Slide 1

Introduction

Data: set of n attribute measurements $\{z(\mathbf{s}_i), i = 1, \ldots, n\}$, available at n sample locations $\{\mathbf{s}_i, i = 1, \ldots, n\}$

Objectives:

- quantify spatial *auto-correlation*, or attribute dissimilarity typically expressed as: $\frac{1}{2}[z(\mathbf{s}_i) - z(\mathbf{s}_j)]^2$ as a function of separation distance between sample pairs $\mathbf{s}_i$ and $\mathbf{s}_j$
- introduce the sample semivariogram, its characteristics, and provide some examples
  NOTE: *Spatial auto-correlation is a second-order characteristic of spatial variation, and hence the sample semivariogram should be computed from data whose spatial variation is not explained by first-order effects*
- justify the need of going beyond the sample semivariogram to a semivariogram model
- introduce parametric functions of distance that can be used as formal theoretical semivariogram models
- discuss issues of fitting semivariogram models to sample semivariogram values

Slide 2

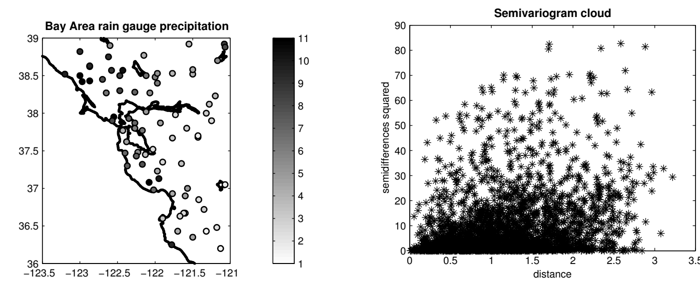
Semivariogram Cloud

Definition: A scatter-plot of *attribute* squared semidifferences between all possible pairs of samples measured at different locations, versus their separation distance

Computational procedure:

1. construct Euclidean distance matrix $\mathbf{D} = [d_{ij}, i = 1, \ldots, n, j = 1, \ldots, n]$ between all n^2 pairs of data locations, where d_{ij} is defined as: $d_{ij} = ||\mathbf{h}_{ij}|| = ||\mathbf{s}_i - \mathbf{s}_j||$
2. construct squared semidifference matrix $\mathbf{E} = [e_{ij}, i = 1, \ldots, n, j = 1, \ldots, n]$ between all n^2 pairs of attribute values, where e_{ij} is defined as: $e_{ij} = \frac{1}{2}[z(\mathbf{s}_i) - z(\mathbf{s}_j)]^2$
3. plot each distance value d_{ij} against the corresponding squared semidifference e_{ij}; in other words, plot $\mathbf{e} = vec(\mathbf{E})$ versus $\mathbf{d} = vec(\mathbf{D})$. The plot of all pairs $\{d_{ij}, e_{ij}\}$ is termed a *semivariogram cloud*

Semivariogram Cloud Example



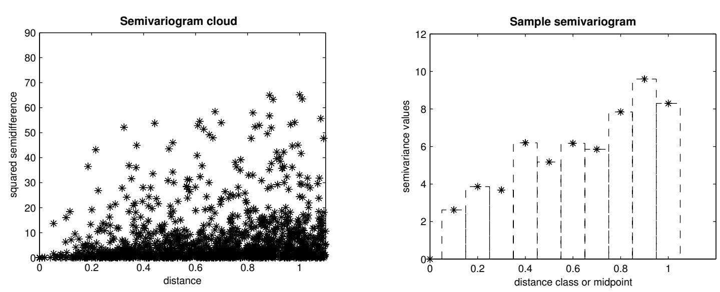
Slide 3

A measure of dissimilarity between attribute values measured at different locations, i.e., a spatial measure of attribute dissimilarity

Expected graph pattern: As the distance d_{ij} between sample pairs increases, the corresponding squared semidifference e_{ij} should also increase

Difficult to interpret, so we consider groups of sample pairs separated by similar distances i.e., average squared semidifferences within distance classes (x-axis bins in the right graph above)

Semivariogram Cloud Versus Plot



Slide 4

Going from the first to the second:

- define a set of L distance classes; the l-th class has limits: $(d_l - t_l, d_l + t_l]$, where d_l is the class midpoint and t_l is half the class width (or distance tolerance)
- for a given distance class $(d_l - t_l, d_l + t_l]$, the semivariogram value $\hat{\gamma}(d_l)$ is the *average* of $n(d_l) << n^2$ squared attribute semidifferences computed from sample pairs whose inter-distances d_{ij} satisfy: $d_l - t_l < d_{ij} \leq d_l + t_l$
- in other words, the semivariogram plot can be regarded as a summary of the semivariogram cloud, according to some distance-based grouping of samples

Computing Sample Semivariograms

Slide 5

1. compute distance matrix $\mathbf{D} = [d_{ij}, i = 1, \ldots, n, j = 1, \ldots, n]$ and squared semidifference matrix $\mathbf{E} = [e_{ij}, i = 1, \ldots, n, j = 1, \ldots, n]$ between n^2 data pairs

$$\mathbf{D} = \begin{bmatrix} 0 & d_{12} & d_{13} & d_{14} & d_{15} \\ d_{12} & 0 & d_{23} & d_{24} & d_{25} \\ d_{13} & d_{23} & 0 & d_{34} & d_{35} \\ d_{14} & d_{24} & d_{34} & 0 & d_{45} \\ d_{15} & d_{25} & d_{35} & d_{45} & 0 \end{bmatrix} \qquad \mathbf{E} = \begin{bmatrix} 0 & e_{12} & e_{13} & e_{14} & e_{15} \\ e_{12} & 0 & e_{23} & e_{24} & e_{25} \\ e_{13} & e_{23} & 0 & e_{34} & e_{35} \\ e_{14} & e_{24} & e_{34} & 0 & e_{45} \\ e_{15} & e_{25} & e_{35} & e_{45} & 0 \end{bmatrix}$$

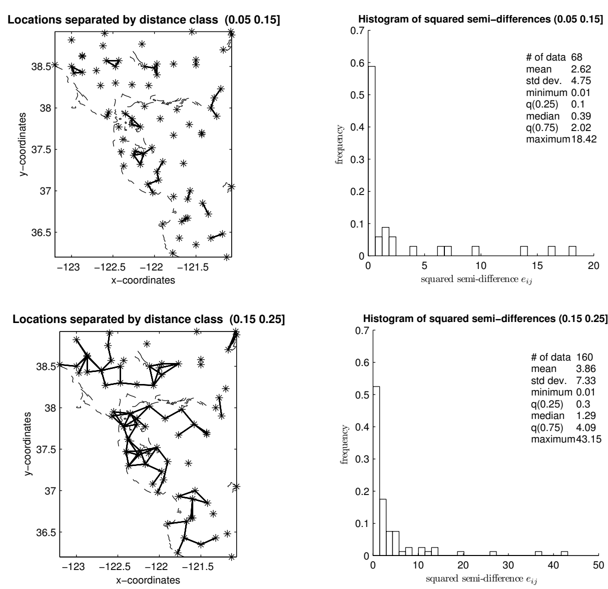
2. for a given distance class $(d_l - t_l, d_l + t_l]$, find entries of $\mathbf{E}$ that correspond to entries of $\mathbf{D}$ falling in that distance class, e.g.:

$$\begin{bmatrix} 0 & \boxed{d_{12}} & d_{13} & d_{14} & d_{15} \\ \boxed{d_{12}} & 0 & d_{23} & \boxed{d_{24}} & d_{25} \\ d_{13} & d_{23} & 0 & d_{34} & \boxed{d_{35}} \\ d_{14} & \boxed{d_{24}} & d_{34} & 0 & d_{45} \\ d_{15} & d_{25} & \boxed{d_{35}} & d_{45} & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & \boxed{e_{12}} & e_{13} & e_{14} & e_{15} \\ \boxed{e_{12}} & 0 & e_{23} & \boxed{e_{24}} & e_{25} \\ e_{13} & e_{23} & 0 & e_{34} & \boxed{d_{35}} \\ e_{14} & \boxed{e_{24}} & e_{34} & 0 & e_{45} \\ e_{15} & e_{25} & \boxed{e_{35}} & e_{45} & 0 \end{bmatrix}$$

3. sample semivariogram $\hat{\gamma}(d_l)$ for that class is the *average* of the $n(d_l)$ squared semidifferences, e-values, whose corresponding distances, d-values, fall in class $(d_l - t_l, d_l + t_l]$; i.e., the mean of all e-values in boxes in the matrix on the right above

Examples of Semivariogram Computation

Slide 6



$\hat{\gamma}((0.05\ 0.15]) = 2.62$, $\hat{\gamma}((0.15\ 0.25]) = 3.86$ = *averages of values displayed in histograms*

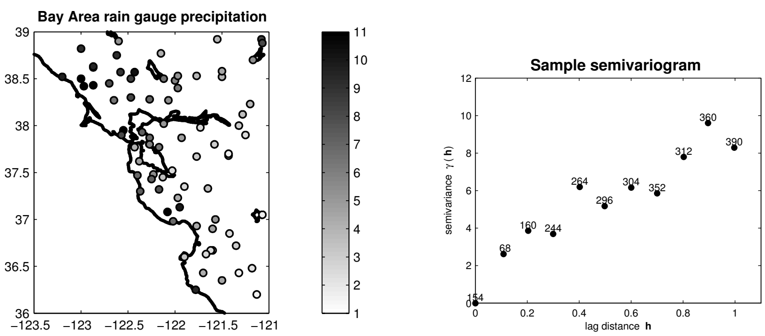
Map views linking sample pairs that contribute to such histograms are extremely informative

Sample Semivariogram Plots

Consider a set of L distance classes with midpoints $\{d_l, l = 1, \ldots, L\}$ and tolerances $\{t_l, l = 1, \ldots, L\}$. The plot of semivariance values $\{\hat{\gamma}(d_l), l = 1, \ldots, L\}$ versus the average sample inter-distance for each class is called a sample semivariogram

$$\hat{\gamma}(d_l) = \frac{1}{n(d_l)} \sum_{c=1}^{n(d_l)} e_c = \frac{1}{2n(d_l)} \sum_{d_{ij} \in (d_l - t_l, d_l + t_l]}^{n(d_l)} [z(\mathbf{s}_i) - z(\mathbf{s}_j)]^2$$

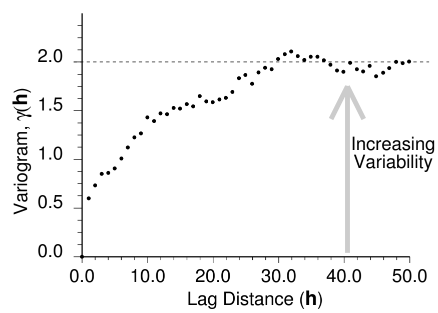
Slide 7



numbers above bullets denote # of sample pairs contributing to $\hat{\gamma}(d_l)$ at each lag distance

could also graph variances of e-values within the distance classes; $\hat{\gamma}(\mathbf{0}) = 0$, always

Semivariogram Characteristics



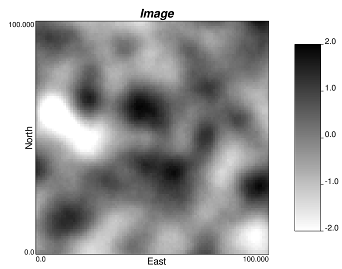
Slide 8

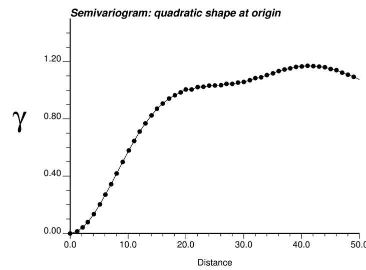
- *sill*: limit semivariogram value (plateau) is approximately equal to sample variance (for representative sample)
- *range*: distance at which semivariogram reaches (or starts oscillating around) sill = distance of influence of any datum on another
- *nugget effect*: discontinuity at origin ($\hat{\gamma}(\epsilon) > \epsilon$); sum of measurement error and micro-structures (variability at scales smaller than sampling interval)
  watch out for sparse data, outliers and positional or attribute errors
- transformation of Euclidean distance into statistical "distance" bearing imprint of specific phenomenon

Sample Semivariogram Shape & Interpretation (1)

Quadratic shape near origin:

Slide 9





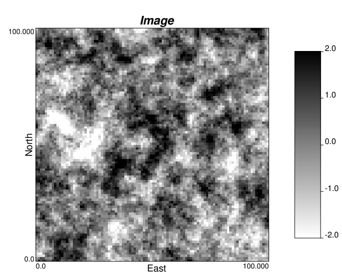
Interpretation:

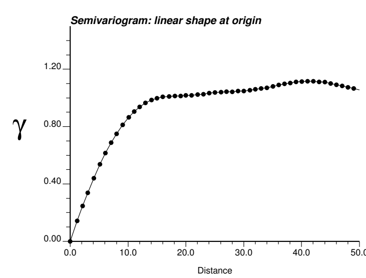
- highly continuous (extremely smooth) spatial attribute variability
- spatial attribute is differentiable
- typical variables: elevation, temperature, ...

Sample Semivariogram Shape & Interpretation (2)

Linear shape near origin:

Slide 10





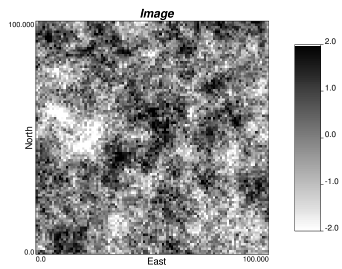
Interpretation:

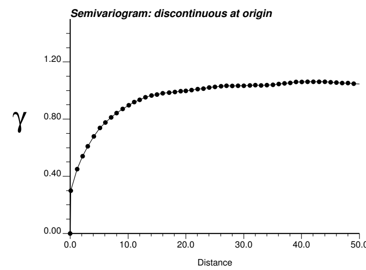
- continuous variability (not extremely smooth) of spatial attribute
- attribute is not differentiable
- typical variables: ore grades, ...

Sample Semivariogram Shape & Interpretation (3)

Discontinuous near origin:

Slide 11





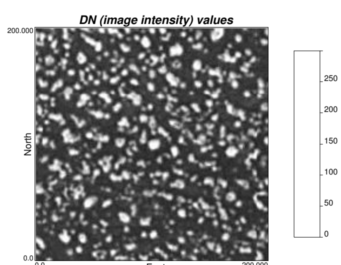
Interpretation:

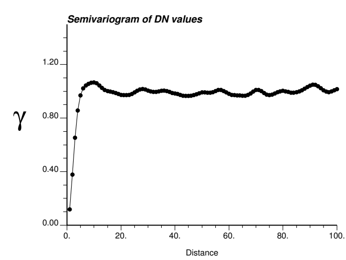
- highly irregular (quasi-random) spatial variability at small scales
- typical variables: precipitation, . . .

Sample Semivariogram Shape & Interpretation (4)

Oscillating (around sill):

Slide 12





Interpretation:

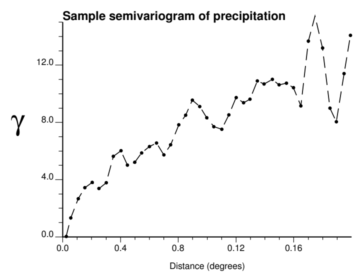
- periodic variability of spatial attribute yields sinusoidal semivariogram
- semivariogram shape possibly due to limited sampling
- need to provide physical evidence for periodicity
- frequently encountered in time series

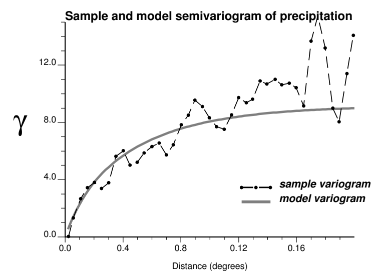
The Need for Semivariogram Models

Problems: (i) sill, range, and relative nugget, cannot be determined directly from the sample semivariogram plot, (ii) a continuum of semivariogram values $\gamma(d)$ for any distance vector d is required in interpolation, but *sample* semivariogram values $\{\hat{\gamma}(d_l), l = 1, \ldots, L\}$ are typically calculated only for few (L) distances $\{d_l, l = 1, \ldots, L\}$.

Semivariogram model definition: parametric function $\gamma(d; \boldsymbol{\theta})$ fitted to sample semivariogram values $\{\hat{\gamma}(d_l), l = 1, \ldots, L\}$; $\boldsymbol{\theta}$ denotes parameter vector with, e.g., range,

Slide 13 and sill (for a given semivariogram function)

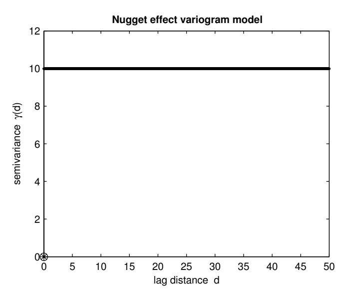




semivariogram modeling is more than a curve fitting exercise;

Warning: *cannot use any curve as semivariogram model !!!*

Valid Semivariogram Models: Pure Nugget Effect



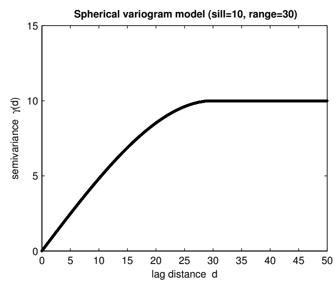
Slide 14

$$\gamma(d; \boldsymbol{\theta}) = \begin{cases} 0, & \text{if } d = 0 \\ \sigma, & \text{if } d > 0 \end{cases}$$

$\boldsymbol{\theta} = [\sigma]$, where σ denotes attribute variance

- indicates complete absence of spatial correlation
- could occur due to measurement error and microstructure, i.e., features occurring at scales smaller than sampling interval

Valid Semivariogram Models: Spherical



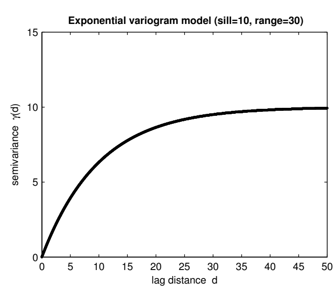
Slide 15

$$\gamma(d;\boldsymbol{\theta}) = \begin{cases} \sigma\left[\frac{3}{2}\left(\frac{d}{r}\right) - \frac{1}{2}\left(\frac{d}{r}\right)^3\right], & \text{if } d < r \\ \sigma, & \text{if } d \geq r \end{cases}$$

$\boldsymbol{\theta} = [\sigma \;\; r]$, where r is the model range

- linear behavior at origin
- clearly defined range parameter r

Valid Semivariogram Models: Exponential



Slide 16

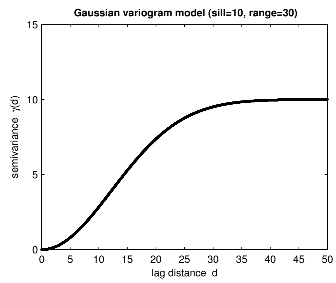
$$\gamma(d;\boldsymbol{\theta}) = \sigma\left[1 - \exp\left(-\frac{3d}{r}\right)\right]$$

$$\boldsymbol{\theta} = [\sigma \;\; r]$$

- linear behavior at origin; rises faster than spherical; reaches sill asymptotically
- effective range parameter r; distance at which 95% of sill reached

Valid Semivariogram Models: Gaussian

Slide 17



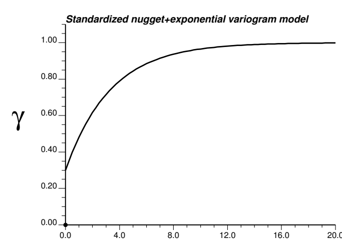
$$\boxed{\gamma(d;\boldsymbol{\theta}) = \sigma\left[1 - \exp\left(-\frac{3d^2}{r^2}\right)\right]}$$

$$\boldsymbol{\theta} = [\sigma \;\; r]$$

- quadratic behavior at origin; implies smooth spatial variability of attribute values; reaches sill asymptotically
- effective range parameter r; distance at which 95% of sill reached

Valid Semivariogram Models: Nugget + Exponential

Slide 18



$$\boxed{\gamma(d;\boldsymbol{\theta}) = \begin{cases} 0, & \text{if } d = 0 \\ a + ([\sigma - a][1 - \exp(\frac{3d}{r})]), & \text{if } d \geq \epsilon \end{cases}}$$

$$\boldsymbol{\theta} = [\sigma \;\; a \;\; r]$$

- discontinuous at origin; reaches sill asymptotically
- practical range parameter r; distance at which 95% of sill reached
- a/σ = relative nugget contribution = proportion (to total sill) of purely random spatial variability
- more complex models can be built by adding or multiplying valid models

Fitting Semivariogram Models to Sample Data

Or fitting valid semivariogram functions (curves) to sample semivariogram values

Manual fitting:

- select number of semivariograms, their type (functional form), sill, and range
- model behavior at origin (nugget effect, shape of semivariogram at distances smaller than first lag) using prior knowledge about phenomenon

Automatic fitting:

Slide 19

- least squares fit (ordinary, generalized, weighted): choose semivariogram model parameters (typically iteratively) so as to minimize discrepancy between model and sample semivariogram values over all lags; other methods also available
- treat with caution, especially with sparse data and outliers

Cross-validation:

- given a proposed parameter set, i.e., a semivariogram model, perform cross-validation using geostatistical interpolation, and record resulting error statistics
- repeat with different model parameters, and select as "optimal" model the one whose parameters yield best cross-validation error statistics

Summary

- Spatial auto-correlation can be quantified by looking at attribute dissimilarity as a function of separation distance
- The semivariogram cloud is "too cloudy" for detecting meaningful patterns
- The semivariogram plot is constructed by averaging squared semidifferences within distance bins to "smooth" out the variability in the semivariogram cloud
  NOTE: *Watch out for trends (first-order effects) in the data; a sample semivariogram quantifies second-order effects and might be contaminated by variations due to trends/drifts*

Slide 20

- A quantitative way to encapsulate a sample semivariogram is through a parametric semivariogram model
- Fitting procedures exist for estimating the parameters of semivariogram models, i.e., for fitting model semivariograms to sample semivariograms
- The final semivariogram model can be used for simulation (pattern generation) and geostatistical interpolation

NOTE: *A semivariogram model is a spatial process model, whose parameters are inferred from the sample data through the sample semivariogram*

Introduction

Data: set of n attribute measurements $\{z(\mathbf{s}_i), i = 1, \ldots, n\}$, available at n sample locations $\{\mathbf{s}_i, i = 1, \ldots, n\}$

Objectives: (i) predict or interpolate unknown attribute value $z(\mathbf{s}_p)$ at location $\mathbf{s}_p$ from the n sample data, and (ii) assess reliability of predicted value

Slide 1 **Geostatistical spatial interpolation:**

- predicted attribute value = weighted linear combination of sample data values + attribute mean, if known (non-linear methods also exist)
- a semivariogram model is used to determine the weights, which account for:
  - spatial auto-correlation between sample data and unknown value
  - spatial auto-correlation between sample data themselves (data redundancy)
- in addition, and contrary to most interpolation algorithms, geostatistics offers a measure of reliability (prediction error variance) regarding the attribute prediction

Simple Kriging (SK)

SK prediction:

$$\hat{z}(\mathbf{s}_p) = m + \sum_{i=1}^{n} w_p(\mathbf{s}_i)[z(\mathbf{s}_i) - m] = \mathbf{w}_p^T \mathbf{r}$$

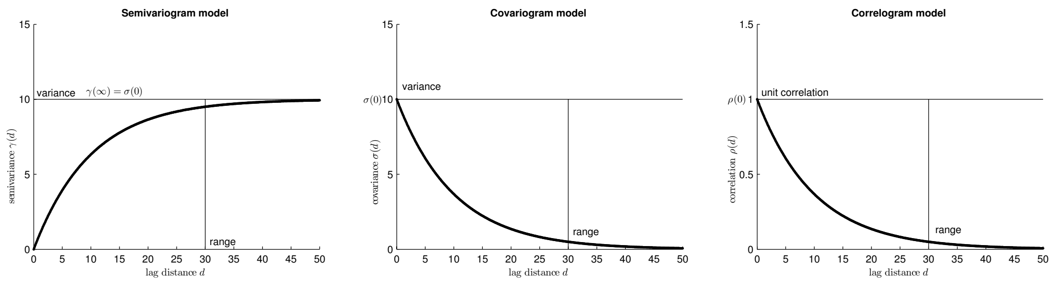
- $\mathbf{w}_p = [w_p(\mathbf{s}_i), i = 1, \ldots, n]^T$: $(n \times 1)$ vector of SK-weights assigned to n sample data for prediction at location $\mathbf{s}_p$; superscript T denotes transposition
- $\mathbf{r} = [z(\mathbf{s}_i) - m, i = 1, \ldots, n]^T$: $(n \times 1)$ vector of residual data from *known* mean m

Slide 2

$$\hat{z}(\mathbf{s}_p) = m + \underbrace{\begin{bmatrix} w_p(\mathbf{s}_1) & \cdots & w_p(\mathbf{s}_i) & \cdots & w_p(\mathbf{s}_n) \end{bmatrix}}_{\mathbf{w}_p^T} \underbrace{\begin{bmatrix} z(\mathbf{s}_1) - m \\ \vdots \\ z(\mathbf{s}_i) - m \\ \vdots \\ z(\mathbf{s}_n) - m \end{bmatrix}}_{\mathbf{r}}$$

use semivariogram model to determine weights at each prediction location; typically, it is the covariogram model that is used due to computational reasons

Semivariogram / Covariogram / Correlogram Model



Slide 3

Conversion between models, with $\sigma(0) = \gamma(\infty)$ being the sill of the semivariogram model:

- Semivariogram $\rightarrow$ covariogram: $\sigma(d) = \sigma(0) - \gamma(d)$
- Covariogram $\rightarrow$ correlogram: $\rho(d) = \frac{\sigma(d)}{\sigma(0)}$
- Semivariogram $\rightarrow$ correlogram: $\rho(d) = 1 - \frac{\gamma(d)}{\sigma(0)}$
- Covariogram $\rightarrow$ semivariogram: $\gamma(d) = \sigma(0) - \sigma(d)$

Requisites for Geostatistical Interpolation I

Data-to-data and data-to-unknown distances:

Slide 4

$$\mathbf{D} = \begin{bmatrix} 0 & \cdots & d_{1j} & \cdots & d_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{i1} & \cdots & 0 & \cdots & d_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{n1} & \cdots & d_{nj} & \cdots & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{d}_p = \begin{bmatrix} d_{1p} \\ \vdots \\ d_{ip} \\ \vdots \\ d_{np} \end{bmatrix}$$

Comments:

- as any other interpolation method, one accounts for the proximity of the n sample locations to the prediction location $\mathbf{s}_p$

  Note: Vector $\mathbf{d}_p$ changes from one prediction location $\mathbf{s}_p$ to another, hence the subscript p
- unlike other interpolation methods, one also accounts for the proximity between sample locations themselves (sample configuration or data layout)

  Note: Matrix $\mathbf{D}$ of sample-to-sample distances is the same for all prediction locations

Requisites for Geostatistical Interpolation II

From distance matrices to model covariance matrices: Take any distance value d_{ij} and d_{ip}, i.e., any entry in $\mathbf{D}$ and $\mathbf{d}_p$, and transform it, via the covariogram model, to a covariance value $\sigma(d_{ij})$ and $\sigma(d_{ip})$

Data-to-data and data-to-unknown model covariances:

Slide 5

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma(0) & \cdots & \sigma(d_{1j}) & \cdots & \sigma(d_{1n}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma(d_{i1}) & \cdots & \sigma(0) & \cdots & \sigma(d_{in}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma(d_{n1}) & \cdots & \sigma(d_{nj}) & \cdots & \sigma(0) \end{bmatrix} \quad \text{and} \quad \boldsymbol{\sigma}_p = \begin{bmatrix} \sigma(d_{1p}) \\ \vdots \\ \sigma(d_{ip}) \\ \vdots \\ \sigma(d_{np}) \end{bmatrix}$$

- data-to-data covariance matrix $\boldsymbol{\Sigma}$: $(n \times n)$ matrix with *model* covariance values $\sigma(d_{ij})$ between any two sample locations separated by distance d_{ij}
- data-to-unknown covariance vector $\boldsymbol{\sigma}_p$: $(n \times 1)$ vector with *model* covariance values $\sigma(d_{ip})$ between the n sample locations and the prediction location $\mathbf{s}_p$
  Note: Vector $\boldsymbol{\sigma}_p$ changes from one prediction location $\mathbf{s}_p$ to another, hence the subscript p

Requisites for Geostatistical Interpolation III

Data-to-data and data-to-unknown model covariances:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma(0) & \cdots & \sigma(d_{1j}) & \cdots & \sigma(d_{1n}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma(d_{i1}) & \cdots & \sigma(0) & \cdots & \sigma(d_{in}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma(d_{n1}) & \cdots & \sigma(d_{nj}) & \cdots & \sigma(0) \end{bmatrix} \quad \text{and} \quad \boldsymbol{\sigma}_p = \begin{bmatrix} \sigma(d_{1p}) \\ \vdots \\ \sigma(d_{ip}) \\ \vdots \\ \sigma(d_{np}) \end{bmatrix}$$

Slide 6 **Comments:**

- data-to-data covariance matrix $\boldsymbol{\Sigma}$: encapsulates the redundancy between the sample data; for positive spatial auto-correlation, the more clustered is the sample layout, the more redundant are the sample data (less information content);
  a clustered sample layout typically translates into larger entries in $\boldsymbol{\Sigma}$
- data-to-unknown covariance vector $\boldsymbol{\sigma}_p$: encapsulates the statistical proximity (correlation) between the sample data and the unknown attribute value $z(\mathbf{s}_p)$ at the prediction location $\mathbf{s}_p$; that correlation is a function of distance between sample and prediction locations, not of the actual (unknown) value $z(\mathbf{s}_p)$;
  The larger the entries of vector $\boldsymbol{\sigma}_p$, the stronger the predictive power of sample data

Simple Kriging (SK) System & Weights

$$\begin{bmatrix} \sigma(0) & \cdots & \sigma(d_{1n}) \\ \vdots & \ddots & \vdots \\ \sigma(d_{n1}) & \cdots & \sigma(0) \end{bmatrix} \begin{bmatrix} w_p(\mathbf{s}_1) \\ \vdots \\ w_p(\mathbf{s}_n) \end{bmatrix} = \begin{bmatrix} \sigma(d_{1p}) \\ \vdots \\ \sigma(d_{np}) \end{bmatrix}$$

$$\boxed{\boldsymbol{\Sigma}\mathbf{w}_p = \boldsymbol{\sigma}_p}$$

Comments:

Slide 7

- the SK system is a (disguised) version of the normal equations for the case of regression with no intercept term: $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$, where $\mathbf{X}$ is the design matrix and $\mathbf{y}$ is the vector of data on the dependent variable; in regression, the data-to-data covariance is estimated as $\mathbf{X}^T\mathbf{X}/n$, and the data-to-unknown covariance as $\mathbf{X}^T\mathbf{y}/n$
- the weights vector $\mathbf{w}_p$ is obtained by solving the SK system, as $\mathbf{w}_p = \boldsymbol{\Sigma}^{-1}\boldsymbol{\sigma}_p$, *anew* at each prediction location $\mathbf{s}_p$ since the entries of $\boldsymbol{\sigma}_p$ change
- entries of $\mathbf{w}_p$ *do not depend* on data values or on sill, $\sigma(0)$, of covariogram model:

$$\sigma(0)\begin{bmatrix} \rho(0) & \cdots & \rho(d_{1n}) \\ \vdots & \ddots & \vdots \\ \rho(d_{n1}) & \cdots & 1 \end{bmatrix} \begin{bmatrix} w_p(\mathbf{s}_1) \\ \vdots \\ w_p(\mathbf{s}_n) \end{bmatrix} = \sigma(0)\begin{bmatrix} \rho(d_{1p}) \\ \vdots \\ \rho(d_{np}) \end{bmatrix}$$

Interpreting the Simple Kriging Weights

$$\begin{bmatrix} w_p(\mathbf{s}_1) \\ \vdots \\ w_p(\mathbf{s}_n) \end{bmatrix} = \frac{1}{\sigma(0)}\begin{bmatrix} 1 & \cdots & \rho(d_{1n}) \\ \vdots & \ddots & \vdots \\ \rho(d_{n1}) & \cdots & 1 \end{bmatrix}^{-1} \sigma(0)\begin{bmatrix} \rho(d_{1p}) \\ \vdots \\ \rho(d_{np}) \end{bmatrix} \Rightarrow \boxed{\mathbf{w}_p = \boldsymbol{\Sigma}^{-1}\boldsymbol{\sigma}_p}$$

- if sample interdistances d_{ij} are larger than correlogram range, then $\rho(d_{ij}) = 0$, and $\boldsymbol{\Sigma} = \sigma(0)\mathbf{I}$, the $(n \times n)$ identity matrix; this entails that $w_p(s_i) = \rho(d_{ip})$, i.e., weights are equal to correlogram values

Slide 8

- but in general, $\boldsymbol{\Sigma} \neq \sigma(0)\mathbf{I}$, i.e., sample interdistances are within correlation range, in which case $\boldsymbol{\Sigma}^{-1}$ modulates $\boldsymbol{\sigma}_p$: influence of samples in clusters is downplayed
- the closer the sample data to the prediction location, and the more spread out the data over the study region, the better the SK prediction is expected to be
- for sample data far away (beyond correlation range) from the prediction location $\mathbf{s}_p$, $\rho(d_{ip}) = 0$ and $w_p(\mathbf{s}_i) = 0$: all weighs are equal to 0
- for prediction at a sample location $\mathbf{s}_p \equiv \mathbf{s}_i$, data-to-unknown covariance vector $\boldsymbol{\sigma}_p = \boldsymbol{\sigma}_i$ is same as i-th column of $\boldsymbol{\Sigma}$; this yields $w_p(\mathbf{s}_i) = 1$ if $\mathbf{s}_i = \mathbf{s}_p$, 0 otherwise: only sample co-located with prediction location receives non-zero $(= 1)$ weight

Simple Kriging Prediction and Error Variance

Once the SK weights are computed as $\mathbf{w}_p = \mathbf{\Sigma}^{-1}\boldsymbol{\sigma}_p$, they are substituted in the following equations to compute the SK prediction $\hat{z}(\mathbf{s}_p)$ and associated error variance $\hat{\sigma}(\mathbf{s}_p)$

SK prediction does not depend on sill $\sigma(0)$ of covariogram model:

Slide 9

$$\hat{z}(\mathbf{s}_p) = m+\mathbf{w}_p^T\mathbf{r} = m+[w_p(\mathbf{s}_1) \ \cdots \ w_p(\mathbf{s}_n)] \begin{bmatrix} z(\mathbf{s}_1) - m \\ \vdots \\ z(\mathbf{s}_n) - m \end{bmatrix} = m+\sum_{i=1}^{n} w_p(\mathbf{s}_i)[z(\mathbf{s}_i)-m]$$

SK prediction error variance does depend on covariogram model sill $\sigma(0)$:

$$\hat{\sigma}(\mathbf{s}_p) = \sigma(0)-\mathbf{w}_p^T\boldsymbol{\sigma}_p = \sigma(0)-[w_p(\mathbf{s}_1) \ \cdots \ w_p(\mathbf{s}_n)] \begin{bmatrix} \sigma(d_{1p}) \\ \vdots \\ \sigma(d_{np}) \end{bmatrix} = \sigma(0)-\sum_{n=1}^{n} w_p(\mathbf{s}_i)\sigma(d_{ip})$$

which can also be written as: $\hat{\sigma}(\mathbf{s}_p) = \sigma(0)\left[1 - \sum_{i=1}^{n} w_p(\mathbf{s}_i)\rho(d_{ip})\right]$

Interpreting the SK Prediction and Error Variance

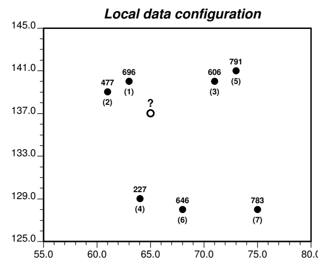
$$\hat{z}(\mathbf{s}_p) = m + \sum_{i=1}^{n} w_p(\mathbf{s}_i)[z(\mathbf{s}_i) - m] \qquad \hat{\sigma}(\mathbf{s}_p) = \sigma(0) - \sum_{i=1}^{n} w_p(\mathbf{s}_i)\sigma(d_{ip})$$

Comments:

Slide 10

- for sample data far away (beyond correlation range) from the prediction location $\mathbf{s}_p$, $w_p(\mathbf{s}_i) = 0, \forall i$: *all weighs are equal to* 0. In this case, the SK prediction equals the known mean m and the SK error variance equals the known covariogram sill: $\hat{z}(\mathbf{s}_p) = m$ and $\hat{\sigma}(\mathbf{s}_p) = \sigma(0)$; *away from the sample data, SK yields back the (assumed known) attribute overall mean and variance*
- for prediction at a sample location $\mathbf{s}_p \equiv \mathbf{s}_i$, $w_p(\mathbf{s}_i) = 1$ if $\mathbf{s}_i = \mathbf{s}_p$, 0 otherwise: the SK prediction identifies the known sample datum and the SK error variance is zero: $\hat{z}(\mathbf{s}_i) = z(\mathbf{s}_i)$ and $\hat{\sigma}(\mathbf{s}_i) = 0$; *SK is an exact interpolation algorithm*
- for all other prediction locations, the SK predictions depend on the sample data configuration and their values, while the SK error variances depend only on the sample data configuration; *both SK predictions and error variances depend on the covariogram model* $\sigma(d)$ *adopted*

Determining the SK Weights: Step 1



(n x n) matrix of data-to-data inter-distances:

Slide 11

$$\mathbf{D} = \begin{bmatrix} 0.00 & 2.24 & 8.00 & 11.05 & 10.05 & 13.00 & 16.97 \\ 2.24 & 0.00 & 10.05 & 10.44 & 12.17 & 13.04 & 17.80 \\ 8.00 & 10.05 & 0.00 & 13.04 & 2.24 & 12.37 & 12.65 \\ 11.05 & 10.44 & 13.04 & 0.00 & 15.00 & 4.12 & 11.05 \\ 10.05 & 12.17 & 2.24 & 15.00 & 0.00 & 13.93 & 13.15 \\ 13.00 & 13.04 & 12.37 & 4.12 & 13.93 & 0.00 & 7.00 \\ 16.97 & 17.80 & 12.65 & 11.05 & 13.15 & 7.00 & 0.00 \end{bmatrix}$$

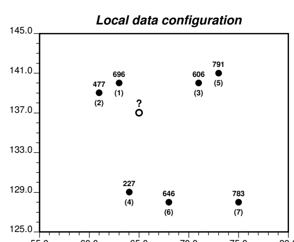
i, j-th element of $\mathbf{D}$: $d_{ij} = ||\mathbf{s}_i - \mathbf{s}_j||$

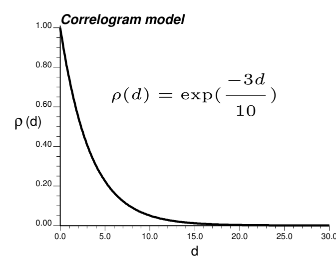
(n x 1) vector of prediction-to-data-location distances:

$$\mathbf{d}_p = \begin{bmatrix} 3.61 & 4.47 & 6.71 & 8.06 & 8.94 & 9.49 & 13.45 \end{bmatrix}^T$$

i-th element of $\mathbf{d}_p$: $d_{ip} = ||\mathbf{s}_i - \mathbf{s}_p||$

Determining the SK Weights: Step 2



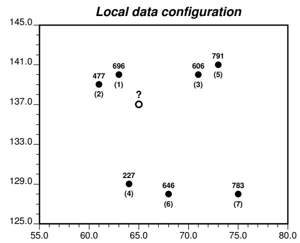


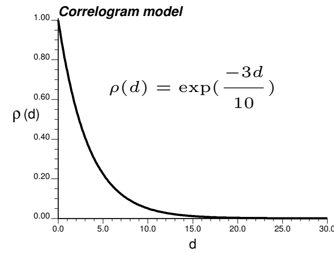
Slide 12

$$\underbrace{\begin{bmatrix} 3.61 \\ 4.47 \\ 6.71 \\ 8.06 \\ 8.94 \\ 9.49 \\ 13.45 \end{bmatrix}}_{\mathbf{d}_p} \rightarrow 1 \underbrace{\begin{bmatrix} \exp(-3 \times 3.61/10) \\ \exp(-3 \times 4.47/10) \\ \exp(-3 \times 6.71/10) \\ \exp(-3 \times 8.06/10) \\ \exp(-3 \times 8.94/10) \\ \exp(-3 \times 9.49/10) \\ \exp(-3 \times 13.45/10) \end{bmatrix}}_{\boldsymbol{\sigma}_p = \text{sill} \exp(-3\mathbf{d}_p/\text{range})} = \begin{bmatrix} 0.34 \\ 0.26 \\ 0.13 \\ 0.09 \\ 0.07 \\ 0.06 \\ 0.02 \end{bmatrix}$$

These would be the weights if one ignored auto-correlation between sample data

Determining the SK Weights: Step 3





SK system:

Slide 13

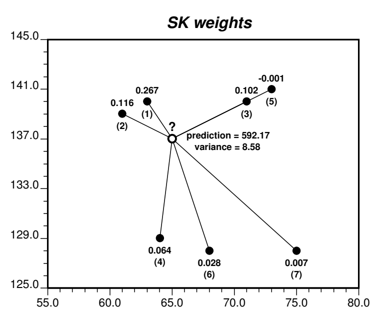
$$\underbrace{\begin{bmatrix} 1.00 & 0.51 & 0.09 & 0.04 & 0.05 & 0.02 & 0.01 \\ 0.51 & 1.00 & 0.05 & 0.04 & 0.03 & 0.02 & 0.00 \\ 0.09 & 0.05 & 1.00 & 0.02 & 0.51 & 0.02 & 0.02 \\ 0.04 & 0.04 & 0.02 & 1.00 & 0.01 & 0.29 & 0.04 \\ 0.05 & 0.03 & 0.51 & 0.01 & 1.00 & 0.02 & 0.02 \\ 0.02 & 0.02 & 0.02 & 0.29 & 0.02 & 1.00 & 0.12 \\ 0.01 & 0.00 & 0.02 & 0.04 & 0.02 & 0.12 & 1.00 \end{bmatrix}}_{\boldsymbol{\Sigma}} \underbrace{\begin{bmatrix} w_p(\mathbf{s}_1) \\ w_p(\mathbf{s}_2) \\ w_p(\mathbf{s}_3) \\ w_p(\mathbf{s}_4) \\ w_p(\mathbf{s}_5) \\ w_p(\mathbf{s}_6) \\ w_p(\mathbf{s}_7) \end{bmatrix}}_{\mathbf{w}_p} = \underbrace{\begin{bmatrix} 0.34 \\ 0.26 \\ 0.13 \\ 0.09 \\ 0.07 \\ 0.06 \\ 0.02 \end{bmatrix}}_{\boldsymbol{\sigma}_p}$$

i, j-th element of matrix $\boldsymbol{\Sigma}$: $\sigma_{ij} = 1 \times \exp(-3 \times d_{ij}/10)$

Determining the SK Weights: Step 4

$$\underbrace{\begin{bmatrix} w_p(\mathbf{s}_1) \\ w_p(\mathbf{s}_2) \\ w_p(\mathbf{s}_3) \\ w_p(\mathbf{s}_4) \\ w_p(\mathbf{s}_5) \\ w_p(\mathbf{s}_6) \\ w_p(\mathbf{s}_7) \end{bmatrix}}_{\mathbf{w}_p} = \underbrace{\begin{bmatrix} 1.36 & -0.69 & -0.09 & -0.02 & 0.00 & 0.00 & -0.01 \\ -0.69 & 1.35 & 0.00 & -0.02 & 0.00 & -0.01 & 0.01 \\ -0.09 & 0.00 & 1.36 & -0.01 & -0.69 & -0.01 & -0.01 \\ -0.02 & -0.02 & -0.01 & 1.09 & 0.00 & -0.32 & -0.01 \\ 0.00 & 0.00 & -0.69 & 0.00 & 1.35 & -0.01 & -0.01 \\ 0.00 & -0.01 & -0.01 & -0.32 & -0.01 & 1.11 & -0.12 \\ -0.01 & 0.01 & -0.01 & -0.01 & -0.01 & -0.12 & 1.02 \end{bmatrix}}_{\boldsymbol{\Sigma}^{-1}} \underbrace{\begin{bmatrix} 0.34 \\ 0.26 \\ 0.13 \\ 0.09 \\ 0.07 \\ 0.06 \\ 0.02 \end{bmatrix}}_{\boldsymbol{\sigma}_p}$$

Slide 14



original weights vector ($\mathbf{w}_p = \boldsymbol{\sigma}_p$) modified by $\boldsymbol{\Sigma}^{-1}$ to account for sample redundancy;
e.g., $w_p(\mathbf{s}_1) = 0.27$ instead of $\rho(d_{1p}) = 0.34$