Computing lower expectations with Kuznetsov's independence condition*

Fabio Gagliardi Cozman
Escola Politécnica, University of São Paulo, Brazil

Abstract

Kuznetsov's condition says that variables X and Y are independent when any product of bounded functions $f(X)$ and $g(Y)$ behaves in a certain way: the interval of expected values $\mathbb{E}[f(X)g(Y)]$ must be equal to the interval product $\mathbb{E}[f(X)] \times \mathbb{E}[g(Y)]$. The main result of this paper shows how to compute lower expectations using Kuznetsov's condition. We also generalize Kuznetsov's condition to conditional expectation intervals, and study the relationship between Kuznetsov's conditional condition and the semi-graphoid properties.

Keywords

Sets of probability distributions, lower expectations, probability intervals, expectation intervals, independence concepts.

1 Introduction

Kuznetsov's condition says that two variables X and Y are independent if, for any two bounded functions $f(X)$ and $g(Y)$, we have

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)] \times \mathbb{E}[g(Y)], \tag{1}$$

where $\mathbb{E}[\cdot]$ denotes an interval of expected values and the product is understood as interval multiplication [8].

Kuznetsov's condition is geared towards models that represent uncertainty through sets of probability measures and expectation intervals. In those models, Kuznetsov's condition is seen to be more general than the standard definition of stochastic independence. The condition can be viewed as a definition of independence, and also as a constraint to be used when building models that involve imprecise beliefs. The relationship between Kuznetsov's condition and other concepts of independence was analyzed in a previous paper [5]; several results from that publication are used in this paper.

*This work has been supported in part by CNPq through grant 300183/98-4, and in part by HP Labs through "Convênio Aprendizado de Redes Bayesianas".

This paper shows how to compute minima and maxima of expected values using Kuznetsov's condition. The main result is a characterization of the largest credal set that complies with Kuznetsov's condition — the "Kuznetsov's extension" of marginal sets. We discuss the computation of lower expectations from Kuznetsov's extensions, and investigate the connection between Kuznetsov's extensions and other extensions used in the literature. Section 4 contains these developments.

We then generalize Kuznetsov's condition to conditional beliefs (Section 5). To clarify the behavior of the resulting condition, we investigate its compliance to the semi-graphoid properties. We show that Kuznetsov's conditional condition satisfies symmetry, redundancy, decomposition and weak union, but fails the contraction property.

Kuznetsov's condition is an interesting tool for modeling independence with imprecise beliefs. This paper provides the basic machinery to manipulate the condition in practice. Section 6 presents our conclusions.

2 Credal sets, lower expectations, extensions

In this section we review the basic concepts necessary for later developments. Consider two random variables X and Y. In this paper all variables have finitely many values. The probability density for X is denoted by $p(X)$, and $E_p[f(X)]$ denotes the expectation of function $f(X)$ with respect to $p(X)$. A non-empty set of probability measures is called a *credal set* [9]; a credal set consisting of densities $p(X)$ is denoted by $K(X)$. A credal set $K(X,Y)$ consisting of joint densities $p(X,Y)$ is called a *joint credal set*. The *lower* and *upper* expectations of function $f(X)$ are respectively $\underline{E}[f(X)] = \min_{p(X)\in K(X)} E_p[f(X)]$ and $\overline{E}[f(X)] = \max_{p(X)\in K(X)} E_p[f(X)]$. The *lower probability* and the *upper probability* of event A are defined similarly. A credal set produces an expectation interval for any bounded function $h(X)$: $\mathbb{E}[h(X)] = \left[\underline{E}[h(X)], \overline{E}[h(X)]\right]$.

There are several concepts of independence that can be applied to credal sets [2, 7]; here we focus on *epistemic independence* and *strong independence*. Variable Y is *epistemically irrelevant* to X if $K(X|y)$ and $K(X)$ have the same convex hull for all possible values of Y (equivalently, $\underline{E}[f(X)|y] = \underline{E}[f(X)]$ for any bounded function $f(X)$ and all possible values of Y). Variables X and Y are *epistemically independent* if X is irrelevant to Y and Y is irrelevant to X. Strong independence focuses instead on decomposition of probability measures [1, 2, 4]: Variables X and Y are *strongly independent* when every extreme point of $K(X,Y)$ satisfies standard stochastic independence of X and Y.

Given marginal credal sets $K(X)$ and $K(Y)$, there may be several credal sets $K(X,Y)$ for which X and Y are independent. Each one of these sets is called an *extension* of $K(X)$ and $K(Y)$. Given marginal sets $K(X)$ and $K(Y)$, their *epistemic extension* (called the *independent natural extension* by Walley) is the largest joint credal set that satisfies epistemic independence with marginals $K(X)$ and $K(Y)$ [13]. Their *strong extension* is the largest joint credal set that satisfies strong independence with marginals $K(X)$ and $K(Y)$ [2, 4]. The term *natural extension* is used to indicate the largest possible extension given whatever constraints on probability and independence are adopted [13].

A credal set $K(X,Y)$ is *finitely generated* when it is a polytope in the space of probability measures — the convex hull of a finite number of probability distributions. Such a set is defined by a finite collection of linear inequalities such as $\sum_{X,Y} h(X,Y)p(X,Y) \geq 0$. In the remainder of this paper, f indicates a function of X, g indicates a function of Y and h indicates a function of X and Y. Similarly, p indicates a density for X, q indicates a density for Y; other densities, such as $p(X,Y)$, are indicated explicitly. We can view functions and probability densities as vectors, so we can write $(fg) \cdot (pq) \geq 0$ instead of $\sum_{X,Y} f(X)g(Y)p(X)q(Y) \geq 0$, using the dot product to produce summation.

Note that any hyperplane $h \cdot p(X,Y) = 0$ goes through the origin. The function/vector h is the normal vector of the hyperplane. If $\underline{E}[h] = 0$, then h defines a *supporting hyperplane* for the credal set. If $\overline{E}[h] = 0$, then $-h$ is a supporting hyperplane. A *face* of a polytope is the intersection of the polytope with a supporting hyperplane; a *facet* is a maximal face distinct of the polytope [11].

To simplify notation, we use the same letter (f, for instance) for a function, a vector (containing the values of a function), a normal vector (orthogonal to an hyperplane), an hyperplane (with the normal vector), or a facet (contained in the hyperplane with the normal vector), depending on the circumstances.

Any function/vector h can be written as $h' + \underline{E}[h]$ or as $-h'' + \overline{E}[h]$, where h' and h'' are supporting hyperplanes that are parallel to h. Consider any supporting hyperplane h' that goes through a vertex V. Take the facets intersecting at V, and the normal vectors to these facets. Then it must be possible to write h' as a linear of these normal vectors.

3 Kuznetsov's condition and Kuznetsov's extension

Kuznetsov's condition is a condition for independence operating on expectations of independent variables [8]. The condition can be expressed either in terms of expectation intervals (Expression (1)), or as

$$\underline{E}[f(X)g(Y)] = \min\left(\begin{array}{c} \underline{E}[f(X)]\,\underline{E}[g(Y)]\,, \underline{E}[f(X)]\,\overline{E}[g(Y)]\,, \\ \overline{E}[f(X)]\,\underline{E}[g(Y)]\,, \overline{E}[f(X)]\,\overline{E}[g(Y)] \end{array} \right). \tag{2}$$

To obtain (2) from (1), we recall that the interval product $[a,b] \times [c,d]$ is equal to

$$[\min(ac,ad,bc,bd), \max(ac,ad,bc,bd)]\,.$$

The following result is used later:

Theorem 1 *For any bounded functions $f(X)$ and $g(Y)$, any extension that satisfies Kuznetsov's condition must contain densities that attain $\underline{E}[f]\,\underline{E}[g]$, $\underline{E}[f]\,\overline{E}[g]$, $\overline{E}[f]\,\underline{E}[g]$, and $\overline{E}[f]\,\overline{E}[g]$.*

Proof. Suppose we have a credal set that satisfies Kuznetsov's condition. Consider a function $h_1 = (f - \underline{E}[f] + \alpha)(g - \underline{E}[g] + \beta)$, where $\alpha, \beta > 0$; then $\underline{E}[h_1] = (\underline{E}[f - \underline{E}[f]] + \alpha)(\underline{E}[g - \underline{E}[g]] + \beta) = \alpha\beta$ for any α, β. But for this to happen, we must have a density $p_1(X,Y)$ such that $E_{p_1}[f] = \underline{E}[f]$ and $E_{p_1}[g] = \underline{E}[g]$ at the same time. The proof can be completed by taking functions $h_2 = (f - \overline{E}[f] + \alpha)(g - \overline{E}[g] + \beta)$, $h_3 = -(f - \overline{E}[f] + \alpha)(g - \underline{E}[g] + \beta)$ and $h_4 = -(f - \underline{E}[f] + \alpha)(g - \overline{E}[g] + \beta)$. $\square$

We can use Kuznetsov's condition to construct credal sets. Suppose we have $K(X)$ and $K(Y)$, and we obtain the information that X and Y satisfy Kuznetsov's condition, without further information on $K(X,Y)$. What can we say about the joint credal set $K(X,Y)$? A reasonable strategy is to focus on the largest joint credal set that satisfies Kuznetsov's condition and has the marginals $K(X)$ and $K(Y)$. This set is referred to as *Kuznetsov's extension* of $K(X)$ and $K(Y)$. It should be noted that a Kuznetsov's extension always exists [5].

Kuznetsov's extensions are smaller than epistemic extensions when all events have positive probability, as in this case Kuznetsov's independence implies epistemic independence — and even when all lower probabilities are larger than zero Kuznetsov's extensions can be strictly smaller than epistemic extensions [5]. A strong extension always satisfies Kuznetsov's condition and is contained in the corresponding Kuznetsov's extension (however, the Kuznetsov's extension can be strictly larger than the strong extension; also, it is possible that a credal set satisfies strong independence but does not satisfy Kuznetsov's condition) [5].

4 Characterizing Kuznetsov's extensions

Suppose we have two binary variables X and Y, and we construct the strong extension of $K(X)$ and $K(Y)$. In this case, it is known that the strong extension and the Kuznetsov's extension of $K(X)$ and $K(Y)$ are identical [5]. A more general result can actually be proved:

Theorem 2 *Consider a binary variable X with credal set $K(X)$, and a variable Y with N values and credal set $K(Y)$ with M vertices; the strong extension and Kuznetsov's extension of $K(X)$ and $K(Y)$ are identical.*

Proof. The strong extension is composed of vertices of the form $p_i(X)q_j(Y)$, where p_i indicates a vertex of $K(X)$ and q_j indicates a vertex of $K(Y)$. If $K(X)$ contains a single point, the result is immediate; suppose that $K(X)$ has two vertices p_1 and p_2 (so there is a function $f_1(X)$ such that $f_1 \cdot p_1 = 0$, and a function

$f_2(X)$ such that $f_2 \cdot p_2 = 0$). The strong extension can have at most $2M$ vertices, all of them with $2N$ components (thus the strong extension lives in $(2N-1)$ dimensional space). Any facet of the strong extension is contained in an hyperplane that is defined by selecting $(2N-1)$ vertices of the strong extension plus the origin. Take a facet and divide its vertices (other than the origin) in two sets:
(i) the set C_1 containing points of the form $p_1 q_j$,
(ii) the set C_2 containing points of the form $p_2 q_k$,
where q_j, q_k are vertices of $K(Y)$. Suppose that C_1 contains more points than C_2. Then we have at most $N-1$ points in C_2; we can always find an hyperplane defined by a function $g(Y)$ that goes through all these points. Thus we can construct a function $f_1(X)g(Y)$ such that

$$\sum_{X,Y} f_1(X)g(Y)p_1(X)q_j(Y) = \left(\sum_X f_1(X)p_1(X)\right)\left(\sum_Y g(Y)q_j(Y)\right) = 0$$

for every point in C_1 and every point in C_2. So the facet is represented by a decomposable function $f_1 g$. The same construction can be followed if C_2 has more elements than C_1, in which case we will arrive at a decomposable function of the form $f_2 g'$ for some $g'(Y)$. Thus, any facet of the strong extension is defined by a decomposable hyperplane and consequently is a valid constraint for Kuznetsov's extension. The strong extension must then contain Kuznetsov's extension, and so both are equal. □

The facets generated in the proof of Theorem 2 are of the form $f(X)g(Y)$. A little reflection shows that this function $g(Y)$ must define a supporting hyperplane of $K(Y)$: If g were not a supporting hyperplane of $K(Y)$, there should be a point q_c such that $\sum_Y g q_c \geq 0$ and a point q_d such that $\sum_Y g q_d \leq 0$. But $g \cdot q_c \geq 0$ would imply $(fg)\cdot(p_1 q_c) \geq 0$ and $g \cdot q_d \leq 0$ would imply $(fg)\cdot(p_1 q_d) \leq 0$, contradicting the fact that fg is a supporting hyperplane for the strong extension. Consequently, the facets of the strong extension in Theorem 2 are defined by decomposable functions that factorize into facets of $K(X)$ and $K(Y)$.

Consider now a more general situation where we have categorical variables X and Y and finitely generated marginal credal sets $K(X)$ and $K(Y)$. Suppose that, instead of trying to compute Kuznetsov's extensions, someone simply constructed the following inequalities:

$$\begin{aligned} \sum_{X,Y} \tilde{f}_i(X)p(X,Y) &\geq 0, \\ \sum_{X,Y} \tilde{g}_j(Y)p(X,Y) &\geq 0, \\ \sum_{X,Y} (\tilde{f}_i(X)\tilde{g}_j(Y))p(X,Y) &\geq 0, \end{aligned} \tag{3}$$

which can be written as

$$\tilde{f}_i \cdot p(X,Y) \geq 0, \quad \tilde{g}_j \cdot p(X,Y) \geq 0, \quad (\tilde{f}_i \tilde{g}_j) \cdot p(X,Y) \geq 0, \tag{4}$$

for all combinations of i and j, where $\tilde{f}_i$ is a facet of $K(X)$ and $\tilde{g}_j$ is a facet of $K(Y)$. Note that any set of densities that satisfies these inequalities will also satisfy $(f'g') \cdot p(X,Y) \geq 0$, where f' and g' are supporting hyperplanes of $K(X)$ and $K(Y)$ respectively.

The next theorem is the main result: it shows how to explicitly construct Kuznetsov's extensions. The proof essentially consists of showing that any inequality required by Kuznetsov's condition is already implied by inequalities (4).

Theorem 3 *Consider a variable X with finitely generated credal set $K(X)$, defined by facets $\tilde{f}_i$, and a variable Y with finitely generated credal set $K(Y)$, defined by facets $\tilde{g}_j$. The Kuznetsov's extension is entirely defined by the facets $\tilde{f}_i$, $\tilde{g}_j$, and $(\tilde{f}_i \tilde{g}_j)$, for all combinations of i and j.*

Proof. Denote by $K_k(X,Y)$ the credal set constructed in the theorem. Every vertex of the strong extension is of the form $p(X)q(Y)$ and consequently satisfies $(\tilde{f}_i \tilde{g}_j) \cdot (pq) \geq 0$. We conclude that the strong extension is contained in $K_k(X,Y)$, thus the expectation intervals generated by the strong extension are contained in the expectation intervals generated by $K_k(X,Y)$. Furthermore, for every decomposable function $f(X)g(Y)$, there is a density in $K_k(X,Y)$ that attains the value prescribed by Kuznetsov's condition, as the strong extension is contained in $K_k(X,Y)$.
Now take two arbitrary bounded functions $f(X)$ and $g(Y)$. There are seven different situations to consider:

1. $\underline{E}[f] \geq 0$, $\underline{E}[g] \geq 0$: Kuznetsov's condition requires that $\underline{E}[fg] = \underline{E}[f]\,\underline{E}[g]$. Write f as $f' + \underline{E}[f]$ (f' is a supporting hyperplane of $K(X)$) and write g as $g' + \underline{E}[g]$ (g' is a supporting hyperplane of $K(Y)$). Then we have: $fg \cdot p(X,Y) = (f' + \underline{E}[f])(g' + \underline{E}[g]) \cdot p(X,Y) = f'g' \cdot p(X,Y) + \underline{E}[f]\,g' \cdot p(X,Y) + \underline{E}[g]\,f' \cdot p(X,Y) + \underline{E}[f]\,\underline{E}[g]$, an expression that is equal to or larger than $\underline{E}[f]\,\underline{E}[g]$ given that $p(X,Y)$ satisfies inequalities (4). This implies that $E_p[fg] \geq \underline{E}[f]\,\underline{E}[g]$ for every $p(X,Y)$ and we obtain $\underline{E}[fg] = \underline{E}[f]\,\underline{E}[g]$ (because the inclusion of the strong extension in $K_k(X,Y)$ guarantees that the equality obtains).

2. $\overline{E}[f] \leq 0$, $\overline{E}[g] \leq 0$: Kuznetsov's condition requires $\underline{E}[fg] = \overline{E}[f]\,\overline{E}[g]$. To show that $E_p[fg] \geq \overline{E}[f]\,\overline{E}[g]$ for every $p(X,Y)$, write f as $-f' + \overline{E}[f]$ and g as $-g' + \overline{E}[g]$ (where f' and g' are appropriate supporting hyperplanes), and then: $fg \cdot p(X,Y) = (-f' + \overline{E}[f])(-g' + \overline{E}[g]) \cdot p(X,Y)$, a quantity that is equal to or larger than $\overline{E}[f]\,\overline{E}[g]$ given inequalities (4).

3. $\underline{E}[f] \geq 0, \underline{E}[g] \leq 0$: Kuznetsov's condition requires $\underline{E}[fg] = \overline{E}[f]\,\underline{E}[g]$. Write $f = f' + \underline{E}[f]$, $f = -f'' + \overline{E}[f]$, and $g = -g' + \overline{E}[g]$ (where f', f'' and g'

are appropriate supporting hyperplanes; note that f is written in two different ways) and then $fg \cdot p(X,Y) = f(g' + \underline{E}[g]) \cdot p(X,Y) = ((f' + \underline{E}[f])g' + (-f'' + \overline{E}[f])\underline{E}[g]) \cdot p(X,Y)$, which that is equal to or larger than $\overline{E}[f]\,\underline{E}[g]$.

4. $\underline{E}[f] \leq 0$, $\overline{E}[f] \geq 0$, $\overline{E}[g] \leq 0$: Kuznetsov's condition requires $\underline{E}[fg] = \overline{E}[f]\,\underline{E}[g]$. Write $f = -f' + \overline{E}[f]$, $g = g' + \underline{E}[g]$, and $g = -g'' + \overline{E}[g]$, and then $fg \cdot p(X,Y) = (-f'(-g'' + \overline{E}[g]) + \overline{E}[f]\,(g' + \underline{E}[g])) \cdot p(X,Y)$, which is equal to or larger than $\overline{E}[f]\,\underline{E}[g]$.

5. $\underline{E}[f] \leq 0, \underline{E}[g] \geq 0$: Kuznetsov's condition requires $\underline{E}[fg] = \underline{E}[f]\,\overline{E}[g]$. Write $f = f' + \underline{E}[f]$, $g = g' + \underline{E}[g]$, and $g = -g'' + \overline{E}[g]$, and then $fg \cdot p(X,Y) = (f'(g' + \underline{E}[g]) + \underline{E}[f]\,(-g'' + \overline{E}[g])) \cdot p(X,Y)$, which is equal to or larger than $\underline{E}[f]\,\overline{E}[g]$.

6. $\overline{E}[f] \leq 0, \underline{E}[g] \leq 0, \overline{E}[g] \geq 0$: Kuznetsov's condition requires that $\overline{E}[fg] = \underline{E}[f]\,\overline{E}[g]$. Write $f = f' + \underline{E}[f]$, $f = -f'' + \overline{E}[f]$, and $g = -g' + \overline{E}[g]$, and $fg \cdot p(X,Y) = (-g'(-f'' + \overline{E}[f]) + \overline{E}[g]\,(f' + \underline{E}[f])) \cdot p(X,Y) = (f''g' + \overline{E}[g]\,f' - \overline{E}[f]\,g') \cdot p(X,Y) + \underline{E}[f]\,\overline{E}[g]$, equal to or larger than $\underline{E}[f]\,\overline{E}[g]$.

7. $\underline{E}[f] \leq 0$, $\overline{E}[f] \geq 0$, $\underline{E}[g] \leq 0$, $\overline{E}[g] \geq 0$: Kuznetsov's condition requires $\underline{E}[fg] = \min(\underline{E}[f]\,\overline{E}[g], \overline{E}[f]\,\underline{E}[g])$. Divide $K_k(X,Y)$ into two sets. Define $K_1(X,Y)$ to contain the distributions in $K_k(X,Y)$ such that $f \cdot p(X,Y) \geq 0$, and $K_2(X,Y)$ to contain the distributions in $K_k(X,Y)$ such that $f \cdot p(X,Y) \leq 0$. The value of $\underline{E}[fg]$ with respect to $K_k(X,Y)$ is the minimum of $\underline{E}[fg]$ with respect to $K_1(X,Y)$ and $K_2(X,Y)$. Following the previous cases, we obtain $\overline{E}[f]\,\underline{E}[g]$ as $\underline{E}[fg]$ with respect to $K_1(X,Y)$, and $\underline{E}[f]\,\overline{E}[g]$ as $\underline{E}[fg]$ with respect to $K_2(X,Y)$. We finally obtain $\underline{E}[fg] = \min(\underline{E}[f]\,\overline{E}[g], \overline{E}[f]\,\underline{E}[g])$.

Thus $K_k(X,Y)$ satisfies Kuznetsov's condition, and Kuznetsov's extension must contain $K_k(X,Y)$ — however Kuznetsov's extension cannot be larger than the set $K_k(X,Y)$, as every inequality (4) is directly required by Kuznetsov's condition. □

Once we know how to construct Kuznetsov's extensions, we can compute $\underline{E}[h(X,Y)]$ for a non-decomposable function $h(X,Y)$:

$$\underline{E}[h(X,Y)] = \min\left(h(X,Y) \cdot p(X,Y)\right), \tag{5}$$

subject to $p(X,Y) \geq 0$, $\sum_{X,Y} p(X,Y) = 1$, and inequalities (4).

The linear program (5) provides the solution to the question, Which (decomposable) constraints to use when computing a lower expectation for Kuznetsov's extension? Theorem 3 proves that inequalities (4) contain all the relevant constraints. Kuznetsov himself seems to have obtained different results, using his condition and additional factorization conditions to define extensions — a framework that led him to prescribe linear programs with infinitely many constraints [8].

Finally, note that we can also use linear programming if we need to compute a conditional lower expectation such as $\underline{E}[h|A]$ for some event A where $\underline{P}(A) > 0$.

The computation of $\underline{E}[h|A]$ requires the solution of a fractional linear program that can be performed using the Charnes-Cooper transformation and linear programming [3], using inequalities (4) as the starting point.

5 Kuznetsov's conditional condition and the semi-graphoid properties

Kuznetsov's condition does not deal with the concept of conditional independence, but it can certainly be extended to do so. Say that two variables X and Y are independent conditional on Z if, for bounded functions $f(X)$ and $g(Y)$,

$$\mathbb{E}[fg|z] = \mathbb{E}[f|z] \times \mathbb{E}[g|z],$$

for any value of Z (we assume that conditioning events have positive lower probability).

How appropriate is Kuznetsov's conditional condition as a concept of conditional independence? One way to study concepts of independence is to verify the *semi-graphoid* properties satisfied by the concept [6, 10, 12]. A relation $(X \perp\!\!\!\perp Y \,|\, Z)$ is called a *semi-graphoid* when it satisfies the following axioms:
Symmetry: $(X \perp\!\!\!\perp Y \,|\, Z) \Rightarrow (Y \perp\!\!\!\perp X \,|\, Z)$
Redundancy: $(X \perp\!\!\!\perp Y \,|\, X)$
Decomposition: $(X \perp\!\!\!\perp (W,Y) \,|\, Z) \Rightarrow (X \perp\!\!\!\perp Y \,|\, Z)$
Weak union: $(X \perp\!\!\!\perp (W,Y) \,|\, Z) \Rightarrow (X \perp\!\!\!\perp Y \,|\, (W,Z))$
Contraction: $(X \perp\!\!\!\perp Y \,|\, Z)$ & $(X \perp\!\!\!\perp W \,|\, (Y,Z)) \Rightarrow (X \perp\!\!\!\perp (W,Y) \,|\, Z)$.

Denote by $(X \perp\!\!\!\perp_K Y \,|\, Z)$ the fact that X and Y satisfy Kuznetsov's condition conditional on Z. The notation $\overline{\underline{E}}[f]$ is used to indicate either $\underline{E}[f]$ or $\overline{E}[f]$, whatever value is required by Kuznetsov's condition. We have:

Theorem 4 *Kuznetsov's conditional condition satisfies symmetry, redundancy, weak union and decomposition when applied to credal sets where no event has zero lower probability.*

Proof. Symmetry is immediate, and redundancy follows from $\mathbb{E}[f(X)g(Y)|x_0] = f(x_0)\mathbb{E}[g(Y)|x_0] = \mathbb{E}[f(X)|x_0] \times \mathbb{E}[g(Y)|x_0]$ for any $f(X)$, $g(Y)$, and any x_0. Decomposition follows from the fact that any function of Y is also a function of Y and W, so we have $\mathbb{E}[f(X)g(Y)|z] = \mathbb{E}[f(X)|z] \times \mathbb{E}[g(Y)|z]$ when $(X \perp\!\!\!\perp_K (W,Y) \,|\, Z)$.
To simplify the proof of the weak union property, the conditioning variable Z is suppressed. What must be shown is that $\underline{E}[fg|w] = \overline{\underline{E}}[f]\overline{\underline{E}}[g|w]$ follows from $\underline{E}[fh] = \overline{\underline{E}}[f]\overline{\underline{E}}[h]$, where h is any function of W and Y (note that $\overline{\underline{E}}[f] = \overline{\underline{E}}[f|w]$ by hypothesis, as events have positive lower probability). Theorem 1 can be easily modified to prove that any credal set satisfying Kuznetsov's condition must contain densities that attain $\underline{E}[f]\,\underline{E}[g|w]$, $\underline{E}[f]\,\overline{E}[g|w]$, $\overline{E}[f]\,\underline{E}[g|w]$ and $\overline{E}[f]\,\overline{E}[g|w]$; thus, there is always a density p in a set that satisfies Kuznetsov's condition

such that $E_p[fg|w] = \underline{E}[fg|w]$, where $\underline{E}[fg|w]$ follows Kuznetsov's condition. Take Kuznetsov's extension of $K(X)$ and $K(W,Y)$, denoted by $K_k(W,X,Y)$. This extension must be equal to or larger than any set satisfying $X \perp\!\!\!\perp_K (W,Y)$. If we determine that $\underline{E}[fg|w] \geq \overline{\underline{E}}[f]\overline{\underline{E}}[g|w]$ for $K_k(W,X,Y)$, then automatically we obtain $\underline{E}[fg|w] = \overline{\underline{E}}[f]\overline{\underline{E}}[g|w]$ for any set satisfying $(X \perp\!\!\!\perp_K (W,Y))$, and weak union follows. The Kuznetsov's extension $K_k(W,X,Y)$ satisfies any inequality $h(W,X,Y) \cdot p(W,X,Y) \geq 0$, and so it satisfies $(f(X)g(Y)I_w(W)) \cdot p(W,X,Y) \geq 0$ for any $f(X)$, $g(Y)$ and w. If we consider the conditional distributions $p(X,Y|w)$ obtained from $K_k(W,X,Y)$, they must satisfy $(f(X)g(Y)) \cdot p(X,Y|w) \geq 0$ as this last inequality is obtained by normalizing the previous one. If we were to construct the Kuznetsov's extension of $K(X)$ and $K(Y|w)$, where $K(Y|w)$ is obtained from $K(W,X)$ by conditioning, then this Kuznetsov's extension would also satisfy any inequality $(f(X)g(Y)) \cdot p(X,Y|w) \geq 0$. So, every inequality constraining the Kuznetsov's extension of $K(X)$ and $K(Y|w)$ is also a constraint for the conditional set obtained from $K_k(W,X,Y)$. Thus the former set is equal to or larger than the latter set. Now notice that, for the Kuznetsov's extension of $K(X)$ and $K(Y|w)$, $\underline{E}[fg|w] = \overline{\underline{E}}[f]\overline{\underline{E}}[g|w]$, and so we must have $\underline{E}[fg|w] \geq \overline{\underline{E}}[f]\overline{\underline{E}}[g|w]$ for $K_k(W,X,Y)$. $\square$

Kuznetsov's condition does not imply the contraction property, as the next example shows.

Example 1 Consider binary variables W, X, and Y, and a credal set $K(W,X,Y)$ with eight vertices such that each vertex decomposes as $p(W|Y)\,p(X)\,p(Y)$. Values of $p(w_0|y_0)$, $p(w_0|y_1)$, $p(x_0)$ and $p(y_0)$ are:

Vertex	$[p(w_0\|y_0), p(w_0\|y_1), p(x_0), p(y_0)]$	Vertex	$[p(w_0\|y_0), p(w_0\|y_1), p(x_0), p(y_0)]$
1	[0.7,0.4,0.3,0.2]	5	[0.7,0.4,0.3,0.3]
2	[0.7,0.5,0.2,0.2]	6	[0.7,0.5,0.3,0.3]
3	[0.8,0.4,0.2,0.2]	7	[0.8,0.4,0.3,0.3]
4	[0.8,0.5,0.2,0.2]	8	[0.8,0.5,0.2,0.3]

It can be verified that the set of marginal densities $K(X,Y)$ contains every combination of $p(x_0)$ and $p(y_0)$, so $K(X,Y)$ is the Kuznetsov's extension for X and Y (Theorem 2). Likewise, $K(W,X|y_0)$ is the Kuznetsov's extension of W and X conditional on y_0, and $K(W,X|y_1)$ is the Kuznetsov's extension of W and X conditional on y_1. Thus the credal set $K(W,X,Y)$ satisfies $(X \perp\!\!\!\perp_K Y)$ and $(X \perp\!\!\!\perp_K W\,|\,Y)$, but it is not true that $X \perp\!\!\!\perp_K (W,Y)$. Take the function $f(X) = [1,2]$ and the function $h(W,Y) = [2,1,1,2]$. Then $\underline{E}[fh] = 2.652$ for $K(W,X,Y)$, but $\underline{E}[f]\,\underline{E}[h] = 1.7 \times 1.54 = 2.61$ — violating Kuznetsov's condition. $\square$

Despite the failure of contraction for generic credal sets, there is an important situation where contraction holds with Kuznetsov's condition.

Theorem 5 *Kuznetsov's conditional condition satisfies the contraction property when applied to credal sets where no events have zero lower probability, and such that the sets $K(X)$, $K(Y)$, and $K(W|Y)$ are separately specified.*

Proof. As the relevant sets are separately specified, minimization can occur separately within each set, so $\underline{E}[f(X)h(W,Y)] = \min E_p[E_p[fh|Y]] = \min E_p[\underline{E}[fh|Y]]$. As we have $(X \perp\!\!\!\perp_K W \,|\, Y)$, $\underline{E}[f(X)h(W,Y)] = \min E_p\left[\overline{\underline{E}}[f|Y]\overline{\underline{E}}[h|Y]\right]$, and because $X \perp\!\!\!\perp_K Y$, $\underline{E}[f(X)h(W,Y)] = \min \overline{\underline{E}}[f] E_p\left[\overline{\underline{E}}[h|Y]\right] = \overline{\underline{E}}[f]\overline{\underline{E}}[h]$. □

6 Conclusion

A Kuznetsov's extension can be viewed as a set that "wraps" a strong extension using decomposable hyperplanes. In fact, there is an interesting duality between these two extensions; while the former is constructed with decomposable hyperplanes, the latter is constructed with decomposable measures.

Kuznetsov's extensions can have complex structures, except when binary variables are present. The fact that the conditional version of Kuznetsov's condition fails the contraction property is troubling. This failure suggests that it may be hard to simplify multivariate models using only judgements of conditional independence (according to Kuznetsov's condition), as these judgements are coupled with the contraction property in traditional multivariate probabilistic models [10].

The challenges for the future are to determine when Kuznetsov's extensions (and derived concepts) are applicable in practice and how to manipulate them efficiently.

References

[1] J. Cano, M. Delgado, and S. Moral. An axiomatic framework for propagating uncertainty in directed acyclic networks. *Int. Journal of Approximate Reasoning*, 8:253–280, 1993.

[2] I. Couso, S. Moral, and P. Walley. Examples of independence for imprecise probabilities. *First Int. Symp. on Imprecise Probabilities and Their Applications*, pages 121–130, Ghent, Belgium, 1999.

[3] F. G. Cozman. Calculation of posterior bounds given convex sets of prior probability measures and likelihood functions. *Journal of Computational and Graphical Statistics*, 8(4):824–838, 1999.

[4] F. G. Cozman. Separation properties of sets of probabilities. *XVI Conf. on Uncertainty in Artificial Intelligence*, pages 107–115, San Francisco, July 2000. Morgan Kaufmann.

[5] F. G. Cozman. Constructing sets of probability measures through Kuznetsov's independence condition. In *Second Int. Symp. on Imprecise Probabilities and Their Applications*, pages 104–111, Ithaca, New York, 2001.

[6] A. P. Dawid. Conditional independence. *Encyclopedia of Statistical Sciences, Update Volume 2*, pages 146–153. Wiley, New York, 1999.

[7] L. de Campos and S. Moral. Independence concepts for convex sets of probabilities. *XI Conf. on Uncertainty in Artificial Intelligence*, pages 108–115, San Francisco, California, United States, 1995. Morgan Kaufmann.

[8] V. P. Kuznetsov. *Interval Statistical Methods*. Radio i Svyaz Publ., (in Russian), 1991.

[9] I. Levi. *The Enterprise of Knowledge*. MIT Press, Cambridge, Massachusetts, 1980.

[10] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, California, 1988.

[11] A. Schrijver. Theory of Linear and Integer Programming. John Wiley and Sons Ltd., New York, 1986.

[12] M. Studeny. Semigraphoids and structures of probabilistic conditional independence. *Annals of Mathematics and Artificial Intelligence*, 21(1):71–98, 1997.

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

Fabio Gagliardi Cozman is with the Engineering School (Escola Politécnica), University of São Paulo, Av. Prof. Mello Moraes, 2231, Cidade Universitária, São Paulo, SP, Brazil, CEP 05508-900. E-mail: fgcozman@usp.br