# CMSE 820 HW10

This HW is due on Nov 28th at 11:59pm.

**Question 1**: Define local charts on $S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ using

(1) polar coordinate representation

(2) Stereographic projection.

**Question 2**: Write a function to perform LLE (Hint: when you solve Problem W, you make need ridge-regression type of trick to invert $C$ matrix) and download HW10dat.csv and HW10color.csv from D2L, where HW10dat.csv contains 2048 data points in $\mathbb{R}^3$ and HW10color.csv contains RGB color value for each data point. Do the following

- Plot the data in 3D together with colors.
- Perform PCA and reduce the data to 2D. Plot the data projected on the 2D subspace.
- Perform Kernel PCA using Gaussian Kernel. Pick the best $\sigma^2$ which can best unfold the data and plot the result.
- Perform LLE using $K = 12$ for $K$ nearest neighors and plot your result. Discuss the results between kernel PCA and LLE.

**Question 3**: Laplacian Eigenmaps (LE) is another useful nonlinear dimension reduction technique. To introduce LE, let us first define $\epsilon-$neighbours and weight matrix. Let $x_1, \ldots, x_n \in \mathbb{R}^p$ be the high dimensional data points. Fix some scalar $\epsilon > 0$, $x_i$ and $x_j$ are called $\epsilon-$neighbours of each other if and only if $\|x_i - x_j\|_2 \leq \epsilon$. Now fix another scalar $\sigma^2 > 0$, for any pair of data points $(x_i, x_j)$, we can define a weight $w_{i,j} = \exp(-\frac{\|x_i - x_j\|_2^2}{\sigma^2})$ if $x_i$ and $x_j$ are $\epsilon-$neighbours, and $w_{i,j} = 0$ otherwise. A reasonable low dimensional embedding $y_1, \ldots, y_n$ minimizes the following objective function

$$\sum_{i,j} w_{i,j} \|y_i - y_j\|^2$$

The exponential weight incurs a heavier penalty than the Euclidean weight if neighbouring points $(x_i, x_j)$ with small distance are mapped far apart. Therefore, minimizing this objective is an attempt to ensure if $x_i$ and $x_j$ are close, then $y_i, y_j$ should be close as well.

a. Prove that $\min \sum w_{i,j} \|y_i - y_j\|^2 = \min \operatorname{Tr}(YLY^T)$, where $Y = [y_1, \ldots, y_n]$ and $L = D - A$ with $A_{i,j} = w_{i,j}$ and $D$ being a diagonal matrix with $D_{i,i} = \sum_j w_{i,j}$. The matrix $L$ is called graph laplacian.

b. To prevent the optimization problem from returning trivial solutions, we add a constraint that normalizes the scaling of the coordinates of $Y$.

$$\min_{Y} \mathrm{Tr}(YLY^T) \text{ subject to } YDY^T = I,$$

Furthermore, we remove the arbitrary shift by adding a second constraint, which ensures $YD$ has 0 mean,

$$\hat{Y} = \arg\min_{Y} \mathrm{Tr}(YLY^T) \text{ subject to } YDY^T = I \text{ and } YD\mathbb{1} = 0. \tag{1}$$

Show that the solutions $\hat{Y} \in \mathbb{R}^{d\times n}$ to (1) are given by the eigenvectors corresponding to the lowest $d$ eigenvalues of the genearlized eigenvalue problem

$$Ly = \lambda Dy.$$

**Question 4**: In the derivation of LLE, we define

$$M = (I_N - W)^T(I_N - W).$$

Prove that $M\mathbf{1} = 0$.