Deliberation, Aggregation, Consensus

Paris, 21–22 October

All agreed
Aumann meets Wagner

⋆

Jan-Willem Romeijn / Olivier Roy

University of Groningen / Bayreuth

Consensus formation

De Groot, Lehrer, Wagner, and others formally describe consensus formation as repeated opinion pooling. We limit attention to the two agents Raquel and Quassim, R and Q for short. On every round $i > 0$, Raquel has

$$P_R^{i+1}(S) = w_R P_Q^i(S) + (1 - w_R) P_R^i(S).$$

and similarly for Quassim. Trusts parameters w_R and w_Q are assumed to be fixed. The result is a sequence of opinion pairs:

$$\langle r_1, q_1 \rangle, \langle r_2, q_2 \rangle, \ldots, \langle r_i, q_i \rangle, \ldots, \langle p, p \rangle$$

Aumann's agreement result

At face value Aumann's result that we cannot agree to disagree is similar. Indeed Aumann writes:

> It seems to me that the Harsanyi doctrine is implicit in much of [the literature on opinion pooling]; reconciling subjective probabilities makes sense if it is a question of implicitly exchanging information, but not if we are talking about "innate" differences in priors. The result of this paper might be considered a theoretical foundation for the reconciliation of subjective probabilities.

Despite this there is, as far as we know, no account of how the Harsanyi doctrine underpins opinion pooling.

Agreement and consensus

This paper provides a reconstruction of an approach to consensus in terms of Aumann's result. More precisely:

- The consensus formation is related to the dynamic approach to common knowledge and agreement.
- The sequence of opinions is used to constrain a probability assignment over the event space of the agents.
- The trust parameters w_R and w_Q can thereby be identified with aspects of the likelihoods for the opinions.

Contents

1 Aumann's agreement result

The setting of Aumann's [1976] theorem is one in which two agents, Raquel and Quassim, are in the following epistemic situation:

- They share a state space Ω and an initial probability assignment P over it.
- Both have their own information partition $\mathcal{R}$ and $\mathcal{Q}$, with elements R_i and Q_i. But they know the partition of the other.
- Each of them gathers private information, in the form of one element from their partition, R_0 and Q_0.

Type space

The agreement theorem is formalized in the context of Harsanyi type spaces. The type space of Raquel consists of:

- A set of states of the world Ω.
- A set of types or epistemic states $t_R \in \mathcal{T}_R$.
- A function that associates types with opinions, $\lambda_R : \mathcal{T}_R \mapsto \mathcal{M}(\Omega \times \mathcal{T}_Q)$.

Since $\mathcal{T}_Q$ is again mapped onto types in $\mathcal{T}_R$, a type is a recursively defined object of infinite depth.

An eventual set of types

Aumann's theorem states that if the posteriors are common knowledge, then they are equal. Assuming common knowledge, Raquel is localized in some final set of types $\mathcal{T}_{R^\star}$, perhaps a singleton. For every $t_R \in \mathcal{T}_{R^\star}$ we have

$$[\lambda_R(t_R)](S) = P(S|[t_R]) = x,$$

where $[t_R]$ denotes the full type space restricted to the type t_R, so $[t_R] \in \Omega \times \mathcal{T}_Q$.

Common knowledge

Common knowledge of the posterior r means that there is no variation of r over the types t_R: if there were, then some higher-order doubt regarding $P_R(S)$ would remain. Hence the marginal probability of S is also r:

$$\sum_{t_R \in \mathcal{T}_{R^\star}} P(S|[t_R])P([t_R]) = r.$$

But this also holds for Quassim, and so $r = q$. It does not matter in what direction we marginalize the probability for S.

2 Dynamic agreement and consensus

Geanakoplos and Polemarchakis [1982] provide an account of how this situation of common knowledge may arise from an exchange of opinions between Raquel and Quassim.

- Both have received private information, R_0 and Q_0, from their own information partition, and an associated set of types.
- At each round i they exchange a note with their posteriors r_i and q_i respectively.
- With this new information they exclude elements from the information partition, and hence types, of the other agent.
- And they update their own set of types accordingly.

Eliminating types

Every round $i > 0$ Raquel eliminates the set $[q_i]$. This set is a union of types of both Raquel and Quassim. It comprises the types associated with those elements from Quassim's information partition that are inconsistent with his pronouncement of q_i. Those sets are:

$$\begin{aligned}[q_i] &= \{t_Q : [\lambda_Q(t_Q)](S) = q_i\} \cup \\ &\quad \{t_R : \exists t_Q \left([\lambda(t_R)]([t_Q]) > 0 \text{ and } [\lambda_Q(t_Q)](S) = q_i\right)\}.\end{aligned}$$

If the revealed probability of Quassim q_i allows Raquel to eliminate some elements from $\mathcal{Q}$, and hence some types of Quassim and herself, within which $P_R(S)$, then this may lead to a new probability value for Raquel. A similar story holds for Quassim.

Fixing the prior

We can now see what kind of prior over the sets R_i and Q_i will need to be in place for the consensus formation process to be manifested in the dynamic approach to agreement. We eliminate types with non-zero probability in every round $i > 0$:

$$P_R^i(S) = r_i = P(S | \cup_{j=0}^{i-1} [q_j] \cap R_0)$$

$$P_Q^i(S) = q_i = P(S | \cup_{j=0}^{i-1} [r_j] \cap Q_0)$$

We set out to establish that this set of constraints can be imposed coherently onto the type space.

3 Pooling as conditioning

Bayesian updating and pooling seem to be different, even opposite operations.

- In conditionalization the prior opinion state is a mixture of posterior opinion states, $P(S) = \sum_j P([x_i])P_{[x_i]}(S)$.
- In pooling the posterior state is a mixture of prior opinions, $P_{[x_i]}(S) = \sum_X w_X P_X(S)$.

How to interpret the weights w_X in terms of Bayesian conditionalization?

Encoding the pooling process in a prior

Note first that the constraints on the prior can easily be met. For Raquel and Quassim we can coherently impose that for all $i > 0$,

$$P(S|\cup_{j=0}^{i-1}[q_j] \cap R_0) = r_i.$$

$$P(S|\cup_{j=0}^{i-1}[r_j] \cap Q_0) = q_i.$$

For this we may employ the freedom in the prior over the sets $[r_i]$ and $[q_i]$: by choosing it small enough, we can pitch the posteriors for Raquel and Quassim at the desired level.

Constructing the likelihoods

We can attempt to construct a set of equations that fully characterizes the likelihoods for the events $[r_i]$ and $[q_i]$. For Raquel's likelihoods the desiderata are:

- At $q_i = r_i$ we must have that $P([q_i]|S \cap [r_i]) = P([q_i]| \neq S \cap [r_i])$.
- The factor used in Bayesian updating must be w_R at the actual value of q_i:
$$\frac{P([q_i]|S \cap [r_i])}{P([q_i]|[r_i])} = w_R.$$
It is preferable if the whole likelihood function is in fact linear.
- Preferably, the density function $P([q_i]|[r_i])$ integrates to 1.

Assumption of linearity

The general problem remains. For simplicity we may assume linear likelihood functions:

$$P([q_i]|S \cap [r_i]) = m_+ + (t_+ - m_+)q_i,$$

$$P([q_i]|\neg S \cap [r_i]) = m_- + (t_- - m_-)q_i,$$

and similar for Quassim. We can then uniquely determine the functions and equate the slope of $P([q_i]|S \cap [r_i])$ with the weight w_R:

$$\frac{t_+ - m_+}{t_- - m_-} = \frac{1-r}{r}, \qquad \frac{t_+ - m_-}{t_- - m_+} = \frac{1-r}{r}, \qquad \Delta_+ = t_+ - m_+ = w_R$$

Most of this fits well with an intuitive understanding of the likelihoods.

Some smallprint

Further remarks on this match between updating and pooling:

- We must assume that the probability P is regular: updating cannot escape extremal values.
- The sets $[r_i]$ and $[q_i]$ have zero measure so conditioning on them needs a little attention.
- At extremal values of x there are interesting connections to Condorcet's theorem.
- Genest and Schervish already provide a link between conditioning and pooling but do not connect this to Aumann.

4 Concluding remarks

We think that these results constitute a bridge between two research traditions.

- The dynamic version of Aumann's result leads to a partition of type space over which we can make the constraints imposed by a pooling process precise.
- Those constraints can be accommodated by the common prior P.
- Based on a few simplifying assumptions we can construct the likelihood functions in detail. The trust parameter in consensus formation thereby gets a Bayesian interpretation.

Further research

It seems natural to try and transport over the bridge just devised. Many challenges remain:

- Providing a general account of the likelihood functions.
- Analysing a lack of consensus in terms of differences in the priors of agents.
- Looking for a natural taxonomy of consensus formation processes in terms of characteristics of the priors.

Thank you

The slides for this talk will be available at http://www.philos.rug.nl/ romeyn.

For comments and questions, email j.w.romeijn@rug.nl.