Chapter 1
Spatio-Temporal Kriging


Benedikt Gräler
http://ifgi.de/graeler
Institute for Geoinformatics
University of Muenster
Spatial Data

From a purely statistical perspective, spatial data is multivariate data with special covariates: the coordinates. Tobler’s first law of Geography states [3]:

*Everything is related to everything else, but near things are more related than distant things.*
Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:
  
  Degrees: 6.88  
  Great Circle: 385 km  
  Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):

  Degrees: 7.32  
  Great Circle: 814 km  
  Rate: 111 km/degree
Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space. Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:
  
  **Degrees:** 6.88
  
  **Great Circle:** 385 km
  
  **Rate:** 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):
  
  **Degrees:** 7.32
  
  **Great Circle:** 814 km
  
  **Rate:** 111 km/degree
Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space. Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:
  
  Degrees: 6.88
  Great Circle: 385 km
  Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):
  
  Degrees: 7.32
  Great Circle: 814 km
  Rate: 111 km/degree
Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space. Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:
  - Degrees: 6.88
  - Great Circle: 385 km
  - Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):
  - Degrees: 7.32
  - Great Circle: 814 km
  - Rate: 111 km/degree
Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space. Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:
  
  Degrees: 6.88  
  Great Circle: 385 km  
  Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):

  Degrees: 7.32  
  Great Circle: 814 km  
  Rate: 111 km/degree
Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:
  
  Degrees: 6.88 
  Great Circle: 385 km 
  Rate: 56 km/degree 

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):

  Degrees: 7.32 
  Great Circle: 814 km 
  Rate: 111 km/degree
Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space. Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:
  - Degrees: 6.88
  - Great Circle: 385 km
  - Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):
  - Degrees: 7.32
  - Great Circle: 814 km
  - Rate: 111 km/degree
Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space. Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:
  
  Degrees: 6.88  
  Great Circle: 385 km  
  Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):
  
  Degrees: 7.32  
  Great Circle: 814 km  
  Rate: 111 km/degree
To distinguish different projections, a well prepared data set comes with its coordinate reference system (CRS) as metadata.

These are often encoded as

- EPSG-codes (by the European Petroleum Survey Group)
- proj4string

They define how the reference surface (sphere, ellipsoid) is fixed to the real world (called the datum) and how the projection (surface in 3D to 2D plane) is made.
Projection

Spatio-Temporal Kriging

Benedikt Gräler

Introduction

Spatio-Temporal Kriging

empirical
metric
separable
product-sum
sum-metric

spatio-temporal
block kriging

Local
Spatio-Temporal Kriging

References
Fields

Fields are understood as continuously spreading over space and/or time (e.g. temperature recordings) and typically observed at a set of distinct locations for a series of time steps. Fields are typically illustrated as interpolated maps and modelled as a realisation of a spatial/spatio-temporal random field.
Stationarity and Isotropy

**stationarity** The process ”looks” the same at each location (e.g. mean and variance do not change from east to west)

**isotropy** The dependence between locations is determined only by their separating distance neglecting the direction (e.g. locations 2 km apart along the north-south axis are as correlated as stations 2 km apart along the east-west axis)

Some tricks exist to weaken these assumptions (e.g. rotating and rescaling coordinates).
The dependence across space of a random field $Z$ is assessed using a variogram $\gamma$:

$$\gamma(h) = \frac{1}{2} \mathbb{E} \left( Z(s) - Z(s + h) \right)^2$$

the empirical estimator looks like

$$\hat{\gamma}(h) = \frac{1}{2|N_h|} \sum_{(i,j) \in N_h} \left( Z(s_i) - Z(s_j) \right)^2$$

while $N_h = \{(i,j) : h - \epsilon \leq \|s_i - s_j\| \leq h + \epsilon\}$
The *sample variogram* is obtained through

```r
vgmMeuse <- variogram(zinc~1, meuse)
```

Variograms II

The *sample variogram* is obtained through

```r
vgmMeuse <- variogram(zinc~1, meuse)
```
And a theoretical *variogram model* can be fitted

```r
> head(vgm())
     short   long
1    Nug     Nug (nugget)
2    Exp     Exp (exponential)
3    Sph     Sph (spherical)
4    Gau     Gau (gaussian)
5  Exclass  (Exponential class)
6      Mat     Mat (Matern)

> vgmModelMeuse <- fit.variogram(vgmMeuse,
                                  vgm(0.6, "Sph", 1000, 0.1))

vgmModelMeuse
     model  psill   range
1      Nug 24813.21 0.0000
2      Sph 134753.99 831.2953
```
Variograms IV

![Variogram Chart]

- **distance**
- **semivariance**

- Distance values: 500, 1000, 1500
- Semivariance values: 50000, 100000, 150000

---

**References**

- **Variograms IV**
- **Empirical metric**
- **Separable product-sum sum-metric**
- **Spatio-temporal block kriging**
Certain variogram models can be used to parametrize a covariance matrix for a Gaussian random field over a finite set of locations $s_1, \ldots, s_n$:

$$Z \sim \text{Gau}(\mu, \Sigma)$$

while $\Sigma = (\sigma^2_{ij})_{ij}$ and $\sigma^2_{ij} = \sigma^2 - \gamma(||s_i - s_j||)$, $1 \leq i, j \leq n$ with $\sigma^2 = \text{Var}(Z(s))$, $\mu = (\mu_1, \ldots, \mu_n)$.

Predictions can be made using matrix inversion and matrix multiplications.
Kriging II

```r
krige(zinc~1, meuse, meuse.grid, model=vgmModelMeuse)
```

obs. zinc concentrations

kriged zinc concentrations
The model quantifies how uncertain it is about the estimates through the kriging variance:
Overview of kriging types

- **simple kriging** the mean value is known
- **ordinary kriging** prediction based on coordinates
- **universal kriging** prediction based on coordinates and additional regressors (distance to the river)
- **co-kriging** the cross-variogram between two variables is as well exploit (zinc and lead)
Overview of kriging types

**Simple kriging** the mean value is known

**Ordinary kriging** prediction based on coordinates

**Universal kriging** prediction based on coordinates and additional regressors (distance to the river)

**Co-kriging** the cross-variogram between two variables is as well exploit (zinc and lead)
Overview of kriging types

**simple kriging**  the mean value is known

**ordinary kriging**  prediction based on coordinates

**universal kriging**  prediction based on coordinates and additional regressors (distance to the river)

**co-kriging**  the cross-variogram between two variables is as well exploit (zinc and lead)
Overview of kriging types

**simple kriging**  the mean value is known

**ordinary kriging**  prediction based on coordinates

**universal kriging**  prediction based on coordinates and
  additional regressors (distance to the river)

**co-kriging**  the cross-viariogram between two variables is as
  well exploit (zinc and lead)
Spatio-Temporal Data

$S \times T$ works as a data structure, but modelling needs to consider special properties of the product of space and time.

**direction** Today’s values influence tomorrow, but will not take effect on yesterday’s values.

**anisotropy** What is the equivalent in terms of dependence of 1 m separation in seconds or minutes?

The easiest way to think of spatio-temporal data is as time slices - but this neglects the temporal dependence.

After modelling temporal trend or periodicities, the residuals might be modelled as a spatio-temporal random field.
Almost Spatio-Temporal approaches

slice wise  the easiest adoption is to do interpolation per slice fitting a variogram model for each time slice

pooled  the variogram is fitted based on all spatio-temporal data and is used to predict each time slice separately with the same model

evolving  models mix the both extremes such that the variogram model adopts to the daily situation (e.g. in terms of overall variability, the sill) but range and the nugget/sill ratio depend on larger data samples.
Almost Spatio-Temporal approaches

Slice wise: the easiest adoption is to do interpolation per slice fitting a variogram model for each time slice.

Pooled: the variogram is fitted based on all spatio-temporal data and is used to predict each time slice separately with the same model.

Evolving models mix the both extremes such that the variogram model adopts to the daily situation (e.g. in terms of overall variability, the sill) but range and the nugget/sill ratio depend on larger data samples.
Almost Spatio-Temporal approaches

slice wise  the easiest adoption is to do interpolation per slice fitting a variogram model for each time slice

pooled  the variogram is fitted based on all spatio-temporal data and is used to predict each time slice separately with the same model

evolving models mix the both extremes such that the variogram model adopts to the daily situation (e.g. in terms of overall variability, the sill) but range and the nugget/sill ratio depend on larger data samples.
The spatio-temporal variogram

Extending the variogram to a two-place function for spatio-temporal random fields $Z(s, t)$:

$$\gamma(h, u) = E(Z(s, t) - Z(s + h, t + u))^2$$

at any location $(s, t)$. And empirical version

$$\hat{\gamma}(h, u) = \frac{1}{2|N_{h,u}|} \sum_{(i,j) \in N_{h,u}} (Z(s_i, t_i) - Z(s_j, t_j))^2$$

while $N_{h,u} = \left\{(i, j) \left| \begin{array}{l}
h - \epsilon_s \leq ||s_i - s_j|| \leq h + \epsilon_s \\
u - \epsilon_t \leq t_i - t_j \leq u + \epsilon_t \end{array} \right. \right\}$
Scenario

We have a set of spatially spread time series of daily measurements and are asked to produce a map of means for the provided time frame. The data is provided by the EEA, the presentation is composed along the lines of [2].
empirical spatio-temporal variogram surface

The idea is the same as in the spatial case: binning of locations according to their separating distance. In the spatio-temporal case, distances are pairs of spatial and temporal distance yielding a variogram surface, not a single line.

\[ \text{empVgm} \leftarrow \text{variogramST}(\text{PM10}\sim1, \text{ger_june}, \text{tlags}=0:4, \text{cutoff}=500e3) \]

# rescaling of distances
\[
\text{empVgm}$\text{dist} \leftarrow \text{empVgm}$\text{dist}/1000
\]

\[
\text{empVgm}$\text{spacelag} \leftarrow \text{empVgm}$\text{spacelag}/1000
\]

# wireframe:
\[ \text{plot(empVgm, wireframe=T, scales=list(arrows=F),} \]
\[ \text{col.regions=bpy.colors(),zlab=list(rot=90),zlim=c(0,20))} \]

# levelplot:
\[ \text{plot(empVgm)} \]
Spatio-Temporal Kriging

Introduction

Spatio-Temporal Kriging

empirical metric separable product-sum sum-metric

spatio-temporal block kriging

Local Spatio-Temporal Kriging

References

empirical spatio-temporal variogram surface - wireframe
Spatio-Temporal Kriging

Benedikt Gräler

Introduction

Spatio-Temporal Kriging

empirical metric
separable product-sum
sum-metric

spatio-temporal block kriging

Local Spatio-Temporal Kriging

References

empirical spatio-temporal variogram surface - levelplot
The **metric kriging** follows the natural idea of extending the 2-dimensional geographic space into a 3-dimensional spatio-temporal one. In order to achieve an isotropic space, the temporal domain has to be rescaled to match the spatial one (spatio-temporal anisotropy correction $\kappa$).

All spatial, temporal and spatio-temporal distances are treated equally resulting in a joint covariance model $C_j$:

$$C_m(h, u) = C_j\left(\sqrt{h^2 + (\kappa \cdot u)^2}\right)$$

The variogram evaluates to

$$\gamma_m(h, u) = \gamma_j\left(\sqrt{h^2 + (\kappa \cdot u)^2}\right)$$

where $\gamma_j$ is any known variogram including some nugget effect.
metric kriging in R

```r
metricModel <- vgmST("metric",
    joint=vgm(0.8,"Exp", 150, 0.2),
    stAni=100)
metricFit <- fit.StVariogram(empVgm,metricModel,
    lower=c(0,10,0,10))

attr(metricFit,"optim.output")$value
>  1.080641
plot(empVgm, metricFit)

predMetric <- krigesT(PM10~1, ger_june,
    STF(ger_gridded,tgrd),
    metricFit)
```
Spatio-Temporal Kriging

Introduction

Spatio-Temporal Kriging

empirical

metric

separable

product-sum

sum-metric

spatio-temporal

block kriging

Local

Spatio-Temporal Kriging

References

metric spatio-temporal variogram surface

Distance vs. Time Lag (days)
Spatio-Temporal Kriging

Introduction

Spatio-Temporal Kriging

empirical
metric
separable
product-sum
sum-metric

spatio-temporal
block kriging

Local
Spatio-Temporal Kriging

References
separable covariance models

In space and under the assumptions of isotropy and stationarity, the covariance is a function $C(h)$ of the separating distance $h$ between two locations. A spatio-temporal covariance function is thought of as a function of a spatial and a temporal distance $C(h, t)$.

A *separable covariance function* is assumed to fulfill $C_{sep}(h, u) = C_s(h)C_t(u)$. This is in general a rather strong simplification. Its variogram is given by

$$\gamma_{sep}(h, u) = \text{nug} \cdot 1_{h>0, u>0} + \text{sill} \cdot (\gamma_s(h) + \gamma_t(u) - \gamma_s(h)\gamma_t(u))$$

where $\gamma_s$ and $\gamma_t$ are spatial and temporal variograms without nugget effect and a sill of 1. The overall nugget and sill parameters are denoted by "nug" and "sill" respectively.
separable covariance model in R

```
sepModel <- vgmST("separable",
    space=vgm(0.8, "Exp", 150, 0.2),
    time = vgm(0.7, "Exp", 6, 0.3),
    sill=18)
sepFit <- fit.StVariogram(empVgm, sepModel,
    lower=c(10, 0, 1, 0, 0))

attr(sepFit, "optim.output")$value
> 0.8001906
plot(empVgm, sepFit)

predSep <- krigeST(PM10~1, ger_june,
    STF(ger_gridded, tgrd),
    sepFit)
```
Spatio-Temporal Kriging

Introduction

Spatio-Temporal Kriging

empirical metric

separable

product-sum

sum-metric

spatio-temporal block kriging

Local Spatio-Temporal Kriging

References

variogram surface of the product-sum model
kriged map for day 15 - separable covariance model

day 15
The \textit{product sum covariance model} extends the simplifying assumption of the separable covariance model to [1]:

\[ C_{ps}(h, u) = k_1 C_s(h) + k_2 C_t(u) + k_3 C_s(h) C_t(u) \]

with \( k_1 > 0, k_2 \geq 0 \) and \( k_3 \geq 0 \) to fulfil the positive-definite condition. The corresponding variogram can be written as

\[ \gamma_{ps}(h, u) = \text{nug} \cdot 1_{h>0,u>0} + \gamma_s(h) + \gamma_t(u) - k\gamma_s(h)\gamma_t(u) \]

where \( \gamma_s \) and \( \gamma_t \) are spatial and temporal variograms without nugget effect and in general different sills. The parameter \( k \) needs to fulfil \( 0 < k \leq 1/(\max(sill_s, sill_t)) \) to let \( \gamma_{ps} \) be a valid model. The overall nugget is denoted by ”nug”.
product-sum covariance model in R

```r
empVgm$dist <- empVgm$dist/10
empVgm$spacelag <- empVgm$spacelag/10

psModel <- vgmST("productSum",
    space=vgm(11,"Exp", 2),
    time =vgm(5,"Sph", 6),
    sill=16, nugget=5)

psFit <- fit.StVariogram(empVgm,psModel)
attr(psFit,"optim.output")$value
> 0.7789366
plot(empVgm, psFit)

predPs <- krigeST(PM10~1, ger_june,
    STF(ger_gridded,tgrd),
    psFit)
```
variogram of the product-sum model
Spatio-Temporal Kriging

Introduction

Spatio-Temporal Kriging
- empirical metric
- separable
- product-sum
- sum-metric

Local Spatio-Temporal Kriging

References
sum-metric covariance model

The *sum-metric covariance model* is given by:

\[ C_{sm}(h, u) = C_s(h) + C_t(u) + C_j(\sqrt{h^2 + (\kappa \cdot u)^2}) \]

Originally, this model allows for spatial, temporal and joint nugget effects, a simplified version may allow only for a joint nugget. The non-simplified variogram is given by

\[ \gamma_{sm}(h, u) = \gamma_s(h) + \gamma_t(u) + \gamma_j(\sqrt{h^2 + (\kappa \cdot u)^2}) \]

where \( \gamma_s, \gamma_t \) and \( \gamma_j \) are spatial, temporal and joint variograms with a separate nugget-effect.
sum-metric covariance model in R

```r
empVgm$dist <- empVgm$dist/10
empVgm$spacelag <- empVgm$spacelag/10

sumMetricModel <- vgmST("sumMetric",
    space=vgm(5,"Exp",10,2),
    time =vgm(5,"Exp", 6,2),
    joint=vgm(5,"Exp",10,2),
    stAni=10)

sumMetricFit <- fit.StVariogram(empVgm,sumMetricModel,
    lower=c(0,1,0,0,1,0,0))

attr(sumMetricFit,"optim.output")$value
> 0.6754955

plot(empVgm, sumMetricFit)

predSumMetric <- krigeST(PM10~1, ger_june,
    STF(ger_gridded,tgrd),
    sumMetricFit)
```
variogram of the sum-metric model
Spatio-Temporal Kriging

Introduction

Spatio-Temporal Kriging
- empirical metric
- separable product-sum sum-metric

Local Spatio-Temporal Kriging

References

kriged map for day 15 - product-sum covariance model
variogram of all spatio-temporal models
Spatio-Temporal Kriging

Benedikt Gräler

Introduction

Spatio-Temporal empirical separable product-sum sum-metric block kriging

Local

References

kriged map for 10 days

2009−06−10 2009−06−11 2009−06−12 2009−06−13 2009−06−14

2009−06−15 2009−06−16 2009−06−17 2009−06−18 2009−06−19

2009−06−10 2009−06−11 2009−06−12 2009−06−13 2009−06−14

2009−06−15 2009−06−16 2009−06−17 2009−06−18 2009−06−19

2009−06−10 2009−06−11 2009−06−12 2009−06−13 2009−06−14

2009−06−15 2009−06−16 2009−06−17 2009−06−18 2009−06−19
In the above scenario and with the presented methods, it is hard to get an uncertainty estimate of the temporally averaged value. Block kriging, with blocks over time, is one way to get such estimates. However, one has to decide on a model beforehand. Here, we will use the metric model again.

Block kriging does not provide estimates for single locations but for areas or volumes. It has the property of providing the correct kriging variance for the block estimate that is typically lower due to the larger area.
monthly mean concentration - block kriged
block kriging variance
Spatio-Temporal Kriging

Benedikt Gräler

Introduction

Spatio-Temporal Kriging

empirical metric
separable product-sum sum-metric

spatio-temporal block kriging

Local Spatio-Temporal Kriging

References

kriging variance day 15
block kriging in R - metric workaround

tmp_pred <- data.frame(cbind(ger_gridded@coords, 15*tmpScale))
colnames(tmp_pred) <- c("x","y","t")
coordinates(tmp_pred) <- ~x+y+t

blockKriging <- krigge(PM10~1,
                        air3d[as.vector(!is.na(air3d@data)),],
                        newdata=tmp_pred, model=model3d,
                        block=c(1,1,15*tmpScale))

ger_grid_time@sp@data <- blockKriging@data
local spatio-temporal kriging

Purely spatial kriging allows to select the n-nearest neighbours and use only these for prediction.

What does nearest mean in a spatio-temporal context?

The idea is to select the most valuable locations, i.e. the strongest correlated ones.

Simply set the argument $n_{\text{max}}$ and a local neighbourhood of the most correlated values is selected from a larger ”metric” neighbourhood.
References 1

