Notes on 'The Geometry of Kernelized Spectral Clustering'

Alden Green

January 5, 2019

1 Analysis of normalized Laplacian embedding.

For a given set of distributions $\mathbb{P}_1, \ldots, \mathbb{P}_m$ and weights $w_1, \ldots, w_K$ in the probability simplex, define the **mixture distribution**

$$\bar{\mathbb{P}} := \sum_{m=1}^{K} w_m \mathbb{P}_m.$$

Given a non-negative, continuous, symmetric kernel function $k(x, y)$ and a distribution $\mathbb{P}$ we introduce the **square root kernelized density** as the function $q \in L^2(\mathbb{P})$ given by

$$q(x) := \sqrt{\int k(x, y) d\mathbb{P}(y)}.$$

In particular, we denote the square root kernelized density of the mixture distribution $\bar{P}$ by $\bar{q}$ and those of the mixture components $\{\mathbb{P}_m\}_{m=1}^K$ by $\{q_m\}_{m=1}^K$.

Normalized densities and the coupling parameter. Because we typically deal with the matrix $L = D^{-1/2} A D^{-1/2}$ when performing spectral embedding, it is useful to define analogous continuum operators. In particular, we define the **normalized kernel function** $\bar{k}$ to be

$$\bar{k}(x, y) = \frac{1}{\bar{q}(x)} k(x, y) \frac{1}{\bar{q}(y)}$$

and the normalized kernel for mixture component k_m to be

$$k_m(x, y) := \frac{k(x, y)}{q_m(x) q_m(y)} \quad \text{for } m = 1, \ldots, K.$$

Now, we introduce the **coupling parameter**

$$\mathcal{C}(\bar{P}) := \max_{m=1,\ldots,K} \left\| k_m - w_m \bar{k} \right\|_{\mathbb{P}_m \bigotimes \mathbb{P}_m}^2$$

This controls the maximum average connection between points generated by any one mixture component $\mathbb{P}_i$ and those generated by a different mixture component $\mathbb{P}_j$.