A Review of

Gaussian Random Fields and Correlation Functions

Second Edition

Petter Abrahamsen

April 1997

Preface

This review presents some properties of Gaussian random field models. The results given are largely taken from the literature and should be correct. The proofs however, are occasionally modified to compile with the general style and notation. So there are surely misprint and errors the reader must be indulgent to. The reader is assumed to be familiar with general concepts such as expectation and covariance and some knowledge of the formalism of probability theory is recommended.

This work is supported by a research fellowship from The Research Council of Norway.

Petter Abrahamsen

January 1994

Preface to the Second Edition

This edition is similar to the preceding one except for two major changes: Section 2.2 on derivatives has been reduced and Section 4.4 containing some examples of sample paths has been added. Also a few examples of spectral densities for correlation functions have been added in Section 3.2.2.

The notation has been slightly adjusted and several errors and misprints have been corrected. I am in particular grateful to Alfhild Lien Eide and Anne-Lise Hektoen who have pointed out several errors.

Petter Abrahamsen

April 1997

"When the going gets weird....
....the weird turns pro!"

Hunter S. Thompson

Contents

Introduction

There are several books and many articles on random fields and correlation functions. The main results however, are often complicated to extract without a major effort. This note is an attempt to collect results important for practical applications as well as results of more theoretical importance. The style is formal with most results stated as theorems. Long or complicated proofs are omitted, but references to the literature are supplied.

Most properties of random fields are related to finite dimensional distributions, that is, the probability distributions of the random field at an arbitrary (finite) set of positions. An exception is geometrical properties requiring characteristics of the probability space.

The study of Gaussian random fields is in many respects a study of covariance functions or correlation functions. The class of covariance functions coincidence with the class of positive definite functions. Thus, the concept of positive definiteness is of monumental importance and will be considered in detail. A general result usually referred to as *Bochner's theorem* states that any positive definite function has a simple spectral representation. A consequence of Bochner's theorem is the *Wiener-Khintchine theorem* which gives the spectral representation of correlation functions. This representation gives explicit methods for construction and validation of permissible correlation and covariance functions.

The first chapter starts by defining what is meant by a random field and in particular a Gaussian random field. Symmetry properties such as stationarity and isotropy are discussed and their implications for covariance and correlation functions are presented. The next chapter on geometrical properties is somewhat technical but the results are simple and of importance to the understanding of how random fields behave. Some examples are given to shed light on the more formal results such as conditions for continuity of random fields. The third chapter on correlation functions gives many results on how to obtain valid correlation functions. General methods for construction are given, and simple restrictions limiting the class of correlation functions are provided. The last chapter gives some examples of correlation functions found in the literature.

So what is left out? A lot! Everything about statistical inference is omitted. Also the topic of simulation of random fields so important in practical applications is hardly mentioned. Moreover, discussions of variograms and generalized covariances are completely excluded. This would require an examination of intrinsic random fields and generalized random fields, considered beyond the scope of this presentation.

The main influence on the contents is from Yaglom (1986*a*) and Matérn (1960). Many results are borrowed from them. The main influence on the chapter on geometrical properties is Adler (1981). Most of the fundamental results on positive definite functions are from Schoenberg (1938*a*, 1938*b*). The reference list is by no means complete but contains at least some of the more significant contributions. For additional references see e.g. Yaglom (1986*b*), Vanmarcke (1983), and Adler (1981).

1 Random Fields

A formal definition of a *random field* reads:

Definition 1.1 (Random field) Let a probability space, $(\Omega, \mathcal{F}, P)^*$, and a parameter set, T, be given. A random field is then a finite or real valued function $X(\mathbf{t}, \omega)$ which, for every fixed $\mathbf{t} \in T$ is a measurable function of $\omega \in \Omega$.

The synonyms *random function* and *stochastic process* are used by some authors.

The n-dimensional Euclidean space, $T = \mathbb{R}^n$, with $n \geq 1$ will be considered in the following. The dependency on the underlying probability space will usually be suppressed throughout the text:

$$X_{\mathbf{t}} = X(\mathbf{t}, \omega); \qquad \mathbf{t} \in \mathbb{R}^n.$$

For a fixed $\omega \in \Omega$, the function $X(\mathbf{t}, \omega)$ is a non-random function of $\mathbf{t}$. This deterministic function is usually called a *sample path* (*sample function*) or a *realization* and is denoted by a lower case letter $x_{\mathbf{t}}$. The parameter $\mathbf{t}$ is called the *coordinate* or *position* by standard terminology. In this context the formal definition simply means:

> *A random field $X_{\mathbf{t}}$ on $\mathbb{R}^n$ (i.e. $\mathbf{t} \in \mathbb{R}^n$) is a function whose values are* random variables *for any $\mathbf{t} \in \mathbb{R}^n$.*

The dimension of the coordinate is usually in the range from one to four, but any $n > 0$ is possible. A one-dimensional random field is usually called a *stochastic process*. The term 'random field' is usually used to stress that the dimension of the coordinate is higher than one. Random fields in two and three dimensions are encountered in a wide range of sciences and especially in the earth sciences such as hydrology, agriculture, and geology. Random fields where $\mathbf{t}$ is a position in space-time are studied in turbulence theory and in meteorology.

Consider the almost intuitive result that functions of random fields are also random fields:

Theorem 1.1 *For measurable functions $\mathbf{f}\colon \mathbb{R}^i \to \mathbb{R}^j$, and for a collection of i random fields $\{X^1_{\mathbf{t}}, \ldots, X^i_{\mathbf{t}}\}$, then $\mathbf{f}(X^1_{\mathbf{t}}, \ldots, X^i_{\mathbf{t}})$ is a collection of j random fields.*

This is a standard result from probability theory; measurable functions, $\mathbf{f}$, of measurable functions, $X^1_{\mathbf{t}}, \ldots, X^i_{\mathbf{t}}$, are again measurable functions. See for instance Loève (1978, pp. 103–111) or Billingsley (1986, pp. 182–184). Measurable functions include continuous functions, sums, products, and maxima. Thus the following corollary follows immediately.

Corollary 1.1.1 *For random fields on $\mathbb{R}^n$:*

(i) *If $X_{\mathbf{t}}$ and $Y_{\mathbf{t}}$ are random fields and $a, b \in \mathbb{R}^1$, then $aX_{\mathbf{t}} + bY_{\mathbf{t}}$ is a random field.*

(ii) *If $X_{\mathbf{t}}$ and $Y_{\mathbf{t}}$ are random fields, then $X_{\mathbf{t}}Y_{\mathbf{t}}$ is a random field.*

*A probability space consists of the sample space Ω of sample points ω, a σ-algebra $\mathcal{F}$ dividing Ω into subsets, and a probability measure P asigning probabilities to all members of $\mathcal{F}$. Each ω corresponds to a possible outcome of an 'experiment' and Ω containes all possible outcomes. The elements of $\mathcal{F}$ corresponds to events. These events must form a σ-algebra for the assignment of probabilities to be consistent and meaningful.

1.1 Finite-Dimensional Distributions

This section concerns the construction of valid probabilistic models for random fields. The properties of the parameter set, T, will not be considered until later. So $\mathbf{t}$ can in principal belong to almost any space T, e.g. $\mathbf{t} \in \mathbb{R}^n \otimes \{1, 2, 3\}$ such that

$$X_{\mathbf{t}} = \mathbf{X}_{\tilde{\mathbf{t}}} = \big(X_1(\tilde{\mathbf{t}}), X_2(\tilde{\mathbf{t}}), X_3(\tilde{\mathbf{t}})\big)\,; \qquad \tilde{\mathbf{t}} \in \mathbb{R}^n.$$

Thus the conditions required for existence are valid for random fields on far more general spaces than $\mathbb{R}^n$.

A random field is normally described by its *finite-dimensional (cumulative) distributions:*

$$F_{\mathbf{t}_1,\dots,\mathbf{t}_k}(x_1,\dots,x_k) = \mathrm{Prob}\big\{X_{\mathbf{t}_1} \le x_1,\dots,X_{\mathbf{t}_k} \le x_k\big\}. \tag{1.1}$$

The cumulative distribution functions are by definition *right-continuous* and *nondecreasing**. Note also that

$$F_{\mathbf{t}_1,\dots,\mathbf{t}_k}(x_1,\dots,-\infty,\dots,x_k) = 0, \qquad F_{\mathbf{t}_1,\dots,\mathbf{t}_k}(\infty,\dots,\infty) = 1.$$

Two consistency requirements on the finite-dimensional distributions must be satisfied. Consider a permutation π of the index set $\{1,\dots,k\}$. Since the events $[X_{\mathbf{t}_1} \le x_1,\dots,X_{\mathbf{t}_k} \le x_k]$ and $[X_{\mathbf{t}_{\pi 1}} \le x_{\pi 1},\dots,X_{\mathbf{t}_{\pi k}} \le x_{\pi k}]$ are identical the *symmetry condition*

$$F_{\mathbf{t}_1,\dots,\mathbf{t}_k}(x_1,\dots,x_k) = F_{\mathbf{t}_{\pi 1},\dots,\mathbf{t}_{\pi k}}(x_{\pi 1},\dots,x_{\pi k}) \tag{1.2}$$

must hold. A similar argument implies that the *compatibility condition*

$$F_{\mathbf{t}_1,\dots,\mathbf{t}_{k-1}}(x_1,\dots,x_{k-1}) = F_{\mathbf{t}_1,\dots,\mathbf{t}_{k-1},\mathbf{t}_k}(x_1,\dots,x_{k-1},\infty) \tag{1.3}$$

must be satisfied. Distribution functions $F_{\mathbf{t}_1,\dots,\mathbf{t}_k}$ originating from a random field according to (1.1) must necessarily satisfy conditions (1.2) and (1.3). The problem is therefore the converse: If there exist distribution functions $F_{\mathbf{t}_1,\dots,\mathbf{t}_k}$ which satisfy conditions (1.2) and (1.3), does there exist a random field having these finite dimensional distributions? The answer is yes according to the following famous theorem:

Kolmogorov's Existence Theorem *If a system of finite-dimensional distributions, $F_{\mathbf{t}_1,\dots,\mathbf{t}_k}$, satisfies the symmetry condition* (1.2) *and the compatibility condition* (1.3)*, then there exists on some probability space $(\Omega, \mathcal{F}, P)$ a random field $[X_{\mathbf{t}} : \mathbf{t} \in T]$ having $F_{\mathbf{t}_1,\dots,\mathbf{t}_k}$ as its finite-dimensional distributions.*

Two proofs are found in Billingsley (1986, pp. 513ff). This means that if it is possible to specify finite-dimensional distributions $F_{\mathbf{t}_1,\dots,\mathbf{t}_k}$ satisfying conditions (1.2) and (1.3) for any set of coordinates $\{\mathbf{t}_1,\dots,\mathbf{t}_k\}$, then a corresponding random field exist. The construction of such probability distributions are in general complicated, but certain special cases become simple.

The role of the underlying probability space, $(\Omega, \mathcal{F}, P)$, is usually irrelevant since the properties of the finite-dimensional distributions, $F_{\mathbf{t}_1,\dots,\mathbf{t}_k}$, determines almost all properties of practical interest. An important exception is the continuity and differentiability of *realizations*, $x_{\mathbf{t}}$, of a random field. See Example 2.1 for an almost trivial example.

*As k-dimensional distributions are considered, the ordinary definitions of 'right-continuous' and 'nondecreasing' must be generalized. The generalization is intuitive but the exact definition is somewhat cumbersome. Billingsley (1986, pp. 176–178) contains the details.

1.2 Expectation and Covariance

The *expectation* of a random field is by definition

$$m(\mathbf{t}) = \mathrm{E}\{X_{\mathbf{t}}\} = \int_{\Omega} X_{\mathbf{t}}(\omega)\,\mathrm{d}P(\omega).$$

This integral can be expressed by using the finite dimensional distribution $F_{\mathbf{t}}(x)$ as Stieltjes integrals in $\mathbb{R}^1$ (Billingsley 1986, p. 280):

$$m(\mathbf{t}) = \int_{\mathbb{R}^1} x\,\mathrm{d}F_{\mathbf{t}}(x). \tag{1.4}$$

The *(auto-)covariance function* is correspondingly expressed as a Stieltjes integral in $\mathbb{R}^2$:

$$\begin{aligned} C(\mathbf{t},\mathbf{s}) &= \mathrm{Cov}\{X_{\mathbf{t}}, X_{\mathbf{s}}\} = \mathrm{E}\{X_{\mathbf{t}}X_{\mathbf{s}}\} - m(\mathbf{t})\,m(\mathbf{s}) \\ &= \iint_{\mathbb{R}^2} x\,y\,\mathrm{d}^2F_{\mathbf{t},\mathbf{s}}(x,y) - m(\mathbf{t})\,m(\mathbf{s}), \end{aligned} \tag{1.5}$$

whereas the *variance* is

$$\sigma^2(\mathbf{t}) = C(\mathbf{t},\mathbf{t}).$$

The *(auto-)correlation function* of a random field is by definition

$$\rho(\mathbf{t},\mathbf{s}) = \mathrm{Corr}\{X_{\mathbf{t}}, X_{\mathbf{s}}\} = \frac{C(\mathbf{t},\mathbf{s})}{\sigma(\mathbf{t})\,\sigma(\mathbf{s})}. \tag{1.6}$$

In the following the existence of the covariance function is assumed.

For continuous distribution functions the *probability density function* is obtained from the partial derivatives (excuse the sloppy notation) as

$$p_{\mathbf{t}_1,\ldots,\mathbf{t}_k}(x_1,\ldots,x_k) = \frac{\partial^k F_{\mathbf{t}_1,\ldots,\mathbf{t}_k}(x_1,\ldots,x_k)}{\partial x_1\cdots\partial x_k}.$$

Then Stieltjes integrals can be written as ordinary Lebesgue (or Riemann) integrals:

$$\begin{aligned} &\int\cdots\int_B g(x_1,\ldots,x_k)\,\mathrm{d}^kF_{\mathbf{t}_1,\ldots,\mathbf{t}_k}(x_1,\ldots,x_k) \\ &\quad = \int\cdots\int_B g(x_1,\ldots,x_k)\,p_{\mathbf{t}_1,\ldots,\mathbf{t}_k}(x_1,\ldots,x_k)\,\mathrm{d}x_1\cdots\mathrm{d}x_k, \qquad B\in\mathcal{B}^k, \end{aligned}$$

where $\mathcal{B}^k$ are the Borel sets* in $\mathbb{R}^k$. So for continuous distribution functions, the expectation and covariance function are readily expressed by using the probability densities:

$$\begin{aligned} m(\mathbf{t}) &= \int_{\mathbb{R}^1} x\,p_{\mathbf{t}}(x)\,\mathrm{d}x, \\ C(\mathbf{t},\mathbf{s}) &= \iint_{\mathbb{R}^2} x\,y\,p_{\mathbf{t},\mathbf{s}}(x,y)\,\mathrm{d}x\,\mathrm{d}y - m(\mathbf{t})\,m(\mathbf{s}). \end{aligned}$$

*The Borel sets $\mathcal{B}^k$ include any subset of $\mathbb{R}^k$ ever encountered in practical applications. So $B\in\mathcal{B}^k$ simply means $B\subset\mathbb{R}^k$ and that the Lebesgue integral can be properly defined.

Integrals of the form (1.4) and (1.5) are called *Stieltjes integrals* (Billingsley 1986, p. 230). The significance of the Stieltjes integral is partly that it incorporates a density and partly that it allows discontinuities in $F_{\mathbf{t}_1,\ldots,\mathbf{t}_k}$. If the number of these are countable, the Stieltjes integral can be decomposed into an ordinary (Lebesgue) integral and a (countable) sum of contributions from the discontinuities. Thus, the Stieltjes integral is a notational construction that includes 'ordinary' integrals, sums, and combinations of these.

1.3 Positive Definiteness

The concept of *positive definiteness* is fundamental. For Gaussian random fields in particular, the positive definiteness of the covariance function is a sufficient and necessary condition for establishing consistent finite-dimensional distributions.

Definition 1.2 Let k be a positive integer, and let $\mathbf{t}_i \in T$ and $c_i \in \mathbb{R}^1$ for $i = 1, \ldots, k$. Then the function C on $T \otimes T$ is said to be positive (semi-)definite on T if

$$\sum_{i=1}^{k} \sum_{j=1}^{k} c_i c_j \, C(\mathbf{t}_i, \mathbf{t}_j) \geq 0 \tag{1.7}$$

for any choice of k, $\{\mathbf{t}_1, \ldots, \mathbf{t}_k\}$, and $\{c_1 \ldots, c_k\}$.

Consider an arbitrary collection $\{X_{\mathbf{t}_1}, \ldots, X_{\mathbf{t}_k}\}$ of random variables from a random field with covariance function C. Then, for arbitrary real numbers $\{c_1, \ldots, c_k\}$

$$\operatorname{Var}\{c_1 X_{\mathbf{t}_1} + \cdots + c_k X_{\mathbf{t}_k}\} = \sum_{i=1}^{k} \sum_{j=1}^{k} c_i c_j \, C(\mathbf{t}_i, \mathbf{t}_j) \geq 0.$$

This coincidence with the definition of positive definiteness so that the following lemma must be true.

Lemma *The class of covariance functions on T must belong to the class of positive definite functions on T.*

This is a strong restriction on the class of possible covariance functions and it is in general hard to verify.

Theorem 1.2 *The class of covariance functions coincidence with the class of positive definite functions.*

Proof. The Lemma says that the class of covariance functions necessarily belongs to the class of positive definite functions. What remains, is to prove the (almost trivial) opposite: That any positive definite function is a covariance function for some random field. [The arguments below follow Gnedenko (1962, pp. 371–372).]

Consider a Gaussian random field (defined in Section 1.4 below) defined such that $\mathrm{E}\{X_{\mathbf{t}}\} = 0$ and $\operatorname{Cov}\{X_{\mathbf{t}}, X_{\mathbf{s}}\} = C(\mathbf{t}, \mathbf{s})$ where C is a positive definite function. All finite dimensional distributions are readily deduced from these properties. Moreover the positive definiteness of C ensures the positive (semi-)definiteness of all covariance matrices, $\mathbf{\Sigma}$, appearing in the exponent of the finite dimensional multivariate normal probability densities. Thus the finite dimensional distributions exist and they satisfy the symmetry and compatibility conditions (see Section 1.4.1). So according to Kolmogorov's existence theorem the corresponding Gaussian random field exists. $\square$

Corollary 1.2.1 *The class of correlation functions coincidence with the class of positive definite functions where* $C(\mathbf{t},\mathbf{t}) = 1$.

Proof. $\rho(\mathbf{t},\mathbf{t}) = 1$ by definition. Since $\rho(\mathbf{t},\mathbf{s}) = C(\mathbf{t},\mathbf{s})/\big(\sigma(\mathbf{t})\,\sigma(\mathbf{s})\big)$ is

$$\sum_{i,j=1}^{k} c_i c_j\, \rho(\mathbf{t}_i,\mathbf{t}_j) = \sum_{i,j=1}^{k} \big(c_i/\sigma(\mathbf{t}_i)\big)\big(c_j/\sigma(\mathbf{t}_j)\big) C(\mathbf{t}_i,\mathbf{t}_j) = \sum_{i,j=1}^{k} c_i' c_j'\, C(\mathbf{t}_i,\mathbf{t}_j) \geq 0$$

for arbitrary $\{c_1 \ldots, c_k\}$ provided $\sigma(\mathbf{t}) > 0$. $\square$

The restriction $\sigma(\mathbf{t}) \neq 0$ is necessary for the correlation function to be properly defined. But this restriction is not necessary for the existence of $C(\mathbf{t},\mathbf{s})$; if $\rho(\mathbf{t},\mathbf{s})$ is positive definite then $C(\mathbf{t},\mathbf{s}) = \sigma(\mathbf{t})\,\sigma(\mathbf{s})\,\rho(\mathbf{t},\mathbf{s})$ is positive definite even though $\sigma(\mathbf{t}) = 0$ in certain regions since

$$\sum_{i,j=1}^{k} c_i c_j\, C(\mathbf{t}_i,\mathbf{t}_j) = \sum_{i,j=1}^{k} \big(c_i\sigma(\mathbf{t}_i)\big)\big(c_j\sigma(\mathbf{t}_j)\big)\rho(\mathbf{t}_i,\mathbf{t}_j) = \sum_{i,j=1}^{k} c_i' c_j'\, \rho(\mathbf{t}_i,\mathbf{t}_j) \geq 0.$$

The significance of positive definiteness will become clear when considering the construction of valid covariance and correlation functions. In particular the spectral representations give explicit ways of constructing such functions and even methods for validating positive definiteness. Many fundamental results on positive definite functions are given by Schoenberg (1938*a*, 1938*b*).

1.4 Gaussian Random Fields

Gaussian random fields play an important role for several reasons: The specification of their finite-dimensional distributions is simple, they are reasonable models for many natural phenomenon, estimation and inference are simple, and the model is specified by expectations and covariances.

Definition 1.3 A Gaussian random field is a random field where all the finite-dimensional distributions, $F_{\mathbf{t}_1,\ldots,\mathbf{t}_k}$, are multivariate normal distributions for any choice of k and $\{\mathbf{t}_1,\ldots,\mathbf{t}_k\}$.

Since multivariate normal distributions are completely specified by expectations and covariances, it suffices to specify $m(\mathbf{t})$ and $C(\mathbf{t},\mathbf{s})$ in such a way that conditions (1.2) and (1.3) holds. The expectation can be arbitrarily chosen, but the covariance function must be positive definite to ensure the existence of all finite-dimensional distributions.

In the following sections the discussion is focused on the expectation and in particular on the covariance function of a random field. Since these determine all stochastic properties of a Gaussian random field, the rest of this paper is in many respects a study of Gaussian random fields.

Be aware that the functions $m(\mathbf{t})$ and $C(\mathbf{t},\mathbf{s})$ do not uniquely specify a random field. For instance Yaglom (1986*a*, p. 88) discuss the *Poisson pulse* *process* in one dimension. It is possible to specify a Gaussian random field with identical expectation and covariance function but the two random fields behave totally different.

1.4.1 Existence

The multinormal probability densities involved are

$$p_{\mathbf{t}_1,\dots,\mathbf{t}_k}(x_1,\dots,x_k) = |2\pi\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^t\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mathbf{m})\right\}, \tag{1.8}$$

where $\mathbf{x}^t = (x_1,\dots,x_k)$, $\mathbf{m}^t = \big(m(\mathbf{t}_1),\dots,m(\mathbf{t}_k)\big)$ are the expectations, and $\Sigma_{ij} = C(\mathbf{t}_i,\mathbf{t}_j)$ are the elements of the covariance matrix.

The symmetry condition (1.2) requires that a simultaneous permutation of the sequence of $\{\mathbf{t}_1,\dots,\mathbf{t}_k\}$ and $\{x_1,\dots,x_k\}$ makes no difference, i.e.

$$\begin{aligned}\int_{-\infty}^{r_1}\cdots\int_{-\infty}^{r_k} & p_{\mathbf{t}_1,\dots,\mathbf{t}_k}(x_1,\dots,x_k)\,\mathrm{d}x_1\cdots\mathrm{d}x_k \\ &= \int_{-\infty}^{r_{\pi 1}}\cdots\int_{-\infty}^{r_{\pi k}} p_{\mathbf{t}_{\pi 1}\dots\mathbf{t}_{\pi k}}(x_{\pi 1},\dots,x_{\pi k})\,\mathrm{d}x_{\pi 1}\cdots\mathrm{d}x_{\pi k}.\end{aligned} \tag{1.9}$$

Too see this consider a permutation matrix $\mathbf{P}$ defined such that $(\pi 1,\dots,\pi k)^t = \mathbf{P}(1,\dots,k)^t$. The permutation of $\mathbf{x}$ becomes: $\mathbf{x}' = \mathbf{P}\mathbf{x} = \mathbf{P}(x_1,\dots,x_k)^t = (x_{\pi 1},\dots,x_{\pi k})^t$. The corresponding permutation of $\{\mathbf{t}_1,\dots,\mathbf{t}_k\}$ amounts to: $\mathbf{m}' = \mathbf{P}\mathbf{m}$ and $\boldsymbol{\Sigma}' = \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^t$. Since $\mathbf{P}^{-1} = \mathbf{P}^t$ (unitary) it follows that $|\mathbf{P}| = \pm 1$ which implies that $|\boldsymbol{\Sigma}'| = |\boldsymbol{\Sigma}|$, i.e. determinants are unchanged by permutations. Moreover $\boldsymbol{\Sigma}'^{-1} = \mathbf{P}^{t^{-1}}\boldsymbol{\Sigma}^{-1}\mathbf{P}^{-1} = \boldsymbol{\Sigma}^{-1'}$. Using this gives

$$\begin{aligned}(\mathbf{x}'-\mathbf{m}')^t\boldsymbol{\Sigma}'^{-1}(\mathbf{x}'-\mathbf{m}') &= (\mathbf{P}\mathbf{x}-\mathbf{P}\mathbf{m})^t\mathbf{P}\boldsymbol{\Sigma}^{-1}\mathbf{P}^t(\mathbf{P}\mathbf{x}-\mathbf{P}\mathbf{m}) \\ &= (\mathbf{x}-\mathbf{m})^t\mathbf{P}^t\mathbf{P}\boldsymbol{\Sigma}^{-1}\mathbf{P}^t\mathbf{P}(\mathbf{x}-\mathbf{m}) \\ &= (\mathbf{x}-\mathbf{m})^t\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mathbf{m}).\end{aligned} \tag{1.10}$$

Thus the quadratic form in the exponent of (1.8), $Q(\mathbf{x}) = (\mathbf{x}-\mathbf{m})^t\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mathbf{m})$, is also unchanged by simultaneous permutation of $\mathbf{x}$ and $\{\mathbf{t}_1,\dots,\mathbf{t}_k\}$. Since both $|\boldsymbol{\Sigma}|$ and $Q(\mathbf{x})$ are invariant under permutations:

$$p_{\mathbf{t}_1,\dots,\mathbf{t}_k}(x_1,\dots,x_k) = p_{\mathbf{t}_{\pi 1}\dots\mathbf{t}_{\pi k}}(x_{\pi 1},\dots,x_{\pi k}).$$

Hence, the symmetry condition (1.9) is satisfied.

The compatibility condition (1.3) requires that the $(k-1)$-dimensional multinormal 'marginal' density is obtained from the k-dimensional multinormal density:

$$p_{\mathbf{t}_1,\dots,\mathbf{t}_{k-1}}(x_1,\dots,x_{k-1}) = \int_{-\infty}^{\infty} p_{\mathbf{t}_1,\dots,\mathbf{t}_{k-1},\mathbf{t}_k}(x_1,\dots,x_{k-1},x_k)\,\mathrm{d}x_k.$$

This is a standard result for multinormal distributions.

So according to Kolmogorov's existence theorem the Gaussian random field is properly defined provided the $p_{\mathbf{t}_1,\dots,\mathbf{t}_k}(x_1,\dots,x_k)$'s involved are well defined. The only restriction is therefore that $C(\mathbf{t},\mathbf{t})$ is positive definite. Thus the major challenge is to specify a valid covariance function.

Note that positive *semi*-definite covariance matrices $\mathbf{\Sigma}$ are possible. For singular covariance matrices a slight change of the multinormal probability densities based on the characteristic function is necessary. See Mardia, Kent & Bibby (1979, pp. 41–44) for details. As the properties needed above are unchanged, Kolmogorov's existence theorem still applies.

1.5 Symmetries

In this section the class of possible covariance functions is restricted to simplify verification of positive definiteness. A symmetry (in the weak sense) is defined by invariance of ρ or C (and m) under some transformation; e.g. stationary correlation functions are invariant under translations and isotropic correlation functions are invariant under translations and rotations.

To define a symmetry the parameter space T must possess certain properties. For instance stationarity requires that T must be closed under addition and subtraction, and isotropy requires a properly defined distance function in T. So from heron the parameter space must be restricted to have certain properties.

1.5.1 Stationarity

Assume that T is a *linear space* (vector space) such that $\mathbf{t}, \mathbf{s} \in T$ implies that $\mathbf{t} + \mathbf{s} \in T$. The obvious example is $\mathbb{R}^n$.

Definition 1.4 (Stationarity in the strict sense) A random field is a *stationary random field in the strict sense* if all its finite-dimensional distributions are invariant under arbitrary translations, i.e.

$$F_{\mathbf{t}_1+\mathbf{s},\cdots,\mathbf{t}_k+\mathbf{s}}(x_1,\ldots,x_k) = F_{\mathbf{t}_1,\ldots,\mathbf{t}_k}(x_1,\ldots,x_k), \quad \text{for all } \mathbf{s} \in T. \tag{1.11}$$

Definition 1.5 (Stationarity in the wide sense) A random field is a *stationary random field in the wide sense* if

$$m(\mathbf{t}) = m \qquad \text{and} \qquad C(\mathbf{t}, \mathbf{s}) = C(\boldsymbol{\tau}), \tag{1.12}$$

where $\boldsymbol{\tau} = \mathbf{t} - \mathbf{s}$ is the separation vector. The corresponding covariance/correlation function is called a *stationary covariance/correlation function*.

Stationarity in the strict sense implies stationarity in the wide sense, whereas the opposite is not necessarily true. The two conditions are equivalent for Gaussian random fields so no distinction will be made in the following. Note that stationary random fields are called *homogeneous random fields* by some authors, e.g. Yaglom (1986*a*) and Vanmarcke (1983).

Any stationary covariance function must have constant variance so that

$$C(\boldsymbol{\tau}) = \sigma^2 \rho(\boldsymbol{\tau}).$$

This means that the attention can be focused on the correlation function or the covariance function without loosing generality. The equivalence of the two functions for stationary random fields has lead to the concept *correlation theory* used by some authors when referring to the study of the two first moments of stationary random fields (Yaglom 1986*a*).

1.5.2 Isotropy

A translation invariant correlation function is still a complex function so verifying positive definiteness is complicated even for these functions. This leads us to consider an even more restricted class of random fields possessing two more symmetries: Rotation and reflection invariance.

Assume that T is a metric space, that is, T is a linear space with a properly defined distance measure or norm usually denoted $d(\mathbf{t}, \mathbf{s})$. A typical example is $\mathbb{R}^n$ equipped with the Euclidean norm:

$$\tau = d(\mathbf{t}, \mathbf{s}) = \|\boldsymbol{\tau}\| = \sqrt{\tau_1^2 + \cdots + \tau_n^2}.$$

Definition 1.6 (Isotropic random field) A stationary random field is an *isotropic random field* (in the wide sense) if the covariance function depends on distance alone, i.e.

$$C(\mathbf{t}, \mathbf{s}) = C(\tau),$$

where $\tau = d(\mathbf{t}, \mathbf{s})$. The corresponding covariance/correlation function is called an *isotropic covariance/correlation function*.

As for stationary random fields no distinction between 'in the wide sense' and 'in the strict sense' will be made. The significance of isotropic correlation functions is the simplicity; it depends on a single variable, the distance between $\mathbf{t}$ and $\mathbf{s}$. When using a Euclidean norm in $\mathbb{R}^n$ the verification of positive definiteness is in principle simple. Verifying positive definiteness on general metric spaces however, is non-trivial. Schoenberg (1938*a*, 1938*b*) deals with this problem in some detail.

Isotropic functions are obviously a subclass of stationary functions.

1.5.3 Anisotropic Random Fields

The usual metric space considered is of course $\mathbb{R}^n$ with the Euclidean norm. In many applications isotropic correlation functions on this space are to restrictive. Instead of the Euclidean norm to evaluate distances, a more general norm of the form

$$\tau = \|\boldsymbol{\tau}\|_{\mathbf{K}} = \sqrt{\boldsymbol{\tau}^t \mathbf{K} \boldsymbol{\tau}} \tag{1.13}$$

is possible. The matrix $\mathbf{K}$ must be positive semi-definite (i.e. $|\mathbf{K}| \geq 0$) to ensure $\|\boldsymbol{\tau}\|_{\mathbf{K}} \geq 0$ for any $\boldsymbol{\tau}$.

Definition 1.7 (Anisotropic correlation function) A stationary correlation function on $\mathbb{R}^n$ is an *anisotropic correlation function* if its dependence on $\boldsymbol{\tau}$ is through a non-Euclidean norm of the form (1.13):

$$\rho(\boldsymbol{\tau}) = \rho(\|\boldsymbol{\tau}\|_{\mathbf{K}}).$$

Using the norm (1.13) gives ellipsoidal symmetry so correlation functions of this form are occasionally called *ellipsoidal correlation functions.*

Theorem 1.3 *The anisotropic correlation function $\rho(\boldsymbol{\tau}) = \rho(\|\boldsymbol{\tau}\|_{\mathbf{K}})$ is positive definite on $\mathbb{R}^n$ provided the isotropic correlation function $\rho(\tau)$ is positive definite on $\mathbb{R}^n$, and $\mathbf{K}$ is a symmetric positive semi-definite n-dimensional matrix.*

Proof. If $\mathbf{K}$ is symmetric and positive definite a unique square root $\mathbf{K}^{1/2}$ exists (Mardia et al. 1979, p. 471). Then

$$\begin{aligned}
\sum_{i,j=1}^{k} c_i c_j\, \rho\big(\|\mathbf{t}_i - \mathbf{t}_j\|^2_{\mathbf{K}}\big) &= \sum_{i,j=1}^{k} c_i c_j\, \rho\big((\mathbf{t}_i - \mathbf{t}_j)^t \mathbf{K} (\mathbf{t}_i - \mathbf{t}_j)\big) \\
&= \sum_{i,j=1}^{k} c_i c_j\, \rho\Big(\big((\mathbf{t}_i - \mathbf{t}_j)^t \mathbf{K}^{1/2}\big)\big(\mathbf{K}^{1/2}(\mathbf{t}_i - \mathbf{t}_j)\big)\Big) \\
&= \sum_{i,j=1}^{k} c_i c_j\, \rho\big((\mathbf{t}'_i - \mathbf{t}'_j)^t (\mathbf{t}'_i - \mathbf{t}'_j)\big) \\
&= \sum_{i,j=1}^{k} c_i c_j\, \rho\big(\|\mathbf{t}'_i - \mathbf{t}'_j\|^2\big) \geq 0.
\end{aligned}$$

This is valid for any choice of k, $\{c_1 \ldots, c_k\}$, and $\{\mathbf{t}_1, \ldots, \mathbf{t}_k\}$, so $\rho\big(\|\mathbf{t}_i - \mathbf{t}_j\|^2_{\mathbf{K}}\big)$ must be positive definite.

Now consider a positive semi-definite matrix $\mathbf{K}$. Then $\mathbf{K}^{1/2}$ is not uniquely defined but a generalized square root of $\mathbf{K}$ can be defined by taking the square root of the non-zero eigenvalues of the spectral decomposition. Using this generalized square root in the calculation makes no difference except that the null space of $\mathbf{K}$ will not contribute so that the transformed coordinates $\mathbf{t}'_i$ will be confined to a subset of $\mathbb{R}^n$. □

1.5.4 Additional Symmetries

Two more symmetries—both subclasses of stationary symmetry—will be considered.

Definition 1.8 (Separable correlation functions) The correlation function $\rho(\boldsymbol{\tau})$ on $\mathbb{R}^n$ is *fully separable* if

$$\rho(\boldsymbol{\tau}) = \rho_1(\tau_1) \cdots \rho_n(\tau_n).$$

The correlation function $\rho(\boldsymbol{\tau})$ on $\mathbb{R}^n$ is *partially separable* if

$$\rho(\boldsymbol{\tau}) = \rho_1(\boldsymbol{\tau}_1) \cdots \rho_m(\boldsymbol{\tau}_m), \quad \text{where } m < n,$$

where $\boldsymbol{\tau}^t = (\boldsymbol{\tau}_1^t, \ldots, \boldsymbol{\tau}_m^t)$, and ρ_i is a correlation function on $\mathbb{R}^{n_i}$ for $i = 1, \ldots, m$, and $n = \sum_{i=1}^m n_i$.

Typical examples are found in geology where the two lateral dimensions are separated from the vertical dimension: $\rho(x, y, z) = \rho(x, y)\rho(z)$, or in space-time contexts, e.g. meteorology, where time is separated: $\rho(x, y, z, t) = \rho(x, y, z)\rho(t)$.

Theorem 1.4 *A function $\rho(\boldsymbol{\tau}) = \rho_1(\boldsymbol{\tau}_1) \cdots \rho_m(\boldsymbol{\tau}_m)$ is a correlation function on $\mathbb{R}^n$ if and only if each ρ_i is a correlation function on $\mathbb{R}^{n_i}$, and $n = \sum_{i=1}^m n_i$.*

Proof. Consider the 'degenerate' function $\rho(\boldsymbol{\tau}) = \rho_l(\boldsymbol{\tau}_l)$ on $\mathbb{R}^n$. This function is positive definite on $\mathbb{R}^n$ since

$$\sum_{i,j=1}^{k} c_i c_j\, \rho(t_i - t_j) = \sum_{i,j=1}^{k} c_i c_j\, \rho_l(t_{li} - t_{lj}) \geq 0.$$

(Since ρ_l positive definite on $\mathbb{R}^{n_l}$.) It is obviously also a correlation function on $\mathbb{R}^n$. Thus, correlation functions $\rho_l(\boldsymbol{\tau}_l)$ on $\mathbb{R}^{n_l}$ are correlation functions on $\mathbb{R}^n$ when the redundant directions on $\mathbb{R}^n$ are ignored. Theorem 3.1 part (iii) states that a finite product of correlation functions on $\mathbb{R}^n$ are a correlation function on $\mathbb{R}^n$. $\square$

Definition 1.9 (Quadrant symmetry) A correlation function $\rho(\boldsymbol{\tau})$ on $\mathbb{R}^n$ is *quadrant symmetric* if

$$\rho(\tau_1, \ldots, \tau_i, \ldots, \tau_n) = \rho(\tau_1, \ldots, -\tau_i, \ldots, \tau_n) \quad \text{for any } i = 1, \ldots, n.$$

Quadrant symmetry does not imply isotropy but isotropic correlation functions are quadrant symmetric. More details on quadrant symmetry is found in Vanmarcke (1983, pp. 80–82).

1.5.5 Non-stationary Gaussian Random Fields

Stationarity and in particular isotropy impose strong restrictions on the possible definition of a Gaussian random field. However, from stationary random fields it is simple to obtain a wide range of non-stationary random fields.

Consider an isotropic Gaussian random field $Y_{\mathbf{t}}$ with properties

$$\mathrm{E}\{Y_{\mathbf{t}}\} = 0 \qquad \text{and} \qquad \mathrm{Cov}\{Y_{\mathbf{t}}, Y_{\mathbf{s}}\} = \rho(\tau),$$

where ρ is an isotropic correlation function. Next consider the random field

$$X_{\mathbf{t}} = \sigma(\mathbf{t})\, Y_{\mathbf{t}},$$

where $\sigma(\mathbf{t})$ is a real valued function. The expectation and covariance function are

$$\mathrm{E}\{X_{\mathbf{t}}\} = 0 \qquad \text{and} \qquad \mathrm{Cov}\{X_{\mathbf{t}}, X_{\mathbf{s}}\} = \sigma(\mathbf{t})\, \sigma(\mathbf{s})\, \rho(\tau),$$

i.e. the non-stationary variance is specified by $\sigma^2(\mathbf{t})$. Now consider the random field

$$X_{\mathbf{t}} = m(\mathbf{t}) + Y_{\mathbf{t}}$$

where $m(\mathbf{t})$ is a real valued function. The expectation and covariance function are

$$\mathrm{E}\{X_{\mathbf{t}}\} = m(\mathbf{t}) \qquad \text{and} \qquad \mathrm{Cov}\{X_{\mathbf{t}}, X_{\mathbf{s}}\} = \rho(\tau)$$

Here the expectation is non-stationary. This can be extended further. Consider the Gaussian random field

$$X_{\mathbf{t}} = \sum_p A_p\, f_p(\mathbf{t}) + Y_{\mathbf{t}},$$

where A_p are Gaussian random variables independent of $Y_{\mathbf{t}}$, and $f_p(\mathbf{t})$ are real valued functions. The expectation and covariance function are

$$\mathrm{E}\{X_{\mathbf{t}}\} = \sum_p \mathrm{E}\{A_p\}\, f_p(\mathbf{t})$$
$$\mathrm{Cov}\{X_{\mathbf{t}}, X_{\mathbf{s}}\} = \sum_{p,q} \mathrm{Cov}\{A_p, A_q\}\, f_p(\mathbf{t})\, f_q(\mathbf{s}) + \rho(\tau).$$

Both the expectation and the covariance function has become non-stationary.

Combinations of these simple extensions are obviously also possible, and moreover, the isotropic correlation function can be replaced by an anisotropic correlation function. The existence of $X_{\mathbf{t}}$ is ensured by Theorem 1.1, also implying the positive definiteness of the covariance function. The random fields are Gaussian random fields since linear combinations of Gaussian random variables are Gaussian random variables.

1.6 Isotropic Multidimensional Random Fields

Consider the *multidimensional random field*

$$\mathbf{X}_{\mathbf{t}} = \big(X_1(\mathbf{t}), \dots, X_m(\mathbf{t})\big); \quad \mathbf{t} \in \mathbb{R}^n .$$

An isotropic multidimensional random field is defined such that each component and all cross-covariance functions are isotropic:

$$\begin{aligned} \mathrm{E}\big\{X_i(\mathbf{t})\big\} &= m_i \\ \mathrm{Var}\big\{X_i(\mathbf{t})\big\} &= \sigma_i^2 \\ \mathrm{Cov}\big\{X_i(\mathbf{t}), X_j(\mathbf{s})\big\} &= C_{ij}(\tau). \end{aligned}$$

The covariance matrix C_{ij} must be a positive definite matrix for all distances. This is the almost trivial extension from ordinary random fields on $\mathbb{R}^n$. Yaglom (1986*a*, pp. 370–372) has more details, e.g. a spectral representation.

1.6.1 Isotropic Random Vector Fields

The definition of isotropy used for multidimensional fields will not do. The axes introduce artificial directions so that each component is not isotropic although the vector is isotropic. Here only some rudimentary results are given. For a thorough discussion on isotropic vector fields see Yaglom (1986*a*, pp. 372–383) or Monin & Yaglom (1975, pp. 35ff).

As opposed to a multidimensional random field a *random vector field* and the coordinates are both in $\mathbb{R}^n$:

$$\mathbf{X}_{\mathbf{t}} = \big(X_1(\mathbf{t}), \dots, X_n(\mathbf{t})\big) \in \mathbb{R}^n \quad \text{and} \quad \mathbf{t} \in \mathbb{R}^n .$$

Moreover, the transformations of the coordinates also applies to the components of the vector field itself.

A random vector field on $\mathbb{R}^n$ is isotropic if expectations and covariances are independent of translations, rotations, and reflections of the coordinates in $\mathbb{R}^n$. Consider an isotropic random vector field $\mathbf{X}_{\mathbf{t}}$. The expectation vector must be invariant under arbitrary rotations implying that $\mathrm{E}\{\mathbf{X}_{\mathbf{t}}\} = \mathbf{0}$. The elements of the covariance matrix

$$\mathrm{Cov}\big\{X_i(\mathbf{t}), X_j(\mathbf{s})\big\} = C_{ij}(\boldsymbol{\tau})$$

must form an *isotropic* second order tensor called the *covariance tensor*. It can be shown that the general form is

$$C_{ij}(\boldsymbol{\tau}) = \big(C_R(\tau) - C_T(\tau)\big)\frac{\tau_i\tau_j}{\tau^2} + C_T(\tau)\,\delta_{ij}, \tag{1.14}$$

where δ_{ij} is the Kronecker delta and C_T and C_R are two isotropic covariance functions. The properties of these are studied by Yaglom (1986*a*, pp. 380–383) and Monin & Yaglom (1975, pp. 35ff).

An immediate observation is that a single component of an isotropic vector field is *not* isotropic since for instance

$$C_{ii}(\boldsymbol{\tau}) = \big(C_R(\tau) - C_T(\tau)\big)\frac{\tau_i^2}{\tau^2} + C_T(\tau)$$

depends on the direction of the coordinate axes.

2 Geometrical Properties

Geometrical properties of random fields is a vast area including any geometrical property of functions in n-dimensional Euclidean spaces. There are several areas important for practical applications; *continuity* and *differentiability* of random fields, *stochastic integration*, and the properties of *excursion sets**. All subjects are found in Adler (1981). Cramér & Leadbetter (1967) also contains a discussion on these topics but it is mainly restricted to random fields on $\mathbb{R}^1$. An even more general treatment of continuity is found in Adler (1990). Excursion sets are also considered in some detail by Vanmarcke (1983).

Excursion sets are not discussed but continuity and differentiability are treated in some detail. Integration is also briefly mentioned. The discussion is slightly technical since properties such as continuity and differentiability of realizations fail to be determined by finite dimensional distributions so reference to the underlying probability space is needed. Therefore, in this section the less compact $X(\mathbf{t}, \omega)$ is used occasionally to stress the dependence on the underlying probability space. The symbol $X(\mathbf{t})$ for $X_{\mathbf{t}}$ is used throughout this section for notational reasons necessary when treating derivatives.

2.1 Continuity

Continuity of a function, f, is a property related to the convergence of a sequence $\{f(\mathbf{t}_n)\}$ when $\|\mathbf{t}_n - \mathbf{t}\| \to 0$ as $n \to \infty$. When studying a function being a random field, continuity is related to the convergence of sequences $\{X(\mathbf{t}_n)\}$ of random variables. As there are different types of convergence for random variables there are corresponding types of convergence for random fields. Three types of continuity will be considered based on *almost sure convergence* (i.e. convergence with probability one) and *mean square convergence*.

Definition 2.1 (Continuity of random fields) Consider a $B \subset \mathbb{R}^n$.

(i) A random field X has *continuous sample paths with probability one* in B if for every sequence $\{\mathbf{t}_n\}$ for which $\|\mathbf{t}_n - \mathbf{t}\| \to 0$ as $n \to \infty$, then

$$\text{Prob}\big\{\omega\colon |X(\mathbf{t}_n, \omega) - X(\mathbf{t}, \omega)| \to 0 \quad \text{as } n \to \infty \quad \text{for all } \mathbf{t} \in B\big\} = 1.$$

(ii) A random field X is *almost surely continuous* in B if for every sequence $\{\mathbf{t}_n\}$ for which $\|\mathbf{t}_n - \mathbf{t}\| \to 0$ as $n \to \infty$, then

$$\text{Prob}\big\{\omega\colon |X(\mathbf{t}_n, \omega) - X(\mathbf{t}, \omega)| \to 0 \quad \text{as } n \to \infty\big\} = 1 \quad \text{for all } \mathbf{t} \in B.$$

(iii) A random field X is *mean square continuous* in B if for every sequence $\{\mathbf{t}_n\}$ for which $\|\mathbf{t}_n - \mathbf{t}\| \to 0$ as $n \to \infty$, then

$$\mathrm{E}\big\{|X(\mathbf{t}_n) - X(\mathbf{t})|^2\big\} \to 0 \quad \text{as } n \to \infty \quad \text{for all } \mathbf{t} \in B.$$

Continuous sample paths with probability one means that there are, with probability one, no discontinuities within the whole domain B. Continuous sample paths with probability one are commonly referred to as *sample path continuity.* Almost sure continuity however, allows discontinuities within B. There is no contradiction here; although a realization has discontinuities, the probability

*An excursion set is the area, volume, etc. where a function has values above a certain pre-specified level.

for finding a discontinuity at a particular position, say $\mathbf{t}$, is zero. Evidently, sample path continuity is a far stronger property than almost sure continuity. An example of an almost surely continuous random field with discontinuous sample paths is given in Example 2.1. In general, mean square continuity is not implied by sample path continuity. Nor does mean square continuity imply sample path continuity. Sample path continuity is a far stronger condition relying on much more specific behavior of the covariance function. For Gaussian random fields however, mean square continuity is a necessary and almost sufficient condition for continuous sample paths. The meaning of 'almost sufficient' will become clear in the subsequent.

Finally note that *continuity in probability*:

$$\lim_{n\to\infty} \mathrm{Prob}\big\{\omega\colon |X(\mathbf{t}_n,\omega) - X(\mathbf{t},\omega)| > \delta\big\} = 0 \qquad \text{for any } \delta > 0$$

is implied by mean square continuity. This follows directly from Chebyshev's inequality:

$$\mathrm{Prob}\big\{\omega\colon |X(\mathbf{t}_n,\omega) - X(\mathbf{t},\omega)| > \delta\big\} \le \frac{\mathrm{E}\big\{|X(\mathbf{t}_n) - X(\mathbf{t})|^2\big\}}{\delta^2}.$$

Two random fields X and Y are said to be *equivalent* or *versions* of each other if

$$\mathrm{Prob}\big\{\omega\colon X(\mathbf{t},\omega) = Y(\mathbf{t},\omega)\big\} = 1 \qquad \text{for any } \mathbf{t}.$$

Two equivalent random fields generate identical finite dimensional distributions but they need not necessarily be identical as the following example shows.

Example 2.1 Consider two random fields $X(t,\omega)$ and $Y(t,\omega)$ where $t \in \mathbb{R}^1$ and $\omega \in [0,1] = \Omega$. That is, the probability space is the unit interval with a uniform distribution of points ω. The two random fields are defined as

$$X(t,\omega) = 0 \quad \text{for all } t \text{ and } \omega,$$
$$Y(t,\omega) = \begin{cases} 1 & \text{for } t = \omega \\ 0 & \text{otherwise.} \end{cases}$$

The finite dimensional distributions for both random fields are identical:

$$F_{t_1,\ldots,t_k}(x_1,\ldots,x_k) = \begin{cases} 1 & \text{for all } x_i \ge 0,\ i = 1,\ldots,k \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, the two random fields are equivalent:

$$\mathrm{Prob}\big\{\omega\colon X(t,\omega) = Y(t,\omega)\big\} = 1,$$

since the probability (measure) of a single point is zero: $\mathrm{Prob}\{\omega\} = 0$. Finally, for the same reason, both X and Y are almost surely continuous and mean square continuous. Nevertheless

$$\mathrm{Prob}\big\{\omega\colon X(t,\omega) \text{ is continuous in } [0,1]\big\} = 1$$
$$\mathrm{Prob}\big\{\omega\colon Y(t,\omega) \text{ is continuous in } [0,1]\big\} = 0.$$

That is, X has continuous sample paths with probability one whereas Y has not. Thus, the finite dimensional distributions do not determine the continuity properties of a random field even for this almost trivial example. Furthermore, surely continuity is not enough to determine continuity of sample paths. This example also shows that there can be two versions having the same finite dimensional distribution although their continuity properties are different. $\diamond$

2.1.1 Separable Random Fields

In practical applications continuity of sample paths are obviously of far more interest than the piecewise continuity implied by almost surely continuity. Therefore it is reasonable to impose additional conditions on the random fields guaranteeing continuous sample paths (with probability one). This is not simple. As a start the concept of *separability* introduced by Doob (1953, p. 53) is needed. A formal definition for random fields is found in Adler (1981, p. 14). Billingsley (1986, pp. 551 ff.) consider the concept in some detail for stochastic processes, i.e. random fields on $\mathbb{R}^1$.

Separability ensures that finite dimensional distributions determine sample path properties by requiring that sample paths are determined by their values on an everywhere dense but *countable* subset of positions in $\mathbb{R}^n$. The definition of separability excludes random fields such as Y above; a separable random field on $\mathbb{R}^1$ satisfies

$$\sup_{t \in D \cap B} X(t) = \sup_{t \in B} X(t), \tag{2.1}$$

where $B \subset \mathbb{R}^1$ and $D \subset \mathbb{R}^1$ and D is *countable.* Since ω in Example 2.1 can be any point in $[0,1]$, it does not necessarily belong to a countable subset of $\mathbb{R}^1$. Moreover, $\mathrm{Prob}\{\omega \in D\} = 0$ so (2.1) will be satisfied with probability zero for the random field Y.

Separability excludes the most pathological random fields, but it does still not guaranty continuous sample paths even though the random field is almost surely continuous. Nevertheless, the significance of separability is a consequence of the two following statements:

> *To any given random field Y, it is always possible to find an equivalent random field X which is separable.*
>
> *Properties of a separable random field is uniquely defined by its finite dimensional distributions.*

The first statement is from Doob (1953, p. 57) and the second is a mere consequence of the definition. Thus it is always possible to claim separability such that random fields behave reasonably regular between two arbitrary close positions. In the following all random fields are assumed to be separable.

2.1.2 Mean Square Continuity

Even though mean square continuity does not imply continuous sample paths, the opposite is almost true for Gaussian random fields: mean square continuity is implied by the sufficient conditions for continuous sample paths given in the next section. The importance of mean square continuity is owing to the following theorem. It makes a simple link between mean square continuity and the continuity of the covariance function.

Theorem 2.1 *Assume* $\mathrm{E}\{X(\mathbf{t})\}$ *is continuous. Then, a random field* $X(\mathbf{t})$ *is mean square continuous at* $\mathbf{t} \in \mathbb{R}^n$ *if and only if its covariance function* $C(\mathbf{s}, \mathbf{s}')$ *is continuous at* $\mathbf{s} = \mathbf{s}' = \mathbf{t}$.

If $C(\mathbf{s}, \mathbf{s}')$ *is continuous at every 'diagonal position',* $\mathbf{s} = \mathbf{s}'$, *then it is everywhere continuous.*

A proof of the first statement will be given. For a proof of the last statement see Adler (1981, p. 83).

Proof. If $\mathrm{E}\{X(\mathbf{t})\} \neq 0$ consider $\tilde{X}(\mathbf{t}) = X(\mathbf{t}) - m(\mathbf{t})$, i.e. $\mathrm{E}\{\tilde{X}(\mathbf{t})\} = 0$. The "if" part of the first statement follows from

$$\mathrm{E}\{|X(\mathbf{t}_n) - X(\mathbf{t})|^2\} = C(\mathbf{t}_n, \mathbf{t}_n) - 2C(\mathbf{t}_n, \mathbf{t}) + C(\mathbf{t}, \mathbf{t}).$$

If $C(\cdot, \cdot)$ is continuous, then the right hand side vanish as $n \to \infty$ and mean square continuity follows.

The "only if" part requires use of the Cauchy-Schwartz inequality. Assuming mean square continuity implies

$$0 = \lim_{n\to\infty} C(\mathbf{t}_n, \mathbf{t}_n) - 2 \lim_{n\to\infty} C(\mathbf{t}_n, \mathbf{t}) + C(\mathbf{t}, \mathbf{t}). \tag{2.2}$$

Squaring gives

$$4\Big(\lim_{n\to\infty} C(\mathbf{t}_n, \mathbf{t})\Big)^2 = \Big(\lim_{n\to\infty} C(\mathbf{t}_n, \mathbf{t}_n) + C(\mathbf{t}, \mathbf{t})\Big)^2.$$

Interchanging the limit and the square on the left hand side and applying the Cauchy-Schwartz inequality, $\mathrm{E}\{X(\mathbf{t})X(\mathbf{t}_n)\}^2 \leq \mathrm{E}\{X(\mathbf{t})^2\}\,\mathrm{E}\{X(\mathbf{t}_n)^2\}$, gives

$$4 \lim_{n\to\infty} C(\mathbf{t}_n, \mathbf{t}_n)\, C(\mathbf{t}, \mathbf{t}) \geq \Big(\lim_{n\to\infty} C(\mathbf{t}_n, \mathbf{t}_n) + C(\mathbf{t}, \mathbf{t})\Big)^2.$$

This inequality is not satisfied unless $\lim_{n\to\infty} C(\mathbf{t}_n, \mathbf{t}_n) = C(\mathbf{t}, \mathbf{t})$. (The parable touch the straight line at $C(\mathbf{t}, \mathbf{t})$). Moreover, (2.2) then implies that $\lim_{n\to\infty} C(\mathbf{t}, \mathbf{t}_n) = C(\mathbf{t}, \mathbf{t})$ so that the "only if" part holds. □

A direct consequence of the theorem is:

Corollary 2.1.1 *A stationary random field* $X(\mathbf{t})$ *is mean square continuous at any* $\mathbf{t} \in \mathbb{R}^n$ *if and only if its correlation function* $\rho(\boldsymbol{\tau})$ *is continuous at* $\mathbf{0}$.

Since separable Gaussian random fields are uniquely specified by expectations and covariances, intuition suggests that mean square continuity must be 'almost' sufficient to guarantee continuous sample paths. Theorem 2.3 below gives the answer: an additional weak restriction on the behavior of $C(\mathbf{t}, \mathbf{s})$ for small τ is necessary.

2.1.3 Sufficient Conditions for Continuous Sample Paths

In this section a few theorems are stated without proof. Most of the proofs are found in Adler (1981, pp. 41–49, pp. 59–62). All random fields are assumed to be separable.

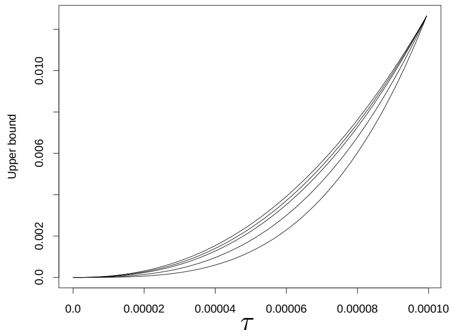

Figure 2.1: The function $c\tau^2/|\log\tau|^{3+\epsilon}$ plotted for different values of ϵ (0, 1, 2, 5, and 10). The value of c is chosen such that the curves coincidence at $\tau = 0.0001$. The upper curve has $\epsilon = 0$ whereas the bottom curve has $\epsilon = 10$.

Theorem 2.2 *Let $X(\mathbf{t})$ be a random field on $\mathbb{R}^n$. Then if, for some $c > 0$, some $\alpha > 0$, and some $\eta > \alpha$,*

$$\mathrm{E}\big\{|X(\mathbf{t}) - X(\mathbf{s})|^\alpha\big\} \le \frac{c\tau^{2n}}{|\log\tau|^{1+\eta}},$$

then the random field $X(\mathbf{t})$ will have continuous sample paths with probability one over any compact subset in $\mathbb{R}^n$.

This result is a corollary to a general theorem by Belyaev (1972). See Adler (1981, pp. 47–49) for a derivation of the result above.

For Gaussian random fields this result can be sharpened. The following theorem is from Adler (1981, p. 60).

Theorem 2.3 *Let $X(\mathbf{t})$ be a zero-mean, Gaussian random field with a continuous covariance function. Then if, for some finite $c > 0$ and some $\epsilon > 0$,*

$$\mathrm{E}\big\{|X(\mathbf{t}) - X(\mathbf{s})|^2\big\} \le \frac{c}{|\log\tau|^{1+\epsilon}}, \tag{2.3}$$

for all $\boldsymbol{\tau}$ with $\tau < 1$, then the random field $X(\mathbf{t})$ will have continuous sample paths with probability one.

For stationary Gaussian random fields condition (2.3) simplifies:

Corollary 2.3.1 *Let $X(\mathbf{t})$ be a stationary Gaussian random field with a continuous correlation function. Then if, for some finite $c > 0$ and some $\epsilon > 0$,*

$$1 - \rho(\boldsymbol{\tau}) \le \frac{c}{|\log\tau|^{1+\epsilon}} \tag{2.4}$$

for all $\boldsymbol{\tau}$ with $\tau < 1$, then the random field $X(\mathbf{t})$ will have continuous sample paths with probability one.

These conditions deserve a few comments. The bounds given by Theorem 2.2 are extremely restricting. Consider for instance a stationary random field on $\mathbb{R}^1$ with known correlation function. Choosing $\alpha = 2$ gives the inequality [compare to (2.4)]

$$1 - \rho(\boldsymbol{\tau}) \le \frac{c\tau^2}{|\log\tau|^{3+\epsilon}}, \tag{2.5}$$

where $\epsilon > 0$. Figure 2.1 shows the behavior of $\tau^2/|\log\tau|^{3+\epsilon}$ for different values of ϵ. The figure shows that the correlation function must approach 0 with zero

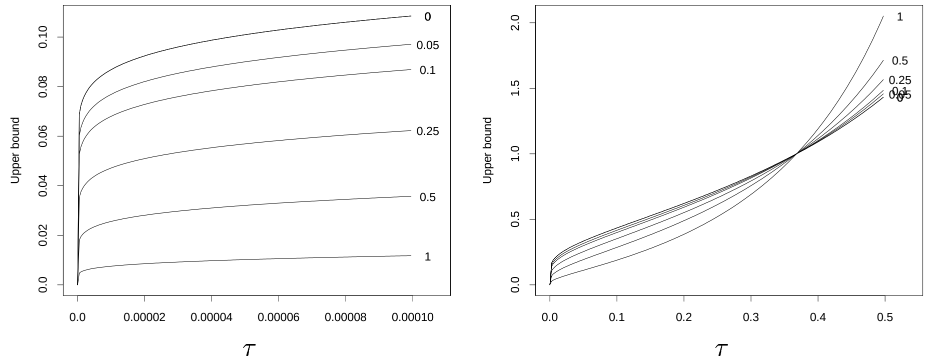

Figure 2.2: The function $|\log \tau|^{-(1+\epsilon)}$ plotted for different values of ϵ (0, 0.05, 0.1, 0.25, 0.5, and 1). The left figure is a magnification of the lower left corner of the right figure.

slope. In higher dimensions the condition becomes even more restrictive since the power of τ in the numerator increases. It is of course possible to choose $\alpha \neq 2$ and hope for less restrictive bounds. In practice however, this can be complicated since the evaluation of the corresponding expectations is difficult. Also note that Theorem 2.2 gives a sufficient but not a necessary condition. It will soon be realized that this condition is abundantly restrictive for Gaussian random fields.

Now consider condition (2.4). The bound is an everywhere increasing function of τ starting at zero at $\tau = 0$ and approaching infinity at $\tau = 1$. Note also that $\tau < 1$ could be replaced by $\tau < \delta$ where δ is some non-zero (but arbitrarily small) constant less than one. This is obvious since a change of scale has no influence on continuity properties. Thus the conditions really impose a restriction on how slow the convergence towards continuity of the covariance function can be. Figure 2.2 shows the behavior of $|\log \tau|^{-(1+\epsilon)}$ for different values of ϵ. Note that the vertical scale is arbitrary since any positive and finite c can be chosen. The figures suggest that the slope of the curves are infinitely steep at $\tau = 0$. A calculation shows that this observation holds: $\mathrm{d}|\log \tau|^{-(1+\epsilon)}/\mathrm{d}\tau = 1/(\tau|\log \tau|^{2+\epsilon})$ which approach* ∞ as $\tau \to 0$. Therefore, it is almost impossible to find a continuous correlation function violating the bound. Thus, I boldly state:

Conjecture 2.1 *Gaussian random fields with continuous expectations and continuous covariance functions possess continuous sample paths with probability one.*

The limiting value $\epsilon = 0$ has the most extreme behavior. Figure 2.3 illustrates this. Thus the most including bound is obtained by using ϵ as small as possible, i.e. infinitesimal. Alternatively, consider $\epsilon = 0$ and require a strict inequality for $\tau > 0$. It is also evident from the figures that covariance functions having a finite derivative at $\tau = 0$ can always be bounded by a proper choice of c.

It is possible to verify the inequality (2.4) analytically. An example illustrates this.

*$\lim_{\tau \to 0} \tau^\alpha \log \tau = 0$ for any $\alpha > 0$ (Abramowitz & Stegun 1972, Eq. 4.1.31). Moreover, $1/(\tau|\log \tau|^{2+\epsilon}) = 1/(\tau^{1/(2+\epsilon)}|\log \tau|)^{2+\epsilon}$. Since $1/(2+\epsilon) > 0$ then $\tau^{1/(2+\epsilon)}|\log \tau| \to 0$ so that $1/(\tau|\log \tau|^{2+\epsilon}) \to \infty$ when $\tau \to 0$.

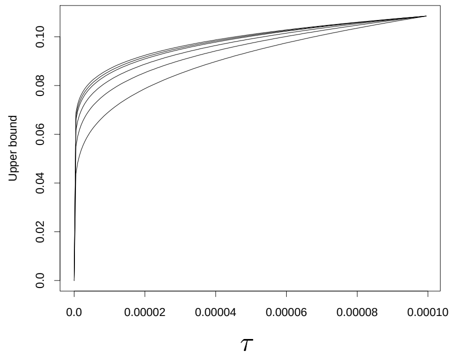

Figure 2.3: The function $c|\log \tau|^{-(1+\epsilon)}$ plotted for different values of ϵ (0, 0.05, 0.1, 0.25, 0.5, and 1). The value of c is chosen such that the curves coincidence at $\tau = 0.0001$. The upper curve has $\epsilon = 0$ whereas the bottom curve has $\epsilon = 1$.

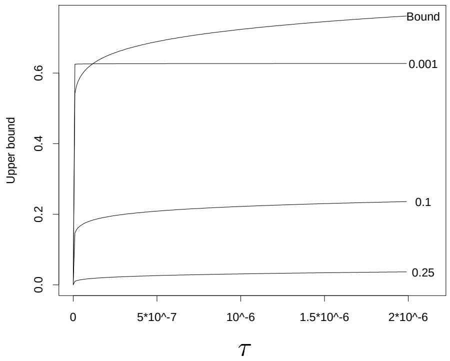

Figure 2.4: The exponential correlation function is plotted as $1-\exp(-\tau^\nu)$ for different values of ν (0.001, 0.1, 0.25). The 'Bound' is $10|\log \tau|^{-1}$. The bound is obviously not satisfied for $\nu = 0.001$.

Example 2.2 Consider the isotropic exponential correlation function (4.5):

$$\rho(\tau) = \exp(-\tau^\nu); \qquad 0 < \nu \leq 2.$$

The corresponding sample paths vary from smooth (analytical at $\nu = 2$) to irregular ($\nu \leq 1$). To ensure continuous sample paths it must be verified that

$$1 - \exp(-\tau^\nu) \leq c|\log \tau|^{-1-\epsilon} \tag{2.6}$$

for some positive c and ϵ. First note that both sides vanish at $\tau = 0$. Thus the crucial point is the 'speed of departure' from $\tau = 0$. The derivative of the left hand side is $\nu\tau^{\nu-1}\exp(-\tau^\nu)$ which is ∞ at $\tau = 0$ when $\nu < 1$. For $\nu \geq 1$ the derivative is finite for any τ, so (2.6) can be satisfied for any ϵ simply by choosing a sufficiently large c. Thus the complicated part is for $\nu < 1$. From Figure 2.4 it is seen that a particular bound can be violated by choosing a sufficiently small ν. Note in particular that the limit $\nu \to 0$ gives a special form of 'white noise':

$$\lim_{\nu\to 0} \rho(\tau) = \begin{cases} 1 & \text{for } \tau = 0 \\ 1/e \approx 0.37 & \text{otherwise.} \end{cases}$$

The corresponding random field is not continuous at any position. Thus to pin down the exact value of ν for which continuous sample paths exist needs more than a quick look at plots. Rewrite inequality (2.6) as

$$\bigl(1 - \exp(-\tau^\nu)\bigr)|\log \tau|^{1+\epsilon} \leq c.$$

For $\tau > 0$ the left hand side is obviously finite since $\log \tau$ and $\exp(-\tau^\nu)$ is finite. For $\tau = 0$ this is also true. Consider

$$\begin{aligned}
\bigl(1 - \exp(-\tau^\nu)\bigr)|\log \tau|^{1+\epsilon} &= \left(-\tau^\nu + \frac{1}{2}\tau^{2\nu} - \frac{1}{2\cdot 3}\tau^{3\nu} + \cdots\right)|\log \tau|^{1+\epsilon} \\
&= \left(-1 + \frac{1}{2}\tau^{\nu} - \frac{1}{2\cdot 3}\tau^{2\nu} + \cdots\right)\tau^\nu|\log \tau|^{1+\epsilon} \\
&= \left(-1 + \frac{1}{2}\tau^{\nu} - \frac{1}{2\cdot 3}\tau^{2\nu} + \cdots\right)\left(\tau^{\frac{\nu}{1+\epsilon}}|\log \tau|\right)^{1+\epsilon} \leq c.
\end{aligned}$$

Since $\lim_{\tau\to 0} \tau^{\frac{\nu}{1+\epsilon}} |\log \tau| = 0$ for any $\frac{\nu}{1+\epsilon} > 0$ (Abramowitz & Stegun 1972, Eq. 4.1.31) the left hand side is finite for any $\tau \geq 0$ provided $\nu > 0$. Thus it is always possible to find a (finite) c satisfying the inequality for all finite $\tau > 0$. So sample paths are continuous with probability one for isotropic Gaussian random field with the exponential correlation functions for all $0 < \nu \leq 2$. $\diamond$

The example shows that sample paths are continuous even though the correlation functions are similar to a white noise correlation function. This strengthens the conjecture that Gaussian random fields with continuous covariance function will possess continuous sample paths.

2.2 Derivatives

Consider a Gaussian random field $X(\mathbf{t})$ on $\mathbb{R}^n$. Assume that X has differentiable sample paths. Then the associated gradient field, $\dot{\mathbf{X}}(\mathbf{t})$, is a *space-vector* in $\mathbb{R}^n$ defined by its *components* in a Cartesian coordinate system*:

$$(2.7) \qquad \dot{X}_i(\mathbf{t}, \omega) = \frac{\partial X(\mathbf{t})}{\partial t_i} = \lim_{\Delta\to 0} \frac{X(\mathbf{t} - \Delta\mathbf{e}_i, \omega) - X(\mathbf{t}, \omega)}{\Delta},$$

where ω is kept fixed and $\mathbf{e}_i$ is a unit vector in the ith direction. A gradient field is commonly called a *potential vector field* and is characterized by zero curl. The components of $\dot{\mathbf{X}}$ are also Gaussian random fields since the differential operator is linear. The moments of $\dot{\mathbf{X}}$ is related to the corresponding moments of X. Assume $\mathrm{E}\{X(\mathbf{t})\} = m(\mathbf{t})$, then

$$\mathrm{E}\{\dot{X}_i(\mathbf{t})\} = \frac{\partial m(\mathbf{t})}{\partial t_i} = \dot{m}_i(\mathbf{t}).$$

Further, assume that the covariance function of $X(\mathbf{t})$ is given:

$$C(\mathbf{t}, \mathbf{s}) = \mathrm{Cov}\{X(\mathbf{t}), X(\mathbf{s})\},$$

and assume that $C(\mathbf{t}, \mathbf{s})$ is simultaneously differentiable in $\mathbf{t}$ and $\mathbf{s}$, that is, X is mean square differentiable. Then, the cross-covariance function between X and a component of $\dot{\mathbf{X}}$ is a space-vector defined by the components:

$$(2.8) \qquad \dot{C}_i(\mathbf{t}, \mathbf{s}) = \mathrm{Cov}\{X(\mathbf{t}), \dot{X}_i(\mathbf{s})\} = \frac{\partial}{\partial s_i} C(\mathbf{t}, \mathbf{s}); \qquad i = 1, \ldots, n.$$

The covariance functions and cross-covariance functions between components of $\dot{\mathbf{X}}$ are

$$(2.9) \qquad \ddot{C}_{ij}(\mathbf{t}, \mathbf{s}) = \mathrm{Cov}\{\dot{X}_i(\mathbf{t}), \dot{X}_j(\mathbf{s})\} = \frac{\partial^2}{\partial t_i \partial s_j} C(\mathbf{t}, \mathbf{s}); \quad i, j = 1, \ldots, n.$$

The matrix $\ddot{\mathbf{C}}$ is the components of a second order tensor and is called the *covariance tensor* of the gradient field. To see that (2.8) and (2.9) are correct simply use the definition (2.7) and use that limits and integration commute. The details can be found in Christakos (1992, pp. 43–46).

*The widely used convention that, e. g. $\dot{\mathbf{m}}$, means a gradient vector whereas, e.g. $\ddot{\mathbf{C}}$, means a tensor of second order derivatives is adopted.

2.2.1 Higher Order Derivatives

Consider the general derivative of X (assuming sufficient regularity conditions are satisfied):

$$\mathbf{X}^{(\kappa)}(\mathbf{t}) = \frac{\partial^{|\kappa|}}{\partial t_1^{\kappa_1} \dots \partial t_d^{\kappa_d}} X(\mathbf{t}),$$

where $\kappa = (\kappa_1, \dots, \kappa_n)$ are n non-negative integers and $|\kappa| = \sum_i \kappa_i$. The cross-covariances form a higher order tensor (Christakos 1992, pp. 44-45):

$$\begin{aligned}\mathbf{C}^{(\kappa,\lambda)}(\mathbf{t},\mathbf{s}) &= \mathrm{Cov}\{X^{(\kappa)}(\mathbf{t}), X^{(\lambda)}(\mathbf{s})\} \\ &= \frac{\partial^{|\kappa|+|\lambda|}}{\partial t_1^{\kappa_1} \cdots \partial t_n^{\kappa_n} \, \partial s_1^{\lambda_1} \cdots \partial s_n^{\lambda_n}} C(\mathbf{t},\mathbf{s}).\end{aligned}$$

By restricting $|\kappa| = k$ and $|\lambda| = l$ the components of $C^{(\kappa,\lambda)}$ form a $k+l$ order tensor. For stationary covariance functions, i. e. $C(\mathbf{t},\mathbf{s}) = C(\boldsymbol{\tau})$, the tensor must satisfy

$$\mathbf{C}^{(\kappa,\lambda)}(\mathbf{t},\mathbf{s}) = (-1)^{|\lambda|}\mathbf{C}^{(\kappa+\lambda)}(\boldsymbol{\tau}).$$

In the case $k = l = 2$ the covariance tensor is a fourth-order tensor containing n^4 components, e.g. 81 components when $n = 3$. Imposing isotropy reduces the degrees of freedom to three. Thus, exploiting symmetries is crucial for reducing complexity.

2.3 Differentiability

As for continuity there are different forms of differentiability based on different forms of convergence.

Definition 2.2 (Differentiability of random fields) Consider a $B \subset \mathbb{R}^n$.

(i) A random field X has *differentiable sample paths with probability one* in B if for every sequence $\{\mathbf{t}_n\}$ for which $||\mathbf{t}_n - \mathbf{t}|| \to 0$ as $n \to \infty$, then

$$\begin{aligned}\mathrm{Prob}\{\omega\colon |\dot{X}_i(\mathbf{t}_n,\omega) - \dot{X}_i(\mathbf{t},\omega)| \to 0 \quad &\text{as } n \to \infty \\ &\forall\, i = 1,\dots,n,\ \forall\, \mathbf{t} \in B\} = 1.\end{aligned}$$

(ii) A random field X is *almost surely differentiable* in B if for every sequence $\{\mathbf{t}_n\}$ for which $||\mathbf{t}_n - \mathbf{t}|| \to 0$ as $n \to \infty$, then

$$\mathrm{Prob}\{\omega\colon |\dot{X}_i(\mathbf{t}_n,\omega) - \dot{X}_i(\mathbf{t},\omega)| \to 0 \quad \text{as } n \to \infty \quad \forall\, i = 1,\dots,n\} = 1 \quad \forall\, \mathbf{t} \in B.$$

(iii) A random field X is *mean square differentiable* in B if for every sequence $\{\mathbf{t}_n\}$ for which $||\mathbf{t}_n - \mathbf{t}|| \to 0$ as $n \to \infty$, then

$$\mathrm{E}\{|\dot{X}_i(\mathbf{t}_n) - \dot{X}_i(\mathbf{t})|^2\} \to 0 \quad \text{as } n \to \infty \quad \forall\, i = 1,\dots,n,\ \forall\, \mathbf{t} \in B.$$

The discussion following Definition 2.1 of continuity of random fields applies to differentiability simply by replacing continuity by differentiability.

2.3.1 Mean Square Differentiability

As for mean square continuity, the significance of mean square differentiability is that it is a necessary condition for differentiable sample paths and it has a simple relation to the covariance function.

Theorem 2.4 *Consider a random field $X(\mathbf{t})$ on $\mathbb{R}^n$ with covariance function C and differentiable expectation. If the derivative $\partial^2 C(\mathbf{s},\mathbf{t})/\partial s_i\,\partial t_i$ exists and is finite for all $i = 1,\dots,n$ at the point $(\mathbf{t},\mathbf{t})$, then $X(\mathbf{t})$ is mean square differentiable at $\mathbf{t}$. The covariance function of $\dot{X}_i(\mathbf{t})$ is then given by $\partial^2 C(\mathbf{s},\mathbf{t})/\partial s_i\,\partial t_i$.*

A proof is given by Cramér & Leadbetter (1967, p. 84). For stationary random fields the theorem simplifies.

Corollary 2.4.1 *Consider a stationary random field $X(\mathbf{t})$ on $\mathbb{R}^n$ with covariance function C. If the derivative $\partial^2 C(\boldsymbol{\tau})/\partial\tau_i^2$ exists and is finite for all $i = 1,\dots,n$ at the point $\mathbf{0}$ then $X(\mathbf{t})$ is mean square differentiable at any $\mathbf{t}$. The covariance function of $\dot{X}_i(\mathbf{t})$ is then given by $-\partial^2 C(\boldsymbol{\tau})/\partial\tau_i^2$.*

The negative sign comes from $\partial C(\mathbf{t}-\mathbf{s})/\partial s_i = -\partial C(\boldsymbol{\tau})/\partial\tau_i$. For a stationary random field the covariance function attains the maximum value at $\mathbf{0}$. Therefore it is necessary that $\partial C(\boldsymbol{\tau})/\partial\tau_i = 0$ for all $i = 1,\dots,n$ for the second order derivatives to exist. For isotropic covariance functions this simplifies even further since all partial derivatives are equal; it is enough to consider $\mathrm{d}C(\tau)/\mathrm{d}\tau^2$.

It is of course possible to continue further in this direction by considering higher order derivatives. By induction (see Section 2.2.1 for notation):

Corollary 2.4.2 *Consider a random field $X(\mathbf{t})$ on $\mathbb{R}^n$ with covariance function C and expectation possessing the necessary derivatives. If the derivative*

$$\frac{\partial^{2|\kappa|}}{\partial t_1^{\kappa_1}\cdots\partial t_n^{\kappa_n}\,\partial s_1^{\kappa_1}\cdots\partial s_n^{\kappa_n}} C(\mathbf{t},\mathbf{s}) \tag{2.10}$$

exists and is finite for all $i = 1,\dots,n$ at the point $(\mathbf{t},\mathbf{t})$ then $X(\mathbf{t})$ is $|\kappa|$ times mean square differentiable at $\mathbf{t}$. The covariance function of

$$\frac{\partial^{|\kappa|} X(\mathbf{t})}{\partial t_1^{\kappa_1}\cdots\partial t_n^{\kappa_n}}$$

is then given by (2.10).

This means that the smoothness of the random field is related to the smoothness of the covariance function.

Example 2.3 Consider once more the isotropic exponential correlation function studied in Example 2.2: $\rho(\tau) = \exp(-\tau^\nu)$ with $0 < \nu \le 2$. The first derivative is

$$\dot{\rho}(\tau) = \frac{\mathrm{d}\rho(\tau)}{\mathrm{d}\tau} = -\nu\tau^{\nu-1}\exp(-\tau^\nu)$$

so that

$$\dot{\rho}(0) = \begin{cases} -\infty & \text{for } 0 < \nu < 1 \\ -1 & \text{for } \nu = 1 \\ 0 & \text{for } 1 < \nu \le 2. \end{cases} \tag{2.11}$$

Thus, there is a potential for differentiable sample paths for $1 < \nu \leq 2$. Also recognize that $\nu = 1$ is special in the sense that the transition from the potentially differentiable domain to the erratic domain is not continuous. The second derivative is

$$\ddot{\rho}(\tau) = \frac{\mathrm{d}^2\rho(\tau)}{\mathrm{d}\tau^2} = \nu\tau^{\nu-2}(1-\nu+\nu\tau^{\nu})\exp(-\tau^{\nu})$$

Since $(1-\nu+\nu\tau^{\nu})\exp(-\tau^{\nu}) \to 1-\nu$ as $t \to 0$ for any $\nu > 1$

$$\lim_{\tau\to 0}\ddot{\rho}(0) = \begin{cases} -\infty & \text{for } 1 < \nu < 2 \\ -2 & \text{for } \nu = 2. \end{cases}$$

This means that the only possibility for differentiable sample paths are $\nu = 2$. For $\nu = 2$ the sample paths are ∞ times mean square differentiable. To verify this it must be shown that all odd numbered derivatives of ρ are zero and all even numbered derivatives are finite. Consider the series expansion for ρ when $\nu = 2$:

$$\rho(\tau) = 1 - \tau^2 + \frac{\tau^4}{1\cdot 2} - \frac{\tau^6}{1\cdot 2\cdot 3} + \frac{\tau^8}{1\cdot 2\cdot 3\cdot 4} + \cdots = \sum_{k=0}^{\infty}(-1)^k\frac{\tau^{2k}}{k!}.$$

It is obvious that odd numbered derivatives will always result in a series of the form $c_1\tau + c_2\tau^3 + c_3\tau^5 + \cdots$ which is zero for $\tau = 0$. Similarly the even numbered (say $2m$) derivatives will always result in a series of the form $c_0 + c_1\tau^2 + c_2\tau^4 + \cdots$ where $c_0 = -\frac{2m!}{m!}$. Thus ρ can be differentiated any number of times. ◇

2.3.2 Sufficient Conditions for Differentiable Sample Paths

Differentiable sample paths means that the partial derivatives of the sample paths are continuous. Thus, applying Theorem 2.2 or 2.3 to the gradient field or correlation functions of the gradient fields gives sufficient conditions. Generalization to higher order derivatives is straight forward; simply consider the higher order partial derivatives of X and C.

2.4 Stochastic Integration

A full account of stochastic integration will not be given; only some preliminary results on the two first moments of a spatially averaged random field are given. For more details on this fundamental subject see e.g. Doob (1953), Cramér & Leadbetter (1967), Adler (1981), Christakos (1992), or Priestley (1992).

The integrals considered here are spatial averages of random fields. Consider a Gaussian random field X on $\mathbb{R}^n$ with an everywhere continuous correlation function*. Then a new 'average' Gaussian random field is defined by the *Riemann integral*

$$Y(\mathbf{t}) = \int_B X(\mathbf{s})\, w(\mathbf{t},\mathbf{s})\,\mathrm{d}\mathbf{s}, \tag{2.12}$$

*Correlation functions with a discontinuity at the origin are not considered since the more general Riemann-Stieltjes integrals would be needed.

where $B \subset \mathbb{R}^n$. The weight function $w(\mathbf{t}, \mathbf{s})$ on $\mathbb{R}^n \otimes \mathbb{R}^n$ is assumed to be piecewise continuous and bounded. The integral $Y(\mathbf{t})$ is defined as the limit of the sequence of random variables $\{Y^i(\mathbf{t})\}$ formed by a refinement of the partition of B leading to the Riemann integral. The existence of $Y(\mathbf{t})$ is ensured provided the sequence $\{Y^i(\mathbf{t})\}$ converges in mean square.

The expectation of Y is

$$\mathrm{E}\{Y(\mathbf{t})\} = \int_B m(\mathbf{s})\, w(\mathbf{t}, \mathbf{s})\, \mathrm{d}\mathbf{s},$$

whereas the covariance is

$$\operatorname{Cov}\{Y(\mathbf{t}), Y(\mathbf{s})\} = \int_B \int_B C(\mathbf{v}, \mathbf{u})\, w(\mathbf{t}, \mathbf{u})\, w(\mathbf{s}, \mathbf{v})\, \mathrm{d}\mathbf{v}\, \mathrm{d}\mathbf{u}. \tag{2.13}$$

These results are obtained directly from the definition (2.12) by interchanging the sequence of integration (Fubini). If $\mathrm{E}\{X(\mathbf{t})\} = 0$, then the requirement of mean square convergence amounts to (Christakos 1992, p. 48)

$$\mathrm{E}\{Y(\mathbf{t})^2\} = \int_B \int_B C(\mathbf{v}, \mathbf{u})\, w(\mathbf{t}, \mathbf{u})\, w(\mathbf{t}, \mathbf{v})\, \mathrm{d}\mathbf{v}\, \mathrm{d}\mathbf{u} \; < \infty.$$

Thus, if the covariance given by (2.13) exists, then $Y(\mathbf{t})$ is properly defined.

Finally assume that the weight function $w(\mathbf{t}, \mathbf{s})$ is everywhere differentiable in $\mathbf{t}$. Then $Y(\mathbf{t})$ will also be differentiable since the components of the gradient field are defined as

$$\dot{Y}_i(\mathbf{t}) = \int_B X(\mathbf{s})\, \frac{\partial w(\mathbf{t}, \mathbf{s})}{\partial t_i}\, \mathrm{d}\mathbf{s}.$$

Moreover, the expectation and covariances of the components of the gradient field are

$$\mathrm{E}\{\dot{Y}_i(\mathbf{t})\} = \int_B m(\mathbf{s})\, \frac{\partial w(\mathbf{t}, \mathbf{s})}{\partial t_i}\, \mathrm{d}\mathbf{s},$$

and

$$\operatorname{Cov}\{\dot{Y}_i(\mathbf{t}), \dot{Y}_j(\mathbf{s})\} = \int_B \int_B C(\mathbf{v}, \mathbf{u})\, \frac{\partial w(\mathbf{t}, \mathbf{u})}{\partial t_i} \frac{\partial w(\mathbf{s}, \mathbf{v})}{\partial s_j}\, \mathrm{d}\mathbf{v}\, \mathrm{d}\mathbf{u}.$$

In this fashion it is possible to obtain arbitrarily smooth Gaussian random fields by using a smooth weight function.

3 Correlation Functions

3.1 General Properties

The two basic properties of a correlation function are its positive definiteness and that $\rho(\mathbf{t}, \mathbf{t}) = 1$. These properties implies that correlation functions are *bounded* by 1*:

$$|\rho(\mathbf{t}, \mathbf{s})| \leq 1. \tag{3.1}$$

This inequality will be made stronger for isotropic correlation functions when $n > 1$ in Section 3.3.3. It also follows directly from the definition of the correlation function (1.6) that it is *symmetric*:

$$\rho(\mathbf{t}, \mathbf{s}) = \rho(\mathbf{s}, \mathbf{t}).$$

Denote by $\mathfrak{B}_n$ the class of correlation functions on $\mathbb{R}^n$. According to Corollary 1.2.1, this class is equivalent to the class of positive definite functions on $\mathbb{R}^n$ where $\rho(\mathbf{t}, \mathbf{t}) = 1$. The following theorem says that $\mathfrak{B}_n$ is closed under addition, multiplication, limits, and integration.

Theorem 3.1 *For* $\mathbf{t}, \mathbf{s} \in \mathbb{R}^n$.

(i) *If* $\rho_1(\mathbf{t}, \mathbf{s}), \rho_2(\mathbf{t}, \mathbf{s}) \in \mathfrak{B}_n$, $a_1, a_2 \geq 0$, *and* $a_1 + a_2 = 1$, *then* $a_1\rho_1(\mathbf{t}, \mathbf{s}) + a_2\rho_2(\mathbf{t}, \mathbf{s}) \in \mathfrak{B}_n$.

(ii) *If* $\rho_i(\mathbf{t}, \mathbf{s}) \in \mathfrak{B}_n$ *and* $a_i \geq 0$ *for* $i = 1, \ldots,$ *and* $\sum_i a_i = 1$, *then* $\sum_i a_i\rho_i(\mathbf{t}, \mathbf{s}) \in \mathfrak{B}_n$.

(iii) *If* $\rho_1(\mathbf{t}, \mathbf{s}), \rho_2(\mathbf{t}, \mathbf{s}) \in \mathfrak{B}_n$ *then* $\rho_1(\mathbf{t}, \mathbf{s})\, \rho_2(\mathbf{t}, \mathbf{s}) \in \mathfrak{B}_n$.

(iv) *If* $\rho_i(\mathbf{t}, \mathbf{s}) \in \mathfrak{B}_n$ *for* $i = 1, \ldots,$ *and* $\rho(\mathbf{t}, \mathbf{s}) = \lim_{i\to\infty} \rho_i(\mathbf{t}, \mathbf{s})$ *exist for all pairs* $\mathbf{t}, \mathbf{s}$, *then* $\rho(\mathbf{t}, \mathbf{s}) \in \mathfrak{B}_n$.

(v) *If* $\rho(\mathbf{t}, \mathbf{s}; a) \in \mathfrak{B}_n$ *for all* $a \in A$, *and* μ *is a measure such that* $\mu(A) = 1$, *then*

$$\rho(\mathbf{t}, \mathbf{s}) = \int_A \rho(\mathbf{t}, \mathbf{s}; a)\, \mathrm{d}\mu(a) \ \in \mathfrak{B}_n.$$

Proof. (i) A positive sum of two positive definite functions must be positive definite. Thus $a_1\rho_1(\mathbf{t}, \mathbf{s}) + a_2\rho_2(\mathbf{t}, \mathbf{s})$ is a covariance function. Since $a_1\rho_1(\mathbf{t}, \mathbf{t}) + a_2\rho_2(\mathbf{t}, \mathbf{t}) = a_1 + a_2 = 1$, the sum is a correlation function.

(ii) Induction from (i).

(iii) Assume that $\mathrm{Corr}\{X_{\mathbf{t}}, X_{\mathbf{s}}\} = \rho_1(\mathbf{t}, \mathbf{s})$, and $\mathrm{Corr}\{Y_{\mathbf{t}}, Y_{\mathbf{s}}\} = \rho_2(\mathbf{t}, \mathbf{s})$, and that $X_{\mathbf{t}}$ and $Y_{\mathbf{t}}$ are independent random fields with expectation zero. The random field $Z_{\mathbf{t}} = X_{\mathbf{t}}Y_{\mathbf{t}}$ is well defined according to Corollary 1.1.1(ii). Independence gives $\mathrm{Corr}\{Z_{\mathbf{t}}, Z_{\mathbf{s}}\} = \mathrm{Corr}\{X_{\mathbf{t}}, X_{\mathbf{s}}\}\, \mathrm{Corr}\{Y_{\mathbf{t}}, Y_{\mathbf{s}}\} = \rho_1(\mathbf{t}, \mathbf{s})\, \rho_2(\mathbf{t}, \mathbf{s})$.

(iv) If $\rho_i(\mathbf{t}, \mathbf{s})$ is positive definite for all i then certainly the limit must be positive definite. The same applies to the condition $\rho_i(\mathbf{t}, \mathbf{t}) = 1$.

(v) This is a consequence of (ii) and the definition of the integral as a finite sum over decompositions of A (Billingsley 1986, p. 203). Since each contribution to the integral must be positive definite the integral must be positive definite according to (ii). Finally, $\rho(\mathbf{t}, \mathbf{t}) = \int_A \rho(\mathbf{t}, \mathbf{t}; a)\, \mathrm{d}\mu(a) = \int_A 1\, \mathrm{d}\mu(a) = \mu(A) = 1$ so that $\rho(\mathbf{t}, \mathbf{s})$ is a correlation function. □

By virtue of Theorem 3.1(i) and (ii), sums of correlation functions give valid models. In geostatistical jargon this is usually referred to as *nested structures.*

*Using $k = 2$ and $c_1 = \pm c_2$ in (1.7) gives $1 \pm \rho(\mathbf{t}, \mathbf{s}) \geq 0$ which is equivalent.

3.2 Stationary Correlation Functions

3.2.1 Continuity

Denote by $\mathfrak{C}_n$ the class of *stationary* correlation functions on $\mathbb{R}^n$. Further, denote by $\mathfrak{C}'_n$ stationary correlation functions being continuous everywhere except possibly at $\mathbf{0}$. Finally denote by $\mathfrak{C}''_n$ stationary correlation functions being continuous everywhere. Thus by definition

$$\mathfrak{C}''_n \subset \mathfrak{C}'_n \subset \mathfrak{C}_n \subset \mathfrak{B}_n.$$

Theorem 3.2 *If $\rho \in \mathfrak{C}'_n$ then*

$$\rho(\boldsymbol{\tau}) = a\rho_{\mathrm{w}}(\boldsymbol{\tau}) + b\tilde{\rho}(\boldsymbol{\tau}) \tag{3.2}$$

where $\rho_{\mathrm{w}} \in \mathfrak{C}'_n$ is the white noise correlation function given by (4.1)*, $\tilde{\rho} \in \mathfrak{C}''_n$, $a, b \geq 0$, and $a + b = 1$.*

A proof is given in Matérn (1960, p. 12). This suggests that a random field $X_{\mathbf{t}}$ on $\mathbb{R}^n$ can be decomposed into a completely chaotic part and a continuous part:

$$X_{\mathbf{t}} = X^{\mathrm{w}}_{\mathbf{t}} + \tilde{X}_{\mathbf{t}}, \tag{3.3}$$

where $\mathrm{Corr}\{X^{\mathrm{w}}_{\mathbf{t}}, X^{\mathrm{w}}_{\mathbf{s}}\} = \rho_{\mathrm{w}}(\boldsymbol{\tau})$ and $\mathrm{Corr}\{\tilde{X}_{\mathbf{t}}, \tilde{X}_{\mathbf{s}}\} \in \mathfrak{C}''_n$, and $X^{\mathrm{w}}_{\mathbf{t}}$ and $\tilde{X}_{\mathbf{t}}$ are independent. Non-continuous correlation functions are seldom reasonable models for natural phenomena. The nugget effect (Journel & Huijbregts 1978, p. 39) considered in geostatistics is hardly an exception. Thus the focus will be on continuous correlation functions, that is $\rho \in \mathfrak{C}''_n$. The lack of generality is minimal since the decompositions (3.2) and (3.3) means that any $\rho \in \mathfrak{C}'_n$ can be obtained from a $\rho \in \mathfrak{C}''_n$ by adding a pure 'white noise' component.

Theorem 3.3 *If $\rho(\boldsymbol{\tau})$ is continuous at $\mathbf{0}$, then $\rho(\boldsymbol{\tau})$ is continuous everywhere, i.e. $\rho \in \mathfrak{C}''$.*

Proof. Consider $U = X_{\mathbf{t}+\boldsymbol{\varepsilon}} - X_{\mathbf{t}}$ and $V = X_{\mathbf{t}-\mathbf{s}}$. Using the Cauchy-Schwartz inequality $(\mathrm{E}\{UV\}^2 \leq \mathrm{E}\{U^2\}\,\mathrm{E}\{V^2\})$ gives $\big(\rho(\boldsymbol{\tau} - \boldsymbol{\varepsilon}) - \rho(\boldsymbol{\tau})\big)^2 \leq 2\big(\rho(\mathbf{0}) - \rho(\boldsymbol{\varepsilon})\big)\rho(\mathbf{0})$. Hence, if $\lim_{\boldsymbol{\varepsilon}\to\mathbf{0}} \rho(\boldsymbol{\varepsilon}) = \rho(\mathbf{0})$ then $\lim_{\boldsymbol{\varepsilon}\to\mathbf{0}} \rho(\boldsymbol{\tau} + \boldsymbol{\varepsilon}) = \rho(\boldsymbol{\tau})$. □

3.2.2 Spectral Representation

Multidimensional Bochner's Theorem *A real function $r(\boldsymbol{\tau})$ on $\mathbb{R}^n$ is positive (semi-)definite if and only if it can be represented in the form*

$$r(\boldsymbol{\tau}) = \int_{\mathbb{R}^n} e^{i\boldsymbol{\tau}\mathbf{k}}\, \mathrm{d}^n F(\mathbf{k}), \tag{3.4}$$

where $F(\cdot)$ is a non-negative bounded measure.

For a proof see for instance Bochner (1959). An even more general proof having Bochner's theorem as a special case is given byRudin (1973, p. 285).

Bochner's theorem says that all positive definite functions have a unique *spectral representation*. The integral (3.4) is the n-dimensional *Fourier transform* of the non-negative F. Such integrals are sometimes called *Fourier-Stieltjes* integrals.

According to Theorem 1.2 the class of positive definite functions coincidence with the class of covariance functions. Thus, the following theorem follows almost immediately from Bochner's theorem:

Wiener-Khintchine's Theorem *A real function $\rho(\boldsymbol{\tau})$ on $\mathbb{R}^n$ is a correlation function if and only if it can be represented in the form*

$$\rho(\boldsymbol{\tau}) = \int_{\mathbb{R}^n} e^{i\boldsymbol{\tau}\mathbf{k}}\,\mathrm{d}^n F(\mathbf{k}), \tag{3.5}$$

where the function $F(\mathbf{k})$ on $\mathbb{R}^n$ has the properties of a n-dimensional distribution function.

Proof. Bochner's theorem says that $\rho(\boldsymbol{\tau})$ is a covariance function if and only if F is bounded and non-negative. Also $\rho(\mathbf{0}) = \int_{\mathbb{R}^n} \mathrm{d}^n F(\mathbf{k}) = F(\infty)$. Thus F is a n-dimensional distribution function if and only if $\rho(\boldsymbol{\tau})$ is a correlation function. □

The n-dimensional distribution function is called the spectral distribution function.

A useful observation linking correlation functions to *characteristic functions* is (Matérn 1960, p. 12):

Theorem 3.4 *A correlation function in $\mathfrak{C}_n$ is the characteristic function of some n-dimensional random variable $\mathbf{X}$. Conversely, the characteristic function of any n-dimensional random variable is a correlation function in $\mathfrak{C}_n$.*

This is obvious since F has the properties of a distribution function so that (3.5) can be written as

$$\rho(\boldsymbol{\tau}) = \mathrm{E}\{e^{i\boldsymbol{\tau}\mathbf{X}}\}.$$

Characteristic functions for many multidimensional distributions are known, so these provide a useful source of valid correlation functions. Thus the Wiener-Khintchine theorem is useful for obtaining valid correlation functions.

When F is continuous, the *spectral density function* exists and is defined as

$$f(\mathbf{k}) = \frac{\partial^n F(\mathbf{k})}{\partial k_1 \cdots \partial k_n}.$$

Then the Fourier-Stieltjes integral in (3.5) becomes

$$\rho(\boldsymbol{\tau}) = \int_{\mathbb{R}^n} e^{i\boldsymbol{\tau}\mathbf{k}}\, f(\mathbf{k})\,\mathrm{d}^n\mathbf{k}. \tag{3.6}$$

Corollary 3.4.1 *A function $f(\mathbf{k})$ on $\mathbb{R}^n$ is the spectral density function of a stationary correlation function on $\mathbb{R}^n$ if and only if $f(\mathbf{k}) \geq 0$ and $\int_{\mathbb{R}^n} f(\mathbf{k})\,\mathrm{d}^n\mathbf{k} = 1$, i.e. $f(\mathbf{k})$ has the properties of a n-dimensional probability density.*

Proof. F is a distribution function if and only if $f(\mathbf{k}) \geq 0$ and $\int_{\mathbb{R}^n} f(\mathbf{k})\,\mathrm{d}^n\mathbf{k} = 1$. □

The spectral density function is obtained from the correlation function by the usual formula for the inversion of an n-dimensional Fourier transform:

$$f(\mathbf{k}) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-i\boldsymbol{\tau}\mathbf{k}}\, \rho(\boldsymbol{\tau})\, \mathrm{d}^n\boldsymbol{\tau}. \tag{3.7}$$

This gives us our first explicit method for verifying the positive definiteness of a (stationary) correlation function on $\mathbb{R}^n$:

> *Evaluate the spectral density,* $f(\mathbf{k})$*, given by* (3.7)*, and check if it is non-negative for any* $\mathbf{k} \in \mathbb{R}^n$ *!*

This is not necessarily a simple procedure; the calculation of n-dimensional Fourier integrals can be tremendously complicated. Even checking that f is non-negative can be difficult. However, for correlation functions possessing additional symmetries the integration simplifies. For instance the spectral density of separable correlation functions are simple to evaluate since each direction is integrated independently. Here are two examples of separable correlation functions and their spectral density functions:

Example 3.1 Consider the separable correlation function

$$\rho(\boldsymbol{\tau}) = a_1^{|\tau_1|} \cdots a_n^{|\tau_n|}; \qquad |a_i| < 1 \text{ for } i = 1, \dots, n.$$

The n-dimensional spectral density can be calculated independently for each direction giving

$$f(\mathbf{k}) = (2\pi)^{-n} \frac{(1-a_1^2)\cdots(1-a_n^2)}{|e^{ik_1} - a_1|^2 \cdots |e^{ik_n} - a_n|^2} \geq 0 \quad \text{for all } \mathbf{k}.$$

Thus, this correlation function is permissible and belongs to $\mathfrak{C}_n''$. ◇

Example 3.2 The stationary exponential correlation function

$$\rho(\boldsymbol{\tau}) = \exp\left(-a_1|\tau_1| - \cdots - a_n|\tau_n|\right); \qquad a_i > 0 \text{ for } i = 1, \dots, n.$$

is also separable and the n-dimensional spectral density is

$$f(\mathbf{k}) = \pi^{-n} \frac{a_1 \cdots a_n}{(k_1^2 + a_1^2)\cdots(k_n^2 + a_n^2)} \geq 0 \quad \text{for all } \mathbf{k}.$$

Once again the spectral density is everywhere non-negative and the correlation function belongs to $\mathfrak{C}_n''$. ◇

3.3 Isotropic Correlation Functions

Isotropic correlation functions on $\mathbb{R}^n$ satisfy: $\rho(\boldsymbol{\tau}) = \rho(\tau)$. It will soon be realized that the requirement of isotropy is restrictive leading to conditions on continuity and constraints on lower bounds. These restrictions become stronger as the dimension increase.

Denote by $\mathfrak{D}_n$ the class of *isotropic* correlation functions on $\mathbb{R}^n$. Since a correlation function valid on $\mathbb{R}^n$ must be valid in any $(n-1)$-dimensional subset of $\mathbb{R}^n$:

$$\mathfrak{D}_1 \supset \mathfrak{D}_2 \supset \mathfrak{D}_3 \supset \cdots \supset \mathfrak{D}_\infty. \tag{3.8}$$

Further, denote by $\mathfrak{D}_n'$ the subclass of isotropic correlation functions that are continuous everywhere except possibly at 0. Finally denote by $\mathfrak{D}_n''$ the subclass of correlation functions being continuous everywhere. Thus by definition

$$\mathfrak{D}_n'' \subset \mathfrak{D}_n' \subset \mathfrak{D}_n.$$

Since isotropic correlation functions are a subclass of stationary correlation functions $\mathfrak{D}_n'' \subset \mathfrak{C}_n''$, $\mathfrak{D}_n' \subset \mathfrak{C}_n'$, and $\mathfrak{D}_n \subset \mathfrak{C}_n$. Thus, according to Theorem 3.2 any $\rho \in \mathfrak{D}_n'$ can be decomposed into a white noise correlation function and an everywhere continuous correlation function.

Schoenberg's Conjecture *The class $\mathfrak{D}_n' - \mathfrak{D}_n$ is empty for all $n > 1$.*

This means that the only discontinuity in $\rho \in \mathfrak{D}_n$ must be at 0. This conjecture from Schoenberg (1938*a*, pp. 822–823) is found in Matérn (1960, pp. 13–14) which gives some arguments strengthening the assumption of correctness. No proof of this statement can been found in the literature—at least by this author—despite its apparent simplicity. The conjecture implies that *any* isotropic correlation function can be decomposed into a white noise correlation function and an everywhere continuous correlation function.

Finally denote by $\mathfrak{D}_{n0}''$ the class of $\rho \in \mathfrak{D}_n''$ which vanish at infinity. Several examples of such isotropic correlation functions will be given in Section 4.2. If Schoenberg's conjecture holds, the following conjecture also holds:

Conjecture 3.1 *For any $\rho \in \mathfrak{D}_n$ and any $n > 1$, then*

$$\rho(\tau) = a + b\rho_{\mathrm{w}}(\tau) + c\tilde{\rho}(\tau),$$

where $a, b, c \geq 0$, $a + b + c = 1$, ρ_{w} is a white noise correlation function, and $\tilde{\rho} \in \mathfrak{D}_{n0}''$.

This is a consequence of Schoenberg's conjecture, Theorem 3.1.(i), Theorem 3.2, and Theorem 3.7 which says that any $\rho \in \mathfrak{D}_n$ for $n > 1$ must approach a limit as τ approach infinity. The significance of this is that any isotropic correlation function can be constructed from correlation functions in $\mathfrak{D}_{n0}''$, i.e. continuous correlation functions approaching zero at infinity.

A preliminary result from Christakos (1984) also found in (Christakos 1992, pp. 73–74) gives a sufficient but not necessary condition for the permissibility of a correlation function on $\mathbb{R}^1$, $\mathbb{R}^2$ or $\mathbb{R}^3$.

Theorem 3.5 *A continuous real function $\rho(\tau)$ is a correlation function if the following conditions hold ($\dot{\rho}$, $\ddot{\rho}$, and $\dddot{\rho}$ are first, second and third order derivatives respectively):*

(i) *The derivative at zero is negative: $\dot{\rho}(0) < 0$.*
(ii) *The long range behavior satisfies [see* (3.20) *below]*

$$\lim_{\tau \to \infty} \frac{\rho(\tau)}{\tau^{(1-n)/2}} = 0.$$

(iii) *One of the following inequalities hold for all $\tau > 0$:*

$$\ddot{\rho}(\tau) \geq 0; \qquad \textit{for } \rho \in \mathfrak{D}_1$$

$$\int_\tau^u u\left(u^2 - \tau^2\right)^{-1/2} \mathrm{d}\ddot{\rho}(u) \geq 0; \qquad \textit{for } \rho \in \mathfrak{D}_2$$

$$\ddot{\rho}(\tau) - \tau\dddot{\rho}(\tau) \geq 0; \qquad \textit{for } \rho \in \mathfrak{D}_3.$$

A proof is found in Christakos (1984, p. 256). This result is useful in practical applications since it imposes restrictions directly on easily accessible derivatives rather than conditions on the spectral representation. On the other hand, it fails to include many useful correlation functions such as those possessing differentiable sample paths.

3.3.1 Spectral Representation

For isotropic correlation functions the Wiener-Khintchine theorem takes a simpler form where the n-dimensional Fourier integral is replaced by a one-dimensional *Bessel transform:*

Theorem 3.6 *A real function $\rho(\tau)$ on $\mathbb{R}^n$ is a correlation function if and only if it can be represented in the form*

$$\rho(\tau) = 2^{(n-2)/2}\Gamma(n/2)\int_0^\infty \frac{J_{(n-2)/2}(k\tau)}{(k\tau)^{(n-2)/2}}\,\mathrm{d}\Phi(k), \tag{3.9}$$

where the function $\Phi(k)$ on $\mathbb{R}$ has the properties of a distribution function and J are Bessel functions of the 1. kind.

This result is obtained by eliminating angular dependencies in (3.5) by using spherical polar coordinates and integrating over all angles. The calculation is mainly straight forward manipulations with known tabulated integrals. For the details consult Yaglom (1986*a*, pp. 349–353) who treats the cases $n = 2, 3$ and general n separately (general n obviously includes the former). For a more compact and rigorous treatment see Adler (1981, pp. 36–37).

The isotropic spectral distribution function Φ is related to the spectral distribution function F by

$$\Phi(k) = \int_{\|\mathbf{k}\|<k} \mathrm{d}F(\mathbf{k}). \tag{3.10}$$

As many integrals of the form (3.9) are tabulated, it is much simpler to verify positive definiteness using this relation rather than the general n-dimensional Fourier representation.

A few special cases of (3.9) are of particular interest:

$$\rho(\tau) = \int_0^\infty \cos k\tau\,\mathrm{d}\Phi(k) \qquad \text{for } \rho \in \mathfrak{D}_1, \tag{3.11a}$$

$$\rho(\tau) = \int_0^\infty J_0(k\tau)\,\mathrm{d}\Phi(k) \qquad \text{for } \rho \in \mathfrak{D}_2, \tag{3.11b}$$

$$\rho(\tau) = \int_0^\infty \frac{\sin k\tau}{k\tau}\,\mathrm{d}\Phi(k) \qquad \text{for } \rho \in \mathfrak{D}_3, \tag{3.11c}$$

$$\rho(\tau) = \int_0^\infty \exp(-k^2\tau^2)\,\mathrm{d}\Phi(k) \qquad \text{for } \rho \in \mathfrak{D}_\infty. \tag{3.11d}$$

All except the last relation are simple to obtain. A proof of the last expression is given in Schoenberg (1938*a*, pp. 817ff).

It is common and convenient to denote the kernel in (3.9) by a special symbol [see Yaglom (1986*a*, p. 355) or Matérn (1960, p. 14)*]:

$$\Lambda_n(x) = 2^{(n-2)/2}\Gamma(n/2)\frac{J_{(n-2)/2}(x)}{x^{(n-2)/2}},$$

*Matérn defines Λ_k where $k = (n-2)/2$.

such that (3.9) reads

$$\rho(\tau) = \int_0^\infty \Lambda_n(k\tau)\,\mathrm{d}\Phi(k). \tag{3.12}$$

It follows that $\Lambda_n(0) = 1$ for any n since $\rho(0) = 1$.

As for stationary correlation functions an inversion formula for the spectral density can be obtained. Assuming

$$\int_0^\infty \tau^{n-1}|\rho(\tau)|\,\mathrm{d}\tau < \infty$$

the spectral density $f(\mathbf{k})$ given by (3.7) exists. Once again passing to spherical polar coordinates and integrating over all angular coordinates results in the following inversion formula for the *n-dimensional isotropic spectral density*:

$$f(k) = \frac{1}{(2\pi)^{n/2}} \int_0^\infty \frac{J_{(n-2)/2}(k\tau)}{(k\tau)^{(n-2)/2}} \tau^{n-1} \rho(\tau)\,\mathrm{d}\tau. \tag{3.13}$$

This gives an explicit method for verifying the positive definiteness of an isotropic correlation function on $\mathbb{R}^n$:

> *Evaluate the n-dimensional isotropic spectral density, $f(k)$, given by* (3.13)*, and check if it is positive everywhere!*

Be aware that $f(k)$ depends on n; a non-negative spectral density on $\mathbb{R}^n$ does not necessarily imply a non-negative spectral density on $\mathbb{R}^{n+1}$. The opposite however, is true, according to (3.8).

The correlation function is given by [compare (3.9)]:

$$\rho(\tau) = (2\pi)^{n/2} \int_0^\infty \frac{J_{(n-2)/2}(k\tau)}{(k\tau)^{(n-2)/2}} k^{n-1} f(k)\,\mathrm{d}k. \tag{3.14}$$

In particular, (3.11a–c) take the forms

$$\begin{aligned}
2\rho(\tau) &= \int_0^\infty \cos k\tau\, f(k)\,\mathrm{d}k && \text{for } \rho \in \mathfrak{D}_1,\\
\rho(\tau) &= \int_0^\infty J_0(k\tau)\, f(k)\, k\,\mathrm{d}k && \text{for } \rho \in \mathfrak{D}_2,\\
\rho(\tau) &= \int_0^\infty \frac{\sin k\tau}{k\tau} f(k)\, k^2\,\mathrm{d}k && \text{for } \rho \in \mathfrak{D}_3.
\end{aligned}$$

Since (3.14) must be a particular form of (3.9), $\Phi(k)$ is related to $f(k)$ by:

$$\Phi(k) = \Sigma_n \int_0^k {k'}^{n-1} f(k')\,\mathrm{d}k',$$

where $\Sigma_n = 2\pi^{n/2}/\Gamma(n/2)$ is the area of an n-dimensional unit sphere (Gradshteyn & Ryzhik 1980, Eq. 4.633). Moreover, provided a spectral density function exists, the identity $x^n J_{n-1}(x) = \mathrm{d}\big(x^n J_n(x)\big)/\mathrm{d}x$ (Abramowitz & Stegun 1972, Eq. 9.1.30) gives

$$\Phi(k) = \frac{1}{2^{(n-2)/2}\Gamma(n/2)} \int_0^\infty J_{n/2}(k\tau)(k\tau)^{n/2} \frac{\rho(\tau)}{\tau}\,\mathrm{d}\tau.$$

It is possible to show that this expression is valid for any isotropic correlation function—even with nonexistent spectral density function—if the half-sum of Φ at points of discontinuity is used.

Here follow some examples of spectral densities. Spectral densities are given up to a proportionality constant since the shape is the most interesting. The cited references give the complete formulas.

Example 3.3 The isotropic exponential (and Gaussian) correlation functions (see Section 4.2.3)

$$\rho(\tau) = e^{-a\tau}; \quad a > 0$$
$$\rho(\tau) = e^{-a\tau^2}; \quad a > 0$$

have the n-dimensional isotropic spectral density functions (Yaglom 1986a, pp. 362, 364)

$$f(k) \propto \frac{1}{(a^2 + k^2)^{(n+1)/2}} > 0 \quad \text{for all } k > 0 \text{ and all } n \tag{3.15}$$
$$f(k) \propto e^{-k^2/4a} > 0 \quad \text{for all } k > 0$$

respectively. Thus, these correlation functions belongs to $\mathfrak{D}''_\infty$. Note that the latter does not depend on n.

Spectral densities for the Gaussian correlation function and the exponential correlation function is plotted in Figure 3.1. $\diamond$

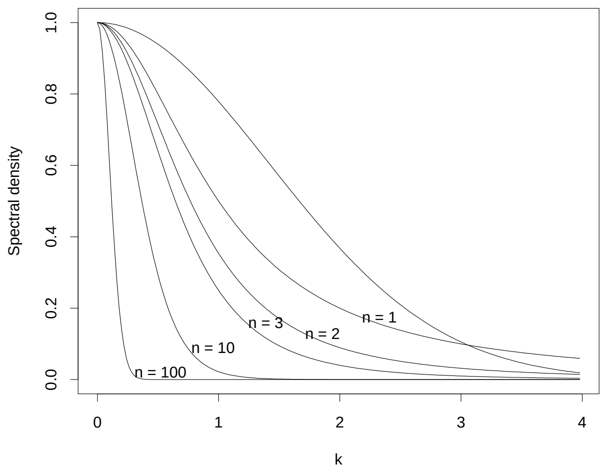

Figure 3.1: The n-dimensional isotropic spectral density functions for the Gaussian correlation function (upper curve) and the exponential correlation function for $n = 1, 2, 3, 10$, and 100. They are all scaled such that $f(0) = 1$ using $a = 1$.

Example 3.4 The modified Bessel correlation function is (see Section 4.2.5)

$$\rho(\tau) \propto (a\tau)^\nu K_\nu(a\tau); \quad a > 0, \nu > 0.$$

The corresponding n-dimensional isotropic spectral density function is (Yaglom 1986a, pp. 363)

$$f(k) \propto \frac{1}{(a^2 + k^2)^{\nu+n/2}} > 0 \quad \text{for all } k > 0 \text{ and all } n,$$

so that the correlation function belongs to $\mathfrak{D}''_\infty$. Note that choosing $\nu = 1/2$ gives (3.15). ◇

Example 3.5 The rational quadratic correlation function (see Section 4.2.4)

$$\rho(\tau) \propto \frac{1}{(a^2+\tau^2)^\nu}$$

has the n-dimensional isotropic spectral density function (Matérn 1960, pp. 17)

$$f(k) \propto (ak)^{\nu-n/2} K_{\nu-n/2}(ak) > 0 \quad \text{for all } k > 0 \text{ and all } n.$$

So once again we find that the correlation function belongs to $\mathfrak{D}''_\infty$.

Note the relationship between rational quadratic and modified Bessel correlation functions and spectral densities. ◇

There is also a Laplace transform representation for correlation functions in $\mathfrak{D}''_\infty$. Consider the isotropic exponential correlation function: $\rho(\tau) = e^{-at} \in \mathfrak{D}''_\infty$ for all finite $a > 0$. Then, according to Theorem 3.1(v)

$$\rho(\tau) = \int_0^\infty e^{-a\tau}\,\mathrm{d}F(a) \quad \in \mathfrak{D}_\infty. \tag{3.16}$$

Thus the *Laplace transform* of any probability measure on $\mathbb{R}^+$, is a valid isotropic correlation function in $\mathfrak{D}_\infty$. (Note that a contribution at $a = 0$ gives an everywhere constant correlation whereas a contribution at $a = \infty$ gives white noise.) In other words, for such probability measures, the *moment generating function* (Billingsley 1986, p. 285) defined as

$$\rho(\tau) = \mathrm{E}\{e^{-\tau X}\},$$

is a correlation function in $\mathfrak{D}_\infty$ for non-negative X.

3.3.2 Turning Bands Representation

The *turning bands method* is a simulation algorithm based on a particular representation suggested by Matheron (1973, pp. 461–462). Consider a random field X on $\mathbb{R}^1$ with an isotropic correlation function $\rho_1 \in \mathfrak{D}_1$. Then consider a random field Y on $\mathbb{R}^n$ defined such that

$$Y_{\mathbf{t}} = X(\mathbf{e}\cdot\mathbf{t}),$$

where $\mathbf{e}$ is a unit random vector in $\mathbb{R}^n$ uniformly distributed on the unit sphere. This construction can be used for simulation of random fields. The correlation function of Y becomes

$$\rho(\boldsymbol{\tau}) = \Sigma_n^{-1} \int_{\|\mathbf{e}\|=1} \rho_1(\mathbf{e}\cdot\mathbf{t})\,\mathrm{d}\sigma(\mathbf{e}); \qquad \rho \in \mathfrak{D}_n,$$

where $\Sigma_n = 2\pi^{n/2}/\Gamma(n/2)$ is the area of the unit sphere in $\mathbb{R}^n$. According to Matheron (1973), this integral can be written as the *turning bands representation*:

$$\rho(\boldsymbol{\tau}) = \frac{2\Gamma(n/2)\pi^{-1/2}}{\Gamma(n/2-1/2)} \int_0^1 \rho_1(v\tau)\,(1-v^2)^{(n-3)/2}\,\mathrm{d}v. \tag{3.17}$$

This representation is related to the spectral representation; it is possible to show that (3.17) is equivalent to (3.9):

$$\rho(\tau) = 2^{(n-2)/2}\Gamma(n/2) \int_0^\infty \frac{J_{(n-2)/2}(k\tau)}{(k\tau)^{(n-2)/2}} \, \mathrm{d}\Phi_1(k) \tag{3.18}$$

where the measure is given by (3.11a):

$$\rho_1(\tau) = \int_0^\infty \cos k\tau \, \mathrm{d}\Phi_1(k).$$

According to Theorem 3.6 the representation (3.18) is the general form of an isotropic correlation function so that the mapping from ρ_1 to ρ must be one-to-one. In particular for $n = 3$, representation (3.17) gives $\rho(\tau) = \int_0^1 \rho_1(v\tau)\,\mathrm{d}v$, or equivalently, $\rho_1(\tau) = \mathrm{d}\big(\tau\rho(\tau)\big)/\mathrm{d}\tau$. Similar relations between the one-dimensional spectral function and the n-dimensional spectral function can be found in Yaglom (1986*a*, pp. 359–360).

An example of the use of the turning bands representation is taken from hydrology (Mantoglou & Wilson 1982, Christakos 1984):

Example 3.6 Consider the correlation function $\rho_1(\tau) = \big(1 - \frac{\tau}{R}\big)\exp\big(-\frac{\tau}{R}\big)$. For $n = 2$ the following correlation function is obtained:

$$\rho(\tau) = \text{const} \cdot \Big(I_0(\tau/R) - L_0(\tau/R) + \frac{\tau}{R}\big(I_1(\tau/R) - L_{-1}(\tau/R)\big)\Big),$$

where I_n are modified Bessel functions and L_n are modified Struve functions (Abramowitz & Stegun 1972, pp. 375ff, 498). ◇

3.3.3 Bounds and Limits

According to (3.10), $\Phi(0) = 0$. However, $\lim_{k\to 0} \Phi(k) = \Phi(0+) = a$ can be non-zero. Since $\Lambda_n(0) = 1$

$$\rho(\tau) = a + \int_0^\infty \Lambda_n(k\tau)\,\mathrm{d}\Phi_1(k), \tag{3.19}$$

where Φ_1 is continuous at 0 and $\Phi_1(\infty) = \int_0^\infty \mathrm{d}\Phi_1(k) = 1 - a$.

Theorem 3.7 *For $\rho \in \mathfrak{D}_n$ and any $n > 1$ the following limit exists*

$$\lim_{\tau\to\infty} \rho(\tau) = a,$$

where $a = \lim_{k\to 0} \Phi(k)$.

Proof. It must be shown that the integral in (3.19) vanishes as $\tau \to \infty$. Assume $n > 1$. It is well known that for some c_n, $|J_n(x)| < c_n x^{-1/2}$ for all x [see (3.21) below]. Thus, $|\Lambda_n(x)| < c_n x^{-(n-1)/2}$ for all x such that $\lim_{x\to\infty} \Lambda_n(x) = 0$. Moreover, since $|\Lambda_n(x)| \le 1$ for all x, the following limit is well defined (bounded convergence theorem):

$$\begin{aligned}
\lim_{\tau\to\infty} \int_0^\infty \Lambda_n(k\tau)\,\mathrm{d}\Phi_1(k) &= \int_0^\infty \lim_{\tau\to\infty} \Lambda_n(k\tau)\,\mathrm{d}\Phi_1(k) \\
&= \Lambda_n(0)\Phi_1(0) + \int_{k>0} \lim_{\tau\to\infty} \Lambda_n(k\tau)\,\mathrm{d}\Phi_1(k).
\end{aligned}$$

Since $\Phi_1(0) = 0$ and $\lim_{\tau\to\infty} \Lambda_n(k\tau) = 0$ for $k > 0$, the limit vanish. □

The next theorem gives a restriction on how slowly this limit can be approached.

Theorem 3.8 *For $\rho \in \mathfrak{D}_n$ where $n > 1$, ρ must behave at infinity according to*

$$\lim_{\tau \to \infty} \frac{\rho(\tau)}{\tau^{(1-n)/2}} = 0. \tag{3.20}$$

Proof. Assume the existence of a spectral density function $f(k)$, that is, assume that $\lim_{\tau \to \infty} \rho(\tau) = 0$. Then the integral (3.13) must be finite for all k. The asymptotic expansion for large positive real arguments of the Bessel function is (Abramowitz & Stegun 1972, Eq. 9.2.1)

$$J_\nu(x) = \sqrt{2/(\pi x)}\big(\cos(x - \nu\pi/2 - \pi/4) + \mathcal{O}(x^{-1})\big). \tag{3.21}$$

Inserting this in (3.13) gives the condition (ignoring insignificant constants)

$$\int_s^\infty \Big(\cos\big(k\tau - \phi(n)\big) + \mathcal{O}\big(|k\tau|^{-1}\big)\Big) \frac{\rho(\tau)}{(k\tau)^{(1-n)/2}} \,\mathrm{d}\tau < \infty \quad \text{for all } k. \tag{3.22}$$

The lower limit can be chosen arbitrarily large but finite. This condition must hold for the spectral density to exist. Thus a sufficient and necessary condition is that $\rho(\tau)/\tau^{(1-n)/2}$ approaches zero for large τ.

For correlation functions not possessing a spectral density function, the result still applies although the condition (3.22) can be violated for certain k's. □

A different proof is given by Christakos (1984, p. 261) who states that this result is a direct consequence of the lemma in Section 14.14 of Watson (1966).

Theorem 3.7 says that the correlation function must approach a specific limit as $\tau \to \infty$. An example of the opposite is $\rho(\tau) = \cos a\tau \in \mathfrak{D}_1$. The following theorem also implies that $\cos a\tau$ is not valid for $n > 1$.

Theorem 3.9 *If $\rho(\tau) \in \mathfrak{D}_n$ with $n > 1$, and $\rho(s) = 1$ for some $s > 0$, then $\rho(\tau) = 1$ for all τ.*

Proof. Suppose $\|\mathbf{t} - \mathbf{t}'\| = s$. To every v in the interval $(0, 2s)$ a point $\mathbf{t}''$ can be found such that $\|\mathbf{t} - \mathbf{t}''\| = s$ and $\|\mathbf{t}' - \mathbf{t}''\| = v$. Since both $X_{\mathbf{t}'}$ and $X_{\mathbf{t}''}$ have correlation 1 with $X_{\mathbf{t}}$, a perfect correlation must also exist between $X_{\mathbf{t}'}$ and $X_{\mathbf{t}''}$. Hence $\rho(v) = 1$, and the theorem follows by induction. □

The following theorem proposed by Matérn (1960, p. 13) strengthens the lower bound given by (3.1). The next theorem will strengthen this bound even further, but the simplicity of the first result makes it worthwhile to mention.

Theorem 3.10 *For any $\rho \in \mathfrak{D}_n$, then $\rho(\tau) \geq -\frac{1}{n}$, for any τ.*

Proof. It is possible to select a set of $n+1$ equidistant points in $\mathbb{R}^n$. By choosing such a set and taking $c_i = 1$ for $i = 1, \ldots, n+1$ in (1.7) gives: $(n+1) + \big((n+1)^2 - (n+1)\big)\rho(\tau) \geq 0$, which is equivalent. □

An absolute lower bound can also be found. To obtain the numerical values however, access to tables of Λ_n is required.

Theorem 3.11 *For any $\rho \in \mathfrak{D}_n$, then $\rho(\tau) \geq \inf_x \Lambda_n(x)$, for any τ.*

Proof. This is a direct consequence of (3.12) since $\Phi(\mathrm{d}k)$ is positive everywhere. □

Tables of Λ_n give (Jahnke, Emde & Lösch 1960, pp. 164–173)

(3.23a)	$\rho(\tau) \geq -1$	for $\rho \in \mathfrak{D}_1$ [from (3.11a)],
(3.23b)	$\rho(\tau) > -0.403$	for $\rho \in \mathfrak{D}_2$,
(3.23c)	$\rho(\tau) > -0.218$	for $\rho \in \mathfrak{D}_3$,
(3.23d)	$\rho(\tau) > -0.133$	for $\rho \in \mathfrak{D}_4$,
(3.23e)	$\rho(\tau) > -0.06$	for $\rho \in \mathfrak{D}_6$,
(3.23f)	$\rho(\tau) > -0.03$	for $\rho \in \mathfrak{D}_8$,
(3.23g)	$\rho(\tau) \geq 0$	for $\rho \in \mathfrak{D}_\infty$.

Thus noticeable negative correlations are restricted to lower dimensions. A consequence of Corollary 3.12.3 below, is that the last inequality must be strict.

3.3.4 Smoothness of Isotropic Correlation Functions

Theorem 3.12 *Any $\rho(\tau) \in \mathfrak{D}''_n$ is at least* $\lfloor((n-1)/2\rfloor$ times differentiable at any $\tau > 0$.*

A proof is found in Schoenberg (1938*a*, pp. 822–823). Three immediate consequences of the theorem are:

Corollary 3.12.1 *Any $\rho \in \mathfrak{D}''_3$ is everywhere differentiable.*

Corollary 3.12.2 *Any $\rho \in \mathfrak{D}''_5$ is everywhere two times differentiable.*

Corollary 3.12.3 *Any $\rho \in \mathfrak{D}''_\infty$ is analytic.*

Thus increasing smoothness is required for correlation functions in higher dimensions. For instance, the widely used spherical (3D) correlation function, (4.3c), is only one time differentiable at the correlation length. Therefore, Corollary 3.12.2 implies that the spherical correlation function is not usable on $\mathbb{R}^5$ or higher dimensions.

Corollary 3.12.3 implies that the inequality (3.23g) must be strict since otherwise $\rho = 0$ everywhere.

* $\lfloor((n-1)/2\rfloor$ means the largest integer not exceeding $(n-1)/2$.

4 Examples of Isotropic Correlation Functions

4.1 Special Correlation Functions

4.1.1 White Noise

The 'white noise' correlation function is defined as

$$\rho_{\mathrm{w}}(\boldsymbol{\tau}) = \begin{cases} 1 & \text{if } \boldsymbol{\tau} = \mathbf{0} \\ 0 & \text{otherwise} \end{cases} \qquad \rho_{\mathrm{w}} \in \mathfrak{D}'_{\infty}. \tag{4.1}$$

It belongs to $\mathfrak{D}'_{\infty}$ (discontinuous isotropic) and describes a completely chaotic phenomenon. The spectrum is uniform in all frequencies and hence a proper normalization of the spectral density function is impossible. This difficulty carries over to the alternative notation $\rho_{\mathrm{w}}(\boldsymbol{\tau}) = c\,\delta(\boldsymbol{\tau})$ where $\delta(\boldsymbol{\tau})$ is the Dirac δ-function. Since $\delta(\mathbf{0}) = \infty$, an appropriate scaling, c, must be zero! White noise models are hardly realistic models for any natural phenomenon. They are sometimes mistakenly considered as a measuring error effect, but measuring errors should be modeled separately. From an estimation and prediction point of view the difference is hardly recognizable, but a simulated sample path with a white noise component has nothing in common with a continuous sample path where conditioning is inexact owing to measuring errors. The nugget effect, see Journel & Huijbregts (1978, p. 39) or Christakos (1992, pp. 272–274), considered in geostatistics is therefore either a mistakenly modeled measuring error, or some erratic micro scale variation not properly modeled by Gaussian random fields.

4.1.2 Everywhere Constant Correlation

The opposite extreme is an everywhere constant correlation:

$$\rho(\boldsymbol{\tau}) = 1 \qquad \rho \in \mathfrak{D}''_{\infty}. \tag{4.2}$$

A random field possessing this correlation function is determined by a single random number, e.g.

$$X_{\mathbf{t}} = A\, f(\mathbf{t}),$$

where A is some random number and f is some deterministic function. A related covariance function was considered in Section 1.5.5 on non-stationary random fields. Consider the random field $X_{\mathbf{t}} = \sum_p A_p\, f_p(\mathbf{t})$ where A_p are Gaussian random variables. The non-stationary covariance function becomes

$$C(\mathbf{t}, \mathbf{s}) = \sum_{p,q} \Sigma_{pq}\, f_p(\mathbf{t})\, f_q(\mathbf{s}),$$

where Σ_{pq} is the covariance matrix of the A_p's.

These are seldom interesting models in themselves, but in combination with other correlation functions, common models appear, e.g. the ordinary or universal kriging model in geostatistics [see e.g. Journel & Huijbregts (1978, pp. 304ff) or Cressie (1991, p. 120)].

4.2 Continuous Non-negative Correlation Functions

In the following sections correlation functions belonging to $\mathfrak{D}''_n$ (everywhere continuous) will be considered. Recall that $\mathfrak{D}''_n \subset \mathfrak{D}''_{n-1}$.

First some examples of non-negative correlation functions are considered. All these are parameterized by the correlation length R. For some a scaling factor is introduced so that $\rho(\tau = R) \approx 0.05$. This is to simplify comparison.

4.2.1 Spherical

The following is taken from Matérn (1960, pp. 28–30) which includes more details. The spherical correlation functions originate from the considerations of a Poisson process, $N(\mathbf{t})$ on $\mathbb{R}^n$ with intensity λ. Consider the stationary random field defined as

$$X_{\mathbf{t}} = \int_{\mathbb{R}^n} I_{\{\tau<r\}} \mathrm{d}N(\mathbf{s}), \quad \text{where} \quad I_{\{\tau<r\}} = \begin{cases} 1 & \text{if } \tau < r, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $X_{\mathbf{t}}$ is the number of points within distance r from $\mathbf{t}$. The covariance function is given by

$$C(\tau) = \lambda \int_{\mathbb{R}^n} I_{\{\tau<r\}} I_{\{\|\mathbf{s}\|<r\}} \,\mathrm{d}^n\mathbf{s}.$$

So $C(\tau)$ is given by the volume of intersection of two n-dimensional spheres of radius r separated by τ. The correlations are zero beyond the correlation length, $R = 2r$. For $\tau \leq R$ the correlation functions are for $n = 1$, 2, 3, and 5 respectively

$$\rho_{\mathrm{sph1}}(\tau; R) = 1 - \tau/R \qquad \in \mathfrak{D}''_1 \tag{4.3a}$$

$$\rho_{\mathrm{sph2}}(\tau; R) = 1 - \frac{2}{\pi}\left(\frac{\tau}{R}\sqrt{1-\left(\frac{\tau}{R}\right)^2} + \arcsin\left(\frac{\tau}{R}\right)\right) \in \mathfrak{D}''_2 \tag{4.3b}$$

$$\rho_{\mathrm{sph3}}(\tau; R) = 1 - \frac{3}{2}\frac{\tau}{R} + \frac{1}{2}\left(\frac{\tau}{R}\right)^3 \qquad \in \mathfrak{D}''_3 \tag{4.3c}$$

$$\rho_{\mathrm{sph5}}(\tau; R) = 1 - \frac{15}{8}\frac{\tau}{R} + \frac{5}{4}\left(\frac{\tau}{R}\right)^3 - \frac{3}{8}\left(\frac{\tau}{R}\right)^5 \qquad \in \mathfrak{D}''_5. \tag{4.3d}$$

The correlation function for $n = 3$ is probably the most used. Figure 4.1 shows a plot of these four spherical correlation functions and Figure 4.9 contains some simulated sample paths.

The correlation function $\rho(\tau; R)_{\mathrm{sph2}}$ is sometimes called the circular correlation function and the correlation function $\rho(\tau; R)_{\mathrm{sph5}}$ is sometimes called the penta-spherical covariance function.

The positive definiteness is guaranteed by construction. Some are possibly valid in dimensions beyond the defining dimension.

4.2.2 Cubic

The cubic correlation function is zero beyond R. For $t \leq R$ the correlation function is (Wackernagel 1995, p. 218)

$$\rho_{\mathrm{cubic}}(\tau; R) = 1 - 7\left(\frac{\tau}{R}\right)^2 + \frac{35}{4}\left(\frac{\tau}{R}\right)^3 - \frac{7}{2}\left(\frac{\tau}{R}\right)^5 + \frac{3}{4}\left(\frac{\tau}{R}\right)^7 \quad \in \mathfrak{D}''_2. \tag{4.4}$$

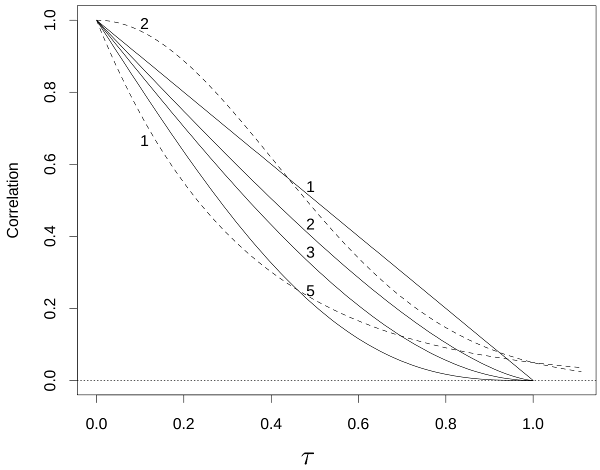

Figure 4.1: Spherical correlation functions (solid lines) obtained with $n = 1$, 2, 3, and 5. Exponential correlation functions with $\nu = 1$ and 2 are added for comparison (broken lines). The correlation length is $R = 1$ for all.

This correlation function is illustrated in Figure 4.2 and a sample path is shown in Figure 4.10.

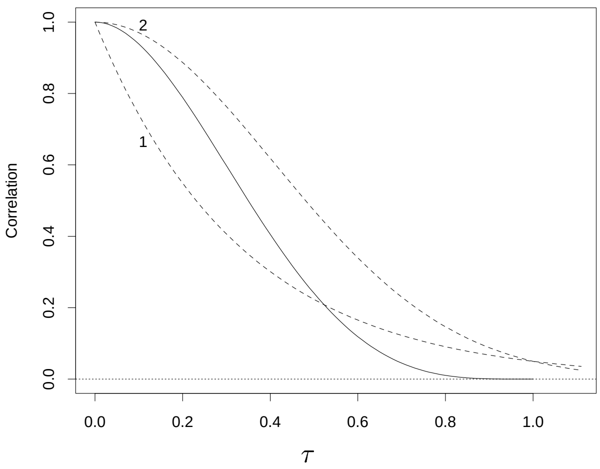

Figure 4.2: Cubic correlation function (solid line). Exponential correlation functions with $\nu = 1$ and 2 are added for comparison (broken lines). The correlation length is $R = 1$ for all.

4.2.3 Exponential

The exponential* correlation function is defined as

$$\rho_{\text{exp}}(\tau; R, \nu) = e^{-3(\tau/R)^{\nu}} \quad \in \mathfrak{D}''_{\infty} \text{ for } 0 < \nu \leq 2. \tag{4.5}$$

*The exponential correlation function is also called the stable correlation function.

The scaling factor '-3' is chosen so that the correlation at the correlation length is $e^{-3} \approx 0.05$. Figure 4.3 illustrates the class of exponential correlation functions. The figure suggests that the exponential class does constitute a ν dependent range of correlation functions with continuously changing properties. This is not so: Examples 2.2 and 2.3 shows that the continuity and differentiability properties are different for the sets $\{0 < \nu < 1\}$, $\{\nu = 1\}$, $\{1 < \nu < 2\}$, and $\{\nu = 2\}$.

Figure 4.11 shows some sample paths for different values of ν.

When $\nu = 2$ the corresponding correlation function is sometimes called the Gaussian correlation function.

Note that $\nu = 0$ corresponds to a white noise process with a constant global correlation of $e^{-3} \approx 0.05$ which belongs to $\mathfrak{D}'_\infty$.

A proof of the positive definiteness is found in Schoenberg (1938*b*, p. 532).

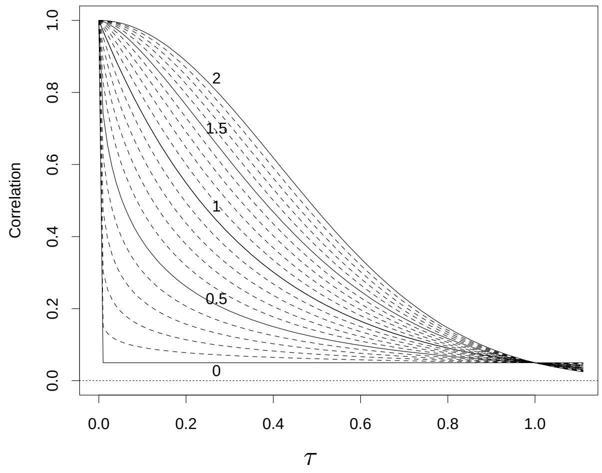

Figure 4.3: Exponential correlation functions for $\nu = 0,\ 0.1,\ 0.2,\ \ldots,\ 1.9$ and 2. (Solid lines for $\nu = 0,\ 0.5,\ 1.0,\ 1.5$, and 2.0.) The correlation length is $R = 1$ for all.

4.2.4 Rational Quadratic

The rational quadratic* correlation function is given by

$$\rho_{\mathrm{rq}}(\tau; R, \nu) = \frac{1}{\big(1 + \mathcal{S}_\nu(\tau/R)^2\big)^\nu} \qquad \rho \in \mathfrak{D}''_\infty \text{ for } \nu > 0. \tag{4.6}$$

Choosing the scaling factor as $\mathcal{S}_\nu = 20^{1/\nu} - 1$ gives correlation 0.05 at the correlation length. Figure 4.4 gives an illustration of the correlation functions and Figure 4.12 shows some sample paths for different ν values.

A proof of the positive definiteness based on representation (3.11d) is found in Matérn (1960, p. 17). See also Example 3.5.

*The rational quadratic correlation function is also called the Cauchy correlation function.

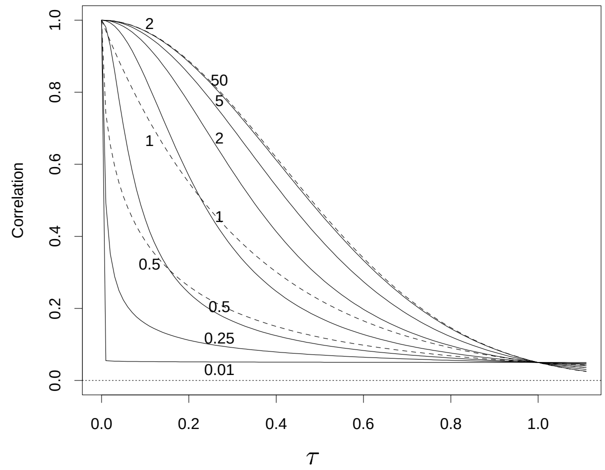

Figure 4.4: Rational quadratic correlation functions (solid lines) for $\nu = 0.01$, 0.25, 0.5, 1, 2, 5, and 50. Exponential correlation functions with $\nu = 0.5$, 1 and 2 are added for comparison (broken lines). The correlation length is $R = 1$ for all.

4.2.5 Modified Bessel Function

The modified Bessel* correlation functions are related to the exponential correlation functions. They are defined as

$$\text{(4.7)} \qquad \rho_K(\tau; R, \nu) = \frac{1}{\Gamma(\nu)2^{\nu-1}} z^\nu K_\nu(z); \quad z = \mathcal{S}_\nu \frac{\tau}{R}, \quad \in \mathfrak{D}''_\infty \text{ for } \nu > 0,$$

where K_ν is the modified Bessel function of order ν. The ν-dependent scaling $\mathcal{S}_\nu$ is chosen such that $\rho(R; R, \nu) \approx 0.05$ for simple comparison; see Appendix A. Note that $\rho_K(\tau; R, \nu = 0.5) = \rho_{\exp}(\tau; R, \nu = 1)$. Some examples of the correlation function is found in Figure 4.5.

The smoothness of the sample paths is determined by the behavior of ρ at $t = 0$. The limiting form for small arguments is

$$K_\nu(z) \sim \Gamma(\nu) 2^{\nu-1} z^{-\nu}; \qquad \nu > 0,$$

(Abramowitz & Stegun 1972, Eq. 9.6.9) such that $\rho(r; R, \nu) \to 1$ as $t \to 0$. (Note that $K_0(z) \sim -\ln z$ so that $\nu = 0$ is not acceptable.) The derivatives satisfy the following recurrence relation:

$$\frac{\mathrm{d}}{\mathrm{d}z}\big(z^\nu K_\nu(z)\big) = -z^\nu K_{\nu-1}(z),$$

(Abramowitz & Stegun 1972, Eq. 9.6.28). According to Corollary 2.4.2 the correlation function must be everywhere $2m$ times differentiable for a random field to be m times mean square differentiable.

For $0 < \nu < \frac{1}{2}$, $\partial\rho(r; R, \nu)/\partial r|_{r=0} = -\infty$ causing extremely erratic surfaces. For $\frac{1}{2} \leq \nu < 1$, $\partial\rho(r; R, \nu)/\partial r|_{r=0} \in (-\infty, 0)$ which gives a wide range of erratic surfaces. For $n \leq \nu$ then $\partial^{2n-1}\rho(r; R, \nu)/\partial r^{2n-1}|_{r=0} = 0$ and

*The modified Bessel correlation function is also called the Basset or the Matérn correlation function.

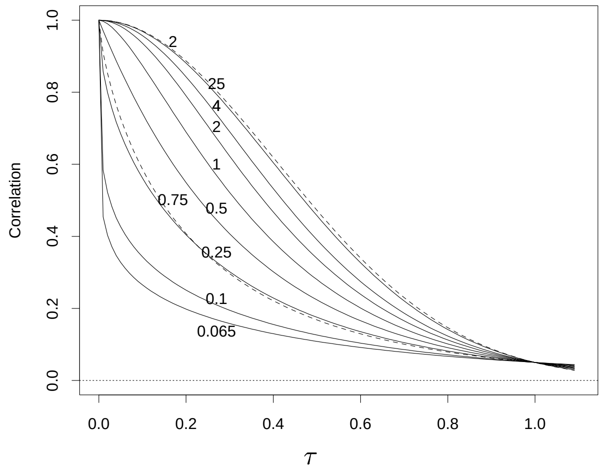

Figure 4.5: Modified Bessel correlation functions (solid lines) for $\nu = 0.065$, 0.1, 0.25, 0.5, 1, 2, 4, and 25. Exponential correlation functions with $\nu = 0.75$ and 2 are added for comparison (broken lines). The correlation length is $R = 1$ for all.

$\partial^{2n}\rho(r; R, \nu)/\partial r^{2n}|_{r=0} \in (-\infty, 0)$ implying that the stochastic surface is n times differentiable. So increasing ν gives increasingly smoother surfaces.

The positive definiteness of modified Bessel correlation functions are established in Yaglom (1986*a*, pp. 362–363) and Matérn (1960, p. 17). See also Example 3.4.

4.3 Continuous Oscillating Correlation Functions

In the following some examples of oscillating correlation functions are given. They are parameterized by the (angular) frequency, ω, or alternatively by the period $\nu = 2\pi/\omega$.

4.3.1 Bessel Function of the First Kind

These Bessel correlation functions are damped oscillations. They are defined as

$$(4.8) \quad \rho_J(\tau; \omega, \nu) = \Gamma(\nu+1)2^{\nu}(\omega\tau)^{-\nu}J_\nu(\omega\tau) \quad \in \mathfrak{D}''_n \quad \text{for } \nu \geq (n-2)/2 \text{ and } \omega > 0.$$

The restriction on ν implies e.g. that $\nu \geq 0.5$ when $n = 3$. Figure 4.6 illustrates some of these correlation functions.

The limiting form of the Bessel functions of the first kind for small arguments are

$$J_\nu(z) \sim (z/2)^\nu/\Gamma(\nu+1), \qquad (\nu \neq -1, -2, -3\ldots)$$

(Abramowitz & Stegun 1972, Eq. 9.1.7). Thus $\rho(r; R, \nu) \to 1$ as $t \to 0$.

For integer ν the positive definiteness is obvious from the representation (3.9). The positive definiteness for arbitrary ν is established in Yaglom (1986*a*, pp. 366–367) and in Matérn (1960, p. 18).

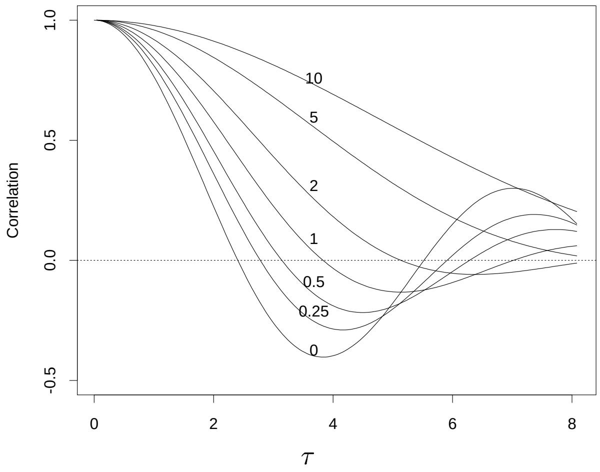

Figure 4.6: Ordinary Bessel correlation functions for $\nu = 0$, 0.25, 0.5, 1, 2, 5, and 10. Note that $\nu = 0.5$ coincidence with the correlation function $\sin \omega\tau/\omega\tau$. The angular frequency is $\omega = 1$ for all.

4.3.2 Exponentially Damped Cosine

The exponentially damped cosine correlation functions are

(4.9)

$$\rho_{\cos}(\tau;\omega,R) = e^{-3\tau/R}\cos\omega\tau \quad \in \begin{cases} \mathfrak{D}_1'' \text{ for } \omega > 0 \text{ and } R \geq 0, \\ \mathfrak{D}_2'' \text{ for } \omega > 0 \text{ and } 3/\omega \geq R \geq 0, \\ \mathfrak{D}_3'' \text{ for } \omega > 0 \text{ and } \sqrt{3}/\omega \geq R \geq 0. \end{cases}$$

The bounds on R for different dimensions are found in Yaglom (1986*a*, p. 366). The period (one cycle) is $\nu = 2\pi/\omega$ so that the bounds requires that $R \lessapprox \frac{1}{2}\nu$ for $\mathbb{R}^2$ and $R \lessapprox \frac{1}{4}\nu$ for $\mathbb{R}^3$.

Figure 4.7 illustrates these oscillatory correlation functions and Figure 4.13 shows some sample paths for different R.

4.3.3 Sine Damped by Inverse Distance

This correlation function is sometimes referred to as the wave or hole effect correlation function (Cressie 1991, p. 62). It is defined as

$$\rho_{\sin}(\tau;\omega) = \frac{1}{\omega\tau}\sin\omega\tau \quad \in \mathfrak{D}_3'' \text{ for } \omega > 0. \tag{4.10}$$

Note that $\rho_{\sin}(\tau;\omega) = \rho_J(\tau;\nu = 0.5,\omega)$ which establish the positive definiteness. Figure 4.8 shows this correlation function and Figure 4.14 shows some sample paths for different periods. The period (one cycle) is once again given by $\nu = 2\pi/\omega$.

A summary of some properties of the correlation functions are given in Table 4.1.

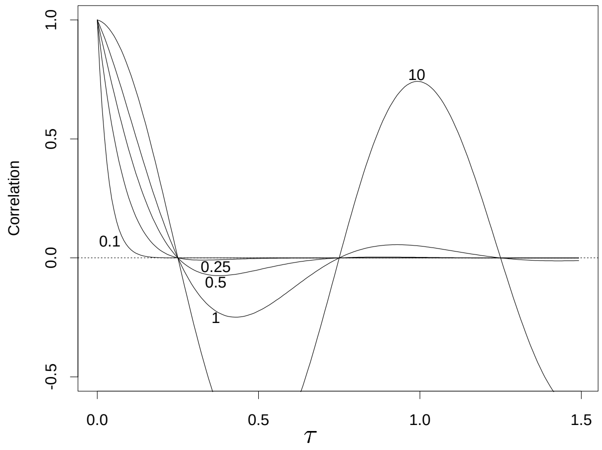

Figure 4.7: The exponentially damped cosine correlation function has been plotted with $R/\nu = 0.1, 0.25, 0.5, 1$ and 10. The period, $\nu = 2\pi/\omega$, is 1 for all.

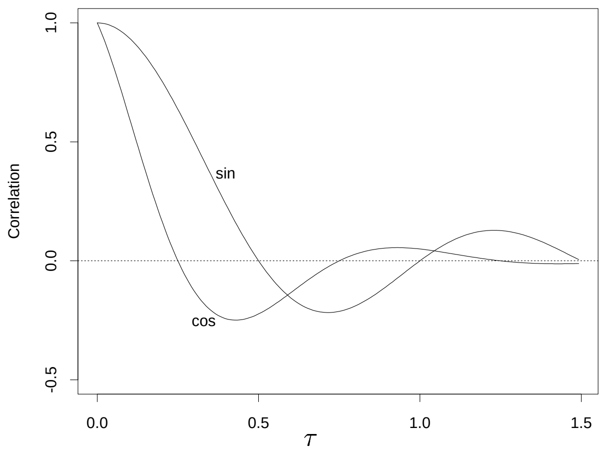

Figure 4.8: The damped sine correlation function with period $\nu = 1$. For comparison the damped cosine correlation function with the same period and $R = 1$ is also plotted. Note that they have a phase shift of $\pi/4$.

4.4 Examples of Sample Paths

Figures 4.9 to 4.14 show some simulated sample paths (realizations) in $\mathbb{R}^1$ for various correlation functions. All have zero expectation and unit variance. The horizontal axes are t and the vertical axes are $X(t)$. The horizontal axes are not labeled but the relevant length scale is marked in the figures.

Each sample path is stored as 200 equidistant points. This is sufficient for most sample paths but the most irregular are slightly smoothed by the lack of resolution. All sample paths are produced using the same 200 pseudo-random numbers. Similar correlation functions will therefore give similar sample paths.

The algorithm for drawing the sample paths consists of constructing the 200×200 covariance matrix for the 200 equidistant points. Then a square root of the covariance matrix is obtained by using the eigenvalue decomposition of

Table 4.1: Summary of properties of correlation functions.

Correlation functions	Eq.	Param.	Class*	Geom.†	Characteristics
White noise	(4.1)		$\mathfrak{D}'_\infty$	0	Zero range
Everywhere constant	(4.2)		$\mathfrak{D}''_\infty$	∞	Infinite range
Spherical	(4.3a–d)	R, n	$\mathfrak{D}''_n$	f	Finite range
Cubic	(4.4)	R	$\mathfrak{D}''_2$	d	Finite range
Exponential	(4.5)	R, ν	$\mathfrak{D}''_\infty$	$\vert$ to ∞	
Rational quadratic	(4.6)	R, ν	$\mathfrak{D}''_\infty$	∞	
Modified Bessel	(4.7)	R, ν	$\mathfrak{D}''_\infty$	$\vert$ to ∞	
Bessel of 1. kind	(4.8)	ω, ν	$\mathfrak{D}''_{n(\nu)}$		Damped osc.
Damped cosine	(4.9)	ω, R	$\mathfrak{D}_{n(\omega,R)}$		Damped osc.
Damped sine (wave)	(4.10)	ω	$\mathfrak{D}''_3$		Damped osc.

*See Section 3.3.
†Differentiability as $t \to 0$:
'0' means discontinuous sample paths,
'|' means an infinite first order derivative at 0,
'f' means finite negative first order derivative at 0,
'd' means differentiable (finite negative second order derivative at 0),
'∞' means analytic sample paths.

the covariance matrix. This is more stable than the standard Cholesky decomposition which fails for very smooth sample paths. Minute negative eigenvalues caused by numerical inaccuracies are set to zero. The simulation algorithm is exact apart from the minor numerical problems caused by the negative eigenvalues.

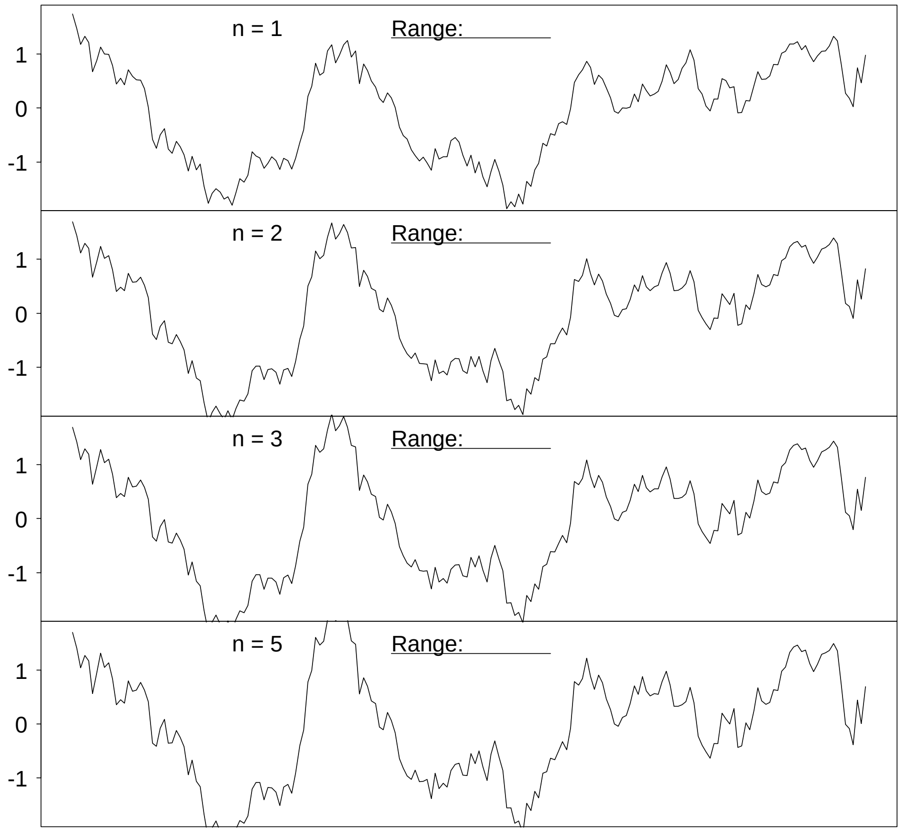

Figure 4.9: Sample paths with spherical correlation functions with $n = 1, 2, 3$ and 5. Note how increasing n gives slightly increasing irregularity. The spherical correlation function with $n = 3$ is the most widely used.

Figure 4.10: Sample path with cubic correlation function. This path looks similar to the path obtained using an exponential correlation function in Figure 4.11 with $\nu = 1.9$. Note however that this sample path is differentible whereas the 'exponential' is not.

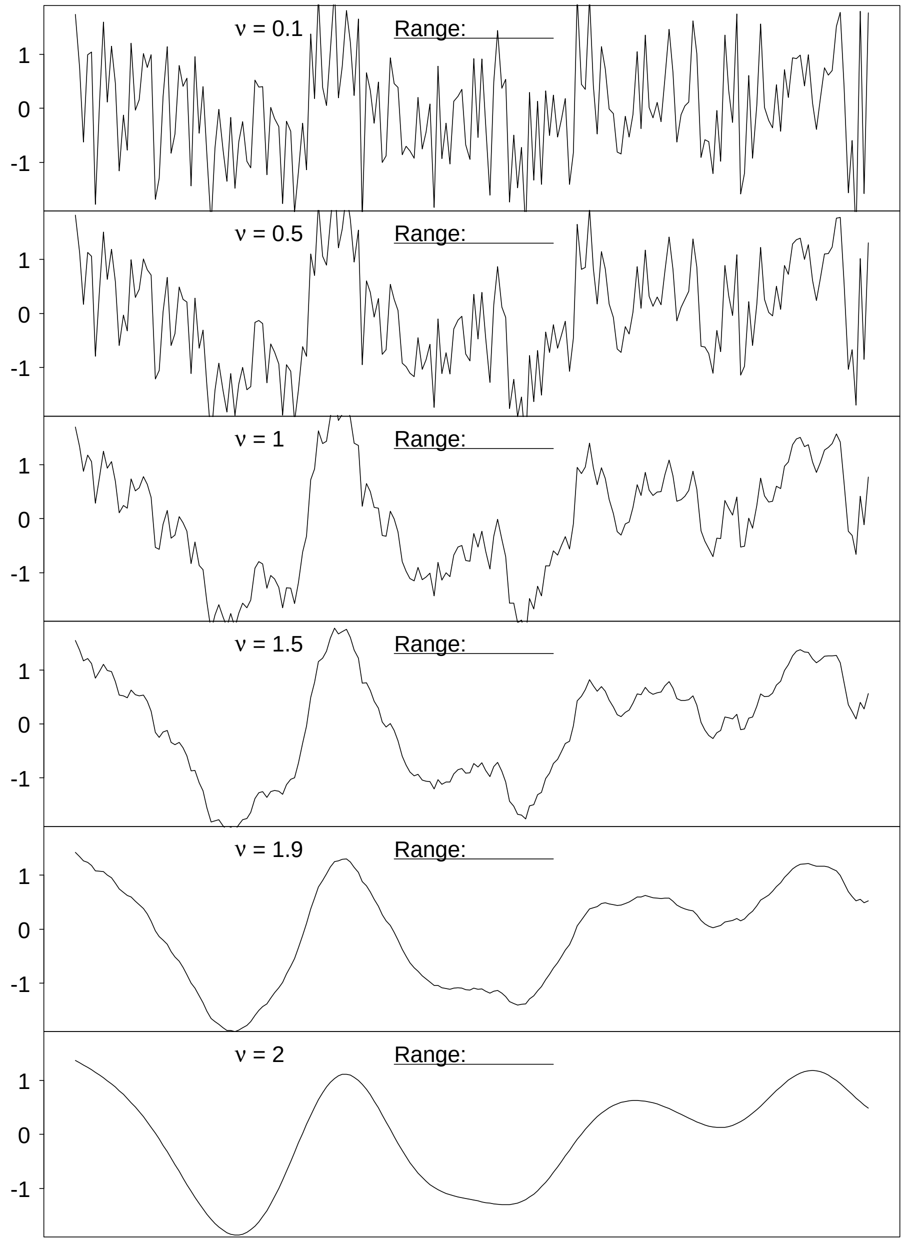

Figure 4.11: Sample paths with exponential correlation functions with $\nu = 0.1, 0.5, 1, 1.5, 1.9$, and 2 (Gaussian). The exponential correlation function with $\nu = 1$ is the most widely used.

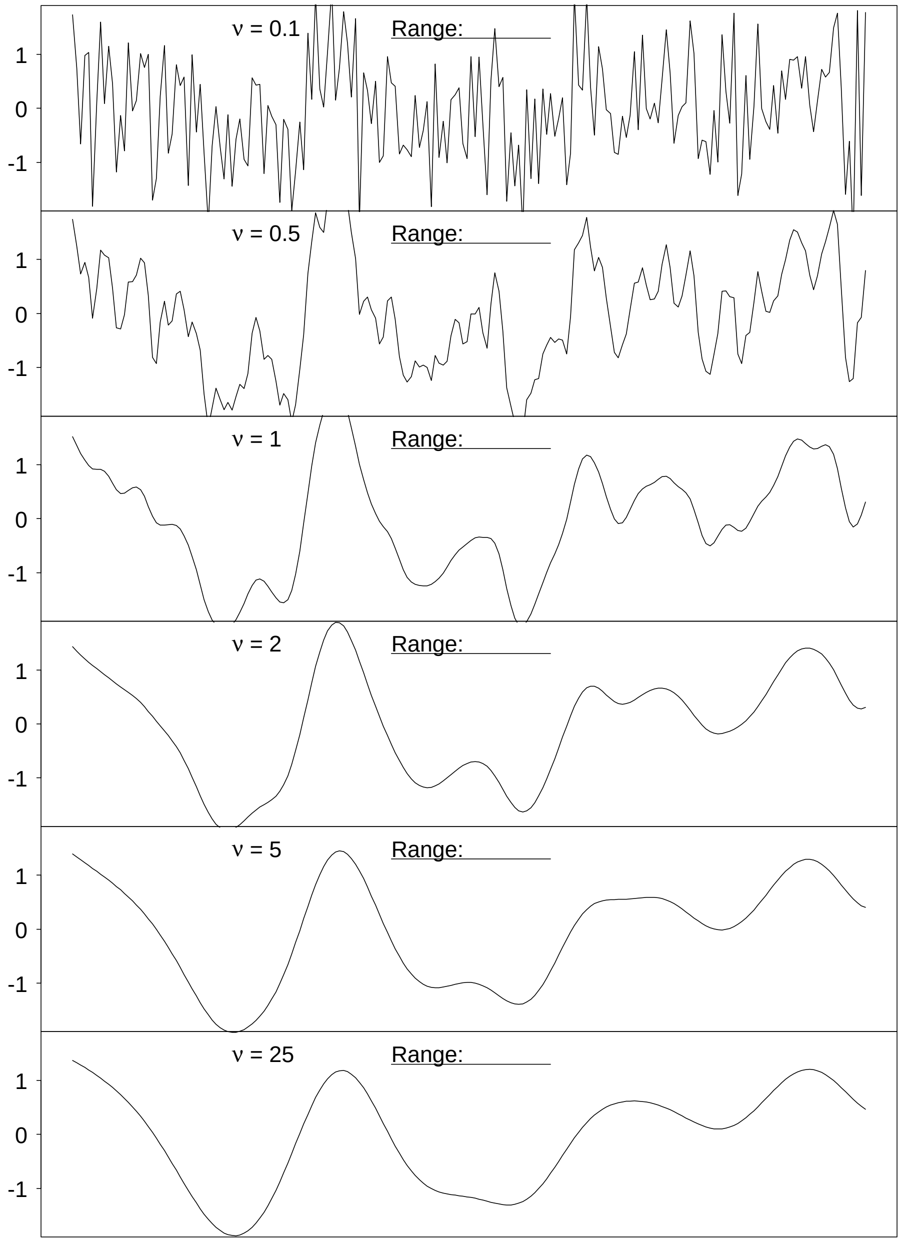

Figure 4.12: Sample paths with rational quadratic correlation functions with $\nu = 0.1, 0.5, 1, 2, 5$, and 25. Note that $\nu = 25$ gives a sample path which is indistinguishable from the path obtained using the Gaussian correlation function in Figure 4.11.

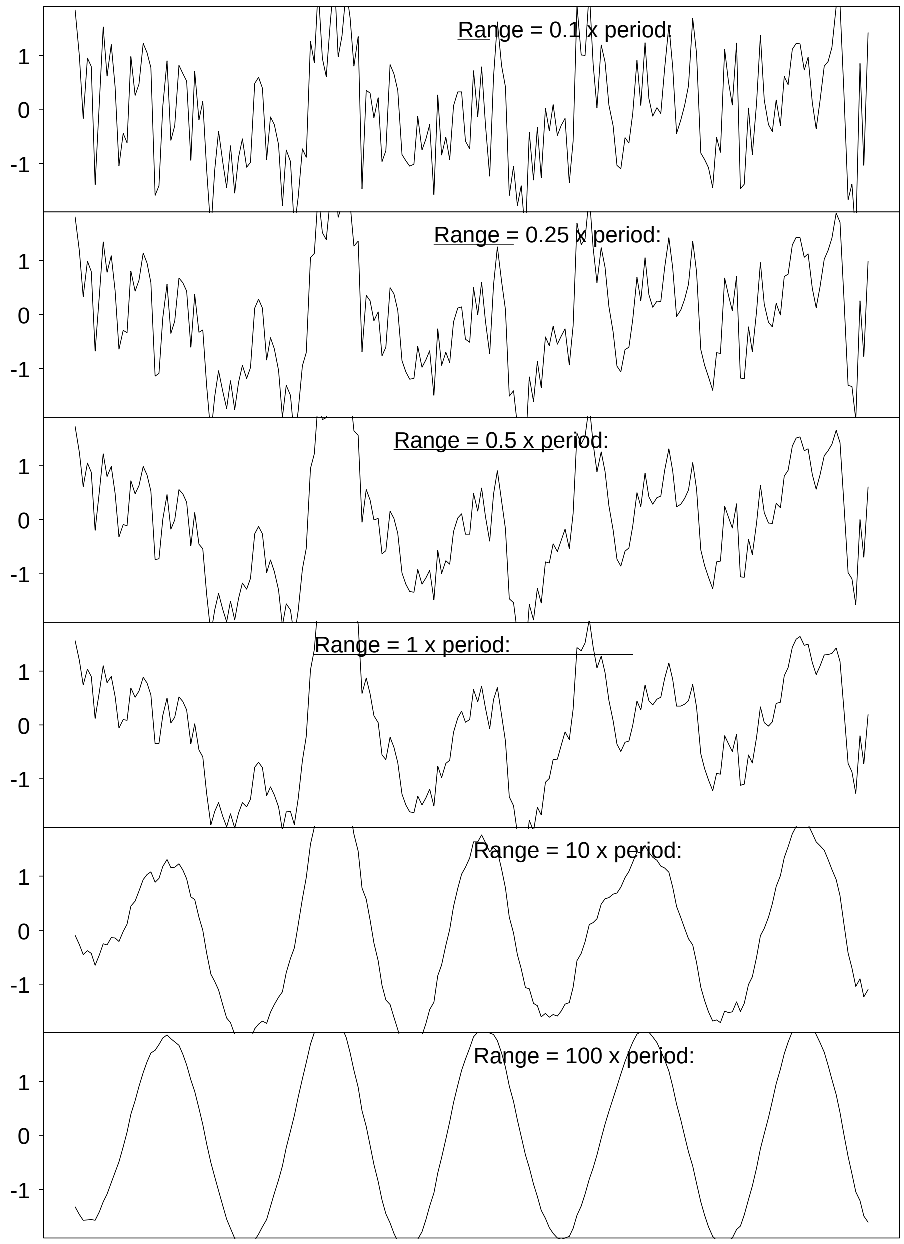

Figure 4.13: Sample paths with damped cosine correlation functions. The period, $\nu = 2\pi/\omega$, is identical in all figures and is clearly seen on the lowest figure.

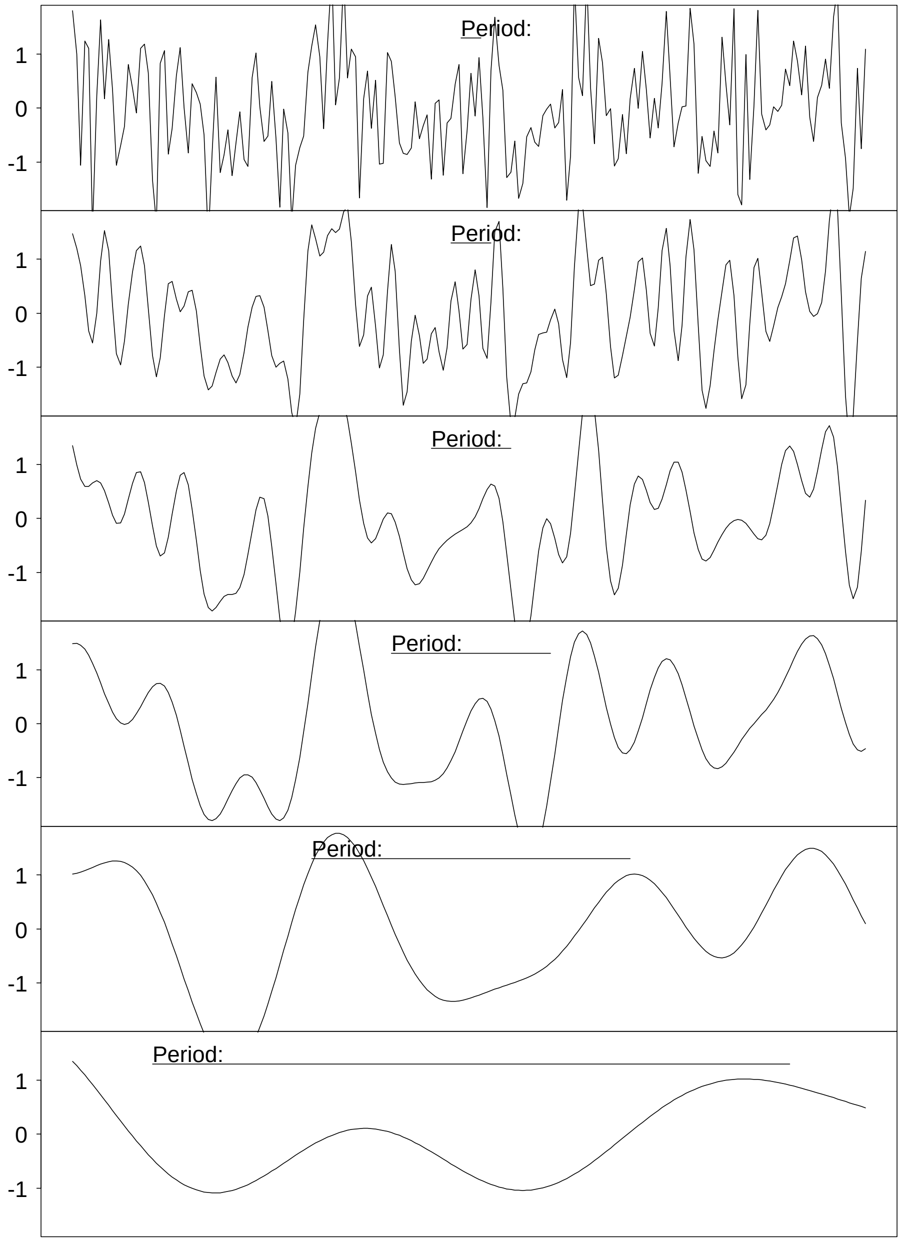

Figure 4.14: Sample paths with damped sine correlation functions. The period is $\nu = 2\pi/\omega$.

A Scaling of Modified Bessel Correlation Functions

The scaling factor $\mathcal{S}_\nu$ in (4.7) is obtained by using Newton's method. Fast convergence is ensured since the correlation functions are monotonically decreasing. Consider the non linear equation $\rho_K(R; R, \nu) = 0.05$ leading to $(z = \mathcal{S}_\nu)$

$$F(z) = z^\nu K_\nu(z) - 0.05 \cdot 2^{\nu-1}\Gamma(\nu) = 0.$$

Newton's method is to iterate

$$z_{i+1} = z_i - \frac{F(z_i)}{\frac{\mathrm{d}}{\mathrm{d}z}F(z_i)}$$

which for the present case gives

$$z_{i+1} = z_i + \frac{z_i^\nu K_\nu(z_i) - 0.05 \cdot 2^{\nu-1}\Gamma(\nu)}{z_i^\nu K_{\nu-1}(z_i)},$$

where

$$\frac{\mathrm{d}}{\mathrm{d}z}\big(z^\nu K_\nu(z)\big) = -z^\nu K_{\nu-1}(z)$$

has been used (Abramowitz & Stegun 1972, Eq. 9.6.28). Any initial point $z_0 > 0$ will do. A reasonable stopping criterion is chosen such that $|F(z_i)| < \delta$ for some reasonably small δ, say $\delta = 0.0001$. The final z_i value is used for $\mathcal{S}_\nu$.

References

ABRAMOWITZ, M. & STEGUN, I. A., eds (1972), *Handbook of Mathematical Functions*, 9th printing, Dover Publications Inc., New York, 1046 pp.

ADLER, R. J. (1981), *The Geometry of Random Fields*, John Wiley & Sons, New York, 280 pp.

ADLER, R. J. (1990), An introduction to continuity, extrema, and related topics for general Gaussian processes, Lecture notes—monograph series 12, Institute of Mathematical Statistics, Hayward, CA, 160 pp.

BELYAEV, YU. K. (1972), Point processes and first passage problems, *in* 'Proc. Sixth Berkeley Symp. Math. Statist. Prob., **2**', Univ. of California Press, Berkeley, pp. 1–17.

BILLINGSLEY, P. (1986), *Probability and Measure*, 2nd edn, John Wiley & Sons, New York, 622 pp.

BOCHNER, S. (1959), *Lectures on Fourier Integrals*, Princeton University Press, Princeton, N.J.

CHRISTAKOS, G. (1984), 'On the problem of permissible covariance and variogram models', *Water Resour. Res.* **20**(2), 251–265.

CHRISTAKOS, G. (1992), *Random Field Models in Earth Sciences*, Academic Press Inc., San Diego, 474 pp.

CRAMÉR, H. & LEADBETTER, M. R. (1967), *Stationary and Related Stochastic Processes*, John Wiley & Sons, New York, 348 pp.

CRESSIE, N. (1991), *Statistics for Spatial Data*, John Wiley & Sons, New York, 900 pp.

DOOB, J. L. (1953), *Stochastic Processes*, John Wiley & Sons, New York, 654 pp.

GNEDENKO, B. V. (1962), *The Theory of Probability*, Chelsea Publishing Co., New York. Translated from Russian by B. D. Seckler.

GRADSHTEYN, I. S. & RYZHIK, I. M. (1980), *Tables of Integrals, Series, and Products: Corrected and Enlarged Edition*, Academic Press Inc., New York.

JAHNKE, E., EMDE, F. & LÖSCH, F. (1960), *Tafeln Höherer Funktionen/Tables of Higher Functions*, 6th edn, B. G. Teubner Verlagsgesellschaft, Stuttgart, 318 pp. (Bilingual).

JOURNEL, A. G. & HUIJBREGTS, C. J. (1978), *Mining Geostatistics*, Academic Press Inc., London, 600 pp.

LOÈVE, M. (1978), *Probability Theory I*, Graduate Texts in Mathematics 45, 4th edn, Springer-Verlag Inc., New York, 425 pp.

MANTOGLOU, A. & WILSON, J. (1982), 'The turning bands method for simulation of random fields using line generation by a spectral method', *Water Resour. Res.* **5**, 1379–1394.

MARDIA, K. V., KENT, J. T. & BIBBY, J. M. (1979), *Multivariate Analysis*, Academic Press Inc., London, 518 pp.

MATÉRN, B. (1960), Spatial variation: Stochastic models and their application to some problems in forest surveys and other sampling investigations, Meddelanden från Statens Skogforskningsinstitut Band 49 No. 5, Swedish Forest Research Institute, Sweden, 144 pp. Ph.D. thesis, rebublished as Matérn (1986).

MATÉRN, B. (1986), *Spatial Variation*, 2nd edn, Springer-Verlag Inc., Berlin, 151 pp. First edition published in 1960 by the Swedish Forest Research Institute.

MATHERON, G. (1973), 'The intrinsic random functions and their applications', *Advances in Appl. Probab.* **5**, 439–468.

MONIN, A. S. & YAGLOM, A. M. (1975), *Statistical Fluid Mechanics: Mechanics of Turbulence*, Vol. 2, The MIT Press, Cambridge, Massachusetts, 874 pp.

PRIESTLEY, M. B. (1992), *Spectral Analysis and Time Series*, Academic Press Inc., San Diego, California, 890 pp. Originally published in 1981 in two volumes.

RUDIN, W. (1973), *Functional Analysis*, McGraw-Hill Book Company Inc., New York, 397 pp.

SCHOENBERG, I. J. (1938*a*), 'Metric spaces and completely monotone functions', *Anal. Math.* **39**(4), 811–841.

SCHOENBERG, I. J. (1938*b*), 'Metric spaces and positive definite functions', *Trans. Amer. Math. Soc.* **44**(3), 522–536.

VANMARCKE, E. (1983), *Random Fields: Analysis and Synthesis*, The MIT Press, Cambridge, Massachusetts, 382 pp.

WACKERNAGEL, H. (1995), *Multivariate Geostatistics*, Springer-Verlag Inc., Berlin Heidelberg, 256 pp.

WATSON, G. N. (1966), *Theory of Bessel Functions*, Cambridge University Press, London.

YAGLOM, A. M. (1986*a*), *Correlation Theory of Stationary and Related Random Functions I: Basic Results*, Springer Series in Statistics, Springer-Verlag Inc., New York, 526 pp.

YAGLOM, A. M. (1986*b*), *Correlation Theory of Stationary and Related Random Functions II: Supplementary Notes and References*, Springer Series in Statistics, Springer-Verlag Inc., New York, 258 pp.

List of Tables

List of Figures

Notation

Abbreviations:
cf. – correlation function(s), cont. – continuous, cov. – covariance, df. – distribution function, iso. – isotropic, rf. – random field.

$\|\cdot\|$ norm of vector.
$|\cdot|$ absolute value of scalar.
$|\cdot|$ determinant of matrix.
$\mathcal{B}^k$ the Borel sets in $\mathbb{R}^k$.
$B \in \mathcal{B}^k$ a Borel set.
$\mathfrak{B}_n$ class of cf. in $\mathbb{R}^n$.
C cov. function.
$\mathbf{C}$ cov. tensor.
$\ddot{\mathbf{C}}$ cov. tensor of gradient field.
$\mathbf{C}^{(\kappa,\lambda)}$ cov. tensor for higher order derivatives.
$\mathrm{Corr}\{\cdot\}$ correlation.
$\mathrm{Cov}\{\cdot\}$ cov.
$\mathfrak{C}_n \subset \mathfrak{B}_n$ class of stationary cf. in $\mathbb{R}^n$.
$\mathfrak{C}'_n \subset \mathfrak{C}_n$ — cont. except at $\mathbf{0}$.
$\mathfrak{C}''_n \subset \mathfrak{C}'_n$ — cont.
$\mathfrak{D}_n \subset \mathfrak{C}_n$ class of iso. cf. in $\mathbb{R}^n$.
$\mathfrak{D}'_n \subset \mathfrak{D}_n$ — cont. except at 0.
$\mathfrak{D}''_n \subset \mathfrak{D}'_n$ — cont.
$\mathrm{E}\{\cdot\}$ expectation.
$\mathcal{F}$ σ-algebra, e.g. Borel sets $\mathcal{B}^k$.
$F_{\mathbf{t}_1,\ldots,\mathbf{t}_k}(\cdot)$ finite-dimensional df.
$F(\mathbf{k})$ spectral df.
$f(\mathbf{k})$ spectral density function.
$\Phi(k)$ iso. spectral df.
$f(k)$ iso. spectral density function.
Γ gamma function.
J_n Bessel function.
K_n modified Bessel function.
$\Lambda_n(x) \propto J_{(n-2)/2}(x)/x^{(n-2)/2}$.
$m(\mathbf{t})$ expectation of rf.
P probability measure.
$p_{\mathbf{t}_1,\ldots,\mathbf{t}_k}(\cdot)$ finite-dimensional probability density function.
$\mathrm{Prob}\{\cdot\}$ probability of event.
$\rho(\cdot)$ correlation function.
$\mathbb{R}^n$ n dimensional Euclidean space.
$\mathbf{\Sigma}$ cov. matrix.
Σ_n area of unit sphere in $\mathbb{R}^n$.
$\sigma^2(\mathbf{t})$ variance of rf.
$\mathbf{\Sigma}^t$ transposed vector or matrix.
T parameter set, e.g. $\mathbb{R}^n$.
$\mathbf{t},\mathbf{s}$ position, coordinate, i.e. $\in T$.
$\boldsymbol{\tau}$ separation vector, e.g. $\boldsymbol{\tau} = \mathbf{t} - \mathbf{s}$.
$\tau = \|\boldsymbol{\tau}\|$ norm of $\boldsymbol{\tau}$.
$\mathrm{Var}\{\cdot\}$ variance.
$X, X_{\mathbf{t}}, X(\mathbf{t}), X(\mathbf{t},\omega)$ rf.
$\mathbf{X}$ multidimensional or vector rf.
$\dot{\mathbf{X}}$ gradient field.
$\ddot{\mathbf{X}}, \dddot{\mathbf{X}}, \mathbf{X}^{(\kappa)}$ higher order derivatives.
$x_{\mathbf{t}}$ sample path.
Ω sample space.
ω sample point in Ω.
$(\Omega, \mathcal{F}, P)$ probability space.

Index