

Spectral representation for extremal graphical model

CAPEL Alexandre

2025-02-24

Contents

1	Spectral representation of max stable process	1
1.1	General case	1
1.2	Finite dimensional max-stable distribution	2
2	Hammersley-Clifford theorem	3
2.1	Classical graphical model	4
2.2	Extremal graphical model	4
3	Husler-Reiss case	5
3.1	No-unicity of the spectral representation	5
3.2	Right matrix indentification	5
3.3	Simulation analysis	6
4	Conditional independance in the degenerate gaussian vector	8

In this document, we would like to find a link between the graphical model for extreme (for MGPD variable) and their spectral representation (with the max-stable).

1 Spectral representation of max stable process

We will use the notation of (“Spatial Extremes and Max-Stable Processes Mathieu Ribatet, Clément Dombry, and Marco Oesting” 2016).

1.1 General case

Let \mathcal{X} a compact subset of \mathbb{R}^p and let's consider a max-stable process $\{Z(x), x \in \mathcal{X}\}$. As the marginal of the process must be a generalized extreme value distribution, we will assume that all the unit margins follows a standard Fréchet distribution.

(Haan 1984) shows that :

Theorem. For a such max-stable process Z with continue sample paths, then :

$$Z(x) \stackrel{d}{=} \max_{i \geq 1} \zeta_i f_i(x), \quad x \in \mathcal{X},$$

where $\{(\zeta_i, f_i), i \geq 1\}$ are the point of a Poisson process on the space $\mathbb{R}_+ \times \mathcal{C}$ with intensity $\zeta^{-2} d\zeta \nu(df)$ for ν , a locally finite measure defined on the space \mathcal{C} of the non-negative continuous function on \mathcal{X} and such that :

$$\int_{\mathcal{C}} f(x) \nu(df) = 1, \quad x \in \mathcal{X}.$$

The integral relation comes from the fact the process Z has standard Fréchet unit margins. Of course there is a first problem which appears here, because the measure ν is not unique.

Of course, in that case where the measure ν is a probability measure, then we can express the f as the set of an independent copies random process $\{W_i(x), x \in \mathcal{X}\}$ such that $\mathbb{E}(W_1(x)) = 1$. In that sense we can thus write the representation as :

$$Z(x) \stackrel{d}{=} \max_{i \geq 1} \zeta_i W_i(x), \quad x \in \mathcal{X}.$$

We can give several examples of parametric max-stable process and their spectral representation :

- the first one is the Smith process that we can represent by :

$$Z(x) = \max_{i \geq 1} \zeta_i \varphi(x - U_i, \Sigma), \quad x \in \mathcal{X},$$

where $\varphi(\cdot, \Sigma)$ is the p -variate centered Gaussian density with covariance Σ and $\{(\zeta_i, U_i), i \geq 1\}$ are the point of a Poisson process on the space $\mathbb{R}_+ \times \mathbb{R}^p$ with intensity $\zeta^{-2} d\zeta du$.

- the next one is very famous, and is a generalization of the previous model. It is called the Brown-Resnik process and it is defined as :

$$Z(x) = \max_{i \geq 1} \zeta_i \exp(W_i(x) - \sigma^2(x)/2), \quad x \in \mathcal{X},$$

where $\{\zeta_i, i \geq 1\}$ is the point of a Poisson process on the space \mathbb{R}_+ with intensity $\zeta^{-2} d\zeta$ and $\{W_i(x), x \in \mathcal{X}\}$ are independent copies of a centered normal process with some covariance matrix $\Sigma(x)$.

The latter have some good properties especially the fact that the law depends only of some variogram Γ . Moreover, all the finite d -dimensional distribution of a such process are Husler-Reiss models.

Now, let's focus on finite dimensional distributions.

1.2 Finite dimensional max-stable distribution

Now assume that \mathcal{X} is a finite subset (of size d) of \mathbb{R}^p .

Of course, the last theorem is still valid, so for a d -dimensional max-stable vector Z , we can write it as :

$$Z_k = \max_{i \geq 1} \zeta_i W_k^{(i)}$$

where $\{\zeta_i, i \geq 1\}$ is the point of a Poisson process on the space \mathbb{R}_+ with intensity $\zeta^{-2} d\zeta$ and $(W^{(i)})_{i \geq 1}$ are independent copies of a random vector of dimension d with density f_W .

I chose to limit for now the case where the measure ν is a probability measure. That why I use the random vector representation.

We want to characterize here the distribution function of the max-stable vector Z using the spectral representation. We can write

$$\mathbb{P}(Z \leq z) = \mathbb{P}(\max_{i \geq 1} \zeta_i W_k^{(i)} \leq z_k) = \mathbb{P}((\zeta, W) \notin A)$$

where $A = \bigcup_{k=1}^d \{(\zeta, W) \in \mathbb{R}_+ \times \mathbb{R}^d \mid \zeta W_k > z_k\}$.

Notice that $\{(\zeta_i, W^{(i)}), i \geq 1\}$ is still a Poisson process with intensity $\zeta^{-2} d\zeta f_W(w) dw$, thus we have :

$$\begin{aligned} \mathbb{P}(Z \leq z) &= \mathbb{P}(N(A) = 0) \\ &= \exp \left[- \int_{(0, \infty) \times \mathbb{R}^d} \mathbb{1}_A(\zeta, w) \zeta^{-2} d\zeta f_W(w) dw \right] \\ &= \exp \left[- \int_0^\infty \mathbb{P}(\zeta > \min_{k=1, \dots, d} \frac{z_k}{W_k}) \zeta^{-2} d\zeta \right] \\ &= \exp \left[- \int_0^\infty \mathbb{P}(\max_{k=1, \dots, d} \frac{W_k}{z_k} > \frac{1}{\zeta}) \zeta^{-2} d\zeta \right] \\ &= \exp \left[- \int_0^\infty \mathbb{P}(\max_{k=1, \dots, d} \frac{W_k}{z_k} > u) du \right] \\ &= \exp \left[- \mathbb{E} \left(\max_{k=1, \dots, d} \frac{W_k}{z_k} \right) \right] \end{aligned}$$

and so we finally get :

$$\mathbb{P}(Z \leq z) = \exp \left[- \int_{\mathbb{R}^d} \max_{k=1, \dots, d} \frac{w_k}{z_k} f_W(w) dw \right]$$

that means we can write the exponent measure function as :

$$H(z) = \int_{\mathbb{R}^d} \max_{k=1, \dots, d} \frac{w_k}{z_k} f_W(w) dw.$$

With this formula, we notice that the density of the random vector of the spectral representation appears explicitly inside the exponent measure function, which characterize the distribution of MGPD corresponding to the max-stable vector Z .

2 Hammersley-Clifford theorem

We identify the space \mathcal{X} as $V = \{1, \dots, d\}$. We will denote by $\mathcal{G} = (V, E)$ the graph of nodes V and the edges E .

We define too the notion of decomposition in a graph : a partition $A, B, C \subset V$ is said to be a decomposition of $\mathcal{G} = (V, E)$ into the components $\mathcal{G}_{A \cup B}$ and $\mathcal{G}_{C \cup B}$ if :

- B separates A from C .
- the restriction to B of the graph is complete.

The decomposition is proper if A and C are not empty.

We say that the graph $\mathcal{G} = (V, E)$ is decomposable if the graph is complete, or if there is a proper decomposition such that the subgraphs are decomposable too.

I wanted to define these notions to formulate the following theorem properly.

2.1 Classical graphical model

When we have the classical notion of conditional independence in graphical models, it exists a theorem which associates all these independences with a factorization of the density : it is the **Hammersley-Clifford theorem**.

Theorem. Let X a graphical model according to a graph $\mathcal{G} = (V, E)$ such that the density f_X is positive and continuous. If the graph \mathcal{G} is decomposable, then we can factorize the density as follow :

$$f_X(x) = \frac{\prod_{C \in \mathcal{C}} f_C(x_C)}{\prod_{D \in \mathcal{D}} f_D(x_D)}.$$

where \mathcal{C} is the set of all cliques and \mathcal{D} is the collection of all the separators of the graph \mathcal{G} .

Actually, we can write this theorem differently to get an equivalence (with the condition for the graph and the density).

2.2 Extremal graphical model

For extremal graphical models, the notion of conditional independence is different because the random vector does not belong to a product space : this notion is defined in (Engelke and Hitz 2020).

In the same article, we can find a similar result that we can call as a Hammersley-Clifford theorem for extreme.

Theorem. Let Y an extremal graphical model according to a graph $\mathcal{G} = (V, E)$ such that the density f_Y is positive and continuous and let λ the density of the exponent measure. If the graph \mathcal{G} is decomposable, then we can factorize the density as follow :

$$f_Y(y) = \frac{1}{H(1)} \frac{\prod_{C \in \mathcal{C}} \lambda_C(y_C)}{\prod_{D \in \mathcal{D}} \lambda_D(y_D)}, \quad y \in \mathcal{L}$$

where \mathcal{C} is the set of all cliques and \mathcal{D} is the collection of all the separators of the graph \mathcal{G} .

Of course the equivalence that I talked previously is still valid here.

The difference here is that the factorization concerns a family of the density of the exponent measure (and its restrictions).

Now we can talk about what I want to do. In fact, we saw there is a link between the density of the spectral representation and the exponent measure, that we also use for the expression of the density of a Multivariate Generalized Pareto distribution. I wonder if we can get a relation of the form that : let Y a MGPD whose spectral representation of the corresponding max stable variable Z is $Z_k = \max_{i \geq 1} \zeta_i W_k^{(i)}$. Then we have :

$$f_W(w) = \frac{\prod_{C \in \mathcal{C}} f_C(w_C)}{\prod_{D \in \mathcal{D}} f_D(w_D)}, \quad \forall w \in \mathbb{R}^d \implies f_Y(y) = \frac{1}{H(1)} \frac{\prod_{C \in \mathcal{C}} \lambda_C(y_C)}{\prod_{D \in \mathcal{D}} \lambda_D(y_D)}, \quad \forall y \in \mathcal{L}$$

But I talked about the fact that there is not a unique spectral representation of the max-stable process : the aim would also be to find some condition on the graphs and the random vectors to get the unicity of the graphical representation.

3 Husler-Reiss case

Before continuing to seek assumption and make proofs, let's study the special case when the random vector is a Husler-Reiss Pareto distribution.

3.1 No-unicity of the spectral representation

Recall the spectral representation of a max-stable Husler-Reiss distribution :

$$Z_k = \max_{i \geq 1} \zeta_i \exp(W_k^{(i)} - \text{diag}(\Sigma)_k/2),$$

where $\{\zeta_i, i \geq 1\}$ is a Poisson process with intensity $\zeta^{-2}d\zeta$ and $W^{(i)}$ are independent copies of a d -dimensional centered gaussian distribution with covariance matrix Σ .

We know that the Husler-Reiss distribution depends only of a strictly conditionally negative definite variogram matrix of W :

$$\Gamma_{ij} = \mathbb{E}[(Y_i - Y_j)^2], \quad i, j \in V.$$

And this implies that, as we saw above, the spectral representation is not unique and there is a lot of possible Gaussian distribution that lead to the same variogram matrix Γ . We can therefore consider the set of all possible matrix for each variogram Γ :

$$S_\Gamma = \{\Sigma \in \mathbb{R}^{d \times d}, \text{semi positive definite} : \mathbb{1} \text{diag}(\Sigma)^T + \text{diag}(\Sigma) \mathbb{1}^T - 2\Sigma = \Gamma\} = \gamma^{-1}(\Gamma)$$

We can notice that γ is a linear application.

So there is an infinite number of matrix, and stochastic representation, but there is one which fit with the Husler-Reiss model.

3.2 Right matrix indentification

In (Hentschel, Engelke, and Segers 2023), they build an extended precision matrix Θ which summarize all the information we need for the conditional independence relationship for the extremal graphical models, in that sense :

$$Y_i \perp\!\!\!\perp_e Y_j \mid Y_{V \setminus \{i,j\}} \iff \Theta_{ij} = 0.$$

This matrix can be obtain by using some applications (which are bijections) that garanties a form of unicity of the spectral representation.

Therefore, let's considere the follwing applications :

$$\sigma : \Gamma \mapsto \Pi(-\frac{1}{2}\Gamma)\Pi, \quad \theta : \Gamma \mapsto \sigma(\Gamma)^+$$

where the matrix A^+ is the general inverse of A , and Π the orthogonal projection matrix in the space $\langle \mathbb{1} \rangle^\perp$.

They show in (Hentschel, Engelke, and Segers 2023) that the above applications are homeomorphisms between the set of the strictly conditionally negative definite variogram matrix \mathcal{D}_d ad the set of symmetric positive semi-definite matrix with kernel equal to $\langle \mathbb{1} \rangle$, denoted by \mathcal{P}_d^1 . More they show that :

$$\sigma^{-1}(\Sigma) = \gamma(\Sigma), \quad \theta^{-1}(\Theta) = \gamma(\Theta^+)$$

From this, we know that for any variogram matrix Γ , we can find a covariance matrix $\Sigma_\Gamma \in \mathcal{P}_d^1$. Thus, by linearity, we can write :

$$S_\Gamma = \{\Sigma_\Gamma + \Sigma_0, \Sigma_0 \in \ker(\gamma)\} \cap S_d^+$$

where S_d^+ is the set of the covariance matrix.

So, to summarise, we get a (*unique*) representative matrix Σ_Γ such that :

- $\Sigma_\Gamma^+ = \Theta$.
- every spectral representation of the max-stable Husler-Reiss distribution is controlled by a covariance matrix of the form $\Sigma_\Gamma + \Sigma_0$.

However, we have a problem : Σ_Γ^+ has no inverse, by definition. . . and so the Gaussian vector is degenerate. . .

To continue, I would like to find if we can get similar properties for degenerate Gaussian vector in graphical models. Then I would like to explore the different possible matrices and the corresponding graphical model and see if there exist a relationship between it and the Husler-Reiss graphical model (graph inclusion etc. . .).

3.3 Simulation analysis

In this section, we would like to see if we can find a matrix $\Sigma \in S_\Gamma$ such that the induced graph is not included in the corresponding Husler-Reiss graphical model.

Let's consider a four dimensional Husler-Reiss graphical model with the graph :

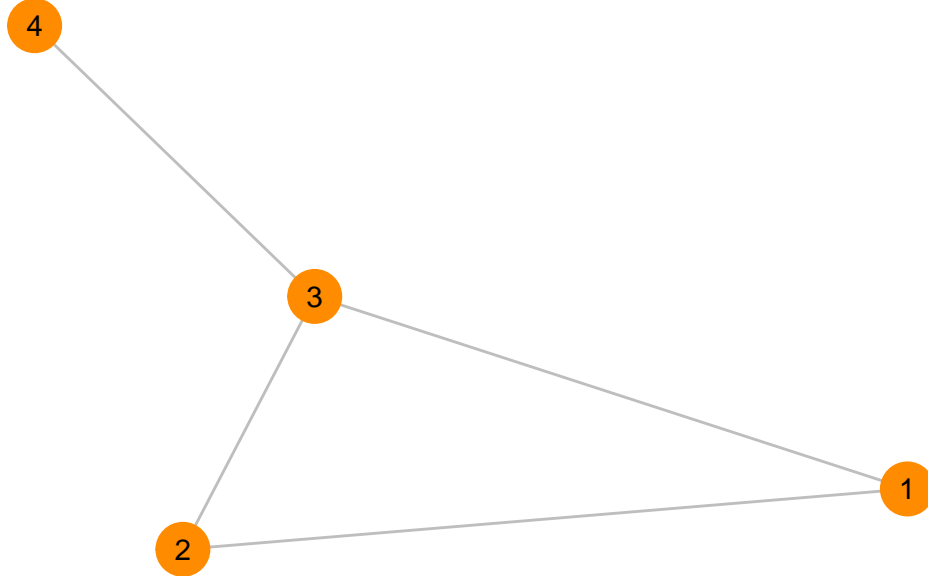


Figure 1: Graph for our variogram simulation

Thanks to the `graphicalExtremes` package, we can generate randomly a variogram matrix such that the induced Husler-Reiss graphical model is the same as the graph above. Thus, we get these following matrix :

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 0.0000000 0.2103068 0.5665426 0.8431731
## [2,] 0.2103068 0.0000000 0.2988806 0.5755112
## [3,] 0.5665426 0.2988806 0.0000000 0.2766306
## [4,] 0.8431731 0.5755112 0.2766306 0.0000000
```

with a precision matrix :

```
##          [,1]      [,2]      [,3]      [,4]
## [1,]  4.8179953 -5.280282  0.4622863  0.000000
## [2,] -5.2802815  9.132742 -3.8524603  0.000000
## [3,]  0.4622863 -3.852460  7.0051035 -3.614929
## [4,]  0.0000000  0.000000 -3.6149295  3.614929
```

(we saw here that we can build the graph from this matrix).

We can even get the representative matrix Σ_Γ as defined at the previous section :

```
##          [,1]      [,2]      [,3]      [,4]
## [1,]  0.23181532  0.05974642 -0.11120205 -0.18035969
## [2,]  0.05974642  0.09798434 -0.04428656 -0.11344421
## [3,] -0.11120205 -0.04428656  0.11232313  0.04316549
## [4,] -0.18035969 -0.11344421  0.04316549  0.25063841
```

We would like to see if there exist a matrix $\Sigma_0 \in \ker \gamma$ such that :

- $\Sigma_\Gamma + \Sigma_0$ is a semi definite positive matrix.
- the precision matrix of $\Sigma_\Gamma + \Sigma_0$ produce a null coefficient in the entry 12 (arbitrary choice).

3.3.1 Kernel characterisation

By definition, we can see that a matrix Σ belongs to the kernel of the linear application γ iff :

$$\Sigma_{ij} = \frac{\Sigma_{ii} + \Sigma_{jj}}{2}, \quad \forall i, j \in V$$

We notice this is a vector space of dimension d and the matrices are characterised by the value on their diagonal.

This characterisation is analytic and easy to compute in the computer.

3.3.2 Research of a good matrix

Let's go back to our case. We set the dimension to $d = 4$. Thus, the space S_Γ can be decribed as :

$$S_\Gamma = \left\{ \Sigma_\Gamma + \frac{1}{2} (\mathbb{I}(a \ b \ c \ d)' + (a \ b \ c \ d)\mathbb{I}') \mid a, b, c, d \in \mathbb{R} \right\} \cap S_d^+$$

So, if we fix the three firsts parameters a, b and c seek a matrix with conditional independence between 1 and 2 is equivalent to find the zero (if it exists) of the function :

$$f : \mathbb{R} \rightarrow \mathbb{R} \tag{1}$$

$$d \mapsto \left(\Sigma_\Gamma + \frac{1}{2} (\mathbb{I}(a \ b \ c \ d)' + (a \ b \ c \ d)\mathbb{I}') \right)_{12}^+ \tag{2}$$

Moreover, for such d_0 , we also need to check if the matrix is semi definite positive.

The calculation are made with a grid between -2 and 2 with 0.2 step for each coefficient (everything is in *alternative_graph_research* target).

We used a Newton algorithm to find the zero, from the *pracma* library.

In this configuration, we find none matrix :

```
alternative_graph_research |>
  rowwise() |>
  filter(semi_def(Sigma_Gamma + ker_gamma(c(a, b, c, d_zero)))) |> # check if the matrix is semi-def po
  ungroup() |>
  count(name = "Number of semi definite positive")
```

```
## # A tibble: 1 x 1
##   'Number of semi definite positive'
##                               <int>
## 1                               0
```

To conclude with this first simulation, we did not find a matrix whose the graphical model induced is not include in the variogram one.

I need to try this with different matrices and different node (for example node 3 and 4 or node 2 and 3) to have a larger point of view.

4 Conditional independance in the degenerate gaussian vector

References

- Engelke, Sebastian, and Adrien S. Hitz. 2020. “Graphical Models for Extremes.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82 (4): 871–932. <https://doi.org/10.1111/rssb.12355>.
- Haan, L. De. 1984. “A Spectral Representation for Max-stable Processes.” *The Annals of Probability* 12 (4): 1194–204. <https://doi.org/10.1214/aop/1176993148>.
- Hentschel, Manuel, Sebastian Engelke, and Johan Segers. 2023. “Statistical Inference for Hüsler-Reiss Graphical Models Through Matrix Completions.” arXiv. <https://doi.org/10.48550/arXiv.2210.14292>.
- “Spatial Extremes and Max-Stable Processes Mathieu Ribatet, Cle´ment Dombry, and Marco Oesting.” 2016. In *Extreme Value Modeling and Risk Analysis*. Chapman and Hall/CRC.