Multiple causation and correlations

Ait-taleb Nabil [*]

Abstract

In the context of multiple causation, I will introduce the causation function. This function is a quadratic form computed from the correlations and serves as a generalization of R-squared, commonly found in machine learning. In this report, the causation function will make the link between the correlations and causal relationship. By examining the causation function through an illustrative example, we will demonstrate how strong or weak correlations between multiple causes and a variable can imply either a highly likely or unlikely causal relationship between the causes and the variable.

[*]Corresponding author: nabiltravail1982@gmail.com

1 Introduction

In this report, we delve into the concept of explained variance, which will play a crucial role in interpreting **the causation** function. This function, denoted as $\mathbb{P}(X|\Omega)$, is computed using a quadratic form computed from correlations. We will trace its origins through a rigorous proof, revealing why this causation function serves as a generalization of R-squared in the context of multiple causation within machine learning.

In what follows, I will prove that the causation function $\mathbb{P}(X|\Omega)$ quantifies **the causal effect** of causes Ω on a variable X. The causation function therefore makes the link between the correlations and the causal relationship.

To facilitate our understanding of correlation and causation, I will present a table that showcases the magnitude of correlations alongside their corresponding causation levels.

Subsequently, we explore a scenario involving two causes acting on a variable. By delineating the correlation pairs associated with highly likely and unlikely causation, we shed light on the intricate relationship between these factors.

The paper concludes with numerical applications, specifically addressing a problem where two causes influence a variable. Through these examples, we establish connections between strong and weak correlations and the likelihood of causation.

2 Conditional variance and explained variance EV for a multivariate Gaussian

We will recall below that for a multivariate Gaussian, we can compute the conditional variance from Shur's complement or the squared deviations between the response X and the Gaussian multiple linear regression of the variable X on the causes Ω:

$$K_{X^2|\Omega} = K_{X^2} - K_{X,\Omega}.K_{\Omega^2}^{-1}.K_{\Omega,X} = \frac{||X - \vec{\mu}_{X|\Omega}||^2}{N} = \frac{||X - \sum_{j=1}^{\#\Omega} \beta_{\omega_j X}\vec{\omega}_j - \beta_X||^2}{N}$$

Where we have:

$$\beta_{\Omega,X} = \begin{pmatrix} \beta_{\omega_1,X} \\ \beta_{\omega_2,X} \\ . \\ . \\ . \\ \beta_{\omega_{\#\Omega},X} \end{pmatrix} = K_{\Omega^2}^{-1}.K_{\Omega,X} = K_{\Omega^2}^{-1}. \begin{pmatrix} K_{\omega_1,X} \\ K_{\omega_2,X} \\ . \\ . \\ . \\ K_{\omega_{\#\Omega},X} \end{pmatrix}$$

$$\beta_X = \mu_X - K_{X,\Omega}.K_{\Omega^2}^{-1}.\mu_\Omega = \mu_X - .K_{X,\Omega}.K_{\Omega^2}^{-1}. \begin{pmatrix} \mu_{\omega_1} \\ \mu_{\omega_1} \\ . \\ . \\ . \\ \mu_{\omega_{\#\Omega}} \end{pmatrix}$$

The explained variance EV, quantifying the predictive quality of the causeses Ω on the variablle X, is defined as follows.

$$EV = 1 - \frac{K_{X^2|\Omega}}{K_{X^2}} = 1 - \frac{K_{X^2} - K_{X,\Omega}.K_{\Omega^2}^{-1}.K_{\Omega,X}}{K_{X^2}} = 1 - \frac{||X - \vec{\mu}_{X|\Omega}||^2}{||X - \mu_X||^2} = 1 - \frac{||X - \sum_{j=1}^{\#\Omega} \beta_{\omega_j X}\vec{\omega}_j - \beta_X||^2}{||X - \mu_X||^2}$$

3 Function of causation for the multiple causation

The causation of multiple causes Ω acting on a single variable X, computed from correlations, is expressed by the function of causation as follows: :

$$\boxed{\mathbb{P}(X|\Omega) = \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X}}$$

Where $\mathbb{P}(X|\Omega)$ is the function of causation.

$0 \leq \mathbb{P}(X|\Omega) \leq 1$ and $\#\Omega \geq 2$

$\tilde{K}_{X,\Omega}$ 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:

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:

$$\mathbb{P}(X|\Omega) = \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X} = 1 - \frac{K_{X^2|\Omega}}{K_{X^2}} = 1 - \frac{||X - \vec{\mu}_{X|\Omega}||^2}{||X - \mu_X||^2} = 1 - \frac{||X - \sum_{j=1}^{\#\Omega}\beta_{\omega_j X}\vec{\omega}_j - \beta_X||^2}{||X - \mu_X||^2}$$

The quadratic form $\mathbb{P}(X|\Omega) = \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X}$ can be considered as a generalization of R squared in machine learning. The relationship $\mathbb{P}(X|\Omega) = 1 - \frac{K_{X^2|\Omega}}{K_{X^2}}$ corresponds to the **explained variance** which quantifies the **prediction quality** for a Gaussian multiple linear regression. The relationship $\mathbb{P}(X|\Omega) = 1 - \frac{||X-\vec{\mu}_{X|\Omega}||^2}{||X-\mu_X||^2}$ shows us that $\mathbb{P}(X|\Omega)$ quantifies the **impact** of the causes Ω on the variable X. If the $\mathbb{P}(X|\Omega)$ value is close to 1 then we can estimate the impact is very important and that the relationship between Ω and X is a quasi-deterministic relationship. In the case where $\mathbb{P}(X|\Omega)$ is close to 0, we can estimate that the variable X is almost independent of the Ω causes. Using the Pythagorean Theorem:

$$||X - E(X)||^2 = ||E(X|\Omega) - E(X)||^2 + ||X - E(X|\Omega)||^2$$

It can be shown that $\mathbb{P}(X|\Omega)$ does indeed express a **causal effect** $||E(X|\Omega) - E(X)||^2$ which will be normalized:

$$0 \leq \mathbb{P}(X|\Omega) = \tilde{K}_{X,\Omega}\tilde{K}_{\Omega^2}^{-1}\tilde{K}_{\Omega,X} = 1 - \frac{||X - E(X|\Omega)||^2}{||X - E(X)||^2} = \frac{||E(X|\Omega) - E(X)||^2}{||X - E(X)||^2} \leq 1$$

4 Correlation value range

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

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

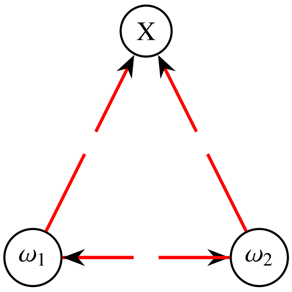
5 Causation value range

From the function of causation $\mathbb{P}(X|\Omega)$, we will present a table containing the magnitudes of causation:

Level of causation	$\mathbb{P}_{min}(X\|\Omega)$	$\mathbb{P}_{max}(X\|\Omega)$
Very unlikely	0	0.25
Unlikely	0.25	0.5
Likely	0.5	0.75
Very likely	0.75	1

6 Problem:Multiple 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:

$\tilde{K}_{\Omega^2} = \begin{pmatrix} 1 & \rho_{\omega_1\omega_2} \\ \rho_{\omega_1\omega_2} & 1 \end{pmatrix}$ 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 very likely causation:

$$0.75 \leq \mathbb{P}(X|\Omega) = \tilde{K}_{X,\Omega}.\tilde{K}_{\Omega^2}^{-1}.\tilde{K}_{\Omega,X} < 1$$

We will also show the representation for a unlikely causation:

$$0.25 \leq \mathbb{P}(X|\Omega) = \tilde{K}_{X,\Omega}.\tilde{K}_{\Omega^2}^{-1}.\tilde{K}_{\Omega,X} < 0.5$$

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. A pair of **strong correlations** between the causes Ω and the variable X that implies a **very likely causation** between the causes and the variable.
2. A pair of **weak correlations** between the causes Ω and the variable X that implies a **very likely causation** between the causes and the variable.
3. A pair of **strong correlations** between the causes Ω and the variable X that implies an **unlikely causation** between the causes and the variable.
4. A pair of **weak correlations** between the causes Ω and the variable X that implies an **unlikely causation** between the causes and variable.

7 Strong correlation, weak correlation and very likely 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 very likely causation $0.75 \leq \mathbb{P}(X|\Omega) < 1$:

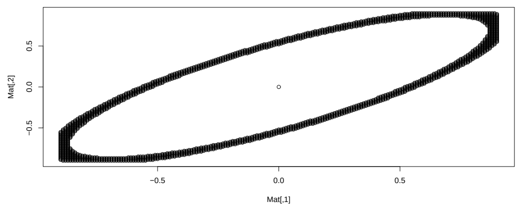


Figure 1: Pairs of correlations $\tilde{K}_{X,\Omega}$ having a very likely causation $0.75 \leq \mathbb{P}(X|\Omega) < 1$

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 function of causation $\mathbb{P}(X|\Omega)$ for the two points:

$$\mathbb{P}(X|\Omega) = (0.76, 0.86)\begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0.76 \\ 0.86 \end{pmatrix} = 0.754$$

$$\mathbb{P}(X|\Omega) = (0.21, -0.34)\begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0.21 \\ -0.34 \end{pmatrix} = 0.7609444$$

We can therefore describe two situations:

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

8 Strong correlation, weak correlation and unlikely 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 unlikely causation $0.25 \leq \mathbb{P}(X|\Omega) < 0.5$:

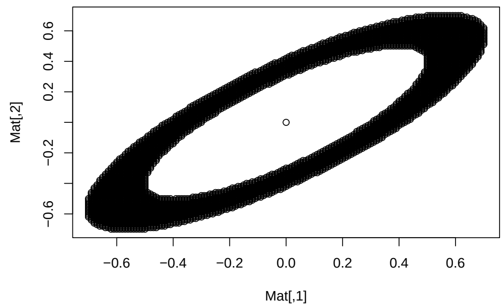


Figure 2: Pairs of correlations $\tilde{K}_{X,\Omega}$ having a unlikely causation $0.25 \leq \mathbb{P}(X|\Omega) < 0.5$

From this graph we will select two points:$\tilde{K}_{X,\Omega} = (0.70, 0.61)$ and $\tilde{K}_{X,\Omega} = (0.22, -0.2)$. We will compute the function of causation $\mathbb{P}(X|\Omega)$ for the two points:

$$\mathbb{P}(X|\Omega) = (0.70, 0.61) \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0.70 \\ 0.61 \end{pmatrix} = 0.4969444$$

$$\mathbb{P}(X|\Omega) = (0.22, -0.2) \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0.22 \\ -0.2 \end{pmatrix} = 0.4411111$$

We can therefore describe two situations:

1. A pair of **strong correlations** between the causes and the variable that implies an **unlikely causation** between the causes and variable.
2. A pair of **weak correlations** between the causes and the variable that implies an **unlikely causation** between the causes and the variable.

9 Conclusion

In this paper, we have explored the relationship between the concepts of causation and correlation for multiple causes acting on a variable. Using the example of two causes acting on a variable, we have illustrated the various scenarios that may arise:

1. A pair of **strong correlations** between the causes Ω and the variable X that implies a **very likely causation** between the causes and the variable.
2. A pair of **weak correlations** between the causes Ω and the variable X that implies a **very likely causation** between the causes and the variable.
3. A pair of **strong correlations** between the causes Ω and the variable X that implies an **unlikely causation** between the causes and the variable.
4. A pair of **weak correlations** between the causes Ω and the variable X that implies an **unlikely causation** between the causes and variable.

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