Efficient Simulation for Tail Probabilities of Gaussian Random Fields

Jingchen Liu
Columbia University

Joint work with
Robert Adler (Technion-Israel Institute of Technology)
Jose Blanchet (Columbia University)

Overview

- Large deviations results on Gaussian random fields
- Euler characteristic of excursion sets of Gaussian random fields
- Simulation of excursion probabilities
- Further discussions

Random Field in a Compact Set

- Gaussian random field,

$$f(t,\omega) : T \times \Omega \longrightarrow R$$

where $T \subset R^d$ is a compact set.

- $E(f(t)) = 0$

- $Var(f(s), f(t)) = C(s,t)$

- $f(t,\omega)$ is almost surely twice differentiable with respect to t.

Excursion Probability

- Interesting probability

$$P(sup_{t \in T} f(t) > b)$$

as b ! 1.

- More generally,

$$E\left(\Gamma\left(f\left(\cdot\right)\right) \middle| \sup_{t \in T} f\left(t\right) > b\right)$$

where Γ(¢) is bounded.

Excursion Probability – Borell-TIS Theorem

Theorem 1 (Borell-TIS) *Let $f(\cdot)$ be a real-valued, separable, continuous Gaussian process. Suppose that*

$$\sigma^2(T) = \sup_T Var(f(t)) < \infty.$$

Let $m = \sup_T Ef(t)$ and choose any a for which

$$P\left(\sup_T (f(t) - m) \geq a\right) \leq \frac{1}{2}.$$

Then, for all b,

$$P\left(\sup_T f(t) > b\right) \leq 2\left(1 - \Phi\left(\frac{b - m - a}{\sigma(T)}\right)\right),$$

where $\Phi(\cdot)$ is the c.d.f. of the standard Gaussian distribution.

Excursion Probability

- Borell-TIS theorem – crude upper bound

$$P\left(\sup_T f(t) > b\right) \leq 2\left(1 - \Phi\left(\frac{b - m - a}{\sigma(T)}\right)\right),$$

- Trivial lower bound

$$P\left(\sup_T f(t) > b\right) \geq 1 - \Phi\left(b/\sigma(T)\right).$$

- Very rough approximation

$$P\left(\sup_T f(t) > b\right) \approx \exp\left(-\frac{b^2}{2\sigma^2(T)}\right)$$

Excursion set

$$f^{-1}(A) = \{t \in T; f(t) \in A\}$$

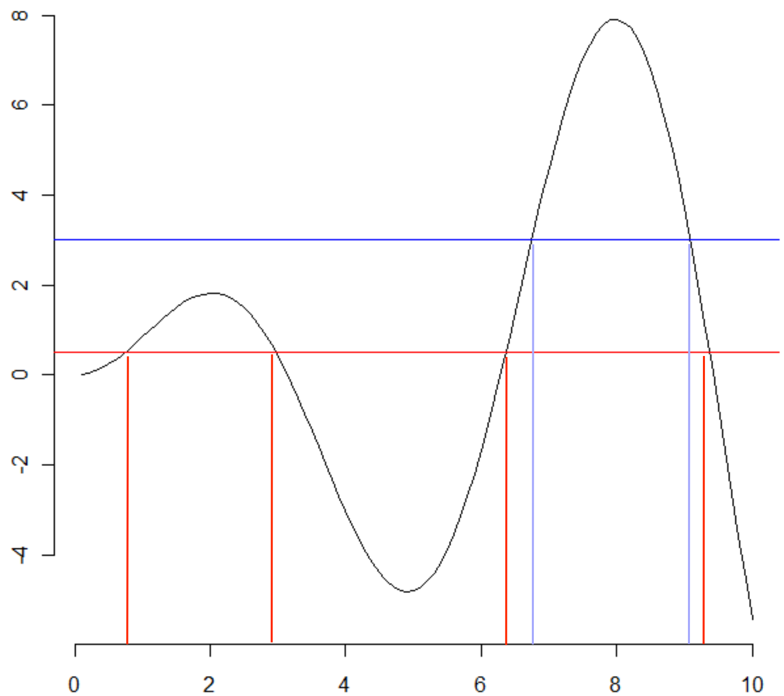
- Excursion set

$$f^{-1}((b, +\infty)) = \{t \in T; f(t) \geq b\}$$

- Equivalently

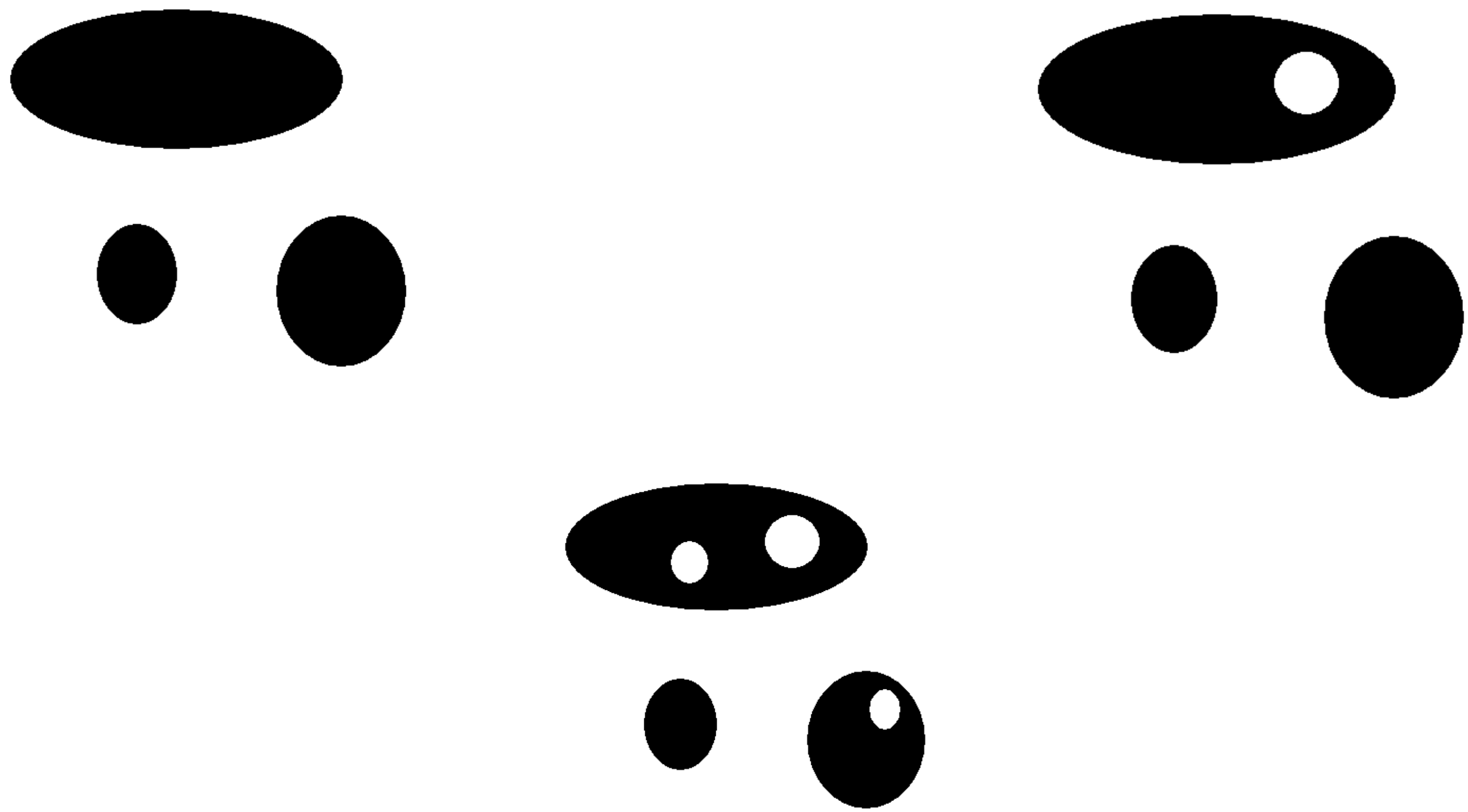
$$P(sup_{t \in T} f(t) > b) = P(f^{-1}((b, +\infty)) \neq \emptyset)$$

Euler Characteristics – One Dimension



$$\chi(f^{-1}((b, +\infty)))$$

Euler Characteristics – Two Dimension



Euler Characteristics of Excursion sets

- Taylor and Adler (2003) for stationary twice differentiable fields

$$E(\chi(f^{-1}((b,+\infty)))) = e^{-b^2/2\sigma^2}\sum_{k=1}^{N}\sum_{J\in O_k}\frac{|J||\Lambda_J|^{1/2}}{(2\pi)^{(k+1)/2}\sigma^k}H_{k-1}(b/\sigma)+\bar{\Phi}(b/\sigma)$$

- Sharp asymptotics

$$P(sup_{t\in T}f(t) > b) = (1+o(1))C(T)b^{k-1}e^{-b^2/2\sigma^2}$$

k=dim(T)

- Pickands (1969), Piterbarg (1995)

Simulation of Random Fields

- Conceptual challenge: infinite object
- Discretize the space T, $\{t_1,\ldots,t_d\}$
- Karhunen-Loeve expansion

$$f_Z(t) = \sum_{i=1}^{n} Z_i \sigma_i(t)$$

- Asmussen, Blanchet, Juneja, and Rojas-Nandayapa (2008)

Maximum of Multivariate Gaussian

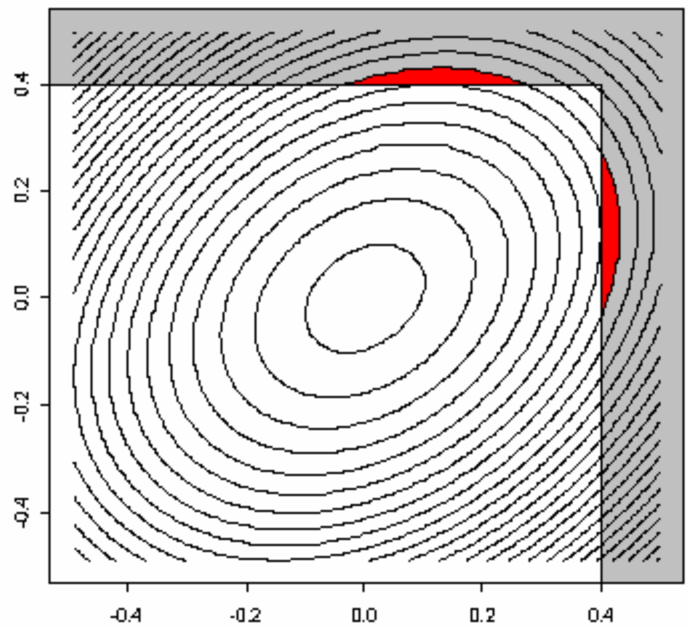
$$\{t_1, ..., t_d\} \subset T$$

$$X_i = f(t_i)$$

$$X = (X_1, ..., X_d) \sim N(0, \Sigma)$$

$$P\left(\max_j X_j > b\right)$$

Two Dimensional Example



The Change-of-measure

- Intuition: the excursion is caused by one component being large

- The proposed change of measure

$$g(x) = \sum_{j=1}^{d} p_j(b) \frac{f_j\left(x_j\right) f\left(x_{-j} | x_j\right)}{P\left(X_j > b\right)} I\left(x_j > b\right)$$

where $p_j(b) = P\left(X_j > b\right) / \sum_{i=1}^{d} P(X_i > b)$.

Efficiency

Let X be a d dimensional random vector with the multivariate Gaussian distribution $N(0, \Sigma)$, where Σ is a positive definite matrix. The sampling distribution

$$g(x) = \sum_{j=1}^{d} p_j(b) \frac{f_j(x_j) f(x_{-j}|x_j)}{P(X_j > b)} I(x_j > b)$$

provides asymptotically zero relative error for the exceedence probability. That is, the corresponding estimator $L = f(x)/g(x)$ in satisfies

$$\frac{Var(L)}{P(\max_j X_j > b)^2} \to 0,$$

as $b \to \infty$. Hence, L is strongly efficient. $f(x)$ is the density function of X.

Finite Karhunen-Loeve expansion

- $f_Z(t) = \sum_{i=1}^{n} Z_i \sigma_i(t)$
- Simulate $Z_1, \ldots, Z_n$
- Propose change-of-measure

$$q_\theta(z_1, \ldots, z_n) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \exp\left(-\frac{1}{2\theta^2}\sum z_i^2\right)$$

$$L_b = I(\|f_z\| > b)\,\theta^d \exp\left(-\frac{1}{2}\left(1 - \frac{1}{\theta^2}\right)\sum z_i^2\right)$$

Efficiency

Take f as defined by,

$$f_Z\left(t\right) = \sum_{i=1}^{n} Z_i \sigma_i\left(t\right)$$

and choose the $\theta = b$, giving the likelihood ratio L_b. Then the corresponding importance sampling estimate for the exceedence probability is asymptotically efficient, that is,

$$\frac{\log E(L_b^2)}{\log P\left(\max_j X_j > b\right)^2} \to 1,$$

as $b \to \infty$.

Example

We compute $P\left(\max_i X_i > b\right)$, where $X = (X_1, ..., X_5)$ is multivariate Gaussian with mean zero, marginal variance one. They share a common correlation 0.5.

b	Est.	Std. Er.
3	6.01E-3	4.70E-5
4	1.53E-4	6.55E-7
5	1.43e-06	2.26E-9

Example

$f(t) = X\cos(t) + Y\sin(t)$ and $T = [0, 3/4]$,
where X and Y are i.i.d. $N(0,1)$.

$$f(\sup_{[0,3/4]} f(t) > b) = 1 - \Phi(b) + \frac{3}{8}e^{-b^2/2}$$

b	True Value	Est.	Std. Er.
1.5	0.1055	0.1057	7.96E-04
2.0	0.0389	0.0392	4.02E-04
2.5	0.0115	0.0118	1.57E-04
3	2.68E-03	2.60E-03	4.31E-05
5	7.31E-07	7.56E-07	2.33E-08
10	3.06E-23	3.15E-23	2.17E-24

The discretization

- Depending on the smoothness of the field
- Discretize the space T, $\{t_1,\ldots,t_d\}$
  - Twice differentiable fields: $d = O(b^k)$
- Karhunen-Loeve expansion

$$f_Z(t) = \sum_{i=1}^{n} Z_i \sigma_i(t)$$

  - $n=O(\log(b))$

Summary

- Large deviations probabilities of Gaussian random fields
- Euler characteristic of excursion sets
- Efficient simulation of high excursion probabilities
  - Discretize the space T
  - Finite Karhunen-Loeve expansion
- Discussion of discretization