Constructing Sets of Probability Measures Through Kuznetsov's Independence Condition

Fabio Gagliardi Cozman
Escola Politécnica, University of São Paulo
Av. Prof. Mello Moraes, 2231 - CEP05508-900 - São Paulo, SP - Brazil
fgcozman@usp.br

Abstract

The purpose of this paper is to investigate a condition, suggested by Kuznetsov, to be required of independent variables. Kuznetsov's condition demands decomposable functions to be associated with decomposable expectation intervals. The paper demonstrates that Kuznetsov's condition does enlarge the universe of models based on sets of probability measures, as the condition is not equivalent to existing concepts of independence.

Keywords. Sets of probability measures, lower expectations, independence concepts, extensions.

1 Introduction

The purpose of this paper is to investigate Kuznetsov's condition on independent variables. Say we have two variables X and Y, and we represent beliefs about X by a set of probability measures $K(X)$; similarly, we represent beliefs about Y by a set $K(Y)$. Kuznetsov's condition demands that, if two variables X and Y are independent, then the lower expectation of any decomposable function $f(X)g(Y)$ must be

$$\underline{E}[f(X)g(Y)] = \min \begin{pmatrix} \underline{E}[f(X)]\,\underline{E}[g(Y)]\,, \\ \underline{E}[f(X)]\,\overline{E}[g(Y)]\,, \\ \overline{E}[f(X)]\,\underline{E}[g(Y)]\,, \\ \overline{E}[f(X)]\,\overline{E}[g(Y)] \end{pmatrix}. \quad (1)$$

This condition may look too complicated at first sight, but in fact it has a simple interpretation in terms of expectation intervals. Kuznetsov's condition demands that the interval that contains $E[f(X)g(Y)]$ must be obtained as the interval-product of intervals for $E[f(X)]$ and $E[g(Y)]$. This interpretation is further discussed in Section 3.

It is interesting to ask whether Kuznetsov's condition can be used as a tool for the construction of sets of probability measures. The idea is to focus on the largest set of probability measures that can be constructed in any given model using Kuznetsov's condition. These sets are called Kuznetsov's extensions. The paper explores the geometry of the resulting sets and shows that Kuznetsov's extensions are *not* identical to existing extensions. In particular, while strong independence is based on decomposition in the space of probability measures, Kuznetsov's condition leads to a decomposition on the space of desirable and almost-desirable decisions. Kuznetsov's condition emerges from this analysis as a new and potentially rich method for the construction of sets of probability measures from independence judgements.

2 Credal sets and their extensions

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 $p(X|y)$ denotes the conditional density of X conditional on the event $\{Y = y\}$. Given a function $f(X)$, $E_p[f(X)]$ denotes the expectation of function $f(X)$ with respect to $p(X)$, and $E_p[f(X)|y]$ denotes the expectation of function $f(X)$ with respect to $p(X|y)$.

A set of probability measures is called a *credal set* [7]. A credal set defined by a collection of densities $p(X)$ is denoted by $K(X)$. Given a credal set $K(X)$ and a function $f(X)$, the *lower* and *upper* expectations of $f(X)$ are defined respectively as $\underline{E}[f(X)] = \min_{p(X)\in K(X)} E_p[f(X)]$ 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 respectively as $\underline{P}(A) = \min_{p(X)\in K(X)} P(A)$ and $\overline{P}(A) = \max_{p(X)\in K(X)} P(A)$. A set of probability measures and its convex hull produce the same lower and upper expectations.

There are several concepts of independence that can be applied to credal sets [3, 5]. Two concepts that have an intuitive appeal and interesting properties are *epistemic independence* and *strong independence*.

Epistemic independence is based on the concept of *epistemic irrelevance*:

Definition 1 ([8]) *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.*

Alternatively, Y is epistemically irrelevant to X if $\underline{E}[f(X)|y] = \underline{E}[f(X)]$ for any function $f(X)$ and all possible values of Y.

Definition 2 ([8]) *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 [2, 3, 4]:

Definition 3 *Variables X and Y are* strongly independent *when every extreme point of $K(X,Y)$ satisfies standard stochastic independence of X and Y (that is, when $p(X,Y) = p(X)\,p(Y)$).*

This definition corresponds to the concept of *independence on the selection* by Couso et al [3]; Couso et al also use the term *strong independence* to indicate the strong extensions of this paper.

Given 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 $K(Y)$. It is reasonable to consider, for every model, the largest possible credal set allowed by the model as the *natural extension* of that model (following terminology by Walley [8, 9]).

Consider two examples of extensions. Suppose $K(X)$ and $K(Y)$ are given. The *independent natural extension* is the largest joint credal set that satisfies epistemic independence with marginals $K(X)$ and $K(Y)$ [8]. The *strong extension* is the largest joint credal set that satisfies strong independence with the marginals $K(X)$ and $K(Y)$ [3, 4]. The strong extension is intuitively the "product" of the sets $K(X)$ and $K(Y)$: Every vertex of $K(X)$ is combined with every vertex of $K(Y)$, and vice-versa — the combinations occur "independently" of both X and Y.

3 Kuznetsov's condition

Kuznetsov's condition is a requirement on the lower expectations of independent variables.[1] The condition can be easily understood in terms of expectation intervals. Note first that a credal set $K(X,Y)$ induces an expectation interval for every function $f(X,Y)$, namely, $\left[\underline{E}[f(X,Y)]\,,\overline{E}[f(X,Y)]\right]$. Denote this interval by $\overline{\underline{E}}[f(X,Y)]$.

Now, given that X and Y are deemed independent, what can we expect from expectation intervals $\overline{\underline{E}}[f(X,Y)]$? We could start by looking at functions of the form $f(X) + g(Y)$, requiring that

$$\overline{\underline{E}}[f(X) + g(Y)] = \overline{\underline{E}}[f(X)] + \overline{\underline{E}}[g(Y)], \qquad (2)$$

where the plus symbol denotes interval summation. The idea is that there is no dependency between the intervals for $f(X)$ and for $g(Y)$. But condition (2) is quite weak: standard probability measures always satisfy this condition, regardless of independence of X and Y. We can then impose a similar condition on products:

$$\overline{\underline{E}}[f(X)g(Y)] = \overline{\underline{E}}[f(X)] \times \overline{\underline{E}}[g(Y)], \qquad (3)$$

where the $\times$ symbol denotes interval multiplication. Now, condition (3) is equivalent to condition (1). Note that condition (3) does for expectation intervals what strong extensions do for sets of probability measures. The condition constructs joint intervals by combination of marginal intervals — and the combinations occur "independently" of both X and Y.

A situation where Kuznetsov's condition may arise naturally is as follows. Suppose that one defines a model for beliefs for variables X and Y using probability intervals. Variables X and Y are then declared to be independent — since the expectations depending on X and Y are all obtained manipulating intervals, it may make sense to declare that "independence" of X and Y means exactly the condition proposed by Kuznetsov.

4 Kuznetsov's extension

It seems that Kuznetsov regarded condition (1) as a test for independence [6]. That is, if $K(X,Y)$ is constructed and seen to verify (1), then X and Y are declared independent.

A more interesting and profitable strategy is to use Kuznetsov's condition to construct credal sets, not just to analyze them. Say that X and Y are Kuznetsov-independent if $K(X,Y)$ satisfies Kuznetsov's condition. Suppose we have $K(X)$ and $K(Y)$, and we obtain the information that X and Y are Kuznetsov-independent without further information on $K(X,Y)$. What can we say about $K(X,Y)$?

A reasonable strategy is to focus on the largest joint credal set that satisfies Kuznetsov's condition and has

[1]The condition was proposed by the Russian researcher V. P. Kuznetsov. Unfortunately, Kuznetsov died in an accident at the beginning of 1998. Most of Kuznetsov's work is contained in the book *Interval Statistical Models*, published in 1991 in Russian and not easily accessible at the moment [6].

	$p(x_0, y_0)$	$p(x_0, y_1)$	$p(x_1, y_0)$	$p(x_1, y_1)$
p_1	$1/4$	$1/4$	$1/4$	$1/4$
p_2	$9/25$	$6/25$	$6/25$	$4/25$

Table 1: The densities $p_1(X, Y)$ and $p_2(X, Y)$ in Example 1.

the marginals $K(X)$ and $K(Y)$. This set is referred to as *Kuznetsov's extension* of $K(X)$ and $K(Y)$.

A Kuznetsov's extension always exists. To see that, note that if $K_1(X, Y)$ and $K_2(X, Y)$ satisfy Kuznetsov's condition and have the correct marginal credal sets, then $\underline{E}_1[f(X)g(Y)]$ and $\underline{E}_2[f(X)g(Y)]$ are equal ($\underline{E}_i[f(X)g(Y)]$ is the lower expectation of $f(X)g(Y)$ with respect to $K_i(X, Y)$). Then the union of $K_1(X, Y)$ and $K_2(X, Y)$ satisfies Kuznetsov's condition because:

$$\begin{aligned}\underline{E}[f(X)g(Y)] &= \min\left(\underline{E}_1[f(X)g(Y)], \underline{E}_2[f(X)g(Y)]\right) \\ &= \underline{E}_1[f(X)g(Y)].\end{aligned}$$

5 Relationships among independence concepts and extensions

Kuznetsov's condition is not equivalent to strong independence as stated in Definition 3, as shown by the next example (inspired by Walley's example [8, 9.3.4]).

Example 1 A binary variable X has values $\{x_0, x_1\}$ and $p(x_1)$ belongs to the interval $[2/5, 1/2]$. Another binary variable Y has values $\{y_0, y_1\}$ and $p(y_1)$ belongs to the interval $[2/5, 1/2]$. Additionally, X and Y are declared strongly independent. Suppose we create an extension with the two vertices $p_1(X, Y)$ and $p_2(X, Y)$ in Table 1.

The density $p_1(X, Y)$ is obtained by multiplying the densities $p(X) = [p(x_0), p(x_1)] = [1/2, 1/2]$ and $q(Y) = [q(y_0), q(y_1)] = [1/2, 1/2]$. The density $p_2(X, Y)$ is obtained by multiplying $p(X) = [3/5, 2/5]$ and $q(Y) = [3/5, 2/5]$. The joint credal set with vertices $p_1(X, Y)$ and $p_2(X, Y)$ satisfies strong independence. Note that the joint credal set is not the strong extension. As an aside, note that the joint credal set satisfies the *repetition independence* concept of Couso et al [3]: each vertex of the credal set is the repeated product of a density.

But the credal set $\{p_1(X, Y), p_2(X, Y)\}$ does not satisfy Kuznetsov's condition. Consider the function $h(X, Y) = [h(x_0, y_0), h(x_0, y_1), h(x_1, y_0), h(x_1, y_1)] = [3, 6, 1, 2]$. This function is the product of functions $f(X) = [f(x_0), f(x_1)] = [3, 1]$ and $g(Y) = [g(y_0), g(y_1)] = [1, 2]$. We have: $\underline{E}[f(X)] = 2$; $\overline{E}[f(X)] = 11/5$; $\underline{E}[g(Y)] = 7/5$; $\overline{E}[g(Y)] = 3/2$. Kuznetsov's condition requires that $\overline{E}[f(X)g(Y)] = \max(2 \times 7/5, 2 \times 3/2, 11/5 \times 7/5, 11/5 \times 3/2) = 33/10 = 3.3$. But $E_{p_1}[h(X, Y)] = 3$ and $E_{p_2}[h(X, Y)] = 77/25$, so $\overline{E}[h(X, Y)] = 77/25 = 3.08$ for the credal set $\{p_1(X, Y), p_2(X, Y)\}$. □

Even though Kuznetsov's condition is not equivalent to strong independence, a strong extension *always* satisfies Kuznetsov's condition. Because the lower expectation of a function is attained at a vertex of a credal set [7, 8], strong extensions have $\underline{E}[f(X)g(Y)] = \min \sum_{X,Y} f(X)g(Y)p(X)p(Y)$, and this leads to condition (1).

Consequently, the strong extension is always smaller than or equal to (in the sense of set inclusion) the Kuznetsov's extension of given sets $K(X)$ and $K(Y)$. Using this result, we can easily prove:

Theorem 1 *Kuznetsov's condition implies the summation independence condition (2).*

Proof. For any function $f(X) + g(Y)$, and any joint credal set $K(X, Y)$, $\underline{E}[f(X) + g(Y)] \geq \underline{E}[f(X)] + \underline{E}[g(Y)]$ (where lower expectations are computed with respect to the given credal set). If the joint credal set is the strong extension of $K(X)$ and $K(Y)$, then $\underline{E}[f(X) + g(Y)] = \underline{E}[f(X)] + \underline{E}[g(Y)]$. Now suppose the joint credal set is Kuznetsov's extension of $K(X)$ and $K(Y)$. Then the joint credal set must be larger than or equal to the strong extension, and consequently we must have $\underline{E}[f(X) + g(Y)] \leq \underline{E}[f(X)] + \underline{E}[g(Y)]$ for a Kuznetsov's extension. As a result, the equality holds for Kuznetsov's extensions. □

Theorem 1 shows that Kuznetsov's condition guarantees "interval-decomposition" for all basic arithmetic operations involving expectation intervals.

Kuznetsov's extensions are possibly larger than strong extensions, but how do Kuznetsov's extensions compare with independent natural extensions? It turns out that Kuznetsov's extensions are smaller than independent natural extensions because Kuznetsov's independence implies epistemic independence. To see that, take the Generalized Bayes Rule, an equation that must be satisfied by any credal set [8]:

$$\underline{E}[(f(X) - \underline{E}[f(X)|y])I_y(Y)] = 0.$$

(The symbol $I_y(Y)$ denotes the indicator function for the event $\{Y = y\}$.) The function $(f(X) - \underline{E}[f(X)|y])I_y(Y)$ is decomposable, so Kuznetsov's condition leads to:

$$\underline{E}[(f(X) - \underline{E}[f(X)|y])]\,\underline{P}(Y = y) = 0.$$

	$p(x_0, y_0)$	$p(x_0, y_1)$	$p(x_1, y_0)$	$p(x_1, y_1)$
p_3	3/10	3/10	1/5	1/5
p_4	3/10	1/5	3/10	1/5
p_5	1/3	2/9	2/9	2/9
p_6	3/11	3/11	3/11	2/11

Table 2: The densities $p_3(X, Y)$ to $p_6(X, Y)$ in Example 2 (from [8]).

When $\underline{P}(Y = y) > 0$, we obtain $\underline{E}[f(X)|y] = \underline{E}[f(X)]$. Epistemic independence is then implied by Kuznetsov's condition, unless $\underline{P}(Y = y) = 0$ (in which case further constraints on $K(f(X)|y)$ may lead to $\underline{E}[f(X)|y] \neq \underline{E}[f(X)]$ [10]). The consequence of these facts is that Kuznetsov's extensions are smaller or equal to independent natural extensions in models with positive lower probability. The next example shows that in general Kuznetsov's extensions are strictly smaller than independent natural extensions.

Example 2 Consider the same binary variables and the same credal sets $K(X)$ and $K(Y)$ as in Example 1. The independent natural extension of $K(X)$ and $K(Y)$ has six vertices, as shown by Walley [8, 9.3.4]. The vertices are $\{p_1, \ldots, p_6\}$, shown in Tables 1 and 2.

Now consider the function $h(X, Y) = [1, 0, 0, 1]$. This function can be written as $(f(X)g(Y) + 1/2)$, where $f(X) = [1/2, -1/2]$ and $g(Y) = [1, -1]$. We have $\underline{E}[f(X)] = 0$; $\overline{E}[f(X)] = 1/10$; $\underline{E}[g(Y)] = 0$; $\overline{E}[g(Y)] = 1/5$. Kuznetsov's condition demands that $\underline{E}[h(X, Y)] = \min(0 \times 0, 0 \times 1/5, 1/10 \times 0, 1/10 \times 1/5) + 1/2 = 1/2$. But $E_{p_6}[h(X, Y)] = 5/11 < 1/2$, violating Kuznetsov's condition. □

6 Are Kuznetzov and strong extensions identical?

An attractive property of Kuznetsov's condition is that it is stated solely in terms of lower expectations. Because lower expectations are the ultimate guides for decision-making, it is convenient to base any concept of independence on lower expectations. Most arguments for strong independence are entirely based on properties of individual probabilities in the joint credal set [4], a significant disadvantage of the concept.[2] So we are led to the question: Are Kuznetsov's extensions equivalent to strong extensions, so that Kuznetsov's condition leads to strong independence? This section examines this question and related topics.

Several arguments regarding Kuznetsov's and strong extensions are best understood through geometric images. Denote by $\mathcal{U}(X, Y)$ the hyperplane $\sum_{X,Y} p(X, Y) = 1$. Every function $h(X, Y)$ defines an hyperplane in the space of probability densities $p(X, Y)$. This hyperplane, denoted by $\mathcal{H}(h(X, Y))$, is defined as $h(X, Y) \cdot p(X, Y) = \underline{E}[h(X, Y)]$. The symbol $\cdot$ indicates dot product; the symbols $h(X, Y)$ and $p(X, Y)$ are understood (abusing notation) as vectors containing the values of the related functions. The hyperplane contains the origin and extends through the region of positive $p(X, Y)$, eventually intersecting $\mathcal{U}(X, Y)$. The hyperplane can then be written as $(h(X, Y) - \underline{E}[h(X, Y)]) \cdot p(X, Y) = 0$. Note that any credal set $K(X, Y)$ belongs to the hyperplane $\mathcal{U}(X, Y)$; more pictorially, any credal set is "drawn" on a flat surface provided by $\mathcal{U}(X, Y)$. This "drawing" is made by carving the hyperplane with other hyperplanes — the hyperplanes $\mathcal{H}(h(X, Y))$ that go through the origin and intersect $\mathcal{U}(X, Y)$.

To illustrate these ideas, take Example 2. There the strategy was to create a function $h(X, Y)$ that is decomposable and whose hyperplane $\mathcal{H}(h(X, Y))$ "cuts" the independent natural extension.

It is interesting to note that, in 4 dimensions, it is possible to write a vector $[\theta_1, \theta_2, \theta_3, \theta_4]$ in the form $[\beta_1\gamma_1, \beta_1\gamma_2, \beta_2\gamma_1, \beta_2\gamma_2] - \alpha$, provided that $\theta_1 + \theta_4 \neq \theta_2 + \theta_3$. Such a decomposition suggests that Kuznetsov's extensions and strong extensions are identical in binary models for the following reason.[3] Take a function $h(X, Y)$ that defines a supporting hyperplane of the Kuznetsov's extension. If $h(X, Y)$ can be written as a product of functions of X and Y, the lower expectation $\underline{E}[h(X, Y)]$ is specified by Kuznetsov's condition *and* it must be the lower expectation obtained using the strong extension. So, the hyperplane defined by $h(X, Y)$ must support the strong extension. This conclusion can be proved more formally:

Theorem 2 *If X and Y are binary variables that are Kuznetsov-independent, the Kuznetsov's extension of $K(X)$ and $K(Y)$ is identical to the strong extension of $K(X)$ and $K(Y)$.*

Proof. The strong extension $K_s(X, Y)$ of a binary model has at most four vertices (there are one or two vertices for $K(X)$, and one or two vertices for $K(Y)$). If both $K(X)$ and $K(Y)$ are singletons, then

[2] Ongoing work by Moral and Cano has found arguments for strong independence based on combination of lower expectations.

[3] This observation, with the implication that Kuznetsov's extensions are strong extensions in binary models, was brought to my attention by Peter Walley.

there is a single joint probability density $p(X,Y)$ that satisfies strong independence and Kuznetsov's condition. In all other cases, it is necessary to construct a number of supporting hyperplanes that define $K_s(X,Y)$. Each one of these hyperplanes is defined by three points, either vertices of the strong extension or suitably selected points. Take a vertex $p(X)$ of $K(X)$ and a vertex $p(Y)$ of $K(Y)$. Each one of the points that define the supporting hyperplanes can be written as $[p_i q_i, p_i q_i^c, p_i^c q_i, p_i^c q_i^c]$, where $p_i = p_i(x_0)$, $p_i^c = p_i(x_1)$, $q_i = q_i(y_0)$ and $q_i^c = q_i(y_1)$. Consider the hyperplane $\mathcal{H}(h(X,Y))$ containing a point defined by (p_1, q_1), a point defined by (p_1, q_2), and a point defined by (p_2, q_2). Note that p_1 appears in the first two points and q_2 appears in the last two points — such an arrangement must necessarily be possible. The hyperplane is of the form $\theta_1 p(x_0,y_0) + \theta_2 p(x_0,y_1) + \theta_3 p(x_1,y_0) + \theta_4 p(x_1,y_1) = 0$, so:

$$
\begin{aligned}
\theta_1 p_1 q_1 + \theta_2 p_1 q_1^c + \theta_3 p_1^c q_1 + \theta_4 p_1^c q_1^c = 0,\\
\theta_1 p_1 q_2 + \theta_2 p_1 q_2^c + \theta_3 p_1^c q_2 + \theta_4 p_1^c q_2^c = 0,\\
\theta_1 p_2 q_2 + \theta_2 p_2 q_2^c + \theta_3 p_2^c q_2 + \theta_4 p_2^c q_2^c = 0.
\end{aligned}
$$

The function $h(X,Y)$ is defined by any vector $[\theta_1, \theta_2, \theta_3, \theta_4]$ that satisfies this linear system. Now assume a solution of the form $\theta_1 = \beta_1\gamma_1$, $\theta_2 = \beta_1\gamma_2$, $\theta_3 = \beta_2\gamma_1$, and $\theta_4 = \beta_2\gamma_2$. Rearranging terms, we obtain:

$$
\begin{aligned}
(\beta_1 p_1 + \beta_2 p_1^c)(\gamma_1 q_1 + \gamma_2 q_1^c) = 0,\\
(\beta_1 p_1 + \beta_2 p_1^c)(\gamma_1 q_2 + \gamma_2 q_2^c) = 0,\\
(\beta_1 p_2 + \beta_2 p_2^c)(\gamma_1 q_2 + \gamma_2 q_2^c) = 0.
\end{aligned}
$$

It is apparent that these equations can indeed have a decomposable solution by selecting $[\beta_1, \beta_2]$ as a function that has zero expectation under $p_1(X)$ and $[\gamma_1, \gamma_2]$ as a function that has zero expectation under $q_2(Y)$. Such functions always exist as they are needed to generate $K(X)$ and $K(Y)$. The same operations can be reproduced for any supporting hyperplane of the strong extension. The result is that the strong extension is entirely specified in terms of decomposable functions. Now, each one of these "decomposable supporting hyperplanes" must also be a supporting hyperplane for the Kuznetsov's extension — as imposed by Kuznetsov's condition on decomposable functions. Consequently, the Kuznetsov's extension must be smaller or equal to the strong extension. Because the strong extension must be smaller or equal to the Kuznetsov's extension, the extensions must be equal. $\square$

The proof of Theorem 2 indicates that the supporting hyperplanes of the strong extension (in the binary case) are of the form $\mathcal{H}(f(X)g(Y))$, where $f(X)$ and $g(Y)$ generate supporting hyperplanes for $K(X)$ and $K(Y)$ respectively. If the strong extension contains 4 non-coplanar vertices, then the four facets of the strong extension must be generated by all combinations of supporting hyperplanes from $K(X)$ and $K(Y)$. Table 3 indicates the structure of the situation: there are four facets, each containing three vertices; for each facet, there is a single combination of marginal supporting hyperplanes that satisfies the joint supporting hyperplane. Note that the hyperplanes of the joint credal set decompose in the hyperplanes of the marginal credal sets. This surprising fact is illustrated by the next example.

Example 3 Consider the same binary variables and the same credal sets as in Example 2. The marginal credal set $K(X)$ is defined by the intersection of two halfspaces:

$$
\begin{aligned}
p(x_0) - p(x_1) \geq 0,\\
-2p(x_0) + 3p(x_1) \geq 0.
\end{aligned}
$$

If we take $f_1(X) = [f_1(x_0), f_1(x_1)] = [1, -1]$ and $f_2(X) = [f_2(x_0), f_2(x_1)] = [-2, 3]$, we can write:

$$\sum_X f_1(X)p(X) \geq 0, \quad \sum_X f_2(X)p(X) \geq 0.$$

The marginal credal set $K(Y)$ is defined by similar halfspaces:

$$
\begin{aligned}
q(y_0) - q(y_1) \geq 0,\\
-2q(y_0) + 3q(y_1) \geq 0.
\end{aligned}
$$

Again, if we take $g_1(Y) = [g_1(y_0), g_1(y_1)] = [1, -1]$ and $g_2(Y) = [g_2(y_0), g_2(y_1)] = [-2, 3]$, we can write:

$$\sum_Y g_1(Y)q(Y) \geq 0, \quad \sum_Y g_2(Y)q(Y) \geq 0.$$

The strong extension $K_s(X,Y)$ has the four vertices $\{p_1, \ldots, p_4\}$, shown in Tables 1 and 2. The hyperplanes that support the strong extension are:[4]

$$
\begin{aligned}
-2p(x_0,y_0) + 2p(x_0,y_1) + 3p(x_1,y_0) - 3p(x_1,y_1) = 0\\
-2p(x_0,y_0) + 3p(x_0,y_1) - 2p(x_1,y_0) - 3p(x_1,y_1) = 0\\
4p(x_0,y_0) - 6p(x_0,y_1) - 6p(x_1,y_0) + 9p(x_1,y_1) = 0\\
p(x_0,y_0) - p(x_0,y_1) - p(x_1,y_0) + p(x_1,y_1) = 0
\end{aligned}
$$

These hyperplanes are of the form $\sum_{X,Y}(f_i(X)g_i(X))p(X,Y) = 0$, for $i,j \in \{1,2\}$. $\square$

The proof of Theorem 2 relies on the fact that Kuznetsov's condition only deals with decomposable

[4] Supporting hyperplanes can be computed from vertices (and vice-versa) using the software lrs, freely distributed by David Avis.

Facets:	1	2	3	4
Combinations of marginal vertices in the facet:	$p_1(X), q_1(Y)$ $p_1(X), q_2(Y)$ $p_2(X), q_1(Y)$	$p_1(X), q_1(Y)$ $p_1(X), q_2(Y)$ $p_2(X), q_2(Y)$	$p_1(X), q_1(Y)$ $p_2(X), q_1(Y)$ $p_2(X), q_2(Y)$	$p_1(X), q_2(Y)$ $p_2(X), q_1(Y)$ $p_2(X), q_2(Y)$
Combinations of marginal supporting hyperplanes:	$f_1(X), g_1(Y)$	$f_1(X), g_2(Y)$	$f_2(X), g_1(Y)$	$f_2(X), g_2(Y)$

Table 3: Facets in the strong extension of binary variables X and Y:
•$K(X)$ with vertices $\{p_1(X), p_2(X)\}$ and supporting hyperplanes $\{\sum_X f_1(X)p(X) = 0, \sum_X f_2(X)p(X) = 0\}$;
•$K(Y)$ with vertices $\{q_1(Y), q_2(Y)\}$ and supporting hyperplanes $\{\sum_Y g_1(Y)q(Y) = 0, \sum_Y g_2(Y)q(Y) = 0\}$.

functions. So, *if* strong extensions were fully specified by decomposable functions, *then* Kuznetsov's condition would be enough to specify the lower expectation of each one of these functions. And then, Kuznetsov's extensions would be equal to or smaller than strong extensions — with the implication that both extensions would be equal. A geometric picture of the situation is that we must carve a region of $\mathcal{U}(X,Y)$ by using a special chisel, one that can only produce decomposable hyperplanes. Anything that can be carved only with decomposable hyperplanes complies with Kuznetsov's condition. If we can carve strong extensions with this special chisel, then strong extensions and Kuznetsov's extensions are identical. So we must ask: Are strong extensions fully defined by decomposable functions? The answer is negative, as shown by the next example.

Example 4 Consider variables X and Y, which have three values and are declared independent. The marginal credal sets $K(X)$ and $K(Y)$ were obtained by generating random triplets and normalizing them. All operations were conducted using the Maple^{TM} symbolic algebra system and the software lrs, in multiple-precision integer arithmetic. Table 4 shows the vertices of each marginal credal set. Each one of these credal sets is a polygon "drawn" in the hyperplane $\sum_{X,Y} p(X,Y) = 1$. Figure 1 shows both credal sets. In clockwise manner, the vertices of $K(X)$ are connected as $p_1(X)-p_2(X)-p_3(X)-p_6(X)-p_5(X)-p_4(X)$. Likewise, the connectivity structure of $K(Y)$ is $q_1(Y)-q_2(Y)-q_3(Y)-q_4(Y)-q_7(Y)-q_6(Y)-q_5(Y)$.

The strong extension $K_s(X,Y)$ has 42 vertices, obtained by multiplying vertices of $K(X)$ and vertices of $K(Y)$. Using the software lrs, it is possible to obtain the supporting hyperplanes for the strong extension. Several of these hyperplanes are decomposable, but not all of them. One of these non-decomposable hyperplanes is the hyperplane $\mathcal{H}'(X,Y)$:

$$\begin{aligned}
(-245229852636038981951)\ p(x_0,y_0) &\ + \\
(-568204502977532539551)\ p(x_0,y_1) &\ + \\
292282410309769330349\ p(x_0,y_2) &\ + \\
246491733322506835749\ p(x_1,y_0) &\ + \\
205737250272154798149\ p(x_1,y_1) &\ + \\
(-186724207313882978195 1)\ p(x_1,y_2) &\ + \\
(-262219178342211417951)\ p(x_2,y_0) &\ + \\
930634499710831024449\ p(x_2,y_1) &\ + \\
150696025699440289434 9\ p(x_2,y_2) &\ = 0
\end{aligned}$$

This hyperplane contains the origin and the vertices $(p_1(X)q_7(Y))$, $(p_2(X)q_7(Y))$, $(p_3(X)q_6(Y))$, $(p_4(X)q_4(Y))$, $(p_5(X)q_1(Y))$, $(p_5(X)q_5(Y))$, $(p_6(X)q_5(Y))$, $(p_6(X)q_6(Y))$.

What we must prove is that the hyperplane $\mathcal{H}'(X,Y)$ cannot be written as $\sum_{X,Y} f(X)g(Y)p(X,Y) = \alpha$. The parameter α allows us to consider families of parallel hyperplanes. Now, any hyperplane of the form $\theta_1 p(x_0,y_0) + \theta_2 p(x_0,y_1) + \theta_3 p(x_0,y_2) + \theta_4 p(x_1,y_0) + \theta_5 p(x_1,y_1) + \theta_6 p(x_1,y_2) + \theta_7 p(x_2,y_0) + \theta_8 p(x_2,y_1) + \theta_9 p(x_2,y_2) = 0$ can be written in decomposable form if and only if the following equalities are simultaneously true:

$$\frac{\theta_1-\alpha}{\theta_4-\alpha} = \frac{\theta_2-\alpha}{\theta_5-\alpha} = \frac{\theta_3-\alpha}{\theta_6-\alpha},$$

$$\frac{\theta_4-\alpha}{\theta_7-\alpha} = \frac{\theta_5-\alpha}{\theta_8-\alpha} = \frac{\theta_6-\alpha}{\theta_9-\alpha}.$$

These equalities cannot be simultaneously satisfied for the hyperplane $\mathcal{H}'(X,Y)$. For example, if we choose α such that $(\theta_1-\alpha)/(\theta_4-\alpha) = (\theta_2-\alpha)/(\theta_5-\alpha)$, then

$$\begin{aligned}
(\theta_1-\alpha)/(\theta_4-\alpha) &= 7.924886446 \\
(\theta_2-\alpha)/(\theta_5-\alpha) &= 7.924886446 \\
(\theta_3-\alpha)/(\theta_6-\alpha) &= -1.192509156 \\
(\theta_4-\alpha)/(\theta_7-\alpha) &= 0.1224867836 \\
(\theta_5-\alpha)/(\theta_8-\alpha) &= -0.1822802441 \\
(\theta_6-\alpha)/(\theta_9-\alpha) &= -1.836749901.
\end{aligned}$$

This demonstrates that the hyperplane $\mathcal{H}'(X,Y)$ cannot be produced only by application of Kuznetsov's condition. □

Vertices of $K(X)$: $[p(x_0), p(x_1), p(x_2)]$	Vertices of $K(Y)$: $[q(y_0), q(y_1), q(y_2)]$
$p_1 = [3/100, 43/100, 27/50]$	$q_1 = [9/100, 33/100, 29/50]$
$p_2 = [39/100, 9/50, 43/100]$	$q_2 = [139/500, 31/500, 33/50]$
$p_3 = [9/20, 19/50, 17/100]$	$q_3 = [401/1000, 19/1000, 29/50]$
$p_4 = [7/100, 47/100, 23/50]$	$q_4 = [7/10, 1/50, 7/25]$
$p_5 = [339/1000, 59/100, 71/1000]$	$q_5 = [1/10, 69/100, 21/100]$
$p_6 = [201/500, 11/20, 6/125]$	$q_6 = [59/100, 9/25, 1/20]$
	$q_7 = [861/1000, 49/1000, 9/100]$

Table 4: Vertices of $K(X)$ and $K(Y)$ in Example 4.

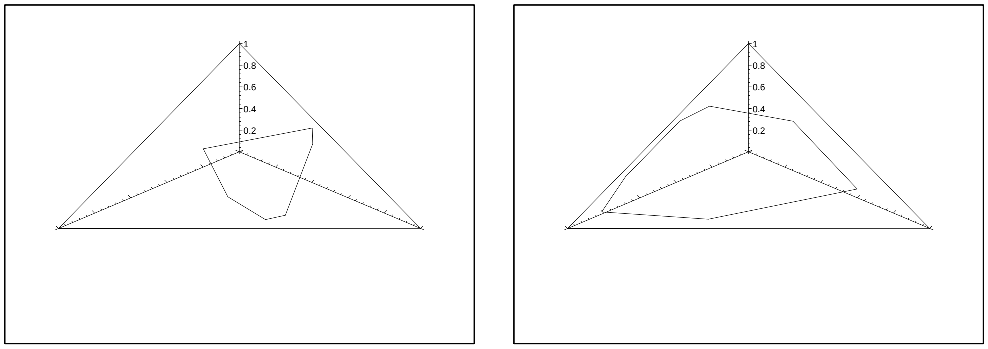

Figure 1: Credal sets $K(X)$ (left) and $K(Y)$ (right). Both sets belong to the hyperplane $\mathcal{U}(X, Y)$.

We cannot generate the necessary hyperplanes to make Kuznetsov's extensions identical to strong extensions: Kuznetsov's chisel is not flexible enough to carve the delicate structure of strong extensions.

7 The geometry of Kuznetsov's extensions

Given credal sets $K(X)$ and $K(Y)$, we can make a mental picture of the Kuznetsov's extension $K_k(X, Y)$: it is the smallest set that contains the strong extension $K_s(X, Y)$ with decomposable supporting hyperplanes. We can imagine the strong extension as a core that is "wrapped" using decomposable hyperplanes; the final package is the Kuznetsov's extension.

Note that for any vertex $p(X)q(Y)$ of a strong extension, a decomposable function $f(X)g(Y)$ can be constructed such that $\sum_{X,Y} f(X)g(Y)p(X)q(Y) = 0$ (it is enough to select either $f(X)$ or $g(Y)$ from the supporting hyperplanes of $K(X)$ or $K(Y)$). So, every vertex of the strong extension belongs to the boundary of the Kuznetsov's extension. We can go one step further and notice that for any edge of a strong extension, we can construct a decomposable hyperplane through the edge. For any edge goes from a point $p_1(X)q_1(Y)$ to a point $p_2(X)q_2(Y)$, and we can select $f(X)$ such that $\sum_X f(X)p_1(X) = 0$ and $g(Y)$ such that $\sum_Y g(Y)q_2(Y) = 0$. The hyperplane $\mathcal{H}(f(X)g(Y))$ then contains the points $p_1(X)q_1(Y)$ and $p_2(X)q_2(Y)$. Hence the discrepancies between a Kuznetsov's extension and the corresponding strong extension occur only in the non-decomposable facets of the strong extension.

Now, suppose we have $K(X)$ and $K(Y)$ and we must compute $\underline{E}[h(X, Y)]$ for a non-decomposable function $h(X, Y)$. Using Walley's theory of natural extensions, we can write [8, 9, 10]:

$$\underline{E}[h(X, Y)] = \max \mu,$$

subject to $\lambda_k \geq 0$ and

$$\sum_k \lambda_k \left(h_k(X, Y) - \underline{E}[h_k(X, Y)]\right) \leq h(X, Y) - \mu, \quad (4)$$

where $h_k(X, Y)$ represents a collection of functions with fixed lower expectations. There are N inequalities in (4), where N is the number of values of $\{X, Y\}$. (This linear program is the dual of a linear program on the space of probability measures.)

Kuznetsov's condition specifies the lower expectation for an infinite number of decomposable functions, so we cannot use inequalities (4) directly. Using a property of linear programs, we need only N decomposable functions $h_k(X,Y)$ to fully determine the values of the λ_k. We can take N functions $f_k(X)g_k(Y)$ and form the following non-linear program:

$$\underline{E}[h(X,Y)] = \max \mu,$$

subject to $\lambda_k \geq 0$, $f_k(X) \in K(X)$, $g_k(Y) \in K(Y)$ for $k = \{1,\ldots,N\}$, and

$$\sum_{k=1}^{N} \lambda_k \left(f_k(X)g_k(Y) - \underline{E}[f_k(X)g_k(Y)]\right) = h(X,Y) - \mu,$$

where $\underline{E}[f_k(X)g_k(Y)]$ is computed using Kuznetsov's condition. Note that here the $f_k(X)$ and $g_k(Y)$ must also be searched for, in addition to λ_k and μ.

8 Discussion

Given credal sets $K(X)$ and $K(Y)$, the Kuznetsov's extension $K_k(X,Y)$ fills a region of space between the strong extension and the independent natural extension. A Kuznetsov's extension can be viewed as a set that wraps a strong extension using decomposable hyperplanes. In fact, there is a subtle "duality" between these two extensions. While strong extensions rely on the decomposition of probability measures, Kuznetsov's extensions rely on the decomposition of supporting hyperplanes. Note that the space of probability measures and the space of hyperplanes are dual spaces [1, page 167], but decomposition of elements in one space does not lead to decomposition of elements in the other space.

Despite the simplicity of Kuznetsov's condition, it seems that Kuznetsov's extensions have complex structures — except in binary models. Derived concepts, such as conditional independence, should be investigated. A natural definition for a conditional Kuznetsov's condition is to require

$$\overline{\underline{E}}[f(X)g(Y)|A] = \overline{\underline{E}}[f(X)|A] \times \overline{\underline{E}}[g(Y)|A],$$

for any conditioning event A. 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.

Acknowledgements

I thank Peter Walley for bringing my attention to Kuznetsov's condition; Lev Utkin for discussing the condition and generously translating large pieces of Kuznetsov's book to English; Serafín Moral for discussing the condition and Kuznetsov's extension and sharing ongoing work on strong extensions; and Igor Kozine for translating additional material from Kuznetsov's book to English. Thanks also to David Avis for freely distributing the lrs package (used in the examples).

This work has been partially supported by CNPq through grant 300183/98-4(RN).

References

[1] D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization.* Athena Scientific, Belmont, Massachusetts, 1997.

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

[3] I. Couso, S. Moral, and P. Walley. Examples of independence for imprecise probabilities. In G. de Cooman, F. G. Cozman, S. Moral, and P. Walley, editors, *Proc. of the First Int. Symp. on Imprecise Probabilities and Their Applications*, pages 121–130, Ghent, Belgium, 1999. Imprecise Probabilities Project, Universiteit Gent.

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

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

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

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

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

[9] P. Walley. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83:1–58, 1996.

[10] P. Walley, R. Pelessoni, and P. Vicig. Direct algorithms for checking coherence and making inferences from conditional probability assessments. Technical report, Dipartimento di Matematica Applicata B. de Finetti, Università di Trieste, Trieste, Italy, 1999.