Rice method for the maximum of Gaussian fields

IWAP, July 8, 2010

Jean-Marc Azaïs

Institut de Mathématiques, Université de Toulouse

1. Examples
2. The record method $d = 2$ or 3
3. The maxima method
   - Second order

Signal + noise model

Spatial Statistics often uses " signal + noise model", for example :

- precision agriculture
- neuro-sciences
- sea-waves modelling

Signal + noise model

Spatial Statistics often uses " signal + noise model", for example :

- precision agriculture
- neuro-sciences
- sea-waves modelling

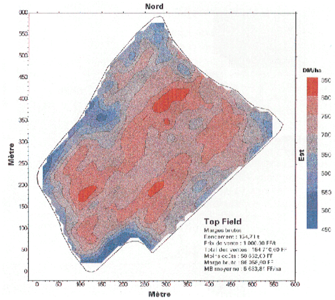
Signal + noise model

Spatial Statistics often uses “ signal + noise model”, for example :

- precision agriculture
- neuro-sciences
- sea-waves modelling

Precision agriculture

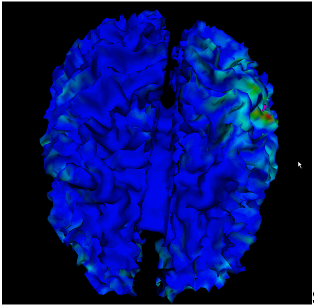
Representation of the yield per unit by GPS harvester .



Is there only noise or some region with higher fertility ? ?

Neuroscience

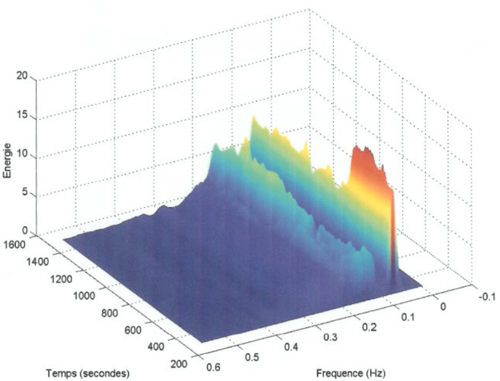
The activity of the brain is recorded under some particular action and the same question is asked



source : Maureen CLERC

Sea-waves spectrum

Locally in time and frequency the spectrum of waves is registered.
We want to localize transition periods.



In all these situations a good statistics consists in observing the **maximum** of the (absolute value) of the random field for deciding if it is **typically** (Noise) or **too large** (signal).

The Rice method

M is the maximum of a smooth random process $X(t), t \in \mathbb{R}^d$ $(d = 1)$ or field $(d > 1)$.
We want to evaluate

$$\mathbb{P}\{M > u\}$$

- In dimension 1 : count the number of up-crossing
- In larger dimension search for an other geometrical characteristic
  - number of connected component of the excursion set :open problem
  - Euler characteristic an alternative to the preceding : Conceptually complicated, computationally easy
  - Number of maxima above the considered level : difficult to compute the determinant but gives bounds
  - Particular points on the level set : the record method

The Rice method

M is the maximum of a smooth random process $X(t), t \in \mathbb{R}^d$ ($d = 1$) or field ($d > 1$) .
We want to evaluate

$$\mathbb{P}\{M > u\}$$

- In dimension 1 : count the number of up-crossing
- In larger dimension search for an other geometrical characteristic
  - number of connected component of the excursion set :open problem
  - Euler characteristic an alternative to the preceding : Conceptually complicated, computationally easy
  - Number of maxima above the considered level : difficult to compute the determinant but gives bounds
  - Particular points on the level set : the record method

The Rice method

M is the maximum of a smooth random process $X(t), t \in \mathbb{R}^d$ ($d = 1$) or field ($d > 1$) .
We want to evaluate

$$\mathbb{P}\{M > u\}$$

- In dimension 1 : count the number of up-crossing
- In larger dimension search for an other geometrical characteristic
  - number of connected component of the excursion set :open problem
  - Euler characteristic an alternative to the preceding : Conceptually complicated, computationally easy
  - Number of maxima above the considered level : difficult to compute the determinant but gives bounds
  - Particular points on the level set : the record method

The Rice method

M is the maximum of a smooth random process $X(t), t \in \mathbb{R}^d$ $(d = 1)$ or field $(d > 1)$.
We want to evaluate

$$\mathbb{P}\{M > u\}$$

- In dimension 1 : count the number of up-crossing
- In larger dimension search for an other geometrical characteristic
  - number of connected component of the excursion set :open problem
  - Euler characteristic an alternative to the preceding : Conceptually complicated, computationally easy
  - Number of maxima above the considered level : difficult to compute the determinant but gives bounds
  - Particular points on the level set : the record method

The Rice method

M is the maximum of a smooth random process $X(t), t \in \mathbb{R}^d$ $(d = 1)$ or field $(d > 1)$.
We want to evaluate

$$\mathbb{P}\{M > u\}$$

- In dimension 1 : count the number of up-crossing
- In larger dimension search for an other geometrical characteristic
  - number of connected component of the excursion set :open problem
  - Euler characteristic an alternative to the preceding : Conceptually complicated, computationally easy
  - Number of maxima above the considered level : difficult to compute the determinant but gives bounds
  - Particular points on the level set : the record method

The Rice method

M is the maximum of a smooth random process $X(t), t \in \mathbb{R}^d$ $(d = 1)$ or field $(d > 1)$.
We want to evaluate

$$\mathbb{P}\{M > u\}$$

- In dimension 1 : count the number of up-crossing
- In larger dimension search for an other geometrical characteristic
  - number of connected component of the excursion set :open problem
  - Euler characteristic an alternative to the preceding : Conceptually complicated, computationally easy
  - Number of maxima above the considered level : difficult to compute the determinant but gives bounds
  - Particular points on the level set : the record method

The maxima method

Let us forget boundaries

$$\{M > u\} = \text{ There is a maximum above } u =: \{M_u > 0\}$$

$$\mathbb{P}\{M > u\} \leq \mathrm{E}(M_u)$$
$$= \int_u^{+\infty} dx \int_S \mathrm{E}\big[|\det(X''(t))\ \mathbf{1}_{X''(t) \prec 0}|\ \big|X(t) = x, X'(t) = 0\big] P_{X(t),X'(t)}(x, 0) dt$$

Very difficult to compute the expectation of the determinant

1. Examples
2. The record method $d = 2$ or 3
3. The maxima method
   - Second order

Under some conditions, Roughly speaking the event

$$\{M > u\}$$

is almost equal to the events

"The level curve at level u is non-empty"

"The point at the southern extremity of the level curve exists"

"There exists a point on the level curve :
$X(t) = u; X_1'(t) = \frac{\partial X}{\partial t_1} = 0; X_2'(t) = \frac{\partial X}{\partial t_2} > 0$
$X_{11}''(t) = \frac{\partial^2 X}{\partial t_1^2} < 0$"

Forget the boundary

and define

$$Z(t) := \begin{pmatrix} X(t) \\ X_1'(t) \end{pmatrix}$$

The probability above is bounded by the expectation of the number of roots of $Z(t) - (u,0) \Rightarrow$ Rice formula

$$\mathbb{P}\{M > u\} \leq \text{Boundary terms}$$

$$+ \int_S \mathrm{E}(|\det(Z'(t)\ \mathbf{1}_{X_1''(t)<0}\ \mathbf{1}_{X_2'(t)>0}|\big|X(t) = u, X_1'(t) = 0) p_{X(t),X_1'(t)}(u,0)dt,$$

The difficulty lies in the computation of the expectation of the determinant

The key point is that under the condition $\{X(t) = u, X_1'(t)\} = 0$, the quantity

$$|\det(Z'(t)| = \begin{vmatrix} X_1' & X_2' \\ X_{11}'' & X_{12}'' \end{vmatrix}$$

is simply equal to $|X_2' X_{11}''|$. Taking into account conditions, we get the following expression for the second integral

$$+\int_S \mathrm{E}(|X_{11}''(t)^- X_2'(t)^+ \big| X(t) = u, X_1'(t) = 0) p_{X(t),X_1'(t)}(u,0)dt.$$

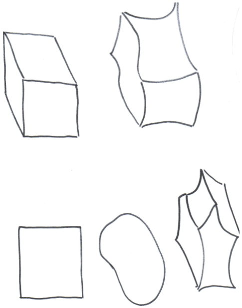
Moreover under stationarity or some more general hypotheses, these two random variables are independent.

Extension to dimension 3

Using Fourier method (Li and Wei 2009) we are able to compute the expectation of the absolute value of a determinant in dimension 2.
It is a simple quadratic form.
We are able to extend the result to dimension 3 (Pham 2010).

1. Examples
2. The record method $d = 2$ or 3
3. The maxima method
   - Second order

Consider a realization with $M > u$, then necessarily there exist a local maxima or a border maxima above U
Border maxima : local maxima in relative topology
If the consider sets that are union of manifolds of dimension 1 to d of this kind



In fact result are simpler (and stronger) in term of the density $p_M(x)$ of the maximum. Bound for the distribution are obtained by integration.

Theorem

$$p_M(x) \leq \widehat{p}_M(x) := \frac{1}{2}[\overline{p}_M(x) + p_M^{EC}(x)] with$$

$$\overline{p}_M(x) := \int_S \mathrm{E}\big(|\det(X''(t))|/X(t) = x, X'(t) = 0\big) p_{X(t),X'_j(t)}(x,0)dt + \textit{boundary term}$$

and

$$p_M^{EC}(x) := (-1)^d \int_S \mathrm{E}\big(\det(X''(t))/X(t) = x, X'(t) = 0\big) p_{X(t),X'_j(t)}(x,0)dt + \textit{boundary}$$

Quantity $p_M^{EC}(x)$ is easy to compute using the work by Adler and properties of symmetry of the order 4 tensor of variance of X'' (under the conditional distribution))

Lemma

$$\mathrm{E}\Big(\det(X''(t))/X(t) = x, X'(t) = 0\Big) = \det(\Lambda) H_d(x)$$

where $H_d(x)$ *is the* d*th Hermite polynomial and* $\Lambda := \mathrm{Var}(X'(t))$

main advantage of Euler characteristic method lies in this result.

computation of $\overline{p}_m$

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

$$X''(t) = \sqrt{8\rho"}G + \xi\sqrt{\rho" - \rho'^2}Id + xId$$

Where G is a GOE matrix (Gaussian Orthogonal Ensemble), and ξ a standard normal independent variable. We use recent result on the the characteristic polynomials of the GOE. Fyodorov(2004)

Theorem

Assume that the random field $\mathcal{X}$ is centered, Gaussian, stationary and isotrpic and is "regular" Let S have polyhedral shape. Then,

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

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

Theorem (continued)

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

$$v := -(2)^{-1/2}\left((1-\gamma^2)^{1/2}y - \gamma x\right) \text{ with } \gamma := |\rho'|(\rho'')^{-1/2}, \qquad (2)$$

$$T_j(v) := \left[\sum_{k=0}^{j-1} \frac{H_k^2(v)}{2^k k!}\right] e^{-v^2/2} - \frac{H_j(v)}{2^j (j-1)!} I_{j-1}(v), \qquad (3)$$

$$I_n(v) = 2e^{-v^2/2} \sum_{k=0}^{[\frac{n-1}{2}]} 2^k \frac{(n-1)!!}{(n-1-2k)!!} H_{n-1-2k}(v) \qquad (4)$$

$$+ \mathbf{1}_{\{n \text{ even}\}} 2^{\frac{n}{2}} (n-1)!! \sqrt{2\pi}(1-\Phi(x))$$

Second order study

Using an exact implicit formula

Theorem

Under conditions above + $\mathrm{Var}(X(t)) \equiv 1$ *Then*

$$\underline{\lim}_{x\to+\infty} \; - \frac{2}{x^2}\log\left[\widehat{p}_M(x) - p_M(x)\right] \geq 1 + \inf_{t\in S}\frac{1}{\sigma_t^2 + \overline{\lambda}(t)\kappa_t^2}$$

$$\sigma_t^2 := \sup_{s\in S\setminus\{t\}} \frac{\mathrm{Var}\big(X(s)\big|X(t), X'(t)\big)}{(1-r(s,t))^2}$$

and κ_t *is some geometrical characteristic et* $\Lambda_t = GEV(\Lambda(t))$

The right hand side is finite and > 1

Adler R.J. and Taylor J. E. *Random fields and geometry*. Springer.

Azaïs,J-M. and Wschebor M. *Level set and extrema of random processes and fields*, Wiley (2009).

Mercadier, C. (2006), Numerical Bounds for the Distribution of the Maximum of Some One- and Two-Parameter Gaussian Processes, *Adv. in Appl. Probab.* **38**, pp. 149–170.

THANK-YOU

GRACIAS