Homework 8
Due Thursday, March 15, 2012

1. (20 points) Recall from Math 51 that if \( A \) and \( B \) are \( n \times n \) matrices, \( \det(AB) = \det(A) \det(B) \). Use this to show that if \( A \) is diagonalizable, then \( \det(A) \) is equal to the product of its eigenvalues.

2. Bonus problem worth 10 points of extra credit. In fact, the result above is always true whether \( A \) is diagonalizable or not. If we factor the characteristic polynomial as
\[
\chi(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda),
\]
we have \( \det(A) = \lambda_1 \lambda_2 \ldots \lambda_n \). Show why this is so.

3. I have created the matrix
\[
a_1 = \begin{bmatrix} 1 \\ \epsilon \\ 0 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 1 \\ 0 \\ \epsilon \\ 0 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \epsilon \end{bmatrix},
\]
where \( \epsilon \) is considerably smaller than machine precision. (So for example, the computer reads \( \epsilon \) as \( \epsilon \), but \( 1 + \epsilon \) as 1.)

   (a) (10 points) In MATLAB, I ran the QR algorithm and got this:
   \[
   [Q,R] = qr(A,0)
   \]
   \[
   Q =
   \begin{bmatrix}
   1.0000 & 0.0000 & 0.0000 \\
   0.0000 & -0.7071 & -0.4082 \\
   0 & 0.7071 & -0.4082 \\
   0 & 0 & 0.8165
   \end{bmatrix}
   \]
   Does it look right to you? Please explain.

   (b) (10 points) What would happen if we were to run the classical Gram-Schmidt algorithm we have seen in class (algorithm 7.1 in your book)? Would we get something that looks like the above? What would we get?

4. Suppose you have \( n \) vectors \( x_1, \ldots, x_n \) in \( \mathbb{R}^m \). In class, we have seen that the first principal component is the unit-normed vector \( u \in \mathbb{R}^m \) so that the projections of those vectors onto \( u \) have maximum variance.

   Another way to look at this is as follows: consider a line \( \mathcal{L} \) going through some point \( x_0 \in \mathbb{R}^m \) and with some orientation \( u \in \mathbb{R}^m, \|u\| = 1 \) (the equation of this line is \( x_0 + tu \) where \( t \) is a scalar). Now consider the line that is closest to the point in the sense that it minimizes
\[
\sum_{i=1}^n |\text{distance}(x_i, \mathcal{L})|^2
\]
(the sum of squares of the distances between the \( x_i \)'s and the line).

   (a) (10 points) Show that the slope of the closest line is the first principal component.

   (b) (10 points) Show that this line goes through the average vector \( \bar{x} = \sum_{i=1}^n x_i \).