GEOG574 Introduction to Geostatistics

Kriging II

Daoqin Tong

School of Geography & Development
408 Harvill Building
Email: daoqin@email.arizona.edu

Simple Kriging

- $\mu(s)$ is known
- Given sample data $y(s_1),\ldots,y(s_n)$, the observed values of residuals can be calculated by $u(s_i) = y(s_i) - \mu(s)$
- Assume stationarity for these residuals: zero mean, known variance, and known covariance function C(h)
- Find an estimate of U(s) at location *s* given observed $u(s_i)$ of the random variable $U(s_i)$ using

$$\hat{u}(s) = \sum_{i=1}^{n} \lambda_i(s) u(s_i)$$

Simple kriging

- Determines the optimal weights as those minimizing the mean square prediction error (MSPE)

$$E\left(\hat{U}(s)-U(s)\right)^2=\lambda^T(s)C\lambda(s)+\sigma^2-2\lambda^T(s)c(s)$$

- Differentiating the MSE with respect to $\lambda(s)$

$$\lambda(s)=C^{-1}c(s)$$

- The simple Kriging predictor is $\hat{U}(s)=\lambda^T(s)U=c^T(s)C^{-1}U$
- Corresponding MSPE is $E\left(\hat{U}(s)-U(s)\right)^2=\sigma^2-c^T(s)C^{-1}c(s)$

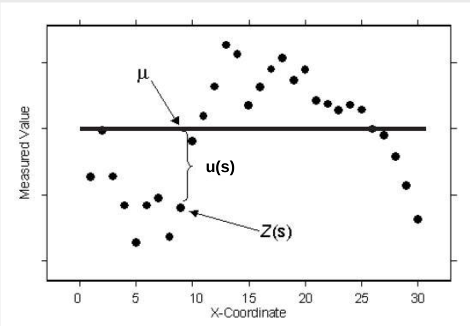
which is referred to as *mean square prediction error* (*MSPE*), or *kriging variance*, denoted by σ_e^2

Simple Kriging

- Unbiased
- The optimal method of spatial prediction (under squared error loss) in a Gaussian random field
- Useful in that it determines the benchmark for other kriging methods

Ordinary Kriging

- Unknown constant mean $Y(s) = \mu(s) + U(s)$
- Predict Y(s) using a weighted linear combination of observed data
- Avoid estimate first order effect directly
- Weights are determined optimally
  - Unbiaseness
  - Minimum MSE (mean square error)



unknown

$$z(s) = \mu + U(s)$$

There is no way to decide , based upon the data alone, whether the observed pattern is the result of autocorrelation alone or a trend ($\mu(s)$changing with s)

Ordinary Kriging

- Mean of the random field is not known
- Ordinary kriging predictor minimizes the mean square prediction error subject to an unbiasedness constraint

- Unbiasedness

$$E\left(\hat{Y}(s)\right) = E\left(\sum_{i=1}^{n}\omega_i(s)Y(s_i)\right) = \mu$$

$$E\left(Y(s_i)\right) = \mu \qquad \Longrightarrow \qquad \sum_{i=1}^{n}\omega_i(s) = 1$$

- Minimum MSE

$$\mathrm{MSE} = E\left[\left(\hat{Y}(s) - Y(s)\right)^2\right] = \omega^T(s)C\omega(s) + \sigma^2 - 2\omega^T(s)c(s)$$

$$\min_{\omega(s)}\left\{\omega^T(s)C\omega(s) + \sigma^2 - 2\omega^T(s)c(s)\right\} \quad \text{subject to } \omega^T(s)\times\vec{1} = 1$$

Minimize MSE

- To minimize MSE, subject to constraints

$$\min_{\omega(s)}\left\{\omega^T(s)C\omega(s)+\sigma^2-2\omega^T(s)c(s)\right\} \quad \text{subject to } \omega^T(s)\times\vec{1}=1$$

- Introduce Lagrange multiplier *v*(*s*)

$$\min_{\omega(s)}\left\{\omega^T(s)C\omega(s)+\sigma^2-2\omega^T(s)c(s)+2\left(\omega^T(s)\times\vec{1}-1\right)v(s)\right\}$$

- Differentiating with respect to *v*(*s*) and *w*(*s*)

$$\begin{cases}\omega^T(s)\times\vec{1}=1\\ C\omega(s)+\vec{1}v(s)=c(s)\end{cases}$$

Solving the Equations

- The Equations

$$\begin{cases}\omega^T(s)\times\vec{1}=1\\ C\omega(s)+\vec{1}v(s)=c(s)\end{cases}$$

- Using augmented matrix and vectors

$$\overset{C_+}{\Downarrow}\qquad\qquad \overset{\omega_+(s)}{\Downarrow} \;=\; \overset{c_+(s)}{\Downarrow}$$

$$\begin{pmatrix} C(s_1,s_1) & \cdots & C(s_1,s_n) & 1 \\ & \ddots & & \\ C(s_n,s_1) & \cdots & C(s_n,s_n) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}\begin{pmatrix}\omega_1(s)\\ \vdots \\ \omega_n(s) \\ v(s)\end{pmatrix}=\begin{pmatrix}c(s,s_1)\\ \vdots \\ c(s,s_n) \\ 1\end{pmatrix} \Rightarrow \boxed{\begin{aligned}&\omega_+(s)=C_+^{-1}c_+(s)\\ &\sigma_{OK}^2=\sigma^2-c_+^T(s)C_+^{-1}c_+(s)\end{aligned}}$$

Ordinary Kriging

- Based on covariogram C(h), calculate

$$C_+ = \begin{pmatrix} C(s_1,s_1) & \cdots & C(s_1,s_n) & 1 \\ & \ddots & & \\ C(s_n,s_1) & \cdots & C(s_n,s_n) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}, \quad c_+(s) = \begin{pmatrix} c(s,s_1) \\ \vdots \\ c(s,s_n) \\ 1 \end{pmatrix}$$

- Calculate the weights

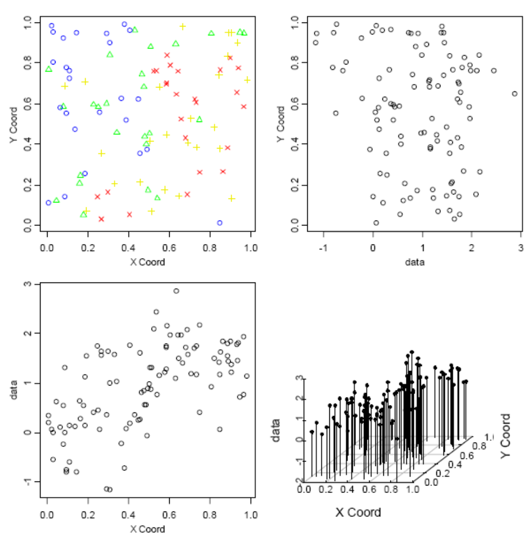
$$\begin{pmatrix} \omega_1(s) \\ : \\ \omega_n(s) \\ v(s) \end{pmatrix} = \omega_+(s) = C_+^{-1} c_+(s)$$

- Predict the value of y at location s using the first *n* elements in the vector

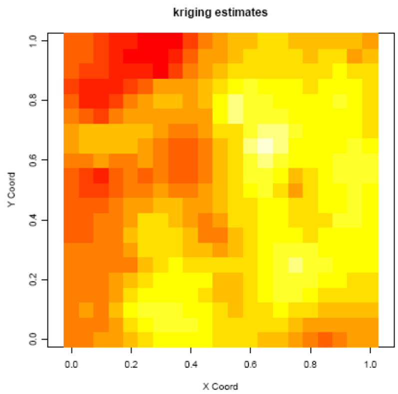
$$\hat{Y}(s) = \sum_{i=1}^{n} \omega_i(s) Y(s_i)$$

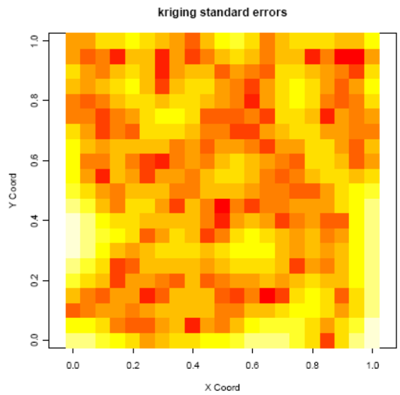
- Calculate the kriging variance $\sigma_{OK}^2 = \sigma^2 - c_+^T(s) C_+^{-1} c_+(s)$

Example



Example: Kriging Estimates and Variances

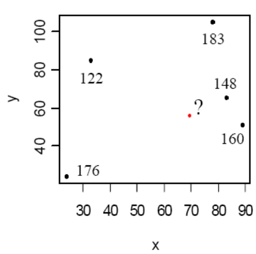




Previous Example

- Data
  - Known covariogram $C(h) = 20e^{-3h/100}$

ID	X	Y	Z
1	33	85	122
2	78	105	183
3	83	65	148
4	89	51	160
5	24	24	176



Previous Example

- Distance matrix

Distances	1	2	3	4	5
1	0	49.2	53.9	65.2	60.795
2		0	40.3	55.9	97.1
3			0	15.8	72.09
4				0	69.6
5					0

- Calculate the covariances

C

Covariances	1	2	3	4	5
1	20.000	4.571	3.970	2.828	3.228
2	4.571	20.000	5.970	3.739	1.086
3	3.970	5.970	20.000	12.450	2.300
4	2.828	3.739	12.450	20.000	2.479
5	3.228	1.086	2.300	2.479	20.000

$c(s)$

6.895
5.032
10.15
8.083
4.079

Previous Example

- Calculate the augmented matrix C_+ and vector $c_+(s)$

$$C_+ = \begin{bmatrix} 20.000 & 4.5710 & 3.9700 & 2.2800 & 3.2280 & 1 \\ 4.5710 & 20.000 & 5.9700 & 3.7390 & 1.0860 & 1 \\ 3.9700 & 5.9700 & 20.000 & 12.4500 & 2.3000 & 1 \\ 2.8280 & 3.7390 & 12.4500 & 20.000 & 2.4790 & 1 \\ 3.2280 & 1.0860 & 2.3000 & 2.4790 & 20.000 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$c_+(s) = \begin{bmatrix} 6.895 \\ 5.032 \\ 10.15 \\ 8.083 \\ 4.079 \\ 1.000 \end{bmatrix}$$

Previous Example

- Calculate the inverse of C_+

$$C_+^{-1} = \begin{bmatrix} 0.0483 & -0.0170 & -0.0107 & -0.0046 & -0.0160 & 0.2183 \\ -0.0168 & 0.0495 & -0.0177 & -0.0066 & -0.0084 & 0.2260 \\ -0.0093 & -0.0181 & 0.0864 & -0.0534 & -0.0056 & 0.0949 \\ -0.0066 & -0.0059 & -0.0528 & 0.0767 & -0.0115 & 0.1883 \\ -0.0156 & -0.0085 & -0.0052 & -0.0122 & 0.0415 & 0.2725 \\ 0.2125 & 0.2271 & 0.0897 & 0.1979 & 0.2728 & -7.0849 \end{bmatrix}$$

- Calculate the weights

$$\omega_+(s) = C_+^{-1} c_+(s) = \begin{bmatrix} 0.0483 & -0.0170 & -0.0107 & -0.0046 & -0.0160 & 0.2183 \\ -0.0168 & 0.0495 & -0.0177 & -0.0066 & -0.0084 & 0.2260 \\ -0.0093 & -0.0181 & 0.0864 & -0.0534 & -0.0056 & 0.0949 \\ -0.0066 & -0.0059 & -0.0528 & 0.0767 & -0.0115 & 0.1883 \\ -0.0156 & -0.0085 & -0.0052 & -0.0122 & 0.0415 & 0.2725 \\ 0.2125 & 0.2271 & 0.0897 & 0.1979 & 0.2728 & -7.0849 \end{bmatrix} \times \begin{bmatrix} 6.895 \\ 5.032 \\ 10.15 \\ 8.083 \\ 4.079 \\ 1.000 \end{bmatrix} = \begin{bmatrix} \boxed{\begin{matrix} 0.2547 \\ 0.0922 \\ 0.3621 \\ 0.1508 \\ 0.1402 \end{matrix}} \\ -0.8540 \end{bmatrix}$$

Previous Example

- The predicted value $y(s)$

$$\hat{y}(s) = \omega^T(s) y = \sum_{i=1}^{n} \omega_i(s) y(s_i)$$
$$= 0.251 \times 122 + 0.094 \times 183 + 0.363 \times 148 + 0.151 \times 160 + 0.141 \times 176$$
$$\simeq 150.5$$

- The kriging variance

$$\sigma_{OK}^2 = \sigma^2 - c_+^T(s) C_+^{-1} c_+(s) = 20.0 - 6.81 = 13.19$$

- The 95% confidence interval

$$\hat{y}(s) \pm 1.96 \sigma_{OK} = 150.5 \pm 7.12$$

Ordinary kriging in terms of variogram $\gamma(h)$

- In practice, kriging is usually implemented using variogram
- It has better statistical properties (unbiased and consistent)The predicted value $y(s)$

$$MSE = \sum_i \sum_j w_i w_j (\sigma^2 - \gamma_{ij}) + \sigma^2 - 2\sum w_i (\sigma^2 - \gamma_{i0}) + 2\lambda(\sum w_i - 1)$$
$$= -\sum_i \sum_j w_i w_j \gamma_{ij} + 2\sum w_i \gamma_{i0} + 2\lambda(\sum w_i - 1).$$

- Similarly, weights can be determined

$$\underbrace{\begin{bmatrix} 0 & \gamma_{12} & \cdots & \gamma_{1n} & 1 \\ \gamma_{21} & 0 & \cdots & \gamma_{2n} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}}_{\substack{\boldsymbol{\Gamma} \\ (n+1)\times(n+1)}} \bullet \underbrace{\begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_n \\ -\lambda \end{bmatrix}}_{\substack{\boldsymbol{w} \\ (n+1)\times 1}} = \underbrace{\begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \cdots \\ \gamma_{n0} \\ 1 \end{bmatrix}}_{\substack{\boldsymbol{D} \\ (n+1)\times 1}} \Longrightarrow \boldsymbol{w} = \boldsymbol{\Gamma}^{-1}\boldsymbol{D}$$

Universal Kriging

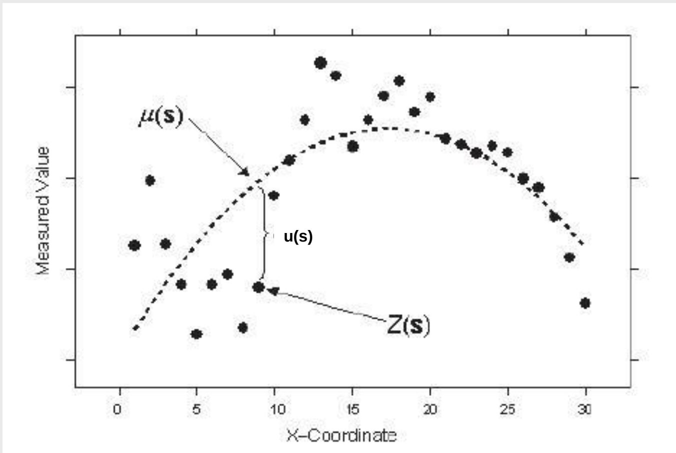
- Unknown non-constant mean in $Y(s) = \mu(s) + U(s)$

$$\mu(s) = x(s)^T \beta, \quad x(s) = \left(x_1(s), x_2(s), \ldots, x_p(s)\right)^T$$

- Predict Y(s) using a weighted linear combination of the observed data $Y(s_1),\ldots,Y(s_2)$
- Avoid estimate first order effect directly
- Weights are determined optimally
  – Unbiaseness
  – Minimum MSE

Deterministic function

$$Z(s) \;= \mu(s) + U(s)$$



Universal Kriging

- Unbiasness

$$E\left(\hat{Y}(s)\right) = E\left(\sum_{i=1}^{n} \omega_i(s) Y(s_i)\right) = \mu(s) = x(s)^T \beta$$

$$E\left(Y(s_i)\right) = \mu(s_i) = x(s_i)^T \beta$$

$$\Rightarrow \sum_{i=1}^{n} \omega_i(s) x(s_i) = x(s)$$

- Minimum MSE

$$\text{MSE} = E\left[\left(\hat{Y}(s) - Y(s)\right)^2\right] = \omega^T(s) C \omega(s) + \sigma^2 - 2\omega^T(s) c(s)$$

$$\min_{\omega(s)} \left\{\omega^T(s) C \omega(s) + \sigma^2 - 2\omega^T(s) c(s)\right\} \quad \text{subject to} \sum_{i=1}^{n} \omega_i(s) x(s_i) = x(s)$$

Universal Kriging

- The solution

$$\omega_+(s) = C_+^{-1} c_+(s)$$

$$\sigma_{UK}^2 = \sigma^2 - c_+^T(s) C_+^{-1} c_+(s)$$

- The augmented matrix and vectors

$$\omega_+(s) = \begin{pmatrix} \omega_1(s) \\ \vdots \\ \omega_n(s) \\ v_1(s) \\ \vdots \\ v_1(s) \end{pmatrix}; \quad C_+ = \begin{pmatrix} C(s_1,s_1) & \cdots & C(s_1,s_n) & x_1(s_1) & \cdots & x_p(s_1) \\ & \ddots & & & \ddots & \\ C(s_n,s_1) & \cdots & C(s_n,s_n) & x_1(s_n) & \cdots & x_p(s_n) \\ x_1(s_1) & \cdots & x_1(s_n) & 0 & \cdots & 0 \\ & \ddots & & & \ddots & \\ x_p(s_1) & \cdots & x_p(s_n) & 0 & \cdots & 0 \end{pmatrix}; \quad c_+(s) = \begin{pmatrix} c(s,s_1) \\ \vdots \\ c(s,s_n) \\ x_1(s) \\ \vdots \\ x_p(s) \end{pmatrix}$$

Calculation Steps

- Based on covariogram *C*(*h*)

$$C_+ = \begin{pmatrix} C(s_1,s_1) & \cdots & C(s_1,s_n) & x_1(s_1) & \cdots & x_p(s_1) \\ & \ddots & & & \ddots & \\ C(s_n,s_1) & \cdots & C(s_n,s_n) & x_1(s_n) & \cdots & x_p(s_n) \\ x_1(s_1) & \cdots & x_1(s_n) & 0 & \cdots & 0 \\ & \ddots & & & \ddots & \\ x_p(s_1) & \cdots & x_p(s_n) & 0 & \cdots & 0 \end{pmatrix}; \quad c_+(s) = \begin{pmatrix} c(s,s_1) \\ \vdots \\ c(s,s_n) \\ x_1(s) \\ \vdots \\ x_p(s) \end{pmatrix}$$

- Calculate the weights

$$\begin{pmatrix} \omega_1(s) \\ : \\ \omega_n(s) \\ v_1(s) \\ : \\ v_p(s) \end{pmatrix} = \omega_+(s) = C_+^{-1} c_+(s)$$

- Predict the value of y at location s using the first n elements in the vector $w_+(s)$ $\hat{Y}(s) = \sum_{i=1}^{n} \omega_i(s) Y(s_i)$

- Calculate the kriging variance $\sigma_{UK}^2 = \sigma^2 - c_+^T(s) C_+^{-1} c_+(s)$

Kriging Variance

- Variance-covariance matrix C is usually unknown
- An estimate C is substituted
- However, the uncertainty associated with the estimation of the semi-variances or covariances is typically not accounted for in the determination of the mean square prediction error
- Therefore, the kriging variance obtained is an underestimate of the mean square prediction error

Evaluation of Kriging Models

- Kriging is an exact interpolator so no residuals for each sample points
- Additional observations can be obtained to validate the kriging models
- Another common approach is cross-validation
  - Set aside some portion of the sample data as test data
  - Compare the observed values of test data with their predicted values which are estimated from the rest of the sample data

Cross Validataion

- Cross validation can be used to
  - Evaluate the fitness of Kriging model
  - Determine the suitable parametric model for variogram
- Procedure
  - For $i = 1 : n$
    - Remove y_i from the sample data $\{y_1, \ldots, y_n\} \rightarrow y_{-i}$
    - Predict y_i using the remaining points y_{-i} with the model $\rightarrow \hat{y}_{-i}$
    - Compare $\hat{y}_{-i}$ with y_i
  - End
- Statistics $\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_{-i} - y_i)^2$