ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Semester project in mathematics, MA1

Simulation by circular embedding method of max-stable random fields

Author:
Nathalie Pellet

Supervisors:
Anthony Davison
Mathieu Ribatet

Autumn 2009

Abstract

Theory of random fields or random processes is presented with an emphasis on the possible choices of correlation functions. Then two methods for generating Gaussian random fields in $\mathbb{R}^d$ are presented. The first method is a direct one but does not work for a too large number of locations. The second one is called the circular embedding method, works with large number of locations and had been introduced independently by Wood and Chan [1994] and Dietrich and Newsam [1993]. It is exact in principle for correlation functions with compact support and an approximation method is given otherwise. Those methods have been implemented in **R** and validated by means of variogram comparisons. Finally Max-stable processes are introduced, the simulation procedure from Schlather's model is presented and max-stable process is simulated with help of the circular embedding method.

Contents

Chapter 1

Introduction

In geostatistics, the data consist on finite samples of measured values at different locations in the space. It is wished to find a model that is valid in the space. Thus, we aim to have method to simulate random fields. In this report, we will focus on Gaussian random fields, that have nice properties and are often used in practice. In a second time, we will focus on max-stable processes, which are limiting processes and are used in the modelling of extremes.

We will present in Chapter 2 the mathematical definitions of random fields or random processes. We will give some nice properties of random fields which are stationarity and isotropy. Since random fields and especially Gaussian random fields are mainly determined by their correlation functions, that have to be positive semidefinite functions, we will discuss some conditions for a function to be positive semidefinite. Then we will expose some examples of valid correlation functions for isotropic and stationary Gaussian random fields.

In Chapter 3, we will present two methods for generating stationary and isotropic Gaussian random fields with known correlation functions, at given locations in $\mathbb{R}$. The first methods we will propose are direct methods. They are simple and exact, but have the drawback that they are strongly dependent of the number of locations and even not defined when this number is too large. Indeed they are based on the Cholesky factorization and the Singular value decomposition respectively, the computational time of which increases sensibly as the number of locations increases. Thus we will propose a second method which is less sensible to the number of locations. This second method is called the circular embedding method. It had been introduced independently by Wood and Chan [1994] and Dietrich and Newsam [1993]. It is exact in principle for correlation functions with compact support and Wood and Chan [1994] give an approximation procedure for the other correlation functions. The only restriction of this method is that the locations

have to be on a regular grid. This method is based on the idea that the correlation matrix is Toepliz and that we can embed this matrix in a bigger one which is circulant. Then nice properties of the circulant matrices are used to generate a Gaussian random field, as for example that the eigenvalues of a circulant matrix can be obtainned by a fast Fourier transform on the first row of the matrix. For both methods we will give the algorithm that were implemented in **R**. We will also describe the approximation procedure for the circular embedding method.

We will use those methods to simulate Gaussian random fields at given locations with different correlation functions. We will validate our simulation by means of variogram comparisons. We will conclude our third chapter on Gaussian random fields simulation by a comparison of the computation time needed by the methods depending on the number of locations.

Finally, in Chapter 4, we will introduce max-stable processes. We will first present a short review of Poisson point processes and the way one can simulate such processes. Then we will present the basic ideas behind max-stable processes and come to a simulation procedure. The procedure we will present will be the simulation procedure for a max-stable process according to the Schlather model, which is based on a simulation of a Gaussian random field and a Poisson process. Finally, we will simulate a max-stable process with help of the circular embedding method and validate our results by means of the $F-$madogram. This is a generalisation of the variogram, that applies also when we have infinite means and variances. We will finally present a max-stable process simulated by the procedure we suggested and comment on this procedure.

General definitions

2.1 Random Fields

In geostatistics one of the basic tools to model a problem is a random field.

Definition 2.1 (Random field)**.**
Given a parameter set T and a probability space $(\Omega, \mathcal{F}, P)$, we define a *random field* $Y(\mathbf{t}, \omega)$ as a real valued function which is measurable on Ω for every fixed $\mathbf{t} \in T$.

We will assume for the following that $T = \mathbb{R}^d$ with $d \geq 1$, that is we consider the d–dimensional Euclidean space. We will use the abuse of notation $Y(\mathbf{t})$ instead of writing $Y(\mathbf{t}, \omega)$. In this context, a random field $Y(\cdot)$ on $\mathbb{R}^d$ is seen as a function whose values are random variables for any $\mathbf{t}$ in $\mathbb{R}^d$.

A particular case of random fields, which will be considered all along this report are Gaussian random fields. They are interesting since many natural phenomena can be modelled by a Gaussian field. For example soil data or surface elevations may be modelled by a Gaussian random fields, e.g. see [Diggle and Ribeiro Jr., 2007, pages 1 to 15].

Definition 2.2 (Gaussian random field)**.**
A random field $Y(\cdot)$ in $\mathbb{R}^d$ is called a *Gaussian random field* if the joint distributions of $Y = \{Y(\mathbf{x_1}), \ldots, Y(\mathbf{x_k})\}$ are multivariate normal distributions for any choice of k and $\mathbf{x_1}, \ldots, \mathbf{x_k} \in \mathbb{R}^d$.

Analogously to the multivariate normal distributions, a Gaussian random field is completely specified by its expectations $\mu(\mathbf{x})$, see Definition (2.3), its variances $\sigma^2(\mathbf{x}) = \mathrm{Var}\{Y(\mathbf{x})\}$, see Definition (2.5) and its correlation function $\rho(u) = \mathrm{Corr}\{Y(\mathbf{x}), Y(\mathbf{y})\}$ with $u = \|\mathbf{x} - \mathbf{y}\|$, see Definition (2.6), which has to be positive semidefinite to ensure the existence of all finite-dimensional distributions.

2.2 Expectation and Covariances

We now will recall briefly the definitions of expectation and correlation since they are of crucial importance for Gaussian random fields.

Definition 2.3 (Expectation).
The *expectation* of a random field is defined as

$$\mu(\mathbf{x}) = \mathrm{E}\left[Y(\mathbf{x})\right] = \int_{\mathbb{R}^d} \mathbf{y} f_{Y(\mathbf{x})}(\mathbf{y}) \mathrm{d}\mathbf{y}.$$

We will often assume the expectation being equal to 0, since for any random field it only means a translation, in the sense that

$$Y^*(\mathbf{x}) = Y(\mathbf{x}) - \mu(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d,$$

where $Y^*(\cdot)$ is a Gaussian random field with null expectation and $Y(\cdot)$ has expectation μ.

Definition 2.4 (Covariance).
In $\mathbb{R}^2$, the *covariance* is defined as follows

$$\begin{aligned}\mathrm{Cov}\left\{Y(\mathbf{x}), Y(\mathbf{y})\right\} &= \mathrm{E}\left[Y(\mathbf{x})Y(\mathbf{y})\right] - \mu(\mathbf{x})\mu(\mathbf{y}) \\ &= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \mathbf{x}\mathbf{y} f_{Y(\mathbf{x}),Y(\mathbf{y})}(\mathbf{x},\mathbf{y}) \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} - \mu(\mathbf{x})\mu(\mathbf{y}).\end{aligned}$$

The diagonal elements of the covariance matrix are called the *variances*.

Definition 2.5 (Variance).
The variance is defined as the number

$$\sigma^2(\mathbf{x}) = \mathrm{Var}\left\{Y(\mathbf{x})\right\} = \mathrm{Cov}\left\{Y(\mathbf{x}), Y(\mathbf{x})\right\}.$$

Finally we need to introduce the correlation which is a ratio of the covariance and the standard deviations which are the square root of the variances.

Definition 2.6 (Correlation).
We denote the *correlation* by ρ and it is defined as

$$\rho(\mathbf{x},\mathbf{y}) = \mathrm{Corr}\left\{Y(\mathbf{x}), Y(\mathbf{y})\right\} = \frac{\mathrm{Cov}\left\{Y(\mathbf{x}), Y(\mathbf{y})\right\}}{\sigma(\mathbf{x})\sigma(\mathbf{y})}.$$

When we say that a random field $Y(\cdot)$ on $\mathbb{R}^d$ is of second order, we mean that it has finite expectation and variance. In that case, when $Y(\cdot)$ is of second order, we can write the covariance as follows

$$\mathrm{Cov}\left\{Y(\mathbf{x}), Y(\mathbf{y})\right\} = \mathrm{E}\left[\{Y(\mathbf{x}) - \mathrm{E}\left[Y(\mathbf{x})\right]\}\{Y(\mathbf{y}) - \mathrm{E}\left[Y(\mathbf{y})\right]\}\right].$$

2.3 Stationarity

Let us now look at two invariance properties of random fields that can ease the computations. The first one that we will consider is the stationarity, which concerns the invariance on translations.

Definition 2.7 (Stationarity in the wide sense)**.**
Let $Y(\cdot)$ be a random field of second order on $\mathbb{R}^d$, thus $\mathrm{E}\left[Y(\mathbf{x})\right] < \infty$ and $\mathrm{Var}\left\{Y(\mathbf{x})\right\} < \infty$ for any $\mathbf{x} \in \mathbb{R}^d$. We call $Y(\cdot)$ *stationary* (*in the wide sense*) if its expectation is constant on $\mathbb{R}^d$ and its correlation function is translation invariant, that is

$$\mu(\mathbf{x}) = \mu \quad \text{and} \quad \rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x} + \mathbf{h}, \mathbf{y} + \mathbf{h}), \quad \text{for all } \mathbf{x}, \mathbf{y}, \mathbf{h} \in \mathbb{R}^d.$$

It follows from the definition that on a stationary random field, we have that $\rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x} - \mathbf{y}, 0)$ and therefore we can define, for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ such that $\mathbf{h} = \mathbf{x} - \mathbf{y}$, the correlation function as

$$\rho(\mathbf{h}) = \rho(\mathbf{x}, \mathbf{y}).$$

When $Y(\cdot)$ is stationary the variance is constant, indeed

$$\begin{aligned}\mathrm{Var}\left\{Y(\mathbf{x})\right\} &= \mathrm{Cov}\left\{Y(\mathbf{x}), Y(\mathbf{y})\right\} \\ &= \mathrm{Cov}\left\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x} + \mathbf{h})\right\} = \mathrm{Var}\left\{Y(\mathbf{x} + \mathbf{h})\right\}.\end{aligned}$$

Therefore the following relation holds

$$\mathrm{Cov}\left\{Y(\mathbf{x}), Y(\mathbf{y})\right\} = \sigma^2 \, \rho(\mathbf{h}), \quad \text{where } \mathbf{h} = \mathbf{x} - \mathbf{y}.$$

2.4 Isotropy

We will now consider the second invariance property of random fields, isotropy. In the same sense as stationarity is invariance to translations, isotropy is invariance to rotations.

Definition 2.8 (Isotropic in the wide sense)**.**
We say that a random field $Y(\cdot)$ is *isotropic* (*in the wide sense*) if

$$\mathrm{E}\left[Y(A\mathbf{x})\right] = \mathrm{E}\left[Y(\mathbf{x})\right] \quad \text{and} \quad \rho(A\mathbf{x}, A\mathbf{y}) = \rho(\mathbf{x}, \mathbf{y}),$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ and all rotation (orthogonal) matrices A. We say that both the expectation and the covariance function are rotation invariant.

When $Y(\cdot)$ is stationary this rotation invariance condition reduces to the condition that the correlation function only depends on the distance, that is

$$\rho(\mathbf{x}, \mathbf{y}) = \rho(h) \quad \text{where } h = \|\mathbf{x} - \mathbf{y}\|.$$

Note that we gave here for stability and isotropy only the wide sense definitions, but there also exist strict sense definitions of both those properties. Since we will consider Gaussian random fields, it is sufficient to know the wide sense definitions. Indeed, in the case of Gaussian random fields, the strong sense and the wide sense definitions coincide, see [Vanmarcke, 1983, Chapter 2] for details.

2.5 Positive semi-definiteness of functions

Let us introduce the notion of positive semidefiniteness which has a relation with the covariance functions on stationary random fields as Theorem (2.1) says.

Definition 2.9 (Positive semidefiniteness)**.**
A function φ on $\mathbb{R}^d$ is said to be *positive semidefinite* if

$$0 \leq \sum_{k=1}^{n} \sum_{j=1}^{n} a_k a_j \ \varphi(\mathbf{t}_k - \mathbf{t}_j),$$

for any choice of $(\mathbf{t}_1, \ldots, \mathbf{t}_n)$ and $(a_1, \ldots, a_n)$ with $\mathbf{t}_k \in \mathbb{R}^d$ and $a_k \in \mathbb{R}$ and for all $n \in \mathbb{N}$.

Let us now present a theorem that links positive semidefinite functions and stationary random fields.

Theorem 2.1. [Abrahamsen, 1997]
The class of positive semidefinite functions on $\mathbb{R}^d$ coincidence with the class of correlation functions of stationary random fields on $\mathbb{R}^d$.

Indeed, we know that the correlation function of a stationary random field (in the wide sense) must be positive semidefinite and that for any positive semidefinite correlation function, we can generate a stationary Gaussian random field with this correlation function.

Checking the positive semi-definiteness of a function by means of Definition (2.9) is often not easy to achieve and therefore it is simpler to use some necessary and sufficient criterion. We will present some of them, namely Bochner's theorem (2.2), Riesz's theorem (2.3) and Schönberg's theorem (2.4). Let us first introduce the characteristic function.

Definition 2.10 (Characteristic function)**.**
We define the *characteristic function* χ of a random variable X on $\mathbb{R}$ with probability distribution F_X, as

$$\chi(x) = \mathrm{E}\left[e^{itX}\right] = \int e^{ixt} \mathrm{d}F_X(t),$$

where $i = \sqrt{-1}$.

We will now present Bochner's theorem.

Theorem 2.2 (Bochner's theorem)**.**
A function $\varphi\colon \mathbb{R} \to \mathbb{C}$ is continuous and positive semidefinite if and only if there exists a non-negative bounded measure F such that we can write φ as a characteristic function

$$\varphi(x) = \int_{\mathbb{R}} e^{ixt} \mathrm{d}F(\mathbf{t}) = \chi(x) \quad \textit{for all } x \in \mathbb{R}.$$

A characteristic function is positive semidefinite, so $\varphi(x)$ is positive semidefinite. The converse is less easy to prove but for instance S. Bochner gives the proof in Bochner [1933].

This theorem can be extended to $\mathbb{R}^d$, that is, a function φ on $\mathbb{R}^d$ is continuous and positive semidefinite if and only if there exists a non-negative bounded measure F such that we can write

$$\varphi(\mathbf{x}) = \int_{\mathbb{R}^n} e^{i\sum_{j=1}^{d} x_j t_j} \mathrm{d}F(\mathbf{t}) = \chi(\mathbf{x}), \quad \text{for all } \mathbf{x} \in \mathbb{R}^d, \tag{2.5.1}$$

This integral (2.5.1) is the d-dimensional Fourier Transform of F and is sometimes called *Fourier-Stieltjes* integral.

We now present Riesz's theorem which does not require the continuity of the function, so that together Riesz and Bochner's theorems will give a complete characterisation of the positive semidefinite measurable functions.

Theorem 2.3 (Riesz's theorem)**.**
Any positive semidefinite and measurable function φ on $\mathbb{R}^d$ can be written as

$$\varphi = \varphi_c + \varphi_0,$$

where φ_c and φ_0 are positive semidefinite functions on $\mathbb{R}^d$ and φ_c is continuous whereas φ_0 equals zero Lebesgue almost everywhere.

For the proof see the article written by Riesz [1933] or a version in english by Bruzual and Dominguez [2001].

Notice that in practice we mostly use for φ_0 the *nugget effect* function stated as

$$\varphi_0(\mathbf{x}) = \alpha \mathbf{1}_{\{0\}}(\mathbf{x}) = \begin{cases} \alpha, & \mathbf{x} = 0 \\ 0, & \text{otherwise} \end{cases}$$

where α is a non-negative constant. It has even been shown by Gneiting and Sasvári [1999] that when φ is a function on $\mathbb{R}^d$ with $d \geq 2$, φ_0 can only be the nugget effect.

In random fields of higher dimensions, we often require both stationarity and isotropy in the wide sense. So it is quite normal to use these properties when checking for positiveness. The following theorem requires the function to be isotropic in the wide sense.

Theorem 2.4 (Schönberg's theorem)**.**
The function φ is a continuous, rotation invariant and positive semidefinite function on $\mathbb{R}^d$ if and only if $\varphi(\mathbf{x}) = \phi(\|\mathbf{x}\|)$ where ϕ is the Hankel transform of F. Thus ϕ has the following form

$$\phi(h) = \int_{[0,\infty)} \Gamma(\frac{d}{2}) \left(\frac{2}{rh}\right)^{(d-2)/2} J_{(d-2)/2}(rh) \mathrm{d}F(r), \quad \text{for all } h \in [0,\infty),$$

with J representing the first Bessel function.

A proof can be found in Schönberg [1938, pages 815-816].

We can summarize Bochner (2.2), Riesz (2.3) and Schönberg's (2.4) theorems in the following way.

Theorem 2.5.
A measurable positive semidefinite function φ on $\mathbb{R}^d$ can be written as

$$\varphi(\mathbf{x}) = \varphi_0(\mathbf{x}) + \int e^{i\sum_{j=1}^d x_j t_j} \mathrm{d}F(\mathbf{t}), \tag{2.5.2}$$

where φ_0 equals zero Lebesgue almost everywhere and is positive semidefinite, F is a non-negative bounded measure. In particular, when φ is rotation invariant and $d \geq 2$, it has the special form

$$\varphi(\mathbf{x}) = \alpha \mathbf{1}_{\{0\}}(\mathbf{x}) + \int_{[0,\infty)} \Gamma(\frac{d}{2}) \left(\frac{2}{\mathbf{t}\|\mathbf{x}\|}\right)^{(d-2)/2} J_{(d-2)/2}(\mathbf{t}\|\mathbf{x}\|) \mathrm{d}F(\mathbf{t}), \tag{2.5.3}$$

with α being a non-negative constant and F a non-negative bounded measure.

The converse holds in the sense that a function on the form (2.5.2) or (2.5.3) is up to a constant a correlation function on suitable stationary and isotropic, random field.

A detailed proof for $d = 2, 3$ and the general case with any d is given by Yaglom [1987, pages 349 to 353]. From Equation (2.5.3), we can derive some special cases of correlation functions listed by Abrahamsen [1997, page 32]. We here just specify the correlation function of an isotropic random field in $\mathbb{R}$, that is

$$\rho(x) = \int_0^\infty \cos(xt) f(t) \mathrm{d}t. \tag{2.5.4}$$

2.6 The variogram

In application, it often appears that handling the correlation function is inconvenient and a related function is used instead, namely the variogram. We will use the term variogram meaning in fact centred semi-variogram.

Definition 2.11 (centred semi-variogram)**.**
The *Centred semi-variogram* is a function

$$\gamma_c(\mathbf{x}-\mathbf{y}) = \frac{1}{2}\mathrm{Var}\left\{Y(\mathbf{x}) - Y(\mathbf{y})\right\}, \quad \mathbf{x},\mathbf{y} \in \mathbb{R}^d$$

defined for an intrinsically stationary random field $Y(\cdot)$ on $\mathbb{R}^d$. By intrinsically stationary, we intend that $Y^*(\mathbf{x}) = Y(\mathbf{x}+\mathbf{h}) - Y(\mathbf{x})$ is stationary in the wide sense for any $\mathbf{h} \in \mathbb{R}^d$. Thus $\mathrm{Var}\left\{Y(\mathbf{x}) - Y(\mathbf{y})\right\} < \infty$, because we can write $\mathbf{x} = \mathbf{y} + \mathbf{h}$ for some $\mathbf{h} \in \mathbb{R}^d$.

We are interested in simulating Gaussian random fields that are stationary or isotropic, thus we will restrict to the case of weakly stationary random fields $Y(\cdot)$ and therefore we will assume that

$$\mathrm{Var}\left\{Y(\mathbf{x})\right\} < \infty.$$

The variogram and the correlation function are related in the following way.

Proposition 2.6.
Assume γ_c denotes a centred semi-variogram on a stationary random field $Y(\cdot)$ in the wide sense and that ρ is a correlation function on $Y(\cdot)$, then

$$\gamma_c(\mathbf{h}) = \sigma^2\{\rho(0) - \rho(\mathbf{h})\} \quad \textit{for all} \quad \mathbf{h} \in \mathbb{R}^d.$$

Proof. By weak stationarity of $Y(\cdot)$, we have that

$$\mathrm{Var}\left\{Y(\mathbf{x})\right\} < \infty, \quad \text{for any } \mathbf{x} \in \mathbb{R}^d$$

and

$$\begin{aligned}
\gamma_c(\mathbf{h}) = \gamma_c(\mathbf{x}-\mathbf{y}) &= \frac{1}{2}\mathrm{Var}\left\{Y(\mathbf{x}) - Y(\mathbf{y})\right\} \\
&= \frac{1}{2}\left[\mathrm{Var}\left\{Y(\mathbf{x})\right\} + \mathrm{Var}\left\{Y(\mathbf{y})\right\} - 2\mathrm{Cov}\left\{Y(\mathbf{x}), Y(\mathbf{y})\right\}\right] \\
&= \frac{1}{2}\{2\sigma^2\rho(0) - 2\sigma^2\rho(\mathbf{x}-\mathbf{y})\} = \sigma^2\{\rho(0) - \rho(\mathbf{x}-\mathbf{y})\} \\
&= \sigma^2\{\rho(0) - \rho(\mathbf{h})\}.
\end{aligned}$$

□

2.7 Examples of valid correlation functions

We will here list some interesting correlation functions ρ of isotropic and stationary random fields such that $\rho(h) = \rho(\|\mathbf{x}-\mathbf{y}\|)$. We will give the expression of the function ρ rescaled such that $\rho(0) = 1$, but we omit the scale parameter.

Example 2.1 (Cosine)**.**

$$\rho(h) = cos(h).$$

The cosine model is valid only on $\mathbb{R}$ *and has a big importance in a theoretical point of view since it is suggested by the real part of Bochner's theorem (2.2). It corresponds to the special case listed as Equation (2.5.4). But this model is not really helpful in practice partly because it is valid only in* $\mathbb{R}$*.*

Another model which has an importance in a theoretical point of view is the following.

Example 2.2 (Nugget effect)**.**

$$\rho(h) = \mathbf{1}_{\{0\}}(h) = \begin{cases} 1, & h = 0 \\ 0, & otherwise. \end{cases}$$

The nugget effect model is valid for $\mathbb{R}^d$ *with* $d \in \mathbb{N}$*. This model plays an important role since it can be added to a more complex model to form a correlation function as seen in Riesz's theorem (2.3).*

We have now seen some models that have theoretical importance, but we are more interested in the models that are used in practice. One of those models is the Gaussian model.

Example 2.3 (Gaussian model)**.**

$$\rho(h) = e^{-h^2}.$$

The Gaussian model is frequently used in practice and a Gaussian random field with this model can be simulated by the spectral method according to Schlather [1999, section 4.6.3]. But this model has some drawbacks, e.g. the numerical instability of the covariance matrices involved in the simulations due to their almost singularity. An other example of drawback is that this correlation function has the theoretical property that the realisation on the whole real line is determined by the realisation on an arbitrary small, continuous interval as explained by Diggle and Ribeiro Jr. [2007, chapter 3], which is unrealistic for most applications.

The Gaussian model is a special case of a wider family called symmetric stable family, namely the case with parameter $\nu = 2$.

Example 2.4 (Symmetric stable family)**.**

$$\rho(h) = e^{-h^\nu}, \qquad \nu \in]0, 2] .$$

The symmetric stable is called so, since for any parameter ν*, the value at 1 is the same. As said by Diggle and Ribeiro Jr. [2007, chapter 3], this family is comparable to another family called the Whittle-Matérn family, but is less flexible. It has as special cases the Gaussian model and the Exponential family.*

Let us now introduce the second special case of stable family, namely the Exponential model, which corresponds to $\nu = 1$.

Example 2.5 (Exponential family)**.**

$$\rho(h) = e^{-h}.$$

The Exponential model is also a special case of another family of correlation functions called the Whittle-Matérn family. The interest of the Exponential model is that it gives fast and simple simulations on the real axis due to its nice Markov property, see [Schlather, 1999, section 4.4].

Let us define the Whittle-Matérn family, which with $\nu = 1/2$ becomes the Exponential model and when $\nu \to \infty$ is the Gaussian model.

Example 2.6 (Whittle-Matérn family)**.**

$$\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)}\ h^{\nu}\ K_{\nu}(h), \qquad \nu > 0,$$

where K_ν stands for the modified Bessel function of order ν of the second kind. The Whittle-Matérn family is also called in different literatures the Basset model or the Modified Bessel family. This family is much used in practice, since the degree of differentiability of the underlying random field is to be specified by the choice of ν. For instance, ρ is $2d$ times differentiable when $\nu = \frac{2d+1}{2}$. Since it is not very realistic to have a surface that is more than twice differentiable, one will prefer to choose $\nu \leq \frac{3}{2}$. For more details about this family see [Matérn, 1960, page 17].

We will now give a family of models that is related to the Whittle-Matérn in the sense that it also uses some Bessel functions.

Example 2.7 (Bessel family)**.**

$$\rho_\nu(h) = 2^\nu \Gamma(\nu+1) h^{-\nu} J_\nu(h), \qquad \nu \geq \frac{d-2}{2},$$

where J_ν denotes a first kind Bessel function of order ν and d is the dimension of $\mathbb{R}^d$. The Bessel family is an example of oscillating correlation functions. These oscillating correlation functions are parametrized by their period $\nu = 2\pi/\omega$, where ω is the angular frequency. The particular case ρ_3 is called the Hole effect that is

$$\rho_3(h) = \frac{1}{h}\sin(h).$$

Let us now define a family which is said to behave poorly for the circular embedding method when the dimension d is high.

Example 2.8 (Cauchy family)**.**

$$\rho(h) = \left(1 + h^2\right)^{-\nu}, \qquad \nu > 0.$$

The Cauchy family is positive semidefinite by construction as explained in Schlather [1999]. We note that the special case $\nu = 1$ is called rational quadratic model.

Finally we give a model that is behaving well in the simulation of Gaussian random field by circular embedding method since it has compact support.

Example 2.9 (Spherical family)**.**

$$\rho(h) = \lambda \int_{\mathbb{R}^n} \mathbf{1}_{\{h \le r\}} \mathbf{1}_{\{\|\mathbf{t}\| < r\}} \mathrm{d}\mathbf{t}_1 \cdots \mathrm{d}\mathbf{t}_n$$

In this family $\rho(h)$ is the volume of the intersection of two d-dimensional spheres of radius r separated by a distance h, whose particular case in dimension $d = 3$ is given by

$$\rho_3(h) = \left(1 - \frac{3}{2}h + \frac{1}{2}h^3\right) \mathbf{1}_{\{0 \le h \le 1\}}.$$

This case with $d = 3$ is practically the most used case. Often it is implicitly meant $d = 3$, when we say spherical family. The spherical family is interesting for simulation of Gaussian random fields since the have finite range, that is $\rho(h) = 0$ sufficiently large h.

For more example we refer to Schlather [1999], Abrahamsen [1997] or Gneiting [1997], but we consider the previous list sufficient for this report.

Finally we notice that we can change the range of a correlation function by redefining $\rho^*(h) = \rho(\frac{h}{\phi})$ where we have that the range $\phi > 0$. A linear combination of two correlation function with non-negative scalars is also a correlation function. Indeed if $Y_1(\cdot)$ and $Y_2(\cdot)$ are independent random fields with correlation function ρ_1 and ρ_2 respectively and if $a_1,\ a_2 \ge 0$, then $\sqrt{a_1}Y_1(\cdot) + \sqrt{a_2}Y_2(\cdot)$ has the correlation function $\rho = a_1\rho_1 + a_2\rho_2$. We can also introduce a nugget effect, which corresponds to a discontinuity at $h = 0$. Finally Theorem (2.5) suggests to generalize in the sense that any covariance function can be written as

$$\mathrm{Cov}^*(h) = \tau^2 \mathbf{1}_{\{0\}}(h) + \mathrm{Cov}\left(\frac{h}{\phi}\right)\left[1 - \mathbf{1}_{\{0\}}(h)\right].$$

Thus in terms of correlation function, normalized to 1 at $h = 0$, it gives

$$\rho^*(h) = \mathbf{1}_{\{0\}}(h) + \frac{\sigma^2}{\sigma^2 + \tau^2}\rho\left(\frac{h}{\phi}\right)\left[1 - \mathbf{1}_{\{0\}}(h)\right],$$

where the scale or range parameter ϕ is a positive constant, the nugget parameter τ^2 is a constant, σ^2 is a constant corresponding to a variance and ρ is a correlation function.

The circular embedding method

We aim to simulate a Gaussian random field $Y(\cdot)$ in $\mathbb{R}^d$ at a set of n locations $\mathbf{x}_i \in \mathbb{R}^d$, with mean vector μ and correlation function ρ. With n not too large we can use direct methods such as Cholesky factorization or singular value decomposition, but as n gets too large, their computation time increases too much. Therefore we need to find other methods when n is large. The one we will present in this chapter is called circular embedding method.

We will first consider the simulation of a Gaussian random field by direct methods. We will consider random fields in $\mathbb{R}$ since it is easier to compute and less sensitive to the size of n.

3.1 Simulation by direct approach

As said before, we aim to simulate a Gaussian random field $Y(\cdot)$ in $\mathbb{R}$ at a set of n locations $x_i \in \mathbb{R}$, with mean vector μ and correlation function ρ. We assume without loss of generality that $\mu \equiv 0$; else $Y^*(\cdot) = Y(\cdot) - \mu$, where $Y^*(\cdot)$ is a Gaussian Random Field with mean 0 and correlation function ρ.

In order to simulate $Y(\cdot)$, we use the fact that $Y = \{Y(x_1), \ldots, Y(x_k)\}$ has a multivariate Gaussian distribution with mean 0 and covariance matrix Σ and thus

$$Y \stackrel{d}{=} \Sigma^{1/2} N,$$

where N is vector of n independent $N(0,1)$ and $\Sigma = \Sigma^{1/2}(\Sigma^{1/2})^T$. Indeed Y has a multivariate Gaussian distribution and we have that

$$\mathrm{E}\left[\Sigma^{1/2} N\right] = \mathrm{E}\left[Y\right]$$

and

$$\begin{aligned}
\mathrm{Cov}\left(\Sigma^{1/2}N, \Sigma^{1/2}N\right) &= \mathrm{E}\left[(\Sigma^{1/2}N)(\Sigma^{1/2}N)^T\right] = \mathrm{E}\left[\Sigma^{1/2} N N^T (\Sigma^{1/2})^T\right] \\
&= \Sigma^{1/2}\mathrm{E}\left[NN^T\right](\Sigma^{1/2})^T = \Sigma^{1/2}\mathrm{I}_n(\Sigma^{1/2})^T \\
&= \Sigma^{1/2}(\Sigma^{1/2})^T = \Sigma = \mathrm{Cov}\left(Y, Y\right).
\end{aligned}$$

Thus $Y \stackrel{d}{=} \Sigma^{1/2} N$.

There are two direct ways of getting the matrix $\Sigma^{1/2}$, which are a Cholesky factorisation or a singular value decomposition. Both of these methods have the drawback that they need a long computation time, so they should not be applied to large n.

The Cholesky factorization works only for positive semidefinite matrices and returns a lower triangular matrix Q such that $\Sigma = QQ^T$. Since the covariance matrix is by definition positive semidefinite, this method can be applied. Therefore, we can use $\Sigma^{1/2} = Q$ for simulating Y.

The singular value decomposition (SVD) works for any matrix and returns two unitary matrices U and V and a vector λ such that $\Sigma = U\Lambda V^*$ with Λ being the diagonal matrix formed with the elements in λ and V^* being the conjugate transpose of V. The elements in λ are the singular values of Σ, given in a decreasing order. This decomposition method ensures that $\lambda \geq 0$ and when the covariance matrix Σ is positive semidefinite, it strengthens this to $\lambda > 0$. As we work with real numbers we have that the conjugate transpose is equal to the transpose, thus we have that $\Sigma^{1/2} = U\Lambda^{1/2}V^T$. Moreover symmetry of Σ implies that $U = V$. So we can use $\Sigma^{1/2} = U\Lambda^{1/2}U^T$ for the simulation of Y. Indeed we have that

$$\begin{aligned}
\Sigma = \Sigma^{1/2}(\Sigma^{1/2})^T &= U\Lambda^{1/2}V^T(U\Lambda^{1/2}V^T)^T = U\Lambda^{1/2}V^TV\Lambda^{1/2}U^T \\
&= U\Lambda^{1/2}\Lambda^{1/2}U^T = U\Lambda U^T = U\Lambda V^T \\
= (\Sigma^{1/2})^T\Sigma^{1/2} &= (U\Lambda^{1/2}V^T)^TU\Lambda^{1/2}V^T = V\Lambda^{1/2}U^TU\Lambda^{1/2}V^T \\
&= V\Lambda^{1/2}\Lambda^{1/2}V^T = V\Lambda V^T = U\Lambda V^T.
\end{aligned}$$

We want to generate $Y(\cdot)$ a Gaussian random field given n locations, the mean μ, the variance σ^2 and the correlation function ρ. We assume the Gaussian random field being stationary, that is μ and σ^2 are constant. Since we are given the correlation function ρ, we can first compute the covariance matrix Σ, corresponding to the multivariate Gaussian distribution of the random field at locations x_i. We know that with $h_{ij} = \|x_i - x_j\|$,

$$\text{Cov}\{Y(x_i), Y(x_j)\} = \sigma^2\rho(h_{ij}); \qquad \text{thus } \Sigma = \sigma^2 D,$$

where D is the correlation matrix. We define the correlation matrix as

$$D = (d_{ij})_{i,j=1,\ldots,n} = \rho(\|x_i - x_j\|)_{i,j=1,\ldots,n} = \rho(h_{ij})_{i,j=1,\ldots,n},$$

where $x_1, \ldots, x_n$ are the given locations. We finally derived the following method to generate $Y(\cdot)$ at locations x_i.

The algorithm for direct methods in $\mathbb{R}$

1. Define the vector $x = (x_1, \ldots, x_n)$ of given location of size n.

2. Compute the covariance matrix Σ at the locations x_i by calculating $\Sigma = \sigma^2 \rho(\|x_i - x_j\|)_{i,j=1,\dots,n}$.

3. Generate a vector N of n independent $\mathrm{N}(0,1)$ random variables.

4. If the selected method is SVD,

   Apply the SVD method to get U and Λ.

   Define $\Sigma^{1/2} = U\Lambda^{1/2}U^T$.

   Else,

   Apply the Cholesky factorization to get Q.

   Define $\Sigma^{1/2} = Q$.

5. Return $Y = \mu + \Sigma^{1/2}N$.

3.1.1 Validation of the method

Theoretically our code should work, but we should find a way to attest that in practice it also does.

We will validate our code by means of the variogram, that is by comparing the empirical variogram and the theoretical variogram. Let us first define the empirical variogram.

Definition 3.1 (Empirical variogram)**.**
For a random field $Y(\cdot)$, we call *empirical variogram* the quantity

$$\hat{\gamma}_c(h) = \frac{1}{\mid N(h) \mid} \sum_{(i,j)\in N(h)} \mid Y(\mathbf{x}_i) - Y(\mathbf{x}_j) \mid^2,$$

where $\mathbf{x}_i$ are the location parameters at which the Gaussian random field $Y(\cdot)$ will be evaluated and $N(h) = \{(i,j) \mid \|\mathbf{x}_i - \mathbf{x}_j\| = h\}$.

The empirical variogram, given by $\hat{\gamma}_c(h) = \frac{1}{|N(h)|}\sum_{N(h)} \mid Y(\mathbf{x}_i) - Y(\mathbf{x}_j) \mid^2$, is the estimated variogram of $Y(\cdot)$. It is an unbiased estimate for the true variogram $\gamma_{c_Y}(h)$, where $h = \|\mathbf{x}_i - \mathbf{x}_j\|$.

We have seen that the variogram, see Definition (2.11), is defined as

$$\gamma_c(h) = \sigma^2\{\rho(0) - \rho(h)\}.$$

So to compute the theoretical variogram, we only need to specify the covariance function, that is the type of correlation function and the chosen parameters, then the variogram is known.

Table 3.1: *Correlation families and parameters used for the simulations.*

	Parameter ν	**Scale ϕ**	**Nugget τ^2**
Exponential		1	0
Whittle-Matérn	3/2	1	0
Cauchy	1/2	1	0
Stable	3/2	1	0

3.1.2 Presentation of the results

We will now present the results obtainned by simulating Gaussian random fields with the correlation families and parameters presented in Table 3.1.

On each graphic, we have plotted the true variogram (solid lines) and the box plot of the empirical variograms for the Gaussian random field generated by Cholesky factorization or singular value decomposition. The left figure (red) corresponds to the Gaussian random field simulated with a Cholesky factorization method and the other one (blue) corresponds to the singular value decomposition method. We plotted the boxplots in ten equidistant intervals. In order to obtain the boxplots, we used a Monte Carlo experience with 500 repetitions of the simulation of a Gaussian random fields, where one simulation includes 50 replications of the Gaussian random field with the same parameters, correlation function and the same 100 locations. On each boxplot we added a cross (purple) representing the mean value.

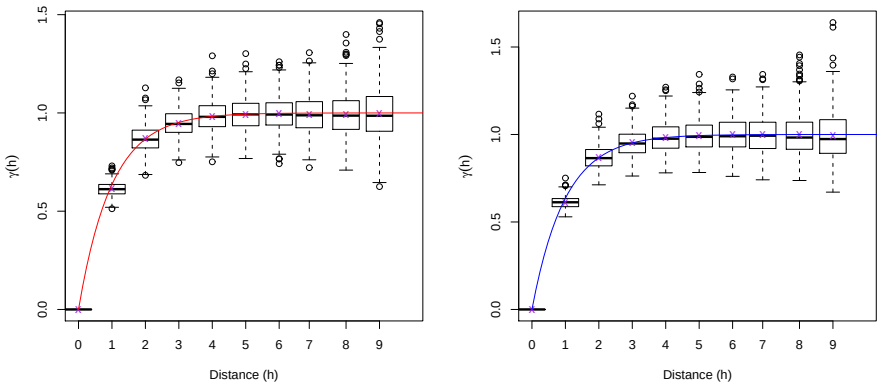

Figure 3.1: *Variogram comparison for a Gaussian random field with correlation following an exponential family and simulated by Cholesky factorization on the left and singular value decomposition on the right. We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field at 100 locations. Crosses represent the mean values.*

In Figure 3.1, we see that the average empirical variograms for the direct methods are really close to the true variogram for both direct simulation methods, especially when the distances are small, i.e. $h \leq 3$. Moreover the true variogram goes through the mean on the boxplot of the empirical variogram for any $h \leq 9$ for the Singular value decomposition method except for $h = 1$. For the Cholesky factorization method, we see that the true variogram is slightly to the right of the mean at $h = 1$, is again good for $h \geq 5$. We globally see that the empirical variogram overestimates the theoretical variogram, but the theoretical curve still remains in the confidence intervals of the boxplots.

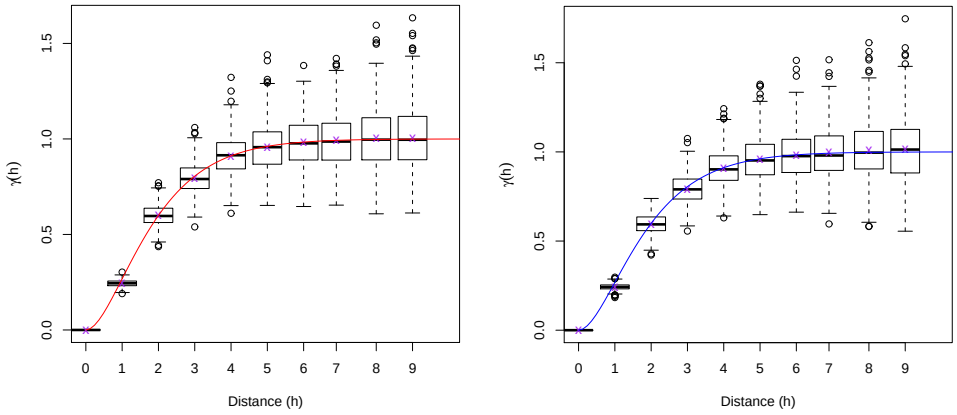

Figure 3.2: *Variogram comparison for a Gaussian random field with a Whittle-Matérn family function with $\nu = 1.5$ as correlation simulated by direct methods, Cholesky factorization on the left and singular value decomposition on the right. We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field at 100 locations. The crosses represent the mean values.*

In Figure 3.2, we see that, as for the exponential family, the two empirical variograms are extremely close to the true variogram. Indeed for the Cholesky factorization method the theoretical variogram goes quite always through the mean values of the empirical variograms. For the singular value decomposition method, the variograms are also close to each other, especially for small distances, i.e. $h \leq 8$. We again see that for both methods the theoretical value at $h = 1$ is overestimated by the empirical variogram. We see that this time with larger distances the theoretical variogram is underestimated by the empirical variogram, but remains in the confidence intervals of the boxplots.

In Figure 3.3, we see again that empirical variograms are pretty good estimates of the theoretical variogram. We remark that both the Cholesky

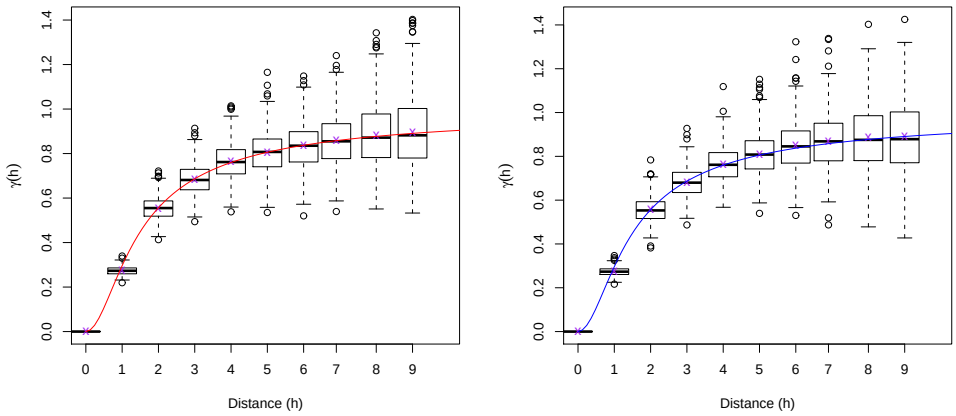

Figure 3.3: *Variogram comparison for a Gaussian random field having a Cauchy family with parameter $\nu = 0.5$ as correlation function, simulated by direct methods. The left figure corresponds to Cholesky factorization method and the right one to singular value decomposition method. We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field at 100 locations. The crosses represent the mean values.*

factorization method and the Singular value decomposition method seem to slightly underestimate the theoretical variogram for high values $h \geq 8$ and $h \geq 6$ respectively. They also overestimate the theoretical variogram at $h = 1$. We see that the Cholesky factorization method seems to lead to a closer estimate of the true variogram than the singular value decomposition method. Indeed there are more mean value points crossed by the theoretical variogram when we use Cholesky factorisation method than with the singular value decomposition method. We have that the empirical variograms remain in the confidence intervals of the boxplots.

In Figure 3.4, we see that both empirical variograms are really close to the theoretical variogram. Indeed we remark that the true variogram (in blue) passes through all mean values of the boxplots of the empirical variogram for both the Cholesky factorization and the Singular value decomposition method except at $h = 1$. At that point $h = 1$ we have that the empirical variogram is overestimating the theoretical one.

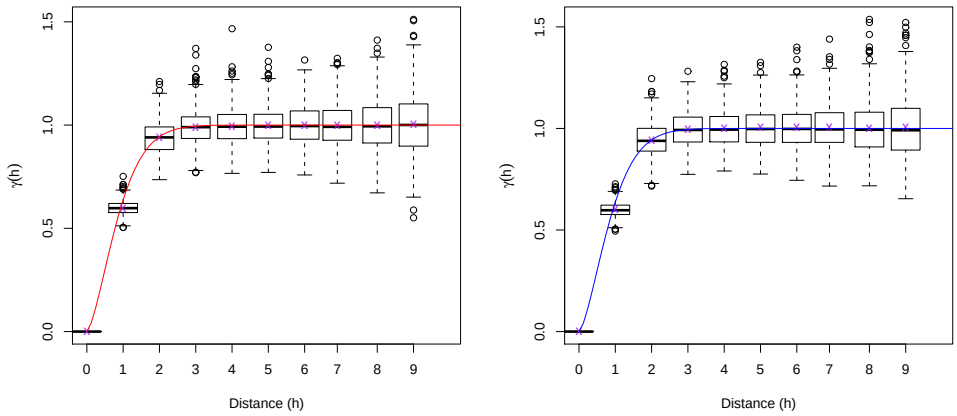

Figure 3.4: *Variogram comparison for a Gaussian random field with correlation function following a stable family with parameter $\nu = 1.5$ simulated by direct methods, left figure by Cholesky factorization and right figure by singular value decomposition. We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field at 100 locations. The crosses represent the mean values.*

To conclude we have seen firstly that both methods globally give average empirical variograms that are well estimating the theoretical variogram for all those choices of correlation functions, but we have difficulties to get the right slope at the origin. Indeed we notice that the true variogram is not passing through the mean values of the boxplots at $h = 1$, point at which the empirical variogram tends to overestimate slightly the theoretical variogram. We remark that the best estimate of the theoretical variogram is achieved when the correlation function is a stable family with parameter $\nu = 1.5$ and that the Cauchy model with $\nu = 0.5$ leads to worse estimations than the other correlation function families.

3.2 Presentation of the circular embedding method

Now that we have seen how to simulate a Gaussian random field by direct methods at n locations, where n is not too large, we want to present another method that can deal with larger n, namely the circular embedding method. We will first describe this method theoretically and then explain how we can implement it in **R**. We saw that we had really good convergence with direct method and would like to keep this property. We will in fact, with the circular embedding method, have a similar property called exactness in

principle. Let us though introduce it.

Definition 3.2 (Exact in principle)**.**
A realization that would have exactly the required distribution if

- the computer arithmetic had no inaccurencies;
- genuinely independent and random numbers could be used instead of creating pseudo-random numbers,

is said to be *exact in principle*.

The circular embedding method is exact in principle for correlation functions that have compact support. This method was developed independently by Dietrich and Newsam [1993] and Wood and Chan [1994] and published quite simultaneously. This method is fast and the idea behind it is to embed the covariance matrix in a circulant matrix and use fast Fourier transforms to treat it. The major restriction of this method is that the location points have to be points on a regular grid. Wood and Chan [1994] showed that the algorithm of the circular embedding method is always exact in principle for the correlation functions that have compact support. Wood and Chan [1994] also propose an approximation procedure otherwise.

3.2.1 The circular embedding method

First, we suggest to have a look at Chan [1999] for an overview of the circular embedding method and for more details on the results we give in this section, we refer to Wood and Chan [1994].

As previously with the direct methods, we aim to simulate a Gaussian Random Field $Y(\cdot)$ in $\mathbb{R}^d$ at a set of n locations $\mathbf{x}_j$, this time restricted on a regular grid in $\mathbb{R}^d$, with mean vector μ and correlation function ρ. We will restrict ourself to $d = 1$ that is to a Gaussian random field in $\mathbb{R}$ to get simpler calculations. We will choose the locations x_j such that they are equidistant in $[0, 1]$, that is $x_j = 0, \frac{1}{n}, \ldots, \frac{n-1}{n}$. We recall that μ and ρ define $Y(\cdot)$ uniquely, see Definition (2.2). We assume without loss of generality that $\mu \equiv 0$. We will finally assume that the Gaussian random field is stationary and isotropic. Therefore we have that

$$\Sigma = \sigma^2 \cdot \begin{pmatrix} \rho(0) & \rho(\frac{1}{n}) & \cdots & \cdots & \rho(\frac{n-1}{n}) \\ \rho(\frac{1}{n}) & \rho(0) & \ddots & \cdots & \rho(\frac{n-2}{n}) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \rho(0) & \rho(\frac{1}{n}) \\ \rho(\frac{n-1}{n}) & \rho(\frac{n-2}{n}) & \cdots & \rho(\frac{1}{n}) & \rho(0) \end{pmatrix},$$

where σ^2 is the variance, which is constant by isotropy assumption, and ρ is the correlation function. We notice that the matrix Σ is Toeplitz, see Definition (3.3).

Definition 3.3 (Toeplitz)**.**
We say that a matrix T is *Toeplitz* or *diagonal-constant* if each descending diagonal of T from left to right is constant.

We will now define a special case of Toeplitz matrix, namely a circulant matrix.

Definition 3.4 (Circulant matrix)**.**
We say that a matrix C is *circulant* if C is Toeplitz and each row vector $(c_{j,k})_{k=1}^{n}$ is a permutation from one element to the right of the preceding row vector, that is

$$c_{j,k} = c_{j-1,k-1} \quad \text{and} \quad c_{j,1} = c_{j-1,n}, \quad 2 \leq k \leq n,\ 2 \leq j \leq n.$$

We remark that a circulant matrix C is fully specified by one vector c, for example the first column of C. Indeed, the remainning columns of C are each cyclic permutation of the vector c. These matrices have the following nice property.

Proposition 3.1. [Golub and Van Loan, 1996, §4.7.7]
A circulant matrix C is such that its eigenvalues are given by applying a fast Fourier transform (FFT) on its first column c and the eigenvectors of C do not depend on C.

Let us now recall the fact that

$$Y(\cdot) \stackrel{d}{=} \Sigma^{1/2} N,$$

where N is vector of n independent $N(0,1)$ and n the number of locations. The key of the circulant embedding method is that instead of using a direct method to find $\Sigma^{1/2}$, we embed Σ in a bigger matrix C of size $m \times m$ which is circulant, with m defined as

$$m = 2^g \geq 2(n-1), \quad g \in \mathbb{N}.$$

We choose m in such a way since the radix-2 Cooley-Tukey FFT algorithm, that deals with matrices of size $m \times m$ with m being a power of 2, is known to compute the FFT with only $O\{m\ log_2(m)\}$ operations. This is faster than classical FFT algorithms, which run in $O\{m\ log(m)\}$. For more details, see [Cooley and Tukey, 1965]. The value of m must also be chosen such that C is positive semidefinite. If the correlation function has compact support, then the existence of a m such that C is positive definite is ensured. This is a consequence of the Theorem (3.2), which as been proved by Wood and Chan [1994]. Indeed if ρ has a compact support it follows that

$$\sum_h |\rho(h)| = \sigma^2 \sum_h |\mathrm{Cov}(h)| < \infty,$$

which is a requirement in the following theorem.

Theorem 3.2. Wood and Chan [1994]
Suppose that

$$\sum_h |Cov(h)| < \infty$$

and the spectral density

$$g(t) = (2\pi)^{-d} \sum_{j \in \mathbb{Z}^d} Cov\left(\frac{j}{n}\right) \exp\left(-2\pi i j^T t\right)$$

is strictly positive for all $t \in [0,1]^d$. *Then* C *in positive definite.*

In practice, we search for the smallest integer g such that

$$2^g = m \geq 2(n-1)$$

and check for the positive semi-definiteness of C. If C is not positive semidefinite, we increase g by one and repeat until C is positive semidefinite. This matrix C is defined as

$$C = \begin{pmatrix} c_0 & c_1 & \cdots & c_{m-1} \\ c_{m-1} & c_0 & \cdots & c_{m-2} \\ \vdots & \vdots & & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{pmatrix}$$

with

$$c_j = \begin{cases} \sigma^2 \rho(\frac{j}{n}) & 0 \leq j \leq \frac{m}{2}, \\ \sigma^2 \rho(\frac{m-j}{n}) & \frac{m}{2} < j \leq m-1. \end{cases}$$

Note that since C is circulant, it is uniquely defined by its first column c and since C is symmetric we have that c is also the first row of C. Proposition (3.1) implies that the eigenvalues of C are obtainned by applying a fast Fourier transform on c. We remark that the top left corner of C is equal to Σ by the definition of C. We will now state another property of circulant matrices, namely.

Proposition 3.3. [Brockwell and Davis, 1991, §4.5]
For any circulant and symmetric matrix C, *there exists an unitary matrix* Q *such that* $C = Q\Lambda Q^*$ *and* $C^{1/2} = Q\Lambda^{1/2}Q^*$, *where* Λ *is a diagonal matrix with the eigenvalues of* C *down the diagonal and where* Q^* *stands for the transpose conjugate of* Q. *Moreover* Q *is such that its columns are the eigenvector of* C.

Let us finally state a property of the symmetric circulant matrices that is a consequence of Proposition (3.3).

Proposition 3.4. [Wood and Chan, 1994]
For any circulant and symmetric matrix C, *written as* $C = Q\Lambda Q^*$, *we can compute* $Q\mathbf{u}$ *by applying a Fast Fourier Transform to* $\mathbf{u}$, *for any vector* $\mathbf{u}$.

In particular, this is true for $\mathbf{u} = \Lambda^{1/2}Q^*N$, where N is a vector of m independent $N(0,1)$. Thus we can generate the Gaussian random field $Y(\cdot)$ such that $Y(\cdot) \stackrel{d}{=} C^{1/2}N$, given by

$$Y(\cdot) \stackrel{d}{=} Q\Lambda^{1/2}Q^*N,$$

where N is a vector of m independent $N(0,1)$. Finally, if we consider only the n first elements of $Y(\cdot) \stackrel{d}{=} Q\Lambda^{1/2}Q^*N$, we get

$$Y(\cdot) \stackrel{d}{=} \Sigma^{1/2}N,$$

where N is this time a vector of n independent $N(0,1)$. Thus we have that $Y(\cdot)$ is a Gaussian random field with mean $\mu = 0$ and correlation function ρ evaluated at the n locations x_j as aimed.

So in order to simulate a Gaussian random field $Y(\cdot)$ in $\mathbb{R}$, it remains to compute $Q\Lambda^{1/2}Q^*N$, where N is a vector of m independent $N(0,1)$, as efficiently as possible. By Proposition (3.4), we know that only $\Lambda^{1/2}Q^*N$ needs to be computed ingeniously. We also have that $\Lambda^{1/2}$ can easily be calculated, since it is given by taking the square root of the elements in Λ, which is obtainned by a FFT of the first row of C by Proposition (3.1). Thus, it remains to simulate Q^*N. In order to have the best efficiency, we want to simulate Q^*N directly.

Proposition 3.5. [Wood and Chan, 1994]
We can write Q^*N as

$$Q^*N = S + iT,$$

where S and T are vectors of m independent $N(0,1)$. Moreover S and T are independent and their covariances are known at the location points x_j.

We finally derived the following method to generate $Y(\cdot)$ when the correlation function has compact support.

The algorithm for the circular embedding method in $\mathbb{R}$

1. Find the smallest integer g such that $m = 2^g \geq 2(n-1)$.

2. Compute c, the first row of C, the symmetric circulant matrix obtainned by embedding Σ, which is given by

$$c_j = \begin{cases} \sigma^2\rho(\frac{j}{n}) & 0 \leq j \leq \frac{m}{2}, \\ \sigma^2\rho(\frac{m-j}{n}) & \frac{m}{2} < j \leq m-1. \end{cases}$$

3. Compute λ the vector of eigenvalues of C by a fast Fourier transform on c.

4. If λ is negative,

Set $g \leftarrow g+1$ and $m \leftarrow 2^g$;

Go back to 2.

Else,

Calculate $\lambda^{1/2}$.

5. Generate S and T two independent $N(0,1)$.

6. Generate $\mathbf{S} = (S_1, \ldots, S_{m/2})$ and $\mathbf{T} = (T_1, \ldots, T_{m/2})$ two independent random vectors of $m/2$ independent $N(0,1)$.

7. Generate $\mathbf{u} = \Lambda^{1/2} Q^* N$ by calculating

$$\mathbf{u}(0) = \sqrt{\frac{\lambda_0}{m}} S, \quad \mathbf{u}(m/2) = \sqrt{\frac{\lambda_{m/2}}{m}} T,$$
$$\mathbf{u}(j) = \sqrt{\frac{\lambda_j}{2m}} (S_j + iV_j) \quad \text{and} \quad \mathbf{u}(m-j) = \overline{\mathbf{u}(j)}, \ 1 \leq j < m/2.$$

8. Apply a fast Fourier transform on $\mathbf{u} = \Lambda^{1/2} Q^* N$ to get $Q\mathbf{u}$, which is equal in distribution to $Y(\cdot)$ and redefine $Y(\cdot)$ as the n first elements of $Y(\cdot)$.

9. Return $Y(\cdot)$.

3.2.2 Approximate version

Suppose we have to simulate a Gaussian random field with a correlation function that does not have compact support. Thus the algorithm may not work since the existence of a m such that the circulant matrix C is positive definite is not ensured. We want to find a way to adapt the method for those cases even if we will loose the exactness in principle property. First we need to detect those "failure" cases. A simple way to do this is to put a higher bound on m as well. If the higher bound is reached and C is still not positive definite we say that we have a "failure" case.

The approximate circulant embedding approach suggests to consider only the part of C that corresponds to its positive eigenvalue values, with given m. Let us fix m to the smallest value such that $m = 2^g \geq 2(n-1)$ with g being an integer, that is

$$m = 2^{1+\lceil \log_2(n-1) \rceil}.$$

Then the corresponding matrix C is circulant, so it can be decomposed as

$$C = Q\Lambda Q^* = Q(\Lambda_+ - \Lambda_-)Q^* = C_+ - C_-,$$

where

$$\Lambda_+ = \text{diag}\{\ \max(0, \lambda_j), \forall j\}\ \text{ and}$$
$$\Lambda_- = \text{diag}\{\ -\min(0, \lambda_j), \forall j\}\,,$$

with λ_j being the j^{th} eigenvalue of C. Then we use instead of C the symmetric, positive semi-definite approximate embedding matrix $\varrho^2 C_+$, with suitable choice of $\varrho \neq 0$. Wood and Chan [1994] suggest two choices of ϱ, which are

$$\varrho_1 = \frac{\text{tr}(\Lambda)}{\text{tr}(\Lambda_+)} \quad \text{and} \quad \varrho_2 = \left\{ \frac{\text{tr}(\Lambda)}{\text{tr}(\Lambda_+)} \right\}^{1/2}.$$

They justify those choices by the fact that ϱ_2 leads to the correct one-dimensional marginal distribution and ϱ_1 is the minimizer to the lower bound of the random error incurred by setting the negative eigenvalues of C to zero. Thus we derive an approximate method with the following changes

1. Find the smallest integer g such that $m = 2^g \geq 2(n-1)$ and initialize $k = 1$.

4. If λ is negative and $k \leq 6$,

Set $g \leftarrow g + 1$ and $m \leftarrow 2^g$,

Set $k \leftarrow k + 1$ and

Go back to 2.

Else if λ is non-negative,

Calculate $\lambda^{1/2}$.

Else

Set $m = 2^{1+\lceil \log_2(n-1) \rceil}$.

Run Step 2.

Compute λ' by a fast Fourier transform on $\frac{\text{tr}(\Lambda)}{\text{tr}(\Lambda_+)} c$ with λ being the eigenvalues of C obtainned by a fast Fourier transform on c, $\Lambda = \text{diag}\{\lambda_j, \forall j\}$ and $\Lambda_+ = \text{diag}\{\ \max(0, \lambda_j), \forall j\}$.

Set $\lambda^{1/2} \leftarrow (\lambda')^{1/2}$.

3.2.3 Validation of the algorithm and presentation of the results

Theoretically our code should work, but we should find a way to attest that in practice it also does. As we did for the direct methods, we will validate our code by means of the variograms, that is by comparing the empirical variogram, see Definition (3.1), and the theoretical variogram, see Definition (2.11).

We will now present the results we obtainned by simulating Gaussian random fields with different correlation functions. We will first consider the spherical correlation family, which has a compact support and should behave pretty well with the circulant embedding method in the sense that the method is ensured to be exact in principle. We will then consider some examples we saw by direct simulation to compare both methods.

On each graphic, we have plotted the true variogram (solid line) and the boxplot of the empirical variograms for the Gaussian random field generated by the circulant embedding method. The boxplots are taken in ten equidistant intervals of the empirical variogram for 500 simulations of the Gaussian random field with the circular embedding method, where one simulation includes 50 replications of the random field on a regular grid at 500 locations. On each boxplot we added a cross representing the mean value.

In Figure 3.5, we see that the convergence of the circular embedding method for spherical correlation family is pretty good. Indeed the true variogram, that is the red curve, remains in the confidence intervals of the boxplots. We see that the empirical variogram is in average slightly overestimating the true variogram for $h \geq 3$. For small distances, that is $h \leq 2$ we have a perfect estimation.

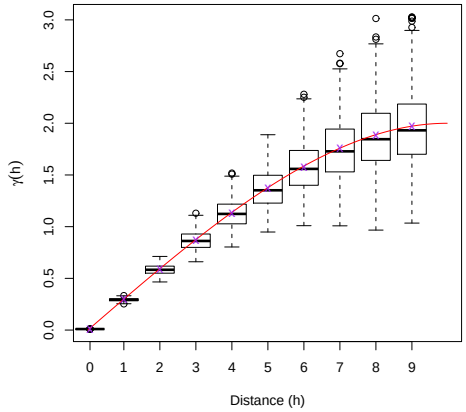

Figure 3.5: *Variogram comparison for a Gaussian random field having a spherical family as correlation function, simulated by circular embedding methods on* 500 *locations on the regular grid* $[0, 1]$. *We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field for the boxplots of the empirical variogram. Crosses represent the mean values.*

We will now present the variogram comparisons for the same correlation function families as with direct methods whose parameters can be seen in Table 3.2.

Table 3.2: *Correlation families and parameters used for the simulations.*

	μ	σ^2	**Parameter** ν	**Scale** ϕ	**Nugget** τ^2
Stable	3	2	3/2	1	0
Exponential	3	2		1	0
Whittle-Matérn	3	2	3/2	1	0
Cauchy	3	2	1/2	1	0

But now the Gaussian random field will be generated by the circulant embedding method and we will consider 500 locations on a regular grid on $[0, 1]$. The boxplots of the empirical variograms are done over 500 simulations of the same Gaussian random field replicated 50 times.

In Figure 3.6, we see that the average empirical variogram of the Gaussian random field is overestimating the true variogram. We see that the shape of the true variogram is quite similar to the one of the empirical variogram, which suggests a shifting of the true variogram from around 0.1 to the top. Indeed we think there is some bias. However we see that the method is not bad since the theoretical variogram globally remains in the confidence intervals of the boxplots, except for low values of h.

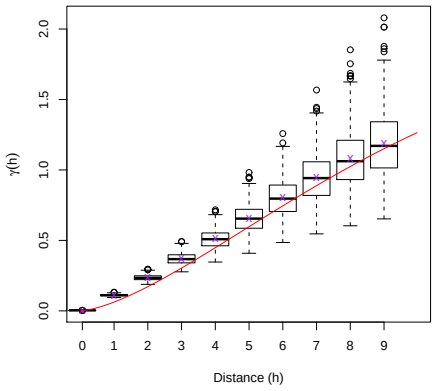

Figure 3.6: *Variogram comparison for a Gaussian random field having a stable family with parameter* $\nu = 1.5$ *as correlation function, simulated by circulant embedding methods on* 500 *locations on the regular grid* $[0, 1]$. *We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field for the boxplots of the empirical variogram. Crosses represent the mean values.*

We saw in Figure 3.6 that we had a slight bias, so we were interested to know if this happens only for the choice of parameter $\nu = 1.5$ or not. We see in Figure 3.7 that this does not only happen with a choice of parameter

$\nu = 1.5$, but that $\nu = 1.5$ is in a sense a limiting case. Indeed we see in Figure 3.7 on the left that with $\nu = 1.4$ we get no bias and a pretty good estimation of the theoretical variogram by the empirical one. In Figure 3.7 on the right, we also plotted the variogram comparison when $\nu = 2$, which corresponds to the limiting case (Gaussian model). As we mentionned in Example 2.3, this model is known to be numerically unstable. Therefore we can suggest an idea to explain the bias with $\nu = 1.5$. The intuition would say that with a parameter too close to the limiting case with $\nu = 2$, we get some numerical instability. Another explanation for this bad estimation could be that the spherical family has not the same shape for small values, $h \leq 1$. This could be a reason why the shape of the theoretical variogram at the origin seems to be wrong with $\nu = 2$. This also reflects this difficulty of our methods to have a good estimation even for small distances.

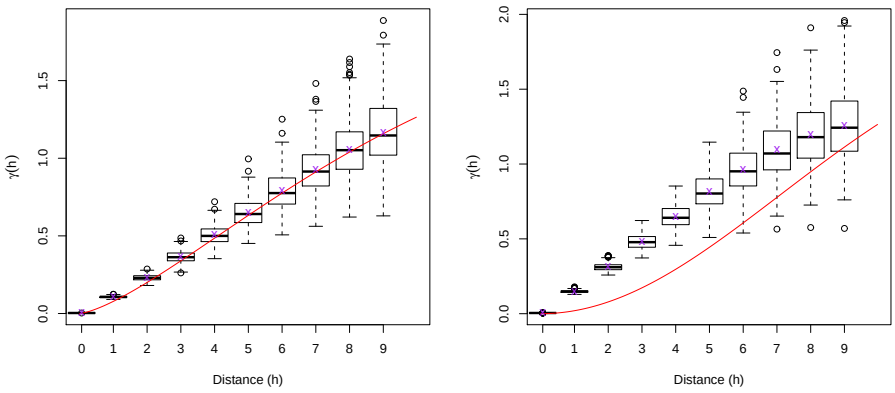

Figure 3.7: *Variogram comparison for a Gaussian random field having a stable family with parameter $\nu = 1$ on the left and $\nu = 2$ on the right as correlation function, simulated by circulant embedding methods on 500 locations on the regular grid $[0, 1]$. We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field for the boxplots of the empirical variogram. Crosses represent the mean values.*

In Figure 3.7, on the left, we see that the empirical variogram globally is a good estimate for the true variogram. Indeed the theoretical variogram passes quite through the mean values for large distances, i.e. $h \geq 6$. We see that there is a tendency of the empirical variogram to overestimate the true variogram. We have a bad estimation at $h = 1$, as the 95% confidence interval of the empirical variogram does not contain the value of the theoretical variogram.

In Figure 3.8, we see that there is a good estimation of the theoretical variogram by the empirical variogram. Indeed the red curve, which is the

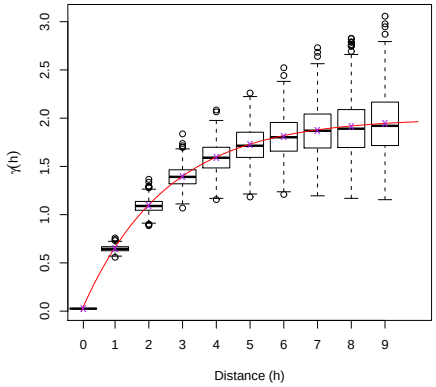

Figure 3.8: *Variogram comparison for a Gaussian random field having its correlation following an exponential family, simulated by circulant embedding methods on* 500 *locations on the regular grid* $[0, 1]$. *We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field. The crosses represent the mean values.*

theoretical variogram, quite passes through the mean values. When the theoretical variogram does not pass through the mean values of the boxplots, the true variogram is slightly overestimated by the empirical variogram.

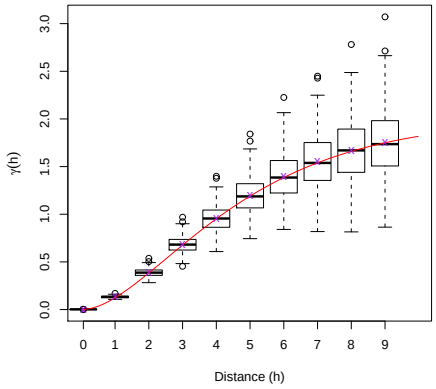

Figure 3.9: *Variogram comparison for a Gaussian random field with Whittle-Matérn correlation function with parameter* $\nu = 1.5$, *simulated by circulant embedding methods on* 500 *locations on the regular grid* $[0, 1]$. *We used 500 repetitions of the simulation, where one simulation includes 50 replications of the random field. Crosses represent the mean values.*

In Figure 3.9, we see that the empirical variogram is slightly overestimating the true variogram for large distances, i.e. $h \geq 7$, but the theoretical variogram remains in the 95 % confidence interval of the boxplots. We also remark that for $h = 1$ we do not have perfect estimation. We can therefore say that the circular embedding method does really well for simulating a Gaussian random field with exponential family correlation function.

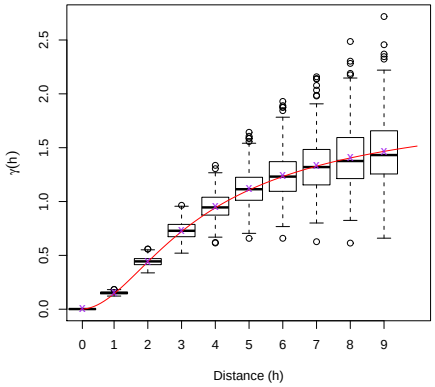

Figure 3.10: *Variogram comparison for a Gaussian random field having a Cauchy family with parameter $\nu = 0.5$ as correlation function, simulated by circulant embedding methods on 500 locations on the regular grid $[0, 1]$. We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the Gaussian random field for the boxplots of the empirical variogram. The crosses represent the mean values.*

In Figure 3.10, we see that the theoretical variogram remains in the confidence interval of the boxplots, but the empirical variogram is slightly overestimating the true variogram. For any value of h except for $h \geq 3$, we even have that the true variogram passes through the mean values of the boxplots of the empirical variograms. We can therefore say that the circular embedding method does really well for simulating a Gaussian random field with Cauchy family with parameter $\nu = 0.5$ as correlation function.

We briefly mention that similarly to the direct methods the circular embedding method gives average empirical variograms that are well estimating the theoretical variogram for any choice of correlation functions, but we have difficulties to get the right slope at the origin, except with the spherical family as correlation function.

3.3 Computation time comparison

Now that we presented two different kind of methods for simulating a Gaussian random field and we are interested in their computation time. Computation times for each of the methods are reported in Table 3.3.

Table 3.3: *Computation time in seconds for the three methods: Cholesky factorisation(chol), singular value decomposition (svd) and circular embedding (c.e.). The times are the mean over 10 estimations with standard errors where one observation contains 20 replications of the process.*

	n=50	n=100	n=200	n=500	n=1000
Exponential					
chol	0.01(0.01)	0.01(0.01)	0.04(0.01)	0.37(0.02)	1.75(0.10)
svd	0.02(0.01)	0.05(0.00)	0.24(0.01)	2.84(0.03)	19.73(0.23)
c.e.	0.02(0.00)	0.03(0.01)	0.04(0.01)	0.08(0.00)	0.16(0.01)
Whittle-Matérn					
chol	0.02(0.01)	0.06(0.01)	0.19(0.01)	1.41(0.08)	5.67(0.19)
svd	0.02(0.01)	0.08(0.01)	0.40(0.01)	3.94(0.12)	22.70(0.35)
c.e.	0.01(0.01)	0.03(0.01)	0.03(0.01)	0.10(0.04)	0.17(0.02)
Cauchy					
chol	0.01(0.01)	0.02(0.01)	0.06(0.01)	0.55(0.02)	2.32(0.08)
svd	0.01(0.01)	0.02(0.01)	0.06(0.00)	0.50(0.07)	2.25(0.07)
c.e.	0.01(0.01)	0.03(0.00)	0.03(0.01)	0.08(0.01)	0.15(0.01)
Stable					
chol	0.00(0.00)	0.02(0.01)	0.05(0.01)	0.48(0.02)	2.02(0.07)
svd	0.00(0.01)	0.01(0.01)	0.05(0.01)	0.41(0.02)	1.98(0.07)
c.e	0.01(0.01)	0.03(0.00)	0.04(0.01)	0.26(0.59)	0.16(0.03)

We see in Table 3.3 that the computation time increases when the number of locations increases, but not every method increases the same. We see that the increase of computation time is also depending on the correlation function of the Gaussian random field. We notice that with low number of locations, that is $n \leq 50$, the methods are all equivalent. We see that when the number of locations starts to become high, the computation time of the direct methods also become quite high, especially for the singular value decomposition method.

To confirm this idea we tried to apply our methods for 5000 and $10'000$ locations, but we did this only for the correlation function being exponential as a matter of example. We got the following results. Already with 5000 locations it is impossible to simulate a random field with the singular value decomposition method and it takes around 2 minutes with the Cholesky factorisation method whereas the circular embedding method needs around

38 seconds. With $10'000$ locations, both the direct methods fail, that is we can generate a random field only with the circular embedding method which takes around 3.5 seconds. And finally with $50'000$ it took around 7.5 seconds.

Chapter 4

From Gaussian random fields to max-stable processes

In the same way random fields are an infinite-dimensional generalisation of multivariate distribution theory, we can view the max-stable processes as an infinite-dimensional extension of the multivariate extreme value theory. Those max-stable random fields are used to model for example the maximal precipitation level at different sites of a spatial process. Smith [1990] gives two arguments that advantage the max-stable processes based approach over the multivariate extreme value approach for the problem of spatial rainfall collected on a grid of points in space. The general representation of max-stable processes is due to de Haan [1984]. We will see that max-stable processes spectral representation and max-stable processes simulation are highly related to Poisson point processes. Therefore, we start this chapter by some reminders about Poisson point processes.

4.1 Poisson point processes

Point processes can be used in practice for modelling a wide range of natural phenomenes such as earthquake epicentres, trees in a forest or population in settlements.

Definition 4.1 (Random point process)**.**
A random set in $\mathbb{R}^d$ whose realisations are made up of a finite or countable number of points is called a *random point process* in $\mathbb{R}^d$.

More simply said a point process $\mathcal{P}$ is a collection of points. We can then define a random point measure $\mathcal{P}_n$ as

$$\mathcal{P}_n(A) = \sum_{j=1}^{n} \delta_{X_j}(A)$$

for some random n which is the number of points in a suitable set A, where δ_x puts unit mass at $x \in A$ and X_j represent the positions of the points.

Among the point processes there are two important ones that are locally finite point processes. Those are the Cox processes and the Poisson point processes. We will here focus only on Poisson point processes, for the Cox processes we refer for example to Lantuéjoul [2002].

Definition 4.2 (Poisson point process)**.**
Let λ be an intensity function of S in $[0, \infty)$ which is locally integrable for all compact sets $B \subseteq S$. A Point process X on $S \subseteq \mathbb{R}^d$, is a *Poisson point process* with intensity function λ if the following properties are verified, where Λ is an intensity measure such that

$$\Lambda(B) = \int_B \lambda(\xi) \mathrm{d}\xi, \quad B \subseteq S.$$

Moreover Λ is locally finite, that is $\Lambda(B) < \infty$ for compact set $B \subseteq S$, and Λ is diffuse, that is $\Lambda(\xi) = 0$ for $\xi \in S \setminus B$.

- For any compact set $B \subseteq S$, the number of points N in B is a Poisson random variable with mean $\Lambda(B)$, that is

  $$\mathrm{P}\left\{N(B) = n\right\} = \frac{\{\Lambda(B)\}^n}{n!} \, e^{-\Lambda(B)}$$

  with the convention that

  $$\mathrm{P}\left\{N(B) = n\right\} = 0, \quad \text{if } \Lambda(B) = +\infty.$$

- For any finite family of pairwise disjoint compact sets $B_1, \ldots, B_m \subseteq S$, the random variables $N(B_1), \ldots, N(B_m)$ are mutually independent.

We note that a Poisson point process on S with rate or intensity λ is called an homogeneous Poisson point process on S if λ is constant, whereas it is called an inhomogeneous process when λ is not constant. Moreover, an unit rate or standard Poisson point process is a process with $\lambda \equiv 1$.

In our simulation of max-stable processes we will be interested in the points of a Poisson point process, especially to simulate them. Theorem (4.1) gives way to obtain a standard Poisson point process form exponential variables.

Theorem 4.1. *Let ξ_i be i.i.d. standard exponentially distributed random variables, then*

$$\Pi = \left\{ \sum_{k=1}^{n} \xi_k \colon n = 1, 2, \ldots \right\}$$

is a Poisson point process on the positive real axis with rate one.

Sketch of the proof.
We first note that since ξ are i.i.d. standard exponential random variable

1. $\xi_i \geq 0$ for all i and thus $\sum_{i=1}^{k} \xi_i \leq \sum_{i=1}^{k+1} \xi_i$, for all k;
2. $\mathrm{P}(\xi_i \leq t) = 1 - e^{-t}$ for $t \in \mathbb{R}_+$;
3. $\sum_{i=1}^{k} \xi_i$ is gamma distributed with parameters $(k, 1)$, that is

$$\mathrm{P}\left(\sum_{i=1}^{k} \xi_i = t\right) = \frac{t^{k-1}}{(k-1)!} e^{-t}.$$

Let us denote the number of points in $[0, t]$ by $N(t)$, then

$$\begin{aligned}
\mathrm{P}\left(N(t) = k\right) &\overset{(1)}{=} \mathrm{P}\left(\sum_{i=1}^{k} \xi_i \leq t,\ \sum_{i=1}^{k+1} \xi_i > t\right) \\
&= \int_0^t \mathrm{P}\left(\sum_{i=1}^{k} \xi_i \leq t,\ \sum_{i=1}^{k+1} \xi_i > t \mid \sum_{i=1}^{k} \xi_i = \tau\right) \mathrm{P}\left(\sum_{i=1}^{k} \xi_i = \tau\right) \mathrm{d}\tau \\
&= \int_0^t \mathrm{P}\left(\xi_{k+1} > t - \tau\right) \mathrm{P}\left(\sum_{i=1}^{k} \xi_i = \tau\right) \mathrm{d}\tau \\
&\overset{(2),(3)}{=} \int_0^t e^{-(t-\tau)} \cdot \frac{\tau^{k-1}}{(k-1)!} e^{-\tau} \mathrm{d}\tau = e^{-t} \int_0^t \frac{\tau^{k-1}}{(k-1)!} \mathrm{d}\tau \\
&= e^{-t} \frac{t^{k-1}}{(k-1)!}.
\end{aligned}$$

So N is Poisson distributed and thus Π is a Poisson point process. □

We will now present a theorem that allows the transformation of a point process to another point process, which with help of Theorem (4.1) will give us a way to simulate any Poisson point process.

Theorem 4.2.
Let E_1, E_2 be two Hausdorff spaces. Let ξ_1, ξ_2 be the associated σ-fields. Let $T\colon (E_1, \xi_1) \longrightarrow (E_2, \xi_2)$ be measurable. If N is a Poisson point process of intensity λ on E_1, then

$$\hat{N} \overset{d}{=} N \circ T^{-1}$$

is a Poisson point process with intensity $\lambda \circ T^{-1}$ on E_2. Moreover if we have the representation

$$N(\cdot) = \sum_j \delta_{X_j}(\cdot),$$

then

$$\hat{N}(\cdot) \overset{d}{=} N \circ T^{-1}(\cdot) = \sum_j \delta_{T(X_j)}(\cdot).$$

A proof of this theorem is given by Resnick [1986].

Let us now consider an example of application of Theorem (4.2), that will be useful for simulation of a max-stable process.

Example 4.1. *Let* $\{\chi_i\}_{i\geq 1}$ *be defined as* $\chi_i = \sum_{k=1}^{i} \xi_k$*, where* ξ_k *are i.i.d. standard exponential random variables. Consider the application*

$$T\colon x \longmapsto \frac{1}{x},$$

then we have that

$$\hat{\chi}_i = \frac{1}{\sum_{k=1}^{i} \xi_k}, \quad i \geq 1$$

is a Poisson point process with intensity $1/\xi^2 d\xi$*.*

Thus if we want to simulate an inhomogeneous Poisson point process, we first simulate a standard Poisson point process on a subset in $\mathbb{R}$ and then apply the transformation procedure given in Theorem (4.2).

4.2 Max-stable processes

We will now present the basic theory about max-stable random fields.

Definition 4.3 (Max-stable process)**.**
A random field (or random process) $Z(\cdot)$ on $\mathbb{R}^d$ is *max-stable* if there exist continuous functions $a_n(\mathbf{x}) > 0$ and $b_n(\mathbf{x})$ such that $\{Z(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ is equal in distribution to

$$Z^*(\mathbf{x}) = \frac{\left\{\max\limits_{1\leq i\leq n} Z_i(\mathbf{x})\right\} - b_n(\mathbf{x})}{a_n(\mathbf{x})}, \mathbf{x} \in \mathbb{R}^d,$$

where $Z_i(\cdot)$ are independent and identically distributed copies of $Z(\cdot)$.

We can without loss of generality transform the margins to one particular extreme value distribution, see [Resnick, 1987]. For convenience, we assume that the max-stable process $Z(\cdot)$ has unit Fréchet margins, that is,

$$\mathrm{P}(Z(\mathbf{x}) \leq z) = e^{-1/z}, \quad \mathbf{x} \in \mathbb{R}^d,$$

where

$$Z(\cdot) \stackrel{d}{=} \max_{1\leq i\leq n} \frac{Y_i(\cdot)}{n}.$$

Thus, $a_n(\mathbf{x}) = n$ and $b_n(\mathbf{x}) = 0$. Those processes are interesting. Indeed de Haan [1984] has shown that, provided that the limit exists,

$$Z(\mathbf{x}) = \lim_{n\longrightarrow\infty} \frac{\left\{\max\limits_{1\leq i\leq n} Y_i(\mathbf{x})\right\} - b_n(\mathbf{x})}{a_n(\mathbf{x})}, \mathbf{x} \in \mathbb{R}^d$$

is a max-stable process on $\mathbb{R}^d$, where $Y_i(\mathbf{x})$ are independent copies of a random field $Y(\cdot)$.

L. de Hann has brought a big contribution to the theory of max-stable processes, for example a spectral representation of the max-stable process is given in [de Haan, 1984].

Theorem 4.3 (Spectral representation)**.** [de Haan, 1984]
Let $(\xi_i, \tau_i)_{i\geq 1}$ be an enumeration of the points in the Poisson process on $\mathbb{R}_+ \times [0,1]$ with intensity measure $\mathrm{d}\Lambda(\xi,\tau) = (\xi^{-2}d\xi) \times \nu(d\tau)$, where τ is a finite positive measure on $[0,1]$. Let $\left\{f(\tau_i, \mathbf{x}), \tau_i \in [0,1], \mathbf{x} \in \mathbb{R}^d\right\}$ be a non-negative function with $\int_{[0,1]} f(s,\mathbf{x})\nu(ds) = 1$ for any $\mathbf{x} \in \mathbb{R}^d$. Then

$$Z(\mathbf{x}) = \max_i \left\{\xi_i f(\tau_i, \mathbf{x})\right\}, \mathbf{x} \in \mathbb{R}^d$$

is a max-stable process on $\mathbb{R}^d$.

Since we assume that we have max-stable processes with unit Fréchet margins, we will prefer the spectral representation from Schlather [2002]. We present here a variant of his version where the condition

$$\mathrm{E}\left[\max\left\{0, Y(\mathbf{x})\right\}\right] = \mu \in (0,\infty)$$

is changed to Equation (4.2.1) in Theorem (4.4) and the intensity of the Poisson process accordingly renormalised.

Theorem 4.4.
Let $(\xi_i)_{i\geq 1}$ be the points of a Poisson process on $\mathbb{R}_+^*$ with intensity measure $\mathrm{d}\Lambda(\xi) = \xi^{-2}d\xi$. Let $Y(\cdot)$ be a stationary random field on $\mathbb{R}^d$ such that

$$\mathrm{E}\left[\max\left\{0, Y(\mathbf{x})\right\}\right] = 1 \tag{4.2.1}$$

and

$$\mathrm{E}\left[\sup_{\mathbf{x}\in\mathbb{R}^d}\left\{Y(\mathbf{x})\right\}\right] < \infty.$$

Let $\left\{Y_i(\cdot)\right\}_{i\geq 1}$ be independent and identically distributed copies of $Y(\cdot)$. Then

$$Z(\mathbf{x}) = \max_{i\geq 1}\left[\xi_i \max\left\{0, Y_i(\mathbf{x})\right\}\right], \mathbf{x} \in \mathbb{R}^d$$

is a stationary max-stable process on $\mathbb{R}^d$ with unit Fréchet margins.

Sketch of the proof.

$$\begin{aligned}
\mathrm{P}\left\{Z(\mathbf{x}) \leq z\right\} &= \mathrm{P}\left[\text{no } \left\{\xi, Y(\mathbf{x})\right\} \in \mathbb{R}_+^* \times \mathbb{R}^d \text{ with } \xi_i \max\left\{0, Y(\mathbf{x})\right\} > z\right] \\
&= \exp\left\{-\int_{\mathbb{R}} \int_{z/\max(0,y)}^{\infty} f_{Y(\mathbf{x})}(y)\mathrm{d}y\xi^{-2}\mathrm{d}\xi\right\} \\
&= \exp\left\{-\int_{\mathbb{R}} \max(0,y)\frac{1}{z} f_{Y(\mathbf{x})}(y)\mathrm{d}y\right\} \\
&= \exp\left\{-\frac{1}{z}\int_{\mathbb{R}_+} y f_{Y(\mathbf{x})}(y)\mathrm{d}y\right\} \\
&= \exp\left\{-\frac{1}{z}\mathrm{E}\left[\max\left\{0, Y(\mathbf{x})\right\}\right]\right\} = \exp\left\{-\frac{1}{z}\right\},
\end{aligned}$$

the superposition of n i.i.d. Poisson point processes form a Poisson point process with its intensity multiplied by n and

$$\begin{aligned}
&\mathrm{P}\left\{\tilde{Z}(\mathbf{x}_1) \leq t_1, \ \ldots, \tilde{Z}(\mathbf{x}_k) \leq t_k\right\} \\
&\qquad = \mathrm{P}\left[\max_{1\leq i\leq n}\left\{Z_i(\mathbf{x}_1)\right\} \leq nt_1, \ldots, \max_{1\leq i\leq n}\left\{Z_i(\mathbf{x}_k)\right\} \leq nt_k\right] \\
&\qquad \overset{\text{ind}}{=} \prod_{i=1}^{n} \mathrm{P}\left\{Z_i(\mathbf{x}_1) \leq nt_1, \ldots, Z_i(\mathbf{x}_k) \leq nt_k\right\} \\
&\qquad \overset{\text{i.d.}}{=} \mathrm{P}\left\{Z_1(\mathbf{x}_1) \leq nt_1, \ldots, Z_1(\mathbf{x}_k) \leq nt_k\right\}^n \\
&\qquad = \exp\left\{-V(nt_1, \ldots, nt_k)\right\}^n = \exp\left\{-\frac{1}{n}V(t_1, \ldots, t_k)\right\}^n \\
&\qquad = \mathrm{P}\left\{Z(x_1) \leq t_1, \ldots, Z(x_k) \leq t_k\right\}.
\end{aligned}$$

□

We see in those two representation theorems (4.3) and (4.4) that the construction of a max-stable process $Z(\cdot)$ involves the maximum over an infinite number of copies of a random field $Y(\cdot)$, but in practice we can only simulate finitely many realisations of $Y(\cdot)$. The next theorem presents conditions under which we may get nonetheless exact simulations for $Z(\cdot)$ with a finite number of realisations.

Theorem 4.5.
Let $Y(\cdot)$ be a stationary random field on $\mathbb{R}^d$, let Π be a Poisson point process on $\mathbb{R}^d \times \mathbb{R}_+^*$ with intensity measure $\mathrm{d}\Lambda(\mathbf{y}, \xi) = \mu^{-1}\mathrm{d}\mathbf{y}\xi^{-2}\mathrm{d}\xi$ and let $Z(\cdot)$ be defined as

$$Z(\mathbf{x}) = \sup_{(\mathbf{y},\xi)\in\Pi} \left[\xi \max\left\{0, Y_\xi(\mathbf{x} - \mathbf{y})\right\}\right].$$

Assume that $Y(\cdot)$ is uniformly bounded by $C \in \mathbb{R}_+^*$ and has support in the ball $\mathrm{b}(0, r)$ for some $r \in \mathbb{R}_+^*$. Let B be a compact set in $\mathbb{R}^d$. Let

$Y_i(\cdot)$ be i.i.d. replications of $Y(\cdot)$, let U_i be i.i.d. uniformly distributed on $B_r = \cup_{x\in B} \mathrm{b}(x,r)$ and let ξ_i be i.i.d. standard exponential random variables. Finally assume that Π, $Y_i(\cdot)$, U_i and ξ_i are all mutually independent. Then, on B,

$$Z^*(\mathbf{x}) = \frac{|B_r|}{\mu} \sup_{1\leq i\leq m} \left\{ \frac{Y_i\left(\mathbf{x}-U_i\right)}{\sum_{k=1}^{i}\xi_k} \right\}, \quad \mathbf{x}\in B$$

almost surely equals the max-stable process $Z(\cdot)$, where m is such that

$$\frac{C}{\sum_{k=1}^{m}\xi_k} \leq \max_{1\leq i\leq m} \left\{ \frac{Y_i\left(\mathbf{x}-U_i\right)}{\sum_{k=1}^{i}\xi_k} \right\}.$$

A proof of this theorem is given by Ribatet [2009] or by Schlather [2002]. It is also mentionned in Schlather [2002] that, for random fields whose support is not included in a ball $\mathrm{b}(0,r)$ or which are not uniformly bounded by a constant C, we can nevertheless use approximations for r and C. He considers $Y(\cdot)$ being a Gaussian random field but $Y(\cdot)$ is not uniformly bounded by $C<\infty$. He suggests that $C=3$ is large enough to get good approximations. Corollary (4.6) is the adaptation of Theorem (4.5) to Gaussian random fields.

Corollary 4.6.
Let $Y(\cdot)$ be a stationary random field on $\mathbb{R}^d$, let Π be a Poisson point process on $\mathbb{R}_+^*$ with intensity measure $\mathrm{d}\Lambda(\xi) = \xi^{-2}\mathrm{d}\xi$ and let $Z(\cdot)$ be defined as

$$Z(\mathbf{x}) = \sup_{i\geq 1} \left[\xi_i \max\left\{0, \sqrt{2\pi}Y_i(\mathbf{x})\right\} \right].$$

Assume that $Y(\cdot)$ is a standard Gaussian process in $\mathbb{R}_+^*$. Let $Y_i(\cdot)$ be i.i.d. replications of $Y(\cdot)$ and let ξ_i be i.i.d. standard exponential random variables. Finally assume that Π, $\sqrt{2\pi}Y_i(\cdot)$ and ξ_i are all mutually independent. Then

$$Z^*(\mathbf{x}) = \sup_{1\leq i\leq m} \left\{ \frac{\sqrt{2\pi}Y_i\left(\mathbf{x}\right)}{\sum_{k=1}^{i}\xi_k} \right\}, \quad \mathbf{x}\in B$$

almost surely equals the max-stable process $Z(\cdot)$, where m is such that

$$\frac{3}{\sum_{k=1}^{m}\xi_k} \leq \max_{1\leq i\leq m} \left\{ \frac{Y_i(\mathbf{x})}{\sum_{k=1}^{i}\xi_k} \right\}.$$

Note that we take the maximum between 0 and $\sqrt{2\pi}Y(\cdot)$, this is because $Y(\cdot)$ is a Gaussian random field and thus

$$E\left[\max\left\{0, Y(\mathbf{x})\right\}\right] = \frac{1}{\sqrt{2\pi}},$$

which violates Condition (4.2.1) in Theorem (4.4). By redefining $Y(\cdot)$ to $\sqrt{2\pi}Y(\cdot)$, we have that

$$E\left[\max\left\{0, \sqrt{2\pi}Y(\mathbf{x})\right\}\right] = 1.$$

By means of Corollary (4.6), we are now ready to present M. Schlather's method for simulating a max-stable random field $Z(\cdot)$ at a given set of locations.

Algorithm for simulation of Max-stable processes

1. Initialize the vector $Z(\mathbf{x}) = \{Z(x_1), \ldots, Z(x_{n_{\text{site}}})\} \leftarrow 0$.
   Set $\zeta \leftarrow 0$ and $C \leftarrow 3 \cdot \sqrt{2\pi}$.

2. While $\kappa \neq 0$, do

   - Initialize a counter κ to $\kappa \leftarrow n_{\text{site}}$;
   - Generate $\xi \sim$Exp(1) and set $\zeta \leftarrow \zeta + \xi$;
   - Set the upper limit $u \leftarrow C\zeta^{-1}$;
   - Generate a standard Gaussian process Y at the locations $\{x_i\}_{i=1}^{n_{\text{site}}}$ with correlation function ρ;
   - For $i = 1$ to n_{site}, do
     - If $u > Z(x_i)$, do
       Update $Z(x_i) \leftarrow \max\left\{Z(x_i), \sqrt{2\pi}\zeta^{-1}Y(x_i)\right\}$;
     - Else
       Update $\kappa \leftarrow \kappa - 1$.

3. Return $Z(\mathbf{x}) = \{Z(x_1), \ldots, Z(x_{n_{\text{site}}})\}$.

4.2.1 Validation of the algorithm

Theoretically our code should work, but we should find a way to attest that in practice it also does.

By analogy to what we did for the other methods, we would like to validate our code by means of the variogram, but the variogram is not defined for max-stable processes with unit Fréchet margins since their expectation and variance are not finite. We will present a variogram-based approach that is valid for max-stable processes, namely when the expectations and variances may not be finite. This estimator based on the variogram concept is called F-madogram and is highly related to the extremal coefficient.

Definition 4.4 (F-madogram)**.**
Let $Z(\cdot)$ be a stationary max-stable random field with unit Fréchet margins, that is

$$F(z) = \exp(-\frac{1}{z}).$$

Then, the *centred semi-F-madogram* is defined as follows

$$\nu(h) = \frac{1}{2}\mathrm{E}\left[\, \left| F\left\{Z(x_1)\right\} - F\left\{Z(x_2)\right\} \right| \,\right],$$

where $h = |x_1 - x_2|$.

We will use the term F-madogram instead of centred semi-F-madogram. As we said the F-madogram and the extremal coefficient are related. Let us recall some basis of extreme value theory and then introduce what the extremal coefficient function for max-stable processes with unit Fréchet margins is. Any multivariate extreme value distribution has the form

$$\mathrm{P}\left\{Z(x_1) \leq z_1, \ldots, Z(x_k) \leq z_k\right\} = \exp\left\{-V(z_1, \ldots, z_k)\right\},$$

where V is an homogeneous function with order depending on the margins. For example when the margins are unit Fréchet, then V is homogeneous of order -1. That is

$$V(z, \ldots, z) = -\frac{1}{z}V(1, \ldots, 1).$$

Definition 4.5 (Extremal coefficient)**.**
Let $Z(\cdot)$ be a max-stable process with unit Fréchet margins then the *extremal coefficient function* is defined as

$$\theta(\mathbf{h}) = -z \log\left[\mathrm{P}\left\{Z(\mathbf{x}_1) \leq z, Z(\mathbf{x}_2) \leq z\right\}\right],$$

where $\mathbf{h} = \mathbf{x}_1 - \mathbf{x}_2$.

The extremal coefficient function $\theta(\mathbf{h})$, which is such that $1 \leq \theta(\mathbf{h}) \leq 2$, is a measure of the dependence between extremes. Indeed,

$$\mathrm{P}\left\{Z(\mathbf{x}_1) \leq z, Z(\mathbf{x}_2) \leq z\right\} = \exp\left\{-\frac{\theta(\mathbf{h})}{z}\right\} = \mathrm{F}(z)^{\theta(\mathbf{h})}.$$

Thus $\theta(\mathbf{h}) = 1$ corresponds to perfect dependence and $\theta(\mathbf{h}) = 2$ to independence.

We note that the bivariate distribution $\mathrm{P}\left\{Z(\mathbf{x}_1) \leq s, Z(\mathbf{x}_2) \leq t\right\}$, for the Schlather model, corresponds to

$$\exp\left\{-\frac{1}{2}\left(\frac{1}{t} + \frac{1}{s}\right)\left[1 + \sqrt{1 - 2\left\{\rho(h) + 1\right\}\frac{st}{(s+t)^2}}\,\right]\right\},$$

where $h = \|\mathbf{x}_1 - \mathbf{x}_2\|$ and $\rho(h)$ is the covariance function of the underlying stationary and isotropic Gaussian random fields. Thus for $s = t$, we get

$$\exp\left[-\frac{1}{t}\left\{1 + \sqrt{1 - \frac{\rho(h) + 1}{2}}\,\right\}\right],$$

This yields an extremal coefficient of

$$\begin{aligned}\theta(h) &= -t \log \left[\mathrm{P} \left\{ Z(\mathbf{x}_1) \leq t, Z(\mathbf{x}_2) \leq t \right\} \right] \\ &= 1 + \sqrt{1 - \frac{\rho(h)+1}{2}} = 1 + \sqrt{\frac{1-\rho(h)}{2}}.\end{aligned}$$

Let us now establish the link between the F-madogram and the extremal coefficient.

Theorem 4.7.
The F-madogram $\nu(h)$ of a stationary max-stable process with unitary Fréchet margins is related to the extremal coefficient function in the following way

$$2\nu(h) = \frac{\theta(h)-1}{\theta(h)+1}$$

and conversely, we have that

$$\theta(h) = \frac{1+2\nu(h)}{1-2\nu(h)}.$$

Proof. We first note that

1. $|x-y| = 2\max(x,y) - (x+y)$,
2. $\mathrm{P}\left[\max\left\{Z(\mathbf{x}_1), Z(\mathbf{x}_2)\right\} \leq z\right] = \exp\left\{-\frac{\theta(h)}{z}\right\}$ by definition of the extremal coefficient and
3. $\mathrm{E}\left[\mathrm{F}\left\{Z(\mathbf{x}_1)\right\}\right] = \mathrm{E}\left[\mathrm{F}\left\{Z(\mathbf{x}_2)\right\}\right] = \frac{1}{2}$, since $\mathrm{F}\left\{Z(\mathbf{x})\right\}$ is uniformly distributed.

Thus we have that

$$\begin{aligned}\nu(h) =& \frac{1}{2}\mathrm{E}\left[\ \left|\mathrm{F}\left\{Z(\mathbf{x}_1)\right\} - \mathrm{F}\left\{Z(\mathbf{x}_2)\right\}\right|\ \right] \\ \overset{(1)}{=}& \mathrm{E}\left[\max\left(\mathrm{F}\left\{Z(\mathbf{x}_1)\right\}, \mathrm{F}\left\{Z(\mathbf{x}_2)\right\}\right)\right] - \frac{1}{2}\left(2\mathrm{E}\left[F\left\{Z(\mathbf{x})\right\}\right]\right) \\ \overset{(3)}{=}& \mathrm{E}\left[\max\left(\mathrm{F}\left\{Z(\mathbf{x}_1)\right\}, \mathrm{F}\left\{Z(\mathbf{x}_2)\right\}\right)\right] - \frac{1}{2} \\ =& \frac{\theta(h)}{\theta(h)+1} - \frac{1}{2} = \frac{\theta(h)-1}{\theta(h)+1},\end{aligned}$$

since

$$\begin{aligned}&\mathrm{E}\left[\max\ \left(\mathrm{F}\left\{Z(\mathbf{x}_1)\right\}, \mathrm{F}\left\{Z(\mathbf{x}_2)\right\}\right)\right] \\ &\overset{(2)}{=} \int_{\mathbb{R}_+} e^{-1/z} \frac{\mathrm{d}}{\mathrm{d}z}\left[\exp\left\{-\frac{\theta(h)}{z}\right\}\right] = \int_{\mathbb{R}_+} \frac{\theta(h)}{z^2}\left[\exp\left\{-\frac{\theta(h)+1}{z}\right\}\right]\mathrm{d}z \\ &= \left(\frac{\theta(h)}{\theta(h)+1}\left[\exp\left\{-\frac{\theta(h)+1}{z}\right\}\right]\right)_0^\infty = \frac{\theta(h)}{\theta(h)+1}\end{aligned}$$

Finally, by solving $2\nu(h) = \frac{\theta(h)-1}{\theta(h)+1}$ for $\theta(h)$, we get $\theta(h) = \frac{1+2\nu(h)}{1-2\nu(h)}$. □

In order to make an analysis of our results, we need an unbiased estimator. A natural choice for the empirical F-madogram would be

$$\hat{\nu}(h) = \frac{1}{2|\mathcal{N}_h|} \sum_{(x_i,x_j)\in\mathcal{N}_h} |Z(x_j) - Z(x_i)|,$$

where $\mathcal{N}_h$ is the set of sample pairs lagged by the distance h with the corresponding extremal coefficient

$$\hat{\theta}(h) = \frac{1 + 2\hat{\nu}(h)}{1 - 2\hat{\nu}(h)}.$$

4.2.2 Presentation of the results

In Figure 4.1 we plotted boxplots of the empirical F-madograms and the true F-madogram (solid red line) of a stationary and isotropic max-stable process with unit Fréchet margins. Simulation is made from Schlather's model, where the standard Gaussian random field has a spherical family as correlation function and has been simulated by the circular embedding method. We consider a process at 500 locations on the regular grid $[0, 1]$. We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the max-stable process. Crosses represent the mean values.

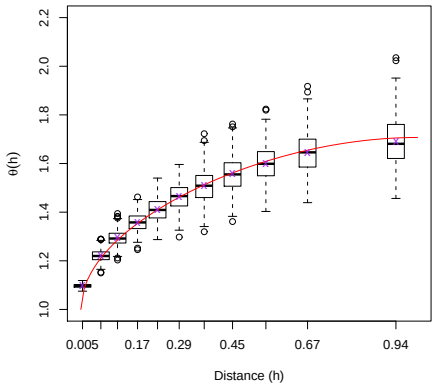

Figure 4.1: *F-madogram comparison for a max-stable process with unit Fréchet margins simulated according to Schlather's model and the standard Gaussian random field having a spherical family as correlation function, simulated by circular embedding methods on* 500 *locations on the regular grid* $[0, 1]$. *We used a Monte Carlo experience with 500 repetitions of the simulation, where one simulation includes 50 replications of the max-stable process for the boxplots of the empirical F-madogram.*

We see in Figure 4.1 that the theoretical F-madogram is fairly well estimated by the empirical F-madogram. The true F-madogram is overestimated by the empirical F-madogram for $h \leq 0.22$ and underestimated for $h \geq 0.45$. Thus we again see this tendency we had when simulating a Gaussian random field to have problem to estimate the true value for small distances. We also see that even if the extremal index is supposed to take values between 1 and 2, but we never reach the value of 2, that is we never get complete independence. This is a default of Schlather's model, whose extremal index is bounded $1 + \sqrt{1/2} \approx 1.7$.

Now that we have seen that our simulation method is reasonably good, let us look how one max-stable process looks like.

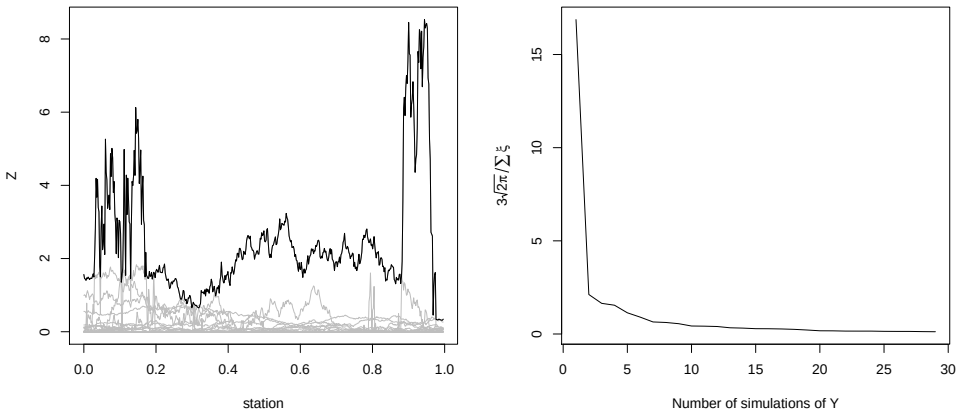

Figure 4.2: *In the left figure, we plotted in black one max-stable process and in light gray* $\max\left\{0, Y_i(\cdot)\sqrt{2\pi}/\xi\right\}$*, where* $Y(\cdot)$ *is a simulation by circular embedding method of a Gaussian random field with spherical correlation function at* 500 *locations. In the right figure we show the evolution of* $\frac{3\sqrt{2\pi}}{\sum_{i=1}^{k} \xi_i}$ *with* k *from 1 to* $m = 29$.

In Figure 4.2 we represented in the left figure one max-stable process (black) and plotted $\max\left\{0, Y_i(\cdot)\sqrt{2\pi}/\xi\right\}$ (light gray), where $Y(\cdot)$ is a simulation by circular embedding method of a Gaussian random field with spherical correlation function at 500 locations. The right figure shows the evolution of $\frac{3\sqrt{2\pi}}{\sum_{i=1}^{k} \xi_i}$ with k from 1 to $m = 29$, since 29 simulations of one Gaussian random field were needed to obtain this max-stable process. We see on the right Figure that the lower limit for the max-stable process quickly gets pretty small and that is why we can be sure that this method works and stops with relatively few iterations.

Conclusion

In conclusion we could simulate a max-stable process with help of the circular embedding mehtod.

In Chapter 2, we defined random fields and the properties of stationarity and isotropy. Then we discussed conditions for positive semidefinitness. We ended by giving some examples of valid correlation functions for isotropic and stationary Gaussian random fields.

In Chapter 3, we presented some methods for simulating stationary and isotropic Gaussian random fields with known correlation functions, at given locations in $\mathbb{R}$. We first gave two direct methods based on the Cholesky factorization and the Singular value decomposition respectively. They were good, but their computational time was sensibly increasing as the number of locations increased. Then we proposed another method, the circular embedding method which is exact in principle for correlation functions with compact support. We also gave an approximation procedure for the other correlation functions. This method was doing good and its only restriction was that the locations had to be on a regular grid. All our methods had the drawback that they had difficulties with too small distances. We ended this chapter by a comparison of the computation time needed by the methods depending on the number of locations and could confirm that the circular embedding method is fast.

Finally, in Chapter 4, we introduced Poisson point processes and the way one can simulate them. Then we described max-stable processes and came to a simulation procedure for a max-stable process according to the Schlather model, which is based on a simulation of a Gaussian random field and a Poisson process. So we could simulate a max-stable process with help of the circular embedding method. We saw that the simulation procedure works well but has the same drawback as the circular embedding method. We ended by presenting a max-stable process simulated by the procedure we suggested and comment on why this procedure was working.

Bibliography

P. Abrahamsen. A Review of Gaussian Random Fields and Correlation Functions. Technical report, Tech. Rep. 917, Norwegian Computing Center, 1997.

S. Bochner. A Theorem on Fourier–Stieltjes Integrals. *Mathematische Annalen*, 108:378–410, 1933.

P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer Series in Statistics. Springer, second edition, 1991. ISBN 0-387-97429-6.

R. Bruzual and M. Dominguez. Representation of operator valued measurable indefinite functions. *Acta Científica Venezolana*, 52(3):223–225, 2001.

G. Chan. An Effective Method for Simulating Gaussian Random Fields. In *Proceedings of the Statistical Computing Section*, pages 133–138. American Statistical Association, 1999.

J.W. Cooley and J.W. Tukey. An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*, 19(90):297–301, 1965.

L. de Haan. A Spectral Representation for Max-Stable Processes. *The Annals of Probability*, 12(4):1194–1204, 1984.

C.R. Dietrich and G.N. Newsam. A Fast and Exact Method for Multidimensional Gaussian Stochastic Simulations. *Water Resources Research*, 29(8): 2861–2869, 1993.

P.J. Diggle and P.J. Ribeiro Jr. *Model-based Geostatistics*. Springer Series in Statistics. Springer, New York, 2007. ISBN 978-0-387-32907-9.

T. Gneiting. *Symmetric Positive Definite Functions with Applications in Spatial Statistics*. PhD thesis, University Bayreuth, 1997.

T. Gneiting and Z. Sasvári. The Characterization Problem for Isotropic Covariance Functions. *Mathematical Geology*, 31(1):105–111, 1999.

G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 1996.

C. Lantuéjoul. *Geostatistical Simulation: Models and Algorithms*. Springer Verlag, Berlin Heidelberg, 2002. ISBN 3-540-42202-1.

B. Matérn. Spatial variation: Stochastic Models and their Application to some Problems in Forest Surveys and Other Sampling Investigations. *Meddelanden från Statens Skogsforskningsinstitut*, 49(5):1–144, 1960.

S.I. Resnick. Point Processes, Regular Variation and Weak Convergence. *Advances in Applied Probability*, 18(1):66–183, 1986.

S.I. Resnick. Extreme Values, Regular Variation, and Point Processes. *Advances in Applied Probability*, 4, 1987.

M. Ribatet. Max-stable processes: Theory and inference. Technical report, École Polytechnique fédérale de Lausanne, 2009. Conference.

F. Riesz. Über Sätze von Stone und Bochner. *Acta Univ. Szeged*, 6:184–198, 1933.

M. Schlather. Introduction to Positive Definite Functions and to Unconditional Simulations of Random Fields. Technical report, Techn. Rep. ST-99-10, 1999.

M. Schlather. Models for Stationary Max-Stable Random Fields. *Extremes*, 5(1):33–44, 2002.

I.J. Schönberg. Metric Spaces and Completely Monotone Functions. *The Annals of Mathematics*, 39(4):811–841, 1938.

R.L. Smith. Max-Stable Processes and Spatial Extremes. *Unpublished manuscript*, 205, 1990.

E. Vanmarcke. Random Fields: Analysis and Synthesis. *Cambridge, MA*, page 394, 1983.

A.T.A. Wood and G. Chan. Simulation of Stationary Gaussian Processes in $[0, 1]^d$. *Journal of Computational and Graphical Statistics*, 3(4):409–432, 1994.

A.M. Yaglom. Correlation Theory of Stationary and Related Random Functions. Volume I: Basic results. pages 1–526, 1987.