Learning sheaf laplacian through direct imputation with KKT conditions

Leonardo Di Nino

May 2024

1 Introduction

Let's consider again the sheaf laplacian learning problem as a minimum total variation problem over the restriction maps:

$$\min_{e \in \mathcal{E}} \sum_{e \in \mathcal{E}} a_e ||\mathcal{F}_{u \triangleleft e} X_u - \mathcal{F}_{v \triangleleft e} X_v||^2 \tag{1}$$

where $a \in \{0, 1\}^{|\mathcal{E}|}$ and we retrieve the structure of the solution setting (or validating) $||a||_0 = q$, and X is the set of 0-cochains observed over the nodes.

We can control the structure of the solution avoiding the trivial one setting a constraint over $\mathcal{F}$ being the block diagonal matrix collecting all the restriction maps:

$$\begin{cases} \min_{e \in \mathcal{E}} \sum_{e \in \mathcal{E}} a_e ||\mathcal{F}_{u \triangleleft e} X_u - \mathcal{F}_{v \triangleleft e} X_v||^2 \\ \operatorname{tr}(\mathcal{F}) = N \end{cases}$$

The lagrangian associated to the problem is:

$$\mathcal{L}(\{\mathcal{F}_{u \triangleleft e}, \mathcal{F}_{v \triangleleft e}\}_{e \in \mathcal{E}}, \mu) = \sum_{e \in \mathcal{E}} a_e ||\mathcal{F}_{u \triangleleft e} X_u - \mathcal{F}_{v \triangleleft e} X_v||^2 + \mu[\operatorname{tr}(\mathcal{F}) - N] \tag{2}$$

The KKT system associated to this convex problem is:

$$\begin{cases} \nabla_{\mathcal{F}_{u \triangleleft e}} \mathcal{L}(\{\mathcal{F}_{u \triangleleft e}, \mathcal{F}_{v \triangleleft e}\}_{e \in \mathcal{E}}, \mu) = 0, & \forall e \in \mathcal{E} \\ \nabla_{\mathcal{F}_{v \triangleleft e}} \mathcal{L}(\{\mathcal{F}_{u \triangleleft e}, \mathcal{F}_{v \triangleleft e}\}_{e \in \mathcal{E}}, \mu) = 0, & \forall e \in \mathcal{E} \\ \operatorname{tr}(\mathcal{F}) = N \end{cases}$$

Expanding we get:

$$\begin{cases} (\mathcal{F}_{u \triangleleft e} X_u - \mathcal{F}_{v \triangleleft e} X_v) X_u^T + \mu \mathbb{I} = 0, & \forall e \in \mathcal{E} \\ -(\mathcal{F}_{u \triangleleft e} X_u - \mathcal{F}_{v \triangleleft e} X_v) X_v^T + \mu \mathbb{I} = 0, & \forall e \in \mathcal{E} \\ \operatorname{tr}(\mathcal{F}) = N \end{cases}$$

Focusing for a moment on the "local" solutions we get:

$$\begin{cases} \mathcal{F}_{u \triangleleft e} = (\mathcal{F}_{v \triangleleft e} X_v X_u^T - \mu \mathbb{I})(X_u X_u^T)^{-1}, & \forall e \in \mathcal{E} \\ \mathcal{F}_{v \triangleleft e} = (\mathcal{F}_{u \triangleleft e} X_u X_v^T - \mu \mathbb{I})(X_v X_v^T)^{-1}, & \forall e \in \mathcal{E} \end{cases}$$

After some math, we get $\forall e \in \mathcal{E}$:

$$\begin{cases} \mathcal{F}_{u \triangleleft e} = \mu\{[(X_u X_u^T)^{-1} X_u X_v^T - \mathbb{I}](X_v X_v^T)^{-1}[X_v X_u^T (X_u X_u^T)^{-1} X_u X_v^T (X_v X_v)^{-1} - \mathbb{I}]^{-1} X_v X_u^T - \mathbb{I}\}(X_u X_u^T)^{-1} \\ \mathcal{F}_{v \triangleleft e} = \mu[(X_u X_u^T)^{-1} X_u X_v^T - \mathbb{I}](X_v X_v^T)^{-1}[X_v X_u^T (X_u X_u^T)^{-1} X_u X_v^T (X_v X_v)^{-1} - \mathbb{I}]^{-1} \end{cases}$$

The two maps as expected are just related to the interactions between the signals over the nodes inciding on the edge:

$$\begin{cases} \mathcal{F}_{u \triangleleft e} = \mu \chi_{u \triangleleft e} \\ \mathcal{F}_{v \triangleleft e} = \mu \chi_{v \triangleleft e} \end{cases}$$

Now in order to retrieve the value for the lagrange multiplier we just use the global constraint imposing the structure to the solution, taking into account the linearity of the trace and the structure of $\mathcal{F}$:

$$\begin{aligned} \operatorname{tr}(\mathcal{F}) &= N \\ \sum_{e \in \mathcal{E}} [\operatorname{tr}(\mathcal{F}_{u \triangleleft e}) + \operatorname{tr}(\mathcal{F}_{v \triangleleft e})] &= N \\ \sum_{e \in \mathcal{E}} [\operatorname{tr}(\mu \chi_{u \triangleleft e}) + \operatorname{tr}(\mu \chi_{v \triangleleft e})] &= N \\ \mu \operatorname{tr}[\sum_{e \in \mathcal{E}} (\chi_{u \triangleleft e} + \chi_{v \triangleleft e})] &= N \\ \mu &= \frac{N}{\operatorname{tr}[\sum_{e \in \mathcal{E}} (\chi_{u \triangleleft e} + \chi_{v \triangleleft e})]} \end{aligned} \tag{3}$$

Substituting back we have:

$$\begin{cases} \mathcal{F}_{u \triangleleft e} = \frac{N}{\operatorname{tr}[\sum_{e \in \mathcal{E}} (\chi_{u \triangleleft e} + \chi_{v \triangleleft e})]} \chi_{u \triangleleft e} \\ \mathcal{F}_{v \triangleleft e} = \frac{N}{\operatorname{tr}[\sum_{e \in \mathcal{E}} (\chi_{u \triangleleft e} + \chi_{v \triangleleft e})]} \chi_{v \triangleleft e} \end{cases}$$