Statistical and dynamical models of Spatio-temporal Processes

Adway Mitra

25 January 2021

Geo-statistical Equation

- $X(s, t) = \mu(s, t) + \eta(s, t) + \epsilon(s, t)$
- $\mu(s, t)$: local mean, spatially or temporally stationery
- $\eta(s, t)$: dynamic process model containing spatial or temporal correlations
- $\epsilon(s, t)$: random noise

Hierarchical model for Spatio-temporal Process

- Hierarchical model: Data model + Process model + Parameter model
- *Parameter model*: parameter values sampled from a prior distribution
- *Process model*: describes the process (including spatio-temporal dynamics) based on the parameters
- *Data model*: describes the observations, in terms of the process
- $Z(s, t)$ is the description of the process (latent variable)
- $X(s, t)$ are the observations

Template of a hierarchical model

$X(s,t) = g(Z(s,t), Y(s,t))$

HYPERPARAMETERS

ENDOGENOUS

EXTRANEOUS

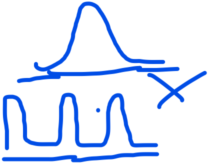
EXOGENOUS

MODEL

- $\theta \sim p(\eta)$ [Parameter Model]
- LATENT $Z(s,t) \sim f(\theta)$ [Process Model]
- $X(s,t) \sim g(Z,Y)$ [Data/observation Model]; Y: co-variates

  Other variables can be measured outside model
- Simulation/forward problem: Generate X by sampling in order
- Inverse problem: estimate Z, θ from X, Y using Bayes Theorem
- How to choose p, f, g?

Designer's choice



Spatial Process

$\eta(s) = a_1 Z(1) + a_2 Z(2) + a_5 Z(5)$
$+ b_1 Y(1) + b_3 Y(3) + b_5 Y(5)$
$+ b_s Y(s)$

$Z(1)$ $Y(5)$
$Z(2)$ $Y(1)$
$Z(5)$ $Y(3)$ $Y(9)$

LOCAL GLOBAL

- Data Model: $X(s) \sim \mathcal{N}(\mu(s) + \eta(s), \sigma)$ RANDOM NOISE
- Observations at each time-point is a realization from this model END EXO.
- $\eta(s) = AZ(s) + BY(s)$
- ϵ is managed by σ
- $\mu(s), Z(s)$: contains covariance between different locations
- Y are co-variates which are extraneous to the model
- A, B represent transformation coefficients

$A = [a_1 \; a_2 \; 0 \; 0 \; a_5 \; \ldots \; 0]$
$B = [b_1 \; 0 \; b_3 \; 0 \; \ldots]$

Spatial Process- vectorized

$C = \begin{bmatrix} \sigma & & 0 & 0 \\ 0 & \sigma & & \\ & & \ddots & \\ 0 & 0 & & \sigma \end{bmatrix}$

$\mathcal{N}(0, \sigma)$

Vector

$X = \begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(S) \end{bmatrix} \quad \begin{bmatrix} \mu(1) \\ \mu(2) \\ \vdots \\ \mu(S) \end{bmatrix} \quad \begin{bmatrix} \eta(1) \\ \vdots \\ \eta(S) \end{bmatrix}$

No. of locations

Multi. Gaussian

- Data Model: $X \sim \mathcal{N}(\underbrace{\mu + \eta}_{\text{vector}}, \sigma I)$ (Identity matrix) σ
- $\eta = \underset{s \times s}{A} Z + \underset{s \times s}{B} Y$
- X, μ, η, Z, Y are vectors of length S
- A, B are transformation matrices
- Vectorization allows the influence of other locations
- If A, B diagonal matrices: back to the previous model

Interpretation

- μ: local effect (eg. all locations have local mean temperature)
- Z: global effect (eg. during a heat wave, temperatures at all locations are affected by different degrees)
- A: transfer the effect of heat wave on the local observations
- Y: covariates (eg. humidity, rainfall etc)
- B: effect of co-variates (eg. role of humidity on temperature)

Temporal Process

- Data Model: $X(t) \sim \mathcal{N}(\mu(t) + \eta(t), \sigma)$
- $\eta(t) = AZ(t) + BY(t)$
- $\mu(t), Z(t)$: contains covariance between different time-points
- Y are co-variates which are extraneous to the model
- A, B represent transformation matrices

Interpretation

- μ: seasonal component (eg. all months have seasonal mean temperature)
- Z: trend component (eg. a heat wave that lasts for a few days, global warming)

Gaussian Process

- Consider a (finite or infinite) set of random variables X_1, X_2, ...
- Consider any random finite subset $\{X_{i1}, X_{i2}, \ldots, X_{iN}\}$
- Then we have $(X_{i1}, X_{i2}, \ldots, X_{iN}) \sim \mathcal{N}(\mu, \Sigma)$
- $\mu(s)$: mean function (a function of s)
- $\Sigma(s, s')$: covariance function (a function of $||s - s'||$)