**Homework/Group Quiz**

1. If I know the column $\mathbf{v}_1$ from the SVD of a matrix $X$, is it possible to compute $\mathbf{u}_1$ directly?

2. On the take-home portion of the exam, we saw that a double-centering of a $p \times 2$ matrix ended up giving us a line. Prove that this is the result using a generic matrix of $p$ data points in the plane.

   (HINT: The work is in setting this problem up!)

3. Given:
$$\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{x}^{(3)} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \qquad \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}$$

   (a) Verify by direct computation that the covariance matrix of $X$ can be written as $\sum_{i=1}^{3} \mathbf{x}^{(i)} \mathbf{x}^{(i)^T}$ (first compute the covariance matrix):

   (b) Write the error (in terms of $\phi_1, \phi_2$) in using the vector $\phi$ as the only basis vector, then find the maximum using Calculus.

   (c) Show that the vector you got in part (b) is the first eigenvector of the covariance matrix (do by hand!).

4. Let $X$ be $p \times n$, where we're thinking of $X$ as having $p$ data points in $\mathbb{R}^n$ (with the mean in $\mathbb{R}^n$ equal zero). Let $\mathbf{v}_1 \in \mathbb{R}^n$ be the first eigenvector of the covariance matrix, or equivalently, the first column vector in $V$ of the SVD of $X$.

   Show that the data, when projected to $\mathbf{v}_1$, has variance $\lambda_1$.

5. Connect the best basis to the SVD:

   (a) Show, by writing $X = U\Sigma V^T$ as the reduced SVD (rank of $X^{m \times n}$ is $k$) that the *best basis* for the columnspace of $X$ is the first $k$ columns of $U$.

   (b) Similarly, show that the *best basis* for the row space of $X$ is the set of the first $k$ columns of $V$.

6. What would the following function if `data` is $[1, 2, 3, 4, 5, 6]$, `lags` is 3 and `shift` is 2?

```
function X = lag(data,lags,shift)

[N,P] = size(data);
lags = lags+1;
X = zeros(lags*N,P-lags*shift+1);

for j=0:lags-1
X(j*N+1:(j+1)*N,:) = data(:,j*shift+1:P-(lags-j)*shift+1);
end
```