ETH zürich

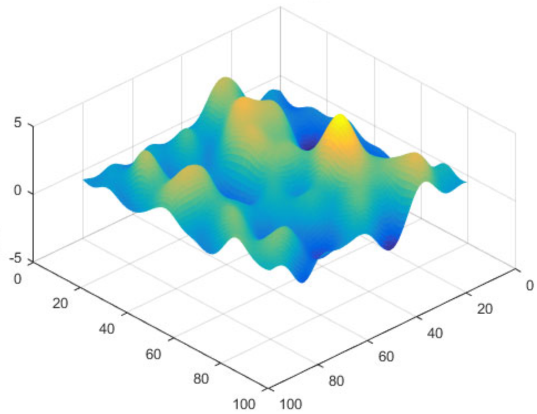
Project Parameter Estimation

Part II: Short introduction to Kriging

Jemil Avers Butt
Prof. Dr. Andreas Wieser

ETH zürich

What we realized last time



Stochastic process:

Family $X_t : \Omega \to \mathbb{R}, t \in \mathbb{R}$ of random variables

Random field:

Family $X_t : \Omega \to \mathbb{R}, t \in \mathbb{R}^n$ of random variables

We denote the spatial random field as a whole as

$$X : \Omega \times \mathbb{R}^n \to \mathbb{R}$$

Object to the left: only one realization of a RF.

- **Estimation between sampled points?**
- **Should respect local structure and should be somehow objective**
- **IDW for example: minimize inverse distance weighted error. Stochastic properties?**
- **Covariance and semivariance are stochastic quantities. Since they describe relations between neighbored data points: important tool for sound analogue of IDW**
- **State: Optimization approach using estimated value of a loss function**
- **Solve: Under certain assumptions, the resulting method is BLUP / Kriging.**

Optimization approach

X is SRF , $y_1,...y_n$ are from specific realization of SRF

Estimate X_S, one of the random variables that constitute X.

$$\text{Minimize } E[L] = \underbrace{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}}}_{\text{sum it all up}} \underbrace{L(\widehat{x_s}(\vec{y}), x_s)}_{\text{take specific loss}} \underbrace{f_{XY}(x_s, y_1, ..., y_n)}_{\text{weight by how probable it is}} dx_s dy_1 ... dy_n$$

$$= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \underbrace{\int_{\mathbb{R}} L(\widehat{x_s}(\vec{y}), x_s) f_{X|Y}(x_s \mid y_1, ..., y_n) dx_s}_{g(\hat{x}_s(\vec{y}), \vec{y})} f_Y(\vec{y}) dy_1 ... dy_n$$

$$= \qquad E[g(\widehat{x_s}(\vec{y}), \vec{y})]$$

Just minimize that g

Example: Square Loss $\qquad L(\widehat{x_s}, x_s) = (\widehat{x_s} - x_s)^2$

Minimize g $\qquad 0 = \int_{\mathbb{R}} \frac{\partial}{\partial \widehat{x_s}} (\widehat{x_s} - x_s)^2 f_X(x_s \mid \vec{Y} = \vec{y}) dx_s$

$$= 2\widehat{x_s} \int_{\mathbb{R}} f_X(x_s \mid \vec{Y} = \vec{y}) dx_s \; - \; 2\int_{\mathbb{R}} x_s f_X(x_s \mid \vec{Y} = \vec{y}) dx_s \stackrel{!}{=} 0$$

$$\Leftrightarrow \widehat{x_s} \qquad = \qquad E[X_s \mid \vec{Y} = \vec{y}]$$

- **Conditional expected value of jointly gaussian distributed random variables:**
- **Just linear combination of observations.** $E[X_s \mid \vec{Y} = \vec{y}] = \sum_{i=1}^{n} \lambda_i y_i$ **(justification of all that is to come)**

Derivation of Kriging equations (Blackboard)

Some rules for expected values and covariances

I. Expected value

a) $\mathrm{E}[X] \qquad = \int_{\mathbb{R}} x f(x) dx$

b) $\mathrm{E}[aX+b] \quad = a\,\mathrm{E}[X] + b \qquad a, b \in \mathbb{R}$

II. Covariance

a) $\sigma(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$

b) $\sigma(X,X) \geq 0$

c) $\sigma(X,Y) = \sigma(Y,X)$

d) $\sigma(\lambda_1 X_1 + \lambda_2 X_2, Y) = \lambda_1 \sigma(X_1, Y) + \lambda_2 \sigma(X_2, Y) \qquad X_1, X_2\, RV's\, ;\, \lambda_1, \lambda_2 \in \mathbb{R}$

e) $\sigma(X + a, Y) = \sigma(X,Y) \qquad a \in \mathbb{R}$

1. Rules for expected values and covariances
2. Derivations of equations from optimality conditions
3. List of assumptions
4. Explanation of interesting special cases
5. Some explicit formulas

Derivation of Kriging equations (Blackboard)

Optimality conditions

$$\widehat{X_s} = \sum_{i=1}^{n} \lambda_i X_i \quad ; \quad \widehat{X_s} - X_s = \epsilon = \text{ the error}$$

a) **Unbiasedness:**

Derivation of Kriging equations (Blackboard)

b) **Minimum Variance:**

Derivation of Kriging equations (Blackboard)

List of assumptions

I. Translation invariance of expected value
II. Translation invariance of covariances
III. Square loss and gaussian distribution

A process with properties I and II is called second order stationary.

I. And II. together are also called «wide sense stationarity up to order 2»

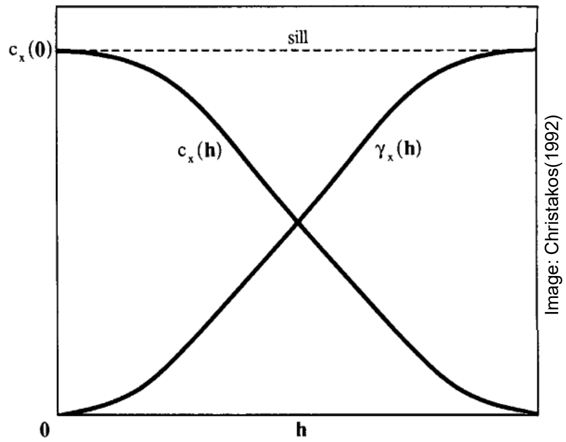
How to estimate those covariances?

$$\gamma(X_s,X_t)=\frac{1}{2}E\left[(X_s-X_t)^2\right]$$

$$\begin{aligned}\sigma(X_s,X_t) &\overset{\text{stationarity}}{=} \sigma(X_s,X_s)-\gamma(X_s,X_t) \quad \text{(Proof:HW)}\\ &= \sigma_0^2 - \gamma(s,t) \quad \text{(Notation)}\end{aligned}$$

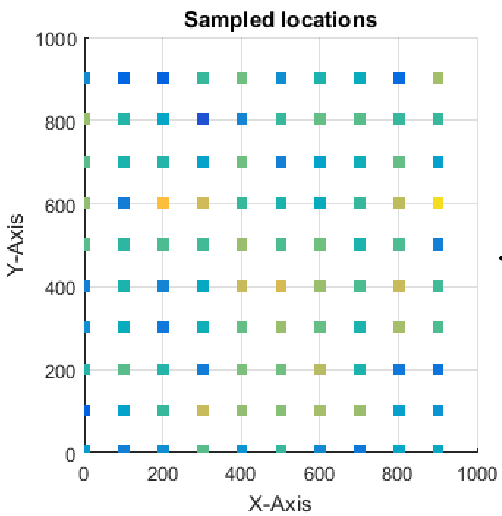
Consider isotropic fields: $\gamma(s,t)=\gamma(\|s-t\|)=\gamma(h)$

$$\hat{\gamma}(h_i)=\frac{1}{2n(h_i)}\sum_{k=1}^{n(h_i)}(X_{s_k}-X_{t_k})^2 \quad \text{where } \|s_k-t_k\|\approx h_i$$



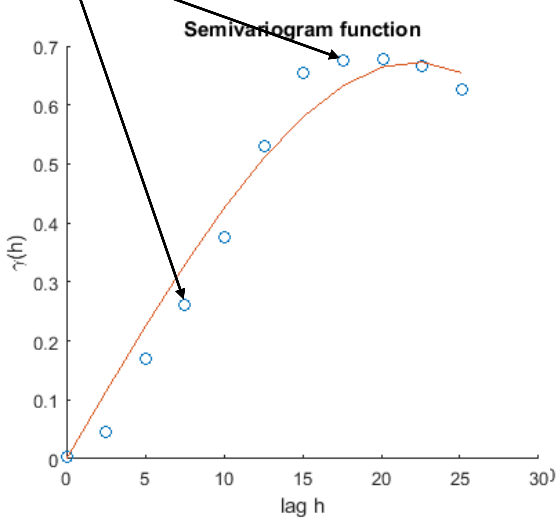
Inference of semivariance

$$\hat{\gamma}(h_i)=\frac{1}{2n(h_i)}\sum_{k=1}^{n(h_i)}(X_{s_k}-X_{t_k})^2 \qquad \|s_k-t_k\|\approx h_i$$



n points

n^2 distances between pairs of points



$\Rightarrow$ Whenever calculation of σ,γ is needed in the future, just use $\gamma(h)$ for that.
It tells you, how two RV's are correlated based on their distance.

Summary of inference

$$\hat{X}_s = \sum_{k=1}^{n} \lambda_k X_k$$

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \bar{\kappa} \end{bmatrix} = \begin{bmatrix} \gamma(X_1,X_1) & \cdots & \gamma(X_1,X_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(X_n,X_1) & \cdots & \gamma(X_n,X_n) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(X_1,X_s) \\ \vdots \\ \gamma(X_n,X_s) \\ 1 \end{bmatrix}$$

1. Estimate $\gamma(h)$ from the data (this is a function)
2. Calculate $\gamma(X_i, X_j)$ based on distance of RV's
3. Solve system and get $\lambda's$
4. Plug in measurements x_k in formula $\hat{X}_s = \sum_{k=1}^{n} \lambda_k X_k$ instead of RV's

$\Rightarrow$ Optimal estimation of X_s. Repeat for different positions you want to estimate RV's for.

Derivation of Kriging equations (compact)

Assumptions:

$i)$ Wide-sense stationarity

$a)\ E[X_t] = const. = \mu_X \qquad \forall t \in \mathbb{R}$

$b)\ \sigma(X_t, X_{t+\tau}) = \sigma(t, t+\tau) = \sigma(\tau) \qquad \forall (t,\tau) \in \mathbb{R}^2$

$ii)$ Underlying distribution is gaussian; quadratic loss function

Linearity: $\widehat{X_s} = \sum_{i=1}^{n} \lambda_i X_i$

Unbiasedness: $E[\widehat{X_s} - X_s] = E[\epsilon] = 0$

$$\Leftrightarrow E[\sum_{i=1}^{n} \lambda_i X_i - X_s] = \sum_{i=1}^{n} \lambda_i E[X_i] - E[X_s] = 0$$

$$\Leftrightarrow \mu_X \left(\sum_{i=1}^{n} \lambda_i - 1 \right) = 0$$

$$\Leftrightarrow \quad \left(\sum_{i=1}^{n} \lambda_i - 1 \right) = 0$$

Optimality: $\sigma(\epsilon,\epsilon) = \sigma(\widehat{X_s} - X_s, \widehat{X_s} - X_s) \rightarrow min$

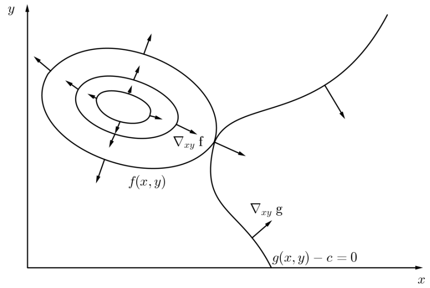
$$\begin{aligned} \sigma(\widehat{X_s} - X_s, \widehat{X_s} - X_s) &= \sigma(\widehat{X_s} - X_s, \widehat{X_s}) - \sigma(\widehat{X_s} - X_s, X_s) \\ &= \sigma(\widehat{X_s}, \widehat{X_s}) - \sigma(X_s, \widehat{X_s}) - \sigma(\widehat{X_s}, X_s) + \sigma(X_s, X_s) \\ &= \sigma(\widehat{X_s}, \widehat{X_s}) - 2\sigma(\widehat{X_s}, X_s) + \sigma(X_s, X_s) \\ &= \sigma\left(\sum_{i=1}^{n} \lambda_i X_i, \sum_{i=1}^{n} \lambda_i X_i \right) - 2\sigma\left(\sum_{i=1}^{n} \lambda_i X_i, X_s \right) + \sigma(X_s, X_s) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \sigma(X_i, X_j) - 2\sum_{i=1}^{n} \lambda_i \sigma(X_i, X_s) + \sigma(X_s, X_s) \end{aligned}$$

- **Now minimize this with respect to the λ_i's. Solution should also respect the unbiasedness-condition.**

Derivation of Kriging equations

$$m := \left\{(x,y)\in\mathbb{R}^2 \mid g(x,y)-c=0\right\}$$

$$\nabla_{xy} g \perp \text{tangent at } m$$



Assume $\nabla_{xy} f = -\kappa \nabla_{xy} g$

$\Rightarrow \nabla_{xy} f \perp$ tangent at m

$\Rightarrow$ No change of f along m in this point

$\Rightarrow$ Potential extremum

$$\nabla_{xy} f = -\kappa\nabla_{xy} g \Leftarrow \nabla_{xy}\underbrace{\left(f+\kappa(g-c)\right)}_{\Lambda}=0$$

$$\Lambda=\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\sigma\left(X_i,X_j\right)-2\sum_{i=1}^{n}\lambda_i\sigma\left(X_i,X_s\right)+\sigma(X_s,X_s)+\kappa\left(\sum_{i=1}^{n}\lambda_i-1\right)$$

$$\frac{\partial}{\partial\lambda_1}\Lambda=2\lambda_1\sigma(X_1,X_1)+\sum_{j\neq 1}\lambda_j\sigma(X_1,X_j)+\sum_{i\neq 1}\lambda_i\sigma(X_i,X_1)-2\sigma(X_1,X_s)+2\kappa^*\overset{!}{=}0$$

$$=2\lambda_1\sigma(X_1,X_1)+2\sum_{j\neq 1}\lambda_j\sigma(X_1,X_j) \qquad -2\sigma(X_1,X_s)+2\kappa^*\overset{!}{=}0$$

$$\frac{\partial}{\partial\lambda_l}\Lambda=2\lambda_1\sigma(X_l,X_l)+2\sum_{j\neq l}\lambda_j\sigma(X_l,X_j) \qquad -2\sigma(X_l,X_s)+2\kappa^*\overset{!}{=}0$$

$$\frac{\partial}{\partial\kappa}\Lambda= \qquad \sum_{i=1}^{n}\lambda_i-1= \qquad \sum_{i=1}^{n}2\lambda_i-2 \qquad \overset{!}{=}0$$

Derivation of Kriging equations

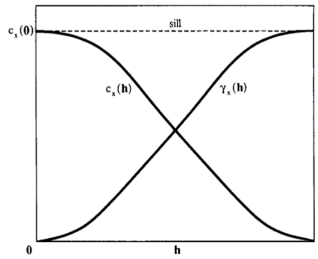


Image: Christakos(1992)

$$\begin{bmatrix}2\sigma(X_1,X_1) & \cdots & 2\sigma(X_1,X_n) & 2\\ \vdots & \ddots & \vdots & \vdots\\ 2\sigma(X_n,X_1) & \cdots & 2\sigma(X_n,X_n) & 2\\ 2 & \cdots & 2 & 0\end{bmatrix}\begin{bmatrix}\lambda_1\\ \vdots\\ \lambda_n\\ \kappa^*\end{bmatrix}=\begin{bmatrix}2\sigma(X_1,X_s)\\ \vdots\\ 2\sigma(X_n,X_s)\\ 2\end{bmatrix}$$

$$\Leftrightarrow \qquad \begin{bmatrix}\lambda_1\\ \vdots\\ \lambda_n\\ \kappa^*\end{bmatrix}=\begin{bmatrix}\sigma(X_1,X_1) & \cdots & \sigma(X_1,X_n) & 1\\ \vdots & \ddots & \vdots & \vdots\\ \sigma(X_n,X_1) & \cdots & \sigma(X_n,X_n) & 1\\ 1 & \cdots & 1 & 0\end{bmatrix}^{-1}\begin{bmatrix}\sigma(X_1,X_s)\\ \vdots\\ \sigma(X_n,X_s)\\ 1\end{bmatrix}$$

$$\begin{aligned}2\sigma(X_1,X_2)=2E[X_1X_2]-2E[X_1]E[X_2] &\overset{\text{stationarity}}{=} 2E[X_1X_2]-E[X_1]^2-E[X_2]^2\\ &= 2\sigma(X_1,X_1)-\left(E[(X_1-X_2)^2]\right)\\ &= 2\sigma(X_1,X_1)-2\gamma(X_1,X_2)\end{aligned}$$

We can estimate $\gamma(X_1,X_2)$ **with the experimental (semi-) variogram.**

$$\Rightarrow \qquad \begin{bmatrix}\lambda_1\\ \vdots\\ \lambda_n\\ \bar{\kappa}\end{bmatrix}=\begin{bmatrix}\gamma(X_1,X_1) & \cdots & \gamma(X_1,X_n) & 1\\ \vdots & \ddots & \vdots & \vdots\\ \gamma(X_n,X_1) & \cdots & \gamma(X_n,X_n) & 1\\ 1 & \cdots & 1 & 0\end{bmatrix}^{-1}\begin{bmatrix}\gamma(X_1,X_s)\\ \vdots\\ \gamma(X_n,X_s)\\ 1\end{bmatrix}$$

Special cases

Case where A has form

$$A=\begin{bmatrix}\sigma(X_1,X_1) & \cdots & \sigma(X_1,X_n) & 1\\ \vdots & \ddots & \vdots & \vdots\\ \sigma(X_n,X_1) & \cdots & \sigma(X_n,X_n) & 1\\ 1 & \cdots & 1 & 0\end{bmatrix}=\begin{bmatrix}\sigma_{11} & & & & 1\\ & \sigma_{22} & & & 1\\ & & \ddots & & \vdots\\ & & & \sigma_{nn} & 1\\ 1 & 1 & \cdots & 1 & 0\end{bmatrix}$$

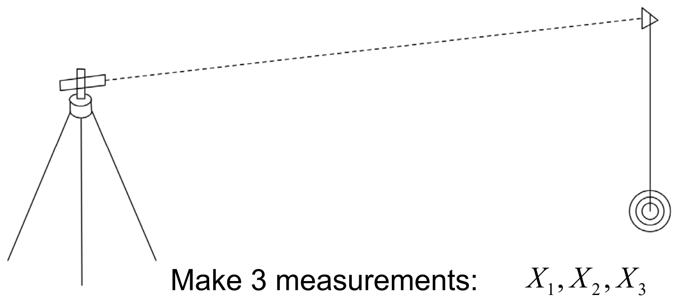
$$A^{-1}=\begin{bmatrix} ? & ? & ? & \dfrac{\prod_{j\neq 1}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ ? & ? & ? & \vdots\\ ? & ? & ? & \dfrac{\prod_{j\neq n}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ ? & ? & ? & ?\end{bmatrix}$$

$$\lambda \quad = \quad A^{-1} \quad c$$

$$\begin{bmatrix}\lambda_1\\ \vdots\\ \lambda_n\\ \kappa^*\end{bmatrix}=\begin{bmatrix}\sigma(X_1,X_1) & \cdots & \sigma(X_1,X_n) & 1\\ \vdots & \ddots & \vdots & \vdots\\ \sigma(X_n,X_1) & \cdots & \sigma(X_n,X_n) & 1\\ 1 & \cdots & 1 & 0\end{bmatrix}^{-1}\begin{bmatrix}\sigma(X_1,X_s)\\ \vdots\\ \sigma(X_n,X_s)\\ 1\end{bmatrix}$$

Example 1

$$A=\underbrace{\begin{bmatrix}\sigma_{00} & 0 & 0 & 1\\ 0 & \sigma_{00} & 0 & 1\\ 0 & 0 & \sigma_{00} & 1\\ 1 & 1 & 1 & 0\end{bmatrix}}_{\text{Our explicit formula holds}} \qquad c=\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

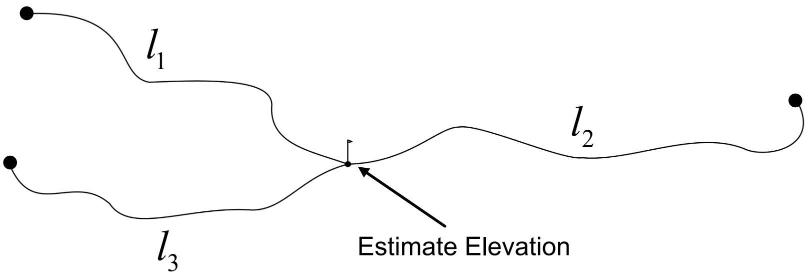


Make 3 measurements: X_1, X_2, X_3
Predict the 4th one: X_4
Measurements uncorrelated

$$A^{-1}=\begin{bmatrix} ? & ? & ? & \dfrac{\prod_{j\neq 1}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ ? & ? & ? & \vdots\\ ? & ? & ? & \dfrac{\prod_{j\neq n}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ ? & ? & ? & ?\end{bmatrix}$$

$$\begin{bmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\\ \kappa^*\end{bmatrix}=\begin{bmatrix} ? & ? & ? & \dfrac{\prod_{j\neq 1}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ ? & ? & ? & \dfrac{\prod_{j\neq 2}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ ? & ? & ? & \dfrac{\prod_{j\neq 3}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ ? & ? & ? & ?\end{bmatrix}\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}\dfrac{\prod_{j\neq 1}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ \dfrac{\prod_{j\neq 2}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ \dfrac{\prod_{j\neq 3}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}\\ ?\end{bmatrix}=\begin{bmatrix}\dfrac{\sigma_{00}*\sigma_{00}}{\sigma_{00}*\sigma_{00}+\sigma_{00}*\sigma_{00}+\sigma_{00}*\sigma_{00}}\\ \dfrac{\sigma_{00}*\sigma_{00}}{\sigma_{00}*\sigma_{00}+\sigma_{00}*\sigma_{00}+\sigma_{00}*\sigma_{00}}\\ \dfrac{\sigma_{00}*\sigma_{00}}{\sigma_{00}*\sigma_{00}+\sigma_{00}*\sigma_{00}+\sigma_{00}*\sigma_{00}}\\ ?\end{bmatrix}=\begin{bmatrix}\frac{1}{3}\\ \frac{1}{3}\\ \frac{1}{3}\\ ?\end{bmatrix}$$

Example 2



$$A=\begin{bmatrix}\sigma_{11} & 0 & 0 & 1\\ 0 & \sigma_{22} & 0 & 1\\ 0 & 0 & \sigma_{33} & 1\\ 1 & 1 & 1 & 0\end{bmatrix} \qquad c=\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

$$\lambda_1=\frac{\prod_{j\neq 1}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}=\frac{\sigma_{22}\sigma_{33}}{\sigma_{11}\sigma_{22}+\sigma_{22}\sigma_{33}+\sigma_{11}\sigma_{33}}=\frac{2}{2+2+1}=0.4$$

$$\lambda_2=\frac{\prod_{j\neq 2}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}=\frac{\sigma_{11}\sigma_{33}}{\sigma_{11}\sigma_{22}+\sigma_{22}\sigma_{33}+\sigma_{11}\sigma_{33}}=\frac{1}{2+2+1}=0.2$$

$$\lambda_3=\frac{\prod_{j\neq 3}\sigma_{jj}}{\sum_{k=1}^{n}\prod_{j\neq k}\sigma_{jj}}=\frac{\sigma_{11}\sigma_{22}}{\sigma_{11}\sigma_{22}+\sigma_{22}\sigma_{33}+\sigma_{11}\sigma_{33}}=\frac{2}{2+2+1}=0.4$$

Our explicit formula holds

$$\lambda_1=\frac{\sigma_{22}\sigma_{33}}{\sigma_{11}\sigma_{22}+\sigma_{22}\sigma_{33}+\sigma_{11}\sigma_{33}}$$

$$=\frac{\sigma_{22}\sigma_{33}}{\sigma_{22}\sigma_{33}}\;\frac{1}{\frac{\sigma_{11}\sigma_{22}}{\sigma_{22}\sigma_{33}}+\frac{\sigma_{22}\sigma_{33}}{\sigma_{22}\sigma_{33}}+\frac{\sigma_{11}\sigma_{33}}{\sigma_{22}\sigma_{33}}}$$

$$=\frac{1}{\frac{\sigma_{11}}{\sigma_{33}}+\frac{\sigma_{11}}{\sigma_{11}}+\frac{\sigma_{11}}{\sigma_{22}}}$$

$$=\frac{1}{\sigma_{11}}\left(\sum_i\frac{1}{\sigma_{ii}}\right)^{-1} \qquad =\frac{p_1}{\sum_i p_i}$$

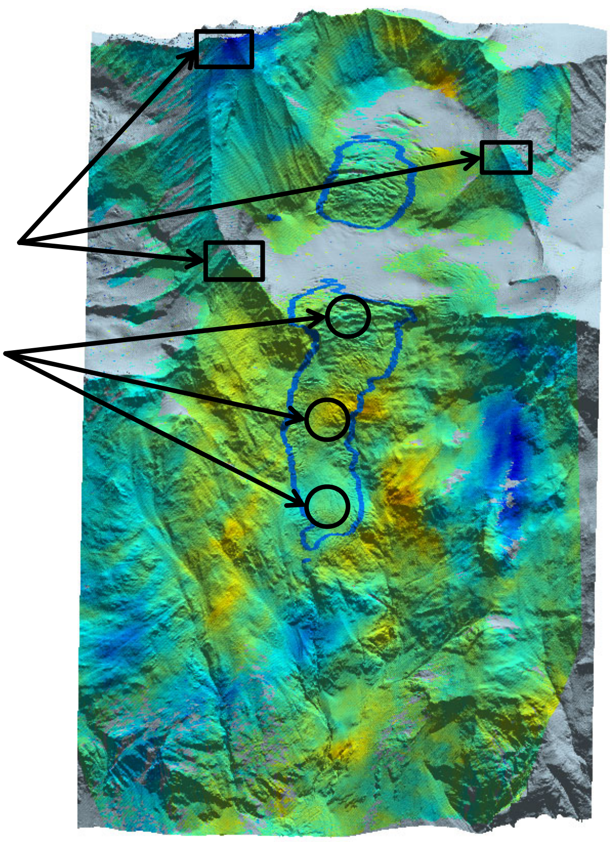
Example 3

X is noise; to be estimated and removed

X is a collection of random variables X(s)

- Realization at some points s_{o_i} known
- Realization on glacier important

  ... but unknown s_u



Estimating APS X at unobserved locations

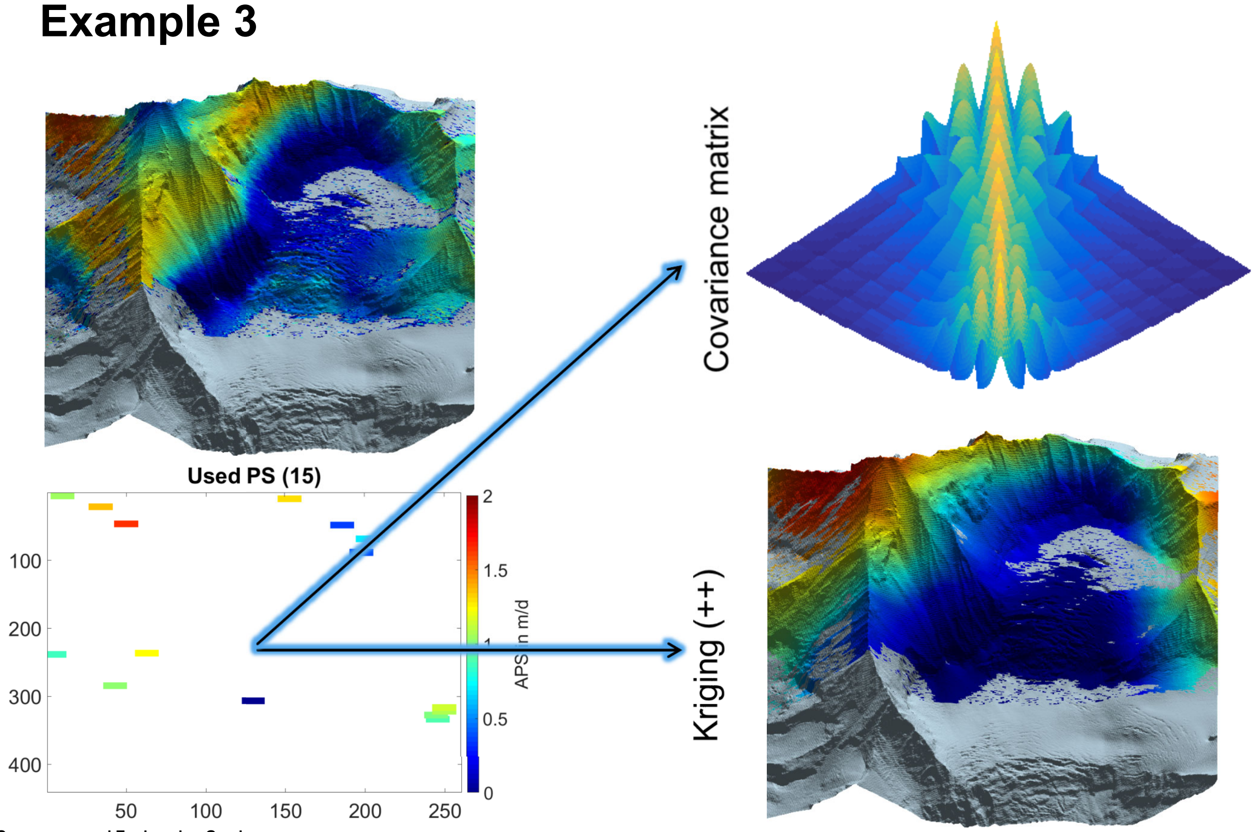
- Statistical inference problem
- Conditional expectation:

$$\mathrm{E}\left[\mathrm{X}(\mathrm{s_u})\;|\mathrm{X}(\mathrm{s_{o_1}}),\,\ldots\,,\mathrm{X}(\mathrm{s_{o_n}})\right]$$

- Jointly gaussian distributed,

$$\mathrm{X}(\mathrm{s_u},\mathrm{t_i})\coprod\mathrm{X}(\mathrm{s_o},\mathrm{t_j})\;|\mathrm{X}(\mathrm{s_o},\mathrm{t_i})_{i\neq j}$$

Example 3



Simulation

Matlab function "mvnrnd": realization=mvnrnd(mu, sigma)

R realization: (n,1) vector where n is nr of dimensions
M mu : (n,1) vector of expected values. Just set to zeros(n,1)
S sigma : (n,n) matrix of covariances. Must be positive definite

Matrix sigma will be generated by a covariance function $c(\cdot,\cdot)$ that determines the smoothness properties of the samples process: $S_{ij} = \sigma(X_{s_i}, X_{s_j}) = c(s_i, s_j)$

Variances always ≥ 0

$$\Rightarrow 0 \leq \sigma(Y,Y) = \sigma\left(\sum_{i=1}^{n} \alpha_i X_{s_i}, \sum_{j=1}^{n} \alpha_j X_{s_j}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j \sigma(X_{s_i}, X_{s_j}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j c(s_i, s_j)$$

If not fulfilled $\Rightarrow$ no valid process (negative variances ...)

Simulation

Here are some valid covariance functions on the unit interval:

$$c(s,t) = \qquad \begin{array}{ccc} \delta_{ts} & \dfrac{1}{t^2+s^2+1} & e^{-\frac{\|t-s\|}{r}} \\ st & \min(t,s) & e^{-\frac{\|t-s\|^2}{r^2}} \end{array}$$

$s,t \in [0,1]$

$r \in \mathbb{R}_+$

$c_{\text{new}}(s,t) = \alpha * c(s,t)$ legal, if $\alpha > 0$

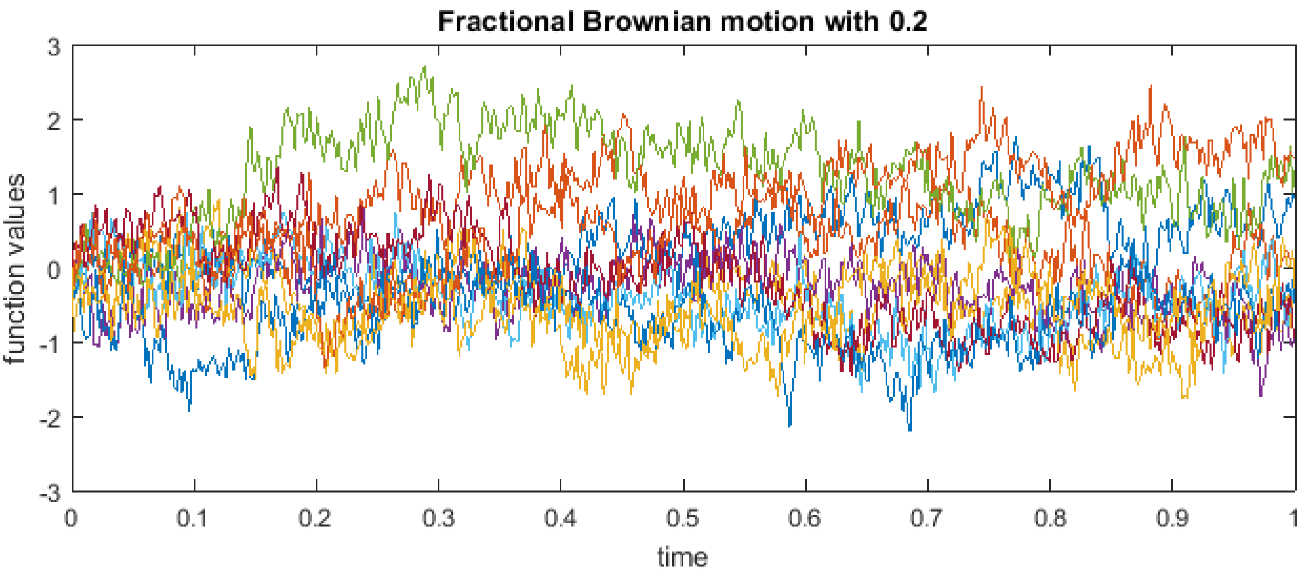
$c_{\text{new}}(s,t) = c_1(s,t) + c_2(s,t)$ legal

$c_{\text{new}}(s,t) = c_1(s,t) * c_2(s,t)$ legal

$c_{\text{new}}(s,t) = A_s A_t c(\cdot,\cdot)$ legal

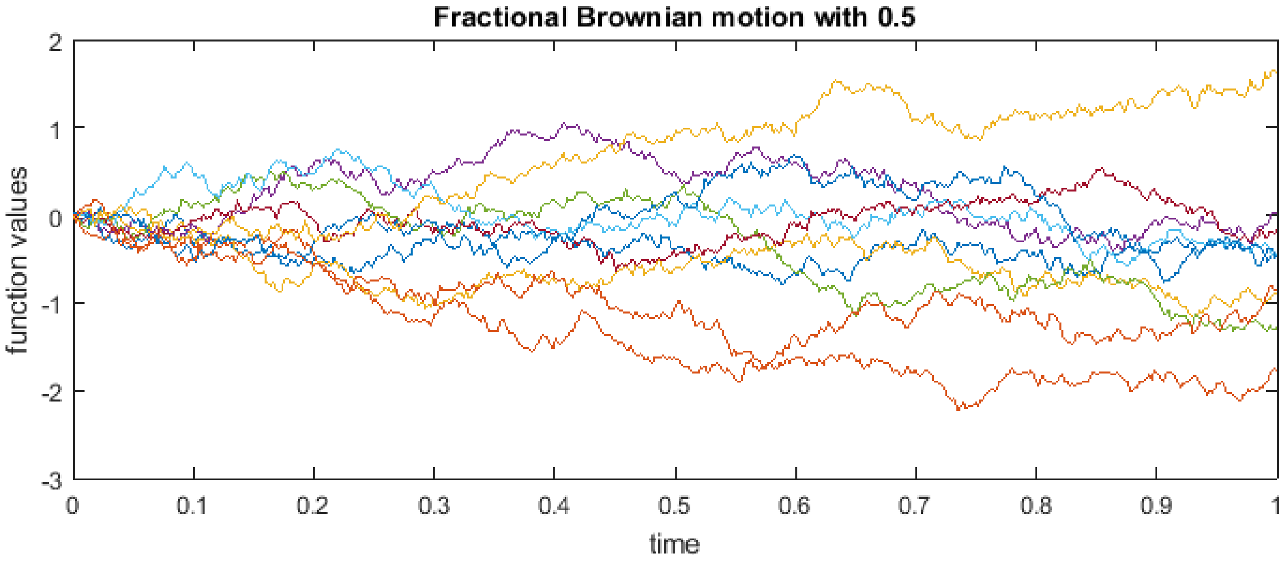
where A is any linear operator (e.g. $c_{\text{new}}(s,t) = \int_0^s \int_0^t c(u,v)dudv$ or A any matrix)

Examples



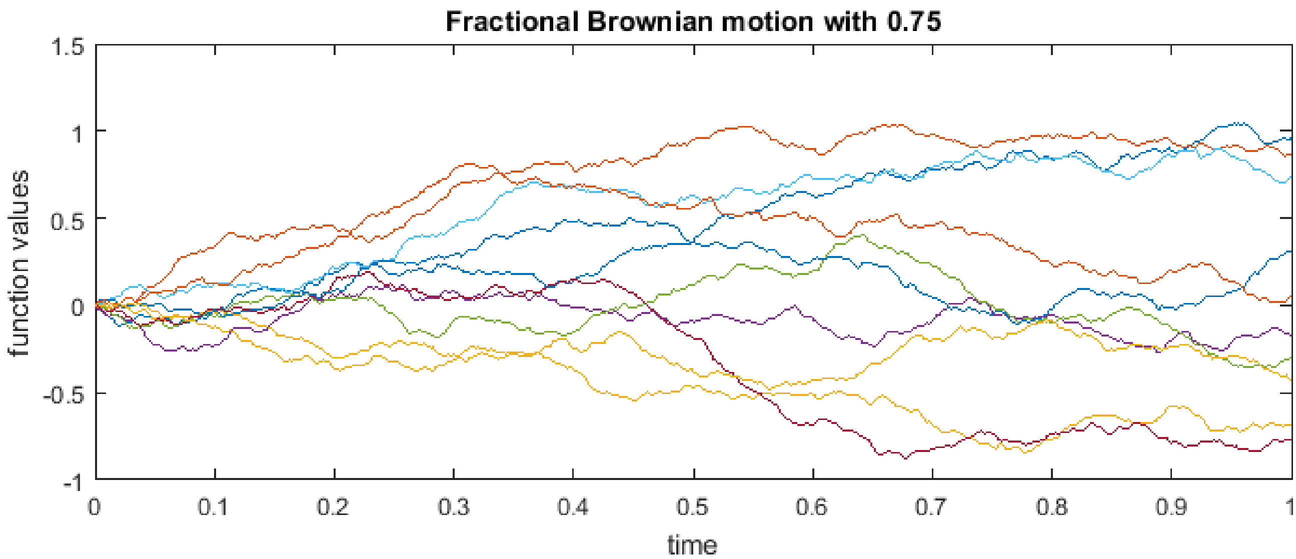
$$c(s,t) = \frac{1}{2}\left(|\, t\,|^{0.4} + |\, s\,|^{0.4} + |\, t-s\,|^{0.4} \right)$$

Examples



$$c(s,t)=\frac{1}{2}\left(|t|^{1}+|s|^{1}+|t-s|^{1}\right)$$

Examples

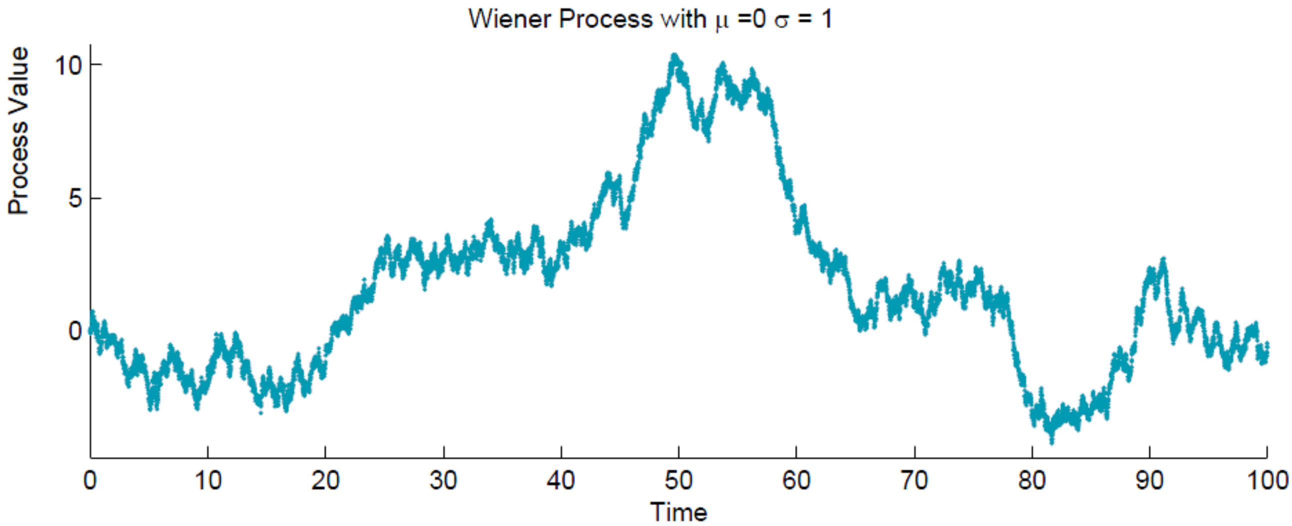


$$c(s,t)=\frac{1}{2}\left(|t|^{1.5}+|s|^{1.5}+|t-s|^{1.5}\right)$$

Homework

2.1 Stationarity and Warm-up

a) The “Wiener Process” is just integrated white noise. Please formulate the equation for a discrete Wiener Process and explain if it is stationary or not.



Homework

2.2 Stochastic Process with dice (I)

Let a Stochastic Process $\{X_t\}_{t=1\ldots n}$ be defined by :

$$X_1 = 1.75 + 0.5D_1$$
$$X_t = 0.5D_{t-1} + 0.5D_t \qquad \text{where } D_i \text{ is the i-th dice roll}$$

a) Calculate the expected value $\mathrm{E}[X_t]$ and the covariance $\sigma(X_t, X_{t+p})$ and draw the covariance as a function of p . Please explain!

b) Is this stochastic process second order stationary?

c) Derive the relationship on slide 7 between the covariance $\sigma(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$ and the semivariance $\gamma(X,Y) = \frac{1}{2}\mathrm{E}[(X-Y)^2]$ for second order stationary stochastic processes

d) Please grab a die and simulate $X_1 \ldots X_{30}$. Then use the relationship from c) to estimate with Matlab or Excel $\sigma(X_t, X_{t+p})$ by using only $\hat{\gamma}(X_t, X_{t+p})$ and $\hat{\sigma}(X_t, X_t)$.

Homework

2.3 Stochastic Process with dice (II)

Let a Stochastic Process $\{X_t\}_{t=1\ldots n}$ be defined by :

$$X_1 = 1.75 + 0.5D_1$$
$$X_t = 0.5X_{t-1} + 0.5D_t \qquad \text{where } D_i \text{ is the i-th dice roll}$$

a) Give a first estimate of what you expect the covariance (as a function of distance p between two random variables) to be. Why do you think this behavior reasonable?

b) Calculate the expected Value $\mathrm{E}[X_t]$. **Hint:** Geometric series, see explanation on slides.

c) Calculate the covariance $\sigma(X_t.X_{t+p})$ and decide if X_t can be regarded as a second order stationary stochastic process. **Hint:** Geometric series, see explanation on slides.

d) After carrying out a simulation of this dice roll game, Bob realizes that he forgot to write down X_{50} . But he has all the other $X's : X_1 \ldots X_{100}$. Please help Bob and write down (schematically) the Kriging system for the best linear predictor for X_{50} .

Homework

Hint: Analysis I, geometric series

$$\alpha < 1$$

$$\Rightarrow \sum_{k=1}^{n} \alpha^k \text{ converges}$$

$$(1\text{-}\alpha)\sum_{k=1}^{n}\alpha^k = \sum_{k=1}^{n}\alpha^k - \sum_{k=2}^{n+1}\alpha^k$$
$$= \alpha - \alpha^{n+1}$$
$$\Leftrightarrow \qquad \sum_{k=1}^{\infty}\alpha^k = \frac{\alpha - \alpha^{n+1}}{1-\alpha}$$

Otherwise the exercise should be straightforward

Homework

2.4 Gaussian Stochastic Processes (SP)

A (discrete) Gaussian process is a collection of random variables $\{X_t\}_{t=1...n}$; all together jointly Gaussian distributed.

a) Assume X_t to be a stationary Gaussian process. Is the expected value $\mathrm{E}[X_t]$ together with the covariance $\sigma(X_t, X_{t+1})$ enough to determine the whole process completely? Please explain (or give a counterexample).

b) Please define two second order stationary Gaussian processes $\{X_t\}_{t=1...300}$, $\{Y_t\}_{t=1...300}$ with:

$$\mathrm{E}[X_t] = 0 \qquad \mathrm{E}[Y_t] = 0$$

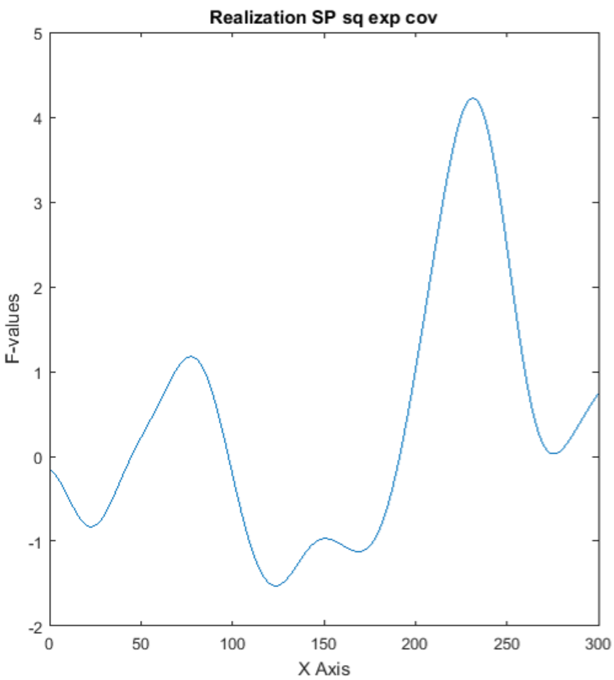
$$\sigma(X_t, X_{t+p}) = \sigma_0^2 e^{-\frac{|p|}{30}} \qquad \sigma(Y_t, Y_{t+p}) = \sigma_0^2 e^{-\left(\frac{|p|}{30}\right)^2} \qquad \sigma_0^2 = 1$$

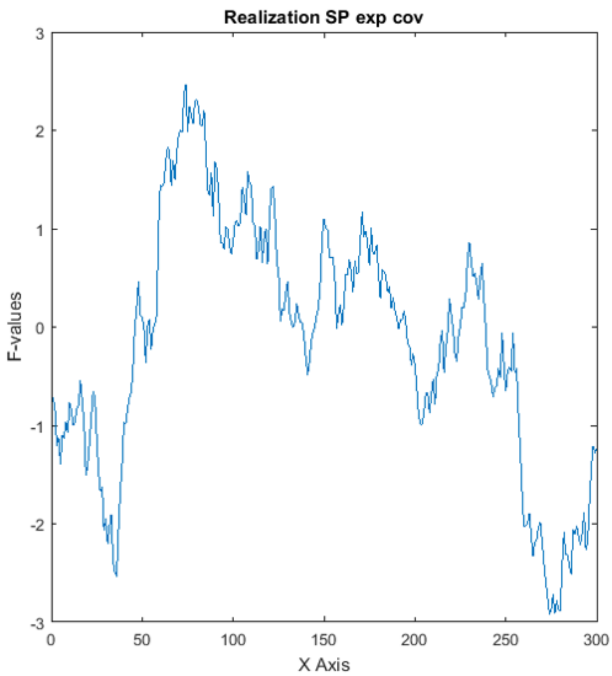
(Hint: Think of what Information you need to fix to completely determine a Gaussian process -> a))

c) Write a Matlab program that simulates the two stochastic processes you just invented! You can use the pre-built Matlab function mvnrnd (generate multivariate Gaussian random numbers – see slides 18, 19). Please plot the values of your Gaussian processes and also the correlation coefficients as a function of the distance between the random variables.

d) Compare the two different stochastic processes by describing and explaining their different behavior by relating it to the two covariance functions.

Homework





Homework

2.5 Gaussian Random Fields (RF)

Let $\{X_s\}_{s\in\{1\ldots30\}\times\{1\ldots30\}}, \{Y_s\}_{s\in\{1\ldots30\}\times\{1\ldots30\}}$ be two 2D discrete Gaussian random fields. The goal is now to simulate drawing realizations from these random fields. This can be done in a straightforward way by considering the 2D matrix to be simulated as a single long vector with appropriate covariance matrix calculated from the respective 2D distances.

a) Write a Matlab program that simulates the two random fields $X_.$ and $Y_.$ with properties as sketched below.

$$\mathrm{E}[X_s]=0 \qquad\qquad \mathrm{E}[Y_s]=0$$

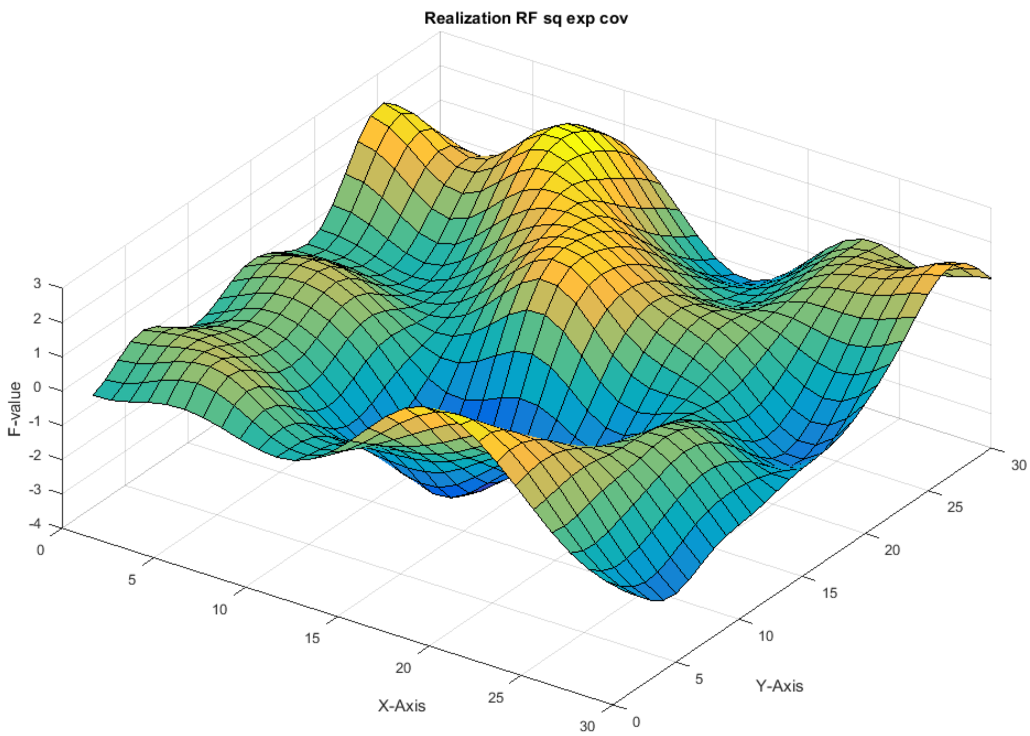
$$\sigma(X_s,X_{s+h})=\sigma_0^2 e^{-\frac{\|h\|}{5}} \qquad\qquad \sigma(Y_s,Y_{s+h})=\sigma_0^2 e^{-\left(\frac{\|h\|}{5}\right)^2} \qquad \sigma_0^2=1$$

where s and h are now vectors and $\|\cdot\|$ denotes the euclidean distance

b) Look at the covariance matrix of the random field you have just sampled from. Can you explain, why it looks how it looks?

c) Try to invent a random field with the weirdest behavior you can come up with.

Homework



Homework

2.6 Spectral Theory of Stochastic Processes (bonus points)

As is well known, Covariance matrices are positive definite and symmetric (see slide 18). Therefore, they have a unique eigenvalue-eigenvector decomposition. Many interesting properties like the mean energy of a stochastic processes and random fields or their smoothness properties can be traced back to these spectral properties.

a) Let $X_.$ be a discrete mean 0 SP with and denote by $\langle X_., X_. \rangle = \| X_. \|^2 = E\left[\sum_{t=1}^{n} X_t^2\right]$ its norm. Show that $\| X_. \|^2 = \sum_{k=1}^{n} \lambda_k$ where λ_k are the eigenvalues of the covariance matrix.

Hint: $\mathrm{trace}(AB) = \mathrm{trace}(BA)$

b) Look at the sequence of eigenvalues of the covariance matrices you used for simulation in 2.4 b). Can you somehow derive some further info about these processes?

c) Let $X_.$ and $Y_.$ be two SP's and let $\langle X_., Y_. \rangle = E\left[\sum_{t=1}^{n} X_t Y_t\right]$ be their inner product. Show that $\langle \cdot, \cdot \rangle$ is bilinear, symmetric, positive definite; i.e. has the same properties as the standard scalar product $a \bullet b = \sum a_i b_i$ that you know from linear algebra and geometry. What implications does this have?

Homework