Multivariate Spatial Modeling

Past, present and future

Will Kleiber

Department of Applied Mathematics
University of Colorado
Boulder, CO

Spatial Statistics, Boulder, CO
July 18, 2023

Module 1: Classical multivariate methods: cross-covariances

Univariate thinking

We're used to thinking about modeling of individual spatial processes

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + Z(\mathbf{s}) + \varepsilon(\mathbf{s})$$

where

- $\mathbf{s} \in \mathbb{R}^d$ is the spatial (or temporal or space-time or something else) index
- $Y(\mathbf{s})$ is the observational process
- $\mu(\mathbf{s})$ is a fixed unknown mean function
- $Z(\mathbf{s})$ is a mean zero correlated stochastic process
- $\varepsilon(\mathbf{s})$ is a mean zero white noise process

Univariate thinking

We're used to thinking about modeling of individual spatial processes

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + Z(\mathbf{s}) + \varepsilon(\mathbf{s})$$

where

- $\mathbf{s} \in \mathbb{R}^d$ is the spatial (or temporal or space-time or something else) index
- $Y(\mathbf{s})$ is the observational process
- $\mu(\mathbf{s})$ is a fixed unknown mean function
- $Z(\mathbf{s})$ is a mean zero correlated stochastic process
- $\varepsilon(\mathbf{s})$ is a mean zero white noise process

Common goals:

- Characterize and estimate the spatial behavior of Z
- Estimate/predict $\mu(\mathbf{s}_0) + Z(\mathbf{s}_0)$ (kriging)
- Stochastic simulation of $Z(\mathbf{s})$

Multivariate examples

In many data examples we have multiple responses that are measured or are of interest:

- Statistical climatology
  - Observational data (minimum temperature, maximum temperature, precipitation and wind speed)
  - Model data (temperature, relative humidity, cloud particle size)
- Social sciences
  - American community survey (household size and type, race, gender)
  - Real estate (asking and selling price)
- Geophysical science
  - Upper atmosphere models (flux and energy of precipitating electrons)
  - Ocean models (density, salinity, velocity and buoyancy)
  - Soil science (concentrations of soil types)

Multivariate thinking

A typical model for p observed spatial processes, $\mathbf{Y}(\mathbf{s})$, is

$$\begin{pmatrix} Y_1(\mathbf{s}) \\ Y_2(\mathbf{s}) \\ \vdots \\ Y_p(\mathbf{s}) \end{pmatrix} = \boldsymbol{\mu}(\mathbf{s}) + \mathbf{Z}(\mathbf{s}) + \boldsymbol{\varepsilon}(\mathbf{s}) = \begin{pmatrix} \mu_1(\mathbf{s}) \\ \mu_2(\mathbf{s}) \\ \vdots \\ \mu_p(\mathbf{s}) \end{pmatrix} + \begin{pmatrix} Z_1(\mathbf{s}) \\ Z_2(\mathbf{s}) \\ \vdots \\ Z_p(\mathbf{s}) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(\mathbf{s}) \\ \varepsilon_2(\mathbf{s}) \\ \vdots \\ \varepsilon_p(\mathbf{s}) \end{pmatrix}$$

where

- $\boldsymbol{\mu}(\mathbf{s})$ is a fixed unknown vector of mean functions
- $\mathbf{Z}(\mathbf{s})$ is a mean zero p-variate correlated stochastic process
- $\boldsymbol{\varepsilon}(\mathbf{s})$ is a mean zero p-variate white noise process

Multivariate thinking

A typical model for p observed spatial processes, $\mathbf{Y}(\mathbf{s})$, is

$$\begin{pmatrix} Y_1(\mathbf{s}) \\ Y_2(\mathbf{s}) \\ \vdots \\ Y_p(\mathbf{s}) \end{pmatrix} = \boldsymbol{\mu}(\mathbf{s}) + \mathbf{Z}(\mathbf{s}) + \boldsymbol{\varepsilon}(\mathbf{s}) = \begin{pmatrix} \mu_1(\mathbf{s}) \\ \mu_2(\mathbf{s}) \\ \vdots \\ \mu_p(\mathbf{s}) \end{pmatrix} + \begin{pmatrix} Z_1(\mathbf{s}) \\ Z_2(\mathbf{s}) \\ \vdots \\ Z_p(\mathbf{s}) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(\mathbf{s}) \\ \varepsilon_2(\mathbf{s}) \\ \vdots \\ \varepsilon_p(\mathbf{s}) \end{pmatrix}$$

where

- $\boldsymbol{\mu}(\mathbf{s})$ is a fixed unknown vector of mean functions
- $\mathbf{Z}(\mathbf{s})$ is a mean zero p-variate correlated stochastic process
- $\boldsymbol{\varepsilon}(\mathbf{s})$ is a mean zero p-variate white noise process

Common goals:

- Characterize and estimate the spatial behavior of $\mathbf{Z}$
- Estimate/predict $\boldsymbol{\mu}(\mathbf{s}_0) + \mathbf{Z}(\mathbf{s}_0)$ (co-kriging)
- Stochastic simulation of $\mathbf{Z}(\mathbf{s})$

To covariance, or to variogram: that is the question

For a univariate process there is a natural connection between the covariance function

$$\mathrm{Cov}(Z(\mathbf{s}), Z(\mathbf{t}))$$

and variogram

$$\mathrm{Var}(Z(\mathbf{s}) - Z(\mathbf{t})).$$

To covariance, or to variogram: that is the question

For a univariate process there is a natural connection between the covariance function

$$\mathrm{Cov}(Z(\mathbf{s}), Z(\mathbf{t}))$$

and variogram

$$\mathrm{Var}(Z(\mathbf{s}) - Z(\mathbf{t})).$$

For a multivariate process, though... What's more (or less) natural?

$$\mathrm{Cov}(Z_i(\mathbf{s}), Z_j(\mathbf{t})) \qquad \text{or} \qquad \mathrm{Var}(Z_i(\mathbf{s}) - Z_j(\mathbf{t}))$$

The former is a covariance, the latter is a...?

These are a cross-covariance and pseudo cross-variogram.

Most authors work with cross-covariances (stay tuned).

Matrix-valued covariance functions

Consider the $p \times p$ matrix-valued covariance function

$$\mathbf{C}(\mathbf{s}-\mathbf{t}) = (C_{ij}(\mathbf{s}-\mathbf{t}))_{i,j=1}^{p}$$

populated by

- (Direct)-Covariance functions $C_{ii}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_i(\mathbf{t}))$
- Cross-covariance functions $C_{ij}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_j(\mathbf{t})), i \neq j$.

Matrix-valued covariance functions

Consider the $p \times p$ matrix-valued covariance function

$$\mathbf{C}(\mathbf{s}-\mathbf{t}) = (C_{ij}(\mathbf{s}-\mathbf{t}))_{i,j=1}^{p}$$

populated by

- (Direct)-Covariance functions $C_{ii}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_i(\mathbf{t}))$
- Cross-covariance functions $C_{ij}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_j(\mathbf{t})), i \neq j$.

Example for $p = 2$:

$$\mathbf{C}(\mathbf{h}) = \begin{pmatrix} C_{11}(\mathbf{h}) & C_{12}(\mathbf{h}) \\ C_{21}(\mathbf{h}) & C_{22}(\mathbf{h}) \end{pmatrix} = \begin{pmatrix} e^{-\frac{\|\mathbf{h}\|}{2}} & 0.6e^{-\frac{\|\mathbf{h}\|}{4}} \\ 0.6e^{-\frac{\|\mathbf{h}\|}{4}} & e^{-\frac{\|\mathbf{h}\|}{10}} \end{pmatrix}$$

so Z_1 and Z_2 are unit variance Gaussian processes with different exponential correlation functions and

$$\mathrm{Cov}(Z_1(\mathbf{s}), Z_2(\mathbf{s})) = 0.6 \quad \text{and} \quad \mathrm{Cov}(Z_1(\mathbf{s}), Z_2(\mathbf{s}+\mathbf{h})) \to_{\mathbf{h}\to\infty} 0.$$

Matrix-valued covariance functions

Consider the $p \times p$ matrix-valued covariance function

$$\mathbf{C}(\mathbf{s}-\mathbf{t}) = (C_{ij}(\mathbf{s}-\mathbf{t}))_{i,j=1}^{p}$$

populated by

- (Direct)-Covariance functions $C_{ii}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_i(\mathbf{t}))$
- Cross-covariance functions $C_{ij}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_j(\mathbf{t})), i \neq j$.

Example for $p = 2$ observed at two locations:

$$\mathrm{Var}\begin{pmatrix}\mathbf{Z}(\mathbf{s})\\ \mathbf{Z}(\mathbf{t})\end{pmatrix} = \begin{pmatrix}\mathbf{C}(\mathbf{0}) & \mathbf{C}(\mathbf{s}-\mathbf{t})\\ \mathbf{C}(\mathbf{s}-\mathbf{t}) & \mathbf{C}(\mathbf{0})\end{pmatrix}$$
$$= \begin{pmatrix} C_{11}(\mathbf{0}) & C_{12}(\mathbf{0}) & C_{11}(\mathbf{s}-\mathbf{t}) & C_{12}(\mathbf{s}-\mathbf{t})\\ C_{21}(\mathbf{0}) & C_{22}(\mathbf{0}) & C_{21}(\mathbf{s}-\mathbf{t}) & C_{22}(\mathbf{s}-\mathbf{t})\\ C_{11}(\mathbf{s}-\mathbf{t}) & C_{12}(\mathbf{s}-\mathbf{t}) & C_{11}(\mathbf{0}) & C_{12}(\mathbf{0})\\ C_{21}(\mathbf{s}-\mathbf{t}) & C_{22}(\mathbf{s}-\mathbf{t}) & C_{21}(\mathbf{0}) & C_{22}(\mathbf{0}) \end{pmatrix}$$

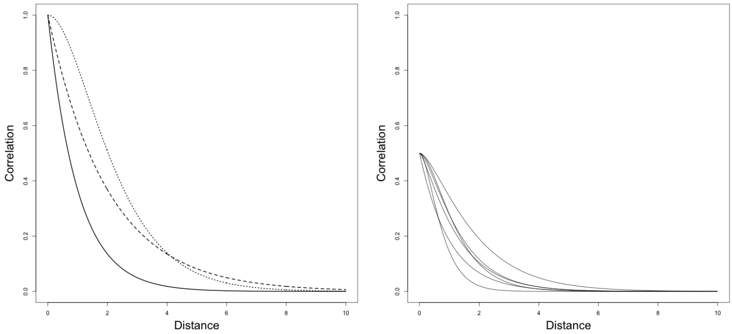
Matrix-valued covariance functions

Consider the $p \times p$ matrix-valued covariance function

$$\mathbf{C}(\mathbf{s}-\mathbf{t}) = (C_{ij}(\mathbf{s}-\mathbf{t}))_{i,j=1}^{p}$$

populated by

- (Direct)-Covariance functions $C_{ii}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_i(\mathbf{t}))$
- Cross-covariance functions $C_{ij}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_j(\mathbf{t})), i \neq j$.



Matrix-valued covariance functions

Consider the $p \times p$ matrix-valued covariance function

$$\mathbf{C}(\mathbf{s}-\mathbf{t}) = (C_{ij}(\mathbf{s}-\mathbf{t}))_{i,j=1}^{p}$$

populated by

- (Direct)-Covariance functions $C_{ii}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_i(\mathbf{t}))$
- Cross-covariance functions $C_{ij}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_j(\mathbf{t})), i \neq j$.

We require $\mathbf{C}$ to be nonnegative definite in that

$$\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ik}a_{j\ell}C_{ij}(\mathbf{s}_k - \mathbf{s}_\ell) \geq 0.$$

Note: the difference between $\mathbf{C}$ being a nonnegative definite function (true) and $\mathbf{C}(\mathbf{h})$ being a nonnegative definite matrix (false).

Matrix-valued covariance functions

Consider the $p \times p$ matrix-valued covariance function

$$\mathbf{C}(\mathbf{s}-\mathbf{t}) = (C_{ij}(\mathbf{s}-\mathbf{t}))_{i,j=1}^{p}$$

populated by

- (Direct)-Covariance functions $C_{ii}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_i(\mathbf{t}))$
- Cross-covariance functions $C_{ij}(\mathbf{s}-\mathbf{t}) = \mathrm{Cov}(Z_i(\mathbf{s}), Z_j(\mathbf{t})), i \neq j$.

We require $\mathbf{C}$ to be nonnegative definite in that

$$\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ik}a_{j\ell}C_{ij}(\mathbf{s}_k-\mathbf{s}_\ell) \geq 0.$$

Note: the difference between $\mathbf{C}$ being a nonnegative definite function (true) and $\mathbf{C}(\mathbf{h})$ being a nonnegative definite matrix (false).

This is a very difficult condition to ensure for some arbitrary proposed model, so most models are constructed to satisfy it.

Properties of cross-covariances

- $\mathbf{C}(\mathbf{0})$ is just a positive definite matrix of co-located cross-covariances

Properties of cross-covariances

- $\mathbf{C}(\mathbf{0})$ is just a positive definite matrix of co-located cross-covariances
- C_{ii} still has to be positive definite

Properties of cross-covariances

- $\mathbf{C}(\mathbf{0})$ is just a positive definite matrix of co-located cross-covariances
- C_{ii} still has to be positive definite
- C_{ij} does not have to be positive definite (it can even be negative)

Properties of cross-covariances

- $\mathbf{C}(\mathbf{0})$ is just a positive definite matrix of co-located cross-covariances
- C_{ii} still has to be positive definite
- C_{ij} does not have to be positive definite (it can even be negative)
- Cross-covariances do not have to be symmetric:

$$\mathrm{Cov}(Z_1(\mathbf{s}), Z_2(\mathbf{t})) \neq \mathrm{Cov}(Z_1(\mathbf{t}), Z_2(\mathbf{s}))$$

Properties of cross-covariances

- $\mathbf{C}(\mathbf{0})$ is just a positive definite matrix of co-located cross-covariances
- C_{ii} still has to be positive definite
- C_{ij} does not have to be positive definite (it can even be negative)
- Cross-covariances do not have to be symmetric:

$$\mathrm{Cov}(Z_1(\mathbf{s}), Z_2(\mathbf{t})) \neq \mathrm{Cov}(Z_1(\mathbf{t}), Z_2(\mathbf{s}))$$

- The cross-correlation function is

$$\rho_{12}(\mathbf{s}, \mathbf{t}) = \mathrm{Cor}(Z_1(\mathbf{s}), Z_2(\mathbf{t})) = \frac{C_{12}(\mathbf{s} - \mathbf{t})}{\sqrt{C_{11}(\mathbf{s} - \mathbf{s}) C_{22}(\mathbf{t} - \mathbf{t})}}$$

Properties of cross-covariances

- $\mathbf{C}(\mathbf{0})$ is just a positive definite matrix of co-located cross-covariances
- C_{ii} still has to be positive definite
- C_{ij} does not have to be positive definite (it can even be negative)
- Cross-covariances do not have to be symmetric:

$$\mathrm{Cov}(Z_1(\mathbf{s}), Z_2(\mathbf{t})) \neq \mathrm{Cov}(Z_1(\mathbf{t}), Z_2(\mathbf{s}))$$

- The cross-correlation function is

$$\rho_{12}(\mathbf{s}, \mathbf{t}) = \mathrm{Cor}(Z_1(\mathbf{s}), Z_2(\mathbf{t})) = \frac{C_{12}(\mathbf{s}-\mathbf{t})}{\sqrt{C_{11}(\mathbf{s}-\mathbf{s})C_{22}(\mathbf{t}-\mathbf{t})}}$$

- Cross-covariance functions often look like covariance functions, but they are not: don't be deceived!

Estimation: empirical cross-covariances

Much like empirical variograms or empirical covariance functions for univariate processes, the empirical cross-covariance matrix is:

$$\hat{\mathbf{C}}(\mathbf{h}) = \frac{1}{|N(\mathbf{h})|} \sum_{(k,\ell)\in N(\mathbf{h})} (\mathbf{Z}(\mathbf{s}_k) - \bar{\mathbf{Z}})(\mathbf{Z}(\mathbf{s}_\ell) - \bar{\mathbf{Z}})^{\mathrm{T}}.$$

In principle you can use a (weighted) least squares fit of $\hat{\mathbf{C}}$ much like in univariate, but there are usually restrictions in the parameter space that result in a valid mode that are more complicated than just $a > 0$.

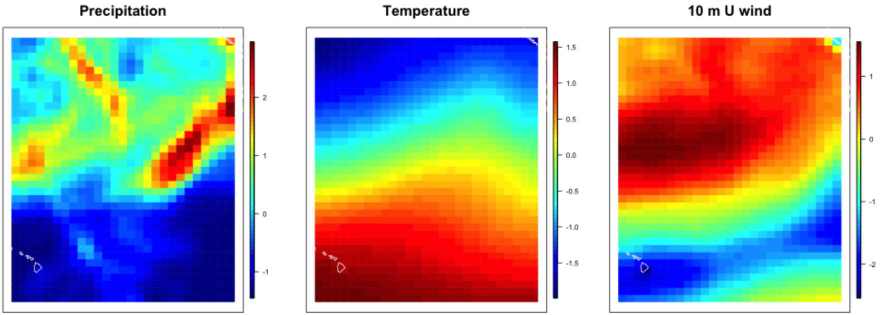
Large Ensemble

National Center for Atmospheric Research (NCAR)'s Large Ensemble (LENS):

- 10 members (think of these as independent realizations of a process)
- 2006 - 2080 monthly means over the North American region
- Resolution is approximately 1 degree
- Variables are surface temperature, large scale precipitation, solar insolation, and surface (10 m) wind speed

Question: Is there any relationship between variables?

LENS data: Pacific ocean



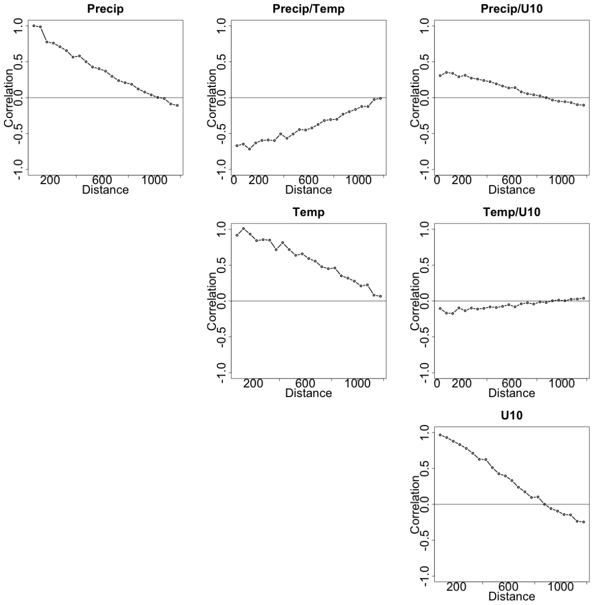
Some R code

```
library(fields)

tsu10.cross <- crossCoVGram(loc1=lon.lat,
  loc2=lon.lat,
  y1=c(TS[,,1]),
  y2=c(U10[,,1]),
  lon.lat=TRUE,
  dmax=1200, # about half the max distance
  type="cross-correlogram") # or cross-covariogram

plot(tsu10.cross,N=30)
```

LENS data: Pacific ocean



Estimation: maximum likelihood or Bayes

Once a parametric model has been chosen, assuming everything's Gaussian, estimation is "easy", just plug into the standard multivariate normal pdf but with being careful about the covariance matrix:

$$\operatorname{Var}\mathbf{Z} = \operatorname{Var}\begin{pmatrix}\mathbf{Z}(\mathbf{s}_1)\\ \mathbf{Z}(\mathbf{s}_2)\\ \vdots\\ \mathbf{Z}(\mathbf{s}_n)\end{pmatrix} = \begin{pmatrix}\mathbf{C}(\mathbf{0}) & \mathbf{C}(\mathbf{s}_1-\mathbf{s}_2) & \cdots & \mathbf{C}(\mathbf{s}_1-\mathbf{s}_n)\\ \mathbf{C}(\mathbf{s}_2-\mathbf{s}_1) & \mathbf{C}(\mathbf{0}) & \cdots & \mathbf{C}(\mathbf{s}_2-\mathbf{s}_n)\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{C}(\mathbf{s}_n-\mathbf{s}_1) & \mathbf{C}(\mathbf{s}_n-\mathbf{s}_2) & \cdots & \mathbf{C}(\mathbf{0})\end{pmatrix}$$

(side note: notationally there is a benefit to grouping by spatial location)
(side note: note the danger of dimensions here!)

Co-kriging

Spatial prediction for multivariate processes follows by co-kriging.

Given observations of $\mathbf{Z}(\mathbf{s})$ at $\mathbf{s} = \mathbf{s}_1, \ldots, \mathbf{s}_n$, the ("simple") co-kriging predictor at $\mathbf{s}_0$ is

$$\hat{\mathbf{Z}}(\mathbf{s}_0) = \mathrm{Cov}(\mathbf{Z}(\mathbf{s}_0), \mathbf{Z})\mathrm{Var}(\mathbf{Z})^{-1}\mathbf{Z}$$

where $\mathrm{Var}(\mathbf{Z})$ is on the previous slide, and

$$\mathrm{Cov}(\mathbf{Z}(\mathbf{s}_0), \mathbf{Z}) = \begin{pmatrix} \mathbf{C}(\mathbf{s}_0 - \mathbf{s}_1) & \mathbf{C}(\mathbf{s}_0 - \mathbf{s}_2) & \cdots & \mathbf{C}(\mathbf{s}_0 - \mathbf{s}_n) \end{pmatrix}$$

(Sanity check the dimensions of these matrices!)

As long as everything is Gaussian all of the standard conditional expectation formulas hold, but with more complicated covariance matrices.

Module 2: models for cross-covariances

Mommy, where do cross-covariances come from?

Just as in univariate modeling, there are essentially two approaches to specifying valid multivariate spatial models:

- Divine a matrix of non-negative definite functions $\mathbf{C}$ (e.g., multivariate Matérn)
- Build $\mathbf{Z}(\mathbf{s})$ from simpler building blocks that then imply a $\mathbf{C}$ (e.g., linear model of coregionalization)

Convolution and separability

Covariance convolution: If c_1 and c_2 are square integrable functions then

$$C_{ij}(\mathbf{h}) = (c_i \star c_j)(\mathbf{h}) = \int_{\mathbb{R}^d} c_i(\mathbf{h} - \mathbf{u}) c_j(\mathbf{u}) \mathrm{d}\mathbf{u}$$

forms a "valid" model.

Kernel convolution: If $C(\mathbf{h})$ is a stationary covariance then

$$C_{ij}(\mathbf{h}) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} c_i(\mathbf{u}) c_j(\mathbf{v}) C(\mathbf{u} - \mathbf{v} + \mathbf{h}) \mathrm{d}\mathbf{u} \mathrm{d}\mathbf{v}$$

forms a valid model.

Separable: If $C(\mathbf{h})$ is a positive definite function and $\mathbf{R}$ is a positive definite matrix, then

$$\mathbf{C}(\mathbf{h}) = \mathbf{R} C(\mathbf{h})$$

forms a valid model.

Conditioning

Cressie and Zammit-Mangion (2016) build multivariate models by conditioning: for $Z_1(\mathbf{s}), Z_2(\mathbf{s})$ specify

$$\begin{aligned} \mathbb{E}(Z_2(\mathbf{s}) \mid Z_1(\cdot)) &= \int b(\mathbf{s}, \mathbf{u}) Z_1(\mathbf{u}) \mathrm{d}\mathbf{u} \\ \mathrm{Cov}(Z_2(\mathbf{s}), Z_2(\mathbf{u}) \mid Z_1(\cdot)) &= C_{2|1}(\mathbf{s}, \mathbf{u}) \end{aligned}$$

where

- b is any integrable function
- $C_{2|1}$ is a univariate covariance
- Z_1 has univariate covariance C_{11}.

This implies some marginal covariance for Z_2 and a cross-covariance.

- Not clear in which order to condition
- Not easily extendable beyond $p = 2, 3$ processes.

Linear model of coregionalization

The linear model of coregionalization or spatial factor model assumes

$$\mathbf{Z}(\mathbf{s}) = A\mathbf{W}(\mathbf{s})$$

where

- $\mathbf{W}(\mathbf{s})$ contain independent spatial processes (easy)
- A is a matrix of coefficients (easy)
- Playing with the dimension of A and $\mathbf{W}$ allows for more or less flexibility, and dimension reduction

Linear model of coregionalization

The linear model of coregionalization or spatial factor model assumes

$$\mathbf{Z}(\mathbf{s}) = A\mathbf{W}(\mathbf{s})$$

where

- $\mathbf{W}(\mathbf{s})$ contain independent spatial processes (easy)
- A is a matrix of coefficients (easy)
- Playing with the dimension of A and $\mathbf{W}$ allows for more or less flexibility, and dimension reduction

Implied cross-covariances are easy:

$$\mathrm{Cov}(Z_1(\mathbf{s}), Z_2(\mathbf{t})) = \sum_k a_{1k}a_{2k}C_k(\mathbf{s}-\mathbf{t})$$

Multivariate Matérn

The multivariate Matérn sets

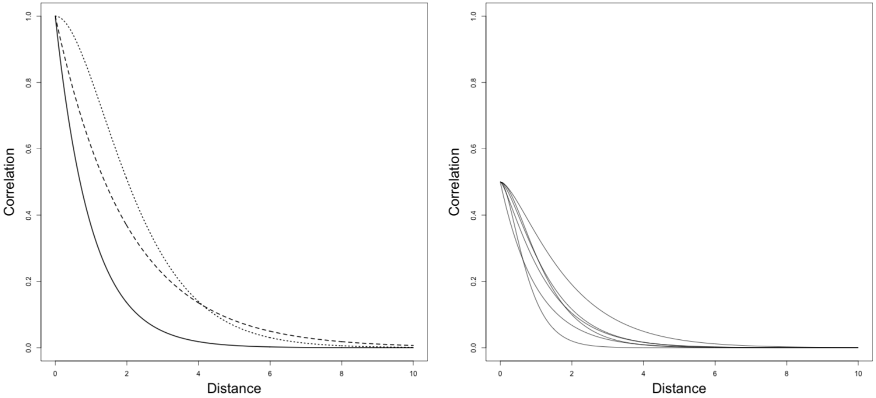
$$C_{ii}(\mathbf{h}) = \frac{2^{1-\nu_{ii}}}{\Gamma(\nu_{ii})}\left(a_{ii}\|\mathbf{h}\|\right)^{\nu_{ii}} \mathrm{K}_{\nu_{ii}}\left(a_{ii}\|\mathbf{h}\|\right)$$
$$C_{ij}(\mathbf{h}) = \rho_{ij}\frac{2^{1-\nu_{ij}}}{\Gamma(\nu_{ij})}\left(a_{ij}\|\mathbf{h}\|\right)^{\nu_{ij}} \mathrm{K}_{\nu_{ij}}\left(a_{ij}\|\mathbf{h}\|\right).$$

Where

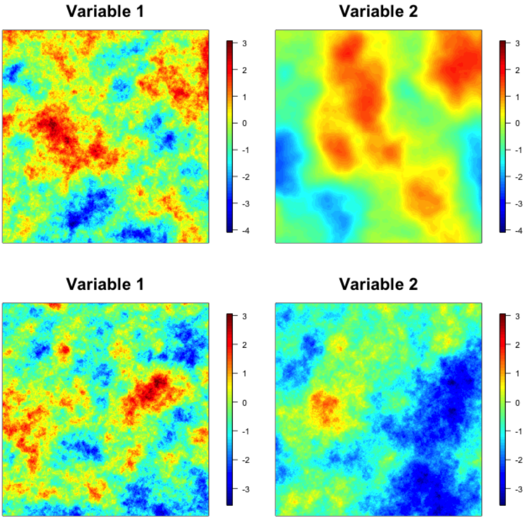
- ν_{ii} and a_{ii} retain marginal interpretations
- ρ_{ij} is cross-correlation coefficient
- ν_{ij} and a_{ij} are the cross-smoothness and cross-range (?).

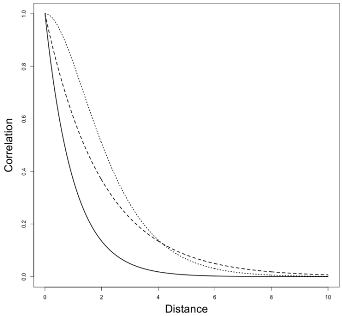
Note restrictions on parameters that result in a valid model are complicated and nontrivial.

Correlations vs. cross-correlations

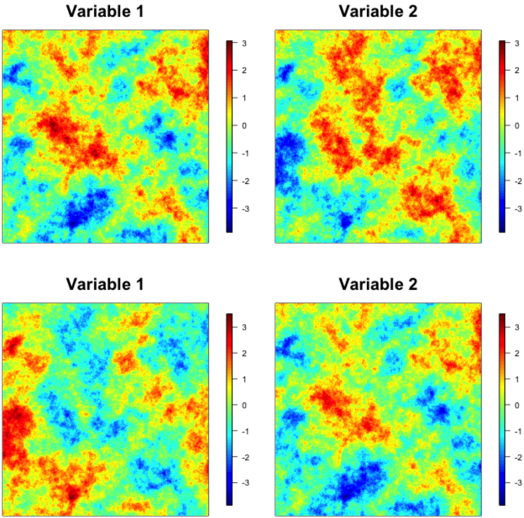


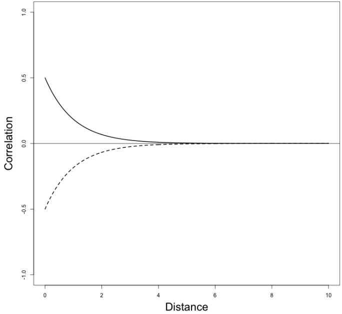
Covariance range, smoothness



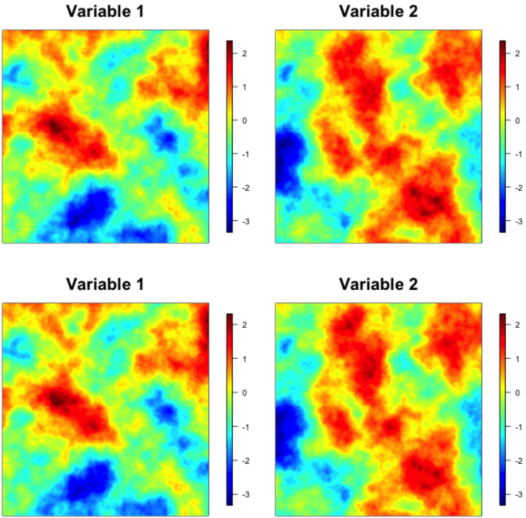


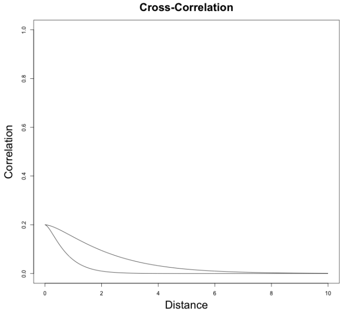
Correlation coefficient



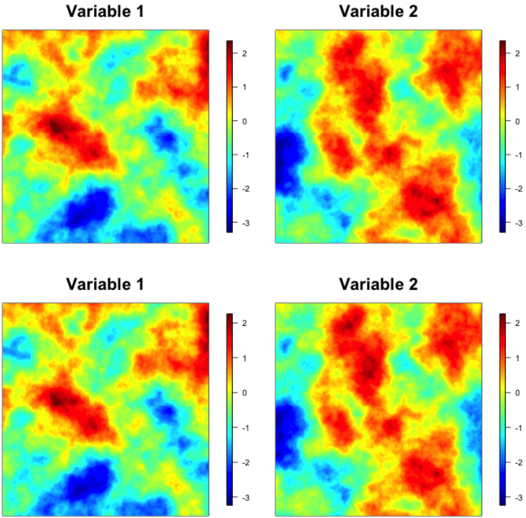


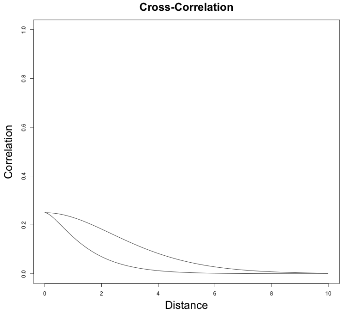
Cross-range


Variable 1
Variable 2
Variable 1
Variable 2
2
1
0
-1
-2
-3


Cross-Correlation
Correlation
Distance
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10

Cross-smoothness





A criticism

Take home message:

Cross-covariances are not necessarily so interpretable

Module 3: beyond covariance

Interlude: univariate spectra

Bochner's theorem: $C(\mathbf{h})$ (univariate!) is nonnegative definite iff

$$0 \leq f(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d} \int C(\boldsymbol{\omega}) \exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{h}) \mathrm{d}\mathbf{h}$$

for all $\boldsymbol{\omega}$ where $f(\boldsymbol{\omega})$ is the spectral density of C.

Interlude: univariate spectra

Bochner's theorem: $C(\mathbf{h})$ (univariate!) is nonnegative definite iff

$$0 \leq f(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d} \int C(\boldsymbol{\omega}) \exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{h}) \mathrm{d}\mathbf{h}$$

for all $\boldsymbol{\omega}$ where $f(\boldsymbol{\omega})$ is the spectral density of C.

Useful for model building, but not so useful for understanding anything (in my experience).

Interlude: univariate spectra

Bochner's theorem: $C(\mathbf{h})$ (univariate!) is nonnegative definite iff

$$0 \leq f(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d} \int C(\boldsymbol{\omega}) \exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{h}) \mathrm{d}\mathbf{h}$$

for all $\boldsymbol{\omega}$ where $f(\boldsymbol{\omega})$ is the spectral density of C.

Useful for model building, but not so useful for understanding anything (in my experience).

Key connection: spectral representation theorem (for stationary processes):

$$Z(\mathbf{s}) = \int \exp(i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{s}) M(\mathrm{d}\boldsymbol{\omega})$$

where $M(\cdot)$ is a mean zero random measure.

What's the deal with spectral densities anyway?

Spectral representation theorem:

$$\begin{aligned} Z(\mathbf{s}) &= \int \exp(i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{s})\, M(\mathrm{d}\boldsymbol{\omega}) \\ &\approx \sum_i \cos(\boldsymbol{\omega}_i^{\mathrm{T}}\mathbf{s} + \phi_i)\, M(\boldsymbol{\omega}_i)\Delta \qquad \text{(Z is real valued)} \\ &= \sum_i \cos(\boldsymbol{\omega}_i^{\mathrm{T}}\mathbf{s} + \phi_i)\, M_i \end{aligned}$$

where

$$\operatorname{Var} M_i = f(\boldsymbol{\omega}_i)\Delta$$

$\Rightarrow f(\boldsymbol{\omega})$ is the variance of the random coefficient at frequency $\boldsymbol{\omega}$.

What's the deal with spectral densities anyway?

Spectral representation theorem:

$$
\begin{aligned}
Z(\mathbf{s}) &= \int \exp(i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{s})\, M(\mathrm{d}\boldsymbol{\omega}) \\
&\approx \sum_i \cos(\boldsymbol{\omega}_i^{\mathrm{T}}\mathbf{s} + \phi_i)\, M(\boldsymbol{\omega}_i)\Delta \qquad \text{(Z is real valued)} \\
&= \sum_i \cos(\boldsymbol{\omega}_i^{\mathrm{T}}\mathbf{s} + \phi_i)\, M_i
\end{aligned}
$$

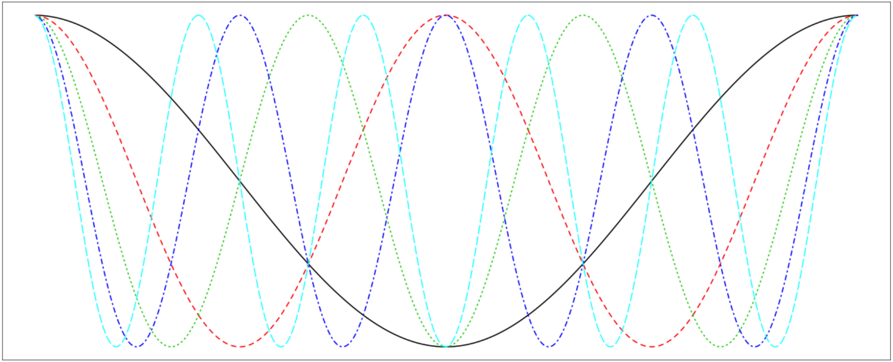
where

$$\operatorname{Var} M_i = f(\boldsymbol{\omega}_i)\Delta$$

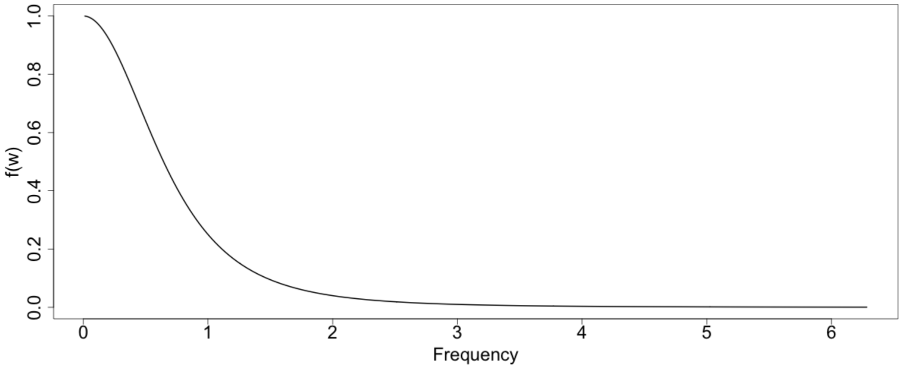
$\Rightarrow f(\boldsymbol{\omega})$ is the variance of the random coefficient at frequency $\boldsymbol{\omega}$.

Another view: instead of decomposing a function as a weighted combination of polynomials as in a power series, here we are decomposing $Z(\mathbf{s})$ using cosines, and M_i is the amplitude for the cosine with frequency $\boldsymbol{\omega}_i$.

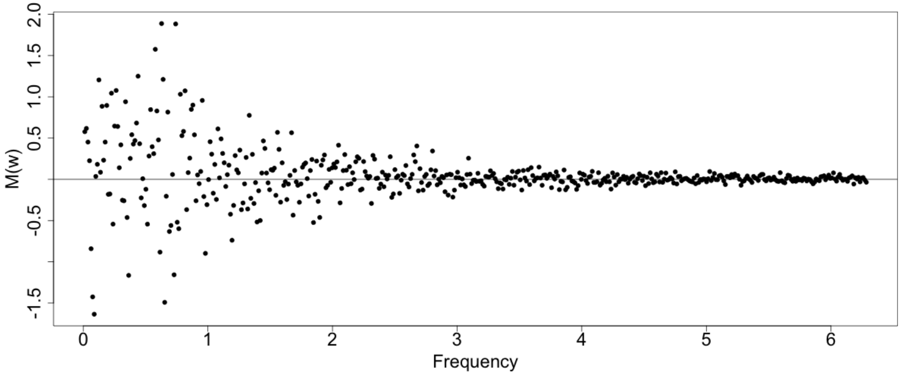
The basis



The spectrum

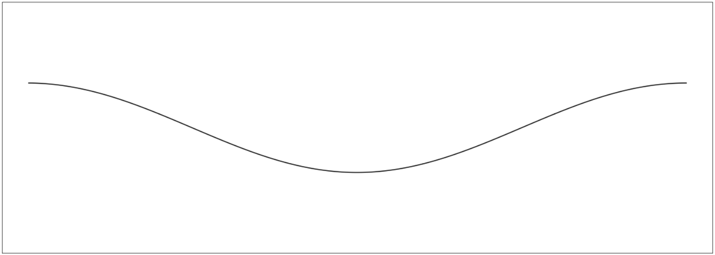


The random amplitudes

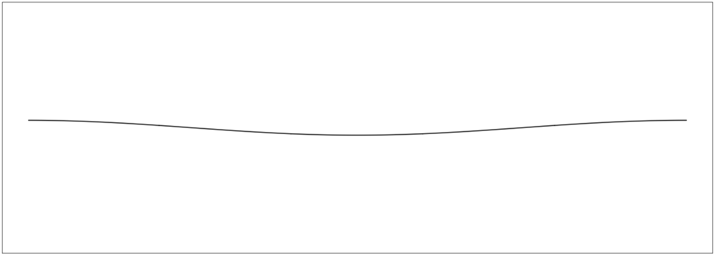


Component and partial sums

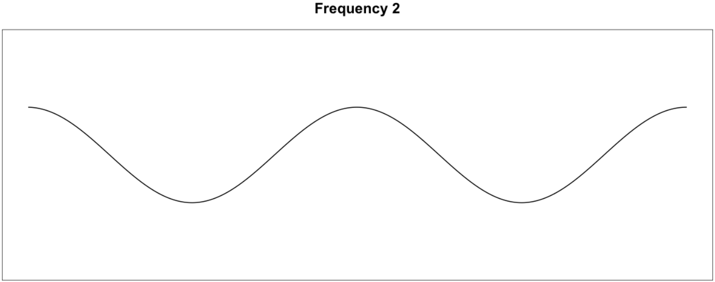
Frequency 1

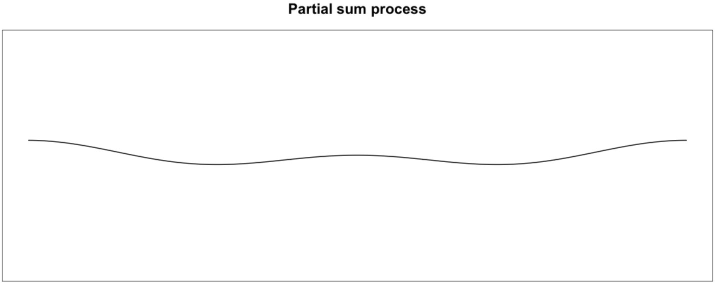


Partial sum process

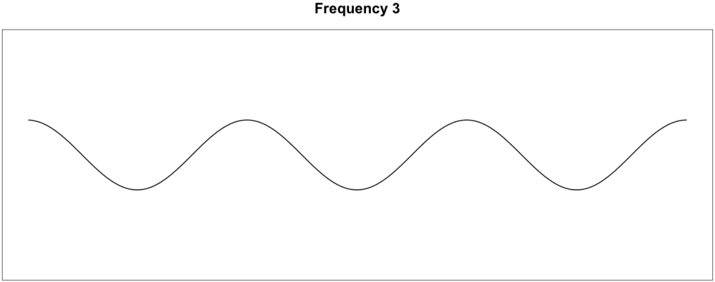


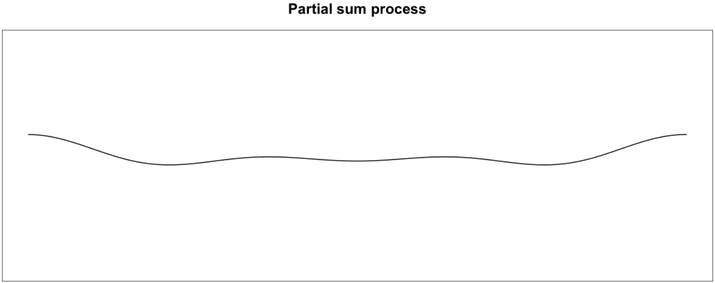
Component and partial sums



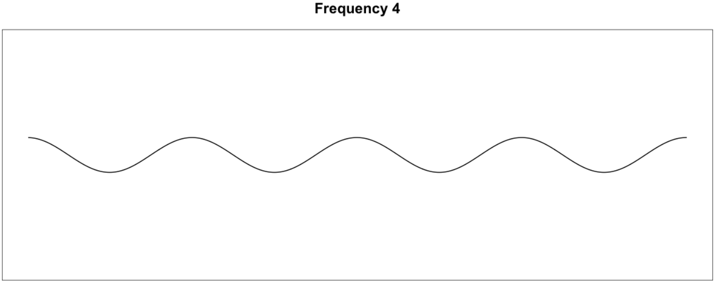


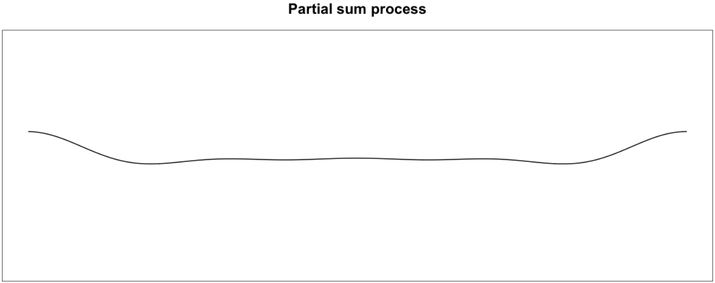
Component and partial sums



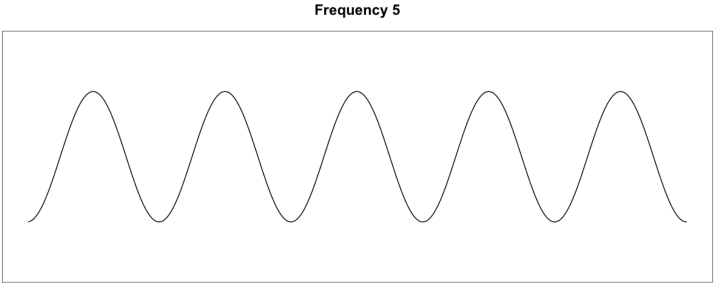


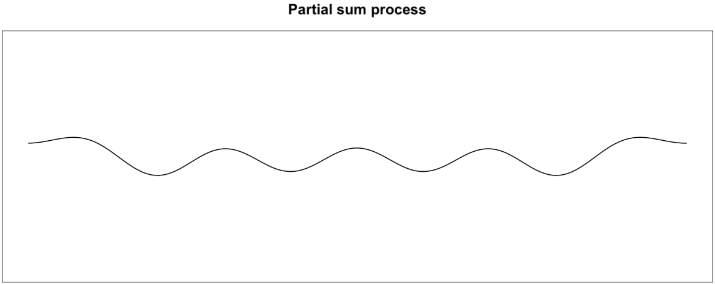
Component and partial sums





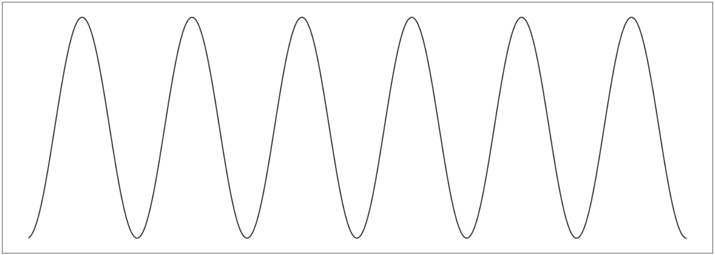
Component and partial sums



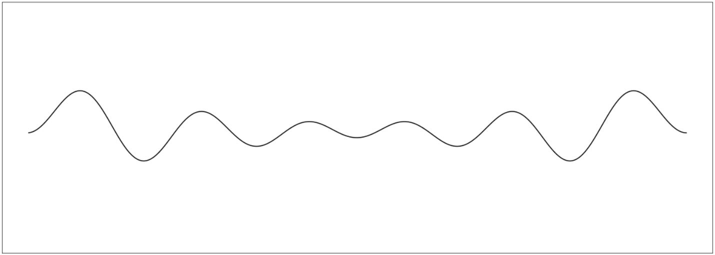


Component and partial sums

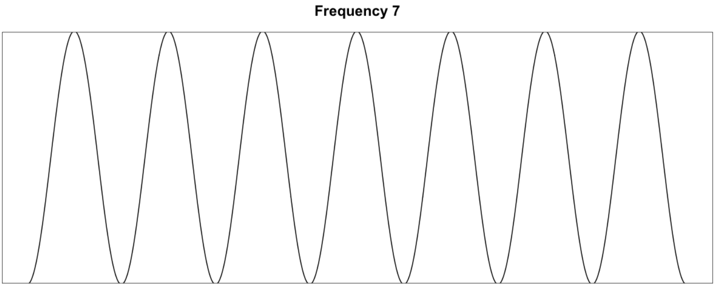
Frequency 6

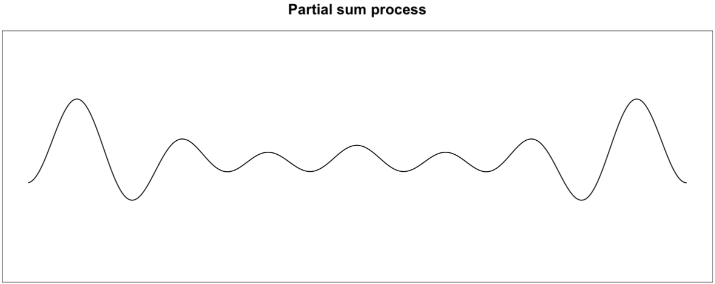


Partial sum process



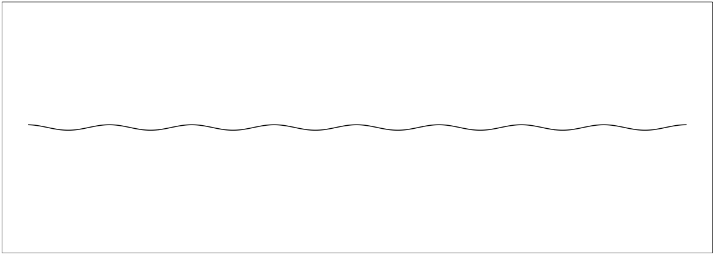
Component and partial sums



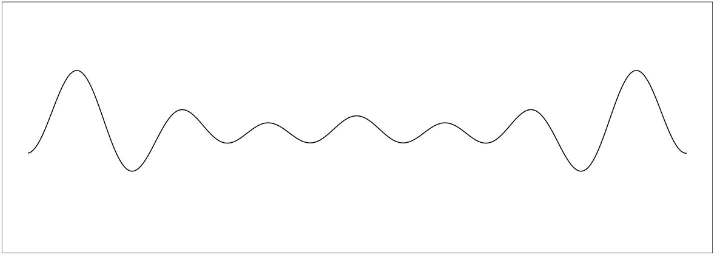


Component and partial sums

Frequency 8

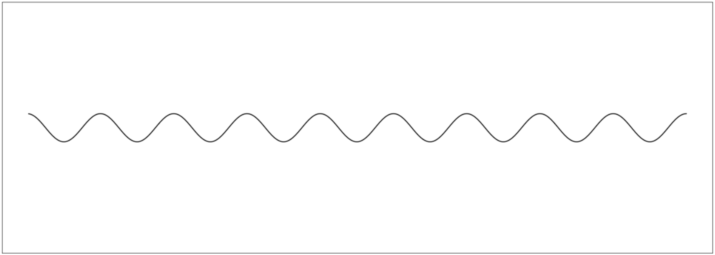


Partial sum process

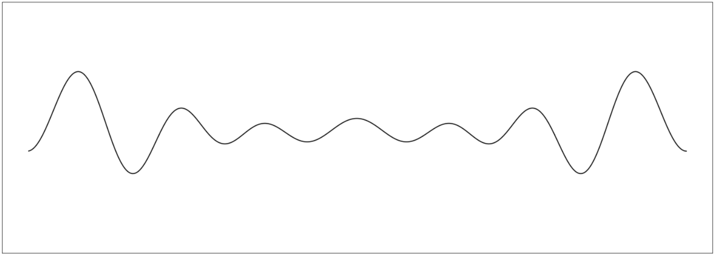


Component and partial sums

Frequency 9

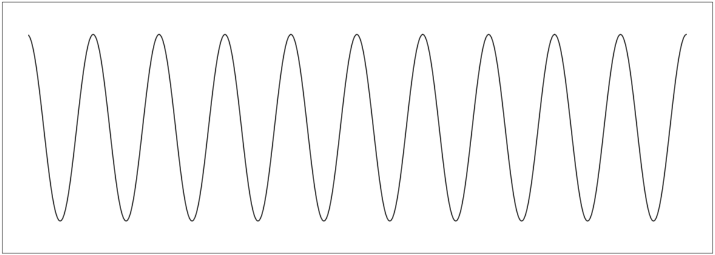


Partial sum process

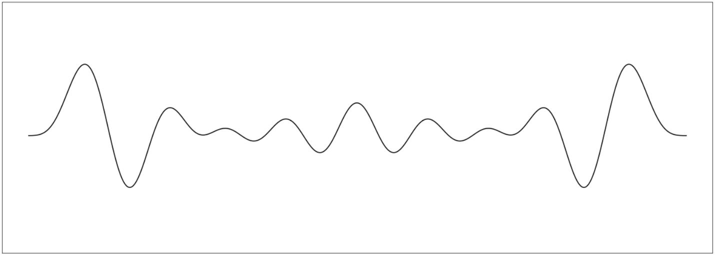


Component and partial sums

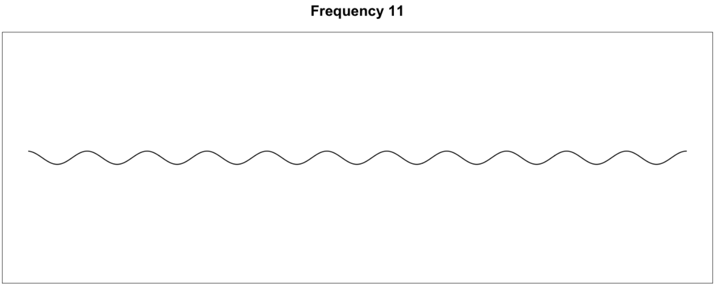
Frequency 10

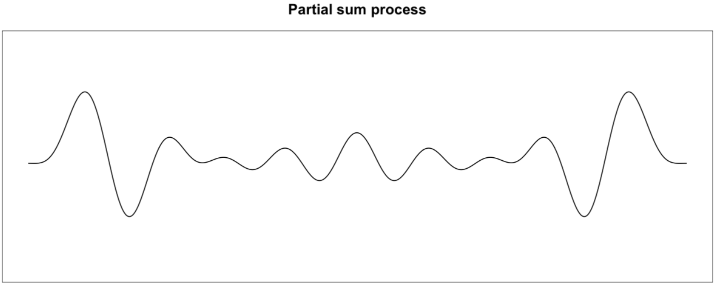


Partial sum process



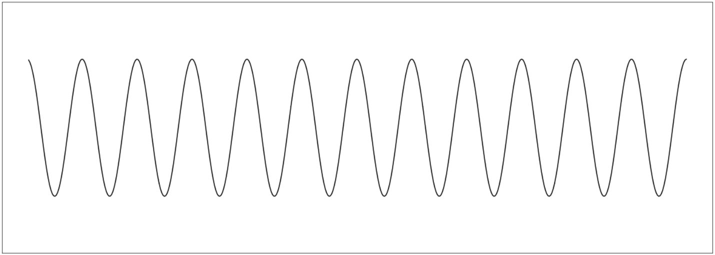
Component and partial sums



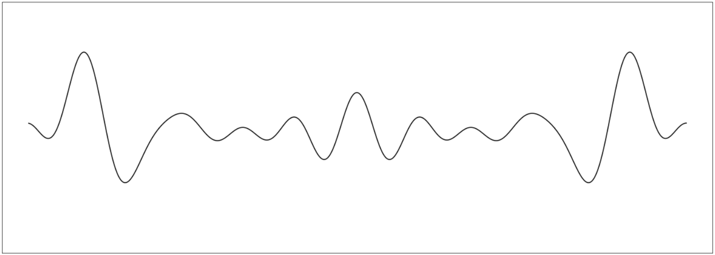


Component and partial sums

Frequency 12

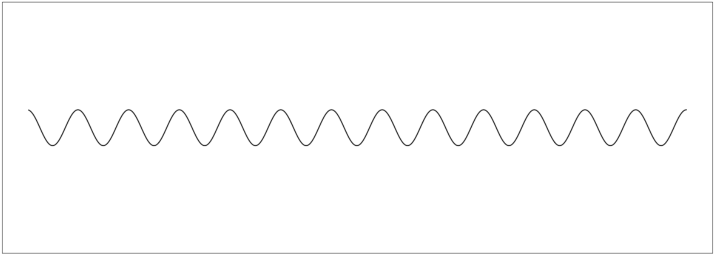


Partial sum process

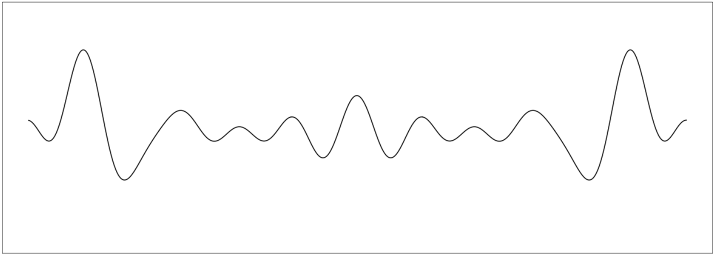


Component and partial sums

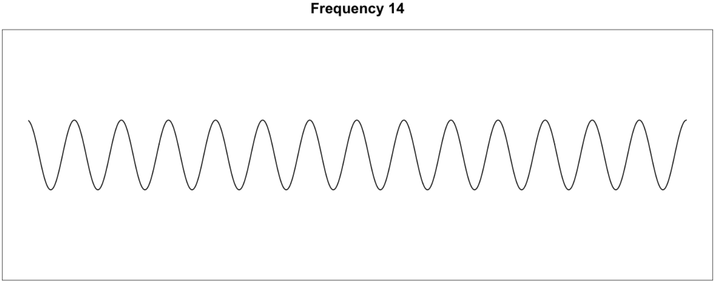
Frequency 13

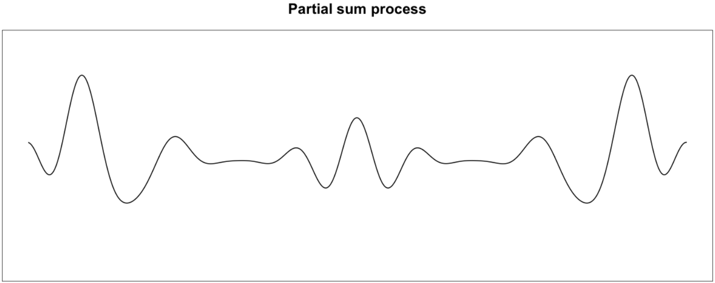


Partial sum process



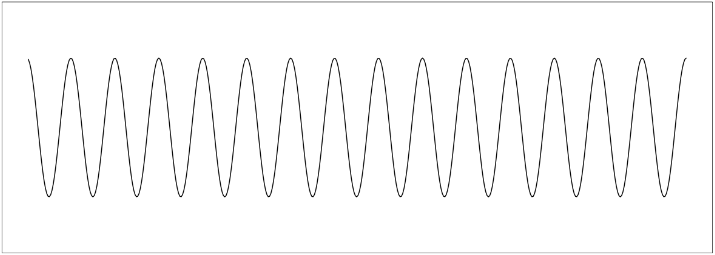
Component and partial sums



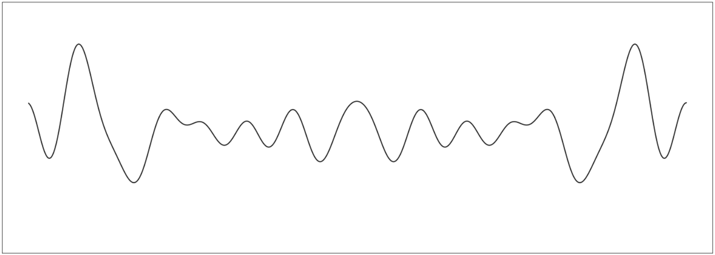


Component and partial sums

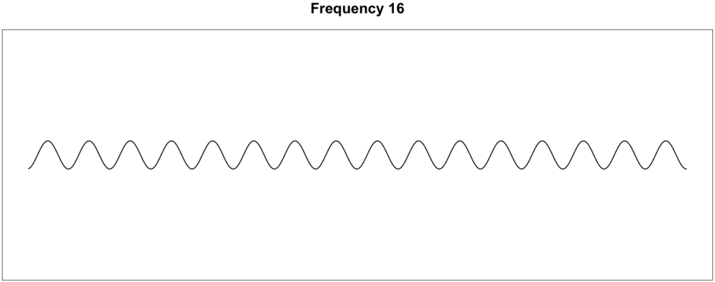
Frequency 15

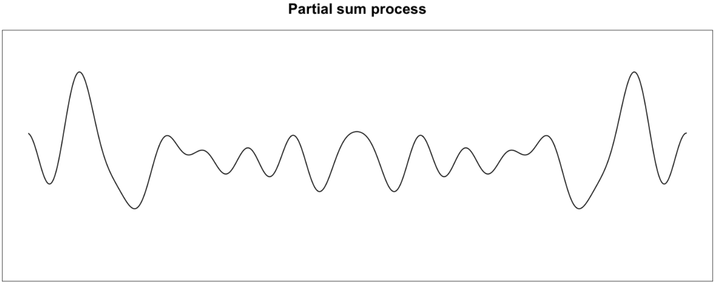


Partial sum process

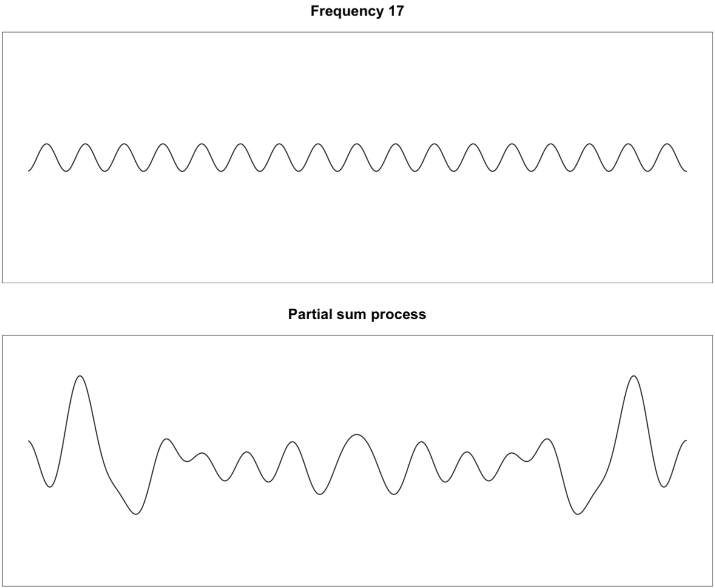


Component and partial sums

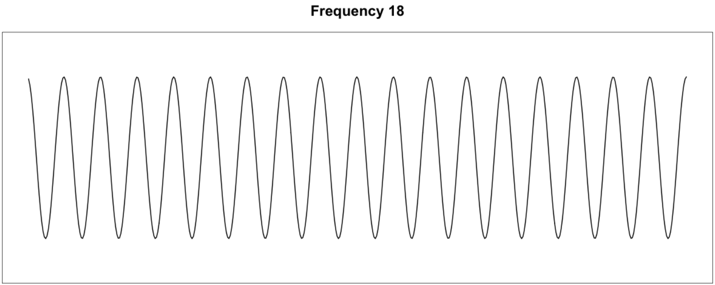


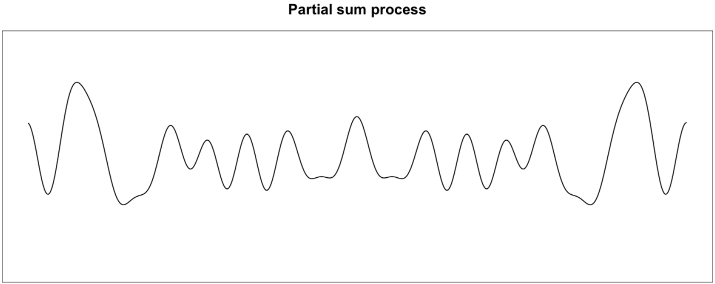


Component and partial sums

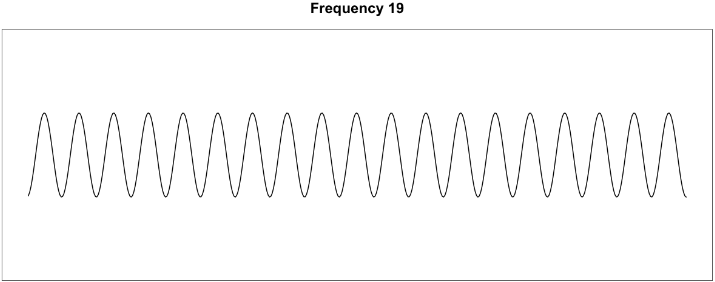


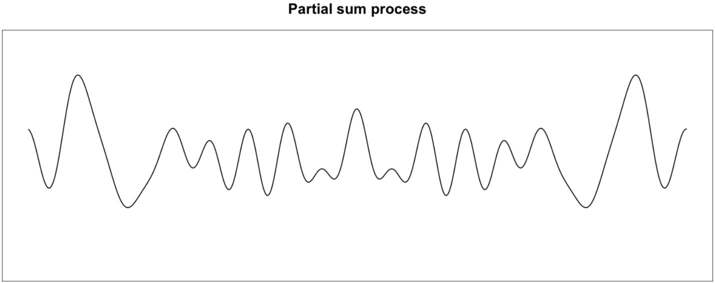
Component and partial sums



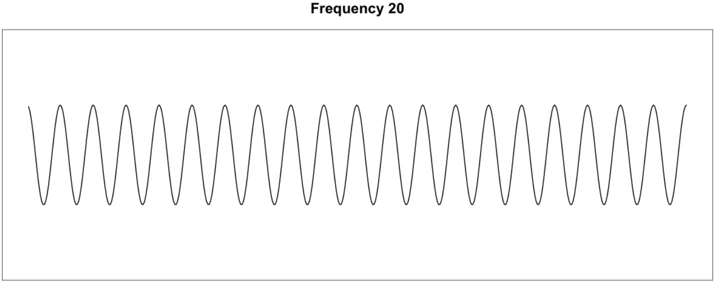


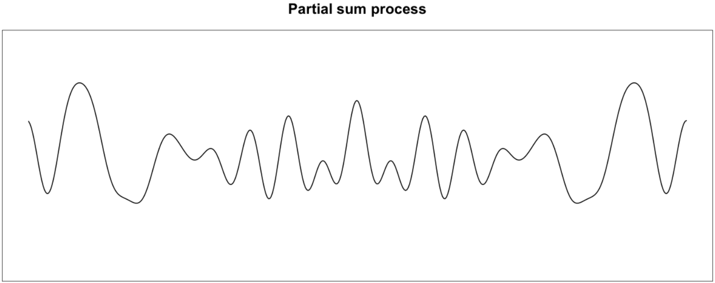
Component and partial sums



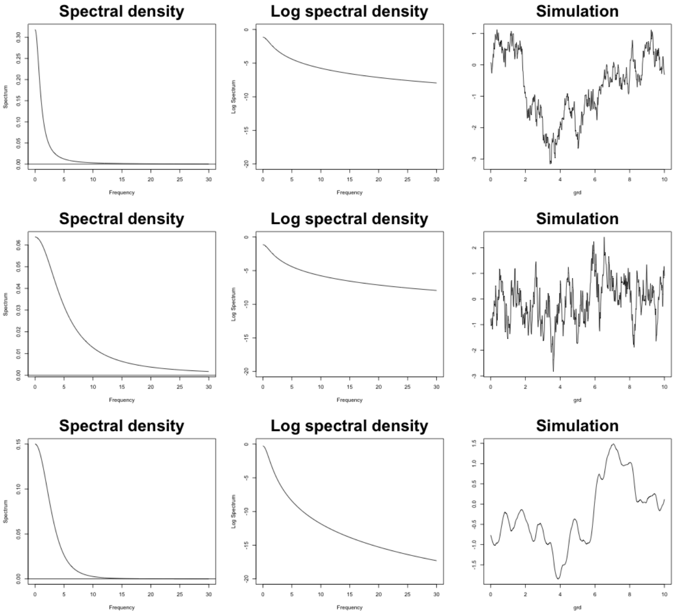


Component and partial sums

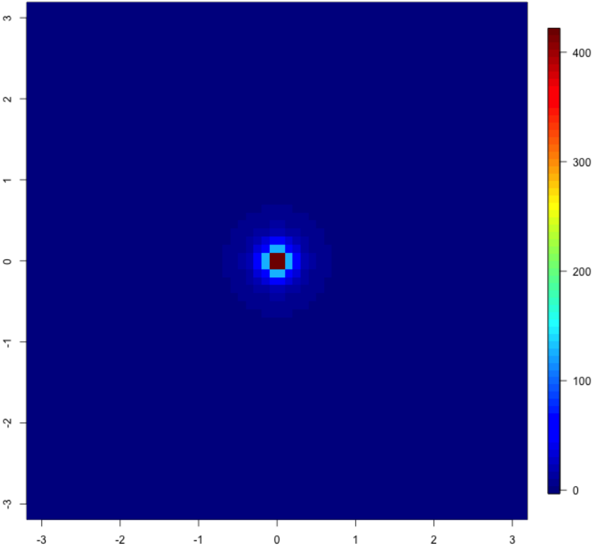




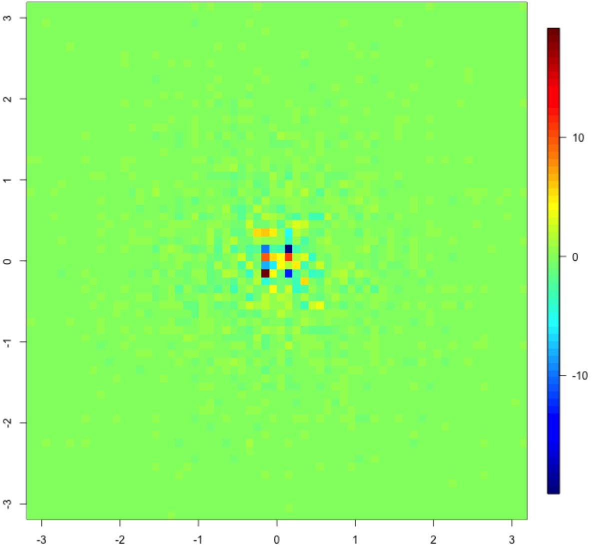
Interpreting spectral densities



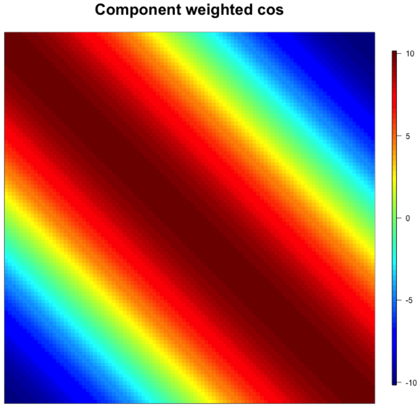
2D spectrum

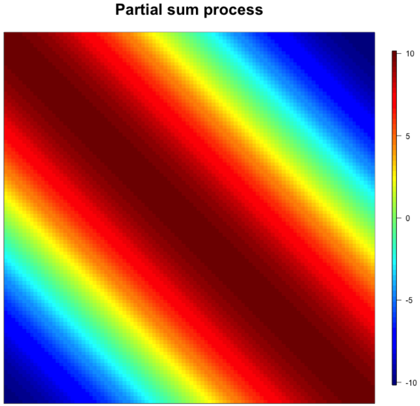


2D random amplitudes

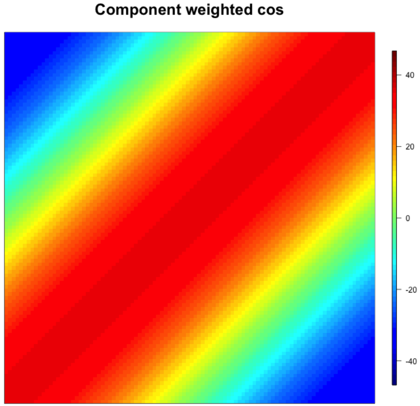


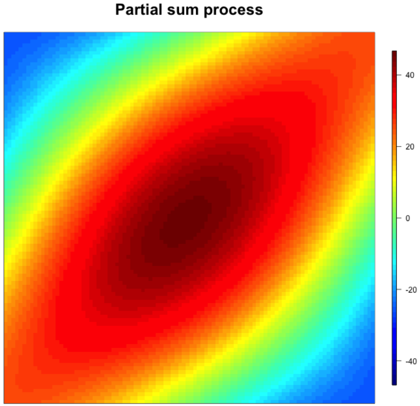
2D component and partial sums



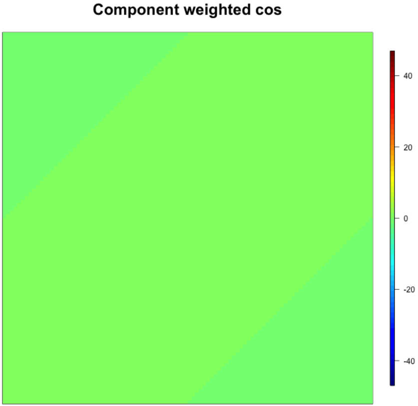


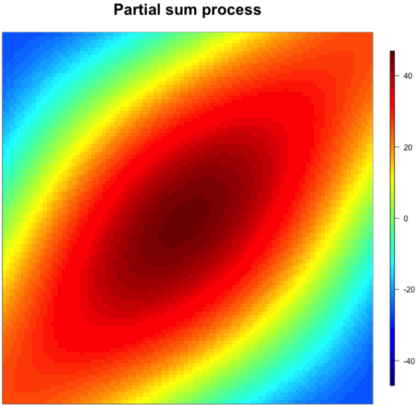
2D component and partial sums



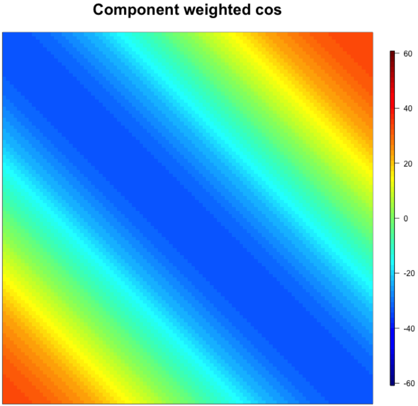


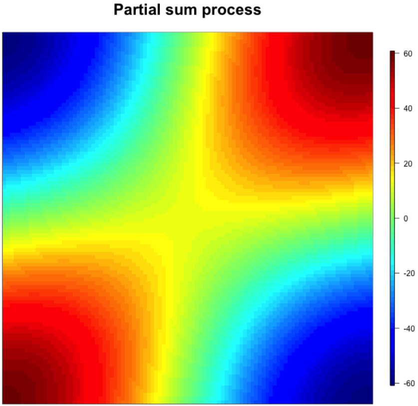
2D component and partial sums



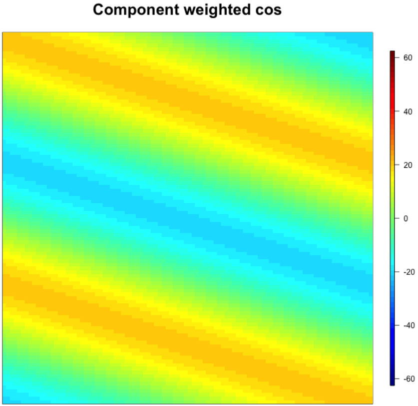


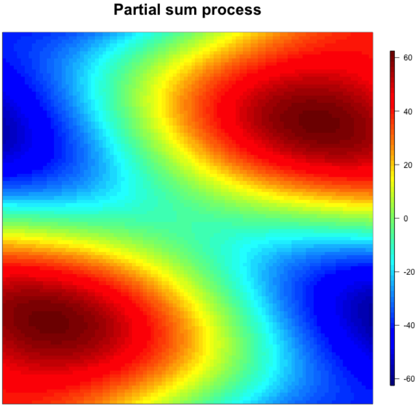
2D component and partial sums



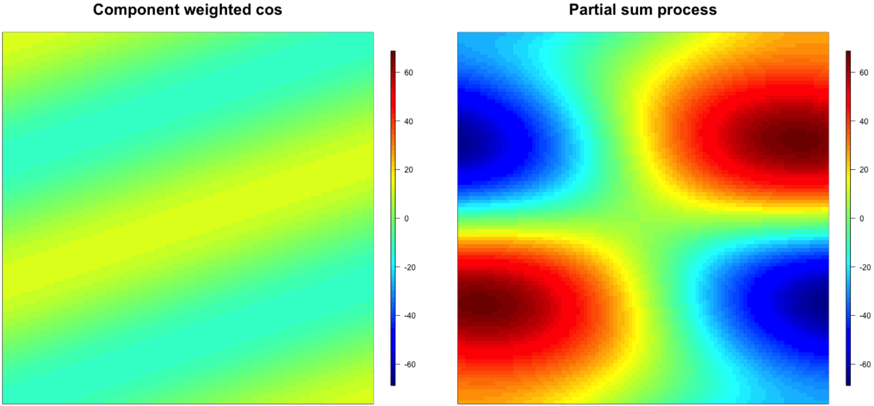


2D component and partial sums

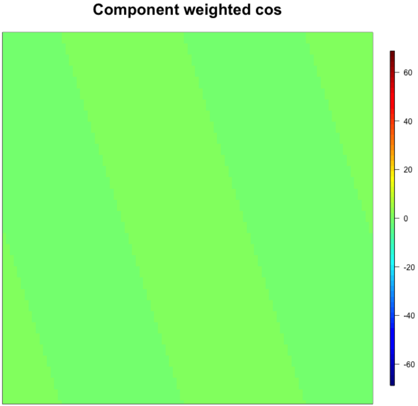


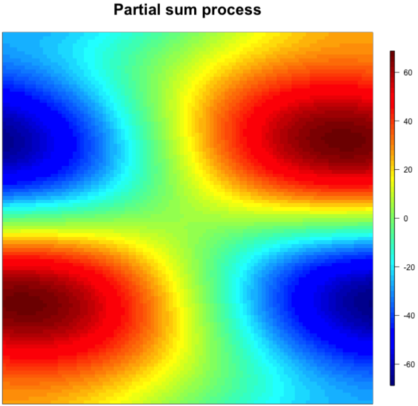


2D component and partial sums



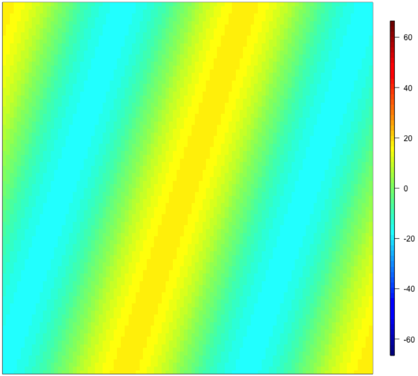
2D component and partial sums



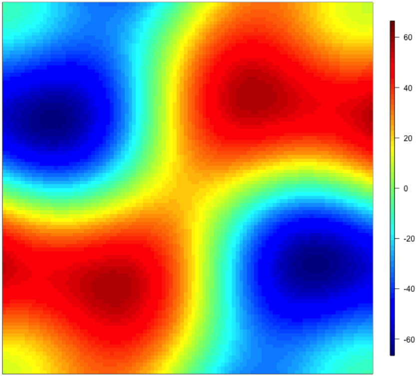


2D component and partial sums

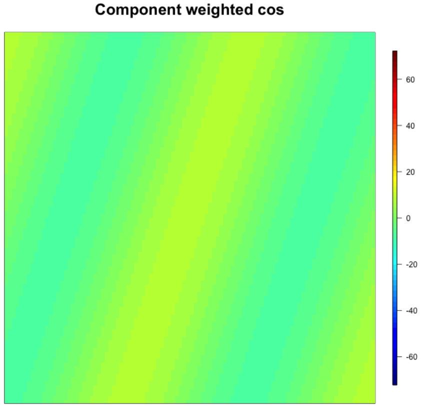
Component weighted cos

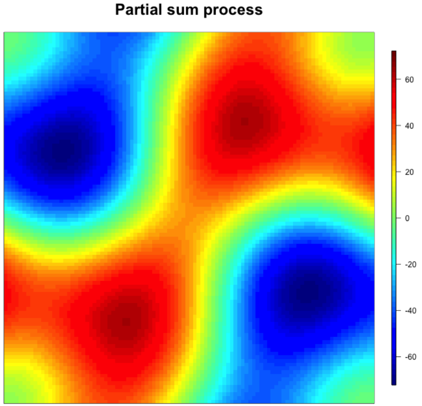


Partial sum process

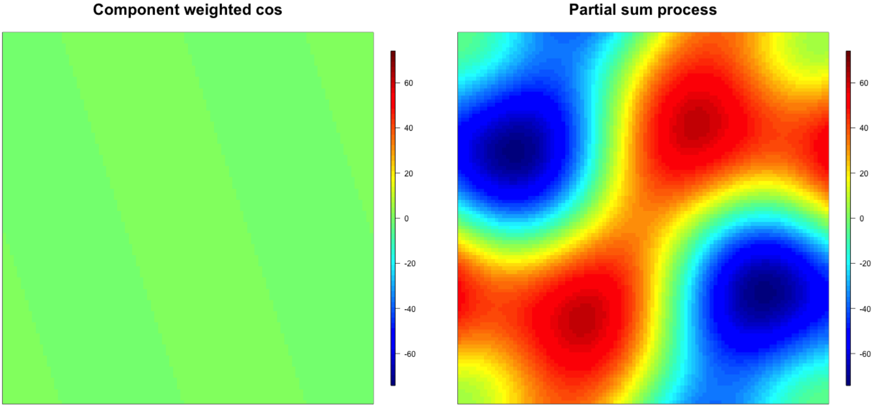


2D component and partial sums

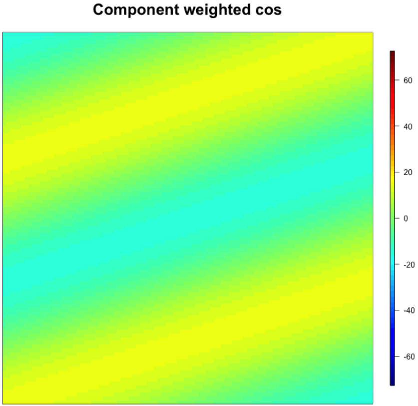


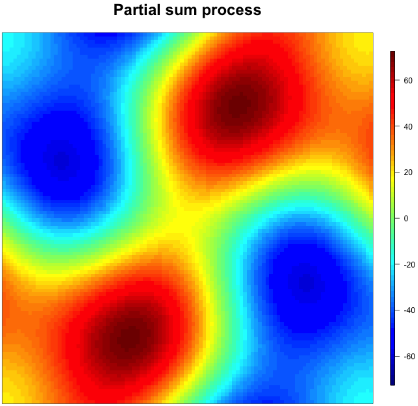


2D component and partial sums

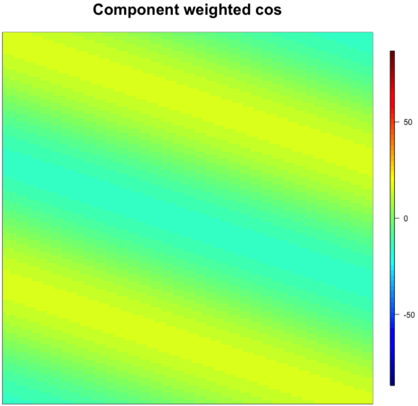


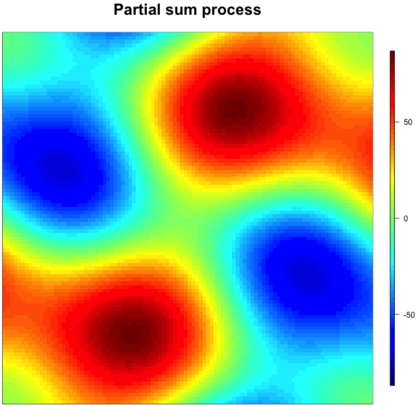
2D component and partial sums



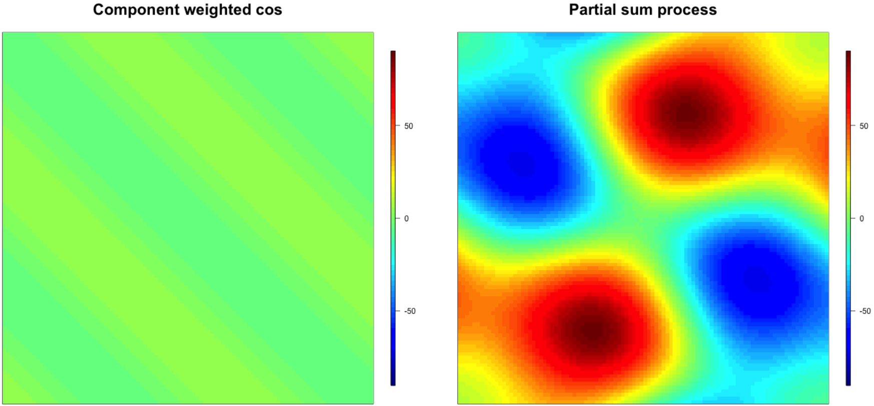


2D component and partial sums

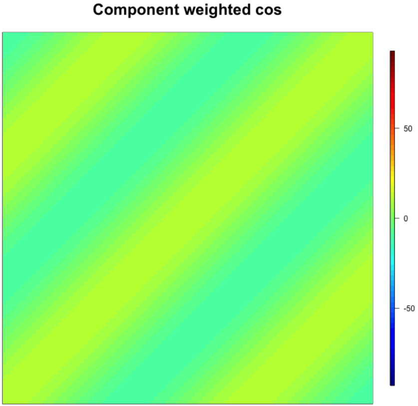


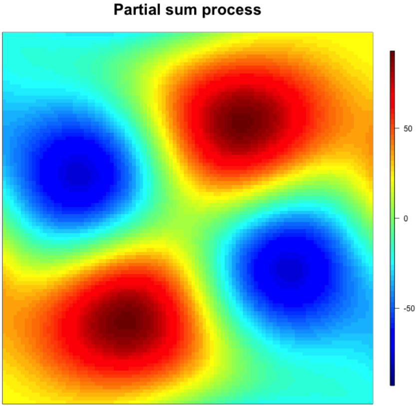


2D component and partial sums

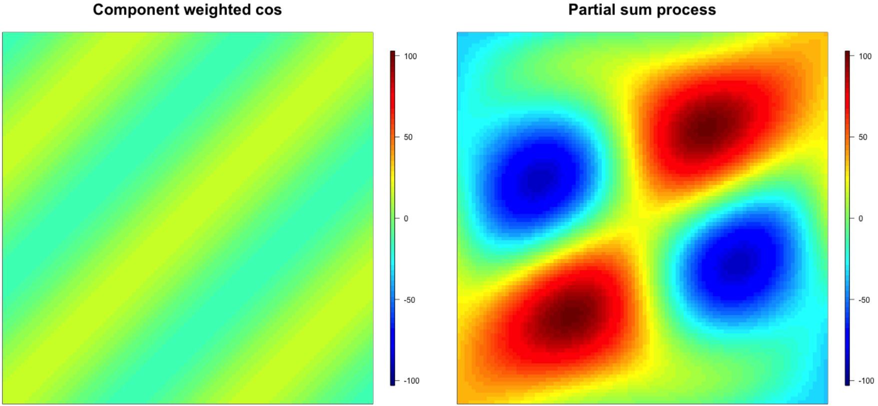


2D component and partial sums

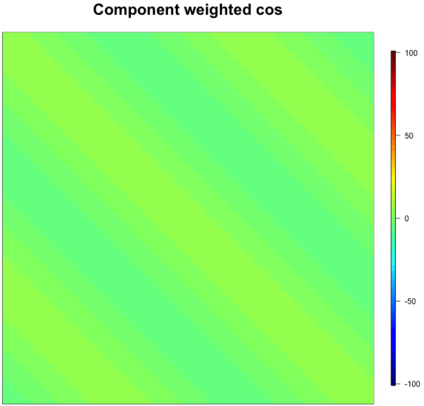


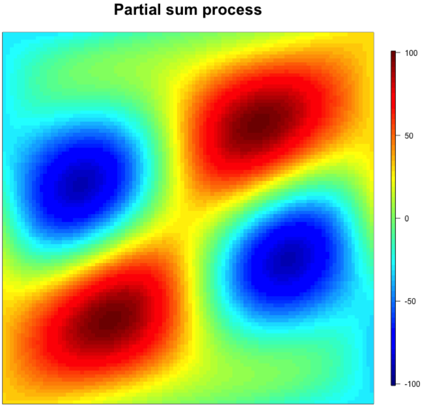


2D component and partial sums

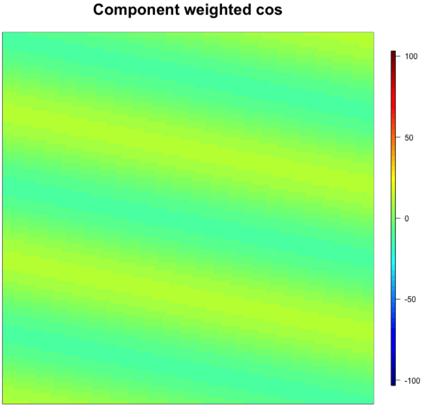


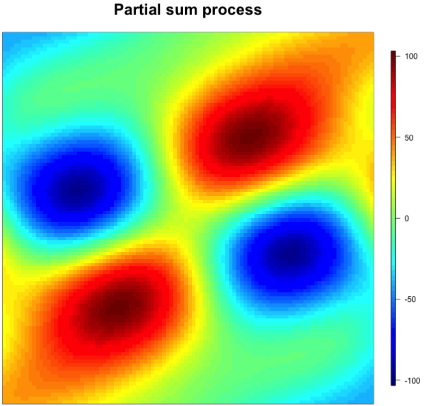
2D component and partial sums



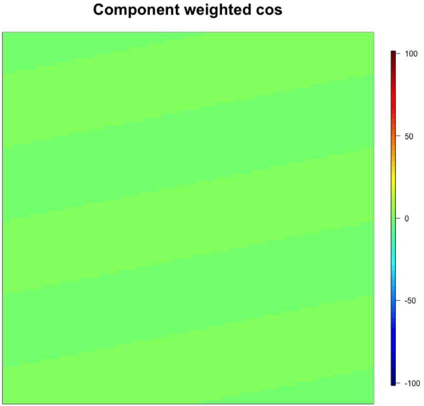


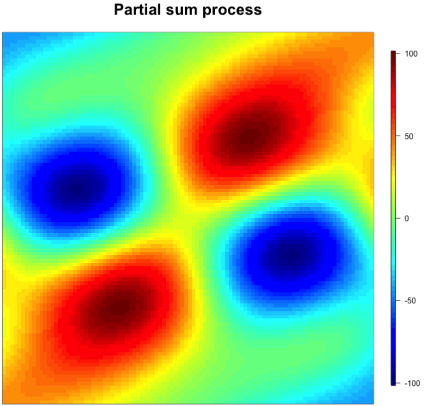
2D component and partial sums



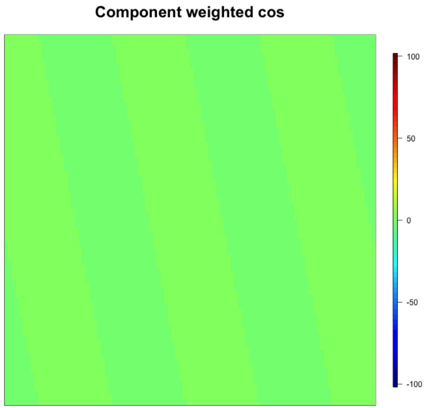


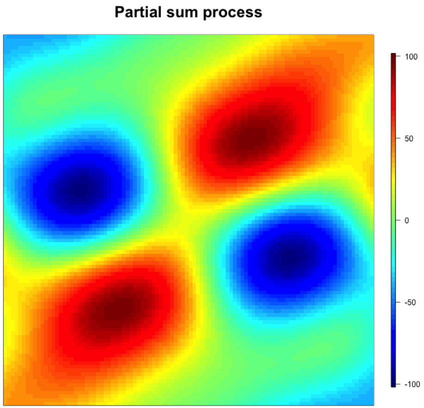
2D component and partial sums



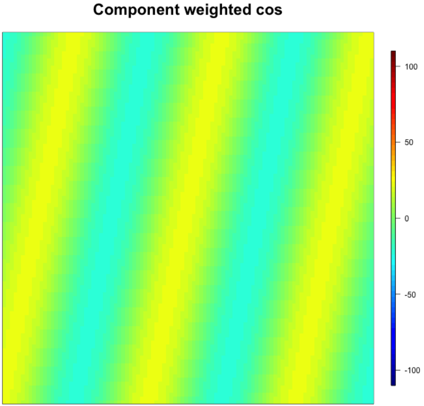


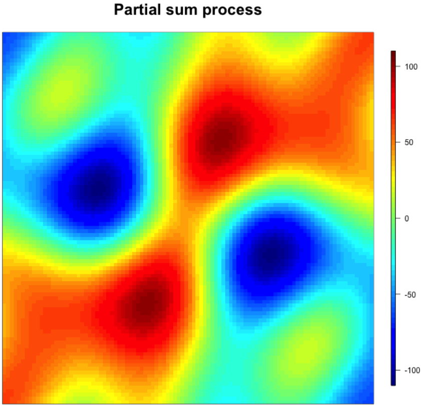
2D component and partial sums



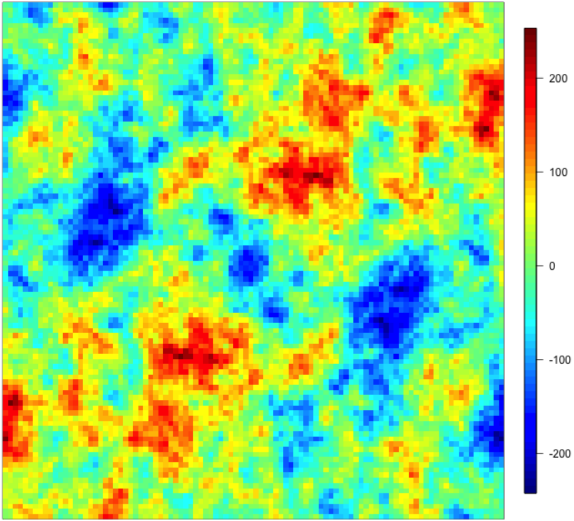


2D component and partial sums





2D final simulation



Spectra for multivariate random fields

Cramér's theorem: $\mathbf{C}(\mathbf{h})$ (multivariate!) is a nonnegative definite function iff

$$\mathbf{f}(\boldsymbol{\omega}) = \begin{pmatrix} f_{11}(\boldsymbol{\omega}) & f_{12}(\boldsymbol{\omega}) & \cdots & f_{1p}(\boldsymbol{\omega}) \\ f_{21}(\boldsymbol{\omega}) & f_{22}(\boldsymbol{\omega}) & \cdots & f_{2p}(\boldsymbol{\omega}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{p1}(\boldsymbol{\omega}) & f_{p2}(\boldsymbol{\omega}) & \cdots & f_{pp}(\boldsymbol{\omega}) \end{pmatrix}$$

is nonnegative definite for every ω where

$$f_{ij}(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} C_{ij}(\mathbf{h}) \exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{h})\mathrm{d}\mathbf{h}$$

- $f_{ii}(\boldsymbol{\omega})$ is the spectral density for $C_{ii}(\mathbf{h})$
- $f_{ij}(\boldsymbol{\omega})$ is the cross-spectral density for $C_{ij}(\mathbf{h})$.

Interpreting matrix spectra

Spectral representation theorem:

$$\begin{aligned}\mathbf{Z}(\mathbf{s}) &= \int \exp(i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{s})\,\mathbf{M}(\mathrm{d}\boldsymbol{\omega}) \\ &\approx \sum_i \cos(\boldsymbol{\omega}_i^{\mathrm{T}}\mathbf{s} + \phi_i)\,\mathbf{M}(\boldsymbol{\omega}_i)\Delta \qquad \text{(}\mathbf{Z}\text{ is real valued)} \\ &= \sum_i \cos(\boldsymbol{\omega}_i^{\mathrm{T}}\mathbf{s} + \phi_i)\,\mathbf{M}_i\end{aligned}$$

where

$$\operatorname{Var}\mathbf{M}_i = \mathbf{f}(\boldsymbol{\omega}_i)\Delta$$

$\Rightarrow \mathbf{f}(\boldsymbol{\omega})$ is the covariance matrix for the random vector coefficient at frequency $\boldsymbol{\omega}$.

(side note: it's slightly more complicated since $\mathbf{M}$ is a complex measure)

Spectral coherence

Something a bit awkward: cross-spectral densities are complex-valued.

Spectral coherence

Something a bit awkward: cross-spectral densities are complex-valued.

Define the (absolute) coherence between Z_1 and Z_2 to be

$$\gamma(\boldsymbol{\omega}) = \frac{|f_{12}(\boldsymbol{\omega})|}{\sqrt{f_{11}(\boldsymbol{\omega}) f_{22}(\boldsymbol{\omega})}}$$

noting

$$\gamma(\boldsymbol{\omega}) \in [0, 1].$$

Spectral coherence

Something a bit awkward: cross-spectral densities are complex-valued.

Define the (absolute) coherence between Z_1 and Z_2 to be

$$\gamma(\omega) = \frac{|f_{12}(\omega)|}{\sqrt{f_{11}(\omega) f_{22}(\omega)}}$$

noting

$$\gamma(\omega) \in [0, 1].$$

$\gamma(\omega)$ is the unsigned correlation coefficient between $Z_1(\mathbf{s})$ and $Z_2(\mathbf{s})$ at frequency ω.

Connection to prediction

Suppose $(Z_1(\mathbf{s}), Z_2(\mathbf{s}))$ is stationary with spectral density matrix $\mathbf{f}(\boldsymbol{\omega}) = (f_{ij}(\boldsymbol{\omega}))_{i,j=1}^2$. The $K(\mathbf{u})$ that minimizes

$$\mathbb{E}\left|Z_1(\mathbf{s}_0) - \int_{\mathbb{R}^d} K(\mathbf{u}-\mathbf{s}_0) Z_2(\mathbf{u})\mathrm{d}\mathbf{u}\right|^2$$

is

$$K(\mathbf{u}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\sqrt{\frac{f_{11}(\boldsymbol{\omega})}{f_{22}(\boldsymbol{\omega})}}\gamma(\boldsymbol{\omega})\exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{u})\mathrm{d}\boldsymbol{\omega}.$$

Connection to prediction

Suppose $(Z_1(\mathbf{s}), Z_2(\mathbf{s}))$ is stationary with spectral density matrix $\mathbf{f}(\boldsymbol{\omega}) = (f_{ij}(\boldsymbol{\omega}))_{i,j=1}^2$. The $K(\mathbf{u})$ that minimizes

$$\mathbb{E}\left|Z_1(\mathbf{s}_0) - \int_{\mathbb{R}^d} K(\mathbf{u}-\mathbf{s}_0)Z_2(\mathbf{u})\mathrm{d}\mathbf{u}\right|^2$$

is

$$K(\mathbf{u}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\sqrt{\frac{f_{11}(\boldsymbol{\omega})}{f_{22}(\boldsymbol{\omega})}}\gamma(\boldsymbol{\omega})\exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{u})\mathrm{d}\boldsymbol{\omega}.$$

Looks a lot like a conditional expectation for a bivariate normal…

Simple coherence example

Working in $d = 1$, suppose

$$Z_1(s) = U_1 \cos(\omega_0 s)$$
$$Z_2(s) = U_1 \cos(\omega_0 s) + U_2 \cos(\omega_1 s)$$

for $\omega_0 \neq \omega_1$ and U_1 and U_2 mean zero and uncorrelated. Then

$$\gamma(\omega) = \begin{cases} 1 & \omega = \omega_0 \\ 0 & \text{otherwise} \end{cases}$$

Exercise: check!

A slightly more complicated coherence example

Build

$$\begin{pmatrix} Z_1(\mathbf{s}) \\ Z_2(\mathbf{s}) \end{pmatrix} = \sum_{j=1}^{L} \begin{pmatrix} c_{1j} \\ c_{2j} \end{pmatrix} \cos(2\pi \boldsymbol{\omega}_j^{\mathrm{T}} \mathbf{s})$$

where let's imagine the frequencies are unique and sorted:

$$\|\boldsymbol{\omega}_1\| < \|\boldsymbol{\omega}_2\| < \cdots < \|\boldsymbol{\omega}_L\|.$$

If we set

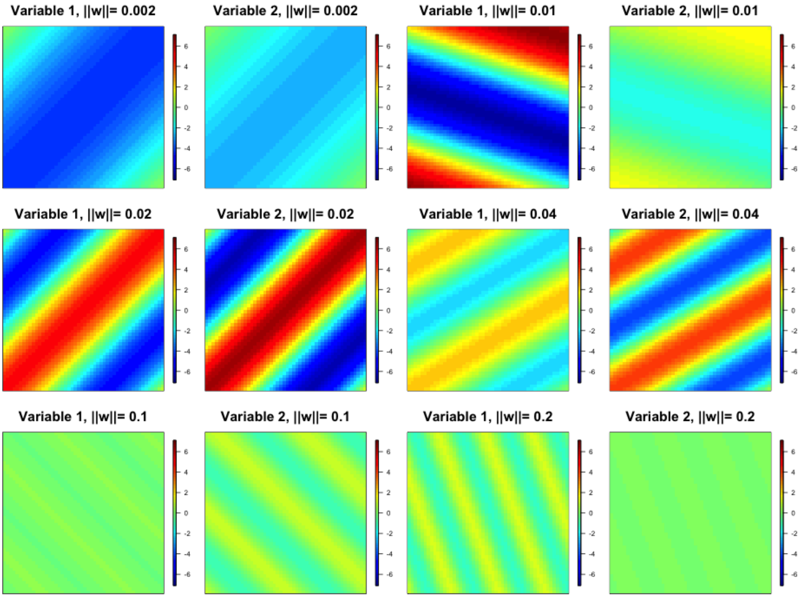
$$\mathrm{Cov} \begin{pmatrix} c_{1j} \\ c_{2j} \end{pmatrix} = \frac{1}{(1 + \|\boldsymbol{\omega}_j\|^2)^{1.5}} \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$$

then

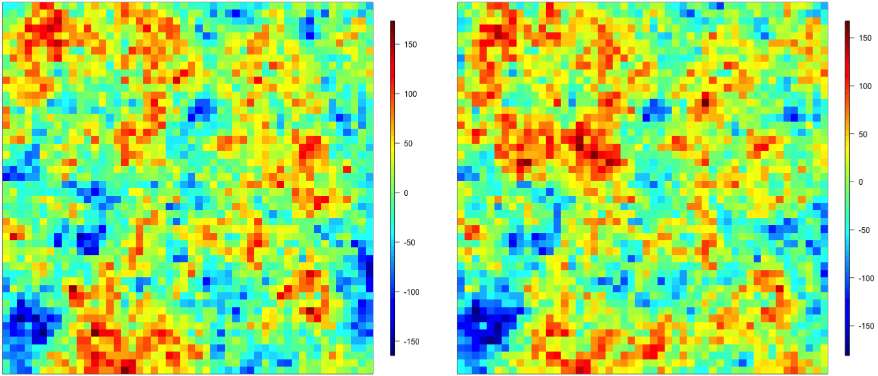
$$\gamma(\boldsymbol{\omega}_j) \equiv 0.8$$

for all j.

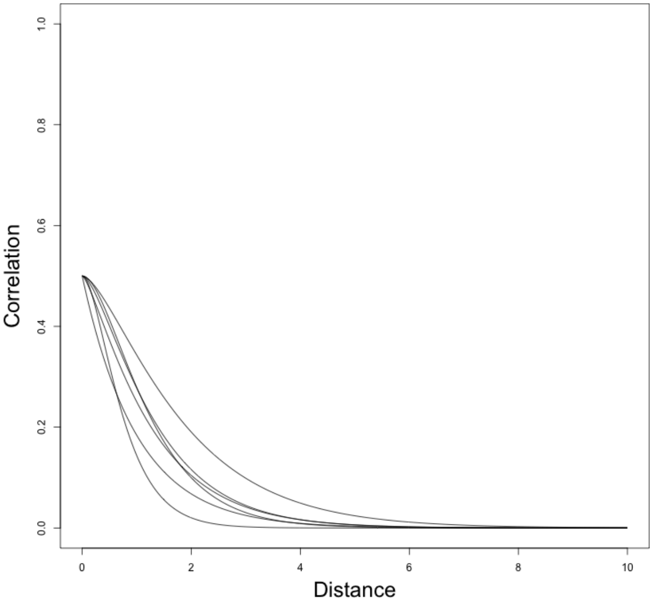
Bivariate spectral building



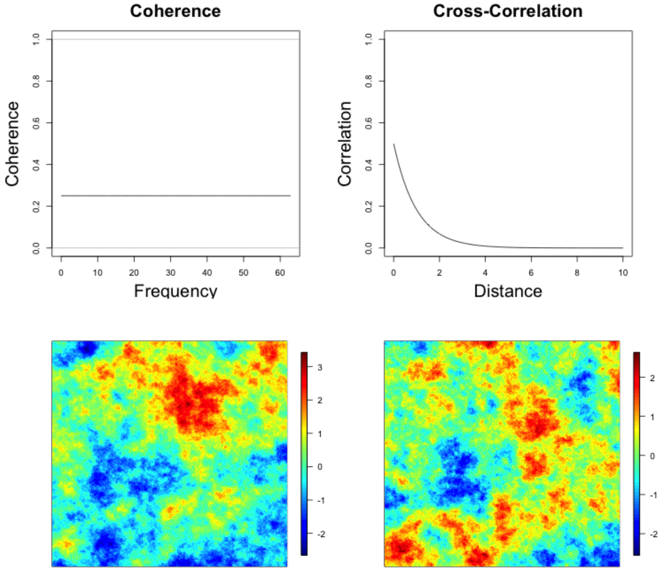
Bivariate spectral building



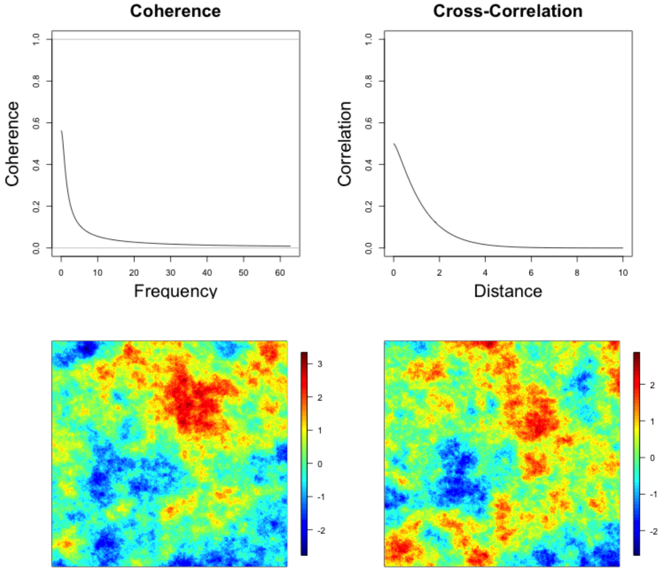
Cross-correlations



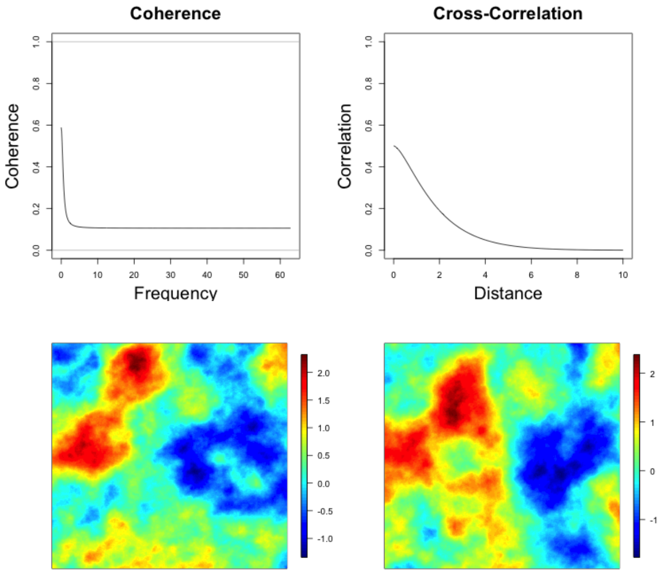
Coherence vs. cross-correlation



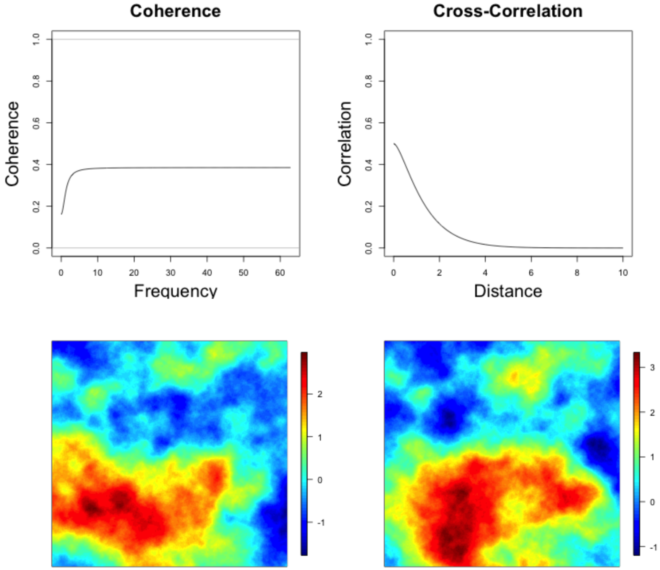
Coherence vs. cross-correlation



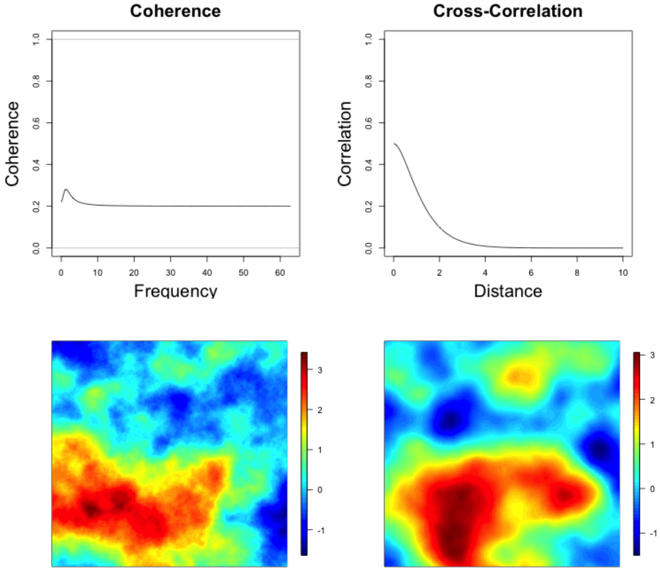
Coherence vs. cross-correlation



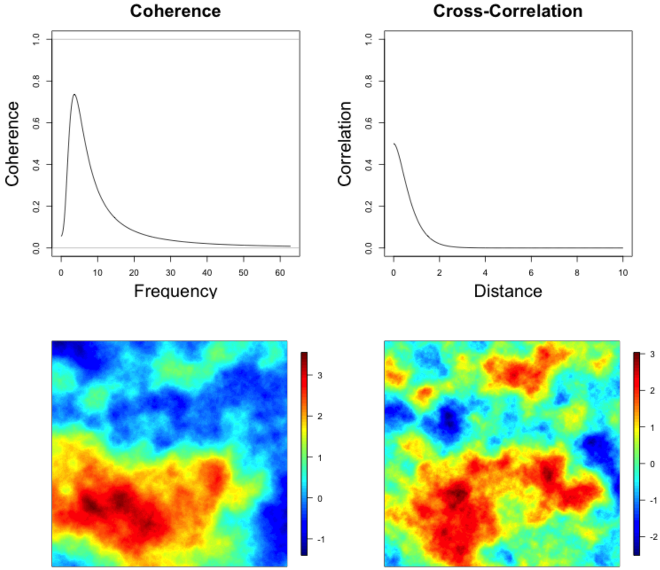
Coherence vs. cross-correlation



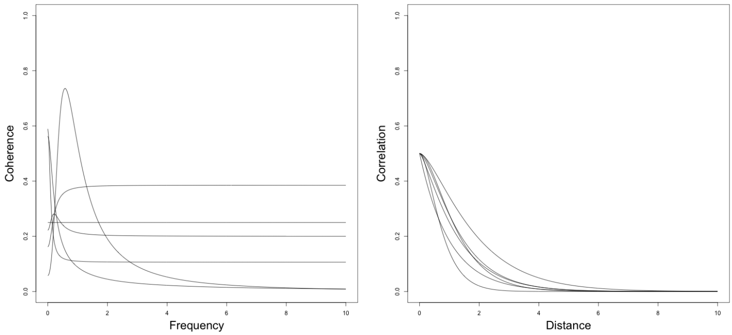
Coherence vs. cross-correlation



Coherence vs. cross-correlation



Coherence vs. cross-correlation



Loosely speaking,

- Cross-covariance "smoothness" controls the rate of decay of coherence at high frequencies
- Cross-covariance "range" controls the frequency of maximal coherence

Spectral estimation setup

Suppose the p-variate process $\mathbf{Z}(\mathbf{s})$ has been observed on a regular integer-valued grid in $\mathbb{R}^2$

$$\mathbf{s} = \begin{pmatrix} j \\ k \end{pmatrix}$$

with $j, k = 1, 2, \ldots, n$ (in particular, n is the number of points in one axial direction).

Spectral estimation setup

Suppose the p-variate process $\mathbf{Z}(\mathbf{s})$ has been observed on a regular integer-valued grid in $\mathbb{R}^2$

$$\mathbf{s} = \begin{pmatrix} j \\ k \end{pmatrix}$$

with $j, k = 1, 2, \ldots, n$ (in particular, n is the number of points in one axial direction).

The next few slides will be a test of patience with:

- i is the imaginary number
- j, k will be integers used for either a spatial or spectral location
- ℓ, m will be processes

Estimation: periodogram

The periodogram matrix is $\mathbf{I}(\boldsymbol{\omega}) = (I_{\ell m}(\boldsymbol{\omega}))_{\ell,m=1}^{p}$ where

$$I_{\ell m}(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d n} \left(\sum_{j=1}^{n} Z_\ell(\mathbf{s}_j) \exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega}) \right) \overline{\left(\sum_{j=1}^{n} Z_m(\mathbf{s}_j) \exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega}) \right)}.$$

Under mixed increasing domain/infill asymptotics,

(i) $\mathbb{E} I_{\ell m}(\boldsymbol{\omega}) \to f_{\ell m}(\boldsymbol{\omega})$,

(ii) $\mathrm{Var} I_{\ell m}(\boldsymbol{\omega}) \to f_{\ell m}(\boldsymbol{\omega})^2$ and

(iii) $\mathrm{Cov}(I_{\ell m}(\boldsymbol{\omega}_1), I_{\ell m}(\boldsymbol{\omega}_2)) \to 0$ for $\boldsymbol{\omega}_1 \neq \boldsymbol{\omega}_2$.

Estimation: periodogram

The periodogram matrix is $\mathbf{I}(\boldsymbol{\omega}) = (I_{\ell m}(\boldsymbol{\omega}))_{\ell,m=1}^{p}$ where

$$I_{\ell m}(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d n}\left(\sum_{j=1}^{n} Z_\ell(\mathbf{s}_j)\exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)\overline{\left(\sum_{j=1}^{n} Z_m(\mathbf{s}_j)\exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)}.$$

Under mixed increasing domain/infill asymptotics,

(i) $\mathbb{E}I_{\ell m}(\boldsymbol{\omega}) \to f_{\ell m}(\boldsymbol{\omega})$,

(ii) $\mathrm{Var}I_{\ell m}(\boldsymbol{\omega}) \to f_{\ell m}(\boldsymbol{\omega})^2$ and

(iii) $\mathrm{Cov}(I_{\ell m}(\boldsymbol{\omega}_1), I_{\ell m}(\boldsymbol{\omega}_2)) \to 0$ for $\boldsymbol{\omega}_1 \neq \boldsymbol{\omega}_2$.

So if we kernel smooth $I_{\ell m}$ over a little window of frequencies then $I_{\ell m}$ is a consistent estimator of $f_{\ell m}$.

Estimating spectra and coherence

If we can calculate $\{I_{\ell m}\}$ and kernel smooth them to get $\{\tilde{I}_{\ell m}\}$ then

$$\hat{\gamma}(\omega) = \frac{\tilde{I}_{12}(\omega)}{\sqrt{\tilde{I}_{11}(\omega)\tilde{I}_{22}(\omega)}}$$

Estimating spectra and coherence

If we can calculate $\{I_{\ell m}\}$ and kernel smooth them to get $\{\tilde{I}_{\ell m}\}$ then

$$\hat{\gamma}(\boldsymbol{\omega}) = \frac{\tilde{I}_{12}(\boldsymbol{\omega})}{\sqrt{\tilde{I}_{11}(\boldsymbol{\omega})\tilde{I}_{22}(\boldsymbol{\omega})}}$$

It's also important to kernel smooth the periodograms if you want to estimate coherence because if you don't,

$$\frac{\left|\left(\sum_{j=1}^{n} Z_1(\mathbf{s}_j)\exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)\overline{\left(\sum_{j=1}^{n} Z_2(\mathbf{s}_j)\exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)}\right|}{\sqrt{\left(\sum_{j=1}^{n} Z_1(\mathbf{s}_j)\exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)\overline{\left(\sum_{j=1}^{n} Z_1(\mathbf{s}_j)\exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)}\left(\sum_{j=1}^{n} Z_2(\mathbf{s}_j)\exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)\overline{\left(\sum_{j=1}^{n} Z_2(\mathbf{s}_j)\exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)}}}$$

is exactly 1.

Fourier frequencies

In practice this is done at the Fourier frequencies

$$\boldsymbol{\omega} = \begin{pmatrix} \frac{2\pi j}{n} \\ \frac{2\pi k}{n} \end{pmatrix} \quad \text{for} \quad j, k = 0, 1, \ldots, n-1$$

which results in n^2 Fourier frequencies.

(Side note: spectral methods are a big pain due to how different authors index their Fourier frequencies and standardize with the 2π)

Your new best friend: the discrete Fourier transform

The first piece in the periodogram is

$$\sum_{j=1}^{n} Z_\ell(\mathbf{s}_j) \exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})$$

is the discrete Fourier transform of $(Z_\ell(\mathbf{s}_1), \ldots, Z_\ell(\mathbf{s}_{n^2}))^{\mathrm{T}}$ when $\boldsymbol{\omega}$ is a Fourier frequency.

This is quickly calculated using the fast Fourier transform (FFT) in `R`, `fft(Z.ell)`.

The second piece is

$$\overline{\left(\sum_{j=1}^{n} Z_m(\mathbf{s}_j) \exp(-i\mathbf{s}_j^{\mathrm{T}}\boldsymbol{\omega})\right)}$$

which is just as easy: `fft(Z.m,inverse=TRUE)`.


Take this object, but beware: it carries a terrible curse.
Ooh, that's bad.
But it comes with a free Frogurt!
That's good.
The Frogurt is also cursed.
That's bad.
But you get your choice of toppings.
That's good.
The toppings contain potassium benzoate.
That's bad.
Can I go now?

Your new worst enemy: `R` and frequency indexing

Unfortunately doing this in `R` is both easy and hard.

It is easy because `fft` is easy to type.

It is hard because `R` does not actually calculate the DFT as written in the previous slide, it includes an extra factor of $\exp(-2\pi i)$ (see the help file).

But it also easy again because we will always be taking the modulus of DFTs, and the extra factor doesn't affect that.

But it's hard because we have to kernel smooth in multiple dimensions.

But that's easy with convolution.

But you end up reindexing frequencies. See `Coherence.R` for example.

Module 4: some practical advice and examples

Some notes on coding

Given p processes $\{Z_i\}$ co-observed at n locations $\{\mathbf{s}_i\}$ it can sometimes be easier, or more difficult, to work with

$$\begin{pmatrix} Z_1(\mathbf{s}_1) \\ Z_1(\mathbf{s}_2) \\ \vdots \\ Z_1(\mathbf{s}_n) \\ Z_2(\mathbf{s}_1) \\ \vdots \\ Z_p(\mathbf{s}_n) \end{pmatrix} \quad \text{vs.} \quad \begin{pmatrix} Z_1(\mathbf{s}_1) \\ Z_2(\mathbf{s}_1) \\ \vdots \\ Z_p(\mathbf{s}_1) \\ Z_1(\mathbf{s}_2) \\ \vdots \\ Z_p(\mathbf{s}_n) \end{pmatrix}$$

i.e., group by space, or group by variable.

(Since we often calculate these in code using a function applied to distance matrices we prefer the first, while in most papers the latter is usually used because the covariance matrix is blocked by $\mathbf{C}$)

NOAA Global Ensemble Forecast System Reforecast

GEFS reforecast project version 2:

- 2012 version of NCEP's GEFS
- 11-member ensemble, daily from 00 UTC initial conditions
- T254 ($\sim$ 50 km) to 8 days, T190 ($\sim$ 70 km) to 16 days

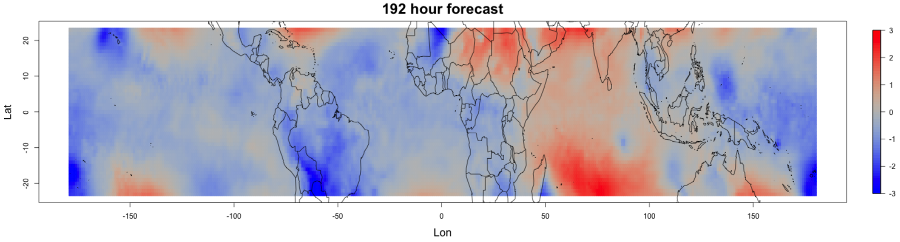
Sea level pressure at forecast horizons:

- 0 hours
- 24 hours, 48 hours, ..., 192 hours (8 days)

over first 90 days of 2014.

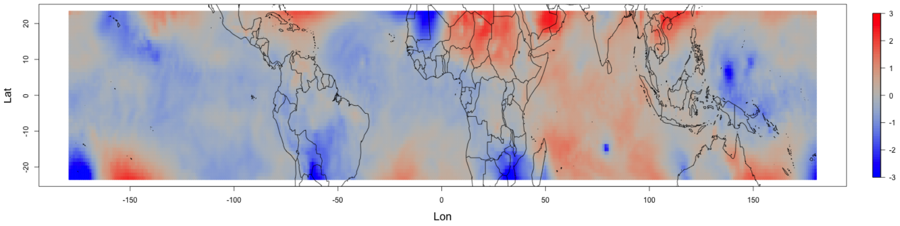
Statistical goal:

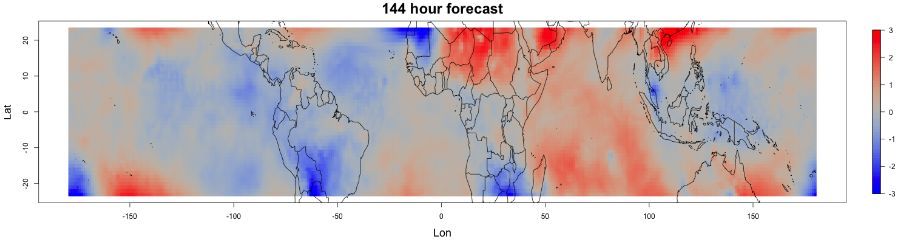
- Quantify the improvement and similarity between forecasts and realizing surfaces

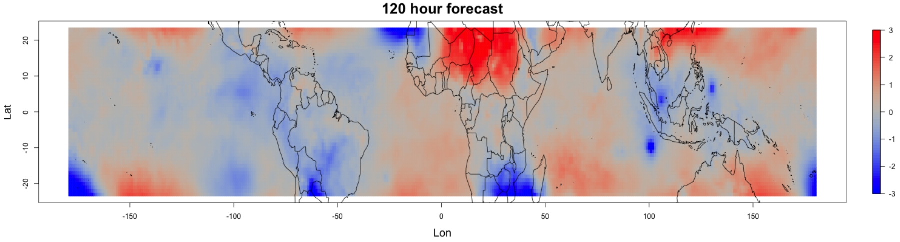


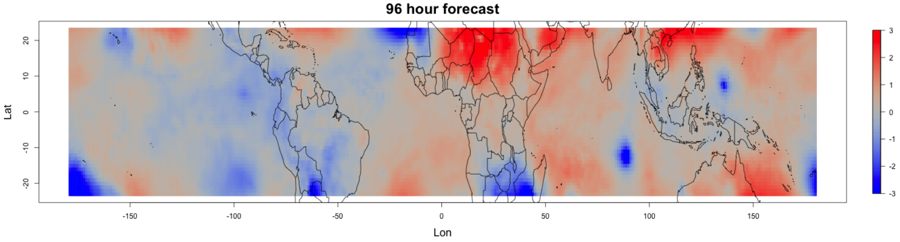
192 hour forecast
Lat
Lon
20
10
0
-10
-20
-150
-100
-50
0
50
100
150
3
2
1
0
-1
-2
-3

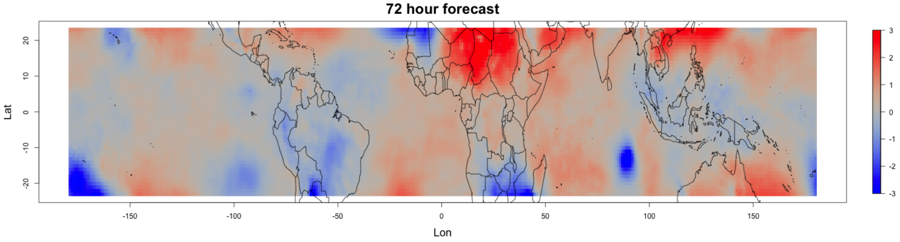
168 hour forecast

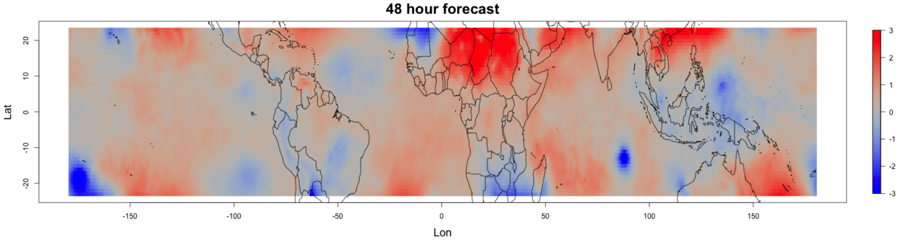



144 hour forecast
Lat
Lon
20
10
0
-10
-20
-150
-100
-50
0
50
100
150
3
2
1
0
-1
-2
-3

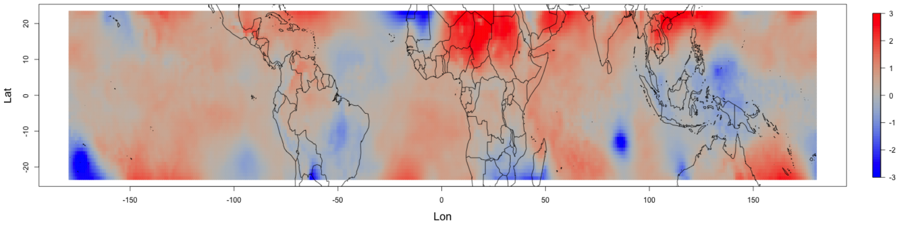

120 hour forecast
Lat
Lon
20
10
0
-10
-20
-150
-100
-50
0
50
100
150
3
2
1
0
-1
-2
-3

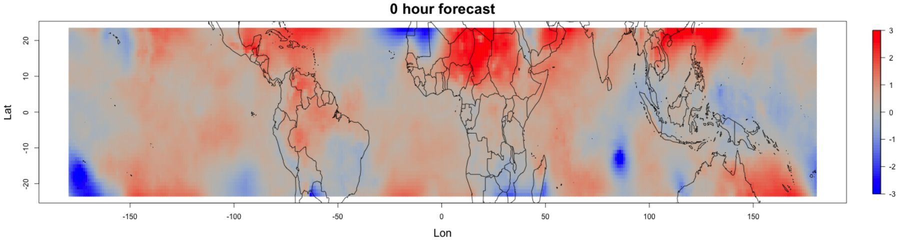

96 hour forecast
Lat
20
10
0
-10
-20
-150
-100
-50
0
50
100
150
Lon
3
2
1
0
-1
-2
-3


72 hour forecast
Lat
20
10
0
-10
-20
-150
-100
-50
0
50
100
150
Lon
3
2
1
0
-1
-2
-3


48 hour forecast
Lat
Lon
20
10
0
-10
-20
-150
-100
-50
0
50
100
150
3
2
1
0
-1
-2
-3

24 hour forecast



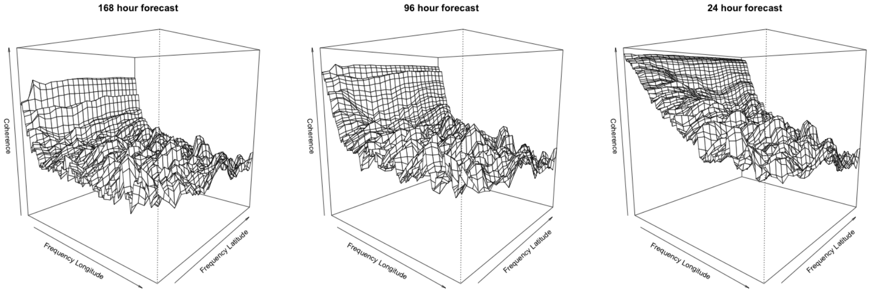

0 hour forecast
Lat
Lon
20
10
0
-10
-20
-150
-100
-50
0
50
100
150
3
2
1
0
-1
-2
-3

Verifying spatial forecasts

Days ahead	8	7	6	5	4	3	2	1
Correlation	0.508	0.663	0.603	0.674	0.737	0.829	0.888	0.918

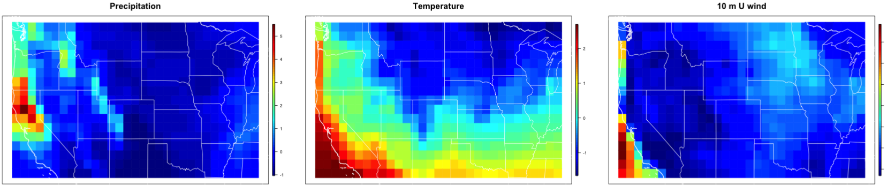
Verifying spatial forecasts

Days ahead	8	7	6	5	4	3	2	1
Correlation	0.508	0.663	0.603	0.674	0.737	0.829	0.888	0.918

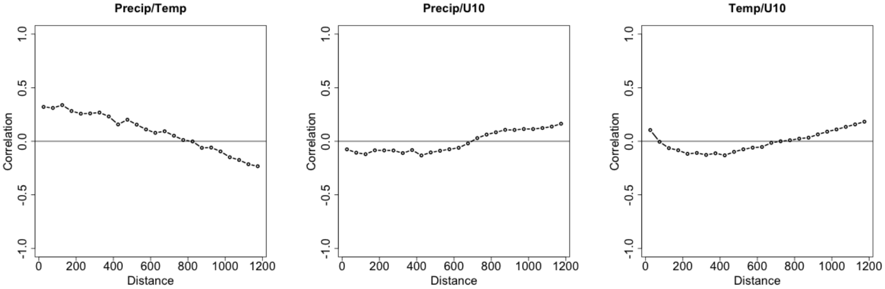


Estimated absolute coherence functions for the GEFS pressure data between (a) 0h and 168h (7 days), (b) 0h and 96h (4 days) and (c) 0h and 24h (1 day).

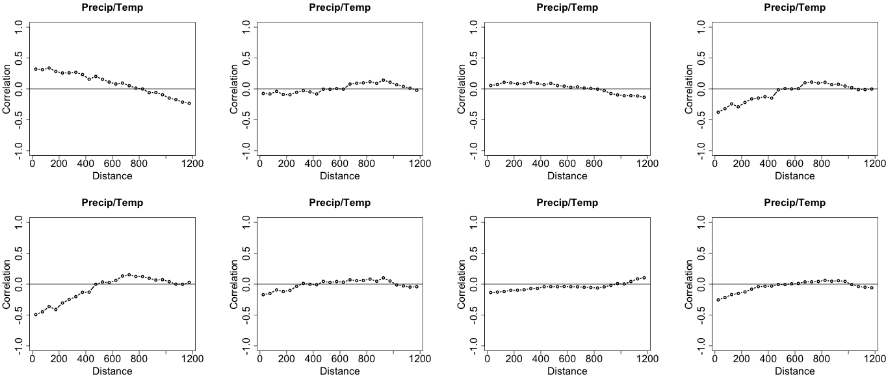
LENS data: Western US



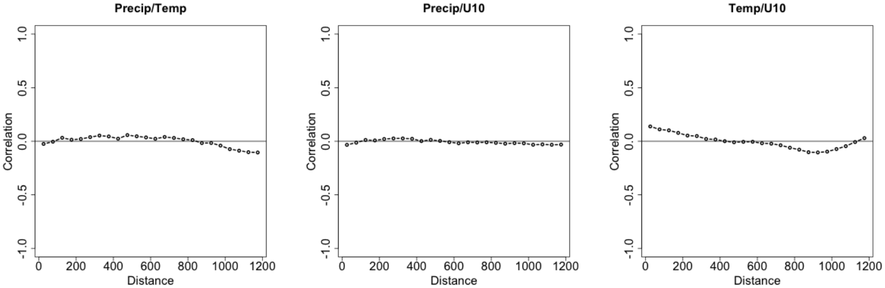
Empirical cross-correlation functions: month 1



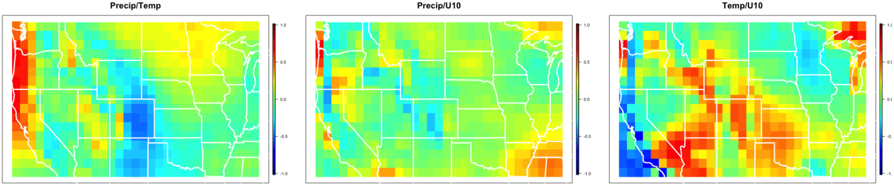
Empirical cross-correlation functions: lots of months



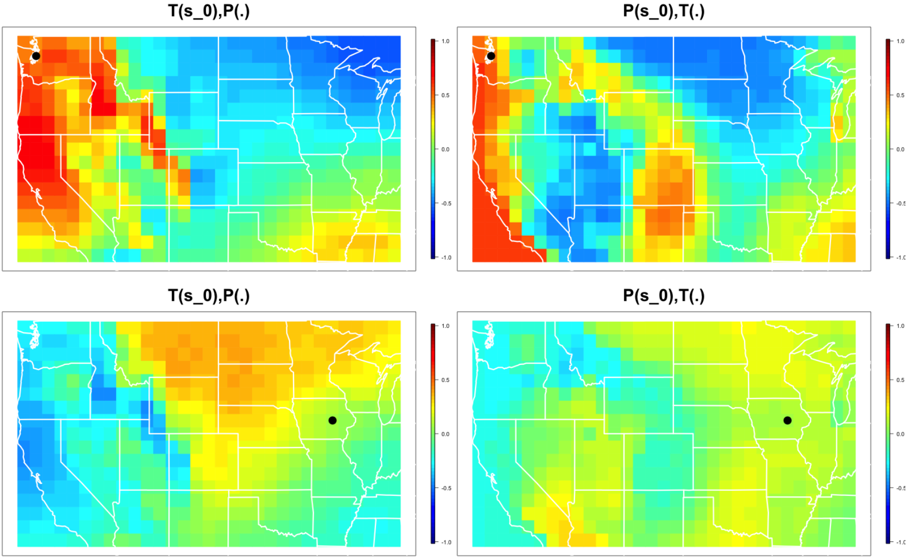
Averaged empirical cross-correlation functions



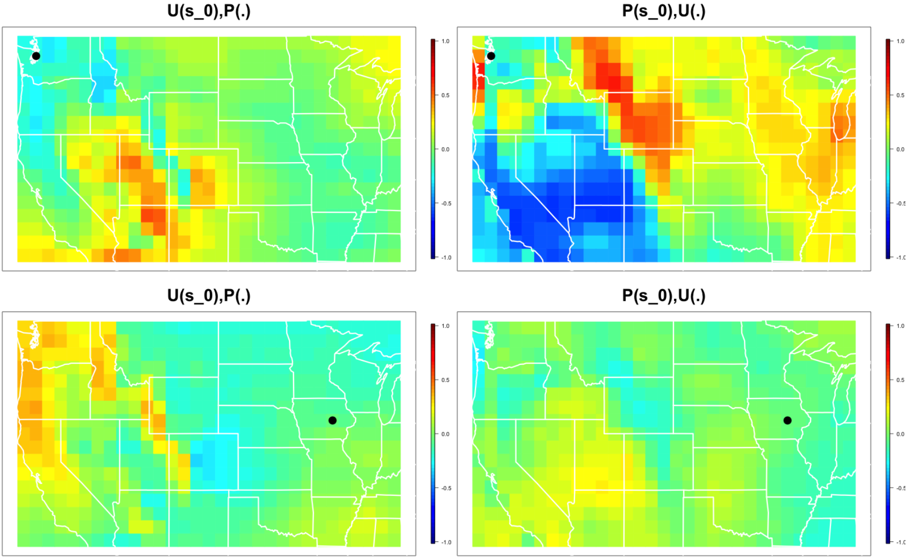
Co-located cross-correlation coefficients



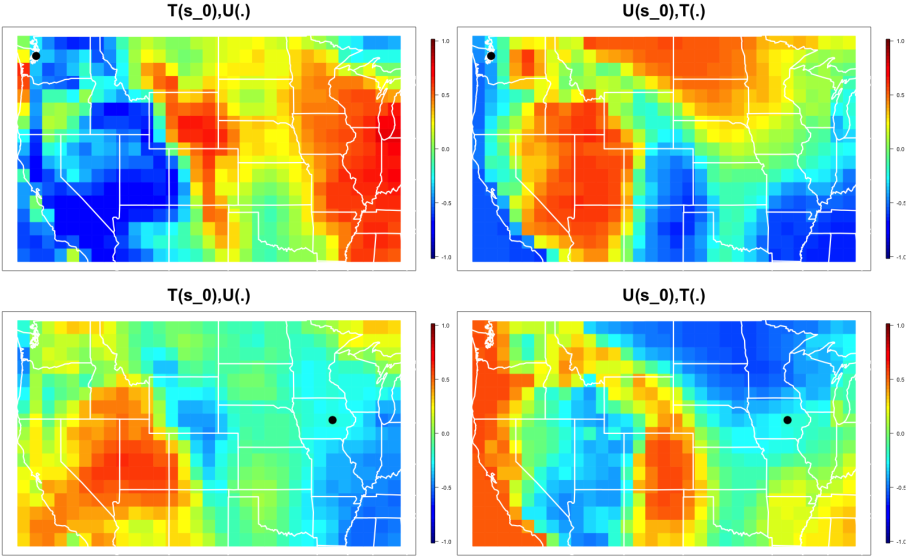
Precipitation/temperature spatial cross correlations



Precipitation/wind spatial cross correlations



Temperature/wind spatial cross correlations



Module 5: the future

A few recent ideas: space-time

Salvaña, Lenzi and Genton (2022, JCGS)

$$\mathbf{Z}(\mathbf{s}, t) = \begin{pmatrix} Z_1(\mathbf{s} - V_{11}t) \\ \vdots \\ Z_p(\mathbf{s} - V_{pp}t) \end{pmatrix}$$

If the $\{V_{ii}\}$ are random vectors then

$$\mathbf{C}(\mathbf{h}; t_1, t_2) = \mathbb{E}[\{C_{ij}(\mathbf{h} - V_{ii}t_1 + V_{jj}t_2\}_{i,j=1}^{p}]$$

A few recent ideas: lots of processes, stationary

Dey, Datta and Banerjee (2022, Biometrika)

- Stitch together univariate stationary processes such that marginal covariance functions are preserved
- Can control conditional dependencies between variables
- Cross-covariances are approximately controlled

A few recent ideas: lots of processes, nonstationary

Krock et al. (2023, JCGS)

$$\mathbf{Z}(\mathbf{s}) = \mathbf{c}_1\phi_1(\mathbf{s}) + \cdots \mathbf{c}_L\phi_L(\mathbf{s})$$

where

- $\phi_1, \ldots, \phi_L$ are fixed basis functions
- $\mathbf{c}_1, \ldots, \mathbf{c}_L$ are Gaussian graphical vectors that are independent

Estimation follows a fused graphical lasso-like idea, the "basis graphical lasso"

What does the future hold?

There are a lot of open research questions in multivariate modeling:

- Variable dependent non-Gaussianity
- Flexible space-time models that work for large datasets
- High-dimensional processes (LENS has hundreds!)
- Nonstationarity of between-variable relationships
- Coherence for nonstationary fields
- Very little theory (e.g., when should co-kriging really improve on kriging?)
- Exploratory data analysis techniques
- *Lots* of low hanging fruit, univariate methods can often be directly imported/used in the multivariate setting

Genton, M.G. and Kleiber, W. (2015). "Cross-covariance functions for multivariate geostatistics." *Statistical Science*, **30**, 147–163.