Spatial Extremes: the Smith Model

Robert L. Wolpert
Department of Statistical Science
Duke University, Durham, NC, USA

May 5, 2017

1 Introduction

2 Model

We are concerned with the so-called “Smith model” (Smith, 1990), one of the earliest max-stable models proposed for studying dependent spatial extremes. Let $\nu_u(du)$ be a $\sigma$-finite positive measure on $\mathbb{R}_+$ and let $S = \mathbb{R}^2$ be “space”, with elements denoted $s$ or $\sigma$ (for later space-time applications we may need to consider a product space $\Omega = S \times \mathcal{F}$ and, for other applications, spaces that are larger or more abstract). Let $N(du d\sigma) \sim \text{Po}(\nu(du d\sigma))$ be a Poisson random measure on the Borel sets of $\mathbb{R}_+ \times S$ with product mean measure

$$\nu(du d\sigma) := \nu_u(du) d\sigma$$

Fix a kernel function $k(s; \sigma) \geq 0$ on $S^2$ and define a stochastic process indexed by $s \in S$ by

$$X_s := \sup_j \nu_j k(s; \sigma_j),$$

where $\{(\nu_j, \sigma_j)\} \subset \mathbb{R}_+ \times S$ comprise the random countably-infinite support of $N$. The CDF for $X_s$ is

$$P[X_s \leq x] = P[ \text{ No Poisson points in } A ] = \exp \left( -\nu(A) \right), \quad \text{where} \quad A := \{ (\nu, \sigma) : \nu k(s; \sigma) > x \}. $$

If $\nu(A) < \infty$ then $X_s$ will be finite almost-surely; if $k(s; \sigma) = k(s - \sigma)$ depends only on the difference $s - \sigma$, then $X_s$ will be stationary, and isotropic.
if \( k(s; \sigma) = k(\|s - \sigma\|) \) depends only on the Euclidean distance. We use the same notation \( k(\cdot) \) for any of these three functions, distinguishing them by their argument(s). Below we will explore the marginal distribution of \( X_s \) and the joint distribution of \( \{X_{s_i}\} \) for a finite set \( S = \{s_i\} \subset \mathcal{S} \).

### 2.1 Dependence

First consider the value of \( X_s \) in one-dimensional setting, with \( \mathcal{S} = [0, 1] \) the unit interval. Figure (1) illustrates that the value of \( X_s \) is, by definition, the maximum value of \( v_jk(s - \sigma) \). But, \( X_s \) is also the smallest value \( y \) such that

\[
\text{Figure 1: } X_s \text{ is the maximum value of } v_jk(s - \sigma) \]

some \((\sigma_j, v_j)\) lies in \( A := \{(\sigma, v) : v > y/k(s - \sigma)\} \), as shown in Figure (2). It follows that the probability \( P[X_s \leq y] \) is the probability \( \exp(-\nu(A)) \) that \( \mathcal{N}(A) = 0 \).

In this example \( \nu_n(\{v\}) = \nu^{-2}1_{[v>0]} dv \) and \( \int_\mathcal{S} k(s) ds = 1 \). It follows that each \( X_s \) has the unit Fréchet distribution with CDF \( P[X_s \leq y] = \exp(-1/y) \),

\[
2
\]
Figure 2: \(X_s\) is also the smallest value \(y\) such that some \((\sigma_j, v_j)\) lies in \(\{(\sigma, v) : v > y/k(s - \sigma)\}\)

a distribution with infinite mean and variance. The values \(X_{s_1}\) and \(X_{s_2}\) of \(X_s\) at two different locations won’t be independent, but the concepts of “correlation” or “covariance” aren’t adequate to describe their dependence since they’re not in \(L_2\). If \(\text{dist}(s_1, s_2)\) is large then they will be nearly independent, however, while if \(\text{dist}(s_1, s_2)\) is small they won’t be. To see this geometrically, compare Figures (3) and (4). These are different realizations of the same random field, evaluated at \(s_1 = 0.5\) and \(s_2 = 0.75\). In Figure (3) the mass points \(\{(\sigma_j, v_j)\}, \{(\sigma_{j'}, v_{j'})\}\) at which the maxima are attained differ, the usual case when \(s_1, s_2\) are distant; in Figure (4), the maxima are both attained at a single point \((\sigma_j, v_j)\), the usual case when \(s_1, s_2\) are close. The point is illustrated in Figures (5) and (6) in a way that offers an avenue to computing the joint CDF \(P[X_{s_1} \leq y_1, X_{s_2} \leq y_2]\).
Figure 3: The maximum values of $v_j k(s_1 - \sigma_j)$ and $v_{j'} k(s_2 - \sigma_{j'})$ with different extremal $(\sigma_j, v_j)$, $(\sigma_{j'}, v_{j'})$

Figure 4: The maximum values of $v_j k(s_1 - \sigma_j)$ and $v_j k(s_2 - \sigma_j)$ with the same extremal $(\sigma_j, v_j)$
Figure 5: The maximum values of $v_j k(s_1 - \sigma_j)$ and $v_j' k(s_2 - \sigma_j')$ with different extremal $(\sigma_j, v_j)$, $(\sigma_j', v_j')$

Figure 6: The maximum values of $k(s_1 - \sigma_j)$ and $k(s_2 - \sigma_j)$ with the same extremal $(\sigma_j, v_j)$
3 Joint Distribution at Two Points

Let \( k(s, \sigma) \) be isotropic and monotone decreasing in the Euclidean distance \(|s - \sigma|\). We will illustrate in \( d = 1 \) dimension using the specific choice of \( k(s, \sigma) = \exp(-\lambda(s - \sigma)^2) \), but the same approach works more generally. For any two locations \( s_1, s_2 \) and levels \( y_1, y_2 \) the joint CDF

\[
F(y_1, y_2) = P[X_{s_1} \leq y_1, X_{s_2} \leq y_2]
\]

can be evaluated as \( \exp(-\nu(A)) \) for the union \( A \) of the sets \( B_1 := \{(v, \sigma) : vk(s_1, \sigma) > y_1 \} \) or, equivalently, of the disjoint sets \( A_1 := B_1 \cap \{\sigma < \tilde{\sigma}\} \) and \( A_2 := B_2 \cap \{\sigma > \tilde{\sigma}\} \) to the left and right of the red line in Figure (7) at \( \sigma = \tilde{\sigma} \), respectively, where

\[
\tilde{\sigma} = (s_1 + s_2)/2 + \log(y_2/y_1)/\lambda(s_1 - s_2).
\]

![Figure 7: The joint CDF of \( X_s \) at \( s_1 \) and \( s_2 \), evaluated at \( y_1 \) and \( y_2 \), is \( \exp(-\nu(A)) \) for the set \( A \) of points \((v, \sigma)\) above the two black curves. We evaluate \( \nu(A) \) as the sum \( \nu(A_1) + \nu(A_2) \) of the portions to the left and right of the red line at \( \sigma = \tilde{\sigma} \), where \( y_1/k(s_1, \tilde{\sigma}) = y_2/k(s_2, \tilde{\sigma}) \).](image-url)
In one dimension with \( s_1 < s_2 \), these measures are

\[
\nu(A_1) = \int_{-\infty}^\sigma \int_{0}^\infty \gamma \alpha v^{-\alpha-1} \, dv \, d\sigma
\]

\[
= \gamma y_1^{-\alpha} \int_{-\infty}^\sigma k_\alpha(\sigma - s_1) \, d\sigma
\]

\[
= \gamma y_1^{-\alpha} \int_{-\infty}^\sigma \exp(-\lambda \alpha (\sigma - s_1)^2) \, d\sigma
\]

\[
= \frac{\gamma \sqrt{2\pi}}{\sqrt{\lambda \alpha}} y_1^{-\alpha}\Phi(\sqrt{\frac{\lambda \alpha}{2}} + \log(y_2/y_1)/\lambda s_1)
\]

\[
\nu(A_2) = \frac{\gamma \sqrt{2\pi}}{\sqrt{\lambda \alpha}} y_2^{-\alpha}\Phi(\sqrt{\frac{\lambda \alpha}{2}} + \log(y_1/y_2)/\lambda s_2)
\]

leading to closed-form expressions for the joint CDF \( F(y_1, y_2) = \exp \left( - \nu(A_1) - \nu(A_2) \right) \) and hence the joint pdf (and likelihood function) \( f(y_1, y_2) \).

### 4 Convex Regions and Tessellations

Let \( C \subset S \) be a convex polygon and define

\[
X_C := \sup_{s \in C} X_s
\]

\[
= \sup_{j \in \mathbb{N}, s \in C} v_j k(s; \sigma_j);
\]  \hspace{1cm} (1)

we will also address the distribution of \( X_C \) and the joint distribution of \( \{X_{C_j}\} \) for disjoint collections of convex sets, like a rectangular grid or hexagonal tiling or a triangulation of a spatial area.

#### 4.1 Example: Fréchet/Square Exponential

Fix \( S = \mathbb{R}^2 \) and numbers \( \alpha > 0, \gamma > 0, \) and \( \lambda > 0 \) and consider the specific example

\[
k(s; \sigma) := e^{-\lambda |s-\sigma|^2/2}
\]

where \( |s-\sigma| \) denotes the Euclidean distance separating \( s \) and \( \sigma \) (the so-called square exponential kernel), and

\[
\nu(dv \, d\sigma) := \gamma \alpha v^{\alpha-1} 1_{\{v>0\}} \, dv \, d^2\sigma
\]
where $d^2\sigma$ denotes Lebesgue measure in $\mathcal{S}$. For $0 < \alpha < 2$, this is the Lévy measure for the stationary $\alpha$-Stable random field. For this kernel and measure,

$$
\nu(A) = \nu \left\{ (v, \sigma) : v > x \exp(\lambda|s - \sigma|^2/2) \right\} \\
= \int_{\mathcal{S}} \gamma \left[ x \exp(\lambda|s - \sigma|^2/2) \right]^{-\alpha} d^2\sigma \\
= x^{-\alpha}2\pi\gamma\int_0^\infty e^{-\alpha\lambda r^2/2} r dr \\
= x^{-\alpha}2\pi\gamma/\alpha\lambda, \quad \text{so}
$$

$$
\Pr[X_s \leq x] = \exp\left(-x^{-\alpha}2\pi\gamma/\alpha\lambda\right), \quad x > 0
$$

$$
f(x | \alpha, \gamma, \lambda) = (2\pi\gamma/\lambda) \exp\left(-x^{-\alpha}2\pi\gamma/\alpha\lambda\right) x^{-\alpha-1}1_{\{x > 0\}}
$$

and $X_s$ has a Fréchet $\text{Fr}(\alpha, 2\pi\gamma/\alpha\lambda)$ distribution. For $|s_i - s_j| \gg (\gamma/\alpha\lambda)^{1/\alpha}$ the random variables $X_{s_i}$ will be nearly independent, while for $|s_i - s_j| \ll (\gamma/\alpha\lambda)^{1/\alpha}$ they will nearly coincide. Their $p$th moments are infinite for $p \geq \alpha$, so in the interesting cases of $0 < \alpha < 2$ the variance is infinite and covariance undefined (and for $\alpha \leq 1$ even the mean is infinite).

### 4.2 Fréchet on Convex Sets

Here we generalize (3) from a single point $s \in \mathcal{S}$ to a convex set $C \subset \mathcal{S}$ and from the squared exponential to a wider class of kernel functions.

**Proposition 1.** Let $\alpha > 0$ and let $k(s; \sigma) = k(|s - \sigma|)$ be an isotropic kernel with finite monotonically-decreasing nonnegative distance function $k(r)$. Set

$$
c_0 := k^\alpha(0) \quad c_1 := \int_0^\infty k^\alpha(r) dr \quad c_2 := \int_0^\infty k^\alpha(r) 2\pi r dr.
$$

If $k \in L_\alpha(\mathbb{R}^2)$ and if $k$ is bounded then each $c_j < \infty$ and for any convex set $C \subset \mathcal{S}$,

$$
X^*_C := \sup_{s \in C} X_s
$$

has a Fréchet distribution $X^*_C \sim \text{Fr}(\alpha, c)$, i.e.,

$$
\Pr[X^*_C \leq x] = e^{-cx^{-\alpha}}, \quad x > 0
$$
with rate

\[ c = \gamma [c_2 + c_1 P + c_0 A] \]

for the perimeter \( P \) and area \( A \) of \( C \). Conversely, if \( k \notin L_\alpha(\mathbb{R}^2) \) or if \( k \) is unbounded, then \( X^*_C = \infty \) almost-surely.

**Proof.** First consider a rectangular region \( C \); without loss of generality we may rotate and translate so that the lower-left corner is the origin, and write \( C = [0, a] \times [0, b] \).

Now from (1) we find

\[
P[X^*_C \leq x] = \exp(-\nu(A)), \text{ where} \]

\[ A = \left\{ (v, \sigma) : v > x/k(|s - \sigma|) \text{ for any } s \in C \right\} \]

\[
\nu(A) = \int_S \gamma \left[ x/k(\text{dist}(\sigma, C)) \right]^{-\alpha} d^2\sigma
\]

\[ = x^{-\alpha} \gamma \int_S k^{\alpha}(\text{dist}(\sigma, C)) \, d^2\sigma
\]

\[ = x^{-\alpha} \gamma [c_2 + c_1 P + c_0 A] \quad (5)
\]

where \( P = 2(a + b) \) is the perimeter of \( C \) and \( A = a \cdot b \) the area. To see this, divide the integral over \( S = \mathbb{R}^2 \) into the nine regions determined by restricting \( \sigma_1 \) to \((\infty, 0), [0, a], (a, \infty)\) and by restricting \( \sigma_2 \) to \((-\infty, 0), [0, b], (b, \infty)\). The four “corner” pieces (with both \( \sigma_1 \) and \( \sigma_2 \) unbounded) sum to the first term in (5); the four semi-bounded strips sum to the second term; and the one bounded term (\( C \) itself) the last term.

The same argument shows that (5) holds for any convex polygon with \( n \geq 3 \) sides; now partition \( S \) into \( 2n + 1 \) regions by drawing two half-lines from each corner orthogonal to the edges that meet there to form \( n \) wedges (whose total contribution is the first term), \( n \) semi-infinite strips (whose total contribution is the second), and \( C \) itself (the third term). Finally, any convex set in \( \mathbb{R}^2 \) can be written as the increasing union of a sequence of convex polygons whose perimeters and areas converge.

\[ \square \]

Equation (5) also holds for the limiting cases of line segments (where \( A = 0 \) and \( P \) is twice the segment length) and points (where \( A = P = 0 \)), where for the squared-exponential kernel it reduces to (2).
For any $\alpha > 0$, the moments $\{c_j\}$ for the squared-exponential kernel $k(s; \sigma) := \exp(-\lambda(s - \sigma)^2)$ are all finite:

\[
\begin{align*}
  c_0 := & \int k(0) dr = e^{-0} = 1 \\
  c_1 := & \int_0^\infty k(r) dr = \int_0^\infty e^{-\alpha \lambda r^2/2} dr = \frac{\pi}{2\alpha \lambda} \\
  c_2 := & \int_0^\infty k(r) r^2 dr = \int_0^\infty e^{-\alpha \lambda r^2/2} 2\pi r dr = \frac{2\pi}{\alpha \lambda}
\end{align*}
\]

and hence (5) holds for all $\alpha > 0$.

Note this gives a way to predict maxima for regions if we can infer the values of $\alpha$, $\gamma$, and whatever parameters determine $k(s; \sigma)$ from observations at points— for example, we might hope to learn the parameters governing precipitation from data at weather stations, then predict extremes over counties or watersheds.

5 Perfect Simulation

It is possible to simulate $\{X_{s_i}\}_{i \in I}$ perfectly for any specified parameters $\alpha$, $\gamma$, $\lambda$ and any collection $\{s_i\}$ of locations (“monitoring stations”) in $\mathbb{R}^2$, using a variation of the Inverse Lévy Measure algorithm of Wolpert and Ickstadt (1998a,b). Here’s how. (Note: recent work by Wang and Stoev (2010) may enable conditional sampling given observations, and hence posterior sampling).

Begin with a ball of some radius $R > 0$ containing the set $\{s_i\}_{i \in I}$. Let $\mathcal{R}$ be a ball with the same center but larger radius $(R + \Delta)$ for some number $\Delta > 0$ (we’ll choose $\Delta$ below— something like $3/\sqrt{\alpha \lambda}$ or so should work). Draw points $\sigma_j$ iid $\text{Un}(\mathcal{R})$, and draw the event times $\tau_j$ from a standard unit-rate Poisson process (i.e., partial sums of independent standard exponential random variables). Set

\[
v_j := \left[ \frac{\gamma \pi (R + \Delta)^2}{\tau_j} \right]^{1/\alpha}.
\]

Then $\{(v_j, \sigma_j)\}$ are distributed exactly as the support of $\mathcal{N}$ restricted to $\mathbb{R}_+ \times \mathcal{R}$, sorted in order of decreasing $v_j$. With each successive $j$, reset each

\[
x_i = \sup_{k \leq j} \left\{ v_k e^{-\lambda |s_i - \sigma_k|^2/2} \right\}
\]
until finally $v_j < \underline{x}_J := \min \{x_i\}_{i \in I}$, whereupon stop.

Figure 8: Simulation of $X_s$ at $50 \times 50$ grid of points $\{s_i\}$. Blue disks represent support points, with area proportional to magnitude. $J = 484$ draws $(v_j, \sigma_j)$ were required to reach $v_J = 1.1687 < \underline{x}_J = 1.1690$, determining 68 support points each supporting from one to 305 points $x_i$.

It is possible that $\mathcal{R}$ wasn’t large enough— that if we could have drawn points from all of $\mathbb{R}^2$, some very-distant point with huge mass would have increased the value of some $x_i$. That event is for some support point of $\mathcal{N}$
to lie in the set

$$A = \left\{ (v, \sigma) : v > \inf_{i \in I} x_i e^{\lambda|\sigma - \sigma_i|^2}, \quad \sigma \notin \mathbb{R} \right\},$$

whose probability is $1 - e^{-\nu(A)}$. This is bounded above by $1 - e^{-\nu(A^*)}$ for the set

$$A^* := \left\{ (v, \sigma) : v > \mathbb{E}_I e^{\lambda|\sigma - R|^2}, |\sigma| \geq R + \Delta \right\}$$

since $A \subset A^*$ (after translating), with

$$\nu(A^*) = \int_{\mathbb{R}^2} \gamma \left[ \mathbb{E}_I e^{\lambda|\sigma - R|^2/2} \right]^{-\alpha} d\sigma$$

$$= 2\pi \gamma \int_{R+\Delta}^{\infty} \exp \left\{ -\alpha \lambda (r - R)^2/2 \right\} r \, dr \, (\mathbb{E}_I)^{-\alpha}$$

$$= \frac{\gamma \sqrt{2\pi}}{\alpha \lambda} \left\{ \phi \left( \Delta \sqrt{\alpha \lambda} \right) + R \sqrt{\alpha \lambda} \Phi \left( -\Delta \sqrt{\alpha \lambda} \right) \right\} (\mathbb{E}_I)^{-\alpha},$$

where $\phi(z)$ and $\Phi(z)$ are the pdf and CDF for the standard Normal distribution.

For sufficiently large $\Delta$ the probability $1 - e^{-\nu(A^*)} \leq \nu(A^*)$ will be small enough that $\{x_i\}_{i \in I}$ is acceptably close to (and identical to with high probability) a sample draw from the $\{X_{s_i}\}$ (in examples it’s easy to attain $\nu(A^*) < 10^{-10}$). For perfect sampling with no approximation error, simply draw (probably zero) a Poisson Po($\nu(A^*)$) number of points from $A^*$ (easy to do, since it’s a radial set); remove any points that are outside $A$; and, if any support points remains use them to update $\{x_i\}$. Figure (8) shows an example of a simulation, with a $50 \times 50$ regular array of points $\{s_i\}$, using $\alpha = \lambda = \gamma = 1$ on the ball of radius $R = 10$ centered at the origin. For this example $\Delta = 4$ and the error bound was $\nu(A^*) = 0.001347$.

6 Inference

OK, here’s a speculative but interesting (to me) idea. Suppose we know $\alpha$, $\lambda$, and $\gamma$. 

12
6.1 \( X_s \) Observed at One Point

Suppose we are able to observe

\[ X_s := \sup_j \{ v_j k(s; \sigma_j) \} \]

exactly without measurement error. Then the conditional distribution of \( x = X_s \) given the location \( \sigma_j = \sigma \) of the supporting point \((v_j, \sigma_j)\) (the one attaining the supremum above) is:

\[
P[X_s \leq x \mid \sigma] = \exp \left\{ -\gamma x^{-\alpha} e^{-\lambda|s-\sigma|^2/2} \right\}
\]

so the conditional pdf for \( X_s \) given the location is

\[
f(x \mid \sigma) = \alpha \gamma x^{-\alpha - 1} \exp \left\{ -\alpha \lambda|s - \sigma|^2/2 - \gamma x^{-\alpha} e^{-\lambda|s-\sigma|^2/2} \right\}.
\]

By Bayes’ rule,

\[
f(\sigma \mid X_s = x) = c \exp \left\{ -\alpha \lambda|s - \sigma|^2/2 - \gamma x^{-\alpha} e^{-\lambda|s-\sigma|^2/2} \right\},
\]

\[
c = \frac{\alpha \gamma \lambda x^{-\alpha}}{2\pi [1 - e^{-\gamma x^{-\alpha}}]}
\]

and of course \( v = x \exp(\lambda|s - \sigma|^2/2) \) is determined by \( x \) and \( \sigma \) because, geometrically, the support point for \( X_s \) lies on the bowl-shaped set

\[
M_1 := \left\{ (v, \sigma) : v = x \exp(\lambda|s-\sigma|^2/2) \right\}.
\]

For large \( x \) we have \( \sigma \approx \mathcal{N}(s, \frac{1}{\alpha} f) \), while for small \( x \) we have the same asymptotic behavior as \( |\sigma - s| \rightarrow \infty \) but the density falls off to zero as \( \sigma \rightarrow s \). The marginal density for \( X_s \) (and hence the likelihood for \( \alpha, \gamma, \lambda \)) is given in (3); together that and (6) determine the joint distribution for \( \alpha, \gamma, \lambda, \sigma, v \) and \( x \) for any specified prior distribution.

6.2 Two Points

Suppose we know \( \alpha, \lambda, \) and \( \gamma \) and observe

\[ X_{s1} := \sup_j \{ v_j k(s_1; \sigma_j) \} = x_1 \quad \text{and} \quad X_{s2} := \sup_j \{ v_j k(s_2; \sigma_j) \} = x_2. \]
Then two cases arise—either both suprema are attained at the same \( j \in \mathbb{N} \) (i.e., at the same element \((u, \sigma)\) in the support of \(N\), or not. Any dependence between \(X_{s_1}\) and \(X_{s_2}\) arises from the possibility of both maxima arising from the same support point—a point necessarily on the manifold

\[
\mathcal{M}_{12} := \mathcal{M}_1 \cap \mathcal{M}_2 = \{(u, \sigma) : u = x_1/k(s_1; \sigma) = x_2/k(s_2; \sigma)\} \subset \mathbb{R}_+ \times S.
\]

For example, with \(k(s; \sigma) = \exp \left(-\lambda|s - \sigma|^2\right)\) and \(s_1 = (0, 0), \ s_2 = (d, 0)\),

\[
\mathcal{M}_{12} := \left\{(u, \sigma) : \sigma_x = \frac{d}{2} + \frac{1}{2\lambda d} \log \frac{x_2}{x_1}, \ \sigma_y \in \mathbb{R}, \ u = x_1 e^{\lambda|s_1 - \sigma|^2/2}\right\}, \quad (7)
\]

whose projection \(\mathcal{M}_{12}^S\) onto \(S\) is a line perpendicular to the segment connecting \(s_1\) and \(s_2\), offset from the midpoint toward the point with the larger value of \(X_s\) by a distance that shrinks as \(\lambda\) or \(d := |s_1 - s_2|\) grows.

If \(X_{s_1}\) and \(X_{s_2}\) do not share a common support point, then each must have its own point in the appropriate component of the set \(\mathcal{M}_{1\land 2} := \mathcal{M} \setminus \mathcal{M}_{12}\), where

\[
\mathcal{M} := \left\{(u, \sigma) : u = \min \left\{x_1 e^{\lambda|s_1 - \sigma|^2}, \ x_2 e^{\lambda|s_2 - \sigma|^2}\right\}\right\}.
\]

Notice the projection \(\mathcal{M}_{1\land 2}^S\) of \(\mathcal{M}_{1\land 2}\) onto \(S\) partitions \(S\) into two parts, separated by line \(\mathcal{M}_{12}^S\) described below (7).

SO— a posteriori, the point(s) \((u, \sigma)\) leading to the observed \(x_1, x_2\) are either a pair of points, one in each component of \(\mathcal{M}_{1\land 2}\), or a single point in \(\mathcal{M}_{12}\). The likelihood might be available for each of these; whether it is or not, it may be possible to draw MCMC samples of the support point(s) and parameters. I don’t know yet how to find, e.g., the probability of one support point vs. two, or the distribution of the latent support points, or the likelihood function.

One idea: Use (6) to get cond’l dist’n of \(\sigma, u\) for the support point for \(X_{s_1}\),... then for each \(u\) and \(\sigma\), find the probability that this is ALSO the support point for \(X_{s_2}\) as \(e^{-u(B)}\) where \(B\) is the set of \(\{(u, \sigma)\}\) that are not in the (already known to be empty) bowl above \(s_1\), but are in the cone above \(s_2\).

### 6.3 Three Points

Now suppose we observe \(x_i = X_{s_i}\) at three locations \(s_1, s_2, s_3\). Now there are three possibilities—either each supremum \(X_s\) arises from a different point \((u, \sigma) \in \mathbb{R}_+ \times S\) in the support of \(N\); or one of the three pairs \(s_i, s_j\)
share a single support point in one of three manifolds $M_{ij}$ similar to that of (7); or all three points share the same support point, necessarily at the intersection $M_{123} = \cap M_i$ of the three manifolds.

Figure 9: Example of $\{M_i\}$ for $n = 3$ points $s_i$. Points $s_i$ are indicated by small circles, connected by line segments to a label $i$ at the center of $M_i$.

Geometrically, the three lines $M_{12}^S$, $M_{13}^S$, and $M_{23}^S$ all intersect in a single point $M_{123}^S$. Each pair $M_{ij}^S$, $M_{ik}^S$ bounds a wedge-shaped region, the intersection of two half-spaces, which often (but not always, if $X_{s_j}$ is far from $X_{s_j}$ or $X_{s_k}$) contains the associated point $s_i$. Those three wedges partition $S$. The posterior distribution for the support of $\{X_{s_i}\}$ is concentrated on three sets: either each open wedge contains a single point $(v_i, \sigma_i)$; or one of the three open wedges contains such a point supporting one of the three $X_{s_i}$, and the half-line bounding the other two wedges contains a point supporting the other two; or all three points are supported by a single point at $\sigma = M_{123}^S$.

Again it’s possible that the likelihood function would be available. And, again, even if the likelihood is unavailable it may be possible to draw MCMC samples; at worst, we could introduce a measurement-error model and draw
MCMC samples of all the parameters and the support points \( \{(v_j, \sigma_j)\} \).

### 6.4 Four Points

Something interesting happens beginning with \( n = 4 \) points \( \{x_i\} \): different topological possibilities arise. Two are shown in Figure (10)—one with a bounded component, and one without. A third possibility is for four lines \( M_{ij} \) to intersect in a single point \( M_{ijkl} \); in this latter case the likelihood will be concentrated on a single support point for all four \( X_{s_i} \).

![Spex Tessellation Demo](image1)

![Spex Tessellation Demo](image2)

Figure 10: Two possible patterns for \( n = 4 \) points \( x_i \), one with a bounded component and one without.

### 6.5 \( n \) Points

The same approach can be explored for \( n \) points \( \{s_i\} \), perhaps with RJMCMC, but the combinatorial problems in evaluating a likelihood will become unwieldy quickly as \( n \) grows. Given the observed values \( x_i = X_{s_i} \) the support point(s) must lie on the lower boundary of a collection of bowls, the manifold

\[
M := \left\{(v, \sigma) : \ v = \min_i \left\{ x_i e^{\lambda|\sigma - s_i|^2/2} \right\} \right\}.
\]

Each \( s_i \) lies in a (possibly unbounded) polygonal region formed by the intersection of all the half-spaces \( M_{ij}^\delta \). Those polygons and their boundaries
Figure 11: Example of $\{M_i\}$ for $n = 10$ points $s_i$.

are the only places where support points can possibly lie. Any polygon corner where four or more edges intersect must be a support point for all the associated $\{s_i\}$ whose polygons meet at that corner. Almost surely, no polygon will have more than one such corner. The combinatorial challenge comes from adjacent polygons without such a corner—which might and might not share a support point, necessarily on the edge they share.

If we don’t know $\alpha$, $\lambda$, and $\gamma$ then the likelihood surface will be even more complicated for $n \geq 3$, since the topology of the possible support sets will vary with the values of $\alpha$ and $\lambda$, but a measurement-error approach (using Metropolis/Hastings MCMC) might work.

7 Extensions

As yet undeveloped but (to me) promising ideas include:

1. Let $S$ be $\mathbb{R}^F$ instead of $\mathbb{R}^2$;

2. Replace space $S$ with space-time $S \times \mathcal{T}$;
3. Replace $\lambda$ with a positive-definite matrix $\Lambda$, so

$$ k(s; \sigma) := e^{-(s-\sigma)^T \Lambda (s-\sigma)/2} $$

4. Add “attributes” to points $\sigma \in S$—altitude, aspect, etc. Even $\lambda$ (or $\Lambda$) could be attributes, hence locus-specific;

5. Replace Fréchet marginal distribution with Gumbel, by replacing density $\nu_u(u) = \gamma u^{-\alpha-1}$ on $\mathbb{R}_+$ with $\nu_u(u) = \gamma e^{-\alpha u}$ on all of $\mathbb{R}$.

6. Replace points $\sigma \in S = \mathbb{R}^p$ with translated subspaces $\ell \in S = \mathbb{R}^{p-k} \times g_{p,k}$, where $g_{p,k}$ is the Grassmann manifold of $k$-dimensional subspaces of $\mathbb{R}^p$ (e.g., lines through the origin if $k = 1$); we build a Poisson field on $\mathbb{R}_+ \times \mathbb{R}^{p-k} \times g_{p,k}$ and construct $X_s$ from it.

8 End Notes

- If $k(s; \sigma) := \exp(-\lambda(s-\sigma)^2/2)$ is replaced with the $t$ density $k(s; \sigma) := (1 + (s-\sigma)^2/\nu)^{-\nu(p+\nu)/2}$ in $S = \mathbb{R}^p$ rescaled to have maximum value one, (2) changes only slightly to

$$ \nu(A) = x^{-\alpha} \gamma(\pi \nu)^{p/2} \Gamma\left(\frac{\alpha(p+\nu)-p}{2}\right) / \Gamma(\alpha(p+\nu)),$$

so again $X_s$ has an $\alpha$-Fréchet distribution with intensity available in closed form, and each $c_j$ of (4) is available explicitly.

References


