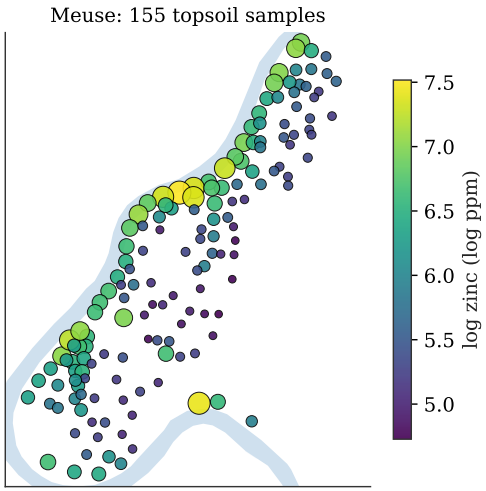
Lecture 1: Random Fields and Covariance

By the end of this lecture you should be able to:

- define a random field and its mean and covariance functions;
- state strict, weak and intrinsic stationarity and prove the implications between them;
- prove the relationship $\gamma(h) = \sigma(0) - \sigma(h)$ and the conditional negative definiteness of γ;
- characterise valid covariograms (positive semi-definiteness, Bochner's theorem) and use the closure properties to build new ones;
- simulate a Gaussian random field.

The Meuse floodplain



- 155 topsoil samples on the floodplain of the river Meuse near Stein (NL); heavy-metal concentrations measured at each site.
- Zinc is high near the river and decays away from it; *nearby samples are similar*.
- Questions we will answer this semester:
  - How strongly, and over what distances, are values related?
  - What is the zinc concentration at an *unsampled* site, and how uncertain is that prediction? (Lecture 5)
  - Does the dependence look the same everywhere, and in all directions? (Lectures 2, 4)
- Key modelling step: regard the 155 values as *one partially observed realisation* of a random function over the whole floodplain.

Random fields

Definition 1 (Random field)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $D \subseteq \mathbb{R}^d$. A *random field* on D is a collection $\{X(s) : s \in D\}$ of real random variables on Ω.

- For fixed $\omega \in \Omega$, $s \mapsto X(s, \omega)$ is a *realisation* (a function on D); for fixed s, $X(s)$ is a random variable.
- The law of X is described by its *finite-dimensional distributions*

$$F_{s_1,\dots,s_n}(x_1, \dots, x_n) = \mathbb{P}\big(X(s_1) \le x_1, \dots, X(s_n) \le x_n\big), \qquad n \in \mathbb{N},\ s_i \in D.$$

- **Kolmogorov's extension theorem** (stated): any family of finite-dimensional distributions that is consistent under permutation of the sites (e.g. $F_{s_1,s_2}(x_1, x_2) = F_{s_2,s_1}(x_2, x_1)$) and under marginalisation (e.g. $F_{s_1}(x_1) = \int_{\mathbb{R}} F_{s_1,s_2}(x_1, \mathrm{d}x_2)$) defines a random field.
- Data: $X_i := X(s_i)$, $i = 1, \dots, n$ — a single realisation, observed at finitely many sites.

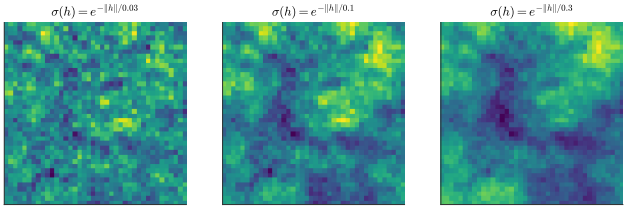
Gaussian random fields

Definition 2 (Gaussian random field)

X is *Gaussian* if every finite-dimensional distribution is multivariate Gaussian:

$$(X(s_1), \ldots, X(s_n))^\top \sim \text{Gaussian}(\mu, \Sigma), \qquad \mu_i = \mu(s_i), \quad \Sigma_{ij} = \sigma(s_i, s_j).$$

Realisations of zero-mean Gaussian random fields (same white noise)



- A Gaussian field is specified *completely* by its mean and covariance functions.
- Non-Gaussian data (counts, presences, point patterns) can be handled through a *latent* Gaussian field (Lecture 8, Sections 3–4).

Mean and covariance functions

Definition 3 (Mean and covariance functions)

For a field with $\mathbb{E}X(s)^2 < \infty$: the *mean function* $\mu(s) := \mathbb{E}X(s)$ and the *covariance function* $\sigma(s, s') := \mathbb{C}(X(s), X(s'))$.

Basic properties. $\sigma(s, s') = \sigma(s', s)$; $\sigma(s, s) = \mathbb{V}X(s) \geq 0$; σ is positive semi-definite (slide 13).

The identifiability problem. From n observations of a *single* realisation we cannot estimate without further assumptions

- n unknown means $\mu(s_1), \ldots, \mu(s_n)$, *and*
- $n(n+1)/2$ unknown covariances $\sigma(s_i, s_j)$.

Is a high value at s_i a high mean, or a large positive fluctuation? Without structure, the question has no answer.

Two sources of structure.

- a *model for the mean*, e.g. $\mu(s) = x(s)^\top \beta$ (Lecture 6);
- *stationarity* (and often isotropy) of the dependence: the covariance depends on the sites only through their separation. Then pairs of sites at similar separations act as *replicates*.

Strict stationarity

Definition 4 (Strict stationarity)

A random field X on $\mathbb{R}^d$ is *strictly stationary* if for all $n \in \mathbb{N}$, $s_1, \dots, s_n \in \mathbb{R}^d$ and $h \in \mathbb{R}^d$,

$$\big(X(s_1), \dots, X(s_n)\big) \overset{d}{=} \big(X(s_1 + h), \dots, X(s_n + h)\big).$$

- All finite-dimensional distributions are invariant under translation: the probabilistic structure "looks the same" everywhere.
- It is a property of the *generating mechanism*, not of a dataset; with one realisation it cannot be verified, only made plausible (or implausible).
- Defined on the whole of $\mathbb{R}^d$. Data on a bounded region D are treated as the restriction of a stationary field on $\mathbb{R}^d$ to D.
- Strict stationarity says nothing about moments existing; when they do, it constrains μ and σ strongly (next slide).

Consequences of strict stationarity

Proposition 5

If X is strictly stationary with $\mathbb{E}X(s)^2 < \infty$, then μ is constant and $\sigma(s, s')$ depends on (s, s') only through $s' - s$.

Proof. Take $n = 1$: $X(s) \overset{d}{=} X(s+h)$, so $\mu(s) = \mu(s+h)$ for all s, h; hence μ is constant.

Take $n = 2$: $(X(s), X(s')) \overset{d}{=} (X(s+h), X(s'+h))$, so $\sigma(s, s') = \sigma(s+h, s'+h)$ for all h. Choosing $h = -s$, $\sigma(s, s') = \sigma(0,\, s' - s)$. □

Definition 6 (Covariogram)

For such a field the *covariogram* is $\sigma(h) := \mathbb{C}\big(X(s), X(s+h)\big)$, $h \in \mathbb{R}^d$, which does not depend on s.

We overload notation: $\sigma(s, s') = \sigma(s' - s)$. The covariogram is a function of one vector argument, the lag h.

Weak stationarity

Definition 7 (Weak (second-order) stationarity)

X is *weakly stationary* if $\mathbb{E}X(s)^2 < \infty$, $\mu(s) = \mu$ is constant, and $\mathbb{C}(X(s), X(s+h)) = \sigma(h)$ for all s, h.

Proposition 8

The covariogram of a weakly stationary field satisfies (i) $\sigma(0) = \mathbb{V}X(s) \geq 0$*; (ii)* $\sigma(-h) = \sigma(h)$*; (iii)* $|\sigma(h)| \leq \sigma(0)$.

Proof. (i) definition. (ii) $\sigma(-h) = \mathbb{C}(X(s), X(s-h)) = \mathbb{C}(X(s-h), X(s)) = \sigma(h)$. (iii) Cauchy–Schwarz: $|\mathbb{C}(X(s), X(s+h))| \leq \sqrt{\mathbb{V}X(s)\,\mathbb{V}X(s+h)} = \sigma(0)$. □

Relationship to strict stationarity.

- Strict stationarity with finite second moments ⇒ weak stationarity (previous slide).
- The converse fails in general: weak stationarity constrains only two moments. [Sheet I, Q1a]
- **Gaussian fields:** weak ⇔ strict. A Gaussian finite-dimensional distribution is determined by its mean vector and covariance matrix, and both are translation invariant under weak stationarity.

Definition 9 (Intrinsic stationarity; variogram)

X is *intrinsically stationary* if for all s, h

$$\mathbb{E}\big(X(s+h) - X(s)\big) = 0 \qquad \text{and} \qquad \tfrac{1}{2}\,\mathbb{V}\big(X(s+h) - X(s)\big) = \gamma(h).$$

γ is the *semivariogram*; 2γ is the *variogram*. (Many often use "variogram" for γ.)

- Only the *increments* need be stationary; $X(s)$ itself need not have a finite or constant variance.
- Properties: $\gamma(0) = 0$, $\gamma(-h) = \gamma(h)$, $\gamma(h) \geq 0$.
- Example: Brownian motion is intrinsically but not weakly stationary. [Sheet I, Q1b]
- A (continuous) variogram grows no faster than $\|h\|^2$, which is attained by a random linear trend. Quadratic growth in an empirical variogram is often a signature of a mean-nonstationarity (Lecture 3).
- The variogram is the traditional object of geostatistics (Matheron, 1962): it requires fewer assumptions and can be estimated without first estimating μ.

Proposition 10

If X is weakly stationary, it is intrinsically stationary with $\gamma(h) = \sigma(0) - \sigma(h)$.

Proof. $\mathbb{E}(X(s+h) - X(s)) = \mu - \mu = 0$, and

$$\tfrac{1}{2}\mathbb{V}\big(X(s+h) - X(s)\big) = \tfrac{1}{2}\big[\sigma(0) - 2\sigma(h) + \sigma(0)\big] = \sigma(0) - \sigma(h). \qquad \square$$

Consequences.

- If $\sigma(h) \to 0$ as $\|h\| \to \infty$, then $\gamma(h) \to \sigma(0)$: the *sill* of the variogram is the variance.
- Conversely, if γ is bounded and $\gamma(h) \to \gamma_\infty$, then $\sigma(h) = \gamma_\infty - \gamma(h)$. An unbounded variogram (e.g. $\gamma(h) = \|h\|$) has *no* covariogram.

The hierarchy.

$$\text{strict} \xrightarrow{\mathbb{E}X^2 < \infty} \text{weak} \implies \text{intrinsic}, \qquad \text{all implications strict;} \quad \text{Gaussian: strict} \Leftrightarrow \text{weak.}$$

Conditional negative definiteness

Proposition 11

If X is intrinsically stationary, then for all n, sites $s_1, \dots, s_n$ and $a \in \mathbb{R}^n$ with $\sum_i a_i = 0$,

$$\mathbb{V}\Big(\sum_i a_i X(s_i)\Big) = -\sum_i \sum_j a_i a_j \, \gamma(s_i - s_j) \;\geq\; 0.$$

Proof. Put $Y_i := X(s_i) - X(s_1)$; since $\sum a_i = 0$, $\sum a_i X(s_i) = \sum a_i Y_i$, which has mean zero. Expanding,

$$\sum_{i,j} a_i a_j (Y_i - Y_j)^2 = 2\Big(\sum_j a_j\Big)\Big(\sum_i a_i Y_i^2\Big) - 2\Big(\sum_i a_i Y_i\Big)^2 = -2\Big(\sum_i a_i Y_i\Big)^2.$$

Take expectations, using $\mathbb{E}(Y_i - Y_j)^2 = \mathbb{E}(X(s_i) - X(s_j))^2 = 2\gamma(s_i - s_j)$. $\square$

- γ is *conditionally negative semi-definite*; not every non-negative even function is a variogram.
- Linear combinations with $\sum a_i = 0$ (*contrasts*) have well-defined variances even when $\mathbb{V}X(s)$ does not exist. Ordinary kriging errors are exactly such contrasts (Lecture 5).

Isotropy and anisotropy

Definition 12 (Isotropy)

A weakly stationary field is *isotropic* if $\sigma(h) = \tilde{\sigma}(\|h\|)$ for some $\tilde{\sigma} : [0, \infty) \to \mathbb{R}$; similarly $\gamma(h) = \tilde{\gamma}(\|h\|)$ for an intrinsically stationary field. We then write $\sigma(r)$ for $\tilde{\sigma}(r)$.

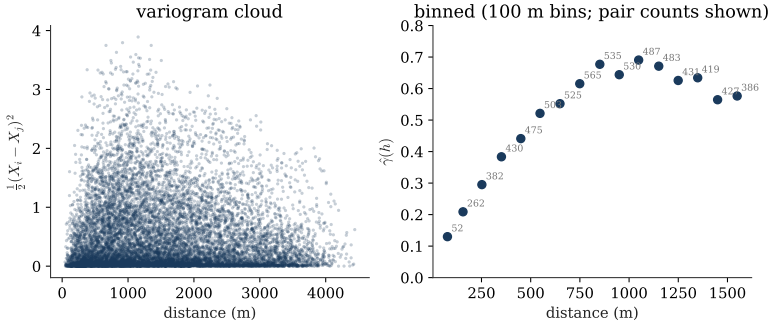
- Dependence decays identically in every direction: the contours of σ are circles (spheres).
- Isotropy is a further assumption, often violated by physical structure: river valleys, prevailing winds, geological layering.

Definition 13 (Geometric anisotropy)

$\sigma(h) = \tilde{\sigma}(\|Ah\|)$ for an invertible matrix A and an isotropic $\tilde{\sigma}$.

- Contours are ellipses; the field becomes isotropic after the linear change of coordinates $s \mapsto As$.
- Parameterisation, estimation and the more general *zonal* anisotropy: Lecture 2.

Meuse: a first look at the variogram



- **Left:** the *variogram cloud*, $\frac{1}{2}(X_i - X_j)^2$ against $\|s_i - s_j\|$ for all $\binom{155}{2} = 11\,935$ pairs of log-zinc values; under intrinsic stationarity each point has expectation $\gamma(s_i - s_j)$.
- **Right:** averaging within distance bins. The curve rises and levels off near 0.6 at about $1\,000$ m: strong spatial dependence that vanishes at large separations. It does not start at 0: a possible nugget (slide 22).
- How to bin, average and fit a model properly: Lecture 3.

Valid covariograms: positive semi-definiteness

Definition 14

An even function $\sigma : \mathbb{R}^d \to \mathbb{R}$ is *positive semi-definite (p.s.d.)* if for all n, sites $s_1, \dots, s_n$ and $a \in \mathbb{R}^n$

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

i.e. every matrix $\Sigma = (\sigma(s_i - s_j))_{ij}$ is p.s.d. We call such σ a *valid covariogram*.

Proposition 15

σ is the covariogram of some weakly stationary field if and only if it is p.s.d.

Proof. ($\Rightarrow$) $0 \leq \mathbb{V}\big(\sum_i a_i X(s_i)\big) = \sum_{i,j} a_i a_j \sigma(s_i - s_j)$.
($\Leftarrow$) For each finite set of sites take $\mathrm{Gaussian}(0, \Sigma)$ with $\Sigma_{ij} = \sigma(s_i - s_j)$; this family is consistent under permutation and marginalisation, so by Kolmogorov it defines a (Gaussian, weakly stationary) field with covariogram σ. $\square$

Every valid covariogram is the covariogram of a Gaussian field. Validity must be checked in the dimension d of the domain.

Why validity is not obvious

A plausible-looking function that fails. Let $\sigma(h) = 1$ for $\|h\| \le 1$ and 0 otherwise ("correlated within distance 1"). Take three sites on a line at $0,\ 0.6,\ 1.2$:

$$\Sigma = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \qquad \det \Sigma = -1 < 0.$$

So Σ has a negative eigenvalue: some linear combination would have *negative variance*.

Further warnings.

- The tent $\sigma(h) = \max\{0, 1 - \|h\|\}$ is valid on $\mathbb{R}^1$ but not on $\mathbb{R}^2$. [Sheet I, Q5]
- A non-valid model leads to kriging systems with negative "variances" and nonsensical predictions.
- Checking p.s.d. directly needs *all* n and *all* configurations, which is impossible in practice.

Strategy. Never invent a covariance by intuition. Either use a *characterisation* (Bochner, next) or a *construction* from known valid pieces (closure properties, slides 20–21; convolution, Lecture 4).

Bochner's theorem

Theorem 16 (Bochner)

A continuous function $\sigma : \mathbb{R}^d \to \mathbb{R}$ is p.s.d. if and only if

$$\sigma(h) = \int_{\mathbb{R}^d} e^{i\langle \omega, h\rangle} \, F(\mathrm{d}\omega)$$

for a finite, non-negative, symmetric measure F on $\mathbb{R}^d$, the spectral measure.

- If F has a density f (the *spectral density*), $\sigma(h) = \int e^{i\langle \omega, h\rangle} f(\omega)\,\mathrm{d}\omega$; if moreover σ is integrable,

$$f(\omega) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i\langle \omega, h\rangle} \sigma(h)\,\mathrm{d}h.$$

- **Practical test:** a continuous, integrable σ is valid iff its Fourier transform is non-negative everywhere.

Proof of "$\Leftarrow$". $\displaystyle\sum_{j,k} a_j a_k \sigma(s_j - s_k) = \int \Big| \sum_j a_j e^{i\langle \omega, s_j\rangle} \Big|^2 F(\mathrm{d}\omega) \;\geq 0.$ "$\Rightarrow$" is more involved.

Equivalently, $\sigma/\sigma(0)$ is the characteristic function of the probability measure $F/F(\mathbb{R}^d)$.

The spectral view of a stationary field

Theorem 17 (Spectral representation)

A weakly stationary, mean-square continuous field with spectral measure F can be written

$$X(s) = \mu + \int_{\mathbb{R}^d} e^{i\langle \omega, s\rangle}\, Z(\mathrm{d}\omega),$$

where Z is a random measure with orthogonal increments and $\mathbb{E}|Z(\mathrm{d}\omega)|^2 = F(\mathrm{d}\omega)$.

- X is a superposition of sinusoids with random amplitudes and phases; $F(\mathrm{d}\omega)$ is the variance contributed by frequencies near ω.
- Total variance: $\sigma(0) = F(\mathbb{R}^d)$.
- Simplest case, one frequency: $X(s) = A\cos(\omega s) + B\sin(\omega s)$ has F concentrated at $\pm\omega$. [Sheet I, Q1a]
- Much mass at high frequencies (heavy spectral tails) means much small-scale variation and rough realisations. This links spectral tails to smoothness (Lecture 2).

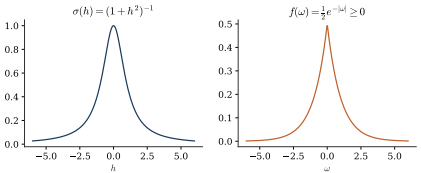
Example: the Cauchy covariogram

Claim. $\sigma(h) = (1+h^2)^{-1}$ is a valid covariogram on $\mathbb{R}^1$, with spectral density $f(\omega) = \frac{1}{2}e^{-|\omega|}$.

Proof. Compute the inverse transform of f:

$$\int_{-\infty}^{\infty} e^{i\omega h}\,\tfrac{1}{2}e^{-|\omega|}\,\mathrm{d}\omega$$
$$= \tfrac{1}{2}\Big[\int_0^{\infty} e^{-(1-ih)\omega}\,\mathrm{d}\omega + \int_0^{\infty} e^{-(1+ih)\omega}\,\mathrm{d}\omega\Big]$$
$$= \tfrac{1}{2}\Big[\tfrac{1}{1-ih} + \tfrac{1}{1+ih}\Big] = \frac{1}{1+h^2}.$$

Since $f \geq 0$, Bochner gives validity. □



- Duality: exponential covariogram ↔ Cauchy spectral density, and Cauchy covariogram ↔ exponential spectral density.
- You will compute the spectral densities of the exponential, Gaussian and Matérn ($\nu = 3/2$) covariograms. [Sheet I, Q2]

Theorem 18 (Schoenberg)

A continuous function $\sigma(h) = \sigma(\|h\|)$ is p.s.d. on $\mathbb{R}^d$ iff

$$\sigma(r) = \int_0^\infty \Omega_d(rt)\, G(\mathrm{d}t), \qquad \Omega_d(x) = \Gamma(\tfrac{d}{2})\Big(\frac{2}{x}\Big)^{\frac{d-2}{2}} J_{\frac{d-2}{2}}(x),$$

for a finite non-negative measure G on $[0,\infty)$; J_ν is the Bessel function of the first kind.

- $\Omega_1(x) = \cos x, \quad \Omega_2(x) = J_0(x), \quad \Omega_3(x) = \sin x / x.$
- Since σ is a mixture of Ω_d's: $\sigma(r)/\sigma(0) \geq \inf_x \Omega_d(x)$, which equals -1 ($d=1$), -0.403 ($d=2$), -0.218 ($d=3$). Strong negative correlation is impossible in higher dimensions.
- The isotropic spectral density is a *Hankel transform*:

$$f(\omega) = \frac{1}{(2\pi)^{\frac{d}{2}} \|\omega\|^{\frac{d-2}{2}}} \int_0^\infty \sigma(r)\, J_{\frac{d-2}{2}}(\|\omega\| r)\, r^{\frac{d}{2}}\, \mathrm{d}r. \qquad \text{[used in Sheet I, Q5b]}$$

Validity depends on dimension

Proposition 19

If $\sigma(\|\cdot\|)$ is a valid covariogram on $\mathbb{R}^d$, it is valid on $\mathbb{R}^{d'}$ for every $d' < d$.

Proof. Embed $\mathbb{R}^{d'} \hookrightarrow \mathbb{R}^d$ by $s \mapsto (s, 0)$, which preserves distances. For sites $s_1, \dots, s_n \in \mathbb{R}^{d'}$,

$$\sum_{i,j} a_i a_j \sigma(\|s_i - s_j\|) = \sum_{i,j} a_i a_j \sigma\big(\|(s_i, 0) - (s_j, 0)\|\big) \geq 0. \qquad \square$$

Write Φ_d for the isotropic functions valid on $\mathbb{R}^d$. Then

$$\Phi_1 \supset \Phi_2 \supset \Phi_3 \supset \cdots \supset \Phi_\infty := \bigcap\nolimits_d \Phi_d.$$

- In Φ_∞ (valid in every dimension): exponential, Gaussian, Matérn, Cauchy.
- Spherical: in Φ_3 but not Φ_4 (Lecture 2).
- Tent: in Φ_1 but not Φ_2. It arises naturally as the covariogram of the increment process $B(s) - B(s+1)$ of Brownian motion. [Sheet I, Q1b(iii), Q5]

Moral: a model taken from a one-dimensional (time-series) context may be invalid on the plane.

Building covariograms I: combinations and limits

Proposition 20

If σ_1, σ_2 are valid covariograms on $\mathbb{R}^d$ and $a_1, a_2 \geq 0$, then $a_1\sigma_1 + a_2\sigma_2$ is valid.

Proof. $\sum_{i,j} c_i c_j (a_1\sigma_1 + a_2\sigma_2)(s_i - s_j) = a_1 \sum c_i c_j \sigma_1(\cdot) + a_2 \sum c_i c_j \sigma_2(\cdot) \geq 0$. $\square$
Construction: if X_1, X_2 are independent with covariograms σ_1, σ_2, then $\sqrt{a_1}X_1 + \sqrt{a_2}X_2$ has covariogram $a_1\sigma_1 + a_2\sigma_2$.

- **Limits:** if σ_n are valid and $\sigma_n \to \sigma$ pointwise, σ is valid, since the inequalities pass to the limit.
- **Mixtures:** $\int \sigma_\theta(h)\, \mu(\mathrm{d}\theta)$ is valid for any finite measure $\mu \geq 0$ (limits of non-negative combinations).
- **Nested structures** are common in practice: variation at two scales,

$$\sigma(h) = \sigma_0^{(1)} \rho(\|h\|/\phi_1) + \sigma_0^{(2)} \rho(\|h\|/\phi_2), \qquad \phi_1 \ll \phi_2.$$

Building covariograms II: products and transformations

Proposition 21

If σ_1, σ_2 are valid covariograms on $\mathbb{R}^d$, so is $\sigma_1\sigma_2$.

Proof. Let X_1, X_2 be independent, zero-mean, weakly stationary, with covariograms σ_1, σ_2 (they exist by slide 13). Put $X := X_1 X_2$. Then $\mathbb{E}X(s) = 0$ and

$$\mathbb{C}\big(X(s), X(s+h)\big) = \mathbb{E}\big[X_1(s)X_1(s+h)\big]\,\mathbb{E}\big[X_2(s)X_2(s+h)\big] = \sigma_1(h)\sigma_2(h). \qquad \square$$

Spectral view: the spectral measure of $\sigma_1\sigma_2$ is the convolution $F_1 * F_2 \geq 0$.

Proposition 22

If σ is valid on $\mathbb{R}^d$ and A is any $d \times d$ matrix, then $h \mapsto \sigma(Ah)$ is valid. In particular $\sigma(h/\phi)$ is valid for $\phi > 0$.

Proof. $\sum_{i,j} a_i a_j \sigma(As_i - As_j) \geq 0$, because it is the p.s.d. sum for the sites $As_1, \ldots, As_n$. $\square$

- Rescaling introduces the *range parameter*; linear maps give *geometric anisotropy*.
- Spectral version of rescaling, and products over separate coordinates (separable models), in [Sheet I, Q3].

The nugget effect and the standard parametric form

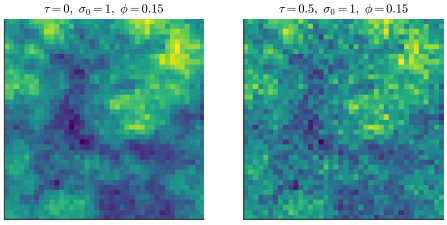
White noise. $\sigma(h) = \mathbb{I}(h = 0)$ is p.s.d. (every Σ is the identity): the covariogram of i.i.d. noise. It is *discontinuous* at 0, so Bochner's theorem does not apply.

Standard form. Combining a nugget with a rescaled correlation function ρ ($\rho(0) = 1$):

$$\sigma(h) = \tau\, \mathbb{I}(h = 0) + \sigma_0\, \rho(\|h\|/\phi).$$

- **nugget** $\tau \geq 0$: measurement error and/or variation below the sampling resolution;
- **semisill** $\sigma_0 > 0$: variance of the spatially structured part;
- **range parameter** $\phi > 0$: the distance scale;
- **sill** $\tau + \sigma_0 = \mathbb{V}X(s)$.

Variogram, *practical range*, and a labelled sketch: [Sheet I, Q4].



Same white noise, exponential ρ. The nugget adds unstructured noise on top of the smooth spatial pattern.

Simulating a Gaussian random field

Cholesky method. To simulate $X = (X(s_1), \dots, X(s_n))^\top \sim \mathrm{Gaussian}(\mu, \Sigma)$:

1. form $\Sigma_{ij} = \sigma(s_i - s_j)$ and its Cholesky factor $\Sigma = LL^\top$ (L lower triangular);
2. draw $Z \sim \mathrm{Gaussian}(0, I_n)$ and set $X = \mu + LZ$.

Then $\mathbb{E}X = \mu$ and $\mathbb{C}(X) = L\,\mathbb{C}(Z)\,L^\top = LL^\top = \Sigma$; X is Gaussian as a linear map of Z.

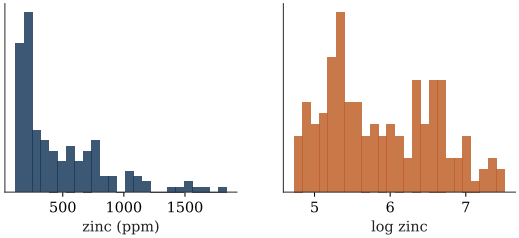
Costs and pitfalls.

- $O(n^3)$ operations and $O(n^2)$ memory: $n \approx 10^4$ is feasible, a 1000×1000 grid ($n = 10^6$) is not.
- Smooth covariances give nearly singular Σ; add a tiny "jitter" ϵI.
- Large grids: circulant embedding (FFT), spectral methods, or sparse (Markov) approximations (Section 3).

In R.

```
g  <- seq(0, 1, length = 40)
s  <- expand.grid(x = g, y = g)
D  <- as.matrix(dist(s))
S  <- exp(-D / 0.1)          # exponential
L  <- t(chol(S + 1e-8 * diag(nrow(s))))
x  <- L %*% rnorm(nrow(s))
image(g, g, matrix(x, 40, 40))
```

Meuse: what we have learned, and what comes next



Raw zinc is strongly right-skewed; log zinc is roughly symmetric, so a Gaussian model for $\log(\text{zinc})$ is a reasonable working assumption.

Working model: $\log$ zinc is one realisation of a Gaussian random field

$$X(s) = \mu(s) + Z(s), \qquad Z \text{ weakly stationary.}$$

Evidence so far.

- The binned variogram levels off near 0.6: a finite sill, consistent with weak stationarity.
- It is positive at short lags: a nugget is plausible.

Open questions.

- Which correlation function ρ, and how smooth is the field? (L2)
- How should γ be estimated and a model fitted? (L3)
- Is μ really constant, given the distance to the river? (L4, L6)

Summary of Lecture 1

- A random field is specified by its finite-dimensional distributions; a Gaussian field by its mean and covariance functions.
- One realisation means we need structure: strict $\Rightarrow$ weak $\Rightarrow$ intrinsic stationarity, with $\gamma(h) = \sigma(0) - \sigma(h)$ and γ conditionally negative semi-definite.
- Valid covariograms are exactly the p.s.d. functions. Bochner: valid $\Leftrightarrow$ non-negative spectral measure. Validity depends on the dimension.
- Closure: non-negative combinations, limits, mixtures, products and linear transformations of valid covariograms are valid. The standard form is $\tau\mathbb{I}(h = 0) + \sigma_0\rho(\|h\|/\phi)$.

Exercise Sheet I (next session): stationarity counterexamples, Brownian motion, spectral densities, closure via Fourier transforms, the nugget and practical range, and the tent covariogram in $\mathbb{R}^2$.

Next: Lecture 2 — smoothness and parametric covariance families.