Mathematical Methods of Uncertainty Quantification

Oliver Ernst

Professur Numerische Mathematik

Sommersemester 2016

Mathematik!
TU Chemnitz

Contents

Contents

Contents

Random Fields

- Similar to a stochastic process, a random field is a family of random variables indexed by a parameter. The former concept is often tied to a parameter set which is totally ordered (e.g. $\mathbb{N}$ or $\mathbb{R}_0^+$), whereas for random fields the parameter is a spatial coordinate, typically from subsets of $\mathbb{R}^2$ or $\mathbb{R}^3$.
- Random fields first arose in the field of geostatistics to model phenomena in Earth Sciences such as hydrology, agriculture or geology.
- Since the data for PDE models often consists of one or more functions of space, it is natural to specify the uncertain or random data for PDEs as random fields.
- The alternative view of random fields is as random variables with values in abstract sets, such as spaces of functions, equivalence classes of functions or distributions.
- Naturally, there are extensions to spatio-temporal random fields featuring an additional (ordered) parameter used to model, e.g., turbulence or meteorological phenomena.

Definition 3.1

Given a set $D \subset \mathbb{R}^d$, $d \in \mathbb{N}$ and a probability space $(\Omega, \mathfrak{A}, \mathbf{P})$, a (real-valued) random field is a mapping

$$a : D \times \Omega \to \mathbb{R}$$

such that each function $a(\boldsymbol{x}, \cdot) : \Omega \to \mathbb{R}$, $\boldsymbol{x} \in D$, is a random variable.

Definition 3.2

For each fixed $\omega \in \Omega$ the associated function $a(\cdot, \omega) : D \to \mathbb{R}$ is called a realization of the random field.

- Denote by $\mathbb{R}^D$ the set of all real-valued functions $f : D \to \mathbb{R}$. In particular, realizations of a real-valued random field belong to $\mathbb{R}^D$ by Definition 3.2.
- Denote further by $\mathfrak{A}(\mathbb{R}^D)$ the smallest σ-algebra containing all sets

$$A = \{f \in \mathbb{R}^D : (f(\boldsymbol{x}_1), \dots, f(\boldsymbol{x}_n)) \in B\}$$

for any $B \in \mathfrak{B}(\mathbb{R}^n), \boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in D, n \in \mathbb{N}$.

Proposition 3.3

Let a be a random field on $D \subset \mathbb{R}^d$ with underlying probability space $(\Omega, \mathfrak{A}, \mathbf{P})$. Then the mapping $\omega \mapsto a(\cdot, \omega)$ from $(\Omega, \mathfrak{A})$ to the measurable space $(\mathbb{R}^D, \mathfrak{A}(\mathbb{R}^D))$ is measurable and hence a random variable with values in $\mathbb{R}^D$.

Definition 3.4

(a) Two real-valued random fields $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ and $\{b(\boldsymbol{y}), \boldsymbol{y} \in D\}$ on $D \subset \mathbb{R}^d$ are said to be independent if the associated $(\mathbb{R}^D, \mathfrak{A}(\mathbb{R}^D))$-valued random variables are independent.

(b) We call $f_i : D \to \mathbb{R}$, $i = 1, 2$, independent realizations of a real-valued random field a on $D \subset \mathbb{R}^d$ if $f_i(\boldsymbol{x}) = a_i(\boldsymbol{x}, \omega)$ for some $\omega \in \Omega$, where a_i are i.i.d. random fields with the same distribution as a.

Definition 3.5

For a real-valued random field $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ defined on $D \subset \mathbb{R}^d$ the probability distributions of all random vectors $(a(\boldsymbol{x}_1), \dots, a(\boldsymbol{x}_n))$ with $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in D$ on $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$ are known as the finite-dimensional distributions of a.

The Daniell-Kolmogorov theorem states consistency conditions for defining the distribution of a random field by a family of finite-dimensional probability measures. We denote by $\mathbf{P}_a$ the probability distribution of a random field a on the measurable space $(\mathbb{R}^D, \mathfrak{A}(\mathbb{R}^D))$.

Theorem 3.6 (Daniell & Kolmogorov)

Suppose that for each set $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\} \subset D$ there exists a probability measure $\mu_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_n}$ on $\mathbb{R}^n$ such that

(i) For any permutation σ of $\{1, \dots, n\}$ and any $B \in \mathfrak{B}(\mathbb{R}^n)$ there holds

$$\mu_{\boldsymbol{x}_{\sigma(1)},\dots,\boldsymbol{x}_{\sigma(n)}}(\sigma(B)) = \mu_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_n}(B),$$

where $\sigma(B) = \{(\boldsymbol{x}_{\sigma(1)}, \dots, \boldsymbol{x}_{\sigma(n)}) : (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in B\}$ and

(ii) for $m < n$ and any $B \in \mathfrak{B}(\mathbb{R}^m)$

$$\mu_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_n}(B \times \mathbb{R}^{n-m}) = \mu_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_m}(B).$$

Then there exists a random field $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ with finite-dimensional distributions $\mu_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_n}$. If $a(\boldsymbol{x})$ and $b(\boldsymbol{x})$ are two such random fields, then $\mathbf{P}_a(A) = \mathbf{P}_b(A)$ for any $A \in \mathfrak{A}(\mathbb{R}^D)$.

Definition 3.7

A random field a on $D \subset \mathbb{R}^d$ is said to be of second order if for all $\boldsymbol{x} \in D$ there holds $a(\boldsymbol{x}) = a(\boldsymbol{x}, \cdot) \in L^2(\Omega; \mathbb{R})$. We say a second-order random field a has mean function $\overline{a}(\boldsymbol{x}) := \mathbf{E}\left[a(\boldsymbol{x})\right]$ and covariance function

$$c(\boldsymbol{x}, \boldsymbol{y}) = c_a(\boldsymbol{x}, \boldsymbol{y}) := \mathbf{Cov}(a(\boldsymbol{x}), a(\boldsymbol{y})), \qquad \boldsymbol{x}, \boldsymbol{y} \in D.$$

Definition 3.8

A function $f : D \times D \to \mathbb{R}$ is called positive semidefinite if for any n-tuple $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in D^n$ and vector $\boldsymbol{z} = [z_1, \ldots, z_n]^\mathsf{T} \in \mathbb{R}^n$ there holds

$$\sum_{j,k=1}^{n} z_j z_k f(\boldsymbol{x}_j, \boldsymbol{x}_k) \geq 0.$$

Note: In the stochastics literature this property is often called simply *positive definite*.

Theorem 3.9

Let $D \subset \mathbb{R}^d$. The following statements are equivalent:

(a) There exists a real-valued second-order random field $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ with covariance function $c : D \times D \to \mathbb{R}$.

(b) $c \in \mathbb{R}^{D \times D}$ is symmetric and positive semidefinite.

Definition 3.10

A real-valued random field on $D \subset \mathbb{R}^d$ is called Gaussian if each random vector $[a(\boldsymbol{x}_1), \dots, a(\boldsymbol{x}_n)]$ follows an n-variate normal distribution for any $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in D$ and any $n \in \mathbb{N}$.

Corollary 3.11

The probability distribution $\mathbf{P}_a$ on $(\mathbb{R}^D, \mathfrak{A}(\mathbb{R}^D))$ of a real-valued Gaussian random field a is uniquely determined by its mean and covariance function.

Contents

Given a second-order random field a on $D \subset \mathbb{R}^d$ with mean $\overline{a}$, consider the centered random field $a - \overline{a}$. Given a CONS $\{\psi_m\}_{m\in\mathbb{N}}$ of $L^2(D)$, we have for each realization of a:

$$a(\cdot,\omega) - \overline{a} = \sum_{m=1}^{\infty} \xi_m(\omega)\psi_m,$$

where the ξ_m are random variables defined by

$$\xi_m(\omega) := (a(\cdot,\omega) - \overline{a}, \psi_m)_{L^2(D)}.$$

The Karhunen-Loève expansion of a results from choosing as a particular CONS the eigenfunctions of the covariance operator $C = C_a : L^2(D) \to L^2(D)$ of a, which is given by

$$u \mapsto Cu, \qquad (Cu)(\boldsymbol{x}) = \int_D u(\boldsymbol{y}) c(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}, \quad \boldsymbol{x} \in D. \tag{3.1}$$

Lemma 3.12

If $a \in L^2(\Omega; L^2(D))$, then $\overline{a} \in L^2(D)$ and $a(\cdot,\omega) \in L^2(D)$ **P**-a.s.

By Definition A.23, if $a \in L^2(\Omega; L^2(D))$ for $D \subset \mathbb{R}^d$, we have for any $\phi, \psi \in L^2(D)$ by Fubini's theorem

$$
\begin{aligned}
(C\phi, \psi)_{L^2(D)} &= \mathbf{Cov}\big((\phi, a)_{L^2(D)}, (\psi, a)_{L^2(D)}\big) \\
&= \mathbf{E}\left[\int_D \phi(\boldsymbol{x})[a(\boldsymbol{x}) - \overline{a}(\boldsymbol{x})]\,\mathrm{d}\boldsymbol{x} \int_D \psi(\boldsymbol{y})[a(\boldsymbol{y}) - \overline{a}(\boldsymbol{y})]\,\mathrm{d}\boldsymbol{y}\right] \\
&= \int_D \int_D \mathbf{Cov}(a(\boldsymbol{x}), a(\boldsymbol{y}))\,\phi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\;\psi(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},
\end{aligned}
$$

from which we infer that the covariance operator C_a of the $L^2(D)$-valued random variable a is the linear integral operator $C_a : L^2(D) \to L^2(D)$ with kernel function $c_a \in L^2(D \times D)$ given by $c_a(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{Cov}(a(\boldsymbol{x}), a(\boldsymbol{y}))$ (not pointwise, in general !)

Example 3.13

For $d = 1$ and $D = [-b, b]$, $b > 0$, the exponential covariance function is defined by

$$c(x, y) = e^{\frac{-|x-y|}{\ell}}, \qquad \ell > 0.$$

The eigenvalues of the associated covariance operator are given by

$$\lambda_m = \frac{2\ell}{\ell^2 \omega_m^2 + 1},\ (m \text{ even}), \qquad \lambda_m = \frac{2\ell}{\ell^2 \tilde{\omega}_m^2 + 1},\ (m \text{ odd})$$

where ω_m and $\tilde{\omega}_m$ denote the solutions of the transcendental equations

$$1 - \omega\ell \tan(\omega b) = 0 \quad \text{and} \quad \omega\ell + \tan(\omega b) = 0,$$

respectively. The associated eigenfunctions are given by

$$f_m(x) = \frac{\cos(\omega_m x)}{\sqrt{b + \frac{\sin(2\omega_m b)}{2\omega_m}}}, \qquad \tilde{f}_m(x) = \frac{\sin(\tilde{\omega}_m x)}{\sqrt{b - \frac{\sin(2\tilde{\omega}_m b)}{2\tilde{\omega}_m}}},$$

Theorem 3.14

If $(\lambda_m, a_m)_{m\in\mathbb{N}}$ denotes the sequence of eigenpairs (in descending order, $\|a_m\|_{L^2(D)} = 1$) of the covariance operator C_a associated with the random field $a \in L^2(\Omega; L^2(D))$ with mean function $\overline{a}(\boldsymbol{x})$, then

$$a(\boldsymbol{x}, \omega) = \overline{a}(\boldsymbol{x}) + \sum_{m=1}^{\infty} \sqrt{\lambda_m}\, a_m(\boldsymbol{x})\, \xi_m(\omega), \tag{3.2}$$

where the series converges in $L^2(\Omega; L^2(D))$, the random variables

$$\xi_m(\omega) = \frac{1}{\sqrt{\lambda_m}} (a(\cdot, \omega) - \overline{a}, a_m)_{L^2(D)}$$

have mean zero, unit variance and are pairwise uncorrelated.
If the random field is, in addition, Gaussian, then $\xi_m \sim \mathrm{N}(0, 1)$ are i.i.d.

- The KL expansion suggests a convenient approach for approximating a random field to a specified accuracy by truncation:

$$a(\boldsymbol{x},\omega) \approx a_M(\boldsymbol{x},\omega) := \bar{a}(\boldsymbol{x}) + \sum_{m=1}^{M} \sqrt{\lambda_m}\, a_m(\boldsymbol{x})\, \xi_m(\omega). \tag{3.3}$$

- The truncated RF a_M has the same mean as a and the covariance function

$$c(\boldsymbol{x},\boldsymbol{y}) \approx c_M(\boldsymbol{x},\boldsymbol{y}) = c_{a,M}(\boldsymbol{x},\boldsymbol{y}) = \sum_{m=1}^{M} \lambda_m a_m(\boldsymbol{x}) a_m(\boldsymbol{y}), \qquad \boldsymbol{x},\boldsymbol{y} \in D. \tag{3.4}$$

- Since $\{\phi_j(\boldsymbol{x})\phi_k(\boldsymbol{u})\}_{j,k\in\mathbb{N}}$ is a CONS for $L^2(D\times D)$, $c_M \to c$ in $L^2(D\times D)$.
- For a bounded domain D and continuous covariance function, its series expansion (3.4) and the MSE of the truncated KL series (3.3) both converge uniformly.

Theorem 3.15

Let a_M denote the truncated approximation defined in (3.3) of a real-valued random field $a \in L^2(\Omega; L^2(D))$ for a compact domain $D \subset \mathbb{R}^d$ with covariance function $c \in C(D \times D)$. Then the KL eigenfunctions $\{a_m\}_{m\in\mathbb{N}}$ are continuous on D and the series expansion of c converges uniformly, i.e.,

$$\sup_{\boldsymbol{x},\boldsymbol{y}\in D} |c(\boldsymbol{x},\boldsymbol{y}) - c_M(\boldsymbol{x},\boldsymbol{y})| \le \sup_{\boldsymbol{x}\in D} \sum_{m=M+1}^{\infty} \lambda_m a_m(\boldsymbol{x})^2 \to 0, \qquad M \to \infty. \tag{3.5}$$

In addition,

$$\sup_{\boldsymbol{x}\in D} \mathbf{E}\left[(a(\boldsymbol{x}) - a_M(\boldsymbol{x}))^2\right] \to 0 \qquad M \to \infty.$$

- For the variance of the truncated KL expansion, we have

$$\mathbf{Var}(a(\boldsymbol{x})) - \mathbf{Var}(a_M(\boldsymbol{x})) = \mathbf{E}\left[(a(\boldsymbol{x}) - a_m(\boldsymbol{x}))^2\right] = \sum_{m=M+1}^{\infty} \lambda_m a_m(\boldsymbol{x})^2 \geq 0,$$

  hence a_M always underestimates the variance of a.
- Viewed as a random variable $a \in L^2(\Omega; L^2(D))$, we have for the truncation error

$$\|a - a_M\|^2_{L^2(\Omega;L^2(D))} = \sum_{m=M+1}^{\infty} \lambda_m.$$

- In addition,

$$\begin{aligned} \|a - a_M\|^2_{L^2(\Omega;L^2(D))} &= \mathbf{E}\left[\|a - \bar{a}\|^2_{L^2(D)} - \|a_M - \bar{a}\|^2_{L^2(D)}\right] \\ &= \int_D \mathbf{Var}\, a(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum_{m=1}^{M} \lambda_m. \end{aligned}$$

- This allows an assessment of the truncation error w.r.t. $\|\cdot\|_{L^2(\Omega;L^2(D))}$ provided the first M eigenvalues can be calculated as well as the integral of **Var** a over D.
- Eigenvalue approximations can be obtained by solving the covariance eigenproblem numerically. If **Var** $a \equiv \sigma^2$ on D, this yields

$$\|a - a_M\|^2_{L^2(\Omega;L^2(D))} = \sigma^2|D| - \sum_{m=1}^{M} \lambda_m,$$

where $|D| = \int_D \mathrm{d}\boldsymbol{x}$ is the Lebesgue measure of D.
In this case the error can always be estimated once a number of leading eigenvalues are available.

Contents

Regularity of Random Fields

Definition 3.16

A random field $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ is said to be mean-square continuous if, for all $\boldsymbol{x} \in D$,

$$\|a(\boldsymbol{x}+\boldsymbol{h}) - a(\boldsymbol{x})\|_{L^2(\Omega)} = \mathbf{E}\left[(a(\boldsymbol{x}+\boldsymbol{h}) - a(\boldsymbol{x}))^2\right] \to 0 \quad \text{as } \boldsymbol{h} \to \boldsymbol{0}.$$

We assume centered random fields a in the remainder of this section, i.e., $\overline{a} \equiv 0$.

Theorem 3.17

Let $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ be a centered random field. Then its covariance function c is continuous at $(\boldsymbol{x}, \boldsymbol{x})$, $\boldsymbol{x} \in D$, if and only if $\mathbf{E}\left[(a(\boldsymbol{x}+\boldsymbol{h}) - a(\boldsymbol{x}))^2\right] \to 0$ as $\boldsymbol{h} \to \boldsymbol{0}$. In particular, if $c \in C(D \times D)$, then a is mean-square continuous.

Corollary 3.18

Let $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ be a centered random field. If its covariance function is continuous along the 'diagonal' $\{(\boldsymbol{x}, \boldsymbol{x}) : \boldsymbol{x} \in D\}$, then it is continuous throughout $D \times D$.

Theorem 3.19

Let $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ be a centered second-order random field. If its covariance function $c \in C^2(D \times D)$, then a is mean-square differentiable, i.e., there exists a random field $\{\partial_{x_j} a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ such that for all $j = 1, 2, \dots, d$,

$$\left\| \frac{a(\boldsymbol{x} + h\boldsymbol{e}_j) - a(\boldsymbol{x})}{h} - \partial_{x_j} a(\boldsymbol{x}) \right\|_{L^2(\Omega)} \to 0 \quad \text{as } h \to 0$$

and $\partial_{x_j} a(\boldsymbol{x})$ has covariance function

$$c_j(\boldsymbol{x}, \boldsymbol{y}) = \frac{\partial^2 c(\boldsymbol{x}, \boldsymbol{y})}{\partial x_j \partial y_j}.$$

Analogous relations hold for higher order mean-square derivatives of a random field given higher order differentiability of the covariance function.

- Mean-square continuity and differentiability depend on the expectation, i.e., on an average over all realizations.
- A related issue is the regularity of each individual realization.
- Even for Gaussian fields, it is not possible to show that each realization is continuous.
- Even though the distribution of a Gaussian random field is uniquely defined on $\mathfrak{A}(\mathbb{R}^D)$, realization-wise continuity cannot hold in general.
- $\mathfrak{A}(\mathbb{R}^D)$ is contructed from a countable set of conditions, whereas statements about the continuity of functions involve conditions on a continuum of points, i.e., uncountably many conditions.
- However, given a condition of the moments of the 'increments' $a(\boldsymbol{x}) - a(\boldsymbol{y})$, a version of $a(\boldsymbol{x})$ with continuous realizations can be shown to exist.

Theorem 3.20

Let D be a bounded domain in $\mathbb{R}^d$ and $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ be a centered Gaussian random field such that, for some $L, s > 0$

$$\mathbf{E}\left[|a(x) - a(y)|^2\right] \le L\|\boldsymbol{x} - \boldsymbol{y}\|_2^s \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \overline{D}.$$

Then for any $p \ge 1$ there exists a random variable K such that $e^K \in L^p(\Omega)$ and

$$|a(\boldsymbol{x}) - a(\boldsymbol{y})| \le K(\omega)\|\boldsymbol{x} - \boldsymbol{y}\|_2^{(s-\epsilon)/2} \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \overline{D} \text{ a.s.},$$

i.e., realizations of a are Hölder continuous with exponent $s/2$.

Definition 3.21

A random field $\{b(\boldsymbol{x}), \boldsymbol{x} \in D\}$ is called a version of a random field $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ if

$$\mathbf{P}\left(a(\boldsymbol{x}) = b(\boldsymbol{x})\right) = 1 \qquad \forall \boldsymbol{x} \in D.$$

A random field $\{a(\boldsymbol{x}), \boldsymbol{x} \in D\}$ is said to have a continuous version if there exists a version of a with continuous realizations.

Theorem 3.22 (cf. [Kallenberg (1997)], Thm. 2.23)

Let a be a random field on $D \subset \mathbb{R}^d$ with values in a Banach space and assume for some $a, b > 0$ that

$$\mathbf{E}\left[\|a(\boldsymbol{x}) - a(\boldsymbol{y})\|^a\right] \lesssim \|\boldsymbol{x} - \boldsymbol{y}\|^{d+b}, \qquad \boldsymbol{x}, \boldsymbol{y} \in D.$$

Then a has a continuous version, and for any $c \in (0, b/a)$ the latter is a.s. locally Hölder continuous with exponent c.

Contents

Definition 3.23

(a) A random field a is strictly stationary or homogeneous if its finite-dimensional distributions are invariant under translation, i.e., if the multivariate distribution of $(a(\boldsymbol{x}_1), \dots, a(\boldsymbol{x}_n))$ is the same as that of $(a(\boldsymbol{x}_1 + \boldsymbol{h}), \dots, a(\boldsymbol{x}_n + \boldsymbol{h}))$, for all $\boldsymbol{h}$.

(b) A random field a is (wide-sense) stationary or (wide-sense) homogeneous if its mean is constant and its covariance function satisfies $c(\boldsymbol{x}, \boldsymbol{y}) = c(\boldsymbol{x} - \boldsymbol{y})$. Such a covariance function is known as a stationary covariance.

Example 3.24

The separable exponential covariance function is given by

$$c(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j=1}^{d} e^{\frac{-|x_j - y_j|}{\ell_j}}, \qquad \ell_j > 0,$$

where ℓ_j is a correlation length parameter in the j-th Cartesian direction, is an example of a stationary covariance function.

Theorem 3.25 (Wiener-Khintchine)

The following two statements are equivalent:

(a) There exists a mean-square continuous stationary random field $\{a(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^d\}$ with stationary covariance function c.

(b) The function $c : \mathbb{R}^d \to \mathbb{R}$ is such that

$$c(\boldsymbol{x}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\lambda}\cdot\boldsymbol{x}} \,\mathrm{d}F(\boldsymbol{\lambda})$$

for some measure F on $\mathbb{R}^d$ with $F(\mathbb{R}^d) < \infty$,

The measure F is called the spectral distribution. If it exists, the density f of F is called the spectral density. Alternatively, given $c : \mathbb{R}^d \to \mathbb{R}$, we may compute

$$f(\boldsymbol{\lambda}) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\lambda}\cdot\boldsymbol{x}} c(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.$$

If f is nonnegative and integrable then c is a valid covariance function.

Example 3.26 (Separable exponential covariance)

The Fourier transform of the separable covariance function is obtained as the rpoduct of the transforms of its factors, i.e.,

$$f(\boldsymbol{\lambda}) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\lambda}\cdot\boldsymbol{x}} c(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \prod_{j=1}^{d} (2\pi)^{-1} \int_{\mathbb{R}} e^{-ix_j\lambda_j} e^{-|x_j|/\ell_j}\,\mathrm{d}x_j,$$

yielding

$$f(\boldsymbol{\lambda}) = \prod_{j=1}^{d} \frac{\ell_j}{\pi(\lambda_j^2 + \ell_j^2)}.$$

Since $\ell_j > 0$ for all j, f is nonnegative and is the density of a measure F with $F(\mathbb{R}^d) < \infty$. By the Wiener-Khintchine theorem c is thus the covariance kernel for some mean-square continuous random field.

Example 3.27 (Gaussian covariance)

For a symmetric positive definite matrix $A \in \mathbb{R}^{d\times d}$ the function

$$c(\boldsymbol{x}) = e^{-\boldsymbol{x}^\mathsf{T} A \boldsymbol{x}}, \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

has the Fourier transform

$$f(\boldsymbol{\lambda}) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{x}\cdot\boldsymbol{\lambda}} e^{-\boldsymbol{x}^\mathsf{T} A \boldsymbol{x}} \,\mathrm{d}\boldsymbol{x} = \frac{1}{(2\pi)^{d/2} 2^{d/2} \sqrt{\det A}} = e^{-\boldsymbol{\lambda}^\mathsf{T} A^{-1} \boldsymbol{\lambda}/4}.$$

f is nonnegative and is the density of a measure F – in fact the Gaussian distribution $\mathrm{N}(\mathbf{0}, 2A)$. Again, the Wiener-Khintchine theorem asserts that c is the covariance function of a random field.

Definition 3.28

A stationary random field $\{a(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^d\}$ is said to be isotropic if its covariance function is invariant under rotations, i.e.,

$$c(\boldsymbol{x}, \boldsymbol{y}) = c(r), \qquad r = \|\boldsymbol{x} - \boldsymbol{y}\|_2.$$

Example 3.29 (Isotropic Gaussian covariance)

A simple example of an isotropic covariance function is $c(r) = e^{-r^2}$, arising from a Gaussian covariance with $A = I_d$. In fact, $c(\boldsymbol{x}) = e^{-\boldsymbol{x}^\mathsf{T} A \boldsymbol{x}}$ is isotropic whenever $A = \sigma I_d$ for some $\sigma > 0$.

Example 3.30 (Bessel covariance)

Another isotropic covariance function, proposed by Whittle as a generalization of the exponential covariance to higher dimensions, is given by $c(r) = rK_1(r)$, where K_1 is the modified Bessel function of second kind with index 1.

For isotropic functions the Fourier transform in the Wiener-Khintchine theorem becomes a Hankel transform (cf. Theorem D.1). For $f(s) = f(\|\boldsymbol{\lambda}\|_2)$, we obtain for $d = 1, 2, 3$

$$c(r) = \begin{cases} 2\int_0^\infty \cos(rs)\, f(s)\,\mathrm{d}s, & d = 1, \\ 2\pi \int_0^\infty J_0(rs) f(s) s\,\mathrm{d}s, & d = 2, \\ 4\pi \int_0^\infty \frac{1}{rs}\sin(rs) f(s) s^2\,\mathrm{d}s, & d = 3. \end{cases}$$

Theorem 3.31

Let $\{a(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^d\}$ be an isotropic random field with mean-square continuous covariance function c. There exists a finite measure F on $\mathbb{R}^+$ known as the radial spectral distribution such that

$$c(r) = \Gamma\left(\tfrac{d}{2}\right) \int_0^\infty \frac{J_\nu(rs)}{\left(\frac{rs}{2}\right)^\nu} \, \mathrm{d}F(s), \qquad \nu = \tfrac{d}{2} - 1.$$

If the spectral density exists, $f(s) = f(\boldsymbol{\lambda})$ for $s = \|\boldsymbol{\lambda}\|_2$ is called the radial spectral density function. Then

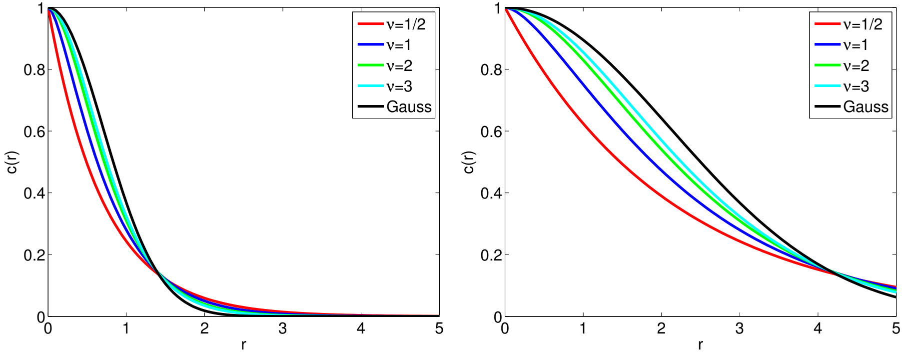
$$\mathrm{d}F(s) = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} s^{d-1} f(s) \, \mathrm{d}s \quad \text{and} \quad f(s) = (2\pi)^{-d/2} \int_0^\infty \frac{J_\nu(rs)}{(rs)^\nu} c(r) r^{d-1} \, \mathrm{d}r.$$

The Matérn class is a family of isotropic covariance functions named after the Swedish forestry statistician Bertil Matérn and is very popular in geostatistics as well as machine learning etc.
The covariance function is given by

$$c(r) = \frac{\sigma^2}{2^{\nu-1}\,\Gamma(\nu)} \left(\frac{2\sqrt{\nu}\,r}{\rho}\right)^{\nu} K_\nu\left(\frac{2\sqrt{\nu}\,r}{\rho}\right), \qquad r = \|\boldsymbol{x} - \boldsymbol{y}\|_2, \tag{3.6}$$

where

K_ν	is the modified (second-kind) Bessel function of order ν,
ν	is known as the smoothness parameter,
σ^2	is the variance parameter,
ρ	is the correlation length parameter,
Γ	denotes the Gamma-function.



$\rho = 1$ $\rho = 3$

- With smaller correlation length ρ the Matérn covariance function becomes more strongly concentrated near $r = 0$.
- With increasing values of the smoothness parameter ν the Matérn covariance function becomes smoother at $r = 0$. (It is analytic everywhere else.)

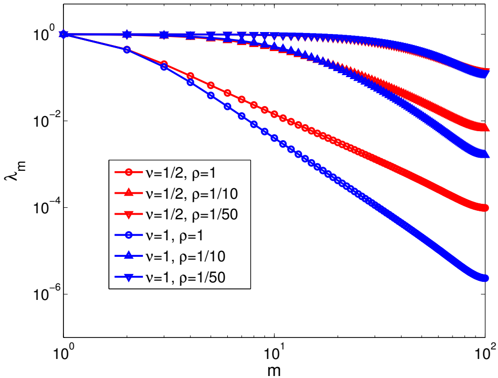
The Matérn family has a number of attractive features:

- It contains the exponential, Bessel and Gaussian covariance functions as special cases:

$\nu = \frac{1}{2}$:	$c(r) = \sigma^2 \exp(-\sqrt{2}r/\rho)$	exponential covariance
$\nu = 1$:	$c(r) = \sigma^2 \left(\frac{2r}{\rho}\right) K_1 \left(\frac{2r}{\rho}\right)$	Bessel covariance
$\nu \to \infty$:	$c(r) = \sigma^2 \exp(-r^2/\rho^2)$	Gaussian covariance

- Smoothness of realizations: a random field with Matérn covariance function is s times mean-square differentiable if and only if $\nu > s$.
- The flexibility of the parametrization allows its application to many statistical situation, the parameters may be estimated from observed data using statistical techniques.

Before asymptotic decay sets in (determined by the smoothness of the kernel), there is a preasymptotic plateau whose length is determined by the correlation length parameter ρ.



Eigenvalue decay, Matérn covariance kernel, $D = [-1, 1]$.

In a paper published in 1963[4], Harold Widom analyzed linear integral operators of the form

$$u \mapsto Ku, \qquad (Ku)(\boldsymbol{x}) = \int_{\mathbb{R}^d} V(\boldsymbol{x})^{1/2} k(\boldsymbol{x} - \boldsymbol{y}) V(\boldsymbol{y})^{1/2} u(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}. \tag{3.7}$$

We obtain the covariance operator for an isotropic covariance function on a bounded domain $D \subset \mathbb{R}^d$ by setting $V(\boldsymbol{x}) = \mathbb{1}_D$ and $k(\boldsymbol{x} - \boldsymbol{y}) = c(\|\boldsymbol{x} - \boldsymbol{y}\|_2)$.

Definition 3.32

Two functions $f : E \to \mathbb{R}$ and $g : F \to \mathbb{R}$, $E \subset \mathbb{R}^n$ and $F \subset \mathbb{R}^m$ are said to be equimeasurable if, for all $t \in \mathbb{R}$,

$$|\{\boldsymbol{x} \in E : f(\boldsymbol{x}) > t\}| = |\{\boldsymbol{y} \in F : g(\boldsymbol{y}) > t\}|$$

where $|\cdot|$ denotes Lebesgue measure.

[4] Widom, H., Asymptotic behavior of the eigenvalues of certain integral equations. *Trans. Amer. Math. Soc.* 109, 278–295 (1963).

We denote the spectral density (Fourier transform) of $c = c(\boldsymbol{x}) = c(\|\boldsymbol{x}\|_2)$ by

$$\hat{c}(\boldsymbol{\lambda}) = (2\pi)^{-d} \int_{\mathbb{R}^d} c(\boldsymbol{x}) e^{-i\boldsymbol{x}\cdot\boldsymbol{\lambda}} \,\mathrm{d}\boldsymbol{x} = f(s)$$

and set $K(\boldsymbol{\lambda}) := (2\pi)^d \, \hat{c}(\boldsymbol{\lambda})$.

Theorem 3.33 (Widom, 1963)

For the integral operator K in (3.7) let V be a bounded, nonnegative function with bounded support, let k be integrable over $\mathbb{R}^d$ with an ultimately positive Fourier transform and let $\{\lambda_m\}_{m\in\mathbb{N}}$ denote its (nonincreasing) sequence of eigenvalues. If the function $\phi_0 : \mathbb{R}_0^+ \to \mathbb{R}$ is equimeasurable to $V(\boldsymbol{x})K(\boldsymbol{\lambda}) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, then

$$\lambda_m \asymp \phi_0((2\pi)^d m) \qquad \text{as } m \to \infty.$$

Note: The theorem still holds when the integral operator leads to a $K'(\boldsymbol{\lambda})$ such that

$$K(\boldsymbol{\lambda}) \asymp K'(\boldsymbol{\lambda}), \qquad \text{as } \|\boldsymbol{\lambda}\|_2 \to \infty.$$

Corollary 3.34

Let $c = c(r)$ be an isotropic covariance function on $\mathbb{R}^d$ with radial spectral density $f = f(s)$. Assume that $f(s) \asymp bs^{-\rho}$ as $s \to \infty$, for some $b, \rho > 0$. Let D be a bounded domain in $\mathbb{R}^d$ and let $\{\lambda_m\}_{m\in\mathbb{N}}$ denote the (nonincreasing) eigenvalues of the covariance operator C given by (3.1). Then

$$\lambda_m \asymp K(D, d, \rho, b)\, m^{-\rho/d}, \qquad m \to \infty,$$

with $K(D, d, \rho, b) := (2\pi)^{d-\rho} b(|D|V_d)^{\rho/d}$, where $V_d = \frac{2\pi^{d/2}}{d\,\Gamma(d/2)}$ denotes the volume of the unit sphere in $\mathbb{R}^d$.

Corollary 3.35

If the spectral density $f(s)$ of an isotropic random field satisfies $f(s) \asymp bs^{-\rho}$, then

$$\lambda_m \asymp Km^{-\rho/d}, \qquad m \to \infty,$$

with $K = (2\pi)^{d-\rho} b\,(|D|V_d)^{\rho/d}$ where V_d denotes the volume of the d-dimensional unit sphere.

The Fourier transform of the Matérn covariance function with smoothness parameter ν, variance σ^2 and correlation length parameter ℓ is given by

$$f(s;\nu,\sigma,\ell) = \sigma^2 \pi^{-d/2} \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} \frac{\alpha^{2\nu}}{(s^2+\alpha^2)^{\nu+d/2}}.$$

where $\alpha := 2\sqrt{\nu}/\ell$.

Corollary 3.36

For the Matérn covariance in d dimensions with smoothness parameter ν the covariance eigenvalues decay asymptotically like

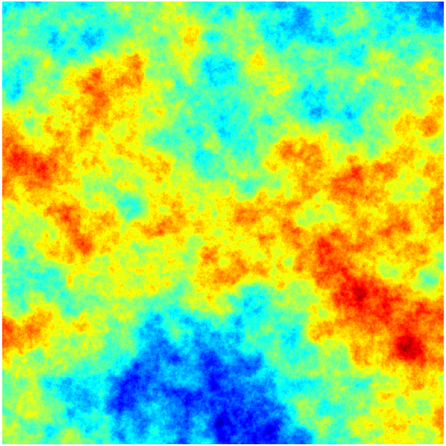
$$\lambda_m \asymp K m^{-(1+2\nu/d)}, \qquad m \to \infty,$$

where

$$K = (2\pi)^{-2\nu-d/2} \sigma^2 \frac{\Gamma(\nu+d/2)}{\Gamma(\nu)} \left(\frac{2\sqrt{\nu}}{\ell} \right)^{2\nu} (|D| V_d)^{1+2\nu/d}.$$

Covariance Eigenvalue Decay

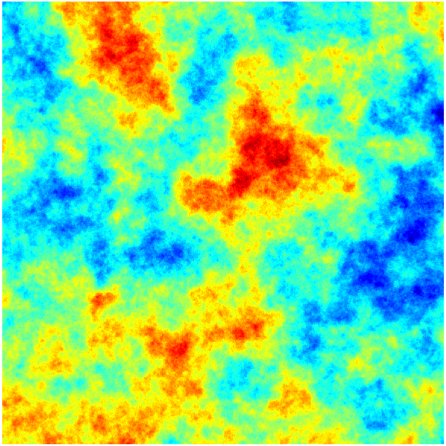
Realizations of Gaussian RF



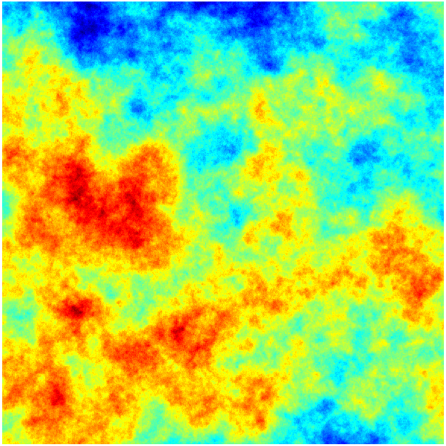
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.5$

Covariance Eigenvalue Decay

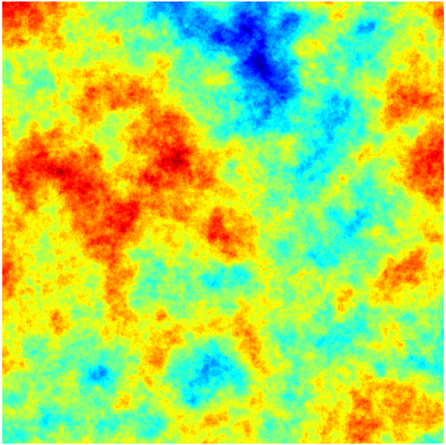
Realizations of Gaussian RF



Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.5$



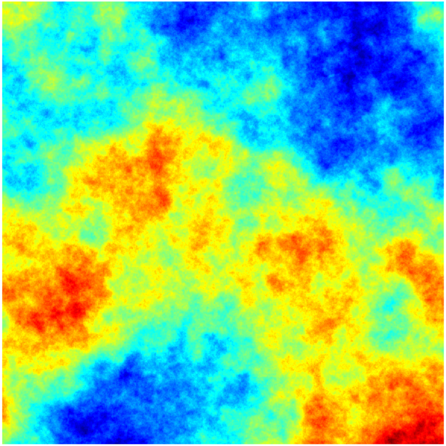
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.5$



Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.5$

Covariance Eigenvalue Decay

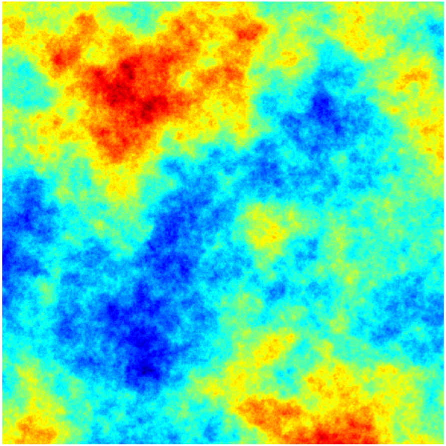
Realizations of Gaussian RF



Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.5$

Covariance Eigenvalue Decay

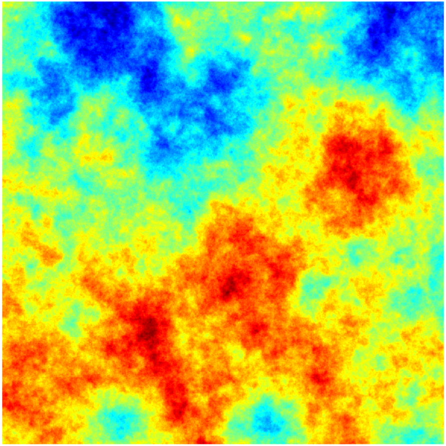
Realizations of Gaussian RF



Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.5$

Covariance Eigenvalue Decay

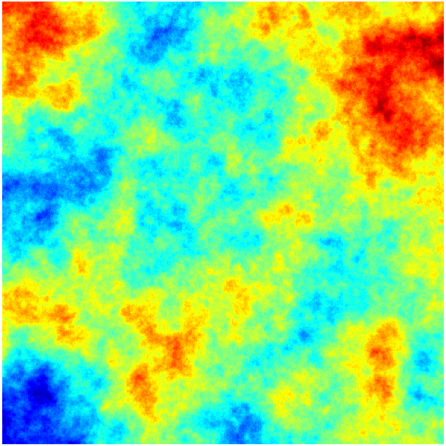
Realizations of Gaussian RF



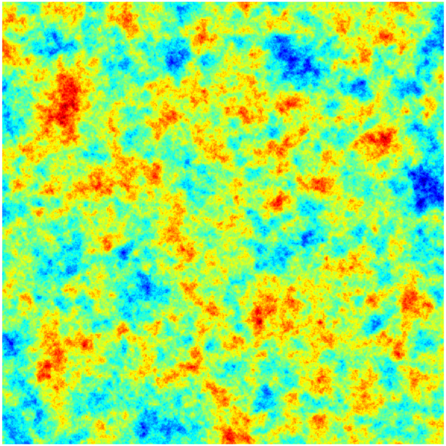
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.5$

Covariance Eigenvalue Decay

Realizations of Gaussian RF



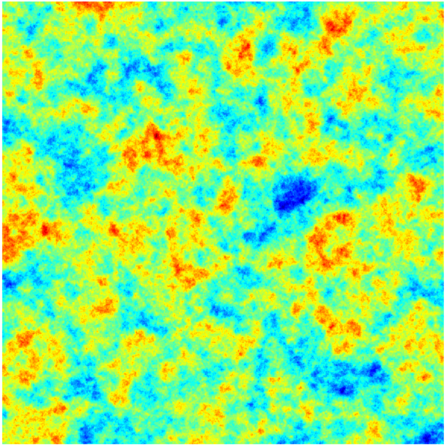
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.5$



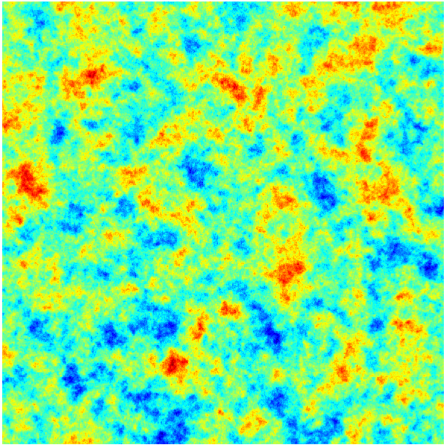
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.05$

Covariance Eigenvalue Decay

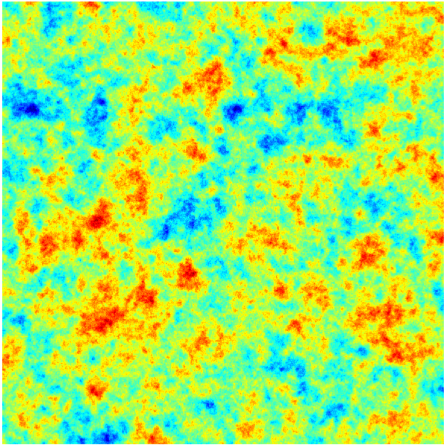
Realizations of Gaussian RF



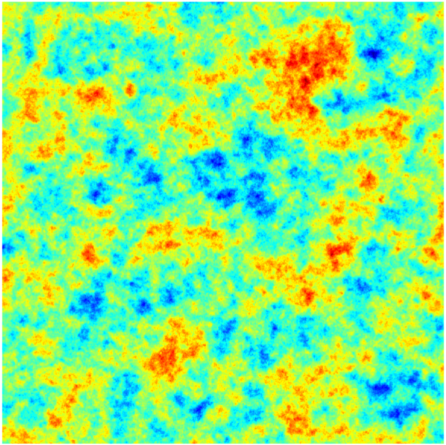
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.05$



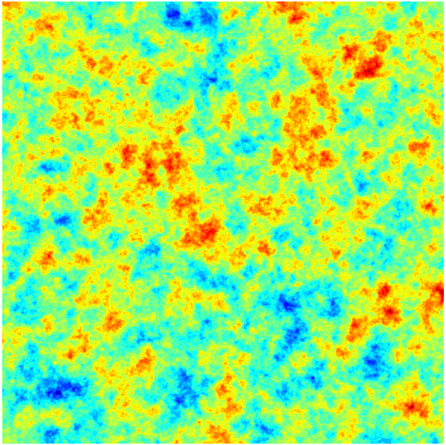
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.05$



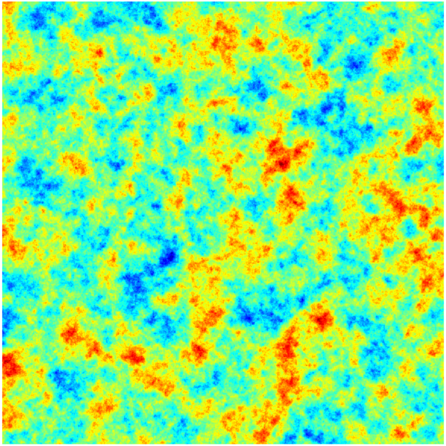
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.05$



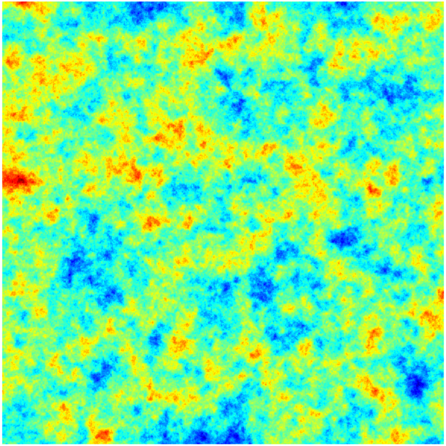
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.05$



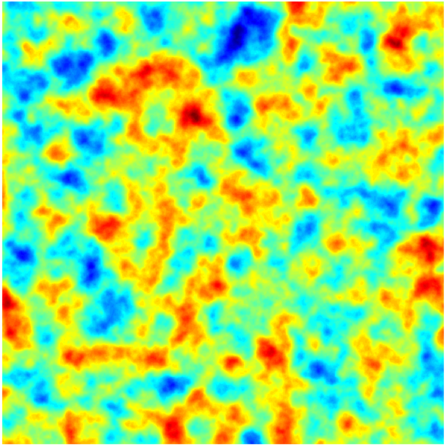
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.05$



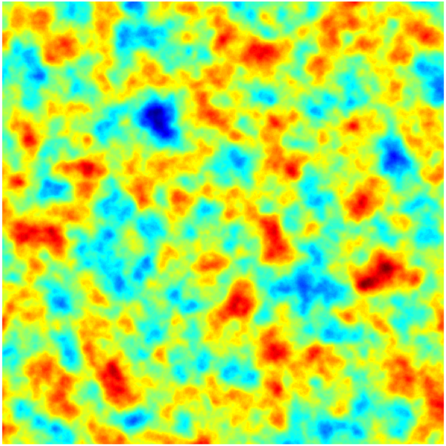
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.05$



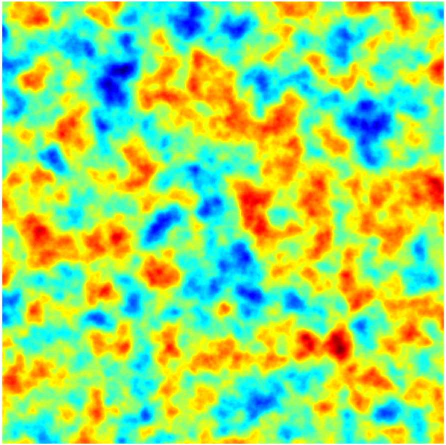
Matérn covariance, $\sigma = 1$, $\nu = \frac{1}{2}$, $\ell = 0.05$



Matérn covariance, $\sigma = 1$, $\nu = \frac{3}{2}$, $\ell = 0.05$



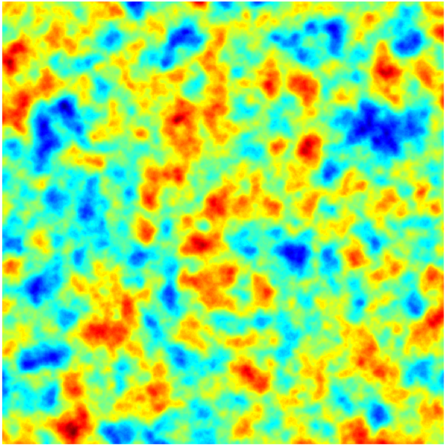
Matérn covariance, $\sigma = 1$, $\nu = \frac{3}{2}$, $\ell = 0.05$



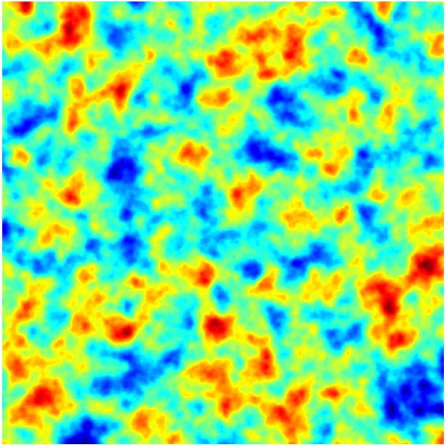
Matérn covariance, $\sigma = 1$, $\nu = \frac{3}{2}$, $\ell = 0.05$

Covariance Eigenvalue Decay

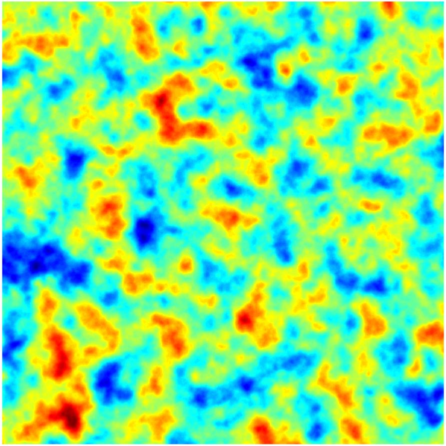
Realizations of Gaussian RF



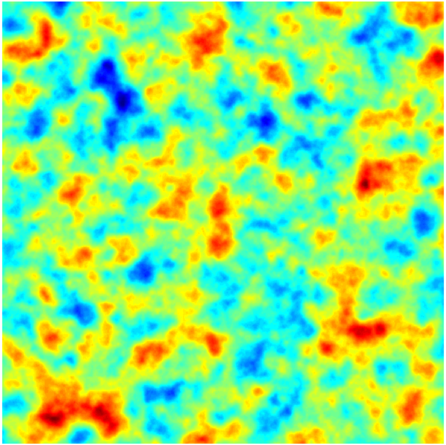
Matérn covariance, $\sigma = 1$, $\nu = \frac{3}{2}$, $\ell = 0.05$



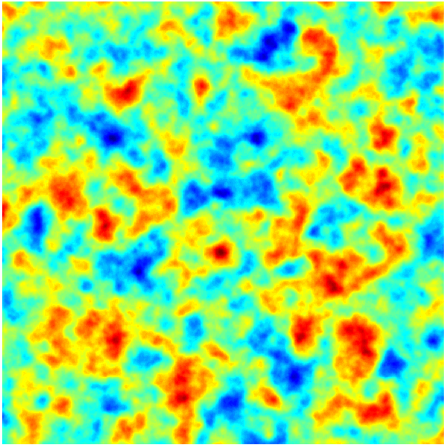
Matérn covariance, $\sigma = 1$, $\nu = \frac{3}{2}$, $\ell = 0.05$



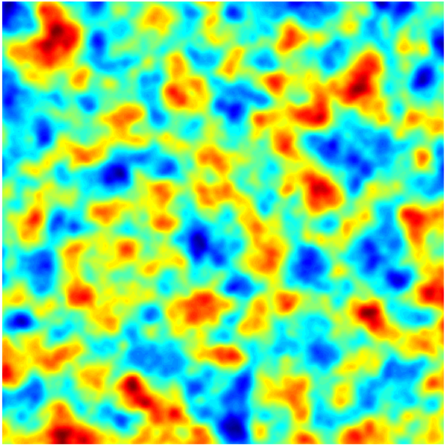
Matérn covariance, $\sigma = 1$, $\nu = \frac{3}{2}$, $\ell = 0.05$



Matérn covariance, $\sigma = 1$, $\nu = \frac{3}{2}$, $\ell = 0.05$



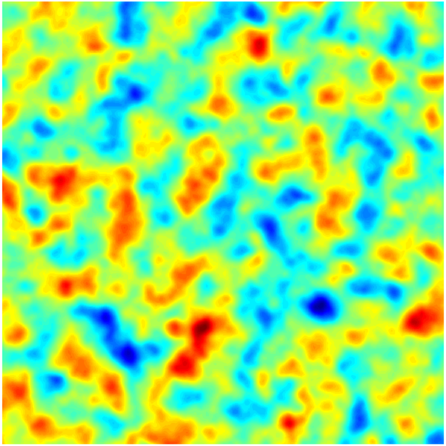
Matérn covariance, $\sigma = 1$, $\nu = \frac{3}{2}$, $\ell = 0.05$



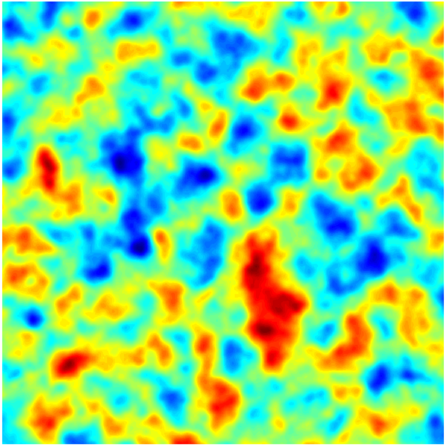
Matérn covariance, $\sigma = 1$, $\nu = \frac{5}{2}$, $\ell = 0.05$

Covariance Eigenvalue Decay

Realizations of Gaussian RF

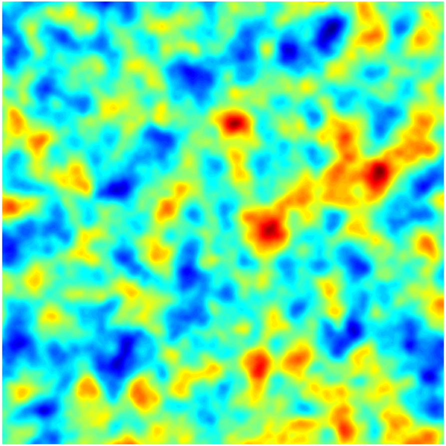


Matérn covariance, $\sigma = 1$, $\nu = \frac{5}{2}$, $\ell = 0.05$

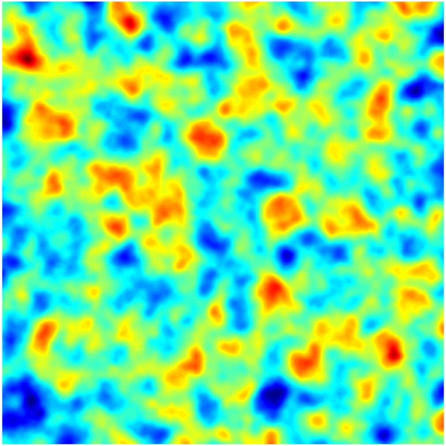


Matérn covariance, $\sigma = 1$, $\nu = \frac{5}{2}$, $\ell = 0.05$

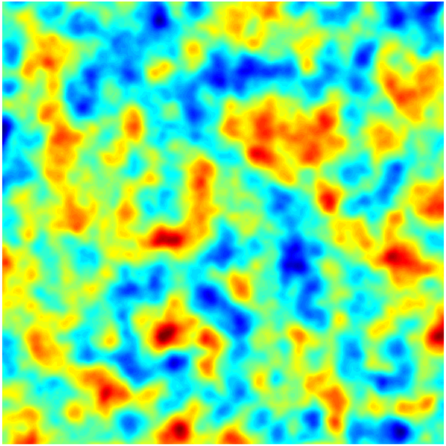
Covariance Eigenvalue Decay



Matérn covariance, $\sigma = 1$, $\nu = \frac{5}{2}$, $\ell = 0.05$



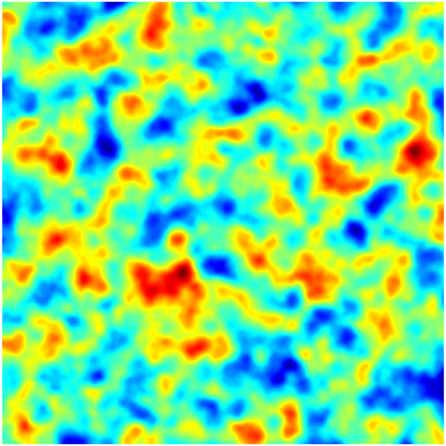
Matérn covariance, $\sigma = 1$, $\nu = \frac{5}{2}$, $\ell = 0.05$



Matérn covariance, $\sigma = 1$, $\nu = \frac{5}{2}$, $\ell = 0.05$

Covariance Eigenvalue Decay

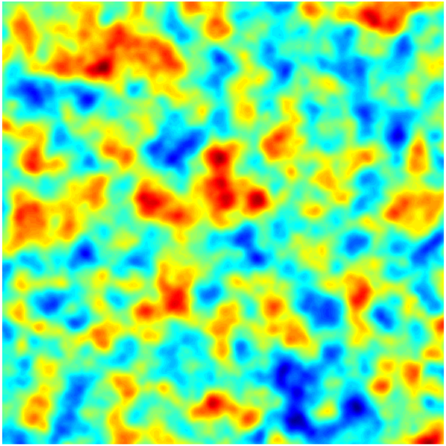
Realizations of Gaussian RF



Matérn covariance, $\sigma = 1$, $\nu = \frac{5}{2}$, $\ell = 0.05$

Covariance Eigenvalue Decay

Realizations of Gaussian RF



Matérn covariance, $\sigma = 1$, $\nu = \frac{5}{2}$, $\ell = 0.05$