Data Generation and Estimation for Axially Symmetric Processes on the Sphere

Chris Vanlangenberg[1]
Advisor: Dr. Haimeng Zhang[2]

University of North Carolina at Greensboro

August 19, 2016

[1]cdvanlan@uncg.edu
[2]h_zhang5@uncg.edu

Abstract

Review and update later
Global-scale processes and phenomena are of utmost importance in the geophysical sciences. Data from global networks and satellite sensors have been used to monitor a wide array of processes and variables, such as temperature, precipitation, etc. In this dissertation, we are planning to achieve explicitly the following objectives,

1. Develop both non-parametric and parametric approaches to model global data dependency.

2. Generate global data based on given covariance structure.

3. Develop kriging methods for global prediction.

4. Explore one or more of the popularly discussed global data sets in literature such as MSU (Microwave Sounding Units) data, the tropospheric temperature data from National Oceanic and Atmospheric and TOMS (Total Ozone Mapping Spectrometer) data, total column ozone from the Laboratory for Atmospheres at NASA's Goddard Space Flight Center Administration satellite-based Microwave Sounding Unit.

Global scale data have been widely studied in literature. A common assumption on describing global dependency is the second order stationarity. However, with the scale of the Earth, this assumption is in fact unrealistic. In recent years, researchers have focused on studying the so-called axially symmetric processes on the sphere, whose spatial dependency often exhibit homogeneity on each latitude, but not across the latitudes due to the geophysical nature of the Earth. In this research, we have obtained some results on the method of non-parametric estimation procedure, in particular, the method of moments, in the estimation of spatial dependency. Our initial result shows that the spatial dependency of axially symmetric processes exhibits both anti-symmetric and symmetric characteristics across latitudes. We will also discuss detailed methods on generating global data and finally we will outline our methodologies on kriging techniques to make global prediction.

Chapter 1

Introduction

In this chapter we have given a brief introduction to some of the basic concepts in spatial statistics which are necessary to follow other chapters in this dissertation. Moreover, we have discuss about stationarity, isotropy, intrinsic stationarity, covarince function and it properties, variogram, continuty and differentiability, spectral representations, Bochner's theorem, spectral densities, circulant matrices and it's properties with special cases.

1.1 Spatial random field

A real-valued spatial proces Z in d dimensions or a spatial random field can be denoted as $\{Z(x) : x \in D \subset \mathbb{R}^d\}$ where x is the location of process $Z(x)$ and x varies over the set D which is fixed and discrete. The distribution of the random vector $\mathbf{Z}(\mathbf{x}) = \{Z(x_1), \ldots, Z(x_n)\}$ is given by the associated finite-dimensional joint distributions

$$F\{Z(x_1), \ldots, Z(x_n)\} = P\{Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n\} \tag{1.1.1}$$

1.1.1 Stationary and Isotropy

A spatial random field is strict stationarity, for all finite n, $x_1, \ldots, x_n \in \mathbb{R}^d$, $h_1, \ldots, h_n \in \mathbb{R}$ and $x \in \mathbb{R}^d$, if the random field is invariant under translation. that is,

$$P\{Z(x_1 + x) \leq h_1, \ldots, Z(x_n + x) \leq h_n\} = P\{Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n\} \tag{1.1.2}$$

Strict stationary is a too strong condition as it involves the distibution of the random field but many spatial methods are based moments. Therefore, it is sufficeint to use weak assumptions and we could say a random process $Z(x)$ is weakly stationary if,

$$\begin{aligned} E(Z(x)) &= \mu \\ E^2(Z(x) &< \infty \\ C(h) &= Cov(Z(x), Z(x+h)) \end{aligned} \tag{1.1.3}$$

if $Z(x)$ has a finite second moment with constant mean and $C(h)$ the covarince (also referred as auto-covariance) function depends on the spatial distance only. Futher strictly stationary random fields with finite second moment is also weakly stationary, but weak stationarity does not imply strict stationarity. However, in the case of Gaussian random fields that weakly stationary are also strict stationarity because the first two moments (μ, σ) will explain the distribution.

Suppose $Z(x)$ is weakly stationary on $\mathbb{R}^d$ with autocavariance function $C(h)$ then it has the following properties,

(i) $C(0) \geq 0$

(ii) $C(h) = C(-h)$

(iii) $|C(h)| \leq C(0)$

(iv) If C_1, C_2 valid covariance functions then,

   (a) $C(x) = a_1 C_1 + a_2 C_2$, $\forall a_1, a_2 \geq 0$ is also a valid covariance function.
   (b) $C(x) = C_1(x) C_2(x)$ is also a valid covariance function.

If the variance between two locations soley depends on the distance between the two locations then the process is said to be intrinsically stationary. Semivariogram is an alternative to the covariance function proposed by Matheron. For an intrinsically stationary random field $Z(s)$,

$$\begin{aligned} E[Z(s)] &= \mu, \\ \gamma(h) &= \frac{1}{2} Var(Z(s+h) - Z(s)), \end{aligned} \tag{1.1.4}$$

Where γ is the semivariogram and $\gamma(h) = C(0) - C(h)$ for a weakly stationary process with covariance function $C(h)$. Intrinsic stationary is defined in terms of variogram and it is more general than weak stationary which is defined in terms of covariance. Clearly, when $C(h)$ is known we can get $\gamma(h)$ but not $C(h)$ when $\gamma(h)$ is known. For example consider a linear semi variogram function,

$$\gamma(h) = \begin{cases} a^2 + \sigma^2 h & h > 0 \\ 0 & otherwise \end{cases}$$

when $\underset{h \to \infty}{lim} \gamma(h) \to \infty$ thus this is not weak stationary and $C(h)$ does not exist.

A weakly stationary process with a covarince function $C(||h||)$ which is free from direction is called isoropic. The random field, $Z(x)$, on $\mathbb{R}^d$ is strictly isotropic if the joint distributions are invariant under all rigid motions. *i.e.,* for any orthogonal $d \times d$ matrix H and any $x \in \mathbb{R}^d$

$$P\{Z(Hx_1 + x) \leq h_1, \ldots, Z(Hx_n + x) \leq h_n\} = P\{Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n\} \tag{1.1.5}$$

Family	C(h)	Parameters	Validity
Matérn	$\frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\frac{h}{\phi})^{\nu}Y_{\nu}(\frac{h}{\phi})$	ν, σ^2, ϕ	R^3, S^2 when $\nu \leq 0.5$
Spherical	$\sigma^2(1 - \frac{3h}{2\phi} + \frac{1}{2}(\frac{h}{\phi})^3)I_{0\leq h\leq\phi}$	ϕ, σ^2	R^3, S^3
Exponential	$\sigma^2 exp\{-(h/\phi)\}$	ϕ, σ^2	R^3
Gaussian	$\sigma^2 exp\{-(h/\phi)^2\}$	ϕ, σ^2	R^3
Power	$\sigma^2(C_0 - (h/\phi)^{\alpha}$	ϕ, σ^2	$R^3\alpha \in [0,2], S^2\alpha \in (1,2]$

Table 1.1: Commonly used isotropic covariance models

Isotropy assumes that it is not required to distinguish one direction from another for the random field $Z(x)$.

A covariance function $C(\cdot)$ on $\mathbb{R}^d$ is positive definite if and only if

$$\sum_{i,j=1}^{N} a_i a_j C(x_i - x_j) \geq 0, \tag{1.1.6}$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$, and any locations $x_1, x_2, \ldots, x_N \in \mathbb{R}^d$.
Similarly, the variogram is conditionally negative definite

$$\sum_{i,j=1}^{N} a_i a_j 2\gamma(x_i - x_j) \leq 0, \tag{1.1.7}$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$ with $\sum a_i = 0$, and any locations $x_1, x_2, \ldots, x_N \in \mathbb{R}^d$.

1.1.2 Mean square continuity & differentiability

There is no simple relationship between $C(h)$ and the smoothness of $Z(x)$. For a sequence of random variables $X_1, X_2, \ldots$ and a random variable X defined on a common probability space. Define,$X_n \xrightarrow{L^2} X$ if, $E(X^2) < \infty$ and $E(X_n - X)^2 \to 0$ as $n \to \infty$. We can say, $\{X_n\}$ converges in L^2 if there exists such a X.

Suppose $Z(x)$ is a random field on $\mathbb{R}^d$, Then $Z(x)$ is mean square continuous at x if,

$$\lim_{h\to 0} E(Z(x+h) - Z(x))^2 = 0$$

If $Z(x)$ is weak stationary and $C(\cdot)$ is the covariance function then $E(Z(x+h) - Z(x))^2 = 2(C(0) - C(h))$. Therefore $Z(x)$ is mean square continuous iff $C(\cdot)$ is continuous at the origin.

1.1.3 Spectral methods

Sometimes it is convenient to use complex valued random functions, rather than real valued random functions.

We say, $Z(x) = U(x) + iV(x)$ is a complex random field if $U(x), V(x)$ are real random fields. If $U(x), V(x)$ are weakly stationary so does $Z(x)$.The covariance function can be defined as,

$$C(h) = cov(Z(x+h), \overline{Z(x)}), \quad C(-x) = \overline{C(x)},$$

for any complex constants $c_1, \ldots, c_n$, and any locations $x_1, x_2, \ldots, x_n$,

$$\sum_{i,j=1}^{n} c_i \bar{c_j} C(x_i - x_j) \geq 0 \tag{1.1.8}$$

1.1.4 Spectral representation of a random field

Suppose $\omega_1, \ldots, \omega_n \in \mathbb{R}^d$ and let $Z_1, \ldots, Z_n$ be mean zero complex random variables with $E(Z_i \bar{Z}_j) = 0, i \neq j \textit{ and } E|Z_i|^2 = f_i$. Then the random sum

$$Z(x) = \sum_{k=1}^{n} Z_k e^{i\omega_k^T x}. \tag{1.1.9}$$

Then $Z(x)$ given above is a weakly stationary complex random field in $\mathbb{R}^d$ with covariance function $C(x) = \sum_{k=1}^{n} f_k e^{i\omega_k^T x}$

Further, if we think about the integral as a limit in L^2 of the above random sum, then the covariance function can be represented as,

$$C(x) = \int_{\mathbb{R}^d} e^{i\omega^T x} F(d\omega) \tag{1.1.10}$$

where F is the so-called spectral distribution. There is a more general result from Bochner.

Theorem 1.1.1 (Bochner's Theorem)

A complex valued covariance function $C(\cdot)$ on $\mathbb{R}$ for a weakly stationary mean square continuous complex-valued random field on $\mathbb{R}^d$ iff it can be represented as above, where F is a positive measure.

If F has a density with respect to Lebesgue measure (spectral density) denoted by f, ($i.e.$ if such f exists) we can use the inversion formula to obtain f

$$f(\omega) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\omega^T x} C(x) dx \tag{1.1.11}$$

1.1.5 Septral densities

(i) Rational Functions that are even, nonnegative and integrable the corresponding covariance functions can be expressed in terms of elementary functions. For example if $f(\omega) = \phi(\alpha^2 + \omega^2)^{-1}$, then $C(h) = \pi\phi\alpha^{-1}e^{-\alpha|h|}$ (obtained by contour integration).

(ii) Gaussian are the most commonly used covariance function for a smooth process on $\mathbb{R}$ where the covariance function is given by $C(h) = ce^{-\alpha h^2}$ and the corresponding spectral density is $f(\omega) = \frac{1}{2\sqrt{\pi\alpha}}ce^{\frac{-\omega^2}{4\alpha}}$.

(iii) $Mat\acute{e}rn$ class has more practical use and more frequently used in spatial statistics. The spectral density of the form $f(\omega) = \frac{1}{\phi(\alpha^2+\omega^2)^{\nu+1/2}}$ where $\phi, \nu, \alpha > 0$ and the corresponding covariance function given by,

$$C(h) = \frac{\pi^{1/2}\phi}{2^{\nu-1}\Gamma(\nu+1/2)\alpha^{2\nu}}(\alpha|h|)^{\nu}Y_{\nu}(\alpha|t|) \tag{1.1.12}$$

where Y_ν is the modified Bessel function, the larger the ν smoother the Y. Further, Y will be m times square differntiable iff $\nu > m$. When ν is in the form of $m + 1/2$ with m a non negative integer, the spectral density is rational and the covariance function is in the form of $e^{-\alpha|h|}\cdot$ polynomial($|h|$)

$$\begin{aligned}
\nu = 1/2 \quad &: \quad C(h) = \pi\phi\alpha^{-1}e^{-\alpha|h|} \\
\nu = 3/2 \quad &: \quad C(h) = \frac{1}{2}\pi\phi\alpha^{-3}e^{-\alpha|h|}(1+\alpha|h|)
\end{aligned}$$

1.2 Circulant matrix

A square matrix $A_{n\times n}$ is a circulant matrix if the elements of each row (except first row) has the previous row shifted by one place to the right.

$$A = circ[a_o, a_1, \cdots, a_{n-1}] \quad = \quad \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{bmatrix}. \tag{1.2.1}$$

The eigenvalues of A are given by

$$\begin{aligned}
\lambda_l \quad &= \quad \sum_{k=0}^{n-1} a_k e^{-i2lk\pi/n} \\
&= \quad \sum_{k=0}^{n-1} c_k \rho_l^k, \qquad l = 0, 1, 2, \cdots, n-1,
\end{aligned}$$

where $\rho_l = e^{-i2\pi l/n}$ represents the lth root of 1), and the corresponding (unitary) eigenvector is given by

$$\psi_l = \frac{1}{\sqrt{n}}(1, \rho_l, \rho_l^2, \cdots, \rho_l^{n-1})^T.$$

If matrix A is real symmetric then its eigen values are real; for even $n = 2h$ the eigen values $\lambda_j = \lambda_{n-j}$ or there are either two eigen values or none with odd multipilicity, for odd $n = 2h - 1$ the eigen value λ_0 equal to any λ_j for $1 \leq j \leq h - 1$ or λ_0 occurs with odd multiplicity.

A square matrix B is Hermitian, if and only if $B^* = B$ where B^* is the complex conjugate. If B is real then $B^* = B^T$. Hermitian matrices has a full set of orthogonal eigen vectors with corresponsding real eigen values.

1.2.1 Block circulant matrices

should we talk about finding the inverse?

should we add any references in the introduction

Chapter 2

Literature Review (due August 28)

1. Axially symmetry, which means that a process is invariant to rotations about the Earth's axis. The idea was first proposed by Jones (1963), where the covariance function depends on the longitudes only through their difference.

2. In the study of a random process on a sphere, homogeneity (covariance depends solely on distance between locations) was assumed. However, this assumption may not be reasonable for actual data. Stein (2007) argued that Total Ozone Mapping Spectrometer (TOMS) data varies strongly with latitudes and homogeneous models are not suitable. Further, Cressie and Johannesson (2008), Jun and Stein (2008), Bolin and Lindgren (2011) pointed out that homogeneity assumption is not reasonable.

3. There are no methods to test axially symmetry in real data. However, this assumption is more plausible and reasonable when modeling spatial data. For example, temperature, moisture, etc. most likely symmetric on longitudes rather than latitudes. Stein (2007) propose a method to model axially symmetric process on a sphere (the fitted model is not the best, but this was a good start).

4. There are no practically useful parametric models available, for our knowledge only models available so far, Stein (1999) with 170 parameters to estimate and Cressie and Johannesson (2008) more than 396 parameters to estimate.

5. When modeling spatial data stationary models are less useful;Jun and Stein (2008) has proposed flexible class of parametric covariance models to capture the non-stationarity of global data. They used Discrete Fourier Transform (DFT) to the data on regular grids and calculated the exact likelihood for large data sets. Furthermore, they used Legendre polynomials to remove the spatial trends when fitting models to global data.

6. Lindgren et al. (2011) analyzed global temperature data with a non-stationary model defined on a sphere using Gaussian Markov Random Fields (GMRF) and Stochastic Partial Differential Equations (SPDE)

7. Monte Carlo Markov Chain (MCMC) is another approach to model non-stationary covariance models on a sphere. Bolin and Lindgren (2011) (continuation of the work proposed in Lindgren et al. (2011)) constructed a class of stochastic field models using

Stochastic Partial Differential Equations (SPDEs). Non stationary covariance models were obtained by spatially varying the parameters in the SPDEs, they argue that this method is more efficient than standard MCMC procedures. There are many articles followed this techniques but we will not discuss more details about these methods.

8. Spatio-temporal mixed-effects model for dimension reduction was proposed by Katzfuss and Cressie (2011). They used MOM parameter estimation method (similar approach in FRS). This work is also based on Cressie and Johannesson (2008) spatial only Fixed Rank model. These methods are eventually focused on Bayesian approach and are less interested about topic.

9. The previous studies have argued that many processes on a sphere are not homogeneous, especially in latitude direction. Huang et al. (2012) proposed a class of statistical processes that are axially symmetric and covariance functions that depend on longitudinal differences. Moreover, they have proposed longitudinally reversible processes and some motivations to construct axially symmetric processes. The covariance models implemented in this dissertation are modified versions of the covariance models proposed by Huang et al. (2012).

10. Hitczenko and Stein (2012) discuss about the properties of an existing class of models for axially symmetric Gaussian processes on the sphere. They applied first-order differential operators to an isotropic process. draw conclusions about the local properties of the processes. Under some restrictions they derived explicit forms for the spherical harmonic representation of these processes covariance functions, and make conclusions about the local properties of the processes.

11. The issues associated when modeling axially symmetric spatial random fields on a sphere was discussed by Li (2013). They proposed convolution methods to generate random fields with a class of $Mat\acute{e}rn$-type kernel functions by allowing the parameters in the kernel function to vary with location. Moreover, they were able to generate flexible class of covariance functions and capture the non-stationary properties on a sphere. Used FFT to get the determinant and the inverse efficiently. Further, semi-parametric variogram estimation method using spectral representation was proposed for intrinsically stationary random fields on S^2.

12. $Mat\acute{e}rn$ covariance models are widely used when modeling spatial data, but when the smoothness parameter (ν) is greater than 0.5 it is not valid for the homogeneous processes on the Earth surface with great circle distance. Jeong and Jun (2015) proposed $Mat\acute{e}rn$-like covariance functions for smooth processes on the earth surface that are valid with great circle distance (models were tested on sea levels pressure data).

Family	C(h)	Parameters	Validity
$Matérn$	$\frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\frac{h}{\phi})^\nu Y_\nu(\frac{h}{\phi})$	ν, σ^2, ϕ	R^3, S^2 when $\nu \leq 0.5$
Spherical	$\sigma^2(1 - \frac{3h}{2\phi} + \frac{1}{2}(\frac{h}{\phi})^3)I_{0\leq h\leq\phi}$	ϕ, σ^2	R^3, S^3
Exponential	$\sigma^2 exp\{-(h/\phi)\}$	ϕ, σ^2	R^3
Gaussian	$\sigma^2 exp\{-(h/\phi)^2\}$	ϕ, σ^2	R^3
Power	$\sigma^2(C_0 - (h/\phi)^\alpha$	ϕ, σ^2	$R^3\alpha \in [0,2], S^2\alpha \in (1,2]$

Table 2.1: Commonly used covariance and variogram models

Chapter 3

Covariance and Variogram Estimation on the Cricle

3.1 Stationary process on a circle

Let $C(\theta), \theta \in (0, \pi)$ denote a stationary covariance function on the circle, then

$$C(\theta) = \sum_{n=0}^{\infty} a_n \cos(n\theta), \tag{3.1.1}$$

with $\sum_{n=0}^{\infty} a_n < \infty$, and $a_n \geq 0$. Note that

$$a_n = \frac{2}{\pi} \int_0^{\pi} C(\theta) \cos(n\theta) d\theta. \tag{3.1.2}$$

Now if a random process on the circle, with continuity in the quadratic means sense, can be represented as

$$X(t) = \sum_{n=0}^{\infty} (A_n \cos(nt) + B_n \sin(nt), \quad t \in (0, 2\pi). \tag{3.1.3}$$

Note that if $X(t)$ is stationary on the circle with covariance function $C(\theta)$, then

$$cov(A_n, A_m) = a_n \delta(n, m) = cov(B_n, B_m), \quad \text{for } n, m \geq 0. \tag{3.1.4}$$

Let $\{X(t_k), k = 1, 2, \cdots, n\}$ be a collection of gridded observations on a circle, with $t_k = (k-1) * 2\pi/n, k = 1, 2, \cdots, n$. Lets assume $E(X(t)) = \mu$ is unknown, the unbiased estimator of μ is given by $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X(t_i)$. The underlying process is stationary, if it's covariance function solely depends on the distance θ,

$$C(\theta) = cov(X(t + \theta), X(t)), \qquad \theta \in [0, \pi]. \tag{3.1.5}$$

3.2 Estimation

3.2.1 Estimation of covaraince on a cricle

We used method of moments (MOM) to estimate the covariance $C(\theta)$ on a circle, the estimator can be given by

$$\hat{C}(\Delta\lambda) = \frac{1}{n}\sum_{i=1}^{n}(X(t_i+\Delta\lambda)-\bar{X})(X(t_i)-\bar{X}), \tag{3.2.1}$$

where $\Delta\lambda = 0, 2\pi/n, 4\pi/n, \cdots, 2(N-1)\pi/n$.

Now we will show that above estimator is not unbiasedness.

$$\begin{aligned}
E(\hat{C}(\Delta\lambda)) &= \frac{1}{n}\sum_{i=1}^{n}E((X(t_i+\Delta\lambda)-\bar{X})(X(t_i)-\bar{X})) \\
&= \frac{1}{n}\sum_{i=1}^{n}E((X(t_i+\Delta\lambda)-\mu-(\bar{X}-\mu))(X(t_i)-\mu-(\bar{X})-\mu)) \\
&= \frac{1}{n}\sum_{i=1}^{n}cov(X(t_i+\Delta\lambda),X(t_i)) - \frac{1}{n}\sum_{i=1}^{n}E((X(t_i+\Delta\lambda)-\mu)(\bar{X}-\mu)) \\
&\quad -\frac{1}{n}\sum_{i=1}^{n}E((X(t_i)-\mu)(\bar{X}-\mu)) + \frac{1}{n}\sum_{i=1}^{n}E((\bar{X}-\mu)(\bar{X}-\mu)) \\
&= C(\Delta\lambda) - E((\bar{X}-\mu)(\bar{X}-\mu)) - E((\bar{X}-\mu)(\bar{X}-\mu)) + E((\bar{X}-\mu)(\bar{X}-\mu)) \\
&= C(\Delta\lambda) - var(\bar{X}).
\end{aligned} \tag{3.2.2}$$

Suppose the variance-covariance matrix of the sample vector $\underset{\sim}{X} = (X_1, X_2, \cdots, X_n)^T$ is given by Σ. Further, we can denote $\bar{X}$ in the following form,

$$\bar{X} = \frac{1}{n}\mathbf{1}_n^T\underset{\sim}{X}$$

then

$$\begin{aligned}
var(\bar{X}) &= cov(\frac{1}{n}\mathbf{1}_n^T\underset{\sim}{X}, \frac{1}{n}\mathbf{1}_n^T\underset{\sim}{X}) \\
&= \frac{1}{n^2}\mathbf{1}_n^T\Sigma\mathbf{1}_n \\
&= \frac{1}{n}\left(C(0)+C(\pi)+2\sum_{m=1}^{N-1}C(m2\pi/n)\right)
\end{aligned}$$

When $n\to\infty$, (assuming $C(\theta)$ is a continuous function on $[0,\pi]$)

$$\frac{1}{n}\left(2\sum_{m=0}^{N}C(m2\pi/n)\right) = \frac{1}{\pi}\frac{\pi}{N}\left(\sum_{m=0}^{N}C(m2\pi/n)\right) \to \frac{1}{\pi}\int_0^{\pi}C(\theta)d\theta = a_0.$$

Hence

$$var(\bar{X}) = \frac{1}{n}\left(2\sum_{m=0}^{N} C(m2\pi/n)\right) - \frac{1}{n}(C(0) + C(\pi)) \to a_0, \qquad \text{as } n \to \infty.$$

We can conclude that

$$var(\bar{X}) \to \frac{1}{\pi}\int_0^{\pi} C(\theta)d\theta \quad \text{as } n \to \infty.$$

Now, we will show that $\bar{X}$ will never be a consistent estimator for μ, mean on a circle.

$$P(|\bar{X} - \mu| > \varepsilon) \to 0.$$

If $var(\bar{X}) \to 0$, then from Chebyshev's inequality we have

$$P(|\bar{X} - \mu| > \varepsilon) \leq \frac{var(\bar{X})}{\varepsilon^2} \to 0, \quad \text{for any } \varepsilon > 0.$$

Therefore, $var(\bar{X}) \to 0$ is a sufficient condition for consistency, but it is not necessary. However, if we assume $\underset{\sim}{X}$ is multivariate normally distributed, then $\bar{X}$ follows normal distribution with mean μ and approximate variance a_0. Then $(Z \sim N(0,1))$

$$P(|\bar{X} - \mu| > \varepsilon) = P\left(\frac{|\bar{X} - \mu|}{\sqrt{a_0}} > \frac{\varepsilon}{\sqrt{a_0}}\right) \approx P\left(|Z| > \frac{\varepsilon}{\sqrt{a_0}}\right) \not\to 0$$

since $\frac{\varepsilon}{\sqrt{a_0}}$ is a fixed constant for each fixed $\varepsilon > 0$.

That is, the MOM estimator $\hat{C}(\Delta\lambda)$ of the covariance function is actually a biased estimator with the shift amount of a_0. Therefore, if $a_0 = 0$ for a covariance function, we have the unbiased estimator $\hat{C}(\Delta\lambda)$.

If the gridded points were on a line, for example in time series, $E(\bar{X} - \mu)^2 \to 0$ as $n \to \infty$ under the assumption that the covariance function $C(\theta) \to 0$ when $\theta \to \infty$ (which is practically feasible), that is, $\bar{X}$ is consistent in the case of points on a line. In the case of circle, we might not have $C(\theta)$ close to 0 since θ is within a bounded region ($(0, \pi)$ for the circle) and we normally assume $C(\theta)$ is continuous for θ.

consistency of the cross covarince estimator?

3.2.2 Estimation of variogram on a circle

The theoretical variogram function is given by,

$$\gamma(\theta) = C(0) - C(\theta). \tag{3.2.3}$$

and the MOM estimator for the variogram is given by,

$$\hat{\gamma}(\Delta\lambda) = \frac{1}{2n}\sum_{i=1}^{n}(X(t_i + \Delta\lambda) - X(t_i))^2. \tag{3.2.4}$$

We can show that variogram estimator through MOM is an unbiased estimator,

$$
\begin{aligned}
E(\hat{\gamma}(\Delta\lambda)) &= \frac{1}{2n}\sum_{i=1}^{n} E(X(t_i+\Delta\lambda) - X(t_i))^2 \\
&= \frac{1}{2n}\sum_{i=1}^{n} E((X(t_i+\Delta\lambda) - \mu) - (X(t_i) - \mu))^2 \\
&= \frac{1}{2n}\sum_{i=1}^{n} cov(X(t_i+\Delta\lambda) - X(t_i), X(t_i+\Delta\lambda) - X(t_i)) \\
&= \frac{1}{2n}\sum_{i=1}^{n} \{cov(X(t_i+\Delta\lambda), X(t_i+\Delta\lambda)) + cov(X(t_i), X(t_i)) \\
&\quad -2cov(X(t_i+\Delta\lambda), X(t_i))\} \\
&= \frac{1}{2n}\sum_{i=1}^{n} (C(0) + C(0) - 2C(\Delta\lambda)) \\
&= C(0) - C(\Delta\lambda) = \gamma(\Delta\lambda).
\end{aligned}
$$

need to prove consistency

3.3 Data generation on a circle

First, we will discuss how to generate correlated data at n grided points on a circle when the covarince function is defined and compare above covariance and variogram estimators. Since the observed data are correlated, the covaince function can be written as a function of distance (angle). For similicity we will use exponential family covarince function as given below,

$$C(\theta) = C_1 e^{-a|\theta|}, \tag{3.3.1}$$

where $\theta = i * \Delta\lambda = \pm i * 2\pi/n, i = 1, 2, \cdots, n/2$

Clearly, each location is correlated with other $n-1$ locations and $C(\theta) = C(-\theta)$ the variance-covariance matrix Σ is circulant and will be in the following form,

$$
\begin{aligned}
\Sigma &= \begin{pmatrix}
C(0) & C(2\pi/n) & \cdots & C((N-1)2\pi/n) & C(\pi) & C((N-1)2\pi/n) & \cdots & C(2\pi/n) \\
C(2\pi/n) & C(0) & \cdots & C((N-2)2\pi/n) & C((N-1)2\pi/n) & C(\pi) & \cdots & C(4\pi/n) \\
C(4\pi/n) & C(2\pi/n) & \cdots & C((N-3)2\pi/n) & C((N-2)2\pi/n) & C((N-1)2\pi/n) & \cdots & C(6\pi/n) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
C(2\pi/n) & C(4\pi/n) & \cdots & C(\pi) & C((N-1)2\pi/n) & C((N-2)2\pi/n) & \cdots & C(0)
\end{pmatrix} \\
&= circ(C(0), C(2\pi/n), C(4\pi/n), \cdots, C((N-1)2\pi/n), C(\pi), C((N-1)2\pi/n), \cdots, C(2\pi/n)). \\
&= Q\Lambda Q^T,
\end{aligned}
$$

where $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ and $Q = \{\psi_1, \psi_2, \cdots, \psi_n\}$ are the respective eigen values and eigen vectros of the above circulant matrix. Now using singular value decomposssion (SVD) we can obtain the correlated data {X(t)} on a circle as follows,

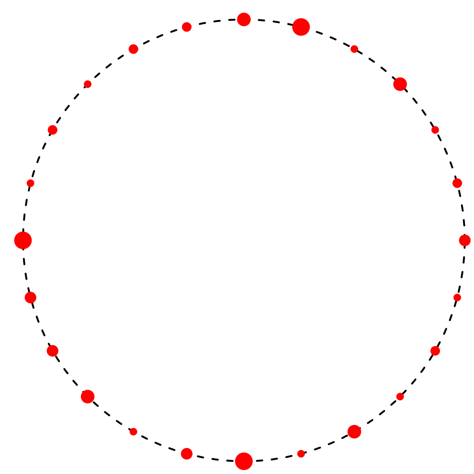

Figure 3.1: Random process on a circle at 24 points ($\Delta\lambda = 30^0$), the red dots represent the observed values at a given time and each red dot is a random process of it own.

$$X(t) = \Sigma^{1/2} * Z = Q\Lambda^{1/2}Q^T * Z$$

where $Z \sim N(\underset{\sim}{0}, 1_n)$.

3.3.1 Compare covarince estimator

- Using the exponential covariance function given by 3.3.1

  In section 3.2.1 we proved that, in general the covarince estimator (3.2.1) on a circle is biased, with a bias of $var(\bar{X})$. In order to compare this estiamtor to it's theoretical covarince given by equation 3.3.1. We computed the MOM estimator $\hat{C}(\theta)$ with 48 gridded observations on the circle from 500 simulations.

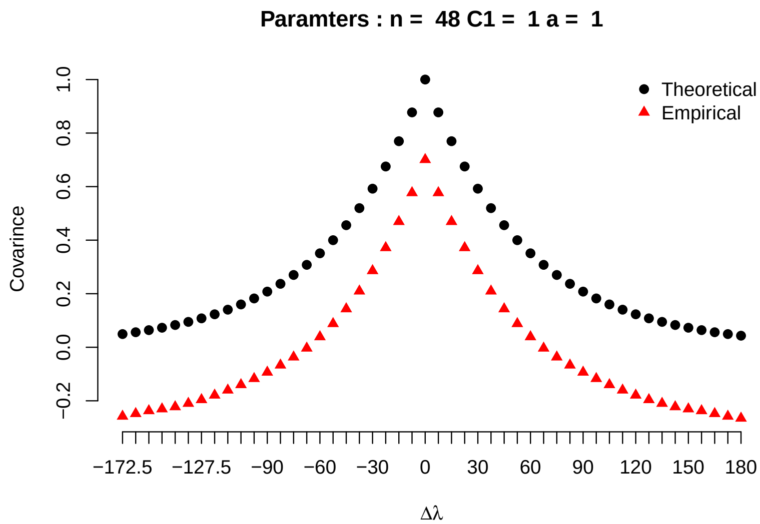

Figure 3.2: Theoretical and empirical covariance (with bias) comparison on a circle

We have noticed that the shift between theoretical and empirical values were approximately equal to a_0.

should we talk about $\hat{var}(\bar{X}) = var(\bar{X}_1, \cdots, \bar{X}_k)$*??*, $a_0 = 0.3045545$ $\hat{var}(\bar{X})$ *will be close to* a_0 *as we increase the number of simulations*

- However, we can obtain a_0 for the above exponential covarince as follows,

$$a_0 = \frac{C_1}{a\pi}(1 - e^{-a\pi})$$

We consider the following covariance function, after subtracting a_0 from $C(\theta)$. The theoretical and empirical values match perfectly.

$$D(\theta) = C(\theta) - a_0.$$

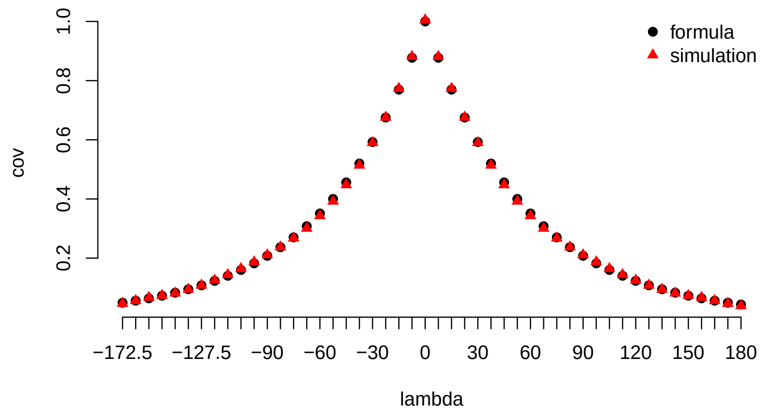

Figure 3.3: Theoretical and empirical covariance comparison on a circle

Note, if the process is a zero mean process the covariance estimator will be unbiased ($i.e. Var(\bar{X}) = 0$) hence we will get a perfect match between theoretical and empirical values.

3.3.2 Compare variogram estimator

We proved that in general the variogram estimator is unbiased. The theoretical variogram based on exponential covariance is given below,

$$\gamma(\theta) = C(0) - C(\theta) = C_1(1 - e^{-a|\theta|})$$

Again we computed the variogram estimator $\hat{\gamma}(\theta)$ with 48 gridded observations on the circle from 4000 simulations and there is a perfect match between theoretical and empirical values.

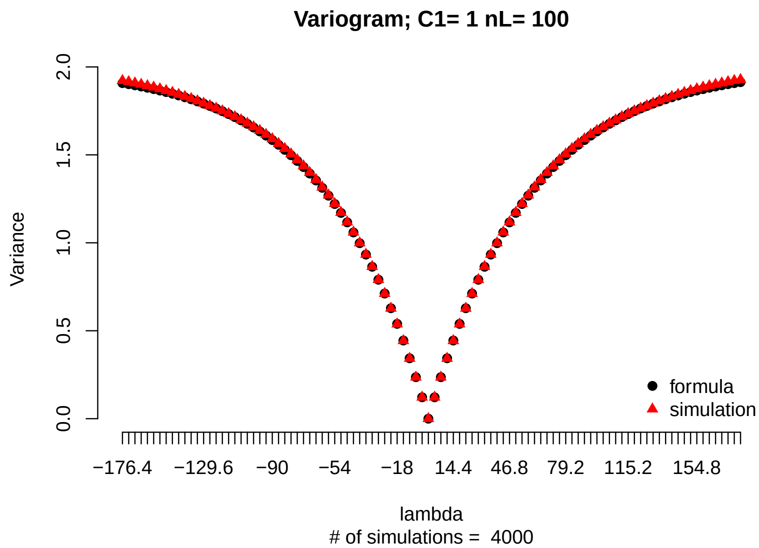

Figure 3.4: Theoretical and empirical comparison for variogram on a circle

should make any argument why we need more simulations

Chapter 4

Random Process on a Sphere (due August 21)

A random process is a collection of random variables $X \in \{X(s) : s \in D\}$, defined in a common probability space. In general, if

- $s \in N$: random sequence, time series
- $s \in R^1$: random process which is a stochastic process
- $s \in R^d$: random filed, spatial process if $d > 1$
- $s \in S^2$: random process on the sphere
- $s \in R^d \times R$: spatio-temporal process

In this dissertation, we will focus only on spatial processes on a sphere, where s represents a location.

4.1 Random process on a sphere

We consider a complex-valued random process $X(P)$ on a unit sphere S^2, where $P = (\lambda, \phi) \in S^2$ with longitude $\lambda \in [-\pi, \pi)$ and latitude $\phi \in [0, \pi]$. Assume the process is continuous in quadratic mean with respect to the location P, and has finite second moment, then $X(P)$ can be represented by spherical harmonics, with convergence of the series in quadratic mean, Li and North (1997).

$$X(P) = X(\lambda, \phi) = \sum_{\nu=0}^{\infty} \sum_{m=-\nu}^{\nu} Z_{\nu,m} e^{im\lambda} P_\nu^m(\cos\phi),$$

where $P_\nu^m(\cdot)$ is a normalized associated Legendre polynomial so that its squared integral on $[-1, 1]$ is 1, and $Z_{\nu,m}$ are the coefficients satisfying

$$Z_{\nu,m} = \int_{S^2} X(P) e^{-im\lambda} P_\nu^m(\cos\phi) dP.$$

Without loss of generality, the process is assumed to have mean zero, i.e., $\mathrm{E}(X(P)) = 0$, which implies $\mathrm{E}(Z_{\nu,m}) = 0$. Then, the covariance function of the process at two locations $P = (\lambda_P, \phi_P)$ and $Q = (\lambda_Q, \phi_Q)$ is given by,

$$R(P,Q) = \mathrm{E}(X(P)\overline{X(Q)}) = \sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty}\sum_{m=-\nu}^{\nu}\sum_{n=-\mu}^{\mu} \mathrm{E}(Z_{\nu,m}\overline{Z}_{\mu,n})e^{im\lambda_P}P_\nu^m(\cos\phi_P)e^{-in\lambda_Q}P_\mu^n(\cos\phi_Q).$$

where $\bar{Z}$ denotes the complex conjugate of Z. Note that the continuity of $X(P)$ on every point P implies that $R(P,Q)$ is continuous on all pairs (P,Q) (Leadbetter, 1967, page 83). If the covariance function depends solely on the spherical distance between these two locations, the process is homogeneous. That is, (Obukhov, 1947, Yaglom et al. (1961))

$$R(P,Q) = R(\theta(P,Q)) = \sum_{\nu=0}^{\infty} \frac{(2\nu+1)f_\nu}{2} P_\nu(\cos\theta(P,Q)),$$

where the spherical distance $\theta(P,Q) = \cos^{-1}(\cos\phi_P\cos\phi_Q + \cos\phi_P\cos\phi_Q\cos(\lambda_P - \lambda_Q))$, $P_\nu(\cdot)$ is the Legendre polynomial of order ν, $f_\nu \geq 0$, and $\sum_{\nu=0}^{\infty}(2\nu+1)f_\nu < \infty$. Here, the random variable $Z_{\nu,m}$ satisfies

$$\mathrm{E}(Z_{\nu,m}\overline{Z}_{\mu,n}) = \delta_{\nu,\mu}\delta_{n,m}f_\nu,$$

where $\delta_{a,b} = 1$ if $a = b$, and 0 otherwise.

4.2 Axially symmetry

The idea was introduced by Jones (1963), if the covarince between two spatial points depends only on the longitudes only through their difference between two points then process is said to be axially symmetric.

Under the assumption of axial symmetry , where the covariance function depends on longitudinal differences, one has

$$\mathrm{E}(Z_{\nu,m}\overline{Z}_{\mu,n}) = \delta_{n,m}f_{\nu,\mu,m}.$$

Hence, the covariance function is of the form

$$R(P,Q) = R(\phi_P, \phi_Q, \lambda_P - \lambda_Q) = \sum_{m=-\infty}^{\infty}\sum_{\nu=|m|}^{\infty}\sum_{\mu=|m|}^{\infty} f_{\nu,\mu,m}e^{im(\lambda_P-\lambda_Q)}P_\nu^m(\cos\phi_P)P_\mu^m(\cos\phi_Q). \quad (4.2.1)$$

Further conditions on $f_{\nu,\mu,m}$ are imposed in order to have the covariance function valid. In particular, $f_{\nu,\mu,m} = \overline{f}_{\mu,\nu,m}$ and for each fixed integer m, the matrix $F_m(N) = \{f_{\nu,\mu,m}\}_{\nu,\mu=|m|,|m|+1,\ldots,N}$ must be positive definite for all $N \geq |m|$. A detailed discussion of parallel conditions on $f_{\nu,\mu,m}$ under the real-valued case is given in Jones (1963).

Chapter 5

Global Data Generation on the Sphere (due August 21)

5.1 Circularly-symmetry Gaussian random vectors

In general, a normal family has two parameters, location parameter μ and scale parameter Σ. But when we are dealing with complex normal family there is one additional parameter, the relation matrix also referred as pseudo-covariance matrix (for real normal family pseudo-covariance matrix is equivalent to the covariance matrix).

Following the notes provided by Gallager (2008), a complex random variable $Z = Z^{Re} + iZ^{Im}$ is Gaussian, if Z^{Re}, Z^{Im} both are real and jointly Gaussian. Then Z is circularly-symmetric if both Z and $e^{i\phi}Z$ has the same probability distribution for all real ϕ. Since $E[e^{i\phi}Z] = e^{i\phi}E[Z]$, any circularly-symmetric complex random vector must have $E[Z] = 0$, in other words its mean must be zero.

Let the covariance matrix K_Z and the pseudo-covariance matrix M_Z of a zero mean $2n$ complex random vector $\underset{\sim}{Z} = (Z_1, Z_2, \ldots, Z_n)^T$, where $Z_j = (Z_j^{Re}, Z_j^{Im})^T$ and $j = 1, 2, \ldots, n$ can be defined as follows,

$$K_{\underset{\sim}{Z}} = E[\underset{\sim}{Z}\underset{\sim}{Z}^*] \quad M_{\underset{\sim}{Z}} = E[\underset{\sim}{Z}\underset{\sim}{Z}^T],$$

where $\underset{\sim}{Z}^*$ is the conjugate transpose of $\underset{\sim}{Z}$.

The covariance matrix of real $2n$ random vector $\underset{\sim}{Z} = (\underset{\sim}{Z}^{Re}, \underset{\sim}{Z}^{Im})^T$ is determined by both $K_{\underset{\sim}{Z}}$ and $M_{\underset{\sim}{Z}}$ as follows,

$$\begin{aligned}
E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Re}] &= \frac{1}{2}Re(K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}}), \\
E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Im}] &= \frac{1}{2}Re(K_{\underset{\sim}{Z}} - M_{\underset{\sim}{Z}}), \\
E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Im}] &= \frac{1}{2}Im(-K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}}), \\
E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Re}] &= \frac{1}{2}Im(K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}})
\end{aligned} \tag{5.1.1}$$

We can get the covariance of $\underset{\sim}{Z} = (\underset{\sim}{Z}^{Re}, \underset{\sim}{Z}^{Im})^T$ as follows,

$$\begin{aligned} Cov(\underset{\sim}{Z}) &= E(\underset{\sim}{Z}\underset{\sim}{Z}^T) \\ &= \begin{pmatrix} E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Re}] & E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Im}] \\ E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Re}] & E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Im}] \end{pmatrix} \end{aligned}$$

Theorem 5.1.1 (Gallager, 2008)
Let $\underset{\sim}{Z}$ be a zero mean Gaussian random vector then $M_{\underset{\sim}{Z}} = 0$ if and only if $\underset{\sim}{Z}$ is circularly-symmetric.

5.1.1 Theoretical development

Let $X(P)$ be a complex-valued random process on a unit sphere S^2, where $P = (\lambda, \phi) \in S^2$ with longitude $\lambda \in [-\pi, \pi)$ and latitude $\phi \in [0, \pi]$.

A covariance function for continuous axially symmetric processes on a sphere given by (Huang et al., 2012, proposition 1):

$$R(P, Q) = R(\phi_P, \phi_Q, \Delta\lambda) = \sum_{m=-\infty}^{\infty} e^{im\Delta\lambda} C_m(\phi_P, \phi_Q) \qquad (5.1.2)$$

where $\Delta\lambda \in [-\pi, \pi]$, and $C_m(\phi_P, \phi_Q)$ is Hermitian and *p.d.* with $\sum_{-\infty}^{\infty} |C_m(\phi_P, \phi_Q)| < \infty$.

One can derive C_m based on an axially symmetric covariance function $R(P, Q)$ defined on a sphere, as we have

$$C_m(\phi_P, \phi_Q) = \frac{1}{2\pi} \int_{-\pi}^{\pi} R(\phi_P, \phi_Q) e^{-im\Delta\lambda} d\Delta\lambda$$

since $C_m(\phi_P, \phi_Q)$ is continuous and both Hermitian and positive definite, by Mercer's theorem, there exists an orthonormal basis $\{\psi_{m,\nu}, \nu = 0, 1, \ldots\}$ in L^2, a complex-valued functional Hilbert space on $[o, \pi]$, such that

$$C_m(\phi_P, \phi_Q) = \sum_{\nu=0}^{\infty} \eta_{m,\nu} \psi_{m,\nu}(\phi_P) \overline{\psi_{m,\nu}(\phi_Q)},$$

Where $\eta_{m,\nu} \geq 0$ are the eigen values and $\psi_{m,\nu}(\cdot)$ are the eigen functions.

Further, according to Huang et al. (2012)[remark 2.5] a continuous axially symmetric process, $X(P)$, is given as:

$$X(P) = X(\phi, \lambda) = \sum_{m=-\infty}^{\infty} W_{m,\nu}(\phi) e^{im\lambda} \psi_{m,\nu}(\phi), \qquad (5.1.3)$$

where λ is the longitude, ϕ is the latitude and $\psi_{m,\nu}(\cdot)$ is a orthonormal basis. When the process is real and Gaussian, $W_{m,\nu}$ are independent normal random variables. In addition, this process can be viewed as a homogeneous random process on the circle with angular distance given by $\Delta\lambda$. That is, for each ϕ, one can expand $X(P)$ in a Fourier series that is convergent in quadratic mean Roy (1972):

$$X(\phi, \lambda) = \sum_{m=-\infty}^{\infty} W_m(\phi) e^{im\lambda} \tag{5.1.4}$$

,

where

$$W_m(\phi) = \frac{1}{2\pi} \int_0^{2\pi} X(\phi, \lambda) e^{-im\lambda} d\lambda,$$

with $\mathrm{E}(W_m(\phi_P)\overline{W_n(\phi_Q)}) = \delta_{m,n} C_m(\phi_P, \phi_Q)$.

5.1.2 Generalization of parametric models

The $R(P,Q)$ given in equation 5.1.2, is clearly a function of both longitude and latitude. The simplest model is the separable model where,

$$R(P,Q) = \tilde{C}(\Delta\lambda) C_m(\phi_P, \phi_Q)$$

In order to make things easier one could assume that $C_m(\phi_P, \phi_Q) = \tilde{C}_m(\phi_P - \phi_Q)$ only depends on the difference of ϕ_P and ϕ_Q, Huang et al. (2011) proposed a simple separable covariance function when both covariance components are exponential

$$R(P,Q) = c_0 e^{-a|\Delta\lambda|} e^{-b|\phi_P - \phi_Q|}.$$

Lets assume,

$$C_m(\phi_P, \phi_Q) = c_m e^{-a_m|\phi_P - \phi_Q|}(\cos\omega_m(\phi_P - \phi_Q) + i \sin\omega_m(\phi_P - \phi_Q)), \quad C_m \geq 0, a_m \geq 0, \omega_m \in R.$$

(Huang et al., 2012, Remark 2.4)states that for real-valued process, the covariance function $R(P,Q)$ is also real-valued, and

$$R(P,Q) = C_{0,R}(\phi_P, \phi_Q) + 2\sum_{m=1}^{\infty} cos(m\Delta\lambda) C_{m,R}(\phi_P, \phi_Q) - sin(m\Delta\lambda) C_{m,I}(\phi_P, \phi_Q),$$

where $C_m(\phi_P, \phi_Q)$=$C_{m,R}(\phi_P, \phi_Q)$+$iC_{m.I}(\phi_P, \phi_Q)$.
$C_{0,R}(\phi_P, \phi_Q)$=$c_0 e^{-a_0|\phi_P - \phi_Q|} \cos\omega_0(\phi_P - \phi_Q)$, $C_{m,R}(\phi_P, \phi_Q)$=$c_m e^{-a_m|\phi_P - \phi_Q|} \cos\omega_m(\phi_P - \phi_Q)$, $C_{m,I}(\phi_P, \phi_Q)$=$c_m e^{-a_m|\phi_P - \phi_Q|} \sin\omega_m(\phi_P - \phi_Q)$

If one takes $a_m = a$, $\omega_m = mu$ we can get the following form for $R(P,Q)$,

$$R(P,Q) = c_0 e^{-a|\phi_P - \phi_Q|} + 2e^{-a|\phi_P - \phi_Q|} \sum_{m=1}^{\infty} c_m \cos[m\theta(P,Q,u)],$$

where $\theta(P,Q,u) = \Delta\lambda + u(\phi_P - \phi_Q) - 2k\pi$, and k is chosen such that $\theta(P,Q,u) \in [0, 2\pi]$.

Moreover, by carefully choosing functions for $C_m(\phi_P, \phi_Q)$ Huang et al. (2012) proposed some nonseparable covariance models $(R(P,Q))$ models valid on the sphere,

$$
\begin{aligned}
R(P,Q) &= Ce^{-a|\phi_P-\phi_Q|}\frac{1-p^2}{1-2p\cos\theta(P,Q,u)+p^2} && (5.1.5)\\
R(P,Q) &= Ce^{-a|\phi_P-\phi_Q|}\log\frac{1}{(1-2p\cos\theta(P,Q,u)+p^2)} && (5.1.6)\\
R(P,Q) &= 2Ce^{-a|\phi_P-\phi_Q|}\left(\frac{\pi^4}{90}-\frac{\pi^2\theta^2(P,Q,u)}{12}+\frac{\pi\theta^3(P,Q,u)}{12}-\frac{\theta^4(P,Q,u)}{48}\right). && (5.1.7)
\end{aligned}
$$

There is one big disadvantage for all of them. They are assumed not only stationarity on longitudes, but stationarity on latitudes as well.

Modifying the covariance models to include non-stationarity

1. We have noticed that when $\phi_P = \phi_Q$, the first model reduces to

$$R(P,Q) = C\frac{1-p^2}{1-2p\cos(\Delta\lambda)+p^2},$$

   and if $\Delta\lambda = 0$, the variance over all latitudes would be constant. This is not supposed to be the case, since both MSU data and TOMS data in figures **??** and **??** shows that variance is highly depending on the latitude.

2. A modification of the above approach is to replace the function

$$C(\phi_P,\phi_Q) = Ce^{-a|\phi_P-\phi_Q|}$$

   by a non-stationary covariance function, which depends on the latitudes, even when $\phi_P = \phi_Q$. Here are the two functions that have been used in our analysis.

$$
\begin{aligned}
\tilde{C}(\phi_P,\phi_Q) &= C_1(C_2 - e^{-a|\phi_P|} - e^{-a|\phi_Q|} + e^{-a|\phi_P-\phi_Q|}),\\
\tilde{C}(\phi_P,\phi_Q) &= C_1\left(C_2 - \frac{1}{\sqrt{a^2+\phi_P^2}} - \frac{1}{\sqrt{a^2+\phi_Q^2}} + \frac{1}{\sqrt{a^2+(\phi_P-\phi_Q)^2}}\right).
\end{aligned}
$$

   Here $C_1, a > 0$, and $C_2 \geq 1$ to ensure the positive definiteness of the above function. When $\phi_P = \phi_Q$, both functions are actually the function of ϕ_P.

$$
\begin{aligned}
\tilde{C}(\phi_P,\phi_P) &= C_1(C_2 - 2e^{-a|\phi_P|} + 1),\\
\tilde{C}(\phi_P,\phi_P) &= C_1\left(C_2 - \frac{2}{\sqrt{a^2+\phi_P^2}} + \frac{1}{a}\right).
\end{aligned}
$$

3. A more general non-stationary covariance function is given as following. If $C(\cdot) = C(x-y)$ is the stationary covariance function and $f(\omega) \geq 0$ is the corresponding spectral density, then

Proposition 5.1.1 *A more general non stationary covariance function is given as following. If $C(\cdot) = C(x-y)$ is the stationary covariance function and $f(\omega) \geq 0$ is the corresponding spectral density, then*

$$\tilde{C}(x,y) = C_2 - C(x) - C(y) + C(x-y), \qquad C_2 \geq \int_{-\infty}^{\infty} dF(\omega) = \int_{-\infty}^{\infty} f(\omega) d\omega > 0$$

is the non stationary covariance function. Note that the covariance function $C(\cdot)$ implies that, by Bochner's theorem, there exists a bounded measure F such that

$$C(x) = \int_{-\infty}^{\infty} e^{-ix\omega} dF(\omega).$$

When $F(\cdot)$ is absolutely continuous, there exists a spectral density $f(\cdot) \geq 0$ such that

$$C(x) = \int_{-\infty}^{\infty} e^{-ix\omega} f(\omega) d\omega.$$

Now we choose a sequence of complex numbers $a_i, i = 1, 2, \cdots, n$, and any sequence of real numbers $t_i, i = 1, 2, \cdots, n$, taking $C_2 = \int_{-\infty}^{\infty} f(\omega) d\omega$,

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} a_i \overline{a}_j \tilde{C}(t_i, t_j) &= \sum_{i}\sum_{j} a_i \overline{a}_j (C_2 - C(t_i) - C(-t_j) + C(t_i - t_j)) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i \overline{a}_j \int_{-\infty}^{\infty} (1 - e^{-it_i\omega} - e^{it_j\omega} + e^{-i(t_i - t_j)\omega}) f(\omega) d\omega \\
&= \int_{-\infty}^{\infty} f(\omega) d\omega \left| \sum_{i=1}^{n} a_i (e^{-it_i\omega} - 1) \right|^2 \geq 0.
\end{aligned}$$

So we propose a five-parameter model for the covariance on a sphere

$$R(P,Q) = \tilde{C}(\phi_P, \phi_Q) C(\theta(P,Q,u)),$$

where $C_1 > 0, C_2 > 0, a, u, p$ can be estimated from the data.

5.1.3 Method development

We can construct normal independent (complex) random variate $W_m(\phi)$ associated with the variance-covariance matrix $C_m(\phi_P, \phi_Q)$ to construct an axially symmetric process. Then finite summation can be used to approximate above (5.1.4) infinite summation as given below,

$$X(P) = X(\phi, \lambda) = \sum_{m=-N}^{N} W_m(\phi) e^{im\lambda} \tag{5.1.8}$$

where this would provide the gridded data. Since W_m's are independent for $m = 1, 2, \cdots$, we have

$$
\begin{aligned}
Cov(X(P), X(Q)) &= Cov\left(\sum_{m=-N}^{N} W_m(\phi_P)e^{im\lambda_P}, \sum_{j=-N}^{N} W_j(\phi_Q)e^{ij\lambda_Q}\right) \\
&= \sum_{m,j} e^{im\lambda_P}e^{-ij\lambda_Q}Cov(W_m(\phi_P), W_j(\phi_Q)) \\
&= \sum_{m} e^{im(\lambda_P-\lambda_Q)}C_m(\phi_P, \phi_Q)
\end{aligned}
$$

The above generated data will be complex random variates. Therefore to have the real-valued data observations or to obtain a real process, we need to have

$$C_{-m}(\phi_P, \phi_Q) = \overline{C_m(\phi_P, \phi_Q)}, \quad \text{for } m = 1, 2, \cdots, N \tag{5.1.9}$$

Lets write $W_m(\phi) = W^r_m(\phi) + iW^i_m(\phi)$ in terms of a real component and an imaginary component. We also write $C_m(\phi_P, \phi_Q) = C^r_m(\phi_P, \phi_Q)+iC^i_m(\phi_P, \phi_Q)$ and with the relationship 5.1.9 above, we have

$$C^r_{-m}(\phi_P, \phi_Q) = C^r_{-m}(\phi_P, \phi_Q), \quad C^i_{-m}(\phi_P, \phi_Q) = -C^i_{-m}(\phi_P, \phi_Q).$$

Now,

$$
\begin{aligned}
Cov(W_m(\phi_P), W_m(\phi_Q)) &= Cov(W^r_m(\phi_P) + iW^i_m(\phi_P), W^r_m(\phi_Q) + iW^i_m(\phi_Q)) \\
&= \left[Cov(W^r_m(\phi_P), W^r_m(\phi_Q)) + Cov(W^i_m(\phi_P), W^i_m(\phi_Q))\right] \\
&\quad +i\left[-Cov(W^r_m(\phi_P), W^i_m(\phi_Q)) + Cov(W^i_m(\phi_P), W^r_m(\phi_Q))\right] \\
&= C^r_m(\phi_P, \phi_Q) + iC^i_m(\phi_P, \phi_Q).
\end{aligned}
$$

If we let $W_{-m}(\phi) = \overline{W_m(\phi)}$, then the covariance function would satisfy the above relationship 5.1.9. In addition, we will set the following,

$$Cov(W^r_m(\phi_P), W^r_m(\phi_Q)) = Cov(W^i_m(\phi_P), W^i_m(\phi_Q)) = \frac{1}{2}C^r_m(\phi_P, \phi_Q), \tag{5.1.10}$$

$$Cov(W^i_m(\phi_P), W^r_m(\phi_Q)) = -Cov(W^r_m(\phi_P), W^i_m(\phi_Q)) = \frac{1}{2}C^i_m(\phi_P, \phi_Q). \tag{5.1.11}$$

Therefore, if we denote $\underset{\sim}{W}_m(\phi) = (W^r_m(\phi), W^i_m(\phi))^T$, then the variance-covariance matrix for $\underset{\sim}{W}_m(\phi)$ is given by

$$\frac{1}{2}\begin{pmatrix} C^r_m(\phi_P, \phi_Q) & -C^i_m(\phi_P, \phi_Q) \\ C^i_m(\phi_P, \phi_Q) & C^r_m(\phi_P, \phi_Q) \end{pmatrix}.$$

However, we cannot have a vector of random variables $\underset{\sim}{W}_m(\phi)$ with a non-symmetric variance-covariance matrix unless $C^i_m(\phi_P, \phi_Q) = 0$. In the next section we will demonstrate how to generate $\underset{\sim}{W}_m(\phi)$ with a symmetric variance-covariance

The process given by (5.1.4) is now simplified as the following (real) process,

$$\begin{aligned}
X(P) &= \sum_{m=-N}^{N} W_m(\phi)e^{im\lambda} = W_0(\phi) + \sum_{m=1}^{N} W_m(\phi)e^{im\lambda} + \sum_{m=-1}^{-N} W_m(\phi)e^{im\lambda} \\
&= W_0(\phi) + \sum_{m=1}^{N} W_m(\phi)e^{im\lambda} + \sum_{m=1}^{N} \overline{W_m(\phi)}e^{-im\lambda} \\
&= W_0(\phi) + \sum_{m=1}^{N} \big[(W_m^r(\phi) + iW_m^i(\phi))(\cos(m\lambda) + i\sin(m\lambda)) \\
&\quad + (W_m^r(\phi) - iW_m^i(\phi))(\cos(m\lambda) - i\sin(m\lambda))\big] \\
&= W_0(\phi) + 2\sum_{m=1}^{N} \left[W_m^r(\phi)\cos(m\lambda) - W_m^i(\phi)\sin(m\lambda)\right]. \qquad (5.1.12)
\end{aligned}$$

5.1.4 Data generation

Now for each fixed $m = 0, 1, 2, \cdots, N$, we consider $W_m(\phi) = W_m^r(\phi) + iW_m^i(\phi)$ then $W_m^*(\phi) = W_m^r(\phi) - iW_m^i(\phi)$ (where $W_m^*(\phi)$ is the conjugate of $W_m(\phi)$). We may assume that $W_m^r(\phi)$ and $W_m^i(\phi)$ are independent, each following a (normal) distribution with mean zero and the same variance $\sigma_m^2(\phi) = \frac{1}{2}C_m^r(\phi,\phi)$, ($C_m^i(\phi,\phi) = 0$ implies $W_m^r(\phi)$ and $W_m^i(\phi)$ are uncorrelated, or independent for Gaussian). From 5.1.1, $W_m(\phi)$ is the circularly-symmetric complex random variable (Gallager (2008)).

Now for a set of distinct latitudes $\Phi = \{\phi_1, \phi_2, \cdots, \phi_{n_l}\}$, we consider a sequence of complex random variables $\{W_m(\phi) : \phi \in \Phi\}$, which forms a multivariate complex random vector $\underset{\sim}{W}_m = (W_m(\phi_1), W_m(\phi_2), \cdots, W_m(\phi_n))^T$ where $W_m(\phi_i) = W_m^r(\phi_i) + iW_m^r(\phi_i)$ with associated $2 \times n_l$-dimensional real random vector

$$\underset{\sim}{V}_m = (W_m^r(\phi_1), W_m^i(\phi_1), W_m^r(\phi_2), W_m^i(\phi_2), \cdots, W_m^r(\phi_{n_l}), W_m^i(\phi_{n_l}))^T.$$

Now we calculate the covariance matrix $K_W = E(\underset{\sim}{W}_m \underset{\sim}{W}_m^*)$ (where $\underset{\sim}{W}_m^*$ is the conjugated transpose) and pseudo-covariance $M_W = E(\underset{\sim}{W}_m \underset{\sim}{W}_m^T)$. Further, from 5.1.1 a complex random vector is circularly-symmetric if and only if M_W is zero.

$$\begin{aligned}
M_W &= \begin{pmatrix}
E[W_m(\phi_1)W_m(\phi_1)] & E[W_m(\phi_1)W_m(\phi_2)] & \cdots & E[W_m(\phi_1)W_m(\phi_{n_l})] \\
E[W_m(\phi_2)W_m(\phi_1)] & E[W_m(\phi_2)W_m(\phi_2)] & \cdots & E[W_m(\phi_2)W_m(\phi_{n_l})] \\
\vdots & \vdots & \ddots & \vdots \\
E[W_m(\phi_{n_l})W_m(\phi_1)] & E[W_m(\phi_{n_l})W_m(\phi_2)] & \cdots & E[W_m(\phi_{n_l})W_m(\phi_{n_l})]
\end{pmatrix} \\
&= \mathbf{0}
\end{aligned}$$

We can easily show the above for $\forall\ i, j$,

$$\begin{aligned}
& E[W_m(\phi_i)W_m(\phi_j)] \\
= \; & E[(W_m^r(\phi_i)+iW_m^i(\phi_i))(W_m^r(\phi_j)+iW_m^i(\phi_j))] \\
= \; & E(W_m^r(\phi_i)W_m^r(\phi_j)) - E(W_m^i(\phi_i)W_m^i(\phi_j)) + i[E(W_m^r(\phi_i)W_m^i(\phi_j)) + E(W_m^i(\phi_i)W_m^r(\phi_j))] \\
& for\ i \neq j \\
= \; & \frac{1}{2}(C_m^r(\phi_i,\phi_j) - C_m^r(\phi_i,\phi_j)) + i[-\frac{1}{2}C_m^i(\phi_i,\phi_j) + \frac{1}{2}C_m^i(\phi_i,\phi_j)] = 0 \\
& for\ i = j \\
= \; & \frac{1}{2}(C_m^r(\phi_i,\phi_i) - C_m^r(\phi_i,\phi_i)) + i[0+0] = 0 \quad ; W_m^r(\phi_i), W_m^i(\phi_i) \text{ are independent}
\end{aligned}$$

Therefore, $\underset{\sim}{W}_m$ is circularly-symmetric. In addition,

$$\begin{aligned}
K_W &= E(\underset{\sim}{W}_m \underset{\sim}{W}_m^*) \\
&= \begin{pmatrix}
E[W_m(\phi_1)W_m^*(\phi_1)] & E[W_m(\phi_1)W_m^*(\phi_2)] & \cdots & E[W_m(\phi_1)W_m^*(\phi_{n_l})] \\
E[W_m(\phi_2)W_m^*(\phi_1)] & E[W_m(\phi_2)W_m^*(\phi_2)] & \cdots & E[W_m(\phi_2)W_m^*(\phi_{n_l})] \\
\vdots & \vdots & \ddots & \vdots \\
E[W_m(\phi_{n_l})W_m^*(\phi_1)] & E[W_m(\phi_{n_l})W_m^*(\phi_2)] & \cdots & E[W_m(\phi_{n_l})W_m^*(\phi_{n_l})]
\end{pmatrix} \\
&= \begin{pmatrix}
C_m^r(\phi_1,\phi_1) & C_m^r(\phi_1,\phi_2)+iC_m^i(\phi_1,\phi_2) & \cdots & C_m^r(\phi_1,\phi_{n_l})+iC_m^i(\phi_1,\phi_{n_l}) \\
C_m^r(\phi_2,\phi_1)-iC_m^i(\phi_2,\phi_1) & C_m^r(\phi_2,\phi_2) & \cdots & C_m^r(\phi_2,\phi_{n_l})+iC_m^r(\phi_2,\phi_{n_l}) \\
\vdots & \vdots & \ddots & \vdots \\
C_m^r(\phi_{n_l},\phi_1)-iC_m^i(\phi_{n_l},\phi_1) & C_m^r(\phi_{n_l},\phi_2)-iC_m^r(\phi_{n_l},\phi_2) & \cdots & C_m^r(\phi_{n_l},\phi_{n_l})
\end{pmatrix} \\
&= \begin{pmatrix}
C_m^r(\phi_1,\phi_1) & C_m^r(\phi_1,\phi_2) & \cdots & C_m^r(\phi_1,\phi_{n_l}) \\
C_m^r(\phi_2,\phi_1) & C_m^r(\phi_2,\phi_2) & \cdots & C_m^r(\phi_2,\phi_{n_l}) \\
\vdots & \vdots & \ddots & \vdots \\
C_m^r(\phi_{n_l},\phi_1) & C_m^r(\phi_{n_l},\phi_2) & \cdots & C_m^r(\phi_{n_l},\phi_{n_l})
\end{pmatrix} \\
&\quad +i \begin{pmatrix}
0 & C_m^i(\phi_1,\phi_2) & \cdots & C_m^i(\phi_1,\phi_{n_l}) \\
-C_m^i(\phi_2,\phi_1) & 0 & \cdots & C_m^i(\phi_2,\phi_{n_l}) \\
\vdots & \vdots & \ddots & \vdots \\
-C_m^i(\phi_{n_l},\phi_1) & -C_m^i(\phi_{n_l},\phi_2) & \cdots & 0
\end{pmatrix} \\
&= Re(K_W) + iIm(K_W)
\end{aligned}$$

Now,

$$K_V = E(\underset{\sim}{V}_m \underset{\sim}{V}_m^*) = E(\underset{\sim}{V}_m \underset{\sim}{V}_m^T)$$

In order to generate K_V for n_l-tuple case, we reorganize the vector $\underset{\sim}{V}_m$ into the following form.

$$\underset{\sim}{V}_m = (W_m^r(\phi_1), W_m^r(\phi_2), \cdots, W_m^r(\phi_{n_l}), W_m^i(\phi_1), W_m^i(\phi_2), \cdots, W_m^r(\phi_{n_l}))^T = (Re(\underset{\sim}{W}_m), Im(\underset{\sim}{W}_m))^T$$

that is, we grouped all real components and imaginary components together. Hence,

$$\begin{aligned} K_V &= E(\underset{\sim}{V}_m \underset{\sim}{V}_m^T) \\ &= \begin{pmatrix} E[Re(\underset{\sim}{W}_m)Re(\underset{\sim}{W}_m)^T] & E[Re(\underset{\sim}{W}_m)Im(\underset{\sim}{W}_m)^T] \\ E[Im(\underset{\sim}{W}_m)Re(\underset{\sim}{W}_m)^T] & E[Im(\underset{\sim}{W}_m)Im(\underset{\sim}{W}_m)^T] \end{pmatrix}_{2n_l \times 2n_l} \end{aligned}$$

Since $\underset{\sim}{W}_m$ is circularly-symmetric from 5.1.1 we can get the following results,

$$E[Re(\underset{\sim}{W}_m)Re(\underset{\sim}{W}_m)^T] = E[Im(\underset{\sim}{W}_m)Im(\underset{\sim}{W}_m)^T] = \frac{1}{2}(Re(K_W))_{n_l \times n_l}$$

$$E[Re(\underset{\sim}{W}_m)Im(\underset{\sim}{W}_m)^T] = -E[Im(\underset{\sim}{W}_m)Re(\underset{\sim}{W}_m)^T] = \frac{1}{2}(Im(K_W))_{n_l \times n_l}$$

$$K_V = \frac{1}{2}\begin{pmatrix} Re(K_W) & Im(K_W)^T \\ Im(K_W) & Re(K_W) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} Re(K_W) & -Im(K_W) \\ Im(K_W) & Re(K_W) \end{pmatrix}$$

Since K_V is a non-negative definite and matrix, it can be represented as follows,

$$K_V = Q\Lambda Q^T,$$

where Λ is a diagonal matrix with eigen values (real-positive) of K_V and Q are the corresponding orthonormal eigenvectors. We can choose $A = Q\Lambda^{1/2}Q^T$ to obtain,

$$\underset{\sim}{V}_m = A_{2n_l \times 2n_l} Z_{2n_l \times 1},$$

where $Z = \{z_1, z_2, \ldots, z_{n_l}, z_1^*, z_2^*, \ldots, z_{n_l}^*\}$ and each $z_i \sim N(0,1)$ hence we can get $\underset{\sim}{W}_m$. Now for each latitude $\phi_l, l = 1, 2, \cdots, n_l$ and $\lambda_k, k = 1, 2, \cdots, n_L$ $(N = n_L/2)$, we denote the axially symmetric data (real) as $X(\phi_l, \lambda_k)$. These random variates can be obtained from the equation (5.1.12), let's rewrite the equation as follows,

$$X(\phi_l, \lambda_k) = W_0(\phi_l) + 2\sum_{m=1}^{N} \left[W_m^r(\phi_l)\cos(m\lambda_k) - W_m^i(\phi_l)\sin(m\lambda_k)\right] \tag{5.1.13}$$

Pseudo-code

- Choose a cross covariance function, $R(P, Q)$
- Initialize the parameters (C_1, C_2, a, u, p) and choose a resolution $\phi_1, \ldots, \phi_{n_l}, \lambda_1, \ldots, \lambda_{n_L}$ (or $n_l \times n_L$),
- Derive $C_m(\phi_P, \phi_Q)$ based on $R(P, Q)$ where $m = 0, 1, \ldots, n_L/2$,
  1. for each m get $Re(K_W)$ and $Im(K_W)$ hence obtain K_V
  2. use SVD to get $\underset{\sim}{V}_m$ $(n_l - tuples)$
  3. get $\underset{\sim}{W}_m$'s from $\underset{\sim}{V}_m$
- apply the equation (5.1.13) to generate grid data.

5.1.5 Property of MOM

we might have to omit this section

We should conform the validity of proposed covariance functions $R(P, Q)$. Since $R(P, Q)$ functions are cross covariance functions, we can compute the empirical covariance, by method of moments and compare.

We can estimate the cross covariance between any two arbitrary latitudes at each longitudinal difference (empirical covariance) based on method of moments and compare it with the theoretical $R(P, Q)$ values. According to Wackernagel (2013) a cross covariance function is not an even function so does the proposed $R(P, Q)$ functions, *i.e.* $R(P, Q, \Delta\lambda) \neq R(P, Q, -\Delta\lambda)$. The empirical (cross) covariance between any two latitudes ϕ_P and ϕ_Q with a zero mean processes can be given as follows,

$$\hat{R}(\phi_P, \phi_Q, \theta) = \frac{1}{n_L} \sum_{k=1}^{n_L} (X(\phi_P, \theta + \lambda_k) \cdot X(\phi_l, \lambda_k)), \tag{5.1.14}$$

where $\theta = m\Delta\lambda$.

In general, the spatial processes are stationary at a given latitude but not within latitudes. If the processes are not zero mean once could simply substract the product of means from the above MOM estimator. The above estimate is clearly unbiased as we have, for fixed two locations $P, Q \in S^2$,

$$E[\hat{R}(\phi_P, \phi_Q, \theta)] = R(P, Q).$$

Later we will prove this estimator is consistent, *i.e.,* for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|\hat{R}(\phi_P, \phi_Q, \theta) - R(P, Q)| > \varepsilon) = 0.$$

5.1.6 Results

Simulated data sample:

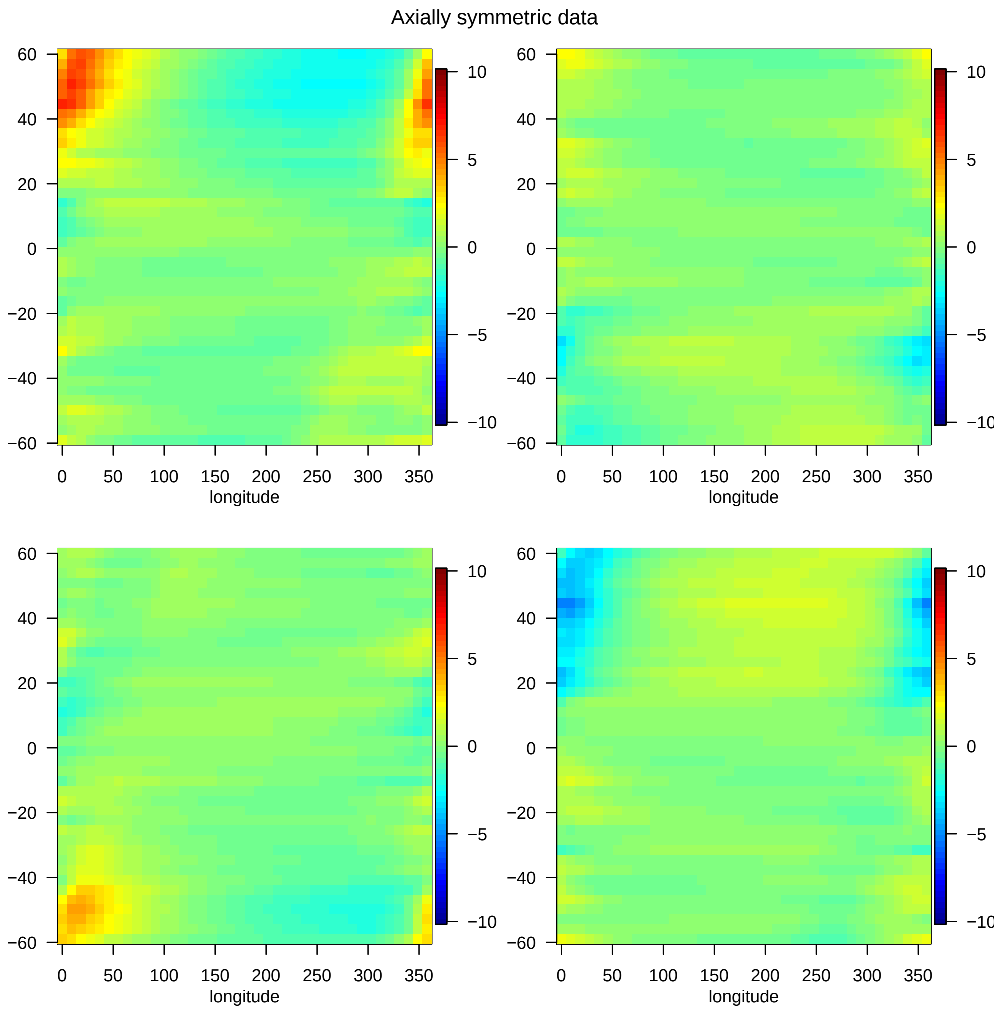

Figure 5.1: Four consecutive axially symmetric data snapshots based on model 2, grid resolution $2^0 \times 1^0$ (data scale -10 and 10).

Comprison of the proposed models with MOM estimates:

- Model 1

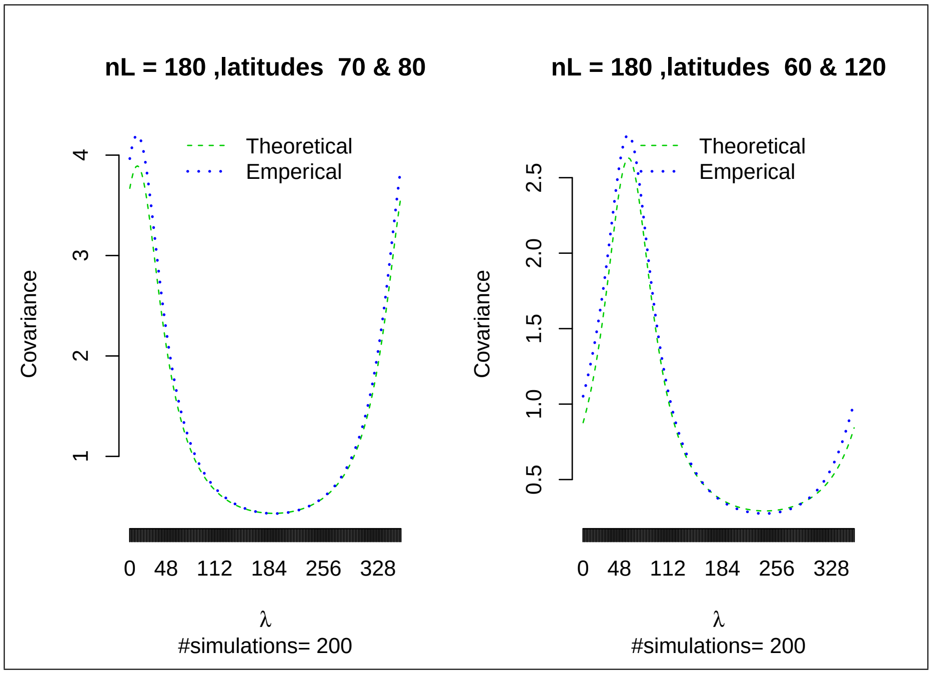

Figure 5.2: Cross covariance comparison of model1

- Model 2

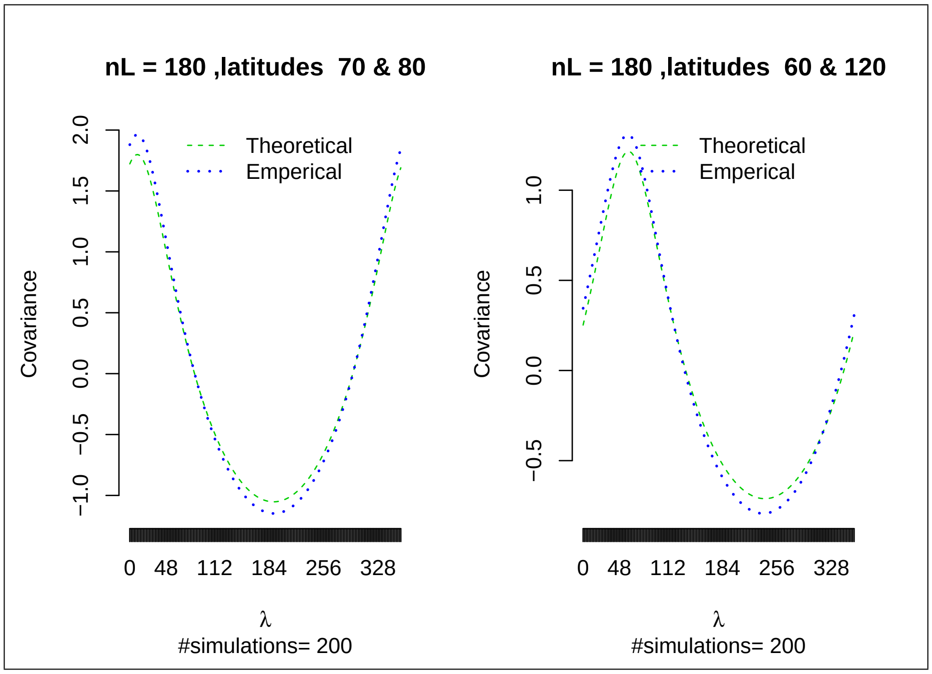

Figure 5.3: Cross covariance comparison of model2

- Model 3

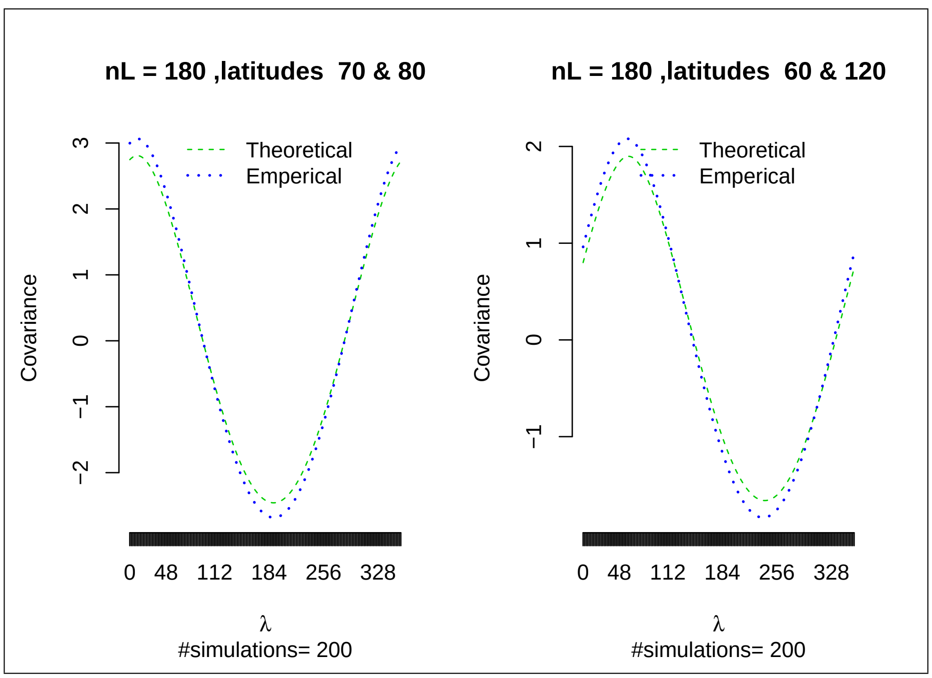

Figure 5.4: Cross covariance comparison of model3

Chapter 6

Future Research (due August 28)

Future research work !!

Bibliography

Bolin, D. and Lindgren, F. (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *The Annals of Applied Statistics*, pages 523–550.

Cressie, N. and Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):209–226.

Gallager, R. G. (2008). Circularly-symmetric gaussian random vectors. *preprint*, pages 1–9.

Hitczenko, M. and Stein, M. L. (2012). Some theory for anisotropic processes on the sphere. *Statistical Methodology*, 9(1):211–227.

Huang, C., Zhang, H., and Robeson, S. M. (2011). On the validity of commonly used covariance and variogram functions on the sphere. *Mathematical Geosciences*, 43(6):721–733.

Huang, C., Zhang, H., and Robeson, S. M. (2012). A simplified representation of the covariance structure of axially symmetric processes on the sphere. *Statistics & Probability Letters*, 82(7):1346–1351.

Jeong, J. and Jun, M. (2015). A class of matérn-like covariance functions for smooth processes on a sphere. *Spatial Statistics*, 11:1–18.

Jones, R. H. (1963). Stochastic processes on a sphere. *The Annals of mathematical statistics*, 34(1):213–218.

Jun, M. and Stein, M. L. (2008). Nonstationary covariance models for global data. *The Annals of Applied Statistics*, pages 1271–1289.

Katzfuss, M. and Cressie, N. (2011). Spatio-temporal smoothing and em estimation for massive remote-sensing data sets. *Journal of Time Series Analysis*, 32(4):430–446.

Leadbetter, H. C.-M. (1967). Stationary and related stochastic processes. *New York*.

Li, T.-H. and North, G. R. (1997). Aliasing effects and sampling theorems of spherical random fields when sampled on a finite grid. *Annals of the Institute of Statistical Mathematics*, 49(2):341–354.

Li, Y. (2013). Non-parametric and semi-parametric estimation of spatial covariance function.

Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498.

Roy, R. (1972). Spectral analysis for a random process on the circle. *Journal of Applied Probability*, pages 745–757.

Stein, M. L. (1999). *Interpolation of spatial data: some theory for kriging*. Springer Science & Business Media.

Stein, M. L. (2007). Spatial variation of total column ozone on a global scale. *The Annals of Applied Statistics*, pages 191–210.

Wackernagel, H. (2013). *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media.

Yaglom, A. M. et al. (1961). Second-order homogeneous random fields. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob*, volume 2, pages 593–622.