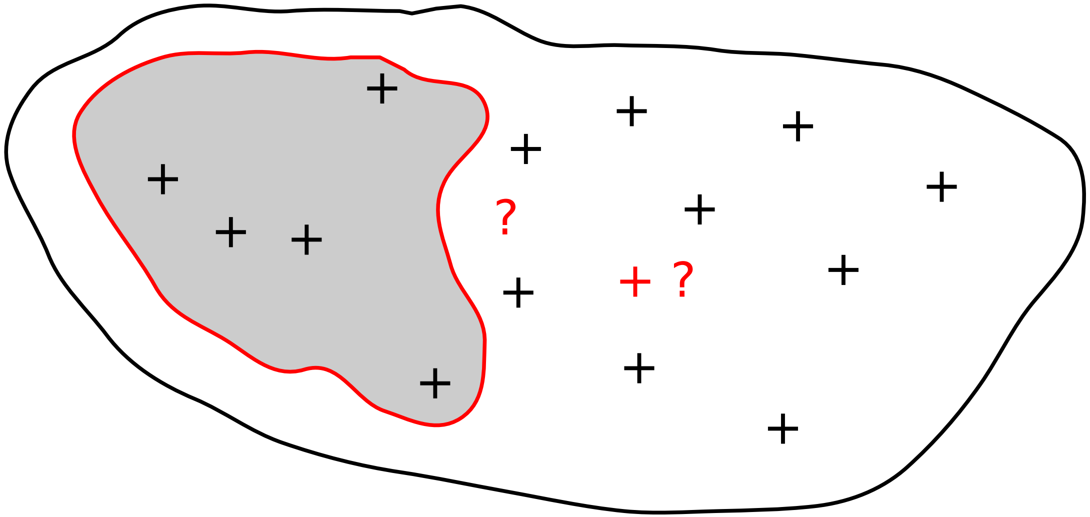
Regionalized variable Z:
- Temperature.
- Pollutant concentration.
- Content of an ore.
- ...



Main questions:

- What are the values where no measurements (interpolation / mean estimation)?
- What is the error associated with these estimates?

→ kriging (from G. Matheron after D. Krige, SA mining engineer)

Outline:

1. Simple kriging
2. Ordinary kriging
3. Universal kriging + kriging with an external drift
4. Properties of kriging
5. Kriging data with uncertainty

Interpolation of spatial data (evenly spaced or not) based on existing knowledge of spatial variability of studied process, as quantified by the variogram (or covariance).

Notations:
Random function Z and regionalized variable z
(probabilistic calculations involve Z, numerical estimations involve z)

$$\{Z(x,\omega) : x \in D \subset \mathbb{R}^n; \omega \in \Omega\}$$

$$\{z(x) = Z(x,\omega_0) : x \in D\} \; \omega_0 \in \Omega$$

S = set of points where Z has been sampled

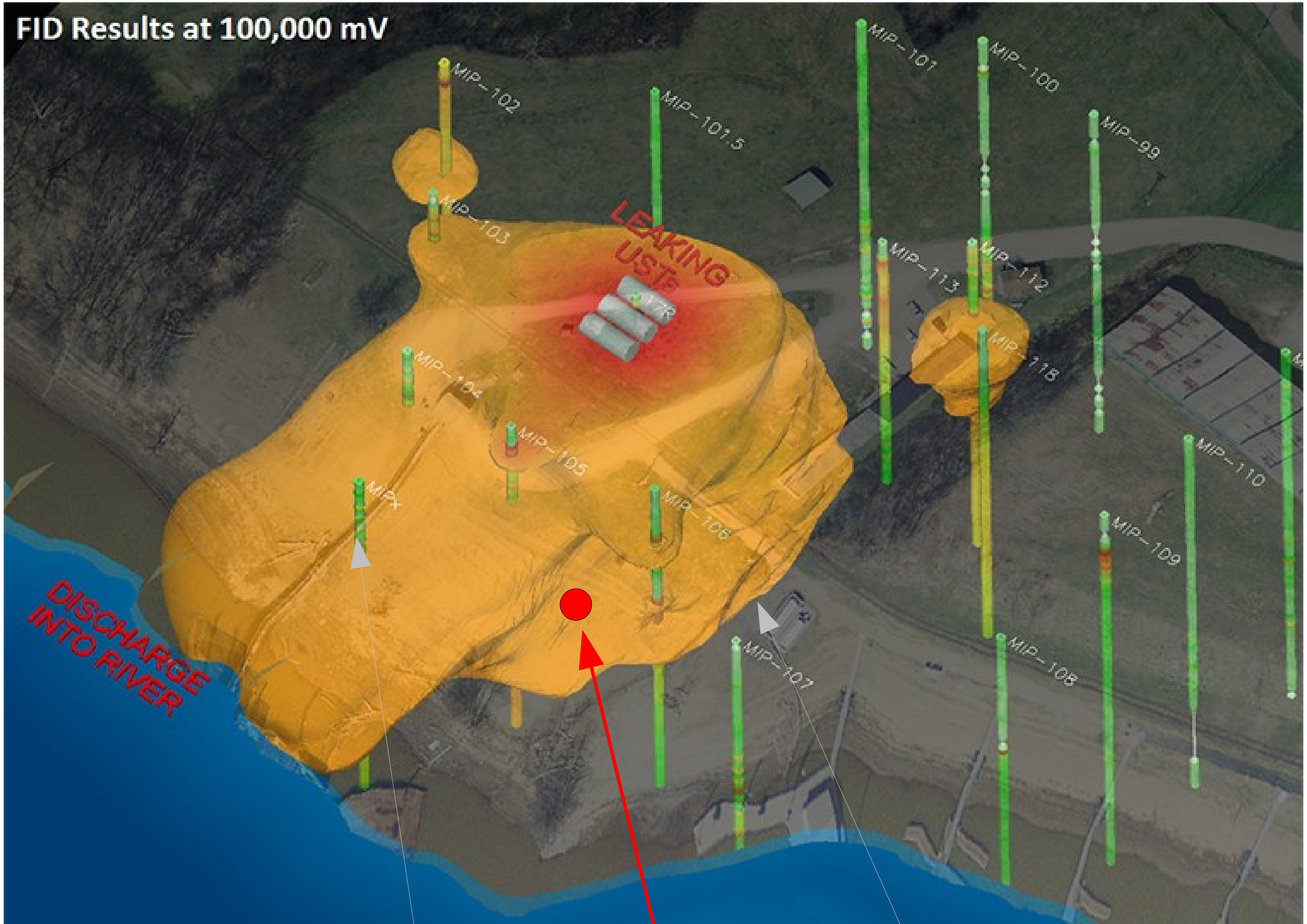
$$S = x_\alpha : \alpha = 1, .., N$$

To simplify notations, in the following: $Z_\alpha = Z(x_\alpha, \omega)$

Objective:
Estimate Z* of Z at x_0 (unknown) as a linear combination of measured $Z(x_\alpha)$

$$Z_0^* = \sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha$$

Conditions: Z* is (1) unbiased and with (2) minimal estimation variance https://www.ctech.com/user-showcase/st-john-mittelhauser-associates/



$Z(x_\alpha)$ $Z(x_0)$ $Z(x)$

$$Z_0^* = \sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha$$

We suppose that Z is a 2^{nd}-order SRF

Condition (1): unbiased estimator:

$$\mathrm{E}[Z_0^*] = \mathrm{E}\left[\sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha\right] = \sum_{\alpha=1}^{N} \lambda_\alpha \mathrm{E}[Z_\alpha] = \mathrm{E}[Z_0]$$

Condition (2): minimal estimation variance:

$$\mathrm{Var}[Z_0^* - Z_0] \text{ is min}$$

$$\begin{aligned} \epsilon_0 &= Z_0^* - Z_0 \\ \sigma^2_{\epsilon_0} &= \mathrm{Var}[Z_0^* - Z_0] \\ \sigma^2_{\epsilon_0} &= \sigma^2_Z + \sum_{\alpha=1}^{N} \sum_{\beta=1}^{N} \lambda_\alpha \lambda_\beta C_{\alpha\beta} - 2 \sum_{\alpha=1}^{N} \lambda_\alpha C_{0\alpha} \end{aligned} \quad \text{eq(1)}$$

$$(C_{\alpha\beta} = \mathrm{Cov}[Z_\alpha, Z_\beta])$$

Z is assumed to be 2nd order stationary, and m = E[Z] is known

m known is equivalent to m = 0: $Y(x) = Z(x) - m$

$$Y_0^* = \sum_{\alpha=1}^{N} \lambda_\alpha Y(x_\alpha) \Leftrightarrow Z_0^* = m + \sum_{\alpha=1}^{N} \lambda_\alpha \left[Z_\alpha - m\right]$$

→ in this section, E[Z] = 0

Condition (1): unbiased estimator: $\mathrm{E}[Z_0^* - Z_0] = 0 \;\; \forall \lambda_\alpha \, , \, \alpha \in [1..N]$

Condition (2): minimal estimation variance: $\mathrm{Var}[Z_0^* - Z_0]$ is min

λ_α are calculated so that $\sigma_{\varepsilon 0}^2$ is minimum → $\dfrac{\partial \sigma_{\epsilon_0}^2}{\partial \lambda_\alpha} = 0 \quad \forall \alpha \in [1..N]$

$$\sum_{\beta=1}^{N} \lambda_\beta C_{\alpha\beta} = C_{\alpha 0} \;\; \forall \alpha \in [1..N] \qquad \text{eq(2)}$$

Eq(2) in eq(1) → estimation variance: $\sigma_{\epsilon_0}^2 = \sigma_Z^2 - \sum_{\alpha=1}^{N} \lambda_\alpha C_{\alpha 0}$

$$\begin{bmatrix} \sigma_Z^2 & C_{12} & \cdots & C_{1N} \\ C_{21} & \sigma_Z^2 & \cdots & C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{N1} & C_{N2} & \ldots & \sigma_Z^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{N0} \end{bmatrix}$$

C λ C_0

Linear system of N equations with N unknowns

Unique solution if C is definite positive and sampled points are distinct

$$\lambda = \mathbf{C}^{-1}\mathbf{C}_0$$

Estimation variance: $\sigma_{\epsilon_0}^2 = \sigma_Z^2 - \lambda^t \mathbf{C_0}$

Example: RF Z, known at (x_1, x_2). We want to estimate Z at x_0

Simple kriging (m=0)

$$\begin{cases} C_{11}\lambda_1 + C_{12}\lambda_2 = C_{01} \\ C_{21}\lambda_1 + C_{22}\lambda_2 = C_{02} \end{cases}$$

By definition: $C_{12} = C_{21}$

Stationarity: $C_{11} = C_{22}$

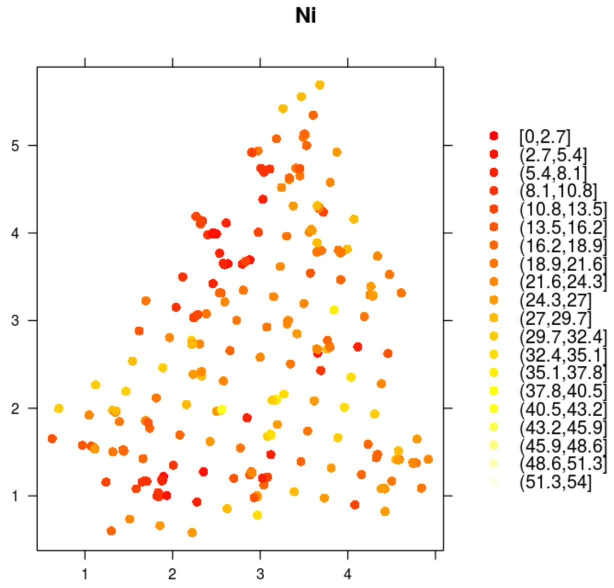
$$\begin{cases} \lambda_1 = \dfrac{C_{01}C_{11} - C_{02}C_{12}}{C_{11}^2 - C_{12}^2} \\ \lambda_2 = \dfrac{C_{02}C_{11} - C_{01}C_{12}}{C_{11}^2 - C_{12}^2} \end{cases}$$

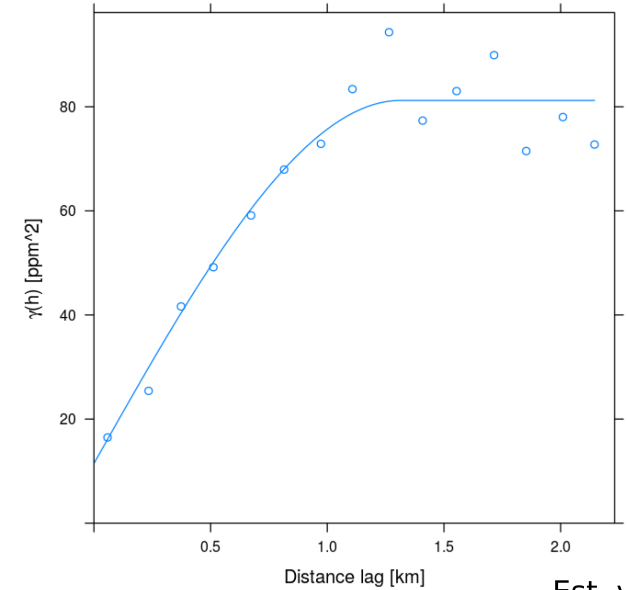
$$\begin{aligned} Z_0^* &= \lambda_1 Z_1 + \lambda_2 Z_2 \\ &= \frac{1}{C_{11}^2 - C_{12}^2}\left[C_{01}(C_{11}Z_1 - C_{12}Z_2) + C_{02}(-C_{12}Z_1 + C_{11}Z_2)\right] \end{aligned}$$

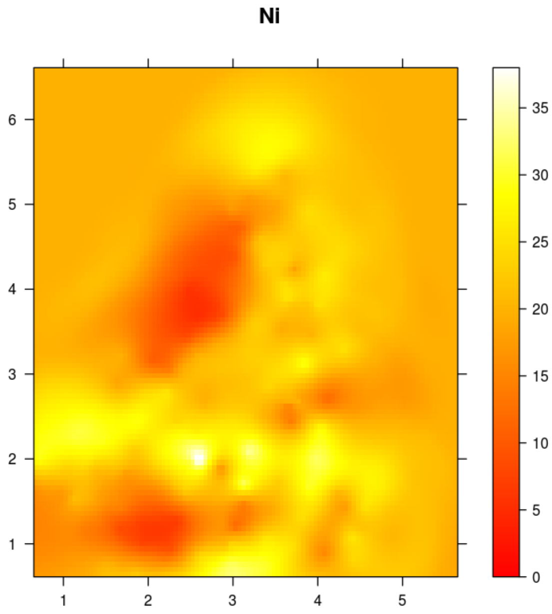
$$\sigma_{\epsilon_0}^2 = \sigma_Z^2 - \frac{1}{C_{11}^2 - C_{12}^2}(C_{01}^2 C_{11} + C_{02}^2 C_{11} - 2C_{01}C_{02}C_{12})$$

If Z_0 is not correlated with both Z_1 and Z_2 (i.e. $C_{01}=C_{02}=0$) then $\begin{cases} Z_0^* = 0 \\ \sigma_{\epsilon_0}^2 = \sigma_Z^2 \end{cases}$

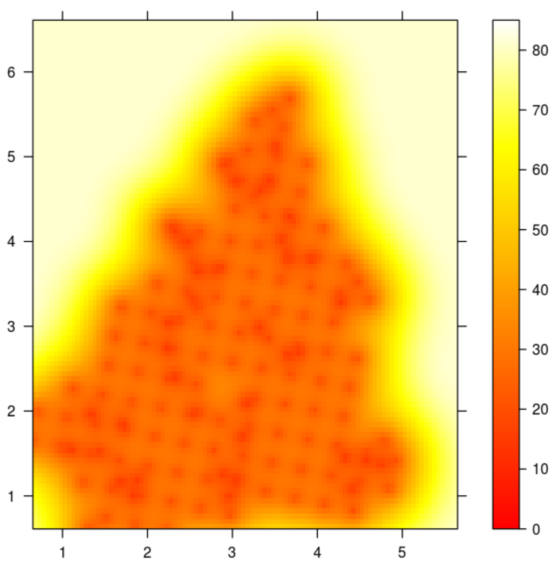
Example simple kriging (2)







Est. variance



Prescribed mean = sample mean

Z is assumed to be 2nd order stationary, and m = E[Z] is unknown

Z(x_0) is estimated as combination of known Z values: $Z_0^* = \sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha$

Condition (1): unbiased estimator

$$\mathrm{E}[Z_0^* - Z_0] = \mathrm{E}\left[\sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha - Z_0\right] = \mathrm{E}[Z]\left(\sum_{\alpha=1}^{N} \lambda_\alpha - 1\right)$$

$$\mathrm{E}[Z_0^* - Z_0] = 0 \;\; \forall \mathrm{E}[Z] \Leftrightarrow \sum_{\alpha=1}^{N} \lambda_\alpha = 1$$

Condition (2): minimal estimation variance $\lambda_\alpha \;/\; \left| \begin{array}{l} \sigma^2_{\epsilon_0} \text{ (given by eq(1)) is min} \\ \sum_{\alpha=1}^{N} \lambda_\alpha = 1 \end{array} \right.$

→ optimization with constraint: Lagrange multiplier

Minimization of an expression $G(\lambda_1,\lambda_2,\ldots,\lambda_N)$
With the constraint $F(\lambda_1,\lambda_2,\ldots,\lambda_N)$
Equivalent to minimization with additional unknown parameter ν

$$Q = G(\lambda_1, \ldots, \lambda_N) + 2\nu F(\lambda_1, \ldots, \lambda_N)$$

→ for OK

$$\mathrm{Q} = \sigma_Z^2 + \sum_{\alpha=1}^{N} \sum_{\beta=1}^{N} \lambda_\alpha \lambda_\beta C_{\alpha\beta} - 2 \sum_{\alpha=1}^{N} \lambda_\alpha C_{\alpha 0} + 2\nu \left(\sum_{\alpha=1}^{N} \lambda_\alpha - 1 \right)$$

→ derivatives with respect to λ_α AND ν

$$\frac{\partial Q}{\partial \lambda_\alpha} = 0 \quad \forall \alpha \in [1..N] \qquad \text{and} \qquad \frac{\partial Q}{\partial \nu} = 0$$

$$\begin{cases} \dfrac{\partial Q}{\partial \lambda_\alpha} = 2\sum_{\beta=1}^{N} \lambda_\beta C_{\alpha\beta} - 2C_{\alpha 0} + 2\nu \quad \forall \alpha \in [1..N] \\ \dfrac{\partial Q}{\partial \nu} = 2\left(\sum_{\alpha=1}^{N} \lambda_\alpha - 1\right) \end{cases}$$

which results in
$$\begin{cases} \sum_{\beta=1}^{N} \lambda_\beta C_{\alpha\beta} = C_{\alpha 0} - \nu \quad \forall \alpha \in [1..N] & \text{eq(3a)} \\ \sum_{\alpha=1}^{N} \lambda_\alpha = 1 & \text{eq(3b)} \end{cases}$$

Estimation variance: $\sigma^2_{\epsilon_0} = \sigma^2_Z - \sum_{\alpha=1}^{N} \lambda_\alpha C_{\alpha 0} - \nu$ eq(3ab) in eq(1)

$$\underbrace{\begin{bmatrix} \sigma_Z^2 & C_{12} & \dots & C_{1N} & 1 \\ C_{21} & \sigma_Z^2 & \dots & C_{2N} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{N1} & C_{N2} & \dots & \sigma_Z^2 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}}_{\text{A}} \underbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \nu \end{bmatrix}}_{\text{X}} = \underbrace{\begin{bmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{N0} \\ 1 \end{bmatrix}}_{\text{B}}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{C} & \mathbf{F} \\ \mathbf{F}^t & 0 \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} \lambda \\ \nu \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} \mathbf{C_0} \\ 1 \end{bmatrix}$$

Linear system of N+1 equations with N+1 unknowns

Unique solution if C is positive definite and sample points are distinct

Estimation variance: $\sigma_{\epsilon_0}^2 = \sigma_Z^2 - \mathbf{X}^t\mathbf{B} \qquad \sigma_{\epsilon_0}^2(SK) < \sigma_{\epsilon_0}^2(OK)$

Ordinary kriging (m unknown)

$$\begin{cases} C_{11}\lambda_1 + C_{12}\lambda_2 + \nu = C_{01} \\ C_{21}\lambda_1 + C_{22}\lambda_2 + \nu = C_{02} \\ \lambda_1 + \lambda_2 = 1 \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = \frac{1}{2}\left[1 + \frac{C_{01} - C_{02}}{C_{11} - C_{12}}\right] \\ \lambda_2 = \frac{1}{2}\left[1 + \frac{C_{02} - C_{01}}{C_{11} - C_{12}}\right] \\ \nu = \frac{1}{2}(C_{01} + C_{02} - C_{11} - C_{12}) \end{cases}$$

By definition: $C_{12} = C_{21}$

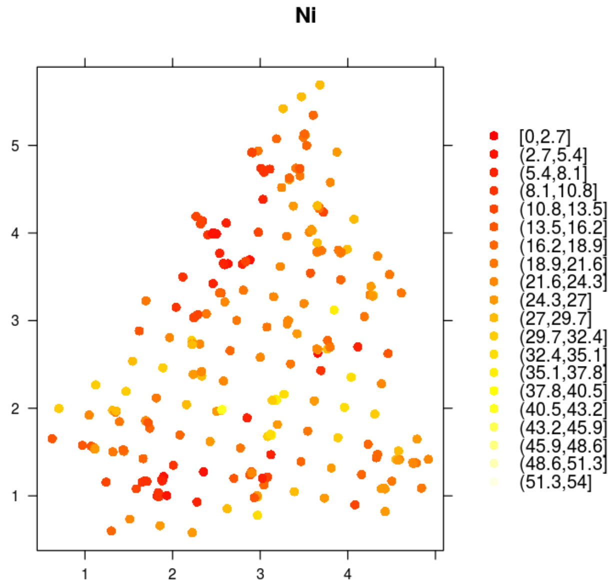
Stationarity: $C_{11} = C_{22}$

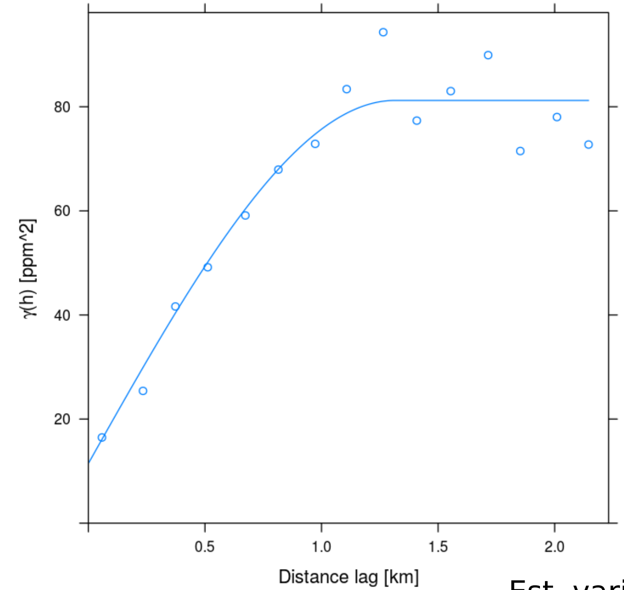
$$Z_0^* = \frac{Z_1 + Z_2}{2} + \frac{C_{01} - C_{02}}{2(C_{11} - C_{12})}(Z_1 - Z_2)$$

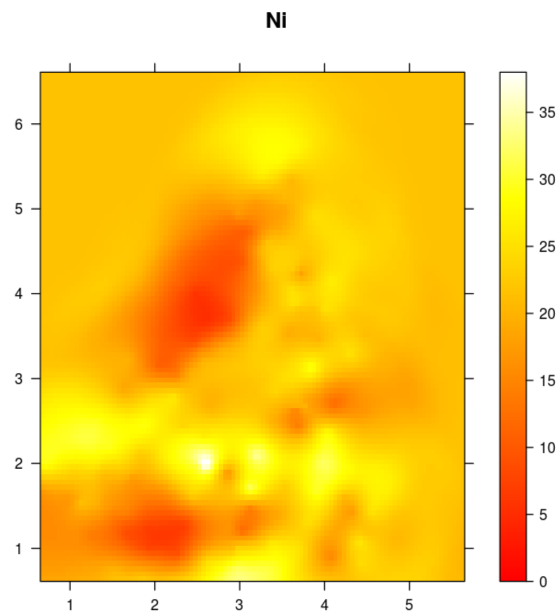
$$\sigma_{\epsilon_0}^2 = \sigma_Z^2 - C_{01} - C_{02} + \frac{1}{2}(C_{11} + C_{12}) - \frac{(C_{01} - C_{02})^2}{2(C_{11} - C_{12})}$$

If $C_{01}=C_{02}=0$ then $Z_0^* = \frac{Z_1 + Z_2}{2}$ $\sigma_{\epsilon_0}^2 = \sigma_Z^2 + \frac{C_{11} + C_{12}}{2}$

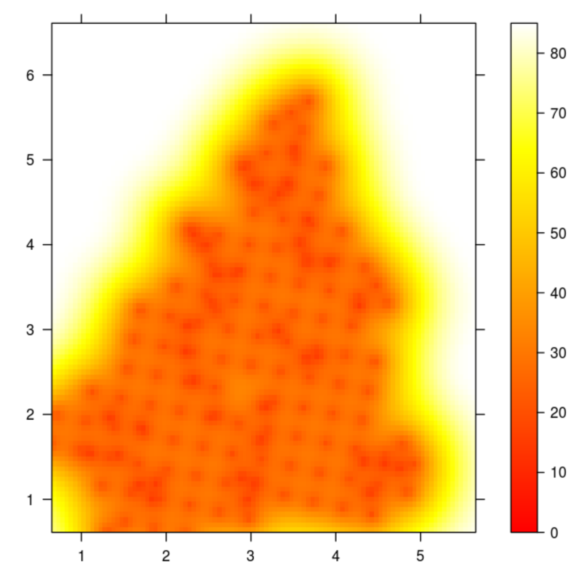
Example ordinary kriging (2)







Est. variance



Large sample (259 values)
→ sample mean is reliable
→ OK ~ SK

Z is assumed to be: $Z(x) = m(x) + Y(x)$

where *m* is a deterministic smooth function, called the drift
Y is a 0-mean 2nd order SRF, called the residual, describing the erratic fluctuations of Z

→ Z is not stationary in the mean and Cov_Z does not exist

We assume that *m* can be written as $m(x) = \sum_{l=0}^{L} a_l f^l(x)$

$f^l(x)$ are a priori known basis functions (e.g., polynomials)

a_l are unknown constant coefficients

$Z(x_0)$ is estimated as combination of known Z values: $Z_0^* = \sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha$

Similar approach to ordinary kriging:
- no bias.
- minimum estimation variance (least square sense).

Condition (1): unbiased estimator

$$\begin{aligned} \mathrm{E}[Z_0^* - Z_0] &= \sum_{\alpha=1}^{N} \lambda_\alpha m_\alpha - m_0 \\ &= \sum_{l=0}^{L} a_l \left(\sum_{\alpha=1}^{N} \lambda_\alpha f_\alpha^l - f_0^l \right) \qquad \text{where } f_\alpha^l = f^l(\alpha) \\ &= 0 \end{aligned}$$

no bias whatever the unknown a_l are, hence

$$\sum_{\alpha=1}^{N} \lambda_\alpha f_\alpha^l = f_0^l \quad l = 0..L$$

This equation defines the universality conditions

Condition (2): minimal estimation variance

$$\mathrm{Var}\left[Z_0^* - Z_0\right] = \mathrm{Var}\left[\sum_{\alpha=0}^{N} \lambda_\alpha Y_\alpha\right] \quad \text{where } \lambda_0 = -1$$

Minimal estimation variance $\lambda_\alpha \; / \; \left| \begin{array}{l} \sigma^2_{\epsilon_0} \text{ (given by eq(1) for } Y) \text{ is min} \\ \sum_{\alpha=1}^{N} \lambda_\alpha f^l_\alpha = f^l_0 \quad l = 0..L \end{array} \right.$

→ optimization with constraints: Lagrange multipliers

Minimization of an expression $G(\lambda_1, \lambda_2, \ldots, \lambda_N)$

With (L+1) constraints $F_l(\lambda_1, \lambda_2, \ldots, \lambda_N)$ $l=0..L$

Equivalent to minimize with additional unknown parameters ν_l

$$Q = G(\lambda_1, ..., \lambda_N) + 2\sum_{l=0}^{L} \nu_l F_l(\lambda_1, ..., \lambda_N)$$

→ for UK $Q = \sigma^2_Y + \sum_{\alpha=1} \sum_{\beta=1} \lambda_\alpha \lambda_\beta C^Y_{\alpha\beta}$

$$-2\sum_{\alpha=1}^{N} \lambda_\alpha C^Y_{\alpha 0} + 2\sum_{l=0}^{L} \nu_l \left(\sum_{\alpha=1}^{N} \lambda_\alpha f^l_\alpha - f^l_0 \right)$$

→ derivatives with respect to λ_α AND ν_l

$$\frac{\partial Q}{\partial \lambda_\alpha} = 0 \quad \forall \alpha \in [1..N] \qquad \text{and} \qquad \frac{\partial Q}{\partial \nu_l} = 0 \quad \forall\, l \in [0..L]$$

$$\begin{cases} \dfrac{\partial Q}{\partial \lambda_\alpha} = 2\sum_{\beta=1}^{N} \lambda_\beta C^Y_{\alpha\beta} - 2C^Y_{\alpha 0} + 2\sum_{l=0}^{L} \nu_l f^l_\alpha \quad \forall \alpha \in [1..N] \\ \dfrac{\partial Q}{\partial \nu_l} = 2\left[\sum_{\alpha} \lambda_\alpha f^l_\alpha - f^l_0\right] \quad \forall l \in [0..L] \end{cases}$$

which results in
$$\begin{cases} \sum_{\beta=1}^{N} \lambda_\beta C^Y_{\alpha\beta} = C^Y_{\alpha 0} - \sum_{l=0}^{L} \nu_l f^l_\alpha \quad \forall \alpha \in [1..N] & \text{eq(4a)} \\ \sum_{\alpha=1}^{N} \lambda_\alpha f^l_\alpha = f^l_0 \quad \forall l \in [0..L] & \text{eq(4b)} \end{cases}$$

Estimation variance: $\sigma^2_{\epsilon_0} = \sigma^2_Y - \sum_{\alpha=1}^{N} \lambda_\alpha C^Y_{\alpha 0} - \sum_{l=0}^{L} \nu_l f^l_0$ eq(4ab) in eq(1)

Matrix notation:

$$\underbrace{\begin{bmatrix} \sigma_Y^2 & C_{12}^Y & \cdots & C_{1N}^Y & f_1^0 & f_1^1 & \cdots & f_1^L \\ C_{21}^Y & \sigma_Y^2 & \cdots & C_{2N}^Y & f_2^0 & f_2^1 & \cdots & f_2^L \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{N1}^Y & C_{N2}^Y & \cdots & \sigma_Y^2 & f_N^0 & f_N^1 & \cdots & f_N^L \\ f_1^0 & f_2^0 & \cdots & f_N^0 & 0 & 0 & \cdots & 0 \\ f_1^1 & f_2^1 & \cdots & f_N^1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_1^L & f_2^L & \cdots & f_N^L & 0 & 0 & \cdots & 0 \end{bmatrix}}_{\text{A}} \underbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \nu_0 \\ \nu_1 \\ \vdots \\ \nu_L \end{bmatrix}}_{\text{X}} = \underbrace{\begin{bmatrix} C_{10}^Y \\ C_{20}^Y \\ \vdots \\ C_{N0}^Y \\ f_0^0 \\ f_0^1 \\ \vdots \\ f_0^L \end{bmatrix}}_{\text{B}}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{C^Y} & \mathbf{F} \\ \mathbf{F}^t & \mathbf{0} \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} \lambda \\ \nu \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} \mathbf{C^Y_0} \\ \mathbf{f_0} \end{bmatrix}$$

Linear system of N+L+1 equations with N+L+1 unknowns

Unique solution if C^Y is positive definite, sample points are distinct
F is of full rank (i.e., all columns are linearly independent)

Estimation variance: $\sigma^2_{\epsilon_0} = \sigma^2_Y - \mathbf{X}^t\mathbf{B}$

Example of basis functions in 1D

m(x) unknown but constant

$$\begin{cases} f_x^0 & = & 1 & \forall x \in D \\ f_x^1 & = & 0 & \forall x \in D \text{ and } \forall l \in [1..L] \end{cases}$$

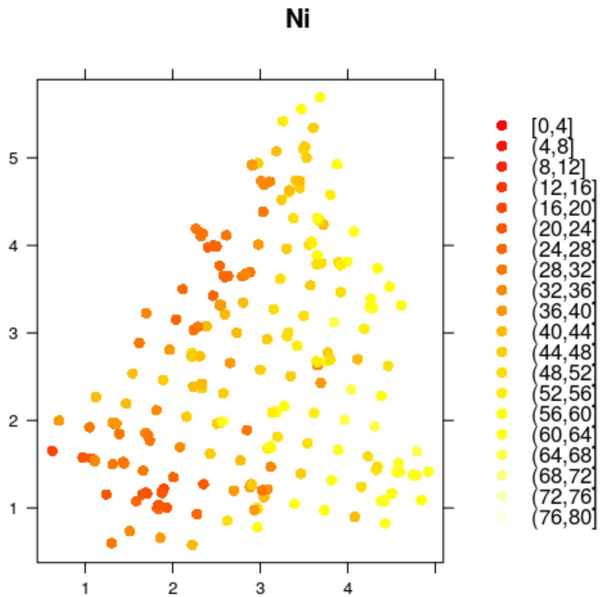
→ Ordinary kriging $m(x) = a_0$

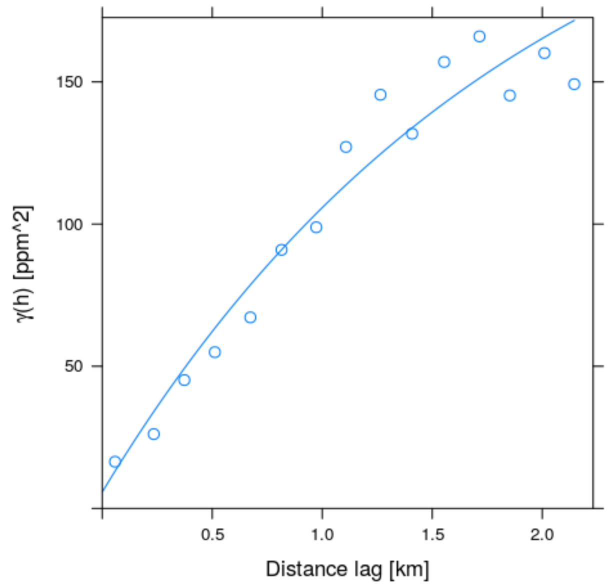
m(x) = linear trend

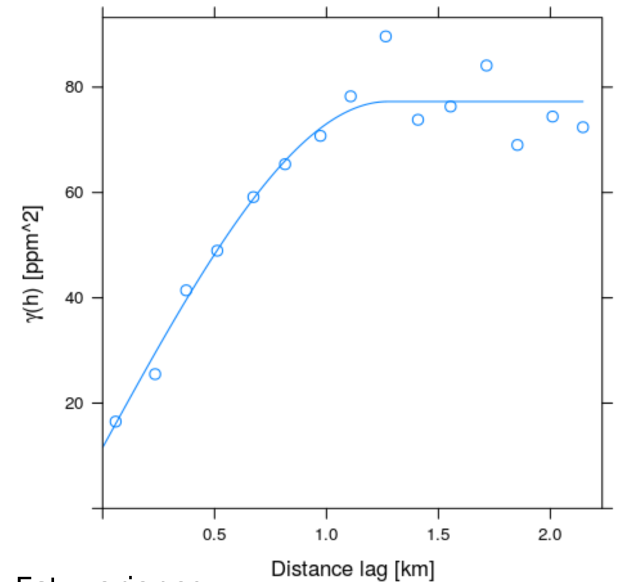
$$\begin{cases} f_x^0 & = & 1 & \forall x \in D \\ f_x^1 & = & x & \forall x \in D \\ f_x^l & = & 0 & \forall x \in D \text{ and } \forall l \in [2..L] \end{cases}$$

$$m(x) = a_0 + a_1 x$$

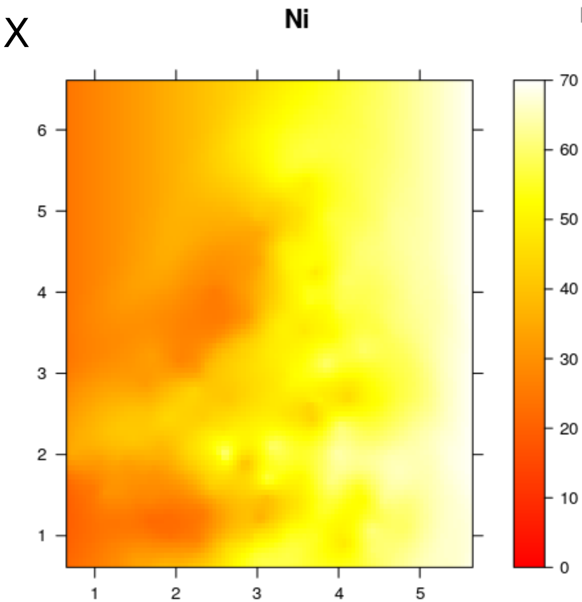
Example universal kriging



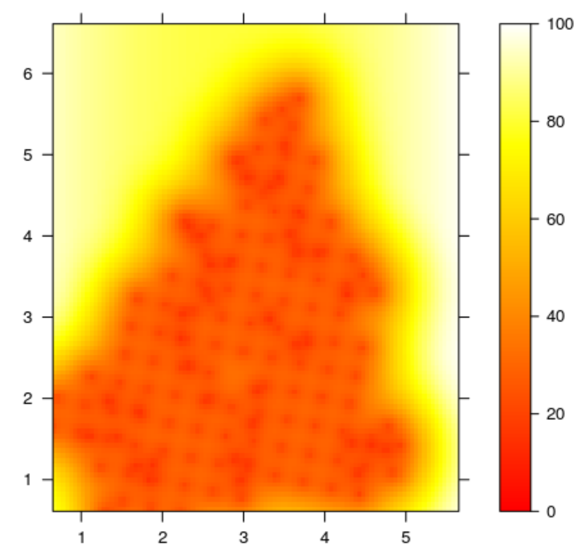




Ni + drift = 8X



Est. variance



Questions

- Why is kriging a linear interpolator?

- What are the differences between simple, ordinary and universal kriging?

- How is evolving the estimation variance when the distance to measurements increases?

So far, we have assumed that we perfectly know the covariance function.

But variogram estimate is less affected by sampling effects than covariance

→ Can we reformulate the kriging equations in terms of variogram?

Simple kriging → no constraints on weights from unbiased conditions
→ but we must have allowable linear comb. for var of increments
→ no formulation of simple kriging in variogram.

O + U kriging → we consider SRF (OK) or SRF residual (UK)
→ direct link between covariance and variogram
→ substitute $-\gamma$ in previous equations for OK (p.13) and UK (p.20).

Matrix notation:

$$
\begin{bmatrix}
\gamma^Y_{11} & \gamma^Y_{12} & \dots & \gamma^Y_{1N} & 1 & f^1_1 & \dots & f^L_1 \\
\gamma^Y_{21} & \gamma^Y_{22} & \dots & \gamma^Y_{2N} & 1 & f^1_2 & \dots & f^L_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\gamma^Y_{N1} & \gamma^Y_{N2} & \dots & \gamma^Y_{NN} & 1 & f^1_N & \dots & f^L_N \\
1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\
f^1_1 & f^1_2 & \dots & f^1_N & 0 & 0 & \dots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
f^L_1 & f^L_2 & \dots & f^L_N & 0 & 0 & \dots & 0
\end{bmatrix}
\begin{bmatrix}
\lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ -\nu_0 \\ -\nu_1 \\ \vdots \\ -\nu_L
\end{bmatrix}
=
\begin{bmatrix}
\gamma^Y_{10} \\ \gamma^Y_{20} \\ \vdots \\ \gamma^Y_{N0} \\ 1 \\ f^1_0 \\ \vdots \\ f^L_0
\end{bmatrix}
$$

$$\gamma^Y_{\alpha\alpha} = 0$$

Kriging with an external drift (KED)

KED is a variant of UK using external auxiliary variables S_l known over the domain to estimate the drift, instead of the f_l functions of the coordinates.

Basis function f^l_α replaced by S^l_α in matrix system on p.21/27

Data → external drift → residuals → estimation of C^Y or γ^Y

$$\begin{bmatrix} \gamma^Y_{11} & \gamma^Y_{12} & \dots & \gamma^Y_{1N} & 1 & S^1_1 & \dots & S^L_1 \\ \gamma^Y_{21} & \gamma^Y_{22} & \dots & \gamma^Y_{2N} & 1 & S^1_2 & \dots & S^L_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma^Y_{N1} & \gamma^Y_{N2} & \dots & \gamma^Y_{NN} & 1 & S^1_N & \dots & S^L_N \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ S^1_1 & S^1_2 & \dots & S^1_N & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ S^L_1 & S^L_2 & \dots & S^L_N & 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ -\nu_0 \\ -\nu_1 \\ \vdots \\ -\nu_L \end{bmatrix} = \begin{bmatrix} \gamma^Y_{10} \\ \gamma^Y_{20} \\ \vdots \\ \gamma^Y_{N0} \\ 1 \\ S^1_0 \\ \vdots \\ S^L_0 \end{bmatrix}$$

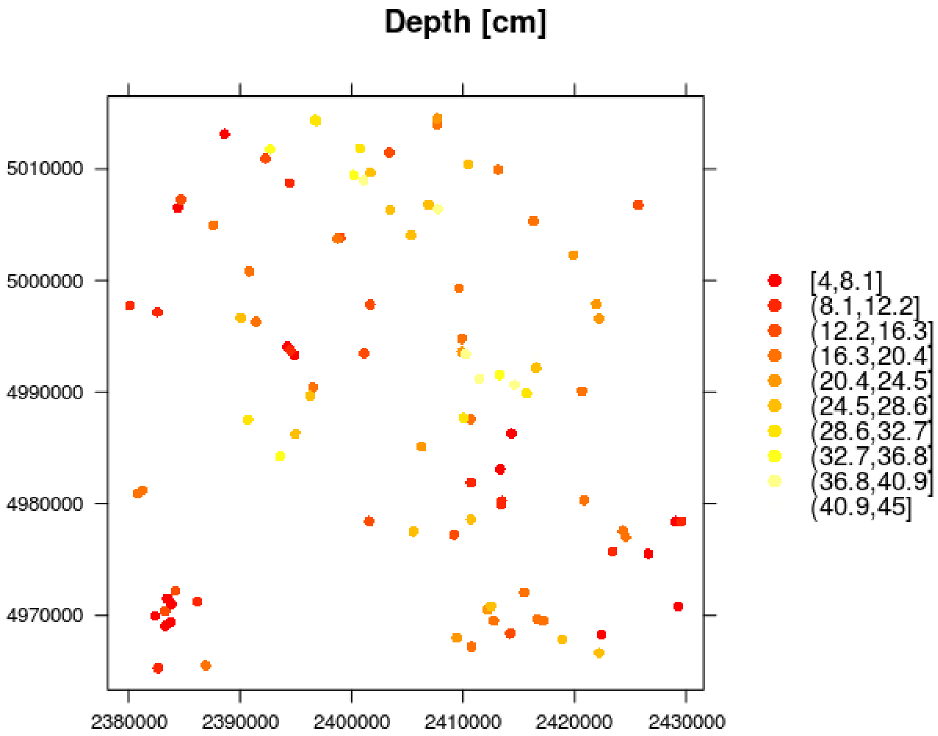
S^l_α must be linearly independent for matrix inversion!

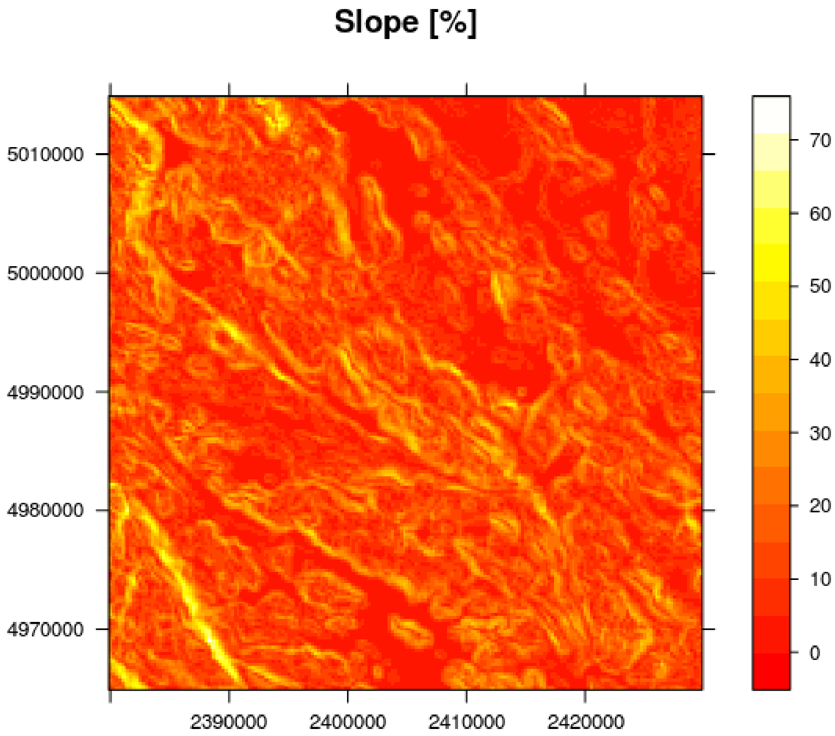
Example: drift in soil depth with slope (from Hengl, Geoderma, 2004)

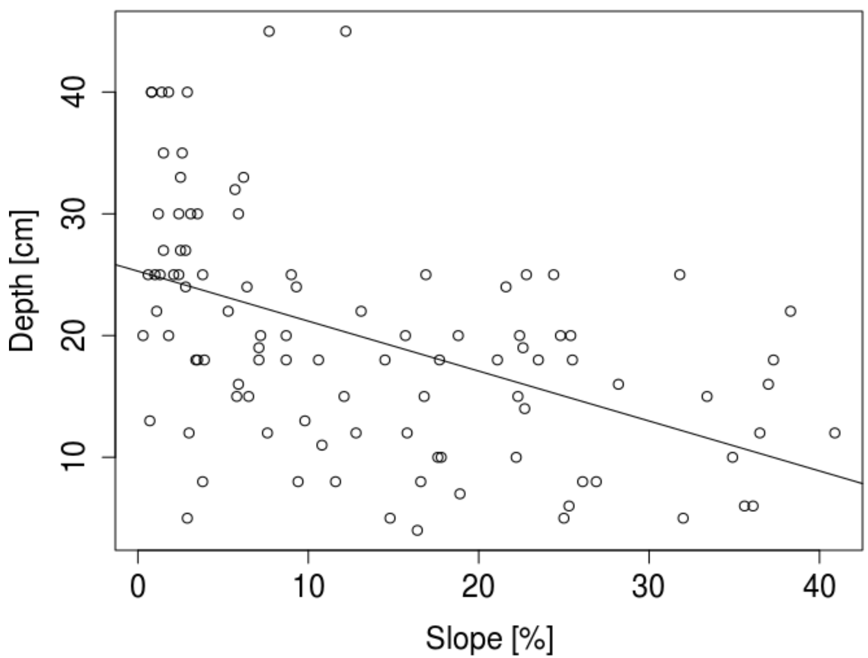
Domain = 50x50 km^2, central Croatia.

100 observations of soil depth (cm).

Auxiliary variable: slope (%) from DEM (100x100 m^2).



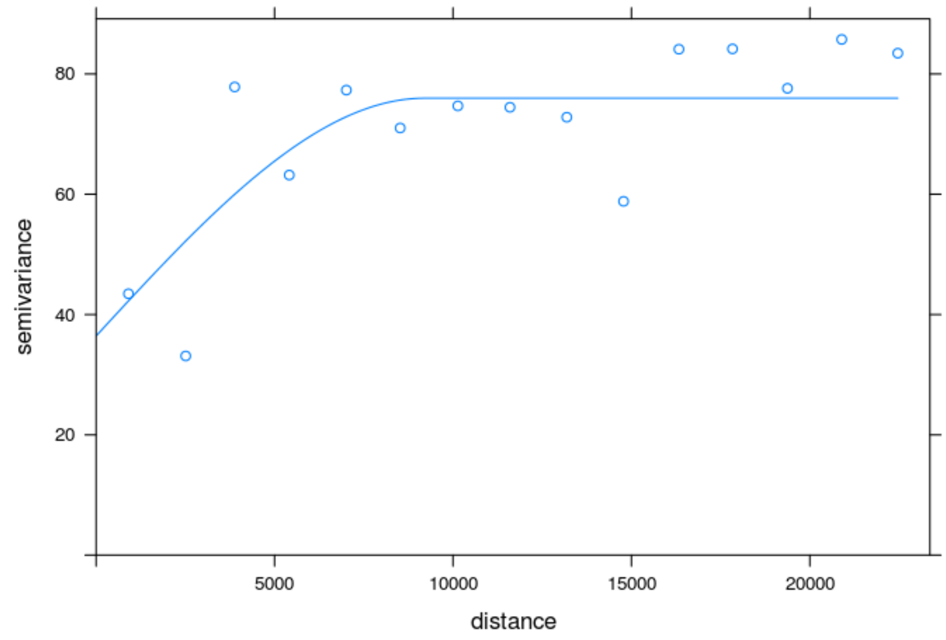




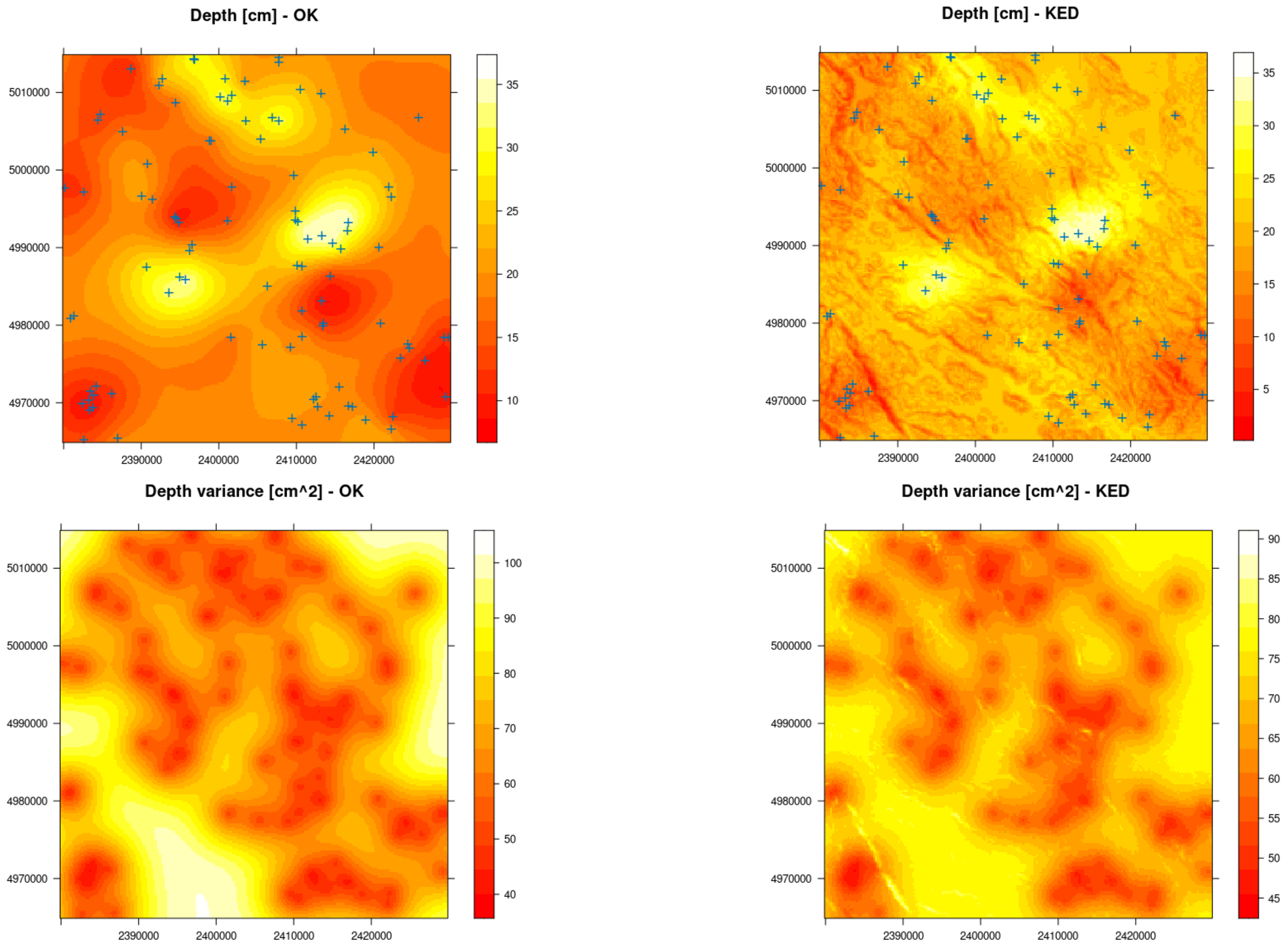
External drift provided by s as

$$m(x) = a_0 + a_1 s(x)$$

Variogram after drift removed



Maps of interpolated values and estimation variance



Warning!

In previous equations, we assume the true cov/vario of Z/Y to be known.

But - covariance/variogram estimated from sample;
- difficulties due to the drift (see Chap.2).

Possible ways to infer underlying covariance/variogram:

- In case of unidirectional drift: estimate the cov/vario in direction $\perp$ to the drift.

- In case of mild drift: bias due to the drift negligible at short distance lags.
Choice between OK or UK in this case can be sorted by cross-validation.

Questions

- Why is there no formulation of simple kriging in variogram?

- How can the cov/variogram needed for universal kriging be estimated?

- What is the main difference between universal kriging and kriging with an external drift?

We will examine the following properties/characteristics of kriging:

- Linear, unbiased and optimal estimator.
- Exact interpolator.
- Quantify uncertainty in estimates (estimation variance).
- Smoothing effect.
- Kriging weights: screen and relay effects, negative weights.
- Kriging neighbourhood.

By definition, the estimate Z^* is:

- a linear combination of measured values Z_α. For SK, OK and UK:

$$Z_0^* = \sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha$$

- unbiased, as imposed by condition 1: $\mathrm{E}[Z^*] = \mathrm{E}[Z]$

- optimal, in the least square sense as imposed by condition 2:

$\mathrm{Var}[Z^* - Z]$ is minimum

"Deterministic interpolator": λ fixed for given Z_α and variogram/covariance model

Simple kriging equation
$$\sum_{\beta=1}^{N} \lambda_\beta C_{\alpha\beta} = C_{\alpha 0} \;\; \forall \alpha \in [1..N]$$

If $x_0 = x_{\alpha_0}$, $\alpha_0 \in [1..N]$
$$\Rightarrow \begin{cases} \lambda_{\alpha_0} & = & 1 \\ \lambda_\alpha & = & 0 \quad \forall \alpha \neq \alpha_0 \end{cases}$$

because α are distinct

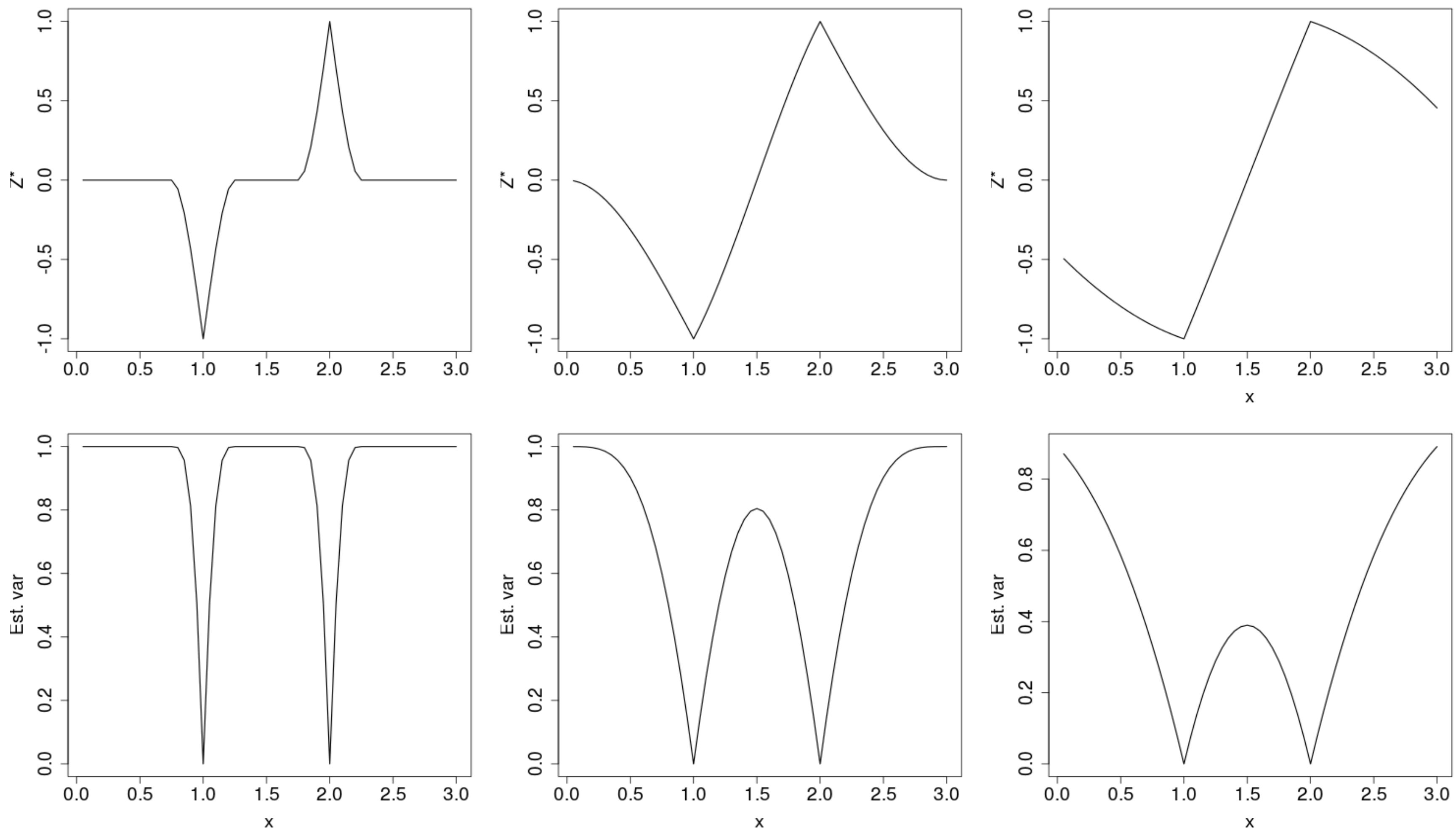
$$\Rightarrow Z^*_{\alpha_0} = \sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha = Z_{\alpha_0}$$

Estimation variance
$$\sigma^2_{\epsilon_0} = \sigma^2_Z - \sum_{\alpha=1}^{N} \lambda_\alpha C_{\alpha 0}$$

$$\sigma^2_{\epsilon_{\alpha_0}} = \sigma^2_Z - \lambda_{\alpha_0} C_{\alpha_0 \alpha_0} = 0$$

$Z_1 = -1$; $Z_2 = 1$; spherical variogram (sill=1,nugget = 0)

Simple kriging (m is known and assumed =0 in this example)



range = 0.25 range = 1 range = 2

Universal kriging equation
$$\sum_{\beta=1}^{N} \lambda_\beta \gamma^Y_{\alpha\beta} = \gamma^Y_{\alpha 0} + \sum_{l=0}^{L} \nu_l f^l_\alpha \quad \forall \alpha \in [1..N]$$
$$\sum_{\alpha=1}^{N} \lambda_\alpha f^l_\alpha = f^l_0 \quad \forall l \in [0..L]$$

If $x_0 = x_{\alpha_0}\ ,\ \alpha_0 \in [1..N] \quad \Rightarrow \begin{cases} \lambda_{\alpha_0} & = \quad 1 \\ \lambda_\alpha & = \quad 0\ \forall \alpha \neq \alpha_0 \\ \sum_{l=0}^{L} \nu_l f^l_{\alpha_0} & = \quad 0 \end{cases}$

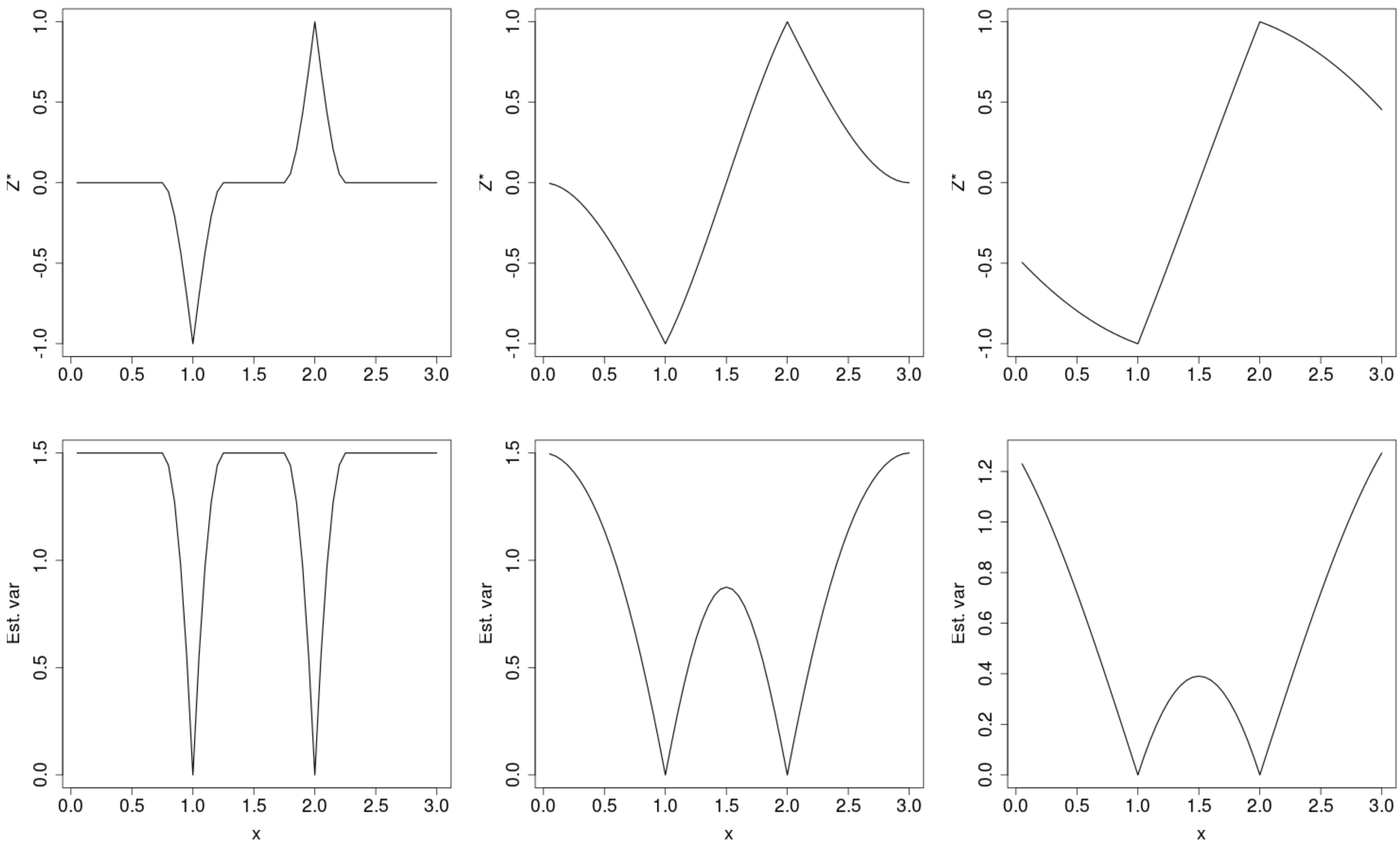
because α are distinct

$$\Rightarrow Z^*_{\alpha_0} = \sum_{\alpha=1}^{N} \lambda_\alpha Z_\alpha = Z_{\alpha_0}$$

Estimation variance $\sigma^2_{\epsilon_0} = \sum_{\alpha=1}^{N} \lambda_\alpha \gamma^Y_{\alpha 0} - \sum_{l=0}^{L} \nu_l f^l_0 \quad \rightarrow \sigma^2_{\epsilon_{\alpha_0}} = \gamma^Y_{\alpha_0 \alpha_0} = 0$

$Z_1 = -1$; $Z_2 = 1$; spherical variogram (sill=1,nugget = 0)

Ordinary kriging



range = 0.25 range = 1 range = 2

Interpolated values exhibit smoother behavior than real values

Variance of interpolated value Z*:

SK $$\sigma^2_{Z_0^*} = \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\lambda_\alpha\lambda_\beta C_{\alpha\beta} = \sum_{\alpha=1}^{N}\lambda_\alpha C_{\alpha 0} = \sigma^2_Z - \sigma^2_{\epsilon_0}$$

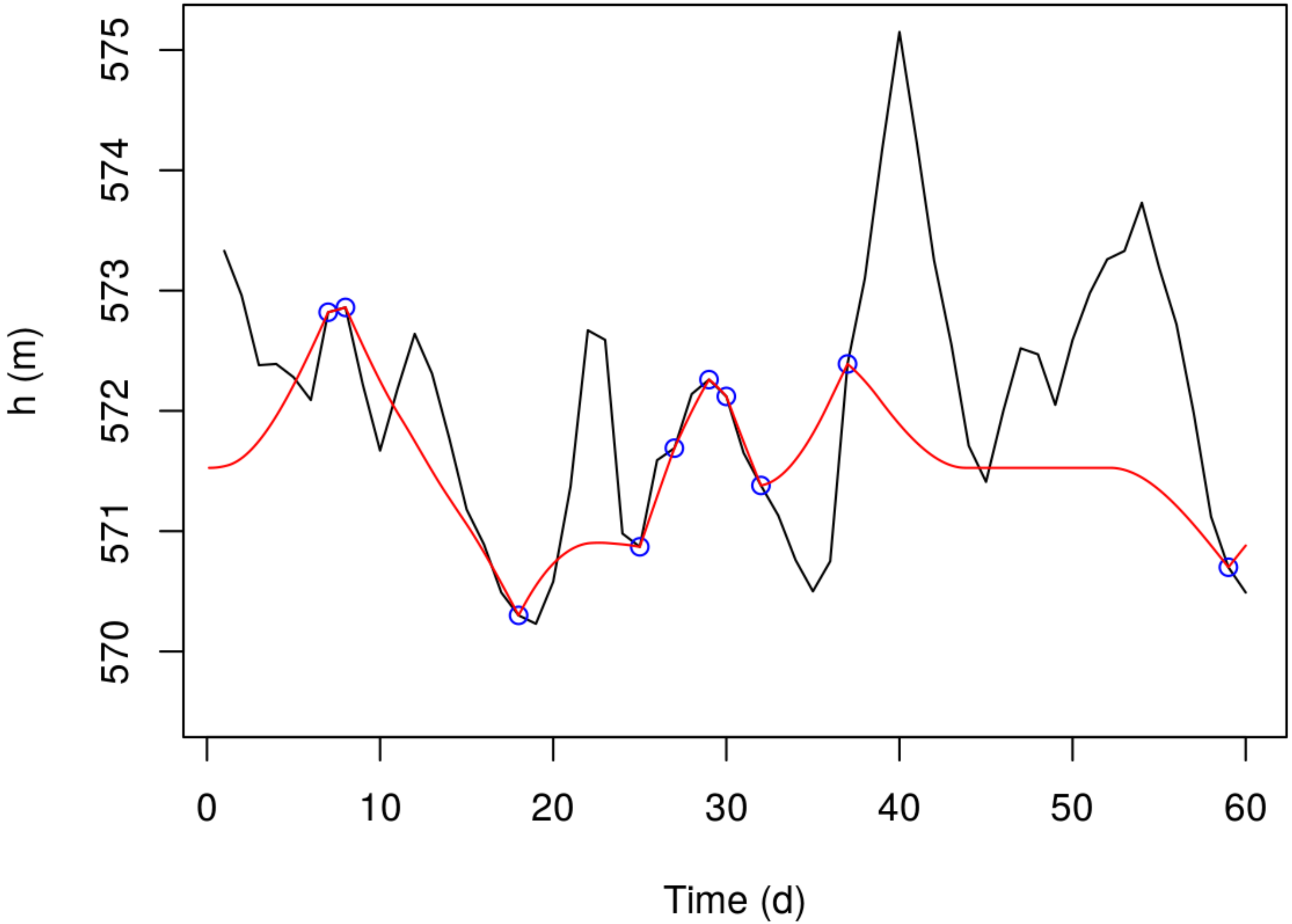
$$\Rightarrow \sigma^2_{Z_0^*} \leq \sigma^2_Z$$

UK $$\sigma^2_{Y_0^*} = \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\lambda_\alpha\lambda_\beta C^Y_{\alpha\beta} \;\; ; \begin{cases} f_0^l = \sum_{\alpha=1}^{N}\lambda_\alpha f_\alpha^l,\ l = 0..L \text{ (p.17)} \\ C^Y_{0\alpha} = \sum_{\beta=1}^{N}\lambda_\beta C^Y_{\alpha\beta} + \sum_{l=0}^{L}\nu_l f_\alpha^l,\ \alpha = 1..N \text{ (p.19)} \end{cases}$$

$$\sigma^2_{\epsilon_0} = \sigma^2_Y - \sum_{\alpha=1}^{N}\lambda_\alpha C^Y_{\alpha 0} - \sum_{l=0}^{L}\nu_l f_0^l$$

$$\Rightarrow \sigma^2_{Y_0^*} = \sigma^2_Y - \sigma^2_{\epsilon_0}$$

Example with piezometric height (at Biolay-Orjulaz)



Kriging provides estimation of uncertainty associated with interpolated values

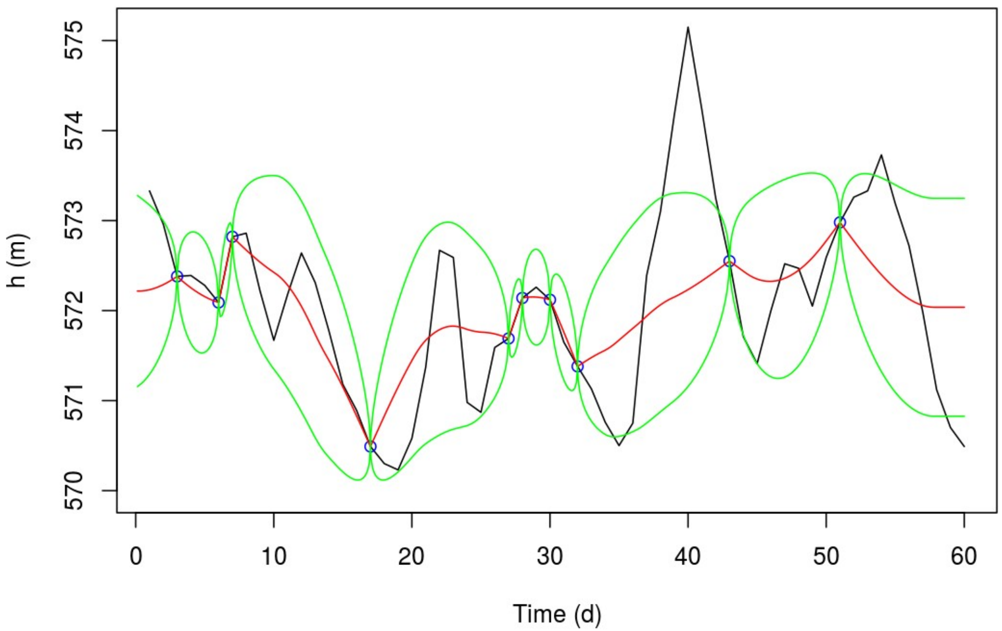
Estimation variance:

SK $$\sigma^2_{\epsilon_0} = \sigma^2_Z - \sum_{\alpha=1}^{N} \lambda_\alpha C_{\alpha 0}$$

UK $$\sigma^2_{\epsilon_0} = \sigma^2_Y - \sum_{\alpha=1}^{N} \lambda_\alpha C^Y_{\alpha 0} - \sum_{l=0}^{L} \nu_l f^l_0$$

If Z Gaussian
→ 80% confidence interval = $\pm 1.28\sigma_{\varepsilon o}$
→ 95% confidence interval = $\pm 2\sigma_{\varepsilon o}$

If Z continuous unimodal
→ 95% confidence interval ≤ $\pm 3\sigma_{\varepsilon o}$
(Vysochanskij-Petunin, 1980)



Screening effect

Data points close to estimated point can “screen” data points beyond (in particular when interdistances << range)

$$\sigma^2_{OK} = 1.14\sigma^2_Z$$

$$\sigma^2_{OK} = 0.87\sigma^2_Z$$

From Wackernagel, 1998, p.93

Interpolation methods based on distance cannot take into account such effects due to data point correlation.

Relay effect

Data points beyond the range may influence estimated point via their correlation with points within range

Interpolation methods based on distance cannot take into account such effects due to data point correlation.

From Chilès and Delfiner, 2012, p.208

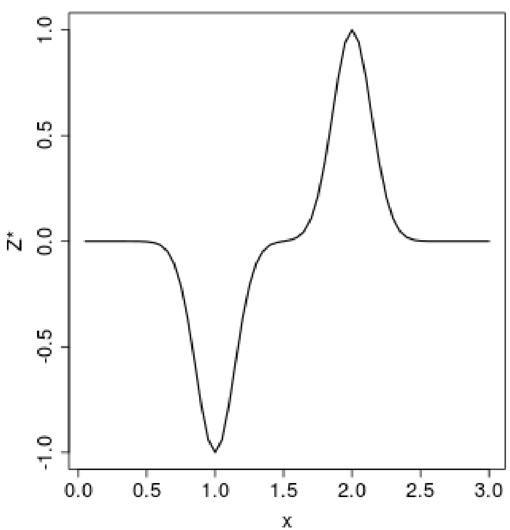
Kriging weights can be negative!

Interesting feature because it allows extrapolation beyond the range of measured values.

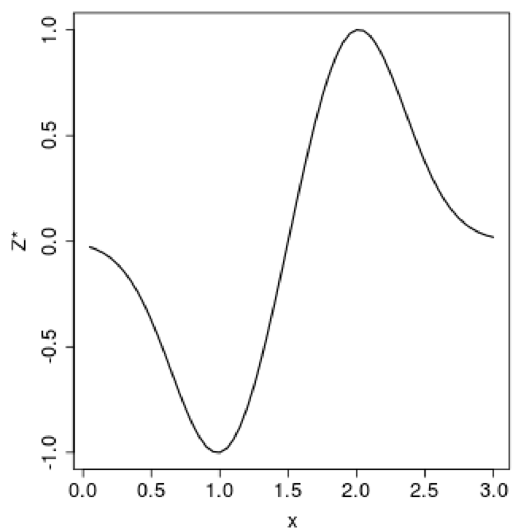
Can be a problem when dealing with variables that must be positive!
(ex: concentration, rainfall amount, hydraulic conductivity, etc...)

$Z_1 = -1$; $Z_2 = 1$; Gaussian variogram (sill=1,nugget = 0)

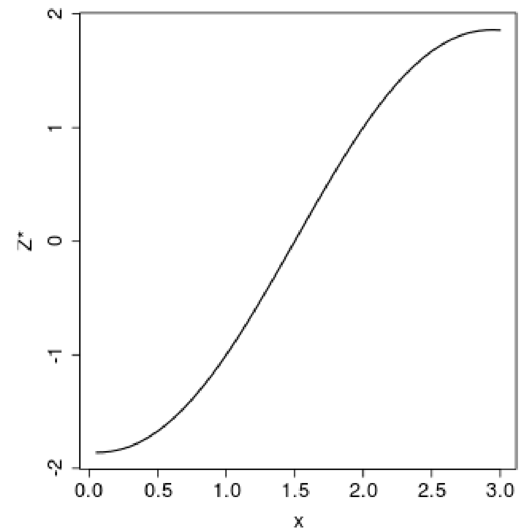
Ordinary kriging



range = 0.2



range = 1



range = 2

- Why is kriging an exact interpolator?

- What is the smoothing effect?

- What do negative kriging weights mean in terms of min/max interpolated values?

Formulation so far has involved all data points measured over the studied domain. This can be a large number, and matrix computations can become too cumbersome.

Because far points are less correlated, it is possible to select a neighborhood around the point to estimate, and perform calculations only on this subset.

However, the selection of this neighborhood is usually not straightforward in practice. Even if the variogram exhibits a clear range, using a circle with the range as radius may not be correct, because of screening and relay effects.

A “rule of thumb” is to check if the results are significantly changing when additional points are considered (larger neighborhood).

Measured values can be affected by uncertainty (e.g., due to sensor)
Y = measured value ; Z = true value ; e = error

$$Y(x) = Z(x) + e(x)$$

Z with unknown constant mean value (ordinary kriging)

Simplifying assumptions: white noise error

$$\rightarrow \mathrm{E}[e] = 0 \quad \sigma^2_{e_\alpha} = \sigma^2_e$$

$$C^{eZ}_{\alpha\beta} = 0 \ \forall(\alpha, \beta) \in [1..N]^2$$

$$C^{ee}_{\alpha\beta} = 0 \ \forall(\alpha \neq \beta) \in [1..N]^2$$

$$\mathrm{E}[Y] = \mathrm{E}[Z + e] = \mathrm{E}[Z]$$

$$\sigma^2_Y = \sigma^2_Z + \sigma^2_e \Leftrightarrow C^{YY}_{\alpha\alpha} = C^{ZZ}_{\alpha\alpha} + \sigma^2_e \quad \forall\alpha \in [1..N]$$

$$\begin{aligned} C^{YY}_{\alpha\beta} &= \mathrm{E}[(Y_\alpha - \mathrm{E}[Y])(Y_\beta - \mathrm{E}[Y])] \\ &= \mathrm{E}[(Z_\alpha + e_\alpha - \mathrm{E}[Z])(Z_\beta + e_\beta - \mathrm{E}[Z])] \\ &= C^{ZZ}_{\alpha\beta} + \delta_{\alpha\beta}\sigma^2_e \quad \begin{cases} \delta_{\alpha\alpha} = 1 \\ \delta_{\alpha\beta} = 0 \end{cases} \end{aligned}$$

Estimated value Z* at x_0: $$Z_0^* = \sum_{\alpha=1}^{N} \lambda_\alpha Y_\alpha$$

Condition 1: unbiased estimation

$$\mathrm{E}[Z_0^*] = \mathrm{E}\left[\sum_{\alpha=1}^{N} \lambda_\alpha Y_\alpha\right] = \sum_{\alpha=1}^{N} \lambda_\alpha \mathrm{E}[Z] \qquad \Rightarrow \sum_{\alpha=1}^{N} \lambda_\alpha = 1$$

Condition 2: optimal estimation

$$\sigma_{\epsilon_0}^2 = \mathrm{E}\left[(Z_0^* - Z_0)^2\right] = \mathrm{E}\left[\left(\sum_{\alpha=1}^{N} \lambda_\alpha Y_\alpha - Z_0\right)^2\right]$$

$$= \sigma_Z^2 - 2\sum_{\alpha=1}^{N} \lambda_\alpha C_{\alpha 0}^{YZ} + \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N} \lambda_\alpha \lambda_\beta C_{\alpha\beta}^{YY}$$

$$C_{\alpha\beta}^{YZ} = \mathrm{E}[(Y_\alpha - E[Y])(Z_\beta - E[Z])] = \mathrm{E}[(Z_\alpha + e_\alpha - E[Z])(Z_\beta - E[Z])]$$

$$= C_{\alpha\beta}^{ZZ} + C_{\beta\alpha}^{Ze} = C_{\alpha\beta}^{ZZ} \quad \forall(\alpha,\beta) \in [1..N]^2$$

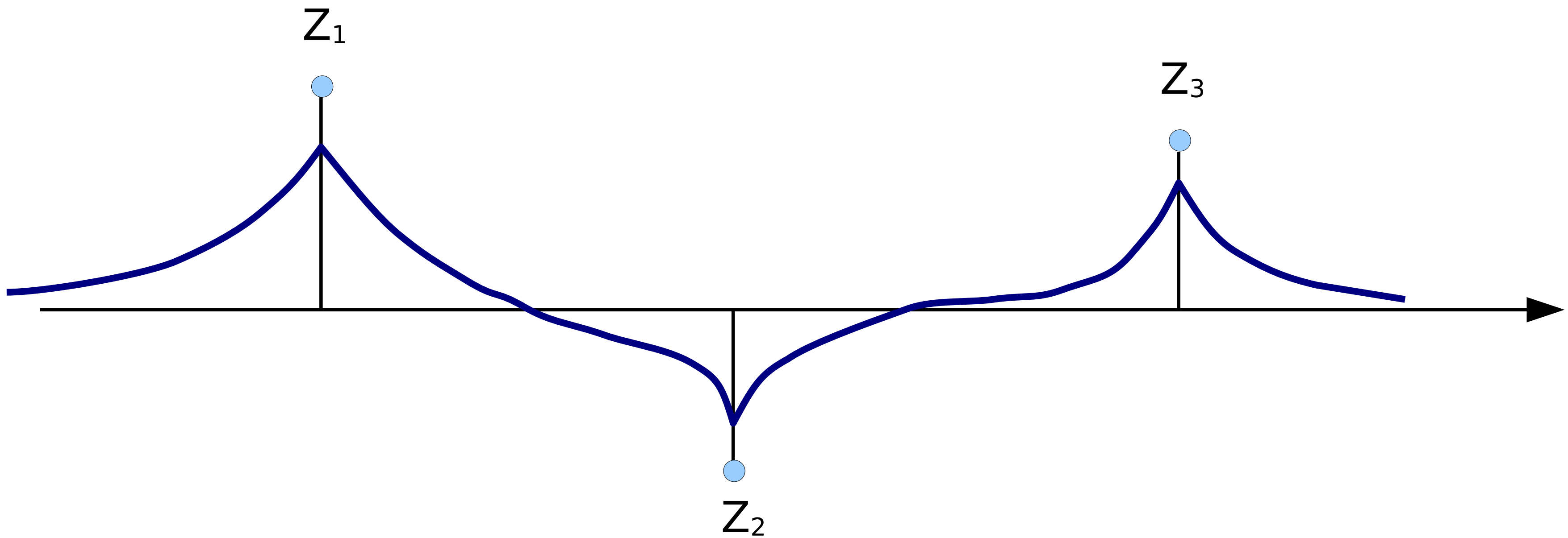
$$\sigma^2_{\epsilon_0} = \sigma^2_Z - 2\sum_{\alpha=1}^{N} \lambda_\alpha C^{ZZ}_{\alpha 0} + \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N} \lambda_\alpha \lambda_\beta (C^{ZZ}_{\alpha\beta} + \delta_{\alpha\beta}\sigma^2_e)$$

→ optimization with constraint (Lagrange multiplier)

$$\sum_{\beta=1}^{N} \lambda_\beta (C^{ZZ}_{\alpha\beta} + \delta_{\alpha\beta}\sigma^2_e) = C^{ZZ}_{\alpha 0} \quad \text{and} \quad \sum_{\alpha=1}^{N} \lambda_\alpha = 1$$

$$\begin{bmatrix} C^{ZZ}_{11} + \sigma^2_e & C^{ZZ}_{12} & \cdots & C^{ZZ}_{1N} & 1 \\ C^{ZZ}_{21} & C^{ZZ}_{11} + \sigma^2_e & \cdots & C^{ZZ}_{2N} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C^{ZZ}_{N1} & C^{ZZ}_{N2} & \cdots & C^{ZZ}_{NN} + \sigma^2_e & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \nu \end{bmatrix} = \begin{bmatrix} C^{ZZ}_{10} \\ C^{ZZ}_{20} \\ \vdots \\ C^{ZZ}_{N0} \\ 1 \end{bmatrix}$$

Same system as OK, except for term σ_e^2



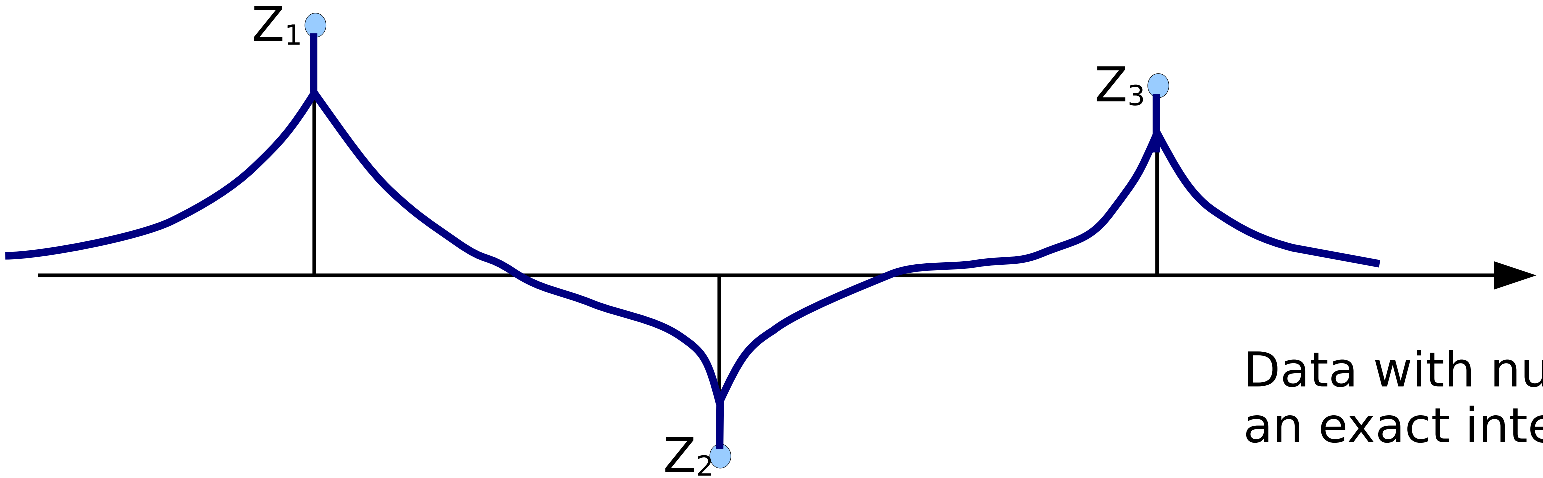
Uncertain data: kriging no more exact interpolator!

Nugget variance = σ_n^2 $$C_{\alpha\beta} = C(|x_\alpha - x_\beta|) + \delta_{\alpha\beta}\sigma_n^2$$

$$\begin{bmatrix} C_{11} + \sigma_n^2 & C_{12} & \cdots & C_{1N} & 1 \\ C_{21} & C_{11} + \sigma_n^2 & \cdots & C_{2N} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{N1} & C_{N2} & \cdots & C_{NN} + \sigma_n^2 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \nu \end{bmatrix} = \begin{bmatrix} C_{10} + \delta_{10}\sigma_n^2 \\ C_{20} + \delta_{20}\sigma_n^2 \\ \vdots \\ C_{N0} + \delta_{N0}\sigma_n^2 \\ 1 \end{bmatrix}$$

Same system as uncertain data, except additional nugget variance in right-hand term

→ exact interpolator



Data with nugget: kriging is an exact interpolator!

1. Kriging

- Interpolated value = linear combination of measurements: $$Z^*(x) = \sum_{\alpha=1}^{N} \lambda_\alpha Z(x_\alpha)$$
- Weights λ are estimated by fulfilling 2 conditions:
  (i) unbiased estimation and and (ii) minimal estimation variance

2. Simple kriging

- Mean of Z is known (Z is SRF2, no formulation in variogram)
- No conditions on weights

3. Ordinary kriging

- Mean of Z is constant but unknown (Z SRF or IRF)
- Formulation in covariance and variogram
- Conditions on weight λ: $$\sum_{\alpha=1}^{N} \lambda_\alpha = 1$$

4. Universal kriging

- Z is not SRF or IRF but Z(x) = m(x) + Y(x), m=determ. trend and Y = SRF residual
- Formulation in covariance and variogram
- OK = UK with constant trend m.
- Difficulty: estimating the covariance or variogram of Y...