Some applications of an implicit formula for the maximum of a Gaussian random field

Joint work with Mario Wschebor
Bannf febr. 23 2009

Jean-Marc Azaïs

Institut de Mathématiques, Université de Toulouse
1. The distribution of the maximum
   - The implicit formula

2. The regularity of the density

3. Non-asymptotic bounds

4. Second order study
Signal + noise model

Spatial Statistics often uses “signal + noise model”, for example:
Representation of the yield per unit by GPS harvester.

Is there only noise or some region with higher fertility? A good statistics is the maximum (of the absolute value)
Gaussian sea modeling
General results

Let $X(t)$ a real-valued (often Gaussian) random field and

$$F_M(x) = P(M_T \leq x)$$

the distribution function of its maximum.

The computation of $F_M(x)$ by means of a closed formula is known only in a very restricted number of cases: Brownian (Bridge), Ornstein-Ulhenbeck etc...
General inequalities Borell, Sudakov, Tsirelson are fundamental for the mathematical theory but numerically weak.

An example in the simplest case The Brownian motion where some parameters are known

<table>
<thead>
<tr>
<th>$u$</th>
<th>true values of $\mathbb{P}(M_W &gt; u)$</th>
<th>Borell’s b. mean</th>
<th>Borell’s b. median</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.045</td>
<td>0.4855</td>
<td>0.2077</td>
</tr>
<tr>
<td>3</td>
<td>0.0027</td>
<td>0.0885</td>
<td>0.0347</td>
</tr>
<tr>
<td>4</td>
<td>6.33 $10^{-5}$</td>
<td>5.93 $10^{-3}$</td>
<td>1.98 $10^{-3}$</td>
</tr>
<tr>
<td>5</td>
<td>5.73 $10^{-7}$</td>
<td>1.46 $10^{-4}$</td>
<td>4.32 $10^{-5}$</td>
</tr>
</tbody>
</table>
The implicit formula

Consider a realization with $M > u$, then necessarily there exists a local maximum or a border maximum above $u$.

Border maximum: maximum in relative topology that is located on the border. Can be local or global.
We consider parameter sets that are union of manifolds of dimension 1 to $d$ + additional conditions
In fact results are simpler (and stronger) in term of the density $p_M(x)$ of the maximum. Bounds for the distribution are obtained by integration.

**Theorem**

Let $M = \max_{t \in S} X(t)$. Under assumptions above, the distribution of $M$ has the density

$$
p_M(x) = \sum_{t \in S_0} E\left( \mathbf{I}_{A_x} \big| X(t) = x \right) p_{X(t)}(x)
$$

$$
+ \sum_{j=1}^d \int_{S_j} E\left( |\det(X_j''(t))| \mathbf{I}_{A_x} \big| X(t) = x, X_j'(t) = 0 \right) p_{X(t),X_j'(t)}(x, 0) \sigma_j(dt),
$$

(1)

where $A_x = \{M \leq x\}$.
What can we do with this implicit formula?

The formula (1) is only implicit: $M$ appears unfortunately on both sides.

Moreover terms like the expectation of the modulus of the determinant is hard to compute.
1. The distribution of the maximum
   - The implicit formula

2. The regularity of the density

3. Non-asymptotic bounds

4. Second order study
Dimension 1

When the parameter $t$ is a scalar variable that varies in $[0, 1]$, the formula (1) reads

**Theorem**

Suppose that $X$ is a Gaussian process with $C^2$ paths and such for all $s, t, s \neq t \in [0, 1]$, $X(s), X(t), X'(t)$ and $X(t), X'(t), X''(t)$ admit a joint density. Then $M$ has a continuous density $p_M$ given for every $u$ by

$$p_M(u) = \mathbb{P}(M \leq u | X(0) = u)p_{X(0)}(u) + \mathbb{P}(M \leq u | X(1) = u)p_{X(1)}(u)$$

$$+ \int_0^1 \mathbb{E}(X''- (t) \mathbb{I}_{M \leq u} | X(t) = u, X'(t) = 0)p_{X(t), X'(t)}(u, 0)dt \quad (2)$$
Using **induction**

**Theorem**

Assume that $X(t)$ has $C^{2k}$ sample paths and satisfies a non-degeneracy condition. Then, $F_M$ is of class $C^k$ and its successive derivatives can be computed by induction using an extension of the preceding theorem.

This goes far beyond the general result given by Tsirelson’s Theorem that who proved a general theorem (1975) on the density $P_M$ for general processes.
Proof

Differentiating, the “Bad Guy” is the term

\[ \mathbb{P}\{M \leq u \mid X(t) = u, X'(t) = 0\} \]

We use Regression and Desingularization arguments

\[ X(s) = b^t(s)X(t) + c^t(s)X'(t) + \frac{(s - t)^2}{2}X^t(s) \quad s \in [0, 1] \quad s \neq t. \]

Under our hypotheses \( X^t(s) \) is a “nice” process that admits a differentiable extension at zero.
\[ \{ M \leq u \mid X(t) = u, X'(t) = 0 \} \]

can be translated as

\[ X^t(s) \leq b^t(s)u \quad \text{where } b^t(s) \text{ is some function} \]

Induction begins using a generalization of the first theorem.
The distribution of the maximum
Non-asymptotic bounds
Second order study

Some applications of an implicit formula for the maximum of a Gaussian random field
The regularity of the density

After Deformation

Barrier

Typical Path
In dimension greater than 1 the desingularization argument is not so nice and results are weaker but still stronger than those of Tsirolelson’s Theorem.
The implicit formula can be turned into a bound by replacing the condition \( \{ M \leq x \} \) by \( \{ X''_T(t) \text{ definite negative} \} \), and \( \{ X'_N(t) \text{ extended outwards} \} \)

\[
E \left[ | \det(X''(t)) | \mathbf{1}_{X''(t) \text{ d. negative}} \right] \leq \frac{1}{2} \left[ E \left[ | \det(X''(t)) | \right] + (-1)^d E \left[ \det(X''(t)) \right] \right]
\]

The term \( E \left[ \det(X''(t)) \right] \) appears in the computation of the expectation of the Euler Characteristic
Lemma

\textit{(Adler)}

\[ \mathbb{E}( \det(X''(t))/X(t) = x, X'(t) = 0) = \det(\text{Var}(X'(t)) ) H_d(x) \]

where \( H_d(x) \) is the \( d \)th Hermite polynomial and \( \Lambda := \text{Var}(X'(t)) \)
The computation of expectations of modulus of quadratic forms and determinants has received some attention in the recent years.

- By Fourier method: Berry and Dennis (2000)
- using Fourier and other methods Li and Wei (2009)

making it possible, in some cases, to do the computations.
The distribution of the maximum
The regularity of the density
Second order study

Some applications of an implicit formula for the maximum of a Gaussian random field

Non-asymptotic bounds

computation of $p_M$

\[
p_M(x) = \sum_{t \in S_0} \mathbb{E}(\mathbf{1}_{X(t) = x} p_X(t)(x))
\]

\[
+ \sum_{j=1}^d \int_{S_j} \mathbb{E}(\det(X_j''(t)) \mathbf{1}_{X_j(t) = x, X_j'(t) = 0} p_{X(t), X_j'(t)}(x, 0) \sigma_j(dt),
\]

A key point is the following
If $X$ is stationary and isotropic with covariance $\rho(||t - s||^2)$ normalized by $\text{Var}(X(t)) = 1$ and $\text{Var}(X'(t)) = \text{Id}$
Then under the condition $\{X(t) = x, X'(t) = 0\}$

\[
X''(t) = \sqrt{8\rho''} G + \xi \sqrt{\rho''} - \rho'^2 \text{Id} + x\text{Id}
\]

Where $G$ is a GOE matrix (Gaussian Orthogonal Ensemble), and $\xi$ a standard normal independent variable.
Theorem

Assume that the random field $\mathcal{X}$ is centered, Gaussian, stationary and isotropic and is “regular” Let $S$ have polyhedral shape (the faces are flat). Then,

$$p(x) = \varphi(x) \left\{ \sum_{t \in S_0} \hat{\sigma}_0(t) + \sum_{j=1}^{d_0} \left[ \left( \frac{|\rho'|}{\pi} \right)^{j/2} H_j(x) + R_j(x) \right] g_j \right\}$$

(3)

- $g_j = \int_{S_j} \hat{\theta}_j(t) \sigma_j(dt)$, $\hat{\theta}_j(t)$ is the normalized solid angle of the cone of the extended outward directions at $t$ in the normal space with the convention $\theta_d(t) = 1$.
  - For convex or other usual polyhedra $\hat{\theta}_j(t)$ is constant for $t \in S_j$,
- $H_j$ is the $j$th (probabilistic) Hermite polynomial.
Theorem

\[ R_j(x) = \left( \frac{2 \rho''}{\pi |\rho'|} \right)^{\frac{j}{2}} \frac{\Gamma(j+1/2)}{\pi} \int_{-\infty}^{+\infty} T_j(v) \exp\left(-\frac{v^2}{2}\right) \, dy \]

\[ v := -(2)^{-1/2}((1 - \gamma^2)^{1/2}y - \gamma x) \quad \text{with} \quad \gamma := |\rho'|(\rho'')^{-1/2}, \]

\[ T_j(v) := \left[ \sum_{k=0}^{j-1} \frac{H_k^2(v)}{2^k k!} \right] e^{-v^2/2} - \frac{H_j(v)}{2^j(j-1)!} I_{j-1}(v), \]

where

\[ I_n(v) = 2 e^{-v^2/2} \sum_{k=0}^{\left[\frac{n-1}{2}\right]} 2^k \frac{(n-1)!!}{(n-1-2k)!!} H_{n-1-2k}(v) \]

\[ + \mathbf{I}_{\{n \text{ even}\}} 2^{\frac{n}{2}} (n-1)!! \sqrt{2\pi} (1 - \Phi(x)) \]
1. The distribution of the maximum
   - The implicit formula

2. The regularity of the density

3. Non-asymptotic bounds

4. Second order study
Second order study

We go back to the implicit formula giving $P_M$ and we study the bound $\overline{P}_M$. Let us look, for simplicity, to the term of dimension $j = d$. When we remove $\mathbf{1}_{A_x}$ in

$$\mathbb{E}(\left| \det(X''(t)) \right| \mathbf{1}_{A_x} \left| X(t) = x, X_j'(t) = 0 \right.)$$

it is easy to show that the logarithmic behavior (as defined below) is the same as

$$\mathbb{P}(M \geq x \left| X(t) = x, X_j'(t) = 0 \right.)$$

and using Borel Sudakov Tsirelson inequality we get
Theorem

Under conditions above + \( \text{Var}(X(t)) \equiv 1 \) Then

\[
\lim_{x \to +\infty} - \frac{2}{x^2} \log [\hat{p}_M(x) - p_M(x)] \geq 1 + \inf_{t \in S} \frac{1}{\sigma_t^2 + \lambda(t) \kappa_t^2}
\]

\[
\sigma_t^2 := \sup_{s \in S \setminus \{t\}} \frac{\text{Var}(X(s)/X(t), X'(t))}{(1 - r(s, t))^2}
\]

and \( \kappa_t \) is some geometrical characteristic et \( \Lambda_t = \text{GEV}(\Lambda(t)) \)

The right hand side is finite and \( > 1 \)

\( \hat{p}_M \) stands for \( p_M^E, \hat{p}_M \) or the mean of the two quantities
Two examples of new results

Suppose that the process is stationary and isotropic with covariance \( \rho(\|t - s\|^2) \) with normalization \( \text{Var}(X(t)) = 1 \) \( \text{Var}(X'(t)) = Id \) and suppose that the parameter set is convex if the covariance \( \rho \) is monotone

\[
\lim_{x \to +\infty} - \frac{2}{x^2} \log [\bar{p}_M(x) - p_M(x)] = 1 + \frac{1}{12 \rho''(0) - 1}
\]

general case

\[
\lim_{x \to +\infty} - \frac{2}{x^2} \log [\hat{p}_M(x) - p_M(x)] \geq 1 + 1/Z_\Delta
\]

with

\[
Z_\Delta := \sup_{z \in (0, \Delta]} \frac{1 - \rho^2(z^2) - 4\rho''(z^2)z^2}{[1 - \rho(z^2)]^2} + \max_{z \in (0, \Delta]} \frac{4[\rho'(z^2) + z]^2}{[1 - \rho(z^2)]^2},
\]

and \( \Delta \) is the diameter of \( S \).
The distribution of the maximum
The regularity of the density
Non-asymptotic bounds
Second order study

Some applications of an implicit formula for the maximum of a Gaussian random field
Second order study
References

Level Sets and Extrema of Random Processes and Fields

Jean-Marc Azaïs and Mario Wschebor


Li W. and Wej A. Gaussian integrals involving absolute value functions. IMS Lecture NotesMonograph Series.