Hierarchical spectral clustering on scattering coefficients

Carmine-Emanuele Cella

École normale supérieure - Paris
carmine.emanuele.cella@ens.fr

Abstract

In this short note we will define a method to group together several vectors of scattering coefficients that represent quasi-stationary signals at different scales. The main idea consists in embedding the scattering coefficients into a diffusion map that can be further used for clustering. This process can be done recursively by splitting in halves the input signal, thus creating a full decomposition tree up to a given order.

I. Diffusion embedding

Let $x \in \mathbb{C}^d$ be an observation and let $\{\Phi x\}_{n \leq N}$ be the corresponding set of vectors of scattering coefficients of first and second order at all scales up to a given scale 2^J; the set of vectors is computed by patching x at scale 2^J, overlapping the patches by half. Let $\mathcal{F}_{i,i}$ be the *scattering flow* of $\{\Phi x\}_{n \leq N}$, a symmetric matrix whose entries are the sum of the modulus of the pairwise difference of the coefficients:

$$\mathcal{F}_{i,i} = \begin{pmatrix} \sum |\Phi x_1 - \Phi x_2| & \times & \times & \times & \sum |\Phi x_1 - \Phi x_i| \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & = & & \times & \times \\ & & & & \sum |\Phi x_i - \Phi x_i| \end{pmatrix}. \tag{1}$$

Let $\mathcal{A}_{i,i}$ be the adjacency matrix of $\mathcal{F}_{i,i}$, created by applying to it an affinity metric with a Gaussian kernel:

$$\mathcal{A}_{i,i} = e^{-\left(\frac{\mathcal{F}_{i,i}}{2\sigma^2}\right)} \tag{2}$$

where σ^2 is the variance of the kernel which determines the scale of the affinity metric; it depends on the nature of x and must be carefully set. In this context we decided to set this value equal to the centroid of the histogram of $\mathcal{F}_{i,i}$:

$$\sigma = \sum_h \left(\frac{h \cdot p(h)}{\sum_h p(h)} \right) \tag{3}$$

where $p(h)$ is the value of the histogram at position h. Let then $\mathcal{L}_{i,i}$ be the normalized *Laplacian matrix* of x defined as:

$$\mathcal{L}_{i,i} = \mathcal{D}_{i,i}^{-1/2} \mathcal{A}_{i,i} \mathcal{D}_{i,i}^{-1/2} \tag{4}$$

where $\mathcal{D}_{i,i} = \sum_j \mathcal{A}_{i,j}$ is the degree matrix.

The matrix $\mathcal{L}_{i,j}$ is symmetric and can be decomposed in principal eigenvalues $\{\lambda_l\}_{\lambda \leq \Lambda}$ and biorthogonal left and right eigenvectors $\{\psi_l\}_{\lambda \leq \Lambda}, \{\phi_l\}_{\lambda \leq \Lambda}$ respectively.

Due to the fast decay of the eigenvalues, only a few terms are necessary to achieve a given relative accuracy in the decomposition. Thus, using the first two largest eigenvectors it is possible to create a 2-dimensional space, called *diffusion embedding*, which is embedded in the original space. Given the low dimensionality of the this space, it is possible to apply a binary clustering to partition it into two clusters; the whole process (with minor variations) is often called *spectral clustering*.

With labels provided by the clustering, we then created two *streams* $\{S^1_p\}_{p \leq P}, \{S^2_q\}_{q \leq Q}$ by selecting the corresponding elements from $\{\Phi x\}_{n \leq N}$. Each stream has an independent time scale but contains information that can be considered quasi-stationary. The process is then recursively applied on each stream thus creating a *hierarchical* decomposition tree up to a given order W.

II. Kernel learning and signal-dependent representations

Once the vectors of scattering coefficients are consistently grouped together, it is possible to estimate a representative kernel per group in several ways. In this context we decided to take the maximum of each vector belonging to the stream, thus creating a set of coefficients whose length depends on the time scale of the stream. Given the logarithmic nature of the hierarchical decomposition, the total number of kernels $\{f^x_k\}_{k \leq K}$ is $K = 2^W$.

By convolving each kernel with the whole original set of vectors, it is possible to create a set of *feature maps* that creates a signal-dependent representation of x:

$$\tilde{\Phi} x = \{\Phi x * f^x_k\}_{k \leq K}. \tag{5}$$