In this lecture, we show that although the VC-hull classes might be considerably larger than the VC-classes, they are small enough to have finite uniform entropy integral.

**Theorem 18.1.** Let \((X, \mathcal{A}, \mu)\) be a measurable space, \(\mathcal{F} \subset \{f|f : X \to \mathbb{R}\}\) be a class of measurable functions with measurable square integrable envelope \(F\) (i.e., \(\forall x \in X, \forall f \in \mathcal{F}, |f(x)| < F(x)\), and \(\|F\|_2 = (\int F^2 d\mu)^{1/2} < \infty\), and the \(\varepsilon\)-net of \(\mathcal{F}\) satisfies \(N(\mathcal{F}, \varepsilon\|F\|_2, \|\cdot\|) \leq C (\frac{1}{\varepsilon})^V\) for \(0 < \varepsilon < 1\). Then there exists a constant \(K\) that depends only on \(C\) and \(V\) such that \(\log N(\text{conv}\mathcal{F}, \varepsilon\|F\|_2, \|\cdot\|) \leq K (\frac{1}{\varepsilon})^{\frac{V+2}{2}}\).

**Proof.** Let \(N(\mathcal{F}, \varepsilon\|F\|_2, \|\cdot\|) \leq C (\frac{1}{\varepsilon})^V \triangleq n\). Then \(\varepsilon = C^{1/n} n^{-1/V}\), and \(\|F\|_2 = C^{1/V} \|F\|_2 \cdot n^{-1/V}\). Let \(L = C^{1/V} \|F\|_2\). Then \(N(\mathcal{F}, L n^{-1/V}, \|\cdot\|) \leq n\) (i.e., the \(L \cdot n^{-1/V}\)-net of \(\mathcal{F}\) contains at most \(n\) elements).

Construct \(\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n \subset \cdots\) such that each \(\mathcal{F}_n\) is a \(L \cdot n^{-1/V}\)-net, and contains at most \(n\) elements. Let \(W = \frac{1}{2} + \frac{1}{n}\). We proceed to show that there exists constants \(C_k\) and \(D_k\) that depend only on \(C\) and \(V\) and are upper bounded (sup\(_k C_k \vee D_k < \infty\)), such that

\[
(18.1) \quad \log N(\text{conv}\mathcal{F}_{n, k}, C_k L \cdot n^{-W}, \|\cdot\|) \leq D_k \cdot n
\]

for \(n, k \geq 1\), and \(q \geq 3 + V\). This implies the theorem, since if we let \(k \to \infty\), we have \(\log N(\text{conv}\mathcal{F}, C_k L \cdot n^{-W}, \|\cdot\|) \leq D_k \cdot n\). Let \(\varepsilon = C^{1/n} C_n^{-1/V} n^{-W}\), and \(K = D_\infty C^{1/2} C^{1/2 + 2}\), we get \(C_k L \cdot n^{-W} = C_k C^{1/V} \|F\|_2 n^{-W} = \varepsilon\|F\|_2, n = \left(\frac{C_k C^{1/V}}{\varepsilon}\right)^{1/W}\) and \(\log N(\text{conv}\mathcal{F}, \varepsilon\|F\|_2, \|\cdot\|) \leq K \cdot \left(\frac{1}{\varepsilon}\right)^{\frac{V+2}{2}}\). Inequality 18.1 will proved in two steps: (1)

\[
(18.2) \quad \log N(\text{conv}\mathcal{F}_n, C_1 L \cdot n^{-W}, \|\cdot\|) \leq D_1 \cdot n
\]

by induction on \(n\), using Kolmogorov’s chaining technique, and (2) for fixed \(n\),

\[
(18.3) \quad \log N(\text{conv}\mathcal{F}_{n, k}, C_k L \cdot n^{-W}, \|\cdot\|) \leq D_k \cdot n
\]

by induction on \(k\), using the results of (1) and Kolmogorov’s chaining technique.

For any fixed \(n_0\) and any \(n \leq n_0\), we can choose large enough \(C_1\) such that \(C_1 L n_0^{-W} \geq \|F\|_2\). Thus \(N(\text{conv}\mathcal{F}_n, C_1 L \cdot n^{-W}, \|\cdot\|) = 1\) and 18.2 holds trivially. For general \(n\), fix \(m = n/d\) for large enough \(d > 1\). For any \(f \in \mathcal{F}_n\), there exists a projection \(\pi_m f \in \mathcal{F}_m\) such that \(\|f - \pi_m f\| \leq C_\pi m^{-\frac{1}{2}} \|f\| = L m^{-\frac{1}{2}}\) by definition of \(\mathcal{F}_m\). Since \(\sum_{f \in \mathcal{F}_n} \lambda_f \cdot f = \sum_{f \in \mathcal{F}_m} \mu_f \cdot f + \sum_{f \in \mathcal{F}_n} \lambda_f \cdot (f - \pi_m f)\), we have \(\text{conv}\mathcal{F}_n \subset \text{conv}\mathcal{F}_m + \text{conv}\mathcal{G}_n\), and the number of elements \(|\mathcal{G}_n| \leq |\mathcal{F}_n| \leq n\), where \(\mathcal{G}_n = \{f - \pi_m f : f \in \mathcal{F}_n\}\). We will find \(\frac{1}{2} C_1 L n^{-\frac{1}{2}}\)-nets for both \(\mathcal{F}_m\) and \(\mathcal{G}_n\), and bound the number of elements for them to finish to induction step. We need the following lemma to bound the number of elements for the \(\frac{1}{2} C_1 L n^{-\frac{1}{2}}\)-net of \(\mathcal{G}_n\).

**Lemma 18.2.** Let \((X, \mathcal{A}, \mu)\) be a measurable space and \(\mathcal{F}\) be an arbitrary set of \(n\) measurable functions \(f : X \to \mathbb{R}\) of finite \(L_2(\mu)\)-diameter \(\text{diam}\mathcal{F} (\forall f, g \in \mathcal{F}, |f - g| \leq \mathbb{R}) < \infty\). Then \(\forall \varepsilon > 0, N(\text{conv}\mathcal{F}, \varepsilon\text{diam}\mathcal{F}, \|\cdot\|) \leq (e + \varepsilon e^2)^{2/\varepsilon^2}\).
Proof. Let $\mathcal{F} = \{f_1, \cdots, f_n\}$. \(\forall \sum_{i=1}^{n} \lambda_i f_i\), let $Y_1, \cdots, Y_k$ be i.i.d. random variables such that $P(Y_i = f_j) = \lambda_j$ for all $j = 1, \cdots, n$. It follows that $\mathbb{E} Y_i = \sum \lambda_j f_j$ for all $i = 1, \cdots, k$, and

$$
\mathbb{E} \left( \frac{1}{k} \sum_{i=1}^{k} Y_i - \frac{1}{k} \sum_{j=1}^{n} \lambda_j f_j \right) \leq \frac{1}{k} \mathbb{E} \left( \sum_{i=1}^{k} Y_i - \sum_{j=1}^{n} \lambda_j f_j \right) \leq \frac{1}{k} (\text{diam} \mathcal{F})^2.
$$

Thus at least one realization of $\frac{1}{k} \sum_{i=1}^{k} Y_i$ has a distance at most $k^{-1/2} \text{diam} \mathcal{F}$ to each $\lambda_i f_i$. Since all realizations of $\frac{1}{k} \sum_{i=1}^{k} Y_i$ has the form $\frac{1}{k} \sum_{i=1}^{k} f_{j_i}$, there are at most $(\frac{n+k}{k} - 1)$ of such forms. Thus

$$
N(k^{-1/2} \text{diam} \mathcal{F}, \text{conv} \mathcal{F}, \| \cdot \|_2) \leq \left( \frac{n+k-1}{k} \right)
\leq \left( \frac{k+n}{k} \right)^{k/n} = \left( \frac{k+n}{k} \right)^{k/n} \leq e^k \left( \frac{k+n}{k} \right)^k = (e + en^2)^2/n^2
$$

\(\square\)

By triangle inequality and definition of $\mathcal{G}_n$, \(\text{diam} \mathcal{G}_n = \sup_{g_1, g_2 \in \mathcal{G}_n} ||g_1 - g_2||_2 \leq 2 \cdot Lm^{-1/V} \). Let $\epsilon \cdot \text{diam} \mathcal{G}_n = \epsilon \cdot 2Lm^{-1/V} = 1/2C_1 Ln^{-W}$. It follows that $\epsilon = 1/4C_1 m^{1/V} \cdot n^{-W}$, and

$$
N(\text{conv} \mathcal{G}_n, \epsilon \text{diam} \mathcal{G}_n, \| \cdot \|_2) \leq \left( e + en \cdot \frac{1}{16}C_1^2 m^{2/V} \cdot n^{-2W} \right)^{32C_1^{-2} m^{2/V} n^{2W}}
= \left( e + \frac{e}{16}C_1^2 d^{-2/V} \right)^{32C_1^{-2} d^{2/V} n}
$$

By definition of $\mathcal{F}_m$ and and induction assumption, $\log N(\text{conv} \mathcal{F}_m, C_1 L \cdot m^{-W}, \| \cdot \|_2) \leq D_1 \cdot m$. In other words, the $C_1 L \cdot m^{-W}$-net of $\text{conv} \mathcal{F}_m$ contains at most $e^{D_1 m}$ elements. This defines a partition of $\text{conv} \mathcal{F}_m$ into at most $e^{D_1 m}$ elements. Each element is isometric to a subset of a ball of radius $C_1 Lm^{-W}$. Thus each set can be partitioned into $(\frac{3C_1 Lm^{-W}}{2C_1 Lm^{-W}})^m = (6d^W)^{n/d}$ sets of diameter at most $\frac{1}{2}C_1 Ln^{-W}$ according to the following lemma.

Lemma 18.3. The packing number of a ball of radius $R$ in $\mathbb{R}^d$ satisfies $D(B(0, r), \epsilon, \| \cdot \|) \leq (\frac{3R}{r})^d$ for the usual norm, where $0 < \epsilon \leq R$.

As a result, the $C_1 Ln^{-W}$-net of $\text{conv} \mathcal{F}_n$ has at most $e^{D_1 n/d} (6d^W)^{n/d} (e + \epsilon C_1^2 d^{-2/V} \cdot 8d^2/V C_1^{-2} n)$ elements. This can be upper-bounded by $e^n$ by choosing $C_1$ and $d$ depending only on $V$, and $D_1 = 1$. For $k > 1$, construct $\mathcal{G}_{n,k}$ such that $\text{conv} \mathcal{F}_{nk^q} \subset \text{conv} \mathcal{F}_{n(k-1)^q} + \text{conv} \mathcal{G}_{n,k}$ in a similar way as before. $\mathcal{G}_{n,k}$ contains at most $nk^q$ elements, and each has a norm smaller than $L(n(k-1)^q)^{-1/V}$. To bound the cardinality of a $Lk^{-2} n^{-W}$-net, we set $\epsilon \cdot 2L(n(k-1)^q)^{-1/V} = Lk^{-2} n^{-W}$, get $\epsilon = \frac{1}{2} n^{-1/2} (k-1)^{q/V} k^{-2}$, 46
and
\[ N(\text{conv} \mathcal{G}_{n,k}, \epsilon \text{diam} \mathcal{G}_{n,k}, \| \cdot \|_2) \leq (e + enk^q e^2)^{2/\epsilon^2} \Rightarrow \]
\[ N(\text{conv} \mathcal{G}_{n,k}, \epsilon \text{diam} \mathcal{G}_{n,k}, \| \cdot \|_2) \leq \left( e + \frac{e}{4} k^{-4+q+2q/V} \right)^{8 \cdot n \cdot k^4 (k-1)^{-2q/V}}. \]

As a result, we get
\[
C_k = C_{k-1} + \frac{1}{k^2}
\]
\[
D_k = D_{k-1} + 8k^4 (k-1)^{-2q/V} \log(e + \frac{e}{4} k^{-4+q+2q/V}).
\]

For \(2q/V - 4 \geq 2\), the resulting sequences \(C_k\) and \(D_k\) are bounded. \(\square\)