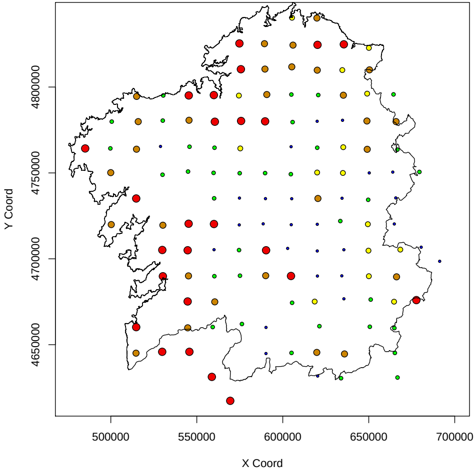
Lecture 2

The linear geostatistical model

1. Formulation of linear geostatistical models and assumptions.
2. Brief introduction to Gaussian processes.
3. Understanding the nugget effect.
4. Parameter estimation via the maximum likelihood method.

Where we observe matters



Geostatistical lead pollution in Galicia

Geostatistical lead pollution in Galicia

- Y_i = lead concentration in a moss sample (ppm)

Geostatistical lead pollution in Galicia

- Y_i = lead concentration in a moss sample (ppm)

$$Y_i = \alpha + S(x_i) + Z_j$$

Geostatistical lead pollution in Galicia

- Y_i = lead concentration in a moss sample (ppm)

$$Y_i = \alpha + S(x_i) + Z_j$$

- $Z_i \sim N(0, \tau^2)$

Geostatistical lead pollution in Galicia

- $Y_i =$ lead concentration in a moss sample (ppm)

$$Y_i = \alpha + S(x_i) + Z_j$$

- $Z_i \sim N(0, \tau^2)$
- We assume that $S(x)$ is a stationary and isotropic Gaussian process, i.e.

$$Cov(S(x), S(x')) = \sigma^2 \rho(u), \quad u = \|x - x'\|.$$

Geostatistical lead pollution in Galicia

- Y_i = lead concentration in a moss sample (ppm)

$$Y_i = \alpha + S(x_i) + Z_j$$

- $Z_i \sim N(0, \tau^2)$
- We assume that $S(x)$ is a stationary and isotropic Gaussian process, i.e.

$$Cov(S(x), S(x')) = \sigma^2 \rho(u), \quad u = \|x - x'\|.$$

- How do we choose $\rho(\cdot)$?

Geostatistical lead pollution in Galicia

- Y_i = lead concentration in a moss sample (ppm)

$$Y_i = \alpha + S(x_i) + Z_j$$

- $Z_i \sim N(0, \tau^2)$
- We assume that $S(x)$ is a stationary and isotropic Gaussian process, i.e.

$$Cov(S(x), S(x')) = \sigma^2 \rho(u), \quad u = \|x - x'\|.$$

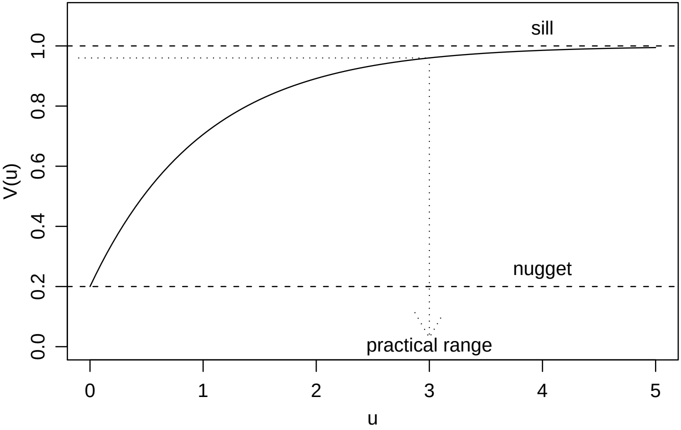
- How do we choose $\rho(\cdot)$?
- Example: $\rho(u) = \exp\{-u/\phi\}$

The theoretical variogram (continued)

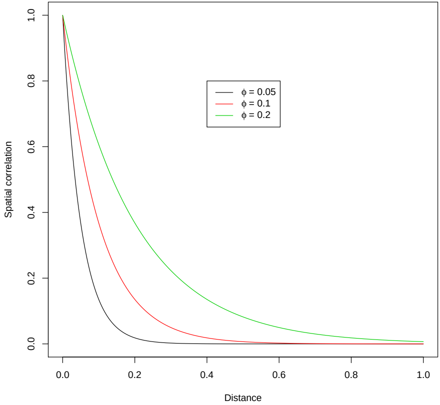
$$v(u) = \tau^2 + \sigma^2(1 - \rho(u))$$

The theoretical variogram (continued)

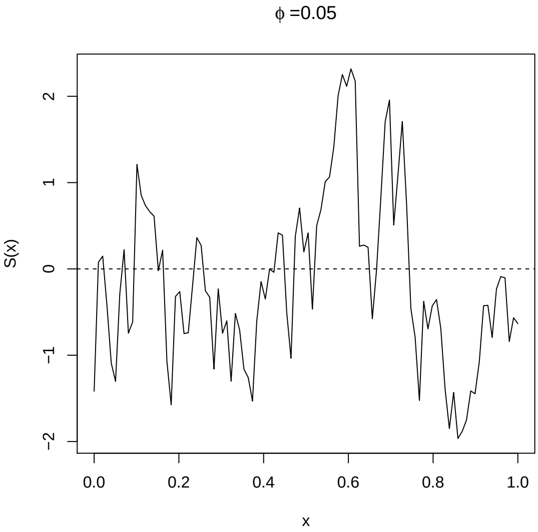
$$v(u) = \tau^2 + \sigma^2(1 - \rho(u))$$



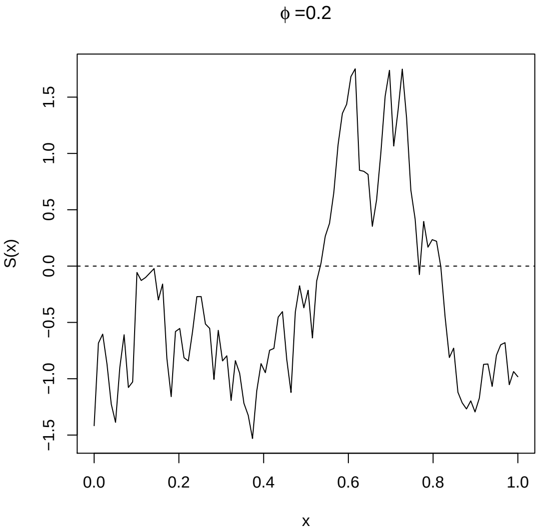
The scale of the spatial correlation



The scale of the spatial correlation



The scale of the spatial correlation



The Matérn correlation function

The Matérn correlation function

$$\rho(u) = 2^{\kappa-1}(u/\phi)^{\kappa} K_{\kappa}(u/\phi)$$

The Matérn correlation function

$$\rho(u) = 2^{\kappa-1}(u/\phi)^{\kappa} K_{\kappa}(u/\phi)$$

- $K_{\kappa}(\cdot)$: modified Bessel function of order κ

The Matérn correlation function

$$\rho(u) = 2^{\kappa-1}(u/\phi)^{\kappa} K_{\kappa}(u/\phi)$$

- $K_{\kappa}(\cdot)$: modified Bessel function of order κ
- Interpretation

The Matérn correlation function

$$\rho(u) = 2^{\kappa-1}(u/\phi)^{\kappa} K_{\kappa}(u/\phi)$$

- $K_{\kappa}(\cdot)$: modified Bessel function of order κ
- Interpretation
  - κ determines smoothness of underlying Gaussian process

$$\kappa > r \rightarrow S(x) \text{ is } r \text{ times differentiable}$$

The Matérn correlation function

$$\rho(u) = 2^{\kappa-1}(u/\phi)^{\kappa} K_{\kappa}(u/\phi)$$

- $K_{\kappa}(\cdot)$: modified Bessel function of order κ
- Interpretation
  - κ determines smoothness of underlying Gaussian process

    $\kappa > r \rightarrow S(x)$ is r times differentiable

  - ϕ determines scale of the spatial correlation

The Matérn correlation function

$$\rho(u) = 2^{\kappa-1}(u/\phi)^{\kappa} K_{\kappa}(u/\phi)$$

- $K_{\kappa}(\cdot)$: modified Bessel function of order κ
- Interpretation
  - κ determines smoothness of underlying Gaussian process

    $$\kappa > r \rightarrow S(x) \text{ is } r \text{ times differentiable}$$

  - ϕ determines scale of the spatial correlation
- Special cases
  - $= 0.5$ gives $\rho(u) = \exp\{u/\phi\}$
  - $k \rightarrow \infty$ gives $\rho(u) = \exp\{(u/\phi)^2\}$

The Matérn correlation function

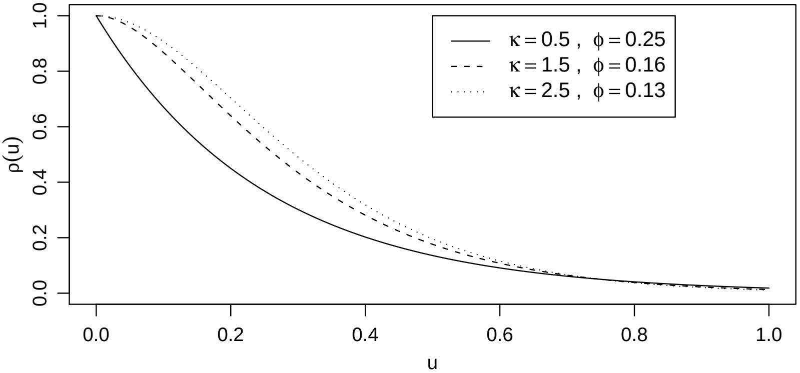
$$\rho(u) = 2^{\kappa-1}(u/\phi)^{\kappa} K_{\kappa}(u/\phi)$$

- $K_{\kappa}(\cdot)$: modified Bessel function of order κ
- Interpretation
  - κ determines smoothness of underlying Gaussian process

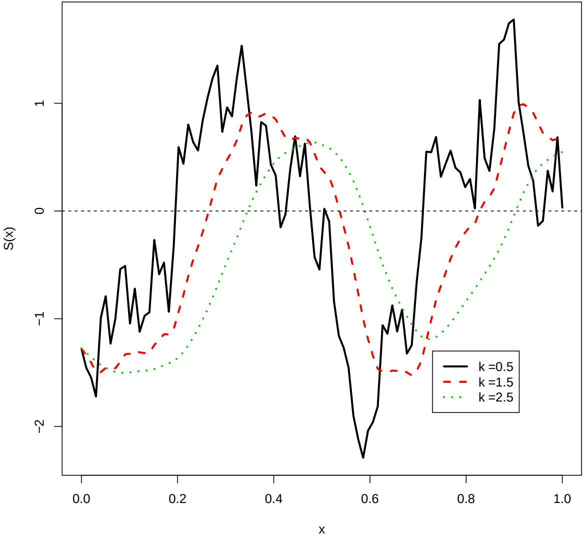
$$\kappa > r \rightarrow S(x) \text{ is } r \text{ times differentiable}$$

  - ϕ determines scale of the spatial correlation
- Special cases
  - $= 0.5$ gives $\rho(u) = \exp\{u/\phi\}$
  - $k \rightarrow \infty$ gives $\rho(u) = \exp\{(u/\phi)^2\}$
- Often sufficient to choose amongst $\kappa = 0.5, 1.5, 2.5$

The Matérn correlation function



The Matérn correlation function



Interpreting the nugget effect

Interpreting the nugget effect

1. Measurement error

Interpreting the nugget effect

1. Measurement error
2. Small scale spatial correlation

Interpreting the nugget effect

1. Measurement error
2. Small scale spatial correlation

$$\begin{aligned} Y_i &= \alpha + S(x_i) + Z(x_i) \\ \mathrm{cor}\{Z(x), Z(x')\} &= \begin{cases} \delta & \text{if } \|x - x'\| < u^* \\ 0 & \text{if } \|x - x'\| \geq u^* \end{cases} \end{aligned}$$

Interpreting the nugget effect

1. Measurement error
2. Small scale spatial correlation

$$
\begin{aligned}
Y_i &= \alpha + S(x_i) + Z(x_i) \\
\mathrm{cor}\{Z(x), Z(x')\} &= \begin{cases} \delta & \text{if } \|x - x'\| < u^* \\ 0 & \text{if } \|x - x'\| \geq u^* \end{cases}
\end{aligned}
$$

If $u^* < u_{min}$ then $Z(x_i)$ is pure noise.

Getting initial parameter estimates

- Widely used, but not recommended except for initial analysis.
- $\theta = (\sigma^2, \phi, \tau^2)$
- Weighted least squares criterion:

$$W(\theta) = \sum_k n_k [\hat{v}(u_k) - v(u_k; \theta)]^2$$

Getting initial parameter estimates

- Widely used, but not recommended except for initial analysis.
- $\theta = (\sigma^2, \phi, \tau^2)$
- Weighted least squares criterion:

$$W(\theta) = \sum_k n_k [\hat{v}(u_k) - v(u_k; \theta)]^2$$

- Arbitrary binning and upper limit for u_k.

Getting initial parameter estimates

- Widely used, but not recommended except for initial analysis.
- $\theta = (\sigma^2, \phi, \tau^2)$
- Weighted least squares criterion:

$$W(\theta) = \sum_k n_k[\hat{v}(u_k) - v(u_k;\theta)]^2$$

- Arbitrary binning and upper limit for u_k.
- Standard errors not available.

Maximum likelihood estimation

Maximum likelihood estimation

- Multivariate Gaussian distribution: $Y \sim MVN(D\beta, \sigma^2 R + \tau^2 I)$.
  - D matrix of covariates: $[D]_{ik} = d_k(x_i)$
  - R matrix of spatial correlation: $[R]_{ij} = \rho(u_{ij})$, with $u_{ij} = \|x_i - x_j\|$.

Maximum likelihood estimation

- Multivariate Gaussian distribution: $Y \sim MVN(D\beta, \sigma^2 R + \tau^2 I)$.
  - D matrix of covariates: $[D]_{ik} = d_k(x_i)$
  - R matrix of spatial correlation: $[R]_{ij} = \rho(u_{ij})$, with $u_{ij} = \|x_i - x_j\|$.
- Fitting process
  1. Initialise β, e.g. using ordinary least squares.
  2. Initialise θ, e.g. using the empirical variogram
  3. Maximize

     $$l(\theta) = \log\{f(y; \beta, \theta)\}$$

     where $f(\cdot; \beta; \theta)$ denotes the density of the multivariate Gaussian distribution.