An inequality concerning the expected values of row-sum and column-sum products in Boolean matrices

J.B.Paris*and A.Vencovská†
School of Mathematics
The University of Manchester
Manchester M13 9PL

jeff.paris@manchester.ac.uk, alena.vencovska@manchester.ac.uk

August 4, 2015

Abstract

We give a proof of a matrix inequality and indicate how it can be applied to establish a natural analogy principle in Pure Inductive Logic.

Key words: Boolean Matrices, Row Sums, Inequalities, Counterpart Principle, Analogy, Inductive Logic.

Introduction

The aim of this paper is to give a proof of a matrix inequality which arises in investigations of the role of analogy in inductive logic. The background and specific

*Supported by a UK Engineering and Physical Sciences Research Council Research Grant.

†Supported by a UK Engineering and Physical Sciences Research Council Research Grant.

details can be found in [2], [3] respectively; here we merely briefly outline the context.

Consider an agent who aims to assigns probability values to sentences involving potentially infinite number of individuals $a_1, a_2, \ldots$ in a strictly rational manner. In other words, an agent who wishes to pick a rational probability functions on the set of sentences of some language, involving these individuals. It is usually assumed that in a situation when the agent knows nothing about these individuals, his/her probabilities should be independent of the particular choice of individuals featuring in a sentence. We call this property of probability functions Constant Exchangeability, Ex.

Amongst other arguably rational principles there is a principle of analogy, called the Counterpart Principle, CP, (see [1] or [2]), which says that the agent should not give less probability to a sentence in the event of receiving evidence consisting of that same sentence except that some individuals and/or relations in it are replaced with new ones. Under another very natural assumption of Pure Inductive Logic, the Principle of Language Invariance, LI, CP follows from Ex.

This leads to the question of whether allowing two sentences in the evidence, both of which are obtained by replacing some individuals and/or relations in the original sentence by different ones, should also not decrease the probability the agent gives to the sentence. A basic case of this is the question whether for some binary relation R of the language the conditional probability of $R(a_1, a_2)$ given both $R(a_1, a_3)$ and $R(a_4, a_2)$ must, under the assumption of Ex and LI, be at least as great as the unconditional probability of $R(a_1, a_2)$.

It turns out that this basic case is a key to a much more general result, and that to prove this basic case we need to show the following property of square matrices with 0,1 entries. Namely that the expected value of the product of the sum of entries in a row and in a column given that the entry at their intersection is 1, is bigger or equal to the expected value of the product of the sum of entries in a row and in a column given that the entry at their intersection is 0. We prove this result in the next section.

The Inequality

Let $(e_{i,j})$ be an $N \times N$ matrix with entries 0 or 1 (i.e. a Boolean matrix) and such that $N^2 > \sum_{i,j} e_{i,j} = T > 0$. Let $A_i = \sum_r e_{i,r}$, $B_j = \sum_s e_{s,j}$. Then the expected value of the product of the sum of entries in a row and in a column given that the entry at their intersection is 1 or 0 are,

$$\frac{\sum_{i,j} e_{i,j} A_i B_j}{T}, \qquad \frac{T^2 - \sum_{i,j} e_{i,j} A_i B_j}{N^2 - T}$$

respectively, since

$$\sum_{i,j} A_i B_j = T^2.$$

Hence the above claimed result holds by virtue of the following theorem:

Theorem 1. *Let $(e_{i,j})$ be an $N \times N$ $\{0,1\}$-matrix such that $\sum_{i,j} e_{i,j} = T > 0$. Then*

$$\sum_{i,j} e_{i,j} A_i B_j \geq T^3 N^{-2} \tag{1}$$

where $A_i = \sum_r e_{i,r}$, $B_j = \sum_s e_{s,j}$.

Proof. Let $M = T/N$. We may assume that $M \geq 1$ otherwise $(e_{i,j})$ will have to have a zero column and a zero row and we can simply remove these as they would only make the right hand side of (1) smaller.

It is enough to show that

$$\sum_{i,j} \left(MN^{-1}(A_i - M)^2 + MN^{-1}(B_j - M)^2 + e_{i,j}(A_i - M)(B_j - M) \right) \geq 0 \tag{2}$$

since multiplying this out and summing it with the use of

$$\sum_{i,j} MN^{-1} A_i^2 = \sum_i M A_i^2 = \sum_{i,j} e_{i,j} A_i M,$$

$$\sum_{i,j} MN^{-1} B_j^2 = \sum_j M B_j^2 = \sum_{i,j} e_{i,j} B_j M,$$

$$\sum_{i,j} -2MN^{-1} A_i M = -2M^2 T = -2M^3 N = \sum_{i,j} -2MN^{-1} B_j M,$$

$$\sum_{i,j} MN^{-1} M^2 = M^3 N = \sum_{i,j} e_{i,j} M^2$$

reduces (2) to

$$\sum_{i,j} e_{i,j} A_i B_j \geq N M^3,$$

which is exactly what we want.

To show (2) let $X_i = A_i - M, Y_j = B_j - M$. Substituting this in (2) and mutiplying it by 2, we obtain

$$\sum_{i,j} \left(2MN^{-1}X_i^2 + 2MN^{-1}Y_j^2 + 2e_{i,j}X_iY_j\right) \geq 0. \tag{3}$$

Adding and substracting

$$\sum_{i,j} e_{i,j}(X_i^2 + Y_j^2)$$

to the left hand side of (3) yields

$$\sum_{i,j} \left(X_i^2(2MN^{-1} - e_{i,j}) + Y_j^2(2MN^{-1} - e_{i,j}) + e_{i,j}(X_i + Y_j)^2\right) \geq 0. \tag{4}$$

Upon noting that

$$\sum_{i,j} 2MN^{-1}X_i^2 = \sum_i 2MX_i^2,$$

$$\sum_{i,j} e_{i,j}X_i^2 = \sum_i A_iX_i^2 = \sum_i X_i^2(X_i + M)$$

(and similarly for j), (4) simplifies to give

$$\sum_i X_i^2(M - X_i) + \sum_j Y_j^2(M - Y_j) + \sum_{i,j} e_{i,j}(X_i + Y_j)^2 \ \geq 0\,. \tag{5}$$

To make the proof easier to follow we shall at this point first prove the result in the special case when each $B_j \leq 2M$, equivalently $Y_j \leq M$, so

$$Y_j^2(M - Y_j) \geq 0.$$

Using this fact and the fact that for a fixed j with $B_j \neq 0$ (that is, $Y_j > -M$, equivalently $e_{i,j} = 1$ for some i) we have $\sum_i e_{i,j}B_j^{-1} = 1$, to show (5) it is enough to show that

$$\sum_i X_i^2(M - X_i) + \sum_{\substack{i,j:\, Y_j<0,\\ e_{i,j}=1}} e_{i,j}Y_j^2(M - Y_j)B_j^{-1} + \sum_{i,j} e_{i,j}(X_i + Y_j)^2 \ \geq 0\,. \tag{6}$$

For a *fixed* i consider

$$MX_i^2 - X_i^3 + \sum_{\substack{j:\, Y_j<0\\ e_{i,j}=1}} e_{i,j}Y_j^2(M - Y_j)B_j^{-1} + \sum_j e_{i,j}(X_i + Y_j)^2 \ \geq \ 0. \tag{7}$$

To show (6), it clearly suffices to show that (7) holds for each i. This is obvious when i is such that $X_i \le M$ so consider i with $X_i > M$. Keeping such an i fixed, we will write $\sum_{e_{i,j}=1} Y_j$ for $\sum_j e_{i,j} Y_j$ etc. Since we have

$$\sum_j e_{i,j} X_i^2 = \sum_{e_{i,j}=1} X_i^2 = (M + X_i) X_i^2 \, ,$$

(7) is equivalent to

$$2MX_i^2 + \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j^2 (M - Y_j) B_j^{-1} + \sum_{e_{i,j}=1} \left(2X_i Y_j + Y_j^2\right) \ \ge \ 0.$$

Since

$$\sum_{e_{i,j}=1} Y_j \ \ge \ \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j \qquad\qquad \sum_{e_{i,j}=1} Y_j^2 \ \ge \ \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j^2,$$

and $X_i > M > 0$ it suffices to show that

$$2MX_i^2 + \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} 2X_i Y_j + \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j^2 + \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j^2 (M - Y_j) B_j^{-1} \ \ge \ 0. \tag{8}$$

If

$$\sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j \ge -MX_i$$

then for the only possibly negative factor in (8),

$$\sum_{\substack{e_{i,j}=1 \\ Y_j<0}} 2X_i Y_j \ge -2MX_i^2 \tag{9}$$

and this cancels with the first term in (8) to produce a non-negative sum.

So assume from now on that

$$\sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j < -MX_i. \tag{10}$$

Using Hölder's inequality and the fact that

$$\sum_{\substack{e_{i,j}=1 \\ Y_j<0}} 1 \ \le \ M + X_i \, ,$$

we can see that

$$\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j^2 \;\geq\; \frac{\left(\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j\right)^2}{M+X_i}\,. \tag{11}$$

Since $MB_j^{-1} > 1$ when $Y_j < 0$ and $e_{i,j} = 1$, we also have

$$\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j^2 MB_j^{-1} \;\;\geq\;\; \frac{\left(\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j\right)^2}{M+X_i}\,. \tag{12}$$

Furthermore, using Chebychev's sum inequality and (11),

$$-\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j^3 = \sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j^2(-Y_j) \;\geq\; \frac{\left(\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j^2\right)\left(-\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j\right)}{(M+X_i)} \;\geq\; -\frac{\left(\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j\right)^3}{(M+X_i)^2}.$$

so since $B_j \leq M$ when $Y_j < 0$ and $e_{i,j} = 1$, $B_j^{-1} > M^{-1}$ and

$$-\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j^3 B_j^{-1} \;\geq\; -\frac{\left(\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j\right)^3}{M(M+X_i)^2}. \tag{13}$$

It now follows from (11), (12) and (13) that to show (8) it suffices to show that

$$2MX_i^2 + 2X_i\left(\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j\right) + \frac{2\left(\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j\right)^2}{(M+X_i)} - \frac{\left(\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j\right)^3}{M(M+X_i)^2} \;\geq\; 0. \tag{14}$$

Putting

$$Z = \frac{-\sum_{\substack{e_{i,j}=1\\Y_j<0}} Y_j}{M+X_i}$$

this amounts to showing that

$$\frac{2M^2X_i^2}{M+X_i} - 2MX_iZ + 2MZ^2 + Z^3 \geq 0\,. \tag{15}$$

Noting that $-Y_j \leq M-1$ when $e_{i,j}=1$ (since $Y_j = B_j - M \geq e_{i,j} - M$), we have from (10) that

$$\frac{MX_i}{(X_i+M)} \;\leq\; Z \;\leq M-1\,. \tag{16}$$

Inequality (15) does hold at the two extreme points $MX_i(M+X_i)^{-1}$ and $M-1$ in (16) so it is a matter of checking that it holds between them too. The right hand side of (15) is a cubic polynomial, so it suffices to check that if its local minimum lies between $MX_i(M+X_i)^{-1}$ and $M-1$ then then its value at this point is non-negative.

The local minimum occurs at

$$Z = \frac{-2M + \sqrt{4M^2 + 6MX_i}}{3}$$

and this point is between the limits given in (16) just when $X_i \leq -M/2$ or

$$2M \leq X_i \leq 7M/2 + 3/(2M) - 5 \leq 4M. \tag{17}$$

We have assumed that $X_i \geq M$ so only (17) needs to be considered.

Following multiplying out and squaring to remove the square root it can be seen that for (15) to hold at this local minimum point reduces to the requirement that

$$324M^4 + 972M^3X_i + 1161M^2X_i^2 - 216MX_i^3 \geq 0.$$

This does hold when $2M \leq X_i \leq 4M$ since then the only negative term is $-216MX_i^3$ and

$$1161M^2X_i^2 - 216MX_i^3 \geq MX_i^2(1161M - 216\cdot(4M)) > 0.$$

The result in the special case that each $B_j \leq 2M$ now follows.

We now turn to the general case where some of the B_j are greater than $2M$, equivalently some of the Y_j are greater than M. For this we consider

$$\sum_{e_{i,j}=1} e_{i,j}A_i^{-1}X_i^2(M-X_i) + \sum_{e_{i,j}=1} e_{i,j}B_j^{-1}Y_j^2(M-Y_j) + \sum_{e_{i,j}=1} e_{i,j}(X_i+Y_j)^2 \;\geq\; 0 \tag{18}$$

in place of (5), observing again that A_i^{-1} and B_j^{-1} are well defined when $e_{i,j}=1$. It suffices to show (18) since the left hand side equals that of (5) except that the terms $X_i^2(M-X_i)$ or $Y_j^2(M-Y_j)$ (equal to $2M^3$) are missing when $A_i = 0$ or $B_j = 0$ respectively. In other words we aim to show

$$\sum_{e_{i,j}=1} e_{i,j}\left(A_i^{-1}X_i^2(M-X_i)+B_j^{-1}Y_j^2(M-Y_j)+(X_i+Y_j)^2\right) \geq 0\,. \tag{19}$$

Define

$$f(i,j)=\begin{cases} 0 & \text{if } X_i\leq M,\\ 1 & \text{if } X_i>M.\end{cases} \qquad \tilde{f}(i,j)=1-f(i,j)=\begin{cases} 1 & \text{if } X_i\leq M,\\ 0 & \text{if } X_i>M.\end{cases}$$

$$g(i,j)=\begin{cases} 1 & \text{if } Y_j\leq M,\\ 0 & \text{if } Y_j>M.\end{cases} \qquad \tilde{g}(i,j)=1-g(i,j)=\begin{cases} 0 & \text{if } Y_j\leq M,\\ 1 & \text{if } Y_j>M.\end{cases}$$

$$h(i,j)=\begin{cases} 2^{-1} & \text{if } X_i\leq M \text{ and } Y_j\leq M,\\ 0 & \text{if } X_i\leq M \text{ and } Y_j>M,\\ 1 & \text{if } X_i>M \text{ and } Y_j\leq M,\\ X_i(X_i+Y_j)^{-1} & \text{if } X_i>M \text{ and } Y_j>M.\end{cases}$$

$$\tilde{h}(i,j)=1-h(i,j)=\begin{cases} 2^{-1} & \text{if } X_i\leq M \text{ and } Y_j\leq M,\\ 1 & \text{if } X_i\leq M \text{ and } Y_j>M,\\ 0 & \text{if } X_i>M \text{ and } Y_j\leq M,\\ Y_i(X_i+Y_j)^{-1} & \text{if } X_i>M \text{ and } Y_j>M.\end{cases}$$

Then the left hand side of (19) is the sum of

$$\sum_{e_{i,j}=1} e_{i,j}\left(f(i,j)A_i^{-1}X_i^2(M-X_i)+g(i,j)B_j^{-1}Y_j^2(M-Y_j)+h(i,j)(X_i+Y_j)^2\right) \tag{20}$$

and

$$\sum_{e_{i,j}=1} e_{i,j}\left(\tilde{f}(i,j)A_i^{-1}X_i^2(M-X_i)+\tilde{g}(i,j)B_j^{-1}Y_j^2(M-Y_j)+\tilde{h}(i,j)(X_i+Y_j)^2\right). \tag{21}$$

So it suffices to show that (20) is non-negative for any *fixed* i

For a fixed i such that $X_i\leq M$, (20) is

$$\sum_{\substack{e_{i,j}=1\\ Y_j\leq M}} 2^{-1}(X_i+Y_j)^2+\sum_{\substack{e_{i,j}=1\\ Y_j\leq M}} Y_j^2(M-Y_j)B_j^{-1} \tag{22}$$

and for $X_i>M$ it is

$$\sum_{\substack{e_{i,j}=1\\ Y_j\leq M}}(X_i+Y_j)^2+\sum_{\substack{e_{i,j}=1\\ Y_j>M}} X_i(X_i+Y_j)+X_i^2(M-X_i)+\sum_{\substack{e_{i,j}=1\\ Y_j\leq M}} Y_j^2(M-Y_j)B_j^{-1}. \tag{23}$$

Similarly for a fixed j such that $Y_j \le M$, (21) is

$$\sum_{\substack{e_{i,j}=1 \\ X_i \le M}} 2^{-1}(X_i + Y_j)^2 + \sum_{\substack{e_{i,j}=1 \\ X_i \le M}} X_i^2(M - X_i)A_i^{-1} \tag{24}$$

and for $Y_j > M$ it is

$$\sum_{\substack{e_{i,j}=1 \\ X_i \le M}} (X_i + Y_j)^2 + \sum_{\substack{e_{i,j}=1 \\ X_i > M}} Y_j(X_i + Y_j) + Y_j^2(M - Y_j) + \sum_{\substack{e_{i,j}=1 \\ X_i \le M}} X_i^2(M - X_i)A_i^{-1}. \tag{25}$$

Thus it is enough (by symmetry) to show that both (22) and (23) are non-negative. For (22) this is clear so we assume now that $X_i > M$ and show it for (23). Expanding and simplifying (23) gives

$$\begin{aligned}(M + X_i)X_i^2 + \sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} 2X_iY_j + \sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} Y_j^2 \\ + \sum_{\substack{e_{i,j}=1 \\ Y_j > M}} X_iY_j + MX_i^2 - X_i^3 + \sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} Y_j^2(M - Y_j)B_j^{-1}.\end{aligned}$$

which simplifies further to

$$2MX_i^2 + \sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} 2X_iY_j + \sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} Y_j^2 + \sum_{\substack{e_{i,j}=1 \\ Y_j > M}} X_iY_j + \sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} Y_j^2(M - Y_j)B_j^{-1}. \tag{26}$$

In this sum the only potentially negative term is the second one,

$$\sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} 2X_iY_j. \tag{27}$$

If

$$\sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} Y_j \ge -MX_i \tag{28}$$

then (27) is canceled out by the first term in (26), so we may henceforth assume that (28) does not hold. Notice that we now have

$$MX_i < \; -\sum_{\substack{e_{i,j}=1 \\ Y_j \le M}} Y_j \; \le -\sum_{\substack{e_{i,j}=1 \\ Y_j < 0}} Y_j \; \le \; (M + X_i)(M - 1)\,.$$

Clearly to show that the sum in (26) is non-negative it is enough to show that the following sum is non-negative:

$$2MX_i^2 + \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} 2X_iY_j + \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j^2 + \sum_{\substack{e_{i,j}=1 \\ Y_j<0}} Y_j^2(M-Y_j)B_j^{-1}. \tag{29}$$

But that follows, as required, exactly as in the special case of (8) considered earlier. □

For an $N \times N$ matrix $(e_{i,j})$ let

$$\parallel (e_{i,j}) \parallel = \sum_{i,j} |e_{i,j}|.$$

Then a restatement of Theorem 1 gives:

Corollary 2. *For an $N \times N$ $\{0,1\}$-matrix E,*

$$N^2 \parallel EE^TE \parallel \geq \parallel E \parallel^3 .$$

References

[1] Hill, A. & Paris, J.B., The Counterpart Principle of Analogical Support by Structural Similarity, *Erkenntnis,* On Line First, October 2013, DOI: 10.1007/s10670-013-9546-0.

[2] Paris, J.B. & Vencovská, A., *Pure Inductive Logic*, in the Association of Symbolic Logic Perspectives in Mathematical Logic Series, Cambridge University Press, April 2015.

[3] Paris, J.B and Vencovská, A. , Combining Analogical Support in Pure Inductive Logic, to be submitted to Erkenntnis.