Spatio-temporal statistics (MATH4341) Michaelmas term, 2024

Lecture notes part 2: Point referenced data modeling / Geostatistics

Lecturer & author: Georgios P. Karagiannis georgios.karagiannis@durham.ac.uk

Aim. To introduce point referenced data modeling (geostatistics) with particular focus on concepts spatial variables, random fields, semi-variogram, kriging, change of support, multivariate geostatistics, for Bayesian and classical inference.

Reading list & references:

[1] Cressie, N. (2015; Part I). Statistics for spatial data. John Wiley & Sons.
[2] Kent, J. T., & Mardia, K. V. (2022). Spatial analysis (Vol. 72). John Wiley & Sons.
[3] Chiles, J. P., & Delfiner, P. (2012). Geostatistics: modeling spatial uncertainty (Vol. 713). John Wiley & Sons.
[4] Wackernagel, H. (2003). Multivariate geostatistics: an introduction with applications. Springer Science & Business Media.
[5] Gaetan, C., & Guyon, X. (2010; Ch 2 & 5.1). Spatial statistics and modeling (Vol. 90). New York: Springer.

Part 1. Basic stochastic models & related concepts for model building

Note 1. We discuss basic stochastic models and concepts for modeling point referenced data in the Geostatistics framework.

1. Random fields (or Stochastic processes)

Definition 2. A random field (or stochastic process, or random function) $Z = (Z(s); s \in \mathcal{S})$ taking values in $\mathcal{Z} \subseteq \mathbb{R}^q$, $q \geq 1$ is a family of random variables $\{Z(s) := Z(s;\omega); s \in \mathcal{S}, \omega \in \Omega\}$ defined on the same probability space $(\Omega, \mathfrak{F}, \mathrm{pr})$ and taking values in $\mathcal{Z}$. The label $s \in \mathcal{S}$ is called site, the set $\mathcal{S} \subseteq \mathbb{R}^d$ is called the (spatial) set of sites at which the random field is defined, and $\mathcal{Z}$ is called the state space of the field.

Note 3. Given a set of sites $\{s_1, ..., s_n\}$, with $s_i \in S$ and $n \in \mathbb{N}$, the random vector $(Z(s_1), ..., Z(s_n))^\top$ has a well-defined probability distribution that is completely determined by its joint CDF

$$\mathrm{F}_{s_1,...,s_n}(z_1, ..., z_n) := \mathrm{pr}(Z(s_1) \leq z_1, ..., Z(s_n) \leq z_n)$$

The family of all finite-dimensional distributions (or fidi's) of Z is called the spatial distribution of the process .

Note 4. According to Kolmogorov Theorem 5, to define a random field model, one must specify the joint distribution of $(Z(s_1), ..., Z(s_n))^\top$ for all of n and all $\{s_i \in S; i = 1, ..., n\}$ in a consistent way.

Proposition 5. *(Kolmogorov consistency theorem) Let $pr_{s_1,...,s_n}$ be a probability on $\mathbb{R}^n$ with join CDF $F_{s_1,...,s_n}$ for every finite collection of points $s_1, ..., s_n$. If $F_{s_1,...,s_n}$ is symmetric w.r.t. any permutation $\mathfrak{p}$*

$$F_{\mathfrak{p}(s_1),...,\mathfrak{p}(s_n)}\left(z_{\mathfrak{p}(1)}, ..., z_{\mathfrak{p}(n)}\right) = F_{s_1,...,s_n}(z_1, ..., z_n)$$

for all $n \in \mathbb{N}$, $\{s_i \in S\}$, and $\{z \in \mathbb{R}\}$, and all if all permulations $\mathfrak{p}$ are consistent in the sense

$$\lim_{z_n \to \infty} F_{s_1,...,s_n}(z_1, ..., z_n) = F_{s_1,...,s_{n-1}}(z_1, ..., z_{n-1})$$

or all $n \in \mathbb{N}$, $\{s_i \in S\}$, and $\{z_i \in \mathbb{R}\}$, then there exists a random field Z whose fidi's coincide with those in F.

Example 6. Let $n \in \mathbb{N}$, let $\{X_i : T \to \mathbb{R};\ i = 1, ..., n\}$ be a set of constant functions, and let $\{Z_i \sim \mathrm{N}(0, 1)\}_{i=1}^n$ be a set of independent random variables. Then

$$\tilde{Z}(s) = \sum_{i=1}^n Z_i X_i(s), \quad s \in S \tag{1.1}$$

is a well defined random field as it satisfies Theorem 5.

1.1. **Mean and covariance functions.**

Definition 7. The mean function $\mu(\cdot)$ and covariance function $c(\cdot, \cdot)$ of a random field $(Z(s); s \in \mathcal{S})$ are defined as

$$\mu(s) = \mathrm{E}(Z(s)), \qquad \forall s \in S \tag{1.2}$$

$$c(s, s') = \mathrm{Cov}(Z(s), Z(s')) = \mathrm{E}\left((Z(s) - \mu(s))(Z(s') - \mu(s'))^\top\right), \quad \forall s, s' \in S \tag{1.3}$$

Example 8. For (1.1), the mean function is $\mu(s) = \mathrm{E}\left(\tilde{Z}_s\right) = 0$ and covariance function is

$$\begin{aligned} c(s, s') &= \mathrm{Cov}(Z(s), Z(s')) = \mathrm{Cov}\left(\sum_{i=1}^n Z_i X_i(s), \sum_{j=1}^n Z_j X_j(s')\right) \\ &= \sum_{i=1}^n X_i(s) \sum_{j=1}^n X_j(s') \overset{1(i=j)}{\mathrm{Cov}(Z_i, Z_j)} = \sum_{i=1}^n X_i(s) X_i(s') \end{aligned}$$

1.1.1. *Construction of covariance functions.*

Note 9. What follows provides the means for checking and constructing covariance functions.

Proposition 10. *The function* $c : S \times S \to \mathbb{R}$, $S \subseteq \mathbb{R}^n$ is a covariance function iff $c(\cdot, \cdot)$ is semi-positive definite; i.e.

$$\forall n \in \mathbb{N}, \, \forall a \in \mathbb{R}^n \text{ and} \, \forall (s_1, ..., s_n) \in \mathbb{R}^n : \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j c(s_i, s_j) \geq 0$$

or in other words, the Gram matrix $(c(s_i, s_j))_{i,j=1}^{n}$ is non-negative definite for any $\{s_i\}_{i=1}^{n}$, $n \in \mathbb{N}$.

Example 11. $c(s, s') = 1(\{s = s'\})$ is a proper covariance function because

$$\sum_{i} \sum_{j} a_i a_j c(s_i, s_j) = \sum_{i} a_i^2 \geq 0, \;\; \forall a$$

Note 12. One way to construct a c.f c is to set $c(s, s') = \psi(s)^{\top} \psi(s')$, for a given vector of basis functions $\psi(\cdot) = (\psi_1(\cdot), ..., \psi_n(\cdot))$.

Proof. From Proposition 10, as

$$\sum_{i} \sum_{j} a_i a_j c(s_i, s_j) = (\psi a)^{\top} (\psi a) \geq 0, \;\; \forall a \in \mathbb{R}^n$$

$\square$

2. Second order random fields (or Second order processes)

Note 13. We introduce a particular class of random fields whose mean and covariance functions exist and which can be used for spatial data modeling.

Definition 14. Second order random field (or second order process) $(Z(s); s \in \mathcal{S})$ is called the random field where $\mathrm{E}\left((Z(s))^2\right) < \infty$ for all $s \in S$.

Example 15. In second order random field $(Z(s); s \in \mathcal{S})$ the associated mean function $\mu(\cdot)$ and covariance function $c(\cdot, \cdot)$ exist, because $c(s, t) = \mathrm{E}(Z(s) Z(t)) - \mathrm{E}(Z(s)) \mathrm{E}(Z(t))$ for $s, t \in \mathcal{S}$.

3. Gaussian random field (or Gaussian process)

Note 16. Gaussian random field is a particular class of second order random field which is widely used in spatial data modeling due to its computational tractability.

Also Example of Proposition

Definition 17. $(Z(s); s \in \mathcal{S})$ is a Gaussian random field (GRF) or Gaussian process (GP) on $\mathcal{S}$ if for any $n \in \mathbb{N}$ and for any finite set $\{s_1, ..., s_n; s_i \in \mathcal{S}\}$, the random vector $(Z(s_1), ..., Z(s_n))^{\top}$ follows a multivariate normal distribution.

Proposition 18. *A GP* $(Z(s); s \in \mathcal{S})$ *is fully characterized by its mean function* $\mu : S \to \mathbb{R}$ *with* $\mu(s) = E(Z(s))$, *and its covariance function with* $c(s, s') = Cov(Z(s), Z(s'))$.

Notation 19. Hence, we denote the GP as $Z(\cdot) \sim \mathcal{GP}(\mu(\cdot), c(\cdot,\cdot))$.

Note 20. When using GP for spatial modeling we just need to specify its functional parameters i.e. the mean and covariance functions.

Note 21. Popular forms of mean functions are polynomial expansions, such as $\mu(s) = \sum_{j=0}^{p-1} \beta_j s^j$ for tunable unknown parameter β. A popular form of covariance functions (c.f.), for tunable unknown parameters $\phi > 0$, and $\sigma^2 > 0$, are

(1) Exponential c.f. $c(s, s') = \sigma^2 \exp\left(-\phi \|s - s'\|_1\right)$
(2) Gaussian c.f. $c(s, s') = \sigma^2 \exp\left(-\phi \|s - s'\|_2^2\right)$
(3) Nugget c.f. $c(s, s') = \sigma^2 1 (s = s')$

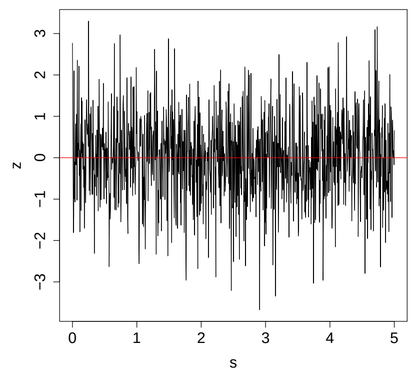
Example 22. Recall your linear regression lessons where you specified the sampling distribution to be $y_x|\beta, \sigma^2 \overset{\text{ind}}{\sim} \mathrm{N}\left(x^\top \beta, \sigma^2\right)$, $\forall x \in \mathbb{R}^d$. Well that can be considered as a GP $Z(\cdot) \sim \mathcal{GP}(\mu(\cdot), c(\cdot,\cdot))$ with $\mu(x) = x^\top \beta$ and $c(x, x') = \sigma^2 1 (x = x')$ in (3).

Example 23. Figures 3.1 & 3.2 presents realizations of GRF $Z(\cdot) \sim \mathcal{GP}(\mu(\cdot), c(\cdot,\cdot))$ with $\mu(s) = 0$ and differently parameterized covariance functions in 1D and 2D. In 1D the code to simulate the GP is given in Algorithm 1. Note that we actually discretize it and simulate it from the fidi.

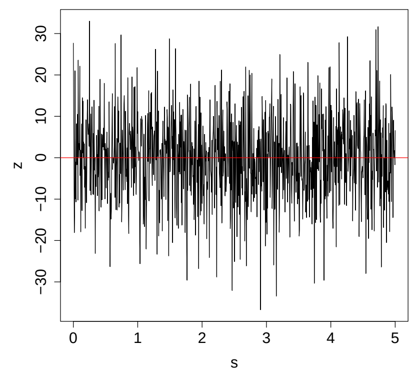
Algorithm 1 R script for simulating from a GP $(Z(s); s \in \mathbb{R}^1)$ with $\mu(s) = 0$ and $c(s,t) = \sigma^2 \exp\left(-\phi \|s-t\|_2^2\right)$

```
# set the GP parameterized mean and covariance function
mu_fun <- function(s) { return (0) }
cov_fun_gauss <- function(s,t,sig2,phi) {
    return ( sig2*exp(-phi*norm(c(s-t),type="2")**2) )
    }
# discretize the problem in n = 100 spatial points
n <- 100
s_vec <- seq(from = 0, to = 5, length = n)
mu_vec <- matrix(nrow = n, ncol = 1)
Cov_mat <- matrix(nrow = n, ncol = n)
# compute the associated mean vector and covariance matrix of the n=100 dimensional
Normal r.v.
sig2_val <- 1.0 ;
phi_val <- 5
for (i in 1:n) {
    mu_vec[i] <- mu_fun(s_vec[i])
    for (j in 1:n) {
        Cov_mat[i,j] <- cov_fun_gauss(s_vec[i],s_vec[j],sig2_val,phi_val)
    }
}
# simulate from the associated distribution
z_vec <- mu_vec + t(chol(Cov_mat))%*%rnorm(n, mean=0, sd=1)
# plot the path (R produces a line plot)
plot(s_vec, z_vec, type="l")
abline(h=0,col="red")
```

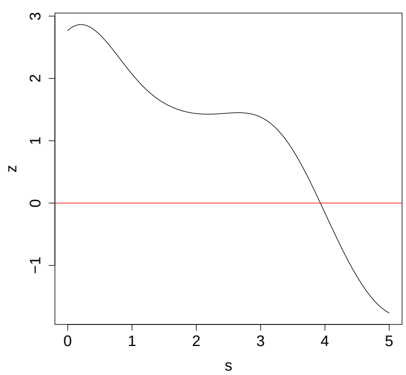
Nugget c.f. is the usual noise where the height of ups and downs are random and controlled by σ^2 (Figures 3.1a & 3.1b ; Figures 3.2a & 3.2b). In Gaussian c.f. the height of ups and downs are random and controlled by σ^2 (Fig.3.1c & 3.1d ; Figures 3.2c & 3.2d), and the spatial dependence / frequency of the ups and downs is controlled by β (Figures 3.1d & 3.1e ; Figures 3.2d & 3.2e). Realizations with different c.f. have different behavior (Figures 3.1a, 3.1d & 3.1e ; Figures 3.2a, 3.2d & 3.2e)



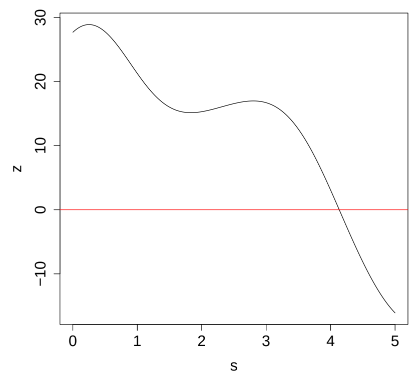
(A) Nugget c.f
($\sigma^2 = 1$)



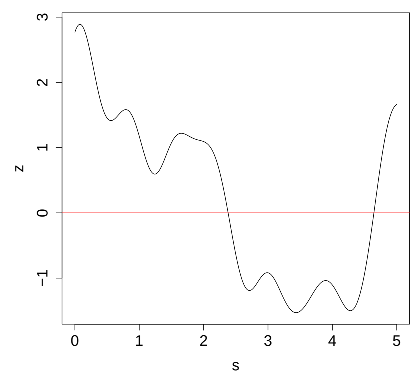
(B) Nugget c.f
($\sigma^2 = 100$)



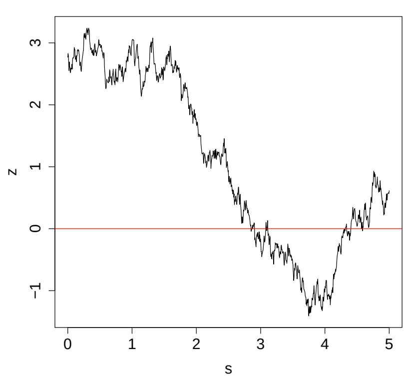
(C) Gauss c.f
($\sigma^2 = 1$,$\phi = 0.5$)



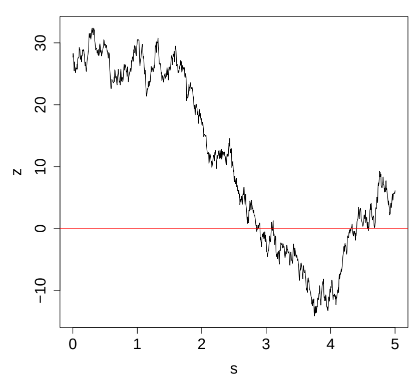
(D) Gauss c.f
($\sigma^2 = 100$,$\phi = 0.5$)



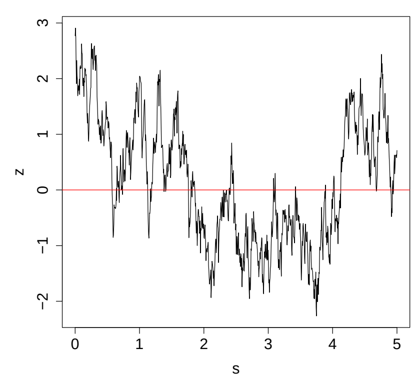
(E) Gauss c.f
($\sigma^2 = 1$,$\phi = 5$)



(F) Exp c.f
($\sigma^2 = 1$,$\phi = 0.5$)

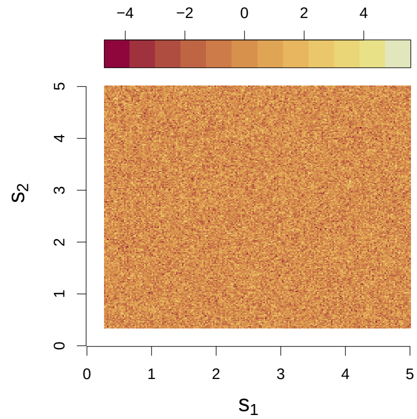


(G) Exp c.f
($\sigma^2 = 100$,$\phi = 0.5$)

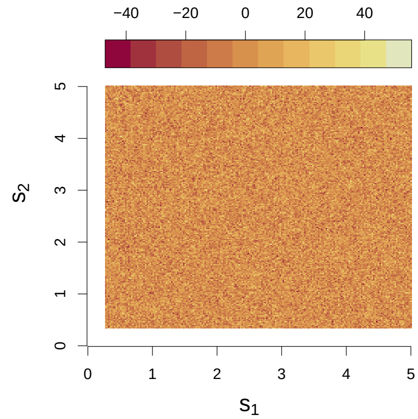


(H) Exp c.f
($\sigma^2 = 1$,$\phi = 5$)

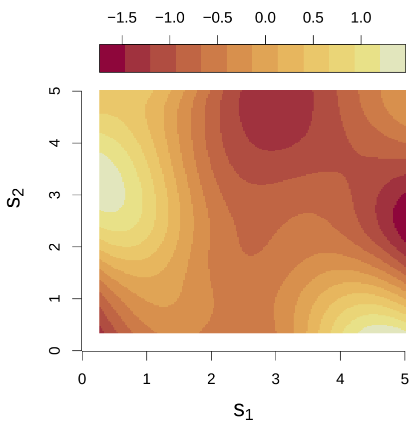
FIGURE 3.1. Realizations of GRF $Z(\cdot) \sim \mathcal{GP}(\mu(\cdot), c(\cdot,\cdot))$ when $s \in [0, 5]$ (using same seed)



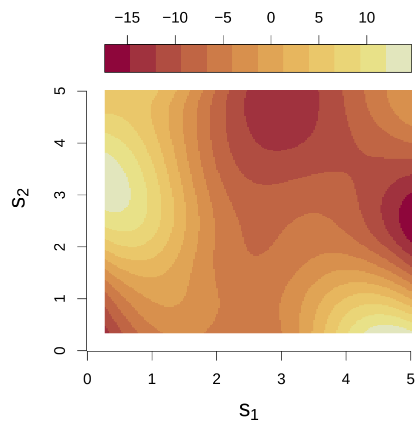
(A) Nugget c.f
($\sigma^2 = 1$)



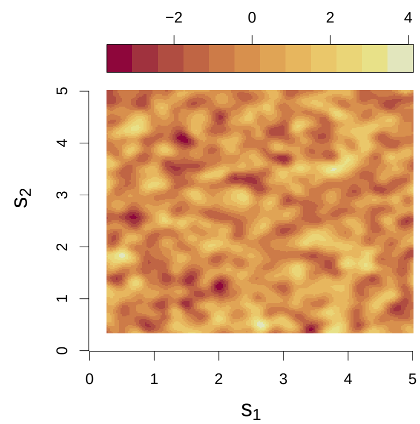
(B) Nugget c.f
($\sigma^2 = 100$)



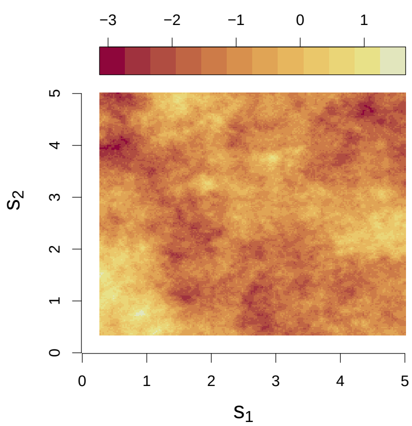
(C) Gauss c.f
($\sigma^2 = 1$,$\phi = 0.5$)



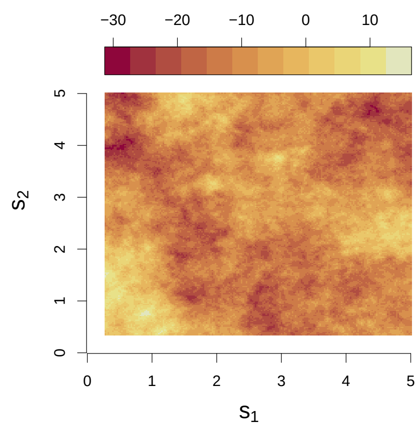
(D) Gauss c.f
($\sigma^2 = 100$,$\phi = 0.5$)



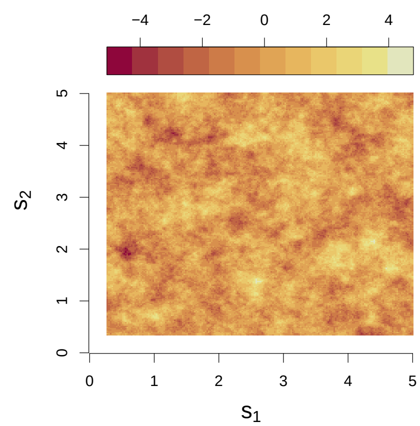
(E) Gauss c.f
($\sigma^2 = 1$,$\phi = 5$)



(F) Exp c.f
($\sigma^2 = 1$,$\phi = 0.5$)



(G) Exp c.f
($\sigma^2 = 100$,$\phi = 0.5$)



(H) Exp c.f
($\sigma^2 = 1$,$\phi = 5$)

FIGURE 3.2. Realizations of GRF $Z(\cdot) \sim \mathcal{GP}(\mu(\cdot), c(\cdot,\cdot))$ when $s \in [0,5]^2$ (using same seed)

4. Strong stationarity

Note 24. We introduce a specific behavior of random field to build our models.

Note 25. Assume $\mathcal{S} = \mathbb{R}^d$ for simplicity. [1]

Definition 26. A random field $(Z(s); s \in \mathcal{S})$ is strongly stationary on $\mathcal{S}$ if for all finite sets consisting of $s_1, ..., s_n \in \mathbb{R}^d$, $n \in \mathbb{N}$, for all $k_1, ..., k_n \in \mathbb{R}$, and for all $h \in \mathbb{R}^d$

$$\mathrm{pr}\left(Z(s_1 + h) \leq k_1, ..., Z(s_n + h) \leq k_n\right) = \mathrm{pr}\left(Z(s_1) \leq k_1, ..., Z(s_n) \leq k_n\right)$$

Note 27. Yuh... strong stationary may represent a behavior being too "restrictive" to be used for spatial data modeling as it is able to represent only limiting number of spatial dependencies.

5. Weak stationarity (or second order stationarity)

Note 28. We introduce another weaker random field behavior which represents a larger class of spatial dependencies.

Note 29. Instead of working with the "restrictive" strong stationarity, we could just properly specify the behavior of the first two moments only; notice that Definition 26 implies that, given $\mathrm{E}(Z_s^2) < \infty$, it is $\mathrm{E}(Z(s)) = \mathrm{E}(Z(s+h)) = $ contst... and $\mathrm{Cov}(Z(s), Z(s')) = \mathrm{Cov}(Z(s+h), Z(s'+h)) \overset{h=-s'}{=} \mathrm{Cov}(Z(s-s'), Z(0)) =$ funct of lag...

Definition 30. A random field $(Z(s); s \in \mathcal{S})$ is weakly stationary (or second order stationary) if it has constant mean and translation invariant covariance; i.e. for all $s, s' \in \mathbb{R}^d$,

1. $\mathrm{E}\left((Z(s))^2\right) < \infty$ (finite)
2. $\mathrm{E}(Z(s)) = \mu$ (constant)
3. $\mathrm{Cov}(Z(s), Z(s')) = c(s' - s)$ for some even function $c : \mathbb{R}^d \to \mathbb{R}$ (lag dependency)

Definition 31. Weakly (or second order) stationary covariance function is called the c.f. of a weakly stationary random field.

6. Covariogram

Note 32. We introduce the covariogram function able to express many aspects of the behavior of a weakly stationary random field and hence be used as statistical descriptive tool.

Definition 33. The covariogram function of a weakly stationary random field $\left(Z(s); s \in \mathbb{R}^d\right)$ is defined by $c : \mathbb{R}^d \to \mathbb{R}$ with

$$c(h) = \mathrm{Cov}(Z(s), Z(s+h)), \ \forall s \in \mathbb{R}^d.$$

[1] Otherwise, we should set $s, s' \in \mathcal{S}$, $h \in \mathcal{H}$, such as $\mathcal{H} = \left\{h \in \mathbb{R}^d : s + h \in \mathcal{S}\right\}$.

Example 34. For the Gaussian c.f. $c(s,t)=\sigma^2\exp\left(-\phi\|s-t\|_2^2\right)$ in (Ex. 20(2)), we may denote just

$$c(h)=c(s,s+h)=\sigma^2\exp\left(-\phi\|h\|_2^2\right) \tag{6.1}$$

Observe that, in Figures 3.1 &3.2, the smaller the ϕ, the smoother the realization (aka slower changes). One way to justify this observation is to think that smaller values of ϕ essentially bring the points closer by re-scaling spatial lags h in the c.f.

Proposition 35. *If $c:\mathbb{R}^d\rightarrow\mathbb{R}$ is the covariogram of a weakly stationary random field $\left(Z(s);s\in\mathbb{R}^d\right)$ then:*

(1) $c(h)=c(-h)$ for all $h\in\mathbb{R}^d$

(2) $|c(h)|\leq c(0)=Var(Z(s))$ for all $h\in\mathbb{R}^d$

(3) $c(\cdot)$ is semi-positive definite; i.e. for all $n\in\mathbb{N}$, $a\in\mathbb{R}^n$, and $\{s_1,...,s_n\}\subseteq S$

$$\sum_{i=1}^n\sum_{j=1}^n a_i a_j c(s_i-s_j)\geq 0$$

Note 36. Given there is some knowledge of the characteristic functions of a suitable distribution, the following spectral representation theorem helps in the specification of a suitable covariogram.

Theorem 37. *(Bochner's theorem) Let $c:\mathbb{R}^d\rightarrow\mathbb{R}$ be a continuous even real-valued function. Then $c(\cdot)$ is positive semi-definete (hence a covariogram of a stationary random field) if and only if it can be represented as*

$$c(h)=\int_{\mathbb{R}^d}\exp\left(i\omega^\top h\right)dF(\omega)$$

where F is a *symmetric positive finite measure on $\mathbb{R}^d$ called spectral measure.*

Note 38. In our course, we focus on cases where F has a density $f(\cdot)$ i.e. $\mathrm{d}F(\omega)=f(\omega)\,\mathrm{d}\omega$. $f(\cdot)$ is called spectral density of $c(\cdot)$, it is

$$c(h)=\int_{\mathbb{R}^d}\exp\left(i\omega^\top h\right)f(\omega)\,\mathrm{d}\omega,$$

and it dies as $\lim_{|h|\rightarrow\infty}c(h)=0$

Theorem 39. *If $c(\cdot)$ is integrable, the spectral density $f(\cdot)$ can be computed by inverse Fast Fourier transformation*

$$f(\omega)=\left(\frac{1}{2\pi}\right)^d\int_{\mathbb{R}^d}\exp\left(-i\omega^\top h\right)c(h)\,dh$$

Example 40. Consider the Gaussian c.f. $c(h) = \sigma^2 \exp\left(-\phi \|h\|_2^2\right)$ for $\sigma^2, \beta > 0$ and $h \in \mathbb{R}^d$. Then, by using Theorem 37, the spectral density is

$$
\begin{aligned}
f(\omega) &= \left(\frac{1}{2\pi}\right)^d \int_{\mathbb{R}^d} \exp\left(-i\omega^\top h\right) \sigma^2 \exp\left(-\phi \|h\|_2^2\right) \mathrm{d}h \\
&= \sigma^2 \left(\frac{1}{2\pi}\right)^d \prod_{j=1}^d \int_{\mathbb{R}} \exp\left(-i\omega_j h_j - \phi h_j^2\right) \mathrm{d}h \\
&= \sigma^2 \left(\frac{1}{2\pi}\right)^d \prod_{j=1}^d \int_{\mathbb{R}} \exp\left(-\phi\left(h_j - \left(-i\omega/\left(2\phi\right)\right)\right)^2\right) \exp\left(-\omega_j^2/\left(4\phi\right)\right) \mathrm{d}h_j \\
&= \sigma^2 \left(\frac{1}{4\pi\phi}\right)^{d/2} \exp\left(-\|\omega\|_2^2/\left(4\phi\right)\right)
\end{aligned}
$$

i.e. it has a Gaussian form.

Definition 41. Let $\left(Z(s); s \in \mathbb{R}^d\right)$ be a weakly stationary random field with covariogram function $c : \mathbb{R}^d \to \mathbb{R}$ and $c(h) = \mathrm{Cov}\left(Z(s), Z(s+h)\right)$. The correlogram function $\rho : \mathbb{R}^d \to [-1, 1]$ is defined as

$$\rho(h) = \frac{c(h)}{c(0)}.$$

7. Intrinsic stationarity (of order zero)

Note 42. Weakly stationary random fields may not be sufficiently general for all of our applications. This class of random fields can be extended by considering intrinsic stationary instead. For instance in certain applications it has been noticed that the "real process" may be sch that the variance of the increments

$$\mathrm{Var}\left(Z(s+h) - Z(s)\right) = \mathrm{Var}\left(Z(s+h)\right) + \mathrm{Var}\left(Z(s)\right) - 2\mathrm{Cov}\left(Z(s+h), Z(s)\right)$$

increases indefinitely with $|h|$, however this real process cannot be modeled by a weakly stationary random field which implies a bounded $\mathrm{Var}\left(Z(s+h) - Z(s)\right) = 2\left(c(0) - c(h)\right) < 2c(0)$.

Definition 43. A random field $\left(Z(s) : s \in \mathcal{S}\right)$ is called intrinsic random field (or intrinsically stationary r.f.) if, for all $h \in \mathcal{H}$, $\mathcal{H} = \{h : s \in \mathcal{S}, s + h \in \mathcal{S}\}$,

1. $\mathrm{E}\left(Z(s+h) - Z(s)\right)^2 < \infty$
2. $\mathrm{E}\left(Z(s+h) - Z(s)\right) = \mu(h)$ for some function $\mu : \mathcal{H} \to \mathbb{R}$ (lag dependent)
3. $\mathrm{Var}\left(Z(s+h) - Z(s)\right) = 2\gamma(h)$ for some function $\gamma : \mathcal{H} \to \mathbb{R}$ (lag dependent)

Definition 44. Intrinsic covariance function is called the c.f. of an intrinsically stationary stochastic process.

Example 45. The covariance function is not weakly

$$c(s,t)=\frac{1}{2}\left(\|s\|^{2H}+\|t\|^{2H}-\|t-s\|^{2H}\right),\ H\in(0,1)$$

is not weakly because for $h\in\mathcal{H}$

$$c(s,s+h)=\frac{1}{2}\left(\|s\|^{2H}+\|s+h\|^{2H}-\|h\|^{2H}\right)$$

but intrinsically stationary because and

$$\frac{1}{2}\mathrm{Var}\left(Z(s+h)-Z(s)\right)=\frac{1}{2}\left(\mathrm{Var}\left(Z(s)\right)+\mathrm{Var}\left(Z(s+h)\right)-2\mathrm{Cov}\left(Z(s),Z(s+h)\right)\right)=\frac{1}{2}\|h\|^{2H}$$

8. Incremental mean function

Definition 46. Incremental mean function of the intrinsic random field $(Z(s))_{s\in\mathcal{S}}$ is defined by $\mu:\mathcal{H}\to\mathbb{R}$ with $\mu(h)=\mathrm{E}\left(Z(s+h)-Z(s)\right)$.

Note 47. The incremental mean function $\mu(h)$ of the intrinsic random field $(Z(s))_{s\in\mathcal{S}}$ has a form

$$\mu(h)=h^{\top}\beta \tag{8.1}$$

for some $\beta\in\mathbb{R}^{d}$. Because,

$$\mu(s+h)=\mathrm{E}\left(Z(s+h)-Z(s)\right)-\mathrm{E}\left(Z(s)-Z(0)\right)=\mu(s)+\mu(h),\ \forall s,h$$

$\mu(0)=0$, and $\mu(\cdot)$ is continuous, $\mu(\cdot)$ can be is a linear expansion under intrinsic assumption.

9. Semi variogram and Variogram

Note 48. The definition of the semi-variogram function requires the random field to be intrinsic stationarity; which is weaker assumption than weak stationary required by covariogram.

Definition 49. The semivariogram of an intrinsic random field $(Z(s))_{s\in\mathcal{S}}$ is defined by $\gamma:\mathcal{H}\to\mathbb{R}$, $\mathcal{H}=\{h:s\in\mathcal{S},s+h\in\mathcal{S}\}$ with

$$\gamma(h)=\frac{1}{2}\mathrm{Var}\left(Z(s+h)-Z(s)\right)$$

Definition 50. Semivariogram of an intrinsic random field $(Z(s):s\in\mathcal{S})$ is called the quantity $2\gamma(h)$.

Note 51. A weakly stationary random field $(Z(s))_{s\in\mathcal{S}}$ with covariogram $c(\cdot)$ and mean μ is intrinsic stationary as well with semi-variogram

$$\gamma(h)=c(0)-c(h), \tag{9.1}$$

and constant incremental mean $\mu(h)=\mu$.

Example 52. For the Gaussian covariance function (Ex. 34) the semi-variogram is

$$\gamma(h) = c(0) - c(h) = \sigma^2 \left(1 - \exp\left(-\beta \|h\|_2^2\right)\right)$$

Proposition 53. *Properties of semivariogram. Let $(Z(s))_{s\in\mathcal{S}}$ be an intrinsically stationary process.*

(1) It is $\gamma(h) = \gamma(-h)$, $\gamma(h) \geq 0$, and $\gamma(0) = 0$
(2) Semi-variogram is conditionally negative definite (c.n.d.): for all $a \in \mathbb{R}^n$s.t. $\sum_{i=1}^n a_i = 0$, and for all $\{s_1, ..., s_n\} \subseteq S$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \gamma(s_i - s_j) \leq 0$$

10. Behavior of semivariogram under intrinsic stationarity

Note 54. The semivariogram $\gamma(h)$ is very informative when plotted against the lag h, below we discuss some of the characteristics of it, using Figure 10.1

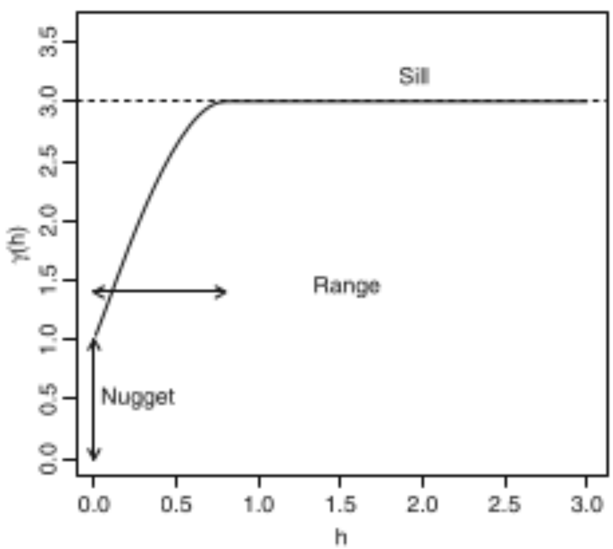


Figure 10.1. Semi Variogram's characteristics

Note 55. A semivariogram tends to be an increasing function of the lag $\|h\|$. Recall that for weakly stationary random fields with c.f. $c(\cdot)$, it is $\gamma(h) = c(0) - c(h)$ where common logic suggests that $c(h)$ is decreases with $\|h\|$.

Note 56. If $\gamma(h)$ is a positive constant for all lags $h \neq 0$, then $Z(s_1)$ and $Z(s_2)$ are uncorrelated regardless of how close $s_1$and s_2 are; and $Z(\cdot)$ is often called white noise.

Note 57. Conversely, a non zero slope of the variogram indicates structure.

Nugget Effect.

Note 58. Nugget effect is the semivariagram limiting value

$$\sigma_{\varepsilon}^{2} = \lim_{\|h\|\to 0} \gamma\left(h\right)$$

In particular when $\sigma_{\varepsilon}^{2} \neq 0$.

Note 59. Nugget effect $\sigma_{\varepsilon}^{2} \neq 0$ may be expected or assumed to appear due to (1) measurement errors (e.g., if we collect repeated measurements at the same location s) or (2) due to some microscale variation causing discontinuity in the origin that cannot be detected from the data i.e. the spatial gaps because we collect a finite set of measurements at spatial locations. Hence theoretically, we could consider a more detailed decomposition $\sigma_{\varepsilon}^{2} = \sigma_{\text{MS}}^{2} + \sigma_{\text{MS}}^{2}$ where σ_{MS}^{2} refers to the microscale and σ_{MS}^{2} refers to the measurement error; however (my experience) this leads to non-identifiability.

Sill.

Definition 60. Sill is the semivariagram limiting value $\lim_{\|h\|\to\infty} \gamma\left(h\right)$.

Note 61. For intrinsic processes, the sill may be infinite or finite. For weakly random field, the sill is always finite.

Partial sill .

Definition 62. Partial sill is $\lim_{\|h\|\to\infty} \gamma\left(h\right) - \lim_{\|h\|\to 0} \gamma\left(h\right)$ which takes into account the nugget.

Range .

Note 63. Range is the distance at which the semivariogram reaches the Sill. It can be infinite or finite.

Other.

Note 64. An abrupt change in slope indicates the passage to a different structuration of the values in space. This is often modeled via decomposition of processes with different semivariograms. E.g., let independent random fields $Y\left(\cdot\right)$ and $X\left(\cdot\right)$ with different semivariograms γ_{Y} and γ_{X},then random field $Z\left(\cdot\right)$ with $Z\left(s\right) = Y\left(s\right) + X\left(s\right)$ has semivariogram $\gamma_{Z}\left(h\right) = \gamma_{Y}\left(h\right) + \gamma_{X}\left(h\right)$ which may present such a behavior.

11. Isotropy

Note 65. Isotropy as a concept imposes the assumption of “rotation invariance”.

Note 66. Isotropy applies to both intrinsic stationary and weakly stationary processes.

Definition 67. An intrinsic random field $(Z(s))_{s\in\mathcal{S}}$ is isotropic iff

$$\forall s,t\in\mathcal{S},\ \frac{1}{2}\mathrm{Var}\left(Z(s)-Z(t)\right)=\gamma\left(\|t-s\|\right),\ \text{for some function }\gamma{:}\mathbb{R}^{+}\to\mathbb{R}. \tag{11.1}$$

Definition 68. Isotropic semivariogram $\gamma:\mathcal{H}\to\mathbb{R}$ is the semivariogram of the isotropic random field (sometimes for simplicity of notation we use $\gamma:\mathbb{R}^{+}\to\mathbb{R}$ with $\gamma(\|h\|)=\frac{1}{2}\mathrm{Var}\left(Z(s)-Z(s-h)\right)$).

Definition 69. Isotropic covariance function $C:\mathcal{S}\times\mathcal{S}\to\mathbb{R}$ is called the covariance function satisfying (11.1).

Definition 70. Isotropic covariogram $c:\mathcal{H}\to\mathbb{R}$ of a weakly stationary process is the covariogram associated to an isotropic semi-variogram (sometimes for simplicity of notation we use $c:\mathbb{R}^{+}\to\mathbb{R}$ with $c(\|h\|)$ from (11.1)).

11.1. Popular isotropic covariance functions.

Note 71. Given the covariogram $c(\cdot)$, and the semi-variogram can be computed from $\gamma(h)=c(0)-c(h)$ for any h.

11.1.1. *Nugget-effect.*

Note 72. For $\sigma^2>0$,

$$c(h)=\sigma^2 1_{\{0\}}\left(\|h\|\right)$$

is the nugget-effect covariogram. It is associate to white noise. It is used to model a discontinuity in the origin of the covariogram / sem-variogram.

11.1.2. *Matern c.f.*

Note 73. For $\sigma^2>0$, $\phi>0$, and $\nu\geq 0$ No need to memorize (11.2)

$$c(h)=\sigma^2\frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\|h\|}{\phi}\right)^{\nu}K_{\nu}\left(\frac{\|h\|}{\phi}\right) \tag{11.2}$$

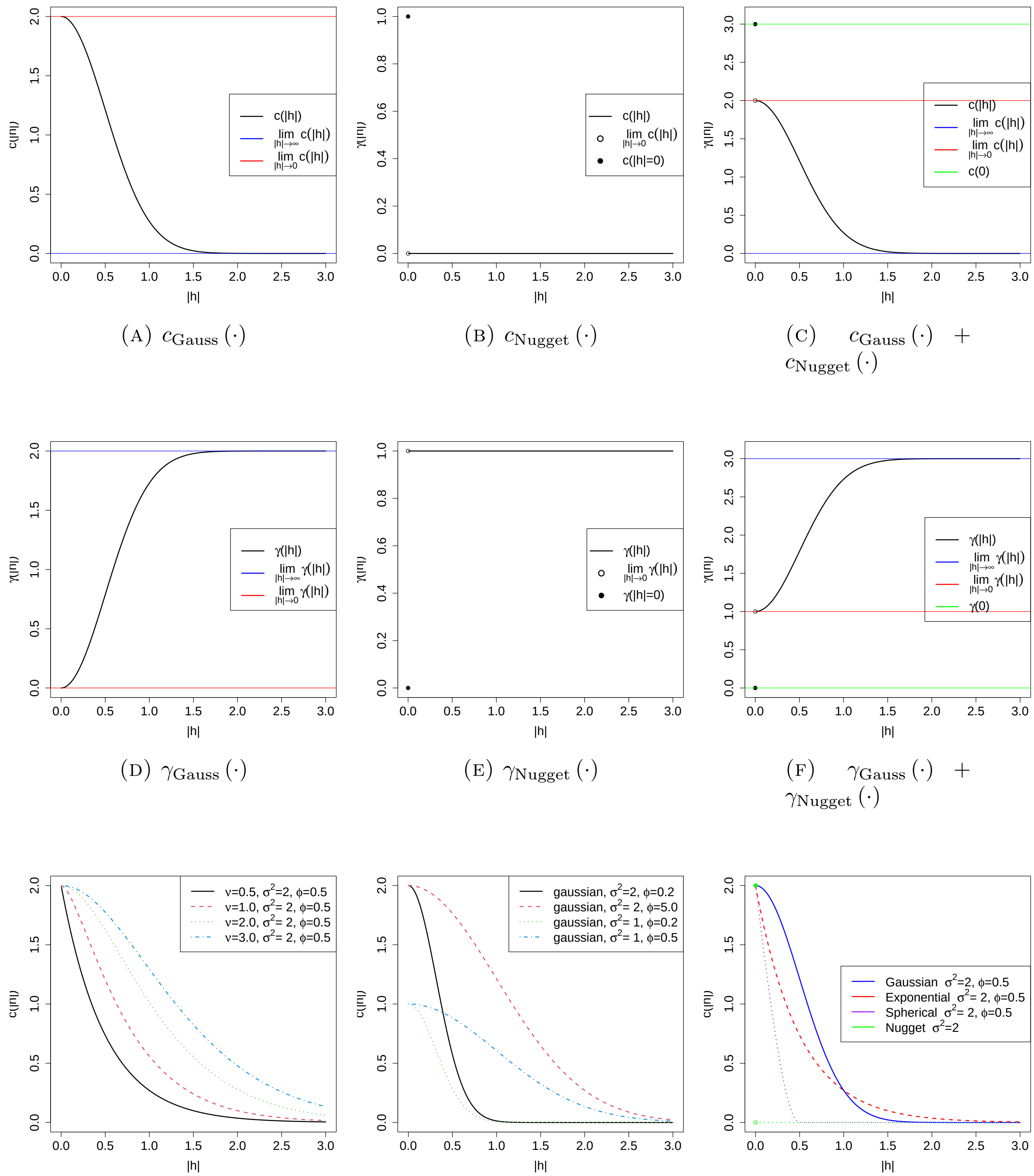
is the Matern covariogram. Parameter ν controls the variogram's regularity at 0 which in turn controls the quadratic mean (q.m.) regularity of the associated process. For $\nu=1/2$, we get the exponential c.f.,

$$c(h)=\sigma^2\exp\left(-\frac{1}{\phi}\|h\|_1\right)$$

which is not differentiable at $h=0$, while for $\nu\to\infty$, we get the Gaussian c.f.

$$c(h)=\sigma^2\exp\left(-\frac{1}{\phi}\|h\|_2^2\right)$$

which is infinite differentiable. ϕ is a range parameter, and σ^2 is the (partial) sill parameter.



(A) $c_{\text{Gauss}}\left(\cdot\right)$ (B) $c_{\text{Nugget}}\left(\cdot\right)$ (C) $c_{\text{Gauss}}\left(\cdot\right) + c_{\text{Nugget}}\left(\cdot\right)$

(D) $\gamma_{\text{Gauss}}\left(\cdot\right)$ (E) $\gamma_{\text{Nugget}}\left(\cdot\right)$ (F) $\gamma_{\text{Gauss}}\left(\cdot\right) + \gamma_{\text{Nugget}}\left(\cdot\right)$

(G) c_{Matern} (H) $c_{\text{Gauss}}\left(\cdot;\sigma^2,\phi\right)$ (I) $c_{\cdot}\left(\cdot\right)$

FIGURE 11.1. Covariogrames $c\left(\cdot\right)$ and semivariogrames $\gamma\left(\cdot\right)$

11.1.3. *Spherical c.f.*

Note 74. [2]For $\sigma^2 > 0$ and $\phi > 0$ No need to memorize (11.3)

$$c(h) = \begin{cases} \sigma^2 \left(1 - \frac{3}{2}\frac{\|h\|_1}{\phi} + \frac{1}{2}\left(\frac{\|h\|_1}{\phi}\right)^3\right) & \|h\|_1 \leq \phi \\ 0 & \|h\|_1 > \phi \end{cases}, \quad h \in \mathbb{R}^3. \tag{11.3}$$

The c.f. starts from its maximum value σ^2 at the origin, then steadily decreases, and finally vanishes when its range ϕ is reached. ϕ is a range parameter, and σ^2 is the (partial) sill parameter.

12. Anisotropy

Note 75. Dependence between $Z(s)$ and $Z(s+h)$ is a function of both the magnitude and the direction of separation h. This can be caused by the underlying physical process evolving differently in space (e.g., vertical and horizontal axes).

Definition 76. The semivariogram $\gamma : \mathcal{H} \to \mathbb{R}$ is anisotropic if there are h_1 and h_2 with same length $\|h_1\| = \|h_2\|$ but different direction $h_1/\|h_1\| \neq h_2/\|h_2\|$ that produce different semivariograms $\gamma(h_1) \neq \gamma(h_2)$.

Definition 77. The intrinsically random field $(Z(s))_{s\in\mathcal{S}}$ is anisotropic if its semivariogram is anisotropic.

Definition 78. The covariogram $c : \mathcal{H} \to \mathbb{R}$ is anisotropic if there are h_1 and h_2 with same length $\|h_1\| = \|h_2\|$ but different direction $h_1/\|h_1\| \neq h_2/\|h_2\|$ that produce different covariogram $c(h_1) \neq c(h_2)$.

Definition 79. The weakly random field $(Z(s))_{s\in\mathcal{S}}$ is anisotropic if its covariogram is anisotropic.

Note 80. For brevity, below we discuss about intrinsic random fields and semivariograms, however the concepts/definitions apply to weakly stationary process and covariograms when defined, as in Defs 76 & 78.

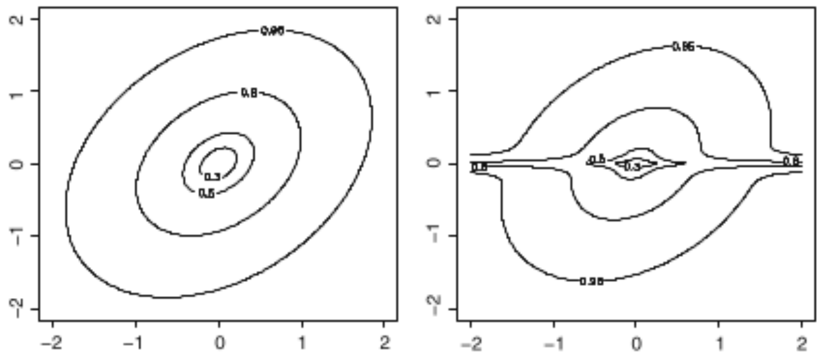
12.1. Geometric anisotropy.

Definition 81. The semivariogram $\gamma_{\text{g.a.}} : \mathcal{H} \to \mathbb{R}$ exhibits geometric anisotropy if it results from an A-linear deformation of an isotropic semivariogram with function $\gamma_{\text{iso}}(\cdot)$; i.e.

$$\gamma_{\text{g.a.}}(h) = \gamma_{\text{iso}}(\|Ah\|_2)$$

Note 82. Such semivariograms have the same sill in all directions but with ranges that vary depending on the direction. See Figure 12.1a.

[2]For it's derivation see Ch 8 in [4]



(A) Geometric anisotropy

(B) Zonal anisotropy

FIGURE 12.1. Isotropy vs Anisotropy

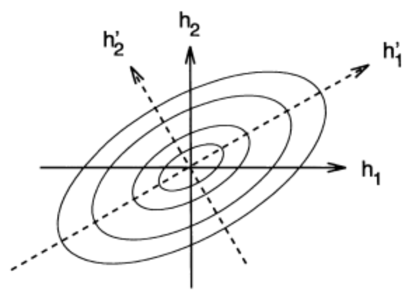


FIGURE 12.2. Rotation of the 2D coordinate system

Example 83. For instance, if $\gamma_{\text{g.a.}}(h) = \gamma_{\text{iso}}\left(\sqrt{h^\top Q h}\right)$, where $Q = A^\top A$.

Example 84. [Rotating and dilating an ellipsoid in 2D] Consider a coordinate system for $h = (h_1, ..., h_n)^\top$. We wish to find a new coordinate system for h in which the iso-semivariogram lines are spherical.

(1) [Rotate] Apply rotation matrix R to h such as $h' = Qh$. In 2D, it is

$$R = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}, \text{ for } \theta \in (0, 2\pi), \text{ is the rotation angle.}$$

(2) [Dilate] Apply a dilation of the principal axes of the ellipsoid using a diagonal matrix $\sqrt{\Lambda} = \text{diag}\left(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n}\right)$, as $\tilde{h} = \sqrt{\Lambda} h'$.

Now the ellipsoids become spheres with radius $r = \left\|\tilde{h}\right\|_2 = \sqrt{\tilde{h}^\top \tilde{h}}$. This yields the equation of an ellipsoid in the h coordinate system

$$h^\top \left(R^\top \Lambda R\right) h = r^2$$

where the diameters d_j (principal axes) of the ellipsoid along the principal directions are

$$d_j = 2r/\sqrt{\lambda_j}$$

and the principal direction is the j-th column of the rotation matrix $R_{:,j}$.

Hence the anisotropic semivariogram is $\gamma_{\text{g.a.}}(h) = \gamma_{\text{iso}}\left(\sqrt{h^\top Q h}\right)$ with $Q = R^\top \Lambda R$. This derivation extends to d dimensions.

12.2. **Zonal (or stratified) anisotropy.**

Definition 85. Support anisotropy is called the type of anysotropy when the semivariogram $\gamma(h)$ depends only on certain coordinates of h.

Example 86. If it is $\gamma(h = (h_1, h_2)) = \gamma(h_1)$, then I ve support anisotropy

Definition 87. Zonal anisotropy occurs when the semivariogram $\gamma(h)$ is the sum of several components each with a support anisotropy.

Example 88. Let γ' and γ'' be semivariogram with sills v' and v'' correspondingly. If it is $\gamma(h = (h_1, h_2)) = \gamma'(\|h_1\|) + \gamma''\left(\sqrt{\|h_1\| + \|h_2\|}\right)$, then I 've Zonal anisotropy because γ has a sill $v' + v''$ is direction $(0,1)$ and a sill v' in direction $(1,0)$.

Note 89. We have Zonal anisotropy then the semivariogram calculated in different directions suggest a different value for the sill (and possibly the range).

Note 90. If in 2D case, the sill in h_1 is larger than that in h_2, we can model zonal anysotropy of random field $(Z(s))_{s\in\mathcal{S}}$ by assuming $Z(s) = I(s) + A(s)$, where $I(s)$ is an isotropic random field with isotropic semivariogram γ_I along dimension of h_1 and $A(s)$ is an process with anisotropic semivariogram γ_I without effect on dimension h_1; i.e. $\gamma_Z(h) = \gamma_I(h) + \gamma_A(h)$.

12.3. **Non-linear deformations.**

Note 91. A (rather too general) non-stationary model can be specified by considering semivariogram $2\text{Var}(Z(s) - Z(t)) = 2\gamma_o(\|G(s) - G(t)\|)$ where we have performed a bijective non-linear (function) deformation $G(\cdot)$ of space $\mathcal{S}$ and applied on the isotropic semivariogram γ_o. For instance, $\gamma_o(h) = \sigma^2 \exp(-\|h\|/\phi)$ and $G(s) = s^2$ as a deterministic function. Now, if function $G(\cdot)$ is considered as unknown, one can model it as a random field $(G(s))_{s\in\mathcal{S}}$, and then we will be talking about deep learning modeling stuff.

13. Geometrical properties of random fields

Note 92. We discuss basic geometric properties of the basic models we will use for modeling, as it can give us a deeper intuition on how to design appropriate spatial statistical models.

Definition 93. (Continuity in quadratic mean (q.m.)) Second-order process $(Z(s))_{s\in\mathcal{S}}$ is q.m. continuous at $s\in\mathcal{S}$ if

$$\lim_{h\to 0}\mathrm{E}\left(Z(s+h)-Z(s)\right)^2=0.$$

Note 94. Consider random field $(Z(s))_{s\in\mathcal{S}}$.Then

$$\mathrm{E}\left(Z(s+h)-Z(s)\right)^2=\left(\mathrm{E}\left(Z(s+h)\right)-\mathrm{E}\left(Z(s)\right)\right)^2+\mathrm{Var}\left(Z(s+h)-Z(s)\right)$$

- If Z is intrinsically stationary, then

$$\mathrm{E}\left(Z(s+h)-Z(s)\right)^2=\frac{1}{2}\gamma(h)$$

  and hence Z is q.m. continuous iff $\lim_{\|h\|\to 0}\gamma(h)=\gamma(0)$.
- If Z is weakly stationary, then

$$\mathrm{E}\left(Z(s+h)-Z(s)\right)^2=\frac{1}{2}\left(c(0)-c(h)\right)$$

  and hence Z is q.m. continuous iff $\lim_{\|h\|\to 0}c(h)=c(0)$ (i.e. ,c is continuous).

Definition 95. Differentiable in quadratic mean (q.m.)) Second-order process $(Z_s)_{s\in\mathcal{S}}$ is q.m. differentiable at $s\in\mathcal{S}$ there exist

$$\dot{Z}(s)=\lim_{h\to 0}\frac{Z(s+h)-Z(s)}{h}.\text{ in } q.m. \tag{13.1}$$

Proposition 96. *Let $c(s,t)$ be the covariance function of $(Z_s)_{s\in S}$. Then Z is everywhere differentiable if $\frac{\partial^2}{\partial s\partial t}c(s,t)$ exists and it is finite. Also, $\frac{\partial^2}{\partial s\partial t}c(s,t)$ is the covariance function of (13.1).*

Example 97. The process with Gaussian c.f. $c(h)=\sigma^2\exp\left(-|h|/\phi\right)$ is continuous because $\lim_{h\to 0}c(h)=\sigma^2=c(0)$ but not differentiable because $\frac{\partial^2}{\partial h^2}c(h)$ does not exist at $h=0$.