcal{X})$ by means of the formula

$$\underline{P}(f) = \sup\{\mu : f - \mu \in \mathcal{D}\}.$$

Corollary: compatibility for coherent lower previsions

Consider subsets $S_1, \ldots, S_r$ of $\{1, \ldots, r\}$ satisfying RIP and for every j let $\underline{P}_j$ be a coherent lower prevision on $\mathcal{X}_{S_j}$.

- There exists a coherent lower prevision $\underline{P}$ on $\mathcal{X}^n$ such that $\underline{P}(f) = \underline{P}_j(f)\ \forall f \in \mathcal{K}_j, \forall j \iff \underline{P}_i(f) = \underline{P}_j(f)\ \forall f \in \mathcal{K}_i \cap \mathcal{K}_j$, and for every $i \neq j \in \{1, \ldots, r\}$.

...so in particular we obtain the classical result.

Conditional information

More generally, we may have unconditional and conditional information.

However, the meaning of compatibility is not as clear as in our previous results, in the sense that such a joint may necessarily induce additional assessments that are not in the original ones.

Taking this into account, given $\mathcal{D}_1, \ldots, \mathcal{D}_r$, we shall investigate to which extent these sets avoid partial loss, meaning that they have a joint coherent superset; but we are not requiring anymore that $\mathcal{D} \cap \mathcal{K}_j = \mathcal{D}_j$ for every $j = 1, \ldots, r$.

First simplification: remove isolated variables

- For every $i = 1, \ldots, r$, let $\mathcal{D}_i^*$ be the restriction of $\mathcal{D}_i$ to $\mathcal{K}_{S_i \cap (\cup_{j \neq i} S_j)}$.

$$\cup_{i=1}^r \mathcal{D}_i \text{ avoids partial loss } \iff \cup_{i=1}^r \mathcal{D}_i^* \text{ avoids partial loss.}$$

We may try to simplify further to pairwise compatibility:

$$\cup_{i=1}^r \mathcal{D}_i \text{ avoid partial loss } \Rightarrow \cup_{i \neq j} \mathcal{D}_i^j \text{ avoid partial loss,}$$

where $\mathcal{D}_i^j$ is the restriction of $\mathcal{D}_i$ to $\mathcal{K}_{S_i \cap S_j}$.

Just look at pairwise intersections?

....but it will not work:

$$\cup_{i\neq j}\mathcal{D}_i^j \text{ avoid partial loss} \not\Rightarrow \cup_{i=1}^r \mathcal{D}_i \text{ avoid partial loss.}$$

Second simplification: coherence graphs

It was proven by Miranda and Zaffalon (2009) that the verification of coherence can be simplfied by means of coherence graphs:

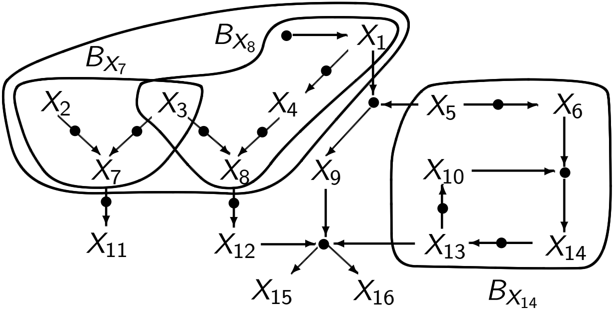


Figure: Example of a coherence graph.

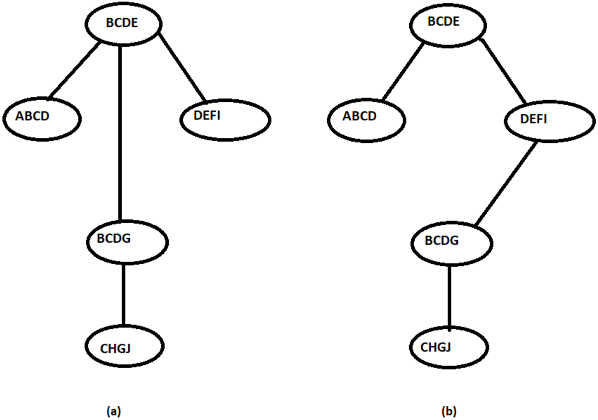
The assessments in different superblocks are automatically coherent, so can focus on each superblock separately.

Join trees

Assume we have conditional information on sets of variables $O_1|I_1, \dots, O_r|I_r$. We make a graphical representation of these templates so that we put the variables $O_j \cup I_j$ in one node, for $j = 1, \dots, r$, and connect two nodes when their associated sets of variables have non-empty intersection.

From this graphical representation, and after triangulation, it is always possible to make a tree of cliques called join tree, so that the sets of variables present in the different cliques satisfy RIP: for any two nodes V,W, all the nodes in the path between V and W contain $V \cap W$.

Example



(a) A join tree. (b) Not a join tree.

Our setting

We assume that:

- On each of the cliques of the join tree we have a coherent set of desirable gambles $\mathcal{D}_j$ on the corresponding set of variables.
- $\mathcal{D}_j$ is coherent relative to the set $\mathcal{K}_j$ of $\mathcal{X}_{S_j}$-measurable gambles.
- The sets are pairwise compatible.

Since the join tree satisfies RIP, the previous result guarantees that there is a compatible joint; the smallest one is the natural extension $\mathcal{E}$. We look for an efficient manner of computing it.

Iterative procedure

- We pick any node as a root. There is a partition of its set of nodes $\{1, \ldots, r\}$ into sets $A_0, A_1, \ldots, A_k$, $k < r$, where A_i includes those nodes that are at a distance i from the root. Thus, A_0 includes only the root.
- Step 1. We consider the nodes in A_k. For each of them, we take its associated set of desirable gambles.
- Step 2. We consider the nodes in A_{k-1}. For each node j of them, we have two possibilities:
  - If it has no adjacent nodes in A_k, we define $\mathcal{D}'_j$ as its set $\mathcal{D}_j$ of desirable gambles.
  - Otherwise, we take the set A of adjacent nodes, and define $\mathcal{D}'_j$ as the natural extension of $\mathcal{D}_j \cup \bigcup_{l \in A} \mathcal{D}'_{l|S_j \cap S_l}$.
- We proceed iteratively until we end up with a set of desirable gambles $\mathcal{D}'_0$ on the root node.

Main result, part 1

- $\mathcal{D}'_0$ is the restriction of the natural extension $\mathcal{E}$ of $\mathcal{D}_1, \ldots, \mathcal{D}_r$ to $\mathcal{K}_{S_0}$.
- $\mathcal{D}_1, \ldots, \mathcal{D}_r$ avoid partial loss if and only if $\mathcal{D}'_0$ is coherent.

Main result, part 1

- $\mathcal{D}'_0$ is the restriction of the natural extension $\mathcal{E}$ of $\mathcal{D}_1, \ldots, \mathcal{D}_r$ to $\mathcal{K}_{S_0}$.
- $\mathcal{D}_1, \ldots, \mathcal{D}_r$ avoid partial loss if and only if $\mathcal{D}'_0$ is coherent.

Ok, but do we need to repeat this for each node so as to get the natural extension everywhere?

Reverse procedure

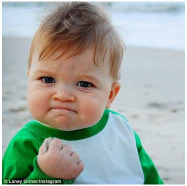
With the same root node as before and the sets $\mathcal{D}'_0, \ldots, \mathcal{D}'_{r-1}$ we generated above, we define iteratively $\mathcal{D}''_0, \ldots, \mathcal{D}''_{r-1}$ as follows:

- We make $\mathcal{D}''_0 := \mathcal{D}'_0$.
- Step 1: if a node i belongs to A_1, we define $\mathcal{D}''_i := \text{posi}(\mathcal{D}'_i \cup \mathcal{D}'_{0|S_i \cap S_0} \cup \mathcal{L}^+(\mathcal{X}_{S_i}))$.
- Step 2: for any $i \in A_2$, we let B_i denote its neighbours in A_1, and let $\mathcal{D}''_i := \text{posi}(\mathcal{D}'_i \cup \bigcup_{j \in B_i} \mathcal{D}''_{j|S_j \cap S_i} \cup \mathcal{L}^+(\mathcal{X}_{S_i}))$.

We proceed iteratively until we get to the nodes in A_k.

Main result, part 2

Let $\mathcal{E}$ be the natural extension of $\mathcal{D}_1, \ldots, \mathcal{D}_r$. If we follow the procedure above, then $\mathcal{D}_i'' = \mathcal{E} \cap \mathcal{K}_i \forall i = 1, \ldots, r$.



Conclusions

- In the unconditional case, RIP is a device that allows to apply the law of total probability.
- Because of that, we can use its extension to the imprecise case: the marginal extension theorem.
- In the conditional case, we can simplify the verification of coherence using join trees and coherence graphs.

Open problems:

- Infinite spaces: conglomerability?
- Clarify the process inside the cliques.
- Results in terms of conditional lower previsions?