# Notes for Week 11/1/19 - 11/8/19

Alden Green

December 11, 2019

Let $P$ and $Q$ be distributions on $\mathbb{R}^d$, absolutely continuous with respect to Lebesgue measure with densities $p$ and $q$, and let $\mu = (P+Q)/2$. For fixed points $Z = \{z_1, \ldots, z_N\} \subset \mathcal{Z}$ for some subset $\mathcal{Z}$ of the support of $\mu$, suppose we observe $\ell = (\ell_1, \ldots, \ell_N) \in \mathbb{R}^n$, where $\ell_1, \ldots, \ell_N$ are independent Rademacher random variables with means

$$\mathbb{E}[\ell_i] = \frac{p(z_i) - q(z_i)}{p(z_i) + q(z_i)}.$$

Let $\theta_p = (p(z_1), \ldots, p(z_N))$ and similarly $\theta_q = (q(z_1), \ldots, q(z_N))$; let $\theta_\mu = \frac{1}{2}\theta_p + \frac{1}{2}\theta_q$. Our goal: use the data $\ell, Z$ to test the following hypothesis:

$$\mathbf{H}_0 : \theta_p = \theta_q \quad \text{versus} \quad \mathbf{H}_a : d(\theta_p, \theta_q) > \epsilon$$

for some metric $d$ on the space $\mathbb{R}^n$, and some $\epsilon > 0$. For a given test $\phi : \{-1,1\}^N \times \mathcal{Z} \to \{0,1\}$, we will evaluate our test error using the sum of Type I and Type II error

$$\mathcal{R}(\phi; \theta_p, \theta_q) = \mathbb{E}_{\theta_\mu, \theta_\mu}(\phi) + \mathbb{E}_{\theta_p, \theta_q}(1 - \phi),$$

and the worst case risk

$$\mathcal{R}(\phi, \epsilon, \Theta) = \sup_{\theta_p, \theta_q \in \Theta} \mathcal{R}(\phi; \theta_p, \theta_q)$$

where $\Theta$ is a subset of $\mathbb{R}^n$ and the supremum is over all vectors $\theta_p, \theta_q$ such that $d(\theta_p, \theta_q) \geq \epsilon$.

In particular, we propose and analyze the following projection-based test. Letting $G_{N,r}$ be the $r$-neighborhood graph over $Z$ with combinatorial Laplacian $L$, we write the spectral decomposition of $L$ as $L = VSV^T$, where $V = (v_1, \ldots, v_N)$ is $N \times N$ orthonormal matrix, and $S$ is a diagonal matrix with entries $s_1 \leq s_2 \leq \ldots \leq s_N$. Our test statistic $T$ depends on a tuning parameter $C > 0$ in the following manner:

$$T = T(\ell, Z) = \frac{1}{N} \sum_{k : s_k \leq C} \left(\langle \ell, v_k \rangle\right)^2$$

and our test $\phi := \phi(\ell, Z)$ will simply be $\phi = \mathbf{1}\{T \geq \frac{\kappa}{N} + b\tau\}$ for a threshold $\tau > 0$ to be specified later.

## 0.1 Moments of $T$.

To upper bound the test error we must show that the fluctuations of $T$ (under null or alternative) are small compared to the difference in means $\mathbb{E}_{\theta_p, \theta_q}(T) - \mathbb{E}_{\theta_\mu, \theta_\mu}(T)$. We therefore must compute the first and second moments of $T$, under null and alternative. To compute the moments of $T$, we decompose the vector $\ell$ into a mean term and a noise term. Letting $\Delta_{P,Q} = \frac{\theta_P - \theta_q}{2\theta_\mu}$, we have $\ell = \Delta_{P,Q} + \varepsilon$ where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ consists of $n$ independent mean-zero random variables, distributed

$$\varepsilon_i = \begin{cases} \frac{\theta_{Q,i}}{\theta_{\mu,i}}, & \text{with probability } \frac{\theta_{P,i}}{\theta_{\mu,i}} \\ -\frac{\theta_{P,i}}{\theta_{\mu,i}}, & \text{with probability } \frac{\theta_{Q,i}}{\theta_{\mu,i}}. \end{cases}$$

### 0.1.1 Mean of $T$.

Let $\kappa$ be the largest integer $k \in [N]$ such that $s_k \leq C$. (Here $\kappa$ is a fixed quantity.) Then,

$$\begin{aligned}
\mathbb{E}(T) &= \frac{1}{N}\sum_{k=1}^{\kappa}\mathbb{E}\left[\langle v_k, \ell\rangle^2\right] \\
&= \frac{1}{N}\sum_{k=1}^{\kappa}\mathbb{E}\left[\langle \Delta_{P,Q}+\epsilon, v_k\rangle^2\right] \\
&= \frac{1}{N}\sum_{k=1}^{\kappa}\left\{\langle \Delta_{P,Q}, v_k\rangle^2 + \mathbb{E}\left[\langle \epsilon, v_k\rangle^2\right]\right\}
\end{aligned}$$

For a given $k \in [\kappa]$, we have

$$\begin{aligned}
\mathbb{E}\left[\langle \epsilon, v_k\rangle_2^2\right] &= \sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left[\varepsilon_i\varepsilon_j v_{k,i}v_{k,j}\right] \\
&= \sum_{i=1}^{N} v_{k,i}^2\mathbb{E}\left[\varepsilon_i^2\right]
\end{aligned}$$

and by a direct computation of $\mathbb{E}\left[\varepsilon_i^2\right]$ we obtain

$$\begin{aligned}
\mathbb{E}(T) &= \frac{1}{N}\left(\sum_{k=1}^{\kappa}\langle \Delta_{P,Q}, v_k\rangle^2 + \sum_{k=1}^{\kappa}\sum_{i=1}^{N} v_{k,i}^2\frac{\theta_{P,i}\theta_{Q,i}}{\theta_{\mu,i}^2}\right) \\
&= \frac{1}{N}\left(\sum_{k=1}^{\kappa}\langle \Delta_{P,Q}, v_k\rangle^2 + \kappa - \sum_{k=1}^{\kappa}\sum_{i=1}^{N} v_{k,i}^2\Delta_{P,Q,i}^2\right) \\
&= \frac{1}{N}\left(\sum_{k=1}^{\kappa}\langle \Delta_{P,Q}, v_k\rangle^2 + \kappa - \sum_{i=1}^{N}\Pi_{\kappa,ii}\Delta_{P,Q,i}^2\right) \\
&\geq \frac{1}{N}\left(\sum_{k=1}^{\kappa}\langle \Delta_{P,Q}, v_k\rangle^2 + \kappa - \Pi_{\max}\|\Delta_{P,Q}\|^2\right)
\end{aligned}$$

Under the null, $\Delta_{P,Q} = 0$ and therefore

$$\mathbb{E}_{\theta_\mu,\theta_\mu}(T) = \frac{\kappa}{N}. \tag{1}$$

Under the alternative, we have the following lower bound on the expectation of

$$\begin{aligned}
\sum_{k=1}^{\kappa}\langle \Delta_{P,Q}, v_k\rangle^2 &= \sum_{k=1}^{N}\langle \Delta_{P,Q}, v_k\rangle^2 - \sum_{k=\kappa+1}^{N}\langle \Delta_{P,Q}, v_k\rangle^2 \\
&= \|\Delta_{P,Q}\|_2^2 - \sum_{k=\kappa+1}^{N}\frac{1}{s_k}\langle \Delta_{P,Q}, v_k\rangle^2 s_k \\
&\geq \|\Delta_{P,Q}\|_2^2 - \frac{1}{C}\langle L\Delta_{P,Q}, \Delta_{P,Q}\rangle.
\end{aligned}$$

leading to the lower bound

$$\mathbb{E}_{\theta_p,\theta_q}(T) \geq \frac{\kappa}{N} + \frac{(1-\Pi_{\max})}{N}\|\Delta_{P,Q}\|^2 - \frac{1}{CN}\langle L\Delta_{P,Q}, \Delta_{P,Q}\rangle. \tag{2}$$

### 0.1.2 Variance of $T$.

Let $V_\kappa = (v_1 \dots v_\kappa)$ be the $n \times \kappa$ matrix containing the first $\kappa$ eigenvectors of $L$, and let $\Pi_\kappa = V_\kappa V_\kappa^T$. Observing that $T = \frac{1}{N}\langle \Pi_\kappa(\Delta_{P,Q} + \varepsilon), \Delta_{P,Q} + \varepsilon\rangle$, we have

$$\begin{aligned}\mathrm{Var}(T) &= \frac{1}{N^2}\mathrm{Var}(2\langle \Pi_\kappa \Delta_{P,Q}, \varepsilon\rangle + \langle \Pi_\kappa \varepsilon, \varepsilon\rangle) \\ &= \frac{1}{N^2}\Big\{4\underbrace{\mathrm{Var}(\langle \Pi_\kappa \Delta_{P,Q}, \varepsilon\rangle)}_{=:V_1} + 2\underbrace{\mathrm{Cov}(\langle \Pi_\kappa \Delta_{P,Q}, \varepsilon\rangle, \langle \Pi_\kappa \varepsilon, \varepsilon\rangle)}_{=:K_1} + \underbrace{\mathrm{Var}(\langle \Pi_\kappa \varepsilon, \varepsilon\rangle)}_{=:V_2}\Big\},\end{aligned} \tag{3}$$

and we now upper bound each of the three terms on the right hand side of the previous display.

**Upper bound on $V_1$:** Let $\Sigma := \mathrm{Cov}(\varepsilon)$ be the covariance matrix of $\varepsilon$. Noting that $\Sigma \preceq I$, we have

$$\mathrm{Var}(\langle \Pi_\kappa \Delta_{P,Q}, \varepsilon\rangle) = \Delta_{P,Q}^T \Pi_\kappa^T \Sigma \Pi_\kappa \Delta_{P,Q} \le \|\Pi_\kappa \Delta_{P,Q}\|^2. \tag{4}$$

**Upper bound on $K_1$:** Noting that $\mathbb{E}(\langle \Pi_\kappa \Delta_{P,Q}, \varepsilon\rangle) = 0$, we have that

$$\begin{aligned}K_1 &= \mathbb{E}\left[\langle \Pi_\kappa \Delta_{P,Q}, \varepsilon\rangle \langle \Pi_\kappa \varepsilon, \varepsilon\rangle\right] \\ &= \sum_{i=1}^N \sum_{j=1}^N \sum_{i'=1}^N \mathbb{E}\left[\varepsilon_i \varepsilon_j \Pi_{\kappa,ij} \varepsilon_{i'} (\Pi_\kappa \Delta_{P,Q})_i\right] \\ &= \sum_{i=1}^N \mathbb{E}\left[\varepsilon_i^3\right] \Pi_{\kappa,ii} (\Pi_\kappa \Delta_{P,Q})_i.\end{aligned}$$

Standard computations yield $\mathbb{E}\left[\varepsilon_i^3\right] = \frac{1}{2}(1 - \Delta_{P,Q,i}^2)(\Delta_{Q,P,i})$, and plugging in to the previous expression by Cauchy-Schwarz we have

$$\begin{aligned}K_1 = \sum_{i=1}^N \mathbb{E}\left[\varepsilon_i^3\right] \Pi_{\kappa,ii} (\Pi_\kappa \Delta_{P,Q})_i &= \frac{1}{2}\sum_{i=1}^N (1 - \Delta_{P,Q,i}^2)(\Delta_{Q,P,i}) \Pi_{\kappa,ii} (\Pi_\kappa \Delta_{P,Q})_i \\ &\le \|\Pi_\kappa \Delta_{P,Q}\| \cdot \left(\sum_{i=1}^N \Delta_{P,Q,i}^2 \Pi_{\kappa,ii}^2\right)^{1/2} \\ &\le \Pi_{\max} \cdot \|\Pi_\kappa \Delta_{P,Q}\| \cdot \|\Delta_{P,Q}\|\end{aligned} \tag{5}$$

**Upper bound on $V_2$:** $V_2$ is a variance of a sum, which we re-express as the sum of covariances:

$$\begin{aligned}V_2 &= \mathrm{Var}(\langle \Pi_\kappa \varepsilon, \varepsilon\rangle) \\ &= \sum_{i,j,i',j'=1}^N (\Pi_\kappa)_{ij} (\Pi_\kappa)_{i'j'} \mathrm{Cov}(\varepsilon_i \varepsilon_j, \varepsilon_{i'} \varepsilon_{j'}).\end{aligned}$$

This covariance will be non-zero only when $i = i' \neq j = j'$, $i = j' \neq j = i'$, or $i = i' = j = j'$, and therefore

$$\begin{aligned}V_2 &= 2\sum_{i=1}^N \sum_{j=1}^N (\Pi_\kappa)_{ij}^2 \mathrm{Var}(\varepsilon_i \varepsilon_j) + \sum_{i=1}^N (\Pi_\kappa)_{ii}^2 \mathrm{Var}(\varepsilon_i^2) \\ &\le 2\sum_{i=1}^N \sum_{j=1}^N (\Pi_\kappa)_{ij}^2 + \sum_{i=1}^N (\Pi_\kappa)_{ii}^2 \le 3\mathrm{tr}(\Pi_\kappa^2) = 3\kappa.\end{aligned} \tag{6}$$

**Putting the pieces together.** Plugging (4), (5), and (6) back into (3), we obtain

$$\operatorname{Var}(T) \le \frac{1}{N^2}\left\{4\|\Pi_\kappa \Delta_{P,Q}\|^2 + \Pi_{\max}\cdot\|\Pi_\kappa\Delta_{P,Q}\|\cdot\|\Delta_{P,Q}\| + 3\kappa\right\}. \tag{7}$$

## 0.2 Type I and Type II error.

We translate our bounds on the mean and variance of $T$ to bounds on the Type I and Type II error of our test $\phi$ using Chebyshev's inequality. Recall that our test statistic is of the form $\phi = \mathbf{1}\{T \ge \frac{\kappa}{N} + b\tau\}$, and let $\tau = \frac{\sqrt{3\kappa}}{N}$.

**Type I error.** We apply Chebyshev's inequality, and obtain

$$\begin{aligned}\mathbb{P}_{\theta_\mu,\theta_\mu}\left(T \ge \frac{\kappa}{N} + b\tau\right) &= \mathbb{P}_{\theta_\mu,\theta_\mu}\left(T - \frac{\kappa}{N} \ge b\tau - \mathbb{E}(T)\right)\\ &= \mathbb{P}_{\theta_\mu,\theta_\mu}\left((T - \frac{\kappa}{N})^2 \ge b^2\tau^2\right)\\ &\le \frac{\operatorname{Var}_{\theta_\mu,\theta_\mu}(T)}{b^2\tau^2} = \frac{1}{b^2}.\end{aligned}$$

**Type II error.** To obtain meaningful bounds on Type I error, we must assume some separation between $\theta_P$ and $\theta_Q$. In particular, letting $d(\theta_P,\theta_Q) := \|\Delta_{P,Q}\|^2$ we assume

$$\frac{1}{N}d(\theta_P,\theta_Q) \ge \frac{1}{(1-\Pi_{\max})}\left(\frac{1}{CN}\|D\Delta_{P,Q}\|^2 + 2b\tau\right). \tag{8}$$

Note that this implies

$$\mathbb{E}_{\theta_P,\theta_Q}(T) - \left(\frac{\kappa}{N} + b\tau\right) \ge \frac{(1-\Pi_{\max})}{N}\|\Delta_{P,Q}\|^2 - \frac{1}{CN}\|D\Delta_{P,Q}\|^2 - b\tau \ge b\tau \vee \frac{(1-\Pi_{\max})}{2N}\|\Delta_{P,Q}\|^2.$$

Now, we apply Chebyshev's inequality:

$$\begin{aligned}\mathbb{P}\left(T \le \frac{\kappa}{N} + b\tau\right) &= \mathbb{P}\big(T - \mathbb{E}[T] \le \frac{\kappa}{N} + b\tau - \mathbb{E}[T]\big)\\ &= \mathbb{P}\big((T-\mathbb{E}[T])^2 \le (\mathbb{E}[T] - \frac{\kappa}{N} - b\tau)^2\big)\\ &\le \frac{\operatorname{Var}(T)}{(\mathbb{E}[T] - \frac{\kappa}{N} - b\tau)^2}\\ &\le \frac{1}{N^2}\frac{4\|\Pi_\kappa\Delta_{P,Q}\|^2 + \Pi_{\max}\cdot\|\Pi_\kappa\Delta_{P,Q}\|\cdot\|\Delta_{P,Q}\| + 3\kappa}{b^2\tau^2 \vee \frac{(1-\Pi_{\max})^2}{4N^2}\|\Delta_{P,Q}\|^4}\\ &\le \frac{16}{(1-\Pi_{\max})^2\|\Delta_{P,Q}\|^2} \vee \frac{4\Pi_{\max}}{(1-\Pi_{\max})^2\|\Delta_{P,Q}\|^2} \vee \frac{1}{b^2}\\ &\le \frac{3}{b^2}\end{aligned}$$

where the last line follows by (8) whenever $N \ge \sqrt{48}$.

# 1 Upper bounds on $\Pi_{\max}$.

Clearly, the critical radius established by inequality (8) is only meaningful if $1 - \Pi_{\max} > 0$. In fact, we will want to show that $\Pi_{\max} < \frac{1}{2}$ so the factor of $(1-\Pi_{\max})^{-1}$ does not affect the rate at which the critical radius shrinks. We review what $\Pi_{\max}$ looks like on some common graphs.

### 1.1 Chain.

Suppose $G := G_{\mathrm{1N}}$ is a length-$N$ chain, i.e. a 1d grid graph on vertices $V = 1 : N$. The eigenvectors of $G$ are given by the discrete Fourier transform;

$$v_{1,i} = \frac{1}{\sqrt{N}}, \quad v_{k,i} = \sqrt{\frac{2}{N}} \cos\left(\frac{(i-.5)(k-1)\pi}{N}\right) \text{ for } i = 1, \ldots, N, \text{ k} = 2, \ldots, N.$$

Clearly $\Pi_{\max} \leq \frac{2\kappa}{N}$, so that whenever $\kappa < \frac{N}{4}$ we have $\Pi_{\max} < \frac{1}{2}$ as desired.

### 1.2 Grid.

Now suppose $G := G_{\mathrm{dN}}$ is a $d$-dimensional grid graph on $N$ nodes, i.e. $G = G_{1M} \otimes \cdots \otimes G_{1M}$ for $M = \frac{N}{d}$ (assume $M$ is an integer for simplicity). The product structure of $G$ yields an expression for its eigenvectors from the eigenvectors of $G_{\mathrm{1M}}$. In particular, for a given $i = (i_1, \ldots, i_d) \in [M]^d$ and $k = (k_1, \ldots, k_d) \in [M]^d$, we have

$$v_{k,i} = \prod_{j=1}^{d} \sqrt{\frac{2}{M}} \cos\left(\frac{(i_j - .5)(k_j - 1)\pi}{M}\right)$$

and therefore $v_{k,i}^2 \leq \frac{2}{M^d} = \frac{2}{N}$ for all $i, k \in [M]^d$. So we arrive at the same bound $\Pi_{\max} \leq \frac{2\kappa}{N}$ and again $\kappa < \frac{N}{4}$ implies $\Pi_{\max} < \frac{1}{2}$ as desired.

### 1.3 Spectral similarity.

Suppose we have graphs $G$ and $H$ with Laplacians $L_G$ and $L_H$, such that $G$ and $H$ are $\epsilon$ spectrally-similar, meaning

$$(1-\epsilon)x^T L_H x \leq x^T L_G x \leq (1+\epsilon) x^T L_H x, \quad \text{for all } x \in \mathbb{R}^N. \tag{9}$$

Write $L_G = V \Lambda V^T$ and $L_H = U S U^T$ for the spectral decompositions of $G$ and $H$, respectively, and assume that we have

$$\max_{(i,l) \in [N]^2} u_{l,i}^2 \leq \frac{2}{N},$$

which we have already shown holds for $H = G_{dN}$. For each $k$, and for any $L(k), R(k) \in [N]$, write

$$v_k = \sum_{l=R(k)}^{L(k)} \alpha_{k,l} u_l + \mathrm{proj}_{U_{L(k)}^{\perp}}(v_k);$$

as a result, we have that

$$\begin{aligned}
v_{k,i}^2 &\leq 2 \sum_{h=R(k)}^{L(k)} \sum_{l=R(k)}^{L(k)} \alpha_{k,h} \alpha_{k,l} u_{h,i} u_{l,i} + 2(\mathrm{proj}_{U_{R(k):L(k)}^{\perp}}(v_k))_i^2 \\
&\leq \frac{4}{N} \|\alpha\|_1^2 + 2\|\mathrm{proj}_{U_{L(k)}^{\perp}}(v_k)\|^2 \\
&\leq \frac{4}{N}(L(k) - R(k)) + 2\|\mathrm{proj}_{U_{R(k):L(k)}^{\perp}}(v_k)\|^2
\end{aligned}$$

where the last inequality follows since $\|\alpha\|_2 \leq 1$, and the inequality $\|\alpha\|_1 \leq \sqrt{L-R}\|\alpha\|_2$, since $\alpha$ is a length $L-R$-vector.

TODO: Upper bound $\|\mathrm{proj}_{U_{L(k)}^{\perp}}(v_k)\|^2$ using (9), and choose $L(k), R(k)$ to make the sum in the previous display as small as possible.

## 2 Notation.

We write $\Pi_{\max} = \max_{i=1,\dots,n} \Pi_{\kappa,ii}$ where we recall $\Pi_\kappa = V_\kappa V_\kappa^T$ and therefore $\Pi_{\kappa,ii} = \sum_{k=1}^{\kappa} v_{k,i}^2$.

We write $D$ for the incidence matrix of $L$.