Lemmas from Past Literature used in "Policy Targeting under Network Interference"

Davide Viviano *

January 28, 2023

This document contains a list of lemmas from past literature used in the main derivations.

Appendix E Lemmas from past literature

E.1 Notation

We recall here the notation used in our derivations. Following Devroye et al. (2013)'s notation, for $x_1^n = (x_1, ..., x_n)$ being arbitrary points in $\mathcal{X}^n$, for a function class $\mathcal{F}$, with $f \in \mathcal{F}$, $f : \mathcal{X} \mapsto \mathbb{R}$, let $\mathcal{F}(x_1^n) = \{f(x_1), ..., f(x_n) : f \in \mathcal{F}\}$.

For a class of functions $\mathcal{F}$, with $f : \mathcal{X} \mapsto \mathbb{R}$, $\forall f \in \mathcal{F}$ and n data points $x_1, ..., x_n \in \mathcal{X}$ define the l_q-covering number $\mathcal{M}_q\Big(\eta, \mathcal{F}(x_1^n)\Big)$ to be the cardinality of the smallest cover $\{s_1, ..., s_N\}$, with $s_j \in \mathbb{R}^n$, such that for each $f \in \mathcal{F}$, there exist an $s_j \in \{s_1, ..., s_N\}$ such that $(\frac{1}{n}\sum_{i=1}^n |f(x_i) - s_j^{(i)}|^q)^{1/q} < \eta$. For $\bar{F}$ the envelope of $\mathcal{F}$, define the Dudley's integral as $\int_0^{2\bar{F}} \sqrt{\log\Big(\mathcal{M}_1(\eta, \mathcal{F}(x_1^n))\Big)} d\eta$.

For random variables $X = (X_1, ..., X_n)$, denote $\mathbb{E}_X[.]$ the expectation with respect to X, conditional on the other variables inside the expectation operator. Let $X_1, ..., X_n$ be arbitrary random variables. Let $\sigma = \{\sigma_i\}_{i=1}^n$ be $i.i.d$ Rademacher random variables ($P(\sigma_i = -1) = P(\sigma_i = 1) = 1/2$), independent of $X_1, ..., X_n$. The empirical Rademacher complexity is $\mathcal{R}_n(\mathcal{F}) = \mathbb{E}_\sigma\Big[\sup_{f\in\mathcal{F}} |\frac{1}{n}\sum_{i=1}^n \sigma_i f(X_i)|\Big| X_1, ..., X_n\Big]$.

E.2 Lemmas

Lemma E.1. *(Vershynin (2018), Lemma 6.4.2 and Boucheron et al. (2013), Lemma 11.4) Let $\sigma_1, ..., \sigma_n$ be Rademacher sequence independent of $X_1, ..., X_n$. Suppose that $X_1, \cdots, X_n$ are independent. Then*

$$\mathbb{E}\Big[\sup_{f\in\mathcal{F}}\Big|\sum_{i=1}^n f(X_i) - \mathbb{E}[f(X_i)]\Big|\Big] \le 2\mathbb{E}\Big[\sup_{f\in\mathcal{F}}\Big|\sum_{i=1}^n \sigma_i f(X_i)\Big|\Big].$$

*Stanford Graduate School of Business, and Department of Economics, Harvard University. Correspondence: dviviano@fas.harvard.edu.

Lemma E.2. *(Brook's Theorem, Brooks (1941)) For any connected undirected graph G with maximum degree Δ, the chromatic number of G is at most Δ unless G is a complete graph or an odd cycle case the chromatic number is $\Delta + 1$.*

Lemma E.3. *(Theorem 29.6, Devroye et al. (2013)) Let $\mathcal{F}_1, ..., \mathcal{F}_k, k \geq 1$ be classes of real functions on $\mathcal{X}$. For n arbitrary points $x_1^n = (x_1, ..., x_n) \in \mathcal{X}^n$, define $\mathcal{F} = \{f_1 + ... + f_k; f_j \in \mathcal{F}_j, \quad j = 1, \cdots, k\}$. Then for every $\eta > 0$, $x_1^n, n \geq 1, k \geq 1$, $\mathcal{M}_1\Big(\eta, \mathcal{F}(x_1^n)\Big) \leq \prod_{j=1}^k \mathcal{M}_1\Big(\eta/k, \mathcal{F}_j(x_1^n)\Big)$.*

Lemma E.4. *(Pollard, 1990) Let $\mathcal{F}$ and $\mathcal{G}$ be classes of real valued functions on $\mathcal{X}$ bounded by M_1 and M_2 respectively. For arbitrary fixed points $x_1^n \in \mathcal{X}^n$, let*

$$\mathcal{J}(x_1^n) = \{(h(x_1), ..., h(x_n); h \in \mathcal{J}\}, \quad \mathcal{J} = \{fg; f \in \mathcal{F}, g \in \mathcal{G}\}.$$

Then for every $\eta > 0$ and $x_1^n, n \geq 1$, $\mathcal{M}_1\Big(\eta, \mathcal{J}(x_1^n)\Big) \leq \mathcal{M}_1\Big(\frac{\eta}{2M_2}, \mathcal{F}(x_1^n)\Big)\mathcal{M}_1\Big(\frac{\eta}{2M_1}, \mathcal{G}(x_1^n)\Big)$.

Lemma E.5. *(Wenocur and Dudley, 1981) Let $g : \mathbb{R}^d \mapsto \mathbb{R}$ be an arbitrary function and consider the class of functions $\mathcal{G} = \{g + f, f \in \mathcal{F}\}$. Then $VC(\mathcal{G}) = VC(\mathcal{F})$ where $VC(\mathcal{F})$, $VC(\mathcal{G})$ denotes the VC dimension respectively of $\mathcal{F}$ and $\mathcal{G}$.*

Lemma E.6. *(From Theorem 5.22 in Wainwright (2019)) For a function class $\mathcal{F}$, with functions $f : \mathcal{X} \mapsto [-\bar{F}, \bar{F}]$, where $\bar{F} > 0$ denotes the envelope of $\mathcal{F}$, for fixed points $x_1^n \in \mathcal{X}^n$, and i.i.d. Rademacher random variables $\sigma_1, ..., \sigma_n \in \{-1, 1\}$,*

$$\mathbb{E}_\sigma\Big[\sup_{f \in \mathcal{F}}\Big|\sum_{i=1}^n \sigma_i f(x_i)\Big|\Big] \leq 32 \int_0^{D_q} \sqrt{n \log\Big(\mathcal{M}_q\Big(\eta, \mathcal{F}(x_1^n)\Big)\Big)} d\eta,$$

for any $n \geq 1$, where D_q denotes the maximum diameter of $\mathcal{F}(x_1^n)$ according to the metric $d_q(f, g) = \Big(\frac{1}{n}\sum_{i=1}^n (f(x_i) - g(x_i))^q\Big)^{1/q}$.

Theorem 5.22 in Wainwright (2019) provides a general version of Lemma E.6. Equivalent versions of Lemma E.6 are also in Van Der Vaart and Wellner (1996).

The following three lemmas control the Rademacher complexity in the presence of unbounded random variables and follow similarly to Kitagawa and Tetenov (2019).

Lemma E.7. *(Lemma A.1, Kitagawa and Tetenov (2019)) Let $t_0 > 1$, define*

$$g(t) := \begin{cases} 0, & \text{for } t = 0 \\ t^{-1/2}, & t \geq 1 \end{cases}, \quad h(t) = t_0^{-1/2} - \frac{1}{2}t_0^{-3/2}(t - t_0) + t_0^{-2}(t - t_0).$$

Then $g(t) \leq h(t)$, for $t = 0$ and all $t \geq 1$.

Lemma E.8. *Let $b_1, ..., b_n$ be independent Bernoulli random variables with $b_i \sim Bern(p_i)$. Let $\bar{p} = \frac{1}{n}\sum_{i=1}^n p_i$ with $n\bar{p} > 1$ and $g(.)$ as defined in the Lemma E.7. Then* $\mathbb{E}\Big[g\Big(\sum_{i=1}^n b_i\Big)\Big] < 2(n\bar{p})^{-1/2}$

Proof of Lemma E.8. The proof follows similarly to Kitagawa and Tetenov (2019), Lemma A.2, where p is replaced by $\bar{p}$ and omitted for brevity. □

Lemma E.9 (From Lemma A.5 in Kitagawa and Tetenov (2019))**.** *For any $i \in \{1, \cdots, n\}$, let $X_i \in \mathcal{X}$ be an arbitrary random variable and $\mathcal{F}$ a class of uniformly bounded functions with envelope $\bar{F}$. Let $\Omega_i | X_1, \cdots, X_n$ be random variables independently but not necessarily identically distributed, where $\Omega_i \geq 0$ is a scalar. Assume that for some $u > 0$, $\mathbb{E}[\Omega_i^{2+u}|Z] < B, \quad \forall i \in \{1, \cdots, n\}$. In addition assume that for any fixed points $x_1^n \in \mathcal{X}^n$, for some $V_n \geq 0$, for all $n \geq 1$, $\int_0^{2\bar{F}} \sqrt{\log\Big(\mathcal{M}_1\Big(\eta, \mathcal{F}(x_1^n)\Big)\Big)} d\eta < \sqrt{V_n}$. Let σ_i be i.i.d Rademacher random variables independent of $(\Omega_i)_{i=1}^n, (X_i)_{i=1}^n$. Then there exist a constant $0 < C_{\bar{F}} < \infty$ that only depend on $\bar{F}$ and u, such that for all $n \geq 1$*

$$\int_0^\infty \mathbb{E}\Big[\sup_{f \in \mathcal{F}} \Big|\frac{1}{n}\sum_{i=1}^n \sigma_i f(X_i) 1\{\Omega_i > \omega\}\Big| |X_1, \cdots, X_n\Big] d\omega \leq C_{\bar{F}} \sqrt{\frac{BV_n}{n}}.$$

Proof of Lemma E.9. The proof follows closely the one in Kitagawa and Tetenov (2019) (Lemma A.5), with the difference that we express the bound in terms of the covering number here. We refer to V_n as V for brevity, and $X = (X_i)_{i=1}^n$. Let

$$\xi_n(\omega) = \sup_{f \in \mathcal{F}} \Big|\frac{1}{n}\sum_{i=1}^n f(X_i) 1\{\Omega_i > \omega\}\sigma_i\Big|, \quad \bar{p}(\omega) = \frac{1}{n}\sum_{i=1}^n P(\Omega_i > \omega | X).$$

Case 1 Consider first values of ω for which $n\bar{p}(\omega) = \sum_{i=1}^n P(\Omega_i > \omega|X) \leq 1$. Due to the envelope condition, and the definition of Rademacher random variables,

$$\Big|\frac{1}{n}\sum_{i=1}^n f(X_i) 1\{\Omega_i > \omega\}\sigma_i\Big| \leq \bar{F}\frac{1}{n}\sum_{i=1}^n 1\{\Omega_i > \omega\}, \forall f \in \mathcal{F}.$$

Taking expectations we have $\mathbb{E}[\xi_n(\omega)|X] \leq \bar{F}\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^n 1\{\Omega_i > \omega\}\Big|X\Big] = \bar{F}\bar{p}(\omega)$ and the right hand side is bounded by $\bar{F}\frac{1}{n}$ for this particular case.

Case 2 Consider now values of ω such that $n\bar{p}(\omega) > 1$. Define the random variable $N_\omega = \sum_{i=1}^n 1\{\Omega_i > \omega\}$. Then I can write

$$\frac{1}{n}\sum_{i=1}^n f(Z_i) 1\{\Omega_i > \omega\}\sigma_i = \begin{cases} 0 \text{ if } N_\omega = 0 \\ \frac{N_\omega}{n}\frac{1}{N_\omega}\sum_{i=1}^n f(X_i) 1\{\Omega_i > \omega\}\sigma_i \text{ if } N_\omega \geq 1. \end{cases}$$

If $N_\omega \geq 1$, then

$$\begin{aligned}
\xi_n(y) &= \sup_{f\in\mathcal{F}} \Big| \frac{N_\omega}{n} \frac{1}{N_\omega} \sum_{i=1}^n f(X_i)\sigma_i 1\{\Omega_i > \omega\} \Big| \\
&= \sup_{f\in\mathcal{F}} \Big| \frac{N_\omega}{n} \frac{1}{N_\omega} \sum_{i=1}^n f(X_i)\sigma_i 1\{\Omega_i > \omega\} - \bar{p}(\omega) \frac{1}{N_\omega} \sum_{i=1}^n f(X_i) 1\{\Omega_i > \omega\}\sigma_i + \\
&+ \bar{p}(\omega) \frac{1}{N_\omega} \sum_{i=1}^n f(X_i) 1\{\Omega_i > \omega\}\sigma_i \Big| \\
&\leq \Big| \frac{N_\omega}{n} - \bar{p}(\omega) \Big| \sup_{f\in\mathcal{F}} \Big| \frac{1}{N_\omega} \sum_{i=1}^n f(X_i) 1\{\Omega_i > \omega\}\sigma_i \Big| + \bar{p}(\omega) \sup_{f\in\mathcal{F}} \Big| \frac{1}{N_\omega} \sum_{i=1} f(X_i)\sigma_i 1\{\Omega_i > \omega\} \Big|.
\end{aligned}$$

Denote $\mathbb{E}_\sigma$ the expectation only with respect to the Rademacher random variables σ. Conditional on N_ω, $\xi_n(\omega)$ sums over N_ω terms. Therefore, for a constant $0 < C_1 < \infty$ that only depend on $\bar{F}$, by Theorem 5.22 in Wainwright (2019) (Lemma E.6)

$$\mathbb{E}_\sigma \Big[\bar{p}(\omega) \sup_{f\in\mathcal{F}} | \frac{1}{N_\omega} \sum_{i=1}^n f(X_i)\sigma_i 1\{\Omega_i > \omega\} \Big| \Big] \leq C_1 \bar{p}(\omega) \sqrt{V} g(N_\omega)$$

where $g(.)$ is defined in Lemma E.7. Similarly,

$$\mathbb{E}_\sigma \Big[\Big| \frac{N_\omega}{n} - \bar{p}(\omega) \Big| \sup_{f\in\mathcal{F}} \Big| \frac{1}{N_\omega} \sum_{i=1}^n f(X_i) 1\{\Omega_i > \omega\}\sigma_i \Big| |X \Big] \leq \Big| \frac{N_\omega}{n} - \bar{p}(\omega) \Big| C_1 \sqrt{V} g(N_\omega).$$

For $N_\omega \geq 1$, it follows by the law of iterated expectations,

$$\mathbb{E}[\xi_n(\omega)|N_\omega, X] \leq |\frac{N_\omega}{n} - \bar{p}(\omega)| C_1 \sqrt{V} g(N_\omega) + C_1 \bar{p}(\omega) \sqrt{V} g(N_\omega) \leq |\frac{N_\omega}{n} - \bar{p}(\omega)| C_1 \sqrt{V} + C_1 \bar{p}(\omega) \sqrt{V} g(N_\omega) \tag{E.1}$$

where the last inequality follows by the definition of the $g(.)$ function and the fact that $N_\omega \in \{1, 2, \cdots\}$. For $N_\omega = 0$ instead, we have

$$\mathbb{E}[\xi_n(\omega)|N_\omega, X] \leq \Big| \frac{N_\omega}{n} - \bar{p}(\omega) \Big| C_1 \sqrt{V} g(N_\omega) + C_1 \bar{p}(\omega) \sqrt{V} g(N_\omega) = 0$$

by the definition of $g(.)$. Hence, the bound in Equation (E.1) always holds.

Unconditional expectation We are left to bound the unconditional expectation with respect to N_ω. Notice first that

$$\mathbb{E}\Big[\Big| \frac{N_\omega}{n} - \bar{p}(\omega) \Big| |X \Big] \leq C_1 \sqrt{V} \sqrt{\mathbb{E}\Big[\Big| \frac{N_\omega}{n} - \bar{p}(\omega) \Big|^2 \Big| X \Big]} = C_1 \sqrt{V} \sqrt{\mathrm{Var}\Big(\frac{1}{n} \sum_{i=1}^n 1\{\Omega_i > \omega\} \Big| X \Big)}.$$

By independence assumption,

$$C_1 \sqrt{V} \sqrt{\mathrm{Var}\Big(\frac{1}{n} \sum_{i=1}^n 1\{\Omega_i > \omega\} \Big| X \Big)} \leq C_1 \sqrt{V} \sqrt{\frac{1}{n^2} \sum_{i=1}^n P(\Omega_i > \omega | X)}.$$

In addition, since $n\bar{p}(\omega) > 1$, by Lemma E.8 , $\mathbb{E}[g(N_\omega)|Z] \leq \frac{2}{\sqrt{n}\sqrt{\bar{p}(\omega)}}$. Combining the inequalities, it follows

$$\mathbb{E}[\xi_n(\omega)|X] \leq C_1\sqrt{V}\sqrt{\frac{1}{n^2}\sum_{i=1}^n P(\Omega_i > \omega|X)} + \bar{p}(\omega)C_1\sqrt{V}\frac{2}{\sqrt{n}\sqrt{\bar{p}(\omega)}} \leq 2(1+C_1)\sqrt{\frac{V}{n}}\sqrt{\frac{1}{n}\sum_{i=1}^n P(\Omega_i > \omega|X)}.$$

This bound is larger than the bound derived for $n\bar{p}(\omega) < 1$, up to a constant factor $C_{\bar{F}} < \infty$. Therefore, $\mathbb{E}[\xi_n(\omega)|X] \leq C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\frac{1}{n}\sum_{i=1}^n P(\Omega_i > \omega|X)}$.

Integral bound I can now write

$$\int_0^\infty \mathbb{E}[\xi_n(\omega)|X]d\omega \leq \int_0^\infty C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^n \frac{P(\Omega_i > \omega|X)}{n}}d\omega = \int_0^1 C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^n \frac{P(\Omega_i > \omega|X)}{n}}d\omega + \int_1^\infty C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^n \frac{P(\Omega_i > \omega|X)}{n}}d\omega.$$

The first term is bounded as follows. $\int_0^1 C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^n \frac{P(\Omega_i>\omega|X)}{n}}d\omega \leq C_{\bar{F}}\sqrt{\frac{V}{n}}$. The second term is bounded instead as follows.

$$\int_1^\infty C_{\bar{F}}\sqrt{\frac{V}{n}}\sqrt{\sum_{i=1}^n \frac{P(\Omega_i > \omega|Z)}{n}}d\omega \leq C_{\bar{F}}\sqrt{\frac{V}{n}}\int_1^\infty \sqrt{\sum_{i=1}^n \frac{\mathbb{E}[\Omega_i^{2+u}|X]}{n\omega^{2+u}}}d\omega \leq C_{\bar{F}}\sqrt{\frac{V}{n}}\int_1^\infty \sqrt{\frac{B}{\omega^{2+u}}}d\omega \leq C'_{\bar{F}}\sqrt{\frac{VB}{n}}$$

for a constant $C'_{\bar{F}} < \infty$ that only depend on $\bar{F}$ and u. The proof completes. □

References

Boucheron, S., G. Lugosi, and P. Massart (2013). *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press.

Brooks, R. L. (1941). On colouring the nodes of a network. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 37, pp. 194–197. Cambridge University Press.

Devroye, L., L. Györfi, and G. Lugosi (2013). *A probabilistic theory of pattern recognition*, Volume 31. Springer Science & Business Media.

Kitagawa, T. and A. Tetenov (2019). Equality-minded treatment choice. *Journal of Business & Economic Statistics*, 1–14.

Pollard, D. (1990). Empirical processes: theory and applications. In *NSF-CBMS regional conference series in probability and statistics*, pp. i–86. JSTOR.

Van Der Vaart, A. W. and J. A. Wellner (1996). Weak convergence. In *Weak convergence and empirical processes*, pp. 16–28. Springer.

Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science*, Volume 47. Cambridge university press.

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*, Volume 48. Cambridge University Press.

Wenocur, R. S. and R. M. Dudley (1981). Some special vapnik-chervonenkis classes. *Discrete Mathematics 33*(3), 313–318.