Statistics 745

Lecture 12

1. In graph-based learning, we define the graph and response as:

$$A = \begin{pmatrix} A_{LL} & A_{LU} \\ A_{UL} & A_{UU} \end{pmatrix} \qquad Y = \begin{pmatrix} Y_L \\ Y_U \end{pmatrix}, \tag{1}$$

where specifically Y_U is missing. Let $S = D^{-1}A$ where D is the diagonal row sum matrix, i.e. $D = \text{diag}(A\mathbf{1})$. Define the partition diagonal matrices as: $D_{LL} = \text{diag}(A_{LL}\mathbf{1})$, $D_{LU} = \text{diag}(A_{LU}\mathbf{1})$, $D_{LU} = \text{diag}(A_{LU}\mathbf{1})$, and $D_{UU} = \text{diag}(A_{UU}\mathbf{1})$. From this,we have that the diagonal row sum matrix for D can be re-expressed as:

$$D = \begin{pmatrix} D_{LL} + D_{LU} & \mathbf{0} \\ \mathbf{0} & D_{UU} + D_{UL} \end{pmatrix}. \tag{2}$$

From this, we employ propagation to fit $\hat{f}$. As shown in class $\hat{f}_L = S_{LL}Y_L + S_{LU}(I - S_{UU})^{-1}S_{UL}Y_L \equiv M_{LL}Y_L$ and $\hat{f}_U = (I - S_{UU})^{-1}S_{UL}Y_L$. Denote T_{LL} as the labeled only smoother (supervised) specifically defined only from A_{LL}. Define a matrix Q and S_{LUL} so that

$$M_{LL} = QT_{LL} + (I - Q)S_{LUL}$$

(a) What is the exact form of T_{LL} in terms of D_{LL} and A_{LL}?

(b) Suppose that the matrix A were a kernel matrix ($A_{ij} = K_\lambda(x_i, x_j)$) then $\widetilde{f}_L = T_{LL}Y_L$ would be loess of an intercept. In terms of partitions in D and A what is the form of the predictions from supervised loess when applied to observations in U, i.e. what is $\widetilde{f}_U$?

(c) What is the form of Q in terms of partitions in D and/or A?

(d) Show that S_{LUL} is *right stochastic*. That is $S_{LUL}\mathbf{1} = \mathbf{1}$ and $S_{LUL_{ij}} \geq 0$ for all $i,\ j$.

(e) Describe a circumstance where the propagation result would be mathematically equivalent to the supervised one. How about when the supervised estimator is close to propagation?