Manifold Learning
(OM for ML-II)

(*As discussed in the last lecture, we will have some integrated learning modules where we use all the feasible components (like online meeting, internet resources, practices at your end, evaluations etc. into one thread.*)

In this integrated learning module, we look at the optimization problems that arise when one tries to visualize or cluster the data in higher dimension. Some of the related ideas came into the lectures when we talked out PCA and Spectral Methods in the past.

Many of the intuitions and ideas from our Euclidean space will not naturally extend a number of practical data sets/problems. For example, we asked the quesion: "Is a spring in 1D, 2D or 3D?" or "a folder paper is 2D or 3D ?". We can now understand these in more detail.

1. To start with see two excellent lecture videos from:

   (a) https://www.youtube.com/watch?v=RPjPLlGefzw

   (b) https://www.youtube.com/watch?v=V680Ev0MNvs

   This also overlaps with what we did in the class on spectral clustering and ratio cut. We also had seen eigen vector based optimization in the class. (please see your notes from OM for ML-I)

   At places, you have to go mildly beyond the videos to be familiar with certain key words or notations or concepts. We also plan to provide support, if required.

2. Download and play with the notebooks from

   - Link1 : https://www.dropbox.com/s/l0uxoqfrsd7s27o/manifold
   - Link2 (html): https://www.dropbox.com/s/z2mh9ub9957pie0/manifold

   (I assume you know how to use/run it.) Play with the notebook and appreciate how these algorithms work in practice. Try on your own data sets. Modify the code/examples etc. as appropriate. (Aditya Arun has created this at a short notice)

3. At the end of this, you should be able to appreciate the wide range optimization problems (and their practical utilities) in:

   - K Means, (Ratio) Graph cuts, Normalized Cuts
   - PCA and MDS
   - LLE, ISOMAP, Laplacian Eigenmap, t-SNE

   Many of them lead to eigen vector based optimization (not all!!)

4. You are given exercises and questions to think and work. We will also ask you to submit answers to such questions separately.

5. We plan to arrange a special online clinic for help in understanding the associated issues. There will be regular chat/thread for regular discussions too. Please post your questions in the channel on "manifold Learning and Spectral Methods"

6. Optional/additional resources, minor changes to the content (eg. any errata or low level details) will be posted over the next few days. Please start soon, if not today.