Dynamic Imprecise Probability Kinematics*

Michele Caprio MICHELE.CAPRIO@DUKE.EDU
Statistical Science, Duke University, U.S.A.

Ruobin Gong RUOBIN.GONG@RUTGERS.EDU
Statistics, Rutgers University, U.S.A.

In this work, we first describe a probability kinematics update of a probability measure P on a state space Ω. We then generalize this update to apply to a set of probabilities $\mathscr{P}$ on Ω. The sequence of successive generalized updates can be completely characterized by a sequence of upper and lower probability measures. We describe an exhaustive procedure to update these latter, and we provide bounds to avoid the bottleneck in the exhaustive procedure. We conclude by studying contraction, dilation, and sure loss.

The updating procedure for P is the following. We first collect data $x_1,\dots,x_n$. These data points are images of random variables, that is, there exist $E_1,\dots,E_n \subset \Omega$ such that for all $\omega \in E_j$, $X_j(\omega)=x_j$, for all $j \in \{1,\dots,n\}$. Notice that E_j can be written as $E_j := \{\omega \in \Omega : X_j(\omega)=x_j\}$. $E_1,\dots,E_n$ induce a partition, $\mathscr{E}=\left\{E_1,\dots,E_n,E_{n+1} := \left(\bigcup_{j=1}^n E_j\right)^c\right\}$. We make three main assumptions: the range $\mathscr{X}$ of $X_1,\dots,X_n$ is not uncountable, $X_1,\dots,X_n$ are identically distributed (this avoids problems with the induced partition $\mathscr{E}$), and the data points we collect are different from one another, that is, $x_1 \neq x_2 \neq \cdots \neq x_n$ (this corresponds to only keeping new data points and discard those that we have already seen, since the information they convey is already encoded in some $E_j \in \mathscr{E}$). We then update P to $P_{\mathscr{E}}$ defined as $P_{\mathscr{E}}(A) := \sum_{E_j \in \mathscr{E}} P(A \mid E_j) P_{\mathscr{E}}(E_j)$, for all $A \subset \Omega$. This is a well defined probability measure, and it is a Jeffrey's posterior – that is, a probability kinematics update – for P. We propose a mechanical way of computing $P_{\mathscr{E}}(A)$, for any A. For all $j \in \{1,\dots,n\}$, we let $P_{\mathscr{E}}(E_j)=P_{X_j}(x_j) \equiv P_X(x_j)$, where the last equivalence comes form having assumed the X_j's to be identically distributed. Also, we let $P_{\mathscr{E}}(E_{n+1}) = 1-\sum_{j=1}^n P_{\mathscr{E}}(E_j)$. In this way, updating from P to $P_{\mathscr{E}}$ becomes a mechanical procedure that does not require subjective assessment of the probabilities of the E_j's. Once we obtain new data points, we update $P_{\mathscr{E}} \equiv P_{\mathscr{E}_1}$ to $P_{\mathscr{E}_2}$ in a similar fashion: $P_{\mathscr{E}_1}$ is our new prior, and $x_1,\dots,x_n,x_{n+1},\dots,x_m$ are the data points that drive the update of $P_{\mathscr{E}_1}$ to $P_{\mathscr{E}_2}$. The sequence $(P_{\mathscr{E}_n})$ converges superlinearly in the TV metric to $P_{\tilde{\mathscr{E}}}$, the probability measure associated with $\tilde{\mathscr{E}}$. This latter is the partition having $\emptyset$ as an element, that is, there exists an $E_j \in \tilde{\mathscr{E}}$ such that $E_j=\emptyset$.

After noting that our suggested procedure is prior sensitive, we present a robust way of performing it. We require that the agent specifies a finite set $\mathscr{P}$ of probability measures on Ω. Their initial opinion, then, is represented by the convex hull $\mathscr{P}^{\text{co}}$ of the elements of $\mathscr{P}$. This is required because upper and lower probabilities associated with $\mathscr{P}$ completely characterize $\mathscr{P}^{\text{co}}$, but not $\mathscr{P}$. To update $\mathscr{P}^{\text{co}}$ to $\mathscr{P}^{\text{co}}_{\mathscr{E}_1}$, we first update all the elements in $\mathscr{P}$ using the procedure we explained earlier, so obtaining $\mathscr{P}_{\mathscr{E}_1}$, and then we compute its convex hull. Repeating this procedure, we build a sequence $(\mathscr{P}^{\text{co}}_{\mathscr{E}_n})$ that converges in the Hausdorff metric to $\mathscr{P}^{\text{co}}_{\tilde{\mathscr{E}}}$, the convex hull of the set of limit probability measures $P_{\tilde{\mathscr{E}}}$. Lower and upper probability measures $\underline{P}_{\mathscr{E}_n}$ and $\overline{P}_{\mathscr{E}_n}$ encode all the information contained in $\mathscr{P}^{\text{co}}_{\mathscr{E}_n}$, for all n. The exhaustive procedure to update $\underline{P}_{\mathscr{E}_n}$ and $\overline{P}_{\mathscr{E}_n}$ is given by the following: (1) We collect data $\{x_j\}$; (2) We update all the elements in $\mathscr{P}_{\mathscr{E}_n}$ according to the mechanical procedure, so to obtain $\mathscr{P}_{\mathscr{E}_{n+1}}$; (3) We compute $\mathscr{P}^{\text{co}}_{\mathscr{E}_{n+1}}$; (4) We calculate $\underline{P}_{\mathscr{E}_{n+1}}(A) = \inf_{P \in \mathscr{P}^{\text{co}}_{\mathscr{E}_{n+1}}} P(A)$, and $\overline{P}_{\mathscr{E}_{n+1}}(A) = \sup_{P' \in \mathscr{P}^{\text{co}}_{\mathscr{E}_{n+1}}} P'(A)$, for all $A \subset \Omega$. $\underline{P}_{\mathscr{E}_{n+1}}$ and $\overline{P}_{\mathscr{E}_{n+1}}$ are the updates to $\underline{P}_{\mathscr{E}_n}$ and $\overline{P}_{\mathscr{E}_n}$, respectively. In practice, step (2) appears to be the bottleneck of the algorithm. To overcome this shortcoming, we give a lower bound for the updated lower probability and an upper bound for the updated upper probability, that can be calculated without updating all of the elements in $\mathscr{P}_{\mathscr{E}_n}$. These bounds can also be used to build an interval around $P_{\mathscr{E}_{n+1}}(A)$, for all A and all $P_{\mathscr{E}_{n+1}} \in \mathscr{P}_{\mathscr{E}_{n+1}}$. Sharper bounds are possible at the price of an extra assumption. We also give conditions for $\mathscr{P}^{\text{co}}_{\mathscr{E}_n}$ to contract, dilate, and exhibit sure loss. $\mathscr{P}^{\text{co}}_{\mathscr{E}_n}$ *contracts with respect to* $\mathscr{P}^{co}_{\mathscr{E}_{n-1}}$ *for some* $A \subset \Omega$ if $\underline{P}_{\mathscr{E}_n}(A) \geq \underline{P}_{\mathscr{E}_{n-1}}(A)$ and $\overline{P}_{\mathscr{E}_n}(A) \leq \overline{P}_{\mathscr{E}_{n-1}}(A)$. *Dilation* is defined analogously, by inverting the inequality signs. Finally, we say that $\mathscr{P}^{\text{co}}_{\mathscr{E}_n}$ exhibits *sure loss with respect to* $\mathscr{P}^{co}_{\mathscr{E}_{n-1}}$ *for some* $A \subset \Omega$ if $\underline{P}_{\mathscr{E}_n}(A) > \overline{P}_{\mathscr{E}_{n-1}}(A)$ or $\overline{P}_{\mathscr{E}_n}(A) < \underline{P}_{\mathscr{E}_{n-1}}(A)$.

We conclude by giving an example in which our method can be useful in a real-world situation, namely in a survey sampling study.

* The authors thank Sayan Murkherjee for inspiring this project and helpful discussions, and an anonymous referee for insightful comments. Research of R. Gong is supported in part by the National Science Foundation (DMS-1916002). M. Caprio would like to acknowledge funding from NSF CCF-1934964.