Causation of multiple causes acting on a single variable computed from correlations

Ait-taleb Nabil *

Abstract

In this paper, we will expose the causation of multiple causes acting on a single variable computed from correlations. Using an example, we will show when strong or weak correlations between multiple causes and a variable imply a strong or weak causation between the causes and the variable.

*Corresponding author: nabiltravail1982@gmail.com

1 Introduction

In this paper we will present the relationship between the causation and correlations for multiple causes acting on a variable. We will present the table of magnitudes for correlations and causation in order to further interpret the notion of strong and weak correlations and causation. We will then explain the way in which we will study a problem of two causes acting on a variable. We will describe the fields of correlation pairs having strong and weak causation. The paper ends with numerical applications for a problem of two causes acting on a variable. From these numerical applications, we will be able to relate the notion of strong and weak correlations with the notion of strong and weak causation.

2 Causation of multiple causes acting on a single variable computed from correlations

The causation of multiple causes acting on a single variable computed from correlations is expressed as follows:

$$\epsilon_{\Omega \longrightarrow X} = -\frac{1}{2}\ln(1 - \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X})$$

Where $\tilde{K}_{X,\Omega}$ is a correlation's vector between the causes Ω and the variable X and $\tilde{K}_{\Omega^2}$ is a correlation matrix of causes Ω.

Proof:

$\epsilon_{\Omega \longrightarrow X} = h(X) - h(X|\Omega) = \frac{1}{2}\ln(2.\pi.e.K_{X^2}) - \frac{1}{2}\ln(2.\pi.e.K_{X^2|\Omega})$

Where $h()$ is the differential entropy Gaussian (see Appendix). K_{X^2} is the variance and $K_{X^2|\Omega}$ is the conditional variance.

In this proof, $\tilde{K}_{\Omega,X}$ is a correlation vector between X and the set of causes Ω. $\tilde{K}_{\Omega^2}$ corresponds to the correlation matrix of causes Ω.

In what follows, we will factorize the variance K_{X^2} of the conditional variance $K_{X^2|\Omega}$:

$K_{X^2|\Omega} = K_{X^2} - K_{X,\Omega}.K_{\Omega^2}^{-1}.K_{X,\Omega}$

$K_{X^2|\Omega} = K_{X^2} - K_{X,\Omega}.(diag^{-1}(K_{\Omega^2}))^{\frac{1}{2}}.\tilde{K}_{\Omega^2}^{-1}.(diag^{-1}(K_{\Omega^2}))^{\frac{1}{2}}.K_{\Omega,X}$

$K_{X^2|\Omega} = K_{X^2} - K_{X^2}^{\frac{1}{2}}.\tilde{K}_{X,\Omega}.\tilde{K}_{\Omega^2}^{-1}.K_{X^2}^{\frac{1}{2}}.\tilde{K}_{\Omega,X}$

$K_{X^2|\Omega} = K_{X^2}.(1 - \tilde{K}_{X,\Omega}.\tilde{K}_{\Omega^2}^{-1}.\tilde{K}_{\Omega,X})$

We obtain:

$\epsilon_{\Omega \longrightarrow X}$

$= h(X) - h(X|\Omega)$

$= \frac{1}{2}\ln(2.\pi.e.K_{X^2}) - \frac{1}{2}\ln(2.\pi.e.K_{X^2|\Omega})$

$= \frac{1}{2}\ln(2.\pi.e.K_{X^2}) - \frac{1}{2}\ln(2.\pi.e.K_{X^2}.(1 - \tilde{K}_{X,\Omega}.\tilde{K}_{\Omega^2}^{-1}.\tilde{K}_{\Omega,X}))$

$= -\frac{1}{2}\ln(1 - \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X})$

3 Correlation value range

We will explain the importance of correlations in order to be able to interpret the order of magnitude in what will follow:

Level of correlation	ρ_{min}	ρ_{max}
Very strong positive correlation	0.8	1
Fairly strong positive correlation	0.6	0.79
Moderate positive correlation	0.4	0.59
Weak positive correlation	0.2	0.39
Very Weak positive correlation	0	0.19
Very strong negative correlation	-1	-0.8
Fairly strong negative correlation	-0.79	-0.6
Moderate negative correlation	-0.59	-0.4
Weak negative correlation	-0.39	-0.2
Very Weak negative correlation	-0.19	0

4 Causation value range

In this section we will give an order of magnitude to the causation. To do this we will divide the quadratic form $0 \leq \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X} < 1$ by 4:

1. $0 \leq \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X} \leq 0.25$ which means we have negligible causation

2. $0.25 < \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X} \leq 0.5$ which means we have weak causation

3. $0.5 < \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X} \leq 0.75$ which means we have moderate causation

4. $0.75 < \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X} < 1$ which means we have strong causation

By using the relationship:

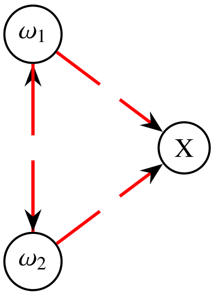
$$\epsilon_{\Omega \longrightarrow X} = -\frac{1}{2}\ln\left(1 - \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X}\right)$$

We can build the following table:

Level of causation	$\epsilon^{min}_{\Omega \longrightarrow X}$	$\epsilon^{max}_{\Omega \longrightarrow X}$
negligible causation	0	0.14
Weak causation	0.15	0.35
moderate causation	0.36	0.69
Strong causation	0.7	$\longrightarrow$

5 Problem: Causation of two causes acting on a single variable computed from correlations

In what follows, we will consider a set of two causes $\Omega = \{\omega_1, \omega_2\}$ acting on a variable X as follows:



To this graph we attribute a matrix of correlations of the causes $\tilde{K}_{\Omega^2}$ and a weight vector of correlations $\tilde{K}_{X,\Omega}$ between the causes Ω and the variable X:

$$\tilde{K}_{\Omega^2} = \begin{pmatrix} 1 & \rho_{\omega_1\omega_2} \\ \rho_{\omega_1\omega_2} & 1 \end{pmatrix} \text{ and } \tilde{K}_{X,\Omega} = (\rho_{\omega_1 X}, \rho_{\omega_2 X})$$

Then we will present a field of correlations $\tilde{K}_{X,\Omega} = (\rho_{\omega_1 X}, \rho_{\omega_2 X})$ for which there is a strong causation:

$$0.7 \leq \epsilon_{\Omega \longrightarrow X} = -\frac{1}{2} \ln(1 - \tilde{K}_{X,\Omega} \tilde{K}_{\Omega^2}^{-1} \tilde{K}_{\Omega,X}) \leq 1.15$$

We will also show the representation for a weak causation:

$$0.15 \leq \epsilon_{\Omega \longrightarrow X} = -\frac{1}{2} \ln(1 - \tilde{K}_{X,\Omega} \tilde{K}_{\Omega^2}^{-1} \tilde{K}_{\Omega,X}) \leq 0.35$$

For correlation's field $\tilde{K}_{X,\Omega} = (\rho_{\omega_1 X}, \rho_{\omega_2 X})$, we select correlation pairs to expose the following situations:

1. When we have a pair of **very strong and fairly strong correlations** between the causes Ω and the variable X implies a **strong causation** between the causes and the variable.
2. When we have a pair of **weak correlations** between the causes Ω and the variable X implies a **strong causation** between the causes and the variable.
3. When we have a pair of **fairly strong correlations** between the causes Ω and the variable X implies a **weak causation** between the causes and the variable.
4. When a pair of **weak correlations** between the causes Ω and the variable X implies a **weak causation** between the causes and variable.

6 Correlation and strong causation between two causes and a single variable

In what follows we will consider the matrix of causes $\tilde{K}_{\Omega^2}$:

$$\tilde{K}_{\Omega^2} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$$

From the previous matrix, we will now represent the pairs of correlations $\tilde{K}_{X,\Omega}$ having a strong causation $0.7 \leq \epsilon_{\Omega \longrightarrow X} \leq 1.15$:

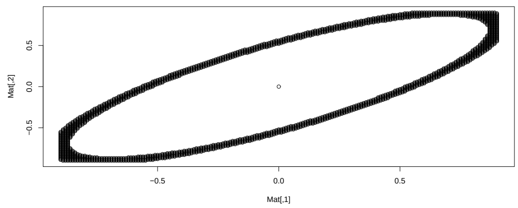


Figure 1: Pairs of correlations $\tilde{K}_{X,\Omega}$ having a strong causation $0.7 \leq \epsilon_{\Omega \longrightarrow X} \leq 1.15$

From this graph we will select two points:$\tilde{K}_{X,\Omega} = (0.76, 0.86)$ and $\tilde{K}_{X,\Omega} = (0.21, -0.34)$. We will compute the causation value $\epsilon_{\Omega \longrightarrow X}$ for the two points:

$$\epsilon_{\Omega \longrightarrow X} = -\frac{1}{2} \ln\{1 - (0.76, 0.86).\begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}^{-1}.\begin{pmatrix} 0.76 \\ 0.86 \end{pmatrix}\} = 0.7012119$$

$$\epsilon_{\Omega \longrightarrow X} = -\frac{1}{2} \ln\{1 - (0.21, -0.34).\begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}^{-1}.\begin{pmatrix} 0.21 \\ -0.34 \end{pmatrix}\} = 0.7155297$$

We can therefore describe two situations:

1. A pair of **fairly and very strong correlations** between the causes and the variable implies a **strong causation** between the causes and variable.
2. A pair of **weak correlations** between the causes and the variable implies a **strong causation** between the causes and the variable.

7 Correlation and weak causation between two causes and a single variable

In what follows we will consider the same matrix of causes $\tilde{K}_{\Omega^2}$:

$$\tilde{K}_{\Omega^2} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$$

From the previous matrix, we will now represent the pairs of correlations $\tilde{K}_{X,\Omega}$ having a weak causation $0.15 \leq \epsilon_{\Omega \longrightarrow X} \leq 0.35$:

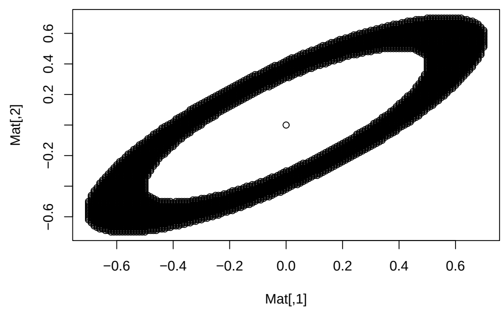


Figure 2: Pairs of correlations $\tilde{K}_{X,\Omega}$ having a weak causation $0.15 \leq \epsilon_{\Omega \longrightarrow X} \leq 0.35$

From this graph we will select two points:$\tilde{K}_{X,\Omega} = (0.70, 0.62)$ and $\tilde{K}_{X,\Omega} = (0.22, -0.20)$. We will compute the causation value $\epsilon_{\Omega \longrightarrow X}$ for the two points:

$$\epsilon_{\Omega \longrightarrow X} = -\frac{1}{2} \ln\{1 - (0.7, 0.62). \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}^{-1} . \begin{pmatrix} 0.7 \\ 0.62 \end{pmatrix}\} = 0.3465736$$

$$\epsilon_{\Omega \longrightarrow X} = -\frac{1}{2} \ln\{1 - (0.22, -0.20). \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}^{-1} . \begin{pmatrix} 0.22 \\ -0.20 \end{pmatrix}\} = 0.2909023$$

We can therefore describe two situations:

1. A pair of **fairly strong correlations** between the causes and the variable implies a **weak causation** between the causes and the variable.

2. A pair of **weak correlations** between the causes and the variable implies a **weak causation** between the causes and the variable.

8 Conclusion

In this paper we have exposed the existing relationship between the notion of causation and correlation for multiple causes acting on a variable. Using an example of two causes acting on a variable, we have shown the different situations that can occur:

1. When we have a pair of **very strong and fairly strong correlations** between the causes Ω and the variable X implies a **strong causation** between the causes and the variable.

2. When we have a pair of **weak correlations** between the causes Ω and the variable X implies a **strong causation** between the causes and the vaiable.

3. When we have a pair of **fairly strong correlations** between the causes Ω and the variable X implies a **weak causation** between the causes and the variable.

4. When a pair of **weak correlations** between the causes Ω and the variable X implies a **weak causation** between the causes and variable.

A Appendix

A.1 Differential entropy of a Gaussian random vector

Theorem: *Given random vector $\vec{x} = (x_1, x_2, ..., x_n)$ with a multivariate Gaussian distribution:*

$$P_X(\vec{x}) = \mathcal{N}(\mu_X, K_{X^2}) = (2\pi)^{-\frac{n}{2}}|K_{X^2}|^{-\frac{1}{2}} \exp\{-\frac{(\vec{x} - \vec{\mu_X})^t K_{X^2}^{-1}(\vec{x} - \vec{\mu_X})}{2}\}$$

with a mean vector μ_X and a covariance matrix K_{X^2} then the differential entropy is equal to:

$$h(X) = \frac{1}{2}\ln(2\pi e)^n |K_{X^2}|$$

Proof:

$h(X)$

$= -\int_{-\infty}^{+\infty} p_X(\vec{x}) \ln\{p_X(\vec{x})\}\overrightarrow{dx}$

$= -\int_{-\infty}^{+\infty} p_X(\vec{x})[-\frac{1}{2}(\vec{x} - \mu_X)^t K_{X^2}^{-1}(\vec{x} - \mu_X) - \ln(\sqrt{2\pi})^n |K_{X^2}|^{\frac{1}{2}}]\overrightarrow{dx}$

$= \frac{1}{2} E_X[\sum_{ij}(\vec{x}_i - \mu_{X_i})^t (K_{X^2}^{-1})_{ij}(\vec{x}_j - \mu_{X_j})] + \frac{1}{2}\ln(2\pi)^n |K_{X^2}|$

$= \frac{1}{2} E_X[\sum_{ij}(\vec{x}_i - \mu_{X_i})^t (\vec{x}_j - \mu_{X_j})(K_{X^2}^{-1})_{ij}] + \frac{1}{2}\ln(2\pi)^n |K_{X^2}|$

$= \frac{1}{2}\sum_{ij} E_X[(\vec{x}_j - \mu_{X_j})^t (\vec{x}_i - \mu_{X_i})](K_{X^2}^{-1})_{ij} + \frac{1}{2}\ln(2\pi)^n |K_{X^2}|$

$= \frac{1}{2}\sum_{ij}[(K_{X^2})_{ji}(K_{X^2}^{-1})_{ij}] + \frac{1}{2}\ln(2\pi)^n |K_{X^2}|$

$= \frac{1}{2}\sum_{j}[(K_{X^2})_{jj}(K_{X^2}^{-1})_{jj}] + \frac{1}{2}\ln(2\pi)^n |K_{X^2}|$

$= \frac{1}{2}\sum_{j} I_{jj} + \frac{1}{2}\ln(2\pi)^n |K_{X^2}|$

$= \frac{n}{2} + \frac{1}{2}\ln(2\pi)^n |K_{X^2}|$

$= \frac{1}{2}\ln(2\pi e)^n |K_{X^2}|$

A.2 Conditional differential entropy of two Gaussian random vectors

Theorem: *Given two concatenated Gaussian random vectors $\vec{x} = (\vec{x}_1, \vec{x}_2)$, of size k_1 and k_2 respectively, with a multivariate Gaussian distribution:*

$$P_X(\vec{x}) = \mathcal{N}(\mu_X, K_{X^2}) = (2\pi)^{-\frac{n}{2}} |K_{X^2}|^{-\frac{1}{2}} \exp\{-\frac{(\vec{x} - \vec{\mu_X})^t K_{X^2}^{-1} (\vec{x} - \vec{\mu_X})}{2}\}$$

with a mean vector μ_X and a covariance matrix K_{X^2}.

In this case, the conditional differential entropy $h(X_1|X_2)$ is equal to :

$$h(X_1|X_2) = \frac{1}{2} \ln(2\pi e)^{k_1} |K_{X_1^2|X_2}|$$

Proof:

$h(X_1|X_2) = -\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_{X_1X_2}(\vec{x_1}, \vec{x_2}) \ln\{p_{X_1|X_2}(\vec{x_1}, \vec{x_2})\} \overrightarrow{dx_1}\overrightarrow{dx_2}$

We know the conditional probability $P_{X_1|X_2}$ can be expressed as follows:

$P_{X_1|X_2}(\vec{x_1}, \vec{x_2}) = (2\pi)^{-\frac{k_1}{2}} |K_{X_1^2|X_2}|^{-\frac{1}{2}} \exp\{-\frac{(\vec{x_1} - \vec{\nu_{X_1/X_2}})^t K_{X_1^2|X_2}^{-1} (\vec{x_1} - \vec{\nu_{X_1/X_2}})}{2}\}$

So we can write:

$h(X_1|X_2)$

$= -\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_{X_1X_2}(\vec{x_1}, \vec{x_2})[-\frac{1}{2}(\vec{x_1} - \nu_{X_1|X_2})^t K_{X^2}^{-1} (\vec{x_1} - \nu_{X_1|X_2}) - \ln(\sqrt{2\pi})^{k_1} |K_{X_1^2|X_2}|^{\frac{1}{2}}] \overrightarrow{dx_1}\overrightarrow{dx_2}$

$= \frac{1}{2} E_{X_1X_2}[\sum_{ij} \{(\vec{x_1})_i - \nu_{(X_1|X_2)_i}\}^t (K_{X_1^2|X_2}^{-1})_{ij} \{(\vec{x_1})_j - \nu_{(X_1|X_2)_j}\}] + \frac{1}{2} \ln(2\pi)^{k_1} |K_{X_1^2|X_2}|$

$= \frac{1}{2} E_{X_1X_2}[\sum_{ij} \{(\vec{x_1})_i - \nu_{(X_1|X_2)_i}\}^t \{(\vec{x_1})_j - \nu_{(X_1|X_2)_j}\} (K_{X_1^2|X_2}^{-1})_{ij}] + \frac{1}{2} \ln(2\pi)^{k_1} |K_{X_1^2|X_2}|$

$= \frac{1}{2} \sum_{ij} E_{X_1X_2}[\{(\vec{x_1})_j - \nu_{(X_1|X_2)_j}\}^t \{(\vec{x_1})_i - \nu_{(X_1|X_2)_i}\}] (K_{X_1^2|X_2}^{-1})_{ij} + \frac{1}{2} \ln(2\pi)^{k_1} |K_{X_1^2|X_2}|$

$= \frac{1}{2} \sum_{ij} (K_{X_1^2|X_2})_{ji}) (K_{X_1^2|X_2}^{-1})_{ij}) + \frac{1}{2} \ln(2\pi)^{k_1} |K_{X_1^2|X_2}|$

$= \frac{1}{2} \sum_{j} (K_{X_1^2|X_2})_{jj}) (K_{X_1^2|X_2}^{-1})_{jj}) + \frac{1}{2} \ln(2\pi)^{k_1} |K_{X_1^2|X_2}|$

$= \frac{1}{2} \sum_{j} I_{jj} + \frac{1}{2} \ln(2\pi)^{k_1} |K_{X_1^2|X_2}|$

$= \frac{k_1}{2} + \frac{1}{2} \ln(2\pi)^{k_1} |K_{X_1^2|X_2}|$

$= \frac{1}{2} \ln(2\pi e)^{k_1} |K_{X_1^2|X_2}|$

[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.