Directed dependency graph obtained from a correlation matrix by the highest successive conditionings method

Ait-taleb Nabil*

December 26, 2021

Abstract

In this paper we will propose a directed dependency graph obtained from a correlation matrix. This graph will include probabilistic causal sub-models for each node modeled by conditionings percentages $\lambda_{X_j|\Omega_1}$ where X_j and Ω_1 will correspond respectively to a child node and parents nodes set. The directed dependency graph will be obtained using the highest successive conditionings method with a conditioning percentage value of 90% to be exceeded.

*Corresponding author: nabiltravail1982@gmail.com

1 Introduction

In this report, we will propose a directed dependency graph obtained from a correlation matrix. This graph will have probabilistic causal sub-models for each node of the graph which will be evaluated by the conditioning percentage[4].

The article will begin by making a reminder on the notion of conditioning percentage for the General case and Gaussian case [4]. In what follows, we will expose the existing relationship between the conditioning percentage and the correlations in order to show that all the operations can be done from the correlations and not only from the variances-covariances [4] to obtain the directed dependency graph .

We will then define the directed dependency graph and we will explain how to obtain the graph from the highest successive conditionings method with a conditioning percentage value to be exceeded. We will also define the notion of directed dependency graph density to compare the graph obtained from a correlation matrix and the associated fully connected graph.

The report will end with the application of the learning algorithm of the directed dependency graph obtained from a correlation matrix for the *Harman23* and *Ability* matrices that can be found in the R software.

2 General conditioning percentage

We use the inequalities $h(X_3|X_1, X_2) \leq h(X_3|X_1) \leq h(X_3)$ to define the conditioning percentage [4].

Definition: *Given a set of variables $\Omega \equiv X_{i=1,\dots,n}$, the variable $X_j \in \Omega$, the subsets $\Omega_1 \subset \Omega \backslash X_j$ and $\Omega_2 \equiv \Omega \backslash \{X_j, \Omega_1\}$ we can define the conditioning percentage $\lambda_{X_j|\Omega_1}$ of variables Ω_1 which act on the variable X_j as follows:*

$$\lambda_{X_j|\Omega_1} = \frac{h(X_j) - h(X_j|\Omega_1)}{h(X_j) - h(X_j|\Omega_1, \Omega_2)}$$

$$0 \leq \lambda_{X_j|\Omega_1} \leq 1$$

From inequalities $h(X_j|\Omega_1, \Omega_2) \leq h(X_j|\Omega_1) \leq h(X_j)$, if we have the equality: $h(X_j) = h(X_j|\Omega_1, \Omega_2)$ then $h(X_j|\Omega_1, \Omega_2) = h(X_j|\Omega_1) = h(X_j)$. In this case, $X_j \perp \Omega_1 \cup \Omega_2$ and $\lambda_{X_j|\Omega_1} = 0$. (where the symbol $\perp$ corresponds to the independency symbol)

If $h(X_j) \neq h(X_j|\Omega_1, \Omega_2)$ and $h(X_j) = h(X_j|\Omega_1)$ then $X_j \perp \Omega_1$ and $\lambda_{X_j|\Omega_1} = 0$

3 Gaussian conditioning percentage

Definition: *Given a set of Gaussian variables $\Omega \equiv X_{i=1,\dots,n}$, the Gaussian variable $X_j \in \Omega$, the subsets $\Omega_1 \subset \Omega \backslash X_j$ and $\Omega_2 \equiv \Omega \backslash \{X_j, \Omega_1\}$ we can define the Gaussian conditioning percentage $\lambda_{X_j|\Omega_1}$ of the Gaussian variables Ω_1 which act on the Gaussian variable X_j as follows:*

$$\lambda_{X_j|\Omega_1} = \frac{h(X_j) - h(X_j|\Omega_1)}{h(X_j) - h(X_j|\Omega_1, \Omega_2)} = \frac{\frac{1}{2}\ln(2\pi.e.K_{X_j^2}) - \frac{1}{2}\ln(2\pi.e.K_{X_j^2|\Omega_1})}{\frac{1}{2}\ln(2\pi.e.K_{X_j^2}) - \frac{1}{2}\ln(2\pi.e.K_{X_j^2|\Omega_1,\Omega_2})}$$

In what follows, we will consider the gaussian conditioning percentage.

4 Conditioning percentage and correlations

We will now establish the relation between the conditioning percentage and the correlations

$$\lambda_{X_j|\Omega_1} = \frac{\ln(\tilde{K}_{X_j^2|\Omega_1})}{\ln(\tilde{K}_{X_j^2|\Omega_1,\Omega_2})} = \frac{\ln(1 - \tilde{K}^t_{X_j;\Omega_1}.\tilde{K}^{-1}_{\Omega_1^2}.\tilde{K}_{X_j;\Omega_1})}{\ln(1 - \tilde{K}^t_{X_j;\Omega_1,\Omega_2}.\tilde{K}^{-1}_{(\Omega_1,\Omega_2)^2}.\tilde{K}_{X_j;\Omega_1,\Omega_2})}$$

We have:

$$\lambda_{X_j|\Omega_1} = \frac{\frac{1}{2}\ln(2\pi.e.K_{X_j^2}) - \frac{1}{2}\ln(2\pi.e.K_{X_j^2|\Omega_1})}{\frac{1}{2}\ln(2\pi.e.K_{X_j^2}) - \frac{1}{2}\ln(2\pi.e.K_{X_j^2|\Omega_1,\Omega_2})} = \frac{\ln(\frac{K_{X_j^2|\Omega_1}}{K_{X_j^2}})}{\ln(\frac{K_{X_j^2|\Omega_1,\Omega_2}}{K_{X_j^2}})}$$

and

$$K_{X_j^2|\Omega_1} = K_{X_j^2} - K^t_{X_j;\Omega_1}.K^{-1}_{\Omega_1^2}.K_{X_j;\Omega_1}$$

$$K_{X_j^2|\Omega_1} = K_{X_j^2} - K^t_{X_j;\Omega_1}.(diag^{-1}(K_{\Omega_1^2}))^{\frac{1}{2}}.\tilde{K}^{-1}_{\Omega_1^2}.(diag^{-1}(K_{\Omega_1^2}))^{\frac{1}{2}}.K_{X_j;\Omega_1}$$

$$K_{X_j^2|\Omega_1} = K_{X_j^2} - K^{\frac{1}{2}}_{X_j^2}.\tilde{K}^t_{X_j;\Omega_1}.\tilde{K}^{-1}_{\Omega_1^2}.K^{\frac{1}{2}}_{X_j^2}.\tilde{K}_{X_j;\Omega_1}$$

$$K_{X_j^2|\Omega_1} = K_{X_j^2}(1 - \tilde{K}^t_{X_j;\Omega_1}.\tilde{K}^{-1}_{\Omega_1^2}.\tilde{K}_{X_j;\Omega_1})$$

$$K_{X_j^2|\Omega_1,\Omega_2} = K_{X_j^2} - K^t_{X_j;\Omega_1,\Omega_2}.K^{-1}_{(\Omega_1,\Omega_2)^2}.K_{X_j;\Omega_1,\Omega_2}$$

$$K_{X_j^2|\Omega_1,\Omega_1} = K_{X_j^2} - K^t_{X_j;\Omega_1,\Omega_2}.(diag^{-1}(K_{(\Omega_1,\Omega_2)^2}))^{\frac{1}{2}}.\tilde{K}^{-1}_{(\Omega_1,\Omega_2)^2}.(diag^{-1}(K_{(\Omega_1,\Omega_2)^2}))^{\frac{1}{2}}.K_{X_j;\Omega_1,\Omega_2}$$

$$K_{X_j^2|\Omega_1,\Omega_2} = K_{X_j^2} - K^{\frac{1}{2}}_{X_j^2}.\tilde{K}^t_{X_j;\Omega_1,\Omega_2}.\tilde{K}^{-1}_{(\Omega_1,\Omega_2)^2}.K^{\frac{1}{2}}_{X_j^2}.K_{X_j;\Omega_1,\Omega_2}$$

$$K_{X_j^2|\Omega_1,\Omega_2} = K_{X_j^2}(1 - \tilde{K}^t_{X_j;\Omega_1,\Omega_2}.\tilde{K}^{-1}_{(\Omega_1,\Omega_2)^2}.\tilde{K}_{X_j;\Omega_1,\Omega_2})$$

If we put:

$$\tilde{K}_{X_j^2|\Omega_1} = 1 - \tilde{K}^t_{X_j;\Omega_1}.\tilde{K}^{-1}_{\Omega_1^2}.\tilde{K}_{X_j;\Omega_1}$$

$$\tilde{K}_{X_j^2|\Omega_1,\Omega_2} = 1 - \tilde{K}^t_{X_j;\Omega_1,\Omega_2}.\tilde{K}^{-1}_{(\Omega_1,\Omega_2)^2}.\tilde{K}_{X_j;\Omega_1,\Omega_2}$$

we obtain:

$$\lambda_{X_j|\Omega_1} = \frac{\ln(\frac{K_{X_j^2|\Omega_1}}{K_{X_j^2}})}{\ln(\frac{K_{X_j^2|\Omega_1,\Omega_2}}{K_{X_j^2}})} = \frac{\ln(\tilde{K}_{X_j^2|\Omega_1})}{\ln(\tilde{K}_{X_j^2|\Omega_1,\Omega_2})} = \frac{\ln(1 - \tilde{K}^t_{X_j;\Omega_1}.\tilde{K}^{-1}_{\Omega_1^2}.\tilde{K}_{X_j;\Omega_1})}{\ln(1 - \tilde{K}^t_{X_j;\Omega_1,\Omega_2}.\tilde{K}^{-1}_{(\Omega_1,\Omega_2)^2}.\tilde{K}_{X_j;\Omega_1,\Omega_2})}$$

5 Directed dependency graph

Definition:

The directed dependency graph is a directed graph to which we attribute for each node a random variable and the conditioning percentage linked to the edges going from the set of nodes Ω_1 *to the node* X_j*:*

$$\lambda_{X_j|\Omega_1} = \frac{h(X_j) - h(X_j|\Omega_1)}{h(X_j) - h(X_j|\Omega_1, \Omega_2)} = C_1 E_\Omega[\ln P_{X_j|\Omega_1}(x_j, \vec{\omega}_1)] + C_2$$

$$0 \leq \lambda_{X_j|\Omega_1} \leq 1$$

Where $C_1 = \frac{1}{h(X_j) - h(X_j|\Omega_1,\Omega_2)} > 0$ *and* $C_2 = \frac{h(X_j)}{h(X_j) - h(X_j|\Omega_1,\Omega_2)} > 0$ *are positive constants for each node if* $h(X_j) \neq h(X_j|\Omega_1, \Omega_2)$.

If $\lambda_{X_j|\Omega_1} = 0$, $h(X_j) = h(X_j|\Omega_1)$ *and* $h(X_j) \neq h(X_j|\Omega_1, \Omega_2)$ *then* $X_j \perp \Omega_1$

If $\lambda_{X_j|\Omega_1} = 0$ *and* $h(X_j) = h(X_j|\Omega_1) = h(X_j|\Omega_1, \Omega_2)$ *then* $X_j \perp \Omega_1 \cup \Omega_2$.

For each node X_j**, we can see that the conditioning percentage** $\lambda_{X_j|\Omega_1}$ **is an affine application of the average conditional probability** $E_\Omega[\ln P_{X_j|\Omega_1}(x_j, \vec{\omega}_1)]$**, this is the reason why we can say that we have probabilistic causal sub-models evaluated by the conditioning percentage for each node of the directed dependency graph. In what follows, we will compute the conditionings percentage from correlations:**

$$\lambda_{X_j|\Omega_1} = \frac{\ln(\tilde{K}_{X_j^2|\Omega_1})}{\ln(\tilde{K}_{X_j^2|\Omega_1,\Omega_2})} = \frac{\ln(1 - \tilde{K}^t_{X_j;\Omega_1}.\tilde{K}^{-1}_{\Omega_1^2}.\tilde{K}_{X_j;\Omega_1})}{\ln(1 - \tilde{K}^t_{X_j;\Omega_1,\Omega_2}.\tilde{K}^{-1}_{(\Omega_1,\Omega_2)^2}.\tilde{K}_{X_j;\Omega_1,\Omega_2})} = C_1 E_\Omega[\ln P_{X_j|\Omega_1}(x_j, \vec{\omega}_1)] + C_2$$

For each node, the relation above relates the correlations to the probabilistic causal sub-model evaluated by the conditioning percentage.

6 The highest successive conditionings method

Among all the dependency graphs learned from the correlation matrix, we will choose the one using the highest successive conditionings method for each node. This method allows to obtain the smallest subset of parent nodes that most strongly condition the child nodes with a conditioning percentage value to be exceeded. For each node X_i, this method can be described by the following optimization problem:

$$\max_{X_j} \lambda_{X_i|X_j}$$

$$\max_{X_k} \lambda_{X_i|X_j,X_k}$$

$$\max_{X_l} \lambda_{X_i|X_j,X_k,X_l}$$

$$\vdots \qquad\qquad \vdots$$

For this optimization problem, we will define a threshold to be exceeded by the conditioning percentage for each node in order to consider the dependencies model as good. This threshold value will be set to 90% ($\lambda = 0.9$).

In what follows, we will propose the learning method applied to the *Harman23* and *Ability* correlation matrices that can be found in the R software.

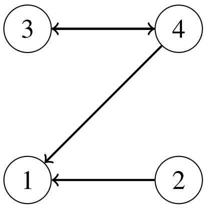
7 Directed dependency graph density

Definition:

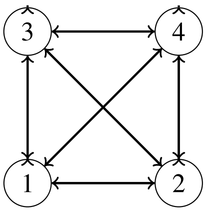
The directed dependency graph density is the number of the directed dependency graph's edges divided by the number of the fully connected graph's edges assigned to the directed dependency graph.

Example:

for example, the directed depedency graph below has 4 edges



and the fully connected graph assigned to the directed dependency graph has 12 edges



the directed dependency graph density is therefore equal to:

$$Density = \frac{4}{12} = 0.3333333$$

8 Directed dependency graph obtained from the Harman23 correlation matrix

In what follows we will consider the *Harman23* correlation matrix which can be found in R software. In the following, we will use the highest successive conditionings method. We will set a percentage value of $90\%(\lambda = 0.9)$ to be exceeded. We will show step by step the operations that allow us to obtain the directed dependency graph from a correlation matrix.

$$\tilde{K}_{(X_1,X_2,X_3,X_4,X_5,X_6,X_7,X_8)^2} = \begin{pmatrix} 1.000 & 0.846 & 0.805 & 0.859 & 0.473 & 0.398 & 0.301 & 0.382 \\ 0.846 & 1.000 & 0.881 & 0.826 & 0.376 & 0.326 & 0.277 & 0.415 \\ 0.805 & 0.881 & 1.000 & 0.801 & 0.380 & 0.319 & 0.237 & 0.345 \\ 0.859 & 0.826 & 0.801 & 1.000 & 0.436 & 0.329 & 0.327 & 0.365 \\ 0.473 & 0.376 & 0.380 & 0.436 & 1.000 & 0.762 & 0.730 & 0.629 \\ 0.398 & 0.326 & 0.319 & 0.329 & 0.762 & 1.000 & 0.583 & 0.577 \\ 0.301 & 0.277 & 0.237 & 0.327 & 0.730 & 0.583 & 1.000 & 0.539 \\ 0.382 & 0.415 & 0.345 & 0.365 & 0.629 & 0.577 & 0.539 & 1.000 \end{pmatrix}$$

$X_1 = Height$

$X_2 = Arm.span$

$X_3 = Forearm$

$X_4 = Lower.leg$

$X_5 = Weight$

$X_6 = Bitro.diameter$

$X_7 = Chest.girth$

$X_8 = Chest.width$

8.1 Conditioning percentage for the node X_8

$$\lambda_{X_8|X_1} = \frac{\ln(\tilde{K}_{X_8^2|X_1})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.2427854$$

$$\lambda_{X_8|X_2} = \frac{\ln(\tilde{K}_{X_8^2|X_2})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.2909297$$

$$\lambda_{X_8|X_3} = \frac{\ln(\tilde{K}_{X_8^2|X_3})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.1950563$$

$$\lambda_{X_8|X_4} = \frac{\ln(\tilde{K}_{X_8^2|X_4})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.220068$$

$$\lambda_{X_8|X_5} = \frac{\ln(\tilde{K}_{X_8^2|X_5})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.7751195$$

$$\lambda_{X_8|X_6} = \frac{\ln(\tilde{K}_{X_8^2|X_6})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.6231575$$

$$\lambda_{X_8|X_7} = \frac{\ln(\tilde{K}_{X_8^2|X_7})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.52829$$

The maximum conditioning percentage is $\lambda_{X_8|X_5} = 0.7751195$

$$\lambda_{X_8|X_5,X_1} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_1})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.7987161$$

$$\lambda_{X_8|X_5,X_2} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_2})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.8726499$$

$$\lambda_{X_8|X_5,X_3} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_3})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.808921$$

$$\lambda_{X_8|X_5,X_4} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_4})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.8012409$$

$$\lambda_{X_8|X_5,X_6} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_6})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.8342121$$

$$\lambda_{X_8|X_5,X_7} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_7})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.8102651$$

The maximum conditioning percentage is $\lambda_{X_8|X_5,X_2} = 0.8726499$

$$\lambda_{X_8|X_5,X_2,X_1} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_2,X_1})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.9043183$$

$$\lambda_{X_8|X_5,X_2,X_3} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_2,X_3})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.9005535$$

$$\lambda_{X_8|X_5,X_2,X_4} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_2,X_4})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.8926957$$

$$\lambda_{X_8|X_5,X_2,X_6} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_2,X_6})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.9255975$$

$$\lambda_{X_8|X_5,X_2,X_7} = \frac{\ln(\tilde{K}_{X_8^2|X_5,X_2,X_7})}{\ln(\tilde{K}_{X_8^2|X_1,X_2,X_3,X_4,X_5,X_6,X_7})} = 0.9096271$$

We consider the maximum conditioning percentage $\lambda_{X_8|X_5,X_2,X_6} = 0.9255975$ for the node X_8.

8.2 Conditioning percentage for the node X_7

$$\lambda_{X_7|X_1} = \frac{\ln(\tilde{K}_{X_7^2|X_1})}{\ln(\tilde{K}_{X_7^2|X_1,X_2,X_3,X_4,X_5,X_6,X_8})} = 0.1149814$$

$$\lambda_{X_7|X_2} = \frac{\ln(\tilde{K}_{X_7^2|X_2})}{\ln(\tilde{K}_{X_7^2|X_1,X_2,X_3,X_4,X_5,X_6,X_8})} = 0.09665281$$

$$\lambda_{X_7|X_3} = \frac{\ln(\tilde{K}_{X_7^2|X_3})}{\ln(\tilde{K}_{X_7^2|X_1,X_2,X_3,X_4,X_5,X_6,X_8})} = 0.06998806$$

$$\lambda_{X_7|X_4} = \frac{\ln(\tilde{K}_{X_7^2|X_4})}{\ln(\tilde{K}_{X_7^2|X_1,X_2,X_3,X_4,X_5,X_6,X_8})} = 0.1369166$$

$$\lambda_{X_7|X_5} = \frac{\ln(\tilde{K}_{X_7^2|X_5})}{\ln(\tilde{K}_{X_7^2|X_1,X_2,X_3,X_4,X_5,X_6,X_8})} = 0.9215957$$

$$\lambda_{X_7|X_6} = \frac{\ln(\tilde{K}_{X_7^2|X_5})}{\ln(\tilde{K}_{X_7^2|X_1,X_2,X_3,X_4,X_5,X_6,X_8})} = 0.502859$$

$$\lambda_{X_7|X_8} = \frac{\ln(\tilde{K}_{X_7^2|X_8})}{\ln(\tilde{K}_{X_7^2|X_1,X_2,X_3,X_4,X_5,X_6,X_8})} = 0.4155402$$

We consider the maximum conditioning percentage $\lambda_{X_7|X_5} = 0.9215957$ for the node X_7.

8.3 Conditioning percentage for the node X_6

$$\lambda_{X_6|X_1} = \frac{\ln(\tilde{K}_{X_6^2|X_1})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5,X_7,X_8})} = 0.1860926$$

$$\lambda_{X_6|X_2} = \frac{\ln(\tilde{K}_{X_6^2|X_2})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5,X_7,X_8})} = 0.1212434$$

$$\lambda_{X_6|X_3} = \frac{\ln(\tilde{K}_{X_6^2|X_3})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5,X_7,X_8})} = 0.1158057$$

$$\lambda_{X_6|X_4} = \frac{\ln(\tilde{K}_{X_6^2|X_4})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5,X_7,X_8})} = 0.1236185$$

$$\lambda_{X_6|X_5} = \frac{\ln(\tilde{K}_{X_6^2|X_5})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5,X_7,X_8})} = 0.9377567$$

$$\lambda_{X_6|X_7} = \frac{\ln(\tilde{K}_{X_6^2|X_7})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5,X_7,X_8})} = 0.4481922$$

$$\lambda_{X_6|X_8} = \frac{\ln(\tilde{K}_{X_6^2|X_8})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5,X_7,X_8})} = 0.4368743$$

We consider the maximum conditioning percentage $\lambda_{X_6|X_5} = 0.9377567$ for the node X_6.

8.4 Conditioning percentage for the node X_5

$$\lambda_{X_5|X_1} = \frac{\ln(\tilde{K}_{X_5^2|X_1})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.1833009$$

$$\lambda_{X_5|X_2} = \frac{\ln(\tilde{K}_{X_5^2|X_2})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.1103222$$

$$\lambda_{X_5|X_3} = \frac{\ln(\tilde{K}_{X_5^2|X_3})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.1128758$$

$$\lambda_{X_5|X_4} = \frac{\ln(\tilde{K}_{X_5^2|X_4})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.1526023$$

$$\lambda_{X_5|X_6} = \frac{\ln(\tilde{K}_{X_5^2|X_6})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.6289935$$

$$\lambda_{X_5|X_7} = \frac{\ln(\tilde{K}_{X_5^2|X_7})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.5509529$$

$$\lambda_{X_5|X_8} = \frac{\ln(\tilde{K}_{X_5^2|X_8})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.3644881$$

The maximum conditioning percentage is $\lambda_{X_5|X_6} = 0.6289935$

$$\lambda_{X_5|X_6,X_1} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_1})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.6906199$$

$$\lambda_{X_5|X_6,X_2} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_2})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.6611338$$

$$\lambda_{X_5|X_6,X_3} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_3})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.6659439$$

$$\lambda_{X_5|X_6,X_4} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_4})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.6987019$$

$$\lambda_{X_5|X_6,X_7} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_7})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.8819717$$

$$\lambda_{X_5|X_6,X_8} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_8})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.7282387$$

The maximum conditioning percentage is $\lambda_{X_5|X_6,X_7} = 0.8819717$

$$\lambda_{X_5|X_6,X_7,X_1} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_7,X_1})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.9417737$$

$$\lambda_{X_5|X_6,X_7,X_2} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_7,X_2})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.9047784$$

$$\lambda_{X_5|X_6,X_7,X_3} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_7,X_3})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.9190013$$

$$\lambda_{X_5|X_6,X_7,X_4} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_7,X_4})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.9290041$$

$$\lambda_{X_5|X_6,X_7,X_8} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_7,X_8})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6,X_7,X_8})} = 0.9250194$$

We consider the maximum conditioning percentage $\lambda_{X_5|X_6,X_7,X_1} = 0.9417737$ for the node X_5.

8.5 Conditioning percentage for the node X_4

$$\lambda_{X_4|X_1} = \frac{\ln(\tilde{K}_{X_4^2|X_1})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.8620462$$

$$\lambda_{X_4|X_2} = \frac{\ln(\tilde{K}_{X_4^2|X_2})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.7381854$$

$$\lambda_{X_4|X_3} = \frac{\ln(\tilde{K}_{X_4^2|X_3})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.6606284$$

$$\lambda_{X_4|X_5} = \frac{\ln(\tilde{K}_{X_4^2|X_5})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.1357426$$

$$\lambda_{X_4|X_6} = \frac{\ln(\tilde{K}_{X_4^2|X_6})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.07375554$$

$$\lambda_{X_4|X_7} = \frac{\ln(\tilde{K}_{X_4^2|X_7})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.07280901$$

$$\lambda_{X_4|X_8} = \frac{\ln(\tilde{K}_{X_4^2|X_8})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.09205062$$

The maximum conditioning percentage is $\lambda_{X_4|X_1} = 0.8620462$.

$$\lambda_{X_4|X_1,X_2} = \frac{\ln(\tilde{K}_{X_4^2|X_1,X_2})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.9534023$$

$$\lambda_{X_4|X_1,X_3} = \frac{\ln(\tilde{K}_{X_4^2|X_1,X_3})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.9516869$$

$$\lambda_{X_4|X_1,X_5} = \frac{\ln(\tilde{K}_{X_4^2|X_1,X_5})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.864842$$

$$\lambda_{X_4|X_1,X_6} = \frac{\ln(\tilde{K}_{X_4^2|X_1,X_6})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.8625307$$

$$\lambda_{X_4|X_1,X_7} = \frac{\ln(\tilde{K}_{X_4^2|X_1,X_7})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.8748237$$

$$\lambda_{X_4|X_1,X_8} = \frac{\ln(\tilde{K}_{X_4^2|X_1,X_8})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6,X_7,X_8})} = 0.8659659$$

We consider the maximum conditioning percentage $\lambda_{X_4|X_1,X_2} = 0.9534023$ for the node X_4.

8.6 Conditioning percentage for the node X_3

$$\lambda_{X_3|X_1} = \frac{\ln(\tilde{K}_{X_3^2|X_1})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6,X_7,X_8})} = 0.64763$$

$$\lambda_{X_3|X_2} = \frac{\ln(\tilde{K}_{X_3^2|X_2})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6,X_7,X_8})} = 0.9283618$$

$$\lambda_{X_3|X_4} = \frac{\ln(\tilde{K}_{X_3^2|X_4})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6,X_7,X_8})} = 0.6364122$$

$$\lambda_{X_3|X_5} = \frac{\ln(\tilde{K}_{X_3^2|X_5})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6,X_7,X_8})} = 0.09672463$$

$$\lambda_{X_3|X_6} = \frac{\ln(\tilde{K}_{X_3^2|X_6})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6,X_7,X_8})} = 0.06656138$$

$$\lambda_{X_3|X_7} = \frac{\ln(\tilde{K}_{X_3^2|X_7})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6,X_7,X_8})} = 0.03585373$$

$$\lambda_{X_3|X_8} = \frac{\ln(\tilde{K}_{X_3^2|X_8})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6,X_7,X_8})} = 0.07859793$$

We consider the maximum conditioning percentage $\lambda_{X_3|X_2} = 0.9283618$ for the node X_3.

8.7 Conditioning percentage for the node X_2

$$\lambda_{X_2|X_1} = \frac{\ln(\tilde{K}_{X_2^2|X_1})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.6645944$$

$$\lambda_{X_2|X_3} = \frac{\ln(\tilde{K}_{X_2^2|X_3})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.7909034$$

$$\lambda_{X_2|X_4} = \frac{\ln(\tilde{K}_{X_2^2|X_4})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.6058329$$

$$\lambda_{X_2|X_5} = \frac{\ln(\tilde{K}_{X_2^2|X_5})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.08053883$$

$$\lambda_{X_2|X_6} = \frac{\ln(\tilde{K}_{X_2^2|X_6})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.05936856$$

$$\lambda_{X_2|X_7} = \frac{\ln(\tilde{K}_{X_2^2|X_7})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.04218238$$

$$\lambda_{X_2|X_8} = \frac{\ln(\tilde{K}_{X_2^2|X_8})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.09987233$$

The maximum conditioning percentage is $\lambda_{X_2|X_3} = 0.7909034$.

$$\lambda_{X_2|X_3,X_1} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_1})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.9341878$$

$$\lambda_{X_2|X_3,X_4} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_4})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.8960546$$

$$\lambda_{X_2|X_3,X_5} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_5})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.795612$$

$$\lambda_{X_2|X_3,X_6} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_6})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.7962427$$

$$\lambda_{X_2|X_3,X_7} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_7})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.8026673$$

$$\lambda_{X_2|X_3,X_8} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_8})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.8250287$$

We consider the maximum conditioning percentage $\lambda_{X_2|X_3,X_1} = 0.9341878$ for the node X_2.

8.8 Conditioning percentage for the node X_1

$$\lambda_{X_1|X_2} = \frac{\ln(\tilde{K}_{X_1^2|X_2})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.7425306$$

$$\lambda_{X_1|X_3} = \frac{\ln(\tilde{K}_{X_1^2|X_3})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.61644$$

$$\lambda_{X_1|X_4} = \frac{\ln(\tilde{K}_{X_1^2|X_4})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.7904523$$

$$\lambda_{X_1|X_5} = \frac{\ln(\tilde{K}_{X_1^2|X_5})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.1495081$$

$$\lambda_{X_1|X_6} = \frac{\ln(\tilde{K}_{X_1^2|X_6})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.1018088$$

$$\lambda_{X_1|X_7} = \frac{\ln(\tilde{K}_{X_1^2|X_7})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.05606627$$

$$\lambda_{X_1|X_8} = \frac{\ln(\tilde{K}_{X_1^2|X_8})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.09311882$$

The maximum conditioning percentage is $\lambda_{X_1|X_4} = 0.7904523$.

$$\lambda_{X_1|X_4,X_2} = \frac{\ln(\tilde{K}_{X_1^2|X_4,X_2})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.9398735$$

$$\lambda_{X_1|X_4,X_3} = \frac{\ln(\tilde{K}_{X_1^2|X_4,X_3})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.8833257$$

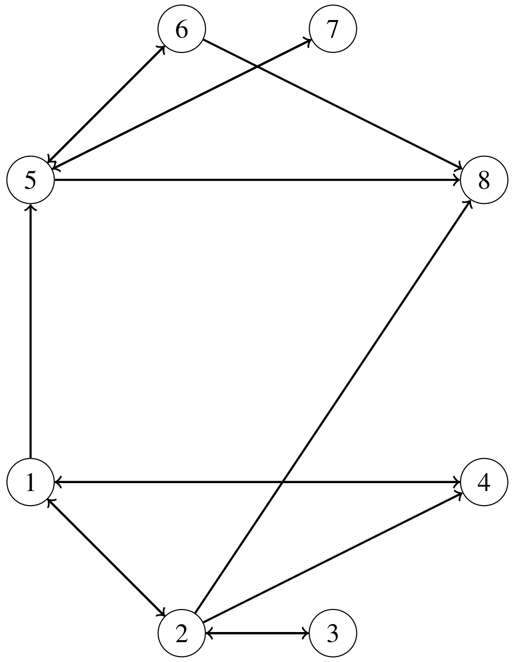
$$\lambda_{X_1|X_4,X_5} = \frac{\ln(\tilde{K}_{X_1^2|X_4,X_5})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.818055$$

$$\lambda_{X_1|X_4,X_6} = \frac{\ln(\tilde{K}_{X_1^2|X_4,X_6})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.8250753$$

$$\lambda_{X_1|X_4,X_7} = \frac{\ln(\tilde{K}_{X_1^2|X_4,X_7})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.7914728$$

$$\lambda_{X_1|X_4,X_8} = \frac{\ln(\tilde{K}_{X_1^2|X_4,X_8})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6,X_7,X_8})} = 0.8027596$$

We consider the maximum conditioning percentage $\lambda_{X_1|X_4,X_2} = 0.9398735$ for the node X_1.



$X_1 \not\perp \{X_2, X_4\}$ $\lambda_{X_1|X_2,X_4} = 0.9398735$

$X_2 \not\perp \{X_1, X_3\}$ $\lambda_{X_2|X_1,X_3} = 0.9341878$

$X_3 \not\perp \{X_2\}$ $\lambda_{X_3|X_2} = 0.9283618$

$X_4 \not\perp \{X_1, X_2\}$ $\lambda_{X_1|X_2} = 0.9534023$

$X_5 \not\perp \{X_1, X_6, X_7\}$ $\lambda_{X_5|X_1,X_6,X_7} = 0.9417737$

$X_6 \not\perp \{X_5\}$ $\lambda_{X_6|X_5} = 0.9377567$

$X_7 \not\perp \{X_5\}$ $\lambda_{X_7|X_5} = 0.9215957$

$X_8 \not\perp \{X_2, X_5, X_6\}$ $\lambda_{X_8|X_2,X_5,X_6} = 0.9255975$

$$Density = \frac{15}{56} = 0.2678571$$

9 Directed dependency graph obtained from the Ability correlation matrix

In what follows we will consider the *Ability* correlation matrix which can be found in R software. In the following, we will use the highest successive conditionings method. We will set a percentage value of $90\%(\lambda = 0.9)$ to be exceeded. We will show step by step the operations that allow us to obtain the directed dependency graph from a correlation matrix.

$$\tilde{K}_{(X_1,X_2,X_3,X_4,X_5,X_6)^2} = \begin{pmatrix} 1.0000000 & 0.4662649 & 0.5516632 & 0.3403250 & 0.5764799 & 0.5144058 \\ 0.4662649 & 1.0000000 & 0.5724364 & 0.1930992 & 0.2629229 & 0.2392766 \\ 0.5516632 & 0.5724364 & 1.0000000 & 0.4450901 & 0.3540252 & 0.3564715 \\ 0.3403250 & 0.1930992 & 0.4450901 & 1.0000000 & 0.1839645 & 0.2188370 \\ 0.5764799 & 0.2629229 & 0.3540252 & 0.1839645 & 1.0000000 & 0.7913779 \\ 0.5144058 & 0.2392766 & 0.3564715 & 0.2188370 & 0.7913779 & 1.0000000 \end{pmatrix}$$

$X_1 = General$

$X_2 = Picture$

$X_3 = Blocks$

$X_4 = Maze$

$X_5 = Reading$

$X_6 = Vocab$

9.1 Conditioning percentage for the node X_6

$$\lambda_{X_6|X_1} = \frac{\ln(\tilde{K}_{X_6^2|X_1})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5})} = 0.3038817$$

$$\lambda_{X_6|X_2} = \frac{\ln(\tilde{K}_{X_6^2|X_2})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5})} = 0.05829067$$

$$\lambda_{X_6|X_3} = \frac{\ln(\tilde{K}_{X_6^2|X_3})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5})} = 0.1343647$$

$$\lambda_{X_6|X_4} = \frac{\ln(\tilde{K}_{X_6^2|X_4})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5})} = 0.04851915$$

$$\lambda_{X_6|X_5} = \frac{\ln(\tilde{K}_{X_6^2|X_5})}{\ln(\tilde{K}_{X_6^2|X_1,X_2,X_3,X_4,X_5})} = 0.97311$$

We consider the maximum conditioning percentage $\lambda_{X_6|X_5} = 0.97311$ for the node X_6.

9.2 Conditioning percentage for the node X_5

$$\lambda_{X_5|X_1} = \frac{\ln(\tilde{K}_{X_5^2|X_1})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.3670778$$

$$\lambda_{X_5|X_2} = \frac{\ln(\tilde{K}_{X_5^2|X_2})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.06509369$$

$$\lambda_{X_5|X_3} = \frac{\ln(\tilde{K}_{X_5^2|X_3})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.1216866$$

$$\lambda_{X_5|X_4} = \frac{\ln(\tilde{K}_{X_5^2|X_4})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.03128545$$

$$\lambda_{X_5|X_6} = \frac{\ln(\tilde{K}_{X_5^2|X_6})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.8943826$$

The maximum conditioning percentage is $\lambda_{X_5|X_6} = 0.8943826$.

$$\lambda_{X_5|X_6,X_1} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_1})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.9945806$$

$$\lambda_{X_5|X_6,X_2} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_2})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.9084487$$

$$\lambda_{X_5|X_6,X_3} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_3})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.9089063$$

$$\lambda_{X_5|X_6,X_4} = \frac{\ln(\tilde{K}_{X_5^2|X_6,X_4})}{\ln(\tilde{K}_{X_5^2|X_1,X_2,X_3,X_4,X_6})} = 0.8946795$$

We consider the maximum conditioning percentage $\lambda_{X_5|X_6,X_1} = 0.9945806$ for the node X_5.

9.3 Conditioning percentage for the node X_4

$$\lambda_{X_4|X_1} = \frac{\ln(\tilde{K}_{X_4^2|X_1})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.480832$$

$$\lambda_{X_4|X_2} = \frac{\ln(\tilde{K}_{X_4^2|X_2})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.148435$$

$$\lambda_{X_4|X_3} = \frac{\ln(\tilde{K}_{X_4^2|X_3})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.8623935$$

$$\lambda_{X_4|X_5} = \frac{\ln(\tilde{K}_{X_4^2|X_5})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.1344845$$

$$\lambda_{X_4|X_6} = \frac{\ln(\tilde{K}_{X_4^2|X_6})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.191692$$

The maximum conditioning percentage is $\lambda_{X_4|X_3} = 0.8623935$.

$$\lambda_{X_4|X_3,X_1} = \frac{\ln(\tilde{K}_{X_4^2|X_3,X_1})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.9258143$$

$$\lambda_{X_4|X_3,X_2} = \frac{\ln(\tilde{K}_{X_4^2|X_3,X_2})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.8900614$$

$$\lambda_{X_4|X_3,X_5} = \frac{\ln(\tilde{K}_{X_4^2|X_3,X_5})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.8662743$$

$$\lambda_{X_4|X_3,X_6} = \frac{\ln(\tilde{K}_{X_4^2|X_3,X_6})}{\ln(\tilde{K}_{X_4^2|X_1,X_2,X_3,X_5,X_6})} = 0.8826522$$

We consider the maximum conditioning percentage $\lambda_{X_4|X_3,X_1} = 0.9258143$ for the node X_4.

9.4 Conditioning percentage for the node X_3

$$\lambda_{X_3|X_1} = \frac{\ln(\tilde{K}_{X_3^2|X_1})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.5182638$$

$$\lambda_{X_3|X_2} = \frac{\ln(\tilde{K}_{X_3^2|X_2})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.5670259$$

$$\lambda_{X_3|X_4} = \frac{\ln(\tilde{K}_{X_3^2|X_4})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.3153114$$

$$\lambda_{X_3|X_5} = \frac{\ln(\tilde{K}_{X_3^2|X_5})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.1912523$$

$$\lambda_{X_3|X_6} = \frac{\ln(\tilde{K}_{X_3^2|X_6})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.1940932$$

The maximum conditioning percentage is $\lambda_{X_3|X_2} = 0.5670259$.

$$\lambda_{X_3|X_2,X_1} = \frac{\ln(\tilde{K}_{X_3^2|X_2,X_1})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.8060573$$

$$\lambda_{X_3|X_2,X_4} = \frac{\ln(\tilde{K}_{X_3^2|X_2,X_4})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.838182$$

$$\lambda_{X_3|X_2,X_5} = \frac{\ln(\tilde{K}_{X_3^2|X_2,X_5})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.66482$$

$$\lambda_{X_3|X_2,X_6} = \frac{\ln(\tilde{K}_{X_3^2|X_2,X_6})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.6799385$$

The maximum conditioning percentage is $\lambda_{X_3|X_2,X_4} = 0.838182$.

$$\lambda_{X_3|X_2,X_4,X_1} = \frac{\ln(\tilde{K}_{X_3^2|X_2,X_4,X_1})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.9855479$$

$$\lambda_{X_3|X_2,X_4,X_5} = \frac{\ln(\tilde{K}_{X_3^2|X_2,X_4,X_5})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.9095238$$

$$\lambda_{X_3|X_2,X_4,X_6} = \frac{\ln(\tilde{K}_{X_3^2|X_2,X_4,X_6})}{\ln(\tilde{K}_{X_3^2|X_1,X_2,X_4,X_5,X_6})} = 0.9117135$$

We consider the maximum conditioning percentage $\lambda_{X_3|X_2,X_4,X_1} = 0.9855479$ for the node X_3.

9.5 Conditioning percentage for the node X_2

$$\lambda_{X_2|X_1} = \frac{\ln(\tilde{K}_{X_2^2|X_1})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.5301722$$

$$\lambda_{X_2|X_3} = \frac{\ln(\tilde{K}_{X_2^2|X_3})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.8586699$$

$$\lambda_{X_2|X_4} = \frac{\ln(\tilde{K}_{X_2^2|X_4})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.08218523$$

$$\lambda_{X_2|X_5} = \frac{\ln(\tilde{K}_{X_2^2|X_5})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.1549267$$

$$\lambda_{X_2|X_6} = \frac{\ln(\tilde{K}_{X_2^2|X_6})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.1275109$$

The maximum conditioning percentage is $\lambda_{X_2|X_3} = 0.8586699$.

$$\lambda_{X_2|X_3,X_1} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_1})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.9659894$$

$$\lambda_{X_2|X_3,X_4} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_4})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.8739891$$

$$\lambda_{X_2|X_3,X_5} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_5})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.8720691$$

$$\lambda_{X_2|X_3,X_6} = \frac{\ln(\tilde{K}_{X_2^2|X_3,X_6})}{\ln(\tilde{K}_{X_2^2|X_1,X_3,X_4,X_5,X_6})} = 0.8632458$$

We consider the maximum conditioning percentage $\lambda_{X_2|X_3,X_1} = 0.9659894$ for the node X_2.

9.6 Conditioning percentage for the node X_1

$$\lambda_{X_1|X_2} = \frac{\ln(\tilde{K}_{X_1^2|X_2})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.34696$$

$$\lambda_{X_1|X_3} = \frac{\ln(\tilde{K}_{X_1^2|X_3})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.5136137$$

$$\lambda_{X_1|X_4} = \frac{\ln(\tilde{K}_{X_1^2|X_4})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.1742261$$

$$\lambda_{X_1|X_5} = \frac{\ln(\tilde{K}_{X_1^2|X_5})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.5717525$$

$$\lambda_{X_1|X_6} = \frac{\ln(\tilde{K}_{X_1^2|X_6})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.4350263$$

The maximum conditioning percentage is $\lambda_{X_1|X_5} = 0.5717525$.

$$\lambda_{X_1|X_5,X_2} = \frac{\ln(\tilde{K}_{X_1^2|X_5,X_2})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.8174173$$

$$\lambda_{X_1|X_5,X_3} = \frac{\ln(\tilde{K}_{X_1^2|X_5,X_3})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.8997852$$

$$\lambda_{X_1|X_5,X_4} = \frac{\ln(\tilde{K}_{X_1^2|X_5,X_4})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.6976076$$

$$\lambda_{X_1|X_5,X_6} = \frac{\ln(\tilde{K}_{X_1^2|X_5,X_6})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.5910925$$

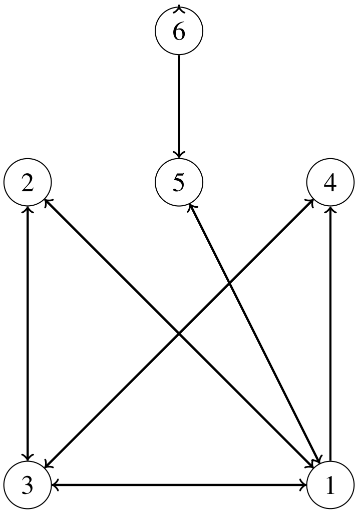
The maximum conditioning percentage is $\lambda_{X_1|X_5,X_3} = 0.8997852$.

$$\lambda_{X_1|X_5,X_3,X_2} = \frac{\ln(\tilde{K}_{X_1^2|X_5,X_3,X_2})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.9628853$$

$$\lambda_{X_1|X_5,X_3,X_4} = \frac{\ln(\tilde{K}_{X_1^2|X_5,X_3,X_4})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.9231183$$

$$\lambda_{X_1|X_5,X_3,X_6} = \frac{\ln(\tilde{K}_{X_1^2|X_5,X_3,X_6})}{\ln(\tilde{K}_{X_1^2|X_2,X_3,X_4,X_5,X_6})} = 0.9054526$$

We consider the maximum conditioning percentage $\lambda_{X_1|X_5,X_3,X_2} = 0.9628853$ for the node X_1.



$$X_1 \not\perp \{X_2, X_3, X_5\} \quad \lambda_{X_1|X_2,X_3,X_5} = 0.9628853$$

$$X_2 \not\perp \{X_1, X_3\} \quad \lambda_{X_2|X_1,X_3} = 0.9659894$$

$$X_3 \not\perp \{X_1, X_2, X_4\} \quad \lambda_{X_3|X_1,X_2,X_4} = 0.9855479$$

$$X_4 \not\perp \{X_1, X_3\} \quad \lambda_{X_4|X_1,X_3} = 0.9258143$$

$$X_5 \not\perp \{X_1, X_6\} \quad \lambda_{X_5|X_1,X_6} = 0.9945806$$

$$X_6 \not\perp \{X_5\} \quad \lambda_{X_6|X_5} = 0.97311$$

$$Density = \frac{13}{30} = 0.4333333$$

10 Conclusion

In this paper, we have shown how to obtain a directed dependency graph from a correlation matrix. The method was detailed using two practical examples with the *Harman23* and *Ability* correlation matrices.

References

[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] Directed dependency graph obtained from a continuous data matrix by the highest successive conditionings method. Year: 2021, Published:viXra. Author: Ait-Taleb Nabil.