**[HW4](https://ualearn.blackboard.com/webapps/assignment/uploadAssignment?content_id=_4622809_1&course_id=_164945_1&group_id=&mode=view)**

Attached Files:

* + [[File](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42721263_1/xid-42721263_1) pca.m](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42721263_1/xid-42721263_1) (1.794 KB)
  + [[File](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42721264_1/xid-42721264_1) 264\_optdigits.mat](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42721264_1/xid-42721264_1) (47.222 KB)
  + [[File](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42721265_1/xid-42721265_1) Mystery\_DataSet(1).mat](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42721265_1/xid-42721265_1) (3.056 MB)
  + [[File](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42722094_1/xid-42722094_1) pca\_faces.m](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42722094_1/xid-42722094_1) (533 B)
  + [[File](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42722095_1/xid-42722095_1) YaleB\_32x32.mat](https://ualearn.blackboard.com/bbcswebdav/pid-4622809-dt-content-rid-42722095_1/xid-42722095_1) (2.112 MB)

**Assignment: Learn how to use Dimensionality Reduction.**

**Undergraduate and Graduate part of the assignment:  
  
1.** Study the attached function pca.m. Copy the function to your homework directory.  
2. Use pca.m function to investigate the reconstruction error on Optdigits dataset (attached below and discussed in the textbook). In this dataset the first 64 columns represent a gray scale 8x8 bitmap of a handwritten digit. The last column represents the actual class of a digit from 0 to 9. Build a graph of error vs. number of principal components.  
3. Use pca.m to replicate figure 6.3 from the book but in 3 dimensions, make sure each data point on the graph has a distinct color to represent what digit it represents. Explain the results.

4. Plot the mean and first 8 principal components as images (8x8 grayscale bitmaps) - eigendigits. The process is explained here: <https://en.wikipedia.org/wiki/Eigenface>. The sample data from the Wikipedia example and the code to produce the eigenfaces are attached to the assignment.

**Graduate part of the assignment:**

1. Download and install [Dimensionality Reduction Toolbox.](https://lvdmaaten.github.io/drtoolbox/) Toolbox documentation is accessible in Matlab. Description of the dimensionality reduction methods is available [here](http://lvdmaaten.github.io/drtoolbox/#usage).  
2. Use Dimensionality Reduction Toolbox and build a figure similar to figure 6.3 for each of the following methods: Principle Component Analysis, Linear Discriminant Analysis, Multi-Dimensional Scaling.  In the report discuss results: what you observe, why you observe it, and how methods compare for this particular problem.  
3. Pick and study any other method available in the toolbox. Apply this method to the problem. If possible, graphically represent results and compare to the results of pca. Email me the method you chose. It will be posted on the assignment page. Only one person in class can use a given method. **If the method "does not work" please explain what properties of the dataset may be the root cause of the problem (use the method description to learn about assumptions it makes on the data).**

4. Use the attached mystery dataset. Use PCA and the non-linear method that you discovered. Observe and compare results. Is the reduction in dimensionality similar or different between two methods? What may be the reason?   
  
Taken methods will be posted here:

-  Laplacian Eigenmaps

-  t-Distributed Stochastic Neighbor Embedding (t-SNE)

-  Gaussian Process Latent Variable Model (GPLVM)

-  Stochasitic Neighbor Embedding (SNE)

-  Kernel PCA

- Landmark Isomap

-  Diffusion maps

- Probabilistic PCA

 -  Deep Autoencoders

-  Local Tangent Space Alignment

- Large-Margin Nearest Neighbor

- Maximally Collapsing Metric Learning (MCML)

- Neighborhood Components Analysis (NCA)

- Local Linear Coordination (LLC)