Machine Learning for Graphs and Sequential Data Exercise Sheet 13

Graphs: Semi-Supervised Learning

Label Propagation

Problem 1: The goal in Label Propagation is to find a labeling $\boldsymbol{y} \in \{0,1\}^N$ that minimizes the energy $\min_{\boldsymbol{y}} \frac{1}{2}\sum_{ij} \boldsymbol{w}_{ij}(y_i - y_j)^2$ subject to $y_i = \hat{y}_i \; \forall i \in S$ where the set of nodes V has been partitioned into the labeled nodes S and the unlabeled nodes U, $w_{ij} \geq 0$ is the non-negative edge weight and $\hat{y}_i$ are the observed labels.

Following from the first observation regarding the Laplacian, the minimization problem can be rewritten and then relaxed to $\min_{\boldsymbol{y}\in\mathbb{R}^N} \boldsymbol{y}^T\boldsymbol{L}\boldsymbol{y}$ subject to the same constraints. Show that the closed form solution is

$$\boldsymbol{y}_U = -\boldsymbol{L}_{UU}^{-1} \cdot \boldsymbol{L}_{US} \cdot \hat{\boldsymbol{y}}_S$$

where w.l.o.g. we assume that the Laplacian matrix is partitioned into blocks for labeled and unlabeled nodes as

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{L}_{SS} & \boldsymbol{L}_{SU} \\ \boldsymbol{L}_{US} & \boldsymbol{L}_{UU} \end{pmatrix}.$$

We begin by plugging the block partitioned form of $\boldsymbol{L}$ into the minimization term.

$$\begin{aligned}\boldsymbol{y}^T\boldsymbol{L}\boldsymbol{y} &= \boldsymbol{y}^T \begin{pmatrix} \boldsymbol{L}_{SS} & \boldsymbol{L}_{SU} \\ \boldsymbol{L}_{US} & \boldsymbol{L}_{UU} \end{pmatrix} \boldsymbol{y} \\ &= \hat{\boldsymbol{y}}_S^T \boldsymbol{L}_{SS}\hat{\boldsymbol{y}}_S + \hat{\boldsymbol{y}}_S^T\boldsymbol{L}_{SU}\boldsymbol{y}_U + \boldsymbol{y}_U^T\boldsymbol{L}_{US}\hat{\boldsymbol{y}}_S + \boldsymbol{y}_U\boldsymbol{L}_{UU}\boldsymbol{y}_U\end{aligned}$$

The laplacian is symmetric and therefore $\boldsymbol{L}_{US} = \boldsymbol{L}_{SU}^T$.

$$= \hat{\boldsymbol{y}}_S^T \boldsymbol{L}_{SS}\hat{\boldsymbol{y}}_S + 2\boldsymbol{y}_U^T\boldsymbol{L}_{US}\hat{\boldsymbol{y}}_S + \boldsymbol{y}_U\boldsymbol{L}_{UU}\boldsymbol{y}_U =: f(\boldsymbol{y}_U)$$

We can find the minimizer of f by finding the root of its first derivative with respect to $\boldsymbol{y}_U$ because f is quadratic.

$$\frac{\partial f}{\partial \boldsymbol{y}_U} = 2\boldsymbol{L}_{US}\hat{\boldsymbol{y}}_S + \left(\boldsymbol{L}_{UU} + \boldsymbol{L}_{UU}^T\right)\boldsymbol{y}_U = 2\boldsymbol{L}_{US}\hat{\boldsymbol{y}}_S + 2\boldsymbol{L}_{UU}\boldsymbol{y}_U = 0 \Leftrightarrow \boldsymbol{y}_U = -\boldsymbol{L}_{UU}^{-1} \cdot \boldsymbol{L}_{US} \cdot \hat{\boldsymbol{y}}_S$$

PPNP

Problem 2: The iterative equation of PPNP is given by

$$\boldsymbol{H}^{(l+1)} = (1-\alpha)\hat{\boldsymbol{A}}\boldsymbol{H}^{(l)} + \alpha\boldsymbol{H}^{(0)}$$

where $\hat{\boldsymbol{A}} = \tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}$ is the propagation matrix. Derive the closed form solution for infinitely many propagation steps.

Hint: If we have for a matrix $\boldsymbol{T}$ that all its eigenvalues λ are strictly between -1 and 1, an equivalent matrix formulation of the geometric series formula holds and

$$\sum_{k=0}^{\infty} \boldsymbol{T}^k = (\boldsymbol{I} - \boldsymbol{T})^{-1}.$$

Hint: The eigenvalues λ_i of the normalized Laplacian $\boldsymbol{L} = \boldsymbol{I} + \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{D}^{-\frac{1}{2}}$ are $0 \leq \lambda_i \leq 2$.

We start with $H^{(1)}$ and expand for a few steps.

$$\boldsymbol{H}^{(1)} = (1-\alpha)\hat{\boldsymbol{A}}\boldsymbol{H}^{(0)} + \alpha\boldsymbol{H}^{(0)}$$

$$\begin{aligned}
\boldsymbol{H}^{(2)} &= (1-\alpha)\hat{\boldsymbol{A}}\boldsymbol{H}^{(1)} + \alpha\boldsymbol{H}^{(0)} \\
&= (1-\alpha)\hat{\boldsymbol{A}}\left((1-\alpha)\hat{\boldsymbol{A}}\boldsymbol{H}^{(0)} + \alpha H^{(0)}\right) + \alpha\boldsymbol{H}^{(0)} \\
&= (1-\alpha)^2\hat{\boldsymbol{A}}^2\boldsymbol{H}^{(0)} + (1-\alpha)\hat{\boldsymbol{A}}\alpha\boldsymbol{H}^{(0)} + \alpha\boldsymbol{H}^{(0)}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{H}^{(3)} &= (1-\alpha)\hat{\boldsymbol{A}}\boldsymbol{H}^{(2)} + \alpha\boldsymbol{H}^{(0)} \\
&= (1-\alpha)\hat{\boldsymbol{A}}\left((1-\alpha)^2\hat{\boldsymbol{A}}^2\boldsymbol{H}^{(0)} + (1-\alpha)\hat{\boldsymbol{A}}\alpha\boldsymbol{H}^{(0)} + \alpha\boldsymbol{H}^{(0)}\right) + \alpha\boldsymbol{H}^{(0)} \\
&= (1-\alpha)^3\hat{\boldsymbol{A}}^3\boldsymbol{H}^{(0)} + (1-\alpha)^2\hat{\boldsymbol{A}}^2\alpha\boldsymbol{H}^{(0)} + (1-\alpha)\hat{\boldsymbol{A}}\alpha\boldsymbol{H}^{(0)} + \alpha\boldsymbol{H}^{(0)}
\end{aligned}$$

We can see the following pattern emerge.

$$\boldsymbol{H}^{(k)} = \left((1-\alpha)\hat{\boldsymbol{A}}\right)^k\boldsymbol{H}^{(0)} + \left(\sum_{i=0}^{k-1}\left((1-\alpha)\hat{\boldsymbol{A}}\right)^i\right)\alpha\boldsymbol{H}^{(0)}$$

If we let k grow to infinty, the first term converges to 0 because $\alpha \in (0,1)$ and in the second term we can apply the geometric series formula to get

$$\boldsymbol{H}^{(\infty)} = \alpha\left(\boldsymbol{I} - (1-\alpha)\hat{\boldsymbol{A}}\right)^{-1}\boldsymbol{H}^{(0)}$$

as the closed form solution as long as the eigenvalues of $(1-\alpha)\hat{\boldsymbol{A}}$ are strictly between -1 and 1.

We know from the hint that the eigenvalues of the normalized Laplacian $\boldsymbol{L} = \boldsymbol{I} + \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{D}^{-\frac{1}{2}}$ for any graph structure A are in $[0,2]$. So it is also true for the amended graph structure with self-loops $\hat{\boldsymbol{A}}$ where $\boldsymbol{L} = \boldsymbol{I} + \hat{\boldsymbol{A}}$. Let λ be an eigenvalue of $\hat{\boldsymbol{A}}$ with eigenvector $\boldsymbol{v}$.

$$(1+\lambda)\boldsymbol{v} = \boldsymbol{v} + \hat{\boldsymbol{A}}\boldsymbol{v} = \left(\boldsymbol{I} + \hat{\boldsymbol{A}}\right)\boldsymbol{v} = \boldsymbol{L}\boldsymbol{v}$$

So $1+\lambda$ is also an eigenvector $\boldsymbol{L}$ and must therefore be in $[0,2]$. Consequently, the eigenvalues of $\hat{\boldsymbol{A}}$ are in $[-1,1]$ and the eigenalues of $(1-\alpha)\hat{\boldsymbol{A}}$ are in $(-1,1)$ because $0 < (1-\alpha) < 1$.