MATH 5473 Homework6 LUO Yuanhui

3. *Nyström method:* In class, we have shown that every manifold learning algorithm can be regarded as Kernel PCA on graphs: (1) given N data points, define a neighborhood graph with N nodes for data points; (2) construct a positive semidefinite kernel K; (3) pursue spectral decomposition of K to find the embedding (using top or bottom eigenvectors). However, this approach might suffer from the expensive computational cost in spectral decomposition of K if N is large and K is non-sparse, e.g. ISOMAP and MDS.

To overcome this hurdle, Nyström method leads us to a scalable approach to compute eigenvectors of low rank matrices. Suppose that an N-by-N positive semidefinite matrix $K \succeq 0$ admits the following block partition

$$K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}. \tag{1}$$

where A is an n-by-n block. Assume that A has the spectral decomposition $A = U\Lambda U^T$, $\Lambda = \mathrm{diag}(\lambda_i)$ ($\lambda_1 \geq \lambda_2 \geq \ldots \lambda_k > \lambda_{k+1} = \ldots = 0$) and $U = [u_1, \ldots, u_n]$ satisfies $U^T U = I$.

(a) Assume that $K = XX^T$ for some $X = [X_1; X_2] \in \mathbb{R}^{N\times k}$ with the block $X_1 \in \mathbb{R}^{n\times k}$. Show that X_1 and X_2 can be decided by:

$$X_1 = U_k \Lambda_k^{1/2}, \tag{2}$$

$$X_2 = B^T U_k \Lambda_k^{-1/2}, \tag{3}$$

where $U_k = [u_1, \ldots, u_k]$ consists of those k columns of U corresponding to top k eigenvalues λ_i ($i = 1, \ldots, k$).

Proof: Since $A = U\Lambda U^T$, $\Lambda = \mathrm{diag}(\lambda_i)$, $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_k > \lambda_{k+1} = \ldots = 0$, we have $A = U_k \Lambda_k U_k^T$, $\Lambda_k = \mathrm{diag}(\lambda_1 \ldots \lambda_k)$

$$XX^T = \begin{pmatrix} X_1X_1^T & X_1X_2^T \\ X_2X_1^T & X_2X_2^T \end{pmatrix} = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

On the one hand, $X_1X_1^T = A = U_k\Lambda_k U_k^T = (U_k\Lambda_k^{\frac{1}{2}})(U_k\Lambda_k^{\frac{1}{2}})^T$,

then $X_1 = U_k \Lambda_k^{\frac{1}{2}}$

On the other hand, $X_2 X_1^T = X_2 \Lambda_k^{\frac{1}{2}} U_k^{-1} = B^T$, then

$X_2 = B^T U_k \Lambda_k^{-\frac{1}{2}}$

(b) Show that for general $K \succeq 0$, one can construct an approximation from (2) and (3),

$$\hat{K} = \begin{bmatrix} A & B \\ B^T & \hat{C} \end{bmatrix}. \tag{4}$$

where $A = X_1 X_1^T$, $B = X_1 X_2^T$, and $\hat{C} = X_2 X_2^T = B^T A^\dagger B$, $A^\dagger$ denoting the Moore-Penrose (pseudo-) inverse of A. Therefore $\|\hat{K} - K\|_F = \|C - B^T A^\dagger B\|_F$. Here the matrix $C - B^T A^\dagger B =: K/A$ is called the (generalized) *Schur Complement* of A in K.

Proof: Based on $X_1 = U_k \Lambda_k^{\frac{1}{2}}$, $X_2 = B^T U_k \Lambda_k^{-\frac{1}{2}}$ from (a), we can construct $\hat{K} = XX^T = \begin{pmatrix} X_1 X_1^T & X_1 X_2^T \\ X_2 X_1^T & X_2 X_2^T \end{pmatrix} = \begin{pmatrix} A & B \\ B^T & X_2 X_2^T \end{pmatrix}$

and $X_2 X_2^T = B^T U_k \Lambda_k^{-1} U_k^T B$

Let $G = U_k \Lambda_k^{-1} U_k^T$, then it's easy to verify that

$GAG = G$, $AGA = A$, $(AG)^T = I = AG$, $(GA)^T = I = GA$

By the definition of Moore-Penrose inverse, we have

$A^+ = G$ and it's unique

Therefore, $\|K - \hat{K}\|_F = \sqrt{\|A - A\|_F^2 + \|B - B\|_F^2 + \|B^T - B^T\|_F^2 + \|C - B^T A^+ B\|_F^2}$

$= \|C - B^T A^+ B\|_F$

(c) Explore Nyström method on the Swiss-Roll dataset (`http://yao-lab.github.io/data/swiss_roll_data.mat` contains 3D-data X; `http://yao-lab.github.io/data/swissroll.m` is the matlab code) with ISOMAP. To construct the block A, you may choose either of the following:

n random data points;

*n landmarks as minimax k-centers (`https://yao-lab.github.io/data/kcenter.m`);

Some references can be found at:

[dVT04] Vin de Silva and J. B. Tenenbaum, "Sparse multidimensional scaling using landmark points", 2004, downloadable at `http://pages.pomona.edu/~vds04747/public/papers/landmarks.pdf`;

[P05] John C. Platt, "FastMap, MetricMap, and Landmark MDS are all Nyström Algorithms", 2005, downloadable at `http://research.microsoft.com/en-us/um/people/jplatt/nystrom2.pdf`.

Please see the matlab codes in the folder.