DCG, Wintersemester 2022/23	Lecture 21 (5 pages)

The ε-Net Theorem

Anne Driemel

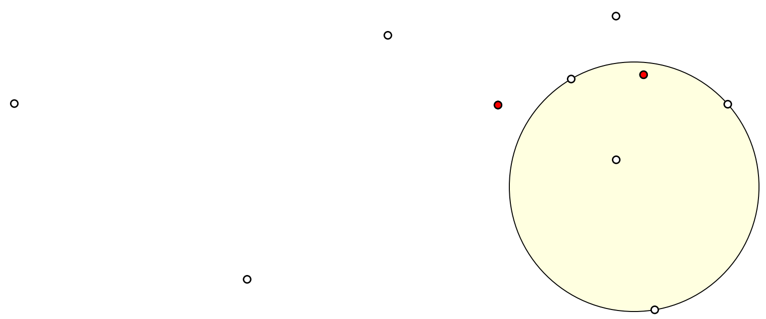
updated: January 17, 2025

In the previous lecture we introduced set systems and we showed that the VC-dimension characterizes the growth of subsystems, which we formalized with the growth function $\Pi_{\mathcal{R}}(m)$. In this lecture we will introduce the concept of ε-nets and show that a small VC-dimension implies the existence of ε-nets of size independent of the size of the ground set.

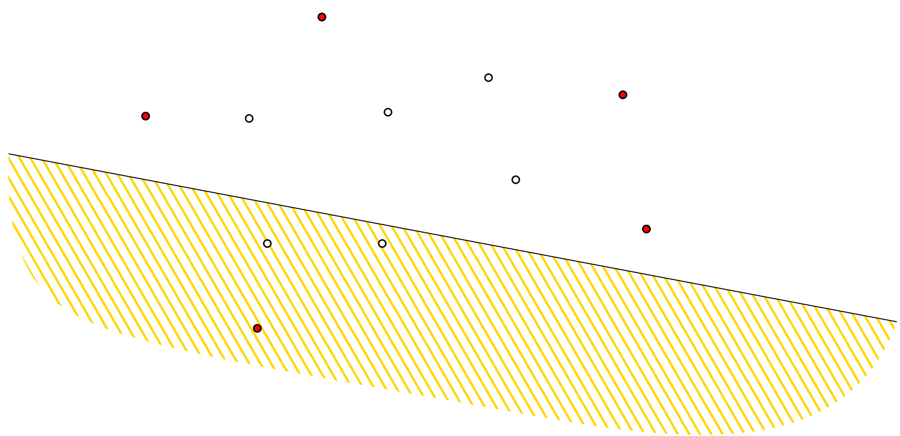
1 Definitions

Definition 21.1 (ε-net)**.** *Let $\mathcal{R}$ be a set system with finite ground set X and let $\varepsilon \in [0, 1]$ be a real value. A subset $A \subseteq X$ is an ε-net for this set system, if $A \cap r \neq \emptyset$ for all $r \in \mathcal{R}$ with $|r| \geq \varepsilon|X|$.*

Example 21.2. *The figure below shows a set of 10 points in the plane. Let X denote this set and consider the set system $\mathcal{R}$ where every set is defined by a disk D and contains the subset of X inside D. Let $\varepsilon = \frac{1}{2}$, an ε-net of $\mathcal{R}$ is shown with red (filled) dots. It can be verified that any disk containing at least 5 points also contains a point of the ε-net.*



Example 21.3. *Let X be a set of n points in $\mathbb{R}^d$. Let $\mathcal{R}$ be a set system, where every set is defined by a vector $a \in \mathbb{R}^d$ and a real value $b \in \mathbb{R}$, by $r = \{x \in X \mid \langle a, x\rangle \leq b\}$. That is, $\mathcal{R}$ is a set system of halfspaces in $\mathbb{R}^d$ intersected with a finite groundset. Let $\varepsilon = \frac{1}{n}$, then the set of points that lie on the boundary of the convex hull of X serves as an example of an ε-net for this set system.*



2 The ε-net theorem

If $\mathcal{R}$ is a finite set system that consists of all possible subsets of the ground set X, then no small ε-nets can exist for any $\varepsilon < 1/2$: no matter what set $A \subset X$ one chooses, if $|A| < |X|/2$, there will be always be a set $r \in \mathcal{R}$, namely $X \setminus A$, such that $|r| \geq |X|/2 > \varepsilon|X|$ while $r \cap A = \emptyset$. However, if $\mathcal{R}$ is a finite set system of small VC-dimension, small ε-nets do exist:

Theorem 21.4. *Let $\mathcal{R}$ be a finite set system with ground set X and VC-dimension at most d and let $0 < \varepsilon < \frac{1}{2}$ be a parameter, then there exists an ε-net of size in $O(\frac{\mathrm{d}}{\varepsilon} \ln \frac{1}{\varepsilon})$ for this set system.*

Proof. In our calculations, we will assume $\mathrm{d} \geq 2$ (the result will also hold for $\mathrm{d} = 1$, since $O(\frac{1}{\varepsilon} \ln \frac{1}{\varepsilon}) = O(\frac{2}{\varepsilon} \ln \frac{1}{\varepsilon})$).

Let $s = \frac{C\mathrm{d}}{\varepsilon} \ln \frac{1}{\varepsilon}$ for some value of C, which will be determined later, and assume for simplicity that s is a natural number. Let N be a random sample of X generated by making s independent draws from X. We claim that N is a ε-net with positive probability.

For simplicity, assume that all $r \in \mathcal{R}$ satisfy $|r| \geq \varepsilon|X|$. (Otherwise, consider the set system that only contains such sets). Let E_0 be the event that there exists a set $r \in \mathcal{R}$ with $N \cap r = \emptyset$ (so N is not an ε-net). We want to bound $\mathbf{Pr}[E_0]$ from above.

To this end, draw another sample M in the same way as above. Let E_1 be the event that there exists a set $r \in \mathcal{R}$ with $N \cap r = \emptyset$ and $|M \cap r| \geq k$, where $k = \frac{1}{2}\varepsilon s$. For simplicity assume that k is a natural number. (Here, we treat N and M as multisets, so an element drawn multiple times is also counted the appropriate number of times.)

The remainder of the proof is based on two claims.

Claim 21.5. $\mathbf{Pr}[E_1] \geq \frac{1}{2}\mathbf{Pr}[E_0]$.

Claim 21.6. $\mathbf{Pr}[E_1] < \frac{1}{2}$.

Together, the two claims imply $\mathbf{Pr}[E_0] < 1$, so N is an ε-net with positive probability. $\square$

It remains to prove the two claims.

To prove Claim 21.5, we will use the following lemma.

Lemma 21.7. *Let $Y = Y_1 + Y_2 + \cdots + Y_s$, where Y_i are independent random variables with $\mathbf{Pr}[Y_i = 1] = \alpha$ and $\mathbf{Pr}[Y_i = 0] = 1 - \alpha$. Then*

$$\mathbf{Pr}\left[Y \geq \frac{s \cdot \alpha}{2}\right] \geq \frac{1}{2}$$

provided that $s \cdot \alpha \geq 8$.

Proof. Recall that the variance of a random variable is defined as the mean squared difference with its mean value:

$$Var(Y) := \mathbf{E}\left[(Y - \mathbf{E}[E])^2\right].$$

For any $t \geq 0$ the following holds:

$$Var(Y) = \mathbf{E}\left[(Y - \mathbf{E}[Y])^2\right] \geq t^2 \cdot \mathbf{Pr}\left[(Y - \mathbf{E}[Y])^2 \geq t^2\right] = t^2 \cdot \mathbf{Pr}\left[|Y - \mathbf{E}[Y]| \geq t\right],$$

hence:

$$\mathbf{Pr}\left[|Y - \mathbf{E}[Y]| \geq t\right] \leq \frac{Var(Y)}{t^2} \quad \text{(Chebyshev's inequality)}$$

In our case,

$$\mathbf{E}[Y] = \sum_{i=1}^{s} \mathbf{E}[Y_i] = s \cdot \alpha$$

and, using the fact that the variance of the sum of independent variables is the sum of their variances:

$$Var(Y) = \sum_{i=1}^{s} Var(Y_i) = \sum_{i=1}^{s} \mathbf{E}\left[(Y_i - \mathbf{E}[Y_i])^2\right] = \sum_{i=1}^{s} (\alpha(1-\alpha)) \leq s \cdot \alpha$$

So, for $t = s \cdot \alpha/2$ with $s \cdot \alpha \geq 8$ we have by Chebyshev's inequality

$$\mathbf{Pr}\left[Y < \frac{s \cdot \alpha}{2}\right] \leq \mathbf{Pr}\left[|Y - \mathbf{E}[Y]| \geq \frac{s \cdot \alpha}{2}\right] \leq \frac{s \cdot \alpha}{(s \cdot \alpha/2)^2} = \frac{4}{s \cdot \alpha} \leq \frac{1}{2}$$

Indeed, the first inequality holds, since

$$Y < \frac{s \cdot \alpha}{2} \implies (s \cdot \alpha) - Y > \frac{s \cdot \alpha}{2} \implies |\mathbf{E}[Y] - Y| > \frac{s \cdot \alpha}{2}$$

□

Proof of Claim 21.5. By the definition of conditional probability

$$\mathbf{Pr}[E_1|E_0] = \frac{\mathbf{Pr}[E_1 \cap E_0]}{\mathbf{Pr}[E_0]} = \frac{\mathbf{Pr}[E_1]}{\mathbf{Pr}[E_0]}$$

since $E_1 \subseteq E_0$. Therefore, to show $\mathbf{Pr}[E_1] \geq \frac{1}{2} \cdot \mathbf{Pr}[E_0]$, it suffices to show $\mathbf{Pr}[E_1|E_0] \geq \frac{1}{2}$.

So assume that E_0 occurs. In this case, there exists an $r \in \mathcal{R}$ with $N \cap r = \emptyset$. Let's fix one such set r and denote it by r_N.

We have

$$\mathbf{Pr}[E_1|E_0] \geq \mathbf{Pr}[|M \cap r_N| \geq k]$$

by the definition of these events. Note that we have an inequality here since there could be more than one set that is hit by M many times.

Now define random variables $Y_1, \ldots, Y_s$ with $Y_i = 1$ if and only if the ith sample point of M falls into r_N. Observe that

$$\mathbf{Pr}[Y_i = 1] = \frac{|r_N|}{|X|} \geq \varepsilon$$

Let $Y = \sum_{i=1}^{s} Y_i$ and observe that $Y = |M \cap r_N|$.

Lemma 21.7 implies

$$\mathbf{Pr}\left[|M \cap r_N| \geq \frac{s \cdot \alpha}{2}\right] \geq \frac{1}{2}$$

where α denotes the probability $\mathbf{Pr}[Y_i = 1]$. (To satisfy the conditions of Lemma 21.7 we need to choose C sufficiently large)

Since $k = \varepsilon s/2$ and since $\alpha \geq \varepsilon$ we have that

$$|M \cap r_N| \geq \frac{s \cdot \alpha}{2} \implies |M \cap r_N| \geq k$$

Therefore,

$$\mathbf{Pr}[E_1|E_0] \geq \mathbf{Pr}[|M \cap r_N| \geq k] \geq \mathbf{Pr}\left[|M \cap r_N| \geq \frac{s \cdot \alpha}{2}\right] \geq \frac{1}{2}$$

□

Proof of Claim 21.6. Instead of choosing N and M directly as above, we generate the samples as follows. We first draw a sequence $A = (z_1, \ldots, z_{2s})$ of $2s$ independent random draws from X. Then in the second step, we randomly choose s positions in A (each subset of s positions out of $2s$ having the same probability) and put the elements at the chosen positions into N and the remaining elements into M. The resulting distribution of N and M is the same as before.

Now, consider A to be fixed and consider the subsystem $\mathcal{R}|_A$. Fix a set $r \in \mathcal{R}|_A$. Let E_r be the event that $N \cap r = \emptyset$ and $|M \cap r| \geq k$ (Note that this is like E_1, but for a fixed set r). We distinguish two cases: (i) $|r| = |A \cap r| < k$, and (ii) $|r| = |A \cap r| \geq k$.

In the first case, E_r cannot occur, so $\mathbf{Pr}[E_r] = 0$.

In the second case, $N \cap r = \emptyset$ implies $|M \cap r| \geq k$, so $\mathbf{Pr}[E_r] = \mathbf{Pr}[N \cap r = \emptyset]$. This is the probability that a random sample of s positions out of $2s$ avoids the at least k positions occupied by $A \cap r$. Therefore we have

$$\mathbf{Pr}[N \cap r = \emptyset] \leq \left(1 - \frac{k}{2s}\right)^s \leq \mathrm{e}^{-(k/2s)\cdot s} = \mathrm{e}^{-k/2}$$

where the last inequality follows from the fact that $1 - x \leq \mathrm{e}^{-x}$ for $x \in [0,1]$.

Recall that $k = \frac{1}{2}\varepsilon s = \frac{1}{2}C\mathrm{d}\ln\frac{1}{\varepsilon}$, so

$$\mathrm{e}^{-k/2} = \mathrm{e}^{-\frac{1}{4}C\mathrm{d}\ln(1/\varepsilon)} = \varepsilon^{\frac{1}{4}C\mathrm{d}}$$

Thus we have, for any fixed A and fixed $r \in \mathcal{R}|_A$:

$$\mathbf{Pr}[E_r] \leq \varepsilon^{\frac{1}{4}C\mathrm{d}}.$$

By Theorem 23.9 the number of distinct sets $r \in \mathcal{R}_A$ is at most

$$\Pi_{\mathcal{R}}(|A|) \leq \left(\frac{\mathrm{e}\cdot 2s}{\mathrm{d}}\right)^{\mathrm{d}} = \left(\frac{\mathrm{e}\cdot 2C}{\varepsilon}\ln\frac{1}{\varepsilon}\right)^{\mathrm{d}} \leq \left(\frac{\mathrm{e}\cdot 2C}{\varepsilon^2}\right)^{\mathrm{d}}$$

Thus, for any fixed A we can use a union bound over all $r \in \mathcal{R}|_A$ and obtain

$$\underset{\text{fixed } A}{\mathbf{Pr}}[E_1] \leq \sum_{r \in \mathcal{R}|_A} \mathbf{Pr}[E_r] \leq \left|\mathcal{R}|_A\right| \cdot \varepsilon^{\frac{1}{4}C\mathrm{d}} \leq \left(\frac{\mathrm{e}\cdot 2C}{\varepsilon^2}\right)^{\mathrm{d}} \cdot \varepsilon^{\frac{1}{4}C\mathrm{d}} = \left(\mathrm{e}\cdot 2C \cdot \varepsilon^{\frac{1}{4}C-2}\right)^{\mathrm{d}} < \frac{1}{2}$$

if C is sufficiently large (50 will do), $\mathrm{d} \geq 1$, and $\varepsilon \leq \frac{1}{2}$.

Now, since this bound holds as a worst-case bound for any sample A, it also holds as a bound on the probability of E_1 over all possible samples A. In particular, we can define F_A as the event that the set A is chosen in the first draw of $2s$ elements from X and we can think of the above probability as conditioned on the event F_A for a fixed set A. Now, by the law of total probability we have

$$\mathbf{Pr}[E_1] = \sum_{F_A} \mathbf{Pr}[E_1|F_A] \cdot \mathbf{Pr}[F_A] \leq \max_{F_A}(\mathbf{Pr}[E_1|F_A]) \cdot \sum_{F_A} \mathbf{Pr}[F_A] = \max_{F_A} \mathbf{Pr}[E_1|F_A] < \frac{1}{2}.$$

$\square$

3 Extension to infinite set systems

Definition 21.8 (ε-net)**.** *Let $\mathcal{R}$ be a set system with ground set X and let $\varepsilon \in [0,1]$ be a real value. Let $\mathcal{D}$ be a probability distribution defined on X. A subset $A \subseteq X$ is an ε-net for this set system, if $A \cap r \neq \emptyset$ for all $r \in \mathcal{R}$ with $\mathbf{Pr}\,[x \in r] \geq \varepsilon$ for a random sample x drawn from $\mathcal{D}$.*

Note that this is an extension of the definition of ε-net that we used above, since for finite set systems we can choose the uniform distribution over X for $\mathcal{D}$ and then the two definitions are equivalent. Using this definition and the above proof one can show with minor modifications that the following extended theorem holds true.

Theorem 21.9. *Let $\mathcal{R}$ be a set system with ground set X and VC-dimension at most d and let $0 < \varepsilon < \frac{1}{2}$ be a parameter. Let $\mathcal{D}$ be a probability distribution defined on X. There exists an ε-net of size in $O(\frac{\mathrm{d}}{\varepsilon} \ln \frac{1}{\varepsilon})$ for this set system.*

References

- Sariel Har-Peled, Chapter 5 in *Geometric Approximation Algorithms.* AMS Mathematical Surveys and Monographs. 2011.
- Jiří Matoušek, Chapters 10.2 and 10.3 in *Lectures on Discrete Geometry*, Springer Graduate Texts in Mathematics.