AESB2440: Geostatistics & Remote Sensing

Lecture 7: Stochastic Interpolation

Wednesday, May 7, 2015, Roderik Lindenbergh

Dept. of Geoscience & Remote Sensing

Delft University of Technology

Overview

A. Motivating example: A different mean

B. Deriving the Simple Kriging equations

C. Obtaining a covariance function from observations

D. Dissimilarities instead of similarity: the variogram

After conclusions: some extra slides.

Dept. of Geoscience & Remote Sensing

References

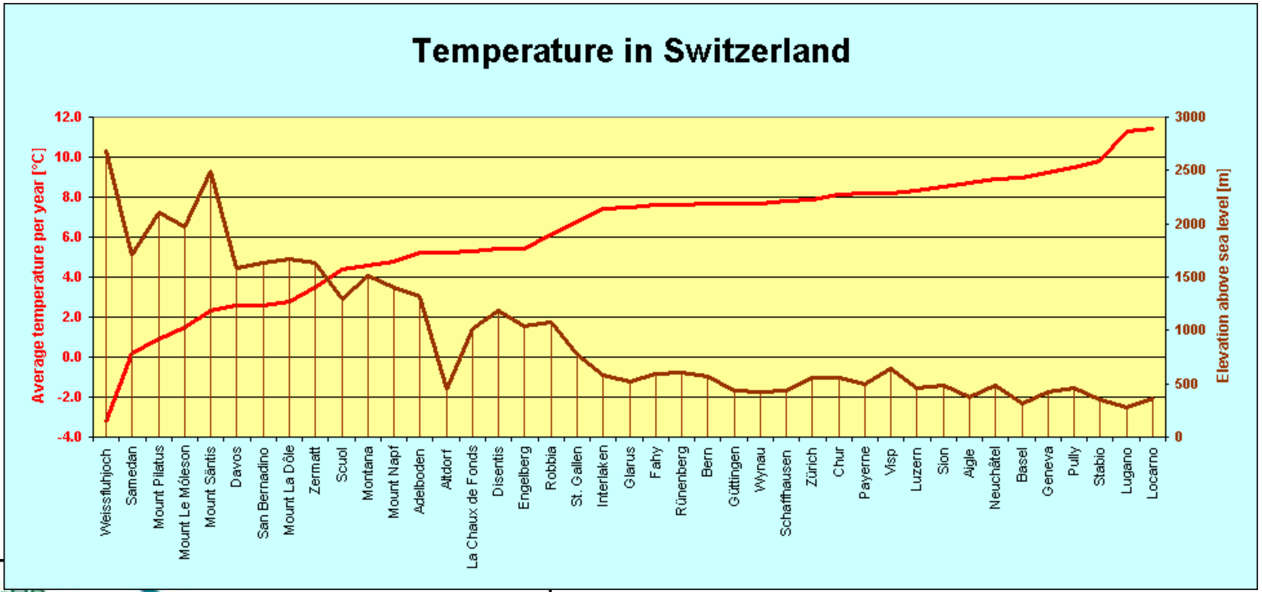
H. Wackernagel
Multivariate Geostatistics
Third Edition, Springer, 2003
Chapter 3, Linear Regression and Simple Kriging
Chapter 4, Kriging the mean
Chapters 6,7, Variograms
Chapter 11, 12 & 13 Ordinary Kriging (properties)

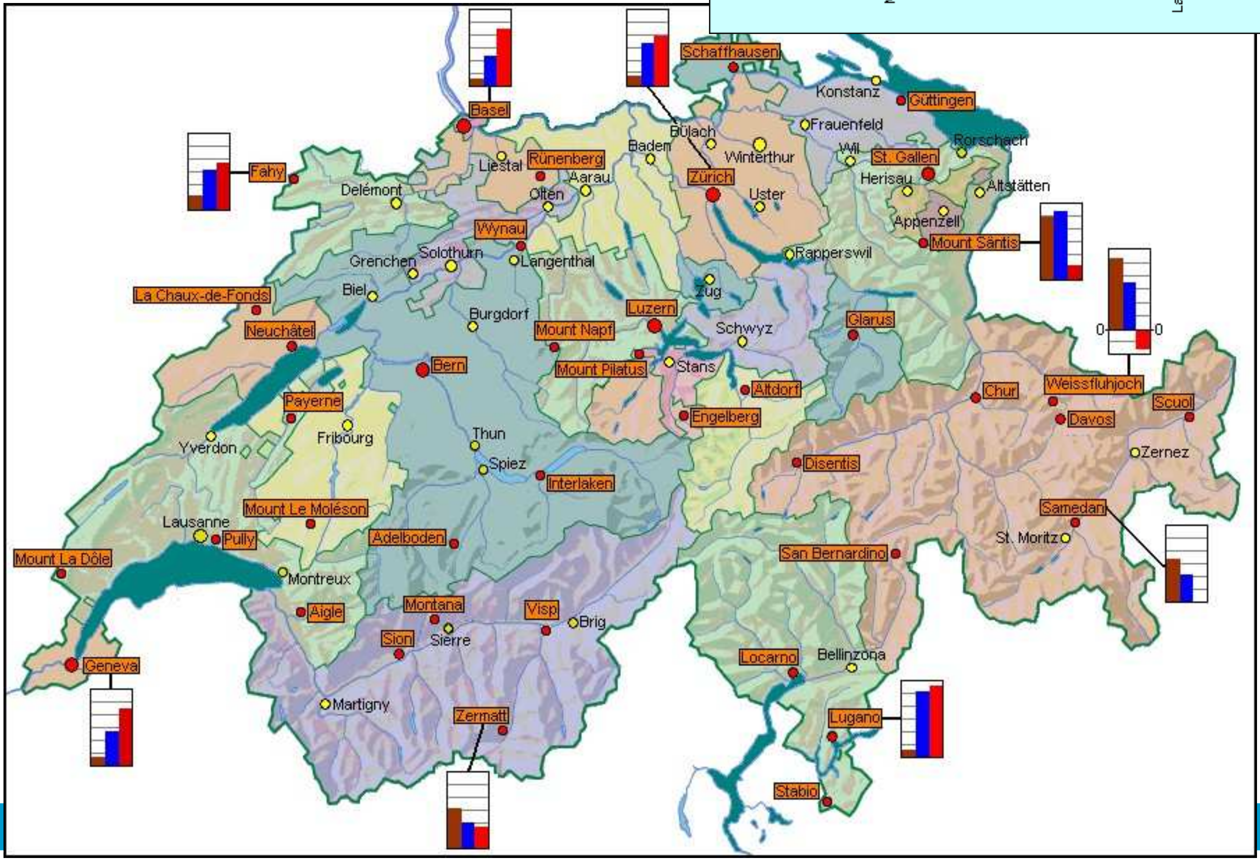
Matlab Recipes for Earth Sciences
Chapter 7, Spatial data,
Notably Section 7.9, Geostatistics

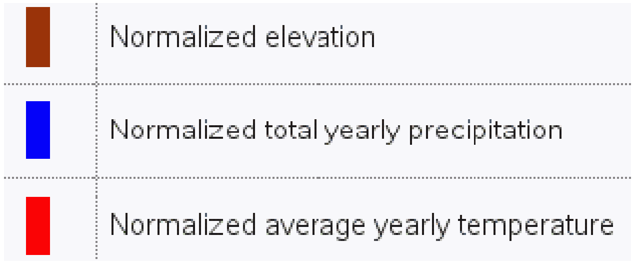
D.D. Sarma
Geostatistics with Applications in Earth Sciences
Chapter: Kriging Variance and Kriging Procedure

Busy trains

Most people experience trains to be busier than they really are, because most people travel when its busy (that's why it's busy)







http://www.about.ch/geography/climate/index.html

Bias and correlation problems

If you ask 1000 people how busy the trains are, the answer is biased towards busy times.

If you only measure temperature

- in the valleys of Switzerland (where it is warmer)
- and not on the peaks (where it is colder)

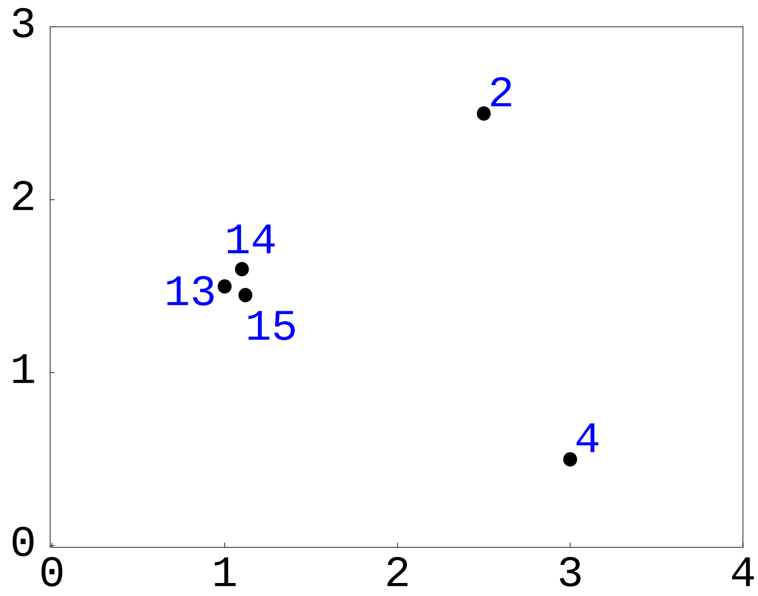
you get biased observations as temperature is correlated to height.

Possible solution:
Assess and incorporate correlation into your method:

- by decreasing the influence of correlated observations
- by taking additional attributes (like height) into account

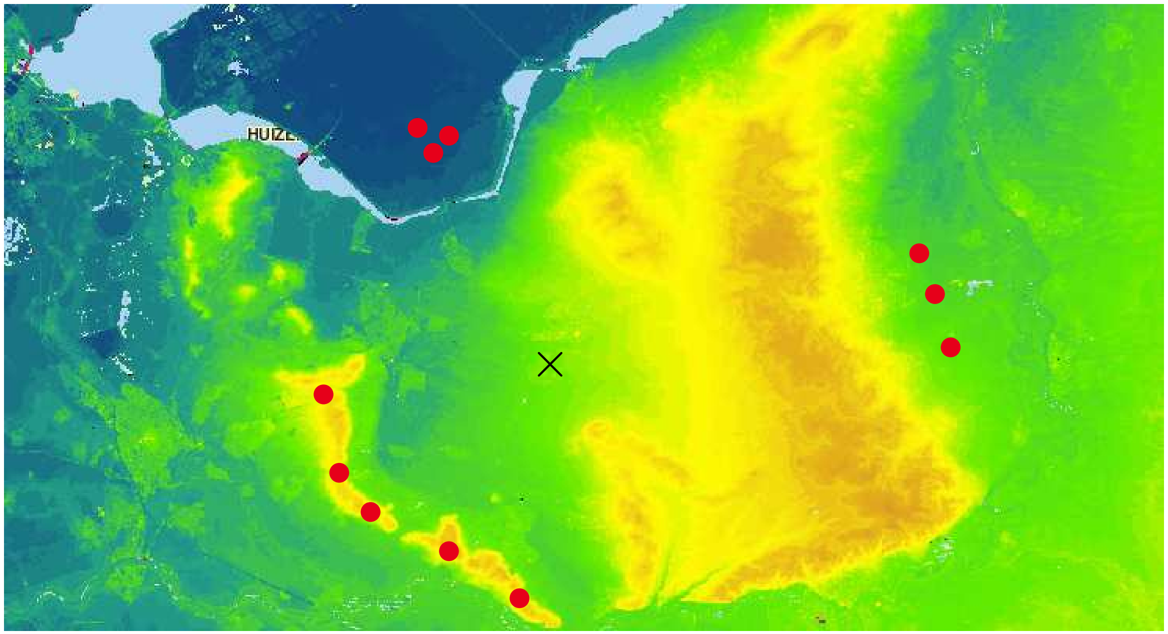
A. A different mean

Traditional Mean

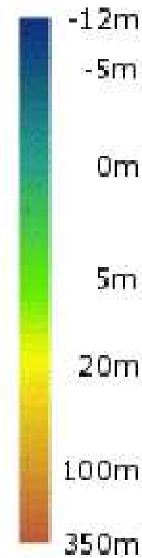


$$\bar{z}_T = \frac{1}{5}\cdot 2 + \frac{1}{5}\cdot 4 + \frac{1}{5}\cdot 13 + \frac{1}{5}\cdot 14 + \frac{1}{5}\cdot 15 = 9.6$$

Correlation while interpolating



Height in m. above NAP



(Source: www.ahn.nl/kaart)

Some red observation points are spatially correlated.

There are three correlated clusters: Utrechtse Heuvelrug, Flevoland and Ijsselvallei

This correlation should be taken into account when using such observations for interpolating at, say, location $\times$.

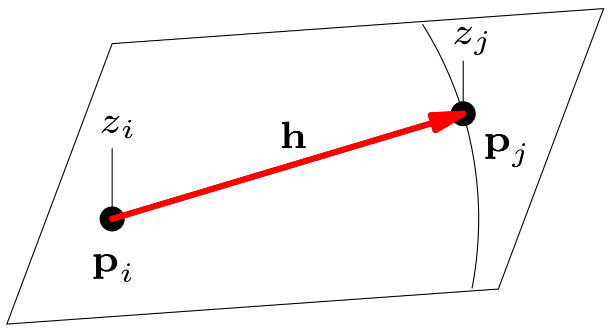
Mean, variance, covariance, correlation

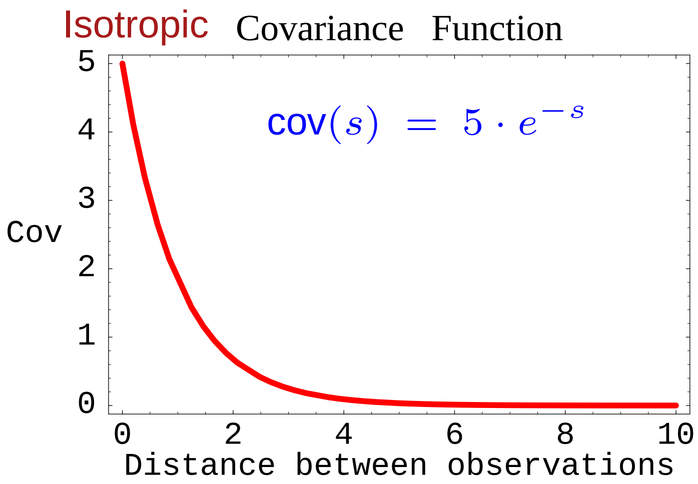
	theoretical	experimental
mean	$E\{z\} = \int_{z \in I\!R} z f(z) dz = \mu$	$\overline{z} = \frac{1}{n} \sum_{i=1}^{n} z_i$
variance	$\text{var}\{z\} = E\{(z - E\{z\})^2\}$	$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (z_i - \overline{z})^2$
covariance	$c(z_i, z_j) = E\{(z_i - \mu)(z_j - \mu)\}$	$c_{ij} = (z_i - \bar{z})(z_j - \bar{z})$
correlation		$\rho_{ij} = \frac{c_{ij}}{c_i \cdot c_j}$

Note that:

- $f(z)$: probability density function
- Variance σ^2 is always positive
- Covariance assumes that mean $\bar{z}$ exists
- Correlation $\rho_{ij} \in [-1, 1]$ is scaled covariance

Covariance function





Assumption. Stationarity of the first two moments

Assume the following holds for the height Z over some domain $\mathcal{D}$

1. The (expected) mean of the height is the same, all over $\mathcal{D}$.
2. The covariance between two locations $\mathbf{p_i}$ and $\mathbf{p_j}$ only depends on the difference vector $\mathbf{h} := \mathbf{p_j} - \mathbf{p_i}$, not on the locations itself

If these assumptions hold, the covariance function is defined and describes the covariance as function of the 2D difference

Isotropy: covariance function depends only on distance, not on direction.

Covariance function, theoretical.

Covariance function definition.

$$\text{cov}(\mathbf{h}) \quad = \quad E\{Z(\mathbf{p}) \cdot Z(\mathbf{p}+\mathbf{h})\} - \mu^2$$

Covariance function properties.

1. $\text{cov}(\mathbf{0}) = \text{var}(Z(\mathbf{x}))$.
2. $|\text{cov}(\mathbf{h})| \ \leq \ \text{cov}(\mathbf{0})$.
3. $\text{cov}(\mathbf{h}) \ = \ \text{cov}(-\mathbf{h})$.
4. Covariances can become negative. (How??)

The covariance function gives a measure of similarity

Dept. of Geoscience & Remote Sensing

Correlated mean

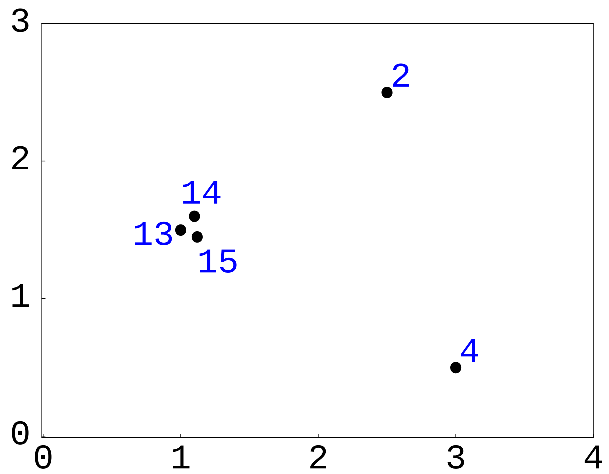
$z_1 = 13$

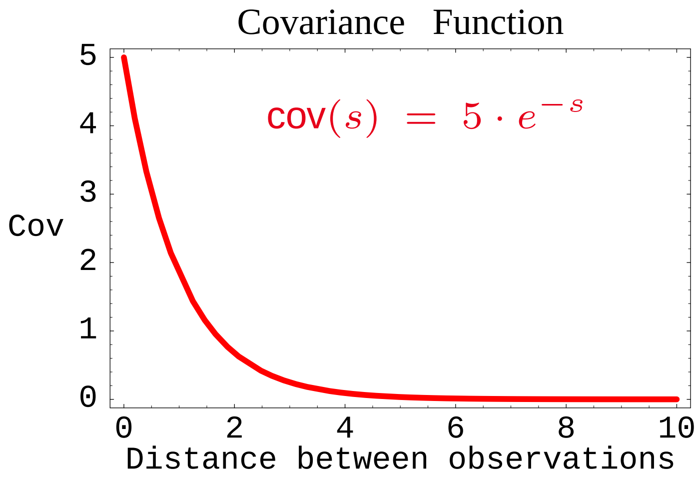
$z_2 = 14$

$z_3 = 15$

$z_4 = 4$

$z_5 = 2$



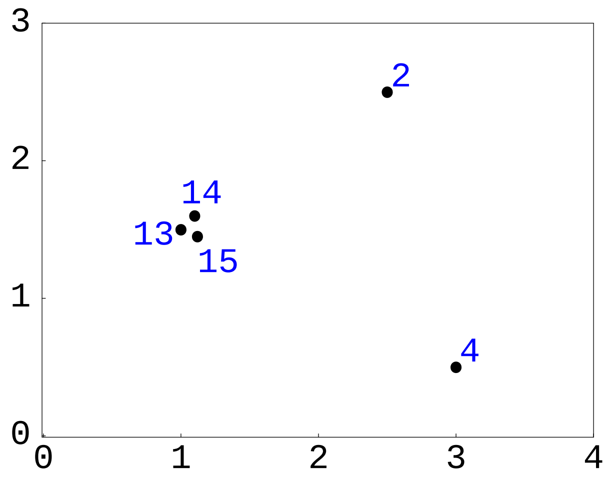


$$\begin{pmatrix} 5 & 4.34 & 4.39 & 0.53 & 0.82 & -1 \\ 4.34 & 5 & 4.30 & 0.56 & 0.95 & -1 \\ 4.39 & 4.30 & 5 & 0.61 & 0.88 & -1 \\ 0.53 & 0.56 & 0.61 & 5 & 0.64 & -1 \\ 0.82 & 0.95 & 0.88 & 0.64 & 5 & -1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ \lambda_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$\Rightarrow$ Correlated mean, $\bar{z}_C = \underbrace{0.16}_{w_1} \cdot 13 + \underbrace{0.10}_{w_2} \cdot 14 + \underbrace{0.09}_{w_3} \cdot 15 + \underbrace{0.34}_{w_4} \cdot 4 + \underbrace{0.31}_{w_5} \cdot 2 = 6.8$

Uncorrelated mean: again

$z_1 = 13$
$z_2 = 14$
$z_3 = 15$
$z_4 = 4$
$z_5 = 2$



$$\text{cov}(s) = \begin{cases} 5, & s = 0 \\ 0, & s > 0 \end{cases}$$

$$\begin{pmatrix} 5 & 0 & 0 & 0 & 0 & -1 \\ 0 & 5 & 0 & 0 & 0 & -1 \\ 0 & 0 & 5 & 0 & 0 & -1 \\ 0 & 0 & 0 & 5 & 0 & -1 \\ 0 & 0 & 0 & 0 & 5 & -1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ \lambda_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Question. What is the solution of this system of linear equations?
Question. Does the solution change when we change $\text{cov}(0) = 5$ into, say, $\text{cov}(0) = 1$?

Redundancy matrix

Correlation/covariance between observations:

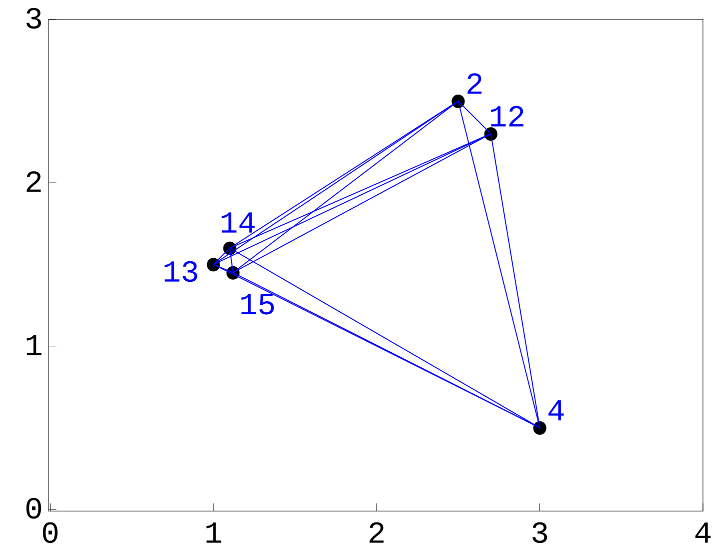
$\Rightarrow$ Observations are not independent

$\Rightarrow$ Redundancies exist between observations

Redundancy matrix C contains all these redundancies.

Remark.
The redundancy matrix, given a covariance function, only depends on the locations of the observations

and not on the actual values of the observations



$$\downarrow \ \text{cov}(s) = 5 \cdot e^{-s}$$

$$C := \begin{pmatrix} 5 & 4.34 & 4.39 & 0.53 & 0.82 & -1 \\ 4.34 & 5 & 4.30 & 0.56 & 0.95 & -1 \\ 4.39 & 4.30 & 5 & 0.61 & 0.88 & -1 \\ 0.53 & 0.56 & 0.61 & 5 & 0.64 & -1 \\ 0.82 & 0.95 & 0.88 & 0.64 & 5 & -1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Dept. of Geoscience & Remote Sensing

First Conclusion and Further Questions

Conclusion: Redundancy matrix enables the division of weights over clusters of observations.

Question. How to obtain a covariance function?
or/and, how to determine if correlation exist?

Question. Is any covariance function we come up with suited for our needs?

Question. How to derive such linear systems, what assumptions are made, what optimization is aimed for?

Question. Can we get a quality description of the solution?

Question. What is the bad news?

Dept. of Geoscience & Remote Sensing

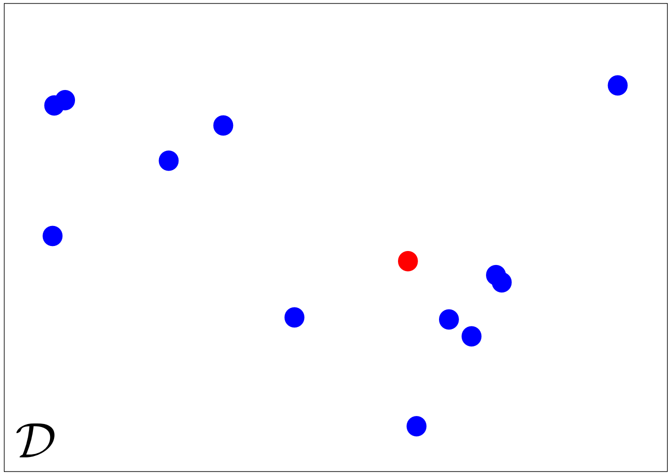
B. The (Simple) Kriging equations

Stochastic interpolation framework

Given are

- n (height) observations $z_1, \ldots, z_n$,
- at locations $\mathbf{p_1}, \ldots, \mathbf{p_n}$

in a spatial region of interest $\mathcal{D}$.



Assumption:
The height signal $z(\mathbf{p})$ for $\mathbf{p}$ in $\mathcal{D}$ is described by a random function

That is, the height at any location $\mathbf{p} \in \mathcal{D}$ is described by a random variable $z_{\mathbf{p}}$ corresponding to a location dependent height signal distribution.

The random function $z(\mathbf{p})$ for $\mathbf{p} \in \mathcal{D}$ is the set of all local random variables together.

Conditions on the random function:

1) On points. Mean exists and is known, and its (expected) value is independent of the location in $\mathcal{D}$.

2) On pair of points. Covariance exists, is known and is location invariant

Wish:
A) We want to obtain an interpolation value $\hat{z}_0$ at location $\mathbf{p}_0$ as a linear combination

$$\hat{z}_0 \; = \; w_1 \cdot z_1 + w_2 \cdot z_2 + \cdots + w_n \cdot z_n$$

of the n observations, interpreted as deviations of the mean.

B) This interpolation should moreover be optimal in the following sense: the expected error variance is minimal.

Remark. Mean exists $\Rightarrow$ no trend allowed in the data!

Interpolation deviations of the mean

Deviation of the mean. Let μ be the given mean.

$$d_i := z_i - \mu$$

We are looking for weights $w_1, \ldots, w_n$ to obtain a height estimation $\hat{z}_0$ at location $\mathbf{p}_0$ as

$$\begin{aligned} \hat{z}_0 &= \mu + \sum_{i=1}^{n} w_i \cdot (z_i(\mathbf{p}_i) - \mu) \\ &= \mu + \sum_{i=1}^{n} w_i \cdot d_i \end{aligned}$$

Interpolation error

The (signed) interpolation error, r_0, is the difference between the real, but unknown, height z_0 and our estimation $\hat{z}_0$:

$$
\begin{aligned}
r_0 &= \hat{z}_0 - z_0 = (\hat{z}_0 - \mu) - (z_0 - \mu) \\
&= (\mu + \sum_{i=1}^{n} w_i \cdot d_i - \mu) - (z_0 - \mu) = \sum_{i=1}^{n} w_i \cdot d_i - d_0 \\
&= \hat{d}_0 - d_0
\end{aligned}
$$

The variance of the interpolation error

The variance of the error is given by

$$
\begin{aligned}
\text{var}(r_0) &= \text{var}(\hat{d}_0 - d_0) \qquad \text{(Lecture 5?: Variance of the sum)} \\
&= \text{var}(\hat{d}_0) + \text{var}(d_0) - 2 \cdot \text{cov}(\hat{d}_0, d_0) \\
&= (\text{Expand } \hat{d}_0) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \cdot w_j \cdot \text{cov}(s_{ij}) + \text{cov}(0) - 2 \cdot \sum_{i=1}^{n} w_i \cdot \text{cov}(s_{ij})
\end{aligned}
$$

where

- cov: $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ indicates the (given) covariance function,
- $s_{ij} := \|\mathbf{p}_i - \mathbf{p}_j\|$ denotes the horizontal distance between points $\mathbf{p}_i$ and $\mathbf{p}_j$

Interpolation goal

Remark. The following is known (given to us):

1. the observation locations, $\mathbf{p_1}, \ldots, \mathbf{p_n}$,
2. the estimation location, $\mathbf{p_0}$,
3. the covariance function, cov$(.)$

Therefore the error variance is just a function of the weights $w_1, \ldots, w_n$.

Goal.
Determine those values of $w_1, \ldots, w_n$ that minimize the error variance var(r_0)

The Simple Kriging system

If the error variance var(r_0) is minimal, all partial derivatives vanish:

$$(\frac{1}{2}) \cdot \frac{\partial\, \text{var}(r_0)}{\partial\, w_i} \quad = \quad 0, \qquad \text{for } i = 1, \ldots, n$$

This gives us the Simple Kriging system

$$\sum_{j=1}^{n} w_j \cdot \text{cov}(s_{ij}) - \text{cov}(s_{i0}) = 0, \qquad \text{for } i = 1, \ldots, n$$

or, in matrix notation

$$\begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} c_{01} \\ c_{02} \\ \vdots \\ c_{0n} \end{pmatrix}$$

with $c_{ij} = \text{cov}(s_{ij})$, where $s_{ij} = \|\mathbf{p}_i - \mathbf{p}_j\|$.

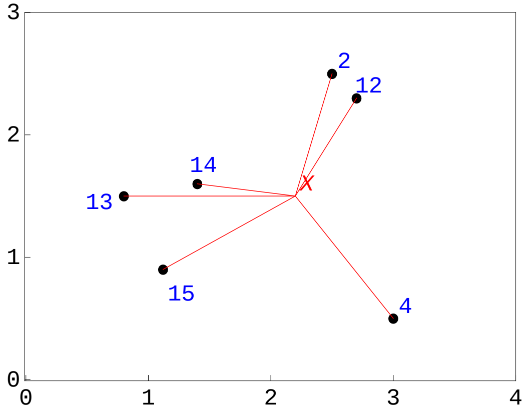
Solving the Simple Kriging system

To determine: weight for each observation: $(w_1, w_2, \ldots, w_n)^T$

Known, given a covariance function:

1) Redundancy matrix, as for "Kriging the mean", except now last row and column are missing

2) Proximity vector, $d_n = (c_{01}, c_{02}, \ldots, c_{0n})$, the covariances between interpolation location and observations.



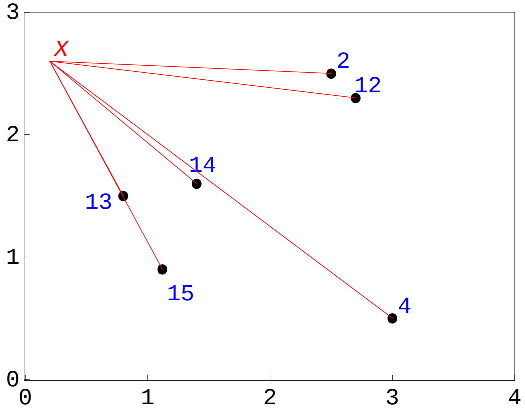
$\leftarrow d_n = (1.2, 2.2, 1.5, 1.4, 1.8, 1.9)$

cov$(s) = 5 \cdot e^{-s}$

$z_1 = 13, z_2 = 14, z_3 = 15$

$z_4 = 4, z_5 = 2, z_6 = 12$

$d_n = (1.4, 1.0, 0.7, 0.2, 0.5, 0.4) \rightarrow$



Question. Why has "Kriging the mean" no proximity vector?

The Simple Kriging variance

With $\mathbf{C}$ the redundancy vector and $\mathbf{d_n}$ the proximity vector, the vector of optimal weights $\mathbf{w_n} = \{w_1, \ldots, w_n\}$ is obtained as:

$$\mathbf{w_n} \;=\; \mathbf{C}^{-1} \cdot \mathbf{d_n}$$

Substituting this optimal solution for the weights in the error variance expression gives us a second result:

Simple Kriging error variance

$$\text{var}_{SK}(\mathbf{p_0}) \;=\; \text{cov}(0) - \sum_{i=1}^{n} w_i \cdot \text{cov}(\mathbf{p_i} - \mathbf{p_0}) \;=\; \text{cov}(0) - \mathbf{w_n} \cdot \mathbf{d_n}$$

Remarks

Mr. Krige

The Simple Kriging variance only depends on the observation locations, not on the actual values of the observations,

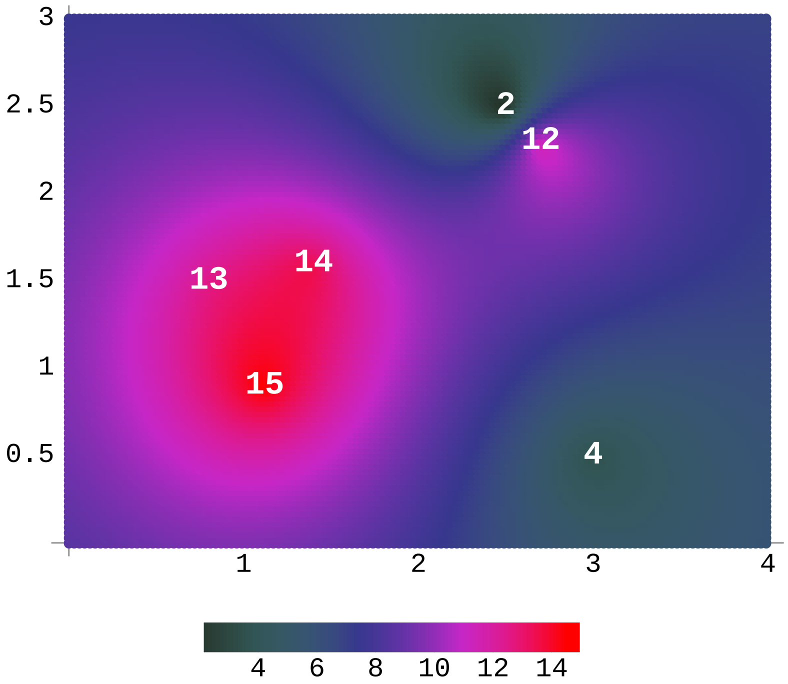
The SK variance is low near observations while it increases with increasing distance to the observations

The SK variance is maximal when no correlation with the observations exists

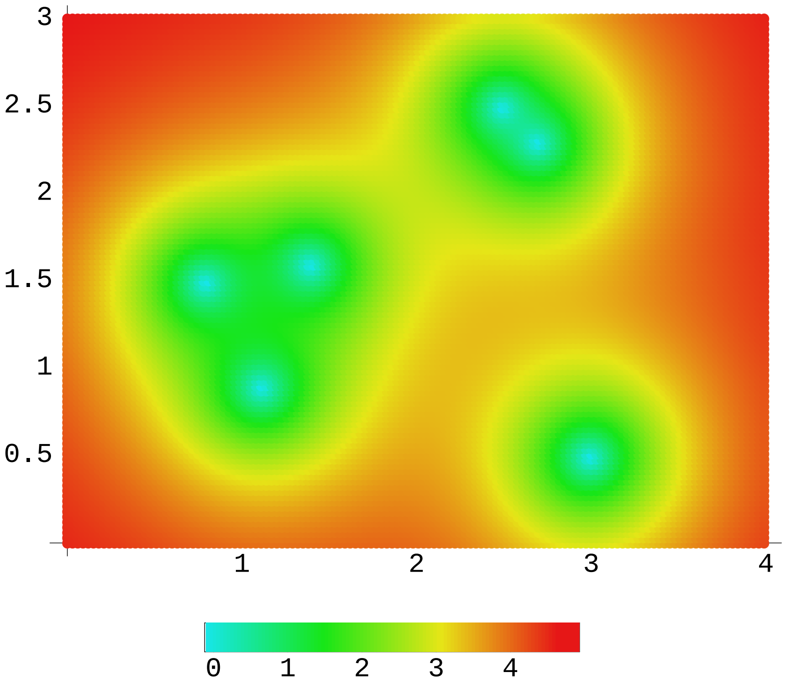
Method only works if the variance-covariance matrix is invertible. This is guaranteed (Linear Algebra) if we use a positive definite covariance function to fill the VC-matrix.

SK interpolation and variance results

Simple Kriging interpolation

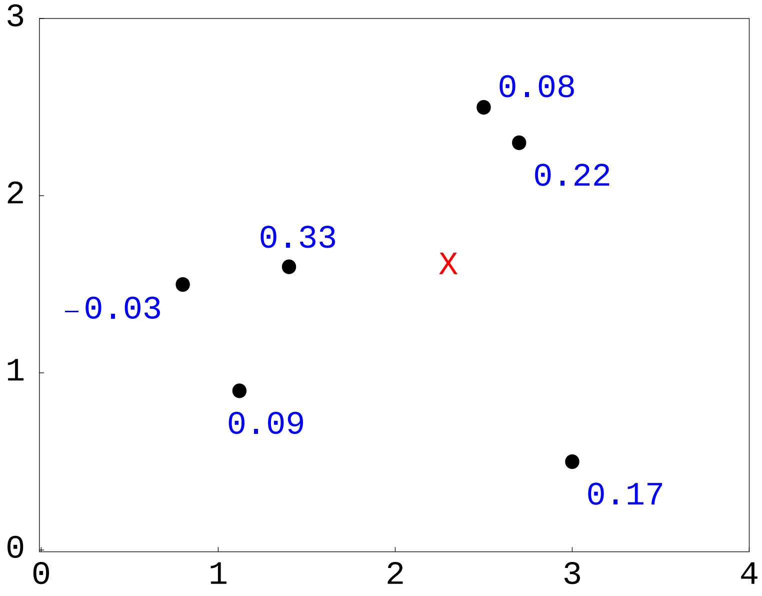


Simple Kriging variance

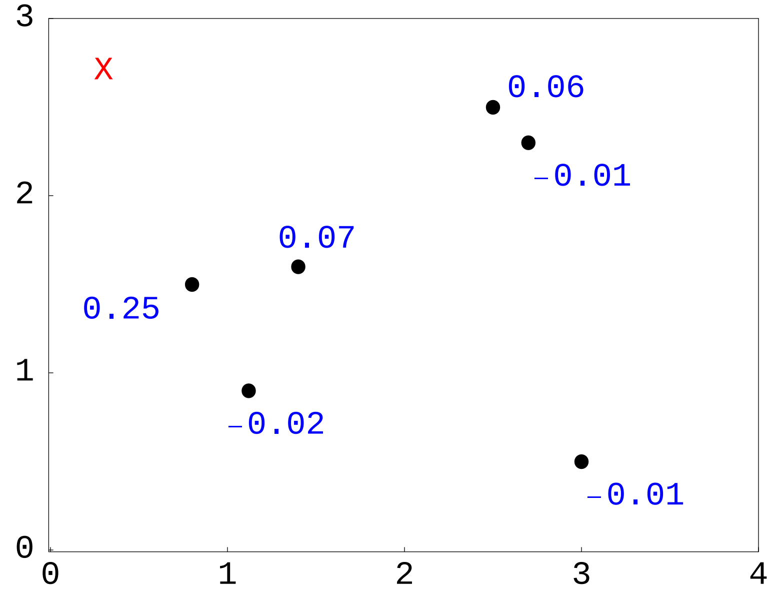


Simple Kriging weights

Total weight = 0.86



Total weight = 0.35

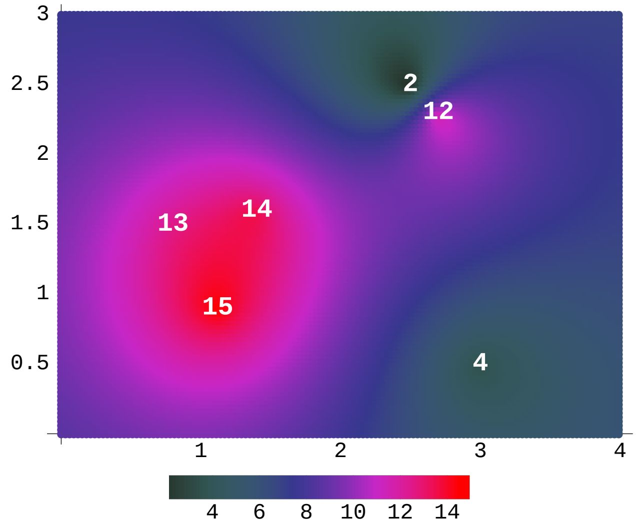


- Weights can be negative (esp. for observations ‘in the back‘)
- Weight is divided between the observations and the mean

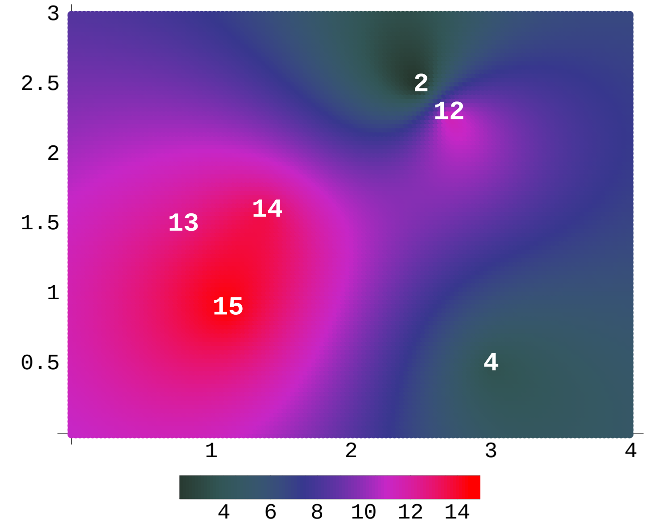
(X: interpolation location)

Influence covariance function

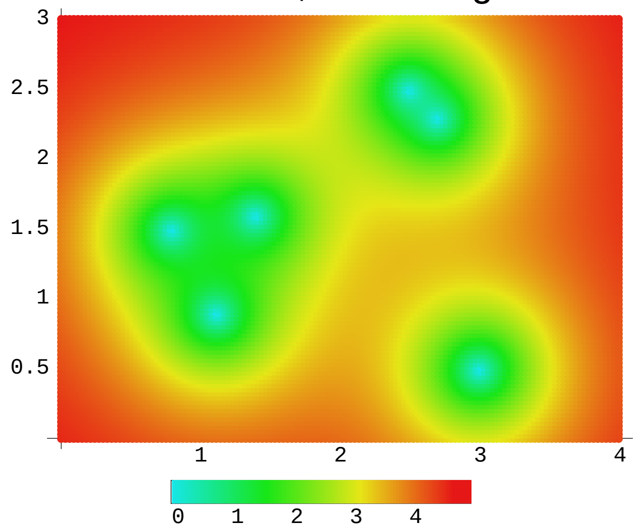
SK interpolation, short range



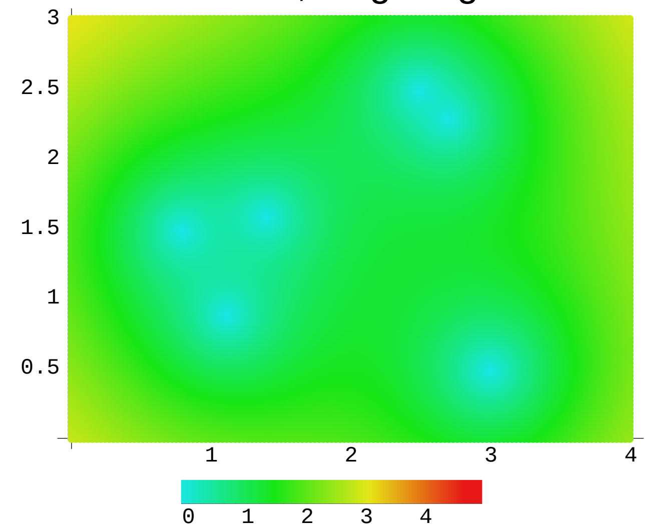
SK interpolation, long range



SK variance, short range



SK variance, long range



Ordinary Kriging

Setting Simple Kriging:

1. Mean is known
2. Aim for deviations from the mean

In case mean is unknown: solve Ordinary Kriging system:

$$\begin{pmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \lambda \end{pmatrix} = \begin{pmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \\ 1 \end{pmatrix}$$

Last row ensures $\sum_{i=1}^{n} w_i = 1$; This condition is added to ensure that the solution is unbiased.

Locations as on Slide 23

Unbiasedness: $E\{z\} = \bar{z}$ for all z in domain $\mathcal{D}$.

Expected value of z equals the (unknown) mean of z over $\mathcal{D}$.

Ordinary Kriging interpolation & Variance

For an estimation of the height $\hat{z}_0$ from observations $z_1, \ldots, z_n$.

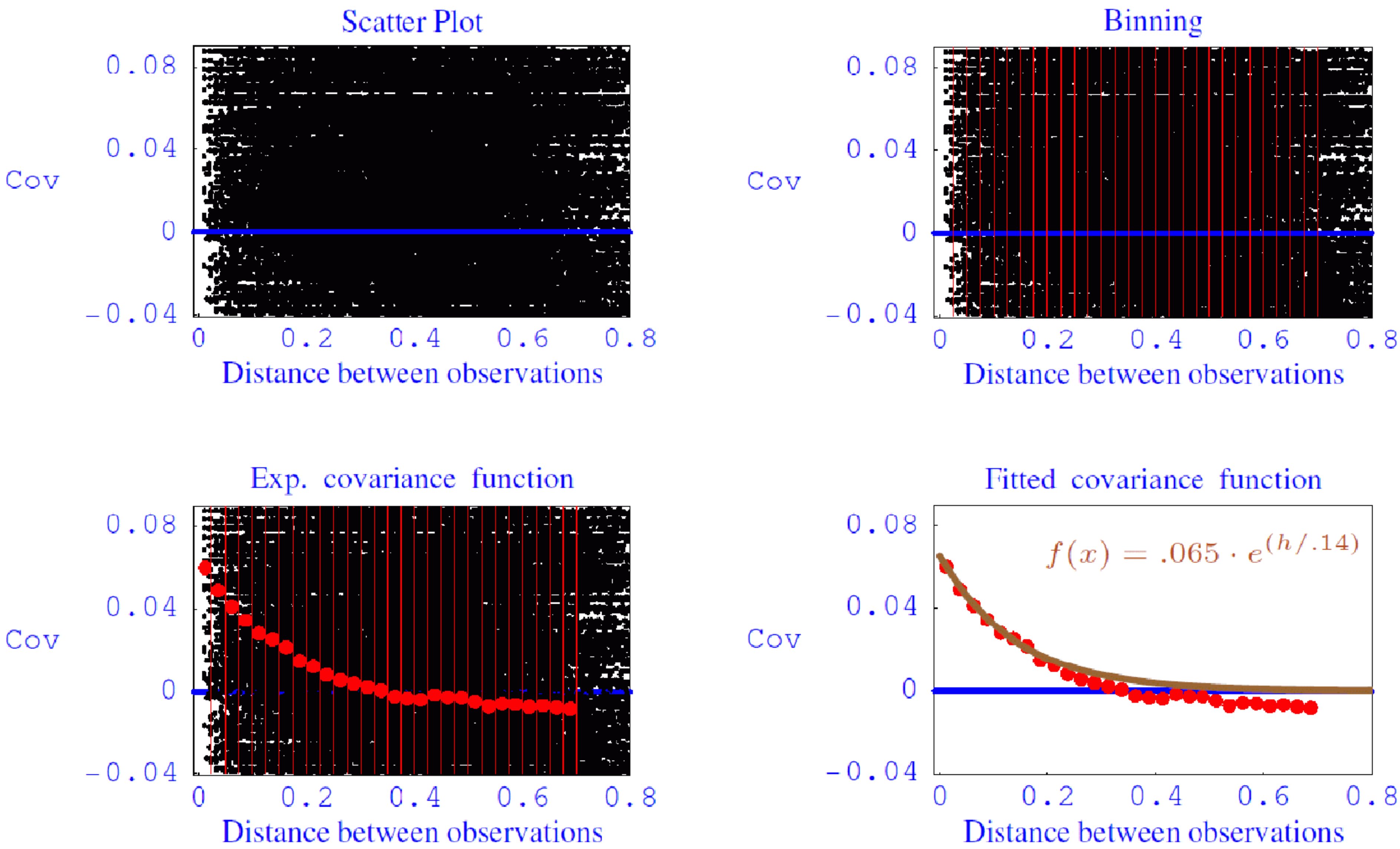
$$\begin{cases} \hat{z}_0 & = & w_1 \cdot z_1 + \cdots + w_n \cdot z_n \\ \text{var}(\hat{z}_0) & = & \text{cov}(0) - \mathbf{w}_n \cdot \mathbf{d}_n \end{cases}$$

Here $\mathbf{w}_n$ is the vector of weights, and $\mathbf{d}_n$ the proximity vector.

C. Obtaining a covariance function

Dept. of Geoscience & Remote Sensing

Experimental covariance function, I



Experimental covariance function, II

1) Determine experimental covariances

$$[\|\mathbf{p_i} - \mathbf{p_j}\|, (z_i - \mu)(z_j - \mu)]$$

as function of distance, for (a suitable subset of) all pairs of observations

2) Average the obtained experimental covariances per suited bin $\longrightarrow$ experimental covariogram

3) Fit an admissible covariance model to the experimental covariogram $\longrightarrow$ covariance function

Definition. A covariance model is admissible if for any choice of model parameters the obtained covariance function is positive definite

Covariance function models

Admissable covariance models:

Exponential model with practical range a

$$\text{cov}(h) = \exp(\frac{-3h}{a})$$

Spherical model with range a

$$\text{cov}(h) = \begin{cases} 1 - 1.5\frac{h}{a} + 0.5(\frac{h}{a})^3 & \text{if } h \leq a \\ 0 & \text{else} \end{cases}$$

Gaussian model with practical range a

$$\text{cov}(h) = \exp(-\frac{3h^2}{a^2})$$

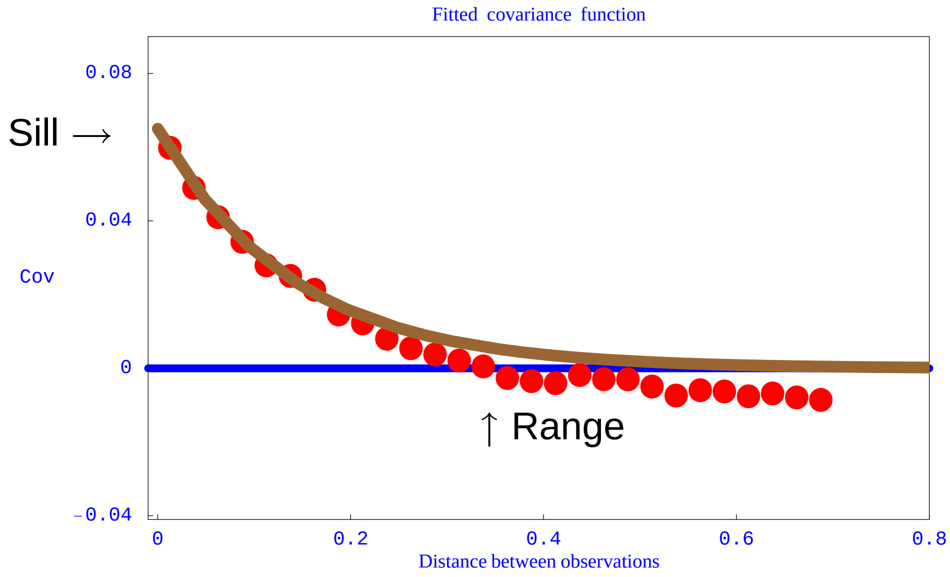
Nugget model

$$\text{cov}(h) = \begin{cases} 1 & \text{if } h = 0 \\ 0 & \text{else} \end{cases}$$

Covariance function parameters

Range: distance at which covariance/correlation is vanished.

Sill: initial covariance, i.e. covariance at very short distance. Scale the functions of the previous pages vertically using the sill!



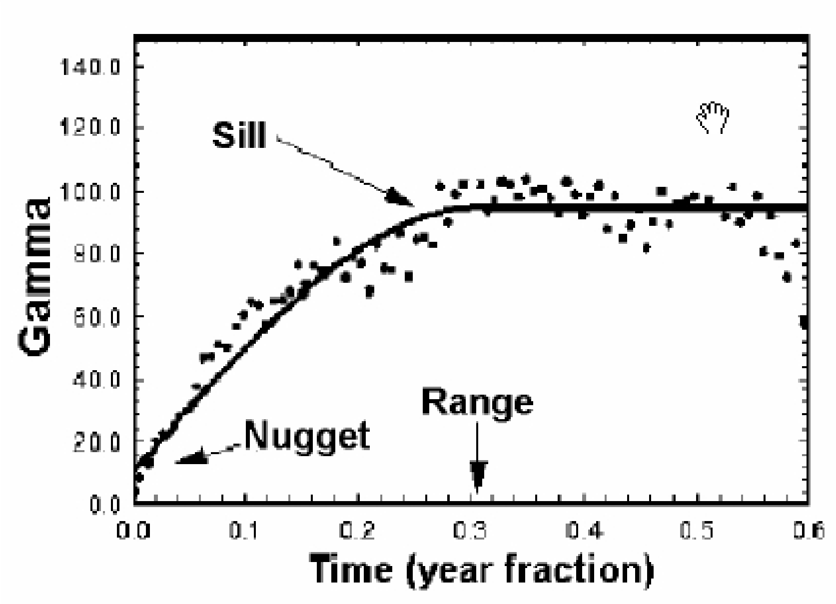
Variogram: increasing dissimilarity

Variogram definition.

$$\gamma(\mathbf{h}) \quad = \quad \frac{1}{2}E\{(Z(\mathbf{p}+\mathbf{h})-Z(\mathbf{p}))^2\}$$

Variogram properties.

1. $\gamma(\mathbf{0}) \ = \ 0.$
2. $\gamma(\mathbf{h}) \geq 0.$
3. $\gamma(\mathbf{h}) = \gamma(-\mathbf{h}).$
4. $\lim_{|\mathbf{h}|\to\infty} \frac{\gamma(\mathbf{h})}{|\mathbf{h}|^2} = 0.$



The variogram measures the average dissimilarity

Relation Variogram/Covariance function

Claim. Uniform mean and covariance $\Rightarrow \gamma(\mathbf{h}) = \text{cov}(0) - \text{cov}(\mathbf{h})$.

Proof.

$$
\begin{aligned}
\gamma(\mathbf{h}) &= \frac{1}{2}E\{(Z(\mathbf{p}) - Z(\mathbf{p}'))^2\} \\
&= \frac{1}{2}E\{((Z(\mathbf{p}) - \mu) - (Z(\mathbf{p}') - \mu))^2\} \\
&= \frac{1}{2}E\{(Z(\mathbf{p}) - \mu)^2\} + \frac{1}{2}E\{(Z(\mathbf{p}') - \mu)^2\} \\
&\quad -E\{(Z(\mathbf{p}) - \mu)(Z(\mathbf{p}') - \mu)\} \\
&= E\{(Z(\mathbf{p}) - \mu)^2\} - E\{(Z(\mathbf{p}) - \mu)(Z(\mathbf{p}') - \mu)\} \\
&= \text{cov}(0) - \text{cov}(s)
\end{aligned}
$$

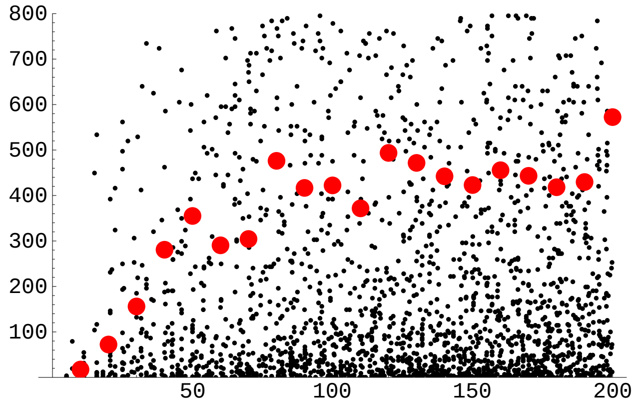
$\square$

The Variogram is a well-known alternative for the covariance function.

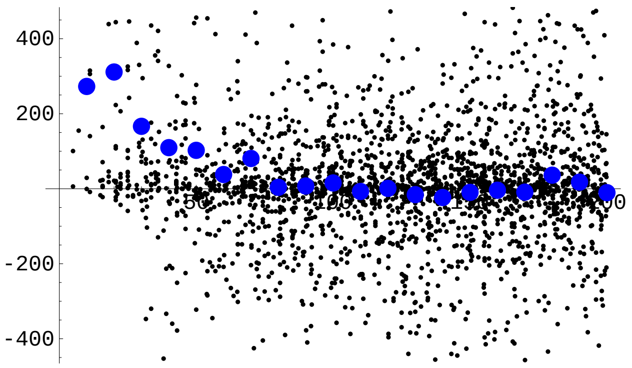
Experimental variogram + Cov. function

Example

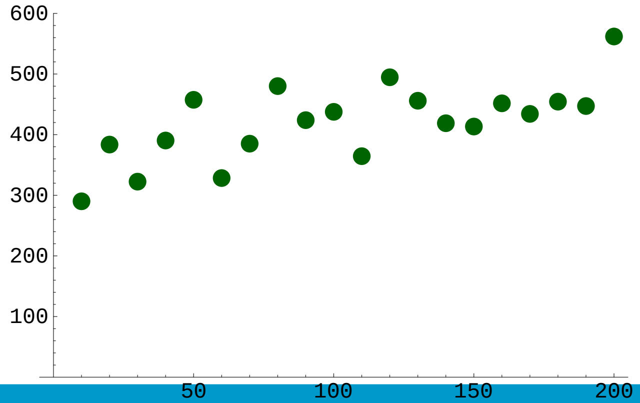
Variogram



Covariance function



Covariance function - variogram



Conclusions

Kriging is the best interpolation technique in a particular sense, provided that the variances of and covariances between observations are known.

Kriging takes correlation between (nearby) observations into account, using

- A covariance function, or,
- A variogram

Kriging not only provides an estimate but also a variance of that estimate

Different Kriging dialects exist, used depending on the available information. Discussed here:

1. Kriging the mean (estimating the mean)
2. Simple Kriging (mean is known)
3. Ordinary Krigng (mean is unknown)

Dept. of Geoscience & Remote Sensing

Appendix: some Bonus slides

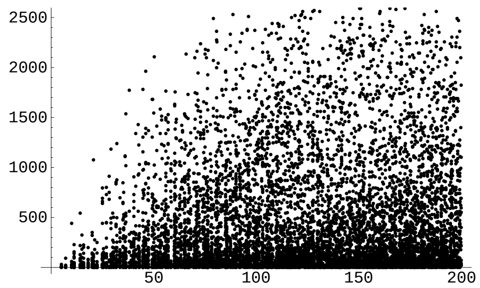
Dept. of Geoscience & Remote Sensing

Experimental variation and variogram.

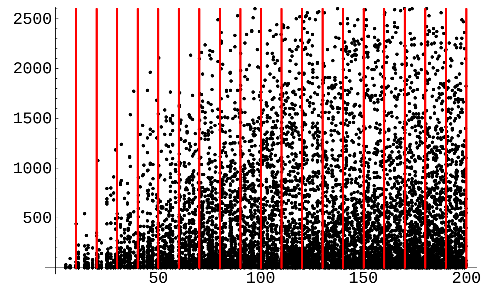
Dissimilarity between two observations z_i and z_j:

$$\gamma_{ij} = \frac{(z_i - z_j)^2}{2} = \frac{(z_i(\mathbf{p}_i) - (z_j(\mathbf{p}_j))^2}{2}$$

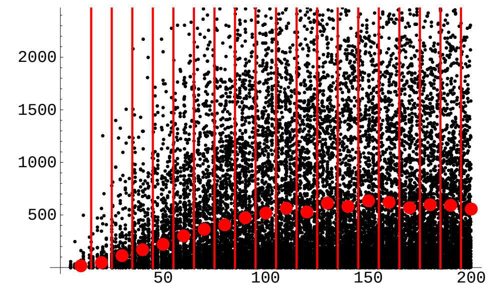
Determine all dissimilarities



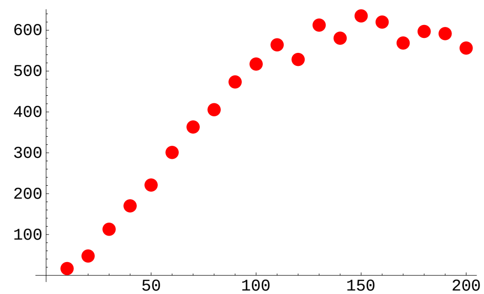
Group them



Take group average



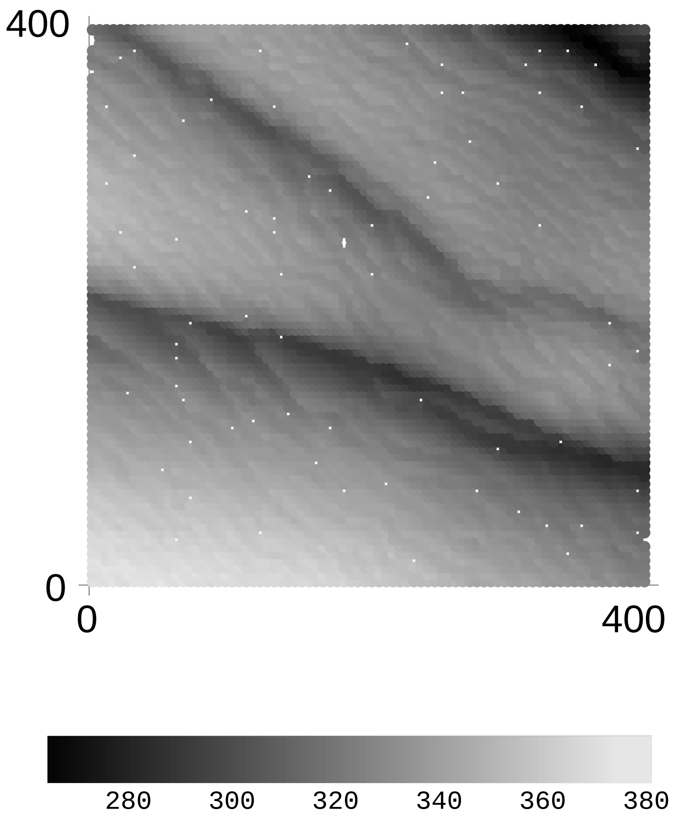
Result: experimental variogram



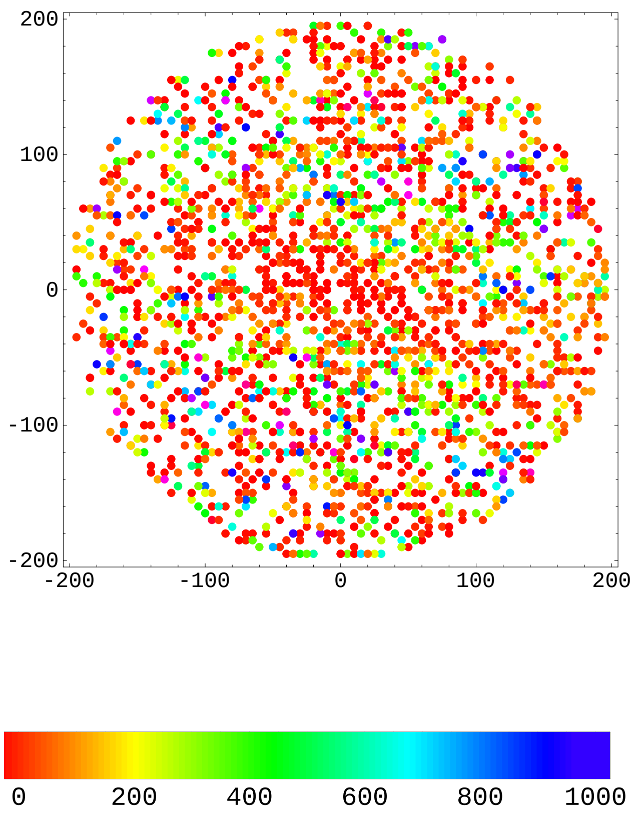
Dept. of Geoscience & Remote Sensing

Anisotropic case.

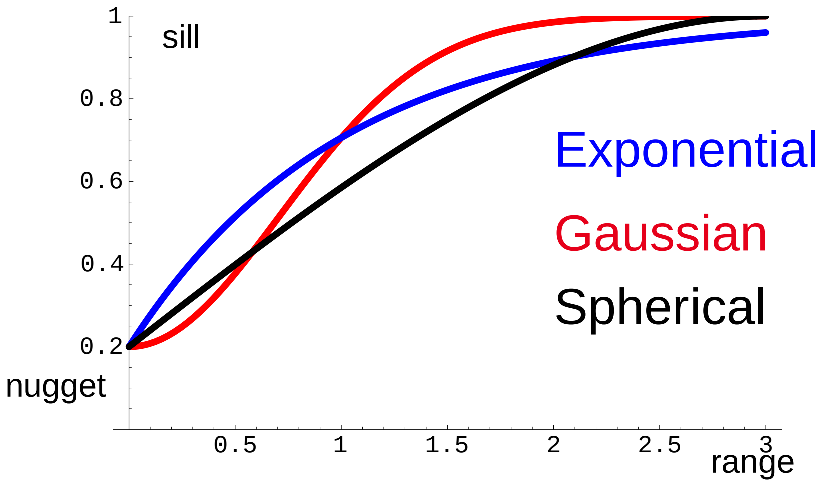
Topography



Variogram cloud: experimental dissimilarities



Variogram models



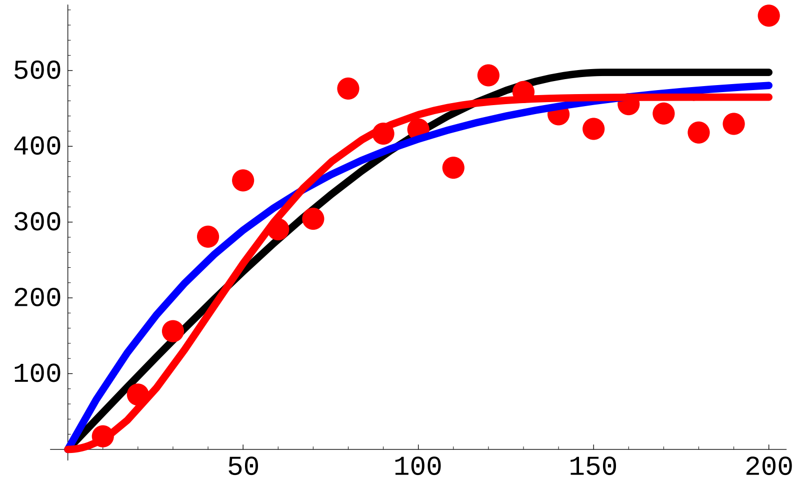
$$\gamma_{\text{expo}}(h) = \sigma^2(1 - e^{-\frac{3h}{R}})$$

$$\gamma_{\text{gauss}}(h) = \sigma^2(1 - e^{-\frac{(3h)^2}{R^2}})$$

$$\begin{aligned} \gamma_{\text{spher}}(h) &= \sigma^2(\frac{3h}{2R} - \frac{h^3}{2R^3}), & h \leq R \\ &= \sigma^2, & h > R \end{aligned}$$

Moreover: linear model, nugget model.

Our example.



Exactness

Claim.

1. Ordinary Kriging is exact, that is, OK respects observations.
2. The variance at an observation location equals zero.

Proof.

1. W.l.o.g. assume that $p_1 = p_0$. Then the OK system reads

$$\begin{pmatrix} C_{00} & C_{02} & \dots & C_{0n} & 1 \\ C_{20} & C_{22} & \dots & C_{0n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n0} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \lambda \end{pmatrix} = \begin{pmatrix} C_{00} \\ C_{20} \\ \vdots \\ C_{n0} \\ 1 \end{pmatrix}$$

   Clearly, $\mathbf{w} = \{1, 0, 0, \dots, 0\}$ is a solution. As the redundancy matrix is non-singular, this is the unique solution. One word alternative proof??

2. Var$\{\hat{z}_0 - z_0\}$ = cov$(0) - \sum_{i=1}^{n} w_i C_{i0} - \lambda$ = cov$(0) - C_{00} = 0$.

OK as interpolator

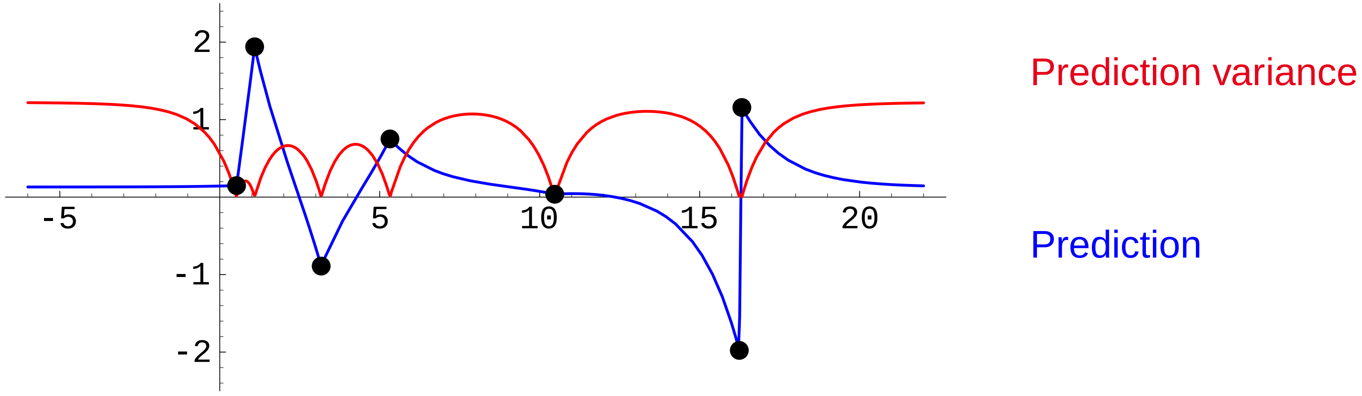
1. **Input:**
   - observations $\mathcal{O}_z = \{z_1, \ldots, z_n\}$ at obs. locations $\mathcal{O}_{x,y} = \{p_1, \ldots, p_n\}$.
   - interpolation locations $\mathcal{Q}_{x,y} = \{q_1, \ldots, q_m\}$.
2. Determine an exp. covariance function `expcov` from $\mathcal{O}_z$ and $\mathcal{O}_{x,y}$.
3. Fit a positive definite covariance model `cov` to `expcov`.
4. Determine the redundancy matrix $\mathcal{C}_n$, by applying `cov` to the distances between the observation locations.
5. For every location q_j in $\mathcal{Q}_{x,y}$:
   - Determine the proximity vector $\mathbf{d}_n$ by applying `cov` to the distances $\{\|q_j - p_1\|, \ldots, \|q_j - p_n\|\}$.
   - Find weights $w_1(q_j), \ldots, w_n(q_j)$ by solving the OK-system $\mathcal{C}_n \cdot \mathbf{w}_n = \mathbf{d}_n$.
   - Get estimation and estimation variance by

$$\begin{cases} \hat{z}_j & = & w_1(q_j) \cdot z_1 + \cdots + w_n(q_j) \cdot z_n \\ \text{var}(\hat{r}_j) & = & \text{cov}(0) - \mathbf{w}_n \cdot \mathbf{d}_n \end{cases}$$

Question: assumption???

Interpolation example.

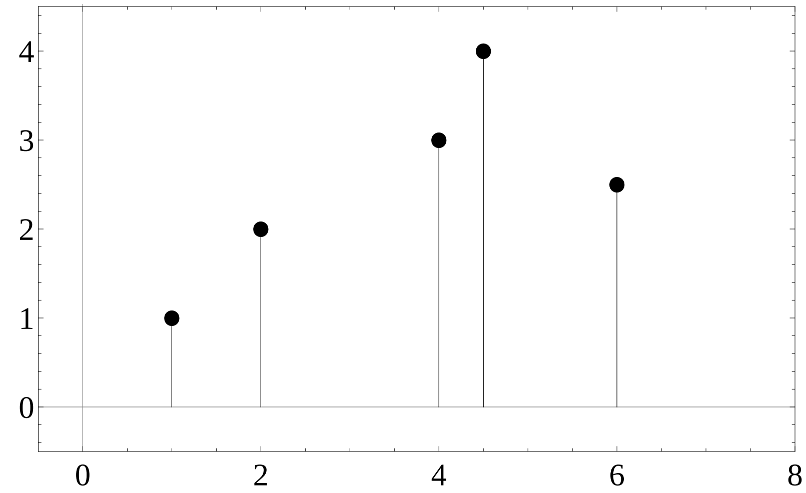
$$\mathcal{C}_n = \begin{pmatrix} 1 & 0.66 & 0.14 & 0.03 & 0 & 0 & 0 & 1 \\ 0.66 & 1 & 0.21 & 0.04 & 0 & 0 & 0 & 1 \\ 0.14 & 0.21 & 1 & 0.20 & 0 & 0 & 0 & 1 \\ 0.03 & 0.04 & 0.20 & 1 & 0.02 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.02 & 1 & 0.01 & 0.01 & 1 \\ 0 & 0 & 0 & 0 & 0.01 & 1 & 0.94 & 1 \\ 0 & 0 & 0 & 0 & 0.01 & 0.94 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$



Remarks. Exactness, vanishing variance, Kriging the mean.

Exercises

Exercise, 1D Kriging



Exercise 7.1 In the figure, the points $p_1 = (1,1)$, $p_2 = (2,2)$, $p_3 = (4,3)$, $p_4 = (4.5,4)$, and $p_5 = (6,2.5)$ are shown. Think of these points as the result of measuring some signal as function of time in seconds. In addition a covariance function $c(t) = e^{-2t}$ is given

a). Is this covariance function admissable? Why?

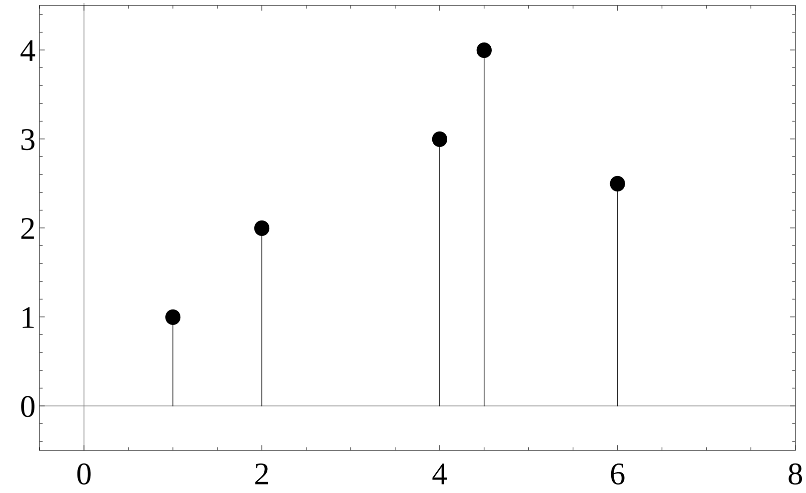
b). Sketch $c(t)$. What are its extrema?

c). What is the covariance between two observations that are 1 second apart?

d). After what time drops $c(t)$ below one tenth of its maximal value?

e). How would you define the range and sill for this covariance function?

Exercise, continued



Exercise 7.2 Continuation of Exercise 3.1

a). Determine the covariances between each of the five measurements in the figure. What is the maximum covariance?

b). Convert these covariance into correlations.

c). Write down the Ordinary Kriging (OK) redundancy matrix of the five measurements in the figure.

d). What is the OK proximity vector for an interpolation at $t = 3$?

e). And what is the OK proximity vector for an interpolation at $t = 4$?

f). Use Ordinary Kriging to interpolate at $t = 4$. How are the weights distributed over the observations? What is the sum of the weights?

g). (Matlab) Use Ordinary Kriging to interpolate at $t = 3$. What are the weights now?

h). (Matlab) Use Ordinary Kriging to interpolate the interval $t = [0, 8]$.

Answers, Exercise 7.1

a) The function is admissable, the exponential function has been tested by mathematicians and is positive definite. This means that when you use $c(t)$ for filling a redundancy matrix, this matrix will always be invertable.

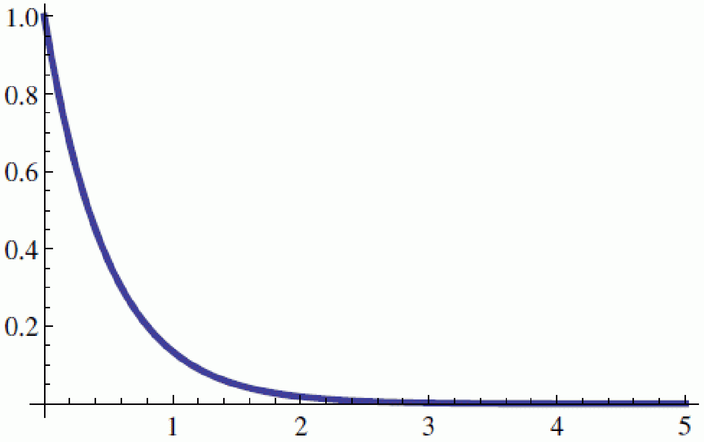
b) $\rightarrow$

c) Fill in $t = 1$ in the covariance function:
$c(1) = 1/e^2 = 0.135$

d) Read from the graph (or use computer):If $t = 1.15$, then $c(t) = 0.1$

e) Range: about 2, as after $t = 2$, $c(t)$ is very small. Sill: 1 as $c(0) = 1$

```
c[t_] := E^(-2 t);
Plot[c[t], {t, 0, 5}, PlotRange → All,
 LabelStyle → Larger, PlotStyle → {Thickness[.01]}]
```



Answers, Exercise 7.2

f) See matrix on the right. Observations are ordered from left to right. Max. covariance is 1 (on the diagonal, the self-covariances or variances). The max. covariance between different observations is 0.37, between obs 3 and obs 4.

```
cov = Table[Table[N[c[Abs[pp[[i, 1]] - pp[[j, 1]] ]] ], {i, 1, 5}], {j, 1, 5}];
cov // MatrixForm
```

ixForm=

1.	0.135335	0.00247875	0.000911882	0.0000453999
0.135335	1.	0.0183156	0.00673795	0.000335463
0.00247875	0.0183156	1.	0.367879	0.0183156
0.000911882	0.00673795	0.367879	1.	0.0497871
0.0000453999	0.000335463	0.0183156	0.0497871	1.

g) These are allready correlations, variance on the diagonal is 1.

h) To get OK redundancy matrix: add extra row & column with notably ones to the result from f).

```
redmat = Transpose[
   Append[Transpose[Append[cov, {1, 1, 1, 1, 1}]], {1, 1, 1, 1, 1, 0}]];
redmat // MatrixForm
```

ixForm=

1.	0.135335	0.00247875	0.000911882	0.0000453999	1
0.135335	1.	0.0183156	0.00673795	0.000335463	1
0.00247875	0.0183156	1.	0.367879	0.0183156	1
0.000911882	0.00673795	0.367879	1.	0.0497871	1
0.0000453999	0.000335463	0.0183156	0.0497871	1.	1
1	1	1	1	1	0

i) Proximity vector for $t = 3$, first the general function, followed by the proximity vector for $t = 3$. The last entry, 1, belongs to Ordinary Kriging.

```
proc[t_] := Append[Table[ N[c[Abs[t - pp[[i, 1]] ] ]], {i, 1, Length[pp]}], 1]
proc[3]
{0.0183156, 0.135335, 0.135335, 0.0497871, 0.00247875, 1}
```

j) Proximity vector for $t = 3$. Note an observation has been done at $t = 4$, so the 3rd entry is now 1 as well.

```
proc[4]
{0.00247875, 0.0183156, 1., 0.367879, 0.0183156, 1}
```

Answers, Ex. 7.2 (continued)

k) Ordinary Kriging for interpolation at $t = 4$. Easy, as we have an observation at $t = 4$ and because OK is exact, it simply reproduces the observation. Therefore the interpolated value equals 3 (= value of observation). The weight of this observtaion is 1, the weights for other observations are zero.

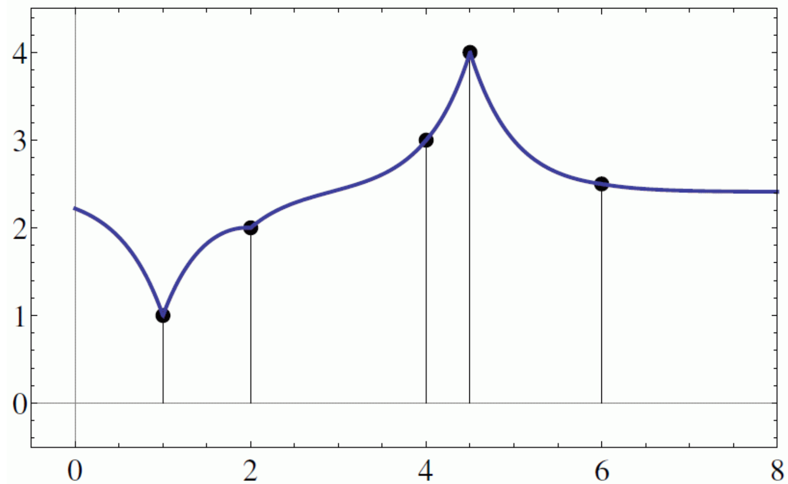
l) Now we do have to work: Multiply the inverse of the redundancy matrix with the proximity vector for $t = 3$

```
weights3 = Inverse[redmatD].proc[3]
```

Weights are appr. $w_1 = 0.16; w_2 = 0.29; w_3 = 0.26; w_4 = 0.12; w_5 = 0.17;$ To get the interpolation result: multiply weights with observations. Final result: 2,43. Note we didn't use the last entry of the weight vector, the -0.18.

m) Same procedure as for parts k) and l), but now for every value of t. Note that the interpolation result indeed passes through the observations (exactness):

```
plkr = Plot[Take[Inverse[redmat].proc[t], 5].waarnemingen,
    {t, 0, 8}, PlotStyle → {Thickness[.005]}];
Show[{shwrn, plkr}]
```



Dept. of Geoscience & Remote Sensing