CS 322 Homework 6

out: Tuesday 27 March 2007
due: Wed 11 April 2007

Principal components analysis: Eigenfaces

The data file `hw6faces.mat` contains a $320 \times 200 \times 40$ matrix which comprises 40 images: 10 faces under 4 lighting conditions each. (The images are from the Yale Face Database B.) These images have some similarities, since they are all faces, and we can use PCA to try to capture these similarities using a set of basis images. (This approach was pioneered by Turk and Pentland in 1991, who dubbed the resulting principal components "eigenfaces.")

1 Compute the principal components of the set of images (i.e. the eigenfaces) using the SVD. In this context you should treat the pixels in each image as a single vector: the PCA process is just taking linear combinations of pixel, without regard for where the pixels are in the images. Describe the first three eigenfaces in terms of what they capture about the variations among the images.

2 If I want to approximate the j^{th} image by a linear combination of the first k Eigenfaces, what are the weights to use?

3 Try approximating the faces in the dataset with different numbers of eigenfaces. How many do you need before the faces become recognizable?

⋆4 Take a picture of yourself or a friend with a neutral expression and soft lighting, so that it looks similar to the faces in the dataset. Using `cpselect`, `cp2tform`, and `imtransform`, align it (by a "linear conformal" mapping, as `cp2tform` calls it) to the images in the dataset. Show the approximations for different numbers of eigenfaces How well can you approximate this new image?

5 How many eigenfaces do you need to approximate the original data matrix to within 1%? (That is, the F-norm error is less than 1% of the F-norm of the data matrix.) How can you tell this from looking at the singular values?