Lecture 6 Multivariate Spatial Modeling

Shiwei Lan[1]

[1]School of Mathematical and Statistical Sciences
Arizona State University

STP598 Spatiotemporal Analysis
Fall 2020

Multivariate spatial modeling

- *Multivariate*: multiple (i.e. more than one) outcomes are measured at each spatial unit.
- Multivariate point-referenced data:
  - Levels of pollutants including ozone, nitric oxide, carbon monoxide, $PM_{2.5}$ etc. are measured at monitoring station
  - Surface temperature, precipitation, and wind speed in atmospheric modeling.
  - In examining real estate markets, both selling price and total rental income observed for individual property...
- Multivariate areal data:
  - In public health, supplies counts or rates for a number of diseases for each county or administrative unit.

Table of Contents

Multivariate spatial modeling for point-referenced data

- We model multivariate point-referenced data by either a *conditioning approach (kriging with external drift)* or a *joint approach (co-kriging)*.
- Inference focuses upon three major aspects:
  1. estimate associations among the processes
  2. estimate the strength of spatial association for each process
  3. predict the processes at arbitrary locations
- Let $\mathbf{Y}(\mathbf{s}) = (Y_1(\mathbf{s}), \cdots, Y_p(\mathbf{s}))^T$ be a $p \times 1$ vector of process referenced at $\mathbf{s} \in \mathcal{D}$.
- We seek to capture the association both within components of $\mathbf{Y}(\mathbf{s})$ and across $\mathbf{s}$.

Cross- variograms and covariances

- Assume $\mathrm{E}(Y(\mathbf{s}+\mathbf{h})-Y(\mathbf{s}))=0$. The joint second order (weak) stationarity hypothesis defines the *cross-variogram* as

$$\gamma_{ij}(\mathbf{h})=\frac{1}{2}\mathrm{E}(Y_i(\mathbf{s}+\mathbf{h})-Y_i(\mathbf{s}))(Y_j(\mathbf{s}+\mathbf{h})-Y_j(\mathbf{s})) \tag{1}$$

  - $\gamma_{ij}(\mathbf{h})=\gamma_{ij}(-\mathbf{h})$.
  - $|\gamma_{ij}(\mathbf{h})|^2\le\gamma_{ii}(\mathbf{h})\gamma_{jj}(\mathbf{h})$.
- The *cross-covariance* function is defined as

$$C_{ij}(\mathbf{h})=\mathrm{E}(Y_i(\mathbf{s}+\mathbf{h})-\mu_i)(Y_j(\mathbf{s})-\mu_j) \tag{2}$$

  - $C_{ij}(\mathbf{h})\ne C_{ji}(\mathbf{h})$.
  - $|C_{ij}(\mathbf{h})|^2\le C_{ii}(0)C_{jj}(0)$. $\quad |C_{ij}(\mathbf{h})|^2\le C_{ii}(\mathbf{h})C_{jj}(\mathbf{h})$?
- Eg: spatial delay models (Wackernagel, 2003): $Y_2(\mathbf{s})=aY_1(\mathbf{s}+\mathbf{h}_0)+\epsilon(\mathbf{s})$.

Cross- variograms and covariances

- How to express $\gamma_{ij}(\mathbf{h})$ in terms of C_{ij}?

Cross- variograms and covariances

- How to express $\gamma_{ij}(\mathbf{h})$ in terms of C_{ij}?

$$\gamma_{ij}(\mathbf{h}) = C_{ij}(0) - \frac{1}{2}(C_{ij}(\mathbf{h}) + C_{ij}(-\mathbf{h})) \tag{3}$$

- Cross-variogram only captures the even term of the cross-covariance function!
- *Pseudo* cross-variogram:
  - Clark et al. (1989) proposed $\pi_{ij}^c(\mathbf{h}) = \mathrm{E}(Y_i(\mathbf{s}+\mathbf{h}) - Y_j(\mathbf{s}))^2$
  - Myers (1991) defined $\pi_{ij}^m(\mathbf{h}) = \mathrm{Var}(Y_i(\mathbf{s}+\mathbf{h}) - Y_j(\mathbf{s}))$
  - $\pi_{ij}^c(\mathbf{h}) = \pi_{ij}^m(\mathbf{h}) + (\mu_i - \mu_j)^2$
- Positive, may not be even. Co-kriging uses $\pi_{ij}^m(\mathbf{h})$.

Co-kriging

- Given $\mathbf{Y} = (\mathbf{Y}(\mathbf{s}_1), \cdots, \mathbf{Y}(\mathbf{s}_p))^T$, we want to know $\mathbf{Y}(\mathbf{s}_0)$.
- Different from multi-output kriging for a univariate spatial process at multiple locations!
- In the regression framework, we could require the predicted value $\hat{\mathbf{Y}}(\mathbf{s}_0)$

$$\hat{\mathbf{Y}}(\mathbf{s}_0) = \sum_{i=1}^{n} \Lambda_i \mathbf{Y}(\mathbf{s}_i), \quad \sum_{i=1}^{n} \Lambda_i = I \tag{4}$$

$$\min_{\Lambda} \quad \mathrm{tr}\mathrm{E}(\hat{\mathbf{Y}}(\mathbf{s}_0) - \mathbf{Y}(\mathbf{s}_0))(\hat{\mathbf{Y}}(\mathbf{s}_0) - \mathbf{Y}(\mathbf{s}_0))^T \tag{5}$$

- $\mathrm{tr}\mathrm{E}(\hat{\mathbf{Y}}(\mathbf{s}_0) - \mathbf{Y}(\mathbf{s}_0))(\hat{\mathbf{Y}}(\mathbf{s}_0) - \mathbf{Y}(\mathbf{s}_0))^T = \mathrm{E}(\hat{\mathbf{Y}}(\mathbf{s}_0) - \mathbf{Y}(\mathbf{s}_0))^T(\hat{\mathbf{Y}}(\mathbf{s}_0) - \mathbf{Y}(\mathbf{s}_0))$.

Co-kriging

- Assume a multivariate Gaussian spatial process $\mathbf{Y}(\mathbf{s})$ with zero mean.
- Suppose we have a finite cross-covariance function (*permissible* cross-variogram).
- Denote $\mathbf{Y} = (\mathbf{Y}(\mathbf{s}_1)^T, \cdots, \mathbf{Y}(\mathbf{s}_p)^T)^T$. Then we have $np \times np$ covariance matrix $\Sigma_{\mathbf{Y}}$.
- Denote $np \times 1$ vector $\mathbf{c}_0$ with jl-th element $c_{0j,l} = \mathrm{Cov}(Y_1(\mathbf{s}_0), Y_l(\mathbf{s}_j))$. Then

$$\mathrm{E}(Y_1(\mathbf{s}_0)|\mathbf{Y}) = \mathbf{c}_0^T \Sigma_{\mathbf{Y}}^{-1} \mathbf{Y} \tag{6}$$

$$\mathrm{Var}(Y_1(\mathbf{s}_0)|\mathbf{Y}) = C_{11}(0) - \mathbf{c}_0^T \Sigma_{\mathbf{Y}}^{-1} \mathbf{c}_0 \tag{7}$$

Co-kriging

- *Intrinsic* co-kriging assumes $C(\mathbf{h}) = \rho(\mathbf{h})T$ with a valid correlation function $\rho(\cdot)$ and a positive definite covariance matrix T.
- Therefore $\Sigma_{\mathbf{Y}} = R \otimes T$, and

$$\mathrm{E}(Y_1(\mathbf{s}_0)|\mathbf{Y}) = \mathbf{c}_0^T \Sigma_{\mathbf{Y}}^{-1} \mathbf{Y} = t_{11} \mathbf{r}_0^T R^{-1} \tilde{\mathbf{Y}}_1 \tag{8}$$

  where $\mathbf{r}_0 = (\rho(\mathbf{s}_0 - \mathbf{s}_j))$ and $\tilde{\mathbf{Y}}_1$ is formed by the first components of $\mathbf{Y}(\mathbf{s}_j)$'s.
- Data availability (missing data):
  - *isotopy*: data is available for each variable at all sampling points
  - partial *heterotopy*: some variables share some sample locations
  - entirely *heterotopic*: the variables have no sample locations in common
- *Collocated* co-kriging makes use of $Y_l(\mathbf{s}_j)$ to help predict $Y_1(\mathbf{s}_0)$.

Cross-covariance function

- Consider a vector-valued spatial process $\{\mathbf{w}(\mathbf{s}) \in \mathbb{R}^p : \mathbf{s} \in \mathcal{D}\}$. Assume $\mathrm{E}[bw(\mathbf{s})] = 0$.
- The *cross-covariance function* is a matrix-valued function $\mathbf{C}(\mathbf{s}, \mathbf{s}')$ with (i, j)-th entry

$$C_{ij}(\mathbf{s}, \mathbf{s}') = \mathrm{Cov}(w_i(\mathbf{s}), w_j(\mathbf{s}')) = \mathrm{E}[w_i(\mathbf{s}) w_j(\mathbf{s}')] \tag{9}$$

- Let $w_i(\mathbf{s}) = Y_i(\mathbf{s}) - \mu_i$. Then $C(\mathbf{s}, \mathbf{s}') = \mathrm{Cov}(\mathbf{w}(\mathbf{s}), \mathbf{w}(\mathbf{s}')) = \mathrm{E}[\mathbf{w}(\mathbf{s})\mathbf{w}(\mathbf{s}')^T]$.
- We require $C(\mathbf{s}, \mathbf{s}') = C(\mathbf{s}', \mathbf{s})^T$.
- $\mathbf{w}(\mathbf{s})$ is *stationary* if $C_{ij}(\mathbf{s}, \mathbf{s}') = C(\mathbf{h})$ is a function of $\mathbf{h} = \mathbf{s} - \mathbf{s}'$. Symmetric cross-covariance implies $C(-\mathbf{h}) = C(\mathbf{h})$.
- $\mathbf{w}(\mathbf{s})$ is *isotropy* if further $C_{ij}(\mathbf{s}, \mathbf{s}') = C_{ij}(\|\mathbf{h}\|)$, which directly implies symmetry in cross-covariance function.

Separable models

- Separable models for p-dimensional $\mathbf{Y}(\mathbf{s})$ assume the following cross-covariance function

$$C(\mathbf{s},\mathbf{s}') = \rho(\mathbf{s},\mathbf{s}')\cdot T \tag{10}$$

- The covariance matrix for $\mathbf{Y}$ has the following Kronecker product structure

$$\Sigma_{\mathbf{Y}} = H \otimes T \tag{11}$$

where $H_{ij} = \rho(\mathbf{s}_i, \mathbf{s}_j)$.

- Pros: $|\Sigma_{\mathbf{Y}}| = |H|^p \cdot |T|^n$, $\Sigma_{\mathbf{Y}}^{-1} = H^{-1} \otimes T^{-1}$.
- Cons: *coherence* $\frac{\mathrm{Cov}(Y_\ell(\mathbf{s}), Y_{\ell'}(\mathbf{s}+\mathbf{h}))}{\sqrt{\mathrm{Cov}(Y_\ell(\mathbf{s}), Y_\ell(\mathbf{s}+\mathbf{h}))\mathrm{Cov}(Y_{\ell'}(\mathbf{s}), Y_{\ell'}(\mathbf{s}+\mathbf{h}))}} = \frac{T_{\ell\ell'}}{T_{\ell\ell}T_{\ell'\ell'}}$ regardless of $\mathbf{s}$ and $\mathbf{h}$: identical spatial dependence for each component of $\mathbf{Y}(\mathbf{s})$!

Interpolation, (spatial) prediction and regression

- Consider response process $Z(\mathbf{s})$ and a vector of covariates $\mathbf{x}(\mathbf{s})$.
- Partition our set of sites into three mutually disjoint groups
  1. S_Z: the sites where only the response $Z(\mathbf{s})$ has been observed
  2. S_X: the the set of sites where only the covariates have been observed
  3. S_{ZX}: the set where both $Z(\mathbf{s})$ and the covariates have been observed
  4. S_U: the set of sites where no observations have been taken.
- Formalize three types of inference questions:
  1. *interpolation*: concerns $Y(\mathbf{s})$ when $\mathbf{s} \in S_X$
  2. *prediction*: concerns $Y(\mathbf{s})$ when $\mathbf{s} \in S_U$
  3. *spatial regression*: concerns the functional relationship between $X(\mathbf{s})$ and $Y(\mathbf{s})$ at an arbitrary site $\mathbf{s}$, along with other covariate information $\mathrm{U}(\mathbf{s})$, $\mathrm{E}[Y(\mathbf{s})|X(\mathbf{s}), \mathrm{U}(\mathbf{s})]$.

Interpolation, (spatial) prediction and regression

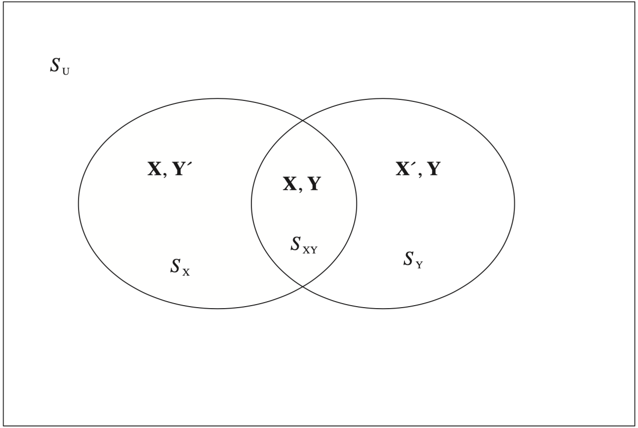

Figure 9.1 *A graphical representation of the S sets. Interpolation applies to locations in S_X, prediction applies to locations in S_U, and regression applies to all locations. $\mathbf{X}_{aug} = (\mathbf{X}, \mathbf{X}')$, $\mathbf{Y}_{aug} = (\mathbf{Y}, \mathbf{Y}')$.*

Table of Contents