Chapter 2.7: Bracketing Numbers

Geerten Koers

Reading group *Weak Convergence and Empirical Processes*

2020-02-17

Outline

α-smoothness

For $\alpha > 0$, let $\underline{\alpha}$ be the greatest integer smaller than α. For $k \in \mathbb{N}^d$, let $k. = \sum k_i$ and define

$$D^k = \frac{\partial^{k.}}{\partial x_1^{k_1} \cdots \partial x_d^{k_d}}.$$

α-smoothness

For $\alpha > 0$, let $\underline{\alpha}$ be the greatest integer smaller than α. For $k \in \mathbb{N}^d$, let $k. = \sum k_i$ and define

$$D^k = \frac{\partial^{k.}}{\partial x_1^{k_1} \cdots \partial x_d^{k_d}}.$$

Definition

For $\alpha > 0$, and $M \geq 0$, a function $f : \mathcal{X} \to \mathbb{R}$ is an element of $C_M^\alpha(\mathcal{X})$ if for

$$\|f\|_\alpha \equiv \max_{k. \leq \underline{\alpha}} \sup_x |D^k f(x)| + \max_{k. = \underline{\alpha}} \sup_{x,y} \frac{|D^k f(x) - D^k f(y)|}{\|x - y\|^{\alpha - \underline{\alpha}}},$$

we have $\|f\|_\alpha \leq M$.

α-smoothness

Theorem (2.7.1)

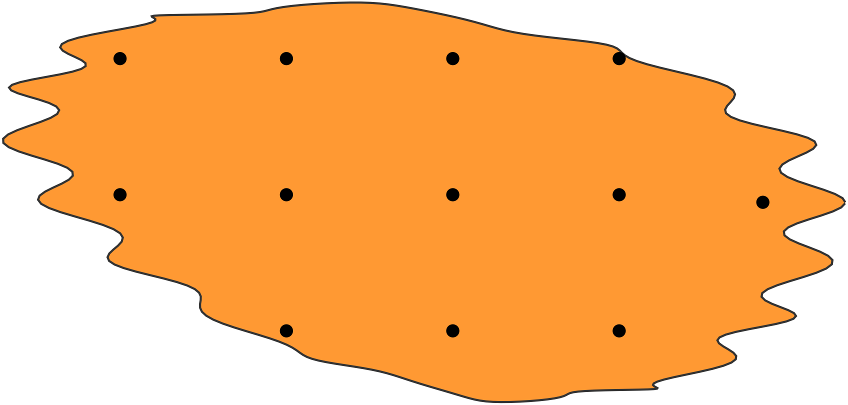
Let $\mathcal{X}$ be a bounded, convex subset of $\mathbb{R}^d$ with nonempty interior. There exists a constant K depending only on α and d such that

$$\log N(\epsilon, C_1^\alpha(\mathcal{X}), \|\cdot\|_\infty) \leq K\lambda(\mathcal{X}^1)\left(\frac{1}{\epsilon}\right)^{d/\alpha}$$

for every $\epsilon > 0$, where $\lambda(\mathcal{X}^1)$ is the Lebesgue measure of the set $\{x : \|x - \mathcal{X}\| < 1\}$.

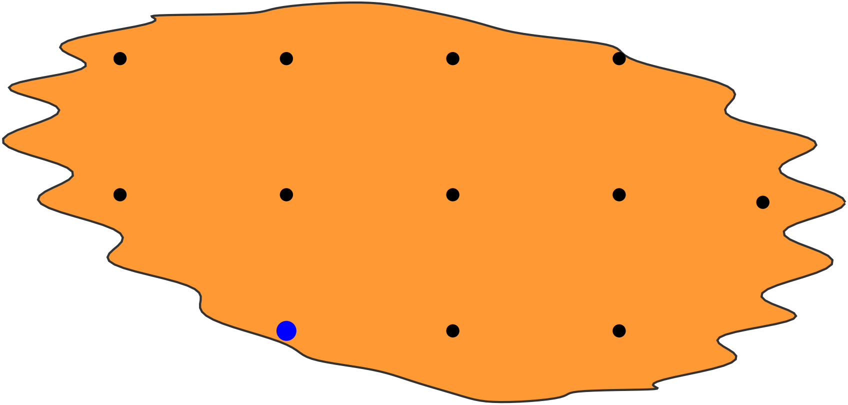
α-smoothness

Sketch of proof:



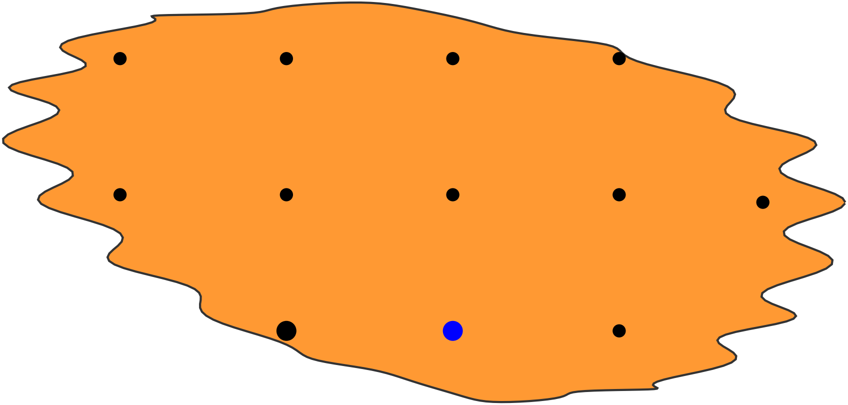
α-smoothness

Sketch of proof:

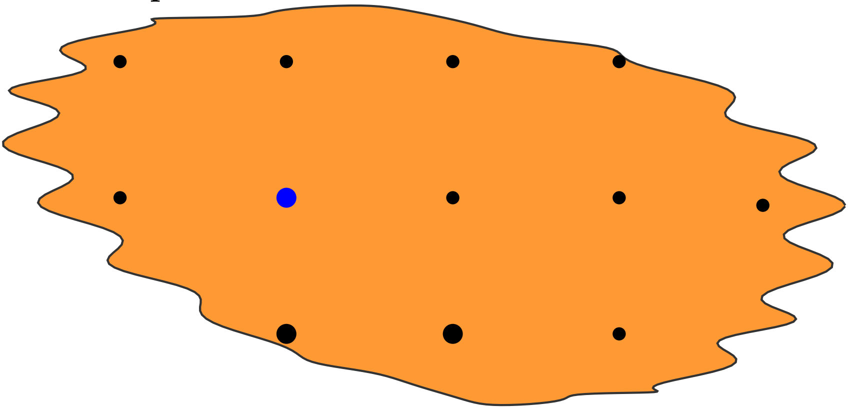


α-smoothness

Sketch of proof:

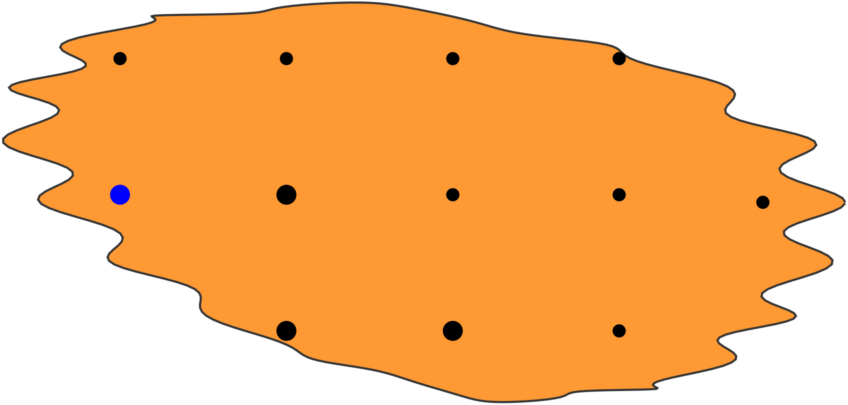


Sketch of proof:



α-smoothness

Sketch of proof:



Proof:

1. **For $\delta = \epsilon^{1/\alpha}$, fix a δ-net $x_1, \ldots, x_m$ of $\mathcal{X}$.**

α-smoothness

Proof:

1. **For $\delta = \epsilon^{1/\alpha}$, fix a δ-net $x_1, \ldots, x_m$ of $\mathcal{X}$.**
2. $m \lesssim \lambda(\mathcal{X}^1)/\delta^d$.

Proof:

1. **For $\delta = \epsilon^{1/\alpha}$, fix a δ-net $x_1, \ldots, x_m$ of $\mathcal{X}$.**
2. $m \lesssim \lambda(\mathcal{X}^1)/\delta^d$.
3. **For $k = (k_1, \ldots, k_d)$, define**

$$A_k f = \left(\left\lfloor \frac{D^k f(x_1)}{\delta^{\alpha - k.}} \right\rfloor, \ldots, \left\lfloor \frac{D^k f(x_m)}{\delta^{\alpha - k.}} \right\rfloor \right)$$

Proof:

1. **For $\delta = \epsilon^{1/\alpha}$, fix a δ-net $x_1, \ldots, x_m$ of $\mathcal{X}$.**
2. $m \lesssim \lambda(\mathcal{X}^1)/\delta^d$**.**
3. **For $k = (k_1, \ldots, k_d)$, define**

$$A_k f = \left(\left\lfloor \frac{D^k f(x_1)}{\delta^{\alpha - k.}} \right\rfloor, \ldots, \left\lfloor \frac{D^k f(x_m)}{\delta^{\alpha - k.}} \right\rfloor \right)$$

4. $\delta^{\alpha - k.} A_k f$ **is the discretization of the $(k_1, \ldots, k_d)$-partial derivative of f evaluated at the net.**

α-smoothness

1. **Claim: if $A_k f = A_k g$ for all k, then $\|f - g\|_\infty \lesssim \epsilon$.**

α-smoothness

1. **Claim: if $A_k f = A_k g$ for all k, then $\|f - g\|_\infty \lesssim \epsilon$.**
2. **For each x there is an x_i with $\|x - x_i\| \leq \delta$. Then**

$$(f-g)(x) = \sum_{k. \leq \beta} D^k (f-g)(x_i) \frac{(x-x_i)^k}{k!} + R$$

α-smoothness

1. **Claim: if $A_k f = A_k g$ for all k, then $\|f - g\|_\infty \lesssim \epsilon$.**
2. **For each x there is an x_i with $\|x - x_i\| \leq \delta$. Then**

$$(f-g)(x) = \sum_{k. \leq \beta} D^k (f-g)(x_i) \frac{(x-x_i)^k}{k!} + R$$

3. $|R| \lesssim \|x - x_i\|^\alpha$.

α-smoothness

1. **Claim: if $A_k f = A_k g$ for all k, then $\|f-g\|_\infty \lesssim \epsilon$.**
2. **For each x there is an x_i with $\|x - x_i\| \leq \delta$. Then**

$$(f-g)(x) = \sum_{k. \leq \beta} D^k(f-g)(x_i)\frac{(x-x_i)^k}{k!} + R$$

3. $|R| \lesssim \|x - x_i\|^\alpha$.
4. $$|f-g|(x) \lesssim \sum_{k. \leq \beta} \delta^{\alpha - k.} \frac{\delta^{k.}}{k!} + \delta^\alpha.$$

α-smoothness

1. **Claim: if $A_k f = A_k g$ for all k, then $\|f - g\|_\infty \lesssim \epsilon$.**
2. **For each x there is an x_i with $\|x - x_i\| \leq \delta$. Then**

$$(f-g)(x) = \sum_{k.\leq\beta} D^k(f-g)(x_i)\frac{(x-x_i)^k}{k!} + R$$

3. $|R| \lesssim \|x - x_i\|^\alpha$.
4. $$|f-g|(x) \lesssim \sum_{k.\leq\beta} \delta^{\alpha-k.}\frac{\delta^{k.}}{k!} + \delta^\alpha.$$
5. **Thus**

$$|f-g|(x) \leq \delta^\alpha(e^d + 1).$$

α-smoothness

1. **The covering number is bounded by the number of matrices**

$$Af = \begin{pmatrix} A_{0,0,\ldots,0}f \\ A_{1,0,\ldots,0}f \\ \vdots \\ A_{0,0,\ldots,\beta}f \end{pmatrix}$$

α-smoothness

1. **The covering number is bounded by the number of matrices**

$$Af = \begin{pmatrix} A_{0,0,\ldots,0}f \\ A_{1,0,\ldots,0}f \\ \vdots \\ A_{0,0,\ldots,\beta}f \end{pmatrix}$$

2. **Number of rows is bounded by** $(\beta+1)^d$**.**

α-smoothness

1. **The covering number is bounded by the number of matrices**

$$Af = \begin{pmatrix} A_{0,0,\ldots,0}f \\ A_{1,0,\ldots,0}f \\ \vdots \\ A_{0,0,\ldots,\beta}f \end{pmatrix}$$

2. **Number of rows is bounded by $(\beta + 1)^d$.**
3. **Since $|D^k f(x_i)| \leq 1$, the number of possible values of $A_k f$ is bounded by $2/\delta^{\alpha - k.} + 2$.**

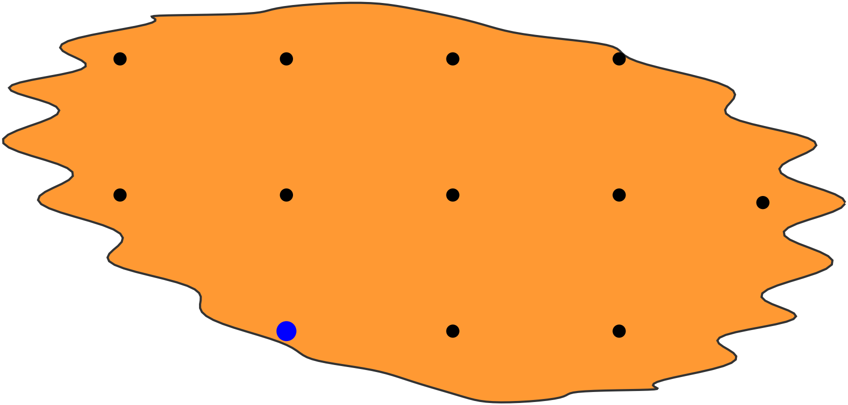
α-smoothness

1. **The covering number is bounded by the number of matrices**

$$Af = \begin{pmatrix} A_{0,0,\ldots,0}f \\ A_{1,0,\ldots,0}f \\ \vdots \\ A_{0,0,\ldots,\beta}f \end{pmatrix}$$

2. **Number of rows is bounded by $(\beta+1)^d$.**
3. **Since $|D^k f(x_i)| \leq 1$, the number of possible values of $A_k f$ is bounded by $2/\delta^{\alpha-k.}+2$.**
4. **Therefore each column has at most $(2\delta^{-\alpha}+2)^{(\beta+1)^d}$ different values.**

At most $(2\delta^{-\alpha}+2)^{(\beta+1)^d}$ different values for Af in



α-smoothness

1. **Let the x_i be such that for all j there is an index $i < j$ with $\|x_i - x_j\| < 2\delta$.**

α-smoothness

1. **Let the x_i be such that for all j there is an index $i < j$ with $\|x_i - x_j\| < 2\delta$.**
2. **For such x_i, x_j:**

$$D^k f(x_j) = \sum_{k.+l.\leq\beta} D^{k+l} f(x_i) \frac{(x_i - x_j)^l}{l!} + R.$$

1. **Let the x_i be such that for all j there is an index $i < j$ with $\|x_i - x_j\| < 2\delta$.**
2. **For such x_i, x_j:**

$$D^k f(x_j) = \sum_{k.+l.\leq\beta} D^{k+l} f(x_i) \frac{(x_i - x_j)^l}{l!} + R.$$

3. $|R| \lesssim \|x_i - x_j\|^{\alpha - k.}$.

α-smoothness

1. **Let the x_i be such that for all j there is an index $i < j$ with $\|x_i - x_j\| < 2\delta$.**
2. **For such x_i, x_j:**

$$D^k f(x_j) = \sum_{k.+l.\leq\beta} D^{k+l} f(x_i) \frac{(x_i - x_j)^l}{l!} + R.$$

3. $|R| \lesssim \|x_i - x_j\|^{\alpha - k.}$.
4. $B_k f := \delta^{\alpha - k.} A_k f$.

α-smoothness

1. **Let the x_i be such that for all j there is an index $i < j$ with $\|x_i - x_j\| < 2\delta$.**
2. **For such x_i, x_j:**

$$D^k f(x_j) = \sum_{k.+l.\leq\beta} D^{k+l} f(x_i) \frac{(x_i - x_j)^l}{l!} + R.$$

3. $|R| \lesssim \|x_i - x_j\|^{\alpha - k.}$.
4. $B_k f := \delta^{\alpha - k.} A_k f$.
5.

$$\left| D^k f(x_j) - \sum_{k.+l.\leq\beta} B_{k+l} f(x_i) \frac{(x_i - x_j)^l}{l!} \right| \leq \delta^{\alpha - k.}.$$

α-smoothness

1. **Given ith column of Af, $D^k f(x_j)$ ranges over an interval of length proportional to $\delta^{\alpha-k}$.**

α-smoothness

1. **Given** i**th column of** Af**,** $D^k f(x_j)$ **ranges over an interval of length proportional to** $\delta^{\alpha-k.}$**.**
2. **Thus the** j**th column of** Af **ranges over an interval proportional to** $\delta^{k.-\alpha}\delta^{\alpha-k.} = 1$**.**

α-smoothness

1. **Given ith column of Af, $D^k f(x_j)$ ranges over an interval of length proportional to $\delta^{\alpha - k.}$.**
2. **Thus the jth column of Af ranges over an interval proportional to $\delta^{k. - \alpha}\delta^{\alpha - k.} = 1$.**
3. **Thus there exists a constant C (depending only on α and d) such that**

$$\#Af \leq (2\delta^{-\alpha} + 1)^{(\beta+1)^d} C^{m-1}.$$

α-smoothness

Corollary (2.7.2)

Let $\mathcal{X}$ be a bounded, convex subset of $\mathbb{R}^d$ with nonempty interior. There exists a constant K depending only on α, diam $\mathcal{X}$, and d such that

$$\log N_{[]}(\epsilon, C_1^\alpha(\mathcal{X}), L_r(Q)) \leq K(\frac{1}{\epsilon})^{d/\alpha},$$

for every $r \geq 1$, $\epsilon > 0$, and probability measure Q on $\mathbb{R}^d$.

α-smoothness

Corollary (2.7.2)

Let $\mathcal{X}$ be a bounded, convex subset of $\mathbb{R}^d$ with nonempty interior. There exists a constant K depending only on α, diam $\mathcal{X}$, and d such that

$$\log N_{[]}(\epsilon, C_1^\alpha(\mathcal{X}), L_r(Q)) \leq K(\frac{1}{\epsilon})^{d/\alpha},$$

for every $r \geq 1$, $\epsilon > 0$, and probability measure Q on $\mathbb{R}^d$.

Proof.

Consider brackets $[f_i - \epsilon, f_i + \epsilon]$ for $f_1, \ldots, f_p$ the centers of $\|\cdot\|_\infty$ balls of radius ϵ that cover $C_1^\alpha(\mathcal{X})$. □

α-smoothness

Corollary (2.7.3)

Let $\mathcal{C}_{\alpha,d}$ be the collection of subgraphs of $C_1^{\alpha}[0,1]^d$. There exists a constant K depending only on α and d such that

$$\log N_{[]}(\epsilon, \mathcal{C}_{\alpha,d}, L_r(Q)) \leq K\|q\|_{\infty}^{d/\alpha}(\frac{1}{\epsilon})^{dr/\alpha},$$

for every $r \geq 1$, $\epsilon > 0$, and probability measure Q with bounded Lebesge density q on $\mathbb{R}^d$.

α-smoothness

1. $f_1, \ldots, f_p$ **centers of** $\| \cdot \|_\infty$**-balls of radius** ϵ **covering** $C_1^\alpha[0,1]^d$**.**

α-smoothness

1. $f_1, \ldots, f_p$ **centers of** $\| \cdot \|_\infty$**-balls of radius** ϵ **covering** $C_1^\alpha[0,1]^d$**.**
2. C_i **and** D_i **subgraphs of** $f_i - \epsilon$ **and** $f_i + \epsilon$**.**

α-smoothness

1. $f_1, \ldots, f_p$ **centers of** $\|\cdot\|_\infty$**-balls of radius** ϵ **covering** $C_1^\alpha[0,1]^d$**.**
2. C_i **and** D_i **subgraphs of** $f_i - \epsilon$ **and** $f_i + \epsilon$**.**
3. **Then** $[C_i, D_i]$ **form brackets that cover** $\mathcal{C}$**.**

α-smoothness

1. $f_1, \ldots, f_p$ **centers of** $\|\cdot\|_\infty$**-balls of radius** ϵ **covering** $C_1^\alpha[0,1]^d$**.**
2. C_i **and** D_i **subgraphs of** $f_i - \epsilon$ **and** $f_i + \epsilon$**.**
3. **Then** $[C_i, D_i]$ **form brackets that cover** $\mathcal{C}$**.**
4. **Their** $L_1(Q)$ **size is**

$$\int_{[0,1]^d} \int_{\mathbb{R}} 1\{f_i(x) - \epsilon \leq t < f_i(x) + \epsilon\} dQ(t,x) \leq 2\epsilon \|q\|_\infty.$$

α-smoothness

1. $f_1, \ldots, f_p$ **centers of** $\|\cdot\|_\infty$**-balls of radius** ϵ **covering** $C_1^\alpha[0,1]^d$**.**
2. C_i **and** D_i **subgraphs of** $f_i - \epsilon$ **and** $f_i + \epsilon$**.**
3. **Then** $[C_i, D_i]$ **form brackets that cover** $\mathcal{C}$**.**
4. **Their** $L_1(Q)$ **size is**

$$\int_{[0,1]^d} \int_{\mathbb{R}} 1\{f_i(x) - \epsilon \leq t < f_i(x) + \epsilon\} dQ(t,x) \leq 2\epsilon \|q\|_\infty.$$

5. **Thus**

$$N_{[]}((2\epsilon\|q\|_\infty)^{1/r}, \mathcal{C}_{\alpha,d}, L_r(Q))$$

is bounded by p**.**

α-smoothness

Theorem (2.7.4)

Let (I_j) be a partition for $\mathbb{R}^d$. If $\mathcal{F}$ is such that $f \in \mathcal{F}$ implies that $f|_{I_j} \in C^{\alpha}_{M_j}(I_j)$, we have

$$\log N_{[]}(\epsilon, \mathcal{F}, L_r(Q)) \leq K \left(\frac{1}{\epsilon}\right)^{d/\alpha}$$

with

$$K = C \left(\sum_{j=1}^{\infty} \lambda(I_j^1)^{\frac{r}{V+r}} M_j^{\frac{Vr}{V+r}} Q(I_j)^{\frac{V}{V+r}} \right)^{\frac{V+r}{r}}.$$

α-smoothness

Theorem (2.7.4)

Let (I_j) be a partition for $\mathbb{R}^d$. If $\mathcal{F}$ is such that $f \in \mathcal{F}$ implies that $f|_{I_j} \in C^{\alpha}_{M_j}(I_j)$, we have

$$\log N_{[]}(\epsilon, \mathcal{F}, L_r(Q)) \leq K \left(\frac{1}{\epsilon}\right)^{d/\alpha}$$

with

$$K = C \left(\sum_{j=1}^{\infty} \lambda(I_j^1)^{\frac{r}{V+r}} M_j^{\frac{Vr}{V+r}} Q(I_j)^{\frac{V}{V+r}} \right)^{\frac{V+r}{r}}.$$

If $\alpha > 1/2$ and $\sum_{j=1}^{\infty} P([j, j+1))^s < \infty$ for some $s < 1/2$, then

$$\log N_{[]}(\epsilon, C_1^{\alpha}(\mathbb{R}), L_2(P)) \leq K \left(\frac{1}{\epsilon}\right)^{2-\delta}$$

Monotone Functions

Theorem (2.7.5)

The class $\mathcal{F}$ of monotone functions $f : \mathbb{R} \mapsto [0,1]$ satisfies

$$\log N_{[]}(\epsilon, \mathcal{F}, L_r(Q)) \leq K(\frac{1}{\epsilon}),$$

for every probability measure Q, every $r \geq 1$, and a constant K that depends on r only.

Monotone Functions

1. **Suffices to check for $\lambda = \textbf{Unif}(0,1)$:**

Monotone Functions

1. **Suffices to check for $\lambda = $ Unif$(0, 1)$:**
2. $Q^{-1}(u) = \inf\{x : Q(x) \geq u\}$.

Monotone Functions

1. **Suffices to check for $\lambda = \mathbf{Unif}(0,1)$:**
2. $Q^{-1}(u) = \inf\{x : Q(x) \geq u\}$.
3. $f \circ Q^{-1} : [0,1] \to [0,1]$ **is monotone.**

Monotone Functions

1. **Suffices to check for $\lambda = \textbf{Unif}(0,1)$:**
2. $Q^{-1}(u) = \inf\{x : Q(x) \geq u\}$.
3. $f \circ Q^{-1} : [0,1] \to [0,1]$ **is monotone.**
4. $Q^{-1} \circ Q(x) \leq x$ **and** $u \leq Q \circ Q^{-1}(u)$.

Monotone Functions

1. **Suffices to check for $\lambda = \mathbf{Unif}(0,1)$:**
2. $Q^{-1}(u) = \inf\{x : Q(x) \geq u\}$.
3. $f \circ Q^{-1} : [0,1] \to [0,1]$ **is monotone.**
4. $Q^{-1} \circ Q(x) \leq x$ **and** $u \leq Q \circ Q^{-1}(u)$.
5. **Thus** $l \circ Q \leq f \circ Q^{-1} \circ Q \leq f$.

Monotone Functions

1. **Suffices to check for $\lambda = \textbf{Unif}(0,1)$:**
2. $Q^{-1}(u) = \inf\{x : Q(x) \geq u\}$.
3. $f \circ Q^{-1} : [0,1] \to [0,1]$ **is monotone.**
4. $Q^{-1} \circ Q(x) \leq x$ **and** $u \leq Q \circ Q^{-1}(u)$.
5. **Thus** $l \circ Q \leq f \circ Q^{-1} \circ Q \leq f$.
6. **For** $[l,u]$ **a bracket for** $f \circ Q^{-1}$**, then**

$$\begin{aligned} \|f - l \circ Q\|_{Q,r} &= \|f \circ Q^{-1} - l \circ Q \circ Q^{-1}\|_{\lambda,r} \\ &\leq \|f \circ Q^{-1} - l\|_{\lambda,r} \\ &< \epsilon \end{aligned}$$

Compact convex sets

Definition

For subsets C and D of a metric space, the Hausdorff distance is

$$h(C, D) = \sup_{x \in C} d(x, D) \vee \sup_{x \in D} d(x, C).$$

Compact convex sets

Definition

For subsets C and D of a metric space, the Hausdorff distance is

$$h(C,D) = \sup_{x\in C} d(x,D) \vee \sup_{x\in D} d(x,C).$$

$h(C,D) \leq \epsilon$ **is equivalent with**

$$\begin{cases} \forall x \in C : \exists y \in D : d(x,y) \leq \epsilon \\ \forall y \in D : \exists x \in C : d(y,x) \leq \epsilon \end{cases}$$

Compact convex sets

Lemma (2.7.8)

For the class $\mathcal{C}$ of all compact, convex subsets of a fixed, bounded subset of $\mathbb{R}^d$, with $d \geq 2$, one has

$$\log N(\epsilon, \mathcal{C}, h) \asymp \left(\frac{1}{\epsilon}\right)^{(d-1)/2},$$

with a constant depending only on d and the bounded set.

Compact convex sets

Corollary (2.7.9)

For the class $\mathcal{C}$ of all compact, convex subsets of a fixed, bounded subset of $\mathbb{R}^d$, with $d \geq 2$, one has

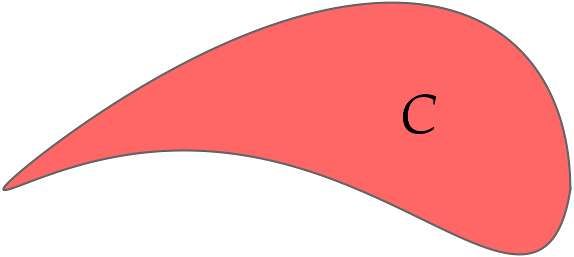
$$\log N_{[]}(\epsilon, \mathcal{C}, L_r(Q)) \leq K \left(\frac{1}{\epsilon}\right)^{(d-1)r/2}$$

Compact convex sets

1. ${}_{\epsilon}C = \{x : d(x, C^c) > \epsilon\}$.
2. $C^{\epsilon} = \{x : d(x, C) < \epsilon$.

Compact convex sets

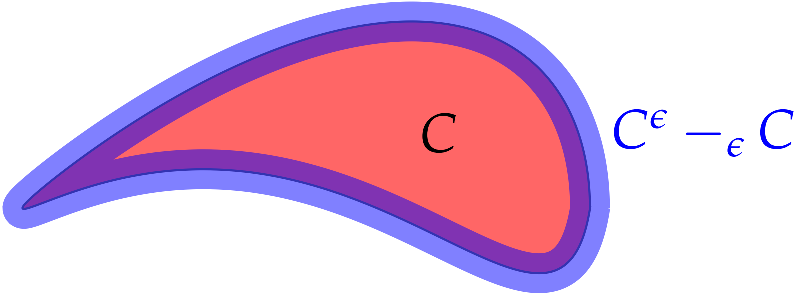
1. ${}_{\epsilon}C = \{x : d(x, C^c) > \epsilon\}$.
2. $C^{\epsilon} = \{x : d(x, C) < \epsilon$.



3.

Compact convex sets

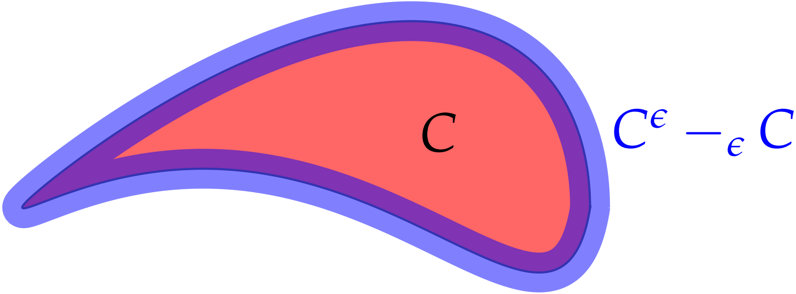
1. ${}_{\epsilon}C = \{x : d(x, C^c) > \epsilon\}$.
2. $C^{\epsilon} = \{x : d(x, C) < \epsilon$.



3.

Compact convex sets

1. ${}_{\epsilon}C = \{x : d(x, C^c) > \epsilon\}$.
2. $C^{\epsilon} = \{x : d(x, C) < \epsilon$.



3.
4. **There is a $K > 0$:**

$$\lambda(C^{\epsilon} -_{\epsilon} C) \leq K\epsilon.$$

Compact convex sets

1. **If $h(C, D) < \epsilon$, then ${}_{\epsilon}C \subset D \subset C^{\epsilon}$.**

Compact convex sets

1. **If** $h(C, D) < \epsilon$**, then** ${}_{\epsilon}C \subset D \subset C^{\epsilon}$**.**
2. **For** $C_1, \ldots, C_p$ **centers of Hausdorff balls of radius** ϵ**, then** $\left[{}_{\epsilon}C_i, C_i^{\epsilon}\right]$ **are covering brackets.**

Compact convex sets

1. **If $h(C, D) < \epsilon$, then ${}_{\epsilon}C \subset D \subset C^{\epsilon}$.**
2. **For $C_1, \ldots, C_p$ centers of Hausdorff balls of radius ϵ, then $[{}_{\epsilon}C_i, C_i^{\epsilon}]$ are covering brackets.**
3. **The size of the bracket is bounded by $\|q\|_{\infty}^{1/r}(K\epsilon)^{1/r}$.**

Compact convex sets

1. **If** $h(C, D) < \epsilon$**, then** ${}_{\epsilon}C \subset D \subset C^{\epsilon}$**.**
2. **For** $C_1, \ldots, C_p$ **centers of Hausdorff balls of radius** ϵ**, then** $[{}_{\epsilon}C_i, C_i^{\epsilon}]$ **are covering brackets.**
3. **The size of the bracket is bounded by** $\|q\|_{\infty}^{1/r}(K\epsilon)^{1/r}$**.**
4. p **is bounded by Lemma** 2.7.8**.**

Supergraphs

Corollary (2.7.10)

Let $\mathcal{F}$ be the class of all convex functions $f : C \to [0,1]$ defined on a compact, convex subset $C \subset \mathbb{R}^d$ such that $|f(x) - f(y)| \leq L\|x - y\|$. Then

$$\log N(\epsilon, \mathcal{F}, \|\cdot\|_\infty) \leq K(1+L)^{d/2} \left(\frac{1}{\epsilon}\right)^{d/2}.$$

Supergraphs

1. **For x with $f(x) < g(x)$, the closest point from $(x, f(x))$ to the supergraph of g is $(y, g(y))$.**

Supergraphs

1. **For x with $f(x) < g(x)$, the closest point from $(x, f(x))$ to the supergraph of g is $(y, g(y))$.**
2. **Their distance is bounded by**

$$|f(x) - g(y)| + \|x - y\| \leq h(C_f, C_g).$$

Supergraphs

1. **For x with $f(x) < g(x)$, the closest point from $(x, f(x))$ to the supergraph of g is $(y, g(y))$.**
2. **Their distance is bounded by**
$$|f(x) - g(y)| + \|x - y\| \leq h(C_f, C_g).$$
3. **Furthermore**
$$\begin{aligned} |f(x) - g(x)| &\leq |f(x) - g(y)| + |g(y) - g(x)| \\ &\leq |f(x) - g(y)| + L\|x - y\| \\ &\leq (1 + L)(|f(x) - g(y)| + \|x - y\|). \end{aligned}$$

Supergraphs

1. **For x with $f(x) < g(x)$, the closest point from $(x, f(x))$ to the supergraph of g is $(y, g(y))$.**
2. **Their distance is bounded by**
$$|f(x) - g(y)| + \|x - y\| \leq h(C_f, C_g).$$
3. **Furthermore**
$$\begin{aligned} |f(x) - g(x)| &\leq |f(x) - g(y)| + |g(y) - g(x)| \\ &\leq |f(x) - g(y)| + L\|x - y\| \\ &\leq (1 + L)(|f(x) - g(y)| + \|x - y\|). \end{aligned}$$
4. **Hence**
$$\|x - y\| + |f(x) - g(y)| \geq (1 + L)^{-1}|f(x) - g(x)|.$$

Supergraphs

1. **From**

$$(1+L)^{-1}|f(x)-g(x)| \leq |f(x)-f(y)| + \|x-y\|.$$

Supergraphs

1. **From**

$$(1+L)^{-1}|f(x)-g(x)| \leq |f(x)-f(y)| + \|x-y\|.$$

and

$$|f(x)-g(y)| + \|x-y\| \leq h(C_f, C_g)$$

Supergraphs

1. **From**

$$(1+L)^{-1}|f(x)-g(x)| \leq |f(x)-f(y)| + \|x-y\|.$$

and

$$|f(x)-g(y)| + \|x-y\| \leq h(C_f, C_g)$$

we get

$$|f(x)-g(x)| \leq (1+L)h(C_f, C_g).$$

Supergraphs

1. **From**

$$(1+L)^{-1}|f(x)-g(x)| \leq |f(x)-f(y)| + \|x-y\|.$$

   and

$$|f(x)-g(y)| + \|x-y\| \leq h(C_f, C_g)$$

   we get

$$|f(x)-g(x)| \leq (1+L)h(C_f, C_g).$$

2. **Thus** $\|f-g\|_\infty \leq (1+L)h(C_f, C_g)$**.**

Supergraphs

1. **From**

$$(1+L)^{-1}|f(x)-g(x)| \leq |f(x)-f(y)| + \|x-y\|.$$

and

$$|f(x)-g(y)| + \|x-y\| \leq h(C_f, C_g)$$

we get

$$|f(x)-g(x)| \leq (1+L)h(C_f, C_g).$$

2. **Thus** $\|f-g\|_\infty \leq (1+L)h(C_f, C_g)$.
3. **Now apply Lemma** 2.7.8.

Lipschitz in a parameter

Consider functions $x \mapsto f_t(x)$ indexed by $t \in T$ such that

$$|f_s(x) - f_t(x)| \leq d(s,t)F(x)$$

for some metric d and a function F.

Lipschitz in a parameter

Consider functions $x \mapsto f_t(x)$ indexed by $t \in T$ such that

$$|f_s(x) - f_t(x)| \le d(s,t)F(x)$$

for some metric d and a function F.

Theorem (2.7.11)

Let $\mathcal{F} = \{f_t : t \in R\}$ be a class of functions satisfying the preceding display for every s and t and some fixed function F. Then, for any norm $\|\cdot\|$,

$$N_{[]}(2\epsilon\|F\|, \mathcal{F}, \|\cdot\|) \le N(\epsilon, T, d).$$

Lipschitz in a parameter

Consider functions $x \mapsto f_t(x)$ indexed by $t \in T$ such that

$$|f_s(x) - f_t(x)| \leq d(s,t)F(x)$$

for some metric d and a function F.

Theorem (2.7.11)

Let $\mathcal{F} = \{f_t : t \in R\}$ be a class of functions satisfying the preceding display for every s and t and some fixed function F. Then, for any norm $\|\cdot\|$,

$$N_{[]}(2\epsilon\|F\|, \mathcal{F}, \|\cdot\|) \leq N(\epsilon, T, d).$$

Proof.

For $t_1, \ldots, t_p$ an ϵ-net of T, the brackets $[f_{t_i} - \epsilon F, f_{t_i} + \epsilon F]$ cover $\mathcal{F}$. $\square$