Summary STK 4150/9150

Odd Kolbjørnsen

May 22 2017

Scope

You are expected to know and be able to use basic concepts introduced in the book.

You knowledge is expected to be larger than than the content of the book, i.e.

You are supposed to be able to derive equations not directly written in the Book, but with similar complexity.

Course @ a slide

Prelim/Jan

- Statistics preliminaries (Chapter 2)
- Temporal processes (Chapter 3)
  - Deterministic models
  - Stochastic models
  - Spectral representation

January-Mach

- Spatial processes
  - Geostatistical processes, sec 4.1
  - Lattice processes, sec 4.2
  - Point processes, sec 4.3

Mach-May

- Spatio-temporal processes
  - Exploratory methods (Chapter 5)
  - Models (Chapter 6)
  - Hierarchical models (Chapters 7 and 8)

- Focus on
  - Modeling
  - Analysis in practice (using R)
  - Theoretical aspects

Temporal processes

Data collected over time

- Past, present, future, change
- Temporal aspect important?

Two (separate) approaches to modelling

- Statistical
  - Variability through randomness
  - Learning dynamic structure from data
- Dynamic system theory
  - Typically deterministic

Characterization of temporal processes

Denoted $\mathbf{Y}(\cdot)$ or $\{\mathbf{Y}(r) : r \in D_t\}$

- r time index
- $\mathbf{Y}(r)$ possibly multivariate, deterministic or stochastic
- $D_t \in \mathcal{R}^1$

D_t specify type of process

- Continuous time process: $D_t \in (-\infty, \infty)$ or $[0, \infty)$: $\mathbf{Y}(t)$
- Discrete time process: $D_t \in \mathcal{N}$ or $\mathcal{N}^+$: $\mathbf{Y}_t$
- Point process: Random times
  Example: Time of tornado. Mark: Severity of tornado

Continuous time:

- Measured discrete
- Modelled discrete?
- Dynamic system theory: Often modelled continuous (PDE's)

Spatial processes- Geostatistics

- Spatial correlation
  - Stationary
  - Isotropic/Anisotropic
  - Permutation test for independence (Morans I)
- Variogram
  - Link to covariance
  - Nugget effect
  - Sill
- Spatial prediction
  - Interpolation
  - Minimum Mean Squared Prediction Error (Kriging)
  - Prediction in a Gaussian model
- Kriging
  - Simple Kriging (known mean)
  - Bayesian Kriging (prior on mean)
  - Universal Kriging (unknown mean)
  - Ordenary Kriging (Special case of UK)

Spatial processes- Geostatistics

- General set up for "Kriging type problems"
  - Kriging case
  - General case
  - Change of support
- Spatial moving average models
  - Construction
  - Correlation function
- Computations in multigaussian settings
  - Conditional distribution
  - Deriving Mean and co-variance
  - Conditional simulation using Kriging equations

Hierarchical model

	Variable	Densities	Notation in book
Data model:	$\mathbf{Z}$	$p(\mathbf{Z}\|\mathbf{Y},\boldsymbol{\theta})$	$[\mathbf{Z}\|\mathbf{Y},\boldsymbol{\theta}]$
Process model:	$\mathbf{Y}$	$p(\mathbf{Y}\|\boldsymbol{\theta})$	$[\mathbf{Y}\|\boldsymbol{\theta}]$
Parameter:	$\boldsymbol{\theta}$		

Hierarchical (statistical) models

Hierarchical model

	Variable	Densities	Notation in book
Data model:	$\mathbf{Z}$	$p(\mathbf{Z}\mid\mathbf{Y},\boldsymbol{\theta})$	$[\mathbf{Z}\mid\mathbf{Y},\boldsymbol{\theta}]$
Process model:	$\mathbf{Y}$	$p(\mathbf{Y}\mid\boldsymbol{\theta})$	$[\mathbf{Y}\mid\boldsymbol{\theta}]$
Parameter:	$\boldsymbol{\theta}$		

Simultaneous model: $p(\mathbf{y},\mathbf{z}|\boldsymbol{\theta}) = p(\mathbf{z}|\mathbf{y},\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})$
Marginal model: $L(\boldsymbol{\theta}) = p(\mathbf{z}|\boldsymbol{\theta}) = \int_{\mathbf{y}} p(\mathbf{z},\mathbf{y}|\boldsymbol{\theta})d\mathbf{y}$

Inference: $\widehat{\boldsymbol{\theta}} = \text{argmax}_{\boldsymbol{\theta}} L(\boldsymbol{\theta})$

Hierarchical (statistical) models

Hierarchical model

	Variable	Densities	Notation in book
Data model:	$\mathbf{Z}$	$p(\mathbf{Z}\mid\mathbf{Y},\boldsymbol{\theta})$	$[\mathbf{Z}\mid\mathbf{Y},\boldsymbol{\theta}]$
Process model:	$\mathbf{Y}$	$p(\mathbf{Y}\mid\boldsymbol{\theta})$	$[\mathbf{Y}\mid\boldsymbol{\theta}]$
Parameter:	$\boldsymbol{\theta}$		

Simultaneous model: $p(\mathbf{y},\mathbf{z}\mid\boldsymbol{\theta}) = p(\mathbf{z}\mid\mathbf{y},\boldsymbol{\theta})p(\mathbf{y}\mid\boldsymbol{\theta})$
Marginal model: $L(\boldsymbol{\theta}) = p(\mathbf{z}\mid\boldsymbol{\theta}) = \int_{\mathbf{y}} p(\mathbf{z},\mathbf{y}\mid\boldsymbol{\theta})d\mathbf{y}$
Inference: $\widehat{\boldsymbol{\theta}} = \text{argmax}_{\boldsymbol{\theta}} L(\boldsymbol{\theta})$

Bayesian approach: Include model on $\boldsymbol{\theta}$

	Variable	Densities	Notation in book
Data model:	$\mathbf{Z}$	$p(\mathbf{Z}\mid\mathbf{Y},\boldsymbol{\theta})$	$[\mathbf{Z}\mid\mathbf{Y},\boldsymbol{\theta}]$
Process model:	$\mathbf{Y}$	$p(\mathbf{Y}\mid\boldsymbol{\theta})$	$[\mathbf{Y}\mid\boldsymbol{\theta}]$
Parameter model:	$\boldsymbol{\theta}$	$p(\boldsymbol{\theta})$	$[\boldsymbol{\theta}]$

Simultaneous model: $p(\mathbf{y},\mathbf{z},\boldsymbol{\theta})$
Marginal model: $p(\mathbf{z}) = \int_{\boldsymbol{\theta}}\int_{\mathbf{y}} p(\mathbf{z},\mathbf{y}\mid\boldsymbol{\theta})d\mathbf{y}d\boldsymbol{\theta}$
Inference: $\widehat{\boldsymbol{\theta}} = \int_{\boldsymbol{\theta}} \boldsymbol{\theta} p(\boldsymbol{\theta}\mid\mathbf{z})d\boldsymbol{\theta}$

Hyper parameters parameters and non-liearities

- Linear parameteres approach (for mean parameters)
  - Ordinary kriging
  - Universal Kriging
  - Bayesian Kriging
- Hyper parameters: Plug-in estimate/Empirical Bayes
- Hyper parameters: Bayesian approach
  - MCMC
  - Laplace approximation
  - INLA
- Non Gaussian observations
  - Monte Carlo
  - Laplace approximation
  - INLA

Lattice models

- AR(1) as lattice process
- MRF – Markov random field
  - Undirected graph (MRF)
  - Neighborhood
  - Clique
- Negpotential function
- Gibbs distribution
- Besag's lemma (conditional vs joint distribution)
- Hammersley- Clifford theorem (clique vs Neg potential function)
- Lattice models
  - Gaussian CAR
  - Precision matrices
  - Latent Gaussian process
  - Auto logistic model (Ising model)
  - Auto Poisson model
- Gaussian CAR
  - Relation to precision matrix
  - When is it well defined?

Point processes

- Point process (random locations, random count, random marks)
- Density of a point process (given the count)
- Poisson process
  - Intensity
  - Homogeneous/ In homogeneous
- Other processes
  - Random intensity (Log Gaussian Cox Process)
  - Clustering (Parent-child,e.g. Neyman-Scott)
  - Repulsion (Markov point process, Strauss - Hard core process)
- Inference
  - Kernel estimate of intensity
  - L-function (K-function)
  - Edge effects

Exploratory methods

- Visualization
- Presentation of results
- Empirical Orthogonal functions
  $\mathbf{C}_Z = \mathbf{\Psi}\mathbf{\Lambda}^2\mathbf{\Psi}^T, \quad \mathbf{\Lambda}^2 = \text{diag}\{\lambda_i^2\} \quad \mathbf{\Psi} = [\boldsymbol{\psi}_i, ..., \boldsymbol{\psi}_m]$
  - Efficient computation
  - Estimation of space functions (time coefficients)
  - Estimation of time functions (space coefficients)
- Space-Time Index (STI) used for permutation test in time space dependency.

Space time process models

Spatio-temporal covariance functions

- Stationarity: spatio-temporal, Spatial, temporal
- Spatio-temporal - Kriging
- Seperable, additive , multiplicative correlation functions

Stochastic differential/difference equations

- Integro-difference equation models
- Using (partial) differential equations (e.g. Matern correlation)
- Diffusion-injection models (interpretation of terms)
- Blurring space: Hard in space-time Simple in Fourier domain
- Discretization

Time series of spatial processes

- AR(q) process in time
- Stationary transitions
- Stationary distributions
- Discretization in space as well gives Vector-AR

Hierarchical Dynamical Spatio-Temporal Models framework

- Data in Process models
- Observation types
  - number of observations
  - Incidence matrix
  - Change of support
  - Multiple data sources
- Linear observations
- Kalman filter
- Kalman smoother
- nonlinear/non Gaussian
- Conditioning to nonlinear functions of a random field

Bayesian approach: Also include model for parameters

Methodology for inference in Hierarchical Dynamical Spatio-Temporal Models

- General Problem
- Sequential vs non sequential
- Kalman filter
- EM-algorithm
- MCMC
- Sequential Monte Carlo, particle filter
- INLA

Data Examples

- Exercises
- Boreal - data
- Potts model
- Scottish lip cancer
- Pacific SST
- Eurasian Dove