STOR 834, p.107 4/10/25

Background on spatial statistics:

Typical model has $\underset{\sim}{Z} = (Z_1, \dots, Z_n)^T$, n spatial locations

$$Z_i = \sum_{k=1}^{p} x_{ik}\beta_k + \varepsilon_i$$

$\varepsilon_i \sim N[0, \sigma^2]$, $\quad \mathrm{cov}(\varepsilon_i, \varepsilon_j) = \sigma^2 c_{ij}$

+ joint normality

c_{ij}: correlation between sites i and j

e.g. $c_{ij} = e^{-d_{ij}/\rho}$ or $e^{-(d_{ij}/\rho)^2}$ where

d_{ij} is distance, ρ is "range".

Many other possibilities, e.g. Matérn.

Write as : $\underset{\sim}{Z} \sim \mathcal{N}_n[X\beta, \sigma^2 C(\underset{\sim}{\phi})]$

Parameters $\beta, \sigma^2, \underset{\sim}{\phi}$

Estimation method:

MLE min $(2\pi\sigma^2)^{-n/2} |C(\underset{\sim}{\phi})|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(Z-X\beta)^T C(\phi)^{-1}(Z-X\beta)\right]$

REML: $(2\pi\sigma^2)^{-(n-p)/2} |X^T C(\phi)^{-1} X|^{-1/2} |C(\phi)|^{-\frac{1}{2}} \exp[- \cdots]$

or Bayes

Multivariate, e.g. $\underset{\sim}{Z}_q$, $q=1,\dots,r$, r spatial processes

Possible model is "co-regionalization"

$$\underset{\sim}{Z}_q = \sum_{s=1}^{q} a_{qs} \underset{\sim}{\eta}_s$$

where $\underset{\sim}{\eta}_s$, $s=1,\dots,\cancel{q}\, r$ are r independent univariate processes and $A = (a_{qs},\ 1 \le s \le q \le r)$ is a lower-triangular (LT) matrix.

Go back to R20 paper

$$\{Y_t(\underset{\sim}{s}),\ t=1,\dots,T,\ \underset{\sim}{s} \in \mathcal{S}\} = GEV(\underset{\sim}{\theta}(\underset{\sim}{s})) \qquad (\mathrm{I})$$

↑ 5-dim process

Assume each $X^T\beta$ process is just a common mean μ

$$\underset{\sim}{\Theta}(\underset{\sim}{s}) \sim MVN[\underbrace{A}_{\substack{5\times 5 \\ \text{LT matrix}}},\ \underbrace{\mu_1,\dots,\mu_5,\ \sigma_1^2,\dots,\sigma_5^2,\ \phi_1,\dots\phi_5}_{\text{pars of } \eta_1,\dots,\eta_5}] \qquad (\mathrm{II})$$

Hierarchical model: alternate between updating $\underset{\sim}{\Theta}$ itself (eqn I) or updating the parameters conditional on Θ (II)

Alternative "measurement error" approach:

Assume "true Θ" given by (II)

$\hat{\Theta}$ = "estimated Θ" are the MLEs from Step I.

Then assume $\hat{\Theta} = \Theta + \underset{\sim}{e}$

vector of estimation errors (dimension $5n$)

Assume $\underset{\sim}{e} \sim N[0, W]$

to be determined

In this approach, we treat W as <u>known</u>

Rewrite in the form

$$\Theta \sim N[X\beta, V] \quad \leftarrow (5n)\text{-dim process}$$

$$\underset{\sim}{\hat{\Theta}} \mid \underset{\sim}{\Theta} \sim N[\underset{\sim}{\Theta}, W]$$

Combine as $\hat{\Theta} \sim N[X\beta, V+W]$

Estimate β and V by MLE or REML, then do

$$\Theta \mid \hat{\Theta} \sim N\left[X\beta + V(V+W)^{-1}(\hat{\Theta} - X\beta), \; V(V+W)^{-1}W \right]$$

"Kriging formula"

Estimation of W: use bootstrap

[This is the method in R20. However, we could consider alternatives, e.g. a parametric model for W and then include in the estimation process]

Application:

Precipitation extremes example in text

My climate examples in my Cardiff talk on my website.

Second approach: max-stable processes

Some background on point process.

Assume N is a point process on some "nice space" $\mathbb{E}$.

i.e. $N(A)$ is a non-neg integer for each measurable set A

"simple" means no double points

Defn 1: N is Poisson with mean measure μ if

(a) $P\{N(A)=k\} = \dfrac{e^{-\mu(A)}(\mu(A))^k}{k!}$ if $k \ge 0$, $\mu(A)<\infty$

(b) For disjoint $A_1, \dots, A_k$, $N(A_1), \dots, N(A_k)$ are indep

Examples:

1. Simple uniform process on $\mathbb{R}_+$:
$$\mu((0,x)) = x \text{ for any } 0<x<\infty$$

2. Inverse Poisson process (IPP):
$$\nu((x,\infty)) = \frac{1}{x} \quad \text{any } x>0.$$

Mappings of Poisson processes

$$T : \mathbb{E} \to \mathbb{E}'$$

Inverse map: $T^{\leftarrow}(A') = \{ e \in \mathbb{E} : T(e) \in A' \}$
any $A' \subseteq \mathbb{E}'$

Prop 2: Suppose $T : \mathbb{E} \to \mathbb{E}'$ measurable map s.t. if $K' \subseteq \mathbb{E}'$ is compact then $T^{\leftarrow}(K')$ is also compact

If N is a PRM(μ) on $\mathbb{E}$, then $N' = N \circ T^{\leftarrow}$ is a PRM(μ') on $\mathbb{E}'$, where $\mu' = \mu \circ T^{-1}$.

In words: ~~if $x \in \mathbb{E}$ is~~ each x in N on $\mathbb{E}$ is transformed into $T(x)$ in N' on $\mathbb{E}'$, then N' is a PRM with ~~$\mu' =$~~ $\mu'(A') = \mu(T^{\leftarrow} A')$
each $A' \subseteq \mathbb{E}'$.

Prop. 3: Suppose N is a IPP, with points $\{R_i, i = 1, 2, \ldots\}$ on $(0, \infty)$ and $\{W_i, i = 1, 2 \ldots\}$ are indep. RVs with mean 1. Then $N' = \{(R_i W_i), i = 1, 2, \ldots\}$ is also an IPP.

Proof: Expand N as a 2-d PP on ~~$\times(0,\infty)$~~ $(0,\infty)^2$ with points at (R_i, W_i). Then for any mble $A \subseteq (0,\infty)$

$$\mu((x,\infty) \times A) = \frac{F(A)}{x}$$

Consider map: $T((R_i, W_i)) = R_i W_i$.

$$\mu'((z,\infty)) = \mu\{(x,w) : xw > z\}$$

$$= \int_{(0,\infty)} F(dw) \cdot \int_{(0,\infty)} I\left(x > \frac{z}{w}\right) \cancel{d}\, \nu(dx)$$

$$= \int_{(0,\infty)} F(dw) \cdot \frac{w}{z} = \frac{1}{z}$$

Prop. 4 Suppose N is a IPP with points $\{R_i\}$.

Let $M = \max_i R_i$. Then $M \sim$ unit Fréchet

Pf: $M < z$ is equivalent to $N((z,\infty)) = 0$

Prob is $e^{-\mu((z,\infty))} = \cancel{1}\, e^{-1/z}$ for $z > 0$.