GEOMETRY OF CERTAIN STOCHASTIC PROCESSES

Michel Talagrand

C.N.R.S.

Overview of talk

Overview of talk

- Fundamental Problem: how to bound the supremum of a collection of random variables (r.v.s) $(X_t)_{t \in T}$.

Overview of talk

- Fundamental Problem: how to bound the supremum of a collection of random variables (r.v.s) $(X_t)_{t \in T}$.
- Basic method: Kolmogorov's chaining

Overview of talk

- Fundamental Problem: how to bound the supremum of a collection of random variables (r.v.s) $(X_t)_{t \in T}$.
- Basic method: Kolmogorov's chaining
- Kolmogorov's chaining is optimal for Gaussian processes when properly used.

Overview of talk

- Fundamental Problem: how to bound the supremum of a collection of random variables (r.v.s) $(X_t)_{t \in T}$.
- Basic method: Kolmogorov's chaining
- Kolmogorov's chaining is optimal for Gaussian processes when properly used. The chaining takes a few steps of a few lines each.

Overview of talk

- Fundamental Problem: how to bound the supremum of a collection of random variables (r.v.s) $(X_t)_{t \in T}$.
- Basic method: Kolmogorov's chaining
- Kolmogorov's chaining is optimal for Gaussian processes when properly used. The chaining takes a few steps of a few lines each. It took over 50 years to find them.

Overview of talk

- Fundamental Problem: how to bound the supremum of a collection of random variables (r.v.s) $(X_t)_{t \in T}$.
- Basic method: Kolmogorov's chaining
- Kolmogorov's chaining is optimal for Gaussian processes when properly used. The chaining takes a few steps of a few lines each. It took over 50 years to find them.
- An important class of processes is defined by random series of functions:

$$X_t = \sum_{i \geq 1} f_i(t)\xi_i,$$

where ξ_i are independent r.v.s.

Overview of talk

- Fundamental Problem: how to bound the supremum of a collection of random variables (r.v.s) $(X_t)_{t \in T}$.
- Basic method: Kolmogorov's chaining
- Kolmogorov's chaining is optimal for Gaussian processes when properly used. The chaining takes a few steps of a few lines each. It took over 50 years to find them.
- An important class of processes is defined by random series of functions:

$$X_t = \sum_{i \geq 1} f_i(t) \xi_i,$$

  where ξ_i are independent r.v.s. Published bounds on these processes always use a combination of Kolmogorov's chaining and "trivial" bounds.

Overview of talk

- Fundamental Problem: how to bound the supremum of a collection of random variables (r.v.s) $(X_t)_{t \in T}$.
- Basic method: Kolmogorov's chaining
- Kolmogorov's chaining is optimal for Gaussian processes when properly used. The chaining takes a few steps of a few lines each. It took over 50 years to find them.
- An important class of processes is defined by random series of functions:

$$X_t = \sum_{i \geq 1} f_i(t) \xi_i,$$

  where ξ_i are independent r.v.s. Published bounds on these processes always use a combination of Kolmogorov's chaining and "trivial" bounds.
- Conjectures: There is actually no other way to bound such random processes.

Bounding a supremum of r.v.s $(X_t)_{t \in T}$

Bounding a supremum of r.v.s $(X_t)_{t \in T}$

Main issue:

Bounding a supremum of r.v.s $(X_t)_{t \in T}$

Main issue:

- Each r.v. X_t has a small chance to be much larger than its typical value.

Bounding a supremum of r.v.s $(X_t)_{t \in T}$

Main issue:

- Each r.v. X_t has a small chance to be much larger than its typical value.
- There are many such variables, and there is a fair chance that at least one of them is much larger than its typical value.

Bounding a supremum of r.v.s $(X_t)_{t \in T}$

Main issue:

- Each r.v. X_t has a small chance to be much larger than its typical value.
- There are many such variables, and there is a fair chance that at least one of them is much larger than its typical value.

To quantify the size of $\sup_{t \in T} X_t$ we consider an arbitrary point t_0 of T and the r.v.

$$Y = \sup_{t \in T} (X_t - X_{t_0}) \geq 0.$$

Bounding a supremum of r.v.s $(X_t)_{t\in T}$

Main issue:

- Each r.v. X_t has a small chance to be much larger than its typical value.
- There are many such variables, and there is a fair chance that at least one of them is much larger than its typical value.

To quantify the size of $\sup_{t\in T} X_t$ we consider an arbitrary point t_0 of T and the r.v.

$$Y = \sup_{t\in T}(X_t - X_{t_0}) \geq 0.$$

It is technically convenient that $Y \geq 0$ since then

$$\mathsf{E}Y = \int_0^\infty \mathsf{P}(Y \geq u)\mathrm{d}u \ .$$

Bounding a supremum of r.v.s $(X_t)_{t\in T}$

Main issue:

- Each r.v. X_t has a small chance to be much larger than its typical value.
- There are many such variables, and there is a fair chance that at least one of them is much larger than its typical value.

To quantify the size of $\sup_{t\in T} X_t$ we consider an arbitrary point t_0 of T and the r.v.

$$Y = \sup_{t\in T}(X_t - X_{t_0}) \geq 0.$$

It is technically convenient that $Y \geq 0$ since then

$$\mathsf{E}Y = \int_0^\infty \mathsf{P}(Y \geq u)\mathrm{d}u \, .$$

The issue is then to bound

$$\mathsf{P}(Y \geq u) = \mathsf{P}\Big(\bigcup_{t\in T}\{X_t - X_{t_0} \geq u\}\Big)$$

for $u > 0$.

The basic idea of chaining

The basic idea of chaining

- It is to use the "union bound"

$$P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i)$$

The basic idea of chaining

- It is to use the "union bound"

$$\mathrm{P}\Big(\bigcup_i A_i\Big) \le \sum_i \mathrm{P}(A_i)$$

so that

$$\mathrm{P}(Y \ge u) = \mathrm{P}\Big(\bigcup_{t\in T}\{X_t - X_{t_0} \ge u\}\Big) \le \sum_{t\in T} \mathrm{P}(X_t - X_{t_0} \ge u).$$

The basic idea of chaining

- It is to use the "union bound"

$$\mathsf{P}\Big(\bigcup_i A_i\Big) \le \sum_i \mathsf{P}(A_i)$$

so that

$$\mathsf{P}(Y \ge u) = \mathsf{P}\Big(\bigcup_{t\in T}\{X_t - X_{t_0} \ge u\}\Big) \le \sum_{t\in T} \mathsf{P}(X_t - X_{t_0} \ge u).$$

This is terribly ineffective if the r.v.s X_t are nearly identical!

The basic idea of chaining

- It is to use the "union bound"
$$\mathsf{P}\Big(\bigcup_i A_i\Big) \le \sum_i \mathsf{P}(A_i)$$
so that
$$\mathsf{P}(Y \ge u) = \mathsf{P}\Big(\bigcup_{t\in T}\{X_t - X_{t_0} \ge u\}\Big) \le \sum_{t\in T} \mathsf{P}(X_t - X_{t_0} \ge u).$$
This is terribly ineffective if the r.v.s X_t are nearly identical!
- Remedy: For $t \in T$ consider an approximation $\pi_1(t)$ of t,

The basic idea of chaining

- It is to use the "union bound"
$$\mathsf{P}\Big(\bigcup_i A_i\Big) \le \sum_i \mathsf{P}(A_i)$$
so that
$$\mathsf{P}(Y \ge u) = \mathsf{P}\Big(\bigcup_{t\in T}\{X_t - X_{t_0} \ge u\}\Big) \le \sum_{t\in T} \mathsf{P}(X_t - X_{t_0} \ge u).$$
This is terribly ineffective if the r.v.s X_t are nearly identical!
- Remedy: For $t \in T$ consider an approximation $\pi_1(t)$ of t, with $\pi_1(t) \in T_1 \subset T$ and card T_1 not too large,

The basic idea of chaining

- It is to use the "union bound"

$$\mathrm{P}\Big(\bigcup_i A_i\Big) \le \sum_i \mathrm{P}(A_i)$$

  so that

$$\mathrm{P}(Y \ge u) = \mathrm{P}\Big(\bigcup_{t\in T}\{X_t - X_{t_0} \ge u\}\Big) \le \sum_{t\in T}\mathrm{P}(X_t - X_{t_0} \ge u).$$

  This is terribly ineffective if the r.v.s X_t are nearly identical!
- Remedy: For $t \in T$ consider an approximation $\pi_1(t)$ of t, with $\pi_1(t) \in T_1 \subset T$ and card T_1 not too large, and write

$$X_t - X_{t_0} = X_t - X_{\pi_1(t)} + X_{\pi_1(t)} - X_{t_0}\ .$$

The basic idea of chaining

- It is to use the "union bound"

$$\mathrm{P}\Big(\bigcup_i A_i\Big) \le \sum_i \mathrm{P}(A_i)$$

so that

$$\mathrm{P}(Y \ge u) = \mathrm{P}\Big(\bigcup_{t\in T}\{X_t - X_{t_0} \ge u\}\Big) \le \sum_{t\in T} \mathrm{P}(X_t - X_{t_0} \ge u).$$

This is terribly ineffective if the r.v.s X_t are nearly identical!

- Remedy: For $t \in T$ consider an approximation $\pi_1(t)$ of t, with $\pi_1(t) \in T_1 \subset T$ and card T_1 not too large, and write

$$X_t - X_{t_0} = X_t - X_{\pi_1(t)} + X_{\pi_1(t)} - X_{t_0}\ .$$

• $\sup_{t\in T}(X_t - X_{\pi_1(t)})$ should be easier to bound than Y because it consists of smaller r.v.s. since $\pi_1(t)$ is an approximation of t.

The basic idea of chaining

- It is to use the "union bound"

$$\mathsf{P}\Big(\bigcup_i A_i\Big) \le \sum_i \mathsf{P}(A_i)$$

so that

$$\mathsf{P}(Y \ge u) = \mathsf{P}\Big(\bigcup_{t\in T}\{X_t - X_{t_0} \ge u\}\Big) \le \sum_{t\in T} \mathsf{P}(X_t - X_{t_0} \ge u).$$

This is terribly ineffective if the r.v.s X_t are nearly identical!

- Remedy: For $t \in T$ consider an approximation $\pi_1(t)$ of t, with $\pi_1(t) \in T_1 \subset T$ and card T_1 not too large, and write

$$X_t - X_{t_0} = X_t - X_{\pi_1(t)} + X_{\pi_1(t)} - X_{t_0}\ .$$

• $\sup_{t\in T}(X_t - X_{\pi_1(t)})$ should be easier to bound than Y because it consists of smaller r.v.s. since $\pi_1(t)$ is an approximation of t.

• $\sup_{t\in T}(X_{\pi_1(t)} - X_{t_0})$ should be easier to bound than Y because there are not so many different r.v.s $X_{\pi_1(t)}$.

Organizing chaining, I

Organizing chaining, I

We iterate the previous method by considering for $n \geq 0$ successive approximations $\pi_n(t) \in T_n$, with $\pi_0(t) = t_0$.

Organizing chaining, I

We iterate the previous method by considering for $n \geq 0$ successive approximations $\pi_n(t) \in T_n$, with $\pi_0(t) = t_0$. Assuming T finite, we have $\pi_n(t) = t$ for n large enough.

Organizing chaining, I

We iterate the previous method by considering for $n \geq 0$ successive approximations $\pi_n(t) \in T_n$, with $\pi_0(t) = t_0$. Assuming T finite, we have $\pi_n(t) = t$ for n large enough.
We write the basic chaining identity

$$X_t - X_{t_0} = \sum_{n \geq 1} (X_{\pi_n(t)} - X_{\pi_{n-1}(t)}) .$$

Organizing chaining, I

We iterate the previous method by considering for $n \geq 0$ successive approximations $\pi_n(t) \in T_n$, with $\pi_0(t) = t_0$. Assuming T finite, we have $\pi_n(t) = t$ for n large enough.
We write the basic chaining identity

$$X_t - X_{t_0} = \sum_{n \geq 1} (X_{\pi_n(t)} - X_{\pi_{n-1}(t)}) .$$

Plan of action:

Organizing chaining, I

We iterate the previous method by considering for $n \geq 0$ successive approximations $\pi_n(t) \in T_n$, with $\pi_0(t) = t_0$. Assuming T finite, we have $\pi_n(t) = t$ for n large enough.
We write the basic chaining identity

$$X_t - X_{t_0} = \sum_{n \geq 1} (X_{\pi_n(t)} - X_{\pi_{n-1}(t)}) \; .$$

PLAN OF ACTION: Prove that with probability close to one each of the differences $X_{\pi_n(t)} - X_{\pi_{n-1}(t)}$ is not too large and therefore that $Y = \sup_{t \in T}(X_t - X_{t_0})$ is not too large.

Organizing chaining, I

We iterate the previous method by considering for $n \geq 0$ successive approximations $\pi_n(t) \in T_n$, with $\pi_0(t) = t_0$. Assuming T finite, we have $\pi_n(t) = t$ for n large enough.
We write the basic chaining identity

$$X_t - X_{t_0} = \sum_{n \geq 1} (X_{\pi_n(t)} - X_{\pi_{n-1}(t)}) \ .$$

PLAN OF ACTION: Prove that with probability close to one each of the differences $X_{\pi_n(t)} - X_{\pi_{n-1}(t)}$ is not too large and therefore that $Y = \sup_{t \in T}(X_t - X_{t_0})$ is not too large.

REVISED PLAN OF ACTION: Prove that with probability close to one, for each n, each $s \in T_n$, each $s' \in T_{n-1}$ one can suitably control the difference $X_s - X_{s'}$ and therefore that Y is not too large.

Organizing chaining II

Organizing chaining II

Specifically, (following X. Fernique) we try to acheive that with probability close to 1, for certain numbers $c(n, s, s')$,

$$\forall\, n \geq 1\ ,\ \forall\, s \in T_n\ ,\ \forall\, s' \in T_{n-1}\ ,\ |X_s - X_{s'}| \leq c(n, s, s')\ , \qquad (1)$$

Organizing chaining II

Specifically, (following X. Fernique) we try to acheive that with probability close to 1, for certain numbers $c(n, s, s')$,

$$\forall\, n \geq 1 \,,\ \forall\, s \in T_n \,,\ \forall\, s' \in T_{n-1} \,,\ |X_s - X_{s'}| \leq c(n, s, s') \,, \qquad (1)$$

so that then

$$X_t - X_{t_0} = \sum_{n \geq 1} (X_{\pi_n(t)} - X_{\pi_{n-1}(t)}) \leq \sum_{n \geq 1} c(n, \pi_n(t), \pi_{n-1}(t))$$

Organizing chaining II

Specifically, (following X. Fernique) we try to acheive that with probability close to 1, for certain numbers $c(n, s, s')$,

$$\forall\, n \geq 1\ ,\ \forall\, s \in T_n\ ,\ \forall\, s' \in T_{n-1}\ ,\ |X_s - X_{s'}| \leq c(n, s, s')\ , \qquad (1)$$

so that then

$$X_t - X_{t_0} = \sum_{n \geq 1} (X_{\pi_n(t)} - X_{\pi_{n-1}(t)}) \leq \sum_{n \geq 1} c(n, \pi_n(t), \pi_{n-1}(t))$$

and consequently

$$(1) \Rightarrow Y = \sup_{t \in T} (X_t - X_{t_0}) \leq \sup_{t \in T} \sum_{n \geq 1} c(n, \pi_n(t), \pi_{n-1}(t))$$

with the same probability close to 1.

Chaining for Gaussian processes I

Chaining for Gaussian processes I

A Gaussian process is such that the family $(X_t)_{t \in T}$ is jointly Gaussian. Then $d(s, s') = (\mathrm{E}(X_s - X_{s'})^2)^{1/2}$ is a distance on T. Basic intuition: the "geometry" of the metric space (T, d) determines the "size" of the Gaussian process.

Chaining for Gaussian processes I

A Gaussian process is such that the family $(X_t)_{t\in T}$ is jointly Gaussian. Then $d(s,s') = (\mathsf{E}(X_s - X_{s'})^2)^{1/2}$ is a distance on T. Basic intuition: the "geometry" of the metric space (T,d) determines the "size" of the Gaussian process. The fundamental fact (which reflects the size of the tails of the Gaussian r.v. $X_s - X_{s'}$): If $v > 0$,

$$\mathsf{P}(|X_s - X_{s'}| \geq v) \leq 2\exp\left(-\frac{v^2}{2d(s,s')^2}\right) . \tag{2}$$

Chaining for Gaussian processes I

A Gaussian process is such that the family $(X_t)_{t\in T}$ is jointly Gaussian. Then $d(s,s') = (\mathrm{E}(X_s - X_{s'})^2)^{1/2}$ is a distance on T.
Basic intuition: the "geometry" of the metric space (T,d) determines the "size" of the Gaussian process.
The fundamental fact (which reflects the size of the tails of the Gaussian r.v. $X_s - X_{s'}$): If $v > 0$,

$$\mathrm{P}(|X_s - X_{s'}| \geq v) \leq 2\exp\left(-\frac{v^2}{2d(s,s')^2}\right) . \tag{2}$$

The inverse function of $\exp(x^2)$ is $\sqrt{\log x}$.

Chaining for Gaussian processes I

A Gaussian process is such that the family $(X_t)_{t\in T}$ is jointly Gaussian. Then $d(s,s') = (\mathrm{E}(X_s - X_{s'})^2)^{1/2}$ is a distance on T.
Basic intuition: the "geometry" of the metric space (T,d) determines the "size" of the Gaussian process.
The fundamental fact (which reflects the size of the tails of the Gaussian r.v. $X_s - X_{s'}$): If $v > 0$,

$$\mathrm{P}(|X_s - X_{s'}| \geq v) \leq 2\exp\left(-\frac{v^2}{2d(s,s')^2}\right) . \qquad (2)$$

The inverse function of $\exp(x^2)$ is $\sqrt{\log x}$. It is then convenient to impose the condition that

$$T_n := \{\pi_n(t)\ ;\ t \in T\}$$

satisfies card $T_n \leq 2^{2^n}$ for $n \geq 1$. (Note than $T_0 = \{t_0\}$.)

Chaining for Gaussian processes I

A Gaussian process is such that the family $(X_t)_{t\in T}$ is jointly Gaussian. Then $d(s,s') = (\mathrm{E}(X_s - X_{s'})^2)^{1/2}$ is a distance on T.
Basic intuition: the "geometry" of the metric space (T,d) determines the "size" of the Gaussian process.
The fundamental fact (which reflects the size of the tails of the Gaussian r.v. $X_s - X_{s'}$): If $v > 0$,

$$\mathrm{P}(|X_s - X_{s'}| \geq v) \leq 2\exp\Big(-\frac{v^2}{2d(s,s')^2}\Big)\,. \qquad (2)$$

The inverse function of $\exp(x^2)$ is $\sqrt{\log x}$. It is then convenient to impose the condition that

$$T_n := \{\pi_n(t)\ ;\ t \in T\}$$

satisfies card $T_n \leq 2^{2^n}$ for $n \geq 1$. (Note than $T_0 = \{t_0\}$.) Then $\sqrt{\log \mathrm{card}\, T_n}$ grows geometrically.

Chaining for Gaussian processes II

Chaining for Gaussian processes II

We want to impose the condition

$$\forall\, n \geq 1\ ,\ \forall\, s \in T_n\ ,\ \forall\, s' \in T_{n-1}\ ,\ |X_s - X_{s'}| \leq c(n, s, s') \qquad (1)$$

for a certain well-chosen $c(n, s, s')$.
By the union bound the probability that (1) **fails** is at most

$$\sum_{n \geq 1} \sum_{s \in T_n, s' \in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n, s, s')).$$

Chaining for Gaussian processes II

We want to impose the condition

$$\forall\, n \geq 1 \;,\; \forall\, s \in T_n \;,\; \forall\, s' \in T_{n-1} \;,\; |X_s - X_{s'}| \leq c(n, s, s') \qquad (1)$$

for a certain well-chosen $c(n, s, s')$.
By the union bound the probability that (1) **fails** is at most

$$\sum_{n \geq 1} \sum_{s \in T_n, s' \in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n, s, s')).$$

Chaining for Gaussian processes II

We want to impose the condition

$$\forall\, n \geq 1 \;,\; \forall\, s \in T_n \;,\; \forall\, s' \in T_{n-1} \;,\; |X_s - X_{s'}| \leq c(n, s, s') \qquad (1)$$

for a certain well-chosen $c(n, s, s')$.
By the union bound the probability that (1) **fails** is at most

$$\sum_{n \geq 1} \sum_{s \in T_n, s' \in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n, s, s')).$$

In the next slide we show that a magic choice is

$$c(n, s, s') = v 2^{n/2} d(s, s') \;,$$

where $v > 0$ is a parameter.

Chaining for Gaussian processes II

We want to impose the condition

$$\forall\, n \geq 1\ ,\ \forall\, s \in T_n\ ,\ \forall\, s' \in T_{n-1}\ ,\ |X_s - X_{s'}| \leq c(n, s, s') \qquad (1)$$

for a certain well-chosen $c(n, s, s')$.
By the union bound the probability that (1) **fails** is at most

$$\sum_{n \geq 1} \sum_{s \in T_n, s' \in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n, s, s')).$$

In the next slide we show that a magic choice is

$$c(n, s, s') = v 2^{n/2} d(s, s')\ ,$$

where $v > 0$ is a parameter.

Chaining for Gaussian processes III

Chaining for Gaussian processes III

Now we are going to bound

$$\sum_{n \geq 1} \sum_{s \in T_n, s' \in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n, s, s')). \tag{3}$$

Chaining for Gaussian processes III

Now we are going to bound

$$\sum_{n\geq 1}\sum_{s\in T_n,s'\in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n,s,s')). \tag{3}$$

Using the "increment condition" (2) for s and s' yields

$$\mathrm{P}(|X_s - X_{s'}| > c(n,s,s')) \leq 2\exp\left(-\frac{c(n,s,s')^2}{2d(s,s')^2}\right)$$

Chaining for Gaussian processes III

Now we are going to bound

$$\sum_{n \geq 1} \sum_{s \in T_n, s' \in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n, s, s')). \qquad (3)$$

Using the "increment condition" (2) for s and s' yields

$$\begin{aligned} & \mathrm{P}(|X_s - X_{s'}| > c(n, s, s')) \leq 2 \exp\left(-\frac{c(n, s, s')^2}{2d(s, s')^2}\right) \\ = \quad & 2 \exp\left(-\frac{(v 2^{n/2} d(s, s'))^2}{2d(s, s')^2}\right) = 2 \exp(-v^2 2^{n-1}). \end{aligned}$$

Chaining for Gaussian processes III

Now we are going to bound

$$\sum_{n\geq 1} \sum_{s\in T_n, s'\in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n,s,s')). \tag{3}$$

Using the "increment condition" (2) for s and s' yields

$$\begin{aligned} & \mathrm{P}(|X_s - X_{s'}| > c(n,s,s')) \leq 2\exp\Big(-\frac{c(n,s,s')^2}{2d(s,s')^2}\Big) \\ = \quad & 2\exp\Big(-\frac{(v2^{n/2}d(s,s'))^2}{2d(s,s')^2}\Big) = 2\exp(-v^2 2^{n-1}). \end{aligned}$$

Now $\operatorname{card} T_n \operatorname{card} T_{n-1} \leq \left(2^{2^n}\right)^2 = 2^{2^{n+1}}$.

Chaining for Gaussian processes III

Now we are going to bound

$$\sum_{n\geq 1}\sum_{s\in T_n, s'\in T_{n-1}} \mathrm{P}(|X_s - X_{s'}| > c(n,s,s')). \qquad (3)$$

Using the "increment condition" (2) for s and s' yields

$$\begin{aligned} \mathrm{P}(|X_s - X_{s'}| > c(n,s,s')) &\leq 2\exp\Big(-\frac{c(n,s,s')^2}{2d(s,s')^2}\Big) \\ = \quad 2\exp\Big(-\frac{(v2^{n/2}d(s,s'))^2}{2d(s,s')^2}\Big) &= 2\exp(-v^2 2^{n-1}). \end{aligned}$$

Now $\operatorname{card} T_n \operatorname{card} T_{n-1} \leq \left(2^{2^n}\right)^2 = 2^{2^{n+1}}$. Therefore, the quantity (3) is at most

$$\sum_{n\geq 1} 2\cdot 2^{2^{n+1}} \exp(-v^2 2^{n-1}) \ ,$$

and for $v \geq 100$ this is $\leq C\exp(-v^2/2)$.

Chaining for Gaussian processes IV

So (1) holds with probability $\geq 1 - C\exp(-v^2/2)$. Moreover

$$
\begin{aligned}
(1) \Rightarrow Y = \sup_{t\in T}(X_t - X_{t_0}) &\leq \sup_{t\in T}\sum_{n\geq 1} c(n, \pi_n(t), \pi_{n-1}(t)) \\
&= v \sup_{t\in T}\sum_{n\geq 1} 2^{n/2} d(\pi_n(t), \pi_{n-1}(t)) \ .
\end{aligned}
$$

Since $\mathrm{E}Y = \int_0^\infty \mathrm{P}(Y \geq u)\mathrm{d}u$ we obtain

$$
\mathrm{E}Y = \mathrm{E}\sup_{t\in T}(X_t - X_{t_0}) \leq C \sup_{t\in T}\sum_{n\geq 1} 2^{n/2} d(\pi_n(t), \pi_{n-1}(t)) \ . \quad (4)
$$

Chaining for Gaussian processes IV

So (1) holds with probability $\geq 1 - C\exp(-v^2/2)$. Moreover

$$\begin{aligned}
(1) \Rightarrow Y = \sup_{t\in T}(X_t - X_{t_0}) &\leq \sup_{t\in T}\sum_{n\geq 1} c(n, \pi_n(t), \pi_{n-1}(t)) \\
&= v\sup_{t\in T}\sum_{n\geq 1} 2^{n/2} d(\pi_n(t), \pi_{n-1}(t)) .
\end{aligned}$$

Since $\mathrm{E}Y = \int_0^\infty \mathrm{P}(Y \geq u)\mathrm{d}u$ we obtain

$$\mathrm{E}Y = \mathrm{E}\sup_{t\in T}(X_t - X_{t_0}) \leq C\sup_{t\in T}\sum_{n\geq 1} 2^{n/2} d(\pi_n(t), \pi_{n-1}(t)) . \quad (4)$$

Since $\pi_n(t) \in T_n$ is an approximation of t it is natural require

$$d(t, \pi_n(t)) = d(t, T_n) = \inf\{d(t,s)\ ;\ s \in T\} .$$

Chaining for Gaussian processes IV

So (1) holds with probability $\geq 1 - C\exp(-v^2/2)$. Moreover

$$\begin{aligned}(1) \Rightarrow Y = \sup_{t\in T}(X_t - X_{t_0}) &\leq \sup_{t\in T}\sum_{n\geq 1} c(n, \pi_n(t), \pi_{n-1}(t)) \\ &= v\sup_{t\in T}\sum_{n\geq 1} 2^{n/2} d(\pi_n(t), \pi_{n-1}(t)) \ .\end{aligned}$$

Since $\mathrm{E}Y = \int_0^\infty \mathrm{P}(Y \geq u)\mathrm{d}u$ we obtain

$$\mathrm{E}Y = \mathrm{E}\sup_{t\in T}(X_t - X_{t_0}) \leq C\sup_{t\in T}\sum_{n\geq 1} 2^{n/2} d(\pi_n(t), \pi_{n-1}(t)) \ . \qquad (4)$$

Since $\pi_n(t) \in T_n$ is an approximation of t it is natural require

$$d(t, \pi_n(t)) = d(t, T_n) = \inf\{d(t,s) \ ; \ s \in T\} \ .$$

Since $d(\pi_n(t), \pi_{n-1}(t)) \leq d(\pi_n(t), t) + d(t, \pi_{n-1}(t))$, (4) yields

$$\mathrm{E}\sup_{t\in T} X_t = \mathrm{E}\sup_{t\in T}(X_t - X_{t_0}) \leq C\sup_{t\in T}\sum_{n\geq 1} 2^{n/2} d(t, T_n) \ .$$

The Majorizing Measure Theorem

Since this bound holds for each choice of T_n (with card $T_n \leq 2^{2^n}$) we have proved

$$\mathsf{E} \sup_{t \in T} X_t \leq C\eta(T, d) , \qquad (5)$$

where

$$\eta(T, d) := \inf \sup_{t \in T} \sum_{n \geq 1} 2^{n/2} d(t, T_n) ,$$

for an infimum over all possible choices of the sets T_n.

The Majorizing Measure Theorem

Since this bound holds for each choice of T_n (with card $T_n \leq 2^{2^n}$) we have proved

$$\mathsf{E}\sup_{t\in T} X_t \leq C\eta(T,d)\ , \tag{5}$$

where

$$\eta(T,d) := \inf\sup_{t\in T}\sum_{n\geq 1} 2^{n/2} d(t, T_n)\ ,$$

for an infimum over all possible choices of the sets T_n. ((5) is called the *generic chaining* bound).

The Majorizing Measure Theorem

Since this bound holds for each choice of T_n (with card $T_n \le 2^{2^n}$) we have proved

$$\mathrm{E}\sup_{t\in T} X_t \le C\eta(T,d)\,, \tag{5}$$

where

$$\eta(T,d) := \inf \sup_{t\in T} \sum_{n\ge 1} 2^{n/2} d(t, T_n)\,,$$

for an infimum over all possible choices of the sets T_n. ((5) is called the *generic chaining* bound). How good is this bound?

The Majorizing Measure Theorem

Since this bound holds for each choice of T_n (with card $T_n \leq 2^{2^n}$) we have proved

$$\mathrm{E} \sup_{t \in T} X_t \leq C\eta(T, d) \,, \tag{5}$$

where

$$\eta(T, d) := \inf \sup_{t \in T} \sum_{n \geq 1} 2^{n/2} d(t, T_n) \,,$$

for an infimum over all possible choices of the sets T_n. ((5) is called the *generic chaining* bound). How good is this bound? It is the best possible:

The Majorizing Measure Theorem

Since this bound holds for each choice of T_n (with card $T_n \leq 2^{2^n}$) we have proved

$$\mathrm{E}\sup_{t\in T} X_t \leq C\eta(T,d)\;, \tag{5}$$

where

$$\eta(T,d) := \inf \sup_{t\in T} \sum_{n\geq 1} 2^{n/2} d(t,T_n)\;,$$

for an infimum over all possible choices of the sets T_n. ((5) is called the *generic chaining* bound). How good is this bound? It is the best possible:

Theorem (The Majorizing measure theorem): *For Gaussian processes*,

$$\frac{1}{C}\eta(T,d) \leq \mathrm{E}\sup_{t\in T} X_t \leq C\eta(T,d)\;.$$

Moral?

Moral?

The majorizing measure theorem does not tell you how to find the sets T_n to get a good bound for a Gaussian process.

Moral?

The majorizing measure theorem does not tell you how to find the sets T_n to get a good bound for a Gaussian process. Rather, it tells you that there is no other way to get a good bound than finding these sets.

Moral?

The majorizing measure theorem does not tell you how to find the sets T_n to get a good bound for a Gaussian process. Rather, it tells you that there is no other way to get a good bound than finding these sets.
Finding these sets requires understanding the geometry of the situation.

Moral?

The majorizing measure theorem does not tell you how to find the sets T_n to get a good bound for a Gaussian process. Rather, it tells you that there is no other way to get a good bound than finding these sets.
Finding these sets requires understanding the geometry of the situation.
Still, the "generic chaining" bound completely explains the size of Gaussian processes.

Moral?

The majorizing measure theorem does not tell you how to find the sets T_n to get a good bound for a Gaussian process. Rather, it tells you that there is no other way to get a good bound than finding these sets.

Finding these sets requires understanding the geometry of the situation.

Still, the "generic chaining" bound completely explains the size of Gaussian processes.

Now the generic chaining bound is just a clever use of the union bound.

Moral?

The majorizing measure theorem does not tell you how to find the sets T_n to get a good bound for a Gaussian process. Rather, it tells you that there is no other way to get a good bound than finding these sets.

Finding these sets requires understanding the geometry of the situation.

Still, the "generic chaining" bound completely explains the size of Gaussian processes.

Now the generic chaining bound is just a clever use of the union bound.

HOW FAR DOES THIS GO?

Representation of Gaussian processes as random series

Representation of Gaussian processes as random series

Gaussian processes can always be represented as series $X_t = \sum_{i \geq 1} f_i(t) g_i$, where (g_i) is an i.i.d. standard Gaussian sequence.

Representation of Gaussian processes as random series

Gaussian processes can always be represented as series $X_t = \sum_{i \geq 1} f_i(t) g_i$, where (g_i) is an i.i.d. standard Gaussian sequence. Only the set of coefficients matters to understand $\sup_{t \in T} X_t$.

Representation of Gaussian processes as random series

Gaussian processes can always be represented as series $X_t = \sum_{i \geq 1} f_i(t) g_i$, where (g_i) is an i.i.d. standard Gaussian sequence. Only the set of coefficients matters to understand $\sup_{t \in T} X_t$. Therefore it is natural to consider this set of coefficients as the parameter.

Representation of Gaussian processes as random series

Gaussian processes can always be represented as series $X_t = \sum_{i \geq 1} f_i(t) g_i$, where (g_i) is an i.i.d. standard Gaussian sequence. Only the set of coefficients matters to understand $\sup_{t \in T} X_t$. Therefore it is natural to consider this set of coefficients as the parameter. That is we consider processes of the type

$$X_t = \sum_{i \geq 1} t_i g_i$$

where the parameter

$$t = (t_i)_{i \geq 1}$$

is a sequence and T is a set of such sequences.

Representation of Gaussian processes as random series

Gaussian processes can always be represented as series $X_t = \sum_{i \geq 1} f_i(t) g_i$, where (g_i) is an i.i.d. standard Gaussian sequence. Only the set of coefficients matters to understand $\sup_{t \in T} X_t$. Therefore it is natural to consider this set of coefficients as the parameter. That is we consider processes of the type

$$X_t = \sum_{i \geq 1} t_i g_i$$

where the parameter

$$t = (t_i)_{i \geq 1}$$

is a sequence and T is a set of such sequences.The canonical distance on T induced by the process (X_t) is then simply the ℓ^2 distance.

Representation of Gaussian processes as random series

Gaussian processes can always be represented as series $X_t = \sum_{i \geq 1} f_i(t) g_i$, where (g_i) is an i.i.d. standard Gaussian sequence. Only the set of coefficients matters to understand $\sup_{t \in T} X_t$. Therefore it is natural to consider this set of coefficients as the parameter. That is we consider processes of the type

$$X_t = \sum_{i \geq 1} t_i g_i$$

where the parameter

$$t = (t_i)_{i \geq 1}$$

is a sequence and T is a set of such sequences.The canonical distance on T induced by the process (X_t) is then simply the ℓ^2 distance.

The map $t \mapsto X_t$ is linear.

Representation of Gaussian processes as random series

Gaussian processes can always be represented as series $X_t = \sum_{i \geq 1} f_i(t) g_i$, where (g_i) is an i.i.d. standard Gaussian sequence. Only the set of coefficients matters to understand $\sup_{t \in T} X_t$. Therefore it is natural to consider this set of coefficients as the parameter. That is we consider processes of the type

$$X_t = \sum_{i \geq 1} t_i g_i$$

where the parameter

$$t = (t_i)_{i \geq 1}$$

is a sequence and T is a set of such sequences.The canonical distance on T induced by the process (X_t) is then simply the ℓ^2 distance.

The map $t \mapsto X_t$ is linear. THIS HAS PROFOUND CONSEQUENCES.

Digression: the hidden power of linearity

Digression: the hidden power of linearity

Some simple looking facts are mysterious: As a consequence of linearity,

Digression: the hidden power of linearity

Some simple looking facts are mysterious: As a consequence of linearity,

$$\mathsf{E} \sup_{t \in T} X_t = \mathsf{E} \sup_{t \in \operatorname{conv} T} X_t \,.$$

Digression: the hidden power of linearity

Some simple looking facts are mysterious: As a consequence of linearity,

$$\mathsf{E}\sup_{t\in T} X_t = \mathsf{E}\sup_{t\in \mathrm{conv}\, T} X_t \ .$$

As a consequence of the majorizing measure theorem, one obtain the following *geometrical* result about Hilbert space:

$$\eta(\mathrm{conv}\, T, d) \le C\eta(T, d) \ .$$

Digression: the hidden power of linearity

Some simple looking facts are mysterious: As a consequence of linearity,

$$\mathsf{E}\sup_{t\in T} X_t = \mathsf{E}\sup_{t\in \mathrm{conv}\,T} X_t\ .$$

As a consequence of the majorizing measure theorem, one obtain the following *geometrical* result about Hilbert space:

$$\eta(\mathrm{conv}\,T, d) \le C\eta(T, d)\ .$$

Potentially important problem: find a geometrical proof of this result.

Bernoulli Processes

Bernoulli Processes

It is natural to consider processes of the type $X_t = \sum_{i \geq 1} t_i \xi_i$ where $(\xi_i)_{i \geq 1}$ are independent r.v.s, and the parameter $t = (t_i)_{i \geq 1}$ is a sequence of coefficients.

Bernoulli Processes

It is natural to consider processes of the type $X_t = \sum_{i \geq 1} t_i \xi_i$ where $(\xi_i)_{i \geq 1}$ are independent r.v.s, and the parameter $t = (t_i)_{i \geq 1}$ is a sequence of coefficients.

In this case the map $t \mapsto X_t$ is linear.

Bernoulli Processes

It is natural to consider processes of the type $X_t = \sum_{i \geq 1} t_i \xi_i$ where $(\xi_i)_{i \geq 1}$ are independent r.v.s, and the parameter $t = (t_i)_{i \geq 1}$ is a sequence of coefficients.

In this case the map $t \mapsto X_t$ is linear.

The most important case is where $\xi_i = \varepsilon_i$ are Bernoulli r.v.s,

$$\mathrm{P}(\varepsilon_i = \pm 1) = \frac{1}{2}.$$

Bernoulli Processes

It is natural to consider processes of the type $X_t = \sum_{i \geq 1} t_i \xi_i$ where $(\xi_i)_{i \geq 1}$ are independent r.v.s, and the parameter $t = (t_i)_{i \geq 1}$ is a sequence of coefficients.

In this case the map $t \mapsto X_t$ is linear.

The most important case is where $\xi_i = \varepsilon_i$ are Bernoulli r.v.s,

$$\mathsf{P}(\varepsilon_i = \pm 1) = \frac{1}{2}.$$

A Bernoulli process is then a collection of r.v.s $X_t = \sum_{i \geq 1} t_i \varepsilon_i$, where the parameter t is a sequence $t = (t_i)_{i \geq 1}$.

Bernoulli Processes

It is natural to consider processes of the type $X_t = \sum_{i \geq 1} t_i \xi_i$ where $(\xi_i)_{i \geq 1}$ are independent r.v.s, and the parameter $t = (t_i)_{i \geq 1}$ is a sequence of coefficients.

In this case the map $t \mapsto X_t$ is linear.

The most important case is where $\xi_i = \varepsilon_i$ are Bernoulli r.v.s,

$$\mathsf{P}(\varepsilon_i = \pm 1) = \frac{1}{2}.$$

A Bernoulli process is then a collection of r.v.s $X_t = \sum_{i \geq 1} t_i \varepsilon_i$, where the parameter t is a sequence $t = (t_i)_{i \geq 1}$. They occur in many circumstances when using symmetrization techniques (e.g in the study of random Fourier series).

Two bounds for Bernoulli Processes.

If (g_i) are independent standard Gaussian r.v.s, the generic chaining bound relies on the inequality

$$P\left(\left|\sum_{i\geq 1} t_i g_i\right| \geq v\right) \leq 2\exp\left(-\frac{v^2}{2\sum_{i\geq 1} t_i^2}\right) .$$

Two bounds for Bernoulli Processes.

If (g_i) are independent standard Gaussian r.v.s, the generic chaining bound relies on the inequality

$$\mathrm{P}\Big(\Big|\sum_{i\geq 1} t_i g_i\Big| \geq v\Big) \leq 2\exp\Big(-\frac{v^2}{2\sum_{i\geq 1} t_i^2}\Big) \,.$$

This inequality remains true if one replace g_i by a random sign ε_i (The subgaussian inequality).

Two bounds for Bernoulli Processes.

If (g_i) are independent standard Gaussian r.v.s, the generic chaining bound relies on the inequality

$$\mathrm{P}\Big(\Big|\sum_{i\geq 1} t_i g_i\Big| \geq v\Big) \leq 2\exp\Big(-\frac{v^2}{2\sum_{i\geq 1} t_i^2}\Big) \ .$$

This inequality remains true if one replace g_i by a random sign ε_i (The subgaussian inequality). Consequently, since $\mathrm{E}X_{t_0} = 0$, for a Bernoulli process,

$$\mathrm{E}\sup_{t\in T} X_t = \mathrm{E}\sup_{t\in T}(X_t - X_{t_0}) \leq C\eta(T,d) \ ,$$

where d is the distance induced by ℓ^2.

Two bounds for Bernoulli Processes.

If (g_i) are independent standard Gaussian r.v.s, the generic chaining bound relies on the inequality

$$\mathsf{P}\Big(\Big|\sum_{i\geq 1} t_i g_i\Big| \geq v\Big) \leq 2\exp\Big(-\frac{v^2}{2\sum_{i\geq 1} t_i^2}\Big) .$$

This inequality remains true if one replace g_i by a random sign ε_i (The subgaussian inequality). Consequently, since $\mathsf{E}X_{t_0} = 0$, for a Bernoulli process,

$$\mathsf{E}\sup_{t\in T} X_t = \mathsf{E}\sup_{t\in T}(X_t - X_{t_0}) \leq C\eta(T, d) ,$$

where d is the distance induced by ℓ^2.

There is a completely different method to bound Bernoulli processes, namely

$$|X_t| = \Big|\sum_{i\geq 1} t_i\varepsilon_i\Big| \leq \sum_{i\geq 1} |t_i| .$$

Interpolation between bounds

Interpolation between bounds

Having two different methods to bound a Bernoulli process we can interpolate between them using linearity, i.e. that $X_{u+v} = X_u + X_v$.

Interpolation between bounds

Having two different methods to bound a Bernoulli process we can interpolate between them using linearity, i.e. that $X_{u+v} = X_u + X_v$. For two sets of sequences U and V, let

$$U + V = \{u + v \; ; \; u \in U, v \in V\} \; .$$

Interpolation between bounds

Having two different methods to bound a Bernoulli process we can interpolate between them using linearity, i.e. that $X_{u+v} = X_u + X_v$. For two sets of sequences U and V, let

$$U + V = \{u + v \; ; \; u \in U, v \in V\} \; .$$

Thus if $T \subset U + V$,

$$\sup_{t \in T} X_t \leq \sup_{t \in U+V} X_t \leq \sup_{t \in U} X_t + \sup_{t \in V} X_t$$

Interpolation between bounds

Having two different methods to bound a Bernoulli process we can interpolate between them using linearity, i.e. that $X_{u+v} = X_u + X_v$. For two sets of sequences U and V, let

$$U + V = \{u + v \; ; \; u \in U, v \in V\} \, .$$

Thus if $T \subset U + V$,

$$\sup_{t \in T} X_t \leq \sup_{t \in U+V} X_t \leq \sup_{t \in U} X_t + \sup_{t \in V} X_t$$

Consequently if $T \subset U + V$,

$$\mathrm{E} \sup_{t \in T} X_t \quad \leq \quad \mathrm{E} \sup_{t \in U} X_t + \mathrm{E} \sup_{t \in V} X_t$$

Interpolation between bounds

Having two different methods to bound a Bernoulli process we can interpolate between them using linearity, i.e. that $X_{u+v} = X_u + X_v$. For two sets of sequences U and V, let

$$U + V = \{u + v\ ;\ u \in U, v \in V\}\ .$$

Thus if $T \subset U + V$,

$$\sup_{t \in T} X_t \le \sup_{t \in U+V} X_t \le \sup_{t \in U} X_t + \sup_{t \in V} X_t$$

Consequently if $T \subset U + V$,

$$\begin{aligned} \mathrm{E} \sup_{t \in T} X_t &\le \mathrm{E} \sup_{t \in U} X_t + \mathrm{E} \sup_{t \in V} X_t \\ &\le C\eta(U, d) + \sup_{v \in V} \sum_{i \ge 1} |v_i| \end{aligned}$$

The Bernoulli Conjecture and the \$ 5000 prize

If $T \subset U + V$ then $\mathsf{E}\sup_{t\in T} X_t \le C\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i|$, so that

$$\mathsf{E}\sup_{t\in T} X_t \le \inf_{T\subset U+V}\left(C\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i|\right) .$$

The Bernoulli Conjecture and the \$ 5000 prize

If $T \subset U + V$ then $\mathsf{E} \sup_{t \in T} X_t \le C\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i|$, so that

$$\mathsf{E} \sup_{t \in T} X_t \le \inf_{T \subset U+V} \left(C\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i| \right) .$$

The Bernoulli Conjecture states that this bound can be reversed: given T we can find U and V with $T \subset U + V$ and

$$\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i| \le C \mathsf{E} \sup_{t \in T} X_t .$$

The Bernoulli Conjecture and the \$ 5000 prize

If $T \subset U + V$ then $\mathsf{E}\sup_{t\in T} X_t \le C\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i|$, so that

$$\mathsf{E}\sup_{t\in T} X_t \le \inf_{T\subset U+V}\left(C\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i|\right) .$$

The Bernoulli Conjecture states that this bound can be reversed: given T we can find U and V with $T \subset U + V$ and

$$\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i| \le C\mathsf{E}\sup_{t\in T} X_t \ .$$

The prize is for proving this.

The Bernoulli Conjecture and the \$ 5000 prize

If $T \subset U + V$ then $\mathsf{E} \sup_{t \in T} X_t \le C\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i|$, so that

$$\mathsf{E} \sup_{t \in T} X_t \le \inf_{T \subset U+V} \left(C\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i| \right) .$$

The Bernoulli Conjecture states that this bound can be reversed: given T we can find U and V with $T \subset U + V$ and

$$\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i| \le C \mathsf{E} \sup_{t \in T} X_t .$$

The prize is for proving this. Difficulty: The decomposition is not canonical in any way.

The Bernoulli Conjecture and the \$ 5000 prize

If $T \subset U + V$ then $\mathsf{E}\sup_{t\in T} X_t \le C\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i|$, so that

$$\mathsf{E}\sup_{t\in T} X_t \le \inf_{T\subset U+V}\left(C\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i|\right) .$$

The Bernoulli Conjecture states that this bound can be reversed: given T we can find U and V with $T \subset U + V$ and

$$\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i| \le C\mathsf{E}\sup_{t\in T} X_t .$$

The prize is for proving this. Difficulty: The decomposition is not canonical in any way. Caveat: to get the prize you have to solve the problem before I am too senile to understand the solution.

The Bernoulli Conjecture and the $ 5000 prize

If $T \subset U + V$ then $\mathsf{E}\sup_{t\in T} X_t \le C\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i|$, so that

$$\mathsf{E}\sup_{t\in T} X_t \le \inf_{T\subset U+V}\left(C\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i|\right) .$$

The Bernoulli Conjecture states that this bound can be reversed: given T we can find U and V with $T \subset U + V$ and

$$\eta(U,d) + \sup_{v\in V}\sum_{i\ge 1}|v_i| \le C\mathsf{E}\sup_{t\in T} X_t .$$

The prize is for proving this. Difficulty: The decomposition is not canonical in any way. Caveat: to get the prize you have to solve the problem before I am too senile to understand the solution.
My thanks to R. Latała for working on this question (and proving a beautiful partial result)

The Bernoulli Conjecture and the \$ 5000 prize

If $T \subset U + V$ then $\mathsf{E}\sup_{t \in T} X_t \le C\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i|$, so that

$$\mathsf{E}\sup_{t \in T} X_t \le \inf_{T \subset U+V} \Big(C\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i| \Big) \ .$$

The Bernoulli Conjecture states that this bound can be reversed: given T we can find U and V with $T \subset U + V$ and

$$\eta(U,d) + \sup_{v \in V} \sum_{i \ge 1} |v_i| \le C\mathsf{E}\sup_{t \in T} X_t \ .$$

The prize is for proving this. Difficulty: The decomposition is not canonical in any way. Caveat: to get the prize you have to solve the problem before I am too senile to understand the solution. My thanks to R. Latała for working on this question (and proving a beautiful partial result) (Rafal started long before there was a prize).

My Conjecture Generator

My Conjecture Generator

- Find an nice class of random series $X_t = \sum_{i \geq 1} t_i \xi_i$ where ξ_i are independent r.v.s, possibly with the same distribution.

My Conjecture Generator

- Find an nice class of random series $X_t = \sum_{i \geq 1} t_i \xi_i$ where ξ_i are independent r.v.s, possibly with the same distribution.
- Carefully collect all the methods you can think about to bound the supremum of a family $(X_t)_{t \in T}$ of such r.v.s

My Conjecture Generator

- Find an nice class of random series $X_t = \sum_{i \geq 1} t_i \xi_i$ where ξ_i are independent r.v.s, possibly with the same distribution.
- Carefully collect all the methods you can think about to bound the supremum of a family $(X_t)_{t \in T}$ of such r.v.s
- Think of all the ways you can to interpolate between these methods.

My Conjecture Generator

- Find an nice class of random series $X_t = \sum_{i \geq 1} t_i \xi_i$ where ξ_i are independent r.v.s, possibly with the same distribution.
- Carefully collect all the methods you can think about to bound the supremum of a family $(X_t)_{t \in T}$ of such r.v.s
- Think of all the ways you can to interpolate between these methods.
- Conjecture that the resulting bound is best possible.

My Conjecture Generator at work: Positive Selector Processes

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$\mathsf{P}(\xi_i = 1) = \delta \; ; \; \mathsf{P}(\xi_i = 0) = 1 - \delta \; .$$

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$\mathsf{P}(\xi_i = 1) = \delta \; ; \; \mathsf{P}(\xi_i = 0) = 1 - \delta \; .$$

Try to understand the processes $(X_t)_{t \in T}$ where $t = (t_i)_{i \geq 1}$, $t_i \geq 0$.

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$\mathsf{P}(\xi_i = 1) = \delta \; ; \; \mathsf{P}(\xi_i = 0) = 1 - \delta \; .$$

Try to understand the processes $(X_t)_{t \in T}$ where $t = (t_i)_{i \geq 1}$, $t_i \geq 0$.
To bound these:

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$P(\xi_i = 1) = \delta \; ; \; P(\xi_i = 0) = 1 - \delta \; .$$

Try to understand the processes $(X_t)_{t \in T}$ where $t = (t_i)_{i \geq 1}$, $t_i \geq 0$. To bound these:

- Linearity:

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$\mathsf{P}(\xi_i = 1) = \delta \; ; \; \mathsf{P}(\xi_i = 0) = 1 - \delta \; .$$

Try to understand the processes $(X_t)_{t \in T}$ where $t = (t_i)_{i \geq 1}$, $t_i \geq 0$. To bound these:

- Linearity: $\mathsf{E} \sup_{t \in \mathrm{conv}\, T} X_t = \mathsf{E} \sup_{t \in T} X_t$.

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$\mathrm{P}(\xi_i = 1) = \delta \; ; \; \mathrm{P}(\xi_i = 0) = 1 - \delta \; .$$

Try to understand the processes $(X_t)_{t \in T}$ where $t = (t_i)_{i \geq 1}$, $t_i \geq 0$. To bound these:

- Linearity: $\mathrm{E} \sup_{t \in \mathrm{conv}\, T} X_t = \mathrm{E} \sup_{t \in T} X_t$.
- Positivity:

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$\mathsf{P}(\xi_i = 1) = \delta \; ; \; \mathsf{P}(\xi_i = 0) = 1 - \delta \; .$$

Try to understand the processes $(X_t)_{t \in T}$ where $t = (t_i)_{i \geq 1}$, $t_i \geq 0$. To bound these:

- Linearity: $\mathsf{E} \sup_{t \in \mathrm{conv}\, T} X_t = \mathsf{E} \sup_{t \in T} X_t$.
- Positivity: If each element of T is smaller than an element of S, i.e.

$$\forall\, t \in T \; , \; \exists s \in S \; , \; \forall\, i \geq 1 \; , \; t_i \leq s_i$$

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$\mathsf{P}(\xi_i = 1) = \delta \; ; \; \mathsf{P}(\xi_i = 0) = 1 - \delta \; .$$

Try to understand the processes $(X_t)_{t \in T}$ where $t = (t_i)_{i \geq 1}$, $t_i \geq 0$. To bound these:

- Linearity: $\mathsf{E} \sup_{t \in \mathrm{conv}\, T} X_t = \mathsf{E} \sup_{t \in T} X_t$.
- Positivity: If each element of T is smaller than an element of S, i.e.

  $$\forall\, t \in T \;,\; \exists s \in S \;,\; \forall\, i \geq 1 \;,\; t_i \leq s_i$$

  Then $\mathsf{E} \sup_{t \in T} X_t \leq \mathsf{E} \sup_{t \in S} X_t$.

My Conjecture Generator at work: Positive Selector Processes

Consider a number $0 < \delta \ll 1$ and independent r.v.s $\xi_i \in \{0, 1\}$ with

$$\mathsf{P}(\xi_i = 1) = \delta \; ; \; \mathsf{P}(\xi_i = 0) = 1 - \delta \; .$$

Try to understand the processes $(X_t)_{t \in T}$ where $t = (t_i)_{i \geq 1}$, $t_i \geq 0$. To bound these:

- Linearity: $\mathsf{E} \sup_{t \in \mathrm{conv}\, T} X_t = \mathsf{E} \sup_{t \in T} X_t$.
- Positivity: If each element of T is smaller than an element of S, i.e.

  $$\forall\, t \in T \; , \; \exists s \in S \; , \; \forall\, i \geq 1 \; , \; t_i \leq s_i$$

  Then $\mathsf{E} \sup_{t \in T} X_t \leq \mathsf{E} \sup_{t \in S} X_t$.
- Combining the previous: $\mathsf{E} \sup_{t \in \mathrm{sol} S} X_t = \mathsf{E} \sup_{t \in S} X_t$ where $\mathrm{sol} S$ is the solid convex hull of S:

  $$\mathrm{sol} S = \{ t \; ; \; \exists s \in \mathrm{conv}\, S \; , \; \forall\, i \geq 1, t_i \leq s_i \} \; .$$

A substitute for chaining for positive selector processes

A substitute for chaining for positive selector processes

This is needed because is difficult to use chaining since the tails of the r.v.s X_t are complicated.

A substitute for chaining for positive selector processes

This is needed because is difficult to use chaining since the tails of the r.v.s X_t are complicated. We detail a possible use of the union bound.

A substitute for chaining for positive selector processes

This is needed because is difficult to use chaining since the tails of the r.v.s X_t are complicated. We detail a possible use of the union bound. Define

$$\mathcal{F}(S) = \inf\left\{A\,;\;\int_A^\infty \sum_{t\in S} \mathrm{P}(X_t \geq u)\mathrm{d}u \leq A\right\}.$$

A substitute for chaining for positive selector processes

This is needed because is difficult to use chaining since the tails of the r.v.s X_t are complicated. We detail a possible use of the union bound. Define

$$\mathcal{F}(S) = \inf\Big\{A\,;\,\int_A^\infty \sum_{t\in S} \mathrm{P}(X_t \ge u)\mathrm{d}u \le A\Big\}\,.$$

Notice that for $A \ge 0$

$$\mathrm{E}\sup_{t\in S} X_t = \int_0^\infty \mathrm{P}\Big(\sup_{t\in S} X_t \ge u\Big)\mathrm{d}u \le A + \int_A^\infty \sum_{t\in S} \mathrm{P}(X_t \ge u)\mathrm{d}u$$

A substitute for chaining for positive selector processes

This is needed because is difficult to use chaining since the tails of the r.v.s X_t are complicated. We detail a possible use of the union bound. Define

$$\mathcal{F}(S) = \inf\Big\{A\,;\ \int_A^\infty \sum_{t\in S} \mathrm{P}(X_t \geq u)\mathrm{d}u \leq A\Big\}\,.$$

Notice that for $A \geq 0$

$$\mathrm{E}\sup_{t\in S} X_t = \int_0^\infty \mathrm{P}\Big(\sup_{t\in S} X_t \geq u\Big)\mathrm{d}u \leq A + \int_A^\infty \sum_{t\in S} \mathrm{P}(X_t \geq u)\mathrm{d}u$$

and thus

$$\mathrm{E}\sup_{t\in S} X_t \leq 2\mathcal{F}(S).$$

Note: $T \subset S \not\Rightarrow \mathcal{F}(T) \leq \mathcal{F}(S)$

Wishful thinking at its best (its worse?)

Consequently, if $T \subset \mathrm{sol} S$ then

$$\mathsf{E} \sup_{t \in T} X_t \le \mathsf{E} \sup_{t \in S} X_t \le 2\mathcal{F}(S)$$

Wishful thinking at its best (its worse?)

Consequently, if $T \subset \mathrm{sol} S$ then

$$\mathsf{E} \sup_{t \in T} X_t \le \mathsf{E} \sup_{t \in S} X_t \le 2\mathcal{F}(S)$$

Thus

$$\mathsf{E} \sup_{t \in T} X_t \le 2 \inf\{\mathcal{F}(S) \; ; \; T \subset \mathrm{sol} S\} \; .$$

Wishful thinking at its best (its worse?)

Consequently, if $T \subset \mathrm{sol}S$ then

$$\mathsf{E}\sup_{t\in T} X_t \le \mathsf{E}\sup_{t\in S} X_t \le 2\mathcal{F}(S)$$

Thus

$$\mathsf{E}\sup_{t\in T} X_t \le 2\inf\{\mathcal{F}(S)\ ;\ T\subset \mathrm{sol}S\}\ .$$

Problem: Can this bound be reversed?

Wishful thinking at its best (its worse?)

Consequently, if $T \subset \mathrm{sol} S$ then

$$\mathsf{E} \sup_{t \in T} X_t \leq \mathsf{E} \sup_{t \in S} X_t \leq 2\mathcal{F}(S)$$

Thus

$$\mathsf{E} \sup_{t \in T} X_t \leq 2 \inf\{\mathcal{F}(S)\ ;\ T \subset \mathrm{sol} S\}\ .$$

Problem: Can this bound be reversed? i.e given T can one find S with $T \subset \mathrm{sol} S$ and $\mathcal{F}(S) \leq C \mathsf{E} \sup_{t \in T} X_T$?

Wishful thinking at its best (its worse?)

Consequently, if $T \subset \mathrm{sol} S$ then

$$\mathsf{E} \sup_{t \in T} X_t \leq \mathsf{E} \sup_{t \in S} X_t \leq 2\mathcal{F}(S)$$

Thus

$$\mathsf{E} \sup_{t \in T} X_t \leq 2 \inf\{\mathcal{F}(S) \ ; \ T \subset \mathrm{sol} S\} \ .$$

Problem: Can this bound be reversed? i.e given T can one find S with $T \subset \mathrm{sol} S$ and $\mathcal{F}(S) \leq C \mathsf{E} \sup_{t \in T} X_T$?

It should be very easy to disprove this! You *simply* have to invent a new method to bound these processes.

Wishful thinking at its best (its worse?)

Consequently, if $T \subset \mathrm{sol} S$ then

$$\mathsf{E}\sup_{t\in T} X_t \le \mathsf{E}\sup_{t\in S} X_t \le 2\mathcal{F}(S)$$

Thus

$$\mathsf{E}\sup_{t\in T} X_t \le 2\inf\{\mathcal{F}(S)\ ;\ T \subset \mathrm{sol} S\}\ .$$

Problem: Can this bound be reversed? i.e given T can one find S with $T \subset \mathrm{sol} S$ and $\mathcal{F}(S) \le C\mathsf{E}\sup_{t\in T} X_T$?

It should be very easy to disprove this! You *simply* have to invent a new method to bound these processes. Of course you can then make a new conjecture....

Will these problems will ever be solved?

Will these problems will ever be solved?

- They are possibly very difficult.

Will these problems will ever be solved?

- They are possibly very difficult. They may require a major effort from very good people.

Will these problems will ever be solved?

- They are possibly very difficult.They may require a major effort from very good people.
- At first the payoff appears little.

Will these problems will ever be solved?

- They are possibly very difficult.They may require a major effort from very good people.
- At first the payoff appears little.
- Still, arguably they are at a fundamental level

Will these problems will ever be solved?

- They are possibly very difficult.They may require a major effort from very good people.
- At first the payoff appears little.
- Still, arguably they are at a fundamental level

THANK YOU FOR YOUR ATTENTION.

Will these problems will ever be solved?

- They are possibly very difficult. They may require a major effort from very good people.
- At first the payoff appears little.
- Still, arguably they are at a fundamental level

THANK YOU FOR YOUR ATTENTION.

I hope I soon have to mail this check.