STAT 745 – Fall 2014
Assignment 12

Group 1: Francene Cicia, Doug Raffle, Melissa Smith

December 8, 2014

a. Graph Based Learning

In graph based learning, we define and adjacency graph and response as:

$$A = \begin{bmatrix} A_{LL} & A_{LU} \\ A_{UL} & A_{UU} \end{bmatrix}, \quad Y = \begin{bmatrix} Y_L \\ Y_U \end{bmatrix}$$

where Y_U is missing. Let $S = D^{-1}A$, where D is the diagonal row sum matrix. So we have that the Laplcian matrix and D can be expressed as:

$$D = \begin{bmatrix} D_{LL} + D_{LU} & 0 \\ 0 & D_{UU} + D_{UL} \end{bmatrix}, \quad \Delta = \begin{bmatrix} \Delta_{LL} & \Delta_{LU} \\ \Delta_{UL} & \Delta_{UU} \end{bmatrix} = \begin{bmatrix} D_{LL} + D_{LU} - A_{LL} & -A_{LU} \\ -A_{UL} & D_{UU} + D_{UL} - A_{UU} \end{bmatrix}$$

From this, we employ propagation to fit $\hat{f}$. We have shown previously that $\hat{f}_L = S_{LL}Y_L + S_{LU}(I - S_{UU})^{-1}S_{UL}Y_L = M_{LL}Y_L$ and $\hat{f}_U = (I - S_{UU})^{-1}S_{UL}Y_L$. Let T_{LL} be the labeled only smoother. Also, define matrix Q and S_{LUL} so that $M_{LL} = QT_{LL} + (I - Q)S_{LUL}$.

b. Exact form of T_{LL}

Notice, we know that $S = D^{-1}A$. Using the matrices defined above, we can solve for the partitions of S.

$$S = D^{-1}A = \begin{bmatrix} (D_{LL} + D_{LU})^{-1} & 0 \\ 0 & (D_{UU} + D_{UL})^{-1} \end{bmatrix} \begin{bmatrix} A_{LL} & A_{LU} \\ A_{UL} & A_{UU} \end{bmatrix} = \begin{bmatrix} (D_{LL} + D_{LU})^{-1}A_{LL} & (D_{LL} + D_{LU})^{-1}A_{LU} \\ (D_{UU} + D_{UL})^{-1}A_{UL} & (D_{UU} + D_{UL})^{-1}A_{UU} \end{bmatrix}$$

We know that the supervised only smoother is S_{LL}, from the above matrix $S_{LL} = (D_{LL} + D_{LU})^{-1}A_{LL} = D_{LL}^{-1}A_{LL}$. Thus, $T_{LL} = D_{LL}^{-1}A_{LL}$.

c. Finding $\widetilde{f}_U$

We have that $f_L = S_{LL}Y_L$. To estimate the responses for the unlabeled nodes, we need to use the information we have that connects the labeled with the unlabeled cases. We connect these cases with A_{UL} (and hence the smoother S_{UL}), so it follows that $\widetilde{f}_U$ has the form:

$$\widetilde{f}_U = S_{UL}Y_L$$

d. Solving for Q and S_{LUL}

From $\hat{f}_L$, we can start with the fact that $M_{LL} = S_{LL} + S_{LU}(I - S_{UU})^{-1}S_{UL}$. So, we can start by putting M_{LL} in terms of partitions of D and A:

$$
\begin{aligned}
M_{LL} =& S_{LL} + S_{LU}\left(I - S_{UU}\right)^{-1} S_{UL} \\
=& \left(D_{LL} + D_{LU}\right)^{-1} A_{LL} + \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \left(I - \left(D_{UU} + D_{UL}\right)^{-1} A_{UU}\right)^{-1} \left(D_{UU} + D_{UL}\right)^{-1} A_{UL} \\
=& \left(D_{LL}\left(I - D_{LL}^{-1} D_{LU}\right)\right)^{-1} A_{LL} + \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \left(I - \left(D_{UU} + D_{UL}\right)^{-1} A_{UU}\right)^{-1} \left(D_{UU} + D_{UL}\right)^{-1} A_{UL} \\
=& \left(I + D_{LL}^{-1} D_{LU}\right)^{-1} D_{LL}^{-1} A_{LL} + \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \left(\left(D_{UU} + D_{UL}\right)^{-1} \left(\left(D_{UU} + D_{UL}\right) - A_{UU}\right)\right)^{-1} \left(D_{UU} + D_{UL}\right)^{-1} A_{UL} \\
=& (I + D_{LL}^{-1} D_{LU})^{-1} D_{LL}^{-1} A_{LL} + \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \left(\left(D_{UU} + D_{UL}\right)^{-1} \Delta_{UU}\right)^{-1} \left(D_{UU} + D_{UL}\right)^{-1} A_{UL} \\
=& \left(I + D_{LL}^{-1} D_{LU}\right)^{-1} D_{LL}^{-1} A_{LL} + \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \Delta_{UU}^{-1} \left(D_{UU} + D_{UL}\right) \left(D_{UU} + D_{UL}\right)^{-1} A_{UL} \\
=& \left(I + D_{LL}^{-1} D_{LU}\right)^{-1} D_{LL}^{-1} A_{LL} + \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \Delta_{UU}^{-1} A_{UL}
\end{aligned}
$$

We know that:

$$M_{LL} = S_{LL} + S_{LU}\left(I - S_{UU}\right)^{-1} S_{UL} = QT_{LL} + \left(I - Q\right) S_{LUL}$$

From the above derivation we found,

$$M_{LL} = \left(I + D_{LL}^{-1} D_{LU}\right)^{-1} D_{LL}^{-1} A_{LL} + \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \Delta_{UU}^{-1} A_{UL}$$

It follows that,

$$\left(I + D_{LL}^{-1} D_{LU}\right)^{-1} D_{LL}^{-1} A_{LL} + \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \Delta_{UU}^{-1} A_{UL} = QT_{LL} + \left(I - Q\right) S_{LUL}$$

From this, we can see

$$\left(I + D_{LL}^{-1} D_{LU}\right)^{-1} D_{LL}^{-1} A_{LL} = QT_{LL}$$

From above, we know that$T_{LL} = D_{LL}^{-1} A_{LL}$ so it follows that:

$$Q = \left(I + D_{LL}^{-1} D_{LU}\right)^{-1}$$

From the above equations for M_{LL} we also have,

$$
\begin{aligned}
\left(I - Q\right) S_{LUL} &= \left(D_{LL} + D_{LU}\right)^{-1} A_{LU} \Delta_{UU}^{-1} A_{UL} \\
&= \left(D_{LU}\left(D_{LU}^{-1} D_{LL} + I\right)\right)^{-1} A_{LU} \Delta_{UU}^{-1} A_{UL} \\
&= \left(D_{LU}^{-1} D_{LL} + I\right)^{-1} D_{LU}^{-1} A_{LU} \Delta_{UU}^{-1} A_{UL}
\end{aligned}
$$

In order for this to hold, we need that $(I - Q) = \left(D_{LU}^{-1} D_{LL} + I\right)^{-1}$. Notice, $(I - Q) + Q = I$, we can see if this holds for the $(I - Q)$ and Q that we found.

$$
\begin{aligned}
(I - Q) + Q &= \left(D_{LU}^{-1} D_{LL} + I\right)^{-1} + \left(I + D_{LL}^{-1} D_{LU}\right)^{-1} \\
&= \left(D_{LU}^{-1}\left(D_{LL} + D_{LU}\right)\right)^{-1} + \left(D_{LL}^{-1}\left(D_{LL} + D_{LU}\right)\right)^{-1} \\
&= \left(D_{LL} + D_{LU}\right)^{-1} \left(D_{LU} + D_{LL}\right) \\
&= I
\end{aligned}
\tag{1}
$$

Therefore, we know that:

$$S_{LUL} = D_{LU}^{-1} A_{LU} \Delta_{UU}^{-1} A_{UL}$$

Also, by the partitioning of the graphs above,

$$
\begin{aligned}
D_{LU}^{-1} A_{LU} \Delta_{UU}^{-1} A_{UL} &= D_{LU}^{-1}\left(-\Delta_{LU}\right) \Delta_{UU}^{-1}\left(-\Delta_{UL}\right) \\
&= D_{LU}^{-1} \Delta_{LU} \Delta_{UU}^{-1} \Delta_{UL}
\end{aligned}
$$

e. Showing S_{LUL} is right stochastic

We will begin by showing some important identities. By properties of Laplacian matrices, we know $\Delta\vec{1} = \vec{0}$ so,

$$\begin{bmatrix} \Delta_{LL} & \Delta_{LU} \\ \Delta_{UL} & \Delta_{UU} \end{bmatrix} \vec{1} \quad = \begin{bmatrix} D_{LL} + D_{LU} - A_{LL} & -A_{LU} \\ -A_{UL} & D_{UU} + D_{UL} - A_{UU} \end{bmatrix} \vec{1} \quad = \vec{0}$$

From above, we can see,

$$\begin{aligned} \Delta_{UU}\vec{1} + \Delta_{UL}\vec{1} = \vec{0} \\ \Delta_{UU}\vec{1} = -\Delta_{UL}\vec{1} \end{aligned}$$

Similarly, we have,

$$\begin{aligned} \Delta_{LL}\vec{1} + \Delta_{LU}\vec{1} &= 0 \\ (D_{LL} + D_{LU} - A_{LL})\,\vec{1} + \Delta_{LU}\vec{1} &= 0 \\ D_{LL}\vec{1} - A_{LL}\vec{1} + D_{LU}\vec{1} + \Delta_{LU}\vec{1} &= 0 \\ \Delta_{LU}\vec{1} &= -D_{LU}\vec{1} \end{aligned}$$

Notice the above holds because $D_{LL}\vec{1} = A_{LL}\vec{1}$.

In order to show S_{LUL} is right stochastic, we first need to show $S_{LUL}\vec{1} = \vec{1}$.

$$\begin{aligned} S_{LUL}\vec{1} &= \left(D_{LU}^{-1}\Delta_{LU}\Delta_{UU}^{-1}\Delta_{UL}\right)\vec{1} \\ &= D_{LU}^{-1}\vec{1}\Delta_{LU}\vec{1}\Delta_{UU}^{-1}\vec{1}\Delta_{UL}\vec{1} \\ &= D_{LU}^{-1}\vec{1}\,(-D_{LU})\,\vec{1}\Delta_{UU}^{-1}\vec{1}\,(-\Delta_{UU})\,\vec{1} \\ &= (-I)\,\vec{1}\,(-I)\,\vec{1} \\ &= \vec{1} \end{aligned} \tag{2}$$

In order for S_{LUL} to be right stochastic, $S_{LUL_{ij}} \geq 0$ for all i, j. We know that A is an adjacency matrix, so all of its elements, $A_{ij} \geq 0$. D is the diagonal row sum matrix of A, so $D_{ij} \geq A_{ij}$ for all i, j. This implies that $\Delta_{ij} \geq 0$ for all i, j. S_{LUL} is a product of matrices containing all non-negative elements, thus $S_{LUL_{ij}} \geq 0$ for all i, j. Thus, S_{LUL} is right stochastic.

f. Supervised vs. Semi-Supervised

If we have no edges connecting the labeled to unlabed cases, A_{UL} and hence S_{UL} would provide no information about Y_U and $\widetilde{f}_U$ would not exist. In this case supervised and semi-supervised techniques should be equivalent, since both cases would make the same predictions for the labeled cases and no predictions for the unlabeled.

In the case that the data is linearly seperable, the boundary found by both methods would similar and the results would be close.