Weak Convergence and Empirical Processes

Chapter 2.2: Maximal Inequalities and Covering Numbers

Bart Eggen

Leiden University

October 21, 2019

What are we going to discuss?

1. Introduction
2. Orlicz norm
3. Finite Maximal Inequalities
4. Maximal Inequalities

What are maximal inequalities?

$$\left\| \max_{1 \le i \le m} X_i \right\| \le K \max_{1 \le i \le m} \|X_i\|$$

What are maximal inequalities?

$$\left\| \max_{1 \le i \le m} X_i \right\| \le K \max_{1 \le i \le m} \|X_i\|$$

Can we do the same for an infinite amount of random variables?

Orlicz norm

Definition (Orlicz norm)

Let ψ be a nondecreasing, convex function with $\psi(0) = 0$ and X a random variable. Then the *Orlicz norm* $\|X\|_\psi$ is defined as

$$\|X\|_\psi = \inf\left\{C > 0 : E\psi\left(\frac{|X|}{C}\right) \leq 1\right\}.$$

Orlicz norm

Definition (Orlicz norm)

Let ψ be a nondecreasing, convex function with $\psi(0) = 0$ and X a random variable. Then the *Orlicz norm* $\|X\|_\psi$ is defined as

$$\|X\|_\psi = \inf\left\{C > 0 : E\psi\left(\frac{|X|}{C}\right) \leq 1\right\}.$$

Definition (Norm)

Given a vector space V over $\mathbb{R}$, a *norm* on V is a function $\|\cdot\| : V \to \mathbb{R}_+$ with the following properties:
For all $a \in \mathbb{R}$ and all $\mathbf{u}, \mathbf{v} \in V$,

1) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.
2) $\|a\mathbf{v}\| = |a| \cdot \|\mathbf{v}\|$.
3) If $\|\mathbf{v}\| = 0$, then $\mathbf{v} = 0$.

Orlicz norm

Definition (Orlicz norm)

Let ψ be a nonzero, nondecreasing, convex function with $\psi(0) = 0$ and X a random variable. Then the *Orlicz norm* $\|X\|_\psi$ is defined as

$$\|X\|_\psi = \inf\left\{C > 0 : E\psi\left(\frac{|X|}{C}\right) \leq 1\right\}.$$

Definition (Norm)

Given a vector space V over $\mathbb{R}$, a *norm* on V is a function $\|\cdot\| : V \to \mathbb{R}_+$ with the following properties:
For all $a \in \mathbb{R}$ and all $\mathbf{u}, \mathbf{v} \in V$,

1) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.
2) $\|a\mathbf{v}\| = |a| \cdot \|\mathbf{v}\|$.
3) If $\|\mathbf{v}\| = 0$, then $\mathbf{v} = 0$.

Examples

Example

For $p \geq 1$, let $\psi(x) = x^p$. Then for any random variable X

$$\|X\|_\psi = \inf\left\{C > 0 : E\left(\frac{|X|}{C}\right)^p \leq 1\right\}$$

Examples

Example

For $p \geq 1$, let $\psi(x) = x^p$. Then for any random variable X

$$\|X\|_\psi = \inf\left\{C > 0 : E\left(\frac{|X|}{C}\right)^p \leq 1\right\} = \|X\|_p$$

Examples

Example

For $p \geq 1$, let $\psi(x) = x^p$. Then for any random variable X

$$\|X\|_\psi = \inf\left\{C > 0 : E\left(\frac{|X|}{C}\right)^p \leq 1\right\} = \|X\|_p$$

Example

For $p \geq 1$, let $\psi(x) = e^{x^p} - 1$

Examples

Example

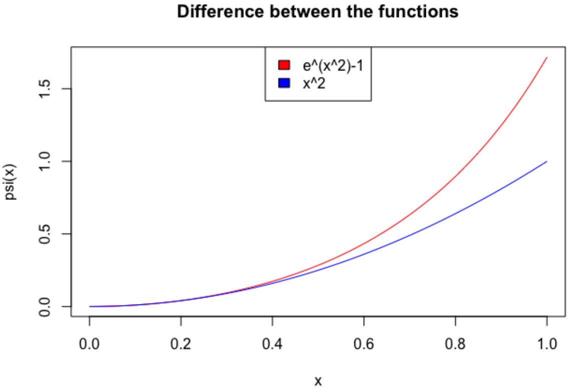
For $p \geq 1$, let $\psi(x) = x^p$. Then for any random variable X

$$\|X\|_\psi = \inf\left\{C > 0 : E\left(\frac{|X|}{C}\right)^p \leq 1\right\} = \|X\|_p$$

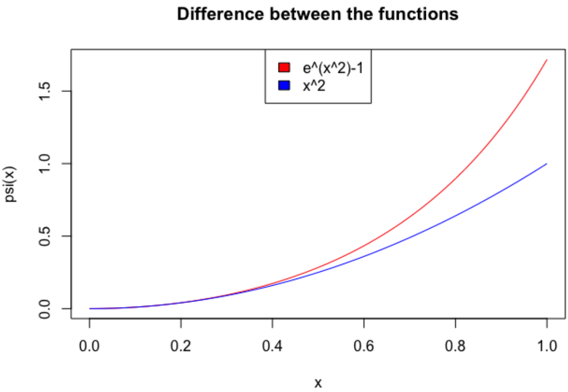
Example

For $p \geq 1$, let $\psi_p(x) = e^{x^p} - 1$

Examples


Difference between the functions
e^(x^2)-1
x^2
psi(x)
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0
x

Examples



Inequality

For all $p \geq 1$ we have

$$\|X\|_p \leq \|X\|_{\psi_p}$$

Tails of distributions

Tail bound

Let X be a random variable and suppose $\|X\|_\psi$ exists. Then for any $x \in \mathbb{R}$:

$$
\begin{aligned}
P(|X| > x) &\leq P\left(\psi(|X|/\|X\|_\psi) \geq \psi(x/\|X\|_\psi)\right) \\
&\leq \frac{E\psi(|X|/\|X\|_\psi)}{\psi(x/\|X\|_\psi)} \\
&\leq \frac{1}{\psi(x/\|X\|_\psi)}
\end{aligned}
$$

Tails of distributions

Tail bound

Let X be a random variable and suppose $\|X\|_\psi$ exists. Then for any $x \in \mathbb{R}$:

$$
\begin{aligned}
P(|X| > x) &\leq P\left(\psi(|X|/\|X\|_\psi) \geq \psi(x/\|X\|_\psi)\right) \\
&\leq \frac{E\psi(|X|/\|X\|_\psi)}{\psi(x/\|X\|_\psi)} \\
&\leq \frac{1}{\psi(x/\|X\|_\psi)}
\end{aligned}
$$

Taking ψ_p

$$P(|X| > x) \lesssim e^{-Cx^p}$$

Tails of distributions

Lemma 2.2.1.

Let X be a random variable with $P(|X| > x) \leq Ke^{-Cx^p}$ for every x, for constants K and C, and for $p \geq 1$. Then its Orlicz norm satisfies $\|X\|_{\psi_p} \leq ((1+K)/C)^{1/p}$.

Tails of distributions

Lemma 2.2.1.

Let X be a random variable with $P(|X| > x) \leq Ke^{-Cx^p}$ for every x, for constants K and C, and for $p \geq 1$. Then its Orlicz norm satisfies $\|X\|_{\psi_p} \leq ((1+K)/C)^{1/p}$.

Proof Idea

Our goal is to find the smallest D' such that

$$E\psi_p\left(\frac{|X|}{D'}\right) = E\left(e^{D|X|^p} - 1\right) \leq 1.$$

Tails of distributions

Lemma 2.2.1.

Let X be a random variable with $P(|X| > x) \leq Ke^{-Cx^p}$ for every x, for constants K and C, and for $p \geq 1$. Then its Orlicz norm satisfies $\|X\|_{\psi_p} \leq ((1+K)/C)^{1/p}$.

Proof Idea

Our goal is to find the smallest D' such that

$$E\psi_p\left(\frac{|X|}{D'}\right) = E\left(e^{D|X|^p} - 1\right) \leq 1.$$

Idea: write

$$e^{D|X|^p} - 1 = \int_0^{|X|^p} De^{Ds} ds$$

and use Fubini's theorem.

Finite maximal inequality

$$\left\| \max_{1 \leq i \leq m} X_i \right\|_{\psi}$$

Finite maximal inequality

$$\begin{aligned}
\left\| \max_{1\le i\le m} X_i \right\|_p &= \left(E \max_{1\le i\le m} |X_i|^p \right)^{1/p} \\
&\le \left(E \sum_{i=1}^{m} |X_i|^p \right)^{1/p} \\
&\le \left(m \cdot \max_{1\le i\le m} E|X_i|^p \right)^{1/p} \\
&= m^{1/p} \cdot \max_{1\le i\le m} \|X_i\|_p
\end{aligned}$$

Finite maximal inequality

$$\left\| \max_{1 \le i \le m} X_i \right\|_p \le m^{1/p} \cdot \max_{1 \le i \le m} \|X_i\|_p$$

Finite maximal inequality

$$\left\| \max_{1\le i\le m} X_i \right\|_p \le \psi^{-1}(m) \cdot \max_{1\le i\le m} \|X_i\|_p$$

Finite maximal inequality

Lemma 2.2.2.

Let ψ be a convex, nondecreasing, nonzero function with $\psi(0) = 0$ and

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

for some constant c . Then, for any random variables $X_1, \dots, X_m$,

$$\left\| \max_{1\leq i\leq m} X_i \right\|_\psi \leq K\psi^{-1}(m) \max_{1\leq i\leq m} \|X_i\|_\psi,$$

for a constant K depending only on ψ.

Finite maximal inequality

Lemma 2.2.2.

Let ψ be a convex, nondecreasing, nonzero function with $\psi(0) = 0$ and

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

for some constant c . Then, for any random variables $X_1, \dots, X_m$,

$$\left\| \max_{1\le i\le m} X_i \right\|_\psi \le K\psi^{-1}(m) \max_{1\le i\le m} \|X_i\|_\psi,$$

for a constant K depending only on ψ.

Proof Lemma 2.2.2.

$$\left\|\max X_i\right\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

1) Suppose there exists a constant c such that

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

Proof Lemma 2.2.2.

$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$\psi\left(\frac{\max |X_i|}{Cy}\right) \leq \max \psi\left(\frac{|X_i|}{Cy}\right)$$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$\psi\left(\frac{\max |X_i|}{Cy}\right) \leq \max \psi\left(\frac{|X_i|}{Cy}\right)$$

$$\leq \max\left[\frac{\psi(c|X_i|/C)}{\psi(y)} \mathbb{1}\left\{\frac{|X_i|}{Cy} \geq 1\right\} + \psi\left(\frac{|X_i|}{Cy}\right) \mathbb{1}\left\{\frac{|X_i|}{Cy} < 1\right\}\right]$$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$\psi\left(\frac{\max |X_i|}{Cy}\right) \leq \max \psi\left(\frac{|X_i|}{Cy}\right)$$

$$\leq \max\left[\frac{\psi(c|X_i|/C)}{\psi(y)} \mathbb{1}\left\{\frac{|X_i|}{Cy} \geq 1\right\} + \psi\left(\frac{|X_i|}{Cy}\right) \mathbb{1}\left\{\frac{|X_i|}{Cy} < 1\right\}\right]$$

Proof Lemma 2.2.2.

$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$\psi\left(\frac{\max |X_i|}{Cy}\right) \leq \max \psi\left(\frac{|X_i|}{Cy}\right)$$

$$\leq \max\left[\frac{\psi(c|X_i|/C)}{\psi(y)}\right] + \psi(1)$$

Proof Lemma 2.2.2.

$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$\psi\left(\frac{\max |X_i|}{Cy}\right) \leq \sum \frac{\psi(c|X_i|/C)}{\psi(y)} + \psi(1)$$

$$C = c \max \|X_i\|_\psi$$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$E\left[\psi\left(\frac{\max |X_i|}{cy \max \|X_i\|_\psi}\right)\right] \leq \sum \frac{E\psi(|X_i|/\max \|X_i\|_\psi)}{\psi(y)} + \psi(1)$$

Proof Lemma 2.2.2.

$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$E\left[\psi\left(\frac{\max |X_i|}{cy \max \|X_i\|_\psi}\right)\right] \leq \sum \frac{E\psi(|X_i|/\|X_i\|_\psi)}{\psi(y)} + \psi(1)$$

Proof Lemma 2.2.2.

$\|\max X_i\|_\psi \leq K\psi^{-1}(m)\max\|X_i\|_\psi$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$E\left[\psi\left(\frac{\max|X_i|}{cy\max\|X_i\|_\psi}\right)\right] \leq \frac{m}{\psi(y)} + \psi(1)$$

Proof Lemma 2.2.2.

$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$E\left[\psi\left(\frac{\max |X_i|}{cy \max \|X_i\|_\psi}\right)\right] \leq \frac{m}{\psi(y)} + \psi(1)$$

$\psi(1) \leq \frac{1}{2}$ $\qquad$ $y = \psi^{-1}(2m)$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1 \text{ and } \psi(1) \leq 1/2$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$E\left[\psi\left(\frac{\max |X_i|}{c\psi^{-1}(2m)\max \|X_i\|_\psi}\right)\right] \leq 1$$

Proof Lemma 2.2.2.

$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1 \text{ and } \psi(1) \leq 1/2$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$\|\max |X_i|\| \leq \psi^{-1}(2m) c \max \|X_i\|_\psi$$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1 \text{ and } \psi(1) \leq 1/2$$

1) Then also $\psi(x/y) \leq \psi(cx)/\psi(y)$ for all $x \geq y \geq 1$

$$\| \max |X_i|\| \leq 2\psi^{-1}(m) c \max \|X_i\|_\psi$$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

1) OK when

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1 \text{ and } \psi(1) \leq 1/2$$

Proof Lemma 2.2.2.

$$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$$

Assumption

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

1) OK when

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1 \text{ and } \psi(1) \leq 1/2$$

2) There exist $\sigma \leq 1$ and $\tau > 0$ such that $\phi(x) = \sigma\psi(\tau x)$

Proof Lemma 2.2.2.

$\|\max X_i\|_\psi \leq K\psi^{-1}(m) \max \|X_i\|_\psi$

Assumption

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

1) OK when

$$\psi(x)\psi(y) \leq \psi(cxy) \text{ for all } x, y \geq 1 \text{ and } \psi(1) \leq 1/2$$

2) There exist $\sigma \leq 1$ and $\tau > 0$ such that $\phi(x) = \sigma\psi(\tau x)$

$$\|X\|_\psi \leq \frac{\|X\|_\phi}{\sigma\tau} \leq \frac{\|X\|_\psi}{\sigma}$$

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$?

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK
- $\psi_p(x) = e^{x^p} - 1$?

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK
- $\psi_p(x) = e^{x^p} - 1$? OK

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK
- $\psi_p(x) = e^{x^p} - 1$? OK
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 \cdot \psi_2$ (pointwise)?

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK
- $\psi_p(x) = e^{x^p} - 1$? OK
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 \cdot \psi_2$ (pointwise)? OK, take $c = \max\{c_1, c_2\}$.

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK
- $\psi_p(x) = e^{x^p} - 1$? OK
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 \cdot \psi_2$ (pointwise)? OK, take $c = \max\{c_1, c_2\}$.
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 + \psi_2$?

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK
- $\psi_p(x) = e^{x^p} - 1$? OK
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 \cdot \psi_2$ (pointwise)? OK, take $c = \max\{c_1, c_2\}$.
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 + \psi_2$? ?

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK
- $\psi_p(x) = e^{x^p} - 1$? OK
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 \cdot \psi_2$ (pointwise)? OK, take $c = \max\{c_1, c_2\}$.
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 + \psi_2$? ?
- For $a \in \mathbb{R}_{>0}$, and ψ satisfying the condition: $a\psi$?

Assumption

What does this mean?

$$\limsup_{x,y\to\infty} \frac{\psi(x)\psi(y)}{\psi(cxy)} < \infty$$

- $\psi(x) = x^p$? OK
- $\psi_p(x) = e^{x^p} - 1$? OK
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 \cdot \psi_2$ (pointwise)? OK, take $c = \max\{c_1, c_2\}$.
- For ψ_1, ψ_2 satisfying the condition: $\psi_1 + \psi_2$? ?
- For $a \in \mathbb{R}_{>0}$, and ψ satisfying the condition: $a\psi$? OK

Maximal Inequality

$$\left\| \max_{1 \le i \le m} X_i \right\|_{\psi} \le K \psi^{-1}(m) \max_{1 \le i \le m} \|X_i\|_{\psi}$$

Maximal Inequality

$$\left\| \sup_{t \in T} |X_t| \right\|_{\psi} \leq ?$$

Covering numbers

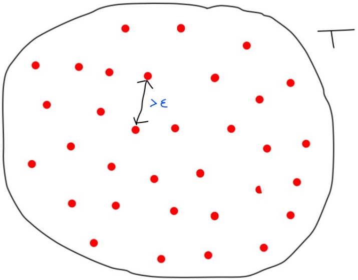
Definition

Let (T, d) be an arbitrary semi-metric space. Then the covering number $N(\epsilon, d)$ is the minimal number of balls of radius ϵ needed to cover T. Call a collection of points ϵ-*separated* if the distance between each pair of points is strictly larger than ϵ. The packing number $D(\epsilon, d)$ is the maximum number of ϵ-separated points in T.

Covering numbers

Definition

Let (T, d) be an arbitrary semi-metric space. Then the covering number $N(\epsilon, d)$ is the minimal number of balls of radius ϵ needed to cover T. Call a collection of points ϵ-*separated* if the distance between each pair of points is strictly larger than ϵ. The packing number $D(\epsilon, d)$ is the maximum number of ϵ-separated points in T.



Maximal Inequalities

Theorem

Let ψ be a convex, nondecreasing, nonzero function with $\psi(0) = 0$ and $\limsup_{x,y\to\infty} \psi(x)\psi(y)/\psi(cxy) < \infty$, for some constant c. Let $\{X_t : t \in T\}$ be a separable stochastic process with

$$\|X_s - X_t\|_\psi \leq C\, d(s,t), \qquad \text{for every } s,t,$$

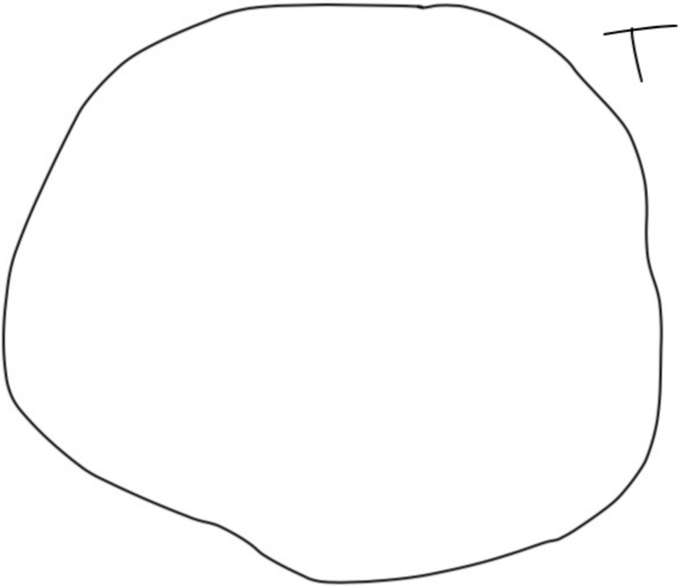
for some semimetric d on T and a constant C. Then, for any $\eta, \delta > 0$,

$$\Big\| \sup_{d(s,t)\leq\delta} |X_s - X_t| \Big\|_\psi \leq K \left[\int_0^\eta \psi^{-1}\left(D(\epsilon, d)\right) d\epsilon + \delta \psi^{-1}\left(D^2(\eta, d)\right)\right],$$

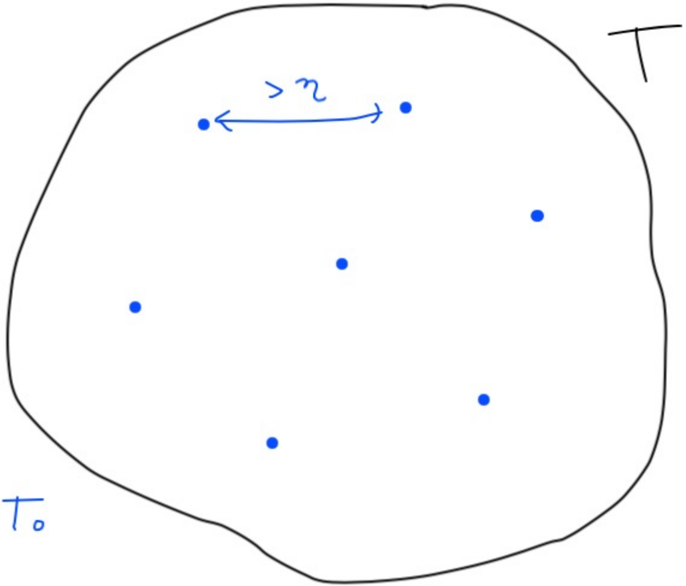
for a constant K depending on ψ and C only.

Separable may be understood in the sense that $\sup_{d(s,t)<\delta} |X_s - X_t|$ remains almost surely the same if the index set T is replaced by a suitable countable subset.

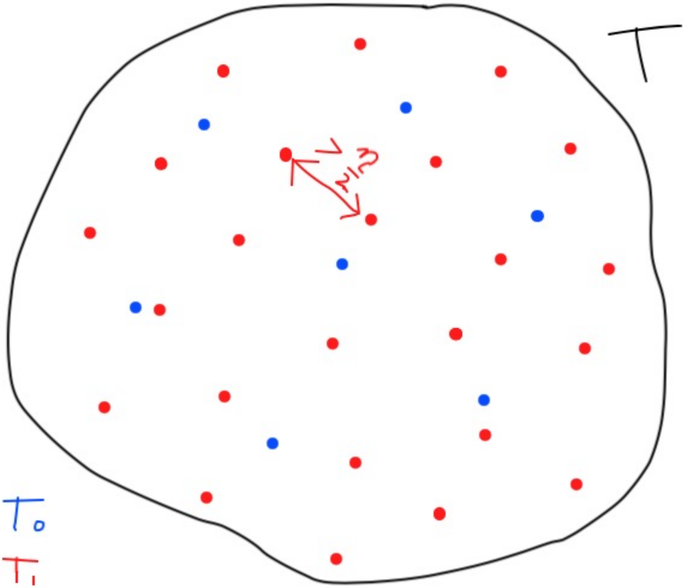
Chaining



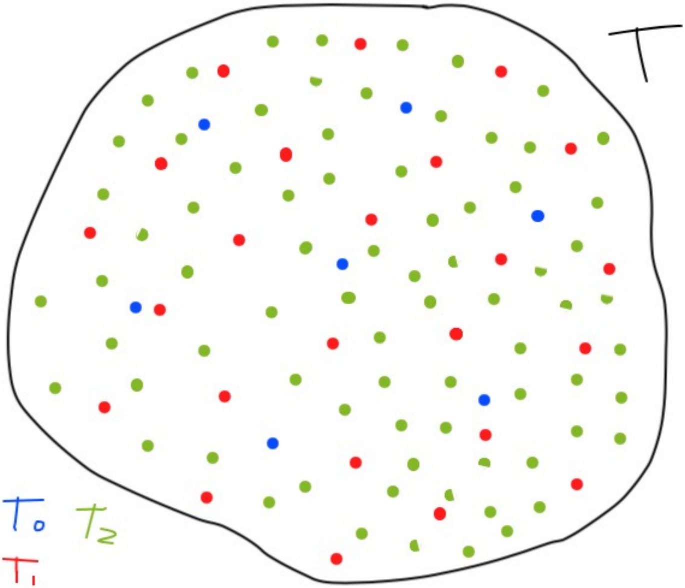
Chaining



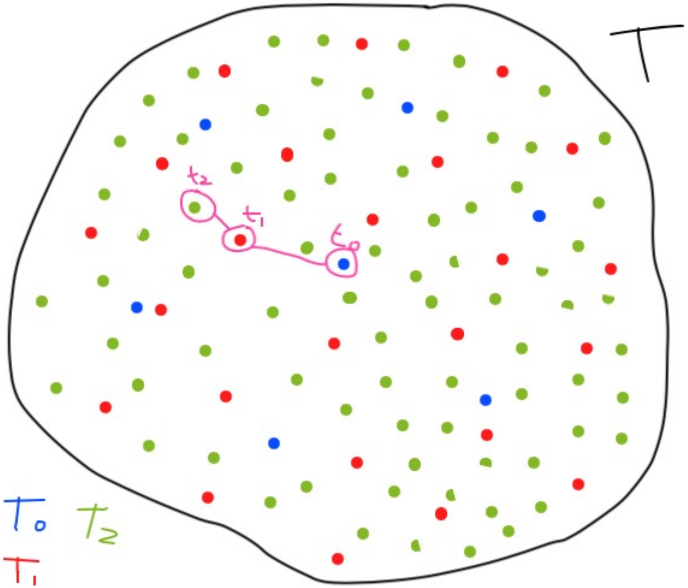
Chaining



Chaining



Chaining



Chaining

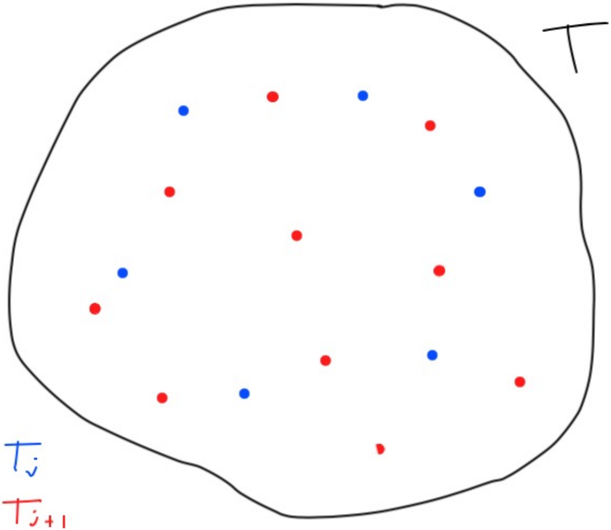
- For every $s, t \in T_j$, we have $d(s,t) > \eta 2^{-j}$
- The T_j are maximal, meaning no point can be added without destroying the validity of the inequality
- $|T_j| \leq D(\eta 2^{-j}, d)$
- Every $t_{j+1} \in T_{j+1}$ is linked to a unique $t_j \in T_j$ such that $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

Chaining

For every t_{j+1}, does there exist a t_j with $d(t_j, t_{j+1}) \leq \eta 2^{-j}$?

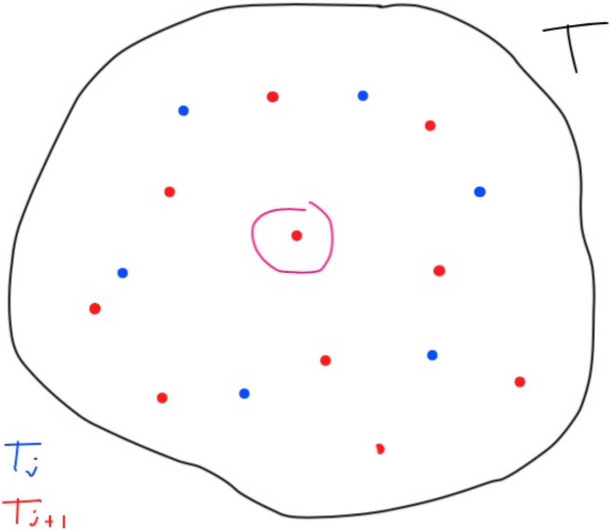
Chaining

For every t_{j+1}, does there exist a t_j with $d(t_j, t_{j+1}) \leq \eta 2^{-j}$?



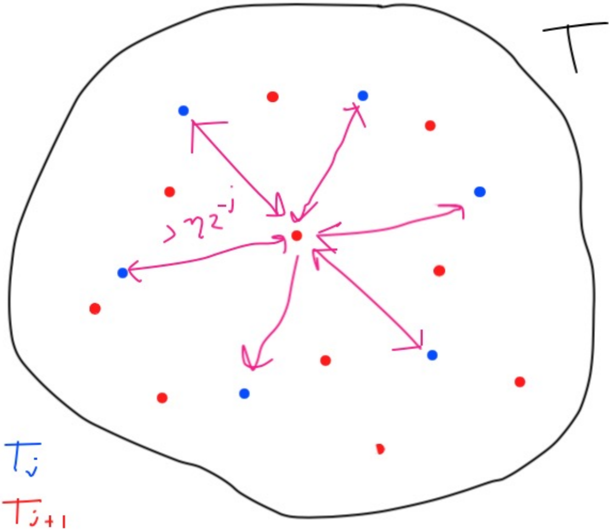
Chaining

For every t_{j+1}, does there exist a t_j with $d(t_j, t_{j+1}) \leq \eta 2^{-j}$?



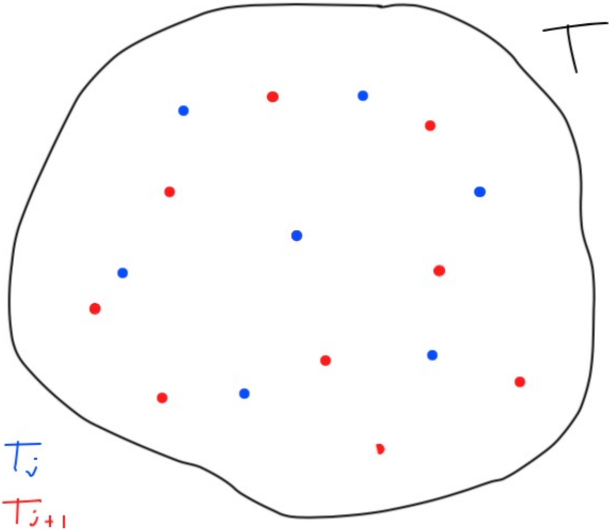
Chaining

For every t_{j+1}, does there exist a t_j with $d(t_j, t_{j+1}) \leq \eta 2^{-j}$?



Chaining

For every t_{j+1}, does there exist a t_j with $d(t_j, t_{j+1}) \leq \eta 2^{-j}$?



Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

Let $s_{k+1}, t_{k+1} \in T_{k+1}$, then we have two chains $t_{k+1}, t_k, \ldots, t_0$ and $s_{k+1}, s_k, \ldots, s_0$.

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

Let $s_{k+1}, t_{k+1} \in T_{k+1}$, then we have two chains $t_{k+1}, t_k, \ldots, t_0$ and $s_{k+1}, s_k, \ldots, s_0$.

$$\begin{aligned}
|(X_{s_{k+1}} - X_{s_0}) - (X_{t_{k+1}} - X_{t_0})| &= \left| \sum_{j=0}^{k} (X_{s_{j+1}} - X_{s_j}) - \sum_{j=0}^{k} (X_{t_{j+1}} - X_{t_j}) \right| \\
&\leq \sum_{j=0}^{k} |X_{s_{j+1}} - X_{s_j}| + |X_{t_{j+1}} - X_{t_j}| \\
&\leq 2 \sum_{j=0}^{k} \max |X_u - X_v|,
\end{aligned}$$

where for every j the maximum is taken over all links (u, v) from T_{j+1} to T_j. At most $|T_{j+1}| \leq D(\eta 2^{-j-1}, d)$ links!

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

$$|(X_{s_{k+1}} - X_{s_0}) - (X_{t_{k+1}} - X_{t_0})| \leq 2 \sum_{j=0}^{k} \max |X_u - X_v|$$

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

$$\max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \leq 2 \sum_{j=0}^{k} \max |X_u - X_v|$$

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_{\psi} \leq 2 \sum_{j=0}^{k} \| \max |X_u - X_v| \|_{\psi}$$

Chaining

- $|T_j| \le D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \le \eta 2^{-j}$

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_\psi \le 2 \sum_{j=0}^{k} \|\max |X_u - X_v|\|_\psi$$

$$\left\| \max_{1 \le i \le m} X_i \right\|_\psi \le K \psi^{-1}(m) \max_{1 \le i \le m} \|X_i\|_\psi$$

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_\psi \leq$$

$$2 \sum_{j=0}^{k} K \psi^{-1} \left(D(\eta 2^{-j-1}, d) \right) \max \|X_u - X_v\|_\psi$$

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_\psi \leq$$

$$2 \sum_{j=0}^{k} K \psi^{-1} \left(D(\eta 2^{-j-1}, d) \right) \max \|X_u - X_v\|_\psi$$

$$\|X_s - X_t\|_\psi \leq C\, d(s,t), \text{ for every } s, t$$

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_\psi \leq$$

$$2KC \sum_{j=0}^{k} \psi^{-1} \Big(D(\eta 2^{-j-1}, d) \Big) \max d(u, v)$$

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_\psi \leq$$

$$2KC \sum_{j=0}^{k} \psi^{-1}\Big(D(\eta 2^{-j-1}, d)\Big) \max d(u, v)$$

Chaining

- $|T_j| \leq D(\eta 2^{-j}, d)$
- $d(t_j, t_{j+1}) \leq \eta 2^{-j}$

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_\psi \leq K \sum_{j=0}^{k} \psi^{-1} \left(D(\eta 2^{-j-1}, d) \right) \eta 2^{-j}$$

Chaining

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_\psi \leq 2K \sum_{j=0}^{k} \psi^{-1}\left(D(\eta 2^{-j-1}, d) \right) \eta 2^{-j-1}$$

Chaining

$$\left\| \max_{s,t \in T_{k+1}} |(X_s - X_{s_0}) - (X_t - X_{t_0})| \right\|_\psi \leq 2K \sum_{j=0}^{k} \psi^{-1}\left(D(\eta 2^{-j-1}, d)\right) \eta 2^{-j-1}$$
$$\leq 2K \int_0^{k+1} \psi^{-1}\left(D(\eta 2^{-j-1}, d)\right) \eta 2^{-j-}$$

Chaining

$$
\begin{aligned}
\left\| \max_{s,t \in T_{k+1}} \left| (X_s - X_{s_0}) - (X_t - X_{t_0}) \right| \right\|_{\psi} &\leq 2K \sum_{j=0}^{k} \psi^{-1}\left(D(\eta 2^{-j-1}, d) \right) \eta 2^{-j-1} \\
&\leq 2K \int_0^{\infty} \psi^{-1}\left(D(\eta 2^{-j-1}, d) \right) \eta 2^{-j-1} \\
&\leq \frac{2}{\log 2} K \int_0^{\frac{\eta}{2}} \psi^{-1}\left(D(\epsilon, d) \right) d\epsilon
\end{aligned}
$$

Chaining

$$\left\| \max_{s,t \in T_{k+1}} \left| (X_s - X_{s_0}) - (X_t - X_{t_0}) \right| \right\|_\psi \leq 2K \sum_{j=0}^{k} \psi^{-1} \Big(D(\eta 2^{-j-1}, d) \Big) \eta 2^{-j-1}$$

$$\leq 2K \int_0^\infty \psi^{-1} \Big(D(\eta 2^{-j-1}, d) \Big) \eta 2^{-j-1}$$

$$\leq 4K \int_0^\eta \psi^{-1} \left(D(\epsilon, d) \right) d\epsilon$$

Chaining

We want to bound

$$\left\| \sup_{d(s,t)\leq\delta} |X_s - X_t| \right\|_\psi$$

Chaining

By $\{X_t : t \in T\}$ being separable, we can bring it to

$$\left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_s - X_t| \right\|_\psi ,$$

letting $k \to \infty$.

Chaining

By $\{X_t : t \in T\}$ being separable, we can bring it to

$$\left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_s - X_t| \right\|_\psi ,$$

letting $k \to \infty$. Note that we have

$$\begin{aligned} |X_s - X_t| &= |(X_s - X_{s_0}) - (X_t - X_{t_0}) + X_{s_0} - X_{t_0}| \\ &\leq |(X_s - X_{s_0}) - (X_t - X_{t_0})| + |X_{s_0} - X_{t_0}| \end{aligned}$$

Chaining

Two terms to bound

$$\left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} \left|(X_s - X_{s_0}) - (X_t - X_{t_0})\right| \right\|_{\psi}, \qquad \left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} \left|X_{s_0} - X_{t_0}\right| \right\|_{\psi}$$

Chaining

Two terms to bound

$$4K\int_0^{\eta}\psi^{-1}(D(\epsilon,d))d\epsilon \qquad \left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_{s_0}-X_{t_0}| \right\|_{\psi}$$

Chaining

Two terms to bound

$$4K\int_0^\eta \psi^{-1}(D(\epsilon, d))d\epsilon \qquad\qquad \left\| \max_{\substack{s,t,\in T_{k+1}\\ d(s,t)\le\delta}} |X_{s_0} - X_{t_0}| \right\|_\psi$$

Circular argument: For every pair of endpoints s_0, t_0 of chains starting at two points $s, t \in T_{k+1}$ with $d(s,t) \le \delta$, choose exactly one pair $s_{k+1}, t_{k+1} \in T_{k+1}$ with $d(s_{k+1}, t_{k+1}) \le \delta$, whose chains end at s_0, t_0. At most $D(\eta, d)^2$ pairs!

Chaining

Two terms to bound

$$4K\int_0^\eta \psi^{-1}(D(\epsilon, d))d\epsilon \qquad \left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_{s_0} - X_{t_0}| \right\|_\psi$$

Circular argument: For every pair of endpoints s_0, t_0 of chains starting at two points $s, t \in T_{k+1}$ with $d(s,t) \leq \delta$, choose exactly one pair $s_{k+1}, t_{k+1} \in T_{k+1}$ with $d(s_{k+1}, t_{k+1}) \leq \delta$, whose chains end at s_0, t_0. At most $D(\eta, d)^2$ pairs! Then we have

$$|X_{s_0} - X_{t_0}| \leq |(X_{s_{k+1}} - X_{s_0}) - (X_{t_{k+1}} - X_{t_0})| + |X_{s_{k+1}} - X_{t_{k+1}}|$$

Chaining

$$\left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_s - X_t| \right\|_\psi \leq 8K \int_0^\eta \psi^{-1}(D(\epsilon, d))d\epsilon + \left\| \max |X_{s_{k+1}} - X_{t_{k+1}}| \right\|_\psi ,$$

where the maximum is taken over the pairs $s_{k+1}, t_{k+1} \in T_{k+1}$ uniquely attached to the pairs s_0, t_0 as before (at most $D(\eta, d)^2$).

Chaining

$$\left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_s - X_t| \right\|_\psi \leq 8K \int_0^\eta \psi^{-1}(D(\epsilon, d))d\epsilon + \left\| \max |X_{s_{k+1}} - X_{t_{k+1}}| \right\|_\psi ,$$

where the maximum is taken over the pairs $s_{k+1}, t_{k+1} \in T_{k+1}$ uniquely attached to the pairs s_0, t_0 as before (at most $D(\eta, d)^2$).

Lemma 2.2.2

$$\left\| \max_{1\leq i\leq m} X_i \right\| \leq K\psi^{-1}(m) \max_i \|X_i\|_\psi$$

Chaining

$$\left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_s - X_t| \right\|_\psi \leq 8K \int_0^\eta \psi^{-1}(D(\epsilon, d))d\epsilon + K\psi^{-1}(D(\eta, d)^2)$$

$$\cdot \max \|X_{s_{k+1}} - X_{t_{k+1}}\|_\psi$$

Chaining

$$\left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_s - X_t| \right\|_\psi \leq 8K \int_0^\eta \psi^{-1}(D(\epsilon, d))d\epsilon + K\psi^{-1}(D(\eta, d)^2) \cdot C\delta$$

Chaining

$$\left\| \max_{\substack{s,t,\in T_{k+1} \\ d(s,t)\leq\delta}} |X_s - X_t| \right\|_\psi \leq K \left[\int_0^\eta \psi^{-1}(D(\epsilon, d))d\epsilon + \delta\psi^{-1}(D(\eta, d)^2) \right].$$

Let $k \to \infty$ to finish the proof.

General bound

Corollary 2.2.5

The constant K can be chosen such that

$$\left\| \sup_{s,t} |X_s - X_t| \right\|_{\psi} \leq K \int_0^{\operatorname{diam} T} \psi^{-1}(D(\epsilon, d))d\epsilon$$

General bound

Corollary 2.2.5

The constant K can be chosen such that

$$\left\| \sup_{s,t} |X_s - X_t| \right\|_\psi \leq K \int_0^{\operatorname{diam} T} \psi^{-1}(D(\epsilon, d))\,d\epsilon$$

Proof

$$\left\| \sup_{d(s,t)\leq\delta} |X_s - X_t| \right\|_\psi \leq K \left[\int_0^{\eta} \psi^{-1}(D(\epsilon, d))\,d\epsilon + \delta\psi^{-1}(D(\eta, d)^2) \right].$$

Take $\eta = \operatorname{diam} T$.

Corollary of the corrolary

$$\left\| \sup_t |X_t| \right\|_\psi \leq \|X_{t_0}\|_\psi + K \int_0^{\text{diam } T} \psi^{-1}(D(\epsilon, d))d\epsilon$$

Assumptions

Assummption Theorem 2.2.4

- Standard assumptions on ψ
- $\{X_t : t \in T\}$ separable
- $\|X_s - X_t\|_\psi \leq Cd(s, t)$, for every s, t