Chapter 6 - spatio-temporal models

Odd Kolbjørnsen and Geir Storvik

27. March 2017

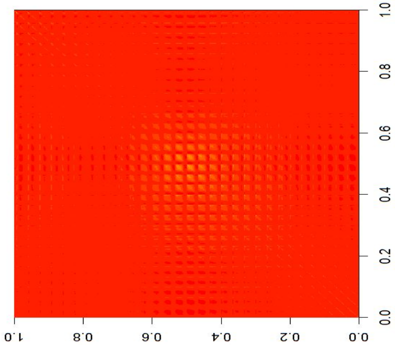
Last time Exploratory methods

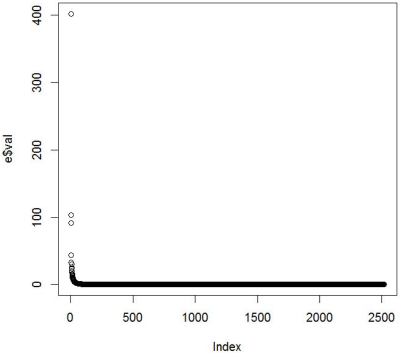
- Visualization
- Presentation of results
- Empirical Orthogonal functions
  $\mathbf{C}_Z = \mathbf{\Psi}\mathbf{\Lambda}^2\mathbf{\Psi}^T, \quad \mathbf{\Lambda}^2 = \text{diag}\{\lambda_i^2\} \quad \mathbf{\Psi} = [\psi_i, ..., \psi_m]$
  - Why ...
  - Efficient computation
  - Estimation of space functions (time coefficients)
  - Estimation of time functions (space coefficients)
- Space-Time Index (STI) used for permutation test in time space dependency.
- Conditional simulation using Kriging equations
- Conditioning to nonlinear functions of a random field
- Application of marked point processes

Empirical Orthogonal functions Why

Empirical Orthogonal functions

$$\mathbf{C}_Z = \mathbf{\Psi}\mathbf{\Lambda}^2\mathbf{\Psi}^T, \quad \mathbf{\Lambda}^2 = \text{diag}\{\lambda_i^2\} \quad \mathbf{\Psi} = [\psi_i, ..., \psi_m]$$

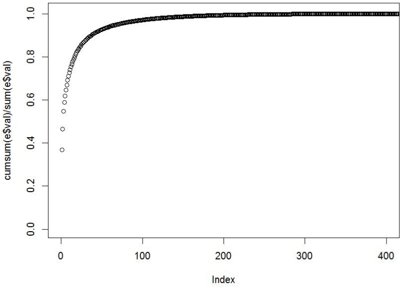


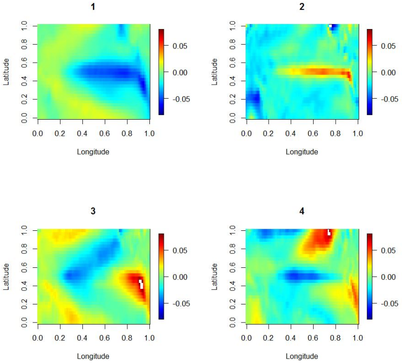


Empirical Orthogonal functions Why

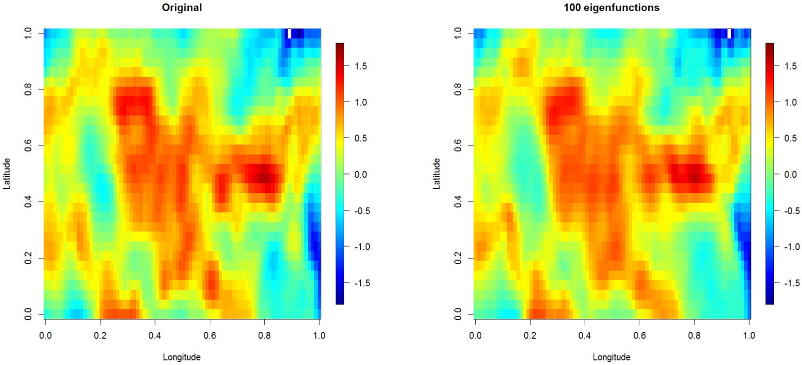
Empirical Orthogonal functions

$$\mathbf{C}_Z = \mathbf{\Psi}\mathbf{\Lambda}^2\mathbf{\Psi}^T, \quad \mathbf{\Lambda}^2 = \text{diag}\{\lambda_i^2\} \quad \mathbf{\Psi} = [\psi_i, ..., \psi_m]$$

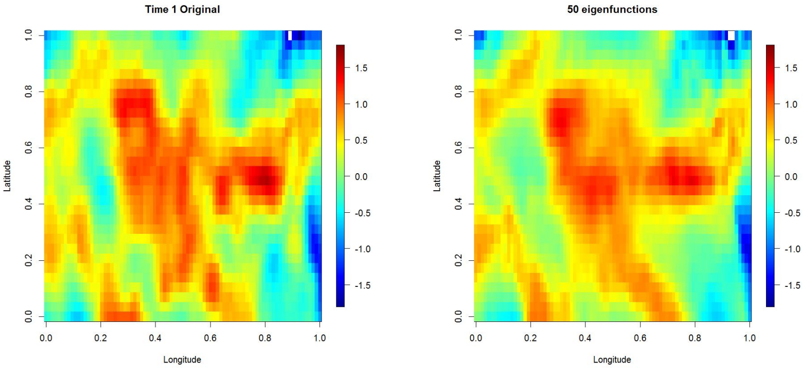




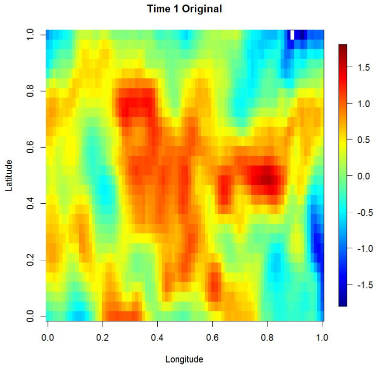
Empirical Orthogonal functions Why

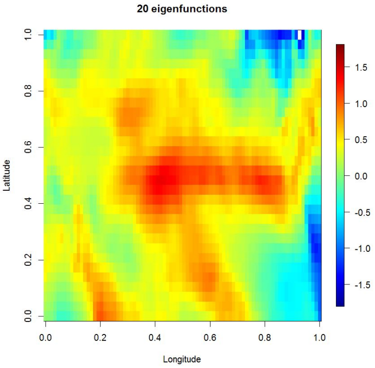


Empirical Orthogonal functions Why

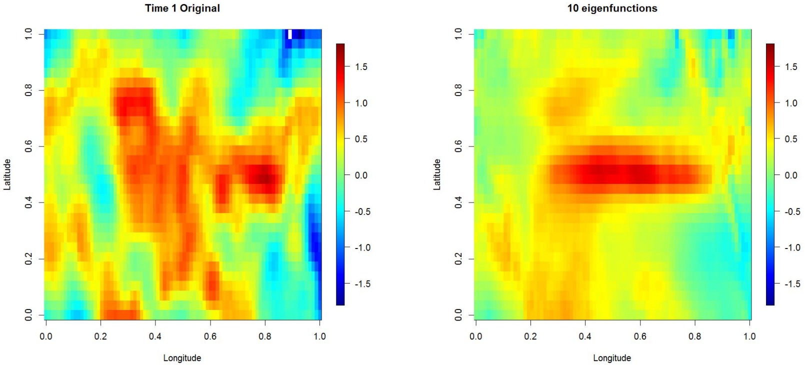


Empirical Orthogonal functions Why

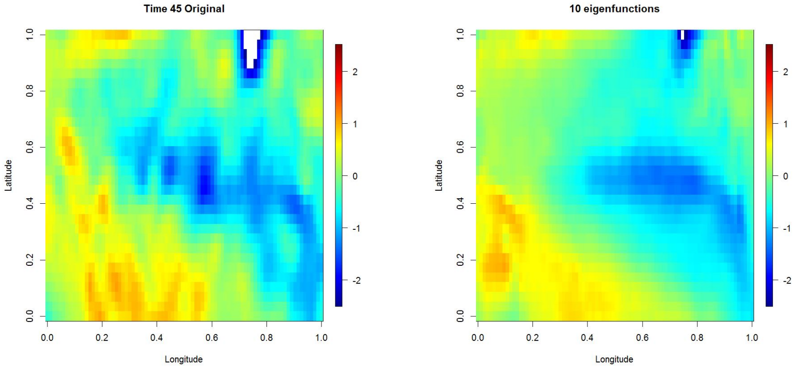




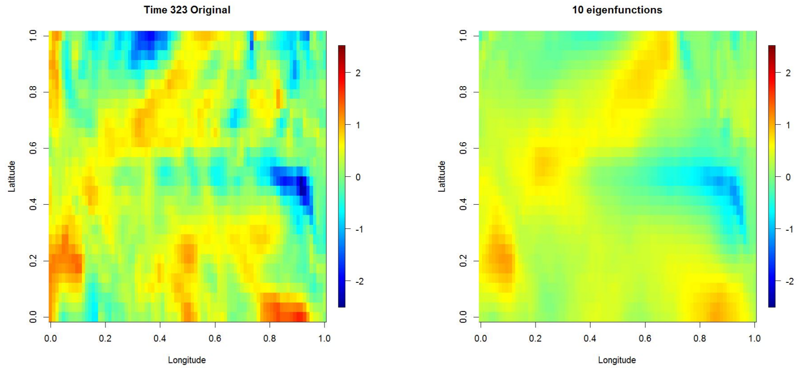
Empirical Orthogonal functions Why



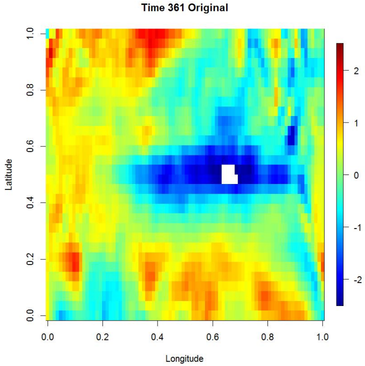
Empirical Orthogonal functions Why

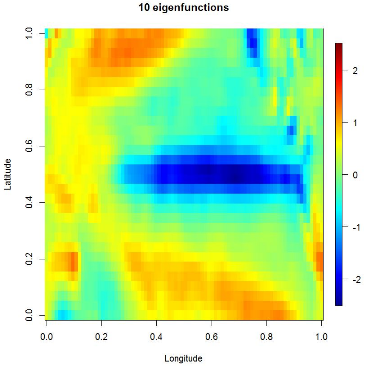


Empirical Orthogonal functions Why

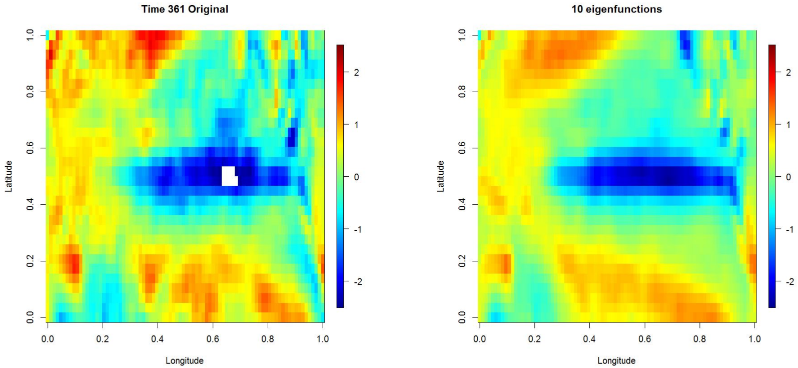


Empirical Orthogonal functions Why

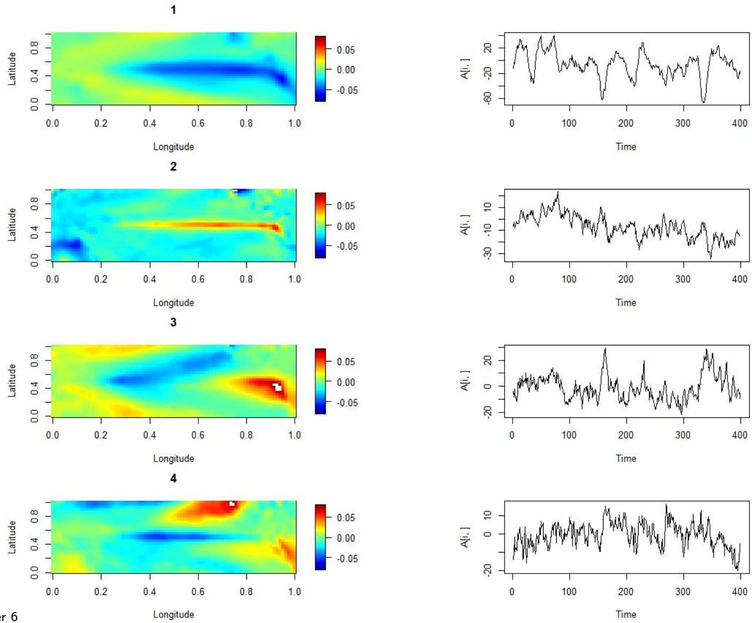




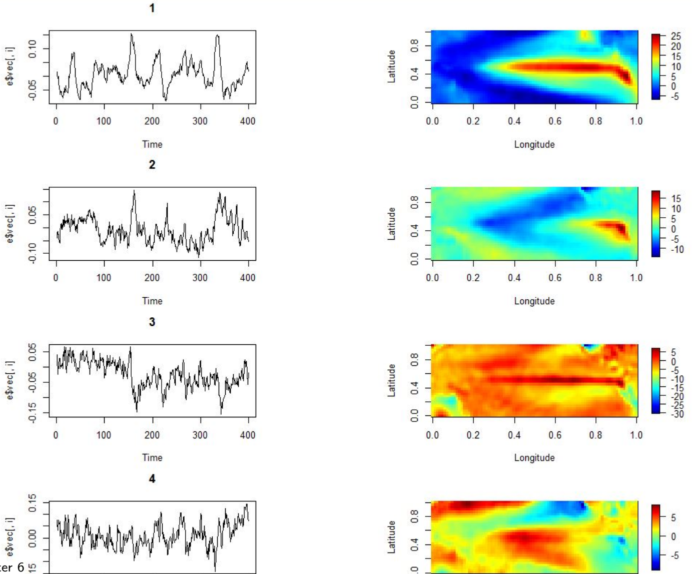
Empirical Orthogonal functions Why



EOF: Coefficients of spatial eigenvectors create timeseries



EOF: Coefficients of temporal eigenvectors create spatial variations



Spatio-temporal data

In mathematical terms there is no difference between $\mathcal{R}^{d+1}$ "Spatial" random fields and $\mathcal{R}^d \times [0, \ \infty)$ "Time-space". The modeling will still be different because the features have different characteristics in time and space.

- Similar: Nearby values tend to be more alike than those far apart. (Near both in time and space)
- Difference: Order in time, no general order in space
- Time series: Two options
  - Describing covariance structure
  - Specifying dynamical model
- Same possibilities with spatio-temporal models
- Book: Emphasis on dynamical models
  - Use *conditional* distributions

$$[Y(\mathbf{s};t)|\{Y(\mathbf{x};r):\mathbf{x}\in D_s, r<t\}\cup\{Y(\mathbf{x},t):\mathbf{x}\neq s\}$$

Time ahead

Today modeling strategies:

- Spatio-temporal covariance functions (sec 6.1,6.2)
- Stochastic differential/difference equations (sec 6.3)
- Time series of spatial processes (sec 6.4)

3. April Ch 7 & 8

- Hierarchical spatio-temporal processes
- Data inference

24. April Ch 8 & 9

- Hierarchical spatio-temporal processes
- Data inference
- Example

8.May Ch 9. Project assignment available online.

- Hierarchical spatio-temporal processes
- Example

Time ahead

8.May Project assignment available online.
15.May no lecture
22.May Project assignment due.
22.May Course summary

30 May at 09:00 (4 hours). Written Exam

Spatio-temporal covariance functions

General model

$$\begin{aligned}
Y(\mathbf{s};t) &\equiv \mu(\mathbf{s};t) + \beta(\mathbf{s}) + \gamma(t) + \kappa(\mathbf{s};t) + \delta(\mathbf{s};t), \quad \mathbf{s}\in D_s, t\in D_t\\
\mu(\mathbf{s};t) &\equiv \text{fixed (covariate) term}\\
\text{cov}[\beta(\mathbf{s}),\beta(\mathbf{x})] &\equiv C_\beta(\mathbf{s},\mathbf{x})\\
\text{cov}[\gamma(t),\gamma(r)] &\equiv C_\gamma(t,r)\\
\text{cov}[\kappa(\mathbf{s};t),\kappa(\mathbf{x};r)] &\equiv C_\kappa(\mathbf{s},\mathbf{x};t,r)\\
\text{cov}[\delta(\mathbf{s};t),\delta(\mathbf{x};r)] &\equiv \sigma_\delta^2 I(\mathbf{s}=\mathbf{x}, t=r)
\end{aligned}$$

Many important special cases

Main focus now: Specification of $C_\kappa(\mathbf{s},\mathbf{x};t,r)$.

Spatio-temporal covariance functions

- Positive definiteness:
  - $\text{var}[\sum_i a_i Y(\mathbf{s}_i, t_i)] \geq 0$ imply
  - requirement $\sum_i \sum_j a_i a_j C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j)) \geq 0$

  A function $C()$ satisfying this for any collection $\{Y(\mathbf{s}_i, t_i)\}$ and all $\{a_i\}$ is positive definite.
- Stationarity:

$$\begin{aligned} \text{cov}[Y(\mathbf{s};t), Y(\mathbf{x};r)] =& C(\mathbf{s}-\mathbf{x}, t, r), && \text{spatial stationarity} \\ \text{cov}[Y(\mathbf{s};t), Y(\mathbf{x};r)] =& C(\mathbf{s}, \mathbf{x}, t-r), && \text{temporal stationarity} \\ \text{cov}[Y(\mathbf{s};t), Y(\mathbf{x};r)] =& C(\mathbf{s}-\mathbf{x}, t-r), && \text{spatio-temporal stationarity} \end{aligned}$$

- Constructing positive definite functions are difficult (even under stationarity)

Separability

$$\text{cov}[Y(\mathbf{s};t),Y(\mathbf{x};r)] = C(\mathbf{s},t,\mathbf{x},r) = C^{(s)}(\mathbf{s},\mathbf{x})\cdot C^{(t)}(t,r)$$

$C^{(s)}(\mathbf{s},\mathbf{x})$ is legal spatial covariance function and $C^{(t)}(t,r)$ is a legal temporal covariance function imply product also legal
(exercise)
Assuming stationarity

$$C(\mathbf{s},t,\mathbf{x},r) = C(\mathbf{s}-\mathbf{x},t-r) = C^{(s)}(\mathbf{s}-\mathbf{x})\cdot C^{(t)}(t-r)$$

gives

$$\rho(\mathbf{h};\tau) = C(\mathbf{h};\tau)/C(\mathbf{0};0) = \rho(\mathbf{h};0)\rho(\mathbf{0};\tau)$$

so $\rho(\cdot;\tau)$ is proportional to $\rho(\cdot;0)$ for all τ.
Much used because of computational simplifications.

Separability and computation

Assume $C(\mathbf{s}-\mathbf{t},t-r)=C^{(s)}(\mathbf{s}-\mathbf{x})\cdot C^{(t)}(t-r), C^{(t)}(0)=1$
Assume observed $\{Y(s_i,t_j), i=1,...,m, j=1,...,n\}$.
$\mathbf{Y}_j=(Y(s_1,t_j),...,Y(s_m,t_j))'$, $\mathbf{Y}'=(\mathbf{Y}_1',...,\mathbf{Y}_n')$: We have

$$\begin{aligned}\text{Cov}[\mathbf{Y}] =&\boldsymbol{\Sigma}_0\otimes\boldsymbol{\Gamma}_0\\ \boldsymbol{\Sigma}_0 =&\text{Var}[\mathbf{Y}_j]\\ \boldsymbol{\Gamma}_0 =&\text{Var}[(Y(s_i,t_1),...,Y(s_i,t_n))']\end{aligned}$$

Gives

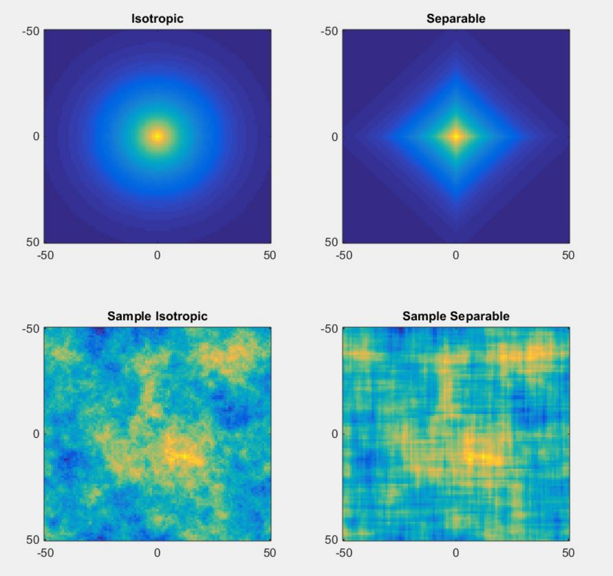
$$\mathbf{Q}=\boldsymbol{\Sigma}_0^{-1}\otimes\boldsymbol{\Gamma}_0^{-1}$$

Only need to invert

- $\boldsymbol{\Sigma}_0$, an $m\times m$ matrix
- $\boldsymbol{\Gamma}_0$, an $n\times n$ matrix

Direct: Need to work with an $nm\times nm$ matrix

Isotropic vs separable Exponential correlation , $\mathcal{R}^2$



Additive structures

Additive independent parts:

$$\begin{aligned} Y(\mathbf{s};t) =& \beta(\mathbf{s}) + \gamma(t) \\ \text{cov}[Y(\mathbf{s};t), Y(\mathbf{x};r)] =& C_\beta(\mathbf{s},\mathbf{x}) + C_\gamma(t,r) \end{aligned}$$

Typically used as part of a larger model

$$Y(\mathbf{s};t) = \mu(\mathbf{s};t) + \beta(\mathbf{s}) + \gamma(t) + \kappa(\mathbf{s};t) + \delta(\mathbf{s};t),$$

Additive structures

Additive independent parts:

$$\begin{aligned} Y(\mathbf{s};t) =& \beta(\mathbf{s}) + \gamma(t) \\ \text{cov}[Y(\mathbf{s};t), Y(\mathbf{x};r)] =& C_\beta(\mathbf{s},\mathbf{x}) + C_\gamma(t,r) \end{aligned}$$

Typically used as part of a larger model

$$Y(\mathbf{s};t) = \mu(\mathbf{s};t) + \beta(\mathbf{s}) + \gamma(t) + \kappa(\mathbf{s};t) + \delta(\mathbf{s};t),$$

General:

$$\sum_j w_j C^j(\mathbf{s},t;\mathbf{x},r), w_j > 0$$

is a valid covariance function if each $C^j(\mathbf{s},t;\mathbf{x},r)$ are.

Additive structures

Additive independent parts:

$$\begin{aligned} Y(\mathbf{s};t) =& \beta(\mathbf{s}) + \gamma(t) \\ \text{cov}[Y(\mathbf{s};t), Y(\mathbf{x};r)] =& C_\beta(\mathbf{s},\mathbf{x}) + C_\gamma(t,r) \end{aligned}$$

Typically used as part of a larger model

$$Y(\mathbf{s};t) = \mu(\mathbf{s};t) + \beta(\mathbf{s}) + \gamma(t) + \kappa(\mathbf{s};t) + \delta(\mathbf{s};t),$$

General:

$$\sum_j w_j C^j(\mathbf{s},t;\mathbf{x},r), w_j > 0$$

is a valid covariance function if each $C^j(\mathbf{s},t;\mathbf{x},r)$ are.
Example: Adding two separable covariance functions

$$pC^{s1}(\mathbf{h})C^{t1}(\tau) + qC^{s2}(\mathbf{h})C^{t2}(\tau)$$

Gives non-separable covariance function

Multiplication of correlation functions

Given two valid stationary correlation functions : $C_1(\mathbf{h}, \delta)$ and $C_2(\mathbf{h}, \delta)$ Then the product: $C_{\text{prod}}(\mathbf{h}, \delta)) = C_1(\mathbf{h}, \delta)C_2(\mathbf{h}, \delta)$ is a legal correlation function as well.

Pf:
$\mathcal{F}$ Denote Fourier transform
$\mathcal{F}(C_{\text{prod}}) = \mathcal{F}(C_1) * \mathcal{F}(C_2)$
Since $\mathcal{F}(C_1)$ and $\mathcal{F}(C_2)$ are non negative then also the convolution of these functions is non negative. Thus Bochner's Theorem holds.

Iterate the argument such that the product of valid correlation functions is a valid correlation function.

Spatio-temporal kriging

Assume observed $\{Z(\mathbf{s}_i; t_{ij}, i = 1, ..., m, j = 1, ..., T_i\}$

$$Z(\mathbf{s}_i; t_{ij}) = Y(\mathbf{s}_i; t_{ij}) + \varepsilon(\mathbf{s}_i; t_{ij})$$

Aim: Predict $Y(\mathbf{s}_0; t_0)$

Spatio-temporal kriging

Assume observed $\{Z(\mathbf{s}_i; t_{ij}, i = 1, ..., m, j = 1, ..., T_i\}$

$$Z(\mathbf{s}_i; t_{ij}) = Y(\mathbf{s}_i; t_{ij}) + \varepsilon(\mathbf{s}_i; t_{ij})$$

Aim: Predict $Y(\mathbf{s}_0; t_0)$

- Mathematics as for pure spatial processes
- Notionally more complex
- Matrices involved much larger

Spatio-temporal kriging

Assume observed $\{Z(\mathbf{s}_i; t_{ij}, i = 1, ..., m, j = 1, ..., T_i\}$

$$Z(\mathbf{s}_i; t_{ij}) = Y(\mathbf{s}_i; t_{ij}) + \varepsilon(\mathbf{s}_i; t_{ij})$$

Aim: Predict $Y(\mathbf{s}_0; t_0)$

- Mathematics as for pure spatial processes
- Notionally more complex
- Matrices involved much larger

In general

$$[Y(\mathbf{s}_0; t_0)|\mathbf{Z}] \sim N(\mu(\mathbf{s}_0; t_0) + \mathbf{c}_0'\mathbf{C}_Z^{-1}(\mathbf{Z} - \boldsymbol{\mu}), C_{0,0} - \mathbf{c}_0'\mathbf{C}_Z^{-1}\mathbf{c}_o)$$

Spatio-temporal kriging

Assume observed $\{Z(\mathbf{s}_i; t_{ij}, i = 1, ..., m, j = 1, ..., T_i\}$

$$Z(\mathbf{s}_i; t_{ij}) = Y(\mathbf{s}_i; t_{ij}) + \varepsilon(\mathbf{s}_i; t_{ij})$$

Aim: Predict $Y(\mathbf{s}_0; t_0)$

- Mathematics as for pure spatial processes
- Notionally more complex
- Matrices involved much larger

In general

$$[Y(\mathbf{s}_0; t_0)|\mathbf{Z}] \sim N(\mu(\mathbf{s}_0; t_0) + \mathbf{c}_0'\mathbf{C}_Z^{-1}(\mathbf{Z} - \boldsymbol{\mu}), C_{0,0} - \mathbf{c}_0'\mathbf{C}_Z^{-1}\mathbf{c}_o)$$

Separable covariance functions: $\mathbf{C}_Z$ simplifies considerable (if $T_i = T$ and $t_{ij} = t_j$)

- Similar extensions when parameters are to be estimated

Time-dynamic modelling

- Integro-difference equation models (AR-type)
- Using (partial) differential equations

Common question:

- What covariance structures do we obtain?
- How can we make the models operational (do computation)?

Integro-difference equation models

Assume

$$Y_t(\mathbf{s}) = \int_{D_s} m(\mathbf{v}) Y_{t-1}(\mathbf{s}+\mathbf{v}) d\mathbf{v} + \eta_t(\mathbf{s})$$

Storvik et al(2002)

$$f_{Y,\tau}(\boldsymbol{\omega}) \text{ :Spatial Fourier transform of } C_Y(\mathbf{h},\tau)$$

$$\begin{aligned} f_{Y,0}(\boldsymbol{\omega}) =& \frac{f_\eta(\boldsymbol{\omega})}{1 - f_m(\boldsymbol{\omega}) f_m(-\boldsymbol{\omega})} \\ f_{Y,\tau}(\boldsymbol{\omega}) =& f_{Y,0}(\boldsymbol{\omega}) f_m(-\boldsymbol{\omega})^{|\tau|} \end{aligned}$$

- Non-separable covariance function
- Separable in Fourier domain!
- AR-process in Fourier domain (all frequencies independent)!
- Can also look at properties when time-difference gets smaller

Modelling through (partial) differential equations

Diffusion-injection models (1D in space)

$$\frac{\partial Y(s;t)}{\partial t} = \beta \frac{\partial^2 Y(s;t)}{\partial s^2} - \alpha Y(s;t) + \delta(s;t)$$

Rate of change in Y equal to

- "spread" of Y in space (diffusion) and
- an offset (loss) proportional to Y

Modelling through (partial) differential equations

Diffusion-injection models (1D in space)

$$\frac{\partial Y(s;t)}{\partial t} = \beta \frac{\partial^2 Y(s;t)}{\partial s^2} - \alpha Y(s;t) + \delta(s;t)$$

Rate of change in Y equal to

- "spread" of Y in space (diffusion) and
- an offset (loss) proportional to Y

Example:

- $Y(s;t)$: concentration of some nutrient in soil
- $\beta \frac{\partial^2 Y(s;t)}{\partial s^2}$: physical diffusion
- $-\alpha Y(s;t)$: Removement due to oxidation, leaching or extraction by plants
- $\delta(s;t)$: Injection part, small random additions/removals of the nutrient

Diffusion-injection models

$$\frac{\partial Y(s;t)}{\partial t} = \beta \frac{\partial^2 Y(s;t)}{\partial s^2} - \alpha Y(s;t) + \delta(s;t)$$

If $\{\delta(s;t)\}$ is white noise, one can show:

$$\begin{aligned}\rho_Y(|h|;\tau) =& \text{Cor}[Y(s;t), Y(s+h;t+\tau)]\\ =& 0.5 e^{-h(\alpha/\beta)^{1/2}} \text{Erfc}\left(\frac{2\tau(\alpha/\beta)^{1/2} - h/\beta}{2(\tau/\beta)^{1/2}}\right) +\\ & 0.5 e^{h(\alpha/\beta)^{1/2}} \text{Erfc}\left(\frac{2\tau(\alpha/\beta)^{1/2} + h/\beta}{2(\tau/\beta)^{1/2}}\right)\\ \text{Erfc}(z) \equiv& (2/\pi)^{1/2} \int_z^\infty e^{-v^2} dv, \quad z \geq 0\\ \text{Erfc}(z) =& 2 - \text{Erfc}(-z), \quad z < 0\end{aligned}$$

Diffusion-injection models

$$\frac{\partial Y(s;t)}{\partial t} = \beta \frac{\partial^2 Y(s;t)}{\partial s^2} - \alpha Y(s;t) + \delta(s;t)$$

If $\{\delta(s;t)\}$ is white noise, one can show:

$$\begin{aligned}
\rho_Y(|h|;\tau) =& \mathrm{Cor}[Y(s;t), Y(s+h;t+\tau)] \\
=& 0.5e^{-h(\alpha/\beta)^{1/2}} \mathrm{Erfc}\left(\frac{2\tau(\alpha/\beta)^{1/2} - h/\beta}{2(\tau/\beta)^{1/2}}\right) + \\
& 0.5e^{h(\alpha/\beta)^{1/2}} \mathrm{Erfc}\left(\frac{2\tau(\alpha/\beta)^{1/2} + h/\beta}{2(\tau/\beta)^{1/2}}\right) \\
\mathrm{Erfc}(z) \equiv& (2/\pi)^{1/2} \int_z^\infty e^{-v^2} dv, \quad z \geq 0 \\
\mathrm{Erfc}(z) =& 2 - \mathrm{Erfc}(-z), \quad z < 0
\end{aligned}$$

- SPDE: α, β easily interpretable
- Covariance function: Make density for observations computable

Discretization of diffusion-injection models

Model

$$\frac{\partial Y(s;t)}{\partial t} = \beta\frac{\partial^2 Y(s;t)}{\partial s^2} - \alpha Y(s;t)$$

Discretization:

$$\begin{aligned}\frac{Y(s;t+\Delta_t)-Y(s,t)}{\Delta_t} =&\beta\left\{\frac{Y(s+\Delta_s;t)-2Y(s;t)+Y(s-\Delta_s;t)}{\Delta_s^2}\right\}\\&-\alpha Y(s;t)\end{aligned}$$

or

$$\begin{aligned}Y(s;t+\Delta_t) =&\theta_1 Y(s;t)+\theta_2 Y(s+\Delta_s;t)+\theta_2 Y(s-\Delta_s;t)\\ \theta_1 =&(1-\alpha\Delta_t-2\beta\Delta_t/\Delta_s^2)\\ \theta_2 =&\beta\Delta_t/\Delta_s^2\end{aligned}$$

Discretization of diffusion-injection models

$$Y(s;t+\Delta_t) = \theta_1 Y(s;t) + \theta_2 Y(s+\Delta_s;t) + \theta_2 Y(s-\Delta_s;t)$$

Let $D_s = \{s_0, s_1, ..., s_{n+1}\}, s_j = s_0 + j\Delta_s$. Then

$$\begin{aligned}
Y(s_1;t+\Delta_t) =& \theta_1 Y(s_1;t) + \theta_2 Y(s_2;t) + Y(s_0;t)\\
Y(s_2;t+\Delta_t) =& \theta_1 Y(s_2;t) + \theta_2 Y(s_3;t) + Y(s_1;t)\\
&\vdots\\
Y(s_n;t+\Delta_t) =& \theta_1 Y(s_n;t) + \theta_2 Y(s_{n+1};t) + Y(s_{n-1};t)
\end{aligned}$$

Defining $\mathbf{Y}_t = (Y(s_1;t), ..., Y(s_n;t))', \mathbf{Y}_t^{(b)} = (Y(s_0;t), Y(s_{n+1};t))'$ gives

$$\mathbf{Y}_{t+\Delta_t} = \mathbf{M}\mathbf{Y}_t + \mathbf{M}^{(b)}\mathbf{Y}_t^{(b)}$$

$\mathbf{M}$ sparse, $\mathbf{Y}_t^{(b)}$ are boundary conditions

Discr of diffusion-injection models - adding stochasticity

$$\mathbf{Y}_{t+\Delta_t} = \mathbf{M}\mathbf{Y}_t + \mathbf{M}^{(b)}\mathbf{Y}_t^{(b)}$$

Stochastic version

$$\mathbf{Y}_{t+\Delta_t} = \mathbf{M}\mathbf{Y}_t + \mathbf{M}^{(b)}\mathbf{Y}_t^{(b)} + \boldsymbol{\delta}_t$$

Multivariate AR(1) process (assuming $\{\mathbf{Y}_t^{(b)}\}$ given)

$$\mathbf{\Sigma}_Y^{(k)} = \text{Cov}(\mathbf{Y}_{t+k}, \mathbf{Y}_t) = \mathbf{M}^k\mathbf{\Sigma}_Y, \mathbf{\Sigma}_Y = \mathbf{\Sigma}_Y^{(0)} = \text{Var}(\mathbf{Y}_t)$$

Note:

$$\mathbf{\Sigma}_Y = \mathbf{M}\mathbf{\Sigma}_Y\mathbf{M}' + \mathbf{\Sigma}_\delta$$

Book: Plots showing that discretization typically quite good.
Inference:

- Use Kalman filter if observations also Gaussian
- Other sequential methods possible in more general cases

SPDE's and blurring - space $\mathcal{R}^p$

$$\frac{\partial Y(\mathbf{s};t)}{\partial t} = -\frac{1}{2}\left[2\boldsymbol{\mu}'\frac{\partial Y(\mathbf{s};t)}{\partial \mathbf{s}} - \text{tr}\left\{\left\{\frac{\partial^2 Y(\mathbf{s};t)}{\partial \mathbf{s}\partial \mathbf{s}'}\right\}\boldsymbol{\Sigma}\right\} + 2\lambda Y(\mathbf{s};t)\right] + \delta(\mathbf{s};t)$$

$$\frac{\partial Y(\mathbf{s};t)}{\partial \mathbf{s}} = \left(\frac{\partial Y(\mathbf{s};t)}{\partial s_1}, \cdots, \frac{\partial Y(\mathbf{s};t)}{\partial s_p}\right)'$$

$$\frac{\partial^2 Y(\mathbf{s};t)}{\partial \mathbf{s}\partial \mathbf{s}'} = \left\{\frac{\partial^2 Y(\mathbf{s};t)}{\partial s_i \partial s_j}\right\}$$

$\{\delta(\mathbf{s};t)\}$:Gaussian, independent in time, stationary in space

SPDE's and blurring - space $\mathcal{R}^p$

$$\frac{\partial Y(\mathbf{s};t)}{\partial t} = -\frac{1}{2}\left[2\boldsymbol{\mu}'\frac{\partial Y(\mathbf{s};t)}{\partial \mathbf{s}} - \text{tr}\left\{\left\{\frac{\partial^2 Y(\mathbf{s};t)}{\partial \mathbf{s}\partial \mathbf{s}'}\right\}\boldsymbol{\Sigma}\right\} + 2\lambda Y(\mathbf{s};t)\right] + \delta(\mathbf{s};t)$$

$$\frac{\partial Y(\mathbf{s};t)}{\partial \mathbf{s}} = \left(\frac{\partial Y(\mathbf{s};t)}{\partial s_1},\cdots,\frac{\partial Y(\mathbf{s};t)}{\partial s_p}\right)'$$

$$\frac{\partial^2 Y(\mathbf{s};t)}{\partial \mathbf{s}\partial \mathbf{s}'} = \left\{\frac{\partial^2 Y(\mathbf{s};t)}{\partial s_i \partial s_j}\right\}$$

$\{\delta(\mathbf{s};t)\}$:Gaussian, independent in time, stationary in space

Brown et al (2000):

$$Y(\mathbf{s};t) = \int_0^\infty \exp(-\lambda v)\int_x \phi(\mathbf{x};v\boldsymbol{\mu},v\boldsymbol{\Sigma})\delta(\mathbf{s}-\mathbf{x};t-v))d\mathbf{x}dv$$

Gaussian since $\{\delta(\mathbf{s},t)\}$ Gaussian, nonseparable

$$C_Y(\mathbf{h};\tau) = \int_0^\infty \exp(-\lambda(2v+|\tau|)\int \phi(\mathbf{x};\tau\boldsymbol{\mu},(2v+|\tau|\boldsymbol{\Sigma}))C_\delta(\mathbf{h}-\mathbf{x})d\mathbf{x}dv$$

Fourier transforms

Complicated covariance function:

$$C_Y(\mathbf{h};\tau) = \int_0^\infty \exp(-\lambda(2v + |\tau|) \int \phi(\mathbf{x}; \tau\boldsymbol{\mu}, (2v + |\tau|\boldsymbol{\Sigma}))C_\delta(\mathbf{h} - \mathbf{x})d\mathbf{x}dv$$

Simpler in Fourier domain

$$f_Y(\boldsymbol{\omega};\xi) = |f_G(\boldsymbol{\omega};\xi)|^2 f_\delta(\boldsymbol{\omega})$$

$$\begin{aligned}
f_Y(\boldsymbol{\omega};\xi) &: \text{Fourier transform of } C_Y(\mathbf{h};\tau)\\
f_\delta(\boldsymbol{\omega}) &: \text{Fourier transform of } C_\delta(\mathbf{h})\\
f_G(\boldsymbol{\omega};\xi) &: \text{Fourier transform of } G(\mathbf{s};t)\\
G(\mathbf{s};t) &= \exp(-\lambda t)\phi(\mathbf{s}; t\boldsymbol{\mu}, t\boldsymbol{\Sigma})I(0 \le t < \infty)
\end{aligned}$$

Discretization

Discretizing in time:

$$\frac{\partial Y(\mathbf{s};t)}{\partial t} = -\frac{1}{2}\left[2\boldsymbol{\mu}'\frac{\partial Y(\mathbf{s};t)}{\partial \mathbf{s}} - \text{tr}\left\{\left\{\frac{\partial^2 Y(\mathbf{s};t)}{\partial \mathbf{s}\partial \mathbf{s}'}\right\}\boldsymbol{\Sigma}\right\} + 2\lambda Y(\mathbf{s};t)\right] + \delta(\mathbf{s};t)$$

leads to

$$Y(\mathbf{s};t) \approx \{(\mathbf{I} - (\Delta_t/2)\mathbf{A}_{\Delta_t})Y(\cdot;t-\Delta_t)\}(\mathbf{s}) + \delta(\mathbf{s};t)$$

where $\mathbf{A}_{\Delta_t}$ is linear operator defined by

$$\{(\mathbf{I} - (\Delta_t/2)\mathbf{A}_{\Delta_t})X(\cdot)\}(\mathbf{s}) = \exp(-\lambda\Delta_t)\int \phi(\mathbf{x};\Delta_t\boldsymbol{\mu},\Delta_t\boldsymbol{\Sigma})X(\mathbf{s}-\mathbf{x})d\mathbf{x}$$

Discretizing spatial domain as well:

$$\begin{aligned}
\mathbf{Y}_t \approx& \mathbf{M}\mathbf{Y}_{t-\Delta_t} + \boldsymbol{\delta}_t\\
\mathbf{Y}_t =& (Y(\mathbf{s}_1,t),...,Y(\mathbf{s}_n;t))'\\
\boldsymbol{\delta}_t =& (\delta(\mathbf{s}_1,t),...,\delta(\mathbf{s}_n;t))'
\end{aligned}$$

More general PDE's

(u, v): Spatial or spatio-temporal coordinates
Partial differential equations (PDE)s (nonstochastic):

$$\left\{a\frac{\partial^2}{\partial u^2} + 2h\frac{\partial^2}{\partial u \partial v} + b\frac{\partial^2}{\partial v^2} + 2g\frac{\partial}{\partial u} + 2f\frac{\partial}{\partial v} + c\right\} Y(u, v) = 0$$

More general PDE's

(u, v): Spatial or spatio-temporal coordinates
Partial differential equations (PDE)s (nonstochastic):

$$\left\{a\frac{\partial^2}{\partial u^2} + 2h\frac{\partial^2}{\partial u \partial v} + b\frac{\partial^2}{\partial v^2} + 2g\frac{\partial}{\partial u} + 2f\frac{\partial}{\partial v} + c\right\} Y(u, v) = 0$$

Stochastic PDE

$$\left\{a\frac{\partial^2}{\partial u^2} + 2h\frac{\partial^2}{\partial u \partial v} + b\frac{\partial^2}{\partial v^2} + 2g\frac{\partial}{\partial u} + 2f\frac{\partial}{\partial v} + c\right\} Y(u, v) = \delta(u, v)$$

$\{\delta(u, v)\}$ a zero-mean, two-dimensional stochastic process,

- random impulses from smaller-order contributions
- uncertainty in the scientific relationship expressed through the PDE

Stochastic PDE

$$\left\{a\frac{\partial^2}{\partial u^2} + 2h\frac{\partial^2}{\partial u\partial v} + b\frac{\partial^2}{\partial v^2} + 2g\frac{\partial}{\partial u} + 2f\frac{\partial}{\partial v} + c\right\} Y(u,v) = \delta(u,v)$$

Question: Given model above, what is

- $E[Y(u,v)]$?
- $\text{var}[Y(u,v)]$?
- $\text{cov}[Y(u,v), Y(u',v')]$?

Zero mean δ and linearity of PDE: $E[Y(u,v)] = 0$
(Can always add expectation structure)
Covariance structure difficult to derive in general

Whittle (1963)

$Y(\mathbf{s}), \mathbf{s} \in \mathcal{R}^d$, $\{\delta(\mathbf{s})\}$ zero-mean white noise:

$$\left\{\frac{\partial^2}{\partial s_1^2} + \cdots + \frac{\partial^2}{\partial s_d^2} - \alpha^2\right\}^p Y(\mathbf{s}) = \delta(\mathbf{s})$$

has covariance function

$$\begin{aligned} C_Y(\mathbf{h}) \propto & \{||\mathbf{h}||/\theta_1\}^{\theta_2} K_{\theta_2}(||\mathbf{h}||/\theta_1) \\ \theta_1 =& 1/\alpha > 0 \\ \theta_2 =& 2p - d/2 > 0 \end{aligned}$$

$K_{\theta_2}(\cdot)$ modified Bessel function of the second kind
The Matern covariance function

Geostatistical and SPDE models

- Equivalent models in some cases

Geostatistical and SPDE models

- Equivalent models in some cases
- Different interpretations
  - of model
  - of parameters

Geostatistical and SPDE models

- Equivalent models in some cases
- Different interpretations
  - of model
  - of parameters
- Different ways of doing inference
  - Geostatistical models: Density directly given in Gaussian case
  - For large models: SPDEs can have computational benefits

Time series of spatial processes

Assume $D_t = \{0, 1, 2, ...\}$, discrete-time model
$\mathbf{Y}_t(\cdot) = \{Y(\mathbf{s}; t), \mathbf{s} \in D_s\}$ (possible infinite dimension)
AR-type model:

$$\mathbf{Y}_t(\mathbf{s}) = \mathcal{M}_t(\mathbf{s}, \mathbf{Y}_{t-1}) + \delta_t(\mathbf{s})$$

$\{\delta_t(\mathbf{s})\}$ independent in time.

Time series of spatial processes

Assume $D_t = \{0, 1, 2, ...\}$, discrete-time model
$\mathbf{Y}_t(\cdot) = \{Y(\mathbf{s}; t), \mathbf{s} \in D_s\}$ (possible infinite dimension)
AR-type model:

$$\mathbf{Y}_t(\mathbf{s}) = \mathcal{M}_t(\mathbf{s}, \mathbf{Y}_{t-1}) + \delta_t(\mathbf{s})$$

$\{\delta_t(\mathbf{s})\}$ independent in time.
Special case

$$\mathbf{Y}_t(\mathbf{s}) = \int_{R^d} m(\mathbf{s}, \mathbf{x})\mathbf{Y}_{t-1}(\mathbf{x}) + \delta_t(\mathbf{s})$$

- Similar to discretized models considered earlier
- Now: Do modelling directly as AR-type models

Linear spatial time series

$$\mathbf{Y}_t = \mathbf{M}_t \mathbf{Y}_{t-1} + \boldsymbol{\delta}_t$$

gives

$$\begin{aligned}\boldsymbol{\Sigma}_{Y,t} =& \mathrm{Var}(\mathbf{Y}_t) = \mathbf{M}_t \boldsymbol{\Sigma}_{Y,t-1} \mathbf{M}_t' + \boldsymbol{\Sigma}_{\delta,t} \\ \mathrm{cov}(\mathbf{Y}_{t+\tau} \mathbf{Y}_t) =& \mathbf{M}_{t+\tau} \mathbf{M}_{t+\tau-1} \cdots \mathbf{M}_{t+1} \boldsymbol{\Sigma}_{Y,t}\end{aligned}$$

Allow for

- nonstationarity in time ($\mathbf{M}_t$ change with time)
- nonstationarity in space ($M_t(\mathbf{s}, \mathbf{x})$ depend on spatial locations)

Example

$D_s = [0, 40], \Delta_s = 1$ and $D_t = [0, 0.8], \Delta_t = 0.01$

$$\mathbf{M}_t = h_t \mathbf{M},$$

$$\mathbf{M} = \begin{pmatrix} \widetilde{\mathbf{M}}_1 & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{M}}_2 \end{pmatrix} \qquad h_t = \begin{cases} 1, & \text{if } 1\Delta_t \leq t \leq 30\Delta_t \\ -1, & \text{if } 31\Delta_t \leq t \leq 60\Delta_t \\ 1, & \text{if } 61\Delta_t \leq t \leq 80\Delta_t \end{cases}$$

$$\widetilde{\mathbf{M}}_1 = \begin{pmatrix} .90 & .01 & 0 & \cdots & 0 \\ .01 & .90 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & .90 & .01 \\ 0 & 0 & & .01 & .90 \end{pmatrix}, \quad \widetilde{\mathbf{M}}_2 = \begin{pmatrix} .20 & .01 & 0 & \cdots & 0 \\ .01 & .20 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & .20 & .01 \\ 0 & 0 & & .01 & .20 \end{pmatrix}$$

Example

$D_s = [0, 40], \Delta_s = 1$ and $D_t = [0, 0.8], \Delta_t = 0.01$

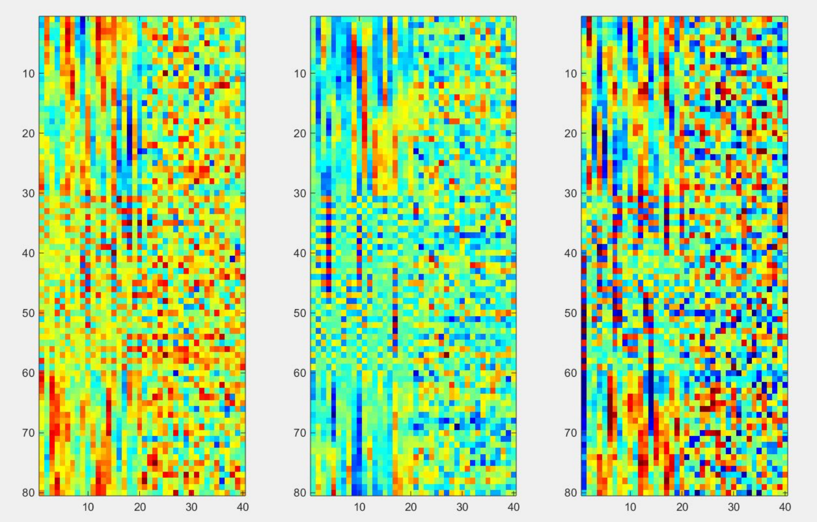
$$\mathbf{M}_t = h_t \mathbf{M},$$

$$\mathbf{M} = \begin{pmatrix} \widetilde{\mathbf{M}}_1 & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{M}}_2 \end{pmatrix} \qquad h_t = \begin{cases} 1, & \text{if } 1\Delta_t \le t \le 30\Delta_t \\ -1, & \text{if } 31\Delta_t \le t \le 60\Delta_t \\ 1, & \text{if } 61\Delta_t \le t \le 80\Delta_t \end{cases}$$

$$\widetilde{\mathbf{M}}_1 = \begin{pmatrix} .90 & .01 & 0 & \cdots & 0 \\ .01 & .90 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & .90 & .01 \\ 0 & 0 & & .01 & .90 \end{pmatrix}, \quad \widetilde{\mathbf{M}}_2 = \begin{pmatrix} .20 & .01 & 0 & \cdots & 0 \\ .01 & .20 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & .20 & .01 \\ 0 & 0 & & .01 & .20 \end{pmatrix}$$

- 6 distinct “regions”
- Can specify $\mathbf{\Sigma}_{\delta,t}$ such that $\mathbf{\Sigma}_{Y,t} = \mathbf{\Sigma}$, i.e. $\mathbf{\Sigma}_{\delta,t} = \mathbf{\Sigma} - \mathbf{M}_t \mathbf{\Sigma} \mathbf{M}_t'$

Example - simulations



Higher order models

In general

$$[\mathbf{Y}_0, \mathbf{Y}_1, ..., \mathbf{Y}_T] = [\mathbf{Y}_0][\mathbf{Y}_1|\mathbf{Y}_0][\mathbf{Y}_2|\mathbf{Y}_0, \mathbf{Y}_1]\cdots[\mathbf{Y}_T|\mathbf{Y}_0, ..., \mathbf{Y}_{T-1}]$$

Need simplifications for

- simpler modelling
- possibility of estimating parameters
- computation

Typical simplifications:

- $[\mathbf{Y}_t|\mathbf{Y}_{t-1}, ..., \mathbf{Y}_0] = [\mathbf{Y}_t|\mathbf{Y}_{t-1}, ..., \mathbf{Y}_{t-r}]$, Vector AR-structure
- $[\mathbf{Y}_t|\mathbf{Y}_{t-1}, ..., \mathbf{Y}_{t-r}] = [\mathbf{Y}_r|\mathbf{Y}_{r-1}, ..., \mathbf{Y}_0]$, stationarity in time
- spatial sparsity in transitions
- stationarity in space

Vector autoregressive model

Linear model

$$\mathbf{Y}_t = \mathbf{M}_t\mathbf{Y}_{t-1} + \boldsymbol{\delta}_t$$

In general $T \times n^2$ parameters in $\mathbf{M}$'s

Simplifications

- $\mathbf{M}_t = \mathbf{M}$: Vector autoregressive model
- $M_{ij} = 0$ for $||\mathbf{s}_i - \mathbf{s}_j|| > h$

Similarly:

- $\boldsymbol{\Sigma}_{\delta,t} = \boldsymbol{\Sigma}_\delta$
- $\boldsymbol{\Sigma}_\delta$ sparse