Causation and correlation

Ait-taleb Nabil *

Abstract

In this paper, I will introduce the causation's magnitude allowing to compute the importance of causes in the cause-and-effect relationship from correlation matrix.

---
*Corresponding author: nabiltravail1982@gmail.com

1 Introduction

This paper is based on the results obtained in papers [5] and [6]. The purpose of this paper is to relate the notions of causality and correlation. The paper begins by introducing the information gain/impact and the causation's magnitude computing the contribution of the causes in the cause-and-effect relationship which will be computed from a correlation matrix only. The paper ends with the application of the causation's magnitude in order to obtain a causal graph from a correlation matrix.

2 Information gain/impact

Definition:

*Consider a set of events $E = \{X_j, \Omega_1, \Omega_2\}$ and the correlation matrix $\tilde{K}_{(X_j, \Omega_1, \Omega_2)^2}$. We define **the information gain/impact** $\epsilon_{\Omega_1 \longrightarrow X_j}$ as being the information brought from a set of events Ω_1 to an event X_j:*

$$\epsilon_{\Omega_1 \longrightarrow X_j} = -\frac{1}{2} \ln\left(1 - \tilde{K}_{X_j; \Omega_1} . K_{\Omega_1^2}^{-1} . \tilde{K}_{\Omega_1; X_j}\right)$$

3 The causation's magnitude obtained from correlations

In this section, we will consider the result obtained in paper [6] page 4 and interpret the percentage of conditioning as the causation's magnitude.

Definition:

*Consider a set of events $E = \{X_j, \Omega_1, \Omega_2\}$ and the correlation matrix $\tilde{K}_{(X_j, \Omega_1, \Omega_2)^2}$. We define **the causation's magnitude** $\lambda_{\Omega_1 \longrightarrow X_j}$ of a events's set Ω_1 which acts on the event X_j as follows:*

$$\lambda_{\Omega_1 \rightarrow X_j} = \frac{\epsilon_{\Omega_1 \longrightarrow X_j}}{\epsilon_{\Omega_1 \cup \Omega_2 \longrightarrow X_j}} = \frac{\ln\left(1 - \tilde{K}_{X_j; \Omega_1} \tilde{K}_{\Omega_1^2}^{-1} . \tilde{K}_{\Omega_1; X_j}\right)}{\ln\left(1 - \tilde{K}_{X_j; \Omega_1 \Omega_2} \tilde{K}_{(\Omega_1, \Omega_2)^2}^{-1} . \tilde{K}_{\Omega_1 \Omega_2; X_j}\right)}$$

<u>Properties:</u>

1. $0 \leq \lambda_{\Omega_1 \longrightarrow X_j} \leq 1$ with $\lambda_{\varnothing \longrightarrow X_j} = 0$ and $\lambda_{\Omega_1 \cup \Omega_2 \longrightarrow X_j} = 1$

2. We can consider a **strong effect** of Ω_1 on X_1 when $\lambda_{\Omega_1 \longrightarrow X_j} \geq 0.75$

3. if $0 \leq -\frac{1}{2} \ln\left(1 - \tilde{K}_{X_j; \Omega_1 \Omega_2} \tilde{K}_{(\Omega_1, \Omega_2)^2}^{-1} . \tilde{K}_{\Omega_1 \Omega_2; X_j}\right) \leq 0.1$, we can consider the event X_j to be independent of the set $\Omega_1 \cup \Omega_2$. Above a value of 0.1 we can compute the contribution of the set causes Ω_1 which acts on the event X_j.

4 Causal graph obtained from correlations matrix

To obtain the causal graph from a correlation matrix, we will use the method of successive highest causation's magnitudes (see paper [6] page 6).

We will present the contribution of the causes on the form of a bar graph and then will set a causation's magnitude value of 0.75 to be exceeded for each node assigned to the events in order to obtain the causal graph.

Once the causal graph is obtained, we will compute the density of the graph as in paper (see paper [6] page 7)

5 Correlation matrix Attitude and causal graph

We will consider in the following the attitude 's correlation matrix found in the R software.

$$
\tilde{K}_{(X_1,X_2,X_3,X_4,X_5,X_6,X_7)^2} = \begin{pmatrix} 1.0000000 & 0.8254176 & 0.4261169 & 0.6236782 & 0.5901390 & 0.1564392 & 0.1550863 \\ 0.8254176 & 1.0000000 & 0.5582882 & 0.5967358 & 0.6691975 & 0.1877143 & 0.2245796 \\ 0.4261169 & 0.5582882 & 1.0000000 & 0.4933310 & 0.4454779 & 0.1472331 & 0.3432934 \\ 0.6236782 & 0.5967358 & 0.4933310 & 1.0000000 & 0.6403144 & 0.1159652 & 0.5316198 \\ 0.5901390 & 0.6691975 & 0.4454779 & 0.6403144 & 1.0000000 & 0.3768830 & 0.5741862 \\ 0.1564392 & 0.1877143 & 0.1472331 & 0.1159652 & 0.3768830 & 1.0000000 & 0.2833432 \\ 0.1550863 & 0.2245796 & 0.3432934 & 0.5316198 & 0.5741862 & 0.2833432 & 1.0000000 \end{pmatrix}
$$

where the events are:

1. X_1 = *rating*

2. X_2 = *complaints*

3. X_3 = *privileges*

4. X_4 = *learning*

5. X_5 = *raises*

6. X_6 = *critical*

7. X_7 = *advance*

5.1 Causation's magnitude and total information gain/impact for the node X_1

The total information gain/impact value is equal:

$$\epsilon_{X_2X_3X_4X_5X_6X_7 \longrightarrow X_1} = -\frac{1}{2}\ln\left(1 - \tilde{K}_{X_1;X_2X_3X_4X_5X_6X_7}.\tilde{K}^{-1}_{(X_2X_3X_4X_5X_6X_7)^2}.\tilde{K}_{X_2X_3X_4X_5X_6X_7;X_1}\right) = 0.65 \geq 0.1$$

We can compute the cause magnitude of this value:

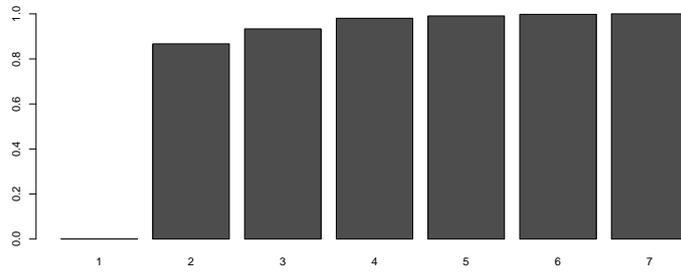Figure 1: Causation's magnitude

1. $\lambda_{\varnothing \longrightarrow X_1} = 0$

2. $\lambda_{X_2 \longrightarrow X_1} = 0.866971$

3. $\lambda_{\{X_2,X_4\} \longrightarrow X_1} = 0.9333113$

4. $\lambda_{\{X_2,X_4,X_7\} \longrightarrow X_1} = 0.9803893$

5. $\lambda_{\{X_2,X_4,X_7,X_3\} \longrightarrow X_1} = 0.9908109$

6. $\lambda_{\{X_2,X_4,X_7,X_3,X_5\} \longrightarrow X_1} = 0.997756$

7. $\lambda_{\{X_2,X_4,X_7,X_3,X_5,X_6\} \longrightarrow X_1} = 1$

5.2 Causation's magnitude and total information gain/impact for the node X_2

The total information gain/impact value is equal:

$$\epsilon_{X_1X_3X_4X_5X_6X_7 \longrightarrow X_2} = -\frac{1}{2}\ln\left(1 - \tilde{K}_{X_2;X_1X_3X_4X_5X_6X_7} . \tilde{K}^{-1}_{(X_1X_3X_4X_5X_6X_7)^2} . \tilde{K}_{X_1X_3X_4X_5X_6X_7;X_2}\right) = 0.73 \geq 0.1$$

We can compute the cause magnitude of this value:

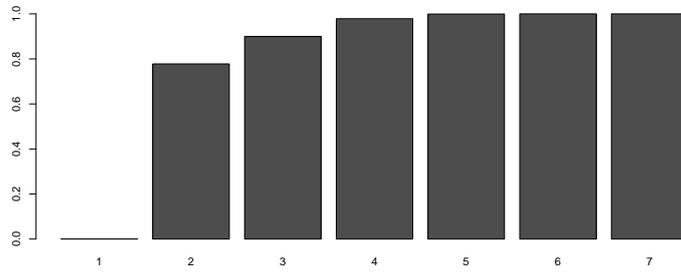Figure 2: Causation's magnitude

1. $\lambda_{\varnothing \longrightarrow X_2} = 0$

2. $\lambda_{X_1 \longrightarrow X_2} = 0.7778966$

3. $\lambda_{\{X_1,X_3\} \longrightarrow X_2} = 0.8994167$

4. $\lambda_{\{X_1,X_3,X_5\} \longrightarrow X_2} = 0.9784996$

5. $\lambda_{\{X_1,X_3,X_5,X_7\} \longrightarrow X_2} = 0.9991813$

6. $\lambda_{\{X_1,X_3,X_5,X_7,X_6\} \longrightarrow X_2} = 0.9999381$

7. $\lambda_{\{X_1,X_3,X_5,X_7,X_6,X_4\} \longrightarrow X_2} = 1$

5.3 Causation's magnitude and total information gain/impact for the node X_3

The total information gain/impact value is equal:

$$\epsilon_{X_1 X_2 X_4 X_5 X_6 X_7 \longrightarrow X_3} = -\frac{1}{2} \ln\left(1 - \tilde{K}_{X_3; X_1 X_2 X_4 X_5 X_6 X_7} . \tilde{K}^{-1}_{(X_1 X_2 X_4 X_5 X_6 X_7)^2} . \tilde{K}_{X_1 X_2 X_4 X_5 X_6 X_7; X_3}\right) = 0.24 \geq 0.1$$
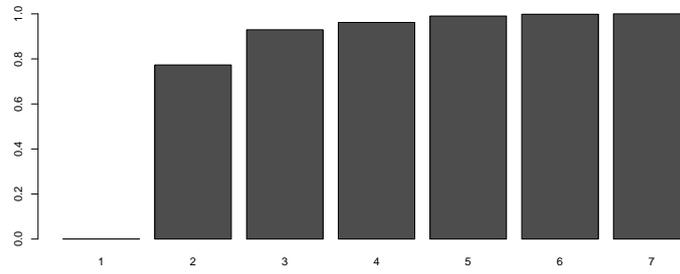
We can compute the cause magnitude of this value:

Figure 3: Causation's magnitude

1. $\lambda_{\varnothing \longrightarrow X_3} = 0$

2. $\lambda_{X_2 \longrightarrow X_3} = 0.773189$

3. $\lambda_{\{X_2, X_7\} \longrightarrow X_3} = 0.9293297$

4. $\lambda_{\{X_2, X_7, X_4\} \longrightarrow X_3} = 0.961864$

5. $\lambda_{\{X_2, X_7, X_4, X_1\} \longrightarrow X_3} = 0.9903198$

6. $\lambda_{\{X_2, X_7, X_4, X_1, X_5\} \longrightarrow X_3} = 0.9982807$

7. $\lambda_{\{X_2, X_7, X_4, X_1, X_5, X_6\} \longrightarrow X_3} = 1$

5.4 Causation's magnitude and total information gain/impact for the node X_4

The total information gain/impact value is equal:

$$\epsilon_{X_1 X_2 X_3 X_5 X_6 X_7 \longrightarrow X_4} = -\frac{1}{2} \ln\left(1 - \tilde{K}_{X_4; X_1 X_2 X_3 X_5 X_6 X_7} . \tilde{K}^{-1}_{(X_1 X_2 X_3 X_5 X_6 X_7)^2} . \tilde{K}_{X_1 X_2 X_3 X_5 X_6 X_7; X_4}\right) = 0.49 \geq 0.1$$
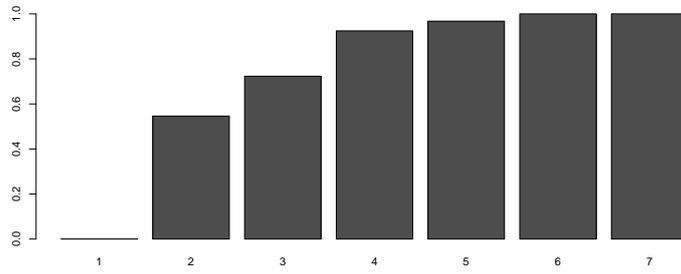
We can compute the cause magnitude of this value:

Figure 4: Causation's magnitude

1. $\lambda_{\varnothing \longrightarrow X_4} = 0$

2. $\lambda_{X_5 \longrightarrow X_4} = 0.5461211$

3. $\lambda_{\{X_5, X_1\} \longrightarrow X_4} = 0.7230502$

4. $\lambda_{\{X_5, X_1, X_7\} \longrightarrow X_4} = 0.924563$

5. $\lambda_{\{X_5, X_1, X_7, X_6\} \longrightarrow X_4} = 0.9674856$

6. $\lambda_{\{X_5, X_1, X_7, X_6, X_3\} \longrightarrow X_4} = 0.9999058$

7. $\lambda_{\{X_5, X_1, X_7, X_6, X_3, X_2\} \longrightarrow X_4} = 1$

5.5 Causation's magnitude and total information gain/impact for the node X_5

The total information gain/impact value is equal:

$$\epsilon_{X_1 X_2 X_3 X_4 X_6 X_7 \longrightarrow X_5} = -\frac{1}{2} \ln \left(1 - \tilde{K}_{X_5; X_1 X_2 X_3 X_4 X_6 X_7} . \tilde{K}^{-1}_{(X_1 X_2 X_3 X_4 X_6 X_7)^2} . \tilde{K}_{X_1 X_2 X_3 X_4 X_6 X_7; X_5} \right) = 0.56 \geq 0.1$$
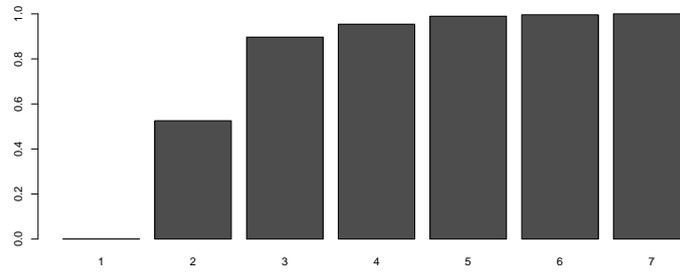
We can compute the cause magnitude of this value:

Figure 5: Causation's magnitude

1. $\lambda_{\varnothing \longrightarrow X_5} = 0$

2. $\lambda_{X_2 \longrightarrow X_5} = 0.5254475$

3. $\lambda_{\{X_2, X_7\} \longrightarrow X_5} = 0.896712$

4. $\lambda_{\{X_2, X_7, X_6\} \longrightarrow X_5} = 0.9541725$

5. $\lambda_{\{X_2, X_7, X_6, X_4\} \longrightarrow X_5} = 0.9895531$

6. $\lambda_{\{X_2, X_7, X_6, X_4, X_1\} \longrightarrow X_5} = 0.9959638$

7. $\lambda_{\{X_2, X_7, X_6, X_4, X_1, X_3\} \longrightarrow X_5} = 1$

5.6 Causation's magnitude and total information gain/impact for the node X_6

The total information gain/impact value is equal:

$$\epsilon_{X_1 X_2 X_3 X_4 X_5 X_7 \longrightarrow X_6} = -\frac{1}{2} \ln\left(1 - \tilde{K}_{X_6; X_1 X_2 X_3 X_4 X_5 X_7} . \tilde{K}^{-1}_{(X_1 X_2 X_3 X_4 X_5 X_7)^2} . \tilde{K}_{X_1 X_2 X_3 X_4 X_5 X_7; X_6}\right) \approx 0.1$$

We can consider the event X_6 to be independent other events $X_1 X_2 X_3 X_4 X_5 X_7$

5.7 Causation's magnitude and total information gain/impact for the node X_7

The total information gain/impact value is equal:

$$\epsilon_{X_1 X_2 X_3 X_4 X_5 X_6 \longrightarrow X_7} = -\frac{1}{2} \ln\left(1 - \tilde{K}_{X_7; X_1 X_2 X_3 X_4 X_5 X_6} \cdot \tilde{K}^{-1}_{(X_1 X_2 X_3 X_4 X_5 X_6)^2} \cdot \tilde{K}_{X_1 X_2 X_3 X_4 X_5 X_6; X_7}\right) = 0.36 \geq 0.1$$
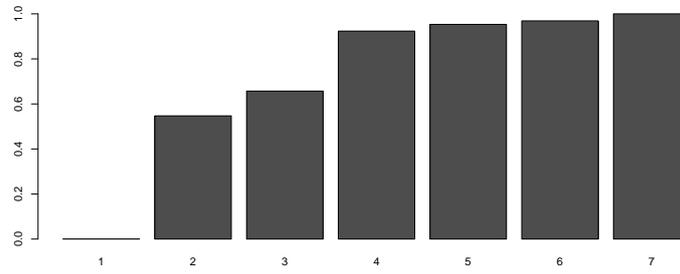
We can compute the cause magnitude of this value:

Figure 6: Causation's magnitude

1. $\lambda_{\varnothing \longrightarrow X_7} = 0$

2. $\lambda_{X_5 \longrightarrow X_7} = 0.5471036$

3. $\lambda_{\{X_5, X_1\} \longrightarrow X_7} = 0.657135$

4. $\lambda_{\{X_5, X_1, X_4\} \longrightarrow X_7} = 0.9234174$

5. $\lambda_{\{X_5, X_1, X_4, X_6\} \longrightarrow X_7} = 0.9532082$

6. $\lambda_{\{X_5, X_1, X_4, X_6, X_2\} \longrightarrow X_7} = 0.9687386$

7. $\lambda_{\{X_5, X_1, X_4, X_6, X_2, X_3\} \longrightarrow X_7} = 1$

5.8 Causal graph obtained from the attitude's correlation matrix

By using the method of successive highest causation's magnitude (see paper [6] page 6) and by setting a causation's magnitude value of 0.75 to exceed, we get the following causal graph:
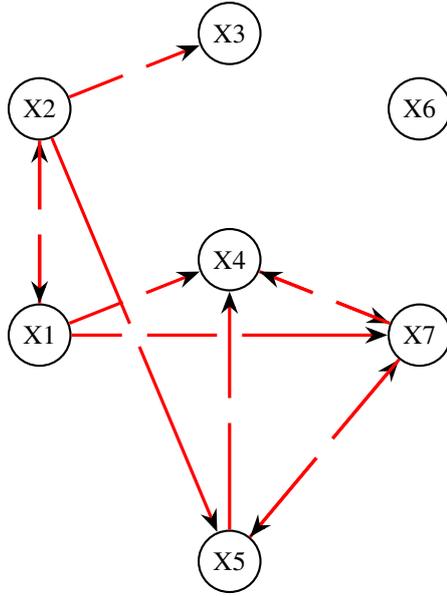
We have as causation's magnitude value and the information gain/impact:

1. $\lambda_{X_2 \longrightarrow X_1} = 0.866971$ $\epsilon_{X_2 \longrightarrow X_1} = 0.57 (nats)$

2. $\lambda_{X_1 \longrightarrow X_2} = 0.7778966$ $\epsilon_{X_1 \longrightarrow X_2} = 0.57 (nats)$

3. $\lambda_{X_2 \longrightarrow X_3} = 0.773189$ $\epsilon_{X_2 \longrightarrow X_3} = 0.19 (nats)$

4. $\lambda_{\{X_5, X_1, X_7\} \longrightarrow X_4} = 0.924563$ $\epsilon_{X_5 X_1 X_7 \longrightarrow X_4} = 0.45 (nats)$

5. $\lambda_{\{X_2, X_7\} \longrightarrow X_5} = 0.896712$ $\epsilon_{X_2 X_7 \longrightarrow X_5} = 0.5 (nats)$

6. $\lambda_{X_6} = 0$ $\epsilon_{X_6} = 0 (nats)$

7. $\lambda_{\{X_5, X_1, X_4\} \longrightarrow X_7} = 0.9234174$ $\epsilon_{X_5 X_1 X_4 \longrightarrow X_7} = 0.33 (nats)$

and we can compute the graph's density:

$$Density = \frac{11}{7.6} = \frac{11}{42} \approx 0.26$$

12

6   Conclusion

In this paper we have exposed the causation's magnitude computed from correlations. From an example, we have shown the steps to follow to obtain a causal graph from a correlation matrix.

[1]Elements of information theory. Author: Thomas M.Cover and Joy A.Thomas. Copyright 1991 John Wiley and sons.

[2]Optimal stastical decisions. Author: Morris H.DeGroot. Copyright 1970-2004 John Wiley and sons.

[3]Matrix Analysis. Author: Roger A.Horn and Charles R.Johnson. Copyright 2012, Cambridge university press.

[4]Causality: Models, reasoning and inference. Author:Judea Pearl .Copyright 2000, Cambridge university press.

[5]Directed dependency graph obtained from a continious data matrix by the highest successive conditionings method. Author: Ait-taleb Nabil.Years:2021. Published: Vixra

[6]Directed dependency graph obtained from a correlation matrix by the highest succesive conditionings method. Author: Ait-taleb Nabil. Years:2021.Published: Vixra