Notes for the week of 4/8/19 - 4/12/19

Alden Green

April 29, 2019

Let $\{x_1, x_2, \ldots\}$ be an infinite sequence of points sampled independently from probability measure $\mathbb{P}$ with density function f. For each n, write $\mathbf{X}_n = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$. Given some $\lambda, \sigma > 0$, let

$$\mathcal{U} = \{x : f(x) \geq \lambda\}\,,\ \mathcal{C} = \text{one connected component of } \mathcal{U},\ \text{ and } \mathcal{C}_\sigma = \mathcal{C} + B(0,\sigma)$$

We will be concerned with the *normalized cut* over the subgraph induced by $\mathcal{C}_\sigma$. For convenience, denote $\widetilde{\mathbf{X}}_n = \mathcal{C}_\sigma[\mathbf{X}_n]$, $\widetilde{n} = \left|\widetilde{\mathbf{X}}_n\right|$, and let

$$\widetilde{E}_n = \left\{(i,j) : x_i, x_j \in \widetilde{\mathbf{X}}_n, \|x_i - x_j\|_2 \leq r\right\}, \widetilde{G}_{n,r} = \left(\widetilde{\mathbf{X}}_n, \widetilde{E}_n\right).$$

For a set $S \subseteq \widetilde{\mathbf{X}}_n$, the normalized cut of S within $\widetilde{G}_{n,r}$ can be defined as

$$\widetilde{\Phi}_{n,r}(S) := \frac{\widetilde{\mathrm{cut}}(S)}{\min\left\{\widetilde{\mathrm{vol}}(S), \widetilde{\mathrm{vol}}(S^c)\right\}},\ \widetilde{\mathrm{cut}}(S) = \left|\left\{(i,j) \in \widetilde{E}_n : x_i \in S, x_j \notin S\right\}\right|,\ \widetilde{\mathrm{vol}}(S) = \left|\left\{(i,j) \in \widetilde{E}_n : x_i \in S\right\}\right|$$

where in this context $S^c = \widetilde{\mathbf{X}}_n \setminus S$ denotes the complement of S within $\widetilde{G}_{n,r}$.

Convergence under the TL^1 metric Let

$$\widetilde{\mathbb{P}}(\mathcal{S}) = \frac{\mathbb{P}(\mathcal{S})}{\mathbb{P}(\mathcal{C}_\sigma)},\ \widetilde{\mathbb{P}}_n(\mathcal{S}) := \frac{1}{\widetilde{n}} \sum_{x_i \in \widetilde{\mathbf{X}}_n} \mathbf{1}(x_i \in \mathcal{S}) \qquad (\mathcal{S} \in \mathfrak{B}(\mathcal{C}_\sigma))$$

be the (empirical) probability measures, conditional on $x \sim \mathbb{P}$ lying within $\mathcal{C}_\sigma$ (Here $\mathfrak{B}(\mathcal{C}_\sigma)$ is the Borel σ-algebra of $\mathcal{C}_\sigma$). A Borel map $T : \mathcal{C}_\sigma \to \widetilde{\mathbf{X}}_n$ is said to be a *transportation map* between $\widetilde{\mathbb{P}}$ and $\widetilde{\mathbb{P}}_n$ if for arbitrary $\mathcal{S} \in \mathfrak{B}(\mathcal{C}_\sigma)$,

$$\widetilde{\mathbb{P}}(\mathcal{S}) = \widetilde{\mathbb{P}}_n(T(\mathcal{S})).$$

The following lemma shows that, under suitable conditions, transportation maps convergence to the identity mapping at rate $\left(\frac{\log n}{n}\right)^{1/d}$.

Lemma 1 (Adaptation of Proposition 5 of Garcia Trillos 2016)**.** *With probability one, there exists a sequence of transportation maps $\{T_n\}_{n\in\mathbb{N}}$, $T_n : \mathcal{C}_\sigma \to \widetilde{\mathbf{X}}_n$ such that the following statement holds:*

$$\limsup_{n\to\infty} \frac{\widetilde{n}^{1/d}\left\|\mathrm{Id} - T_n\right\|_{L^\infty(\widetilde{\mathbb{P}})}}{(\log \widetilde{n})^{p_d}} \leq C$$

where $Id(x) = x$ is the identity mapping over $\mathcal{C}_\sigma$, C is a universal constant and $p_d = 3/4$ for $d = 2$ and $1/d$ for $d \geq 3$.

If a sequence of transportation maps $\{T_n\}_{n\in\mathbb{N}}$ satisfies $\|\mathrm{Id} - T_n\|_{L^1(\widetilde{\mathbb{P}})} = o_P(1)$, we refer to it as a sequence of *stagnating transportation maps*. Lemma 1 establishes that with probability one, such a sequence of stagnating transportation maps will exist.

Definition 0.1. For a sequence $\{u_n\}_{n\in\mathbb{N}} \subseteq L^1(\widetilde{\mathbb{P}}_n)$ and $u \in L^1(\widetilde{\mathbb{P}})$, we say that $\{u_n\}_{n\in\mathbb{N}}$ converges TL^1 to u if there exists a sequence of stagnating transportation maps $\{T_n\}_{n\in\mathbb{N}}$ such that

$$d^{TL^1}(u, u_n) := \int_{\mathcal{C}_\sigma} |u(x) - u_n \circ T_n(x)|\, d\widetilde{\mathbb{P}}(x) \overset{n}{\to} 0 \tag{1}$$

and denote it $u_n \overset{TL^1}{\to} u$.

Remark 1. Note that this definition does not make sense on its face, as u and u_n lie in different spaces. Technically, we can resolve this by writing

$$d^{TL^1}((\widetilde{\mathbb{P}}, u), (\widetilde{\mathbb{P}}_n, u_n)) = \inf_{\pi\in\Gamma(\widetilde{\mathbb{P}},\widetilde{\mathbb{P}}_n)} \iint_{\mathcal{C}_\sigma\times\mathcal{C}_\sigma} |x-y| + |f(x) - g(y)|\, d\pi(x,y) \tag{2}$$

where γ is the space of couplings over the measures $\widetilde{\mathbb{P}}, \widetilde{\mathbb{P}}_n$. However, it can be shown that (2) converges to zero if and only if (1) is satisfied and

$$\widetilde{\mathbb{P}}_n \overset{w}{\to} \widetilde{\mathbb{P}}.$$

Since this additional condition will clearly be satisfied with probability one, we simplify things by hereafter referring only to the condition in (1). See Garcia Trillos 15 for more details.

0.1 Desired result.

Consider a sequence $S_1, S_2, \ldots, S_n, \ldots$ of sets with $S_n \subseteq \widetilde{\mathbf{X}}_n$ for all n, with characteristic functions $u_n : \widetilde{\mathbf{X}}_n \to \{0,1\}, u_n(x_i) = \mathbf{1}(x_i \in S_n)$ for $i \in \{1, \ldots, n\}$. We have already established the behavior of normalized cut when the sequence $\{u_n\}_{n\in\mathbb{N}}$ converges TL^1 to some $u \in L^1(\widetilde{\mathbb{P}})$. We require a complementary statement, of the form of Lemma 2.

Lemma 2 (Precompactness)**.** *Let $\{S_n\}_{n\in\mathbb{N}}$ and $\{u_n\}_{n\in\mathbb{N}}$ be as in the preceding paragraph. Suppose u_n does not converge TL^1 to any $u \in L^1(\mathbb{P})$. Then, with probability one:*

$$\liminf_{n\to\infty} \widetilde{\Phi}_{n,r}(S_n) \geq c_d r^1 \omega_r(1)$$

where c_d is a constant which does not depend on r but may depend on d, λ, σ and f, and $\omega_r(1)$ represents a term which goes to infinity as $r \to 1$.

However, we wish to replace the asymptotic term $\omega_r(1)$ of Lemma 2 with an explicit function of r.

1 Supporting Theory.

Definition 1.1. For a sequence $\{y_n\}_{n\in\mathbb{N}}$ over a metric space Y equipped with metric d_Y, the sequence $\{y_n\}_{n\in\mathbb{N}}$ is *precompact* if every subsequence $(y_{n_k})_{k\in\mathbb{N}}$ has an accumulation point in Y.

For $u_n \in L^1(\widetilde{\mathbb{P}}_n)$, let the *graph total variation* be given by

$$GTV_{n,r_n}(u_n) = \frac{1}{n^2 r_n^{d+1}} \sum_{x_i\in\widetilde{\mathbf{X}}_n} \sum_{x_j\in\widetilde{\mathbf{X}}_n} \mathbf{1}\big(\|x_i - x_j\| \leq r\big)\, |u_n(x_i) - u_n(x_j)|$$

To see the relation between GTV and $\widetilde{\Phi}$, introduce the balance term

$$B_n(u_n) = \min_{m\in\mathbb{R}} \frac{1}{\widetilde{\mathrm{vol}}(\widetilde{\mathbf{X}}_n)} \sum_{x_i\in\widetilde{\mathbf{X}}_n} \widetilde{\deg}(u_n)\,|u_n(x_i) - m|\,.$$

Note that if $u_n(x) = \mathbf{1}(x \in S_n)$ is the characteristic function for some $S_n \subseteq \widetilde{\mathbf{X}}_n$, then (assuming S_n is not the empty set)

$$B_n(u_n) = \frac{\min\left\{\widetilde{\mathrm{vol}}(S_n), \widetilde{\mathrm{vol}}(S_n^c)\right\}}{\widetilde{\mathrm{vol}}(\widetilde{\mathbf{X}}_n)}, \text{ and } GTV_{n,r_n}\left(\frac{u_n}{B_n(u_n)}\right) = \frac{\mathrm{vol}(\widetilde{\mathbf{X}}_n)}{n^2 r^{d+1}} \widetilde{\Phi}_{n,r_n}(S_n).$$

We introduce an energy functional E_n as shorthand for $GTV_{n,r_n}\left(\frac{u_n}{B_n(u_n)}\right)$:

$$E_n(v_n) := \begin{cases} GTV_{n,r_n}(v_n), & \text{if there is } S_n \subseteq \widetilde{X}_n, u_n(x) = 1(x\in S_n) \text{ with } B_n(u_n) > 0 \text{ such that } v_n = \frac{u_n}{B_n(u_n)} \\ \infty, & \text{otherwise} \end{cases}$$

Lemma 3 (Lemma 23 of Garcia Trillos 16)**.** *Let $\{r_n\}_{n\in\mathbb{N}}$ be a sequence of positive numbers converging to 0 and satisfying*

$$\lim_{n\to\infty} \frac{(\log n)^{p_d}}{n^{1/d}} \frac{1}{r_n} = 0$$

where $p_d = 3/4$ if $d = 2$ and $1/d$ for $d \geq 3$.

With probability one the following statement holds: for any sequence $\{v_n\}_{n\in\mathbb{N}}$ with $v_n \in L^1(\widetilde{\mathbb{P}}_n)$, if

$$\limsup_{n\to\infty} E_n(v_n) < \infty$$

then v_n is precompact in TL^1.

Remark 2. As outlined in Remark 1, to be technically precise about TL^1 convergence we need to work with distances between $(\widetilde{\mathbb{P}}, v)$ and $(\widetilde{\mathbb{P}}_n, v_n)$. Here, when we say "$v_n$ is precompact in TL^1" we mean that for every subsequence $(v_{n_k})_{k\in\mathbb{N}}$ there exists some $v \in L^1(\mathbb{P})$ such that $(\widetilde{\mathbb{P}}, v)$ is an accumulation point of $\big((\widetilde{\mathbb{P}}_{n_k}, v_{n_k})_{k\in\mathbb{N}}\big)$ with respect to TL^1 distance.

Lemma 3 builds straightforwardly from Lemma 4. Lemma 4 makes a similar statement to Lemma 3 but with respect to the graph total variation functional, meaning it does not take into account the balance term.

Lemma 4 (Theorem 1.2 of Garcia Trillos 16b)**.** *Let $\{r_n\}_{n\in\mathbb{N}}$ be a sequence of positive numbers converging to 0 and satisfying*

$$\lim_{n\to\infty} \frac{(\log n)^{p_d}}{n^{1/d}} \frac{1}{r_n} = 0 \tag{3}$$

where $p_d = 3/4$ if $d = 2$ and $1/d$ for $d \geq 3$. Consider a sequence of functions $u_n \in L^1(\widetilde{\mathbb{P}}_n)$. If

$$\sup_{n\in\mathbb{N}} \|u_n\|_{L^1(\widetilde{\mathbb{P}}_n)} < \infty \tag{4}$$

and

$$\sup_{n\in\mathbb{N}} GTV_{n,r_n}(u_n) < \infty \tag{5}$$

then $\{u_n\}_{n\in\mathbb{N}}$ is precompact in TL^1.

We will apply Lemma 4 to the sequence $\{v_n\}_{n\in\mathbb{N}}$ specified in the setup of Lemma 3. To do so, we will need to show $\sup_{n\in\mathbb{N}} \|v_n\|_{L^1(\widetilde{\mathbb{P}}_n)} < \infty$ is implied by $\limsup_{n\to\infty} E_n(v_n) < \infty$. This holds thanks to the gamma convergence of the functionals E_n to a continuous analogue. For $u \in L^1(\widetilde{\mathbb{P}})$, recall that the total variation of u is given by

$$TV(u) := \sup\left\{\int_{\mathcal{C}_\sigma} u(x)\mathrm{div}(\Psi(x))dx : \Psi \in C_c^1(\mathcal{C}_\sigma : \mathbb{R}^d), |\Psi(x)| \le f^2(x)\right\}$$

where $C_c^1(\mathcal{C}_\sigma : \mathbb{R}^d)$ represents the set of C^1-functions from $\mathcal{C}_\sigma$ to $\mathbb{R}^d$ whose support is compactly contained in $\mathcal{C}_\sigma$.

For $u \in L^1(\widetilde{\mathbb{P}})$, let $B(u)$ be a balance term

$$B(u) := \min_{m\in\mathbb{R}} \frac{\int_{\mathcal{C}_\sigma} |u(x) - m|\, f^2(x)dx}{\int_{\mathcal{C}_\sigma} f^2(x)dx}.$$

Then, analogously to the discrete case, let the continuous energy functional $E(v)$ be given by

$$E(v) := \begin{cases} TV(v), & \text{if there exists } \mathcal{S} \subseteq \mathcal{C}_\sigma, u(x) = \mathbf{1}(x \in \mathcal{S}) \text{ with } B(u) > 0 \text{ such that } v = \frac{u}{B(u)} \\ \infty, & \text{otherwise.} \end{cases}$$

Lemma 5 (Proposition 21 of Garcia Trillos 16b)**.** *For any $v \in L^1(\widetilde{\mathbb{P}})$ and any sequence $\{v_n\}_{n\in\mathbb{N}}$ with $v_n \in L^1(\widetilde{\mathbb{P}}_n)$ that converges to v in TL^1,*

$$c_1 E(v) \le \liminf_{n\to\infty} E_n(v_n)$$

where c_1 is a universal constant.

2 Proofs

2.1 Proof of Lemma 3

By Lemma 4, it is sufficient to show that (4) and (5) are satisfied for $v_n = \frac{1}{B_n(u_n)} u_n$, where $u_n(x) = \mathbf{1}(x \in S_n)$ is the characteristic function for S_n. Of course, as we have noted

$$GTV_{n,r_n}\left(\frac{u_n}{B_n(u_n)}\right) = \frac{\mathrm{vol}(\widetilde{\mathbf{X}}_n)}{n^2 r^{d+1}} \widetilde{\Phi}_{n,r_n}(S_n)$$

and therefore by hypothesis (5) holds for v_n. We turn now to showing (4).

To begin, let

$$w_n := \begin{cases} v_n, & \text{if } \mathrm{vol}(S_n) \le \mathrm{vol}(S_n^c) \\ (1 - u_n)/B_n(u_n), & \text{if } \mathrm{vol}(S_n) > \mathrm{vol}(S_n^c). \end{cases}$$

Note that

$$\|u_n\|_{L^1(\mathbb{P}_n)} = \frac{|S_n|}{\widetilde{n}} \le \frac{\widetilde{\mathrm{vol}}(S_n)}{\widetilde{d}_{\min}\widetilde{n}}$$

and similarly $\|1 - u_n\|_{L^1(\mathbb{P}_n)} \le \frac{\widetilde{\mathrm{vol}}(S_n^c)}{\widetilde{d}_{\min}\widetilde{n}}$. Therefore

$$\|w_n\|_{L^1(\mathbb{P}_n)} \le \frac{\min\left\{\widetilde{\mathrm{vol}}(S_n), \widetilde{\mathrm{vol}}(S_n^c)\right\}}{\widetilde{d}_{\min}\widetilde{n}B_n(u_n)} \le \frac{\widetilde{\mathrm{vol}}(\widetilde{\mathbf{X}}_n)}{\widetilde{n}\widetilde{d}_{\min}} \le \frac{\widetilde{d}_{\max}}{\widetilde{d}_{\min}}$$

There exists constant c_σ which depends only on σ such that

$$\limsup_{n\to\infty} \frac{\widetilde{d}_{\max}}{\widetilde{d}_{\min}} \le c_\sigma \frac{\Lambda_\sigma}{\lambda_\sigma} < \infty.$$

Moreover, $GTV_{n,r_n}(w_n) = GTV_{n,r_n}(v_n)$. Therefore, by Lemma 4, w_n is precompact in TL^1, meaning any subsequence of w_n has a further convert subsequence; let $w_{n_k} \overset{TL^1}{\to} w$ denote this convergent subsequence.

Therefore by Lemma 5,

$$\infty > \liminf_{n\to\infty} E_n(w_{n_k}) \ge c_1 E(w)$$

which in turn implies that $B(w) > 0$. But then

$$\begin{aligned}
\|v_{n_k}\|_{L^1(\widetilde{\mathbb{P}})} &= \frac{1}{B_n(u_{n_k})} \|u_{n_k}\|_{L^1(\widetilde{\mathbb{P}})} \\
&\le \frac{1}{B_n(u_{n_k})} \\
&= \frac{1}{B_n(w_{n_k})} \to \frac{1}{B(w)} < \infty
\end{aligned}$$

where the proof of convergence $B_n(w_{n_k}) \to B(w)$ is omitted but is straightforward. So every subsequence of $\{v_n\}_{n\in\mathbb{N}}$ has a further subsequence $(v_{n_k})_{k\in\mathbb{N}}$ which satisfies (4) and (5) and therefore is precompact in TL^1.

2.2 Proof of Lemma 4

By Lemma 1, there exists a sequence of transportation maps $\{T_n\}_{n\in\mathbb{N}}$ such that for a.e. $z, y \in \mathcal{C}_\sigma$ with $\|T_n(z) - T_n(x)\| > r_n$,

$$\|z - y\| > r_n - 2\|\mathrm{Id} - T_n\|_\infty =: \widetilde{r}_n.$$

and as an immediate implication for a.e. $z, y \in \mathcal{C}_\sigma$,

$$\|z - y\| \le \widetilde{r}_n \Longrightarrow \|T_n(z) - T_n(y)\| \le r_n$$

Note that by the lower bound on r_n implied by (3), $\frac{r_n}{\widetilde{r}_n} \to 1$; in particular, $\widetilde{r}_n$ will be positive for sufficiently large n with probability 1. Therefore for n sufficiently large,

$$\begin{aligned}
&\frac{1}{r_n^{d+1}} \iint_{\mathcal{C}_\sigma\times\mathcal{C}_\sigma} \mathbf{1}(\|z-y\| \le \widetilde{r}_n) \left| u_n\circ T_n(z) - u_n \circ T_n(y)\right| f(z) f(y) dz dy \\
&\qquad \le \frac{1}{r_n^{d+1}} \iint_{\mathcal{C}_\sigma\times\mathcal{C}_\sigma} \mathbf{1}(\|T_n(z) - T_n(y)\| \le r_n) \left| u_n\circ T_n(z) - u_n \circ T_n(y)\right| f(z) f(y) dz dy \\
&\qquad = GTV_{n,r_n}(u_n)
\end{aligned}$$

and by hypothesis $\limsup_{n\to\infty} GTV_{n,r_n}(u_n) < \infty$. But as $\frac{r_n}{\widetilde{r}_n} \to 1$, this implies

$$\limsup_{n\to\infty} \frac{1}{\widetilde{r}_n} \iint_{\mathcal{C}_\sigma\times\mathcal{C}_\sigma} \frac{\mathbf{1}(\|z-y\| \le \widetilde{r}_n)}{\widetilde{r}_n^d} \left| u_n\circ T_n(z) - u_n \circ T_n(y)\right| f(z) f(y) dz dy < \infty$$

and so by Proposition 1 $\{u_n \circ T_n\}_{n\in\mathbb{N}}$ is precompact in $L^1(\widetilde{\mathbb{P}})$, which is equivalent to $\{u_n\}_{n\in\mathbb{N}}$ being precompact in TL^1.

Proposition 1 – which we prove in Section 2.3 – is stated with respect to a nonlocal functional $TV_r(u)$, given by

$$TV_r(u) := \frac{1}{r} \iint_{\mathcal{C}_\sigma \times \mathcal{C}_\sigma} \frac{\mathbf{1}(\|x-y\| \le r)}{r^d} |u(x) - u(y)| f(x)f(y)dxdy$$

Proposition 1. *Let $\{v_{r_n}\}_{n\in\mathbb{N}}$ be a sequence in $L^1(\widetilde{\mathbb{P}})$ such that*

$$\sup_{n\in\mathbb{N}} \|v_{r_n}\|_{L^1(\widetilde{\mathbb{P}})} < \infty$$

and

$$\sup_{n\in\mathbb{N}} TV_{r_n}(v_{r_n}) < \infty.$$

Then, $\{v_{r_n}\}_{n\in\mathbb{N}}$ is precompact in $L^1(\widetilde{\mathbb{P}})$.

2.3 Proof of Proposition 1

Without loss of generality, let $f \equiv 1$ (since otherwise it is bounded above and below by positive constants over its support). We begin by extending each function v_{r_n} to $\mathbb{R}^d$. By assumption, $\forall x \in C^{2\sigma}$ there exists a unique closest point on ∂C^σ; let Px denote this point, and define the local reflection mapping

$$\widehat{x} = 2Px - x.$$

Letting $\xi(s)$ be a smooth function satisfying

$$\xi(s) = 1, s \le \frac{\sigma}{8}, \text{ and } \xi(s) = 0, s \ge \frac{\sigma}{4}$$

we define an auxiliary function

$$\widetilde{v}_{r_n}(x) = \xi(|x - Px|)v_{r_n}(\widehat{x})$$

and claim that for some constant C,

$$\sup_{n\in\mathbb{N}} \frac{1}{r_n} \iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{\mathbf{1}(\|x-y\| \le r)}{r^d} |\widetilde{v}_{r_n}(x) - \widetilde{v}_{r_n}(y)|\, dydx <$$
$$\le C \sup_{n\in\mathbb{N}} \left(\frac{1}{r_n} \iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{\mathbf{1}(\|x-y\| \le r)}{r^d} |v_{r_n}(x) - v_{r_n}(y)|\, dydx + \|v_{r_n}\|_{L^1(\widetilde{\mathbb{P}})} \right)$$

which by hypothesis is less than ∞.

3 Additional Notation

- $c_1 = \int_{B(0,1)} |\langle x, e_1 \rangle|\, dx$ where $B(0,1)$ is a d-dimensional unit ball centered at the origin.