Spatial Statistics (Spring 2013) – Peter Craigmile

Stationary geostatistical processes

• Introduction
  – Why do we model dependencies?
  – A geostatistical model

• The mean function

• Review: The covariance between two RVs

• The covariance function
  – Why nonnegative definite functions?
  – The correlation function

• Strictly and weakly stationary processes
  – The IID process and white noise process.
  – (Weakly) stationary processes
  – The covariance and correlation function of a stationary process
  – Bochner’s theorem
  – Norms; Isotropic and anisotropic processes

• Gaussian processes
Introduction

- Remember we defined a **geostatistical process** to be the stochastic process

\[ \{Z(s) : s \in D\}, \]

where the spatial domain \( D \subset \mathbb{R}^p \).

- We next study the statistical properties of such processes.
  
  - We do this so that we can build statistical models for our dependent geospatial data.
Why do we model dependencies?

• In Statistics, simple (parsimonious) models are often the norm.
  – Why do we need more complex models?

• If we ignore the dependencies that we observe in spatial data, then we can be led to incorrect statistical inferences.

• But, we will still try to keep our models as simple as possible by assuming stationarity.
  – Stationarity means that some characteristic of the distribution of a spatial process does not depend on the spatial location, only the displacement between the locations.
  – If you shift the spatial process, that characteristic of the distribution will not change.

• While most spatial data are not stationary,
  – there are ways to either remove or model the non-stationary parts (the components that depend on the spatial location),

so that we are only left with a stationary component.
A geostatistical model

- A **geostatistical model** is a specification or summary of the probabilistic distribution of a collection of random variables (RVs) \( \{Z(s) : s \in D\} \).

(The observed data is a **realization** of these RVs).

- The simplest geostatistical model is **IID noise**, where \( \{Z(s) : s \in D\} \) are independent and identically distributed random variables.
  - There is **no dependence** in this model.
  - We **cannot predict** at other spatial locations with IID noise as there are no dependencies between the random variables at different locations.
The mean function

- The **mean function** of \( \{ Z(s) \} \) is

\[
\mu_Z(s) = E(Z(s)), \quad s \in D.
\]

Think of \( \mu_Z(s) \) are being the theoretical mean/expectation at location \( s \), taken over the possible values that could have generated \( Z(s) \).

- When \( Z(s) \) is a **continuous RV**,

\[
\mu_Z(s) = \int_{-\infty}^{\infty} z f_{Z(s)}(z) \, dz,
\]

where \( f_{Z(s)}(\cdot) \) is the probability density function (pdf) for \( Z(s) \).

- When \( Z(s) \) is a **discrete RV**,

\[
\mu_Z(s) = \sum_{z=-\infty}^{\infty} z f_{Z(s)}(z),
\]

where \( f_{Z(s)}(\cdot) \) is the probability mass function (pmf) for \( Z(s) \).
Review: The covariance between two RVs

- The **covariance** between the RVs $X$ and $Y$ is

$$\text{cov}(X, Y) = E\{(X - \mu_X)(Y - \mu_Y)\}$$

$$= E(XY) - \mu_X \mu_Y.$$ 

It is a measure of linear dependence between the two RVs.

- When $X = Y$ we have

$$\text{cov}(X, X) = \text{var}(X).$$

- For constants $a$, $b$, $c$, and RVs $X$, $Y$, $Z$:

$$\text{cov}(aX + bY + c, Z) = \text{cov}(aX, Z) + \text{cov}(bY, Z)$$

$$= a \text{cov}(X, Z) + b \text{cov}(Y, Z).$$

- Thus

$$\text{var}(X + Y) = \text{cov}(X, X) + \text{cov}(X, Y) + \text{cov}(Y, X) + \text{cov}(Y, Y)$$

$$= \text{var}(X) + \text{var}(Y) + 2 \text{cov}(X, Y).$$
The covariance function

- The covariance function of \( \{Z(s) : s \in D\} \) is

\[
C_Z(s, t) = \text{cov}(Z(s), Z(t)) = \text{E}\{[Z(s) - \mu_Z(s)][Z(t) - \mu_Z(t)]\}.
\]

- The covariance measures the strength of linear dependence between the two RVs \( Z(s) \) and \( Z(t) \).

- Properties:

1. \( C_Z(s, t) = C_Z(t, s) \) for each \( s, t \in D \).

2. When \( s = t \) we obtain

\[
C_Z(s, s) = \text{cov}(Z(s), Z(s)) = \text{var}(Z(s)) = \sigma^2_Z(s),
\]

the variance function of \( \{Z(s)\} \).

3. \( C_Z(s, t) \) is a nonnegative definite function.
Why nonnegative definite functions?

• Consider the following weighted average of the geostatistical process \( \{Z(s)\} \) at \( n \geq 1 \) locations \( s_1, \ldots, s_n \):

\[
Y = \sum_{j=1}^{n} a_j Z(s_j),
\]

where \( a_1, \ldots, a_n \) are real constants.

• The variance of \( Y \) is

• A function \( f(\cdot, \cdot) \) is nonnegative definite if

\[
\sum_{j=1}^{n} \sum_{k=1}^{n} a_j f(s_j, s_k) a_k \geq 0,
\]

for all positive integers \( n \) and real-valued constants \( a_1, \ldots, a_n \).

• Thus \( C_Z(s, t) \) must be a nonnegative definite function.

Why?
The correlation function

• The correlation function of \( \{Z(s)\} \) is

\[
\rho_Z(s, t) = \text{corr}(Z(s), Z(t)) = \frac{C_Z(s, t)}{\sqrt{C_Z(s, s)C_Z(t, t)}}.
\]

• The correlation measures the strength of linear association between the two RVs \( Z(s) \) and \( Z(t) \).

• Properties:

1. \(-1 \leq \rho_Z(s, t) \leq 1\) for each \( s, t \in D \)
2. \( \rho_Z(s, t) = \rho_Z(t, s) \) for each \( s, t \in D \)
3. \( \rho_Z(t, t) = 1 \) for each \( t \in D \).
4. \( \rho_Z(s, t) \) is a nonnegative definite function.
Strictly stationary processes

• In **strict stationarity** the joint distribution of a set of RVs are unaffected by spatial shifts.

• A geostatistical process, \( \{Z(s) : s \in D\} \), is **strictly stationary** if

\[
(Z(s_1), \ldots, Z(s_n)) =_d (Z(s_1 + h), \ldots, Z(s_n + h))
\]

for all \( n \geq 1 \), spatial locations \( \{s_j : j = 1, \ldots, n\} \) and a displacement or spatial lag \( h \).

• Then:

1. \( \{Z(s) : s \in D\} \) is identically distributed
   
   – Not necessarily independent!

2. \( (Z(s), Z(s + h)) =_d (Z(0), Z(h)) \) for all \( s \) and \( h \);

3. When \( \mu_Z \) is finite, \( \mu_Z(s) = \mu_Z \) is independent of spatial location.

4. When the variance function exists,

\[
C_Z(s, t) = C_Z(s + h, t + h) \text{ for any } s, t \text{ and } h.
\]
(Weakly) stationary processes

- $\{Z(s) : s \in D\}$ is (weakly) stationary if

  1. $E(Z(s)) = \mu_Z(s) = \mu_Z$ for some constant $\mu_Z$ which does not depend on $s$.

  2. $\text{cov}(Z(s), Z(s + h)) = C_Z(s, s + h) = C_Z(h)$, a finite constant that can depend on $h$ but not on $s$.

- The quantity $h$ is called the spatial lag or displacement.

- Other terms for this type of stationarity include second-order, covariance, wide sense.

- Relating weak and strict stationarity:

  1. A strictly stationary process $\{Z(s)\}$ is also weakly stationary as long as $\sigma_Z^2(s)$ is finite for all $s$.

  2. **Weak** stationarity does not imply **strict** stationarity! (unless we have a Gaussian process – see later).
The covariance and correlation function of a stationary process

- The covariance function of a stationary process \( \{Z(s) : s \in D\} \) is defined to be

\[
C_{Z}(h) = \text{cov}(Z(s), Z(s + h)) = \mathbb{E}((Z(s) - \mu_Z)(Z(s + h) - \mu_Z)),
\]

and measures the amount of linear dependence between \( Z(s) \) and \( Z(s + h) \).

- Properties of the covariance function:

  1. \( C_{Z}(0) = \text{var}(Z(s)) > 0 \).
  2. \( |C_{Z}(h)| \leq C_{Z}(0) \) for all \( h \).
  3. \( C_{Z}(-h) = C_{Z}(h) \) for each \( h \) (\( C_{Z}(\cdot) \) is an even function).
  4. \( C_{Z}(s - t) \) as a function of \( s \) and \( t \) is nonnegative definite.
    (\( C_{Z}(\cdot) \) is nonnegative definite).

- Similarly we can define the correlation function of a stationary process \( \{Z(s)\} \) – denoted \( \rho_{Z}(h) \).

  - Has the same properties as the covariance function, except

    1. \( \rho_{Z}(0) = 1 \) and 2. \( -1 \leq \rho_{Z}(h) \leq 1 \) for all \( h \).
The white noise process

- Assume $E(Z(s)) = \mu$, $\text{var}(Z(s)) = \sigma^2 < \infty$.

- $\{Z(s)\}$ is a white noise process if

$$C_Z(h) = 0,$$

for $h \neq 0$.

- $\{Z(s)\}$ is clearly stationary.

- But, distributions of $Z(s)$ and $Z(t + h)$ can be different!

- All IID noise with finite variance is white noise

  (but, the converse need not be true).
What suffices to be a valid covariance or correlation function?

**Bochner’s theorem:**

- A real-valued function $C_Z(\cdot)$ defined on $\mathbb{R}^p$ is the covariance function of a stationary process if and only if
  
it is even and nonnegative definite.

or

- A real-valued function $\rho_Z(\cdot)$ defined on $\mathbb{R}^p$ is the correlation function of a stationary process if and only if
  
it is even and nonnegative definite AND $\rho_Z(0) = 1$.

- We can use Bochner’s theorem to construct stationary processes, using the literature of what is known about nonnegative definite functions.
Norms; Isotropic and anisotropic processes

- Let $||·||$ be some norm.

  - The most common norm is the Euclidean norm in $\mathbb{R}^p$,
    $$
    ||x|| = \sqrt{\sum_{j=1}^{p} x_j^2},
    $$
    where $x = (x_1, \ldots, x_p)^T$.

  - See Banerjee [2005] for a discussion of distances in different coordinate systems.

- A stationary process for which $C_Z(s, t)$ only depends on the distance between the locations, $||s - t||$, is called an isotropic process.

  - With a spatial lag $h$, the covariance function of an isotropic process can be written in terms of its length $||h||$:
    $$
    C_Z(s, s + h) = C_Z(||h||).
    $$

- If the covariance of a stationary process depends on the direction and distance between the locations, then the process is called anisotropic.
Gaussian processes

- \{Z(s) : s \in D\} is a **Gaussian process** if the joint distribution of any collection of the RVs has a **multivariate normal (Gaussian)** distribution.

- In this case, the distribution is completely characterized by \(\mu_Z(\cdot)\) and \(C_Z(\cdot, \cdot)\).

- The joint probability density function of \(Z = (Z(s_1), \ldots, Z(s_n))^T\) at a finite set of locations is

\[
    f_Z(z) = (2\pi)^{-n/2} \det(\Sigma)^{-1/2} \times 
    \exp\left(-\frac{1}{2}(z - \mu)^T \Sigma^{-1}(z - \mu)\right),
\]

where \(\mu = (\mu(s_1), \ldots, \mu(s_n))^T\) and the \((j, k)\) element of the covariance matrix \(\Sigma\) is \(C_Z(s_j, s_k)\).

- If the Gaussian process \(\{Z(s) : s \in D\}\) is (weakly) stationary then the process is also strictly stationary with mean \(\mu_Z\) and covariance \(C_Z(\cdot)\).

References