On the rank of $2 \times 2 \times 2$ probability tables

Iván Pérez perez@utia.cas.cz
Faculty of Mathematics and Physics, Charles University

Jiří Vomlel vomlel@utia.cas.cz
Institute of Information Theory and Automation, Czech Academy of Sciences

Abstract

Bayesian networks for real-world problems typically satisfy the property of positive monotonicity (in the context of educational testing, it is commonly assumed that answering correctly a question A increases the probability of answering correctly another question B). In this paper, we focus on the study of relations between positive monotonic influences on three-variable patterns and a family of $2 \times 2 \times 2$ tensors. In this study, we use the Kruskal polynomial, well-known in the psychometrics community, which is equivalent to Cayley's hyperdeterminant (homogeneous polynomial of degree 4 in the 8 entries of a $2 \times 2 \times 2$ tensor). It is known that when the Kruskal polynomial is positive, the rank of the tensor is two. We show that when a probability table associated with three random variables obeys the positive monotonicity property, its corresponding $2 \times 2 \times 2$ tensor has rank two. Moreover, it can be decomposed using only nonnegative tensors, which can each be given a statistical interpretation. We study two concepts of monotonicity in sets of three random variables, strong monotonicity (any two variables have a positive influence on the third one), and weak monotonicity (just one pair of variables that have a positive influence on the third one), and we give an example to show they do not coincide. Furthermore, we proved that the strong monotonicity property implies that the tensor rank is at most two. We also performed experiments with real data to test the monotonicity properties. The real datasets were formed by information from the Czech high school final exam from the years 2016 to 2022. These datasets are representative since the sample size (number of students) for each year is very large ($N > 10000$) and information comes from students of all regions of the Czech Republic. In this datasets, we observed that almost all $2 \times 2 \times 2$ tensors are monotone and all their corresponding $2 \times 2 \times 2$ tensors have nonnegative decomposition.

Keywords: Tensor rank; Conditional probability tables; Monotonicity; Educational testing.

1. Introduction

Tensors and their decompositions are applied in many domains, like statistics, data science, and machine learning. In this paper, we focus on the study of relations between a family of $2 \times 2 \times 2$ tensors and some qualitative (monotonic) influences on three-variable patterns. The rest of this paper is organized as follows. In the next section we present the definition of tensor and tensor rank, explaining in particular our interest on probability tensors of order three. In Section 3 we briefly review the concept of monotonicity properties in the context of Bayesian networks, and we described this properties in detail for two and three variables. For the case of three variables, we proved that the strong monotonicity property implies the tensor rank is at most two. In Section 4 we first describe the origin and details of the datasets used in our experiment, in a second moment, we describe the steps of our analysis, contrasting the results obtained from real data with randomly generated data. We end in Section 5 with some concluding remarks.

2. Tensor Rank

A tensor is a mapping $X : \mathbb{I} \to \mathbb{A}$, where $\mathbb{A} = \mathbb{R}$ or $\mathbb{A} = \mathbb{C}$, $\mathbb{I} = I_1 \times \ldots \times I_k$, k is a natural number called the order of tensor X, and $I_j, j = 1, \ldots, k$ are index sets. Typically, I_j are sets of integers of cardinality n_j. Then we can say that tensor X has dimensions $n_1, \ldots, n_k$. In this paper all index sets will be $\{0, 1\}$ and all considered tensors will be real valued and will have order three.

Tensor $X : \{0, 1\}^3 \to \mathbb{R}$ has rank one if it can be written as an outer product of vectors:

$$X = a \otimes b \otimes c \ .$$

The outer product is defined for all $(i, j, k) \in \{0, 1\}^3$ as

$$X_{i,j,k} = a_i \cdot b_j \cdot c_k \ ,$$

where $a = (a_0, a_1)$, $b = (b_0, b_1)$, and $c = (c_0, c_1)$, are real valued vectors. Each tensor can be decomposed as a linear combination of rank-one tensors:

$$X = \sum_{r=1}^{R} a^r \otimes b^r \otimes c^r \ ,$$

where for $r = 1, \ldots, R$ $a^r = (a^r_0, a^r_1)$, $b^r = (b^r_0, b^r_1)$, and $c^r = (c^r_0, c^r_1)$, are real valued vectors. The rank of a tensor X, denoted $rank(X)$, is the minimal R over all such decompositions (Kruskal, 1989).

Kruskal proved that typical ranks for 2×2×2 tensors are two and three, and performed numerical simulations to obtain the approximate values 0.79 and 0.21 for the probability of ranks two and three respectively. Bergqvist (Bergqvist, 2013) provided exact values, for a 2×2×2 tensor with elements from a standard normal distribution, the probability to be of rank two is $\pi/4$, and the probability to be of rank three is $1 - \pi/4$.

Ten Berge (Berge, 1991) showed that a necessary and sufficient condition for a 2×2×2 tensor having rank two is that the Kruskal's polynomial must be positive. Kruskal polynomial, well-known in the psychometrics community, is equivalent to Cayley's hyperdeterminant:

$$\begin{aligned} \Delta_{222} = {} & x_{000}^2 x_{111}^2 + x_{010}^2 x_{101}^2 + x_{001}^2 x_{110}^2 + x_{011}^2 x_{100}^2 \\ & +4 \left(x_{000} x_{011} x_{101} x_{110} + x_{001} x_{010} x_{100} x_{111} \right) \\ & -2 \left(\begin{array}{l} x_{000} x_{001} x_{110} x_{111} + x_{000} x_{010} x_{101} x_{111} \\ + x_{000} x_{011} x_{100} x_{111} + x_{001} x_{010} x_{101} x_{110} \\ + x_{001} x_{011} x_{100} x_{110} + x_{010} x_{011} x_{100} x_{101} \end{array} \right) \ . \end{aligned}$$

Probability tables can be understood as tensors with their values being from $[0, 1]$. We performed a numerical experiment, in which we created randomly one million of 2×2×2 tensors with elements drawn from a uniform distribution on $[0, 1]$. We identified that rank two and rank three are also typical in this context, the probabilities for a tensor to have rank two or three are approximately 0.84 and 0.16, respectively.

3. Monotonicity

In the context of Bayesian networks the concept of monotonicity properties has been discussed in the literature for a long time (Wellman, 1990) and (Druzdzel and Henrion, 1993). More recent papers in this topic are (Restificar and Dietterich, 2013), (Masegosa et al., 2016) and (Plajner and Vomlel, 2020).

A Bayesian network models the probabilistic influences between its variables. The concept of qualitative influence has been designed to describe these influences in a qualitative way (Wellman, 1990). A qualitative influence between two variables expresses how observing a value for the one variable affects the probability distribution for the other variable.

A positive qualitative influence of a variable B on a variable A along an arc $B \to A$ in the network means that the occurrence of B increases the probability of A occurring, assuming that the values of the other parents of A remain the same. It means that

$$P(A = 1|B = 1, \mathbf{c}) \quad \geq \quad P(A = 1|B = 0, \mathbf{c})$$

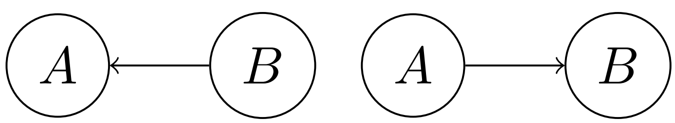
for any combination of values $\mathbf{c}$ for the set of parents of A other than B (Masegosa et al., 2016). The definition of negative influence between B and A along the arc $B \to A$ is analogous.

In the context of educational testing, a positive influence is commonly assumed since answering correctly a question A increases the probability of answering correctly another question B. In this manuscript we will refer to the positive influence (also called the positive monotonic influence).

We start our analysis in the simplest case (using patterns with two variables).

Definition 1 (Pairwise monotonicity) *Let $\boldsymbol{X}$ be a set of two binary random variables A and B. We say that $\boldsymbol{X}$ has the pairwise monotonicity property if B has a positive qualitative influence on A, and A has a positive qualitative influence on B, i.e., if*

$$P(A = 0|B = 0) \quad \geq \quad P(A = 0|B = 1) \tag{1}$$
$$P(B = 0|A = 0) \quad \geq \quad P(B = 0|A = 1) \ . \tag{2}$$



(a) Influence of B on A (b) Influence of A on B

Figure 1: Influence patterns on A and B

Remark 2 *Inequalities* (1) *and* (2) *are equivalent. To show this, it is enough to construct the joint probability table of A and B (see Table 1), substitute these values into the inequalities, and compute the conditional probabilities:*

$$P(A = 0|B = 0) \geq P(A = 0|B = 1) \quad \Leftrightarrow \quad \frac{x_{00}}{x_{00} + x_{10}} \geq \frac{x_{01}}{x_{01} + x_{11}} \quad \Leftrightarrow \quad x_{00}x_{11} \geq x_{01}x_{10}$$

$$P(B = 0|A = 0) \geq P(B = 0|A = 1) \quad \Leftrightarrow \quad \frac{x_{00}}{x_{00} + x_{01}} \geq \frac{x_{10}}{x_{10} + x_{11}} \quad \Leftrightarrow \quad x_{00}x_{11} \geq x_{01}x_{10} \ .$$

A	B	$P(A,B)$
0	0	x_{00}
0	1	x_{01}
1	0	x_{10}
1	1	x_{11}

Table 1: Probability table of A and B.

Then, a simple way to check whether $\boldsymbol{X}$ satisfies the pairwise monotonicity property is to rearrange its probability table in a 2×2 matrix and test whether its corresponding determinant is nonnegative:

$$\begin{vmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{vmatrix} \ = \ x_{00}x_{11} - x_{01}x_{10} \ \geq \ 0 \ .$$

This is equivalent to test whether $x_{00}x_{11}/x_{01}x_{10}$ is greater than 1. This ratio is known as *cross-product ratio* or *odds ratio* and is a well-known measure of association in 2×2 tables. Note that if we create a random 2×2 matrix with each entry element drawn from a uniform distribution on $[0,1]$ then the probability of having the pairwise monotonicity property (equivalent to nonnegative determinant) is only 0.5.

Now, we continue our analysis using patterns with three variables.

Definition 3 (Strong three-wise monotonicity) *Let $\boldsymbol{X}$ be a set of three random variables: A, B and C. We say that $\boldsymbol{X}$ satisfies the strong three-wise monotonicity property if any two variables have a positive influence on the third one, i.e. if*

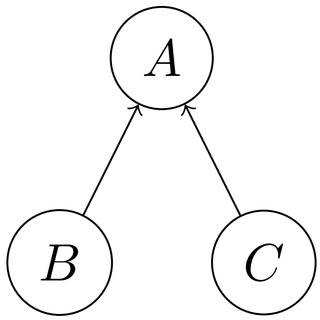
$$\begin{aligned} P(A=0|B,C=0) &\geq P(A=0|B,C=1) \\ P(A=0|B=0,C) &\geq P(A=0|B=1,C) \end{aligned}$$

$$\begin{aligned} P(B=0|A,C=0) &\geq P(B=0|A,C=1) \\ P(B=0|A=0,C) &\geq P(B=0|A=1,C) \end{aligned}$$

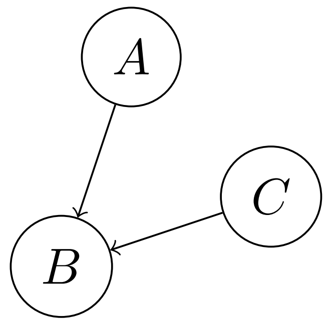
$$\begin{aligned} P(C=0|B=0,A) &\geq P(C=0|B=1,A) \\ P(C=0|B,A=0) &\geq P(C=0|B,A=1) \ . \end{aligned}$$

Considering that all variables are binary, these conditions translate to 12 inequalities. Analogously to the pairwise monotonicity, we can construct the joint probability table of A, B, and C (Table 2), substitute these values into the inequalities, and compute the conditional probabilities:

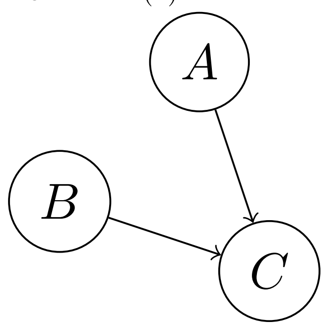
$$P(A=0|B=0,C=0) \geq P(A=0|B=1,C=0) \Leftrightarrow x_{000}x_{110} \geq x_{010}x_{100} \quad (3)$$
$$P(A=0|B=0,C=1) \geq P(A=0|B=1,C=1) \Leftrightarrow x_{001}x_{111} \geq x_{011}x_{101} \quad (4)$$
$$P(A=0|B=0,C=0) \geq P(A=0|B=0,C=1) \Leftrightarrow x_{000}x_{101} \geq x_{001}x_{100} \quad (5)$$
$$P(A=0|B=1,C=0) \geq P(A=0|B=1,C=1) \Leftrightarrow x_{010}x_{111} \geq x_{011}x_{110} \quad (6)$$



(a) Influence of B and C on A



(b) Influence of A and C on B



(c) Influence of A and B on C

Figure 2: Influence patterns between A, B and C

A	B	C	$P(A,B,C)$
0	0	0	x_{000}
0	0	1	x_{001}
0	1	0	x_{010}
0	1	1	x_{011}
1	0	0	x_{100}
1	0	1	x_{101}
1	1	0	x_{110}
1	1	1	x_{111}

Table 2: Probability table of A, B and C.

$$
\begin{array}{rclcrcll}
P(B=0|A=0,C=0) & \geq & P(B=0|A=1,C=0) & \Leftrightarrow & x_{000}x_{110} & \geq & x_{010}x_{100} & (7)\\
P(B=0|A=0,C=1) & \geq & P(B=0|A=1,C=1) & \Leftrightarrow & x_{001}x_{111} & \geq & x_{011}x_{101} & (8)\\
P(B=0|A=0,C=0) & \geq & P(B=0|A=0,C=1) & \Leftrightarrow & x_{000}x_{011} & \geq & x_{001}x_{010} & (9)\\
P(B=0|A=1,C=0) & \geq & P(B=0|A=1,C=1) & \Leftrightarrow & x_{100}x_{111} & \geq & x_{101}x_{110} & (10)
\end{array}
$$

$$
\begin{array}{rclcrcll}
P(C=0|B=0,A=0) & \geq & P(C=0|B=1,A=0) & \Leftrightarrow & x_{000}x_{011} & \geq & x_{001}x_{010} & (11)\\
P(C=0|B=0,A=1) & \geq & P(C=0|B=1,A=1) & \Leftrightarrow & x_{100}x_{111} & \geq & x_{101}x_{110} & (12)\\
P(C=0|B=0,A=0) & \geq & P(C=0|B=0,A=1) & \Leftrightarrow & x_{000}x_{101} & \geq & x_{001}x_{100} & (13)\\
P(C=0|B=1,A=0) & \geq & P(C=0|B=1,A=1) & \Leftrightarrow & x_{010}x_{111} & \geq & x_{011}x_{110}\ . & (14)
\end{array}
$$

It can be seen that the right inequalities come in pairs ((3) $\equiv$ (7), (4) $\equiv$ (8), (5) $\equiv$ (13), (6) $\equiv$ (14), (9) $\equiv$ (11), and (10) $\equiv$ (12)). Then, to test if $\boldsymbol{X}$ has the strong three-wise monotonicity property, it is enough to check only the next six inequalities:

$$
\begin{array}{rcll}
x_{000}x_{110} & \geq & x_{010}x_{100} & (15)\\
x_{001}x_{111} & \geq & x_{011}x_{101} & (16)\\
x_{000}x_{101} & \geq & x_{001}x_{100} & (17)\\
x_{010}x_{111} & \geq & x_{011}x_{110} & (18)\\
x_{000}x_{011} & \geq & x_{001}x_{010} & (19)\\
x_{100}x_{111} & \geq & x_{101}x_{110}\ . & (20)
\end{array}
$$

From the pairs of inequalities {(15),(18)}, {(17),(20)} and {(19),(16)} We can derive another three inequalities:

$$
\begin{array}{rcll}
x_{000}x_{111} & \geq & x_{011}x_{100} & (21)\\
x_{000}x_{111} & \geq & x_{001}x_{110} & (22)\\
x_{000}x_{111} & \geq & x_{010}x_{101}\ . & (23)
\end{array}
$$

These nine binomial inequalities (15) - (23) satisfy the supermodularity property described in the next definition based on (Allman et al., 2015).

Definition 4 *Let P be a $2\times2\times2$ tensor. Then P is π-supermodular if*

$$p_{i_1i_2i_3} \cdot p_{j_1j_2j_3} \quad \leq \quad p_{k_1k_2k_3} \cdot p_{l_1l_2l_3}$$

whenever $\{i_r, j_r\} = \{k_r, l_r\}$ and $\pi_r(k_r) \leq \pi_r(l_r)$ holds for $r = 1, 2, 3$. We call a tensor P supermodular if it is π-supermodular for some $\pi = (\pi_1, \pi_2, \pi_3)$ where π_i is a permutation of $(1, 2, 3)$.

Then, the $2\times2\times2$ tensor formed from Table 2 that satisfies the nine monotonic inequalities is supermodular (using $\pi = (id, id, id)$ where id stands for the identity permutation). The following lemma holds.

Lemma 5 *Let $\boldsymbol{X}$ be a set of three binary random variables. If $\boldsymbol{X}$ has the strong three-wise monotonicity property, then the nonnegative rank of the corresponding $2{\times}2{\times}2$ tensor is at most two.*

The proof is immediate from

Proposition 6 (Allman et al. (2015)) *Let P be a nonnegative $2{\times}2{\times}2$ tensor. Then P has nonnegative rank ≤ 2 if and only if P is supermodular.*

It is important to stress that the strong three-wise monotonicity property can be too restrictive for some applications. A relaxed definition we present next may be more appropriate for some situations.

Definition 7 (Weak three-wise monotonicity) *Let $\boldsymbol{X}$ be a set of three random variables: A, B and C. We say that $\boldsymbol{X}$ has the weak three-wise monotonicity property if there is just one pair of variables that have a positive influence on the third one.*

If $\boldsymbol{X}$ satisfies the weak three-wise monotonicity property, then one set of four inequalities (from the three sets of four inequalities presented above) is satisfied. From these twelve possible inequalities the following lemma is derived.

Lemma 8 *Let $\boldsymbol{X}$ be a set of three random variables: A, B and C. If $\boldsymbol{X}$ has the weak three-wise monotonicity property for two different configurations of variables, then $\boldsymbol{X}$ has the strong three-wise monotonicity property.*

This result comes from the fact that the union of any two sets of four inequalities equals the six inequalities required for the strong three-wise monotonicity property. Next we present an example of three variables with weak monotonicity property.

Example 1 *Consider the probability distribution of three variables A, B, and C presented in table 3.*

A	B	C	$P(A,B,C)$
0	0	0	$x_{000} = 0.20$
0	0	1	$x_{001} = 0.15$
0	1	0	$x_{010} = 0.05$
0	1	1	$x_{011} = 0.10$
1	0	0	$x_{100} = 0.10$
1	0	1	$x_{101} = 0.05$
1	1	0	$x_{110} = 0.20$
1	1	1	$x_{111} = 0.15$

Table 3: Probability distribution of A, B and C.

Inequalities (15), (16), (19) *and* (20) *are satisfied, while the inequalities* (17) *and* (18) *are not, direct substitutions leads to*

$$\begin{aligned} 0.20 \cdot 0.05 &< 0.15 \cdot 0.10 \quad \textit{and} \\ 0.05 \cdot 0.15 &< 0.10 \cdot 0.20 \ . \end{aligned}$$

Then $\boldsymbol{X} = \{A, B, C\}$ *satisfies the weak three-wise monotonicity property because* A *and* C *have a positive influence on* B*, but it does not have the strong three-wise monotonicity property because there is no positive influence between* A *and* C *(when* B *is fixed). Graphical representation of this example is in Figure 2b.*

We performed a numerical experiment by creating randomly one million of $2\times2\times2$ tensors with each tensor element drawn randomly from a uniform distribution on $[0, 1]$. We tested different monotonicity properties for the whole sample. Most tensors do not satisfy any monotonicity property (around 77.7%), about 21% satisfies the weak monotonicity, and just around 1.3% satisfy the strong monotonicity property.

4. A real data experiment

4.1 CERMAT datasets

The Ministry of Education, Youth and Sports of the Czech Republic, have established an experimental verification of knowledge and skills in secondary school mathematics according to the catalogue of requirements for the optional selective examination in mathematics. Complete information about the examination is published on the website of the Center for the Determination of Educational Results (CERMAT): maturita.cermat.cz.

The catalogue of requirements for the Mathematics exam contains nine main topics, their corresponding representation is specified by CERMAT. The percentages of subjects representation do not refer to the number of questions, but represent the proportion of points that can be obtained in the questions for the nine topics. (*Numerical sets*, *Algebraic expressions*, *Equations and inequalities*, *Functions*, *Sequences and series*, *Planimetry*, *Stereometry*, *Analytic geometry*, and *Combinatorics, probability and statistics)*.

The tests consists of about 30 questions (open and closed). A maximum of 50 points can be obtained in the test, half of which are for the open and half for the closed questions.

The datasets used in this study are publicly available on the statistical section of the CERMAT website: vysledky.cermat.cz/statistika. There are two evaluation periods: Spring and Autumn.

4.2 Experiment description

We are interested in analysis of the monotonic properties and tensor ranks studied in previous sections on real datasets. Our expectation was that the monotonic properties are more common in real contexts.

In this experiment we use all publicly available CERMAT datasets (2016-2022) corresponding to Spring periods because they are considerably bigger than Autumn periods. These datasets are representative since the sample size for each year is very large ($N > 10000$) and information comes from students of all regions of the Czech Republic. We binarized all the responses before performing our analysis, so the value 1 means that the question was answered correctly, and 0 means that the answer was incorrect. The data are described in Table 4.

First, we test pairwise monotonicity on all pairs of variables in each dataset, and found that all pairs have this property. As we mentioned before, creating a random model of two variables, the probability to have pairwise monotonicity property is just 0.5.

Year	N	Variables	Triplets	NND	NNND	P(NND)	SMonot	WMonot
2016	19695	29	3654	3649	5	0.9986	3649	5
2017	18545	29	3654	3650	4	0.9989	3650	4
2018	16957	30	4060	4018	42	0.9897	4018	42
2019	15702	30	4060	4029	31	0.9924	4029	30
2020	15020	29	3654	3653	1	0.9997	3653	1
2021	14456	28	3276	3265	11	0.9966	3265	11
2022	12709	30	4060	4041	19	0.9953	4041	19

Table 4: Czech high school final exam. We use the following abbreviations: N: Number of students, NND: Nonnegative Decomposition, NNND: No Nonnegative Decomposition, P(NND): Probability of NND, SMonot: Strong Monotonicity, WMonot: Weak Monotonicity.

Second, we compute the Kruskal polynomial for each 2×2×2 tensor generated by each possible triplet (in each dataset), and found that it was always positive, this implies that all of these tensors have rank two (Kruskal, 1989). As mentioned in section 2, we made a numerical experiment with one million of 2×2×2 tensors (elements drawn from a uniform distribution [0,1]) and found that around 84% of them were positive.

Third, due to the fact that all 2×2×2 tensors generated by each possible triplet have rank two, we use the algorithm proposed by (Berge, 1991) to decompose and analyze them in detail. Almost all of these tensors have nonnegative decomposition (99.57%), and moreover, all these tensors satisfy the strong monotonicity property (Columns NND and SMonot of Table 4, respectively). As we mentioned before, when 2×2×2 tensors were created randomly, just 1.3% of them satisfied the strong monotonicity property.

Fourth, regarding the tensors that do not have a nonnegative decomposition (0.43% of the whole sample), all except one of them satisfy the weak monotonicity property (Columns NNND and WMonot of Table 4, respectively). The exceptional case was found in the dataset corresponding to the year 2019, while 31 tensors are NNND, just 30 of them satisfy the weak monotonicity property. This case is described in detail in the next subsection.

4.3 An exceptional case

Among all possible triplets generated in our experiment (they add up 26418 in the seven years), just one triplet does not have any monotonicity property. This case corresponds to the triplet generated by questions (or tasks) *q9*, *q10* and *q28* in the dataset of year 2019. In this test, *q9* and *q10* are theoretically related in the test, they belong to the same task and are related with topic called *Functions*:

q9. Compute the coordinate a_2 of the point $A(4, a_2)$ of the function graph $y = \log_2 x$.

q10. Compute the coordinate b_1 of the point $B(b_1, -1)$ of the function graph $y = \log_2 x$.

On the other hand, question *q28* belongs to topic *Algebraic expressions*:

q28. The expression $\frac{12(a-2)^2}{12-6a}$ with real variable a is given. Which statement is true?

- For $a = 101^8$ the expression is positive.
- For $a = 2$, the value of the expression is 0.
- The value of the expression can never be zero.
- For all $a \neq \frac{1}{6}$ the expression is equal to $\frac{(a-2)^2}{1-6a}$.
- For some a, the expression is equal to $2(a-2)$.

The following array summarizes the data we obtained

$$\left[\begin{array}{cc|cc} u_{000} & u_{001} & u_{100} & u_{101} \\ u_{010} & u_{011} & u_{110} & u_{111} \end{array}\right] = \left[\begin{array}{cc|cc} 2579 & 966 & 978 & 359 \\ 172 & 57 & 5601 & 4990 \end{array}\right]$$

while its corresponding probability distribution is presented in table 5.

$q9$	$q10$	$q28$	$P(q9, q10, q28)$
0	0	0	$x_{000} = 0.1642$
0	0	1	$x_{001} = 0.0615$
0	1	0	$x_{010} = 0.0110$
0	1	1	$x_{011} = 0.0036$
1	0	0	$x_{100} = 0.0623$
1	0	1	$x_{101} = 0.0229$
1	1	0	$x_{110} = 0.3567$
1	1	1	$x_{111} = 0.3178$

Table 5: Probability distribution of $q9$, $q10$ and $q28$.

When we evaluate the six inequalities to test monotonicity property for the triplet (*q9*, *q10*, *q28*), we see that inequality (17) and (19) are not satisfied, direct computation shows that

$$0.1642 \cdot 0.0229 = \mathbf{0.00376} \quad < \quad \mathbf{0.00383} = 0.0615 \cdot 0.0623$$

that contradicts inequality (17), $x_{000}x_{101} \;\geq\; x_{001}x_{100}$, and

$$0.1642 \cdot 0.0036 = \mathbf{0.00059} \quad < \quad \mathbf{0.00067} = 0.0615 \cdot 0.0110$$

that contradicts inequality (19), $x_{000}x_{011} \;\geq\; x_{001}x_{010}$. Inequality (17) is required to have weak monotonicity on A (Figure 2a) and C (Figure 2c), while inequality (19) is required to have weak monotonicity on B (Figure 2b) and C (Figure 2c).

Analyzing the values that affect the inequalities that were not fulfilled to satisfy monotonicity conditions, we noticed that the differences between these values are relatively small, for this reason, we decided to slightly manipulate some values just to know "how far" this triplet is from fulfilling monotonicity conditions.

If we transfer 7 cases from state u_{001} to state u_{011}, inequality (19) is satisfied (without altering others), i.e., the triplet fulfills weak monotonicity. And if we also transfer 10 cases from state u_{100} to state u_{000}, inequality (17) is also satisfied, and in general all the inequalities to satisfy the conditions of strong monotonicity.

$q9'$	$q10'$	$q28'$	$P(q9', q10', q28')$
0	0	0	$x_{000} = 0.1649$
0	0	1	$x_{001} = 0.0611$
0	1	0	$x_{010} = 0.0110$
0	1	1	$x_{011} = 0.0041$
1	0	0	$x_{100} = 0.0616$
1	0	1	$x_{101} = 0.0229$
1	1	0	$x_{110} = 0.3567$
1	1	1	$x_{111} = 0.3178$

Table 6: Probability distribution of $q9'$, $q10'$ and $q28'$.

The probability distribution of this slight manipulation is presented in Table 6.

It can be noted that the values are almost the same as the original distribution presented in Table 5 (since only 17 cases were moved out of the 15702 that make up the dataset of year 2019), however this modified distribution meets strong monotonicity conditions.

In summary, the exceptional case found, indeed, does not meet monotonicity conditions (neither weak nor strong), but it is "not far" from it.

5. Concluding Remarks

In the context of 2×2×2 tensors, we proved that if a triplet has the strong monotonicity property, then the nonnegative rank of its corresponding tensor is at most two (Lemma 5).

In our experiment, we compute the value of the Kruskal polynomial for each 2×2×2 tensor generated by each possible triplet, and found that all of them were positive, this implies that all of these tensors have rank two.

We realized that all tensors with nonnegative decompositions correspond to triplets that have the strong monotonicity property. We performed numerical experiments to contrast our observation, we create randomly one million of 2×2×2 tensor (tensor elements drawn from a uniform distribution [0,1]), and decompose all of them that satisfied strong monotonicity property (around 1.3%). The same result was obtained, all of these decompositions were nonnegative. This observation allowed us to conjecture that if a 2×2×2 tensor is associated to a triplet that has the strong monotonicity property, then it has a nonnegative decomposition.

Finally, we showed that pairwise monotonicity property does not imply any type of three-wise monotonicity, the exceptional case described in section 4.3 works as an example due to all pairs in all datasets have the monotonicity property, particularly in the three possible pairs generated by the questions $q9$, $q10$, and $q28$, however, this triplet does not have any three-wise monotonicity property.

References

E. S. Allman, J. A. Rhodes, B. Sturmfels, and Z. Piotr. Tensors of nonnegative rank two. *Linear Algebra and its Applications*, 473:37–53, 2015.

J. T. Berge. Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays. *Psychometrika*, 56:631–636, 1991.

G. Bergqvist. Exact probabilities for typical ranks of $2 \times 2 \times 2$ and $3 \times 3 \times 2$ tensors. *Linear Algebra and its Applications*, 438, 2:663–667, 2013.

J. Druzdzel and M. Henrion. Efficient Reasoning in Qualitative Probabilistic Networks. *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 548–553, 1993.

J. B. Kruskal. Rank, Decomposition, and Uniqueness for 3-way and N-way Arrays. *Multiway Data Analysis*, 1989.

A. R. Masegosa, A. J. Feelders, and L. C. van der Gaag. Learning from Incomplete Data in Bayesian Networks with Qualitative Influences. *International Journal of Approximate Reasoning*, 69:18–34, 2016.

M. Plajner and J. Vomlel. Learning bipartite Bayesian networks under monotonicity restrictions. *International Journal of General Systems*, 49:1:88–111, 2020.

A. C. Restificar and T. G. Dietterich. Exploiting Monotonicity via Logistic Regression in Bayesian Network Learning. *Technical Report. Corvallis*, 2013.

M. P. Wellman. Fundamental Concepts of Qualitative Probabilistic Networks. *Artificial Intelligence*, 44 (3):257–303, 1990.