Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Chair for Mathematical Information Science

Prof. Dr. H. Bölcskei
Sternwartstrasse 7
CH-8092 Zürich

Mathematics of Information

Spring semester 2022

Problem Set 9

Problem 1 Metric entropy and minimax code length ☕

Let $d \in \mathbb{N}, \Omega \subset \mathbb{R}^d$, $\mathcal{C} \subset L^2(\Omega)$ compact and $\rho(f_1, f_2) = \|f_1 - f_2\|_{L^2(\Omega)}$. Show that

$$L(\epsilon, \mathcal{C}) = \lceil \log_2 N(\epsilon; \mathcal{C}, \rho) \rceil \quad \forall \epsilon > 0 \ ,$$

where $L(\epsilon, \mathcal{C})$ and $N(\epsilon; \mathcal{C}, \rho)$ are given in Definition 10.1 and 10.2 of the script.

Problem 2 Covering of unit cube. ☕

It has been proved in class that the covering number of the interval $\mathcal{C} := [-1, 1]$ scales as

$$\log_2 N(\varepsilon; \mathcal{C}, |\cdot|) \asymp \log\left(\varepsilon^{-1}\right).$$

We want to generalize this result to arbitrary dimensions. Namely, fix an integer $d \geq 1$, and prove that the covering number of the d-dimensional unit cube $\mathcal{C}^d := [-1, 1]^d$ scales as

$$\log_2 N(\varepsilon; \mathcal{C}^d, \|\cdot\|_\infty) \asymp d \log\left(\varepsilon^{-1}\right).$$ [1]

Problem 3 Metric entropy of parametric class of functions ☕☕

a) Let $(\mathcal{X}, \rho)$ be a metric space and $\mathcal{C}$ a compact set in $\mathcal{X}$. Order the following quantities

$$N(\epsilon; \mathcal{C}, \rho), N(2\epsilon; \mathcal{C}, \rho), M(\epsilon; \mathcal{C}, \rho), M(2\epsilon; \mathcal{C}, \rho).$$

b) Consider the following parametric class of functions

$$\mathcal{F} = \{f_\theta : [0, 1] \to \mathbb{R} \mid \theta \in [0, 1]\},$$

where for any fixed $\theta \in [0, 1]$, we set $f_\theta(x) := \ln(1 + \theta x), x \in [0, 1]$. We take the ∞-norm of functions defined on $[0, 1]$ to be given by

$$\|f - g\|_\infty = \max_{x \in [0,1]} |f(x) - g(x)|.$$

[1] One writes $f \asymp g$, if $f = O(g)$ and $g = O(f)$. One writes $f = O(g)$, if $\limsup_{\epsilon \to 0} \left|\frac{f(\epsilon)}{g(\epsilon)}\right| < \infty$.

i) For $\epsilon < 1/2$, construct an ϵ-covering $A(\epsilon)$ for the class $\mathcal{F}$ as follows.

Set $T = \lfloor \frac{1}{2\epsilon} \rfloor$, and for $i = 0, 1, \ldots, T$, define $\theta_i = 2\epsilon i$. We also add the point $\theta_{T+1} = 1$, thereby forming a collection $\{\theta_0, \ldots, \theta_T, \theta_{T+1}\}$ contained in $[0, 1]$. Show that the associated functions $\{f_{\theta_0}, \ldots, f_{\theta_T}, f_{\theta_{T+1}}\}$ constitute an ϵ-covering of $\mathcal{F}$. Find an upper bound on the ϵ-covering number $N(\epsilon; \mathcal{F}, \|\cdot\|_\infty)$ in terms of ϵ.

Hint: *You can use without proof that* $\frac{x}{1+x} \leq \ln(1 + x) \leq x$, *for all* $x > -1$.

ii) For $\epsilon < 1/3$, construct an ϵ-packing $P(\epsilon)$ for the class $\mathcal{F}$. Find a lower bound on the ϵ-packing number $M(\epsilon; \mathcal{F}, \|\cdot\|_\infty)$ in terms of ϵ.

Hint: *You can use without proof that* $\frac{x}{1+x} \leq \ln(1 + x) \leq x$, *for all* $x > -1$.

iii) Show that the metric entropy of the class $\mathcal{F}$ w.r.t. the norm $\|\cdot\|_\infty$ satisfies

$$\log N(\epsilon; \mathcal{F}, \|\cdot\|_\infty) \asymp \log(1/\epsilon), \quad \text{as } \epsilon \to 0.^2$$

Problem 4 Optimality with Johnson-Lindenstrauss.☕☕☕

Fix $\delta \in (0, 1/2)$. Throughout, 'log' denotes logarithm to the base 2. Fix an integer $m \geq 1$ and take $\{x_j\}_{j=1}^m$ to be an orthonormal basis for $\mathbb{R}^m$. Also fix an integer $k \geq 1$ together with a function $f : \mathbb{R}^m \to \mathbb{R}^k$ which, for all $1 \leq i, j \leq m$, satisfies

$$(1 - \delta)\|x_i - x_j\|_2^2 \leq \|f(x_i) - f(x_j)\|_2^2 \leq (1 + \delta)\|x_i - x_j\|_2^2. \qquad (1)$$

In this problem, we prove a converse to the Johnson-Lindenstrauss Lemma discussed in the lecture, namely that there exists a constant $C > 0$ independent of k, m, and δ such that, if $C \log(m) > k$, there does not exist a function f satisfying (1).

a) Prove that $\{f(x_j)\}_{j=1}^m \subseteq \mathcal{B}(y, 2)$, where $y := (1/m)\sum_{j=1}^m f(x_j)$ and $\mathcal{B}(y, 2)$ is the open ball with respect to the $\|\cdot\|_2$-norm centered at y and of radius 2.

b) Prove that $\{f(x_j)\}_{j=1}^m$ is a 1-packing of $(\mathcal{B}(y, 2), \|\cdot\|_2)$.

c) Prove that the 1-packing number $M(1; \mathcal{B}(y, 2), \|\cdot\|_2)$ of $(\mathcal{B}(y, 2), \|\cdot\|_2)$ satisfies

$$C \log (M(1; \mathcal{B}(y, 2), \|\cdot\|_2)) \leq k,$$

where $C > 0$ is a constant that does not depend on any of k, m, δ.

Hint: Use the volume ratio estimate provided in the Handout (Lemma 10.5).

d) Conclude that there exists a constant $C > 0$ independent of k, m, and δ such that, if $C \log(m) > k$, there does not exist a function $f : \mathbb{R}^m \to \mathbb{R}^k$ satisfying (1).

---

[2] One writes $f \asymp g$, if $f = O(g)$ and $g = O(f)$. One writes $f = O(g)$, if $\limsup_{\epsilon \to 0} \left| \frac{f(\epsilon)}{g(\epsilon)} \right| < \infty$.

Problem 5 Volume ratio estimate for the Hamming cube ☕☕

In this problem, we want to generalize the volume ratio estimate to the Hamming cube. Fix the integer $n \geq 1$, define the Hamming cube as $\mathbb{H}^n := \{0,1\}^n$, and consider the map

$$\begin{aligned} d\colon \mathbb{H}^n \times \mathbb{H}^n &\to \mathbb{N}_0 \\ (x,y) &\mapsto \#\{i \in \{1,\dots,n\} \,|\, x_i \neq y_i\}. \end{aligned}$$

We use the notation $[n]$ to designate the set of integers $\{1,\dots,n\}$ and $\mathbb{N}_0$ stands for the non-negative integers.

a) Prove that d is a metric on $\mathbb{H}^n$.

b) Given $x \in \mathbb{H}^n$ and an integer $m \in [n]$, we define the ball $\mathcal{B}(x,m)$, centered at x and of radius m with respect to the metric d, to be the subset of $\mathbb{H}^n$ given by

$$\mathcal{B}(x,m) := \{y \in \mathbb{H}^n \,|\, d(x,y) \leq m\}.$$

Compute the cardinality of the ball $\mathcal{B}(x,m)$.

c) Fix $m \in [n]$. An m-covering of $\mathbb{H}^n$ with respect to the metric d is a set $\{x_1,\dots,x_N\} \subset \mathbb{H}^n$ such that for all $x \in \mathbb{H}^n$, there exists an $i \in \{1,\dots,N\}$ so that $d(x,x_i) \leq m$. The m-covering number $N(m;\mathbb{H}^n,d)$ is the cardinality of the smallest m-covering. Prove that

$$N(m;\mathbb{H}^n,d) \geq \frac{2^n}{\sum_{k=0}^{m} \binom{n}{k}}.$$

d) Fix $m \in [n]$. An m-packing of $\mathbb{H}^n$ with respect to the metric d is a set $\{x_1,\dots,x_M\} \subset \mathbb{H}^n$ such that $d(x_i,x_j) > m$, for all distinct i,j. The m-packing number $M(m;\mathbb{H}^n,d)$ is the cardinality of the largest m-packing. Prove that, for a maximal m-packing $\{x_j\}_{j=1}^M$, the balls $\{\mathcal{B}(x_j,\lfloor m/2 \rfloor)\}_{j=1}^M$ are, indeed, disjoint subsets of $\mathbb{H}^n$.

e) Deduce from the statement in subproblem (d) that

$$M(m;\mathbb{H}^n,d) \leq \frac{2^n}{\sum_{k=0}^{\lfloor m/2 \rfloor} \binom{n}{k}}.$$

f) Prove that

$$\frac{2^n}{\sum_{k=0}^{m} \binom{n}{k}} \leq N(m;\mathbb{H}^n,d) \leq M(m;\mathbb{H}^n,d) \leq \frac{2^n}{\sum_{k=0}^{\lfloor m/2 \rfloor} \binom{n}{k}}.$$

Hint: You can use, without proof, that, for the Hamming cube, $N(m;\mathbb{H}^n,d) \leq M(m;\mathbb{H}^n,d)$.