You know the rules from the first problem set. This one is due in two weeks from now (Friday Oct 2). The assignment descriptions are intentionally vague, with enough missing information to help you think critically. That's a feature, not a bug!

Problem Set 2

Lecture 5: Recreate the figures in slides:

- Slide 25, plot the eigenvectors of the covariance instead of the arrows in my plots. The right plot has to be the result of a PCA, not a new data set. You cannot use the **`pca()`** function from the stats toolbox. You need to implement your own PCA routine using an eigen-decomposition or the singular value decomposition
- Slide 37, use the data in the file **`faces.mat`** Your output might not look exactly the same, but should qualitatively look similar
- Slide 39
- Slide 52, use the face data again. Use 10×10 pixel patches

Lecture 6: Recreate the figures in slides:

- Slide 22, the right figures should be the result of your analysis. Use the ICA algorithm in slide 17 with $\mathrm{f}(x) = x^3$. Note that this algorithm is formulated for making updates one input vector $\mathbf{y}$ at a time. You can instead get the average gradient for all input vectors at once by using the matrix $\mathbf{Y}$, but in this case you need to multiply $\mathbf{I}$ by the number of columns of $\mathbf{Y}$
- Slide 32, your results will not be exactly the same. Use the same algorithm as above, feel free to do PCA ahead of time so that you only compute ICA on 16 dimensions
- Slide 45
- Slides 51, 52, 54, your results might not look exactly the same. For PCA do not remove the mean, we simply want to find features for the data as is (as opposed to finding the principal axes of variance). The first principal component will effectively act as the mean. Use the same ICA algorithm from above, but precede it with PCA so that you get the dimensions down to three. Implement your own NMF, don't use the function from a MATLAB toolbox. The movie is named **`hands.mp4`**

Lecture 7: Recreate the figures in slides:

- Slide 36, load the distance matrix from **`cities.mat`** It will be faster and easier to perform the mean removal using simpler MATLAB code as opposed to the matrix version I've shown in class
- Slide 51 or 56 or 60, your results will be a little different. There is no need to color the points. For all of these algorithms you need to make sure that you use distances/affinities from a local neighborhood only otherwise they won't see the manifold structure. Experiment until you get the swiss roll to unfold.
- Slide 68, or 69, or 70, your results will be a little different. No need to plot the images in the low-dimensional space, just plot the low-dimensional data in 2D space. The data is in the file **`hotlips.mp4`** (your local neighborhood should be in the order of 4 to 8 points)

Extra credit: Load the image data from **`ones.mat`**, which contains a set of handwritten ones. Use a manifold method to project down to two dimensions. Show some of the ones at the point where they are projected in that 2d space. Do you see a pattern? If so what?

Enjoy!