Proceedings of Machine Learning Research 103:324–326, 2019 ISIPTA 2019

Dilation and Asymmetric Relevance

Arthur Paul Pedersen APP@ARTHURPAULPEDERSEN.ORG

Gregory Wheeler G.WHEELER@FS.DE

Centre for Human and Machine Intelligence, Frankfurt School of Finance & Management, Germany

Abstract

A characterization result of dilation in terms of positive and negative association admits an extremal counterexample, which we present together with a minor repair of the result. Dilation may be asymmetric whereas covariation itself is symmetric. Dilation is still characterized in terms of positive and negative covariation, however, once the event to be dilated has been specified.

Keywords: dilation, sets of probabilities

1. Introduction and Preliminaries

A characterization result specifying necessary and sufficient conditions for dilating sets of probabilities expressed in terms of witnesses to positive and negative association in lower and upper conditional probability neighborhoods was given in [5] and generalized in [6]. This result admits an extremal counterexample, presented in Section 2. A minor modification to the conditions is shown to re-establish the characterization result in Section 3.

A *lower probability space* is a quadruple $(\Omega, \mathscr{A}, \mathbb{P}, \underline{\mathbf{P}})$ such that Ω denotes a set of states, $\mathscr{A}$ denotes an algebra over Ω, $\mathbb{P}$ denotes a set of probability functions on $\mathscr{A}$, and $\underline{\mathbf{P}}$ denotes the *lower probability function* over $\mathscr{A}$ determined by $\mathbb{P}$ by the requirement that $\underline{\mathbf{P}}(A) = \inf\{p(A) : p \in \mathbb{P}\}$ for each $A \in \mathscr{A}$. The value $\underline{\mathbf{P}}(A)$ is called the *lower probability* of A. The *upper probability function* $\overline{\mathbf{P}}$ over $\mathscr{A}$ is accordingly defined, as usual, by stipulating that $\overline{\mathbf{P}}(A) = 1 - \underline{\mathbf{P}}(A^c)$ for each $A \in \mathscr{A}$; the value $\overline{\mathbf{P}}(A)$ is called the *upper probability* of A. Given $B \in \mathscr{A}$ for which $\underline{\mathbf{P}}(B) > 0$, conditional lower and upper probabilities are defined as $\underline{\mathbf{P}}(A \mid B) = \inf\{p(A \mid B) : p \in \mathbb{P}\}$ and $\overline{\mathbf{P}}(A \mid B) = \sup\{p(A \mid B) : p \in \mathbb{P}\}$, respectively. In the following, call a subcollection of events $\mathcal{B}$ from $\mathscr{A}$ a *positive measurable partition* (of Ω) if $\mathcal{B}$ is a partition of Ω such that $\underline{\mathbf{P}}(B) > 0$ for each $B \in \mathcal{B}$.

Let $\mathcal{B}$ be a positive measurable partition of Ω. Say that $\mathcal{B}$ *dilates* A if each $B \in \mathcal{B}$:

$$\underline{\mathbf{P}}(A \mid B) < \underline{\mathbf{P}}(A) \leq \overline{\mathbf{P}}(A) < \overline{\mathbf{P}}(A \mid B).^{1}$$

In other words, $\mathcal{B}$ dilates A just in case the closed interval $[\underline{\mathbf{P}}(A), \overline{\mathbf{P}}(A)]$ is contained within the open interval $(\underline{\mathbf{P}}(A \mid B), \overline{\mathbf{P}}(A \mid B))$ for each $B \in \mathcal{B}$. Examples of dilation are discussed in [7, 5, 6] and [9, §6.4.3]

1. While this terminology agrees with that of [3, p. 252], it differs from that of [8, p. 1141] and [4, p. 412], who call dilation in this sense *strict dilation*.

1.1. Measures of Dependence

Given a probability function p on algebra $\mathscr{A}$ and events $A, B \in \mathscr{A}$, define:

$$\mathsf{S}_p(A,B) := \begin{cases} \dfrac{p(A \cap B)}{p(A)p(B)} & \text{if } p(A)p(B) > 0; \\ 1 & \text{otherwise.} \end{cases}$$

Thus the quantity S_p is an index of deviation from stochastic independence between events. The value $\mathsf{S}_p(A,B)$ expresses in ratio form the covariance between events A and B, $\operatorname{cov}(A,B) = p(A \cap B) - p(A)p(B)$. Events A and B are stochastically independent if $\mathsf{S}_p(A,B) = 1$; positively correlated if $\mathsf{S}_p(A,B) > 1$, and negatively correlated if $\mathsf{S}_p(A,B) < 1$.

Given a set of probabilities $\mathbb{P}$ on $\mathscr{A}$ and events $A, B \in \mathscr{A}$, define:

$$\begin{aligned} \mathsf{S}^{+}_{\mathbb{P}}(A,B) &:= \{p \in \mathbb{P} : \mathsf{S}_p(A,B) > 1\}; \\ \mathsf{S}^{-}_{\mathbb{P}}(A,B) &:= \{p \in \mathbb{P} : \mathsf{S}_p(A,B) < 1\}; \\ \mathsf{I}_{\mathbb{P}}(A,B) &:= \{p \in \mathbb{P} : \mathsf{S}_p(A,B) = 1\}. \end{aligned}$$

The set of probability functions $\mathsf{I}_{\mathbb{P}}$ for which A and B are stochastically independent is called *the surface of independence* for A and B with respect to $\mathbb{P}$. In what follows, subscripts are dropped when there is no danger of confusion.

1.2. Characterizing Dilation

Given lower probability space $(\Omega, \mathscr{A}, \mathbb{P}, \underline{\mathbf{P}})$, events $A, B \in \mathscr{A}$ with $\underline{\mathbf{P}}(B) > 0$, and $\varepsilon > 0$, define:

$$\begin{aligned} \underline{\mathbb{P}}(A \mid B, \varepsilon) &:= \{p \in \mathbb{P} : |p(A \mid B) - \underline{\mathbf{P}}(A \mid B)| < \varepsilon\}; \\ \overline{\mathbb{P}}(A \mid B, \varepsilon) &:= \{p \in \mathbb{P} : |p(A \mid B) - \overline{\mathbf{P}}(A \mid B)| < \varepsilon\}. \end{aligned}$$

Call the sets $\underline{\mathbb{P}}(A \mid B, \varepsilon)$ and $\overline{\mathbb{P}}(A \mid B, \varepsilon)$ *lower* and *upper neighborhoods* of A *conditional* on B, respectively, with radius ε. A probability function p is a member of the lower neighborhood of A conditional on B with radius ε if $p(A \mid B)$

is within ε of $\underline{\mathbf{P}}(A \mid B)$, and similarly for an upper neighborhood.

Corollary 5.2 of [5] reports that $\mathcal{B}$ dilates A just in case there is $(\varepsilon_B)_{B\in\mathcal{B}} \in \mathbb{R}_+^{\mathcal{B}}$ such that $\underline{\mathbb{P}}(A \mid B, \varepsilon_B) \subseteq \mathsf{S}^-(A,B)$ and $\overline{\mathbb{P}}(A \mid B, \varepsilon_B) \subseteq \mathsf{S}^+(A,B)$, which Theorem 1 of [6] generalizes. The right-to-left implication admits a counterexample to be presented in the next section.

2. Counterexample

The following example, due to Michael Nielsen and Rush Stewart, was conveyed to us in correspondence.

Suppose $\Omega := \{\omega_1, \omega_2, \omega_3, \omega_4\}$ supports two probability functions, p_0 and p_1, such that:

	ω_1	ω_2	ω_3	ω_4
p_0	$1/8$	$1/2$	$1/3$	$1/24$
p_1	$5/24$	$1/24$	$1/24$	$17/24$

Let $\mathbb{P}$ be the convex hull of $\{p_0, p_1\}$.

Consider events $A := \{\omega_1, \omega_2\}$ and $B := \{\omega_1, \omega_3\}$ and partition $\mathcal{B} := \{B, B^c\}$. Observe that $\mathbb{P} = (p_\alpha)_{\alpha\in[0,1]}$, where $p_\alpha := (1-\alpha)p_0 + \alpha p_1$ for each α in $[0,1]$.

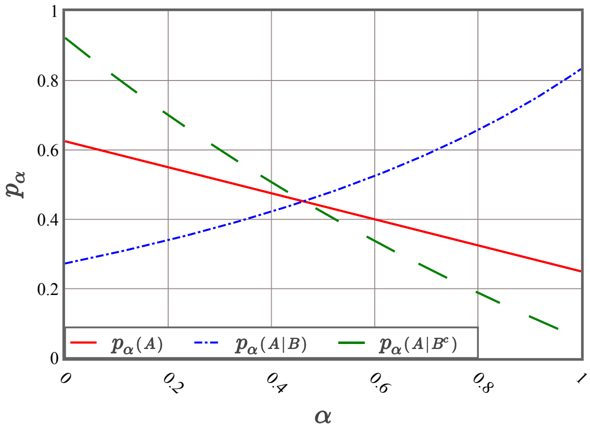

Figure 1: Graphing p_α's against values of α from 0 to 1.

Hence:

- $p_\alpha(A) = \frac{5-3\alpha}{8}$
- $p_\alpha(A \mid B) = \frac{3+2\alpha}{11-5\alpha}$, and
- $p_\alpha(A \mid B^c) = \frac{12-11\alpha}{13+5\alpha}$.

Figure 1 plots the values thus parametrized by α in $[0,1]$. It is readily established that there are positive real numbers ε_B and ε_{B^c} satisfying the requirements $\underline{\mathbb{P}}(A \mid B, \varepsilon_B) \subseteq \mathsf{S}^-(A,B)$ and $\overline{\mathbb{P}}(A \mid B^c, \varepsilon_{B^c}) \subseteq \mathsf{S}^+(A,B^c)$, while $\underline{\mathbf{P}}(A) = 1/4 < 3/11 = \underline{\mathbf{P}}(A \mid B)$, so $\mathcal{B}$ does not dilate A.

3. Repaired Result

Return to the example from Section 2 and observe that the partition $\mathcal{C} := \{A, A^c\}$ nevertheless dilates event B. That is, $\underline{\mathbf{P}}(B) = 1/4$ and $\overline{\mathbf{P}}(B) = 11/24$, while $\underline{\mathbf{P}}(B \mid A) = 1/5 < 1/4$ and $\underline{\mathbf{P}}(B \mid A^c) = 1/8 < 1/4$, as well as $11/24 < 5/6 = \overline{\mathbf{P}}(B \mid A)$ and $11/24 < 8/9 = \overline{\mathbf{P}}(B \mid A^c)$.

The foregoing example illustrates a key insight. While the results reported in [5] and [6] do indeed identify conditions which suffice to establish dilation between variables associated with $\mathcal{A}$ and $\mathcal{C}$, they not provide for conditions determining its direction. Yet since relevance might be asymmetric in this setting [1, 2], the indices S^- and S^+ of association are symmetric, so specifying the target event for dilation is important to rule out cases, like this one, where asymmetric relevance is in play.

Given a probability function p, a set of probability functions $\mathbb{P}$, and events A and B, define:

$$\underline{\mathsf{S}}_p(A,B) := \frac{p(A\cap B)}{\underline{\mathbf{P}}(A)p(B)} \quad \text{and} \quad \overline{\mathsf{S}}_p(A,B) := \frac{p(A\cap B)}{\overline{\mathbf{P}}(A)p(B)},$$

Likewise define:

$$\overline{\mathsf{S}}^+_{\mathbb{P}}(A,B) := \{p \in \mathbb{P} : \overline{\mathsf{S}}_p(A,B) > 1\};$$
$$\underline{\mathsf{S}}^-_{\mathbb{P}}(A,B) := \{p \in \mathbb{P} : \underline{\mathsf{S}}_p(A,B) < 1\}.$$

The following result is easily established:

Theorem 1 *Let A be an event and $\mathcal{B} = (B_i)_{i\in I}$ be a positive measurable partition for a given set of probability functions $\mathbb{P}$ over an algebra. The following statements are equivalent*

(i) $\mathcal{B}$ dilates A;

(ii) There exists $\varepsilon > 0$ such that for every $i \in I$:

$$\underline{\mathbb{P}}(A \mid B_i, \varepsilon) \subseteq \underline{\mathsf{S}}^-_{\mathbb{P}}(A,B_i) \text{ and } \overline{\mathbb{P}}(H \mid B_i, \varepsilon) \subseteq \overline{\mathsf{S}}^+_{\mathbb{P}}(A,B_i)$$

Proof For $(i) \Rightarrow (ii)$, suppose that $\mathcal{B}$ dilates A. Select $\varepsilon := \min(|\underline{\mathbf{P}}(A) - \underline{\mathbf{P}}(A \mid B_i)| : i \in I)$. For $i \in I$, suppose $|p(A \mid B_i) - \underline{\mathbf{P}}(A \mid B_i)| < \varepsilon$. Then by hypothesis it follows that $p(A \mid B_i) < \underline{\mathbf{P}}(A)$. So $p(A\cap B_i)/\underline{\mathbf{P}}(A)p(B_i) < 1$, thus $\underline{\mathsf{S}}_p(A,B) < 1$. Therefore, $\underline{\mathbb{P}}(A \mid B_i, \varepsilon) \subseteq \underline{\mathsf{S}}^-_{\mathbb{P}}(A,B_i)$. Similarly, $\overline{\mathbb{P}}(A \mid B_i, \varepsilon) \subseteq \overline{\mathsf{S}}^+_{\mathbb{P}}(A,B_i)$.

For $(ii) \Rightarrow (i)$, suppose that condition (ii) holds for some positive ε and assume for *reductio ad absurdum* that $\mathcal{B}$ fails to dilate A. Without loss of generality, suppose $\underline{\mathbf{P}}(A) \leq \underline{\mathbf{P}}(A \mid B_i)$ for some $i \in I$. Then there is a $p \in \underline{\mathbb{P}}(A \mid B_i, \varepsilon) \subseteq \underline{\mathsf{S}}^-_{\mathbb{P}}(A,B_i)$ such that $\underline{\mathsf{S}}_p(A,B_i) < 1$ and $\underline{\mathbf{P}}(A) \leq p(A \mid B_i) < \underline{\mathbf{P}}(A)$, yielding a contradiction. ■

Acknowledgments

We would like to thank Michael Nielsen and Rush Stewart for sharing their example with us. Gregory Wheeler's research is supported in part by the joint Agence Nationale de la Recherche (ANR) & Deutsche Forschungsgemeinschaft (DFG) project "Collective Attitudes Formation" ColAForm, award RO 4548/8-1,

References

[1] Inés Couso, Serafín Moral, and Peter Walley. Examples of independence for imprecise probabilities. In Gert de Cooman, editor, *Proceedings of the First Symposium on Imprecise Probabilities and Their Applications (ISIPTA)*, Ghent, Belgium, 1999.

[2] Fabio Cozman. Sets of probability distributions, independence, and convexity. *Synthese*, 186(2):577–600, 2012.

[3] Timothy Herron, Teddy Seidenfeld, and Larry Wasserman. The extent of dilation of sets of probabilities and the asymptotics of robust bayesian inference. In *PSA 1994 Proceedings of the Biennial Meeting of the Philosophy of Science Association*, volume 1, pages 250–259, 1994.

[4] Timothy Herron, Teddy Seidenfeld, and Larry Wasserman. Divisive conditioning: further results on dilation. *Philosophy of Science*, 64:411–444, 1997.

[5] Arthur Paul Pedersen and Gregory Wheeler. Demystifying dilation. *Erkenntnis*, 79(6):1305–1342, 2014.

[6] Arthur Paul Pedersen and Gregory Wheeler. Dilation, disintegrations, and delayed decisions. In *Proceedings of the 9th Symposium on Imprecise Probabilities and Their Applications (ISIPTA)*, pages 227–236, Pescara, Italy, 2015.

[7] Teddy Seidenfeld. When normal and extensive form decisions differ. In D. Prawitz, B. Skyrms, and D. Westerstahl, editors, *Logic, Methodology and Philosophy of Science*. Elsevier Science B. V., 1994.

[8] Teddy Seidenfeld and Larry Wasserman. Dilation for sets of probabilities. *The Annals of Statistics*, 21(9): 1139–154, 1993.

[9] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.