

Conditional independance on extremal linear latent model

CAPEL Alexandre

2025-02-06

Contents

1	The latent linear model	1
1.1	Presentation	1
1.2	Induced properties	2
2	Conditionnal independance over Z	2
2.1	Heuristic and definition	2
2.2	Properties	3
2.3	Extremal independance	4
2.4	Extremal conditionnal independance	6

In this document, we want to define a new notion on conditional independence for a particular model and see if it corresponds to the “classic” extremal conditional independence.

Let $V = \{1, \dots, d\}$.

1 The latent linear model

1.1 Presentation

Let's consider a random vector X of \mathbb{R}^V (indexed by V) such that we have the following representation :

$$X = MZ + \varepsilon$$

where Z represents a unobservable random vector of \mathbb{R}^K (identified as the latent variable), M is a $d \times K$ matrix and $\varepsilon \in \mathbb{R}^d$ a random noise.

For the next assumptions, we will use the same context as (Boulin 2024), and so :

- K is not known and the dimensional parameters d and K can increase and be larger than n , the number of observation.
- all the components of the random vector Z are asymptotically independent with a tail index equal to one. It means that we can express the exponential measure as below :

$$\Lambda_Z = \sum_{k=1}^K \delta_0 \otimes \cdots \otimes \Lambda_{Z_k} \otimes \cdots \otimes \delta_0, \quad \Lambda_{Z_k}(dy) = y^{-2} dy$$

- the random noise ε possess a distribution with a tail that is lighter than the factors (*what does it mean*?).

1.2 Induced properties

Theses assumptions give to the random vector X the regular variation property and bring also a spectral measure Φ which have a discrete representation :

$$\Phi(\cdot) = \sum_{k=1}^K \|M_{\cdot,k}\| \delta_{\frac{M_{\cdot,k}}{\|M_{\cdot,k}\|}}(\cdot),$$

with $\delta_x(\cdot)$ the Dirac measure on x .

I don't really see why all of this is true...

Thus, we can compute the limit distribution for the maxima n replications of X :

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\bigvee_{j=1}^n c_n^{-1} X_n < x\right) = \exp\left[\sum_{k=1}^K \bigvee_{j=1}^d \left(\frac{M_{jk}}{x_j}\right)\right], \quad x \in \mathbb{R}_+^d$$

The maxima of vectors is defined as componentwise maxima.

I am sure this comes from the expression of the spectral measure Φ I need to understand that...

2 Conditionnal independance over Z

2.1 Heuristic and definition

In the definition of our model, we can see that for each component X_j , or each subvector X_A , we can build a list containing all the components of Z which define the latter component (or set of component).

For instance, if we define the model as below :

$$\begin{cases} X_1 = Z_1 + 2Z_4 + \varepsilon_1 \\ X_2 = 4Z_2 + Z_4 + \varepsilon_2 \\ X_3 = Z_2 - 3Z_3 + \varepsilon_3 \end{cases}$$

One way to define a model of conditional independence is to look if the variable

TODO Heuristic

Let $j \in V$ and $A \subset V$, let's define the index sets :

$$I_j = \{k \in [K] \mid m_{jk} \neq 0\} \quad \text{and} \quad I_A = \{k \in [K] \mid \forall j \in A, m_{jk} \neq 0\}$$

These sets represent the list of the component of Z as presented above.

Definition. Let $A, B, C \subset V$. We say that A and B are conditionally independent over Z according to C if and only if :

$$I_A \cap I_B \subset I_C \quad (\Leftrightarrow (I_A \cap I_B) \setminus I_C = \emptyset)$$

We will note $X_A \perp\!\!\!\perp_Z X_B \mid X_C$.

In the case where $C = \emptyset$, then we note that $X_A \perp\!\!\!\perp_Z X_B$ iif $I_A \cap I_B = \emptyset$.

2.2 Properties

For this new notion of independence, we would like to verify if some general properties on the ternary relation. In particular, we would like to know if the model of independence is a **semi-graphoid**, or even a **graphoid**.

2.2.1 Semi-graphoid case

So, we want to verify :

- the **Symmetry** : $X_A \perp\!\!\!\perp_Z X_B \mid X_C \Rightarrow X_B \perp\!\!\!\perp_Z X_A \mid X_C$.

Indeed, if $I_A \cap I_B \subset I_C$ then $I_B \cap I_A \subset I_C$.

- the **Decomposition** : $X_A \perp\!\!\!\perp_Z X_{B \cup D} \mid X_C \Rightarrow X_A \perp\!\!\!\perp_Z X_B \mid X_C$.

We have $I_{B \cup D} = I_B \cup I_D$ and thus $I_A \cap I_B \subset (I_A \cap I_B) \cup (I_A \cap I_D) = I_A \cap I_{B \cup D} \subset I_C$.

- the **Weak Union** : $X_A \perp\!\!\!\perp_Z X_{B \cup D} \mid X_C \Rightarrow X_A \perp\!\!\!\perp_Z X_B \mid X_{C \cup D}$.

By decomposition we know already that $I_A \cap I_B \subset I_C$ and $I_{C \cup D} = I_C \cup I_D$. Therefore $I_A \cap I_B \subset I_C \subset I_{C \cup D}$.

- the **Contraction** : $X_A \perp\!\!\!\perp_Z X_B \mid X_C$ and $X_A \perp\!\!\!\perp_Z X_D \mid X_{B \cup C} \Rightarrow X_A \perp\!\!\!\perp_Z X_{B \cup D} \mid X_C$.

We have $I_A \cap I_{B \cup D} = (I_A \cap I_B) \cup (I_A \cap I_D)$. First, we already have $I_A \cap I_B \subset I_C$ and $I_A \cap I_D \subset I_{B \cup C}$. These two relations induce that $I_A \cap I_D \subset I_C$ and thus $I_A \cap I_{B \cup D} \subset I_C$.

So our model of independence verify the four axioms of the semi-graphoid that is a very good point.

2.2.2 Graphoid case

Now we want to know if the model is a graphoid, which means that the relation verify :

Intersection : $X_A \perp\!\!\!\perp_Z X_B \mid X_{C \cup D}$ and $X_A \perp\!\!\!\perp_Z X_C \mid X_{B \cup D} \Rightarrow X_A \perp\!\!\!\perp_Z X_{B \cup C} \mid X_D$.

Our model is strongly linked to set theory. Let's see the following figure :

We can verify visually that the intersection property is not true in this type of configuration. In our context, we can also build an example which is inspired of the configuration to show that our model of independence is not always a graphoid.

$$\begin{cases} X_1 = Z_1 + 2Z_2 - Z_3 + \varepsilon_1 \\ X_2 = 4Z_2 + -2Z_3 + Z_5 + \varepsilon_2 \\ X_3 = Z_2 - 3Z_3 + Z_4 + \varepsilon_3 \\ X_4 = 10Z_4 + \varepsilon_4 \end{cases}$$

In this example, we can build the index set for each component :

$$I_1 = \{1, 2, 3\}, \quad I_2 = \{2, 3, 5\}, \quad I_3 = \{2, 3, 4\}, \quad I_4 = \{4\}$$

One can verify that the intersection property is not true for this model.

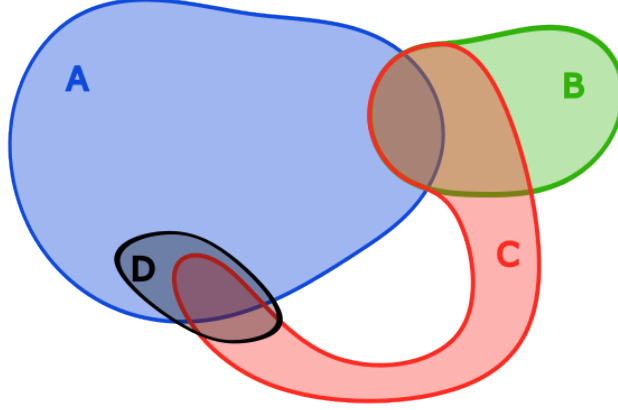


Figure 1: Intersection property not verified

2.3 Extremal independance

Assume that $C = \emptyset$. We want to show that the independence over Z is equivalent to the asymptotic independence for the corresponding Multivariate Generalized Pareto Distribution.

First, let's take a look to the exponential measure of X using the expression from above :

$$\Lambda_X(\mathcal{E} \setminus [0, x]) = \sum_{k=1}^K \bigvee_{j=1}^d \left(\frac{M_{jk}}{x_j} \right)$$

We can see quickly that this expression is not derivable and the exponent measure does not have a density function over the Lebesgue measure.

Thus, the model of independence that we will consider here is the conditional independence over a Levy measure Λ as defined in (Engelke, Ivanovs, and Stokorb 2024).

In particular, they show that there exist an equivalence between independence and the decomposition of the exponent measure.

Theorem. Let A, B a partition of V and assume that the exponent measure follows the explosiveness assumption. Then the following statements are equivalent :

- $A \perp\!\!\!\perp B \mid [\Lambda]$.
- $\Lambda(y_A \neq 0 \text{ and } y_B \neq 0) = 0$.

Thus, we need to show two things : first that our spectral measure verifies the explosiveness assumption, and then, that we can show the second statement of above is true.

2.3.1 Explosiveness assumption

We can recall the explosiveness assumption before show that our model verify it.

Let Λ an exponential measure defined on \mathbb{R}^V . We say that the measure Λ follows the explosiveness assumption if :

$$\Lambda(y_v \neq 0) = \Lambda_{\{v\}}((0, \infty)) \in \{0, \infty\}, \quad \forall v \in V$$

Fortunately, the one-dimensional marginal is easy to compute with a good choice of set :

$$\Lambda_{\{v\}}\left(\left(\frac{1}{n}, \infty\right)\right) = n \sum_{k=1}^K m_{jk}.$$

Thus if we consider the set $(0, \infty) = \bigcup_{i=1}^{\infty} (\frac{1}{n}, \infty)$, we can deduce by monotony of the measure that :

$$\Lambda(y_v \neq 0) = \lim_{n \rightarrow \infty} \Lambda_{\{v\}}\left(\left(\frac{1}{n}, \infty\right)\right) = \lim_{n \rightarrow \infty} n \sum_{k=1}^K m_{jk} = \begin{cases} 0, & \text{if } \sum_{k=1}^K m_{jk} = 0 \\ \infty, & \text{else} \end{cases}$$

□

2.3.2 Same independence

Now, we can use the theorem from (Engelke, Ivanovs, and Strokorb 2024) to show that our new independence relation is the same as them.

So, let's assume that $X_A \perp\!\!\!\perp_Z X_B$. Then we have $I_A \cap I_B = \emptyset$. Then we can rewrite the expression of the exponent measure as :

$$\Lambda(\mathcal{E} \setminus [0, x]) = \sum_{k \in I_A} \bigvee_{j \in A} \left(\frac{M_{jk}}{x_j}\right) + \sum_{k \in I_B} \bigvee_{j \in B} \left(\frac{M_{jk}}{x_j}\right) = \Lambda_A(\mathcal{E}_A \setminus [0, x_A]) + \Lambda_B(\mathcal{E}_B \setminus [0, x_B])$$

However, we also have the general relation :

$$\begin{aligned} \Lambda(\mathcal{E} \setminus [0, x]) &= \Lambda\left[(\mathcal{E}_A \setminus [0, x_A]) \times [0, x_B]^B\right] + \Lambda\left[[0, x_A]^A \times (\mathcal{E}_B \setminus [0, x_B])\right] + \Lambda\left[(\mathcal{E}_A \setminus [0, x_A]) \times (\mathcal{E}_B \setminus [0, x_B])\right] \\ &= \Lambda\left[(\mathcal{E}_A \setminus [0, x_A]) \times \mathcal{E}_B\right] + \Lambda\left[\mathcal{E}_A \times (\mathcal{E}_B \setminus [0, x_B])\right] - \Lambda\left[(\mathcal{E}_A \setminus [0, x_A]) \times (\mathcal{E}_B \setminus [0, x_B])\right] \\ &= \Lambda_A\left[\mathcal{E}_A \setminus [0, x_A]\right] + \Lambda_B\left[\mathcal{E}_B \setminus [0, x_B]\right] - \Lambda\left[(\mathcal{E}_A \setminus [0, x_A]) \times (\mathcal{E}_B \setminus [0, x_B])\right] \end{aligned}$$

So, by using the alternative formula of the exponent measure, we can deduce that :

$$\Lambda\left[(\mathcal{E}_A \setminus [0, x_A]) \times (\mathcal{E}_B \setminus [0, x_B])\right] = 0, \quad \forall (x_A, x_B) \in \mathcal{E}_A \times \mathcal{E}_B$$

Let's consider the family :

$$R_n = (\mathcal{E}_A \setminus [0, \frac{1}{n}]) \times (\mathcal{E}_B \setminus [0, \frac{1}{n}])$$

Of course, we have :

$$\bigcup_{n=1}^{\infty} R_n = \{y_A \neq 0 \text{ and } y_B \neq 0\}$$

Thus, by monotony of the measure and the fact that $\Lambda(R_n) = 0$, for all $n \in \mathbb{N}^*$, we get :

$$\Lambda(y_A \neq 0 \text{ and } y_B \neq 0) = \lim_{n \rightarrow \infty} \Lambda(R_n) = 0$$

□

Everything seems to be good to continue and we would like to see if our new conditional independence relation fits well with the conditional independence over a measure.

2.4 Extremal conditionnal independance

The goal of this section is to see if we have :

$$X_A \perp\!\!\!\perp_Z X_B \mid X_C \iff A \perp\!\!\!\perp B \mid C [A]$$

2.4.1 Simulation study

We would like to see if it is the case in a very simple case.

Let's consider a graph $\mathcal{G} = (V, E)$ which is a path. For $d = 10$, it corresponds to :

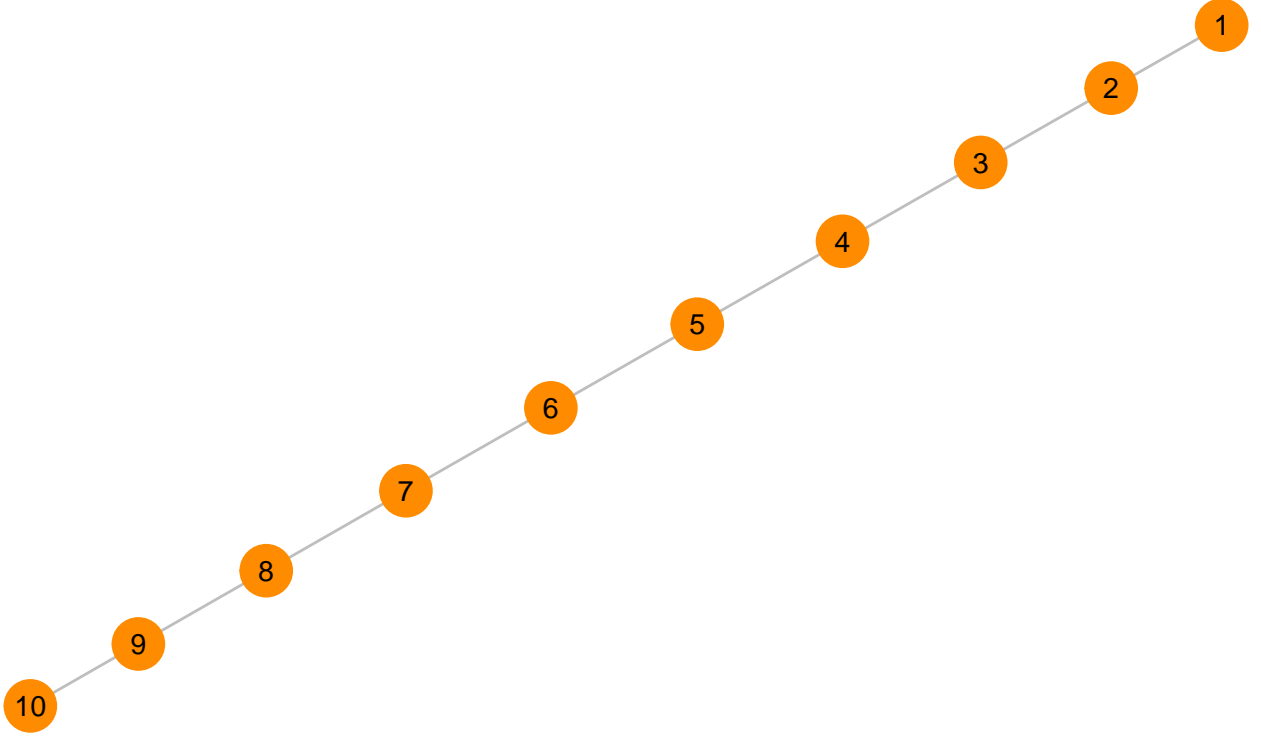


Figure 2: Path with 10 vertices.

We would like to simulate a simple matrix M such that the graphical model defined in the beginning respects the upper graph structure. We recall that the conditional independence relationship is based on the no-zero coefficient of the M matrix.

We can easily verify that a matrix with this type of format respect a graph structure of a path.

Setup and Methods

- for $m = 10$ vertices we need to simulate a latent random vector Z of size $K = 5 + 2(m - 1) = 23$ to respect the matrix format as presented above.
- we simulated a sample of $n = 10^5$.
- the seeking of the classic graphical model is done using the MST algorithm using empirical extremal coefficient as (Engelke and Volgushev 2022) (*we need to verify if we are in the context of the article*).

$$\begin{pmatrix}
 * & * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \square & \square & \square \\
 0 & 0 & * & * & * & * & * & 0 & 0 & 0 & 0 & 0 & \square & \square & \square \\
 0 & 0 & 0 & 0 & * & * & * & * & * & 0 & 0 & 0 & \square & \square & \square \\
 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & * & 0 & \square & \square & \square \\
 & & & \diamond & & & & & \diamond & & & & \diamond & & \\
 & & & & \diamond & & & & & \diamond & & & & \diamond & \\
 & & & & & \diamond & & & & & \diamond & & & & \diamond
 \end{pmatrix}$$

Figure 3: Matrix M format (stars are non zero values).

We can recall that the empirical estimator for χ_{ij} as :

$$\hat{\chi}_{ij} = \frac{1}{k} \sum_{l=1}^n \mathbb{1}_{\{\tilde{F}_i(X_{li}) > 1 - \frac{k}{n}, \tilde{F}_i(X_{lj}) > 1 - \frac{k}{n}\}}.$$

where \tilde{F}_i is the empirical distribution function of X_i .

At the end, we built the following graph :

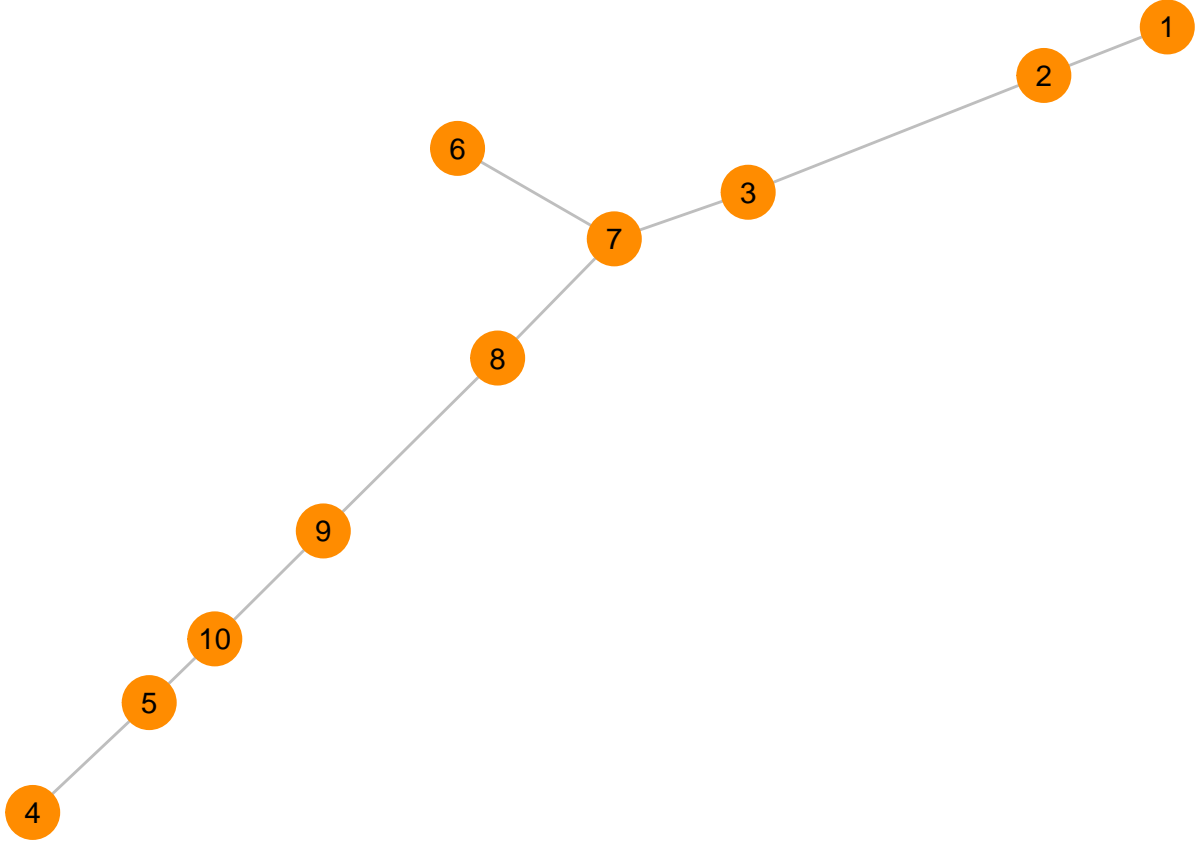


Figure 4: Estimated graph using MST on empirical extremal coefficient.

Unfortunately, it does not match with what we expected. Perhaps we need to add something to our conditional model to be able to conclude anything about the classic extremal conditional independence.

References

- Boulin, Alexis. 2024. “Variable Clustering of Multivariate Time Series According to the Dependence of Their Extremes.” PhD thesis, Université Côte d’Azur.
- Engelke, Sebastian, Jevgenijs Ivanovs, and Kirstin Stokorb. 2024. “Graphical Models for Infinite Measures with Applications to Extremes.” arXiv. <https://doi.org/10.48550/arXiv.2211.15769>.
- Engelke, Sebastian, and Stanislav Volgushev. 2022. “Structure Learning for Extremal Tree Models.” arXiv. <https://doi.org/10.48550/arXiv.2012.06179>.