

Lecture 11: Kriging Extensions

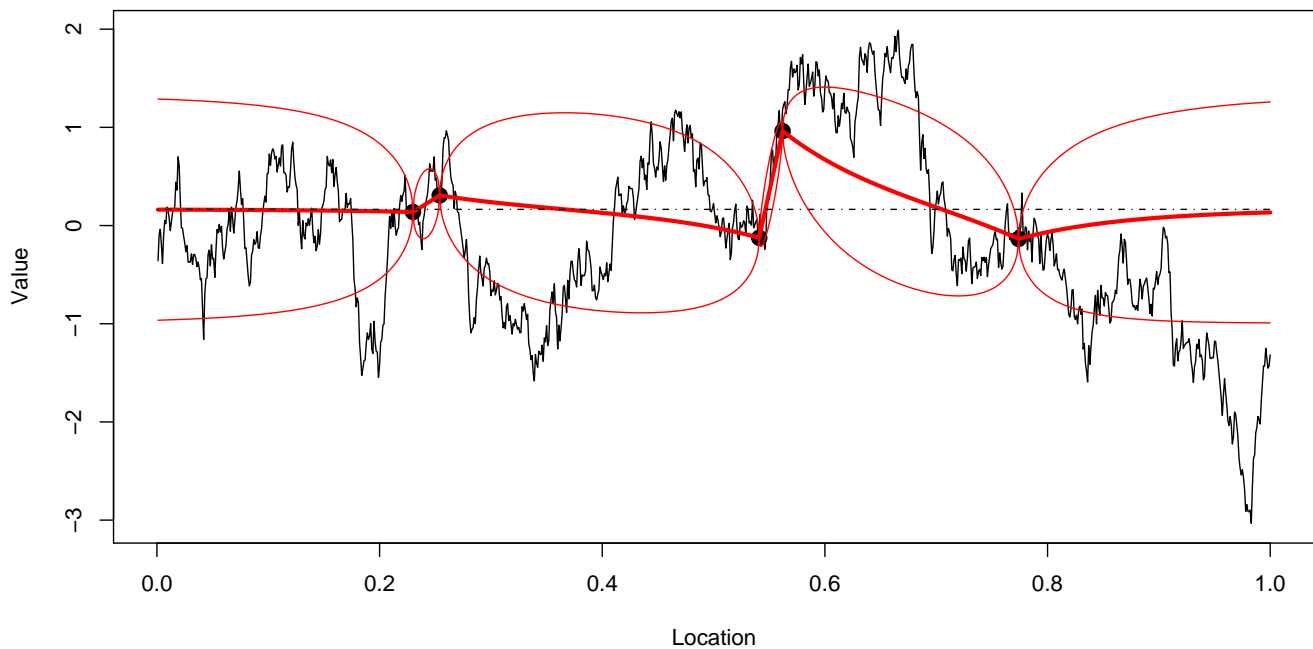
Math 586

April 5, 2009

Example for Ordinary Kriging.

Consider 5 points sampled from a 1-d random field with known covariance function $C(h) = \exp(-|h|/\ell)$, with scale parameter $\ell = 0.1$

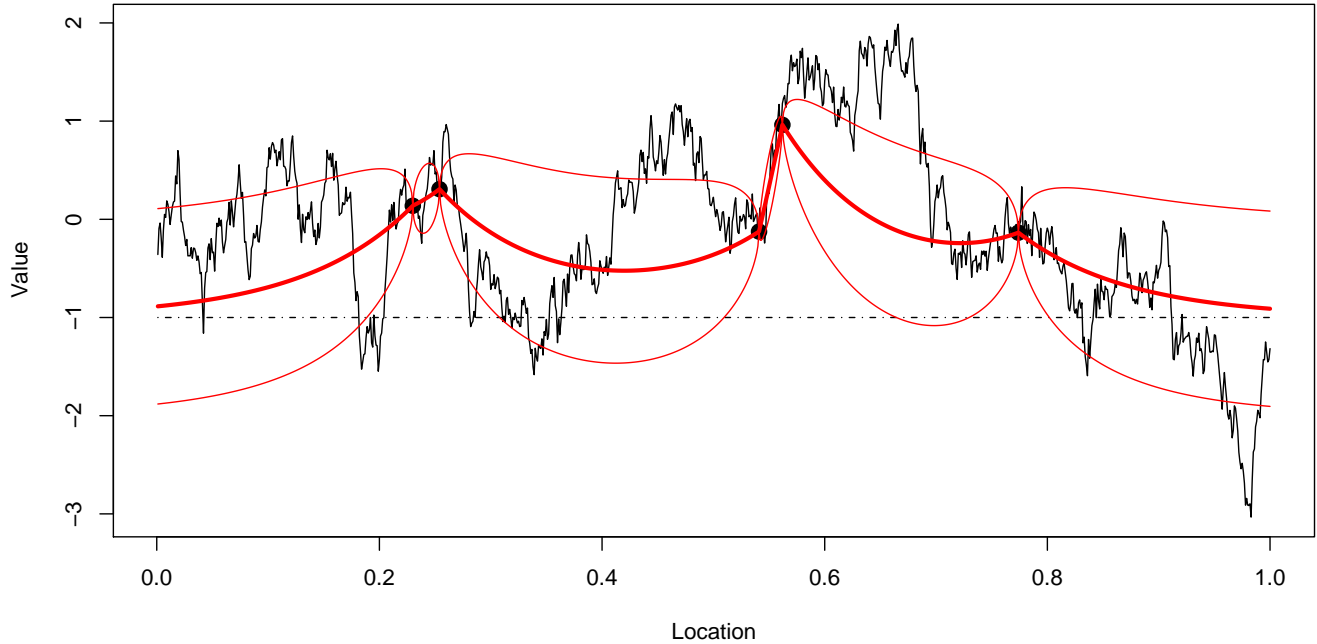
Ordinary kriging, cov. length = 0.1



Plot shows kriging mean \pm one st.dev.; broken line = estimate of the mean.

However, if we knew the value of the mean, we could use simple kriging. This will decrease kriging variance σ_K^2 , but what happens when the mean is incorrect? See the next Figure.

Simple kriging, mean = -1, cov. length = 0.1



Other relevant data:	locations	0.230	0.254	0.541	0.562	0.774
	observations	0.138	0.307	-0.125	0.963	-0.136

C-matrix:

1.0000	0.7857	0.0439	0.0356	0.0042
0.7857	1.0000	0.0559	0.0453	0.0054
0.0439	0.0559	1.0000	0.8097	0.0962
0.0356	0.0453	0.8097	1.0000	0.1188
0.0042	0.0054	0.0962	0.1188	1.0000

Estimate of the mean: $\hat{m} = 0.1648$ with weights $\lambda_j = 0.1926, 0.1744, 0.1719, 0.1536, 0.3075$.
 Note the higher weight for the observation #5 that stands alone.

I. Extension to IRF-0

Covariance may not exist.

To the definition of IRF-0 add one more requirement:

- (i) $\mathbb{E} V(\mathbf{x}) = \text{const}$
- (ii) $\gamma(\mathbf{h}) = \frac{1}{2} \mathbb{E} [V(\mathbf{x} + \mathbf{h}) - V(\mathbf{x})]^2$ is independent of \mathbf{x}
- (iii) $\text{Var} [\sum_{j=0}^J \alpha_j V(\mathbf{x}_j)]$ is finite

for every J , set of \mathbf{x}_j and every set of α_j such that $\sum \alpha_j = 0$.

For example, if $\sum_{j=1}^n \lambda_j = 1$ then we require

$$\text{Var}[V(\mathbf{x}_0) - \sum_{j=1}^n \lambda_j V(\mathbf{x}_j)] \text{ is finite.}$$

Consider

$$\text{Cov}[V(\mathbf{x}_i) - V(\mathbf{x}_0), V(\mathbf{x}_j) - V(\mathbf{x}_0)] = \mathbb{E} \{ [V(\mathbf{x}_i) - V(\mathbf{x}_0)][V(\mathbf{x}_j) - V(\mathbf{x}_0)] \}.$$

First,

$$\begin{aligned} \gamma(\mathbf{x}_i - \mathbf{x}_j) &= \frac{1}{2} \mathbb{E} [V(\mathbf{x}_i) - V(\mathbf{x}_j)]^2 = \frac{1}{2} \mathbb{E} [\{V(\mathbf{x}_i) - V(\mathbf{x}_0)\} - \{V(\mathbf{x}_j) - V(\mathbf{x}_0)\}]^2 = \\ &= \frac{1}{2} \mathbb{E} [\{V(\mathbf{x}_i) - V(\mathbf{x}_0)\}^2 + \{V(\mathbf{x}_j) - V(\mathbf{x}_0)\}^2 - 2\{V(\mathbf{x}_i) - V(\mathbf{x}_0)\}\{V(\mathbf{x}_j) - V(\mathbf{x}_0)\}]^2 = \\ &= \gamma(\mathbf{x}_i - \mathbf{x}_0) + \gamma(\mathbf{x}_j - \mathbf{x}_0) - \mathbb{E} \{ [V(\mathbf{x}_i) - V(\mathbf{x}_0)][V(\mathbf{x}_j) - V(\mathbf{x}_0)] \} \end{aligned}$$

Therefore,

$$\boxed{\text{Cov}[V(\mathbf{x}_i) - V(\mathbf{x}_0), V(\mathbf{x}_j) - V(\mathbf{x}_0)] = \gamma(\mathbf{x}_i - \mathbf{x}_0) + \gamma(\mathbf{x}_j - \mathbf{x}_0) - \gamma(\mathbf{x}_i - \mathbf{x}_j)} \quad (1)$$

Now consider kriging equations to find best linear unbiased predictor:

$$\text{Linear} \quad \hat{V} = \sum_{j=1}^n \lambda_j V(\mathbf{x}_j)$$

$$\text{Unbiased} \quad \mathbb{E}(\hat{V}) = m \quad : \quad \sum_{j=1}^n \lambda_j m = m \quad \text{or} \quad \sum_{j=1}^n \lambda_j = 1.$$

$$\text{MSE} = \mathbb{E} [\hat{V} - V(\mathbf{x}_0)]^2 = \mathbb{E} \left\{ \sum_{j=1}^n \lambda_j [V(\mathbf{x}_j) - V(\mathbf{x}_0)] \right\}^2 =$$

$$\begin{aligned}
&= \sum_{k=1}^n \sum_{j=1}^n \lambda_k \mathbb{E} \{ [V(\mathbf{x}_k) - V(\mathbf{x}_0)][V(\mathbf{x}_j) - V(\mathbf{x}_0)] \} \lambda_j = \quad \text{using (1)} \\
&= \sum_{k=1}^n \sum_{j=1}^n \lambda_k [\gamma(\mathbf{x}_k - \mathbf{x}_0)] \lambda_j + \sum_{k=1}^n \sum_{j=1}^n \lambda_k [\gamma(\mathbf{x}_j - \mathbf{x}_0)] \lambda_j - \sum_{k=1}^n \sum_{j=1}^n \lambda_k [\gamma(\mathbf{x}_k - \mathbf{x}_j)] \lambda_j
\end{aligned}$$

Finally (argh), we obtain

$$\text{MSE} = 2 \sum_{j=1}^n \lambda_j \gamma(\mathbf{x}_j - \mathbf{x}_0) - \sum_{j=1}^n \sum_{i=1}^n \lambda_i \lambda_j \gamma(\mathbf{x}_i - \mathbf{x}_j) \quad (2)$$

Again, minimize with the constraint that $\sum_{j=1}^n \lambda_j = 1$. Using Lagrange multiplier μ , minimize

$$H(\lambda_1, \dots, \lambda_n; \mu) = \text{MSE} - 2\mu \left[\sum_{j=1}^n \lambda_j - 1 \right],$$

take partials with respect to λ_i, μ and equate to 0:

$$2\gamma(\mathbf{x}_i - \mathbf{x}_0) - 2 \sum_{j=1}^n \lambda_j \gamma(\mathbf{x}_i - \mathbf{x}_j) - 2\mu = 0, \quad \text{therefore}$$

$$\begin{aligned}
\sum_{j=1}^n \lambda_j \gamma(\mathbf{x}_j - \mathbf{x}_i) &= \gamma(\mathbf{x}_i - \mathbf{x}_0) - \mu, \quad i = 1, \dots, n \quad \text{and} \\
\sum_{j=1}^n \lambda_j &= 1
\end{aligned}$$

Kriging error, using (2), is $\sigma_{\text{OK}}^2 = \sum_{j=1}^n \lambda_j \gamma(\mathbf{x}_j - \mathbf{x}_0) + \mu$

Note: If we use $\gamma(\mathbf{x}_i - \mathbf{x}_j) = C(0, 0) - C(i, j)$ in stat. homog. case, then get back to ordinary kriging equations from last Lecture.

In matrix form: let $\mathbf{\Gamma}$ be a matrix with entries $\gamma_{ij} = \gamma(\mathbf{x}_i - \mathbf{x}_j)$. Let also $\mathbf{1} = (1, 1, \dots, 1)'$ an n -vector of ones, and $\mathbf{a} =$ vector of $\gamma(\mathbf{x}_i - \mathbf{x}_0)$.

Then

OK using variogram

$$\begin{bmatrix} \mathbf{\Gamma} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ 1 \end{pmatrix}$$

OK using covariance

$$\begin{bmatrix} \mathbf{C} & -\mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ 1 \end{pmatrix},$$

keeping in mind $\mathbf{\Gamma} = \sigma^2 \mathbf{I} - \mathbf{C}$ and $\mathbf{a} = \sigma^2 \mathbf{1} - \mathbf{b}$.

II. Nugget effect

Now: allow for measurement error.

First, let $V(\mathbf{x}_j)$ stat. homog., $Cov[V(\mathbf{x} + \mathbf{h}), V(\mathbf{x})] = C_V(\mathbf{h})$.

$$\text{Measurements: } \tilde{V}_j := V(\mathbf{x}_j) + W_j, \quad j = 1, \dots, n$$

where W_j represent measurement errors (think residuals for regression) and

$$\mathbb{E}(W_j) = 0, \quad Var(W_j) = \sigma_W^2,$$

and W_j are independent of each other and everything else.

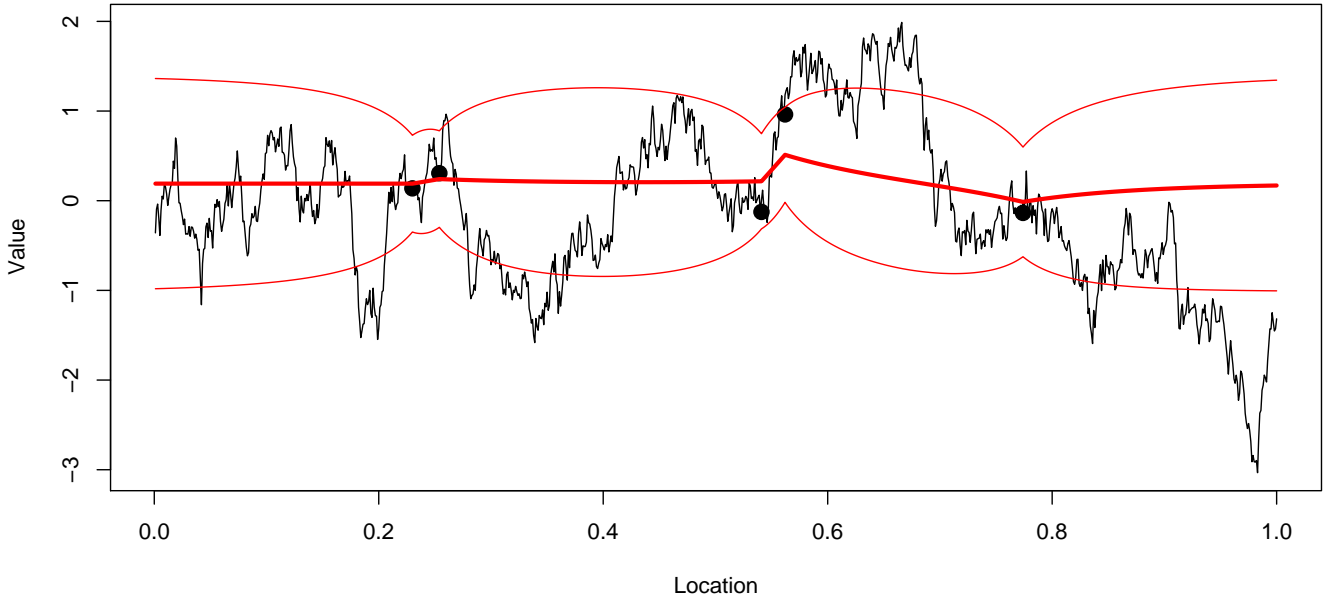
Goal: predict $V(\mathbf{x}_0)$ (and not $V(\mathbf{x}_0) + W_0$). Now

$$\tilde{C}(i, j) \equiv C_{\tilde{V}}(i, j) = Cov[V(\mathbf{x}_i) + W_i, V(\mathbf{x}_j) + W_j] = \begin{cases} C_V(\mathbf{x}_i - \mathbf{x}_j) & \text{if } i \neq j \\ C_V(\mathbf{0}) + \sigma_W^2 & \text{if } i = j \end{cases}$$

Ordinary kriging equations apply, but the \mathbf{C} -matrix changes.

Not an exact interpolator, smoothing effect. Also note the increase in kriging variance.

Ordinary kriging, cov. length = 0.1, nugget Var = 0.5



Result for IRF-0 is similar. The effect of adding noise on (semi)variogram is

$$\tilde{\gamma}(\mathbf{h}) = \gamma(\mathbf{h}) + \sigma_W^2, \quad \mathbf{h} \neq \mathbf{0}$$

III. Block kriging

Sometimes we get weighted averages of the data:

$$V_{ave}(x) = \int_a^b g(x-y)V(y) dy$$

(say in 1-d), g is some averaging function.

E.g. pump test (circle picture); or ore grade data averaged over some volume etc.

Say, observations are

$$W_j = \int_{a_j}^{b_j} V(y) dy, \quad j = 1, \dots, n$$

and we need to predict

$$W_0 = \int_{a_0}^{b_0} V(y) dy$$

Let $E[V(x)] = const = m$ and $Cov[V(x+\xi), V(x)] = C_V(\xi)$.

Consider $C_W(i, k) = Cov(W_i, W_k)$ - will depend on C_V , and

$$\mathbb{E}(W_j) = \int_{a_j}^{b_j} \mathbb{E}[V(y)] dy = m[b_j - a_j]$$

Consider predictor

$$\hat{W}_0 = \sum_{j=1}^n \lambda_j W_j,$$

unbiasedness condition is

$$\mathbb{E}(W_0) = m(b_0 - a_0) = \mathbb{E}(\hat{W}_0) \quad \text{hence} \quad \sum_{j=1}^n \lambda_j [b_j - a_j] = b_0 - a_0.$$

Kriging equations are simple, constraint differs:

$$\sum_{j=1}^n \lambda_j C_W(j, k) - \mu [b_k - a_k] = C_W(0, k), \quad k = 1, \dots, n$$

$$\sum_{j=1}^n \lambda_j [b_j - a_j] = b_0 - a_0.$$

To compute C_W , use double integral

$$Cov[W_j, W_k] = \int_{a_j}^{b_j} \int_{a_k}^{b_k} C_V(x, y) dx dy$$

may be messy to find.

In higher dimensions: \mathbf{x} is l -dim. vector, B_j is a region in l -space ($l = 1, 2, 3$).

$$W_j = \iint_{B_j} V(\mathbf{y}) d\mathbf{y},$$

$$Cov[W_j, W_k] = \iint_{B_j} \iint_{B_k} C_V(\mathbf{x} - \mathbf{y}) d\mathbf{x} d\mathbf{y}$$

Now the kriging equations become

$$\begin{cases} \sum_{j=1}^n \lambda_j C_W(j, k) - \mu |B_k| = C_W(0, k), & k = 1, \dots, n \\ \sum_{j=1}^n \lambda_j |B_j| = |B_0|, \end{cases}$$

where $|B_k|$ is the size of region k .

Also, kriging variance

$$\sigma_{\text{BK}}^2 = C_W(0, 0) - \sum_{j=1}^n \lambda_j C_W(j, 0) + \mu |B_0|$$

Can also kriging $V(\mathbf{x}_0)$ based on W_j 's: consider later under co-kriging.