Clustering

Lecturer: Xiuyuan Cheng   *Scribes: Dev Dabke and Andrew Cho*

1   Introduction

The lecture covered the following on the consistency of spectral clustering

- Setup of $\mathcal{L}_{un}$ and $\mathcal{L}_n$ and their limit operators U and U'

- First r spectral convergence

2   Consistency of Spectral Clustering

This section is a high-level overview of the ideas presented by von Luxburg, including the details of the proofs. [vLBB08]

2.1   The Problem Setup

Suppose $X_i \sim P$ where P is some distribution on $\Omega \subset \mathbb{R}^D$. W_{ij} is the affinity matrix where, as an example,

$$W_{ij} = e^{\frac{-|x_i - x_j|^2}{\varepsilon}}$$

where $\varepsilon > 0$.

As $n \to \infty$, we want to show the convergence of the graph Laplacian $\mathcal{L}$.

1. $\mathcal{L}_n = D - W$ where $D_{ij} = \sum_{j=1}^n W_{ij}$ ($\to U$), i.e. unnormalized

2. $\mathcal{L}'_n = D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}$ where ($\to U'$), i.e. symmetric

3. $\mathcal{L}''_n = D^{-1}(D - W) = I - P$, i.e. random walk

2.2   Limit operators U and U'

We construct linear limit operators U and U' on $C(X)$ which are the limit of the discrete operators $\mathcal{L}_n$ and $\mathcal{L}'_n$. We prove that the first "r" eigenvectors of the discrete operators converge to eigenfuction of the limit operators.

Definition 2.1: Limit Operators

We define U as

$$U : C(\Omega) \to C(\Omega)$$

$$Uf(x) = f(x)d(x) - \int k(x,y)f(y)dP(y)$$

---

where

$$dP(x) = p(x)dx$$

$$dx = \int k(x, y)dP(y)$$

$$x \in \Omega$$

Theorem 2.1: U'

$$U'f(x) = f(x) - \int \frac{k(x, y)}{\sqrt{d(x)}\sqrt{d(y)}}f(y)dP(y)$$

Proof 2.1: U', Theorem 2.1

$$
\begin{aligned}
(D^{-\frac{1}{2}}WD^{-\frac{1}{2}})_{ij} &= \frac{1}{\sqrt{D_{ii}}}W_{ij}\frac{1}{\sqrt{D_{jj}}} \\
&= \frac{\frac{1}{n}k(x_i, x_j)}{\sqrt{\frac{1}{n}\sum_{j'}k(x_i, x'_j)}\sqrt{\frac{1}{n}\sum_{j'}k(x_j, x'_j)}} \\
&\approx \frac{\frac{1}{n}k(x_i, x_j)}{\sqrt{d(x_i)}\sqrt{d(x_j)}}
\end{aligned}
$$

2.3 First r spectral convergence

Let's discuss what "first r spectral convergence" means. M_n first r spectral convergence to T if first r eigenvalues of M_n converge to those of T, and the associated eigenvectors converge to the eigenfunctions of T, the first smallest r eigenvalues.

Theorem 2.2: Convergence

For fixed $r > 0$, $n \to \infty$, and mild conditions,

1. (**Unnormalized**) $\mathcal{L}_n$ first r spectral converge to U if the first r eigenvalues of U lie outside of the range of the degree function $d(x)$. We need extra constraints for convergence since U might coincide with the range of $d(x)$.

2. (**Normalized Symmetric**) $\mathcal{L}'_n$ first r spectral converge to the operator U'.

Proof 2.2: Convergence, Theorem 2.2

(Case 2 $\mathcal{L}_{sym}$) Have M_n converge to T where

$$M_n : \mathbb{R}^n \to \mathbb{R}^n$$
$$T : C(\Omega) \to C(\Omega)$$
$$M_n : I - D^{\frac{-1}{2}} W D^{\frac{-1}{2}}$$
$$T = U'$$

$$Tf(x) = f(x) - \int h(x,y)f(y)dP(y)$$
$$= f(x) - \int h(x,y)f(y)dP_n(y)$$
$$= f(x) - \frac{1}{n}\sum_{i=1}^{n} h(x,x_i)f(x_i)$$

where

$$h(x,y) = \frac{k(x,y)}{\sqrt{d(x)}\sqrt{d(y)}}$$
$$dP_n(x) = \frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}(x)dx$$

Lemma 2.2: Spectral Equivalence between M_n and T_n

1. If $T_n\varphi = \lambda\varphi$, let $v \in \mathbb{R}^n, i = 1,...,n v_i = \varphi(x_i)$, then $M_n v = \lambda v$.

2. If $M_n v = \lambda v$ and $\lambda \neq 1$, then let $\varphi = \frac{\frac{1}{n}\sum h(x,x_j)v_j}{1-\lambda}$ and so then $T_n\varphi = \lambda\varphi$.

Lemma 2.2: Spectral Convergence

Replacing $dP_n(y)$ to be $dP(y)$, T_n spectral converges to T.

Proof 2.2: Spectral Convergence, Lemma 2.2

For all f, $T_n \to Tf$ by the Law of Large Numbers. $\|T_n f - Tf\|_{\inf} \to 0$ simultaneously for "sufficiently many" f such that for each eigenvalue of $T(\lambda \neq 1)$, the associated eigenvalue of T_n converge to λ and the associated eigenfunction of T_n converges. In other words, $T_n\varphi_n = \lambda_n\varphi_n$ so $\lambda_n \to$ asymptotically $T\varphi = \lambda\varphi$ and $\|\varphi_n - \varphi\|_\infty \to 0$

References

[vLBB08] Ulrike von Luxburg, Mikhail Belkin, and Olivier Bousquet. Consistency of spectral clustering. *Ann. Statist.*, 36(2):555–586, 04 2008.